Implementing the Freight Transportation Data Architecture: Data Element Dictionary

DETAILS

158 pages | 8.5 x 11 | PAPERBACK
ISBN 978-0-309-30889-2 | DOI 10.17226/21910

AUTHORS

Walton, C. Michael; Seedah, Dan P. K.; Choubassi, Carine; Wu, Hui; Ehlert, Andy; Harrison, Robert; Loftus-Otway, Lisa; Harvey, Jim; Meyer, Joel; Calhoun, Jacob; Maloney, Lucia; Cropley, Stephen; and Ford Annett

NATIONAL COOPERATIVE FREIGHT RESEARCH PROGRAM

NCFRP REPORT 35

Implementing the Freight Transportation Data Architecture: Data Element Dictionary

C. Michael Walton
Dan P. K. Seedah
Carine Choubassi
Hui Wu
Andy Ehlert
Robert Harrison
Lisa Loftus-Otway
THE UNIVERSITY OF TEXAS AT AUSTIN CENTER FOR TRANSPORTATION RESEARCH
Austin, TX

Jim Harvey
Joel Meyer
Jacob Calhoun
Lucia Maloney
ALLIANCE TRANSPORTATION GROUP
Austin, TX

Stephen Cropley
Ford Annett
TRANSMETRIC AMERICA INC.
Austin, TX

Subscriber Categories
Data and Information Technology • Freight Transportation

Research sponsored by the Office of the Assistant Secretary for Research and Technology

TRANSPORTATION RESEARCH BOARD

WASHINGTON, D.C.
2015
www.TRB.org

NATIONAL COOPERATIVE FREIGHT RESEARCH PROGRAM

America's freight transportation system makes critical contributions to the nation's economy, security, and quality of life. The freight transportation system in the United States is a complex, decentralized, and dynamic network of private and public entities, involving all modes of transportation—trucking, rail, waterways, air, and pipelines. In recent years, the demand for freight transportation service has been increasing fueled by growth in international trade; however, bottlenecks or congestion points in the system are exposing the inadequacies of current infrastructure and operations to meet the growing demand for freight. Strategic operational and investment decisions by governments at all levels will be necessary to maintain freight system performance, and will in turn require sound technical guidance based on research.

The National Cooperative Freight Research Program (NCFRP) is a cooperative research program sponsored by the Office of the Assistant Secretary for Research and Technology under Grant No. DTOS59-06-G-00039 and administered by the Transportation Research Board (TRB). The program was authorized in 2005 with the passage of the Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (SAFETEA-LU). On September 6, 2006, a contract to begin work was executed between the Research and Innovative Technology Administration, which is now the Office of the Assistant Secretary for Research and Technology, and The National Academies. The NCFRP will carry out applied research on problems facing the freight industry that are not being adequately addressed by existing research programs.

Program guidance is provided by an Oversight Committee comprised of a representative cross section of freight stakeholders appointed by the National Research Council of The National Academies. The NCFRP Oversight Committee meets annually to formulate the research program by identifying the highest priority projects and defining funding levels and expected products. Research problem statements recommending research needs for consideration by the Oversight Committee are solicited annually, but may be submitted to TRB at any time. Each selected project is assigned to a panel, appointed by TRB, which provides technical guidance and counsel throughout the life of the project. Heavy emphasis is placed on including members representing the intended users of the research products.

The NCFRP will produce a series of research reports and other products such as guidebooks for practitioners. Primary emphasis will be placed on disseminating NCFRP results to the intended end-users of the research: freight shippers and carriers, service providers, suppliers, and public officials.

NCFRP REPORT 35

Project NCFRP-47
ISSN 1947-5659
ISBN 978-0-309-30889-2
Library of Congress Control Number 2015947135

NOTICE

The project that is the subject of this report was a part of the National Cooperative Freight Research Program, conducted by the Transportation Research Board with the approval of the Governing Board of the National Research Council.

The members of the technical panel selected to monitor this project and to review this report were chosen for their special competencies and with regard for appropriate balance. The report was reviewed by the technical panel and accepted for publication according to procedures established and overseen by the Transportation Research Board and approved by the Governing Board of the National Research Council.

The opinions and conclusions expressed or implied in this report are those of the researchers who performed the research and are not necessarily those of the Transportation Research Board, the National Research Council, or the program sponsors.

The Transportation Research Board of the National Academies, the National Research Council, and the sponsors of the National Cooperative Freight Research Program do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of the report.

Published reports of the

NATIONAL COOPERATIVE FREIGHT RESEARCH PROGRAM

are available from:

Transportation Research Board
Business Office
500 Fifth Street, NW
Washington, DC 20001

and can be ordered through the Internet at:

http://www.national-academies.org/trb/bookstore

Printed in the United States of America

COOPERATIVE RESEARCH PROGRAMS

CRP STAFF FOR NCFRP REPORT 35

Christopher W. Jenks, *Director, Cooperative Research Programs*
William C. Rogers, *Senior Program Officer*
Charlotte Thomas, *Senior Program Assistant*
Eileen P. Delaney, *Director of Publications*
Sharon Lamberton, *Editor*

NCFRP PROJECT 47 PANEL
Freight Research Projects

Diane Jacobs, *California DOT, Los Angeles, CA* (Chair)
Patrick R. Anater, *CDM Smith, Pittsburgh, PA*
Jean-Francois Arsenault, *CPCS Transcom Limited, Ottawa, ON*
Peter S. Lindquist, *University of Toledo, Toledo, OH*
Anne Strauss-Wieder, *A. Strauss-Wieder, Inc., Westfield, NJ*
John H. Taylor, *Tallahassee, FL*
Krishnan Viswanathan, *CDM Smith, Tallahassee, FL*
Rolf R. Schmitt, *RITA Liaison*
Amanda J. Wilson, *RITA Liaison*
Christopher Smith, *AASHTO Liaison*
Lisa Loyo, *TRB Liaison*
Thomas Palmerlee, *TRB Liaison*

FOREWORD

By William C. Rogers
Staff Officer
Transportation Research Board

NCFRP Report 35: Implementing the Freight Transportation Data Architecture: Data Element Dictionary provides the findings of the research effort to develop a freight data dictionary for organizing the myriad freight data elements currently in use. The research identifies differences in data element definitions and methods for bridging those differences where appropriate. A product of this research effort is a searchable and sustainable web-based freight data element dictionary for transportation analysis that will be hosted by the U.S. Department of Transportation's Bureau of Transportation Statistics (BTS). A temporary link to the freight data dictionary web application is currently available at http://freightdatadictionary.com.

NCFRP Report 9: Guidance for Developing a Freight Data Architecture articulates the value of establishing architecture for linking data across modes, subjects, and levels of geography to obtain essential information for decision making. Central to the architecture is a catalog of data elements currently being collected and the definitions of those elements. Lack of a sound freight data dictionary can cause problems within and across organizations, with organizations calling the same freight data element by different names or different data elements by the same name. Worse, an organization may combine freight data elements it thinks are equivalent and make incorrect investment decisions from invalid data.

In NCFRP Project 47, The University of Texas at Austin Center for Transportation Research was asked to (1) identify readily available databases associated with freight for inclusion in the dictionary, including their key characteristics; (2) organize and classify these databases (e.g., by type and level of aggregation, attribute definitions, and spatial and temporal characteristics); (3) organize and classify the elements into a typology (with rationale) across databases and provide terms and definitions used for each element, taking into account the intended uses (e.g., land use, planning, environmental impacts, economic development, supply chain analysis, safety, and security); (4) develop and test a user interface for a searchable and sustainable web-based freight data element dictionary and make updates based on findings from the testing of the user interface; (5) identify differences in definitions and assess whether crosswalks or other bridges are adequate and relevant; (6) recommend new harmonization or statistical bridges as appropriate for resolving differences in definitions; and (7) prepare a production-ready, BTS-hosted searchable and sustainable web-based freight data element dictionary, with full documentation (including data structures, data requirements, source codes, and maintenance and updating guidelines).

AUTHORS' PREFACE

Slight or subtle variants in data definitions and metadata structures across datasets, and sometimes temporally within the same data sources, pose challenges to the compilation and use of freight data. Data analysts, regulators, and policy analysts frequently face challenges when combining data from multiple sources into a single national or state-level analysis, or when using the data for program development and administration that spans multiple geographic areas. Organizations may call the same freight data element by different names or different data elements by the same name. In some cases, freight data elements thought to be equivalent are combined, leading to incorrect investment decisions based on invalid information.

A dictionary that organizes the many current freight data elements, provides a method for identifying differences in definitions, and offers a set of homogeneous approaches for bridging gaps between definitions would serve as a critical tool for developing a National Freight Transportation Data Architecture and strengthen freight planning across agencies.

CONTENTS

1 Chapter 1 Background
1 1.1 Introduction
2 1.2 Research Objective
2 1.3 Study Approach

4 Chapter 2 The Web-Based Freight Data Element Dictionary
4 2.1 Discussion Wall
5 2.2 Data Dictionary
9 2.3 Glossary Terms

10 Chapter 3 Freight Data Uses
10 3.1 Introduction
10 3.2 Methodology
12 3.3 Congestion Management
13 3.4 Operations/Services
14 3.5 Safety Planning and Analysis
15 3.6 Freight Mobility Planning
16 3.7 Emergency Preparedness and Security Planning
16 3.8 Economic Development Planning
17 3.9 Freight Transportation and Land Use Planning
18 3.10 Environmental Planning
19 3.11 Regulation and Enforcement Planning
20 3.12 Intermodal Trade Corridor Planning
20 3.13 Terminal and Border Access Planning
21 3.14 Hazardous Materials Planning
21 3.15 Roadway Pavement and Bridge Maintenance Planning
22 3.16 Modal Shift Analysis
23 3.17 Freight Performance Measurement
23 3.18 Sustainable Transportation Investment
25 3.19 Findings from the Literature Review on Freight Data Uses

26 Chapter 4 Inventory of Freight Data Sources, Dictionaries, and Glossary Terms
26 4.1 Introduction
26 4.2 Data Dictionaries and Glossary Terms

34 Chapter 5 Classifying Data Elements Across Databases
34 5.1 Introduction
34 5.2 Background
36 5.3 Methodology
37 5.4 The RBCS Primary and Secondary Level Classifications
38 5.5 Validation
42 5.6 Model Limitation
42 5.7 Application

43 Chapter 6 Differences in Data Element Definitions
43 6.1 Introduction
43 6.2 Methodology
44 6.3 Differences in Origin and Destination Data Elements
48 6.4 Differences in Commodity Data Elements
54 6.5 Import and Export Data Elements
63 6.6 Industry Data Elements
66 6.7 Mode of Transport Data Elements
80 6.8 Safety Data Elements
83 6.9 Units of Measurement Data Elements
91 6.10 Geospatial Data

92 Chapter 7 Resolving Differences in Data Element Definitions
92 7.1 Introduction
92 7.2 Methodology
93 7.3 Place Name Bridges
98 7.4 Units of Measurement Bridges
106 7.5 Commodity Classification Bridges
113 7.6 Industry Classification Bridges
117 7.7 Mode of Transport Bridges

130 Chapter 8 Conclusions and Suggested Future Steps
130 8.1 Addressing the Barriers Identified by NFAC
131 8.2 Future Steps

134 Sources

138 References

148 Abbreviations

Note: Photographs, figures, and tables in this report may have been converted from color to grayscale for printing. The electronic version of the report (posted on the web at www.trb.org) retains the color versions.

CHAPTER 1

Background

1.1 Introduction

U.S. state and metropolitan planning agencies are now expected to incorporate freight demand into their strategic transportation policies. While many had always recognized trucking demand in their highway needs, the inclusion of other modes was stimulated by federal legislation that began in 1991 with the passage of the Intermodal Surface Transportation Efficiency Act (referred to as ISTEA). This began the process of switching the planning focus from highway departments and networks to transportation agencies and modal systems. Subsequent legislative actions include the Transportation Equity Act for the 21st century (TEA-21) in 1998 and Moving Ahead for Progress in the 21st Century Act (MAP-21) in 2012, which further strengthened the critical role played by freight transportation in supporting the goals of economic competitiveness, safety, and sustainability.

MAP-21 includes provisions to improve the condition and performance of the multi-modal national freight network and also requires that all state departments of transportation (DOTs) direct resources toward improving freight movement through several initiatives, such as the following:

- Assessing the condition and performance of the national freight network
- Identifying highway bottlenecks that cause significant freight congestion
- Forecasting freight volumes
- Identifying major trade gateways and national freight corridors
- Reducing barriers that impact freight transportation performance (FHWA 2012)

These initiatives require an understanding of both current and future freight demand and the different modal transportation networks utilized, which is arguably most efficiently determined using robust models accessing accurate, consistently defined freight data (Walton et al. 2014). Moreover, current modeling effectiveness and potential is already limited by data constraints related to focus, structure, definitions, sampling designs, timing, and relevance. Improved and more robust data sets are needed to allow freight models to adequately capture the determinants of freight demand, accurately measure the impact of freight on the transportation infrastructure, and effectively support the decision-making processes of public and private stakeholders at the national, state, regional, and local levels (Chase et al. 2013).

A national freight data architecture linking various freight data sources across modes, subjects, and levels of geography has been proposed to enhance data inputs, thus improving current modeling (Chase et al. 2013). The many current challenges in linking multiple freight data sources (as identified in the literature) include the following:

- Different origin and destination definitions and geographic units that do not directly correlate
- Different commodity classifications
- Different assumptions to estimate data or deal with missing data

- Different expansion factors and control totals
- Differing procedures used for data aggregation or disaggregation
- Difficulty in obtaining proprietary data from private sources
- Inconsistency of data across different modes of transport
- Inconsistency of data collection efforts across different modes of transport (rail versus highway versus air cargo versus intermodal)
- Inaccurate or nonexistent local-level commodity flow data
- Different vehicle classifications
- Different data storage formats and dictionary definitions

Slight or subtle variants in data definitions and metadata structures across datasets, and sometimes temporally within the same data sources, pose challenges to the compilation and use of freight data. Data analysts, regulators, and policy analysts frequently face challenges when combining data from multiple sources into a single national or state-level analysis, or when using the data for program development and administration that spans multiple geographic areas.

Organizations may call the same freight data element by different names or call different data elements by the same name. In some cases, freight data elements thought to be equivalent are combined, leading to incorrect investment decisions based on invalid information. A dictionary that organizes the many current freight data elements, provides a method for identifying differences in definitions, and offers a set of homogeneous approaches for bridging gaps between definitions would constitute a critical tool to strengthen freight planning. The need for such a dictionary identifies the primary focus of this research project.

1.2 Research Objective

The key objective of the research was to produce a searchable and sustainable web-based freight data element dictionary for transportation analysis, with an accompanying set of recommendations for identifying differences in definitions and developing statistical and harmonization bridges between definitions, as appropriate for resolving differences. The dictionary, designed for a wide range of potential users, is capable of supporting a variety of future freight planning initiatives at metropolitan, state, regional, and national levels. It is structured to benefit from user feedback and database updates as well as helping frame greater consistency in terms of definitions, content, and temporal sampling of current databases when they are updated.

1.3 Study Approach

The study consisted of the following tasks:

- Identify "readily available" data sources associated with freight.
- Provide examples of freight data uses and applications.
- Compile and classify an inventory of data elements and glossary terms found in the selected sources into a uniform typology.
- Identify differences in data element definitions.
- Provide metadata tools and resources to guide data users on the appropriate steps and procedures for combining data from multiple freight data sources.
- Develop a searchable and sustainable web-based application containing the study findings, an inventory of freight data dictionaries, and a discussion feature to be used by practitioners to exchange ideas and information.

NCFRP Report 35 presents the findings from each of the above tasks. More than 40 U.S. freight-related data sources were identified in the literature, and their data elements were organized into a typology across databases so that similar data elements could be identified. Classifying similar data elements facilitated the identification of differences in their definitions and aided in the development of harmonization or statistical bridges, as appropriate, for resolving those differences. Examples of freight data uses were also compiled from the literature to demonstrate how freight data sources are currently being utilized by agencies and the research community.

All the information contained in this report is available on the searchable web-based Freight Data Dictionary application. The purpose of this web-based tool is to provide an avenue where the information gathered from this study can be updated as newer data sources and methods for resolving data heterogeneity become available. The web-based application provides an opportunity for practitioners to exchange ideas and information to support the effective and accurate use of freight data. It also widens the utility of freight data sources by assisting less experienced planners to derive more accurate output and widen data use.

This report is organized into eight chapters, including the introduction. Chapter 2 describes the development of a web-based data dictionary framework that would result in a searchable and sustainable product. Chapter 3 identifies the wide range of activities in which freight data are used, including operations, congestion, safety, security, economic development, and land use. Chapter 4 provides an inventory of freight data sources and dictionaries and provides a glossary of terms. Chapter 5 considers the challenge of classification and validation of data elements across databases—factors that limit models and their application. Chapter 6 examines differences in data element definitions across a wide set of databases, while Chapter 7 addresses the challenge of resolving the differences, which is critical to meeting the prime research objective of the project. Chapter 8 provides suggestions for implementing the product of this work and undertaking a variety of additional activities related to strengthening the model and extending its use to analyze and support freight planning programs across a wide range of public and private uses.

CHAPTER 2

The Web-Based Freight Data Element Dictionary

The objective of the research in NCFRP Project 47 was to produce a searchable and sustainable web-based freight data element dictionary. The information contained in the Freight Data Element Dictionary (Freight Data Dictionary) is illustrated in Figure 2-1.

The Freight Data Dictionary is made up of three major databases: the Discussion Wall, the Data Dictionaries, and the list of Glossary Terms. The application's default search bar is shown in Figure 2-2.

The search bar comes with several features, including autocomplete, exact phrase search using quotes, and alternate search suggestions for misspelled words. An advanced search function is also available to narrow down search results to specific entries that meet the criteria provided by the user. However, the advanced search option is available only for the Freight *Data Dictionaries* and the *Glossary Terms*.

The balance of this chapter provides additional information on the content contained in each database.

2.1 Discussion Wall

The Discussion Wall contains all of the information in this report (Figure 2-1). It facilitates ***interlinking*** of information across multiple topics and seamlessly integrates with the Freight Data Dictionary search engine. For example, names of databases are linked to their original discussions as provided in Appendix A of this report. This is a useful feature considering the limitations of paper-based reports. Searching for keywords in the default search bar will also result in a search being performed on the Discussion Wall (Figure 2-3).

When a user clicks on the relevant topic, the web application navigates to the page containing this information. References to related topics, relevant documents, and reports are included on each page. Registered users also have the ability to post comments on topics of interest by clicking on the "Discussion" tab shown next to the "Page" tab (Figure 2-4). These external comments will be regulated by an administrator to ensure their appropriateness to a particular topic.

In addition to the topics presented in this report, the Discussion Wall contains the *Freight Data Dictionary User Guide*, a list of frequently asked questions, and a contact page for communicating with the administrator.

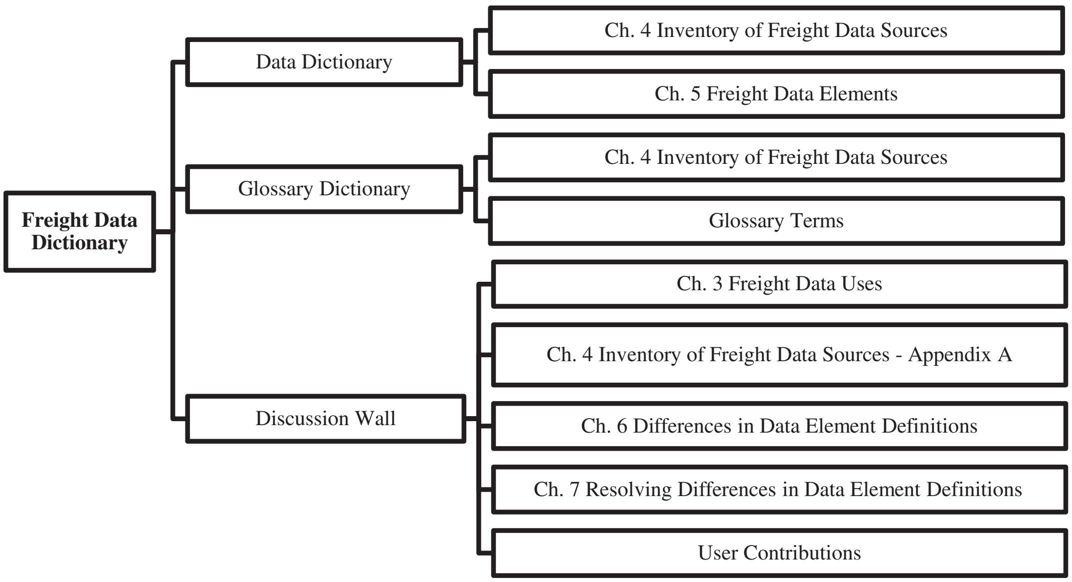


Figure 2-1. Freight Data Dictionary framework.

2.2 Data Dictionary

The Data Dictionary provides a way to search all of the data dictionaries discussed in Chapter 4 and Chapter 5 of this report. The tool works in a manner similar to the way it works for the Discussion Wall: it queries the available data element names and descriptions to provide results relevant to the search keyword. Each page of results represents data dictionary elements from a particular data source. For example, a search for the word "origin" will yield results from 10 data sources, as shown in Figure 2-5. Using the dropdown menu (currently showing "Vehicle Inventory and Use Survey"), users can navigate between the 10 data sources. The elements on the page are divided into sub-databases and tables, based on the table or sub-database to which the element belongs.

Users can view a data element's description, type, and other additional information by clicking on the "Info" icon (a lowercase "i" in a circle). Data elements that are similar to the search results can also be obtained by clicking on the "Similar Elements" link at the bottom of the page. Users seeking to retrieve the entire dictionary generated for that particular data table can click on the "Complete Table Profile" link.

Users also can employ Role-Based Classification Schema (RBCS) queries to identify all Data Dictionary elements that have the same element role. To use this feature, simply type "RBCS" (case insensitive) followed by the desired element role. Autocomplete can be useful for entering acceptable RBCS queries. An example of an acceptable input is shown in Figure 2-6.

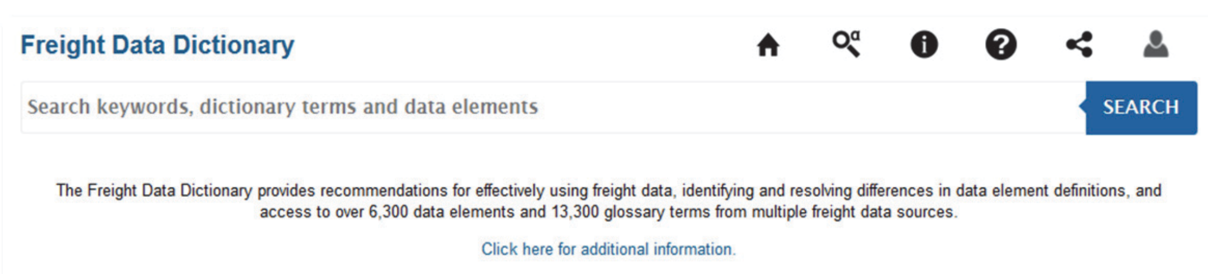


Figure 2-2. Freight Data Dictionary search bar.

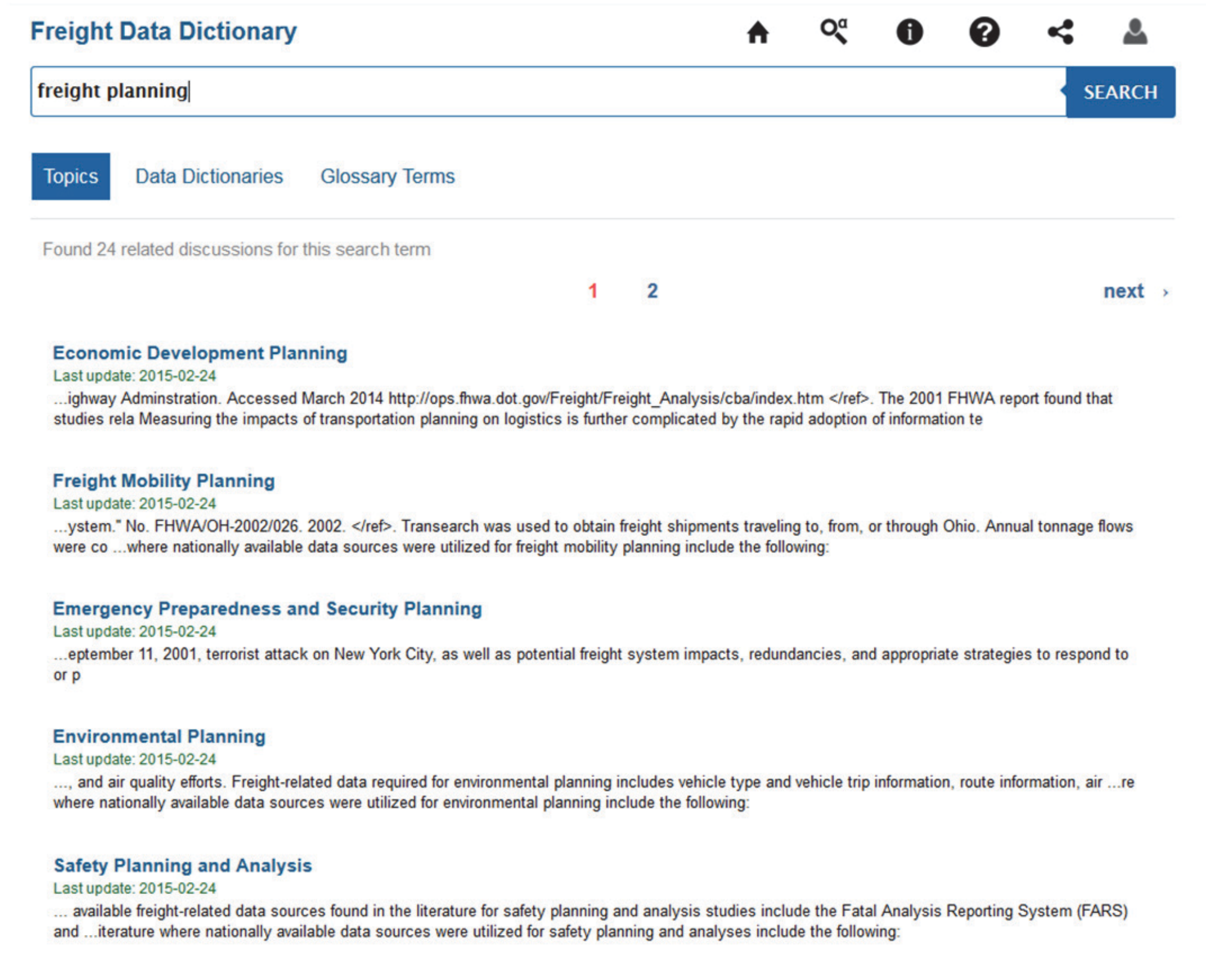


Figure 2-3. Discussion Wall topic search.

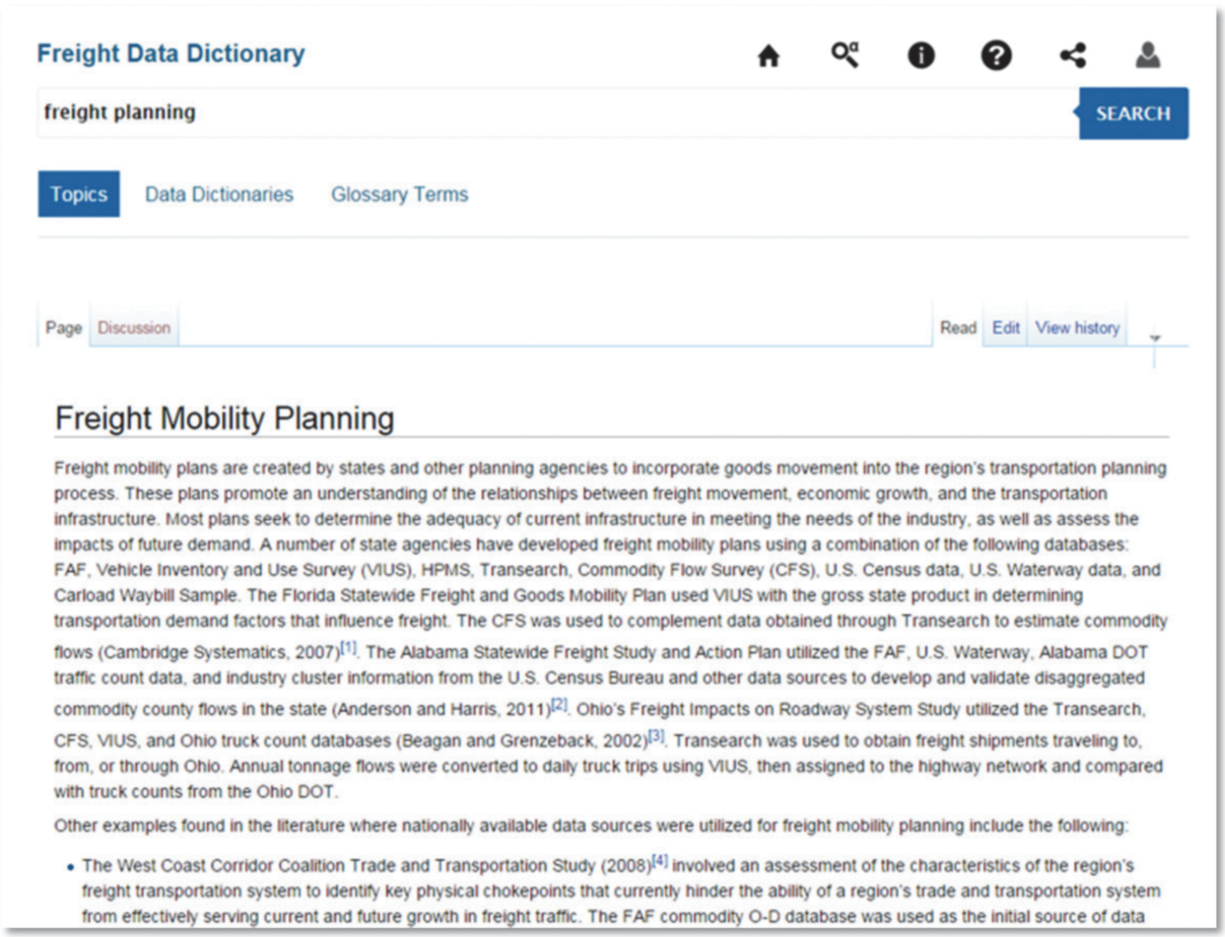


Figure 2-4. Discussion Wall topic screenshot.

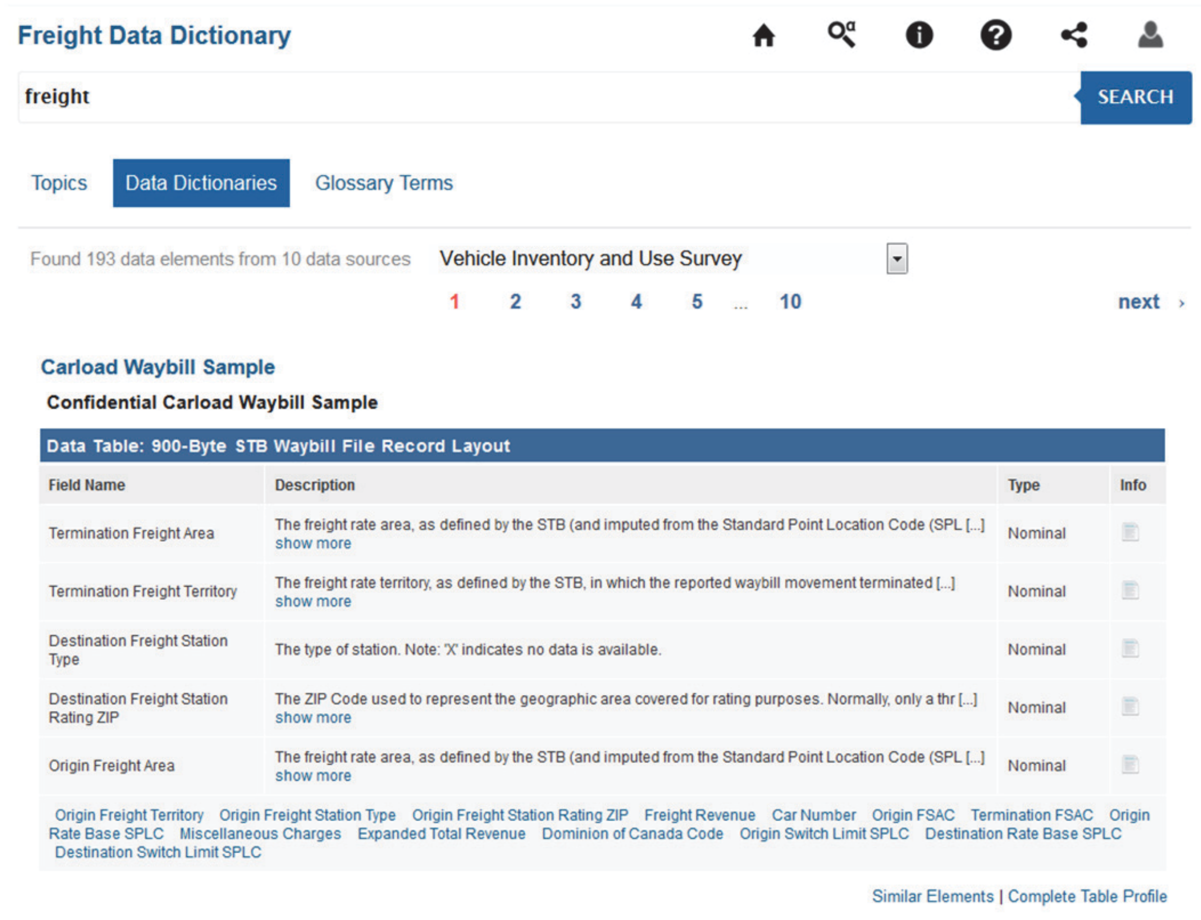


Figure 2-5. Data Dictionaries search results.

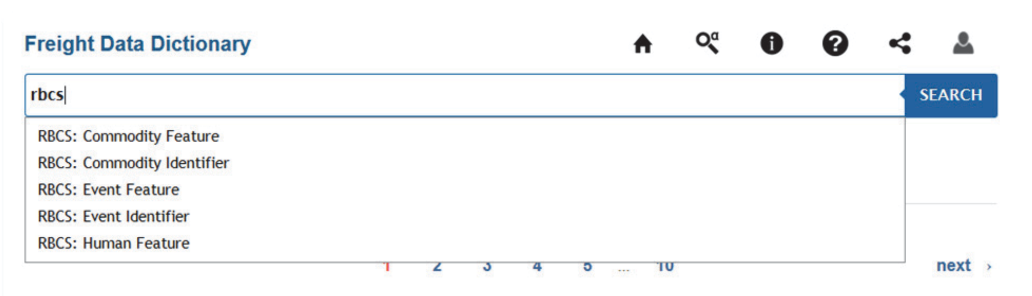


Figure 2-6. RBCS search feature.

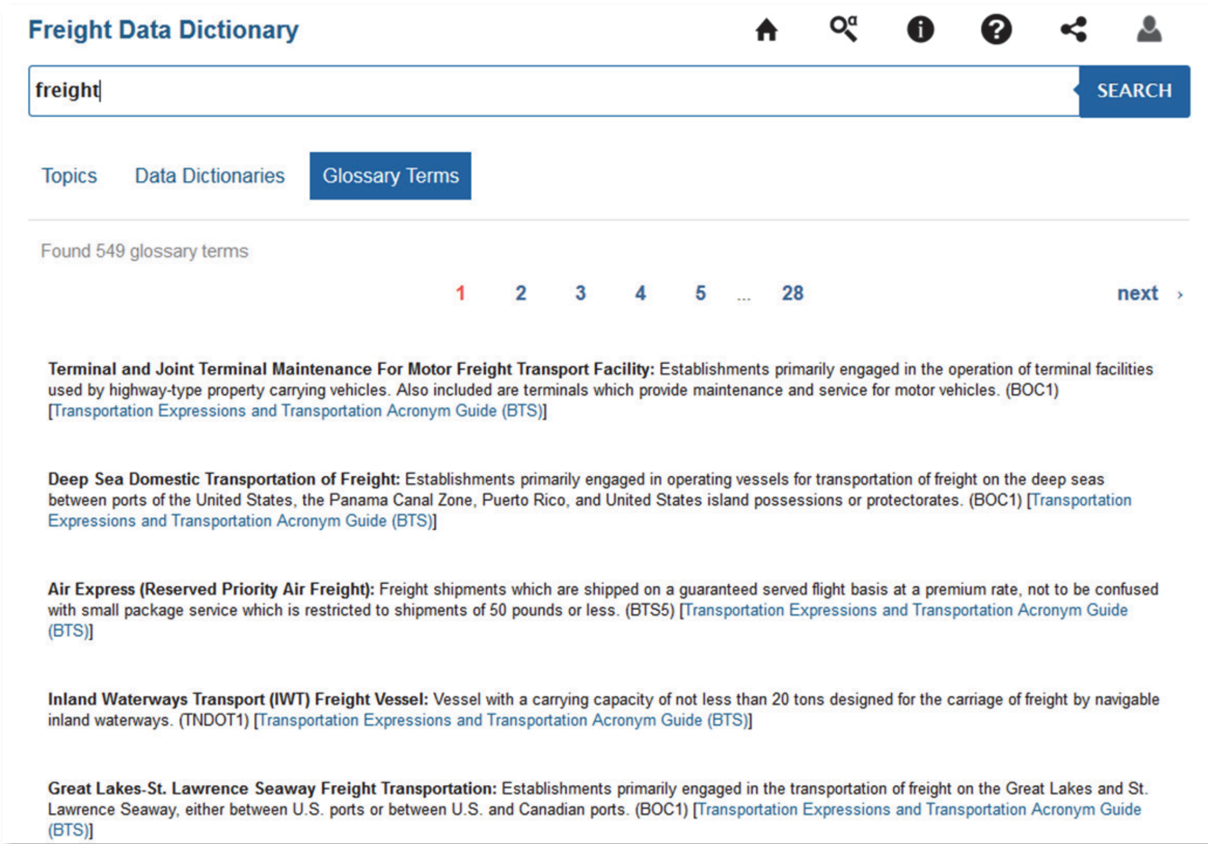


Figure 2-7. Glossary Terms search results.

2.3 Glossary Terms

The Glossary Terms option searches the compiled list of glossary terms discussed in Chapter 4 of this report. As with searches of the Data Dictionaries, searches of the Glossary Terms work by querying the available glossary terms and their definitions (see Figure 2-7). Links to the sources of the glossary terms appear in brackets at the end of each term definition. Clicking on a link loads the respective glossary data source.

CHAPTER 3

Freight Data Uses

3.1 Introduction

NCFRP Report 22: Freight Data Cost Elements compiled a comprehensive list of 18 public-sector planning and decision-making functions from a review of research publications, government documents, and other sources. Public-sector organizations captured in the identification and definition of the functions include federal and state departments of transportation (state DOTs), metropolitan planning organizations (MPOs), and port/airport/railroad authorities, as well as economic development and environmental agencies (Holguín-Veras et al. 2013).

To ensure consistency between the definitions of public-sector functions found in the study for *NCFRP Report 22* and in this study, a literature review was conducted using the initial list of public-sector functions identified in *NCFRP 22* plus two additional functions: modal shift analysis and freight performance measurements. Therefore, 20 freight-related public-sector functions were used in conducting the literature review. On completion of the literature review, the final set of public-sector functions was reduced to 16, given that some examples of freight data uses were found to be captured in other public-sector functions, as shown in Table 3-1 and Table 3-2.

This chapter summarizes the results of the literature search by the NCFRP Project 47 research team on how freight data is being used in an innovative or unique way to perform a function. Given the limitations on the scope of the study, this chapter provides examples that illustrate how freight databases are being used. It is suggested that over time, data elements from new or additional studies be added to and cited in the web-based freight data dictionary.

3.2 Methodology

TRB's TRID database was used in the literature search (TRB 2014). The various functions were separately searched using combinations of keywords, index terms, and subject headings derived from the descriptions provided of each function. Results were limited to nationally based studies written in the English language. Research in progress and international studies were not included.

Searches were honed and publication years were limited until a manageable number of relevant results (100–300 results) were achieved. All searches began with the publication year limit of 1994–2014, but many searches were limited to the most recent decade or even the past 5 years, depending on the depth of the topic. The review covered approximately 1,000 publications found in the transportation literature. From the manageable number of relevant results, the most recently published items with readily available full-text PDFs were selected. The literature review then examined how freight-related sources were utilized for that study.

Table 3-1. Freight planning and decision-making public-sector functions.

	Function	Description (adapted from *NCFRP Report 22*)
1	Congestion Management	Identify and monitor recurring and non-recurring congestion along road corridors and evaluate and recommend mitigation strategies
2	Operations/Services	Develop, operate, and maintain transportation modes; improve the movement of goods and people and increase the safety and efficiency of the transportation system through enhanced management and operations coordination
3	Safety Planning and Analysis	Implement and maintain integrated multimodal safety and transportation planning; the ultimate goal is to reduce crashes, injuries, and fatalities
4	Freight Mobility Planning	Incorporate goods movement into the regional transportation planning process
5	Emergency Preparedness and Security Planning	Increase the safety and security of the transportation system through enhanced coordination and communications among emergency responders
6	Economic Development Planning	Estimate the impacts of transportation planning on local population and employment
7	Freight Transportation and Land Use Planning	Coordinate regional freight transportation planning and land use development
8	Environmental Planning	Investigate activities involving mobile emissions planning, environmental protection, land use management, and air quality efforts
9	Regulation and Enforcement	Conduct activities such as licensing, inspection, size and load specifications, work hours regulation, and taxes/fares
10	Intermodal Trade Corridor Planning	Develop intermodal corridors to ensure efficient freight movement and reduce congestion
11	Terminal and Border Access Planning	Manage terminals and borders to ensure efficient movement of people and goods across modes
12	Hazardous Materials Planning	Improve safe movement and monitoring of hazardous materials transported using the freight system
13	Roadway Pavement and Bridge Maintenance Planning	Study the effects of fleet use on infrastructure, such as expected pavement deterioration
14	Modal Shift Analysis	Investigate policies and incentives that foster modal shift changes, including measuring the impact of shifting from one mode to another
15	Freight Performance Measurements	Develop measures to monitor the performance of the freight transportation system, including its subsystems and components
16	Sustainable Transportation Investment	Investigate ways to fund the existing transportation system and future projects

Table 3-2. Additional functions identified but only partially covered in *NCFRP Report 35*.

	Function	Description (from *NCFRP Report 22*)
1	Financial Planning	Investigate grants, loans, and subsidies to support the transportation system; also involves tax policy, road user fee assessment, and other activities such as public-private partnerships (*partially captured in Economic Development Planning)*
3	Interregional Connectivity	Develop intermodal corridors to ensure efficient freight movement and reduce congestion (*captured in Intermodal Trade Corridor Planning)*
4	Security Planning	Integrate emergency response and other calculations into transportation planning (*captured in Emergency Preparedness Planning)*
5	Transportation Equity Planning	Incorporate transit equity principles and legislation such as SAFETEA-LU into regional transportation planning *(excluded as no examples of freight data use were found)*

The initial strategy was to develop a weighting system by which freight data uses could be categorized as serving one of three functions:

1. Cleaning data and using simple graphs and tables to show relevant information, such as using the Freight Analysis Framework (FAF) for commodity flows.
2. Combining a database with another database to perform a selected function; for example, integrating place-based databases (e.g., the Commodity Flow Survey) with network-based databases (e.g., the Highway Performance Monitoring System).
3. Combining a database with other databases to perform extensive transformation through statistical analysis and assumptions to perform a selected function, such as using the FAF to determine roadway emissions.

On completing the literature review, the research team found that many of the studies fell into the first category. Though relevant in validating the need for performing a public-sector function, studies from the first category were excluded from this report because of the sheer volume of such publications. It was also thought to be more beneficial to present a less redundant and more diverse set of examples of how currently available freight databases are being utilized.

The literature selected for discussion in *NCFRP Report 37* is therefore based on how freight data was used in an innovative or unique way to perform a public-sector function. The research team recognizes that modeling attempts in some of these studies may have been limited; however, these studies were included in the report to serve as examples and to demonstrate the current limitations of most of our databases, especially in areas such as freight modeling at disaggregate county and sub-county levels. The idea is that the methodologies used in some of these studies can be adopted and enhanced as newer and richer data becomes available to researchers. The current design of the Freight Data Dictionary enables future studies to be easily incorporated into the system using the Discussion Wall feature.

3.3 Congestion Management

In this report, *congestion management* is defined as the identification and monitoring of recurring and non-recurring congestion along road corridors. It involves evaluating and recommending strategies that mitigate traffic congestion and facilitate a reliable and efficient flow of personal and commercial vehicles. Freight-related congestion management studies found in the literature mostly focused on mitigating the effects of truck movements along urban corridors. Recurring topics included examining the impact of truck lane restrictions, roadway pricing strategies, and the cost of urban freight congestion.

The most commonly cited nationally available freight-related database for congestion management studies is the Highway Performance Monitoring System (HPMS) dataset. In this database, truck traffic counts are combined with other data sources, such as roadway geometry data, incident data, weather data, vehicle registration data, and truck GPS data. In some instances, the HPMS was combined with the FAF to estimate future truck flows and truck freight value. However, most congestion management studies utilized field data collected specifically for the study area or data provided by local traffic agencies. Data from these studies were sometimes complemented with state DOT traffic data. Sources of field data collected for most of the studies include manual traffic counts, video monitoring, and surveys specifically designed for a study area. The data from these individual studies, though relevant, is rarely available to or accessible by others on completion of a study. Furthermore, no central data collection repository is available where the locally collected data can be stored or shared with other data users in the transportation community.

These examples from the literature demonstrate the use of nationally available data sources for congestion management studies:

- Eisele et al. (2013a) combined traffic volume and roadway inventory data from the HPMS with historical speed data from INRIX (a traffic information provider) speed data to estimate urban truck freight delay and diesel fuel consumption, and the associated costs for trucks in urban congestion. A geographic matching process was performed to assign traffic speed data from INRIX to each HPMS road section, and traffic volumes for each hour time interval from daily volume data were estimated. Congestion performance measures were calculated using calculated average travel speed and total delay for each hourly interval.
- Guo et al. (2010) utilized the HPMS, National Highway Planning Network (NHPN), and the Lock Performance Monitoring System (LPMS) from the U.S. Waterway database to identify and assess transportation infrastructure bottlenecks in the Mississippi Valley Region. The HPMS database was used in analyzing regional highway traffic conditions and freight bottlenecks, and the LPMS database was used in identifying the location of lock delays on the inland waterway system in the region. HPMS data was mapped onto the NHPN network through a dynamic segmentation process, and detailed traffic information on sampled sections was extrapolated to universe sections for freeways. Truck unit delay, measured in hours of delay for trucks per 1,000 miles, was used in identifying bottleneck locations on the network. A congestion corridor growing method was also incorporated in the analysis framework to account for the systematic congestion caused by interchange bottlenecks.
- Cambridge Systematics (2005) developed a methodology to identify freight bottlenecks using HPMS, NHPN, and FAF data. Highway bottlenecks were located by scanning the HPMS database for highway sections that were highly congested, as indicated by a high volume of traffic in proportion to the available roadway capacity (the volume-to-capacity ratio). Using the FAF and the HPMS sample databases, the volume of trucks passing through the identified bottlenecks were also estimated, and truck-hours of delay was calculated.
- Eisele et al. (2013b) also developed estimated state and urban-area commodity values by integrating the commodity value supplied by FAF with truck vehicle-miles traveled (VMT) calculated from the HPMS roadway inventory database. Truck VMT is computed as the product of average daily traffic, percent trucks, and link length. To obtain the truck VMT-based commodity values, predetermined state and urban truck VMT percentages were multiplied by the U.S. truck commodity values from the FAF.

3.4 Operations/Services

Operations and services functions involve the development, management, and maintenance of transportation modes to improve the movement of goods and people and increase the safety and efficiency of the transportation system. The studies within the literature relating to this function also included infrastructure planning, prioritizing needs, and assessing network vulnerability and resiliency. In some instances, FAF data were combined with data from the HPMS, the U.S. Bureau of Economic Analysis (BEA), and the National Transportation Atlas Database (NTAD).

These examples from the literature show how nationally available data sources were utilized for operations/services:

- Jansuwan et al. (2010) developed a decision support tool to assess the vulnerability of the transportation network and conducted a case study based on disruption scenarios of deficient highway bridges on the Utah highway network. State-specific commodity flows within, out of, into, and through Utah were extracted from FAF version 2.2 (FAF2.2) and converted into truck origin-destination (O-D) trips. To generate the case study scenarios, data from the National Bridge Inventory (NBI) database and Utah's seismic hazard map were utilized.

Disruption scenarios from an earthquake, based on assumed impassable status of bridges after a strong earthquake, were selected for structurally deficient bridges in or near high seismic hazard areas. Changes in travel distance and VMT as a result of trucks using alternative routes were measured. Applications of the tool include developing recommendations for prioritizing bridges for maintenance, retrofitting, and detour route planning for freight movements, among others.

- Schroeder et al. (2012) used BEA and FAF data to develop a freight-based prioritization framework to identify freight infrastructure needs critical to maintaining economic vitality by incorporating economic metrics associated with infrastructure performance and roadway level of service. The framework first evaluated infrastructure needs on a specified highway network, then prioritized those needs using a decision model to balance developed economic metrics that estimate regional corridor-wide benefits of the local improvement with severity of needs as quantified with conditional performance measures. The BEA input-output model was used to identify the most transportation-dependent industrial sectors, which were then linked with commodity flows using the FAF. A set of conditional performance measures was selected to identify critical locations meriting improvements, including National Bridge Investment Analysis System (NBIAS) outputs, International Roughness Index (IRI), truck fatality crash rate and truck crash rate, and deficiencies in geometric standards.
- Kersh et al. (2012) developed a risk-based approach to identifying and prioritizing Interstate segments for planning alternate route diversions for trucks, and a method for selecting preferred alternative truck routes when diversion is required. The methodology used traffic data from the Tennessee DOT's travel information system, the Tennessee Department of Safety, and the NTAD-NBI to rank all Tennessee Interstate segments on the basis of route restrictions, relative truck traffic, history of severe accidents, and congestion levels. Alternate routes were generated in a GIS environment that considered both trucks and passenger vehicles and took into account criteria for roadway grades, clearances, bridge design loads, school zones, capacity, and demand.

3.5 Safety Planning and Analysis

Safety planning and analysis are defined here as implementing and maintaining integrated multimodal safety and transportation planning. The ultimate goal is to reduce crashes, injuries, and fatalities. Nationally available freight-related data sources found in the literature for safety planning and analysis studies include the Fatal Analysis Reporting System (FARS) and the Federal Railroad Administration (FRA) Safety Database. FHWA uses injury and fatality data from the FARS database, combined with VMT data from HPMS, to report the number and rate of injuries and fatalities involving large trucks in its safety performance measures criteria (FHWA 2000).

These examples from the literature demonstrate the use of nationally available data sources for safety planning and analyses:

- Hall and Mukherjee (2008) carried out analytical and statistical analyses to identify and quantify the factors that contribute to freight-related crashes using FARS and an additional dataset called Trucks Involved in Fatal Accidents (TIFA), which provides coverage of all medium and heavy trucks recorded in FARS (Jarossi et al. 2011). Researchers linked the crash time, date, day, month, year, and age of driver from the FARS database, as well as the number of hours driven and the trip type from TIFA, to study the safety impact of the length of time drivers have been operating their vehicles and the effect of hour-of-service regulations on enhancing safety.
- Liu et al. (2013a) developed a methodology for quantifying the relationship between train derailment severities and their associated affecting factors, such as residual train length, derailment speed, train power distribution, and proportion of loaded railcars in the train, using the Rail Equipment Accident (REA) database maintained by the FRA.

- Liu et al. (2013b) also used the REA data on broken-rail-caused car derailments to develop a statistical model that considers a combination of risk-reduction strategies to assist decision-makers in improving the safety of transporting hazardous materials by rail.

3.6 Freight Mobility Planning

Freight mobility plans are created by states and other planning agencies to incorporate goods movement into the region's transportation planning process. These plans promote an understanding of the relationships between freight movement, economic growth, and the transportation infrastructure. Most plans seek to determine the adequacy of current infrastructure in meeting the needs of the industry and to assess the impacts of future demand. State agencies have developed freight mobility plans using a combination of the following databases: FAF, Vehicle Inventory and Use Survey (VIUS), HPMS, IHS Global Insight's Transearch (Transearch), Commodity Flow Survey (CFS), U.S. Census data, U.S. Waterway data, and the Carload Waybill Sample. The Florida Statewide Freight and Goods Mobility Plan used VIUS with the gross state product in determining transportation demand factors that influence freight. The CFS was used to complement data obtained through Transearch to estimate commodity flows (Cambridge Systematics 2007). The Alabama Statewide Freight Study and Action Plan utilized the FAF, U.S. Waterway, Alabama DOT traffic count data, and industry cluster information from the U.S. Census Bureau and other data sources to develop and validate disaggregated commodity flows in the state (Anderson and Harris 2011). Ohio's Freight Impacts on Roadway System Study utilized the Transearch, CFS, VIUS, and Ohio truck count databases (Beagan and Grenzeback 2002). Transearch was used to obtain freight shipments traveling to, from, or through Ohio. Annual tonnage flows were converted to daily truck trips using VIUS, then assigned to the highway network and compared with truck counts from the Ohio DOT.

These examples from the literature demonstrate the use of nationally available data sources for freight mobility planning:

- The West Coast Corridor Coalition Trade and Transportation Study (Cambridge Systematics 2008) involved an assessment of the characteristics of the region's freight transportation system to identify key physical chokepoints that currently hinder the ability of a region's trade and transportation system from effectively serving current and future growth in freight traffic. The FAF commodity O-D database was used as the initial source of data for estimating international commodity flow demand through West Coast seaports and inland movements of international shipments. The study team compared FAF base-year and forecast estimates with existing port demand estimates to determine and address inconsistences in individual port demand estimates. The FAF also was used for the estimation of base-year and forecast North American Free Trade Agreement (NAFTA) freight demand between the United States and Canada/Mexico through border-crossing locations in the study area. FAF NAFTA freight flow estimates were compared against the North American Transborder Freight Database (Transborder) and other border-crossing traffic flow data from Canada and Mexico. The FAF highway network was used for the analysis of base-year and forecast highway system characteristics in the study area pertaining to highway network capacity constraints and bottlenecks. HPMS truck traffic data was used as the base-year truck traffic count. Internal and external truck growth rates for forecasting were developed from the FAF and converted to truck trips using payload factors from VIUS data. Base-year air cargo demand through major airports in the study area relied on cargo data reported by airlines to individual airports. The airline-reported data was then compared and vetted against the Air Carrier Statistics database and the U.S. Foreign Trade Data. Forecast air cargo demand data was derived from the airport master plans available from major airports in the study area. Rail network demand and forecast were based on data from available regional/statewide rail studies in the study area.

- *NCFRP 14: Guidebook for Understanding Urban Goods Movement* (Rhodes et al. 2012) provides information on how multiple freight data sources can be used to address freight issues at the local level. Examples of issues discussed include safety, congestion, land use planning, emissions, environmental justice, commercial vehicle routing, and travel demand modeling. Data sources cited in the guidebook and grouped by geographical coverage include the following:
  - **Freight node data,** which represent consolidated or individual endpoints that generate or receive freight flows and are the key points of production, consumption, or intermediate handling for goods. Example data sources are the NTAD, InfoUSA™, Harris InfoSource, or ThomasNet®.
  - **Freight network data,** which define major route patterns and critical infrastructure being used to convey freight shipments through the various modal systems. Examples include the HPMS, NTAD, and NHPN.
  - **Freight flow data,** which provide information on commodity flows and provides insight on the economic and trade environment of regions. Typical commodity flow records will contain information on the O-D of shipments, type of commodity, weight, and/or value of the commodity shipment, and mode of shipment. Example data sources include CFS, FAF, Carload Waybill Sample, and Transearch.
  - **Neighborhood freight data,** which provide information on safety, congestion, land use, and emissions. Example data sources are HPMS and FARS.

3.7 Emergency Preparedness and Security Planning

For NCFRP Project 47, emergency preparedness and security planning were defined as increasing the safety and security of the transportation system through enhanced coordination and communications among emergency responders. Few specific freight-related emergency preparedness planning studies were found in the literature, as most studies focused on first responders, mass transit, and natural disaster response (e.g., earthquake and strategic military response). Moreover, many studies were overview reports that did not undertake data analysis. This finding may reflect both post-9/11 priorities set by the U.S. Congress and the paucity of reliable data available, as noted in the studies.

A nationally available data source was utilized for emergency preparedness and security planning in the Freight Planning Support System for Northern New Jersey study (Fallat et al. 2003). This study examined interruptions in freight movement caused by the September 11, 2001, terrorist attack on New York City, as well as potential freight system impacts, redundancies, and appropriate strategies to respond to or prevent system failure in the event of another major disaster within the northern New Jersey region. The study utilized Transearch data for commodity flows, traffic and highway network data from the New Jersey DOT, the Center for Transportation Analysis (CTA) national rail network, Port Import/Export Reporting Service (PIERS) Maritime Database, FRA Rail Waybill Sample, and other public and private data sources, which were all integrated into GIS for analyses.

3.8 Economic Development Planning

Economic development planning seeks to tie the impacts of freight-related infrastructure projects to economic growth—specifically, to increases in employment opportunities, resource consumption, property values, wealth accumulation, and productivity (Litman 2010). Though it is possible to measure the direct impact of transportation improvements on travel time, difficulties arise when attempting to estimate “the indirect nature and relevance” of transportation improvements to logistics operations, inventory management, and overall business decision-making (AECOM 2001). The 2001 FHWA report found that studies relating highway improvements to logistics decision-making have been found to be more qualitative than empirical because of a lack of data resulting from privacy concerns and operational competitiveness (AECOM 2001). Following are some examples of the qualitative evidence of highway improvements on logistics operations cited by Jack Faucett Associates (1994):

- Reduced inventory costs resulting from faster and more reliable replenishment delivery times.
- Economies of scale in larger volumes of output per plant given access to wider distribution markets.
- Reductions in regional warehouse operations resulting from more direct deliveries from plants to retailers, wholesale distributors, and customers as a result of more reliable delivery times direct from manufacturers (Jack Faucett Associates 1994).

Measuring the impacts of transportation planning on logistics is further complicated by the rapid adoption of information technology tools in the supply chain. As stated by AECOM (2001), "the impacts of highway improvements on transit time as well as technological changes in the trucking industries suggest that distinguishing causal relationships of highway improvements on logistics has become more complex." Therefore, attempts to empirically quantify the "explicit linkages" between infrastructure projects and economic growth "are [often] characterized by assumptions or hypothetical situations" (AECOM 2001).

These examples from the literature demonstrate the use of nationally available data sources for economic development planning:

- AECOM (2001) developed a methodology to relate the demand for freight transportation to freight transport charges and highway performance. This methodology is based on the assumption that freight charges are dependent on highway performance since average vehicle speed and speed cycling directly affect carrier's costs and, presumably, shipping rates. The study utilized data from several sources, including performance and traffic volume data from the HPMS, commodity flow data from the FAF, and regional economic activity data from the Bureau of Labor Statistics and the BEA.
- Sage et al. (2013) developed a process to address the need for an improved method to analyze freight benefits associated with proposed highway and truck intermodal improvements. Regional travel demand models (TDMs) are used in calculating transportation benefits associated with freight investments, including truck travel time savings, truck operating cost savings, and truck emission changes. The freight transportation-related benefits from the TDM are then used in performing regional economic impacts analysis with IMPLAN's Input-Output and Washington State computable general equilibrium models, which were generated with IMPLAN data.
- *NCFRP Report 12: Framework and Tools for Estimating Benefits of Specific Freight Network Investments* (Cambridge Systematics et al. 2011) also developed the Freight Evaluation Framework, which seeks to (1) enhance public planning and decision-making processes regarding freight; (2) supplement benefit/cost assessment with distributional impact measures; and (3) advance public-private cooperation for infrastructure facility financing, development, operation, and maintenance. Though no specific analysis or calculations were performed as part of the study, databases identified for implementing the framework include utilizing HPMS for estimating VMT, and the Carload Waybill Sample, Air Carrier Statistics, and waterborne data for estimating mode specific services and market shares.

3.9 Freight Transportation and Land Use Planning

For NCFRP Project 47, freight transportation and land use planning has been defined as the coordination of regional transportation plans with land use development. An effective and well-integrated freight and land use plan results in both public and private sector benefits, such as reduced congestion, improved air quality and safety, enhanced community livability, improved operational efficiency, reduced transportation costs, and greater access to facilities and markets (Hartshorn and Lamm 2012). Examples of data required in performing this function include data on truck trip generation, delivery tours, transportation network characteristics, and economic characteristics and spatial distribution of participating agents. The most commonly cited databases found in the literature were CFS, U.S. Customs and Border Protection (CBP), Dun and Bradstreet, and industry-related databases. These data sources often were combined with locally collected data (often via surveys).

For example, nationally available data sources were cited for freight transportation and land use planning in *NCHRP Report 739/NCFRP Report 19: Freight Trip Generation and Land Use*. This study found that some freight trip generation data relating to land use was collected over the years, but most of the data was either outdated or insufficient for current planning needs (Holguín-Veras et al. 2012). As part of the study, shipper and carrier surveys were conducted and used in the development of the freight trip generation models. *NCHRP Report 739/NCFRP Report 19* provides additional information on the freight and land use data needs and suggests data collection techniques to address those needs. The report cites CFS and zip code business patterns as useful sources of data for freight trip generation modeling. It further recommends the use of the CFS micro-data to estimate commodity movement parameters for freight demand models.

3.10 Environmental Planning

Environmental planning involves activities such as mobile emissions planning, environmental protection, land use management, and air quality efforts. Freight-related data required for environmental planning includes vehicle type and vehicle trip information, route information, air quality data, and network information. EPA's Motor Vehicle Emission Simulator (MOVES) model is generally utilized in performing regional emission modeling studies (EPA 2014). Input data required by MOVES includes vehicle population, age distribution, VMT by vehicle type, and average speed distribution, among others. VMT data from HPMS is the primary source; however, the required speed data for MOVES is taken from other sources, such as INRIX or GPS equipment (Eisele et al. 2013c; Boriboonsomsin et al. 2012).

These examples from the literature demonstrate the use of nationally available data sources for environmental planning:

- Ostria (1996) developed a methodology by which intercity trucking emissions can be assessed using emission factors documented in state implementation plans (SIPs) and data from VIUS (formerly TIUS—Truck Inventory and Use Survey). Using the gross vehicle weight classification and area of operation variables housed in VIUS, intercity VMT were calculated, and the disaggregated emission estimates reported in SIP documents were utilized in isolating the intercity freight VMT.
- Vanek and Morlok (1998) estimated total freight energy consumption for a range of commodity groups using an activity-based approach to energy consumption. Total freight activity was decomposed into components by mode and by commodity group, and each component was multiplied by an intensity estimate to calculate total energy use for that commodity group. Fourteen commodity groups as defined in the CFS were used, and total energy use for each commodity group was based on the modal volumes for truck, rail, truck-rail intermodal, marine, and air.
- Ang-Olson and Cowart (2014) explored current and future air quality effects that result from the development of North American trade and transportation corridors, and strategies to mitigate their impacts. The analysis focused on five specific binational corridor segments:

Vancouver–Seattle, Winnipeg–Fargo, Toronto–Detroit, San Antonio–Monterrey, and Tucson–Hermosillo. Current and future levels of trade, transportation, and emissions were estimated for each corridor segment using commodity flow and traffic volume data. Commodity flows, developed from an analysis of the Transborder surface freight data, were used to analyze trade origin and destination patterns, changes in trade levels in particular industries, changes in vehicle size and weight, and shifts in mode share. The Transborder surface freight data was supplemented with traffic volume data for cross-border truck and rail movements from the U.S. Customs Service, the Canada Customs and Revenue Agency, and private bridge and tunnel operating authorities.

- Corbett et al. (2010) developed the California Geospatial Intermodal Freight Transportation (GIFT) model to analyze energy and environmental impacts of goods movement through California's marine, highway, and rail systems. The GIS-based model incorporates information from energy and environmental variables into segments of the national highway, rail, and waterway network, to enable the reporting of environmental performance measures associated with freight flows on the network. It also enables the comparison of alternative cargo flow patterns that minimize energy consumption and emissions when least cost or shortest path routes are considered. Road, rail, and waterway network features and facility locations used in GIFT are from NTAD, U.S. Army Corps of Engineers (USACE) waterways, and other private and public data sources. O-D Freight Flow Data used in this study were from the FAF version 2 (FAF2) and CFS data sets, which were supplemented with USACE Entrance and Clearance data. The Entrance and Clearance data, which contains a vessel's International Maritime Organization identification number, can be used in quantifying the volume of container traffic entering and leaving a port. When linked to data compiled by classification societies such as Lloyd's Registry of Ships, operational characteristics of vessels can be further examined.

3.11 Regulation and Enforcement Planning

Regulation and enforcement planning seeks to improve the safety of freight operations through the implementation and management of activities such as vehicle licensing, inspections, size and weight specifications, work hours regulation, and taxes/fares. Databases found in the literature for performing this function include the Motor Carrier Management Information System (MCMIS) and FARS.

These examples from the literature demonstrate the use of nationally available data sources for regulation and enforcement planning:

- Gillham et al. (2013), in cooperation with FMCSA, developed the Intervention Model to measure the effectiveness of roadside inspections and traffic enforcement in terms of crashes and injuries avoided, and lives saved. Roadside inspections as recorded in the MCMIS database are converted into crash risk probabilities with the assumption that an inspection violation implies a certain degree of crash risk. Thus, for each inspection that is uncovered and corrected, it is assumed that there is a reduced risk of an accident occurring. "By summing the crash risk probabilities for all violations corrected over all inspections, the model estimates the number of crashes avoided as a result of the FMCSA Roadside Inspection and Traffic Enforcement Programs" (Gillham et al. 2013).
- Dang (2007) used FARS and state data files to determine the effectiveness of Electronic Stability Control (ESC) systems in reducing fatal run-off-road crashes and vehicle rollovers. Vehicle identification number (VIN) data from the crash files were matched with VIN data obtained from the Passenger Vehicle Identification Manual (published annually by the National Insurance Crime Bureau) to obtain vehicle make, model, and year information. The final analysis database contained records of each vehicle involved in a crash and the vehicle make, model,

and year. A series of statistical analyses were then performed with an emphasis on testing the effectiveness of ESC systems.

3.12 Intermodal Trade Corridor Planning

For purposes of NCFRP Project 47, intermodal trade corridor planning was defined as the monitoring and development of policies and strategies that facilitate the efficient movement of goods by a variety of modes along key national and international trade corridors. Data sources cited in the literature included Transborder, Foreign Trade Database, USA Trade Online, and the Bureau of Transportation Statistics (BTS).

These examples from the literature demonstrate the use of nationally available data sources for intermodal trade corridor planning:

- Figliozzi et al. (2001) examined alternate methods for estimating loaded NAFTA truck volumes between the United States and Mexico. The first method utilizes truck volume counts reported in the Transborder database and estimates loaded trucks by applying a factor of empty trucks to the total number of trucks crossing the bridge. Corrections for intermodal freight shipments, single-unit trucks, and local trade are also applied. In the second method, truck volumes are estimated using U.S. international trade data from the Foreign Trade Database. Truckload weight per commodity is calculated by multiplying a commodity group's density by the capacity volume of various truck types, based on the assumption that high density commodities will weigh out and low density commodities will cube out.
- Harrison et al. (2010) examined major trends in intermodal shipping that impact Texas intermodal trade corridors, including an analysis of key supply-and-demand forces that underpin intermodal service and routing options. A review of current and future trade corridors used for handling international intermodal trade was performed to show the comparative strengths and weaknesses of different routing options for intermodal cargo shipping. Transborder trade data from BTS and foreign trade data from USA Trade Online were used in examining trade patterns between Texas and its top trading partners.

3.13 Terminal and Border Access Planning

Terminal and border access planning involves the management and maintenance of intermodal freight terminals and border facilities to ensure efficient movement of goods and people. Related topics in this area include border and port congestion, security, terminal access to rail and port facilities, and port efficiency and throughput. Studies in this area tend to utilize project-specific survey data for analytical purposes. Data on cargo volume and terminal traffic can be obtained from terminal operators and port authorities (Shafran and Strauss-Wieder 2003).

These examples from the literature demonstrate the use of nationally available data sources for terminal and border access planning:

- Turnquist and Rawls (2010) developed a multimodal network model that performs a vulnerability assessment of border trade flows to disruptions at one or more of the major bridges and tunnels on the border between the United States and Canada. The model's O-D table was estimated using freight flows from the Transborder database and validated with Canadian data and bridge-specific truck and count information. The assessment is performed by constructing a multimode equilibrium model of international freight flows in the Lake Erie corridor, and the model is subjected to disruptions (closure of one or more border-crossing facilities), which results in shifts in traffic flow and congestion in the remaining facilities. The economic costs of disruptions are then measured.

- Bhamidipati and Demetsky (2009) described a general methodological framework for evaluating the impacts of intermodal terminals on the transportation system and applied it to the highway system of Virginia. Models developed for the framework were calibrated with commodity flow, socioeconomic, and other data for the Commonwealth of Virginia. Data sources included Transearch, Carload Waybill Sample data, the Oak Ridge National Laboratory GIS database (now part of the NHPN), Census data, IMPLAN County Wise Employment Data, and the Virginia DOT's Crash Database and GIS Integrator. Truck diversions were then estimated from the data by using distance/travel time models and discrete choice models to estimate freight demand and drayage activities on the network.

3.14 Hazardous Materials Planning

Hazardous materials planning involves improving safe movement and monitoring of hazardous materials transported on the freight network. Hazardous materials movement data, such as what is reported in the CFS, is used in policy development, the rule-making process, program planning, and identification of emergency response and preparedness needs (Duych et al. 2011). The databases most frequently cited within the literature included the Pipeline and Hazardous Material Safety Administration's (PHMSA) database, the Hazardous Materials Incident Reporting System (HMIRS), and CFS. Estimates of daily hazardous materials shipments can be derived and the safety of the hazardous materials transportation assessed from the CFS data (Duych et al. 2011).

These examples from the literature demonstrate the use of nationally available databases for hazardous materials planning:

- Restrepo et al. (2009) examined the causes and economic consequences of hazardous liquid pipeline accidents in the United States. Data on accidents related to hazardous liquid pipelines from the PHMSA was utilized in the analysis. Regression models were used to determine what factors were associated with product-loss cost, property damage cost, and cleanup and recovery costs. Factors examined included the system part involved in the accident, location characteristics, and type of incident.
- Ellis (2011) examined factors contributing to the release of packaged or dangerous containerized goods during marine transport. Data from the HMIRS database and the UK's Marine Accident Investigation Branch accident databases were utilized in identifying factors contributing to the release of dangerous containerized goods.

3.15 Roadway Pavement and Bridge Maintenance Planning

Roadway pavement and bridge maintenance planning involves examining the effects of freight movement, typically truck traffic and heavy vehicles, on the pavement and bridge infrastructure. Several studies focused on examining the structural integrity of bridges or pavement condition and design and were found to utilize data collected through controlled lab experiments. Databases frequently cited in the literature for roadway pavement and bridge maintenance planning include HPMS for traffic volume and pavement condition information, the NBI database for bridge information (Jansuwan et al. 2010), and FAF and Transearch for commodity flow data.

These examples from the literature demonstrate the use of nationally available data sources for roadway pavement and bridge maintenance planning:

- Fortowsky and Humphreys (2006) examined the cost impact of higher truck weight limits being allowed on Interstate routes in Maine. Two methodologies were developed for the assessment. The first methodology estimated the changes in freight truck traffic volumes as a result of the increased weight limits, and subsequent changes in VMT, truck configurations, and equivalent single-axle loads (ESALs). The second methodology estimated road cost per ESAL using the predetermined ESAL calculations from the first methodology. Truck O-D flows were estimated using Transearch data and supplemented with weigh-in-motion station data, vehicle classification counts, and network data from the Maine DOT's TIDE road database system.

- Schroeder et al. (2012) developed a freight-based prioritization framework to identify freight infrastructure needs critical to maintaining economic vitality. The framework incorporates economic metrics associated with infrastructure performance and level of service. Using BEA data, an input-output model is used to identify transportation-dependent industrial sectors, which are then linked with FAF commodity flows. A set of conditional performance measures is selected to identify critical locations meriting improvements, including (1) the structural integrity of bridges from NBIAS outputs; (2) IRI from HPMS; (3) truck fatality crash rate and truck crash rate from the Virginia DOT crash database; and (4) other geometric standards deficiencies. The framework's outputs are a prioritized list of economically critical highway infrastructure needs selected in consideration with regional economic impacts, safety, and mobility improvements.

3.16 Modal Shift Analysis

Modal shift is recognized to occur "when one mode has a comparative advantage in a similar market over another" (Rodrigue et al. 2013). Incentives for modal shift in freight networks include cost savings, travel time reductions, network reliability, and implementing strategies to mitigate energy usage and greenhouse gas emissions (Nealer et al. 2012; Eisele et al. 2012). Feasible modal shift considerations for freight are mainly composed of shifts between road (trucks), rail, and water modes (Corbett et al. 2010). Modal shift analysis includes not only measuring the impact of shifting from one mode to another but also examining the policies and incentives that foster modal shift changes. The most commonly cited nationally available freight-related databases for modal shift studies are the FAF, CFS, and PIERS.

These examples from the literature demonstrate the use of nationally available data sources for modal shift analysis:

- Nealer et al. (2012) compared energy usage and emissions across multiple freight transportation modes to determine opportunities for modal shift. A transportation flow input-output model was developed for more than 400 U.S. economic sectors using freight transport data from CFS commodity categories and input-output use tables from the BEA. Sector-specific mode choice shifts were analyzed, and large-scale reductions in emissions and fuel consumption also were examined. Multiple scenarios were analyzed, including (1) quantifying the reasonable bounds of energy and emissions to be reduced by a complete modal shift from truck to rail; (2) determining foregone energy and emissions when modal shift occurs for the top 20% supply chain sectors; and (3) the effect of an increased truck efficiency on modal shift.
- The Maritime Administration study, *Impact of High Oil Prices on Freight Transportation: Modal Shift Potential in Five Corridors* (Transportation Economics & Management Systems 2008), sought to evaluate the impact of oil prices on U.S. domestic freight transportation and assess how prices impact transportation logistics chains. The analysis was performed using the GOODS™ demand and supply model, which is calibrated to identify the potential for waterborne transportation to capture containerized traffic. FAF O-D traffic data was used in calibrating the model's demand parameters. FAF data was further augmented with Transborder

data for cross-border flows and with Transport Canada data for Canadian domestic flows. Historical energy price data, short-term outlook data, and long-term growth rate scenarios developed by the Energy Information Administration (EIA) were used in developing the scenarios to be tested.

3.17 Freight Performance Measurement

Performance measurement is defined by the U.S. Government Accountability Office (2011) as "the ongoing monitoring and reporting of program accomplishments, particularly progress toward pre-established goals." Freight performance measures examine the transportation system's efficiency, safety, and condition in meeting freight demand, including the impact on energy use and the environment (Gordon Proctor & Associates et al. 2011). It provides a greater insight into the performance of the current transportation system and allows agencies to "rank capital investments and evaluate alternative programs," "provide a rationale for allocating resources," and "assist in monitoring progress toward achieving specific transportation goals and targets" (Prozzi et al. 2011).

Freight performance measures found in the literature tend to rely on disaggregated data from sources such as GPS devices for monitoring truck movements (Sage et al. 2013). These data sources tend to be proprietary in nature; however, the American Transportation Research Institute (ATRI), in partnership with FHWA, provides aggregated truck GPS data to evaluate travel time and travel time reliability measures along critical freight corridors (ATRI 2014).

McMullen et al. (2010) and *NCFRP Report 10: Performance Measures for Freight Transportation* (Gordon Proctor & Associates et al. 2011) identified several national and readily available freight databases that can assist in developing performance measures (Table 3-3). Though these may not be sufficient in meeting current demands, knowledge of their capabilities and current limitations can inform how future data collection efforts should be tailored to supplement these data sources. Freight network and node data sources also are available from the National Transportation Atlas, Waterways Facilities data, and FAF network databases.

3.18 Sustainable Transportation Investment

Sustainable transportation investment involves investigating ways to fund both the existing transportation system and future projects. For example, nationally available data sources were utilized for sustainable transportation investments in the development of a baseline roadmap by the California Hybrid, Efficient and Advanced Truck Research Center's (CalHEAT) Research and Market Transformation Roadmap for Medium- and Heavy-Duty Trucks. The roadmap is intended to guide the advancement and demonstration of efficient truck technologies and systems to meet or exceed the 2020 goals for California in air quality, energy security, petroleum reduction, and greenhouse gas reductions. Data sources used for the vehicle technologies characterization map and baseline included the R.L. Polk Co., U.S. Census, 2002 VIUS, 2009 Climate Registry Reporting Protocol, a 2008 CARB truck and bus study that used department of motor vehicle (DMV) data, the *Transportation Energy Data Book* (TEDB), and fuel use estimates from Argonne National Lab (California Energy Commission 2013). The initial truck inventory study used a baseline inventory from R.L. Polk consisting of 2009 vehicle registration data from 1.5 million commercial medium- and heavy-duty trucks. The vehicles were then grouped by weight and application. Additional data was gathered on average VMT, fuel consumption, and emissions per mile to determine average fuel use, NO_x, and CO_2e emissions for each of the truck categories. These averages were then multiplied by the vehicle population inventory to develop baseline fuel consumption, CO_2e, and NO_x by average VMT and vehicle category (California Energy Commission 2013).

Table 3-3. Freight performance measurement data sources.

Performance Measure	Potential Source of Data
1. Safety	
Highway	Accident Crash Reporting Systems (state level)
	Fatality Analysis Reporting System
	Motor Carrier Management Information System
	Safety Measurement System
	Safety and Fitness Electronic Records
Rail	FRA State Freight Rail Safety Statistics
Air	Accident/Incident Data System
	Aviation Safety Reporting System
	Near Midair Collision System
	Runway Safety Office Runway Incursion Database
Ports/Marine	Marine Information for Safety and Law Enforcement
2. Maintenance/Preservation	
Highway	Pavement Management System (state level)
	National Bridge Inventory
Rail	Rail Network Data (state level)
Air	Airport Pavement Management System (state level)
Ports/Marine	USACE Navigation Data Center
3. Mobility, Congestion, and Reliability	
Highway	Highway Performance Measurement System
	ATRI N-CAST
	INRIX Probe Vehicle Data
	Weigh-in-motion Data
Rail	Association of American Railroads' Railroad Performance Measures
Air	Air Carrier Statistics
Ports/Marine	USACE Lock Performance Measurement System
	Maritime Safety and Security Information System
	Port Import and Export Reporting System
4. Accessibility and Connectivity	
Highway	State, regional, or MPO-level GIS databases
Rail	Carload Waybill Sample
Ports/Marine	USACE Lock Performance Measurement System
Air	Air Carrier Statistics
Commodity Flow Data	State-level commodity flow models
	Freight Analysis Framework
	Transearch Database
	Commodity Flow Survey
5. Environment	
Highway	The EPA's MOVES2010

Source: Adapted from McMullen et al., 2010, and *NCFRP Report 10.*

3.19 Findings from the Literature Review on Freight Data Uses

The literature review identified key studies showing innovative and unique examples of how available freight data sources are utilized by agencies and the research community. The following points summarize the findings from the review:

- Studies limited by the availability of disaggregated data typically rely on field- or project-specific survey data to perform the task at hand. Data from these project-specific studies, though relevant, is rarely available to or accessible by others on completion of a study.
- Several studies used state and regional databases and/or models to either supplement or replace nationally available freight data sources, where possible.
- Additional sources of data utilized by practitioners include data from local and regional planning agencies, marine port and airport authorities, and industry sources.
- Data sources such as the CFS and FAF, though popular, tend to be outdated for performing specific tasks—a limitation cited in several of the studies reviewed.
- Data reliability and validity of nationally available freight data sources remain items of concern. Thus, there is a shift toward the use of relatively new and more reliable intelligent transportation system (ITS)-related data sources such as GPS data and vehicle-to-infrastructure connected devices.
- Several freight-related models were found to be theoretical in nature, requiring data that is currently either unavailable or insufficient and, therefore, necessitating certain assumptions and transformations for the models to be used.
- A need may exist for a central data collection repository at which locally collected or project-specific data can be stored or shared with other data users in the transportation community. These project-specific data sources could complement currently available freight data sources as well as provide additional opportunities to test or validate freight-related models.

CHAPTER 4

Inventory of Freight Data Sources, Dictionaries, and Glossary Terms

4.1 Introduction

Public and private agencies collect data relating to freight transport to meet their specific needs. This study identified 42 data sources, comprising 31 public sources and 11 commercially available sources (Table 4-1 and 4-2). The following details are provided from 25 of the 31 public sources:

- **Overview:** A brief description of the data source and contents.
- **Coverage:** The extent or degree to which data was collected, analyzed, or reported.
- **Availability:** The time period that each data source covers and the frequency at which data is collected or updated.
- **Uses:** A brief description of how each data source is currently being utilized as reported by the agencies.
- **Data Tables:** A summary of identified databases and data tables for each data source and their years of availability.
- **Data Collection Method and Limitations:** A brief summary of data collection procedures, sample design, statistical estimation, and other related data processing and quality control procedures. Readers are referred to the original data source, user guides, or manuals to acquire additional information.
- **References:** A list of web addresses to user manuals, data download ports, location of data dictionaries, and other useful or recommended reading materials.
- **Data Provider and Contact:** Information about the data-providing agency.

Limited information also is documented for the remaining six public and 11 commercial data sources. Agency reports generated from the actual data sources are excluded from the review. Some data sources were found to contain multiple databases, data tables and associated data dictionaries or glossary terms. This chapter provides additional information on how the data sources were inventoried.

4.2 Data Dictionaries and Glossary Terms

Data elements from data dictionaries and glossaries similar to those shown in Figure 4-1 and Figure 4-2 were identified. Data dictionaries from 28 sources were compiled, including two commercial sources, and the total number of data elements included in the *master data dictionary* for NCFRP Report 35 was 6,322. In addition, 13,554 glossary terms from 13 glossaries were compiled into a *glossary* for this project.

4.2.1 Definition of Terms

- As used in *NCFRP Report 35*, the phrase *data source* refers to the actual name given by an agency to its data. It is important to note that a data source may contain multiple databases. For example,

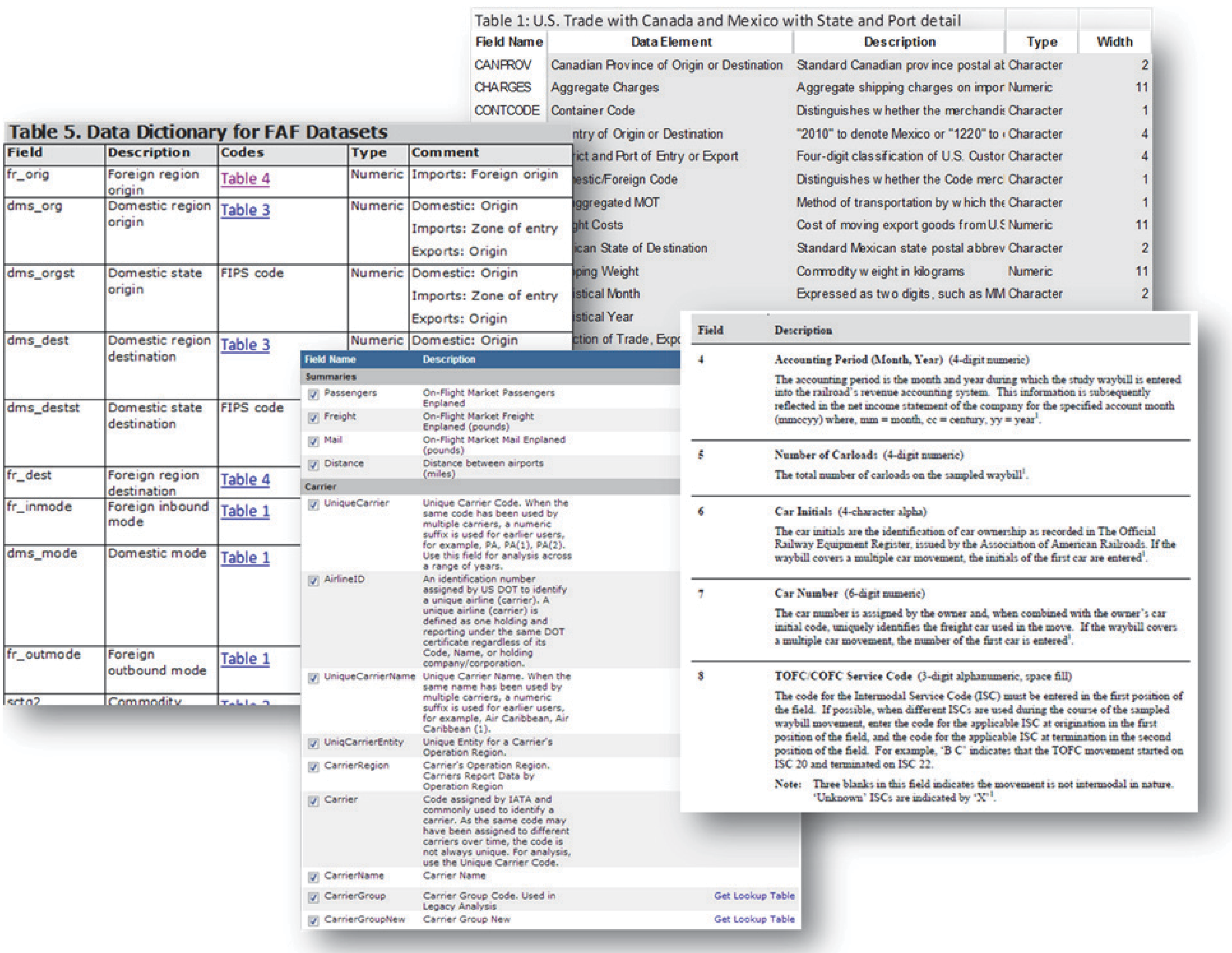
Table 4-1. Identified public freight data sources.

	Public Freight Data Source	Agency
1	Air Carrier Statistics	U.S. DOT - RITA - BTS
2	Annual Survey of Manufacturers	U.S. DOC - Census Bureau
3	Carload Waybill Sample	Surface Transportation Board
4	Commodity Flow Survey	U.S. DOT - RITA - BTS
5	County Business Patterns	U.S. DOC - Census Bureau
6	EIA Data Services	U.S. DOE - EIA
7	Fatality Analysis Reporting System	U.S. DOT - NHTSA
8	Federal Railroad Administration Safety Database	U.S. DOT - FRA
9	Foreign Trade	U.S. DOC - Census Bureau
10	Freight Analysis Framework	U.S. DOT - FHWA
11	Highway Performance Monitoring System	U.S. DOT - FHWA
12	Pipeline and Hazardous Material Safety Administration	U.S. DOT - PHMSA
13	Maritime Statistics	U.S. DOT - MARAD
14	Motor Carrier Management Information System	U.S. DOT - FMCSA
15	Motor Carrier Safety Measurement System	U.S. DOT - FMCSA
16	National Agricultural Statistics Service	USDA - NASS
17	National Highway Planning Network	U.S. DOT - FHWA
18	Survey of Business Owners	U.S. DOC - Census Bureau
19	Service Annual Survey	U.S. DOC - Census Bureau
20	Topologically Integrated Geographic Encoding and Referencing	U.S. DOC - Census Bureau
21	Transborder Freight Database	U.S. DOT - RITA - BTS
22	U.S. Economic Accounts	U.S. DOC - BEA
23	U.S. Waterway Data	USACE - Waterborne Commerce
24	Vehicle Inventory and Use Survey	U.S. DOC - Census Bureau
25	Vehicle Travel Information System	U.S. DOT - FHWA
	Additional Public Freight Data Sources*	**Agency**
26	Air Carrier Financial Reports	U.S. DOT - RITA – BTS
27	Business Dynamic Statistics	U.S. DOC - Census Bureau
28	Statistics of U.S. Businesses	U.S. DOC - Census Bureau
29	Transportation Services Index	U.S. DOT - RITA – BTS
30	U.S. Highway Statistics Series	U.S. DOT – FHWA
31	Workforce Information Database (structure only)	Analyst Resource Center

BTS = Bureau of Transportation Statistics; EIA = Energy Information Administration; MARAD = United States Maritime Administration; USACE = U.S. Army Corps of Engineers; U.S. DOC = U.S. Department of Commerce.
*These publicly available sources were identified but are not included in the discussions.

Table 4-2. Identified commercial freight data sources.

	Commercial Freight Data Source	Agency
1	Dun and Bradstreet Hoovers Database	Dun and Bradstreet
2	FleetSeek	Fleet Owner Magazine
3	IMPLAN Data Files	IMPLAN Group LLC
4	InfoUSA	InfoGroup
5	Intermodal Association of North America Data and Statistics	Intermodal Association of North America
6	Lloyd's Marine Intelligence Unit	Lloyd's List Intelligence
7	Motor Carrier Annual Report	American Trucking Association
8	Port Import Export Reporting Service	United Business Media Global Trade
9	State of Logistics Report	Council of Supply Chain Management Professionals
10	Transearch	IHS Global Insight
11	Woods and Poole Economics	Woods and Poole Economics, Inc.



Clockwise from top: FAF3, Transborder, Carload Waybill Sample, Air Carrier Statistics.

Figure 4-1. Examples of available freight data element dictionaries.

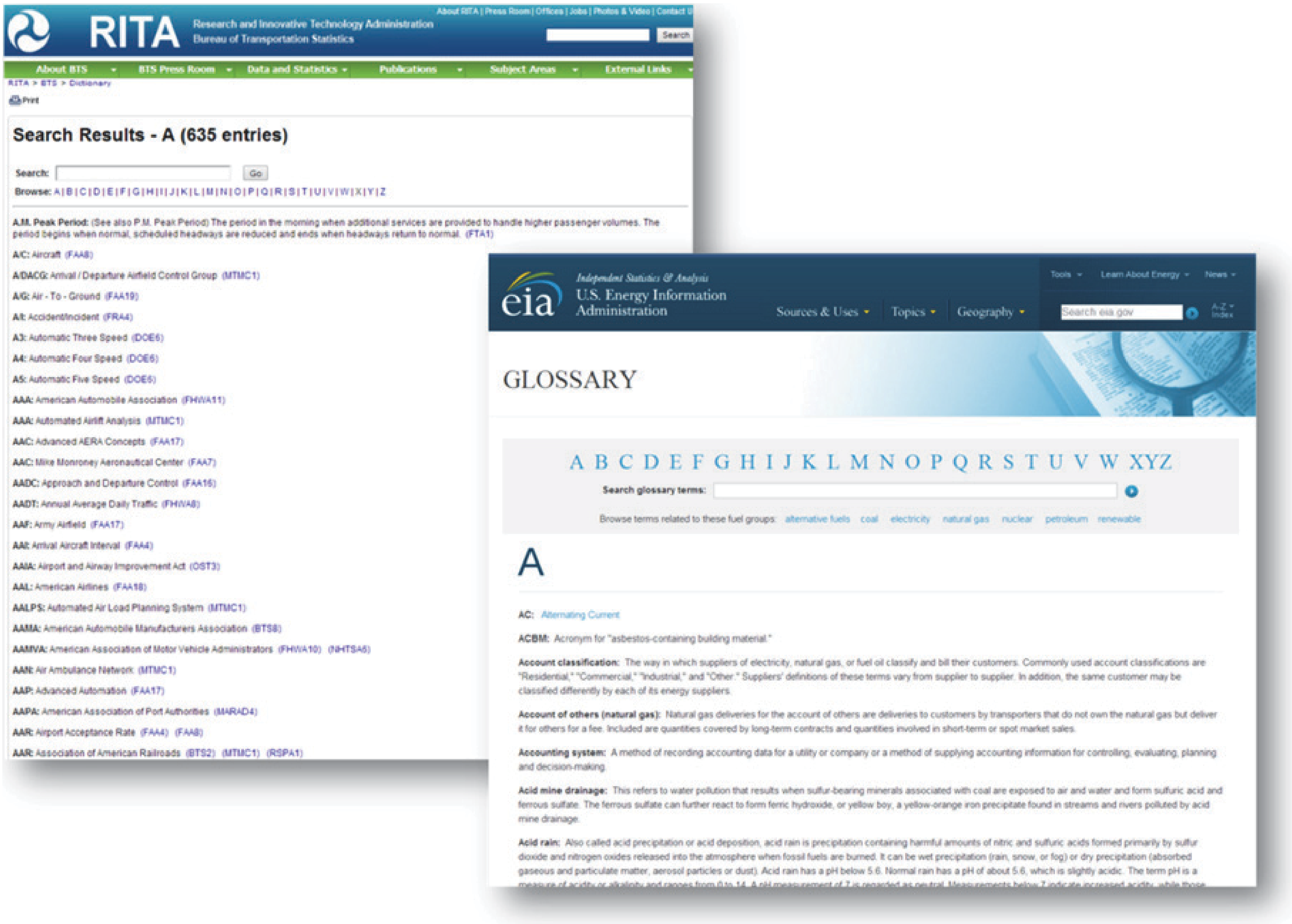


Figure 4-2. BTS and EIA glossaries.

the Freight Analysis Framework (FAF), which is treated in this guide as a data source, is made up of multiple databases (i.e., regional databases, state-level databases and a network database). A database may also have multiple tables, each containing data elements and records. Finally, over time a database may be made available in updated versions. In Figure 4.1, for example, the version of the FAF used to generate the image included at the top of the figure is FAF3).

- As adopted for this report, the term *data element dictionary* (*data dictionary*) is defined in the *IBM Dictionary of Computing* as "a centralized repository of information about data such as meaning, relationships to other data, origin, usage, and format" (IBM 1993). A data dictionary "describes, defines, and lists all of the data elements that are stored in a database" (General Services Administration 1996). A typical representation of a data dictionary will include information such as elements included in the database, type/format (e.g., numeric, text, alphanumeric), description/definition of element, possible values or scope, relationship with other elements or tables, and metadata such as comments on the quality and condition of the data. Examples of data dictionaries include the FAF3, Transborder, Carload Waybill Sample, and Air Carrier Statistics dictionaries (see Figure 4-1).
- As adopted for this report, the term *glossary* is defined in the Oxford dictionary as "an alphabetical list of words relating to a specific subject, text, or dialect, with explanations" (Oxford Dictionaries 2014). Examples include the BTS Dictionary and EIA Data Services Glossary websites, which are shown in Figure 4-2.

```
Data Source A*
 |--- Database A.1
      |--- Sub Database
           |--- Table A.1.1*
                |--- Data Element I*
                     |--- name, definition, type, … *
                |--- Data Element I
                     |--- name, definition, type, …
                |--- Data Element II
                     |--- name, definition, type, …
           |--- Table A.1.2*
                |--- Data Element …
 |--- Database A.2
      |--- Sub Database
           |--- Table A.2.1
                |--- Data Element …
* required fields
```

Figure 4-3. Data structure of master data dictionary and glossary.

- As adopted for this report, the term *data element* is defined by the *Federal Standard 1037C* as “a named identifier of each of the entities and their attributes that are represented in a database” (General Services Administration 1996). For each data element, associated information such as element name, type/format, description of element, example values, reference to other data tables or sources, comments, and other relevant information may be available.

Based on these definitions, the basic structure of the information stored in the master data dictionary and glossary is illustrated in Figure 4-3. For each data source, the minimum required entities are the data source name, a table containing the elements, and the elements themselves. A data source may have multiple databases and tables (as shown in Table 4-3 for the *glossary*); data element name and definition are required. Table 4-4 provides information on the glossaries and the number of elements contained in each glossary.

Tables 4-5 and 4-6 show the data element properties and their definitions as stored in the master data dictionary and glossary tables, respectively.

4.2.2 Recommended Data Types

Given the variability of the data types reported in the various data element dictionaries, the study team developed a uniform set of *recommended data types.* The recommended data types seek to assist data users in determining how to correctly use each data element for research and analysis. The uniformly designated data types and their definitions are listed in Table 4-7.

Recommended data types are specified for each data element in the master data dictionary. The field containing the recommended data types appears next to the originally reported data types, which were sometimes considered ambiguous when describing the context in which a data element can be used (see Figure 4-4). For example, a data dictionary will specify a data element field as numeric but it may differentiate whether the numbers represent a name of a place or an actual measured value. When used with the data element definitions, the recommended data types provide an additional level of clarity on how to correctly apply (or not apply) a chosen form of statistical analysis to a data element set.

Table 4-3. Data dictionaries.

	Public and Commercial Data Sources	Number of Tables	Number of Elements
1	Air Carrier Statistics	12	504
2	Air Carrier Financial Reports	12	478
3	Annual Survey of Manufacturers	4	62
4	Border Crossing/Entry	1*	5
5	CTA Intermodal Terminals Database	2	12
6	Carload Waybill Sample	2	252
7	Commodity Flow Survey	2	18
8	County Business Patterns	7	322
9	Fatality Analysis Reporting System	18	310
10	Federal Railroad Administration Safety Database	6	503
11	Foreign Trade	32	389
12	Freight Analysis Framework	2	70
13	Highway Performance Monitoring System	2	117
14	IHS Transearch	2	30
15	Motor Carrier Management Information System	22	358
16	Motor Carrier Safety Measurement System	1*	32
17	National Agricultural Statistics Service	4*	38
18	National Ballast Information Clearinghouse Database	3*	38
19	National Corridors Analysis and Speed Tool Database	1*	21
20	North American Transborder Freight Database	5	66
21	Pipeline and Hazardous Material Safety Administration	1*	33
22	Service Annual Survey	1	28
23	Survey of Business Owners	1	198
24	Topologically Integrated Geographic Encoding and Referencing	15	483
25	U.S. Waterway Data	11	266
26	Vehicle Inventory and Use Survey	1	242
27	Vehicle Travel Information System	9	207
28	Woods and Poole Economics, Inc.	2	1240
	Total	181	6,322

*Element names were extracted from web forms.

Table 4-4. Glossaries.

	Public and Private Glossaries	Number of Elements
1	Air Carrier Financial Report Glossary	29
2	BEA Glossary	272
3	Border Crossing/Entry Data	12
4	Commercial Vehicle Information Systems and Networks Glossary	453
5	Economic Census Definitions (Census Bureau)	65
6	EIA Glossary	2,579
7	Freight Glossary and Acronyms (FHWA)	166
8	Glossary of Shipping Terms (Maritime Administration)	832
9	IMPLAN Glossary (IMPLAN)	207
10	Intermodal Glossary (IANA)	197
11	State of Logistics Report Glossary (CSCMP)	2,461
12	Topologically Integrated Geographic Encoding and Referencing	12
13	Transportation Expressions and Transportation Acronym Guide	6,069
	Total	*13,354*

Table 4-5. Master data dictionary table.

	Field	Description
1	Data Source	The data source name (e.g., Air Carrier Statistics, FAF, Foreign Trade, etc.).
2	Database	A database contained in the data source, if available (e.g., Air Carrier Statistics has two databases: U.S. Carriers and All Carriers. See Appendix A for examples).
3	Sub-Database	A sub-database of the database, if applicable (e.g., U.S. Waterway Data includes 10 databases and 11 sub-databases. See Appendix A for examples).
4	Table	Each table can be considered as the "data dictionary" for a specific group of elements.
5	Data Element Name	The name of the data element in the dictionary.
6	Alias	Any published secondary name of the data element that slightly differs from the Data Element Name.
7	Definition	A readable phrase or sentence associated with a data element within a data dictionary that describes the meaning or semantics of a data element.
8	Additional Definition	Additional definition information, if available in the data dictionary.
9	Reported Data Type	Data type as reported in data dictionary. Examples include character, variable character (varchar), numeric, and text.
10	Recommended Data Type	A uniformly categorized set of fields (data types) for use in data elements. See Table 4-7.
11	Unit Tag	Unit of measurement as determined from the data element definition.
12	Range of Values	Possible values for this data element as provided in the data dictionary.
13	Comments	Any additional comments concerning the data element field either made in the original data dictionary or included by the study team.
14	Primary Element Role	This field is discussed in Chapter 5.
15	Secondary Element Role	This field is discussed in Chapter 5.

Table 4-6. Glossary table.

	Field	Description
1	Data Source	The data source name (e.g., Air Carrier Statistics, FAF, Foreign Trade, etc.).
2	Database	The database containing the data element, if available.
3	Glossary Term	The name of the term as it appears in the glossary. Glossary terms include acronyms and abbreviations listed in the glossary.
4	Definition	An explanation of the meaning of the glossary term.

Table 4-7. Recommended data types.

	Data Type	Description
1	Nominal	Fields whose values exist in name only and can be counted but not measured. Examples include texts, labels, categories, highway number, city name, commodity code, zip code, and contact information.
2	Binary	Fields whose values are composed of or involve two things. Examples include 0/1, true or false, and yes or no.
3	Date/Time	Fields that report on the time of the day, day of the week, day of the month, year, or time period.
4	Real Number	Fields whose values can be measured. Real numbers are used mainly for fields that can be represented in non-whole numbers (e.g., decimals). Examples include tonnage, miles, accidents per vehicle-mile, etc.
5	Integer	Fields whose values are expressed only in whole numbers (not fractions). Examples include number of trucks, average annual truck traffic, number of containers, number of accidents, etc.
6	Currency	Fields that represent monetary values. An example is the value of commodities moved in U.S. dollars.
7	Ratio	Fields that report on a relationship between two numbers of the same kind. An example is the ratio of passenger miles to available seat miles, which is reported as "Load Factor" in the Air Carrier Statistics database.
8	Percentage	Fields whose values are numbers or ratios expressed as a fraction of 100. Examples include percentage of truck traffic, percentage of total sales, and so forth.
9	Geometry	Fields used to represent data found in GIS databases. Examples include point, line, and polygon.

1	DEFINITION	ADDITIONAL_DEFINITION	REPORTED_DATA_TYPE	RECOMMENDED_DATA_TYPE	UNIT_TAG	Range of Values	PERMALINK
9	Unique Entity for a Carrier's OpeRation Region.		Numeric	Nominal			
10	Carrier's OpeRation Region. Carriers Report Data by OpeRation Region		Text	Nominal			
11	Code assigned by IATA and commonly used to identify a carrier. As the same code ma		VarChar	Nominal			http://www.transt
12	Carrier Name		Text	Nominal			
13	Carrier Group Code. Used in Legacy Analysis		Numeric	Nominal			http://www.transt
14	Carrier Group New		Numeric	Nominal			http://www.transt
15	Origin Airport, Airport ID. An identification number assigned by US DOT to identify a ur		Numeric	Nominal			http://www.transt
16	Origin Airport, Airport Sequence ID. An identification number assigned by US DOT to id		Numeric	Nominal			http://www.transt
17	Origin Airport, City Market ID. City Market ID is an identification number assigned by U		Numeric	Nominal			http://www.transt
18	Origin Airport		VarChar	Nominal			http://www.transt
19	Origin City		Text	Nominal			
20	Origin State Code		Text	Nominal			http://www.transt
21	Origin State FIPS (U.S. Federal Information Processing Standard Codes)		Integer	Nominal			http://www.transt
22	Origin Airport, State Name		Text	Nominal			
23	Origin Airport, World Area Code		Numeric	Nominal			http://www.transt
24	Destination Airport, Airport ID. An identification number assigned by US DOT to identi		Numeric	Nominal			http://www.transt
25	Destination Airport, Airport Sequence ID. An identification number assigned by US DO		Numeric	Nominal			http://www.transt
26	Destination Airport, City Market ID. City Market ID is an identification number assigne		Numeric	Nominal			http://www.transt
27	Destination Airport		VarChar	Nominal			http://www.transt

Figure 4-4. Segment of master data dictionary showing reported data type and recommended data type columns.

CHAPTER 5

Classifying Data Elements Across Databases

5.1 Introduction

Freight data sources tend to be heterogeneous in terms of structure, syntax, and semantics (Buccella et al. 2003). *Structural*, or schematic, heterogeneity deals with differences in how the data is stored in the various databases (e.g., table schemas, primary and foreign keys, etc.). *Syntactic* heterogeneity deals with differences in the representation of the data; in other words, data types and formats (e.g., numeric, text, alpha-numeric values, categorical, etc.). *Semantic* heterogeneity, which is the most challenging to resolve, deals with differences in interpretation of the meaning of the data (Merriam-Webster 2014).

Cui and O'Brien (2014) classify semantic heterogeneity as follows:

- **Semantically Equivalent Concepts:** Different models use the same or synonymous terms to refer to the same concept; however, there may be differences in property types (e.g., the concept *weight* may be presented in tons in one model but in kilograms in another model).
- **Semantically Unrelated Concepts:** Different models use the same terms, but the terms have different meanings (e.g., the concept *channel* may mean ship channel in the U.S. Waterway database but mean traffic channelization device in the Federal Railroad Administration [FRA] Safety database).
- **Semantically Related Concepts:** Concepts may become generalized as they are classified across models; for example, the city "Austin, Texas," in the Air Carrier Statistics database is referenced in the commodity flow survey (CFS) as "Austin-Round Rock, Texas."

Resolving freight data heterogeneity across multiple data sources is required to facilitate the integration of data elements, enable interoperability between multiple systems, and smooth the exchange of data and information. Heterogeneity resolution first involves identifying which elements are related and vice versa. When dealing with more than 6,300 data elements, however, this process can be a tedious and time-consuming task.

To address this problem, a general freight data classification system was developed to categorize similar elements within each database, thus facilitating the identification of related data elements across multiple data sources. By first identifying related data elements, the process of determining the differences in data element definitions and resolving those differences through harmonization or statistical bridges becomes much clearer and more defined.

5.2 Background

A literature review identified practitioners' attempts to classify freight data, although no formal classification system currently exists for freight data elements. Specific applications and examples of how the classification schemes could be utilized for data integration and heterogeneity resolution across multiple freight data sources were not found. The classification schemes found in the review of the literature are described in this section.

To define key attributes of freight-related shipments, the TRB Committee on Freight Transportation Data (2003) coined the mnemonic CODMRT:

- **Commodity**, which describes the type of freight being moved and contains information such as value, weight, and handling characteristics.
- **Origin,** which describes the geographic starting point of a freight trip.
- **Destination,** which describes the geographic ending point of a freight trip.
- **Mode,** which describes the vehicles and infrastructure used to transport goods.
- **Route,** which describes the sequence of specific individual facilities (e.g., sections of roads, railroad tracks, etc.) that are used to transport freight between the origin and destination on a specific mode.
- **Time,** which is defined as the time period for which the freight data was collected (i.e., the freight forecast time period).

Ambite et al. (2004) classified data elements for multiple sources by representing each data item as a measurement that has values along a set of dimensions (e.g., geographic area, type of flow, mode of transport, type of product, time interval, value, and unit of measurement). The classification schemes for both CODMRT (2003) and Ambite et al. (2004) were found, however, to be limited to the commodity flow domain, and they do not capture elements from other freight data sources such as accident data and industry information.

Tok et al. (2011) developed a conceptual data structure for California that identified the relevant data set for a standardized national freight transportation data architecture. The high-level data elements defined in the data structure schema were time periods, time resolutions, zones, facility networks, commodities, modes, socioeconomic data, and logistics. Time-resolution data elements include items such as annual, quarterly, monthly, and daily time periods. Zones include items such as states, gateways, foreign, and trade regions. Facility networks contain information such as highway geography, rail geography, and waterways. The socioeconomic category considers elements such as employment and population, and the logistics category considers elements such as time, emissions, energy consumption, and safety (Tok et al. 2011). *NCFRP Report 14: Guidebook for Understanding Urban Goods Movement* (Rhodes et al. 2012) provides classified freight data sources in the following categories:

- **Freight node data,** which represents consolidated or individual endpoints that generate or receive freight flows and are the key points of production, consumption, or intermediate handling for goods.
- **Freight network data,** which defines major route patterns and critical infrastructure being used to convey freight shipments through the various modal systems.
- **Freight flow data,** which provides information on commodity flows and provides insight on the economic and trade environment of regions. Typical commodity flow records will contain information on the O-D of shipments, type of commodity, weight, and/or value of the commodity shipment, and mode of shipment.
- **Neighborhood freight data,** which provides information on safety, congestion, land use, and emissions.

Although both Tok et al. (2011) and Rhodes et al. (2012) addressed a broader range of freight data types as compared to CODMRT and Ambite et al. (2004), specific applications and examples of how the data structures can be mapped across data sources and utilized to resolve data heterogeneity across multiple data sources were not stated.

An XML schema such as TransXML (which was developed for the exchange of transportation data interoperability and dissemination) also is limited in scope. TransXML currently addresses four key business areas in transportation: (1) survey/roadway design, (2) transportation construction/materials, (3) highway bridge structures, and (4) transportation safety; however, it excludes other areas specific to freight movement. Multimodal freight movements (air, marine, rail, and pipeline), economic and census data, industry information, and commodity flow data schemas cannot be addressed with the current version of TransXML, and other standards, such as LandXML (LandXML.org 2000), Geographic Markup Language (GML), Geographic Information Framework Data Standard (Federal Geographic Data Committee 2008), and International Organization for Standardization (ISO) 14825:2011 Geographic Data Files (ISO 2014), were not developed specifically to address freight data.

As found in the literature search, freight data classification is mostly restricted to the commodity flow domain (e.g., CODMRT), and currently no agreed-upon classification system applies to all data elements from the various freight databases. Current data standards such as TransXML are limited in scope in terms of their representation of freight data, as they were developed mainly to address the exchange of transportation data and facilitate communication across multiple transportation industry stakeholders and agencies. The existing standards are inadequate to serve as a formal representation of the various data elements contained in multiple freight databases. For example, data elements that describe the freight industry, events that may occur during the transport of goods, and the role of human activity are currently not captured in these standards.

A generalized framework for classifying freight data elements across multiple data sources is therefore proposed. The proposed schema, called the **Role-Based Classification Schema (RBCS),** organizes and classifies data elements within their respective parent databases and facilitates the comparison, unification, translation, and fusion of data from multiple databases. The RBCS does not replace the existing standards; rather, it facilitates the process of identifying related data elements across a multitude of existing freight data sources. Data elements captured in the proposed schema can be used in advising the future development of existing standards (e.g., TransXML) to adequately capture all the existing freight data sources in their respective schemas.

5.3 Methodology

In developing the generalized classification system for representing freight data elements across multiple databases, an attempt was made to identify and group elements with similar "roles." For purposes of NCFRP Project 47, a *role* was defined as "the kind of information conveyed by an element or attribute" in its database. The researchers found that the roles of data elements in their respective databases could be used as a means for developing the RBCS classification system.

Two levels of classification groupings were identified: a top-level, primary grouping and a second-level grouping. The top-level grouping was based on an enumeration of multiple freight databases and the literature on freight data classification. Examples of freight data classification schema from the literature that were utilized in developing the top-level primary groups included those of CODMRT (2003), Ambite et al. (2004), and Tok et al. (2011). Data elements from freight data dictionaries were examined and a final list of nine top-level, primary groups was identified.

The second level of classification seeks to differentiate data elements that *identify* objects from data elements that *describe the features* of an object. This distinction became necessary as some elements were found to define entities that tend to be unique, whereas other elements were found to provide additional information about those identified elements. For example, a data element such as "origin ID" refers to or identifies a particular place, and the data element "population" describes the number of inhabitants living in that place. The distinction between these two types of elements is that only one "origin ID" can refer to a particular place in a database, but multiple places can each have "population" numbers, which are not necessarily unique.

RBCS first determines and assigns a **role** to each data element within its database based on the primary and secondary level classification of that data element. Grouping data elements within their respective databases simplifies the process of identifying similar data elements across multiple databases, as similar elements tend to fall within the same group. To validate the generality of RBCS, the classification schema was applied to all the freight data sources included in the master data dictionary to determine the generality of RBCS in successfully classifying data elements across those databases.

5.4 The RBCS Primary and Secondary Level Classifications

Nine top-level, primary groups were identified from examining the databases and reviewing the literature: **commodity, event, humans, industry, link, mode, place, time,** and **unclassified.** Figure 5-1 illustrates the inherent relationships that persist between the various data elements despite their classification into different roles. Commodities (C) generated by the industry (I) are moved by various transport modes (M) from one place (P) to another (P) along the transportation network (L) within a time period (T). During the transport process, a chain of possible events (E) may occur that involve various stakeholders or individuals (H). The last category, "unclassified," forms part of a larger "virtual boundary" that contains elements that do not fit under any of the aforementioned roles but need to be accounted for to preserve data integrity.

The nine primary groupings capture many kinds of information that could potentially be retrieved from a freight database (the validation of which is explained by demonstrating classification efficiency in the next section of this chapter). Considering the possibility of other researchers identifying new roles in the future, however, the outline of the virtual boundary is drawn in dashed lines.

The second level of classification applies to all the above roles except the **time** and **unclassified** roles. This secondary classification seeks to separate elements that *identify* a known object from elements that *describe the features* of the object. For this purpose, data elements that identify objects are defined as *identifiers*, and examples include "origin," "destination," "road name," and "transport mode." Data elements that describe the features of an object are defined as *features*, and examples include "population," "area," "length," "unit train," and "number of carloads."

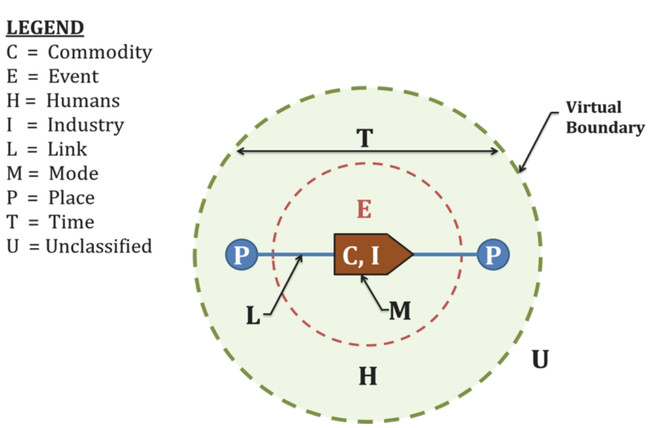


Figure 5-1. Schematic representation of the RBCS.

From the nine primary and two secondary classification groups, the following classifications groups (or roles) were developed:

- **Time elements,** which provide information regarding either the exact time period (e.g., year, month, time, day) or duration (e.g., seasons, quarter, biannual) for which the data is being reported or the freight movement occurred.
- **Place elements,** which identify or describe the O-D of freight movement or the location of an event (e.g., an accident), or which may provide information relating to the characteristics of the place.
  - ***Place identifier*** (e.g., city name, state, origin county name, destination country name, accident location). For geospatial databases, this can either be points or polygons.
  - ***Place feature*** (e.g., area, population).
- **Commodity elements,** which identify or describe a commodity being moved.
  - ***Commodity identifier*** (e.g., Standard Transportation Commodity Codes [STCC], Standard Classification of Transported Goods [SCTG] commodity codes, Harmonized System codes, hazardous material).
  - ***Commodity feature*** (e.g., liquid, bulk, value, weight, trade type).
- **Link elements,** which identify or describe information about the roadways, waterways, routes, etc., on which freight is moving.
  - ***Link identifier*** (e.g., a roadway name, a waterway name).
  - ***Link feature*** (e.g., width, length, from, or to).
- **Mode elements,** which identify or describe the vehicles involved in the movement of freight.
  - ***Mode identifier*** (e.g., truck, rail, air, vessel).
  - ***Mode feature*** (e.g., unit train, vehicle class, number of trucks).
- **Industry elements,** which identify or describe fields that report on economic activities.
  - ***Industry identifier*** (e.g., North American Industry Classification System [NAICS] codes, Standard Industrial Classification [SIC] codes, company name).
  - ***Industry feature*** (e.g., number of employees, sales, annual payroll).
- **Event elements,** which identify or describe occurrences or actions that occur when freight is being moved.
  - ***Event identifier*** (e.g., an accident report number, a dredging operation, or a port call).
  - ***Event feature*** (e.g., number of fatalities as a result of an accident; depth of dredge; or number of port calls).
- **Human elements,** which identify or describe a person involved in a data record.
  - ***Human identifier*** (e.g., investigating officer, reporting agent, or contact person).
  - ***Human feature*** (e.g., drunk driver, driver age, or operator condition).
- **Unclassified elements,** which present a unique proposition in that some databases report additional information about the dataset themselves (e.g., expansion factors applied to dataset, empty fields, etc.). By themselves, these fields do not necessarily describe freight movement, but they can provide information that is useful when performing data analysis. Examples of unclassified elements include record IDs, primary keys, comment fields, record modification dates, metadata, and administrative ID fields.

5.5 Validation

To validate the generality of RBCS, the schema was applied to all 6,322 data elements from the 28 public and commercial freight data dictionary sources. Table 5-1 illustrates the application of RBCS to data tables from five sources using all the possible roles. To quantify the ability of the proposed roles to classify data elements, this process was repeated for all the data sources. For each source, the number of elements that were successfully classified using the defined

Table 5-1. How RBCS groups data elements across databases.

Element Role	FAF3	Public Use Carload Waybill Sample	Air Carrier Statistics (all carriers)	HPMS	U.S. Waterway Foreign Cargo Inbound and Outbound Data
Time	Year	Waybill Date Accounting Period Waybill Reporting Period Length	Year Quarter Month	Year of Last Improvement Year of Last Construction	Year
Place Identifier	Foreign Region Origin Domestic Region Origin Domestic State Origin Domestic Region Destination Domestic State Destination Foreign Region Destination	Inter/intra State Code Origin BEA Area Origin Freight Rate Territory Interchange State #1 Interchange State #2 Interchange State #3	OriginAirportID OriginCityName OriginStateFips OriginStateName	Urban Code County Code Climate Zone	U.S. Port Code U.S. Port Name U.S. State Foreign Port Schedule K Code Foreign Port Code Foreign Port Name U.S. Coastal District
Place Feature	-	-	-	-	Longitude of Foreign Port Latitude of Foreign Port
Link Identifier	-	-	-	Functional System Route Number Alternate Route Name…	Waterway Code
Link Feature	-	Estimated Short Line Miles Number of Interchanges	Distance Between Airports	Facility Type Structure Type Access Control Ownership Speed Limit…	-
Mode Identifier	Foreign Inbound Mode Domestic Mode Foreign Outbound Mode	-	-	-	-

(continued on next page)

Table 5-1. (Continued).

Element Role	FAF3	Public Use Carload Waybill Sample	Air Carrier Statistics (all carriers)	HPMS	U.S. Waterway Foreign Cargo Inbound and Outbound Data
Mode Feature	-	Number of Carloads Car Ownership Category Code AAR Equipment Type Code AAR Mechanical Designation STB Car Type TOFC/COFC Service Code …	CarrierGroup CarrierGroupNew DistanceGroup Class	-	-
Commodity Identifier	Commodity (STCG)	Commodity Code (STCC)	-	-	Lock Performance Monitoring System Commodity Code
Commodity Feature	Type of Trade Value Weight Ton-Miles	Billed Weight Actual Weight Freight Revenue ($) …	-	-	Tonnage Type Processing
Event Feature	-	-	-	-	-
Industry Identifier	-	-	UniqueCarrier AirlineID UniqueCarrierName UniqCarrierEntity …	-	-
Unclassified	-	Subsample Code Exact Expansion Factor Theoretical Expansion Factor	DataSource	-	-

AAR = Association of American Railroads; BEA = U.S. Bureau of Economic Analysis; FAF3 = Freight Analysis Framework, version 3.
Note: A dash (-) indicates "not applicable."

classification roles was counted, and the ratio (classification efficiency) of classified elements to the total number of elements in a data source was calculated, as follows:

$$\textit{Classification efficiency} = \frac{\textit{Classified Elements}}{\textit{Classified Elements} + \textit{Unclassified elements}}$$

Table 5-2 shows the classification efficiency of all 28 data sources. In general, RBCS is found to yield high classification efficiencies. Of the 28 data sources, 12 had a classification efficiency of 100% and six had values ranging between 95% and 100%. Seven data sources had values ranging between 80% and 95%. It is important to note that the lower classification efficiencies can be attributed to the low number of total elements in the respective databases. As an example, the

Table 5-2. Classification efficiency of freight data sources.

Database Name	RBCS		
	Classified	Unclassified	Classification Efficiency
Air Carrier Statistics	500	4	99%
Air Carrier Financial Reports	478	0	100%
Annual Survey of Manufacturers	62	0	100%
Border Crossing/Entry	5	0	100%
CTA Intermodal Terminals Database	11	1	92%
Carload Waybill Sample	252	0	100%
Commodity Flow Survey	18	0	100%
County Business Patterns	190	132	59%
Fatality Analysis Reporting System	310	0	100%
Federal Railroad Administration Safety Database	414	89	82%
Foreign Trade	362	27	93%
Freight Analysis Framework	68	2	97%
Highway Performance Monitoring System	117	0	100%
IHS Transearch	30	0	100%
Motor Carrier Management Information System	314	44	88%
Motor Carrier Safety Measurement System	28	4	88%
National Agricultural Statistics Service	38	0	100%
National Corridors Analysis and Speed Tool Database	17	4	81%
North American Transborder Freight Database	60	6	91%
Pipeline and Hazardous Material Safety Administration	32	1	97%
Service Annual Survey	6	22	21%
Survey of Business Owners	198	0	100%
Topologically Integrated Geographic Encoding and Referencing	475	8	98%
U.S. Waterway Data	263	3	99%
Vehicle Inventory and Use Survey	241	1	100%
Vehicle Travel Information System	154	53	74%
Woods and Poole Economics, Inc.	1240	0	100%

National Corridors Analysis and Speed Tool (N-CAST) database had 17 classified elements out of a total of 21 elements, which resulted in a classification efficiency of 81%. Data sources such as the County Business Patterns, Service Annual Survey, and Vehicle Travel Information System were found to have a significant amount of "noise flag," or metadata-related, data elements.

5.6 Model Limitation

A limitation of RBCS is the need for consistency during the classification process. When ambiguity exists, this is usually resolved by critically examining data element definitions to determine an element's role and ensure it has been classified consistently throughout the process. For example, if the data element "trade type" is assigned to the role "Commodity Feature" in one database, the same role should be applied to similar "trade type" data elements in subsequent databases. This is important because if data elements like "trade type" are assigned to one role "Commodity Feature" in one database but to a different role (e.g., "Event Identifier") in another database, the data elements will not be uniformly grouped—which makes it difficult to find and analyze similar elements.

After carefully grouping the data elements across multiple databases, any decisions concerning changing an element's role can be easily made. Being consistent in the initial classification process facilitates future changes.

Another limitation of RBCS may be the limited number of roles defined. The validation process in this study revealed that the nine primary groupings essentially capture the majority of the data elements from the databases tested. However, future enumerations of other databases may result in the identification of additional roles. For example, the "Place Identifier" role could be further expanded to "Point of Origin" and "Point of Destination" roles, and the "Commodity Feature" role could be further expanded to differentiate between a "Commodity Unit of Measure" (e.g., tons, value) and a commodity feature such as "Trade Type" (e.g., import or export). Considering the possibility of additional primary and subordinate roles, The "Virtual Boundary" described in Figure 5-1 provides an opportunity for future iterations of this classification schema.

5.7 Application

In this study, RBCS was used to identify similar data elements and bridge differences in their definitions. For example, the "Place Identifier" role in Table 5-1, data elements that identify places are defined in the FAF3 data dictionary as "Foreign region origin, Domestic region origin, etc." and in the Carload Waybill Sample dictionary as "Inter/Intra State Code, Origin BEA Area, Origin Freight Rate Territory, etc." This example demonstrates a case of semantic heterogeneity in "Place" between the FAF and Carload Waybill Sample databases. Examination of these elements in isolation helps researchers formalize the process of identifying and addressing semantic heterogeneity. By ascertaining similar data elements, the subsequent process of mediating elements from those databases becomes much clearer, especially when dealing with hundreds of data elements across diverse databases.

CHAPTER 6

Differences in Data Element Definitions

6.1 Introduction

Differences in data element definitions hinder the process of (1) combining elements from individual sources into a single dataset; (2) combining elements within individual data sources from different time periods for analysis; and (3) inferring statistics from joined data elements. Identifying differences in data element definitions is critical to performing freight transportation analysis. This chapter presents the study's findings on the inherent differences among data sources for commonly utilized freight data elements such as origin and destination, commodity, mode of transport, industry, imports and exports, safety, and units of measure. Differences in element definitions were assessed based on a variety of characteristics, including the type of data, level of measurement, attribute definitions, and the spatial and temporal characteristics of each element. The discussions are limited to publicly available data sources. Privately held freight data sources have been excluded because of confidentiality concerns and the unavailability of certain data sources.

6.2 Methodology

The RBCS (role-based classification schema) was first used in identifying similar or related data elements within and across multiple data sources. Several sources from the literature also were referenced. Particularly helpful were user guides, data dictionaries, and the metadata associated with each data source. These documents provided detailed attribute descriptions and caveats for using the data, allowing the research team to compare similar elements across sources to determine whether relevant differences existed.

Differences in element definitions were categorized into three main groups: (1) taxonomic differences, (2) temporal differences, and (3) methodological or analytical differences. Complex topics within each of the three main categories were broken down further into "sub-differences." For example, under "Differences in Origin and Destination Data Elements" (Section 6.3 of this chapter), items categorized as having temporal differences were placed into one of three categories: (a) infrequent data collection, (b) changes in methodology over time, or (c) data elements accounting for temporal differences. This extra level of classification enables data users to clearly identify the interactions and interrelationships of the data elements for each topic.

6.2.1 Taxonomic Differences

Data element definitions may vary by taxonomy in terms of how the individual elements are classified. These differences can be as basic as the definition of a truck (e.g., by size, weight, axles,

44 Implementing the Freight Transportation Data Architecture: Data Element Dictionary

and so forth) or a traffic volume count (e.g., AADT [annual average daily traffic] or AAWDT [average annual weekday traffic], seasonal factors), or they may be slightly more complex, such as geographic differences for which location references have been reported differently (e.g., by point or polygon). Differences in geographic scale also can exist with regard to data from various sources, and data may be statistically valid only at certain levels of geographic detail (e.g., state, county, district).

6.2.2 Temporal Differences

Data elements across data sources or within a single data source may include inherent internal inconsistencies that make data comparisons difficult. The data may have been collected in different reporting years, or the source may have changed the definition of the data element over time. For example, changes may occur in the way industries are classified within a particular industry classification system. If an entity such as Home Depot is reclassified from a wholesale to a retail establishment within the structure of the North American Industry Classification System (NAICS)/Standard Industrial Classification (SIC), failing to account for the change can affect an analysis.

6.2.3 Methodological or Analytic Differences

Often, highly compatible data sources with commonalities in a substantial number of areas will diverge in their reporting or analytical properties. For example, two data sources that both analyze commodity movements may report the results in divergent terms, such as dollar value of cargo, tons of cargo, or units of cargo (e.g., 40-ft. container equivalent units [CEUs] or 20-ft. equivalent units [TEUs]).

Note: In the Freight Data Dictionary web application, in this chapter, and in Chapter 7, data element names given in all capital letters appear as they are represented in the actual data sources (e.g., ORIGINID, DESTINATIONID, and COMMODITYID).

6.3 Differences in Origin and Destination Data Elements

Keywords: origin, destination, place, terminal, terminus

Origin and destination are critical inputs for conducting a wide range of analyses related to freight movement. When working with origin and destination data elements, data users should be aware of taxonomic, temporal, and methodological/analytical differences among and within data sources.

6.3.1 Taxonomic Differences

The taxonomic differences among and within the origin and destination data elements discussed in this report relate to differences in geographical scale and definition.

6.3.1.1 Geography Scale and Definition

Origin and destination data elements often are represented at different geographic scales, making it difficult to perform certain types of analysis between data sources and to disaggregate the data within a single data source to smaller geographic scales. Some examples of differences in scale or definitions among origin and destination data sources are described in the balance of this section.

Commodity Flow Survey (CFS)

- CFS Areas are drawn from a subset of combined statistical areas and metropolitan statistical areas as defined by the Office of Management and Budget. When they include more than one state, however, CFS Areas are divided into their state parts.[1] For example, the Kansas City–Overland Park–Kansas City, MO–KS Metropolitan Area has both a Kansas area and a Missouri area. Given that not all origin-destination data elements disaggregate metropolitan areas into their separate state areas, caution should be used when comparing CFS Area data to other data of similar geography.
- Remainder of State is a unique geographical category used in the CFS to represent those areas of a state not contained within the CFS-defined metropolitan areas. Remainder of State can encompass large geographic zones and may introduce challenges for data users attempting localized analysis or bridging with other data sources, as other origin-destination data elements may not have the ability to identify the Remainder of State category.

Freight Analysis Framework (FAF3)

- FAF3 uses geographic definitions that may not be present within other data sources.[2]
  - FAF REGION—FAF3 contains 123 domestic regions[3] and includes data on eight foreign regions.[4] Although statistical methods exist that allow analysts to disaggregate FAF data from FAF regions to counties or smaller areas, FHWA has not measured any of these methods to establish estimates of reliability or accuracy. FAF estimates of truck tonnage and number of trucks on the network, particularly in regions with multiple routes or significant local traffic between major centers of freight activity, should be supplemented with local data to support local applications.
  - ZONE OF ENTRY—For import shipments, this data element represents the origin of flow (the FAF REGION or state of entry).
  - ZONE OF EXIT—For export shipments, this data element represents the destination of the flow (the FAF REGION or state of exit).

Carload Waybill Sample

- Standard Point Location Code (SPLC)—SPLC is used to identify the origin and destination stations (ORIGIN SPLC and DESTINATION SPLC). Other freight databases do not use this code.
- FREIGHT AREA, FREIGHT RATE AREA, FREIGHT RATE TERRITORY, and other related data elements are imputed from the SPLC.
- FREIGHT STATION ACCOUNTING CODE (FSAC)—FSAC is used to identify origin and destination stations (ORIGIN FSAC and TERMINATIONS FSAC). Other databases do not use this code.
- Business Economic Area (BEA) Codes—BEA codes are used to identify the reported waybill movement's origin and termination location. Other databases included in the review do not use this code.[5]
- Despite revisions made in November 2004 to the U.S. Department of Commerce's BEA codes and their regional boundaries to reflect changes in economic and population growth, as of the 2013 release, the Carload Waybill Sample has continued to use the February 1995 designations. The November 2004 definitions contain 179 economic areas, and the February 1995 definitions contained 172 economic areas.
- In addition to the 172 BEA codes, the Carload Waybill Sample includes 13 codes representing Puerto Rico, Mexico, and provinces in Canada. Data users are advised that the following codes are not recognized by the Department of Commerce:
  - 173: Newfoundland
  - 174: Nova Scotia

  - 175: Prince Edward Island
  - 176: New Brunswick
  - 177: Quebec
  - 178: Ontario
  - 179: Manitoba
  - 180: Saskatchewan
  - 181: Alberta
  - 182: British Columbia
  - 183: Yukon/Northwest Territories
  - 184: Puerto Rico
  - 185: Mexico.
- Princeton Transportation Network Model number—This number is used to identify the node to which the waybill movement's origin location is assigned. The number incorporates the data elements ORIGIN NET3 NUMBER and TERMINATION NET3 NUMBER. Other data sources do not use this code.

Air Carrier Statistics

- ORIGIN and DESTINATION data elements in the Air Carrier Statistics signify airport codes.
- _CITYMARKETID is used as a data element in the Air Carrier Statistics to identify and consolidate airports serving the same city market. Other data sources do not use this code.
- Other data elements unique to the Air Carrier Statistics database include:
  - _AIRPORTID, which is an identification number assigned by U.S. DOT to identify a unique airport. This field is recommended for use when performing airport analysis across a range of years, because airports can change their airport codes and airport codes can be reused.
  - _AIRPORTSEQID, which is an identification number assigned by U.S. DOT to identify a unique airport *at a given point of time.*
  - _WAC, which reports the world area code where an airport is located.
- Data elements within the Air Carrier Statistics database that are similar to those in other databases include _CITYNAME, _STATE, _STATEFIPS, _STATENAME, and _COUNTRY. These data elements represent the location of the originating or destination airport.

6.3.2 Temporal Differences

Temporal differences among and within origin and destination data elements fall under the following categories:

- Infrequent data collection
- Changes in methodology over time
- Data elements accounting for temporal differences

The balance of this section presents examples of temporal differences and how these can pose challenges in data analysis. Sources that collect data infrequently may make trend analysis or data interpolation difficult within a single data source. Similarly, making comparisons *across* data sources can be difficult if the sources use different collection periods.

6.3.2.1 Infrequent Data Collection

Commodity Flow Survey (CFS)

- Since 1993, the CFS has been conducted every 5 years (during years ending in 2 or 7). Temporal gaps in data collection years may create difficulty when filling gaps and making comparisons with other datasets that have more complete temporal coverage.

6.3.2.2 Changes in Methodology Over Time

Commodity Flow Survey (CFS)

- Before the 2007 CFS, a survey was conducted to obtain information on shipping status and value of shipments for the auxiliaries. The U.S. Census Bureau concluded that the advance survey enabled more accurately assigned shipper status for both the warehouse and managing office auxiliaries on the 2007 CFS sampling frame as compared with the 2002 sampling frame; however, the accuracy of shipper status for managing offices on the frame was less than for non-auxiliaries.
- The 2012 survey included the addition of 11 additional metropolitan statistical area geographies.[6]

Carload Waybill Sample

- Despite revisions made in November 2004 to the U.S. Department of Commerce's Business Economic Area (BEA) codes and their regional boundaries to reflect changes in economic and population growth, as of the 2013 release, the Carload Waybill Sample has continued to use the February 1995 designations. The November 2004 definitions contain 179 economic areas, and the February 1995 definitions contained 172 economic areas.[7]

Foreign Trade Statistics (FTS)

- The FTS's Origin of Movement identifier was added in 1985. This identifier indicates the state where the export journey began. It allows the compilation of the Origin of Movement—Based on Origin State series. Available since 1987, this series provides export statistics based on the state from which the merchandise starts its journey to the port of export; that is, the data reflects the transportation origin of exports.

6.3.2.3 Data Elements Accounting for Temporal Differences

Air Carrier Statistics

- ORIGINAIRPORTID and ORIGINAIRPORTSEQID are data elements that address temporal differences within the data source, accounting for the fact that cities can change their airport codes over time. ORIGINAIRPORTSEQID identifies a unique airport at a given point in time.[8]

6.3.3 Methodological/Reporting Differences

Commodity Flow Survey (CFS)

- The methodology for verifying origin changed slightly in 2012 and differs from previous reporting years. Additional information is available in the Freight Data Dictionary web application or in the CFS documentation available at http://www.rita.dot.gov/bts/help_with_data/commodity_flow_survey.html#naics_table.

Freight Analysis Framework (FAF3)

- The FAF3 documentation states that the methods and data sources used have changed compared to those used in developing previous FAF versions, and that versions should not be compared to each other. For example, FAF version 2 (FAF2) has 114 domestic origins and destinations but FAF3 has 123 domestic regions.

6.4 Differences in Commodity Data Elements

Keywords: commodity, HS (Harmonized System) code, STCC (Standard Transportation Commodity Codes), SCTG (Standard Classification of Transported Goods), SITC (Standard International Trade Classification), bulk, break-bulk, hazardous materials, Schedule B, HTS (Harmonized Tariff Schedule)

6.4.1 Taxonomic Differences

Taxonomic differences among and within commodity data elements fall under the following categories:

- Data elements that include or exclude commodities or commodity groups
- Differing classification systems

6.4.1.1 Data Elements that Include or Exclude Commodities or Commodity Groups

Data sources that report commodity information often include certain industries or modes of transport but exclude others, which makes it difficult to directly compare commodities across multiple data sources. The balance of this section discusses which industries or commodities are included or excluded in the definition of a commodity across multiple data sources.

Carload Waybill Sample

- This data source uses the Standard Transportation Commodity Codes (STCC) to identify the product designation for the commodities transported. In the Carload Waybill Sample, this field includes the first five digits of the seven-digit STCC; however, the codes for some commodities, like STCC 19 series commodities (ordnance [guns and artillery] or accessories) are reported only at the two-digit level. Commodities in the STCC 49 series (hazardous materials) and STCC 50 series (bulk materials in boxcars) also have been translated to actual product commodity codes.
- The Public Use Waybill Sample, which is a sub-parent of the Carload Waybill Sample, does not include hazardous materials (STCC series 49xxx) or bulk materials in boxcars (STCC series 50xxx).[9]

Center for Transportation Analysis Intermodal Terminals Database

- Only grain elevators, cement terminals, petroleum tank farms, and liquid bulk storage and transfer terminals with waterway connections are included in the data source. All other data on intermodal terminals will need to be obtained from additional sources.[10]
- This data source excludes many public warehouses served by rail or truck-rail reload centers for lumber, steel, paper, or other break-bulk freight. Additional information on these public warehouses will need to be obtained from other sources.

Commodity Flow Survey (CFS)[11]

- Within the CFS, a commodity is defined as a product that an establishment produces, sells, or distributes. This definition does not include items that are considered excess or operational waste. Survey respondents report the description and the five-digit SCTG (Standard Classification of Transported Goods) code for the commodity contained in the shipment. Shipments having multiple commodities are grouped together, and the commodity with the greatest weight is selected to represent the total shipment. Commodities that are part of the shipment but are not the majority weight are not classified.

- Shipments originating from business establishments located in Puerto Rico and other U.S. possessions and territories are excluded from the data file. Thus, commodities originating from these locations also are excluded.
- Data for government-operated establishments are excluded from the CFS. These establishments include public utilities, publicly operated bus and subway systems, public libraries, and government-owned hospitals.
- The CFS also excludes establishments or firms with no paid employees. Data users should be aware that any commodities imported or exported from or for these establishments are excluded.
- Commodities shipped via containerized cargo are labeled Intermodal. Commodities that move by more than one mode are labeled "multiple modes and mail." These classifications differ from those used in other data sources.

Fatal Analysis Reporting System (FARS)[12]

- To be included in FARS, a crash must involve a motor vehicle traveling on a trafficway customarily open to the public within the 50 states, the District of Columbia, or Puerto Rico that resulted in the death of a person (occupant of a vehicle or a non-motorist) within 30 days of the crash. Notably, any hazardous commodities not meeting the criteria for inclusion in the FARS are excluded from the report. Therefore, FARS data on hazardous materials can only be used with other data on hazardous materials in circumstances that involved a fatal crash.
- NHTSA updates the FARS Analytical User's Manual every year to summarize the evolution of coding. When conducting analysis across years, data users should check every data element of interest in each year's coding manual.

Foreign Trade Statistics (FTS)[13]

- Imports and exports of commodities on vessels moving under their own power or afloat, and on aircraft flown into or out of the United States, are included in the "All Methods" data table but excluded from the "Vessel and Air" statistics. Thus, commodities that are shipped on vessels moving under their own power and via aircraft are included in the former, but excluded from the latter.
- Mail and parcel post shipments (including those transported by vessel or air) are included in the "All Methods" data table but excluded from the "Vessel and Air" statistics.
- Low-value shipments, which are defined as exports valued under $2,501 or imports valued under $1,251, are included in the "All Methods" data table but excluded from the "Vessel and Air" statistics. Commodities that qualify under low-value shipments are included in the "All Methods" data. Low-value shipments are estimated, and may not directly correspond to other data source estimates.[14]

Freight Analysis Framework (FAF3)[15]

- The FAF3 is built primarily on the Commodity Flow Survey (CFS), but the FAF3 does not have the level of commodity detail found in the CFS, nor does it identify hazardous cargo.
- Commodities shipped via containerized cargo are labeled "intermodal." Commodities that move by more than one mode are labeled "multiple modes and mail." These classifications differ from those used in other data sources.

Motor Carrier Management Information System (MCMIS)

- Hazardous materials information is not available to the general public, so authorization is needed to obtain data on these commodities. Without proper authorization, commodity data on hazardous materials cannot be used with other sources.

North American Transborder Freight Database (Transborder)[16]

- Commodities are identified using the two-digit commodity code indicated by Schedule B for U.S. export shipments and the HTS (Harmonized Tariff Schedule) for U.S. import shipments.
- Because of customer requests, the U.S. Bureau of Transportation Statistics discontinued the inclusion of trans-shipment activity in Transborder freight data beginning with the January 1997 data month.
- Air and vessel data by month or year are not available before 2004.
- Import values from Mexican states are not available.

U.S. Waterway Data[17]

- The U.S. Waterway database is the only data source focused exclusively on waterways, including inland waterways, offshore waters, the Great Lakes, and the Saint Lawrence Seaway. Data on commerce, facilities, locks, dredging, imports and exports, and accidents are included, along with the geographic waterway network.
- Commodities reported in this data source include coal, petroleum products, chemicals, crude materials, manufactured goods, farm products, machinery, and waste, and commodities labeled “unknown.”[18] All other commodities are excluded from the data source.

Vehicle Inventory and Use Survey (VIUS)[19]

- The VIUS excludes federal, state, or local government vehicles, as well as ambulances, buses, motor homes, farm tractors, unpowered trailer units, and trucks that have been reported to have been sold, junked, or wrecked before January 1 of the survey year. Any commodities, including hazardous materials, carried via the excluded vehicles also are excluded from the survey, which may create difficulties when bridging with other data sources.

6.4.1.2 Differing Classification Systems

Data sources that report commodity information often utilize commodity classification systems that differ from one data source to another. The most commonly used commodity classification systems are the Standard Classification of Transported Goods (SCTG), Standard Transportation Commodity Codes (STCC), Harmonized System (HS), and Standard International Trade Classification (SITC). Data users are advised to be aware of the different resolutions of these systems and account for them when performing data analysis. A detailed discussion of each classification system appears in Appendix B.

Carload Waybill Sample[20]

- The Carload Waybill Sample contains various resolutions of freight movements reported at the Business Economic Area (BEA)-to-BEA level (or across multi-county BEA areas) and the seven-digit STCC level. Data users are advised to note the following differences in resolution between tables:

- UNIQUE SERIAL NUMBER—To allow for unique identification of waybills, the Association of American Railroads/Railinc assigns a unique, six-digit serial number to all waybills processed.
  - Hardcopy waybills are assigned serial numbers in the 100,000 to 199,999 range.
  - Machine Readable Input (MRI) waybills are assigned serial numbers in the 200,000 to 999,999 range and 000,000 to 099,999.

  This unique serial number does not correspond with commonly used system codes.[21]
- WAYBILL NUMBER—This number is the number an originating railroad document assigns to each waybill. The waybill number gives detailed instructions relating to a shipment, and the codes vary depending on the consignor or consignee, the point of origin, its destination, and route.[22]
- CONFIDENTIAL CARLOAD WAYBILL SAMPLE COMMODITY CODE (STCC)—This data element uses the STCC coding to identify the product designation for the commodity being transported at the seven-digit STCC level.[23]
  - The STCC 48 series (hazardous waste) is part of the regular STCC.
  - The STCC 49 series (hazardous materials) is used only for hazardous materials, in lieu of the regular STCC.
  - The STCC 50 series is used for bulk commodities transported in box cars.
- STCC W/O HAZARDOUS (49) CODES—This data element on the Confidential Carload Waybill Sample takes the hazardous codes (STCC series 49xxxxx) and bulk codes (STCC series 50xxxxx), and translates them to the actual product commodity codes.
- Public Use Waybill Sample Commodity Code (STCC)—The STCC identifies the product designation for the transported commodity. This data field includes the first five digits of the seven-digit STCC; however, STCC 19 series commodities are reported only at the two-digit level.

Center for Transportation Analysis Intermodal Terminals Database[24]

- The intermodal terminals data source contains a list of 3,100 transload facilities in the United States where commodities may be transferred between surface modes. Data users are advised to note the following difference in resolution between tables:
  - CARGO—A three-digit code for the type of cargo or commodity group involved in the intermodal connection. This code does not correspond to other data sources.[25]

Commodity Flow Survey (CFS)[26]

- The CFS contains shipment data using varying resolutions of the Standard Classification of Transported Goods (SCTG) system. Data users are advised to note the following difference in resolution between tables:
  - COMMODITY—A product that an establishment produces, sells, or distributes. Respondents report the description and the five-digit SCTG code for the commodity contained in the shipment. Shipments involving multiple commodities are classified as the commodity with the greatest weight in the total shipment.

Fatal Analysis Reporting System (FARS)[27]

- FARS is a nationwide census providing NHTSA (National Highway Traffic Safety Administration), Congress, and the American public yearly data regarding fatal injuries suffered in motor vehicle traffic crashes. Data users are advised to note the following differences in resolution between tables:

- HAZ_ID—This data element identifies the four-digit hazardous material identification number for a vehicle in transit. These numbers are developed by the United Nations (UN) and used worldwide in international commerce and transportation to identify hazardous materials. Materials without a UN number may be assigned a four-digit North American (NA) number, which usually starts with the number 8 or the number 9.[28]
- HAZ_CNO—This data element identifies the single-digit hazardous material class number for a vehicle in transit. The U.S. DOT has identified nine hazard classes based on the dangers posed in transportation.[29]
- PHAZ_ID—This data element applies to parked and working vehicles and uses the same four-digit hazardous material identification number for a vehicle as HAZ_ID.
- PHAZ_CNO—This data element applies to parked and working vehicles and uses the same single-digit hazardous material class number for a vehicle as HAZ_CNO.

- NHTSA updates the FARS Analytical User's Manual[30] every year to summarize the evolution of coding. When conducting analysis across years, data users should check every data element of interest in each year's coding manual.

Foreign Trade Statistics (FTS)[31]

- The FTS database contains varying resolutions of HS (Harmonized System) industry classifications for different tables. Data users are advised to note the following differences in resolution between tables:
  - U.S. Exports of Merchandise—Monthly—This data table contains commodity details using a variety of codes at varying levels. These include the 10-digit Schedule B code, the 5-digit SITC (Standard International Trade Classification) code, the 10-digit HTS (Harmonized Tariff Schedule) code, the 5-digit End-Use code, and the 6-digit HS code.[32]
  - U.S. Imports of Merchandise—Monthly—This data table contains commodity detail at the 2-, 4-, 6-, and 10-digit HS levels.[33]
  - U.S. Exports and Imports by Port—This data element contains various data fields for HS commodities at the six-digit HS level.

Freight Analysis Framework (FAF3)[34]

- The FAF3 contains varying resolutions of Standard Classification of Transported Goods (SCTG) classifications for different tables. Data users are advised to note the following difference in resolution:
  - SCTG2—This data element contains commodity codes that are based off the SCTG code at the two-digit level.[35]

Motor Carrier Management Information System (MCMIS)[36]

- The MCMIS contains information on the safety fitness of commercial motor carriers and hazardous material shippers subject to the Federal Motor Carrier Safety Regulations and the Hazardous Materials Regulations. Data users are advised to note the following differences in resolution between tables:
  - HAZMAT MATERIAL ID—This identifying code is associated with hazardous materials cargo. The codes correspond with four-digit United Nations/North American (UN/NA) identification numbers.[37]
  - HAZARDOUS MATERIALS CARRIED/SHIPPED—This code identifies the type of hazardous material transported or shipped by the entity and whether bulk (B), non-bulk (N), or all (A). It is important to note that the conversion of the Hazardous Materials Data elements

of the new Census File to the old is as follows: Bulk (B) = Tank (T), Non-Bulk (N) = Package (P), and All (A) = Both (B). These codes do not correspond to other data sources.
- CARGO—Describes the cargo hauled by a particular carrier. A maximum of three cargo types are printed. These codes do not correspond to other data sources.[38]
- HAZMAT C—This code identifies the type of hazardous material carried by interstate and intrastate motor carriers. Up to three hazardous material types may be printed. The letter B indicates that the cargo is carried in bulk quantities. N indicates that the cargo is carried in non-bulk quantities. A indicates cargo that is carried both in bulk and non-bulk quantities. These codes do not correspond to other data sources.
- HAZMAT S—This code identifies the type of hazardous material shipped by interstate and intrastate shippers, with cargo coded the same as for HAZMAT C. Up to three hazardous materials types may be printed. The letter B indicates that the cargo is shipped in bulk quantities. N indicates that the cargo is shipped in non-bulk quantities. A indicates cargo that is shipped both in bulk and non-bulk quantities. These codes do not correspond to other data sources.

National Agricultural Statistics Service (NASS)[39]

- The NASS database contains varying resolutions of agricultural and demographic statistics for different tables within the data source. Data users are advised to note the following differences between tables:
  - SECTOR—In this data source, sectors constitute five high-level, broad categories that are useful in narrowing down choices: Crops, Animals & Products, Economics, Demographics, and Environmental. These codes do not correspond to other data sources.
  - GROUP—These data elements are subsets within a sector (e.g., under the sector Crops, the groups are Field Crops, Fruit & Tree Nuts, Horticulture, and Vegetables). These codes do not correspond to other data sources.
  - COMMODITY—This data element records the primary subject of interest (e.g., Corn, Cattle, Labor, Tractors, Operators). These codes do not correspond to other data sources.

North American Transborder Freight Database (Transborder)[40]

- Transborder contains varying resolutions of freight flow data by commodity type and by mode of transport (rail, truck, pipeline, air, vessel, and other) for U.S. exports to and imports from Canada and Mexico at the North American Free Trade Agreement (NAFTA), national, and state/province level. Data users are advised to note the following differences in resolution between tables:
  - COMMODITY—This data element contains export or import commodity detail and is available at the two-digit HS (Harmonized System) level.
  - U.S. Trade with Canada and Mexico Import Commodity Detail—This data element contains commodity codes indicated by a two-digit Schedule B number for U.S. export shipments and two-digit Harmonized Tariff Schedule (HTS) for U.S. import shipments. Schedule B and HTS codes correspond to HS codes up to the six-digit level.

U.S. Waterway Data[41]

- The U.S. Waterway database contains varying resolutions of data on commerce, facilities, locks, dredging, imports and exports. Accidents are included along with the geographic waterway network. Data users are advised to note the following differences in resolution between tables:

- CONTAINER—This data element indicates whether the vessel carries containers (signified by the letter C) or not (left blank).[42] These codes do not correspond to other data sources.
- PMS_COMM—This data element contains the two-digit Lock Performance Monitoring System (LPMS) commodity code. The LPMS is based off the Standard International Trade Classification (SITC) Revision 3 commodity code.[43]
- PRINC_COMM—This data element describes the principal commodities carried by a Transportation Lines vessel company.[44]

Vehicle Inventory and Use Survey (VIUS)[45]

- The VIUS contains varying resolutions of data on the physical and operational characteristics of the nation's truck population at the national level and state level. Data users are advised to note the following differences in resolution between tables:
  - PRODUCT_PRINCPL—This data element indicates the principle product carried, at the two-digit SCTG (Standard Classification of Transported Goods) level, by this vehicle configuration.[46] Products are recoded to the highest percent; if the highest percent occurs for more than one category, the record is assigned to "multiple categories."[47]

6.5 Import and Export Data Elements

The narrative given in this section is derived from the following documents:

- Border Crossing/Entry Data: FAQ (frequently asked questions). Retrieved from http://Transborder.bts.gov/programs/international/Transborder/TBDR_BC_FAQs.html
- A Description of the FAF3 Regional Database and How It Is Constructed. The Freight Analysis Framework Version 3 (FAF3). FHWA, June 16, 2011. Retrieved from http://faf.ornl.gov/fafweb/Data/FAF3ODDoc611.pdf
- Principal Ports of the United States. U.S. Waterway Data. Navigation Data Center. Retrieved from http://www.navigationdatacenter.us/data/datappor.htm
- Guide to Foreign Trade Statistics. U.S. Census Bureau. Retrieved from http://www.census.gov/foreign-trade/guide/
- Transborder Freight Data Program (Transborder Documentation). U.S. Department of Transportation, Research and Innovative Technology Administration Bureau of Transportation Statistics (RITA/BTS), September 2009. Retrieved from http://Transborder.bts.gov/programs/international/Transborder/PDF/TransborderFreightDataProgram.pdf

The authors recommend reviewing these sources for additional information concerning each data source. This narrative serves only as a summary of information gleaned from these sources.

Keywords: import, export, port of entry, trade, foreign trade

The import and export data elements discussion relates to databases that report on U.S. trade with other countries or geographical regions. The Bureau of Transportation Statistics (BTS) classifies U.S. international trade and transportation data into three primary categories: administrative trade data, carrier-based data, and shipper-based data. These categories are based on how the data is collected and the scope of each data source. The taxonomic, temporal, and methodological/analytical differences that still exist in these data sources are discussed in this section of *NCFRP Report 35.*

- Administrative Trade Statistics—These international trade statistics are captured from administrative documents required by the Department of Homeland Security (DHS). U.S. Customs and Border Protection (CBP) is responsible for collecting this information either in paper form or electronic form at U.S. ports of entry, exit, or clearance. Currently, electronic information is captured through the Automated Broker Interface for imports and through the Automated Export System for exports. Together, the Automated Broker Interface and Automated Export System are known as the Automated Commercial System. The Automated Commercial System is being replaced by the Automated Commercial Environment (ACE), which will serve as the primary system through which the trade community will report imports and exports. CBP has established a schedule for completing development of all trade processing capabilities in ACE by the end of 2016.[48]
- The U.S. Census Bureau's Foreign Trade Division is responsible for verifying, processing, and distributing the data after collection by the CBP. Other federal agencies receive special tabulations from the Census Bureau, based on the official U.S. international trade statistics. These agencies then perform additional quality assurance reviews and analyses for their own purposes and to meet the needs of their customers. Following are the types of administrative trade statistics gathered:
  - Foreign trade statistics
  - North American land trade, disseminated as the North American Transborder Freight Data (Transborder)
  - U.S. international maritime trade, released to the Maritime Administration and the U.S. Army Corps of Engineers (USACE)
  - U.S. transportation-related goods and overall trade data, released to the U.S. Bureau of Economic Analysis
- Carrier-Based Sources—The main sources of carrier-based international trade data are:
  - International air freight data from the Research and Innovative Technology Administration (RITA)/BTS, disseminated as the Air Carrier Statistics
  - Maritime data from the *Journal of Commerce*'s Port Import/Export Reporting Service (PIERS)
  - Special periodic surveys, such as Canada's National Roadside Survey
- Shipper-Based Sources—The Commodity Flow Survey (CFS, conducted in 1993, 1997, 2002, and 2007) is the only publicly available shipper-based survey that provides some information on U.S. international trade and transportation. The export data is limited, however, and not directly comparable to merchandise trade exports released by other sources, including the Census-based Foreign Trade Statistics (FTS).

6.5.1 Taxonomic Differences

Taxonomic differences among various data sources relating to their scope and definitions of import/export data fall under the following categories:

- Sources that report on foreign trade movement origin and destination, including the port of entry (e.g., Foreign Trade Statistics [FTS], Transborder, Freight Analysis Framework)
- Sources that report on foreign trade movement only at the port-of-entry level (e.g., Border Crossing Entry Data, U.S. Waterway Data)

- Sources that report on foreign trade movement origin and destination but exclude the port of entry (e.g., Carload Waybill Sample)

6.5.1.1 Sources that Report on Foreign Trade Movement's Origin and Destination, Including the Port of Entry

Foreign Trade Statistics (FTS)

- The Foreign Trade Statistics (FTS) data compiled by the Census Bureau are the official U.S. import and export statistics and reflect both government and nongovernment shipments of merchandise between foreign countries. The data is made available for subscription in four different formats: (1) Merchandise Trade, (2) State Data, (3) Port Data, and (4) Special Products.[49] Following are the taxonomic differences in the four types of data:
  - Merchandise Trade—These data files provide commodity information for different commodity classification codes, as follows:
    - 10-digit Schedule B
    - 10-digit Harmonized System (HS)
    - Standard International Trade Classification (SITC)
    - End-Use, North American Industry Classification System (NAICS)
    - USDA
    - Advanced Technology Products

    Merchandise Trade data files also include only port district information. Individual ports of entry/exit are not reported in these data files.
  - State Data—These data files report summarized trade statistics by U.S. state. In comparison to the merchandise trade data files, commodity data is available in just two formats: the six-digit HS code and four-digit NAICS code. State Data files do not include information on the port of entry/exit. The State Exports/Port database reports trade data by U.S. state, district, and port of exit. Time data, however, is reported in periods rather than in statistical year and month as in the other databases. For example, the period coded using the number 1 covers January, February, and March; period 2 is April, May, and June; period 3 is July, August, and September; and period 4 is October, November, and December. State Imports/Port data is not published.
  - Port Data—These files report trade information through the individual port of entry/exit but exclude the U.S. state of origin or destination. Commodity data is reported at the six-digit HS commodity code level.
  - Special Products—These data files report on the following:
    - U.S. general imports assembled abroad from components produced in the United States (textile summary)
    - U.S. imports for consumption and general imports for all imports entered under secondary or Census special program indicators
    - Shipments of merchandise from the United States to Puerto Rico and the U.S. Virgin Islands, and shipments from Puerto Rico, the U.S. Virgin Islands, Guam, American Samoa, the Northern Mariana Islands, and U.S. minor outlying islands to the United States.

North American Transborder Freight Database (Transborder)

- Transborder is a subset of the Foreign Trade Statistics (FTS) database. It is the first attempt by the U.S. Census Bureau to disaggregate U.S. foreign trade statistics into the various surface modes of transportation. Transborder contains freight flow data by commodity type and by mode of transport for U.S. exports to and imports from Canada and Mexico. Taxonomic characteristics of data elements in this database that may differ from those in the other data sources include the following:

- Geographic scope:
  - USASTATE—U.S. states (introduced January 2007)
    - This data element is based on the two-digit U.S. Postal code.
    - It identifies the U.S. state of origin for exports to or state of destination for imports from Canada and/or Mexico. The state may not always represent the physical origin or destination of the import or export goods, because the exporter's or importer's address may not be in the same state as the origin or destination of the goods.
  - TRDTYPE—This data element identifies the direction in which the commodity is moved. The USASTATE with TRDTYPE will identify the origin and destination information based on whether TRDTYPE is an import or an export.
  - DEPE—The U.S. Census Bureau is responsible for maintaining the classification of U.S. Customs districts/ports of entry, codes, and descriptions. This classification is known as "Schedule D."
    - For imports, this data element represents where the entry documentation was filed with Customs and the duties paid, and it may not always be where the goods physically entered the United States.
    - For U.S. exports, this data element represents the last port where the shipment is cleared for export.
    - State totals for trade can be based on the state of destination for imports and the state of origin for exports, not on the state Customs port of entry or exit. This is because many border ports serve as national gateways, and not all the goods that enter or exit through a port either originated in or are destined for that particular state.
  - CANPROV—Canadian provinces
    - For U.S. imports from Canada, the Canadian province represents where the goods were grown, manufactured, or otherwise produced. However, the province information may also reflect the province used as the mailing address of the Canadian exporter or the address of an intermediary; therefore, in some instances, the mailing address may not be the actual province of physical origin.
    - For U.S. exports to Canada, the Canadian province represents the Canadian province of clearance. The province of clearance is the province in which Canadian Customs cleared the shipment, and is not necessarily the province of final destination.
  - MEXSTATE—Mexican states
    - The Mexican state of destination is the state in which the ultimate consignee is located in Mexico, and is not necessarily the state of final destination. The Census Bureau captures the data field for MEXSTATE from the ultimate consignee's address. If a Mexican state of destination cannot be identified for a particular shipment, it is considered unknown and coded as OT in the data field.
    - Data for the Mexican state of origin for U.S. imports from Mexico is not captured as part of current trade filing requirements.
- STATMO—Data reported for each calendar month
- COMMODITY—Commodity data reported using the two-digit HTS (Harmonized Tariff Schedule)
- DISAGMOT—Disaggregated Mode of Transport
  - This data element represents only the mode by which shipments enter or exit the United States and does not reflect all the modes of transportation used throughout the entire journey of the shipment, from foreign point of origin to final destination.
  - DISAGMOT uses numerical codes to signify the following modes:
    - 1 = Vessel
    - 3 = Air

- 4 = Mail (U.S. Postal Service)
- 5 = Truck
- 6 = Rail
- 7 = Pipeline
- 9 = Foreign trade zones
- 8 = Other (including unknown)

(Foreign trade zones was added as a mode of transport in 1995.) "Mail" is used as the mode used for U.S. Postal Service shipments and cannot be further divided into either rail or truck shipments. The category "Other" includes "flyaway aircraft, or aircraft moving under their own power (i.e., aircraft moving from the aircraft manufacturer to a customer and not carrying any freight), powerhouse (electricity), vessels moving under their own power, pedestrians carrying freight, unknown, and miscellaneous other."[50] Users should note that the actual mode of transport for a specific shipment into or out of a foreign trade zone is unknown because U.S. Customs and Border Protection (CBP) does not collect this information.

– VALUE and SHIPWT—These data fields report data by value and weight.
  - For imports, the data field VALUE refers to the Customs value or the value of merchandise for duty purposes. It is usually the selling price in the foreign country of origin. VALUE excludes freight costs, insurance, and other charges incurred in bringing the merchandise from the foreign port of export to the United States.
  - For exports, the data field VALUE refers to the value of the merchandise, usually the selling price, plus insurance and freight at the U.S. port of export. The value, as defined, excludes the cost of loading the merchandise aboard the exporting carrier at the port of export and also excludes freight, insurance, and any charges or transportation costs beyond the U.S. port of export.
  - For exports, the weight of U.S. exports by land modes of transportation is not available because this data is not required to be reported on the paper Shipper's Export Declarations documents required by the U.S. Census Bureau. The new electronic filing system for exports, the Automated Export System, does require that export weight be filed for all modes of transport.

    RITA/BTS uses the value-to-weight ratio of U.S. imports at the two-digit commodity code level to calculate the export weights. Although the export weights are not published as tables, RITA/BTS uses these numbers for U.S. Transborder publications.

Freight Analysis Framework (FAF3)

- Taxonomic characteristics of data elements in this database that may differ from those in the other data sources include the following:
  – Geographic scope for imports and exports are either reported at the state level or based on FAF3 zones and regions (e.g., FR_ORG, DMS_ORG, FR_DEST DMS_DESTST), which include FAF zone-specific ports of entry and exit.
  – Historical data is available for a limited number of years (1997, 2002, 2007, and 2012). Projected data is reported in 5-year increments from 2015 to 2040.
  – Two-digit commodity codes (SCTG2 is the data field) are based on the SCTG (Standard Classification of Transported Goods) classification system.
  – Mode of transport is classified as inbound, outbound, or domestic, as follows:
    - FR_INMODE—This data element represents mode of transport from foreign origin to zone of entry.
    - FR_OUTMODE—This data element represents mode of transport from zone of exit to the foreign destination.

  - DMS_MODE—This data element represents the domestic mode of transport from zone of entry to destination zone for imports, and from origin zone to zone of exit port for exports.
- Data is reported by value (VALUE), weight (TONS), and ton-miles (TMILES), and is determined based on mode-specific data modeling procedures. (See Section 6.5.3 Methodological Differences in this chapter for a discussion of import/export data elements.)

Air Carrier Statistics

- Air Carrier Statistics—This data source reports data differently from other data sources. The T-100 Market data tables, which report only on trips from origin to destination, exclude port-of-entry/exit information if the port of entry/exit is an intermediate stop for the shipment. The T-100 Segment data tables, on the other hand, include the port of entry/exit for international shipments but exclude the original origin-destination when a shipment has multiple stops. The Market and Segment data tables report only on the weight of cargo shipped and exclude both value and commodity type data. Unlike other data sources, however, the Air Carrier Statistics reports data at a more disaggregate level (i.e., origin and destination city and airport).

6.5.1.2 Sources that Report on Foreign Trade Movement Only at the Port-of-Entry Level

Border Crossing/Entry Data

- Border Crossing/Entry Data provides summary statistics for incoming crossings at the U.S.–Canadian and the U.S.–Mexican border at the port level. The data element MEASURE provides information on the number of personal vehicles, trucks, buses, containers, trains, passengers, or pedestrians entering the United States. The data, which is reported on a monthly basis (YEAR and MONTH), does not include information on the place of origin or final destination of commodities transported. It provides information only on the port of entry or exit (PORT LOCATION).

U.S. Waterway Data—Principal Ports of the United States

- The U.S. Waterway Principal Port file contains USACE commodity tonnage summaries (total tons, domestic, foreign, imports, and exports), port codes (PORT), geographic locations (LONGITUDE and LATITUDE), and names (PORT_NAME) for the top 150 ports for a particular year. Commodity names and descriptions are not available; neither is there data on place of origin or final destination.

6.5.1.3 Sources that Report on a Shipment's Origin and Destination and Exclude the Port of Entry

Confidential Carload Waybill Sample

The Confidential Carload Waybill Sample contains the data element field TYPE MOVE, whose options include details from which a user can infer whether that particular data record involves an import or export commodity or none. Available field options include the following:

- Neither import nor export
- Imported commodity
- Exported commodity
- Commodity imported and exported (e.g., land bridge traffic)
- Unknown

6.5.2 Temporal Differences

Data users should be aware of temporal differences among and within databases as a result of varying frequency of data collection and changes in the definition of a data element over time.

Foreign Trade Statistics (FTS)

Temporal differences in the FTS database, as documented in the *Guide to Foreign Trade Statistics*, include the following:

- The United States is substituting Canadian import statistics for U.S. exports to Canada in accordance with a 1987 Memorandum of Understanding signed by the Census Bureau, U.S. Customs and Border Protection (CBP), Canadian Customs, and Statistics Canada. This data exchange includes only U.S. exports destined for Canada and does not include shipments destined for third countries by routes passing through Canada or shipments of certain grains and oilseeds to Canada for storage before exportation to a third country, which are reported on and compiled from Electronic Export Information documents.
- The statistical month of importation is the month in which U.S. Customs and Border Protection (CBP) releases the merchandise to the importer.
- The statistical month of exportation is based on the date when the merchandise leaves the United States. (For vessel or air shipments, it is the date when the carrier departs or is cleared from the port of export.)
- The Census Bureau seasonally adjusts the merchandise trade data at the five-digit end-use commodity category level, the most detailed end-use level possible. These detailed data are then summed to the one-digit level for release with the monthly merchandise trade totals.
- Effective with the release of January 2014 statistics on March 7, 2014, the Census Bureau publishes seasonally adjusted selected countries and world areas in FT-900 Exhibit 19. Unlike the commodity-based adjustments, these adjustments are developed and applied directly at the country and world area level.

North American Transborder Freight Database (Transborder)

- For temporal differences in data element definitions for the Transborder database and several significant reporting changes since the first release of data, see the section on "Major Reporting Changes" in the Transborder Freight Data Program Documentation.[51] For example, starting in January 2007, the Bureau of Transportation Statistics (BTS) used a new data structure to release the Transborder data. Twenty previously separate data tables were consolidated into the current form of three data tables:
  - U.S. Imports and Exports with State and Port Detail
  - U.S. Imports and Exports with State and Commodity Detail
  - U.S. Imports and Exports with Port and Commodity Detail
- In addition, trans-shipments data (covering shipments from a third country through Canada or Mexico to the United States or from the United States to a third country through Canada or Mexico) were excluded from the public database beginning with the January 1997 data. (Note: Before January 1997, documentation for this dataset referred to this type of activity as "in-transit shipments.")

Freight Analysis Framework (FAF3)

- For temporal differences, see the discussion on data collection methods and limitations of the FAF database available on the Freight Data Dictionary web application.

Border Crossing/Entry Data

- Temporal differences within the Border Crossing/Entry Data include the following:
  - Data on passenger vehicles and passengers in personal vehicles for the Cape Vincent, NY, ferry are available beginning in 2007. The ferry between Wolfe Island (Canada) and Cape Vincent does not operate in the winter.
  - Until May 2011, truck and rail data for the port of entry of Otay Mesa, CA, was reported by the Bureau of Transportation Statistics (BTS) as Otay Mesa/San Ysidro, CA, which is the same as the CBP's reporting of the data to the BTS. However, San Ysidro has been a passenger crossing for many years and no freight is allowed through this port of entry by truck or rail. Hence, BTS changed the name to Otay Mesa for truck and rail crossings to avoid any confusion to the data user. Thus, Otay Mesa, CA, and San Ysidro, CA, are now reported as separate ports of entry for all data elements.

6.5.3 Methodological Differences

Methodological differences arise among import and export data elements of different databases as a result of the processes by which the data is collected and disseminated.

Foreign Trade Statistics (FTS)

- Electronic Export Information—These mandatory documents are filed by the U.S. Principal Party in Interest or its agents through the Automated Export System and record U.S. exports data for merchandise from all countries except Canada.
- Automated Commercial System—This automated U.S. Customs database compiles U.S. imports data on merchandise. U.S. imports data on merchandise also is compiled from import entry summary forms, warehouse withdrawal forms, and foreign trade zone documents as required by law to be filed with U.S. Customs and Border Protection (CBP). Data on imports of electricity and natural gas (NG) from Canada is obtained from Canadian sources.

North American Transborder Freight Database (Transborder) and Border Crossing/Entry Data

- Transborder Surface Freight Data—This data is extracted from the foreign trade statistics collected by the Census Bureau.
- Border Crossing/Entry Data—This data is based on transportation mode count data collected by U.S. Customs and Border Protection (CBP).

Freight Analysis Framework (FAF3)

- Import and export flows are constructed using mode-specific data sources, each of which is converted from agency specific commodity codes to FAF3's two-digit SCTG (Standard Classification of Transported Goods) codes. In addition, commodity flows from the respective databases are either spatially aggregated or disaggregated to match FAF3 regions.
- This list summarizes the mode-specific data modeling procedures used in developing the FAF3:
  - Waterborne Imports and Exports
    - Main sources of data include the following:
      - FAF3-specific extraction of data from the Port Import/Export Reporting Service (PIERS) maritime database
      - USACE's International Waterborne Commerce database
      - Foreign Trade Statistics (FTS) database

    - PIERS forms the basis of foreign waterborne flows in FAF3 with several adjustments, including the following:
        - Ensuring PIERS total commodity tonnages is consistent with USACE Waterborne tonnages
        - Ensuring PIERS total commodity dollar valued trades is consistent with FTS totals
        - Inferring missing data for zip-code-level reporting of shipment originations and destinations within the continental United States and inland mode of transport within the continental United States
        - Addressing known issues, such as the reporting of origin and destination data in the FTS dataset (i.e., exporting/importing company addresses are reported rather than the actual physical location of the point of departure or arrival of the shipment)
    - Missing or questionable data were allocated across domestic FAF3 zones in proportion to the distribution of shipment volumes in the 2007 U.S. Commodity Flow Survey (CFS).
- Air Freight Imports and Exports
    - Main sources of data include the following:
        - The Air Carrier Statistics T-100 International Market data table provides estimates of total tons shipped annually between an originating airport (where the cargo is first loaded onto an aircraft) and a destination airport (where the cargo is unloaded for final land-based delivery, usually by truck).
        - The FTS database provides information on value, commodity class, quantity, method of transportation, and shipping weights.
    - The Air Carrier Statistics T-100 data tables and FTS database are combined into a single FAF3 air freight dataset by reconciling differences in the level of spatial and commodity detail. If differences exist between the T-100 and FTS flow totals, the T-100 data tables are taken to be definitive for total tons shipped, and the FTS database is used to control the allocation of freight shipments across commodity classes and to assign value-to-weight ratios to these flows.
- Transborder U.S.–Canada and U.S.–Mexico Imports and Exports
    - The main source of data is the North American Transborder Freight Database (Transborder). Shipments were allocated to the most likely counties of origination or destination in each state using the 2007 U.S. County Business Pattern data.
    - Origin-destination (O-D) estimation is done by:
        - Removing vessel, air, and pipeline mode movements from the dataset (leaving truck and rail "land" border shipments)
        - Spatially allocating flows reported at the state level to their most likely FAF3 regions within the United States
        - Converting Harmonized System (HS) commodity classes to FAF3 SCTG classes
    - Shipment weight data for exports to Canada and Mexico is estimated on the basis of average dollar/ton statistics generated from export shipments by specific HS commodity class, mode, and country.
- Imports of Crude Petroleum
    - Monthly reported Energy Information Administration (EIA) data containing company, U.S. seaport of entry/exit, and foreign country information is used to estimate crude petroleum imports in FAF3.
    - O-D flow is represented as movement from foreign country (i.e., source of commodity import) to a U.S. port (domestic FAF3 origin region), then to a U.S. refinery (FAF3 domestic destination region). Allocation of these flows to specific modes of transportation are based on EIA data on crude oil refinery receipts, broken down by mode of transport (ship, pipeline, rail, barge, truck), and further broken down by domestic versus foreign sources of production.

- Imports and Exports of Natural Gas (NG)
  - EIA reports annual movement of liquefied natural gas (LNG), which is carried by larger tanker ships to and from a U.S. seaport of entry/exit. EIA also reports on natural gas (NG) trade by pipeline between the United States and Canada or Mexico. Supporting data used in allocating flows to specific FAF3 O-D pairs came from the U.S. Census Bureau's County Business Patterns dataset. NG flows were allocated to respective FAF3 domestic regions based on U.S. port of entry or exit, and exporting countries also were allocated to their respective FAF3 foreign regions.

- For additional information on the data sources, estimation methods, and data quality issues, please refer to "Estimation of Import and Export Flows" in The Freight Analysis Framework Version 3—A Description of the FAF3 Regional Database and How It Is Constructed.[52]

Air Carrier Statistics

- The Federal Aviation Act of 1958 requires each large certificated air carrier to file Form 41 (reports of financial and operating statistics) monthly, quarterly, semiannually, and annually with the Bureau of Transportation Statistics (BTS). BTS publishes the data submitted on the forms as the Air Carrier Statistics.[53]

U.S. Waterway Data

- In this database, the Principal Ports of the United States data is derived from the Waterborne Commerce Statistics Center.

6.6 Industry Data Elements

Keywords: industry, NAICS, SIC, HS (Harmonized System) code

When working with data elements related to industry classification, data users should be aware of taxonomic and methodological/analytical differences among and within data sources. This section discusses differences between data elements that classify industries, including the North American Industry Classification System (NAICS), HS (Harmonized System), and SIC (Standard Industrial Classification) industry codes.

6.6.1 Taxonomic Differences

Taxonomic differences among and within data elements related to industry classification can be categorized as follows:

- Inclusion/exclusion of industry groups
- Industry definition resolution

6.6.1.1 Inclusion/Exclusion of Industry Groups

Data sources may sometimes include or exclude certain industry groups, making it difficult to directly compare industries across multiple data sources.

Annual Survey of Manufacturers

- Annual Survey of Manufacturers—This database covers only manufacturing establishments with one or more paid employees, along with non-employers that use leased employees for manufacturing (which are classified in NAICS Sector 31-33).[54] All other industries are excluded from this data source, which creates difficulty in making comparisons with other data sources' industry classifications.

County Business Patterns

- County Business Patterns—This database covers most North American Industry Classification System (NAICS) industries, with the following exceptions:
  - NAICS 111 and NAICS 112—Crop and animal production
  - NAICS 482—Rail transportation
  - NAICS 491—Postal Service
  - NAICS 525110, 525120, and 525190—Pension, health, welfare, and vacation funds
  - NAICS 525920—Trusts, estates, and agency accounts
  - NAICS 814—Private households
  - NAICS 92—Public administration[55]

  These exclusions should be noted when using this data with other data sources related to industry classification.

Foreign Trade Statistics (FTS)

- The export statistics contained in the FTS data source consist of goods valued at more than $2,500 per commodity shipped by individuals and organizations (including exporters, freight forwarders, and carriers) from the United States to other countries. Data users are advised to note the exclusion of goods valued at under $2,500 per commodity shipped.

Commodity Flow Survey (CFS) and Freight Analysis Framework (FAF3)

- The CFS does not contain data from the following freight-generating and freight-receiving industries, or it contains insufficient data to cover the industries in a comprehensive manner.
  - Multimodal truck, rail and pipeline flows of crude petroleum, petroleum products, and natural gas (NG)
  - Truck freight shipments associated with farm-based, fishery, logging, construction, retail, services, municipal solid waste, and household and business moves
  - Imported and exported goods transportation by ship, air, and Transborder land (truck, rail) modes
- In FAF3 these industries produce what are called Out-Of-Scope to the CFS freight flows (OOS flows). Their estimation required a great deal of data collection and integration into the larger flow matrix generation process. For the most part, the data sources for these OOS flows are derived from freight-carrier-reported data sources. In some cases they require the use of secondary or indirect data sources, such as location-specific measures of industrial activity, employment, or population, to allocate flows to specific geographic regions. Developing OOS flow estimates represents a good deal of effort, with different commodity classes requiring very different, typically multi-step, treatments, including the use of both spatial and commodity class "crosswalks" that convert mode-specific and industry-class-specific estimates from their native coding categories into FAF3 regional and commodity class breakdowns.[56]

6.6.1.2 Industry Definition Resolution

Data sources containing industry classification data elements frequently use differing industry coding resolutions. These industry classifications include North American Industry Classification System (NAICS) or Harmonized System (HS) resolution. Analysts should be aware of the different resolutions and account for them during analysis.

Annual Survey of Manufacturers

- The Annual Survey of Manufacturers contains varying resolutions of NAICS (North American Industry Classification System) industry classifications for different tables within the data source. Data users are advised to note the following differences in resolution between tables:
  - Statistics for Industry Groups and Industries—Total manufacturing establishments' statistics are presented at the three-, four-, five-, and six-digit NAICS levels at the national level.
  - Value of Product Shipments—This file represents shipments statistics for the 471 six-digit NAICS product groups and approximately 1,384 seven-digit NAICS product classes at the national level.
  - Geographic Area Statistics—This file represents manufacturing establishments' statistics at the three- and four-digit NAICS levels for each state.
  - Supplemental Statistics for the United States—This file represents supplemental manufacturing establishments' statistics at the two-digit NAICS levels for each state.

County Business Patterns

- The County Business Patterns data source contains varying resolutions of North American Industry Classification System (NAICS) industry classifications for different tables within the data source. Data users are advised to note the following differences in resolution between tables:
- County File—Provides data at the six-digit North American Industry Classification System (NAICS) industry code at the county level
  - State File—Provides data at the six-digit NAICS industry code at the state level
  - U.S. File—Provides data at the six-digit NAICS industry code at the national level
  - Metropolitan Area File—Provides data at the six-digit NAICS industry code at the metropolitan area level
  - Zip Code Industry Detail File—Provides data at the six-digit NAICS industry code at the zip code level
  - Commonwealth of Puerto Rico File—Provides data at the six-digit NAICS industry code for Puerto Rico and the Island Areas of American Samoa, Guam, the Commonwealth of the Northern Mariana Islands, and the Virgin Islands of the United States.

Foreign Trade Statistics (FTS)

- The Foreign Trade Statistics (FTS) data source contains varying resolutions of North American Industry Classification System (NAICS) and Harmonized System (HS) industry classifications for different tables within the data source. Data users are advised to note the following differences in resolution between tables:
  - USA Trade Online—This online subscription service provides U.S. export and import statistics of industries at a high level of granularity, up to the 10-digit HS and six-digit NAICS classifications, by state, country, and Customs district. Data categories include the following:
    - District Data (10-digit HS detail)—2003–current
    - Port Data (six-digit HS detail)—2003–current
    - State Export Data (six-digit HS detail and four-digit NAICS detail)—2002–current
    - State Import Data (six-digit HS detail and four-digit NAICS detail)—2002–current
    - NAICS Data (six-digit NAICS detail)—2002–current
  - U.S. Exports and Imports of Merchandise—This data source provides export and import statistics for industry-level HS commodities at the two-, four-, six-, and 10-digit levels.

Country and Customs district data for value and quantity are provided on a monthly, year-to-date, and annual basis.
  - Merchandise Trade Exports—This data source offers multiple files (12 in all)—1989–current
  - Merchandise Trade Imports—This data sources offers multiple files (12 in all)—1989–current
- U.S. Exports and Imports of Merchandise—This data source provides 5 years of historical annual revised export and import statistics for industry-level HS commodities at the two-, four-, six-, and 10-digit levels. Commodity data for value and quantity are provided on an annual basis.
  - Merchandise History Exports—This data source offers multiple files (12 in all)—1989–current
  - Merchandise History Imports—This data source offers multiple files (12 in all)—1989–current
- U.S. Exports and Imports by State—This data source provides export and import statistics by State of Origin of Movement (export) and State of Destination (import) for industry-level commodities at the six-digit HS level or the three- or four-digit NAICS level. Data is provided on a monthly, quarterly, or annual basis.
  - State Exports—1987–current, and
  - State Imports—This data release was discontinued in 1988 but reinstated in 2010; 2008–current data is now available.[57]
- U.S. Exports and Imports by Port—This data source provides export and import statistics by State of Origin of Movement for industry-level commodities at the six-digit HS level on a monthly, quarterly, or annual basis.

Service Annual Survey

- In the Definitions data table, the Service Annual Survey data source classifies businesses into categories using the six-digit North American Industry Classification System (NAICS) code.

6.7 Mode of Transport Data Elements

Keywords: mode, transport, air, rail, pipeline, truck, waterway, vessel, vehicle, multimodal, intermodal, unknown

In this section, differences in data element definitions are broken down by the various modes of transport. If applicable, taxonomic, temporal, and methodological differences across data sources are presented for each mode. The following modes of transport are addressed:

- Air
- Highway
- Rail
- Water
- Pipeline
- Multimodal/intermodal
- Unknown/other

6.7.1 Air

6.7.1.1 Taxonomic Differences

Data sources often use unique data elements to identify and differentiate mode of transport. Data users are advised to note these differences when using air mode data elements from multiple data sources.

Air Carrier Statistics

- T-100 Market (All Carriers)
  - This data table differentiates and reports freight class in four categories using letter codes:
    - F—Scheduled passenger/cargo service
    - G—Scheduled all cargo service
    - L—Non-scheduled civilian passenger/cargo service
    - P—Non-scheduled civilian all cargo service
- T-100 Segment (All Carriers)
  - This data table differentiates mode by aircraft group, aircraft type, and aircraft configuration.
    - AIRCRAFTGROUP—This data element includes codes such as:[58]
      - Piston—Single engine
      - Piston—2-engine
      - Helicopter/Stol
      - Turbo-prop 1- and 2-engine
      - Turbo-prop 4-engine
      - Jet
    - AIRCRAFTTYPE—This data element includes codes such as:[59]
      - Cessna 180
      - Piper PA-32
      - Convair CV-340/440
      - McDonnell Douglas DC-6
    - AIRCRAFTCONFIG—This data element includes codes such as:[60]
      - Passenger configuration
      - Freight configuration
      - Combined passenger and freight on a main deck
      - Seaplane
  - The Air Carrier Statistics T-100 Segment data table also reports air service type, including service class, as follows:
    - F—Scheduled passenger/cargo service
    - G—Scheduled all cargo service
    - L—Non-scheduled civilian passenger/cargo service
    - P—Non-scheduled civilian all cargo service

  The full list of T-100 Segment codes can be found at the url provided in references 57–59 in the endnotes to Chapter 6 and 7.

Fatal Analysis Reporting System (FARS)

- FARS uses a data element called "Transported to Medical Facility By," which reports details on travel mode using numerical codes. For example, the number 1 signifies "EMS air," meaning transport by emergency medical services using air mode.[61]
- NHTSA updates the FARS Analytical User's Manual[62] every year to summarize the evolution of coding. When conducting analysis across years, data users should check every data element of interest in each year's coding manual.

6.7.1.2 Temporal Differences

Temporal differences among and within air mode data sources fall under the following categories:

- Changes in methodology over time
- Data elements that account for temporal variation

6.7.1.2.1 Changes in Methodology over Time. Data sources may change their data collection or reporting methods over time, making it difficult to compare data elements across multiple years within a single data source, or across data sources.

North American Transborder Freight Database (Transborder)

- Beginning in 1997, the Bureau of Transportation Statistics (BTS) restructured the Transborder freight data files to simplify the table structure and improve usability of the data. Under the new reporting methodology, the DISAGMOT data element uses numerical codes to identify mode of transport for shipments entering and exiting the United States. DISAGMOT 3 indicates air mode;[63] however, DISAGMOT 8 ("Other and unknown") includes "flyaway aircraft," defined as aircraft moving under their own power (i.e., aircraft moving from the aircraft manufacturer to a customer and not carrying any freight).
- With the release of the January 2004 data, BTS began incorporating vessel and air data provided by the U.S. Census Bureau into the Transborder data. The vessel and air data provides information on U.S.-North American Transborder trade similar to U.S.-North American Transborder surface freight. Further reporting changes can be found in the Transborder Freight Data Documentation.

Commodity Flow Survey (CFS)

- The following methodological changes were incorporated into the CFS in 2012:
  - Shipments with a respondent-provided mode of "parcel" must weigh 150 pounds or less, while shipments with a respondent-provided mode of "air" are not given a weight restriction.
  - A shipment's mode of transport is imputed whenever a respondent has provided a mode of "other" or "unknown," or has failed to provide a modal response (coded as "missing mode").[64]

6.7.1.2.2 Data Elements that Account for Temporal Variation. Data sources sometimes use multiple data elements to identify locations whose names or codes change over time. Caution should be exercised to ensure that the correct identifier is used.

Air Carrier Statistics

- Over time the code or name of an air carrier may change, and the same code or name may be assumed by multiple airlines. To ensure that data users analyze data from the same airline, Air Carrier Statistics provides four airline-specific variables that identify unique carriers (airlines) or their associated entities:
  - AIRLINEID—Airline ID
  - UNIQUECARRIER—Unique carrier code
  - UNIQUECARRIERNAME—Unique carrier name
  - UNIQCARRIERENTITY—Unique carrier entity

A unique carrier is defined as one holding and reporting under the same department of transportation certificate regardless of its code, name, or holding company/corporation.[65] Notably, the Air Carrier Statistics data includes large certified carriers with annual operating revenues of $20 million or more.[66]

6.7.1.3 Methodological Differences

When working with data elements related to the air mode of transport, data users should be aware that methodological differences exist not only among data sources but also within individual data sources.

Commodity Flow Survey (CFS)

- In the CFS, air mode shipments include shipments carried by truck to or from an airport. For multiple-mode shipments, if the respondent has reported a shipment's mode of transport as both parcel and air, the CFS treats the shipment as parcel only.
- The 2007 CFS classified air shipments as shipments weighing 100 pounds or more. During mileage processing for the 2007 CFS, an "air" shipment was manually converted to "parcel" if the weight of the shipment was less than 100 pounds. However, airlines do not necessarily have minimum weight restrictions when transporting cargo. Hence, for the 2012 CFS, the definition of an air shipment was changed. As a result, an air shipment was acceptable as provided by the respondent, regardless of weight. Furthermore, for the 2012 CFS, parcel shipments conformed to the definition used by the parcel industry that a parcel is a shipment of 150 pounds or less. For shipments submitted by the respondent with mode of Parcel and a weight above 150 pounds, GeoMiler changed the mode to For-Hire Truck during mileage processing.[67]
- In the case of imports and exports by air, domestic moves by ground to and from the port of entry or exit are categorized as shipments by truck.[68]

Foreign Trade Statistics (FTS)

- The FTS data source presents transportation statistics in three categories—vessel, air, and "All Methods"—based on the method of transportation by which the merchandise arrived in or departed from the United States. Some shipments between the United States and other countries will enter or depart the United States through Canada or Mexico. Such shipments are recorded under the method of transportation by which they enter or depart the United States, regardless of the transportation mode between Canada or Mexico and the country of origin or destination.[69]
- Data reported on vessel, air exports, and general imports represents waterborne and airborne shipments only (i.e., merchandise leaving or arriving in the United States aboard a vessel or an aircraft).
- Imports and exports of vessels moving under their own power or afloat, and aircraft flown into or out of the United States, are included in the "All Methods" data table but are excluded from the "Vessel and Air" statistics.
- Mail and parcel post shipments (including those transported by vessel or air) are included in the "All Methods" data but excluded from the "Vessel and Air" statistics.

Freight Analysis Framework (FAF3)

- Because of a modification in the reporting of multimodal and intermodal categories between the 2002 and 2007 Commodity Flow Survey (CFS) on which the FAF is based, there is no direct equivalence in the modal classes implied between these two sets of definitions, with the exception of the truck-only and rail-only modes.
- Air data includes any shipment sent via air mode to its destination. Data users are advised to note that air mode shipments include shipments carried by truck to or from an airport. For multiple-mode shipments, if the respondent reported a shipment's mode of transport as both parcel and air, FAF3 treats the shipment as parcel only.
- This data source includes shipments weighing more than 100 pounds that move by air, or a combination of truck and air, in commercial or private aircraft, including air freight and air express. The CFS/FAF3 does not include shipments weighing 100 pounds or less, which are typically classified as "multiple modes and mail."[70]

6.7.2 Highway

When performing data analysis, data users are advised to be aware of several temporal, taxonomical, and methodological differences associated with data elements related to highway mode.

6.7.2.1 Temporal Differences

Temporal differences among and within highway mode data sources occur because of changes in methodology over time. Data sources may change their data collection or reporting methods over time, making it difficult to compare data elements across multiple years within a single data source, or across data sources.

Commodity Flow Survey (CFS)

- For the 2012 CFS, a change was made relating to mileage processing. Private truck is now considered a short-haul mode (i.e., private trucks not routed more than 500 miles during shipment routing).[71] Data users should be aware of this adjustment.

Motor Carrier Management Information System (MCMIS)

- Beginning in 1994, states participating in the Motor Carrier Safety Assistance Program were required to report through the SAFETYNET system a standard set of data items on all trucks and buses involved in traffic crashes that met a specific severity threshold. Reportable crashes include one or more of the following vehicle types:
  - A truck (used primarily for the transportation of property) having at least six tires in contact with the road surface
  - A vehicle displaying a hazardous material placard
  - A bus with seating for at least nine people (15 people before 2001), including the driver[72]

North American Transborder Freight Database (Transborder)

- Beginning in January 1997, the Bureau of Transportation Statistics (BTS) restructured the Transborder freight data files to simplify the table structure and improve usability of the data. Land mode tables that were previously separate from air and vessel tables have been combined, and now all modes of transport are covered by the data element DISAGMOT. The DISAGMOT data element identifies mode of transport for shipments entering and exiting the United States using numerical codes. DISAGMOT 5 signifies truck mode; DISAGMOT 4, signifying mail mode, represents U.S. Postal Service and courier shipments, and cannot be further subdivided into a mode such as air, rail, or truck.[73]
- Before 1993, the U.S. Census Bureau provided mode of transport information only for air, water, and "Other." No detail was available for surface trade. The current version of the Transborder Freight Database (Transborder) makes freight transportation data available for all modes of transportation.[74]

Vehicle Inventory and Use Survey (VIUS)

- Now discontinued, the VIUS was conducted every 5 years from 1963 until the final release in 2002. Although data releases are available for all of the surveys, public use microdata files are only available for years 1977 and later. Data users should also be aware that before 1997 the survey was known as the Truck Inventory and Use Survey (TIUS).[75]

6.7.2.2 Taxonomic Differences

Commodity Flow Survey (CFS)

The CFS includes the following distinctions between private and for-hire trucks for the highway mode, as follows:

- Private Truck—This data element is defined as a truck operated by employees of the establishment or the buyer/receiver of the shipment, and includes trucks providing dedicated services to the establishment. Shipments via private truck are generally short-haul in nature.[76]
- For-Hire Truck—This data element is defined as a truck operated by common or contract carriers made under a negotiated rate. "For-hire truck" also is used if the shipment mileage was equal to or greater than 500 miles, regardless of the commodity being transported.[77]

Fatal Analysis Reporting System (FARS)

- Transported to Medical Facility By—This data element reports ground travel using the number 5, which signifies EMS ground (i.e., transport by emergency medical services using ground mode).[78]
- NHTSA updates the FARS Analytical User's Manual[79] every year to summarize the evolution of coding. When conducting analysis across years, data users should check every data element of interest in each year's coding manual.

Freight Analysis Framework (FAF3)

- The FAF3, which uses the same definitions as the Commodity Flow Survey (CFS), reports the highway mode using the number 1 (signifying truck mode), which includes both private and for-hire trucks. The truck mode does not include trucks that are part of "multiple modes and mail" (coded using the number 5), or truck moves in conjunction with domestic air cargo.[80]

Vehicle Inventory and Use Survey (VIUS)

- BODYTYPE—This data element describes which body type a vehicle most closely resembles. This field contains 30 options for body type, such as concrete pumper, sport utility, street sweeper, and tow/wrecker. This level of classification is not used in other data sources.[81]

6.7.2.3 Methodological Differences

Commodity Flow Survey (CFS)

- For 2012 CFS mileage processing, if the shipment weighed less than 80,000 pounds, it was routed via highway mode as "for-hire truck"; if the shipment weighed 80,000 pounds or more, it was routed via rail mode.
- The CFS does not report on highway shipments weighing 150 pounds or less, which are typically classified as "multiple modes and mail."[82]

North American Transborder Freight Database (Transborder)

- For this data source, the mode of transport is recorded as the mode in use when the shipment enters or exits the United States. Therefore, if a shipment originates from Dallas, Texas, by

rail but transfers to truck in Austin, Texas, and arrives in the Port of Laredo to cross the U.S.–Mexico border by truck, mode of transport for that shipment is truck.[83]
- Before 2007, data by port and commodity detail were not available for download or analysis for the land modes.

6.7.3 Rail

Several temporal, taxonomical, and methodological differences are associated with data elements related to rail that data users should be aware of when performing data analysis. These differences are discussed in this section.

6.7.3.1 Taxonomic Differences

Carload Waybill Sample

- The Confidential Carload Waybill Sample contains the data element TYPE MOVE (Movement Type), which indicates whether the rail freight is imported, exported, imported and exported (e.g., land bridge traffic), neither, or unknown.
- Both the Confidential Carload Waybill Sample and the Public Use Waybill Sample contain the data element TRANSIT CODE, which indicates whether goods were moved using "all rail," "intermodal" (a continuous movement involving at least one railroad and another mode), or "unknown" mode.
- In accordance with Accounting Rule 260, the Confidential Carload Waybill Sample uses multiple data elements for interline transactions, including "origin railroad," up to eight "bridge" railroads, and a "termination" railroad. This taxonomy is different from the Public Use Waybill Sample, which does not contain this information.
- For additional information, users can refer to Railway Accounting Rules, Association of American Railroads (September 2, 2012).[84]

Federal Railroad Administration (FRA) Safety Database

- For accidents involving rail, the data element VEHICLE indicates whether automobiles, buses, trucks, motorcycles, bicycles, farm vehicles, and all other modes of surface transportation were involved in an incident.[85]

6.7.3.2 Temporal Differences

Temporal differences among and within data sources related to rail are a result of changes in methodology over time. Data sources may change their data collection methodology over time, making it difficult to compare data elements across multiple years within a single data source, or with other data sources.

North American Transborder Freight Database (Transborder)

- Before 1993, the U.S. Census Bureau provided mode of transport information only for air, water, and "Other." No detail was available for surface trade. Currently, North American freight transportation data are available for all modes of transport, including rail.
- Beginning in January 1997, land mode tables that were previously separate from air and vessel tables were combined so that all modes of transport were covered by the data element DISAGMOT. DISAGMOT identifies mode of transport for shipments entering and exiting the United States, using numerical codes. DISAGMOT 6 signifies rail mode. DISAGMOT 4,

signifying mail mode, represents U.S. Postal Service and courier shipments, and cannot be further subdivided into individual modes such as air, rail, or truck.[86]

6.7.3.3 Methodological Differences

Several methodological differences exist within individual data sources, as well as among data sources, that data users should be aware of when working with data elements related to the rail mode.

Commodity Flow Survey (CFS)

- For 2012 CFS mileage processing, if the shipment weighed 80,000 pounds or more, it was routed via rail mode; if the shipment weighed less than 80,000 pounds, it was routed via highway mode as a for-hire truck.
- Rail includes any common carrier or private railroad, regardless of the class. The CFS does not report on shipments weighing 150 pounds or less, which are typically classified with "multiple modes and mail."[87]

Foreign Trade Statistics (FTS)

- Transportation statistics are presented in terms of three categories—vessel, air, and "All Methods"—based on the method of transportation by which shipments arrived in or departed from the United States. Some shipments between the United States and other countries enter or depart the United States through Canada or Mexico. Such shipments are recorded under the method of transportation by which they enter or depart the United States, regardless of the transportation mode between Canada or Mexico and the country of origin or destination.

Freight Analysis Framework (FAF3)

- Because of the redefinition of multimodal and intermodal categories between the 2002 and 2007 Commodity Flow Survey (on which the FAF is based), there is no direct equivalence in the modal classes implied between these two sets of definitions, with the exception of the truck-only and rail-only modes. Appendix A[88] of the FAF3 shows the modal class changes between 2002 and 2007 as well as definitions for the modes.
- In the FAF3 data source, rail mode (coded using the number 2) includes any common carrier or private railroad. Shipments with multiple modes, including those with rail, are identified as "multiple modes and mail" (using the number 5).[89]

North American Transborder Freight Database (Transborder)

- Before 2007, data by port and commodity detail were not available for download or analysis for the land modes, which includes rail. The "Mode of Transport Bridges" page provides a crosswalk from the three new tables, starting January 2007, to all the previous data tables before 2007.[90]
- For this data source, mode of transport is recorded as the mode in use when the shipment enters or exits the United States. For example, if a shipment originates from Dallas, Texas, by rail but transfers to truck in Austin, Texas, and arrives in the Port of Laredo to cross the U.S.–Mexico border by truck, the mode of transport for that shipment is "truck."[91]
- Further reporting changes related to rail can be found in the Transborder Freight Data Documentation.[92]

6.7.4 Water

6.7.4.1 Taxonomic Differences

Taxonomic and methodological differences among and within water mode data sources fall under the following categories:

- Unique data elements
- Inclusion/exclusion of data

6.7.4.1.1 Unique Data Elements. Data sources often use unique data elements to identify and differentiate attributes related to water modes of transport. Data users are advised to note these differences when performing analysis with data elements from multiple data sources.

Carload Waybill Sample

- The Confidential Carload Waybill Sample contains the data element TYPE OF MOVE VIA WATER, which provides data on water movement within the United States. This data element includes the following distinctions, coded by number:
  - 0 = Not a water movement
  - 1 = Ex-Lake (from Great Lakes to reporting railroad)
  - 2 = Lake Cargo (from rail to Great Lakes)
  - 3 = Intercoastal (a continuous movement by U.S. rail that is part of an Atlantic Ocean [or Gulf of Mexico] and Pacific Ocean movement, in either direction)
  - 4 = Coastwise (a continuous movement involving rail at either end of a coastwise movement between ports on the East Coast, including the Gulf of Mexico, or between ports on the West Coast)
  - 5 = Inland Waterways (a rail movement in combination with a barge movement on rivers and canals other than on the Great Lakes that is not considered a part of the rail movement [e.g., rail-car ferry])
  - 9 = Unknown
  - Blank (not reported on hardcopy waybills)[93]

Commodity Flow Survey (CFS)

- The CFS contains the data element MODE within the ORIGIN BY DESTINATION BY MODE data table. This data element reports the mode of transport used to move a shipment to its domestic destination. For water moves, the CFS includes the following mode categories:
  - Inland water—This data element is used to report vessels or barges operating primarily in navigable waters, both within and along the borders of the United States, including rivers, lakes, vessels moving along the shoreline but actually in the ocean (e.g., on the Intracoastal Waterway along the Atlantic and Gulf coasts, Inside Passage of Alaska), canals, harbors, major bays, and inlets.
  - Great Lakes—This data element is used for vessels or barges operating on the Great Lakes.
  - Deep sea—This data element is used for vessels or barges operating primarily in the open waters of the ocean, outside the borders of the United States.
  - Multiple waterways—This data element is used for shipments sent by any combination of Inland water, Great Lakes, and Deep sea, and which usually involve a transfer between vessels.[94]

North American Transborder Freight Database (Transborder)

- The DISAGMOT data element uses numerical codes to identify mode of transport for shipments entering and exiting the United States. DISAGMOT 1 signifies "Vessel" (indicating water mode).[95]

U.S. Waterway Data

- The data element VTCC (Vessel Type, Construction, and Characteristics) contains a four-character alphanumeric code that describes in general terms the vessel type, construction, and characteristics of its use. For example, a VTCC code of 2A22 represents the code for a self-propelled tanker constructed of steel that is being used as a liquid bulk tanker. A full list of vessel types can be found in the Navigation Data Center User's Guide under Appendix 4.[96]

6.7.4.1.2 Inclusion/Exclusion of Data. Certain data elements are included in some water data sources and excluded in others. Data users are advised to note these gaps in the data when using data sources.

Foreign Trade Statistics (FTS)

- Data for "Vessel and Air" exports and for general imports represent waterborne and airborne shipments only (i.e., merchandise actually leaving or arriving in the United States aboard a vessel or an aircraft).
- Imports and exports moved by vessels moving under their own power or afloat and by aircraft flown into or out of the United States are included in the "All Methods" data but excluded from the "Vessel and Air" statistics.
- Mail and parcel post shipments, including those transported by vessel or air, are included in the "All Methods" data, but are excluded from the "Vessel and Air" statistics.
- Low-value shipments are included in the "All Methods" data but are excluded from the "Vessel and Air" statistics.[97]

Freight Analysis Framework (FAF3)

- In the FAF, the water mode includes shallow draft, deep draft, Great Lakes, and intra-port shipments. Data users are advised to note that water mode data does not include shipments that are classified under "multiple modes and mail."[98]

6.7.4.2 Methodological Differences

Foreign Trade Statistics (FTS)

- Statistics related to water modes are based on the method of transportation by which the merchandise arrived in or departed from the United States. Some shipments between the United States and other countries enter or depart the United States through Canada or Mexico. Such shipments are recorded under the method of transportation by which they enter or depart the United States regardless of the transportation mode between Canada or Mexico and the country of origin or destination.[99]

Freight Analysis Framework (FAF3)

- Changes in the way the 2002 versus 2007 Commodity Flow Survey (CFS) assigned water-only versus water-inclusive intermodal shipments (typically, truck-water combinations) make direct comparisons of water-only traffic volumes and modal shares problematic. Appendix A of the FAF3 User Guide shows the modal class changes between 2002 and 2007 and provides definitions for the modes.[100]

North American Transborder Freight Database (Transborder)

- With the release of January 2004 statistics, the Bureau of Transportation Statistics (BTS) began incorporating vessel data provided by the U.S. Census Bureau into the Transborder data. The vessel data provided information on U.S.–North American transborder trade similar to U.S.–North American transborder surface freight. Thus, for the first time, additional information such as U.S.–North American transborder trade by port and commodity became available. Further reporting changes related to vessel data can be found in the Transborder Freight Data Documentation.[101]

6.7.5 Pipeline

6.7.5.1 Temporal Differences

Temporal differences among and within pipeline mode data sources occur as a result of changes in methodology over time.

North American Transborder Freight Database (Transborder)

- Before 1993, the U.S. Census Bureau only provided mode of transport information for air, water, and "Other." No detail was available for surface trade. Since 1993, however, North American freight transportation data has been made available for all modes of transportation, including pipelines.
- Beginning in January 1997, the Bureau of Transportation Statistics (BTS) restructured the Transborder freight data files to simplify the table structure and improve usability of the data. Land mode tables that had previously been separate from the air and vessel tables were combined, and now all modes of transportation are covered by the data element DISAGMOT. DISAGMOT uses numerical codes to identify mode of transport for shipments entering and exiting the United States. For example, DISAGMOT 7 signifies pipeline mode.[102]

6.7.5.2 Methodological Differences

Commodity Flow Survey (CFS)

- Pipeline data in the CFS includes movements of oil, petroleum, gas, slurry, and so forth through pipelines that extend to other establishments or locations beyond the shipper's establishment; however, aqueducts for the movement of water are not included.[103]

Freight Analysis Framework (FAF3)

- The FAF3 definition of pipeline mode (coded using the number 6) includes crude petroleum, natural gas (NG), and product pipelines. Data users are advised to note that products shipped via pipeline include flows from offshore wells to land which USACE counts as water moves. Pipelines that are part of "multiple modes and mail" are not included in the FAF3 pipeline data.[104]

6.7.6 Multimodal/Intermodal

6.7.6.1 Taxonomic Differences

Taxonomical differences among and within multimodal mode data sources involve the inclusion or exclusion of data in element definitions.

Commodity Flow Survey (CFS)

- The CFS does not report on shipments weighing 150 pounds or less, which are typically classified under "multiple modes and mail."[105]

Freight Analysis Framework (FAF3)

- Multiple Modes and Mail—This value includes shipments by multiple modes and by parcel delivery services, U.S. Postal Service, or couriers. This category is not limited to trailer-on-flat-car or container-on-flat-car (TOFC/COFC) shipments.[106]

Intermodal Terminals Database

- Because an intermodal terminal may connect more than one pair of modes or transfer more than one type of cargo between one or more pairs of modes, it may have multiple records in the intermodal connections file.[107]

6.7.6.2 Temporal Differences

Temporal differences among and within data sources related to multimodal freight movement are a result of changes in methodology over time.

Carload Waybill Sample

- To provide more complete data, in 1994 Railinc began flagging privately owned intermodal units; however, although reporting of private intermodal units has since increased, many of these units are still not reported in the Universal Machine Equipment Register (UMLER) computer platform.[108]
- Railinc's UMLER database is used by railroads, rolling stock owners, and repair shops to share a wealth of rail-car information, which is used to interchange cars, pool traffic, and issue blocking requests.[109]
- To reduce the possibility of confusion, the UMLER database maintains only the most recent car initial/number/type assignments for TTX equipment. (The TTX Company assigns car initials and car type by car number and, based on need, frequently and repeatedly reassigns series of car numbers to different initials and car types. The original number assignment usually refers to intermodal flatcars, but subsequent assignments have often related to multi-level flatcars.)
- Because the UMLER locates flatcars by comparing the car number with its assigned car initial and car type, reassignment of series numbers can complicate data analysis and lead to reporting errors in the edited database. For example, the car initial and car type currently assigned to a particular car number are written onto edited waybill records. An error flag, "14" will be appended to the record in UMLER if the car type no longer corresponds to certain codes (P, Q, or S). In many cases, however, at the time of the waybill movement, the car number was most likely assigned to a different car initial and to car type P, Q, or S.
- To reduce the number of waybill errors generated by this issue, intermodal waybills processed after September 1, 1995, have used the dummy car initial number GBRX 091193 in instances of traditional trailer-on-flat-car/container-on-flat-car (TOFC/COFC) movements.[110]

Intermodal Terminals Database

- The amount of effort required to keep the database current depends on the volatility of the data, which in turn depends on how much and how fast the intermodal infrastructure is changing. Changes to the intermodal infrastructure are being driven by the continued growth of containerized traffic, recent and proposed railroad mergers, the formation of ocean carrier alliances, technological advances, freight rate incentives, the availability of federal funding for intermodal projects, and many other events and factors. Although the number of intermodal terminals and intermodal connections added to or removed from the transportation system over the course of a year may be relatively small compared to the number of terminals in existence, it is nevertheless significant.[111]

6.7.6.3 Methodological Differences

Freight Analysis Framework (FAF3)

- Differences in the way the 2002 versus the 2007 Commodity Flow Survey (CFS) assigned water-only versus water-inclusive intermodal shipments (typically, truck-water combinations) make direct comparisons of water-only traffic volumes and modal shares problematic. Appendix A of the FAF3 shows the modal class changes between 2002 and 2007 and provides definitions for the modes.[112]
- For multiple-mode shipments, if a respondent has reported a shipment's mode of transport as both parcel and air, CFS treats the shipment as parcel only.

Vehicle Inventory and Use Survey (VIUS)

- The 2002 VIUS dropped the intermodal question (railroad, maritime, or domestic containers; piggyback trailers; or conventional trailers). The U.S. Census Bureau had requested that questions be considered for deletion to make room for the questions being added to the 2002 VIUS, and data users agreed that this question was either of limited use or the quality was questionable.[113]

6.7.7 Unknown/Other

6.7.7.1 Taxonomic Differences

Taxonomical differences among and within unknown/other mode data sources are a result of unique items included in the data element definitions.

Carload Waybill Sample

- For the data element corresponding to the TOFC/COFC service code, the code for the Intermodal Service Code (ISC) must be entered in the first position of the field. Three blanks in this field indicate that the movement is *not* intermodal in nature. Unknown ISCs are indicated by an X.
- For the data element corresponding to "All Rail/Intermodal," the number 9 indicates an unknown mode.
- For the data element corresponding to "Type of Move Via Water," the number 9 indicates an unknown mode.[114]

Fatal Analysis Reporting System (FARS)

- FARS uses a data element called "Transported to Medical Facility By," which reports details on travel to a medical facility via unknown modes using the following number codes:
  - 3 = EMS (emergency medical services), unknown mode
  - 4 = Transported by unknown sources
  - 9 = Unknown mode.

- Other mode travel is reported using the code "6—Other."[115]
- NHTSA updates the FARS Analytical User's Manual[116] every year to summarize the evolution of coding. When conducting analysis across years, data users should check every data element of interest in each year's coding manual.

Freight Analysis Framework (FAF3)

- The data element corresponding to "Other and Unknown" includes movements not elsewhere classified, such as flyaway aircraft and shipments for which the mode cannot be determined.[117]

North American Transborder Freight Database (Transborder)

- DISAGMOT uses numerical fields to identify the surface mode or other mode of transport of shipments entering or exiting the United States. DISAGMOT 8 ("Other and unknown"), includes "flyaway aircraft, or aircraft moving under their own power (i.e., aircraft moving from the aircraft manufacturer to a customer and not carrying any freight), powerhouse (electricity), vessels moving under their own power, pedestrians carrying freight, unknown, and miscellaneous other."[118]

6.7.7.2 Temporal Differences

Temporal differences among and within data sources related to unknown/other freight movement are a result of changes in methodology over time. Data sources may change their data collection or reporting methods over time, making it difficult to compare data elements across multiple years within a single data source, or across data sources.

Commodity Flow Survey (CFS)

- For the 2012 CFS, a change was made relating to mileage processing. Mode of transportation is now imputed whenever a respondent has provided a mode of "other," or "unknown," or otherwise failed to provide a modal response ("missing mode") for a shipment.
- During the 2007 CFS mileage processing, 2.4% of shipments had a respondent-provided mode of "unknown" or "other," and an additional 2.1% had no reported mode at all. Since all shipments must be properly routed to calculate a distance traveled, imputations were made. For 2012 CFS mileage processing, if the shipment weighed less than 80,000 pounds, it was routed via highway mode as a for-hire truck; if the shipment weighed 80,000 pounds or more, it was routed via rail mode.[119]

Fatal Analysis Reporting System (FARS)

- FARS uses a data element known as "Transported to Medical Facility By," which reported unknown mode travel under different codes before 2010. Data users should be aware of the changes and consult the appropriate FARS analytical reference guide for the proper codes.[120]

North American Transborder Freight Database (Transborder)

- DISAGMOT (the data element corresponding to mode of transport) uses number codes to identify the mode of transport of shipments entering or exiting the United States. Since April 1995, in response to inquiries from data users and a U.S. Census Bureau investigation, the Transborder database added DISAGMOT 9, signifying foreign trade zones, as a mode of transport. Data users are advised to keep in mind two things: (1) before April 1995, such imports were included under DISAGMOT 8 ("Other and unknown"); and (2) The actual mode of transportation is not available for imports coded under "foreign trade zones" using DISAGMOT 9. Although foreign trade zones are treated as a mode of transportation in this dataset, the actual mode for a specific shipment into or out of a foreign trade zone remains unknown because U.S. Customs and Border Protection (CBP) does not collect this information.

6.8 Safety Data Elements

Keywords: safety, risk, accident, incident, crash, collision, fatality, injury, property damage, hazardous material, driver, vehicle unit, first harmful event

6.8.1 Taxonomic Differences

The definitions of "fatal crash" and "fatality" were found to be consistent in the FARS, MCSMS, and FRA Safety Database; no taxonomic differences were identified.

Fatal Analysis Reporting System (FARS)[121]

- A fatal crash is a crash that involves a motor vehicle traveling on a trafficway customarily open to the public, and results in the death of an occupant of a vehicle or a non-occupant within 30 days (720 hours) of the crash.
- NHTSA updates the FARS Analytical User's Manual[122] every year to summarize the evolution of coding. When conducting analysis across years, data users should check every data element of interest in each year's coding manual.

Motor Carrier Safety Measurement Systems (MCSMS)[123]

- A fatality is any person killed in or outside of any vehicle (e.g., truck, bus, car) involved in a crash or who dies within 30 days of a crash as a result of an injury sustained in the crash.

Federal Railroad Administration (FRA) Safety Database[124]

- A fatality is defined as an individual who is confirmed dead within 30 days of a rail-transit-related incident.

6.8.2 Temporal Differences

Temporal differences among and within safety-related data elements are a result of changes in definitions and codes over time. Data sources sometimes change the data element definitions and codes over time to accommodate changes in the type of data collected and the way the data is presented. Caution should be exercised to ensure that the correct definitions and codes are used.

Fatal Analysis Reporting System (FARS)[125]

- FARS, which became operational in 1975, is a nationwide census providing the National Highway Traffic Safety Administration (NHTSA), Congress, and the American public yearly data regarding fatal injuries suffered in motor vehicle traffic crashes. A comprehensive coding manual has been produced each year. In addition, NHTSA updates the FARS Analytical User's Manual[126] every year to summarize the evolution of coding. When conducting analysis across years, data users should check every data element of interest in each year's coding manual.

- P22/NM21—This data element identifies the method of transportation provided to transport a person to a hospital or medical facility. Although this field exists in the 1975 and 1976 files, it is not initialized (i.e., it has no values in those years). This variable was expanded to include non-motorists in 2010.
- HAZ_CARG—From 1982 to 2006, this data element was used to identify the presence of hazardous cargo for a vehicle and to record information about the hazardous cargo when available. Since 2007, however, HAZ_CARG has been replaced with the following five data elements:
  - HAZ_INV—This data element identifies whether the vehicle was carrying hazardous materials.
  - HAZ_PLAC—This data element identifies the presence of hazardous materials for the vehicle and whether the vehicle displayed a hazardous materials placard.
  - HAZ_ID—This data element identifies the four-digit hazardous material identification number for the vehicle.
  - HAZ_CNO—This data element identifies the single-digit hazardous material class number for the vehicle.
  - HAZ_REL—This data element identifies whether any hazardous cargo was released from the cargo tank or compartment of the vehicle.
- Data users should be cautious about changes in attribute codes over time. For instance, the data elements BODY_TYP and TOW_VEH define vehicle categories such as passenger cars, pickups, buses, trucks, and so forth. These fields help differentiate freight-related fatal crash records from fatal crash records of other motor vehicles types. Table 6-1 summarizes temporal differences in truck-related codes.
- For additional examples, please see Appendix C of the FARS Analytical User's Manual,[127] which tabulates changes made to all FARS data elements since 1975.

6.8.3 Methodological/Reporting Differences

Methodological/reporting differences among and within safety-related data elements fall under the following categories:

- Methodological differences within a data source
- Reporting differences within a data source

Table 6-1. NHTSA's vehicle body type classification.

Classification (BODY_TYP)	Data Year and Code		
	1975–1981	1982–1990	1991–Later
Pickups	50	50, 51	30-39
Large Trucks	53-59, or (60 and tow_veh=1)	70-72, 74-76, 78, or (79 and tow_veh in 1-5)	60-64, 66, 67, 71, 72, 78, or (79 and tow_veh in 1-4)
Light Trucks & Vans	43, 50-52, or (60 and tow_veh=0)	12, 40, 41, 48-51, 53-56, 58, 59, 68, 69, or (79 and tow_veh=0 or 9)	14-22, 24, 25, 28-41, 45-49, or (79 and tow_veh =0 or 9)
Medium Trucks	53, 54, 56	70, 71, 75, 78	60-62, 64, 67, 71
Heavy Trucks	55, 57-59, or (60 and tow_veh=1)	72, 74, 76, or (79 and tow_veh in 1-5)	63, 66, 72, 78, or (79 and tow_veh in 1-4)
Combination Trucks	((53-56, 60) and tow_veh=1), or 57-59	((70-72, 75, 76, 78, 79) and tow_veh in 1-5), or 74	((60-64, 71, 72, 78, 79) and tow_veh in 1-4), or 66
Single-Unit Trucks	(53-56, 60) and tow_veh =0	(70-72, 75, 76, 78, 79) and tow_veh in (0,9)	(60-62, 63, 64, 67, 71, 72, 78, 79) and tow_veh in (0,5,6,9)

6.8.3.1 Methodological Differences Within a Data Source

Data users should be aware of methodological differences within a single data source that make certain types of analysis difficult.

Motor Carrier Safety Measurement Systems (MCSMS)[128,129]

- The MCSMS methodology is frequently updated by the Federal Motor Carrier Safety Administration (FMCSA) to include the most current set of violations being recorded from inspections. The original MCSMS methodology was developed based on the SafeStat measurement system. In January 2008, FMCSA started an Operational Model Test of the Compliance, Safety, and Accountability program. Notable milestones of the methodology changes in the Carrier Safety Measurement System (CSMS) portion of the MCSMS (as opposed to the Driver Safety Measurement Systems) are as follows:
  - CSMS[130] Methodology Changes from Version 1.2 to 2.0 (Implemented August 2010)
  - CSMS Methodology Changes from Version 2.0 to 2.1 (Implemented December 2010)
  - CSMS Methodology Changes from Version 2.1 to 2. 2 (Implemented January 2012)
  - CSMS Methodology Changes from Version 2.2 to 2.2.1 (Implemented August 2012)
  - CSMS Methodology Changes from Version 2.2 to 3.0 (Implemented December 2012)
  - CSMS Methodology Changes from Version 3.0 to 3.0.1 (Implemented August 2013)
  - CSMS Methodology Changes from Version 3.0.1 to 3.0.2 (Implemented June 2014)

6.8.3.2 Reporting Differences Within a Data Source

For some data sources, reporting mechanisms might change over time. Caution should be exercised when using or interpreting data in certain types of safety analyses.

Pipeline and Hazardous Material Safety Administration (PHMSA)

- The accident reporting criteria for hazardous liquid pipeline systems were revised in 1990, 1991, 1994, 1996, and 2002. For example, beginning in 1991, a release of carbon dioxide (50 or more barrels) was added as a type of hazardous materials (hazmat) accident. In addition, incident reporting criteria for gas transmission, gas gathering, and gas distribution pipeline systems were revised in 1990 and updated in 2011. For more information on these and other methodological changes, see the PHMSA Reporting Criteria Changes—1990–Current.[131]
- Beginning in 2005, incident reporting criteria were modified to include the discovery of undeclared hazmat. It is reported that these types of incidents consist of approximately 8% of total reported incidents, although about half of them do not indicate a release of hazmat or any other criteria for incident reporting. More information on data quality assessment can be found in the PHMSA publication, A Data Quality Assessment: Evaluating the major safety data programs for pipeline and hazardous materials safety (November 10, 2009).[132]

Federal Railroad Administration (FRA) Safety Database

- The reporting threshold for the Rail Equipment Accident/Incident Reporting Threshold table is updated annually. Starting with $750 for data released between 1957 and 1974, the reporting threshold increased to $6,700 for data released 2002–2005 and to $10,500 in 2014.[133]
- In response to changes associated with the Occupational Safety and Health Act, FRA amended its accident/incident reporting rules so that the data on occupational fatalities, injuries, and illnesses in the railroad industry is comparable with such data for other industries The changes were implemented beginning May 1, 2003.[134,135]

Motor Carrier Management Information System (MCMIS)[136]

- Beginning January 1, 1994, states participating in the Motor Carrier Safety Assistance Program were required to report through the SAFETYNET system a standard set of data items on all trucks and buses involved in traffic crashes that met a specific severity threshold. Reportable crashes include one or more of the following vehicle types:
  - A truck (used primarily for the transportation of property) having at least six tires in contact with the road surface
  - A vehicle displaying a hazardous material placard
  - A bus with seating for at least nine people (15 people before 2001), including the driver
- The Federal Motor Carrier Safety Administration (FMCSA) uses data from both the Fatal Analysis Reporting System (FARS) and MCMIS. The two databases may report different fatal crash counts because of variations in the way they define reportable vehicle configurations. FMCSA provides the FARS/MCMIS Fatal Crash Record Matching Tool to help reconcile differences between the FARS and MCMIS databases. The tool matches fatal large truck and bus crash records between the databases by comparing several key fields (e.g., county, date, time, VIN, DOT #) of large truck or bus fatal crash records.[137,138]

6.9 Units of Measurement Data Elements

Keywords: length, width, volume, depth, height, capacity, distance, monetary, passengers, time, weight

Definitions of these commonly used units of measurement vary among freight data sources:

- Distance
- Monetary data
- Passenger movements
- Time
- Volume: Traffic
- Volume: Water/vessels
- Weight
- Geospatial data

Applicable taxonomic, temporal, and methodological differences identified as part of NCFRP Project 47 are detailed in the following sections.

6.9.1 Distance

6.9.1.1 Taxonomic Differences

Air Carrier Statistics

- DISTANCE GROUP—This data element measures the distance of a flight segment in 500-mile increments using code numbers 1–17. [139]

Federal Railroad Administration (FRA) Safety Database

The FRA Safety Database reports distance using several distinct data elements. Data users should be aware of the differences to ensure that the correct distance measure is being used for analysis.[140]

- LOCOMI—This data element reports the number of locomotive train-miles traveled in the month. A train-mile is defined as the movement of a train for a distance of 1 mile. Data users should note that the presence of multiple locomotives in the train does not affect the mileage calculation.

- MTMI—This data element reports the number of motor train-miles for the month.
- YSMI—This data element reports the number of yard-switching train-miles for the month, which represents the miles traveled while the train is engaged in yard-switching service.
- TOTMI—This data element indicates the total miles as reported on Form FRA F6180.55, Railroad Injury and Illness Summary.
- PASSMI—This data element reports the number of passenger-miles for the month. A passenger-mile is defined as the movement of a passenger for a distance of 1 mile.
- FRTRNMI—This data element reports the number of train-miles in freight service during the month.
- PASTRNMI—This data element reports the number of train-miles in passenger service during the month, defined as the movement of a passenger for a distance of 1 mile.
- OTHERMI—This data element reports any other train-miles not included in freight, passenger, or yard-switching train-miles.

Vehicle Inventory and Use Survey (VIUS)

- MILES_ANNL—This data element reports the number of miles a vehicle was driven in the reporting year without adjusting for partial-year ownership of the vehicle. Data users should be aware that this data element may reflect additional miles traveled when vehicles were not owned by the respondents.
- MILES_ANNLNOIMP—This data element reports the number of miles a vehicle was driven in the reporting year as adjusted for partial-year ownership of the vehicle.
- TAB_MILES—This data element indicates the weighted annual truck-miles driven during 2002 after applying the expansion factor for trucks (the TAB_TRUCKS data element).[141]

6.9.1.2 Methodological Differences

Federal Railroad Administration (FRA) Safety Database

- YSMI—This data element reports the number of yard-switching train-miles. The *FRA Guide for Preparing Accident/Incident Reports* advises that, if actual mileage is not known, YSMI can be computed at the rate of 6 mph for the time actually engaged in yard-switching service.[142]
- FRTRNMI—This data element reports the number of freight train-miles run by a railroad on its own track during the month. Data users should be aware that FRTRNMI reports freight train-miles by railroad, not by track; it does not aggregate train-miles reported by the railroad that owns the track together with train-miles that may be reported by another railroad, which may occur if one railroad's equipment is being operated over the track by a different railroad's crew. In such cases, the railroad of the crew operating the equipment enters the freight train-miles on their own FRA form.[143]

6.9.2 Monetary Data

6.9.2.1 Taxonomic Differences

Air Carrier Financial Report

- Beginning on October 18, 2006, numbers reported in the Schedule B-1, B-1.1, P-1.1, and P-1.2 data tables began following the format of common public financial documents, such as reports filed with the Securities and Exchange Commission or company financial statements. This format reverses signs from the accounting format in which numbers appeared before that date.[144]

6.9.2.2 Methodological Differences

Carload Waybill Sample

- The Surface Transportation Board (STB) classifies railroads based on their annual operating revenues as either Class I ($250 million or more), Class II ($20 million or more), or Class III ($0–$20 million). The average index (deflator factor) is based on the annual average Railroad Freight Price Index for all commodities. The formula below is used to adjust a railroad's operating revenues to eliminate the effects of inflation.

  Current Year's Revenues × (1991 Avg. Index / Current Year's Avg. Index)

- EXPANDED TOTAL REVENUE—This data element indicates the total freight revenue (item 15 in the STB Reference Guide[145]) multiplied by the expansion factor (item 88). Revenue splits are calculated by dividing the waybill's expanded freight revenue figure by the number of 100-mile blocks traveled by each railroad in the route. The origin railroad is apportioned revenue for an additional block to allow for pickup and switching expenses. Likewise, the termination railroad is credited with revenue for an additional block, to allow for delivery expenses.[146]
- TOTAL VARIABLE COST—This data element indicates the expanded variable cost for all railroads in the waybill computed using the Uniform Railroad Costing System (URCS). The URCS produces average variable costs for Class I railroads using railroad-specific accounting and operating data. Costs for local and regional railroads use URCS regional data. See the STB Reference Guide for more details on the methodology used to calculate TOTAL VARIABLE COST.[147]

Commodity Flow Survey (CFS)

- VALUE (MILLION $)—This data element reports the dollar value, in millions of dollars, of the entire shipment. This is defined as the net selling value, exclusive of freight charges and excise taxes. Data users are advised to note that the total value of shipments as measured by the CFS and the U.S. gross domestic product (GDP) provide different measures of economic activity in the United States and are not directly comparable. GDP is the value of all goods produced and services performed by labor and capital located in the United States. As measured by the CFS, the value of shipments is the market value of goods shipped from manufacturing, mining, wholesale, and select retail and service establishments, as well as warehouses and managing offices of multiunit establishments.[148] Table 6-2 highlights three important differences between GDP and CFS value of shipments.

Table 6-2. Differences between GDP and CFS value of shipments.

GDP	CFS
Captures goods produced by all establishments located in the United States	Measures goods shipped from a subset of all goods-producing establishments
Measures the value of goods produced and of services performed	Measures the value of goods shipped
Counts for only the value added at each step in the production of a product	Captures the value of shipments of materials used to produce or manufacture a product, as well as the value of shipments of the finished product itself*

*This means that the value of the materials used to produce a particular product contributes multiple times to the value of the commodity in the CFS.

North American Transborder Freight Data (Transborder)

- Although Transborder contains data on exports to and imports from Canada and Mexico, all data elements that report monetary information (e.g., FREIGHT, VALUE) are reported in U.S. dollars.[149]

6.9.3 Passenger Movements

6.9.3.1 Methodological Differences

Air Carrier Statistics

- In the T-100 Market Data, a passenger is "enplaned" and is counted only once as long as he or she remains on the same flight. In the T-100 Segment Airline Traffic Data, a passenger is "transported" and is counted for each leg of the trip. Therefore, the numbers in the segment data will tend to be higher than those in the market data (except for international flights). The Bureau of Transportation Statistics (BTS) generally uses market data for passenger, freight, or mail totals, as shown in this example provided by the U.S. DOT:[150]

> For example, 250 people take a flight from JFK (Point A) to BWI (Point B), where 200 passengers deplane and the other 50 passengers, along with 70 additional passengers, continue on to MIA (Point C), where all passengers deplane.
>
> In [the market dataset, (T-100 Market Data)], Point A to Point B would be counted as one market of 200 and Point A to Point C would be counted as another market of 50. Point B to Point C would be a market of 70 people.
>
> In [the segment dataset (T-100 Segment Airline Traffic Data)], Point A to Point B would be counted as one segment of 250 and Point B to Point C would be counted as another segment of 120. A passenger from A to B to C would be counted for both legs.
>
> A to B: 200 Market, 250 Segment
> B to C: 70 Market, 120 Segment
> A to C: 50 Market (no Segment)[151]

Border Crossing/Entry Data

- A *passenger* is defined as a person entering the United States at a particular port in a privately owned vehicle, pickup truck, motorcycle, recreational vehicle, taxi, ambulance, hearse, tractor, snowmobile, or other motorized private ground vehicle.
- A *pedestrian* is a person arriving on foot or by certain conveyance (such as a bicycle, moped, or wheelchair) requiring U.S. Customs processing.[152]

6.9.4 Time

6.9.4.1 Taxonomic Differences

Air Carrier Statistics

- Air Carrier Statistics reports quarterly data using specific timeframes; however, as shown in Table 6-3, calendar quarters as defined by Air Carrier Statistics may differ from calendar quarters as defined by other data sources, including the U.S. government:

Table 6-3. Calendar quarter definitions.

	Air Carrier Statistics[153]	U.S. Government[154]
Q1	January 1–March 31	October 1–December 31
Q2	April 1–June 30	January 1–March 31
Q3	July 1–September 30	April 1–June 30
Q4	October 1–December 31	July 1–September 30

- To ensure that they are using the desired unit of measurement for analysis, data users need to recognize the difference between the data elements RAMPTORAMP (or RAMPTIME) and AIRTIME.
  - RAMPTORAMP (or RAMPTIME)—This data element reports the time computed from the moment the aircraft first moves under its own power for purposes of flight until it comes to rest at the next point of landing.[155]
  - AIRTIME—This data element, on the other hand, reports the airborne hours of the aircraft, computed from the moment it leaves the ground until it touches the ground at the end of a flight stage.[156]

Carload Waybill Sample

- Some data elements in the Carload Waybill Sample report dates using different codes, which could create difficulty in making direct joins with data element within the data source, as well as with data elements from other data sources. The data element WAYBILL DATE, for example, uses the data coding system *mmddccyy* (month, day, century, year), while the data element DEREGULATION DATE uses the coding system *ccyymmdd.*[157]

6.9.4.2 Temporal Differences

Fatal Analysis Reporting System (FARS)

- The 2010 FARS incorporated many changes, most of which resulted from efforts by the National Highway Traffic Safety Administration (NHTSA) to standardize variables in FARS and the National Automotive Sampling System's General Estimates System (GES).[158] Three substantial changes regarding FARS data elements related to time were:
  - CRASH DATE—This data element added GES element information, including new GES Special Instructions. The new reporting system removed Attribute 98, "Not Reported for Both Month and Day."
  - CRASH TIME—This data element added GES element information, including new GES Special Instructions. The new reporting system removed Attribute 9988, "Not Reported."
  - DEATH_TM—This data element records the hour and minute of a person's death using a four-digit coding system and the 24-hour clock format. Data from 1975 to 2008, however, followed a slightly different reporting format than did data from 2009 and later, as shown in Table 6-4:
- NHTSA updates the FARS Analytical User's Manual[159] every year to summarize the evolution of coding. When conducting analysis across years, data users should check every data element of interest in each year's coding manual.

Table 6-4. Changes in format of DEATH_TM records.

	1975–2008	2009 and Later
Midnight	2400	0000
Time of Death (*hhmm* Format)	0001–2359	0001–2359
Not Applicable (Non-Fatal)	--	8888
Unknown	9999	9999

6.9.5 Volume: Traffic

6.9.5.1 Taxonomic Differences

Freight Analysis Framework (FAF3)

- Data users should be aware that local truck traffic that is not part of FAF3.1 (FAF version 3.1) truck estimates is provided under two data elements:
  - NONFAF07 This data element is used for current traffic.
  - NONFAF40 This data element is used for forecast traffic.[160]

Highway Performance Monitoring System

- AADT_SINGLE_UNIT—This data element represents the annual average daily traffic (AADT) for single-unit trucks and buses, which are defined as vehicle classes 4 through 7 (buses through single-unit trucks with four or more axles).
- AADT_COMBINATION—This data element represents the AADT volume for combination unit trucks. Combination trucks are defined as vehicle classes 8 through 13 (single-trailer trucks with four or fewer axles through multi-trailer trucks with seven or more axles).[161]

6.9.5.2 Temporal Differences

Freight Analysis Framework (FAF3)

- FAF3 uses 2008 Highway Performance Monitoring System (HPMS) data to determine annual average daily traffic (AADT) for the year 2007. Data users are advised to note the temporal difference between the HPMS data and the FAF3 reporting year.[162]

Highway Performance Monitoring System (HPMS)

- FUTURE_AADT—This data element represents a 20-year forecast annual average daily traffic (AADT), which may cover a period of 18 to 25 years from the year of the data submittal.[163]

6.9.5.3 Methodological Differences

Highway Performance Monitoring System (HPMS)

- Data users are advised to note that the annual average daily traffic (AADT) for the National Highway System, Interstate, Principal Arterials (OFE, OPA), and HPMS Sample Panel sections are typically based on traffic counts taken on a minimum 3-year cycle, while AADT for the Non-Principal Arterial System and Non-Sample Panel sections are typically based on a minimum 6-year counting cycle.[164]
- HPMS guidance requires that growth factors be applied if the AADT is not derived from current year counts. For specific guidance on factor development recommended for HPMS data, see the *Traffic Monitoring Guide.*[165]

- HPMS requires that vehicle classification counts be adjusted to represent average conditions as recommended in the FHWA's *Traffic Monitoring Guide;* see that guide for specific guidance on count adjustments used in the HPMS.[166,167,168]

6.9.6 Volume: Water/Vessels

6.9.6.1 Taxonomic Differences

U.S. Waterway Data

- ACTUALCY—This data element reports the actual cubic yards dredged.
- NRT—This data element reports vessel net tonnage, defined as the volume of space available for the accommodation of passengers and the stowage of cargo, expressed in units of 100 cubic feet for each net ton. Data users are advised to note the difference between NRT and tonnage capacity, which simply expresses a volume capacity for passengers and cargo. For a more detailed discussion of how to calculate vessel net tonnage, see the 2012 *Waterborne Transportation Lines of the United States.*[169]

6.9.7 Weight

6.9.7.1 Taxonomic Differences

Foreign Trade Statistics (FTS)

- AIR_SWT_MO—This data element, representing air shipping weight, reports the gross weight in kilograms of shipments made by air, including the weight of moisture content, wrappings, crates, boxes, and containers (other than cargo vans and similar substantial outer containers).[170]

North American Transborder Freight Data (Transborder)

- SHIPWT—This data element, representing shipping weight, reports the gross weight of shipments of imports (and some exports) in kilograms, including the weight of moisture content, wrappings, crates, boxes, and containers (other than cargo vans and similar substantial outer containers). SHIPWT does not include data for exports shipped by land modes of transportation and reported using paper Shipper's Export Declarations documents; however, export weight (SHIPWT) is required to be filed for all modes of transportation using the Automated Export System.

Vehicle Travel Information System Documentation

- TOTAL WEIGHT OF VEHICLE—This data element reports the gross vehicle weight to the nearest tenth of a metric ton (100 kilograms). Data users are advised to note that this measurement differs from measurements based on a short ton (2,000 pounds), which are often used by other similar data sources.[171]

Vehicle Inventory and Use Survey (VIUS)

- WEIGHT_SIZE—This data element reports the average weight of the vehicle or vehicle/trailer combination grouped into the following ranges:
  - Light—The average vehicle weight is 10,000 pounds or less.
  - Medium—The average vehicle weight is 10,001 to 19,500 pounds

- Light-heavy—The average vehicle weight is 19,501 to 26,000 pounds
- Heavy-heavy—The average vehicle weight is 26,001 pounds or more.[172]

This classification may be different from that used by other data sources.

6.9.7.2 Temporal Differences

Fatal Analysis Reporting System (FARS)

- GVWR—This data element reports the gross vehicle weight rating. In 2007, GVWR was modified to allow gross combination weight rating (GCWR) to be recorded for combination vehicles to match the nationally accepted reporting criteria for GVWR (i.e., FMCSA's SAFETYNET and Model Minimum Uniform Crash Criteria). Use of GCWR instead of GVWR will impact only these vehicles:
  - Light trucks, 10,000 lbs. or less, pulling trailers (truck/trailers) (greater than 10,000 pounds GCWR)
  - Single-unit trucks, less than 26,000 lbs., pulling trailers (truck/trailers) (greater than 26,000 pounds GCWR)[173]
- NHTSA updates the FARS Analytical User's Manual[174] every year to summarize the evolution of coding. When conducting analysis across years, data users should check every data element of interest in each year's coding manual.

6.9.7.3 Methodological Differences

Carload Waybill Sample

- TARE WEIGHT OF CAR—This data element reports the light weight for each car (i.e., not an average) in hundreds of pounds. Data users are advised to note that, if articulated, the tare weight represents the sum of the light weight vehicles for the total number of units of the consist (the set of vehicles forming a complete train).[175]
- Freight weight statistics in the Carload Waybill Sample are based on billed rather than actual lading weights. Even though the overall difference between billed and actual weights is small, statistically significant variation does exist among many individual commodities. Consequently, the use of billed weights in certain types of waybill analysis can lead to biased conclusions for a variety of reasons. The Surface Transportation Board (STB) therefore advises that it is unwise to extrapolate weight-related calculations to multiple decimal point levels of precision.[176]
- EXACT EXPANSION FACTOR—Each waybill uses an expansion factor (EXACT EXPANSION FACTOR) to expand car, ton, trailer/container, and revenue statistics to 100% levels. For example, the data element EXPANDED TONS reports the billed weight in tons multiplied by the expansion factor. The expansion factor is calculated according to the following formula:

  Factor = (Population count / Sample count)

Commodity Flow Survey (CFS)

- Data users are advised to note that the ton totals in the CFS represent the sum of separate shipments of a commodity as it moves through the production and consumption segments of the supply chain; hence, the tonnage of goods may be counted more than once in the production life cycle (e.g., goods that are moved through distribution centers).[177]

North American Transborder Freight Data (Transborder)

- SHIPWT—This data element reports shipping weight for all imports but only certain exports in the Transborder database. Historically, shipping weight information from the U.S. Census Bureau has been available for shipments by vessel and air only. In the Transborder database, shipping weight data is available for all import modes. For exports, Transborder SHIPWT data is available for air and vessel modes but not for surface modes.[178]

6.10 Geospatial Data

The two main sources of geospatial transportation data are the National Transportation Atlas Database (NTAD) and the Topologically Integrated Geographic Encoding and Referencing (TIGER).

The NTAD is a compilation of multiple transportation data sources provided by the U.S. DOT and other federal agencies. The Bureau of Transportation Statistics (BTS) maintains and distributes the NTAD. However, the contributing agencies are responsible for the maintenance and accuracy of the data.

TIGER, which is maintained by the U.S. Census Bureau, is made up of severable file types containing census geographic data and information such as geographical boundaries, roads, rivers, lakes, cities, census blocks groups, and census tracts.

Geographical features contained in these two data sources may sometimes overlap; however, the attributes (or geographical information) contained in each data source may vary. Some of the geospatial data provided by TIGER is made available in the NTAD and vice versa. Data sources compiled within the NTAD and TIGER are available at their respective websites and the Freight Data Dictionary web application. For additional information, please review the database's metadata. Identifying the differences within the attributes of each geospatial data sources was beyond the scope of NCFRP Project 47, as was merging and combining of the data elements contained in these data sources. The study team recommends that additional information on geospatial data integration be sought from other well-versed sources.

Other publicly available freight-related data sources not included in the NTAD and TIGER are:

- Cropscape,[179] which is provided and maintained by the National Agricultural Statistics Service (NASS), and
- National Corridors Analysis and Speed Tool (N-CAST),[180] which is administered by the American Transportation Research Institute (ATRI) through an agreement with FHWA.

Private-sector geospatial data sources containing commodity flow information include:

- Transearch by I Global Insight[181]
- vFreight by the Economic Development Research Group[182]

Additional information on these data sources is available on their respective websites and the Freight Data Dictionary web application.

Endnotes for both Chapter 6 and Chapter 7 are listed in the References section.

CHAPTER 7

Resolving Differences in Data Element Definitions

7.1 Introduction

The simple fact that the data elements contained in the various datasets may be defined, measured, and reported differently does not indicate that the datasets cannot be used in tandem or combined in a single analysis. The differences simply require that some effort be expended to normalize the data for the intended application. This chapter provides guidelines to reconcile, harmonize, and create statistical bridges and crosswalks to resolve differences in data element definitions when combining those elements for an analysis. It also provides guidance on circumstances in which crosswalks may not be statistically sound.

7.2 Methodology

Five main topics were chosen for bridge development because of their importance in many different facets of freight data analyses. These include place names, units of measurement, commodity and industry classification systems, and modes of transport. The reconciliation process involved identifying the nature of the differences, identifying commonalities *within* the differences, and determining whether the differences are statistically significant, or whether they are inconsequential for freight data analysis. For statistically significant differences, recommendations are provided on whether a bridge should or should not be applied, the parameters required for applying each bridge, and limitations of the crossover methodologies.

Several sources were used in developing bridges or crosswalks for each topic, including the RBCS, user guides, and metadata associated with each data source. These documents provided detailed attribute descriptions and caveats for using the data. Along with these primary sources, various secondary sources such as academic papers and online data guides were also utilized. An example is the crosswalk between the North American Industry Classification System (NAICS) and the Standard Industrial Classification (SIC) system developed by the NAICS Association.[183]

Each discussion follows the same general format:

1. Topic (e.g., "[Place Name] Bridges").
2. Keywords—Search terms related to the topic of discussion.
3. Type of Bridge, which may include:
   - Taxonomic Bridges (if applicable)—These bridges apply to data differences that result from how the data elements are classified.
   - Temporal Bridges (if applicable)—These bridges apply to data differences that result from how the definitions of data elements vary over time.
   - Methodological/Analytical Bridges (if applicable)—These bridges apply to data differences that result from how the data is collected, processed, and disseminated by the various reporting agencies.

4. Data Sources Discussed—A list of data sources included in each bridge discussion for taxonomic, temporal, and methodological bridges. This list can be expanded in future work to include additional data sources not cited in a topic's discussion.

Disclaimers

- Privately held freight data sources are excluded from the bridging discussions at this time because of confidentiality concerns and the unavailability of certain data sources. Incorporating private data into the data discussions in the future will enhance the value of the Freight Data Dictionary in addressing the full range of potential problems data users might encounter when working with freight data.
- Data element names as they exist in the actual data sources are represented in ALL CAPS: for example, ORIGINID, DESTINATIONID, and COMMODITYID.

7.3 Place Name Bridges

Keywords: country, state, county, city, place, metropolitan area

Place-identifier data elements identify the origin-destination of freight movement or the location of an event (e.g., an accident). Place identifiers from one data source often cannot be used with data elements from other data sources given the taxonomic or methodological differences between data sources. The tables presented in this discussion can be used to bridge such differences so that place identifiers from different data sources can be used with one another for freight data analysis.

Within tables 7-1, 7-2, and 7-3, place-identifier data elements for each data source are categorized into columns based on their geographic classification system (e.g., two-letter postal code vs. FIPS [Federal Information Processing Standard] code vs. full text name). Data elements from different sources that are located in the same column can be used with one another with no bridging required. If data elements from multiple sources are in different columns (i.e., different classification systems), however, users can determine which type of bridge needs to be performed and then implement the bridge using the appropriate conversion tables, which are included at the end of this discussion.

Assume a data user wants to compare U.S. Waterway data for the state of Texas with County Business Pattern data for Texas. The state data elements table (Table 7-2) shows that the U.S. Waterway data element "STATE" uses a two-letter postal code, whereas the County Business Pattern data element "FIPSSTATE" uses the FIPS system. To bridge these two state data elements, users can consult the State Name Crosswalk (Table 7-3) to determine which FIPS state numeric code corresponds with the U.S. Postal Service's Texas postal code.

State and county FIPS codes can be further paired with census tracts and block groups. These two subsequent geographies form part of the larger regions using a 12-digit numeric code (see Figure 7-1). In the figure, the first two digits signify the state of Texas (48), followed by Travis County (453), the tract—in this case, Tract 7 (0007.00), and the block group (1). The tract number is a six-digit number with two digits after a decimal. The first four digits identify the tract number and the digits after the decimal identify changes in subdivisions.[184,185]

7.3.1 Country Data Elements

Table 7-1 identifies country data elements across multiple databases with similar reporting codes.

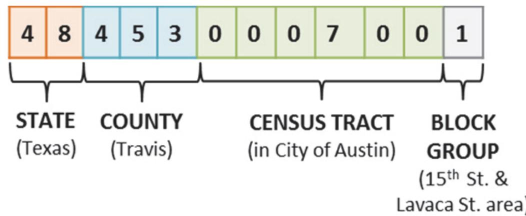


Figure 7-1. Sample 12-digit place identifier using numeric code.

The following resources provide information on how to bridge data elements that identify the same geographic unit but use different classification systems or codes:

- FIPS (PUB 10-4) Country Code to County Name (full text)
  To convert between FIPS country codes and country name (full text), users can consult the Geopolitical Entities and Codes resource developed by the National Geospatial-Intelligence Agency. Users should note that FIPS publication 10-4 was withdrawn by the National Institute of Standards and Technology (NIST) in 2008 as a Federal Information Processing Standard.
- Four-digit Schedule C or ISO Code to Country Name (full text)
  To convert between four-digit Schedule C or ISO Country Codes and the full text country name, users can consult the Schedule C Country Codes and Descriptions page maintained by the US Census Bureau.

Table 7-1. Country data elements.

	Two-letter Abbreviation (FIPS PUB 10-4)	Four-digit Country Code (Schedule C or ISO Code)	Full Name (text)
Air Carrier Statistics	ORIGINCOUNTRY DESTCOUNTRY		OriginCountryName DestCountryName
Carload Waybill Sample	ORIGIN RAILROAD COUNTRY CODE FIRST (SECOND, THIRD FOURTH, FIFTH, SIXTH) INTERCHANGE RAILROAD COUNTRY CODE TERMINATION RAILROAD COUNTRY ROAD		
U.S. Waterway Data		CTRYCODE CTRY_F CTRY_C ITCTRY	
National Agriculture Statistics Service (NASS)		COUNTRY_CODE	
North American Transborder Freight Data (Transborder)		COUNTRY	
National Ballast Information Clearinghouse Database			Last Country

7.3.2 State Data Elements

Table 7-2 identifies state data elements across multiple databases with similar reporting schemas. Consult Table 7-3 for a crosswalk between State Names, FIPS/ANSI/GSA Numeric State Codes, and two-character Postal Codes.[186]

7.3.3 County Data Elements

Table 7-4 identifies county data elements across multiple databases with similar reporting codes. The Census 2010 FIPS Codes for Counties and County Equivalent Entities online database provides a crosswalk between County FIPS codes and counties and equivalent entities. Users should also consult the Census-published *Substantial Changes to Counties and County*

Table 7-2. State data elements.

	Two-character Postal Code	Two-digit FIPS/ANSI/ GSA Code	Full Name (text)
Air Carrier Statistics	ORIGINSTATE DESTSTATE	ORIGINSTATEFIPS DESTSTATEFIPS	OriginStateName DestStateName
Carload Waybill Sample	ORIGIN STATE ALPHA TERMINATION STATE ALPHA		
County Business Patterns		FIPSTATE	
Federal Railroad Administration Safety Database		STATE	
Freight Analysis Framework	STATE	STATEFIPS DMS_ORGST DMS_DESTST	
U.S. Waterway Data	STATE ORIGIN DEST		
Survey of Business Owners		FIPST	
National Agricultural Statistics Service (NASS)	STATE_ALPHA	STATE_ANSI STATE_FIPS_CODE[a]	state_name (State)
Motor Carrier Management Information System (MCMIS)	COUNTY_CODE_STATE ORIG_REPORT_STATE INSP_CARRIER_STATE SHIPPER_STATE REPORT_STATE STATE STATE_ISSUING_NUMBER	PHY_ST	
Vehicle Travel Information System Documentation	ABBREV	STATECODE STATE FIPS CODE	NAME
Vehicle Inventory and Use Survey (VIUS)	HB_STATE		
North American Transborder Freight Data (Transborder)	USASTATE		U.S. State
CTAA Intermodal Terminals Database		STFIPS	

Note: For state identifiers, FIPS, ANSI, and GSA codes are interchangeable. For a description of the relationship between American National Standards Institute (ANSI) Codes and FIPS Codes, users should consult the following resource: http://www.census.gov/geo/reference/ansi.html.

[a] For this data source, GSA two-digit state codes also include "99" and "98" for US TOTAL and OTHER STATES, respectively; otherwise they match ANSI codes.

Table 7-3. State name crosswalk.

Name	FIPS/ANSI/GSA Numeric Code	Postal Code (USPS)	Name	FIPS/ANSI/GSA Numeric Code	Postal Code (USPS)
Alabama	01	AL	Montana	30	MT
Alaska	02	AK	Nebraska	31	NE
Arizona	04	AZ	Nevada	32	NV
Arkansas	05	AR	New Hampshire	33	NH
California	06	CA	New Jersey	34	NJ
Colorado	08	CO	New Mexico	35	NM
Connecticut	09	CT	New York	36	NY
Delaware	10	DE	North Carolina	37	NC
District of Columbia	11	DC	North Dakota	38	ND
Florida	12	FL	Ohio	39	OH
Georgia	13	GA	Oklahoma	40	OK
Hawaii	15	HI	Oregon	41	OR
Idaho	16	ID	Pennsylvania	42	PA
Illinois	17	IL	Rhode Island	44	RI
Indiana	18	IN	South Carolina	45	SC
Iowa	19	IA	South Dakota	46	SD
Kansas	20	KS	Tennessee	47	TN
Kentucky	21	KY	Texas	48	TX
Louisiana	22	LA	Utah	49	UT
Maine	23	ME	Vermont	50	VT
Maryland	24	MD	Virginia	51	VA
Massachusetts	25	MA	Washington	53	WA
Michigan	26	MI	West Virginia	54	WV
Minnesota	27	MN	Wisconsin	55	WI
Mississippi	28	MS	Wyoming	56	WY
Missouri	29	MO			

Table 7-4. County data elements.

	FIPS/ANSI/GSA Code	Full Name (text)
Carload Waybill Sample	ORIGIN FIPS CODE TERMINATION FIPS CODE	
Fatality Analysis Reporting System (FARS)	COUNTY	
Federal Railroad Administration Safety Database	CNTYCD COUNTY	
Freight Analysis Framework (FAF3)	CTFIPS	
U.S. Waterway Data		COUNTY1
National Agricultural Statistics Service (NASS)	COUNTY_ANSI	COUNTY_NAME (COUNTY)
Motor Carrier Management Information System (MCMIS)	COUNTY CODE (COUNTY_CODE)	COUNTY NAME
Highway Performance Monitoring System	COUNTY_CODE	
Vehicle Travel Information System Documentation	COUNTYCODE COUNTY FIPS CODE	

Note: For state identifiers, FIPS, ANSI, and GSA codes are interchangeable. For a description of the relationship between American National Standards Institute (ANSI) codes and FIPS codes, users should consult the following resource: http://www.census.gov/geo/reference/ansi.html.

Equivalent Entities[187] and the Missouri Census Data Center Geographic Correspondence Engine (MABLE/Geocorr)[188] when working with county place-identifier data elements to be aware of county boundary changes that have occurred since 1970 and may influence the accuracy of bridging county definitions between multiple years. In 2010, for example, Petersburg Borough (02-195) in Alaska was created from part of former Petersburg Census Area (02-195) and part of Hoonah-Angoon Census Area (02-105). In Virginia, Bedford (independent) city (51-515) was changed to town status and added to Bedford County (51-019).[189]

7.3.4 Statistical/Economic Area Data Elements

Table 7-5 identifies statistical/economic area data elements across multiple databases with similar reporting codes.

Below are recommendations for bridging data elements that identify the same geographic unit but use different classification systems or codes:

- For a crosswalk between the five-digit U.S. DOT codes for ORIGINCITYMARKETID and DESTCITYMARKETID used in the Air Carrier Statistics database and corresponding metropolitan areas, users can consult the data definitions provided at BTS.gov.[190]
- For a crosswalk between three-digit Business Economic Area (BEA) codes used in the Carload Waybill Sample and corresponding metropolitan areas, see the STB BEA Codes (Table 7-5).[191]
- For appropriate concordances between Metropolitan Statistical Areas, Micropolitan Statistical Areas, Metropolitan Divisions, and Combined Statistical Areas, see the BEA Statistical Areas online conversion tool.[192]

7.3.5 City/Other Place Name Data Elements

Table 7-6 identifies city and other place name data elements across multiple databases with similar reporting codes.

Following are recommendations for bridging data elements that identify the same geographic unit but use different classification systems or codes:

- For a crosswalk applicable to FIPS/ANSI place codes, users can consult the 2010 ANSI Codes for Places online conversion tool.[193]
- Standard Point Location Codes use a six-digit system of nested two-digit codes with the following pattern: STATE–COUNTY–CITY (POINT). These two-digit codes are based off FIPS codes; therefore, to derive the correct FIPS City code from the SPLC code, users can simply look at the last two digits.[194]

Table 7-5. Statistical/economic area data elements.

	Metropolitan Statistical Area (Name)	Combined Statistical Areas	Business Economic Area (BEA) Codes	Other
Air Carrier Statistics				ORIGINCITYMARKETID[a] DESTCITYMARKETID
Carload Waybill Sample		ORIGIN SMSA TERMINATION SMSA	ORIGIN BEA AREA TERMINATION BEA AREA	
County Business Patterns	MSA			

[a] Origin Airport, City Market ID. City Market ID uses an identification number assigned by US DOT to identify a city market. This field can be used to consolidate airports serving the same city market.

Table 7-6. City and other place name data elements.

	Full Name (text)	Standard Point Location (SPLC) Code	FIPS Place Code
Air Carrier Statistics	ORIGINCITYNAME DESTCITYNAME		
Carload Waybill Sample		ORIGIN SPLC [a] DESTINATION SPLC	
County Business Patterns	CITY		
Federal Railroad Administration Safety Database	CITYNAM		city CITYCD
Motor Carrier Management Information System (MCMIS)	SHIPPER_CITY CITY		CITY_CODE

[a] City can be inferred from the SPLC database published by the National Motor Freight Traffic Association (NMFTA) – Information available at http://www.nmfta.org/pages/splc.

7.4 Units of Measurement Bridges

Keywords: area, distance, value, monetary, volume, weight

Temporal, methodological, or taxonomic differences exist between units of measurements in the various data sources. A user may wish, for example, to compare the shipping weight of a commodity in one data source that uses metric tonnage as the unit of measurement with that from another data source that uses hundredweight. Section 7.4.1 provides guidance on situations in which it is appropriate to use similar data sources with disparate units of measurement, and presents several bridges and crosswalks that users can employ when conducting freight data analysis.

7.4.1 Area

Foreign Trade Statistics (FTS)

In 1989 the United States adopted the Harmonized Commodity Description and Coding System, commonly referred to as the Harmonized System (HS), to classify exports and imports. Given that this system collects information based on the metric standard, Table 7-7 can assist users in converting measures of area from other data sources to metric quantities and values for use with foreign trade statistics.[195]

Table 7-7. Converting units of quantity to HS (metric) units.

Reported Units of Quantity Name (Abbreviation)	HS Units of Quantity Name (Abbreviation)	Multiplication Factor to Convert Reported Units to HS Units
Square centimeter (SCM)	Square meter (SQM)	0.0001
Square meter (SQM)	Square centimeter (SCM)	10000
Square feet (SFT)	Square meter (SQM)	0.0929
Square inch (SQT)	Square meter (SQM)	0.0006452
Square inch (SQT)	Square centimeter (SCM)	6.452
Square yard (SYD)	Square meter (SQM)	0.8361
Thousand square ft (MSF)	Square meter (SQM)	92.9

Table 7-8. Crosswalk between square footage and other measures of area.

Area Unit of Measurement	Area Equivalent
1 square mile	27,878,400 square feet
1 square mile	640 acres
1 square mile	258.999 hectares
1 acre	43,560 square feet
1 acre	0.0015625 square miles
1 acre	0.404686 hectares
1 hectare	2.47 acres
1 hectare	0.99386 square miles

Topologically Integrated Geographic Encoding and Referencing (TIGER)

The data elements ALAND and AWATER, which report land area and water area, respectively, report area in square feet. Table 7-8 provides a crosswalk between square footage and other commonly used measures of area found in freight data sources.[196] (Note: use Table 7-7 to convert to metric units of measurement).

7.4.2 Distance

Air Carrier Statistics

The Air Carrier Statistics database uses distance intervals to classify flight segment distances under the data element DISTANCEGROUP. Users wishing to compare Air Carrier Statistics flight segments distances to distance measurements from other data sources can use Table 7-9. Also included in this table is the equivalent distance measurement in kilometers.[197]

Table 7-9. Comparing flight segment distances.

"DistanceGroup" Code	Distance Interval	Metric Equivalent	"DistanceGroup" Code	Distance Interval	Metric Equivalent
1	< 500 miles	805 km	11	5,000 - 5,499 miles	8,046.7 - 8,850 km
2	500 - 999 miles	805 - 1,608 km	12	5,500 - 5,999 miles	8,851 - 9,654 km
3	1,000 - 1,499 miles	1,609.34 - 2,412 km	13	6,000 - 6,499 miles	9,656 - 10,459 km
4	1,500 - 1,999 miles	2,414 - 3,218 km	14	6,500 - 6,999 miles	10,461 - 11,264 km
5	2,000 - 2,499 miles	3,218.69 - 4,022 km	15	7,000 - 7,499 miles	11,265 - 12,068 km
6	2,500 - 2,999 miles	4,023 - 4,826 km	16	7,500 - 7,999 miles	12,070 - 12,873 km
7	3,000 - 3,499 miles	4,828 - 5,631 km	17	8,000 - 8,499 miles	12,875 - 13,678 km
8	3,500 - 3,999 miles	5,633 - 6,436 km	18	8,500 - 8,999 miles	13,679 - 14,482 km
9	4,000 - 4,499 miles	6,437 - 7,240 km	19	9,000 - 9,499 miles	14,484 - 15,287 km
10	4,500 - 4,999 miles	7,242 - 8,045 km	20	9,500 - 9,999 miles	15,289 - 16,092 km

Table 7-10. Converting measures of distance for use with FTS distance measures.

Reported Units of Quantity Name (Abbreviation)	HS Units of Quantity Name (Abbreviation)	Multiplication Factor to Convert Reported Units to HS Units
Centimeter (CM)	Meter (MTR)	0.01
Foot (FT)	Meter (MTR)	0.3048
Linear feet (LFT)	Meter (MTR)	0.3048
Meters (MTR)	Thousand meters (THM)	0.001
Thousand meters (THM)	Meters (MTR)	1000
Thousand linear feet (MLF)	Linear meter (LNM)	304.8
Yard (YD)	Meter (MTR)	0.9144

Foreign Trade Statistics (FTS)

In 1989 the United States adopted the Harmonized Commodity Description and Coding System, commonly referred to as the Harmonized System (HS), to classify exports and imports. Given that this system collects information based on the metric standard, Table 7-10 can assist users in converting measures of distance from other data sources to metric quantities and values for use with foreign trade statistics.[198]

7.4.3 Dimensions (length, width, and depth)

Carload Waybill Sample

The Carload Waybill Sample reports dimensions using a coding system unique to this data source. For the data elements OUTSIDE WIDTH, OUTSIDE HEIGHT, and EXTREME OUTSIDE HEIGHT, dimensions are reported using a four-digit code, where the first two digits represent feet and the last two digits represent inches, rounded up to the next inch in the case of a fraction. The data element OUTSIDE LENGTH reports length using a five-digit code, with the first three digits representing feet and the last two digits represent inches. Users can consult Table 7-11

Table 7-11. Carload waybill sample measures of dimension.

Data Element	Definition	Example dimension (length, width, height, etc.)	Corresponding Carload Waybill Sample code
OUTSIDE WIDTH	Measurement of outside width of car, including attachments projecting to greatest extent.	2 feet, 7 inches (0.79 meters)	0207 0208
OUTSIDE HEIGHT	Measurement from top of rail to top of eaves at side of car.	2 feet, 7 1/2 inches (0.80 meters)	0208
EXTREME OUTSIDE HEIGHT	Measurement from top of rail to location where extreme height occurs.	2 feet, 8 inches (0.81 meters)	0208
OUTSIDE LENGTH	Distance between pulling faces of the couplers in normal position.	22 feet, 3 inches (6.78 meters) 22 feet, 3 1/2 inches (6.79 meters) 22 feet, 4 inches (6.81 meters)	02203 02204 02204

as a guide for comparing measurements such as length, width, and height from other data sources with those reported by the Carload Waybill Sample.

Vehicle Inventory and Use Survey (VIUS)[199]

The data element LENGTHTOTAL reports the total length of the vehicle or vehicle/trailer combination using a coding system unique to VIUS. Table 7-12 provides a crosswalk between the VIUS codes for vehicle length and length ranges that can be compared with vehicle lengths from other data sources; also included is the metric unit equivalent.

7.4.4 Monetary

Carload Waybill Sample

The Carload Waybill Sample classifies railroads based on their annual operating revenues as either Class I ($250 million or more), Class II ($20 million or more), or Class III ($0–$20 million). Users can apply the following formula to adjust a railroad's operating revenues to eliminate the effects of inflation:

$$\textit{Annual Operating Revenue} = \textit{Current Year's Revenues} \times (\textit{1991 Avg. Index}/\textit{Current Year's Avg. Index})$$

The average index (deflator factor) is based on the annual average Railroad Freight Price Index for all commodities.[200]

Commodity Flow Survey (CFS)

Users should note that the total value of shipments—as measured by the CFS with the data element VALUE (MILLION $) and by the U.S. Gross Domestic Product (GDP)—provides different measures of economic activity in the United States. **These measures are not directly comparable, and no bridge exists to directly correlate the two.** The value of shipments, as measured by the CFS, is the market value of goods shipped from manufacturing, mining, wholesale, and select retail and service establishments, as well as warehouses and managing offices of multiunit

Table 7-12. VIUS codes and vehicle length measurements.

LENGTHTOTAL code	Length	Metric Equivalent
01	< 16.0 feet	< 4.88 meters
02	16.0 to 19.9 feet	4.877 to 6.07 meters
03	20.0 to 27.9 feet	6.10 to 8.50 meters
04	28.0 to 35.9 feet	8.53 to 10.94 meters
05	36.0 to 40.9 feet	10.97 to 12.47 meters
06	41.0 to 44.9 feet	12.50 to 13.69 meters
07	45.0 to 49.9 feet	13.72 to 15.21 meters
08	50.0 to 54.9 feet	15.24 to 16.73 meters
09	55.0 to 59.9 feet	16.76 to 18.26 meters
10	60.0 to 64.9 feet	18.29 to 19.78 meters
11	65.0 to 69.9 feet	19.812 to 21.31 meters
12	70.0 to 74.9 feet	21.34 to 22.83 meters
13	75.0 to 79.9 feet	22.86 to 24.35 meters
14	> 80.0 feet	> 24.38 meters

establishments. Broader in scope, the GDP is the value of all goods produced and services performed by labor and capital located in the United States.[201]

Foreign Trade Statistics (FTS)

Users should note that the Census constant dollar series data used in FTS data **does not match** U.S. Bureau of Economic Analysis constant dollar series data because of the underlying coverage differences between the current dollar National Income and Product Accounts (NIPA) and Census data.[202]

Users should note that adjustments are required to use Canadian import data to produce U.S. export data to make the two comparable with one another. Canadian imports are recorded at their U.S. point of origin and do not include inland freight to the port of exit in the United States. On the other hand, U.S. exports include inland freight to the U.S. port of exit and are recorded at the U.S. seaport, airport, or border port of export inside the United States. Canada adds an estimated 4.5% of the value to each transaction to cover inland freight to compensate for this discrepancy.

Average monthly exchange rates as quoted by the Federal Reserve Board are applied to adjust the Canadian import data to U.S. dollars. A formula for converting U.S. total exports to corresponding Canadian imports is provided at the U.S. International Trade in Goods and Services (FT900) web page.

Air Carrier Statistics

Beginning in 2006, numbers in the Schedule B-1, B-11, P-11, and P-12 data tables began following the format of common public financial documents, such as reports filed with the Securities and Exchange Commission (SEC) or company financial statements.[203] When using data from Air Carrier Statistics, use Table 7-13 to reconcile reporting differences between the pre- and post-2006 financial reports:[204]

7.4.5 Time

7.4.5.1 Quarterly Reporting

To compare quarterly data across data sources, users must ensure that the definition of quarters is consistent between each data source. Table 7-14 provides a crosswalk between the definitions and coding systems for freight data sources that report quarterly data. Note that most freight data quarter definitions are different from those used by the Federal government for its fiscal year.

Table 7-13. Reconciling financial data before and after October 2006 in the Air Carrier Statistics.

Attribute	After October 2006	Before October 2006
Revenues	Report revenues as positive	Report revenues as negative
Expenses	Report expenses as negative	Report expenses as positive
Profits	Report profits as positive	Report profits as negative
Losses	Report losses as negative	Report losses as positive
Net Income	Report profits as positive Report losses as negative	Report profits as negative Report losses as positive

Table 7-14. Crosswalk of time definitions and coding systems used in freight data sources.

	Air Carrier Statistics[205]		Carload Waybill Sample[206]		County Business Patterns[207]	Service Annual Survey[208]	U.S. Government Fiscal Year[209]
	Quarter Definition	Code	Quarter Definition	Code[210]	Quarter Definition	Quarter Definition	Quarter Definition
1st Quarter	Jan. 1 – March 31	1	Jan. 1 – March 31	13	Jan. 1 – March 31*	Jan. 1 – March 31	Oct. 1 – Dec. 31
2nd Quarter	April 1 – June 30	2	April 1 – June 30	14	–	April 1 – June 30	Jan. 1 – March 31
3rd Quarter	July 1 – Sept. 30	3	July 1 – Sept. 30	15	–	July 1 – Sept. 30	April 1 – June 30
4th Quarter	Oct. 1 – Dec. 31	4	Oct. 1 – Dec. 31	16	–	Oct. 1 – Dec. 31	July 1 – Sept. 30

*County Business Patterns reports quarterly payroll estimates for the first quarter only.

7.4.5.2 Military Time

Fatality Analysis Reporting System (FARS)

FARS records the time of death using two data elements, DEATH_HR and DEATH_MN, indicating the hour of death and minute of death, respectively. These times are reported using military time, also known as the 24-hour clock format. Data users can consult Table 7-15 for a conversion between military time (24-hour time) and the 12-hour time. As an example for the FARS database, if DEATH_HR = 14 and DEATH_MN = 55, the time of death would be 1455, or 2:55 p.m.

Federal Railroad Administration (FRA) Safety Database

The FRA Safety Database uses a similar convention to identify the hour and minute of highway-rail crossing accidents using the data elements TIMEHR and TIMEMIN, respectively. Consequently, users can use Federal Railroad Administration (FRA) Safety data with data from

Table 7-15. Converting military time to 12-hour time.

12-Hour Time	Military Time (24 Hour)	12-Hour Time	Military Time (24 Hour)
Midnight	0000 or 0000 hours	Noon	1200 or 1200 hours
1:00 a.m.	0100 or 0100 hours	1:00 p.m.	1300 or 1300 hours
2:00 a.m.	0200 or 0200 hours	2:00 p.m.	1400 or 1400 hours
3:00 a.m.	0300 or 0300 hours	3:00 p.m.	1500 or 1500 hours
4:00 a.m.	0400 or 0400 hours	4:00 p.m.	1600 or 1600 hours
5:00 a.m.	0500 or 0500 hours	5:00 p.m.	1700 or 1700 hours
6:00 a.m.	0600 or 0600 hours	6:00 p.m.	1800 or 1800 hours
7:00 a.m.	0700 or 0700 hours	7:00 p.m.	1900 or 1900 hours
8:00 a.m.	0800 or 0800 hours	8:00 p.m.	2000 or 2000 hours
9:00 a.m.	0900 or 0900 hours	9:00 p.m.	2100 or 2100 hours
10:00 a.m.	1000 or 1000 hours	10:00 p.m.	2200 or 2200 hours
11:00 a.m.	1100 or 1100 hours	11:00 p.m.	2300 or 2300 hour

the Fatal Analysis Reporting System (FARS) with no conversion necessary. Consult Table 7-15 for a conversion between military time (24-hour time) and the 12-hour time.

7.4.6 Volume

Foreign Trade Statistics (FTS)

In 1989 the United States adopted the Harmonized Commodity Description and Coding System, commonly referred to as the Harmonized System (HS), to classify exports and imports. As this system collects information based on the metric standard, Table 7-16 can assist users in converting measures of volume from other data sources to metric quantities and values for use with foreign trade statistics.[211]

U.S. Waterway Data[212]

The data element NRT or "Vessel Net Tonnage" reports the volume of space available for the accommodation of passengers and the stowage of cargo. NRT can be determined using the following formula, expressed in units of 100 cubic feet for each net ton:

$$\textit{Vessel Net Tonnage} = \textit{gross tonnage} - \textit{volume of space used for accommodating vessel master, officers, crew, navigation and propelling machinery (in 100 cubic feet per ton)}$$

Users should note that **<u>NRT should not be confused with a tonnage capacity,</u>** as it simply expresses a volume capacity for passengers and cargo.

7.4.7 Weight

Foreign Trade Statistics (FTS)

In 1989 the United States adopted the Harmonized Commodity Description and Coding System, commonly referred to as the Harmonized System (HS), to classify exports and imports.

Table 7-16. Conversion measures for volume.

Reported Units of Quantity Name (Abbreviation)	HS Units of Quantity Name (Abbreviation)	Multiplication Factor to Convert Reported Units to HS Units
Cord (CD)	Cubic meter (CBM)	2.550
Cubic centimeter (CC)	Liter (LTR)	0.001
Cubic meter (CBM)	Liter (LTR)	1000.
Cubic meter (CBM)	Thousand cubic meters (TCM)	0.001
Gallon (GAL)	Liter (LTR)	3.785
Gallon (GAL)	Barrel (BBL)	0.02381
Liter (LTR)	Cubic meter (CBM)	0.001
Proof gallon (PFG)	Proof liter (PFL)	3.785
Thousand cubic meters (TCM)	Cubic meters (CBM)	1000.
Thousand cubic feet (MCF)	Thousand cubic meters (TCM)	0.02832
Wine gallon (WG)	Liter (LTR)	3.785

Table 7-17. Conversion measures for weight.

Reported Units of Quantity Name (Abbreviation)	HS Units of Quantity Name (Abbreviation)	Multiplication Factor to Convert Reported Units to HS Units
Barrel (BBL)	Thousand cubic meters (TCM)	0.000159
Barrel (BBL)	Liter (LTR)	159
Clean yield pound (CYP)	Kg dry rubber content (KDR)	0.4536
Content pound (CLB)	Content kilogram (CKG)	0.4536
Content pound (CLB)	Clean yield kilogram (CYK)	0.4536
Clean yield pound (CYP)	Clean yield kilogram (CYK)	0.4536
Cubic foot (CF)	Cubic meter (CBM)	0.02832
Cubic yard (CYD)	Cubic meter (CBM)	0.7646
Content ton (CTN)	Content metric ton (CTN)	0.9072
Content short ton (CST)	Content metric ton (CTN)	0.9072
Gram (GM)	Kilogram (KG)	0.001
Gross pound (GLB)	Gram (GM)	453.6
Gross pound (GLB)	Gross kilogram (GKG)	0.4536
Gross pound (GLB)	Kilogram (KG)	0.4536
Gross metric ton (GTN)	Metric Ton (TON)	1
Hundredweight (CWT)	Kilogram (KG)	45.36
Hundredweight (CWT)	Metric Ton (TON)	0.04536
Kilogram (KG)	Metric ton (TON)	0.001
Kilogram (KG)	Gram (GM)	1000
Long ton (LTN)	Kilogram (KG)	1016
Long ton (LTN)	Metric Ton (TON)	1.016
Metric ton (TON)	Kilogram (KG)	1000
Metric ton (TON)	Gross metric ton (GTN)	1
Metric ton (TON)	Barrel (BBL)	7.33331
Ounces (OZ)	Kilogram (KG)	0.02835
Ounces (OZ)	Grams (GM)	28.35

Given that this system collects information based on the metric standard, Table 7-17 can assist users in converting measures of weight from other data sources to metric quantities and values for use with foreign trade statistics.[213]

Vehicle Inventory and Use Survey (VIUS)

The data elements VIUS_GVW (Gross Vehicle Weight Based on Reported Average Weight) and WEIGHTAVG (Average Weight of Vehicle or Vehicle/Trailer Combination) use the coding system shown in Table 7-18 to report average weights. In addition, the data element WEIGHT_SIZE classifies the average weight of the vehicle or vehicle/trailer combination in four categories: "Light," "Medium," "Light-Heavy," and "Heavy-Heavy."[214] Use Table 7-18 to find the equivalent VIUS codes, Weight Range, and VIUS Weight Size class, or to compare other weight-related measurements when using data elements from other data sources.

Table 7-18. VIUS codes for weight range and weight size class.

VIUS CODE	Weight Range	Weight Size Class
01	Less than 6,001 pounds	**Light** (10,000 pounds or less)
02	6,001 to 8,500 pounds	
03	8,501 to 10,000 pounds	
04	10,001 to 14,000 pounds	**Medium** (10,001 to 19,500 pounds)
05	14,001 to 16,000 pounds	
06	16,001 to 19,500 pounds	
07	19,501 to 26,000 pounds	**Light-Heavy** (19,501 to 26,000 pounds)
08	26,001 to 33,000 pounds	**Heavy-Heavy** (26,001 pounds or more)
09	33,001 to 40,000 pounds	
10	40,001 to 50,000 pounds	
11	50,001 to 60,000 pounds	
12	60,001 to 80,000 pounds	
13	80,001 to 100,000 pounds	
14	100,001 to 130,000 pounds	

7.5 Commodity Classification Bridges

Keywords: bridge, crosswalk, Harmonized System (HS), NAICS (North American Industry Classification System), Standard Transportation Commodity Code (STCC), Standard Classification of Transported Goods (SCTG), Standard International Trade Classification (SITC), Hazardous Materials, time, temporal

7.5.1 Commodity Code Resolution

Users wishing to compare commodity data from one data source with another source may have difficulties because different data sources often report commodities at varying levels of resolution, even when they use the same classification system. Although a commodity from data source A may be reported at the two-digit level, data source B may report that commodity at the six-digit level. This section provides guidance on when it is appropriate to use data elements that use the same commodity classification systems with one another, and presents methods for bridging data resolution discrepancies.

Within each table, commodity codes for each data source are categorized into columns by data resolution. Data elements from different sources that are located in the same column can be used with one another with no bridging required. If data elements from multiple sources are in different columns (i.e., different data resolutions), these data elements can be bridged at the lowest data resolution (i.e., two-digit).

Harmonized System (HS) Codes

HS codes provide an increasing level of detail about a given commodity as the number of digits increases. Use the table below to identify which HS codes from the listed data sources can be used with one another for data analysis. As an example, Table 7-19 shows that Foreign Trade Statistics (FTS) commodity data can be bridged with Transborder freight commodity data at the six-digit level with no data manipulation required. If a user wants to bridge 10-digit FTS data

Table 7-19. Bridging foreign trade commodity data with Transborder freight commodity data.

	Harmonized System (HS)			
Data Source	**Two-Digit**	**Four-Digit**	**Six-Digit**	**Ten-Digit**
Foreign Trade Statistics (FTS)[215]	available*	available*	COMMODITY	COMMODITY
North American Transborder Freight Database[216]	COMMODITY	available*	COMMODITY	unavailable

* Data at this resolution can be derived by truncating the longer commodity codes.

with six-digit Transborder freight data, however, they would need to remove the last four digits from the FTS commodity code to bridge the two data sets.

Standard Classification of Transported Goods (SCTG) Codes

SCTG codes provide an increasing level of detail about a given commodity as the number of digits increases. Use Table 7-20 to identify which SCTG data elements from the listed data sources can be used with one another for data analysis.

As an example, the table shows that CFS commodity data can be bridged with FAF2 and VIUS commodity data at the two-digit level by truncating the five-digit data element COMMODITY down to two digits.

Standard International Trade Classification (SITC) Codes

SITC codes provide an increasing level of detail about a given commodity as the number of digits increases. Use Table 7-21 to identify which SITC data elements from the listed data sources can be used with one another for data analysis.

As an example, the table shows that commodities reported in the Foreign Trade Statistics (FTS) database can only be used in conjunction with U.S. Waterway Data commodities at the two-digit level.

Hazardous Material Codes

Hazardous Material codes provide an increasing level of detail about a given commodity as the number of digits increases. Use Table 7-22 to identify which Hazardous Material code data elements from the listed data sources can be used with one another for data analysis.

As an example, the table shows that MCMIS data can be bridged with FARS data at both the one- and four-digit resolutions after truncating the MCMIS four-digit codes.

Table 7-20. Using Standard Classification of Transported Goods (SCTG) data elements from various sources.

	Standard Classification of Transported Goods (SCTG)			
Data Source	**Two-Digit**	**Three-Digit**	**Four-Digit**	**Five-Digit**
Commodity Flow Survey (CFS)[217]	available*	available*	available*	COMMODITY
Freight Analysis Framework (FAF2)[218]	SCTG2	unavailable	unavailable	unavailable
Vehicle Inventory and Use Survey (VIUS)[219]	PRODUCT_PRINCPL	unavailable	unavailable	unavailable

* Data at this resolution can be derived by truncating the longer commodity codes.

Table 7-21. Using SITC data elements from various sources.

	Standard International Trade Classification (SITC)			
Data Source	**Two-Digit**	**Three-Digit**	**Four-Digit**	**Five-Digit**
Foreign Trade Statistics (FTS)[220]	available*	available*	available*	SITC SITC_CODE
U.S. Waterway Data[221]	PMS_COMM	unavailable	unavailable	unavailable

* Data at this resolution can be derived by truncating the longer commodity codes.

Miscellaneous Codes Related to Commodities

The data elements presented in Table 7-23 provide specific details for commodities that are beyond what the Harmonized System (HS), Standard Transportation Commodity Codes (STCC), Standard Classification of Transported Goods (SCTG), Standard International Trade Classification (SITC), and hazardous material codes provide. Users should note that it is inappropriate to compare these data elements directly with similar data elements from other sources without further investigation.

7.5.2 Temporal Bridges within Classification Systems

Data related to commodities may change over time as classification systems are refined and updated. This page provides methods for bridging temporal differences within the Harmonized System (HS), North American Industry Classification System (NAICS), Standard Classification of Transported Goods (SCTG), and Standard International Trade Classification (SITC) commodity classification systems. The applicable freight data sources are listed under each classification system heading.

Harmonized System (HS)

United Nations HS Conversion Tables

The HS is regularly updated by the World Customs Organization (WCO) to accommodate the emergence of new and disappearance of previously existing products, with major revisions occurring in 1996, 2002, 2007, and 2012. The United Nations (UN) Comtrade database provides concordance tables between current HS codes and previous versions, which are available online at http://unstats.un.org.[222] Figure 7-2 shows which concordance tables are available for each HS version-pair.

The concordance tables, which are available in separate Microsoft Excel files, provide direct conversions for newer codes with codes for earlier versions. In addition to showing the corresponding

Table 7-22. Using Hazardous Material Code data elements from various sources.

	Hazardous Material Codes	
Data Source	**One-Digit**	**Four-Digit**
Fatality Analysis Reporting System (FARS)[223]	HAZ_CNO PHAZ_CNO	HAZ_ID PHAZ_ID
Motor Carrier Management Information System (MCMIS)[224]	available*	HAZMAT MATERIAL ID

* Data at this resolution can be derived by truncating the longer commodity codes.

Table 7-23. Detailed data elements that require further investigation before making comparisons.

Data Source	Data Element	Definition
Carload Waybill Sample[225]	UNIQUE SERIAL NUMBER	To allow for unique identification of waybills, the AAR/Railinc assigns a six-digit number to all waybills processed. Hardcopy waybills are assigned serial numbers in the 100,000 to 199,999 range. MRI waybills are assigned serial numbers in the 200,000 to 999,999 range and the 000,000 to 099,999 range.
	WAYBILL NUMBER	The waybill number is the number an originating railroad document assigns to each waybill.
Center for Transportation Analysis Intermodal Terminals Database[226]	CARGO	A three-digit code for the type of cargo or commodity group involved in the intermodal connection
Foreign Trade Statistics (FTS)[227]	USDA	One-digit agriculture or non-agriculture product code
Motor Carrier Management Information System (MCMIS)[228]	CARGO	Description of cargo hauled by this carrier. A maximum of three cargo types are printed.
	HAZARDOUS MATERIALS CARRIED/SHIPPED	Identifies the type of hazardous material transported or shipped by the entity and whether bulk (B), non-bulk (N), or all (A).Note: The conversion of the Hazardous Materials Data elements of the new Census File to the old is as follows: Bulk (B) = Tank (T), Non-Bulk (N) = Package (P), and All (A) = Both (B).
	HAZMAT S	Type of hazardous material shipped by interstate and intrastate shippers. Coded same as HAZMAT C. Up to three hazardous materials may be printed. "B" indicates that the cargo is shipped in Bulk quantities. "N" indicates that the cargo is shipped in Non-Bulk. "A" indicates cargo is shipped both in Bulk and Non-Bulk quantities.
	HAZMAT C	Type of hazardous material carried by interstate and intrastate motor carriers. Up to three hazardous materials may be printed. "B" indicates that the cargo is carried in Bulk quantities. "N" indicates that the cargo is carried in Non-Bulk quantities. "A" indicates cargo is carried both in Bulk and Non-Bulk quantities.
National Agricultural Statistics Service (NASS)[229]	SECTOR	Five high level, broad categories useful to narrow down choices (CROPS, ANIMALS & PRODUCTS, ECONOMICS, DEMOGRAPHICS, and ENVIRONMENTAL).
	GROUP	Subsets within sector (e.g., under sector = CROPS, the groups are FIELD CROPS, FRUIT & TREE NUTS, HORTICULTURE, and VEGETABLES).
	COMMODITY	The primary subject of interest (e.g., CORN, CATTLE, LABOR, TRACTORS, OPERATORS).
U.S. Waterway Data[230]	CONTAINER	Container Indicator
	PRINC_COMM	Principal Commodity List

AAR = Association of American Railroads

from / to	HS 2007	HS 2002	HS 1996	HS 1992
HS 2012		-	-	-
HS 2007	-			
HS 2002	-	-		
HS 1996	-	-	-	
HS 1992	-	-	-	-

Figure 7-2. Concordance tables, HS 2007–2012.

HS codes between given years, the tables also indicate the relationship between the two HS versions that informed the method by which the conversions were performed (see Column D in Figure 7-3).

The four types of relationships are as follows:

- For a 1:1 relationship, the HS subheading is correlated with one and only one subheading in the previous HS.
- For a 1:n relationship, the HS subheading is the result of merging several subheadings in the previous classification.
- For an n:1 relationship, the HS subheading is a result of a split of one subheading in the previous classification into several subheadings.
- For an n:n relationship, the subheading is the result of a split and merge of several subheadings in the previous classification.

A more detailed discussion on the methodology used to create the concordance tables, along with the potential shortcomings of these conversions, is available from Comtrade in an explanatory document, Correlation and Conversion Tables Used in UN Comtrade.[231]

Concording U.S. Harmonized System (HS) Categories Over Time

In Concording U.S. Harmonized System Categories Over Time, Pierce and Schott (2010) developed an algorithm to track changes in product codes to construct a comprehensive concordance of HS codes over time.[232] Concordance files for HS codes from 1989–2009 are provided in an appendix that accompanies the paper and is available online.[233]

With sufficient knowledge of data analysis and statistical software, data users can use the algorithm code to customize or extend it to incorporate future revisions of HS categories. The state code used to build the concordance also is provided in the appendix to the paper by Pierce and Schott, and the data used in the algorithm are available within a .ZIP file located at Schott's International Economics Resource Page, Trade Data, and Concordances.[234]

	A	B	C	D
1			Between	
2	HS 2007	ex.	HS 2002	Relationship
3	010110		010110	1:1
4	020890		020820	1:n
5	030194	ex	030199	n:1
6	282739		282733	n:n

Figure 7-3. Conversion detail from concordance table.

from / to	SITC 4	SITC 3	SITC 2	SITC 1
SITC 4	-	-	-	-
SITC 3	-	-		
SITC 2	-	-	-	
SITC 1	-	-	-	-

Figure 7-4. Conversion and correlation tables, SITC 4–SITC 1.

Standard International Trade Classification (SITC)

The SITC system was introduced in 1950 by the United Nations (UN).[235] The UN Comtrade database provides concordance tables between current SITC codes and previous versions, which are available online at unstats.un.org.[236] Figure 7-4 shows which concordance tables are available for each SITC version-pair.

The conversion tables, which are available in separate Microsoft Excel files, provide direct conversions to newer codes from codes used with earlier versions. In addition to showing the corresponding SITC codes for the given years, the tables indicate the relationship between the two SITC versions that informed the method by which the conversions were created (see the "Relationship" column in Figure 7-5).

As with the HS, four types of relationships are possible in the SITC, as follows:

- For a 1 to 1 (1:1) relationship, the SITC subheading is correlated with one and only one subheading in the previous classification.
- For an n to 1 (n:1) relationship, the SITC subheading is a result of a split of one subheading in the previous classification into several subheadings.
- For a 1 to n (1:n) relationship, the SITC subheading is the result of merging several subheadings in the previous classification.
- For the n to n (n:n) relationship, the subheading is the result of a split and merge of several subheadings in the previous classification.

A more detailed discussion of the methodology used to create the conversion tables, along with the potential shortcomings of these conversions, is provided in the document Correlation and Conversion Tables used in UN Comtrade.[237]

Correlation table between SITC, Rev.3 and SITC, Rev.2

Correlation/ Correspondence table between the basic headings, sub-groups and groups the Standard International Trade Classification (SITC), Revision 3 and Revision 2

SITC, Rev. 3	SITC, Rev. 2	Relationship
00111	00111	1 to 1
00119	00119	1 to 1
00121	00121	1 to 1
00122	00122	1 to 1
00131	0013	n to 1
00139	0013	n to 1

Figure 7-5. Screenshot from correlation table between SITC revision 3 and SITC revision 2.

7.5.3 Bridges across Classification Systems

The discussions below provide methods for reconciling data elements across data sources that use different commodity classification systems.

Harmonized System (HS) to Standard Classification of Transported Goods (SCTG)

The SCTG has not been updated since 1996, when it was first introduced as a replacement to the Standard Classification of Goods system.[238] The Freight Analysis Framework (FAF) reports annual tonnage and dollar valued freight flows using the same 43 two-digit SCTG classes used by the 2007 U.S. Commodity Flow Survey (CFS). Commodities reported using the 10-digit Harmonized Tariff Schedule (Schedule B for exports) must be translated to SCTG using a crosswalk developed for the purpose: Users can consult the crosswalk provided in Appendix D of *The Freight Analysis Framework Version 3 (FAF3): A Description of the FAF3 Regional Database and How It Is Constructed* (2011).[239]

Harmonized System (HS) to NAICS to Standard International Trade Classification (SITC) Bridge

The U.S. Census Bureau provides a concordance table that allows for quick bridging between NAICS and HS commodity codes.[240] Users can follow the instructions provided below to bridge these two systems for use in freight data analysis.

Step 1. Open a web browser and go to: http://censtats.census.gov/.

Step 2. Once at the Censtats Databases website, go to the **International Trade Data** subheading and click on the link to the **Concordances** (see Figure 7-6).

Step 3. Choose either the **Import Concordance** table or the **Export Concordance** table (see Figure 7-6).

Step 4. Notice that the concordance tables contain the following dropdown menus (see Figure 7-7).

- Classification system: End-Use,[241] NAICS, SITC (Standard International Trade Classification, shown in Figure 7-7), HI-TECH[242] categories, HS Codes
- Year: 2008–2014
- Options to "Browse SITC code" or "Search code/description for SITC" for specific codes based on the chosen classification system

Step 5. From the dropdown menus, select the classification system and year and press "Go," then select the commodity of interest (e.g., under "Browse SITC code" as shown in the figure) and press "Go." The table automatically generates a concordance table showing the corresponding

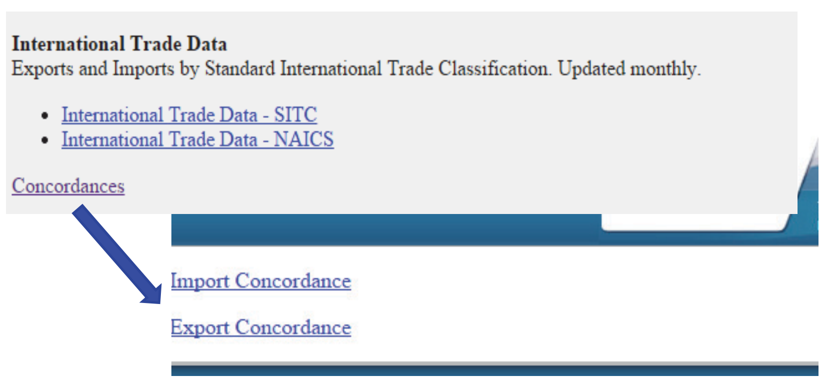


Figure 7-6. Screenshot showing links on Censtats website.

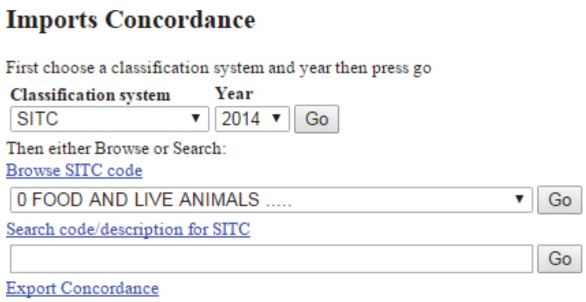


Figure 7-7. Screenshot of Imports Concordance table showing dropdown menus.

codes and descriptions used in the other classification systems. Figure 7-8 shows an example concordance table for soybean imports for each of the classification systems in 2013.

Standard Classification of Transported Goods (SCTG) to Standard Transportation Commodity Code (STCC)

A concordance bridge between SCTG and STCC classification codes is available for purchase as an online subscription through Railinc, a for-profit subsidiary of the Associate of American Railroads (AAR).[243] A discussion about the difficulties associated with bridging SCTG with other commodity classification systems is available at the Statistics Canada website.[244]

7.6 Industry Classification Bridges

Keywords: bridge, NAICS (North American Industry Classification System), concordance table, time, temporal, SIC, County Business Patterns, Commodity Flow Survey (CFS), industry, revision

Users wishing to compare industry data from one data source with another source may have difficulties because different data sources often use different industry classification systems. The temporal differences within the same industry classification system across multiple years can also make freight data analysis difficult for users. This section provides guidance on when it is appropriate to bridge temporal differences across data sources using the same classification system, as well as crosswalks for converting between classification systems.

7.6.1 Temporal Bridges within Industry Classification Systems

Data related to freight-related industries may change over time as classification systems are refined and updated. This page provides methods for bridging temporal differences in the North American Industry Classification System (NAICS).

HS	HS Description	Quantity 1	Quantity 2	END_USE	END_USE Description	SITC	SITC Description	NAICS	NAICS Description	HITECH
1201100000	SOYBEAN SEEDS OF A KIND USED FOR SOWING	KG		10140	AGRIC. FARMING-UNMANUFACTURED	22220	SOYBEANS	111110	SOYBEANS	00

Figure 7-8. Sample concordance table for soybean imports in 2013.

North American Industry Classification System (NAICS)

The North American Industry Classification System (NAICS) is reviewed every 5 years for potential revisions so that the classification system can keep pace with the changing economy. The U.S. Census Bureau provides concordance tables in spreadsheet form to bridge changes in NAICS codes over time.[245]

These tables provide detailed descriptions of the direct relationships between classification systems for each version of the NAICS from 1987 to 2012. Table 7-24 provides links to the available concordance tables from the US Census website.[246] Data users should note that not all versions of NAICS can be bridged with one another (e.g., the 2012 NAICS can only be bridged directly with the 2007 NAICS), as additional concordances are needed to bridge larger gaps in time between NAICS versions.

Figure 7-9 presents an example of how the concordance tables appear when opened. This figure shows the relationship between 2012 and 2007 NAICS codes for the farming industry. Note

Table 7-24. Links to concordance tables from U.S. Census website.*

2012 NAICS
2012 NAICS to 2007 NAICS[247*]
2007 NAICS
2007 NAICS to 2012 NAICS[248]
2007 NAICS to 2002 NAICS[249]
2002 NAICS
2002 NAICS to 2007 NAICS[250]
2002 NAICS to 1997 NAICS[251]
2002 NAICS to 1987 NAICS[252]
1997 NAICS
1997 NAICS to 2002 NAICS[253]

* Print readers are referred to the endnotes for Chapter 6 and Chapter 7, which include urls for the online documents that contain the concordances.

2012 NAICS U.S. Matched to 2007 NAICS U.S. (Full Concordance)

(Note: 2012 NAICS codes in bold indicate pieces of the 2012 industry came from more than one 2007 NAICS industry; 2007 NAICS codes in italics indicate the 2007 industry split to two or more 2012 NAICS industries.)

2012 NAICS Code	2012 NAICS Title	2007 NAICS Code	2007 NAICS Title *(and specific piece of the 2007 industry that is contained in the 2012 industry)*
111110	Soybean Farming	111110	Soybean Farming
111120	Oilseed (except Soybean) Farming	111120	Oilseed (except Soybean) Farming
111130	Dry Pea and Bean Farming	111130	Dry Pea and Bean Farming
111140	Wheat Farming	111140	Wheat Farming
111150	Corn Farming	111150	Corn Farming
111160	Rice Farming	111160	Rice Farming
111191	Oilseed and Grain Combination Farming	111191	Oilseed and Grain Combination Farming
111199	All Other Grain Farming	111199	All Other Grain Farming
111211	Potato Farming	111211	Potato Farming
111219	Other Vegetable (except Potato) and Melon Farming	111219	Other Vegetable (except Potato) and Melon Farming

Figure 7-9. Screenshot showing a sample concordance table.

Table 7-25. County Business Patterns datasets, 1998, 2003, 2008.

Year Range of Data	Data Classified
2012 to present	NAICS 2012
2008 to 2011	NAICS 2007
2003 to 2007	NAICS 2002
1998 to 2002	NAICS 1997

that the concordance table includes six-digit NAICS codes for both 2012 and 2007, as well as columns labeled "NAICS Title," describing the industry/piece of the industry (e.g., potato farming).

NAICS Changes in County Business Patterns

As Table 7-25 shows, the 1998, 2003, and 2008 County Business Patterns datasets lagged by 1 year in terms of the NAICS classification system used.

Adoption of the 2012 NAICS system changed the way industries are classified in the County Business Patterns dataset. The update added several new industries, and realigned a significant number of other industries. The major changes to the 2012 NAICS are discussed in the balance of this section, along with methods that users can employ to bridge the new and old classification systems when using County Business Patterns data.[254]

New Industries

For the 2012 NAICS, five new industries were derived from 2007 NAICS Code 22119, Other Electric Power Generation (see Table 7-26).

Realignment and Consolidation of Industries

The 2012 NAICS included a comprehensive review of the manufacturing sector, which resulted in the consolidation of more than 20% of the manufacturing industries coded in 2007. Industry realignment, consolidation, and other changes also affected other industrial sectors, including Construction, Wholesale Trade, Retail Trade, and Accommodation and Food Services. Table 7-27 shows how several 2007 NAICS electronics store categories were combined into a single 2012 NAICS industry.

A more detailed discussion about the classification changes and realignment of industries introduced in the 2012 County Business Patterns can be found at the Economic Census, Industry Classification Updates web page.[255]

Table 7-26. Multiple 2012 NAICS industries derived from one 2007 NAICS code.

2007 NAICS	2012 NAICS	NAICS Industry Title
22119		*Other Electric Power Generation*
	221114	Solar Electric Power Generation
	221115	Wind Electric Power Generation
	221116	Geothermal Electric Power Generation
	221117	Biomass Electric Power Generation
	221118	Other Electric Power Generation

Table 7-27. Consolidation reflected in changes from the 2007 NAICS to 2012 NAICS.

2007 NAICS	2012 NAICS	NAICS Industry Title
	443142	Electronics Stores
443112		*Radio, Television, and Other Electronics Stores*
443120		*Computer and Software Stores*
443130		*Camera and Photographic Supplies Stores*
451220		*Prerecorded Tape, Compact Disc, and Record Stores*

Bridging 2012 NAICS and 2007 NAICS

To bridge 2012 NAICS industries with 2007 NAICS industries when using County Business Pattern data, users should consult the **2012 to 2007 NAICS Concordance File,** which is provided on the NAICS website.[256]

7.6.2 Bridging across Industry Classification Systems

Different freight data sources employ different classification systems to identify industries. It is often unclear, however, whether disparate systems can be used with one another for freight data analysis. This page helps fill this gap by providing conditions when it is appropriate to apply crosswalks and bridges for linking different industry classification systems with one another. Particular attention is paid to data sources that have undergone changes related to the North American Industry Classification System (NAICS) and the Standard Industrial Classification (SIC) system.

County Business Patterns

Since 1998, County Business Patterns industry data has been tabulated based on the North American Industry Classification System (NAICS), whereas prior releases were tabulated according to the Standard Industrial Classification (SIC). Table 7-28 provides the industry classification system used for specific versions of County Business Patterns data.

Users can consult the *Bridge between NAICS and SIC (1997)*, which was published as part of the 1997 Economic Census, to bridge 1998–2002 County Business Pattern data with prior releases that used the SIC. The sample table in Figure 7-10 shows how this resource bridges NAICS and SIC codes for elements of the construction industry. NAICS codes appear in bold

Table 7-28. Industry classification system for County Business Patterns data.

Year Range of Data	Data Classified
2012 to present	NAICS 2012
2008 to 2011	NAICS 2007
2003 to 2007	NAICS 2002
1998 to 2002	NAICS 1997
1988 to 1997	SIC 1987
1974 to 1987	SIC 1972

23	**CONSTRUCTION**
233110	**Land subdivision & land development**
655200	Land subdividers & developers,except cemeteries
233210	**Single-family housing construction**
152100	General contractors—single-family houses
153110	Operative builders (pt)
874121	Management services (pt)
233220	**Multifamily housing construction**
152220	General contr—residential bldgs, other than single-family (pt)
153120	Operative builders (pt)
874122	Management services (pt)
233310	**Mfg & industrial building construction**
153130	Operative builders (pt)
154120	General contractors—industrial buildings & warehouses (pt)
874123	Management services (pt)

Figure 7-10. Bridging NAICS and SIC data.

type, with the corresponding SIC codes appearing in regular type just below. Tables from this resource also include the number of establishments, sales/receipts/revenue/shipments, annual payroll, and paid employees for both NAICS and SIC industries.[257]

To compare 2012 County Business Pattern industries with SIC industries from prior releases, users can consult the NAICS to SIC Crosswalk provided by the NAICS Association.[258] The crosswalk directly compares industries between the two classification systems. Users should note that for Puerto Rico County Business Patterns data, the change from SIC to NAICS occurred in 2003.[259]

Commodity Flow Survey (CFS)

The CFS has been updated four times since it was introduced in 1993. In addition to improvements to the design of the survey, sample size, survey methodology, and modes of transport, these updates have involved changes in industry classification related to the switch from the Standard Industrial Classification (SIC) system to NAICS beginning after 1997. Table 7-29 compares industry coverage between different versions of the CFS. A full description of changes between the various CFS versions is provided at the Bureau of Transportation Statistics CFS website.[260]

NAICS to SIC Crosswalk

To use NAICS industries from the 2002, 2007, and 2012 CFS with those from the 1993 and 1997 CFS for freight data analysis, users can consult the *NAICS to SIC Crosswalk* provided by the NAICS Association.[261] These tables provide direct comparisons of industries between the two classification systems.

7.7 Mode of Transport Bridges

Keywords: mode, transport, air, rail, pipeline, truck, waterway, vessel, vehicle, multimodal, intermodal, unknown

Definitions of freight transport modes tend to be consistent among freight data sources; however, taxonomic and temporal differences still exist in their naming and scope. For example, within the truck mode, trucks may be classified as "commercial trucks" or "large trucks" to differentiate them from "passenger pickup trucks." Similarly, vessels may be referred to as "carriers," "ships," or "water mode of transport"; and rail mode may be referred to as "railroad" or "train."

Table 7-29. Industry coverage between versions of the CFS.

1993 CFS and 1997 CFS	2002 CFS	2007 CFS	2012 CFS
Establishments classified based on the 1987 Standard Industrial Classification System (SIC)	Establishments classified based on 1997 North American Industry Classification System (NAICS)	Establishments classified based on 2002 North American Industry Classification System (NAICS)	Establishments classified based on 2007 North American Industry Classification System (NAICS)
Publishers in Manufacturing Sector	Not covered [a]	Publishers in Information Sector [a]	Publishers in Information Sector [a]
Logging in Manufacturing Sector	Not in scope. Classified in Agriculture (NAICS 113)	Not in scope. Classified in Agriculture (NAICS 113)	Not in scope. Classified in Agriculture (NAICS 113)
Other Manufacturing (excluding Printing Trade Services [SIC 279])	Other Manufacturing (excluding Prepress Services [NAICS 323122])	Other Manufacturing (excluding Prepress Services [NAICS 323122])	Other Manufacturing (excluding Prepress Services [NAICS 323122])
Mining (except mining services [SICs 108, 124, 138, 148] and oil and gas extraction [SICs 131 and 132])	Mining (except support activities [NAICS 213] and oil and gas extraction [NAICS 211])	Mining (except support activities [NAICS 213] and oil and gas extraction [NAICS 211])	Mining (except support activities [NAICS 213] and oil and gas extraction [NAICS 211])
Wholesale (merchants and manufacturers' sales branches and government-owned liquor stores)	Wholesale (merchants' and manufacturers' sales branches and government liquor wholesales)	Wholesale (merchants' and manufacturers' sales branches and government liquor wholesales)	Wholesale (merchants' and manufacturers' sales branches and government-owned liquor wholesales)
Retail - catalog and mail order houses	Retail (electronic shopping and mail order houses)	Retail (electronic shopping and mail order houses, fuel dealers)	Retail (electronic shopping and mail order houses, fuel dealers)
Auxiliaries (e.g., warehouses)	Auxiliaries (e.g., warehouses)	Auxiliaries (e.g., warehouses) [b]	Auxiliaries (e.g., warehouses) [b]
Generalized and Specialized Freight Trucking			General Freight Trucking (NAICS 4841)[c] and Specialized Freight Trucking (NAICS 4842)[c]

[a] Under NAICS, publishers were reclassified from Manufacturing (SIC 2711, 2721, 2731, 2741, and part of 2771) to Information (NAICS 5111 and 51223) and were excluded in the 2002 CFS. However, for the 2007 CFS, publishers were restored as an in-scope industry.

[b] Although they are included in all surveys, the procedures for identifying in-scope auxiliary establishments have changed over the years. For the 1997 CFS, a managing office was considered in scope only if it had sales or end-of-year inventories in the 1992 Census. Research conducted prior to the 2002 CFS showed that not all managing offices with shipping activity in the 1997 CFS indicated sales or inventories in the 1997 Economic Census. Consequently the 1997 Economic Census results were not used to determine scope for managing offices in the 2002 CFS. For 2002, an auxiliary was included if it supported an in-scope or retail company. For the 2007 CFS, an advance survey of approximately 40,000 auxiliary establishments was conducted in 2006 to identify auxiliary establishments with shipping activity. Those that indicated that shipping was performed (as well as non-respondents) were included in the CFS sample universe.

[c] Includes only captive warehouses that provide storage and shipping support to a single company. Warehouses offering their services to the general public and other businesses are excluded. For tabulation and publication purposes, NAICS 484 is grouped with NAICS 4931.

To illustrate these differences, Table 7-30 presents a summary of data element values related to mode of transport as reported in the most recent versions of the respective databases. For each mode of transport, the various names used in the database are listed.

7.7.1 Taxonomic Mode of Transport Bridges

Different databases use different names for the same transport mode. Sometimes information on subcategories of transport modes is provided; the information provided may also vary among

Table 7-30. **Summary of freight mode of transport values in the various freight data sources.**

	Highway	Rail	Water	Air	Pipeline	Intermodal [a]	Multimodal [a]	Other/ Unknown
Air Carrier Statistics				Air [b]				
Border Crossing/Entry Data	Truck	Train				Containers		
Carload Waybill Sample			• Ex-lake • Lake cargo • Intercoastal • Coastwise • Inland waterways			Intermodal trailer-on-flat-car (TOFC) and container-on-flat-car (COFC)		Unknown
Commodity Flow Survey (CFS)	• Truck • Private truck • For-hire truck	Railroad	• Water • Inland water • Deep sea • Great lakes • Multiple waterways	Air (includes truck to/from airport)	Pipeline		• Multiple modes • Parcel, USPS, or courier	Other and unknown modes
Freight Analysis Framework (FAF3)	Truck	Rail	Water	Air	Pipeline		Multiple modes and mail	• Other and unknown • No domestic mode
Fatal Analysis Reporting System (FARS)	• Medium truck • Heavy truck • Single unit truck • Combination truck			EMS air				• EMS unknown mode • Transported unknown source • Other • Not transported • Not reported

(continued on next page)

Table 7-30. (Continued).

	Highway	Rail	Water	Air	Pipeline	Intermodal [a]	Multimodal [a]	Other/ Unknown
Foreign Trade Statistics (FTS)			Vessel	Air				All methods
North American Transborder (Transborder)	Truck	Rail	Vessel	Air	Pipeline		• Mail • Free trade zones	Other
U.S. Waterway Data			• Self-propelled [c] • Non-self-propelled [d]					
Vehicle Inventory and Use Survey (VIUS)	• Straight truck • Truck-tractor							
Vehicle Travel Information System (VTRIS)	• Single unit trucks • Single trailer trucks • Multi-trailer trucks							

[a] For purposes of Table 7-30, the category "multimodal" means freight movement by multiple modes (which may sometimes include transit), and the category "intermodal" means movement of a container, as defined by MARAD.

[b] For a complete list of aircraft types, see the "aircraft type" field in the Air Carrier Statistics database.

[c] Self-propelled vessel types include dry bulk carrier, container ship, general cargo carrier, specialized carrier, tanker, push boat, and tugboat.

[d] Non-self-propelled vessel types include dry covered barge, dry open barge, deck barge lash /Seabee barge, other dry barge, single hull tank barge, double hull tank barge, and other tank barge.

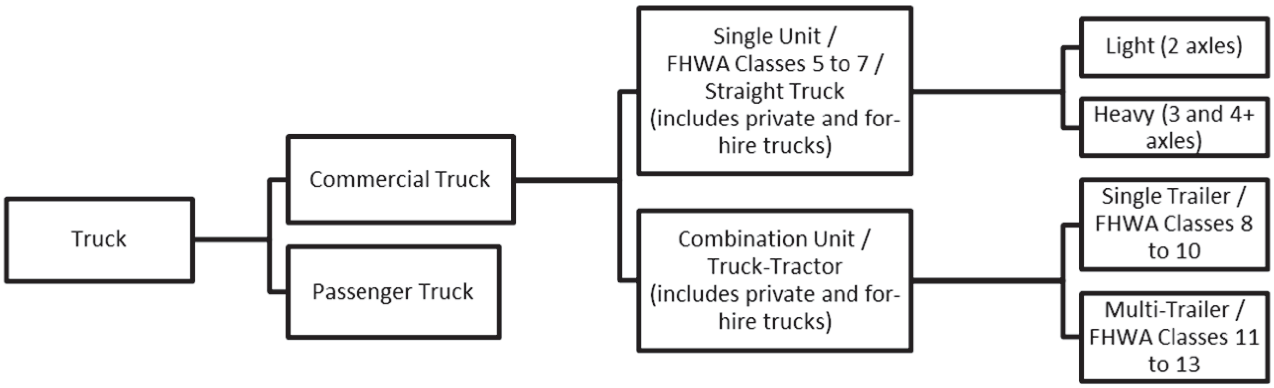


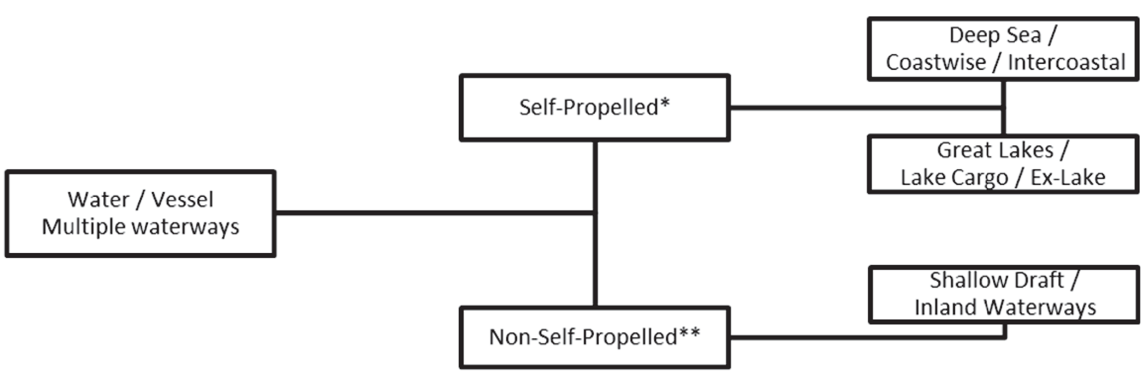
Figure 7-11. Relationships between highway modes of transport as provided in the various data sources.

databases. The following sections provide further discussions on the data values that exist for the freight mode of transport data elements. Figures 7-11 and 7-12 show the inherent relationships between data value definitions for truck and vessel transport modes, respectively, as defined in the databases. These figures serve as a guide to identifying and bridging the different data values as used in their respective databases.

Data users are advised to note that these values are restricted to freight modes of transport. Individual databases may contain other modes of transport such as buses, privately owned vehicles, and pedestrians. In addition, only the most recent versions of the databases are used in the following table. Users should consult the discussion of temporal differences in Chapter 6 for changes in individual databases over time.

Air Carrier Statistics

The Air Carrier Statistics database provides information for a single mode of transport: air. The data element AIRCRAFT TYPE, available only in the T-100 Domestic Segment, T-100 International Segment, and T-100 Segment data tables, contains a list of more than 250 aircraft models.



* Self-propelled vessel types include dry bulk carrier, container ship, general cargo carrier, specialized carrier, tanker, push boat, and tugboat.

** Non-self-propelled vessel types include dry covered barge, dry open barge, deck barge lash /Seabee barge, other dry barge, single hull tank barge, double hull tank barge, and other tank barge.

Figure 7-12. Relationships between "water" mode of transport definitions as provided in the various data sources.

Carload Waybill Sample

The data element TYPE OF MOVE VIA WATER is a classification of water movement in the Carload Waybill Sample. Users should take note of the following water movement definitions as provided in the STB Waybill Reference Guide:[262]

- 0—Not a water movement
- 1—Ex-Lake (from Great Lakes to reporting railroad)
- 2—Lake Cargo (from rail to Great Lakes)
- 3—Intercoastal (a continuous movement by U.S. rail that involves an Atlantic Ocean [or Gulf of Mexico] and Pacific Ocean movement, in either direction)
- 4—Coastwise (a continuous movement involving rail at either end of a coastwise movement between ports on the East Coast [including the Gulf of Mexico] or between ports on the West Coast)
- 5—Inland Waterways (a rail movement in combination with a barge movement on rivers and canals [waterways other than the Great Lakes] that is not considered a part of the rail movement [e.g., rail-car ferry])
- 9—Unknown

Border Crossing/Entry

The Border Crossing/Entry database contains several freight-related surface mode definitions of which users should be aware:

- Container—Any conveyance entering the United States that is used for commercial purposes, either full or empty (including containers moving in-bond for the port initiating the bonded movements)
- Rail container crossings (loaded and empty)—The number of full or empty rail containers arriving at a port
- Train crossings—The number of arriving trains at a particular port
- Truck container crossings (loaded and empty)—The number of full or empty truck containers arriving at a port
- Truck crossings—The number of arriving trucks (does not include privately owned passenger pickup trucks)

Commodity Flow Survey (CFS)

The CFS provides information on mode of transport for a variety of modes and at various levels of detail. Below are the modes included in the database. Users should pay attention to the mode definitions when using mode-related data from other data sources.[263]

- Single mode shipments—Shipments transported by only one of the following modes: private truck, for-hire truck, rail, any water mode, pipeline, or air
  - Private truck—Trucks operated by employees of the establishment or the buyer/receiver of the shipment; includes trucks providing dedicated services to the establishment
  - For-hire truck—Trucks operated by common or contract carriers made under a negotiated rate
  - Rail—Any common carrier or private railroad
  - Inland water—Vessels or barges operating primarily in navigable waters, both within and along the borders of the United States, such as:
    - Rivers (Mississippi River, Saint Lawrence Seaway, etc.)
    - Lakes (excluding Great Lakes)
    - Along the shoreline but actually in the ocean (e.g., Intracoastal Waterway along the Atlantic and Gulf coasts, Inside Passage of Alaska, etc.)
    - Canals, harbors, major bays, and inlets

- Great Lakes—Vessels or barges operating on the Great Lakes
- Deep sea—Vessels or barges operating primarily in the open waters of the ocean, outside the borders of the United States
- Multiple waterways—Shipments sent by any combination of Inland water, Great Lakes, and Deep sea; usually involving a transfer between vessels
- Pipeline—Movement (of oil, petroleum, gas, slurry, etc.) through pipelines that extend to other establishments or locations beyond the shipper's establishment; does not include aqueducts for the movement of water
- Air—Any shipment sent via air mode to its destination, including shipments carried by truck to or from an airport

- Multiple mode shipments—Shipments for which two or more of the following modes of transportation were used AND parcel delivery/courier/U.S. Parcel Post shipments:
  - Private truck or for-hire truck
  - Railroad
  - Water (inland water, Great Lakes, deep sea, and multiple waterways)
  - Pipeline
  - Air
  - Other mode
  - Parcel delivery/courier/U.S. Parcel Post—Includes ground and air shipments of packages and parcels that weigh 150 pounds or less, and were transported by a for-hire carrier. (Parcel delivery/courier/U.S. Parcel Post shipments are all considered multiple mode because this category includes all parcel shipments [whether via ground or air] tendered to a parcel or express carrier. In defining this mode, these shipments were not combined with any other reported mode because, by their nature, parcel delivery/courier/U.S. Parcel Post shipments are already multimodal. For example, if a respondent has reported a shipment's mode of transportation as parcel and air, the shipment is treated as parcel only)
  - Other multiple modes—Shipments sent by any other mode combinations not specifically listed in the tables
- Other mode(s)—Shipments for which no mode of transportation were reported, or were reported by the respondent as "other" or "unknown"; also includes shipments with a mode other than any of the listed modes (e.g., conveyor belt, animal power, and so forth).

Foreign Trade Statistics (FTS)

Mode of transport information in the FTS database is reported based on the method of transportation by which the merchandise arrives in or departs from the United States. Modes contained in the database include:[264]

- Air—Shipments leaving or arriving in the United States only by air
- Vessel—Shipments leaving or arriving in the United States only by vessel
- All Methods—Exports and general imports leaving or arriving in the United States by vessel, air, truck, rail, air mail, parcel post, and other methods of transport, including the following (which are excluded from the vessel and air statistics):
  - Mail and parcel post shipments (including those transported by vessel or air)
  - Imports and exports transported by (a) vessels moving under their own power or afloat and (b) aircraft flown into or out of the United States
  - Low-value shipments

Freight Analysis Framework (FAF3)

The FAF3 reports three main mode categories: domestic mode (DMS_MODE), foreign inbound mode (FR_INMODE), and foreign outbound mode (FR_OUTMODE). As the names imply, domestic mode signifies the mode of transport used only within the United States. Foreign inbound and outbound modes represent the modes of transport for shipments entering the United States and exiting the United States, respectively. The mode of transport definitions used in the FAF3 include:[265]

- Truck—Private and for-hire truck; does not include truck moves categorized under "multiple modes and mail" or truck moves in conjunction with domestic air cargo.
- Rail—Any common carrier or private railroad; does not include rail moves categorized under "multiple modes and mail."
- Water—Shallow draft, deep draft, Great Lakes, and intra-port shipments; does not include water moves categorized under "multiple modes and mail."
- Air (includes truck-air)—Shipments typically weighing more than 100 pounds that move by air or by a combination of truck and air in commercial or private aircraft; includes air freight and air express, but does not include shipments weighing 100 pounds or less, which are typically categorized under "multiple modes and mail." In the case of imports and exports by air, domestic moves by ground to and from the port (airport) of entry to or exit from the United States are categorized under truck mode.
- Multiple modes and mail—Shipments by multiple modes and by parcel delivery services, U.S. Postal Service, or couriers. This category is not limited to containerized or trailer-on-flatcar shipments.
- Pipeline—Crude petroleum, natural gas, and product pipelines; includes flows from offshore wells to land (which are counted as water moves by the U.S. Army Corps of Engineers); does not include pipeline moves categorized under "multiple modes and mail."
- Other (and unknown)—Shipments not classified elsewhere, such as flyaway aircraft, and shipments for which the mode cannot be determined.
- No domestic mode—Shipments that have an international mode but no domestic mode; limited to import shipments of crude petroleum transferred directly from inbound ships to a U.S. refinery at the zone of entry. This is done to ensure a proper accounting of import flows while avoiding assigning flows to the domestic transportation network that are not used.

North American Transborder Freight Database (Transborder)

Transborder uses the DISAGMOT data field to identify mode of transportation for shipments entering and exiting the United States. The specific number codes for mode of transportation are:[266]

- 1 = Vessel
- 3 = Air
- 4 = Mail (U.S. Postal Service and courier shipments; cannot be further subdivided into specific modes such as air, rail, or truck)
- 5 = Truck
- 6 = Rail
- 7 = Pipeline
- 8 = Other, a category that includes:
  - Flyaway aircraft (aircraft moving under their own power from the aircraft manufacturer to a customer and not carrying any freight)
  - Powerhouse (electricity)
  - Vessels (moving under their own power)
  - Pedestrians carrying freight
  - Unknown and miscellaneous "other"
- 9 = Foreign trade zones

Before April 1995, the actual modes of transport for imports into foreign trade zones (FTZs) were unknown and were therefore categorized under DIGAMOT 8 (other). Beginning in April 1995, as the result of inquiries from users, DIGAMOT 9 (foreign trade zones) was added as a mode of transport. Although FTZs are treated as a mode of transport, the actual mode for a specific shipment into or out of the FTZ remains unknown because U.S. Customs does not collect this information.

Vehicle Inventory and Use Survey (VIUS)

VIUS contains multiple data element fields describing highway transport modes.

- BODYTYPE—This data element distinguishes between truck tractors and non-truck tractors.
- TRUCK_SORTER—This data element distinguishes between small trucks (pickups, minivans, other light vans, and sport utilities) and large trucks.
- AXLE_CONFIG—This data element provides the best option to determine truck types based on the number of axles on the power unit and the number of axles on any trailer(s) pulled. The following parent categories are available:
  1. Straight Trucks (not pulling a trailer) and Truck Tractors (not pulling a trailer - not in use)
  2. Straight Trucks (pulling a trailer)
  3. Truck tractors (pulling a trailer)

Data element fields that provide additional information on vehicle and trailer types are:

- TRAILER—Single trailer pulled, double trailers pulled, or triple trailers pulled
- TRAILERTYPE and TRUCKTYPE—Tractor or other truck
- WEIGHT_SIZE—Average weight of vehicle or vehicle/trailer combination (light, medium, light-heavy, heavy-heavy)
- VIUS_GVW—Gross vehicle weight based on reported average weight

Vehicle Travel Information System (VTRIS)

The VTRIS uses the FHWA vehicle classification system for highway transport modes. The categories described in the database are:

1. Motorcycles
2. Passenger cars
3. Single unit trucks (2-axle, 4-tire)
4. Buses
5. Single unit trucks (2-axle, 6-tire)
6. Single unit trucks (3-axle)
7. Single unit trucks (4-axle or more)
8. Single trailer trucks (4-axle or less)
9. Single trailer trucks (5-axle)
10. Single trailer trucks (6-axle or more)
11. Multi-trailer trucks (5-axle or less)
12. Multi-trailer trucks (6-axle)
13. Multi-trailer trucks (7-axle or more)

7.7.2 Temporal Mode of Transport Bridges

Temporal mode of transport bridges are generally provided by the individual databases in an effort to reconcile changes in data collection efforts and reporting over time. The balance of this section provides additional information on these bridges.

Air Carrier Statistics

Numerous airline mergers and acquisitions have occurred in the history of air freight. Such changes can alter how a particular airline is labeled in the Air Carrier Statistics. These mergers and acquisitions affect the data elements related to carrier identification within the Air Carrier Statistics (e.g., AIRLINEID, UNIQUECARRIER, and UNIQUECARRIERNAME). Users are advised to consult documents containing information on airline mergers. An example is the "List of airline mergers and acquisitions" webpage on Wikipedia.[267]

Commodity Flow Survey (CFS)

The CFS has undergone changes over time in the way it describes water mode of transport. Table 7-31 provides a temporal bridge.[268] The main changes over the years are the definition of the water mode of transport values. In 1993, these were classified as "inland water and/or Great Lakes" and "deep sea water." As of 2012, the new names are "inland water" and "deep sea" modes of transport.

Fatality Analysis Reporting System (FARS)

In the FARS database, multiple data elements in the VEHICLE data file can be used to identify the highway transport mode of the commercial vehicle that was involved in a fatal crash.

- V_CONFIG—This data element describes the general configuration of this vehicle.
- BODY_TYP—This data element identifies a classification of the vehicle based on its general body configuration, size, shape, doors, and so forth.

These data elements have undergone changes throughout the years to provide additional detail information of vehicle configuration and classification. Table 7-32 summarizes changes to the attribute codes of the vehicle configuration (V_CONFIG) data element over the years as provided in the *FARS Analytical User's Guide*:[269]

Table 7-33 summarizes NHTSA's vehicle body type classifications. The data elements BODY_TYP (body type) and TOW_VEH (vehicle trailing) are used to determine vehicle categories, and their attribute codes can be found in the FARS Analytical User's Guide.[270]

The data element "Transported to Medical Facility By" has undergone changes throughout the years to provide additional detail on the method of transportation provided to move an individual to a hospital or medical facility. The FARS Analytical Reference Guide lists these changes.

Table 7-31. Bridging temporal changes in CFS mode names.

1993	1997, 2002, and 2007	2012
For-hire truck	For-hire truck	For-hire truck
Private truck	Private truck	Private truck
Rail	Rail	Rail
Air	Air	Air
Inland water and/or Great Lakes	Shallow draft vessel	Inland water
Deep sea water	Deep draft vessel	Deep sea
Pipeline	Pipeline	Pipeline
Parcel delivery, courier, or U.S. Parcel Post	Parcel delivery, courier, or U.S. Parcel Post	Parcel delivery, courier, or U.S. Parcel Post
Other mode	Other mode	Other mode
Unknown	Unknown	Unknown

Table 7-32. Temporal changes to attribute codes of V_CONFIG in FARS.

Attribute Codes				
1991–1994	1995–2000	2001–2009	2010–later	
0	0	--	--	Not applicable, not a medium/heavy truck or bus
--	--	00	--	Not applicable, not a medium/heavy truck or bus or vehicle displaying a hazardous material placard
--	--	--	00	Not applicable
1	1	01	--	Single unit truck (2 axles, 6 tires)
--	--	--	01	Single unit truck (2 axles and GVWR more than 10,000 pounds)
2	2	02	02	Single unit truck (3 or more axles)
--	3	03	--	Single unit truck (unknown number of axles, tires)
3	4	04	--	Truck/trailer(s)
--	--	--	04	Truck pulling trailer(s)
4	5	05	05	Truck-tractor (bobtail, i.e., tractor only, no trailer)
5	6	--	--	Truck-tractor/semi-trailer
--	--	06	--	Truck-tractor/semi-trailer (one trailer)
--	--	--	06	Truck-tractor/semi-trailer
--	--	07	--	Truck-tractor/doubles (two trailers)
--	--	--	07	Truck-tractor/double
--	--	08	--	Tractor/triples (three trailers)
--	--	--	08	Truck-tractor/triple
--	--	--	10	Vehicle 10,000 pounds or less placarded for hazardous materials
6	7	19	--	Medium/heavy truck, cannot classify
--	--	--	19	Truck more than 10,000 pounds, cannot classify
7	8	--	--	Bus
--	--	20	--	Bus (seats for 9–15 occupants, including driver)
--	--	--	20	Bus/large van (seats for 9–15 occupants, including driver)
--	--	21	--	Bus (seats for more than 15 people, including driver, 2001–2006)
--	--	21	--	Bus (seats for 16 or more people, including driver, 2007–2009)
--	--	--	21	Bus (seats for more than 15 occupants, including driver, 2010 and later)
--	--	70	--	Light truck (van, mini-van, panel, pickup, sport utility vehicle displaying a hazardous material placard)
--	--	80	--	Passenger car (only when displaying a hazardous material placard)
--	--	--	98	Not reported
9	--	--	99	Unknown
--	9	99	--	Unknown if light or medium/heavy truck/bus

The following list summarizes changes to the data element definitions based on the years the data was released:

- **1977–2000**
  - 0 = No
  - 1 = Yes
  - 7 = Died at the scene (1999–2000)
  - 8 = Died en route (1999–2000)
  - 9 = Unknown
- **2001–2006**
  - 0 = No
  - 1 = Yes
  - 9 = Unknown

Table 7-33. NHTSA's vehicle body type classifications.

Classification (BODY_TYP)	Data Year and Code		
	1975–1981	1982–1990	1991–Later
Passenger Cars	9-Jan	01-11, 67	01-11, 17 (since 2010)
Light Trucks & Vans[d]	43, 50-52, or (60 and tow_veh=0)	12, 40, 41, 48-51, 53-56, 58, 59, 68, 69, or (79 and tow_veh=0 or 9)	14-22, 24[a, f], 25[b,f], 28-41, 45-49, or (79 and tow_veh =0 or 9)
Large Trucks	53-59, or (60 and tow_veh=1)	70-72, 74-76, 78, or (79 and tow_veh in 1-5[h])	60-64, 66, 67[e], 71, 72, 78, or (79 and tow_veh[g] in 1-4)
Motorcycles	15-18	20-29	80-89
Buses	25-29	30-39	50-59 (55 van-based >10k lb. since 2011)
Other/Unknown Vehicles	35-42, 44, 45, 99	13, 14, 42, 52, 73, 77, 80, 81, 82, 83, 88, 89, 90, 99	12, 13, 23[f], 42, 65, 73, 90, 91, 92, 93, 94[c], 95 (since 2012), 97, 99 Also, since 2004 (79 and tow_veh[g] =5 or 6) or 98 (since 2010)
Passenger Vehicles	01-09, 43, 50-52, or (60 and tow_veh=0)	01-12, 40, 41, 48-51, 53-56, 58, 59, 67-69, or (79 and tow_veh-0 or 9)	01-11, 14-22, 24[a], 25[b], 28-41, 45-49, or (79 and tow_veh=0 or 9), or 17 (since 2010)
Utility Vehicles (a.k.a. On/Off Road)	43	12, 56, 68	14-16, 19
Pickups	50	50, 51	30-39
Vans	51	40, 41, 48, 49	20-22, 24[a, f], 25[b,f], 28, 29
Medium Trucks	53, 54, 56	70, 71, 75, 78	60-62, 64, 67[e], 71
Heavy Trucks	55, 57-59, or (60 and tow_veh=1)	72, 74, 76, or (79 and tow_veh in 1-5[h])	63, 66, 72, 78, or (79 and tow_veh[g] in 1-4)
Combination Trucks	((53-56, 60) and tow_veh=1), or 57-59	((70-72, 75, 76, 78, 79) and tow_veh in 1-5[h]) or 74	((60-64, 71, 72, 78, 79) and tow_veh[g] in 1-4) or 66
Single Unit Trucks	(53-56, 60) and tow_veh =0	(70-72, 75, 76, 78, 79) and tow_veh in (0,9)	(60-62,63,64,67,71,72,78,79) and tow_veh in (0,5,6[g], 9)

Notes:

[a] Body type code 24 (van-based school bus) was added in 1993. When solely defining School Buses, be sure to include body type code 24.

[b] Body type code 25 (van-based transit bus) was added in 1993. When solely defining Transit Buses, be sure to include body type code 25.

[c] Body type code 94 (motorized wheelchair) was added in 1997 and deleted in 1998.

[d] "Light trucks & vans" is frequently referred to as just "light trucks."

[e] Body type code 67 (medium/heavy pickup [e.g., Ford Super Duty 450/550]) was added in 2001. For the purposes of medium and heavy truck classifications, this body type will be considered a medium truck.

[f] Body type codes for van-based bus (24, 25) and van-based motor home (23) were deleted in 2003. These attributes were removed because a review of coding used by FARS analysts revealed that these body types were rarely being captured.

[g] New code was added in 2004 for Vehicle Trailing (tow_veh) - 5 (vehicle towing another motor vehicle). In 2009 the attribute was split into two to distinguish between fixed and non-fixed linkages (5 and 6). This attribute is not part of the selection criteria for the classifications "light," "large," "heavy," or "combination truck." Beginning with 2004, an unknown truck type (light/medium/heavy) that was towing another vehicle (BODY_TYP=79 and TOW_VEH=5,6) should be classified as Other/Unknown. This classification is subject to change.

[h] From 1982 to 1990, Vehicle Trailing (TOW_VEH) attribute value 5 (yes, two or more trailing units) existed in 1982 only. Including "5" in the range from 1982 to 1990 does not affect the classification.

- **2007–2009**
  - 0 = Not transported
  - 1 = Yes, EMS
  - 2 = Yes, law enforcement
  - 3 = Yes, other
  - 4 = Yes, transported by unknown source
  - 9 = Unknown

- **2010–Current**
  - 0 = Not transported
  - 1 = EMS air
  - 2 = Law enforcement
  - 3 = EMS unknown mode
  - 4 = Transported unknown source
  - 5 = EMS ground
  - 6 = Other
  - 8 = Not reported
  - 9 = Unknown

North American Transborder Freight Database (Transborder)

Temporal changes to the DISAGMOT mode of transport field in the Transborder freight database include the following:

- July 1995—U.S. foreign trade zones (FTZs) were added as a mode of transport, recognizing the increased activity of FTZs with regard to imports from Mexico and Canada. Although FTZs are treated as a mode, the actual mode of transportation for a specific shipment into or out of the FTZ is unknown because the data is not collected by U.S. Customs. Before July 1995, FTZ shipments had been incorrectly included as rail shipments.
- January 2004—Air and vessel modes of transport were added.
- January 2007—The database was consolidated from 12 tables to three tables to improve import and export reporting by U.S. state, port of entry/exit, and commodity. The consolidation did not affect reporting on mode of transport.

Vehicle Inventory and Use Survey (VIUS)

The Vehicle Inventory and Use Survey (VIUS) program documentation provides a list of changes made to the data source over time with specific descriptions of the change and why it was implemented.[271] The U.S. Census Bureau provides a "comparability" Excel spreadsheet that allows users to compare each data release to the previous data release by variable and valid response.[272] Users should consult these references when bridging temporal differences within the data source. The following bullet points describe an example from the document showing the changes made to the AXLE CONFIGURATION and BODY/TRAILER TYPE data element fields from 1997 compared to 2002:[273]

> Axle Configuration—The 2002 VIUS broke out additional axle response options and collapsed "utility" and "full" trailer. Truck tractors were allowed to indicate no trailer (or trailer axles) in the 2002 VIUS, whereas in the 1997 VIUS truck tractors were required to have a trailer and trailer axles present.
>
> *Reason for Change*—The additional axle and utility/full trailer changes were done at the request of data users. The Census Bureau attempted to correct erroneous 1997 VIUS editing by allowing truck tractors not in use to not report a trailer (and trailer axles).
>
> Body/Trailer Type—The 1997 VIUS asked respondents to classify their truck by selecting from a list of body types. If the vehicle was a truck-tractor, the respondent was asked to make their selection based on the trailer type most often pulled. The 2002 VIUS separated these, allowing single units to report both a body type and a trailer type (if applicable). Response options for both questions were modified.
>
> *Reason for Change*—Some body and trailer types are not interchangeable, so using separate questions reduced respondent error. The response option changes for both questions were based on data user input and questionnaire testing.

Endnotes for both Chapter 6 and Chapter 7 are listed in the References section.

CHAPTER 8

Conclusions and Suggested Future Steps

NCFRP Project 47 resulted in the development of an interactive web-based Freight Data Dictionary to organize data elements from multiple freight data sources, provide a method to identify differences in the data element definitions, and offer a set of homogeneous approaches for bridging gaps between the definitions. The study also identified examples from the literature of how freight data currently is utilized by agencies and researchers to perform public-sector functions.

This study also supports four key recommendations made in the 2014 National Freight Advisory Committee (NFAC) findings on current barriers to obtaining available, adequate, and useful U.S. freight data.

8.1 Addressing the Barriers Identified by NFAC

The 2014 NFAC recommendations to the U.S. DOT related to the development of a National Freight Strategic Plan identify barriers and corrective courses of action to improve current freight data. One critical barrier identified is that "certain types of data are reported differently depending on the mode of transportation" (NFAC 2014). *NCFRP Report 35* provides guidelines on bridging data gaps so that differences in data reporting do not hinder the integration and use of the databases. Chapter 7, which offers methods to bridge data gaps, includes a specific sub-section (Section 7.7) that deals with mode-of-transport bridges and the differences within related definitions.

- **NFAC Recommendation B30** states that "data collection needs to be comprehensive, coordinated among federal agencies (especially with the Department of Homeland Security (DHS) (TSA, USCG, CBP), and complete by including information from all freight infrastructure owners and freight carriers to the extent that proprietary data is protected" (NFAC 2014). The research for NCFRP Project 47 suggests that it may be both feasible and beneficial for all publicly available data sources to be compiled into a single source, using a tool such as the web-based Freight Data Dictionary. Accordingly, *NCFRP Report 35* offers guidelines on bridging the inherent differences of the various data sources. Such a compilation would serve to facilitate any future coordination among agencies by providing them with a single, common, and complete source of information derived from all the different databases.
- **NFAC Recommendation B31** supports the need to "strengthen data collection, including multimodal origin-destination freight flows, ports of entry performance, import bottlenecks and the repositioning of empty containers for exports. U.S. DOT should evaluate the benefit of purchasing third-party aggregator data to fill critical gaps" (NFAC 2014). The Freight Data Dictionary facilitates searching data elements and identifying gaps among the various sources. A third-party aggregator can utilize the Freight Data Dictionary's application program

interfaces (APIs), along with the information contained in this report, to automate the process of bridging critical data gaps.

- **NFAC Recommendation P12** states that "data collection efforts should be tailored to performance measures that are in line with specific outcomes that the U.S. DOT and Congress want to obtain with the increased emphasis on the multimodal national freight system" (NFAC 2014). This recommendation suggests that data collection efforts should focus on the nation's multimodal freight needs. Chapter 3 of *NCFRP Report 35* focuses on freight data uses and demonstrates how current, publicly available, multimodal data sources are being used to address freight planning issues. The Freight Data Dictionary also offers researchers and freight data users a medium to exchange knowledge of and experience with the adequate use of data for performance assessment and areas where upgrades can be targeted to improve data effectiveness and relevance.
- **NFAC Recommendation C8** states that the "U.S. DOT should continue to support the development of best practices toolkits for urban and rural freight transportation planning that seek to reduce freight-related congestion, air emissions, parking issues, and impacts on the health and safety of transportation professionals and the public" (NFAC 2014). The Freight Data Dictionary provides a medium for freight data users to discuss their findings and experiences with their use of freight data and related toolkits. Such discussions could include newly discovered best practices that can be incorporated into the website. Additionally, Chapter 7 of *NCFRP Report 35* serves to highlight selected best practices in the use and integration of available freight databases to serve freight needs, and Chapter 3 provides some examples of previous uses of integrated and non-integrated data sources and the challenges faced by practitioners.

8.2 Future Steps

To ensure the long-term sustainability and usefulness of a web-based Freight Data Dictionary, the research team suggests that the following steps be taken:

8.2.1 Updates to Data Sources, Dictionary Elements, and Methodologies

The purpose of the Freight Data Dictionary is to provide an avenue by which information gathered from NCFRP Project 47 can be updated as newer data sources and methods for resolving data heterogeneity become available.

The Freight Data Dictionary developed in conjunction with the research for NCFRP Project 47 is currently housed at http://freightdatadictionary.com. The Dictionary will be permanently hosted on the BTS website, where it will be updated and maintained by the National Transportation Library of the Office of the Assistant Secretary for Research and Technology. Changes made over time in data element names and definitions will need to be tracked and documented. Recommendations by practitioners for resolving data heterogeneity will also need to be vetted before inclusion in the main web pages. Updates should be documented and published to enable others to keep track of data gaps that may still exist and also provide guidance on additional areas for research.

8.2.2 Inclusion of Private and Big Data Sources

Data reliability and validity of nationally available freight data sources remain a concern. It is suggested that future updates consider the shift toward the use of relatively new and more reliable ITS-related data sources, such as GPS data and vehicle-to-infrastructure connected devices. The private sector also continues to invest in systems to improve the collection of highly disaggregate information on cargo movement, mode of transport operations, and other facets of

132 Implementing the Freight Transportation Data Architecture: Data Element Dictionary

their operations. Dissemination of this information is, however, restricted due to privacy concerns and the competitive nature of private-sector businesses.

The need for effective freight planning has led to discussions on data sharing partnerships between the private sector and the public sector. *NCFRP Report 25: Freight Data Sharing Guidebook* explored and addressed some of the barriers to freight data sharing and cited case studies demonstrating instances in which the private sector has been willing to work with the public sector to make this possible. Progress is also being made in some states and regions to cost-effectively procure more accurate truck travel data using technological applications instead of the traditional survey methods. In the case of truck data, several ITS technologies are able to collect salient truck travel attributes not previously available.

As new sources of data become available, the issue of data heterogeneity will persist. Strategies for resolving differences in data element definitions from these new sources with the traditional sources will need to be further examined.

8.2.3 Expansion of Discussion Topics

NCFRP Report 33 focuses on taxonomic, methodological, and temporal differences in the data, both within *individual* data sources and *across* different data sources. Seven main types of data elements have been examined: origin and destination, commodity, mode of transport, industry, imports and exports, safety, and units of measure. These topics serve as a starting point for data users to become familiar with the particular issues they may encounter when working with the specific data elements. Similarly, concerning reconciling, harmonizing, and creating statistical bridges to address differences in data element definitions, five main topics have been examined in this report: place name bridges, commodity classification bridges, industry classification bridges, mode-of-transport bridges, and units of measure bridges. These topics also serve as a starting point for data users to begin to reconcile differences in data elements that are commonly encountered in freight data analysis.

Although extensive, these topics do not cover all aspects of freight data. Additional issues, such as addressing heterogeneity in GIS data sources, resolving the issue of information reported at different time frequencies, determining which data source is most suitable for a particular task, and dealing with error propagation within the data sources themselves, remain to be addressed. It is suggested that future research further examine these topics and other topics yet to be identified.

8.2.4 Updates to the Literature on Freight Data Uses

The literature review identified key studies showing innovative and unique examples of how available freight data sources are utilized by agencies and the research community. Most studies were, however, found to be limited either by outdated data or the availability of disaggregated data. It is recommended that information on newly published studies be added to the Freight Data Dictionary website as the information becomes available in order to provide novice users with more recent examples of how freight data is being used to perform public-sector functions such as transportation planning, congestion management, economic development analysis, safety related studies, and mode-of-transport operations and services.

8.2.5 Need for a Centralized Freight Data Repository

A review of the literature on freight data identified additional sources of data utilized by practitioners from local and regional planning agencies, marine port and airport authorities, and industry sources. Data from these project-specific studies, though relevant, is rarely available to or accessible by others on completion of a study. There may be an opportunity to develop a central data collection repository in which locally collected or project-specific data can be stored or shared with other users in the transportation community. These project-specific data sources will complement currently available freight data sources and will provide additional opportunities to test or validate freight-related models.

8.2.6 Evaluate and Track Agency Data Needs to Meet MAP-21 Objectives

As MAP-21 performance measures and targets are being finalized, a targeted review of database profiles, crosswalks, and statistical bridges could be conducted to identify key areas of data needs for setting and assessing agency targets. A review of case studies on how currently available data sources are being used to develop MAP-21–mandated state and metropolitan planning organization (MPO) targets for supporting freight movement could also be beneficial. In addition, it is suggested that pilot methodologies for compiling and analyzing data that speaks directly to the evaluation of performance measures be developed and further examined in future research.

8.2.7 Software Updates

As in any software development cycle, there is a need to ensure that the web-based tool developed in this project is kept up-to-date with current technologies. Issues such as browser compatibility, spam blocking, and security vulnerabilities will need to be addressed through frequent system updates. It is suggested that a dedicated system administrator be assigned to keep track of system bugs, perform minor fixes, and address requests from users.

8.2.8 Promote the Use of the Web-based Freight Data Dictionary

The intended audiences for the Freight Data Dictionary are data analysts, modelers, planners, regulators, and policy analysts and organizations responsible for the use, development, and implementation of freight models and tools. Research engineers and technicians at universities and private research organizations who frequently utilize data also may benefit from the availability of a unified freight data element dictionary. Suggestions for dissemination include promotion of the use of this application in classroom settings to educate future engineers and planners on the appropriate use, limitations, and sources of freight data. A session during the TRB Annual Meeting also could be used to promote use of the application and to gather recommendations for future updates from professional societies and organizations such as state DOTs, city planners, and MPO staff.

Sources

AECOM. 2001. Freight Benefit/Cost Study—Compilation of the Literature. Freight Management and Operations, Federal Highway Administration. http://ops.fhwa.dot.gov/Freight/Freight_Analysis/cba/index.htm (accessed March 2014).

Anderson, Michael, and Gregory Harris. 2011. Development of Statewide Freight Plan for Alabama Using Integrated Freight Planning Framework. Transportation Research Circular E-C158: Commodity Flow Survey Workshop. Transportation Research Board of the National Academies, Washington, D.C.

Ang-Olson, Jeffrey, and Bill Cowart. 2014. Freight Activity and Air Quality Impacts in Selected North American Free Trade Agreement Trade Corridors. In Transportation Research Record: Journal of the Transportation Research Board, No. 1815, Transportation Research Board of the National Academies, Washington, D.C., pp. 86–95.

Ambite, José Luis, Genevieve Giuliano, Peter Gordon, Stefan Decker, Andreas Harth, Karanbir Jassar, Qisheng Pan, and LanLan Wang. 2004. Argos: An Ontology and Web Service Composition Infrastructure for Goods Movement Analysis. In Proceedings of the 2004 Annual National Conference on Digital Government Research, Digital Government Society of North America, p. 5.

ATRI. 2014. A National Corridors Analysis and Speed Tool (N-CAST). http://atri-online.org/n-cast/. Accessed March 2014.

Beagan, Daniel, and Lance Grenzeback. 2002. Freight Impacts on Ohio's Roadway System. No. FHWA/OH-2002/026.

Bhamidipati, Chiranjivi Sarma, and Michael J. Demetsky. 2009. Evaluation of Impacts of Intermodal Terminals on the Highway System. In Transportation Research Record: Journal of the Transportation Research Board, No. 2097, Vol. 1, Transportation Research Board of the National Academies, Washington, D.C., pp. 97–108.

Boriboonsomsin, Kanok, Ross Sheckler, and Matthew Barth. 2012. Generating Heavy-Duty Truck Activity Data Inputs for MOVES Based on Large-Scale Truck Telematics Data. In Transportation Research Record: Journal of the Transportation Research Board, No. 2270, Vol. 1, Transportation Research Board of the National Academies, Washington, D.C. pp. 49–58.

Buccella, Agustina, Alejandra Cechich, and Nieves Rodríguez Brisaboa. 2003. An Ontology Approach to Data Integration. Journal of Computer Science & Technology, Vol. 3, No. 2, pp. 62–68.

California Energy Commission. 2013. CalHEAT Research and Market Transformation Roadmap for Medium- and Heavy-Duty Trucks. Report prepared by the California Hybrid, Efficient and Advanced Truck Research Center. Available at www.calstart.org/projects/CalHEAT.aspx.

Cambridge Systematics. 2005. An Initial Assessment of Freight Bottlenecks on Highways. Cambridge Systematics, Incorporated.

Cambridge Systematics. 2007. Florida Statewide Freight and Goods Mobility Plan. http://www.seacip.com/FDOT/FDOT_Freight_Plan_Report.pdf (accessed March 2014).

Cambridge Systematics et al. 2011. NCFRP Report 12: Framework and Tools for Estimating Benefits of Specific Freight Network Investments. Transportation Research Board of the National Academies, Washington, D.C.

Cambridge Systematics, HDR, and Sarah Catzz. 2008. West Coast Corridor Coalition Trade and Transportation Study. West Coast Corridor Coalition.

Chase, Keith M., Patrick Anater, and Thomas Phelan. 2013. SHRP 2 Report S2-C20-RW-2: Freight Demand Modeling and Data Improvement Strategic Plan. Transportation Research Board of the National Academies, Washington, D.C.

Corbett, James J., J. Scott Hawker, and James J. Winebrake. 2010. Development of a California Geospatial Intermodal Freight Transport Model with Cargo Flow Analysis. California Air Resources Board, Research Division, California Environmental Protection Agency.

Cui, Zhan and P. O'Brien. 2014. Domain Ontology Management Environment, Vol.1:9. IEEE Comput. Soc. Accessed March 12, 2014. doi:10.1109/HICSS.2000.926977.

Dang, Jennifer N. 2007. Statistical Analysis of the Effectiveness of Electronic Stability Control (ESC) Systems-Final Report. NHTSA Technical Report DOT HS-810 794, Washington, D.C.

Duych, Ron, Chester Ford, and Hossain Sanjani. 2011. Hazardous Materials Highlights: 2007. Commodity Flow Survey. No. SR-026. U.S. Department of Transportation, Research and Innovative Technology Administration, Bureau of Transportation Statistics, Washington, D.C.

Eisele, William, David Schrank, T. Lomax and R. Schuman. 2012. Developing and Applying Truck Mobility Measures for Decision-making. In Proceedings of the 2012 National Conference on Intermodal Transportation: Problems, Practices, and Policies. Hampton, Virginia, October 2012.

Eisele, William L., David L. Schrank, Rick Schuman, and Timothy J. Lomax. 2013a. Estimating Urban Freight Congestion Costs: Methodologies, Measures, and Applications. In TRB 92nd Annual Meeting Compendium of Papers, No. 13-1344, Transportation Research Board of the National Academies, Washington, D.C.

Eisele, William L., David L. Schrank, Jason Bittner, and Gregory Larson. 2013b. Incorporating Urban-Area Truck Freight Value into the Urban Mobility Report. Transportation Research Record: Journal of the Transportation Research Board, Vol. 2378, No. 1, Transportation Research Board of the National Academies, Washington, D.C., pp. 54–64.

Eisele, William L., Tyler Fossett, David L. Schrank, Mohamadreza Farzaneh, Paul J. Meier, and Scott P. Williams. 2013c. Estimating and Incorporating CO_2 Emissions and Associated Fuel Consumption into the Urban Mobility Report. No. CFIRE 05-16. National Center for Freight and Infrastructure Research and Education, Madison, WI.

Ellis, Joanne. 2011. Analysis of Accidents and Incidents Occurring During Transport of Packaged Dangerous Goods by Sea. Safety Science, Vol. 49, No. 8–9, Elsevier, pp. 1231–1237.

Fallat, George, Keir Opie, Joshua Curley, Jakub Rowinski, Rongfang Liu. 2003. Freight Planning Support System for Northern New Jersey. Prepared by New Jersey Institute of Technology National Center for Transportation and Industrial Productivity, July 2003.

Federal Geographic Data Committee. 2008. Geographic Information Framework Data Standard. Accessed March 2014 at http://www.fgdc.gov/standards/projects/FGDC-standards-projects/framework-data-standard/framework-data-standard.

Federal Highway Administration. 2012. Prioritization of Projects to Improve Freight Movement Guidance, https://www.fhwa.dot.gov/map21/guidance/guidesec1116.cfm. Published October 19, 2012.

Federal Highway Administration. 2000. FHWA FY 2000 Performance Plan, Appendix I Performance Measures, http://www.fhwa.dot.gov/policy/apptitle.htm. Accessed March 2014.

Figliozzi, Miguel A., Robert Harrison, and John P. McCray. 2001. Estimating Texas-Mexico North American Free Trade Agreement Truck Volumes. Transportation Research Record: Journal of the Transportation Research Board, No. 1763, Vol. 1, Transportation Research Board of the National Academies, Washington, D.C., pp. 42–47.

Fortowsky, J. Keith, and Jennifer Humphreys. 2006. Estimating Traffic Changes and Pavement Impacts from Freight Truck Diversion Following Changes in Interstate Truck Weight Limits. Transportation Research Record: Journal of the Transportation Research Board, No. 1966, Vol. 1, Transportation Research Board of the National Academies, Washington, D.C., pp. 71–79.

General Services Administration. 1996. “Data Element” entry [Def 2]. Federal Standard 1037C. Retrieved February 3, 2014 from http://www.its.bldrdoc.gov/fs-1037/fs-1037c.htm.

Gillham, Olivia, Suzanne Horton, Judith Schwenk. 2013. FMCSA Safety Program Effectiveness Measurement: Intervention Model Fiscal Year 2009. Federal Motor Carrier Safety Administration.

Gordon Proctor & Associates et al. 2011. NCFRP Report 10: Performance Measures for Freight Transportation. Transportation Research Board of the National Academies, Washington, D.C.

Guo, Jessica Y., Qi Gong, and Andrew Obernesser. 2010. Assessment of Multimodal Freight Bottlenecks and Alleviation Strategies for the Upper Midwest Region. No. MVFC 05. National Center for Freight and Infrastructure Research and Education.

Hall, Randolph W., and Aviroop Mukherjee. 2008. Bounds on Effectiveness of Driver Hours-of-Service Regulations for Freight Motor Carriers. Transportation Research Part E: Logistics and Transportation Review, Vol. 44, No. 2, Elsevier pp. 298–312.

Harrison, Robert, Nathan Hutson, Dan Seedah, Jim Kruse, and Curtis Morgan. 2010. Emerging Trade Corridors and Texas Transportation Planning. No. FHWA/TX-10/0-5973-2.

Hartshorn, Sophie, and Christopher Lamm. 2012. FHWA Freight and Land Use Handbook. No. FHWA-HOP-12-006.

Holguín-Veras, José, Miguel Jaller, Ivan Sanchez-Diaz, Jeffrey Wojtowicz, Shama Campbell, Herbert Levinson, Catherine Lawson, Erica Levine Powers, and Lorant Tavasszy. 2012. NCHRP Report 739/NCFRP Report 19: Freight Trip Generation and Land Use. Transportation Research Board of the National Academies, Washington, D.C.

Holguín-Veras, José, et al. 2013. NCFRP Report 22: Freight Data Cost Elements. Transportation Research Board of the National Academies, Washington, D.C.

IBM. 1993. IBM Dictionary of Computing, 10th ed. Retrieved February 3, 2014 from http://www-03.ibm.com/ibm/history/documents/pdf/glossary.pdf.

ISO. 2014. "ISO 14825:2011 - Intelligent Transport Systems Geographic Data Files (GDF 5.0)." Accessed March 2014.

Jack Faucett Associates. 1994. Industry Studies of the Relationship between Highway Transportation and Productivity. Federal Highway Administration, U.S. DOT.

Jansuwan, Sarawut, Thitima Wonginta, and Anthony Chen. 2010. Development of a Decision Support Tool for Assessing Vulnerability of Transportation Networks. No. UT-10.13. Utah Department of Transportation, Research Division.

Jarossi, Linda, Daniel Hershberger, and John Woodrooffe. 2011. Trucks Involved in Fatal Accidents Codebook 2009, Center for National Truck and Bus Statistics, University of Michigan Transportation Research Institute.

Kersh, Erin, James P. Dobbins, and Mark D. Abkowitz. 2012. Truck Diversion Routing Using Geographic Information Systems. Transportation Research Record: Journal of the Transportation Research Board, Vol. 2288, No. 1, Transportation Research Board of the National Academies, Washington, D.C., pp. 74–82.

LandXML.org. 2000. LandXML. Accessed March 2014 at http://www.landxmlproject.org/.

Litman, Todd. 2010. Evaluating Transportation Economic Development Impacts. Victoria Transport Policy Institute, Victoria, BC, Canada.

Liu, Xiang, M. Rapik Saat, and Christopher PL Barkan. 2013b. Integrated Risk Reduction Framework to Improve Railway Hazardous Materials Transportation Safety. Journal of Hazardous Materials, Vol. 260, Elsevier, pp. 131–140.

Liu, Xiang, M. Rapik Saat, Xiao Qin, and Christopher PL Barkan. 2013a. Analysis of U.S. Freight-Train Derailment Severity Using Zero-Truncated Negative Binomial Regression and Quantile Regression. Accident Analysis & Prevention, Vol. 59, Elsevier, pp. 87–93.

McMullen, B. Starr, and Christopher M. Monsere. 2010. Freight Performance Measures: Approach Analysis. No. FHWA-OR-RD-10-18.

Merriam-Webster Online. 2014. "Semantics" entry. An Encyclopedia Britannica Company, http://www.merriam-webster.com/dictionary/semantics (accessed: March, 2014).

National Freight Advisory Committee. 2014. Recommendations to U.S. Department of Transportation for the Development of the National Freight Strategic Plan. U.S. Department of Transportation, Washington, D.C.

Nealer, Rachael, H. Scott Matthews, and Chris Hendrickson. 2012. Assessing the Energy and Greenhouse Gas Emissions Mitigation Effectiveness of Potential U.S. Modal Freight Policies. Transportation Research Part A: Policy and Practice, Vol. 46, No. 3, Elsevier, pp. 588–601.

Ostria, Sergio J. 1996. Assessing Emissions Contribution of Intercity Trucking. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1520, No. 1, Transportation Research Board of the National Academies, Washington, D.C., pp. 35–43.

Oxford Dictionaries. 2014. "Glossary" entry [Def. 1]. (n.d.) Retrieved February 3, 2014 from http://www.oxforddictionaries.com/definition/american_english/glossary?q=glossary.

Prozzi, Jolanda; Seedah, Dan; Carrion, Migdalia; Perrine, Ken; Hutson, Nathan; Bhat, Chandra; Walton, Michael C. 2011. Freight Planning Factors Impacting Texas Commodity Flows, Texas Department of Transportation Report No. 6297-1, FHWA/TX-11/0-6297-1.

Restrepo, Carlos E., Jeffrey S. Simonoff, and Rae Zimmerman. 2009. Causes, Cost Consequences, and Risk Implications of Accidents in U.S. Hazardous Liquid Pipeline Infrastructure. International Journal of Critical Infrastructure Protection, Vols. 1–2, No. 1, pp. 38–50.

Rhodes, Suzann S. et al. 2012. NCFRP Report 14: Guidebook for Understanding Urban Goods Movement. Transportation Research Board of the National Academies, Washington, D.C.

Rodrigue, Jean-Paul, Claude Comtois, and Brian Slack. 2013. The Geography of Transport Systems. Routledge.

Sage, Jeremy, Ken Casavant, Anne Goodchild, Ed McCormack, Zun Wang, B. Starr McMullen, and Daniel Holder. 2013. Development of a Freight Benefit/Cost Methodology for Project Planning. No. WA-RD 815.1. p. 24.

Schroeder, Jeremy L., Michael Demetsky, Terry Friesz, and Tao Yao. 2012. Infrastructure Management: Project A: Developing a Framework for Prioritizing Infrastructure Improvements on Critical Freight Corridors, and Project B: Developing a Market-Based Framework for Freight Infrastructure Management. No. MAUTC-2010-01, Mid-Atlantic Universities Transportation Center.

Shafran, Isaac, and Anne Strauss-Wieder. 2003. NCHRP Report 497: Financing and Improving Land Access to U.S. Intermodal Cargo Hubs Transportation Research Board of the National Academies, Washington, D.C.

Tok, Andre Y. C., Miyuan Zhao, Joseph Y. J. Chow, Stephen Ritchie, and Dmitri Arkhipov. 2011. Online Data Repository for Statewide Freight Planning and Analysis. Transportation Research Record: Journal of the Transportation Research Board, Vol. 2246, No. 1, Transportation Research Board of the National Academies, Washington, D.C., pp. 121–29.

Transportation Economics & Management Systems. 2008. Impact of High Oil Prices on Freight Transportation: Modal Shift Potential in Five Corridors. Technical Report.

Transportation Research Board. 2014. TRID Database. Accessed January 2014 at http://www.trb.org/InformationServices/AboutTRID.aspx.

Turnquist, Mark A., and Carmen Rawls. 2010. Multimodal Network Analysis and Vulnerability Assessment of US-Canadian Trade in Lake Erie Corridor. Transportation Research Record: Journal of the Transportation Research Board, Vol. 2168, No. 1, Transportation Research Board of the National Academies, Washington, D.C., pp. 9–16.

United States Environmental Protection Agency (EPA). 2014. MOVES (Motor Vehicle Emission Simulator). Accessed March 2014 at http://www.epa.gov/otaq/models/moves/.

United States Government Accountability Office. 2011. Performance Measurement and Evaluation—Definitions and Relationships. GAO-11-646SP. Government Printing Office, Washington, D.C.

Vanek, Francis M., and Edward K. Morlok. 1998. Freight Energy Use Disaggregated by Commodity: Comparisons and Discussion. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1641, No. 1, Transportation Research Board of the National Academies, Washington, D.C., pp. 3–8.

Walton, C. Michael et al. 2014. Integration of Data Sources to Optimize Freight Transportation in Texas, Center for Transportation Research, Texas Department of Transportation.

References

Endnotes to Chapter 6

1. Commodity Flow Survey Metropolitan Areas. (n.d.) U.S. Department of Transportation. Retrieved July 2014 from http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/commodity_flow_survey/2007/metropolitan_areas/index.html#dc
2. Comparison between FAF1, FAF2, and FAF3. *Freight Analysis Framework 3 User Guide.* (June 2012). Retrieved July 2014 from http://www.ops.fhwa.dot.gov/freight/freight_analysis/faf/faf3/userguide/index.htm#s14
3. Table 3 FAF Domestic Regions. Freight Analysis Framework 3 User Guide. (June 2012). Retrieved July 2014 from http://www.ops.fhwa.dot.gov/freight/freight_analysis/faf/faf3/userguide/index.htm#t3
4. Table 4 FAF Foreign Regions. Freight Analysis Framework 3 User Guide. (June 2012). Retrieved July 2014 from http://www.ops.fhwa.dot.gov/freight/freight_analysis/faf/faf3/userguide/index.htm#t4
5. Table 4-4 and 4-5. Carload Waybill Sample Reference Guide. (October 18, 2013). Retrieved July 1, 2014, from http://www.stb.dot.gov/stb/docs/Waybill/2012 STB Waybill Reference Guide - FINAL.pdf
6. Difference between T-100 Market and T-100 Segment Airline Traffic Data. (2011, March 21). Retrieved from http://www.rita.dot.gov/bts/help_with_data/commodity_flow_survey.html#naics_table.
7. Table 4-4 and 4-5. Carload Waybill Sample Reference Guide. (October 18, 2013). Retrieved July 1, 2014, from http://www.stb.dot.gov/stb/docs/Waybill/2012 STB Waybill Reference Guide - FINAL.pdf
8. Difference between T-100 Market and T-100 Segment Airline Traffic Data. (2011, March 21). Retrieved from https://ntl.custhelp.com/app/answers/detail/a_id/398/~/difference-between-t-100-market-and-t-100-segment-airline-traffic-data
9. Reference Guide for the 2010 Surface Transportation Board. (January 26, 2012). Retrieved from http://www.stb.dot.gov/stb/docs/Waybill/2010%20STB%20Waybill%20Reference%20Guide_JN.pdf
10. Middendorf, D. (January 1998). Intermodal Terminals Database: Concepts, Design, Implementation, and Maintenance. Retrieved from http://cta.ornl.gov/transnet/terminal_doc/
11. Commodity Flow Survey Methodology. Bureau of Transportation Statistics (2012). Retrieved from http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/commodity_flow_survey/html/methodology_2012html
12. Fatal Crash Data Overview. (April 2005). Retrieved from http://www-nrd.nhtsa.dot.gov/Pubs/FARS Brochure.pdf
13. Description of the Foreign Trade Statistical Program. (July 14, 2004). Retrieved from http://www.census.gov/foreign-trade/guide/sec2.html
14. About the Foreign Trade Programs and Products. Retrieved from http://www.census.gov/foreign-trade/about/
15. Freight Analysis Framework 3 User Guide. (June 2012). Retrieved from http://www.ops.fhwa.dot.gov/freight/freight_analysis/faf/faf3/userguide/#t2
16. North American Transborder Freight Data. Bureau of Transportation Statistics. Retrieved from http://Transborder.bts.gov/programs/international/Transborder/
17. Data Dictionary. Waterway Network Link Commodity Data. (September 2, 2014). Retrieved from http://www.navigationdatacenter.us/data/dictionary/ddlink.htm
18. Waterway Network Link Commodity Data. (September 2, 2014). Retrieved from http://www.navigation datacenter.us/data/datalink.htm
19. Program Documentation. Vehicle Inventory and Use Survey. (April 12, 2006). Retrieved from https://www.census.gov/svsd/www/vius/ProgramDocumentation.pdf
20. Reference Guide for the 2010 Surface Transportation Board. Carload Waybill Sample. (January 26, 2012). Retrieved from http://www.stb.dot.gov/stb/docs/Waybill/2010%20STB%20Waybill%20Reference%20Guide_JN.pdf

21. Rail Waybill Data. Freight and Trade Data. Center for Transportation Research and Education. Retrieved from http://www.ctre.iastate.edu/research/bts_wb/cd-rom/freight/rail.htm
22. Waybill. Business Dictionary. (2014). Retrieved from http://www.businessdictionary.com/definition/waybill.html
23. Washington State Rail Plan. Washington State Department of Transportation. (August 2013). Retrieved from http://www.wsdot.wa.gov/NR/rdonlyres/A8A477E1-2BB1-4399-8443-32BEBAE631BB/0/RailPlanFindings_Aug2013.pdf
24. Middendorf, D. (January 1998). Intermodal Terminals Database: Concepts, Design, Implementation, and Maintenance. Retrieved from http://cta.ornl.gov/transnet/terminal_doc/
25. Table 3-3. Middendorf, D. (January 1998). Intermodal Terminals Database: Concepts, Design, Implementation, and Maintenance. Retrieved from http://cta.ornl.gov/transnet/terminal_doc/#tbl3-3
26. Commodity Flow Survey Methodology. Bureau of Transportation Statistics (2012). Retrieved from http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/commodity_flow_survey/html/methodology_2012html
27. Analytical User's Manual 1975–2012. Fatality Analysis Reporting System (FARS). National Highway Traffic Safety Administration. (January 2014). Retrieved from http://www-nrd.nhtsa.dot.gov/Pubs/811855.pdf
28. UN and NA Numbers. Safety Emporium. (May 9, 2010). Retrieved from http://www.ilpi.com/msds/ref/unna.html
29. The Hazardous Materials Regulations CFR 49 Parts 100 to 185. Pipeline and Hazardous Materials Safety Administration. Retrieved from https://hazmatonline.phmsa.dot.gov/services/publication_documents/HowToUse0507.pdf
30. Analytical User's Manual 1975–2012. Fatality Analysis Reporting System (FARS). National Highway Traffic Safety Administration. (January 2014). Retrieved from http://www-nrd.nhtsa.dot.gov/Pubs/811855.pdf
31. Guide to Foreign Trade Statistics. U.S. Census Bureau. Retrieved from http://www.census.gov/foreign-trade/guide/sec2.html
32. Commodity Classifications. Description of the Foreign Trade Statistics Program. Retrieved from https://www.census.gov/foreign-trade/www/sec2.html#classification
33. U.S. Imports of Merchandise. Foreign Trade Statistics. U.S. Census Bureau. Retrieved from https://www.census.gov/foreign-trade/reference/products/catalog/impDVD.html
34. Table 2–FAF Commodity Codes. Freight Analysis Framework 3 User Guide. Federal Highway Administration. (June 2012) Retrieved from http://www.ops.fhwa.dot.gov/freight/freight_analysis/faf/faf3/userguide/#t2
35. Type of Commodity. Freight Analysis Framework 3 User Guide. Federal Highway Administration. (June 2012) Retrieved from http://www.ops.fhwa.dot.gov/freight/freight_analysis/faf/faf3/userguide/#s22
36. Description and Available Reports. MCMIS Catalog and Documentation. Federal Motor Carrier Safety Administration. Retrieved from http://mcmiscatalog.fmcsa.dot.gov/c_chap3.asp
37. Inspection HAZMAT Table. Inspection File Data Element Definitions. MCMIS Catalog and Documentation. Federal Motor Carrier Safety Administration. Retrieved from http://mcmiscatalog.fmcsa.dot.gov/d_inspection3.asp#HM
38. Census Master Table. Census File–Data Element Definitions. MCMIS Catalog and Documentation. Federal Motor Carrier Safety Administration. Retrieved from http://mcmiscatalog.fmcsa.dot.gov/d_census_daEleDef.asp
39. Quick Stats Tools. National Agricultural Statistics Service. United States Department of Agriculture. Retrieved from http://www.nass.usda.gov/Quick_Stats/
40. North American Transborder Freight Data. Bureau of Transportation Statistics. Retrieved from http://Transborder.bts.gov/programs/international/Transborder/
41. National Waterway Network. U.S. Waterway Data. Navigation Data Center. Retrieved from http://www.navigationdatacenter.us/data/dictionary/ddnwn.htm
42. Vessel Entrances and Clearances. U.S. Waterway Data. Navigation Data Center. Retrieved from http://www.navigationdatacenter.us/data/dictionary/ddclen.htm
43. Foreign Cargo Inbound and Outbound. U.S. Waterway Data. Navigation Data Center. Retrieved from http://www.navigationdatacenter.us/data/dictionary/ddimex.htm
44. Vessel Characteristics. U.S. Waterway Data. Navigation Data Center. Retrieved from http://www.navigationdatacenter.us/data/dictionary/ddvess.htm
45. Vehicle Inventory and Use Survey Documentation (April 12, 2006). Retrieved from https://www.census.gov/svsd/www/vius/ProgramDocumentation.pdf
46. Ibid.
47. 2002 Microdata Data Dictionary. Vehicle Inventory and Use Survey. Retrieved from https://www.census.gov/svsd/www/vius/datadictionary2002.pdf
48. Information available from http://www.cbp.gov/document/guidance/ace-development-and-deployment-schedule

140 Implementing the Freight Transportation Data Architecture: Data Element Dictionary

49. Foreign Trade Statistics, Report Descriptions, U.S. Census Bureau. Available at www.census.gov/foreign-trade/reference/products/catalog/ftdproducts.html
50. North American Transborder Freight Data. Bureau of Transportation Statistics. Retrieved from http://www.rita.dot.gov/bts/help_with_data/Transborder_surface_freight_data.html
51. Information available from http://Transborder.bts.gov/programs/international/Transborder/PDF/TransborderFreightDataProgram.pdf
52. A Description of the FAF3 Regional Database and How It Is Constructed. Freight Analysis Framework Version 3 (FAF3) User Guide. Freight Management and Operations. Federal Highway Administration. (June 16, 2011). Retrieved from http://faf.ornl.gov/fafweb/Data/FAF3ODDoc611.pdf
53. ER-755, 37 FR 19726. (Sept. 21, 1972). Available at http://www.law.cornell.edu/cfr/text/14/part-241
54. Annual Survey of Manufacturers. U.S. Census Bureau. Retrieved from http://www.census.gov/manufacturing/asm/about_the_surveys/index.html
55. How the Data are Collected (Coverage and Methodology). County Business Patterns. U.S. Census Bureau. Retrieved from https://www.census.gov/econ/cbp/methodology.htm
56. A Description of the FAF3 Regional Database and How It Is Constructed. The Freight Analysis Framework Version 3 (FAF3) User Guide, p. 6. Freight Management and Operations. Federal Highway Administration. (June 16, 2011). Retrieved from http://faf.ornl.gov/fafweb/Data/FAF3ODDoc611.pdf
57. State Data Series. U.S. Census Bureau. Retrieved from https://www.census.gov/foreign-trade/aip/elom.html
58. T-100 Domestic Segment (All Carriers). Aircraft Group. TranStats. Bureau of Transportation Statistics. Retrieved from http://www.transtats.bts.gov/
59. Ibid.
60. Ibid.
61. Analytical Data Users Manual 1975–2010. Fatality Analysis Reporting System (FARS). (September 2013). Retrieved from http://www-nrd.nhtsa.dot.gov/Pubs/811529.pdf
62. Analytical User's Manual 1975–2012. Fatality Analysis Reporting System (FARS). National Highway Traffic Safety Administration. (January 2014). Retrieved from http://www-nrd.nhtsa.dot.gov/Pubs/811855.pdf
63. Table Structure and Data Fields for Raw Data Time Series. North American Transborder Freight Data. Bureau of Transportation Statistics. Retrieved from http://Transborder.bts.gov/programs/international/Transborder/TBDR_DataFields.html
64. 2012 Commodity Flow Survey Overview and Methodology. Bureau of Transportation Statistics. Retrieved from http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/commodity_flow_survey/html/methodology_2012.html
65. Air Carrier Statistics (Form 41 Traffic)–All Carriers. TranStats. Bureau of Transportation Statistics. Retrieved from http://www.transtats.bts.gov/Tables.asp?DB_ID=111
66. Air Carrier Statistics (Form 41 Traffic)–All Carriers. TranStats. Bureau of Transportation Statistics. Retrieved from http://www.transtats.bts.gov/DatabaseInfo.asp?DB_ID=111&Link=0
67. Information retrieved from: http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/commodity_flow_survey/html/methodology_2012.html
68. Key CFS Terms. Commodity Flow Survey. Bureau of Transportation Statistics. Retrieved from http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/commodity_flow_survey/html/def_terms.html
69. Section 2–Description of the Foreign Trade Statistics Program. Foreign Trade. Retrieved from https://www.census.gov/foreign-trade/www/sec2.html
70. Freight Analysis Framework 3 User Guide. Freight Management and Operations. Federal Highway Administration. (June 2012). Retrieved from http://www.ops.fhwa.dot.gov/freight/freight_analysis/faf/faf3/userguide/
71. 2012 Commodity Flow Survey Overview and Methodology. Bureau of Transportation Statistics. Retrieved from http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/commodity_flow_survey/html/methodology_2012.html
72. Blower, D. Evaluation of the Motor Carrier Management Information System Crash File–Phase One. Federal Motor Carrier Safety Administration. (March 2003) Retrieved from http://deepblue.lib.umich.edu/bitstream/handle/2027.42/1542/97644.0001.001.pdf?sequence=2
73. Transborder Freight Data Program. Bureau of Transportation Statistics. Retrieved from http://Transborder.bts.gov/programs/international/Transborder/PDF/TransborderFreightDataProgram.pdf
74. Ibid.
75. Vehicle Inventory and Use Survey Documentation (April 12, 2006). Retrieved from https://www.census.gov/svsd/www/vius/ProgramDocumentation.pdf
76. 2012 Commodity Flow Survey Overview and Methodology. Bureau of Transportation Statistics. Retrieved from http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/commodity_flow_survey/html/methodology_2012.html

77. Ibid.
78. Analytical Data Users Manual 1975–2010. Fatality Analysis Reporting System (FARS). National Highway Traffic Safety Administration. (September 2013) Retrieved from http://www-nrd.nhtsa.dot.gov/Pubs/811529.pdf
79. Analytical User's Manual 1975–2012. Fatality Analysis Reporting System (FARS). National Highway Traffic Safety Administration. (January 2014). Retrieved from http://www-nrd.nhtsa.dot.gov/Pubs/811855.pdf
80. Table 1. Freight Analysis Framework 3 User Guide. Freight Management and Operations. (June 2012). Retrieved from http://www.ops.fhwa.dot.gov/freight/freight_analysis/faf/faf3/userguide/#t1
81. 2002 Microdata Data Dictionary. Vehicle Inventory and Use Survey. Retrieved from https://www.census.gov/svsd/www/vius/datadictionary2002.pdf
82. 2012 Commodity Flow Survey Overview and Methodology. Bureau of Transportation Statistics. Retrieved from http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/commodity_flow_survey/html/methodology_2012.html
83. Transborder Freight Data Program. Bureau of Transportation Statistics. Retrieved from http://Transborder.bts.gov/programs/international/Transborder/PDF/TransborderFreightDataProgram.pdf
84. Retrieved from https://www.railinc.com/rportal/documents/18/260749/RAR.pdf
85. FRA Guide for Preparing Accident/Incident Reports. Office of Railroad Safety. (July 1, 2011).
86. Question 27, North American Transborder Freight Data: FAQ. Bureau of Transportation Statistics. Retrieved from http://Transborder.bts.gov/programs/international/Transborder/TBDR_FAQs.html#27
87. Key CFS Terms. Commodity Flow Survey. Bureau of Transportation Statistics. Retrieved from http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/commodity_flow_survey/html/def_terms.html
88. A Description of the FAF3 Regional Database and How It Is Constructed. Freight Analysis Framework 3 User Guide. Freight Management and Operations. Federal Highway Administration. Retrieved from http://faf.ornl.gov/fafweb/Data/FAF3ODDoc611.pdf
89. Table 1. Freight Analysis Framework 3 User Guide. Freight Management and Operations. Federal Highway Administration. Retrieved from http://www.ops.fhwa.dot.gov/freight/freight_analysis/faf/faf3/userguide/index.htm#t1
90. Transborder Freight Data Program. Bureau of Transportation Statistics. Retrieved from http://Transborder.bts.gov/programs/international/Transborder/PDF/TransborderFreightDataProgram.pdf
91. Ibid.
92. Ibid.
93. Rail Waybill Data. Center for Transportation Research and Education. Retrieved from http://www.ctre.iastate.edu/research/bts_wb/cd-rom/freight/rail.htm
94. Key CFS Terms. Commodity Flow Survey. Bureau of Transportation Statistics. Retrieved from http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/commodity_flow_survey/html/def_terms.html
95. Question 27, North American Transborder Freight Data: FAQ. Bureau of Transportation Statistics. Retrieved from http://Transborder.bts.gov/programs/international/Transborder/TBDR_FAQs.html#27
96. Navigation Data Center User's Guide. U.S. Army Corps of Engineers. (June 1996). Retrieved from http://www.navigationdatacenter.us/pdf/userguid.pdf
97. Transportation Statistics–Method of Transportation. Guide to Foreign Trade Statistics. U.S. Census Bureau. Retrieved from http://www.census.gov/foreign-trade/guide/sec2.html#mot
98. Freight Analysis Framework 3 User Guide. Freight Management and Operations. Federal Highway Administration. Retrieved from http://www.ops.fhwa.dot.gov/freight/freight_analysis/faf/faf3/userguide/
99. Transportation Statistics–Method of Transportation. Guide to Foreign Trade Statistics. U.S. Census Bureau. Retrieved from http://www.census.gov/foreign-trade/guide/sec2.html#mot
100. Freight Analysis Framework 3 User Guide. Freight Management and Operations. Federal Highway Administration. Retrieved from http://www.ops.fhwa.dot.gov/freight/freight_analysis/faf/faf3/userguide/
101. North American Transborder Freight Data: Major Changes. Bureau of Transportation Statistics. Retrieved from http://Transborder.bts.gov/programs/international/Transborder/TBDR_MajorChanges.html
102. Ibid.
103. Key CFS Terms. Commodity Flow Survey. Bureau of Transportation Statistics. Retrieved from http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/commodity_flow_survey/html/def_terms.html
104. Freight Analysis Framework 3 User Guide. Freight Management and Operations. Federal Highway Administration. Retrieved from http://www.ops.fhwa.dot.gov/freight/freight_analysis/faf/faf3/userguide/
105. Key CFS Terms. Commodity Flow Survey. Bureau of Transportation Statistics. Retrieved from http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/commodity_flow_survey/html/def_terms.html
106. Freight Analysis Framework 3 User Guide. Federal Highway Administration. (June 2012). Retrieved from http://www.ops.fhwa.dot.gov/freight/freight_analysis/faf/faf3/userguide/faf3_guide.pdf
107. Middendorf, D. (January 1998). Intermodal Terminals Database: Concepts, Design, Implementation, and Maintenance. Retrieved from http://cta.ornl.gov/transnet/terminal_doc/

108. Reference Guide for the 2010 Surface Transportation Board. (January 26, 2012). Retrieved from http://www.stb.dot.gov/stb/docs/Waybill/2010%20STB%20Waybill%20Reference%20Guide_JN.pdf
109. Stagl, J. New UMLER system will enable car owners to communicate more clearly and quickly, Railinc says. Progressive Railroading. (July 2009). Retrieved from http://www.progressiverailroading.com/mechanical/article/New-UMLER-system-will-enable-car-owners-to-communicate-more-clearly-and-quickly-Railinc-says--20810
110. Reference Guide for the 2010 Surface Transportation Board. (January 26, 2012). Retrieved from http://www.stb.dot.gov/stb/docs/Waybill/2010%20STB%20Waybill%20Reference%20Guide_JN.pdf
111. Middendorf, D. (1998, January). Intermodal Terminals Database: Concepts, Design, Implementation, and Maintenance. Retrieved from http://cta.ornl.gov/transnet/terminal_doc/
112. A Description of the FAF3 Regional Database and How It Is Constructed. The Freight Analysis Framework Version 3 (FAF3). Federal Highway Administration. (June 16, 2011). Retrieved from http://faf.ornl.gov/fafweb/Data/FAF3ODDoc611.pdf
113. Vehicle Inventory and Use Survey Documentation (April 12, 2006). Retrieved from https://www.census.gov/svsd/www/vius/ProgramDocumentation.pdf
114. Reference Guide for the 2010 Surface Transportation Board. (January 26, 2012). Retrieved from http://www.stb.dot.gov/stb/docs/Waybill/2010 STB Waybill Reference Guide_JN.pdf
115. Analytical Data Users Manual 1975–2010. Fatality Analysis Reporting System (FARS). National Highway Traffic Safety Administration. (September 2013) Retrieved from http://www-nrd.nhtsa.dot.gov/Pubs/811529.pdf
116. Analytical User's Manual 1975–2012. Fatality Analysis Reporting System (FARS). National Highway Traffic Safety Administration. (January 2014). Retrieved from http://www-nrd.nhtsa.dot.gov/Pubs/811855.pdf
117. Freight Analysis Framework 3 User Guide. Freight Management and Operations. Federal Highway Administration. Retrieved from http://www.ops.fhwa.dot.gov/freight/freight_analysis/faf/faf3/userguide/
118. North American Transborder Freight Data. Bureau of Transportation Statistics. Retrieved from http://www.rita.dot.gov/bts/help_with_data/Transborder_surface_freight_data.html
119. 2012 Commodity Flow Survey Overview and Methodology. Bureau of Transportation Statistics. Retrieved from http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/commodity_flow_survey/html/methodology_2012.html
120. Analytical Users Manual 1975–2010. Fatality Analysis Reporting System (FARS). National Highway Traffic Safety Administration. (September 2013) Retrieved from http://www-nrd.nhtsa.dot.gov/Pubs/811529.pdf
121. Fatality Analysis Reporting System (FARS) Official Website. National Highway Traffic Safety Administration. Retrieved from http://www.nhtsa.gov/FARS
122. Analytical User's Manual 1975–2012. Fatality Analysis Reporting System (FARS). National Highway Traffic Safety Administration. (January 2014). Retrieved from http://www-nrd.nhtsa.dot.gov/Pubs/811855.pdf
123. Frequently Asked Questions. Federal Motor Carrier Safety Administration. Retrieved from https://csa.fmcsa.dot.gov/FAQs.aspx?faqid=1203
124. Federal Railroad Administration Official Website. Retrieved from https://www.fra.dot.gov/
125. Fatality Analysis Reporting System (FARS) Official Website. National Highway Traffic Safety Administration. Retrieved from http://www.nhtsa.gov/FARS
126. Analytical User's Manual 1975–2012. Fatality Analysis Reporting System (FARS). National Highway Traffic Safety Administration. (January 2014). Retrieved from http://www-nrd.nhtsa.dot.gov/Pubs/811855.pdf
127. Ibid.
128. Carrier Safety Measurement System (CSMS) Methodology–Version 3.0.3. CSA. (September 2014). Retrieved from https://csa.fmcsa.dot.gov//Documents/SMSMethodology.pdf
129. Driver Safety Measurement System (DSMS) Methodology–Version 3.0.1. CSA. (August 2013). Retrieved from http://csa.fmcsa.dot.gov/Documents/Driver_SMSMethodology.pdf
130. Two measurement systems were designed in the SMS: the Carrier Safety Measurement System (CSMS) and the Driver Safety Measurement Systems (DSMS).
131. Reporting Criteria Changes–1990-Current. Pipeline and Hazardous Materials Safety Administration. (March 2011). Retrieved from http://primis.phmsa.dot.gov/comm/reports/safety/docs/IncidentReportingCriteriaHistory1990-2011.pdf
132. Retrieved from http://phmsa.dot.gov/pv_obj_cache/pv_obj_id_BA1B26D970823D0EF3A2A5BE71A5E0DF4B691800/filename/DQA Report.pdf
133. More Railroad Equipment Accident/Incident Reporting Threshold can be found at the following source: Railroad Equipment Accident/Incident Reporting Threshold. Federal Railroad Administration. (January 31, 2012). Retrieved from https://www.fra.dot.gov/eLib/details/L03622
134. FRA Guide Preparing Accident Incident Reports. Federal Railroad Administration. Office of Safety. (May 2011). Retrieved from https://safetydata.fra.dot.gov/OfficeofSafety/publicsite/Newregulation.aspx?doc=FRAGuideforPreparingAccIncRe portspubMay2011.pdf

135. Conforming the Federal Railroad Administration's Accident/Incident Reporting Requirements to the Occupational Safety and Health Administration's Revised Reporting Requirements; Other Amendments. Federal Railroad Administration. *Federal Register,* Vol. 68, No. 41. (March 3, 2003). Retrieved from http://www.gpo.gov/fdsys/pkg/FR-2003-03-03/pdf/03-4633.pdf
136. Blower, D. Evaluation of the Motor Carrier Management Information System Crash File, Phase One. Federal Motor Carrier Safety Administration. (March 2003). Retrieved from http://deepblue.lib.umich.edu/bitstream/handle/2027.42/1542/97644.0001.001.pdf?sequence=2
137. Data Quality. Federal Motor Carrier Safety Administration. Retrieved from http://ai.fmcsa.dot.gov/DataQuality/DataQuality.asp?redirect=about_datatools_desc.asp
138. Crash Data Sources. Crash Statistics. Retrieved from http://ai.volpe.dot.gov/CrashProfile/datasource.asp
139. T-100 Domestic Market (U.S. Carriers). TranStats. Bureau of Transportation Statistics. Retrieved from http://www.transtats.bts.gov/Fields.asp?Table_ID=258
140. FRA Guide for Preparing Accident/Incident Reports. Office of Railroad Safety. (July 1, 2011).
141. 2002 Microdata Data Dictionary. Vehicle Inventory and Use Survey. Retrieved from https://www.census.gov/svsd/www/vius/datadictionary2002.pdf
142. FRA Guide for Preparing Accident/Incident Reports. Office of Railroad Safety. (July 1, 2011).
143. Ibid.
144. Air Carrier Financial Reports (Form 41 Financial Data). TranStats. Bureau of Transportation Statistics. Retrieved from http://www.transtats.bts.gov/Tables.asp?DB_ID=135
145. Reference Guide for the 2010 Surface Transportation Board. (January 26, 2012). Retrieved from http://www.stb.dot.gov/stb/docs/Waybill/2010%20STB%20Waybill%20Reference%20Guide_JN.pdf
146. Ibid.
147. Ibid.
148. 2002 Commodity Flow Survey. Bureau of Transportation Statistics. Retrieved from http://apps.bts.gov/publications/commodity_flow_survey/2002/hazardous_materials/html/2002_commodity_flow_survey.html
149. Transborder Freight Data Program. Bureau of Transportation Statistics. Retrieved from http://Transborder.bts.gov/programs/international/Transborder/PDF/TransborderFreightDataProgram.pdf
150. Description of the Foreign Trade Statistical Program. (July 14, 2004). Retrieved from http://www.census.gov/foreign-trade/guide/sec2.html
151. Difference between T-100 Market and T-100 Segment Airline Traffic Data. U.S. Department of Transportation. (March 21, 2011). Retrieved from https://ntl.custhelp.com/app/answers/detail/a_id/398
152. Border Crossing/Entry Data. Bureau of Transportation Statistics. Retrieved from http://Transborder.bts.gov/programs/international/Transborder/TBDR_BC/TBDR_BC_Index.html#howto
153. Air Carrier Statistics (Form 21 Traffic). TranStats. Bureau of Transportation Statistics. Retrieved from http://www.transtats.bts.gov/Tables.asp?DB_ID=110&DB_Name=Air%20Carrier%20Statistics%20%28Form%2041%20Traffic%29-%20%20U.S.%20Carriers&DB_Short_Name=Air%20Carriers
154. Fiscal Year. United States Senate. Retrieved from http://www.senate.gov/reference/glossary_term/fiscal_year.htm
155. Data Library: Aviation. TranStats. Bureau of Transportation Statistics. Retrieved from http://www.transtats.bts.gov/databases.asp?mode_id=1&mode_desc=aviation&subject_id2=0
156. Data Profile. TranStats. Bureau of Transportation Statistics. Retrieved from http://www.transtats.bts.gov/databaseinfo.asp?DB_ID=111
157. Reference Guide for the 2010 Surface Transportation Board. (January 26, 2012). Retrieved from http://www.stb.dot.gov/stb/docs/Waybill/2010%20STB%20Waybill%20Reference%20Guide_JN.pdf
158. Analytical User's Manual 1975–2012. Fatality Analysis Reporting System (FARS). National Highway Traffic Safety Administration. Retrieved from http://www.nber.org/fars/ftp.nhtsa.dot.gov/FARS/FARS-DOC/FARS%20Analytical%20Data users%20Manual%201975-2012_FINAL-2013-11-01.pdf
159. Analytical User's Manual 1975–2012. Fatality Analysis Reporting System (FARS). National Highway Traffic Safety Administration. (January 2014). Retrieved from http://www-nrd.nhtsa.dot.gov/Pubs/811855.pdf
160. Freight Analysis Framework Data Dictionary. Federal Highway Administration. Retrieved from http://www.ops.fhwa.dot.gov/freight/freight_analysis/faf/faf3/netwkdbflow/database/fafdb_datadictionary.pdf
161. Highway Performance Monitoring System Field Manual. Office of Highway Policy Information. (March 2014). Retrieved from https://www.fhwa.dot.gov/policyinformation/hpms/fieldmanual/HPMS_2014.pdf
162. Freight Analysis Framework 3 User Guide. Freight Management and Operations. Federal Highway Administration. (June 2012). Retrieved from http://www.ops.fhwa.dot.gov/freight/freight_analysis/faf/faf3/userguide/index.htm
163. Highway Performance Monitoring System Field Manual. Office of Highway Policy Information. (March 2014). Retrieved from https://www.fhwa.dot.gov/policyinformation/hpms/fieldmanual/HPMS_2014.pdf
164. Ibid.

165. Traffic Monitoring Guide. Highway Administration Traffic Monitory Guide. Federal Highway Administration. (September 2013). Retrieved from http://www.fhwa.dot.gov/policyinformation/tmguide/tmg_fhwa_pl_13_015.pdf
166. Ibid.
167. Highway Performance Monitoring System Field Manual. Office of Highway Policy Information. (March 2014). Retrieved from https://www.fhwa.dot.gov/policyinformation/hpms/fieldmanual/HPMS_2014.pdf
168. Item 2: Urban_Code. Field Manual Chapter 4: Data Requirements and Specifications. Office of Highway Policy Information. (March 2014). Retrieved from https://www.fhwa.dot.gov/policyinformation/hpms/fieldmanual/chapter4b.cfm#item2
169. *Waterborne Transportation Lines of the United States.* Volume 1–National Summaries. U.S. Army Corps of Engineers. (2012). Retrieved from http://www.navigationdatacenter.us/veslchar/pdf/wtlusvl1_12.pdf
170. Trade Definitions. Foreign Trade. U.S. Census Bureau. Retrieved from https://www.census.gov/foreign-trade/reference/definitions/
171. Heavy Vehicle Travel Information System Field Manual. Federal Highway Administration. (May 2001). Retrieved from http://www.fhwa.dot.gov/ohim/tvtw/hvtis.htm
172. 2002 Microdata Data Dictionary. Vehicle Inventory and Use Survey. Retrieved from https://www.census.gov/svsd/www/vius/datadictionary2002.pdf
173. 2012 Fatality Analysis Reporting System (FARS) and National Automotive Sampling System General Estimates System (GES) Coding and Validation Manual. National Highway Traffic Safety Administration. (November 2013). Retrieved from http://www-nrd.nhtsa.dot.gov/Pubs/811854.pdf
174. Analytical User's Manual 1975–2012. Fatality Analysis Reporting System (FARS). National Highway Traffic Safety Administration. (January 2014). Retrieved from http://www-nrd.nhtsa.dot.gov/Pubs/811855.pdf
175. Reference Guide for the 2010 Surface Transportation Board. (January 26, 2012). Retrieved from http://www.stb.dot.gov/stb/docs/Waybill/2010%20STB%20Waybill%20Reference%20Guide_JN.pdf
176. Ibid.
177. Key CFS Terms. Commodity Flow Survey. Bureau of Transportation Statistics. Retrieved from http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/commodity_flow_survey/html/def_terms.html
178. Transborder Freight Data Program. Bureau of Transportation Statistics. Retrieved from http://Transborder.bts.gov/programs/international/Transborder/PDF/TransborderFreightDataProgram.pdf
179. CropScape–Cropland Data Layer. National Agricultural Statistics Service. Retrieved from http://nassgeodata.gmu.edu/CropScape/
180. National Corridors Analysis and Speed Tool (N-CAST). American Transportation Research Institute. Retrieved from http://atri-online.org/n-cast/
181. Transearch. Economics and Country Risk. HIS. Retrieved from http://www.ihs.com/products/global-insight/industry/commerce-transport/database.aspx
182. TREDIS–Freight. TREDIS. Retrieved from http://www.tredis.com/products/tredis-freight

Endnotes to Chapter 7

183. 2012 NAICS to SIC Crosswalk. North American Industry Classification System (NAICS). (2012). Retrieved from http://www.naics.com/naicswp2014/wp-content/uploads/2014/10/NAICS-to-SIC-Crosswalk.pdf
184. U.S. Census Bureau, Geography–Geographic Terms and Concepts–Census Tract. Retrieved March 4, 2015, from https://www.census.gov/geo/reference/gtc/gtc_ct.html
185. Karp, David, A Quick Guide to Geographic Place Codes (PART III), The Policy Map, Retrieved March 4, 2015, from http://www.policymap.com/blog/2012/08/tips-on-fips-a-quick-guide-to-geographic-place-codes-part-iii/
186. American National Standards Institute (ANSI) Codes for States, the District of Columbia, Puerto Rico, and the Insular Areas of the United States. U.S. Census Bureau. Retrieved from http://www.census.gov/geo/reference/ansi_statetables.html
187. US Census Bureau, Geography–Substantial Changes to Counties and County Equivalent Entities: 1970–Present. Retrieved March 4, 2015, from https://www.census.gov/geo/reference/county-changes.html
188. Missouri Census Data Center, MABLE/Geocorr12: Geographic Correspondence Engine. Retrieved March 2015 from http://mcdc.missouri.edu/websas/geocorr12.html
189. Missouri Census Data Center, MABLE/Geocorr12: Geographic Correspondence Engine. Retrieved March 2015 from http://mcdc.missouri.edu/websas/geocorr12.html
190. Air Carrier Statistics (Form 21 Traffic). TranStats. Bureau of Transportation Statistics. Retrieved from http://www.transtats.bts.gov/Tables.asp?DB_ID=110&DB_Name=Air%20Carrier%20Statistics%20%28Form%2041%20Traffic%29-%20%20U.S.%20Carriers&DB_Short_Name=Air%20Carriers

191. Reference Guide for the 2010 Surface Transportation Board. (January 26, 2012). Retrieved from http://www.stb.dot.gov/stb/docs/Waybill/2010%20STB%20Waybill%20Reference%20Guide_JN.pdf
192. Statistical Areas. Bureau of Economic Analysis. (November 21, 2013). Retrieved from http://www.bea.gov/regional/docs/msalist.cfm
193. 2010 ANSI Codes for Places. U.S. Census Bureau. Retrieved from https://www.census.gov/geo/reference/codes/place.html
194. Reference Guide for the 2010 Surface Transportation Board. (January 26, 2012). Retrieved from http://www.stb.dot.gov/stb/docs/Waybill/2010%20STB%20Waybill%20Reference%20Guide_JN.pdf
195. Conversion Tables and Units of Quantity. Guide to Foreign Trade Statistics. U.S. Census Bureau. (July 14, 2004). Retrieved from http://www.census.gov/foreign-trade/guide/sec4.html
196. 2014 TIGER Geodatabase Record Layouts. U.S. Census Bureau. Retrieved from https://www.census.gov/geo/maps-data/data/pdfs/tiger/tgrshp2014/2014_TIGER_GDB_Record_Layouts.pdf
197. T-100 Domestic Segment (All Carriers). TranStats. Bureau of Transportation Statistics. Retrieved from http://transtats.bts.gov/Fields.asp?Table_ID=311
198. Conversion Tables and Units of Quantity. Guide to Foreign Trade Statistics. U.S. Census Bureau. (July 14, 2004). Retrieved from http://www.census.gov/foreign-trade/guide/sec4.html
199. 2002 Microdata Data Dictionary. Vehicle Inventory and Use Survey. Retrieved from https://www.census.gov/svsd/www/vius/datadictionary2002.pdf
200. FAQs. Surface Transportation Board. Retrieved from http://www.stb.dot.gov/stb/faqs.html
201. 2002 Commodity Flow Survey. Bureau of Transportation Statistics. Retrieved from http://apps.bts.gov/publications/commodity_flow_survey/2002/hazardous_materials/html/2002_commodity_flow_survey.html
202. Trade Definitions. Foreign Trade. U.S. Census Bureau. Retrieved from http://www.census.gov/foreign-trade/reference/definitions/index.html
203. Air Carrier Financial Reports (Form 41 Financial Data). TranStats. Bureau of Transportation Statistics. Retrieved from http://www.transtats.bts.gov/Tables.asp?DB_ID=135
204. Notice on B1, B11, P11, P12 Examples. TranStats. Bureau of Transportation Statistics. Retrieved from http://www.transtats.bts.gov/showHelp2.asp#_SIGNEXAMPLES
205. Air Carrier Statistics (Form 21 Traffic). TranStats. Bureau of Transportation Statistics. Retrieved from http://www.transtats.bts.gov/Tables.asp?DB_ID=110&DB_Name=Air%20Carrier%20Statistics%20%28Form%2041%20Traffic%29-%20%20U.S.%20Carriers&DB_Short_Name=Air%20Carriers
206. Guide for Railroads. Association of American Railroads. Railinc. (May 30, 2013). Retrieved from https://www.railinc.com/rportal/documents/18/260641/GuideforRailroads.pdf
207. County Business Patterns (CBP) / Zip Code Business Patterns (ZBP). County Business Patterns. U.S. Census Bureau. Retrieved from http://www.census.gov/econ/cbp/
208. Quarterly Services Survey Technical Documentation. Annual & Quarterly Services. U.S. Census Bureau. Retrieved from http://www.census.gov/services/qss/qsstechdoc.html
209. United States Senate. Retrieved from http://www.senate.gov/reference/glossary_term/fiscal_year.htm
210. 2012 Surface Transportation Board Reference Guide. Carload Waybill Sample. (October 18, 2013). Retrieved from http://www.stb.dot.gov/stb/docs/Waybill/2012%20STB%20Waybill%20Reference%20Guide%20-%20FINAL.pdf
211. Conversion Tables and Units of Quantity. Guide to Foreign Trade Statistics. U.S. Census Bureau. (2004, July 14). Retrieved from http://www.census.gov/foreign-trade/guide/sec4.html
212. Vessel Characteristics. U.S. Waterway Data. Navigation Data Center. Retrieved from http://www.navigationdatacenter.us/data/dictionary/ddvess.htm
213. Conversion Tables and Units of Quantity. Guide to Foreign Trade Statistics. U.S. Census Bureau. (July 14, 2004). Retrieved from http://www.census.gov/foreign-trade/guide/sec4.html
214. 2002 Microdata Data Dictionary. Vehicle Inventory and Use Survey. Retrieved from https://www.census.gov/svsd/www/vius/datadictionary2002.pdf
215. Trade Definitions. Foreign Trade. U.S. Census Bureau. Retrieved from http://www.census.gov/foreign-trade/reference/definitions/
216. Table Structure and Data Fields for Raw Data Time Series. North American Transborder Freight Data. Bureau of Transportation Statistics. Retrieved from http://Transborder.bts.gov/programs/international/Transborder/TBDR_DataFields.html
217. SCTG Commodity Codes. 2012 Commodity Flow Survey. U.S. Census Bureau. Retrieved from https://bhs.econ.census.gov/bhs/cfs/Commodity%20Code%20Manual%20(CFS-1200).pdf
218. An Overview of the 2002 Commodity Origin-Destination Database: Methodology and Data. Freight Management and Operations. Federal Highway Administration. Retrieved from http://ops.fhwa.dot.gov/freight/freight_analysis/faf/faf2_reports/report1.htm
219. 2002 Microdata Data Dictionary. Vehicle Inventory and Use Survey. Retrieved from https://www.census.gov/svsd/www/vius/datadictionary2002.pdf

220. Trade Definitions. Foreign Trade. U.S. Census Bureau. Retrieved from http://www.census.gov/foreign-trade/reference/definitions/
221. Waterborne Commerce of the United States. U.S. Waterway Data. Navigation Data Center. (September 2, 2014). Retrieved from http://www.navigationdatacenter.us/data/dictionary/ddwcus.htm
222. Complete HS and SITC Conversion and Correspondence Tables along with Detailed Note on its Conversion Methodology. Department of Economic and Social Affairs–Statistics Division. United Nations. (2010). Retrieved from http://unstats.un.org/unsd/trade/conversions/HS%20Correlation%20and%20Conversion%20tables.htm
223. Analytical User's Manual 1975–2012. Fatality Analysis Reporting System (FARS). National Highway Traffic Safety Administration. (January 2014). Retrieved from http://www-nrd.nhtsa.dot.gov/Pubs/811855.pdf
224. Federal Motor Carrier Safety Administration. Retrieved from http://mcmiscatalog.fmcsa.dot.gov/
225. Reference Guide for the 2010 Surface Transportation Board. (January 26, 2012). Retrieved from http://www.stb.dot.gov/stb/docs/Waybill/2010%20STB%20Waybill%20Reference%20Guide_JN.pdf
226. Middendorf, D. (January 1998). Intermodal Terminals Database: Concepts, Design, Implementation, and Maintenance. Retrieved from http://cta.ornl.gov/transnet/terminal_doc/
227. Trade Definitions. Foreign Trade. U.S. Census Bureau. Retrieved from http://www.census.gov/foreign-trade/reference/definitions/
228. Federal Motor Carrier Safety Administration. Retrieved from http://mcmiscatalog.fmcsa.dot.gov/
229. General Explanation and Census of Agriculture Report Form. 2007. USDA. National Agricultural Statistics Service. Retrieved from http://www.agcensus.usda.gov/Publications/2007/Full_Report/Volume_1,_Chapter_1_US/usappxb.pdf
230. Waterborne Commerce of the United States. U.S. Waterway Data. Navigation Data Center. (September 2, 2014). Retrieved from http://www.navigationdatacenter.us/data/dictionary/ddwcus.htm
231. Correlation and Conversion Tables Used in the UN Comtrade. Department of Economic and Social Affairs–Statistics Division. United Nations. (February 2012). Retrieved from http://unstats.un.org/unsd/trade/conversions/Correlation%20and%20conversion%20tables%20used%20in%20UN%20Comtrade%20-%20General%20Note.pdf
232. Pierce, Justin R. and Peter K. Schott. (2010). Concording U.S. Harmonized System Codes Over Time, Retrieved from http://faculty.som.yale.edu/peterschott/files/research/papers/readme_hs_over_time_33.pdf
233. Schott's International Economics Resource Page, Trade Data and Concordances, Retrieved from http://faculty.som.yale.edu/peterschott/sub_international.htm
234. Ibid.
235. Standard International Trade Classification. Series M No. 34/Rev. 4. Department of Economic and Social Affairs–Statistics Division. United Nations. (2006). Retrieved from http://unstats.un.org/unsd/publication/SeriesM/SeriesM_34rev4e.pdf
236. Complete HS and SITC Conversion and Correspondence Tables along with Detailed Note on its Conversion Methodology. Department of Economic and Social Affairs–Statistics Division. United Nations. (2010). Retrieved from http://unstats.un.org/unsd/trade/conversions/HS%20Correlation%20and%20Conversion%20tables.htm
237. Correlation and Conversion Tables Used in the UN Comtrade. Department of Economic and Social Affairs–Statistics Division. United Nations. (February 2012). Retrieved from http://unstats.un.org/unsd/trade/conversions/Correlation%20and%20conversion%20tables%20used%20in%20UN%20Comtrade%20-%20General%20Note.pdf
238. Standard Classification of Transported Goods (SCTG). Statistics Canada. (June 17, 2003). Retrieved from http://www.statcan.gc.ca/subjects-sujets/standard-norme/sctg-ctbt/sctgintro-ctbtintro-eng.htm
239. A Description of the FAF3 Regional Database and How it is Constructed. The Freight Analysis Framework Version 3 (FAF3). Federal Highway Administration (June 16, 2011). Retrieved from http://faf.ornl.gov/fafweb/Data/FAF3ODDoc611.pdf
240. Exports Concordance. Foreign Trade Statistical Program. U.S. Census Bureau. Retrieved from http://censtats.census.gov/cgi-bin/concordance/concordance.pl
241. End-use commodity codes provide detailed commodity information with export and import data aggregated into broader economic categories according to the commodities' uses.
242. Hi-Tech commodity codes consist solely of high-technology products (products that are the most technology-intensive).
243. 2014 Publication Product List. Railinc Business Services. Retrieved from https://www.railinc.com/rportal/documents/18/260755/Publications_PriceList.pdf
244. Standard Classification of Transported Goods (SCTG). Statistics Canada. (June 17, 2003). Retrieved from http://www.statcan.gc.ca/subjects-sujets/standard-norme/sctg-ctbt/sctgintro-ctbtintro-eng.htm
245. Concordances. North American Industry Classification System. U.S. Census Bureau. Retrieved from http://www.census.gov/eos/www/naics/concordances/concordances.html

246. Ibid.
247. 2012 to 2007 NAICS. Concordances. North American Industry Classification System U.S. Census Bureau. Retrieved from http://www.census.gov/eos/www/naics/concordances/2012_to_2007_NAICS.xls
248. Ibid.
249. 2007 to 2002 NAICS. Concordances. North American Industry Classification System U.S. Census Bureau. Retrieved from www.census.gov/eos/www/naics/concordances/2007_to_2002_NAICS.xls
250. Ibid.
251. 2002 to 1997 NAICS. Concordances. North American Industry Classification System U.S. Census Bureau. Retrieved from www.census.gov/eos/www/naics/concordances/2002_NAICS_to_1997_NAICS.xls
252. 2002 to 1987 NAICS. Concordances. North American Industry Classification System U.S. Census Bureau. Retrieved from http://www.census.gov/eos/www/naics/concordances/2002_NAICS_to_1987_SIC.xls
253. 1997 to 2002 NAICS. Concordances. North American Industry Classification System U.S. Census Bureau. Retrieved from http://www.census.gov/eos/www/naics/concordances/1997_NAICS_to_2002_NAICS.xls
254. County Business Patterns, Historical Data. Retrieved from http://www.census.gov/econ/cbp/historical.htm
255. Economic Census, Industry Classification Updates. Retrieved from http://www.census.gov/econ/census/about/industry_class_updates.html
256. North American Industry Classification System, Concordances, Retrieved from https://www.census.gov/eos/www/naics/concordances/concordances.html
257. Bridge between NAICS and SIC. 1997 Economic Census. U.S. Census Bureau. (June 2000). Retrieved from https://www.census.gov/prod/ec97/97x-cs3.pdf
258. 2012 NAICS to SIC Crosswalk. North American Industry Classification System (NAICS). (2012). Retrieved from http://www.naics.com/naicswp2014/wp-content/uploads/2014/10/NAICS-to-SIC-Crosswalk.pdf
259. Historical Data. County Business Patterns. U.S. Census Bureau. Retrieved from http://www.census.gov/econ/cbp/historical.htm
260. Commodity Flow Survey, Bureau of Transportation Statistics, http://www.rita.dot.gov/bts/help_with_data/commodity_flow_survey.html
261. 2012 NAICS to SIC Crosswalk. North American Industry Classification System (NAICS). (2012). Retrieved from http://www.naics.com/naicswp2014/wp-content/uploads/2014/10/NAICS-to-SIC-Crosswalk.pdf
262. Reference Guide for the 2010 Surface Transportation Board Carload Waybill Sample. (January 26, 2012). Retrieved from http://www.stb.dot.gov/stb/docs/Waybill/2010%20STB%20Waybill%20Reference%20Guide_JN.pdf
263. Key CFS Terms. Commodity Flow Survey. Bureau of Transportation Statistics. Retrieved from http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/commodity_flow_survey/html/def_terms.html
264. Guide to Foreign Trade Statistics. Foreign Trade Statistics. U.S. Census Bureau. Retrieved from http://www.census.gov/foreign-trade/guide/sec2.html#mot
265. Freight Analysis Framework 3 User Guide. Federal Highway Administration. (June 2012). Retrieved from http://www.ops.fhwa.dot.gov/freight/freight_analysis/faf/faf3/userguide/faf3_guide.pdf
266. Transborder Freight Data Program Documentation. North American Transborder Freight Database. Bureau of Transportation Statistics. Retrieved from http://Transborder.bts.gov/programs/international/Transborder/PDF/TransborderFreightDataProgram.pdf
267. List of Airline Mergers and Acquisitions. Wikipedia. Retrieved from http://en.wikipedia.org/wiki/List_of_airline_mergers_and_acquisitions
268. Comparison of Industry Coverage. Commodity Flow Survey (CFS). Bureau of Transportation Statistics. Retrieved from http://www.rita.dot.gov/bts/help_with_data/commodity_flow_survey.html#comparison_of_industry_coverage
269. Analytical Users Manual 1975–2010. Fatality Analysis Reporting System (FARS). National Highway Traffic Safety Administration. (September 2013) Retrieved from http://www-nrd.nhtsa.dot.gov/Pubs/811529.pdf
270. Ibid.
271. Program Documentation. Vehicle Inventory and Use Survey. (April 12, 2006). Retrieved from https://www.census.gov/svsd/www/vius/ProgramDocumentation.pdf
272. 2002 Data Releases. Vehicle Inventory and Use Survey–Discontinued. U.S. Census Bureau. Retrieved from http://www.census.gov/svsd/www/vius/2002.html
273. Ibid.

Abbreviations

These abbreviations supplement the list of abbreviations and acronyms used without definitions in TRB publications and are used through most of NCFRP Report 35. Text in Chapters 6 and 7 limits abbreviations, however, because that text corresponds directly to the text that will appear on the Freight Data Dictionary website (http://freightdatadictionary.com). To make the Freight Data Dictionary environment as user friendly as possible, the full names of organizations and key terms are spelled out in each instance.

AADT	annual average daily traffic
AAR	Association of American Railroads
AAWDT	annual average weekday traffic
API	application program interface
ATRI	American Transportation Research Institute
BEA	Business Economic Area (used with “BEA code”)
BTS	Bureau of Transportation Statistics
CBP	U.S. Customs and Border Protection
CFS	Commodity Flow Survey
CO_2e	carbon dioxide equivalent
CODMRT	commodity, origin, destination, mode, route, and time
CTA	Center for Transportation Analysis
DOT	department of transportation
EIA	Energy Information Administration
ESAL	equivalent single-axle loads
ESC	Electronic Stability Control
FAF	Freight Analysis Framework (general)
FAF2	Freight Analysis Framework (version 2)
FAF3	Freight Analysis Framework (version 3)
FARS	Fatal Analysis Reporting System
GIFT	Geospatial Intermodal Freight Transportation
GIS	geographic information system
HMIRS	Hazardous Materials Incident Reporting System
HPMS	Highway Performance Monitoring System
ISO	International Organization for Standardization
ITS	intelligent transportation system
IRI	International Roughness Index
LPMS	Lock Performance Monitoring System
MARAD	United States Maritime Administration
MCMIS	Motor Carrier Management Information System
MOVES	Motor Vehicle Emission Simulator

MPO	metropolitan planning organization
MRI	machine-readable input
NAFTA	North American Free Trade Agreement
NAICS	North American Industry Classification System
NASS	National Agricultural Statistics Service
NBI	National Bridge Inventory
NBIAS	National Bridge Investment Analysis System
N-CAST	National Corridors Analysis and Speed Tool
NHPN	National Highway Planning Network
NO_x	nitrogen oxide (emissions)
NTAD	National Transportation Atlas Database
O-D	origin-destination
PIERS	Port Import/Export Reporting Service
RBCS	Role-Based Classification Schema
REA	Rail Equipment Accident
SIC	Standard Industrial Classification
SIP	state implementation plan
SCTG	Standard Classification of Transported Goods
SPLC	Standard Point Location Code
STCC	Standard Transportation Commodity Codes
TDM	travel demand model
TIFA	Trucks Involved in Fatal Accidents
USACE	U.S. Army Corps of Engineers
VIN	vehicle identification number
VIUS	Vehicle Inventory and Use Survey
VMT	vehicle-miles traveled

Abbreviations and acronyms used without definitions in TRB publications:

A4A	Airlines for America
AAAE	American Association of Airport Executives
AASHO	American Association of State Highway Officials
AASHTO	American Association of State Highway and Transportation Officials
ACI–NA	Airports Council International–North America
ACRP	Airport Cooperative Research Program
ADA	Americans with Disabilities Act
APTA	American Public Transportation Association
ASCE	American Society of Civil Engineers
ASME	American Society of Mechanical Engineers
ASTM	American Society for Testing and Materials
ATA	American Trucking Associations
CTAA	Community Transportation Association of America
CTBSSP	Commercial Truck and Bus Safety Synthesis Program
DHS	Department of Homeland Security
DOE	Department of Energy
EPA	Environmental Protection Agency
FAA	Federal Aviation Administration
FHWA	Federal Highway Administration
FMCSA	Federal Motor Carrier Safety Administration
FRA	Federal Railroad Administration
FTA	Federal Transit Administration
HMCRP	Hazardous Materials Cooperative Research Program
IEEE	Institute of Electrical and Electronics Engineers
ISTEA	Intermodal Surface Transportation Efficiency Act of 1991
ITE	Institute of Transportation Engineers
MAP-21	Moving Ahead for Progress in the 21st Century Act (2012)
NASA	National Aeronautics and Space Administration
NASAO	National Association of State Aviation Officials
NCFRP	National Cooperative Freight Research Program
NCHRP	National Cooperative Highway Research Program
NHTSA	National Highway Traffic Safety Administration
NTSB	National Transportation Safety Board
PHMSA	Pipeline and Hazardous Materials Safety Administration
RITA	Research and Innovative Technology Administration
SAE	Society of Automotive Engineers
SAFETEA-LU	Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (2005)
TCRP	Transit Cooperative Research Program
TEA-21	Transportation Equity Act for the 21st Century (1998)
TRB	Transportation Research Board
TSA	Transportation Security Administration
U.S.DOT	United States Department of Transportation

TRANSPORTATION RESEARCH BOARD
500 Fifth Street, NW
Washington, DC 20001

ADDRESS SERVICE REQUESTED

NON-PROFIT ORG.
U.S. POSTAGE
PAID
COLUMBIA, MD
PERMIT NO. 88

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

The nation turns to the National Academies of Sciences, Engineering, and Medicine for independent, objective advice on issues that affect people's lives worldwide.
www.national-academies.org

ISBN 978-0-309-30889-2
90000
9 780309 308892