U.S. Department
of Transportation
National Highway
Traffic Safety
Administration

DOT HS 813 342 August 2022

Pedestrian/Bicyclist Safety in Numbers Program Evaluation

DISCLAIMER

Suggested APA Format Citation:

Jackson, S., Miller, S., Goughnour, E., Boller, N., Johnson, K., Kehoe, N., & Symoun, J. (2022, August). *Pedestrian/bicyclist safety in numbers program evaluation* (Report No. DOT HS 813 342). National Highway Traffic Safety Administration.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT HS 813 342		

4. Title and Subtitle	5. Report Date
Pedestrian/Bicyclist Safety in Numbers Program Evaluation	August 2022
	6. Performing Organization Code
7. Authors Steve Jackson (Toxcel), Sheryl Miller (Toxcel), Elissa Goughnour (VHB), Nadia Boller (VHB), Kristie Johnson (NHTSA), Nicholas Kehoe (Toxcel), and Jennifer Symoun (Toxcel)	8. Performing Organization Report No.
9. Performing Organization Name and Address Toxcel LLC 7140 Heritage Village Plaza Gainesville, VA 20155	10. Work Unit No. (TRAIS) 11. Contract or Grant No. DTNH2217F00146
12. Sponsoring Agency Name and Address National Highway Traffic Safety Administration Office of Behavioral Safety Research 1200 New Jersey Avenue SE Washington, DC 20590	13. Type of Report and Period Covered Final 14. Sponsoring Agency Code NHTSA

15. Supplementary Notes

Kristie Johnson, Ph.D., was the contracting officer's representative (task order) for this project.

16. Abstract

Safety in Numbers (SIN) posits that there is an inverse relationship between the extent of walking/bicycling and the probability of a motorist collision with a pedestrian/bicyclist (Jacobsen, 2003). The research team developed a plan to evaluate the relationships between pedestrian and bicyclist programs and respective volumes, and the relationship between these volumes and crashes. Three cities emerged as suitable sites for the evaluation (Fort Collins, Colorado; Philadelphia, Pennsylvania; and Anchorage, Alaska). Programs were evaluated on how effective they were at increasing bicyclist and pedestrian volumes using established statistical models. Results for Fort Collins were mixed and raised questions about the nature of the programs evaluated and the quality of the underlying data. The bikeshare program in Philadelphia was found to positively affect bicyclist volumes with no effect on pedestrian volumes. Program data from Anchorage were insufficient for analysis. Robust, multifaceted data are required to evaluate program effectiveness and SIN. The literature and the analysis described in this report demonstrate how these are challenging to obtain. The research team encountered some challenges concerning the data and developed solutions from which future researchers may benefit.

17. Key Words	18. Distribution Statement
pedestrian, bicyclist, safety, safety in numbers, volume, crashes, behavior change, walking, biking, cyclist, risk	This document is available to the public from the DOT, BTS, National Transportation Library, Repository & Open Science Access Portal, rosap.ntl.bts.gov.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	74	N/A

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Table of Contents

List of Figures

List of Tables

List of Acronyms

AAD	annual average daily
AADB	annual average daily bicyclists
AADP	annual average daily pedestrians
AADM	annual average daily motor vehicles
BAP	Bicycle Ambassador Program
BTWD	Bike to Work Day
DVRPC	Delaware Valley Regional Planning Commission
FARS	Fatality Analysis Reporting System
MPO	metropolitan planning organization
NACTO	National Association of City Transportation Officials
OS	Open Streets
PII	personally identifiable information
SRTS	Safe Routes to School
SIN	Safety in Numbers
VMT	vehicle miles traveled

Executive Summary

This report describes an evaluation to investigate the concept of Safety in Numbers in the context of programs meant to increase walking and biking. SIN posits that there is an inverse relationship between walking/bicycling volumes and the probability of a motorist collision with a pedestrian or bicyclist (Jacobsen, 2003).

Prior to this evaluation, the research team conducted a literature review synthesizing the evidence related to the SIN theory across many fields of study and areas of practice including engineering, planning and land use, sociology, psychology, education, public health, enforcement, human factors, and others (Kehoe et al., 2022). The review concluded that there was indeed evidence that the SIN phenomenon is real. However, it also pointed out that there are many challenges in investigating SIN, including the variety, amount, and quality of data available to researchers.

With this understanding of SIN, the research team developed a plan to evaluate the relationships between pedestrian and bicyclist programs and respective volumes, and the relationship between said volumes and crashes. For purposes of this study, programs were defined as ongoing or repetitive efforts directed toward the behavior and well-being of pedestrians and bicyclists. A program scan identified 230 candidate programs, 15 of which were determined to have data suitable for analysis. After discussions with program managers and additional investigation, three cities emerged as the most suitable sites to conduct evaluations of their programs. More details on site/program selection are available in the report section *Site Selection*.

- Fort Collins, Colorado, and its Safe Routes to School program, Open Streets events, Bicycle Ambassador Program, and Bike to Work Day
- Philadelphia, Pennsylvania, and its Indego Bikeshare Initiative
- Anchorage, Alaska, and its Bikeology program

The research team worked to acquire the relevant data sets, including program metrics (e.g., numbers of participants or attendees at program events), crash data, and traffic volume data. Each of the data sets required individualized plans for data preparation. These included converting short term volume counts into annual average daily volumes, connecting and interpolating data from single-mode counters, geocoding count and crash locations, and determining appropriate crash zone sizes.

Programs were evaluated on how effective they were at increasing bicyclist and pedestrian volumes using the established statistical models noted in the work of Elvik (2013) to analyze SIN effects. Results for Fort Collins were mixed and raised questions about the nature of the programs evaluated and the quality of the underlying data. The bikeshare program in Philadelphia was found to positively affect bicyclist volumes with no effect on pedestrian volumes. Program data from Anchorage were insufficient for analysis.

The SIN phenomenon was investigated with similar statistical models. Complete SIN is said to occur when bicyclist/pedestrian crashes increase at a rate less than proportional to simultaneous increases in bicyclist/pedestrian *and* motor vehicle volumes. In contrast, partial SIN occurs when bicyclist/pedestrian crashes increase at a rate less than proportional to increases in bicyclist/pedestrian *or* motor vehicle volumes. Results indicated complete SIN for bicyclists and partial SIN for pedestrians in both Fort Collins and Anchorage, but no evidence of SIN in Philadelphia.

An ad-hoc analysis was conducted to investigate the role of infrastructure. The presence of 17 pedestrian/bicyclist facilities (bike lanes, sidewalks, crosswalks, pedestrian hybrid beacons, etc.) was coded for each crash observed in Philadelphia. This new information was added to previously described models and yielded statistically significant results for volumes and crash rates among both modes.

Robust, multifaceted data are required to evaluate program effectiveness and SIN. The literature and the analysis described in this report demonstrate how these data are challenging to obtain. The research team encountered several challenges concerning the data and developed solutions from which future researchers may benefit.

Introduction

Pedestrian and bicyclist safety remain a complex problem in the United States. In 2020 an estimated 6,205 pedestrians and 891 bicyclists and other cyclists were killed in traffic crashes (NCSA, 2021); there were 76,000 and 49,000 injuries for these groups respectively in 2019 (NCSA, 2020). Over the course of the decade 2010 to 2019, pedestrian fatalities in urban areas increased by 62%; they decreased by 4.8% in rural areas. Over the same time period, bicyclist fatalities in urban areas increased by 49%; they increased by 4.6% in rural areas. Despite a 13.2% decrease in vehicle miles traveled (VMT) in 2020 during the COVID-19 pandemic (FHWA, 2021), pedestrian fatalities did not significantly increase from 2019 to 2020 and bicyclist fatalities increased 5% (NCSA, 2021).

Transportation engineers, planners, policymakers, and advocates seek to increase walking and biking in service of public health and equity. Understanding how the amount of walking, biking, and driving affects safety is an issue that must be understood to develop effective programs, policies, and infrastructure. However, understanding this issue can be complicated by the amount and quality of data available in evaluating program implementation, changes in behavior, and safety outcomes.

One effort to understand these relationships revolves around the concept of Safety in Numbers, which seeks to explain an individual person's chance of avoiding negative consequences depending on the number of people who are also walking or biking. NHTSA undertook research to assess the existence and potential impact of the SIN effect. This research included a literature review (Kehoe et al., 2022); identification of programs designed to increase walking and/or biking; and a data-driven evaluation of any relationship between program implementation and SIN.

Safety in Numbers: What Does the Literature Say?

As part of the literature review, 250 sources were critically reviewed, including 93 domestic sources and 141 international sources. These sources span over 15 years and include older sources that are foundational to SIN. The review spanned various fields of study and areas of practice, including engineering, planning and land use, sociology, psychology, education, public health, enforcement, human factors, and others.

Overall, the literature review found that the majority of the available literature affirms that there is a SIN effect for both bicyclists and pedestrians, supported by a non-linear relationship between pedestrian and bicyclist exposure and crash risk. The effect differs by mode, and bicyclists appear to have a stronger effect than pedestrians. What follows is a brief description of the SIN concept through a description of several key studies. More information can be found in the Literature Review report (Kehoe et al., 2022).

Smeed (1949) completed groundbreaking research that later became known as Smeed's Law, proposing that increases in traffic volumes led to a decrease in fatalities per vehicle. With Smeed's research serving as a basis of his hypothesis, Jacobsen (2003) questioned whether the relationship between the number of pedestrians or bicyclists and motor vehicle traffic volumes was linear. He studied five data sets representing multiple countries to compare the amount of walking or bicycling and the injuries resulting from collisions with motor vehicles. He found that as the number of people walking or bicycling increased, the relative risk of a motor vehicle and pedestrian or bicyclist crash decreased. Jacobsen calculated that at the population level, the number of motorists colliding with people walking or bicycling will increase at roughly 0.4 power of the number of people walking or bicycling and coined the term “Safety in Numbers” to describe this effect. This means that if the number of people walking or bicycling doubled, the number of crashes between them and motor vehicles would increase by a factor of $2^{0.41}$=1.33, or only 33%.

The term “Safety in Numbers” began to take hold in the research community as various researchers investigated the concept, and an uptick in the number of studies related to the SIN theory began in 2009. In many cases, these studies referenced one or more SIN papers such as Jacobsen’s work, but they may or may not have included the term “Safety in Numbers” within the body of the paper. With this increased focus on the SIN theory, researchers began to critically review and even challenge the idea.

Some of the most influential investigations into SIN were conducted by Elvik. His 2009 literature review studied the non-linear relationship between exposure and risk to understand how a SIN effect could exist when vulnerable road users have a much higher risk of injury or death compared to motor vehicle drivers. He concluded that there is a non-linear relationship; however, he found that the SIN effect was only demonstrated when there was a large transfer of motorized trips to a non-motorized mode. He did note several concerns with the SIN effect, listed here.

- Crashes involving vulnerable road users are poorly reported in official statistics.
- The exact shape of the non-linear relationship for risk is unknown. It is possible that the SIN effect strengthens, weakens, or ceases to exist as a function of the number of vulnerable road users present.
- Data regarding injuries between vulnerable roads users (crashes involving bicyclists with pedestrians) are largely nonexistent, so it is unclear how increasing the volume of vulnerable road users will affect safety amongst themselves.
- The percentage of motorized trips transferred to a non-motorized mode required to make a SIN effect present may be unrealistic in many situations.

Elvik (2013) reviewed common crash prediction models to better understand if these models can completely confirm the existence of a SIN effect. In this paper Elvik critically reviewed the model used in Jacobsen’s 2003 research that predicted the relative risk for a unit of walking or cycling. Elvik showed that there are inherent flaws with trying to provide a SIN effect using this model due to the mathematical relationships between the model’s variables. He focused on the variables for risk and exposure. In some cases, risk can be measured as number of crashes or injuries per kilometer walked or biked. Exposure can also be measured by number of kilometers walked or biked per resident.

If risk and exposure are both measured in this fashion, there is an inherent mathematical relationship between the two – distance traveled. Elvik demonstrated this relationship using random numbers for motor vehicle volumes, pedestrian volumes, and number of crashes with realistic upper and lower limits based on a data set from Oslo, Norway.

The other crash prediction model Elvik (2013) discussed uses the number of crashes as the dependent variable and estimates the relationships between crashes and volumes using negative binomial regression. Elvik showed that this model is more appropriate for researching SIN, as it uses actual exposure data (e.g., counts) rather than proxy values and shared components, while also allowing for researchers to control for confounding variables. This model is discussed in greater detail in the *Statistical Analysis* section of this report.

SIN can be *complete* or *partial* (Elvik, 2013). Recall that SIN does not predict a *decrease* in total crashes, but rather a smaller increase in crashes than a corresponding increase in road users. *Complete* SIN is said to occur when bicyclist/pedestrian crashes increase at a rate less than proportional to simultaneous increases in bicyclist/pedestrian *and* motor vehicle volumes. In contrast, *partial* SIN occurs when bicyclist/pedestrian crashes increase at a rate less than proportional to increases in bicyclist/pedestrian *or* motor vehicle volumes.

To evaluate the research conducted thus far, Elvik and Bjørnskau (2017) conducted a systematic review and meta-analysis. The researchers compiled a list of 26 studies researching the relationship between pedestrian and bicyclist safety and volume to use as the basis of the meta-analysis. The researchers further culled this list of studies to 15, as the remaining 11 had methodological shortcomings or were lacking details that prevented them from being included in the meta-analysis. Results indicated the existence of a clear SIN effect.

While results of this meta-analysis support the existence of a SIN effect, Elvik (2017) discusses several challenges with understanding and applying the effect. One potential issue is the crash prediction model used most commonly in SIN analysis: the negative binomial regression. Elvik states that this model does not allow for turning points; it is possible that without anything else changing, negative safety implications could emerge if the percentage of vulnerable road users became too great, but these models have no way to depict this turning point given their format. Elvik's research found that no study in his review controlled sufficiently for human behavior and for quality of infrastructure. He also pointed out that most of the studies were conducted using vulnerable road user crash data from official data sets and that these data sets greatly underrepresent minor crashes. Finally, Elvik's research included papers published prior to 2016, and given the state of research at this time, he concluded that it is still not possible to determine if the SIN effect is a causal relationship or merely a statistical relationship.

Building on the results of the systematic review, Elvik (2017) further explored the strength of the SIN effect by considering additional factors that, theoretically, could affect the strength of a SIN effect: characteristics of the pedestrians and bicyclists, and characteristics of the built environment, such as infrastructure availability and design. Interestingly, characteristics of the motor vehicle driver were not mentioned. Much like his previous meta-analysis, this study was a cross-sectional study that looked at results of other existing studies. Findings indicated that cross-sectional data show a tendency for the SIN effect to weaken as the number of pedestrians or bicyclists increases. Second, although this finding was not statistically significant, Elvik highlighted this weakening trend in the data and hypothesized that it may be due to the ratio of motor vehicles to bicyclists or pedestrians. Ultimately, he was unable to find a clear relationship between the strength of the effect and the ratio of the groups.

Objective of the Evaluation

Following the literature review, the next phase of research was to examine whether programs to increase walking and biking can demonstrate a SIN effect. To investigate this, a scan of walking and biking programs across the United States was conducted to understand the purpose, content, and metrics associated with program implementation. Particular sites were sampled, and existing data were analyzed to investigate the effect of the program on walking and biking and any associated safety impacts that could provide support for SIN.

Methodology

The evaluation methodology involved site selection, data acquisition, data preparation, and data analysis. Each of these is described in the following sections.

Site Selection

As part of the overall program scan, the research team identified 230 programs from all 50 States and Washington, DC. This also included several national-level programs. Of the 230 programs identified, 48% focused on bicyclists only, 15% focused on pedestrians only, and 37% focused on both bicyclists and pedestrians.

A mix of program sponsors were identified including advocacy groups, local agencies, State agencies, Federal agencies, metropolitan planning organizations, nonprofits, universities, associations, employers, public schools, public health departments, and faith-based groups.

Each program was classified by its messaging focus. These focus areas included: community, environment, health, money savings, planning, recreation, safety, and time savings.

This information was compiled during 2019 and 2020. Criteria were established to select sites to work with in the evaluation phase of the project. A summary of the programs is included in Appendix A.

> What Is a Program?
>
> For the purposes of this project, a "program" clearly has an initiative directed toward the behavior and well-being of pedestrians and bicyclists; occurs on an ongoing or repetitive basis rather than just a one-time effort; potentially has data available regarding increasing the number of pedestrians and/or bicyclists and/or safety outcomes; and has program elements or activities that are either unique or can be duplicated in other communities.

Site Selection Criteria

In order to identify which of the 230 programs had the greatest potential for a SIN evaluation, the research team assessed information on two major topics: a) the effectiveness of local bicycling and walking promotion programs at increasing participation in these modes; and b) any potential linkages between changes in bicycling and walking rates in the localities, and pedestrian and bicyclist crash rates in the cities.

The research team developed the following criteria for the assessment to screen candidate bicycle and pedestrian programs for further study using several criteria.

- Walking and bicycling culture (supportive of walking and biking; availability of sidewalks, trails, bike paths)
- Weather (temperate climate, conducive to walking and biking)

- Geographic diversity (range of landscapes)
- Target audience (e.g., school children, adults, healthcare workers, older adults, corporate employees)
- Scale (the research team determined that smaller scale programs may provide higher quality data and therefore were preferred – city or community versus an entire State)
- Available program participation, crash, and volume data (e.g., outreach/media exposure data, community surveys, pedestrian and/or bicyclist counts, enforcement data, exposure data)

Using the criteria listed above, the research team identified 15 locations as promising areas for further study of their pedestrian and/or bicycle education or encouragement programs. In this case promising meant that the area appeared to meet the criteria based on publicly available information, but this had not been confirmed by program officials. Appendix A: Preliminary Program Site List provides more details on the 15 initially selected locations and is organized as follows.

- Statewide programs.
- Local/regional programs run by advocacy groups or other non-governmental organizations, often in partnerships with the local/regional governments.
- Local/regional government agencies that are known to be heavily involved in pedestrian and bicyclist safety.

For the locations that have more than one education or encouragement program listed, the research team planned to evaluate how these programs work together as a whole to improve pedestrian and bicycle safety. For all other locations, the team used internet resources and worked through the State or local agencies to identify appropriate contacts within specific programs to make an inquiry to obtain data.

The research team conducted outreach with the candidate program agencies to determine their interest in participating in the study and ability to provide relevant data. The research team held preliminary phone conversations with the programs identified in Table 1 to determine which of these programs have the type of data and information needed to conduct this project and identify which participants were willing to contribute and share the information with the research team for their program evaluation.

During the initial screening call, the research team shared the intent of the research, outcomes of the literature review, and the list of program measures and possible data sources needed for the analysis. The program candidates provided available data samples to the research team for the initial review and determination for the final site selection.

Selected Sites

After the initial screening, the research team examined the data samples received from the potential participants, evaluated collecting information, and carefully eliminated candidates that were not able to meet the study criteria. It is important to mention that, generally, staff from most agencies were very excited to participate in the study and provide all of the information that was available at the time of initial outreach. The enthusiasm of the staff from each of the programs was very encouraging, although, in many cases, the needed data were insufficient or unavailable for participation in the study.

The collection of program participation data was one of the most challenging steps of this research study. Some of these educational programs are run with limited staff and volunteers and participation data collection is not a priority for many. Some do not have the capacity and/or equipment to record attendance or registration data. Some programs are new and while they had some participation data, the collection period was often not sufficient to be included in the analysis. The research team explored the possibility of using website analytics and social media metrics for the selected programs but was not successful in obtaining those.

The initial larger list was reduced to five cities. Upon closer review of data samples, the following three sites/programs were selected for the analysis.

- Fort Collins, Colorado – Safe Routes to School, Open Streets, the Bicycle Ambassador Program, and Bike to Work Day
- Philadelphia, Pennsylvania – Indego Bikeshare Initiative
- Anchorage, Alaska – Bikeology

Fort Collins, Colorado

SRTS is a nationwide program to increase the number of students safely walking and biking to school. The target audience is school-aged children. The main goal of the SRTS program in Fort Collins is to achieve 50% of local K-12 students biking and walking to school safely on a regular basis. This program aims to shift children's travel from their parent's car to walking and biking to school through education, training, and encouragement activities. Students and their families are reached through bike clubs, bike rodeos, and other SRTS initiatives. While the SRTS program's target audience includes K-12 students, the program encourages whole family participation.

The OS initiative creates temporary routes for bicycling, walking, and other active modes of transportation. These streets feature reduced car traffic, slower and safer traffic speeds, and additional traffic calming measures. The initiative creates safer places to exercise, to be mobile, and to maintain social connections with neighbors and friends. This program targets participants of all ages.

The BAP in Fort Collins consists of a group of community members and, through community events, aims to educate all drivers on the best and safest ways to share the road with people on bicycles. The target audience includes both bicyclists and drivers. BAP events include classes, group rides, webinars, and hands-on training in schools.

BTWD is a semi-annual event designed to encourage city residents to commute by bicycle. Breakfast, water, and promotional materials (T-shirts, lanyards, etc.) are provided to participants. This program primarily targets the adult, working population.

Philadelphia, Pennsylvania

One program was identified in Philadelphia, the Indego Bikeshare Initiative. The city launched Indego in 2015, offering affordable access to bicycles for city residents and visitors. The city's Office of Transportation, Infrastructure, and Sustainability owns the bicycles and docking stations and manages the program. In addition to the physical bicycles and docking stations, the program includes education and encouragement elements such as neighborhood ride guides, bicycle riding classes, bicycle safety tips and videos, and ridership rewards.

Anchorage, Alaska

Anchorage's Bikelogy program provides bicycles to students to teach them bicycle safety skills. In order to develop this program, the Alaska Injury Prevention Center received funding from Alaska's SRTS program to put together a bike fleet for the Anchorage School District. In addition, to house and transport the bikes, the school district was able to obtain a trailer to allow teaching bicycle safety at different locations. Classroom and on-the-bike instruction are being taught using the Bikeology curriculum designed by the American Alliance for Health, Physical Education, Recreation, and Dance with funding from NHTSA.

Data Acquisition

This section describes the data acquired for analysis, as well as the necessary data preparation steps and statistical models employed to evaluate program effectiveness and investigate the SIN effect. Table 1 provides a summary.

Building on the sample data sets requested earlier in the site selection process, the research team worked with each of the agencies to collect full data sets needed for SIN analysis. In each instance, a person from the research team reviewed the data for personally identifiable information before sharing it with the analysts. In some instances, the program agency may not have had complete data sets available, and so the project team identified other sources for the data. The following sections describe each of the data sets that were used.

Table 1. Summary of data acquired and necessary preparation

Site	***Program data***	***Crash data***	***Traffic volume data***	***Data preparation***
Fort Collins	Safe Routes to School (adults reached and trained, students reached and educated) Open Streets (participants) Bicycle Ambassador Program (event attendees) Bike To Work Day (participants)	Individual crash records, provided by the Fort Collins Traffic Operations department	Short-term counts, made available to the public by the City of Fort Collins	Converting short-term counts to AAD volumes Determining an appropriate crash zone size
Philadelphia	Indego Bikeshare (city-wide stations and trips, localized trip origins)	Individual crash records, made available to the public by the Pennsylvania DOT	Average Annual Daily (AAD) volumes, collected and estimated by the Delaware Valley Regional Planning Commission	Connecting and interpolating data from single-mode counters Determining an appropriate crash zone size
Anchorage	Bikeology (city-wide education program that provided bicycles and training to local students)	Individual crash records, made available to the public by the Alaska DOT	AAD volumes, made available to the public by the Alaska DOT	Geocoding crash and volume locations Connecting and interpolating data from single-mode counters Determining an appropriate crash zone size

Program Data

Each participating program's data came from the city's staff and/or information posted on the program website. The summary description and events with evaluation activities for each program selected to participate in this study is provided in Table 2.

Table 2. Summary of programs and metrics analyzed

Site	*Program Name*	*Metrics*
Fort Collins	Safe Routes to School	Students reached Students educated Adults reached Adults trained
	Open Streets	Participants
	The Bicycle Ambassador Program	Attendees
	Bike to Work Day	Participants
Philadelphia	Indego Bikeshare	City-wide trips City-wide stations Localized trip origins
Anchorage	None	None

Fort Collins, Colorado

Fort Collins provided a total of seven metrics associated with four distinct programs: Safe Routes to School, Open Streets, the Bicycle Ambassador Program, and Bike to Work Day. The following are a description of those metrics.

- Four SRTS program metrics: 1) students reached, 2) students educated, 3) adults reached, and 4) adults trained. SRTS defines students reached as the number of students exposed to the program, distinct from those who received education by participating in a class. Adults reached are similarly the number of adults exposed to the program; adults trained refers to the number of adult volunteers trained in leading groups of children to schools, serving as crossing guards, and other duties that support the SRTS mission.
- The OS program metric was event attendance.
- The BAP metric was also measured by attendance at each of the community events.
- The BTWD program metric was attendance on the day of the event.

Philadelphia, Pennsylvania

Individual metrics for each of the education and encouragement efforts – such as the bicycle riding classes – were not identified. Instead, as these efforts all likely contribute to increased ridership, Indego ridership was the metric selected. Publicly available trip summary data were used to calculate the program's total annual number of trips and docking stations. The number of trips originating from each station was also used as a more localized measure of the program's reach.

Anchorage, Alaska

While this program provided robust opportunities for students to learn how to bike; unfortunately, the research team was not able to obtain any specific program metrics to use in the analysis.

Crash Data

Many States maintain public-facing crash databases to enable research and promote safety. Crash data for Philadelphia and Anchorage were extracted from respective State DOTs. Fort Collins crash data were provided by the city's Traffic Operations department. All crash data used in this analysis were sourced from each site's respective State DOT. Records described the location, date, and severity of all crashes involving at least one bicyclist or pedestrian. Single-vehicle bicycle and pedestrian crashes (i.e., those crashes with a single bicyclist or a single pedestrian and not involving motorized vehicles) were not excluded as they may reflect bicyclists' and pedestrians' attempts to avoid contact with motor vehicles. Fort Collins and Philadelphia records used geographical coordinates; Anchorage crash data were identified by intersecting roadways and geocoded using the method described in the *Geocoding Crash and Volume Locations* section.

Traffic Volume Data

Traffic volume data were provided in very different formats. Fort Collins volume data were extracted from publicly available turning movement reports. These short-term counts were converted to AAD volumes prior to analysis (see *Data Preparation: Converting Short-Term Counts to AAD Volumes* for details). Philadelphia volume data were extracted from publicly available counts collected by the Delaware Valley Regional Planning Commission (DVRPC). DVRPC conducted counts using pneumatic tubes and converted them to AAD volumes. However, each count reflected individual road user groups (e.g., a count of only bicyclists or only pedestrians), necessitating geographical matching and temporal interpolation to be made suitable for analysis (see *Data Preparation: Connecting and Interpolating Data From Single-Mode Counters* for details). Anchorage volume data were in a similar format and required an additional step to convert intersecting roadways into geographical coordinates (see *Data Preparation: Geocoding Crash and Volume Locations* for details).

Infrastructure Data

Data on the presence of 17 pedestrian/bicyclist facilities were coded for the city of Philadelphia to support an ad-hoc investigation into the role of infrastructure. For each observation, the research team used Google Streetview to determine if these facilities were present for each year represented in the data. Google Streetview provides three-dimensional photos along many of the country's roadways and allows users to select among photography from various years. Photography in cities is updated more often than rural areas. For Philadelphia, direct observations were made the majority of the time; when photography was unavailable for a particular year, the team relied on context to infer the presence of facilities. For example, if a bike lane is visible in years 2015 and 2017, it was assumed to also be present in 2016. More ambiguous cases were left uncoded. For example, if 2018 photos showed a bike line but 2015 photos did not, the first year with the bike lane is unclear, thus its presence would not be coded and the 2016 and 2017 data points would not be included in statistical models.

Data Preparation

The statistical models employed in this analysis impose certain spatial and temporal requirements on the input data. Namely, one complete observation has at least two volumes (Average Annual Daily Motor Vehicle *and* Average Annual Daily Bicyclist *or* Average Annual Daily Pedestrian), crashes, and program metrics for one location for one year. As the data gathered for this analysis were not originally recorded with this in mind, each site underwent a data preparation phase, tailored to the format and content of the available data. For Fort Collins, one-hour turning movement reports had to be converted to AAD volumes. Counts in Philadelphia were provided as AAD volumes, but each count focused on just one road user group and thus had to be matched geographically to produce a full observation. All data from Anchorage (both volumes and crashes) were identified by intersecting roadways and thus had to be geocoded to geospatial coordinates. Finally, an appropriate crash radius had to be determined for each site.

Converting Short-Term Counts to AAD Volumes

Fort Collins conducts turning movement studies and makes the reports available to the public. These reports provide the results of counts of vehicles, bicyclists, and pedestrians at intersections. Each study consists of three 1-hour counts: in the morning (7:30 a.m.– 8:30 am), at midday (12 p.m. – 1 p.m.) and in the evening (4:30 p.m. – 5:30 p.m.). For this study, the morning counts were used, as they were thought to be the most consistent in terms of volume over time, as the morning data collection period coincides with the typical morning rush hour.

Figures 1 and 2 provide examples of one turning movement report from which volume data were extracted. Figure 1 provides the number of motor vehicles observed on each leg of the intersection during the morning count, while Figure 2 provides the number of bicyclists and pedestrians observed on each leg of the same intersection during the same time. This particular report indicates that 2,320 motor vehicles (Figure 1), 167 bicyclists, and 66 pedestrians (Figure 2) passed through this intersection during the peak one-hour AM period.

	Center Southbound					Prospect Westbound					Centre Northbound					Prospect Eastbound					
Start Time	Right	Thru	Left	Peds	App. Total	Right	Thru	Left	Peds	App. Total	Right	Thru	Left	Peds	App. Total	Right	Thru	Left	Peds	App. Total	Int. Total
07:30 AM	2	13	5	0	20	32	122	69	0	223	44	50	10	1	105	23	180	31	1	235	583
07:45 AM	1	20	6	0	27	35	135	62	0	232	60	51	7	0	118	25	253	60	0	338	715
Total	3	33	11	0	47	67	257	131	0	455	104	101	17	1	223	48	433	91	1	573	1298
08:00 AM	1	15	4	0	20	11	110	64	0	185	49	33	3	0	85	19	185	20	0	224	514
08:15 AM	6	14	3	0	23	24	124	43	0	191	38	26	18	0	82	18	182	12	0	212	508
Grand Total	10	62	18	0	90	102	491	238	0	831	191	160	38	1	390	85	800	123	1	1009	2320
Apprch %	11.1	68.9	20	0		12.3	59.1	28.6	0		49	41	9.7	0.3		8.4	79.3	12.2	0.1		
Total %	0.4	2.7	0.8	0	3.9	4.4	21.2	10.3	0	35.8	8.2	6.9	1.6	0	16.8	3.7	34.5	5.3	0	43.5	

Figure 1. Example turning movement report: motor vehicles

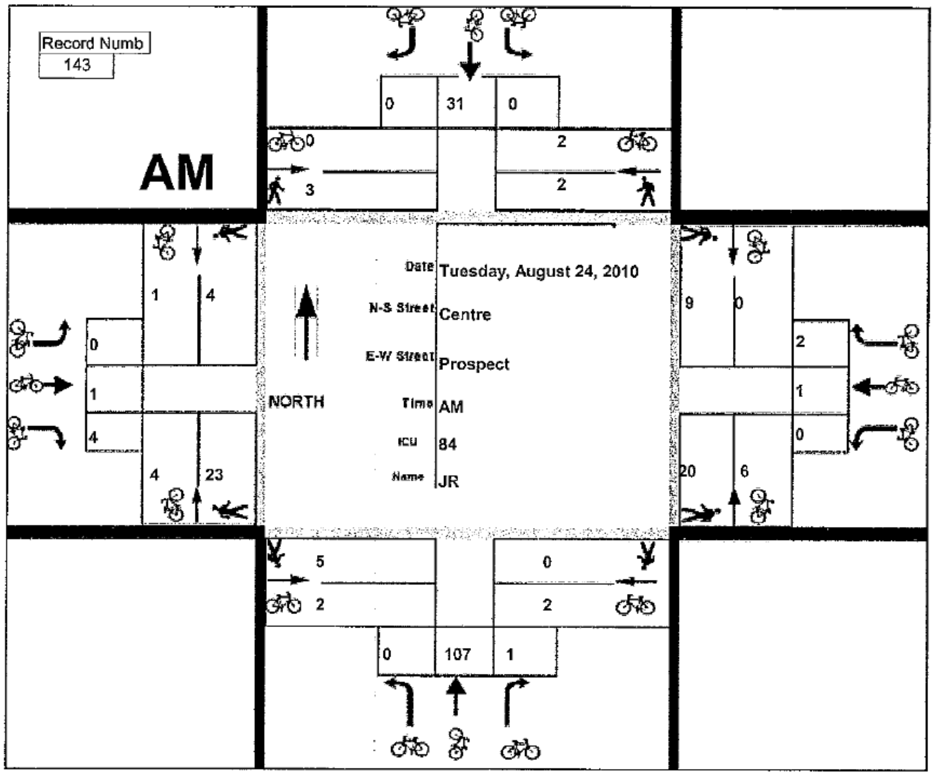

Figure 2. Example turning movement report: bicyclists and pedestrians

Short-term counts must be converted to AAD volumes to conform with the statistical model described in the *Statistical Analysis* section. A method was developed to leverage data from continuous counters to convert one-hour counts to AAD volumes. The method essentially determines the relationship between short-term counts and annual totals from continuous counters and applies a factor to the counts made during the turning movement studies. It is similar to the methodology outlined in section 4.5 of the FHWA's Traffic Monitoring Guide (FHWA, 2016).

Step 1: Acquire Continuous Count Data

Continuous count data for at least one year are necessary for converting short-term counts to AAD volumes due to the multiple layers of seasonality. Road user volumes have been shown (confirmed below) to vary by time of day, day of week, month, and other factors.

Ideally, continuous counters would be located in the same community as the short-term studies were conducted. The research team acquired bicyclist and pedestrian counts from five continuous counting stations in Fort Collins, but these were found to be problematic in several ways. Two of the counters failed to differentiate between pedestrians and bicyclists, and two more only counted bicyclists. The fifth counted bicyclists and pedestrians separately, but was located on a trail, and was thus considered potentially nonrepresentative of the number of bicyclists exposed to motor vehicles. Furthermore, none of these stations counted motor vehicles; if these counters were used for bicyclist and pedestrian volumes, counts from elsewhere would be required for motor vehicles, and the use of different data sources could introduce bias or error.

Data from counters in Denver were used in place of counters from Fort Collins. Denver is approximately 60 miles south of Fort Collins. Although these cities differ in terms of size, density, population demographics, and other factors, they experience remarkably similar climates. Table 3 compares the monthly average temperatures, precipitation, and snowfall in the two cities. These three weather phenomena affect the choice to walk or bicycle. The similarity of Denver's climate to that of Fort Collins makes Denver a suitable proxy for converting short-term counts to AAD volumes.

Table 3. Comparison of climates in Fort Collins and Denver

Metric	***Location***	***Jan***	***Feb***	***Mar***	***Apr***	***May***	***Jun***	***Jul***	***Aug***	***Sep***	***Oct***	***Nov***	***Dec***
Average high (°F)	Fort Collins	44	47	55	62	71	80	87	84	76	64	51	43
	Denver	44	46	54	61	71	81	88	86	77	65	52	43
Average low (°F)	Fort Collins	18	21	28	35	44	53	58	57	47	36	26	18
	Denver	17	20	26	34	44	53	59	57	47	36	25	17
Average precipitation (inches)	Fort Collins	0.40	0.40	1.59	2.06	2.43	2.17	1.71	1.60	1.33	1.15	0.76	0.50
	Denver	0.47	0.47	1.25	1.74	2.30	1.69	2.05	2.06	1.06	1.08	0.82	0.59
Average snowfall (inches)	Fort Collins	8	7	13	6	1	0	0	0	1	4	9	8
	Denver	7	6	11	7	1	0	0	0	1	4	9	9

Source: US Climate Data, 1981-2010.

Three continuous counting stations in Denver were identified. Each counter measured one road user type: motor vehicles (MVs) or bicyclists (Bikes) or pedestrians (Peds). Hourly aggregated counts for 2017 were acquired from the Colorado DOT's Online Transportation Information System. Figure 3 plots raw hourly counts for each road user group for all of 2017 while Figure 4 provides a closer look at one week beginning on Saturday, July 1, 2017. These plots highlight outliers and missing data patterns. The proceeding steps address these issues.

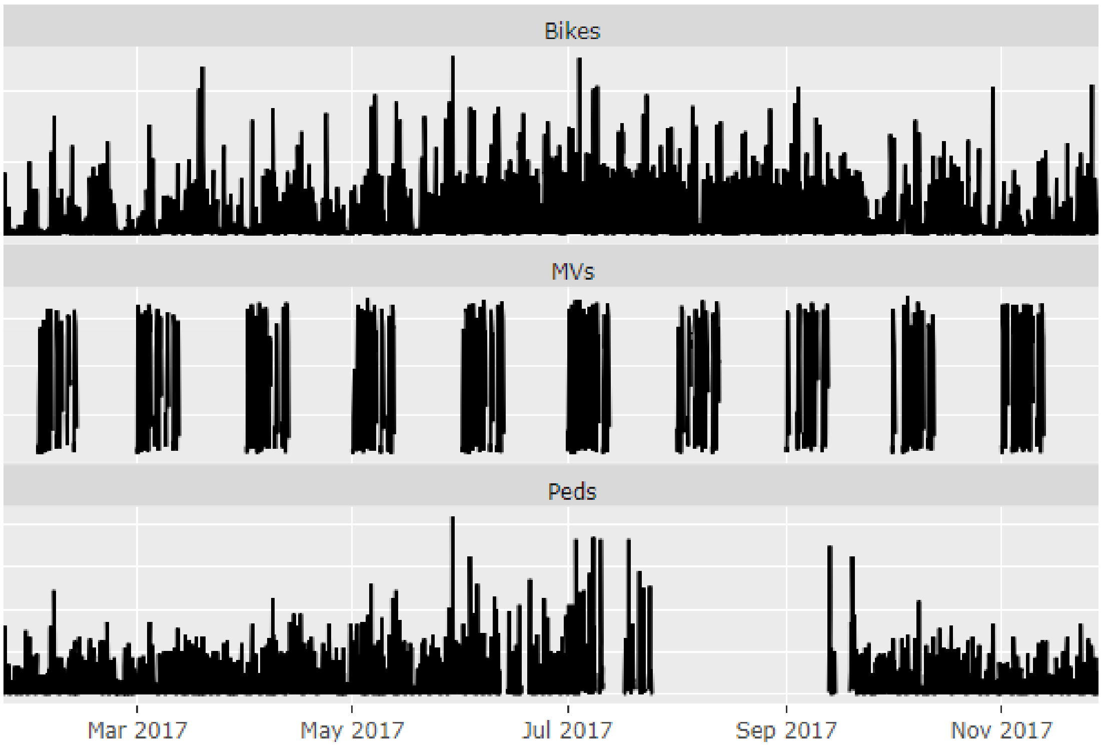

Figure 3. Raw hourly counts from Denver's continuous counting stations, 2017

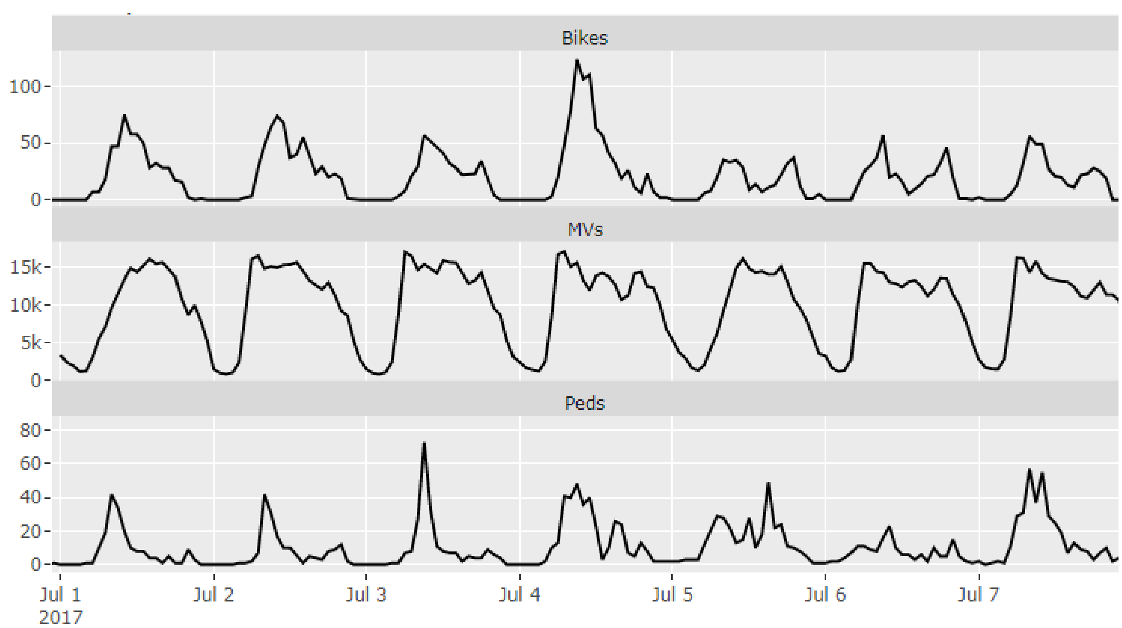

Figure 4. Raw hourly counts from Denver's continuous counting stations, July 1-8

Step 2: Adjust Outliers

A visual inspection of Figure 4 reveals the presence of outliers, particularly with bicyclists. One outlier occurred during the 9:00 a.m. hour of July 4, when 124 bicyclists were observed, compared to 43 exactly one week earlier. This was likely due to a holiday bike ride, but the source or motivation underlying many outliers is largely unknown and unknowable. Adjusting such outliers will lead to more stable estimates in Step 3 and thus more reliable factors to convert short-term counts to AAD volumes.

Outliers were adjusted using a method known as *seasonal and trend decomposition using locally weighted least squares regression*. This method essentially identifies the trend and seasonal components of a time series. The *remainder* (original data, minus trend, minus seasonal components) is then examined to identify outliers, defined as $r_t < Q1 - 3IQR$ or $r_t > Q3 + 3IQR$ where r_t represents the remainder at time t, $Q1$ and $Q3$ represent the first and third quartiles (25th and 75th percentiles) of the remainder, respectively, and $IQR = Q3 - Q1$. Outliers are then replaced via linear interpolation applied to the seasonally adjusted data (Hyndman & Khandakar, 2008). Figure 5 shows the observed (in black) and adjusted (blue) counts to illustrate the effects of this process; the two lines overlap at all points shown except for several peak hours on July 4, where several bicyclist counts were deemed outliers and adjusted downward.

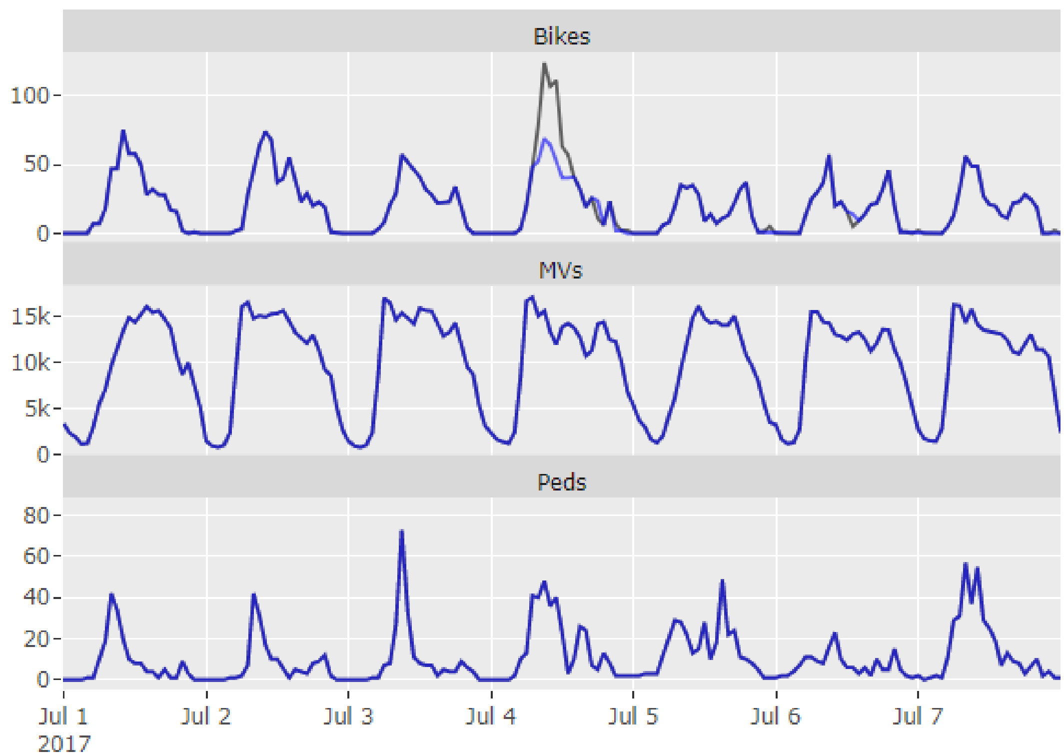

Figure 5. Results of outlier adjustment

Step 3: Impute Missing Data

Converting short-term counts to AAD volumes requires an annual total for each road user type. Missing counts bias the annual totals downward. The goal of this step is to accurately estimate missing counts.

The data acquired in Step 1 exhibits two missing data patterns: motor vehicle counts are missing in regularly spaced (two-week) intervals, perhaps to minimize costs; while pedestrian counts are missing from mid-July to mid-September, perhaps due to a hardware malfunction.

Generalized linear regressions were used to estimate missing motor vehicle and pedestrian volumes, separately. Bicyclist volumes were not modelled because there were no missing data points. The two models were similar in many ways. Both used Poisson response distributions, as is appropriate when analyzing count data. Both used second-degree polynomials to represent the month ($month, month^2$ where January = 1, February = 2, etc.). Using categorical terms would enable a simpler calculation of monthly volume factors, but the pedestrian data were missing the entire month of August, necessitating a numerical stand-in. Further, the monthly trend in non-motorized volumes is known to be parabolic in shape, peaking in the warmer summer months, hence the *second* degree term. Both regressions used categorical terms for hour (allowing each hour to exhibit a mean independent of the others and irrespective of a linear trend), a dummy variable to indicate the weekend (Saturday and Sunday), and included an interaction term for weekend and hour.

The pedestrian model also included a term for bicyclist volume. Counts for the two modes exhibit a Spearman rank correlation of 0.80, thus the inclusion of bicyclist volumes on a model of pedestrian volumes was considered appropriate. Indeed, the model produces a positive and statistically significant relationship. Figure 6 shows observed (in black) and imputed (blue) counts to illustrate the high degree of accuracy produced by this process. There were zero missing observations for bicyclists. Only missing values were replaced with imputed ones. Figure 7 extends the view to the entire month of July to illustrate the values imputed during longer stretches of missing data.

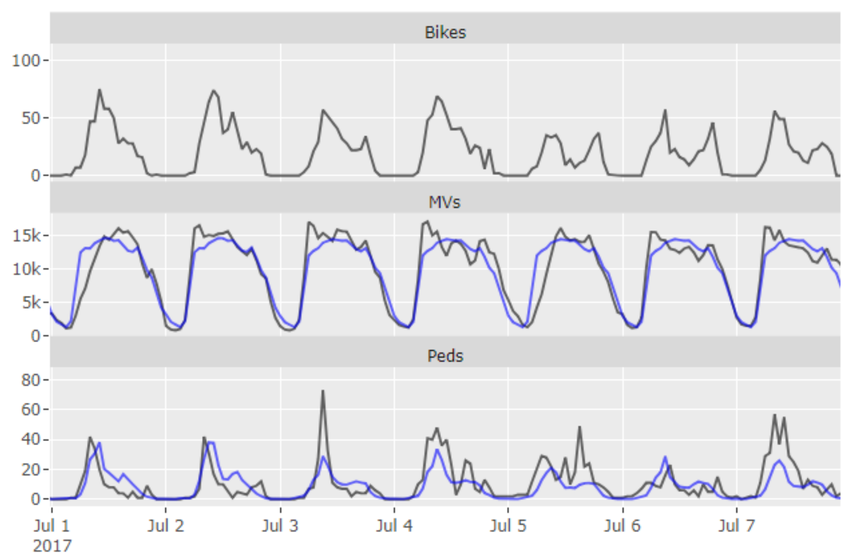

Figure 6. Results of missing data imputation

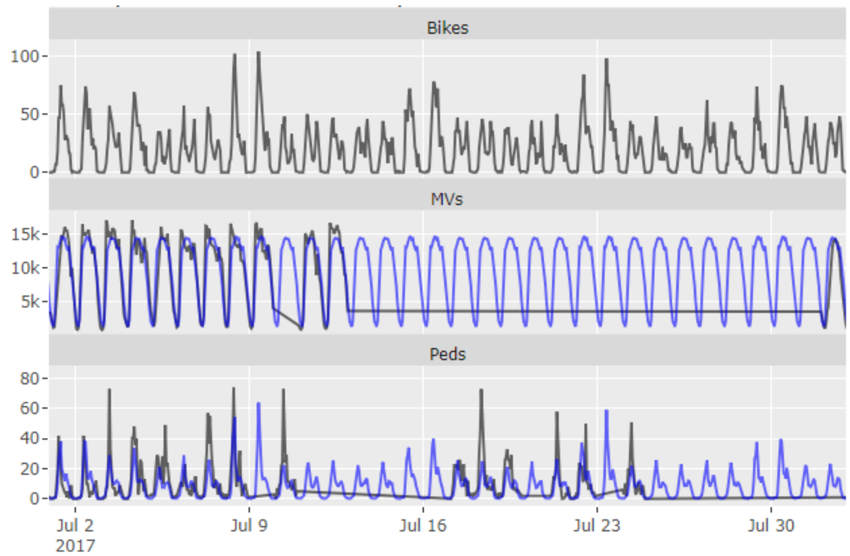

Figure 7. Extended results of missing data imputation

Step 4: Calculate Factors

The goal of this process is to develop factors to convert short-term counts to AAD volumes. Having completed Step 3, the hourly continuous counts have been adjusted for outliers and rid of missing data via imputation. Now factors can be developed to relate the short-term observations (in the continuous data) to AAD volumes (also in the continuous data). Thirty-six factors were calculated, one for each month for each of three road user groups. Mathematically, these factors can be expressed as:

$$F_{u,m} = \frac{AAD_u}{s_{u,m}}$$

where $F_{u,m}$ represents the factor for road user group u during month m, AAD_u represents the AAD volume for each road user group (i.e., AADB, AADM, AADP) and $s_{u,m}$ represents the short-term count for road user group u during month m. AAD volume is calculated simply as the sum of all road users of a given group in a single year divided by 365. For the counters in Denver during 2017, AADB = 255, AADM = 235,755, and AADP = 148. The relevant short-term counts are those that coincide with the turning movement studies in Fort Collins, all of which took place between 7:30 and 8:30 a.m. on weekdays.[1] This reduces the relevant counts to approximately 20 (5 weekdays in each of 4 weeks) observations per month per road user group, which were then averaged. Table 4 shows the resulting factors for each month and road user group.

[1] Note that the turning movement reports cover the period 7:30-8:30 a.m. while the continuous counts are given for the entire 7:00 and 8:00 hour. To overcome this scale difference, the 7:00 and 8:00 hour counts were averaged.

Table 4. Calculated factors to convert short-term counts to AAD volumes

Month	***Bicyclists***	***Motor Vehicles***	***Pedestrians***
Jan	150.3	18.0	32.1
Feb	71.0	19.2	21.9
Mar	56.4	18.4	16.5
Apr	41.5	19.2	12.4
May	25.4	19.3	12.1
Jun	10.1	17.5	8.8
Jul	7.7	17.7	6.6
Aug	12.4	18.2	11.1
Sep	16.6	18.6	9.9
Oct	36.1	18.4	18.8
Nov	44.1	17.9	22.3
Dec	94.6	17.6	31.0

Note how the factors are nearly constant for motor vehicles, but much higher during winter months relative to summer months for bicyclists and pedestrians. This reflects the climate-driven seasonality of non-motorized traffic and the result of sampling: if bicyclists (or any group of road users) are less active during a certain time of year, then the probability of observing them during a one-hour count is also lower. Some constant value of AADB exists for every day of the year, but the one-hour counts must be adjusted differently in each month to bring the count to AADB.

Turning movement study counts were converted to AAD volumes by multiplying the count by the appropriate factor. For example, the turning movement study conducted in January 2011 at the intersection of Boardwalk Drive and Harmony Road (in Fort Collins, Colorado) reported 7 bicyclists, 2,790 motor vehicles, and 7 pedestrians. Therefore:

$$\widehat{AADB} = 150.3 \times 7 = 1{,}052$$
$$\widehat{AADT} = 18.0 \times 2{,}790 = 50{,}220$$
$$\widehat{AADP} = 32.1 \times 7 = 225$$

Estimated AAD volumes are mapped in Figure 8.

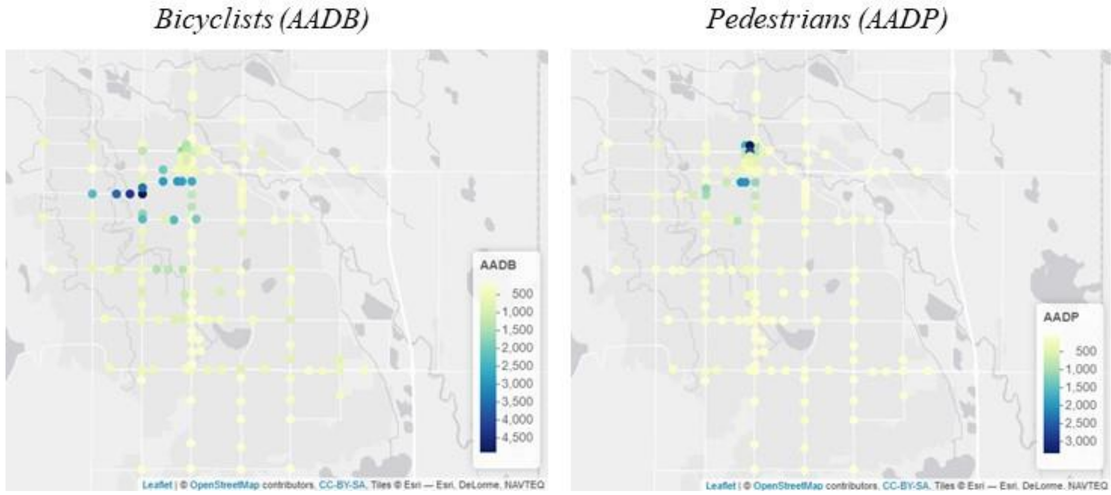

Figure 8. Maps of estimated AAD volumes in Fort Collins, Colorado

Note: All map figures were developed using [Leaflet] | Map street data © [OpenStreetMap] contributors, CC-BY-SA, Tiles ©Esri – Esri, DeLorme, NAVTEQ

Connecting and Interpolating Data From Single-Mode Counters

Philadelphia and Anchorage provided single-mode count data: locations where just one mode of transportation was measured. SIN models require co-located counts. This section describes how the single-mode counters in Philadelphia were connected to form one observation with AADM and either AADB or AADP per year and interpolated across missing years. The same methodology was applied to Anchorage.

Philadelphia provided single-mode count data from 5,156 unique locations between 2005 and 2019. These counts were conducted by various organizations (Pennsylvania DOT, private consultants, volunteers) and consolidated into a central database managed by the DVRPC, the Federally designated MPO for the city. DVRPC converts these counts to AAD volumes using an unknown methodology.

DVRPC count data included highly localized GPS coordinates (in decimal form, to the fifth decimal place). Figure 9 provides several examples of nearby locations; the six points within the circle must be consolidated. Without a consolidation step, none of the provided counts would appear to be co-located, as is required by the model described in a previous section. Locations within 30m of one another were consolidated into one location. In doing so, one-way counts on two-way roads were summed to yield the total AAD volume observed at a given location.

Figure 9. Map of selected count locations

The above process yields one data point for each count, but counts are still single-modal. That is, the six points highlighted in Figure 9 have been reduced to three and must be further reduced to one to be included in the statistical model previously described. However, the timing of many counts presents a problem; counts of different modes have been co-located but were conducted in different years. To mitigate the timing issue, annual AAD volumes were linearly interpolated *between* years. For example, if AADP at one location was 70 in the year 2010, missing for 2011, and 90 in the year 2012, the 2011 value would be interpolated as 80. Note that AAD volumes were not *extra*polated, only *inter*polated. The interpolation step makes more complete observations possible, but it does not extend trends beyond existing observation windows.

The consolidated, interpolated, single-modal AAD volumes were then connected using two criteria: the distance between the two locations, with matches being no more than 300 m apart;[2] and whether they were located along the same road. Specifically, locations were matched to minimize the distance between them, giving preference to locations along the same travel corridor, as long as they were still within 300 m of one another. Matched pairs retained the location of the non-motorized AAD volume for crash counting purposes.

To mitigate the error introduced by the matching process, statistical weights were calculated to give more weight to closer matches. Weights were calculated as $w_i = 1/\log(d + 3)$ where w_i represents the weight for observation i, d is the distance between the matched locations in meters, and 3 is a correction factor to lower the maximum weight to 0.9 for observations where $d = 0$.

Geographic matches were then matched by year. The overlapping years of data combine to form complete observations. Note that some statistical models require volumes for all three modes, while others require only two (AADM+AADB or AADM+AADP).

[2] A match radius of 500 m was used for Anchorage due to its greater geographical dispersion.

Figure 10 illustrates the methodology. Point A is a pedestrian count location to be matched to a motor vehicle count location, represented by points B and C. The circle surrounding point A has a radius of 300m. Points B and C are both within the 300 m-radius. Point B is closer to Point A than is Point C, but Points A and C are on the same road. Thus, the pedestrian count location (Point A) matches to the motor vehicle count location labelled Point C. If Point A has AADP for 2005-2016 and Point C has AADM for 2009-2018, then complete observations (AADM+AADP) would exist for 2009-2016.

Figure 10. Illustration of count location connection methodology

Figure 11 provides examples of AADM+AADB matches. Note that a small amount of random noise has been added to visually separate nearly overlapping points. All points shown belong to a pair. Of the 803 AADB locations, 681 (85%) were matched to AADM locations; of the 305 AADP locations, 225 (74%) were matched to AADM locations.

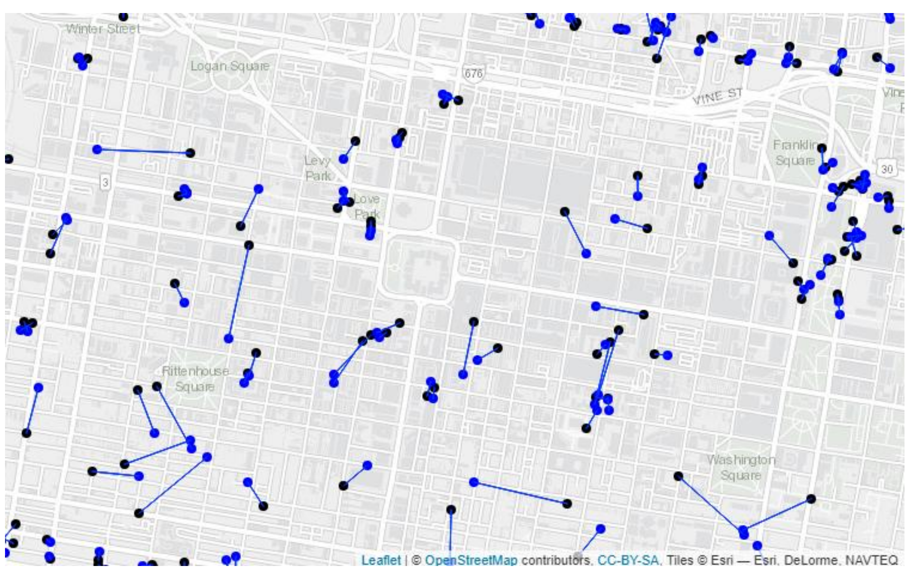

Figure 11. Examples of AADM+AADB matches

Geocoding Crash and Volume Locations

Geocoding is the process of converting descriptions of locations into geographical coordinates (longitude and latitude). Both crash and volume data from Anchorage were geographically identified by intersecting roadways (e.g., Abbott Road and Golovin Street). Other sites provided geographical coordinates. These coordinates are important for counting crashes that occurred near volume count locations.

Geocoding was accomplished using *googleway* (Cooley, 2020). *Googleway* is an add-on for the *R* language for statistical computing (R Core Team, 2019). It uses an application programming interface (API) to pass intersecting street names to Google Maps, which then returns the desired coordinates. Figure 12 provides examples of the resulting geocoded crashes. Note that a small amount of random noise has been added to visually separate overlapping points.

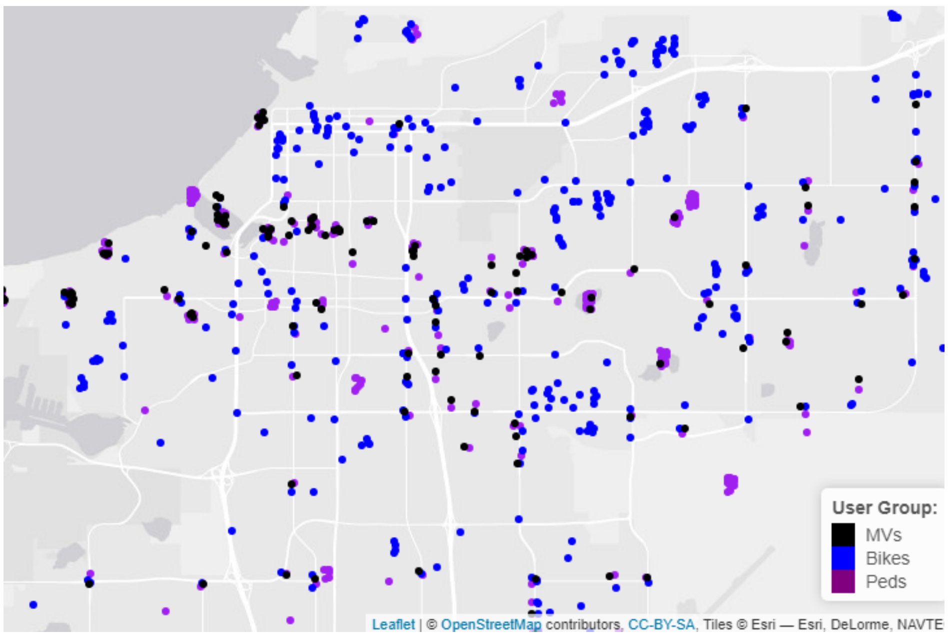

Figure 12. Examples of geocoded crashes in Anchorage

Determining an Appropriate Crash Zone Size

Analyses of the SIN effect and outreach programs require data points with various volumes as well as relevant crashes (those involving bicyclists or pedestrians). However, count locations are very small: on one-way roads they depend on road users to cross a two-dimensional line and counts at intersections require road users to pass through the intersection. Crashes are unlikely to occur within these bounds, but rather in the surrounding areas. Determinations of crash zone radius are inconsistent and vary by location.

Wang et al. (2008) conducted a nationwide survey to review how police officers, crash records technicians, and State safety engineers identify intersection crashes. Although the intent here is to identify an appropriate area in which to associate crashes to volume count locations, this guidance was considered. Authors report that Alaska uses a default radius of 61.0 m (200 ft) while Colorado uses 76.2 m (250 ft) and Pennsylvania has no default, adding that Colorado's radius "represents one-half of a typical urban block, which might help explain why 250 ft is commonly used" (p. 87). Crash zones with radii of 76.2 m (250 ft) were used for Fort Collins and Philadelphia. In the case of Anchorage, however, 76.2 m radii omitted most crashes; ultimately 250 m-radius zones were used to account for the greater degree of geographical dispersion. Figure 13 serves as a scale reference for the bicyclist crash zones used in the analysis of Fort Collins. The map represents the boundaries of 76.2 m-radius crash zones with blue lines and all crashes with red dots.

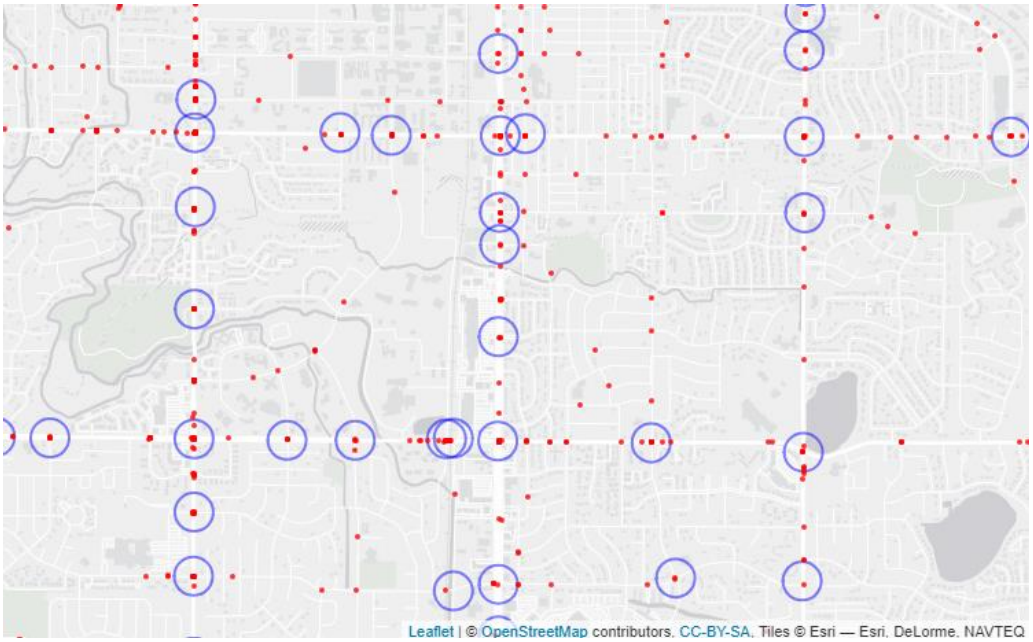

Figure 13. Examples of bicyclist crash zones in Fort Collins

Statistical Analysis

This program evaluation follows an extensive literature review documenting the evolution of the analytical methods used to quantify SIN. The evaluation seeks to quantify a SIN effect as well as any effects on bicyclist and pedestrian volumes attributable to *local program* accomplishments. As the data are focused on local programs and their available data, random assignment was not possible. Rather, pre-post comparisons using archival data sources were used to look for differences and whether the data fit the mathematical definition of SIN. No claims are made about the generalizability of the results to other programs or localities.

Elvik's (2013) research identified flaws in prior methodologies and proposed the model used here to both quantify SIN and examine program effects. The SIN model can be expressed as:

$$\log(C_B) = \beta_0 + \beta_1 \log(AADB) + \beta_2 \log(AADT) + \beta_3(AADB \times AADT) + \sum(\beta_i X_i) \qquad (1)$$

$$\log(C_P) = \beta_0 + \beta_1 \log(AADP) + \beta_2 \log(AADT) + \beta_3(AADP \times AADT) + \sum(\beta_i X_i) \qquad (2)$$

where $\log(\dots)$ denotes a logarithm, C_B and C_P represent the number of crashes involving bicyclists and pedestrians, respectively; AADB, AADP and AADM represent the annual average daily volume of bicyclists, pedestrians, and (motorized) traffic, respectively; X_i represent other known factors that may influence safety; and all βs are coefficients estimated via negative binomial regression. The non-log-transformed AAD volume interaction terms $(AADB \times AADT)$

and ($AADP \times AADT$) were found by Elvik, Sørensen, and Nævestad (2013) to improve the model fit and were thus included in this analysis.

Two crash metrics were used in analysis: the total number of crashes involving each road user group, and a severity-adjusted total. The former treats crashes of all severities equally, while the latter assigns values of 1.0, 0.8, 0.6, 0.4 and 0.2 to fatal, incapacitating, non-incapacitating, possible injury and property-damage-only crashes, respectively.

By using the log-transformed AAD volumes, coefficients β_1 and β_2 yield elasticities, or the percentage change in the number of accidents associated with a 1% increase in each AAD volume. For example, if Equation (1) produces $\beta_1 = 0.55$, one would expect a 0.55% increase in bicyclist crashes if AADB increases by 1%.

Values for β_1 and β_2 determine the presence of SIN. Both coefficients are considered simultaneously because the two modes are often correlated. If they sum to exactly 1.0, then bicyclist/pedestrian risk remains constant amid increases in both bicyclist/pedestrian and motor vehicle volumes. A sum greater than 1.0 indicates a *hazard* in numbers and a sum less than 1.0 indicates SIN. If the sum exceeds 1.0 while either coefficient is less than 1.0 the data are said to exhibit a *partial* SIN, where SIN occurs if *only* one mode's volume increases (Elvik, 2013).

Equations (1) and (2) were adapted to model program effects by changing the roles of some variables and adding program metrics ($PROG$) as independent variables:

$$\log(AADB) = \beta_0 + \beta_1 \log(AADP) + \beta_2 \log(AADT) + \beta_3 C_B + \beta_4 C_P + \beta_5 \log(PROG) \quad (3)$$

$$\log(AADP) = \beta_0 + \beta_1 \log(AADB) + \beta_2 \log(AADT) + \beta_3 C_B + \beta_4 C_P + \beta_5 \log(PROG) \quad (4)$$

These equations include terms for the number of crashes involving bicyclists and pedestrians (C_B and C_P respectively) as they were believed to potentially affect bicyclist and pedestrian volumes as well. All variables are expressed as annual quantities.

Results

Several variations of the models described in the *Statistical Analysis* section were fit to available data to investigate the effects of outreach programs and quantify SIN. Poisson, negative binomial, and zero-inflated models were estimated when appropriate, and compared to identify the ideal models for each site. Initial models included all available variables and were then narrowed down using a stepwise selection process. All analysis and data manipulation were conducted using the *R* language for statistical computing (R Core Team, 2019) version 3.6.1. Various functions included in the *MASS* package were employed for modeling (Venables & Ripley, 2002).

Fort Collins, Colorado

Four programs in Fort Collins were examined for their effects on bicyclist and pedestrian volumes: Safe Routes to School, Open Streets, the Bicycle Ambassador Program, and Bike to Work Day. Seven program metrics were derived from various data sources, described below. Programs were assumed to be non-existent beyond the data provided, and any associated metrics were given a value of zero for such times.

Figure 14 shows the metrics related to the Safe Routes to School program in Fort Collins. Note that the number of students *reached* was highly correlated with the number of students *educated* (Pearson correlation of 0.95) and was thus excluded from further analysis.

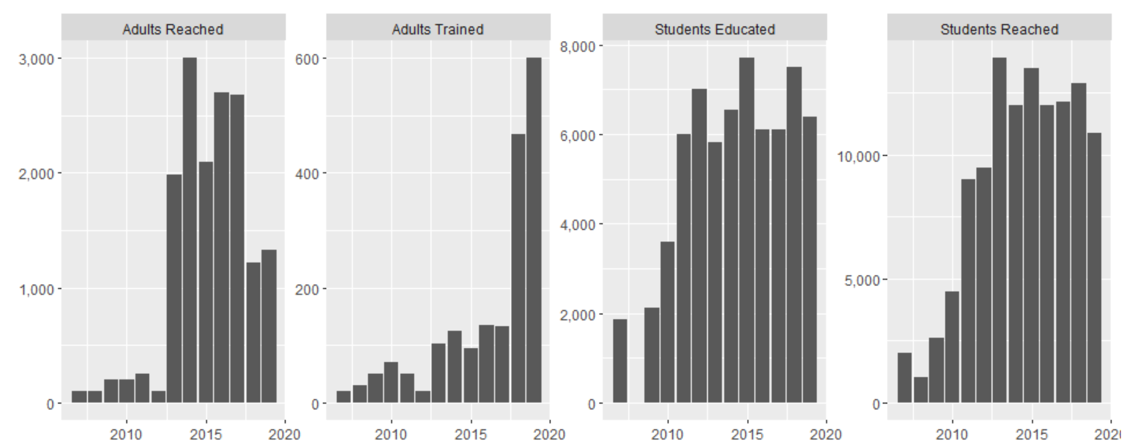

Figure 14. Program metrics, Safe Routes to School, Fort Collins

Figure 15 shows annual participants in the Open Streets events in Fort Collins.

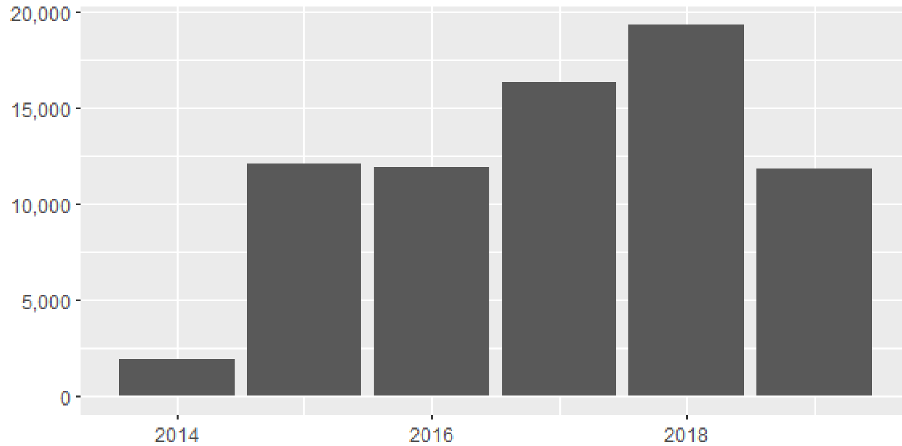

Figure 15. Program metrics, Open Streets participants, Fort Collins

Figure 16 shows the total number of Bicycle Ambassador Program event attendees in Fort Collins. The BAP holds many kinds of events targeted toward different audiences. Bicycle maintenance classes and rides target bicyclists, while some educational classes target drivers. Other events such as board meetings and webinars target the general public. Figure 16 distinguishes between target audiences, but the total annual attendance was used in analysis.

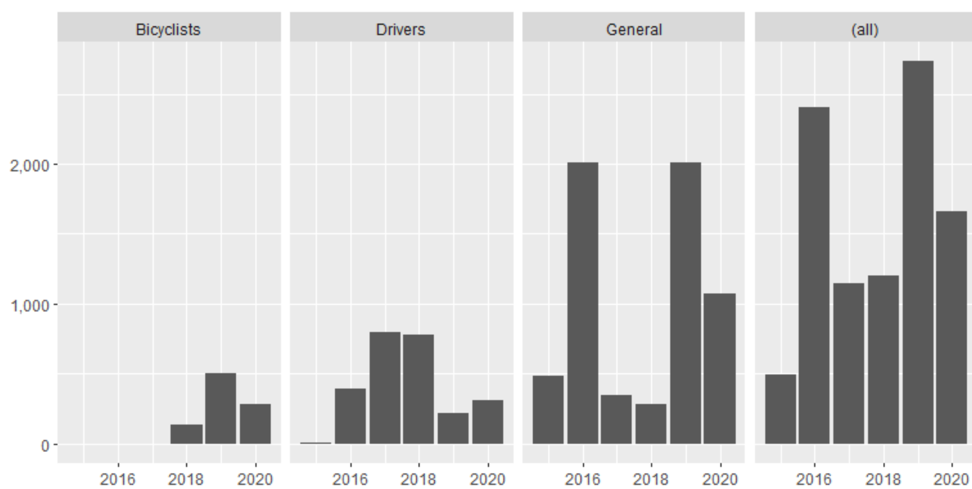

Figure 16. Program metrics, Bicycle Ambassador Program event attendees, Fort Collins

Figure 17 shows the annual number of participants in Bike to Work Day in Fort Collins, in both summer and winter. Participation in both seasons exhibits an upward trend. Only the summer participation was used in analysis as it was believed to be more representative of program attendance and including both seasons could double-count bicyclists who participated in both seasons.

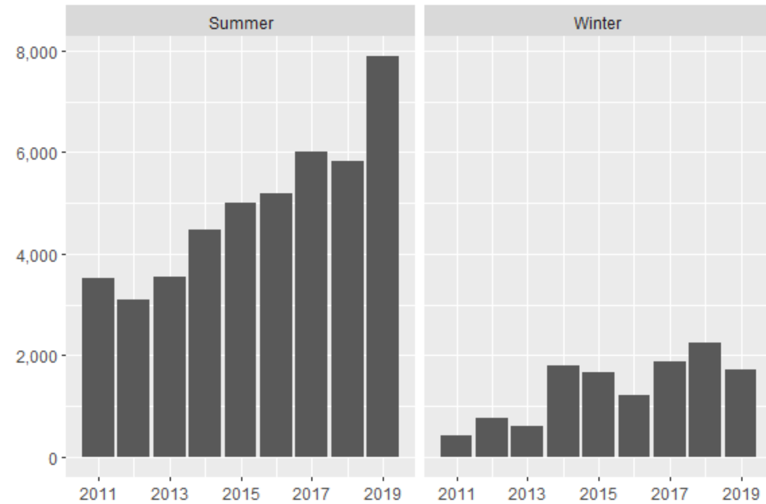

Figure 17. Program metrics, Bike to Work Day participants, Fort Collins

All volumes for Fort Collins were extracted from publicly available turning movement study reports (City of Fort Collins, 2020) and converted to annual AAD volumes using the method described in *Data Preparation: Converting Short-Term Counts to AAD Volumes.* Figure 18 provides histograms of AAD volumes for each road user group. Note that these histograms include all years and locations for which data were available.

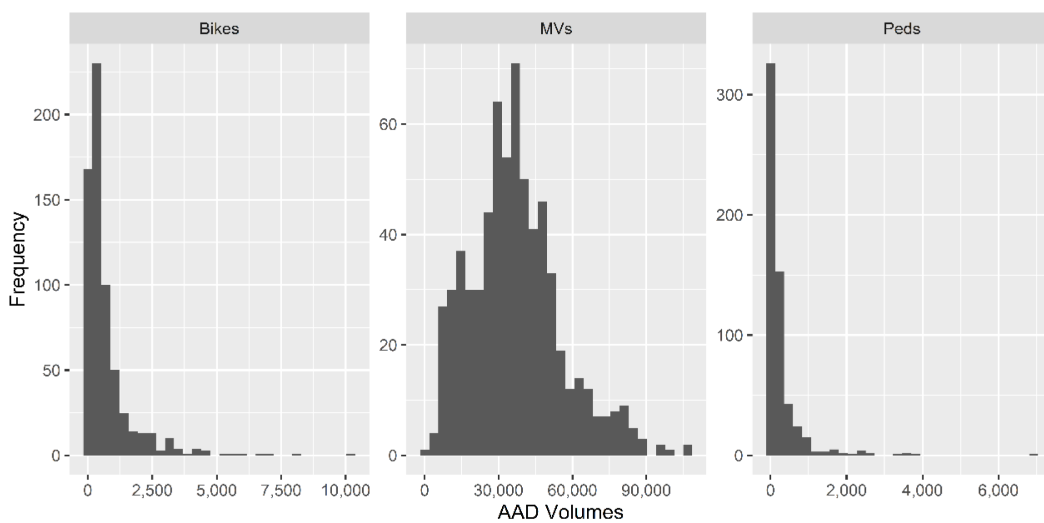

Figure 18. Histogram of AAD volumes by road user group, Fort Collins

Crash data for Fort Collins were provided by the city's Traffic Operations department. Figure 19 shows annual crashes by severity and the non-motorized road user group. Although 2017 appears to be an incomplete year, one crash record is dated December 5, 2017, suggesting completeness.

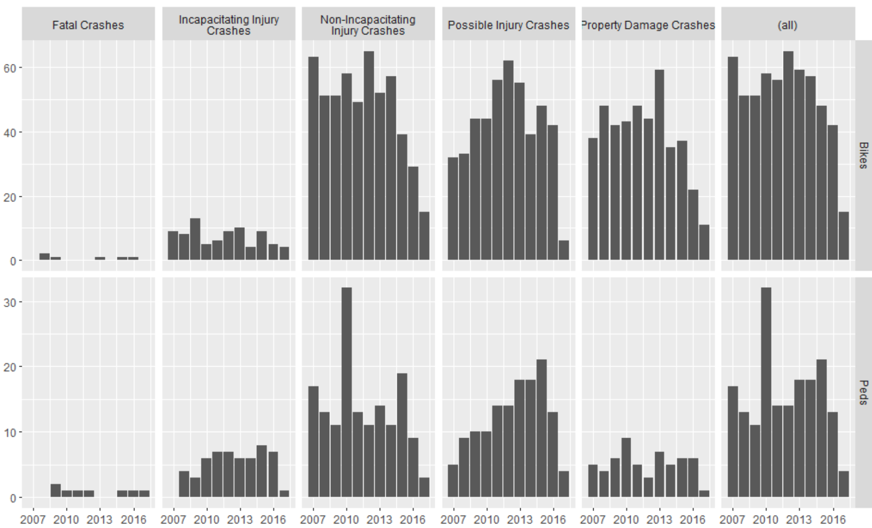

Figure 19. Annual crashes by severity and affected non-motorized road user group, Fort Collins

Program Effectiveness

Program impacts on bicyclist and pedestrian volumes were evaluated using the models described in the *Statistical Analysis* section.

Table 5 provides summary statistics for all variables included in initial models of program effectiveness. The analysis data set consisted of 571 observations between 2009 and 2017, with average AADB=786, AADM=36,894, and AADP=277. The numbers of crashes within each crash zone ranged from 0 to 4 for pedestrians (mean=0.1) and 0 to 8 for bicyclists (mean=0.5).

Table 5. Summary statistics, Fort Collins

Statistic	*N*	*Mean*	*St. Dev.*	*Min*	*Pctl(25)*	*Pctl(75)*	*Max*
AADB	571	786.3	1,092.6	10	211.4	859.2	10,220
AADM	571	36,893.7	18,829.2	840.0	24,536.4	46,419.5	107,675.4
AADP	571	277.0	549.1	6.6	44.6	256.4	6,933.6
Ped. Crashes (sev.-adj.)	571	0.1	0.2	0	0	0	2
Ped. Crashes	571	0.1	0.4	0	0	0	4
Bic. Crashes (sev.-adj.)	571	0.2	0.4	0.0	0.0	0.4	3.4
Bic. Crashes	571	0.5	0.9	0	0	1	8
Students Reached (SRTS)	571	10,062.7	3,725.6	2,600	9,000	12,129	13,907
Students Educated (SRTS)	571	5,652.4	1,553.7	2,121	5,828	6,544	7,700
Adults Reached (SRTS)	571	1,551.7	1,169.5	100	200	2,679	3,000
Adults Trained (SRTS)	571	90.4	36.1	20	50	124	134
Participants (OS)	571	4,691.7	6,334.2	0	0	11,888	16,312
Attendees (BAP)	571	433.1	763.2	0	0	492	2,404
Participants (BTWD)	571	3,457.8	2,035.0	0	3,082	4,995	6,009

Notes:
Ped. = pedestrian, Bic. = bicyclist, sev.-adj. = severity-adjusted
N = observation count, St. Dev. = standard deviation, Min = minimum, Pctl(25) = 25th percentile, Pctl(75) = 75th percentile, Max = maximum

Table 6 presents the program effectiveness model results for Fort Collins. Bicyclist and pedestrian volumes were modelled separately, each with full and reduced models (reduced via stepwise selection using the Akaike information criterion [AIC]). Both Poisson and negative binomial models were estimated, but the AIC indicated that the Poisson distribution was inappropriate; therefore, only the negative binomial model results are provided. The value of each coefficient is provided, with statistical significance indicated by asterisks (* indicates $p<0.10$, ** $p<0.05$, *** $p<0.01$) and 95% confidence intervals in parentheses.

Table 6. Program effectiveness model results, Fort Collins

	AADB		**AADP**	
	Full	**Stepwise**	**Full**	**Stepwise**
Constant	1.745	3.469***	14.017***	12.762***
	(-2.611, 6.101)	(1.929, 5.010)	(9.478, 18.557)	(8.588, 16.936)
log(AADP)	0.525***	0.543***		
	(0.466, 0.585)	(0.487, 0.600)		
log(AADB)			0.599***	0.606***
			(0.534, 0.665)	(0.540, 0.671)
log(AADM)	0.078	0.098*	-0.690***	-0.689***
	(-0.042, 0.199)	(-0.018, 0.215)	(-0.810, -0.570)	(-0.809, -0.568)
Ped. Crashes (sev.-adj.)	0.208		0.658**	0.374***
	(-0.335, 0.751)		(0.083, 1.233)	(0.206, 0.543)
Ped. Crashes	0.086		0.201	0.112***
	(-0.150, 0.322)		(-0.049, 0.451)	(0.028, 0.195)
Bic. Crashes (sev.-adj.)	-0.264		-0.613	
	(-1.314, 0.785)		(-1.728, 0.501)	
Bic. Crashes	-0.086		-0.191	
	(-0.605, 0.434)		(-0.742, 0.361)	
log(Students Educated (SRTS))	0.242		-0.691***	-0.557**
	(-0.252, 0.736)		(-1.216, -0.166)	(-1.050, -0.064)
log(Adults Reached (SRTS))	0.176	0.104	-0.105	
	(-0.084, 0.435)	(-0.024, 0.232)	(-0.381, 0.172)	
log(Adults Trained (SRTS))	-0.377	-0.290*	0.102	
	(-0.869, 0.114)	(-0.599, 0.020)	(-0.422, 0.626)	
log(Participants (OS))	-0.006		0.041*	
	(-0.046, 0.033)		(-0.001, 0.083)	
log(Attendees (BAP))	-0.025	-0.032**	0.022	0.055***
	(-0.065, 0.016)	(-0.057, -0.007)	(-0.021, 0.065)	(0.031, 0.079)
log(Participants (BTWD))	-0.030		0.083**	0.066**
	(-0.097, 0.037)		(0.012, 0.154)	(0.014, 0.118)
Observations	571	571	571	571

	AADB		AADP	
	Full	**Stepwise**	**Full**	**Stepwise**
Log Likelihood	-4,208.771	-4,210.726	-3,509.385	-3,511.513
Theta	1.501*** (0.081)	1.492*** (0.081)	1.335*** (0.072)	1.327*** (0.072)
AIC	8,443.542	8,433.451	7,044.771	7,039.026

Notes:
Coefficients and 95% confidence intervals shown.
*p<0.10; **p<0.05; ***p<0.01

Results indicate that bicyclist volumes are positively correlated with pedestrian and motor vehicle volumes, but negatively correlated with some program metrics. The latter relationship may not be accurate and may instead reflect a geographical and/or temporal disconnect between the program metrics and volume counts. Again, while the relationships could be spurious, it is possible that outreach programs exerted a positive impact on volumes that were either not picked up by the turning movement studies or were washed out during the AAD volume estimation process. It may also be the case that programs increased in activity in response to lower volumes.

Model results also indicated several unexpected relationships concerning pedestrians. Pedestrian volume was found to be positively correlated with bicyclist volumes, pedestrian crashes, and two bicyclist-oriented program metrics; and negatively correlated with motor vehicle volumes and one SRTS metric. Again, it seems highly unlikely that outreach programs would cause decreases in pedestrian volumes; rather they are simply correlated. The inverse relationship between AADP and AADM is indicative of mode shifts.

Safety in Numbers

SIN models were generated separately for bicyclists and pedestrians using the models described in the *Statistical Analysis* section. The program metrics acquired for Fort Collins describe various counts (attendees, participants, etc.). Missing annual counts were therefore assumed to be zero. For example, the number of Bike to Work Day participants begins in 2011, so a zero was used for all preceding years. The same data set used to evaluate the effectiveness of active transportation programs – or those programs supporting walking and bicycling – and was therefore used to investigate SIN. See Table 5 for summary statistics.

Table 7 presents the SIN model results for Fort Collins. Bicyclist and pedestrian crashes were modelled separately, each with two negative binomial models: one with and one without the product of AADM and the appropriate non-motorized AAD volume. Note that these values are very large and were not log-transformed in accordance with Elvik, Sørensen, and Nævestad (2013). Both severity-adjusted and unadjusted crash counts were estimated, but the AIC indicated that the unadjusted counts produced better model fit; therefore, only unadjusted crash count model results are provided. The value of each coefficient is provided, with statistical significance indicated by asterisks (* indicates p<0.10, ** p<0.05, *** p<0.01) and 95% confidence intervals in parentheses.

Table 7. Model results, SIN, Fort Collins

	Bicyclist Crashes		Pedestrian Crashes	
	Simple	**Full**	**Simple**	**Full**
Constant	-9.642***	-8.495***	-11.720***	-11.809***
	(-12.711, -6.574)	(-12.213, -4.778)	(-17.116, -6.323)	(-18.141, -5.477)
log(AADB)	0.314***	0.231**		
	(0.181, 0.446)	(0.029, 0.434)		
log(AADP)			0.469***	0.476***
			(0.266, 0.672)	(0.140, 0.812)
log(AADM)	0.665***	0.597***	0.711***	0.717***
	(0.396, 0.934)	(0.300, 0.894)	(0.244, 1.179)	(0.205, 1.229)
AADM x AADB		0.000		
		(-0.000, 0.000)		
AADM x AADP				-0.000
				(-0.000, 0.000)
Observations	571	571	571	571
Log Likelihood	-511.801	-511.218	-234.957	-234.955
Theta	1.199*** (0.299)	1.235*** (0.315)	0.782* (0.404)	0.783* (0.405)
AIC	1,029.602	1,030.437	475.913	477.911

Notes:
Coefficients and 95% confidence intervals shown.
*p<0.10; **p<0.05; ***p<0.01

Results indicate SIN for bicyclists, but only partial SIN for pedestrians. The full bicyclist crash model produced coefficients of 0.231 and 0.597 for $\log(AADB)$ and $\log(AADM)$, respectively, which sum to 0.828, indicative of complete SIN. These coefficients suggest that a 1% increase in AADB is associated with a 0.231% increase in bicyclist crashes, that a 1% increase in AADM is associated with a 0.597% increase in bicyclist crashes, and that a simultaneous 1% increase in *both* modes is associated with a 0.828% increase in bicyclist crashes. The $AADB \times AADM$ interaction was not statistically significant, but its inclusion decreased the AAD volume coefficients relative to the simple model.

For pedestrians, model results indicate partial SIN. The full pedestrian crash model produced coefficients of 0.476 and 0.717 for $\log(AADP)$ and $\log(AADM)$, respectively, which sum to 1.193. These coefficients suggest that a 1% increase in AADP is associated with a 0.476% increase in pedestrian crashes, that a 1% increase in AADM is associated with a 0.717% increase in pedestrian crashes, and a simultaneous increase in both modes is associated with a 1.193% increase in pedestrian crashes. The $AADP \times AADM$ interaction was not statistically significant, and its inclusion caused minimal changes to AAD volume coefficients relative to the simple model. As with bicyclists, the weighted crash counts produced similar results.

Philadelphia, Pennsylvania

The Indego Bikeshare program was examined in Philadelphia. Publicly available trip summaries were analyzed to determine the annual number of bikeshare stations and trips. These metrics served as broad measures of the bikeshare program's presence in the city over time. For more localized measures, the annual number of trips originating from each station was also calculated, shown in Figure 20. A small amount of random noise has been added to this map to visually separate multiple years of data at one station. As expected, more trips originated closer to the city center than the outskirts.

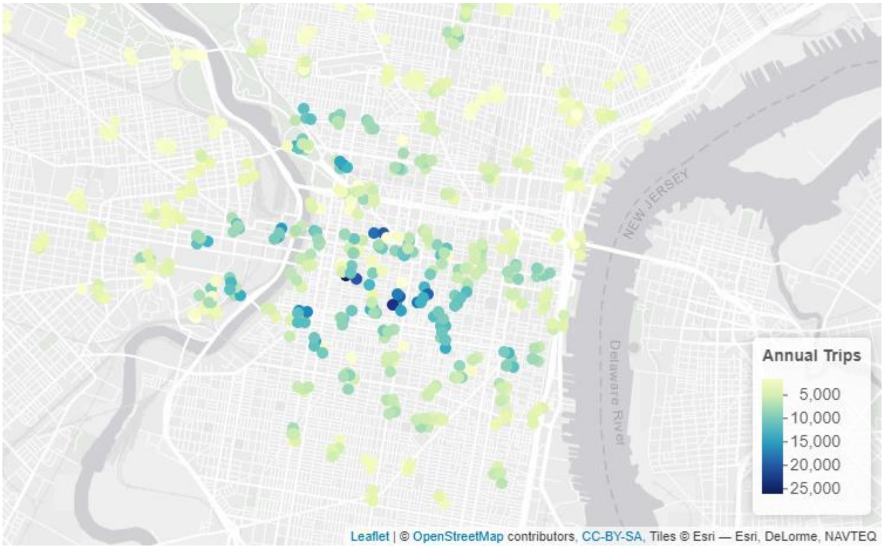

Figure 20. Map of annual Indego Bikeshare trips by origin, Philadelphia

All volumes for Philadelphia were provided by DVRPC, the federally designated MPO in Philadelphia. These volumes required extensive data manipulation, as described in *Data Preparation: Connecting and Interpolating Data From Single-Mode Counters.*

Figure 21 provides histograms of AAD volumes for each road user group. Note that these histograms include all years and locations for which data were available.

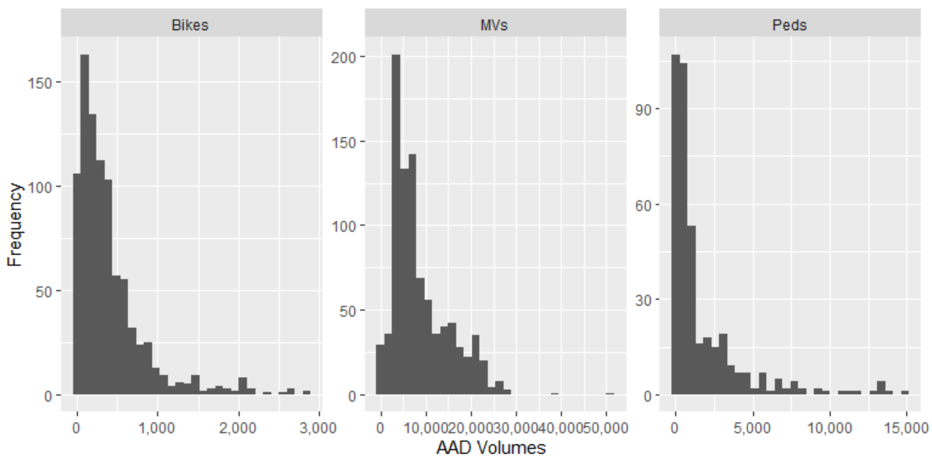

Figure 21. Histogram of AAD volumes by road user group, Philadelphia

Many of Philadelphia's bicyclist volume counts were accompanied by additional variables describing bicycle facilities present at the location (bike lane, multi-use trail, various treatments that remove bicyclists from motorized traffic, "sharrows," and striped shoulder) and a location type (low volume, mixed, and recreation). This information was included in analyses of AADB and bicyclist crashes, but not sufficiently present to use in pedestrian analyses.

Crash data for Philadelphia were sourced from the Pennsylvania DOT's Crash Information Tool (Pennsylvania DOT, n.d.). Figure 22 shows annual crashes by severity and the non-motorized road user group.

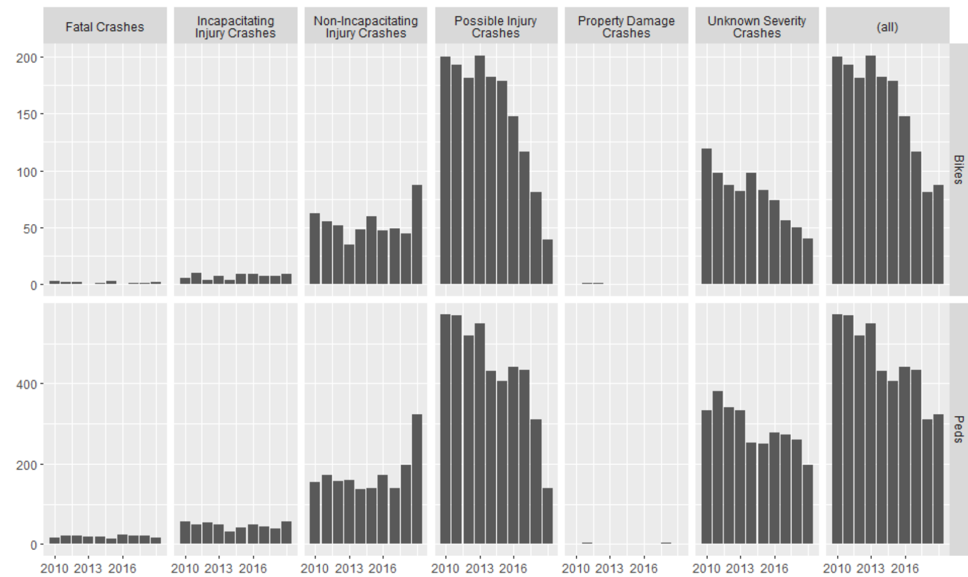

Figure 22. Annual crashes by severity and affected non-motorized road user group, Philadelphia

Program Effectiveness

Program impacts on bicyclist and pedestrian volumes were evaluated using the models described in the *Statistical Analysis* section. Because bicyclist and pedestrian volumes were rarely co-located, the two road user groups were analyzed separately.

Table 8 provides summary statistics for all variables included in initial models of program effectiveness regarding bicyclists. The analysis data set consisted of 736 observations between 2010 and 2019, with average AADB=429 and AADM=7,707. The number of crashes within each crash zone ranged from 0 to 2 (mean=0.1). The average distance between connected volume locations was 81m (see *Data Preparation: Connecting and Interpolating Data From Single-Mode Counters*). Ninety percent of the observations were in mixed traffic, 20% with bike lanes, and 20% at low volume locations.

Table 8. Summary statistics for program effectiveness, bicyclists, Philadelphia

Statistic	***N***	***Mean***	***St. Dev.***	***Min***	***Pctl(25)***	***Pctl(75)***	***Max***
AADB	736	428.6	449.0	6	134.5	544	2,850
AADM	736	7,707.4	5,443.0	59	3,995.1	9,720	38,226
Indego Trips (localized)	736	1,733.5	3,621.6	0	0	0	19,090
Ped. Crashes (sev.-adj.)	736	0.1	0.2	0	0	0	1
Ped. Crashes	736	0.2	0.4	0	0	0	3
Bic. Crashes (sev.-adj.)	736	0.1	0.2	0	0	0	1
Bic. Crashes	736	0.1	0.4	0	0	0	2
Indego Trips (global)	736	241,781.3	309,271.1	0	0	650,239	782,556
Indego Stations	736	58.6	80.3	0	0	139	210
Bike Lane Present	736	0.2	0.4	0	0	0	1
Mixed Traffic	736	0.9	0.3	0	1	1	1
Trail Location	736	0.1	0.2	0	0	0	1
Physical Separation from Motorized Traffic	736	0.1	0.2	0	0	0	1
Pavement Markings Present	736	0.1	0.3	0	0	0	1
Low Volume Corridor	736	0.2	0.4	0	0	0	1
Recreation Location	736	0.1	0.3	0	0	0	1

Notes:
Ped. = pedestrian, Bic. = bicyclist, sev.-adj. = severity-adjusted
N = observation count, St. Dev. = standard deviation, Min = minimum, Pctl(25) = 25th percentile, Pctl(75) = 75th percentile, Max = maximum

Table 9 provides summary statistics for all variables included in initial models of program effectiveness regarding pedestrians. The analysis data set consisted of 284 observations between 2010 and 2019, with average AADP=2,136 and AADM=11,765. The number of crashes within each crash zone ranged from 0 to 9 (mean=0.3). The average distance between connected volume locations was 137 m (see *Data Preparation: Connecting and Interpolating Data From Single-Mode Counters* for more information).

Table 9. Summary statistics for program effectiveness, pedestrians, Philadelphia

Statistic	***N***	***Mean***	***St. Dev.***	***Min***	***Pctl(25)***	***Pctl(75)***	***Max***
AADM	284	11,765.0	7,069.5	114	6,540.1	15,704	50,888
AADP	284	2,135.8	2,875.3	0	446.0	2,838	14,920
Indego Trips (localized)	284	735.5	2,347.4	0	0	0	13,322
Ped. Crashes (sev.-adj.)	284	0.1	0.4	0.0	0.0	0.0	4.5
Ped. Crashes	284	0.3	0.8	0	0	0	9
Bic. Crashes (sev.-adj.)	284	0.1	0.2	0.0	0.0	0.0	1.1
Bic. Crashes	284	0.1	0.3	0	0	0	2
Indego Trips (global)	284	233,205.0	309,438.6	0	0	650,239	782,556
Indego Stations	284	58.2	82.3	0	0	139	210

Bicyclist and pedestrian volumes were modelled separately, using Poisson and negative binomial response distributions, with and without statistical weights derived from the distance between volume counting locations. The AIC statistic was used to compare different model specifications as well as weighting schemes (Ingdal et al., 2019), and indicated that negative binomial, distance-weighted models were superior.

Table 10 provides the full and reduced (via stepwise selection) negative binomial, distance-weighted models. The value of each coefficient is provided, with statistical significance indicated by asterisks (* indicates $p<0.10$, ** $p<0.05$, *** $p<0.01$) and 95% confidence intervals in parentheses.

The reduced model results indicate that bicyclist and pedestrian volumes are positively associated with motor vehicle volumes: a 1% increase in AADM is associated with a 0.234% increase in AADB and a 0.405% increase in AADP. Volumes are generally correlated, so this relationship is to be expected. Crashes involving bicyclists were also statistically significant, though interestingly severity-adjusted crashes were associated with AADB while unadjusted crashes were associated with relatively larger increases in AADP. This could suggest that bicyclists are able to perceive the severity of crashes involving other bicyclists and continue to increase in number, while pedestrians perceive all crashes involving bicyclists as deterrents, and respond by increasing in number. However, a more likely explanation is that the two data sets simply include different locations.

Two of the three program metrics were found to be statistically significantly associated with AADB. Increases in Indego stations were negatively correlated with AADB, but this is likely due to a disconnect between localized volumes and city-wide stations. Notably, the localized number of Indego trips was positively associated with AADB. This suggests that the program successfully increased bicyclist volumes in Philadelphia. Further, the lack of significance between Indego program metrics and AADP suggests that the program *created* new bicyclists rather than converting them from pedestrians.

All indicator variables in Table 10 use the absence of the listed characteristic as the reference. Thus, these coefficients can be interpreted as follows: bike lanes are associated with $\exp(0.276) = 1.3$ times more AADB than corridors without bike lanes. Bike lanes, mixed traffic, trails, and physical separations (buffered and protected bike lanes) were associated with increased AADB, while low volume corridors and recreation areas were associated with decreased AADB.

Table 10. Model results, program effectiveness, Philadelphia

	AADB		AADP	
	Full	**Reduced**	**Full**	**Reduced**
Constant	3.602***	3.549***	4.093***	3.752***
	(2.658, 4.547)	(2.614, 4.484)	(1.566, 6.621)	(1.215, 6.290)
log(AADM)	0.229***	0.234***	0.378***	0.405***
	(0.132, 0.326)	(0.138, 0.330)	(0.106, 0.651)	(0.131, 0.679)
log(Bic. Crashes, sev.-adj.)	1.198	0.514**	0.728	
	(-1.284, 3.680)	(0.016, 1.012)	(-8.508, 9.964)	
log(Bic. Crashes)	-0.432		0.522	1.088**
	(-1.973, 1.109)		(-5.298, 6.342)	(0.179, 1.997)
log(Ped. Crashes, sev.-adj.)	0.620		0.638	
	(-1.468, 2.709)		(-3.566, 4.843)	
log(Ped. Crashes)	-0.357		-0.268	
	(-1.655, 0.942)		(-2.982, 2.447)	
log(Indego Trips, global)	0.053		-0.034	
	(-0.066, 0.172)		(-0.428, 0.359)	
log(Indego Stations)	-0.210	-0.069***	0.031	
	(-0.528, 0.109)	(-0.111, -0.026)	(-1.015, 1.077)	
log(Indego Trips, localized)	0.037***	0.038***		
	(0.009, 0.065)	(0.010, 0.066)		
Bike Lane Present	0.275**	0.276**		
	(0.040, 0.510)	(0.046, 0.507)		
Mixed Traffic	0.398*	0.410*		
	(-0.042, 0.838)	(-0.023, 0.843)		
Trail Location	1.383***	1.385***		
	(0.840, 1.927)	(0.845, 1.926)		
Physical Separation from Motorized Traffic	0.400**	0.392**		
	(0.048, 0.751)	(0.044, 0.740)		
Pavement Markings Present	-0.013			
	(-0.364, 0.338)			
Low Volume Corridor	-0.952***	-0.935***		
	(-1.161, -0.742)	(-1.143, -0.727)		
Recreation Location	-0.666***	-0.650***		
	(-1.083, -0.249)	(-1.063, -0.236)		
Observations	736	736	284	284
Log Likelihood	-2,498.461	-2,499.243	-752.366	-753.354
Theta	2.050*** (0.141)	2.042*** (0.141)	0.891*** (0.119)	0.876*** (0.117)
AIC	5,028.922	5,020.486	1,520.732	1,512.708

Notes:
Coefficients and 95% confidence intervals shown.
**p*<0.10; ***p*<0.05; ****p*<0.01

Safety in Numbers

SIN models were generated separately for bicyclists and pedestrians using the models described in the *Statistical Analysis* section. The program metrics acquired for Philadelphia describe various counts (stations and trips). Missing annual counts were therefore assumed to be zero. The same data sets used to evaluate the effectiveness of the Indego bikeshare program were therefore used to investigate SIN. See Table 8 for summary statistics on bicyclists and Table 9 for summary statistics on pedestrians. Note that the "trail location" indicator variable was omitted from SIN models due to zero crashes being recorded in such locations.

Remarkably few crashes occurred within the 76.2 m-radius crash zones: 90% of locations experienced zero bicyclist crashes, and 86% experienced zero pedestrian crashes. In an attempt to increase the number of crashes analyzed, the crash radius was doubled, but it did not produce meaningfully different results. The original 76.2 m-radius crash zones were used in the models discussed below. Zero-inflated models were also employed to account for this dispersion.

Bicyclist and pedestrian volumes were modelled separately, using Poisson and negative binomial response distributions, with and without zero-inflation, and with and without statistical weights derived from the distance between volume counting locations. Similar models were also estimated for severity-weighted crash counts. AIC statistics indicated that Poisson, distance-weighted models of unadjusted crash counts were superior. Table 11 provides the full and reduced (via stepwise selection) Poisson, distance-weighted SIN models for Philadelphia. The value of each coefficient is provided, with statistical significance indicated by asterisks (* indicates $p<0.10$, ** $p<0.05$, *** $p<0.01$) and 95% confidence intervals in parentheses.

Results did not indicate safety, nor hazard, in numbers for bicyclists or pedestrians. The full bicyclist crash model yielded negative coefficients for log(AADB) and log(AADM). These results, although not statistically significantly different from zero, would indicate *decreases* in crashes associated with increases in volumes. The reduced model omitted both terms, keeping only the presence of a bike lane, and the product of AADM and AADB. Locations with bike lanes were associated with approximately one-third as many bicyclist crashes compared to locations without ($e^{-1.020} = 0.36$). Notably, none of the Indego bikeshare program metrics were found to be statistically significantly associated with bicyclist crashes. Based on the variables used in the model, this indicates that the bikeshare program neither positively nor negatively affected safety in Philadelphia. Note that additional information such as infrastructure features over time and the actual paths taken by bikeshare users may shed new light on the process by which the program influenced volumes.

The full pedestrian crash model also produced unlikely results: an extreme, complete SIN effect due to the low sum of the coefficients on log(AADP) and log(AADM). These terms were also omitted in the reduced model.

Table 11. Model results, SIN, Philadelphia

	Bicyclist Crashes		Pedestrian Crashes	
	Full	**Reduced**	**Full**	**Reduced**
Constant	-1.217	-2.177***	-2.437	-1.654***
	(-6.266, 3.833)	(-2.511, -1.843)	(-8.615, 3.742)	(-2.193, -1.115)
log(AADB)	-0.134			
	(-0.550, 0.282)			
log(AADP)			0.116	
			(-0.325, 0.556)	
log(AADM)	-0.185		0.004	
	(-0.551, 0.180)		(-0.593, 0.600)	
Indego Trips (global)	-0.205			
	(-0.717, 0.307)			
Indego Trips (localized)	-0.061			
	(-0.177, 0.055)			
Indego Stations	0.600			
	(-0.757, 1.958)			
Bike Lane Present	-1.039	-1.020*		
	(-2.333, 0.255)	(-2.187, 0.148)		
Mixed Traffic	1.435			
	(-1.987, 4.857)			
Physical Separation from Roadway	-1.056			
	(-3.161, 1.049)			
Pavement Markings Present	-0.245			
	(-1.716, 1.227)			
Low Volume Corridor	-0.286			
	(-1.330, 0.758)			
Recreation Location	-0.250			
	(-2.359, 1.860)			
AADM x AADB	0.000***	0.000**		
	(0.000, 0.000)	(0.000, 0.000)		
AADM x AADP			0.000	0.000***
			(-0.000, 0.000)	(0.000, 0.000)
Observations	736	736	284	284
Log Likelihood	-135.531	-139.268	-57.413	-57.557
AIC	297.063	284.536	122.825	119.115

Notes:
Coefficients and 95% confidence intervals shown.
*p<0.10; **p<0.05; ***p<0.01

The Role of Infrastructure

The presence of 17 pedestrian/bicyclist facilities was added to the statistical models described above to investigate the role of infrastructure. Because bicyclist and pedestrian volumes were rarely co-located, the two road user groups were analyzed separately. Table 12 lists the facilities and their prevalence among complete observations for each road user group. Recall that complete observations are those with at least two volumes (AADM and AADB or AADP), crashes, and program metrics for a given location and year.

Table 12. Percentage of complete observations with each infrastructure element

Infrastructure Element	***Bicyclists (N=695)***	***Pedestrians (N=257)***
Bike box	0.4	0.0
Bike lane, buffered	8.9	1.6
Bike lane, grade-separated	1.7	4.7
Bike lane, protected	0.3	0.0
Bike lane, standard	28.8	37.7
Bikeshare station	2.0	1.9
Bump out	4.3	5.8
Crosswalk, high-visibility	2.5	45.1
Crosswalk, standard	4.5	09.3
Median refuge	4.0	12.5
Pedestrian hybrid beacon	1.0	1.2
Pedestrian signal	22.2	30.0
Sidewalk, buffered	7.6	16.0
Sidewalk, standard	8.0	96.9
Signalized intersection	3.0	41.2
Stop sign	4.5	14.0
Street parking	5.8	51.0

The new models described below cannot address whether the installation of infrastructure elements is responsible for changes in volumes or crashes. Such conclusions would require numerous observations per location; 75% of the locations available for analysis had 4 or fewer annual observations. Rather, these models attempt to explain some of the variation in volumes and crashes that was previously more coarsely quantified with other model terms.

As before, models were fit using all of the new variables (the “full” models) and then simplified using stepwise selection (the “reduced” models). Table 13 and Table 14 provide succinct comparisons of the changes in the reduced models resulting from the addition of new infrastructure data, and Table 15 summarizes the findings related to each infrastructure element.

Table 13. Comparison of model results after adding infrastructure data, program effectiveness, Philadelphia

	AADB		**AADP**	
	Without Infrastructure	**With Infrastructure**	**Without Infrastructure**	**With Infrastructure**
Constant	3.549***	0.762	3.752***	4.948***
log(AADM)	0.234***	0.435***	0.405***	0.324**
log(Bic. Crashes, sev.-adj.)	0.514**			
log(Bic. Crashes)			1.088**	1.406***
log(Indego Stations)	-0.069***	-0.040*		
log(Indego Trips, localized)	0.038***	0.048***		

Notes:
Coefficients shown (95% confidence intervals omitted). Only terms appearing in the models without infrastructure data are shown in order to compare the differences in results. Findings associated with new infrastructure elements are provided in a separate table.
*p<0.10; **p<0.05; ***p<0.01
Bic. = bicyclist, sev.-adj. = severity-adjusted

Adding infrastructure elements to the AADB model results in a shift of statistical significance from the constant term to several infrastructure elements (see Table 15). The constant term can generally be thought of as an overall average, with other model terms indicating increases or decreases from this baseline. This shift suggests that the new model has decomposed what the previous model lumped into an average (the constant). The new model thus produces more precise estimates of AADB.

The coefficient for log(AADM) increased for AADB and decreased for AADP. This indicates that, after adjusting for infrastructure elements, a 1% increase in AADM is associated with a 0.435% increase (previously 0.234%) in AADB and a 0.324% increase (previously 0.405%) in AADP.

The effect of the number of severity-adjusted bicyclist crashes on AADB became insignificant. This suggests that infrastructure elements are now explaining what was previously attributed to bicyclist crashes. On the contrary, the effect of the number of (non-adjusted) bicyclist crashes increased after adding infrastructure elements, indicating a stronger increase in AADP with increases in bicyclist crashes than previously predicted.

The coefficients for the number of nearby Indigo stations and trips did not change significantly after adding infrastructure elements. This suggests that the effect of this program was already adequately modeled.

Table 14. Comparison of model results after adding infrastructure data, SIN, Philadelphia

	Bicyclist Crashes		**Pedestrian Crashes**	
	Without Infrastructure	**With Infrastructure**	**Without Infrastructure**	**With Infrastructure**
Constant	-2.177***	-1.131	-1.654***	-1.296***
log(AADB)		-0.300*		

Notes:
Coefficients shown (95% confidence intervals omitted). Only terms appearing in the models without infrastructure data are shown in order to compare the differences in results. Findings associated with new infrastructure elements are provided in a separate table.
*p<0.10; **p<0.05; ***p<0.01

As with AADB, the constant term for bicyclist crashes became statistically insignificant with the addition of infrastructure data, while several new infrastructure terms became significant (see

Table 15). Notably, log(AADB) became significant and negative, indicating that a 1% increase in log(AADB) is associated with a 0.3% *decrease* in bicyclist crashes. SIN predicts a less-than-proportional increase in crashes associated with a given change in volumes, but this relationship suggests a decrease in crash rates associated with an increase in bicyclist volume (all else equal). Table 15 summarizes the effects of infrastructure elements on both volumes and safety. The values in the table represent the ratio between the metric listed at the top of the column with the infrastructure and without it. For example, the presence of bike lanes is associated with roughly half (0.53) as much pedestrian volume as locations without bike lanes. Values above 1.0 generally indicate more volume or more crashes and values less than one indicate the opposite.

Table 15. Summary of the effects of infrastructure elements on volumes and safety

Infrastructure Element	***AADB***	***AADP***	***Bicyclist Crashes***	***Pedestrian Crashes***
Bike box				
Bike lane, buffered		0.03		
Bike lane, grade-separated	5.12			
Bike lane, protected	0.19			
Bike lane, standard		0.53	0.26	0.21
Bikeshare station				
Bumpout				
Crosswalk, standard			13.82	
Crosswalk, high-visibility	1.41		10.23	
Median refuge	0.55			
Pedestrian hybrid beacon	2.83			
Pedestrian signal				4.42
Sidewalk, standard	2.37			
Sidewalk, buffered				
Signalized intersection		0.62	0.10	
Stop sign	0.62	0.43	0.04	
Street parking	1.39			

Notes:
Values shown represent the ratio of volume or crashes among complete observations with each infrastructure element relative to those without.
Only coefficients determined to be at least 95% statistically significant are shown.

Some infrastructure elements were associated with higher bicyclist volumes, while others were associated with lower volumes. Observations made near grade-separated bike lanes were associated with the greatest positive difference in AADB (5.12) relative to observations without, followed by pedestrian hybrid beacons (2.83), standard sidewalks (2.37), high-visibility crosswalks (1.41), and street parking (1.39). Interestingly, protected bike lanes were associated with the greatest negative difference (0.19) followed by median refuges (0.55) and stop signs (0.62). Pedestrian volumes were roughly half as high near bike lanes, signalized intersections and stop signs relative to locations without, while buffered bike lanes are associated with just 3% of the pedestrian volume observed in the absence thereof.

Three infrastructure elements (standard bike lanes, signalized intersections, and stop signs) were associated with significantly fewer bicyclist crashes while both standard and high-visibility crosswalks were associated with alarmingly higher rates of bicyclist crashes. These findings underscore the importance of predictability on the road. Bike lanes, traffic signals, and stop signs may be effective at helping drivers detect bicyclists and encouraging road users to obey a right-of-way. Note that the crosswalks are not significant predictors of pedestrian crashes, only bicyclist crashes. Some bicyclists opt to ride on the sidewalk instead of in the traffic lane; when they approach intersections these bicyclists often cross in crosswalks, as they are natural extensions of sidewalks. This can surprise drivers by defying their expectations of how various road users pass through the intersection.

Locations with pedestrian signals were associated with 4.42 times as many pedestrian crashes compared to locations without. It may be that pedestrian signals were installed at high-pedestrian-volume locations, and this coefficient is the result of the elevated opportunity for incidents. Alternatively, this may suggest a pattern of misuse or noncompliance at pedestrian signals.

Anchorage, Alaska

Despite the research team's efforts, insufficient program metrics were acquired from Anchorage for analysis. As such, the program analysis was excluded for this location.

All volume and crash data for Anchorage were extracted from the Alaska DOT's Traffic Data Management System (Alaska DOT, 2021). Single-mode volumes were geocoded using the method described in *Data Preparation: Geocoding Crash and Volume Locations* and then connected using the method described in *Data Preparation: Connecting and Interpolating Data From Single-Mode Counters*. This process yielded 100 observations for bicyclists and 111 observations for pedestrians. The majority of these observations (87% for bicyclists, 83% for pedestrians) took place in 2018. Figure 23 provides histograms of AAD volumes for each road user group in the final data sets.

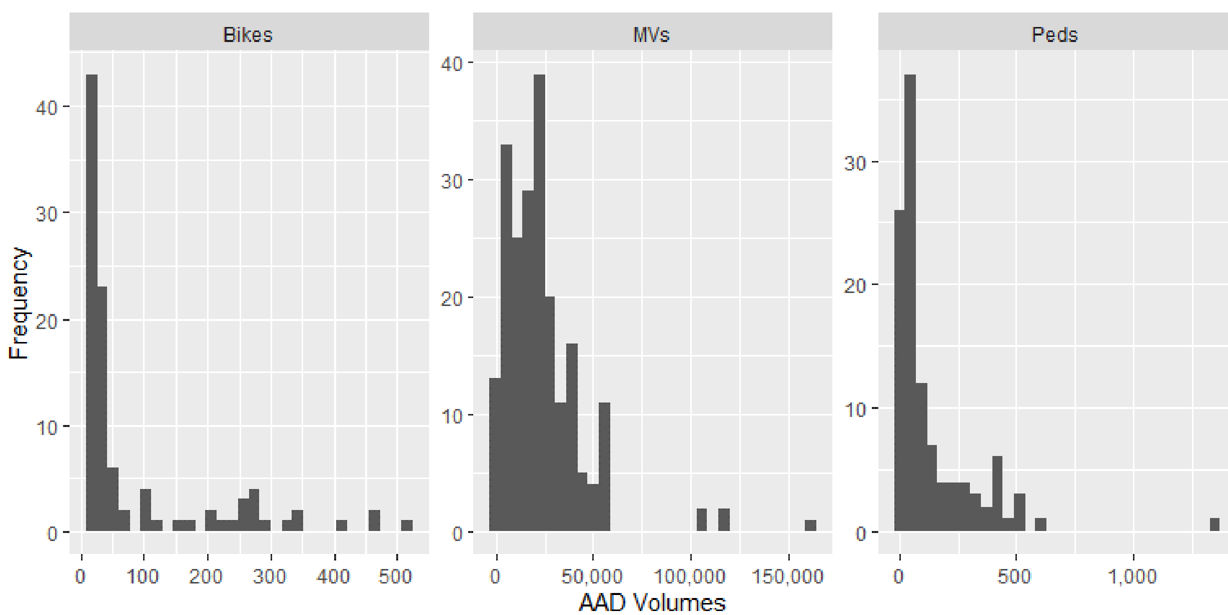

Figure 23. Histogram of AAD volumes by road user group, Anchorage

Figure 24 shows annual crashes by severity and the non-motorized road user group. Note that, although twenty years of crashes are shown, the final data set primarily consists of observations in 2018. During that time, 77% of locations experienced zero bicyclist crashes and 79% experienced zero pedestrian crashes within the 250 m-radius crash zones.

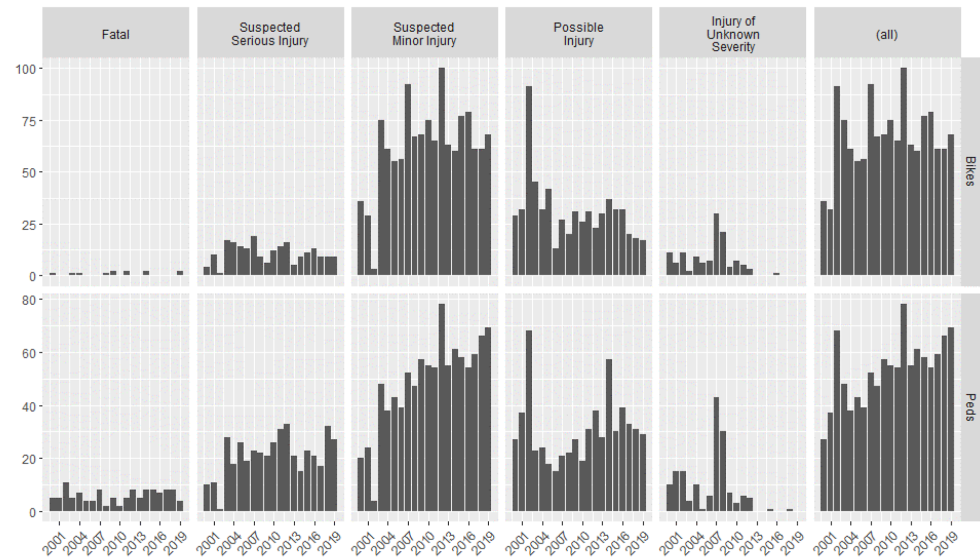

Figure 24. Annual crashes by severity and affected non-motorized road user group, Anchorage

Table 16 provides summary statistics for the bicyclist data set. The 100 observations took place between 2016 and 2019 (with 87% occurring in 2018 alone), with average AADM=21,879 and AADB=85. The number of crashes within each crash zone ranged from 0 to 3 (mean=0.3). The average distance between connected volume locations was 88m (see *Data Preparation: Connecting and Interpolating Data From Single-Mode Counters* for more information).

Table 16. Summary statistics, bicyclists, Anchorage

Statistic	***N***	***Mean***	***St. Dev.***	***Min***	***Pctl(25)***	***Pctl(75)***	***Max***
AADM	100	21,878.5	16,463.6	700	10,050	28,150	103,700
AADB	100	85.2	119.9	10	10	100	508
Bic. Crashes (sev.-adj.)	100	0.2	0.4	0.0	0.0	0.0	1.8
Bic. Crashes	100	0.3	0.6	0	0	0	3

Table 17 provides summary statistics for the pedestrian data set. The 111 observations also took place between 2016 and 2019 (with 83% occurring in 2018 alone), with average AADM=24,970 and AADP = 136. The number of crashes within each crash zone ranged from 0 to 5 (mean=0.3). The average distance between connected volume locations was 114m (see *Data Preparation: Connecting and Interpolating Data From Single-Mode Counters* for more information).

Table 17. Summary statistics, pedestrians, Anchorage

Statistic	*N*	*Mean*	*St. Dev.*	*Min*	*Pctl(25)*	*Pctl(75)*	*Max*
AADM	111	24,969.7	24,425.2	700	10,750	30,900	162,000
AADP	111	135.7	186.6	10	30	170	1,360
Ped. Crashes (sev.-adj.)	111	0.2	0.4	0	0	0	2
Ped. Crashes	111	0.3	0.7	0	0	0	5

Bicyclist and pedestrian volumes were modelled separately, using Poisson and negative binomial response distributions, with and without zero-inflation, and with and without statistical weights derived from the distance between volume counting locations. Similar models were also estimated for severity-weighted crash counts. AIC statistics indicated that Poisson, distance-weighted models of unadjusted crash counts were superior. Models were also fit with and without volume interaction terms. Stepwise selection was not applied, as the only variables available were those required for the SIN model described in the Statistical Analysis section.

Table 18 provides the Poisson, distance-weighted SIN models for Anchorage. The value of each coefficient is provided, with statistical significance indicated by asterisks (* indicates p<0.10, ** p<0.05, *** p<0.01) and 95% confidence intervals in parentheses.

Table 18. Model results, SIN, Anchorage

	Bicyclist Crashes		**Pedestrian Crashes**	
	Simple	**Full**	**Simple**	**Full**
Constant	-10.679^{***}	-22.793^{***}	-9.066^{***}	-9.504^{**}
	(-18.086, -3.272)	(-38.688, -6.897)	(-15.228, -2.903)	(-17.058, -1.950)
log(AADB)	0.075	1.147^{*}		
	(-0.341, 0.490)	(-0.076, 2.370)		
log(AADP)			0.547^{***}	0.589^{**}
			(0.204, 0.891)	(0.053, 1.125)
log(AADM)	0.933^{**}	1.837^{***}	0.552^{*}	0.580^{*}
	(0.182, 1.685)	(0.527, 3.146)	(-0.046, 1.150)	(-0.079, 1.239)
AADM x AADB		-0.000		
		(-0.000, 0.000)		
AADM x AADP				-0.000
				(-0.000, 0.000)
Observations	100	100	111	111
Log Likelihood	-49.211	-45.651	-48.025	-48.005
AIC	104.421	99.301	102.051	104.011

Notes:
Coefficients and 95% confidence intervals shown.
*p<0.10; **p<0.05; ***p<0.01

Results indicate *complete* SIN for bicyclists and *partial* SIN for pedestrians. The $AADM \times AADB$ and $AADM \times AADP$ terms were found statistically insignificant, while increasing the values of the parameters of interest. Hence, the “simple” model is discussed here. According to the “simple” model of bicyclist crashes, a 1% increase in AADM was associated with a 0.933% increase in bicyclist crashes. The coefficient for log (AADB) was not statistically significantly different from 0, supporting the conclusion that increases in AADB were not associated with any change in bicyclist crashes. The sum of the two coefficients is thus effectively $0.933 + 0 = 0.933 < 1$ which satisfies the condition for complete SIN.

For pedestrians, a 1% increase in AADP was associated with a 0.547% increase in pedestrian crashes, and a 1% increase in AADM was associated with a 0.552% increase in pedestrian crashes, implying that a 1% increase in both volumes would be associated with a 1.069% increase in pedestrian crashes. This is known as a *partial* SIN because each exponent is individually less than one, but their sum is greater than one.

Conclusions

This report is a culmination of an extensive investigation into SIN. Simply put, SIN is the idea that more walking and biking is associated with a safer road environment for pedestrians and bicyclists. Specifically, as the number of people who walk and bike increase, pedestrian and bicyclist crashes will increase at a lower rate. As more agencies are focusing on ways to reduce crashes, including efforts such as the Road to Zero or Vision Zero, measures that might increase pedestrian and bicycle crashes are a concern. SIN would help to provide the understanding of potential outcomes associated with increases the number of people walking and bicycling. A literature review conducted as part of this project provides a detailed history of the SIN concept, including when the term was first coined and challenges to and refinement of the concept (Kehoe et al., 2022).

What has been clear is that the cause of SIN is not well-understood. Jacobson (2003) hypothesized that SIN is a result of driver adaptation to increasing numbers of pedestrians and bicyclists in the driving environment. However, there are many other factors that might affect the presence of SIN, including population density (Coughenour et al., 2013), land use (Geng, 2014), infrastructure and countermeasure elements (Dumbaugh et al., 2013), policies and enforcement activities (Bhatia & Wier, 2011), cultural perspectives and attitudes toward non-motorized traffic (Fyhri et al., 2014), and education and outreach efforts (Johnson et al., 2014).

The evaluation described herein was undertaken to investigate the effect of pedestrian- and bicyclist-focused programs in increasing walking and biking and if implementing such programs creates a demonstratable SIN effect. Program evaluation results are often not formally published, do not describe the relationship between an increase in pedestrian and bicyclist volume and crashes, or discuss the other factors that may influence SIN. This project sought to use existing data from established programs to investigate SIN.

Findings and Discussion

Table 19 provides a summary of the findings for each program site evaluated in terms of program effectiveness and SIN, for both pedestrians and bicyclists.

Table 19. Summary of findings

Site	***Program effectiveness***		***Safety in Numbers***	
	Bicyclists	***Pedestrians***	***Bicyclists***	***Pedestrians***
Fort Collins	Unclear	Unclear	Safety (complete)	Safety (partial)
Philadelphia	Success	Not Applicable	No relation	No relation
Anchorage	Insufficient Data	Insufficient Data	Safety (complete)	Safety (partial)

The effect that selected programs exerted on bicyclist and pedestrian volumes remains unclear. The negative correlation between volumes and some program metrics in Fort Collins more likely reflects a geographical and/or temporal disconnect between the program metrics and volumes. Safe Routes to School and other programs may influence bicyclist and pedestrian volumes on a (geographic) scale too small to be measured by volume counts. Regardless of the size of the area of influence, if counts are not conducted in the highly localized communities where programs are implemented (e.g., within walking distance of schools), there is little chance that any change will be measured. The positive correlation between bicyclist volumes and localized bikeshare trips – and lack of correlation between volumes and city-wide trips – provides further evidence of this claim.

The analysis used annualized statistics for nine years in Fort Collins, 10 years in Philadelphia, and essentially one year in Anchorage. Annualized statistics may not provide the level of granularity required to detect meaningful changes, and these time spans may not be long enough for slower changes to manifest. The nature of some of the programs studied may cause immediate volume shifts while others may take several years to produce changes in volumes. For example, a group ride that some participants subsequently adopt, and repeat could cause an immediate increase in volume, but Safe Routes to School may provide younger road users with the skills and confidence to bicycle or walk for years to come.

Of course, many other relevant factors could be in flux over the course of the observation period, such as infrastructure improvements and culture shifts. Some transportation agencies develop inventories of infrastructure features, but infrastructure changes over time, and inventories must be updated to support meaningful research. An ad-hoc analysis of Philadelphia revealed several statistically significant predictors of volumes and crashes. Most notably, both standard and high-visibility crosswalks were associated with higher rates of crashes among bicyclists, and pedestrian signals were associated with higher crash rates among pedestrians; while this seems counterintuitive, the findings likely result from increased exposure and higher volumes of pedestrians and bicyclists at these locations. Crosswalks and signals are typically installed in areas that exhibit more potential for conflicts between pedestrian, bicyclist, and vehicle movements and in areas with higher pedestrian and bicyclist volumes. While exposure may explain the results, the underlying mechanisms for these relationships are unclear. More detailed and targeted analysis of infrastructure elements and crash circumstances may produce valuable insight into these results.

This analysis produced differing results for pedestrians and bicyclists concerning SIN. Bicyclists in Fort Collins and Anchorage benefit from complete SIN, while pedestrians experienced only partial SIN. Recall that *complete* SIN is said to occur when bicyclist/pedestrian crashes increase at a rate less than proportional to simultaneous increases in bicyclist/pedestrian *and* motor vehicle volumes. In contrast, *partial* SIN occurs when bicyclist/pedestrian crashes increase at a rate less than proportional to increases in bicyclist/pedestrian *or* motor vehicle volumes. The outcome of complete SIN for bicyclists and partial SIN for pedestrians may indicate an important difference between the two travel modes. Perhaps – because bicyclists generally ride in the road, whether in a bike lane or mixed with motorized traffic, while pedestrians are more often removed from the road by a sidewalk and mostly interact with motorized traffic through traffic control devices – an increase in bicyclist volumes elicits a greater response from drivers than a comparable increase in pedestrian volumes. This possibility speaks to the larger question surrounding the exact mechanism behind SIN.

Challenges and Opportunities

Robust, multifaceted data are required to evaluate program effectiveness and SIN. The literature and the analysis described in this report demonstrate how these are challenging to obtain. Comprehensive count or volume data are rarely readily available for pedestrians and bicyclists, and collecting these data is often resource-intensive. The research team encountered four major challenges concerning the data and developed solutions from which future researchers can benefit. Short-term counts are essentially meaningless unless they can be converted into more standardized statistics (i.e., AAD volumes). These converted statistics would remain unsuitable for analysis if they were single-mode counts and consolidating single-mode counts would be laborious and error-prone if their geographical locations were not expressed in latitude and longitude. See the *Data Preparation* section for more information on how these challenges were overcome.

There are other challenges with pedestrian and bicycle safety data including incomplete crash data sets due to underreporting of pedestrian and bicycle injury data. Also, the built environment can have a substantial impact on crashes. While some researchers have been successful in introducing variables describing the built environment and behavioral characteristics, these topics are a current gap in SIN research and are often covered only briefly, if at all, by current research. Overcoming that gap in SIN research will require robust data sets related to those environmental and behavioral characteristics; however, these data have practical limitations. For example, organizations involved in transportation (Departments of Transportation at the State and local levels, regional planning organizations, public health agencies, transportation service providers (such as bike share organizations), schools, law enforcement organizations, etc.) have different data collection capabilities based on their organizational goals at the time and their resources. This is even more apparent in program data collection. In many instances, pedestrian and bicycle programs dedicate their resources to implementation efforts and less often on record keeping on activities. This can ultimately limit knowledge about the reach and impact of the program.

One question that still lingers is: what causes SIN? This research can point to correlations in the data but not causation. As discussed with the data challenges, with increased and improved data sets, researchers can come closer to understanding the factors influencing SIN. Some education and encouragement programs – such as Safe Routes to School – have better data sets because of funding requirements tied to program elements and participation. Regardless, there are many opportunities for improvements in data sources, such as program participation, robust pedestrian and bicyclist volume and crash data collection, monitoring changes in pedestrian/bicyclist/driver behaviors, improved infrastructure documentation (such as installation dates), and surveys to determine the impact of education and encouragement programs. Advance planning for data collection efforts can facilitate analyses to gain a better understanding of what works to improve safety for pedestrians and bicyclists.

References

Alaska Department of Transportation. (2021). *Traffic data management system.* https://traffic.muni.org/TSReports.aspx

Bhatia, R., & Wier, M. (2011, January). "Safety in Numbers" re-examined: Can we make valid or practical inferences from available evidence? *Accident Analysis & Prevention, 43*(1), 235-240. www.sciencedirect.com/science/article/pii/S0001457510002484

City of Fort Collins. (2020). *Public records.* https://records.fcgov.com/WebLink/Welcome.aspx?dbid=0&repo=FortCollins

Climate Denver – Colorado. (2021). *U.S. climate data.* www.usclimatedata.com/climate/denver/colorado/united-states/usco0105

Climate Fort Collins – Colorado. (2021). *U.S. climate data.* www.usclimatedata.com/climate/fort-collins/colorado/united-states/usco0140

Cooley, D. (2020). *Googleway: Accesses Google maps apis to retrieve data and plot maps. R package version 2.7.3.* https://CRAN.R-project.org/package=googleway

Coughenour, C., Pharr, J., & Bungum, T. (2013). Environmental, human, and socioeconomic characteristics of pedestrian injury and death in Las Vegas, NV. *International Journal of Sciences*, 31-39. www.semanticscholar.org/paper/Environmental%2C-Human-and-Socioeconomic-of-Injury-in-Pharr-Coughenhour/0d6ecb401b1ede6f02259ba5f6119e4aff295337

Dumbaugh, E., Li, W., & Joh, K. (2013). The built environment and the incidence of pedestrian and cyclist crashes. *Urban Design International*, *18*(3), 217-228. http://link.springer.com/article/10.1057/udi.2013.2

Elvik, R. (2009). The non-linearity of risk and the promotion of environmentally sustainable transport. *Accident Analysis & Prevention, 41*(4), 849-855. www.sciencedirect.com/science/article/abs/pii/S0001457509000876

Elvik, R. (2013, November). Can a safety-in-numbers effect and a hazard-in-numbers effect co-exist in the same data? *Accident Analysis & Prevention*, *60*, 57-63. www.sciencedirect.com/science/article/pii/S0001457513003229

Elvik, R. (2016, June). Safety-in-numbers: Estimates based on a sample of pedestrian crossings in Norway. *Accident Analysis & Prevention*, *91*, 175-182. www.sciencedirect.com/science/article/abs/pii/S0001457516300768

Elvik, R. (2017, March). Exploring factors influencing the strength of the safety-in-numbers effect. *Accident Analysis & Prevention, 100,* 75-84.

Elvik, R., & Bjørnskau, T. (2017, February). Safety-in-numbers: A systematic review and meta-analysis of evidence. *Safety Science, 92*, 274-282. www.sciencedirect.com/science/article/pii/S0925753515001812

Elvik, R., Sørensen, M., & Nævestad, T. (2013, October). Factors influencing safety in a sample of marked pedestrian crossings selected for safety inspections in the city of Oslo. *Accident Analysis & Prevention*, *59*, 64-70. www.sciencedirect.com/science/article/abs/pii/S0001457513002029

Federal Highway Administration (2016, October). *Traffic monitoring guide* (FHWA Report FHWA-PL-17-003). www.fhwa.dot.gov/policyinformation/tmguide/tmg_fhwa_pl_17_003.pdf

FHWA. (2021). *December 2020 traffic volume trends*. www.fhwa.dot.gov/policyinformation/travel_monitoring/20dectvt/

Fyhri, A., Sundfør, H., & Bjørnskau, T. (2014, November 18-19). Safety-in-numbers—combining a panel design and cross cultural survey to examine the suggested mechanisms. *3rd International Cycling Safety Conference*, Gothenburg, Sweden. www.researchgate.net/profile/Aslak_Fyhri/publication/271131002_Safety_in_Numbers_-_combining_a_panel_design_and_cross-cultural_survey_to_examine_the_suggested_mechanisms/links/54be21630cf27c8f2814df18.pdf

Geng, S. (2014). *Environmental characteristics around hotspots of pedestrian-automobile collision in the city of Austin*. University of Texas at Austin. https://repositories.lib.utexas.edu/handle/2152/25873

Hauer, E. (2013, January 13-17). *Even perfect regressions may not tell the effect of interventions.* Transportation Research Board 92nd Annual Meeting, Washington DC.

Hijmans, R. J. (2019). Geosphere: Spherical trigonometry. R package version 1.5-10. https://CRAN.R-project.org/package=geosphere

Hyndman, R. J., & Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software*, *27*(3), 1-22.

Ingdal, M., Johnsen, R., & Harrington, D. A. (2019). The Akaike information criterion in weighted regression of immittance data. *Electrochimica Acta, 317*, 648-653.

Jacobsen, P. (2003). Safety in numbers: more walkers and bicyclists, safer walking and bicycling. *Injury Prevention, 9*(3), 205-209. https://injuryprevention.-bmj.com/content/9/3/205

Johnson, M., Oxley, J., Newstead, S., & Charlton, J. (2014, September). Safety in numbers? Investigating Australian driver nehaviour, knowledge and attitudes towards cyclists. *Accident Analysis & Prevention*, *70*, 148-154. www.sciencedirect.com/science/article/abs/pii/S0001457514000499

Kehoe, N. P., Goughnour, E., Jackson, S., Sykes, K., Miller, S., & Blackburn, L. (2022, June). *Safety in numbers: A literature review* (Report No. DOT HS 813 279). National Highway Traffic Safety Administration.

National Center for Statistics and Analysis. (2020, December). *Overview of motor vehicle crashes in 2019* (Traffic Safety Facts Research Note. Report No. DOT HS 813 060). National Highway Traffic Safety Administration. https://crashstats.nhtsa.dot.gov/Api/Public/Publication/813060

NCSA. (2021, June (revised)). *Early estimates of motor vehicle traffic fatalities and fatality rate by sub-categories in 2020* (Crash•Stats Brief Statistical Summary. Report No. DOT HS 813 118). National Highway Traffic Safety Administration. https://crashstats.nhtsa.dot.gov/Api/Public/Publication/813118

Pennsylvania Department of Transportation. (n.d.). *Pennsylvania crash information tool.* https://crashinfo.penndot.gov/PCIT

R Core Team (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computinga. www.R-project.org/

Smeed, R. (1949). Some statistical aspects of road safety research. *Journal of the Royal Statistical Society. Series A (General), 112*(1), 1-34. www.jstor.org/stable/2984177

Venables, W. N. & Ripley, B. D. (2002). Modern applied statistics with S. Fourth edition. Springer. ISBN 0-387-95457-0.

Wang, X., Abdel-Aty, M., Nevarez, A., & Santos, J. B. (2008). Investigation of safety influence area for four-legged signalized intersections: Nationwide survey and empirical inquiry. *Transportation Research Record, 2083*(1), 86-95.

Appendix A: Preliminary Program Site List

This section provides information on programs considered for this project. Symbols to the left of each program name indicate the program's target audience. Parentheticals indicate the program's location and first year of operation.

Watch for Me NC **(North Carolina, 2012)**

Unique features	Statewide - there are both urban and rural areas, numerous universities, and heavy tourism.
Known data available	There are annual progress reports between 2012 and 2018 as well as a crash-based evaluation by Saleem et al. (2018). Progress reports and evaluation reports associated with the program describe data collection through the program, including number of workshops held, number of workshop participants, amount of paid and earned media delivered, quantity and type of law enforcement activities, number of pedestrian- or bicyclist-related crashes.
Basis for selection	This program is still active, has large reach, and appears that data may be accessible based on the progress reports and evaluation report.
Points of contact	Senior Research Associate at the UNC Highway Safety Research Center; Former NCDOT Bicycle and Pedestrian Coordinator

Colorado Pedals **(Colorado, 2015)**

Unique features	High tourism, residents are known for having an active lifestyle, the majority of residents own a bicycle.
Known data available	The 2016 Economic and Health Benefits of Biking and Walking Study reveals the who, what, where, how, and why of bicycling, walking and health in the State of Colorado. It is a first of its kind and could set a national standard for gathering and legitimizing such data.
Basis for selection	The Colorado Governor initiated this program to increase bicycle safety. In the League of American Bicyclists' 2018 progress report, Colorado ranked as the sixth most bicycle-friendly state. However, Colorado wants to do better. From 2007-17, Colorado experienced a sharp decrease in people riding to work on their bicycles and has continued to see threats to people riding their bikes.
Point of contact	Bicycle Colorado

NJ Ambassadors in Motion **(New Jersey, 2012), and** ***Street Smarts NJ*** **(New Jersey, 2013)**

Unique features	While these programs are statewide programs, New Jersey is a geographically small State. Activities completed within this program range in audiences, with some examples including college students and children. Given the statewide coverage, there is a range of locations where program activities were completed. These include urban and suburban areas, as well as tourist locations such as the New Jersey shore.
Known data available	Ambassadors in Motion measures performance using the following metrics. Data for these metrics are included in the program's annual reports: # of Statewide Events; # of Outreach Events at Target Locations; # of People Educated On-Street; # of Training and Educational In-Class Workshops; # of People trained and Educated In-Class; # of Walking and Biking Events Attended; # of People Educated at Biking and Walking Events; # of People Indirectly Contacted; Total Persons Reached (Weighted); Complete Streets Policies Adopted Street Smarts conducts evaluations, but it is unknown what data are collected.
Basis for selection	Ambassadors in Motion is one of multiple pedestrian safety programs initiated by New Jersey as the result of being named a "Focus State" by FHWA and has history, reach, and appears to be collecting data. Evaluations of Street Smarts have been conducted, at least some data have been collected, and the program has history/significant outreach.
Points of contact	Ambassadors in Motion-NJ Bicycle and Pedestrian Resource Center Street Smarts NJ

Heads Up (Maine, 2016), *Bicycle Coalition of Maine* (Maine, 1992), and *Portland Area Comprehensive Transportation System* (Maine, 1992)

Unique features	Heavy tourism, far northeast with snow and coastal conditions. Rural in nature with one small city (Portland). Portland has strong bicycle and pedestrian infrastructure.
Known data available	The Heads Up program began in 2016 by identifying focus communities that experienced the highest number of pedestrian crashes between 2011 and 2015. These locations often represent the State's most densely populated and "urbanized" areas. The *Bicycle Coalition of Maine* collects data on crashes, collisions, close calls, and incidents of bicyclist harassment. They collect this data because they believe it assists them in identifying issues faced by their constituents, as well as advocating for better infrastructure, education, and enforcement efforts.
Basis for selection	Maine has strong pedestrian and bicycle infrastructure and the State DOT is dedicated to improving ped and bike safety. In 2007 the Coalition led the successful effort to win passage of a major revision of Maine bicycling laws. A key provision requires motorists to give three feet of clearance when passing bicycles. This might assist with collection of data before and after the new laws. The Heads Up program collection of data between 2011-2015 would be useful for the evaluation.
Points of contact	Maine DOT Active Transportation Planner Bicycle Coalition of Maine PACTS, Senior Planner

BikeArlington, and WalkArlington (Arlington, Virginia, 2009)

Unique features	Urban area with tourists, commuters, and residents as well as a large transit system and the possibility that people are biking or walking into Washington, DC. Residents include both lower and higher income groups.
Known data available	Rackspotter, counter data, and Bikeometer. Rackspotter is a free, crowdsourced tool available for web browsers and smartphones that can be used to identify bike parking locations throughout the greater Washington DC region. The counter data are count data for Arlington County. Resources for the count data as well as an online dashboard are provided. The Rosslyn Bikeometer is the first of its kind on the East Coast and sixth in the United States, preceded by versions in Portland, Seattle, and California. It provides a highly visible, engaging, and fun view of the volume of bike usage on the Custis Trail in Arlington. Ride past the Bikeometer and notice the daily bike count tick off another digit and acknowledge another car-free trip in Arlington County. The information displayed is real time and includes month and year-to-date data. There was a survey asking residents what type of rider they are, what type of trips they make, how comfortable they feel on different types of bike lanes and trails and what ways the county could encourage them to choose to bike more often.
Basis for selection	BikeArlington is an active program that combines events, training, education, and bikesharing with the availability of tools and data. Program use of tools like Rackspotter and the Bikeometer are fairly limited, so studying these tools and related outreach efforts may provide insight not found in other agencies. Arlington County is recognized as a Gold Walk Friendly Community by the Pedestrian and Bicycle Information Center. The County is one of only 15 communities across the country to have received a Gold rating. Arlington is one of nine communities profiled and featured in America Walks and the Every Body Walk! Collaborative's book, *America's Walking Renaissance*. An entire chapter of the book is dedicated to Arlington County's success, calling Arlington "America's Most Walkable Suburb."
Points of contact	Arlington County Program Director and Design Team Supervisor

Bikeology (Anchorage, Alaska, 2015)

Unique features	None
Known data available	Bikeology program participant records. Crash records from Alaska DOT and AAD volumes from Alaska DOT.
Basis for selection	The city has used the Bikeology program developed by the Society of Health and Physical Educators America with funding and technical support from NHTSA for some time. They felt as though the program data were robust and also had quite a bit of traffic volume data. Additionally, there are unique environmental characteristics, such as the long stretches of darkness.
Point of contact	None

Boston Bikes (Massachusetts, 2007), *WalkBoston* (Massachusetts, 1990)

Unique features	Urban area with tourists, commuters, university students, and residents. Residents include both lower and higher income groups.
Known data available	Boston Bikes - By 2030 Boston Bikes' goal is increasing bicycling fourfold. They collect and analyze data each year to understand their progress toward that goal. In 2016 they began a new annual bicycle count program using automated technology. They collect data at more than 60 locations over a 48-hour period in late September and early October. Over time, they can use this data to better understand how many people are already biking in Boston, and what they can do to encourage more people to go by bike. They also collect commuting data and bike share data. WalkBoston tracks pedestrian fatalities across Massachusetts through a robust monitoring of news reports. This information is more timely and more detailed than the tracking that is available through state or municipal sources.
Basis for selection	Boston Bikes is very much focused on data collection and evaluation. Much of what they describe on their web site aligns with what we are trying to identify for this project. WalkBoston has existed for over 25 years and works with over 115 cities and towns across the State and may have data to share from other locations.
Points of contact	Boston Bikes WalkBoston, Executive Director

BikePGH (2002), *OpenStreetsPGH, Walk Works* (Pittsburg, Pennsylvania)

Unique features	Urban area with tourists, commuters, university students, and residents. Topographical challenges with steep terrain and separated neighborhoods connected through city steps. Rainy and snowy environment. There is a possible rural component as well through SPC's efforts in the region. They have multiple Walk Works groups in Indiana and Fayette Counties.
Known data available	Permanent bike counters placed in locations through the city. Recently completed a large survey/data analysis effort on mode choices. Counts are developed for Open Streets events.
Basis for selection	The city has been encouraging mode shifts through complete streets designs and has a mayor who strongly supports bike/ped initiatives so there appears to be a general swell of support and encouragement for walking and biking. The city and regional staff support data/information sharing.
Points of contact	BikePGH Director Southwest Planning Commission Transportation Planner

Great Rides Bike Share (Fargo, North Dakota, 2015)

Unique features	Colder weather, flat and open terrain, university students, operations on a seasonal schedule (open 8 months of the year, closed in fall and winter).
Known data available	Web site includes ride data both real time and historical, data includes start and end location of trips and duration.
Basis for selection	While the research team is unsure if there is enough data available to support an evaluation, this program would be interesting to study because of the unique university partnership, as well as the different climate and geography of Fargo as compared to other selected areas.
Point of contact	Great Rides Fargo

***Bike Austin*, and *City of Austin Bike/Ped Program* (Austin, Texas, 1975)**

Unique features	Urban area with tourists, commuters, university students, and residents.
Known data available	TXDOT project, “Evaluation of Bicycle and Pedestrian Monitoring Equipment to Establish Collection Database Methodologies for Estimating Non-Motorized Transportation” was conducted in Austin and Houston. Results may be helpful for this effort. Ride Report App has crowdsourced bicycle trip data. The city has permanent bike/ped counters.
Basis for selection	Austin is recognized as a GOLD Level Bicycle Friendly Community by The League of American Bicyclists. Austin was selected through national competitions by PeopleForBikes as one of 10 leading cities to join the Big Jump Project (2017) and as a groundbreaker city for its Green Lane Project. Austin was selected as the host city for the 2015 NACTO Designing Cities conference, the leading national event on the multimodal design of city streets. Austin was a leading city in the U.S. Department of Transportation Mayor’s Challenge for Safer People and Safer Streets. Austin received the program’s Ladders of Opportunity Award in 2016.
Points of contact	Active Transportation and Street Design Division Manager, City of Austin Bike Austin, President of Board of Directors

Look Alive (Baltimore, Maryland, unknown year)

Unique features	Urban area with tourists, commuters, university students, and residents, low and high-income groups.
Known data available	Basic crash statistics (e.g., number of death and injuries for pedestrians and cyclists in the region) are listed on the program's website. A social media toolkit is available that lists the dates and content of social media posts for the program. Look Alive conducts pre- and post-campaign surveys to measure awareness and attitudes among drivers, cyclists, and pedestrians. Detailed post-campaign reporting also includes impressions and engagement via paid media, donated media, news coverage, digital efforts, and outreach.
Basis for selection	This is a fairly comprehensive program integrating outreach and enforcement and it appears that they have conducted evaluations. This program seems to have a larger enforcement component than some of the others. Pedestrian and bicyclist fatalities in the Baltimore region accounted for 50% of the pedestrian and bicyclist fatalities in the state in 2018.
Point of contact	Baltimore Metropolitan Council's Principal Transportation Engineer

***Pima Association of Governments (PAG) Bicycle and Pedestrian Planning Program* (Pima County, Arizona, unknown year)**

Unique features	Arizona is one of the top states in the country for pedestrian fatalities and is involved in efforts led by FHWA to focus on pedestrian safety. While there are advances in Tucson, the State is generally considered to have less bike/ped friendly design/ atmosphere. Weather conditions – hot arid environment.
Known data available	PAG conducts a bicycle and pedestrian count every fall, relying on the support of jurisdiction staff and community volunteers to count at approximately 80 locations through the entire region. PAG began the program in 2008 to: Better understand trends/characteristics of cyclists by collecting data on: direction of travel, gender, age, helmet usage, sidewalk riding and wrong-way riding; help evaluate planning efforts; and help guide investments. PAG updates a Regional Bicycle Crash Analysis annually, with data dating back to 2001. The analysis is used to help identify mitigation strategies, such as wrong-way signs and pavement markings, as well as help identify enforcement education needed. It also quantifies the number of total crashes, crashes per population and fatal crashes.
Basis for selection	In 2006 and again in 2008 the League of American Bicyclists (LAB) recognized the Tucson - Pima Eastern Region as a Gold Level "Bicycle Friendly Community," the first and only such regional designation in the United States. Bicycling Magazine has ranked the City of Tucson as the 2nd best bicycling city in the United States in 1995, 1999, and more recently, in 2006. The City of Tucson, Pima County, Oro Valley, and PAG all have full-time staff working on bicycle issues. There are also a variety of active, involved citizens, bike clubs and advocate groups working to support and improve cycling in this region.
Points of contact	Lead Planner, Bicycle & Pedestrian Program Tucson, Arizona Bicycle & Pedestrian Program Coordinator Tucson, Arizona Tohono O'odham Nation

City of Davis, California, Bike and Pedestrian Program **(unknown year)**

Unique features	Urban area with tourists, commuters, university students, and residents
Known data available	In April 2018 The City of Davis inventoried all the bike parking in downtown Davis and collected occupancy data. There are two active bike counters in the City of Davis. There is an Eco-Counter PYRO Box on the shared use path west of Pole Line Rd between Drexel Dr and Loyola Dr. This counter uses a passive infrared pyroelectric sensor to detect bicycle and pedestrian traffic. This counter has a directional sensor.
Basis for selection	The City of Davis is one of the top bicycling cities in the country and is considered the bicycle capital of the United States. With over 100 miles of on-street and Class 1 bicycle lanes, the City of Davis provides bicyclists and pedestrians safe access to and from school, thereby eliminating the need for the school district to provide school buses.
Points of contact	Caltrans Pedestrian and Bicycle Safety Branch Chief Caltrans Senior Transportation Planner, Active Transportation Program

***North Central, Texas, Council of Governments Bicycle and Pedestrian Advisory Committee* (unknown year)**

Unique features	Includes one of the largest metropolitan regions in the United States – however both cities are less dense than other similarly sized regions. Generally unfriendly bicycling and pedestrian environment.
Known data available	TXDOT has been collecting and analyzing pedestrian and bicycle data throughout the state. NCTCOG also conducts regional bike/ped counts and analyzes regional bike/ped crash data. Has a bicycle opinion survey, as well.
Basis for selection	A notoriously poor environment for bicyclists and pedestrians that is working to improve infrastructure and encourage more walking and biking. The region and TXDOT have data that may prove to be beneficial for this task. The research team has been working with the region on bicycle and pedestrian efforts and have strong connections with NCTCOG, the cities, and TXDOT.
Points of contact	Program Manager, Sustainable Development NCTCOG Transportation Planner, City of Dallas Senior Planner, City of Fort Worth

***Delaware Valley Regional Planning Commission Bicycle and Pedestrian Planning Programs; Philadelphia Office of Transportation, Infrastructure, & Sustainability Pedestrian and Bicycle Program – Indego Bikeshare Initiative* (unknown year)**

Unique features	Urban area with tourists, commuters, university students, and residents, low and high income groups, and a robust transit system.
Known data available	DVRPC has an ongoing program to collect bicycle and pedestrian counts on roadways and trails throughout the region using infrared equipment.
Basis for selection	Known to have robust bicycle and pedestrian data. Recent Vision Zero programs in the city of Philadelphia might have information to use for this effort.
Points of contact	Director of Complete Streets, City of Philadelphia
	Director of Policy & Strategic Initiatives, City of Philadelphia
	Associate Manager, Office of Transit, Bicycle, and Pedestrian Planning

***Fort Collins, Colorado, Safe Routes to School program, Open Streets events, Bicycle Ambassador Program, and Bike to Work Day Program* (unknown year)**

Unique features	Small city, more spread out with larger block sizes. Strong culture of walking and biking, platinum-level Bicycle Friendly City with over 285 miles of bike lanes and bikeable trails.
Known data available	Individual crash records, short-term counts, program data on participants trained and outreach numbers.
Basis for selection	Number of programs focused on pedestrians and bicyclists.
Points of contact	Safe Routes to School Program Coordinator, City of Fort Collins
	Director of Policy & Strategic Initiatives, City of Fort Collins

DOT HS 813 342
August 2022

U.S. Department
of Transportation

National Highway
Traffic Safety
Administration

15666-080422-v3