

Exploring Crowdsourced Monitoring Data for Safety - StreetLight Data Bicycle Count Estimates vs Texas Permanent Bicycle Counts Dataset

Dataset available at: <https://doi.org/10.15787/VTTI/OBV82F>

(This dataset supports report **Exploring Crowdsourced Monitoring Data for Safety**, https://martrec.uark.edu/research/vu_development_final_report_accessible.pdf)

This U.S. Department of Transportation-funded dataset is preserved by the Safety through Disruption (Safe-D) University Transportation Center (UTC) in the digital repository Virginia Tech Transportation Institute (VTTI) Dataverse (<https://dataverse.vtti.vt.edu/>), and is available at <https://doi.org/10.15787/VTTI/OBV82F>

The related final report **Exploring Crowdsourced Monitoring Data for Safety**, is available from the National Transportation Library's Digital Repository at <https://rosap.ntl.bts.gov/view/dot/50717>

Metadata from the VTTI Dataverse record:

Description:

- **Project Description:** These data represent monthly average weekday and weekend bicycle counts and count index values at specific points locations across eight cities in Texas: Allen, Austin, Dallas, Fort Worth, Houston, North Richland Hills, Plano, and San Antonio. The bicycle counts were produced from permanent counters with complete sets of 15-minute counts per day. The bicycle count index values were provided by StreetLight Data and represent a normalized estimate that controls for variation in the passive data sample penetration rates each month. StreetLight uses passive data from mobile devices, i.e. crowdsourced location-based services, to algorithmically derive mode of travel and route. Depending on data availability, the data represent April, May, and June of 2018. These data were used to evaluate crowdsourced bicyclist activity data from StreetLight Data against actual bicycle counts. The researchers found promising correlations (R² of 62% and 69% for monthly weekday and weekend daily averages) when the StreetLight bicyclist index values were compared to bicyclist counts from 32 permanent counter locations in eight Texas cities, and even better correlation (R² of 94%) when compared with countywide Strava data expanded to represent total bicycling activity estimates.
- **Data Scope:**
 - Monthly average weekday and weekend bicycle counts and crowdsourced bicycle count index values.
 - 32 unique count locations.
 - 8 cities across Texas.
 - 58 monthly observations.
- **Data Specification:**
 - KEY - character varying; unique ID; no missing values
 - City - text; Allen, Austin, Dallas, Fort Worth, Houston, North Richland Hills, Plano, and San Antonio; no missing values

- Station ID TMG - character varying; no missing values
- Month - numeric; 4-6; 4 = April, 5 = May, 6 = June; no missing values
- StL Monthly Weekday Average - numeric; 5-611; no missing values
- TxDOT Monthly Weekday Average - numeric; 17-1518; no missing values
- StL Monthly Weekend Average - numeric; 11-800; no missing values
- TxDOT Monthly Weekend Average - numeric; 34-2507; no missing values

Subject: Engineering

Keyword: bicycle count, passive data, LBS, StreetLight Data

Recommended citation:

Turner, Shawn; Martin, Michael, 2019, "Exploring Crowdsourced Monitoring Data for Safety - StreetLight Data Bicycle Count Estimates vs Texas Permanent Bicycle Counts (TTI-Student-05)", <https://doi.org/10.15787/VTI1/OBV82F>, VTTI, V1, UNF:6:6JVCAnxeQjxHCF1bnVLO4g== [fileUNF]

Dataset description:

This dataset contains 1 .csv file described below.

StL vs TxDOT Counts - Safe-D Upload.csv:

The .csv, Comma Separated Value, file is a simple format that is designed for a database table and supported by many applications. The .csv file is often used for moving tabular data between two different computer programs, due to its open format. The most common software used to open .csv files are Microsoft Excel and RecordEditor, (for more information on .csv files and software, please visit <https://www.file-extensions.org/csv-fileextension>).

National Transportation Library (NTL) Curation Note:

As this dataset is preserved in a repository outside U.S. DOT control, as allowed by the U.S. DOT's Public Access Plan (<https://ntl.bts.gov/public-access>) Section 7.4.2 Data, the NTL staff has performed **NO** additional curation actions on this dataset. NTL staff last accessed this dataset at <https://doi.org/10.15787/VTI1/OBV82F> on 2020-10-01. If, in the future, you have trouble accessing this dataset at the host repository, please email NTLDataCurator@dot.gov describing your problem. NTL staff will do its best to assist you at that time.