

NO ACCIDENT

HOW POPULATION CHARACTERISTICS AND LAND USE INFLUENCE STREET SAFETY FOR PEDESTRIANS IN SAN FRANCISCO

MIKE JACOBSON
MAY 2018



No Accident: How Population Characteristics and Land Use Influence Street Safety for Pedestrians in San Francisco

A Planning Report Presented to The Faculty of the Department of Urban and Regional Planning
San José State University

In Partial Fulfillment Of the Requirements for the Degree Master of Urban Planning

By

Mike Jacobson

May 2018

ACKNOWLEDGMENTS

This report would not have been possible without the dedicated support of several individuals. I would like to thank the following people:

- Advisors Richard Kos, Asha Weinstein Agrawal, and Kevin Fang. Their guidance was paramount in refining a once broad sketch of a research proposal into a focused and impactful planning report.
- Professor Shishir Mathur for reviewing my dataset and steering me towards completing a logistic regression.
- City staff at the San Francisco Municipal Transportation Agency and the San Francisco Department of Public Health for providing publically accessible data and providing additional datasets that were not available online.
- My 298B study group for the academic and emotional support.
- Bonnie Kaufman and Tracy Minicucci for helping edit the document.
- All of my family and friends for their constant and unwavering support.

CONTENTS

TABLE OF CONTENTS

Executive Summary X-XV **Chapter 4: Methodology** 44-81

Chapter 1: Introduction 16-21

- 1.1 Research Questions and Introduction of the Methods
- 1.2 How to Read This Report

- 4.1 Prior Research
- 4.2 Population Density, Employment Density, and Land Use Mix Workflow
- 4.3 Trip Generator Typology
- 4.4 Trip Generator Metrics
- 4.5 Trip Generator Workflow
- 4.6 Logistic Regression

Chapter 2: Vision Zero 22-31

- 2.1 The Origins of Vision Zero in San Francisco
- 2.2 The High Injury Network and "Focus on the Five"

Chapter 5: Findings 82-111

- 5.1 Population Density and Crashes
- 5.2 Employment Density and Crashes
- 5.3 Land Use Mix and Crashes
- 5.4 Trip Generator Typology
- 5.5 Logistic Regression

Chapter 3: Mapping Clusters of Automobile-Pedestrian Crashes 32-43

- 3.1 Aggregating Automobile-Pedestrian Crashes
- 3.2 Hot Spot Analysis and Kernel Density Estimation

Chapter 6: Discussion and Suggestions for Future Research 112-117

Appendix A: Trip Generator Typology Maps
Appendix B: Data Sources
References

TABLES

FOCUS LIST

Table 1: Aggregated Zoning Districts	50
Table 2: Acreage by Zoning Category	51
Table 3: Land Use Mix by Block Group	52
Table 4: Gross Crashes and Crashes per Capita by Block Group Population Density	87
Table 5: Gross Crashes and Crashes per Capita by Block Group Employee Density	91
Table 6: Crash per Acre by Zoning Category	95
Table 7: Crash per Acre by Land Use Mix	96
Table 8: Count, Area, and Street Miles for Each Trip Generator Category	98
Table 9: Crash Frequency	98
Table 10: Crash Density	99
Table 11: Crash Rate (pedestrian)	101
Table 12: Crash Rate (vehicular)	103
Table 13: Sum-of-Ranks Scores	105
Table 14: Crash Score Ranks	107
Table 15: Logistic Regression Output	110

FIGURES

LIST

Figure 1: Vehicle Speeds and Pedestrian Safety	25	FIGURE 18: Locations of Senior Living Facilities	71
Figure 2: 2017 High Injury Network	27	FIGURE 19: Locations of Police and Fire Stations	73
Figure 3: Focus on the Five	28	Figure 20: Detail View of School Locations, Walksheds, and Street Miles	77
Figure 4: Fatal, Severe and non-Severe Injury Collisions in San Francisco, 2013-15	29	Figure 21: Histogram of Population Density	83
Figure 5: Mode Share for All Trips in San Francisco	33	Figure 22: Population Density by Block Group, 2011-2015	85
Figure 6: Average Nearest Neighbor Output	36	Figure 23: A Comparison of Crashes per Acre and per Capita (population)	88
Figure 7: Incremental Spatial Autocorrelation Output	37	Figure 24: Employee Density by Block Group, 2011-2015	89
Figure 8: Hot Spot Analysis of Automobile-Pedestrian Collisions, 2011-2015	39	Figure 25: A Comparison of Crashes per Acre and per Capita (employee)	90
Figure 9: Kernel Density Estimation Output of Automobile-Pedestrian Collisions, 2011-2015	42	Figure 26: Zoning Districts in San Francisco	93
Figure 10: Locations of San Francisco Unified School District Schools	55	Figure 27: Land Use Mix by Block Group	94
Figure 11: Locations of Parks and Open Spaces	57	Figure 28: Sum-of-Ranks Equation	104
Figure 12: Locations of Libraries	59	Figure 29: Crash Score Equation	106
Figure 13: Locations of Medical Centers	61		
Figure 14: Locations of Rapid Transit Nodes	63		
Figure 15: Location of NCT Districts	65		
Figure 16: Locations of NCD Districts	67		
Figure 17: Locations of Parking Lots and Garages	69		

SUMMARY

EXECUTIVE



IDENTIFYING THE RELATIONSHIP BETWEEN LAND USE AND PEDESTRIAN SAFETY IN SAN FRANCISCO

This report uses a geographic information system (GIS) approach to better understand pedestrian safety in San Francisco between 2011 and 2015. Whereas pedestrian safety research has largely been focused on identifying the relationship between safety and right-of-way characteristics, the research in this report instead focuses on land uses adjacent to the right-of-way, as well as adding a demographic component to the analysis.

Two primary questions were addressed by the research:

What is the relationship between population density, employment density, and land use mix with pedestrian safety in San Francisco?

Is there a correlation between a pedestrian's proximity to specific land uses and the risk of being involved in a collision?

Three spatial statistical methods were used. Method one consisted of aggregating the totality of automobile-pedestrian crashes in San Francisco between 2011 and 2015. This dataset was integrated into a GIS shapefile that quantified the population density, employment density, and land use mix at the block group level. Population density and employment density was measured per acre, and land use mix was quantified by dissolving San Francisco's complex zoning designations into five distinct categories: residential, commercial, mixed use, industrial, and public. Block groups scored between a one and five according to the number of unique land use categories that are inside their boundaries.

Method two consisted of generating a trip attractor typology and measuring several individual and composite metrics to find correlations between land uses and pedestrian safety. Ten categories were included in the typology. Examples include schools, transit nodes, and senior living facilities. There were three individual metrics: crash frequency, crash density, and crash rate. Crash frequency

quantifies the number of crashes, crash density normalizes crash frequency at the per mile basis, and crash rate normalizes crash frequency by pedestrian exposure. Two composite metrics were also included. Sum-of-ranks combined the individual metrics to create a single score and crash rate normalized the individual metrics before combining them.

buffers to count the number of locations within each crash's service area. Additional independent variables were re-coded in SPSS.

Walksheds, or service areas, of 250 feet were created around each trip generator site. The walksheds were created using Network Analyst to more accurately model true walking distances. By intersecting the crash dataset with these walksheds, this report counted the number of crashes that occur within one short city block from each location.

Method three introduced a binary logistic regression model. Since official police records include injury severity, the dataset was re-coded to create a binary injury severity dependent variable. Injuries were coded as either "no visible injury" or "visible injury." Nearly twenty independent variables were tested. These variables included the trip generator typology and zoning designations, as well as crash location details, citations written by the responding officer, and time of day.

To create the dataset for the regression model, 1/8th mile walksheds were drawn around the location of each crash site. Land uses and zoning district shapefiles were intersected with these

Findings

The methodology described above yielded significant findings. They illustrated some striking relationships between certain land uses and pedestrian safety, as well as clearly demonstrated the impact of normalizing data on a study's findings.

Population Density, Employee Density, and Land Use Mix

When measured at the per capita level, block groups at either density level extreme were shown to be the least safe to pedestrians, although the densest block groups had the highest crash frequency. The positive relationship was tempered at the higher density levels, though. This means that even if crash rates rise as density rise, they do not rise as quickly once a “terminal density” is reached.

Employee density showed a similar relationship as population density, however employees were clearly safest at the per capita level in the densest block groups. There are temporal and demographic factors that may explain these findings. Central business districts (CBDs) during working hours are primarily composed of mobile, risk adverse seniors, unlike areas that attract young people or seniors.

Block groups with completely homogenous or heterogenous land use mixes had the most crashes. This means that block groups with moderate land use mix were safest. One can only infer as to why this was. One explanation is that diverse block groups create the greatest number of interactions between cars and walkers, and the monotony of single land use block groups lull travelers into a false sense of security.

Trip Generator Typology

Different conclusions can be drawn when different metrics are considered. For example, parks have a far higher crash frequency than any other land use. This finding is less disconcerting when one realizes that at the same time parks have the lowest crash density. This is because there are lots of parks, and opportunities for crashes near them, but the ubiquity of parks means that one sees fewer crashes at the per mile basis.

Neighborhood Commercial Districts (NCDs) were the most unsafe for pedestrians according to the crash score metric. Neighborhood Transit Districts (NCTs) fared significantly better, so more should be done to understand the differences between these two zoning districts. One potential factor is that NCTs call for less parking supply than NCDs.

Senior living facilities, libraries, and police and fire stations were considered the safest land uses. Only senior living facilities that offered independent living housing solutions were included in the analysis. Even so, perhaps few seniors walk near the facility and instead rely on other forms of transportation. Libraries attract a diverse clientele and are located across the city in a variety of settings. It is heartening that libraries scored as well as they did. Police and fire stations generate a high number of heavy vehicle trips, but were still considered to be safe. This is probably due to on-vehicle safety measures, as well as changes to the right-of-way that mitigate pedestrian interactions with trucks and fast-moving cars.

Logistic Regression

Non-land use variables had the strongest predictive ability in the model. Driver intoxication and a pedestrian being cited at the scene both correlated in a statistically significant way with a pedestrian being involved in a crash that results in visible injury. Unfortunately, a driver being cited for a “Focus on the Five” violation did not withstand mathematical scrutiny. Focus on the Five, a Vision Zero-related program, calls for at least half of the moving violations written by each police station be for one of five driving behaviors: red light running, stop sign running, not yielding while turning, not yielding to a pedestrian in a

crosswalk, and speeding.

Other significant non-land use related variables were intersection crashes and traffic calming measures. Mid-block crashes, more so than intersection crashes resulted in visible injury. And the introduction of traffic calming measures, like speed humps and bulb outs, within 1/8th mile of a crash location correlated with safer conditions.

For land use independent variables, as the share of public and industrial land increases so does the likelihood of being in a crash that results in visible injury. Once again NCD also correlated with unsafe conditions, although unlike public and industrial zoning districts, NCDs were not statistically significant.

Recommendations

This report has several recommendations:

1. **Schools** see a disproportionately high crash rate, having more crashes per pedestrian or per vehicle than other trip generators. Because this study’s findings supported the “Strength in Numbers” theory, Safe Routes to Schools programs should be expanded and children should be encouraged to walk to school in groups with their peers.

2. **NCD** districts scored poorly across every

metric. This report recommends that on-street parking supply should be minimized in NCD districts. San Francisco's demand responsive pricing scheme should be aggressively expanded in NCD districts to minimize circling the block looking for on-street parking which results in distracted driving.

3. Although **parking lots** did not fare much better than NCD districts according to the metrics, they did score safer. Transportation Network Companies (TNCs) have already relaxed the demand for parking in the city's garages, so the city would be well-served to re-route on-street parking demand to off-street locations. Parking lot ingress and egress must be located off of high-demand pedestrian streets, but close enough to still access them by foot.

4. **Parks** require more study for San Francisco's parks have massive variation in size, amenities, and location. A pocket park in a dense downtown area will generate different kinds of trips than vast open spaces to the south or west. Future research should separate parks and open spaces into sub-categories for a more refined analysis.

5. NCT districts scored safer than NCD districts. They already forbid the expansion of off-street parking supply, but more must be done to address on-street parking. Past transit projects have acquiesced to the populist demand for on-

street parking. San Francisco's policymakers must take a safety first approach to counteract these supply side tendencies.

6. The city should be looking to expand its supply of **rapid transit nodes** in and outside of NCT districts. Rapid transit nodes, often located within NCT districts, scored safer than NCT districts as a whole. This means that pedestrians in NCT districts are safer the closer they are to a rapid node.

CHAPTER ONE



INTRODUCTION: WHY IS PEDESTRIAN SAFETY WORTHY OF STUDY?

Road safety is a topic of vital importance for researchers across a vast spectrum of professional fields. Transportation planners and traffic engineers have traditionally considered the relationship between crash propensities and road network characteristics, while public health researchers have attempted to better understand the numerous social and economic costs related to crashes. As the body of knowledge grows, new policies have been formulated and introduced to protect human life. The advent of the seatbelt, the introductions of airbags, public education campaigns, and re-thinking the relationship between speed and safety have all brought down the rate of fatal automobile crashes.¹

Although the frequency and severity of automobile crashes have declined, the same is not true of automobile-pedestrian collisions. The United States has seen a 27 percent increase in pedestrian fatalities in the past decade. Walkers now account for 16 percent of all crash deaths, up from 11 percent in 2007.² Pedestrians, the most vulnerable of all travelers, are injured every six minutes and killed every 107 minutes in the United States.³ Researchers are hesitant to make causal links, but the growth in pedestrian death has mirrored the country's climb out of the Great Recession. As the economy has rebounded, vehicle miles travelled (VMT) has risen and cell phone use has grown in ubiquity, perhaps increasing the proportion of distracted drivers. American citizens are also getting older. Reflexes slow and mobility suffers as one ages.⁴ Research has routinely shown that seniors are the highest risk group to be victims of a fatal crash.⁵

1 Patricia C. Dischinger, et al., "Declining Statewide Trends in Motor Vehicle Crashes and Injury-Related Hospital Admissions" *Annals of Advances in Automotive Medicine* 57, (2013): 247-256.

2 Richard Retting, "Pedestrian Traffic Fatalities by State: 2017 Preliminary Data," *Spotlight on Highway Safety* (2017): 1- 38.

3 National Highway Traffic Safety Administration, *National Pedestrian Crash Report*, 2008.

4 Isabelle Tournier, Aurelie Dommès, and Viola Cavallo, "Review of Safety and Mobility Issues Among Older Pedestrians," *Accident Analysis & Prevention* 91, (2016): 24-35.

5 George Dunbar, "The Relative Risk of Nearside Accidents is High for the Youngest and Oldest Pedestrians," *Accident Analysis & Prevention* 45, (2012): 517-521; Aurelie Dommès, et al., "Crossing a Two-Way Street: Comparison of Young and Old Pedestrians," *Journal of Safety Research* 50, (2014): 37-34; Robert J. Schneider, Jason Vargo, and Aida Sanatizadeh, "Comparison of US Metropolitan Region Pedestrian and Bicyclist Fatality Rates," *Accident Analysis & Prevention* 106, (2017): 82-98.

In concert with Euclidean zoning, the automobile has been instrumental in shaping the modern American landscape. Turn of the century cities and their nascent suburbs were largely built along streetcar trunks, until the car opened up development in the once vacant interstitial zones between streetcar lines. The diffuse development patterns encouraged vehicle use in these ever-expanding areas. VMT, car speeds and volumes increased as new auto-centric infrastructure was built across the country. The performance of a road was (and in most cases still is) measured by vehicle throughput, or a road's level-of-service (LOS).

Progressive Californian planners have slowly started to consider a road's performance using new metrics. LOS, long the metric of choice for traffic engineers, skews toward recommending supply side solutions to solving congestion problems, which in turn induces more travel, kicking in a feedback loop of ever growing congestion. Further, LOS does little to consider the environmental impacts of travel. VMT is a new metric that has replaced LOS across California.⁸ When using VMT, planners can consider greenhouse gas emissions (GHGs) and the performance of active and public transportation along corridors and through intersections.

Like GHG, pedestrian deaths and serious injuries were once thought of as unfortunate

yet inevitable byproducts of the transportation system. Public education campaigns and legal solutions were common responses to pedestrian death. With its roots in the 1970's-era "Stop de Kindermoord" ("Stop the Child Murder") campaign in the Netherlands, which culminated in the Vision Zero movement, pedestrian safety is now conceptualized as a public health issue. Proponents of the public health approach argue that pedestrian fatalities are preventable and that the effects of crashes permeate more than just the transportation system. This fact becomes clear when one considers that in San Francisco alone, collisions cost San Francisco General Hospital \$35 million per year and half of all patients treated there are victims of a crash.⁹

8 Melanie Curry, "San Jose Becomes Fourth California City to Adopt VMT as Metric for Traffic Impacts," Streetsblog California, <https://cal.streetsblog.org/2018/03/06/san-jose-becomes-fourth-california-city-to-adopt-vmt-as-metric-for-traffic-impacts/> (accessed March 3, 2018).

9 Nuala Sawyer, "Traffic injuries cost patients \$35 million per year at San Francisco General," SF Examiner, November 17, 2016; Population Health Division of the San Francisco Department of Public Health. "2017 High Injury Network Update." Policy and Governance Committee hearing, San Francisco Municipal Transportation Agency, San Francisco, July 21, 2017.

1.1 Research Questions and Introduction of Methods

Pedestrian safety has largely been considered in the context of built form characteristics and human behavior. Although urban planners understand the interrelatedness of land use and transportation, most research has failed to consider the link between land uses adjacent to a right-of-way and collisions. This report endeavors to answer the following questions:

What is the relationship between population density, employment density, and land use mix with pedestrian safety in San Francisco?

Is there a correlation between a pedestrian's proximity to specific land uses and the risk of being involved in a collision?

To answer these questions, three distinct methods, each utilizing several spatial and statistical techniques, will be deployed in this report.

- Method one is a comparison of pedestrian crash locations with demographic and built form characteristics at the block group level in San Francisco. Specifically, a block group's population density, employment density, and land use mix will be quantified. A dataset containing every automobile-pedestrian collision in San Francisco between 2011 and 2015 will be created and integrated with the population density, employment density, and land use mix calculations.

- Method two utilizes a geographic information system (GIS) to explore the relationship between land use and pedestrian safety. For this method, a trip generator typology was created. Land use types were selected for the trip generator typology if they are known to produce pedestrian trips in heterogeneous, mixed-traffic areas. Using Esri's ArcGIS software suite, walksheds were created around each trip generator location. By assessing the relationship between crashes and trip generator walksheds, this report endeavors to glean if particular land use types were correlated with more crashes in San Francisco between 2011 and 2015.

- Method three is a logistical regression. Several built form and demographic variables will be included to further elucidate the relationship between land use and pedestrian safety in San Francisco.

1.2 How to Read This Report

This planning report is laid out as follows: chapter two provides a brief overview of Vision Zero – an international framework that strives for the elimination of all traffic-related serious injuries and deaths. The concept originated in Sweden, but San Francisco, along with numerous American cities, has introduced the program at the local level. Beyond discussing the Vision Zero approach to road safety, this report describes the context in which San Francisco started the program, as well as the specific policies and tools that the city has deployed. Chapter two concludes by explaining how the original research in this report can complement existing traffic safety findings.

Chapter three begins the spatial statistical discussion. Several tools, including Average Nearest Neighbors and Spatial Autocorrelation, are explained. After describing the purpose of the relevant tools, Hot Spot and Kernel Density analyses are completed. These workflows help illustrate the spatial orientation of pedestrian crashes in San Francisco.

Chapter four is a detailed description of the three methods described above. Prior research is integrated into the discussion to support the substantial methodological choices made throughout the study. The workflow for each of the three methods is described individually before moving on to a discussion of the findings.

Using a combination of tables, figures, and text, chapter five presents the findings from the three methods in chapter four. The raw data and its implications are considered.

Chapter six concludes the report with a summary and further discussion of the findings. Study limitations are identified in a larger discussion of suggestions for future research.

TWO

CHAPTER



VISION ZERO IN SAN FRANCISCO: A HOLISTIC APPROACH TO ROAD SAFETY

Vision Zero was adopted by the Swedish parliament in 1997 as a response to the “violence” of severe injuries or deaths on the nation’s transportation infrastructure.¹⁰ The program aims to re-frame how society treats traffic safety. Crashes should no longer be thought of as “accidents,” and human error should not be thought of as the primary cause for these incidents. Creating safer streets requires a systems thinking approach: the marriage of a public health framework with a re-tooling of traffic engineering orthodoxy.¹¹ Injuries and deaths should be deemed unacceptable and no longer considered an unavoidable byproduct of traffic engineering that intends to maximize vehicle throughput.

Twenty years after its genesis, the Vision Zero toolkit has grown into a holistic, data-driven response to unsafe streets. Planners and engineers have numerous methods to reduce injuries or deaths that result from a crash. These include physical changes to the right-of-way. Pedestrian bulb outs, raised crosswalks, roundabouts, and the strategic placement of bollards are just a few design options.¹² Innovative technologies have

also been introduced. Red light cameras and remote speed enforcement have both proven capable of reducing vehicle speeds and promoting road safety.¹³

10 Sarah Goodyear, “The Swedish Approach to Road Safety: ‘The Accident is not the Major Problem,’” CityLab, <https://www.citylab.com/transportation/2014/11/the-swedish-approach-to-road-safety-the-accident-is-not-the-major-problem/382995/> (Accessed February 20, 2018).

11 Ellen Kim, Peter Muennig, and Zohn Rosen, “Vision Zero: A Toolkit for Road Safety in the Modern Era,” *Injury Epidemiology*, 4 (2017).

12 Roger Johansson, “Vision Zero – Implementing a policy for traffic safety.” *Safety Science* 47 (2009): 826–831.

13 New York City Department of Transportation, “Vision Zero Action Plan, 2014,” New York City, <http://www.nyc.gov/html/visionzero/pdf/nyc-vision-zero-action-plan.pdf> (accessed March 4, 2018).

AN ASSORTMENT OF VISION ZERO SAFETY ENHANCEMENTS



Curb bulb outs shorten the crossing distance for pedestrians.



Soft hit posts and contrasting tan paint daylight the intersection. Daylighting treatments are added to intersections to increase visibility for both pedestrians and motorists.



Center medians act as pedestrian refuge on wide streets.



Raised intersections slow down vehicular traffic and create level crossing for pedestrians.

Traditional road safety frameworks only consider the actions of road users when assessing a crash. To counteract unsafe behaviors, new laws are passed or education campaigns are deployed. The crash is considered in a vacuum, rather than as an outcome of the transportation system as a whole. Vision Zero thinking is radically different. Vision Zero considers the role of both the system and road users.¹⁴ It is incumbent that engineers design roads to promote safety, that manufacturers create safer vehicles, and that traffic engineers respond to dangerous collisions by making physical changes to the roadway.

As Figure 1 below shows, managing vehicle speeds is paramount to protecting pedestrians. An increase of only ten miles per hour can mark the difference between life and death.

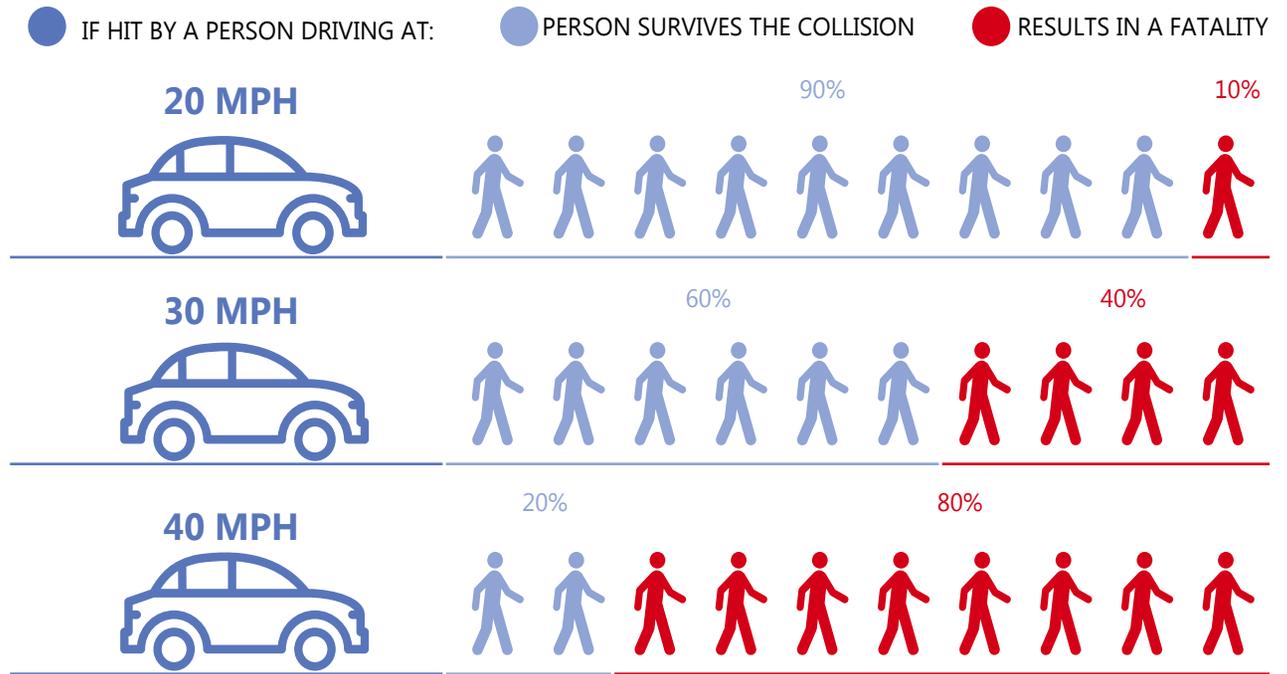


Figure 1: Vehicle Speeds and Pedestrian Safety

Source: <https://www.sfmta.com/projects/automated-speed-enforcement> (3/1/2017)

14 Claes Tingvall and Narelle Haworth, "Vision Zero- An Ethical Approach to Road Safety," International Conference Road Safety and Traffic Enforcement, (1999).

2.1 Origins of Vision Zero in San Francisco

The City and County of San Francisco enacted its first coordinated Vision Zero effort in 2014 when officials created a two-year action plan. The movement towards safer streets was further galvanized by the tragic deaths of Heather Miller and Katherine Slattery, two cyclists who were struck by automobiles in separate locations in the city within two hours of one another on June 22, 2016.¹⁵ Mayor Ed Lee, incensed by both events, responded by saying, “[t]hese aren’t accidents. They are tragedies that can be prevented.”¹⁶ Fewer than two months after the cyclists’ deaths, Mayor Lee announced an Executive Directive that would hasten the implementation of Vision Zero projects throughout the city.

2.2 High Injury Network and “Focus on the Five”

The San Francisco Department of Public Health (SFDPH) and the San Francisco Municipal Transportation Agency (SFMTA) lead the joint task force that addresses street safety. In concert with over forty local organizations that include San Francisco Unified School District and non-profit advocacy organizations such as Walk SF and the San Francisco Bicycle Coalition, the Vision Zero task force has been instrumental in enacting forward thinking initiatives that support safer streets. The task force’s influential strategies, such as the High Injury Network Map, have since been introduced in several American cities.¹⁷ San Francisco’s High Injury Network Map is an annually updated tool that highlights the corridors accountable for a disproportionate number of the crashes that result in severe injury or death. The city has focused its capital improvement projects on high injury road segments.. The most recent update from 2017 identifies 13 percent of city streets that have accounted for 75 percent of serious injuries and deaths.¹⁸

15 Office of the Mayor, “Mayor Lee Issues Executive Directive on Bicycle and Pedestrian Safety,” Office of the Mayor, August 4, 2016.

16 Kevin Schultz and Lizzie Johnson, “SF Mayor Outraged by Hit-and-Run Death of 2 Bicyclists,” SFGate, <https://www.sfgate.com/bayarea/article/Two-bicyclists-killed-in-separate-SF-hit-and-run-8321105.php> (accessed October 20, 2017).

17 “Pedestrian Safety and Vision Zero in San Francisco: State Climate Action Team Public Health Workgroup,” <https://www.arb.ca.gov/cc/ab32publichealth/meetings/100814/wier.pdf> (accessed February 20, 2018).

18 “San Francisco Department of Public Health, “VISION ZERO HIGH INJURY NETWORK: 2017 UPDATE,” https://www.sfdph.org/dph/files/EHSdocs/PHES/VisionZero/Vision_Zero_High_Injury_Network_Update.pdf (accessed March 10, 2018).



Figure 2: 2017 High Injury Network

Map created by author. Adapted from: <http://visionzerosf.org/maps-data/> (accessed January 2018)

In addition to the High Injury Network, San Francisco has also implemented its “Focus on the Five” campaign: a traffic enforcement program that calls for local police to take extra care to cite drivers for infractions that are correlated with severe injuries or fatalities, rather than infractions that are more administrative in nature. Refer to Figure 3 for the five driver behaviors police are to be aware of.

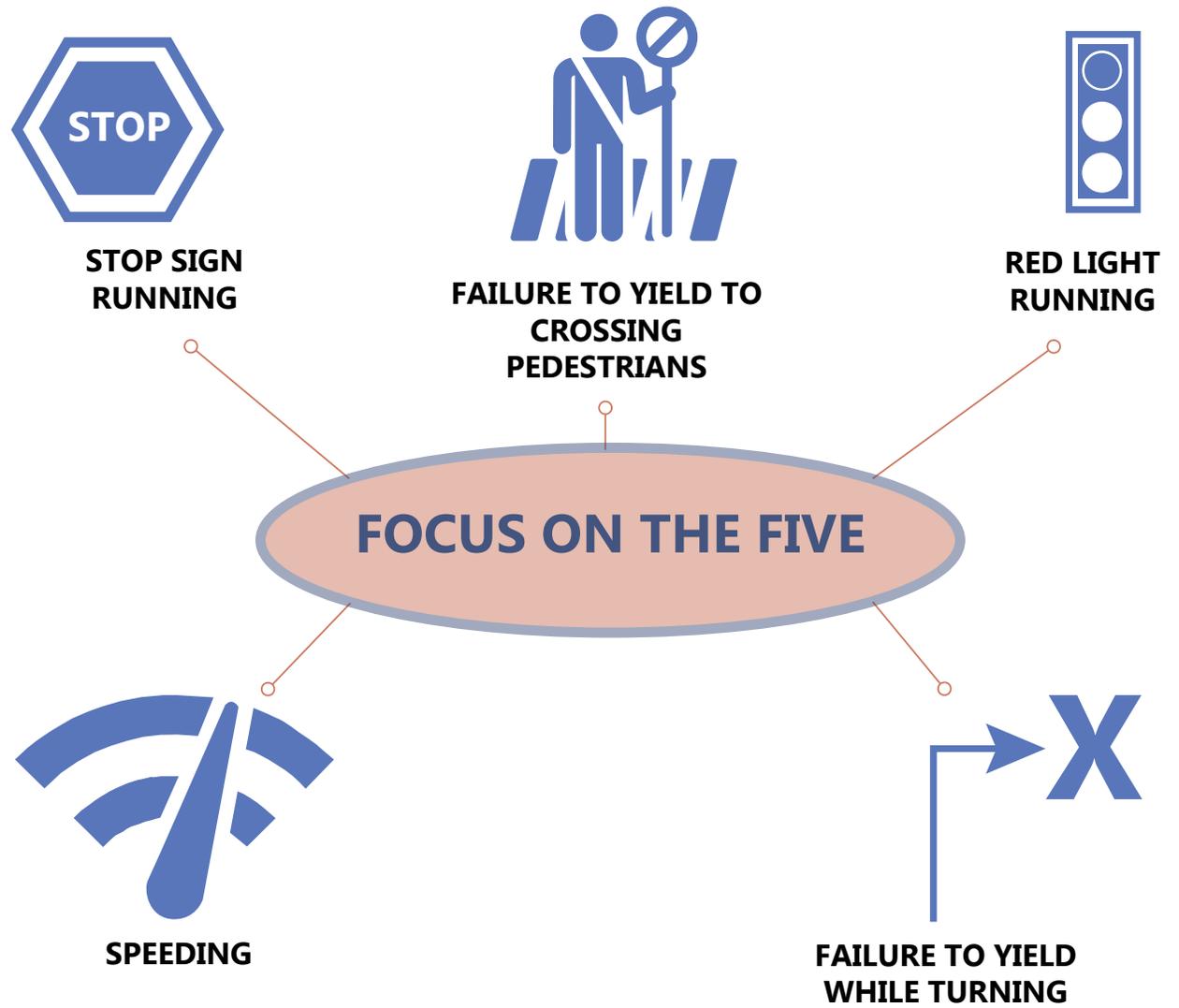


Figure 3: Focus on the Five
Graphic created by the author

The Focus on the Five initiative directs officers to patrol high injury segments more regularly than streets where fewer collisions occur. Although it is prudent to minimize behaviors that most often result in serious injuries, the Vision Zero methodology emphasizes street design over law enforcement solutions. Material changes to the roadway that make driving, walking, and cycling safer will free up an already taxed police force to focus on other crimes. Additionally, directing enforcement to the High Injury Network means that most city streets are not being patrolled at the same rates.

The San Francisco Office of the Controller concurs that focusing enforcement on the High Injury Network is detrimental to overall safety.¹⁹ The department analyzed collisions between 2013 and 2015 and found that several parts of the city saw significant clusters of crashes but were not included on the High Injury Network. Figure 4 shows where these clusters of collisions have occurred. The map was created by the Office of the Controller as part of a public presentation in the summer of 2017. The presenters pressed that more must be done to understand the context of crash clusters that occur off high injury corridors. By studying crashes at a localized scale rather than at the corridor level, this report aims to fill in gaps of understanding that the Office of the Controller has identified.

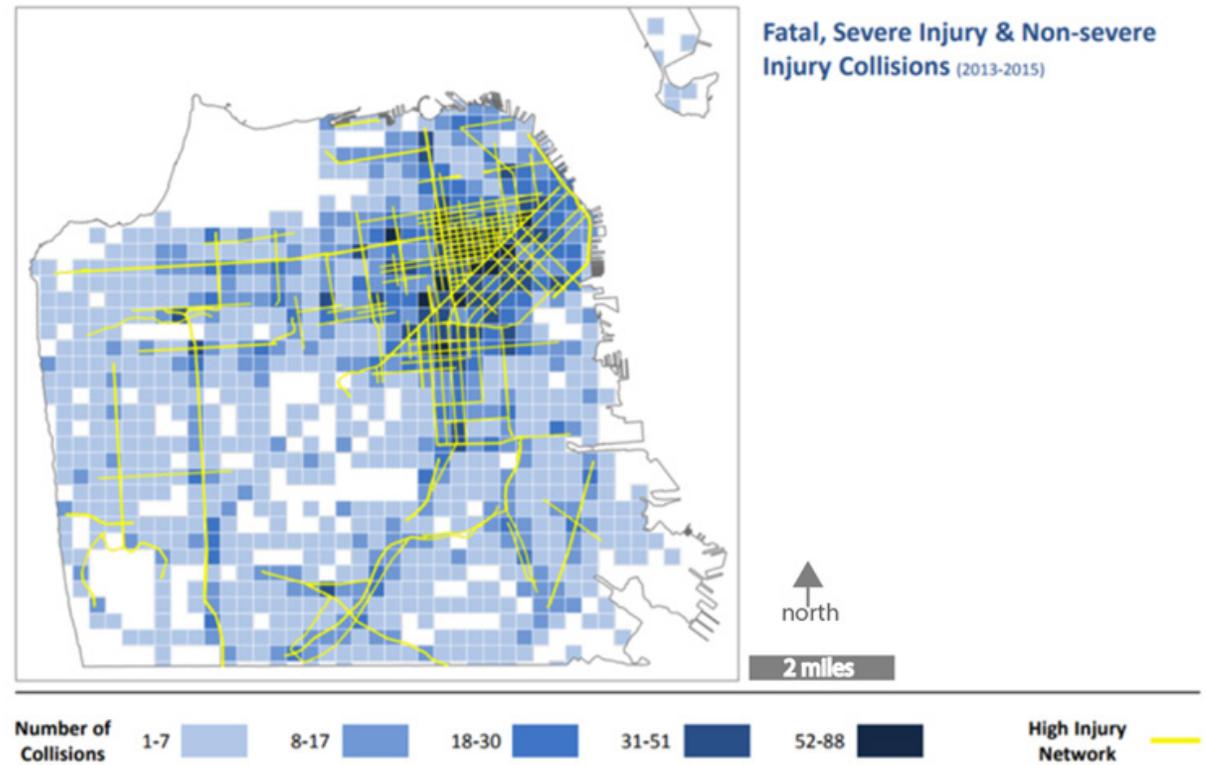


Figure 4: Fatal, Severe and non-Severe Injury Collisions in San Francisco, 2013-15

Scale bar and north arrow not in original graphic. and have been added as elements to the map.

Source: San Francisco Office of the Controller, "Focus on Enforcement: Insights from research and analysis in support of San Francisco's Vision Zero plan to eliminate traffic deaths." Policy and Governance Committee hearing, San Francisco Municipal Transportation Agency, San Francisco, July 21, 2017.

¹⁹ San Francisco Office of the Controller, "Focus on Enforcement: Insights from research and analysis in support of San Francisco's Vision Zero plan to eliminate traffic deaths." Policy and Governance Committee hearing, San Francisco Municipal Transportation Agency, San Francisco, July 21, 2017.

Like the High Injury Network and the Focus on the Five initiative, this report is the result of a data-driven approach to better understanding road violence. Beyond adding to the local Vision Zero discussion, the findings in this report complement existing pedestrian safety literature. Although existing research is thorough, there are gaps. So far, researchers have primarily focused on road and network characteristics and their relationship to road safety.²⁰ This area of research may consider, for example, how road width, intersection density, or travel speeds correlate with automobile-pedestrian collisions. For researchers more concerned with demographic variables than those associated with built form, much research has also been undertaken to find correlations between at-risk groups and crash rates.²¹ For example, studies have explored how the age, income, gender, and physical ability of crash victims have influenced crash likelihood and severity. From these studies, researchers and advocates have isolated the young, the elderly, and lower-income people as at-risk groups. Compared to road characteristics and demographic analysis, less research has been done that considers the relationship between land uses adjacent to the right-of-way and pedestrian safety.

For the purpose of this study, land uses adjacent to the right-of-way are physical locations that generate trips. Land uses can include

buildings like schools, libraries, or medical centers; undeveloped parcels like parks, open spaces, or surface level parking lots; or zoning designations. Specific zoning designations, such as Neighborhood Commercial District (NCD), influence the kinds of physical development within the zoning district's boundaries. By focusing on land uses adjacent to the right-of-way, this report endeavors to address gaps in the research and add to the pedestrian safety conversation.

Now that the context is clear, the following chapter will describe the study area, the data utilized in the study, and a first cut analysis of where crashes have happened in San Francisco between 2011 and 2015. It is important to first understand broadly where crashes occur before exploring the granular details of the land use and demographic variables. The following chapter will also ascertain whether crashes have been clustered together or have been spread out evenly throughout the study area. If crashes are randomly dispersed around San Francisco, it will be more difficult to find correlations between specific land uses and the likelihood of a crash.

20 Peter Swift, Dan Painter, and Matthew Goldstein, "Residential Street Typology and Injury Accident Frequency." *Congress for the New Urbanism* (2006): 1-8; Kelly J. Clifton, Carolina V. Burnier, and Gulsah Akar, "Severity of injury resulting from pedestrian-vehicle crashes: What can we learn from examining the built environment?" *Transportation Research Part D: Transport and Environment* 14 (2009): 425-436; Roger Johansson, "Vision Zero – Implementing a policy for traffic safety." *Safety Science* 47 (2009): 826-831.

21 Dilum Dissanayake, James Aryajaja, and D.M. Priyantha Wedagama, "Modelling the Effects of Land Use and Temporal Factors on Child Pedestrian Casualties," *Accident Analysis and Prevention* 41, (2009): 1016-1024; Aurelie Dommès, et al., "Crossing a Two-Way Street: Comparison of Young and Old Pedestrians," *Journal of Safety Research* 50, (2014): 37-34; Caitlin D. Cottrill and Piyushimita (Vonu) Thakuriah, "Evaluating Pedestrian Crashes in Areas with High Low-Income or Minority Populations," *Accident Analysis & Prevention* 42, no. 6 (2010): 1718-1728; Kathryn M. Neckerman, et al. "Disparities in Urban Neighborhood Conditions: Evidence from GIS Measures and Field Observation in New York City," *Journal of Public Health Policy* 30, (2009): 264-285.

THREE

CHAPTER



MAPPING CLUSTERS OF AUTOMOBILE-PEDESTRIAN CRASHES IN SAN FRANCISCO

Chapter three orients the reader to the study area and describes the preparation for, and execution of, several spatial analyses. These techniques were used to assess where pedestrian crashes have occurred in San Francisco between 2011 and 2015 and if these crashes show statistically significant spatial clustering.

After concluding that crashes do show spatial clustering, chapter four will describe the three primary methods that were used to further clarify the relationship between demographic factors, land use, and collisions.

San Francisco, California is a densely settled population and employment center that is located on the northern tip of the San Francisco Peninsula. Forty-seven square miles in area, and with a resident population near 900,000, it is second only to New York City in population density amongst America's major cities. As the tourist center in a region of five million people, San Francisco's daytime population reaches over one million people each weekday.²² Because of its density, diverse land use mix, and plentiful public transportation options, the city has a

high pedestrian mode share for all intracity trips. With numerous conflict points between automobiles and pedestrians, San Francisco is a prime location to better understand pedestrian safety.

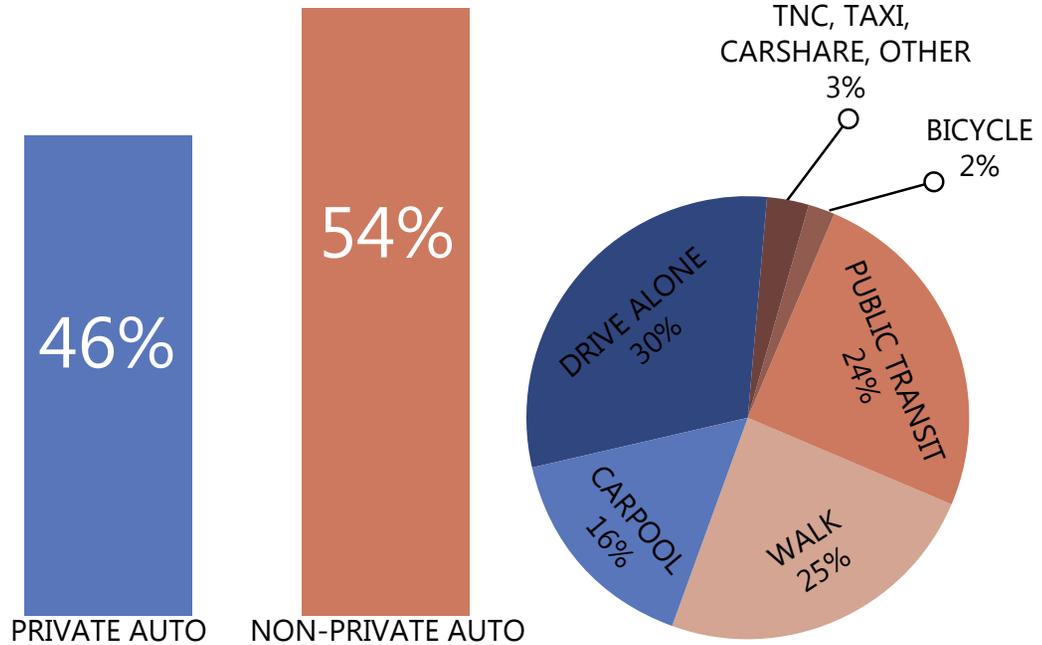


Figure 5: Mode Share for All Trips in San Francisco

Graphic created by author

Source: SFMTA 2013-2018 Strategic Plan

22 San Francisco Office of the Controller, Citywide Benchmarking Report: Demographics, Livability, Public Safety, Transportation, Finance, Safety Net, Population Health, February 16, 2017.

3.1 Aggregating Automobile-Pedestrian Crashes

This study uses a geographic information system (GIS) to better understand automobile-pedestrian collisions between 2011 and 2015 in the City and County of San Francisco. Treasure Island and the San Francisco International Airport, both within the purview of city and county government, were excluded from the analysis since they are physically separated from the core city and are not representative of San Francisco's development patterns.

Esri's ArcGIS software suite was used to analyze crashes in San Francisco. A GIS allows one to visualize data in its spatial form and makes spatial patterns more discernable than viewing the same data on a spreadsheet. In addition to utilizing common geoprocessing tools, this study utilized the Network Analyst extension and Esri's Spatial Statistics toolset.

A crash dataset was integrated into ArcMap. The collision data was downloaded from www.transbasesf.org, a public-facing website that is organized and regularly updated by the San Francisco Department of Public Health. TransbaseSF was created after the adoption of Vision Zero in San Francisco to inform the public where collisions have occurred and

where safety-promoting capital improvement projects are being planned and implemented. TransbaseSF does not allow direct exporting or downloading of data, so a database server connection in ArcMap was created to export the city's data directly to an external hard drive.

TransbaseSF includes collisions involving all travel modes dating back to 2005, so considerable cleaning of the data was completed before settling on a final dataset. This involved removing non-pedestrian crashes, and merging multiple datasets to fully aggregate all automobile-pedestrian crashes between 2011 and 2015. 15,000 collisions were identified.

Since this study is only concerned with pedestrian crashes, non-pedestrian collisions had to be removed from the dataset. The "type_of_collision" field in the dataset's attribute table highlighted all "vehicle/pedestrian" collisions. There were 3,546 such collisions. Because the data is from police records, which can be prone to error²³, not every pedestrian collision was cataloged as a "vehicle/pedestrian" crash. Because of this, 456 misidentified crashes were found and added to the dataset. A total of 4,002 crashes between 2011 and 2015 were ultimately identified. Of these, 3,987 were successfully geocoded, meaning over 99 percent of the crashes were included in the analysis.

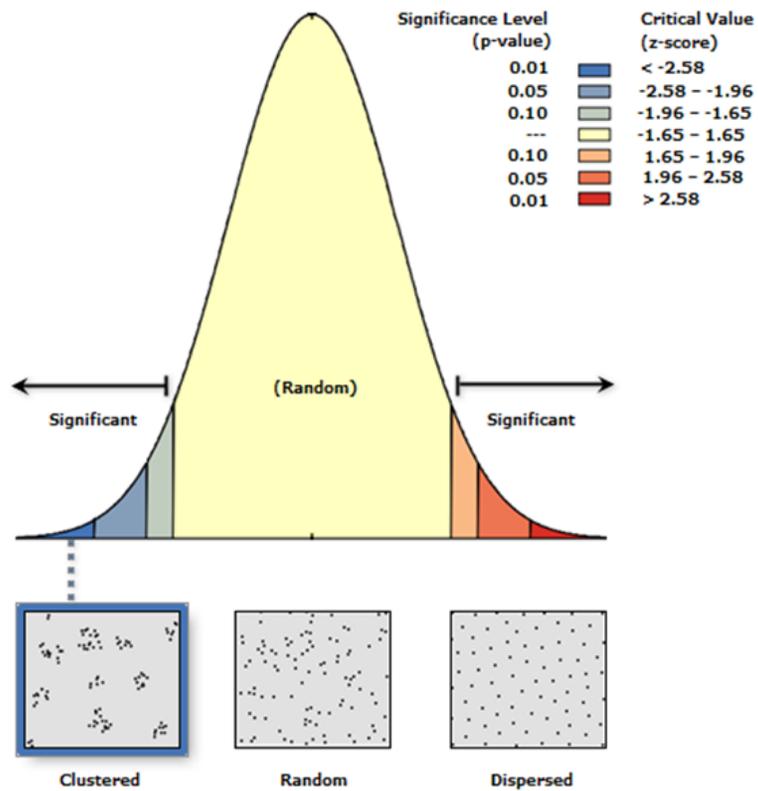
²³ Errors include the intervening officer incorrectly filling out the report at the scene or inaccurately transferring the physical report to a digital format. It should also be noted that estimates point to fewer than half of all pedestrian-automobile crashes being reported to the police. Source: Rodney Tolley, "Sustainable Transport: Planning for Walking and Cycling in Urban Environments," Woodhead Publishing in Environmental Management, 2003.

The original research in this report includes quantifying the correlation between crashes and population density, employment density, and land use mix, as well as explaining the relationship between crashes and a trip generator typology. Before considering these relationships, it is important to understand the spatial distribution of crashes in San Francisco. One must know if the universe of crashes shows significant clustering, dispersion, or randomness before being able to make any claims about specific built form and demographic relationship to crashes. If crashes are randomly distributed, claims of a positive relationship between density and crashes cannot be made.

3.2 Hot Spot Analysis and Kernel Density Estimation

There are a variety of techniques to find statistically significant relationships between physical phenomena using GIS. This report utilizes two of them: Hot Spot Analysis and Kernel Density Estimation. Hot Spot Analysis determines if crashes show statistically significant clustering, dispersion, or randomness, and Kernel Density compares nearby features to calculate the density of the features. Hot Spot Analysis affixes a significance level to each feature itself, whereas Kernel Density creates an interpolated surface. This interpolated surface is represented as a raster output on the map. The processes and outcomes of the Hot Spot Analysis and Kernel Density Estimation techniques will be described below.

There are several prerequisite steps before one can confidently carry out a Hot Spot Analysis, for a Hot Spot Analysis must be set up with the correct parameters. The Average Nearest Neighbor tool in ArcMap was used in preparation. This tool creates an output that states if the input features (crashes) are clustered, dispersed, or random. It shows *if* there is clustering rather than *where* the clustering may occur. The tool calculates the mean distance between the features and compares this calculation to the expected distance. A



Nearest Neighbor Ratio:
0.423179

z-score: -69.677872

p-value: 0.000000

Observed Mean Distance:
119.6592 US_Feet

Expected Mean Distance:
282.7624 US_Feet

Given the z-score of -69.6778719523, there is a less than 1 percent likelihood that this clustered pattern could be the result of random chance.

Figure 6: Average Nearest Neighbor Output

24 A p-score is a measure of statistical significance. It ranges from zero to one, however for most studies in the realm of the social sciences, only p-values less than or equal to 0.05 are accepted to reject the null hypothesis. A p-value of 0.05 means that one can conclude with 95 percent certainty that there is a statistically significant relationship between two phenomena.

25 A z-score, or standard score, measures a data point's value in relation to the mean. It is expressed in terms of standard deviations from the mean. Therefore, a score of one states that a value is one standard deviation above the mean.

p-score²⁴ and z-score²⁵ is calculated based on the observed and actual calculations.

The Average Nearest Neighbor tool was used for the San Francisco pedestrian crash dataset. After running the tool, crashes show clustering with less than one percent chance that the distribution of crashes is the result of randomness. A low negative z-score coupled with a low p-score means that there is significant clustering. The Average Nearest Neighbor report is shown in Figure 6.

After concluding that clustering has occurred, a new tool, Spatial Autocorrelation, is used. Expressed as Global Moran's I, spatial autocorrelation also measures if the data is clustered, dispersed, or random, but the Spatial Autocorrelation tool considers both the location and the features of each record.²⁶ The tool also measures at which distance clustering is most intense. This distance band is necessary to properly carry out a Hot Spot Analysis.

The Spatial Autocorrelation tool measures one distance band at a time and outputs a z-score for that distance. Because only one z-score is calculated at a time, finding the correct distance band is an iterative process. Rather than run the Spatial Autocorrelation tool multiple times, ArcMap includes the Incremental Spatial Autocorrelation tool. This tool creates z-scores for a range of distance bands, in effect automating the iterative process.

Incremental Spatial Autocorrelation requires accurate input parameters to generate accurate z-scores. These input parameters include the maximum distance between nearby features, as well as the average distance between incidents. This means that the tool will start at the maximum value and iterate by the average distance between features. ArcMap calculates the maximum distance between incidents with the Calculate Distance Band from Neighbor Count tool in ArcToolbox. The maximum distance was

found to be 4,000 feet. This distance was put in to the Incremental Spatial Autocorrelation tool. The tool was run at 120-foot iterations to calculate the distance band that yielded the highest z-score. 120 feet was the observed mean distance between incidents from the Average Nearest Neighbor tool. The Incremental Spatial Autocorrelation tool outputs a line chart that plots the z-score for each distance band. The highest peak correlates with the highest z-score. The higher the z-score, the more intense the clustering. The tool returned a value of 5,360 feet. The output is shown above.

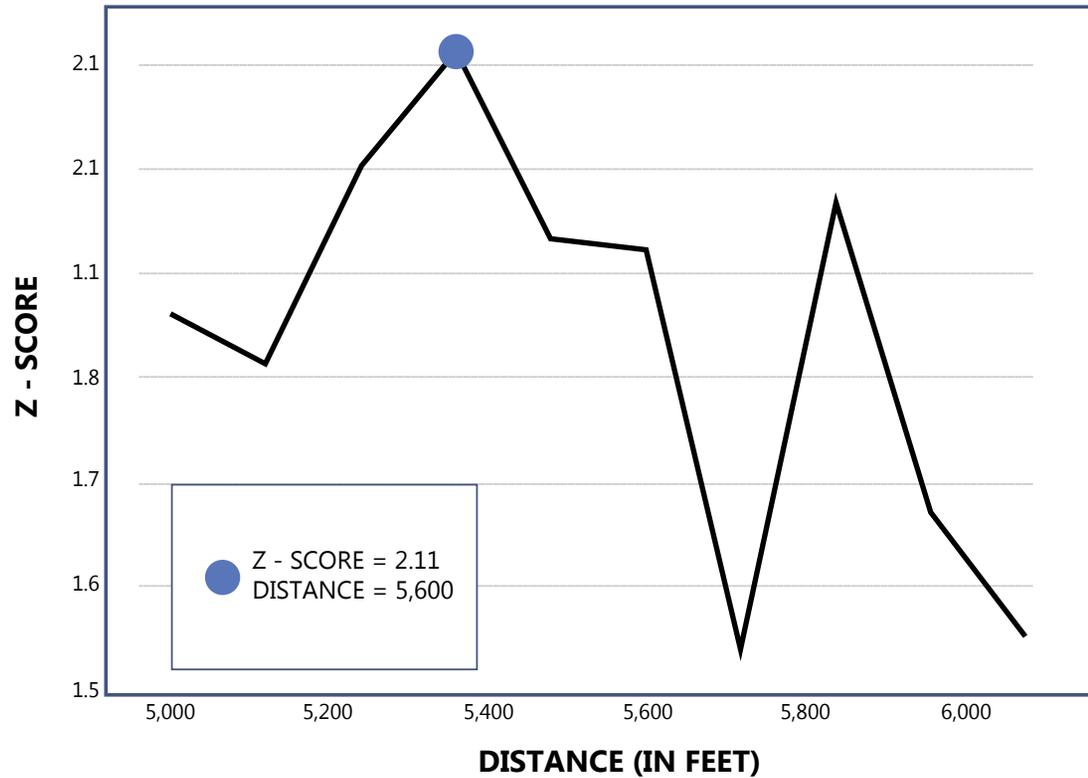


Figure 7: Incremental Spatial Autocorrelation Output

²⁶ For example, if one is measuring the spread of disease, the Spatial Autocorrelation tool considers both the locations of where people are sick, as well as the number of sick people in each location.

A Hot Spot Analysis measures the likelihood that a distribution of values happened randomly. To accomplish this, the tool requires a set of features, each with its own numeric value. In this case, each feature (a crash location) has a value that represents the number of pedestrians injured in the crash. There was a maximum of five people injured in a single crash in San Francisco during the period of study. 92 crashes had zero injuries because said crashes resulted in fatalities. Since the Hot Spot Analysis for this study does not differentiate between the severities of injuries, the dataset was manipulated to include these fatal crashes in the Hot Spot Analysis.

The result of the Hot Spot Analysis is a map layer that shows if a feature is a cold spot or a hot spot and ascribes a confidence interval to each feature. The confidence intervals are 90, 95, and 99 percent. A hot spot is the result of a feature that is in a geographic area, or neighborhood, with a value higher than the overall study area. Conversely, a cold spot is where the neighborhood in which a feature is in is a lower value than the study area.

The Hot Spot Analysis tool was used with a 5,360-foot distance band: the distance in which the most intense clustering occurs. This is the distance in which each neighborhood will be created. The results of the Hot Spot Analysis are below. Notice the intense clustering of hot spots in the mid-Market and South of Market

areas, as well as the city's southern Ingleside neighborhood. There looks to be many incidents in the southeastern section of the city, but ArcMap's analysis found those crashes to be less clustered than elsewhere in the city.

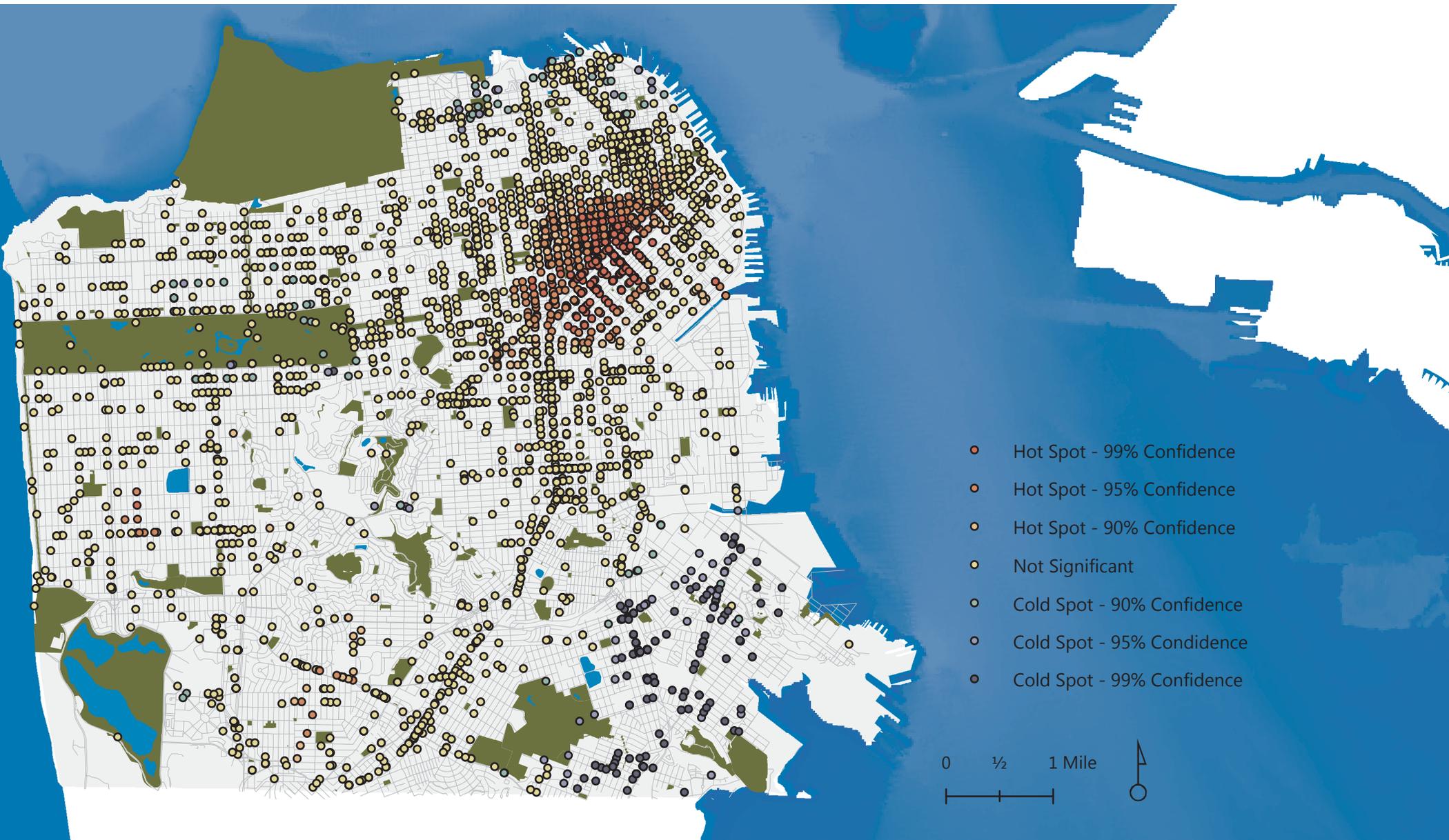


Figure 8: Hot Spot Analysis of Automobile-Pedestrian Collisions, 2011-2015

See Appendix B for list of sources

Kernel Density Estimation, like Hot Spot Analysis, considers the spatial relationship between features. It differs from Hot Spot Analysis in that the value of each feature is *not* included in the analysis. And unlike Hot Spot Analysis, Kernel Density Estimation does not affix confidence intervals to each feature. Finally, the Hot Spot Analysis layer shown above was represented as point features of varying colors, while a Kernel Density Estimation output layer is a raster representing an interpolated surface. In other words, Kernel Density Estimation creates new features based on the values of the nearby observed features. As an example, envision a city block with crash incidents at each end. A Hot Spot Analysis will not make any claims about the middle of the block for there are no data points there. However, a Kernel Density Estimation will create values for the midblock that are based on the observed data points at each end of the block.

The granularity of the Kernel Density analysis will depend on the parameters (for raster cell size and search radius) that the user sets for the tool. The result of the Kernel Density Estimation is a map layer that is easier to comprehend than a Hot Spot Analysis. The smooth gradations between concentrations in the Kernel Density Estimation allow for patterns to visually “pop” whereas the colored points in the Hot Spot Analysis require closer inspection to find meaningful patterns.

Borrowing from the methods used by Pulugurtha, et al. in their study of crashes in Las Vegas, this report used the Kernel Density method to locate concentrations of crashes.²⁷ As compared to Simple Density Analysis, the Kernel method is the best way to perform a fine-grained analysis of crash concentration. This is because Kernel analysis works by creating a buffer around each incident. All nearby incidents are aggregated in the buffer to measure how concentrated the incidents are. The Kernel method, as compared to Simple Density Analysis, refines the analysis by applying weights to the aggregated incidents. An incident’s score decreases as its distance increases relative to the central feature. This creates smoother gradations between the measurements than one would find with Simple Density Analysis.

As stated above, the granularity of the analysis is dependent upon the tool’s parameters. The GIS user can identify the output cell size to be used in the analysis, effectively dictating how fine or coarse the analysis will be. The output cell size relates to the search radius to be used for the incidents and will affect the output raster’s resolution. A large cell size will result in greater pixelization, whereas a small cell size will be less pixelated but will create a large output file. Put another way, selecting too large of a cell size will not capture the details in the data, while too small of a cell size will not properly consider the

27 Srinivas S. Pulugurtha, et al, “New Methods to Identify and Rank High Pedestrian Crash Zones: An Illustration.” *Accident Analysis & Prevention* 39, no. 4 (2007): 800-811.

relationship between nearby incidents. Therefore, Kernel Density is an iterative process. After running numerous analyses, an output cell size of 125 feet was chosen because it represents the approximate length of one-half of a short city block. 125 feet can consider mid-block and intersection level crashes together, and resulted in the smoothest raster output.

The map on the following page is the result of the Kernel Density Estimation. The deepest red polygons represent locations with the highest crash concentrations and blue areas show low concentrations. Like the Hot Spot Analysis result shown in Figure 8, mid-Market, the Tenderloin, and sections of SoMa show the most intense concentration of crashes. There are crash concentrations along numerous major corridors, too. For example, Mission Street, Taraval Street, 19th Avenue, and sections of Third Street all show varying levels of crash concentrations.

Although the map above has value, we can clearly see *where* crashes cluster, but we cannot say with any certainty *why* crashes cluster. Is the high concentration area near Civic Center simply a result of high population densities? Are crash concentrations along major thoroughfares due to high vehicle volumes? The following chapter will describe the methods used to clarify the connection between demography and land use with pedestrian crashes.

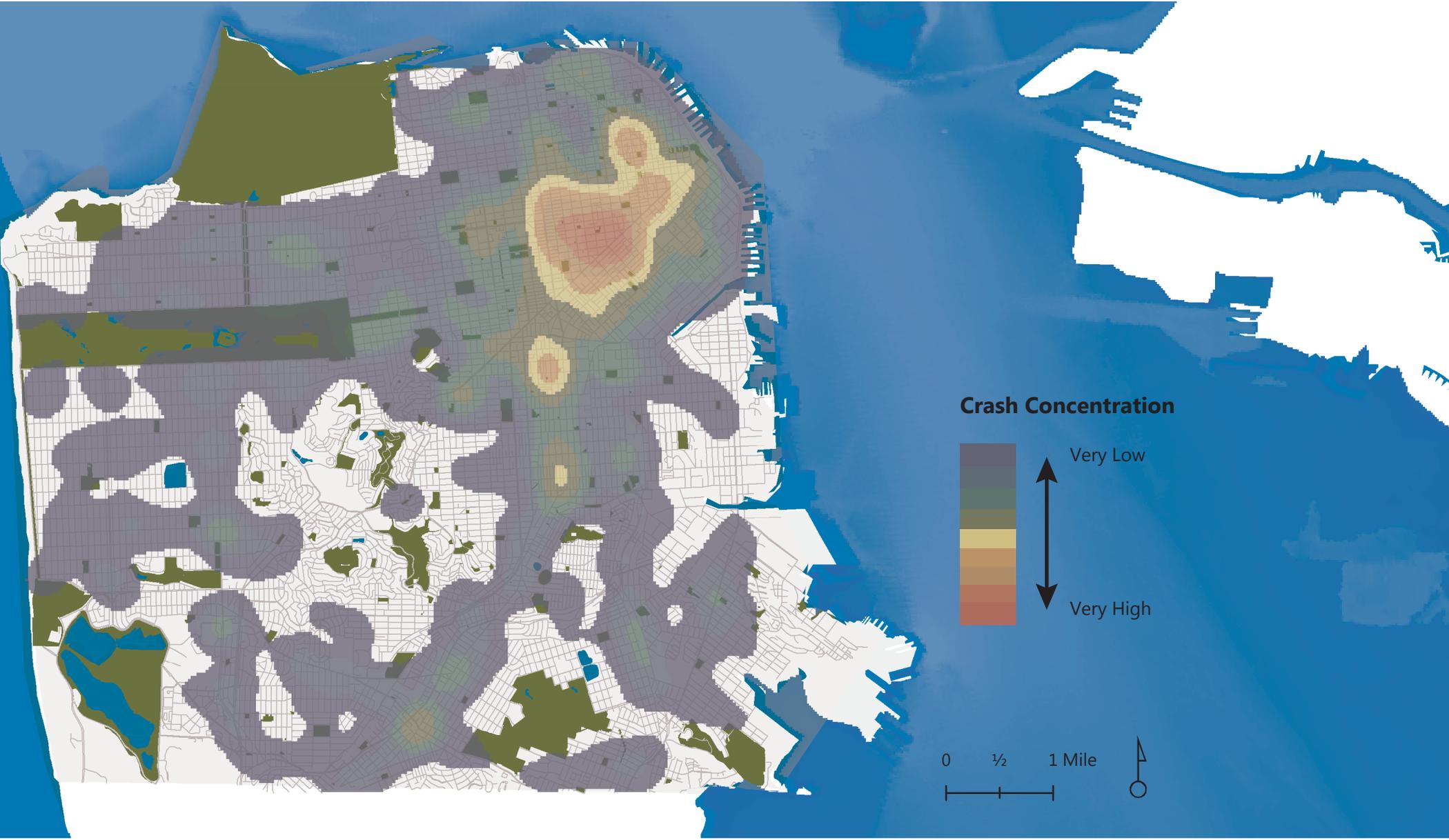


Figure 9: Kernel Density Estimation Output of Automobile-Pedestrian Collisions, 2011-2015

See Appendix B for list of sources

FOUR

CHAPTER



METHODOLOGY AND UNDERPINNINGS IN PAST RESEARCH STUDIES

While the workflows in the preceding chapter aimed to understand if spatial patterns of crashes exist at a citywide scale, the three distinct methods that will be explained below were utilized to better understand the relationship between demography, land uses, and crashes in a more localized context. From the Hot Spot Analysis and Kernel Density Estimation results, it should be clear that crashes exhibit clustering. This is a valuable insight, but it lacks the detail necessary to comprehend the issue and be able to make policy prescriptions to mitigate the likelihood of a crash.

The three methods in this chapter are:

- A comparison of crashes to population density, employment density, and land use mix at the census block group level across several metrics. The purpose of this method is to quantify if a denser or more complex urban environment correlates with more or fewer crashes than a less dense or more homogenous block group.
- Generating a trip generator typology, and investigating if crashes cluster near certain trip generators more often than others. The trip generators will be measured by individual metrics

as well as composite metrics. Using the metrics, one can rank the categories by how safe they are in order to assist with prioritization of capital improvement safety projects.

- Creating a logistic regression model to predict if the presence of several demographic and land use variables makes it more or less likely to be a victim of a crash that results in visible injury or death.

Before taking a close look at the methodology, it is prudent to briefly describe the data sources that were used in this study. TransbaseSF contained the crash data and several websites were utilized to collect the rest of the necessary data. San Francisco's public data repository, www.data.sf.org, as well as the United States Census Bureau, American Community Survey, Longitudinal Employer-Household Dynamics, and the Metropolitan Transportation Commission's ArcGIS online page were vital in the collection of precise and accurate data. The full list of data categories and where they are found online is included in Appendix (letter).

The first research question is:

What are the relationships between population density, employment density, and land use mix with pedestrian safety in San Francisco?

This is a vital connection to understand because as cities increasingly reconsider the outcomes of Euclidean zoning techniques and strive towards creating more walkable, compact communities, our neighborhoods will see more residents and employees interacting with each other in an increasingly complex built form. As more vehicles and pedestrians are introduced to an urban environment, the opportunities for a crash increase.

4.1 Prior Research

Results from existing studies are mixed in showing if there is a positive relationship between crashes and population density. For example, researchers studying pedestrian collisions in Buffalo, New York found a clear positive, linear relationship between population density and crashes.²⁸ They found that at the highest population densities, crashes were more common than at lower densities. These findings were not shared by all researchers. Wedagama, et al. concurred that crashes increased as population density increased however, once a critical density threshold was reached, the pattern reversed.²⁹ The idea is that as congestion increases and travel speeds slow, there are fewer dangerous pedestrian-automobile interactions. Other hypotheses explain this phenomenon differently. Rather than reducing travel speeds, Ewing and Dumbaugh claim that compact development leads to high population densities and lower per capita vehicle miles travelled (VMT). Shorter trips help minimize the likelihood of a crash.³⁰

How one normalizes crash data can make a dramatic difference in the findings. Pedestrian exposure is one metric to normalize crash rates. Pedestrian exposure is simply the degree to which a pedestrian interacts with vehicular traffic. Depending on the data available, pedestrian exposure can be quantified using pedestrian

counts, population density, employment density, the number of people that walk to work, or VMT.

Measuring crash rates as a function of pedestrian volumes generally supports the “strength in numbers” premise.³¹ Strength in numbers posits that as pedestrian volumes increase, the number of crashes grow more slowly than the proportion of pedestrians. That is, per capita crash rates decrease as pedestrian density increases. Burnier, Graham, and Glaister, as well as Bhatia and Wier confirmed this idea.³²

Prior research is less clear on employment density than it is on population density. First, there are unique temporal and spatial factors as they relate to employment density. High-rise office settings create different conditions than the same employment density in spread-out industrial areas like office parks. High intensity employment areas also attract a specific kind of person: someone of working age. This may explain why Wedagama, et al. and Dissanayake, et al. found that age was a mitigating factor of crash rates in employment zones. Adults were more likely than children to be a crash victim during normal work hours in these areas.³³

Most of the existing research has considered land use, defined by zoning designation, in isolation rather than in terms of land use mix. However, Burnier and Clifton, Burnier, and Akar both studied the relationship of land

28 Hoe Hun Ha and Jean-Claude F. Thill, “Analysis of Traffic Hazard Intensity: A Spatial Epidemiology Case Study of Urban Pedestrians,” *Computers, Environment and Urban Systems* 35, (2011): 230-240.

29 D.M. Priyantha Wedagama, Roger N. Bird, and Andrew V. Metcalfe, “The Influence of Urban Land-Use on Non-Motorized Transport Casualties,” *Accident Analysis & Prevention* 38, no. 6 (2006): 1049-1057.

30 Reid Ewing and Eric Dumbaugh, “The Built Environment and Traffic Safety: A Review of Empirical Evidence,” *Journal of Planning Literature* 23, no. 4 (2009): 347-367.

31 Peter Lyndon Jacobsen, “Safety in Numbers: More Walkers and Bicyclists, Safer Walking and Bicycling,” *Injury Prevention* 9, (2003): 205-209.

32 Bhatia and Wier, “Safety in Numbers’ Re-examined”; Burnier, “Pedestrian-Vehicular Crashes”; Daniel J. Graham and Stephen Glaister, “Spatial Variation in Road Pedestrian Casualties: The Role of Scale, Density and Land-use Mix,” *Urban Studies Journal Limited* 40, no. 8 (2003): 1591-1607.

33 D.M. Priyantha Wedagama, Roger N. Bird, and Andrew V. Metcalfe, “The Influence of Urban Land-Use on Non-Motorized Transport Casualties”; Dilum Dissanayake, James Aryaija, and D.M. Priyantha Wedagama, “Modelling the Effects of Land Use and Temporal Factors on Child Pedestrian Casualties.”

use complexity and crashes, but found no agreement.³⁴ Although they found no consensus, when measured in isolation, research has routinely found that commercial and industrial land uses correlate with the highest crash rates. Children were most often crash victims in residential zoning districts, but the finding could not be generalized to adults.³⁵

The following section will explain how population density, employment density, and land use mix were examined in this report, including a description of the GIS workflow and rationale for methodological choices.

34 Carolina V. Burnier, "Pedestrian-Vehicular Crashes: The Influence of Personal and Environmental Factors." Master's thesis, University of Maryland, College Park, 2005; Kelly J. Clifton, Carolina V. Burnier, and Gulsah Akar, "Severity of Injury Resulting from Pedestrian-Vehicle Crashes: What Can We Learn from Examining the Built Environment?" *Transportation Research Part D: Transport and Environment* 14, no. 6 (2009): 425-436.

35 Loukaitou-Sideris, et al., "Death on the Crosswalk."

4.2 Population Density, Employment Density, and Land Use Mix Workflow

The first step was to choose an appropriate geographic unit of analysis. Since this report endeavors to explore crashes in San Francisco in granular detail, the smallest reasonable census designation was chosen. Although block level data is finer grained than block *group* level data, fewer demographic variables are collected at the block level. Additionally, the American Community Survey, an annual sampling of Americans, is collected at the block group level rather than the block level. Since the crash dataset included crashes between 2011 and 2015, the density calculations were also created from data collected during the same time span.

After choosing to aggregate crashes at the block group level, the correct shapefile was downloaded from the United States Census Bureau's website. The City and County of San Francisco include many block groups outside of the city's core. These block groups include the Farallon Islands, which are void of human population, Angel Island, Alcatraz Island, and Treasure Island. Other block group boundaries consist primarily of ocean or the San Francisco Bay. These block groups were removed from the analysis so as not to skew the results. In the end, 578 block groups were included in the analysis.

After preparing the block group, feature class, population, employment, and zoning data were joined before the crash data was overlaid for the analysis.

An American Community Survey population estimate at the block group level for San Francisco between 2011 and 2015 was downloaded. The table was then joined by common attributes to the block group feature class and checked for accuracy using a variety of methods. These included creating a choropleth map of population by block group and spot-checking the table join with the Census Bureau's online tools. There existed a wide range of population sizes amongst the block groups. For example, block group 60750176011 had zero residents. This peculiar finding was confirmed when it was discovered that the block group is made up entirely of the Westfield Mall near the city's Union Square. The most populated block group was 60750607001, which makes up much of the Mission Bay neighborhood. This block group is significantly larger in area than most block groups in San Francisco, so it must be noted that these figures consider total population, not population density.

A new field, acres, was added to the attribute table of the block group feature class. This allowed the population data to be normalized by the acreage of each block group. After creating a choropleth map showing resulting population

density, the table was exported to Excel for more analysis. While the Mission Bay block group had the highest population, block group 60750125022 in the city's Tenderloin neighborhood was its most dense, with 310 people per acre (or 123,000 people per square mile). After successfully quantifying the population density of each block group and adding the information to the table, employment density was considered.

The number of employees at the block group level for 2015 was quantified by the U.S. Bureau of Labor Statistics and shared on its "On the Map" digital application. The tabular information was exported and joined to a copy of the block group layer. Once again, a new field, acres, was added to the attribute table. Employees at the block group level were normalized by the acreage of each block group to derive the employment density at the block group level. A choropleth map was created and the attribute table was exported to Excel.

Before moving forward, the data was once again checked for quality. Using a combination of personal knowledge of the city and comparing the employment output with what zoning designations allow, the employment density calculations were confirmed. Block group 60750427001 in the largely residential Sea Cliff neighborhood had the fewest employees. Five people work within that block group. Block group 60750615001 had the most employees -

this block group boundary stretches from the Ferry Building to Fourth Street along Market Street and south towards Howard Street. 73,403 employees work in this block group. Block group 60750159002 had the highest employment density. This Japantown location is home to a large mall and high-density office buildings and has a density of 840 workers per acre.

Zoning laws dictate the legal uses of a parcel. These detailed, often-dense regulations have been instrumental in how cities have developed over the last 100 years in the United States. For example, the concentration of single-family residential development away from noxious industrial uses in San Francisco was not created organically. Rather, zoning districts have been created to ensure the safety and welfare of residents by isolating them from potentially harmful externalities of industrial uses.

Cities treat zoning in dramatically different ways. While more homogenous locales may be content with minimizing the types of zoning districts, complex urban areas such as San Francisco have myriad zoning designations. At minimum, most municipalities zone for residential, commercial, industrial, and public or institutional uses. Each of these broad categories can be further refined, such as creating multiple residential zoning districts, each with its own height, bulk, or density requirements. San Francisco has 13 residential zoning districts. These range from fairly uniform, low-density districts to neighborhood or detailed corridor-specific zoning districts.

For the purpose of this study the districts were aggregated into five categories: commercial, industrial, mixed-use, residential, and public. A zoning shapefile was downloaded from the San Francisco Planning Department and added to ArcMap. As noted, the city’s zoning code is complex. While there are nearly 60 unique zoning districts, all of these districts can easily be organized into the five primary categories. Using the “dissolve” geoprocessing tool in ArcMap, the zoning designations were aggregated into the five primary groups, and the total acreage that each group accounts for was summed. Table 1 shows the districts that were included in each category.

Table 1: Aggregated Zoning Districts

Category	Zoning Districts
Commercial	C-2, C-3-G, C-3-O, C-3-R, C-3-S, CCB, CVR, MB-O, RCD, SSO
Industrial	M-1, M-2, PDR-1-B, PDR-1-D, PDR-1-G, PDR-2, SALI, SLI
Mixed-Use	CRNC, HP-RA, MB-RA, MUG, MUO, MUR, NC-1, NC-2, NC-3, NC-S, NCD, NCT-1, NCT-2, NCT-3, PM-MU1, PM-MU2, RC-3, RC-4, RED-MX, SPD, UMU, WMUG, WMUO
Residential	RH-1, RH-2, RH-3, RH-DTR, RM-1, RM-2, RM-3, RM-4, RTO-M, PM-R, RED, SB-DTR, TB-DTR
Public	MB-OS, P, PM-CF, PM-OS, PM-S,

Source: San Francisco Planning Department

It should be noted that San Francisco’s zoning code has been continuously updated. Longtime conforming uses may become non-conforming after a zoning change is made, and in certain cases non-conforming uses have been “grandfathered” in. This means that one may find a laundromat or corner convenience store in an RH-1 district, two uses that are normally prohibited in low-density residential areas. For the purpose of this study, non-conforming uses were not considered when quantifying the acreage of each zoning designation. The breakdown of the acreage of the five districts is based off the Planning Department’s official shapefile.

Over half of San Francisco is zoned for residential use, the majority of which are located in the city’s west and south. Public use, which includes parks and open spaces, public schools, and governmental facilities, is the second most common category. Mixed-use districts, often found along commute corridors and neighborhood commercial centers and industrial uses largely located along the city’s east coast, account for nearly 17 percent of San Francisco. The city’s compact downtown core contains the majority of its commercial zoning districts.

Table 2: Acreage by Zoning Category

Category	Area (acres)
Commercial	954
Industrial	2,057
Mixed-Use	3,375
Public	8,054
Residential	15,115
Total	29,555

Source: San Francisco Planning Department

Because this study is concerned with understanding land use mix at the block group level, once the zoning districts were organized into five core categories, they were joined to a block group feature class. ArcMap has a “spatial join” feature, which allows for the joining of data from one feature class to another based on shared geographic attributes. Through the spatial join process, one can quantify the number of unique zoning types in each block group, as well as calculate the acreage for the different levels of land use mix. The breakdown is shown in Table 3.

As stated above, many cities are moving towards more complex development patterns rather than homogenous, rigidly separated ones. This has manifested itself in an increased production of “mixed-use” housing and commercial stock located near common amenities that are connected by transit. By comparing crash rates of block groups with different levels of land use complexity, one will better understand the relationship between land use mix and crashes.

The results of the analyses appear in chapter five.

Table 3: Land Use Mix by Block Group

Number of Land Uses	Number of Block Groups	Area (acres)
1	22	944
2	180	7,113
3	313	16,380
4	50	4,516
5	13	648
Totals	578	29,601

Source: San Francisco Planning Department

4.3 Trip Generator Typology

In addition to analyzing crashes as they relate to population density, employment density, and land use mix at the block group level, this report also analyzed crashes as they relate to proximity to specific land use categories. The purpose is to answer the second research question:

Is there a correlation between a pedestrian's proximity to specific land uses and the risk of being involved in a collision?

To answer this question, land uses that generate trips were organized into a trip generator typology. The typology was created for this study but was informed by prior research. The land uses that make up the trip generator typology are:

- **Schools**
- **Parks**
- **Libraries**
- **Medical centers**
- **Rapid transit nodes**
- **Neighborhood Commercial Transit districts**
- **Neighborhood Commercial districts**
- **Public parking garages and lots**
- **Senior living facilities**
- **Fire and police stations**

The trip generators were defined as follows.

Schools

Ukkusuri and Rothman, et al. clearly showed in their research that of all built environment factors, the presence of a school was the strongest predictor of a crash. Further, it was found that half of all crashes that involve a child pedestrian occurred during school commute times, most often within 150 meters of the school.³⁶

A shapefile of all San Francisco Unified School District schools was downloaded. Schools that do not serve students ten years or older were removed from the list. This meant that a pre-kindergarten to fifth grade school was included, but a pre-kindergarten *exclusive* facility was not included. Prior research shows that students younger than ten tend to be accompanied by an adult and have a relatively low crash rate. As students reach ten years old, crash rates increase dramatically.³⁷ Private schools were not included, for San Francisco has a considerable number of private schools and the dataset did not include how many students each school served. Not knowing the size of the school, it was decided to include only public schools, as they tend to be larger than private schools. 132 schools in the city were analyzed.

36 Satish Ukkusuri, et al., "The Role of Built Environment on Pedestrian Crash Frequency," *Safety Science* 50, no. 4 (2012): 1141-1151; Rothman, et al., "Walking and Child Pedestrian Injury: A Systematic Review of Built Environment correlates of Safe Walking," *Injury Prevention* 20 (2014): 41-49.

37 Dilum Dissanayake, James Aryajja, and D.M. Priyantha Wedagama, "Modelling the Effects of Land Use and Temporal Factors on Child Pedestrian Casualties," *Accident Analysis and Prevention* 41, (2009): 1016-1024.



Figure 10: Locations of San Francisco Unified School District Schools

See Appendix B for list of sources

Parks

Research has regularly found that the presence of parks is negatively correlated with crash risk at both the aggregate and disaggregate levels.³⁸ In Baltimore and Los Angeles, block groups with the lowest percentage of park land had the highest crash rates and children were safer near parks than in places without open space.³⁹

A zipfile containing a shapefile of parks and open spaces in San Francisco was downloaded and extracted. According to the Recreation and Parks Department, San Francisco has 220 parks and open spaces. In fact, San Francisco's park system is so expansive that it is the only city in the country where every resident is within a 10-minute walk from a park.

Because San Francisco has so disproportionately more parks, it was necessary to shorten the list. Parks that did not foster congregation or recreation were removed. For example, planted medians along major thoroughfares such as Sunset Boulevard were included in the city's list of parks and thus were discarded from the analysis. To make sure that only parks that generated pedestrian trips were included, a careful visual audit was undertaken using Google Streetview. If the Google Streetview image did not show anyone at the park, if the park had no suitable areas for play, or if the park lacked seating, it was removed from the analysis. As a result, of the 220 parks on the original list, this analysis included 166 of them.

38 Burnier, "Pedestrian-Vehicular Crashes"; Loukaitou-Sideris, et al., "Death on the Crosswalk"; Rothman et al., "Walking and Child Pedestrian Injury."

39 Burnier, "Pedestrian-Vehicular Crashes"; Loukaitou-Sideris, et al., "Death on the Crosswalk."



Figure 11: Locations of Parks
See Appendix B for list of sources

Libraries

Since libraries have not been studied in detail by past researchers, there is no consensus as to the relationship between libraries and crash rates. They were chosen for the trip generator typology since they serve a diverse clientele and are located citywide in a variety of settings.

The San Francisco Public Library webpage has a list of every library branch in the city. The list was transcribed in an Excel spreadsheet and geocoded in ArcMap. San Francisco has 28 public libraries.



Figure 12: Locations of Libraries
See Appendix B for list of sources

Medical Centers

Medical centers have not necessarily been correlated with an increased risk of being a crash victim, however it was still included in the trip generator typology.⁴⁰ This is because it has been found that ambulances and other large emergency vehicles are involved in crashes that result in injury or death at higher rates than lighter duty vehicles.⁴¹

Forty hospitals and major medical centers, as defined by the San Francisco Planning Department, were included in the analysis. The list of medical centers was gleaned from a shapefile called “San Francisco Facilities” that was found on SF Open Data.

40 Schneider, Robert J., Rhonda M. Ryznar, and Asad J. Khattak. “An Accident Waiting to Happen: A Spatial Approach to Proactive Pedestrian Planning.” *Accident Analysis & Prevention* 36, no. 2 (2004): 193-211.

41 Kelly J. Clifton, Carolina V. Burnier, and Gulsah Akar, “Severity of Injury Resulting from Pedestrian-Vehicle Crashes: What Can We Learn from Examining the Built Environment?”



Figure 13: Locations of Medical Centers
See Appendix B for list of sources

Rapid Transit Nodes

Transit riders often begin and end each transit trip on foot, so it was reasonable to include it as a category in the trip generator typology. Researchers in New York City, Baltimore, Los Angeles, and Montreal have all studied the relationship between crashes and proximity to transit, but have arrived at a range of conclusions. Although the data has suggested that transit stops are inherently unsafe for pedestrians due to an increased mixing of traffic modes, research has also found that when measured at a per capita level, areas near transit stops are quite safe.⁴²

San Francisco has a plethora of local and regional transit choices, so it was paramount to decide which locations to include and which to leave out. Access to San Francisco Transportation Agency (SFMTA) ridership at the stop level was not publicly available, so it was impossible to parse the high-use transit nodes from ones that generated fewer boardings and alightings.

For the purpose of this study, rapid transit nodes are defined in two ways. First, a rapid transit node is any subway station. This includes the totality of Bay Area Rapid Transit (BART) stations inside city limits, as well as the nine SFMTA subway stations that facilitate the city's light rail transit vehicles. It should be noted that SFMTA and BART share four of the downtown stations.

Rapid transit nodes are also defined as any at-grade stop where a traveler can transfer between rapid routes. The SFMTA operates myriad bus routes and light rail routes throughout the city. The city is “mode agnostic” when referring to rapid transit lines. Headways and stop spacing rather than vehicle type defines if a route is “rapid”. Additionally, rapid stops tend to have more substantial shelters than lesser-used stops, adding to the stop's visibility. When considering subway stations and transfer points between rapid lines, San Francisco has 26 rapid nodes.

⁴² Ukkusuri, et al., “The Role of Built Environment.”



Figure 14: Locations of Rapid Transit Nodes

See Appendix B for list of sources

Neighborhood Commercial Transit (NCT) Districts

NCT districts are sections in the city where transit access and mixed uses are promoted. They tend to be organized along major commercial corridors. These districts have stronger parking controls to ensure that transit vehicles can move more quickly along routes. The combination of plentiful transit access and a variety of commercial and residential uses made NCT zoning designations an attractive land use type to explore.



Figure 15: Locations of NCT Districts
See Appendix B for list of sources

Neighborhood Commercial District (NCD)

NCD districts look to retain the “village” feel of having lower intensity, neighborhood-serving retail uses near residential and public districts. There is less of an emphasis on transit access, but parking meters and time-limited curbs are abundant in NCDs to promote more frequent parking turnover.



Figure 16: Locations of NCD Districts
See Appendix B for list of sources

Parking Lots and Garages

Just as the transit rider is a pedestrian at the beginning and end of most transit trips, so is one who parks at a parking lot or garage. For this reason, and because geography of parking lots require conflicts between pedestrians and cars, parking lots and garages were included in the trip generator typology.

The SFMTA keeps track of all parking lots and garages under its and the Port of San Francisco's purview. Seventy-four public lots and garages were listed and included in the analysis. They tended to be along the waterfront on the city's eastern edge, but there are still plentiful lots spread throughout the city.

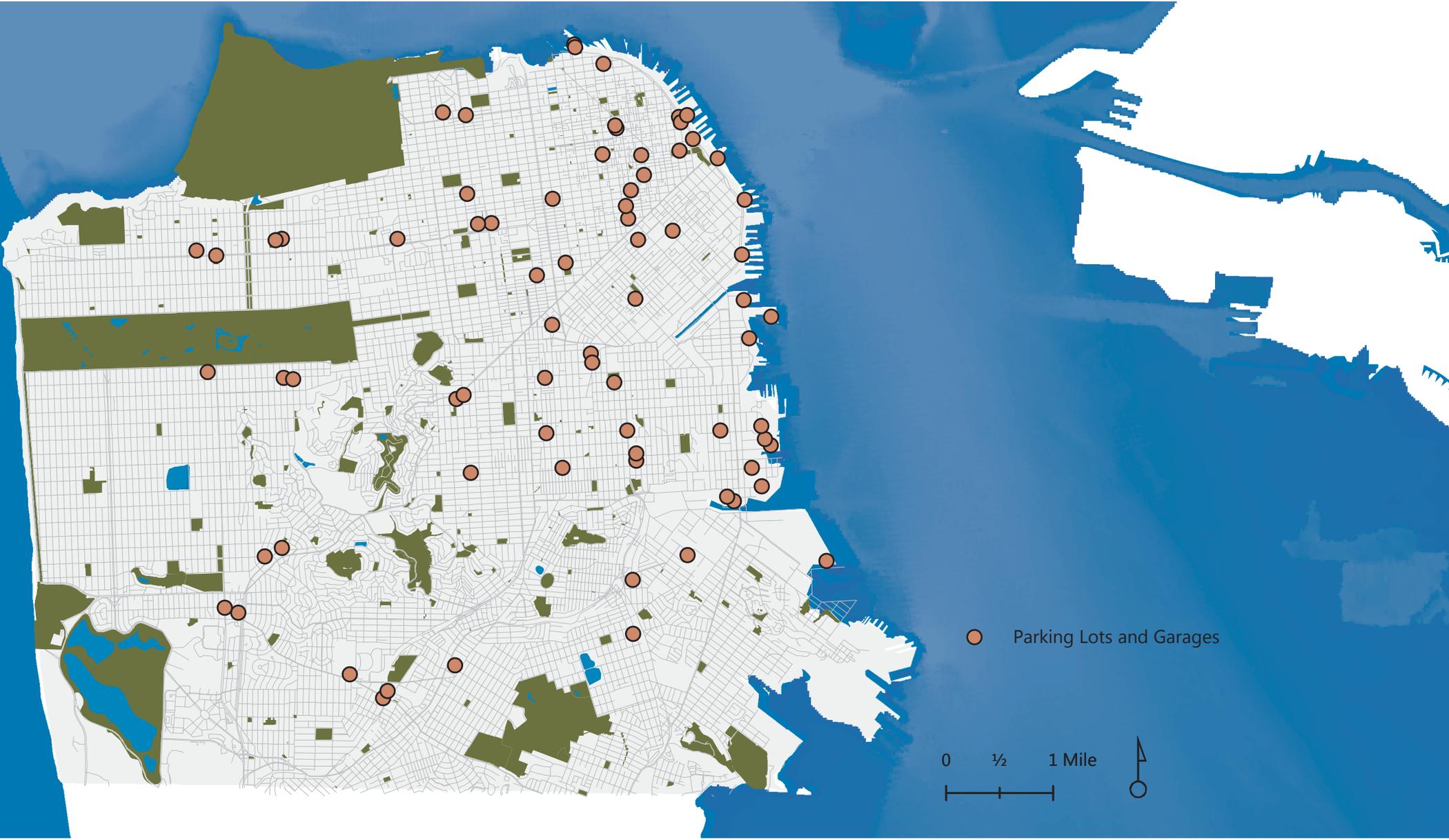


Figure 17: Locations of Parking Lots and Garages
See Appendix B for list of sources

Senior Living Facilities

Seniors and children are the two age cohorts most likely to be involved in a collision, but crash rates among the elderly were not uniform according to prior research.⁴³ In fact, an existing study of pedestrian safety for seniors in San Francisco showed that those older than 65 are unlikely to be victims of a crash.⁴⁴ One explanation is that seniors are less often exposed to vehicular traffic compared to more mobile age groups.

There are a plethora of living arrangements for seniors. From independent living, to assisted living, memory care, and hospice, the elderly population has many choices. Since seniors are a vulnerable population, senior living facilities were added to the trip generator typology. Since no singular list exists, Google.com and Yelp.com were used to generate a directory of senior living facilities. For a senior living facility to make the list, it must clearly state on its website that it is designed for independent living. Seniors must be able to walk freely outside. More intensive arrangements, such as hospice, can be included on-site, however to make the cut, independent living must be an option. Seventeen such facilities were found through the Internet search.

43 Schneider, et al., "Comparison of United States Metropolitan Regions."

44 Wier, et al., "An Area-Level Model of Vehicle-Pedestrian Injury Collisions."



Figure 18: Locations of Senior Living Facilities

Source: Google.com and Yelp.com search for "Senior Living Home," "Senior Living Facility," "Assisted Living," "Senior Housing," "Elder Care." Each search result's homepage was then checked.

Fire and Police Stations

As noted above, emergency and other large vehicles are more often involved in collisions that result in injury or death than smaller vehicles. Firehouses and police stations often include larger vehicles that regularly travel at high speeds to respond to an emergency. Although prior research has not identified either land use as a predictor for crashes, they were included in the trip generator typology to test these existing findings.

San Francisco's police and fire stations were extracted from the "San Francisco Facilities" shapefile. Prior research has shown that when ambulances and other emergency vehicles are involved in a collision, injuries tend to be more severe or fatal than when passenger vehicles are involved. There are 49 police and fire stations in the city.



Figure 19: Locations of Police and Fire Stations
See Appendix B for list of sources

4.4 Trip Generator Metrics

Five primary metrics were used to properly assess each of the land uses listed above. They were chosen based on the findings from prior research. Pulugurtha, et al. tested the efficacy of each of the metrics when studying vehicular crashes in Las Vegas. They identified three individual metrics, and two composite scores that utilized the results of the individual metrics. The goal of metrics is to normalize and quantify crash rates using pedestrian volumes, vehicle volumes, and built form characteristics. Through normalizing the data, one will be able to ultimately rank each category in terms of relative danger.

The three metrics below are all individual counts.

- **Crash frequency:** Crash frequency is the gross number of crashes. One simply sums the number of crashes within every service area of each trip generator category. Service areas represent walksheds for each trip generator and were created using Network Analyst. The process to create a service area will be described in more detail below.

- **Crash density:** Crash density considers the number of crashes in relation to the length, in street miles, of each service area polygon. To measure crash density, one quantifies the mileage of streets within each service area polygon, then

divides the gross crashes by the total street miles to quantify crashes per street mile.

- **Crash rate:** Crash rate measures the number of crashes in relation to pedestrian exposure⁴⁶ and vehicle exposure. Fehr and Peers created a pedestrian volume model at the intersection level for the SFMTA in 2011.⁴⁷ This model was used to quantify pedestrian exposure. SFMTA traffic counts taken between 1995 and 2016 were aggregated to quantify vehicle exposure at the intersection level.

The following two metrics are composite scores:

- **Sum-of-ranks:** Sum-of-ranks adds crash frequency, crash density, crash rate (pedestrian exposure), and crash rate (vehicular exposure) and then divides the total by four. The trip generator categories are then ranked according to the sum-of-ranks output.

- **Crash score:** A crash score is similar to sum-of-ranks in that combines several individual metrics. Whereas sum-of-ranks combines raw data, a crash score further normalizes the findings to thereby rank each land use variable. The process will be described in more detail below.

⁴⁶ Pedestrian exposure is the rate at which pedestrians must interact with vehicles. VMT, intersection density, and population density can all be used to quantify pedestrian exposure.

⁴⁷ Fehr and Peers, "San Francisco Pedestrian Volume Model," May 2011.

4.5 Trip Generator Workflow

Now that the trip generator typology and crash metrics are clear, the following text will describe the workflow that was undertaken to find correlations between crashes and a variety of land uses. The workflow includes several GIS processes, including utilizing ArcMap's Network Analyst extension and making calculations using the Field Calculator.

The process is as follows:

1. Create the trip generator typology:

The trip generator typology consists of ten land use variables. The list is the result of an extensive literature review and includes land uses that generate pedestrian trips. The full list and their definitions are listed below.

2. Create service area polygons for each trip generator:

Network Analyst was used to create service area polygons around each land use. The buffers measured 250 feet from each feature. 250 feet was chosen to mimic an approximately one block walk from each land use.

3. Intersect the pedestrian feature class to the service area polygons:

Intersecting the two feature classes will calculate the total number of crashes that occur within a 250 foot walk from each feature.

4. Quantify the mileage of streets (street miles) within each land uses service areas:

The intersect geoprocessing tool was run to find the overlap between the street centerlines feature class and the service area polygons. The mileage of streets was summed to quantify the total length of the streets for each land use category. Quantifying street miles is necessary to calculate the crash density metric.

5. Quantify pedestrian exposure:

A feature class representing pedestrian volumes was added to the map. The feature class represents point data. Each intersection has one point that has a corresponding pedestrian volume. The pedestrian volume feature class was intersected with the service area polygons to sum the total number of pedestrians in for each land use. This sum was then divided by the total number of intersections for each land use type to represent the average pedestrian exposure. The output was then used to calculate the crash rate (pedestrian).

6. Quantify vehicle exposure:

The same process described above was used to quantify vehicle volumes for a given land use's service areas. The average vehicle volume per land use was used to calculate the crash rate (vehicle).

7. Calculate each land use's "sum-of-ranks" score:

Sum the crash frequency, crash density, and crash rate metrics to get a composite score. Each score was weighted evenly.

8. Calculate each land use's "crash score":

Another composite metric, a crash score normalizes the individual metrics, combines them, and gives an output that compares the relative safety of each land use in the trip generator typology.

Since the first step, create the trip generator typology, was described above, the following section will commence with a discussion of creating service area polygons using Network Analyst followed by a detailed description of the remaining workflow.

Esri's Network Analyst extension was used to create a service area around the location of each trip generator. A simple radial buffer would suffice, but a service area more closely represents the built form environment. Whereas a simple buffer creates a perfect circle around a central feature, a service area polygon is the result of measuring the distance from a central feature along the street network. Therefore, the service area boundary can be irregular in shape.

One must create a network dataset to utilize Network Analyst's functionality. A network dataset was created from a shapefile of San Francisco's street network. Freeway segments or any other street that pedestrians do not walk along were removed. One-way restrictions were removed since pedestrians can walk against traffic on one-way streets.

The Create New Service Area feature in Network Analyst was used for each trip generator category. After loading a trip generator type, a 250-foot boundary was created around each trip generator location. This radius was chosen because it measures the approximate length of

one short block city block.

To illustrate, the trip generator category, schools, was added to the map. One service area polygon was created for each school, resulting in a total of 132 service area polygons, each representing a 250-foot walk from the school. The totality of crashes that occurred within the service area polygons will be introduced later in the workflow.

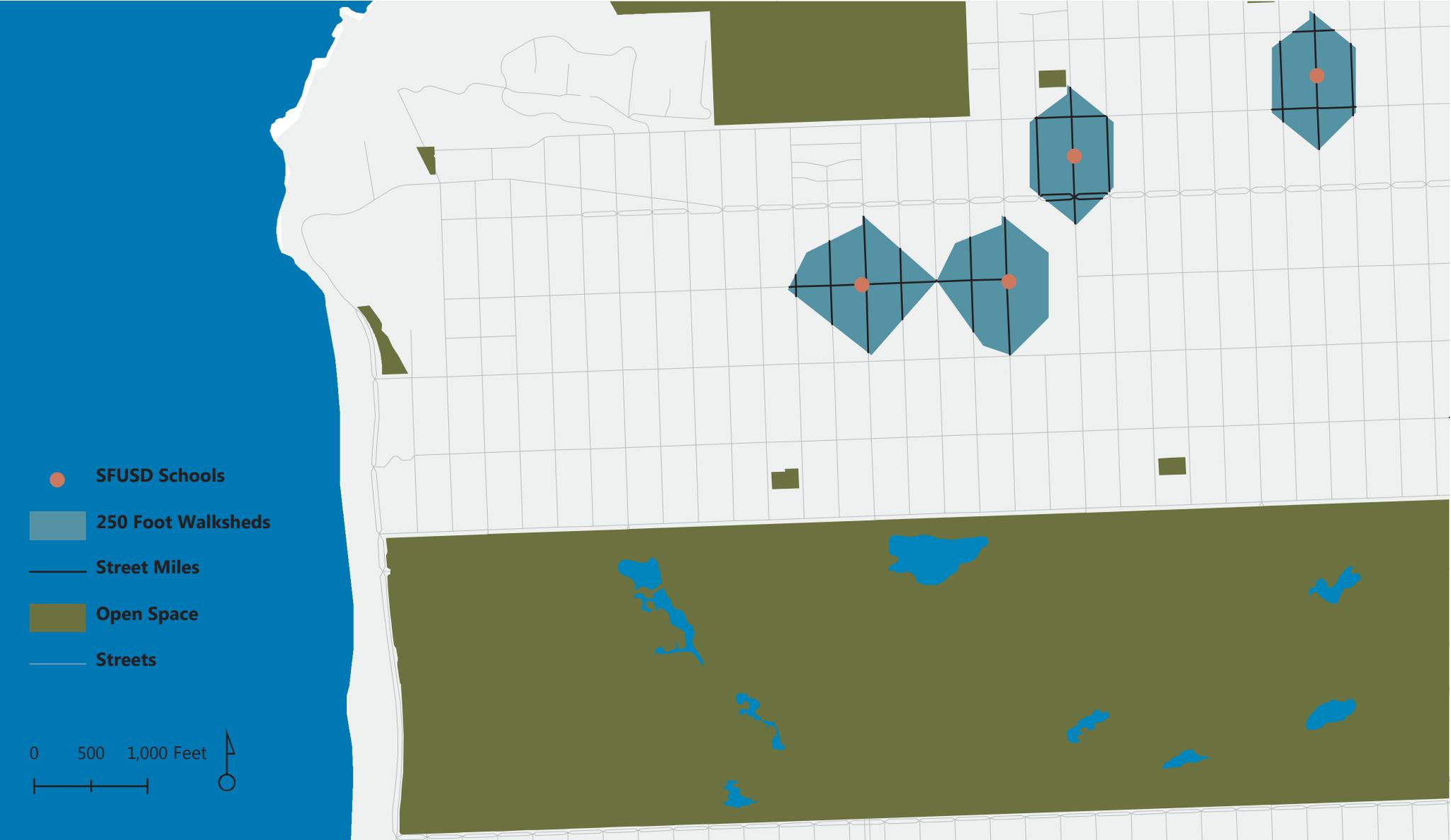


Figure 20: Detail View of School Locations, Walksheds, and Street Miles
 As described, Network Analyst service areas can be irregular in shape.

The same process described above was repeated for eight of the ten trip generator categories. NCD and NCT districts did not need service area polygons. Since both features are already polygons, only the streets contained inside the polygons, rather than the streets that are within 250 feet of the polygons' borders were deemed important for study.

After creating the service areas, the mileage of streets within the polygons was calculated. The same streets feature class that was used to create the network dataset was used. Using a "Select by Location" query, the roads that were completely contained in each polygon were highlighted. These roads were exported to a new feature class called "SF_street_miles_[trip generator]". The length of roads, in miles, was summed and the figures added to an Excel spreadsheet listing each trip generator category and the metrics listed above.

Pedestrian exposure, represented by pedestrian volume, was measured at the intersection level and is represented as point data. After adding the pedestrian volume feature class to the map, the process described above to quantify the total street miles within each service area was repeated to locate every intersection inside each of the polygons. The selected features were exported to a new feature class called, "SF_PED_volume_[trip generator]". The total number of pedestrians as well as the number of

intersections inside the polygons were summed. The average pedestrian volume was calculated by dividing the total number of pedestrians by the number of intersections. The average pedestrian volume was then used to normalize the gross number of crashes and calculate crash rates.

Whereas pedestrian volumes were used to quantify pedestrian exposure, vehicle volumes at the intersection level were used to calculate vehicular exposure. Both exposure metrics were used to calculate crash rate and included in the composite metrics. To quantify vehicle volumes at the intersection level, a large dataset was exported from the SFMTA's geospatial database. The tabular data was organized and geocoded using ArcMap's Address Locator functionality. Vehicular volumes for nearly 6,000 intersections were successfully geocoded. The data is a result of vehicle counts undertaken between 1995 and 2016.

Once a feature class was created showing vehicle volumes at intersections in San Francisco, an intersect tool was run to find the overlap between service area polygons and the vehicle volume points. The total vehicular volume was summed and divided by the number of intersections that were included in the dataset and added to the Excel table. All five metrics could then be organized and analyzed from the Excel table.

4.6 Logistic Regression

A regression model was created to further investigate land use and crashes at a localized level. Both the logistic regression to be described shortly and the trip generator method just explained used Network Analyst to draw service areas around a feature. Whereas a service area was created around each trip generator location in the prior method, for the purpose of the regression model, service areas were created around each crash site. Several GIS layers representing demographic and built form variables were added to the map and intersected with each crash location service area polygon.

Via an iterative process, 660 feet (or 1/8th mile) was ultimately chosen as the service area radius for each crash location. 660 feet was short enough of a distance to retain the localized nature of the study, but created a service area vast enough to capture sufficient data to differentiate each crash.

After the 660 feet service area polygons were created, the following GIS layers were added to map:

- Each trip generator typology category
- Each zoning district category
- Communities of concern (CoC)⁴⁸
- Completed traffic calming projects through FY 2014

Additional variables were added to the regression model. The crash dataset organized by TransbaseSF includes several important details for each crash. The following factors were gleaned from the crash dataset:

- Collision severity
- Street lighting
- Road conditions
- Intersection control (ie. presence of stop sign or traffic lights)
- Crash location (mid-block vs intersection)
- Driving under the influence of drugs or alcohol
- Driver ticketed for a Focus on the Five infraction
- Pedestrian cited for being at fault

A logistic regression is a well-known regression model that is effective in predicting and explaining the relationship between a dependent variable and several independent, explanatory variables. A logistic regression requires a binary dependent variable but can include nominal, ordinal, interval, or ratio level independent variables. In plain English, a logistic regression model states that for every one unit increase of the independent variable, one would expect the odds of the dependent variable to increase or decrease by a specific rate.

⁴⁸ A community of concern (CoC) is defined by the Bay Area's metropolitan planning organization (MPO) the Metropolitan Planning Commission (MTC). It is measured at the tract level and is based on eight variables: race, income, English language proficiency, elderly, zero-vehicle households, single-parent households, disability, and number of rent-burdened households. A CoC is any census tract that surpasses the minimum thresholds for income and race or measures above the threshold for income and any three other categories.

For the purpose of this study, the dependent variable was a crash that resulted in physical injury or death. The crash dataset included an injury severity column in which crashes are, per police protocol, categorized into four types at the scene by the officer. The least severe category is “complaint of pain.” The next step is “visible injury” followed by “severe injury” and “fatality.” Since there are four categories and a logistic regression requires a binary dependent variable, the four categories were re-coded into two. The first category, coded with a “0” in the dataset, represented crashes that resulted in only a complaint of pain. The second category, coded with a “1,” represented the remaining crashes, all of which resulted in visible injury or worse.

After creating the dataset in ArcMap and coding several independent variables, the dataset was exported to SPSS, a statistical software package well suited for logistic regression.

Now that three methods have been described in detail, a discussion of the findings will commence in the following chapter. Starting with a review of population density, employment density, and land use mix, the chapter will move on to discussing the trip generator results, and conclude with a summary of the findings from the logistic regression model. The report will conclude with chapter 6 with policy recommendations to mitigate dangers to pedestrians and make recommendations for future research.

FIVE

CHAPTER

← 635

HOWARD



FINDINGS

This chapter will state the findings from each method described above. After recapitulating the results, the implications of each will be discussed. Population density, employment density, and land use mix at the block group level are first to be examined, followed by the results of the trip generator typology method, and lastly the logistic regression.

5.1 Population Density and Crashes

Since population density was measured at the block group level, it was vital to accurately categorize the block groups into different density ranges. ArcMap has several methods to classify the population data. The block groups were organized in four equal quantiles: low density, low-medium density, medium-high density, and

high density. This means the nearly 600 block groups in San Francisco were divided into four equal groups which creates the fairest comparison of the block groups. Figure 21 is a frequency distribution depicting the population of each block group in the city. The vertical blue lines are the breaks for each density level. Although they are not uniformly spaced apart, the cut off points are set to create an equal number of block groups at each density level.

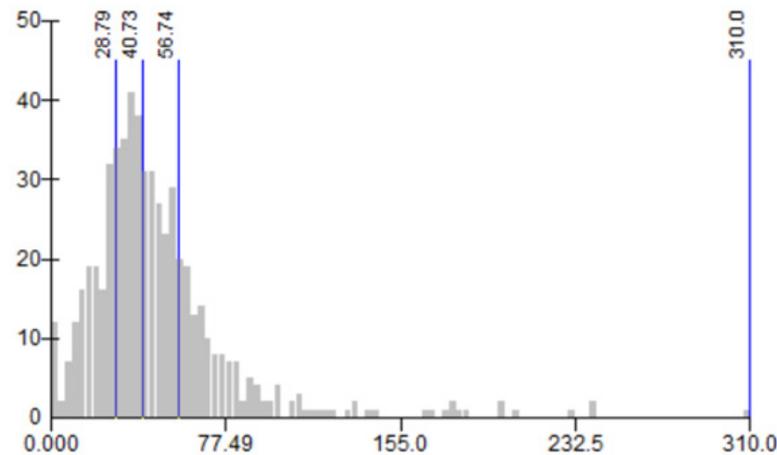


Figure 21: Histogram of Population Density

The X axis represents the population per acre and the Y axis quantifies the number of block groups per density. The blue lines show the quantile break values.

ArcMap calculated the numeric breaks for each category, with the result depicted in Figure 22. The number of crashes as well as the total number of acres for each density classification was summed. The total population at each density level was totaled, too.

San Francisco sees its highest population densities in its historic core: east of Van Ness Avenue, north of Market Street, and west of the Financial District. The block groups encompass the Tenderloin, Nob Hill, Chinatown and North Beach. There are also clusters of high density block groups along the Mission corridor towards Bernal Heights, and in the Lower Haight and Western Addition as one moves westward towards Golden Gate Park. The Richmond and Sunset districts, together forming the western portion of the city, are largely zoned for lower density residential, however the population densities there are moderate.

The most sparsely populated block groups are along San Francisco's eastern waterfront, in the Financial District, and west of Twin Peaks. The waterfront includes the majority of the city's industrial land, the Financial District is zoned for high intensity office and commercial uses, and west of Twin Peaks is a mix of open spaces, undulating topography, and detached single family homes.

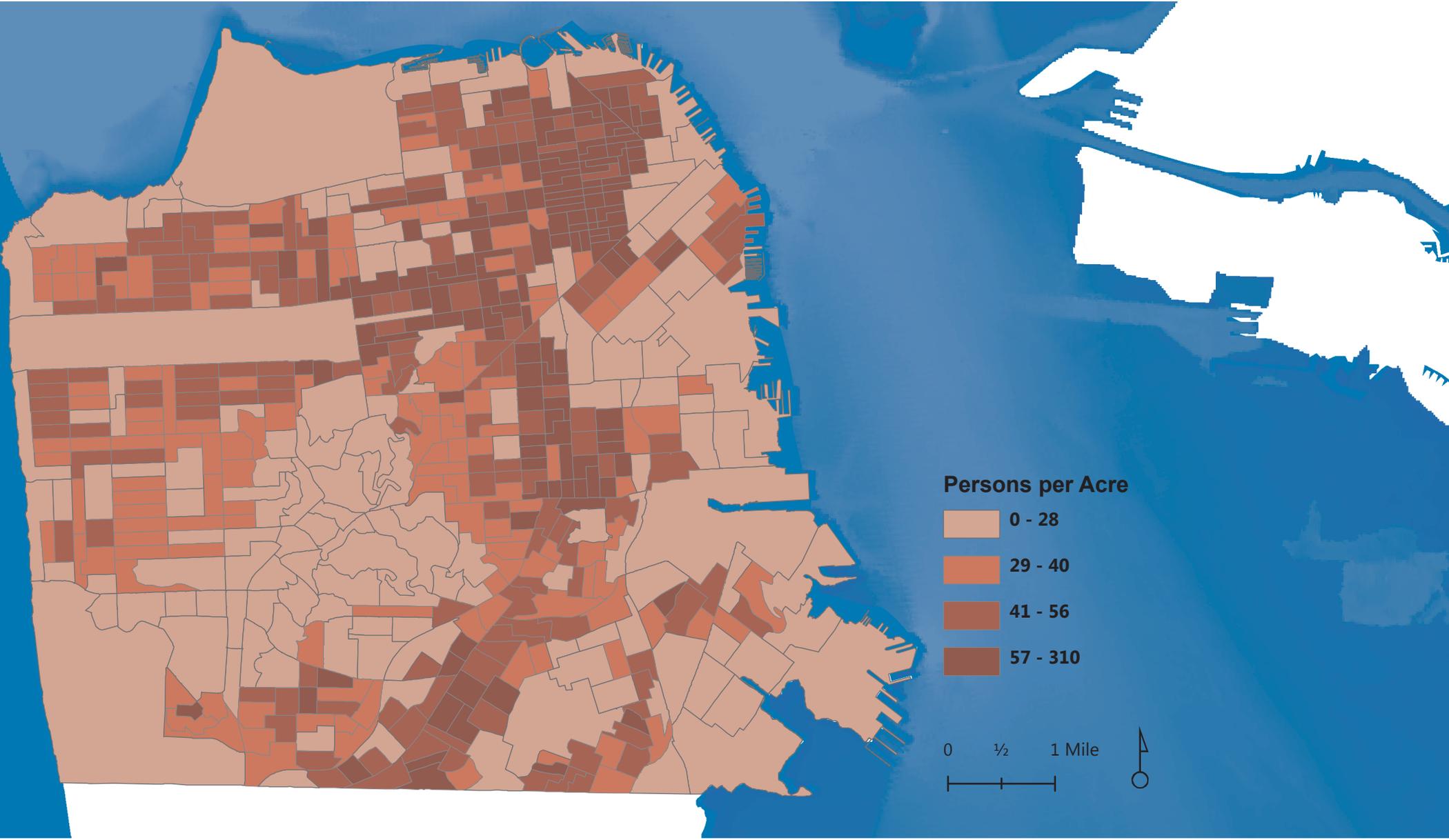


Figure 22: Population Density by Block Group, 2011-2015

See Appendix B for list of sources

The population density findings are included in Table 4 on the following page. From 2011-2015 the data shows that in terms of gross crashes it is most dangerous for pedestrians in block groups with the lowest and highest population densities. Keeping the number of block groups constant between the quantiles, the low density and high-density block groups together see twice the number of collisions as the block groups in the middle. Only six crashes separate the low and high-density block groups, with 1,306 crashes in high density block groups and 1,300 in low density. When looking at gross crashes alone, one can conclude that between 2011 and 2015 in San Francisco's block groups with the densest and least dense populations are the most unsafe for pedestrians.

There is a positive relationship between crashes and population density: as block groups become denser, there are more crashes per acre. In fact, the relationship is nearly exponential in nature with crash concentrations nearly doubling as density climbs. The low-density quantile sees fewer than 1/10th of a crash per acre, whereas the high density quantile has 0.43 crashes per acre, a substantial jump.

It is reasonable to expect that high-density block groups will have a greater number of crashes per acre. Block groups are generally delineated according to population characteristics. With equal populations, a high-density block group

tends to contain fewer acres than a sparsely populated one. In fact, Q1 contains one more block group than Q4, but takes up over four times the area. It would be expected that if comparing crashes per acre, high density block groups would seem most unsafe to pedestrians.

One can induce that the high number of crashes in the lowest density of block groups are due to two primary factors: low-density block groups make up nearly half of the entire acreage of San Francisco. More street miles increases the opportunities for a collision. Another potential explanation is that drivers are less aware and drive faster in less complex environments. Additionally, these more newly developed areas may have wider streets with poor enclosure which can influence travel speeds.

Table 4: Gross Crashes and Crashes per Capita by Block Group Population Density

Quantile	Number of Block Groups	Area (acres)	Population	Crashes	Crashes per Acre	Crashes per Capita
Q1 Low Density 0-28.79 persons/acre	145	16,649	184,360	1,300	0.0781	0.0071
Q2 Low-Medium Density 28.80-40.73 persons/acre	145	5,290	184,691	603	0.1140	0.0033
Q3 Medium-High Density 40.74-56.74 persons/acre	144	4,618	224,364	777	0.1683	0.0035
Q4 High Density 56.75-309.98 persons/acre	144	3,044	244,161	1,306	0.4290	0.0053
Totals	578	29,601	837,576	3,986	n/a	n/a
Averages	144.5	7,400.25	209,394	996.5	0.1482	0.0048

Source: SFDPH and San Francisco Planning Department

Although there are more crashes in total and per acre in high-density block groups, a different conclusion can be reached when considering crashes per capita. The block groups with the lowest population density have the highest per capita crash rate. High density block groups rank second. Low-medium density block groups are the safest per capita. Figure X makes these relationships clear.

The results here somewhat support the “strength in numbers” theory. Strength in numbers posits that as pedestrian volumes increase, so does safety. Even though Q4 is not as safe as Q2 and Q3 at the per capita level, high density block groups are safer than low density ones.

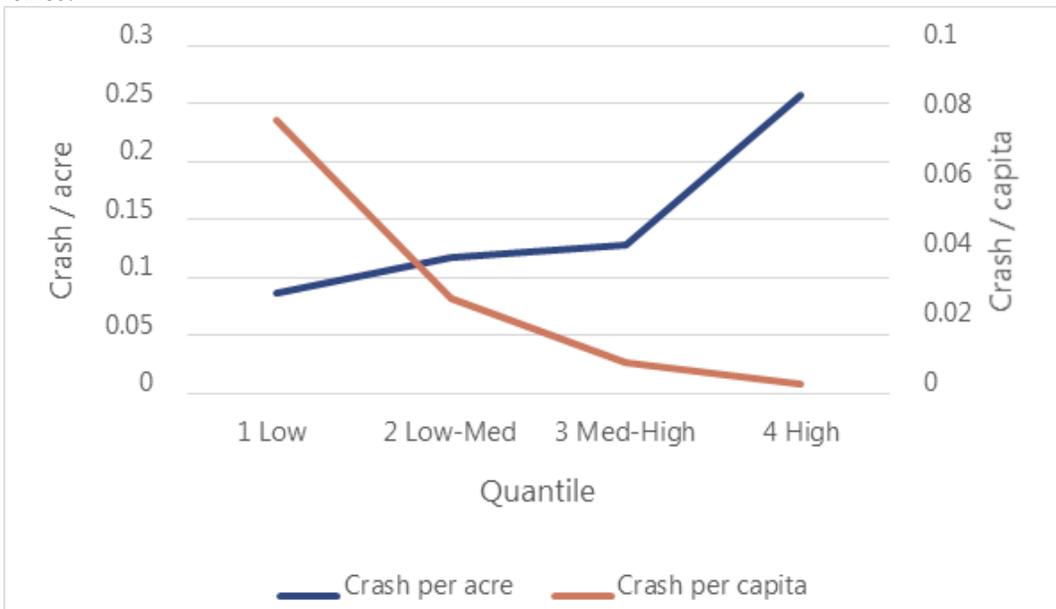


Figure 23: A Comparison of Crashes per Acre and per Capita (population)

5.2 Employment Density and Crashes

The same process described above was used to divide the block groups into four equal quantiles. There is less variation in employment densities than population densities. The high intensity employment centers see as many as 851 workers per square acre, however most block groups have 25 workers or fewer per acre.

It is clear when comparing the map on the following page to the population density map that dramatically different block groups make up each quantile. Unlike population density, San Francisco’s employment patterns show a clear clustering of high density employment in the city’s northeast quadrant. This area includes the Financial District, South of Market, Union Square, and the majority of the city’s regional public transportation nodes. Whereas the city’s east coast saw relatively low population densities, there are far more workers.

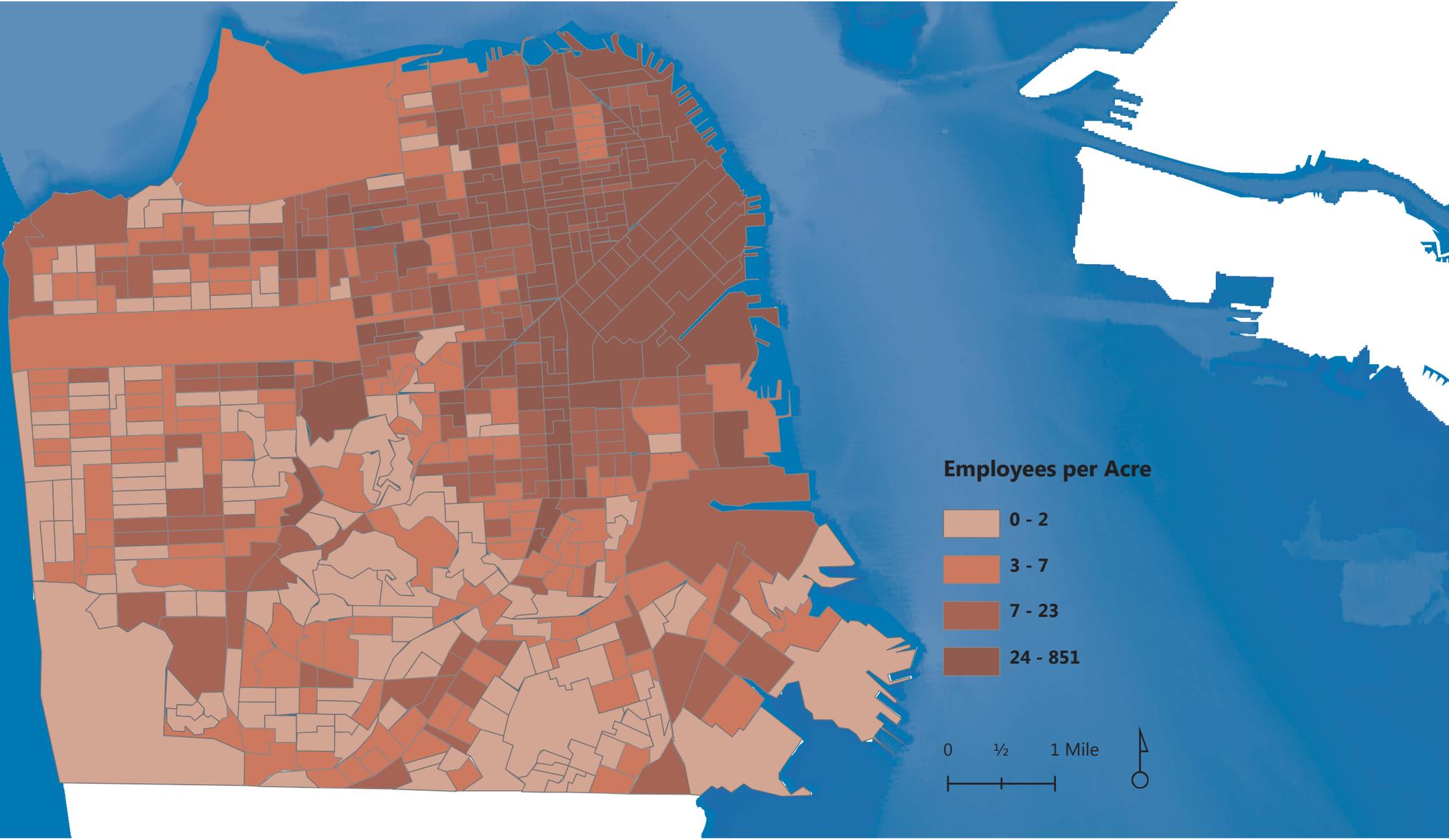


Figure 24: Employee Density by Block Group, 2011-2015

See Appendix B for list of sources

Table 5 summarizes the relationship between employment density and crashes. No clear pattern emerges as one considers the number of crashes in each quantile. Q1 (with low densities of employees) has the fewest number of crashes, and Q4 (with the highest densities) sees the greatest number of crashes. However, the relationship between employment density and crashes is not linear. Block groups in Q2 actually see more crashes than block groups in Q3. With the highest density block groups having nearly twice the number of crashes as the lowest densities, some may conclude that areas with lots of employees are the most unsafe; however, as we will see, the number changes drastically after they are normalized for pedestrian exposure.

Like population density and crashes, employment density and crashes (when normalized by area) shows a positive relationship: as block groups get more dense with employees, the crash rate per acre increases. The findings make *a priori* sense. The larger the at-risk population, the greater the likelihood that a crash will take place. And as block groups get more dense they become increasingly smaller in area which means that there will be a positive relationship between crashes per acre and employment density. Figure 25 makes this clear: the blue line representing crashes per acre steadily climbs with density levels, whereas per capita crash rates are their smallest in the densest block groups.

Crash totals are non-uniform when employment density is considered. Low-medium and high-density quantiles account for the most crashes. Low density employment areas had the fewest. As our example above showed, normalizing the data can have dramatic effects. When acreage is included, there is a linear relationship between crashes and employment density: the higher the density, the higher the crash per acre rate is. Although this may seem as though high employment density is correlated with a more unsafe environment, the results reverse when considered at the per capita rate. There is a linear relationship when crashes are measured per capita, but the relationship reverses. The negative linear relationship means that pedestrians at the per capita level are struck by vehicles more often as employment density declines.

When normalizing crashes by employee density, the relationship between crashes and density reverses. There is a clear negative relationship: this means that as a block group gets more dense, one sees a decline in collisions per capita. This trend is even stronger than it was for population density and crashes. In terms of employment density, the most dense quantile is the safest for pedestrians.

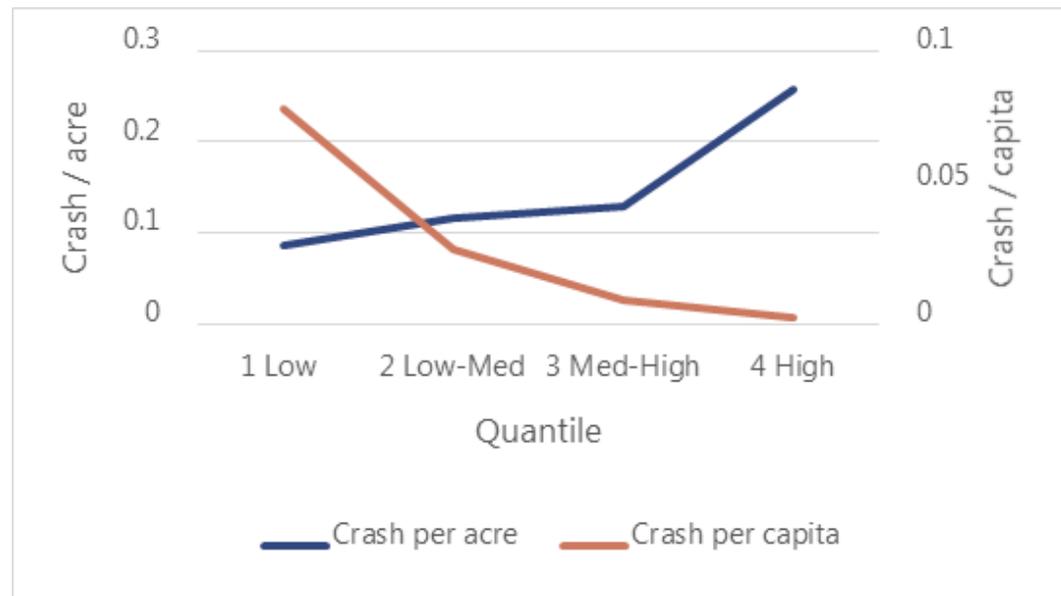


Figure 25: A Comparison of Crashes per Acre and per Capita (employee)

Table 5: Gross Crashes and Crashes per Capita by Block Group Employee Density

Quantile	Number of Block Groups	Area (acres)	Employee Population	Crashes	Crashes per Acre	Crashes per Capita
Q1						
Low Density	144	9,141	10,014	789	0.0863	0.0788
0.00-2.31 persons/acre						
Q2						
Low-Medium Density	144	8,825	37,504	1,029	0.1166	0.0274
2.32-7.83 persons/acre						
Q3						
Medium-High Density	145	6,419	91,402	827	0.1288	0.0090
7.84-23.11 persons/acre						
Q4						
High Density	145	5,217	544,701	1,341	0.2570	0.0025
23.12-851.10 persons/acre						
Totals	578	29,601	683,621	3,986	n/a	n/a
Averages	144.5	7,400.25	170,905.25	996.5	0.1347	0.0058

5.3 Land Use Mix and Crashes

This report analyzed automobile-pedestrian crashes according to five land use categories: commercial, industrial, mixed-use, public, and residential. The number of crashes were quantified for each zoning designation, as well as totaled according to land use mix. To quantify land use mix, the number of unique zoning categories were summed within each block group. This resulted in block groups having a land use mix between one and five. Two maps are below. The first shows the city's zoning districts organized into the five categories listed above. The second map shows the city's land use mix at the block group level.

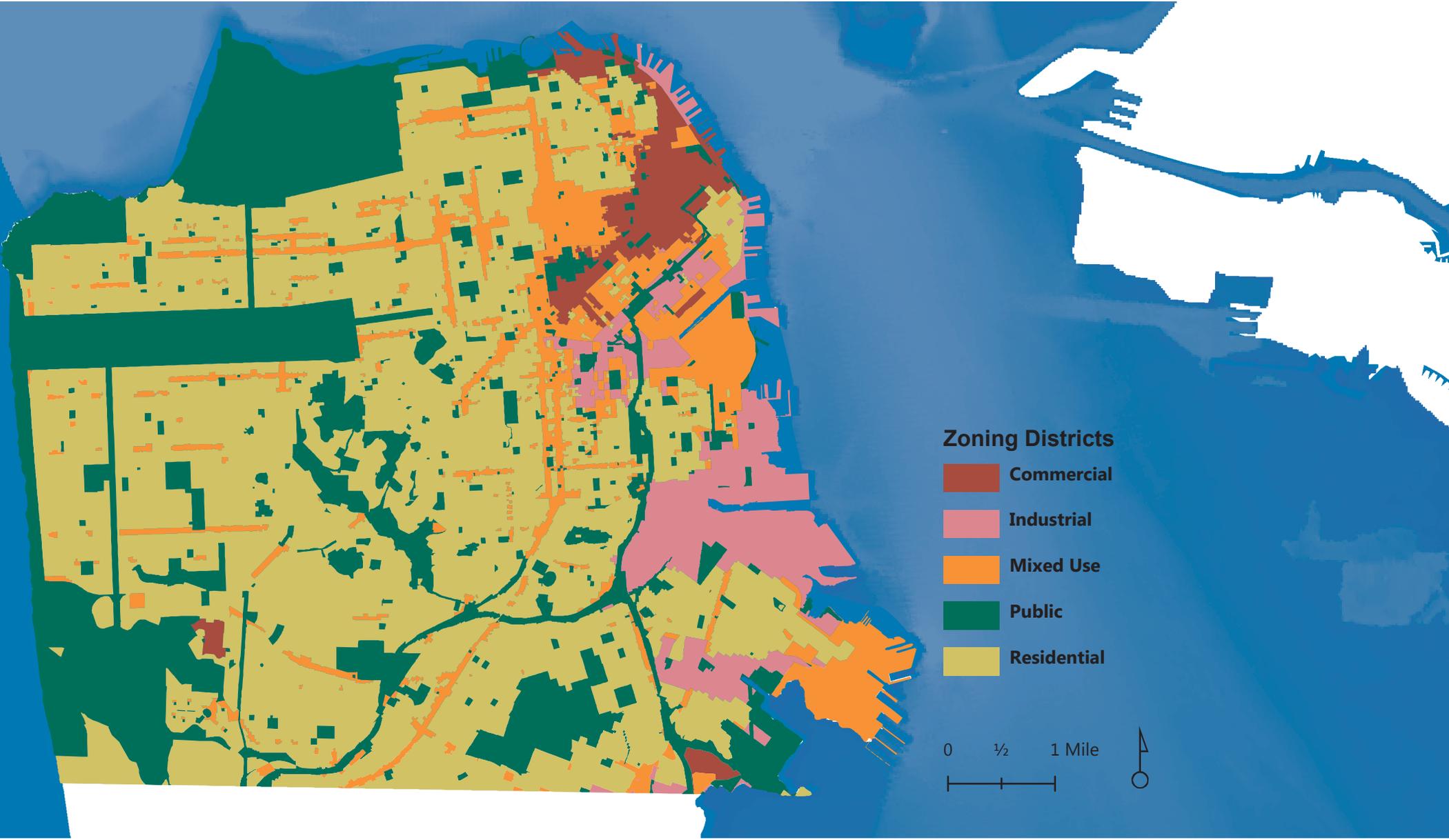


Figure 26: Zoning Districts in San Francisco

See Appendix B for list of sources

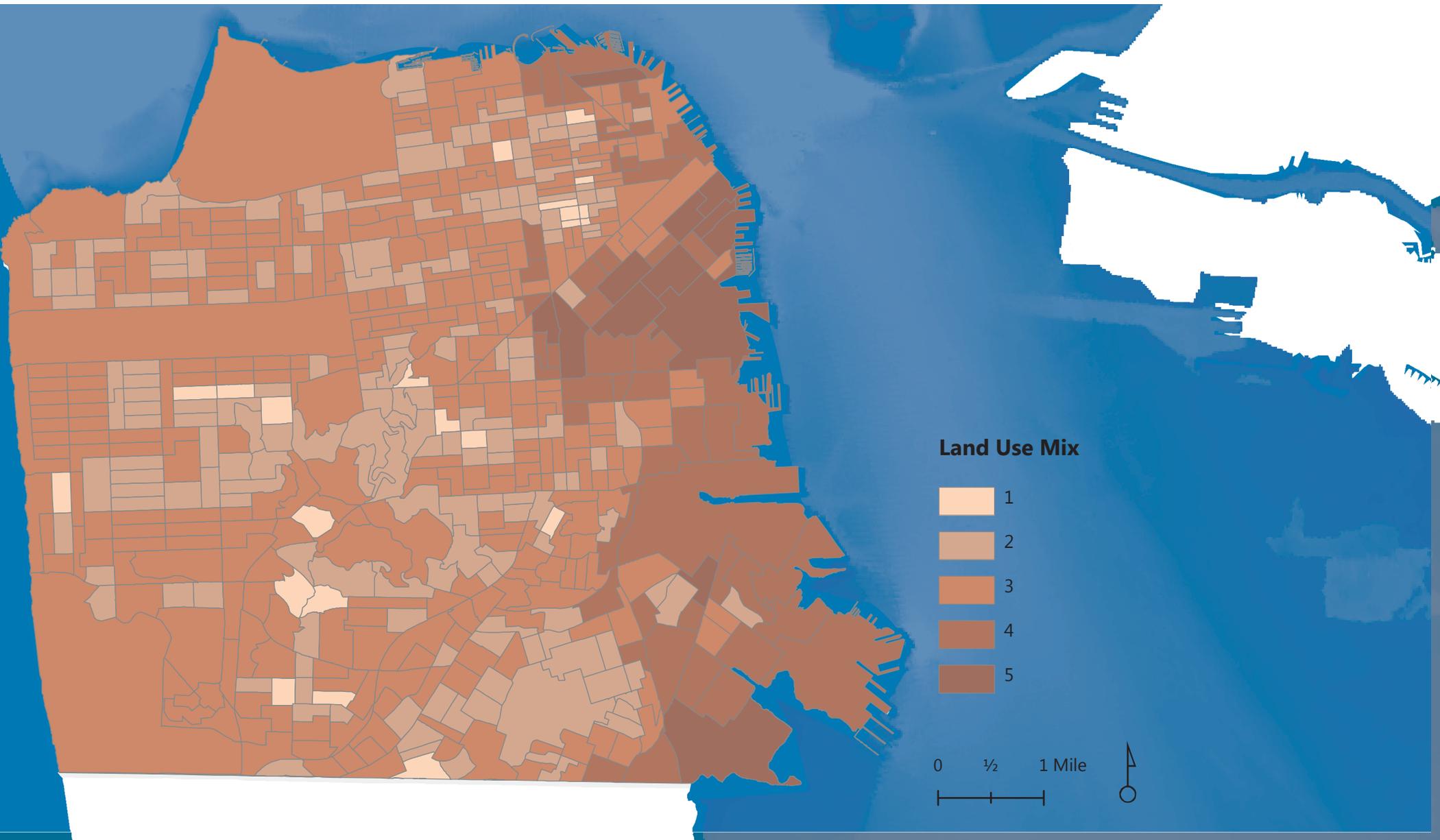


Figure 27: Land Use Mix by Block Group

See Appendix B for list of sources

It is clear that much of the western section of the city is zoned for residential and the city's eastern neighborhoods are its most diverse in terms of land use. One sees large mixed-use districts along the central waterfront as well as in the city's Hunter's Point neighborhood. North of Hunter's Point is zoned for industrial with public, residential, and commercial nearby. The larger geographic area of the eastern block groups also contribute to them being more diverse in land use than the north and west.

The following table totals the number of crashes that have occurred within each zoning category. In terms of gross crashes, mixed-use and residential zoning districts accounted for the most crashes with industrial land uses seeing the lowest crash total. The results change when one normalizes crashes by the acreage of each zoning designation. Although residential zoning districts account for nearly 25 percent of crashes, over 50 percent of San Francisco is zoned for residential uses. So, when measured at crashes per acre, only public districts are safer for pedestrians than residential districts. Commercial designations are the most unsafe for pedestrians at the per-acre level with a crash rate nearly ten times higher than that of residential districts.

Normalizing crashes for mixed-use designations did not temper the initial findings: mixed-use zoning districts account for the greatest share of crashes and saw the second highest per acre crash rate after commercial.

Table 6: Crash per Acre by Zoning Category

Designation	Acres	Crashes	Crash per Acre
Commercial	954	633	0.6632
Industrial	2,057	195	0.0948
Mixed Use	3,375	1,709	0.5064
Public	8,054	423	0.0525
Residential	15,115	1,026	0.0679
Totals	29,556	3,986	n/a
Averages	5,911.2	797.2	0.1349

Table 7: Crash per Acre by Land Use Mix

Number of Land Uses	Number of Block Groups	Area (acres)	Crashes	Crash per Acre
1	22	631	102	0.1616
2	180	6,557	667	0.1017
3	313	16,286	2,207	0.1355
4	50	4,476	669	0.1495
5	13	1,651	341	0.2065
Totals	578	29,601	3,986	n/a
Averages	115.6	5,920.2	797.2	0.1347

San Francisco has a complex and diverse mix of land uses. This report aggregated the nearly 60 zoning districts into five basic categories. The five land use categories were then spatial joined to a block group feature class to quantify the number of zoning districts in each block group. Land use mix ranged from a minimum of one to a maximum of five. It must be noted that this analysis does not take into account Conditional Use authorizations for non-conforming uses, such as the addition of a corner grocer or laundromat to a residential zoning district. There is no doubt that San Francisco's mix of land uses is more complex than the official zoning districts initially suggest. Table 7 organizes crashes, acreage, and crash density according to the number of land uses within each block group.

Eighty-five percent of the city's block groups have two or three land uses. Taken together, these block groups account for nearly 75 percent of collisions in San Francisco. This means that although a moderate land use mix accounts for most the most crashes, these block groups have proportionally fewer crashes than their more and less diverse counterparts.

It is rare for block groups in San Francisco to have a single land use or all five land uses. The lack of block groups with only one land use type illustrates San Francisco's complex and diverse built form. In terms of gross crashes, block groups with only one or all five land uses had the fewest number of crashes. This was a result of there being fewer of these kinds of block groups rather than something inherently more or less safe about them.

The results change when one considers per acre crash rates. Completely homogeneous or heterogeneous block groups are the most unsafe for pedestrians. Block groups with all five land uses had the highest number of crashes per acre and block groups with only one zoning designation ranked just behind. The results signal that homogenous zoning methods should be avoided for pedestrian safety reasons, but as land uses become more mixed, greater caution should be taken to ensure that the environment is safe for pedestrians.

To summarize, between 2011 and 2015 in San Francisco pedestrians were struck most often in block groups at the highest population density, however pedestrians are more unsafe at the per capita level in less dense areas. The same relationship was even stronger in block groups with high employment density. This report offers that diverse land use should not be shunned, but extra safety and traffic calming precautions are prudent in diverse block groups.

5.4 Trip Generator Typology

The following section will discuss the findings from the trip generator methodology. Each of the individual and composite crash metrics will be analyzed followed by a discussion of the implications from the data.

The above table lists the ten trip generator variables and their individual counts. Each category's acreage and street miles are totaled. These figures were derived from the service area polygons that were created using Network Analyst.

There are more parks and schools than any other land use in San Francisco. Although there are only 34 more parks than schools in the city, since several parks are spread across significantly larger areas than schools, acreage of parks is nearly four times larger than that of schools. In fact, there are more miles of streets bordering parks in the city than all the other land uses combined.

After organizing each trip generator category, the individual metrics were measured. Crash frequency is the most basic of the metrics in the study. It simply represents the total number of crashes for each land use category. Table 9 below shows the crash frequency for each category.

Table 8: Count, Area, and Street Miles for Each Trip Generator Category

Category	Number	Area (acres)	Street Miles
Schools	132	2,160	96
NCT	21	464	25
NCD	25	523	17
Medical centers	40	615	8
Parks	166	9,538	485
Senior Living	17	258	22
Rapid Transit Nodes	26	436	24
Libraries	28	477	23
Parking Lot	74	1,162	21
Police and Fire	49	787	38

Table 9: Crash Frequency

Category	Number	Crashes	Crash Frequency
Schools	132	479	479
NCT	21	408	408
NCD	25	392	392
Medical Centers	40	117	117
Parks	166	1,163	1,163
Senior Living	17	104	104
Transit Nodes	26	331	331
Libraries	28	136	136
Parking Lot	74	434	434
Police and Fire	49	234	234

The most crashes occur within 250 feet of a park and open space in San Francisco. This may seem alarming since parks generate trips from diverse populations, but the high levels of crashes are a function of the high number of parks in the city rather than something inherently dangerous about them. Numerous studies have shown that parks and open space are generally more safe than other land use types.

Schools followed by public parking lots account for the second and third most crashes. These numbers should be studied more carefully since vulnerable children frequently walk to school and parking lots necessarily set up conflicts between pedestrians and cars.

NCT and NCD districts are next. These trip generators see a mix of land uses and planned for multimodal travel. It is alarming that crash frequencies are so high as there are fewer NCT and NCD districts than other land use types.

Transit nodes see fewer crashes than the zoning districts listed above. The remaining land use categories are also correlated with fewer crashes than parks, schools, parking lots, NCT, NCD districts.

Crash density normalizes gross crash rates by the length, in street miles, of the service area polygons. Crash density is an improvement over crash frequency, but it still does not factor in pedestrian exposure. The crash density table is shown here.

Table 10: Crash Density

Category	Street Miles	Crashes	Crash Density
Schools	96	479	4.98
NCT	25	408	16.40
NCD	17	392	22.53
Medical Centers	8	117	14.21
Parks	485	1,163	2.40
Senior Living	22	104	4.75
Transit Nodes	24	331	13.63
Libraries	23	136	5.79
Parking Lots	21	434	21.17
Police and Fire	38	234	6.20

Although parks had significantly more crashes in total, when street miles are considered in the calculation, parks have the lowest crash density. On the other hand, NCDs, high on the crash frequency scale, see more crashes per street mile than any other land use category. Conclusions cannot be drawn as to *why* there are more crashes per mile. One can hypothesize that NCDs have diverse land uses, see high levels of pedestrian activity, and have many conflict points between pedestrians, transit vehicles, and cars. Further study can attempt to breakdown the circumstances of each crash to better understand the specific variables that correlate with crashes in NCD districts.

Parking lots see the second highest crash density with 21.17 crashes per street mile. This figure is disconcerting. There are numerous parking lots along the eastern waterfront and the Embarcadero. With few through-streets and lower intensities of development along the eastern edge of the Embarcadero, one would think that there would be fewer crashes within 250 feet of parking lots and garages.

NCTs, medical centers, and rapid transit nodes have the next three highest crash densities. NCTs are similar to NCDs in that they combine a multitude of uses and are multimodal in nature. Medical centers have only eight street miles – the lowest of all categories – but has 14.21 crashes per street mile. Many rapid transit nodes are located

in or near NCT districts, so it is reasonable that its crash density is similar to NCTs. A heartening sign is that crash density near rapid transit nodes is lower than at NCTs, so it may be safer for pedestrians as they get closer to a rapid transit transfer station or underground station.

The remaining categories see similarly low crash densities. Like parks, schools is an interesting case. Although schools had a high crash frequency, since there is a high number of schools in San Francisco there is a high number of street miles surrounding them. Because of this, schools may look unsafe in terms of crash frequency, but actually see relatively few crashes per street mile than other land use categories.

Crash rate is normalized by pedestrian exposure which can be measured by several metrics. This study used two metrics to create two unique crash rates. The first metric normalized crashes by pedestrian volumes. This report used a 2011 model created by Fehr and Peers for the SFMTA to quantify pedestrian volumes at the intersection level. The second metric normalized collisions by vehicle volumes. Data was derived from an SFMTA database of traffic counts. Crash rates normalized by pedestrian volume are shown in Table 11 below.

Category	Pedestrian Volume	Number of Intersections	Average Pedestrian Exposure	Crashes	Crash Rate (ped)
Schools	1,166,900,885	727	1,605,091	479	0.000298
NCT	818,917,332	258	3,174,098	408	0.000129
NCD	760,522,615	287	2,649,905	392	0.000148
Medical Centers	215,482,408	127	1,696,712	117	0.000069
Parks	2,993,355,462	2,558	1,170,194	1,163	0.000994
Senior Living	314,770,541	93	3,384,629	104	0.000031
Transit Nodes	671,392,742	156	4,303,800	331	0.000077
Libraries	349,263,156	178	1,962,153	136	0.000069
Parking Lot	1,212,828,257	345	3,515,444	434	0.000123
Police and Fire	861,822,007	292	2,951,445	234	0.000079

Table 11: Crash Rate (pedestrian)

When normalizing crash frequency by pedestrian exposure, parks have the highest crash rate by a significant margin. Nearly three billion people walk near parks according the SFMTA model – the highest of any land use category, however with 2,558 intersections included in the parks and open spaces service area polygons, the average exposure is quite low. It follows that with a high frequency of crashes and relatively few pedestrians at each intersection, the crash rate will be quite high.

Schools, NCDs, NCTs, and parking lots have the next highest crash rates. Besides parks, schools and parking lots are the only categories that attract over one billion pedestrians annually. What is especially disconcerting about parking lots, however, is that unlike schools and parks, the average pedestrian exposure near parking lots is quite high. This means that “strength in numbers” does not necessarily apply to parking lots like it does for other land use variables.

The remaining categories have the lowest crash rates of the land use variables that were considered. Of these categories transit nodes and police and fire stations see the most foot traffic within 250 feet of each site. Since there are fewer rapid transit nodes in the city as compared to police and fire stations, transit nodes generate the highest average pedestrian exposure of all ten land use variables. This is reasonable since the majority of transit trips start and end by walking,

and rapid transit nodes are located in dense areas that are suitable for mass transit.

Senior living facilities have the lowest of all crash rates. They are located across the city and are not necessarily clustered in quiet areas. One cannot draw conclusions from this research, but the findings mean that senior living facilities have taken care to make sure their pedestrian environments are safe. Or it simply means that seniors, a normally at-risk group, may not walk much outside.

As noted, this report also quantified a second crash rate metric by using vehicle counts at the intersection level. Crash rates as a function of vehicle volumes are represented in Table 12.

Category	Vehicular Volume	Number of Intersections	Average Vehicular Exposure	Crashes	Crash Rate (veh)
Schools	2,966,637	587	5,054	479	0.0948
NCT	3,886,071	407	9,548	408	0.0427
NCD	1,046,124	243	4,305	392	0.0911
Medical Centers	1,483,876	177	8,383	117	0.0140
Parks	11,738,219	2,094	5,606	1,163	0.2075
Senior Living	1,084,178	118	9,188	104	0.0113
Transit Nodes	2,393,312	295	8,113	331	0.0408
Libraries	490,711	125	3,926	136	0.0346
Parking Lot	3,459,176	392	8,824	434	0.0492
Police and Fire	1,445,210	214	6,753	234	0.0346

Table 12: Crash Rate (vehicular)

When normalizing crash frequencies by vehicle volumes, parks are correlated with the highest crash rate. This is because parks see a significantly high number of crashes, but have a low average vehicle exposure of 5,606 vehicles per intersection per day. On the other end of the spectrum is senior living facilities. Service areas within 250 feet of senior living facilities generate the highest average traffic volume, but generate the fewest number of collisions out of all ten land use categories.

Parking lots and NCT districts have the second and third highest crash rates. Unlike parks, parking lots and NCT districts generate significantly high levels of vehicular traffic. One would think that a high average vehicular exposure would mitigate crash frequencies, but in these cases, crash frequency and crash rates are quite high.

Crash frequency, crash density, and crash rate are all individual metrics. Sum-of-ranks, which will be described next, is a composite metric. To identify the sum-of-ranks for the ten categories, the results of the individual metrics are ranked numerically from most dangerous to safest. The sum-of-ranks is the result of adding up how each land use category ranks for each metric and dividing by three. The three individual metrics can be weighted, but they were not in this study. Since a rank of one means that a land use measured as the most dangerous, the lower the sum-of-rank value, the more dangerous the land use is. The equation looks as follows:

$$SR = \frac{\text{Rank}(CD_A) + \text{Rank}(CR_{VV}) + \text{Rank}(CR_{PP})}{3}$$

Figure 28: Sum-of-Ranks Equation

Source: Srinivas S. Pulugurtha, et al, "New Methods to Identify and Rank High Pedestrian Crash Zones."

The results can be ordered from low to high in order to create ranks among the land use variables. Table 13 shows each category's sum of ranks score.

Table 13: Sum-of-Ranks Scores

Category	Sum-of-Ranks Score
NCD	7
Parking Lot	11
NCT	12
Parks	12
Schools	12
Transit Nodes	18
Police and Fire	19
Medical Centers	22
Libraries	23
Senior Living	29

According to the sum-of-ranks composite metric, NCD districts are most highly correlated with more crashes than the other land use variables. Parking lots have a slightly higher score, followed by schools, NCT districts, and parks. It is important to note that parks scored as the most dangerous in two of the three categories, but the least dangerous in crash density. This is due to crashes near parks being diffused among a significantly larger area than the other trip generators.

Like the sum-of-ranks method, a land use's crash score is a composite metric. It differs from a sum-of-ranks by normalizing each metric before comparing them. Further, by normalizing the data, one is able to better diagnose and explain why a certain trip generator is unsafe. For example, if a land use has a disproportionately high crash rate when compared to vehicle volumes, it may mean that drivers are travelling too quickly along relatively uncongested roads. If a trip generator sees disproportionately high crash rates when compared to pedestrian volume, it may mean that motorists are not yielding to pedestrians or the pedestrian realm is lacking in infrastructure.

The crash score equation is as follows:

$$\text{Score } CD_A = \frac{CD_A}{\text{maximum } CD_A} \times 100$$

$$\text{Score } CR_{VV} = \frac{CR_{VV}}{\text{maximum } CR_{VV}} \times 100$$

$$\text{Score } CR_{PP} = \frac{CR_{PP}}{\text{maximum } CR_{PP}} \times 100$$

Figure 29: Crash Score Equation

Source: Srinivas S. Pulugurtha, et al, "New Methods to Identify and Rank High Pedestrian Crash Zones."

The three scores are summed to get a trip generator's final crash score. The land uses can then be ranked from most dangerous to most safe. Table 14 below shows the crash scores.

Table 14: Crash Score Ranks

Category	Crash Score	Rank
Parks	210.64	1
NCD	158.77	2
Parking Lot	130.09	3
NCT	106.31	4
Schools	97.83	5
Transit Nodes	87.89	6
Medical Centers	76.74	7
Police and Fire	52.18	8
Libraries	49.37	9
Senior Living	29.65	10

Whereas NCD districts were deemed most unsafe according to the sum-of-ranks method, parks have the highest crash score. Parks scored a perfect 100 for both crash rate metrics, meaning that it led all other categories in crashes as they relate to pedestrian and automobile volumes. Since parks have a low crash density score, one can confidently claim that parks are not too densely situated in the city. Their ubiquity makes them relatively safer, however crashes are happening at disproportionately high rates as they relate to pedestrian and vehicle volumes.

NCD districts ranked just behind parks. The relatively small footprint of these areas meant crashes were highly concentrated. Because of this, NCDs were the most dangerous of all land uses in terms of crash density. NCD districts also had a disproportionately high crash rate as compared to vehicle volumes. This means that automobile-pedestrian interactions are particularly dangerous in these areas. One can only make educated guesses, but the phenomenon of circling the block looking for on-street car parking spaces is prevalent in NCD districts. Distracted drivers looking for a parking space are more prone to strike a pedestrian than a more aware driver. If this hypothesis is true, an expansion of San Francisco's demand responsive pricing model should be expanded throughout the city.

Parking lots scored third on the list. They scored high on the crash density metric, meaning that there are a high number of crashes as compared to the street miles surrounding parking lots and garages. According to crash rates as a function of vehicle volume, they scored safer than NCD districts. Although care should always be taken when building more parking supply, San Francisco should direct drivers to existing parking lots and garages rather than looking for difficult to find on-street parking in and near NCD districts.

5.5 Logistic Regression

The final method was the creation of a binary logistic regression model. The model was set up to glean the relationship between 18 demographic and land use variables and the odds of there being a crash that results in visible injury or death. Although the totality of the methods until now have only included land use and population variables, the logistic regression model also included driver characteristics, driving conditions at the time of the crash, and if any citations were written at the scene. Although this original research is aimed at learning more about the relationship between land use and safety, by adding extra non-land use variables, one can be more certain that the conclusions reached are valid.

3,985 crashes were ultimately included in the model. Care was taken to remove any records with incomplete data. Of the cases that were included, 2,267 of them resulted in only a complaint of pain. The 1,718 remaining crashes created visible injury or death. The nearly fifty-fifty split lent itself nicely to a logistic regression.

The following table is the regression output. For the purpose of this study, the key columns in the table are the name of each variable, the intercept (B), significance level (Sig.), and odds-ratio (Exp(B)). Several independent variables that had insignificant p-values were removed

from the model. However, since only including statistically significant independent variables could result in a model that is “over-fit,” variables that were integral to this study were included even if they could not disprove the null hypothesis.

Choosing the correct variables in a logistic regression can be more art than science, so this report selected independent variables using a careful, iterative approach. The totality of the variables were included in the model, and the “Backward- Conditional” method was used to iterate. This method works in a stepwise manner: it starts by including every independent variable and, one by one, removing the variable with the highest significance rating. At the end of the process, only the significant variables were included. The 17 iterations informed which variables would ultimately be included in the model.

Table 15: Logistic Regression Output

		Variables in the Equation					80% C.I. for EXP(B)		
		B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
Step 1 ^a	FocusOnTheFive	.092	.098	.880	1	.348	1.096	.967	1.243
	DUI	1.193	.474	6.328	1	.012**	3.298	1.796	6.058
	PedCited	.736	.101	53.429	1	.000**	2.088	1.835	2.376
	IntersectionCrash	-.163	.082	3.935	1	.047**	.850	.765	.944
	Traffic_Calm	-.155	.081	3.702	1	.054*	.856	.772	.949
	Nighttime	.150	.069	4.677	1	.031**	1.162	1.063	1.269
	Libraries	-.271	.141	3.713	1	.054*	.762	.636	.913
	Police_Fire	-.276	.107	6.609	1	.010**	.759	.662	.871
	Transit_Nodes	-.066	.108	.373	1	.541	.936	.814	1.075
	CoC	-.097	.068	2.067	1	.150	.907	.832	.989
	Schools	-.015	.078	.036	1	.850	.985	.892	1.089
	NCT	.096	.058	2.785	1	.095*	1.101	1.023	1.186
	NCD	.112	.072	2.409	1	.121	1.118	1.020	1.226
	Parks	-.010	.048	.044	1	.834	.990	.931	1.053
	PERC_INDUSTRIAL	.494	.274	3.243	1	.072*	1.639	1.153	2.330
	PERC_PUBLIC	.524	.220	5.681	1	.017**	1.689	1.274	2.239
	PERC_COM	-.024	.170	.020	1	.888	.976	.785	1.214
	PERC_RES	.207	.138	2.248	1	.134	1.230	1.030	1.468
	Constant	-.577	.153	14.294	1	.000	.562		

a. Variable(s) entered on step 1: FocusOnTheFive, DUI, PedCited, IntersectionCrash, Traffic_Calm, Nighttime, Libraries, Police_Fire, Transit_Nodes, CoC, Schools, NCT, NCD, Parks, PERC_INDUSTRIAL, PERC_PUBLIC, PERC_COM, PERC_RES.

Variables with an (**) are statistically significant at the 95 percent confidence level.

Variables with an (*) are statistically significant at the 90 percent confidence level.

There were six variables that were statistically significant at the 95 percent confidence level. Ten variables were significant at 90 percent confidence level. If a significance value is equal to or less than 0.05, one can say with 95 percent certainty that there is a statistically significant relationship between the independent and dependent variables and the null hypothesis can be rejected.

Looking at the output, one quickly finds that the type of violation that was written by the police officer at the scene has the strongest positive relationship with a pedestrian being involved in a crash that results in visible injury. If a driver is cited for driving under the influence, the odds of a crash resulting in visible injury increases three-fold. Likewise, if a pedestrian is cited at the scene for an illegal behavior, such as crossing against a light or walking in the street, he is over twice as likely to be involved in a crash that results in injury. Future research needs to be done to learn if this is due to an underreporting of crashes by pedestrian victims, or if responding police officers are more prone to find a pedestrian at fault than a driver when making crash site reports.

Intersection crashes are less likely to result in visible injury than mid-block crashes, and crashes that occur near traffic calming measures such as speed humps, roundabouts, or bulb outs are less likely to result in visible injury. This can be due to intersection crashes being lower speed than

mid-block crashes. These results also support the efficacy of street calming measures in slowing down traffic to make streets safer. This report calls for an increase in mid-block traffic calming measures and the continuation of intersection level pedestrian upgrades. Finally, nighttime crashes were modelled and were statistically significant to increase the odds of a visible injury. A street light audit should be undertaken to address this issue.

Five land use variables were significant at the 90 percent confidence level. Libraries, police and fire stations were all found to decrease the odds of a visible injury. This is heartening because libraries attract a diverse clientele. And there would be serious problems if police and fire stations create dangers rather than mitigate them.

NCT districts, and percentage of public or industrial were all correlated with a slightly higher odds of a crash resulting in visible injury. NCT districts, although statistically significant, had a weaker correlation than public and industrial zoning districts. It should also be noted that, although not statistically significant, the presence of a rapid transit node lessened the odds of a dangerous crash.

Each percentage increase in public and industrial zoning districts, on the other hand, correlated with a 68 percent and 63 percent

greater odds of a crash resulting in visible injury or death. The city should look into additional traffic calming measures when locating new housing developments in and around industrial areas.

Overall, whereas several land use variables were statistically significant, non-land use variables such as driver intoxication and pedestrian citations were stronger predictors of an increased likelihood of a crash resulting in visible injury. This confirms Ukkusuri, et al.'s finding that whereas land use is a useful indirect measure of crash severity, vehicle speeds and street design more directly influence pedestrian safety.⁵¹

51 Satish Ukkusuri, et al., "The Role of Built Environment on Pedestrian Crash Frequency."

SIX

CHAPTER



POLICY RECOMMENDATIONS AND SUGGESTIONS FOR FUTURE RESEARCH

This report endeavored to better understand the relationship between demography and land use with pedestrian crashes. Population, employment, zoning, and land use data was collected and integrated with a crash dataset for the City and County of San Francisco between 2011 and 2015. After an analysis of crashes at a citywide scale, three distinct, localized methods were used. The first method compared crashes with population density, employment density, and land use mix at the block group level. The findings confirm prior research: although there are more crashes in total, *per capita* crash rates are lower in denser block groups than in less dense ones. This relationship was stronger for employment density than population density.

San Francisco's diverse land use mix is represented, and shaped, by its zoning code. Most block groups had a moderate mix of two to four distinct land use types, and fewer were entirely homogenous or heterogeneous. Block groups with a single land use and all five land uses were the two most dangerous land use mix levels.

The second method included the creation of a trip generator typology to closely study the relationship between each trip generator and pedestrian safety. To understand these links, several metrics were calculated. Depending on the metric, different conclusions can be drawn; however, if one uses the normalized, composite crash score metric, parks, NCDs, and parking lots rank as the most dangerous land uses. Police and fire stations, libraries, and senior living facilities were deemed the safest.

A logistic regression model was created to quantify the odds that the presence of an independent variable near a crash site would increase the likelihood of a crash resulting in visible injury or death. The land uses from the typology, as well as demographic and crash location variables were included as independent variables in the model. Although several trip generator land uses were statistically significant, non-land use variables had a stronger predictive relationship for the dependent variable.

There are policy implications from the findings recapitulated above. First, this study has shown the value of addressing pedestrian safety in a localized context rather than at the corridor or citywide level. By highlighting nearby land uses, one is more able to answer the *why* rather than simply the *where* of a crash. It is still impossible to make any causal claims from the research, and care should be taken not to overly generalize the findings, but the methodology can be easily replicated in other municipalities to see if patterns emerge.

From the density and land use calculations, the strength-in-numbers theory was generally supported. Because of this, dense housing or office developments should not be avoided, however extra precautions should be taken to minimize automobile-pedestrian interactions in the densest areas. A congestion pricing scheme in the city's northeast quadrant, its densest, where the densest areas of the city are, would moderate traffic levels and reduce vehicular exposure.

Although denser block groups are safer per capita, one must still reckon with the fact that more people are injured *in total* in dense block groups than less dense ones. Per capita crash rates matter, but San Francisco should be more concerned with gross crash totals when measuring progress in pedestrian safety.

San Francisco should continue its development of mixed-use, walkable communities. Although the block groups with all five land uses were the most unsafe, block groups with multiple land uses were safer than block groups with only one land use. One can only infer as to why this is, but one explanation is that homogenous block groups correlate with higher traffic speeds or less attentive drivers. What urban planning researchers do know is that mixed use developments can better promote short, walkable trips without the need for a private vehicle, so for both safety and quality of life reasons, land use diversity should be promoted. This report recommends further study to see why block groups with five land uses are most unsafe.

Numerous policy recommendations can be made from the trip generator typology.

- **Schools** see a disproportionately high crash rate as compared to other land uses. That is, one sees more crashes per pedestrian or per vehicle near schools than other areas. To counteract this trend, Safe Routes to Schools programs should continue to be funded and children should be encouraged to walk to school with their peers.

Research has shown that parents choose to drive instead of walk with their children to school because of time constraints rather than safety concerns. Rather than have hurried parents converge near schools in vehicles, children should be supported to safely walk in groups to school.

- **NCD** districts routinely scored poorly. Its crash rate suggests that distracted driving is a culprit. NCD districts attract a wide range of interests due to a wide variety of trip destinations, so more time should be taken to study their intricacies. One recommendation is that on-street parking supply should be minimized in favor of curb regulations that promote active uses. San Francisco's demand responsive pricing scheme should be aggressively expanded in NCD districts to minimize circling the block looking for on-street parking which results in distracted driving.

- Although **parking lots** did not fare much better than NCD districts according to

the metrics, they did score safer. Transportation Network Companies (TNCs) have already relaxed the demand for parking in the city's garages, so the city would be well-served to re-route on-street parking demand to off-street locations. Parking lot ingress and egress must be located off of high-demand pedestrian streets, but close enough to still access them by foot.

- **Parks** require more study. This report only partially supported the working theory that parks and open space foster safety. This can be due to the extreme variation in size, amenities, and locations of San Francisco's parks. A pocket park in a dense downtown area will generate different kinds of trips than vast open spaces to the south or west. Future research should separate parks and open spaces into sub-categories for a more refined analysis..

- **NCT** districts scored safer than NCD districts, however they did not perform as well as libraries, senior housing, or emergency services. NCT districts already forbid the expansion of off-street parking supply, but more must be done to address on-street parking. As the Mission Street dedicated bus lane and the Muni Metro L-Taraval projects have shown, San Francisco has not taken an aggressive stance in limiting vehicular exposure in NCT districts. In both of these examples, populist demand for on-street parking trumped safety. Opponents of the each project successfully lobbied for the watering

down of transit improvements in favor of parking retention.

- Rather than sacrifice safety for most for the convenience of a few, the city should be looking to expand its supply of **rapid transit nodes** in and outside of NCT districts. Rapid transit nodes, often located within NCT districts, scored safer than NCT districts as a whole. This means that pedestrians in NCT districts are safer the closer they are to a rapid node.

Each of the above recommendations withstand the results from the logistic regression, but the addition of demographic and crash details to the regression model add more depth to analysis. Although outside the scope of this report, it was found that a pedestrian being cited at the scene of a crash correlated with greater odds that the pedestrian was visibly injured. This is a disconcerting trend and more must be done to explain this finding. Further, a driver being ticketed for a Focus on the Five offense had a statistically insignificant negative relationship with crash severity. This means that when a driver is cited, the crash is less likely to produce a visible injury. This paper recommends the statewide adoption of automated speed enforcement and red light cameras to remove the human element as much as possible from policing traffic infractions. Camera footage can also be used for midblock and intersection level crashes to more objectively measure the culpability of either party.

The logistical regression shows that mid-block crashes more seriously injure pedestrians than crashes in intersections. This finding, coupled with the model clearly stating that traffic calming measures increase safety, means that the city should focus on mid-block safety interventions when possible. This can include mid-block roundabouts, traffic humps, or chicanes.

According to the model, as the share of public zoning within a crash location's service area increases, so does the likelihood of a crash resulting in visible injury or death. As mentioned above, more must be done to understand the relationship of publicly zoned land and safety in San Francisco. A temporal study, as well as a study of crash victim demography may elucidate why public zoning districts scored poorly for safety.

The research in this report is an early step for local planners and safe streets advocates to best understand the conditions for pedestrians in San Francisco. By harnessing current data and using several methods to analyze safety at numerous localized levels, this paper was able to propose achievable and actionable policy recommendations for practitioners and researchers alike. However, as smart-phone enabled shared mobility, managed fleet, and autonomous technologies proliferate, it is paramount that researchers continuously update and reconsider previously held notions. Even though the ways we interact with our built form environments change with technology, we all deserve to be safe.

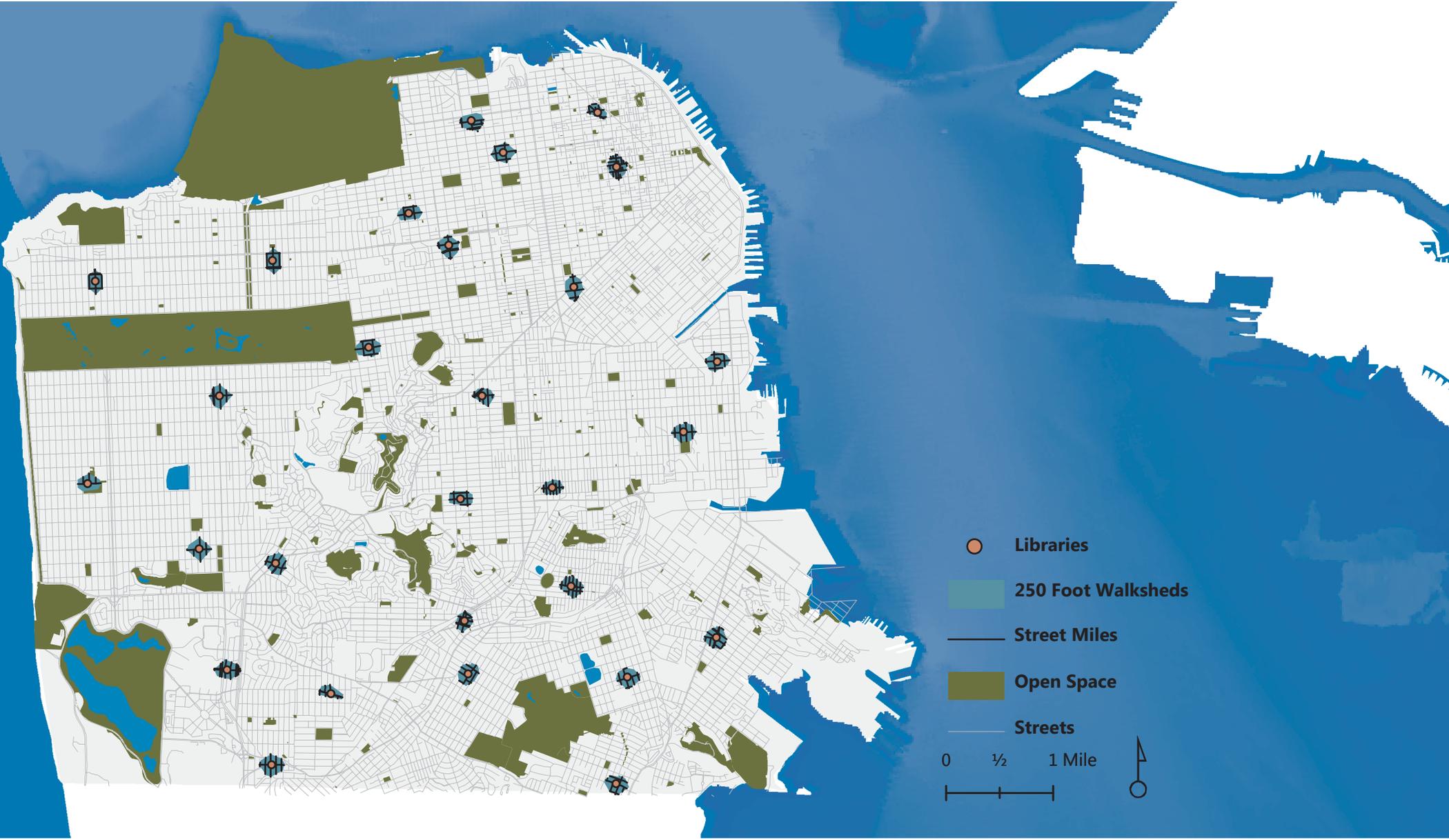
Appendix A

The following series of maps shows the locations, street miles, and 250 foot service areas for each trip generator category. In addition, a map showing the every crash location has been included for reference.



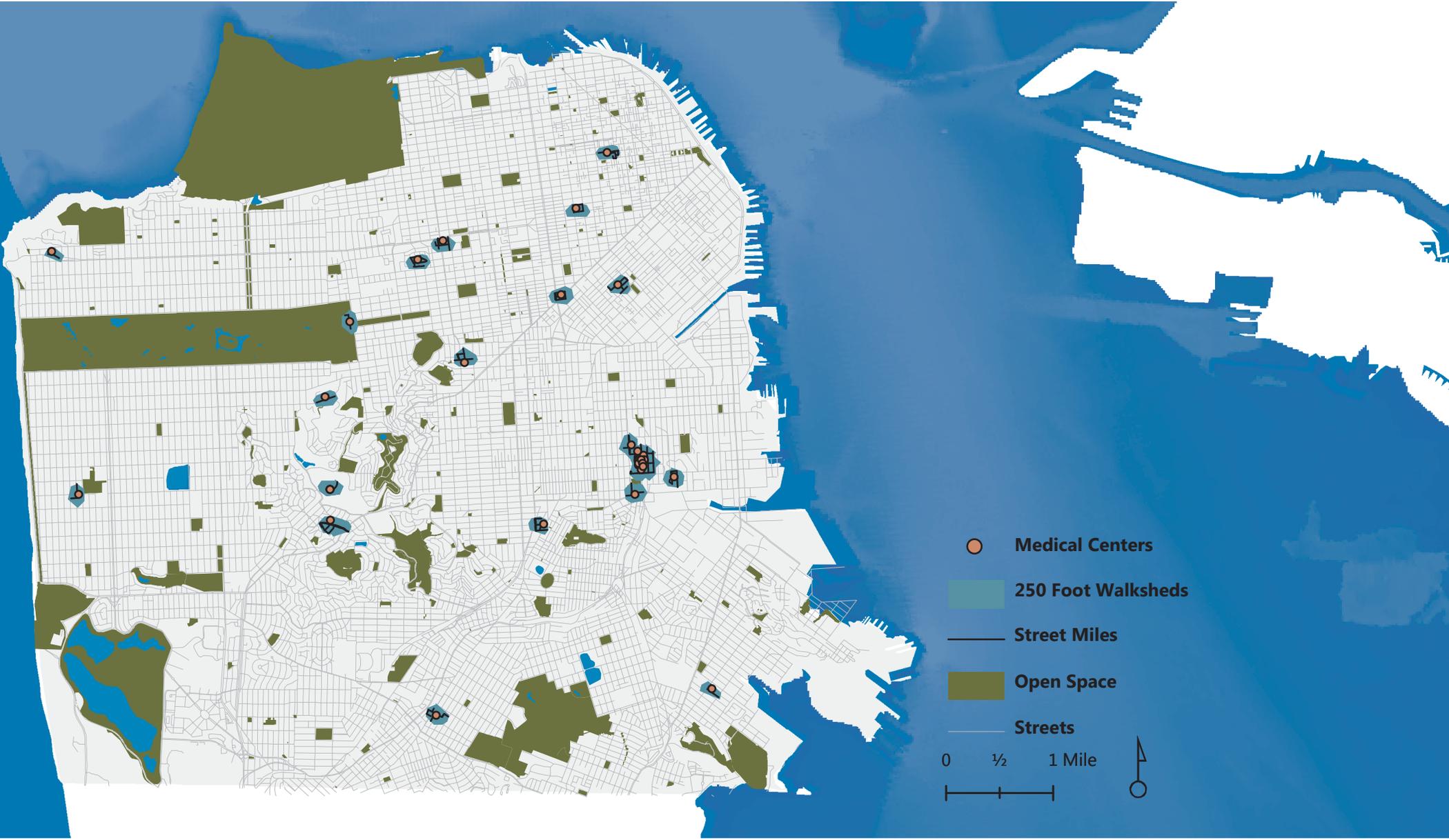
Location of Each Automobile-Pedestrian Crash in San Francisco, 2011-2015

See Appendix B for list of sources



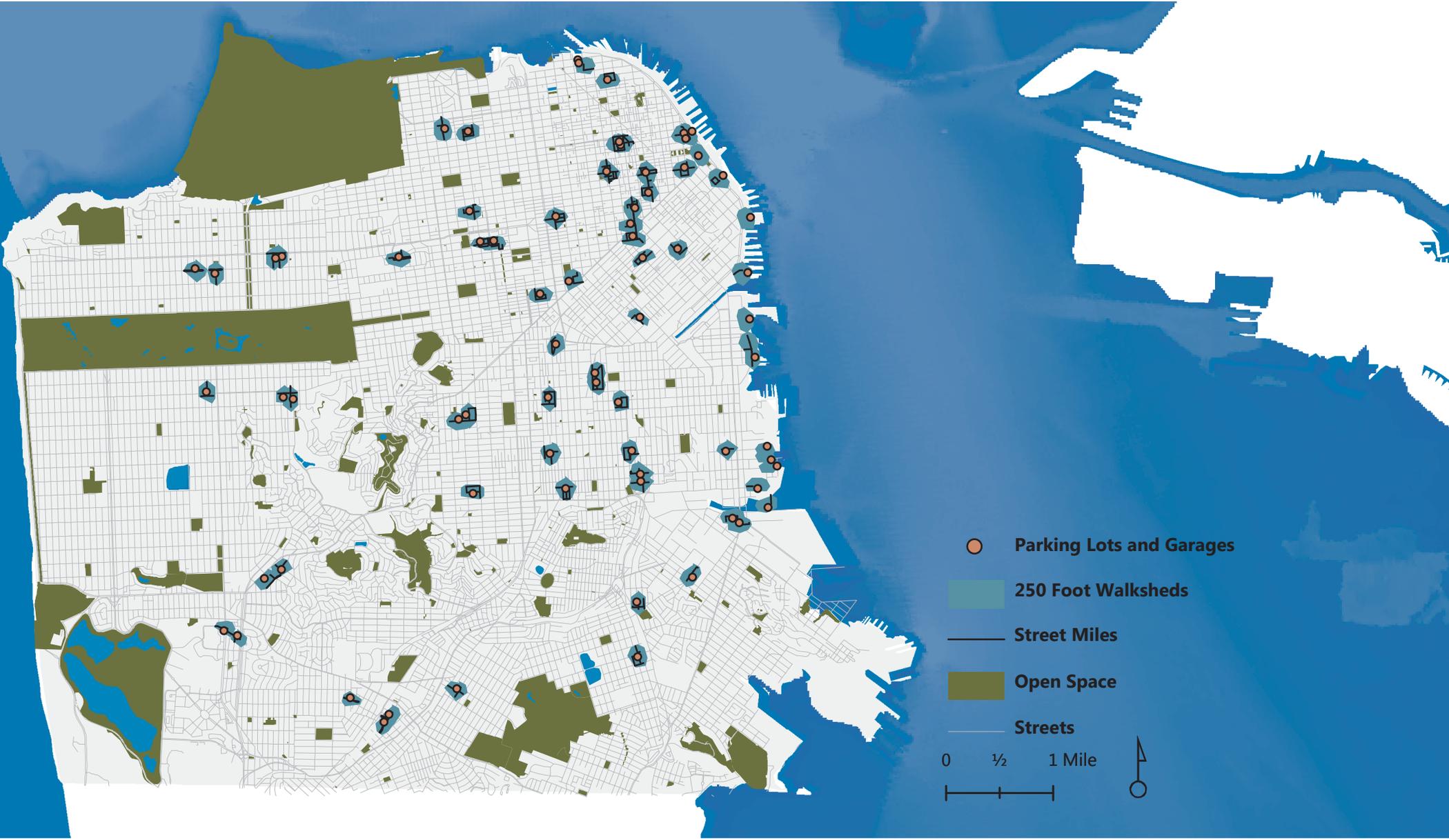
Libraries

See Appendix B for list of sources



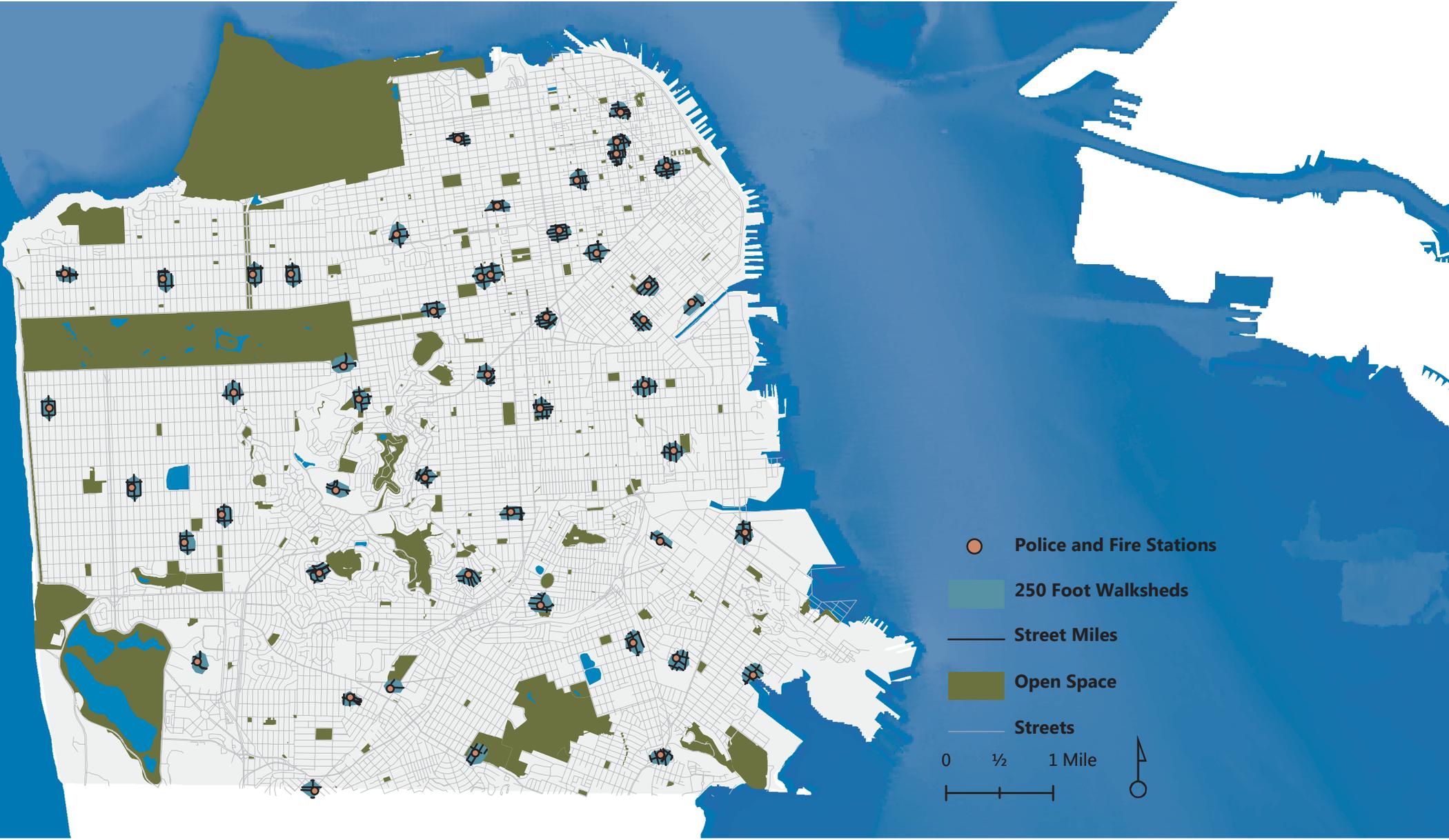
Medical Centers

See Appendix B for list of sources



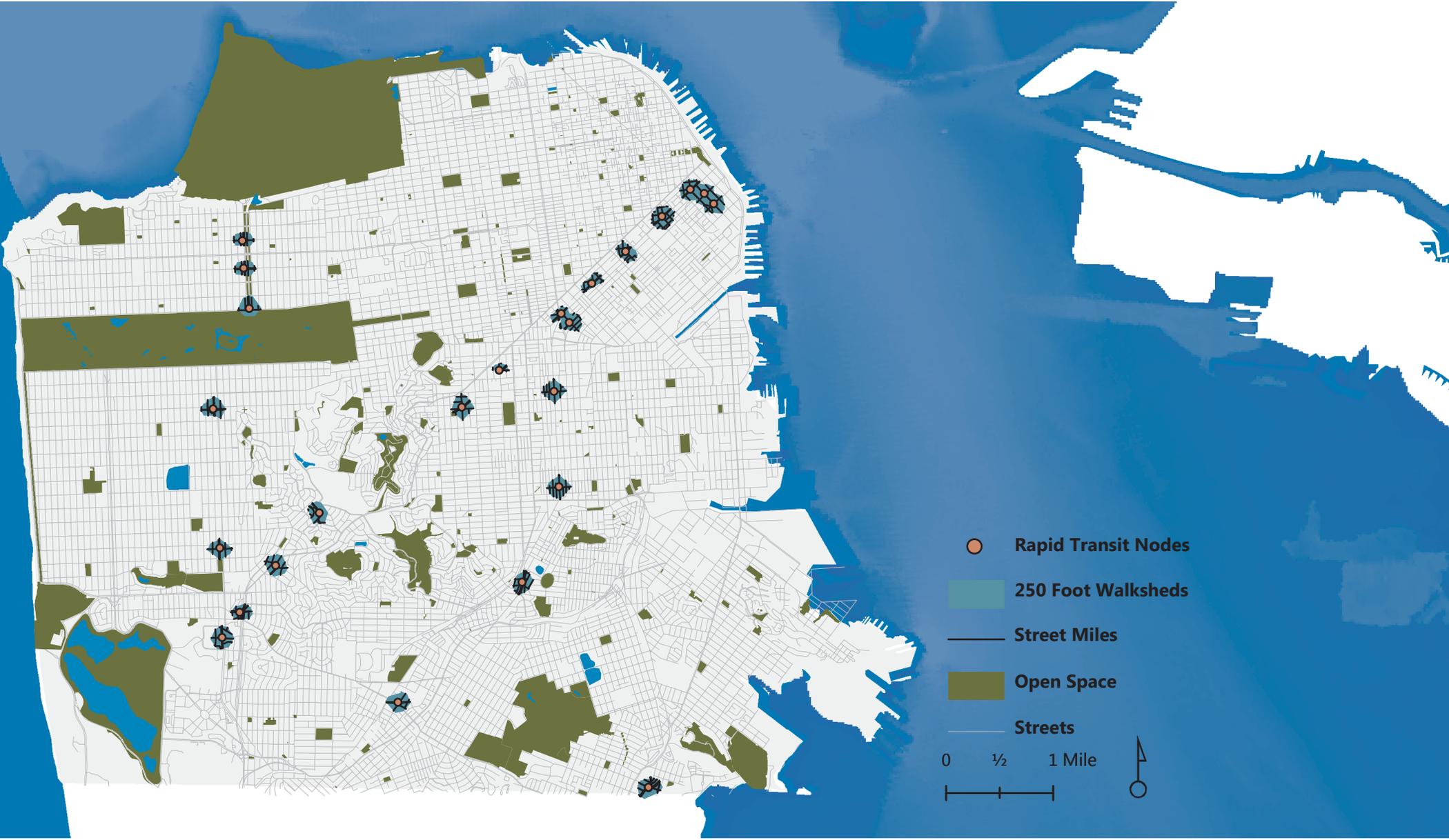
Parking Lots and Garages

See Appendix B for list of sources



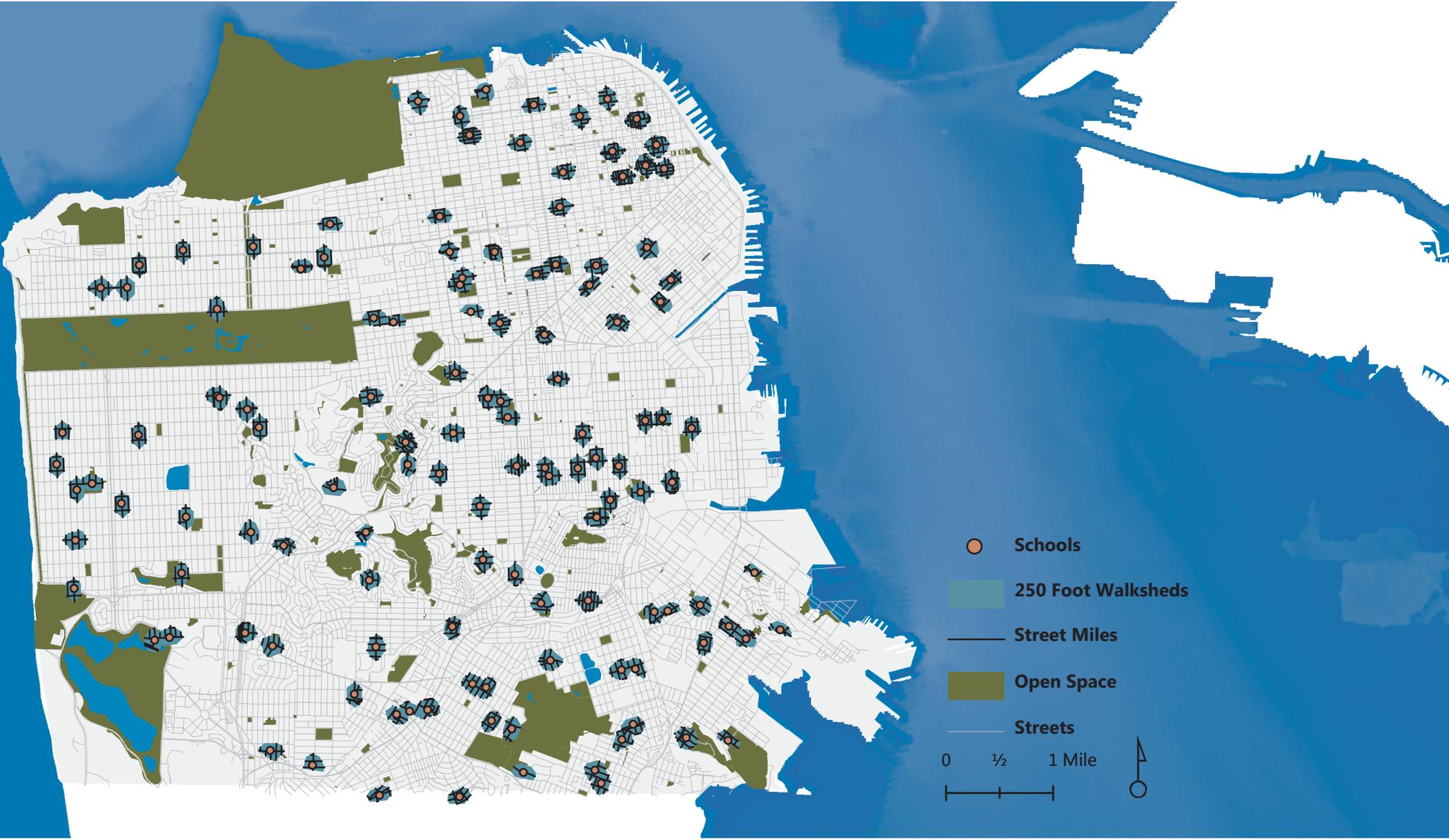
Police and Fire Stations

See Appendix B for list of sources



Rapid Transit Nodes

See Appendix B for list of sources



Schools

See Appendix B for list of sources

Appendix B

The following table lists the categories, online locations, and purpose for each category in the trip generator method.

Data categories and sources

Category	Source	URL	Purpose
Parks	SF Open Data	https://data.sfgov.org/Culture-and-Recreation/Park-Lands-Recreation-and-Parks-Department/42rw-e7xk	Trip Generator Typology
Map of Schools	SF Open Data	https://data.sfgov.org/Economy-and-Community/Map-of-Schools/qb37-w9se	Trip Generator Typology
Library	SF Public Library	https://sfpl.org/pdf/libraries/sfpl421.pdf	Trip Generator Typology
SFMTA/Port	SF Open	https://data.sfgov.org/Transportation/Parking/9qrz-nwix	Trip Generator Typology
Parking Garage	Data		
Fire/Police	SF Open Data	https://data.sfgov.org/Geographic-Locations-and-Boundaries/San-Francisco-Facilities-Current-Zipped-Shapefile-/9vhi-qnsc	Trip Generator Typology
Rapid Transit Nodes	SF Open Data	https://data.sfgov.org/Geographic-Locations-and-Boundaries/San-Francisco-Facilities-Current-Zipped-Shapefile-/9vhi-qnsc	Trip Generator Typology
Annual Pedestrian Volume	SF Open Data	https://data.sfgov.org/Transportation/Map-of-Estimated-Yearly-Pedestrian-Volume-at-Inter/xzu6-znwn	Crash Metrics
Medical Facilities	SF Open Data	https://data.sfgov.org/Geographic-Locations-and-Boundaries/San-Francisco-Facilities-Current-Zipped-Shapefile-/9vhi-qnsc	Trip Generator Typology
Zoning Districts	SF Open Data	https://data.sfgov.org/Geographic-Locations-and-Boundaries/Zoning-Districts/8br2-hhp3	Land Use Mix

Category	Source	URL	Purpose
Neighborhood Commercial District	SF Open Data	https://data.sfgov.org/Geographic-Locations-and-Boundaries/Zoning-Districts/8br2-hhp3	Trip Generator Typology
Neighborhood Commercial Transit District	SF Open Data	https://data.sfgov.org/Geographic-Locations-and-Boundaries/Zoning-Districts/8br2-hhp3	Trip Generator Typology
San Francisco Block Groups	United States Census	https://www.census.gov/geo/maps-data/data/cbf/cbf_blkgrp.html	Density Calculations
San Francisco Population by Block Group, American Community Survey 5-Year Estimate (2011-2015)	American Community Survey	https://factfinder.census.gov/	Population Density
Employment Population	Longitudinal Employer-Household Dynamics	https://lehd.ces.census.gov/	Employment Density
Metropolitan Transportation Commission's Communities of Concern	MTC's ArcGIS Online Data Portal	http://mtc.maps.arcgis.com/home/item.html?id=7ce7b5ba22514340bb7dffdc6bdc4287	Equity Calculation

The following list defines each independent variable in the logistic regression.

Logistic Regression Independent Variables

Category	Description
Focus on the Five	Crashes where a driver was cited for a Focus on the Five infraction was coded as “1.” All other crashes coded with a “0.”
DUI	Crashes where a driver was cited for DUI was coded as “1.” All other crashes coded with a “0.”
Pedestrian Cited	Crashes where a pedestrian was cited at the scene were coded as “1.” If they were not, the crash was coded with a “0.”
Traffic Calming	If a traffic calming feature was located within the service area of a crash, the record was coded with a “1.” All other crashes coded with a “0.”
Nighttime	Nighttime crashes were coded with a “1.” All other crashes coded with a “0.”
Libraries	The count of libraries within the service area of a crash.
Police Fire	The count of police and fire stations within the service area of a crash.
Transit Nodes	The count of transit nodes within the service area of a crash.
Community of Concern	Crashes within the boundary of a CoC coded with a “1.” If not, the crash was coded with a “0.”
Schools	The count of schools within the service area of a crash.
NCT	Crashes within the boundary of an NCT coded with a “1.” If not, crash was coded with a “0.”
NCD	Crashes within the boundary of an NCD coded with a “1.” If not, crash was coded with a “0.”
Parks	The count of parks within the service area of a crash
Percent Industrial	Is the percentage of each service area polygon that is zoned for industrial uses.
Percent Public	Is the percentage of each service area polygon that is zoned for public uses.
Percent Commercial	Is the percentage of each service area polygon that is zoned for commercial uses.
Percent Residential	Is the percentage of each service area polygon that is zoned for residential uses.
Percent Mixed Use	Is the percentage of each service area polygon that is zoned for mixed uses.

References

- Bhatia, Rajiv, and Megan Wier. "Safety in Numbers' Re-examined: Can We Make Valid or Practical Inferences from Available Evidence?" *Accident Analysis & Prevention* 43, no. 1 (2011): 235-240.
- Burnier, Carolina V. "Pedestrian – Vehicular Crashes: The Influence of Personal and Environmental Factors." Master's thesis, University of Maryland, College Park, 2005.
- Clifton, Kelly J., Carolina V. Burnier, and Gulsah Akar. "Severity of Injury Resulting from Pedestrian – Vehicle Crashes: What Can We Learn from Examining the Built Environment?" *Transportation Research Part D: Transport and Environment* 14, no. 6 (2009): 425-436.
- Cottrill, Caitlin D., and Piyushimita (Vonu) Thakuriah. "Evaluating Pedestrian Crashes in Areas with High Low-Income or Minority Populations." *Accident Analysis & Prevention* 42, no. 6 (2010): 1718-1728.
- Dissanayake, Dilum, James Aryaija, and D.M. Priyantha Wedagama. "Modelling the Effects of Land Use and Temporal Factors on Child Pedestrian Casualties." *Accident Analysis & Prevention* 41, no. 15 (2009): 1016-1024.
- Dommes, Aurelie, et al. "Crossing a Two-way Street: Comparison of Young and Old Pedestrians." *Journal of Safety Research* 50, (2014): 27-34.
- Dunbar, George. "The Relative Risk of Nearside Accidents is High for the Youngest and Oldest Pedestrians." *Accident Prevention & Analysis* 45, (2012): 517-521.
- Ewing, Reid, and Eric Dumbaugh. "The Built Environment and Traffic Safety: A Review of Empirical Evidence." *Journal of Planning Literature* 23, no. 4 (2009): 347-367.
- Graham, Daniel J., and Stephen Glaister. "Spatial Variation in Road Pedestrian Casualties: The Role of Scale, Density and Land-use Mix." *Urban Studies Journal Limited* 40, no. 8 (2003): 1591-1607.

Ha, Hoe-Hun, and Jean-Claude Thill. "Analysis of Traffic Hazard Intensity: A Spatial Epidemiology Case Study of Urban Pedestrians." *Computers, Environment and Urban Systems* 35, (2011) 230-240.

Jacobsen, Peter Lyndon. "Safety in Numbers: More Walkers and Bicyclists, Safer Walking and Bicycling." *Injury Prevention* 9, (2003): 205-209.

Kim, Karl, and Eric Yamashita. "Motor Vehicle Crashes and Land Use: Empirical Analysis from Hawaii." *Transportation Research Record* 1784, (2002): 73-79.

LaScala, Elizabeth A., Daniel Gerber, and Paul J. Gruenewald. "Demographic and Environmental Correlates of Pedestrian Injury Collisions: A Spatial Analysis." *Accident Analysis & Prevention* 32, no. 5 (2000): 651-658.

Loukaitou-Sideris, Anastasia, et al. "Death on the Crosswalk: A Study of Pedestrian – Automobile Collisions in Los Angeles." *Journal of Planning and Education Research* 26 (2007): 38-51.

McDonald, Noreen C., and Annette E. Aalborg. "Why Parents Drive Children to School: Implications for Safe Routes to School Programs." *Journal of the American Planning Association* 75, no. 3 (2009): 331-342.

Mirando-Moreno, Luis F., Patrick Morency, and Ahmed M. El-Geneidy. "The Link Between Built Environment, Pedestrian Activity and Pedestrian-Vehicle Collision Occurrence at Signalized Intersections." *Accident Analysis & Prevention* 43, no. 5 (2011): 1624-1634.

Mori, Naglisa, Francisco Armada, and D. Craig Willcox. "Walking to School in Japan and Childhood Obesity Prevention: New Lessons from an Old Policy." *American Journal of Public Health* 102, no. 11 (2012): 2068-2073.

National Highway Traffic Safety Administration. National Pedestrian Crash Report. 2008.

- Neckerman, Kathryn M., et al. "Disparities in Urban Neighborhood Conditions: Evidence from GIS Measures and Field Observation in New York City." *Journal of Public Health Policy* 30, (2009): 264-285.
- Rothman, Linda, et al. "Walking and Child Pedestrian Injury: A Systematic Review of Built Environment Correlates of Safe Walking." *Injury Prevention* 20 (2014): 41-49.
- Schneider, Robert J., Jason Vargo, and Aida Sanatizadeh. "Comparison of US Metropolitan Region Pedestrian and Bicyclist Fatality Rates." *Accident Analysis & Prevention* 106, (2017): 82-98.
- Schneider, Robert J., Rhonda M. Ryznar, and Asad J. Khattak. "An Accident Waiting to Happen: A Spatial Approach to Proactive Pedestrian Planning." *Accident Analysis & Prevention* 36, no. 2 (2004): 193-211.
- Tournier, Isabelle, Aurelie Dommès, and Viola Cavallo. "Review of Safety and Mobility Issues Among Older Pedestrians." *Accident Analysis & Prevention* 91, (2016): 24-35.
- Ukkusuri, Satish, et al. "The Role of Built Environment on Pedestrian Crash Frequency." *Safety Science* 50, no. 4 (2012): 1141-1151.
- Wedagama, D.M. Priyantha, et al., "The Influence of Urban Land-Use on Non-Motorized Transport Casualties." *Accident Analysis & Prevention* 28, no. 6 (2006):1049-1057.
- Wier, Megan, et al. "An Area-Level Model of Vehicle – Pedestrian Injury Collisions with Implications for Land Use and Transportation Planning." *Accident Analysis & Prevention* 41, no. 1 (2009): 137-145.

