VISION ZERØ

CITY OF PHILADELPHIA

PEDESTRIAN SAFETY ACTION PLAN

APPENDICES

May 2021

Appendix A:
Engagement Summary

The survey was developed in SurveyMonkey and included information about Vision Zero policies and programs, pedestrian crash safety trends in Philadelphia, recommended design options to improve traffic safety, a Vision Zero conceptual design toolkit, and optional demographic questions. Survey respondents gave feedback on the specific design improvements and traffic calming solutions they would prefer in their neighborhoods.

The survey was promoted through the Office of Transportation, Infrastructure, and Sustainability's (oTIS) Facebook and Twitter accounts, as well as sent out to stakeholder groups and shared among their networks. The survey was available in the first two weeks of June 2020. Approximately 150 survey responses were collected and analyzed.

The top walkability issues that respondents identified in their neighborhoods included drivers failing to yield, speeding, and failing to obey traffic control devices. Overall, respondents preferred reallocating roadway space to reduce speeding and provide shorter crossing distances, improving visibility at intersections with parking restrictions, and leading pedestrian intervals to increase drivers stopping for pedestrians.

The majority of respondents were residents of South Philadelphia, West Philadelphia, and Northwest Philadelphia, with 60% of respondents between the ages of 25 and 44. Survey participants primarily identified as White/Caucasian. See below for a summary of all survey results. City agencies will take these preferences into consideration when designing future neighborhood street projects, along with additional community input.

Summary of Survey Results

Question 1: What challenges to walking or using a wheelchair have you seen in your neighborhood. (Please check as many that apply).

Answer Choices: Speeding, Sharp turns, Quick lane changes, Not giving right of way to people walking and using wheelchairs, and Failing to obey traffic control devices, Other (please specify)

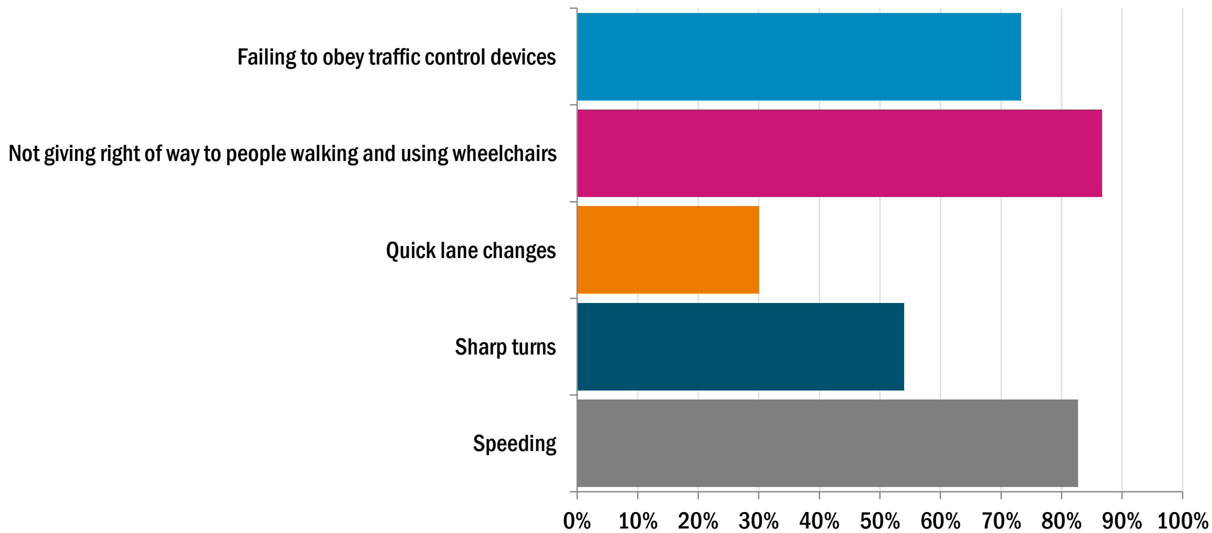
Answers for "Other":

- Too many lanes of traffic (ex. Roosevelt Blvd)
- Poorly maintained or obstructed sidewalks, causing pedestrians to walk in the street
- No sidewalks at secondary street intersections with main roads
- Cars illegally parking in front of curb cuts
- Lack of ADA ramps/curb cuts
- Lack of visibility at intersections due to parking at intersection
- Lack of continuous raised crossings at intersections
- Excessive traffic volumes on minor roads
- Excessive speed limits and lack of traffic calming
- Illegal parking on sidewalks or in pedestrian rights-of-way
- Slip lanes (uncontrolled right turns)

FIGURE 1.

SUMMARY OF RESPONSES TO SURVEY QUESTION 1

The top three responses were drivers not giving the right of way to people walking and using wheelchairs at 87%, drivers speeding at 83%, and drivers failing to obey traffic control devices at 73%.



Source: Survey Monkey, 2020

Question 2: Improving the safety of pedestrians by lowering the speeds of drivers has two major benefits: Lowering injury severity if a crash occurs and Reducing the likelihood of a crash occurring at all. Please rank the following options to share which of these would you like most to see in your neighborhood. (1 - Least interested, 3 - Most interested)

Answer Choices: Roadway Diet, Raised Crossings and Raised Intersections, Automated Enforcement

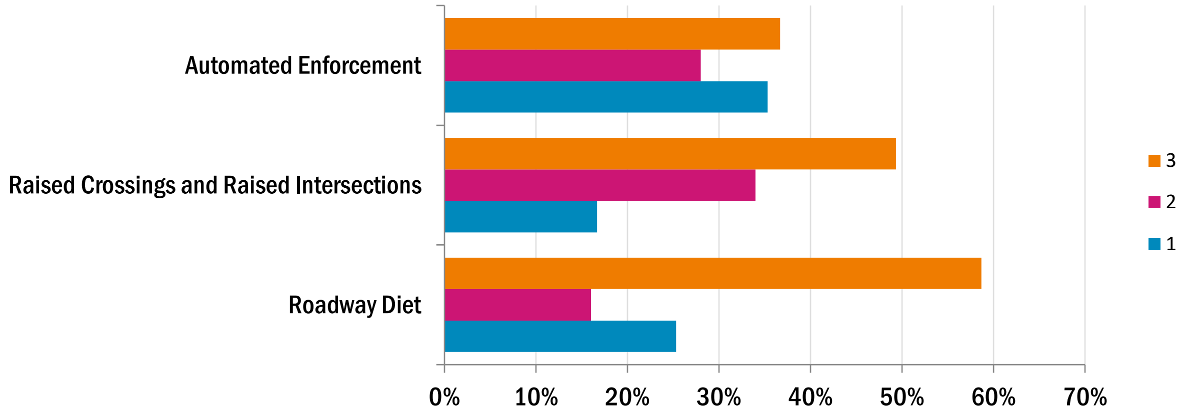
Comments:

- Reduced speed limit in residential areas, including on Roosevelt Blvd
- Make people pay to drive in the city at all
- Automated enforcement cameras should have signs notifying drivers so they are more likely to slow down
- Our immediate neighborhood has stop signs, not lights
- Automated enforcement locations must not be concentrated in certain neighborhoods; equitable placement important
- Would really like bulb-outs installed along York St at the intersections between Frankford Ave and Almond St. The crossings here are just too wide and intersections are only controlled by stop signs
- Most dangerous intersections involve on/off ramps with highways (ex. I-76 with South + Walnut). Raised crosswalks here could be very helpful

FIGURE 2.

SUMMARY OF RESPONSES TO SURVEY QUESTION 2

*To improve the safety of pedestrians by lowering the speeds of drivers, 59% of respondents were most interested in a **Roadway Diet,** 49% of respondents were most interested in **raised crossings and intersections**, and 35% of respondents were most interested in **Automated Enforcement**.*



Source: Survey Monkey, 2020

Question 3: Greater visibility helps everyone see each other with enough time to react and avoid crashes. To make it easier for drivers to see pedestrians, which of these recommendations would you most like to see in your neighborhood? Please rank the following options to share which of these would you like most to see in your neighborhood. (1 - Least interested, 3 - Most interested)

Answer Choices: Roadway Lighting, High Visibility Crosswalks, Intersection Daylighting and Parking Restrictions

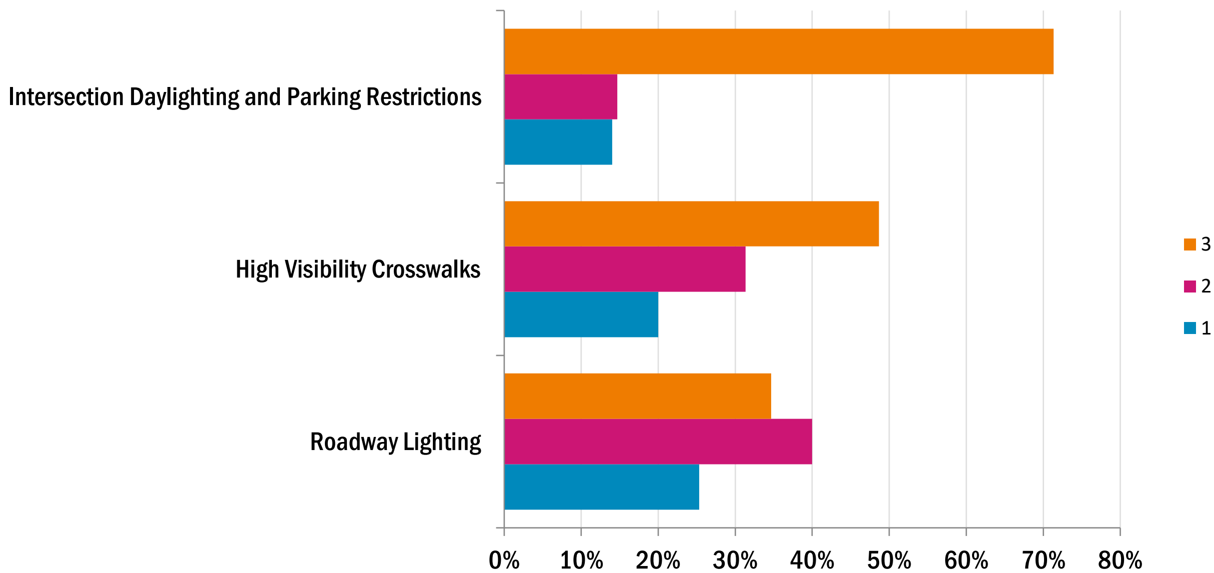
Comments:

- Particularly lighting on crosswalks and high traffic roads
- Doesn't seem like drivers pay attention to high visibility crosswalks when they are there
- Very bright lights are obnoxious and bad for vegetation
- Roadway lighting must take careful account of color temperature and light pollution concerns
- Daylighting seems good on the surface; but overall success also dependent upon no obstructions placed on bump outs, such as elevators to transit stations, newsstands, etc.
- Especially need daylighting design interventions within 19125, though they are probably needed in South Philly and other residential areas. As OTIS is well aware, there is minimal enforcement by the PPD (in residential areas not patrolled by the PPA) of parking violations. These violations are frequent and rarely addressed. It should not be the responsibility of neighbors to call the police on their fellow community members who disregard parking regulations.
- Instead of simply ugly delineators, put in rainwater gardens
- No daylight white lights, please. They make sleep difficult

FIGURE 3.

SUMMARY OF RESPONSES TO SURVEY QUESTION 3

*To make it easier for drivers to see pedestrians, 71% of respondents most interested in **Intersection Daylighting and Parking Restrictions,** 49% were most interested in **High Visibility Crosswalks**, and 25% of respondents were most interested in **Roadway Lighting**.*



Source: Survey Monkey, 2020

Question 4: Shorter crossing distances mean shorter crossing times, reducing the amount of time a person is walking in the street at risk of a crash. To reduce crossing widths for a person walking, which of these recommendations would you most like to see in your neighborhood? Please rank the following options to share which of these would you like most to see in your neighborhood. (1 - Least interested, 3 - Most interested)

Answer Choices: Crossing Islands, Corner Radius Reduction, Road Diets

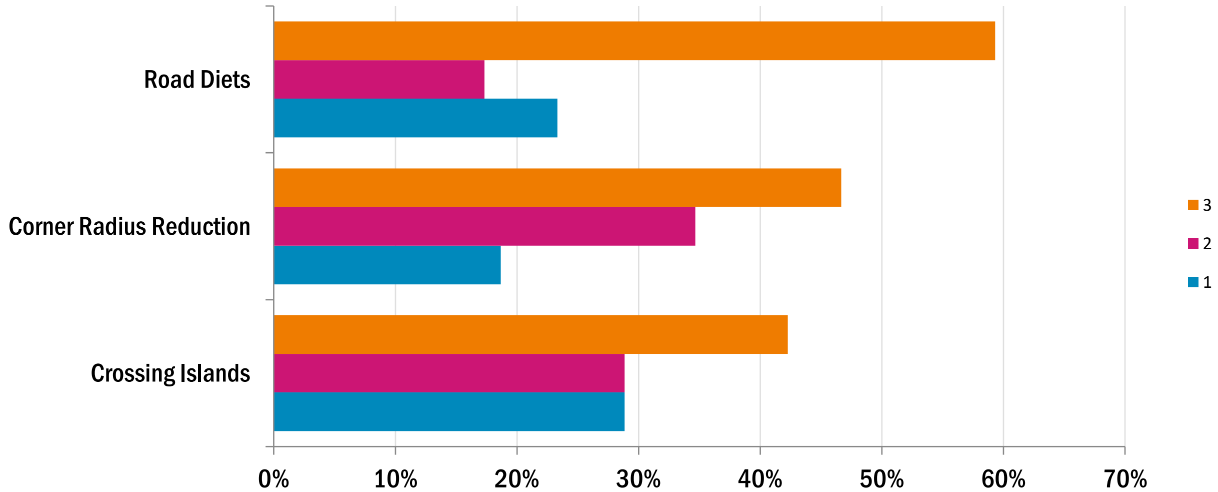
Comments:

- Raised medians with planters and areas level to crosswalks would feel safest
- Why not 'increase' crossing times with longer lights favoring pedestrians? This exists in many cities
- Road narrowing in the 19125 is less of a priority as many of our streets are one-lane and relatively narrow. Other interventions such as pedestrian islands would be useful for the various arterials, and corner radius reduction on those same arterials is also advisable.
- Some of these could incorporate "greening" options that are attractive and good for the environment.
- Paint is not a useful barrier, I think
- There are few opportunities for road diets in South Philly, but daylighting intersections is essential and long overdue. Need PHYSICAL barriers, not just markings that cars can ignore (and ideally not just flex posts that can easily run over either)

- Greatest problem has been drivers who want to make a right turn when light changes and who do not wait for pedestrians to cross first. So they play chicken with pedestrians. Whatever is done should be accompanied by a PR campaign reminding drivers that pedestrians have the right of way as soon as they set foot off of the curb.
- None of these work realistically
- We can’t paint our way out of this problem
- Please pay attention to the way any raised, redefined areas would affect drainage, which is already a new problem because of ADA curb cuts.
- Ideally, corner radius reduction would extend beyond the crosswalk to limit parking in/adjacent to the crosswalk
- Crossing islands feel very car oriented, stranding pedestrians in the middle of a sea of vehicles.
- Only if crossing islands are safe
- I live in South Philly, roadways are already quite narrow. In my opinion, visibility seems to be a huge issue. For example, cars parked up to stop sign leaving no sight lines until vehicle is stopped (if they even do). This ends up being a very high risk for those in wheelchairs, elderly with walkers (who may be hunched over), parents with strollers, and dog walkers, to name a few.
- My neighborhood street is already on a “road diet” (Pine St), which helps a lot but things that would slow turns would be great to make sure drivers are looking for pedestrians about to cross.
- Drivers often stop in crosswalks and move even when pedestrians are in crosswalks. They need to be prevented or strongly discouraged from doing so, which is why fewer lanes and especially crossing bump-ups are needed, among other things.
- I support parking removal when it means safer street for bike and pedestrian users
- Please do not use the metal flapper bars -- the 11th street corridor is a total wreck and after almost a year of installation is still causing havoc for drivers, bikers and pedestrians. The combination of the posts, the zebra areas and the bike lane has made this street MORE dangerous for all types of travel.
- Blocking ability to even park illegally and consistent enforcement
- On streets that are state-controlled by PennDOT, it's been difficult to understand how to change the type of street we are designing for. Since many of these are classified as arterials or even highways, calls for road diets get rejected in a sort of tautological fashion. As we think about the actual engineering and design changes we want to make, I am also interested in seeing changes to the process by which neighborhoods can win approval to change streets, especially when PennDOT is the ultimate decision-maker.
- Reduce crossing widths using GSI to get multiple benefits and real curbs instead of plastic sticks. Crossing islands should be concrete.
- This survey is anti-driver and uses incorrect assumptions. Pull up the National Motorists Association.
- My neighborhood is almost all one lane streets already. These proposals are pretty meaningless.
- Use of sneckdowns to build out walkable areas
- In many blocks in my neighborhood the sidewalks are not at all accessible (Mantua/Belmont), they are in poor repair or non-existent. There are often things blocking the right of way, like parked cars, large trash items, etc. I've reported these issues to 311, which has responded at times.
- Raised crosswalks/intersections can be very beneficial but the grade change was way too minor for the Broad St ones in Center City, so they don't seem accomplish the goal at all. If done correctly, would rate higher.

FIGURE 4.
SUMMARY OF RESPONSES TO SURVEY QUESTION 4

*To reduce crossing widths for a person walking, 59% of respondents were most interested in **Road Diets,** 49% of respondents were most interested in **Corner Radius Reductions**, while 42% of respondents were most interested in **Crossing Islands**.*



Source: Survey Monkey, 2020

Question 5: Reducing the number of potential conflicts between roadway users means reducing the number of eventual crashes. To reduce conflicts between roadway users, which of these recommendations would you most like to see in your neighborhood? Please rank the following options to share which of these would you like most to see in your neighborhood. (1 - Least interested, 3 - Most interested)

Answer Choices: No Turn on Red, Protected Turn Phases, Crossing Islands,

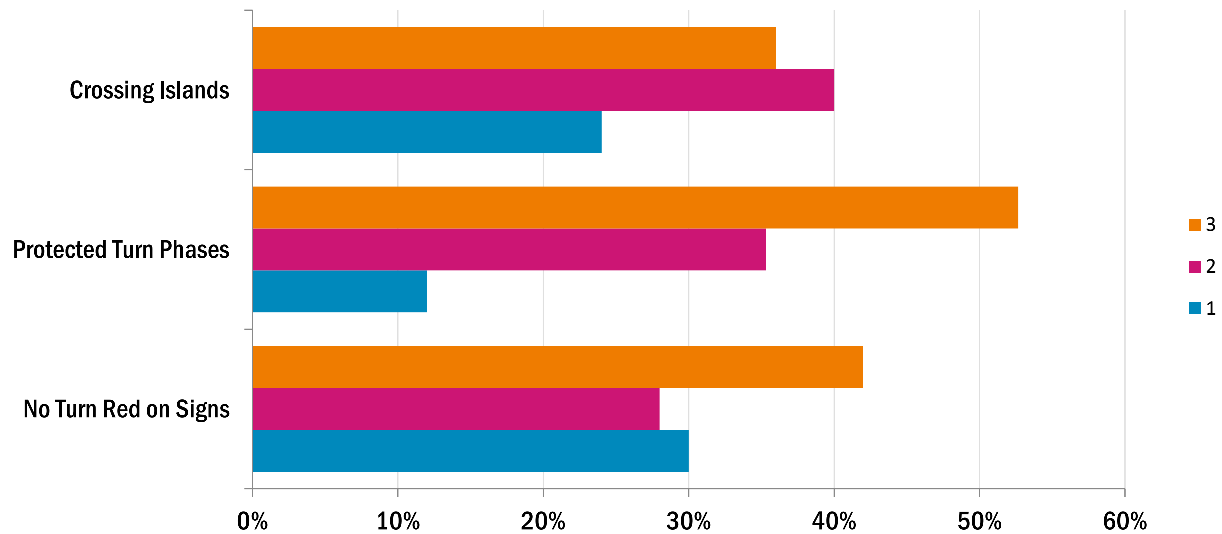
Comments:

- Which is preferred depends on context of the particular intersection
- Engineering is always a better solution than signs
- I think that physical changes (e.g. road diet) are more effective than signs. A cross island means that the road is too wide!
- Should consider city-wide No Turn on Red
- These interventions are all needed for my area's various arterials, including Aramingo Ave and Lehigh Ave (separate study for the latter already underway).

FIGURE 5.

SUMMARY OF RESPONSES TO SURVEY QUESTION 5

To reduce conflicts between roadway users, 53% of respondents were least interested in ***Protected Turn Phases,*** *42% of respondents were most interested in* ***No Turn on Red Signs,*** *while 30% of respondents were most interested in* ***Crossing Islands.***



Source: Survey Monkey, 2020

Question 6: If drivers are not stopping for pedestrians, then they are at a higher risk for hitting them. To increase the number of people driving that yield for pedestrians, which of these recommendations would you like to see in your neighborhood? Please rank the following options to share which of these would you like most to see in your neighborhood. (1 - Least interested, 4 - Most interested)

Answer Choices: Hardened Centerlines and Turn Wedges, Roundabouts, Leading Pedestrian Intervals

Comments:

- Roundabouts can be difficult for pedestrians as vehicles are focused on yielding/merging with other vehicles and navigating/emerging at the right exit.
- Roundabouts make drivers tense and cause confusion for everyone. Drivers ignore flashing lights; steady red lights would be more effective.
- I love roundabouts, having lived abroad for many years, but how practical that is in a city environment with narrow streets is my question
- Roundabouts are incompatible with the fabric of most neighborhoods and often give even more space over to cars
- Roundabouts are really hard to safely navigate as a pedestrian. There isn't a culture of respecting crosswalks or stopping for pedestrians in Philly, so the painted crossings won't help.
- Roundabouts should be considered but only if they are single lane and have separate bike and pedestrian paths.
- Why did your categories change from 3 to 4? This seems to be a fatal flaw for data comparably; also for RRFBs graphic misleading

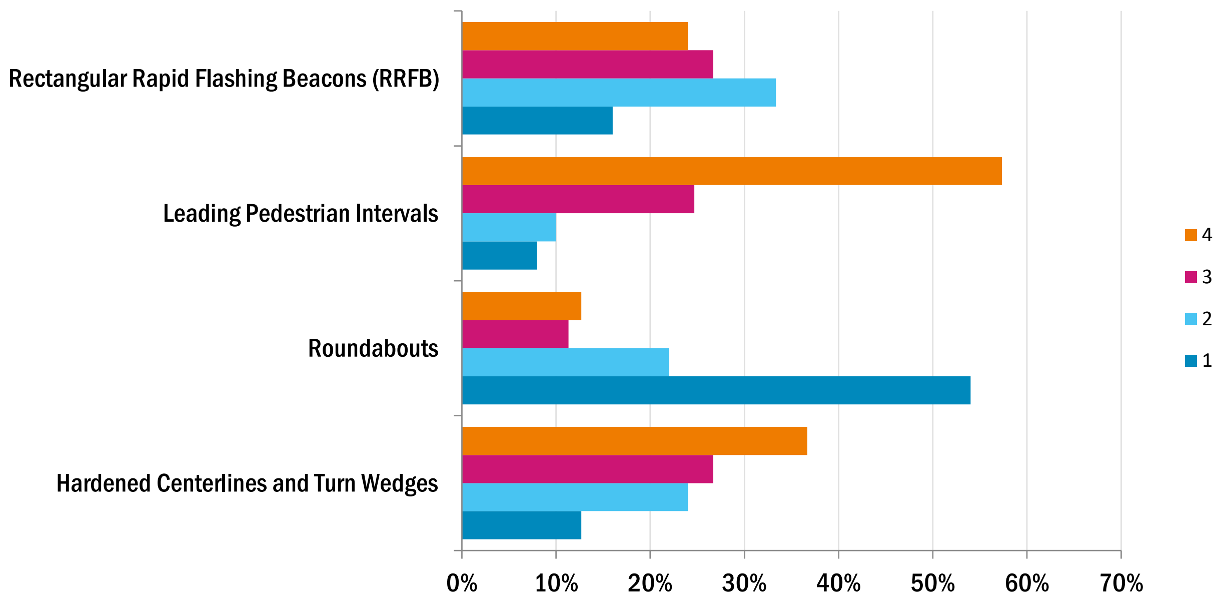
as it also represents a mid-block crossing, which was never mentioned previously.

- Also need audible crossings at commercial intersections for visually impaired
- Less keen on RRFB, despite studies insisting they work. My own observations of RRFBs installed in CC near the Gallery show drivers are not necessarily inclined to stop for pedestrians there. Peds feel uncertain about entering these crossings too. Roundabouts are interesting and one has been proposed for Frankford and York; will be interested to see how that works when/if it is ultimately built. LPIs, on the other hand are greatly needed for Aramingo and Lehigh. Should also be considered for Frankford Ave, which experiences high foot traffic. Also like the turn wedges/centerlines to slow turning motor vehicles.

FIGURE 6.

SUMMARY OF RESPONSES TO SURVEY QUESTION 6

*To increase the number of people driving that yield for pedestrians, 57% of respondents were most interested in **Leading Pedestrian Intervals,** 37% of respondents were most interested in **Hardened Centerlines & Turn Wedges**, and 54% of respondents were least interested in **Roundabouts.***



Source: Survey Monkey, 2020

Question 7: Please indicate the nature of your relationship to the City of Philadelphia (Please check as many that apply)

Answer Choices: I live here, I work here, I own a business here, Other (please specify)

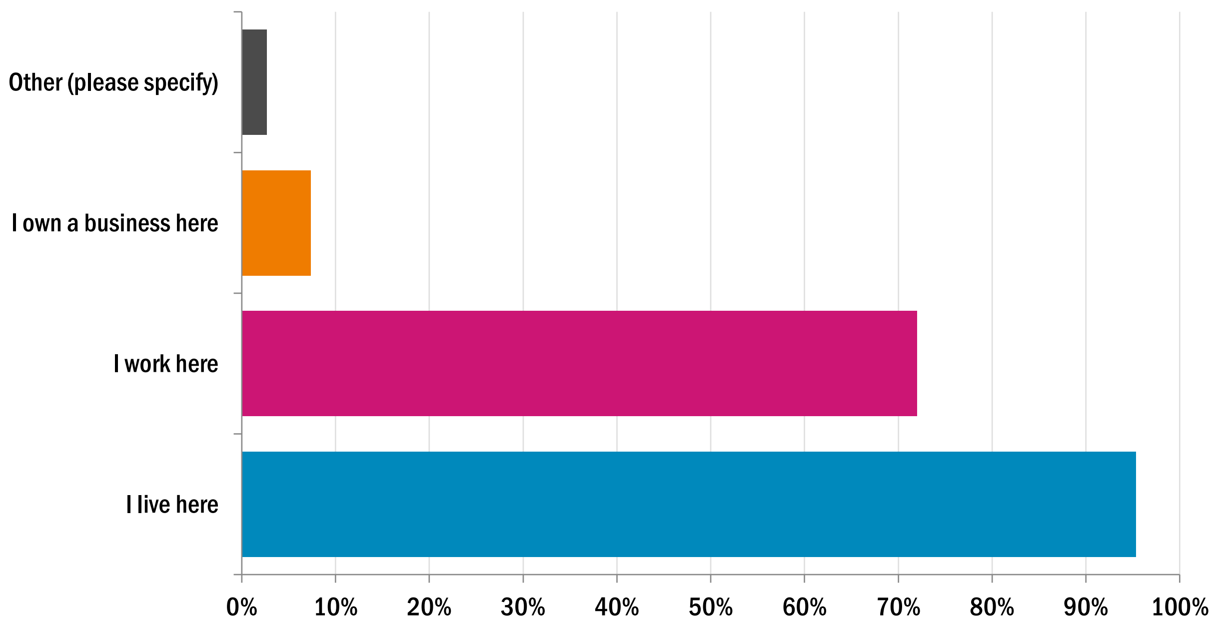
Answers for "Other":

- I'm a student
- I walk all over Center City and parts of West Philly, so I have to cross streets all the time. I've been almost hit by cars too many times to name, and it is getting worse, along with speeding in the city.

FIGURE 7.

SUMMARY OF RESPONSES TO SURVEY QUESTION 7

Approximately 95% of respondents live in Philadelphia and 72% work in the City. Only 7% of the respondents owned a business in the City.



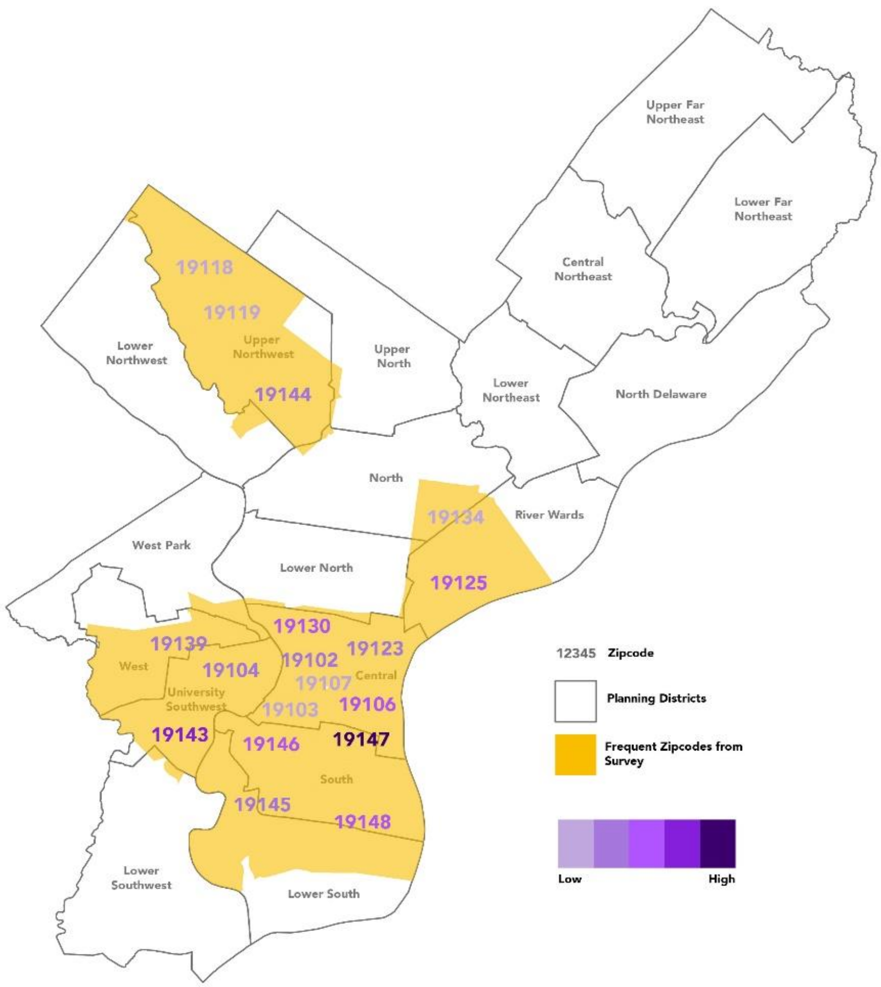
Source: Survey Monkey, 2020

Question 8: What is your home zip code?

FIGURE 8.

MAP OF RESPONSES TO SURVEY QUESTION 8

Approximately 39% of respondents live in South Philadelphia and 16% live in West Philadelphia while only 9% lived in North Philadelphia, 10% lived in Center City, and 2% lived in Northeast. There wasn't a balanced representation of respondents in the priority locations. Priority intersections, areas, and corridors are concentrated in Northeast, North, Center City, West Philadelphia, and South Philadelphia with hotspots of pedestrian fatalities and injuries.



Source: Survey Monkey, 2020

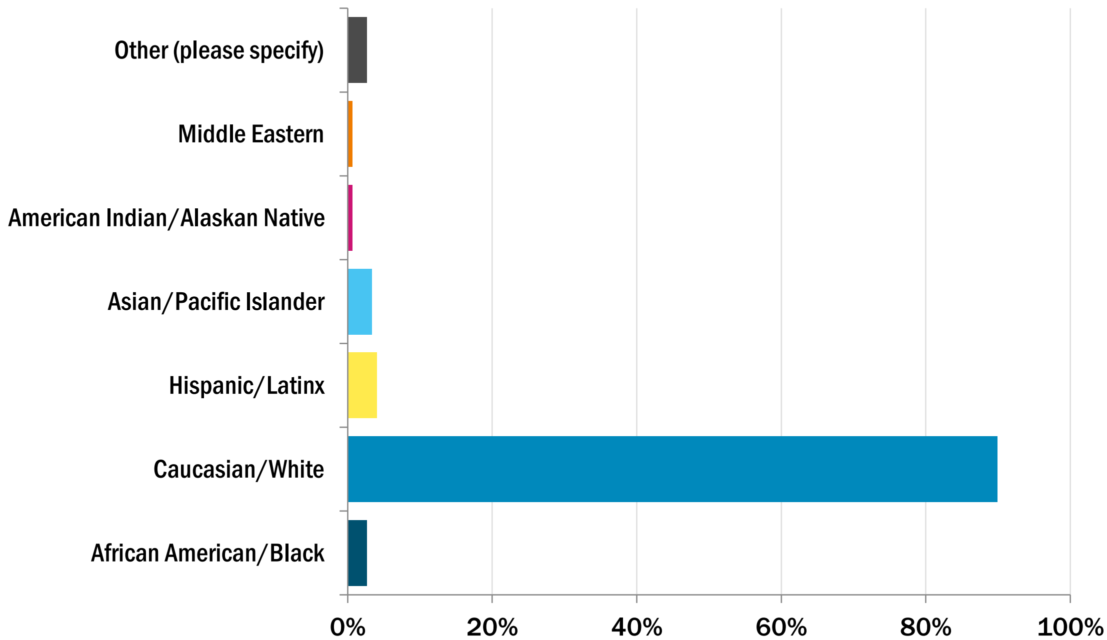
Question 9: What is your race/ethnicity? (Please check as many that apply)

Answer Choices: African American/Black, Caucasian/White, Hispanic/Latinx, Asian/Pacific Islander, American Indian/Alaskan Native, Middle Eastern, Other (please specify)

FIGURE 9.

SUMMARY OF RESPONSES TO SURVEY QUESTION 9

Approximately 90% of respondents identified as Caucasian/White and a little less than 3% of respondents identified as African American/Black. The survey responses are, therefore, not representative of the citywide population overall, which has a 44% African American/Black population.



Source: Survey Monkey, 2020

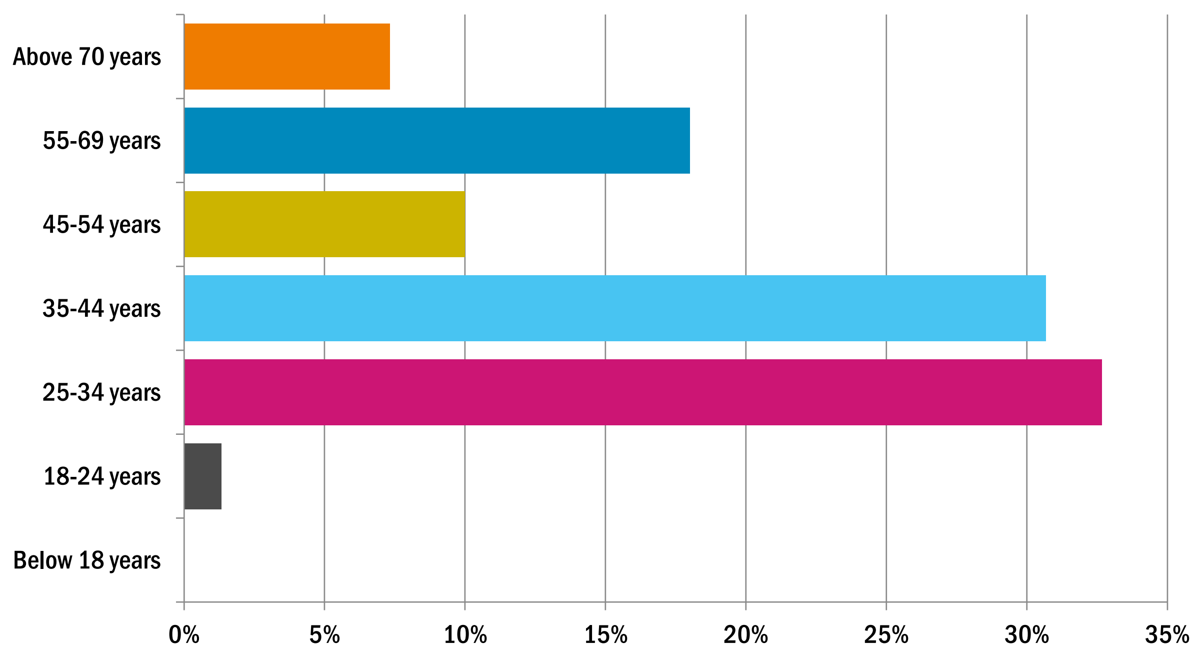
Question 10: What is your age?

Answer Choices: Below 18 years, 18-24 years, 25-34 years, 35-44 years, 45-54 years, 55-69 years, Above 70 years

FIGURE 10.

SUMMARY OF RESPONSES TO SURVEY QUESTION 10

While more than 60% of survey respondents were between the ages of 25-44, this population group comprises only about 30% of Philadelphia residents. Approximately 1% of survey respondents were between the ages of 18-24. Male drivers 20-29 are likely to cause more severe pedestrian crashes. There were no respondents under the age of 18 that participated in the survey. Both age groups were underrepresented in the survey.



Source: Survey Monkey, 2020

Question 11: How did you find out about this survey?

Answer Choices: Social Media, Word of mouth, City website, Other (please specify)

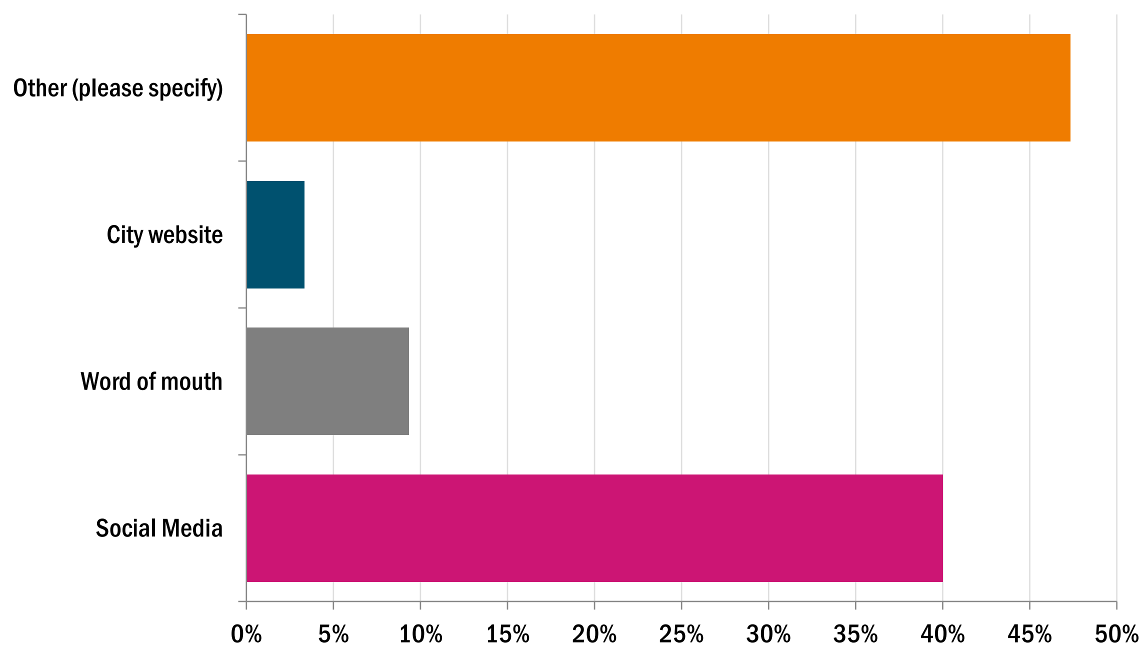
Other:

- Email list
- 5th Square newsletter
- Friends of Cloverly Park
- Community newsletter
- Community garden
- Twitter
- Newspaper
- Local neighborhood Facebook page
- NMA

FIGURE 11.

SUMMARY OF RESPONSES TO SURVEY QUESTION 11

Approximately 40% of survey respondents find out about the survey through social media. *47% of survey respondents found out about the survey from other sources, including a newsletter from 5th Square.*



Source: Survey Monkey, 2020

Question 12: Please provide your email address to receive updates on the Action Plan.

75 respondents provided their email to receive updates on the Action Plan.

Appendix B: Prioritization Top 50 Pedestrian Intersections and Corridors

PRIORITY INTERSECTIONS

To create a list of priority intersections, all pedestrians injured or killed in crashes at each intersection were added together. Pedestrian fatalities were weighted four times larger than injuries. Aligning with the City of Philadelphia's Vision Zero goal of bringing traffic deaths to zero by 2030, this prioritizes intersections with high numbers of pedestrian fatalities. Each intersection was then sorted by its score. For example, at Harbison Ave and Roosevelt Boulevard, there were three pedestrian fatalities (weighted by four, creating a score of 12) and nine pedestrian injuries between 2014-2018, totaling a score of 21 for that intersection. Many intersections have the same score and are listed as ties. For example, the intersection of Allegheny Ave and Germantown Ave has a score of 18 and the intersection of Roosevelt Blvd and Faunce St/Revere St also has a score of 18. Therefore, both intersections have the same rank (#2). The table below documents the top 50 priority intersections.

TABLE 1.

TOP FIFTY PRIORITY PEDESTRIAN INTERSECTIONS IN PHILADELPHIA

Rank	Intersection	Pedestrian Fatalities (people)	Pedestrian Injuries (people)	Total Pedestrian Fatalities and Injuries (people)
1	Bustleton Ave/Levick St & Roosevelt Blvd	4	3	7
2	W Allegheny Ave & Germantown Ave	4	2	6
2	Faunce St/Revere St & Roosevelt Blvd	4	2	6
4	Harbison Av & Roosevelt Blvd	3	9	12
5	N 2nd St & W Lehigh Ave	3	7	10
6	Large St & Roosevelt Blvd	3	0	3
7	Whitaker Ave/Adams Ave & Roosevelt Blvd	2	7	9
8	N 9th St & Roosevelt Blvd	2	6	8
9	Arch St & N Broad St	2	5	7
10	E Allegheny Ave & Aramingo Ave	2	4	6
11	A St & E Lehigh Ave	2	0	2
12	N 16th St & John F Kennedy Blvd	1	14	15
13	N Broad St & W Lehigh Ave	1	13	14

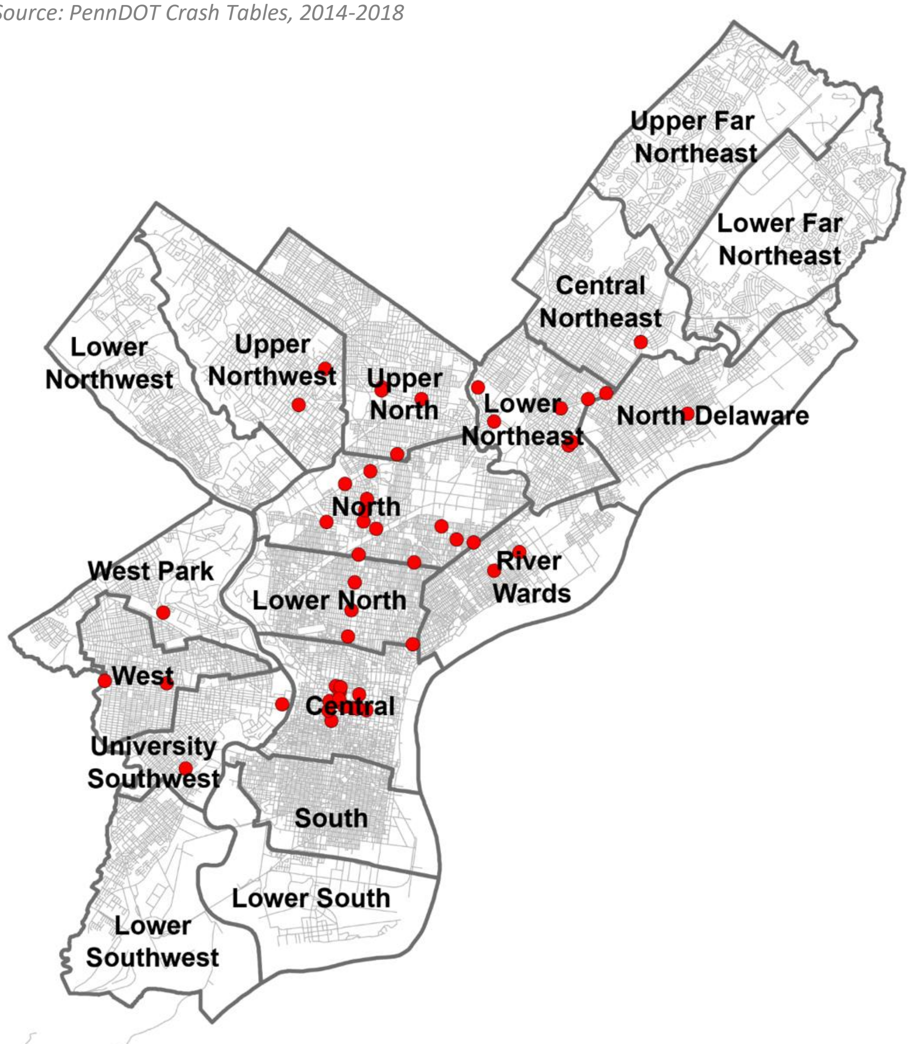
Rank	Intersection	Pedestrian Fatalities (people)	Pedestrian Injuries (people)	Total Pedestrian Fatalities and Injuries (people)
14	N 15th St & Vine St	1	11	12
15	S 15th St & Locust St	1	9	10
15	N 11th St & Market St	1	9	10
17	E Chelten Ave & Germantown Ave	1	8	9
18	B St & E Ontario St	1	7	8
18	S 54th St & Woodland Ave	1	7	8
20	Frankford Ave & Pratt St	1	5	6
20	E Allegheny Ave & E St	1	5	6
20	Aramingo Ave & E Venango St	1	5	6
20	N Broad St & W Tioga St	1	5	6
20	Cottman Ave & Torresdale Ave	1	5	6
25	N 38th St & Market St	1	4	5
25	Welsh Rd & Roosevelt Blvd	1	4	5
27	N 59th St & Market St	1	3	4
27	N 57th St & W Girard Ave	1	3	4
27	S 23rd St & Chestnut St	1	3	4
27	W Lehigh Ave & Mascher St	1	3	4
27	S 15th St & Washington Ave	1	3	4
27	S 62nd St & Chestnut St	1	3	4
27	N 54th St & City Ave	1	3	4
27	Pratt St & Roosevelt Blvd	1	3	4
36	S 21st St & Jackson St	1	2	3
36	Pennypack St & Yale Pl	1	2	3
36	Race St & N Watts St	1	2	3
36	S 11th St & Spruce St	1	2	3
36	N Franklin St & Race St	1	2	3
36	Blue Grass Rd & Welsh Rd	1	2	3
36	E Lehigh Ave & N Front St	1	2	3
36	N 2nd St & W Allegheny Ave	1	2	3
36	S 23rd St & W Passyunk Ave	1	2	3
36	N Broad St & W Glenwood Ave	1	2	3

Rank	Intersection	Pedestrian Fatalities (people)	Pedestrian Injuries (people)	Total Pedestrian Fatalities and Injuries (people)
36	Frankford Ave & Morrell Ave	1	2	3
36	Bustleton Ave & Stanwood St	1	2	3
36	S 11th St & Walnut St	1	2	3
36	Frankford Ave & Sheffield Ave	1	2	3
36	Cottman Ave & Oakland St	1	2	3
36	Bleigh Ave & Castor Ave	1	2	3
36	Cottman Ave & Erdrick St	1	2	3
36	N 63rd St & Lansdowne Ave	1	2	3
36	Island Ave & Lindbergh Blvd	1	2	3

FIGURE 12.

MAP OF TOP FIFTY PRIORITY PEDESTRIAN INTERSECTIONS IN PHILADELPHIA

Source: PennDOT Crash Tables, 2014-2018



PRIORITY CORRIDORS

To create priority corridors, crashes that occurred along each corridor were added together. Corridors are segments of streets that are contiguous, have the same street name, functional classification (e.g. major arterial, minor arterial, expressway), Complete Streets typology (from the City of Philadelphia's 2017 Complete Streets Handbook, which created street typologies such as Urban Arterial, Park Road, City Neighborhood Street), and are longer than 1,000 feet. To create a list of priority corridors, all pedestrians injured or killed in crashes in each corridor were added together. Pedestrian fatalities were given a weight four times larger than an injury. Corridors were then sorted by their "score": pedestrian injuries and pedestrian fatalities (weighted by four) added together.

Below is a Top 50 list of priority corridors ranked by number of pedestrian injuries, then by pedestrian fatalities. This list can be a foundation for plans to improve pedestrian safety.

TABLE 2.

TOP FIFTY PRIORITY PEDESTRIAN CORRIDORS IN PHILADELPHIA

Rank	Corridor	Pedestrian Fatalities (people)	Pedestrian Injuries (people)	Total Pedestrian Fatalities and Injuries (people)	Corridor Length (miles)
1	Roosevelt Blvd - from Schuylkill River to Bucks County Line	31	132	163	14.70
2	N Broad St - from City Hall to Glenwood	5	177	182	3.04
3	N Broad St - from Glenwood to Windrim	5	138	143	2.26
4	S Broad St - from City Hall to Oregon	0	110	110	2.44
5	Market St - from City Hall to 2nd	1	85	86	1.02
6	Allegheny Ave - from Sedgley to Ridge	2	62	64	1.60
7	N Broad St - from Lindley to Montgomery County Line	0	82	82	2.29
8	Chestnut St - from Independence Mall to 20th	0	79	79	1.31
9	Kensington Ave - from Front to Pacific	0	73	73	1.87

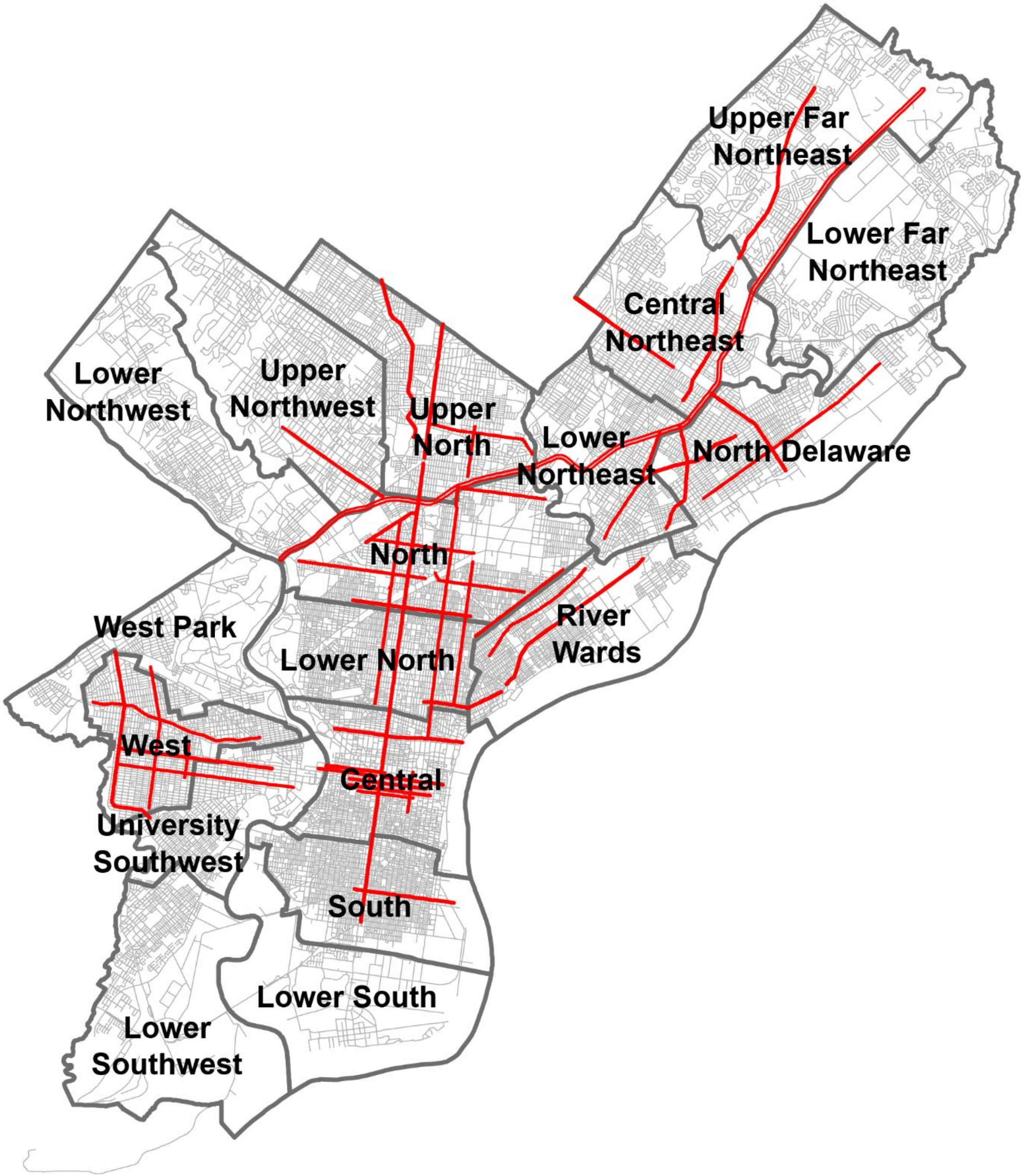
Rank	Corridor	Pedestrian Fatalities (people)	Pedestrian Injuries (people)	Total Pedestrian Fatalities and Injuries (people)	Corridor Length (miles)
10	Chestnut St - from Cobbs Creek to 38th	1	70	71	2.62
11	Allegheny Ave - from Reach to Sedgley	2	63	65	2.18
12	Lehigh Ave - from 24th to 6th	1	63	64	1.54
13	Spruce St - from Cobbs Creek to 33rd	0	67	67	3.02
14	63rd St - from Market to Lancaster	2	55	57	1.56
15	Aramingo Ave - from Westmoreland to Orthodox	5	42	47	1.72
16	Torresdale Ave - from Linden to Benner	3	50	53	3.78
17	Girard Ave - from Franklin to Aramingo	0	60	60	1.38
18	Cottman Ave - from Roosevelt to Wissinoming	2	51	53	1.85
19	Walnut St - from 20th to 4th	1	55	56	1.39
20	Bustleton Ave - from Tyson to Benton	4	42	46	2.45
21	Spring Garden St - from Benjamin Franklin Pkwy to Columbus	0	58	58	2.19
22	Hunting Park Ave - from Pacific to N Broad	0	56	56	1.00
23	Ogontz Ave - from Somerville to Montgomery County Line	0	56	56	2.92
24	Erie Ave - from Roosevelt to 2nd	0	55	55	1.77
25	Frankford Ave - from Bridge to Kensington	1	51	52	1.18
26	Market St - from City Hall to 30th	0	55	55	0.97

Rank	Corridor	Pedestrian Fatalities (people)	Pedestrian Injuries (people)	Total Pedestrian Fatalities and Injuries (people)	Corridor Length (miles)
27	Olney Ave - from N Broad to Tabor	0	54	54	1.91
28	Locust St - from 18th to Washington Square	1	48	49	0.94
29	Frankford Ave - from Bridge to Tyson	0	50	50	1.79
30	Haverford Ave - from 68th to Lancaster	0	50	50	2.97
31	Wayne Ave - from Lincoln to Windrim	1	46	47	2.08
32	Aramingo Ave - from Norris to Westmoreland	2	41	43	1.51
33	John F Kennedy Blvd - from Schuylkill to Juniper	1	45	46	0.99
34	57th St - from Lancaster to Baltimore	1	43	44	2.42
35	2nd St – from Girard to Glenwood	4	30	34	2.50
36	Bustleton Ave - from Hendrix to Winchester	1	42	43	3.22
37	Frankford Ave - from Sedgley to York	0	46	46	2.23
38	Lehigh Ave – 6th to Mutter	4	30	34	0.50
39	16th St - from John F Kennedy to Walnut	1	40	41	0.28
40	Bustleton Ave -	4	28	32	
41	Harbison Ave -	3	32	35	
42	Snyder Ave - from 16th to Columbus	0	44	44	1.70
43	Wyoming Ave - from Roosevelt to H	0	44	44	1.50
44	52nd St - from Arch to Pine	0	43	43	0.56
45	8th St - from Arch to Lombard	0	43	43	0.68

Rank	Corridor	Pedestrian Fatalities (people)	Pedestrian Injuries (people)	Total Pedestrian Fatalities and Injuries (people)	Corridor Length (miles)
46	5th St - from Chew to Rockland	0	42	42	0.85
47	Cobbs Creek Pkwy - from Market to Hoffman	0	42	42	1.77
48	17th St - from Hunting Park to Ridge	0	38	38	3.17
49	6th St - from Roosevelt to Spring Garden	1	33	34	4.32
50	Cottman Ave - from Castor to Montgomery County Line	1	37	38	2.04

FIGURE 13.

MAP OF TOP FIFTY PRIORITY PEDESTRIAN CORRIDORS IN PHILADELPHIA



Source: PennDOT Crash Tables, 2014-2018

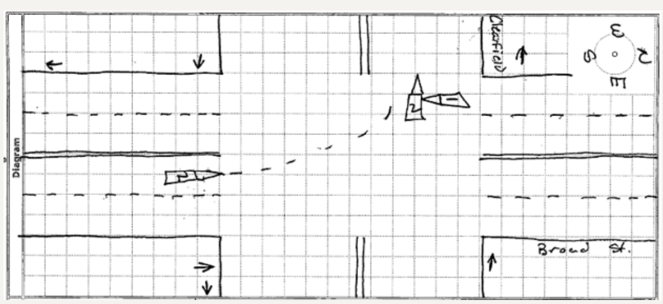
Appendix C: Methodology Overview

Understanding the Crash Data

This study uses crash data from the Pennsylvania Department of Transportation (PennDOT), which is digitized from crash reports prepared on-scene by the Philadelphia Police Department (PPD). Additional data (such as location of crosswalks, bike facilities, and bus stops) was provided by the City of Philadelphia and SEPTA.

Documenting crashes is a process with multiple steps. PennDOT maintains a state-wide crash database to comply with federal traffic safety reporting requirements. The process typically starts when a crash involving a motor vehicle is serious enough that the police are called to the scene. A PPD officer prepares a report documenting the crash and submits it to PennDOT within 15 days. PennDOT receives and processes each crash report, removing personal information and coding crash information in the PPD crash report. PennDOT does not include minor crashes or fender-benders in their database, they only include crashes that result in death, injury, and/or damage to a vehicle that is so significant the vehicle has to be towed from the scene of the crash. PennDOT's database also excludes crashes that occur on private property, crashes that are the result of deliberate intent, and crashes that are the result of a natural disaster.

FIGURE 14.
EXAMPLE OF A DIAGRAM FROM A PPD CRASH REPORT



Methodology

Data Review & Analysis

An analysis of PennDOT crash data over the last five years (2014-2018) highlights crash trends and patterns, detailed in Section III Findings. Those system-wide crash patterns and characteristics were then used to generate systemic solutions to address these trends, details in *Chapter 3: Systemic Solutions.*

TABLE 3.

CRASH DATASET SUMMARY TABLE

Description	Crash data that includes information about the people, vehicles, factors, and roadway involved in the crash
Source	PennDOT publicly available crash tables
Geography	City of Philadelphia
Years	2014-2018 (last five years available)
Distribution	• Total pedestrian injury crashes: 7340 • 151 pedestrian fatal crashes • 322 pedestrian serious suspected injury crashes • 6,394 pedestrian suspected minor injury, possible injury, and unknown injury crashes

FIGURE 15.

DATA USED IN THE ANALYSIS

CRASH DATA

Age

Sex

Pedestrian Location

Injury Severity

Aggressive Driving

Running Red Light

Running Stop Sign

Speeding

Tailgating

Cell Phone

Distracted

Vehicle Failure

Train/Trolley

Season

Weekday/Weekend

Time

Illumination

Wet Road

Bad Weather

Curved Road

Driveway

Type of Striking Vehicle

Hit & Run

Left Turn

Right Turn

STREET CHARACTERISTICS

Type of Street (Complete Street Typology)

Bicycle Facility

Speed Cushions

Mean Average Annual Daily Traffic (AADT)

Mean Average Annual Daily Pedestrians (AADP)

Intersection Density (intersections per corridor)

INTERSECTION CHARACTERISTICS

Crosswalk

Leading Pedestrian Interval (LPI)

Traffic Signal or Stop

Red Light Camera

Type of Intersection

ADDITIONAL FACTORS

Park

School

Subway station

Bus stop (all)

Trolley stop

Regional Rail station

Transit (bus, trolley, subway, and Regional Rail combined)

Systemic Crash Evaluations

The study analyzed crashes to identify circumstances in which pedestrian crashes were more likely to occur and determine whether they resulted in more serious injuries. For example, what is the relationship between speeding and pedestrian injury severity? Instances of statistically significant occurrences (*p* value of .05 and below) are starred with an asterisk (*p**).

The study developed three different geographies – area, corridor, and intersection – to examine pedestrian crashes citywide.

- **Area** identifies large clusters of pedestrian crashes by neighborhood geography.
- **Corridors**, defined as street segments that share similar features, help identify recurring issues along the length of a street that may be related to factors such as land use, transit lines, bicycle lanes, and other roadway characteristics that span many blocks. Corridors also contain information about pedestrian crashes that occur in the midblock, outside of intersections.
- **Intersections** represent the smallest geographic scale, where conflicts occur when pedestrians and drivers' cross paths in the street.

Accounting for Activity Analysis

To help see where pedestrians are the most at risk of being in a potential crash, a calculation was created to compare injury and fatality crashes, land area, and employment and residential activity. This calculation results in a risk value for locations where pedestrians are more likely to be in a crash that results in injury or death. Analysis of where pedestrians or more and less at risk of being injured helps to prioritize locations for pedestrian safety improvements.

Error! Reference source not found. (on page **Error! Bookmark not defined.**) depicts the employment and residential activity by census block group. Controlling all injury crashes by size and activity identifies areas in which high rates of crashes are occurring relative to the size of the analysis area and the composite employment and residential activity index score and thus, exposure, to potential crash incidents. This statistic is calculated with the following equation:

(n / area / activity index) X 1,000,000 = risk of all injury

Where n is all injury crashes, area is the total land area calculated in square feet and the activity index is the sum of the census block's employment activity index (interval scale variable with 5 distinct values) and the census block's residential activity index (interval scale variable with 5 distinct values). The highest activity index value for employment and residential activity is represented by a score of 5 and the lowest by a score of 1. The base equation is multiplied by 1,000,000 to eliminate small values with several decimal digits for statistical comparison purposes.

Scores were then categorized into five risk categories for the legend, with the lowest scores represented by the "Lowest Risk" category and the highest scores represented by the "Highest Risk" category.

Similar analyses were also conducted to calculate pedestrian major injury crash rates by size and activity and pedestrian fatality crash rates by size and activity. These statistics were calculated with the following equations:

(n / area / activity index) X 1,000,000 = risk of pedestrian injury

(n / area / activity index) X 1,000,000 = risk of pedestrian fatality

Best Practices

A review of best practices guided the analysis and prioritization process. The study reviewed eleven national, state, and city guides, plans, and studies on pedestrian crashes:

- FHWA: How to Develop a Pedestrian Safety Action Plan
- Smart Growth America: Dangerous by Design
- New York State DOT: Pedestrian Safety Action Plan
- Virginia DOT: Pedestrian Safety Action Plan
- DVRPC: Crash Analysis Standards & Recommendations
- New York City DOT: Pedestrian Safety Action Plan (Bronx)
- City of Minneapolis: Pedestrian Crash Study
- Chicago DOT: Pedestrian Crash Analysis
- Alamo Area MPO: San Antonio-Bexar County Pedestrian Safety Action Plan
- City of Seattle: Bicycle & Pedestrian Safety Analysis

The study evaluated each plan on a number of factors, including the scale of geographies, variables analyzed, prioritization methodology, and how pedestrian exposure was addressed.

The best practice review informed and guided:

- The selection of three different geographic scales of analysis: area, corridor, and intersection.
- The selection of key variables to include in the analysis (such as injury severity, signal or stop signs at intersections, time of day, age, crosswalks, signal characteristics, transit, and type of roadway).
- The evaluation of population density and employment density in crash clustering patterns. Population density and/or employment density were recommended by the FHWA, Chicago DOT, VDOT, and the Alamo Area MPO as variables to control for the level of activity happening around crash locations.
- Inclusion of equity indicators in the prioritization methodology, as recommended by the Delaware Valley Regional Planning Commission (DVRPC) and the Alamo Area MPO.

What Wasn't Included

Some crashes were excluded from the analysis. Reasons for exclusion include:

- **Non-reportable crashes** are crashes that do not result in death, injury, and/or damage to a vehicle that is so severe it must be towed from the scene. These crashes are not included in PennDOT's data collection. Crashes on private property, such as parking lots of college campuses, are also considered non-reportable and are not included in PennDOT's data collection. Finally, crashes that are the result of deliberate intent or a natural disaster are also non-reportable. The number of non-reportable crashes is unknown.
- **Crashes on highways** (or other rights-of-way where pedestrians are prohibited) accounted for 14.6% of all injury crashes

between 2014 and 2018. These crashes were not included because this analysis is focused on pedestrians, and on those roadways where the City leads maintenance and policy.

- **Crashes that did not result in an injury** of any kind, also known as property-damage only crashes. These crashes were excluded because this analysis is focused on negative physical outcomes for pedestrians. These types of crashes accounted for 17% of all injury crashes between 2014 and 2018.
- **Crashes that were the result of "not normal" driver behavior**, such as driving while under the influence or having a medical emergency while driving. These crashes were excluded because this analysis is focused on identifying problems that can be solved with design or engineering. These types of crashes accounted for 4.8% of all injury crashes between 2014 and 2018.

Excluded crashes total 36.4% of all injury crashes in the City of Philadelphia between 2014 and 2018.

Data Gaps

PennDOT's crash reporting system details each qualifying crash. However, there are several limitations to be aware of:

- **The location of crashes is usually but not always accurate**. The specific location of crashes is an approximation, depending on how the reporting PPD officer diagrammed it at the scene and how PennDOT geocodes the report for inclusion in their database.
- The potential for **unreported crashes** is an important data gap to consider. In some instances, individuals avoid involvement with law enforcement and/or reject medical treatment, leading to incomplete information and underreporting. Further, involving police in the aftermath of a crash is often related to insurance claims and repairing damage to the vehicle involved in the crash. Pedestrian crashes are less likely to involve damage to vehicles, and as the study shows, drivers are much more likely to flee from crashes involving pedestrians. Some cities who have conducted pedestrian crash studies have access to hospital data related to people involved in crashes. Hospital data related to crashes was not available in Philadelphia at the time of this study.
- Safety research tries to distinguish between high numbers of crashes due to specific factors and high numbers of crashes simply because more crashes are occurring in an area due to high volumes of people and vehicles. Where there are high volumes of pedestrians, there are more pedestrian crashes. This problem is often referred to as "exposure." The greater the number of pedestrians is present in an area, the higher the likelihood of conflict with motor vehicles. **Calculating exposure remains a challenge for all pedestrian safety studies.** Due to the lack of volume data at a city-wide level, this analysis focusing on controlling for exposure using residential and employment density.

Future Data Collection and Analysis

During the process of cleaning and analyzing the PennDOT and additional geospatial data, possible datasets were identified that would be helpful to collect and include in future analyses:

- **Number of lanes:** the number of lanes provides information about the capacity of a corridor, in addition to its functional classification or its Complete Streets typology. The number of lanes affects sightlines and the crossing distance for pedestrians. A significant relationship between number of lanes and pedestrian safety could help inform potential engineering solutions, such as road diets.

- **Roadway width:** roadway width provides information about the capacity of a corridor, in addition to its functional classification or its Complete Streets typology. Roadway width affects crossing distance for pedestrians. A significant relationship between roadway width and pedestrian safety could help inform potential engineering solutions, such as road diets.
- **Date of crosswalk installation:** crosswalks and other safety infrastructure is often installed in response to crash incidents. In order to differentiate between crashes that occurred before a crosswalk was installed and those that occurred afterwards is crucial to evaluating the effect of crosswalks.
- **Hospital data:** in the future, the availability and quality of data from hospitals about pedestrians involved in crashes should be explored in Philadelphia to see what it can add to future analyses.

For future crash analysis, studying all crashes is recommended whether injury or not to provide a control group to compare results against. Crashes that do not have any reported injuries have comparable driver and pedestrian behaviors. Because pedestrian crash numbers are relatively low compared to vehicle crashes, expanding to analyze all injury crashes can increase sample size.

Additionally, crashes could be flagged at schools and parks to study the potential effect on midblock crashes. Because this study focused on analyzing a wide range of variables, future analyses could focus on crashes occur near schools. Some relevant research questions could include: how do pedestrian safety outcomes vary by school age? Or how do pedestrian safety outcomes vary by infrastructure provided within the mile around the school? Similarly, future studies could examine more closely the relationship between crashes and parks. Some relevant research questions could include: how do different safety outcomes relate to parks of different types?

To better understand the particularities of pedestrian crashes, additional fields could be added to the police report to include pedestrian movement (to give a better sense of the pedestrian’s direction of travel/intent) and racial information for those involved in crashes to track crash data for equity considerations. Renewed effort should be made to collect speed related data fields for pedestrian crashes. In hit and run crashes where data is limited (a much higher proportion of pedestrian crashes), witnesses, video cameras, pedestrian testimonial, and hospital follow ups are crucial to getting a clearer picture and these sources should be routinely used to investigate pedestrian crashes.

Another element to consider for future studies is a detailed analysis of the relationship between land use and pedestrian crashes. This study used the Complete Streets Typology as a proxy for land use. The Complete Streets Typology combines and generalizes roadway function and land use. There are opportunities in future studies to more closely study land use and pedestrian safety using land use data.

Appendix D: Cut Sheet References

CUT SHEET REFERENCES

Bahar, G., M. Parkhill, E. Hauer, F. Council, B. Persaud, & C. Zegeer. (2007). Parts I and II of a Highway Safety Manual: Knowledge Base for Part II. Unpublished material from *NCHRP Project 17-27*.

Bahar, G., Masliah, M., Wolff, R., & Park, P. (2008). *Desktop Reference for Crash Reduction Factors*. https://safety.fhwa.dot.gov/tools/crf/resources/fhwasa08011/

Chen, L., Chen, C., & Ewing, R. (2012). The Relative Effectiveness of Pedestrian Safety Countermeasures at Urban Intersections - Lessons from a New York City Experience. Presented at the *91st Annual Meeting of the Transportation Research Board*.

De Brabander, B., & Vereeck, L. (2006). Safety Effects of Roundabouts in Flanders: Signal type, speed limits and vulnerable road users. Retrieved from: https://pubmed.ncbi.nlm.nih.gov/17118323/

Elvik, R., Christensen, P., & Amundsen, A. (2004). Speed and Road Accidents An Evaluation of the Power Model. Retrieved from: https://www.toi.no/getfile.php/131013/Publikasjoner/T%C3%98I%20rapporter/2004/740-2004/Repsumm.pdf

Federal Highway Administration. (2017) Corridor Access Management. *Proven Safety Countermeasures*. Retrieved from: https://safety.fhwa.dot.gov/provencountermeasures/corridor_access_mgmt/

Fayish, A.C. & Gross, F. (2010). Safety Effectiveness of Leading Pedestrian Intervals Evaluated by a Before–After Study with Comparison Groups. *Transportation Research Record; Journal of the Transportation Research Board* 2198: 15-22.

Furth, P. G., Halawani, A. T. M., Li, J., Hu, W. (Jake), & Cesme, B. (2018). "Using Traffic Signal Control to Limit Speeding Opportunities on Bidirectional Urban Arterials." *Transportation Research Record, Journal of the Transportation Research Board* 2672(18), 107–116.

Gan, A., Shen, J., & Rodriguez, A. (2005). *Update of Florida Crash Reduction Factors and Countermeasures to Improve the Development of District Safety Improvement Projects*.

Gayah, V., Donnell, E., & Jovanis, P. (2014). *Pennsylvania CMF Guide*. Retrieved from: https://www.penndot.gov/TravelInPA/Safety/Documents/PA-HSM-Tools_and_Data/Pennsylvania%20CMF%20Guide%20(2014).pdf

Gayah, V., Donnell, E., Zhengyao, Y., & Li, L. (2018). Safety and operational impacts of setting speed limits below engineering recommendations. *Accident Analysis and Prevention*, Vol. 121, pp.43-52.

Harkey, D.L., Srinivasan, R., Baek, J., Council, J., Eccles, K., & Lefler, N. (2008). Accident Modification Factors for Traffic Engineering and ITS Improvements. *NCHRP Report 617.*

ITE. (2004). *Toolbox of Countermeasures and Their Potential Effectiveness to Make Intersections Safer.* Briefing Sheet 8.

Johnson, R. S. (2005). Pedestrian Safety Impacts of Curb Extensions: A Case Study. *FHWA Report No. 0R-DF-06-01.*

Li, H., Graham, D.J., & Majumdar, A. (2013). The impacts of speed cameras on road accidents: An application of propensity score matching methods. *Accident Analysis and Prevention*, Vol.60, pp.148-157

McMahon, P., Zegeer, C., Duncan, C., Knoblauch, R., Stewart, R., & Khattak, A. (2002). An Analysis of Factors Contributing to 'Walking Along Roadway' Crashes: Research Study and Guidelines for Sidewalks and Walkways. *FHWA Report No. RD-01-101*.

Mead, J., Zeeger, C., & Bushell, M. (2014). "Evaluation of Pedestrian-Related Roadway Measures: A Summary of Available Research." Retrieved from: http://www.pedbikeinfo.org/cms/downloads/PedestrianLitReview_April2014.pdf#page=104&zoom=100,69,338

Michigan Department of Transportation, FHWA, & Western Michigan University. (2016). *User Guide for R1-6 Gateway Treatment for Pedestrian Crossings*. Retrieved from: http://aii.transportation.org/documents/User%20Guide_2018_0503_Final_UPDATED%20CDM%20Edgeline%20Clarification.pdf

New York City Department of Transportation. (2020). *Left Turn Traffic Calming*. Retrieved from: https://www1.nyc.gov/html/dot/html/pedestrians/left-turn-traffic-calming.shtml

---. (2013). *Making Safer Streets*. Retrieved from: https://www1.nyc.gov/html/dot/downloads/pdf/dot-making-safer-streets.pdf

Pawlovich, M.D., Li, W., Carriquiry, A. & Welch, T. (2006). Iowa's Experience with Road Diet Measures—Use of Bayesian Approach to Assess Impacts on Crash Frequencies and Crash Rates. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1953.

Pennsylvania Department of Transportation. (n.d.) *Flashing Yellow Arrow Fact Sheet*. Retrieved from: https://www.penndot.gov/TravelInPA/TrafficSignalsManagement/Documents/Flashing%20Yellow%20Arrow%20Fact%20Sheet.pdf

Pennsylvania State Transportation Advisory Committee. (2011). *Evaluating the Automatic Red-Light Enforcement Program, Final Report*. Retrieved from: https://www.talkpatransportation.com/assets/TAC/Evaluating%20the%20Automated%20Red%20Light%20Enforcement%20Program%20(ARLE)%20-%20October%202011%20-%20Final%20Report.pdf

Persaud, B., Council, F. M., Lyon, C., Eccles, K. & Griffith, M. (2005). Multi-jurisdictional safety evaluation of red-light cameras. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1922, pp. 29–37.

Schattler, K. & T. Hanson. (2016). Safety Impacts of a Modified Right Turn Lane Design at Intersections. Presented at the *95th Annual Meeting of the Transportation Research Board*, Paper No. 16-0790.

Schepers, J.P., Kroeze, P.A., Sweers, W. & Wust, J.C. (2011). Road Factors and Bicycle-Motor Vehicle

Crashes at Unsignalized Priority Intersections. *Accident Analysis and Prevention*, 43(3): pp. 853-861.

Thomas, L., Thirsk, N.J. & Zegeer, C. (2016). *NCHRP Synthesis 498: Application of Pedestrian Crossing Treatments for Streets and Highways*. Retrieved from: http://www.trb.org/Publications/Blurbs/175419.aspx.

Thomas, L., Srinivasan, R., Worth, M., Parker, M.R. & Miller, R. (2015). Jurisdiction Speed Management Action Plan Development Package. *FHWA Report No. FHWA-SA-15-017.*

Van Houten, R. &. Hochmuth, J. (2017). Evaluation of R1-6 Gateway Treatment Alternatives for Pedestrian Crossings: Follow-Up Report. *CTS 17-05*.

Zeeger, C., Cynecki, M., & McGee, H. (1986). Methods Of Increasing Pedestrian Safety At Right-Turn-On-Red Intersections User's Manual. *FHWA Report No. IP-86-010*.

Zegeer, C., Nabors, D., & Lagerwey, P. (2013). Gateways. *Pedestrian Safety Guide and Countermeasure Selection System.* Retrieved from: http://www.pedbikesafe.org/pedsafe/countermeasures_detail.cfm?CM_NUM=37

Zeeger et al. (2017). Development of Crash Modification Factors for Uncontrolled Pedestrian Crossing Treatments. *NCHRP Report 841*.

Appendix E: Additional Findings

Factors that were tested and did not have a strong relationship to pedestrian injuries or fatalities included:

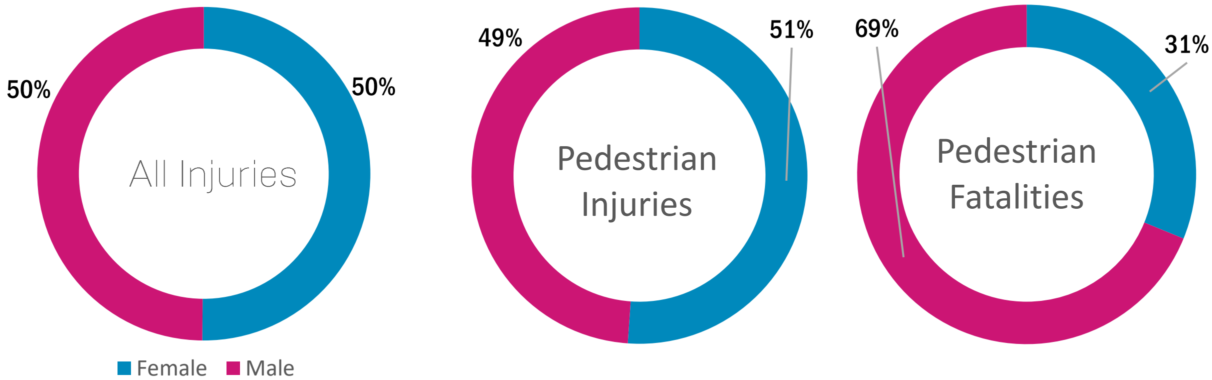
- Sex
- Aggressive Driving
- Running Red Light
- Running Stop Sign
- Tailgating
- Cell Phone
- Distracted
- Vehicle Failure
- Train/Trolley
- Wet Road
- Bad Weather
- Curved Road
- Driveway
- Intersection Density
- Leading Pedestrian Interval (LPI)
- Park
- Regional Rail station

A chart for each of these crash factors and why it was not further studied is included below.

FIGURE 16.

PERCENT OF PEOPLE INJURED BY SEX, 2014-2018

Male pedestrians are a higher share of those killed while walking. Without pedestrian volume numbers it cannot be determined whether men are overrepresented or not. These numbers are consistent with other cities.

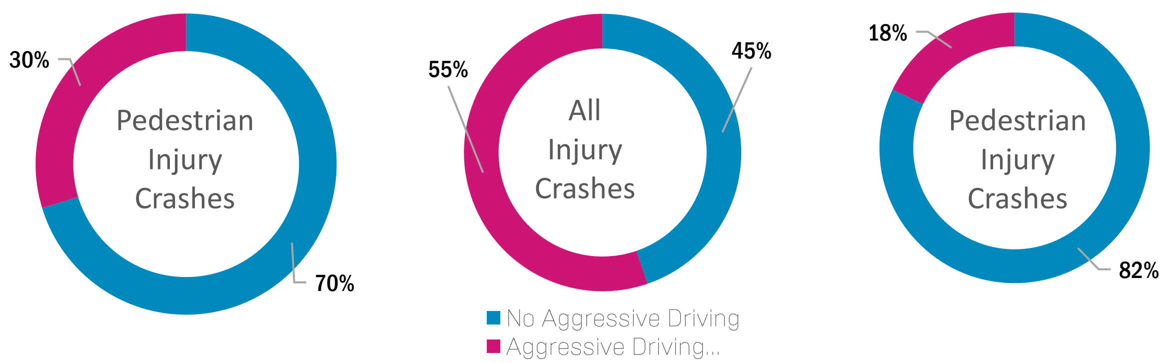


Source: PennDOT Crash Tables, 2014-2018

FIGURE 17.

PERCENT OF CRASHES INVOLVING AGGRESSIVE DRIVING, 2014-2018

Highest share of aggressive driving in all injury crashes. Field completion rate low.

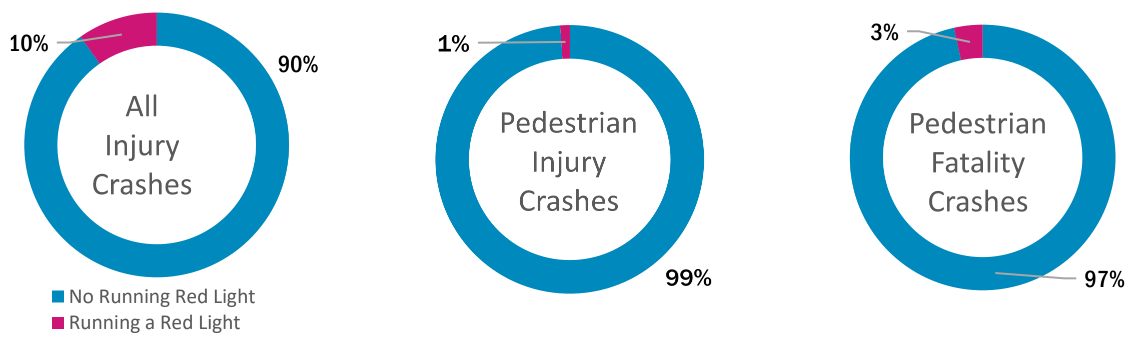


Source: PennDOT Crash Tables, 2014-2018

FIGURE 18.

PERCENT OF CRASHES INVOLVING RUNNING A RED LIGHT, 2014-2018

Highest share of running red light in all injury crashes. Field completion rate low.

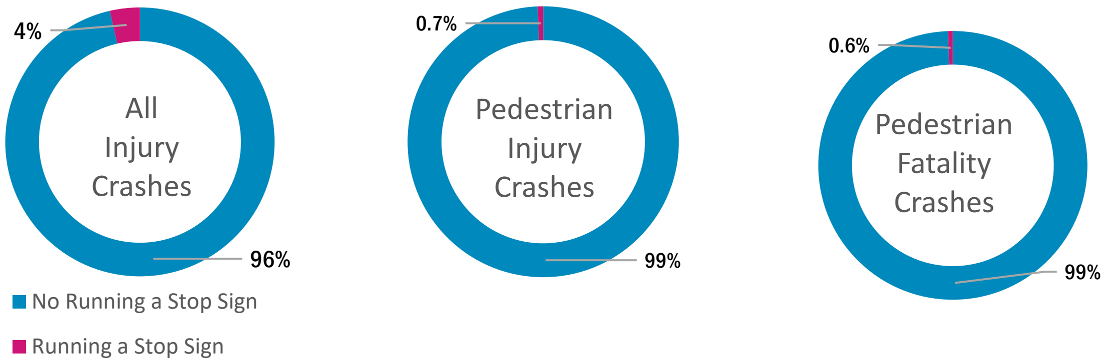


Source: PennDOT Crash Tables, 2014-2018

FIGURE 19.

PERCENT OF CRASHES INVOLVING RUNNING A STOP SIGN, 2014-2018

Highest share of running stop sign in all injury crashes. Field completion rate low.

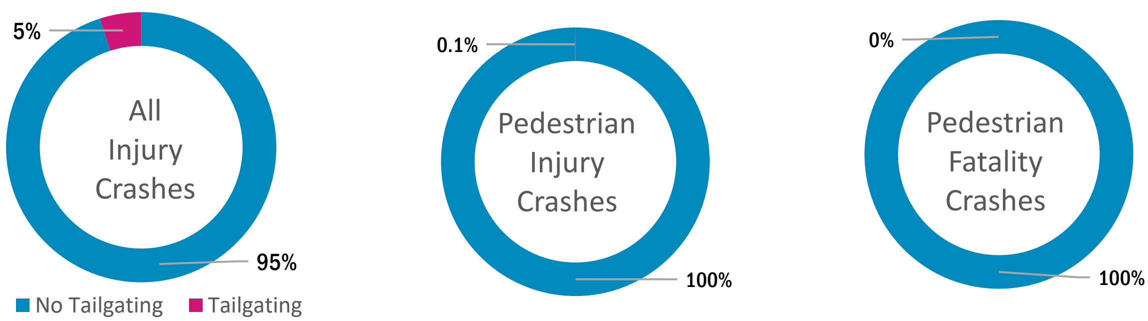


Source: PennDOT Crash Tables, 2014-2018

FIGURE 20.

PERCENT OF CRASHES INVOLVING TAILGATING, 2014-2018

Highest share of tailgating in all injury crashes. Field completion rate low.

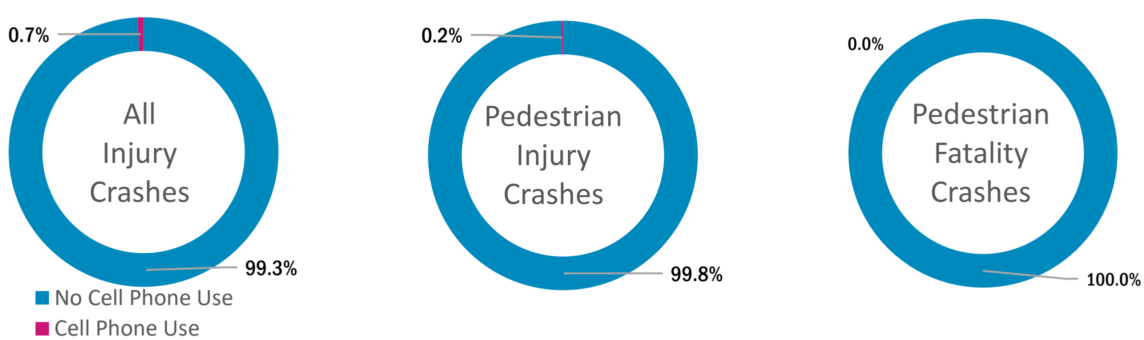


Source: PennDOT Crash Tables, 2014-2018

FIGURE 21.

PERCENT OF CRASHES INVOLVING CELL PHONE USE, 2014-2018

Highest share of cell phone use in all injury crashes. Field completion rate low.

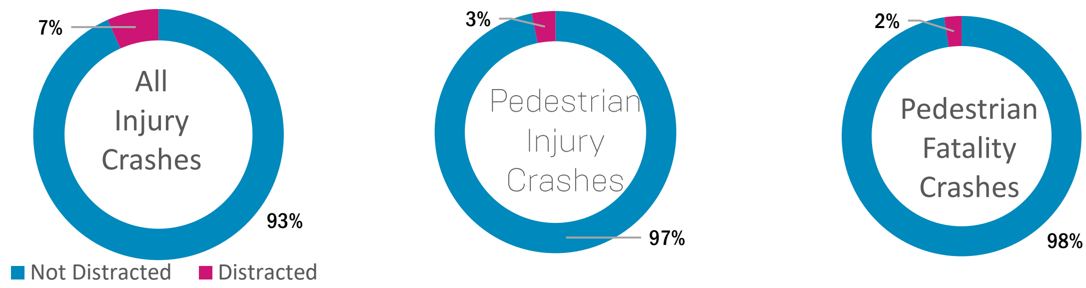


Source: PennDOT Crash Tables, 2014-2018

FIGURE 22.

PERCENT OF CRASHES INVOLVING DISTRACTED DRIVING, 2014-2018

Highest share of distracted driving in all injury crashes. Field completion rate low.

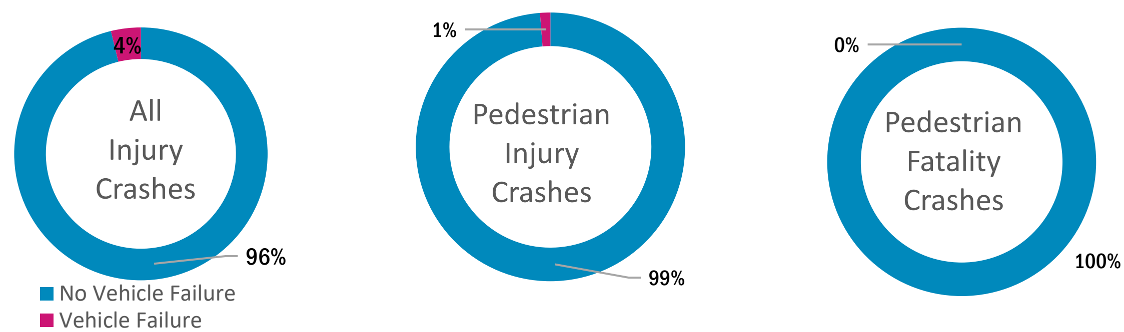


Source: PennDOT Crash Tables, 2014-2018

FIGURE 23.

PERCENT OF CRASHES INVOLVING VEHICLE FAILURE, 2014-2018

Highest share of vehicle failure in all injury crashes. Field completion rate low.

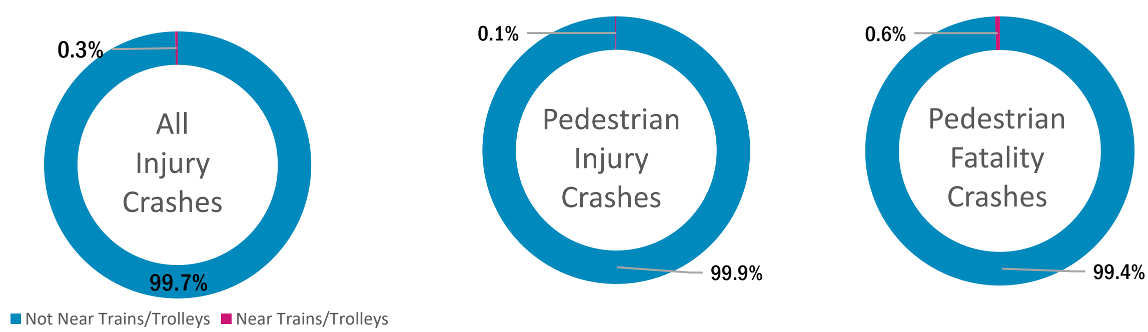


Source: PennDOT Crash Tables, 2014-2018

FIGURE 24.

PERCENT OF CRASHES INVOLVING A TRAIN OR TROLLEY VEHICLE, 2014-2018

Similar share across crash groups.

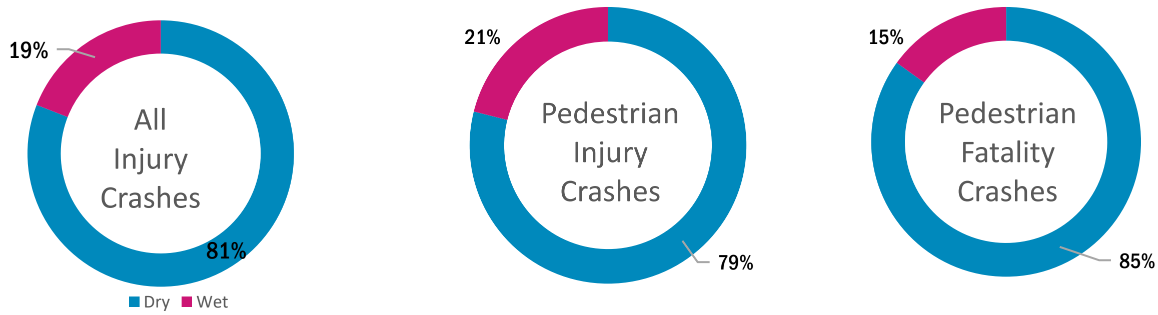


Source: PennDOT Crash Tables, 2014-2018

FIGURE 25.

PERCENT OF CRASHES INVOLVING WET ROADWAY, 2014-2018

Similar share across crash groups.

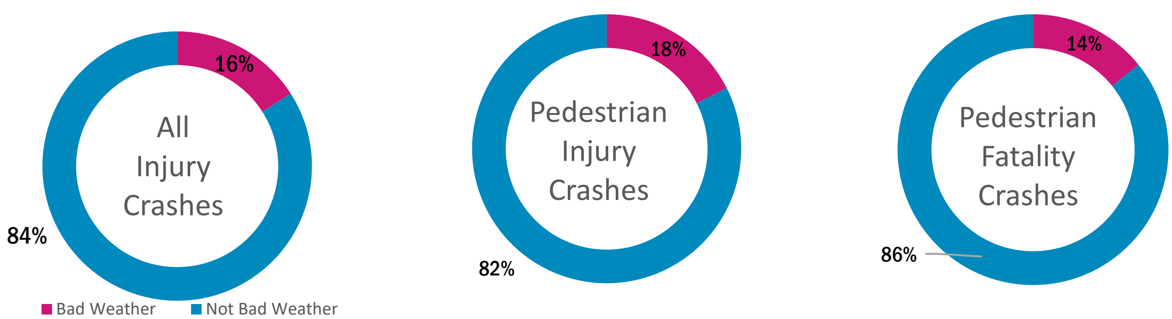


Source: PennDOT Crash Tables, 2014-2018

FIGURE 26.

PERCENT OF CRASHES INVOLVING BAD WEATHER, 2014-2018

Similar share across crash groups.

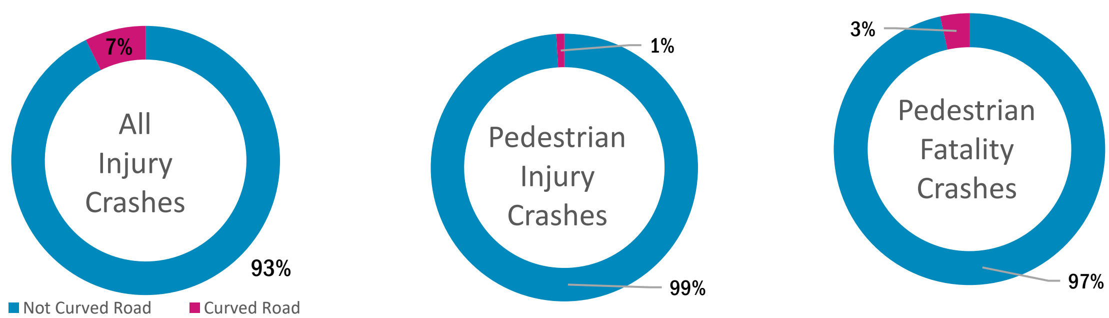


Source: PennDOT Crash Tables, 2014-2018

FIGURE 27.

PERCENT OF CRASHES INVOLVING A CURVED ROADWAY, 2014-2018

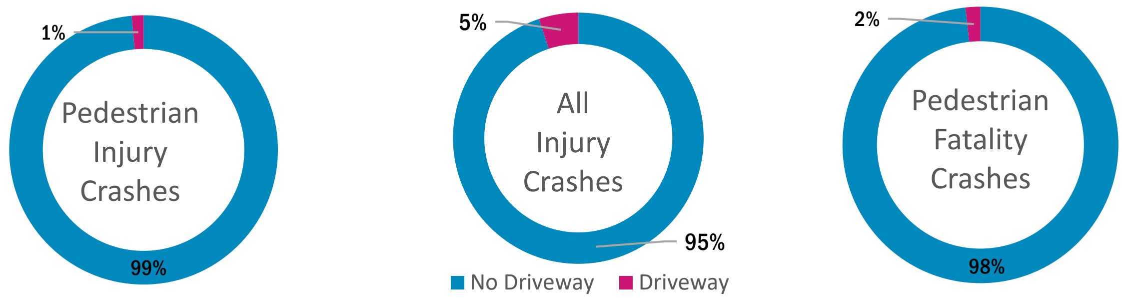
Highest share of crashes on a curved roadway in all injury crashes, numbers are very low.



Source: PennDOT Crash Tables, 2014-2018

FIGURE 28.
PERCENT OF CRASHES INVOLVING A DRIVEWAY, 2014-2018

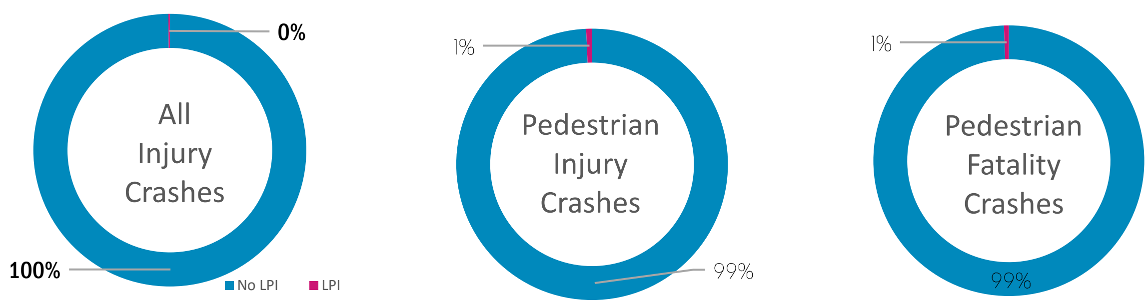
Highest share of vehicle failure in all injury crashes, numbers are very low.



Source: PennDOT Crash Tables, 2014-2018

FIGURE 29.
PERCENT OF CRASHES AT INTERSECTIONS WITH LEADING PEDESTRIAN INTERVALS (LPIS), 2014-2018

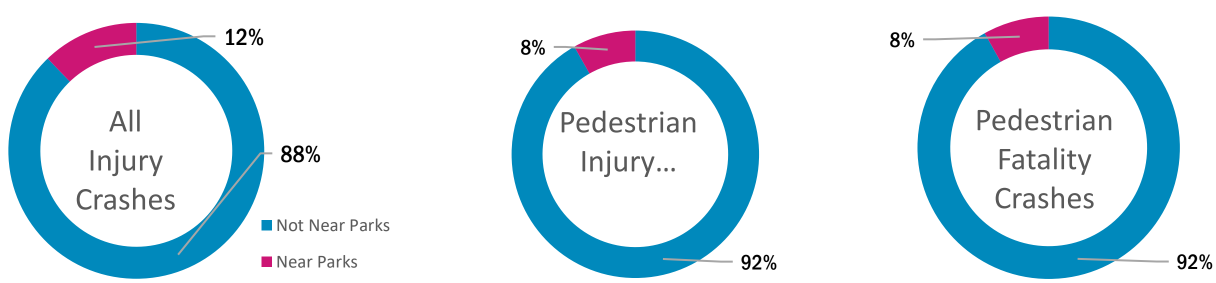
Similar share across crash groups, numbers are very low.



Source: PennDOT Crash Tables, 2014-2018; City of Philadelphia Leading Pedestrian Interval shapefile

FIGURE 30.
PERCENT OF CRASHES NEAR* A PARK, 2014-2018

Highest share of crashes occurring near parks for all injury crashes, numbers are low.

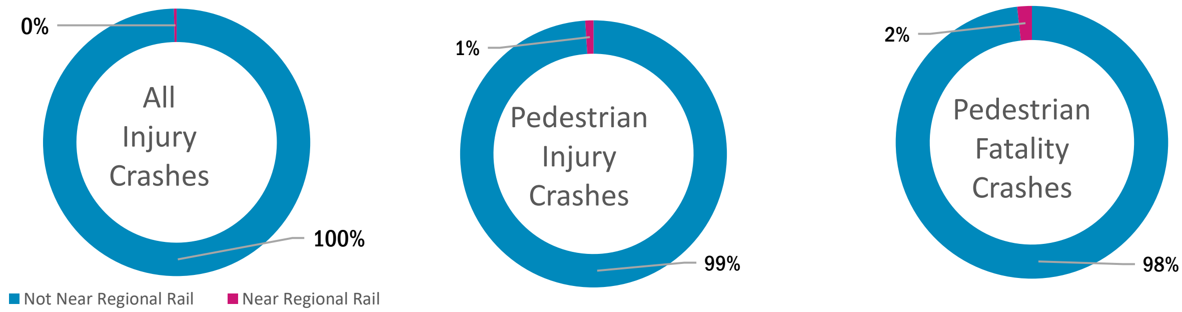


* Near a park is defined as an intersection within 100’ of a park

Source: PennDOT Crash Tables, 2014-2018; City of Philadelphia Parks & Recreation Assets shapefile

FIGURE 31.
PERCENT OF CRASHES NEAR* A REGIONAL RAIL STATION, 2014-2018

Similar share across crash groups, numbers are very low.



* Near a regional rail station is defined as an intersection within 300' of a regional rail station

Source: PennDOT Crash Tables, 2014-2018; SEPTA ArcGIS Data Portal

Appendix F – Technical Appendix

Introduction

The purpose of this appendix is to explain the assumptions and methods used to analyze crashes for the Vision Zero Pedestrian Safety Study and Action Plan.

Audience Disclaimer

This appendix is written for a technically proficient transportation analyst who is:

- Familiar with the tenets of transportation theory and analysis;
- Familiar with statistics and statistical modeling (summary statistics, correlation testing, univariate and multivariate non-linear regression, logit models);
- An intermediate to advanced R user who can use basic R functions (such as create data tables and execute merges), learn functions from external packages, troubleshoot on Stack Overflow or other forums, and can use R to conduct correlation tests, and non-linear and logit regressions (bivariate and multivariate); and
- An intermediate ArcGIS user who can execute basic geospatial functions, such as Buffer, Intersect, Dissolve, Spatial Join, Select by Location, create shapefiles, adjust projections, use Field Calculator.

Process Flow

The overall analysis process consisted of nine steps:

1. Cleaning and merging of the 2009-2018 PennDOT crash tables in the state of Pennsylvania, resulting in a database of injury crashes (excluding property damage only crashes) in the City of Philadelphia (or within 100' of the county line) between 2014-2018. Crashes on interstates and highways that are below- or above-grade were excluded;
2. Cleaning available data to create two sets of geographies for analysis: intersections and corridors;
3. Cleaning available data to merge with the intersection and corridor geographies;
4. Geolocation of crash data and merging with the intersection and corridor geographies;

5. Merging geospatial crash data with PennDOT crash data;
6. Summary statistics at the intersection- and corridor-level in Excel, using Pivot Tables;
7. Correlation tests, univariate modeling, and multivariate non-linear and logit regression modeling at the intersection- and corridor-level in R;
8. Spatial analysis, including kernel density, Moran's I, and hot spot analysis, of all injury crashes, pedestrian injury crashes, and pedestrian fatality crashes; and
9. Identification of key crash findings from summary statistics and statistical tests.

Database Creation

Creating a Unified PennDOT Crash Database (R)

Purpose

The purpose of this step was to take the PennDOT crash data available online in eight tables and to transform them into a single database of crashes that fit the study's parameters: crash data from 2014-2018 at the most detailed level available (person-level).

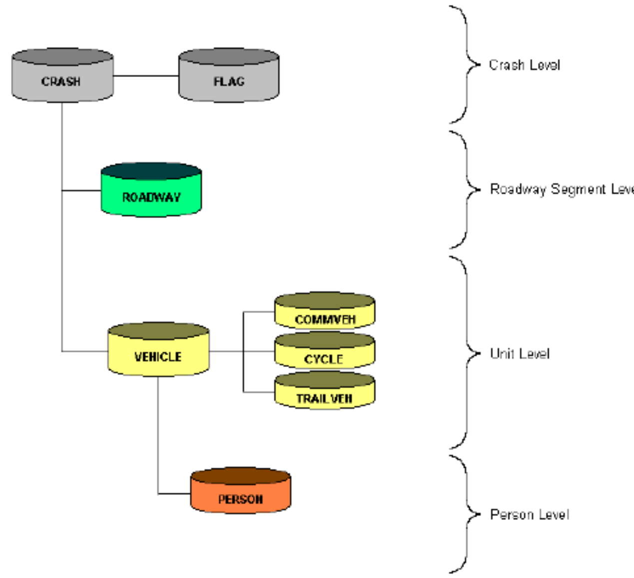
This step was conducted in R and includes code that creates longitudinal tables, unique identifiers, subsets variables, reshapes tables, merges tables, subsets tables, and other necessary transformations.

Creating Longitudinal Tables and Unique Identifiers

PennDOT maintains public crash data for the state of Pennsylvania eight tables, which are at four different units of analysis (see 0 on page 39). The project team downloaded five of the eight crash tables: crash, vehicle, person, flag, and roadway. These five were chosen based on the variables that the team used for analysis. These five crash tables were downloaded for the years 2009 through 2018. PennDOT periodically releases updated tables.

FIGURE 32.

DIAGRAM OF PENNDOT CRASH DATA TABLES BY UNIT LEVEL



Source: *PennDOT, http://pennshare.maps.arcgis.com/sharing/rest/content/items/cae1501f58d842c385ebfaa963098d61/data*

Each table, which was separated by year, was minimally cleaned and appended. Three unique identifiers were used throughout this step to achieve the merges:

1. The CRN is a ten-digit number that PennDOT assigns each crash event. The CRN generally starts with the year the crash happened and is used throughout all eight tables. The CRN is PennDOT's unique identifier within all eight tables.
2. The unit-level identifier, which the project team created during this step and is present in the vehicle- and person-level tables. It is called MERGEID_VEH and is the CRN + "_" + the unit number.
3. The person-level identifier, which the project team created during this step and is

present in the person-level table. It is called MERGEID and is the CRN + "_" + the unit number + "_" + person number.

Table 4. (on page 40) shows the identifiers created for each of the tables.

TABLE 4.

UNIQUE IDENTIFIERS FOR EACH PENNDOT CRASH TABLE

PennDOT Table	CRN	MERGE_VEH	MERGEID
Crash	✔		
Person	✔	✔	✔
Vehicle	✔	✔	
Roadway	✔		
Flag	✔		

Reshaping the Roadway Table

The roadway table contains multiple records for each crash (CRN). Each record is a roadway involved in the crash. Each CRN had from 1-10 roadway records associated with it. The roadway table needed to be consolidated to the crash level to merge with the other tables. The dcast function from the reshape2 package was used to re-shape the roadway table from the roadway level to the crash level.

Merging Longitudinal Tables

To create a person-level database including the variables from the five longitudinal crash tables, the project team executed a merge aiming to retain as many CRNs as possible. The team created a table of all CRNs contained in the five crash tables. Not all CRNs listed in the five crash tables have matching records in the other four tables. The team analyzed incomplete records except for CRNs that did not have crash, person, or vehicle records since these three tables contain essential injury severity and mode-specific data. The team created a binary flag for CRNs that did not contain crash, person, and vehicle records (1,449 out of 1,252,772 CRNs statewide). Less than 1% (0.12%) of CRNs were excluded from the merge due to missing crash, person, and vehicle records.

Using the master list of complete CRNs, all incomplete CRNs were pruned from the five crash tables. Three sets of merges followed:

1. **Merge 1 –** crash-level tables: The crash, flag, and roadway tables were merged by CRN using a full outer join.
2. **Merge 2 –** vehicle- and person-level tables: The person and vehicle tables were merged by the vehicle-level identifier (MERGEID_VEH) using a right outer join to achieve a person-level table.
3. **Merge 3 –** crash-level tables and person-level tables: The merged crash-level tables (crash, flag, and roadway) were merged with the merged person-level tables (vehicle and person) by CRN using a right outer join to achieve a person-level table.

Throughout these three sets of merges, 1,251,323 CRNs were maintained. There were 1,252,772 unique CRNs in all five tables, but only 1,251,323 were complete (contained associated crash, person, and vehicle information).

Subset of Combined Longitudinal Tables

The crash database was then winnowed down based on the study's parameters to crashes that occurred between 2014-2018, did not occur on an interstate, are in the City of Philadelphia (or within 100' of the county line), and include at least one injury. Of the 1,251,323 unique crashes in the all crashes database, 3% crashes remained in the study's database.

Misalignment Between Data Dictionary and Data

There were several variables that did not match or did not have matching entries in the Data Dictionary. For example, the variable travel speed had entries such as 335 mph and 406 mph, which the project team decided were typos (the team converted these to 35 and 40 mph, respectively). Travel speed also had several "junk" entries that were not specified in the Data Dictionary, such as 99, 999, and 997, which are common codes indicating that the variable is unknown.

Creating Geographies and Geospatial Data

Purpose

To conduct an intersection- and corridor-level crash analysis, the project team created intersection and corridor geographies. After creating the geographies, the team prepared data from the City (e.g. crosswalk locations, bus stops) and attached it, when possible, to the intersection and corridor geographies. Once the geographies contained all possible data, they were merged with the PennDOT crash data so that each crash contained information, when available, about the intersection and corridor it was associated with. Once the crash data was merged with the geospatial data, it was loaded back into R and lightly cleaned before analysis.

Creating Intersection Geographies (ArcGIS)

Creating Intersection Polygons

To create the intersection geography, the project team created a polygon shapefile that roughly encompassed intersections city-wide. The project team started by creating intersection points, which were then used to create intersection polygons as buffers.

Using Network Analyst and the Streets Department Street Centerline shapefile, the project team created roughly 40,000 points at the intersection of every line segment in the Centerline shapefile. Simultaneously, the project team edited the Street Centerline shapefile in the following ways:

- Converted divided highways (e.g. Spring Garden, Roosevelt Boulevard) from double centerlines to one centerline;
- Excluded the following line segments that were of the following classes:
  - Expressways
  - Driveway
  - Low speed ramps
  - High speed ramps
  - Non-travelable
  - City Boundary
  - Walking Connector

Due to the irregularities of the Street Centerline file and the nodes generated from it, there were often multiple nodes in the same intersection. The goal was the have one point per intersection that would later be buffered out. Therefore, the project team conducted a one-to-one spatial join of the intersection points and the edited centerlines to understand how many line segments intersected with the node. For example, if an intersection was four-way, 90-degree intersection, the point would be intersected by four lines. However, if a point was intersected by more than four lines, it was likely a multi-leg intersection. The project team focused on points where there were less than four intersecting lines:

- Points with three intersecting lines: mainly T-intersections
- Points with two intersecting lines: mainly L-intersections
- Points with one intersecting line: dead ends
- Points with no intersecting lines: intersection points that were created before the exclusion of street centerline segments, such as expressways

The project team excluded points with one intersecting line and no intersecting lines, which removed less than 1% of the intersection points.

The project team then created four subsets of the edited street centerline file by the four remaining functional classes: major arterials, minor arterials, collectors, and locals. Then the team executed four one-to-one Spatial Joins between the centerline subsets and the intersection points, which created function class-specific counts of line segments that intersected with each intersection point. The project team used the composition of line segments by intersection point to classify each intersection point into twelve preliminary intersection typologies (Table 5.).

TABLE 5.

PRELIMINARY INTERSECTION TYPOLOGIES BASEDON JOINED CENTERLINE CLASS

Preliminary Typology Name	Composition	*n*	%	Notes
A	75% major arterial	578	2%	At or above 75% major arterials
B	75% minor arterial	438	2%	At or above 75% minor arterials
C	75% collectors	3,004	13%	At or above 75% collectors
D	75% locals	4,356	19%	At or above 75% locals
E	Local/collector mix	7,782	34%	Any mix of local/collector that falls below the threshold of 75% for both – excludes any mix that includes major or minor arterials
F	Minor arterial/local mix	2,034	9%	Any mix of minor arterials/locals that falls below the threshold of 75% for both – excludes any mix that includes major arterials and collectors

G	Major arterial/collector mix	1,139	5%	Any mix of major arterials and collectors that falls below the 75% threshold for both – excludes minor arterials and locals
H	Major arterial/minor arterial mix	443	2%	Any mix of major/minor arterials that falls below the 75% threshold for both – excludes collectors and locals

The project team then compared the preliminary intersection typologies and combined them into representative groups (Table 6. on page 43). Roosevelt Boulevard was given its own intersection typology, as explained below.

TABLE 6.

INTERSECTION TYPOLOGIES

Typology Name	Preliminary Typology Name	*n*	%	Notes
Majors	A	579	3%	At or above 75% major arterials
Major Inclined	G, H, I, J	3,451	15%	Major arterial inclined
Minors	B, K, L	1,695	7%	Minor arterials and minor arterial inclined
Minor-Local	F	2,034	9%	Minor arterial and local
Collectors	C	3,004	13%	At or above 75% collectors
Collector-Local	E	7,782	34%	Local/collector mix
Alleys	D	4,356	19%	Alley streets
Roosevelt	-	132	0.6%	Roosevelt

Using the intersection typologies, the project team assigned each buffer sizes appropriate to their size characteristics to create the intersection polygons (Table 7.).

TABLE 7.

INTERSECTION TYPOLOGY BUFFERS

Typology Name	Buffer Size	Notes
Majors	100’	At or above 75% major arterials
Major Inclined	75’	Major arterial inclined
Minors	50’	Minor arterials and minor arterial inclined
Minor-Local	50’	Minor arterial and local
Collectors	50’	At or above 75% collectors
Collector-Local	25’	Local/collector mix
Alleys	25’	Alley streets
Roosevelt	75’	Roosevelt

Roosevelt Boulevard

Roosevelt Boulevard’s unique roadway design made creating intersections that were comparable to the rest of the City challenging. The project team’s previous work on the Vision Zero Roosevelt Boulevard Crash Analysis was repurposed so that the intersection polygons, which had been drawn for all the intersections that crossed all four cross sections of the Boulevard, were used as intersection polygons for this analysis. The polygons were drawn based on the geometry, lane markings, and high speeds of Roosevelt Boulevard.

Midblock Geographies

The project team wrapped up the intersection geographies by creating an approximation of the midblock geography, so that every crash that was not in an intersection was not automatically coded as a midblock crash. The team separated the edited street centerline file into its four functional classes and created a buffer around each:

- Local centerlines – 25’
- Collector centerlines – 50’
- Minor Arterial – 75’
- Major Arterial – 100’

The project team excluded any crashes that fell outside these buffers or the intersection polygons. Only 1,585 crashes (3% of all crashes, property damage only crashes included) were excluded from the intersection crash analysis dataset. Many of these crashes were geolocated to expressways, on pedestrian-only streets (i.e. Locust Walk), or in large strip mall parking lots. In short, these crashes were not able to be associated to the intersection-level and were therefore excluded.

Creating Corridor Geographies (ArcGIS)

For the corridor analysis the project team created a city-wide corridor geography based on the Complete Streets Standards developed by the Office of Transportation, Infrastructure and Sustainability.

Creating Corridor Polygons

To create the corridor geography, the project team dissolved street centerlines based on functional class and Complete Streets typology. Two types of centerlines were removed: centerlines less than 100 feet and centerlines that were less than 1,000 feet that had the following functional classifications: expressways, locals, collectors, and minor arterials. This resulted in 5,274 corridors. The project team then buffered the centerlines, with the following buffers:

- Expressways: 150 feet
- Major arterials: 75 feet
- All other functional classes: 50 feet

Preparing City Data for Geographies (ArcGIS)

Introduction to City Data

The City provided the project team with the following data, some of which was downloaded via Open Data or other public data portals:

- Leading Pedestrian Intervals, separated into Y1 (2017) and Y2 (2019)
- Bus stops and routes, downloaded from SEPTA's ArcGIS data portal
- High Speed lines and stations (MFL, BSL, and NHSL stations), downloaded from SEPTA's ArcGIS data portal
- Regional Rail Stations, downloaded from SEPTA's ArcGIS data portal
- Trolley routes, downloaded from SEPTA's ArcGIS data portal
- Crosswalks
- Speed Cushions
- Traffic Signals
- Stop Signs
- Red Light Cameras
- Bike Network
- DVRPC Pedestrian Counts
- DVRPC Traffic Counts
- Activity Index

Cleaning City Data

Each dataset required cleaning, normalization, and then a set of rules to relate it to each geography:

Leading Pedestrian Intervals

Used for the intersection geographies only. Y2 data was excluded due to incompatibility with crash data date ranges. Intersection polygons were flagged where they intersected with the LPI points in the Y1 shapefile. After the geospatial data was merged with the PennDOT crash data, flags were removed for crashes that occurred at LPI intersections before 2017.

Crosswalks

Used for intersection geographies only. The project team only used "MARKED" crosswalks (in the Status column) and flagged all intersection polygons within 20' of a crosswalk polyline.

Speed Cushions

Used for intersections only. The project team excluded all speed cushions that were not marked as "Permanent" and hand-selected intersections that were adjacent to those speed cushions. After the geospatial data was merged with the PennDOT crash data, flags were removed for crashes that occurred at speed cushion intersections before the year they were installed.

Traffic Signals

Used for intersections only. Flagged intersection polygons within 20’ of a traffic signal point.

Stop Signs

Used for intersections only. Flagged intersection polygons within 20’ of a stop sign point.

Red Light Cameras

Used for intersections only. Because red light camera points were generally positioned outside of intersection polygons, the project team hand-selected and flagged intersection polygons with red light cameras that had been in operation between 2014 and 2018.

Parks

Used for intersections only. Flagged intersection polygons that were within 100’ of parks (as defined by PPR’s publicly available asset shapefile).

Schools

Used for intersections only. Flagged intersection polygons that were within 500’ of schools (as defined by Philadelphia City Planning Commission’s publicly available school shapefile).

Bus Service

For the intersection geography, the project team used the bus stops shapefile. Half of the bus stop shapefile did not contain a stop ID and was excluded. While there was a ridership column, it was not filled out consistently. Therefore, the project team conducted a Spatial Join between the intersection polygons and bus stops, also flagging intersections that had more than four bus stops (an approximation of a high level of service).

For the corridor geography, the project team used the bus routes shapefile. The team ran an Intersect between the bus routes and the corridor polygons, calculated the length of the intersect between the bus route lines and the corridor polygons, and then divided the length of the intersecting bus route and the length of the corridor. The shapefiles were then joined by FID. For intersects that were greater than or equal to 25%, the team counted that as a route that had through-running service on that corridor.

Subway Service

For the intersection geography the project team flagged all intersection polygons within 300’ of a subway station.

For the corridor geography, the project team hand-selected and flagged corridors that coincided with subway service.

Trolley Service

For the corridor geography, the project team used the trolley routes shapefile. The team ran an Intersect between the trolley routes and the corridor polygons, calculated the length of the intersect between the trolley routes lines and the corridor polygons, and then divided the length of the intersection and the length of the corridor. The shapefiles were then joined by FID. For intersects that were greater than or equal to 50%, the team counted that trolley as having through-running service on that corridor.

Regional Rail Service

For the intersection geography, the project team flagged all intersection polygons within 300' of a regional rail station.

Bicycle Facilities

For the intersection geography, the project team used the bike network shapefile. The project team conducted a Spatial Join between the intersection polygons and the bike network, flagging intersections in which bicycle networks passed through the intersection polygon.

For the corridor geography, the project team had to dissolve the bicycle network shapefile by name and street code (allowing for multi-part features) and then fix divided highways manually (i.e. Oregon and Island Ave). Sharrows were excluded. The project team ran an Intersect between the bike network lines and the corridor polygons and then calculated the length of the intersect between the bike facility lines and the corridor polygons, and then divided the length of the intersecting bike facility and the length of the corridor. The shapefiles were then joined by FID. For intersects that were greater than or equal to 25%, the team counted that as a through-running bicycle facility on that corridor.

Note that install dates were not provided with bicycle facilities shapefile. Therefore, it is unclear whether crashes occurred when the bicycle facility was in place or not.

DVRPC Traffic Counts - AADT

Used for the corridor geography only. The project team selected traffic counts taken between 2014 and 2018, counts within corridor polygons, and only used volume or 15-minute volume counts. To normalize AADT across traffic counts, which was recorded both as bi-directional and uni-directional, the project team selected uni-directional counts (see the COUNTDIR column) and multiplied it by two to approximate bi-directional counts. The project team then dissolved the counts by road name, from limit, and to limit to get a mean of AADT at that location. From there, the team used a Spatial Join of the dissolved counts to the corridor polygons and took the mean of the AADT. Of the 5,274 corridor segments, 802 have AADT information (15%).

DVRPC Pedestrian Counts – AADP

Used for the corridor geography only. The project team removed any pedestrian counts that had comments (in the Comments column), which is used when a count may be inaccurate in representing an "average" day (i.e. a holiday, counter malfunction, unusual weather event). Eighty-two counts were excluded during this process (6% of all counts).

The project team then dissolved the counts by road name, from limit, and to limit to get a mean of AADP at that location. Then project team selected all counts within corridors. Of the 371 dissolved counts, 163 dissolved counts were excluded (44% of dissolved pedestrian counts). Most of the pedestrian counts that were excluded in this process were outside of the City of Philadelphia.

From there, the team spatial joined the dissolved counts to the corridor polygons (search radius of 25') and took the mean of the AADP. Of the 5,274 corridor segments, 191 had matching AADP data (4%).

Joining Crashes and Geographies (ArcGIS & R)

Crashes were imported from R and joined to the geographies using a Spatial Join with a 75' search radius. Two to three percent of crashes were not matched to corridors during the intersection and corridor joins due to crashes being outside of the geographies (i.e. in parking lots).

There were several instances of crashes being in two or more intersections or corridors. In these instances, the geography was randomly assigned. Sixteen percent of intersection crashes were assigned randomly. Fifty-one percent of corridor crashes were assigned randomly between two corridors, and twenty-five percent of corridor crashes were assigned between three or more corridors.

Seventy-three corridor crashes were excluded due to having the corridor typology of "1," which is unknown. None of the crashes were KSI crashes.

PennDOT Crash Data Limitations

There are a few notable issues with the PennDOT crash data:

- Geolocating of the crashes is generally accurate to the intersection level. The intersection geographies created for this analysis may or may not capture all crashes that occurred in the intersection, as geolocation is approximate. Previous work with PennDOT crash data has shown that the XY coordinates of crashes do not always match the crash narratives and diagrams contained in the original police reports.
- The PennDOT crash data collection and police reporting process is designed for automotive vehicles, which is understandable since vehicles are involved in almost all crashes. However, the preponderance of crash data involving vehicles and the focus of police reports on data fields related to vehicles produces a skewed understanding of what occurs in pedestrian crashes.
- Pedestrian fatality crashes exhibit patterns that indicate their causes are multifactorial. There may be two or more factors involved, such as speeding, nighttime, pedestrian location outside crosswalk, "not normal", male pedestrians (overrepresented in pedestrian fatalities), vehicles going straight, or location at an intersection with Roosevelt Boulevard. The sequence of events that lead to a crash are often more informative than single variables, such as a left turn. However, the PennDOT crash data does not capture the sequence of events that lead to a crash.
- Potential underreporting – both of crash incidence and the severity of crashes – is a data limitation acknowledged in previous work with PennDOT crash data.
- If someone involved in a crash was transported to a hospital, almost all reports included the name of the hospital. Most police officers also note that they followed up with the attending physician, frequently only including that the patient is "stable," which doesn't give any useful medical information about the severity and type(s) of injury or the potential outcome. Some reports describe important medical information relayed by doctors, such as "broken ribs" or "fractured skull."
  - Many injuries may be missed since they are not immediately outwardly evident (internal damage, broken bones, bruising, and traumatic brain injury). Other injuries may be missed because the police officer doesn't receive or record important medical information from the doctor.

- Some people in crashes do not want to be involved with the police or taken to the hospital, leading to incomplete information and underreporting.

The project team has observed, in previous analyses that included reviewing police crash reports, instances where crash severity was not consistently coded which could lead to underreporting the severity of crashes. For example, crashes where the injury was described as “bleeding in the brain” or “broken leg, neck, and back,” were categorized as “possible injury” or “unknown severity” rather than the more appropriate and more serious injury category. At the scene, it is easier to for the police officer to assess damage to a vehicle than to the human body, as generally the vehicular damage is more visible and doesn’t change over time.

- Historically, there have also been instances where the project team observed that minor injuries were coded as suspected serious injuries.

In general, the project team conducted the analysis of the PennDOT crash data with an attempt to keep data limitations in mind.

Analysis

The purpose of the analysis was to conduct an initial and macro-level exploration into the causation of the incidence of pedestrian crashes and the crash severity of pedestrian crashes between 2014-2018 in the City of Philadelphia. The project team used PennDOT crash data combined with cleaned City data to make preliminary conclusions about incidence and causation. The preliminary conclusions of this analysis are meant as a start to the conversation about causality of pedestrian crashes in Philadelphia and to serve as a reference document and touchstone for future in-depth analyses.

The project team’s analysis was split into three steps:

1. Summary statistics at the intersection- and corridor-level in Excel, using Pivot Tables;
2. Correlation tests, univariate modeling, and multivariate non-linear and logit regression modeling at the intersection- and corridor-level in R;
3. Spatial analysis, including kernel density, Moran’s I, and hot spot analysis, of all injury crashes, pedestrian injury crashes, and pedestrian fatality crashes

Summary Statistics (Excel)

Pivot tables are an Excel tool to easily combine, organize, and analyze data at a basic level. To better understand the overall crash environment, summary statistics were produced in Excel with pivot tables for the intersection and corridor analyses. The summary statistics for the variables were also compared to statistical modeling results as a quality check.

For the intersection and corridor analysis, 37 different variables were explored with pivot tables to identify crash factors that were notable due to overrepresentation, such as vehicle movements, near transit, or near a school.

Statistical Modeling (R)

Purpose

In this section, the project team conducted three types of statistical tests - correlation testing and non-linear based regression modeling (univariate and multivariate) - at the intersection- and corridor-level. Statistical tests were conducted to investigate two separate (and at times, overlapping) phenomena:

1. The relationship between 100+ dependent variables (such as time of day or presence of crosswalks) and the incidence of pedestrian injury crashes as compared to all injury crashes and the incidence of pedestrian fatality crashes as compared to all injury crashes. To test this set of hypotheses, logit modeling was used.
2. The relationship between 100+ dependent variables and the severity of pedestrian injury crashes. To test this set of hypotheses, non-linear regression modeling was used.

The results of these preliminary statistical tests are not meant to be standalone data points but are a complement to all the work conducted throughout the analysis (summary statistics, spatial analysis).

Data Limitations

This analysis is an initial and macro-level exploration into what are complex phenomena. Previous studies using PennDOT crash data has provided a window into what is possible with state-level crash data as it is. However, more than half of the dependent variables tested in this analysis are being normalized, tested, and modeled for the first time in this capacity. Due to the complexity and quirks of each new data source, the results of these tests should be considered preliminary. The hypotheses that were tested were kept as high level as possible, and are meant to indicate future channels of inquiry, rather than conclude on phenomenon. Therefore, the modeling conducted in this analysis should be considered preliminary. There are several additional data limitations that should be considered when interpreting the results of the models presented in this analysis, and in any further modeling.

Foremost among these is exposure. Where there are higher volumes of pedestrians, or cars, more pedestrian crashes (or car crashes) will occur. Therefore, it is essential to normalize crash data against existing volumes to gain a clear view of causation. Throughout the course of this analysis, the project team was unable to explore beyond exposure for many dependent variables. For example, when testing incidence and land use at the corridor level, the project team found that the Central Business District and the Central Business District Adjacent areas had higher odds of pedestrian injury crashes are compared to all injury crashes. However, the Central Business District and adjacent areas have some of the highest volumes of pedestrians in the City. If exposure was controlled for appropriately, tests would more accurately gauge if the odds of the incidence of pedestrian injury crashes in the Central Business District and adjacent areas were significantly higher *in relation to* volumes of pedestrians.

Introduction to Modeling

With a complete crash database, the project team embarked upon a series of statistical tests and models to get a deeper understanding of the statistical significance of relationships between variables and outcomes. Summary statistics are helpful but superficial descriptors of a dataset; however, statistical tests and models are essential to diagnosing the significance of phenomenon. All statistical tests and models are reliant upon the quality and reliability of the dataset; the preceding sections show the steps the project team took to ensure as complete and quality data as possible. However, as the preceding steps document, there are several ongoing issues and challenges with the quality PennDOT crash data.

The testing and modeling processes were as follows:

- Define two sets of hypotheses and independent variables – incidence and severity;
- Vetting and grouping dependent variables to test;
- Diagnostic testing of correlation coefficients between independent and dependent variables;
- Diagnostic bivariate non-linear regression modeling or logit regression modeling between independent and dependent variables; and
- Multivariate non-linear regression modeling or logit regression modeling based on the diagnostic results of the correlation coefficients and bivariate models to create a multivariate non-linear regression for intersection and corridor crashes.

As noted in the audience disclaimer, readers need to be familiar with statistics and statistical modeling. Theoretically supported relationships and statistical outputs (correlation coefficients, multivariate regression coefficients, odds ratios, and p-values) were considered in review of the materials covered in this section. “Significance,” in keeping with the industry and academic standard, was primarily determined by p-values under 0.05 (demonstrating that a phenomenon was consistent at least 95% of the time). Sometimes variables in a model are significant but their effect is negligible or small. This was denoted by descriptors of size, such as a “small” or “moderate” effect. The conclusions drawn from these tests are covered briefly below. Additional detail can be gleaned from the test outputs in the R database.

Defining Indepent Variables

Previous analysis of PennDOT data has focused on injury severity as the independent variable in modeling. In this analysis, two sets of hypotheses – and corresponding independent variables – were developed:

1) **Incidence**: In this set of hypotheses, the project team tested the relationship between the dependent variables and the odds that a pedestrian injury crash or a pedestrian fatality crash had occurred as compared to all injury crashes (mode-inclusive). Dummy variables indicated whether a crash was part of the following groups:

    a. All injury crashes – any crash in the City of Philadelphia between 2014-2018 that resulted in at least one injury;
    b. All pedestrian injury crashes – any crash in the City of Philadelphia between 2014-2018 that resulted in at least one pedestrian injury (inclusive of pedestrian fatalities); and
    c. All pedestrian fatality crashes – any crash in the City of Philadelphia between 2014-2018 that resulted in at least one fatal pedestrian injury.

2) **Injury** Severity: In this set of hypotheses, the project team tested the relationship between dependent variables and the severity of the pedestrian injury. As mentioned above, the injury severity categories used by PennDOT were simplified to four categories: no injury, minor injury (which encompasses Injury/Unknown Severity, Possible Injury, Suspected Minor Injury, and Unknown if

Injured), suspected serious injury, and fatality. The project team converted the simplified injury severity variable into a continuous integer field (no injury = 0, minor injury = 1, etc.). While this reduces each of the four injury severity outcomes to being of equal distance from each other, the project team accepted this as an acceptable method for preliminary modeling.

It is important to understand that these hypotheses cover different phenomena that, at times, overlap.

Grouping Dependent Variables

Refining the dependent variables was important to thorough modeling; the project team tested blanks and NAs and then explored different groupings within each variable. For example, the variable weather has nine categories: no adverse, fog, rain, rain and fog, sleet, sleet and fog, snow, and unknown. From a modeling perspective, it is unlikely that there is a statistical difference between many of these categories. The project team created a second variable, WEATHER_ADVERSE, to capture a binary weather variable to test: adverse or not. For other variables, like collision type, which categorizes crashes as sideswipes, head-on crashes, etc., new columns were created with dummy variables.

Dependent variables that had too many NAs included: travel speed (80% blank) and clothing type (95% blanks). Many important dependent variables had around 20% blanks, such as: vehicle type, vehicle position, vehicle movement, driver/pedestrian condition, and travel direction. Other dependent variables had around 10% blanks, such as age and sex.

Correlation Tables

Correlation tests are a bivariate analysis that measures the strength of associate between two variables. For the first set of hypotheses – incidence of pedestrian crashes – a Point-Biserial correlation was used because the independent variable was a binary value. For the second set of hypotheses – pedestrian injury severity – a Spearman correlation test was used. Correlation tests are included at the beginning of each sub-section in the modeling section of the R code (lines 2169-3756).

Univariate Non-Linear Regression and Logit Regression

Univariate non-linear regression analysis and univariate logit regression analysis measures how closely one dependent variable is related to the independent variables (logit is used when the independent variable is a binary variable, as is the case for the first set of hypotheses that relate to incidence of pedestrian injury crashes). However, as the hypotheses being tests are complex phenomenon, a single dependent variable is unlikely to be satisfactorily explanatory (this would be an example of omitted variable bias). Therefore, the univariate analysis was used as a diagnostic tool in assessing the potential statistical strength and the direction of relationships with the independent variables. Univariate tests are included at in each sub-section in the modeling section of the R code (lines 2169-3756).

Multivariate Non-Linear Regression

Multivariate non-linear regression analysis and multivariate logit regression analysis attempts to explain (or "fit") a phenomenon (y, the independent variable) with two or more dependent variables. This was the final and most important step in drawing conclusions about whether dependent variables – time of day, weather conditions, contributing factors –

were *significantly* related to either incidence or injury severity. Multivariate tests are included at in each sub-section in the modeling section of the R code (lines 2169-3756).

Area-Level Spatial Analysis (ArcGIS)

For the area-level of analysis the project team used ArcGIS. The team conducted four types of spatial analysis (1) Kernel Density, (2) Global Moran's I, (3) Anselin Local Moran's I, and (4) Getis – Ord Gi* (G-I-Star) on all injury crashes, pedestrian injury crashes, and pedestrian fatality crashes.

Kernel Density

Kernel Density analysis calculates the density of all crashes and all pedestrian crashes. A smoothly curved surface is fitted over each point. The surface value is the highest at the location of the points and diminishes with increasing distance. The Kernel Density was used gain insight on the general density distribution.

Global Moran's I

*Getis – Ord Gi**

The Getis – Ord Gi* analysis used the resultant z-score and p-value to determine whether the features are either high or low value cluster spatially. In this case, to be a statistically significant hot spot, all crashes and pedestrian crashes need to have high values surrounded by other high values. Getis – Ord Gi* was used to identify locations of hot spots and cold spots.

Prioritization (ArcGIS)

Global Moran's I measures the spatial autocorrelation based on feature location and feature value simultaneously. In this case, the feature values were all crashes and all pedestrian crashes. The feature location was based upon census block areas. Global Moran's I evaluates whether feature locations and feature values are clustered, dispersed, or random. Spatial distance between areas were used to calculate weight. This process helped to understand the whether there were statistically significant clustered patterns occurring in Philadelphia (at the 99% confidence level).

Anselin Local Moran's I

The Anselin Local Moran's I technique identifies statistically significant hotspots, cold spots, and spatial outliers. This provided a more comprehensive understanding of the spatial distribution of all crashes and pedestrian crashes. Block groups identified as high – high clusters or low – low clusters indicate the feature is surrounded by other features that have similar values (high positive z-score). Areas with high – low outlier or low - high outliers indicate areas that are surrounded by features that have dissimilar values (low negative z-score).

Identifying places to focus pedestrian improvements is the basis for an implementation program. Priority intersections, priority corridors, and priority areas were identified for the City of Philadelphia based on the findings.

Intersection Prioritization

To create a list of priority intersections, all pedestrians injured or killed in crashes at each intersection were added together. Pedestrian fatalities were weighted four times larger than injuries. Each intersection was then sorted by the highest number of pedestrian fatalities, and then the highest number of pedestrian injuries. Aligning with Vision Zero's goal of bringing traffic deaths to zero, this prioritizes intersections with high numbers of pedestrian fatalities.

Corridor Prioritization

To create priority corridors, crashes that occurred along each corridor were added together. Corridors are segments of streets that have the same street name, functional classification (e.g. major arterial, minor arterial, expressway), Complete Streets typology (from the City of Philadelphia's 2017 Complete Streets Handbook, which created street typologies such as Urban Arterial, Park Road, City Neighborhood Street), and are longer than 1,000 feet.

To create a list of priority corridors, all pedestrians injured or killed in crashes in each corridor were added together. Pedestrian fatalities were weighted four times larger than injuries. Corridors were then sorted by their "score."

Area Prioritization

To identify priority areas in Philadelphia for focused safety improvements, a Hot Spot analysis was conducted using all injury, pedestrian injury, and pedestrian fatality crashes. The Hot Spot analysis generated a Getis-Ord Gi* statistic for each crash in the dataset. The scores, taken into consideration with nearby context, identify areas where high or low values cluster spatially within Census Block Groups. For the purposes of this analysis, only clusters of high values were considered, or areas where high numbers of crashes occurred at or above the 90% confidence interval. Area clusters of pedestrian injury crashes were then overlaid with area clusters of pedestrian fatality crashes to identify the most serious pedestrian crash problems in Philadelphia, which form the Priority Areas.