APPLYING GENETIC ANALYSIS TECHNIQUES TO PEDESTRIAN CRASH DATA

by

Megan Rosica

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Master of Civil Engineering

Spring 2017

© 2017 Megan Rosica
All Rights Reserved

APPLYING GENETIC ANALYSIS TECHNIQUES

TO PEDESTRIAN CRASH DATA

by

Megan Rosica

Approved: ____
Earl Lee II, Ph.D.
Professor in charge of thesis on behalf of the Advisory Committee

Approved: ____
Harry W. Shenton III, Ph.D.
Chair of the Department of Civil and Environmental Engineering

Approved: ____
Babatunde A. Ogunnaike, Ph.D.
Dean of the College of Engineering

Approved: ____
Ann L. Ardis, Ph.D.
Senior Vice Provost for Graduate and Professional Education

ACKNOWLEDGMENTS

I would like to show my sincerest gratitude to my advisor, Professor Lee, for his immense knowledge, patience, and direction during my time in graduate school. Through every challenge I was able to count on his help and guidance. Without his dedication and support this thesis would not have been possible.

I would also like to gratefully acknowledge Mindy Laybourne, Gene Donaldson, and the DelDOT TMC for the extensive opportunities, experience, and knowledge I gained while working with them throughout my graduate school career.

This research was also greatly supported by the Department of Civil and Environmental Engineering at the University of Delaware and my utmost thanks goes to the faculty and staff members who have helped me throughout my time here.

I owe my deepest appreciation to my parents, my family, and my friends who have been there for me through everything and who have shown me unconditional love and support. I thank them for their constant motivation and encouragement, without which none of this would be possible.

TABLE OF CONTENTS

Appendix

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

No single collision is identical to another: like the structure of DNA, each crash involves a specific sequence of events and distinct characteristics. Current crash data analysis tools and visualization techniques focus on specific factors of crash events, but difficulties arise in attempts to create a single visualization to show all characteristics. The goal of this research is to analyze and visualize a large data set that includes multiple aspects of pedestrian crash data in the State of Delaware. The data set included individual crash factors as well as roadway traits. Even though this state is one of the smallest in America, Delaware continues to have extremely high rates of pedestrian fatalities year after year. By using a genetic based evaluation of pedestrian crash data in Delaware, conclusions can be drawn about possible causes of pedestrian injuries and fatalities, and can aid in the prevention of these incidents occurring.

Chapter 1

INTRODUCTION

The number of pedestrian fatalities involved in motor vehicle crashes in the United States is growing every year. In 2014, there were 4,884 pedestrians killed and around 65,000 injured on the roads. Of the 32,675 total fatalities due to motor vehicles accidents in America, fifteen percent were pedestrians, and this percentage has increased over time from 11% in 2005 to 15% in 2014 (NHTSA, 2016).

The high rate of pedestrian fatalities is a problem in America, but an even bigger issue in Delaware. Although the state is one of the smallest in the nation, it had the highest rate of pedestrian deaths in 2015, with 3.7 pedestrian fatalities per 100,000 people. It also had the highest rate in 2012 and 2013 with 2.94 and 2.70 fatalities per 100,000 people, respectively (NHTSA, 2016). In its highest populated county, New Castle County, pedestrian crash rates increased 31% from 2005 to 2014 (WILMAPCO, 2016). There have been recent initiatives to try and educate the public and prevent the occurrence of pedestrian incidents on roadways in Delaware. The Office of Highway Safety (OHS) and the Delaware Department of Transportation (DelDOT) have developed a campaign called Walk Smart to promote pedestrian safety throughout the state (OHS).   The program has published creative posters and flyers with catchy slogans and characters encouraging the use of crosswalks and practicing caution when using the roads. Law enforcement officers are also a large part of the program. Local and state police monitor high volume and high risk areas to remind violators of the issues they face if they are not aware of the safety precautions they must take (Markell, 2015).

In order to alleviate the increasingly high pedestrian fatality rate, it is important to first figure out why these incidents are occurring. An extensive analysis of the characteristics of pedestrian related crashes and roadway traits is necessary in order to come up with a reason for why so many crashes are happening each year. If trends in the characteristics of crashes and facilities can be found, then improvements can be made to prevent vehicle-pedestrian collisions from occurring. The objective of this study is to evaluate current methods of crash data analysis and visualization as well as preventative measures that are currently in place; develop a genetic based evaluation that looks at multiple characteristics of each crash and visualize this data so that state agencies and the public can understand the data trends; and draw conclusions to find answers about what could be causing the recent increase of pedestrian fatalities in Delaware. As stated in Karl Kim's research involving modeling fault among pedestrian-vehicle accidents, "increased pedestrian safety requires the development and implementation of effective enforcement, education and engineering strategies (Kim, 2008)." This research will include strategies to improve engineering and design techniques as well as education for pedestrians in Delaware.

The next chapter will review the literature related to current crash data visualization techniques, crash data analysis methods, and GIS applications in traffic safety data. Chapter 3 will cover the methodology of this study and discuss the steps that were taken to organize the data as well as the genetic based approach that was used to analyze the given data. Results of the study will be shown in chapter 4 and will include scored tables, mapping methods and various techniques to visualize the large data sets in a comprehendible way. Chapter 5 will discuss future work to be done following this study in order to evaluate pedestrian injury and fatality data in Delaware.

Chapter 2

LITERATURE REVIEW

This literature review was completed in order to provide insight into previously conducted studies related to crash data analysis, visualization techniques, and GIS applications in traffic safety engineering. Although all of the cited studies and techniques have contributed to crash data analysis and visualization, the research conducted in this thesis will provide new applications into traffic safety and pedestrian related crash data analysis and visualization. The major objective of this research is to develop an analysis and visualization procedure to draw conclusions regarding the high rate of pedestrian fatalities in the State of Delaware. Past projects have focused on pedestrian crash data analysis, vehicle crash data analysis, or more general crash data visualization, but none have focused on combining multiple roadway and incident characteristics in pedestrian crash data in order to develop visualization techniques.

2.1 Crash Data Visualization

When it comes to visualization of crash data, many techniques have been recorded and studied. In Australia, a case study was done to look into crash data and visualize it in an interesting way. Sahar Alian used a statistical approach and focused on road geometry and the connection between road design and crash incidents. A sinuosity index was created to show how road curvature can alter driver behavior and therefore incidents to occur. ArcGIS was also used in the research to look at road segments, as opposed to individual crashes and locations along a roadway. Data visualization included a diagram that showed how the number of accidents that occur vary over segments with a higher sinuosity index both during the day and during the night (Alian, 2016). This original way to show data trends was easily understood and effective in proving where trends occur along the Kings Highway in Australia. Regression analysis was performed to link characteristics to accident occurrences. The only traits this research focused on was road geometry, time of day that the crash occurred (day or night) and which direction the vehicles were traveling in. Fewer characteristics allows for an easier regression analysis to be done so statistically conclusions can be drawn. This data did not focus on specifically pedestrian crash safety data.

Another visualization method was used by Charlotte Plug and her team in Western Australia to look at single vehicle crash occurrences between 1999 and 2008. Spatial, temporal, and spatio-temporal methods were used to visualize data and find patterns in crash traits using spider plots, kernel density estimation (KDE), and comap. Spider plots were used to show when crashes were occurring most frequently, i.e. hours during the day and days of the week. KDE was used to show crash density, like in other sources and regional maps showed where crashes were located, using the "hot spot" technique. Finally, comap was used to combine both the temporal and spatial visualization. Comap is an analysis technique that uses the time of a crash occurrence and relates it to the area that these crashes occurred, using KDE. With information about when the crashes happened and where, conclusions were drawn about specific crash traits and how they were linked. For example, traits like speed and fatigue-related crashes were linked with time of occurrence and location. It was found that speeding happened more often at night, and fatigue-related crashes were most common in non-metropolitan areas, during the times of 1-2AM and 2-4PM (Plug, 2011). Other conclusions included information about what time of day crashes were occurring during certain days of the week, where crashes occurred most often and during what times, and the variation between hitting pedestrians and hitting objects. It was evident in the analysis that pedestrian crashes were most common around 8-9AM, 12PM, and 3-6PM because of the high volumes of pedestrians during this time, but object crashes mostly occurred during 10PM-4AM. Overall this source was interesting in analyzing single vehicle crash data and connecting crash traits spatially and temporally. There was no scoring technique used or main focus on pedestrian injury and fatality in this data.

Many sources used analysis to find “hot spots” along roads and visualized crash data based on density and frequency of crashes. For example, Saffet Erdogan and researchers focused on finding “hot spots” based on spatial characteristics in a case study that took place in the country of Turkey (Erdogan, 2008). Erdogan used similar characteristics to the research presented in this thesis by incorporating accident traits such as the season, day of the week, time of day, and driver behavior. Researchers were able to draw conclusions relating accident hot spots to various accident traits which led to engineering and management recommendations, such as increased signage in certain areas as well as improved driver training. Erdogan used statistical analysis, mainly K-means clustering techniques and Chi-squared tests, for their data analysis, which was how they were able to draw important conclusions about their data. Although similar methods were used, Erdogan did not employ any new data visualization techniques; researchers instead employed the use of charts and tables to show accident occurrences over time and during certain days of the week.

The research shows maps with crash data “hot spots”, some based on density, and some based on different numbers of reliability. The importance of detailed crash reporting is emphasized, which is relevant to any crash data analysis. If the data is not accurate or missing then an accurate analysis is more difficult to construct. Similarly, Brian Hilton and his team of researchers developed a publically accessible map program that shows users where dangerous areas in a community are based on the frequency of vehicle crashes over time (Hilton, 2011). The visualization mostly consisted of a heat map that showed frequency and density of crashes. There was no pedestrian focus and no specific crash characteristics were studied or linked to incidents. Surveys completed by users of this product concluded that even though the heat maps clearly showed where dangerous locations were, driver behavior was not likely to change.

In order to develop an improved incident management program, Asad J Khattak and Nagui Rouphail report on research methods to find where accidents were occurring most often and how patrol operations should be organized dependent on this location data (Khattak, 2005). This idea is similar to finding “hot spots” and focuses mostly on location and frequency, as well as density, of crash occurrences. Although their research did not focus on any other characteristics other than location and traffic volume, it similarly mapped high-risk locations and conclusions were drawn about the best places to have patrol operations, as a way to address crash incidents.

2.2 Crash Data Analysis

Many researchers focus only on analysis techniques to draw conclusions about crash data and do not use visualization data in their studies. An interesting look into crash data analysis was taken by Karl Kim and his research team in a case study done in Hawaii. This research focused on pedestrian data and used regression analysis and statistical techniques to identify common factors in pedestrian related fatalities and serious injuries (Kim, 2008). The research employed many of the same characteristics as were used in this thesis, such as the month, day of the week, time of day, weather conditions, and certain human factors, but Kim's research focused on statistical regression analysis for explanatory models rather than a scoring system and visualization method. They also looked at accident fault, whether it was the pedestrian or the driver who caused the accident. This is a possible future installment in the current thesis. These researchers modeled specific demographics to draw conclusions, such as drunk pedestrians, child pedestrians, and male business commuters. Tables were used to display the statistical analysis results.

Similar to Karl Kim's research in Hawaii and the thesis presented, a study performed by Monica Soares Velloso and Maria Jacques was completed to develop an investigatory procedure to identify factors that contribute to pedestrian crashes in Brazil. The methods involved data mapping, planning, research and data collection and storage, and analysis (Velloso, 2012). All contributory factors of crashes were coded for the data analysis portion, and factors included seven groups: personal details (impairment, distraction, behavior, inattention, etc.), details of the pedestrian (behavior, child vs. elderly, where they crossed from, etc.), details of the driver (speeding, behavior, aggression, etc.), mechanical problems on the vehicle, conditions of the site (road conditions, winds, etc.),obscurity (view, surroundings, weather, etc.) and animals involved. Velloso's researched used much more detailed crash factors than presented in this thesis and were able to code each detail of the crash to draw conclusions. A total of 120 pedestrian crashes were analyzed and results show that causes of crashes were closely associated to pedestrian behavior: negligence, the use of alcohol, and lack of attention were the main contributory factors. Also, there was an association found between pedestrian behavior and highway environment: the combination of “pedestrian” and “highway” as features in the crash data was present in 37.5% of the pedestrian crashes (Velloso, 2012). This research used a statistical coding approach rather than regression analysis or visualization techniques, but conclusions drawn were important and relevant to pedestrian crash data analysis.

In contrast to Kim’s and Velloso’s research, Tibebe Beshah and Shawndra Hill took the focus off of human behavior and used data mining technologies to find connections between accident severity and road characteristics in Ethiopia (Beshah, 2010). They looked at similar categorical attributes as this thesis presents, such as road surface, weather condition, and light condition. They used classification models, including WEKA explorer and various algorithms, to find results about crash data in Ethiopia. This method did not use any visualization or GIS analysis to locate crashes, it only used the characteristics of roadways where different accidents occurred. There was also no focus on pedestrian incidents, but the contributions of this study can be applied to any crash data analysis: it is important to find conclusions in data so that “decision makers can formulate better traffic safety control policies, label roads with necessary signs informing drivers and pedestrians of accident risks, and design better roads (Beshah, 2010).”

2.3 GIS Applications In Traffic Safety Data

ArcGIS is a very important tool used in this research. The program allows for a comprehensive view of crash locations and was crucial in associating locations of crashes with roadway characteristics and pedestrian crash traits for this thesis. As previously stated in the literature above, GIS is often used for traffic safety analysis due to its ability to map and visualize data: Ergodan and the research done in Turkey used GIS to determine “hot spots” and link crash characteristics with location; Khattak focused on incident management assistance patrols based on the use of GIS; and Alian used GIS to create sinuosity indices to learn about road geometry and the occurrence of crashes in Australia.

Another project involved Indira Khanal who used a software approach that supplements AASHTO’s Highway Safety program by providing location of crashes, color-coded information, charts and tables (Khanal, 2014). Researchers used Arc GIS and Google maps as primary visualization methods for base maps. Without GIS their research would not have been possible. Although Khanal’s research accurately emphasized the importance of the spatial nature of traffic safety and provided insight into how to visualize crash location using GIS, they did not use coinciding crash characteristics in the mapping techniques and had no specific focus on pedestrian incidents as this thesis has done.

Another contributor is Aline Aylo who developed a thesis that used GIS to develop a safety analysis application. Similar traits as the present thesis were used, such as day of the week, weather conditions, light conditions, etc. but this application was different in that Aylo developed a specific software program using GIS to analyze traffic accidents (Aylo, 2010). Statistical analysis was used, and there was no crash data visualization technique, except to show crash frequency and density on maps. A macro and micro approach was made available in this program, giving the chance for users to focus on specific road segments or groups of road segments as well as specific intersections or groups of intersections. The application was created to analyze traffic safety data, but it took a statistical approach to create a software tool rather than visualizing data and linking accident traits and locations. This research was helpful in showing how GIS can be used for traffic data analysis and software tool development, but it did not focus on pedestrian data and did not display multiple characteristics of crashes in one visualization.

Chapter 3

METHODOLOGY

The following chapter will cover the methods used in this research to evaluate and visualize pedestrian crash data in the State of Delaware. Methods included the processes of collecting and organizing data, developing a match formula as a scoring system for pedestrian crashes, using ArcGIS to map those scores, and providing unique visualization techniques to identify important pedestrian crash traits.

3.1 Background

Pedestrian crash data for this research came from the Delaware Department of Transportation (DelDOT) and the Wilmington Area Planning Council (WILMAPCO). Research methods began with pedestrian crash data for three years (2013-2015), evaluating one year at a time. This information included many details from vehicle incidents but did not include any personal information about civilians involved in the crashes. Once this data was organized and consolidated, it was converted to an ArcGIS shapefile and was mapped along with associated roadway characteristics. DelDOT and WILMAPCO provided shape files that consisted of up to date roadway characteristic information for the entire State of Delaware. Tools in ArcMap, including spatial join, were used to create one large dataset that combined all pedestrian crash data with corresponding roadway characteristics. This dataset was converted to an excel file which was then organized and cleaned. In excel, pivot tables were created to find the most common attributes for each crash. Each crash was then given a score based on the number of the most common characteristics it displayed, and the scores were plotted in ArcMap to view any significant connection between crash location and higher scores. Visualization techniques were used to show this data in a single Microstation display, which resulted in extended knowledge of pedestrian crash data traits and corresponding road characteristics.

3.2 Pedestrian Crash Data Consolidation and Organization

The data obtained was thorough and included many important characteristics of pedestrian-involved crashes. The raw data held over 190 fields, however many of these were not useful to this analysis. Arranging the data was important so that the most relevant traits were extracted and studied to draw conclusions. Certain fields that were deemed irrelevant included: the name of the officer who was there on the scene, the current agency, whether or not the complaint was approved, what sector the crash occurred in, and many more technical descriptors that were unnecessary to have in the analysis, either because they could be accounted for in another way or because they did not affect the incident data directly. Often in the data set there would be an entire field left blank because the crashes did not involve those characteristics. The information was filtered so that only the following pedestrian crash traits were collected and used in the analysis:

- Complaint ID
- County
- Latitude and longitude
- Fatality or injury
- Driver action
- Road surface

- Light condition
- Weather condition
- Day or night (based on light condition and time)
- First unstable
- Environment
- Whether or not alcohol was involved
- Month
- Day of the week
- If the crash occurred during a peak travel time (7-9 AM or 4-6 PM)
- Roadway circumstance
- Roadway junction
- Whether or not the crash was a hit and run

Some of the above characteristics were combined or extracted from the entirety of the data given. For example, the month was extracted from the date of the crash. It was decided that information about the month that the crash occurred in is more valuable than the actual date of the crash. It also makes more sense to compare crashes month to month rather than date to date. Using the time a crash occurred, two traits were extracted ("Day or night" and "During peak hours") for similar reasons. If the crash was classified as happening during "daylight" or "dawn" and occurred during the hours of daylight, then it was given the trait "Day." Any other light condition, including "dusk," was deemed "Night." Keeping it simple and comparing whether a crash occurred during the day or during the night is more valuable than comparing the specific time at which the crashes happened. Also, examining whether or not the crash occurred during a peak travel time gives more information about if crashes occurred when there were higher numbers of vehicles on the roadways. If more crashes were occurring during peak hours, then more focus could be put on emphasizing pedestrian and vehicle safety during those times of day, rather than just observing an actual time on the clock. In this research there were many instances when it was more valuable to extract specific information from the data than to use it as it was given.

3.3 ArcMap and Spatial Joining Data

Once this pedestrian crash data was consolidated and organized, it was converted to a table in ArcGIS and consequently a shapefile using Excel conversion tools in ArcMap. Because the latitude and longitude coordinates of the crashes were given, it was possible to map these crashes with all of their data as points on a map. Each shapefile that was added to ArcMap from excel uses the geographic coordinate system as a spatial reference based on latitude and longitude coordinates.

The crash points were then separated into intersection and non-intersection locations. Using an intersection buffer shapefile, given by WILMAPCO, crashes were classified as occurring either at an intersection or at a non-intersection. Using the "Select by Location" tool in ArcMap, a crash was "selected" if it fell within an intersection buffer. After this selection tool was applied to all pedestrian crashes, a separate layer was exported that included all intersection-related crashes. By inverting the selection, all of the non-intersection-related crashes were exported into their own layer. This new layer with the non-intersection crashes holds the data that would be used for the duration of this research. Historically, and seen in the Delaware pedestrian data, more pedestrian crashes have occurred at non-intersections (NHTSA, 2016). Also, intersection-related incidents possess different roadway traits than those located at non-intersections and therefore require extended and varying types of analysis; so this research focused mainly on non-intersection related pedestrian accidents.

After the pedestrian crash data was separated into non-intersection events, various roadway characteristics were imported into ArcGIS to be mapped along with the pedestrian crashes in order to spatially join the characteristics with the crash data. Each roadway characteristic was given as a separate shapefile so that all could be joined and viewed together, along with the crashes. The roadway characteristic shapefiles that were used and linked with pedestrian crash data were given by DelDOT and WILMAPCO and are listed below:

- Speed Limit
- Number of Lanes
- Road Width
- General Maintenance Responsibility
- Sidewalk
- Bus stop (within 100 feet)
- Channel
- Median
- Shoulder

Using the "Spatial Join" tool in ArcMap, each trait was linked to any pedestrian crash that fell within a 75-foot radius, not including whether a bus stop was in a 100-foot vicinity. For example, the entire pedestrian crash data layer was spatially joined with the Speed Limit layer, so that each crash was linked with the speed limit of the road that the crash occurred on. A new layer was created that held all of the pedestrian crashes, but now had traits associated with speed limit. This was done continuously for each of the above traits so that as a result, a final layer was created that embodied all pedestrian crash data as well as all data for the roadway characteristics. This method allows for each crash to be spatially associated with the traits they possess.

After the crash data was linked to the associated roadway traits, the attribute table for the final layer with the combined information was exported from GIS into an excel file, again using the Excel conversion tool in ArcMap. Similar to the pedestrian crash data, this newly compiled data had many fields that were deemed irrelevant to the research in this study. Many of the roadway information shapefiles contained classification fields that were unnecessary to the entirety of the data. Specific fields were removed from the data such as codes for certain values, shape lengths, roadway IDs, object IDs, etc. After removing all unnecessary data fields, the remaining traits were those listed above. The final excel file for the entire year included the previously listed pedestrian crash traits with the previously listed roadway characteristic traits, totaling twenty-three fields of study to be examined and compared crash by crash.

3.4 Pedestrian Crash Match-Scoring Technique

Each pedestrian crash was then associated with roadway characteristics, but not all crashes were located on the same types of roadways or possessed the same types of roadway traits. Analysis of the patterns in the crash data and their locations was necessary to find relationships. First, pivot tables were created for each of the twenty-three fields to see which characteristic was most common under each category. For example, when it came to road surface conditions for each crash, the most common was a “dry” road. After finding the most common characteristic throughout all the data, a score was given to each crash if it had the trait that matched the most common characteristic. Continuing with the road surface condition example, if a crash occurred on a “dry” road surface, it was given a “1” for that characteristic. If not, it was given a “0”. Therefore, the highest score a crash could receive would be a 23 if that crash had every common trait.

Binary scores were chosen for this type of analysis. A scaled ranking system was attempted for visual display and severity of crash, but it made the analysis more complicated than necessary. Using binary scores rather than rating a trait based on a scale more easily allows for a simple and objective approach. Binary scores also provide a clearer visual display: either the crash exhibits a common trait, or it does not. The objective of this rating method is to find commonalities between traits and crashes in the simplest way. If a large number of crashes all possess a certain trait, maybe that specific trait should be monitored or focused on by engineers and educators when it comes to prevention of pedestrian crashes.

After scores of either 1 or 0 were given to each crash element, the scores were added for each crash. If a crash had a higher score then it possessed many of the common traits found in the dataset for that year, and perhaps this means that those common traits should be targeted and fixed if they were spread over a large number of the crashes.

As noted above, all pedestrian data (injuries, fatalities, and property damages) was taken and each crash was scored using the common traits. It was also important to separate pedestrian fatalities from the rest of the data to try and observe if the linkages varied based on whether or not the crash was fatal. New pivot tables were created to see how the most common traits varied between fatal crashes and non-fatal crashes and the same processes were used to compare fatal crashes to one another and draw conclusions. More information about fatal crashes vs. total crashes can be seen in the results section of this thesis.

3.5 ArcMap and Plotting Scored Pedestrian Crashes

Once each crash had an associated score, an excel file was created specifically for the crashes and their scores. The file only consisted of the complaint number of the crash (for ID purposes), the county the crash occurred in, the latitude and longitude coordinates, whether the crash was fatal, and the associated score. The crashes were then grouped into low, moderate, and high score categories. If the crash fell between scores of 1-13 it was considered low; 14-17 it was considered moderate; and 18-23 was considered high. Separate sheets were created for each category of score. This file was then converted to a table and eventual shapefile in ArcMap. Each category of score was made into a shapefile and mapped as a blue dot (low score), yellow dot (moderate score), or red dot (high score). Mapping the category of scores allowed for visual display of where the most common traits of pedestrian crashes were occurring over that specific year in the state of Delaware. It should be noted that the same process was done specifically for fatal crashes. Exactly like the full data set, fatal crashes were scored and subsequently grouped into the same categories, based on the most common fatal traits, as opposed to the common traits in the full data set. The scores were also mapped and viewed on a full map of the State of Delaware in ArcMap.

The process of scoring the crashes, inputting them into score categories, and mapping them in ArcMap was completed for all three years of given data (2013-2015). Shapefiles were created for each scored category and for each year of data so that all years could be mapped into a final visualization over the entire state. Fatal crashes were separated for each year as well and viewed in separate ArcMap files.

3.6 Microstation Visualization Technique for Scored Crashes

A unique visualization method was developed in Microstation to view all scored fatal crashes in one display for each year of data. Similar to a genome map, a large circle was drawn with 23 spaces inside, one for each crash/roadway trait. The circle was then cut into “slices”, the number of which depending on the number of fatal crashes in the year of study. For example, in the year 2013, there were 15 fatal crashes that occurred at non-intersections, so the gene-like visualization included 15 “slices” with 23 spaces in each. For every uncommon trait (every “0” score) the associated space was filled in for that crash. The final product resembled a genome map, although on a smaller scale, but allows for a unique visual display of crash data and roadway characteristics as if each trait were a gene in a crash DNA molecule.

Chapter 4

RESULTS

The results obtained from the previously defined research methods are outlined and described in this chapter. After organizing pedestrian crash data, spatially associating the data with roadway characteristics, scoring each crash based on commonalities, and subsequently mapping those scores, conclusions were drawn about pedestrian crash data and their locations in the State of Delaware. Results include percentage of crashes that hold common traits, distributions of crash data scores, maps with scored data, a comparison between fatal and overall data, and a unique genetic based visual to show common traits over a number of fatal crashes over time in Delaware. A discussion of some limitations to this research will also be included in this section.

4.1 Common Crash Traits

Before the data could be scored, the most common traits were established and analyzed for each year. Crash numbers for all three years totaled to 587 pedestrian-related, non-intersection incidents; 472 of which were personal injury crashes, 58 involved property damage only, and 57 were fatal crashes. Results per year varied but there were some strong similarities between the data when it came to the most common crash attributes. It is evident in the data that throughout the years 2013-2015, more than 80% of pedestrian related crashes occurred on dry road surfaces. About 50% of crashes every year occurred during the day, in daylight and over 70% of all crashes occurred during clear weather conditions. Over 80% of crashes each year did not involve alcohol and around 12% of crashes each year occurred during the months of July or August while about 20% of crashes each year happened on a Friday or Saturday. About 50% of crashes occurred on roadways that have a speed limit of 25 mph. Between the years of 2013 and 2015, more than 70% of crashes happened on two-lane roadways and about 60% of all crashes happened on state-maintained roads. For each of the three years, almost 100% of crashes occurred on roads with no presence of channels, medians, or shoulders.

The most common traits were identified for all three years of the total pedestrian crash data as one dataset. Most incidents occurred on dry road surfaces in clear weather, during the summer months and during the day on a 25 mph speed limit two-lane roadway. The only differences between common traits through the years are highlighted in the table below. It should be noted that most of the common traits were shared year to year, or were very similar, for example the most common day of the week for a crash to occur in 2013 and 2014 was Friday but in 2015 it was Saturday. Both can be considered weekend days. Also, in 2013 the most common month for a crash was August but in 2014 and 2015, it was July. These are both summer months that fall in succession. The data shows strong trends in accident traits throughout each year.

Table 1 Most common characteristics per year for incidents involving pedestrians in non-intersection related locations. The notable differences in traits from year to year are highlighted.

		Most common characteristics		
	Characteristic	*2013*	*2014*	*2015*
1	Driver action	Other	Blank	Blank
2	Road surface	Dry	Dry	Dry
3	Light condition	Daylight	Daylight	Daylight
4	Weather condition	Clear	Clear	Clear
5	Alcohol?	N	N	N
6	Month	August	July	July
7	Day of the week	Friday	Friday	Saturday
8	Peak hour?	N	N	N
9	Day or night?	Day	Day	Night
10	First unstable	On Roadway	On Roadway	On Roadway
11	Environment	None	None	None
12	Roadway circumstance	None	None	None
13	Roadway junction	Non-Junction	Non-Junction	Non-Junction
14	Hit and run?	N	N	N
15	Speed limit	25	25	25
16	Number of lanes	2	2	2
17	Road width	26, 38	26	26
18	General maintenance responsibility	State	State	State
19	Sidewalk	Y	N	Y
20	Bus stop (w/in 100')?	N	N	N
21	Channel?	N	N	N
22	Median?	N	N	N
23	Shoulder?	N	N	N

4.2 Crash Score Distributions

Once the most common traits were determined for each year, it was possible to score the crashes. As previously stated, the lowest score a crash could obtain was 0 and the highest was 23. The lowest observed score for any crash in the three years of studied data was 6, in 2014, and the highest was 21, in 2015. This means that there were crashes that possessed almost all of the common traits during that year of data, but there were also crashes with very low numbers of common traits. Scored crashes yielded average values between 15 and 16 over the three years of data. Only thirteen crashes were scored lower than 12 and only forty were scored greater than 18. The histograms below show the distribution of crash scores for each of the three years studied. As the charts below show, the scores are normally distributed each year.

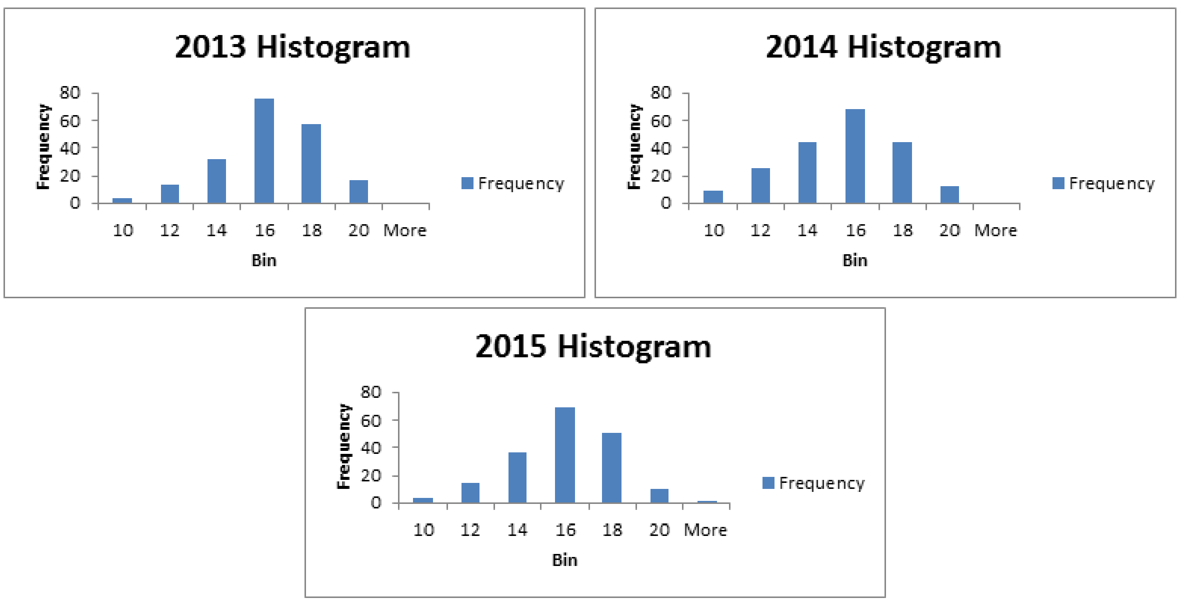


Figure 1 Histograms show the distribution of scores throughout each year, all of which are normally distributed around mean scores that fall in the range of 15-16.

4.3 Fatal Crashes vs. Total Crash Data

Year to year, fatal crashes increased throughout the State of Delaware, even when total non-intersection pedestrian crash numbers decreased. In 2013 there were a total of 199 non-intersection related pedestrian crashes. Of these 199 crashes, 15 were fatal. In 2014, there were 202 crashes, 19 of which were fatal; and in 2015, there were 186 pedestrian crashes, 23 of which were fatal. The results of the scored crash data varied when focusing only on the fatal crash information. These differences made it important to study the fatal data separately and identify patterns. For example, although it was uncommon in the general data for a crash to involve alcohol, in fatal crashes about 50% of them involved alcohol during 2013 and in 2015 78% of fatal crashes involved alcohol. Also, more than 80% of all fatal crashes every year occurred during the night, and about 60% occurred in unlit conditions. About 70% of all fatal crashes occurred on two-lane roads each year and 60% of fatalities occurred in clear weather conditions, similar to the general data. As previously stated, some crash and roadway characteristics were the most common for all recorded crashes, but when focusing on just fatal crashes, differences in common traits were found. Knowing the differences between the fatality and injury or property damage characteristics can help to draw conclusions about the most dangerous characteristics for pedestrians on roadways.

The table below shows the common traits of fatal crashes throughout the three years of data and the differences between them. Similar to the general data, there were strong trends between the years. The most notable difference is that in 2014, the month with the highest number of fatalities was December, which seems uncommon given that in crashes were evenly distributed in 2013, with one or two crashes occurring in each month. Also in 2015 the most common month was July similar to the general data which shows pedestrian crashes were more common during the summer months of July or August. It is, however important to note that a high number of fatal crashes involved common trends. For example, it was consistent that every year more fatalities occurred on dry road surfaces in dark conditions at night, on two-lane roads with speed lmits higher than 25 mph. Although there are less crashes that involve fataltıes to study, these trends are important to notice.

Table 2	Most common characteristics per year for fatality incidents involving pedestrians in non-intersection related locations. The notable differences in traits from year to year are highlighted.

		Most common characteristics		
	Characteristic	*2013*	*2014*	*2015*
1	Driver action	Failed to yield right of way	Blank	Blank
2	Road surface	Dry	Dry	Dry
3	Light condition	Dark-Not Lighted	Dark-Not Lighted	Dark-Not Lighted
4	Weather condition	Clear	Clear	Clear
5	Alcohol?	Y	N	Y
6	Month	no max	December	July
7	Day of the week	Thursday	Saturday	Saturday
8	Peak hour?	N	N	N
9	Day or night?	Night	Night	Night
10	First unstable	On Roadway	On Roadway	On Roadway
11	Environment	None	None	None
12	Roadway circumstance	None	None	None
13	Roadway junction	Non-Junction	Non-Junction	Non-Junction
14	Hit and run?	N	N	N
15	Speed limit	35	50	50
16	Number of lanes	2	2	2
17	Road width	34, 40	32	38
18	General maintenance responsibility	State	State	State
19	Sidewalk	N	N	N
20	Bus stop (w/in 100')?	N	N	N
21	Channel?	N	N	N
22	Median?	N	N	N
23	Shoulder?	N	N	N

After scoring both the general data and the fatal data separately, then classifying scores as either low, moderate, or high for each year, results can be learned about fatal scores versus the total data scores. The figures below show the differences between total data each year and fatal data each year. It is evident that fatal crashes generally received higher scores throughout the three year period. In 2013, 23% of the total crashes yielded high scores, while 47% of the fatal crashes yieled high scores. This was a trend throughout the years: in 2014 only 12% of all crashes received high scores but 37% of fatal crashes received high scores and in 2015 18% of all crashes were high scoring while 39% of fatals were high scoring. This means a high number of fatal crashes shared many of the same accident traits, perhaps making these traits the most dangerous for pedestrians. If these traits can be monitored and changed, perhaps the number of fatal crashes in Delaware will decrease.

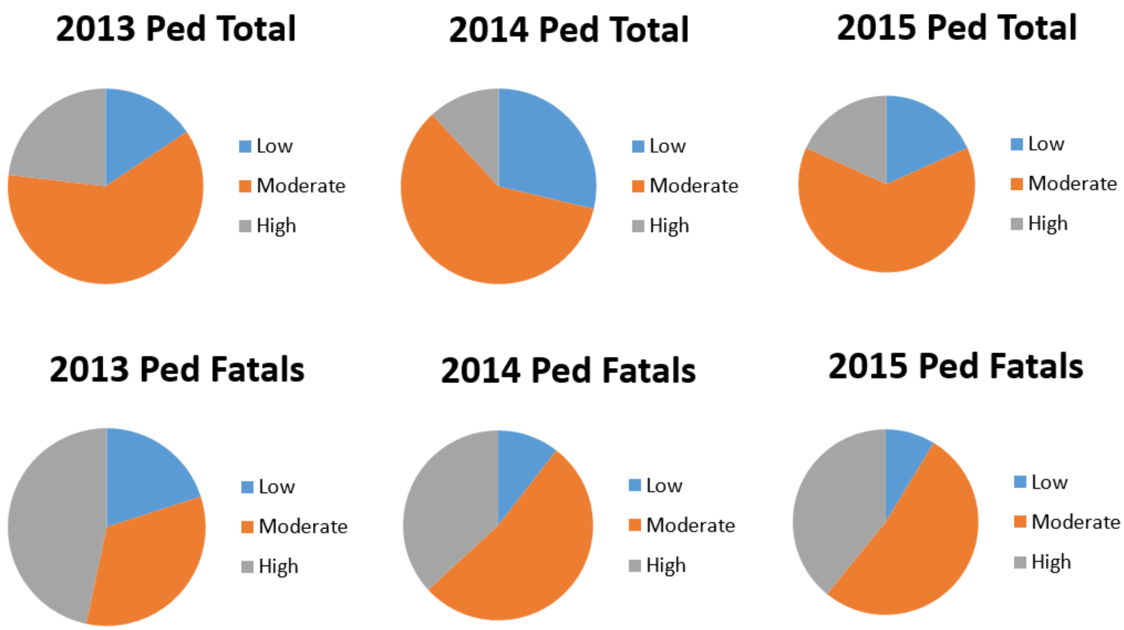


Figure 2 Pie charts show the distribution of low, moderate, and high scores for total data as well as fatal data for each year. Comparisons of scores can be made between total data and fatal data.

4.4 ArcMap Visualization

When the scores were plotted on the map of Delaware, it was evident that the most of the crashes occurred in higher populated areas, i.e. New Castle County, around the City of Wilmington, which can be seen in the map below. Crashes are distributed throughout the state with no obvious trend.

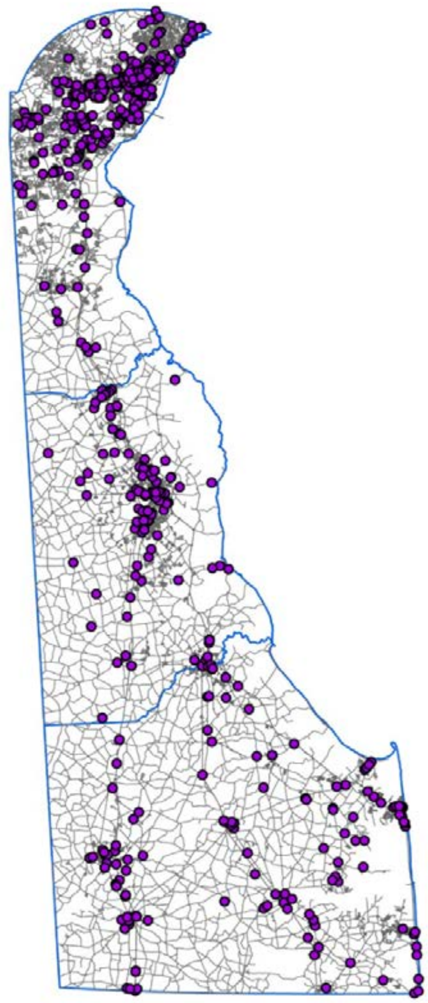


Figure 3 Map of Delaware showing all pedestrian related incidents at non-intersection locations from the year 2013-2015.

In the appendix there are three different maps that show the non intersection, scored pedestrian crashes per year. These maps show no significant difference in crash location from year to year. Crashes tend to be concentrated in the same places, despite a few outliers. The map below shows the locations of all fatal crashes from 2013 to

2015. Fatal crashes are also distributed throughout the state, but concentrated in the higher populated areas, i.e. around the City of Wilmington.

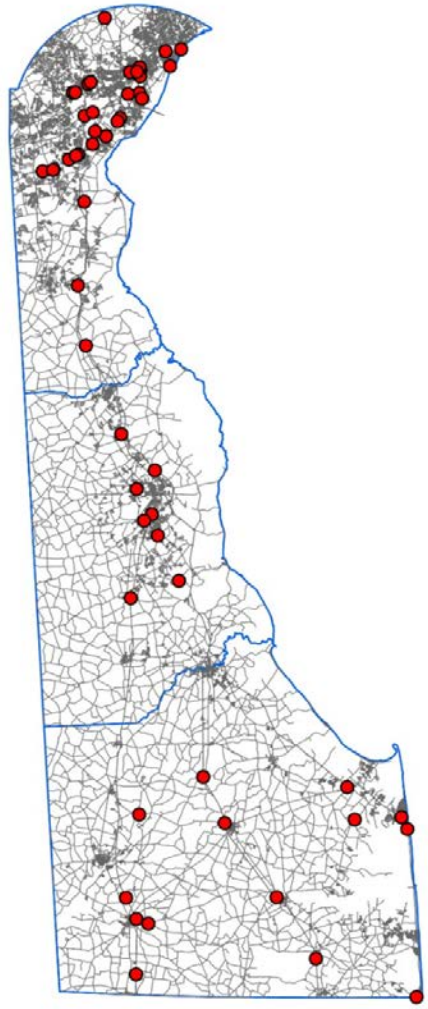


Figure 4 Map of Delaware showing all pedestrian related fatal incidents at non-intersection locations from the year 2013-2015.

Also, the different classifications of scores were plotted on separate maps to show locations of high scores, moderate scores, and low scores throughout the three year period. These maps show that crash locations, even with different scores, do not vary throughout the state. There were more moderately scored crashes than high or low, and therefore there are more yellow dots on the map, but they are all plotted in similar locations from year to year. The maps with the plotted scores may be evidence that crashes are occurring in higher populated areas but that there is no real trend between where crashes are occurring and the number of commonly shared traits. The map below shows the entire state with all scored crashes from 2013-2015.

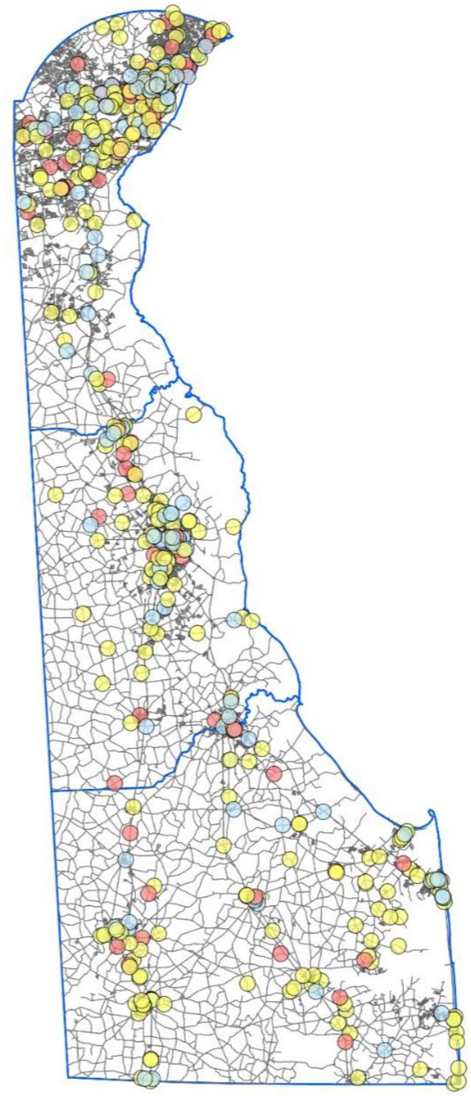


Figure 5 Map of Delaware showing all pedestrian related incidents at non-intersection locations from the year 2013-2015 as scored points. Blue dots indicate low scores, yellow dots indicate moderate scores, and red dots indicate high scores.

4.5 Genetic Based Visualization of Fatal Crashes

Due to the application of genetic analysis techniques in this research, an appropriate visual display of the data was created in Microstation to view all fatal crash data and roadway characteristics in a single gene-like visualization. Each fatal crash is labeled and shows if its traits were the most common for that year of fatal crash data. The white spaces mean that crash involved the most common trait for that incident characteristic. The visualizations present how the binary scores were used. If the crash trait was scored 1, it is a white space in the visualization below. If it scored a 0, it is black. So the figures make it easier to view which crashes traits were common and which were not widespread. Each "slice" of the circle is a crash, so if there are more black spaces in a section then that crash did not involve many of the common traits associated with fatal crashes that year. Below is an example of the genetic based maps showing fatal crash data from 2015. Maps from 2013 and 2014 can be seen in the appendix. The white spaces, each year, are concentrated around crash traits 8-13 which include peak hour, day or night, and various circumstances from the crash reports (all traits are numbered in Tables 1 and 2 in this thesis). The fact that there are white spaces concentrated around these traits means fatal crashes often involved these characteristics. On the contrary, we see more black spaces for crash traits 6 and 7, which are month and day of the week. That means although it was more common for a crash to occur during a certain month or day of the week during that year of data, it may not have been common enough to draw conclusions about that trait, for example, since there were many other times crashes occurred. Overall, this representation of the data helps to view similarities in a simple and distinct way.

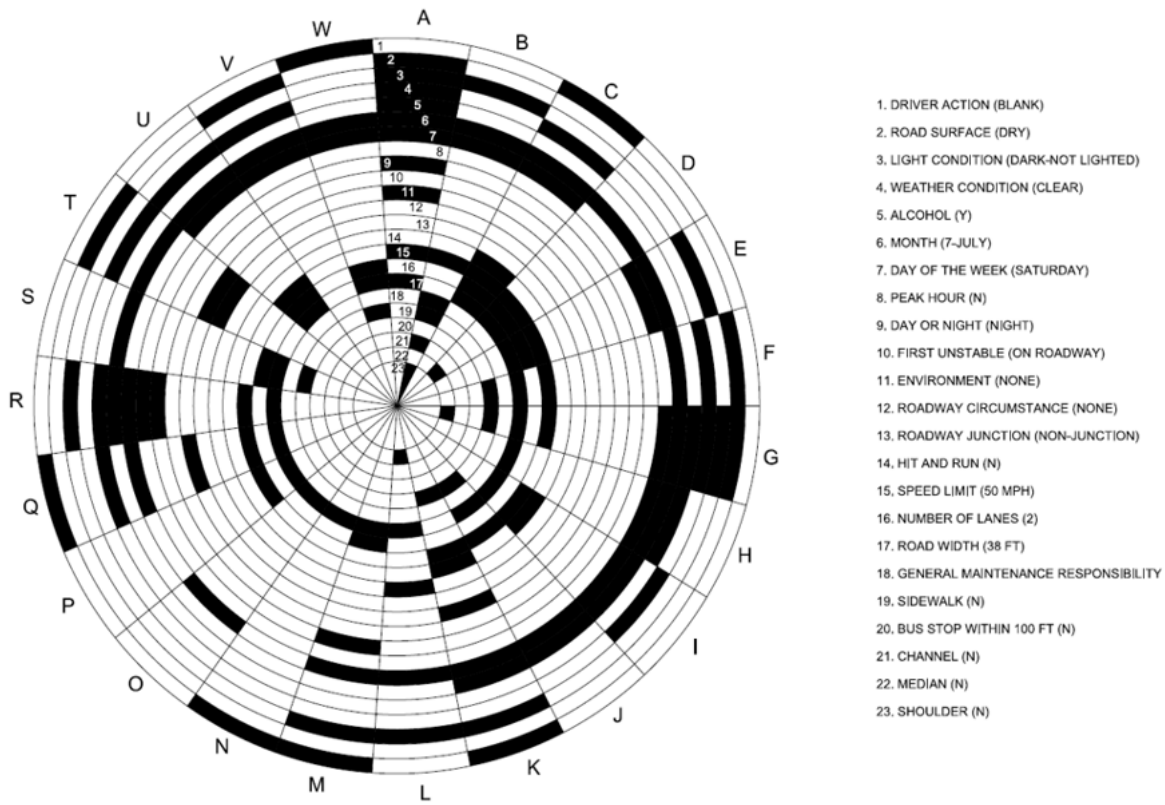


Figure 6 An example of a gene-like map for the fatal pedestrian crash data for the year 2015. Each white space represents the most common characteristic. Each crash is labeled with a letter and each crash trait is labeled with a number which corresponds to the list next to the display.

4.6 Limitations to Research

Like any research project, there were limitations to this study. An immense amount of data was given, and organizing it was a large task. With such a great amount of data, there is always the chance of error either in the way the data was originally coded, the way it was copied to another platform, or the organization of it in a compiled file. Many of the data fields had missing or incomplete information. This could be due to many factors: perhaps the crash with missing data did not possess certain traits, or the data was unknown for that specific crash, but it does cause a certain limit to the amount one can learn from this research. Some data fields had more blank cells than cells with actual information, so in that case the most common trait was “blank.” Also, for example, the road width field had such varying widths associated with each crash; it was difficult to draw conclusions about if the road width had any link to a crash occurrence or if it was too random a variable to tell. For this research, the most common attribute was taken, even if it was a “blank” trait or a number that only occurred a few times. If it was the most common, it was taken and studied, in case the way the information was recorded caused a pattern in crash trait trends.

Another limitation comes from the constraint of the data collection timing and crash incident reporting. The roadway characteristic information is up to the current date, so there is always the possibility that a crash that occurred in 2013 happened due to a condition that has since changed on the roadway. Crash data reporting in general is a concern since the reports come straight from the officers who were on the scene of the incident. If they recorded something wrong or not in accordance with other data, then the pedestrian crash data used in this study could be altered in a way that would alter conclusions. It was assumed that since this data is the most recent, changes in roadway geometry and conditions most likely have not changed enough for it to impact the study.

Personal confidentiality is a limitation associated with any crash data analysis. There is limited knowledge to what researchers can learn about the people involved in a crash incident. For this study, information about the driver/pedestrian age, gender, socio-economic status and other personal traits were not included or provided. If given this information perhaps more conclusions could be drawn about the occurrence of pedestrian related incidents in Delaware, but for the current study only attributes that were easily gathered were used. Although personal information may help to learn more about pedestrian safety, there is an ethical importance to keeping that information confidential.

One of the most limiting factors in any research having to do with safety and human behavior is the fact that people cannot be controlled, no matter the measures taken to attempt to keep them the safest as possible. Law enforcement, engineers, and planners can implement precautions and changes to roads but humans will always make mistakes, take an unmarked path, or ignore these measures, resulting in dangerous accidents. It is important to try and learn as much as possible about the way humans behave and the improvements that can be made on facilities, but there is only so much that can be done to prevent pedestrian-related crashes.

Chapter 5

CONCLUSION

The methods developed in this thesis generated informative results about non-intersection related pedestrian crashes in Delaware and the common traits they involve. Past literature shows evidence of the need for improvement in pedestrian education and facilities. Many studies have been completed that focus on pedestrian safety, crash data analysis, and various visualizations of traffic safety data and the results have been helpful in creating preventative measures. However none have used as comprehensive an approach as this thesis has done to learn about the causes of an increasing number of fatal crashes in the State of Delaware.

The development of the analysis outlined in this thesis involved consolidating DelDOT's and WILMAPCO's pedestrian crash data, spatially joining this data with road characteristics, evaluating this data to find the most common crash traits, scoring this data based on those traits and subsequently mapping these scored crashes in ArcGIS. Evaluating the most common road and crash characteristics made it clear what kind of conditions most of the crashes were occurring in the three years studied. Results from the total data varied when the fatal data was separated and certain crash traits were deemed more dangerous than others since they seemed to occur more often in fatal crashes than in the total dataset. Maps created in ArcMap along with the scores of each crash show the locations of varying levels of scores. No obvious trend was seen in the map data as to where crashes were occurring based on the level of the scores.

Visualization was also developed in Microstation that represents the binary scores assigned to each crash and the traits they involve. The visual is based on the concept of a genetic analysis and allows for results about the most and least common traits to be seen in a single representation.

The purpose of this study was to gain knowledge about the cause of such high numbers of fatal pedestrian crashes that have occurred in the State of Delaware and to develop a way to visualize these data trends in a new and unique way. The data analysis and scoring system was conclusive in that strong links were drawn from crash data records and road characteristics to the most common traits that occur in all pedestrian related non-intersection incidents. The first step to any improvement is figuring out the cause. With evidence that most fatal pedestrian-related incidents in the dataset occurred during the night in unlit conditions, often involving alcohol, the next steps to reducing the amount of fatal pedestrian crashes and alleviating these conditions can be planned. For example, developing campaigns surrounding alcohol use may be helpful. Drunk drivers are not the only issue when it comes to pedestrian incidents: many accidents occur when a pedestrian is under the influence and misusing roadway facilities. In recent years many pedestrian crashes occurred when the pedestrian was walking across the street after drinking too much alcohol. Although it is difficult to control or predict how a human behaves, especially when they are under an influence, educating them as much as possible about how to properly use pedestrian facilities or other means of transportation is a way to prevent these incidents from occurring. Another recommendation would be to increase lighting in dark areas where pedestrians may be walking at night and also encourage the use of reflective gear at night. The results of this data made it clear that making pedestrians as visible as possible during the night is important. Again, controlling and predicting human behavior is close to impossible, but knowing the conditions in which so many fatal pedestrian crashes have occurred is the first step to mitigating and decreasing the number of pedestrian fatalities.

Chapter 6

FUTURE WORK

Extensions of this research could be completed to enhance the comprehensive analysis of pedestrian crash data in Delaware outlined in this thesis. Since this study focused on pedestrian related crashes at non-intersection locations, a next step would be to involve intersection-related crashes and analyze data at these locations. It would require developing a different set of traits but scoring intersection crashes could increase understanding about pedestrian incidents in Delaware.

Similar to what much of the previous literature discussed, a statistical analysis approach could also be taken to learn more about the distributions of crashes and score data in order to view correlations between location, traits, and crash data. Much of the past research took a statistical approach which yielded helpful results in the corresponding studies. This data does involve a larger amount of data and variables than past research shows, so identifying correlations could be more difficult. However, extending this research to a more statistical understanding of the data may be helpful in order to learn more.

Crash rate and crash frequency would also be helpful in learning more about the pedestrian data used in this study. In order to do this sort of analysis, pedestrian volumes as well as vehicle volumes would have to be collected. A limitation to pedestrian safety analysis is the flexibility of pedestrian behavior. Unlike drivers, pedestrians do not need to use roads or specific paths: they can walk anywhere they'd like which can result in dangerous circumstances. It also makes it difficult to collect data about how they are moving. As much as can be learned about the amount of pedestrians and the way they move would be helpful. Crash density would also be helpful in linking crashes to population density in Delaware to find varying trends in the data.

Similar to Kim's research, an enhanced study into accident fault would be a solid addition to this research. Whether it is pedestrian or driver fault, perhaps learning more about this involvement can help with enforcing safety measures. If there is a trend in crash characteristics and who was at fault during these various crashes perhaps a more obvious solution could be learned and implemented.

REFERENCES

Alian, S., Baker, R., & Wood, S. (2016). Rural casualty crashes on the Kings highway: A new approach for road safety studies. *Accident Analysis & Prevention, 95*, 8-19.

Anderson, G. C., Craig; Davis, Kara. (2014). *Dangerous by Design 2014*.

Aylo, A. (2010). *Developing A GIS Safety Analysis System: A Case Study.* University of Dayton.

Beshah, T., & Hill, S. (2010). *Mining Road Traffic Accident Data to Improve Safety: Role of Road-Related Factors on Accident Severity in Ethiopia.* Paper presented at the AAAI Spring Symposium: Artificial Intelligence for Development.

Erdogan, S., Yilmaz, I., Baybura, T., & Gullu, M. (2008). Geographical information systems aided traffic accident analysis system case study: city of Afyonkarahisar. *Accident Analysis & Prevention, 40*(1), 174-181.

Hilton, B. N., Horan, T. A., Burkhard, R., & Schooley, B. (2011). SafeRoadMaps: Communication of location and density of traffic fatalities through spatial visualization and heat map analysis. *Information Visualization, 10*(1), 82-96.

Khanal, I. (2014). Development of a Visualization System for Highway Safety Management Using Safety Analyst.

Khattak, A. J., & Rouphail, N. (2005). *Incident Management Assistance Patrols: Assessment of Investment Benefits and Costs*. Retrieved from

Kim, K., Brunner, I. M., & Yamashita, E. (2008). Modeling fault among accident—Involved pedestrians and motorists in Hawaii. *Accident Analysis & Prevention, 40*(6), 2043-2049.

Markell, J. S., Lewis D.; Simpler, Jana (2014). *FY 2015 Highway Safety Plan*.

Markell, J. S., Lewis D.; Simpler, Jana. (2015). *FY2014 Annual Evaluation Report*.

NHTSA. (2016). *Traffic Safety Facts: 2014 Data*.

Plug, C., Xia, J. C., & Caulfield, C. (2011). Spatial and temporal visualisation techniques for crash analysis. *Accident Analysis & Prevention, 43*(6), 1937-1946.

Velloso, M. S., & Jacques, M. A. P. (2012). On-the-spot study of pedestrian crashes on Brazilian Federal District rural highways crossing urban areas. *Transportation Research Part F: Traffic Psychology and Behaviour, 15*(5), 588-599.

WILMAPCO. (2016). *New Castle County Intersection Crash Analysis 2005-2014*.

Appendix A

SUMMARY TABLES OF PEDESTRIAN CRASH TRAIT INFORMATION

2013				
Total pedestrian related crashes:	370		All Data	Fatal
Non-intersection related pedestrian crashes:	199	Minimum score:	8	8
Fatal pedestrian crashes at non-intersections:	15	Maximum score:	19	16

	All Data		Fatals	
Characteristic	Most common	Percentage	Most common	Percentage
Driver action	Other	33%	Failed to yield right of way	47%
Road surface	Dry	85%	Dry	80%
Light condition	Daylight	60%	Dark-Not Lighted	60%
Weather condition	Clear	72%	Clear	60%
Alcohol?	N	88%	Y	53%
Month	August	13%	no max	
Day of the week	Friday	19%	Thursday	27%
Peak hour?	N	74%	N	93%
Day or night?	Day	62%	Night	87%
First unstable	On Roadway	88%	On Roadway	100%
Environment	None	83%	None	87%
Roadway circumstance	None	84%	None	73%
Roadway junction	Non-Junction	75%	Non-Junction	87%
Hit and run?	N	77%	N	80%
Speed limit	25	53%	35	27%
Number of lanes	2	84%	2	73%
Road width	26, 38	19%	34, 40	27%
General maintenance responsibility	State	70%	State	100%
Sidewalk	Y	58%	N	53%
Bus stop (w/in 100')?	N	90%	N	100%
Channel?	N	99%	N	100%
Median?	N	94%	N	93%
Shoulder?	N	88%	N	100%

2014				
Total pedestrian related crashes:	388		All Data	Fatal
Non-intersection related pedestrian crashes:	202	Minimum score:	6	13
Fatal pedestrian crashes at non-intersections:	19	Maximum score:	20	20
	All Data		Fatals	
Characteristic	Most common	Percentage	Most common	Percentage
Driver action	Blank	38%	Blank	58%
Road surface	Dry	80%	Dry	79%
Light condition	Daylight	56%	Dark-Not Lighted	47%
Weather condition	Clear	74%	Clear	68%
Alcohol?	N	88%	N	58%
Month	July	12%	December	26%
Day of the week	Friday	22%	Saturday	26%
Peak hour?	N	81%	N	100%
Day or night?	Day	56%	Night	79%
First unstable	On Roadway	71%	On Roadway	95%
Environment	None	80%	None	95%
Roadway circumstance	None	86%	None	95%
Roadway junction	Non-Junction	70%	Non-Junction	89%
Hit and run?	N	62%	N	95%
Speed limit	25	49%	50	37%
Number of lanes	2	71%	2	68%
Road width	26	12%	32	16%
General maintenance responsibility	State	59%	State	89%
Sidewalk	N	50%	N	63%
Bus stop (w/in 100')?	N	94%	N	79%
Channel?	N	98%	N	95%
Median?	N	89%	N	95%
Shoulder?	N	88%	N	89%

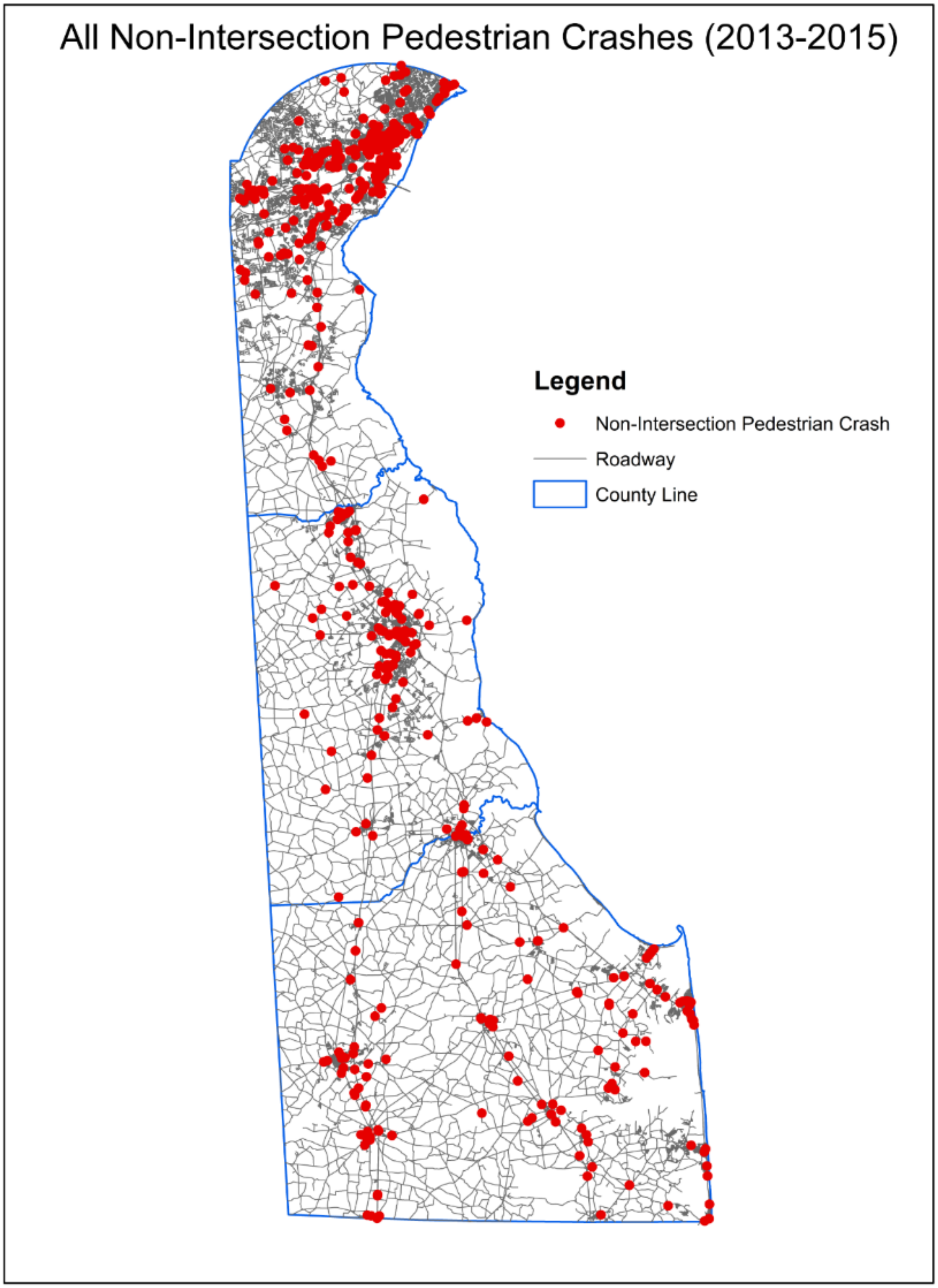
2015				
Total pedestrian related crashes:	376		All Data	Fatal
Non-intersection related pedestrian crashes:	186	Minimum score:	9	13
Fatal pedestrian crashes at non-intersections:	23	Maximum score:	21	21
	All Data		Fatals	
Characteristic	Most common	Percentage	Most common	Percentage
Driver action	Blank	47%	Blank	74%
Road surface	Dry	82%	Dry	78%
Light condition	Daylight	47%	Dark-Not Lighted	65%
Weather condition	Clear	72%	Clear	61%
Alcohol?	N	84%	Y	78%
Month	July	13%	July	26%
Day of the week	Saturday	18%	Saturday	26%
Peak hour?	N	82%	N	96%
Day or night?	Night	53%	Night	87%
First unstable	On Roadway	88%	On Roadway	100%
Environment	None	77%	None	83%
Roadway circumstance	None	83%	None	96%
Roadway junction	Non-Junction	81%	Non-Junction	96%
Hit and run?	N	70%	N	87%
Speed limit	25	51%	50	39%
Number of lanes	2	85%	2	70%
Road width	26	13%	38	17%
General	State	70%	State	96%
Sidewalk	Y	60%	N	70%
Bus stop (w/in 100')?	N	93%	N	100%
Channel?	N	98%	N	96%
Median?	N	92%	N	87%
Shoulder?	N	87%	N	96%

ALL DATA				
Total pedestrian related crashes:	1134		All Data	Fatal
Non-intersection related pedestrian crashes:	587	Minimum score:	7	9
Fatal pedestrian crashes at non-intersections:	57	Maximum score:	20	21

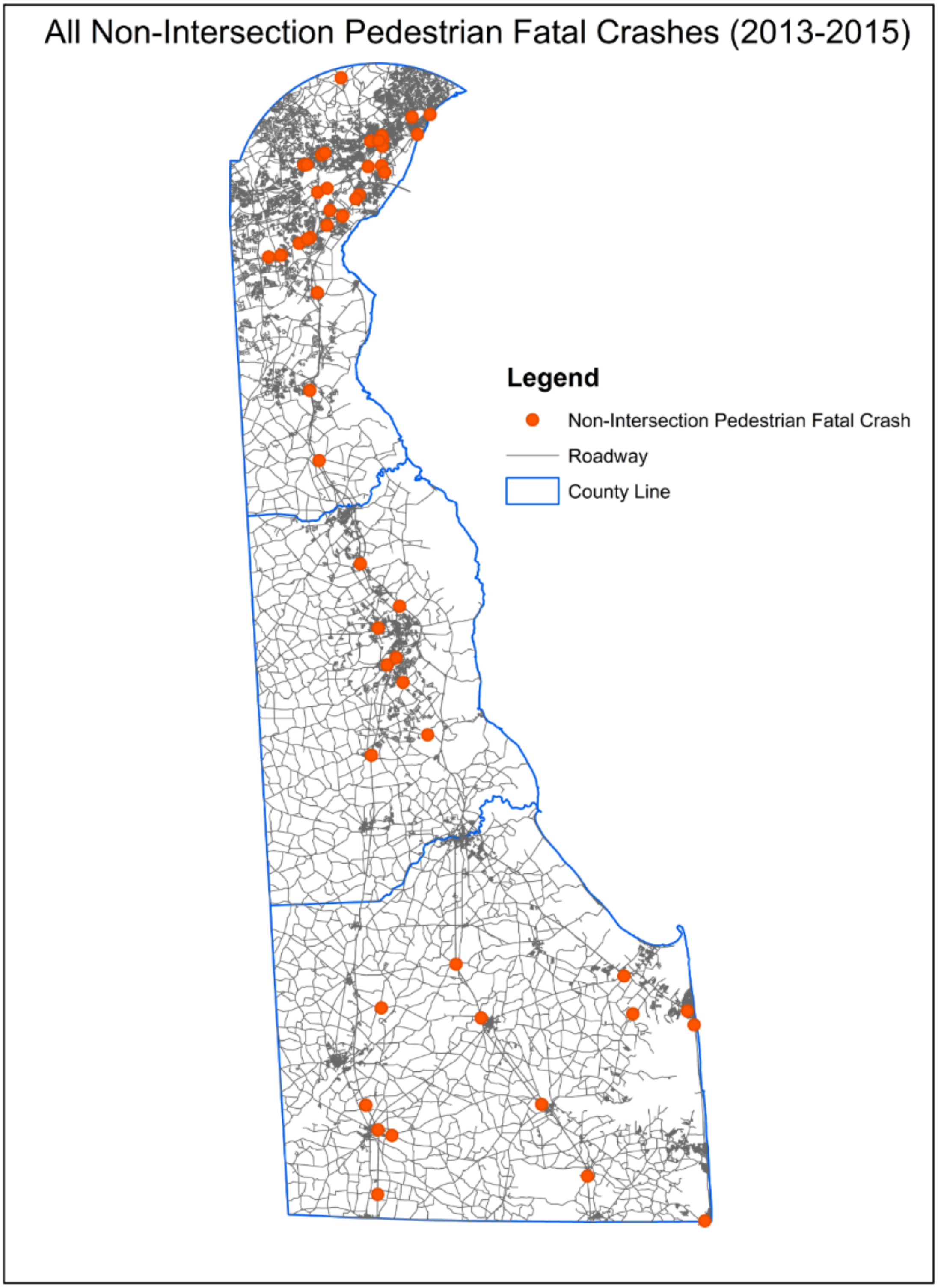
	All Data		Fatals	
Characteristic	Most common	Percentage	Most common	Percentage
Driver action	Blank	34%	Blank	58%
Road surface	Dry	82%	Dry	79%
Light condition	Daylight	55%	Dark-Not Lighted	58%
Weather condition	Clear	73%	Clear	63%
Alcohol?	N	87%	Y	60%
Month	July	12%	July	14%
Day of the week	Friday	19%	Saturday	25%
Peak hour?	N	79%	N	96%
Day or night?	Day	56%	Night	84%
First unstable	On Roadway	82%	On Roadway	98%
Environment	None	80%	None	88%
Roadway circumstance	None	85%	None	89%
Roadway junction	Non-Junction	75%	Non-Junction	91%
Hit and run?	N	70%	N	88%
Speed limit	25	51%	50	32%
Number of lanes	2	80%	2	70%
Road width	26	12%	38	14%
General maintenance responsibility	State	66%	State	95%
Sidewalk	Y	56%	N	63%
Bus stop (w/in 100')?	N	92%	N	93%
Channel?	N	98%	N	96%
Median?	N	92%	N	91%
Shoulder?	N	88%	N	95%

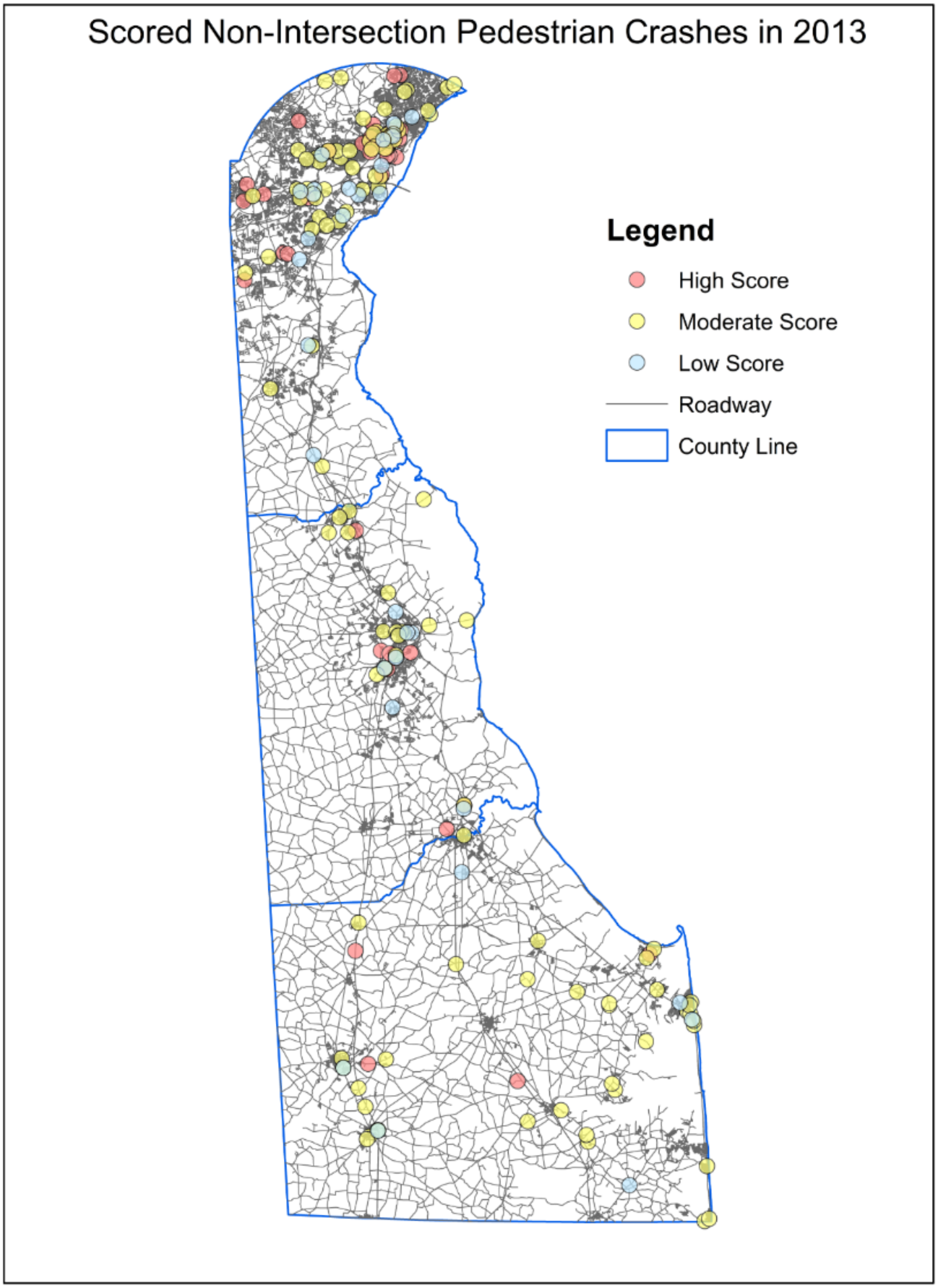
Appendix B

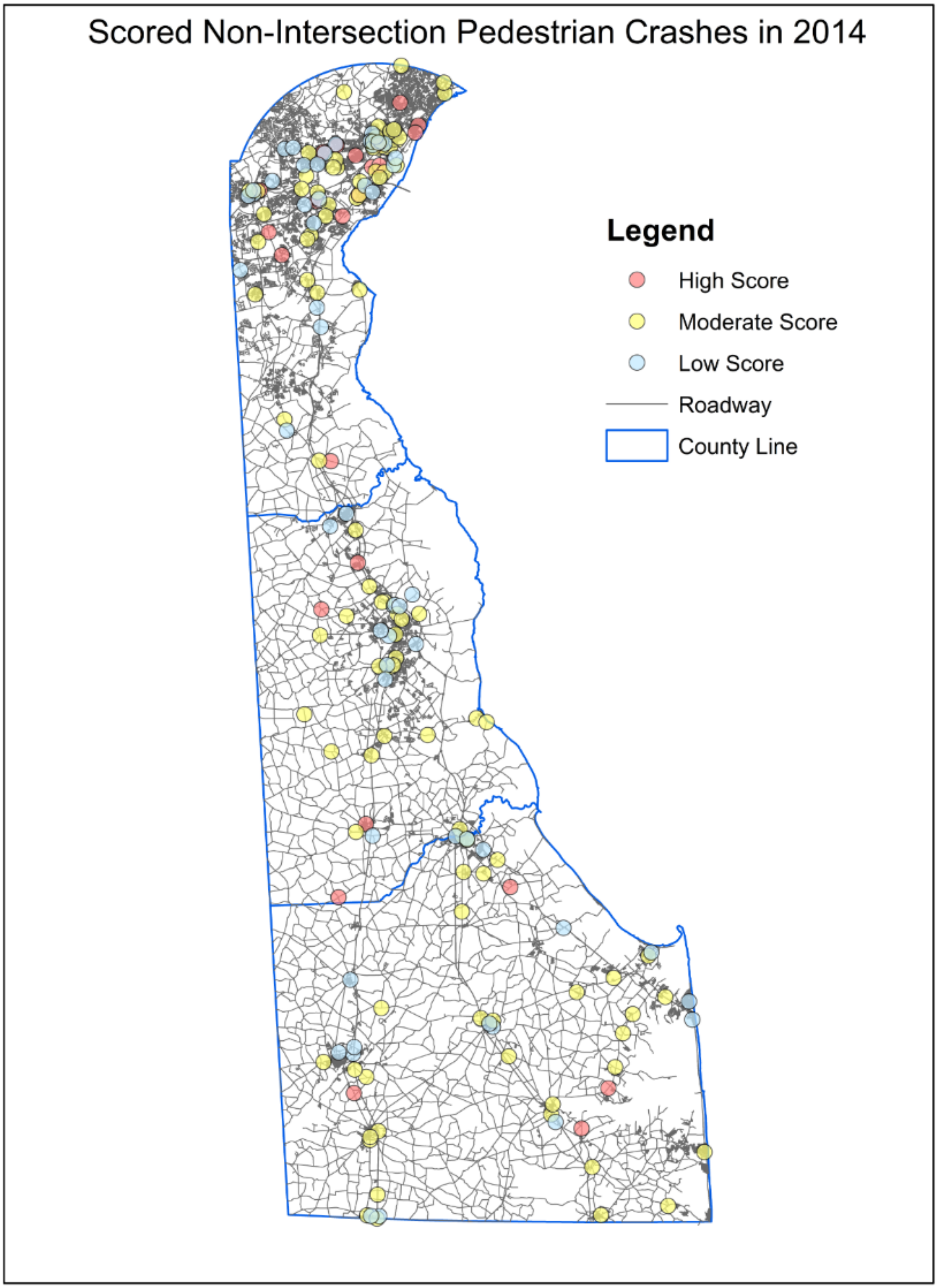
ARCMAP VISUALIZATIONS OF PEDESTRIAN CRASH DATA

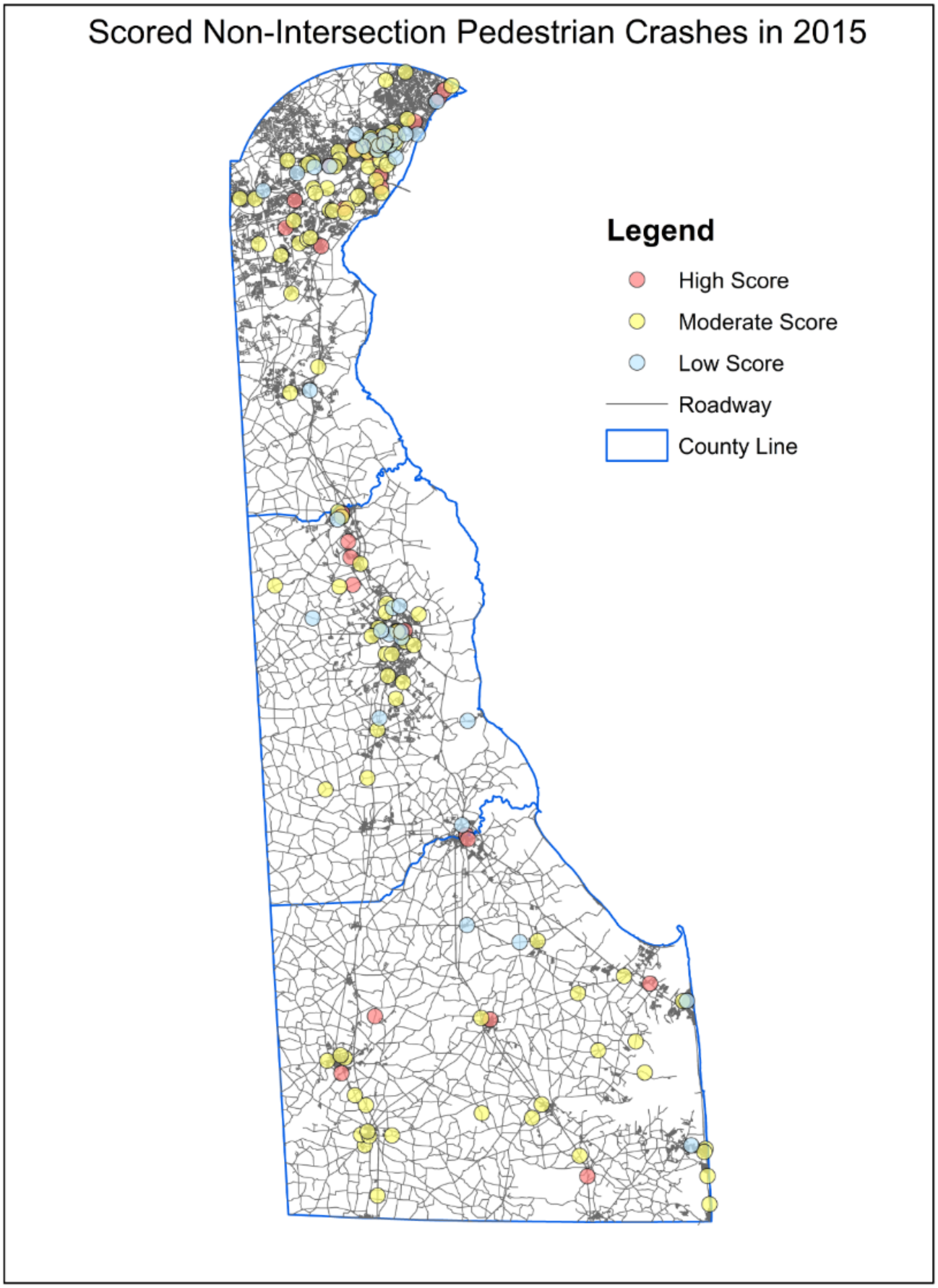


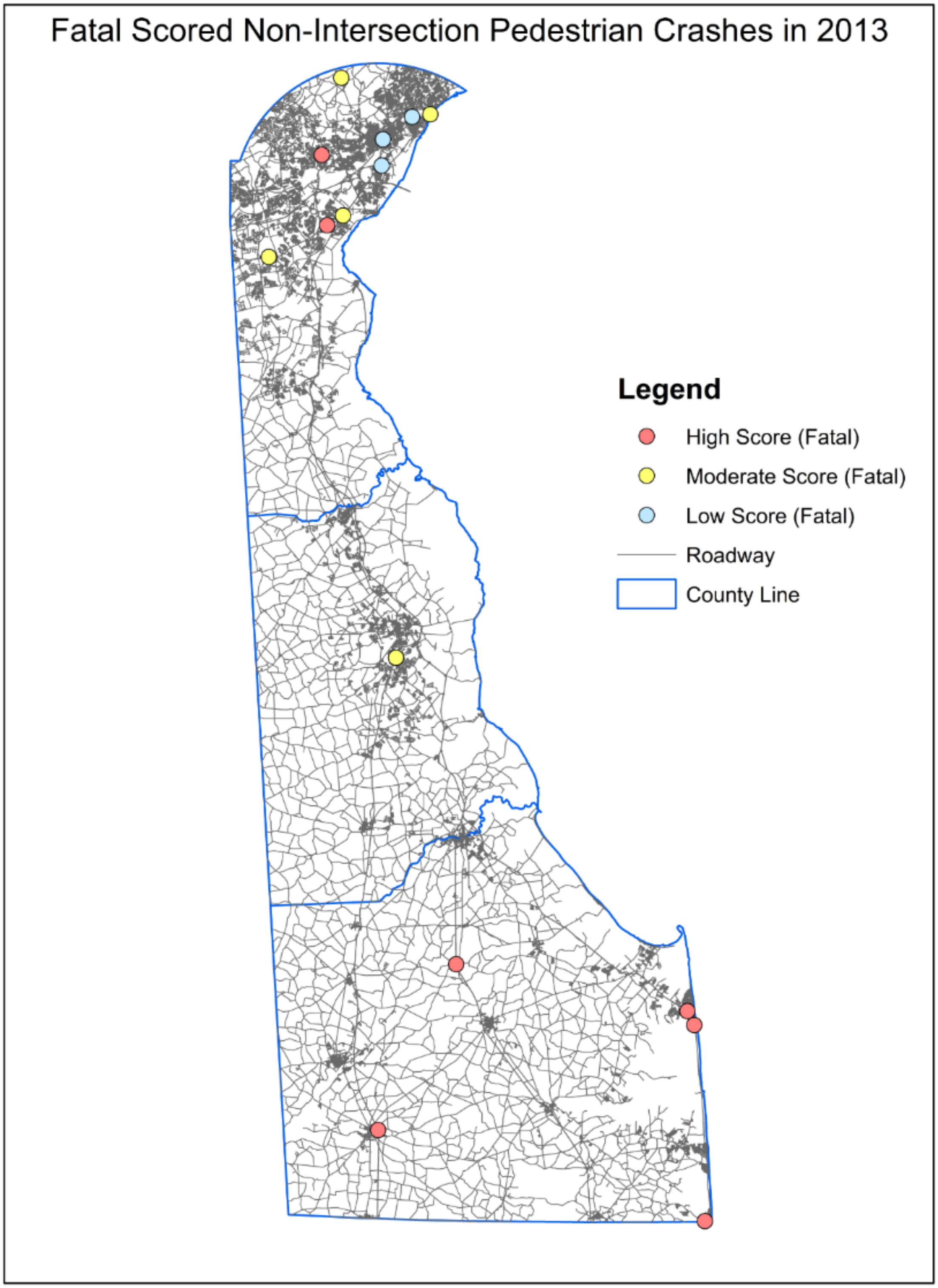
All Non-Intersection Pedestrian Crashes (2013-2015)
Legend
Non-Intersection Pedestrian Crash
Roadway
County Line

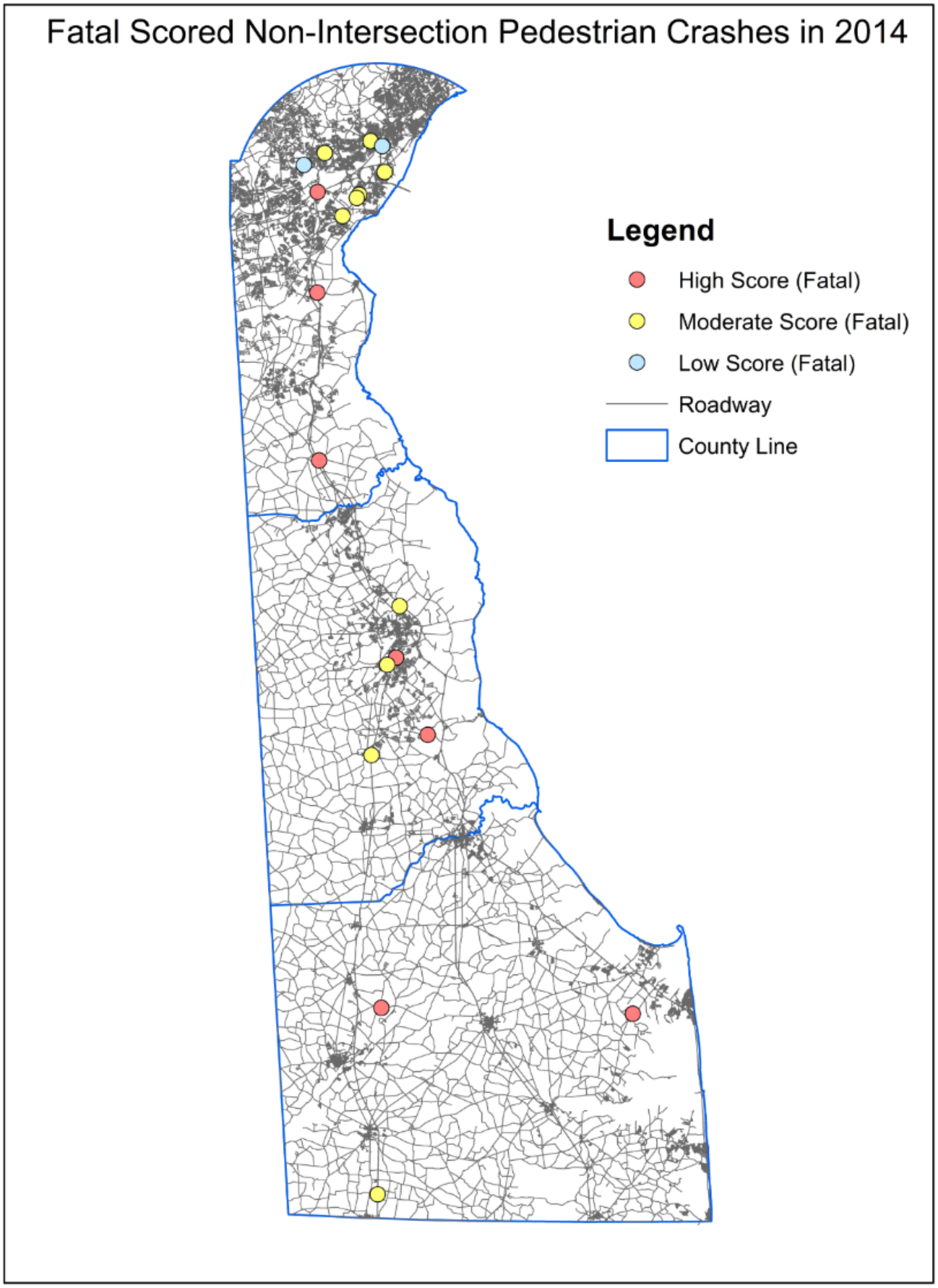

All Non-Intersection Pedestrian Fatal Crashes (2013-2015)
Legend
Non-Intersection Pedestrian Fatal Crash
Roadway
County Line

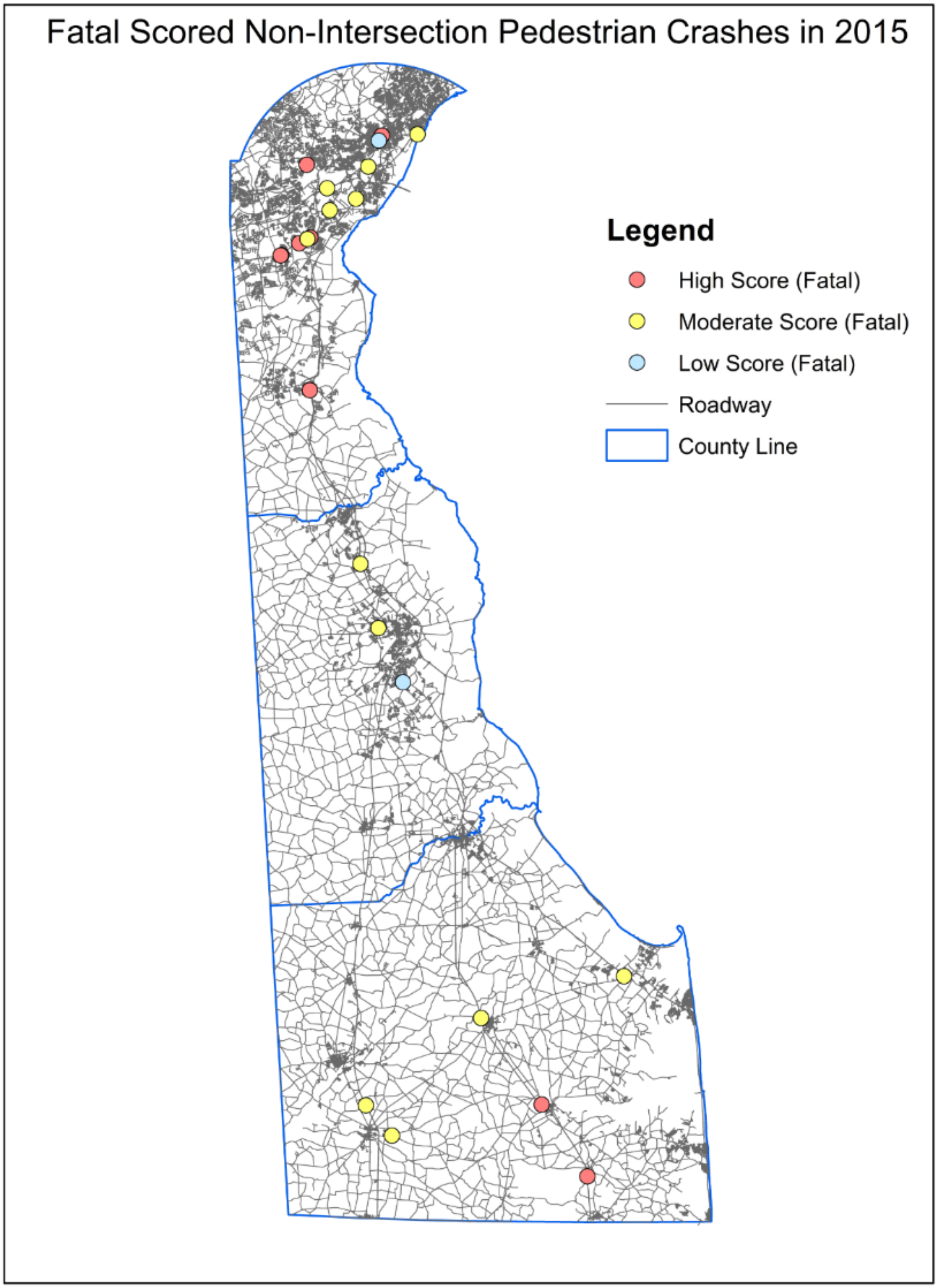

Scored Non-Intersection Pedestrian Crashes in 2013
Legend
High Score
Moderate Score
Low Score
Roadway
County Line


Scored Non-Intersection Pedestrian Crashes in 2014
Legend
High Score
Moderate Score
Low Score
Roadway
County Line


Scored Non-Intersection Pedestrian Crashes in 2015
Legend
High Score
Moderate Score
Low Score
Roadway
County Line


Fatal Scored Non-Intersection Pedestrian Crashes in 2013
Legend
High Score (Fatal)
Moderate Score (Fatal)
Low Score (Fatal)
Roadway
County Line

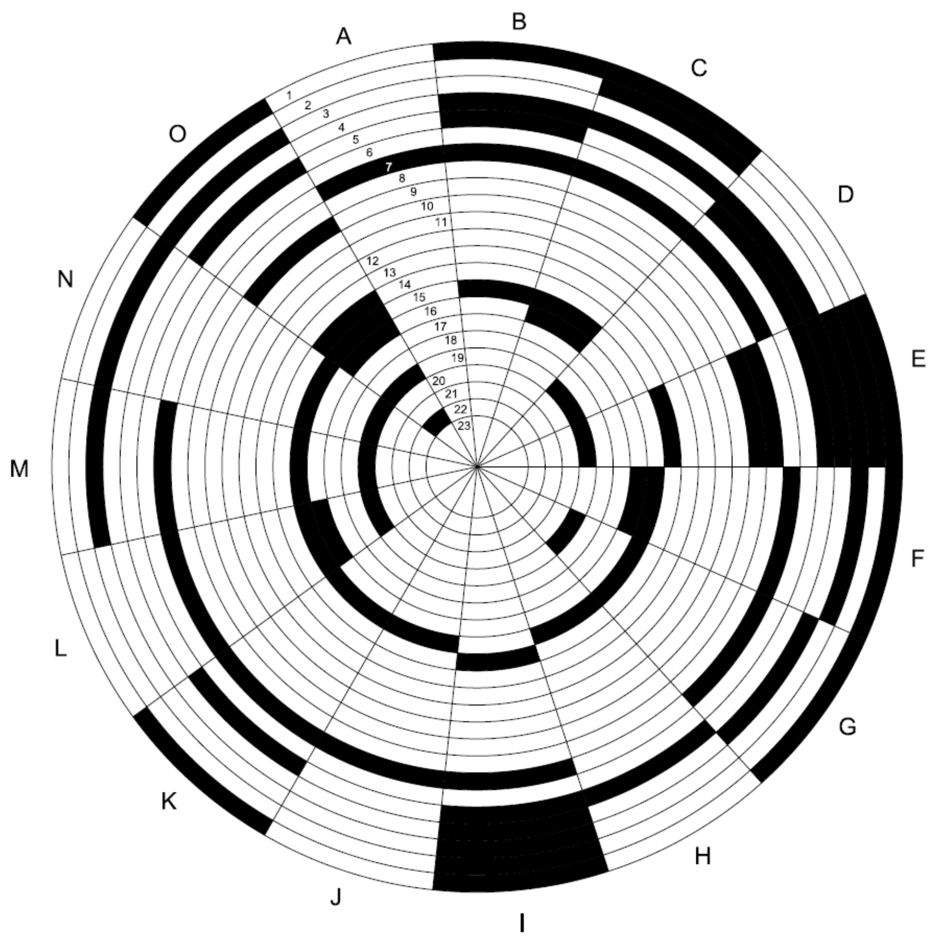

Fatal Scored Non-Intersection Pedestrian Crashes in 2014
Legend
High Score (Fatal)
Moderate Score (Fatal)
Low Score (Fatal)
Roadway
County Line


Fatal Scored Non-Intersection Pedestrian Crashes in 2015
Legend
High Score (Fatal)
Moderate Score (Fatal)
Low Score (Fatal)
Roadway
County Line

Appendix C

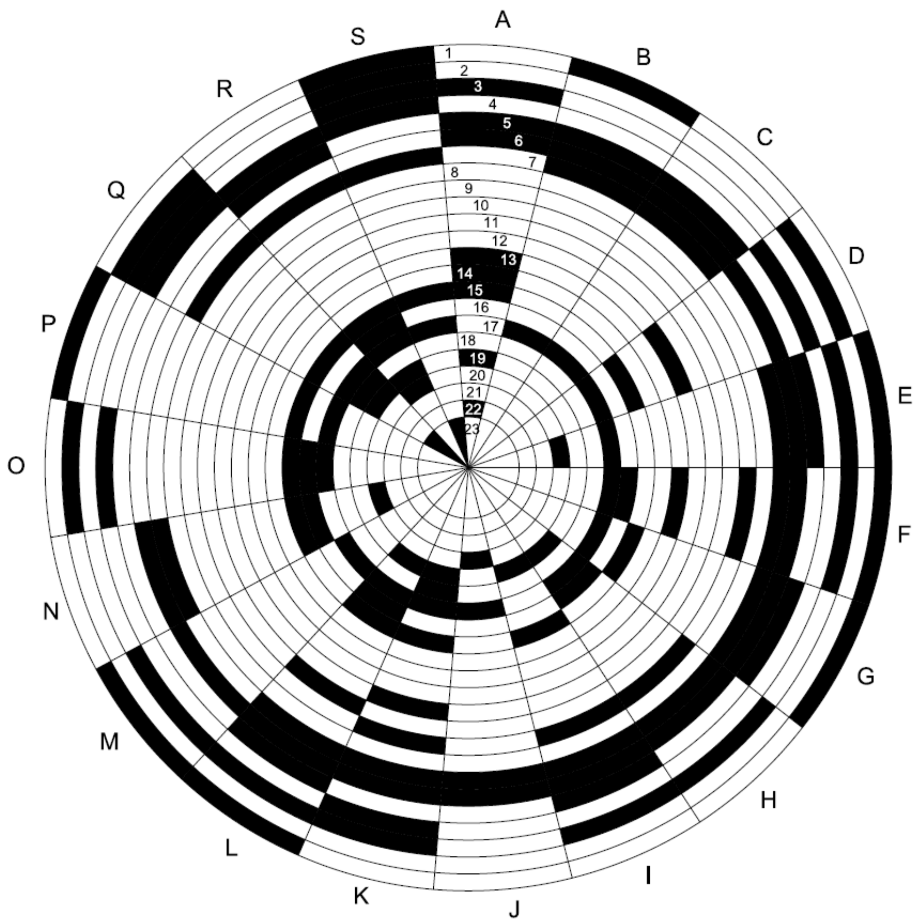
MICROSTATION VISUALIZATION OF PEDESTRIAN FATALITY DATA

2013 Pedestrian Fatality Data



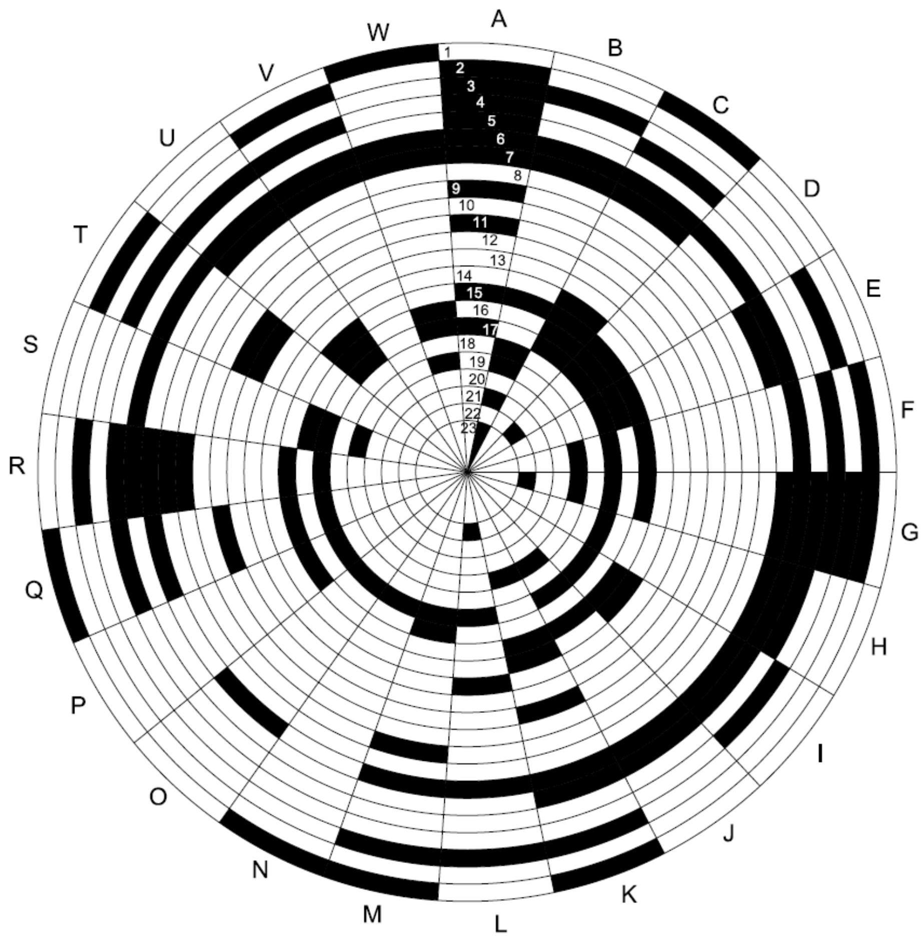
1. DRIVER ACTION (FAILED TO YIELD RIGHT OF WAY)
2. ROAD SURFACE (DRY)
3. LIGHT CONDITION (DARK-NOT LIGHTED)
4. WEATHER CONDITION (CLEAR)
5. ALCOHOL (Y)
6. MONTH (NO MAX)
7. DAY OF THE WEEK (THURSDAY)
8. PEAK HOUR (N)
9. DAY OR NIGHT (NIGHT)
10. FIRST UNSTABLE (ON ROADWAY)
11. ENVIRONMENT (NONE)
12. ROADWAY CIRCUMSTANCE (NONE)
13. ROADWAY JUNCTION (NON-JUNCTION)
14. HIT AND RUN (N)
15. SPEED LIMIT (35 MPH)
16. NUMBER OF LANES (2)
17. ROAD WIDTH (MAX IS BLANK)
18. GENERAL MAINTENANCE RESPONSIBILITY (STATE)
19. SIDEWALK (N)
20. BUS STOP WITHIN 100 FT (N)
21. CHANNEL (N)
22. MEDIAN (N)
23. SHOULDER (N)

2014 Pedestrian Fatality Data



1. DRIVER ACTION (BLANK)
2. ROAD SURFACE (DRY)
3. LIGHT CONDITION (DARK-NOT LIGHTED)
4. WEATHER CONDITION (CLEAR)
5. ALCOHOL (N)
6. MONTH (12- DECEMBER)
7. DAY OF THE WEEK (SATURDAY)
8. PEAK HOUR (N)
9. DAY OR NIGHT (NIGHT)
10. FIRST UNSTABLE (ON ROADWAY)
11. ENVIRONMENT (NONE)
12. ROADWAY CIRCUMSTANCE (NONE)
13. ROADWAY JUNCTION (NON-JUNCTION)
14. HIT AND RUN (N)
15. SPEED LIMIT (50 MPH)
16. NUMBER OF LANES (2)
17. ROAD WIDTH (32 FT)
18. GENERAL MAINTENANCE RESPONSIBILITY (STATE)
19. SIDEWALK (N)
20. BUS STOP WITHIN 100 FT (N)
21. CHANNEL (N)
22. MEDIAN (N)
23. SHOULDER (N)

Pedestrian Fatality Data 2015



1. DRIVER ACTION (BLANK)
2. ROAD SURFACE (DRY)
3. LIGHT CONDITION (DARK-NOT LIGHTED)
4, WEATHER CONDITION (CLEAR)
5. ALCOHOL (Y)
6. MONTH (7-JULY)
7. DAY OF THE WEEK (SATURDAY)
8. PEAK HOUR (N)
9, DAY OR NIGHT (NIGHT)
10. FIRST UNSTABLE (ON ROADWAY)
11. ENVIRONMENT (NONE)
12. ROADWAY CIRCUMSTANCE (NONE)
13. ROADWAY JUNCTION (NON-JUNCTION)
14, HIT AND RUN (N)
15, SPEED LIMIT (50 MPH)
16. NUMBER OF LANES (2)
17. ROAD WIDTH (38 FT)
18. GENERAL MAINTENANCE RESPONSIBILITY (STATE)
19, SIDEWALK (N)
20, BUS STOP WITHIN 100 FT (N)
21. CHANNEL (N)
22. MEDIAN (N)
23. SHOULDER (N)