RISK FACTORS FOR PEDESTRIAN AND BICYCLE CRASHES

FINAL REPORT

SPR 779

Oregon Department of Transportation

RISK FACTORS FOR PEDESTRIAN AND BICYCLE CRASHES

Final Report

SPR 779

by

Chris Monsere, Ph.D., P.E., Associate Professor
Yi Wang, Graduate Research Assistant
Portland State University
Department of Civil and Environmental Engineering
PO Box 751
Portland, OR 97207

Haizhong Wang, Ph.D., Assistant Professor
Chen Chen, Graduate Research Assistant
Oregon State University
School of Civil and Construction Engineering
101 Kearney Hall
Corvallis, OR 97331

for

Oregon Department of Transportation
Research Section
555 13th St. NE
Salem, OR 97301

May 2017

1. Report No. FHWA-OR-RD-17-13	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Risk Factors for Pedestrian and Bicycle Crashes		5. Report Date May 2017 6. Performing Organization Code
7. Author(s) Chris Monsere, Haizhong Wang, Yi Wang, Chen Chen		8. Performing Organization Report No. SPR 779
9. Performing Organization Name and Address Oregon Department of Transportation Research Section 555 13th Street NE Salem, OR 97301		10. Work Unit No. (TRAIS) 11. Contract or Grant No.
12. Sponsoring Agency Name and Address Oregon Department of Transportation Federal Highway Administration Research Section and 400 Seventh Street, SW 555 13th Street NE Washington, D.C. 20590-0003 Salem, OR 97301		13. Type of Report and Period Covered Final Report 14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

The primary goal of this research was to develop a tool for the Oregon Department of Transportation (ODOT) to improve methods to identify and prioritize locations with increased or elevated risk for pedestrian and bicycle crashes. This report includes a comprehensive review of many scientific reports and papers about the pedestrian and bicycle crashes on road segments or intersections. To develop the risk model data were collected from 188 segments and 184 intersections randomly selected following the data collection plan. . These samples included 213 bicycle and pedestrian crashes on the segments and 238 at intersections. Geometric, land use, volume, and crash data were collected from different databases, including Google Maps, EPA's Smart Location Database and the ODOT crash database from 2009-2013. The research team developed logistic regression models for both crash occurrence (crash or not) and crash severity models. The models related to crash severity were not robust, most likely due to the few segments and intersections with severe crashes in the dataset. The crash occurrence models were used to create a risk-scoring tool. The risk-scoring tool was applied to safety projects identified in the 2015 All Roads Transportation Safety (ARTS) project lists from Oregon DOT's Region 1 and 2. The risk scores for the case study applications aligned reasonably well with the project's benefit-costs estimates.

17. Key Words Pedestrian safety; bicycle safety; risk analysis		18. Distribution Statement Copies available from NTIS, and online at http://oregon.gov/ODOT/Programs/Pages/Research-Publications.aspx	
19. Security Classification (of this report) Unclassified	20. Security Classification (of this page) Unclassified	21. No. of Pages 99	22. Price

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

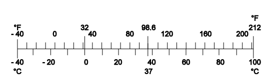
SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	millimeters squared	mm^2
ft^2	square feet	0.093	meters squared	m^2
yd^2	square yards	0.836	meters squared	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	kilometers squared	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	meters cubed	m^3
yd^3	cubic yards	0.765	meters cubed	m^3
NOTE: Volumes greater than 1000 L shall be shown in m^3.				
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams	Mg
		TEMPERATURE (exact)		
°F	Fahrenheit temperature	5(F-32)/9	Celsius temperature	°C

APPROXIMATE CONVERSIONS FROM SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	millimeters squared	0.0016	square inches	in^2
m^2	meters squared	10.764	square feet	ft^2
ha	hectares	2.47	acres	ac
km^2	kilometers squared	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	meters cubed	35.315	cubic feet	ft^3
m^3	meters cubed	1.308	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.205	pounds	lb
Mg	megagrams	1.102	short tons (2000 lb)	T
		TEMPERATURE (exact)		
°C	Celsius temperature	1.8 + 32	Fahrenheit	°F



* SI is the symbol for the International System of Measurement

(4-7-94 jbp)

ACKNOWLEDGEMENTS

The authors would like to thank the members of the Technical Advisory Committee for their help in the preparation of this report and the research. The TAC members consisted of Roger Gutierrez, ODOT Traffic-Roadway Section; Zahidul Siddique, ODOT Traffic-Roadway Section; Amanda Salyer, Senior Traffic Investigator, ODOT Technical Services; Gary Obery, ODOT Alternate Mode Traffic Engineer and Kristie Gladhill, Region 1 Transportation Safety Coordinator. The project was managed by Mark Joerger, ODOT Research.

DISCLAIMER

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

In 2013, there were 52 pedestrian and three bicyclist fatalities – approximately 17% of the total traffic fatalities in Oregon (Bergh and Griffin, 2013). There were an additional 814 pedestrians and 922 bicyclists injured in that same year. These fatalities and injuries are a concern for many communities in Oregon. To help mitigate the high social cost of these events, the Oregon Department of Transportation (ODOT) has identified pedestrian and bicycle crashes as a primary focus area for investing infrastructure funding. ODOT has appropriated approximately $4 million annually in the All Roads Transportation Safety Program (ARTS) to help address this key need. Selecting projects based only on crash performance is challenging because pedestrian and bicycle crashes are few. Predicting where these crashes will occur next is also a challenging task. An alternative to frequency-based selection is to develop a risk-based criteria and method.

In a transportation context, risk is defined as a probability or threat of damage, injury, liability, loss, or any other negative occurrence that is caused by external or internal vulnerabilities, and that may be avoided through preemptive action. The amount of risk can be interpreted by the probability of the outcome and potential severity of the outcome if the event occurs. This can be presented as the risk matrix in Figure 1.1 (Berdica, 2002). For transportation and vulnerable road users, the probability is a function of exposure and consequence is a function of operating conditions (e.g., vehicle speeds and size). Risk scoring should include elements of exposure and expectations of the severity of the outcome.

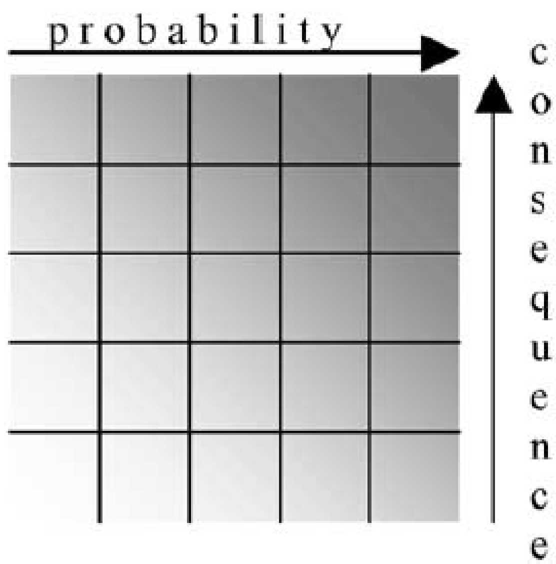


Figure 1.1: Risk Matrix, Darker Shading Indicates Higher Risk (*Berdica, 2002*)

Prior to this research project, Kittelson & Associates, Inc. produced a systemic safety analysis of pedestrian and bicycle safety for ODOT (Bergh et al., 2013). This thorough and detailed report analyzed pedestrian and bicycle crashes following a systemic safety planning approach. The objective of the plan was to "identify corridors with the most potential for reducing frequency and severity of pedestrian and bicycle crashes." The report attempted to match effective safety countermeasures with potential locations for improvements by identifying key patterns of behavior and roadway conditions that create high-risk locations. The results of the plan were presented and discussed with stakeholders from around Oregon. Prioritized lists of corridors, both on state highways and non-state roads, were developed. A frequency-based and risk-based prioritization was developed. The quantification of risk factors and the magnitude of their influence was constrained by limited supplemental data that the project was able to collect. Thus, many of the risk scores were based on best judgment.

The objective of this research was to develop a risk-scoring method with weights derived from a data analysis. To accomplish this, the research assembled a robust dataset merging data elements collected on segments, at intersections. Selection of segments for data collection was random and no attempt was made to identify high-crash locations. Logistic models were developed and the results transformed to a simple tool. Application of the tool to actual ARTS projects was done to demonstrate the applicability.

This final report summarizes the research and is organized in six chapters. Chapter 2.0 presents a brief literature review. Chapter 3.0 describes the process used to identify the data collection methods used to assemble the data. Chapter 4.0 reviews the basic descriptive analysis of the data and the preparation of the modeling dataset. Chapter 5.0 then presents the modeling methodology, and Chapter 6.0 summarizes the findings. In Chapter 7.0, the conversion of the modeling results to a risk-scoring tool is presented, and sample applications are presented in Chapter 8.0. Finally, Chapter 9.0 summarizes the work and discusses the limitations and recommendations for future work.

2.0 LITERATURE REVIEW

This literature review is organized into three sections. First, brief examples of risk-based approaches to network screening for highway safety are presented with examples of methods to prioritize or score non-motorized projects. The following sections provide a detailed review of the factors affecting bicycle and pedestrian safety.

2.1 RISK-BASED METHODS FOR PRIORITIZING SAFETY

Most network screening techniques that are used to identify high-crash locations are crash-based (AASHTO, 2010). The risk-based approaches that are the focus of this literature review (using the potential for crash outcomes combined with severity) are relatively rare in the crash screening methods (especially for motor vehicles). This is primarily due to fact that there are few reported bicycle or pedestrian crashes and there is insufficient exposure data to identify trouble spots. In this section, the systemic safety approach is presented, followed by a brief review of risk-based scoring methods for highways. Finally, since the perceptions of safety and actual safety can be distinctly different, safety prioritization measures for non-motorized safety that help include this factor are presented.

2.1.1 Systemic Safety Approach

Recently, in recognition that many approaches to safety improvement were generally only reactive (i.e., identifying the locations with high crash frequency), the Federal Highway Administration (FHWA) has been promoting an approach to highways safety screening and countermeasure application that is termed “systemic safety.” The systemic safety approach, illustrated in Figure 2.1 taken from *Systemic Safety Project Selection Tool*, can be described as a methodology to identify, diagnose and treat locations that are at high risk for crashes on a system-wide basis through a data-driven approach (Preston et al., 2013). The figure shows the three key elements for the overall program.

Figure 2.1: Systemic Safety Approach

Within Element 1, the steps most relevant to this review are the steps to identify crash types and risk factors. After first finding the facility and crash type to focus on (usually using a crash-tree diagram that shows the percentage of crashes, both total and severe, in various categories of roads/intersections), a more detailed analysis can reveal some measurable characteristics of the facility that may be subject to treatment. For example, after first identifying rural, undivided, two-lane roadways as a focus area, the NYSDOT further compared crashes on curves by radii. As shown in Figure 2.2, they compared curve radii in four categories. When compared with all curves along the focus facility type, overrepresented features become clear. While only representing 7% of the curves reviewed, curves with a curve radius less than 300 feet accounted for 12% of severe crashes. Thus, curve radii less than 300 feet were identified as a risk factor. Through further analysis, the factors were assigned scores that could be combined in an overall risk analysis.

It should be noted that the prior work by Kittelson & Associates, Inc. mentioned in the introduction section a similar approach to identify some of the risk factors for the pedestrian and bicycle crashes in Oregon.

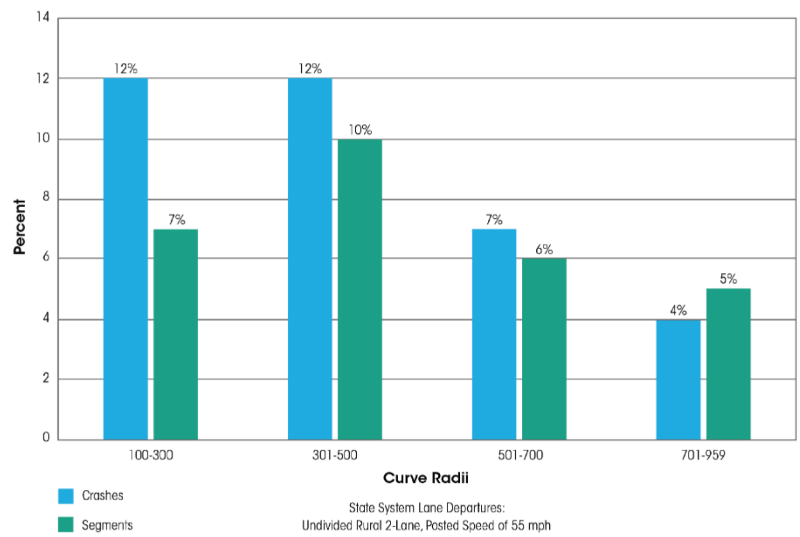


Figure 2.2: Example of NYSDOT: Curve Radii as a Risk-Factor Evaluation for Rural Undivided Highways

2.1.2 U.S. Road Assessment Program

The U.S. Road Assessment Program (USRAP) aims to provide a “systematic road assessment program in North America to inform motorists of the level of safety on the roads they travel.” (Harwood et al., 2008). The program builds on the success of the European Road Assessment Program (RAP), where the methods were developed. The RAP program categorizes roadways in a number of traditional ways (frequency and rate of fatal crashes, the difference from average rate performance), but also includes a method to develop a road protection score. The road protection score is calculated based on the potential for severe outcomes for head-on, run-off-the-road, and intersection crashes. The scoring is based on critical sub-factors, such as the width of the median for head-on crashes. For each of the crash types, there is a relative risk score calculated from a set of sub-factors that is clearly associated with the risk. The methodology prescribes that relative risk scores (converted to a simple index) are weighted together, with 15% for head-on crashes, 19% for run-off-the-road crashes and 66% for intersection crashes for the road segment under consideration.

A sample of the weighting table for head-on crashes is shown in Figure 2.1. The figure shows that the posted speed and degree of separation of opposing vehicle lanes influence the relative risk assigned to a specific roadway segment. For example, for an undivided roadway with a posted speed of 70 mph, the relative risk score is 38. This is translated to a 1 on the star rating on a 1-4 scale, with 4 being the lowest risk. For the same road but with a posted speed of 35 mph, the score is 1- which translates to a star score of 1 (the lowest risk segments). A similar process is done for the other crash types; then an aggregated star score is computed. A sample of the process, shown in map form for state highways in southeastern Iowa, is shown in Figure 2.4. On the map, the higher risk 1-2 starred segments can be seen as the red and black segments. The method is an example of translating relative risk factors to an index to create a ranking of segments.

Relative Risk Scores

	85th percentile or posted speed (mph)							
	70	65	60	55	50	45	40	35
Median treatment	**Relative Risk Score**							
Median width 70 ft or more	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
Median width 50 to 69.9 ft	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
Median barrier	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
Median width 30 to 49.9 ft	1	1	1	1	1	1	1	1
Median width 12 to 29.9 ft	4	4	4	3	2	1	1	1
Median width 3 to 11.9 ft	16	12	9	7	4	2	1	1
Undivided with centerline rumble strips	27	20	14	8	5	2	1	1
Undivided with marked centerline only	38	28	19	12	7	2	1	1

Star Rating Criteria

Number of Stars	Head-on Risk Score
4	0 - 2
3	2.01 - 5
2	5.01 - 10
1	over 10

Figure 2.3: Preliminary Relative Risk Scores and Star-Rating Criteria for Head-On Crashes (USRAP)

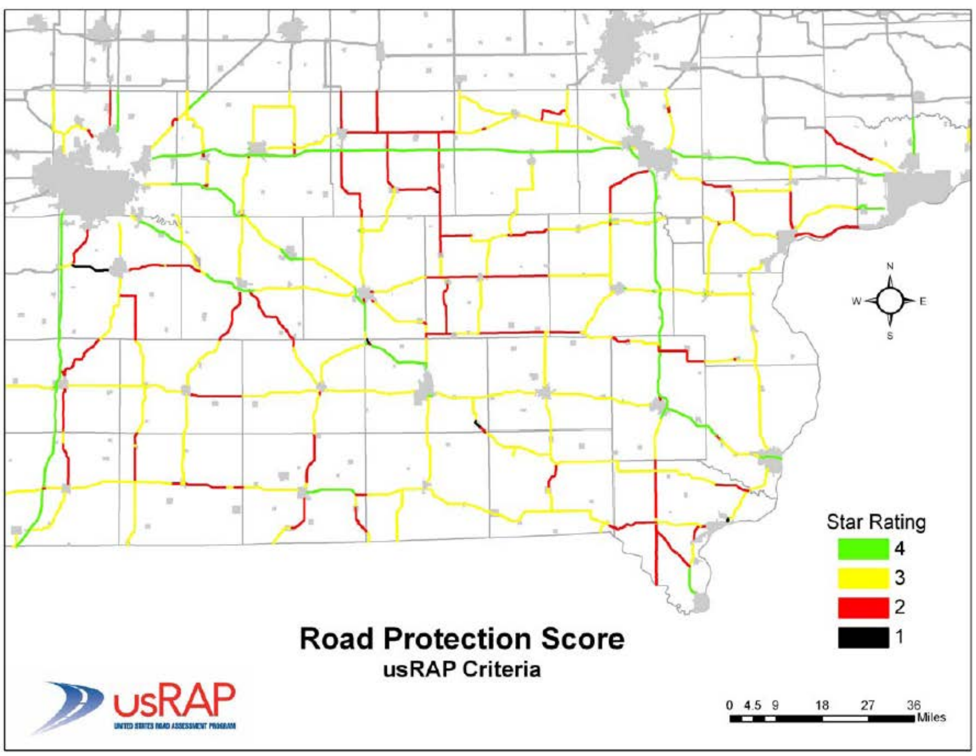


Figure 2.4: Star-Rating Map for Southeast Iowa Using USRAP RPS Criteria

2.1.3 Pedestrian and Bicyclist Intersection Safety Indices

In 2006, the FHWA sponsored research to develop the Pedestrian and Bicycle Safety Indices (Ped ISI and Bike ISI). The purpose of the ISI was to allow engineers and planners to proactively identify intersection crossings and approach legs which should be the greatest priority for undergoing pedestrian and bicycle safety improvements (Carter et al., 2006). In the methodology, the Ped and Bike ISI score is an evaluation of each approach leg of an intersection rather than evaluating the intersection as a whole. For example, a four-leg intersection would have four possible pedestrian safety scores, one for each crossing, and 12 possible bicycle safety scores, three for each leg (through, right lane, left lane). The approach with higher scores indicates a greater priority for an in-depth safety assessment. This tool applies to intersections with the following characteristics:

- Three- or four-leg intersections
- Signalized, two-way stop, and four-way stop
- Traffic volume from 600 to 5,000 vehicle per day
- One-way or two-way roads
- One to four through lanes
- Speed limit from 15 to 45 mph

To develop these indices, the researchers collected video and data from 68 pedestrian crossing intersections and 67 bicycle approaches at intersections in Florida, Pennsylvania, California, and Oregon. Safety ratings (opinion) from experts and bicycle/pedestrian-motorist interactions from a video analysis of each site were used to generate a multivariate linear regression model to explain the safety indices. The resulting models are listed in Table 2.1 and Table 2.2. The variables included in the model could be considered similar to risk factors. In the Bicycle ISI there are separate models depending on the type of maneuver the cyclist would be making (a through movement, a left turn or a right turn). The sign on the model coefficients and the units can be used to interpret the effect on the ISI score (where a larger ISI indicates a higher priority for improvement). For example, in the Ped ISI model the sign on the coefficient on a number of through lanes is positive, indicating that as the number of lanes increases, the relative safety of pedestrian crossing decreases.

Table 2.1: Ped ISI Model and Variable Descriptions

Ped ISI = 2.372—1.867**SIGNAL**—1.807**STOP** + 0.335**THRULNS** + 0.018**SPEED** + 0.006(**MAINADT*SIGNAL)** + 0.238**COMM** where:		
SIGNAL	Signal-controlled crossing	0 = no 1 = yes
STOP	Stop-sign controlled crossing	0 = no 1 = yes
THRULNS	Number of through lanes on street being crossed (both directions)	1, 2, 3, …
SPEED	Eighty-fifth percentile speed of street being crossed	Speed in miles per hour
MAINADT	Main street traffic volume	ADT in thousands
COMM	Predominant land use on surrounding area is commercial development (i.e., retail, restaurants)	0 = not predominantly commercial area 1 = predominantly commercial area

Table 2.2: Bike ISI Model and Variable Descriptions

Through	Bike ISI = 1.13 + 0.019**MAINADT** + 0.815**MAINHISPD** + 0.650**TURNVEH** + 0.470(**RTLANES*****BL**) + 0.023(**CROSSADT*****NOBL**) + 0.428(**SIGNAL*****NOBL**) + 0.200**PARKING**	
Right Turn	Bike ISI = 1.02 + 0.027**MAINADT** + 0.519**RTCROSS** + 0.151**CROSSLNS** + 0.200**PARKING**	
Left Turn	Bike ISI = 1.100 + 0.025**MAINADT** + 0.836**BL** + 0.485**SIGNAL** + 0.736(**MAINHISPD*****BL**) + 0.380(**LTCROSS*****NOBL**) + 0.200**PARKING**	
BL	Bike lane presence	0 = NONE or wide curb lane (WCL) 1 = bike lane (BL) or bike lane crossover (BLX)
CROSSADT	Cross-street traffic volume	ADT in thousands
CROSSLNS	Number of through lanes on cross Street	1, 2, …
LTCROSS	Number of traffic lanes for cyclists to cross to make a left turn	0, 1, 2, …
MAINADT	Main street traffic volume	ADT in thousands
MAINHISPD	Main street speed limit ≥ 56.3 km/h (35 mi/h)	0 = no 1 = yes
NOBL	No bike lane present	0 = BL or BLX 1 = NONE or WCL
PARKING	On-street parking on main street approach	0 = no 1 = yes
RTCROSS	Number of traffic lanes for cyclists to cross to make a right turn	0, 1, 2, …
RTLANES	Number of right-turn traffic lanes on main street approach	0, 1
SIGNAL	Traffic signal at intersection	0 = no 1 = yes
TURNVEH	Presence of turning vehicle traffic across the path of through cyclists	0 = no 1 = yes

2.1.4 ActiveTrans Priority Tool (APT)

Recognizing that for pedestrian and bicycle projects there are many competing project priorities and selection criteria, NCHRP sponsored research to develop the ActiveTrans Priority Tool (APT) (NCHRP Report 803 Lagerwey et al., 2015). The research resulted in a final report as well as a user-adaptable spreadsheet. This tool is a step-by-step methodology used for prioritizing pedestrian and bicycle improvements along the existing roads. To develop and support the research included, a literature review of general pedestrian and bicycle prioritization methodologies, special methodologies by jurisdictions, agency surveys, interviews, and case studies were conducted for more than 450 agencies throughout North America. The background research results indicated that prioritization methodology should consider the balance of need and feasibility of different projects and locations.

To use the tool, the user must select which general categories to consider in the prioritizing scheme. These include input options for stakeholder input; constraints (both cost and legal); opportunities (upcoming projects); safety; existing conditions; demand; connectivity; equity; and compliance. Each of these categories is assigned weights (selected by the user) to reflect the desired priority in the project selection and ranking process. Within each of these categories, there are predefined project scoring criteria that can be selected and entered by the user (or entirely custom criteria can be entered). Scoring criteria data can be converted to scaled scores using a method that best fits the data (e.g., proportionate, inverse proportionate, quantile scaling). For example, the predefined safety variables include total crash frequency, fatal and severe crash frequency, and rates for each of these variables for both pedestrian and bicycle crashes.

To illustrate the usefulness of the APT, the final report includes a number of case studies. Figure 2.5 shows the result of the APT applied to Gastonia, N.C., a community located west of Charlotte with a population of nearly 72,000. To maintain compatibility with the statewide project selection tool used by NCDOT, Gastonia's planners selected to weight the categories as follows: safety =15, demand/density = 10, benefit/cost = 10, access = 10, and constructability = 5. The safety score was computed by considering pedestrian crashes, overall speed limit, and whether the project provides separated facility and/or encourages a reduction in vehicular speeds. These criteria were then evaluated and prioritized using the APT and then mapped. The highlighted corridors were included in the city's long-range pedestrian plan.

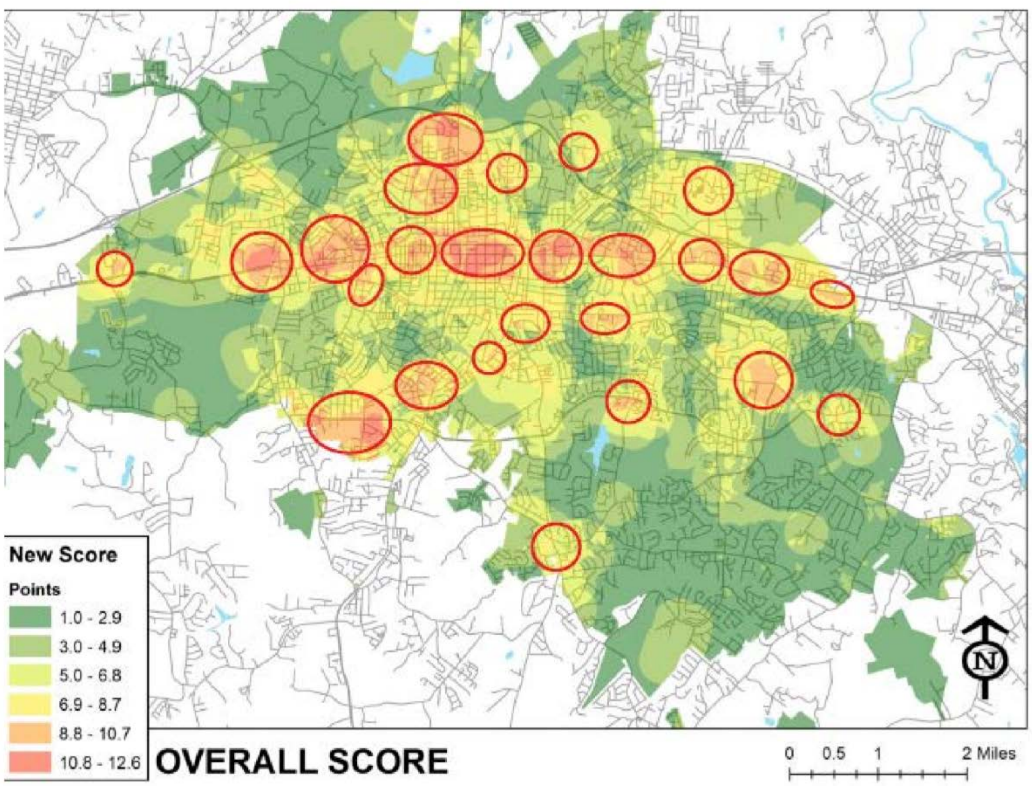


Figure 2.5: Combined Overall Safety Project Prioritization for Gastonia, N.C.

2.1.5 Level of Traffic Stress (LTS)

For bicycle networks, another tool has emerged that defines the Level of Traffic Stress (LTS) for users based on some simple risk factors. The method measures low-stress connectivity, defined as "the ability of a network to connect travelers' origins to their destinations without subjecting them to unacceptably stressful links" (Mekuria and Furth, 2013). The LTS method categorizes streets and intersections from LTS 1 (suitable for children) through LTS 4 (suitable for riders who are comfortable sharing the road with autos traveling at 35 mph or more). LTS 2, which anchors the street standards for all levels, is based on Dutch standards for bicycle facilities, because these have been shown to increase bicycling rates in the overall population (Mekuria and Furth, 2013). The criteria are shown in Table 3-3 for LTS categorization.

Table 2.3:: Criteria for Level of Traffic Stress in Mixed Traffic from Low Stress Bicycling and Network Connectivity (Mekuria et al., 2013)

Level of Traffic Stress			
LTS 1	LTS 2	LTS 3	LTS 4
- Physically separated from traffic or low-volume, mixed-flow traffic at 25 mph or less - Bike lanes 6 ft wide or more - Intersections easy to approach and cross - Comfortable for children	- Bike lanes 5.5 ft wide or less, next to 30 mph auto traffic - Unsignalized crossings of up to 5 lanes at 30 mph - Comfortable for most adults - Typical of bicycle facilities in the Netherlands	- Bicycle lanes next to 35 mph auto traffic, or mixed-flow traffic at 30 mph or less - Comfortable for most current U.S. riders - Typical of bicycle facilities in United States	- No dedicated bicycle facilities - Traffic speeds 40 mph or more - Comfortable for "strong and fearless" riders (vehicular cyclists)

- LTS 1: Anybody would bike on it
- LTS 2: For basic adult cyclists
- LTS 3 or 4: For advanced cyclists

The LTS method and the above-referenced BLTS criteria have been used to develop a GIS-based "low stress" bicycle analysis tool and methodology that uses community-based input data to evaluate the "bikeability" of a roadway network as can be seen from Figure 2.6, which is a coded, low-stress, GIS-based bicycle network for Salem, OR, and Figure 2.7, which shows the geospatial correlation between the bicycle network LTS and where bicycle crashes happened. Preliminary studies show that more than half of the bicycle crashes happened on streets with a higher level of stress.

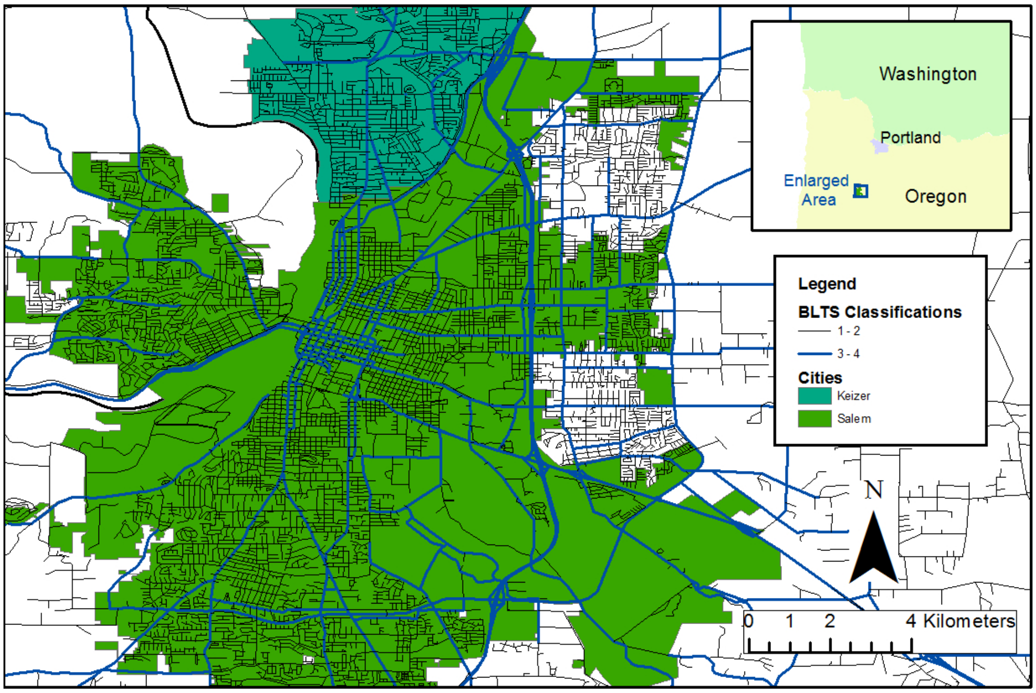


Figure 2.6: Salem Low-Stress Bicycle Network

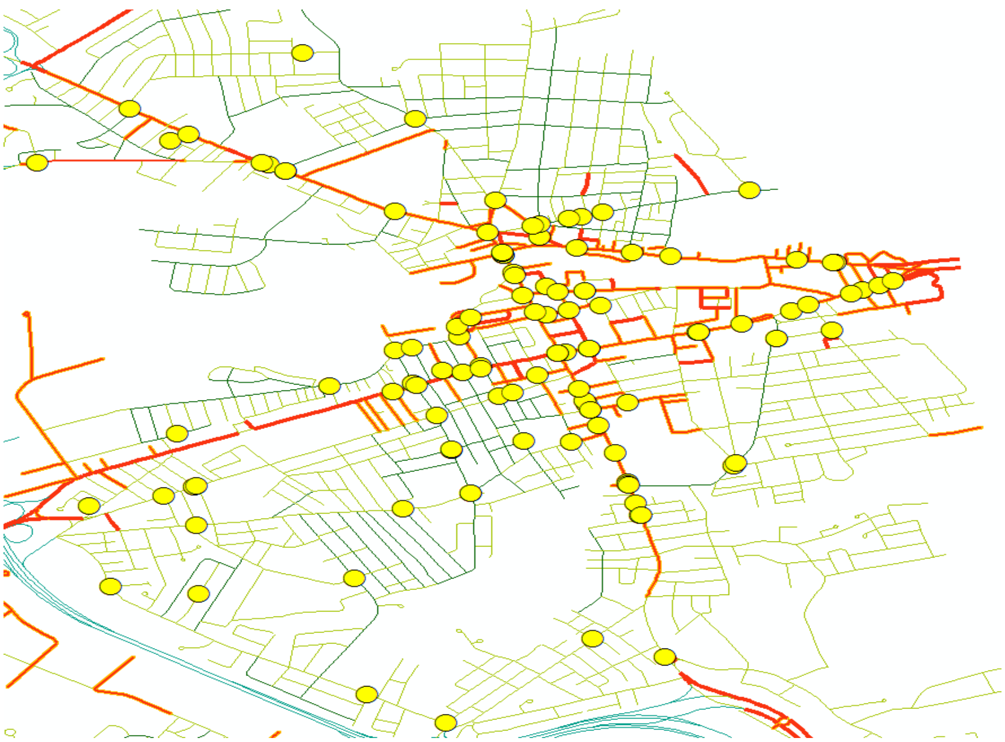


Figure 2.7: Geospatial Correlation of Bicycle LTS (color coded) and Bicycle Crashes (circular dots), Analysis of New Hampshire Data

2.2 QUANTIFYING BICYCLE SAFETY

In order to develop risk-based scoring models, it is important to understand from the literature the factors that increase the risk level for these types of crashes. This section begins by reviewing the modeling approaches that have been used to understand bicycle crashes. The section then reviews the literature on possible risk factors for bicycle crashes, including crash characteristics and the various roadway design features that are associated with increased risk.

2.2.1 Modeling Methods for Bicycle Crashes

This section explores the different models that are used to find which factors influence the severity of bicycle crashes. The most common models used were the Poisson distribution, negative binomial models, linear regression models, logit model, ordered probit model, and multivariable logistic regression. Some studies used simple summary statistics as well as a model, and others created their own model to overcome some model shortcomings. Safety Performance Functions (SPFs) are also used to help describe the mathematical relationships between crash frequency and significant factors of the bicycle crashes.

2.2.1.1 Poisson and Negative Binomial

The Poisson distribution was used to analyze the relationships of the crashes and the variables that influenced the crashes. In their study, Oh et al. (2008) used the Poisson distribution to analyze bicycle collisions at urban signalized intersections. However, only bicycle variables were considered and there could be more risk factors found if driver characteristics had been considered (Oh and Ph, 2008). Nordback et al. (2014) focused on finding SPFs for bicycles in cities in the United States, and used the Poisson distribution because of its ability to create a logical fit for the accident data provided. Finally, the Poisson distribution was used for a study of the largest cycling event held in New Zealand to determine what factors play into risk level for bicyclists from incident rates (Tin Tin, Woodward, and Ameratunga, 2013).

There are several studies that use the negative binomial model or some variation of the model. Oh et al. (2008) considered a negative binomial model when analyzing bicycle collisions at signalized intersections in an urban area. In a study that considered crashes involving a bicycle and motor vehicle at a signalized intersection, Wang et al. (2004) used three different negative binomial models to estimate the risk of such collisions. Noland and Quddus (2004) used a fixed-effect negative binomial model to analyze the risk factors of pedestrian and bicycle casualties for various regions in England. Finally, a negative binomial regression model was used to study various factors, both road and bicycle, which influence bicycle risk factors at unsignalized intersections in order to try and prioritize their safety levels (Schepers et al., 2011).

2.2.1.2 Linear Regression

Linear regression models were used in some studies but were generally accompanied by another modeling approach. Dixon et al. (2012) used SPFs along with two types of linear regression models, urban and rural, in order to quantify SPFs of driveways on state highways. These SPFs were mainly focused and applied to vehicles (Dixon, Avelar et al., 2012). Another study looking at predicting accidents for roads with minor junctions used a linear model in conjunction with an empirical Bayes procedure (Mountain and Jarrett, 1996).

2.2.1.3 Logit Model

Logit models were very commonly used in previous studies concerning bicycle-related crashes due to the models' ability to examine discrete choices, which are the level of severities of the crashes. Eluru et al. (2008) created a variation of the logit model, termed as a mixed generalized ordered response logit model, due to the limitations of a standard ordered response logit model to study pedestrian and bicycle injury severities in crashes. Kim et al. (2007) used a multinomial logit model for predicting the probability of different severity levels for bicycle-motor vehicle crashes in North Carolina. Another study used a mixed multinomial model to investigate three different types of crashes and the factors involved in those crashes (Pai, 2011).

Boufous et al. (2012) used a logit model to determine the risk factors for bicycles in Victoria, Australia, and Schepers & Brinker (2011) used a logit model to determine visual risk factors perceived by bicycles through a questionnaire. In order to find the perceived cycling risks and route acceptability of cyclists, Parkin et al. (2007) also used a logit model and a non-linear least squares model. Finally, Lenguerrand et al. (2006) used three different models – a multilevel logistic model, generalized estimating equation models and logistic models – to model the hierarchical structure of road crashes. However, the results from the logistic models are not consistent with other studies (Lenguerrand, Martin, and Laumon, 2006).

2.2.1.4 Probit Model

A common model used for analyzing crash injury severity levels and risk factors at intersections, both signalized and unsignalized, was a probit model or various forms of the probit model. This model was used because it can account for injury severities as naturally ordered variables (Lee and Abdel-Aty, 2005). Lee and Abdel-Aty used the ordered probit model to analyze vehicle-pedestrian crashes at intersections. Although this study did not consider bicycle crashes, the crash data used is similar to bicycle data and the model could be applied to bicycle incidents to determine risk factors and their significance. Abdel-Aty and Keller (2005) used an ordered probit model, tree-based regression, and probit model to examine signalized intersections and predict the level of injury severity. Another study used two types of probit models, an ordered and binary probit model, to analyze crash severity levels at unsignalized intersections (Haleem and Abdel-Aty, 2010).

2.2.1.5 Other Modeling and Analysis Methods

Several studies have developed their own models or use unusual models to explore bicycle crashes and determine the risk factors. One study chose to use a multivariate regression, which is based on a linear model, to determine the critical factors in fatal bicycle-automobile crashes (Bíl, Bílová, and Müller, 2010). Vandenbulcke et al. (2014) used a spatial Bayesian modeling approach to predict the risk levels of bicyclists along control sites, or a bikeable road network. Another unique model that was used was a quasi-induced exposure approach in order to identify factors that influence the bicycle crashes (Martínez-Ruiz et al., 2013).

Other studies chose to use a new type of model or methodology in order to use the results for planning or prioritization purposes. One such study evaluated projects for sketch-level scenario planning for bicycles in dangerous situations to make improvements to existing infrastructure (Lowry and Cool, 2014). Another study considered separated facilities for bicycles and the impact factors affecting crash rates by using a new side-path safety model (Petritsch et al., 2006).

Although many studies did not present specific risk factors in their results, their methodologies to help better understand bicycle behavior is helpful in determining where certain risk factors are more significant than others. A unique method presented by McDaniel et al. (2013) used origin-destination centrality to determine the importance of each link and node in a street network system and to estimate directional bicycle volumes. Basic statistical analysis was used for some studies in order to better understand various bicycle crash statistics (Bíl et al., 2010; de Geus et al., 2012; Räsänen and Summala, 1998; Summala et al., 1996). One study used basic statistical analysis tools as well and an incident rate method to study commuter cyclists involved in minor accidents and determine potential causes for these crashes (de Geus et al., 2012). Various models such as structural equation models were also considered in studies using R (Dolatsara, 2014; Bíl, Bílová and Müller, 2010).

2.2.1.6 Safety Performance Functions

SPFs describe the mathematical relationship between the frequency of crashes and the most significant factors in bicycle crashes. SPFs for vehicles are relatively well developed and general procedures can be found in the *Highway Safety Manual* (HSM), but SPFs for bicycles are still being developed. One recent study used crash data, AADT, and annual average daily bicyclists (AADD) at intersections in Boulder, CO, to create bicycle SPFs by using a negative binomial generalized linear model (Nordback, Marshall, and Janson, 2014). Specifically in Oregon, Dixon (2012) illustrates a method of site selection, the collection of crash and site-specific data, and analysis method for calibration that is in line with the HSM's methods to quantitatively estimate three facility types: rural two-lane roads, two-way roads, rural multilane roads, and urban and suburban arterial roads.

In Florida, Lu (2013) used SafetyAnalyst default SPFs and developed Floride-specific SPFs for various road types in order to compare the performance of each method, and then apply the Florida-specific SPFs for crash prediction performance identification and explore alternative clustering algorithms to identify high-crash locations. In Michigan, Dolatsara (2014) investigated factors that affect safety at intersections to enhance the development of SPFs and found that bicycle SPFs demonstrate that exposure, the presence of bicycle lanes and bus stops, and the number of left-turn lanes at intersections are positively associated with bicycle crashes.

2.2.2 Critical Factors Associated with Bicycle Crashes

The critical factors are presented in six categories: roadway, intersection, traffic characteristics, land-use, demographic and behavioral, and weather and lighting. Roadway factors include research related to the roadway geometry and cross section. The intersection section includes specific factors related to intersections (geometry and operations). The traffic characteristics include factors such as speed limit, peak-hour traffic, and traffic volumes. In the land-use section, research is reviewed that highlights the increased exposure associated with land-use types. In the demographic and behavioral section, the risk factors include the age of the bicyclist, their conditions while riding, driving conditions, type of vehicle involved in the crash, type of crash, the age of the driver, and the speed of the vehicle. Finally, the last subsection includes weather and lighting conditions at the time of the crash.

2.2.2.1 Roadway Geometry

Geometry plays a huge role in collisions between vehicles and bicycles. In published literature, researchers have analyzed factors such as the number of traffic lanes adjacent to bicycle traffic, road curvature, and the shoulder characteristics or the presence of a bike lane in depth.

Greibe (2003) found that when there were two lanes there were more accidents. In addition, there were more accidents in the same direction on a single lane with no centerline markings. This study also noted that many of the roadways geometry characters had a strong correlation with each other. When considering pedestrian-vehicle crashes, Lee and Abdel-Aty (2005) found that one lane reduced the number of pedestrian crashes when it is the pedestrian's fault by four times and two lanes reduced by nearly 0.75 times. It was also found that more crashes occurred on undivided roads with more lanes than divided roads with fewer lanes (Lee and Abdel-Aty, 2005). Petritsch et al. (2006) considered a side-path safety model design and found that the more lanes that are on the roadway, the more motorists focus on the opposing travel lanes and turning traffic as opposed to the activity on a side path. Additionally, on two-lane roads motorists look for cyclists on the side of the roadway, and cyclists using a side path may only concern themselves with traffic in the nearest travel lanes (Petritsch et al., 2006).

Pai (2011) found that horizontal and vertical curves can contribute to bicycle accidents. Schepers and den Brinker (2011) considered potential visual barriers that different road geometry causes cyclists, and found that cyclists collide with a bollard or road narrowing or ride off the road in a curve. This type of crash was found to occur more than when cyclists hit an obstacle because they were looking at something on the side of the road, but not more than cyclists looking behind them. The biggest takeaway from the study was that focal operations play a more important role in crashes involving a curve. Dixon et al. (2012) found that "no horizontal curves" should be an SPF that is included when calculating the unadjusted crash prediction model for the base conditions for a rural two-lane, two-way road segment. Eluru et al. (2008) found that crashes on curved/non-flat roadways tend to result in more severe injuries. Using a multinomial logit model, Kim et al. (2007) found that curved rounds significantly increase the chance that a fatal or incapacitating injury will occur during a vehicle-bicycle accident.

There are a number of different types of facility designs for bicycles and each has an impact on bicycle safety, such as the presence of bicycle lanes, the grade of the roadways/bicycle track, and if there are any different pavement markings or colors (Oh et al., 2008; Vandenbulcke, Thomas and Int Panis, 2014). Vandenbulcke et al. (2014) considered different cycle facilities and found that there is an increased risk of accidents when associated with a specific type of intersection. This study found that right-of-way intersections equipped with cycle lanes tend to have higher accident risk for cyclists, due to vehicles not respecting the right-of-way (i.e., right-hook crashes). The researchers also found that cyclists riding on marked cycle lanes in roundabouts and signalized intersections with marked cycle lanes had higher accident risk, and attributed the higher risk to the cyclists being in drivers' blind spots (Vandenbulcke, Thomas and Int Panis, 2014).

Schepers et al. (2011) found that more crashes where the bicycle has the right-of-way on a through movement occur at intersections with two-way bicycle tracks that are well marked and are reddish in color. However, this study found a cycle track where the approach is deflected 2-5 meters (6-10 feet) from the intersection decreased the risk for the cyclist. Walker (2007) considered the effect of lanes on how drivers overtake bicyclists on the road and discussed that more narrow roads might lead to vehicles passing cyclists closer, which might cause more risk.

Petritsch et al. (2006) created the Sidepath Safety Model in order to determine if a side path, or separated bicycle track, would be a viable option for a given road segment or how to improve an existing side path with multiple crashes. This model found that the path width has an impact on safety, and recommends that paths be built wide enough to accommodate multiple users along a segment but restricted at conflict points to calm traffic. It also found that the distance between the side path and the roadway, the speed of the adjacent roadway, and the number of lanes on the adjacent roadway were also key safety factors.

2.2.2.2 Intersections

The design of the intersection has an impact on bicycle safety in multiple ways, as concluded by Wang and Nihan (2004). For intersection and network movement, hazardous crossings, right hook, left sneak and complicated interactions are potentially dangerous to cyclists. Intersection safety was influenced by vehicle volume, vehicle speed, the percentage of heavy vehicles, and many other factors for both the major and minor roads (Dixon et al., 2012).

Oh et al. (2008) conducted a study based on surveys collected at 151 signalized intersections and found that average daily traffic volume, presence of bus stops, sidewalk widths, number of driveways, presence of speed restrict devices, and presence of crosswalks are all statistically significant factors that influence the risk level of bicycles. It has also been found that complex intersections (high number of road legs, road users, high number of signs, dense traffic crossings, etc.), and therefore complex traffic situations, increase the risk for bicycles (Vandenbulcke, Thomas and Int Panis, 2014)

Abdel-Aty and Keller (2005) considered three types of variables in different probit models for signalized intersections; (1) based on collision types, (2) based on intersection characteristics, and (3) based on a complete set of significant variables. These models found that the division of the minor road, as well as a higher speed limit on the minor road, was found to lower the expected injury level while a median on the minor road may prevent more head-on crashes, which were found to be more severe crashes. Additionally, a higher speed limit on the minor road may cause the speed differential between vehicles on intersecting roads to be smaller, likely resulting in a decrease in the crash severity level.

Another study looked at two types of crashes across 540 unsignalized intersections; (1) through bicycle-related collisions where the cyclist has the right of way, and (2) through motor vehicle-related collisions where the motorist has the right of way (Schepers et al., 2011). The results showed that Type 1 crashes occurred more when the two-way bicycle tracks are well marked and there is a reddish-colored bicycle crossing. Fewer crashes occur when there are raised bicycle crossings (speed humps) or other speed reduction measures. Haleem and Abdel-Aty (2010) considered the number of lanes for unsignalized intersections and found that the traffic volume on the major approach, the number of through lanes on the minor approach (surrogate measure for traffic volume), the upstream and downstream distance to the nearest signalized intersection, left and right shoulder width, number of left-turn movements on the minor approach, and number of right- and left-turn lanes on the major approach were significant factors that influence bicycle risk.

2.2.2.3 Traffic Characteristics

Many studies have recognized that traffic characteristics such as speed limit, peak-hour traffic, and traffic volumes such as AADT and ADT are risk factors for cyclists. Greibe (2003) found that higher speed limits relate to lower accident risks, but clarify that it does not mean that high speeds, in general, are safer; rather that high-speed roads tend to have few vulnerable road users. Wang and Nihan (2004) also found that speed limit decreases the risk of bicycle accidents, but state that it could be related to the turn maneuvers of right-turning vehicles. Similarly, Abdel-Aty and Keller (2005) determined that higher speed limits on the minor road lowered the expected injury level, and Eluru et al. (2008) found that higher speed limits lead to higher injury severity levels. On the other hand, Haleem and Abdel-Aty (2010) found that lower speed limits (less than 45 mph) reduced fatal injury probability when compared to greater than 45 mph. Kim et al. (2007) found that any speed greater than 20 mph and heavy vehicle traffic increased the risk of fatal injury.

Kim et al. (2007) also considered the peak-hour effects and found that during the a.m. peak hour (6-9:59 a.m.) there is an increased risk of fatal injuries for cyclists. Nordback et al. (2014) found that collisions were equally sensitive to both AADT and AADB (average annual daily bicycles). Haleem and Abdel-Aty (2010) determined that AADT on the major approach decreased the effect on fatal injury when a natural logarithm was used, but that the effect was increased when a surrogate measure for AADT was used to represent one, two and three through lanes on a minor road. Dixon et al. (2012) found that AADT increased the risk for cyclists in an urban environment.

2.2.2.4 Land Use

Land-use impacts, though not very detailed in the literature, do influence the overall safety of bicyclists because it impacts the amount and type of traffic and facilities of the road. Common distinctions of land-use types are urban, rural, residential, industry, farmland, institutional and commercial (Kim et al., 2007; Dixon, Monsere et al., 2012; Haleem and Abdel-Aty, 2010). Dixon et al. (2012)found that land use is a key factor that affects driveway safety, and Schepers et al. (2014) stated that land use has an effect on the distribution of traffic (bicycles included) over time and space. Oh et al. (2008) determined that the presence of industrial areas near intersections was associated with increased bicycle collisions. This is due to the more complicated traffic activities when compared with non-industrial areas.

Nordback et al. (2014) conclude that land use is a variable that might influence cyclist safety and should be considered for SPFs. One study analyzed the descriptive statistics of land use and found that higher severity crashes occurred outside of urban areas and at farm/wood/pasture or residential areas (Kim et al., 2007). Greibe (2003) used a dataset where land use proved to be one of the most important variables in the models generated, and land use and speed limit explain the level of vulnerable road users exposed to a certain extent. In the model used for this study, it was found that shops, blocks of flats (or apartments), and industrial/residential/neighborhood were a significant influence on bicycle safety (Greibe, 2003).

2.2.2.5 Demographic and Behavior

As expected, there are several factors that are specific to bicyclists when considering their risk level. The most impactful factor, according to the literature, is the age of the cyclist. Several studies found that riders over the age of 45 were more likely to be involved in a more severe crash (Kim et al., 2007; Boufous et al., 2012; Schepers and den Brinker, 2011; Tin Tin, Woodward and Ameratunga, 2013; Noland and Quddus, 2004). Bíl et al. (2010) found that cyclists 65 years and older were most at risk. Specifically, Schepers and den Brinker (2011) found that cyclists over 60 years old were more likely to be involved in crashes due to their low visibility. Kröyer (2015) found that fatalities increased for riders above the age of 55 and that there was an extreme increase in fatality risk between the age groups of 55-64 and 65-74. Alternatively, one study found that riders between the age of 10-19 were more likely to be involved in a higher severity crash (Martínez-Ruiz et al., 2013), and another discovered that children 9-11 years old are also at a higher risk (Maring and Van Schagen, 1990). Other studies reported that age was an important factor; however, they did not specify which age group was most at risk (Haleem and Abdel-Aty, 2010).

Kim et al. (2007) found that bicyclists without a helmet were more likely to have an incapacitating or non-incapacitating injury. Several other studies also found that when cyclists were not wearing a helmet, they were at higher risk (Andersson and Bunketorp, 2002; Martínez-Ruiz et al., 2013; Moahn et al., 2006; Noland and Quddus, 2004; Räsänen and Summala, 1998; Tin Tin et al., 2013). Additionally, the location of the crash was an important feature to the risk level, although it was not determined if there was a specific location that led to higher risk levels (Abdel-Aty and Keller, 2005; Eluru, Bhat, and Hensher, 2008). Several studies found that males are more at risk for higher severity of crashes (Boufous et al., 2012; Ekman et al., 2001; Eluru et al., 2008; Kim et al., 2007; Noland and Quddus, 2004; Schepers and den Brinker, 2011; Tin Tin et al., 2013).

Another factor that several studies found to contribute to high-risk levels was if the bicyclist was intoxicated (Olkkonen and Honkanen, 1990; Rodgers, 1995; Boufous et al., 2012; Schepers and den Brinker, 2011; Martínez-Ruiz et al., 2013; Andersson and Bunketorp, 2002; Eluru, Bhat and Hensher, 2008; Kim et al., 2007; Haleem and Abdel-Aty, 2010; Noland and Quddus, 2004). Other factors that were found included failure to follow traffic rules such as right-of-way, cyclist familiarity with the area, brake defects, and if there were two riders (Schepers and den Brinker, 2011; Martínez-Ruiz et al., 2013; Bíl, Bílová and Müller, 2010; Kim et al., 2007).

Driver characteristics also directly impact the risk level for bicyclists. The most influential factor based on several studies is if the driver is intoxicated (Eluru, Bhat and Hensher, 2008; Noland and Quddus, 2004). Additional factors that were found in many studies was that the risk of the bicyclist increased if a truck was involved in the crash (Kim et al., 2007; Walker, 2007; Greibe, 2003; de Geus et al., 2012; Boufous et al., 2012) or if the crash was a head-on collision (Greibe, 2003; Abdel-Aty and Keller, 2005; Lenguerrand, Martin and Laumon, 2006; Bíl, Bílová and Müller, 2010; Dixon, Avelar et al., 2012; Kim et al., 2007).Räsänen and Summala (1998) pointed out that drivers' attention greatly influences accidents or that the improper allocation of attention may lead drivers to ignore a cyclist who comes from an unexpected direction, such as drivers turning right hitting cyclists coming from the left. Drivers do not allocate enough attention to cyclists and, in some cases, cyclists do not feel or notice that they are in danger (Räsänen and Summala, 1998).

Other factors include vehicles speeding; the age of the vehicle; if a bus is involved in the crash; if there are parked vehicles along the side of the road; and if the age of the driver is above 60 years old (Walker, 2007; Vandenbulcke, Thomas and Int Panis, 2014; Parkin, Wardman and Page, 2007; Pai, 2011; Martínez-Ruiz et al., 2013; Bíl, Bílová and Müller, 2010; Eluru, Bhat and Hensher, 2008; Kim et al., 2007; Noland and Quddus, 2004).

2.2.2.6 Weather and Lighting

Bicycle crashes inherently have their own specific factors. The two more impactful factors are bad weather, such as fog, snow or rain, and the lighting of the road when it is dark outside. Moahn et al. (2006) recognize that weather conditions and darkness are risk factors that influence crash involvement. One study found that bad weather increases the probability of fatality by 128%, and darkness with no street lights increases the probability of fatality by 110% (Kim et.al, 2007).

Pai (2011) found that adverse weather, wet roads, and unlit streets were most common in rear-end crashes. Mountain and Jarrett (1996) stated that weather, quality of street lighting, and condition of the road surface used in a regression model will still have different underlying mean accident frequencies due to unique and unmeasured site characteristics. Stone and Broughton (2003) found that darkness increased the accident incidence rates and fatality rates. Martínez-Ruiz et al. (2013) considered bicycle defects and found that bicycles with brake defects were at a higher risk of being involved in a crash with a vehicle.

2.3 QUANTIFYING PEDESTRIAN SAFETY

This section seeks to identify the factors found from the literature that increases the risk level for pedestrian crashes. This section begins by reviewing the modeling approaches that have been used to understand pedestrian crashes. The section then reviews the literature on possible risk factors for pedestrian crashes, including crash characteristics and the various roadway design features that are associated with increased risk.

2.3.1 Modeling Methods for Pedestrian Crashes

This section emphasizes the different methods used by previous researchers to develop models of pedestrian safety. Most studies use statistical models to describe the safety performance of intersections, corridors or other locations. Negative binomial, linear regression and probit models are commonly used models, and GIS is another widely used technology to identify high-risk locations of pedestrian-vehicle crashes (though more from a descriptive sense). In order to identify and prioritize the locations with high-risk factors of pedestrian crashes, historical pedestrian crash data are analyzed. The high-priority locations are invested in for new facilities, and treatments are usually the ones where there is either a high frequency or high severity of pedestrian crashes. Thus, the crash frequency (or frequency rate) and/or severity degree are two common dependent variables used in risk-factor models of pedestrian safety.

Crash frequency is one of the typical variables to evaluate pedestrian safety level and identify the locations with high safety risks. Many researchers have used the number of crashes occurred in historical data to search significant risk factors contributing to vehicle-pedestrian collisions in the United States, Australia, and Canada (Wier et al., 2009; Schneider et al., 2010; Lee and Abdel-Aty, 2005; Chimba et al., 2014; Pulugurtha and Nujjetty, 2011; Torbic et al., 2010; Zegeer et al., 2001; Gårder, 2004; Miranda-Moreno et al., 2011; McMahon et al., 1999; Poch and Mannering, 1996). Frequency rate is another common dependent variable which depicts crash frequency per unit. Zegeer et al. (2006) collected the physical characteristics and behavioral data of 68 sites in California, Pennsylvania, and Florida. A number of pedestrian crashes per site per year were used as a dependent variable. Loukaitou et al. (2007) used a number of fatal crashes per 10,000 population to analyze the influence of urban sprawl degree to pedestrian safety level.

Crash severity can also be modeled. Besides the frequency models, Lee et al. (2005), Palamara and Broughton (2013), and Zegeer et al.(2001) also developed models to identify and evaluate the severity levels of pedestrian crashes in Florida, Perth central business district and 30 cities in the United States, respectively. Jang et al. (2013) divided the crashes into five categories: property damage only, slight injury, visible injury, severe injury and fatal (i.e., the KABCO scale). Fitzpatrick et al. (2014) conducted a very thorough analysis of the crashes in Texas from 2007-2011 where the crashes were divided into fatal and non-fatal categories.

Instead of categorizing crash locations into midblock crossings and intersections, some studies only devoted to the crashes occurred in the divided zones, with the most typical being census tracks. Census tracks are the basic unit of social economy variables; most of this type of studies focuses on the relationship between pedestrian safety and land-use characteristics. Wier et al. (2009); Jang et al. (2013); Chimba et al. (2014); Fitzpatrick et al. (2014); Gårder (2004); Ewing and Cervero (2001); Loukaitou et al. (2007); and McMahon et al. (2002) all developed the models by using the crash data per area or census tracks. Senserrick et al. (2014) and Miranda-Moreno et al. (2011) created a variable to indicate whether the crash occurred in the intersections or not. The results showed that more crashes occurred at non-intersection locations than intersections.

2.3.1.1 Negative Binomial

Negative binomial regression models are appropriate for modeling the number, or count data, as is considered the state of the practice in crash modeling. Several studies used negative binomial models to develop tools to estimate the pedestrian-vehicle crash frequency rate of both intersections and midblock crossings by using the data from Canada, the United States and Brazil (Schneider et al., 2010; Schneider, Ryznar and Khattak, 2004; Chimba et al., 2014; Pulugurtha and Sambhara, 2011; Torbic et al., 2010; Zegeer et al., 2001; Poch and Mannering, 1996).

2.3.1.2 Linear Regression

Wier et al. (2009) used ordinary linear regression to build a model to estimate pedestrian crash frequency based on the data from 2001 to 2005 in San Francisco. Zegeer et al. (2006) used a similar statistical model on pedestrian crash frequency rate at intersections. Ewing and Cervero (2001) made the log transformation of dependent and predictor variables in order to keep normal distributions using a log-linear regression model.

2.3.1.3 Probit

Jang et al. (2013) used the ordered probit model to estimate the severity level of pedestrian crashes by using the data from 2002 to 2007 in San Francisco. Lee et al. (2005) applied a similar model to the crashes in intersections in Florida. The ordered probit model is useful in crash severity research areas.

2.3.2 Critical Factors Associated with Pedestrian Crashes

The critical factors are presented in six categories: roadway, intersection, traffic characteristics, land use, demographic and behavioral, and weather and lighting.

2.3.2.1 Roadway

Diogenes and Lindau (2010) developed the models and indicated that larger road widths lead to higher vehicle-pedestrian crash frequencies at midblock crossings in Brazil. However, they found that the pedestrian crash rate would decrease as sidewalk width decreases. Correspondingly, Chimba et al. (2014); Palamara and Broughton (2013); Fitzpatrick et al. (2014); Garder (2004); and Sandt and Zegeer (2006) verified this conclusion in other countries, both at midblock crossings and intersections.

The presence of public transit stops or other public transit facilities usually indicates high pedestrian activities and high probability of pedestrian-related crash occurrence. Torbic et al. (2010) and Miranda-Moreno et al. (2011) verified this statement in both three-leg and four-leg intersection models with data from Canada. Pulugurtha et al. (2011) pointed out that this variable is only significant in intersections with low pedestrian volume. Schneider et al. (2004) and Diogenes et al. (2010) indicated that the midblock crossings near public transit stops and other public transit system facilities would experience high pedestrian crash rates in Brazil because pedestrians may behave unsafely around the stops.

Sandt and Zegeer (2006) pointed out that most pedestrian crashes occurred at the undivided midblock crossings using a North Carolina crash dataset. This might be due to the fact that a large proportion of the U.S. roadway system consists of undivided roadways. Lee et al. (2005) supported this is a significant variable in the model of crashes at intersections, but only for pedestrian-fault crashes based on the data from Florida.

Palamara and Broughton (2013) highlighted the importance of median refuges because they can reduce pedestrian exposure to traffic flow and provide sufficient walk time for pedestrians to complete their crossing. Schneider et al. (2010) used the proportion of crosswalks across main lanes or cross streets with medians, and made the conclusion that the presence of medians was negatively associated with pedestrian crashes. Medians may offer a refuge for pedestrians and may allow pedestrians to concentrate on crossing one direction of traffic at a time. Zegeer et al. (2006) focused on median islands in the intersections and reported that the presence of a median island was associated with a significantly lower pedestrian crash risk on multilane roads. McMahon et al. (1999) reported that the locations with grassy areas and unpaved shoulders might be less likely to be crash sites using data from North Carolina.

Based on the studies done by Wier et al. (2009), Fitzpatrick et al. (2014) and Miranda-Moreno et al. (2011), road classification (functional class) is associated with the pedestrian safety level. Major arterial roads were found to have a negative effect on safety issues. Wier et al. used the percentage of arterial roads' length to the total length as a factor in their model, and Fitzpatrick et al. used the four categories of road classifications – primary highway, secondary highway, arterial and local street – and treated them as factor variables in the model. Schneider et al. (2004) and Miranda-Moreno et al. indicated that the models explored the effect of the built environment on pedestrian crashes. The overall street length would lead to high pedestrian activities, thus, high pedestrian-related crash frequencies. The log-linear model verified that street length is a significant variable, but the magnitude of its effect is not very large.

Diogenes et al. (2010) indicated that the presence of a crosswalk would make drivers cautious and drive slightly slower than elsewhere. In addition, the crosswalk would channel the pedestrians' flow in both midblock crossings and intersections. Thus, the presence of a marked crosswalk would decrease the pedestrian crash rate. On the other hand, the distance to closed crosswalks or intersections has a positive correlation to the pedestrian crash rate, but the coefficient value is relatively small. Schneider et al. (2001), Palamara and Broughton (2013), and Garder (2004) verified that the locations with marked crosswalks are safer than the locations without crosswalks. Schneider et al. (2004) stated that the presence of sidewalks would increase the pedestrian safety performance, and McMahon et al. (1999) pointed out that the sidewalks have a particularly large safety benefit, especially in residential and mixed residential areas.

2.3.2.2 Intersection

Roadway intersections are critical locations for pedestrian safety. Some researchers have reported that the most common locations for fatal and injury pedestrian crashes are within 50 feet of intersections. Schneider et al. (2010); Lee et al. (2005); Zegeer et al. (2006); Palamara and Broughton (2013); and Poch and Mannering (1996) focus on the vehicle-pedestrian collisions that occurred in intersections in California, Florida and other states in the United States. Pulugurtha et al. (2011) and Torbic et al. (2010) investigated the relationship between pedestrian crashes and predictor variables, such as demographic characteristics, land-use characteristics, and road network characteristics on signalized intersections in North Carolina and select cities in Canada. In those, Pulugurtha et al. emphasized that low pedestrian volume at signalized intersections should be separate from high pedestrian volume ones. On the other hand, Zegeer et al. (2001) focused on the non-signalized intersections. This study emphasized the safety effect of marked and unmarked crosswalks on these locations, as well as the influence of different median types. McMahon et al. (1999) developed separate models for three-leg and four-leg intersections based on the data from North Carolina.

Poch and Mannering (1996) found that the presence of combining left-through lanes and protected left-turn lanes indicated an increasing left-turn volume, thus, increased annual accident frequencies. Left-turn movements cause problems of sight distance and accidents. Schneider et al. (2010) developed a negative binomial regression model and showed that significantly more pedestrian crashes occurred at intersections with more right-turn-only lanes and more nonresidential driveways within 50 feet of intersections. This might indicate that intersections with right-turn lanes tend to have longer crossing distance and a more complex set of interactions between pedestrians and motorists. It could also indicate a tendency for more right-turn-on-red collisions. Moreover, the driveways represent additional conflict points between vehicles and pedestrians near the intersections, and drivers may not look carefully for pedestrians as they exit driveways across the sidewalk.

Martin (2006) pointed out that the long signal cycle at intersections would increase the waiting time and delay for pedestrians as well as the crash frequencies because some people violate the traffic regulations. Poch and Mannering (1996) highlighted the uphill or downhill grades and a horizontal curve on an opposing approach may influence the sight distance and visibility of drivers and increase the probability of pedestrian crash occurrence.

Garder (2004) draws the conclusion that most pedestrian crashes happened at locations without any traffic control devices, and that single-lane roundabouts typically have the lowest crash rates among all types of intersections. Zegeer et al. (2006) also reported that signalized traffic control showed up as a variable with the most effect on safety. Stop-sign control did not show up as significant in the behavior model, but it is might because of the small sample size in the dataset. Lee et al. (2005), Poch and Mannering (1996) and Olga et al. (2010) had similar results that traffic controls at intersections are strongly associated with pedestrian safety levels.

Diogenes et al. (2010) proposed a method to evaluate the potential risk of pedestrian crashes at midblock crossings. Twenty-one midblock crossings with the highest number of pedestrian crashes in Porto Alegre, Brazil, from 1998 to 2006 were selected for evaluation using a Poisson regression model. The results indicated that a combination of interactive risk factors influences the estimated number of pedestrian crashes and that as the number of crossing stages increases the pedestrian crash rate decreases. Sandt and Zegeer (2006) aimed to understand the characteristics of midblock crossing pedestrian crashes to determine appropriate safety treatments. Datasets from Kentucky, Florida and North Carolina were used to determine which variables were significant at midblock crossing locations. Compared to crashes that occurred at intersections, crashes are more likely to occur at the midblock with the following features: no divided road, fewer lanes, urban area, and residential areas and during 2-6 p.m.

2.3.2.3 Traffic Characteristics

Empirically, increases in road facility vehicle volume would increase the probability of vehicle-pedestrian conflicts on that facility. Schneider et al. (2004); Zegeer et al. (2006); Pulugurtha et al. (2011); Palamara and Broughton (2013); Miranda-Moreno et al. (2011); Martin (2006); Loukaitou et al. (2007); and Poch and Mannering (1996) all verified this statement and concluded that higher traffic volumes result in higher pedestrian crash frequency. Wier et al. (2009) made the log transformation and Schneider et al. (2010) made the natural log transformation in order to keep normal distribution. Torbic et al. (2010) pointed out that the traffic volume of major roads and minor roads are both significant variables when predicting pedestrian crashes in intersections. Fitzpatrick et al. (2014) emphasized that high heavy vehicle (trucks) volume would increase the probability of high severity pedestrian crash occurrences. Diogenes et al. (2010) also mentioned that the high public transit vehicle volume would lead to more pedestrian crashes in midblock crossings in Brazil.

Similarly, the probability of pedestrian crashes would be greater when the pedestrian volume increases because pedestrians would have more exposure to other road users. Schneider et al. (2004), Pulugurtha et al. (2011) and McMahon et al. (1999) support this conclusion using data from North Carolina intersections.

In most studies, high vehicle speed would increase both the frequency rate and severity level of vehicle-pedestrian crashes (Lee et al., 2005; Zegeer et al., 2006; Chimba et al., 2014; Garder, 2004; Sandt and Zegeer, 2006; Martin , 2006; McMahon et al., 1999; Poch and Mannering, 1996). Most studies used the speed limit of the target zone as a predictor variable because it was easy to collect the data. Zegeer et al. (2006) collected the 85^{th} percentile of the total vehicle speed, which was closer to the real situation. Senserrick et al. (2014) had different results as most crashes occurred not in low-speed urban areas but in high-speed rural areas.

2.3.2.4 Land Use

Senserrick et al. (2014) found that the vast majority of pedestrian casualties occurred in urban areas, but rural areas represent a higher risk of fatal outcomes based on the data from 2004 to 2008 in Australia. Sandt and Zegeer (2006) and Lee et al. (2005) verified that pedestrian crashes occur predominantly in urban areas but rural crashes more often lead to pedestrian deaths, possibly due to higher vehicle speeds, by using the data from Florida and other cities in the United States.

Land-use type is a common variable which is associated with vehicle-pedestrian collisions. Land-use variables are potential partial proxies for pedestrian activity and pedestrian attractors. Wier et al. (2009) found that the percentage of neighborhood commercial area of land area and the percentage of residential-neighborhood commercial area of land area had a positive association with vehicle-pedestrian collisions. On the contrary, land area has a negative effect on the crash frequency. Pulugurtha et al. (2011) made the conclusion that land-use predictor variables such as a single-family residential area, urban residential-commercial area, commercial center area, and neighborhood service district have a negative effect on the pedestrian crashes in a given area. The negative effect could be attributed to an increased level of pedestrian activity and motorists tend to be more alert and attentive, indirectly resulting in pedestrian safety. Loukaitou et al. (2007) pointed out that the percentage of commercial and high-density residential areas has a positive effect on pedestrian crashes, but the percentage of vacant, industrial and office land use have a negative effect. Schneider et al. (2010); Zegeer et al. (2006); Chimba et al. (2014); Torbic et al. (2010); Miranda-Moreno et al. (2011); and Poch and Mannering (1996) also mentioned that land-use variables are significant in pedestrian safety models.

Chimba et al. (2014) included the presence of schools in their model and found that they were positively associated with pedestrian crashes, likely because of higher exposure when school is in session. Miranda-Moreno et al. (2011) had a similar conclusion by using data from Montreal. The number of schools in the area has a positive effect on pedestrian crash frequencies. Schneider et al. (2004) pointed out that the locations near libraries, stadiums and academic buildings have a higher risk for pedestrian crashes using data from the University of North Carolina, Chapel Hill.

Ewing et al. (2003) indicated that block size, resources inventory net density, and Eigenvalue index also affect pedestrian safety levels in an area. These variables are all related to urban sprawl level, thus higher values correspond to a lesser degree of sprawl and lower fatality rates in pedestrian crashes. McMahon et al. (1999) and Martin Allison (2006) pointed out that older neighborhoods are more likely to contain pedestrian crashes compared with newer neighborhoods.

2.3.2.5 Demographic and Behavior

Wier et al. (2009) indicated that resident population and employment population are both positively associated with pedestrian-related crash frequencies using crash data from San Francisco. Loukaitou et al. (2007) found similar results, with a higher probability of pedestrian collisions in neighborhoods with high population and employment density. Martin (2006) and McMahon et al. (1999) included residential population in their models. Miranda-Moreno et al. (2011) and Ewing et al. (2003) include employment density in their models. Miranda-Moreno also introduced the number of jobs as a variable, and it had the similar effect of employment density.

Chimba et al. (2014) indicated that higher rates of crashes are associated with lower household income and a higher percentage of people living below the poverty level. The potential reason might be that households with high average income have more vehicles and walk less in their daily lives. Torbic et al. (2010), Ewing et al. (2003), and Martin Allison (2006) had similar findings about household income. Wier et al. (2009) focused on the percentage of people living below the poverty level and made the same conclusion.

Chimba et al. (2014) also found that the number of vehicles per housing unit had a negative effect on pedestrian crashes in Tennessee. It was likely because more vehicles in a household could be an indication of decreased pedestrian activity and, hence, less exposure to the risk of pedestrian collisions. Martin (2006) made a similar conclusion in his review of pedestrian safety levels.

Lee et al. (2005); Senserrick et al. (2014); Diogenes et al. (2010); Sandt and Zegeer (2006); and Martin (2006) found that there were proportionally more male casualties than female. In addition, males were shown to sustain more serious injuries, to be hit at higher average vehicle speeds, and to have higher mortality rates than female pedestrians.

Age was found to be a critical factor in relation to the risk of a pedestrian crash and injury. Palamara and Broughton (2013) reported children and older people recorded higher proportions of pedestrian deaths. Senserrick et al. (2014) found that age groups most commonly involved were those aged 18-24, especially on weekends; 75+ especially on weekday days; and 13-17 especially at school commute times. Schneider et al. (2010); Lee et al. (2005); Jang et al.(2013); Chimba et al. (2014); Fitzpatrick et al. (2014); and Sandt and Zegeer (2006) also included an age factor into their analysis, and emphasized children and old people are the most vulnerable pedestrians.

The topic of race and ethnicity is linked with pedestrian safety. Jang et al. (2013), Chimba et al. (2014) and Loukaitou et al. (2007) reported that people who identify as Latino, Black and Hispanic were more likely to be involved in pedestrian-vehicle crashes.

Ewing et al. (2003) included the household size and people's working age into the model of pedestrian fatal rates in 448 counties in the United States. The natural logarithms of sprawl index, average household size, the percentage of the population of working age, and per capita income accounted for 47% of the variance of traffic fatality rate. Chimba et al. (2014) found that the percentage in the labor force has a significant positive influence on the occurrence of pedestrian crashes.

Lee et al. (2005) claimed that pedestrian and drivers' alcohol and drug use is an important factor affecting pedestrian crashes. Alcohol-impaired pedestrians and drivers were more involved in severe crashes, and young intoxicated males are considered to be a high-risk pedestrian group since alcohol and drug use impairs pedestrians' perception. Martin (2006) indicated that pedestrians with disabilities or carrying heavy bags are more likely to be involved in pedestrian crashes because they move slowly and have difficulty avoiding vehicles.

Jang et al. (2013) found that severe injury pedestrian crashes are more likely to happen at night and on weekends because of low visibility and high pedestrian activities, respectively. Fitzpatrick et al. (2014) had a similar conclusion that nighttime is the peak time for pedestrian crash occurrence using Texas crash data from 2007 to 2011. Palamara and Broughton (2013) included time factors for adjustment factors of pedestrian volume at intersections.

2.3.2.6 Weather and Lighting

Lighting is a serious risk factor in pedestrian crashes because it is related to drivers' sight and visibility, especially at nighttime. Fitzpatrick et al. (2014) indicated that 82% of the crashes in Texas from 2007 to 2011 occurred in dark conditions, almost half of which were at locations with no lighting. Senserrick et al. (2014) claimed that males and older pedestrians are more likely to be involved in crashes in poorer lighting conditions, particularly when crossing a road away from an intersection. However, Palamara and Broughton (2013) pointed out that the majority of pedestrian crashes occurred during daylight hours in CBD (Central Business District) areas. Lee et al. (2005), Jang et al. (2013), and Fitzpatrick et al. (2014) all included the weather factor in the models and found that rainy weather has a positive influence on both pedestrian crash frequency and severity level.

2.4 SUMMARY

The objective of this review was to identify potential risk factors for bicycle and pedestrian crashes. The review identified a number of key risk factors. Table 2-3 summarizes the key potential risk factors for bicycles and Table 2-4 for pedestrians. It is clear from the summary table and the review that the two modes share many of the same potential risk factors.

Table 2.4: Summary of Identified Potential Risk Factors for Bicycle Crashes

Roadway: • Horizontal curves • Lane width • Number of driveways • Presence of bicycle lanes • Presence of bicycle paths • Presence of bus stops • Presence of median • Presence of parking • Vertical grade (slope) • Width of bicycle lanes	**Traffic Characteristics:** • Average daily traffic • Functional class • Number of left- and right-turning vehicles • Operating vehicle speed • Percent heavy vehicles • Posted speed limit	**Demographic and Behavior:** • Age • Driver drug/alcohol use • Gender • Inappropriate speed • Inattention • Pedestrian drug/alcohol use • Protective measures (helmet/wear fluorescent color)
Intersections: • Distance to nearest traffic signal • Number of left-turn lanes • Number of right-turn lanes • Number of traffic lanes • Presence of a roundabout • Presence of crosswalk • Presence of traffic signal • Type of traffic control	**Land Use:** • Commercial area • Industrial area • Institutional area • Residential area	**Lighting and Weather:** • Lighting adequacy data • Visibility • Weather conditions

Table 2.5: Summary of Identified Potential Risk Factors for Pedestrian Crashes

Roadway:	**Traffic Characteristics:**	**Demographic and Behavior:**
• Average sidewalk width • Distance to the closest marked crosswalk or intersections • Lane widths • Maximum number of crossing stages • Number of driveways • Number of traffic directions • Number of traffic lanes • Paved shoulder • Presence of bus stop • Presence of marked crosswalk • Presence of median • Presence of paved sidewalk • Total road width	• Average daily traffic • Functional class • Number of left- and right-turning vehicles • Operating vehicle speed • Pedestrian crossings volume • Percent heavy vehicles • Percentage of public transit vehicles • Posted speed limit • Time of a day • Weekend/Weekday	• Age<18 and Age >65Age • Driver drug/alcohol use • Employee population • Gender • Household size • Mean household income • Pedestrian drug/alcohol use • Proportion of people of working age • Race • Resident population • Single-family residential • Vehicles numbers in housing unit
Intersections:	**Land Use:**	**Lighting and Weather:**
• Horizontal curve on opposing approach • Midblock location • Number of lanes • Number of left-turn lanes • Number of right-turn lanes • Number of through lanes being crossed • Pedestrian delay (cycle length) • Presence of enhanced crossing • Presence of intersection • Presence of median (major or minor) • Type of traffic control	• Block size • Natural Resources Inventory net density • Neighborhood business • Neighborhood service district • Presence of school zone • Residential-neighborhood commercial • Urban/rural areas	• Lighting (dark/daylight) • Lighting adequacy data • Visibility • Weather (clear/other/raining)

3.0 DATA COLLECTION

This chapter describes the data collection methods (including sampling approach and selection criteria), the elements collected and methods for data collection, and process for linking to crash data.

3.1 SAMPLING APPROACH

A random sampling approach was used to select segments and intersections for data collection to ensure that any heterogeneous characteristics associated with land use, functional class and other data elements could be captured in the modeling. The criteria for selection, development of the sample and procedure for the random draw is described.

3.1.1 Criteria

The research team collected data on a subset of all public roads (all public roads are eligible for funding in the ARTS program). The research team used the following criteria to develop the sample pool:

- Both state and non-state owned roadways are included;
- Connection ramp segments of state highway systems were filtered out:
- Roadway segments are in an urban area; and
- Roadway segments are arterial (minor or principal) functional classification.

Though people walking and biking are exposed to risk in all areas of Oregon, the majority of crashes occur in urban areas (87% of pedestrian crashes and 92% of bicycle crashes in 2013, based on a review of the crash data GIS files). To define urban areas, the project team used city limit boundaries provided in ODOT GIS files. The use of the city boundaries resulted in some segments that were not necessarily "urban" in character. This definition also excluded roads in unincorporated county jurisdiction that might be urban in nature.

Figure 3.1 shows all roadways meeting the criteria in the state clipped to only arterials in urban areas. As expected, there is a large concentration of roadways in the Willamette Valley but other areas of the state are represented. Figure 3.2 illustrates the pedestrian and bicycle crashes (on all roadways) mapped for Bend, OR, from 2009-2013. In the figure, only arterial roadways are shown and the dots represent crashes on lines. The figure shows that most of the pedestrian and bicycle crashes occurred on these roadways.

There are approximately 1,800 miles of roadway for potential sampling (arterials in urban areas) which are split between principal arterials (800 miles) and minor arterials (1,000 miles). Region 1 and 2 have similar miles available for data collection and have approximately 70% of the roadway mileage available for data collection.

In this GIS file, each roadway is broken up into many smaller segments. After a preliminary inspection of the roadway segments identified in the previous step, the research team considered some segments too short for practical data collection. Thus, the research team eliminated short segments (i.e., less than 0.05 miles) from the potential sample sites. About 200 miles of potential samples were filtered out for length. Table 3.1 presents the breakdown of the candidate roadway miles by region and arterial type (principal and minor) after filtering. Table 3.2 shows the number of potential segments after filtering for this minimum length. The resulting sample pool consisted of approximately 1,600 miles. As shown in Table 3.3, the average length of any one segment is roughly 0.10 mile (528 feet).

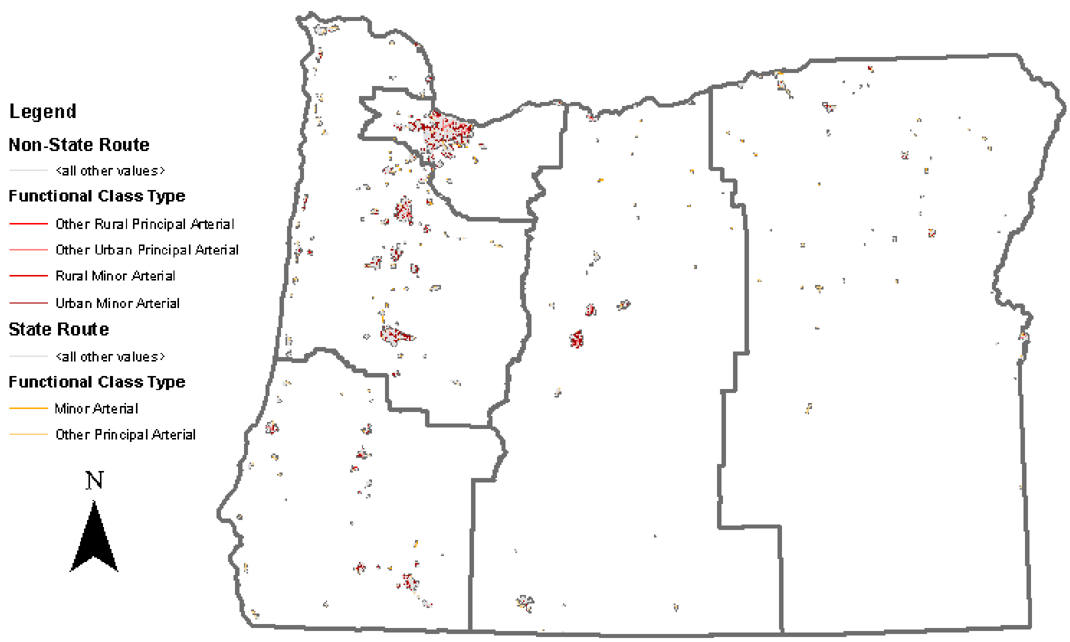


Figure 3.1: Filtered Functionally Classified Roadways Clipped to Urban Areas

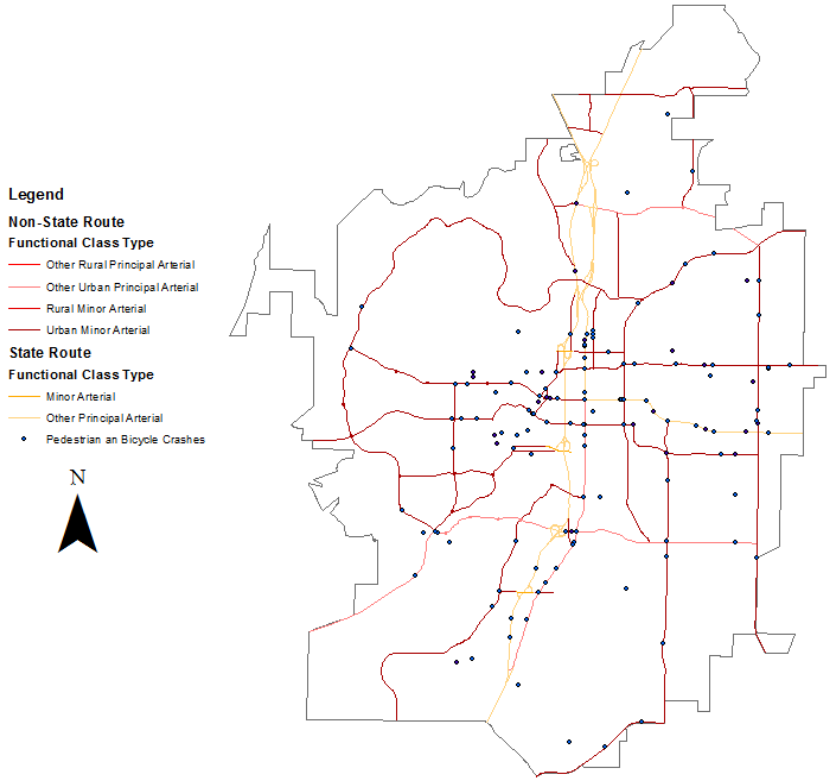


Figure 3.2: Filtered Roadways with Pedestrian and Bicycle Crashes Shown, 2009-2013, Bend OR

Table 3.1: Summary of Potential Segments by ODOT Region, Miles

ODOT Region	Principal Arterial		Minor Arterial		Total	
	Miles	Percent (%)	Miles	Percent (%)	Miles	Percent (%)
1	196.19	27.5	310.43	37.2	506.62	32.8
2	261.47	36.7	292.20	34.9	553.67	35.7
3	96.14	13.5	93.37	11.2	189.51	12.2
4	84.89	11.9	93.02	11.1	177.91	11.5
5	74.15	10.4	47.22	5.6	121.37	7.8
Total	712.84	100	836.24	100	1549.08	100

Table 3.2: Summary of Potential Segments by ODOT Region, Number of Segments

ODOT Region	Principal Arterial		Minor Arterial		Total	
	Count	Percent (%)	Count	Percent (%)	Count	Percent (%)
1	757	38.7	2531	37.8	3288	38.0
2	700	35.7	2326	34.9	3026	35.0
3	233	11.9	778	11.6	1011	11.7
4	196	10.0	712	10.6	908	10.5
5	73	3.7	343	5.1	416	4.8
Total	1959	100	6690	100	8649	100
Total Mileage	712.84		836.24		1549.08	

Table 3.3: Maximum, Minimum and Average Length (Mile) of Segments

ODOT Region	Principal Arterial			Minor Arterial			Average (Both)
	Max	Min	Average	Max	Min	Average	
1	16.11	0.05	0.26	6.11	0.05	0.12	0.15
2	9.88	0.05	0.37	6.57	0.05	0.13	0.18

3	6.83	0.05	0.41	2.79	0.05	0.12	0.19
4	7.56	0.05	0.43	3.35	0.05	0.13	0.20
5	4.80	0.05	1.02	1.99	0.05	0.14	0.29
All	16.11	0.05	0.36	6.57	0.05	0.12	0.18

3.1.2 Random Selection Process

To execute the random sample draw, a data table of segments meeting the criteria was created from an ODOT GIS functional class file. In this GIS file, each roadway is broken up into many smaller segments. Each row in the data file was assigned a randomly generated number. The segments were then sorted by the random number and selected in order for consideration in data collection.

Ideally, segments would be homogenous in key data elements over the entire length. This is desirable for modeling purposes since each segment can then be an observation unit and it is not necessary to control for the partial presence of the variable. After a segment is selected, the research team viewed the segment in Google Earth. If the segment is homogenous in the key data elements (shown in the following section), the segment was included in the sample. If not homogenous and the segment is long enough it was split to create homogenous segments. One of the segments was chosen at random by the data collector. If the resulting segments were too short, they were discarded from the sample.

Intersections were collected concurrently with the segment sampling process. To limit the data collection scope, only the locations with traffic control on the major road (stop or signal) were included in the intersection data collection. If a selected segment contains an intersection with the traffic control on the major road, the intersection was selected for data collection. Note that segments that contain intersections would be divided at the intersection. If the segment is bounded by two signalized or four-way stop intersections, only one intersection was included in the sample.

The sampling process continued until a sufficient number of intersections for modeling was selected. In the initial sample draws, crash history was not considered. As the sampling process progressed the initial sample was analyzed. It became clear that the random process was not generating enough selections with observed crashes. The research team filtered the remaining segments for crash occurrence and continued the random sampling of these segments.

3.2 DATA ELEMENTS

The literature review identified the key variables that should be considered in a risk model. In this section, the data elements collected for segments, intersection, land use, crash data, and motor vehicle and bicycle volumes are described.

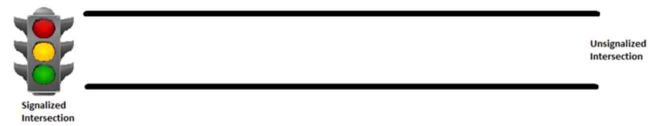
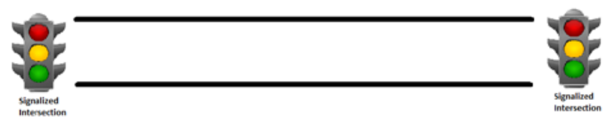
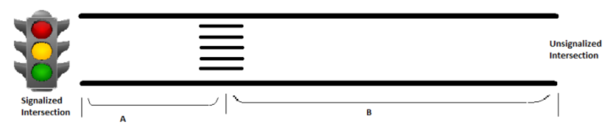
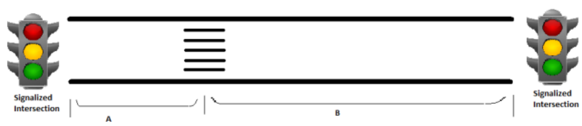
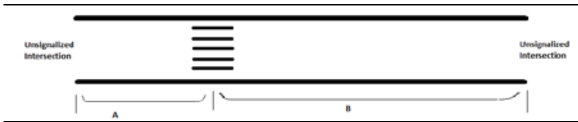
3.2.1 Segments

Table 3.4 summarizes the data elements collected for segments. The elements were primarily collected manually from inspection and measurement of Google Earth aerial photos. Most of the data elements are self-explanatory in description. During the pilot data collection exercise, it became clear that it was important to capture both the presence of the midblock crossing opportunity and the types of traffic control available at the bounded ends of the segment. The classification scheme devised is shown in graphical form in Table 3.5. For segments with no midblock crossing and a signalized intersection on one end (type A), the maximum crossing distance is taken as the segment length. If the segment is bounded by signalized intersections (type B), then the maximum walking distance was calculated as half the segment distance. For types C, D, and E, the placement of the crosswalk dictates the maximum distance a person would have to walk to have a crossing location.

Table 3.4: Segment Data Elements

Segment Data Element	Collection Method or Source
Functional class of roadway	Functional class of roadway (arterial, minor)
Traffic volume (AADT, factored to 2013)	ODOT databases, local files, other sources
Estimated bicycle volume per day (STRAVA)	STRAVA database with expansion
Number of traffic lanes (excluding two-way left-turn lane)	Google Earth/ODOT Digital Video Log
Presence of two-way left-turn lane	Google Earth/ODOT Digital Video Log
Presence of bicycle lanes	Google Earth/ODOT Digital Video Log
Width of bicycle lanes (ft.)	Google Earth/ODOT Digital Video Log
Travel direction (one-way or two-way)	Google Earth/ODOT Digital Video Log
Number of driveways	Google Earth/ODOT Digital Video Log
Posted speed limit (mph)	Google Earth/ODOT Digital Video Log
Width of sidewalk buffer (the space between curb face and sidewalk)	Google Earth/ODOT Digital Video Log
Width of sidewalk buffer (ft.)	Google Earth/ODOT Digital Video Log
Number of marked midblock crosswalks within segment	Google Earth/ODOT Digital Video Log
Presence of on-street parking	Google Earth/ODOT Digital Video Log
Presence of lighting along segment	Google Earth/ODOT Digital Video Log
Surrounding land use (commercial, industrial, institutional, residential)	EPA Smart Location Database
Presence of school area within 1000 feet from midpoint of segment	Google Earth
Number of transit lines with routes on segment	Google Transit
Number of intersections within segment	Google Earth/ODOT Digital Video Log

Table 3.5: Segment Crossing Opportunity Classification Scheme

Segment Configuration	Type	Maximum Walking Distance (ft.)
	A	½ segment length
	B	½ segment length
	C	Distance B in figure
	D	Half of minimum of distance B and A in figure
	E	Distance B in figure
	N	½ segment length

3.2.2 Intersections

Table 3.6 summarizes the data elements collected at intersections. Data were collected on each leg of the intersection approach. However, for modeling purposes, much of the data were condensed to represent the intersection-level rather than each approach. Note that volume of vehicles and bicycles is explained in Section 3.2.4.

Table 3.6: Intersection Data Elements

Segment Data Element	Collection Method or Source
Traffic volume (AADT, factored 2014)	ODOT databases, local files, other sources
Estimated bicycle volume per day (STRAVA)	STRAVA database with expansion
Functional class of roadway	ODOT databases
Number of left-turn lanes	Google Earth/ODOT Digital Video Log
Number of right-turn lanes	Google Earth/ODOT Digital Video Log
Presence of bicycle lanes	Google Earth/ODOT Digital Video Log
Number of total traffic lanes (including left- and right-turn lanes) on all approaches	Google Earth/ODOT Digital Video Log
Posted speed limit (mph)	Google Earth/ODOT Digital Video Log
Presence of lighting by approach	Google Earth/ODOT Digital Video Log
Type of traffic control (four-way stop, signal, roundabout)	Google Earth/ODOT Digital Video Log
Presence of school area within 1,000 feet	Google Earth
Presence of green bicycle markings	Google Earth/ODOT Digital Video Log
Number of bus stops within 1,000 feet	Google Transit
Presence of median	Google Earth/ODOT Digital Video Log

3.2.3 Land Use

Based on work for a complementary ODOT research project, two data sources were used to gather information on the land-use and built environment characteristics. The data were available in a GIS format and spatial buffering/joining was used to associate the data with each segment and intersection. Currans et al. (2014) defined five neighborhood concepts (A-F) on the urban – suburban spectrum using three measures of the built environment – density, diversity, and design, and classified the census blocks based on these categories. Figure 3.3 shows a sample map of the neighborhood concepts.

The research team also gathered -related variables such as total road network density and street intersection density from the Environmental Protection Agency's (EPA) Smart Location Database (2013). The street intersection density variable in the EPA database reflects the weighted density of intersections. This data was available by census tract. Figure 3.4 provides an example of the street network and intersection density variables.

Table 3.7: Land and Built Environment Data

Data Element	Collection Method or Source
Neighborhood concepts	GIS geo-database (Currans et al., 2014)
Three-leg intersection density (per square mile)	EPA's Smart Location Database
Four-leg intersection density (per square mile)	EPA's Smart Location Database
Retail density (per acre)	EPA's Smart Location Database
Total population density (people per square mile)	EPA's Smart Location Database
Household density (per acre)	EPA's Smart Location Database
Household size	EPA's Smart Location Database

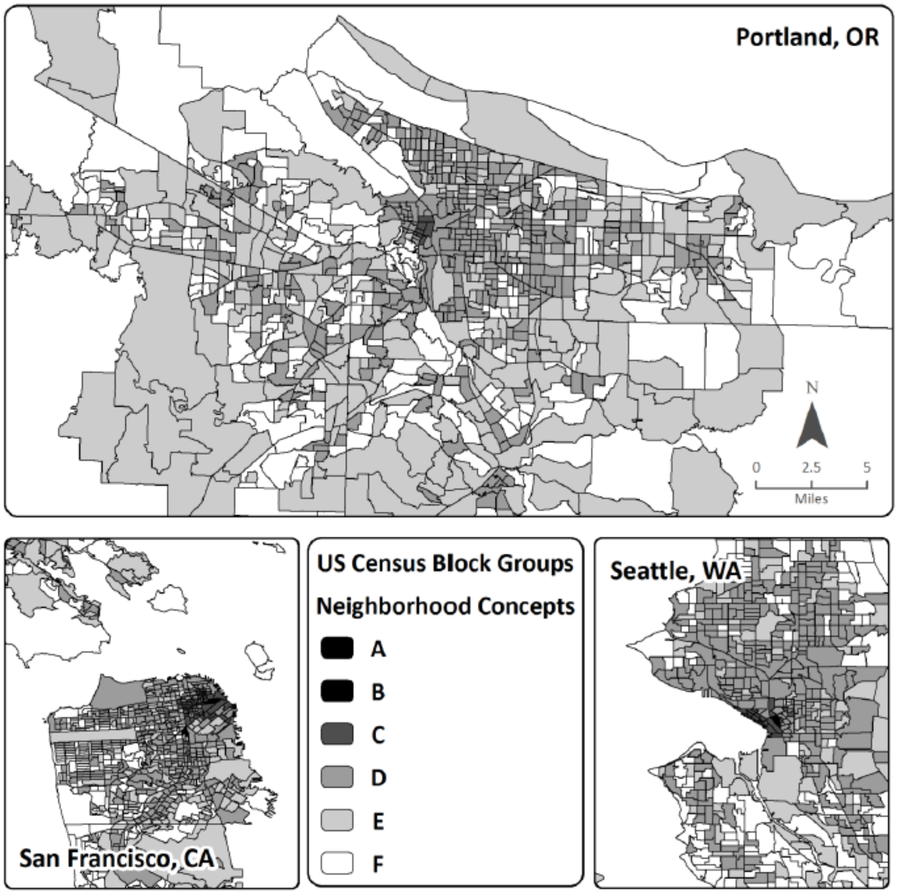


Figure 3.3: Neighborhood Concepts (Curran et al., 2014)

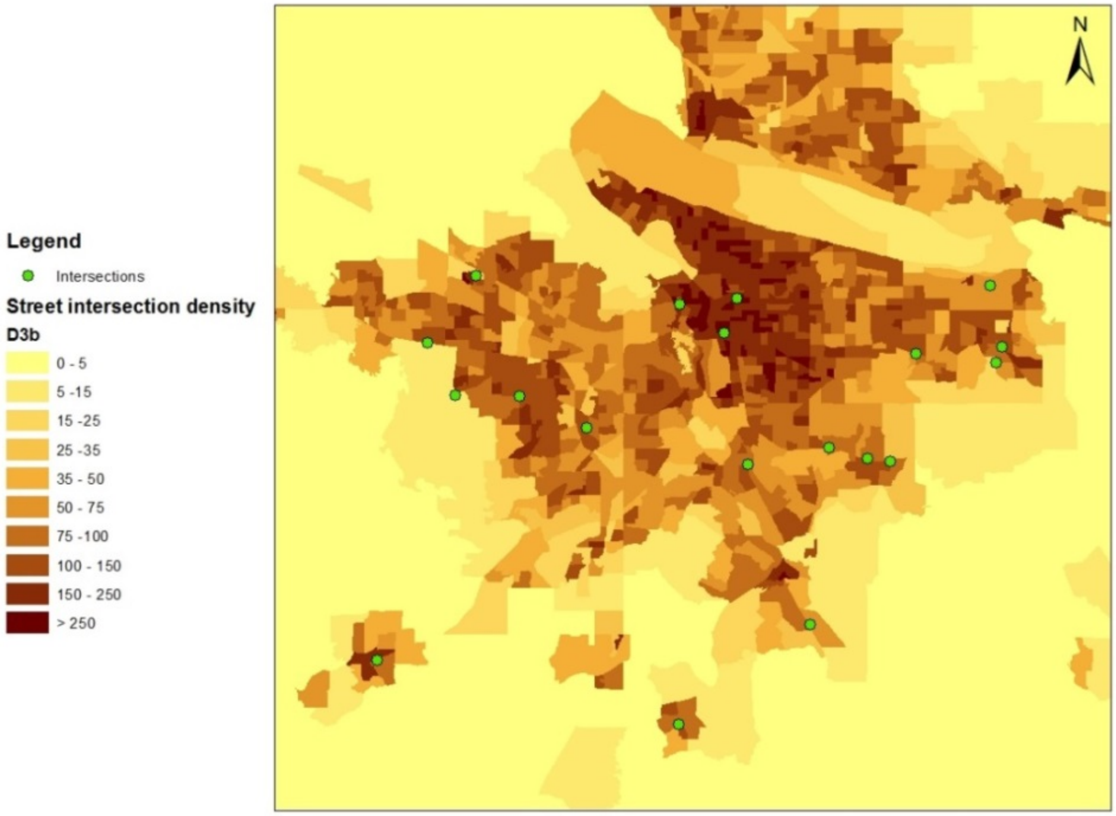


Figure 3.4: Sample Images Showing GIS Network Information

3.2.4 Volume

3.2.4.1 Motor Vehicle Traffic

Traffic volume data were assembled for all segments and intersections. A variety of sources were used to gather these data and convert to AADT values. These methods included internet searches and public record requests by the research team. The counts were either available as AADT (already factored and converted by reporting agency), ADT, or peak-hour counts. All the AADT data were calibrated to 2014 using growth factors obtained from the ODOT ATR station growth factors and following the methods described in Appendix B of ODOT SPR 756 “Improved Safety Performance Functions for Signalized Intersections” (Dixon et al, 2015).

3.2.4.2 Bicycle Traffic (STRAVA)

Systematic bicycle count data were not available. However, STRAVA is a mobile application that can track athletic activities including cycling and running through GPS (STRAVA, 2016a). When athletes and others are doing activities with the app open on their mobile devices, the STRAVA app will record the detailed information such as location and time. STRAVA has social network features by which users can communicate and interact with other users and groups.

STRAVA cooperated with several departments of transportations to create a research product called STRAVA Metro, which aggregates all of the cycling records from

STRAVA members. ODOT has purchased STRAVA Metro for research and project purposes (OSU and PSU researchers were allowed to obtain the data under this agreement). In this product, locations and time frames are aggregated into street networks and compiled to a shape file that can be used in GIS. The GIS map provides the information from cycling records for individual segments, including location, time, month, year, week or weekend, gender, and commuter or cyclist. Figure 3.5 illustrates the bike count on each link in Oregon.

Figure 3.5: STRAVA Cyclist Count in Oregon (left) and in Portland Metropolitan Area (right) (STRAVA, 2016b)

The bike volume can be roughly represented via the STRAVA bike count, but the accuracy of representation is one limitation of the data (even though STRAVA has differentiated the commuter count and cyclist count). ODOT TPAU has been doing tests to compare STRAVA counts to actual bike volumes. Some comparisons indicate that the results show the STRAVA count can represent 1% of total bike volume without considering the difference between commuter and cyclist. Other researchers have used STRAVA data in bicycle safety work and found significant results.

One issue with STRAVA data in GIS is that there can be multiple lines representing the same link on some segments. For example, Figure 3.6 shows that there are three count links (in red) on a bridge in the Portland downtown area, and each of them has a bike count of 3,473 bike trip/year, 5,264 bike trip/year, and 2,983 bike trip/year from top to bottom, respectively. This issue may come from the bike count assignment process since STRAVA built buffers around the GPS signal to assign bike counts to segments. Thus, we manually checked all of the links in our sample and only used the link with the highest bike count.

The unit of bike volume presented in the STRAVA data is bike count per year. However, bicycle ADT is a more commonly used variable in models. In this work, we made the very broad assumption that the STRAVA bike count represents around 1% of the actual bike volume on average. To convert the data to estimated bike volume so that these volumes could be interpreted in the risk scoring tool, we used the following formula to convert STRAVA bike count to bike ADT:

$$Bike\ ADT = \frac{STRAVA\ Bike\ Count}{1\% * 365} = \frac{STRAVA\ Bike\ Count}{3.65} \qquad \textbf{(3-1)}$$

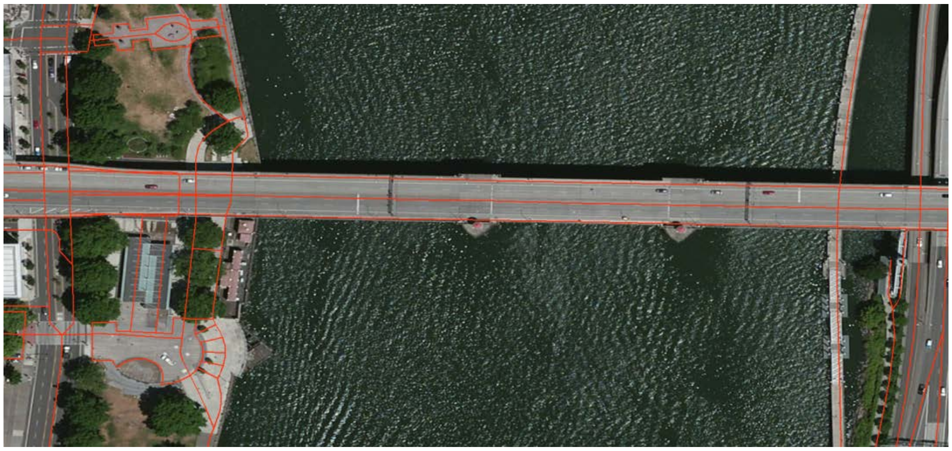


Figure 3.6: Multiple Bike Links on the Same Segment in Portland Downtown Area

3.2.5 Crash Data

Statewide geolocation of reported crashes in Oregon began in the 2007 data year. To merge and extract the crash data for safety analysis, locations of all the crossing enhancements were mapped in ArcGIS® using the latitude and longitude of the location at the center of the crossing. The crashes for each year were also imported into ArcGIS® using the latitude and longitude. Crash data were used from 2010-2013. The data includes severity and other detailed crash information.

One important step of this project is to link crash data to intersections and road segments. ArcGIS 10.2.2 was used to automatically assign crashes into segments and intersections. The spatial relation is used to determine whether a crash happened at an intersection or on a road segment, so it is crucial to identify how large an area an intersection or a segment can cover. An intersection can generally influence an area with a 250-foot diameter. Therefore, a buffer with a 125-foot radius was created at every middle point of intersections to build the influence area of an intersection. The crashes that occurred in this area are assigned to the corresponding intersection, shown in Figure 3.7. Similarly, a buffer with a 50-foot radius was created around a segment (line in ArcGIS) to build the influence area of the road segment, and crashes within the buffer were assigned to the segment.

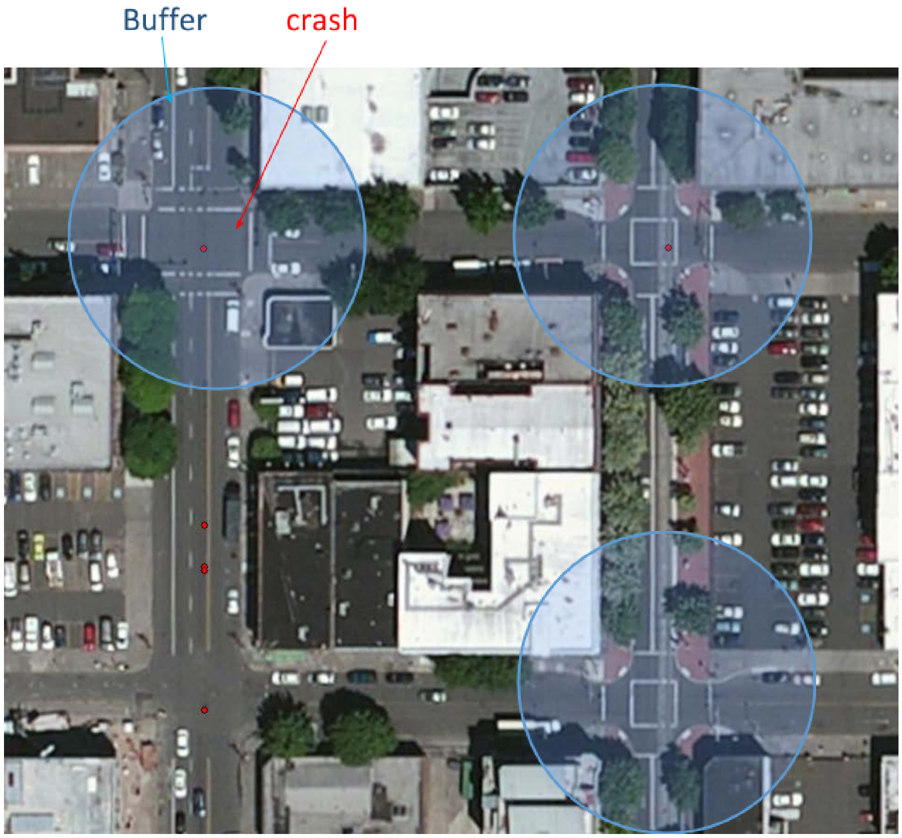


Figure 3.7: Intersection Crashes Counted within a 250-Foot Buffer Distance

3.3 SUMMARY

All of the data described in this chapter were assembled in a consistent format for use in the modeling effort. Descriptive analysis of the data are described in the next chapter.

4.0 DESCRIPTIVE DATA ANALYSIS

This chapter presents the descriptive analysis of crash, geometric and volume count data collected for the modeling process. Data on a total of 188 statewide segments and 184 intersections were collected. The chapter presents descriptive summaries for segment, intersection, and land use and count data.

4.1 SEGMENTS

There are 12 categorical variables (number of total traffic lanes, presence of TWLTL, presence of bike lanes, presence of sidewalk buffer, number of marked crosswalks, presence of on-street parking, presence of lighting, one or two-way travel, posted speed, presence of school, state highway and ODOT Region), and five continuous variables (segment length, width of bike lane (feet), number of driveways, number of intersections and maximum walking distance).

Table 4.1 presents the summary of categorical geometric-related variables of 188 selected segments. Figure 4.1 and Figure 4.2 show the same data but in a graphical format for easier inspection. More than half of the segments have only 2 traffic lanes and about 30% of the segment have 4 or more lanes. Most of the segment samples do not have two-way-left-turn lanes (69%), sidewalk buffer (76%), marked crosswalk (96%) or on-street parking (76%). About 40% of the segments do not have transit line go through it and 45% of the samples have only 1 transit line.

Table 4.2 presents the descriptive statistics for the continuous variables of segment geometric information. The number of driveways within the segment varies between a minimum of zero and a maximum of 26, with a mean of 4.36. The maximum number of transit lines through the segment is five but the mean number is only 0.81. Most segments do not have intersections, as the mean number of intersections is 0.28.

Table 4.1: Summary of Categorical Geometric Variables of Segments

Variable		Frequency	Percentage%
Number of traffic lanes (excluding two-way left turn lane)	1 Lane	2	1.06%
	2 Lane	112	59.57%
	3 Lane	16	8.51%
	4 Lanes or more	58	30.85%
Presence of two-way left-turn lane (TWLTL)	No	130	69.15%
	Yes	58	30.85%
Presence of bike lane	No	97	51.60%
	Yes	91	48.40%
Presence of sidewalk buffer	No	143	76.06%
	Yes	45	23.94%
Presence of marked midblock crosswalks within segment	No	181	96.28%
	Yes	7	3.72%
Presence of on-street parking	No	143	76.06%
	Yes	45	23.94%
Presence of lighting along segment	No	73	38.83%
	Yes	115	61.17%
Traffic direction	One-way	24	12.77%
	Two-way	164	87.23%
Posted speed limit (mph)	20	12	6.38%
	25	40	21.28%
	30	27	14.36%
	35	79	42.02%
	>35	30	15.96%
Presence of school area within 1000 feet from midpoint of segment	No	110	58.51%
	Yes	78	41.49%
State highway	State Highway	8	4.26%
	Not State Highway	180	95.74%
Region	1	66	35.11%
	2	56	29.79%
	3	28	14.89%
	4	19	10.11%
	5	9	4.79%
Number of transit lines go through the segment	0	76	40.43%
	1	85	45.21%
	2	19	10.11%
	3	5	2.66%
	>3	3	1.60%

Table 4.2: Summary of Continuous Geometric Variables of Segments

Variable	Mean	Standard Deviation
Segment length (ft.)	706.49	722.60
Width of bike lane (ft.)	5.47	1.02
Number of driveways	4.43	4.39
Number of intersections within segment	0.21	0.82
Maximum walk distance	350.09	325.48

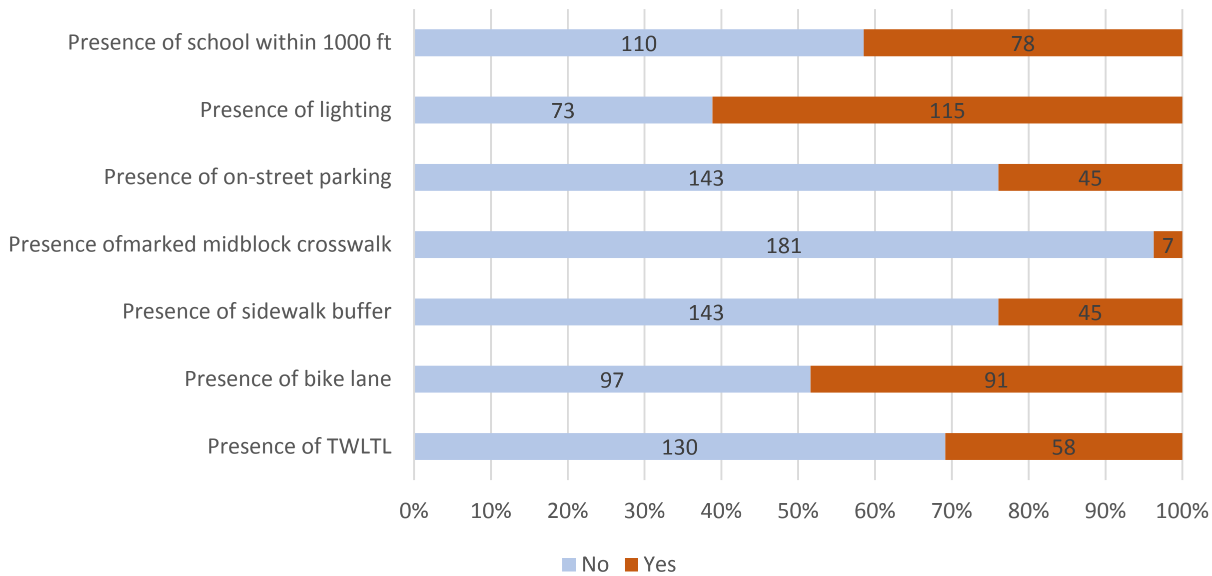


Figure 4.1: Summary of Segment Categorical Data in Each Category (No, Yes)

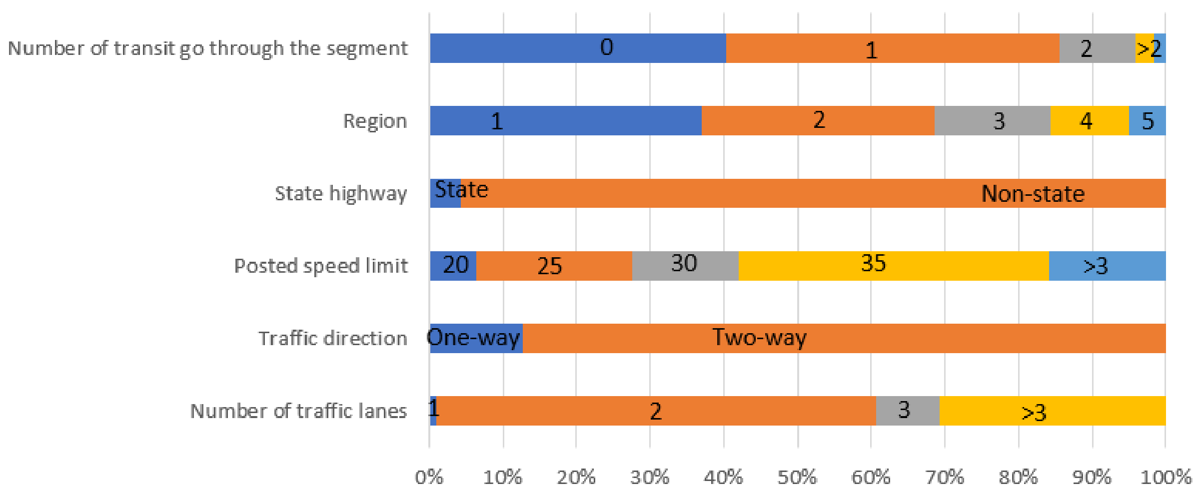


Figure 4.2: Graphical Summary of Segment Categorical Data, Part 2

4.2 INTERSECTIONS

Table 4.3 shows the summary of categorical geometric-related variables of intersection samples. Figure 4.3 and Figure 4.4 show the same data but in a graphical format for easier inspection. shows the summary of categorical geometric-related variables of intersection samples. Most of the samples are two-way, 4-leg signalized intersections with lighting facilities. Almost half of the intersections are located in the area where school presence within 1000 ft. Around 90% of the major roads are arterials while only 50% of the minor roads are arterials. The percentage of presence of left-turn lane and a right-turn lane on the major road and on the minor road are very similar, which is 72% and 27% for major road and 65% and 35% on the minor road. Table 4.4 shows the one continuous variable in the sample.

Table 4.3: Summary of Categorical Geometric Variables of Intersections

Variable		Frequency	Percentage %
Type of traffic control	Signal	172	93.48%
	4-Way Stop	9	4.89%
	Roundabout	3	1.63%
Presence of lighting	No	6	3.26%
	Yes	178	96.74%
Presence of school within 1000ft	No	103	55.98%
	Yes	81	44.02%
Intersection legs	4-Leg	157	85.33%
	3-Leg	27	14.67%
Major road. Posted speed limit (mph)	<=20	7	3.80%
	25	32	17.39%
	30	30	16.30%
	35	80	43.48%
	>35	35	19.02%
Major road. Travel direction	One-Way	30	16.30%
	Two Way	154	83.70%
Major road. Presence of left-turn lane	No	51	27.72%
	Yes	133	72.28%
Major road. Presence of right-turn lane	No	133	72.28%
	Yes	51	27.72%
Major road. Total number of traffic lanes	2	25	13.59%
	3	52	28.26%
	4	30	16.30%
	>4	77	41.85%
Major road. Presence of bicycle lanes	No	73	39.67%
	Yes	111	60.33%
Major road. Presence of median	No	165	89.67%
	Yes	19	10.33%
Minor road. Functional class	Arterial	92	50.00%
	Collector	92	50.00%
Minor road. Posted speed limit (mph)	<=20	12	6.52%
	25	88	47.83%
	30	21	11.41%
	35	44	23.91%
	>35	19	10.33%
Minor road. Traffic Direction	One-Way	16	8.70%
	Two Way	168	91.30%
Minor road. Presence of left-turn lane	No	64	34.78%
	Yes	120	65.22%
Minor road. Presence of right-turn lane	No	120	65.22%
	Yes	64	34.78%
Minor road. Total number of traffic lanes	2	43	23.37%
	3	90	48.91%
	4	26	14.13%
	>4	25	13.59%
Minor road. Presence of bicycle lanes	No	170	92.39%
	Yes	14	7.61%
Minor road. Presence of median	No	169	91.85%
	Yes	15	8.15%
Number of transit lines go through the intersection	0	34	18.48%
	1	90	48.91%
	2	49	26.63%
	3	9	4.89%

	>3	2	1.09%

Table 4.4: Summary of Continuous Geometric Variables of Intersection

Variable	Mean	Standard Deviation
Number of transit lines through intersection	1.36	1.18

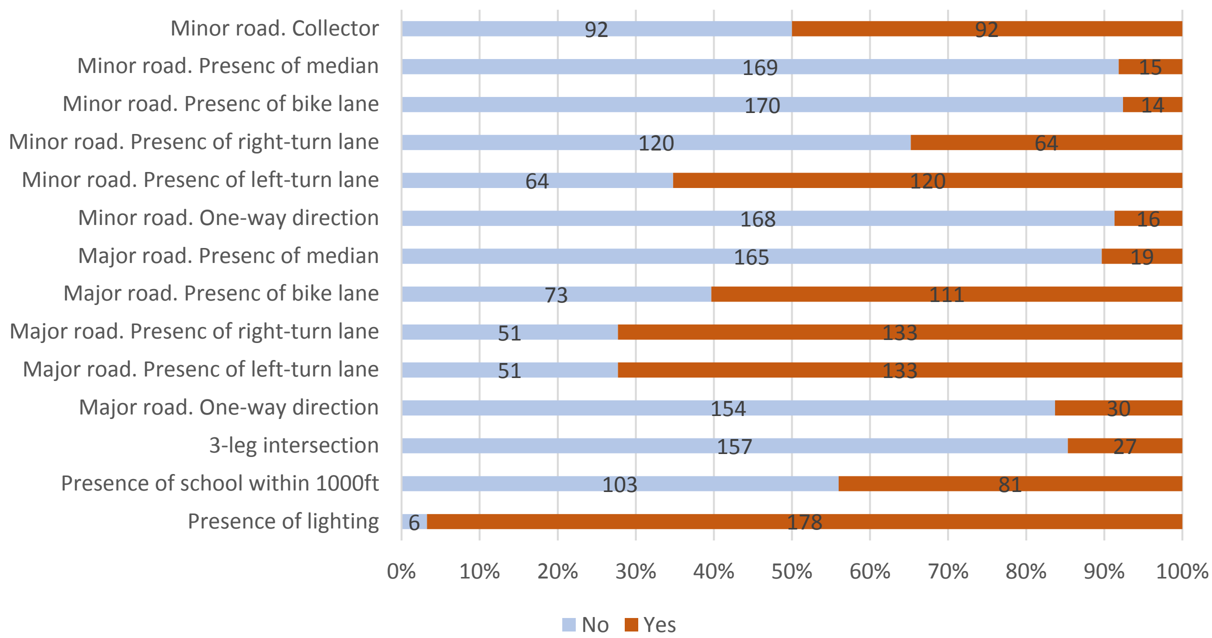


Figure 4.3: Summary of Intersection Categorical Data in Each Category (No, Yes)

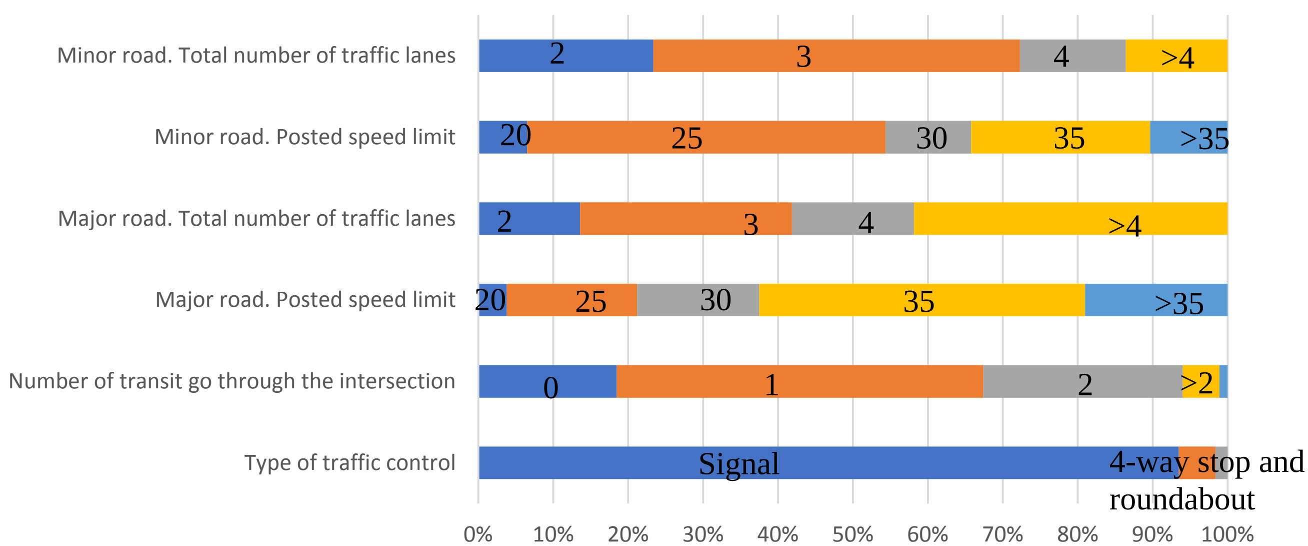


Figure 4.4: Summary of Intersection Categorical Data, Part 2

4.3 LAND USE

Table 4.5 presents the descriptive statistics for the land-use data of segments. The three-leg intersection, four-leg intersection, retail, total population density and household density are listed. All the data resources are from the U.S. Environmental Protection Agency's Smart Location Database (SLD).

Table 4.5: Summary of Continuous Land-Use Variables

Variable	Mean	Standard Deviation
Segment		
3-Leg intersection density (per square mile)	162.66	110.12
4-Leg intersection density (per square mile)	70.28	89.30
Retail density (per acre)	447.93	1314.31
Total population density (People per square mile)	4086.92	5333.08
Household density (per acre)	1916.90	4384.46
Household size	2.38	0.45
Intersection		
3-Leg intersection density (per square mile)	162.89	101.46
4-Leg intersection density (per square mile)	71.69	73.58
Retail density (per acre)	446.46	779.28
Total population density (People per square mile)	3826.29	2907.25
Household density (per acre)	1639.93	1675.44
Household size	2.32	0.45

Table 4.6 presents the summary of the only categorical land-use data, the neighborhood concept, which is assigned to the block-group size based on the accessibility to pedestrian and bicycle activities by using three measures of the built environment collected from nationally available data from 25 metropolitan areas. As described, the neighborhood concept has six levels, A~F, in which A is the best and F is the worst. Most segment and intersection samples are at level D and E (89% of segments and 84% of intersections). Few segment samples have Level C (4%) and Level F (7%). Intersections have a relatively higher proportion of level F (14%) and correspondingly less Level C (2%) than segment samples.

Table 4.6: Summary of Categorical Land-Use Variables

Variable		Frequency	Percentage %
Segment			
Neighborhood concept	C	3	1.60%
	D	69	36.70%
	E	81	43.09%
	F	35	18.61%
Intersection			
Neighborhood concept	C	1	0.54%
	D	72	39.13%
	E	83	45.11%
	F	28	15.22%

4.4 COUNT DATA

4.4.1 Motor Vehicle Traffic

All the AADT data are calibrated to 2014 using growth factors obtained from the ODOT ATR station growth factors. The range of AADT of segment samples is between 235 to 34,278 with the mean as 10,806 and the standard deviation as 7,607. AADT data is collected for both major roads and minor roads for intersection samples. For some signalized intersections, volume data could not be obtained. For these locations, the models developed for ODOT project SPR 756 were used to estimate the AADT of minor roads at signalized intersections. In the data sample, the minimum AADT of major roads is 839 and the maximum value is 50,240, with the mean as 14,080 and the standard deviation as 8143. The AADT of minor roads varies from 500 to 26,130, with the mean value as 7,648 and the standard deviation as 5,480.

Table 4.7: Summary of AADT

Intersection	Minimum	Maximum	Mean	Standard Deviation
Major road	837	50,240	14,080.34	8,143.53
Minor road	500	26,134	7,648.98	5,480.24
Segment	236	34,278	10,806.27	7,607.40

4.4.2 Bicycle Traffic (STRAVA)

STRAVA data is used as daily bike count data in the project. Daily bike volume of selected segments varies from minimum 0 to maximum 1,480 with mean value at 94.32 and 165.37 standard deviation. Intersection daily bike volume is counted as the summation of the volume of major road and minor road. The range is from a minimum 5 to maximum 7,020 with mean value 261.3 and standard deviation 641.2.

Table 4.8: Summary of Bike Volume Data

Location	Minimum	Maximum	Mean	Standard Deviation
Intersection	5	7,020	261.26	641.24
Segment	0	1,480	94.32	165.37

4.5 CRASH DATA

Bicycle and pedestrian crashes share very similar distribution patterns on segments. There are no crashes present from 2009 to 2013 on many segments in the sample. In total, there were 113 pedestrian and 100 bicycle crashes on the segments and 108 pedestrian and 130 bicycle crashes at the intersections. Around 20% of the segments have only one bicycle or pedestrian crash. Fewer segments, around 10%, have two crashes during five years and only a few have more than two crashes. Intersections have similar patterns for pedestrian and crash frequency. Most of the intersections selected do not have crashes and few intersections had more than two crashes from 2009 to 2013. Others have one or two pedestrian or bicycle crashes. Figure 4.6 shows the bar chart of the intersection crash frequency distribution.

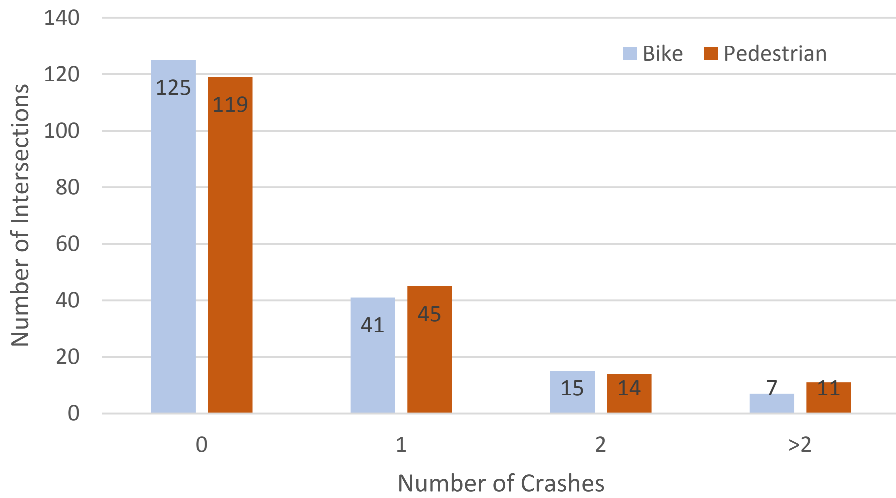


Figure 4.5: Number of Segments by Crash Frequency

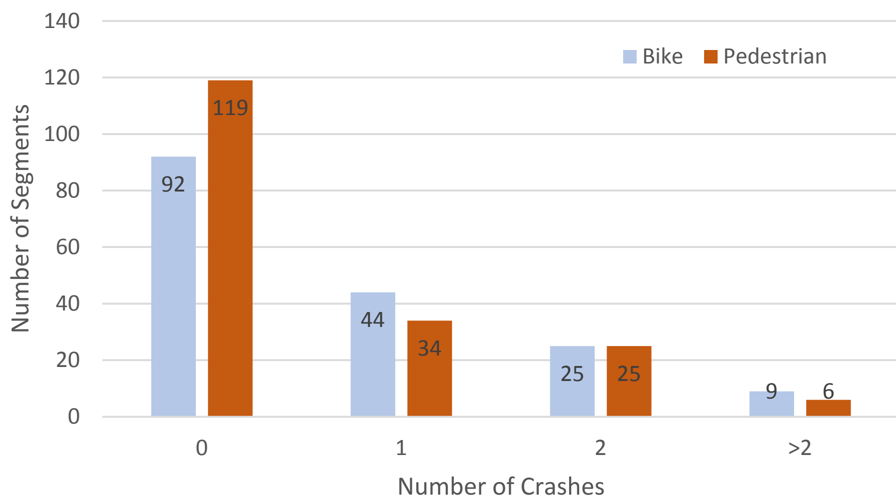


Figure 4.6: Number of Intersections by Crash Frequency

Pedestrian and bicycle crash severity levels on segments show similar distribution patterns.

Figure 4.7 indicates that the majority proportion of severity levels in the five years are Injury B and Injury C, which are 49% and 39% for pedestrians and 53% and 34% for bicycles. Injury A severity crashes make up about 10% of the total crashes of both pedestrians and bicycles (10% for pedestrians and 9% for bicycles). There was only one fatal bicycle crash and no pedestrian fatal crashes within the five years on selected segments. PDO crashes also share a very low proportion (2% for pedestrian and 3% for bicycle crashes.

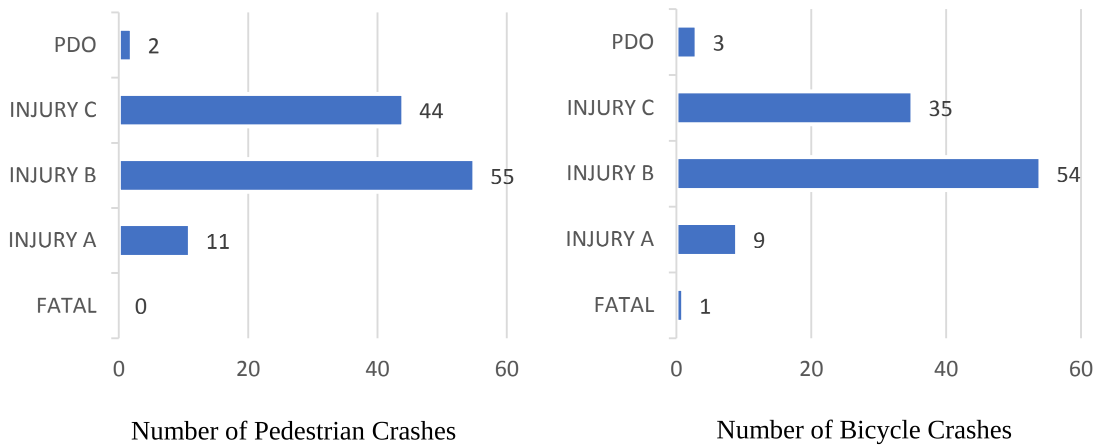


Figure 4.7: Segment Crash Severity Level Distribution, Five Years

Severity levels of pedestrian and bicycle crashes at intersections have a similar distribution pattern as segment crashes. Injury B and Injury C levels make up almost 90% of the total crashes. However, pedestrians have more Injury C-level crashes (48%) than Injury B-level crashes (39%), and bicycles have the opposite pattern – less Injury C crashes (31%) than Injury B crashes (59%). There are 8% of pedestrian intersection crashes and 7% of bicycle intersection crashes that are Injury A level, and 1% and 3% are PDO crashes, respectively. Four pedestrian fatal crashes were reported at intersections from 2009 to 2013.

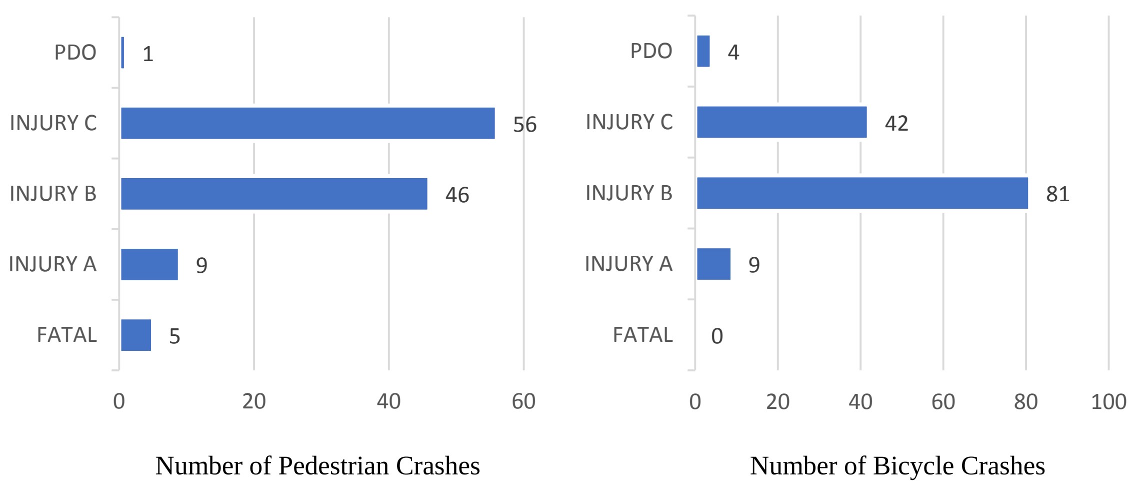


Figure 4.8: Intersection Crash Severity Level Distribution, Five Years

5.0 METHODOLOGY

This chapter presents the statistical models used to create crash occurrence and severity models for pedestrians and bicycles on segments and intersections. The research team explored a multitude of different statistical models before identifying the logistic model as preferable. This chapter summarizes the methods used to build the models. This method for developing the risk scoring tool is presented in Chapter 7.

5.1 LOGISTIC REGRESSION MODEL

As highlighted in Chapter 2 (Literature Review), logistic regression models are frequently used in estimating the risk factors of bicycle and pedestrian crashes and severity. Lenguerrand et al. (2006); Parkin et al. (2007); Kim et al. (2007); Eluru et al. (2008); Boufous et al. (2012); Schepers and Brinker (2011); Pai (2011), etc. used the logistic model in their research to establish either bicycle or pedestrian risk models, although the results are not consistent (Lenguerrand, Martin, and Laumon, 2006). The inconsistency is largely driven by the difference in geometric factors, weather conditions, and other factors.

Logistic regression was initially proposed by Cox in 1958 (Walker, 1967) to measure the categorical dependent variable (Y) and multiple independent variables (X) by using the logistic function. A standard logistic regression function is given by Equation 5-1 and the basic shape is shown in Figure 5.1. A logistic model is also called logit model, and the difference is only the format in which logistic function is the inverse of a logit function. A binomial logistic regression is often used when the dependent variable has only two levels (0 or 1), which are frequently used to represent crash data in the following form:

$$p(x) = \frac{L}{1+e^{-k(x-x_0)}} \qquad \textbf{(5-1)}$$

Where:

e = the natural logarithm base (2.71828),

x_0 = the x-value of the sigmoid function's midpoint

L = the curve's maximum value, L, equals 1 in the binomial model

k = the steepness of the curve

p (x) = the probability of the dependent variable

The binomial logistic model is often used when the independent variable is in binary format (0 or 1). The binomial independent variable can be interpreted as 1 represents an element (or situation) exists whereas 0 represents the opposite.

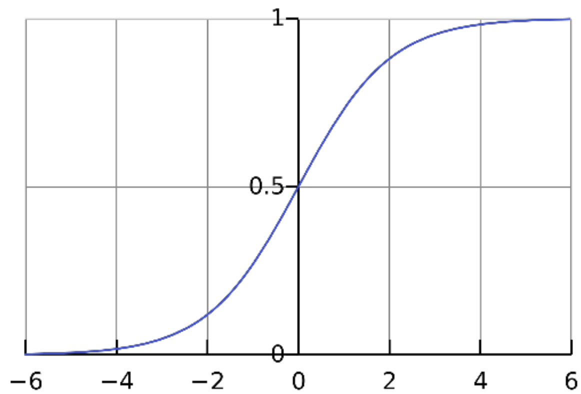


Figure 5.1: Standard Logistic Regression

5.1.1 Coefficients

Logistic regression modeling will generate a coefficient for each variable. This value represents the slope or rate of change of the dependent variable per unit of change in the independent variable (either increasing or decreasing) when other variables are held as constants (Al-Ghamdi, 2002). However, a proper interpretation of logistic model coefficient is the slope which represents the change in the logit for a change of one unit of the independent variable. Odds ratios, explained later, are used to interpret the model coefficients.

5.1.2 Model Building - Variable Selection Process

The research team developed crash occurrence and crash severity models for bicycles and pedestrians on segments or at intersections (eight total models). To develop the models, a backward and forward stepwise procedure was used that tested all combinations to find those significant variables. Backward stepwise started with a full model (all variables included) and insignificant variables were eliminated from the model in an iterative process until all variables are significant. A forward stepwise started with a reduced model (only variable “intercept” included) and significant variables were added into the model in an iterative process until no significant variables can be added further. This process was applied automatically in R (R Core Team, 2013). In addition, we manually tested a subset of variables that might be significant based on the project team’s engineering judgments and relevant project experience.

5.1.3 Interpreting Model Outputs

A coefficient for an independent variable represents the slope or rate of change of the dependent variable per unit change in the independent variable (Al-Ghamdi, 2002). In other words, it represents how much the dependent variable changes when the corresponding variable changes per unit. However, the correct interpretation of a logistic model is different from a standard linear regression model because the link function between independent variables and the mean of dependent variables of the logistic model is different from the standard linear model. Table 5.1 shows the different link functions.

Table 5.1: Link Functions of Standard Linear and Logistic Model

Model Type	Link Function
Standard linear model	g(x) = x
Logistic model	g(x) = log(x/(1-x))

The link function of the standard linear model, in Table 5-1 indicates that the change of the dependent variable has “direct” influence on the mean of the dependent variable by multiplying the corresponding coefficient. Whereas, in a logistic model, there is a relatively complicated link function, indicating that the change of one unit of the dependent variable will change the log odds – g(x) = log(x/(1-x)) – of the dependent variable. More detailed interpretation of logit odds can be found in Section 6.3.

In the modeling results tables in the following sections, the significance level represents how much evidence we have to reject the null hypothesis. In other words, in this logistic model, it represents how much evidence we have to state the influence on log odds of change in an independent variable which is not equal to zero. The R outputs the p-value and standard error for each coefficient. The p-value is the probability of obtaining a result equal or more extreme than observed situations when the null hypothesis is true (Ramsey and Schafer, 2013). We considered any variables with p-values less or equal than 0.05 as significant variables.

Meaningful interpretation of coefficients in logistic models relies on how to interpret the difference between two odds (Al-Ghamdi, 2002). Equation 5-2 represents the change in the logit for a change of one unit of the independent variable, so the appropriate interpretation of the coefficient in a logistic model depends on the meaning of differences between two logits. The odds ratio, which is shown in Equation 5-3, can provide a foundational interpretation for all logistic models. If the odds ratio is greater than 1, it represents the likelihood of an event indicated in nominator (P in Equation 5-3 is greater than the likelihood of an event indicated in the denominator (1-P); if the odds ratio is less than 1 (from 0 to 1), it represents the opposite.

$$Ln\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k \qquad \textbf{(5-2)}$$

$$Odds\ Ratio = \psi = \frac{P}{1-P} = e^{\beta_0 + \beta_1 x_1 \ldots \beta_k x_k} \qquad \textbf{(5-3)}$$

Where:

P - probability that Y=1 (crash or severe crash occurs in the model)

β_i- coefficient of the independent variable,

ψ – odds ratio will be β_i times higher than before when x_i increase 1 unit.

The exponent of this change defines the odds ratio as the ratio of the odds ($\frac{\pi}{1-\pi}$) in which the independent variables will be present to the odds that it will not be present. It should be noted that an odds ratio which is greater than 1 indicates that this variable increases the probability of the odds. In this case, the odds ratio represents the probability of the occurrence of crashes or the severity of crashes Interpreting odds ratio for categorical variables requires additional calculation. As an example, we calculate the odds ratio for the "Presence of TWLTL" variable in the pedestrian segment crash occurrence model. The base condition of this variable is "No TWLTL" and the coefficients of the other level indicate the odds differences compared to the segment with no TWLTL. According to the results of this model, the differences should be calculated as follows, given other variables are held as constants:

Logit (Presence of Crash/ No TWLTL) $=\beta_0 + \beta_{traffic\_direction} + \beta_{on\_street\_parking} + \beta_{speed} + \beta_{total\_population} + \beta_{total\_traffic\_lane}$

Logit (Presence of Crash/ Presence of TWLTL) $=\beta_0 + \beta_{traffic\_direction} + \beta_{on\_street\_parking} + \beta_{speed} + \beta_{total\_population} + \beta_{total_{traffic_{lane}}} + \beta_{TWLTL}$

Logit differences $=\beta_0 + \beta_{traffic\_direction} + \beta_{on\_street\_parking} + \beta_{speed} + \beta_{total\_population} + \beta_{total_{traffic_{lane}}} + \beta_{TWLTL}$ - $(\beta_0 + \beta_{traffic\_direction} + \beta_{on\_street\_parking} + \beta_{speed} + \beta_{total\_population} + \beta_{total\_traffic\_lane}) = \beta_{TWLTL}$

Thus, the odds ratio $\psi = e^{\beta_{TWLTL}} = e^{1.071} = 2.92$ indicates that the odds of crash occurrence at a segment with a two-way left-turn lane are 2.92 times higher than the odds of a segment with no two-way left-turn lanes.

5.2 MODELS

5.2.1 Crash Occurrence

A logistic model was applied to the pedestrian data using crash frequency (crash or not) as the dependent variable (Y). Other geometric data, land-use data, and traffic data were independent variables. Models were developed for:

- Pedestrian, Segment
- Pedestrian, Intersection
- Bicycle, Segment
- Bicycle, Intersection

5.2.2 Crash Severity

The same logistic modeling method was used to develop the crash severity models. To create the dependent variable, grouping by crash severity (Fatal, Injury A, Injury B, Injury C, and PDO) was required. A multinomial logistic model could be a better fit; however, the number of crashes in each category such as "Fatal" was too few to meet the lowest requirement of a multinomial logistic model. Thus, the five crash-severity levels are regrouped into two levels so that the binomial logistic model can be used. Models were attempted for:

- Pedestrian, Segment
- Pedestrian, Intersection
- Bicycle, Segment
- Bicycle, Intersection

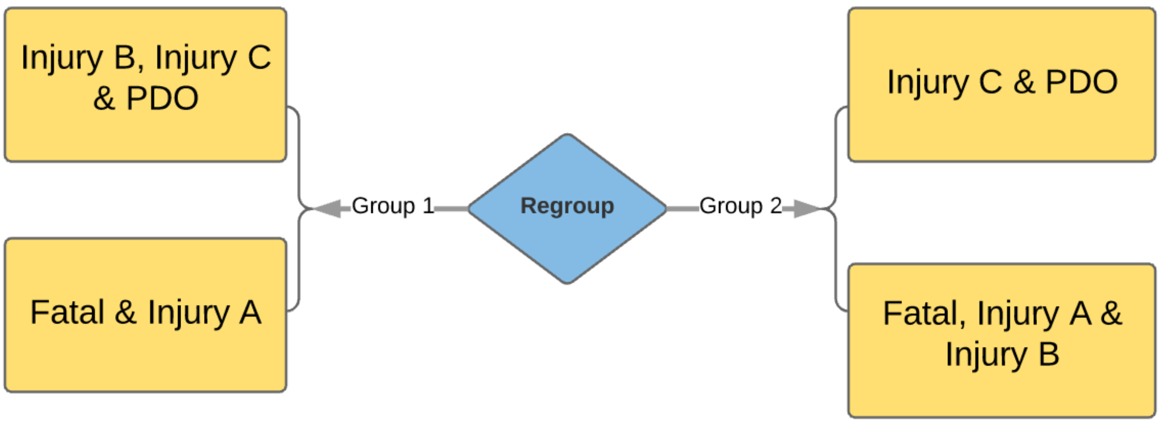


Figure 5.2: Two Regroupings of Crash Severities

6.0 MODELING RESULTS

This chapter documents the risk modeling results and analyses for the eight models. Pedestrian models are presented followed by bicycle models.

6.1 PEDESTRIAN MODELS

This section presents the results of four different pedestrian models: 1) pedestrian segment crash occurrence model; 2) pedestrian intersection crash occurrence model; 3) pedestrian segment crash severity model, and 4) pedestrian intersection crash severity model.

6.1.1 Crash Occurrence

As discussed in the methods section, a combined backward and forward stepwise method was used to determine the significant variables to be included in the final model. The project team also attempted models using a subset of plausible variables identified in the literature by engineering judgments and previous relevant project experience. The variables of the final pedestrian segment crash occurrence model included: traffic directions, the presence of on-street parking, the presence of TWLTL, posted the speed limit, total population density and number of total traffic lanes. The coefficients, standard error, p-value, and significance are shown in Table 6.1. The table also includes the odds ratio calculated from the coefficients of the pedestrian segment crash occurrence model. In the model, the base condition of categorical variables is one-way direction, no presence of on-street parking and no TWLTL.

Table 6.1: Model 1: Pedestrian Segment Crash Occurrence Model Results

Variable	Coefficients	Standard Error	P-value	Significance	Odds Ratio
Travel direction (one-way or two-way) (One-way is base)	-1.289	1.018	0.001	***	0.028
Presence of on-street parking	1.337	0.514	0.012	*	3.808
Presence of two-way left-turn lane	1.071	0.384	0.005	**	2.918
Posted speed limit (mph)	0.047	0.027	0.082	.	1.048
Total population density (people per square mile)	0.0017	0.00006	0.006	**	1.002
Number of traffic lanes (excluding two-way left-turn lane)	0.370	0.016	0.023	*	1.447
Null deviance (measure the goodness of fit of a logit model): 249.16 on 188 degrees of freedom					
Residual deviance (measure the goodness of fit of a logit model): 204.62 on 182 degrees of freedom					
AIC: 218.62					

Significant Code: 0 ‘***’; 0.001 ‘**’; 0.01 ‘*’; 0.05 ‘.’; 0.1 ‘ ’.

To help interpret the results of a binomial model, Figure 6.1 shows the number of segments with and without pedestrian crashes. The left panel of Figure 6.1 shows more crashes occurred on the one-way traffic segments, whereas the right panel of the figure shows fewer crashes occurred on the two-way traffic segments. The negative coefficient (-1.289) reflects that the two-way traffic segment has a smaller pedestrian crash probability than the one-way traffic segment. This is also clear in the odds ratio (0.0276), which indicates the probability of crashes is low.

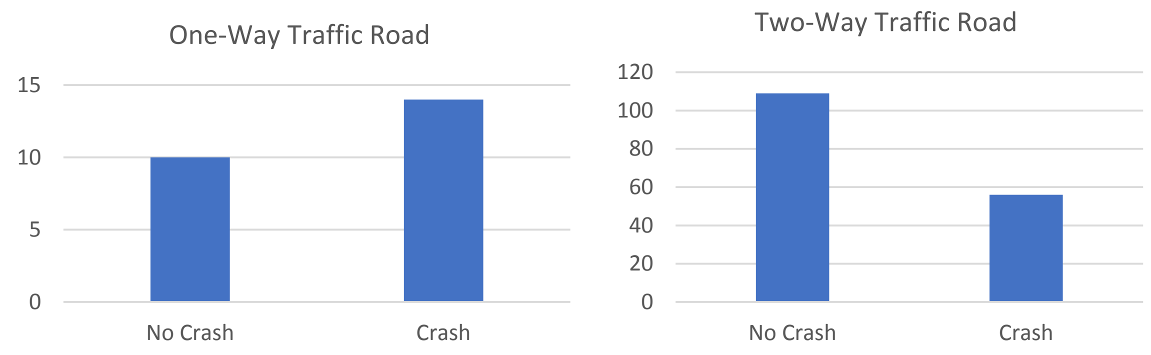


Figure 6.1: Occurrence of Crashes on One-way Roads and Two-way Roads

The effect of increased posted speeds on the probability of crashes is in the expected direction (an increase in speed increases the probability of a crash outcome). The continuous nature of how the speed was modeled also means that the effect is not linear and that larger speeds have a very significant effect on the overall probability prediction. The presence of a TWLTL is also positive. The hypothesis is that this additional width and turning conflicts add additional risk for pedestrians. It could also be capturing the effect of higher volumes associated with TWLTL roads. The modeling results for two-way traffic also likely reflect some association with more traffic and lanes on one-way streets. Finally, the total population density is associated with the potential pedestrian activity density in this area. We hypothesize that as the connectivity of the population density increases, the crash probability increases.

For intersections, the final pedestrian crash occurrence model includes the following variables: total population density; the number of transit lines through the intersection; the number of major road right-turn lanes; the major road AADT in 2014; the presence of a median on the minor road; and the number of right-turn lanes the minor road. The coefficients, standard error, p-value, and significance are shown in Table 6.2. Most of the significant variables were modeled as continuous variables and only the presence of a median on major roads is a categorical variable, with “no presence” as the base condition.

Table 6.2: Model 2: Pedestrian Intersection Crash Occurrence Model Results

Variable	Coefficients	Standard Error	P-value	Significance	Odds Ratio
Total population density (people per square mile)	0.00024	0.000072	0.000	***	1.000
Number of transit lines through intersection	0.383	0.208	0.065	.	1.467
Major road, number of right-turn lanes	0.784	0.432	0.070	.	2.190
Major road, AADT 2014	0.000063	0.000023	0.005	**	1.000
Minor road, presence of median	-1.260	0.664	0.058	.	0.284
Minor road, number of right-turn lanes	-1.312	0.440	0.003	**	0.269
Null deviance (measure the goodness of fit of a logit model): 238.99 on 183 degrees of freedom					
Residual deviance (measure the goodness of fit of a logit model): 195.40 on 177 degrees of freedom					
AIC: 209.4					

Significant Code: 0 '***'; 0.001 '**'; 0.01 '*'; 0.05 '.'; 0.1 ' '.

The variables for population density, major road AADT and the number of transit lines all relate to the exposure of pedestrians at intersections. All of the variables show increased probability of a crash with increases in the variable. The three remaining significant variables relate to right-turn lanes on the major and minor roads and the presence of a median on the minor road. The presence of a median has a negative sign on the coefficient, meaning that it reduces the probability of pedestrian crashes. One interpretation is that the median space might provide a refuge on long pedestrian crossings. The right-turn lane numbers differ by whether they are on the major or minor road. On the major road, the number of right-turn lanes increases the crash probability. The presence of right-turning traffic is a risk for pedestrians (which this variable is a proxy for). Finally, the right-turn lane would increase the pedestrian's crossing distance. On the minor road, the presence of the right-turn lanes has the opposite effect. This variable is possible capturing other effects and it is not easy to explain the difference.

6.1.2 Crash Severity

Table 6.3 and Table 6.4 summarize the results of the pedestrian segment and intersection crash severity model. The modeling explored the independent variable with the two recoded crash groups.

0 = Presence of property damage only (PDO) or Injury C-level crash

1 = Presence of a fatal, Injury A-level and Injury B-level crash

and:

0 = Presence of property damage only (PDO), Injury C-level or Injury B-level crash

1 = Presence of fatal, Injury A-level crash

The second grouping produced fewer significant variables, thus only the first group is reported here. The pedestrian crash severity model has only three significant variables: retail density, AADT, and the presence of lighting. The sign for all variables is not in the expected direction (increases in the variables decrease the crash probability). The intersection severity model has many of the same significant variables as the crash occurrence model, but again many of the coefficient estimates are not expected (e.g., the AADT estimate is negative). We conclude that our modeling dataset was not sufficient to estimate models based on crash severity.

Table 6.3: Model 3: Pedestrian Segment Crash Severity Model Results

Variable	Coefficient s	Standard Error	P-value	Significance	Odds Ratio
Retail density (per acre)	-0.000059	0.000023	0.011	*	0.999
AADT 2014	-0.000019	0.0000072	0.009	**	0.999
Presence of lighting	-0.219	0.106	0.041	*	0.803
Null deviance (measure the goodness of fit of a logit model): 28.475 on 117 degrees of freedom					
Residual deviance (measure the goodness of fit of a logit model): 24.732 on 114 degrees of freedom					
AIC: 160.48					

Significant Code: 0 ‘***’; 0.001 ‘**’; 0.01 ‘*’; 0.05 ‘.’; 0.1 ‘‘.

Table 6.4: Model 4: Pedestrian Intersection Crash Severity Model Results

Variable	Coefficients	Standard Error	P-value	Significance	Odds Ratio
Major road, number of total traffic lanes (including left- and right-turn lanes)	1.905	0.466	0.000	***	6.721
Three-leg intersection density (per square mile)	0.027	0.0076	0.001	***	1.027
Four-leg intersection density (per square mile)	-0.029	0.0086	0.001	***	0.971
Total population density (per square mile)	-0.00073	0.00024	0.003	**	0.999
Minor road, presence of bicycle lanes	-1.4485	0.674	0.032	*	1.00098
Minor road, number of right-turn lanes	2.216	0.977	0.023	*	0.235
Minor road, number of total traffic lanes	-0.516	0.282	0.068	.	9.171
Major road AADT 2014	-0.00013	0.000049	0.010	*	0.597
Household density	0.00098	0.00041	0.016	*	0.999
Major road, number of right-turn lanes	-1.533	0.793	0.053	.	0.216
Minor, road, presence of one-way traffic	1.957	0.954	0.040	*	7.077
Null deviance (measure the goodness of fit of a logit model): 149.68 on 107 degrees of freedom					
Residual deviance (measure the goodness of fit of a logit model): 104.21 on 96 degrees of freedom					
AIC: 128.21					

Significant Code: 0 ‘***’; 0.001 ‘**’; 0.01 ‘*’; 0.05 ‘.’; 0.1 ‘‘.

6.2 BICYCLE MODELS

This section provides the modeling results of four different bicycle models: 1) bicycle segment crash occurrence model; 2) bicycle intersection crash occurrence model; 3) bicycle segment crash severity model; and 4) bicycle intersection crash severity model.

6.2.1 Crash Occurrence

A combined backward and forward stepwise method was used to determine the significant variables to be included in the final model. The project team manually tested a subset of significant variables that it deemed to be important by engineering judgment and previous relevant project experience. The variables of final bicycle segment crash occurrence model include the presence of crossing (no crossing is the base condition); AADT (factored to 2014); three-leg intersection density; and bike volume (per day). The coefficients, standard error, p-value, significance and odds ratios are shown in Table 6.5.

Unlike all the other crash occurrence model, very few variables were found to be significant in the model. The research team selected a final model that was based on judgment and lower confidence level threshold for variable inclusion. The model includes an exposure metric for bicycles per day and the sign is expected (as bicycles per day increase then probability also increases). The included variables significant variables also include the land-use variables related to intersection-type density. Three-leg intersection density (we hypothesize is associated with less connectivity) is associated with a positive increase in crash probability. The presence of crossings decreases the probability of bicycle crashes on the segments. We hypothesize that the presence of pedestrian crossings is related to the overall design of the roadway (i.e. a more non-motorized user-friendly character). Vehicle volume, represented by AADT in the model, has the positive coefficients as expected, indicating that high vehicle volume could lead to high risk for bicyclists on this segment.

Table 6.5: Model 5: Bicycle Segment Crash Occurrence Model Results

Variable	Coefficients	Standard Error	P-value	Significance	Odds Ratio
Presence of marked midblock crosswalks within segment	-1.207	-3.662	0.000976	***	0.2991
AADT 2014	0.00003187	0.00002124	0.1336		1.00003187
Three-leg intersection density (per square mile)	0.002087	0.001486	0.1602		1.002089
Bicycles per day (STRAVA)	0.001007	0.000994	0.3110		1.0010012
Null deviance (measure the goodness of fit of a logit model): 240.60 on 188 degrees of freedom					
Residual deviance (measure the goodness of fit of a logit model): 220.35 on 183 degrees of freedom					
AIC: 230.35					

Significant Code: 0 ‘***’; 0.001 ‘**’; 0.01 ‘*’; 0.05 ‘.’; 0.1 ‘ ’.

A similar backward and forward modeling exercise was conducted to estimate the model for intersections. At the intersection level, the bicycle models the significant variables in the final bicycle intersection crash occurrence model include bicycles per day, the number of transit stops, the minor road functional class, minor road total traffic lanes, and minor road right-turn lanes. The coefficients, standard error, p-value, significance and odds ratios are provided in Table 6-6. Clearly, the number of bicycles per day capture the increased exposure as volumes increase. The number of transit stops indicate a presence of other road users and possibly additional interactions with bus traffic. The number of lanes on the minor road can be interpreted as increasing the total intersection size.

Table 6.6: Model 6: Bicycle Intersection Crash Occurrence Model Results Table

Variable	Coefficients	Standard Error	P-value	Significance	Odds Ratio
Bicycles per day (STRAVA)	0.00146	0.00024	0.0368	*	1.001
Number of transit stops	0.3507	0.1924	0.0683	.	1.420
Minor functional class (arterial as base)	-0.9096	0.3585	0.0112	*	0.4027

Minor road, total number of traffic lanes	0.49698	0.2067	0.0231	*	1.644
Minor road, presence of right-turn lane	-0.7056	0.3581	0.0488	*	0.4938
Null deviance (measure the goodness of fit of a logit model): 232.04 on 167 degrees of freedom					
Residual deviance (measure the goodness of fit of a logit model): 200.04 on 161 degrees of freedom					
AIC: 214.04					

Significant Code: 0 '***'; 0.001 '**'; 0.01 '*'; 0.05 '.'; 0.1 ' '.

6.2.2 Crash Severity

Table 6.7 and Table 6.8 summarize the results of bicycle segment and intersection crash severity models. Since the model is binomial, the dependent variable is assigned as 0 or 1. Zero represents the property damage only (PDO) and Injury C-level crashes. One represents the fatal, Injury A-level and Injury B-level crashes. Table 6.7 shows the odds ratio of the bicycle segment crash severity model. The crosswalk combination type is categorical data with four levels: A, B, C and N. Level A is referred to as the "base condition" in this model. The other three levels all have negative coefficients, which lead to odds ratios of less than 1. The sidewalk buffer width is a continuous variable with an odds ratio of 1.02293 for each increase of 1 foot of width. Due to poor model specification and results, we conclude that the models that try to capture severity are not adequate.

Table 6.7: Model 7: Bicycle Segment Crash Severity Model Results

Variable	Coefficients	Standard Error	P-value	Significance	Odds Ratio
Cross combination type B (type A is base)	-0.26041	0.10743	0.0172	*	0.77074
Cross combination type C (type A is base)	-0.37745	0.27695	0.1761		0.68561
Cross combination type N (type A is base)	-0.30769	0.13241	0.0222	*	0.73514
Width of sidewalk buffer (ft.)	0.02268	0.01284	0.0804	.	1.02293
Null deviance (measure the goodness of fit of a logit model): 23.703 on 100 degrees of freedom					
Residual deviance (measure the goodness of fit of a logit model): 20.587 on 96 degrees of freedom					
AIC: 137.99					

Significant Code: 0 '***'; 0.001 '**'; 0.01 '*'; 0.05 '.'; 0.1 ' '.

Table 6.8: Model 8: Bicycle Intersection Crash Severity Model Results

Variable	Coefficients	Standard Error	P-value	Significance	Odds Ratio
Total population density (per square mile)	-0.0010641	0.0002584	3.83e-05	***	0.99905
Household density (per acre)	0.0017755	0.0004257	3.03e-05	***	1.001621
Household size	3.3000403	0.8386549	8.32e-05	***	20.65588
Minor road, number of right-turn lanes	-0.7263629	0.4346123	0.09466	.	0.48366
Major road, number of left-turn lanes	-1.5006639	0.5498043	0.00634	**	0.216752
Minor road, number of left-turn lanes	1.3020888	0.4949559	0.00852	**	3.575124
Null deviance (measure the goodness of fit of a logit model): 169.31 on 139 degrees of freedom					
Residual deviance (measure the goodness of fit of a logit model): 132.27 on 133 degrees of freedom					
AIC: 146.27					

Significant Code: 0 '***'; 0.001 '**'; 0.01 '*'; 0.05 '.'; 0.1 ''.

6.3 SUMMARY

The research team developed logistic regression models for both crash occurrence (crash or not) and crash severity models. The models related to crash severity were not robust and did not prove useful. This is most likely due to the few segments and intersections with severe crashes in the dataset. The crash occurrence models produced more plausible models and significant variables, including a blend of exposure and geometric/operational variables hypothesized to relate to risk. Table 6.9 and Table 6.10 show the list of the significant variables identified in the modeling effort. The table highlights that there are a number of common explanatory variables. A comparison between Table 6.9 and Table 6.10 shows that generally, the bicycle models capture fewer significant variables than those pedestrian models, which is consistent with other studies. The crash occurrence models are converted to risk scoring tools in the next chapter

Table 6.9: Summary of Significant Variables in the Pedestrian Models

Crash Occurrence		Crash Severity	
Model 1 Segment	**Model 2 Intersection**	**Model 3 Segment**	**Model 4 Intersection**
Traffic directions (one-way is base)	Total population density (per square mile)	Retail density (per acre)	Major road. Presence of total traffic lanes
Presence of on-street Parking	Number of transit lines go through the segment	AADT 2014	3-leg intersection density (per square mile)
Presence of TWLTL	Major road, Presence of right turn lanes	Presence of lighting	4-leg intersection density (per square mile)
Posted speed limit (mph)	Major road. AADT 2014		Total population density (per square mile)
Total population density (per square mile)	Minor road. Presence of median		Minor road, Presence of bicycle lanes
Number of total traffic lane	Minor road. Presence of right turn lanes		Minor road, Presence of right turn lanes
			Minor road, Number of total traffic lanes
			Major road, AADT 2014
			Household density (per acre)
			Major number. Presence of right turn lanes
			Minor road. Traffic direction

Table 6.10: Summary of Significant Variables in the Bicycle Models

Crash Occurrence		Crash Severity	
Model 5 - Segment	**Model 6 - Intersection**	**Model 7 - Segment**	**Model 8 - Intersection**
Presence of crossing	Bicycles per day (STRAVA)	Cross combination type B (type A is base)	Total population density (per square mile)
AADT (2014)	Number of transit lines go through the intersection	Cross combination type C (type A is base)	Household density (per acre)
3-Leg intersection density (per square mile)	Minor road. Functional class	Cross combination type N (type A is base)	Household size
Bikes per day (STRAVA)	Minor road. Number of total traffic lanes	Width of sidewalk buffer	Minor road, Presence of right turn lanes
			Major road. Presence of left turn lanes
			Minor road. Presence of left turn lanes

7.0 RISK-SCORING TOOL

This chapter describes the risk-scoring tool for pedestrians and bicycles at segments and intersections (four scoring tools in total) that was created using the crash occurrence models. This chapter first presents the method used to create the actual risk scores, and then each of the risk-score tables is presented. The risk-scoring method was applied to all segments and intersections in the modeling dataset. The method is applied to the entire modeling set and a distribution of the risk scores is presented.

7.1 METHOD TO DEVELOP RISK SCORE

The results of the logit occurrence models (crash / no crash) were used to develop the risk scores. The purpose of developing a risk score is to assign a value that can be used to prioritize locations with increased or elevated risk. Risk is defined as a probability or threat of damage, injury, liability, loss, or any other negative occurrence that is caused by external or internal vulnerabilities, and that may be avoided through preemptive actions. Thus, the risk scores include both elements of exposure and likely outcomes. For each score matrix, the maximum risk score is set to 100.

To convert the model results to a risk score, the research team used the odds ratio. A higher odds ratio indicates that there is a higher probability of having crashes in the segment. Before calculating risk scores it is necessary to transform the "base condition" to the condition with minimum odds ratio in order to make all risk scores positive. We then assign each variable a component of the risk score based on the overall contribution. For ease of use in the calculation tool, levels or categories are created for continuous variables. This method is demonstrated using the pedestrian segment model.

Step 1: Calculate internal weights for categorical variables. For categorical variables, the value with a coefficient 0 is considered as the base condition in the model. We need to transform the base condition to the level with the lowest odds ratio in order to make all ratios larger than 1. For example, in Table 7-1 "Traffic Direction," the lowest odds ratio is from level "Two-Way." When the level "Two-Way" is selected as the new base condition, a transformed odds ratio will be calculated. The new odds ratio of "one-way traffic" is 1/0.2755=3.6292, which indicates that the probability of having pedestrian crashes on a one-way segment would be 3.6 times more than the two-way segments when the other variable holds.

Step 2: Select levels of continuous variables, calculate the internal weight. Continuous variables are assigned levels for convenience. The level breaks are based on the underlying data distribution of the modeling dataset. First, the median value of the variable is used to calculate the odds ratio for each level. For example, the total population density variable was divided into five categories: 1-1000, 1001-3000, 3001-5000, 5001-7000 and >7000. Similar to categorical variables, the odds ratio of increasing one unit of total population density when other variables holds is $e^{0.0.00024} = 1.00024$, which indicates that the logit difference would be 1.00024 compared to the density is 0. In level 1-1000, a median 500 is used to calculate the odds ratio, meaning that the logit difference is $e^{500*0.00024} = 1.13$ compared to the density is 0. Similarly, the odds ratio of level 1001-3000 is $e^{0.00024*2000} = 1.62$ using density = 2000 and the odds ratio of level 3001-5000 is $e^{0.000248*4000} = 2.61$ using 4000. The odds ratio of level 5001-7000 is

$e^{0.00024*6000} = 4.22$ and density =8000 is used to calculate the odds ratio of level >7000, which equals to $e^{0.00024*8000} = 6.82$ We also needed to convert some continuous variables to change the "base condition," such as number of intersections with negative coefficients. Also note that for posted speed limit, the original base condition is 0, which is not realistic. For example, the base condition was changed to 25 mph (all observations in the data have posted speed limits larger or equal to 25).

Step 3: Assign a component risk score to each variable out of a 100 scale. Each variable is assigned a component risk score out of 100. This is done by the proportion of each variable's highest internal weight to the total maximum internal weight. Table 7-1 shows the maximum internal weight for each variable in this model. Hence, the component risk score of "more than four traffic lanes":

$\frac{4.38}{(3.63+3.81+2.7+2.92+4.40+4.38)} * 100 = 20.04$ which we round to 20.

Step 4: Distribute each component risk score to levels by internal weights. After the calculation in Steps 1-3 is complete, the risk scores for each level are distributed based on the internal weights for each level. For example, we can calculate the risk score of "three or four traffic lanes" by the ratio of the level internal weight to the maximum internal weight:

$\frac{2.09}{4.38} * 20 = 9.5$ which we round to 10. The final assigned risk scores by level are shown in the last column of Table 7-1.

Table 7.1: Sample Calculation of Risk Score by Level

Variable	How Modeled?	Model Coefficient	Odds Ratio	Levels	Internal Weight	Max Internal Weight	Component Risk Score	Risk Score By Level
Traffic direction	Cat.	-1.289	0.28	One-way	3.63	3.63	17	**17**
				Two-way	1.00			0
Presence of on-street parking	Cat.	1.337	3.81	Yes	3.81	3.81	17	**17**
				No	1			0
Posted speed limit (mph)	Cont.	0.0496	1.05	<=25	1.00	2.70	12	0
				30	1.28			6
				35	1.64			8
				>35	2.70			**12**
Presence of TWLTL	Cat.	1.071	2.92	Yes	2.92	2.92	14	**14**
				No	1.00			0
Total population density (per square mile)	Cont.	0.000174	1.0002	<=1000	1.00	4.41	20	0
				1001-3000	1.30			6
				3001-5000	1.84			8
				5001-7000	2.62			12
				>7000	4.41			**20**
Total traffic lane	Cat.	0.3695	1.45	2	1	4.38	20	0
				3 or 4	2.09			10
				>4	4.38			**20**

7.2 PEDESTRIAN RISK SCORE

The conversion of the models to the risk-scoring tools followed the same procedure for all four applications. The scoring-tool results are presented for segments and intersections for pedestrians.

7.2.1 Segment Risk Score

Risk scores of significant variables in intersection pedestrian crash models are presented in Table 7.2. The maximum value for each category is bolded. Three continuous variables – total population density, total traffic lanes and posted speed limit – are divided into distinct levels, as shown. The largest contributing variable to the risk scores (maximum value =20) is the total population density (larger than 7,000) and total traffic lanes (more than four). Traffic direction has a large risk-score component; a score of 17 for one-way roads and presence of on-street parking also has a score of 17. The presence of TWLTL is assigned 17 risk points.

Table 7.2: Pedestrian Segment Risk Scores

Variables	Levels	Internal Weight	Risk Score
Traffic direction	One-way	3.63	**17**
	Two-way	1.00	0
On-street parking	Yes	3.81	**17**
	No	1.00	0
Posted speed limit (mph)	<=25	1.00	0
	30	1.28	6
	35	1.64	8
	>35	2.70	**12**
Presence of TWLTL	Yes	2.92	**14**
	No	1.00	0
Total population density (per square mile)	<=1000	1.00	**0**
	1001-3000	1.30	6
	3001-5000	1.84	8
	5001-7000	2.62	11
	>7000	4.41	**20**
Total traffic lanes	2	1	0
	3 or 4	2.09	10
	>4	4.38	**20**

7.2.2 Intersection Risk Score

Risk scores of significant variables in the intersection pedestrian crash model are presented in Table 7.3. Three continuous variables – total population density, the number of transit lines and AADT of major roads – are divided into several levels as shown. The largest contributing variable to the risk score (maximum value = 26 is assigned to the number of transit routes running through the intersection. Population density in the census block received the next highest weight with a maximum score of 20. Traffic volume, with respect to major road AADT, was next with a maximum score of 18 for volumes exceeding 25,000. The remaining geometric variables related to medians on the major roads, right-turn lanes on the minor roads, and right-turn lanes on the minor roads received 13, 15, and 8 points, respectively.

Table 7.3: Pedestrian Intersection Risk Scores

Variable	Levels	Internal Weight	Risk Score
Total population density (per square mile)	<=1000	1.00	0
	(1001, 3000)	1.44	5
	(3001, 5000)	2.30	8
	(5001, 7000)	3.77	13
	>7000	6.03	**21**
Number of transit lines with routes through intersection	0(base)	1.00	0
	1	1.47	6
	2	2.15	8
	3	3.16	12
	>3	6.79	**25**
Major AADT (2014)	<=5000	1.00	0
	(5001, 10000)	1.37	5
	(10001, 15000)	1.88	7
	(15001, 20000)	2.57	10
	(20001, 25000)	3.52	13
	>25000	4.82	**18**
Presence of median on major road	Yes	1.00	0
	No	3.52	**13**
Minor road, presence of right-turn lanes	Yes	1.00	0
	No	3.71	**15**
Major road, presence of right-turn lanes	No	1.00	0
	Yes	2.19	**8**

7.3 BICYCLE RISK SCORE

Risk-score tools were developed for both segments and intersections for bicycles.

7.3.1 Segment Risk Score

Table 7.4 shows the calculated risk scores of significant variables of the bicycle segment crash occurrence model. Similar to the pedestrian segment occurrence model, the base condition is converted to the presence of a crossing in order to make all risk scores positive.

Table 7.4: Bicycle Segment Risk Scores

Variables	Levels	Internal Weight	Risk Score
Bikes per day (STRAVA)	<=200	1.00	0
	201-800	1.50	15
	>800	2.48	**25**
AADT	<=5000	1.00	0
	5001-10000	1,17	12
	10001-15000	1.38	14
	15001-20000	1.61	16
	20001-25000	1.89	19
	>25000	2.40	**25**
Three-leg intersection density per square mile (EPA Smart Location)	1-150	1.00	0
	151-200	1.23	13
	>200	1.60	**16**
Presence of marked crosswalk	Yes	1.00	**0**
	No	3.34	**34**

7.3.2 Intersection Risk Score

Table 7.5 shows the risk score developed from the bicycle intersection model. In order to make all risk scores positive, two variables' base condition is converted. Variable "Minor road functional class" transferred its base condition from Arterial to Collector and the base condition of variable "Minor road presence of right-turn lane" is "Presence of right-turn lane" instead of No right-turn lane.

Table 7.5: Bicycle Intersection Risk Scores

Variable	Level	Internal Weight	Risk Score
Bikes per day (STRAVA)	<=200 (base)	1	0
	<= 800	2.4	11
	>800	4.3	**20**
Number of transit stops	0(base)	1	0
	1	1.4	7
	2	2	10
	3	2.8	14
	>3	5.7	**27**
Minor functional class	Collector	1	0
	Arterial	2.3	**12**
Minor road total number of traffic lanes	2(base)	1	0
	3	1.6	8
	4	2.7	12
	>4	7.2	**31**
Minor road presence of right-turn lane	Yes (base)	1	0
	No	2.2	**10**

7.4 RISK-SCORE DISTRIBUTION

The risk-scoring tables were applied to all 188 segments and 184 intersections in the modeling dataset. Figure 7.1 shows the distribution of scores for each of the risk tools as applied to these locations. With the exception of the intersection tool for pedestrians, the distributions are skewed. The bicycle segment model is less distributed, reflecting the limited number of significant variables that were in the model. Table 7.6 shows the risk score by percentiles and the mean score. The risk scores are only intended to be evaluated within each context (i.e., the risk score for bicycles at intersections is not comparable to the score for pedestrians on segments). The distributions of the risk scores could be used when making comparisons across the tools by estimating the percentile of the score from Table 7.6. For example, a risk score of 46 would be above the 75th percentile of the calculated scores for the pedestrian, intersection and bicycle segment, but only average for the other two tools.

Table 7.6: Risk-Score Distribution

Risk Tool	25th Percentile	50th Percentile	Mean	75th Percentile
Segment, Pedestrian	16.0	24.0	24.97	34.0
Segment, Bicycle	26.5	46.0	41.02	61.0
Intersection, Pedestrian	39.0	48.0	47.12	55.0
Intersection, Bicycle	20.0	31.0	33.72	43.0

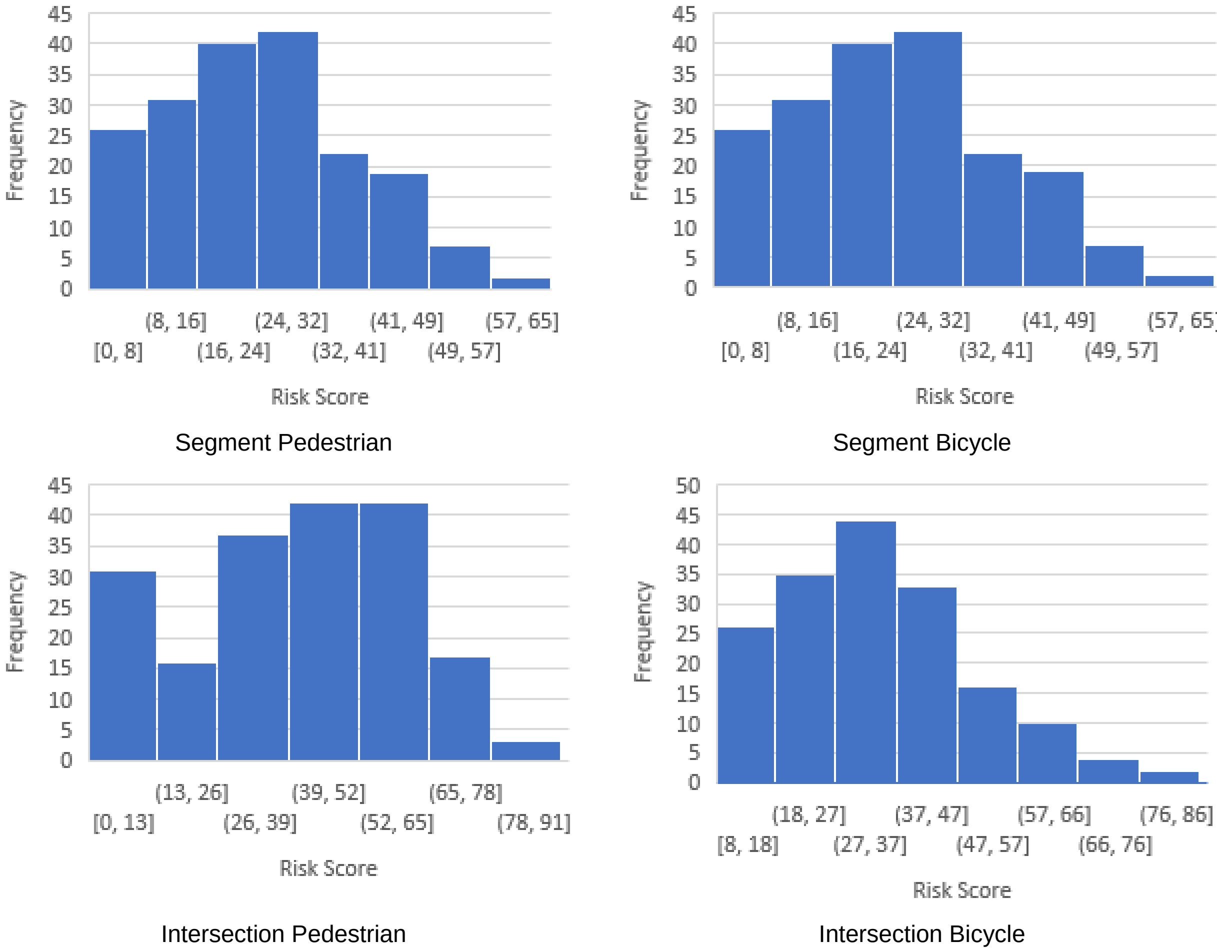


Figure 7.1: Intersection Crash Severity Level Distribution

7.5 SUMMARY

To aid in the implementation of these risk-scoring tables the research team constructed an Excel spreadsheet. The use of the scoring tool is demonstrated in the following chapter using projects identified for funding in the All Roads Transportation Safety (ARTS) final project lists in Region 1 and Region 2. This chapter presented the method used to convert the model results to a weighted risk score for each of the components. Presently, the risk scores are only intended to be evaluated within each context (i.e., the risk score for bicycles at intersections is not comparable to the score for pedestrians on segments). Use of the percentiles can be used to inform comparisons between the scoring tools.

8.0 SAMPLE APPLICATIONS

In this chapter, we apply the risk-scoring tool to safety projects that were recommended in the 2015 All Roads Transportation Safety (ARTS) project lists from Region 1 and 2. The chapter includes screen captures from the risk-scoring spreadsheet tool that was developed for the project and is available for use on the ODOT research website. Appendix A contains examples of the individual data collection elements.

8.1 INTERSECTION PROJECTS

A total of five intersections projects with pedestrian-related countermeasures were selected from ARTS final project lists in Region 1 and Region 2. These are shown in Table 8.1 with the project name, proposed countermeasures and the benefit-cost (B/C) ratio calculated through the ARTS procedures.

Table 8.1: Validation List of Pedestrian Intersection Risk Score

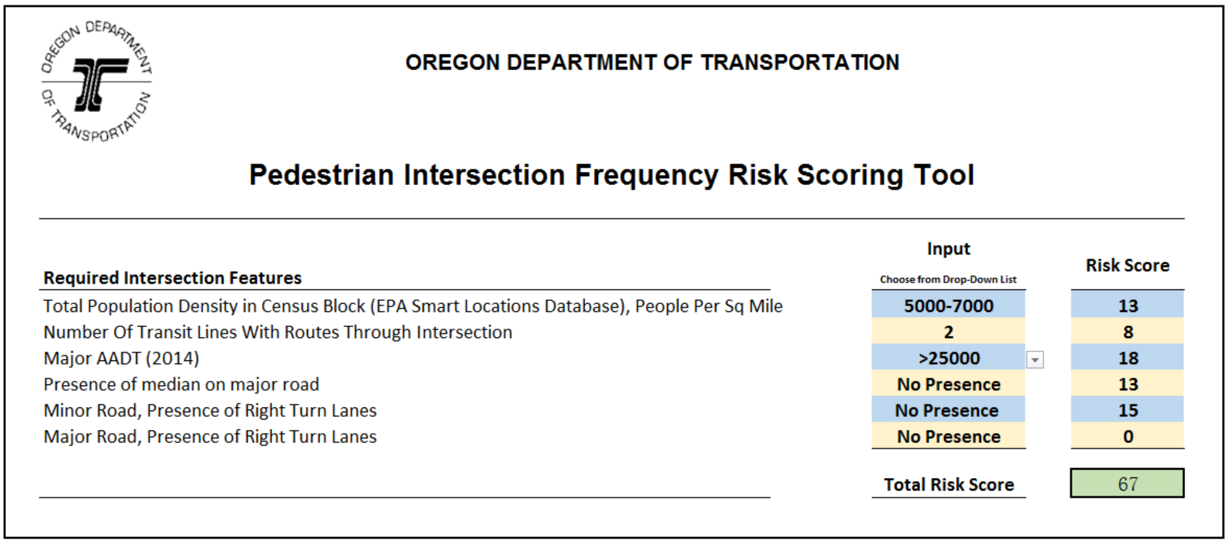
City	Project Name	Countermeasures	Benefit Cost Ratio
Keizer	RIVER RD NE @ SAM ORCUTT WAY NE	BP4 - Install No Pedestrian Phase Feature with Flashing Yellow Arrow	17.16
Eugene	I 105 @ MP 1.8: COBURG RD @ MLK JR BLVD	BP1 - Install Pedestrian Countdown Timer(s)	9.87
Albany	GEARY ST @ QUEEN AVE	BP4 - Install No Pedestrian Phase Feature with Flashing Yellow Arrow BP5 - Install Urban Green Bike Lanes at Conflict Points	7.13
Salem	BROADWAY ST NE @ PINE ST NE	BP1 - Install Pedestrian Countdown Timer(s)	2.45
Beaverton	SW Hall Blvd @ SW Nimbus Ave	BP4 - Install No Pedestrian Phase Feature with Flashing Yellow Arrow	17

The first step to calculate the risk score is to collect the necessary information for the risk-scoring tool. Table 7.3 shows that six variables need to be collected. Total population density can be obtained from the Environmental Protection Agency (EPA)’s Smart Location Database (2013); AADT and number of transit can be collected from ODOT TransGIS website; and Google Maps street view images provided the geometric information such as the presence of a median on a segment and the presence of a right-turn lane on major and minor roads. For the first intersection in Keizer, the detailed information of this intersection is shown in Table 8.2.

Table 8.2: Detailed Information of Intersection River Road @ Sam Orcutt Way

Project name	Total population density (per square mile)	Number of transit lines on go through intersection	AADT on major road	Major road: Presence of median	Minor road: Presence of right-turn lane	Major road: Presence of right-turn lane
RIVER RD NE @ SAM ORCUTT WAY NE	5857	2	25500	No	No	No

The next step is to use the tool/method to calculate the risk score. A user could easily self-score the location using Table 7.3 or by using the spreadsheet tool and selecting the most appropriate one from drop-down lists for each variable. Figure 8.1 shows the risk score of individual input and the total risk score, 67 points.



OREGON DEPARTMENT OF TRANSPORTATION

Pedestrian Intersection Frequency Risk Scoring Tool

Required Intersection Features	Input (Choose from Drop-Down List)	Risk Score
Total Population Density in Census Block (EPA Smart Locations Database), People Per Sq Mile	5000-7000	13
Number Of Transit Lines With Routes Through Intersection	2	8
Major AADT (2014)	>25000	18
Presence of median on major road	No Presence	13
Minor Road, Presence of Right Turn Lanes	No Presence	15
Major Road, Presence of Right Turn Lanes	No Presence	0
	Total Risk Score	67

Figure 8.1: Calculated Risk Score of River Road @ Sam Orcutt Way with Excel Risk-Scoring Tool

Following the same procedure, the risk scores of the remaining four intersections are calculated and listed in Table 8.3. The project order implied by the risk scores corresponds well to the project's final B/C value. Referring to Table 7-6, the risk scores for the projects with the two highest B/C ratios are above the 75th percentile score (55). The lower-ranked B/C projects correspond to lower percentiles of the risk score. As ARTS projects go through a substantial evaluation process, higher risk scores might be expected.

Table 8.3: Risk Scores of Five Pedestrian Intersection Projects

City	Project Name	B/C	Risk Score	Risk Percentile
Keizer	RIVER RD NE @ SAM ORCUTT WAY NE	17.16	67	> 75th
Beaverton	SW Hall Blvd @ SW Nimbus Ave	17	63	> 75th
Eugene	I 105 @ MP 1.8: COBURG RD @ MLK JR BLVD	9.87	58	> 75th
Albany	GEARY ST @ QUEEN AVE	7.13	46	= 50th
Salem	BROADWAY ST NE @ PINE ST NE	2.45	53	= 75th

A similar exercise was conducted for another five intersection projects that relate to bicycles. Following the same steps described above in data gathering and evaluation, the risk scores are listed in Table 8.4. As shown, the higher risk scores align with the higher B/C ratios with the exception of the Albany project. The 75th percentile risk score is 42.75 from Table 7.6; all of the projects are below this value.

Table 8.4: Risk Scores of Five Intersection Projects for Bicycles

City	Project Name	Countermeasures	BC	Risk Score	Risk Percen tile
Salem	D ST NE @ LANCASTER DR NE	BP5 - Install Urban Green Bike Lanes at Conflict Points Protected/Permissive BP4 - Install No Pedestrian Phase Feature with Flashing Yellow Arrow	20.66	40	< 50th
Salem	FAIRVIEW AVE SE @ 12TH ST SE	BP5 - Install Urban Green Bike Lanes at Conflict Points	20.06	18	< 25th
Portlan d	Lombard St @ N Interstate Ave (US 30B)	BP1 - Install Pedestrian Countdown Timer(s)	16.80	28	< 50th
Albany	GEARY ST @ QUEEN AVE	BP4 - Install No Pedestrian Phase Feature with Flashing Yellow Arrow BP5 - Install Urban Green Bike Lanes at Conflict Points	7.13	34	> 50th
Eugene	RIVER RD @ IRVING RD	BP5 - Install Urban Green Bike Lanes at Conflict Points	2.45	18	< 25th

8.2 SEGMENT PROJECTS

Projects were selected from the same ARTS final project lists in Region 1 and Region 2 to demonstrate the use of the segment analysis tool. There are not as many segment-based pedestrian or bicycle projects to select from in the ARTS project list. In addition, the length of projects in the lists cover long distances. To use the risk tool, the segment projects are broken into smaller segments defined by breaks for major street traffic control. As an example, the project on Oatfield Road in Clackamas County was selected as a validation sample for pedestrian crashes. Two shorter segments, shown in Table 8.5, were created and risk scores were calculated using the tool.

Table 8.5: Segment Project for Pedestrian Segment Risk-Score Validation

Agency	Name	Countermeasures	B/C
Clackamas County	Oatfield Rd	BP1 - Install Pedestrian Countdown Timer(s)	13.04
Segment for Validation			
Clackamas County	Oatfield Rd from Roethe Rd to Jennings Ave	BP1 - Install Pedestrian Countdown Timer(s)	
Clackamas County	Oatfield Rd from Roethe Rd to SE Thiessen Rd	BP1 - Install Pedestrian Countdown Timer(s)	

Following the same procedure for intersection validation in the previous section, risk scores could be easily calculated by the tool after collecting all the necessary information. Figure 8.2 shows how the risk-scoring tool works for the segment Oatfield Road from Roethe Road to Jennings Avenue, and Table 8.6 shows more details on the risk scores of two segment samples. The calculated risk scores of 16 and 14 for the project are low (near the 25th percentile of scores) while the B/C for the project is 13.04 (high).

OREGON DEPARTMENT OF TRANSPORTATION

Pedestrian Segment Frequency Risk Scoring Tool

Required Segment Features	Input (Choose from Drop-Down List)	Risk Score
Presence of On-street Parking	No	0
Total Population Density in Census Block (EPA Smart Locations Database), People Per Mile	3001-5000	8
Traffic Direction	Two-Way	0
Posted Speed Limit (MPH)	35	8
Presence of Two-way Left Turn Lane	No Presence	0
Total Traffic Lane	2	0
	Total Risk Score	16

Figure 8.2: Calculated Risk Score for Oatfield Road from Roethe Road to SE Thiessen Road with Excel Risk Tool

Table 8.6: Risk Scores of Validation Pedestrian Segment Projects

Variable	Oatfield Rd from Roethe Rd to Jennings Ave		Oatfield Rd from Roethe Rd to SE Thiessen Rd	
	Data	Risk Score	Data	Risk Score
Presence of on-street parking	No	0	No	0
Total population density in census block (EPA Smart Locations Database), people per mile	3873.6	8	2075	6
Traffic direction	2	0	2	0
Posted speed limit (MPH)	35	8	35	8
Two-way left-turn lane presence	No	0	No	0
Total traffic lanes	2	0	2	0
Total		**16**		**14**
Risk Score Percentile		=25th		< 25th

For bicycles, a project in Salem on Commercial Street was selected for testing of risk score calculations for bicycle crashes. The whole segment was cut off to three smaller segments at signalized intersections and the risk scores are calculated by the tool. Table 8.7 shows the results of the risk scores. The calculated risk scores are well above the 75th percentile score of 61.

Table 8.7: Risk Scores for Validation Segment for Bicycles

Agency	Name	Countermeasures	B/C
Salem	Commercial St	**BP20** - Install Buffered Bike Lanes **BP2** - Provide Intersection Illumination (Bike & Ped) **BP10** - Install Rectangular Rapid Flashing Beacon with Median (3-Lane or More Roadway)	0.32

Variable	Commercial St From Alice Ave to Boice St	Commercial St Vista Ave to Alice Ave	Commercial St Boice St to Hoyt St	Risk Score	Risk Score	Risk Score
Average daily bicycle volume per day	653	529	789	15	15	15
AADT	32400	23300	24200	25	19	19
Three-leg intersection density per square mile (EPA Smart Location)	254.25	174.41	254.25	16	13	16
Presence of crosswalk	No	No	No	34	34	34
Total				**90**	**81**	**81**
Risk Score Percentile				> 75th	> 75th	> 75th

8.3 SUMMARY

This chapter applied the risk-scoring tools to safety projects that were recommended in the 2015 ARTS project lists from Region 1 and 2. The data requirements for the risk-scoring tool are not intensive and the scores aligned reasonably well with the benefit-cost calculations for intersection project. The chapter demonstrated how the risk scores can be interpreted using the percentile from the risk-score distributions.

9.0 CONCLUSIONS

The objective of this research was to develop a tool for ODOT to identify and prioritize locations with increased or elevated risk for pedestrian and bicycle crashes. Risk scoring should include elements of exposure and expectations of the severity of the outcome but not be dependent on crash history. To accomplish this objective, the project team assembled a database of segments and intersections for analysis. The segments were randomly selected from arterials within urban areas in the state of Oregon. Both state and non-state facilities were included in the random sample. Following a detailed literature review, important variables were identified for data collection. In keeping with the project-based application, the effort focused on building tools for intersections and segments separately. The database assembled for analysis included detailed geometric and operational elements as well as broad descriptors of the built environment.

The research team gathered motor vehicle volumes and estimates of bicycle exposure from a third-party data of bicycle use from STRAVA under ODOT's license. The sampled segments and intersections were then linked to crash data. A total of 188 segments and 184 intersections were included in the modeling database. There were 213 segment crashes and 238 intersection crashes included in the model. The research team developed logistic regression models for both crash occurrence (crash or not) and crash severity models. The models related to crash severity were not robust, most likely due to the few segments and intersections with severe crashes in the dataset. The crash occurrence models produced more plausible results, with significant variables that included a blend of exposure and geometric/operational variables hypothesized to relate to risk.

The primary outcome of this research was the development of a risk-score tool for pedestrians and bicycles. Using the results of the modeling effort, a method was developed to create a risk-scoring tool for pedestrians and bicycles at intersections and segments (a total of four scoring tools). These risk tables were incorporated into a spreadsheet for easy application. The risk-scoring tool was applied the tool to safety projects that were recommended in the 2015 All Roads Transportation Safety (ARTS) project lists from Region 1 and 2. The risk scores for the case study applications for intersections aligned reasonably well with the project's benefit-costs estimate developed by the ARTS process. Application to the segments scoring was less aligned.

The primary challenge to quantifying the risk for pedestrian and bicycles on road segments is the missing measures of exposure and the relatively few pedestrian and bicycle crashes observed on most segments and intersections. The inclusion of the bicycle STRAVA data significantly improved the bicycle models, though the data's ability to accurately represent all bicycle travel is still somewhat uncertain. Prior to the inclusion of this variable, there were few surrogate variables for bicycle exposure and the model fits were generally poor.

In this research, the value of each risk score was derived from the modeling output. All models suffer from the limitations of the input dataset. With a larger or different sample for modeling, there is the possibility that the risk scores would be different. Also, the variables identified as significant in the risk-score estimation procedure should not be interpreted as recommendations for engineering-level improvements. The variables are, in many cases, explaining more about the safety of the location than the individual variable. Design-level safety decisions should use more robust tools such as the *Highway Safety Manual*.

9.1 FUTURE WORK

There are a number of potential future areas of research to recommend:

- As with any non-motorized safety research or analysis project, the lack of good exposure data is limiting. For the pedestrians, much promise lies in using area-type models to develop exposure estimates. The addition of the STRAVA data proved very useful in the model's explanatory power. Models of bicycle crashes without some data on bicycle use were not useful. Additional work with the STRAVA data or others like it for estimating bicycle volumes is recommended.

- A larger dataset, perhaps derived from GIS or automated data mining tools, would produce a very robust database for a similar modeling effort. Rather than a statewide focus, a regional or MPO-level analysis would likely yield good results leveraging the more detailed spatial data available.

10.0 REFERENCES

AASHTO. *Highway Safety Manual*. American Association of State Highway and Transportation Officials (AASHTO), Washington, DC, 2010.

Abdel-Aty, M., and J. Keller. Exploring the Overall and Specific Crash Severity Levels at Signalized Intersections. In *Accident Analysis & Prevention*, Vol. 37 No. 3, 2005, pp. 417-425.

Al-Ghamdi, A. S. Using Logistic Regression to Estimate the Influence of Accident Factors on Accident Severity. In *Accident Analysis & Prevention*, Vol. 34 No. 6, 2002, pp. 729-741

Andersson A. L., and O. Bunketorp. Cycling and Alcohol. In *Injury*, Vol. 33 No. 6, 2002, pp. 467-471.

Berdica, K. An Introduction to Road Vulnerability: What Has Been Done, Is Done and Should be Done. In *Transport Policy*, Vol. 9 No. 2, 2002, pp. 117-127.

Bergh, C., E. Ferguson, and B. Ray. *Oregon Pedestrian or Bicycle Safety Implementation Plan: Project Direction and Scope*. Kittelson & Associates, Inc , Portland, 2013.

Bergh, C. and A. Griffin. (2013). ODOT Pedestrian or Bicycle Safety Implementation Plan: Draft Risk-Based Segment Prioritization. Kittelson & Associates, Inc., Bend, 2013.

Bíl, M., M. Bílová, & I. Müller. Critical Factors in Fatal Collisions of Adult Cyclists with Automobiles. In *Accident Analysis & Prevention,* Vol. 42 No. 6, 2010, pp. 1632-1636.

Boufous, S., L. de Rome, T. Senserrick, and R. Ivers. Risk Factors for Severe Injury in Cyclists Involved in Traffic Crashes in Victoria, Australia. In *Accident Analysis & Prevention*, Vol. 49, 2012, pp. 404-409.

Carter, D. L., W.W. Hunter, C.V. Zegeer, J.R. Stewart, and H.F. Huang. *Pedestrian and Bicyclist Intersection Safety Indices: Final Report*. Publication FHWA-HRT-06-125. FHWA, U.S. Department of Transportation, 2006.

Chimba, D., D. Emaasit, C.R. Cherry, and Z. Pannell. Patterning Demographic and Socioeconomic Characteristics Affecting Pedestrian and Bicycle Crash Frequency. No. 14-0600. Presented at Transportation Research Board 93rd Annual Meeting, Washington D.C., 2014.

de Geus, B., G. Vandenbulcke, L. I. Panis, I. Thomas, B. Degraeuwe, E. Cumps, and R. Meeusen. A Prospective Cohort Study on Minor Accidents Involving Commuter Cyclists in Belgium. In *Accident Analysis & Prevention*, Vol. 45, 2012, pp. 683-693.

Diogenes, M., and L. Lindau. (2010). Evaluation of Pedestrian Safety at Midblock Crossings, Porto Alegre, Brazil. In *Transportation Research Record: Journal of the Transportation Research Board No. 2193*, TRB, National Research Council, Washington D.C., 2010, pp. 37-43.

Dixon, K., R. Avelar, L. Brown, M. Mecham, and I. Van Schalkwyk. *Quantifying Safety Performance of Driveways on State Highways* Publication FHWA-OR-RD-13-02, FHWA, U.S Department of Transportation, 2012.

Dixon, K., C. Monsere, F. Xie, and K. Gladhill. *Calibrating the Future Highway Safety Manual Predictive Methods for Oregon State Highways* Publication No. OTREC-RR-12-02. 2012.

Dixon, K., C. Monsere, R. Avelar, J.S. Barnett, P. Escobar, and S.M. Kothuri. *Improved Safety Performance Functions for Signalized Intersections.*2015.

Dolatsara, H.A.. *Development of Safety Performance Functions for Non-Motorized Traffic Safety*. 2014.

Ekman, R., G. Welander, L. Svanström, L. Schelp, and P. Santesson. Bicycle-Related Injuries Among the Elderly—a New Epidemic?. In Public Health, Vol. 115 No. 1, 2001, pp. 38-43.

Eluru, N., C.R. Bhat, and D.A. Hensher. A Mixed Generalized Ordered Response Model for Examining Pedestrian and Bicyclist Injury Severity Level in Traffic Crashes. In *Accident Analysis & Prevention,* Vol. 40 No. 3, 2008, pp. 1033-1054.

Ewing, R., & Cervero, R. (2001). Travel and the built environment: a synthesis. Transportation Research Record: Journal of the Transportation Research Board, (1780), 87-114.

Travel, E. N. Federal Highway Administration University Course on Bicycle and Pedestrian Transportation. 2006.

Fitzpatrick, K., V. Iragavarapu, M.A. Brewer, D. Lord, J. Hudson, R. Avelar, and J. Robertson (2014). *Characteristics of Texas Pedestrian Crashes and Evaluation of Driver Yielding at Pedestrian Treatments*. TxDOT Report FHWA/TX-13/0-6702-1. Texas Department of Transportation. 2014.

Gårder, P.E. The Impact of Speed and Other Variables on Pedestrian Safety in Maine. In *Accident Analysis & Prevention*, Vol. 36 No. 4, 2004, pp. 533-542.

Greibe, P. Accident Prediction Models for Urban Roads. In *Accident Analysis & Prevention*, Vol. 35 No. 2, 2003, pp. 273-285.

Haleem, K., and M. Abdel-Aty. Examining Traffic Crash Injury Severity at Unsignalized Intersections. In *Journal of Safety Research*, Vol. 41 No. 4, 2010, pp. 347-357.

Harwood, D.W., D.K. Gilmore, D.J. Torbic, R. Souleyrette, and Z.N. Hans. *United States Road Assessment Program: Pilot Program: Phase II*. 2008.

Jang, K., S. Park, S. Kang, K. Song, S. Kang, and S. Chung. Evaluation of Pedestrian Safety: Pedestrian Crash Hot Spots and Risk Factors for Injury Severity. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2393*, TRB, National Research Council, Washington D.C., 2013, pp. 104-116.

Kim, J.K., S. Kim, G.F. Ulfarsson, and L.A. Porrello. Bicyclist Injury Severities in Bicycle–Motor Vehicle Accidents. In *Accident Analysis & Prevention*, Vol. 39 No. 2, 2007, pp. 238-251.

Kröyer, H. R. (2015). The Relation Between Speed Environment, Age and Injury Outcome for Bicyclists Struck by a Motorized Vehicle – a Comparison with Pedestrians. In *Accident Analysis & Prevention*, Vol. 76, 2015, pp. 57-63.

Lagerwey, P.A., M.J. Hintze, J.B. Elliott, J.L. Toole, and R.J. Schneider. Pedestrian and Bicycle Transportation along Existing Roads. In *ActiveTrans Priority Tool Guidebook*. Publication No. Project 07-17. 2015.

Lee, C., and M. Abdel-Aty. Comprehensive Analysis of Vehicle–Pedestrian Crashes at Intersections in Florida. In *Accident Analysis & Prevention*, Vol. 37 No. 4, 2005, pp. 775-786.

Lenguerrand, E., J.L. Martin, and B. Laumon. Modelling the Hierarchical Structure of Road Crash Data—Application to Severity Analysis. In *Accident Analysis & Prevention*, Vol. 38 No. 1, 2006, pp. 43-53.

Loukaitou-Sideris, A., R. Liggett, and H.G. Sung. Death on the Crosswalk: A Study of Pedestrian-Automobile Collisions in Los Angeles. In *Journal of Planning Education and Research*, Vol. 26 No. 3, 2007, pp. 338-351.

Lowry, M.B., and S. Cool. (2015). Dangerous Situation Topology and Analysis for Comparing Bicycle Infrastructure Improvement Plans. Paper. Presented at the 94th Annual Meeting of the Transportation Research Board. Paper 15-5989.

Lu, J. *Development of Safety Performance Functions for Safety Analyst Applications in Florida*. 2013.

Maring, W., and I. Van Schagen, I. Age Dependence of Attitudes and Knowledge in Cyclists. In *Accident Analysis & Prevention*, Vol. 22 No. 2, 1990, pp. 127-136.

Martin, A. *Factors Influencing Pedestrian Safety: a Literature Review*. Publication No. PPR241. Wokingham, Berks. TRL. 2006.

Martínez-Ruiz, V., P. Lardelli-Claret, E. Jiménez-Mejías, C. Amezcua-Prieto, J.J. Jimenez-Moleon, and J.D.D.L. del Castillo. Risk Factors for Causing Road Crashes Involving Cyclists: An Application of a Quasi-Induced Exposure Method. In *Accident Analysis & Prevention*, Vol. 51, 2013, pp. 228-237.

McDaniel, S., M. Lowry, and M. Dixon. Using Origin-Destination Centrality to Estimate Directional Bicycle Volumes. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2430*, TRB, National Research Council, Washington D.C., 2014, pp. 12-19.

McMahon P., C. Duncan, J. Stewart, C. Zegeer, and A. Khattak. Analysis of Factors Contributing to" Walking Along Roadway" Crashes. In *Transportation Research Record: Journal of the Transportation Research* Board, *No. 1674*, TRB, National Research Council, Washington D.C., 1999, pp. 41-48.

McMahon, P.J. *An Analysis of Factors Contributing to" Walking Along Roadway" Crashes: Research Study and Guidelines for Sidewalks and Walkways, Vol. 1*. DIANE Publishing. 2002.

Furth, P.G., M.C. Mekuria, and H. Nixon. Network Connectivity for Low-Stress Bicycling. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2587,* TRB, National Research Council, Washington D.C., 2016, pp. 41-49.

Miranda-Moreno, L.F., P. Morency, and A.M. El-Geneidy. The Link Between Built Environment, Pedestrian Activity and Pedestrian–Vehicle Collision Occurrence at Signalized Intersections. In. *Accident Analysis & Prevention*, Vol. 43 No. 5, 2011, pp. 1624-1634.

Moahn, D., Geetam T., Meleckidzedeck K., and F. Muyla Nafukho. (2006). "Risk Factors for Road Traffic Injuries." In Road Traffic Injury Prevention: Training Manual, 21–40. Geneva: World Health Organization.

Mountain, L., B. Fawaz, and D. Jarrett. Accident Prediction Models for Roads with Minor Junctions. In *Accident Analysis & Prevention*, Vol. 28 No. 6, 1996, pp. 695-707.

Noland, R., and M. Quddus, M. Analysis of Pedestrian and Bicycle Casualties with Regional Panel Data. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1897,* TRB, National Research Council, Washington D.C., 2004, pp. 28-33.

Nordback, K., W.E. Marshall, & B.N. Janson. Bicyclist Safety Performance Functions for a U.S. City. In *Accident Analysis & Prevention,* Vol. 65, 2014, pp. 114-122.

Oh, J., J. Jun, E. Kim, and M. Kim. *Assessing Critical Factors Associated with Bicycle Collisions at Urban Signalized Intersections*. Transportation Research Board, TRB, National Research Council, Washington D.C., 2008, pp. 1-17.

Olkkonen, S., and R. Honkanen. The Role of Alcohol in Nonfatal Bicycle Injuries. In *Accident Analysis & Prevention*, Vol. 22 No. 1, 1990, pp. 89-96.

Oregon Department of Transportation (ODOT). *ODOT TransGIS*. 2016. Available on: http://gis.odot.state.or.us/TransGis Access on 2016-10-12

Pai, C.W. Overtaking, Rear-End, and Door Crashes Involving Bicycles: An Empirical Investigation. In *Accident Analysis & Prevention*, Vol. 43 No. 3, 2011, pp. 1228-1235.

Palamara, P., and M. Broughton. An Investigation of Pedestrian Crashes at Traffic Intersections in the Perth Central Business. In *J Public Health*, Vol. 93, 2013, pp. 1456-1463.

Parkin, J., M. Wardman, and M. Page. Models of Perceived Cycling Risk and Route Acceptability. In *Accident Analysis & Prevention*, Vol. 39 No. 2, 2007, pp. 364-371.

Petritsch, T., B. Landis, H. Huang, and S. Challa. (2006). Sidepath Safety Model: Bicycle Sidepath Design Factors Affecting Crash Rates. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1982,* TRB, Washington D.C., 2006, pp. 194-201.

Poch, M., and F. Mannering. Negative Binomial Analysis of Intersection-Accident Frequencies. In *Journal of Transportation Engineering*, Vol. 122 No. 2, 1996, pp. 105-113.

Preston, H., R. Storm, J.D. Bennett, and B. Wemple. *Systemic Safety Project Selection Tool*. Publication No. FHWA-SA-13-019. 2013.

Pulugurtha, S.S., and A.P. Nujjetty. *Crash Estimation Models for Intersections. In Transportation and Development Institute Congress 2011: Integrated Transportation and Development for a Better Tomorrow*. 2011, pp. 1306-1314.

Pulugurtha, S.S., and V.R. Sambhara. Pedestrian Crash Estimation Models for Signalized Intersections. In *Accident Analysis & Prevention*, Vol. 43 No. 1, 2011, pp. 439-446.

Team, R.C. (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. 2013.

Ramsey, F., and D. Schafer. *The Statistical Sleuth: a Course in Methods of Data Analysis*. Cengage Learning. 2012.

Räsänen, M., and H. Summala. Attention and Expectation Problems in Bicycle–Car Collisions: an In-Depth Study. In *Accident Analysis & Prevention*, Vol. 30 No. 5, 1998, pp. 657-666.

Rodgers, G.B. Bicyclist Deaths and Fatality Risk Patterns. In *Accident Analysis & Prevention*, Vol. 27 No. 2, 1995, pp. 215-223.

Schepers, J.P., P.A. Kroeze, W. Sweers, and J.C. Wüst. Road Factors and Bicycle–Motor Vehicle Crashes at Unsignalized Priority Intersections. In *Accident Analysis & Prevention*, Vol. 43 No. 3, 2011, pp. 853-861.

Schepers, P., and B. den Brinker. What Do Cyclists Need to See to Avoid Single-Bicycle Crashes?. In *Ergonomics*, Vol. 54 No. 4, 2011, pp. 315-327.

Schepers, P., M. Hagenzieker, R. Methorst, B. Van Wee, and F. Wegman. A Conceptual Framework for Road Safety and Mobility Applied to Cycling Safety. In *Accident Analysis & Prevention*, Vol. 62, 2014, pp. 331-340.

Schneider, R., M. Diogenes, L. Arnold, V. Attaset, J. Griswold, and D. Ragland. Association Between Roadway Intersection Characteristics and Pedestrian Crash Risk in Alameda County, California. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2198*, TRB, Washington D.C., 2010, pp. 41-51.

Schneider, R.J., R.M. Ryznar, and A.J. Khattak. An Accident Waiting to Happen: a Spatial Approach to Proactive Pedestrian Planning. In *Accident Analysis & Prevention*, Vol. 36 No. 2, 2004, pp. 193-211.

Senserrick, T., S. Boufous, L. De Rome, R. Ivers, and M. Stevenson. Detailed Analysis of Pedestrian Casualty Collisions in Victoria, Australia. In *Traffic Injury Prevention*, Vol. 15 Sup 1, 2014, pp. S197-S205.

Stone, M., and J. Broughton, J. Getting Off Your Bike: Cycling Accidents in Great Britain in 1990–1999. In *Accident Analysis & Prevention*, Vol. 35 No. 4, 2003, pp. 549-556.

STRAVA, *Oregon Bicycle Count GIS Layer*. STRAVA INC. Provided by Oregon Department of Transportation. TPAU Union. 2014.

STRAVA. *About Us*. Available on: https://www.STRAVA.com/about. Retrieved on 2016-11-2.

STRAVA. *Global Heatmap*. Available on: http://labs.STRAVA.com/heatmap/#5/-103.99658/38.44498/blue/bike. Retrieved on 2016-11-2.

Summala, H., E. Pasanen, M. Räsänen, and J. Sievänen. Bicycle Accidents and Drivers' Visual Search at Left and Right Turns. In *Accident Analysis & Prevention*, Vol. 28 No. 2, 1996, pp. 147-153.

Tin, S.T., A. Woodward, and S. Ameratunga. Incidence, Risk, and Protective Factors of Bicycle Crashes: Findings From a Prospective Cohort Study in New Zealand. In *Preventive Medicine*, Vol. 57 No. 3, 2013, pp. 152-161.

Torbic, D., D. Harwood, D., C. Bokenkroger, R. Srinivasan, D. Carter, C. Zegeer, and C. Lyon. Pedestrian Safety Prediction Methodology for Urban Signalized Intersections. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2198*, TRB, Washington D.C., 2010, pp. 65-74.

Vandenbulcke, G., I. Thomas, and L.I. Panis. Predicting Cycling Accident Risk in Brussels: a Spatial Case–Control Approach. In *Accident Analysis & Prevention*, Vol. 62, 2014, pp. 341-357.

Walker, I. Drivers Overtaking Bicyclists: Objective Data on the Effects of Riding Position, Helmet Use, Vehicle Type and Apparent Gender. In *Accident Analysis & Prevention*, Vol. 39 No. 2, 2007, pp. 417-425.

Walker, S.H., and D.B. Duncan. Estimation of the Probability of an Event as a Function of Several Independent Variables. In *Biometrika*, Vol. 54 Nos. 1-2, 1967, pp. 167-179.

Wang, Y., and N.L. Nihan. Estimating the Risk of Collisions between Bicycles and Motor Vehicles at Signalized Intersections. In *Accident Analysis & Prevention*, Vol. 36 No. 3, 2004, pp. 313-321.

Wier, M., J. Weintraub, E.H. Humphreys, E. Seto, and R. Bhatia. An Area-Level Model of Vehicle-Pedestrian Injury Collisions with Implications for Land Use and Transportation Planning. In *Accident Analysis & Prevention*, Vol. 41 No. 1, 2009, pp. 137-145.

Zegeer, C., J. Stewart, H. Huang, and P. Lagerwey. Safety Effects of Marked Versus Unmarked Crosswalks at Uncontrolled Locations: Analysis of Pedestrian Crashes in 30 Cities. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1773*, TRB, Washington D.C., 2001, pp. 56-68.

Zegeer, C., D. Carter, W. Hunter, J. Stewart, H. Huang, A. Do, and L. Sandt. Index for Assessing Pedestrian Safety at Intersections. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1982*, TRB, Washington D.C., 2006, pp. 76-83.

APPENDIX A

EXAMPLE DATA COLLECTION FOR RISK TOOL

APPENDIX A

A.1 EXAMPLE DATA COLLECTION FOR RISK TOOL

A.1.1 Segment Example

Before calculating the risk scores for the spreadsheet tool, the first thing is to collect the data needed. These data are collected from google map, ODOT transgis website, google earth and EPA smart location database. Oatfield Rd form Roethe Rd to Jennings Ave, a segment sample application in the report, is used as an example in this appendix. Table A.1 lists all the variables that can be obtained from Google map images

Table A.1: Risk Score Variables of Oatfield Rd from Roethe Rd to Jennings Ave

Variable	Used for	Value
Presence of On-street Parking	Pedestrian Risk Score	No
Traffic Direction	Pedestrian Risk Score	Two-way
Post Speed Limit (MPH)	Pedestrian Risk Score	35
Presence of TWLTL Lane	Pedestrian Risk Score	No
Total Traffic Lane	Pedestrian Risk Score	2
Presence of Crosswalk	Bicycle Risk Score	No

Figure A.1: Google map image of Oatfield Rd

Social demographic variables could be extracted from Smart location database. The database is geocoded at block-size. Figure A.1 is the screen shot of the smart location database shapefile and the identity box of the corresponding area. The red line is the road segment used as a sample in this example. Table A.1 shows the value of these parameters obtained from Smart location database.

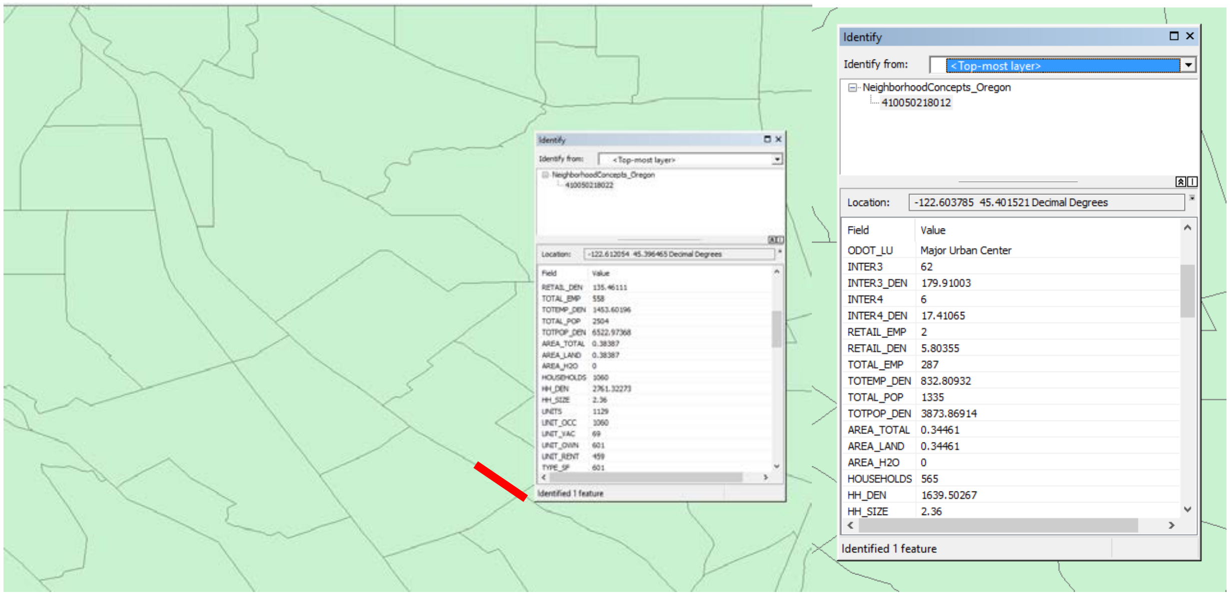


Figure A.2: Data from Smart Location Database

Table A.2: Social Demographic Variables Value

Variable	Used for…	Value
Total population density (People per square mile)	Pedestrian Risk Score	3873.6
Three-leg intersection density (per square mile)	Bicycle Risk Score	254.25

Strava bicycle volume database is also a shape file. Figure A.2 is the screen shot of identity box of sampled road segment (red line is the segment used). The variable “BIKECNT_YR” indicated the bike volume counted by strava. The volume of Oatfield Rd from Roethe Rd to Jennings Ave is 455. According to the formula in Chapter 3.2.4, bike adt could be converted as 455/3.65=124.66.

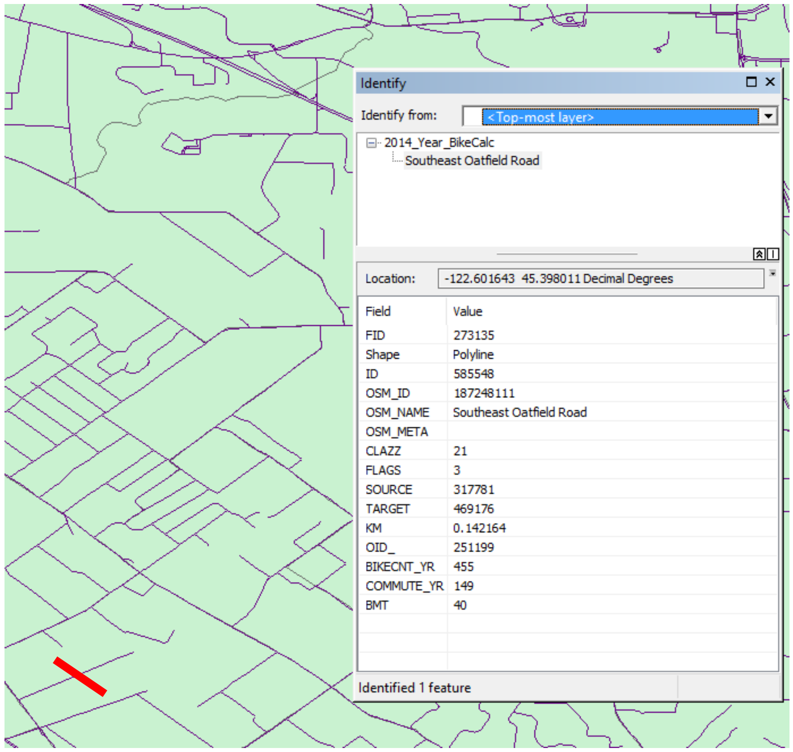


Figure A.3: Data in Strava Shapefile

AADT of the road segment could be extracted from variance resources. In most cases, ODOT Transgis website could be the first try. "Traffic count" should be displayed on the map and find the count station on or very close to the sample segment (shown as "dots" on the map).Figure A.3 showed the layer names needed to be displayed on the website and Figure A.4 showed the counted AADT of target segment, which is 8500 counted at 2015. If there are two count stations which are both close to the segment, average AADT value is calculated as the value of the segment. AADT value should be converted to a year of 2014 by AADT growth factor spreadsheet. If no AADT data is available on ODOT Transgis website, the transportation planning report from local agencies (county, city, etc.) will be the potential AADT resources.

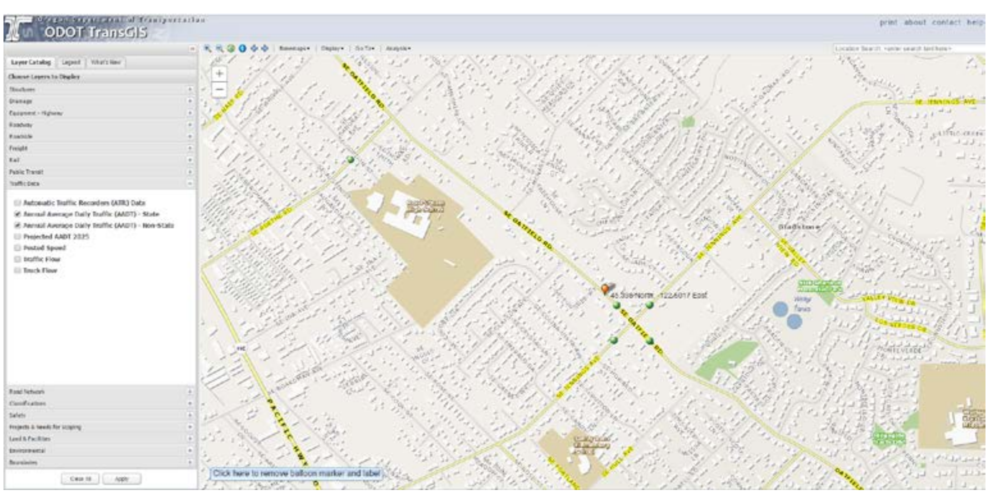


Figure A.4Traffic Count Layer on ODOT TransGIS Website

Identify Map Features	
Annual Average Daily Traffic (AADT) - Non-State - SE Oatfield Rd	
Road Description	SE Oatfield Rd
Location	SE Oatfield Rd., North of SE Jennings Ave
Milepoint	2.06
Begin Milepoint	2.01
End Milepoint	2.66
Site Number	29693
Annual Avg Daily Traffic (AADT)	8500
AADT Range	5,001 - 10,000
ODOT Road Number	825
01-Motorcycle	Null
02-Car	Null
03-Light Truck	Null
04-Bus	Null
05-Sgl Unit Truck-2 Axle	Null
06-Sgl Unit Truck-3 Axle	Null
07-Sgl Unit Truck-4 plus Axle	Null
08-Sgl Trailer Truck-4 or less Axle	Null
09-Sgl Trailer Truck-5 Axle	Null
10-Sgl Trailer Truck-6 plus Axle	Null
11-Multi Trailer-5 or less Axle	Null
12-Multi Trailer-6 Axle	Null
13-Multi Trailer-7 plus Axle	Null
Sgl Unit Truck AADT	Null
Multi Unit Truck AADT	Null
Peak Percent Sgl Unit Truck	Null
Peak Percent Multi Unit Truck	Null
Directional Factor	Null
Design Hour Factor (k Factor)	Null
Latitude	45.397686
Longitude	-122.601186
GIS Process Date	12/28/2016
Effective Date	2015

Figure A.5: AADT Information from Count Point

A.1.2 Intersection Example

Similar to segment example, one intersection in Chapter 8.1 of the report, River Rd Ne and Sam Orcutt Way Ne in City of Keizer, is used here as a demonstration of how and where to collect the data needed. First google map images are used to collect geometric-related information.

Figure A.5 and Figure A.6 show the google map image of major road and minor roads of the intersection. Table A.2 is the summary of the variable values, which can be obtained from these two images.

Figure A.5: Google Map Image of River Rd NE (Major Road)

Figure A.6: Google Image of Sam Orcutt Way NE (Minor Road)

Table A.3: Geometric Variables of Intersection Example

Variable	Used for…	Value
Presence of Median on Major Road	Pedestrian Risk Score	No
Minor Road, Presence Of Right Turn Lanes	Pedestrian Risk Score Bicycle Risk Score	No
Major Road, Presence Of Right Turn Lanes	Pedestrian Risk Score	No
Minor Road Total Number of Traffic Lanes	Bicycle Risk Score	3

Transit information is collected from Google Earth. After pinning the intersection on google earth map, the number of transit lines going through this intersection can be obtained from the information of the closest transit station. Figure A.7 is the Google Earth map of the intersection in the example. The transit information provided by ODOT Transgis website could be used to check when you are not sure what you get from Google Earth.

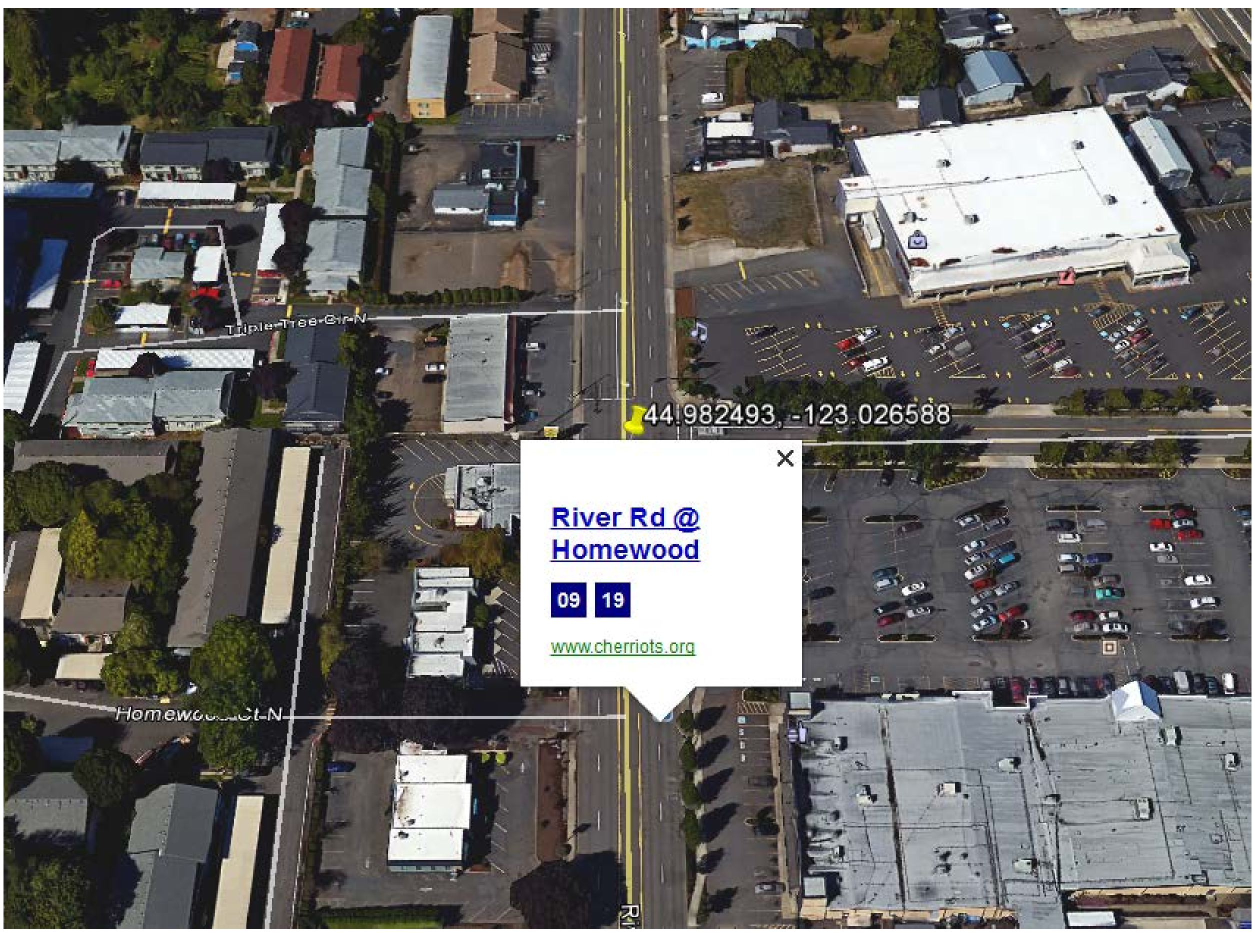


Figure A.7: Google Earth Map of Intersection

Similar to segment example, the value of total population density is collected from the EPA Smart Location Database. AADT of the major road is from ODOT Transgis website and bike volume per day is get from the Strava database (the summation of major road and minor road). Finally, the functional classification of the minor road is extracted from ODOT TransGIS website. Figure A.8 showed we need to display classification layers on the website and the minor road of the target intersection is an Urban Collector Road.

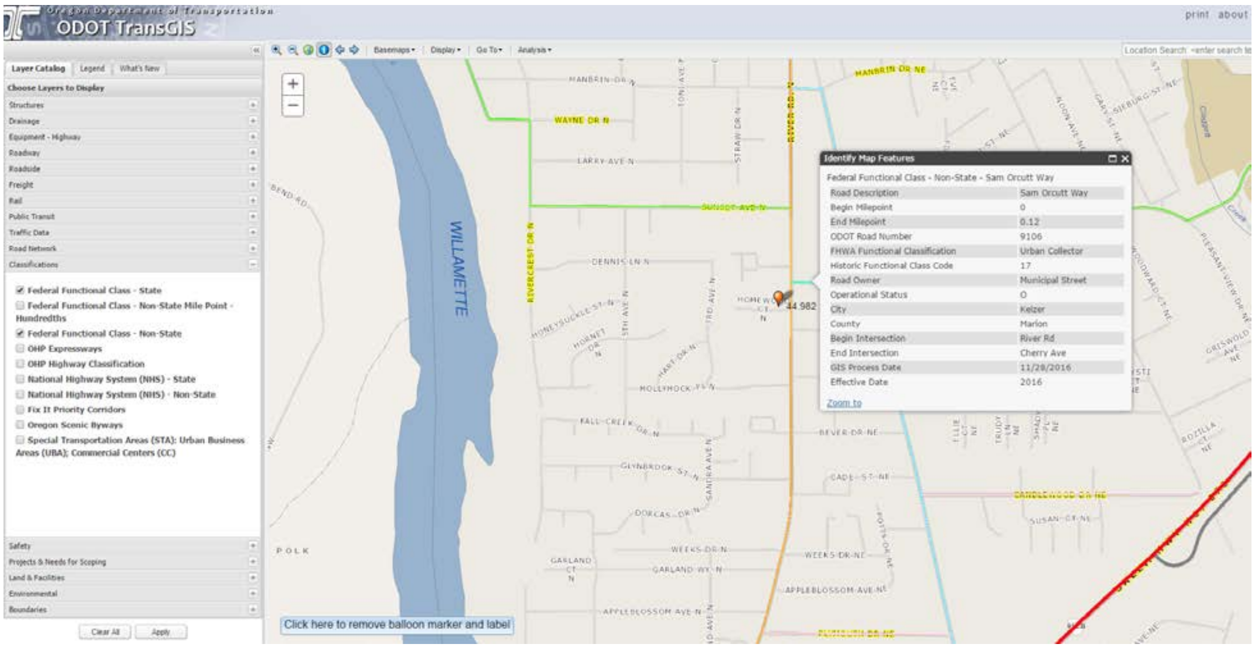


Figure A.8: Classification of Minor Road