DEVELOPING AN INTERACTIVE BASELINE DATA PLATFORM FOR VISUALIZING AND ANALYZING RURAL CRASH CHARACTERISTICS IN RITI COMMUNITIES

Interactive Crash Data Visualization and Analyses in RITI Communities

FINAL PROJECT REPORT

by

Guohui Zhang, Ph.D., Panos D. Prevedourous, Ph.D., David T. Ma, Ph.D.,
Hao Yu, Ph.D., Zhenning Li, Runze Yuan
Department of Civil and Environmental Engineering
University of Hawaii at Manoa

for

Center for Safety Equity in Transportation (CSET)
USDOT Tier 1 University Transportation Center
University of Alaska Fairbanks
ELIF Suite 240, 1764 Tanana Drive
Fairbanks, AK 99775-5910

INE/AUTC 19.15

In cooperation with U.S. Department of Transportation,
Research and Innovative Technology Administration (RITA)

rural · isolated · tribal · indigenous

DISCLAIMER

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No.	**2. Government Accession No.**	**3. Recipient's Catalog No.**

4. Title and Subtitle

Developing an Interactive Baseline Data Platform for Visualizing and Analyzing Rural Crash Characteristics in RITI Communities: Interactive Crash Data Visualization and Analyses in RITI Communities

5. Report Date

Mar. 14, 2019

6. Performing Organization Code

7. Author(s) and Affiliations

Guohui Zhang, Ph.D., P.E., Panos D. Prevedourous, Ph.D., P.E., David T. Ma, Ph.D., Hao Yu, Ph.D., Zhenning Li, Runze Yuan
Department of Civil and Environmental Engineering
University of Hawaii at Manoa

8. Performing Organization Report No.

9. Performing Organization Name and Address

Center for Safety Equity in Transportation
ELIF Building Room 240, 1760 Tanana Drive
Fairbanks, AK 99775-5910

10. Work Unit No. (TRAIS)

11. Contract or Grant No.

12. Sponsoring Organization Name and Address

United States Department of Transportation
Research and Innovative Technology Administration
1200 New Jersey Avenue, SE
Washington, DC 20590

13. Type of Report and Period Covered

Research Report

14. Sponsoring Agency Code

15. Supplementary Notes

Report uploaded to:

16. Abstract

This project focused on developing an interactive baseline crash data platform, termed as Rural Crash Visualization Tool System (RCVTS), to visualize and analyze rural crash characteristics in RITI communities. More than 975 thousand crash records were collected in the state of Alaska, Idaho, and Washington, from 2010 to 2016. Data fusion is applied to unify the collected data. In the proposed RCVTS platform, three main functions are defined: crash data visualization, data analysis, and data retrieval. Crash data visualization includes an on-street map based crash location tool and a graphic query tool. Data analysis involves a number of visualization approaches, including static charts— i.e., the scatter chart—the line chart, the area chart, the bar chart, and interactive graph— i.e., the sunburst chart. Users are allowed to generate customized analytical graphs by specifying the parameters and scale. The three types of authorized users are defined to download crash information in the data retrieval section following corresponding limitations. The proposed RCVTS was illustrated using a sample case with crash records of the State of Alaska. It showed that the proposed RCVTS functions well. Recommendations on future research are provided as well.

17. Key Words

Traffic safety, Data management, Visualization, Web application, Rural area

18. Distribution Statement

19. Security Classification (of this report)	**20. Security Classification (of this page)**	**21. No. of Pages**	**22. Price**
Unclassified.	Unclassified.	73	N/A

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	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	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 (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
^{o}F	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	^{o}C
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

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	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact degrees)		
^{o}C	Celsius	1.8C+32	Fahrenheit	^{o}F
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380.
(Revised March 2003)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

Compared to crashes occurring in urban areas, traffic crashes in rural, isolated, tribal and indigenous (RITI) communities are associated with a series of significant attributes, such as high speed, low seatbelt usage rate, poor weather and pavement conditions, inferior lighting conditions, considerable distractions, etc. Therefore, it is critical to investigate the unique attributes associated with RITI traffic crashes based on data-driven methods. However, the basic data infrastructure needed to develop suitable methodologies is either lacking or incomplete for RITI communities. It is necessary to build up the comprehensive data infrastructure to enhance the ability to develop informed data-driven plans and crash injury mitigation strategies. To address this gap, this project developed an interactive baseline crash data platform, termed Rural Crash Visualization Tool System (RCVTS), to visualize and analyze rural crash characteristics in RITI communities.

Toward this goal, the research objectives include:

1) Gathering and leveraging rural crash data from multiple Departments of Transportation for RITI communities.
2) Designing and building a relational database that stores all crash data and identifying high-risk locations of rural crashes on the statewide selected highways.
3) Conducting data quality control and data consistency check of the relational databases.
4) Developing the onstreetmap-based online rural crash data platform for crash attribute interpretation and visualization.
5) Developing graphic query functions to enhance baseline rural crash data retrievals and analysis.
6) Providing user-friendly interfaces for further studies on safety implications of behavioral characteristics of RITI drivers, passengers, and other roadway users.

The project collected 975,000 crash records from three states, i.e., the State of Alaska, the State of Idaho, and the State of Washington, within the duration from 2010 to 2016. The number of variables applied in these crash records are 266, 262, and 272, respectively. Through a brief description analysis on to selected characteristics in the database, the researchers found that the collected crash records correlate both spatially and temporally. With the collected database, the proposed RCVTS was summarized in both the website workflow and functionality. To provide an intuitive and easy to use virtual environment, the application initializes data processing with a fusion process focusing on three steps: common parameters selection, uniform parameter definition, and crash mapping. It is noted that since the collected crash records were not the raw crash reports, cleaned data was assumed. This study did not consider issues, such as typos, duplications, referential integrity, etc.

The interface design of the RCVTS was quite straightforward, following the guidelines of "overview first, filter, visualization, details-on-demand, and then download." Accordingly, the three main functions, i.e., data visualization, data analysis, and data retrieval, are located under three juxtaposed tags. Under the data visualization tag, RCVTS provides the users a list of filter options including filter type, crash information, environmental condition, passenger condition, and a timeline. A significant feature of RCVTS is that the three functions are tied together via sharing data query results. Once the filtering condition is submitted under the data visualization tag, selected crashes records will be presented in the embedded map. The data analysis and retrieval process are applied to the crash data set presented on the map directly. Visualization approach, including static charts—i.e., the scatter chart—the line chart, the area chart, the bar chart, and interactive graph—i.e., the sunburst chart—are involved in data analysis function. The system allows users to generate customized analytical graphs by specifying the parameters and scale. For data retrieval, only authorized users have the option, with limitations, to download the selected crash data in a comma separated value (CSV) format. Three types of users are defined using different authority with respect to the accessibility of crash information and query frequency.

In summary, the RCVTS provides user-friendly interfaces on safety implications of characteristics of RITI drivers, roads, and environments. This report recommends future research on distributed computing and premium analysis function.

CHAPTER 1. INTRODUCTION

1.1. Problem Statement

Severe traffic crashes have resulted in considerable incapacitating injuries and fatalities, especially in Rural, Isolated, Tribal, or Indigenous (RITI) communities, which have been disadvantaged from a traffic safety perspective across the United States. For example, although rural roads constitute only 40% of Vehicle Mile Traveled (VMT), more 50% of fatalities occur on rural roadways, and about 20,000 people killed annually in rural crashes (NHTSA, 2013). The U.S. Department of Transportation (USDOT) indicates that the fatality rate (fatalities per VMT) for rural crashes is more than twice the fatality rate in urban crashes (NHTSA, 2013). In Hawai'i, the rural crash fatality rate is 195% higher than the urban fatality rate in 2014, and native Hawaiians and other Pacific Islanders are involved in about 26% of motor vehicle traffic fatalities (HDOT, 2014). As clarified by the USDOT Secretary, safety is the highest priority, though comfort, convenience, and cost are all still important for travelers (USDOT, 2014).

The Federal Motor Carrier Safety Administration (FMCSA) developed its 2015-2018 strategic plans to identify four strategic focus areas including "Safety 1st" culture and comprehensive data utilization and leveraging technology (FMCSA, 2017). Compared to crashes occurring in urban areas, traffic crashes in RITI communities are associated with a series of significant attributes, such as high speed, low seatbelt usage rate, poor weather and pavement conditions, inferior lighting conditions and considerable distractions. Therefore, it is critical to investigate the unique attributes associated with RITI traffic crashes based on data-driven methods. However, the basic data infrastructure needed to develop suitable methodologies is either lacking or incomplete for RITI communities. It is necessary, therefore, to build up the comprehensive data infrastructure to better develop informed data-driven plans and crash injury mitigation strategies.

To address this gap, this project aimed to develop an interactive baseline crash data platform to visualize and analyze rural crash characteristics in RITI communities. This research effort has gathered and leveraged existing traffic accident databases with the State of Washington, Idaho, and Alaska, and developed an online system to dynamically retrieve rural traffic crash data and graphically visualize the data for crash attribute analysis. As part of baseline crash data infrastructure establishment, the proposed data platform enabled effective traffic safety program management at all levels in RITI communities to design and implement appropriate countermeasures to mitigate rural crash severities and risks. The proposed interactive baseline crash data platform would be expanded to serve as crash data infrastructure for all the states to set up a solid foundation for the development of effective traffic safety policies and successful public safety campaigns to reduce traffic crash injuries and fatalities in RITI communities.

1.2. General Background

This project aligns well with the CSET Year 1 Project Themes on baseline data establishment in that it develops an interactive, online data platform for rural crash characteristic analysis and visualization to enable the upcoming CSET research, education, and outreach activities in RITI communities. Based on the research tasks, the project team acquired and obtained rural crash data related to RITI transportation safety. The proposed data platform system built up the data infrastructure needed to measure CSET performance and overall contribution to RITI transportation safety over time. This project directly contributes to safety data collection, retrieval, management, visualization, and analysis in the rural and tribal areas. The research tasks clearly address CSET baseline data needs, such as:

- Gather and integrate region-wide multiple-year (2010-2015) RITI community safety-related baseline data;
- Design and implement online data platform and its supporting relational database, such as SQL database to unify data storage and management; and
- Develop methods for RITI community safety data quality control and cleaning.

The developed rural crash data platform will greatly facilitate effective countermeasure development to minimize crash risks and severities in RITI communities. To our best knowledge, based on a thorough literature search, there is no existing literature to document an online baseline crash data platform for safety performance analysis and visualization, which motivated us to develop an interactive baseline crash data platform to visualize and analyze rural crash characteristics in RITI communities.

1.3. Research Objectives

This project aimed to develop a data-driven baseline crash data platform to visualize and analyze rural crash characteristics in RITI communities. Towards this goal, the research objectives were as follows:

- Gather and leverage rural crash data from multiple Departments of Transportation for RITI communities;
- Design and build a relational database that stores all the crash data and identifies high risk locations of rural crashes on the state-wide selected highways;
- Conduct data quality control and a data consistency check of the relational databases;
- Develop the onstreetmap-based online rural crash data platform for crash attribute interpretation and visualization;
- Develop graphic query functions to enhance baseline rural crash data retrievals and analysis;
- Provide user-friendly interfaces for further studies on safety implications of behavioral characteristics of RITI drivers, passengers, and other roadway users.

1.4. Report Organization

Chapter 2 presents a comprehensive review of previous studies that are relevant to this study. Firstly, we reviewed several famous online traffic data management systems: the PeMS by Caltrans, the RITIS by University of Maryland, and the DRIVE-NET by Washington University. Next, we introduced the most frequently applied visualization techniques in the past decade.

Chapter 3 presents the results of the data collection. More specifically, 975 thousand crash records were reported from the states of Alaska, Idaho, and Washington. The time horizon is between 2010 and 2016. Chapter 4 presents the proposed interactive baseline crash data platform in two aspects, i.e. the work flow and the functionality.

Chapter 5 illustrates the crash data analysis and analysis with the help of RCVTS. More specifically, the statistic descriptions were conducted to the key variables associated with crash records in the State of Alaska. Finally, Chapter 6 presents the conclusion of this research and the recommendations for future research.

CHAPTER 2. LITERATURE REVIEW

With the rapid development of networking, data collection and data storage by new technology—traffic big data—representing the new era in data science, is now expanding into the transportation area, especially for crash analysis. In facing the challenges of large data volumes and multiple data sources, we need to develop sufficient ways to manage and analyze the crash data and related information

2.1. Existing Online Traffic Data Management Systems

The majority of research on traffic database management and visualization has focused on traffic mobility performance and vehicle emission impacts based on various sensor data. For example, the online Freeway Performance Measurement System (PeMS) in California is designed to archive and visualize freeway traffic sensor data to provide freeway real-time performance measures (Chen, 2003; Chen et al., 2012; Kwon and Varaiya, 2008; Xuan and Kanafani, 2014). Bertini and Tufte developed the Portland Oregon Regional Transportation Archive Listing system for archiving and analyzing freeway data (Bertini et al., 2005; Tufte, 2010). The researchers in the CATT Laboratory at the University of Maryland, College Park, developed the Regional Integrated Transportation Information System (Hale et al., 2016; Hossan et al., 2016; Jeihani et al., 2015; Jin et al., 2017; Zhang and Haghani, 2015). This database system integrates multiple data sources from different transportation agencies and focuses on freeway applications. Furthermore, Ma and Wang developed a DRIVE-NET system for visualizing real-time traffic conditions and performing online arterial traffic data analysis by using intersection loop data and traffic signal timing data (Ma and Wang, 2014; Ma et al., 2011, 2013; Wei et al., 2016; Zheng et al., 2013).

2.1.1. Freeway Performance Measurement System

The California Department of Transportation (Caltrans) sponsors PeMS. As the first step in effective management and operation of California highway system, PeMS collects real-time data from automatic sensors installed on most freeways, as well as a large amount of historical data. A list of performance measures and other traffic quantities, such as the speed, the vehicle-hours of delay, the vehicle-mile travels, and the travel time statistics, are visualized in plots and summarized in reports. Users get access to the information via a web interface.

PeMS consists of three main functional units: the data collection, the data processing, and the data access. The main data source of PeMS is loop sensors. They measure vehicle counts and freeway occupancy across all lanes and on- and off-ramps. Each sensor sends the measurements every 30 seconds to a computer called the Front End Processor at the traffic management center (TMC). In total, PeMS collects data from six of 12 Caltrans Districts, including Sacramento, Bay Area, Los Angeles, San Bernardino, San Diego, and Orange County. The data processing calculates the quantities not directly measured and derived from the raw data. Key steps are computing derived values, such as speed, VMT, VHT, delay and travel time, diagnosing data errors, aggregating data geographically and temporally, and fusing data.

2.1.2. Regional Integrated Transportation Information System

RITIS is an automated data sharing, dissemination, and archiving system that helps traffic professions and agencies to gain situational awareness, measure performance, and communicate information with the public. It involves various transportation related data available from the public and private agencies, such as traffic flow characteristics, traffic incident information, weather data, traffic infrastructure status, geographical information, and in-route communication data.

In general, RITIS provides 36 web-based analysis tools, classified into six groups, operations, planning, research, developer sources, travel information, and data retrieval. Those tools enable the support of core functions of the agencies or organizations, such as maintaining situational awareness beyond local coverage areas, understanding how transportation influences law enforcement activities and vice-versa, collaborating with and sharing information among peer agencies, demonstrating competent stewardship of public funds, and testing new data, new strategies.

2.1.3. DRIVE-NET System

DRIVE-NET is a regional web-based transportation decision support system. The DRIVE-NET provides data layers for integrating a variety of data sources based on digital roadway map. It demonstrates potential for use as a standard tool for incorporating more data sets from different fields and as a platform for real-time decision-making.

DRIVE-NET archives data via four ways, i.e., direct uploading, periodic downloading, active data acquisition, and direct data archiving. Typical data sources are freeway loop data, INRIX data, HERE data, Washington state’s incident response (WITS) data, weather station data, road geometric data, ferry data, park and ride data, transit data, bicycle and pedestrian count data, Car2go data, and interstate freeway elevation data. Based on the present data, DRIVE-NET provides three aspects functionalities, i.e., database functionality, visualization functionality, and analytical functionality.

2.2. Traffic Data Visualization

Data visualization applied visual means to represent datasets (Chen et al., 2015); transforming various types of data into suitable visual representations, so that data understanding and analysis can be completed efficiently. Data visualization is beneficial for its incorporation of human capabilities into an intuitive visual interface, thereby combining machine intelligence with human intelligence. Specifically, traffic data visualization can enhance understanding of the behavior of moving objects and discovery of traffic, socioeconomic, and geometric patterns. Chen et al. (2015) conducted a survey of traffic visualization and introduced the basic concepts of various traffic data visualization approaches. Moreover, they established state-of-the-art methods for depicting the temporal, spatial, numerical, and categorical properties of traffic data, used by many traffic data analysis projects and intelligent transportation systems (Kloeckl et al., 2016; Pu et al., 2013). Shekhar et al. (2002) proposed a web-based traffic data visualization system, termed CubeView. Traffic data was modeled as a multi-dimensional data warehouse to facilitate the online query processing. Du et al. (2015) also established a visualization tool, termed MetroViz, which explores public transportation data and evaluates the performance of public transportation systems. Typical visualization techniques applied in traffic data ware summarized in Table 2.1 and were designed for time, special, spatiotemporal information, and other properties.

Table 2.1 Existing Visualization Techniques

Object	Type	Techniques	Related Studies
Time	Linear Time	line chart, stack graph, theme river	Byron and Wattenberg, 2008; Ferreira et al., 2013; Havre et al., 2000
	Periodic Time	radial chart	Pu et al., 2013
	Branching Time	storylines	Ogawa and Ma, 2010
Space	Point	dot chart, moving dot, heat map	Barry and Card, 2013; Kloeckl et al., 2016; Xie and Yan, 2008
	Line	line chart, heat map, edge bundling	Crnovrsanin et al., 2009; Ersoy et al., 2011; Lampe and Hauser, 2011; Zhou et al., 2003
	Region	radial metaphor	Zeng et al., 2013
Multiple Properties	Spatiotemporal	Space-time-cube	Kraak, 2003
	Categorical Property	color mapping	Colorbrewer, 2017
	Textual Property	tagCloud, Wordle	Fekete and Plaisant, 1999; Rivadeneira et al., 2007

Despite the numerous data sharing and visualization techniques and platforms that have been developed, a limited transportation platform has been proposed that focuses on traffic crash data sharing, visualization, and analysis.

CHAPTER 3. RURAL CRASH RECORDS

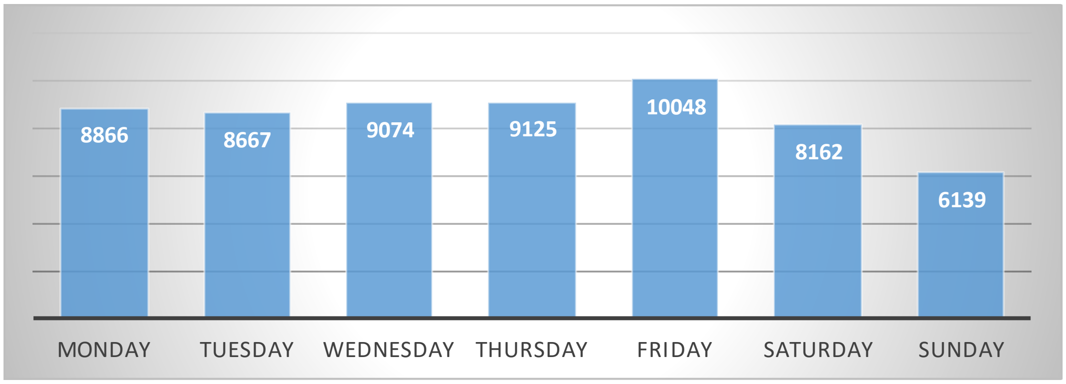
This chapter presents a brief description of the rural crash record collected by the research group, specifically, crash records from three states in northwest regions—i.e., the state of Alaska, Idaho, and Washington—collected and imported into our database. As summarized in Table 3.1, 975 thousand crash records were reported from 2010 to 2016.

Table 3.1 Summary of Crash Record Collection

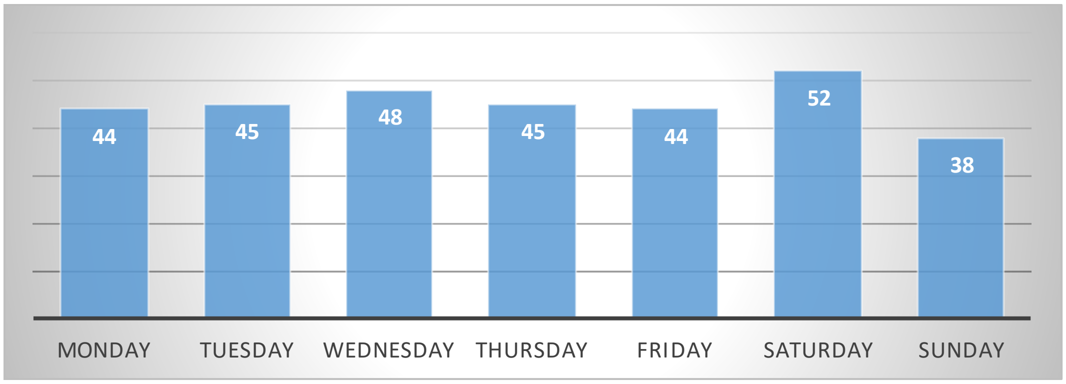
	Year	State		
		Alaska	**Idaho**	**Washington**
Crash Record Count	2010	12399	22633	105175
	2011	12576	21182	102642
	2012	11781	21689	103935
	2013	12422	22586	104127
	2014	10903	22451	112778
	2015	-	24070	123370
	2016	-	-	128154
Parameter count		266	262	272

3.1. Alaska Crash Records

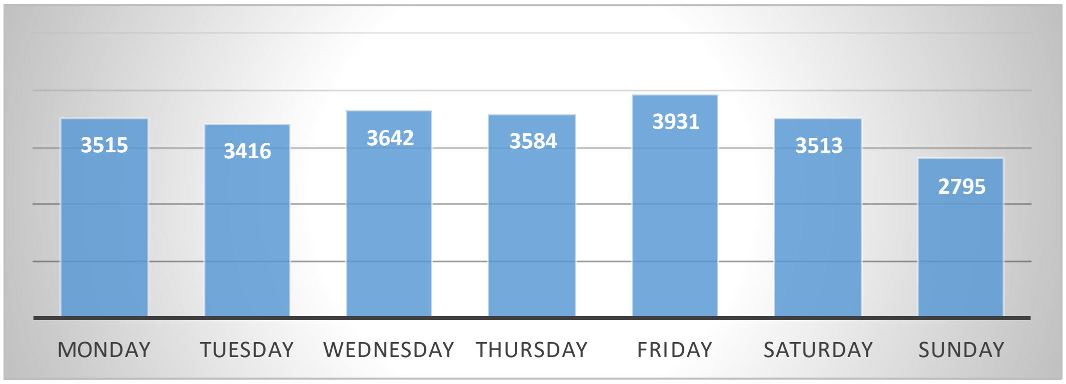
As shown in Table 3.1, 60,081 crashes that occurred in Alaska are in the database. For each record, the analysis employed 266 different variables to present a description for the corresponding crashes. Table A.1 provides the definitions for the non-confidential variables, noting that crash variables for second or more vehicles are omitted. Three hundred sixteen people were killed in the crashes during the five years listed. As shown in Figure 3.1(a), the crash counts increased to the peak on Friday and reached the lowest value on Sunday. The total injury counts shown in Figure 3.1(c) present a similar pattern. However, as indicated in Figure 3.1(b), the highest fatalities occurred on Saturday. The possible explanation is that people injured in Friday crashes eventually die on Saturday. Similarly, Figure 3.2 presents the monthly distribution of crash counts, fatality counts, and total injury counts. Both Figure 3.1 and Figure 3.2 are the temporal visualization for the crash objects.



(a) Crash counts

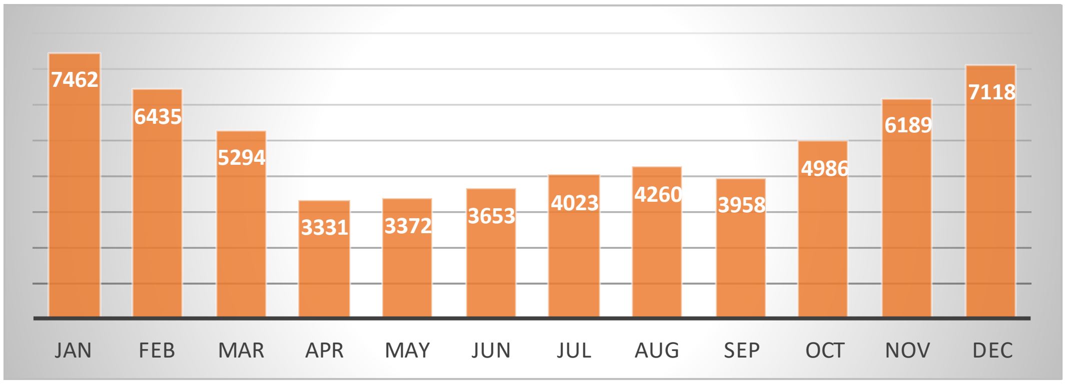


(b) Fatality counts

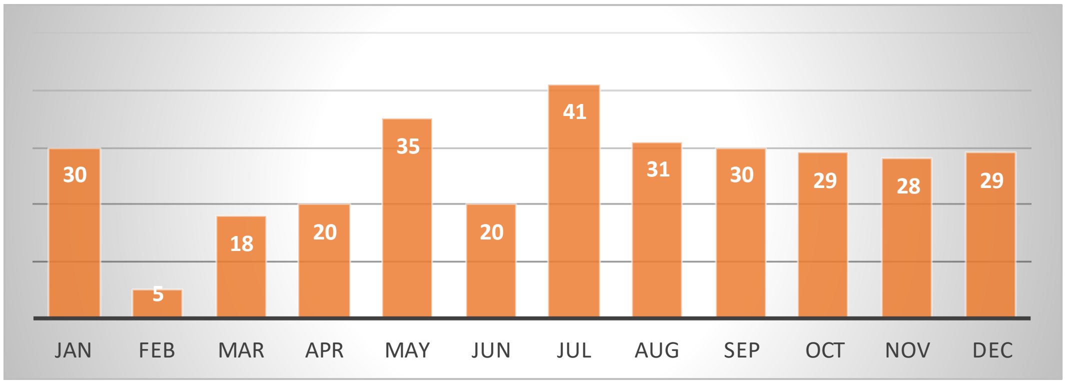


(c) Total injury counts

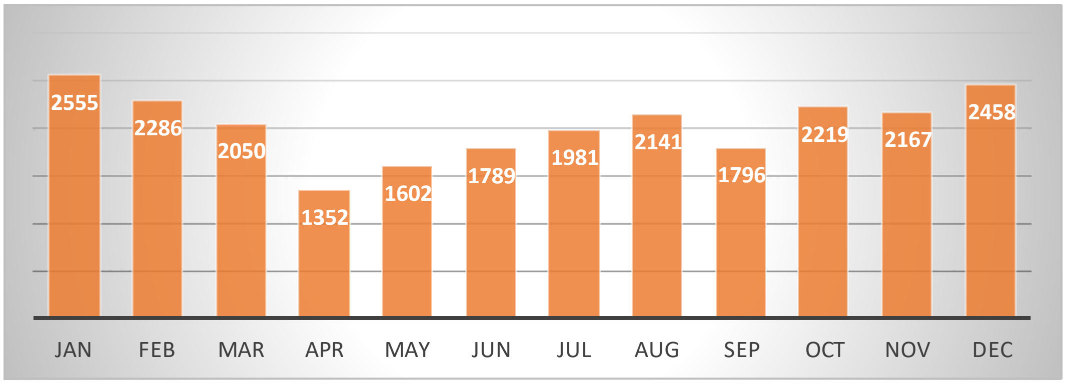
Figure 3.1 Weekly distribution of different crash statistics.



(a) Crash counts



(b) Fatality counts



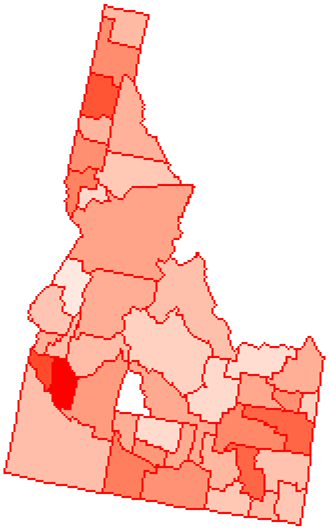
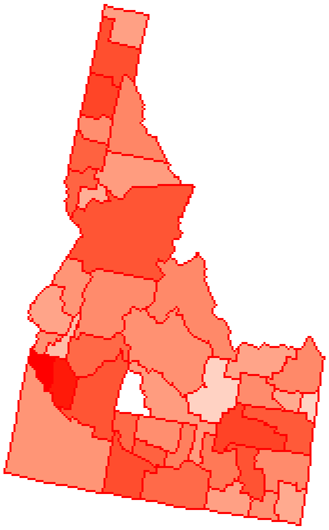
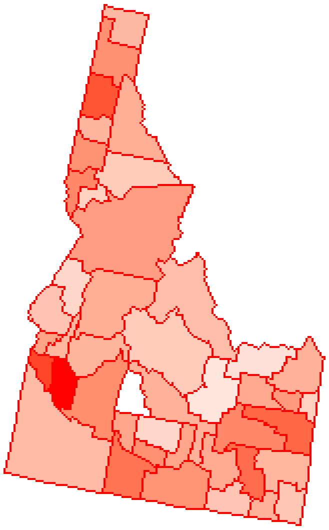
(c) Total injury counts

Figure 3.2 Monthly distribution of different crash statistics.

3.2. Idaho Crash Records

Table 3.1 shows that 136,411 crashes occurred in the State of Idaho from 2010 to 2015. Each crash record includes more than 262 variables, including the crash description, police information, geometric design, and injury conditions. Table A.2 shows detailed information of the non-confidential variables, as well as their definitions.

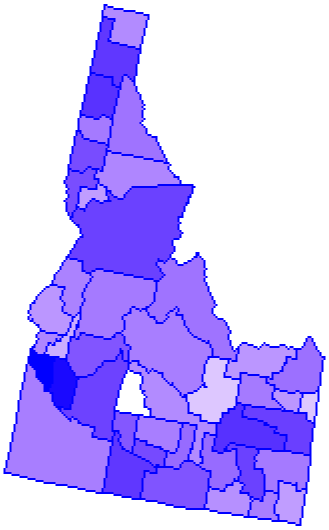
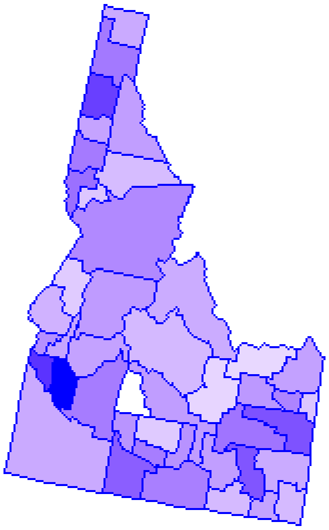
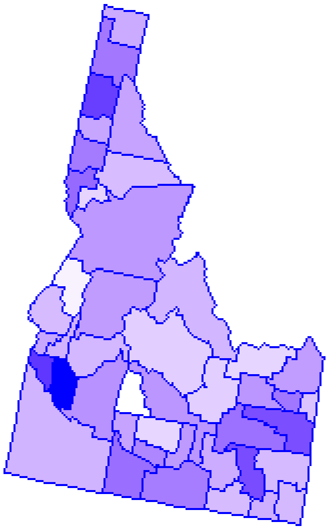
With the information provided in the crash database, Figure 3.3 illustrates the spatial distributions for different crash statistics, presenting crash counts, fatality counts, and total injury counts at the county-level using different shade levels. In different counties, the crash counts range from 111 to 35089; the fatality counts range from 1 to 109; and the total injury counts range from 48 to 19790. It is clear that the total injury counts are spatially distributed in line with the distribution of crash counts, while the fatality crashes are distributed more evenly, as the differences in the shades of different counties are smaller.

(a) Crash counts (b) Fatality counts (c) Total injury counts

Figure 3.3 Spatial distribution of different crash statistics at the county level

Figure 3.4 illustrates the spatial distribution of different crash severities, i.e., the fatality crash, the injury crash, and the property damage only (PDO) crash. In different counties, the fatality crash counts range from 1 to 99; the injury crash counts range from 28 to 13946; and the PDO crash counts range from 83 to 21088.

(a) Fatality crash (b) Injury crash (c) Property damage only crash

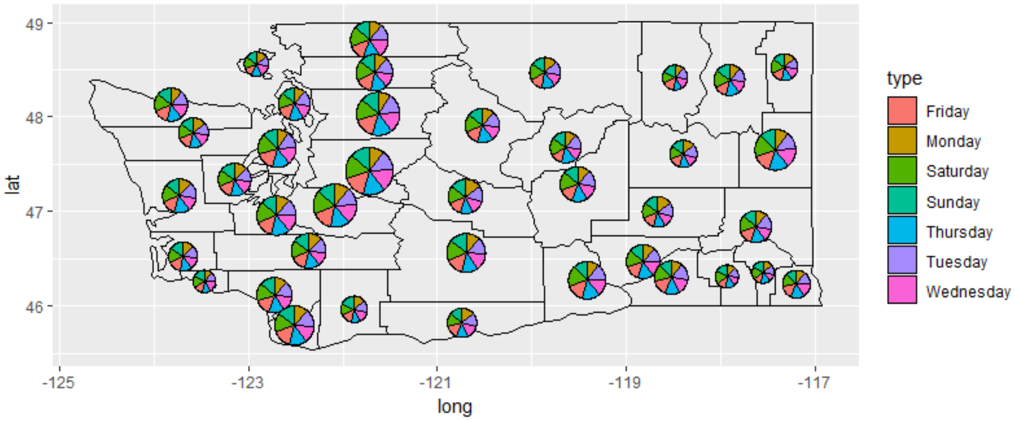
Figure 3.4 Spatial distribution of crashes for different severities at the county level.

3.3. Washington Crash Records

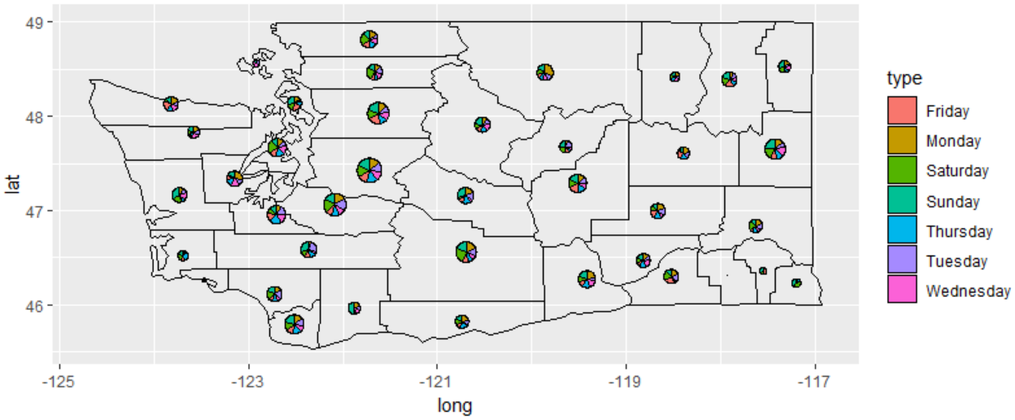
In our database, crash records from the state of Washington made up the greatest proportion, 780,181 crash records were collected. Each record contained at least 272 variables with one involved passenger, among which key variables are summarized in

Table A. 3. It was understandable that the higher population size in the state of Washington resulted in the larger crash record size.

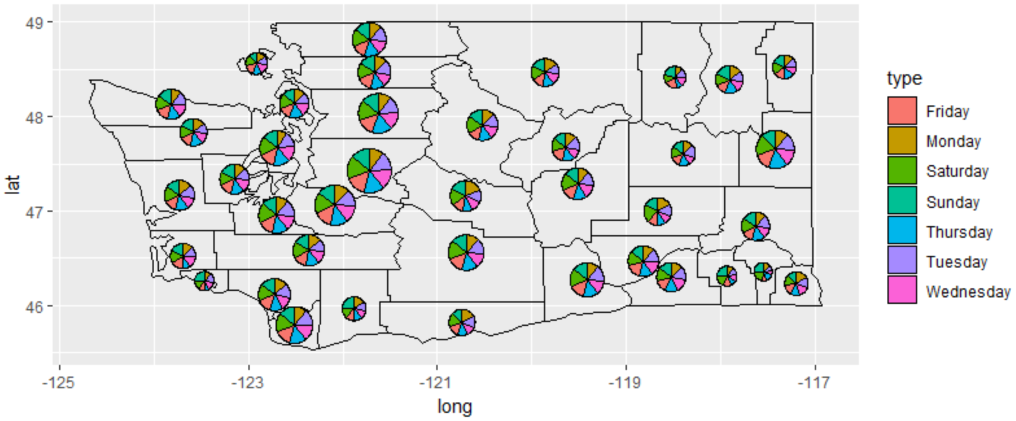
Figure 3.5 illustrated the spatial distribution of crash counts, fatality counts, and injury counts at the county level. As shown in Figure 3.5(a), crash counts for both weekdays and weekends were presented using a pie chart. The radius of each pie chart was equal to the natural logarithm of the total crash count in the corresponding states. In different counties of Washington State, the total crash counts varied from 357 to 284,565. More specifically, Garfield county obtained the lowest crash count, while the King county obtained the highest crash count. The analysis found that the weekly variation in the state of Washington is different from the state of Alaska; i.e., in state of Washington, the highest crash count occurred on Saturday, and the Monday is the safest.



(a) Crash counts

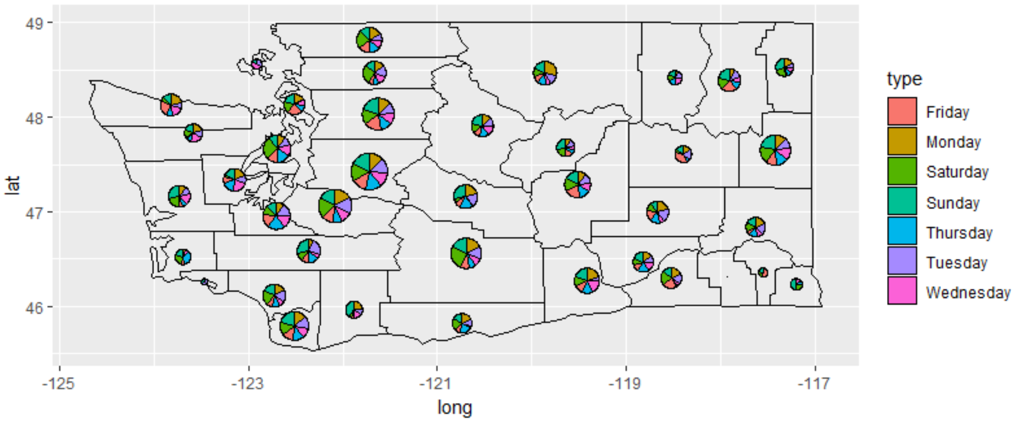


(b) Fatality counts

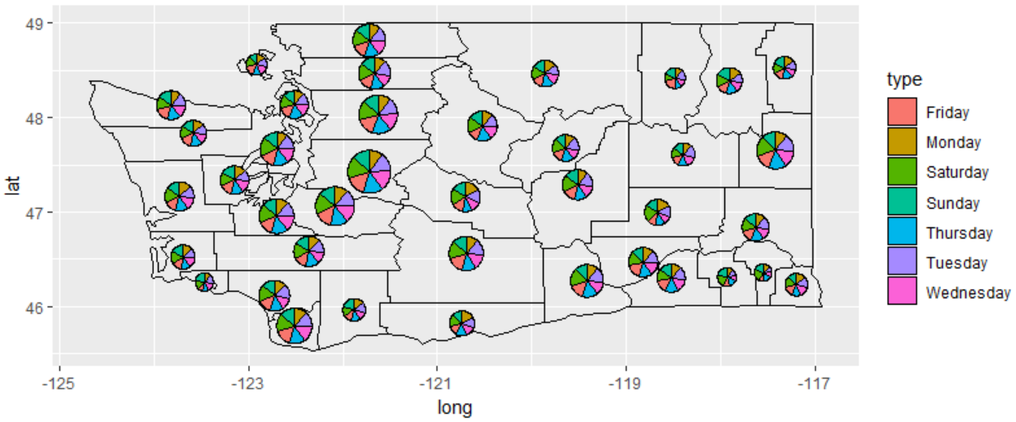


(c) Injury counts

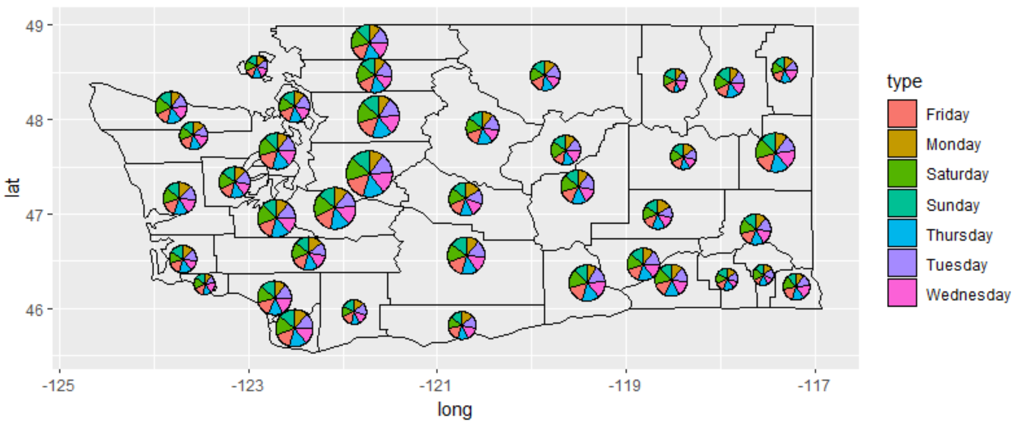
Figure 3.5 Spatial distribution for different crash statistics considering weekly variation.



(a) Fatality crash



(b) Injury crash



(c) PDO crash

Figure 3.6 Spatial distribution of crash severities considering weekly variation.

Figure 3.6 illustrates the spatial distribution of crash severities in the same way. Three different crash severities are presented, i.e., the fatality crash, the injury crash, and the property damage only (PDO) crash. The proportions for different severities in each county are shown in Figure 3.7.

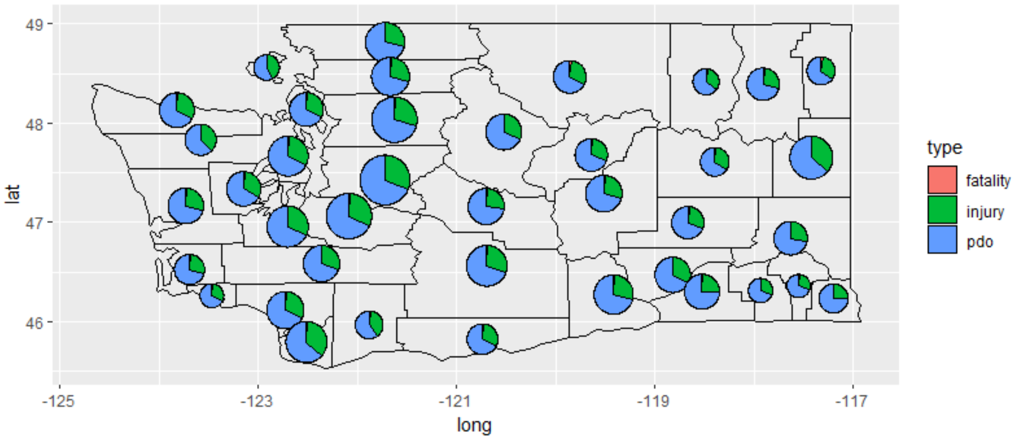


Figure 3.7 Spatial distribution of severity proportions in each county.

CHAPTER 4. INTERACTIVE ONLINE CRASH DATA PLATFORM

In this chapter, we describe the initial development of an interactive baseline crash data platform, i.e., Rural Crash Visualization Tool System (RCVTS). RCVTS is a web-based tool that deals with visualization issues associated with rural crash records, which enables effective traffic safety analysis.

As a part of the baseline, the proposed data infrastructure establishment enables the design and implementation of appropriate countermeasures to mitigate crash severities and risks. This infrastructure will help create a solid foundation for the development of effective traffic safety policies and successful public safety campaigns to reduce traffic crash injuries and fatalities. The remainder of this chapter is organized as follows. Firstly, we present the workflow of the proposed website, including the data preprocessing and interface description. Then, the functionality of our website— i.e., crash data mapping, visualization, and data retrieval—is described in-depth. In the last section, we conclude our work with a discussion on future research.

4.1. Website Workflow

4.1.1. Data Processing

As indicated in CHAPTER 3, more than 350 parameters were employed in a crash report, assuming that three users were involved. By comparing Table A.1-

Table A. 3, it is clear that parameters provided in the crash record were different from state to state. To provide an intuitive and easy to use virtual environment, data processing was initialized with a fusion process focusing on the following steps:

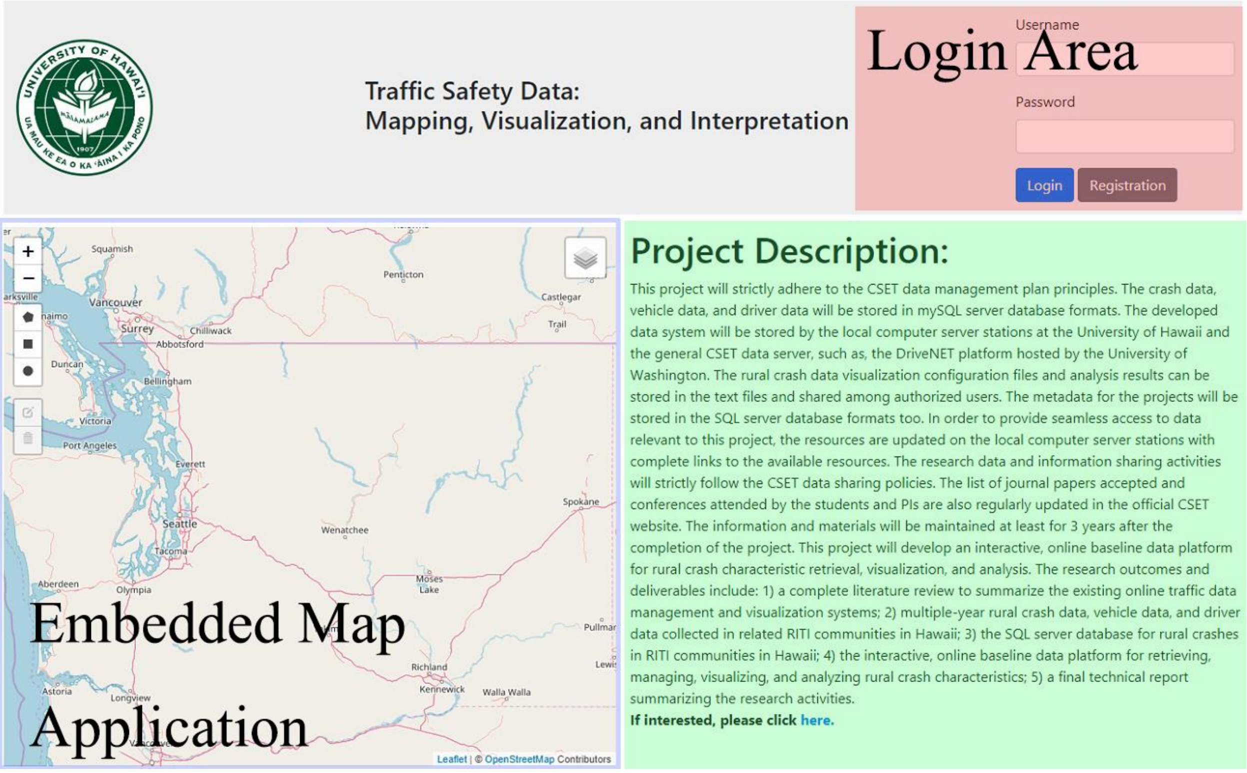
- Comparison of the parameters listed from 2010 to 2016 in the three states' crash records to group them into two types, i.e., the common parameters and the unique parameters. Only the common parameters were used in the proposed visualization platform, while the unique parameters did not enter the analysis process.
- After selection of the common parameters, data formats, such as time label, float or integer variables and the definitions for dummy parameters, such as the contribution factors, crash type descriptions, were unified by reprocessing the combined data.
- Additionally, in terms of the GIS-based map application, the state plane coordinates provided in the crash record were then converted to corresponding latitudes and longitudes. A segment-based matching algorithm was applied to location information—i.e., the latitude and longitude pair—so that the crashes can be exactly located onto the corresponding roads.

It is noted that since the collected crash records were not the raw crash reports, cleaned data was assumed. This study did not consider issues, such as typos, duplications, referential integrity, etc. As a result, more than 270 million entries of crash characteristics were stored in a MySQL database using phpMyAdmin. In order to accelerate the database query for each of these entries, indexes were generated based on the most critical information, including crash location and crash severity. The indexes reduced the query time by 90%.

4.1.2. Description of Interface

Similar to the PeMs system, RCVTS starts with a login page distributed in three sections, which are the login area, project description, and embedded map application, as shown in Figure 4.1. Users are able to log in or to register for an account. Except for the administrator, three types of user authority were regulated (see Table 4.1). The interface asked users at the beginning of the registration process if they were using RCVTS for academic purposes. If so, they registered as a researcher, otherwise as a public user. When a registered researcher used the data retrieve function for the first time, an application of authority would be sent to the administrator and the user would have to wait until approved.

In Figure 4.1, the lower right presents some guideline information describing the whole tool system. The embedded map application located in the lower left part was created using the leaflet JavaScript library, which served as a sandbox for the users who are authorized to log in to see an example of what to be expected. It would contain a sample data set which users could manipulate using the mapping function to change the map preference to Google satellite, street, hybrid, or the original leaflet map shown by default. Users are able to specify the favorite map preference as default.



Traffic Safety Data:
Mapping, Visualization, and Interpretation

Project Description:

This project will strictly adhere to the CSET data management plan principles. The crash data, vehicle data, and driver data will be stored in mySQL server database formats. The developed data system will be stored by the local computer server stations at the University of Hawaii and the general CSET data server, such as, the DriveNET platform hosted by the University of Washington. The rural crash data visualization configuration files and analysis results can be stored in the text files and shared among authorized users. The metadata for the projects will be stored in the SQL server database formats too. In order to provide seamless access to data relevant to this project, the resources are updated on the local computer server stations with complete links to the available resources. The research data and information sharing activities will strictly follow the CSET data sharing policies. The list of journal papers accepted and conferences attended by the students and PIs are also regularly updated in the official CSET website. The information and materials will be maintained at least for 3 years after the completion of the project. This project will develop an interactive, online baseline data platform for rural crash characteristic retrieval, visualization, and analysis. The research outcomes and deliverables include: 1) a complete literature review to summarize the existing online traffic data management and visualization systems; 2) multiple-year rural crash data, vehicle data, and driver data collected in related RITI communities in Hawaii; 3) the SQL server database for rural crashes in RITI communities in Hawaii; 4) the interactive, online baseline data platform for retrieving, managing, visualizing, and analyzing rural crash characteristics; 5) a final technical report summarizing the research activities.

If interested, please click here.

Figure 4.1 Web-based rural crash visualization tool system: login page

Table 4.1 Authorities for different user

User Type	Target User	Available Data	Function
I	Public User	3-year-data (2010-2013)	Static Plot
II	Registered Researcher	All Data	Data Mapping Static & Interactive Graph
III	Authorized Researcher	All Data	Data Mapping

User Type	Target User	Available Data	Function
	Related Officials		Static & Interactive Graph Data Retrieve

After entering the correct login information, the users are presented with the main interface in RCVTS, (see Figure 4.2). The design of RCVTS is quite straightforward, following the guidelines of "overview first, filter, visualization, details-on-demand, and then download" (Shneiderman, 1996). On this page, the description area is replaced with the functional area. The three main functions—i.e., data visualization, data analysis, and data retrieval—are located under different tags. Under the data visualization tag, RCVTS provides the users a comprehensive filter options including filter type, crash information, environmental condition, passenger condition, and a timeline. A significant feature of RCVTS is that all these seemingly independent components are tied together. Once the filtering condition is submitted under data mapping tag, selected crashes records will be presented in the embedded map. The data analysis and retrieval process applied to the crash data set presented on the map directly, i.e., the filter results are shared within the three components.

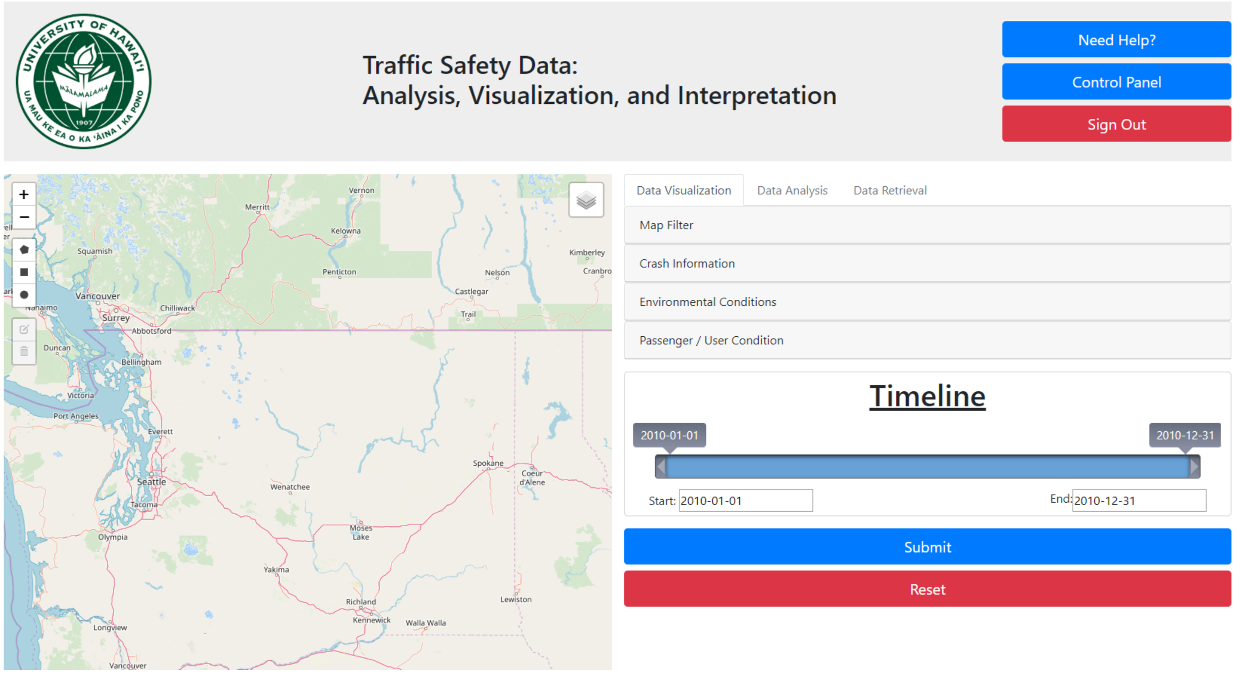


Figure 4.2 Web-based rural crash visualization tool system: functional interface.

4.2. Functionality

In this section, we discuss in detail the three major functions of RCVTS, i.e., the crash visualization, crash data analysis and crash data retrieval.

4.2.1. Crash Visualization

As mentioned before, the RCVTS provides a rich set of filter options. Table 4.2 summaries the filter options in the four categories. Firstly, users are required to choose the filter type, i.e., by area or by road. When users choose by area, they have the option to query the database based on the state of crash, county, and the city town the accident occurred. Otherwise, the user has the option to query the database based on the road type and road name. RCVTS provides three road types, i.e., city/street, state route, and county road. If a user chooses the city/street option, they can query based on the name of the primary roadway. If they choose the state route option, users can query based off the state route id. If the county road option is chosen, the user can query based on the county road number. The RCVTS populates the options using a php program dynamically querying all the options based on the dataset.

Table 4.2 Summary of filter options provided in RCVTS

Filter Group	Subfilter	Filter Options
Map Filter[a]	By Area	State, County, City/Town, etc.
	By Road	Road type, Road name
Crash Information		Severity, First collision type, Second collision type, Number of involved vehicle, Number of involved user, Major contribution, etc.
Environmental Conditions		Weather, Road surface condition, Light condition
Passenger/User Condition	Driver 1	Gender, Age, Vehicle Type, Injury Type, Seat Position, Alcohol test result, etc.
	Driver 2[b]	Gender, Age, Vehicle Type, Injury Type, Seat Position, Alcohol test result, etc.
	Driver 3[b]	…

[a] Users shall choose a filter type from either “By Area” or “By Road.”

[b] The number of Driver information here depends on the number of involved vehicles entered in crash information filter.

Figure 4.3 illustrates typical query results. A successful query indicates that the database contains data that meets the filter conditions. As shown in Figure 4.3, in the desired area, 21046 crash records are identified. If no crash records satisfy the recent query, the interface generates a popup to inform the user. The popup presents a summary of submitted filter conditions for user’s convenience, as shown in Figure 4.4.

Occlusion has been an issue when dealing with nearby crash records because it is difficult to count overlapping points. Wongsuphasawat solved this problem using the hot mode, in which regions with the most influence are colored in reds while less interesting areas are colored in blue (Wongsuphasawat, 2009). In RCVTS, the solution is quite straightforward: crash records are grouped into certain clusters with a label indicating the total number of crashes in this cluster (see Figure 4.3). This solution is beneficial for the following reasons:

- It reduces the overhead cost for creating each individual marker on the map.
- It prevents overlapping of multiple markers.
- The labelled number illustrates the density of crashes directly.

As shown in Figure 4.5, users can zoom in and get a more detailed distribution of crashes in this area. The color for different cluster represents the crash counts in a hot mode. When it cannot be zoomed anymore, each marker presents a crash, and by clicking the crash mark, the interface provides crash-related information to the user, as shown in Figure 4.6. Note that, in RCVTS, zooming in can be achieved either by scroll or by double-clicks.

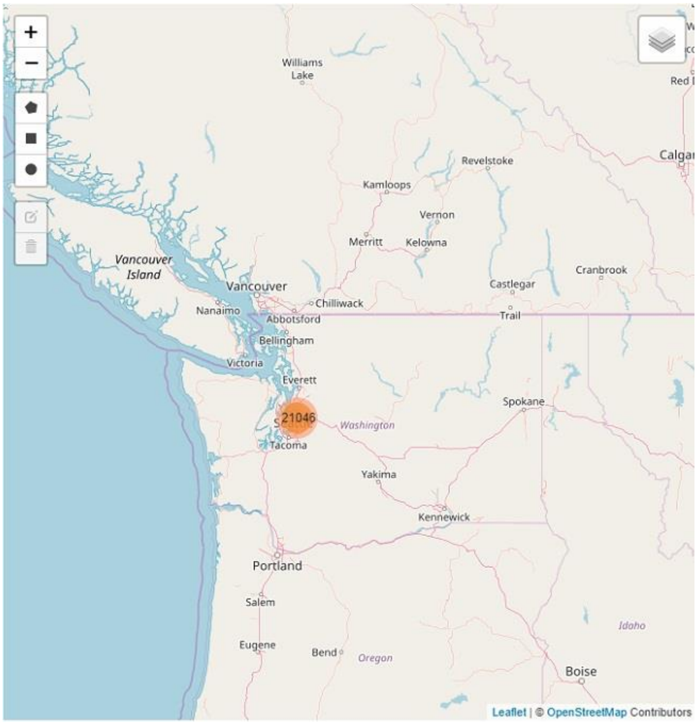
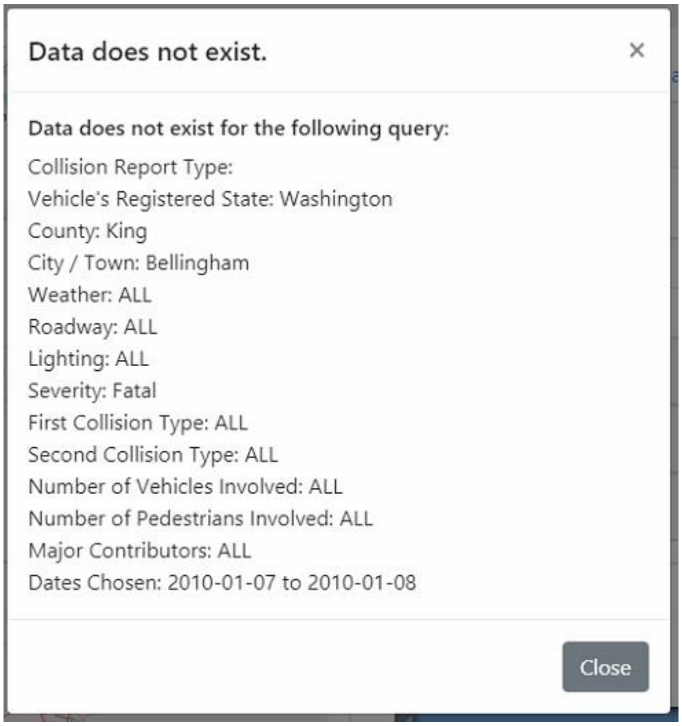


Figure 4.3 Successful query result for RCVTS



Data does not exist.

Data does not exist for the following query:

Collision Report Type:
Vehicle's Registered State: Washington
County: King
City / Town: Bellingham
Weather: ALL
Roadway: ALL
Lighting: ALL
Severity: Fatal
First Collision Type: ALL
Second Collision Type: ALL
Number of Vehicles Involved: ALL
Number of Pedestrians Involved: ALL
Major Contributors: ALL
Dates Chosen: 2010-01-07 to 2010-01-08

Figure 4.4 Pop-up with failure information.

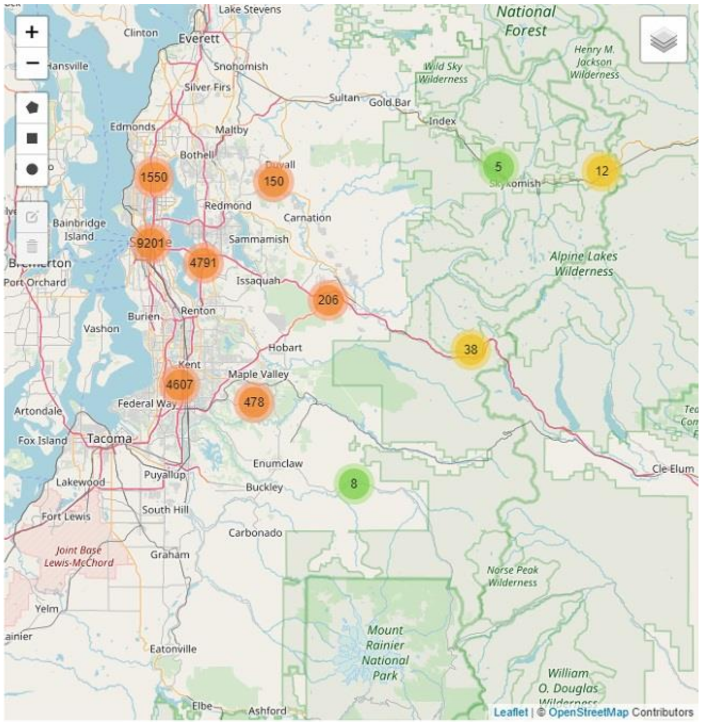
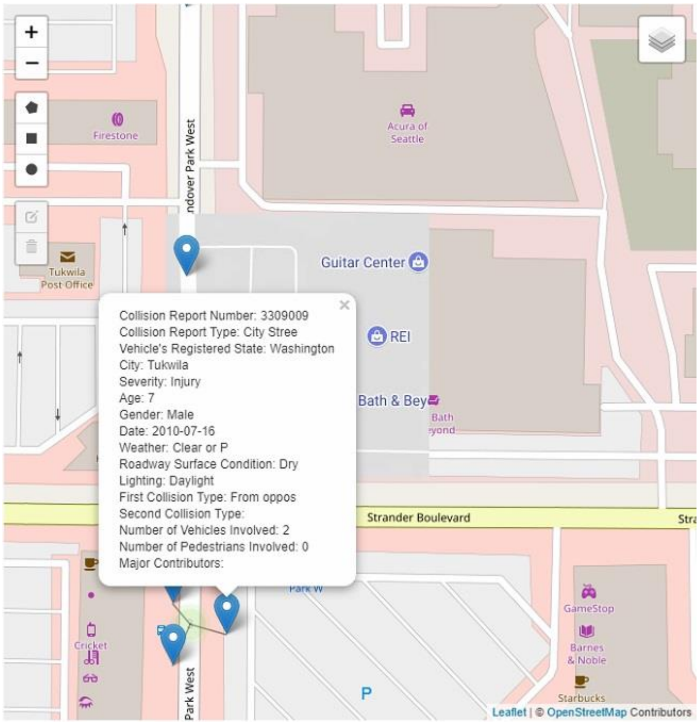


Figure 4.5 Zoom in result in crash query.

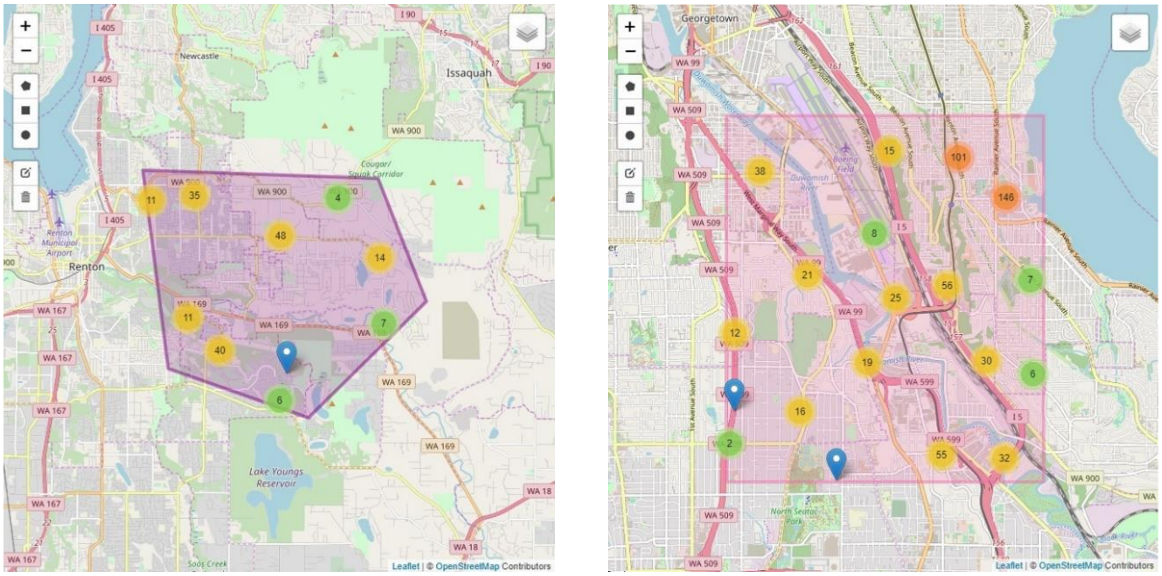


Collision Report Number: 3309009
Collision Report Type: City Stree
Vehicle's Registered State: Washington
City: Tukwila
Severity: Injury
Age: 7
Gender: Male
Date: 2010-07-16
Weather: Clear or P
Roadway Surface Condition: Dry
Lighting: Daylight
First Collision Type: From oppos
Second Collision Type:
Number of Vehicles Involved: 2
Number of Pedestrians Involved: 0
Major Contributors:

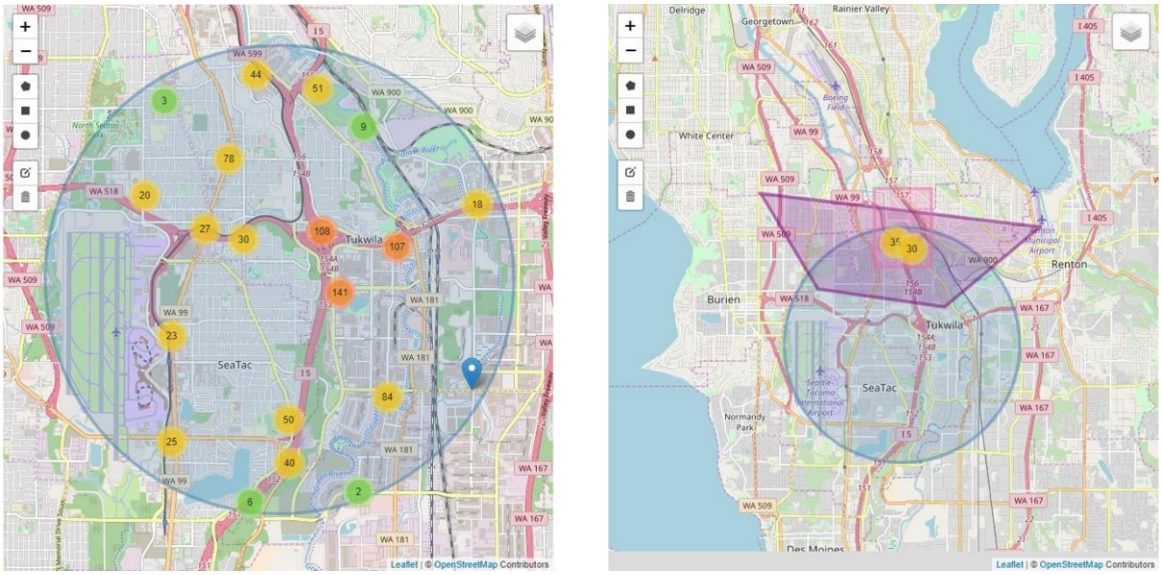
Figure 4.6 Crash detail shown in map-based interface.

To enhance the flexibility of crash data selection, RCVTS also provides a graphic query tool. More specifically, users can choose a specific marker on the map to reshape the area; this allows the user to select crashes in the designated area and remove all crash records outside of that area. Currently three types of marker shapes—i.e., the polygon, the square and the circle—are included, as shown in Figure 4.7.

Furthermore, either users can choose to add more shapes to get an intersection union for a specific area within the designated area already chosen, or they can delete the shape to recall all the markers deleted from the shape, as depicted in Figure 4.7(d).



(a) Polygon (b) Square



(c) Circle (d) Intersection union

Figure 4.7 Result of graph query tool in RCVTS.

4.2.2. Crash Data Analysis

In the past decades, researchers developed a large amount of visualization approaches to illustrate the relationships among human beings, environmental conditions, traffic dynamics, and crashes. RCVTS tends to adopt several widely used visualization approaches to help users explore interesting distributions. As a prototype, visualization approaches, including static charts—i.e., the scatter chart—the line chart, the area chart, the bar chart, and interactive graph—i.e., the sunburst chart—are involved in RCVTS. The interface allows users to generate customized analytical graphs by specifying the parameters and scale. These visualization tools—i.e., the scatter chart, the line chart, the area chart, and the bar chart—can be accessed by selecting the corresponding option located in the lower part under data visualization tag. For example, to generate a line graph users can examine the crash counts, fatality counts, and injury crashes. Users choose the parameter of interest, then they may select the time scale displayed on the graph, e.g., daily, monthly, or yearly. After pressing the create button at the bottom, animation will occur for each point of the line graph, which may reduce the anxiety of waiting.

- Scatter chart

A scatter chart is a type of plot using Cartesian coordinates to display values for typically two variables for a set of data, as show in Figure 4.8. A scatter chart can suggest various kinds of correlations between variables with a certain confidence interval.

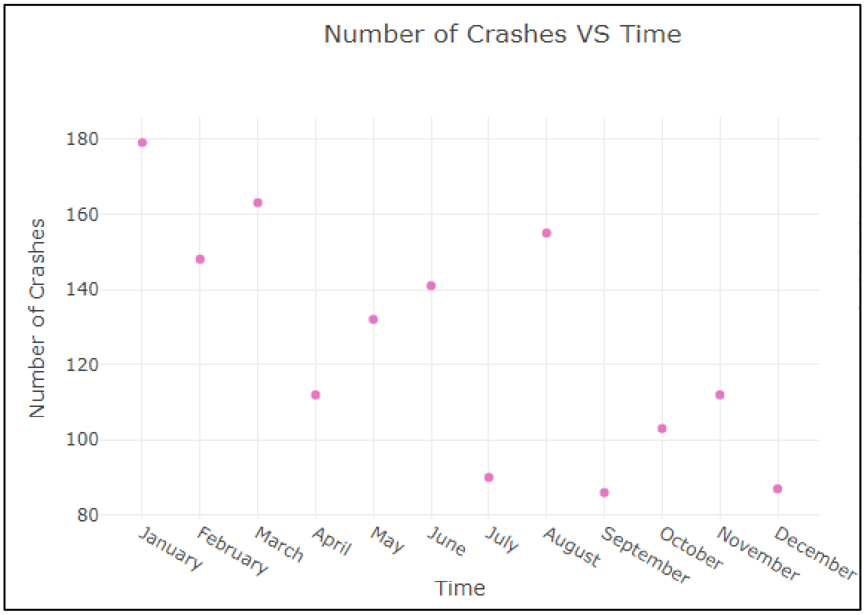


Figure 4.8 Scatter chart sample generated in RCVTS.

- Line chart

A line chart is a type of chart that displays information as a series of data points connected by a straight line. It is a basic type of chart common in many fields. It is similar to a scatter plot except that the measurement points are ordered and joined with straight-line segments. The line chart is often used to visualize a trend in data over intervals of time. Thus, the line is often drawn chronologically. Figure 4.9 presents a sample graph generated in RCVTS.

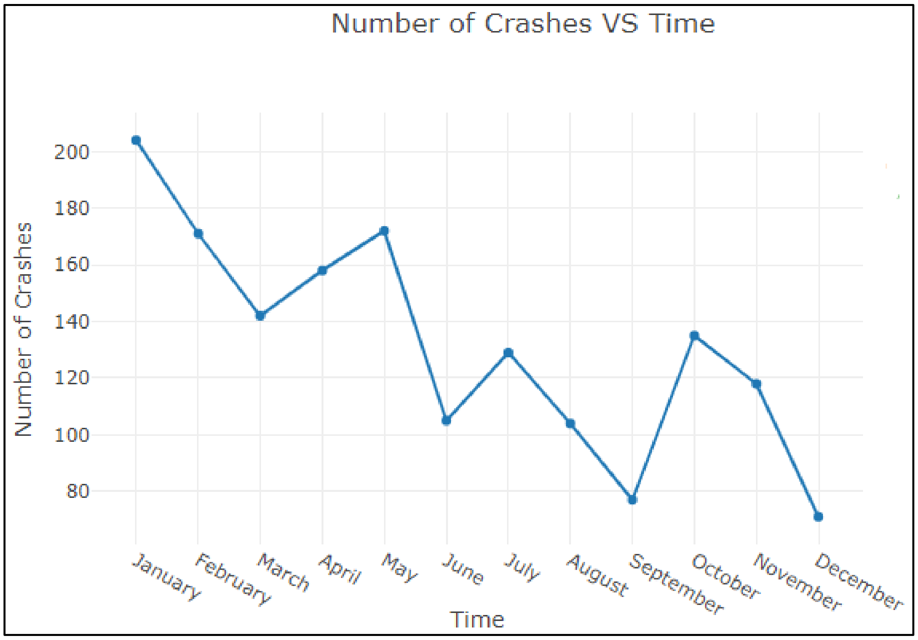


Figure 4.9 Line chart sample generated in RCVTS.

- Area chart

An area chart displays graphically quantitative data. It is based on the line chart. The area between axis and line are commonly emphasized with colors, textures, and hatchings. Commonly one compares two or more quantities with an area chart. Figure 4.10 presents a sample chart.

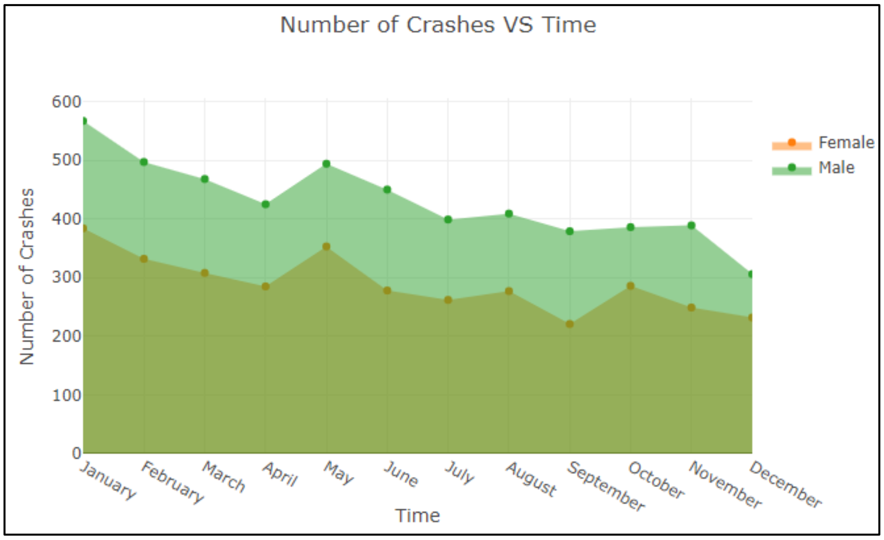
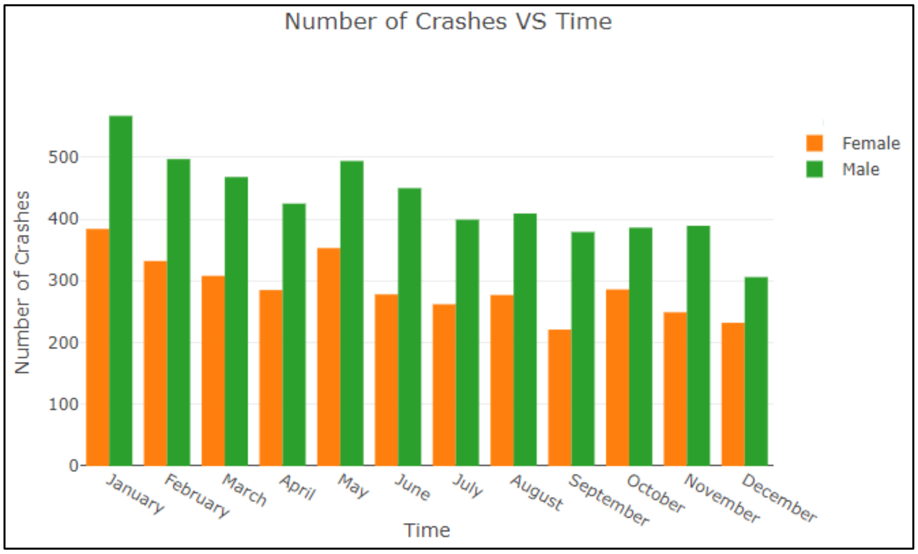


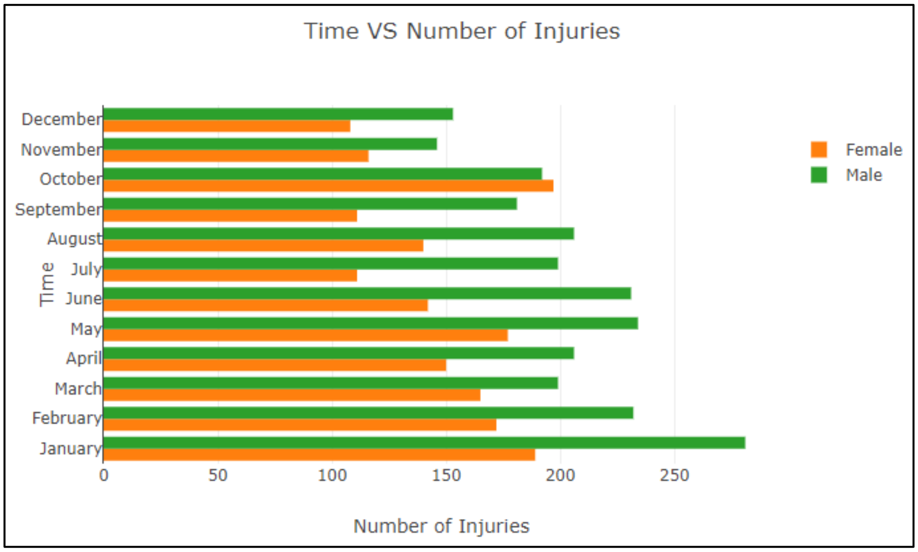
Figure 4.10 Area chart sample generated in RCVTS

- Bar chart

A bar chart is a graph that presents categorical data with rectangular bars with heights proportional to the values that they present. Bar graphs can be used for more complex comparisons of data with grouped bar charts. In a grouped bar chart, for each categorical group there are two or more bars. These bars are color-coded to represent a particular grouping. As Figure 4.11 shows, the number of crashes for different genders makes a grouped bar chart with different colored bars to represent the male and the female: the horizontal axis shows the months of the year and the vertical axis shows the crash counts.



(a) Vertical bar chart



(b) Horizontal bar chart

Figure 4.11 Bar chart sample generated in RCVTS.

- Interactive Graph

Considering the large number of parameters involved in the crash records, the RCVTS uses the interactive sunburst chart to provide the users a directed exploration of the crash dataset (Smith, 2014). It is beneficial for supporting the visualization of various hierarchical relationship and attribute values via a ring-structure without scrolling or other interaction.

In preparing the sunburst chart, crash parameters are counted in a simple top-down scheme for producing a hierarchical pie chart by recursively splitting. Because of this process, each characteristic in a crash record is assigned with a ratio within the upper level characteristic. Figure 4.12 displays a top-level view of the sunburst visualization for a subset of crash records in King County, WA from Mar. 2010 to Aug. 2010. Each characteristic of the crash record is represented as a ring of the sunburst chart. By not labeling each slice within the sunburst, the high-level overview is presented to the user with minimal complexity.

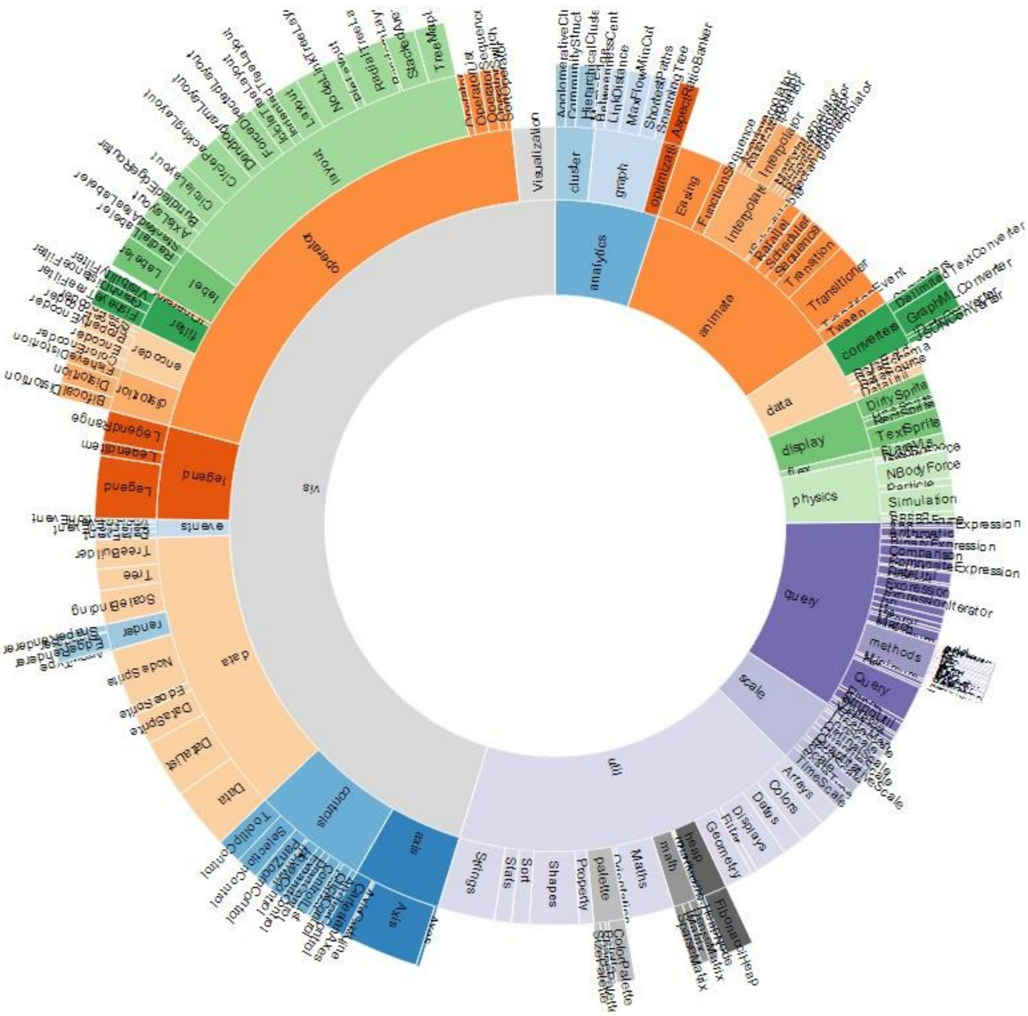
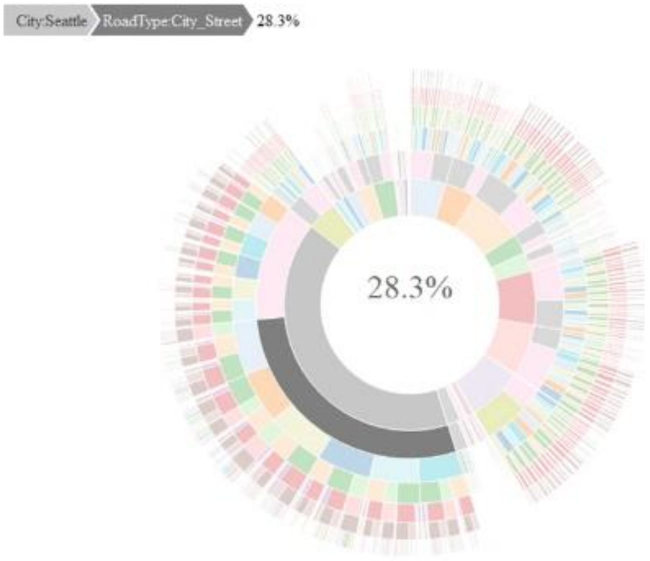
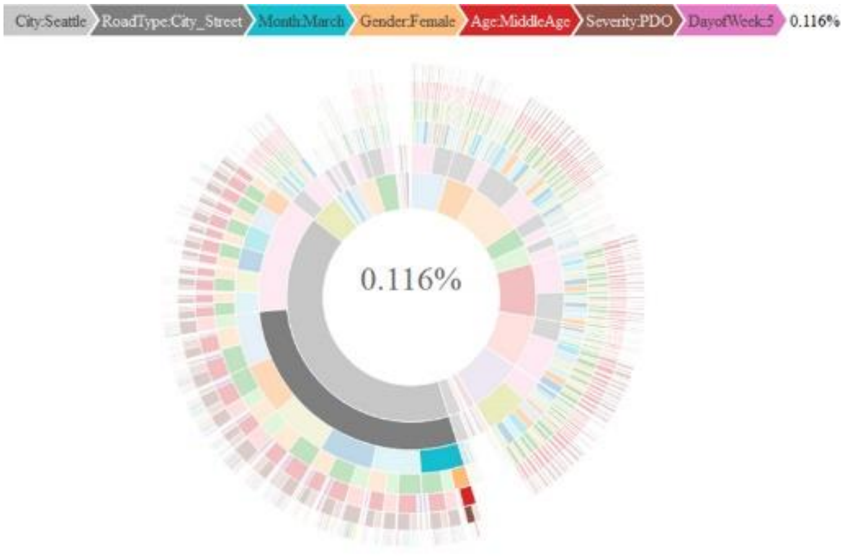


Figure 4.12 Sunburst chart sample generated in RCVTS.

The initial top-level of the sunburst chart, as shown in Figure 4.12, does not display details for every parameter. When a user moves the mouse pointer to a slice of interest, the sunburst presents the detailed information on the top of the graph, as examples presented in Figure 4.13(a) and Figure 4.13(b). Moreover, since the default structure of the sunburst chart is pre-specified, a user is able to select other parameters as the top-level characteristic by clicking the corresponding ring, e.g., Figure 4.14 shows an updated structure if severity is selected as the top-level.



(a) Seattle-City Street Level



(b) Seattle-City Street-March-Female-Middle Age-PDO Crash-Friday Level

Figure 4.13 Sunburst chart in different level of information.

For future work, we will provide more flexibility to the users, such as embedded graphs on a geo-based map, which enables the users to visualize the spatial correlation among different characteristics and multiple graph types in one figure, which enables users to discover different features for specific characteristics. Currently, users can generate figures for different characteristics separately, via submitting query options under data mapping tag.

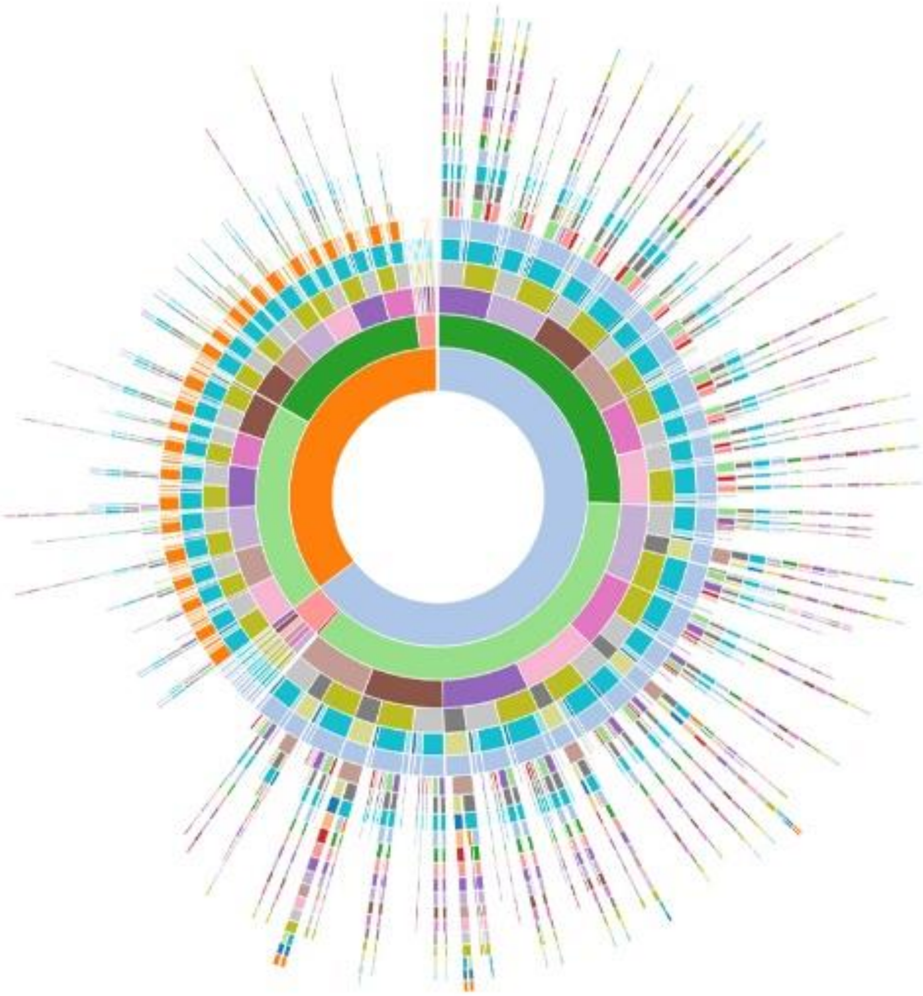


Figure 4.14 Sunburst chart in different sequence.

4.2.3. Data retrieve

As mentioned before, only authorized users have the option to download the selected crash data in a comma separated value (CSV) format with limitations, as shown in Table 4.3.

Table 4.3 Data retrieval limitation

Limitation Type	Description
Frequency	5 queries per day
Quantity	maximum 50000 records per query
Accessible Information	time label, GPS, route name, crash type, severity level, weather condition, lighting condition, major contributing

Currently, access to the raw data is not provided even with authority. On top of the raw data, we plan to enable the access to processed data generated in visualization procedure.

4.3. Summary

The project developed RCVTS, a web-based tool that aims to deal with visualization issues associated with various rural crash characteristics. Our RCVTS features a rich set of filters and various visualization options. It allows users access to traffic data stored in the database, and to create highly customized analytical graphs. Currently, traffic crash data collected in northwest region— i.e., Alaska, Idaho, and Washington—were shared online through a MySQL database using the phpMyAdmin technique. RCVTS regulated three levels of user author with different accessibility to the database and visualization tools. The three major functions provided in RCVTS were traffic data visualization, data analysis, and retrieval of corresponding data. More specifically, in traffic data visualization procedure, a combination of conditional filter and map-based graph query provided the users a flexible data query environment; in the analysis part, different tools were provided based on the type of data.

The researchers hope that the RCVTS application will help transportation professionals to spend less time in crash data analysis and inspire their creativity to investigate the underlying relationships among various parameters. More endeavors are also underway to enhance both the depth and width of the proposed work. The potential improvements include:

- As a prototype of web-based crash visualization tool, only a limited set of graph types are implemented. More functionality should be developed in future study, e.g., spatiotemporal analysis is highly desired in crash data visualization;
- Users will be allowed to upload their own crash data onto the RCVTS. Accordingly, RCVTS would be able to help those professionals as a crash data visualization tool, for not only the data provided in the database, but also their own data.

CHAPTER 5. APPLICATION AND RESULTS

In this chapter, we present the crash data analysis of the Alaska crash record with the help of the proposed RCVTS. More specifically, the data analysis begins with the statistic description, and then all the crashes are plotted onto the map. Further visualization-based analyses are conducted via the RCVTS platform.

5.1. Statistic Description

Among all the 266 variables, 12 variables—i.e., number of involved vehicles, the severity, total injury count, main injury count, minor injury count, fatality count, weather condition, surface condition, light condition, alcohol existing, drivers' age, and gender information—are summarized below.

5.1.1. Number of Involved Units

As shown in Figure 5.1, the number of involved units in crash records from 2010 to 2014 ranged from 1 to 30, and the corresponding frequency ranged from 0 to 40598.

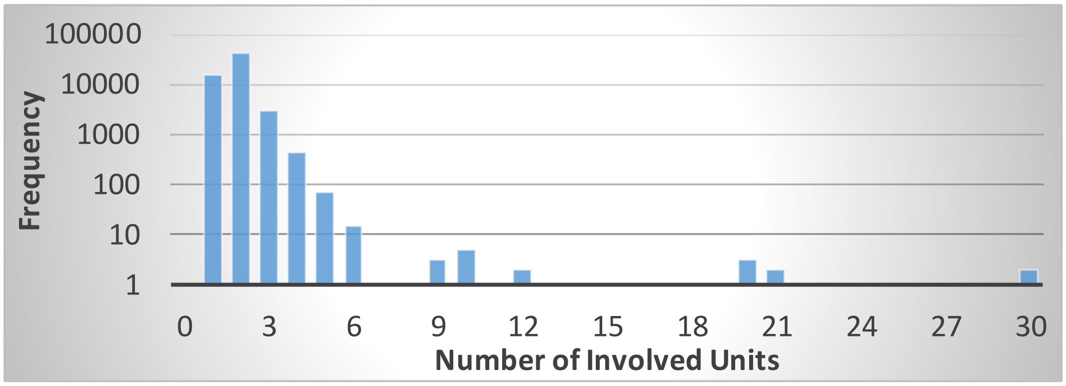


Figure 5.1 Frequency of different number of involved units in Alaska (2010-2014).

Table 5.1 further describes the frequency percentages in the included five years. On average, the most frequent unit counts were two units, which was quite straightforward. Only 1% of crashes occurred with more than four involved units. It was noted that more than a quarter crashes involved only one unit, which indicated that these crashes occurred due to the drivers' themselves.

Table 5.1 Yearly variation of the number of involved units in Alaska (2010-2014)

# of Involved Units	Year					Avg.
	2010	2011	2012	2013	2014	
1	28.78%	24.98%	21.77%	27.84%	28.62%	26.39%
2	65.41%	71.69%	71.69%	66.45%	65.36%	67.57%
3	4.97%	5.64%	5.64%	4.69%	5.15%	5.18%
4	0.71%	0.57%	0.74%	0.81%	0.72%	0.71%
5	0.10%	0.08%	0.14%	0.14%	0.11%	0.11%
6	0.03%	0.02%	0.01%	0.02%	0.03%	0.02%
7	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
8	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
9	0.00%	0.00%	0.01%	0.01%	0.00%	0.00%

# of Involved Units			Year			Avg.
	2010	2011	2012	2013	2014	
10	0.00%	0.01%	0.00%	0.01%	0.02%	0.01%
>10	0.00%	0.00%	0.00%	0.03%	0.01%	0.01%

5.1.2. Crash Severity

In the original database, different classification methods were proposed for different years. In order to uniformly describe the crash records as a whole, four types of crashes severities are applied, i.e., the fatality crash, the serious injury crash, the minor injury crash, and the property damage only crashes, which are defined by the most serious injury occurred in the crash. Figure 5.2 illustrates the crash frequencies for each crash type.

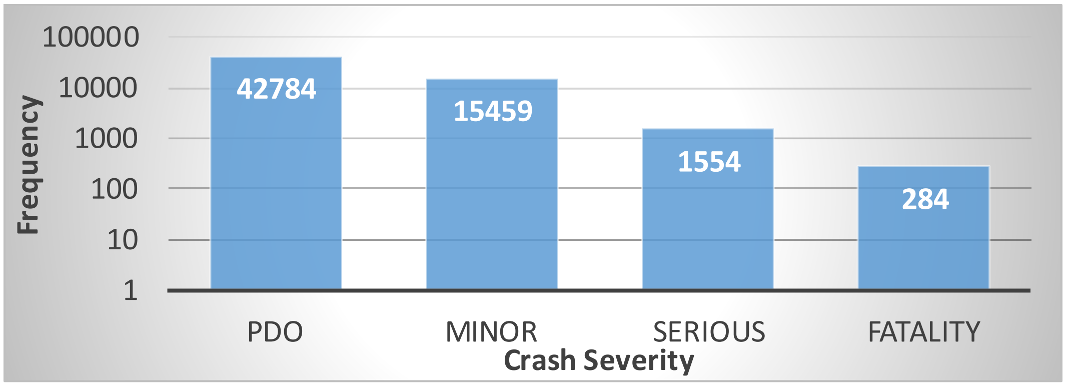


Figure 5.2 Frequency of different crash severity in Alaska (2010-2014).

Figure 5.3 illustrated the yearly variations for different severities during the five years. Unfortunately, analysis found that the crash severities did not improve during the past five years.

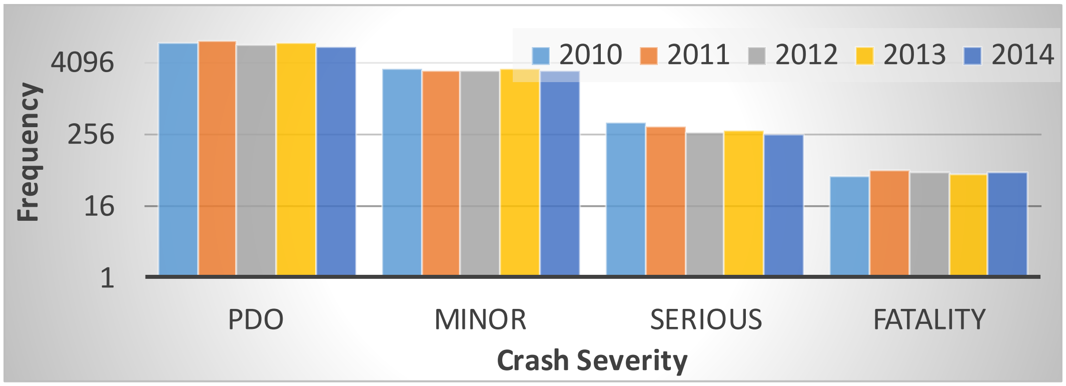


Figure 5.3 Yearly variation of the crash severities in Alaska (2010-2014).

5.1.3. Total Injury Count

By reviewing the whole database, the total injury count ranged from 0 to 19, as shown in Figure 5.4. Those zero injury crashes are also termed as the property damage only events. Table 5.2 described the proportions of different injury counts in each year. It was found that the proportion of crashes with more than 5 injuries is lower than 0.15% in each year. However, the data in Table 5.2 indicated that the number of injuries increased in 2014, up to 12422, but the average value during the past five years was 10903.

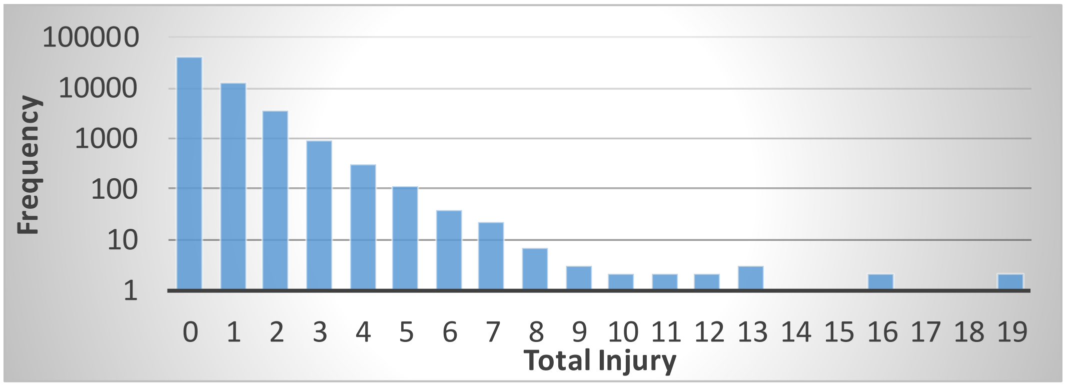


Figure 5.4 Frequency of different total injury counts in Alaska (2010-2014).

Table 5.2 Yearly variation of total injury counts in Alaska (2010-2014)

# of Total Injury	Year					Avg.
	2010	2011	2012	2013	2014	
0	70.57%	73.31%	72.05%	71.94%	68.33%	71.31%
1	20.74%	18.90%	19.84%	20.55%	22.98%	20.55%
2	6.06%	5.60%	5.95%	5.31%	6.15%	5.80%
3	1.60%	1.38%	1.44%	1.44%	1.68%	1.50%
4	0.67%	0.52%	0.45%	0.47%	0.51%	0.52%
5	0.23%	0.15%	0.17%	0.16%	0.20%	0.18%
>5	0.14%	0.14%	0.10%	0.13%	0.15%	0.13%

5.1.4. Minor Injury Count

Minor injury occupied the largest proportion in the crash data sets. For example, in all the 12345 one-injury crashes, there exists 11770 minor injury crashes. Comparison of Figure 5.4 and Figure 5.5 found that crashes with more than six injuries were all minor injury crashes.

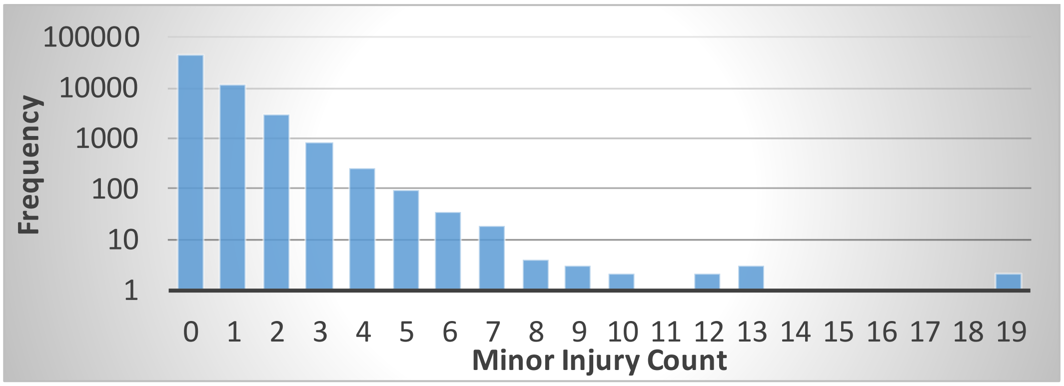


Figure 5.5 Frequency of different minor injury counts in Alaska (2010-2014).

Table 5.3 described the variations on the minor injury crash frequencies in each year. Similar to the total injury crash frequency mentioned before, the minor injury crash frequency increased significantly in 2014, with a sharp reduction on the PDO crashes.

Table 5.3 Yearly variation of minor injury crashes in Alaska (2010-2014)

# of Minor Injury	**Year** **2010**	**2011**	**2012**	**2013**	**2014**	**Avg.**
0	72.75%	75.04%	73.69%	73.57%	70.28%	73.13%
1	19.51%	18.11%	19.14%	19.56%	21.90%	19.59%
2	5.42%	4.92%	5.27%	4.92%	5.67%	5.23%
3	1.51%	1.29%	1.26%	1.30%	1.45%	1.36%
4	0.52%	0.40%	0.41%	0.40%	0.40%	0.43%
5	0.18%	0.13%	0.15%	0.14%	0.17%	0.15%
>5	0.12%	0.11%	0.08%	0.10%	0.12%	0.11%

5.1.5. Serious Injury Crash Frequency

In the past five years, 1608 serious injury crashes occurred in Alaska. Among the 1608 crashes, the number of one-person injury crashes was 1397; the number of two-person injury crashes was 163; the number of three-person injury crashes was 38. Only one crash hurt six people, as illustrated in Figure 5.6.

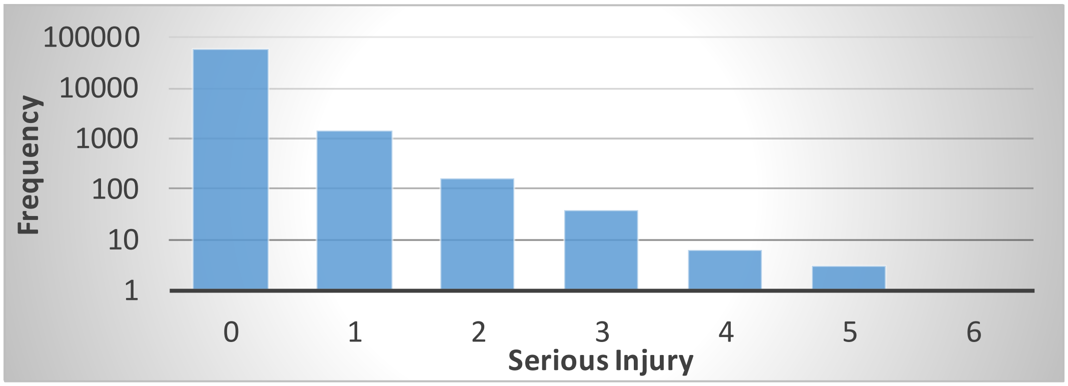


Figure 5.6 Frequency of different serious injury counts in Alaska (2010-2014).

Table 5.4 showed the percentages of serious injury crashes with different injured people counts. I interestingly, the percentages reduced sharply when the serious injury count increased from 0 to 1 and from 1 to 2. More specifically, the percentage of one-person serious injury crashes ranged from 2.05% to 2.87%, with an average of 2.32%, while the percentage of two-person serious injury crashes ranged from 0.21% to 0.33%, with an average of 0.27%.

Table 5.4 Yearly variation of serious injury crashes in Alaska (2010-2014)

# of Serious Injury	**Year** **2010**	**2011**	**2012**	**2013**	**2014**	**Avg.**
0	96.77%	97.25%	97.53%	97.62%	97.49%	97.32%
1	2.87%	2.36%	2.05%	2.11%	2.21%	2.32%
2	0.26%	0.33%	0.32%	0.21%	0.23%	0.27%

# of Serious Injury	Year					Avg.
3	0.08%	0.04%	0.08%	0.06%	0.06%	0.06%
4	0.02%	0.02%	0.01%	0.00%	0.01%	0.01%
5	0.01%	0.00%	0.01%	0.01%	0.00%	0.00%
6	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%

5.1.6. Fatality Count

The number of fatality counts in the whole crash reports ranged from zero to three. As illustrated in Figure 5.7, the number of crashes with one fatality was 264; the number of crashes with two fatalities was 23; and the number of crashes with three fatalities was two. As indicated in Table 5.5, the variation of crash frequency was quite limited in the past five years. The number of one-fatality crashes ranged from 48 to 59, and the number of two-fatality crashes ranged from 3 to 7. The three-fatality crashes occurred in 2011 and 2014 only.

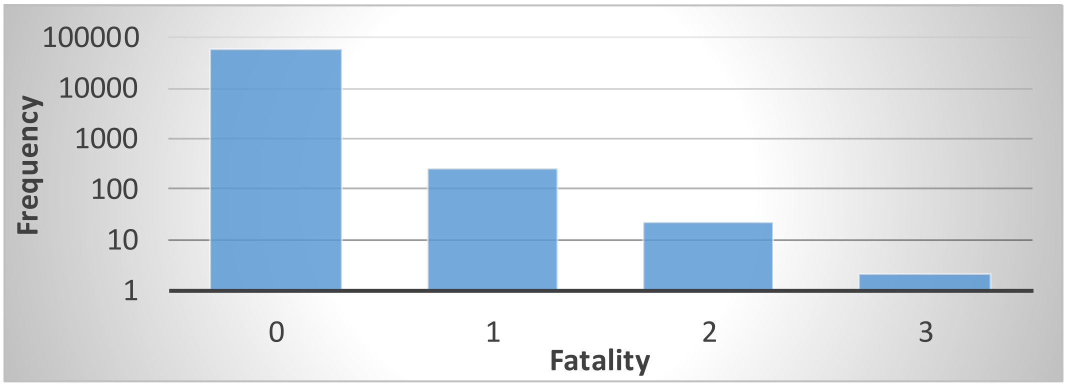


Figure 5.7 Frequency of different fatality counts in Alaska (2010-2014).

Table 5.5 Yearly variation of fatality crashes in Alaska (2010-2014)

# of Fatalities	Year					Avg.
	2010	2011	2012	2013	2014	
0	99.58%	99.51%	99.51%	99.57%	99.42%	99.52%
1	0.39%	0.43%	0.44%	0.41%	0.54%	0.44%
2	0.03%	0.06%	0.05%	0.02%	0.03%	0.04%
3	0.00%	0.01%	0.00%	0.00%	0.01%	0.00%

5.1.7. Weather Condition

The weather condition is also recorded in each crash report, while some missing records are termed as unknown. As shown in Figure 5.8, typical weather conditions are clear; cloudy; snow; rain; blowing sand, soil, dirt, or snow; freezing rain; fog; and severe crosswinds, ordered in their crash frequency. It should be noted that the crash frequency in different weather conditions does not indicate the impacts of different weather conditions on the crash.

As shown in Table 5.6, the percentage of crashes in clear weather ranged from 33.79% to 47.54%; the percentage of crashes in cloudy weather ranged from 18.43% to 36.26%; the percentage of crashes in snowy weather ranked the third, which ranged from 8.30% to 13.73%. The percentage of crashes in rainy weather was around a half of the snow crashes. In the first three years, the missing data took around 1.5% of all the crashes; however, the percentage of missing data increased up to 11.99% in 2013 and 4.68% in 2014. Considering that the percentage of crashes in clear weather reduced sharply from 40% to 33.79%, it is possible that most missing records come from clear data.

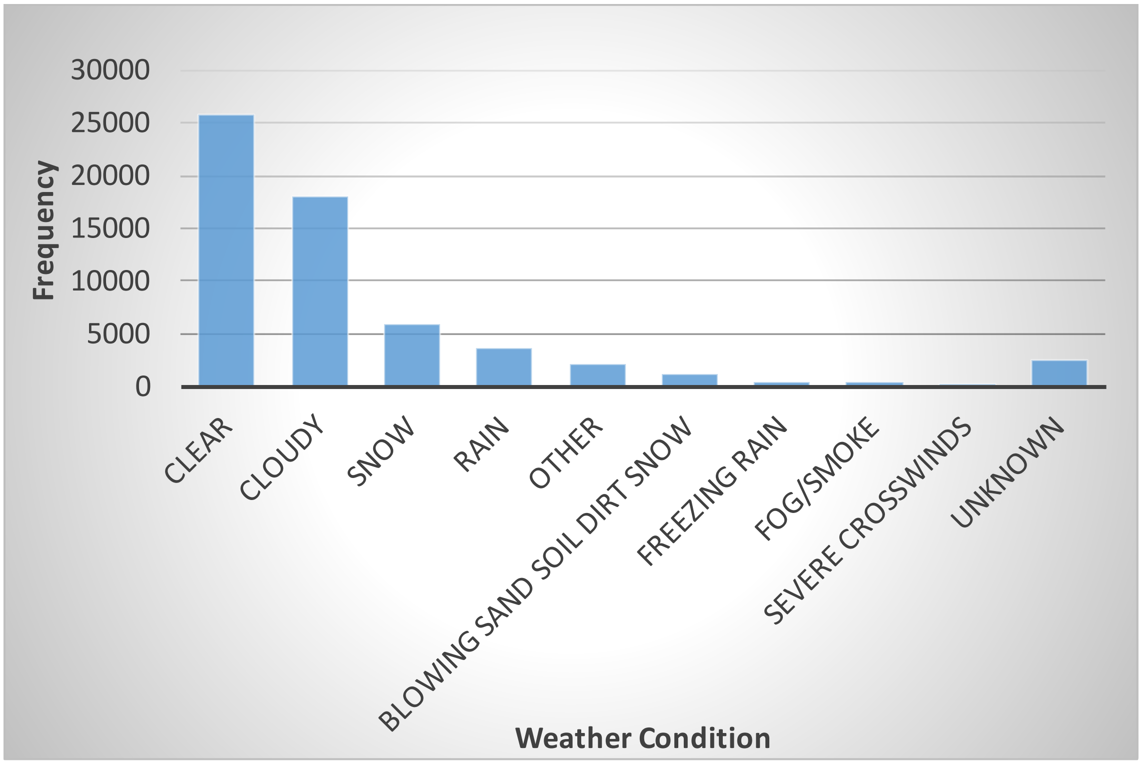


Figure 5.8 Frequency of crashes in different weather in Alaska (2010-2014).

Table 5.6 Yearly variation of crashes in different weather in Alaska (2010-2014)

Weather Condition	Year 2010	2011	2012	2013	2014	Avg.
Clear	43.75%	43.30%	46.72%	33.79%	47.54%	42.87%
Cloudy	36.08%	36.26%	32.68%	18.43%	25.80%	29.93%
Snow	8.51%	8.92%	9.35%	13.73%	8.30%	9.80%
Rain	5.41%	5.41%	5.34%	6.54%	7.21%	5.96%
Other	1.37%	1.19%	0.87%	11.57%	2.09%	3.47%
Blowing Sand Soil Dirt Snow	1.93%	2.21%	2.36%	1.79%	1.97%	2.05%
Freezing Rain	0.65%	0.56%	0.85%	0.70%	1.08%	0.76%
Fog/Smoke	0.62%	0.33%	0.36%	1.28%	1.17%	0.75%
Severe Crosswinds	0.14%	0.25%	0.15%	0.19%	0.15%	0.17%
Unknown	1.54%	1.57%	1.32%	11.99%	4.68%	4.23%

5.1.8. *Surface Condition*

Surface conditions for each record include dry, ice, snow, wet, water, sand, slush, and other surface conditions. Except for dry surface, ice surface condition takes the highest percentage of crashes, as shown in Figure 5.9. Snow condition follows closely behind ice condition, and is then followed by wet, water, sand, slush conditions. 1.55% crash records missed the surface condition information.

Table 5.7 presented the percentages of crashes in different surface conditions in the past five years. The problem in weather condition records occurred again in surface conditions. The missing data for 2013 increased sharply from 0.37% to 6.47%, while the crash records in dry condition are quite lower than the other years.

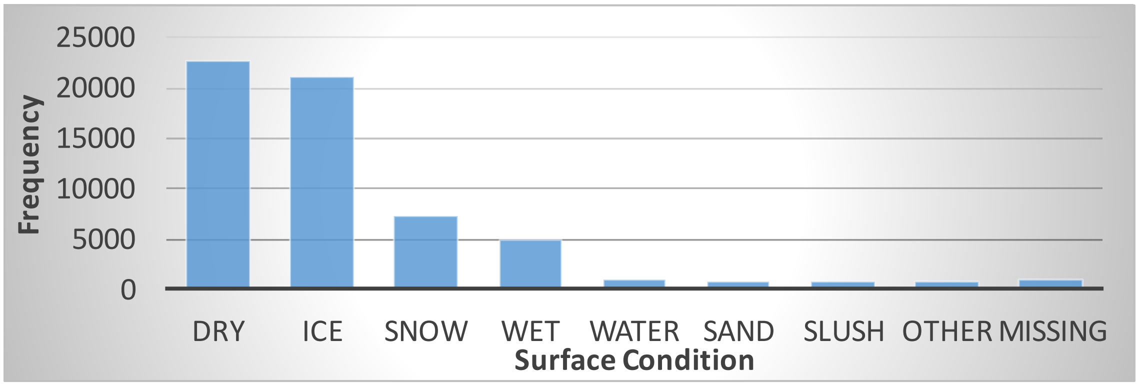


Figure 5.9 Frequency of crashes on different surface in Alaska (2010-2014).

Table 5.7 Yearly variation of crashes on different surface in Alaska (2010-2014)

Surface Condition	**Year** **2010**	**2011**	**2012**	**2013**	**2014**	**Avg.**
Dry	38.54%	36.82%	34.67%	33.75%	46.00%	37.79%
Ice	37.27%	36.78%	38.88%	33.67%	29.29%	35.29%
Snow	9.59%	12.93%	13.64%	14.10%	10.22%	12.13%
Wet	7.73%	6.80%	7.04%	8.98%	9.80%	8.03%
Water	2.19%	1.81%	2.02%	0.54%	0.72%	1.47%
Sand	1.77%	1.64%	0.91%	0.86%	1.35%	1.31%
Slush	1.42%	1.57%	1.07%	1.00%	0.81%	1.18%
Other	1.23%	1.28%	1.40%	0.61%	1.83%	1.25%
Missing	0.26%	0.38%	0.37%	6.49%	0.00%	1.55%

5.1.9. *Light Condition*

Six different types of light conditions are defined in the crash records. They are: daylight, twilight, dark with lighted roadway, dark with unknown lighting, dark with no lighting, and some other lighting conditions. On average, 1.47% of the crash records missed the light conditions. As shown in Figure 5.10, 37803 crashes occurred in the daylight condition, followed by 12698 crashes in a dark and lighted roadway, 5771 crashes in the dark and roadway not lighted category, 2224 crashes in twilight condition, and 534 crashes in the dark with unknown lightings.

Table 5.8 illustrated the light conditions in the past five years as well as the crash proportions. It was found that the proportions of crashes in the dark with no lighted roadway and twilight increased in the latter two years. Furthermore, the proportions of crash in dark with lighted roadway reduced from 22.50% to 19.78% and 20.32%.

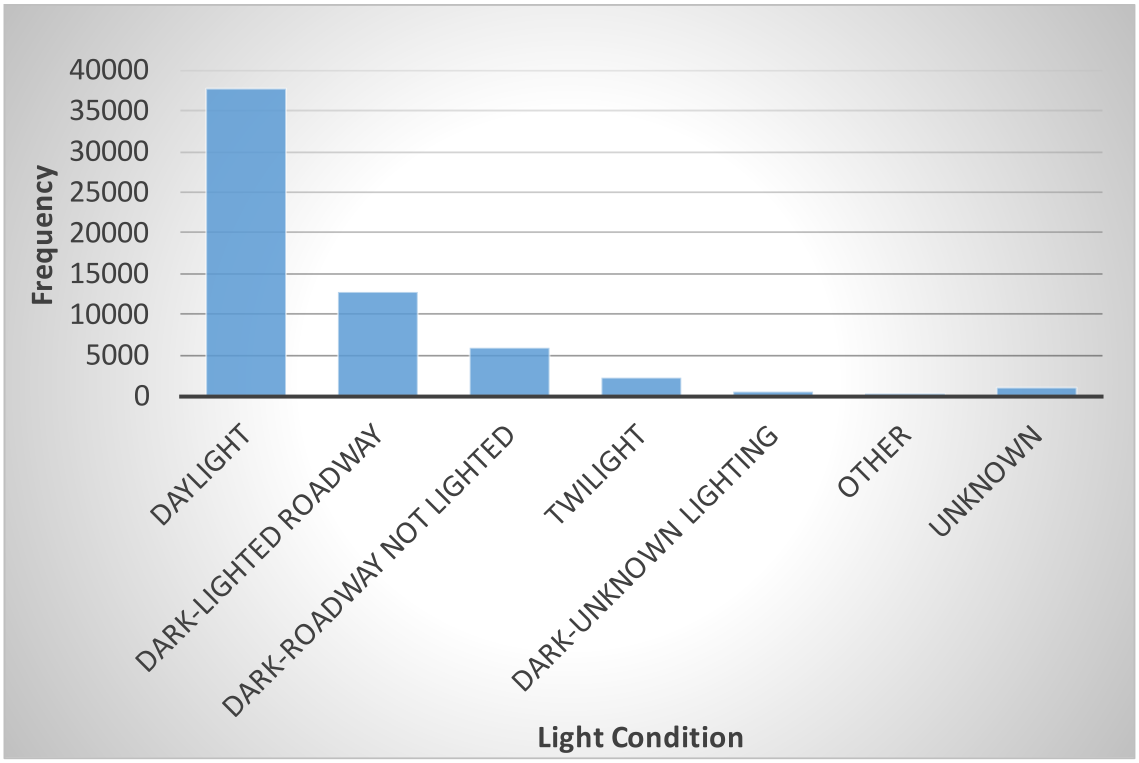


Figure 5.10 Frequency of crashes in different light conditions in Alaska (2010-2014).

Table 5.8 Yearly variation of crashes in different light conditions in Alaska (2010-2014)

Surface Condition	**Year** **2010**	**2011**	**2012**	**2013**	**2014**	**Avg.**
Daylight	63.67%	63.54%	64.08%	61.72%	61.47%	62.92%
Dark Lighted Roadway	21.16%	21.88%	22.50%	19.78%	20.32%	21.13%
Dark Roadway not Lighted	9.78%	9.53%	8.03%	10.24%	10.47%	9.61%
Twilight	3.02%	2.81%	3.00%	5.09%	4.67%	3.70%
Dark Unknown Lighting	0.86%	0.73%	0.87%	1.07%	0.92%	0.89%
Other	0.27%	0.20%	0.25%	0.32%	0.38%	0.28%
Missing	1.23%	1.31%	1.26%	1.78%	1.78%	1.47%

5.1.10. Alcohol Existing

The crash database records the alcohol condition of the crash driver to see if the crash driver was impaired by alcohol during the crash. As shown in Figure 5.11, 6% of crashes occurred directly due to the influence of alcohol.

Also illustrated in Table 5.9, the proportion of alcohol crashes were almost the same during the past five years. The percentages ranged from 5.60% to 6.52%, with an average of 5.91%.

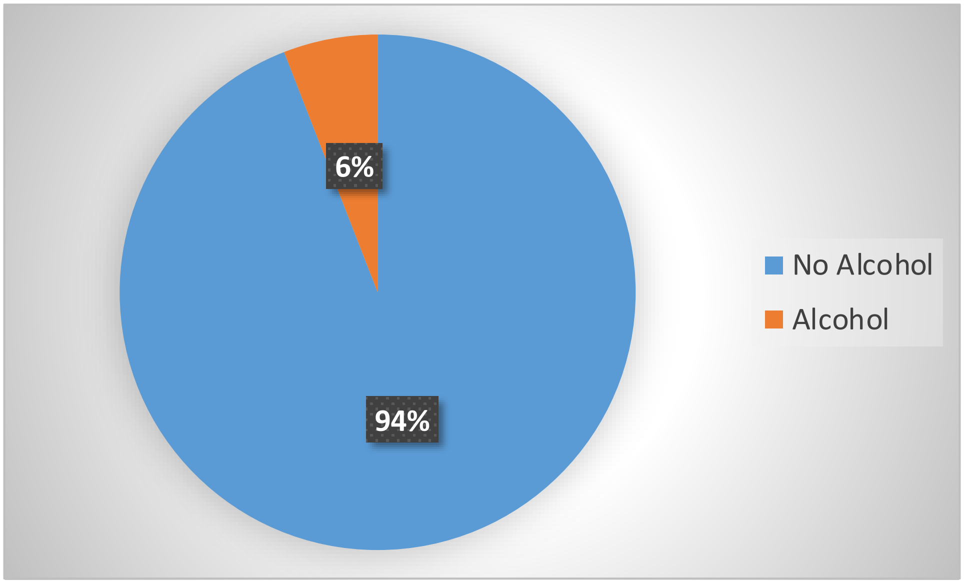


Figure 5.11 Frequency of crashes with different alcohol test results in Alaska (2010-2014).

Table 5.9 Yearly variation of crashes with different alcohol test results in Alaska (2010-2014)

Alcohol Condition	Year					Avg.
	2010	2011	2012	2013	2014	
No	93.48%	94.40%	94.25%	94.31%	94.03%	94.09%
Yes	6.52%	5.60%	5.75%	5.69%	5.97%	5.91%

5.1.11. Drivers' Age

Another interesting characteristic for crashes recorded in the database was the age for crash drivers. As shown in Figure 5.12, drivers of the recorded crashes grouped into five classes: below 16-years-old, from 16 to 35 years-old, between 36-50 years-old, between 50-80 years-old, and above 80 years-old. We find that crashes were most likely to occur in the young-mid age group, i.e., 16-35 year-old.

Figure 5.13 plotted the pie charts for different age ranges in different years. It showed that the percentages were quite consistent. Specifically, the proportion of crashes caused by young drivers (0-16) ranged from 7.46% to 9.41%. The proportion of crashes caused by young-mid drivers (16-35) ranged from 47.11% to 49.03%. The proportion of crashes caused by mid-aged drivers (35-50) ranged from 19.92% to 21.13%. The proportion of crashes caused by old drivers (50-80) ranged from 20.45% to 24.07%, and the remaining crashes were caused by drivers older than 80.

5.1.12. Gender Information

The last characteristic is the gender of crash drivers. It should be noted that more than 27.41% of gender information is missing. As indicated in Figure 5.14, crashes involved male drivers at a rate 50% higher than female drivers. In the recorded 5 years, as shown in Figure 5.15, the proportion of male drivers ranged from 41% to 45%, while the proportion of female drivers ranged from 24% to 31%.

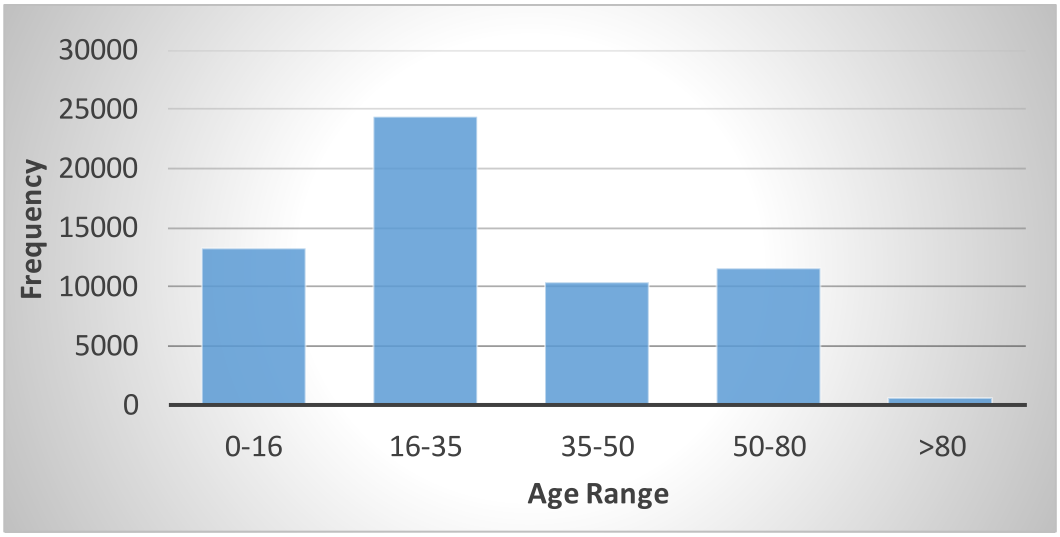
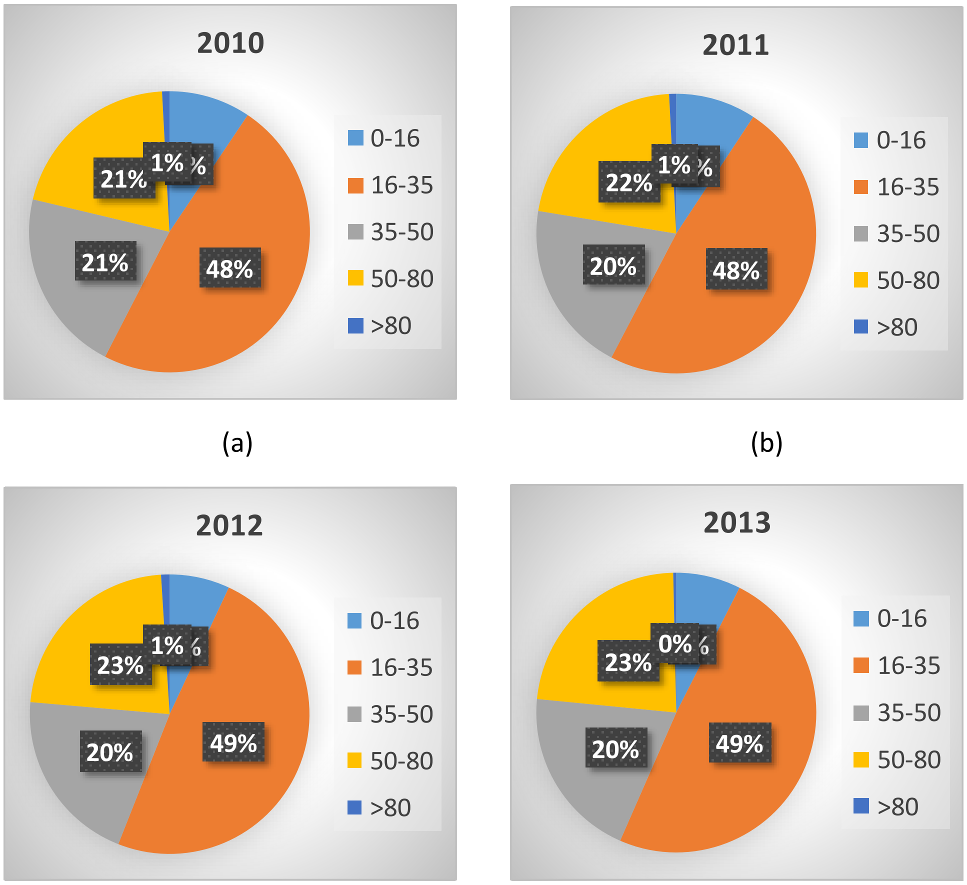
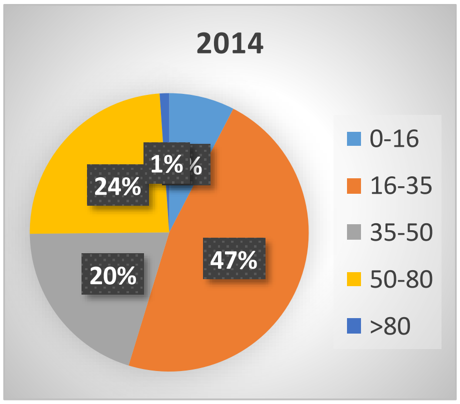


Figure 5.12 Frequency of crashes with different drivers' age range in Alaska (2010-2014).



(a) (b)

(c) (d)



(e)

Figure 5.13 Frequency of crashes with different drivers' age range in Alaska in each year.

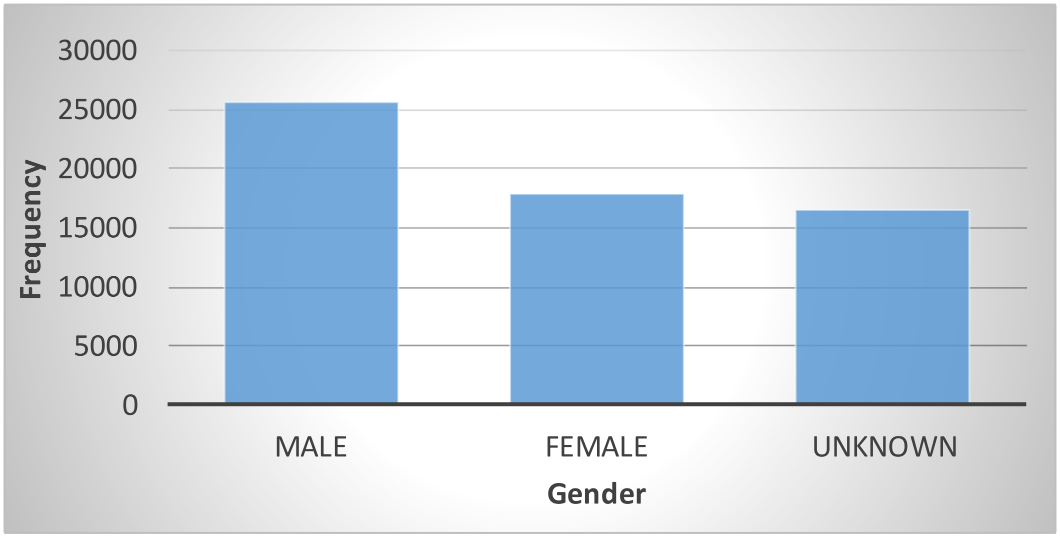
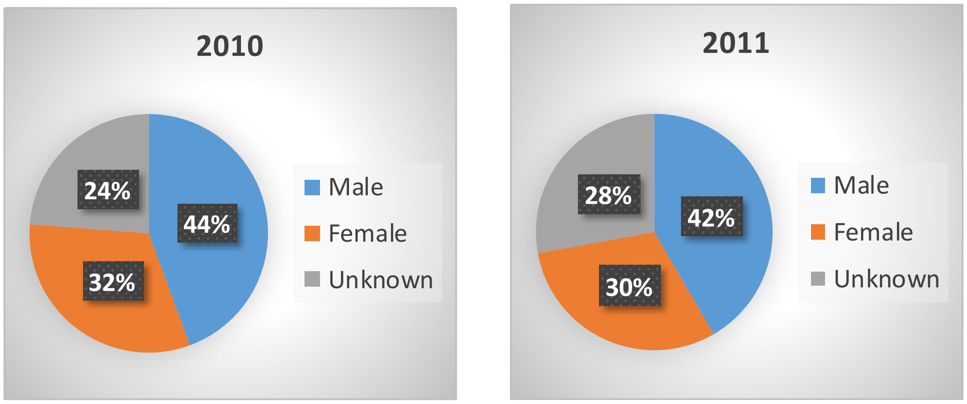
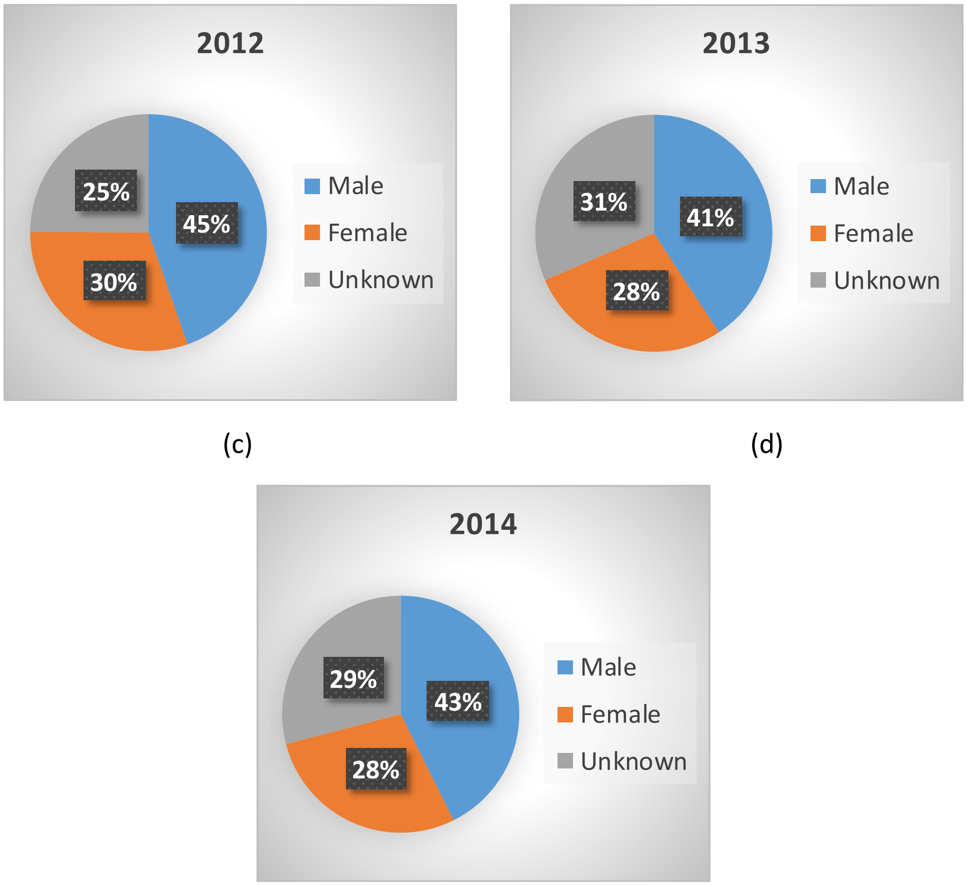


Figure 5.14 Frequency of crashes with different drivers' gender in Alaska (2010-2014).



(a) (b)



(c) (d)

(e)

Figure 5.15 Frequency of crashes with different drivers' gender in Alaska in each year.

5.2. Interactive Visualization

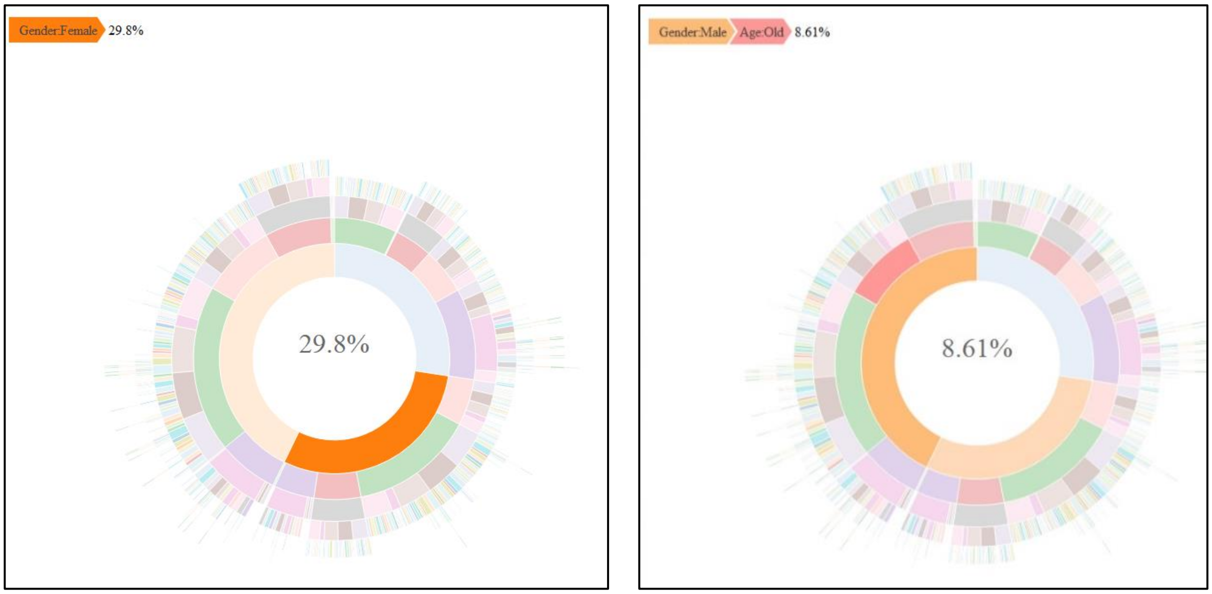
With the help of the RCVTS, five-year crash reports from Alaska State were queried and loaded onto the analysis process. By selecting the corresponding characteristics, the system generated interactive sunburst chart shown in Figure 5.16.

As introduced in Section 4.2.2, the sunburst chart is beneficial for visualizing the hierarchical relationship and attribute value via a ring-structure. The following attributes were picked as samples to demonstrate the usage of the chart. Table 5.10 lists selected samples and Figure 5.17 presents them. All the queries can be obtained by simply selecting the corresponding groups.

Table 5.10 Selected query samples

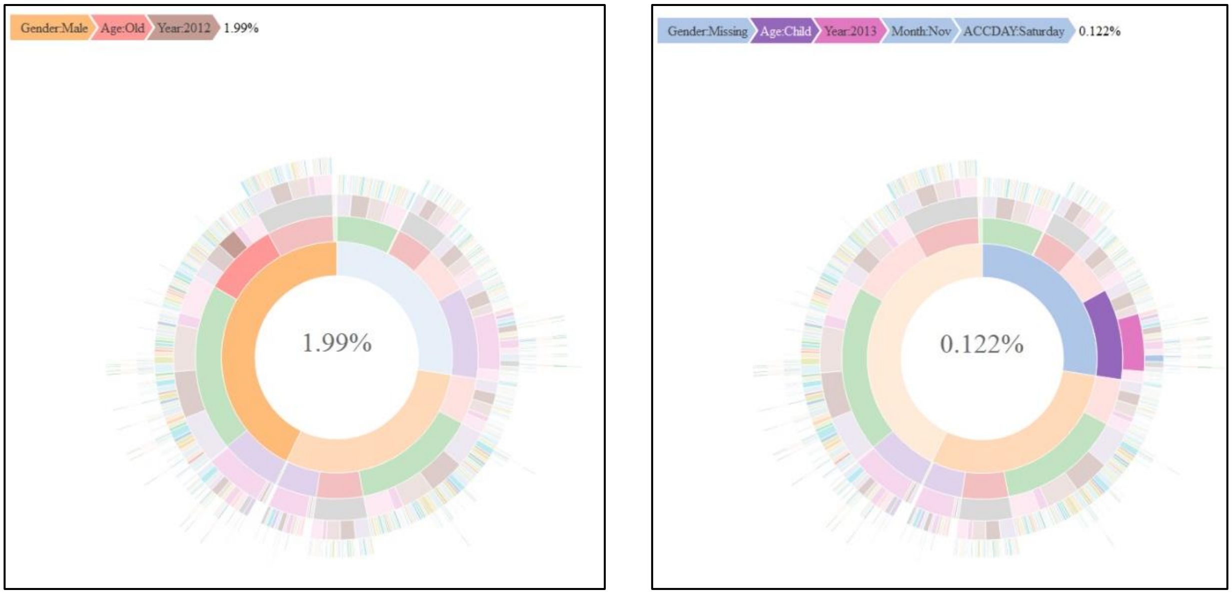
ID	Attributes	Value
I	Female	29.8%
II	Male-Old	8.61%
III	Male-Old-2012	1.99%
IV	Missing Gender-Child-2013-Nov-Saturday	0.12%
V	Male-Young-2010-Dec-Monday-PDO	<0.01%
VI	Male-Young-2012-Jan-Thursday-PDO-No Alcohol	0.12%

Figure 5.16 Sunburst Chart for Interactive Crash Data Analysis in Alaska (2010-2014).



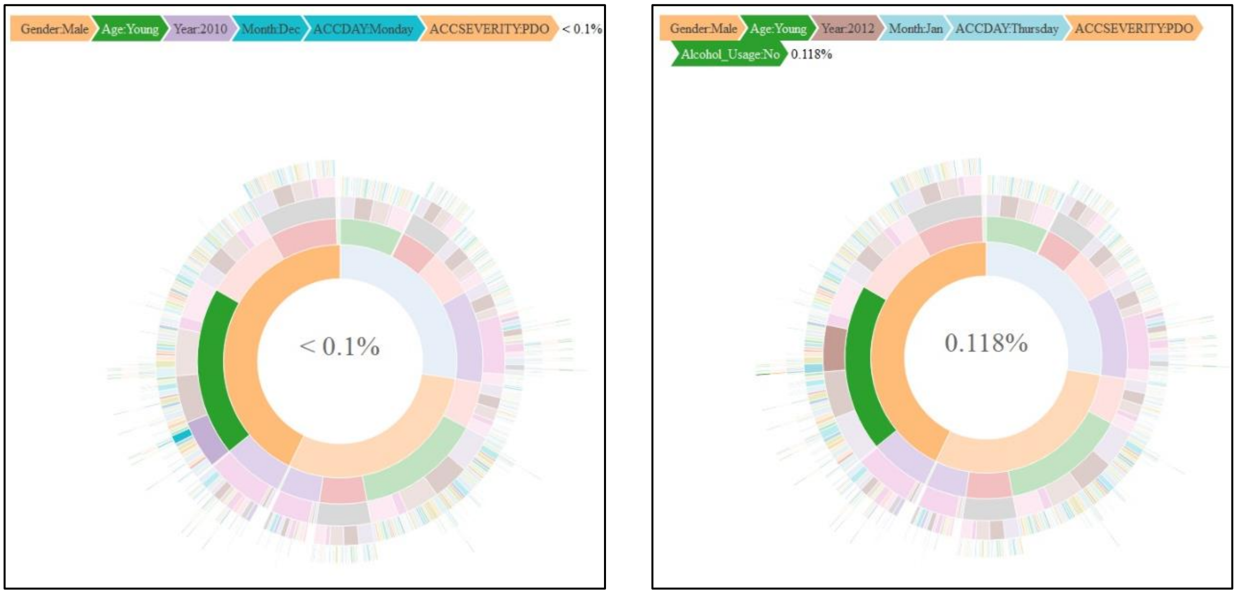
(a) Sample I

(b) Sample II



(c) Sample III

(d) Sample IV



(e) Sample V

(f) Sample VI

Figure 5.17 Query results using sunburst chart.

5.3. Summary

We illustrated a sample case of crash data analysis in this chapter using crash records collected from the State of Alaska. Crash characteristics—including the number of involved units, crash severity, number of total injury, number of minor injury, number of serious injury, number of fatality, weather condition, surface condition, light condition, alcohol condition, drivers' age and, gender information—were analyzed using various static charts and the interactive sunburst chart. The results showed that the proposed RCVTS provides a user-friendly and easy-to-use interface to analyze safety implications of the characteristics of RITI drivers, roads, and environments.

CHAPTER 6. CONCLUSIONS AND RECOMMENDATIONS

6.1. Conclusions

The tendency for smarter transportation management and crash monitoring has increased significantly in recent years. Although the quantity of crash data available to transportation safety engineers and decision makers is rapidly increasing, there is still a significant need for crash management, analysis, and visualization tools to make those data accessible for premium analysis and decision-making support. The challenges associated with managing and integrating numerous large and heterogeneous crash reports are such that the standards in crash reports vary from state to state and from time to time. Consequently, we have a significant need to develop an interactive baseline crash data platform to visualize and analyze rural crash characteristics in RITI communities.

The project developed RCVTS, a powerful, web-based crash visualization platform, to deal with visualization issues associated with rural crash records, which enables effective traffic safety analysis. This version of the platform offers the ability to handle visualization tasks and support data sharing services. RCVTS retains the capability of further improvement as concluded in Section 4.3.

The applications using the RCVTS presented in this report illustrate the best functionalities. Specifically, researchers analyzed the crash data from state of Alaska, including statistic descriptions for multiple crash characteristics and interactive visualization charts. In summary, the RCVTS provides a user-friendly interface on safety implications of characteristics of RITI drivers, roads, and environments.

6.2. Recommendations

To facilitate future research, the researchers make the following recommendations:

(1) Because of the scope and magnitude of new and emerging data sources, distributed computing and data management software should be considered. For example, Apache Spark is a unified analytics engine for large-scale data processing and distributed computing. It achieves high performance for both batch and streaming data, using a query optimizer and a physical execution engine. The application of such technology will likely improve system performance significantly.

(2) Although RCVTS has addressed many of the challenges associated with crash analysis and visualization, one key challenge remains. Specifically, there is a need for enhanced capacity to handle the linear and nonlinear relationships among multiple characteristics. Future work would address this issue by creating an advanced data process structure for safety performance modelling.

REFERENCES

Barry, M. and Card, B., 2009, *Boston's Massachusetts Bay Transit Authority (MBTA)*. [Online]. Available: http://mbtaviz.github.io/

Bertini, R., Hansen, S., Byrd, A., and Yin, T., 2005, Experience Implementing a User Service for Archived Intelligent Transportation System Data. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1917, Transportation Research Board of the National Academies, Washington, D.C., 90–99.

Byron, L., Wattenberg, M., 2008, Stacked graphs-Geometry & aesthetics, *IEEE Trans. Vis. Comput. Graphics*, 14(6): 1245–1252.

Chen, C., 2003, Freeway Performance Measurement System (PeMS). Publication *UCB-ITS-PRR-2003-22*. California Partners for Advanced Transit and Highways (PATH), University of California, Berkeley, Richmond, CA.

Chen, C., Wang, Y., Li, L., Hu, J., Zhang, Z., 2012, The retrieval of intra-day trend and its influence on traffic prediction, *Transportation Research Part C: Emerging Technologies*, 22: 103-118

Chen, W., Guo, F., Wang, F., 2015, A Survey of Traffic Data Visualization, *IEEE Transactions on Intelligent Transportation Systems*, 16(6): 2970-2984.

Crnovrsanin, T., Muelder, C., Correa, C., and Ma, K., 2009, Proximity-based visualization of movement trace data, in *Proc. IEEE Symp. Visual Anal. Sci. Technol.*, 11–18.

Hale, D.,, Hajbabaie, A., Ma, J., Hu, J., Park, H., Bared, J., 2016, Proposed Data-Driven Performance Measures for Comparing and Ranking Traffic Bottlenecks, *Transportation Research Procedia*, 15: 483-494.

Department for Communities and Local Government, 2008. Planning application a faster and more responsive system: final report, *Killian Pretty Review*. London: Communities and Local Government.

Ersoy, O., Hurter, C., Paulovich, F., Cantareiro, G., Telea, A., 2011, Keleton-based edge bundling for graph visualization, *IEEE Trans. Vis. Comput. Graphics*, 17(12): 2364–2373.

Fekete, J., Plaisant, C., 1999, Excentric labeling: Dynamic neighborhood labeling for data visualization, in *Proc. ACM SIGCHI*, 512–519.

Ferreira, N., Poco, J., Vo, H., Freire, J., and Silva, C., 2013, Visual exploration of big spatio-temporal urban data: A study of New York City taxi trips, *IEEE Trans. Vis. Comput. Graphics*, 19(12): 2149–2158.

Havre, S., Hetzler, B., Nowell, L., 2000, ThemeRiver: Visualizing theme changes over time, in *Proc. IEEE Symp. Inf. Vis.*, 115–123.

Hawaii State Department of Transportation, 2014, *Highway Safety Annual Report*, http://hidot.hawaii.gov/highways/files/2013/01/2014-Annual-Report.pdf Accessed in May 2017.

Kwon, J., Varaiya, P., 2008, Effectiveness of California's High Occupancy Vehicle (HOV) system, *Transportation Research Part C: Emerging Technologies*, 16(1): 98-115

Kraak, M., 2003, The space-time cube revisited from a geovisualization perspective, in *Proc. 21st Int. Cartogr. Conf.*, 1988–1996.

Kloeckl, K., Chen, X., Sommer C., Ratti C., and Biderman A. 2016, Trains of data. [Online]. Available: http://senseable.mit.edu/trainsofdata/, accessed on April 5, 2016.

Lampe, O., Hauser, H., 2011, Interactive visualization of streaming data with kernel density estimation, in *Proc. IEEE Pac. Vis. Symp.*, 171–178.

Ma, X., McCormack, E., Wang, Y., 2011, Processing commercial global positioning system data to develop a web-based truck performance measures program. *Transportation Research Record: Journal of the Transportation Research Board*, 2246: 92-100.

Ma, X., Tao, Z., Wang, Y., Yu, H., Wang, Y., 2015, Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transportation Research Part C: Emerging Technologies*, 54: 187-197.

Ma, X., Wang, Y., 2014, Development of a data-driven platform for transit performance measures using smart card and GPS data. *Journal of Transportation Engineering*, 140(12): 04014063.

Ma, X., Wu, Y., Wang, Y., Chen, F., Liu, J., 2013, Mining smart card data for transit riders' travel patterns. *Transportation Research Part C: Emerging Technologies*, 36: 1-12.

Jeihani, M., James, P., Saka, A., Ardeshiri, A., 2015, Traffic recovery time estimation under different flow regimes in traffic simulation, *Journal of Traffic and Transportation Engineering (English Edition)*, 2(5): 291-300.

Hossan, M., Asgari, H., Jin, X., 2016, Investigating preference heterogeneity in Value of Time (VOT) and Value of Reliability (VOR) estimation for managed lanes, *Transportation Research Part A: Policy and Practice*, 94: 638-649.

Ogawa M., Ma, K., 2010, Software evolution storylines, in *Proc. 5th Int. Symp. Softw. Vis.*, 35–42

Pu, J., Liu S., Ding Y., Qu H., and Ni L., 2013, T-Watcher: A new visual analytic system for effective traffic surveillance, in *Proc. IEEE 14th International Conference on MDM*, 1: 127–136.

Rivadeneira, A., Gruen, D., Muller, M., Millen, D., 2007, Getting our head in the clouds: Toward evaluation studies of tagclouds, in *Proc. ACM SIGCHI*, 995–998.

Shashi, S., Tien Lu, C., Liu, R., Zhou. C., 2002, CubeView: a system for traffic data visualization. In *Proceedings IEEE 5th International Conference on Intelligent Transportation Systems*, 674-678.

Shneiderman, B. 1996, The eyes have it: a task by data type taxonomy for information visualization. In *Proceedings IEEE Symposium on Visual Languages,* 336-343.

Smith, A., Hawes, T., and Myers, M., 2014, Hiearchie: Visualization for hierarchical topic models. In *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces*, 71–78.

Traffic Safety Facts – 2013 Data, 2015, National Highway Traffic Safety Administration, DOT HS 812 181, Washington, DC.

Tufte, K. A. Portal 2.0: Toward a Next-Generation Archived Data User Service. Presented at 89th *Annual Meeting of the Transportation Research Board*, Washington, D.C., 2010.

United Sates Department of Transportation (USDOT) Federal Motor Carrier Safety Administration (FMCSA). *New Hours-of-Service Safety Regulations to Reduce Truck Driver Fatigue Begin Today*, https://www.fmcsa.dot.gov/newsroom/new-hours-service-safety-regulations-reduce-truck-driver-fatigue-begin-today. Accessed in May 2017

Wikipedia, 2018. *Executive summary.* https://en.wikipedia.org/wiki/Executive_summary. Accessed January 4.

Wongsuphasawat, K., Pack, M., Filippova, D., VanDaniker, M., Olea, A., 2009, Visual Analytics for transportation incident data sets. *Journal of Transportation Research Board: Transportation research record*, 2138: 135–145

Wu, X., Liu, H., 2014, Using high-resolution event-based data for traffic modeling and control: An overview, *Transportation Research Part C: Emerging Technologies*, 42: 28-43

Jin, X., Hossan, S., Asgari, H., Shams, K., 2018, Incorporating attitudinal aspects in roadway pricing analysis, *Transport Policy*, 62: 38-46, doi.org/10.1016/j.tranpol.2017.03.020

Xie Z., Yan, J., 2008, Kernel density estimation of traffic accidents in a network space, *Comput., Environ. Urban Syst.*, 32(5): 396–406.

Xuan Y., Adib K., 2014, Evaluation of the effectiveness of accident information on freeway changeable message signs: A comparison of empirical methodologies, *Transportation Research Part C: Emerging Technologies*, 48: 158-171.

Wei, Y., Deng, P., Taleb, T., Wan, J., Bi, C., 2016, An unlicensed taxi identification model based on big data analysis. *IEEE Transactions on Intelligent Transportation Systems* 17(6): 1703-1713.

Zeng, W., Fu, C., Arisona, S., Qu, H., 2013, Visualizing interchange patterns in massive movement data, *Comput. Graph. Forum*, 32(3): 271–280.

Zhang, Y., Ali H., 2015, A gradient boosting method to improve travel time prediction, *Transportation Research Part C: Emerging Technologies*, 58(B): 308-324.

Zheng, J., Ma, X., Wu, Y., Wang, Y., 2013, Measuring signalized intersection performance in real-time with traffic sensors. *Journal of Intelligent Transportation Systems*, 17(4): 304-316.

Zhou, H., Xu, P., Yuan, X., Qu, H., 2013, Edge bundling in information visualization, *Tsinghua Sci. Technol.*,18(2): 145–156.

APPENDIX A

Table A.1 Definition of Non-confidential Variables for State of Alaska

Variables	Definitions
ACCNUM/SR NUMBER	State record number
DATASOURCE	Police or Citizen report
FORM TYPE	12-209 (CITIZEN) OR 12-200 (LAW ENFORCEMENT)
POLICEDEPT/REPORTING AGENCY	Name of law enforcement agency providing report
PCASENUM/CASE NUMBER	Identification number given to crash by reporting law enforcement agency
CDSRTE/CDS ROUTE	State DOTPF numerical identification given to road
ACCMIPT/MILEPOINT	Milepoint, in decimal miles, of crash. Not same as milepost.
ROADNAME	Name of the road (PRE 2012)
ACCDATE / CRASH DATE&TIME	Date of Crash
Year/YEAR	Year of Crash
Month/MONTH	Two digit numerical month of crash
Day/DAY	Numerical day of crash
ACCDAY/ACCDAY	Text of day of crash (Monday, Tuesday, Wednesday, …)
ACCTIME/ ACCTIME	Four digit military time
ACCHOURS/ACCHOURS	One hour range, in military time, (00:00-00:59, 01:00-01:59,…)
STREET/STREET HIGHWAY	Name of street or road
CROSSSTREET/INTERSECTING STREET	Name of cross street or road crash is oriented to
INTERDIST/DISTANCE	Distance from cross street
REFUNITS/MEASUREMENT DESC	Units of distance (e.g. feet, miles) or at intersection
INTERDIR/DIRECTION DESC	Cardinal direction from reference cross street
RDJUNCT/JUNCTION DESC	Intersection type
NUMVEH/TOTAL MOTORIZED UNITS	Number of vehicles involved in crash
ACCSEVERITY/INJURY STATUS DESC	Determines level of severity, from property damage only through injury, to fatal
TOTINJ/INJURY COUNTS	Total serious and minor injuries in crash
MAJINJ/MAJOR INJURY COUNTS	Total serious injuries in crash
MININJ/MINOR INJURY COUNT	Total minor injuries in crash
TOTFATAL/FATAL COUNTS	Total fatalities in crash
EVETYPE/FIRST HARMFUL EVENT	First causal event of crash
MANNER OF COLLISION	For 'motor vehicle in transit', the configuration of the collision (POST 2013)
EVELOC/REL TO TRAFFICWAY	Indicates where crash occurred (e.g. roadway, shoulder, median…)
WEATHER/WEATHER DESC	Describes weather conditions at time of crash
RDCHARACTER	Indicates whether road is straight or level (PRE 2012)
SURFACECOND/ROAD SURFACE	Describes the road surface conditions at time of crash
LIGHT/LIGHTING	Describes source of lighting at time of crash

Variables	Definitions
ALC_DRUG	Describes whether officer responding to crash suspected alcohol or drugs (PRE 2012)
ALCOHOL SUSPECTED	Describes whether officer responding to crash suspected alcohol (POST 2013)
DRUGS SUSPECTED	Describes whether officer responding to crash suspected drugs (POST 2013)
TOTOCC/TOTAL PEOPLE IN VEHICLE	Provides count of all people
D1INJSTAT/ DRIVER INJURY STATUS	Injury status (no injury, minor, major, fatality) for driver
D1SAFEQ1/DRIVER RESTRAINT SYS 1	Describes restraint use for driver
D1SAFEQ2/DRIVER RESTRAINT SYS 2	Describes restraint use for driver if more than one system used
D1AGE/DRIVER AGE	The age of driver
D1AGERANGE	Age range of driver of vehicle 1 in five year increments
D1SEX/ DRIVER SEX DESC	The gender of vehicle 1 driver
D1ALCDRGSUSP	Whether or not law enforcement suspected alcohol or drugs in vehicle 1 driver (PRE 2012)
D1ALCDRGTEST	Whether or not law enforcement tested for alcohol or drugs in vehicle 1 driver (PRE 2012)
D1BAC	If law enforcement tested for alcohol with BAC test, the BAC result in vehicle 1 driver (PRE 2012)
DRIVER ALCOHOL SUSPECTED	Whether or not law enforcement suspected alcohol in vehicle 1 driver (POST 2013)
DRIVER ALCOHOL TEST TYPE	If law enforcement tested for alcohol in vehicle 1 driver, what test was used (POST 2013)
DRIVER ALCOHOL LEVEL	If law enforcement tested for alcohol, what was test result in vehicle 1 driver (POST 2013)
DRIVER DRUGS SUSPECTED	Whether or not law enforcement suspected drugs in vehicle 1 driver (POST 2013)
DRIVER DRUGS TEST TYPE	If law enforcement tested for drugs in vehicle 1 driver, what test was used (POST 2013)
RIVER DRUGS RESULT	If law enforcement tested for drugs what was test result in vehicle 1 driver (POST 2013)
ENVCIRC	Indicates whether there were environment contribution to the crash (e.g. glare, obstruction, weather)
RDCIRC/ROAD CONDITION	Indicates whether road conditions contributed to the crash
VEHCIR/VEHICLE CIRCUMSTANCE	Indicates whether vehicle conditions or problems contributed to the crash
HUMCIRC1/ DRIVER CONTRIB CIRCUMSTANCE 1	Indicates whether driving behaviors or other people problems contributed to the crash
HUMCIRC2/DRIVERCONTRIB CIRCUMSTANCE 2	If more than one behavior present, indicates whether driving behaviors or other people problems contributed to the crash
TRFCCONTDEV/TCD DESC	Description of any traffic controls at site of crash specific
TRVDIRECT/ DIRECTION OF TRAVEL	Description of vehicle direction of travel
ACTPREACT/ ACTION	Description of vehicle 1 action at time of crash
SECEVENT	For 2012 and earlier, first event described the crash summary

Variables	Definitions
	data. This field describes any second event that contributed to crash.
FIRST EVENT	This field describes any first event that contributed to crash.
SECOND EVENT	This field describes any second event that contributed to crash.
THIRD EVENT	This field describes any third event that contributed to crash.
FOURTH EVENT	This field describes any fourth event that contributed to crash.
MOST HARMFUL EVENT	Of events contributing to the crash, this field describes the most harmful event that contributed to crash.
TICKETCODE1/ DRIVER CHARGES 1	If citations issued to driver, describes first citation
TICKETCODE2/ DRIVER CHARGES 2	If citations issued to driver, describes second citation
NONCOMCONFIG/ BODY TYPE	Vehicle body type and configuration. Choices changed from 2012 to 2013
COMCONFIG1	If vehicle is a commercial vehicle, then this field describes configuration.
COMCONFIG2	If vehicle is a commercial vehicle, then this field describes configuration if there's more than one descriptor.
BODYTYPE	If vehicle is a commercial vehicle, then this field describes body type.
BODY TYPE	Vehicle body type, applies to commercial or non-commercial
COMM VEH BODY TYPE	Further description of vehicle 1 if a commercial vehicle
DAMAGETYPE/ EXTENT OF DAMAGE	Describes the extent of damage from no damage to disabling damage
REGION/REGION	Alaska Department of Transportation and Public Facilities region where crash occurred
BOROUGH/BOROUGH	Borough (unorganized or named) where crash occurred
CENSUSAREA/CENSUS AREA	Census area where crash occurred
ELECTIONDISTRICT/ELECTION DISTRICT	Alaska election district where crash occurred
CITY/CITY	City where crash occurred
POLICEDETACH/DETACHMENT	Alaska State Trooper detachment oversight of area where crash occurred
NHS/NHSSYS	Indicates if crash occurred on a National Highway System designated road
AHS/AHSSYS	Indicates if crash occurred on a Alaska Highway System designated road
FUNCTIONALCLASS/FUNCTIONAL CLASS	Describes functional class of road where crash occurred
RURALURBAN/URBAN RURAL	Describes whether crash area is designated urban or rural
OWNERSHIP/MAINTENANCE RESPONSIBILITY	Describes ownership of road
PAVEDUNPAVED/PAVED	Describes if point on road where crash occurred was paved or not
MAINTENANCESTATION/MAINTENANCE STATION	Indicates name of maintenance station responsible for portion of road where crash occurred
MAINTENANCECATEGORY/MAINTENAN	Provides maintenance category for reach of road where crash

Variables	Definitions
CE CATEGORY	occurred

Table A.2 Definition of Non-confidential Variables for State of Idaho

Variables	Definition
fldAccidentID	Crash ID
fldDataStamp	Data Stamp
fldAccidentDate	Accident Date
fldAccidentTime	Accident Time
fldAccidentYM	Year Month
fldAgencyCaseID	Report Agency ID
fldAgencyName	Report Agency Name
fldCityName	Crash City Name
fldCountyName	County Name
fldDirFromIntersection	Direction From Intersection
fldEmsArrivalTime	EMS Arrival Time
fldEmsDispatchTime	EMS Depature Time
fldEmsProviderName	EMS Provider Name
InCity	City/Rural
fldIntersectionDistance	Intersection Distance
fldIntersectionDistanceUnitsName	Intersection Distance Units
fldIntersectionID	Intersection ID
fldLaneDirectionName	Lane Direction
fldLaneOfImpact	Lane of Imapact
fldLightCondition	Light Condition
fldLocalityName	Local Name
fldMedicalCareProviderName	Medical Care Provider Name
fldMilepost	Milepost
fldMilepostAccuracy	Milepost Accuracy
fldOfficerID	Officer ID
fldPoliceArrivalTime	Police Arrival Time
fldPoliceDispatchTime	Police Dispatch Time

Variables	Definition
fldReportDistrict	Report District
fldRoadConditionOtherName	Road Condition
fldRoadFuncName	Rural/Urban (detailed)
fldRoadSurfaceConditionName	Road Surface Condition
fldRoadSurfaceTypeName	Road Surface type
fldSegmentCode	Segment ID
fldSerialNbr	Serial Number
fldSpeedLimitPrimary	Speed Limit Primary
fldSpeedLimitSecondary	Speed Limit Secondary
fldStreetIDPrimary	Street ID Primary
fldStreetIDSecondary	Street ID Secondary
fldStreetPrimaryName	Street Primary Name
fldStreetPrimaryPreDir	Street Primary Previous Direction
fldStreetPrimaryPostDir	Street Primary Post Direction
fldStreetSecondaryName	Street Secondary Name
fldStreetSecondaryPreDir	Street Secondary Previous Direction
fldStreetSecondaryPostDir	Street Secondary Post Direction
fldTrafficControlName	Traffic Control Name
fldTrafficControlStatusName	Traffic Control Status Name
fldWeatherCondition1Name	Weather Condition 1 Name
fldWeatherCondition2Name	Weather Condition 2 Name
fldAccidentYYYY	Accident Year
fldAccidentDayOfWeek	Accident Day of Week
fldRoadTypeName	Road type
fldRoadLaneName	Road Lane Name
fldCityDirection	Direction to City
fldCityDistance	Distance to City
fldLatitude	Latitude
fldLongitude	Longitude
fldExtra1Accd	Extra Accident 1 ID

Variables	Definition
fldExtra2Accd	Extra Accident 2 ID
IsReportable	Is Reportable
fldAccidentHourTime	Accident Hour Time
fldAccidentMonth	Accident Month
OfficialMarkDate	Official Mark Date
OfficialRecord	Official Record
fldGeometricsH	Road Curved
fldGeometricsV	Road grade
fldIntersectionType	Intersection Type
fldWorkZoneRelated	Is in Work Zone Area
fldWorkZoneCrashLocation	Work Zone Crash Location
fldWorkZoneType	Work Zone Type
fldWorkZoneWorkersPresent	Work Zone Workers Present Condition
fldWorkZoneLawPresent	Work Zone Law Present Condition
fldLanesBlocked	Is Lanes Blocked
fldLanesClearedDate	Lane Block Cleared Data
fldLanesClearedTime	Lane Block Cleared Time
fldIntersectionDistance2	Distance to Intersection 1
fldIntersectionDistanceUnits2	Distance to Intersection 2
fldDirFromIntersection2	Direction from Intersection 2
fldLocationType	Location Type
fldStateHWYSystem	State Highway
fldAccidentDateTime	Accident Date Time
fldFullSegmentCode	Full Segment Code
fldContribCircID	Contributing ID
fldContributingCircName	Contribution Circumstance
fldCircumstanceOrder	Contribution Circumstance Order
EventCodeName	Event Code Name
EventLocationName	Event Location
fldEventOrder	Event Order

Variables	Definition
fldUnitID1	Unit 1 ID
fldUnitID2	Unit 2 ID
fldUnitNbr1	Unit 1 Number
fldUnitNbr2	Unit 2 Number
fldEventCode	Event Code
fldPersonID	Person ID
fldAccidentID	Accident ID
fldAge	Age
fldAlcoholLevel	Alcohol Level
fldAlcoholTest	Alcohol Test
fldBirthdate	Birth Date
fldCitationName	Citation Name
fldPersonCountyName	Person Country Name
fldDrugTestName	Drug test
fldDrugTestResult	Drug test Result
fldDrugUsedname	Drug User Name
fldEjectionName	Ejection
fldExtractionName	Extraction
fldFatalityID	Fatality ID
fldInjuryName	Injury Name
fldIsOperator	Is Operator
fldLicenseState	License State
fldMiddleInitial	Middle Initial
fldProtectiveDeviceName	Protective Device
fldSeatingName	Seating position
fldSex	Sex
fldSubstanceUseName	Substance Use Name
fldAirBagDeployment	Airbag Deployment
fldAirBagDeploymentLocation	Airbag Deployment Location

Variables	Definition
fldDLCommercial	Is Driver License Commerical
fldFatalities	Fatalities
fldInjuries	Injuries
fldIntersectionRelated	Intersection related
fldPersons	Persons
SeverityName	Severity Name
fldUnits	Units
fldUpdateDate	Updated Date
fldUpdateUser	Updated User
fldApprovalDate	Approval Date
fldApprovalUser	Approval User
fldTransmitDate	Transmit Date
fldExtra1AccdStat	Extra Accident Status
fldEstDamage	Estimate Damage
fldEvent1RelToJunctionName	Event location
fldEventFirstHarmful	First Harmful Event
fldEventMostHarmful	Most harmful event
fldHitAndRun	Hit and Run
fldIsCommercial	Commercial
fldLiabilityInsurance	Liability Insurance
fldLicensePlateState	License Plate State
fldOperatorActionname	Operator Action name
fldTowed	Towed
fldTravelDirection	Travel direction
fldUnitNbr	Unit number
fldUnitTypeName	Unit type
fldImpairedDriverUnit	Impaired Driver
fldContributingCircumstance1	Contributing Circumstance 1
fldContributingCircumstance2	Contributing Circumstance 2

Variables	Definition
fldContributingCircumstance3	Contributing Circumstance 3
fldDistractedBy	Distracted by

Table A. 3 Definition of Non-confidential Variables for State of Washington

Variable	Definition
Collision Report Number	A number used to uniquely identify each Police Traffic Collision Report form, Civilian Vehicle Collision Report form, or electronic SECTOR report.
State Reportable Indicator Type Code	An indicator that denotes the damage of a vehicle or property involved in the collision that meets the states damage threshold of $750 or a person involved is injured or dies as a result of the collision.
Intentional Type Code	An Indicator that denotes the collision was caused by deliberate actions.
Legal Intervention Type Code	An indicator that denotes that the collision resulted from the action or inaction of a law enforcement officer or other official during the pursuit of a suspect that affected the Motor Vehicle Driver or Pedalcyclist and prevented them from operating their vehicle in a safe, controlled manner.
Medically Caused Type Code	An indicator that denotes whether the Collision resulted from medical reasons that affected the Motor Vehicle Driver or Pedalcyclist and prevented them from operating their vehicle in a safe, controlled manner.
County_See Appendix A	The largest administrative sub-division within a state in the United States.
City_See Appendix A	A primarily urban political unit having corporate status and usually powers of self-government established by state charter.
Collision Report Type Code	A classification of the Collision report based on the type of trafficway jurisdiction on which the Collision occurred.
Number of Fatalities	Total number of Persons killed per Collision.
Number of Injuries	Total number of Persons injured per Collision
Number of Pedal cyclists Involved	The total number of Pedalcyclists involved in a Collision
Number of Pedestrians Involved	The total number of Pedestrians involved in a Collision.
Number of Motor Vehicles Involved	Total Number of Motor Vehicles involved in a Collision.
City Primary Trafficway	The roadway that the law enforcement officer or citizen considers to be the principal site of the collision. If the collision occurred where two Roadways cross then the intersecting trafficway will contain a value.
City Block Number	The building site identifier or the range of building site identifiers, in which the collision occurred, as assigned to the city street by the county

Variable	Definition
City Intersecting Trafficway	The roadway that crosses the primary trafficway. An intersecting trafficway is captured when the collision occurs at the point where the primary trafficway crosses the Intersecting trafficway or when these two trafficways are associated due to the presence of channelization.
City Distance From Reference Point	The distance from the collision scene to the reference point.
City Reference Point Miles_Feet Type Code	An indicator that denotes if the distance from a reference point for a city Ssreet collision is in miles or feet.
City Compass Direction From Reference Point Type Code	An abbreviation of the four principal directional indicators, North, South, East , West and the four points midway between, Northeast, Northwest, Southeast, Southwest, referred to as the intercardinal direction that represents the compass direction of the of the city street collision relative to the reference point.
City Reference Point Name	The name of a trafficway, landmark or, a description of a place from which an event or object can be located via a linear referencing method.
City Secondary Trafficway 1	The name of the trafficway that crosses the primary trafficway. It is used to describe the relative location of the collision on the primary trafficway between two secondary trafficways.
City Secondary Trafficway 2	The name of the trafficway that crosses the primary trafficway. It is used to describe the relative location of the collision on the primary trafficway between two secondary trafficways.
State Route ID	The concatenation of state route number, RRT and RRQ.
State Route Milepost	A logical number, assigned by a linear referencing method, to a given point along a state route. This value will always match the Accumulated Route Mileage value unless there has been a realignment or a jurisdictional transfer which changes the length of a route.
State Route Milepost Ahead_Back Type Code	An indicator that denotes if the State Route Milepost is within a back mileage equation area. A back mileage equation area occurs when a segment of a state route is added at any point other than the end of an existing state route, or when a realignment occurs.
State Route Accumulated Route Milepost or ARM	The measure of a point along the length of a route in which the distance is measured as an accrual of mileage from the beginning of the route.
State Route Number	The number assigned to the state route and enacted into law by the Washington State Legislature.
State Route Related Roadway Type Code	A code that represents a classification of a route associated with a mainline state route, examples include: Couplet, Spur, Alternate Route, etc.

Variable	Definition
State Route Related Roadway Qualifier	A physical location reference used to identify the State Route Related Roadway Type. The State Route Collision Current Related Route Qualifier can be based on a State Route Milepost, or the name of a street or city.
State Route History_Suspense Indicator Type Code	An Indicator that denotes whether a collision record has a current State Route location assigned to it or in limbo (Suspense) awaiting further review.
State Route Region Type Code	A textual description of the geographic and administrative areas of responsibility of the Washington State Department of Transportation within the State of Washington as described in the Chart of Accounts.
State Route_State Functional Class Type Code	A code that denotes a hierarchical grouping of trafficways based on the level of service they provide as set forth by WSDOT's Strategic Planning and Programming Office.
State Route Urban_Rural Type Code	An Indicator that determines whether the State Route is in an urban or rural area.
State Route Federal Functional Class Type Code	The category title for hierarchical grouping of trafficways based on the level of service they provide as established by FHWA.
State Route Vehicle 1 or 2 Compass Direction Type Code	A code that describes compass direction of the motor vehicle, pedalcycle or pedestrian involved in a state route collision.
State Route Vehicle 1 or 2 Movement Type Code	A code that identifies a movement type for the first motor vehicle, pedalcycle or pedestrian involved in a state route collision.
State Route Vehicle 1 or 2 Mile Post Direction Type Code	A code that describes the direction relative to the state route’s milepost of the second motor vehicle, pedalcycle or pedestrian involved in a state route collision.
State Route Diagram Data Collision Type Code	A type of impact for a collision based on which a collision unit struck another, the movement of the units, and other factors.
First/Second Impact Location Type Code__Effective Date 1_1_10 for City, County and Misc Traf	A type of position, based on trafficway features and direction of travel, on which the first impact in a state route collision occurred.
County Road Number	The identifier that the county assigned to the county road.
County Road Milepost	A logical number, assigned by a Linear Referencing Method, to a given point along a country road.
County Road Milepost Ahead_Back Type Code	An indicator that denotes if the county road milepost is within a back mileage equation area.
County_Intersecting County Road Number	The county road number assigned to the county road that intersects with the county road on which the collision occurred.
County_Intersecting County Road Milepost	The milepost for the county road intersecting with the county road on which the collision occurred.

Variable	Definition
County_Intersecting County Road Milepost Ahead_Back Type Code	An indicator that denotes if the county road collision intersecting milepost is within a back mileage equation area.
County_Federal Functional Class Type Code	The category title for hierarchical grouping of trafficways based on the level of service they provide as established by FHWA.
Miscellaneous Trafficway Type Code	A classification of miscellaneous trafficways based on the type of government or non-government authority that has jurisdiction over the trafficway.
Miscellaneous Trafficway Primary Trafficway	The roadway that the law enforcement officer or citizen considers to be the principal site of the collision. If the collision occurred where two roadways cross then the intersecting trafficway will contain a value.
Miscellaneous Trafficway Block Number	The building site identifier or the range of building site identifiers, in which the collision occurred, as assigned to the miscellaneous trafficway by the jurisdictional authority.
Miscellaneous Trafficway Intersecting Trafficway	The name of the roadway that crosses the primary trafficway
Miscellaneous Trafficway Distance From Reference Point	The distance from the collision scene to the trafficway reference.
Miscellaneous Trafficway Reference Point Miles_Feet Type Code	An indicator that denotes if the distance from a reference point to the collision is in miles or feet.
Miscellaneous Trafficway Compass Direction From Reference Point Type Code	An abbreviation of the four principal directional indicators, North, South, East , West and the four points midway between, Northeast, Northwest, Southeast, Southwest, referred to as the intercardinal direction that represents the compass direction of the collision relative to the reference point.
Miscellaneous Trafficway Number	A unique identifier for a miscellaneous trafficway assigned by a jurisdictional authority that has responsibility over the trafficway on or near the scene of the collision.
Miscellaneous Trafficway Milepost	A logical number, assigned by a linear referencing method, to a given point along a traveled way.
Miscellaneous Trafficway Reference Point Name	The name of a trafficway, landmark or, a description of a place from which an event or object can be located via a linear referencing method.
Miscellaneous Trafficway Secondary Trafficway 1	The name of the trafficway that crosses the primary trafficway. It is used to describe the relative location of the collision on the primary trafficway between two secondary trafficways.
Miscellaneous Trafficway Secondary Trafficway 2	The name of the trafficway that crosses the primary trafficway. It is used to describe the relative location of the collision on the primary trafficway between two secondary trafficways.

Variable	Definition
Most Severe Injury Type Code	An code that indicates the most harmful level of physical damage suffered by any one person involved in a collision.
Collision Severity Type Code	A code that represents general classification of the highest level of damage or harm that occurred in a collision.
Most Severe Sobriety Type Code	A code that represents a classification of the highest level of alcohol impairment of a motor vehicle driver, a pedestrian or a pedalcyclist involved in a collision.
First/Second Collision Type Code	A description of collision based on what was struck, and/or the orientation of the motor vehicle units in respect to one another.
First/Second Object Struck Type Code	The type of material(s) that the motor vehicle unit or pedal cyclist made impact with during a collision, other than another motor vehicle unit, or pedal cyclist.
Junction Relationship Type Code	A code that represents a type of relationship based on proximity and/or involvement, between a junction (intersection, approach, driveway or ramps) and the collision scene.
Weather Conditions Type Code	A code that represents a state of the atmosphere at the location and time of the collision.
Roadway Surface Condition Type Code	A code that represents type of precipitation or other substance found on a traveled way at or near the collision.
Lighting Condition Type Code	A code type of natural or artificial light that may be available at the scene of the collision.
Location Characteristics Type Code	A description of the immediate vicinity of the collision that further identifies the area.
Roadway Characteristics Type Code	A code that represents the alignment and topography of the traveled way.
Workzone Type Code	A code that represents the classification the work zone by the type of activity.
Investigative Agency Type Code	A code that represents the type of Law enforcement agency that filed the collision report.
ORI#	Original Investigating Agency Number - An identifier for the original agency that investigated the collision.
Reporting Agency Long Name	The full name of the law enforcement agency that reported the collision.
Reporting Agency Short Name	The detachment name of the Washington State Patrol or a short description of the law enforcement agency.
Hazardous Material Type Code	An indicator that denotes whether the commercial vehicle was displaying a hazardous material placard.
Fire Type Code	An indicator that denotes whether there was a fire caused by the collision
Stolen Type Code	An indicator that a motor vehicle unit involved in the collision was taken from its owner without right or permission.

Variable	Definition
Hit and Run Type Code	An indicator that the collision involved at least one motor vehicle which fled the scene of the collision.
Unit Number	The collision unit number of the commercial carrier that was involved in the collision.
Unit Type_Type Code	A code that represents what type person or motor vehicle unit was involved in the collision. Types include Motor Vehicle, Pedestrian, Pedalcyclist or Property Owner.
Damage Threshold Met Type Code	An indicator that identifies whether there was at least $750 worth of damage to the collision unit at the scene of the collision.
Involved Person Type Code	A code that represents the type of involved person.
Age	The number of years the involved person has lived.
Gender Type Code	A code that represents the classification of a person's gender.
Air Bag Type Code	A code that represents the level of operation for an air-inflated restraint device for a motor vehicle passenger or motor vehicle driver.
Ejection Type Code	A code that represents a state of expulsion of a motor vehicle driver or motor vehicle passenger from a motor vehicle unit in a collision.
Restraint System Type Code	A code that represents a type of seat safety device used by a motor vehicle driver or a motor vehicle passenger.
Helmet Use Type Code	A code that represents whether a motor vehicle driver motor vehicle passenger, or pedalcycle driver or pedalcycle passenger used a helmet.
Injury Type Code	A type of physical damage or harm, which a motor vehicle driver, motor vehicle passenger, pedalcyclistdriver, pedalcyclist passenger, or pedestrian incurred in a collision.
Seat Position Type Code	A code that represents type of position of the motor vehicle passenger in relation to the motor vehicle at the scene of a collision.
Sobriety Type Code	A code that represents a classification of the level of alcohol impairment of a motor vehicle driver, a pedestrian, or a pedalcyclist involved in a collision.
Alcohol Test Results	The results, or status of the results, of the blood alcohol content test of the motor vehicle driver, pedalcycle driver or Pedestrian, performed by a law enforcement officer.
Drug Recognition Class 1 or 2 Type Code	A code that represents the type of drug(s) responsible for the impairment of a motor vehicle driver, pedalcyclist driver or pedestrian involved in a collision, as determined by a trained and certified drug recognition expert.
Liability Insurance Type Code	An indicator that denotes whether the motor vehicle driver has valid liability insurance.
Unlicensed Driver Type Code	An indicator that denotes if a driver involved in the collision does not have a valid driver's license.

Variable	Definition
On Duty Police Officer Type Code	An indicator that identifies whether the motor vehicle driver was an on duty law enforcement officer, paramedic or fire personnel at the time of the collision.
Pedestrian_Pedalcyclist Clothing Visibility Type Code	A code that represents the type of visibility for the clothing a pedalcyclist or pedestrian was wearing at the time of the collision.
Pedestrian_Pedalcyclist Was Using Type Code	A code that represents the part of the roadway that the pedestrian or pedalcyclist was using at the time of the collision.
Pedestrian_Pedalyclist Type Code	A classification of the pedestrian's activity or method of movement.
Pedalcyclist Action Type Code	A code that represents a type of interaction with a motor vehicle, action related to traffic on the traveled way or action related to the trafficway of a pedalcycle driver at the time of the collision.
Pedestrian Action Type Code	A code that represents a type of the [pedestrian's activity in relation to a motor vehicle, trafficway, or intersection.
Contributing Circumstance 1, 2 or 3 Type Code	A code that represents a type of event that may have led to the involvement of a motor vehicle driver, pedalcyclist driver, or pedestrian in a collision.
Motor Vehicle Driver Miscellaneous Action 1, 2 or 3 Type Code	A code that represents a type of action taken by or an event or condition affecting the motor vehicle driver that was a factor in the collision.
Vehicle Type Code	A code that represents the category indicating the general configuration or shape of a motor vehicle distinguished by characteristics such as number of doors, rows of seats, windows, or roof line.
Towed Indicator Type Code	An indicator that denotes whether a vehicle involved in a collision had to be towed from the scene.
Government Owned Indicator Type Code	An indicator that denotes whether any vehicle involved in a collision was government owned.
Vehicle Make	A code that identifies the distinctive name applied to a group of motor vehicles by a manufacturer.
Vehicle Model	A code that represents the manufacturer-assigned code denoting a family of motor vehicles (within a make) that have a degree of similarity in construction, such as body, chassis, etc.
Vehicle Style	A code that identifies the model configuration for a motor vehicle involved in a collision.
Vehicle Year	The year which is assigned to a motor vehicle by the manufacturer.

Variable	Definition
Traffic Control Type Code	A code that identifies the traffic control type that controlled the action of the collision unit at the moment of the collision. A code that represents a sign, signal, marking, or other device used to regulate, warn, or guide traffic, placed on, over, or adjacent to a street, highway, pedestrian facility, or shared-use path by authority of a public agency having jurisdiction that controlled the action of the collision unit at the moment of the collision.
Posted Speed Limit	The speed limit posted on the trafficway where the collision occurred.
Roadway Type Code	A code that identifies the roadway type for the trafficway on which the collision unit was traveling at the time of the collision.
Vehicle Classification Type Code	A code that identifies the type of motor vehicle based on its gross vehicle weight, number of passengers carried, or type of use. Each type denotes a requirement by the Washington State Department of Licensing for a motor vehicle driver to have a commercial driver’s license to operate the motor vehicle.
Vehicle Usage Type Code	A code that identifies the functional category for a motor vehicle involved in a collision.
Registered State Type Code	A code that identifies the state for the registration of the motor vehicle.
Vehicle Action Type Code	A code that represents the classification of the movements of each motor vehicle at the moment of the collision.
Vehicle Condition 1, 2 or 3 Type Code	A code that represents a type of defect or other special factor about a motor vehicle involved in a collision.
Sequence of Events 1, 2, 3 or 4 Type Code	A code that represents a type of event for, and the kind of object struck by, a ,motor vehicle or pedalcycle involved in a collision.
Compass Direction From/To Type Code	A code that represents a cardinal direction that identifies what direction the collision unit was moving from at the moment of the collision.
Commercial Carrier Address	A composite of one or more standard address components that describes a low level of geographical/physical description of the owner or operator of a commercial motor carrier vehicle involved in a collision, as indicated on the Officer Commercial Motor Carrier Supplemental Collision form.
Commercial Carrier City Name	The name of a primarily urban political unit having corporate status and usually powers of self-government established by state charter for the commercial carrier's street address.
Commercial Carrier State Code (See Registered State)	A 2 character code that represents a territory occupied by one of the constituent administrative districts of a nation for the street address.

Variable	Definition
Commercial Carrier Zip Code	An alphanumeric code used for mail sorting, established by a postal authority, that identifies a geographic area for the street address.
Commercial Cargo Body Type Code	
Commercial Vehicle Class Type Code	A classification of motor vehicles used for business purposes. This classification scheme is based on whether the motor vehicle is considered a bus or a truck, and the truck's size, number of axles, and articulated units.
Commercial Carrier Name Source Type Code	A code that represents a source for the name of a commercial carrier.
GVWR	The maximum total weight the commercial vehicle is rated to carry.
Hazardous Material Name	The name of a substance or material which has been determined by the US Secretary of Transportation to be capable of posing an unreasonable risk to health, safety, and property when transported in commerce, and which has been so designated.
Interstate/Intrastate Type Code	An identifier that denotes whether the commercial vehicle involved in the collision is engaged in Interstate or Intrastate commerce.
Number of Axles	The number of axles on the commercial vehicle involved in the collision.
Placard Number	A unique four character field used to identify the type of hazardous material the commercial motor carrier vehicle is transporting.
Placard Suffix Type Code	A code that identifies a class of hazardous material described by the HAZMATplacard on the commercial carrier's vehicle.
USDOT Number	A unique identifying number issued by the Federal Motor Carrier Safety Administration (FMCSA) of the United States Dept. of Transportation to identify a commercial carrier.