Comprehensive Roadway Safety Data Visualization and Evaluation Platform for Yakama Nation
FINAL PROJECT REPORT

by

Yinhai Wang, Wei Sun, Hollyanna Littlebull, Meng-Ju Tsai, Sam Ricord, and Yifan Zhuang
University of Washington

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

In cooperation with the Department of Natural Resources, Confederated Tribes and Bands of the Yakama Nation

rural • isolated • tribal • indigenous

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation's Safety Data Initiative Program. This report does not constitute a standard, specification, or regulation.

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Comprehensive Roadway Safety Data Visualization and Evaluation Platform for Yakama Nation		**5. Report Date** July 2024 **6. Performing Organization Code**
7. Author(s) and Affiliations Yinhai Wang, Wei Sun, Hollyanna Littlebull, Meng-Ju Tsai, Sam Ricord, Yifan Zhuang		**8. Performing Organization Report No.** INE/CSET 24.10
9. Performing Organization Name and Address Department of Civil and Environmental Engineering, University of Washington, Box 352700, Seattle, WA 98195 Department of Natural Resources, Confederated Tribes and Bands of the Yakama Nation, 401 Buster Rd, Toppenish, WA 98948		**10. Work Unit No. (TRAIS)** **11. Contract or Grant No.**
12. Sponsoring Organization Name and Address U.S. Department of Transportation 1200 New Jersey Ave, SE Washington, DC 20590		**13. Type of Report and Period Covered** Final Report (October 2020 - April 2022) **14. Sponsoring Agency Code**

15. Supplementary Notes

Conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration

16. Abstract

The Yakama Nation Department of Natural Resources (DNR) Engineering collaborated with the Smart Transportation Application & Research Laboratory (STAR Lab) at the University of Washington to develop a comprehensive roadway safety data visualization and evaluation platform. With the U.S. Department of Transportation's Safety Data Initiative (SDI) fund, this tool will support information for the Yakama Nation government for their decision-making. The safety datasets provided in this tool consist of collision records (collision, vehicle, occupant, pedestrian) and roadway characteristics (roadlog, curve and grade, ramp, traffic information, special-use lane, etc.). The multi-source database supported data collection, quality control, integration, database management, visualization, and analytical results. The safety tools can be utilized for analytical and visualization functions such as crash data visualization, hotspot identification, and network screening. Examples of available safety data include crash type, frequency, severity, and risk estimate, and safety data download.

17. Key Words		**18. Distribution Statement**	
Crash Modeling Analysis, Multi-Source Database Management, Web-based Visualization Platform, Network Screening, Hotspot Identification			
19. Security Classification (of this report) Unclassified.	**20. Security Classification (of this page)** Unclassified.	**21. No. of Pages** 50	**22. Price** N/A

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized.**

S.I.* (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)		
^oF	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	^oC
		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)		
^oC	Celsius	1.8C+32	Fahrenheit	^oF
		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

The disparity between transportation safety in rural and urban areas is enormous due to area characteristics. One of the tribes in Washington State with high statistical road collisions is the Confederated Tribes and Bands of the Yakama Nation. This project aims to improve transportation safety for both rural and Native American populations by developing a comprehensive web-based roadway safety tool. With the fund from the U.S. Department of Transportation's Safety Data Initiative (SDI) and collaboration from the Smart Transportation Application & Research Laboratory (STAR Lab) at the University of Washington, this comprehensive roadway safety data visualization and evaluation platform will support information for the Yakama Nation government for their decision making.

In 2013, the STAR Lab team developed the Digital Roadway Interactive Visualization and Evaluation Network (DRIVE Net), which stores Washington State Department of Transportation (WSDOT) traffic data and incorporates third-party data [1][2]. The DRIVE Net was used as the base for developing this safety tool. The database collected, cleaned, and stored safety-related datasets, which consist of collision records (collision, vehicle, occupant, and pedestrian) and roadway characteristics (roadlog, curve and grade, ramp, traffic information, special-use lane, etc.). By using a multi-source database management system, data management processes such as data collection, quality control, integration, database management, visualization, and analytical results can be performed. In addition, the safety tool can provide analytical and visualization functions, including crash data visualization, hotspot identification, and network screening based on crash type, frequency, severity, estimated risks, reporting, and safety data download as well.

The advanced analytical techniques and powerful visualization provided in the safety tool will be beneficial to support decision-making by providing insight data. The Yakama Nation can prioritize the high-risk roadway segments and apply countermeasures to reduce risk factors that have been analyzed on the selected roadway sections. These analytical methods and safety tools can support other state and local governments with similar safety issues with rural roadways in Yakama Nation.

CHAPTER 1. INTRODUCTION

1.1. Project Background

The Yakama Nation reservation has approximately 1,200 miles of public roads under their responsibility. According to National Highway Traffic Safety Administration, 45% of 36,560 motor vehicle traffic fatalities in 2018 occurred in rural areas [3]. Moreover, the fatality rate on rural roads has been more than two times higher than in urban from 2009 to 2018. Yakima County's 10-year trend ranked it among the worst counties in Washington [4]. There were 330 traffic fatalities between 2011 and 2020, or 13.15 per 100,000 people. In addition, Yakama nation also has the highest pedestrian fatality rates. The Yakama Nation Department of Natural Resources (DNR) engineer program has been working on safety issues to decrease the nation's pedestrian and vehicle fatality rates.

Many risk factors have been recognized as the cause leading to road crashes, such as roadway geometrics (sharp curves, steep hills, rough/no pavement, lack of markers and signs, etc.), lack of pedestrian facilities, and adverse weather conditions (heavy fog, snow, ice, etc.). According to the Safety Management Plan developed by Yakama Nation DNR Engineering program (2018), the fatality and severe injuries were caused by driving under the influence, speeding, unrestrained driving/occupant, and distracted driving. For the past 20 years, speeding has contributed to almost one-third of all fatal car accidents [5]. Evidence shows that speeding makes collisions between vehicles and pedestrians more likely and severe. There is no effective law enforcement on speed in Yakama nation, which results in high speeding on their road network. With all factors that lead to the roadway crashes within the Yakama Nation reservation mentioned above, the government was unclear on how to transform these contributing factors into practical insights, such as how to prioritize and allocate budget for improving safety on roadways/intersections to achieve the most benefit and in an efficient way.

In this research, the roadway safety problem was identified and defined as rural roadway crashes based on frequency and severity. These crashes were analyzed for their variety of contributing factors. The high-risk locations were then identified, and information for supporting vehicles and pedestrians was made.

1.2. Research Objective

The following questions were raised after the roadway safety problem of the Yakama Nation was identified:

1) What are the contributing factors to the roadway crashes in Yakama Nation?
2) What are the contributing factors affecting the crash frequency and severity?
3) Given limited funding and resources, what are the priorities of the Yakama Nation to improve traffic safety conditions on their roadway network?
4) What are the contributing factors of road safety inequality in rural areas, especially for tribal communities? For example, lack of traffic and road safety data for planning and operations

The Yakama nation can turn safety data into actionable insights for decision-making by developing data-driven solutions. The quantitative modeling of roadway crashes and their contributing factors, developing by using safety-related data, will be able to solve the defined safety problem. This model can address the first two questions. Decision-makers can select effective countermeasures with a complete evaluation and comprehension of the crash's contributing factors. The gap between safety data and actionable insights reflects the knowledge gap of decision-makers, which will be addressed in question three. The analytical result with extensive visualization will address the contributing factors, the impacts on roadway crashes, and the vehicle and pedestrian safety ranking on roadway segments. These data will be beneficial for transportation practitioners to make effective and powerful decisions on policies and budget allocations.

In 2017, Federal Highway Administration (FHWA) initiated efforts to improve safety by developing the Tribal Transportation Strategic Safety Plan (TTSP) [6]. The Yakama nation also collaborated with federal, state, county, and city transportation agencies to work on this plan. Besides the initial TTSP, the Safety Management Plan was developed based on Federal Highway Administration's (FHWA) and Washington State's Strategic Highway Safety Plans (SHSP), which launched in 2018. While the safety plan provided a robust platform for developing a policy and decision-making process supported by data-driven safety analysis, the current method of making safety decisions based on basic crash data analysis still needs to be improved. The following areas are the main focuses area of this research that the team identified and worked on:

1) **Data management system for multi-source data:** The system used for safety-related datasets management is still missing. Developing a database management system will allow traffic engineers and planners access to road safety data. The integrated safety-related datasets will be used for performing advanced analysis and data visualization.
2) **Advanced analytical methods for roadway crash modeling and prediction:** Besides the database management system, the advanced analytical methods used for analyzing roadway

crashes will be developed to produce accurate crash models. This model can generate predictive insights.

3) **An efficient platform for safety data visualization:** A comprehensive platform that can provide efficient visualization of the data and the analytical results is needed to illustrate data analysis results. The high quality of visualization enables a user to gain valuable insight into data, leading to efficient interpretation.

CHAPTER 2. LITERATURE REVIEW

2.1. Pedestrian Safety on Rural roads

In 2022, the Governors Highway Safety Association or GHSA released that there is an average of 20 pedestrian deaths every day [7]. Roadway characteristics and design impact pedestrian safety. Pollack et al. investigated traffic-related hazards for pedestrians using multiple data collection strategies [8]. After performing the descriptive epidemiologic study, it was identified that a lack of signage posting the speed limit, faded crosswalks, issues with a traffic light and walk sign synchronization, and limited formal pedestrian crossings are the risk factors that cause road accidents on the street. Stoker et al. found that the built environment can either increase or reduce pedestrian traffic injury; hence, planners must design road facilities to minimize pedestrian risk [9]. Speed has a strong relationship with crash severity. Garder analyzed pedestrian crashes in Maine using prediction models to compare the crash number and predicted outcome [10]. The result shows that high speeds and wide roads lead to more crashes. This study also suggests that safety improvement should be focused on arterials and major collectors. Pedestrian road crossing behavior can result in increasing the risk of collision.

Martin has investigated the pedestrian behavior that might be influenced to reduce the number of casualties on London's roads [11]. It was found that, in general, enforcement measures, such as speed limits, are best applied on strategic routes, where physical traffic calming measures would reduce capacity. An adequately built environment will encourage pedestrians to behave safely while crossing. Weather significantly impacts the likelihood of pedestrians being killed or having a severe injury in a car crash. Zhai et al. developed a mixed logit model to determine the association between pedestrian crash severity and possible risk factor [12]. The study found that drivers are likelier to have less attention and be reckless concerning pedestrians in hot weather. At the same time, the presence of rain increases pedestrian crash severity and changes driver behavior significantly. Many studies investigate the risk of severe injury and death to pedestrians influenced by the type of vehicle. Lefler and Gabler found that pedestrians have a two to three times greater likelihood of dying when struck by a lightweight truck than when stuck by a car since light trucks are heavier, stiffer, and blunter than passenger cars [13]. Qi and Guoguo revealed that the percentage of left-turn trucks and pedestrian volume significantly impact pedestrian safety [14].

Rural areas have different characteristics in the road network, land use, and travel patterns, which consequently differ in collision crash characteristics. The report “America’s Rural Roads: Beautiful and

Deadly" reports that 85,002 people died in crashes on rural roads between 2016 and 2020. Moreover, the risk of dying in a crash was 62% higher on a rural road than on an urban road of the same length. Hence, it is crucial to predict pedestrian collisions on rural roads so that planners can provide a sufficient built environment for pedestrians to cross or use the road safely. In tribal lands, it was found that factors such as limited sight zones, alcohol involvement, lacking sidewalk facilities, and uncontrolled intersections are involved in pedestrian crashes. Vachal performed descriptive statistics on nationally fatal U.S. crashes involving pedestrian motor vehicle crashes (MVC) [15]. This study's result helps decision-makers gain a better understanding of a tribal pedestrian-involved crash. According to Naumann, American Indian, and Alaskan Native populations had the highest age-adjusted pedestrian fatality rate of all races [16]. Iragavarapu reviewed that the main contributing factors that lead to fatality in tribal populations are rural location, poverty, and lack of traffic control devices [17]. Davis developed a model for predicting indicators of a local street's safety for pedestrians [18]. Some of the model's variables can be random since pedestrian-vehicle collisions are rare on a local road. Zajac and Ivan use an ordered probit model to evaluate the effect of the roadway and area-type features on the injury severity of pedestrian crashes in rural Connecticut [19]. It was found that pedestrians in low-density residential areas have experienced higher injury than pedestrians in compact residential when attempting to cross two-lane highways. Hall et al. examined all rural pedestrian collisions in New Mexico for three years [20]. They found that clear weather, hours of darkness, weekends, non-intersection locations, and level, straight roads are the most prominent characteristics of rural pedestrian fatalities. The planner can take corrective action to provide a safer design of rural highways by considering roadway alignment and profile. From this study, rural vehicle crashes are more likely to occur under conditions of adverse geometrics, particularly sharp horizontal curves and steep downgrades. When these factors act together, it was found that they will contribute more to the occurrence of pedestrian crashes. Baireddy et al. found that variables such as the number of lanes, driver condition, and pedestrian condition contributed to pedestrian crashes in rural Illinois [21].

Determining the cause of pedestrian injuries is critical for enhancing the safety of vulnerable road users. The traditional statistical model, such as the regression model, was often used in previous studies to determine the factors that lead to road fatality. However, many pre-assumptions were made to develop those models, adding a large inflexibility scale. An advanced machine learning algorithm issues pedestrian safety to eliminate these limitations. Tao et al. employed a Bayesian neural network model on the Australian Road Death Database to generate sound pedestrian death forecasts based on individuals' characteristics, time, occasions, road characteristics, and crash attributes [22]. Riccardi uncovered relations between road infrastructure, environmental, vehicle, driver-related patterns, and pedestrian crashes using the mixed logit model and machine learning [23]. Das et al. develop a framework for applying machine-learning models to classify crash types [24]. By using the XGBoost model, the advanced machine-learning models can detect patterns and trends of crash mechanisms with the highest accuracy rate of up to 77% for training data. Machine-learning techniques can be applied to address the crash type associated with pedestrians and vehicles.

2.2. Safety Tools

A safety tool is a tool that facilitates practitioners in evaluating safety on the road. Federal highway administration, FHWA, has developed various safety tools in both software and manual [26]. Various safety tools are available from FHWA, such as SafetyAnalyst, Interactive Highway Safety Design Model (IHSDM), and the Crash Modification Factors (CMF) Clearinghouse. Combined with the Highway Safety Manual (HSM), these tools will provide insightful data for state and local highway agencies to plan and support their project development decision-making. Interactive Highway Safety Design Model (IHSDM) is a software analysis tool that evaluates highway geometric design's safety and operational effects [27]. The IHSDM consists of five main evaluation modules: Crash Prediction, Design Consistency, Policy Review, Traffic Analysis, and Driver/Vehicle. AASHTOWare is a Software as a Service (SaaS) platform for highway traffic safety management [28]. This tool aims to facilitate the decision-making process that will identify and manage site-specific highway safety improvements cost-effectively.

Safety tools can be provided to evaluate the effectiveness of safety improvement. Texas A&M Transportation Institute has developed screening tools to predict the number of crashes at road segments or intersections based on historical frequency and road characteristics. They also developed systemic tools that calculate the potential of safety improvement from improving on-site safety, such as installing a median barrier or guardrail. Texas Department of Transportation has developed safety score tools to aid designers in making safety-driven decisions [30]. Geometric design, traffic, and roadside elements are used to calculate the safety score among various configuration designs. Another tool is RSAP, an encroachment-based computer software tool for the cost-effectiveness evaluation of roadside safety improvements. A series of probabilities representing vehicle roadside, conditional of a crash, severity of crashes, and the expected benefit-cost ratios are used for analysis in RSAP. Four modules are available in RSAP; Encroachment module, Crash prediction module, Severity prediction module, and Benefit-cost module.

Visualization is another type of safety tool that can be used to gain insight into safety data. The Pennsylvania Crash Information Tool (PCIT) provides various reports that allow the public to learn about traffic crashes, fatalities, and injuries statewide and in specific counties or municipalities [32]. The reporting tools function enables users to retrieve the number of crashes, people involved, or vehicles involved, which can be filtered by Month and Year, County/Municipality, and various crash characteristics. The crash download function generates reportable crash statistics shown in a multi-year period. The public datasets function can be used to analyze each crash record's details.

CHAPTER 3. METHODOLOGICAL APPROACH

3.1. Data

The safety data tool has integrated crash, roadway characteristics, and weather data. The STAR Lab team collaborated with Yakama Nation DNR Engineering team and checked the available safety-related data sources. Those data are collected and stored in the developed database management system. The steps for data preparations are as follows:

1) **Data Collection**: There are many agencies that the UW research team and Yakama Nation DNR Engineering team have worked with to acquire the data. The Highway Safety Information System (HSIS) provided the safety-related datasets. The Washington State Department of Transportation (WSDOT) provided statewide weather station data, WSDOT crash data portal, and WSDOT crash data portal. The datasets were cleaned before being stored in the database.
2) **Data Integration:** Since multi-source datasets were integrated into the database system, there is a high chance that some of the attributes might be duplicated. The redundant data across the data sources were deleted to reserve space and allow faster database computation. In case the deleted redundant attribute is going to be used, the correlation analysis and monitoring were performed to determine whether the original data can be restored
3) **Data Management:** MS SQL and PostgreSQL databases were used in this research for data management.
4) **Quality Control:** This research has applied advanced data integration and imputation algorithms for controlling data quality in the data management system. It is usual to have an error in the dataset. Hence, the data cleaning process is necessary for maintaining the quality of this database. The data cleaning procedure implemented in this research were redundant data processing, missing data processing, noise data processing, etc.

3.2. Analytic Approach

Besides the visualization of the collision records, the accurate roadway crashes model will be beneficial for transferring data into actionable insights. Several crash modeling methods were explored and investigated in this research and incorporated into the safety data tool. The model will be used for visualization and conveniently identifying high-risk roadways/intersections. The model will be assessed for its accuracy, feasibility, and efficiency in order to be implemented in a web-based environment.

3.2.1. Road segment performance estimation

Figure 1 shows the definition of road segments. “A” represents intersection locations, while “B” represents road segments. Note that the road segment starts at the center of an intersection and ends at either the center of the adjacent intersection, or any changes from one homogeneous segment to another.

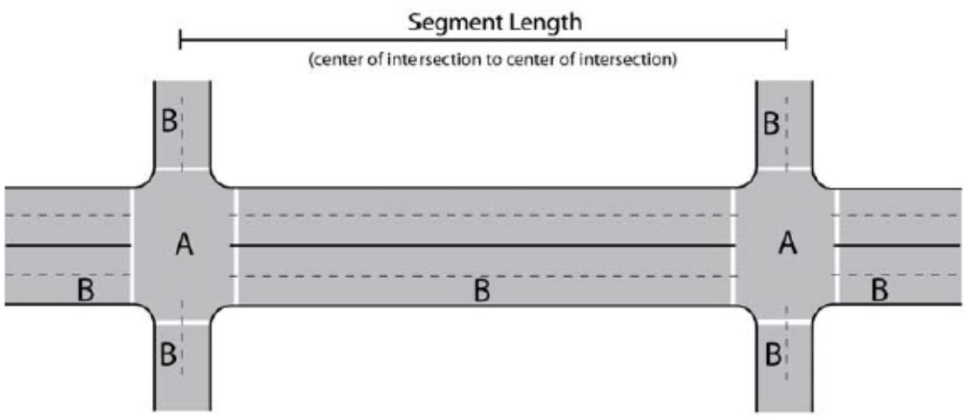

Figure 1 Road Segment Definition based on Highway Safety Manual (HSM) [1]

The single-vehicle driveway collision is selected as the focused fundamental prediction model in this research. This model is based on the Poisson-gamma model, which is denoted in equation 1 as follows:

$$Y_{it} = \theta_{it} \sim Po(\theta_{it}) \tag{1}$$

Where Y_{it} = the number of crashes in the segment i and at the time stamp t

θ_{it} = The mean of the Poisson (assumed to be Poisson distributed and independent overall road segments and time stamps)

The mean of the Possion, θ_{it}, is defined in equation 2 as follows:

$$\theta_{it} = \mu_{it} \exp(\varepsilon_{it}) \tag{2}$$

Where μ_{it} = a function integrating all covariates

ε_{it} = an independent variance and indicates model errors

A function integrating all covariates, μ_{it}, is defined in equation 3 as follows

$$\mu_{it} = \exp(\beta_0 + \beta_1 X_{it1} + \cdots + \beta_k X_{itk}) \tag{3}$$

Where β = unknown coefficients, which should be regressed based on real-world data

The single-vehicle driveway collision prediction model is defined in equation 4 as follows:

$$N_{brvs} = e^{a+b\cdot\ln(AADT)+\ln(L)} \tag{4}$$

Where a, b = coefficients

AADT = Average daily segment volume (veh/day)

L = the road segment length (mile)

Equation 4 can be expanded into the multiple-vehicle driveway-related collisions, which is the sum of the driveway-related collisions from different driveway types as shown in equation 5 as follows:

$$N_{brdwy} = \sum_{all\ types} n_j \cdot N_j \cdot \left(\frac{AADT}{15000}\right)^t \tag{5}$$

Where n_j = the number of driveways within the target road segment of driveway type j

N_j = the number of collisions related to driveway type j per year

t = the coefficient for traffic volume adjustments

Usually, there are several primary road segment types, including major commercial driveways, minor commercial driveways, major industrial/institutional driveways, etc. However, this research only used one type for each road segment type—the original model using regression data from Minnesota and Michigan states. The total number of road segments in the original model is 4,255, totaling 598.3 miles. This research validated the model by using data from Washington state.

Chapter 12 of the Highway Safety Manual (HSM) presents the predictive method for urban and suburban arterial facilities to estimate the expected average crash frequency, severity, and collision types [25]. According to HSM, five primary crash modification factors (CMFs) can be applied to equations 4 and 5. Adding the CMFs to those equations, the crash prediction model is shown in equation 6 as follows [2]:

$$N_{rs} = N_{br} \cdot (CMF_{1r} \cdot CMF_{2r} \cdot \ldots \cdot CMF_{nr}) \tag{6}$$

Where N_{rs} = the predicted number of total crashes in road segments after adding CMFs

$CMF_{ir}\ (i = 1, \ldots, n)$ = CMF for various road segment features, which are limited to primary CMFs mentioned above

The final prediction model is shown in Equation 7. Note that this research took only four significant CMFs into account.

$$N_i = \beta_0 L_i (AADT_i)^{\beta_1} \cdot e^{\beta_2 W_L + \beta_3 W_S + \beta_4 W_M + \beta_5 N_L} \tag{7}$$

Where N_i = the total number of predicted crashes in the road segment i

L_i = the length of the road segment i

AADT = Average daily segment volume (veh/day)

W_L = lane width

W_S = Shoulder width

W_M = Median width

N_L = number of lanes

β = represents unknown coefficients

These coefficients are calculated based on data collected in Washington state. The details can be found in Table 1

Table 1 Variables of Prediction Model in Equation (7)

Variables	Descriptions
AADT	Two-way average traffic volume
Segment Length $\boldsymbol{L}$	Length of homogenous segments
Lane Width $\boldsymbol{W_L}$	The average width of all lanes
Shoulder Width $\boldsymbol{W_S}$	The average width of the shoulder center in the two opposite directions
Median Width $\boldsymbol{W_M}$	Average median width along segments
Number of Lanes $\boldsymbol{N_L}$	Number of total lanes within the segment

CHAPTER 4. SAFETY DATA TOOL

4.1. Development Approach

The STAR Lab team has developed a platform named the Digital Roadway Interactive Visualization and Evaluation Network (DRIVE Net) which is a transportation multi-source data management, fusion, analysis, and visualization platform. The safety tools development was based on this platform. Transportation data analysis and visualization functions, such as crash record filtering and visualization, roadway network safety performance analysis and visualization with incident frequency, speed ratio and traffic flow map, HCM level of service analysis, travel time reliability analysis, elevation data visualization, and emission map can be used for supporting decision making.

The existing roadway safety analysis module on DRIVE Net can be used for developing the safety tool for Yakama Nation as an extension. Integrating multi-source safety-related data, performing quality control on the data, supporting modeling and analysis, and providing visualization functions can be done in the safety tool. The multi-source database system in MS SQL Server and PostgreSQL databases stored and managed collected safety-related datasets. The multi-source database system performed data checking and ensured data consistency and integrity. In addition, a user guide document was developed for Yakama Nation and the safety tool's development. This user guide document consists of the implementation and maintenance guidelines for the safety tool.

4.2. Analytical and Visualization Functions

One of the highest rates of traffic deaths is found in Yakima County. Moreover, the county has the most significant number of statewide pedestrian deaths. The primary safety concerns in Yakima County are roadway geometry, human behavior, weather, the lack of adequate pedestrian crossings, the lack of pedestrian-vehicle separation on major roadways, and inadequate signage to warn vehicles of pedestrians. Based on these concerns, the research team then built a web-based comprehensive roadway safety tool for traffic planners and engineers of Yakama Nation Department of Natural Resources (DNR) Engineering (Figure 2). The safety tool consists of data from various sources which are Yakama Nation DNR Engineering, the Digital Roadway Interactive Visualization and Evaluation Network (DRIVE Net), Washington State Department of Transportation (WSDOT), Highway Safety Information System (HSIS), and OpenStreetMap.

Figure 2 The Homepage of the Safety Tool

The developed safety tool consists of various applications (Figure 3). Users can access each application on the web tool navigational page.

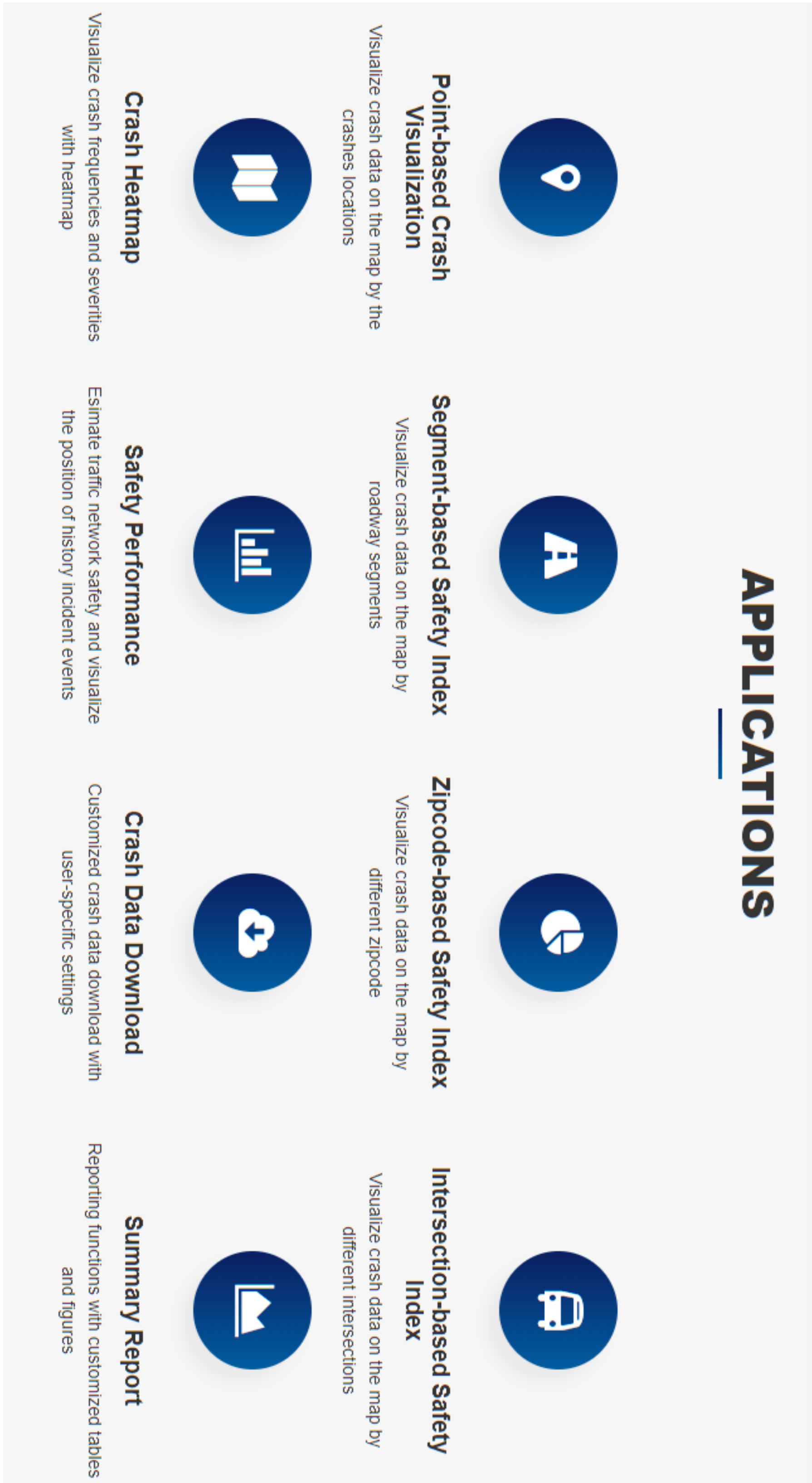

Figure 3 Functions of the Web-based Safety Data Tool

Users can query data based on year, month, day on a weekday, roadway class, and severity. These filters are available for point-based crash visualization, segment-based safety index, zip-code-based safety index, and crash heatmap. Note that multiple filters can be selected.

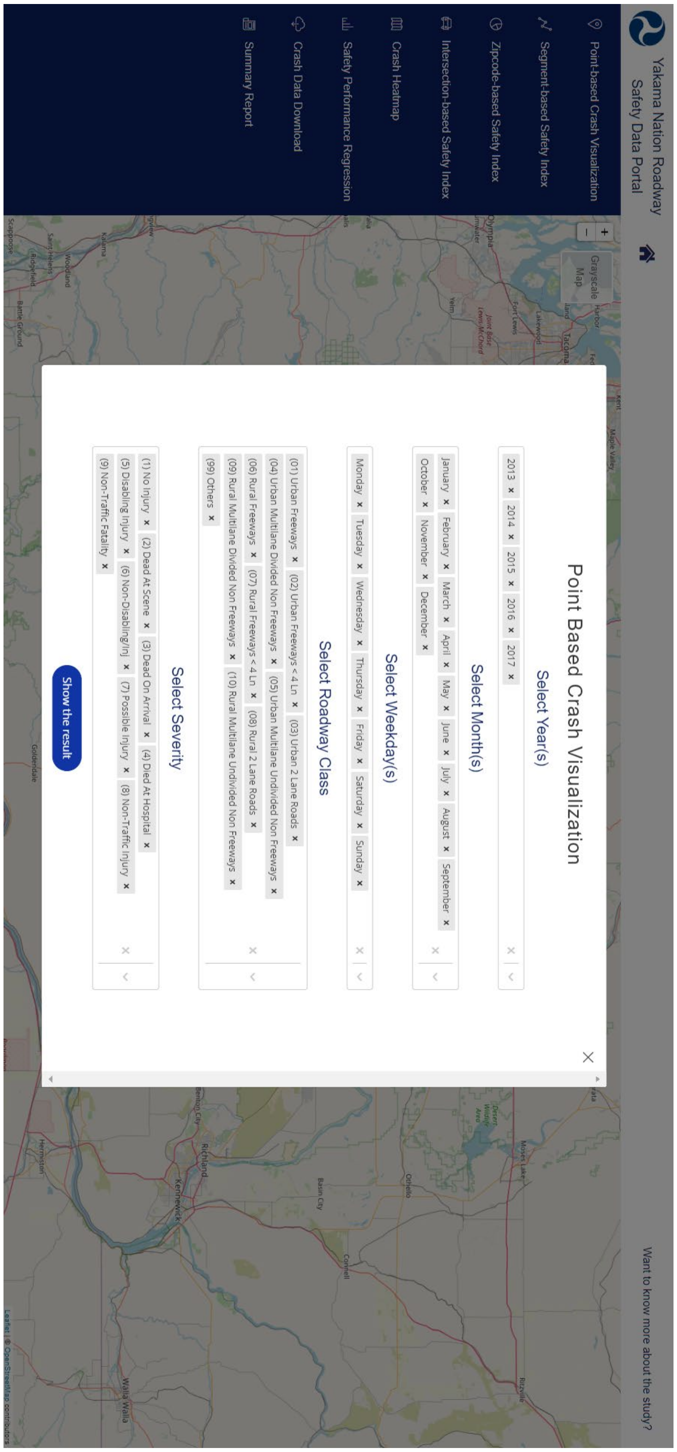

Figure 4 Data Query Available in Safety Tool

Point-Based Crash Visualization shows the collision record on the map by location (Figure 5). The records can be distributed at a finer level when the map expands (Figure 6). Each record has details of year, roadway type, and crash severity (Figure 7).

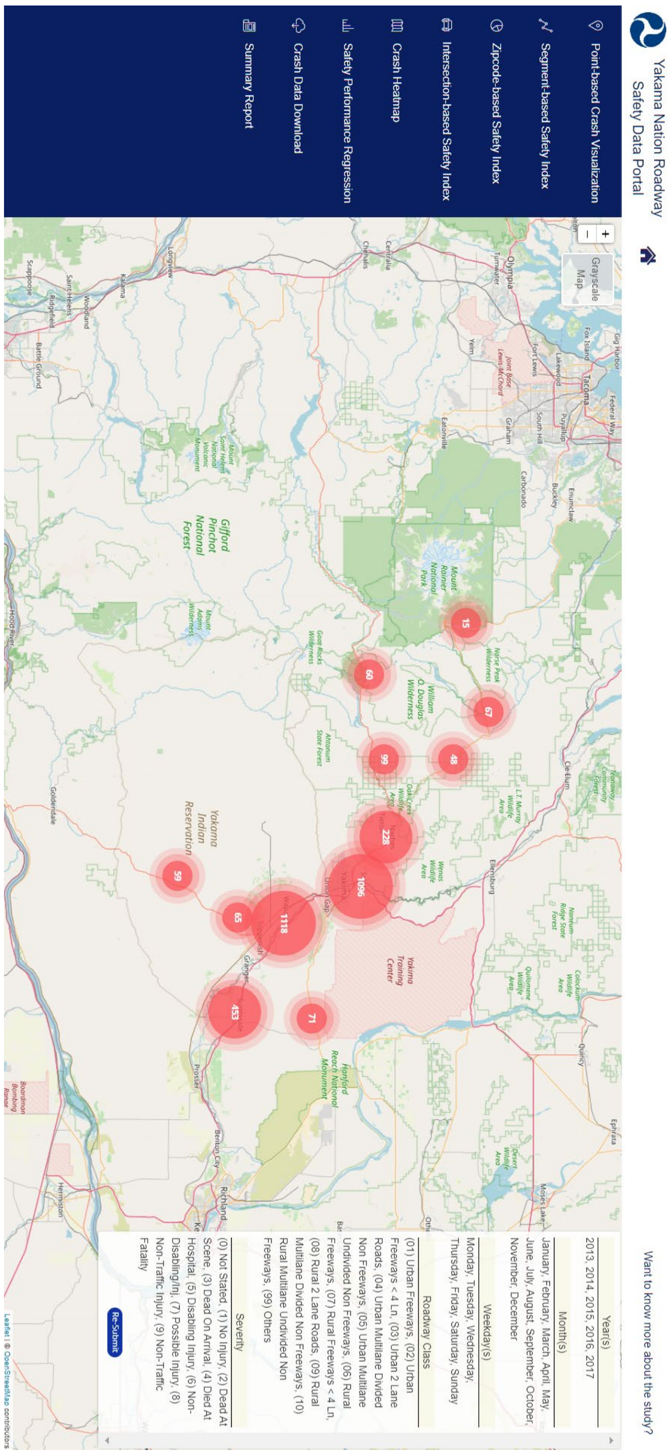

Figure 5 Point-Based Crash Visualization

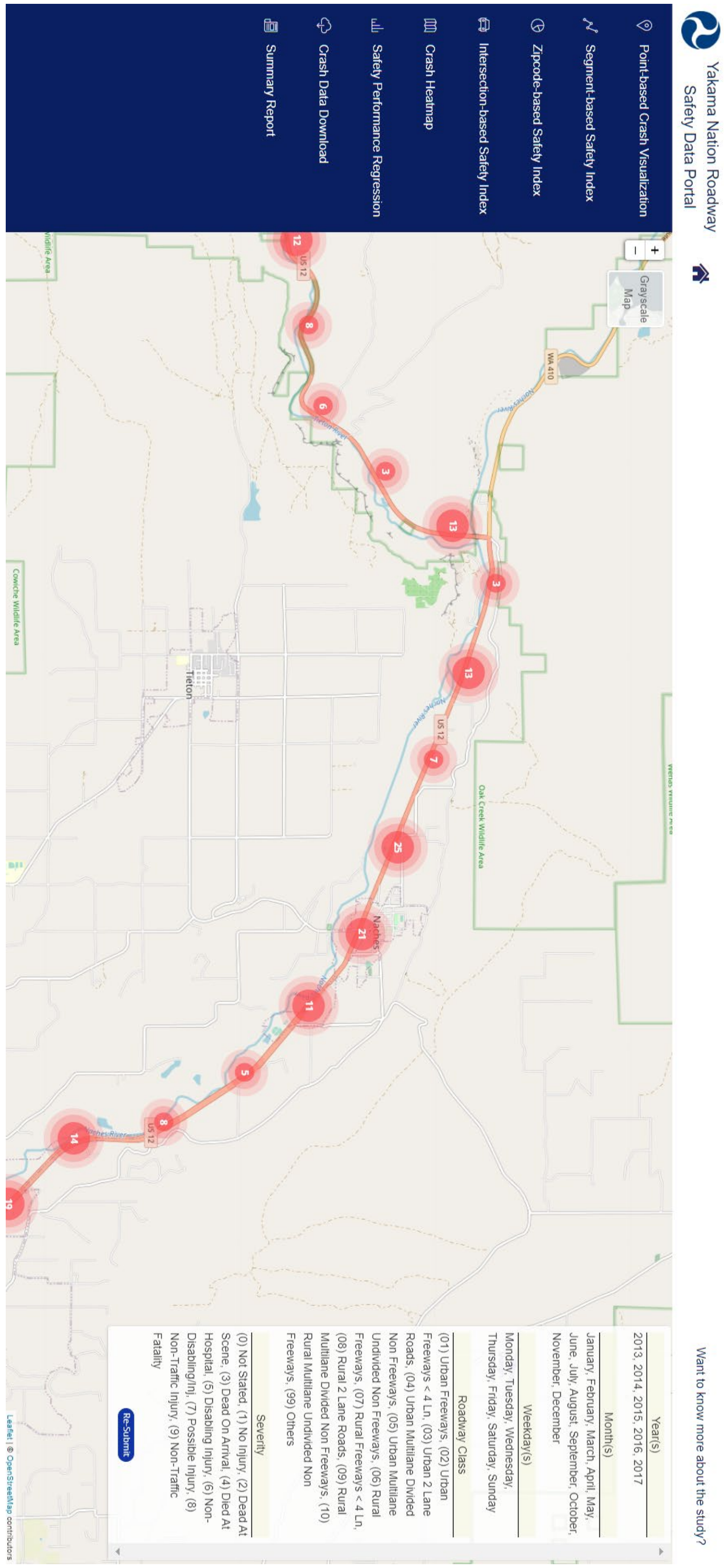

Figure 6 Expanding Map of the Point-Based Crash Visualization

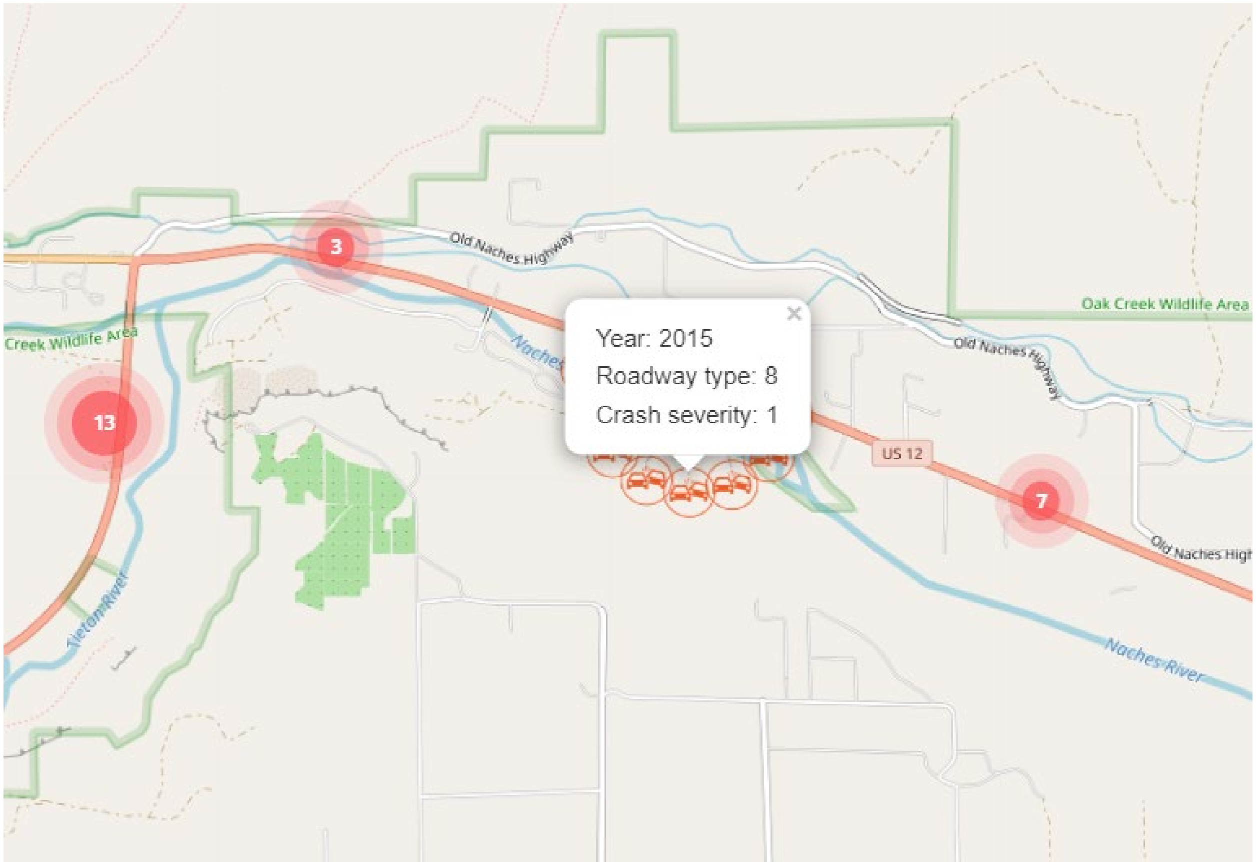

Figure 7 Details of Each Record

Segment-Based Crash Visualization shows a visualization of historical collision frequency or estimated frequency based on crash modeling analysis on the map by the segment (Figure 8). The green line color represents lower cases, while the red line color represents higher cases. The detail of historical collision frequency or estimated frequency based on crash modeling analysis, route number, and begin and end milepost are shown when the user clicks on each segment (Figure 9).

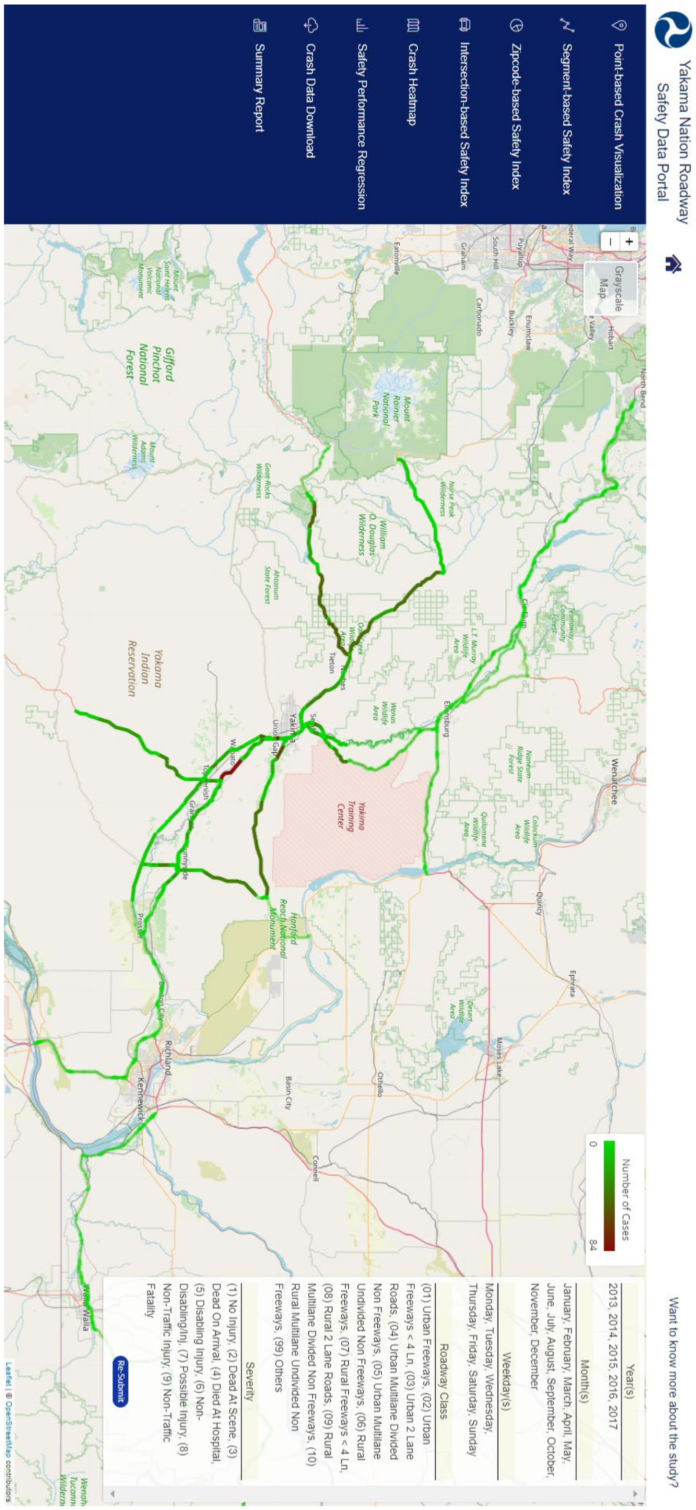

Figure 8 Segment-Based Crash Visualization

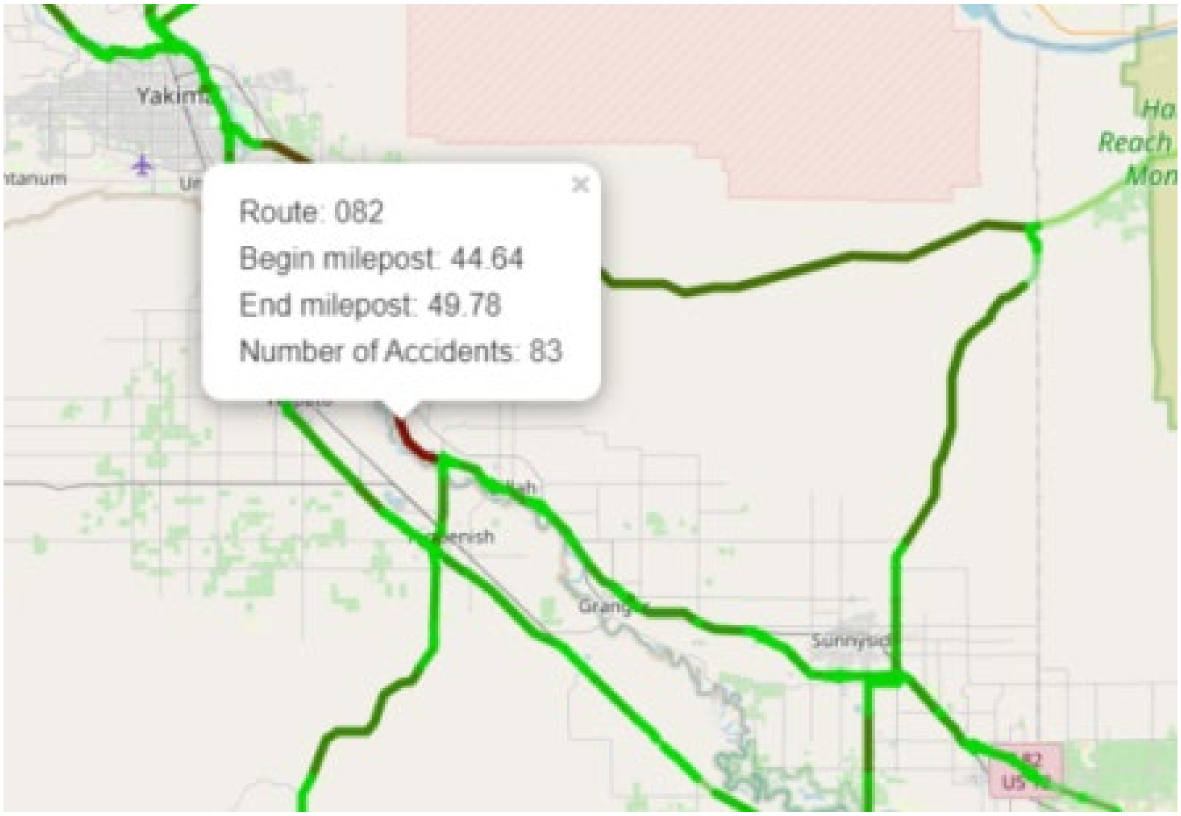

Figure 9 Detail of Segment Collision Frequency

Zip Code-Based Crash Visualization shows historical collision frequency on the map by area of each zip code (Figure 10). Red colors mean higher cases, while green color means fewer cases. When the user clicks the area, the number of accidents and zip code will show up (Figure 11)

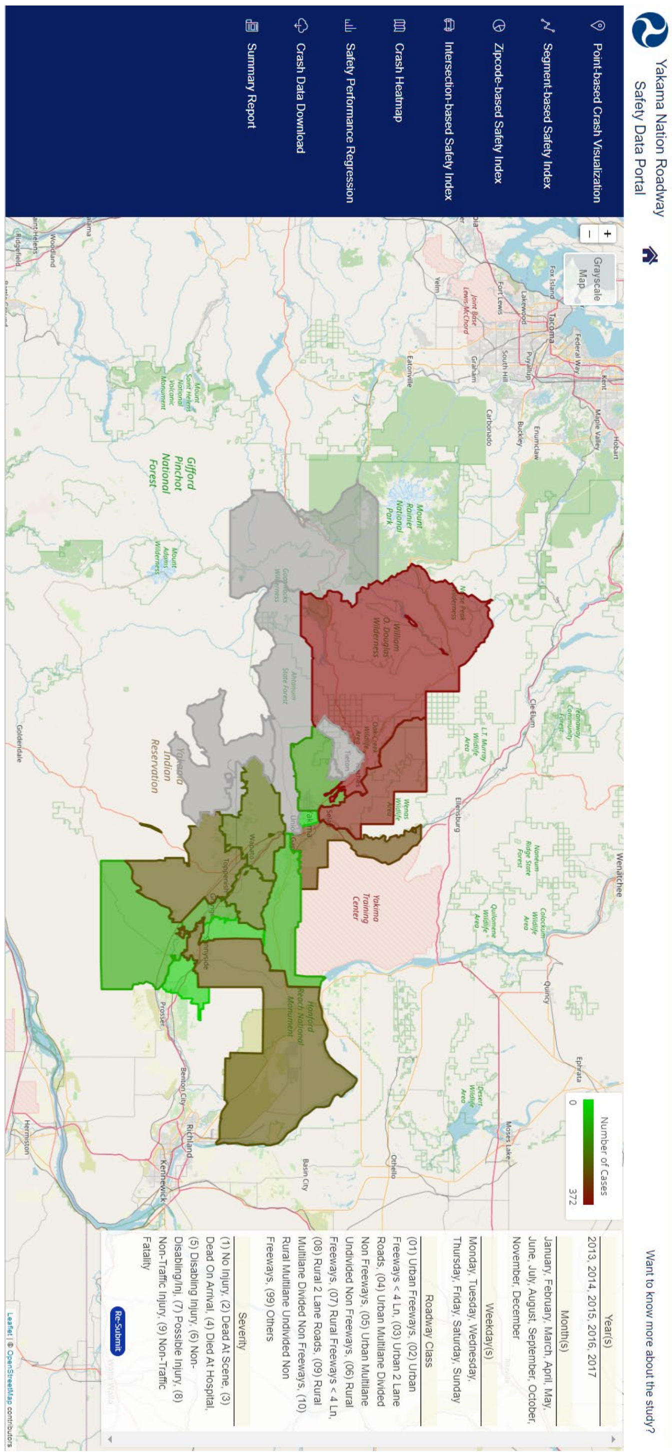

Figure 10 Zip Code-based Crash Visualization

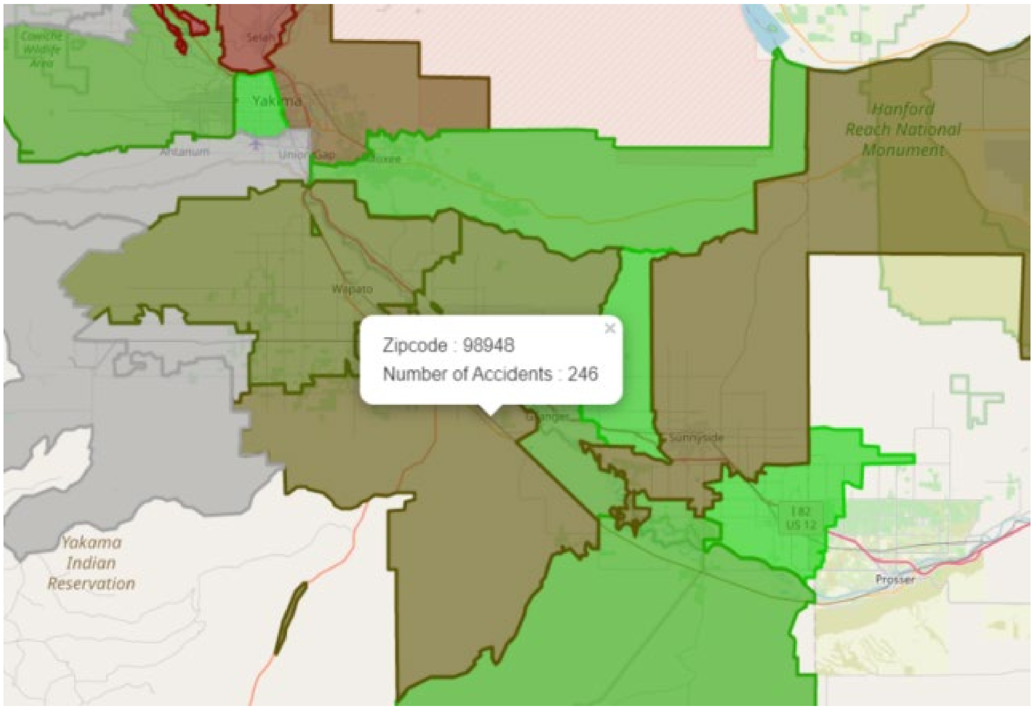

Figure 11 Detail of the Number of Accidents in Each Zip Code Area

Intersection-Based Crash Visualization shows the intersections' location with historical collision frequency (Figure 12). The name of the intersection and the total number of accidents will show up when the user clicks on the location of the intersection (Figure 13). The number of crashes can be classified by severity and shown in a bar graph based on the zip code area that the user selected (Figure 14). The number of accidents in each zip code area are listed by the top five dangerous intersections.

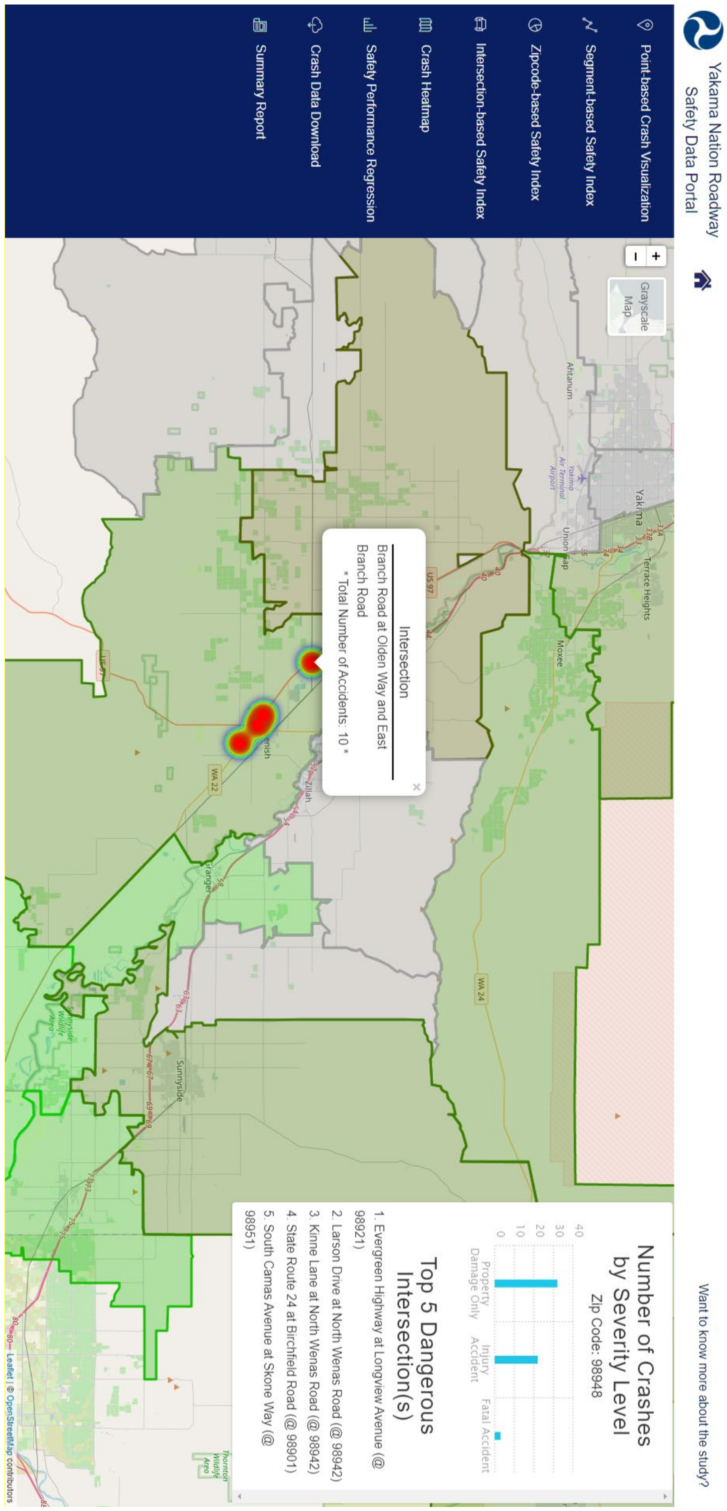

Figure 12 Intersection-based Crash Visualization

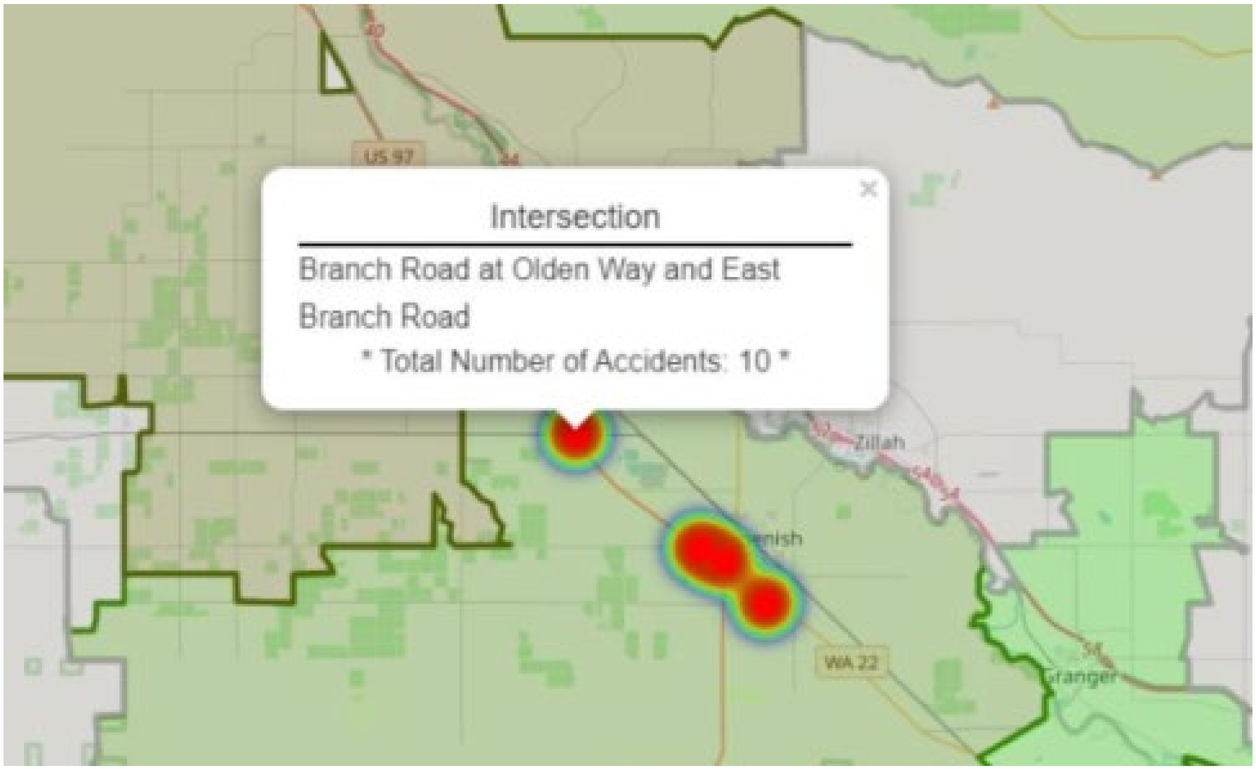

Figure 13 Number of Accidents at the Selected Intersection

Figure 14 Additional Intersection Crashes Data Base on Selected Area

Crash Heatmap shows the intensity of the collision records based on location (Figure 15). The red area means the higher intensity of the frequency of collision. When the map is expanded, the user can obtain a more precise location with high intensity of the collision records.

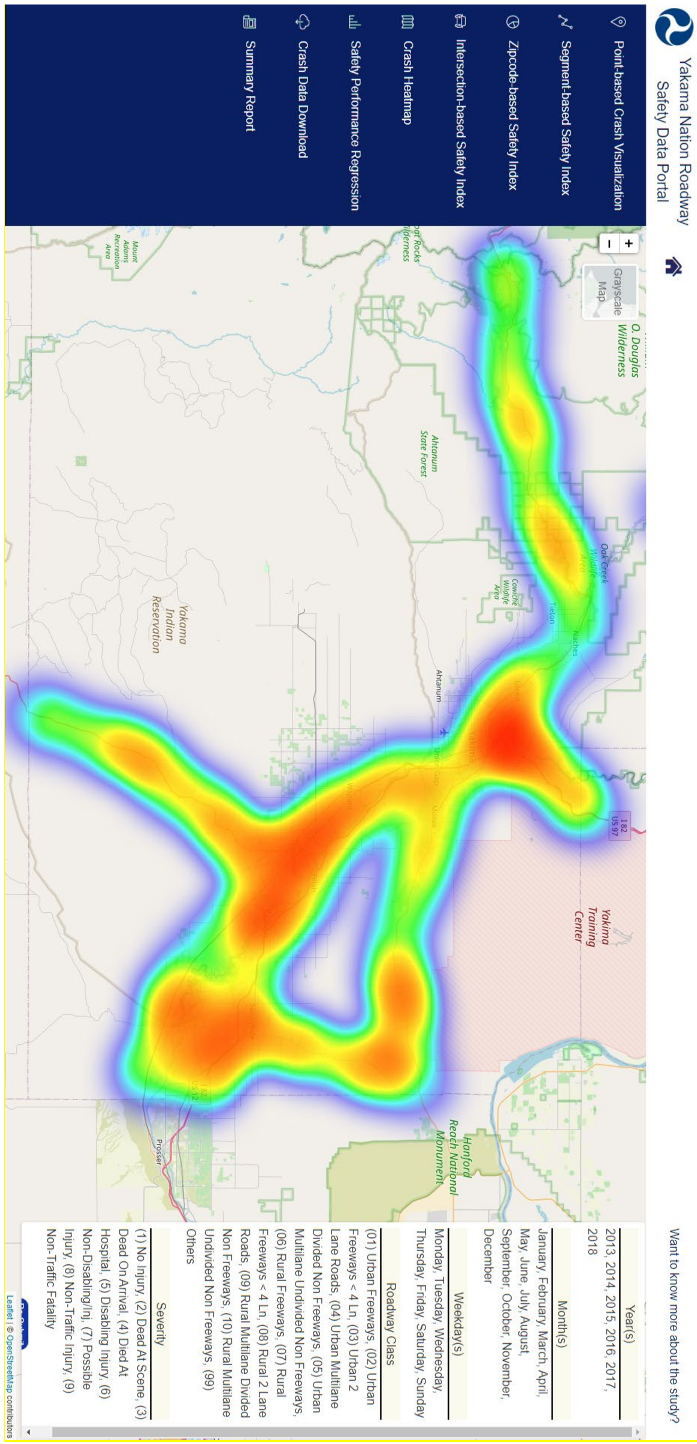

Figure 15 Crash Heatmap

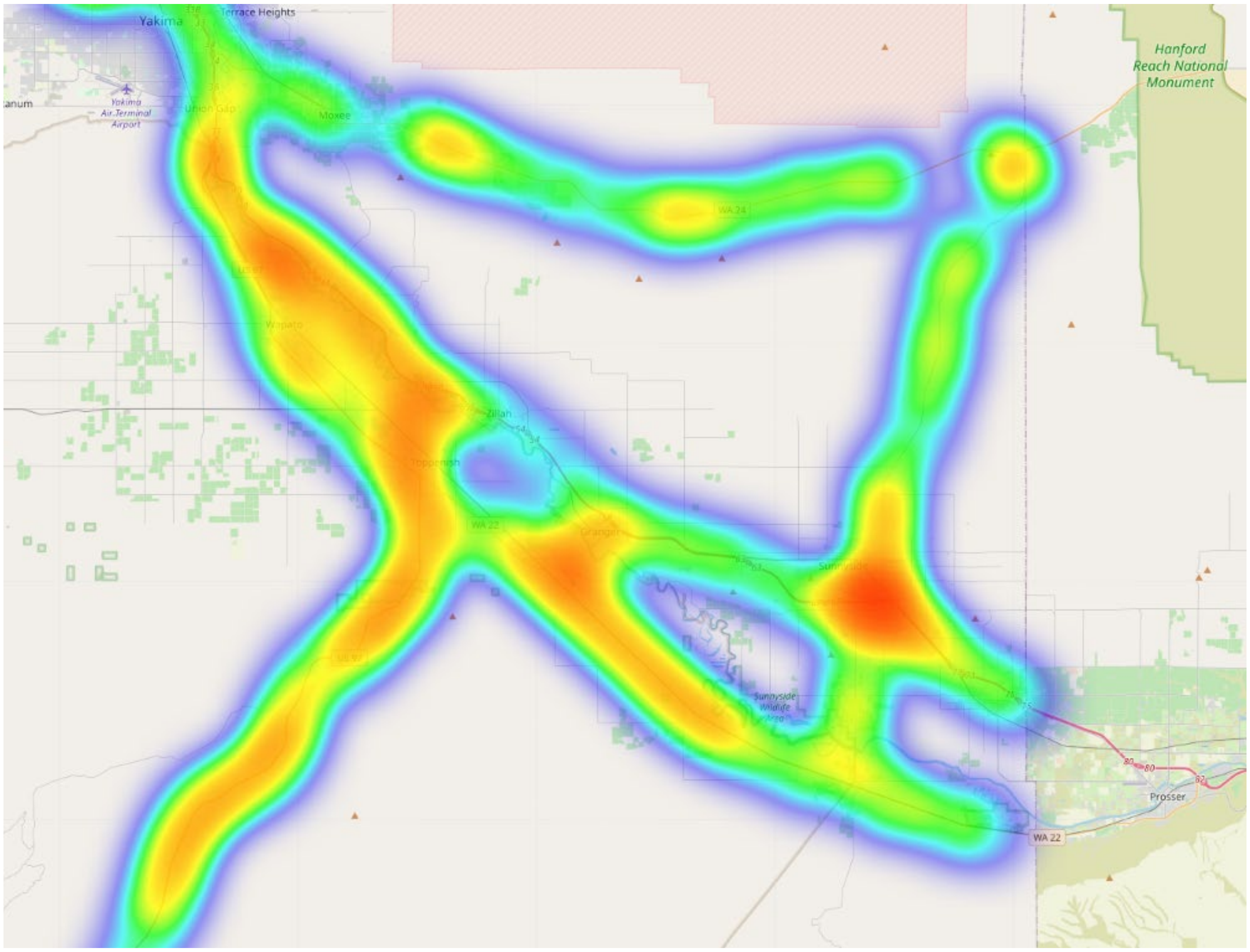

Figure 16 Expanded Crash Heatmap

The safety Performance Regression function allows users to select either incident frequency or estimate crash mean visualization (Figure 17). Users can query data that will be used for calculating safety performance regression based on start date, start time, end date, end time, and route for both functions. The incident frequency visualization function allows users to investigate safety levels (Figure

18). The green color means the level of safety of A, while the red color means the level of safety of F. User can obtain the level of safety at each road segment when the map is expanded (Figure 20).

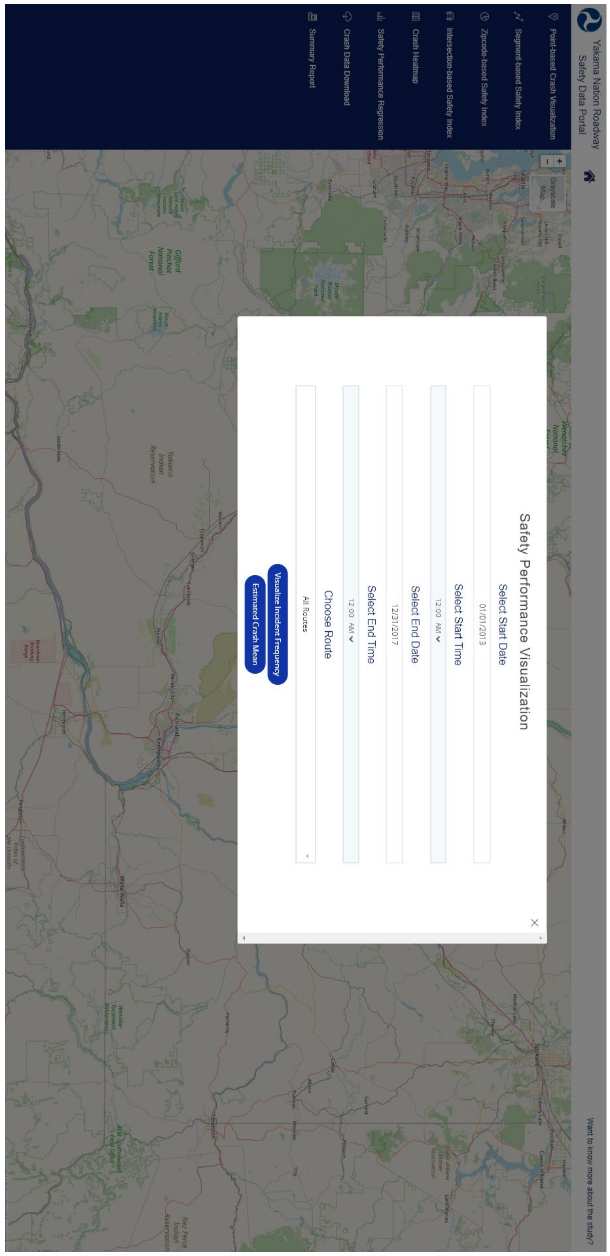

Figure 17 Safety Performance Visualization

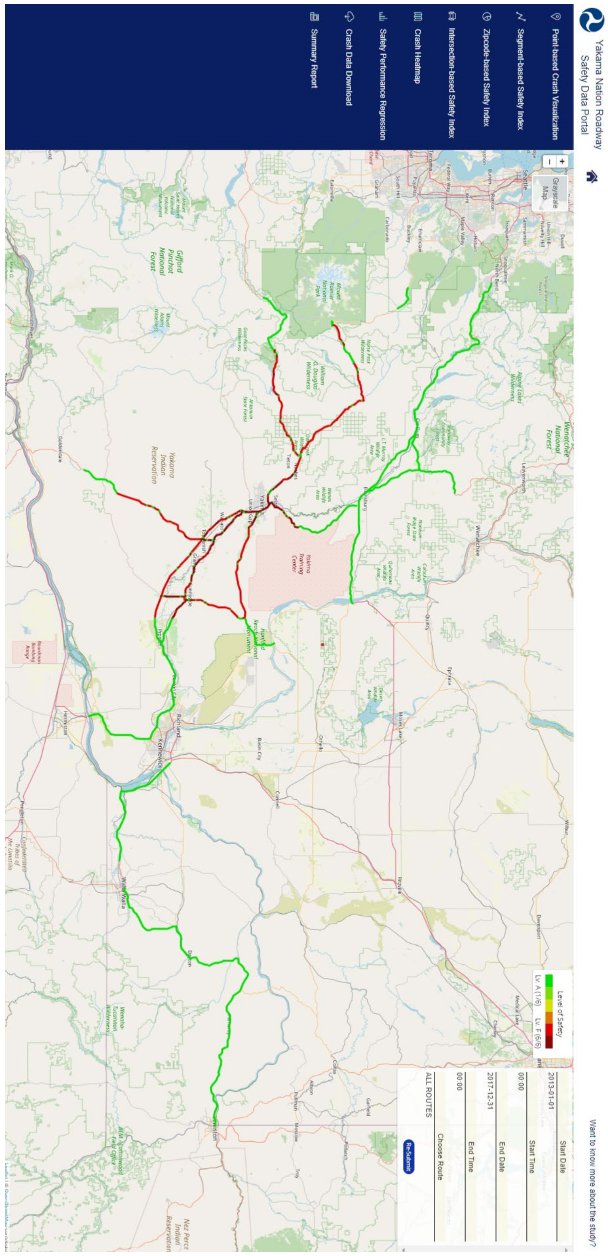

Figure 18 Incident Frequency Visualization

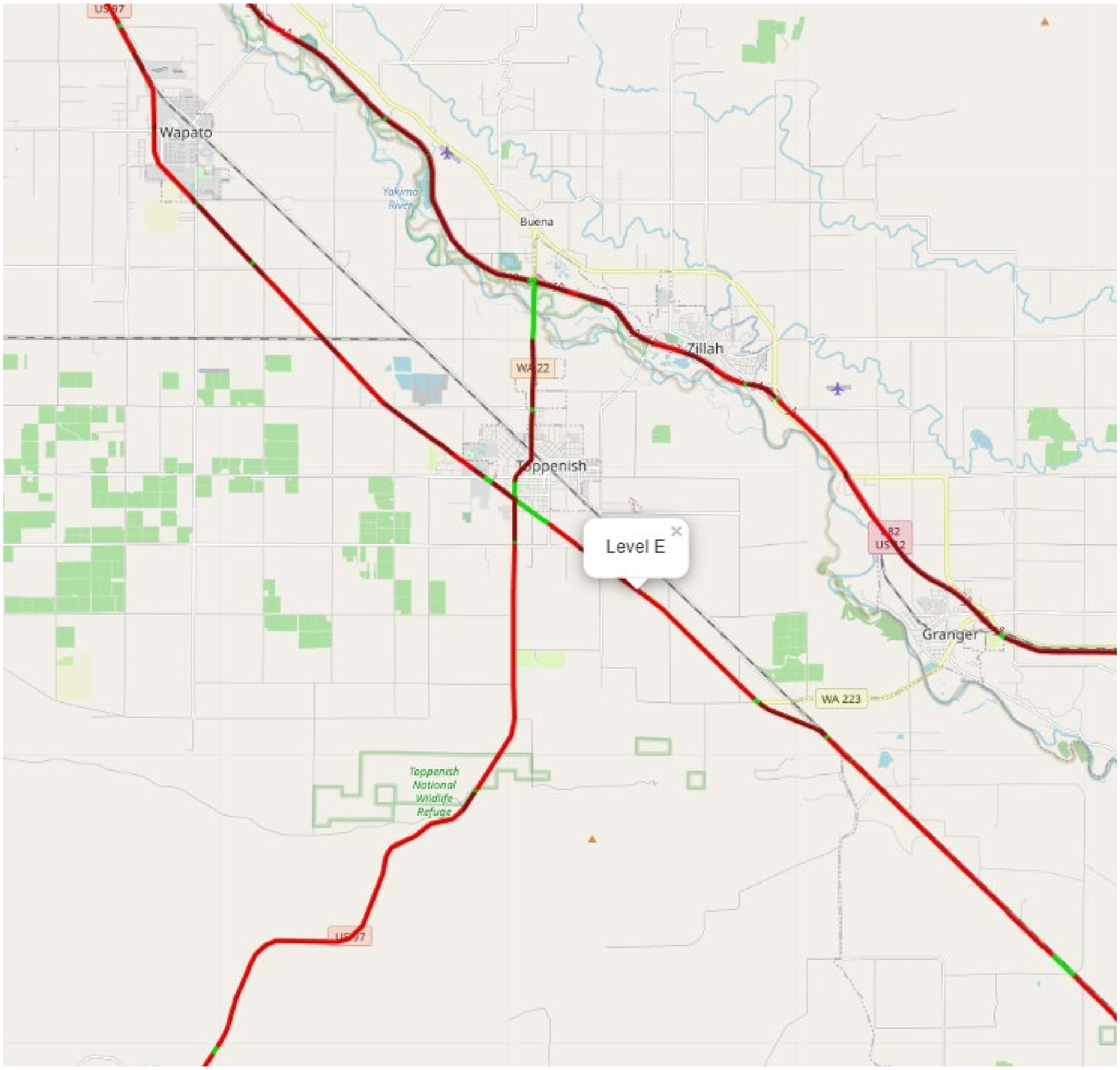

Figure 19 Level of Safety at Each Road Segment

The estimated crash mean function shows the crash mean based on the regression model described in the Analytical Approach section (Figure 20). The color on each segment represents the level of safety;

green represents the level of A, and red represents the level of F. Users can obtain the level of safety at each road segment based on the estimated crash mean when the map is expanded (Figure 21).

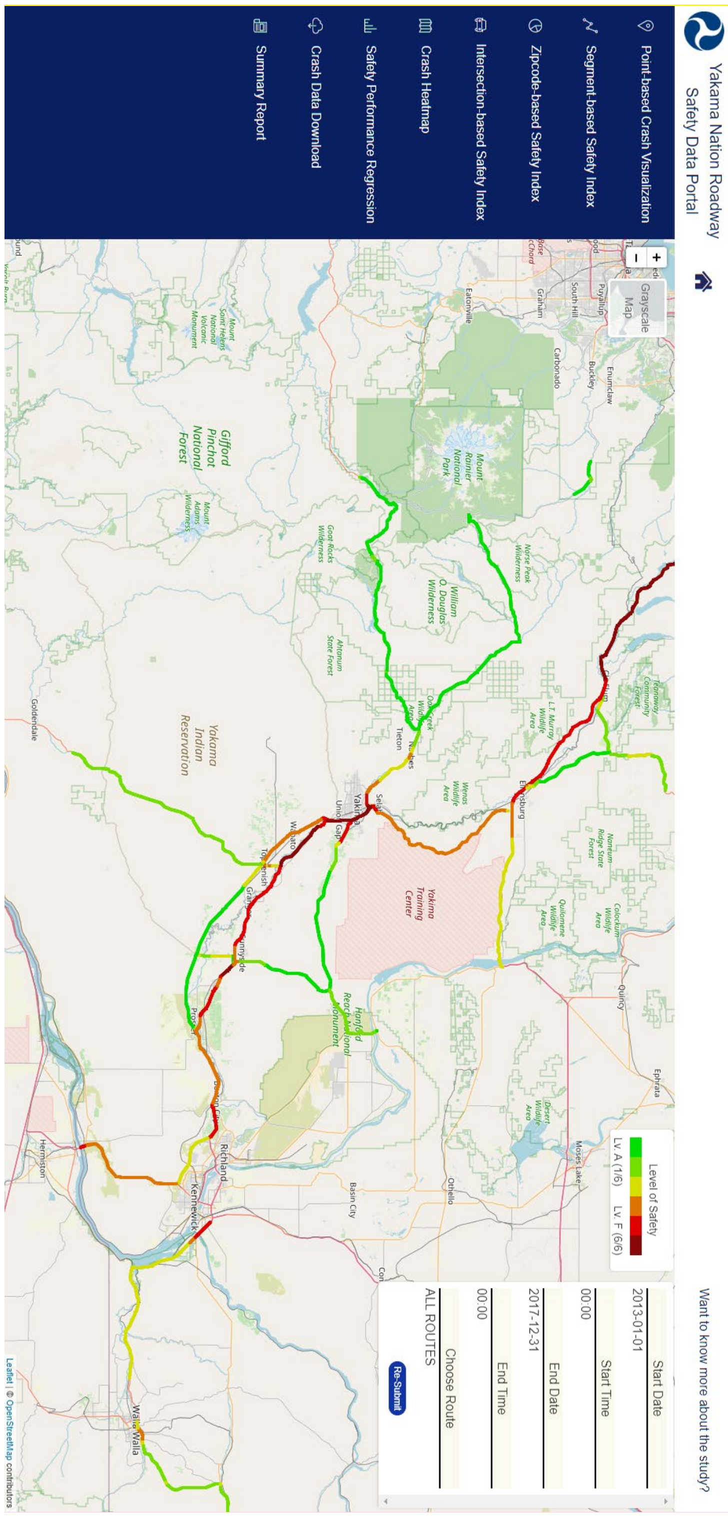

Figure 20 Safety Performance Regression

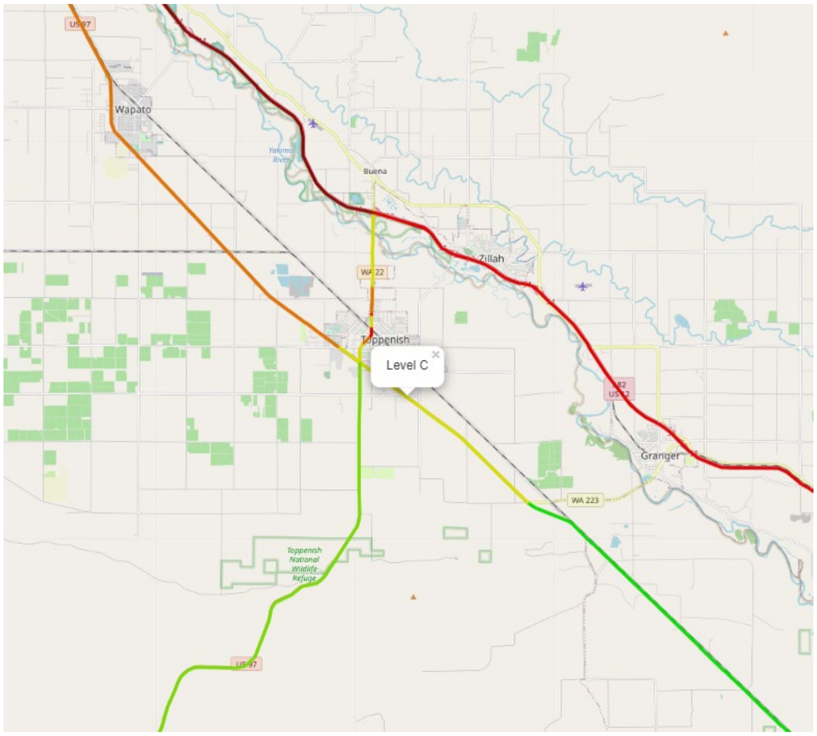

Figure 21 Level of Safety at Each Road Segment Based on Estimated Crash Mean

Crash Data Download allows users to query and download crash data (Figure 22). The downloaded data format is in an excel file. The data consists of accident type, accident year, latitude, longitude, severity class, and roadway class (Figure 23).

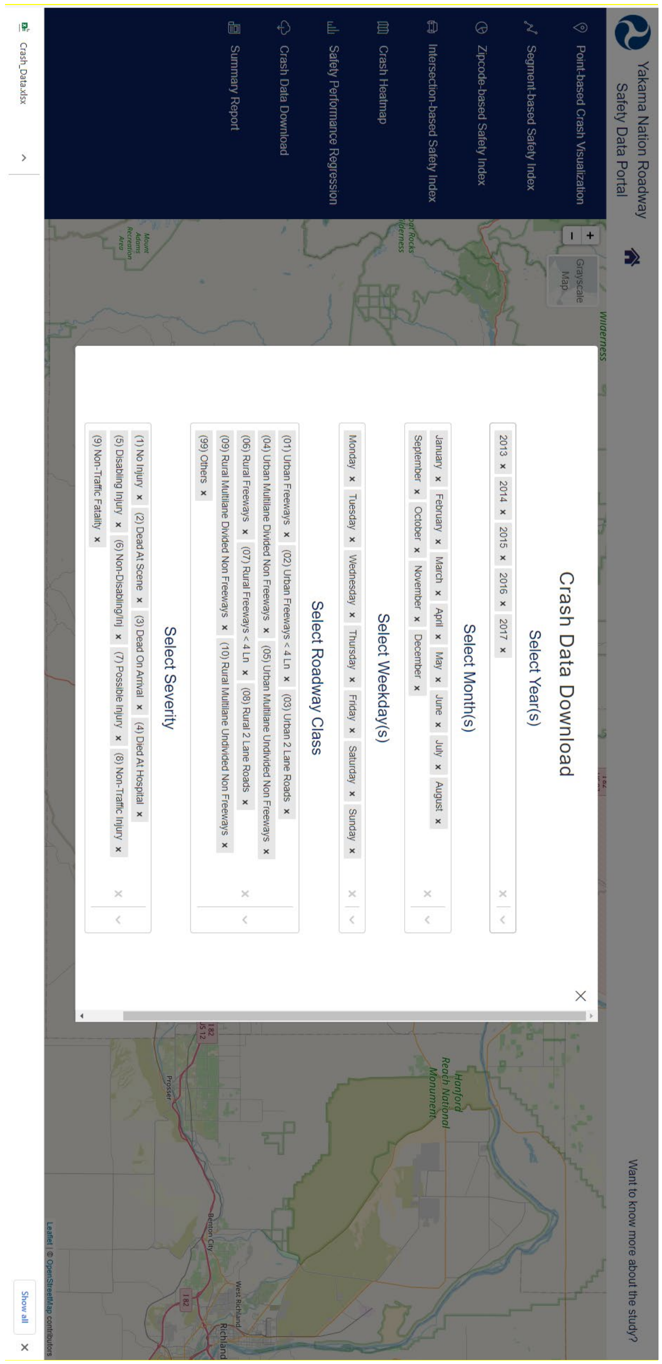

Figure 22 Crash Data Download Application

	A	B	C	D	E	F
1	ACCTYPE	ACCYR	LATITUDE	LONGITUD	SEVERITY	rodwycls
2	34	2013	46.63728	-121.393	1	8
3	34	2013	46.63728	-121.393	1	8
4	6	2013	46.63728	-121.393	1	8
5	34	2013	46.65214	-121.362	1	8
6	33	2013	46.65214	-121.362	1	8
7	34	2013	46.65214	-121.362	1	8
8	33	2013	46.65214	-121.362	1	8
9	33	2013	46.65214	-121.362	1	8
10	4	2013	46.65567	-121.306	1	8
11	33	2013	46.65567	-121.306	2	8
12	34	2013	46.64476	-121.271	6	8
13	34	2013	46.64476	-121.271	5	8
14	33	2013	46.64476	-121.271	1	8
15	50	2013	46.64476	-121.271	1	8
16	33	2013	46.64335	-121.231	1	8
17	60	2013	46.66527	-121.124	6	8
18	62	2013	46.66527	-121.124	6	8
19	34	2013	46.67287	-121.085	1	8
20	60	2013	46.67448	-121.046	1	8

Figure 23 Sample of Downloaded Data

The summary Report function generates three summary charts for selected years (Figure 24). First, the pie chart shows the proportion of crash severity. Second, the area chart shows crash counts by severity and year. This chart allows users to observe the year's trend. Lastly, the bar chart shows crash counts by severity and month. This chart lets the user monitor crashes in monthly aggregate based on severity class.

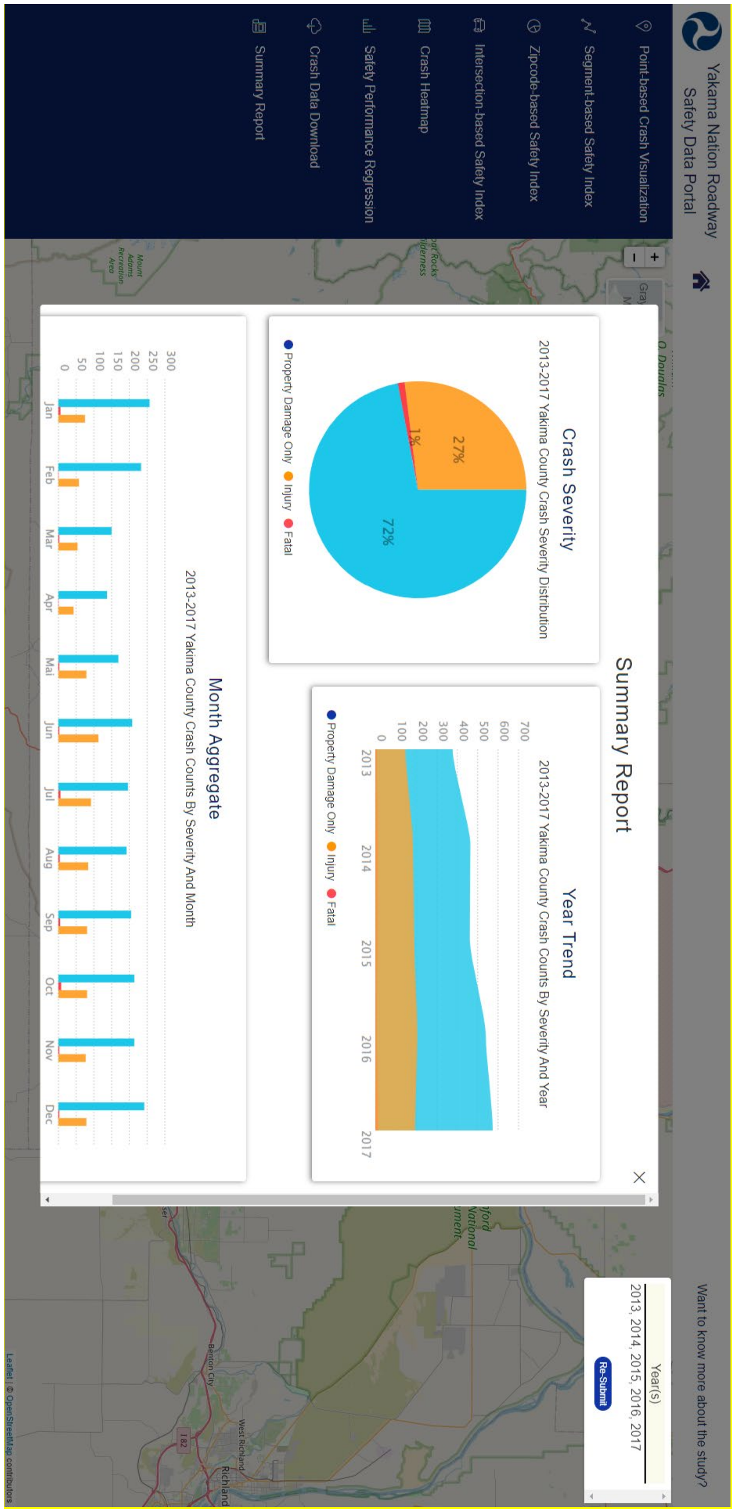

Figure 24 Crash Summary Reporting

4.3. Supporting Documentation and Information

This safety tool was presented to various agencies and conferences such as the Yakama Nation Tribal Council, the Yakama Nation Traffic Safety Committee Meetings, the 2021 National Transportation in Indian Country Conference (NTICC), the 2021 Highway Safety Information System (HSIS) Annual Liaison Meeting, and the 2022 TRB Annual Meeting.

The tool is available as the Yakama Nation Safety Data Portal on an open-source GitHub repository (Figure 25). The supporting documentation can be found at https://github.com/AI-Group-STAR-Lab-UW/yakama-nation-roadway-safety-data-portal. Besides details of data collection and features of the safety tool, this link also guides setting up Express-based client development and deployment environments for Yakama Nation Roadway Safety Data Portal API Server and guidance for setting up React-based client development and deployment environments for Yakama Nation Roadway Safety Data Portal.

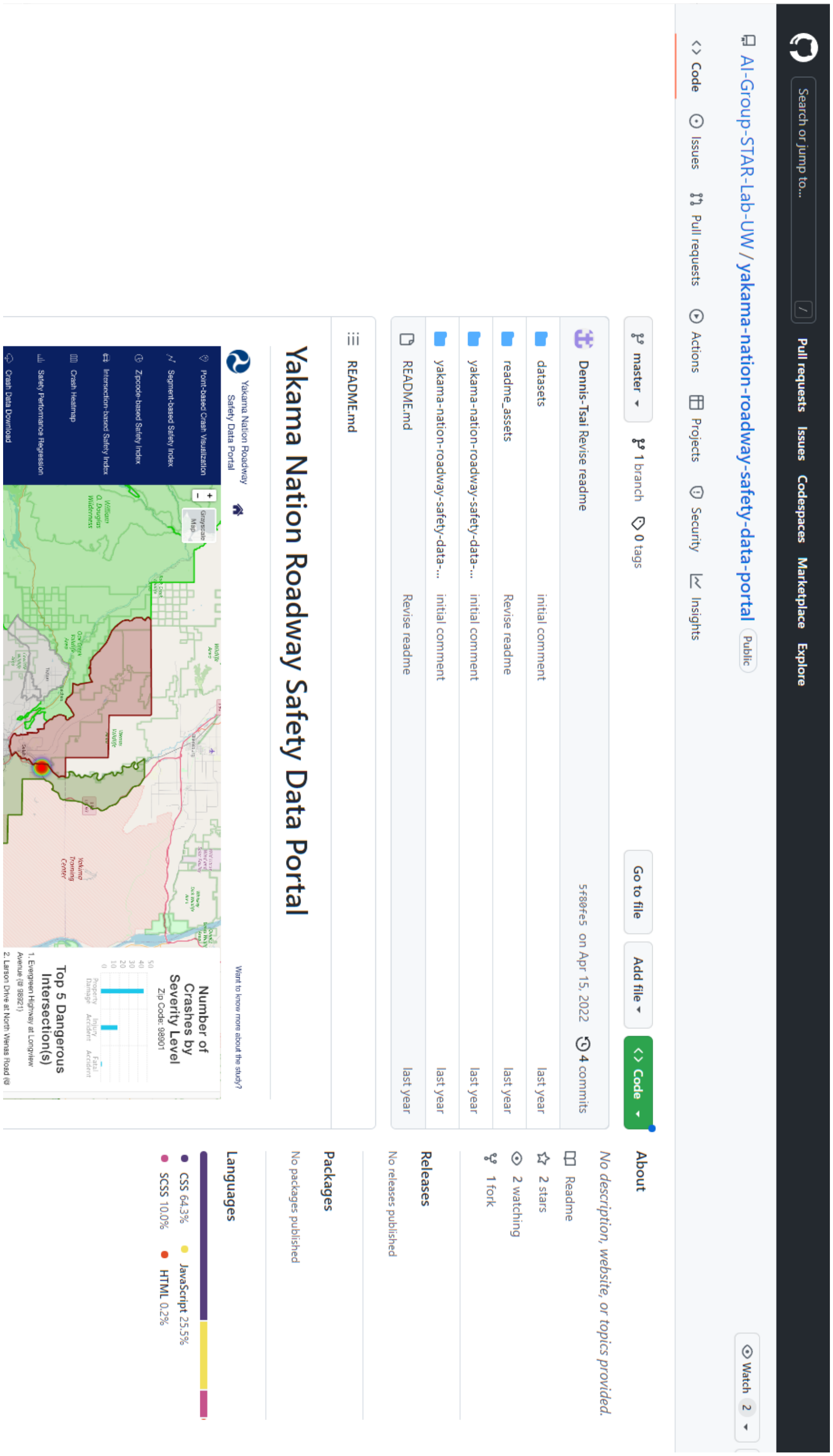

Figure 25 The Yakama Nation Safety Data Portal on an Open-source

4.4. Training on Safety Data Collection, Management, and Analytics

Besides the supporting documentation, the team provided training for engineers and planners from the Yakama Nation Department of Natural Resources (DNR) Engineering team on May 19, 2022, at the STAR Lab at the University of Washington. The training covered details on the safety data portal and Mobile Unit for Sensing Traffic or MUST sensors. The safety data portal session described safety data collection to its sources and attributes, including portal management. Machine learning approaches for traffic safety analysis were described to understand how the portal predicts pedestrian fatalities in rural areas of Washington State. In addition, the teams have discussed the challenge and future work.

Figure 26 Training on Safety Data Collection, Management, and Analytics for Yakama Nation DNR

CHAPTER 5. CONCLUSION AND FUTURE WORK

5.1. Addresses the gaps

The developed safety data, such as a comprehensive roadway safety tool, served as a multi-source safety data management platform for Yakama Nation transportation professionals. The safety tool enables efficient data management and includes various analytical and visualization functions, such as crash data visualization, hotspot identification, and network screening based on crash type, frequency, severity, estimated risks, reporting, and data download. The actionable insights gained from data analysis results produced by the advanced analytical techniques coupled with powerful visualization can assist the decision-maker in making high-quality decisions. Hence, Yakama Nation can allocate its funding based on the analytical results. Moreover, the safety tool can also be used to prioritize high-risk roadway segments. Lastly, suitable countermeasures for each roadway segment can be selected based on the analysis of risk factors.

5.2. Lesson learned and future work

The STAR Lab research team has collaborated with Yakama Nation DNR Engineering team frequently to address their needs and adjusted the platform to be in accordance with the feedback to make sure that the analytical and visualization functions were tailored to the specific needs of Yakama Nation's transportation professionals. The feedback and suggestions on safety tool development were also collected through periodic updates to key constituencies, such as the Yakama Nation Traffic Safety Committee.

Since Yakama Nation has mixed jurisdictions, the multi-source safety datasets were collected and managed by different agencies in different jurisdictions, such as federal, state, counties, and cities. As a result, obtaining interesting datasets, such as the citation records from the tribal police department, was regarded as a challenge in interpreting and managing data to be compatible with this research.

The traffic engineers and planners of the Yakama Nation DNR Engineering team are the targeted user of this safety tool. Yakama Nation will be responsible for the maintenance of the tool, whereas the STAR Lab will provide technical assistance when requested. Moreover, Yakama Nation DNR Engineering team will continue the efforts of incorporating more datasets into the safety data tool. This process is advised in the STAR Lab's user guide and technical assistance.

However, there is a gap that has been identified in this research. This safety tool was developed exclusively for Yakama Nation. In the future, the researcher should include the data management and analytical methods, as well as the safety tool developed in this research to other rural and tribal agencies that are encountering similar traffic safety issues.

CHAPTER 6. REFERENCES

[1] Wang, Y., Weibin, Z., Henrickson, K., Ke, R., and Cui, Z. (2016). Digital roadway interactive visualization and evaluation network applications to WSDOT operational data usage. No. WA-RD 854.1. Washington (State). Dept. of Transportation.

[2] Xiao, S., Liu, X., and Wang, Y. (2015). Data-driven geospatial-enabled transportation platform for freeway performance analysis. IEEE Intelligent Transportation. Systems Magazine 7, no. 2.

[3] NHTSA's National Center for Statistics and Analysis. (2020). 2018 Data Traffic Safety Facts. U.S. Department of Transportation. No. HS 812 957. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812957

[4] Donofrio, J. (2022). Traffic fatalities increasing across Washington, in Yakima County, data show. Yakima Herald-Republic. https://www.yakimaherald.com/news/local/traffic-fatalities-increasing-across-washington-in-yakima-county-data-show/article_e02912b6-7d75-11ed-8556-cbd81b641360.html

[5] NHTSA. (2021). Speeding and Agressive Prevention. U.S. Department of Transportation. https://www.nhtsa.gov/risky-driving/speeding#resources

[6] Whitlock, E., Casebier, M., and Yazzie, M. (2018). Tribal Safety Plans: Development-Data-Implementation. 2018 Tribal Planning & Western Planning Conference. https://static1.squarespace.com/static/570155d17c65e42063eed1f5/t/5b8ed61521c67cfacc1dfcdf/1536087587118/Tribal+Safety+Plans_Casebier%2BWhitlock_Tribal%2BWP+Conf_8-6-18.pdf

[7] GHSA. (2022). America's Rural Roads: Beautiful and Deadly. State Farm. https://www.ghsa.org/sites/default/files/2022-09/America%E2%80%99s%20Rural%20Roads%20-%20Beautiful%20and%20Deadly%20FNL.pdf

[8] Pollack, K. M., Gielen, A. C., Mohd Ismail, M. N., Mitzner, M., Wu, M., & Links, J. M. (2014). Investigating and improving pedestrian safety in an urban environment. Injury epidemiology, 1(1), 11. https://doi.org/10.1186/2197-1714-1-11

[9] Stoker, P., Garfinkel-Castro, A., Khayesi , M., Odero, W., Mwangi, M. N., Peden, M., & Ewing , R. (2015). Pedestrian Safety and the Built Environment: A Review of the Risk Factors. Journal of Planning Literature 2015, Vol. 30(4) 377-392.

[10] Per E Gårder. (2014). The impact of speed and other variables on pedestrian safety in Maine, Accident Analysis & Prevention, Volume 36, Issue 4, 2004, Pages 533-542. https://doi.org/10.1016/S0001-4575(03)00059-9.

[11] Martin, A. (2006). Factors influencing pedestrian safety: a literature review (No. PPR241). Wokingham, Berks: TRL.

[12] Zhai, X., Huang, H., Sze, N. N., Song, Z., & Hon, K. K. (2019). Diagnostic analysis of the effects of weather condition on pedestrian crash severity. Accident Analysis & Prevention, 122, 318-324.

[13] Lefler, D. E., & Gabler, H. C. (2004). The fatality and injury risk of light truck impacts with pedestrians in the United States. Accident Analysis & Prevention, 36(2), 295-304.

[14] Qi, Y., and Guoguo, A. (2017). Pedestrian Safety Under Permissive Left-turn Signal Control. International Journal of Transportation Science and Technology, No. 6 (2017) 53–62.

[15] Vachal, K. (2018). Risk Factors for Pedestrian Injury on Tribal Lands. Journal of Transport & Health, 9, S23.

[16] Naumann, R.B., and L.F. Beck. (2010). Motor Vehicle Traffic-Related Pedestrian Deaths: United States, 2001–2010. Centers for Disease Control and Prevention Morbidity and Mortality Weekly Report, Vol. 62, No. 15, 2013, pp. 277–282. http://www.cdc.gov/mmwr/pdf/other/su6001.pdf.

[17] Iragavarapu, V., Carlson, P., & Schertz, G. (2015). Review of Tribal Transportation Safety. Transportation Research Record, 2531(1), 153-160.

[18] Davis, G. A. (1998). Method for estimating effect of traffic volume and speed on pedestrian safety for residential streets. Transportation Research Record, 1636(1), 110-115.

[19] Zajac, S. S., & Ivan, J. N. (2003). Factors influencing injury severity of motor vehicle–crossing pedestrian crashes in rural Connecticut. Accident Analysis & Prevention, 35(3), 369-379.

[20] J. W. Hall, J. D. Brogan, and M. Kondreddi. (2004). Pedestrian Safety on Rural Highways (FHWA-SA-04-008). U.S. Department of Transportation. https://www.pedbikeinfo.org/cms/downloads/Ped_Safety_RuralHighways.pdf

[21] Baireddy, R., Zhou, H., & Jalayer, M. (2018). Multiple correspondence analysis of pedestrian crashes in rural Illinois. Transportation research record, 2672(38), 116-127.

[22] Tao, W., Aghaabbasi, M., Ali, M., Almaliki, A. H., Zainol, R., Almaliki, A. A., & Hussein, E. E. (2022). An Advanced Machine Learning Approach to Predicting Pedestrian Fatality Caused by Road Crashes: A Step toward Sustainable Pedestrian Safety. Sustainability, 14(4), 2436.

[23] Rella Riccardi, M., Galante, F., Scarano, A., & Montella, A. (2022). Econometric and Machine Learning Methods to Identify Pedestrian Crash Patterns. Sustainability, 14(22), 15471.

[24] Das, S., Le, M., & Dai, B. (2020). Application of machine learning tools in classifying pedestrian crash types: A case study. Transportation safety and environment, 2(2), 106-119.

[25] FHWA. (2010). Highway Safety Manual. FHWA Highway Safety Programs. https://highways.dot.gov/safety/pedestrian-bicyclist/safety-tools/volume-2chapter-12-highway-safety-manual

[26] Office of Safety of FHWA. (2022). Tools and technology – safety. FHWA .https://safety.fhwa.dot.gov/tools/

[27] FHWA. (2023). Interactive Highway Safety Design Model (IHSDM): Overview. Tools and technology – safety. https://highways.dot.gov/research/safety/interactive-highway-safety-design-model/interactive-highway-safety-design-model-ihsdm-overview

[28] AASHTO. (2023). AASHTOWare – Safety Overview. https://www.aashtoware.org/products/safety/safety-overview/

[29] Center for Transportation Safety. (2023). List of Safety Tools. Texas A&M Transportation Institute. https://cts.tti.tamu.edu/safetytools/

[30] Texas Department of Transportation. (2023). Design tools and training. https://www.txdot.gov/business/resources/design-tools-training.html

[31] Roadsafe LLC. (2012). Overview of RSAP. https://www.rsap.roadsafellc.com/

[32] Pennsylvania Department of Transport. (2023). Pennsylvania Crash Information Tool. https://crashinfo.penndot.gov/PCIT/welcome.html