# Comprehensive Roadway Safety Data Visualization and Evaluation Platform for Yakama Nation

# Final Report

Prepared By:

Department of Natural Resources
Confederated Tribes and Bands of the Yakama Nation
401 Buster Rd
Toppenish, WA 98948

Department of Civil and Environmental Engineering
University of Washington, Box 352700
Seattle, WA 98195

Prepared for

US Department of Transportation
Safety Data Initiative

April 2022

# TECHNICAL REPORT DOCUMENTATION PAGE

| 1. Report No. | 2. Government Accession No. | 3. Recipient's Catalog No. |
|---|---|---|
| **4. Title and Subtitle**<br>Comprehensive Roadway Safety Data Visualization and Evaluation Platform for Yakama Nation | | **5. Report Date**<br>April 2022 |
| | | **6. Performing Organization Code** |
| **7. Author(s)**<br>Yinhai Wang, Wei Sun, Hollyanna Littlebull, Meng-Ju Tsai, Sam Ricord, Yifan Zhuang | | **8. Performing Organization Report No.** |
| **9. Performing Organization Name and Address**<br>Department of Civil and Environmental Engineering, University of Washington, Box 352700, Seattle, WA 98195<br><br>Department of Natural Resources, Confederated Tribes and Bands of the Yakama Nation, 401 Buster Rd, Toppenish, WA 98948 | | **10. Work Unit No.** |
| | | **11. Contract or Grant No.**<br>69A34520501020620 |
| **12. Sponsoring Agency Name and Address**<br>US Department of Transportation<br>1200 New Jersey Ave, SE<br>Washington, DC 20590 | | **13. Type of Report and Period Covered**<br>Final Report (October 2020 - April 2022) |
| | | **14. Sponsoring Agency Code**<br>OST-Policy |
| **15. Supplementary Notes**<br>Conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration. | | |

**16. Abstract**

Funded by the US Department of Transportation's Safety Data Initiative (SDI), the Yakama Nation Department of Natural Resources (DNR) Engineering collaborated with the Smart Transportation Application & Research Laboratory (STAR Lab) at University of Washington to develop a comprehensive roadway safety data visualization and evaluation platform to support Yakama Nation government's policy and decision making. Safety related datasets, including collision records (collision, vehicle, occupant, pedestrian), roadway characteristics (roadlog, curve and grade, ramp, traffic information, special-use lane, etc.) were collected, cleaned, and stored to the database. The multi-source database management system supports data collection, quality control, integration, database management, visualization, analytical results, etc. Besides efficient data management, the safety tool includes analytical and visualization functions, including crash data visualization, hotspot identification, and screening the network based on crash type, frequency, severity and estimated risks, reporting, 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. | No restrictions. |

| 19. Security Classif. (of this report) | 20. Security Classif. (of this page) | 21. No. of Pages | 22. Price |
|---|---|---|---|
| Unclassified | Unclassified | 10 | Refers to the price of the report. Leave blank unless applicable. |

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

# 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.

# Executive Summary

There is a great disparity concerning transportation safety in rural areas of the country. The Confederated Tribes and Bands of the Yakama Nation is one of the federally recognized tribes in Washington State. This project aims to address the disparity for both rural and Native American populations through the development of a web-based comprehensive roadway safety tool. Funded by the US Department of Transportation's Safety Data Initiative (SDI), 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 to support Yakama Nation government's policy and decision making.

The development of the safety tool was based on a transportation multi-source data management, fusion, analysis, and visualization platform that the STAR Lab team has developed, called the Digital Roadway Interactive Visualization and Evaluation Network (DRIVE Net) [1] [2]. Safety-related datasets, including collision records (collision, vehicle, occupant, and pedestrian), and roadway characteristics (roadlog, curve and grade, ramp, traffic information, special-use lane, etc.) were collected, cleaned, and stored in the database. The multi-source database management system supports data collection, quality control, integration, database management, visualization, analytical results, etc. Besides efficient data management, the safety tool includes analytical and visualization functions, including crash data visualization, hotspot identification, and screening of the network based on crash type, frequency, severity, estimated risks, reporting, and safety data download.

These advanced analytical techniques coupled with powerful visualization can assist the decision makers' by transferring data analysis results into actionable insights. With the help of the safety tool, Yakama Nation can make effective use of the resources utilizing the safety tool to prioritize the high-risk roadway segments, and apply countermeasures based on the analysis of risk factors of the selected roadway sections. Besides, since many state and local governments are experiencing safety issues with rural roadways similar to Yakama Nation, the analytical methods and safety tools developed in this project can be modified to support the decision-making of these governments as well.

# Background

## Roadway safety problem statement

Within the Yakama Nation reservation, there are approximately 1,200 miles of public roads. Most of the roads are in rural agricultural settings and crashes happen every day on these roadways. Of the land governed by Tribal Governments, Yakama Nation has both the highest pedestrian and vehicle fatality rates. Yakama Nation DNR Engineering program has been deeply concerned by these crashes and high fatality rates and is determined to work on solutions addressing the safety issues and freeing the reservation of serious injuries and fatal accidents.

Particularly, risk factors such as roadway geometrics (harp 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.) have been recognized that could lead to roadway crashes. Besides, according to the Safety Management Plan developed by Yakama Nation DNR Engineering program (2018), driving under the influence, speeding, unrestrained driver/occupant, and distracted driver are among the top fatality and serious injury priorities. Speeding has been documented to increase the likelihood and severity of both vehicle and pedestrian collisions. Yakama nation has a high level of speeding, mostly due to a significant lack of enforcement. While the factors above have been recognized as contributing factors to the roadway crashes within the Yakama Nation reservation, it is not clear to government officials how to turn these contributing factors into actionable insights, such as how to prioritize the safety funding on specific roadways/intersections to improve the traffic safety conditions most effectively.

Therefore, the research team identifies and defines the roadway safety problem as rural roadway crashes (frequency and severity) analysis of various contributing factors, and the identification of high-risk locations for vehicles and pedestrians to support decision making.

## Research questions

After identification of the roadway safety problem of Yakama Nation, the following research questions were raised:

- What are the contributing factors to the roadway crashes in Yakama Nation?
- How are the contributing factors affecting the crash frequency and severity?
- Of all the roads within the Yakama Nation roadway network, what are the priorities to improve the traffic safety conditions given limited funding and resources?
- What are the contributing factors of road safety inequality in rural areas, especially for tribal communities? For instance, lack of traffic and road safety data for planning and operations.

The effective use of the safety data to develop data-driven solutions that turn into actionable insights for decision makers' is critical to Yakama Nation. To solve the defined safety problem, the key is quantitative modeling of roadway crashes and their contributing factors using safety-related data, which will address the first two research questions. A thorough analysis and understanding of the contributing factors to crashes also help decision makers' select effective countermeasures. The third research question identifies the knowledge gap of decision makers', which is the gap between safety data and actionable insights. To assist transportation practitioners in making efficient and effective policies and decisions to invest and apply countermeasures, it is important to provide them with analytical results combined with visualization functions that clearly show the contributing factors, their impacts on roadway crashes, and the ranking of roadway segments regarding vehicles as well as pedestrian safety.

## Identified gaps

Yakama Nation has been collaborating with federal, state, county, and city transportation agencies to work on the development of the Tribal Traffic Safety Plan (TTSP). In 2018, the initial TTSP, Safety Management Plan, was developed based on Federal Highway

Administration's (FHWA) and Washington State's Strategic Highway Safety Plans (SHSP). While the safety plan laid a great foundation for building toward a policy and decision-making process supported by data-driven safety analysis, the current practice for making safety-related decisions based on basic crash data analysis still needs improvement. In particular, the research team identified the following areas to work on, which were the main focuses of this research.

- *Data management system for multi-source data*. There lacked a database system that manages the safety-related datasets so that traffic engineers and planners can have access to and utilize the road safety data for the greater good of Yakama Nation. The Yakama Nation transportation practitioners needed a database system that can manage and integrate the safety-related datasets, which is also the basis for performing advanced analysis and data visualization.
- *Advanced analytical methods for roadway crash modeling and prediction*. In addition to managing the safety-related datasets, there was a need to utilize advanced analytical methods to analyze the roadway crashes and develop accurate crash models that can generate predictive insights.
- *Efficient platform for safety data visualization*. Besides performing advanced modeling and analysis on the crash data, it is also critical to have a platform that can provide efficient visualization of the data and the analytical results.

# Methodological Approach

## Data

The required data to accomplish the vision of the safety tool are crash data, roadway characteristics data, and weather data. The STAR Lab team had worked with Yakama Nation DNR Engineering team to check the available safety-related data sources, and collect, store and manage the collected data into the developed database management system. Specifically, the data preparation included the following steps.

- *Data collection.* The UW research team had worked with the Yakama Nation DNR Engineering team for safety-related datasets including the Highway Safety Information System (HSIS) dataset, statewide weather station data by the Washington State Department of Transportation (WSDOT), WSDOT crash data portal, and Washington state GIS roadway network data. These datasets were then cleaned and stored in the database.
- *Data integration*. Multi-source datasets were integrated into the database system, for redundant data across data sources, if all attribute values are duplicated, then the entire data was deleted. For some cases where redundant attributes were used, correlation analysis and monitoring were performed to determine whether the original data can be restored.
- *Data management.* The datasets used in this research were managed in the database system in MS SQL server and PostgreSQL databases.
- *Quality control*. The data management system applied advanced data integration and imputation algorithms for data quality control. As abnormal data is common in roadway safety datasets, the data cleaning procedure was implemented to perform redundant data processing, missing data processing, noise data processing, etc.

## Analytic approach

In addition to manage and visualize the collision records, it is vital to accurately model the roadway crashes to transfer data into actionable insights. This project explored and investigated several crash modeling methods and incorporated the following method into the safety data tool for visualization to conveniently identify high-risk roadway/intersections considering both modeling accuracy and feasibility/efficiency to implement in a web-based environment.

### Road segment performance estimation

The definition of road segments is shown in Figure 1, where A represents intersections and B represents road segments. The road segment begins at the center of an intersection and ends at either the adjacent intersection center or a change from one homogeneous segment to another.

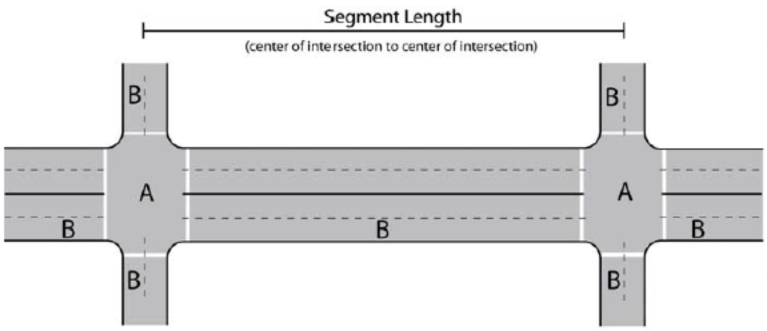


Figure 1. Road Segment Definition based on Highway Safety Manual (HSM) [1]

The fundamental prediction model is focusing on single-vehicle driveway collisions. The prediction model is defined in Equation (1), where $a$ and $b$ are coefficients. $AADT$ implies the average daily road segment volume, and its unit is vehicles per day. $L$ is the road segment length as shown in Figure 1, and its unit is miles.

$$\boldsymbol{N_{brsv} = e^{a+b\cdot\ln(AADT)+\ln(L)}} \quad \textbf{(1)}$$

The model is based on Poisson-gamma model $Y_{it}|\theta_{it}\sim Po(\theta_{it})$, where $Y_{it}$ indicates the number of crashes in the segment $i$ and at the time stamp $t$. $\theta_{it}$ is the mean of the Poisson and is assumed to be Poisson distributed and independent over all road segments and time stamps. $\theta_{it}$ is defined as $\theta_{it} = \mu_{it}\exp(\varepsilon_{it})$, where $\mu_{it}$ represents a function integrating all covariates. Generally, this function is written as $\mu_{it} = \exp(\beta_0 + \beta_1 X_{it1} + \cdots \beta_k X_{itk})$. $\beta$ represents unknown coefficients which should be regressed based on real-world data. $\varepsilon_{it}$ is an independent variance and indicates model errors.

Equation (2) expands Equation (1) to multiple-vehicle driveway-related collisions, which is the sum of the driveway-related collisions from different driveway types. In Equation (2), $n_j$ is the number of driveways within the target road segment of driveway type j. $N_j$ is the number of collisions related to driveway type j per year. $t$ is the coefficient for traffic volume adjustments.

$$\boldsymbol{N_{brdwy} = \sum_{all\ types} n_j \cdot N_j \cdot \left(\frac{AADT}{15000}\right)^t} \quad \textbf{(2)}$$

Several primary road segment types include major commercial driveways, minor commercial driveways, major industrial/institutional driveways, etc. There is only one type for each road segment type in this project. The original model was regressed using data from Minnesota and

Michigan states, including 4,255 road segments with a total length of 598.3 miles. This project validated the model using data from Washington state. Based on HSM Chapter 12, there are five primary CMFs applied to prediction models described in Equation (1) and Equation (2). After adding the CMFs [2], the crash prediction model is shown as Equation (3), where $N_{rs}$ implies the predicted number of total crashes in road segments after adding CMFs. $N_{br}$ represents this crash number for base conditions without considering these CMFs. $CMF_{ir}(i = 1, \dots, n)$ is CMF for various road segment features which are limited to primary CMFs mentioned above.

$$\boldsymbol{N_{rs} = N_{br} \cdot (CMF_{1r} \cdot CMF_{2r} \dots \cdot CMF_{nr})} \quad \textbf{(3)}$$

As a result, the final prediction model is shown in Equation (4). $N_i$ is the total number of predicted crashes in the road segment $i$. $L_i$ is the length of the road segment $i$. $AADT_i$ represents the average traffic volume. In this project, four significant CMFs are taken into account – lane width $W_L$, shoulder width $W_S$, median width $W_M$, and number of lanes $N_L$. $\beta$ represents unknown coefficients. In this project, these coefficients are calculated based on data collected in Washington state. The detailed explanations are described in Table 1.

$$\boldsymbol{N_i = \beta_o L_i (AADT_i)^{\beta_1} e^{\beta_2 W_L + \beta_3 W_S + \beta_4 W_M + \beta_5 N_L}} \quad \textbf{(4)}$$

Table 1. Variables of Prediction Model in Equation (4)

| **Variables** | **Descriptions** |
|---|---|
| $\boldsymbol{AADT}$ | Two-way average traffic volume |
| **Segment Length $\boldsymbol{L}$** | Length of homogenous segments |
| **Lane Width $\boldsymbol{W_L}$** | Average width of all lanes |
| **Shoulder Width $\boldsymbol{W_S}$** | 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 |

# Safety Tool

## Development approach

The development of the safety tool was based on a transportation multi-source data management, fusion, analysis, and visualization platform that the STAR Lab team has developed, called the Digital Roadway Interactive Visualization and Evaluation Network (DRIVE Net). DRIVE Net has various transportation data analysis and visualization functions to support decision making, 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, etc.

The safety tool for Yakama Nation can be regarded as an extension of the existing urban roadway safety analysis module on DRIVE Net. Similarly, the developed safety tool was able to integrate multi-source safety-related data, perform quality control on the data, support modeling and analysis, and provide visualization functions. The safety-related datasets collected were stored and managed by the multi-source database system in MS SQL Server and PostgreSQL databases. Data consistency and integrity were checked and ensured. Along with the development of the safety tool, a detailed user guide was developed for Yakama Nation that documented the implementation and maintenance guidelines of the tool.

## Analytical and visualization functions

To better assist Yakama Nation transportation practitioners for safety data analysis and decision making, the project developed the following analytical and visualization functions (Figure 2).

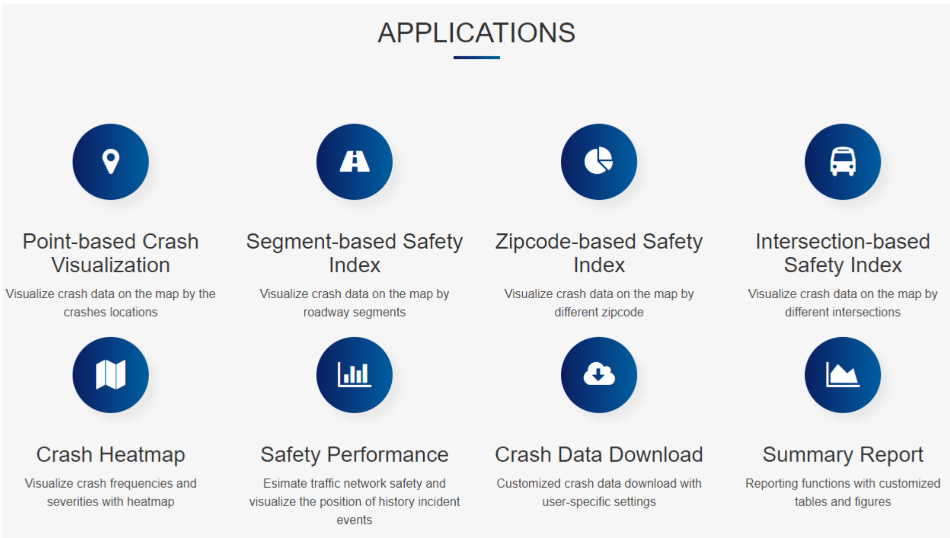


Figure 2. Functions of the web-based safety data tool

As shown in Figure 3, point-based crash visualization provides collision record visualization on the map by the location of the selected collision events based on the specific query settings (year, month, weekday, roadway class, severity, etc.).

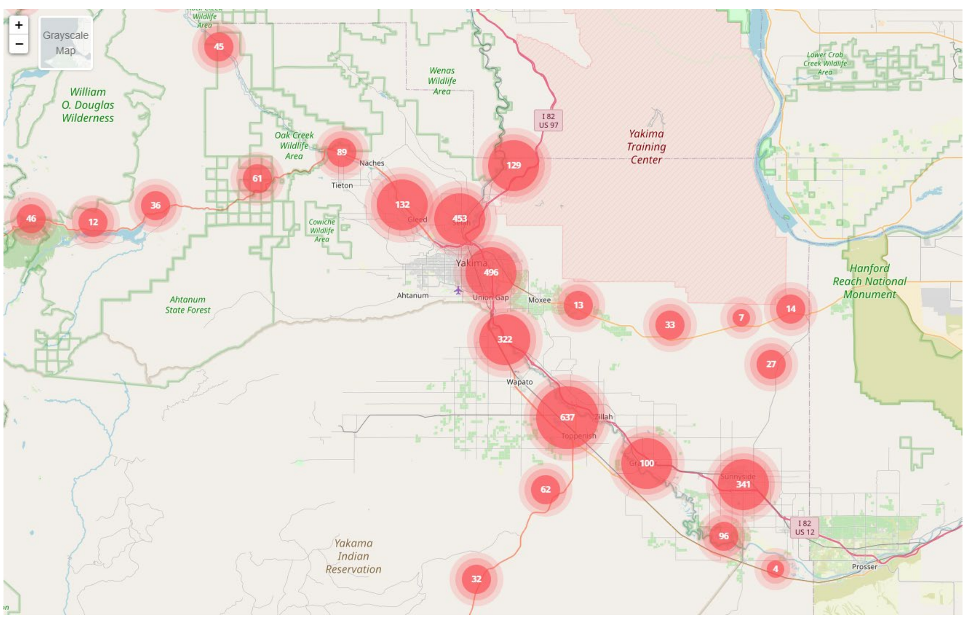


Figure 3. Point-based crash visualization

As shown in Figure 4, segment-based safety index visualization provides visualization of historical collision frequency or estimated frequency based on crash modeling analysis on the map by the segment of the selected collision events based on the specific query settings (year, month, weekday, roadway class, severity, etc.).

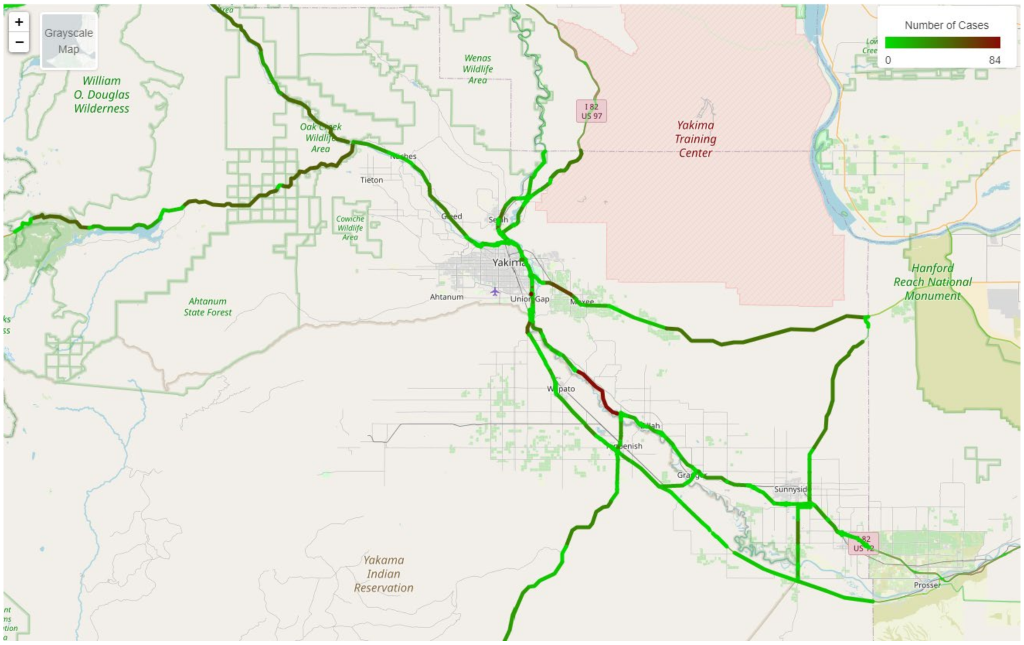


Figure 4. Segment-based safety index visualization

As shown in Figure 5, zip code-based crash visualization provides visualization of historical collision frequency on the map by the segment of the selected collision events based on the specific query settings (year, month, weekday, roadway class, severity, etc.).

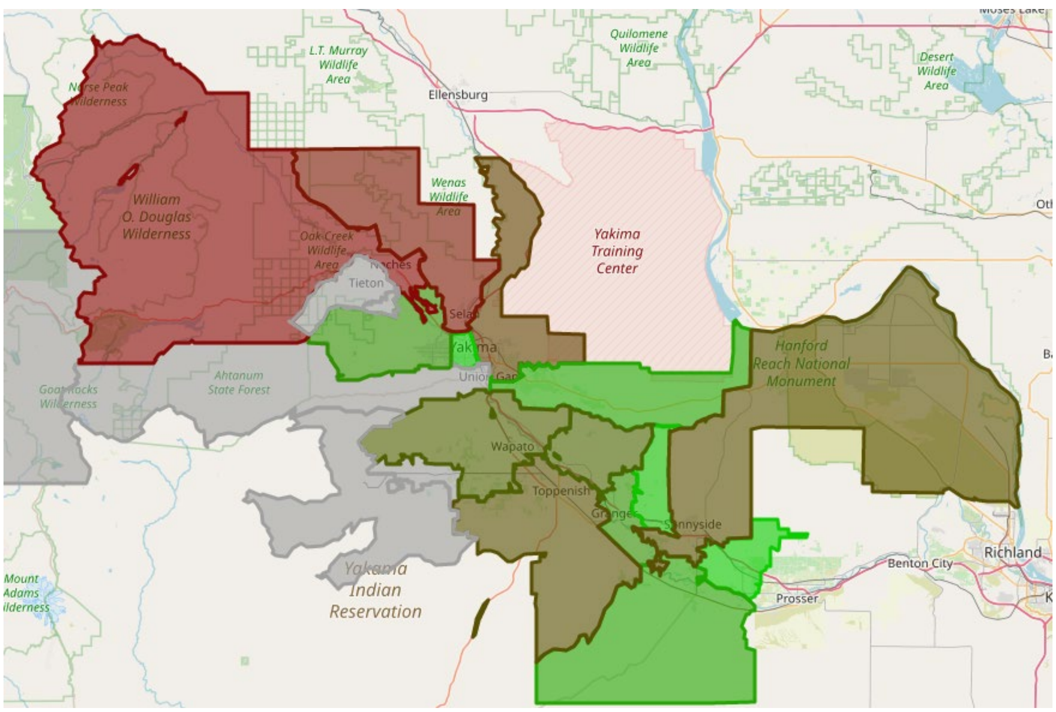


Figure 5. Zip code-based crash visualization

As shown in Figure 6, intersection-based crash visualization provides visualization of historical collision events that happened near a particular intersection on the map based on the specific query settings (year, month, weekday, roadway class, severity, etc.).

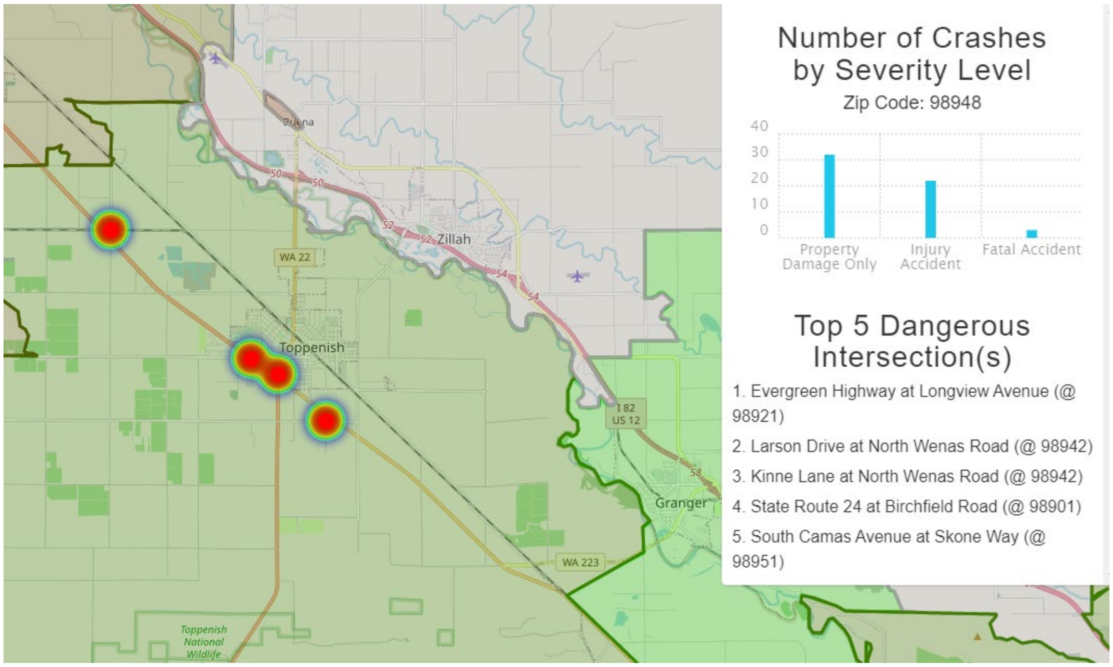


Figure 6. Intersection-based crash visualization

As shown in Figure 7, the safety tool has the capability to visualize the collision records in a heatmap.

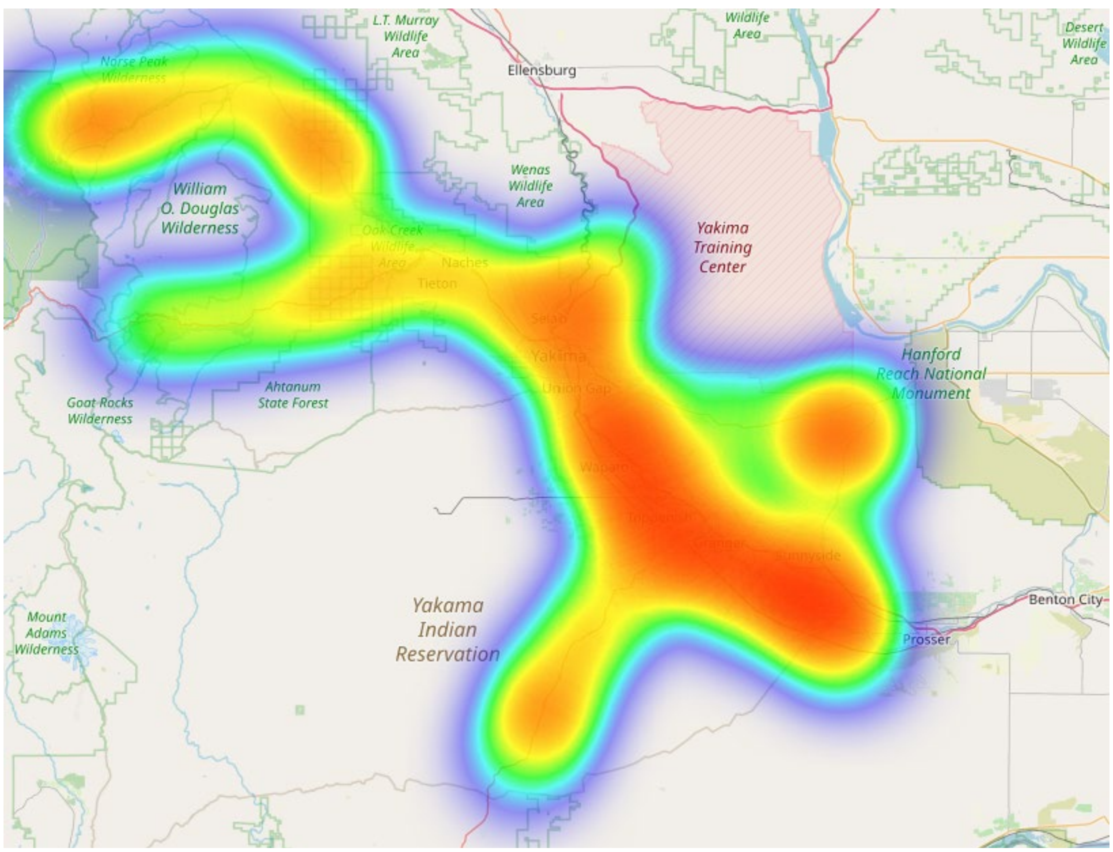


Figure 7. Crash heatmap

As shown in Figure 8, safety performance regression function provides segment-based visualization of crash frequency or estimated crash mean based on the regression model described in the Analytical Approach section.

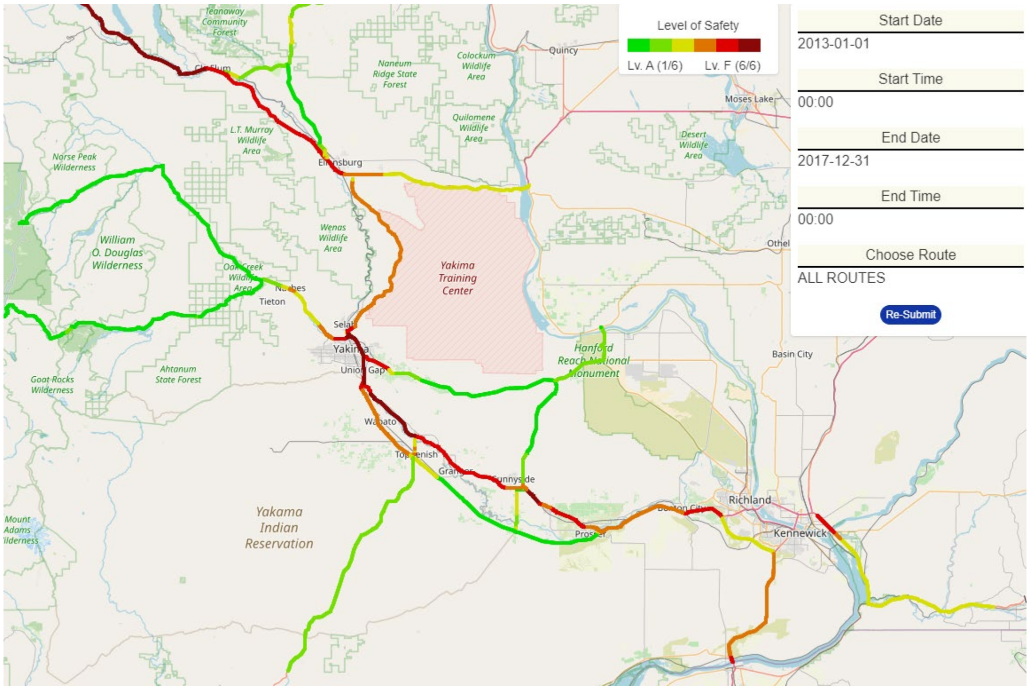


Figure 8. Safety performance regression

In addition, the safety tool has the capability to download crash data into excel/csv file based on the query settings of the users and reporting functions to generate summary figures and charts of the selected crashes (Figure 9).

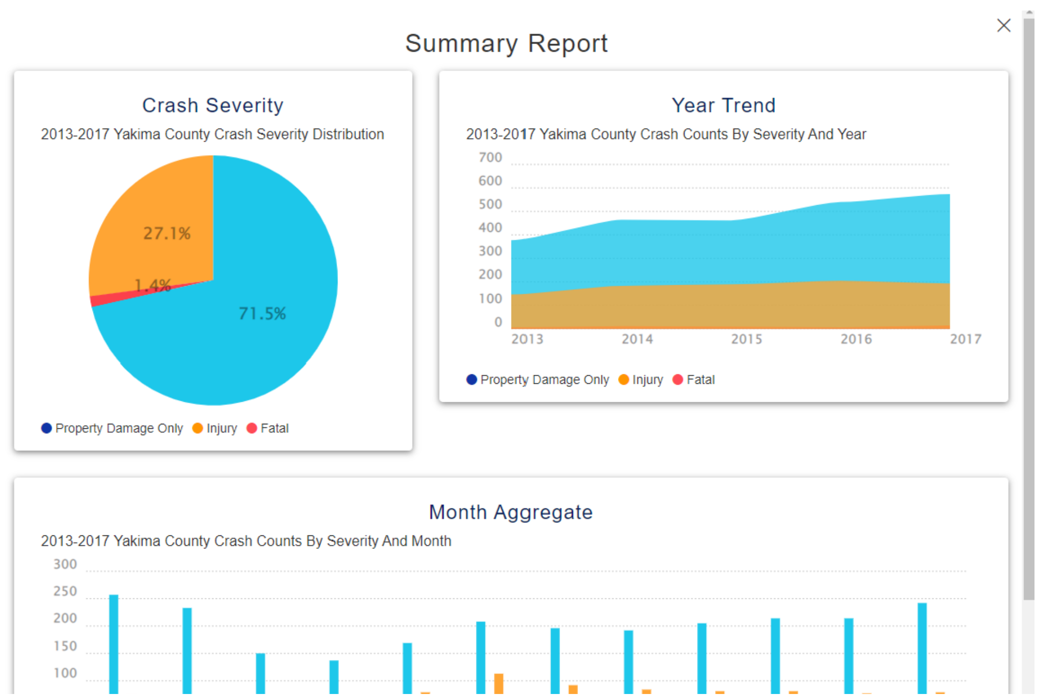


Figure 9. Crash summary reporting

### Supporting documentation and information

The supporting documentation and information can be found at the following link: https://github.com/AI-Group-STAR-Lab-UW/yakama-nation-roadway-safety-data-portal.

## Conclusion and Future Work

### Addresses the gaps

The developed safety data, i.e., comprehensive roadway safety tool, can serve as a multi-source safety data management platform for Yakama Nation transportation professionals. Besides efficient data management, the safety tool included analytical and visualization functions, such as crash data visualization, hotspot identification, screening the network based on crash type, frequency, severity, estimated risks, reporting, data download, etc. The advanced analytical techniques coupled with powerful visualization can assist the decision-makers by transferring data analysis results into actionable insights. With the help of the safety tool, Yakama Nation can obtain funding based on the analytical results of the safety tool, make effective use of the resources utilizing the safety tool to prioritize the high-risk roadway segments, and apply countermeasures based on the analysis of risk factors of the selected roadway sections.

### Lessons learned and future work

To make sure that the analytical and visualization functions were tailored to the specific needs of Yakama Nation's transportation professionals. During the project, the STAR Lab research team had frequent communication with Yakama Nation DNR Engineering team and refined the tool based on their feedback. Besides, periodic updates to key constituencies, such as the Yakama Nation Traffic Safety Committee, helped collect feedback and suggestions on safety tool development.

Yakama Nation has mixed jurisdictions, such as federal, state, counties, and cities. Multi-source safety datasets were collected and managed by different agencies in different jurisdictions. Therefore, this project experienced challenges in obtaining interesting datasets, such as the citation records from the tribal police department.

After development, the safety data tool is expected to be used by the traffic engineers and planners of the Yakama Nation DNR Engineering team. While Yakama Nation will be responsible for the maintenance of the tool, the STAR Lab will provide technical assistance when needed. Particularly, Yakama Nation DNR Engineering team plans to continue the efforts of incorporating more datasets into the safety data tool following the user guide and technical assistance from the STAR Lab.

The gaps identified through this project is not exclusive to Yakama Nation. Future work should also include bringing the data management and analytical methods, as well as the safety tool developed in this project to other rural and tribal agencies that are experiencing similar traffic safety issues.

# References


[1] Wang, Yinhai, Weibin Zhang, Kristian Henrickson, Ruimin Ke, and Zhiyong Cui. *Digital roadway interactive visualization and evaluation network applications to WSDOT operational data usage*. No. WA-RD 854.1. Washington (State). Dept. of Transportation, 2016.

[2] Xiao, Sonia, Xiaoyue Cathy Liu, and Yinhai Wang. "Data-driven geospatial-enabled transportation platform for freeway performance analysis." *IEEE Intelligent Transportation Systems Magazine* 7, no. 2. 2015.

[3] National Research Council (US). Transportation Research Board. Task Force on Development of the Highway Safety Manual, and Transportation Officials. Joint Task Force on the Highway Safety Manual. *Highway safety manual*. Vol. 1. AASHTO, 2010.

[4] Harwood, Douglas W., et al. *Prediction of the expected safety performance of rural two-lane highways*. No. FHWA-RD-99-207, MRI 4584-09, Technical Report. United States. Federal Highway Administration, 2000.