Towards Enhancing Road Safety in South Carolina Using Insights from Traffic and Driver-Education Data

Nitin Gupta[1*], **Bharath Muppasani**[1], **Saina Srivastava**[5†] **Aarohi Goel**[6†], **Ross Hartfield**[2], **Todd Buehrig**[3], **Melissa Reck**[3], **Emma Kennedy**[4], **Kevin Poore**[4], **Karilyn Tremblay**[4], **Biplav Srivastava**[1], **Lucas Vasconcelos**[1]

[1]University of South Carolina, USA [2]South Carolina Department of Public Safety
[3]National Safety Council Southeastern Chapter [4]Department of Health and Environmental Control
[5]Duke University, USA [6]BASIS Independent Silicon Valley High School

Abstract

In this student paper, we report on our project to enhance road safety in South Carolina (SC) by analyzing traffic data provided by the Department of Transportation and evaluating the impact of a school-level student driver education program called Alive@25. We improve the understanding of road safety using these traffic and training data to understand collision patterns and areas for improvement and assess training coverage gaps. Our approach combines geospatial analysis, economic impact assessment, temporal trend analysis, and interactive visualizations while leveraging AI techniques to clean and analyze extensive datasets. Key findings revealed higher collision rates in urban counties and rising collision rates in mostly rural areas, where Alive@25 participation is declining. These insights led to recommendations for improving road infrastructure and expanding safety training programs. This research demonstrates the potential of AI-driven insights to inform timely, cost-effective interventions and promote multi-stakeholder engagement in addressing public safety challenges while teaching students data science and AI skills and civic engagement.

Dashboard — https://ai4society.github.io/Traffic-Data-Analysis
Code — https://github.com/ai4society/Traffic-Data-Analysis
Project Page — https://ai4society.github.io/projects/traffic_page/index.html

1 Introduction

Road safety is critically important worldwide, particularly in the United States (US) which has one of the highest accident rates in the world (Ahmed et al. 2023). We focus on South Carolina (SC), a fast-developing US state that has also experienced a significant increase in traffic-related incidents. The application of artificial intelligence in analyzing complex datasets has opened new avenues for addressing critical societal issues such as road safety. A study utilizing machine learning methods to predict pavement conditions in SC found that factors such as Annual Average Daily Traffic (AADT) and truck percentages are crucial in determining pavement quality (Ahmed et al. 2022). This study similarly showcases the role of AI in cleaning data, revealing insights, and recommending cost-effective and timely interventions to enhance road safety in SC. By engaging multiple stakeholders, including transportation authorities, public safety officials, student educators, and AI researchers, we present a multi-faceted approach to a pressing public concern. AI techniques play a crucial role by:

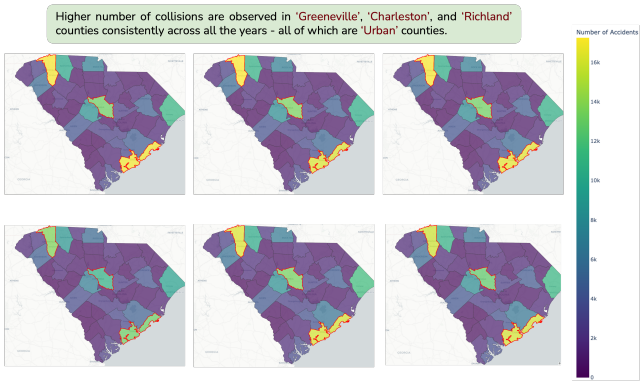

Figure 1: Heat map of South Carolina counties showing collision rates from 2017 to 2022. Darker colors indicate higher collision rates, with Greenville, Charleston, and Richland counties highlighted as high-risk areas.

1. Cleaning and preprocessing large volumes of traffic and education data (130MB size, $\succ$ 700K records)
2. Identifying patterns and trends across different geographical (road, city, county, state) and temporal scales (days, months, years)
3. Generating 3 time-based, interactive visualizations for improved understanding and decision-making
4. Proposing 2 data-driven recommendations to stakeholders for targeted interventions

The economic impact of road accidents is substantial. SC has one of the highest traffic fatality rates in the U.S., with annual losses exceeding two billion dollars due to traffic incidents (Iqbal et al. 2020). Our analysis engages with the traffic and Alive@25 data to make high-priority recommendations and promotes the use of AI for education and civic engagements.

*Contact authors: niting@email.sc.edu, bharat@email.sc.edu, biplav.s@sc.edu. †Work done during a summer internship at USC.

2 Approach and Case Study

Our methodology involved a comprehensive analysis of traffic accident data provided by the SC Department of Transportation, complemented by data from the Alive@25 program (and US census data for context).

2.1 Data Cleaning and Analysis

We encountered significant data quality issues which made it necessary to pre-process, such as filtering out invalid entries and out-of-state records. For example, approximately 26.8% and 18.3% of the data were missing from the traffic and Alive@25 datasets, respectively, for the year 2021. Furthermore, we classified counties as Urban, Mostly Rural, or Rural based on 2015 Census data to provide a framework for comparative analysis across different county types.

2.2 Interactive Dashboard

We created an interactive dashboard that visualizes collision data over time and across different counties. This tool allows stakeholders to explore and understand traffic patterns, identify high-risk areas, and monitor the impact of safety interventions. The main dashboard features are: (a) a *time-based choropleth map* that displays the distribution of collisions across counties from 2017 to 2022, allowing users to observe temporal changes and trends, (b) an *interactive heat map* that highlights areas with high collision frequencies, enabling users to identify hotspots and focus areas for intervention, (c) *county-based cumulative collision score map* that aggregates collision data over the years, providing a score for each county that reflects long-term safety trends.

2.3 Key Insights

Our comprehensive traffic and training data analysis revealed several critical insights that can inform targeted interventions to improve road safety.

Geography-Based Insights: Higher collision rates were consistently observed in urban counties such as Greenville, Charleston, and Richland (Figure 1).

Economic Impact: By combining injury costs from the National Safety Council with unit damage data from our traffic datasets, we calculated total costs per incident. This revealed that mostly rural counties had higher average costs per incident despite lower overall collision numbers.

Road Type Analysis: Two-way undivided roads were associated with more collisions than other road types.

Alive@25 Program: There was a decreasing trend in the number of Alive@25 participants in mostly rural counties, while collision rates were increasing in these areas.

2.4 Recommendations

Based on our analysis, we proposed two recommendations:

Infrastructure Improvement: Increase the count of two-way divided roads with barriers in urban counties, particularly Greenville, Charleston, and Richland. This aims to reduce collision incidents in high-risk areas (Figure 2).

Education and Awareness: Expand the Alive@25 program and conduct awareness camps in mostly rural counties to address the increasing collision rates and declining training participation.

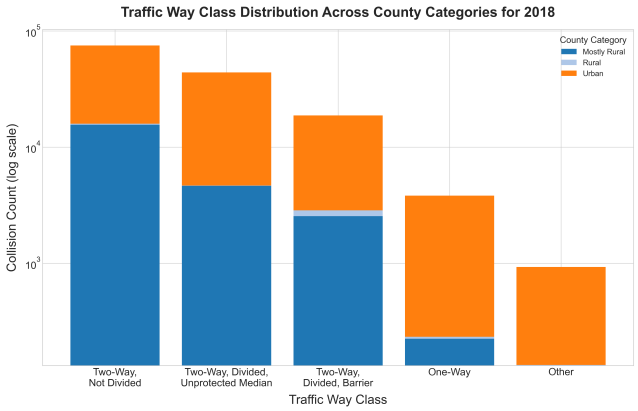

Figure 2: Stacked bar-plot highlighting the need for more two-way divided roads with barriers in urban counties to improve safety.

3 Discussion and Future Work

This study highlights the first step of collecting, cleaning, and visualizing large-scale data that engages and helps multiple stakeholders. We encourage ongoing collaboration and data sharing by making our analysis and dashboard open-source, fostering a community-driven approach to improving road safety.

There are many avenues for future work. First, one can make a *prioritized* set of recommendations using advanced statistical and simulation methods that *increase the possibility of success* given the uncertain and multi-stakeholder nature of the domain. Another promising direction is automating visualization updates to enable real-time data analysis, ensuring stakeholders can access the most current information for decision-making processes. Additionally, it is crucial to investigate the long-term impacts of the Alive@25 program on driving behavior and accident rates, thus evaluating its effectiveness comprehensively and identifying areas for improvement.

References

Ahmed, N. S.; Huynh, N.; Gassman, S.; Mullen, R.; Pierce, C.; and Chen, Y. 2022. Predicting Pavement Structural Condition Using Machine Learning Methods. *Sustainability*, 14(14).

Ahmed, S. K.; Mohammed, M. G.; Abdulqadir, S. O.; El-Kader, R. G. A.; El-Shall, N. A.; Chandran, D.; Rehman, M. E. U.; and Dhama, K. 2023. Road traffic accidental injuries and deaths: A neglected global health issue. In *Health science reports, 6(5), e1240. https://doi.org/10.1002/hsr2.1240)*.

Iqbal, A. M.; Sarasua, W.; Brown, K.; Ogle, J.; Famili, A.; Davis, W. J.; Basnet, S. B.; and Kumar, D. 2020. Assessment of crash location accuracy in electronic crash reporting systems. *Transportation Research Record: Journal of the Transportation Research Board*, 2674: 311–323.