

Abstract: Team IC23058, Level 4 Data Analysis

Fatal car crashes have been increasing in the state of Washington, particularly in the years following the Covid-19 pandemic (2020-onwards). By analyzing the dataset provided by the Washington Traffic Safety Commission, our team sought to answer the problem statement: What kind of drivers are involved in the fatal crashes in Washington and where do they come from? We acquired crash site zip codes from latitude and longitude data using reverse geocoding from ArcGIS and matched them with city boundaries to define local communities. We found that 68.9% of drivers involved in fatal crashes did not reside in the communities where the crashes occurred. We used chi-squared tests to investigate whether certain behaviors (such as drinking, speeding, drowsiness, and distracted driving) and crash types (such as road class, traffic flow, pedestrian involvement, heavy truck involvement, and bike involvement) were statistically different between crashes with visitors and residents. Our analysis revealed that there is a significant difference in crash types between visitors and residents. We computed risk rates for each zip code as the number of crashes per zip code per year to determine which zip codes produce high-risk drivers at a higher rate. Using demographic data from the Census and ArcGIS, we created data visualizations to compare the differences in population statistics between locales that tend to produce high-risk drivers. This information can assist state transit officials in deciding how to promote safety campaigns to best reach target audiences across counties, zip codes, and demographics to reduce road fatalities.

Word Count: 249