Abstract of the analysis of the Washington Fatal Crash Files

The project presents a data-driven analysis of the Washington Fatal Crash Files. By leveraging the Google Map API, we successfully implemented reverse geo-coding of the original dataset and thereby acquired the zip codes of the place where the accident occurred. Based on zip code analysis, we discovered that the proportion of non-resident drivers are significantly higher than that of the resident drivers, suggesting that non-resident drivers are much more likely to be involved in fatal crash accident. Some 75% of the crash accidents were caused by non-resident drivers.

Based on our analysis of crash types, we found that there are 4 crash types which contribute significantly to fatal crashes, namely lateral move, decelerating, slower, and going straight. As of driver behavioral factors, we found that non-resident drivers are more likely to encounter skidding, swerving, sliding, or wet leaves on road, which could be attributed to the unfamiliarity of the road condition.

On the other hand, our analysis of driver zip code suggests that some zip codes account for a significantly higher proportion of fatal crash events.