3 March 2023

From: Gabriel Abreu
Christopher Jones
Tanner Kelly
Caden Chau

Abstract:
Using the data file from the Washington Fatal Crash Survey, our goal was to determine some factors behind fatal car accidents in two areas: crashes in the driver's zip code and crashes outside of the driver's zip code. As the average land area of a zip code is around 90 square miles, a zip code provides an indication of the distance between two land areas. When the two land areas share the same zip code, it indicates that they are at a fairly close geographical distance. Through the use of Python, we were able to use tools, including pandas and geopy, as resources to guide us through the dataset to answer some key topics of discussion. With the exact x and y coordinates of the location where the car accidents occurred, we were able to trace those coordinates back to a specific zip code and match them to the driver's residency zip code. We made two datasets with this information, one dataset showed the car accidents that matched the driver's residency zip code, and a second dataset showed the car accidents that did not match the driver's residency zip code. With the two datasets, we began to analyze some causes of these accidents that occurred in these areas, the types of crashes, and their severity. We intend to determine demographic factors that might have caused the accident as well as external factors that may lead to additional data analysis.