

# Final Report

Group 6

Lucas Anderton, Hannah Brown, Lucas Gorak, Edward Mikkelsen

1/10/2020

## 1 Research Question

In our analysis, we were interested in understanding the relationship, if any, between various factors, like population increase and road conditions, and the crashes that occurred in the District of Columbia. Previous research has indicated that D.C. has the third worst traffic congestion in the United States. However, their ranking has improved since 2011, which D.C. had the worst congestion of all 50 states. In that same period of time, the number of crashes in D.C. has increased significantly. Intuitively, we credited the increase to more traffic from the constantly growing population in the Washington metro area. D.C.'s public datasets on traffic are convoluted and difficult to manipulate for analysis. It is recorded in 15,000+ rows from various traffic-volume meters throughout the city on an annual basis. However, every year, the number of meters and other variables in the data change. Alternately, we chose to compare the number of crashes to the 311 traffic service requests. 311 traffic service requests offer information regarding the number of commuters out on the roads and the state of the city's vehicle infrastructure. In our research, we wanted to understand what the relationship between these 311 traffic requests and the frequency of crashes in D.C. Additionally, these data include timestamp information, like the date they were submitted, the date they were due (most likely based on a city algorithm for how long certain requests should take), and the date they were resolved. We wanted to see if a relationship existed between rising latency in resolution times and the frequency of crashes. The reasoning behind our hypothesis was that if the city was taking longer to respond to poor road and navigation conditions, more incidents may occur. We also wanted to investigate if there was a geographical relationship between crashes and 311 requests. And lastly, we were interested in analyzing what factors best estimate risk of failure in a crash.

## 2 Data Collection Procedure

## 3 Data Processing

## 4 Statistical Methods

## 5 Analysis of Results

## 6 Implications