

Dylan Patel

Aver Mckay

Jianqi Liu

## **Final Project Proposal**

For our final project, group 20 would like to analyze car crash data in the United States. In order to implement this we have already found a data set on the website Kaggle.com. Kaggle is a great website that we came across in PLS202 last semester. It had a ton of highly credible, user made excel files with loads of data for about almost any topic that you could imagine.

Some questions that we hope to answer are what states have the most crashes, how severe are the car crashes per state, is there a time of year where accidents are most prominent, is there a year that had significantly more crashes than another, and also what part of day do most crashes occur.

In order to do this, we will have to subset our data into many different groups, and based on our findings for one question we will dig deeper into it and find out why that is the case. A lot of bar graphs will be implemented, and many libraries will as well such as caret, tidyverse, and readr to name a few.

Link to data set

<https://www.kaggle.com/sobhanmoosavi/us-accidents#US Accidents Dec19.csv>