Abstract for Capstone Project
Najah Moussa
24 August 2020

Automotive accidents result in over 30,000 fatalities in the United States annually. The National Automotive Sampling System (NASS) provides a nationally representative sample of police reported collisions and is made available to researchers and the general public.

The research question is to identify and quantify factors which impact the survivability of various crash types (Rear-end, Sideswipe, etc) using R, and potentially create a web app using the shiny package to predict survivability for given inputs using regression.

Techniques will include web-scraping, xml-parsing and data cleaning of real-world dataset, exploratory analysis to identify relevant factors, feature engineering, and linear regression.