Adam Imran, Thomas Malejko, Douglas Post, and Patrick Aquino

ANLY 512

Final Project Proposal

*Preliminary Topic:*

**Montgomery County Traffic Data:**

**An Exploratory and Predictive Analysis**

*Goal:*

Based on the Montgomery County Traffic Dataset identify any significant trends that exist regarding ticketing, accident, and injury likelihood as they relate to driver demographic, vehicular, and weather data.

*Dataset*:

[Montgomery County Traffic Violations](https://data.montgomerycountymd.gov/Public-Safety/Traffic-Violations/4mse-ku6q)

* Size: 1.66 Million Rows x 43 Features (670 MB)
* Time Frame: 1 January 2012 – 4 March 2020
* Background: “This dataset contains traffic violation information from all electronic traffic violations issued in the County. Any information that can be used to uniquely identify the vehicle, the vehicle owner or the officer issuing the violation will not be published.”

*Detailed Objectives:*

* Explore the dataset, generate summary statistics, and identify any preliminary trends that emerge.
  + Reduce dataset scope to three-year window (1 January 2017-31 December 2019)
  + Generate features (integrate light and weather data, rush-hour, etc.)
* Predict the likelihood of a driver being injured in an accident based on appropriate predictors.
* Predict the likelihood of a driver receiving a citation or warning based on appropriate demographic predictors.

*Specifications:*

1. TBD

*Milestones:*

07 March—Project Proposal Agreed Upon

16 March—Preliminary Analysis Complete (Summary Statistics, etc.)

23 March—Objectives Updated Based on Preliminary Analysis

13 April—Technical Analysis Complete

24 April—Project Presentation Rehearsal

28 April—**Project Presentation**

02 May—Project Submission

03 May—**NLT** **Project Submission**