Car Accident Analysis Project Proposal

Noah Shoap and Jonathan Graham

***Abstract*— A project proposal for the analysis of car accident data sets, with the goal of finding trends we wouldn't expect.**

I. Project Objective

The objective of this project, Car Accident Analysis, is to apply what we've learned as computer scientists and conduct a thorough analysis of a data set containing car accident information. By doing this, we aim to gain insight into what features of the data set correlate to more frequent and serious accidents, specifically, the aim being to find correlations we may not expect.

II. Motivation

Car accidents and the fatalities that often result from them are a serious public safety concern. While some conditions are obviously bad for driving, such as intense rainfall, perhaps there are some other dangerous correlations that are not as obvious. If we could find those, bringing awareness to those correlations could make drivers more aware and increase public safety.

III. Data

The data we're going to be using is a data set from Kaggle. It is a rather large data set (the file is over 3GB) and it contains information pertaining to weather conditions, traffic conditions, time of day, city, state, and more.

The data set is rather large, so we will have a team discussion about if we want to filter it down to a subset of the data, and if so what criteria we will use to form that subset.

IV. Team Responsibility

We only have two team members, so currently the expectation is that both Jon and Noah will often be working together and on the same tasks. Currently, the expected tasks are:

- Data Collection
- Filtering data set down to a sub set (if we decide to)
- Analyzing Data
- Data Visualization
- Write-up of results

V. Milestones

- (Oct. 1) – Decide on any possible data filtering criteria
- (Oct. 8) – Data used for analysis is finalized
- (Oct. 22) – Analysis Complete
- (Nov. 5) – Visualization of Data / Analysis Complete
- (Nov. 12) – Summary of results written

VI. Expected Outcome

We expect to see some obvious correlations, such as heavy rainfall correlating to more frequent accidents. But we also expect that we will find some correlations that we did not expect going into the analysis. The hope being that there is something we can learn from this analysis to help people become safer drivers.