Proposal: Analyzing and Visualizing traffic accidents data

Members:

Mehrnaz Motamed Willy Ma Huaqing Shen Rakshith Manandi

Problem:

Analyzing, finding correlations and visualizing vehicle accident history

Dataset:

Link: FARS dataset

The entire dataset contains csv files from the year 1975 to 2017, which contain various information like the number of vehicle crashes, the number of vehicle fatalities, etc. The data includes accident level, vehicle level, and person level information in addition to alcohol consumption.

Proposed Solution and Real-world Application:

The National Highway Traffic Safety Administration (NHTSA) has interesting public datasets describing statistics of car crashes in each year. The purpose of this study is to explore this dataset, visualize it, understand it, and determine the key factors in vehicle accidents.

This analysis can potentially include the relationship between car accidents and age, gender, and so forth. The analysis can also be used to visualize information from dataset such as the age group with the highest or lowest accident rate.

The real world application of this analysis is that decision-makers can rely upon this data to invest on education for a certain age group in driving, or to improve the factors that yield in lower traffic accident rates. The extracted information can also be used to interpret the factors behind the traffic accident trends in time and to look for possible responsible factors so that we can build a safer environment with fewer fatalities from these kinds of accidents.

Step	Estimated completion time	Person(s) in charge (among the group of 4)
Extracting and cleaning up data	One week	Mehrnaz Motamed (alcohol) Willy Ma (time) Huaqing Shen (location) Rakshith Manandi (circumstances)
2. Data Analysis	Two weeks	Mehrnaz Motamed (alcohol) Willy Ma (time) Huaqing Shen (location) Rakshith Manandi (circumstances)
3. Data visualization (to obtain data statistics and trends)	One week	Mehrnaz Motamed Willy Ma Huaqing Shen Rakshith Manandi