Analysis of Seattle Collision Data

a presentation by Miranda Childs for Coursera's IBM Data Science Professional Certificate capstone project

Collision reduction starts with awareness

▶ In Seattle, WA, from 2004 to present there have been 194,673 collisions reported by the Seattle Police Department (SPD) to the Seattle Department of Transportation (SDOT). 58,188 of those collisions involved an injury.

▶ By determining and focusing on the factors that contribute most heavily to **severe collisions**, we can create meaningful strategies to reduce the number of accidents, especially those with injuries, in order to increase the wellbeing and longevity of our community.

▶ Vision Zero Network is "a collaborative campaign helping communities reach their goals of Vision Zero -- eliminating all traffic fatalities and severe injuries -- while increasing safe, healthy, equitable mobility for all." (https://visionzeronetwork.org). Through our thorough analysis we will make recommendations for the next campaigns and strategies that Vision Zero can execute in collaboration with SDOT.

Data Acquisition and Cleaning

▶ Collision data from 2004 to present, provided by the Traffic Records Group in conjunction with the Seattle Police Department and Seattle Department of Transportation. (collision data set, and metadata)

▶ 194,673 rows and 37 features in the raw dataset

▶ Duplicate columns and null values were dropped

▶ Features were transformed for consistency and accuracy

▶ Undersampling was utilized to balance the labeled data

▶ 4 features were chosen to predict the severity of collisions: Under The Influence, Weather, Road Condition, and Light Condition

▶ The cleaned data contains 111,840 rows and 4 features