**Data**

The data used is Collisions—All Years data and it is provided by by SPD and recorded by Traffic Records.

The data is a collection of records of the accidents that have occurred from 2004 and onwards. This includes all types of collisions. Collisions will display at the intersection or

mid-block of a segment. It is updated weekly.

The data has many Features/attributes 38 in total. But for the study only few are selected for the analysis. These are mostly having information about severity of collision, Type of Address, Road condition, Weather condition, Alertness of driver, Alcohol consumption or not by the driver, and few others.

So to proceed with the study these Attributes will be analyzed to see which features are more frequent and related to which type of collision and causing what level of severe accident.

The findings will be displayed as graphs for better understanding.

Once the features which are strongly linked in somehow leading to an accident are identified, they can be use to build a predictive modeling, in order to know beforehand what are the chances of having an accident if the situations are mentioned.