**Data**

The data used in the analysis is gathered from the Traffic Records Group in the SDOT Traffic Management Division from Seattle, WA. It includes all collisions provided by the Seattle Police Department and recorded on the Traffic Record, displayed at the intersection or mid-block of a segment from 2004 to the present. The project purpose is to analyze and predict the severity of an accident based on some particular features that will be chosen.

**Data Cleaning**

The data set has 194673 observations with a lot of variety. Out of which many has ‘NaN’ values. The target variable here will be SEVIARITY CODE, which shows that the damage is only to the property or it caused injury. For this we need to assign the values to the variables which are in form of text. Then, Data cleaning an balancing can be performed.