**Data**

The data was collected by the Seattle traffic police department from 2004 to present.

The data is raw data set which has inconsistencies and imbalanced , which needs to be cleaned and processed.

The data consists of 37 independent variables and records of 194,673 accident cases occurred in Seattle. The dependent variable is “SEVERITYCODE”, contains numbers that correspond to different levels of severity caused by an accident from 0 to 4.

Severity codes are as follows:

0: Little to no Probability (Clear Conditions)

1: Very Low Probability — Chance or Property Damage

2: Low Probability — Chance of Injury

3: Mild Probability — Chance of Serious Injury

4: High Probability — Chance of Fatality