**A description of the data and how it will be used to solve the problem. (15 marks)**

The data being used is a shared dataset from Seattle city, it shows a record of the occurrence of traffic accidents from 2004 to the present day. Our predictor variable will be accident severity and our independent variables will be weather conditions and speed.

We must prepare the dataset first through normalizing and changing the nature of the variables in use.

Regression analysis will be used to assess the impact of speeding and weather conditions on accident severity. From this, a machine learning model will be constructed (using the most suitable technique) to determine accident severity going forward under the same conditions and the model itself will be deployed and then tested to assess model accuracy.