Data

The data used in this project is hosted on <https://s3.us.cloud-object-storage.appdomain.cloud/cf-courses-data/CognitiveClass/DP0701EN/version-2/Data-Collisions.csv>. It is a csv file containing 38 columns and 194,673 rows. The data captured is various information on accidents occuring in Seattle, Washington over various years. The target values to be predicted are the severity of the accident, found in the headings SEVERITYCODE and SEVERITYDESC.

Some of the factors used to predict severity will be the weather, road conditons and lighting at the time of the accident, given by WEATHER,ROADCOND and LIGHTCOND respectively. Time is another possible predicitng element given in INCDATE, INCDTTM (incident date/time). The amount of people involved, pedistrian/cyclists, and vechiles involved can also be imdicators, given by PERSONCOUNT,PEDCOUNT,PEDCYLCOUNT,VEHCOUNT.