1. **Dataset description:** We have used the weekly updated traffic dataset provided by Seattle government that contains road accident and corresponding severity based on the different attributes. The link of the dataset is given below: https://s3.us.cloud-object-storage.appdomain.cloud/cf-courses-data/CognitiveClass/DP0701EN/version-2/Data-Collisions.csv

   The above dataset has 37 attributes which act as independent variable for us to build the prediction model. The target variable will be the collision severity code which ranges from 0 to 3 namely; 0 for unknown, 1 for prop damage, 2 for injury, 2b for serious injury, and 3 for fatality. Here we will choose the attributes which strengthen our prediction model. In term of Data Science, we should select the attributes which column has fewer empty cells.

2. **Approach to solve the problem:** Our target variable is 'SEVERITYCODE' because it measures the severity of an accident. The important attributes used to measure the severity of accidents are 'WEATHER, 'ROADCOND', and 'LIGHTCOND' as we have more data in these attributes which will help to build the model for collision severity detection. Also, it will be helpful to alert the end user to be careful while driving on roads.