Introduction/Business Problem:

Traffic collisions lead to related deaths and economic loss every year globally, and from a damage prevention point of view, this is an area of critical concern to society. In most cases, the significant causes of injuries that we can avoid by enforcing strict laws are not paying enough attention when driving, consuming drugs, and alcohol or traveling at too high speeds. In addition to the above explanations, the main uncontrollable variables that we can avoid are temperature, visibility, or road conditions by exposing secret trends in the data and issuing warnings on the relevant roads to the municipal authorities, police, and passengers.

For the IBM Data Science Capstone Project, I have planned to build an accident severity prediction model to forecast an accident's likelihood given the current weather, road, and safety conditions to minimize car crashes in a city. By implementing such a model that warns drivers, health services, and police to encourage them to be more vigilant in sensitive circumstances, the Seattle government can deter avoidable car crashes.

Data:

The data was collected by the Seattle Police Department and Accident Traffic Records Department in the SDOT Traffic Management Division from Seattle, WA. It includes all collisions provided by the Seattle Police Department and recorded by the Traffic Record, displayed at the intersection or mid-block of a segment from 2004 to the present.

The data consists of 37 independent variables and 194,673 rows. The dependent variable, "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