Introduction/ Business Problem

The car accident severity dataset is provided which gives the severity of the car accident based on the fatality of the occurred accident. It also has many different types of attributes which could have contributed to the causation of the accident as well as the fatality. For this capstone project, I am going to use the data given to hypothesize which of the conditions during the drive has the highest rate of accident occurring versus which conditions result in the highest number of fatalities. This could be used to caution the drivers on the road when those conditions are present to decrease the amount of accidents that occur in the future. It could also be used to increase the amount of police on the road during these specific conditions to oversee whether the drivers are driving at a lower speed during those hazardous conditions.

Data

The dataset that is given is called car accident severity. It has many attributes as well as the fatality of the car accidents. Some attributes that are included are location, weather condition, car speeding, light conditions, road conditions, whether the place of accident was a junction, number of people involved and number of vehicles involved. I would be using all the above attributes by posing them against the fatality of the occurring accidents as well as the amount of accidents that occurred.