**Description of Data-**

Data Source

Data Source: These data have been collected and shared by the Seattle Police Department (Traffic Records) and are provided by Coursera for downloading through a link. I have uploaded the data in github repository

Data Location: https://github.com/TuesyBharadwaj/Coursera\_Capstone

Data set name: Data-Collisions.csv It has 37 attributes and a little more than 1 million rows with the accidents’ severity data. We will be looking into the below attributes to predict accident severity

The data describes two types of Severity for accidents

-> Injury collision

-> Property Damage by collision

We have to do exploratory data analysis on the data set to analyze patterns and relationships between factors in the table so that we can predict the severity of cases in the future.

We will be looking into the below attributes to predict accident severity

1. Road Condition (Wet, Dry, Sand, Oil, Mud, Snow,)

2. Weather Condition (Clear, Sun, Partly cloudy, rainy,)

3. Light Condition (Daylight, Dark, Dusk)

4. Inattention IND (If driver was distracted - Y/N)

5. UnderInfl (If driver was under influence of alcohol – Y/ N)

6. Speeding (Y/ N)

We will find out answers to the below questions

* What are the factors that have a high impact on road accidents?
* Is there a pattern to them?
* Is there any Correlation?

We will have to analyze the data to get a clearer picture and draw conclusions.

**Objective**

The objective is to define the problem, to find the factors that can have a relevant weight in the quantity and seriousness of the accidents, so that hospitals, emergency services, para medics etc. can be interested in reducing these figures, and focus the resources in points where these conditions converge.

In order to provide greater clarity, I will try to analyze the data, see if there are relationships or patterns, based on severity of accidents, so that measures can focus on these points as a first prevention strategy.

The data will be used so that we can determine which attributes are most common in traffic accidents in order to target prevention at these high-incidence points, facilitate resource allocation and mitigate probability of serious accidents given conditions