**Traffic Accidents Severity Project**

**Introduction of Problem:**

Traffic accidents represent one of the leading causes of death worldwide and damage of economic expenditure. Government is providing various measures and campaigns to raise awareness of the seriousness of the accident problem, it still occurs quite frequently. The impact of road accidents on society and the economy is high, and human losses are compounded by large expenditures on health care, awareness campaigns, mobilization of specialized personnel, etc. The WHO sets the economic impact of road accidents in a developed country at 2 to 3% of GDP, a significant figure for any country. Collaboration to reduce these losses has become an important issue of general interest.

In this project we would be using data, collected from the Seattle Police Department and provided by Coursera via a download link. The time period for this data is from 2004- present and contains such information as severity, location, collision type, weather conditions, road conditions, and light conditions, among others.

**Defining the problem**:

On the basis of Seattle car accident data, we will focus on the below problems:

* What are the factors that have a high impact on road accidents?
* Is there a pattern to them, so that we can take measures to reduce accidents?

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

This analysis would help traffic police and our society to reduce traffic accident incidents. Police can take required precautions to avoid them.

I have primarily focus that how road, light and weather conditions are going to effect the severity of the accidents. Is there any particular circumstances in which there are more accidents or accident severity depends?

I have also tried to find the accident-prone areas in the city.

**Description of Data:**

In this project we would be using data, collected from the Seattle Police Department and provided by Coursera via a download link. The time period for this data is from 2004- present and contains such information as severity, location, collision type, weather conditions, road conditions, and light conditions, among others.

For an accurate prediction of the magnitude of damage caused by accidents, they require a large number of data on traffic accidents with accurate information to train prediction models. The data set provided for this work allows the analysis of a record of ~195k accidents in the state of Seattle, from 2004 to the date it is issued, in which 20+ attributes or variables are recorded and the codification of the type of accident is allowed The information can be extracted from it:

speed information

information on road conditions and visibility

type of collision

affected persons, etc

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.

**Methodology:**

First of all, Panadas library was imported in the Jupyter notebook and data was observed to analyze. After that the data was cleaned for further analysis.

**Cleaning of data:**

Cleaning the data is an important step prior to analyze the data. As I am focused on a few data points to observe the data, hence I have checked if there is any blank cells in case of weather, light condition, address type and road condition. There were a lot cells which are empty in these columns, I removed those cells and reconfirm the same.

**Drawing relation between different parameters:**

I have tried to find the relationship between weather, light condition, road condition and address type with total accidents.

I have used below code to analyse the same:

*# Draw relation between address type and accidents count*

*df\_addresstype\_accident=df2.groupby("ADDRTYPE")["SEVERITYCODE"].count().reset\_index()*

*df\_addresstype\_accident*

*# Draw relation between road condition and accidents count*

*df\_roadcond\_accident=df2.groupby("ROADCOND")["SEVERITYCODE"].count().reset\_index()*

*df\_roadcond\_accident*

*# Draw relation between weather condition and accidents count*

*df\_weather\_accident=df2.groupby("WEATHER")["SEVERITYCODE"].count().reset\_index()*

*df\_weather\_accident*

*# Draw relation between light condition and accidents count*

*df\_lightcond\_accident=df2.groupby("LIGHTCOND")["SEVERITYCODE"].count().reset\_index()*

*df\_lightcond\_accident*

After this observation, I have tried to observe data if there is any specific relation between the above conditions and severity of the accidents. For this I have divided data in two parts according to severity of the collisions, first one is ***injury collisions*** and ***property damage only collisions***.

*# Extract data which we need to draw patter between road condition, weather, light condition, address type, person count in accident and severity of accident include Injury Collision*

*df3=df2.loc[(df2['SEVERITYCODE'] >= 2.0), ['ADDRTYPE','PERSONCOUNT', 'WEATHER', 'ROADCOND', 'LIGHTCOND', 'SEVERITYCODE']]*

*print(df3)*

*# Extract data which we need to draw patter between road condition, weather, light condition, address type, person count in accident and severity of accident include Property Damage Only Collision*

*df4=df2.loc[(df2['SEVERITYCODE'] <= 1.0), ['ADDRTYPE','PERSONCOUNT', 'WEATHER', 'ROADCOND', 'LIGHTCOND', 'SEVERITYCODE']]*

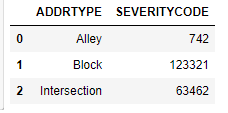
*print(df4)*

Then, again I draw the relation between different conditions of weather, light and road, collision address type and collision type, using the same code.

**Result:**

Results of the relations various with the different parameters.

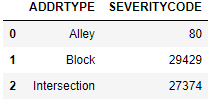
***Address type and collisions:***



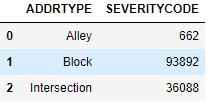
This shows that maximum collisions happened in the block area and intersections is on second. However, the collisions occurred in block is approx. double the same occurred in intersections.

In case of injury collisions, block and intersection areas are almost same. In blocks, in 24% collisions injury happened, however in case of intersections 43% collisions are injury collisions.

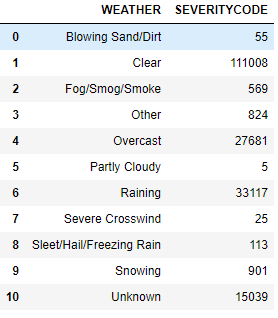
***Injury Collisions***



***Property Damage Only Collisions***

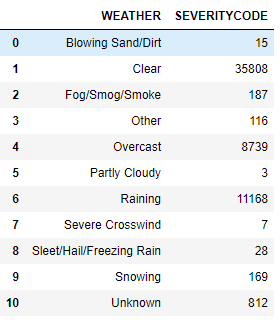


***Weather conditions and collisions:***



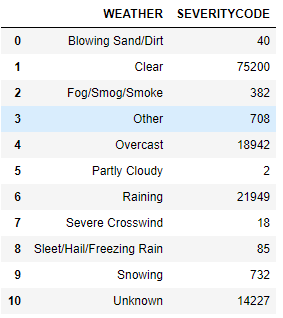
According to the above table, maximum accidents happened in clear environment in comparison to other conditions.

***Injury Collisions***

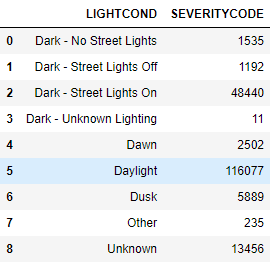


In case of clear, overcast and raining weather conditions, 32%, 32% and 34% respectively collisions were injury collisions.

***Property Damage Only collisions***

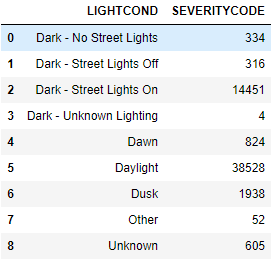


***Light conditions and collisions:***



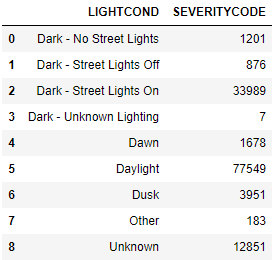
Maximum collisions occurred in daylight circumstances compared to other bad light condition accept dark – street lights on.

***Injury Collisions***

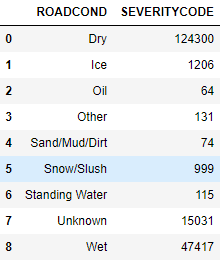


Approximately, 30% to 36% of total collisions are injury collisions, however, dark-no street lights or street lights off are conditions where property damage occurred more in comparison to other light conditions.

***Property Damage Only Collisions***



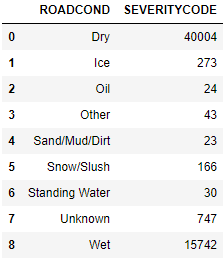
***Road condition and collisions:***



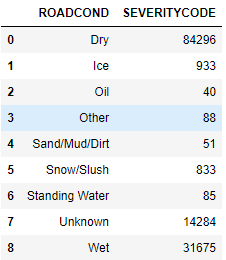
The maximum collisions occurred in the most suitable condition of road., which more than 65% of total collisions during the period. However, in case of wet road conditions, 25% of total collision happened.

In case of dry and wet road conditions, ~32% of collisions are injury collisions, however oil road condition is the most dangerous condition, where 38% of total collisions are injury collisions.

***Injury collisions***



***Property Damage Only Collisions***



**Conclusions:**

The above results shows the block areas more dangerous for collisions in comparisons to other areas, however intersections are also not safe, as in these areas approx. 43% of total collisions are injury collisions, which is higher than block areas collisions.

In case of other conditions, such as road, light and weather, good conditions for driving are more prone to collisions, as people were more careless and over speeding the vehicle.