Analysis of Severity of Accidents in UK

Introduction

Road accidents are a major world economic and social problem as shown by the report of loss of lives and properties in many countries around the world.

On average, five people die every day on the road in Great Britain and countless more are injured.

Britain's road safety record has stagnated in recent years, with the number of road deaths remaining broadly constant for several years.

In recent years, there is an increase in the researchers' attention to determine the significant factors that affect the severity of the injuries which is caused due to the road accidents.

Problem

The objective of this analysis is to know and understand more about the severity of accidents and to see whether different variables considered have any effect to the cause of the accident and what could be done with respect to this.

Target Audience

The purpose of accident analysis is to help decision-makers understand the nature, causes, and injury outcomes of crashes, the mental and emotional injuries after a car accident. This information provides context for the design of strategies and interventions that will reduce accidents and their consequences.

Data Acquisition and Cleaning

The secondary data used for this analysis has been taken from Kaggle consisting of road accidents in UK.

Period of 10 years ranging from 2004 to 2014 and which is further divided into three datasets namely Casualties, Accidents and Vehicles.

Merged the three datasets on the basis of accident index.

Dropped some features like latitude, longitude, accident location to address the missing values.

Target Variable

Questions that were considered for the analysis part:

Is there any relationship between the timing of the accident and number of fatal accidents?

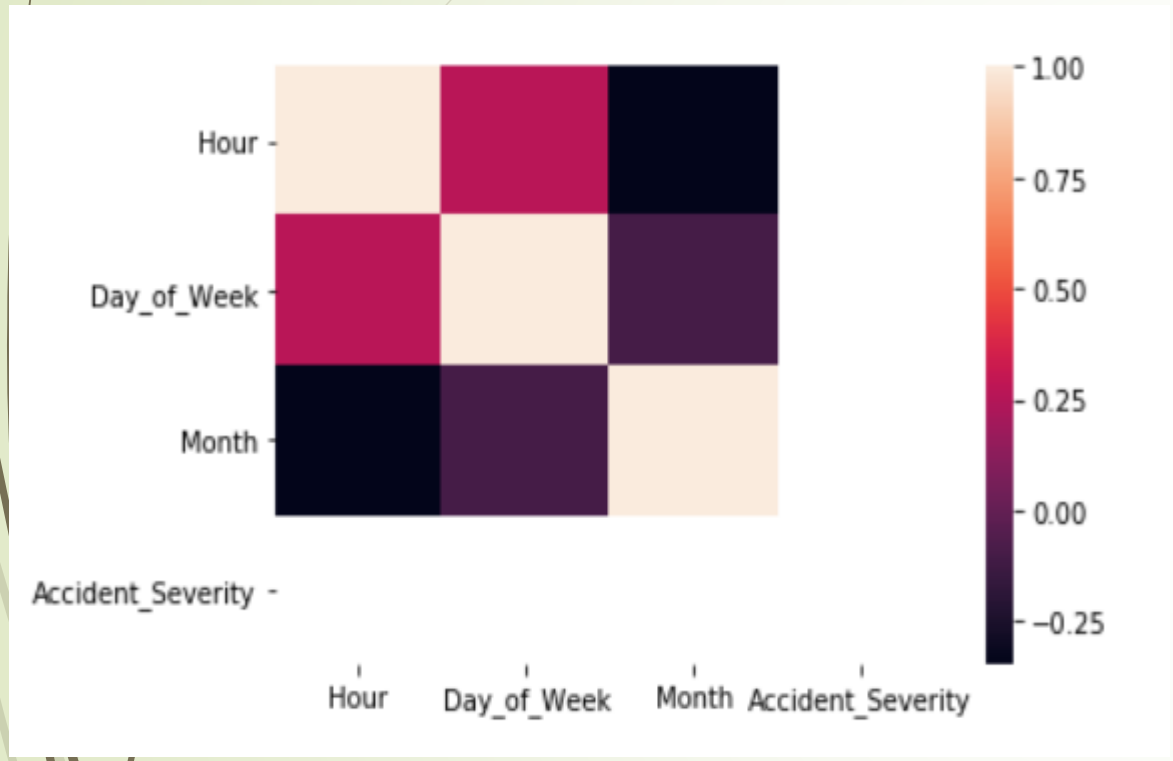
Does the age of driver have any effect on the number of accidents happened?

How does the weather impact the severity of accidents?

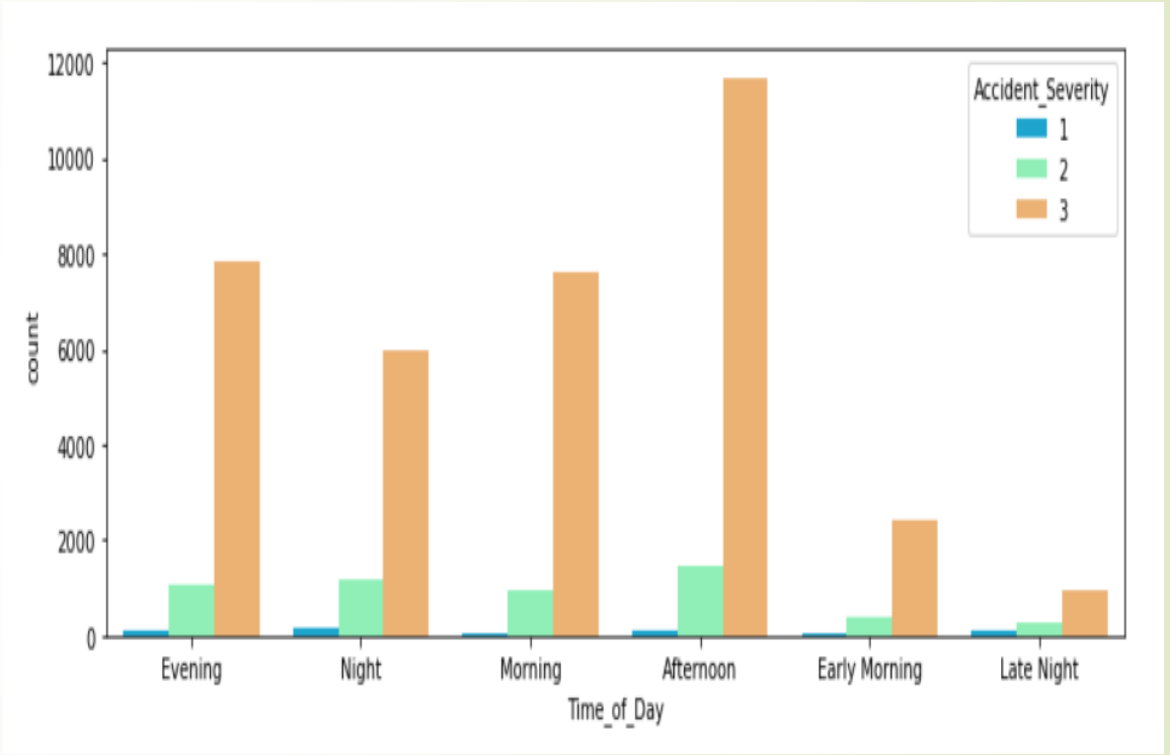
Are certain designs of vehicles safer than others?

Target variable- Severity of accidents

Relationship between Timing and Fatality of Accidents

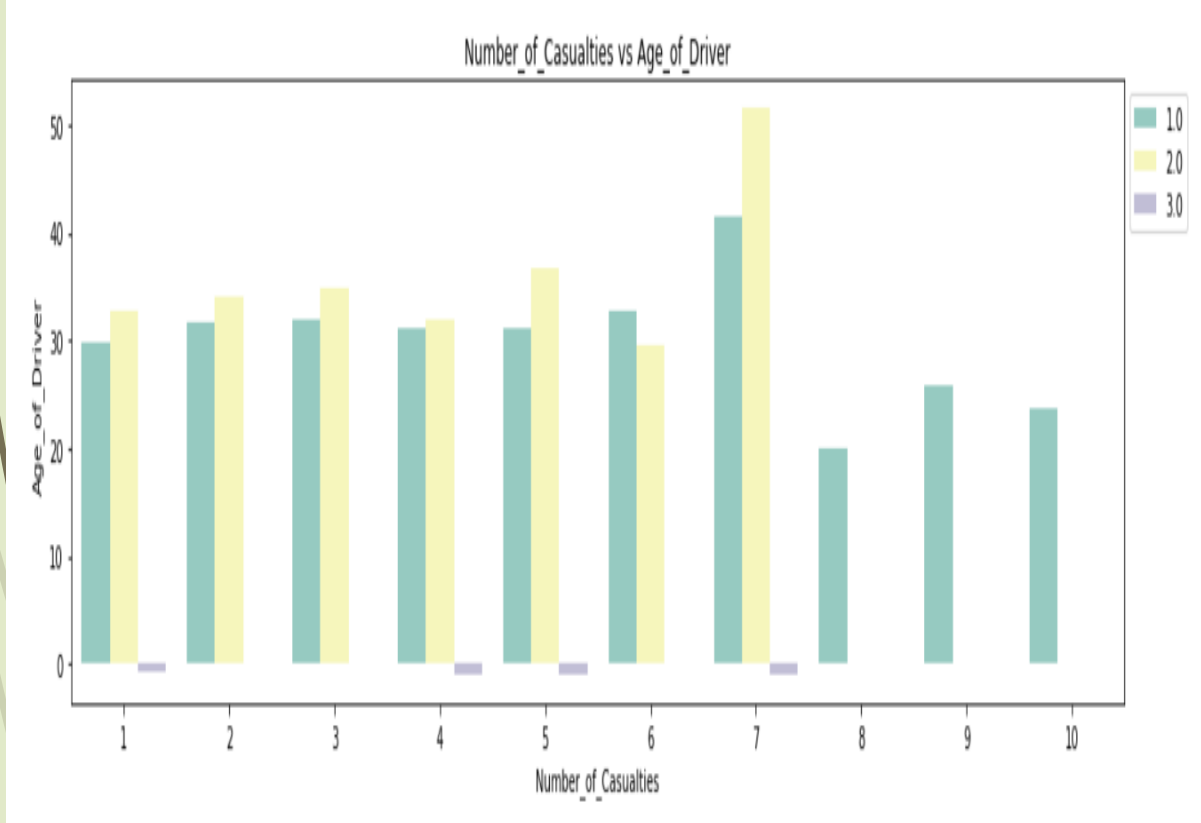


Correlation Matrix between Timing of the Accident and Accident Severity

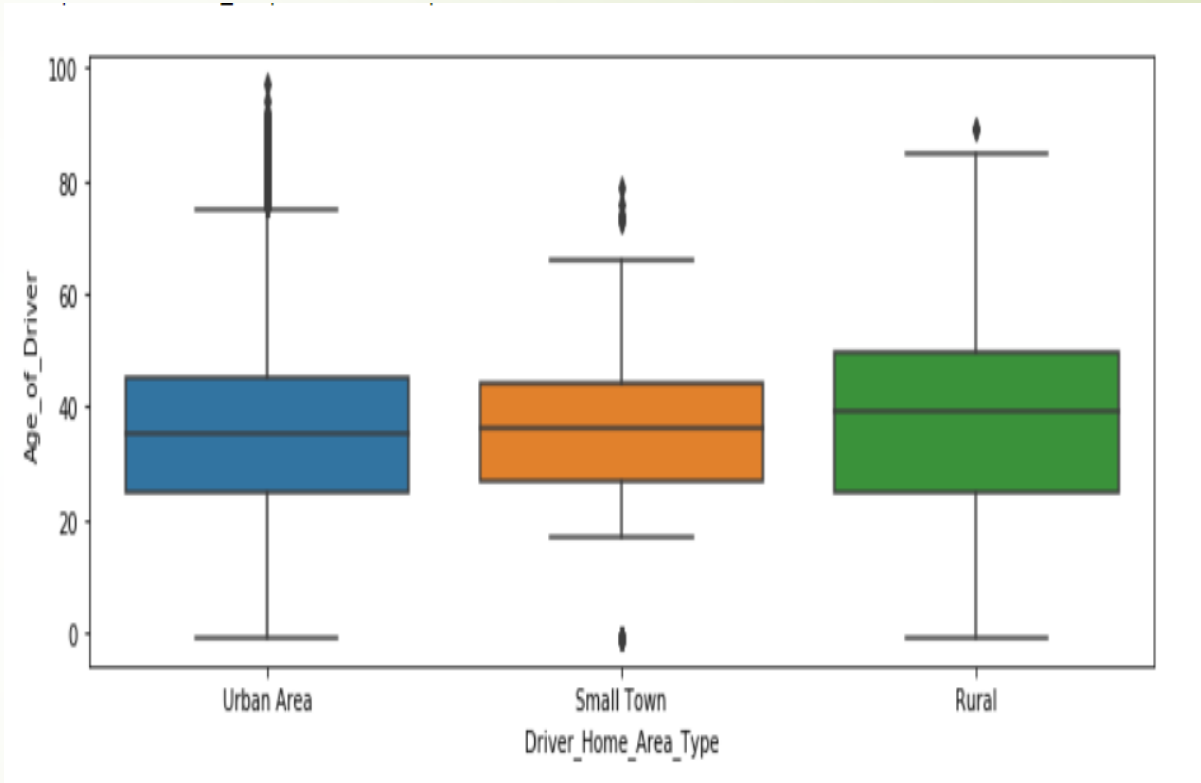


Bar Graph showing Timing of Day and the Number of Accidents

Relationship between Driver and Accidents

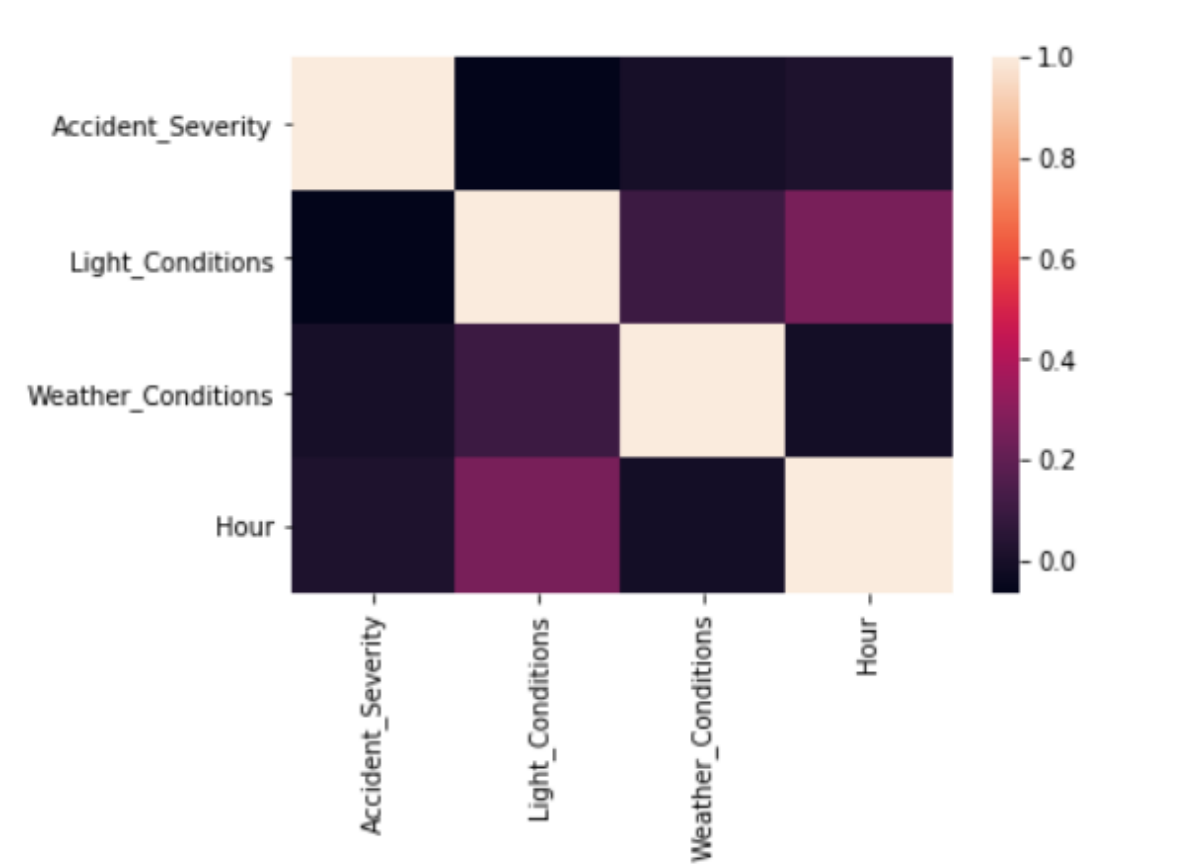


Bar graph showing relationship between Number of Casualties and Age of Driver

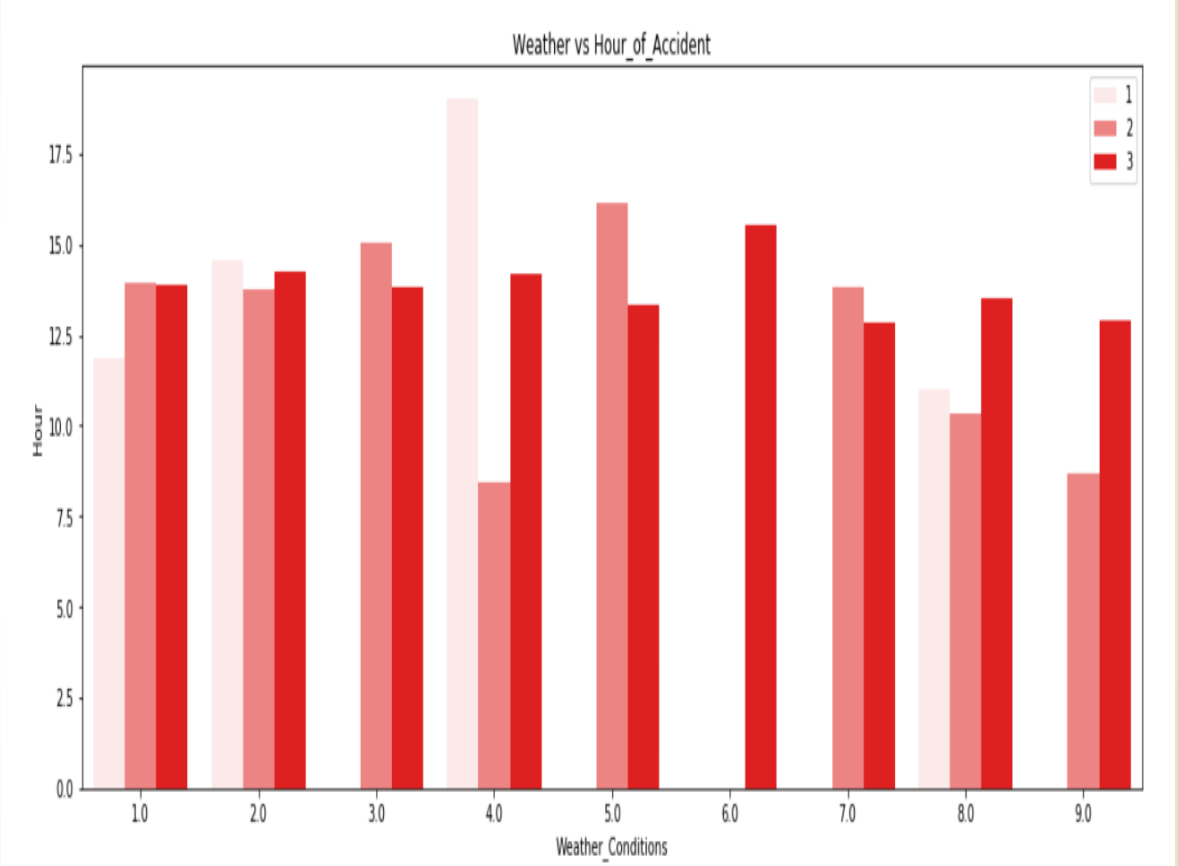


Box Pot showing the distribution of Age of Driver and their residence type

Relationship between Weather and Severity of Accidents

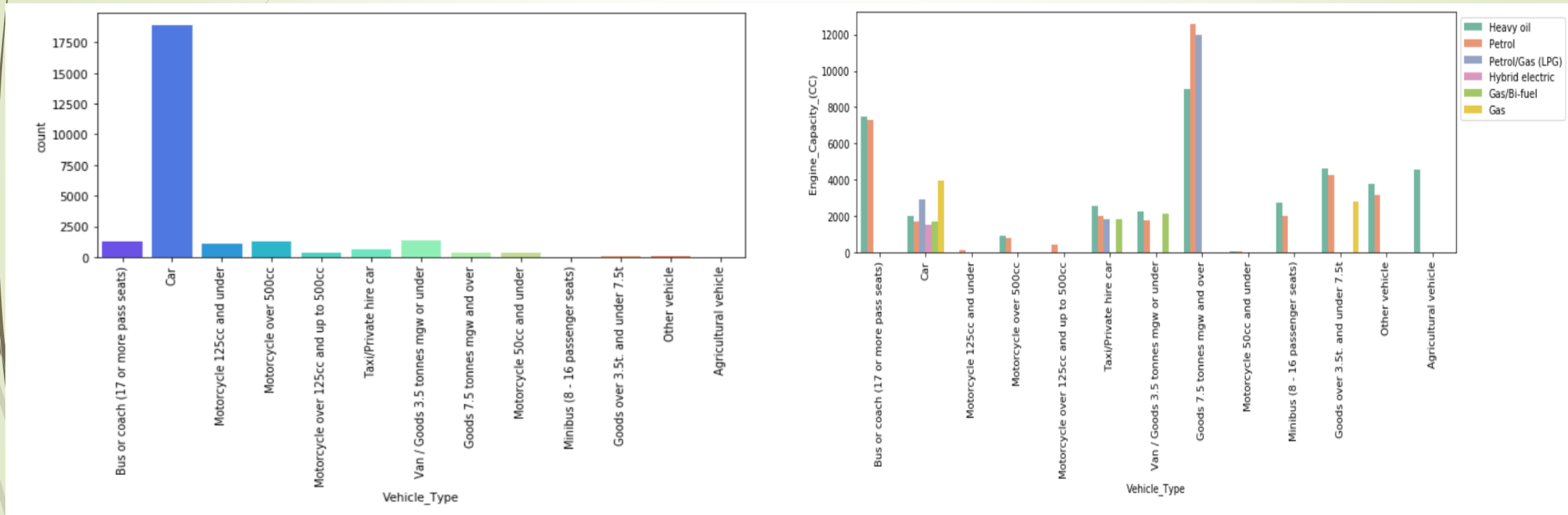


Correlation Matrix between Weather Conditions and Accident Severity



Bar Graph showing relationship between Weather Conditions and Hour of Accident

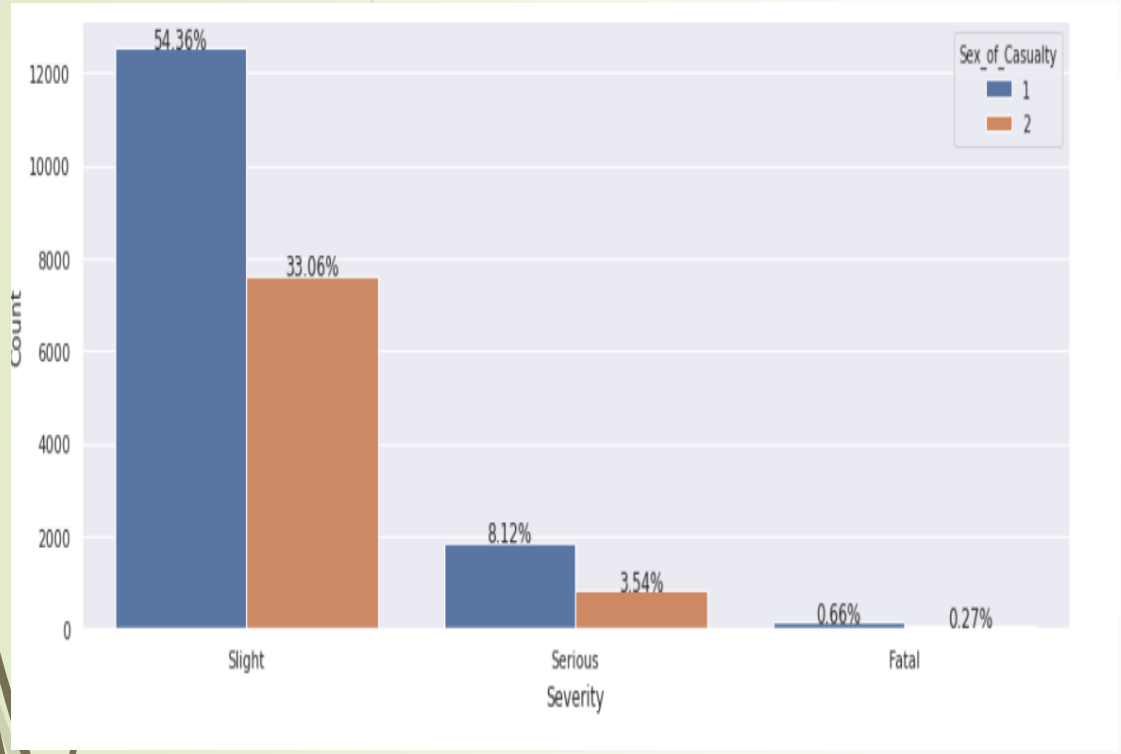
Relationship between Vehicle and Accidents



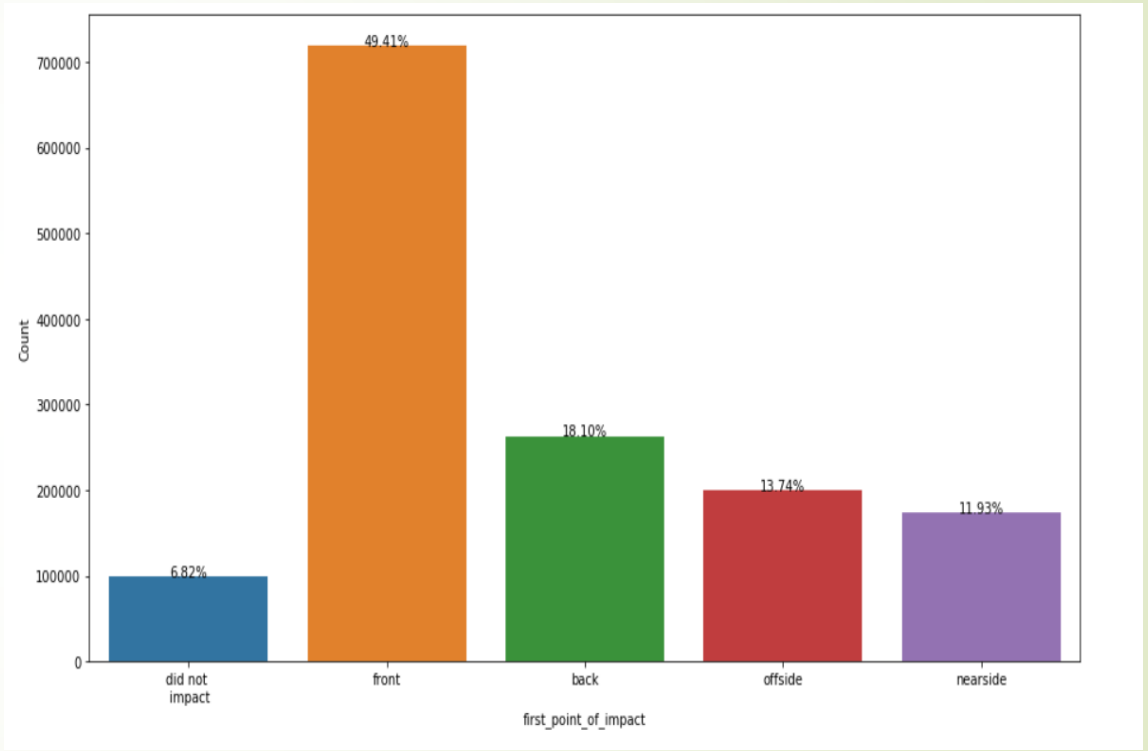
Bar Graph showing the Vehicle Types that have been met with an Accident

Bar Graph showing the Vehicle type and their Engine Capacity

Severity and Impact of Accidents



Bar Graph showing Severity of Accidents on the basis of Gender



Bar Graph showing the point of First Impact of an Accident

Classification Model

CLASSIFICATION REPORT

	precision	recall	f1-score	support
0	0.99	1.00	0.99	9854
1	0.20	0.04	0.06	138
accuracy			0.98	9992
macro avg	0.59	0.52	0.53	9992
weighted avg	0.98	0.98	0.98	9992

CONFUSION MATRIX

```
[[9834   20]
 [ 133    5]]
```

It can be seen that class 0 has a higher precision than class 1 and better recall as well. The support is the number of occurrences of the given class in the dataset, since it is 9854 of class 0 and 138 for class 1, it means that the dataset is imbalanced.

The confusion matrix shows that this model predicted that 9854/9992 people did not suffer fatal injuries when there were actually 9967/9992 people not suffering fatal injuries. This model has an accuracy of 9839/9992 and the recall of this model is equal to 9834/9967, that is, 0.98%.

Result

Drivers who met with an accident were mostly in the age band of 30-40 years and were males.

Weather conditions didn’t effectively contribute to the occurrence of accidents and resulted in slight severity with not much consequences..

The first point of impact during the crash is the front side of the vehicles.

Number of accidents taking place is mostly with cars and cars had low engine capacity with all types of fuel that could be one of possible reason for the greatest number of accidents.

Better data engineering is required.

Thank You !