Car Accident Severity Analysis: Seattle, WA

MILLER PALANISWAMY

IBM-COURSERA-CAPSTONE

15-SEP-20


gettyimages
25 YEARS
400tmax

Introduction

Preciousness of human life is above all. For an individual & the family dependent on him/her, their life and health is more important for survival. It can not be crippled just like that in a freak accident.

There are numerous car accidents happen across the world every day. All car accidents will create damage to cars involved or even worse, take lives. However Transport sector plays an indispensable role in our society.

What should we do to address this unavoidable but critical issue?

Data science can bring about an evolution to this issue, by analyzing the contributing factors/data that lead to car accidents including locations, environmental conditions, etc.

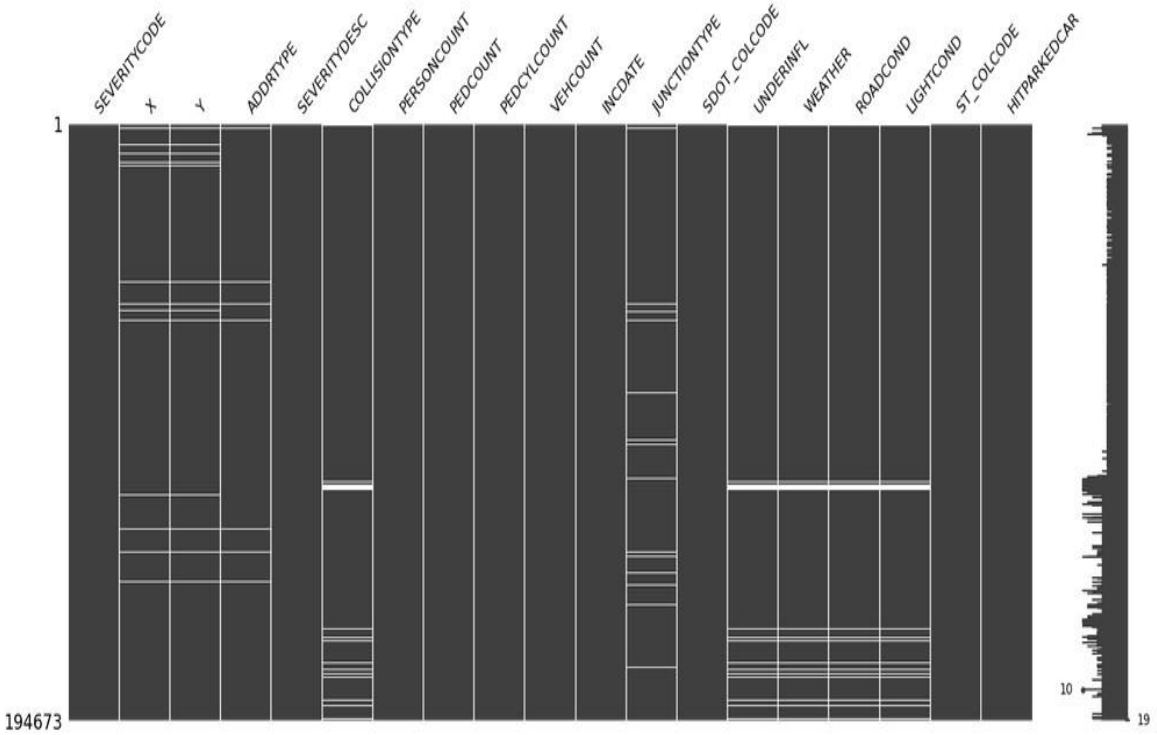
Therefore, it is advantageous for related government departments & stakeholders to accurately predict the severity of accidents under those conditions and take precautionary measures to reduce the number of accidents.

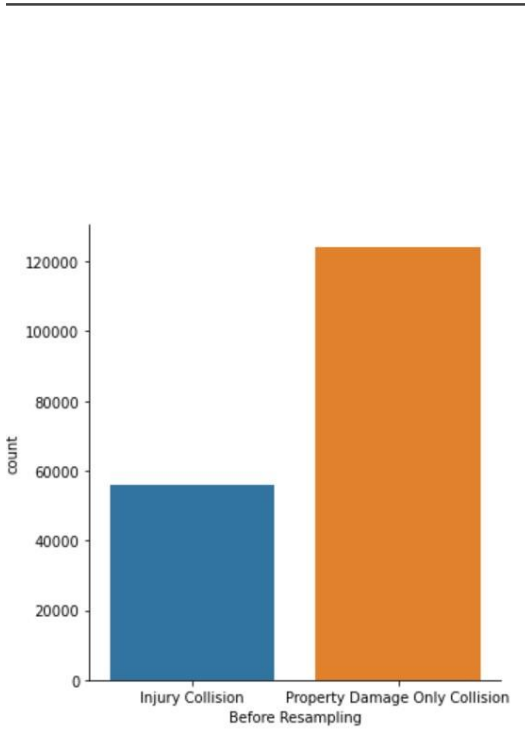
This project aims to predict the car accident severity in Seattle Metro, based on the data available.

Data Analysis - Preprocessing

```
msno.matrix(df3)
```

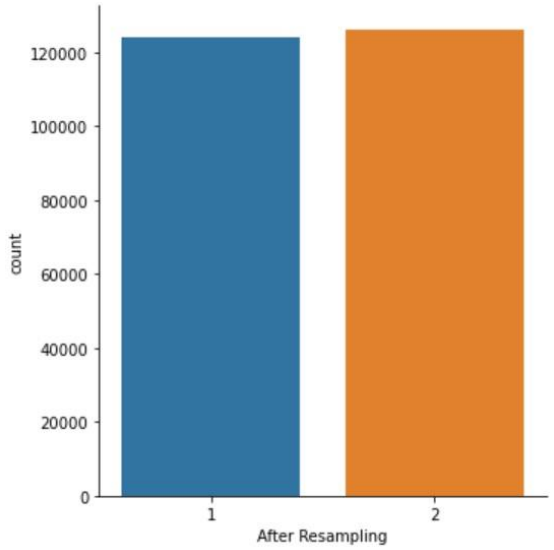
```
<AxesSubplot:>
```





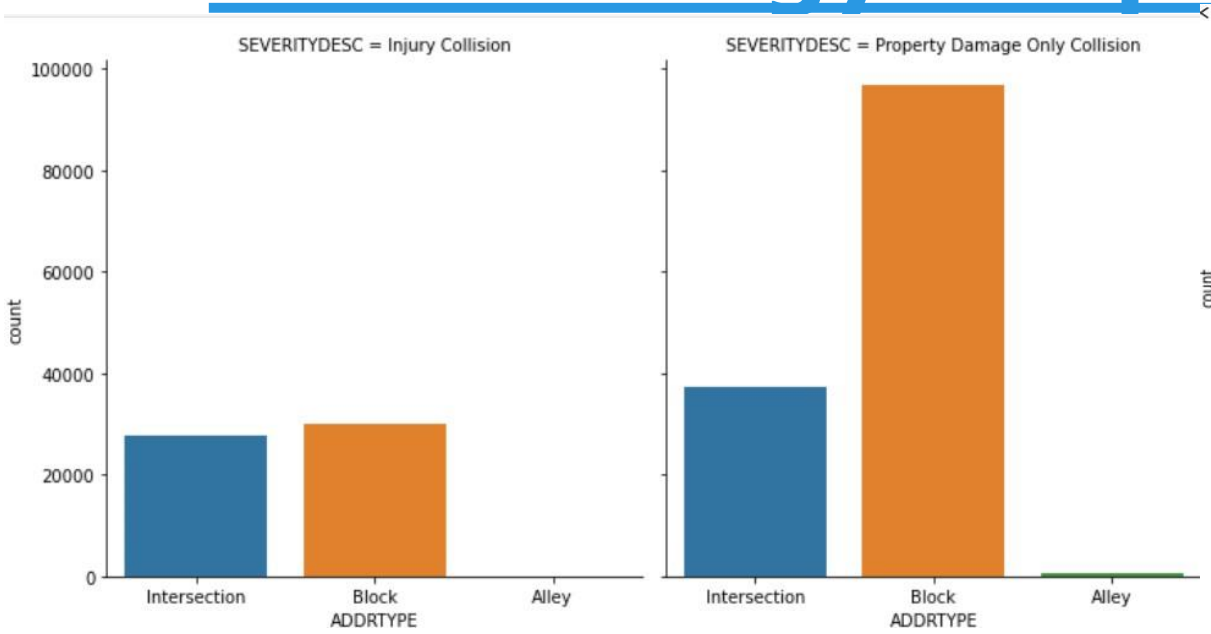

```
ax2 = sns.catplot(x = "SEVERITYCODE", data = df6, kind = "count")
ax2.set(xlabel='After Resampling')
```

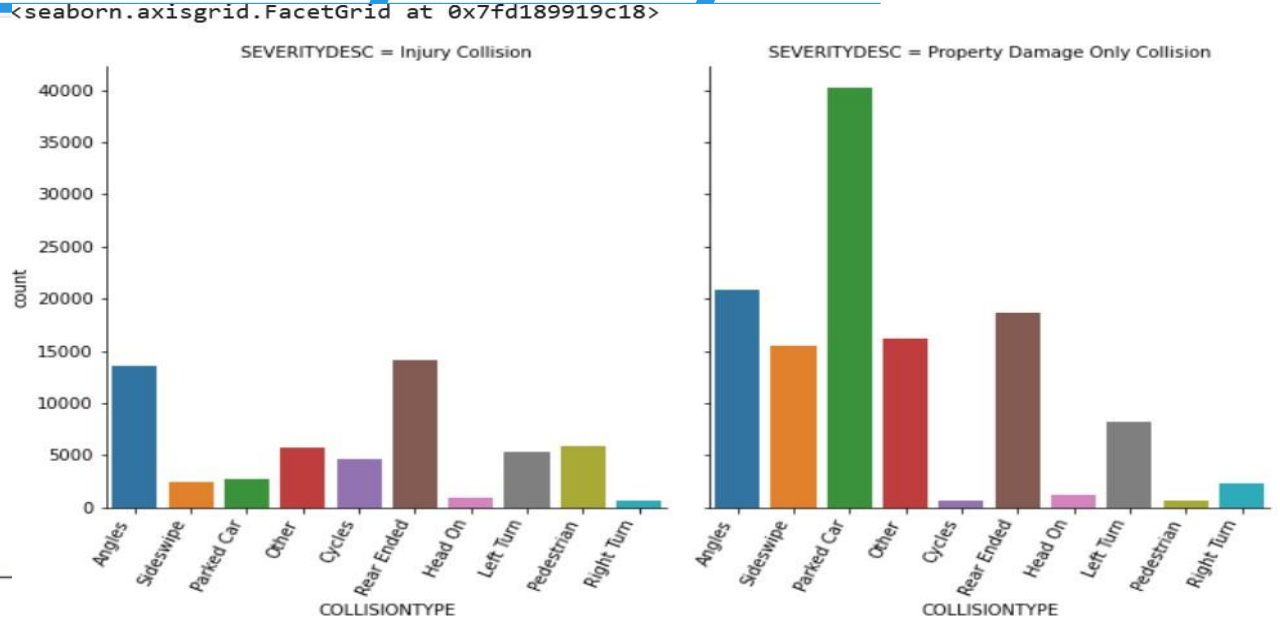
```
<seaborn.axisgrid.FacetGrid at 0x7fd17722a198>
```

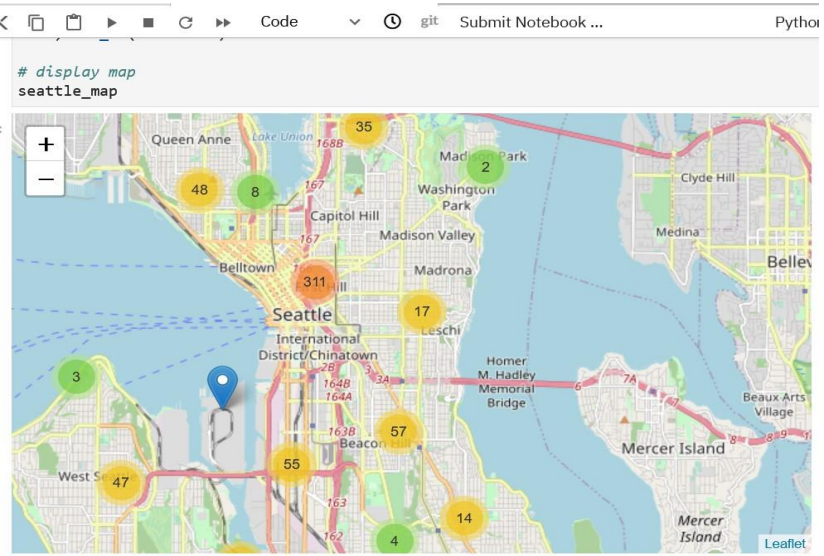


- After the resampling and preprocessing, there are 180067 objects in 22 columns which will be used for exploratory data analysis and modeling.

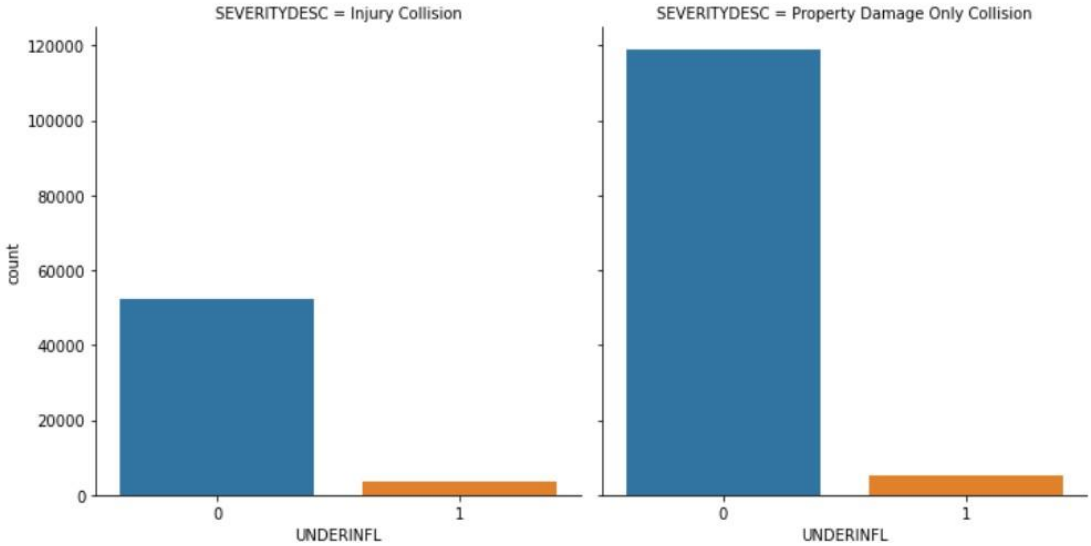
Methodology –Exploratory Analysis







- Please refer the report for the complete list of charts for different conditions.

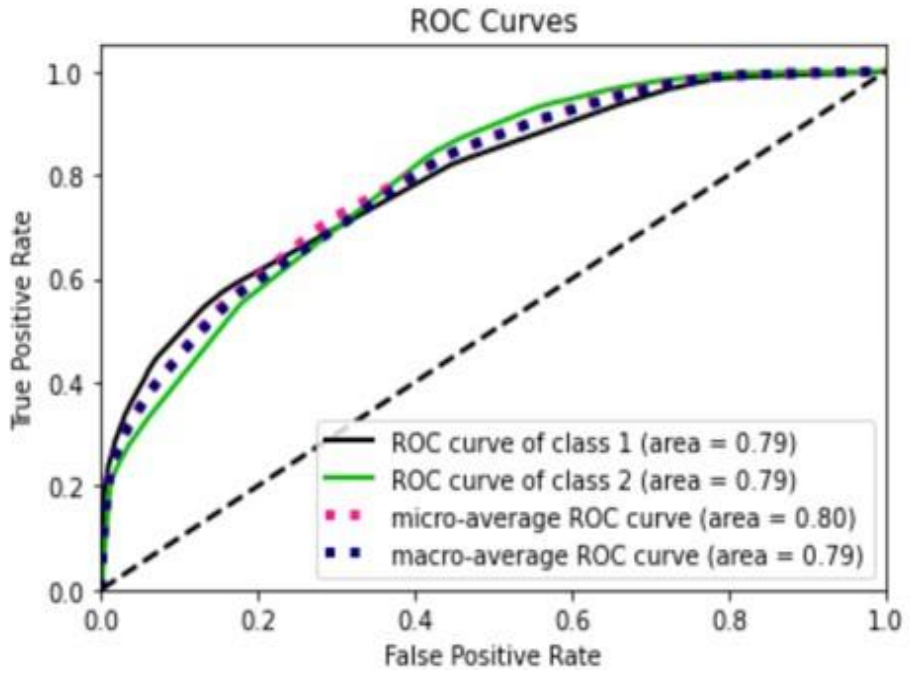


Results

4.1 The Performance of Decision Tree model

Accuracy of Decision Tree: 0.7103740071049375

F1-Score of Decision Tree: 0.70518985924198 64

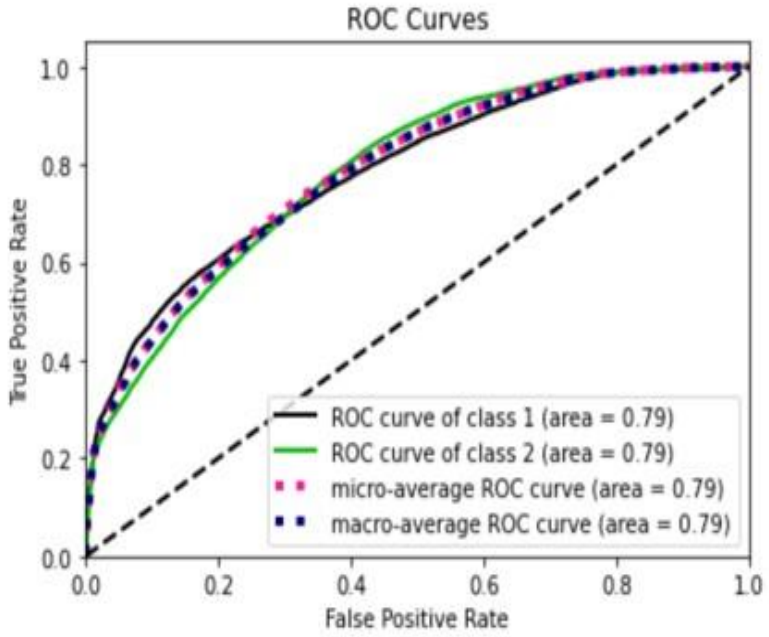


4.2 The Performance of Logistic Regression model

Logistic Regression's Accuracy: 0.7038012746311154

Logistic Regression's F1-Score: 0.7004139410689796

Logistic Regression's LogLoss : 0.5445183510374939



- Results from both methods are similar in predicting the severity of accidents, the Accuracy and F1 score are about 0.7, AUC is about 0.8, which is high for a machine learning model.

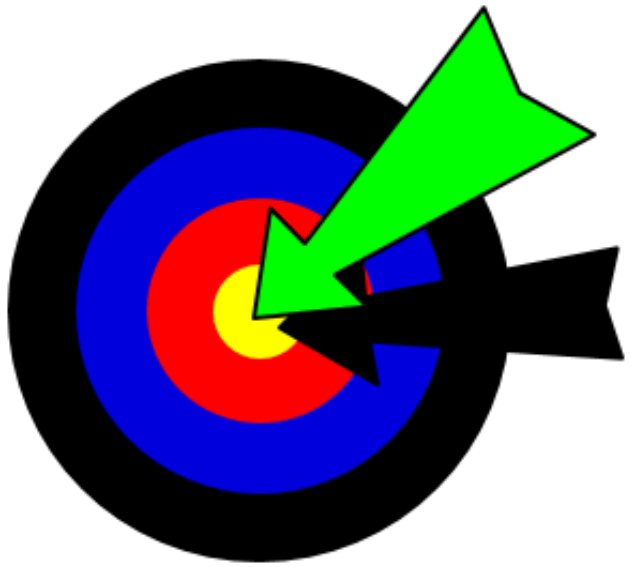
Observations

- Collision accidents happened at intersections have led to human injuries rather than the collisions at block which are high but mostly property damages only.
- As per the recorded data on accidents, environmental conditions such as weather, light & road are not critical factors, as surprisingly the number of accidents are high even in good weather and clear road.
- It is observed that more accidents are on the 4th workday of the week, i.e Thursday which may be due to the high work stress of the drivers/people.
- Collisions on Saturdays are less since people may be taking rest at home.
- More accidents happened downtown, where population density is high.
- Accidents happened along state highways, because lots of people drive along these roads or live close to these roads.
- With the data we have, we can predict the severity of an accident with about 70% accuracy.

* For complete analysis and briefing please check the report.

Conclusion

- This analysis not only will help the general public, but other concerned authorities could also see this and find the areas that are red zones at specific locations and save precious lives. As a first step they can even put the barricades to alert the upcoming traffic.
- Other concerned authorities who keep the condition of road and light could mark the places which have problems and solve them.
- Stakeholders should improve the infrastructure of those communities to bring in those resources.
- In addition, communities with low number of resources but high collisions should be studied to see what sets them apart.



Discussion, Q & A

References

Seattle Metro Data from "https://www.macrotrends.net“

Seattle Car Owners data from "https://www.seattletimes.com“

Call accidents analysis from ASIRT "https://www.asirt.org“

National Highway safety information from "https://www.nhtsa.gov/“

WA - DOT "https://wsdot.wa.gov/"