Predicting Accident Severity

Rosa Ferrusca

Nov2020

Road Accidents Background

- Road injuries are within the 10 top causes of death worldwide, occupying 8th place in 2016 (WHO, 2020)
- Approximately 1.35 million people die each year as a result of road traffic crashes
- As a result of car accidents many people go through time spending, pain, family member losses and economic losses

Problem

- As part of government efforts to reduce deaths and economic impacts due to car accidents, we will look into car collision information and search for possibilities to **predict car accident severity** so this data can be used by governments or organizations in local communities, to provide live information on actual car accidents for people to take precautions or change their mobility decisions, avoiding time spending or possible new accidents cause by existing ones.

Data

- Seattle city with collision data from 2004 to 2020
- It contains 194,673 records with 38 attributes
- Total attributes kept after first cleanup for further analysis were 13 out of the initial 38.

SEVERITYCODE ADDRTYPE WEATHER ROADCOND LIGHTCOND SPEEDING HITPARKEDCAR

COLLISIONTYPE PERSONCOUNT PEDCOUNT PEDCYLCOUNT VEHCOUNT INCDTTM

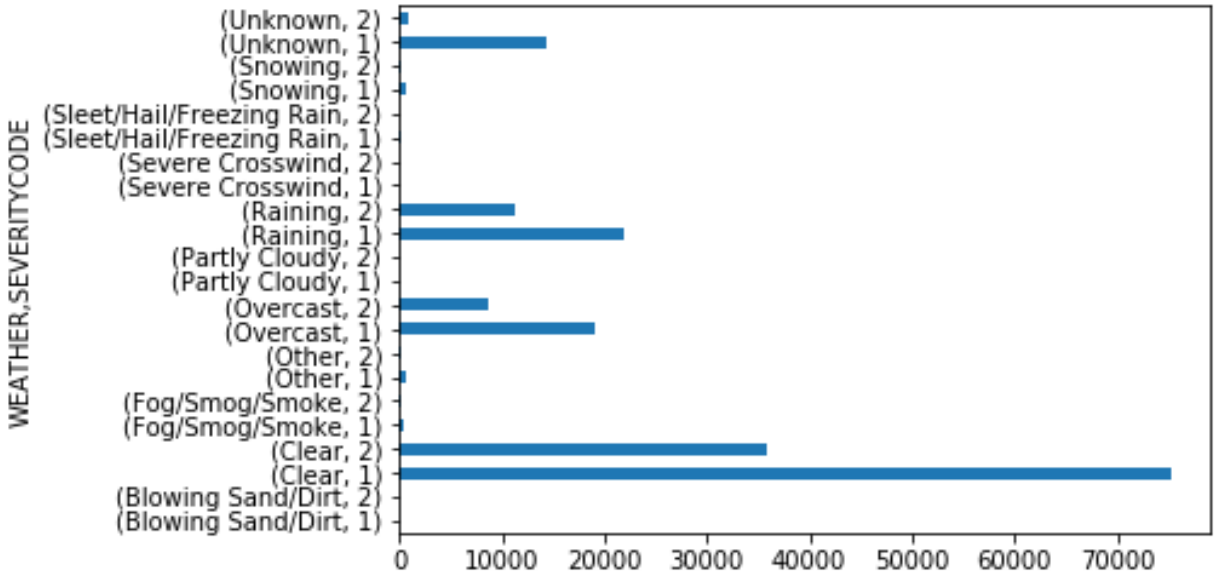
Methodology

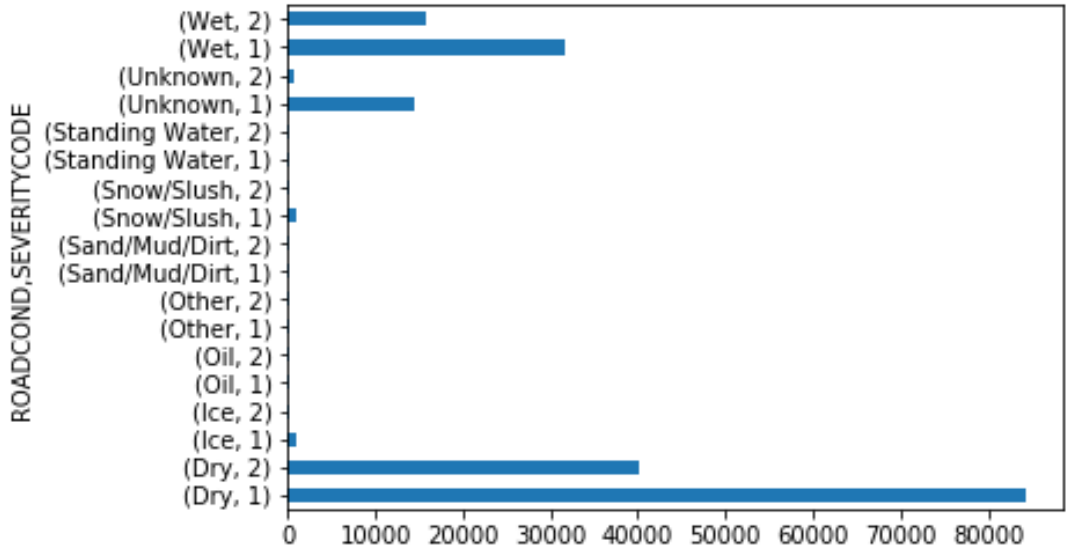
- For each attribute, frequency was plot to drop all the rows with no useful information. After analyzing the 13 attributes and how much representative data they provided for the accident severity, data was reduced to 143,561 records with 7 attributes with highest impact in information related to severity.

	SEVERITYCODE	ADDRTYPE	WEATHER	ROADCOND	LIGHTCOND	COLLISIONTYPE	DAYTIME
0	2	1	3	2	1	2	2
1	1	2	2	2	2	5	2
2	1	2	3	1	1	1	1
3	1	2	1	1	1	4	1
4	2	1	2	2	1	2	1

Results

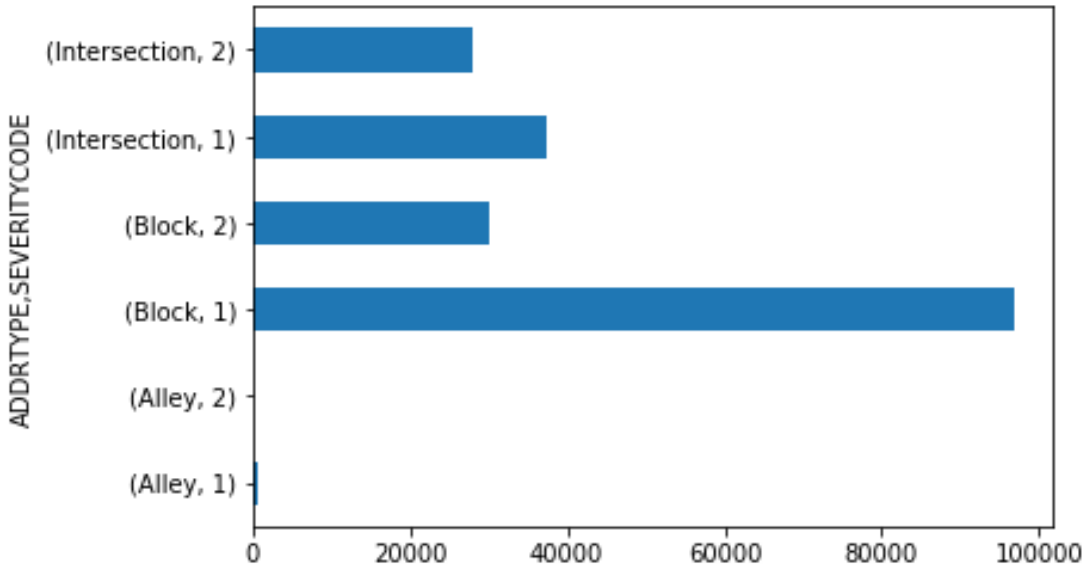
- Findings showed that car accidents' conditions are mostly the same for a severity that causes property damage and the ones ending in injury.

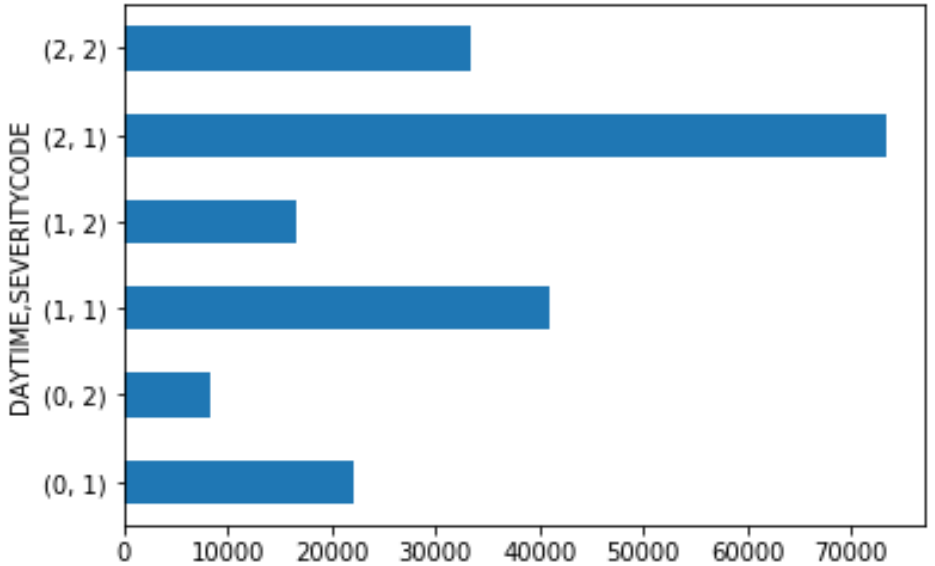




Discussion

- Even though all the final variables selected were all similar for the two severity types, there were 2 conditions that showed higher incidents that might caused injuries and be a good indicator for citizens to avoid the area.





Conclusion

- Based on the results showed by this study, a logistic regression prediction model with a 68% accuracy is good enough for our purpose of helping local governments inform their citizens with live information on actual car accidents, for them to take precautions and make better mobility decisions every day. This way, local communities have better results on efforts of avoiding increasing car accidents.