SEATTLE CAR ACCIDENT SEVERITY PREDICTION

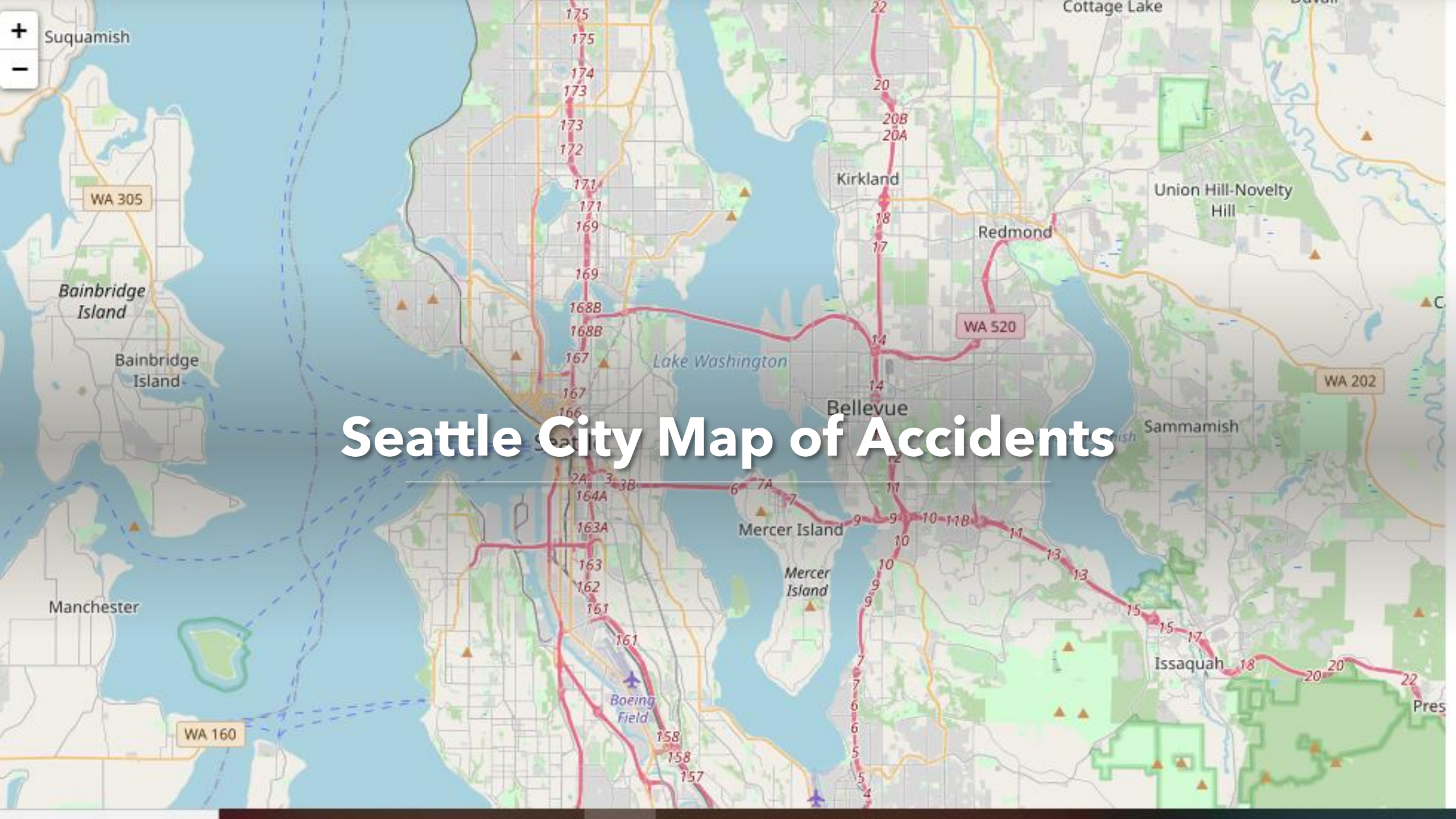
MANOJ KUMAR CHALAMALA

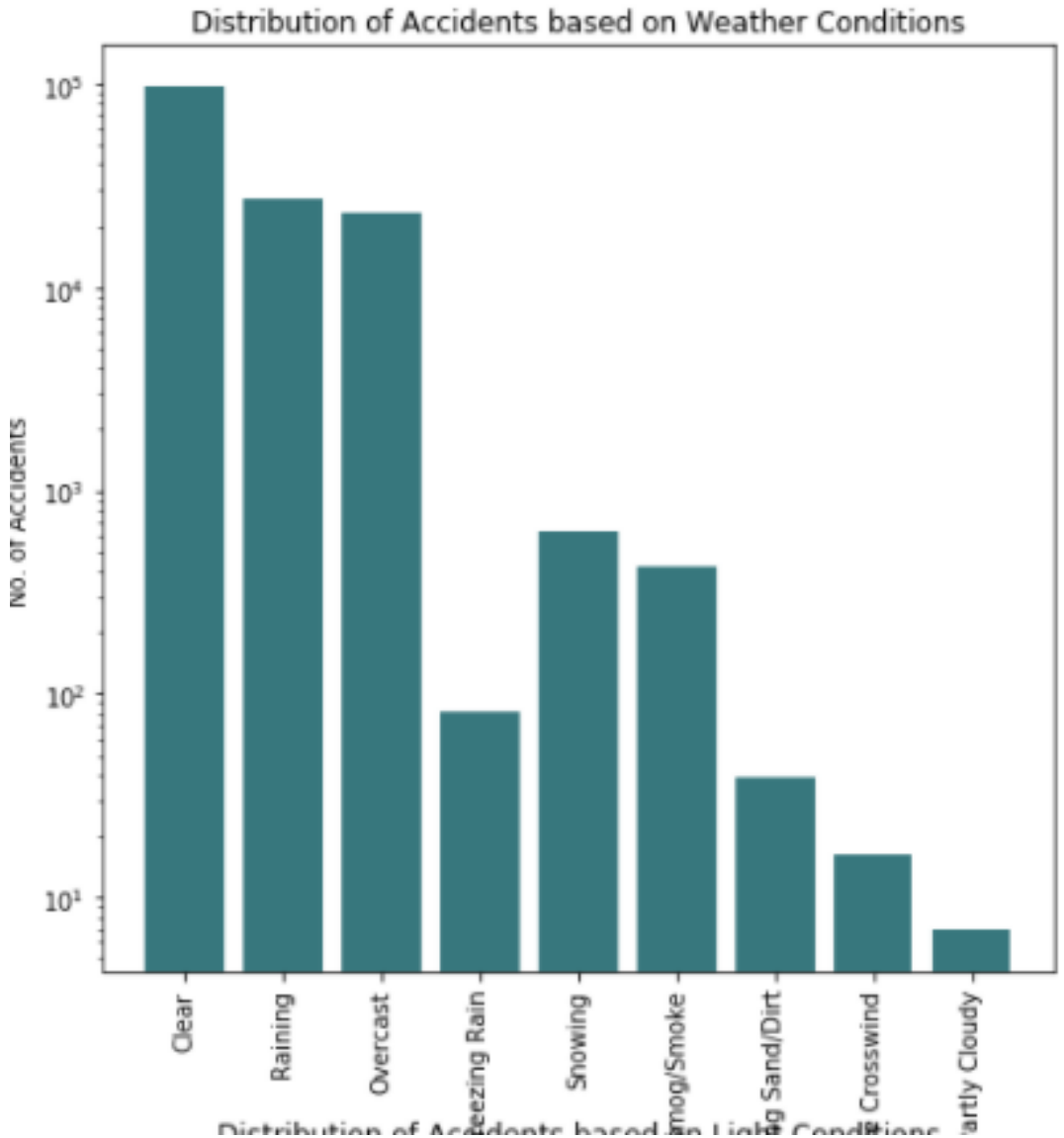
Predicting Accident Severity will reduce Fatalities

- It is estimated that road traffic accidents cost the United States' economy ~ $810 billion per year
- The main objective is to identify the key factors that determine the severity of accidents like Weather, Road Conditions, Light Conditions etc.
- Prevent or reduce the Severe or Fatal accidents in future by taking appropriate preventive measures.

Data Acquisition and Cleaning

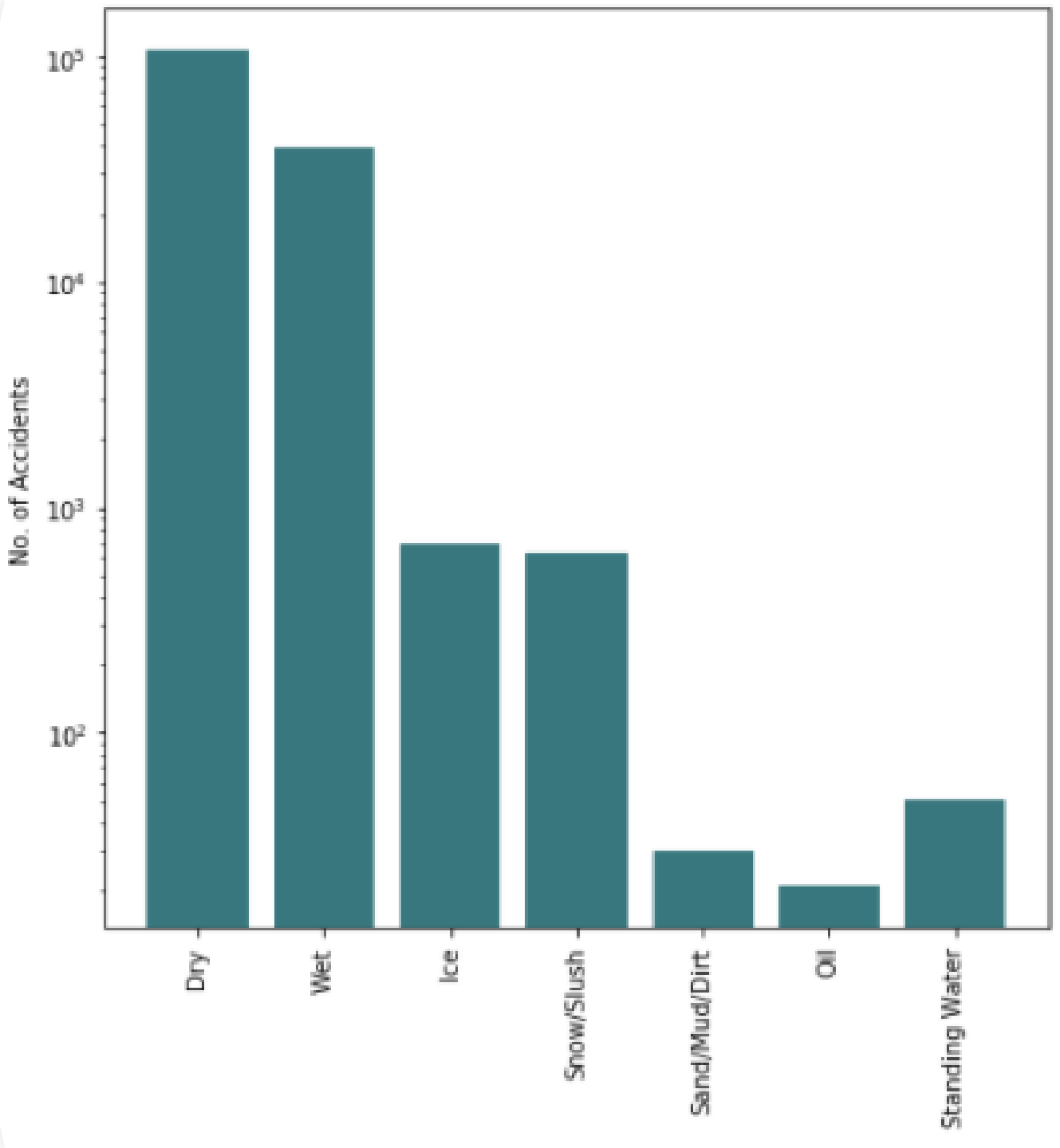
- Data is obtained from all road traffic accidents recorded in the Seattle municipal area between Jan 2004–Aug 2020 by the Seattle Department of Transport (SDOT).
- The Dataset contains 221738 rows (accidents) and 40 columns (attributes)
- Duplicate, highly similar or highly correlated features were dropped.
- Replace all Unknown and missing values

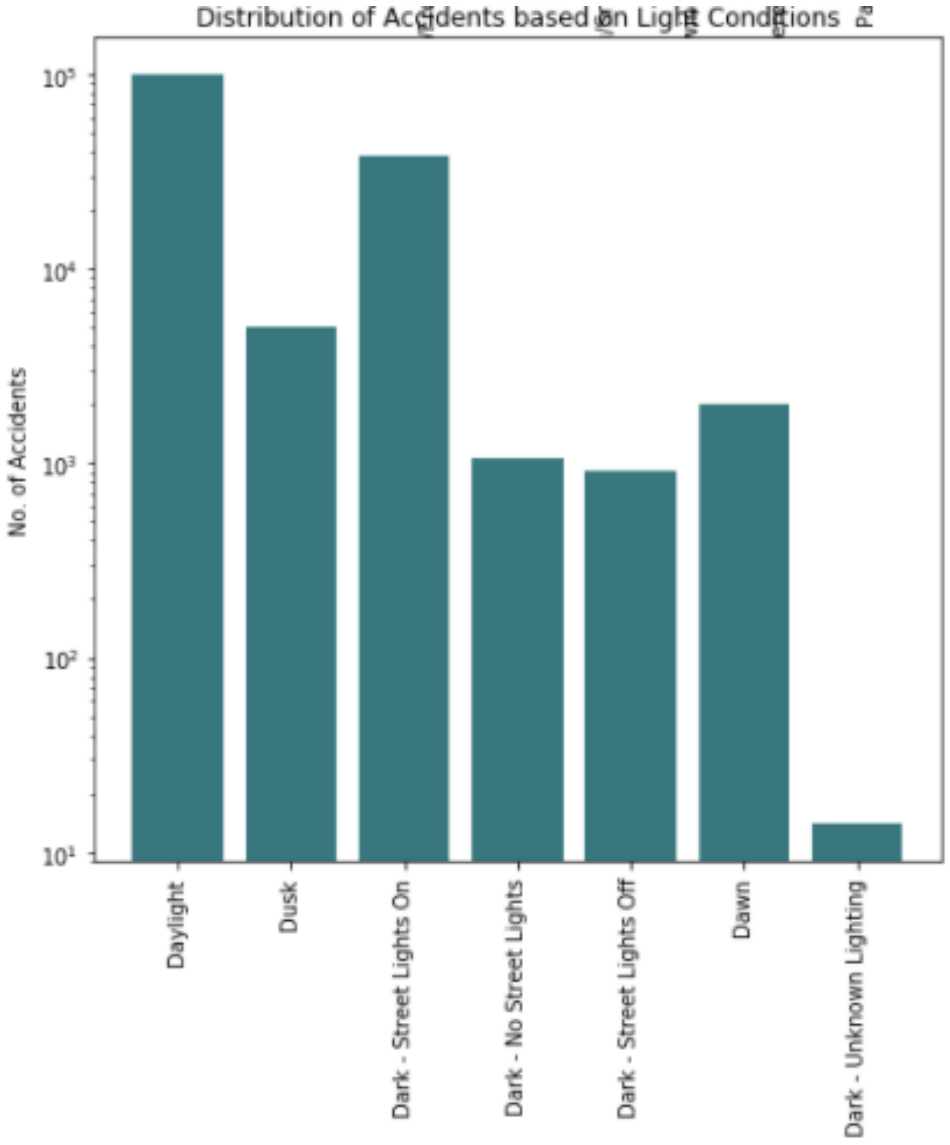

Seattle City Map of Accidents
Suquamish
Cottage Lake
WA 305
Bainbridge Island
Bainbridge Island
Kirkland
Union Hill-Novelty Hill
Redmond
WA 520
Lake Washington
WA 202
Bellevue
Sammamish
Mercer Island
Mercer Island
Manchester
Boeing Field
Issaquah
WA 160



Most accidents (75.6%) occurred in clear or overcast (i.e. dry) weather conditions. The remaining 24.4% took place either in severe conditions (such as severe winds) or during periods of precipitation (rain, snow, fog, etc).

Road conditions at the time of each accident. Clearly the road conditions are related to the prevailing weather at the time (e.g. if there is rain, the roads are likely to be wet), however conditions are not wholly determined by the weather. For instance, 61 accidents occurred on roads where oil was present.





The light conditions at the time of each accident. 62.6% accidents occurred during daylight hours, while 26.2% of accidents occurred at nighttime in areas with streetlights (i.e. urban areas). The remaining 11.2% of accidents include those which happened at dawn/dusk, or on roads with no/faulty streetlights.

Model Performance and Results

Model	Precision	recall	f1-score	Jaccard Index
Decision Tree	1.00	1.00	1.00	1.00
Random Forest	1.00	1.00	1.00	1.00
Logistic Regression	1.00	1.00	1.00	1.00
SVM	1.00	1.00	1.00	1.00

Conclusion

- The accuracy of the classifiers is excellent, i.e. 100%.
- The model has trained well and fits the training data and performs well on the testing set as well as the training set.
- This model can accurately predict the severity of car accidents in Seattle

Future Work

In future, the model could be improved to predict the accident severity on a continuum running from 1–4, rather than simply predicting a binary accident severity of 0 (minor) or 1 (major).

In future, it may be worth revisiting this work and modelling the accident data in five-year chunks, to see if the features which best predict accident severity have changed over time.


Thank you