

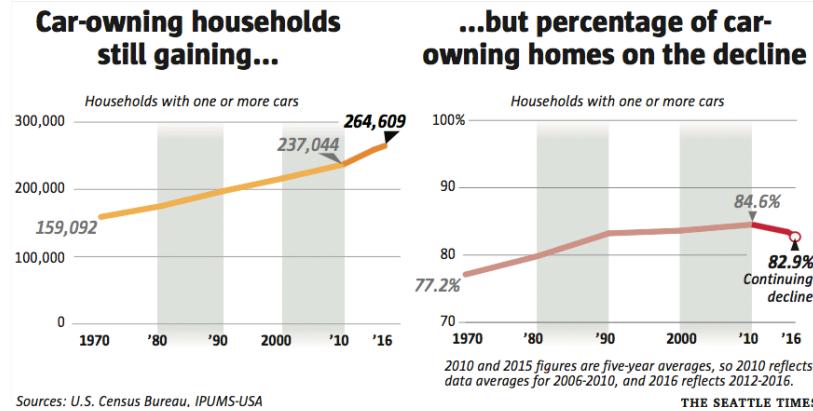
COLLISION SEVERITY PREDICTION

Based on Seattle Police Department Data Set



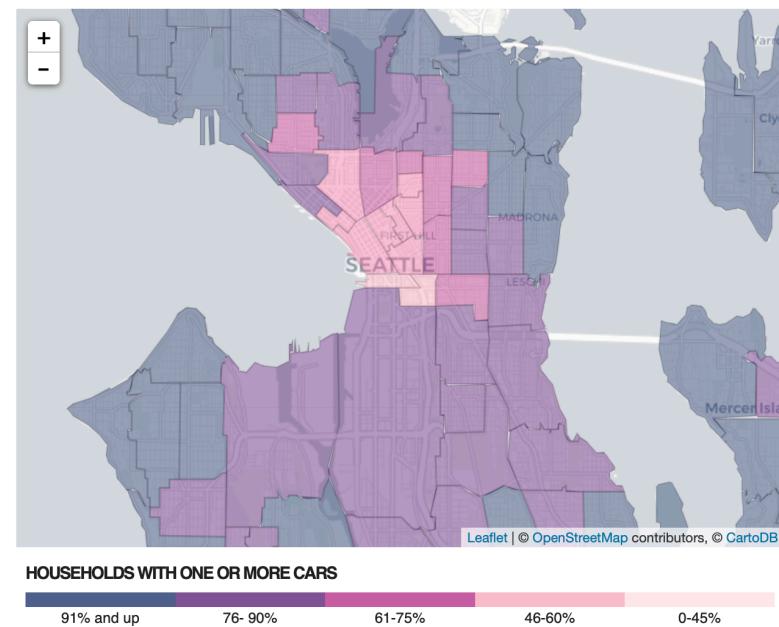
CONTEXT

- Seattle is the largest city in Washington with a population of 3.98M (US Census 2019)
- 444,000 cars, and increasing (Seattle News 2018)
- Motor vehicles crashes costs the country \$871B. (National Highway Traffic Safety Administration, 2014)
- Traffic accidents are a significant source of
 - Deaths, injuries and property damage
 - Congestion and delays



Where are our cars?

Downtown Seattle and the University District contain the only census tracts in the city where the majority of households don't own a car.



Source: U.S. Census

EMILY M. ENG / THE SEATTLE TIMES

BUSINESS VALUE

Knowing the contributing reasons to an accident can enable efficient deployment of public resources to mitigate the risks and probability of accident occurrence.



DATA SET

- Source Seattle Police Department since 2004 to present
- Consists 38 features and 194,674 rows of data
- A total of 18 features were retained and used in analysis and data visualization
- 20 features were dropped

FEATURE SETS RETAINED		
X	Y	ADDRTYPE
LOCATION	SEVERITYCODE.1	SEVERITYDESC
COLLISIONTYPE	PERSONCOUNT	PEDCOUNT
PEDCYLCOUNT	VEHCOUNT	JUNCTIONTYPE
INATTENTIONIND	UNDERINFL	WEATHER
ROADCOND	LIGHTCOND	SPEEDING

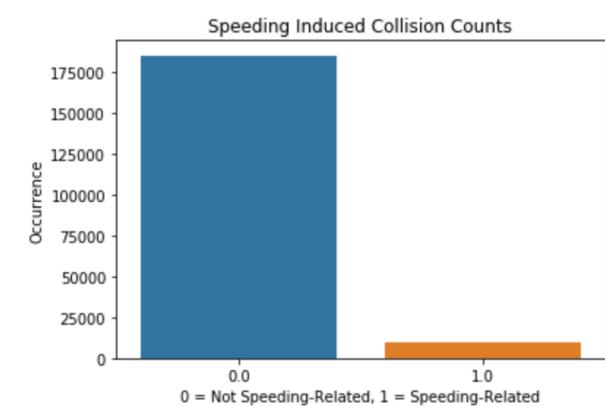
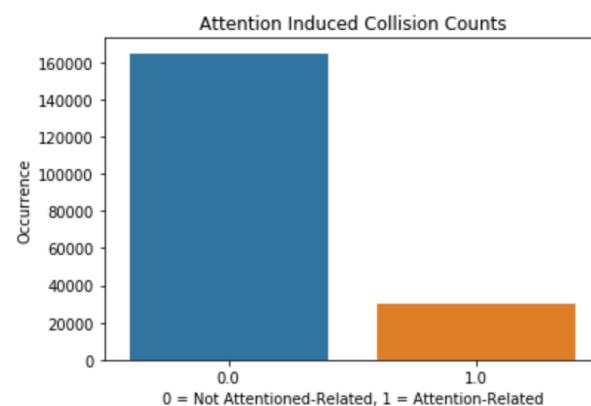
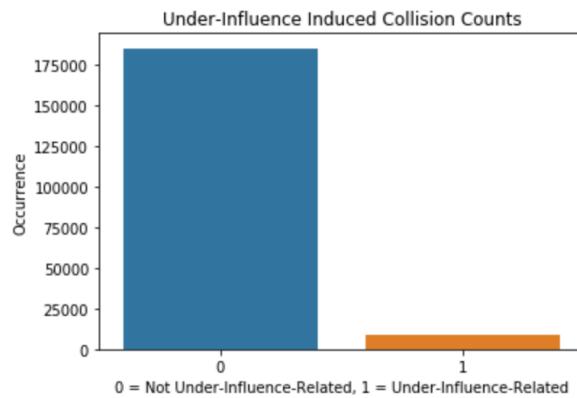
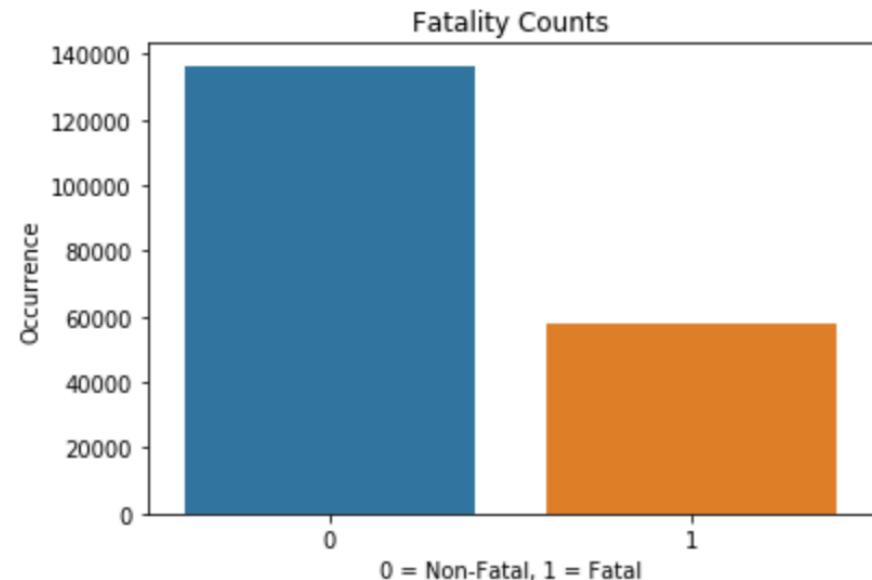
COLUMNS REMOVED		
OBJECTID	INCKEY	COLDETKEY
REPORTNO	STATUS	INTKEY
EXCEPTRSNCODE	EXCEPTRSNDESC	INCDATE
INCDTTM	SDOT_COLCODE	SDOT_COLDESC
PEDROWNOTGRNT	SDOTCOLNUM	ST_COLCODE
ST_COLDESC	SEGLANEKEY	CROSSWALKKEY
HITPARKEDCAR		



DATA VISUALIZATION

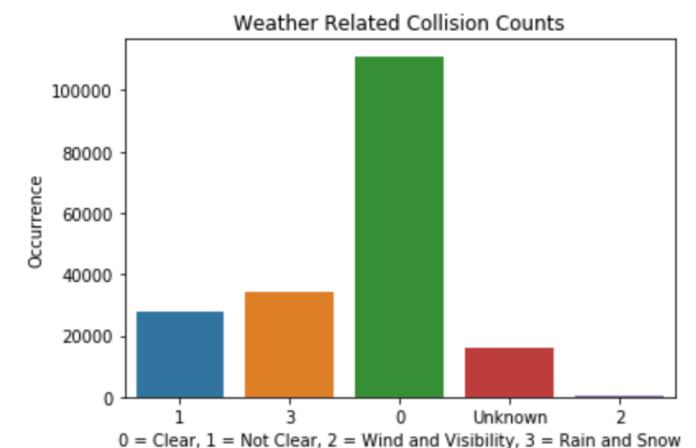
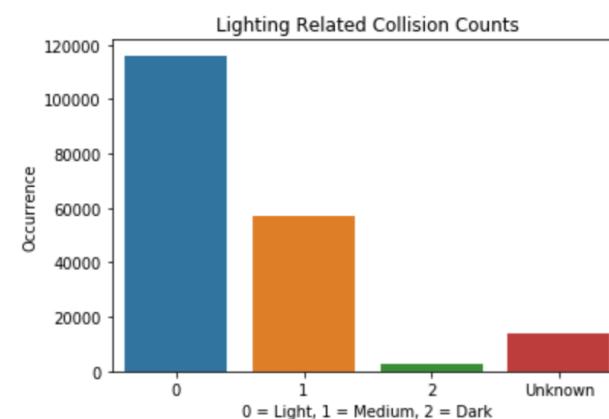
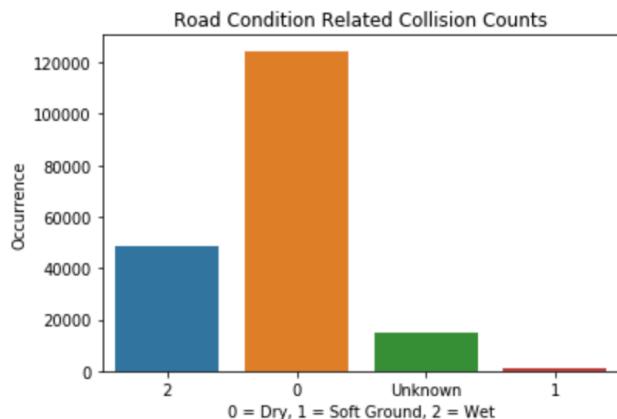
Observations

- Data set is imbalanced, with non-fatal cases outnumbering fatality 2:1
- UNDERINFL, INATTENTIONIND, SPEEDING, they are not significant contributions to fatality



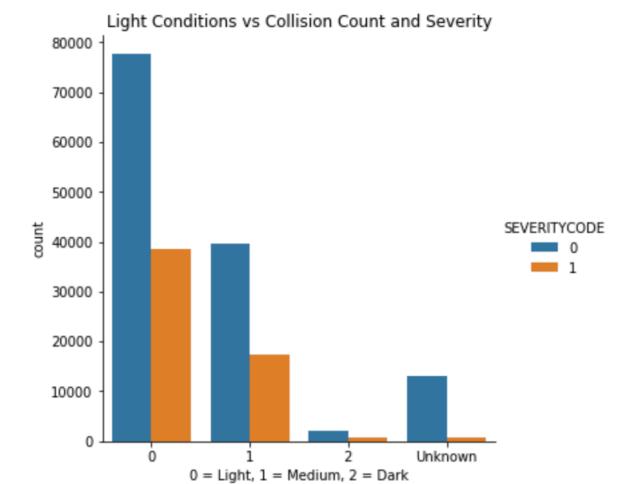
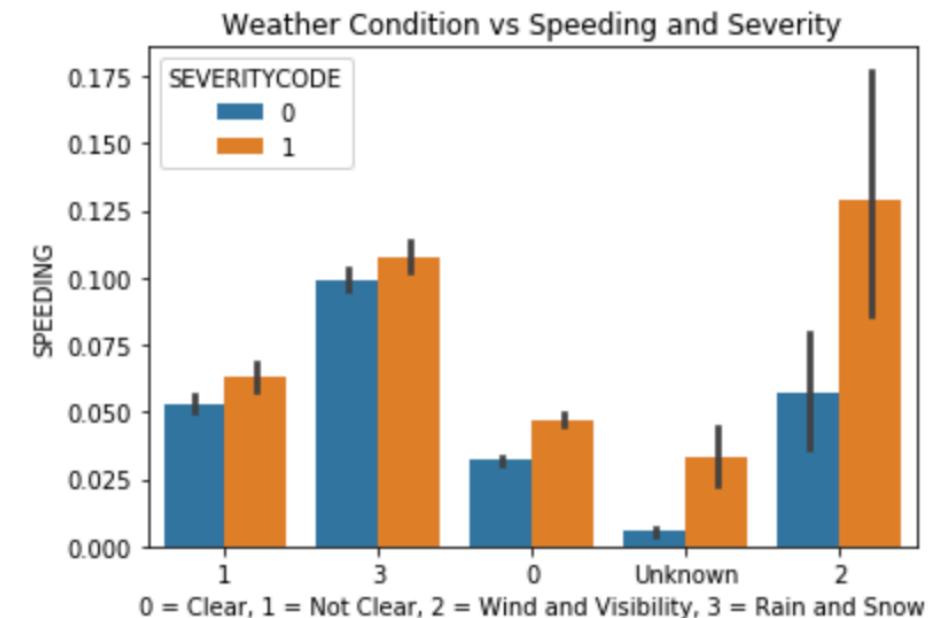
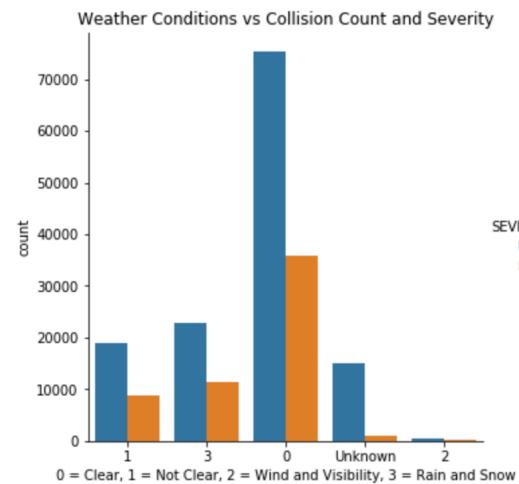
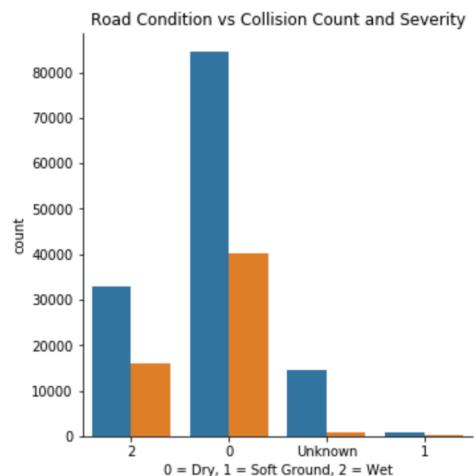
DATA VISUALIZATION

- Contrary to popular belief, most collisions happen when
 - Road condition is dry;
 - Weather is clear, and;
 - In daylight
- Significantly less accidents during night hours, low visibility and when road conditions are non-favorable.



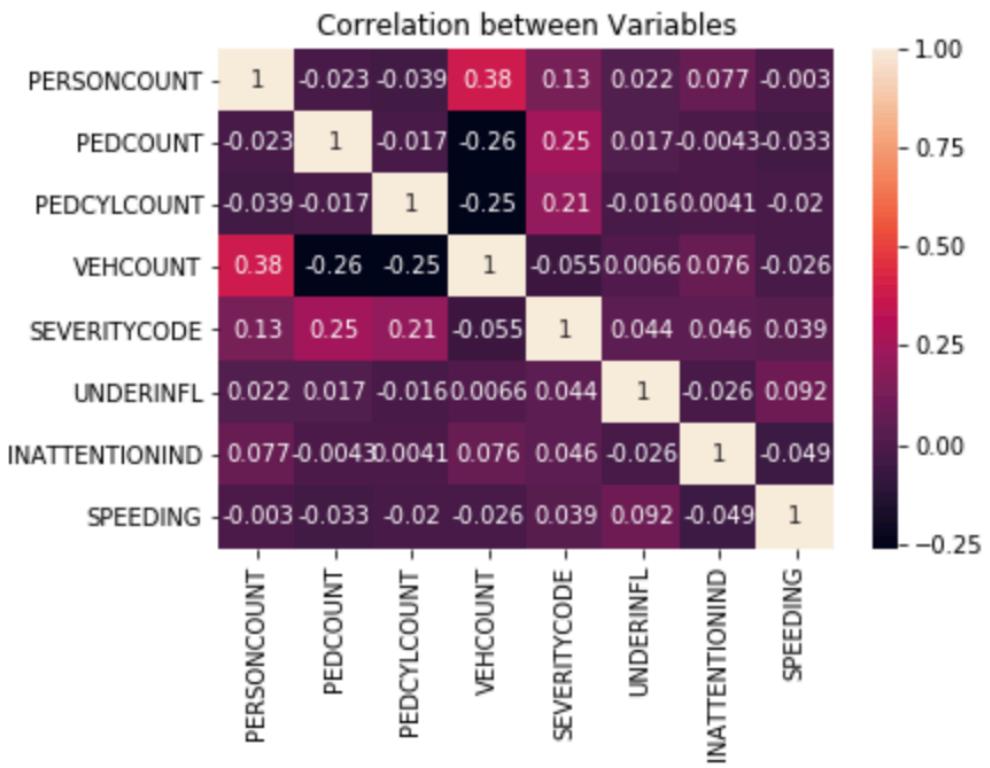
DATA VISUALIZATION

- Speeding had a significant impact on fatality rates when the weather condition was contributed by wind and low visibility.
- Most accidents happen during good conditions.



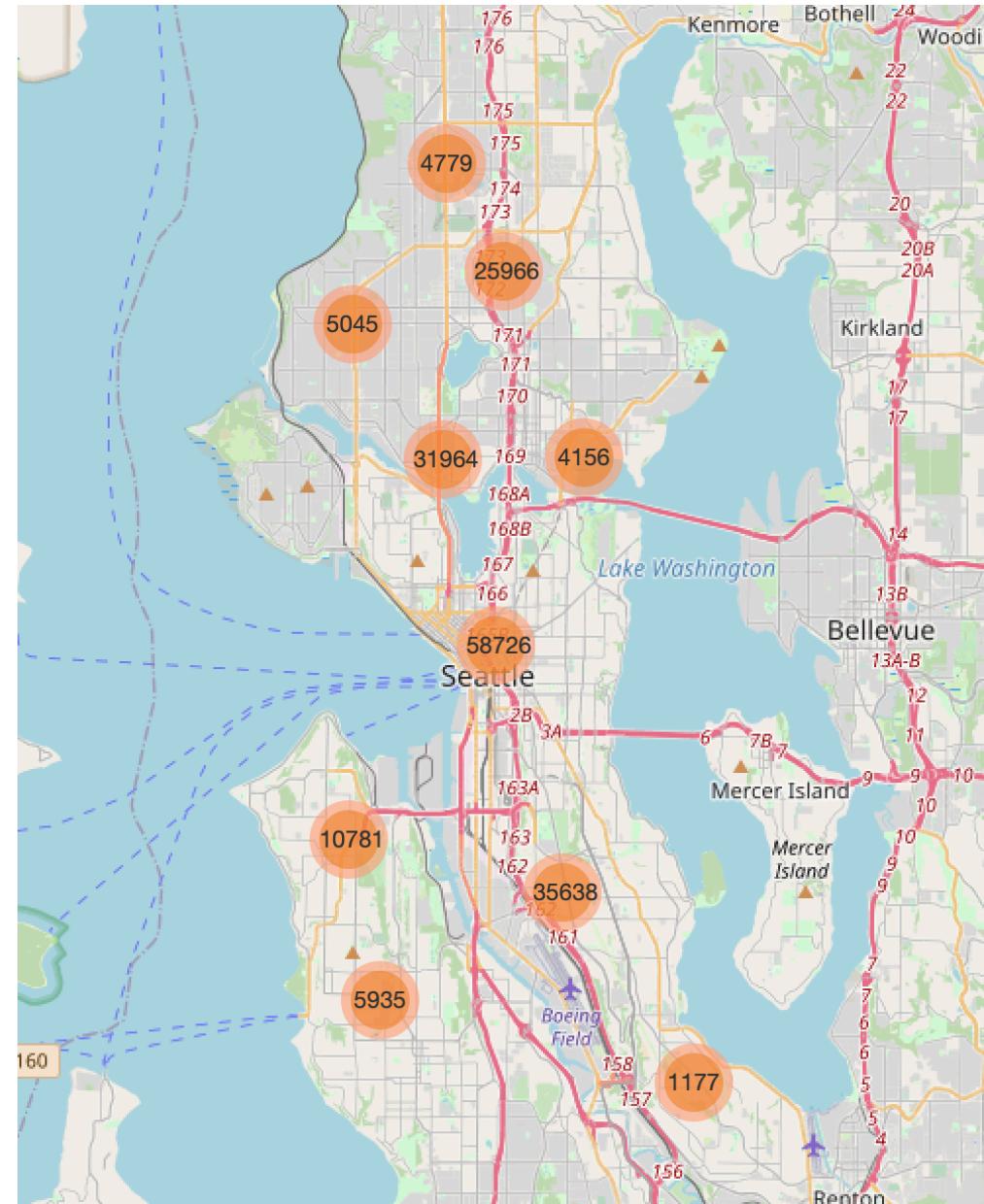
DATA VISUALIZATION

- There were weak correlations across the following variables.
 - PERSONCOUNT
 - PEDCOUNT
 - PEDCYLCOUNT
 - VEHCOUNT
 - SEVERITYCODE
 - UNDERINFL
 - INATTENTIONIND
 - SPEEDING
- The strongest correlation was only 0.38 between person count and vehicle count, which is a logical correlation anyway.



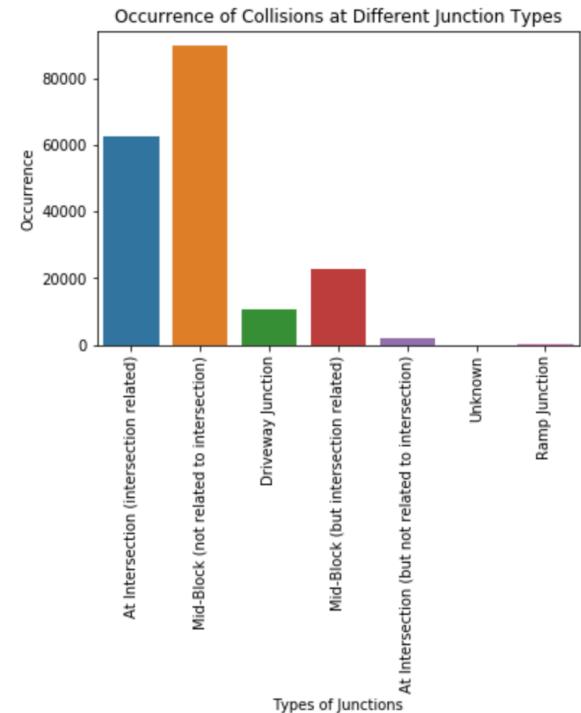
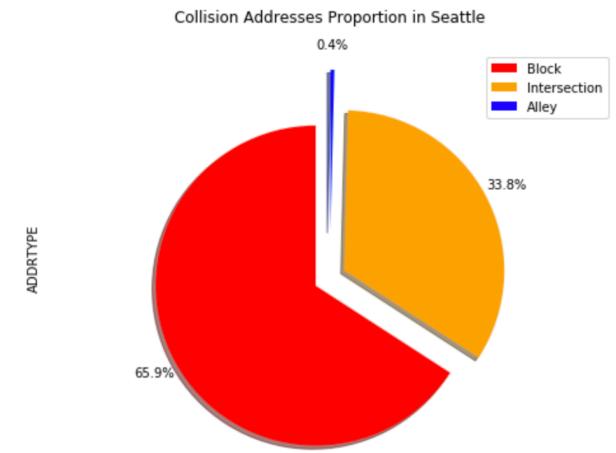
DATA VISUALIZATION

- Most accidents happen in vicinity of the freeway.
- Downtown Seattle boast the highest number at 58,726, likely brought about by higher road density and traffic volume.



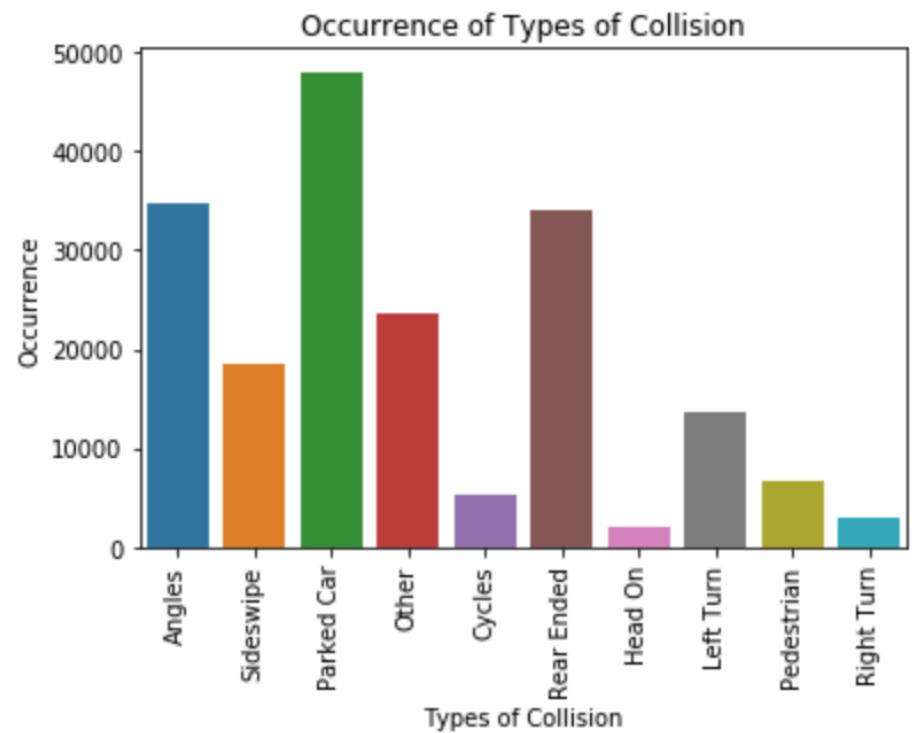
DATA VISUALIZATION

- Interestingly, a significant number of accidents happen at mid-block, with the second highest with intersection-related occurrences.
- Alleys have the lowest number of recorded accidents.



DATA VISUALIZATION

- Another interesting insight is that there are more accidents involving parked cars than other types of collisions.



DATA MODELLING

- The data set was modelled using the following four different models
 - K-Nearest Neighbor
 - Decision Tree
 - Logistic Regression
 - Support Vector Machine



MODEL EVALUATION

Metrics	K-Nearest Neighbor	Decision Tree	Logistic Regression	Support Vector Machine	
Solver	K = 2	D = 7	All solvers	Others	Sigmoid
Jaccard	0.699	0.699	0.699	0.699	0.648
F1	0.823	0.823	0.823	0.823	0.602
Log Loss			0.611		

- The models performed the same for the input data set and test data
- Only SVM with Sigmoid kernel performed less optimally than the rest in comparison.



RECOMMENDATION

- **Drivers Education**

- Likelihood of accidents due to drivers' complacency and inattention on the roads
- Regular road safety campaigns
- Reduce the occurrence of drivers being subjected to speeding, inattention and driving under influence.

- **Improvement to Road Conditions**

- Ensure that there is sufficient drainage for surface runoffs during rain and snow
- Post-accident treatment of road surfaces to reduce the presence of oil

- **Improvement to Lighting Conditions**

- Street lightings should adhere closely to sunrise and sunset timings to enable smooth transition from light to dusk



RECOMMENDATION

- **Parked Cars and Parking Berth**
 - Enforcement of vehicles keeping within parking spaces
 - Checking that the roads are appropriately wide enough for vehicles to pass in vicinity of parked cars
 - Vehicles can also have a mandatory reflector at their corners to improve visibility
- **Placement of Warning Signs**
 - Road signages can be installed in locations of higher traffic accidents



WAY AHEAD

- Implement recommendations
- Monitor subsequent data related to future collisions to assess the effectiveness of recommendations

