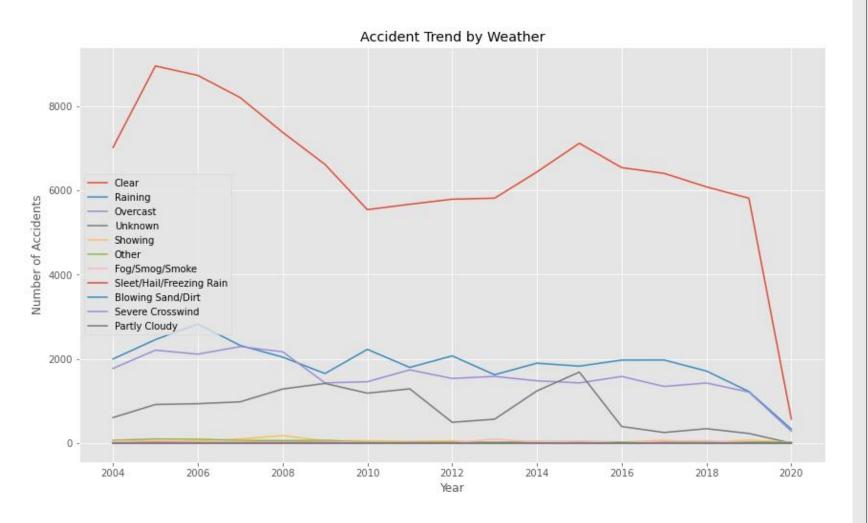REDUCING COMMUNTE TIMES

&

PREDICTING ACCIDENT SEVERITY

The value in reducing delays and predicting accident severity

- Reduced Delays:
  - Less overall congestion when vehicles are not idling at accident locations
  - Reduced overall fuel consumption
  - Reduced probability of follow-on accidents as a result of drivers trying to make-up for delays
  - Efficient commutes equal happy drivers ☺
- Predicting Accident Severity
  - More efficient dispatching of the correct response personnel and equipment
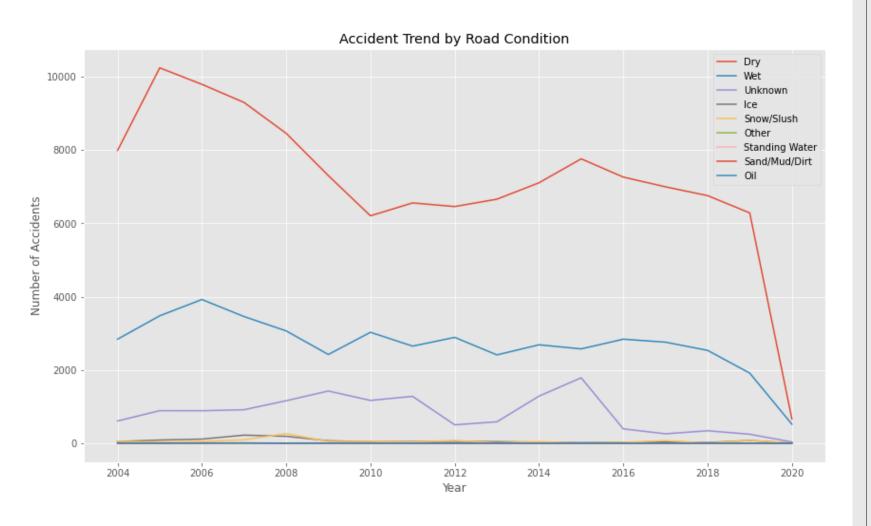 - Reduced traffic delays when accident sites are cleared quickly

Data acquisition and preparation

◦ The source data came from Seattle city as a csv file with 194673 accident spanning the years 2004 to 2020.

◦ Removed all incomplete accident records

◦ Removed all non-vehicle accident records

◦ Create three groups of data

  ◦ Group 1: the complete dataset once cleaned for use in showing trends
  ◦ Group 2: accidents from 2019-01 forward for use in training and testing prediction models
  ◦ Group 3: most recent six accident records for plotting on city map
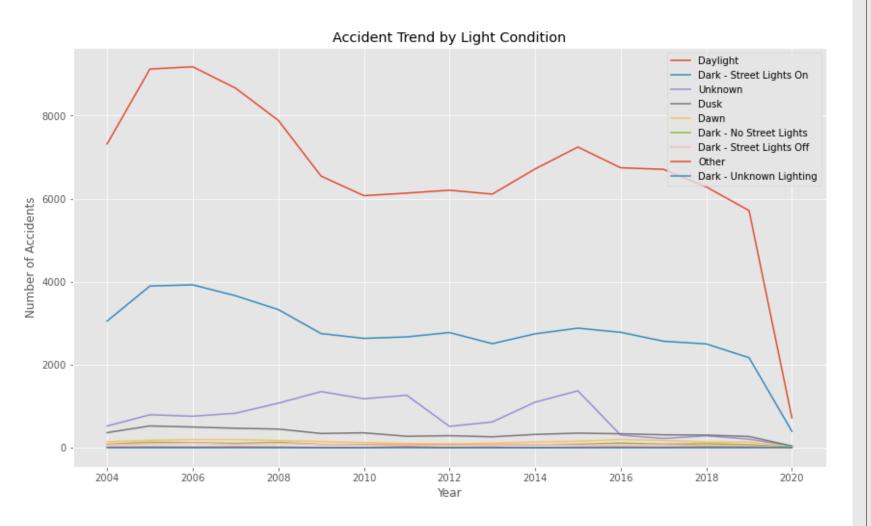
Accident Trend by Weather

Weather trend

Clear days have the most accidents and could be an input to the scheduling of first responders
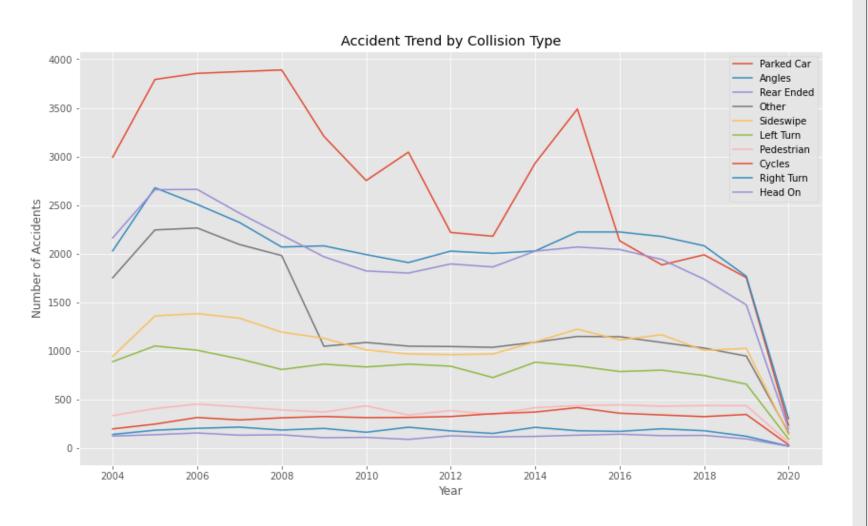
Accident Trend by Road Condition

Road Condition trend

Dry roads may lead to more people choosing to travel and may also indicate that there is a relationship with speed (not explored here)

Accident Trend by Light Condition
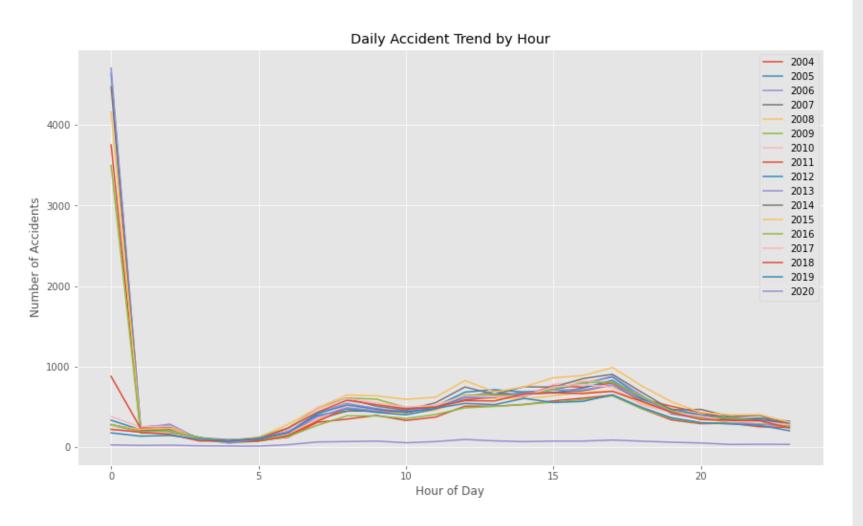
Light trend

Most accidents happen in daylight and there doesn't seem to be a change to the trend.

The Unknown category trend decreased sharply in around 2015 which seems to indicate an improvement in the accident recording process.

Accident Trend by Collision Type
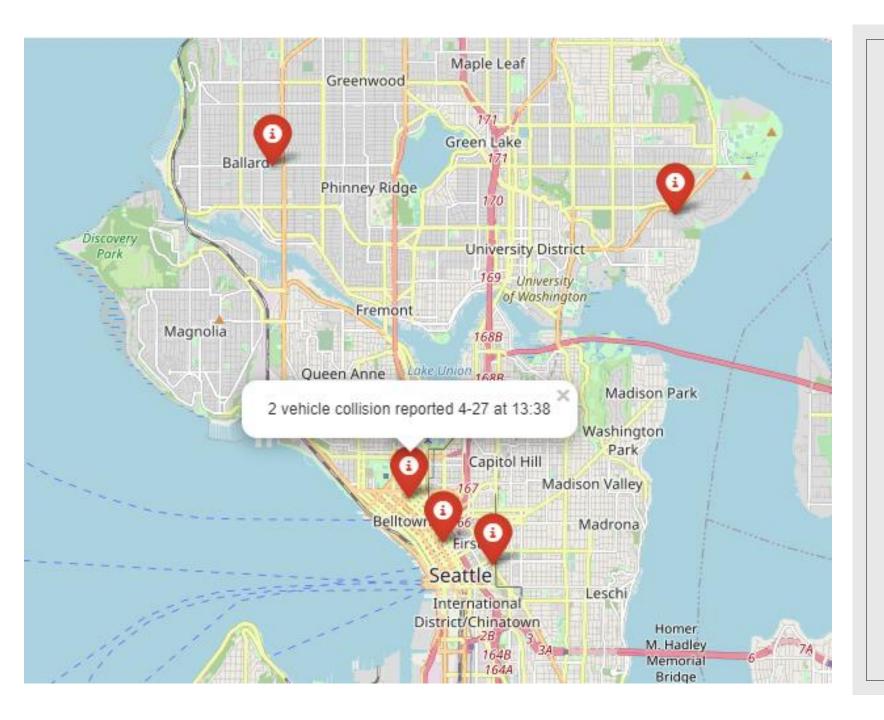
Collision trend

The top 3 types of collisions, in recent years, seems to have equal shares of the accident volume.

Daily Accident Trend by Hour

Daily trend

There are 3 distinct peaks during the day. 8am, noon and 5pm which closely correspond to the 'work day'. A 4th spike appears at midnight and would correspond to the commute home following a night out.
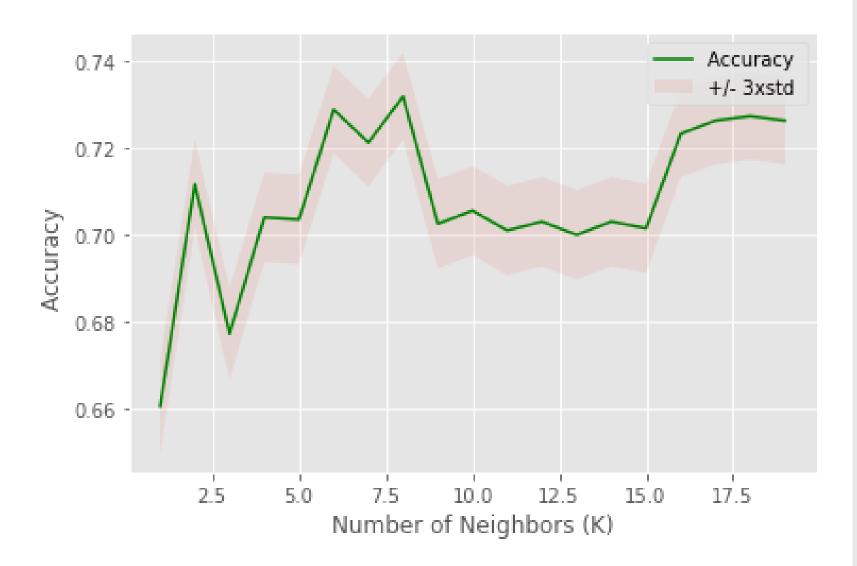
At all 4 peak times traffic volume is probably high.

Sharing Accident Information

Accidents are easily plotted on a city map that is available to all drivers via the city site.

Improved route planning for drivers when accidents can be avoided

KNN model

KNN provides the best model accuracy over other models (DT, SVM, LOG) and when comparing to 8 neighbors the severity prediction is 73%.

Model \ Accuracy	Jaccard	F1-score	Log
KNN (neighbor=8)	73.19	70.09	
Decision Tree	73.14	67.60	
SVM	73.19	68.73	
Log Regression	72.98	68.04	52

Model Evaluation

The accuracy across the four models show how close they are but KNN is slightly more accurate in determining accident severity.

- ◦ Models should be re-evaluated If additional severity codes start to be introduced in the accident collision data.

- ◦ Models are focused on vehicle collision data but can be easily modified to include non-vehicle collisions.

- ◦ Revisit the map refresh rates to ensure it's updated for perk periods

- ◦ Monitor traffic peak periods for any change that my impact the stakeholder peak usage periods

Conclusion