Analysis of Road Incidents in Seattle

IBM Data Science Capstone Project

Problem

- The problem at hand is to identify how certain conditions such as the weather, location and lighting affect the risks of driving. In particular, how likely are they to cause a severe collision.

- This information can be used for navigation system providers who would direct the flow of traffic to less risky areas under certain conditions.

- Navigation companies could also sell the tools to public transportation regulation branches in whose interest it is to increase road safety and reduce traffic incidents.

Data

- Given data is provided by SDOT Traffic Management Division and it is updated on a weekly basis. The data contains 194,673 observations of road incidents in Seattle with various features such as the severity of the incident, number of vehicles and people involved, conditions during the incident, time of the incident, etc.

Distribution of the incidents


Distribution of collissions
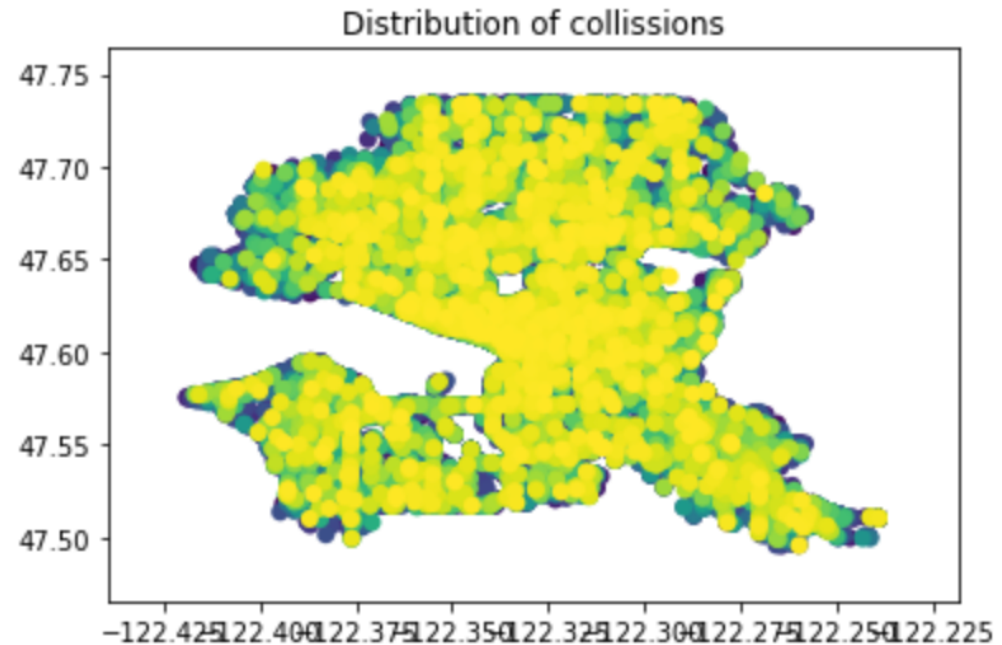
- From the heatmap, it is visible that the distribution of the collissions is quite even across the city with slightly more concentration in center of the city than in the outskirts.

Location Correlations

- From the data, it is evident that there are certain locations that are more prone to collisions than others. In particular, "BATTERY ST TUNNEL NB BETWEEN ALASKAN WY VI NB AND AURORA AVE N" had 276 collisions, most out of all recorded locations. It was closely followed by "BATTERY ST TUNNEL SB BETWEEN AURORA AVE N AND ALASKAN WY VI SB" with 271 collisions and "N NORTHGATE WAY BETWEEN MERIDIAN AVE N AND CORLISS AVE N" with 265.

- Vast majority of the collisions occurred in blocks. Half of that number was recorded at intersections while only a small percentage occurred in alleys.

- Most collisions (47,987) where recorded on parked cars. There were 34,674 collisions at angles and 34,090 accidents where the car was rear-ended.

Hour of the day

Number of collisions per hour of the day
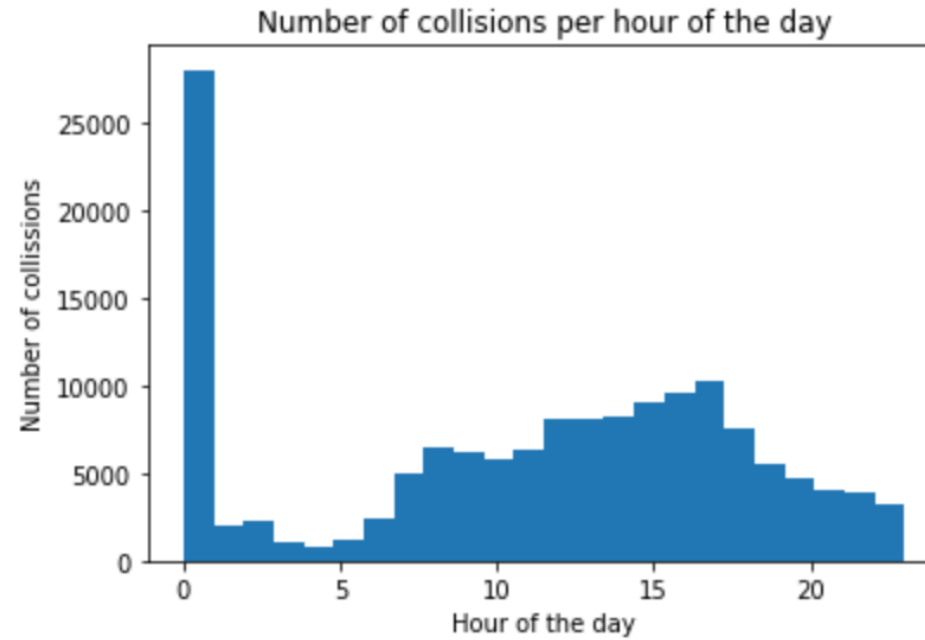
- There was a peak in the number of collisions from 3 to 5 pm (peak at 0 hour is due to null values being converted to 0s).

Hour of the day

Number of collisions per day of the week
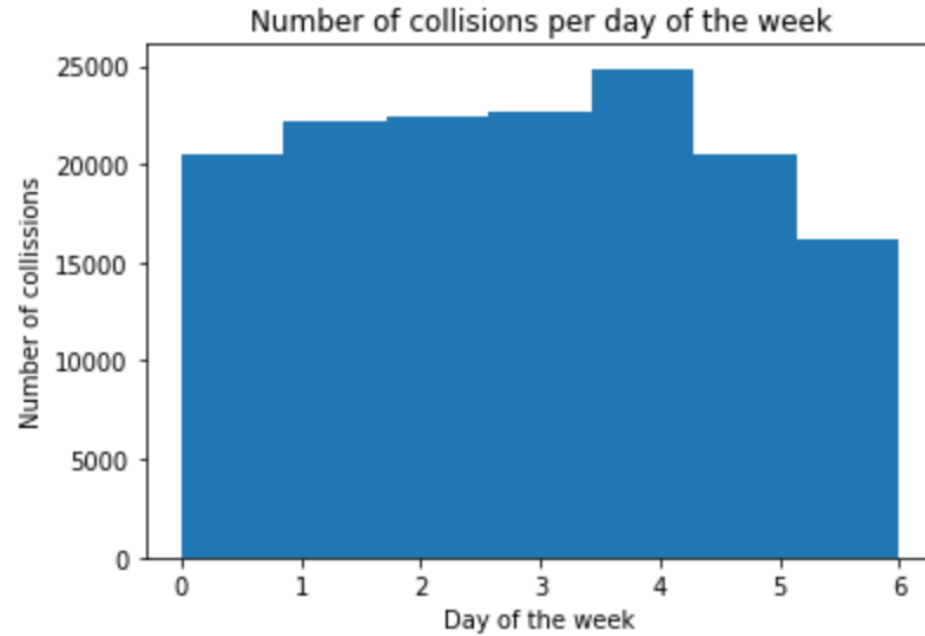
- For the days of the week, Friday was associated with most collisions while the number during the weekends were lower. The number of collisions increased during the weekdays but by a slight margin only.

Conditions

- When it comes to weather conditions, the share of severe collisions was higher during rain, fog and partly cloudy conditions. During snow, the collisions were list likely to be severe.

- For the Light Conditions, more severe collisions were observed during dawn, daylight, and dusk.

- It was also more likely for a severe collision to occur in dry and oily conditions.