Introduction: Business Problem

In this project I tried to show different types of crimes and collisions which have been occurring in Seattle for the last several years and map those crimes and collisions by neighborhoods. The project is also designed to identify the safest place to live in Seattle neighborhoods. In this project I also tried to see the distribution of crimes and car collisions spatially across different neighborhoods in Seattle. I also tried to show the most important venues across each neighborhood in Seattle using the foursquare api. This report targeted stakeholders interested in looking for a safe place in Seattle either to open a business or to find a house to live or to buy.

Data sources used

In order to accomplish this project I will use the crime data from the Seattle police department, the collision data from Seattle department of transportation. I also use the counties and neighborhood data from the Seattle gov.org site and the different venues using the foursquare api. Using the foursquare api and geopy data to map top ten venues for all Seattle neighborhoods and clustered it in groups and to map

the location of crimes and collisions in Seattle neighborhoods. For this project I have used the following data from different sources.

The Seattle crime report data from the Seattle police department

df = pd.read_csv('/Users/alelignfaris/Desktop/spd_reports.csv', index_col=None)

The Seattle collision report data from the Seattle department of Transportation

df = pd.read_csv('/Users/alelignfaris/Desktop/collisions.csv', index_col=None)

And I also use the counties and neighborhood data venues around each neighborhood from the foursquare api.

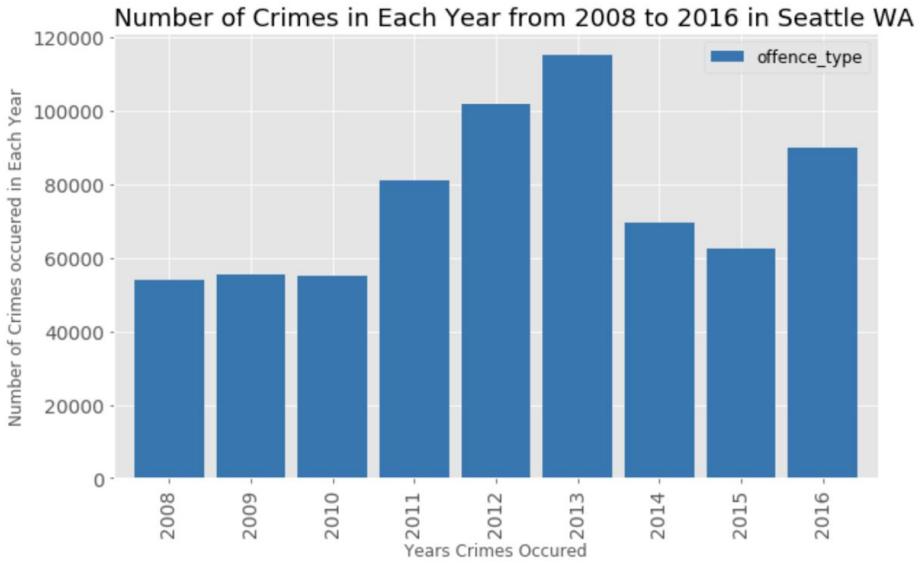
Methodology

In this project I have used different types of data science methodologies in order to finish my project. First I collected the crime, collision and Seattle neighborhood data from different sites in csv format. After collecting the data I have used different data cleaning techniques to clean the data and bring it in the format which is ready for further analysis using different data science tools.

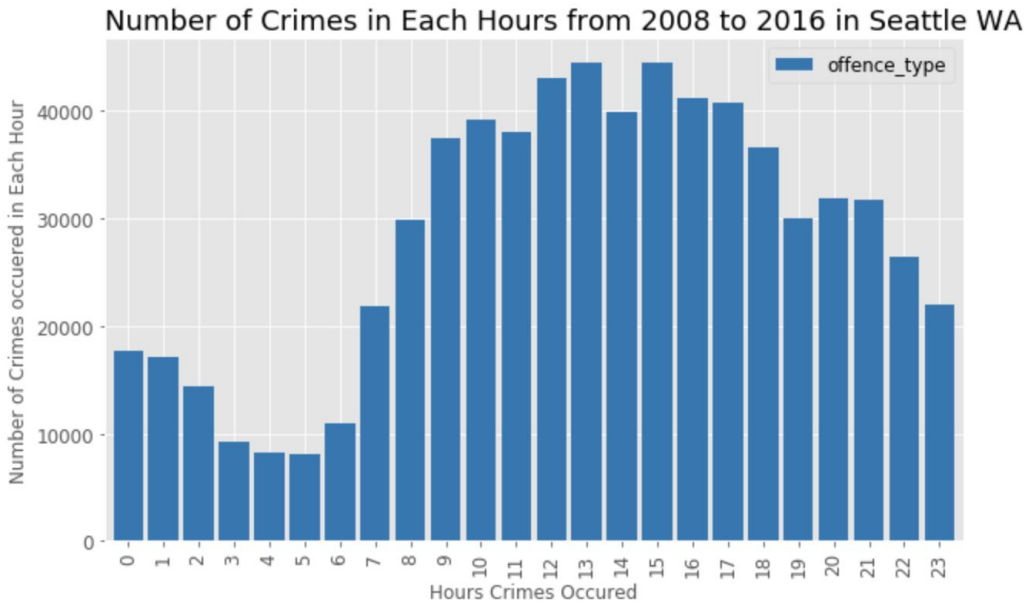
I have used barcharts and area plots to show the distribution of crimes and collisions in different time periods like year, month, day and hour of the day. I also use follium to map the crime data across neighborhoods and k-means clustering to segment the neighborhoods based on different venues.

Analysis

I have used different analysis techniques to accomplish this project. In the first part of the project I have tried to clean and analyze the Seattle crime data using Pandas. After analysing and cleaning the crime data using panda, I tried to show the data in different visualization techniques, mainly bar charts, area plots and finally mapping the most common crimes across different neighborhoods using folium.

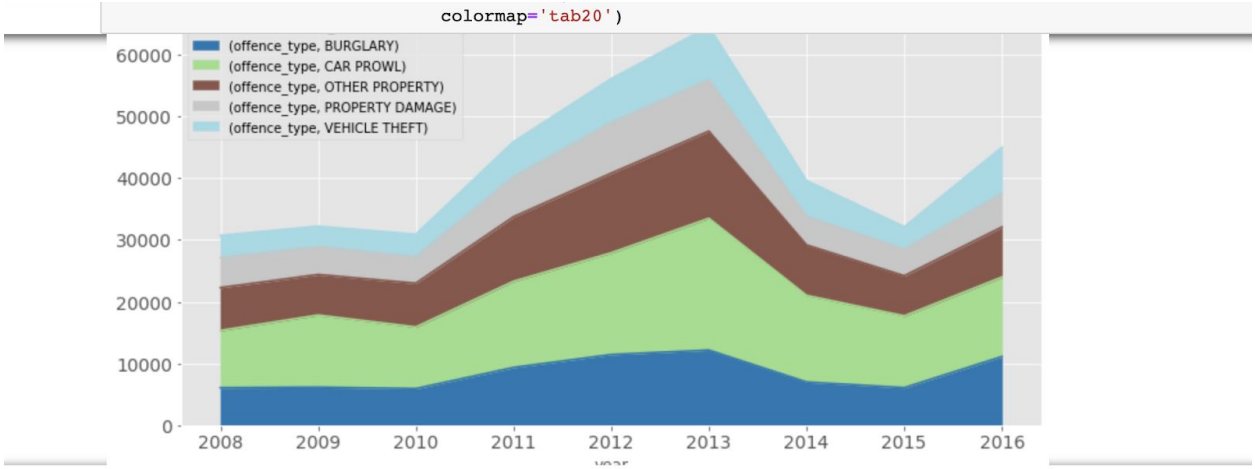


The graph shows the distribution of crimes across the Seattle area from 2008 to 2016. As shown in the graph above the highest crime rate occurred in Seattle is in 2013.

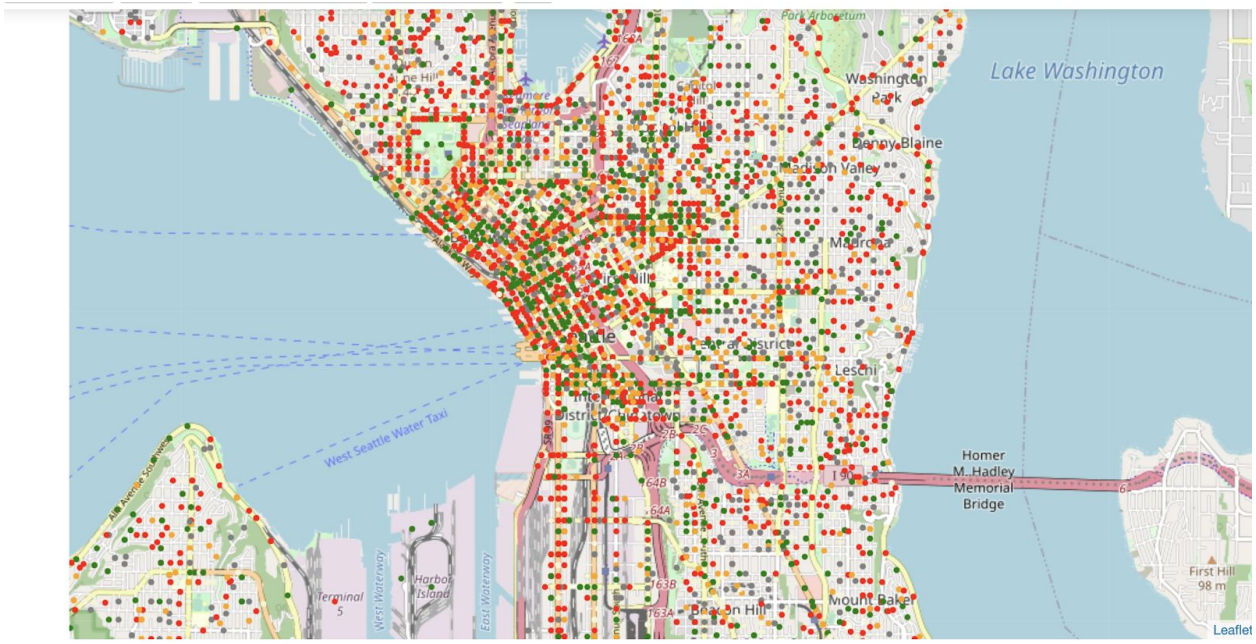


When we see the distribution of crimes in different times of the day, most of the crimes occur in the time range between 9:00 am to 5:00 pm.

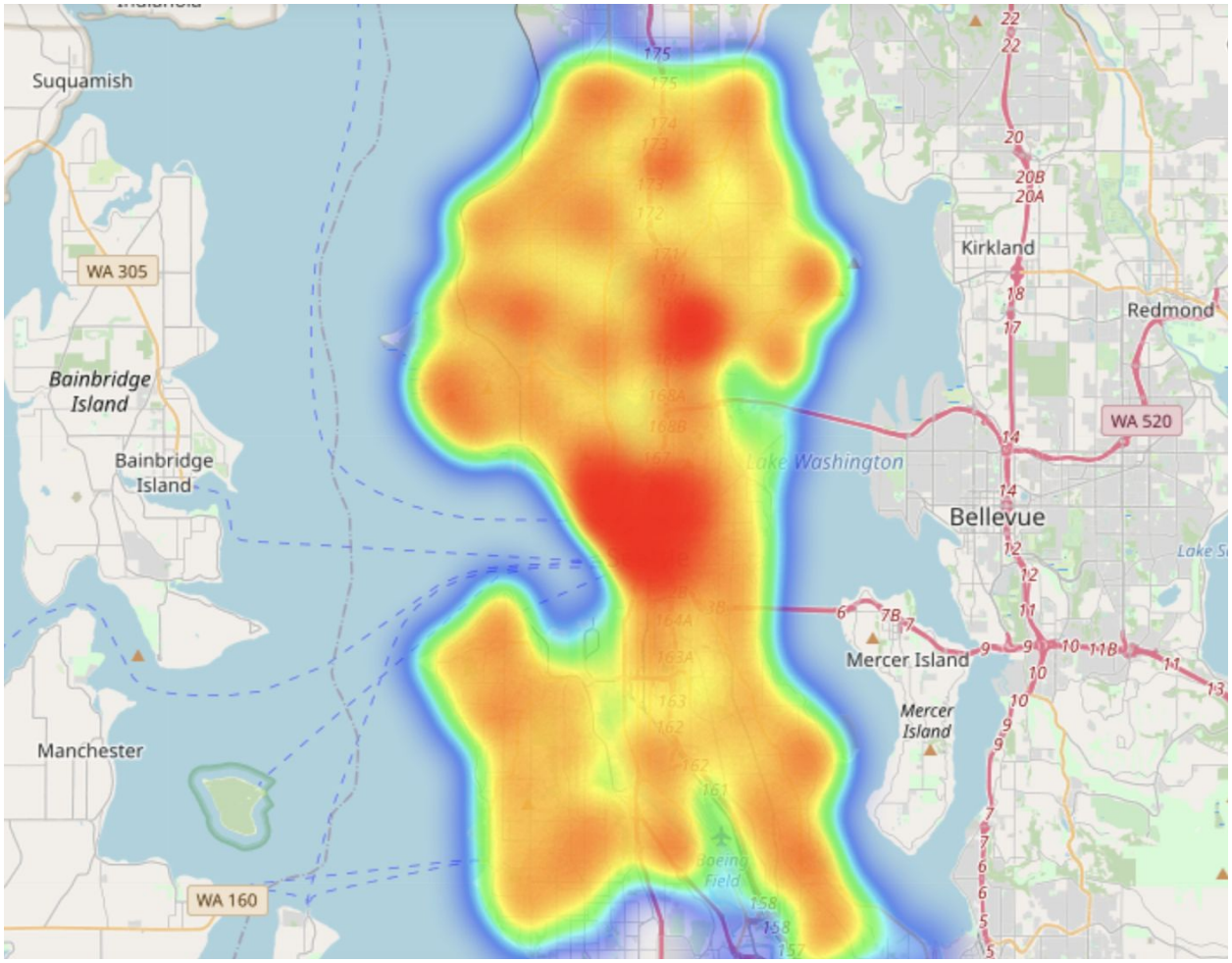
The five top most commonly occurred Crimes in Seattle WA are: CAR PROWL, OTHER PROPERTY, BURGLARY, PROPERTY DAMAGE and VEHICLE THEFT.



The area graph shows the top five crimes in Seattle WA from 2008 to 2016.

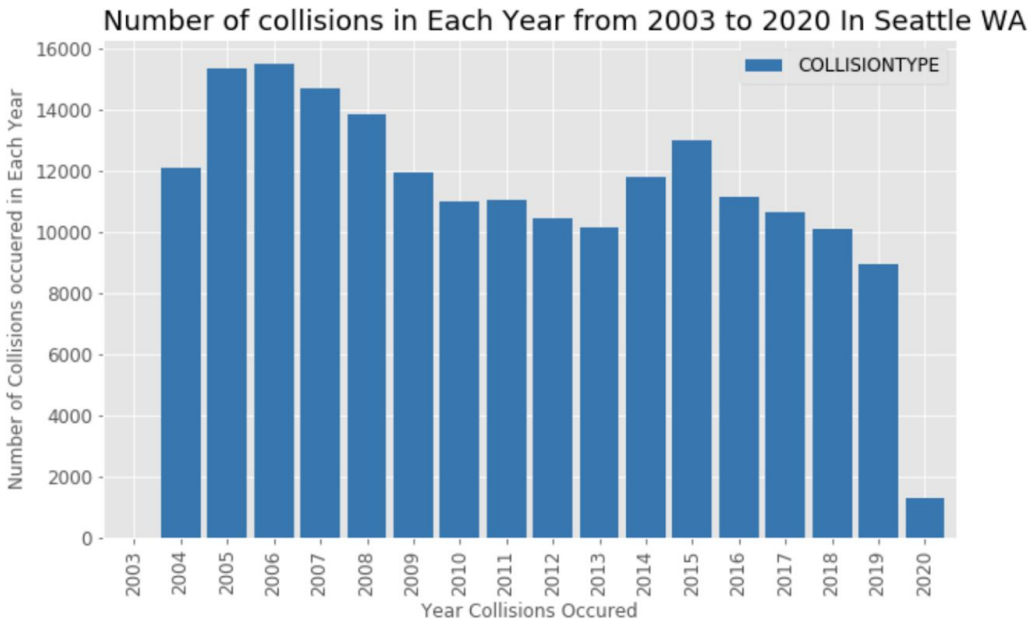


When we see the distribution of crimes across Seattle spatially, most of the crimes are concentrated in downtown Seattle area. As clearly shown in the map, as we go away from downtown, the rate of crime decreases, which means neighborhoods outside downtown Seattle are more safer than downtown Seattle.

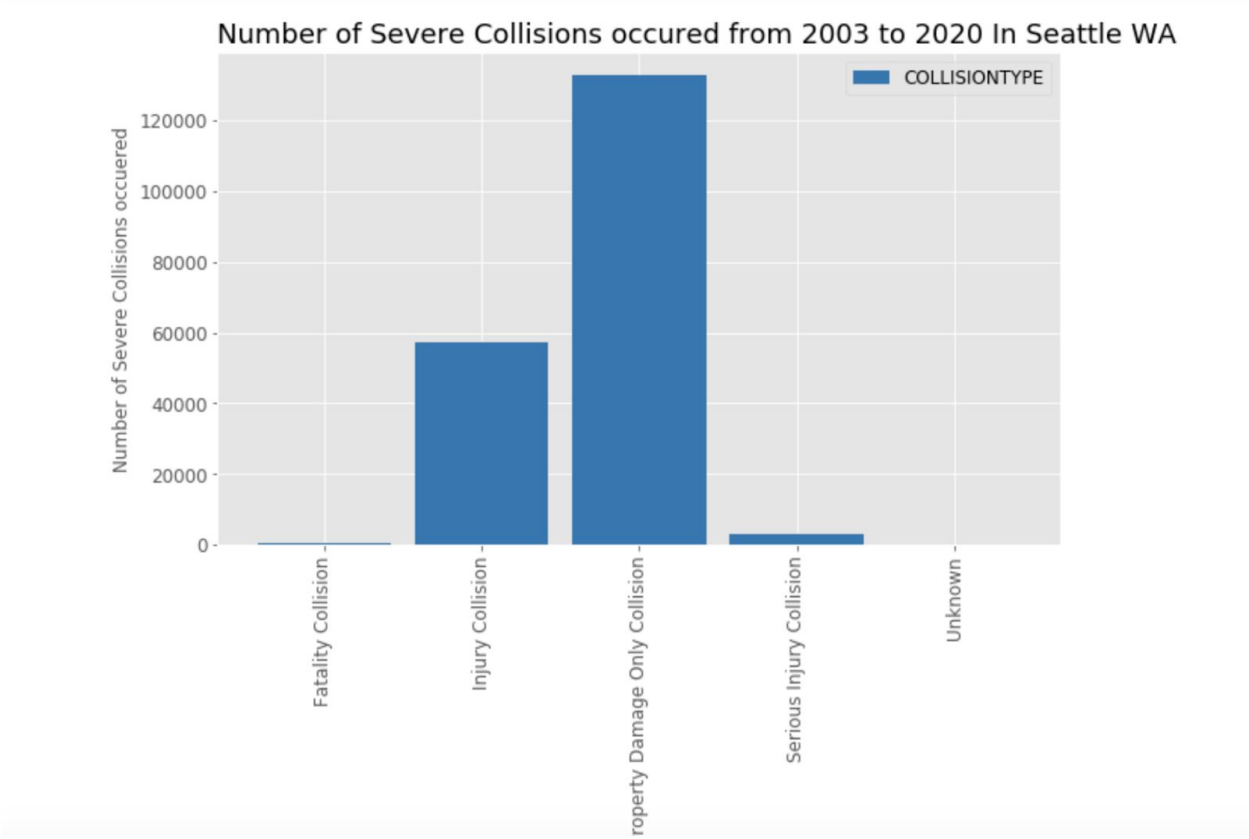


Even the heat map shows downtown Seattle is characterised by more crimes than the surrounding neighborhoods. As clearly shown in the heat map, the crime is more in the center of the city.

The next Part of my project shows the distribution of collisions in Seattle area form 2003 to 2020

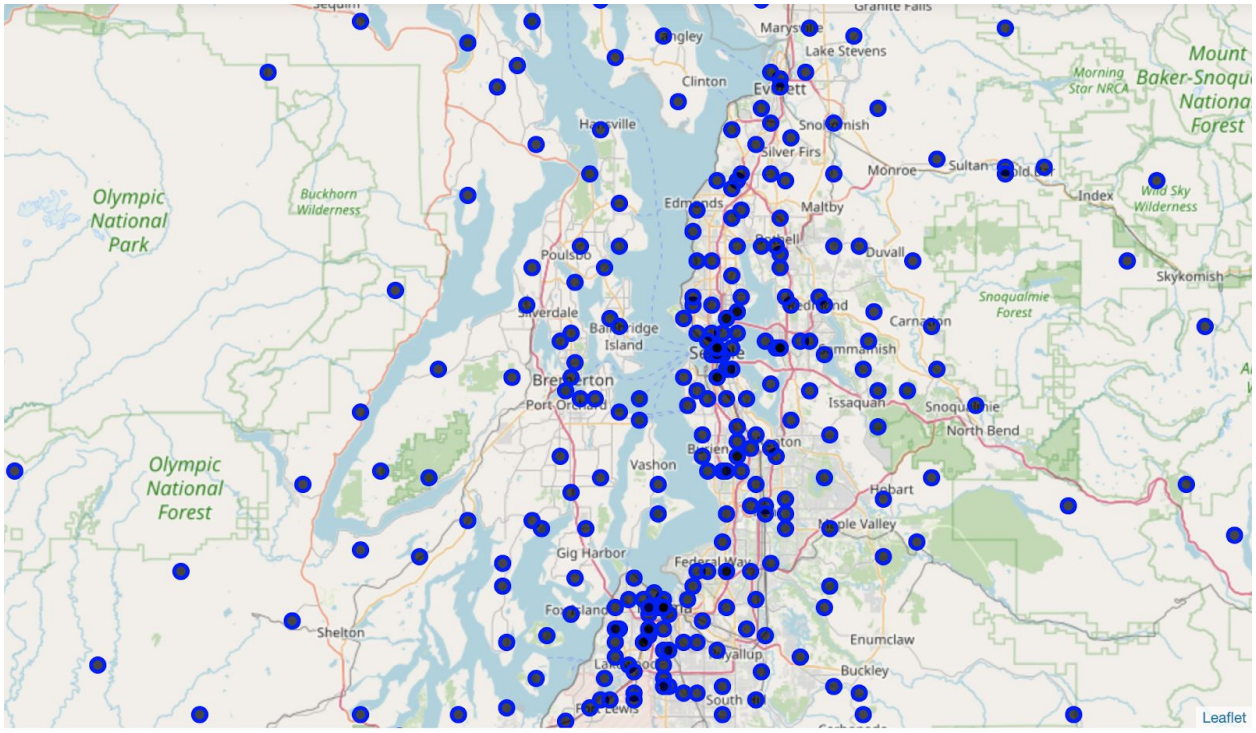


The graph shows the distribution of traffic collisions in Seattle WA from the year 2003 to 2020, as shown in the graph most of the collisions occurred in the year 2005 and 2006. From 2006 the colission gradually decreased until 2013 and then started to increase.



When we see the collisions by severity, the highest collision is property damage only collision.

The third Part of my project shows the mapping the seattle neighborhoods and the corresponding venues around seattle WA.



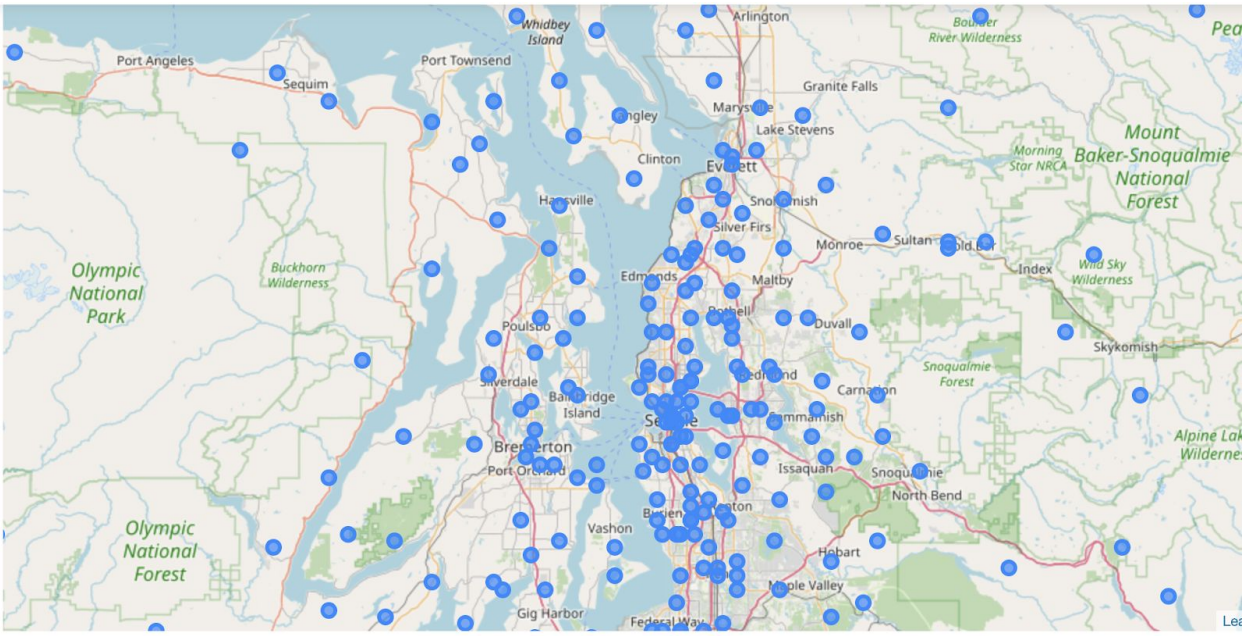
The map shows the distribution of Seattle neighborhoods using the foursquare api.

	venue.name	venue.categories	venue.location.lat	venue.location.lng
0	Biscuit B*tch	[{'id': '4bf58dd8d48988d143941735', 'name': 'B...	47.603237	-122.332010
1	Columbia Tower Club	[{'id': '52e81612bcbc57f1066b7a33', 'name': 'S...	47.604507	-122.330484
2	Smith Tower Observation Bar	[{'id': '4bf58dd8d48988d11e941735', 'name': 'C...	47.602012	-122.331793
3	Tat's Delicatessen	[{'id': '4bf58dd8d48988d1c5941735', 'name': 'S...	47.601901	-122.332423
4	Hole in the Wall Barbecue	[{'id': '4bf58dd8d48988d1df931735', 'name': 'B...	47.602525	-122.332009
5	Flatstick Pub	[{'id': '52e81612bcbc57f1066b79eb', 'name': 'M...	47.600143	-122.331002
6	Elm Coffee Roasters	[{'id': '4bf58dd8d48988d1e0931735', 'name': 'C...	47.600152	-122.330944
7	Smith Tower	[{'id': '4bf58dd8d48988d130941735', 'name': 'B...	47.601914	-122.331831
8	Manu's Bodega	[{'id': '4bf58dd8d48988d144941735', 'name': 'C...	47.601625	-122.329653
9	Cherry Street Public House	[{'id': '4bf58dd8d48988d1e0931735', 'name': 'C...	47.600450	-122.332908

These are the ten most common types of Venues around downtown Seattle.

neighborhood	latitude	longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
Auburn	47.30	-122.26	1.0	Coffee Shop	Pizza Place	Italian Restaurant	Chinese Restaurant	Donut Shop	Grocery Store	Park	Toy / Game Store	Thai Restaurant	Deli / Bodega
Auburn	47.31	-122.22	1.0	Coffee Shop	Pizza Place	Italian Restaurant	Chinese Restaurant	Donut Shop	Grocery Store	Park	Toy / Game Store	Thai Restaurant	Deli / Bodega
Federal Way	47.30	-122.31	1.0	Korean Restaurant	Chinese Restaurant	Coffee Shop	Pizza Place	Gym	Business Service	Comic Shop	Bank	Mexican Restaurant	Automotive Shop
Bellevue	47.61	-122.21	1.0	Coffee Shop	Pizza Place	Italian Restaurant	Vietnamese Restaurant	Seafood Restaurant	Sandwich Place	Asian Restaurant	Clothing Store	Park	Furniture / Home Store
Bellevue	47.62	-122.16	1.0	Coffee Shop	Pizza Place	Italian Restaurant	Vietnamese Restaurant	Seafood Restaurant	Sandwich Place	Asian Restaurant	Clothing Store	Park	Furniture / Home Store

The table shows the ten most common venues around each neighborhood in Seattle WA. The Venues are grouped using the K-means clustering classification method.



The above ten most common venues are shown in the map using the folium library.

Conclusion

The objective of the business problem was to help stakeholders identify the safest neighborhoods in Seattle WA and also to help neighborhoods to easily find the most common Venues near by their neighborhoods.It also helps them to identify roads which have the highest number of road collisions. As I have tried to show in the report in particular and in the notebook in general, most of the crimes and collisions occurred in the downtown area.