Capstone Project - the possible covid-19 outbreak districts in Seattle

1. Introduction

The covid-19 Virus has become a major problem in the world, it is spreading at a fast speed, and can be transmitted thought human to human contact, hence, it is important to find district with lots of places that require human to human contact, such as restaurants. The study is done at a place that I used to live: Seattle.

2. Data

The analysis to find the density of different venues in different district in Seattle:

The sources of data are the following to achieve their respective aims:
- I get the information about the district in seattle through the websitle https://seattle.findwell.com/seattle-neighborhoods/
- I look up the latitudes and longitudes of the districts through the OpenCage Geocoder API
- I get the location information thought the Foursquare API.

for the sake of simplicity, I copied the information on the website and then pasted it to my excel file, and saved it as the csv format.
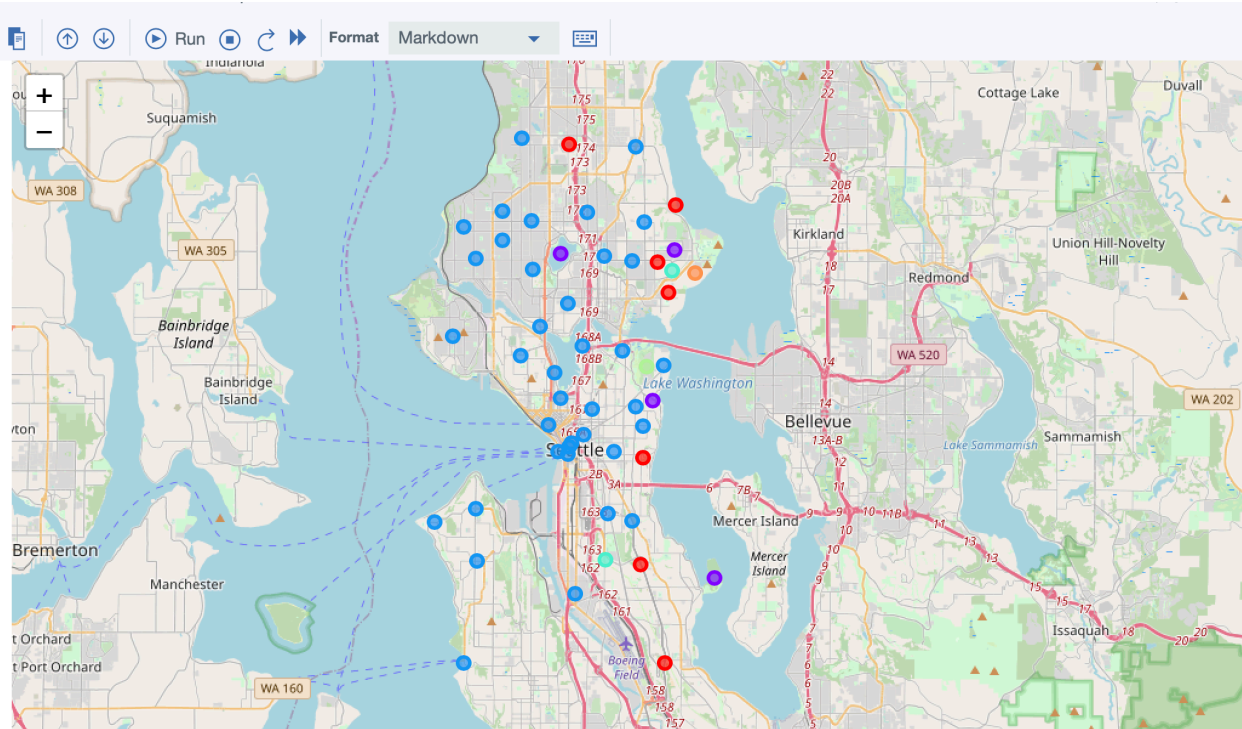
3. Methodology

By converting addresses into their equivalent latitude and longitude values. It is possible to use the Foursquare API to explore neighborhoods in Seattle. To explore function to get the most common venue categories in each neighborhood, and K-means clustering algorithm is used to group the neighborhoods into clusters. Finally, the Folium library is used to visualize the neighborhoods in Seattle and their emerging cluster.

The area with the greatest density of huaman to human contact venues is the most vulnerable district.

4. Analysis

Use geopy library to get the latitude and longitude values of SEATTLE

5. Results

The resulting clusters are as follow :

- Cluster 1 - low probility district
- Cluster 2 - low probility district
- Cluster 3 - High probility district
- Cluster 4 - moderate probility district
- Cluster 5 - low probility district
- Cluster 6 - low probility district

From the result, Cluster 3 has the most density of cafe shop and restrants **Therefore, Cluster 3 has the highest potential to have the covid-19 outbreak.**

Cluster 3 - high probility districts.

```
n [77]: district_merged.loc[district_merged['Cluster Labels'] == 2, district_merged.columns[[1] + [2] + list(range(5, district_merged.s
```

	neighborhood	Latitude	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
0	Central Seattle	47.613231	Hotel	Bakery	Coffee Shop	Breakfast Spot	Bar	Sushi Restaurant	Seafood Restaurant	Italian Restaurant	New American Restaurant	Pizza Place
2	Central Seattle	47.619004	Coffee Shop	Bar	American Restaurant	Mexican Restaurant	Japanese Restaurant	Thrift / Vintage Store	Pizza Place	Cocktail Bar	Café	Taco Place
3	Central Seattle	47.603110	Coffee Shop	Performing Arts Venue	Breakfast Spot	Plaza	Playground	Vietnamese Restaurant	Art Gallery	Pizza Place	Chinese Restaurant	BBQ Joint
5	Central Seattle	47.604949	Coffee Shop	Hotel	Cocktail Bar	Sandwich Place	Seafood Restaurant	Italian Restaurant	Breakfast Spot	Park	Pier	Pizza Place
6	Central Seattle	47.643145	Sandwich Place	Coffee Shop	Mexican Restaurant	Playground	Italian Restaurant	Steakhouse	Bar	Burger Joint	Deli / Bodega	Furniture / Home Store
7	Central Seattle	47.609305	Sandwich Place	Coffee Shop	Hotel	Pharmacy	Pizza Place	Bakery	Asian Restaurant	Restaurant	Poke Place	Indian Restaurant