Presentation – The Battle Of Neighborhoods

Student: Fodyé Cissé

Description of the problem: *Finding the best place to open a Pizza restaurant in the city of Toronto*

- Stakeholders wanted to open a new Pizza chain Restaurants in the city of Toronto
- They were looking for the best place to open their first new Pizza restaurant.

- They would like to avoid unnecessary competition against existing Pizza places.

- So, they wanted to open their first Pizza restaurant in locations where there as less pizza restaurants.

- As a data scientist, I have been asked by these stakeholders to resolve this business problem for them by finding the right places in the city of Toronto where they should start opening their Pizza restaurant.

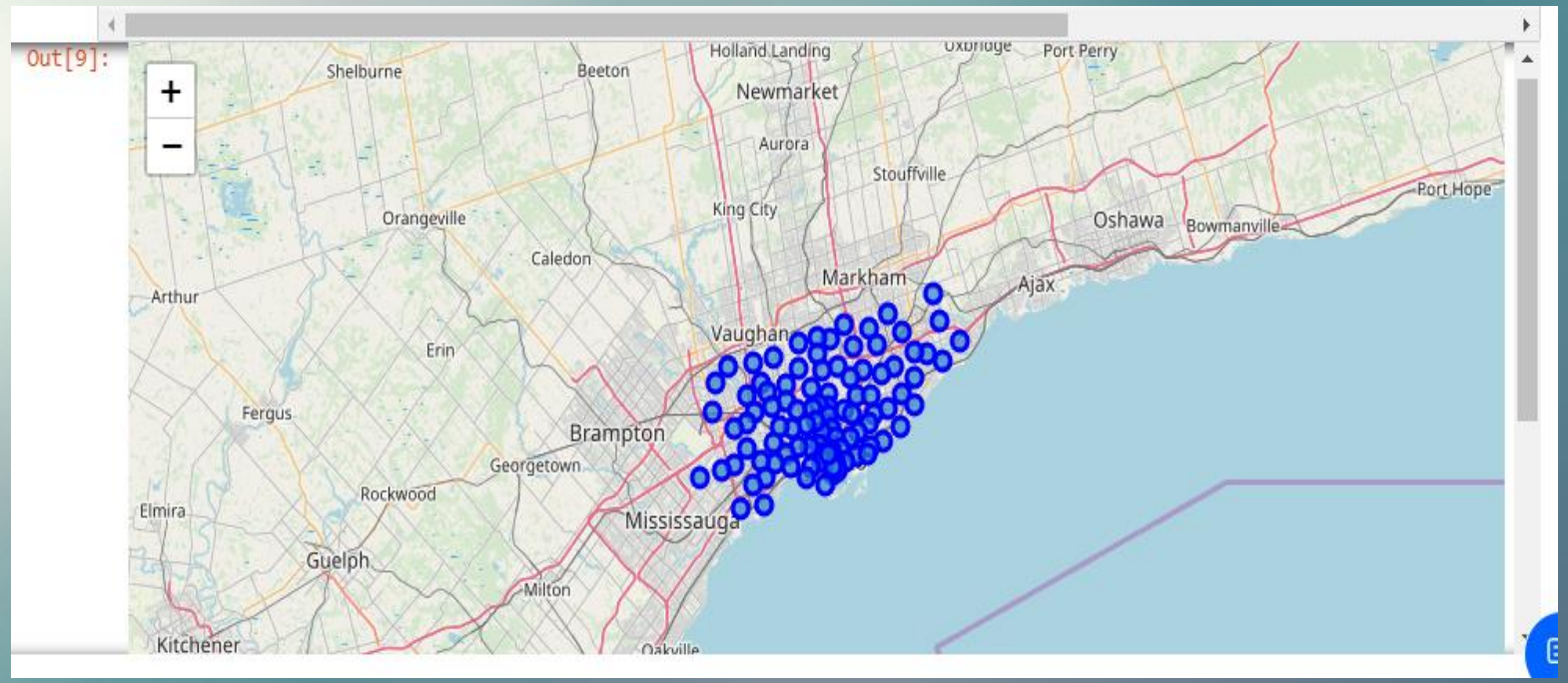
Description of the Data

- The Toronto city data from Wikipedia : https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

- The Toronto Geospace data: https://cocl.us/Geospatial_data. This is a csv file which contains all the geographical coordinates of all the Toronto Neighborhoods. That will be required to know all the latitudes and longitudes of all the Neighborhoods in Toronto.

- The Foursquare API: http://developper.foursquare.com . I will use the Foursquare's explorer API tool to collect all the venues, etc nearby each location.

Cleaning and Exploring the Data



- After Cleaning the Data, I can display in a Folium map all the Neighborhoods in Toronto

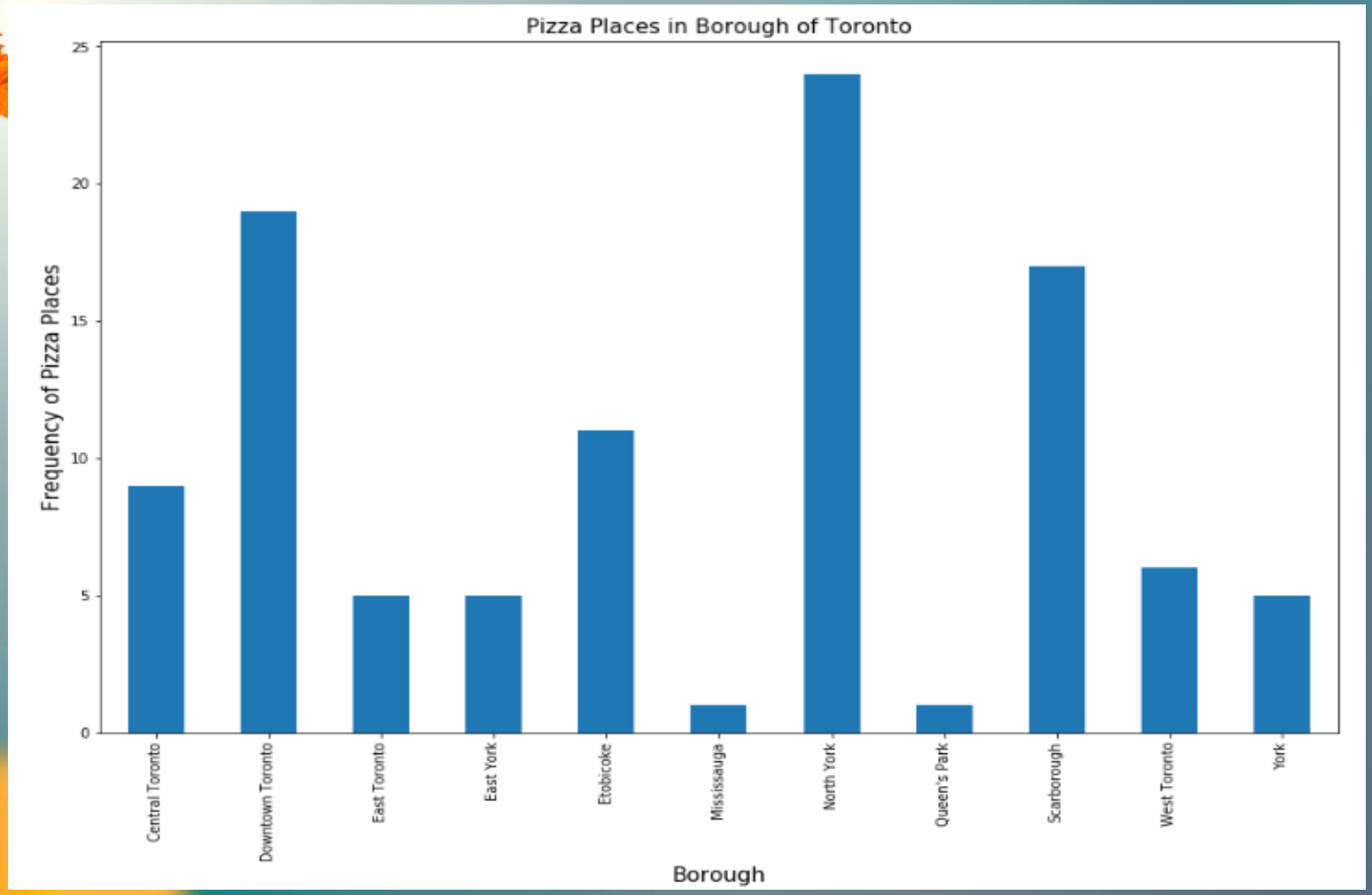
All the Data merged into one single DataFrame

```
# Now let's merge the Toronto Geodat and the Toronto Pizza Places data per Neighborhood
toronto_PizzaPlace_merged = pd.merge(df_Toronto_Geodata, toronto_PizzaPlace, on='Neighborhood')

# Now, let's change the columns order
toronto_PizzaPlace_merged = toronto_PizzaPlace_merged.reindex(columns=['Postalcode','Borough','Neighborhood','Pizza Pl

toronto_PizzaPlace_merged.head()
```

	Postalcode	Borough	Neighborhood	Pizza Place	Latitude	Longitude
0	M1B	Scarborough	Rouge, Malvern	0.058824	43.806686	-79.194353
1	M1C	Scarborough	Highland Creek, Rouge Hill, Port Union	0.000000	43.784535	-79.160497
2	M1E	Scarborough	Guildwood, Morningside, West Hill	0.121212	43.763573	-79.188711
3	M1G	Scarborough	Woburn	0.093750	43.770992	-79.216917
4	M1H	Scarborough	Cedarbrae	0.015152	43.773136	-79.239476

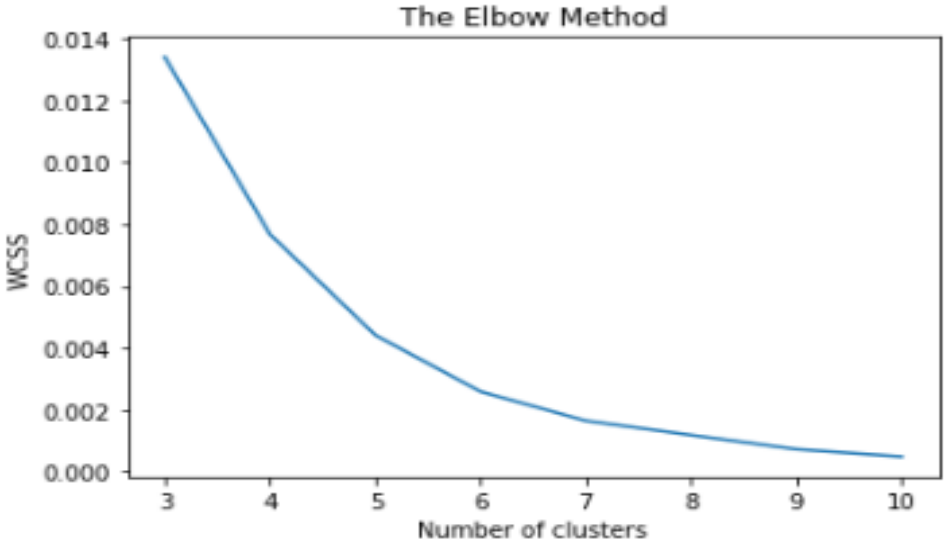
All the Pizza Place data merged with the Latitudes and Longitudes of all the Neighborhoods:

Frequency of Pizza Places in all the Boroughs in Toronto



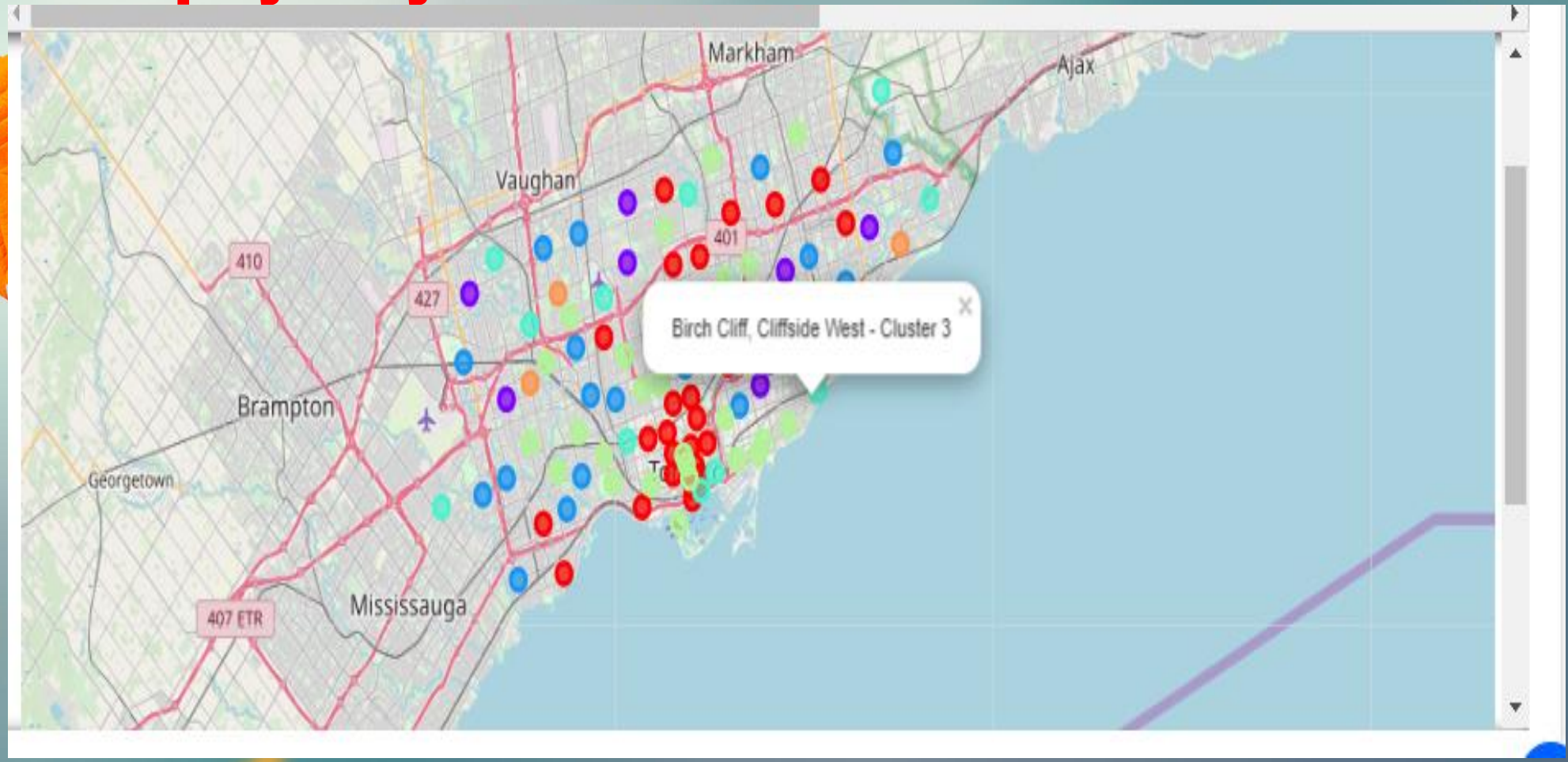
How many Pizza Places we have per Borough?

Clustering: Using the Elbow method to determine the best K value to use



From this map, we can see that 6 is the best K value.

Clustering: Folium map with the Neighborhoods displayed by clusters



Clustering: Cluster 3 has all the Neighborhoods that have less Pizza Places

```
# Cluster 3
toronto_PizzaPlace_Cluster3 = toronto_PizzaPlace_Clustering_merged2.loc[toronto_PizzaPlace_Clustering_merged2['Cluster Labels'] == 3]
print(toronto_PizzaPlace_Cluster3.shape)
toronto_PizzaPlace_Cluster3
```

```
(11, 7)
```

	Postalcode	Latitude	Longitude	Borough	Neighborhood	Pizza Place	Cluster Labels
96	M9L	43.756303	-79.565963	North York	Humber Summit	0.0	3
87	M7R	43.636966	-79.615819	Mississauga	Canada Post Gateway Processing Centre	0.0	3
16	M1X	43.836125	-79.205636	Scarborough	Upper Rouge	0.0	3
19	M2K	43.786947	-79.385975	North York	Bayview Village	0.0	3
97	M9M	43.724766	-79.532242	North York	Emery, Humberlea	0.0	3
1	M1C	43.784535	-79.160497	Scarborough	Highland Creek, Rouge Hill, Port Union	0.0	3
76	M6H	43.669005	-79.442259	West Toronto	Dovercourt Village, Dufferin	0.0	3
9	M1N	43.692657	-79.264848	Scarborough	Birch Cliff, Cliffside West	0.0	3
30	M3K	43.737473	-79.464763	North York	CFB Toronto, Downsview East	0.0	3
53	M5A	43.654260	-79.360636	Downtown Toronto	Harbourfront	0.0	3
56	M5E	43.644771	-79.373306	Downtown Toronto	Berczy Park	0.0	3

Conclusion

So, the best places to open their first Pizza Restaurant are the Neighborhoods identified in Cluster 3.

In the Folium map, Cluster 3 Neighborhoods are neighborhoods displayed in the map in Cyan color.