The_Battle_of_Neighborhoo

Code

Python

1. Introduction

This is the capstone project of the **IBM Data Science Professional Certification.**

For this capstone project, which is named **The Battle of Neighborhoods**, I decided to carry out the idea suggested by this course for this final assignment, which is to compare the cities of New York and Toronto. Both cities are very diverse and are the financial capitals of their respective countries.

Now that we have been equipped with the skills and the tools to use location data to explore a geographical location, we will have the opportunity to explore or compare neighborhoods or cities.

The_Battle_of_Neighborhoo

Code

Python

4. New York City

4.1. Download and Explore Dataset of New York City

Neighborhood of New York City has a total of 5 boroughs and 306 neighborhoods.

Boroughs

Neighborhoods

The_Battle_of_Neighborhoo

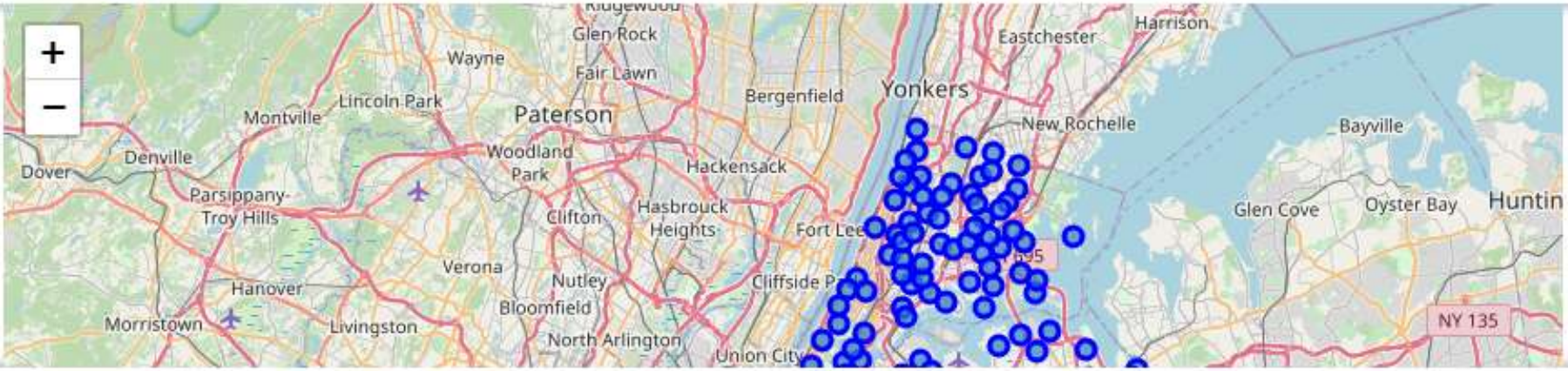
Code

Python

```
    label = folium.Popup(label, parse_html=True)
    folium.CircleMarker(
        [lat, lng],
        radius=5,
        popup=label,
        color='blue',
        fill=True,
        fill_color='#3186cc',
        fill_opacity=0.7,
        parse_html=False).add_to(map_newyork)

map_newyork
```

[106]:



The_Battle_of_Neighborhoo ●

Code

Python

neighborhood and its respective borough.

4.1.5. Clustering and segmentation the neighborhoods in Brooklyn.

For illustration purposes, let's simplify the above map and segment and cluster only the neighborhoods in Brooklyn. So let's slice the original dataframe and create a new dataframe of the Brooklyn data.

Therefore, int the code below, we will sort the Boroughs that contain the word 'Brooklyn'.

[107]:
```
brooklyn_data = neighborhoods[neighborhoods['Borough'] == 'Brooklyn'].reset_index(drop=True)
brooklyn_data.head()
```

The_Battle_of_Neighborhoo

Code

Python

[107]:
```
brooklyn_data = neighborhoods[neighborhoods['Borough'] == 'Brooklyn'].reset_index(drop=True)
brooklyn_data.head()
```

[107]:

	Borough	Neighborhood	Latitude	Longitude
0	Brooklyn	Bay Ridge	40.625801	-74.030621
1	Brooklyn	Bensonhurst	40.611009	-73.995180
2	Brooklyn	Sunset Park	40.645103	-74.010316
3	Brooklyn	Greenpoint	40.730201	-73.954241
4	Brooklyn	Gravesend	40.595260	-73.973471

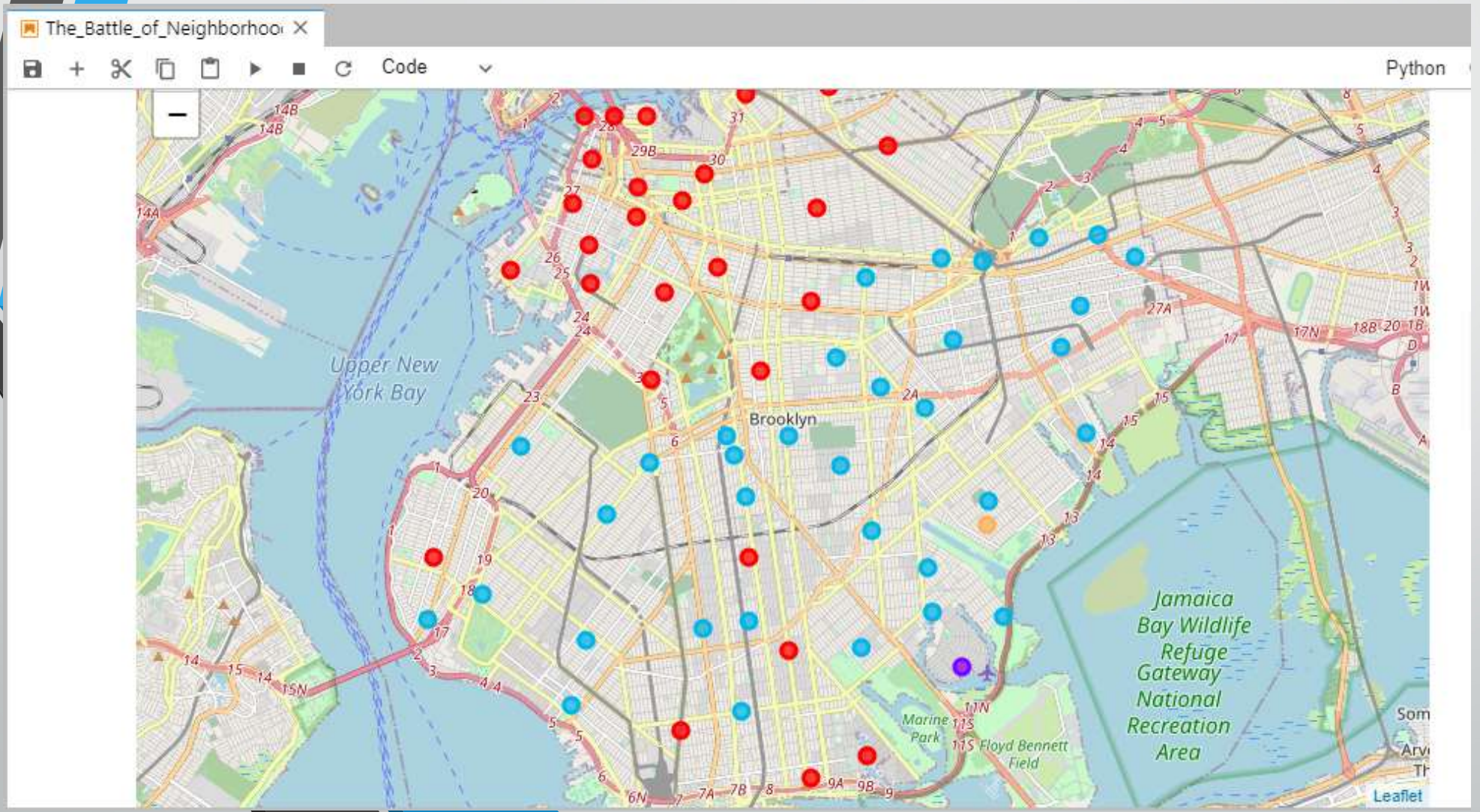
4.1.5.1. Geographical coordinates of Brooklyn.

Let's get the geographical coordinates of Brooklyn.

[108]:
```
address = 'Brooklyn, NY'

geolocator = Nominatim(user_agent="ny_explorer")
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude
print('The geograpical coordinate of Brooklyn are {}, {}.'.format(latitude, longitude))
```

```
The geograpical coordinate of Brooklyn are 40.6501038, -73.9495823.
```


The_Battle_of_Neighborhoo
Code
Python
Upper New York Bay
Brooklyn
Jamaica Bay Wildlife Refuge
Gateway National Recreation Area
Marine Park
Floyd Bennett Field
Leaflet

The_Battle_of_Neighborhoo

Code

Python

5. Toronto

In this section, we will explore, segment, and cluster the neighborhoods in the city of Toronto.

Neighborhood of Toronto has a total of 6 boroughs and 140 neighborhoods.

Boroughs

Neighborhoods

The_Battle_of_Neighborhoo

Code

Python

5.2.9. Sort the Dataframe by Postal Code.

[147]:
```
df1 = df1.sort_values(by ='Postal Code' )
df1 = df1.reset_index(drop = True)
df1.head()
```

[147]:

	Postal Code	Borough	Neighborhood
0	M1B	Scarborough	Malvern, Rouge
1	M1C	Scarborough	Rouge Hill, Port Union, Highland Creek
2	M1E	Scarborough	Guildwood, Morningside, West Hill
3	M1G	Scarborough	Woburn
4	M1H	Scarborough	Cedarbrae

Lets confirm the size of the table of the Postal Codes of Toronto.

[148]:
```
df1.shape
```

[148]:
```
(103, 3)
```

The_Battle_of_Neighborhoo

Code

Python

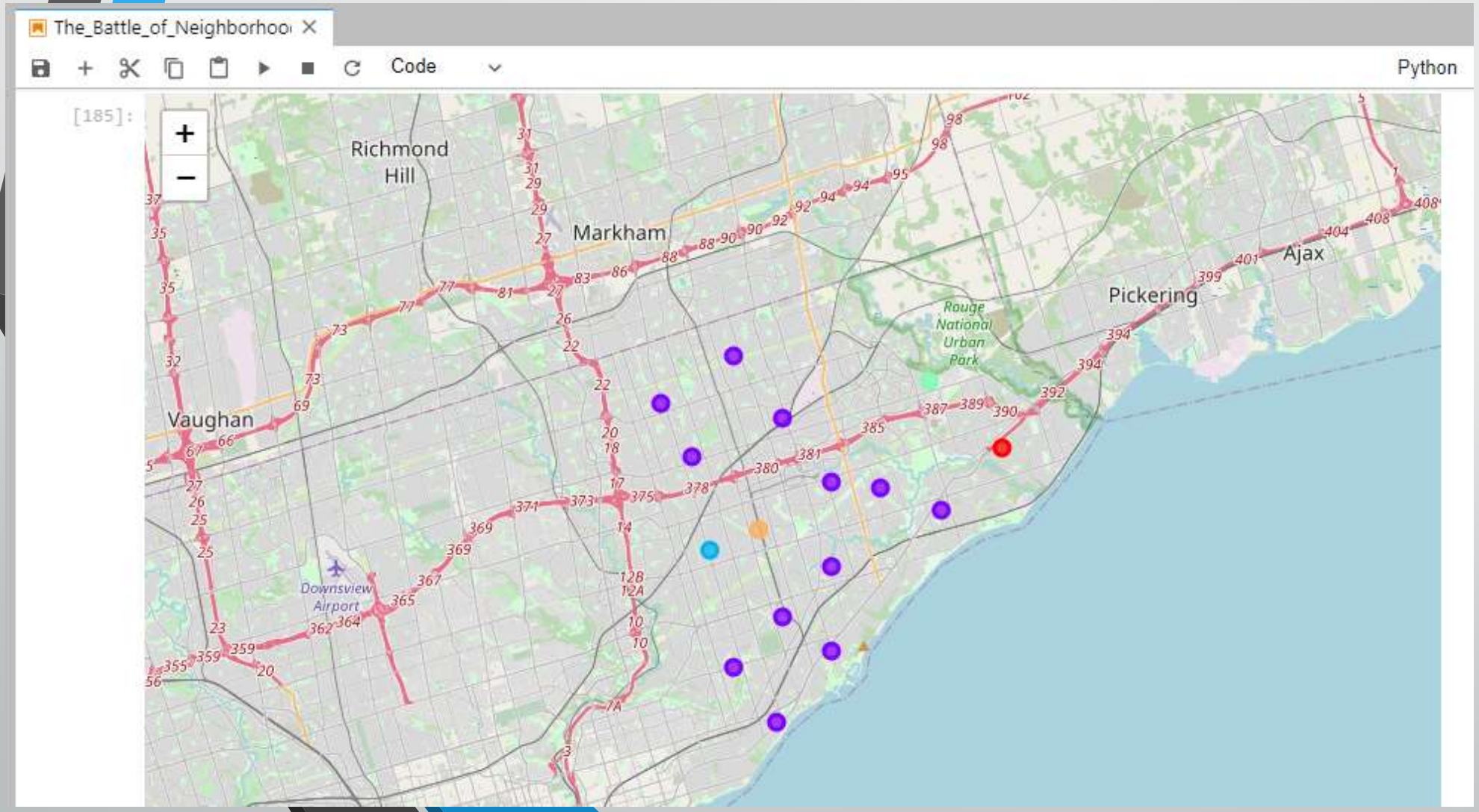
Leaflet

5.4.3. Clustering and segmentation the neighborhoods in Scarborough.

For illustration purposes, let's simplify the above map and segment and cluster only the neighborhoods in Scarborough. So let's slice the original dataframe and create a new dataframe of the Scarborough data.

Therefore, int the code below, we will sort the Boroughs that contain the word 'Scarborough'.

[156]:
```
scarborough_data = df1[df1['Borough'].str.contains('Scarborough')].reset_index(drop = True)
scarborough_data
```


The_Battle_of_Neighborhoo
Code
Python
[185]:
Richmond Hill
Markham
Vaughan
Pickering
Ajax
Rouge National Urban Park
Downsview Airport

The_Battle_of_Neighborhoo

Markdown

Python

6. Results

The first neighborhood of Brooklyn is Bay Ridge. The first neighborhood of Scarborough is Malvern, Rouge. Bay Ridge in Brooklyn has 63 venues. Malvern and Rouge in Scarborough have only 1 venue; which is Wendy's.

---

7. Discussion

Given the above results, I suggest that starting an Italian Food Restaurant in the first neighborhood of Scarborough would be very interesting since this king of food in missing in this place. The people or citizens in this neighborhood will accept this restaurant because they will enjoy different choices of food.

---

8. Conclusion

Foursquare location is a very strong tool for Data Scientists. Among other things, we can explore cities in order to find insights for new or actual business.

I hope you enjoy my work! Thank you!