

# Capstone Project:

Battle of the Neighborhoods  
--- Which place should I choose?

by Yiyao Ma



# Table of Contents

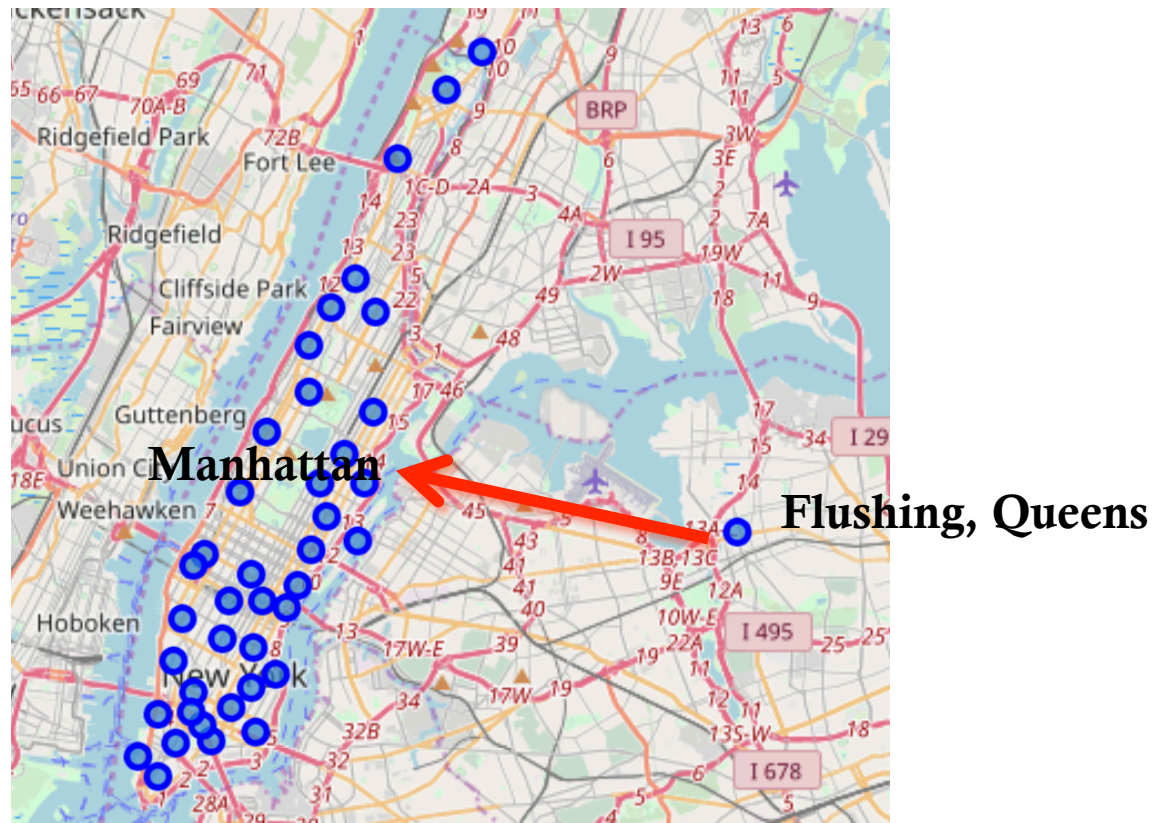
- ◆ A. Introduction/Business Problem
- ◆ B. Data
- ◆ C. Method
- ◆ D. Results
- ◆ E. Discussion
- ◆ F. Conclusion

# A. Introduction/Business Problem

- ◆ Jack Lee is currently living in Flushing, Queens, New York City. He needs to rent a studio in a neighborhood in Manhattan.
- ◆ He has two requirements:
  - ◆ (1) similar neighborhoods/venues and facilities as Flushing
  - ◆ (2) lowest average rent price for studio among the similar neighborhoods.

# A. Introduction/Business Problem

So, the problem is --- Which neighborhood in Manhattan is similar to Flushing, Queens and has the lowest average rent for a studio?



# B. Data

## 💧 1. Geographic Data

['Borough', 'Neighborhood', 'Latitude', 'Longitude']

## 💧 2. Foursquare Location Data

['Neighborhood', 'Neighborhood Latitude', 'Neighborhood Longitude',  
'Venue', 'Venue Latitude', 'Venue Longitude', 'Venue Category']

## 💧 3. Rent Data

['Neighborhood', 'Studio Rent']

# C. Methodology

- 💧 We will leverage the **Foursquare location data** to explore or compare Flushing, Queens, NYC and the neighborhoods of Manhattan, use the **k-means clustering algorithm** to find the similar neighborhood(s), use the **Folium library** to visualize the neighborhoods and their emerging clusters, and finally apply **Beautifulsoup library** to extract the average rent data in Manhattan by neighborhood from online sources to find the similar neighborhoods in Manhattan that has the lowest average rent for a studio.



# C. Methodology

💧 What Machine Learning were used and why?

We use **k-means clustering method**. K-means is vastly used for clustering in many data science applications, especially useful if you need to quickly discover insights from unlabeled data.

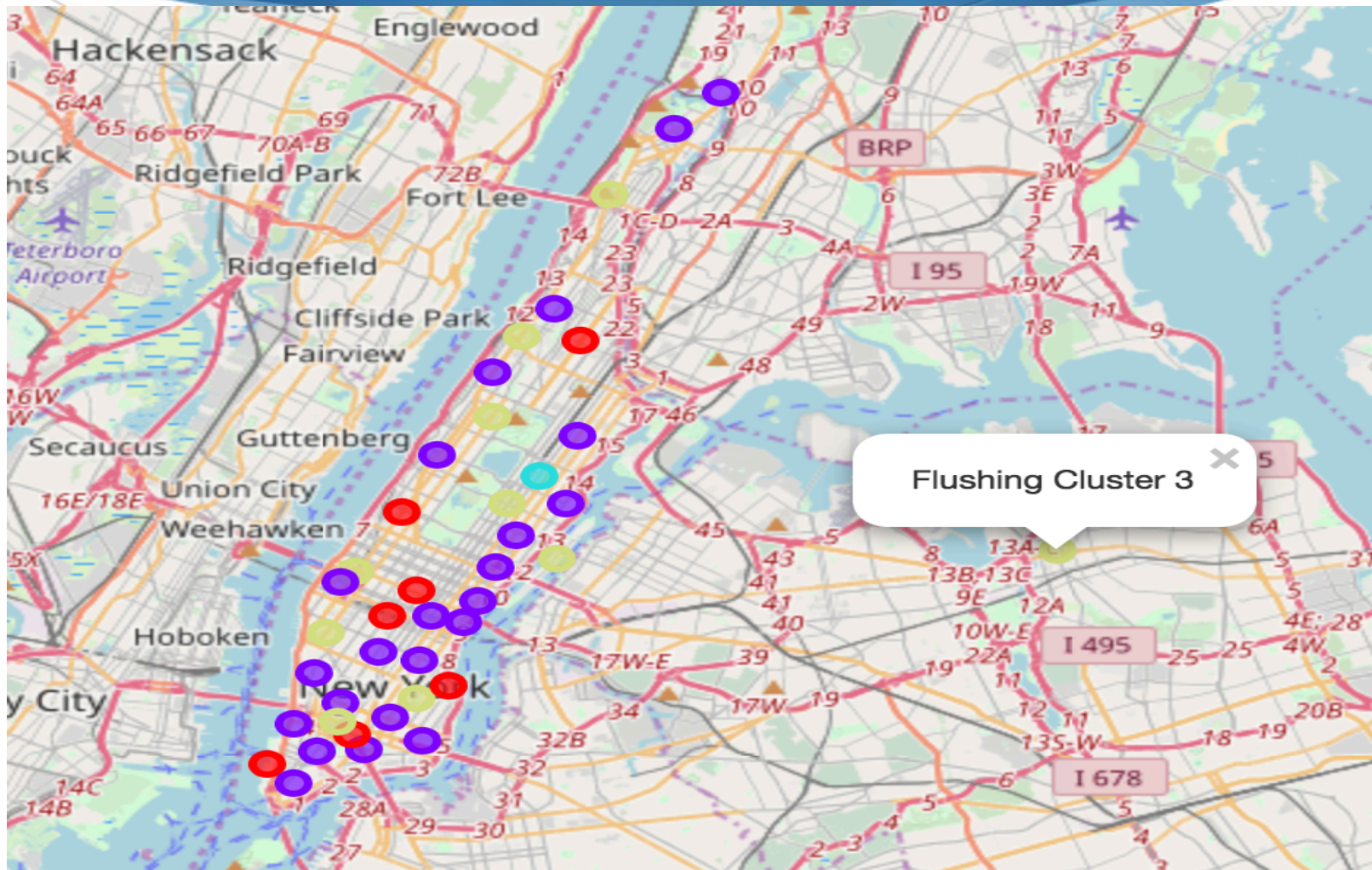
# C. Methodology

- 1. Download and Explore Dataset
- 2. Explore Neighborhoods in New York City
- 3. Analyze Each Neighborhood
- 4. Cluster Neighborhoods --- **K-means clustering**
- 5. Examine Clusters and Find Similar Neighborhoods in Manhattan
- 6. Rank the Similar Neighborhoods by Studio Rent Price



# D. Results

- 1. Here is the clustering map of Flushing and Manhattan after running k-means clustering.



# D. Results

- 2. Examining each cluster, we can see that Flushing, Queens is in Cluster 4 (['Cluster Labels'] == 3)

## Cluster 4

```
[71]: fl_manhattan_merged.loc[fl_manhattan_merged['Cluster Labels'] == 3, fl_manhattan_merged.columns[[1] + list(range(5, fl_manhattan_merged.shape[1])
```

```
[71]:
```

	Neighborhood	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	Flushing	Karaoke Bar	Bubble Tea Shop	Chinese Restaurant	Korean Restaurant	Bakery	Hotpot Restaurant	Szechuan Restaurant	Seafood Restaurant	Asian Restaurant	Construction & Landscaping
3	Washington Heights	Café	Sandwich Place	Mobile Phone Shop	Bakery	Mexican Restaurant	Chinese Restaurant	Caribbean Restaurant	Park	Clothing Store	Supermarket
6	Manhattanville	Deli / Bodega	Mexican Restaurant	Coffee Shop	Italian Restaurant	Seafood Restaurant	Music School	Bike Trail	Sushi Restaurant	Supermarket	Burger Joint
9	Upper East Side	Italian Restaurant	Exhibit	Coffee Shop	Juice Bar	Bakery	Art Gallery	Spa	Gym / Fitness Center	French Restaurant	Boutique
12	Roosevelt Island	Park	Sandwich Place	Liquor Store	Café	Market	Greek Restaurant	School	Gym	Coffee Shop	Metro Station
15	Clinton	Theater	Italian Restaurant	Hotel	American Restaurant	Wine Shop	Coffee Shop	Indie Theater	Gym / Fitness Center	Gym	Spa
18	Chelsea	Italian Restaurant	Coffee Shop	Nightclub	Ice Cream Shop	Bakery	Art Gallery	Seafood Restaurant	Theater	American Restaurant	Hotel
20	East Village	Ice Cream Shop	Bar	Wine Bar	Mexican Restaurant	Chinese Restaurant	Cocktail Bar	Pizza Place	Coffee Shop	Vegetarian / Vegan Restaurant	Italian Restaurant
24	Soho	Clothing Store	Women's Store	Shoe Store	Boutique	Men's Store	Italian Restaurant	Coffee Shop	Mediterranean Restaurant	Spa	Art Gallery
26	Manhattan Valley	Pizza Place	Coffee Shop	Bar	Yoga Studio	Mexican Restaurant	Indian Restaurant	Italian Restaurant	Playground	American Restaurant	Café

# D. Results

- 3. Let's create a dataframe that includes all the neighborhoods in Manhattan in Cluster 4.

The similar neighborhoods in Manhattan are:

	Neighborhood
0	Washington Heights
1	Manhattanville
2	Upper East Side
3	Roosevelt Island
4	Clinton
5	Chelsea
6	East Village
7	Soho
8	Manhattan Valley

# D. Results

- 4. Let's check the average Studio Rent of similar neighborhoods ranking in ascending order.

	Neighborhood	Studio Rent
0	Washington Heights	1678.0
1	Manhattan Valley	2506.0
2	East Village	2933.0
3	Chelsea	3171.0
4	Roosevelt Island	3192.0

We can see that **Washington Heights** has the lowest Studio Rent in the Cluster 4.

Note: Because the rent dataset doesn't include the rent data of all the neighborhoods, some neighborhoods such as Manhattanville, Upper East Side, Clinton do not have rent data for analysis. To simplify the process, we simply delete those neighborhoods which means that Jack Lee won't choose from those neighborhoods.



# E. Discussion

As we can see from the results above:

- From the Cluster 4 table as shown below, we can see that both Washington Heights and Flushing have many various kinds of restaurants such as Chinese restaurant, Asian restaurant, Bakeries and etc. And probably the main reason that Jack Lee enjoys living in Flushing is due to such many diversified restaurants especially Chinese restaurants as he was originally from China, which makes a lot of sense.

	Neighborhood	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	Flushing	Karaoke Bar	Bubble Tea Shop	Chinese Restaurant	Korean Restaurant	Bakery	Hotpot Restaurant	Szechuan Restaurant	Seafood Restaurant	Asian Restaurant	Construction & Landscaping
3	Washington Heights	Café	Sandwich Place	Mobile Phone Shop	Bakery	Mexican Restaurant	Chinese Restaurant	Caribbean Restaurant	Park	Clothing Store	Supermarket

- In the meantime, as Jack just graduated and has to pay back student loans, he needs to be frugal about his living expenses. Therefore he needs to rent a studio in a neighborhood that has the lowest average rental price.

# F. Conclusion

## 💧 Conclusion:

Based on the analysis and results shown above, we conclude that the similar neighborhood in Manhattan as Flushing with Lowest Studio Rental price is Washington Heights with studio rent of 1678 USD.

Therefore, **we recommend Jack Lee to rent a studio in Washington Heights** because it meets his two requirements:

- (1) Washington Heights has the **similar neighborhoods/venues and facilities** as Flushing, and
- (2) Washington Heights has the **lowest average rent price** for studio among the similar neighborhoods.



# F. Conclusion

## ◆ Further Possible Developments:

Due to some limitations, there are several areas that could be developed or improved in the future research:

- (1) Use **DBSCAN** (Density-based spatial clustering of applications with noise) as the machine learning method to cluster neighborhoods for further studies.
- (2) Find more completed rent dataset that includes the average rental costs of all neighborhoods.

Capstone Project:  
Battle of the Neighborhoods  
--- Which place should I choose?

Thanks!