Battle of Neighborhoods

By:
Jigisha Desai

Introduction

People are turning to the new cities/countries for a job. For this reason, before actually moving to the new city, if people can determine neighborhoods in the new city that are exactly the same as their current neighborhood, and if not perhaps similar neighborhoods that are at least closer to their new job through their access to the platforms where the detailed information like transportation, rental places, venues are mentioned together then their life will be easier in the new city.

Data Acquisition

- In this project city of London(source) and Manhattan (destination) are considered.
- In order to make a good choice of a similar apartment in Manhattan NY, we need data about subway stations, rental places and venues in Manhattan.
- Subway Stations Data:

To retrieve subway stations details data scrapped from

https://en.wikipedia.org/wiki/Category:New_York_City_Subway_stations_in_Manhattan

- Rental places Data:

To retrieve rental places details data scrapped from
https://www.realtor.com/
http://www.rentmanhattan.com/
https://www.calibernyc.com/

- Venues Data:

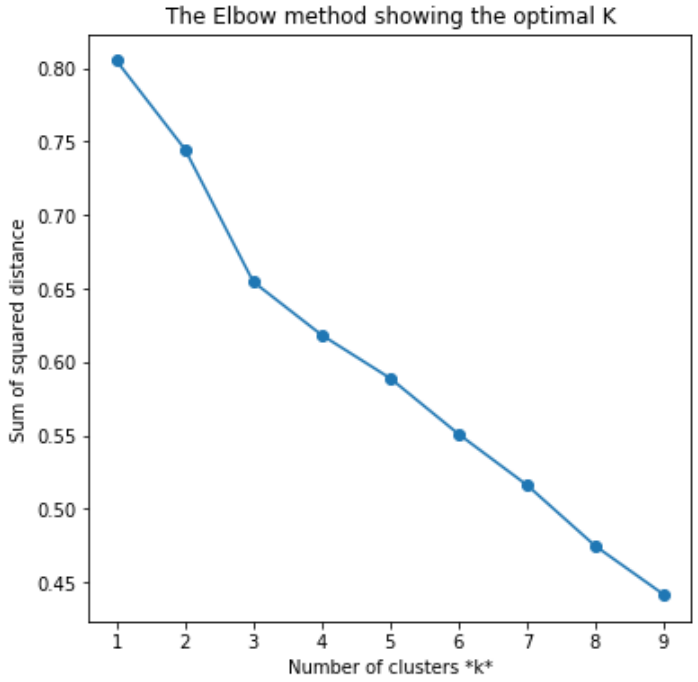
To retrieve different venues in Manhattan, first we need to get the dataset of neighborhoods in Manhattan. This can be retrieved from **"newyork_data.json"** file. Using Foursquare credentials and APIs details about the venues in the neighborhoods of Manhattan is retrieved.

Exploratory data analysis & Selecting Machine Learning Algorithm:

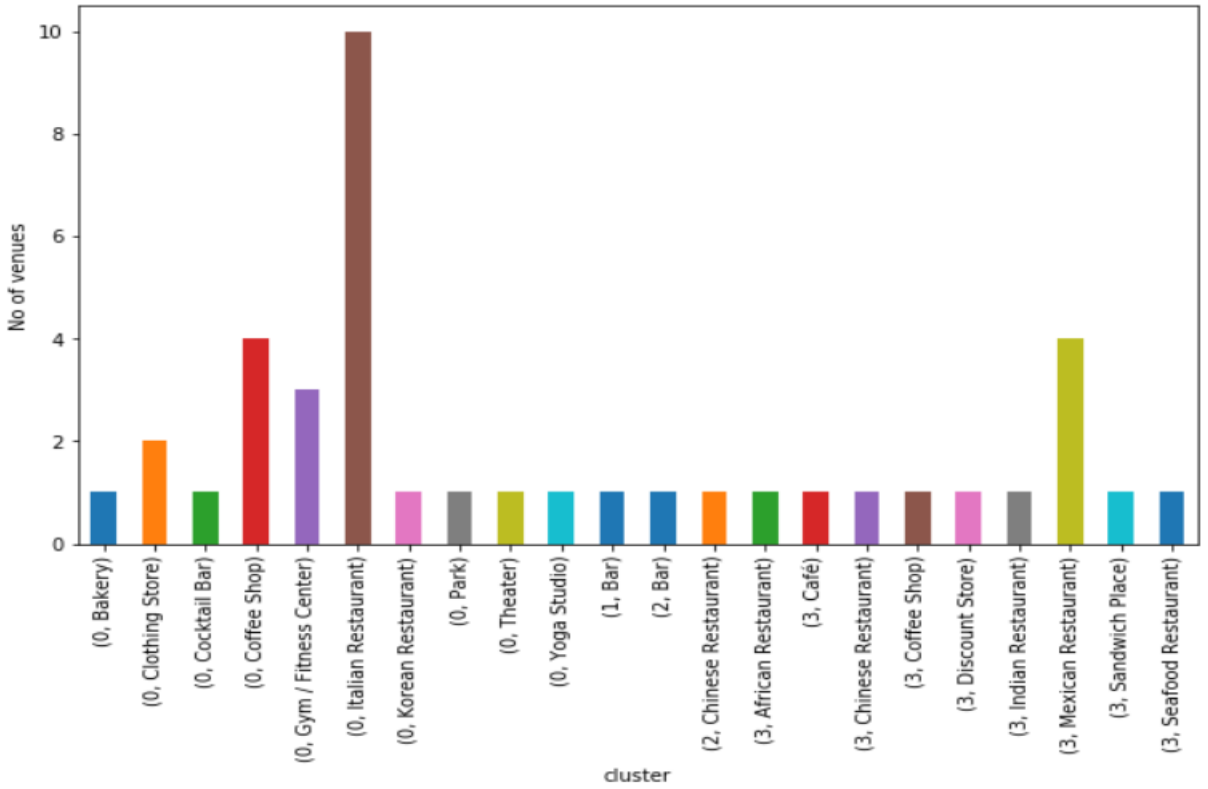
- Using Foursquare API and credentials, venues in Manhattan are explored. I designed the limit as **100 venues** and the radius **500 meter**. Then a table top 10 venue category for each borough is created.

	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	Battery Park City	Park	Coffee Shop	Hotel	Gym	Wine Shop	Shopping Mall	Italian Restaurant	Food Court	Burger Joint	Fountain
1	Carnegie Hill	Coffee Shop	Pizza Place	Bar	Café	Yoga Studio	Spa	Bookstore	Cosmetics Shop	French Restaurant	Gym
2	Central Harlem	African Restaurant	American Restaurant	Public Art	Gym / Fitness Center	French Restaurant	Chinese Restaurant	Seafood Restaurant	Cycle Studio	Bookstore	Ethiopian Restaurant
3	Chelsea	Coffee Shop	Italian Restaurant	Ice Cream Shop	Bakery	American Restaurant	Nightclub	Hotel	Theater	Seafood Restaurant	Asian Restaurant
4	Chinatown	Chinese Restaurant	American Restaurant	Vietnamese Restaurant	Cocktail Bar	Dim Sum Restaurant	Noodle House	Salon / Barbershop	Bakery	Hotpot Restaurant	Ice Cream Shop

- We have some common venue categories in boroughs. In this reason unsupervised learning K-means algorithm to cluster the boroughs is selected.
- Using Elbow method, value of k is selected as 4.



- By estimating the number of **1st Most Common Venue** in each cluster and there by creating a bar chart may help us to find proper labels for each cluster.



- By examining the bar chart we can label each cluster as follows:

Cluster 0: Coffee shop venues & Intensive Italian restaurants

Cluster 1: Bar venues

Cluster 2: Chinese restaurants

Cluster 3: Multi-cuisine restaurants

Result

- The result of cluster analysis is as follows:
- **Cluster 0:**

```
# Cluster 0
#Cluster Label:Coffee shop venues  & Intensive Italian restaurants
manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 0, manhattan_merged.columns[[1] + list(range(5, manhattan_merged.shape
```

	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
8	Upper East Side	Italian Restaurant	Coffee Shop	Exhibit	Art Gallery	Bakery	Juice Bar	Gym / Fitness Center	French Restaurant	Boutique	Hotel
9	Yorkville	Italian Restaurant	Coffee Shop	Gym	Bar	Pizza Place	Diner	Deli / Bodega	Mexican Restaurant	Thai Restaurant	Sushi Restaurant
10	Lenox Hill	Italian Restaurant	Sushi Restaurant	Coffee Shop	Gym / Fitness Center	Gym	Deli / Bodega	Burger Joint	Pizza Place	Sporting Goods Shop	Café
12	Upper West Side	Italian Restaurant	Bar	Vegetarian / Vegan Restaurant	Indian Restaurant	Wine Bar	Mediterranean Restaurant	Bakery	Coffee Shop	Burger Joint	Ice Cream Shop
13	Lincoln Square	Gym / Fitness Center	Theater	Café	Plaza	Concert Hall	Italian Restaurant	French Restaurant	Performing Arts Venue	Park	Opera House
14	Clinton	Theater	American Restaurant	Hotel	Gym / Fitness Center	Indie Theater	Wine Shop	Spa	Coffee Shop	Gym	Italian Restaurant
15	Midtown	Clothing Store	Hotel	Cocktail Bar	Steakhouse	Theater	Bakery	Coffee Shop	Spa	American Restaurant	Bookstore
16	Murray Hill	Coffee Shop	Hotel	Spa	French	American	Italian	Bar	Sandwich	Gym	Japanese

- **Cluster 1:**

```
# Cluster 1
#Cluster Label: Bar venues
manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 1, manhattan_merged.columns[[1] + list(range(5, manhattan_merged.shape
```

	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
37	Stuyvesant Town	Bar	Park	Pet Service	Baseball Field	Harbor / Marina	Farmers Market	Beer Garden	Cocktail Bar	Coffee Shop	Heliport

- **Cluster 2:**

```
# Cluster 2
#Cluster Label: Chinese Restaurants
manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 2, manhattan_merged.columns[[1] + list(range(5, manhattan_merged.shape
```

	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
1	Chinatown	Chinese Restaurant	American Restaurant	Vietnamese Restaurant	Cocktail Bar	Dim Sum Restaurant	Noodle House	Salon / Barbershop	Bakery	Hotpot Restaurant	Ice Cream Shop
19	East Village	Bar	Wine Bar	Ice Cream Shop	Mexican Restaurant	Chinese Restaurant	Pizza Place	Ramen Restaurant	Cocktail Bar	Japanese Restaurant	Vegetarian / Vegan Restaurant

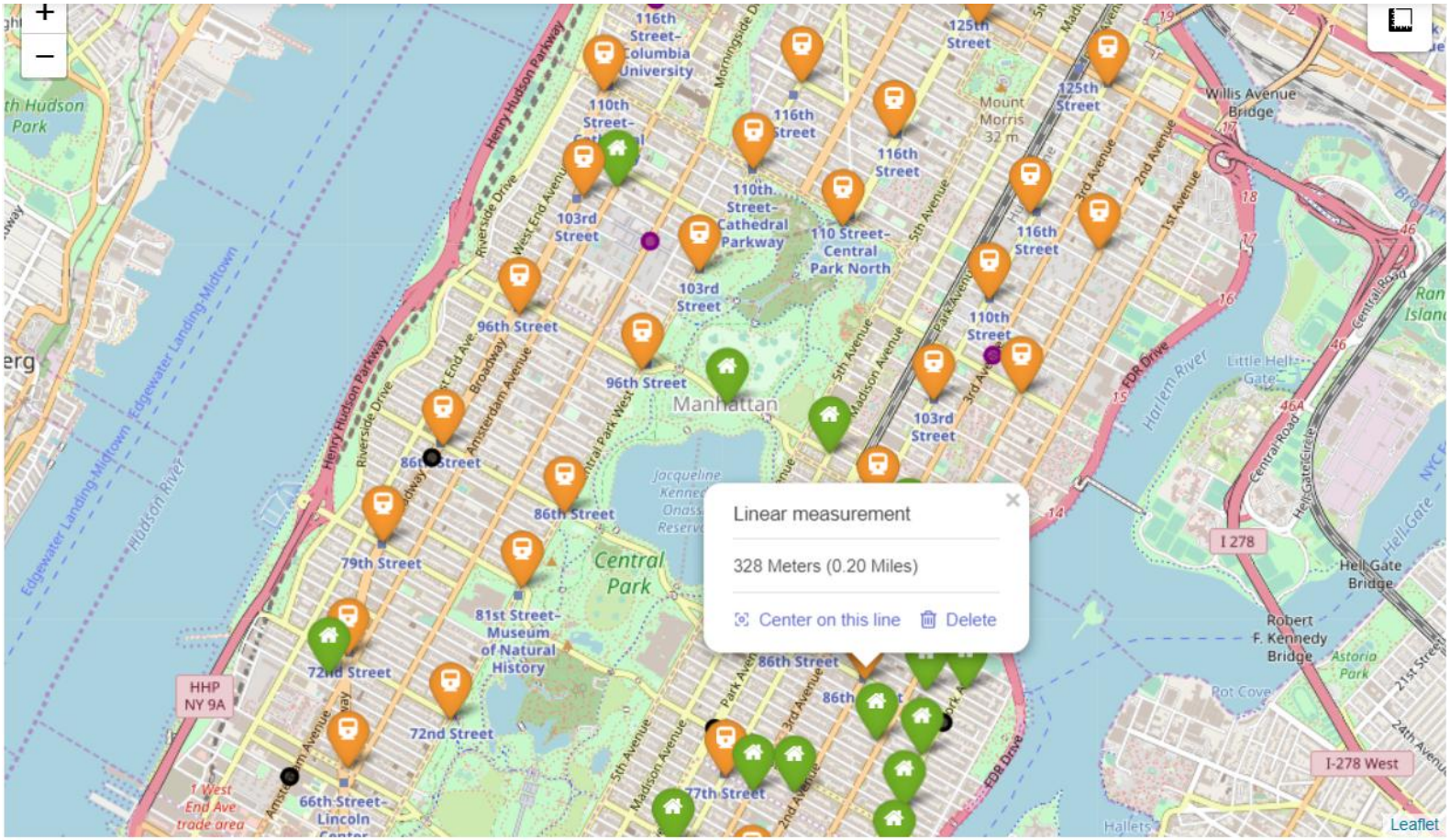
- **Cluster 3:**

```
# Cluster 3
#Cluster Label: Multi-cuisine restaurants
manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 3, manhattan_merged.columns[[1] + list(range(5, manhattan_merged.shape
```

	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	Marble Hill	Discount Store	Coffee Shop	Yoga Studio	Clothing Store	Tennis Stadium	Big Box Store	Supplement Shop	Spa	Shoe Store	Seafood Restaurant
2	Washington Heights	Café	Bakery	Mobile Phone Shop	Deli / Bodega	Gym / Fitness Center	Supermarket	Spanish Restaurant	Latin American Restaurant	Mexican Restaurant	Tapas Restaurant
3	Inwood	Mexican Restaurant	Café	Deli / Bodega	Lounge	Pizza Place	Restaurant	Chinese Restaurant	Frozen Yogurt Shop	Coffee Shop	Bakery
4	Hamilton Heights	Mexican Restaurant	Coffee Shop	Café	Pizza Place	Deli / Bodega	Liquor Store	Indian Restaurant	Sushi Restaurant	Sandwich Place	Yoga Studio
5	Manhattanville	Seafood Restaurant	Mexican Restaurant	Park	Italian Restaurant	Coffee Shop	Chinese Restaurant	Food & Drink Shop	Bike Trail	Sushi Restaurant	Supermarket
6	Central Harlem	African Restaurant	American Restaurant	Public Art	Gym / Fitness Center	French Restaurant	Chinese Restaurant	Seafood Restaurant	Cycle Studio	Bookstore	Ethiopian Restaurant
7	East Harlem	Mexican Restaurant	Bakery	Deli / Bodega	Latin American Restaurant	Thai Restaurant	Pet Store	Convenience Store	Dance Studio	Beer Bar	Pizza Place

- Using Folium library in Python, The final map will be created that shows details about **rental places**, **subway stations** and **clusters of venues** together. It will also provide a facility to check the distance between a subway station and a rental place or a distance between a rental place and a venue. This will help to make a good choice of a similar apartment in Manhattan NY.

Map with subway stations, rental places and clusters of venues



- By comparing the top 5 categories of the venues in London with these 4 clusters, we can say that **Cluster 0** resembles more to the venues in neighborhood in London .

- From cluster 0, "**Yorkville**" is the best choice for selecting a rental apartment as it has many options available for rental places and also subway stations are nearby to these rental places.

Conclusion and future directions

- Both Manhattan and London are the big cities. However, in this project when the venues are explored using the Foursquare API the limit is set to **100 venues** and the radius **500 meter.** For more detailed and accurate result from K-means, the data set can be expanded by increasing these limits.

- Also for getting the rental places details only limited number of websites are explored. By exploring more number of the websites we can provide more options to the audience of this project for rental place selection.

- A project/map where the detailed information about transportation, rental places, venues(that offers characteristics and benefits similar to his current neighborhood) in a particular city are mentioned together will certainly be helpful to the people moving to the new city for a job or for those people who just want to know about facilities available in a particular city.

****

Thank You