

THE BATTLE OF THE NEIGHBORHOODS

**Finding Best Location in New York to open
Ethiopian Restaurant**

GEBRU M W CAPSTONE PROJECT WEEK-2-PART-2

I. INTRODUCTION

- The City of New York is the most populous city in the United States.
- In 2018, population of around 8.4 million people distributed over a land area of about 302.6 square miles (784 km²)
- It is diverse and is the financial capital of USA.
- It is multicultural.
- It provides lot of business opportunities and business friendly environment.
- The market is highly competitive.
- Cost of doing business is also one of the highest.



Problem Description:

- A restaurant is a business which prepares and serves food and drink to customers in return for money.
- The City of New York is famous for its excellent cuisine.
- Its food culture includes an array of international cuisines influenced by the city's immigrant history.
- To survive in such competitive market it is very important to strategically plan.
- Various factors need to be studied to decide on the Location



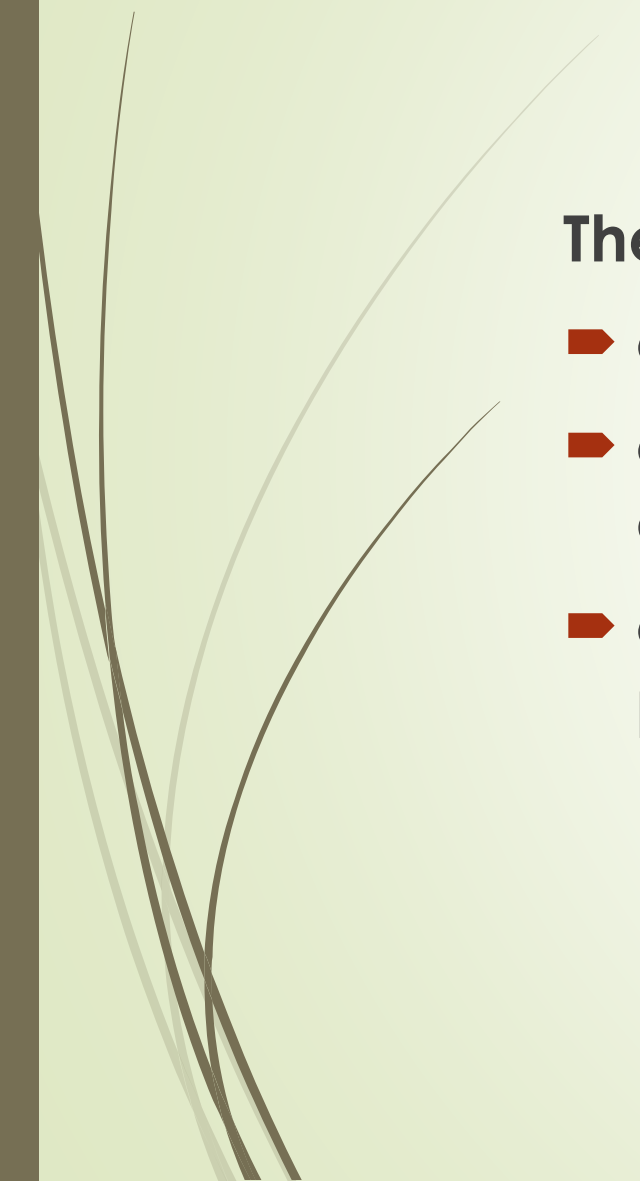


Target location:

- Neighborhood of New York
- Locate and recommend to the management which borough/neighborhood of New York will be best choice to open new Ethiopian restaurant
- The Management also expects to understand the rationale of the recommendations made.



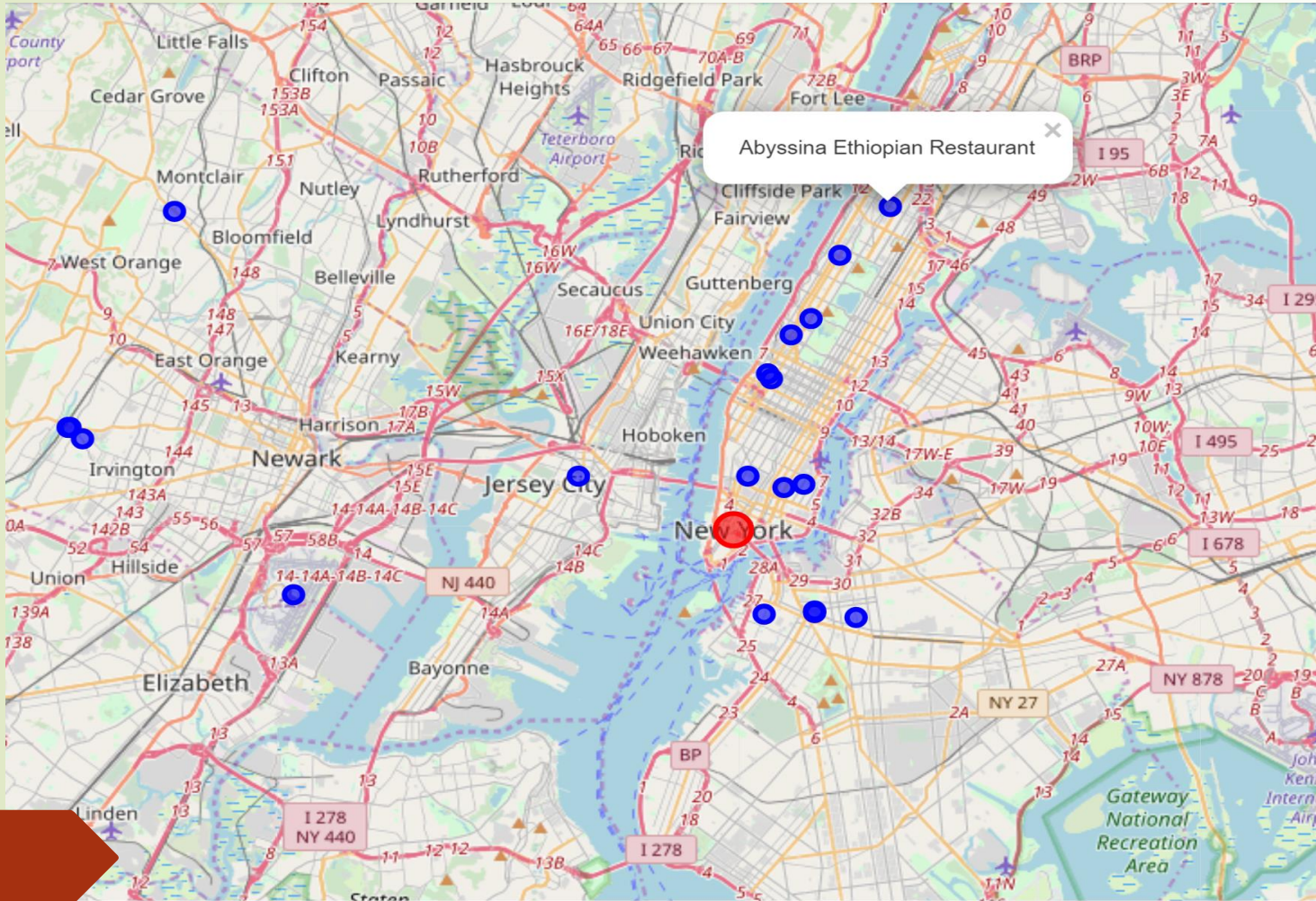
The location

- close enough to other vegan restaurants.
 - consider the distribution other Ethiopian restaurants in different Boroughs
 - close enough to other amenities and subways as well as popular venues.
- 

II. Data Used

- **Data 1:** New York a dataset that contains the 5 boroughs and the neighborhoods that exist in each borough as well as the latitude and longitude coordinates of each neighborhood.
 - https://geo.nyu.edu/catalog/nyu_2451_34572
- **Data 2:** For the New York Population, New York City Demographics, description of Boroughs analysis we will get data from Wikipedia as given below:
 - https://en.wikipedia.org/wiki/Boroughs_of_New_York_City
 - https://en.wikipedia.org/wiki/New_York_City
 - https://en.wikipedia.org/wiki/Economy_of_New_York_City
 - https://en.wikipedia.org/wiki/Portal:New_York_City
 - https://en.wikipedia.org/wiki/List_of_Michelin_starred_restaurants_in_New_York_City
- **Data 3 & 4:** New York city geographical coordinates data Distribution of Ethiopian Restaurants
 - <https://foursquare.com/developers/apps>

III. Methodology and Result

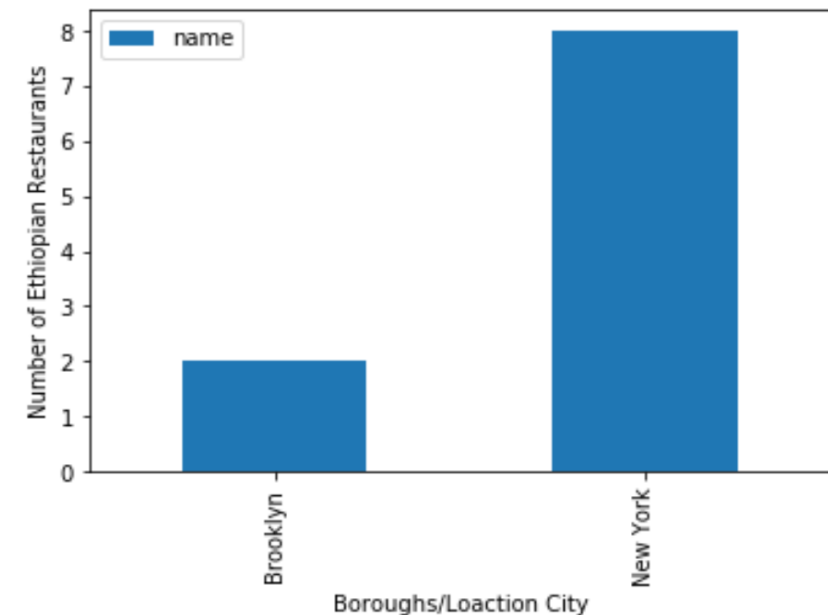


- python libraries and packages were downloaded
- defined Foursquare developer credentials and its Version
- converting the Central New York address to its latitude and longitude coordinates
- Get distribution of Ethiopian Restaurants in New York's five Boroughs

- After cleaning, the restaurant data was reduced from 22 to 10 restaurants.
- The restaurants were only distributed only in two Borough

	name	location.city
0	Meskerem Ethiopian Restaurant	New York
1	Awash Ethiopian Restaurant	Brooklyn
2	Haile Ethiopian Cuisine	New York
3	Bati Ethiopian Restaurant	Brooklyn
4	Ethiopian Meskerem Restaurant	New York
5	Bari Ethiopian Kitchen	New York
6	Nile Ethiopian Restaurant	New York
7	Abyssina Ethiopian Restaurant	New York
8	Meske	New York
9	Awash	New York

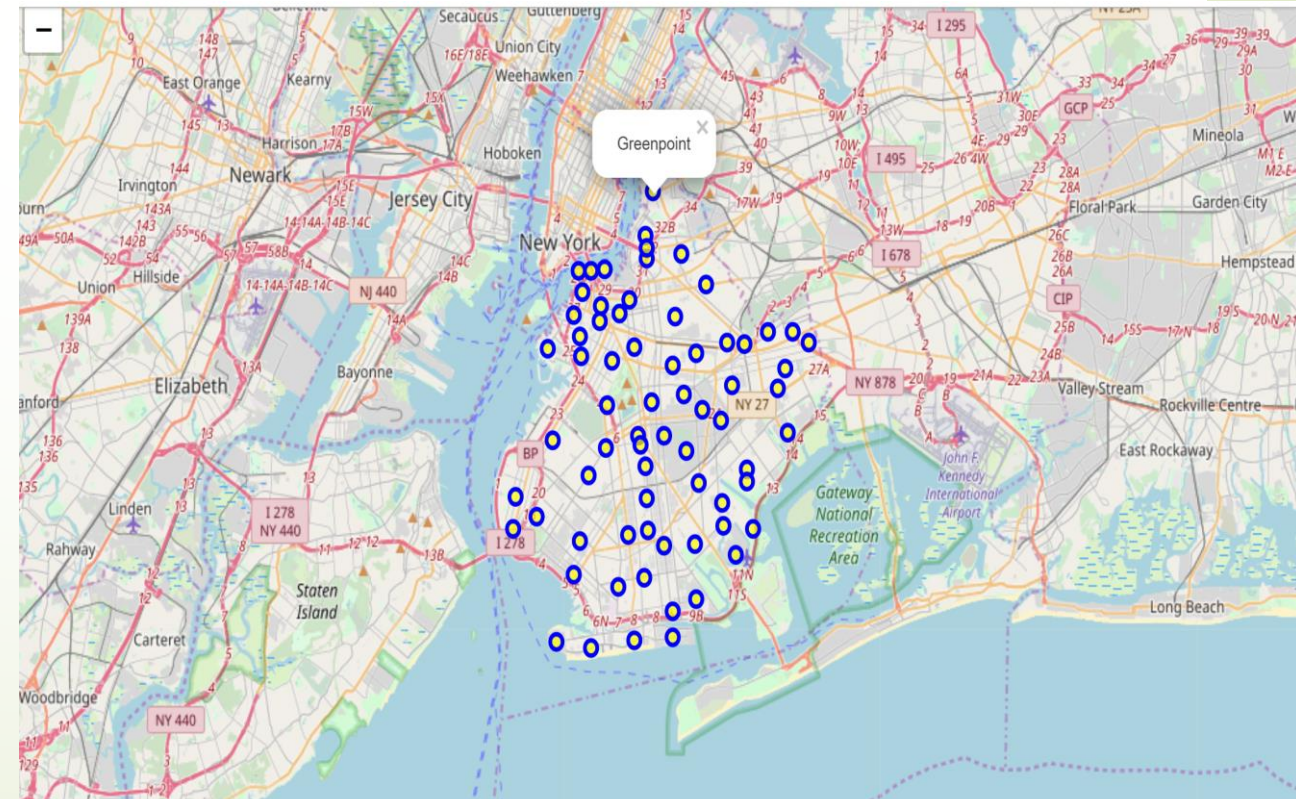
NEW YORK BOROUGH AND NUMBER OF ETHIOPIAN RESTAURANTS



Downloading and Exploring Dataset for New York

- The data was loaded and transformed into a pandas data frame and grouped by Borough.
- Sliced the original data frame and created a new data frame of the Brooklyn data
- Data further segmented and clustered only for the neighborhoods in Brooklyn,
- Brooklyn was chosen because it seem to have the least number of Ethiopian Restaurants compared to Manhattan (New York)

	Neighborhood	Latitude	Longitude
Borough			
Bronx	52	52	52
Brooklyn	70	70	70
Manhattan	40	40	40
Queens	81	81	81
Staten Island	63	63	63



➔ The coordinates of each Neighborhood were retrieved and a function created to get the venues in all the neighborhoods in Brooklyn.

➔ Then created new data frame to display the top 5 venues for each neighborhood

```
#check the size of the resulting frame  
print(Brooklyn_venues.shape)  
Brooklyn_venues.head()
```

(2851, 7)

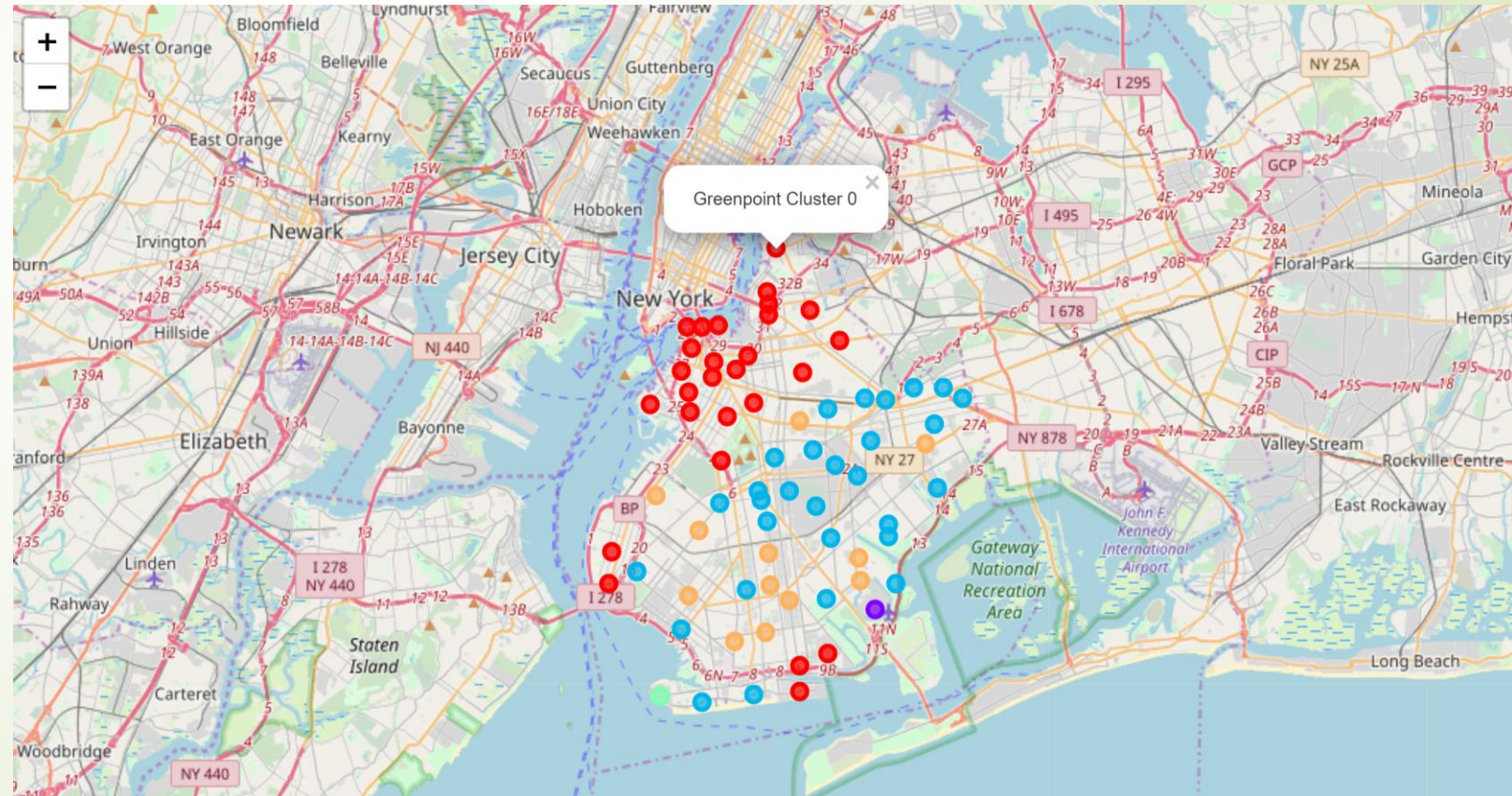
	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Bay Ridge	40.625801	-74.030621	Pilo Arts Day Spa and Salon	40.624748	-74.030591	Spa
1	Bay Ridge	40.625801	-74.030621	Bagel Boy	40.627896	-74.029335	Bagel Shop
2	Bay Ridge	40.625801	-74.030621	Cocoa Grinder	40.623967	-74.030863	Juice Bar
3	Bay Ridge	40.625801	-74.030621	Pegasus Cafe	40.623168	-74.031186	Breakfast Spot
4	Bay Ridge	40.625801	-74.030621	Ho' Brah Taco Joint	40.622960	-74.031371	Taco Place

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Bath Beach	Pharmacy	Pizza Place	Fast Food Restaurant	Sushi Restaurant	Chinese Restaurant
1	Bay Ridge	Italian Restaurant	Spa	Pizza Place	American Restaurant	Thai Restaurant
2	Bedford Stuyvesant	Pizza Place	Coffee Shop	Café	Deli / Bodega	Bar
3	Bensonhurst	Chinese Restaurant	Donut Shop	Sushi Restaurant	Ice Cream Shop	Italian Restaurant
4	Bergen Beach	Harbor / Marina	Athletics & Sports	Baseball Field	Playground	Donut Shop

Clustering Neighborhoods

- Used K means clustering
- set the number of clusters to 5 and create a new data frame that includes the cluster as well as the top 5 venues for each neighborhood.

	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Brooklyn	Bay Ridge	40.625801	-74.030621	0	Italian Restaurant	Spa	Pizza Place	American Restaurant	Thai Restaurant
1	Brooklyn	Bensonhurst	40.611009	-73.995180	4	Chinese Restaurant	Donut Shop	Sushi Restaurant	Ice Cream Shop	Italian Restaurant
2	Brooklyn	Sunset Park	40.645103	-74.010316	4	Pizza Place	Mexican Restaurant	Bakery	Latin American Restaurant	Bank
3	Brooklyn	Greenpoint	40.730201	-73.954241	0	Bar	Coffee Shop	Cocktail Bar	Pizza Place	Bakery
4	Brooklyn	Gravesend	40.595260	-73.973471	4	Pizza Place	Deli / Bodega	Spa	Bakery	Italian Restaurant





IV. Discussion

- used data from different sources to get the distribution of Ethiopian restaurants in New York and its Boroughs in addition to the venues in each Boroughs.
- Some data retrieved from Foursquare are vague, duplicate and even some business are wrongly labeled as Restaurants.
- I used K mean Clustering machine learning algorithm for predictive analysis
- After clustering the best location identified for the new restaurants.
 - The identified location has great venues and other amenities
 - the competition from other similar restaurants is not significant.



V. Conclusion

- This project has used different sources of data, exploratory data analysis and algorithm to identify the best location.
 - Cluster 0 or the first cluster seems to have the most venues in Brooklyn and will be the first choice to open the new Ethiopian Restaurant.
 - This location can be further studied and refined using additional data and analysis.
- 