Recommendation System for places to Open a Food Court Chain in New York City

By S M Azizul Hakim

Background

- The City of New York is the most populous city in the United States.
- New York City has been described as the cultural, financial, and media capital of the world.
- New York City is home to more than 3.2 million residents born outside the United States, the largest foreign-born population of any city in the world.

Business Problem

- New York is the most diverse city in the United States.
- The city manages to bring dishes from all over the world and make them it's own.
- Over the past few decades, food courts have become an important way of experiencing the cities diverse food culture at a reasonable price.
- This project aims to help any business person interested in opening a food court chain in New York City.
- Our goal is to recommend the best places to open a food court in different neighborhoods.

Data preparation

- New York City consists of 5 boroughs and 306 neighborhoods.

- There is a opensource dataset with each neighborhood as well as the latitude and the longitude of each neighborhood.

- We combined this data with the foursquare venue data to create the final dataset.
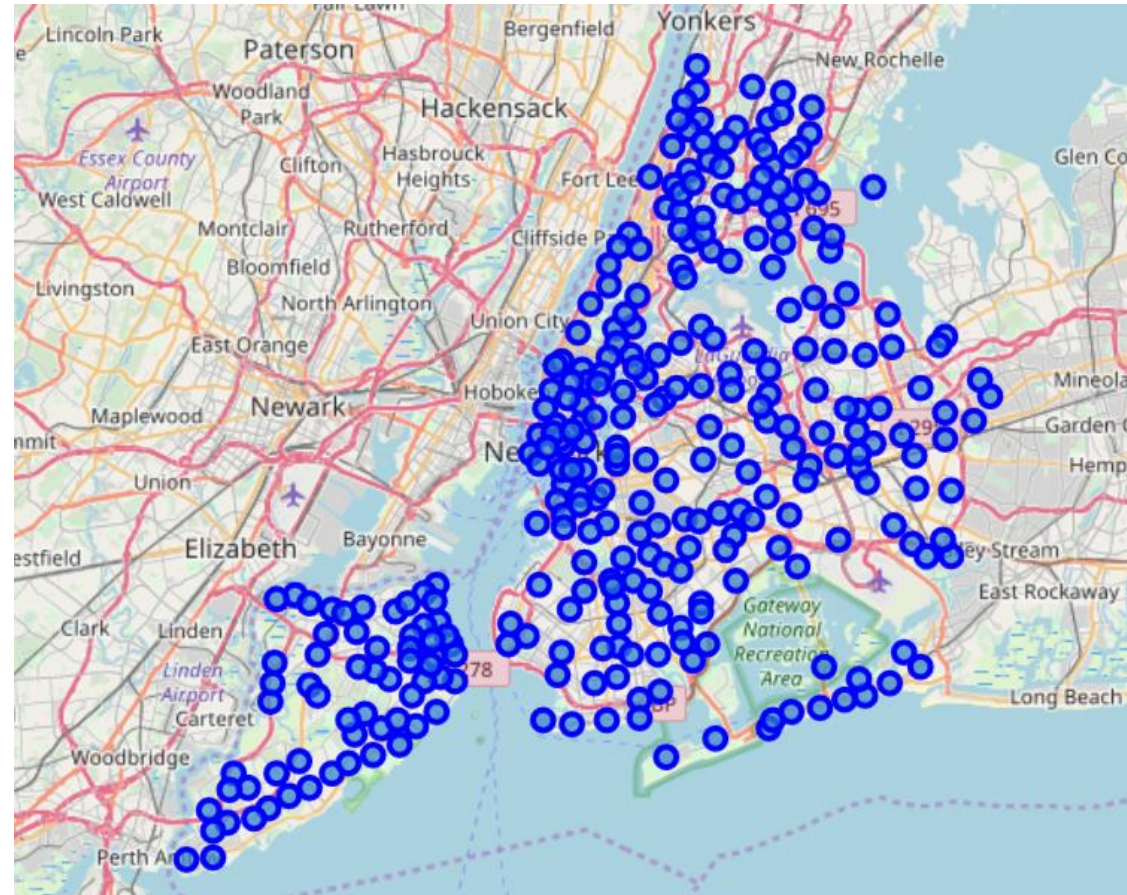
Data preparation(contd.)

Here is the final dataframe

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Wakefield	40.894705	-73.847201	St. Theresa School	40.849301	-73.832999	College Academic Building
1	Wakefield	40.894705	-73.847201	O'Hare Hall	40.859291	-73.882599	College Residence Hall
2	Wakefield	40.894705	-73.847201	Fordham University - Rose Hill	40.861041	-73.885490	University
3	Wakefield	40.894705	-73.847201	Monroe College Jerome Hall & Campus Library	40.863212	-73.900518	General College & University
4	Wakefield	40.894705	-73.847201	Bronxnet	40.871987	-73.896252	College Administrative Building

Data Preparation(contd.)

All the neighborhoods in map:

Exploratory Data Analysis

We have applied soe analysis on the data:

- The dataframe is grouped by 'neighborhood' attribute. This shows how many desired venues located in each neighborhoods.

- Calculated the number of unique categories in the dataframe.

- We grouped rows by neighborhood and by taking the mean of the frequency of occurrence of each category.

- Then we sorted the venues in descending order. Then we created a new dataframe displaying the top 10 venues for each neighborhood.

Exploratory Data Analysis(contd.)

A new dataframe displaying the top 10 venues for each neighborhood:

	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	Allerton	College Academic Building	College Administrative Building	General College & University	College Residence Hall	College Library	College Quad	College Cafeteria	Medical School	College Gym	Student Center
1	Annadale	University	Sorority House	General College & University	College Classroom	College Academic Building	Fraternity House	Student Center	Gym	Trade School	College Theater
2	Arden Heights	University	General College & University	Sorority House	College Academic Building	College Classroom	College Gym	Fraternity House	College Bookstore	Gym	Trade School
3	Arlington	College Academic Building	College Administrative Building	College Science Building	General College & University	Community College	University	Student Center	College & University	College Cafeteria	Gym / Fitness Center
4	Arrochar	General College & University	College Administrative Building	College Residence Hall	Student Center	University	College Academic Building	College Classroom	College Science Building	College Quad	College Football Field

Clustering

We used K-means clustering and the number of clusters is set to 10:

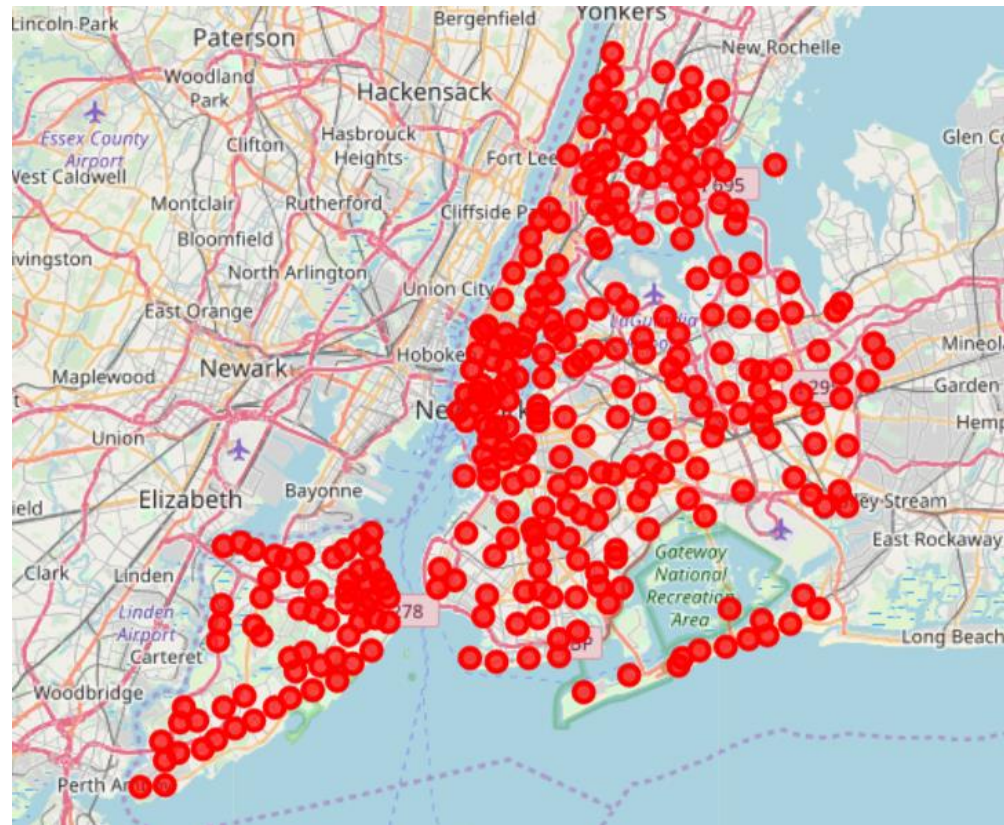|	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Allerton	College Academic Building	College Administrative Building	General College & University	College Residence Hall	College Library	College Quad	College Cafeteria	Medical School	College Gym	Student Center
1	Annadale	University	Sorority House	General College & University	College Classroom	College Academic Building	Fraternity House	Student Center	Gym	Trade School	College Theater
2	Arden Heights	University	General College & University	Sorority House	College Academic Building	College Classroom	College Gym	Fraternity House	College Bookstore	Gym	Trade School
3	Arlington	College Academic Building	College Administrative Building	College Science Building	General College & University	Community College	University	Student Center	College & University	College Cafeteria	Gym / Fitness Center
4	Arrochar	General College & University	College Administrative Building	College Residence Hall	Student Center	University	College Academic Building	College Classroom	College Science Building	College Quad	College Football Field

Clustering(contd.)

Clustered venues with the neighborhood they are located in:

	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	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Bronx	Wakefield	40.894705	-73.847201	6	College Academic Building	General College & University	College Administrative Building	College Residence Hall	College Quad	College Gym	College Cafeteria	College Library	College Classroom	University
1	Bronx	Co-op City	40.874294	-73.829939	6	College Academic Building	College Residence Hall	College Administrative Building	College Cafeteria	College Gym	College Library	Medical School	General College & University	Student Center	University
2	Bronx	Eastchester	40.887556	-73.827806	6	College Academic Building	College Residence Hall	College Gym	General College & University	University	College Cafeteria	College Classroom	Medical School	College Library	College Auditorium
3	Bronx	Fieldston	40.895437	-73.905643	6	College Academic Building	General College & University	College Administrative Building	College Gym	College Library	College Residence Hall	College Quad	Student Center	College Classroom	College Cafeteria
4	Bronx	Riverdale	40.890834	-73.912585	6	College Academic Building	General College & University	College Administrative Building	College Library	College Residence Hall	College Gym	College Quad	Student Center	College Classroom	College Cafeteria

Results

We finally visualize the resulting clusters by adding them to a map of New York City using folium:

Results(contd.)

Here are some of the results of clustering

Cluster-1:

	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
169	Rosedale	College Academic Building	General College & University	College Library	Student Center	College Classroom	Fraternity House	University	College Gym	College Rec Center	College Administrative Building
188	Laurelton	Student Center	College Academic Building	College Library	General College & University	College Gym	College Classroom	College Rec Center	Fraternity House	Trade School	Medical School
193	Brookville	College Academic Building	Student Center	College Library	College Classroom	Fraternity House	General College & University	College Gym	College Rec Center	University	Trade School
213	Great Kills	General College & University	College Classroom	College Academic Building	College Science Building	University	Trade School	Sorority House	College Administrative Building	Fraternity House	College Communications Building
214	Eltingville	General College & University	College Academic Building	College Classroom	College Science Building	University	Trade School	Sorority House	College Administrative Building	College Lab	Fraternity House
215	Annadale	University	Sorority House	General College & Classroom	College Classroom	College Academic	Fraternity House	Student Center	Gym	Trade School	College Theater

Results(contd.)

Cluster-5:

	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
38	Schuylerville	College Academic Building	College Classroom	College Residence Hall	Student Center	College Rec Center	Fraternity House	College Administrative Building	Medical School	University	High School
40	Castle Hill	College Academic Building	College Classroom	Medical School	College Residence Hall	University	College Rec Center	College Administrative Building	Trade School	High School	Student Center
47	Bensonhurst	College Academic Building	General College & University	College Classroom	College Quad	University	Student Center	Elementary School	College Arts Building	College Cafeteria	College Bookstore
48	Sunset Park	College Academic Building	General College & University	College Classroom	College Library	College Rec Center	College Quad	Elementary School	College Cafeteria	Medical School	College Administrative Building
50	Gravesend	College Academic Building	General College & University	College Classroom	College Arts Building	College Cafeteria	College Math Building	College Quad	Elementary School	College Bookstore	College Library
51	Brighton Beach	College Academic Building	College Classroom	General College & University	College Math Building	College Arts Building	College Quad	College Cafeteria	College Library	College Science Building	Medical School

Results(contd.)

Cluster-10:

	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
72	East New York	College Academic Building	General College & University	College Administrative Building	Student Center	College Classroom	College Science Building	Medical School	College Library	University	School
73	Starrett City	College Academic Building	Student Center	General College & University	College Administrative Building	College Science Building	College Library	College Classroom	Medical School	College Cafeteria	School
140	Sunnyside	College Academic Building	College Administrative Building	General College & University	College Library	College Science Building	University	Student Center	College Residence Hall	College Communications Building	College Cafeteria
155	Glen Oaks	General College & University	College Residence Hall	Medical School	College Academic Building	University	College Library	College Arts Building	College Administrative Building	College Science Building	College Lab
156	Bellerose	General College & University	College Administrative Building	College Academic Building	College Residence Hall	College Cafeteria	Medical School	College Library	University	College Bookstore	College Quad
168	Cambria Heights	College Administrative Building	Student Center	College Cafeteria	General College & University	College Residence Hall	College Gym	College Classroom	College Academic Building	College Bookstore	University

Conclusion

- We can use the approximate number of students in a given venue to make better decision.

- As food courts are mobile shops, we can move them from time to time to better exploit the recommendation.

- These will require additional timeframe attribute to figure out places are most occupied by students between which time. This will increase our recommendation result more viable and robust.

Thank You