RANJITHKUMAR PALANIAPPAN

AN ENTREPRENEURIAL SEARCH TO OPEN A CHAIN OF INDIAN RESTAURANTS IN NEW YORK - CAPSTONE PROJECT

I. Introduction:

New York, The City of Dreams is known for its multicultural community of people at all facets of life. The bustling city paves way to a plethora of business opportunities at its dispersal. The major businesses are in the areas of banking and finance, retailing, world trade, transportation, tourism, F&B, real estate, news media, traditional media, advertising, legal services, accountancy, insurance, theater, fashion, and fine arts.

Even though the ease of doing business is almost negligible, there's a neck to neck competition prevailing among companies at various domains. Leveraging modern technologies for increasing the capabilities of every single business became the modus operandi. In regard to that, the utilization of Data Science will help in the increase of ROI, reduction of risk while looking for expanding or venturing new businesses.

II. Business Problem:

The cuisine of New York City comprises many cuisines belonging to various ethnic groups that have entered the United States through the city. Almost all ethnic cuisines are well represented in New York City, both within and outside the various ethnic neighborhoods influenced by the city's immigrant history. The bustling city with a growing number of New Yorkers, Indians and non-Indians, are mulling over the passage to India through spicy delicious foods. There are dozens of pricy fancier Indian restaurants across the city. So to distinguish from them, it is not only necessary to focus on the product offerings of a new franchise, but also to select the right locations for it. Various factors need to be studied in order to decide on the Location, and to strategically open further branches of it.

Target Audience:

Manhattan has a long tradition of formal Indian restaurants, often overseen by cooks who learned to present their country's cuisine in a fine-dining idiom by hard-core training in India's extensive system of hotel kitchens. At the other end of the scale are the no-frills places where value makes style irrelevant. One of my Acquaintance wants to open his business in New York or in the surrounding boroughs, so I focus on developing a concrete evidence. The objective is to locate and recommend to the management which neighborhood of New York city will be best choice to start the first restaurant. The Management also expects to understand the rationale of the recommendations made.

III. Data:

The objective is to find the best locations for opening the new Indian restaurant:

- so we are going to analyze the dataset corresponding to New York Neighborhoods. New York comprises of a total of 5 boroughs and 306 neighborhoods. In order to segment the neighborhoods and explore them, we will essentially need a dataset that contains the 5 boroughs and the representative neighborhoods that exist in each borough as well as the latitude and longitude coordinates.

This dataset exists for free at the NYU Spatial Data Repository. Link to the dataset is:

https://geo.nyu.edu/catalog/nyu_2451_34572

```
neighborhoods.head()
```

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585

- From Foursquare Venues Categories - https://developer.foursquare.com/docs/resources/categories

Category "Indian" cuisine's id: 4bf58dd8d48988d10f941735

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Marble Hill	40.876551	-73.910660	Riverdale Indian Cuisine	40.880886	-73.908800	Indian Restaurant
1	Marble Hill	40.876551	-73.910660	Tazmohol Indian Restaurant	40.879331	-73.903192	Indian Restaurant
2	Marble Hill	40.876551	-73.910660	Cumin Indian Cuisine	40.886459	-73.909816	Indian Restaurant
3	Chinatown	40.715618	-73.994279	Kabab Bites	40.720094	-73.995819	Indian Restaurant
4	Chinatown	40.715618	-73.994279	Hampton Chutney Co.	40.720220	-73.999347	Dosa Place

IV.Methodology

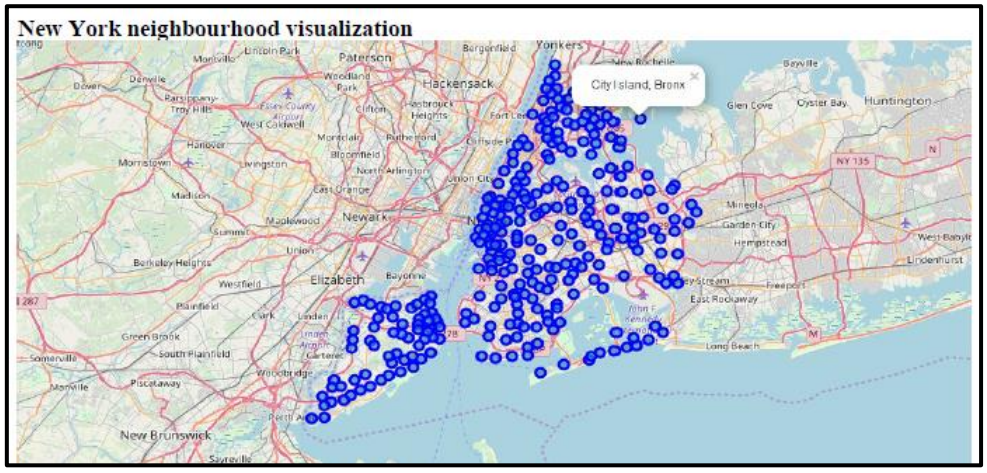
Our main goal is to get optimum location for the first branch of the chain of restaurants in New York City.

Analytic Approach:

New York city neighborhood has a total of 5 boroughs and 306 neighborhoods. In this project clustering of venues will be done as part of the following Exploratory data analysis.

Data 1- New York city Geographical Coordinates Data

1. In this we load the data and explore data from newyork data.json file.
2. Transform the data of nested python dictionaries into a pandas dataframe.
3. This dataframe contains the geographical coordinates of New York city neighborhoods.
4. Addresses are converted into their equivalent latitude and longitude values.
5. We used geopy and folium libraries to create a map of New York city with neighborhoods superimposed on top.



Data 2: New York city geographical coordinates data has been utilized as input for the Foursquare API with explore endpoint, which is leveraged to provision venues information for each neighborhood. We used the Foursquare API data to explore neighborhoods in New York City.

New York Neighborhood Visualization w.r.t Indian restaurants

Then using this feature to group the neighborhoods into clusters K-means clustering algorithm to complete the task. And also, the Folium library to visualize the neighborhoods in Manhattan and its emerging clusters.

TOP VENUES

	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
0	Battery Park City	Indian Restaurant	Food Truck	Tapas Restaurant	Asian Restaurant	Vegetarian / Vegan Restaurant	South Indian Restaurant	Pakistani Restaurant
1	Carnegie Hill	Indian Restaurant	North Indian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	Indian Chinese Restaurant
2	Central Harlem	Indian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant
3	Chelsea	Indian Restaurant	Food Truck	Vegetarian / Vegan Restaurant	Indian Chinese Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant
4	Chinatown	Indian Restaurant	North Indian Restaurant	Dosa Place	Deli / Bodega	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant

V. Results:

i) Very High Competition with 16 groups

Cluster 0 : Saturated Market

```
anhattan_merged[manhattan_merged['Cluster Labels'] == 0].reset_index(drop=True)
```

	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	Manhattan	Chinatown	40.715618	-73.994279	0	Indian Restaurant	North Indian Restaurant	Dosa Place	Deli / Bodega	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	Indian Chinese Restaurant	Hookah Bar
1	Manhattan	Upper East Side	40.775639	-73.960508	0	Indian Restaurant	North Indian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	Indian Chinese Restaurant	Hookah Bar	Himalayan Restaurant	Food Truck
2	Manhattan	Yorkville	40.775930	-73.947118	0	Indian Restaurant	North Indian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	Indian Chinese Restaurant	Hookah Bar	Himalayan Restaurant	Food Truck
3	Manhattan	Lenox Hill	40.768113	-73.958860	0	Indian Restaurant	North Indian Restaurant	Asian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	Indian Chinese Restaurant	Hookah Bar	Himalayan Restaurant

ii) Moderate competition with 8 groups

Cluster 1 : Moderately competetive Market

```
manhattan_merged[manhattan_merged['Cluster Labels'] == 1].reset_index(drop=True)
]:
```

	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	Manhattan	Marble Hill	40.876551	-73.910660	1	Indian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant	Hookah Bar	Himalayan Restaurant	Food Truck
1	Manhattan	Washington Heights	40.851903	-73.936900	1	Indian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant	Hookah Bar	Himalayan Restaurant	Food Truck
2	Manhattan	Hamilton Heights	40.823604	-73.949688	1	Indian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant	Hookah Bar	Himalayan Restaurant	Food Truck
3	Manhattan	Manhattanville	40.816934	-73.957385	1	Indian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant	Hookah Bar	Himalayan Restaurant	Food Truck
4	Manhattan	Central Harlem	40.815976	-73.943211	1	Indian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant	Hookah Bar	Himalayan Restaurant	Food Truck

iii) Moderate Competition with 6 groups

Cluster 2 : Moderately competetive Market

```
nhattan_merged[manhattan_merged['Cluster Labels'] == 2].reset_index(drop=True)
```

	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	Manhattan	Clinton	40.759101	-73.996119	2	Indian Restaurant	Food Truck	South Indian Restaurant	Diner	Vegetarian / Vegan Restaurant	Tapas Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant	Hookah Bar
1	Manhattan	Chelsea	40.744035	-74.003116	2	Indian Restaurant	Food Truck	Vegetarian / Vegan Restaurant	Indian Chinese Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Hookah Bar	Himalayan Restaurant
2	Manhattan	Battery Park City	40.711932	-74.016869	2	Indian Restaurant	Food Truck	Tapas Restaurant	Asian Restaurant	Vegetarian / Vegan Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant	Hookah Bar
3	Manhattan	Financial District	40.707107	-74.010665	2	Indian Restaurant	Food Truck	Tapas Restaurant	Asian Restaurant	Vegetarian / Vegan Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant	Hookah Bar
4	Manhattan	Civic Center	40.715229	-74.005415	2	Indian Restaurant	Food Truck	Tapas Restaurant	Dosa Place	Asian Restaurant	Vegetarian / Vegan	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese

iv) Low competition with only 1 group

Cluster 3 : Untapped Market

```
manhattan_merged[manhattan_merged['Cluster Labels'] == 3].reset_index(drop=True)
```

	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	Manhattan	Stuyvesant Town	40.731	-73.974052	3	Indian Restaurant	Vegetarian / Vegan Restaurant	North Indian Restaurant	Dosa Place	Deli / Bodega	Chaat Place	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	Indian Chinese Restaurant

v) High competition with 9 groups

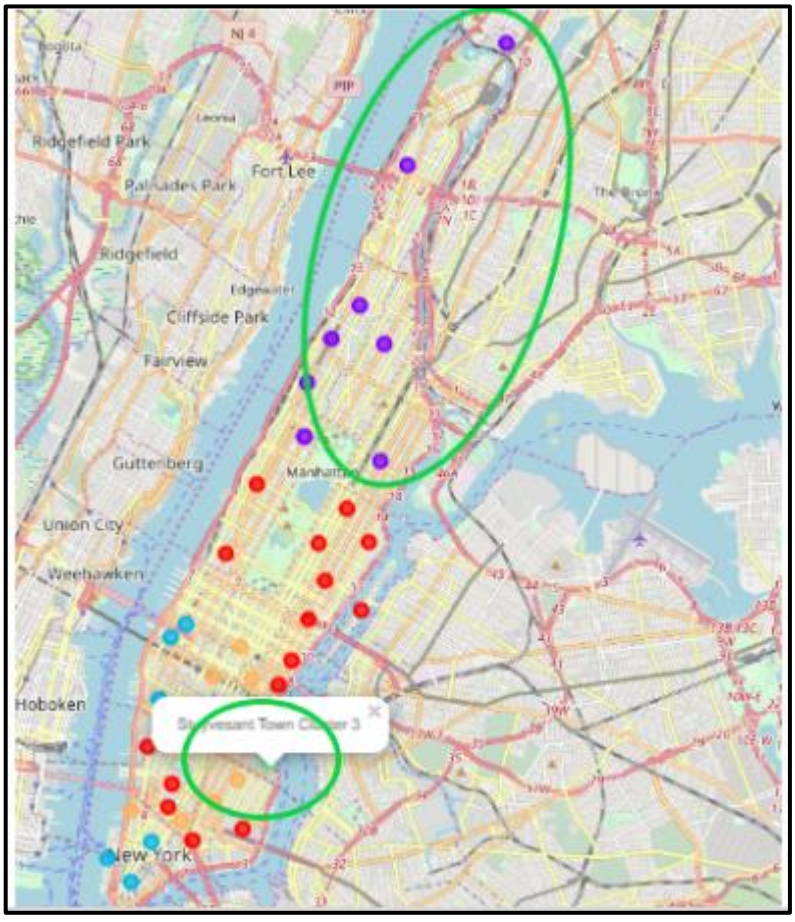
Cluster 4 : Moderately competetive Market

```
manhattan_merged[manhattan_merged['Cluster Labels'] == 4].reset_index(drop=True)
```

	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	Manhattan	Midtown	40.754691	-73.981669	4	Indian Restaurant	Food Truck	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Diner	Vegetarian / Vegan Restaurant	Tapas Restaurant	Indian Chinese Restaurant	Hookah Bar
1	Manhattan	Murray Hill	40.748303	-73.978332	4	Indian Restaurant	South Indian Restaurant	Food Truck	Vegetarian / Vegan Restaurant	North Indian Restaurant	Diner	Tapas Restaurant	Pakistani Restaurant	Indian Chinese Restaurant	Hookah Bar
2	Manhattan	East Village	40.727847	-73.982226	4	Indian Restaurant	Dosa Place	Deli / Bodega	Vegetarian / Vegan Restaurant	North Indian Restaurant	Food Truck	Chaat Place	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant
3	Manhattan	Tribeca	40.721522	-74.010683	4	Indian Restaurant	Food Truck	Tapas Restaurant	Dosa Place	Asian Restaurant	Vegetarian / Vegan Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant

vi) We can notice that within 3 clusters can be opted for opening a new restaurant as follows:

- Stuyvesant Town Cluster 3
- Washington heights Cluster 1
- Civic center Cluster 2



VI. Discussion:

1. There is scope to include Farmers markets data analysis to optimize procurement for the restaurants.
2. There is scope to explore cuisines of Chinese and Japanese to leverage furthermore business opportunities.

VII. Conclusion:

This analysis is performed on limited data to explore the possibilities of analytical approach towards finding the best location for the franchise. To conclude, this analysis proves to be a start of selecting a low risk-high yield location. On further development, it can be rendered into an efficient evidence with better accuracy and precision.