**Capstone Project - The Battle of Neighborhoods**

**Introduction**

**New York City** (**NYC**), often called the **City of New York** or simply **New York** (**NY**), is the most populous city in the United States. With an estimated 2018 population of 8,398,748 distributed over about 302.6 square miles (784 km2), New York City's demographics show that it is a large and ethnically diverse metropolis. It is the largest city in the United States with a long history of international immigration. Over the last decade the city has been

growing faster than the region. The New York region continues to be by far the leading metropolitan gateway for legal immigrants admitted into the United States.  The city and its metropolitan area constitute the premier gateway for legal immigration to the United States. As many as 800 languages are spoken in New York, making it the most linguistically diverse city in the world. New York 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 as of 2016.The term "melting pot" was

coined to describe densely populated immigrant neighborhoods on the Lower East Side.

With its diverse culture, comes diverse food items. There are many restaurants in New York City, each belonging to different cuisines like Italian, Chinese, Indian, French, Japanese etc. A large population search for good Indian cuisines.

We will list and visualize Manhattan, New York City that has great Indian restaurants. We will also recommend where a new India restaurant can be opened based upon distance from other restaurants. The area of interest which is approx. 12x12 kilometers centered around Manhattan.

**Data**

For this project we will use the following data:

1. New York City data that contains list Boroughs, Neighborhoods along with their latitude and longitude.

Data source: <https://cocl.us/new_york_dataset>

Description: This data set contains the required information of New York city data. We will use this data set to explore various neighborhoods of New York city.

1. All restaurants and Indian restaurants in area of interest near to Manhatten of New York city.

Data source: Fousquare API

Description: By using this API we will get all the venues in each neighborhood. We can filter these venues to get only Indian restaurants.

1. GeoSpace data

Data source: **Google Maps API reverse geocoding**

Description: By using this geo space data we will get the New York Borough boundaries that will help us visualize choropleth map.

Based on definition of our problem, factors that will influence our decision are:

* number of existing restaurants in each of the three neighborhoods (any type of restaurant)
* number of and distance to Indian restaurants in each of the neighborhoods, if any
* distance of neighborhood from Manhattan

We decided to use regularly spaced grid of locations, centered around Manhattan, NY, to define our neighborhoods.

## Methodology

In this project we will direct our efforts on detecting areas of New York city that have low restaurant density, particularly those with low number of Indian restaurants. We will limit our analysis to area ~6km around Manhattan.

In first step we have collected the required **data: location and type (category) of every restaurant within 6km from Manhattan** . We have also **identified Indian restaurants** (according to Foursquare categorization).

Second step in our analysis will be calculation and exploration of '**restaurant density**' across different areas of Berlin - we will use **heatmaps** to identify a few promising areas close to center with low number of restaurants in general (and no Italian restaurants in vicinity) and focus our attention on those areas.

In third and final step we will focus on most promising areas and within those create **clusters of locations that meet some basic requirements** established in discussion with stakeholders: we will take into consideration locations with **no more than two restaurants in radius of 250 meters**, and we want locations **without Indian restaurants in radius of 400 meters**. We will present map of all such locations but also create clusters (using **k-means clustering**) of those locations to identify general zones / neighborhoods / addresses which should be a starting point for final 'street level' exploration and search for optimal venue location by stakeholders.

**Results:**

Total number of restaurants: 1492

Total number of Indian restaurants: 64

Percentage of Indian restaurants: 4.29%

Average number of restaurants in neighborhood: 7.3489010989010985

Average number of restaurants in every area with radius=300m: 7.3489010989010985

Average distance to closest Indian restaurant from each area center: 1398.5805344930536

OK, so on average Indian restaurant can be found within ~1400m\*\* from every area center candidate.

Looks like a few pockets of low restaurant density closest to city center can be found new East Harlem area.

The map is not so 'hot' (Indian restaurants represent a subset of ~4.29% of all restaurants in NYC) but it also indicates higher density of existing Indian restaurants between 42nd street and 57th street between 2nd and 3rd ave, with closest pockets of low Indian restaurant density positioned near East Harlem.

Based on this, we recommend East Harlem area as the preferred location for new Indian Restaurant

**Discussion**

Our analysis shows that although there is a great number of restaurants in New York City(~1500 in our initial area of interest which was 12x12km around Manhattan), there are pockets of low restaurant density fairly close to Manhattan. Highest concentration of restaurants was detected most of the Manhattan area with least density around East Harlem. The highest density of Indian restaurants is in between 42nd street and 57th street, close to 2nd and 3rd ave. Again, East Harlem area had low density of Indian restaurants compared to other parts of New York city.

**Conclusion**

Purpose of this project was to identify NYC areas close to Manhattan with low number of restaurants (particularly Indian restaurants) in order to aid stakeholders in narrowing down the search for optimal location for a new Indian restaurant. Based on our analysis of restaurant density in general around Manhattan and Indian restaurant density in particular, we recommend East Harlem as the preferred location for new Indian Restaurant.

Final decision on optimal restaurant location will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness of each location (proximity to park or water), levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.