

# **Capstone Project - Indian Restaurant in Queens,Ny**

**IBM/Coursera**

*Srinivasan Paramasivam*

*21/11/2019*

## **Table of contents**

- Introduction: Business Problem
- Data
- Methodology
- Analysis
- Results and Discussion
- Conclusion

## **Introduction: Business Problem:**

This report we will try to find optional location for restaurant business in Berlin. Since there are lots of restaurants in Queens,Ny

we will try to find :-

1. Locations that is less crowded with restaurants.
2. Locations with no Indian restaurants in vicinity
3. Locations close to city centre (as possible), (if first two conditions are met).

We will use our data science tools and techniques to generate a few most promising neighbourhoods locations based on above criteria. Advantages of each area will then be clearly expressed so that best possible final location can be chosen by stakeholders.

**Target Group** – Targeted of this report will be stakeholders interested in opening an Indian restaurant in Queens,Ny,USA.

## **Data**

To find a solution of above problem we would be looking for the following data:-

1. Number of existing restaurants in the neighbourhood
2. Number of Indian restaurants in the neighbourhood
3. Distance of Indian restaurants in the neighbourhood
4. Distance of neighbourhood from city center

To Define our neighbourhood we will regularly spaced grid of locations centered around city center.

**Data Source:**

1. Centers of candidate areas will be generated through algorithm
2. Google Maps API - Using Google Map API approximate addresses of generated centers will be obtained.
3. Coordinate of Berlin center will be obtained using Google Maps API geocoding of well known Berlin location (Queens)
4. Foursquare API - Number of restaurants, type and location in every neighbourhood will be obtained using Foursquare API
5. For Boundaries Of New York - [https://raw.githubusercontent.com/sj10bsrinivasan/Datascience/master/new\\_york.geojson](https://raw.githubusercontent.com/sj10bsrinivasan/Datascience/master/new_york.geojson)

**Coordinates of City Center:**

Using Google Maps geocoding API Latitude & longitude of Queens,Ny city center found with well known address “Queen, New York, USA”

Coordinate of Queens, NY, USA: [ 40.7282239, -73.7948516 ]

**Creating a Grid:**

A grid of cells covering our area of interest has been created which is approx. 12x12 kilo meters around Queens,Ny city center

**Candidate Area:**

A grid of candidate area, equally spaced, centered on city center and within ~6km from Queens,Ny has been created.

**Defining Neighbourhood:**

our neighbourhoods will be defined as circular areas with a radius of 300 meters, so our neighbourhood centers will be 600 meters apart

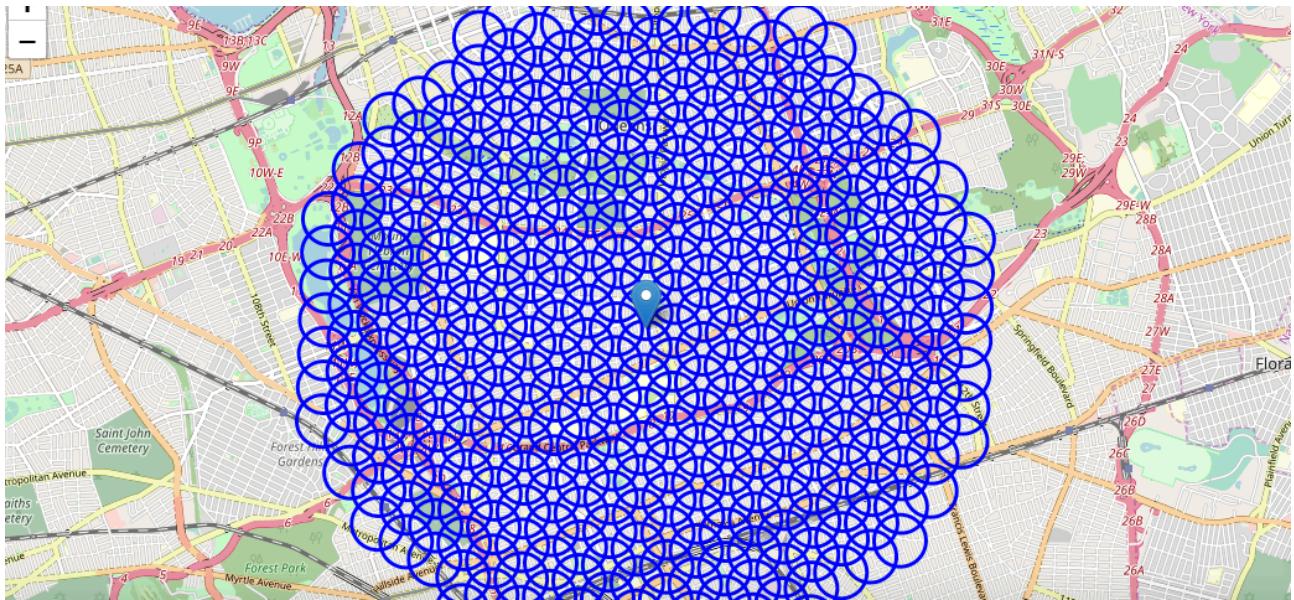
**Distance Calculation:**

To accurately calculate distances we need to create our grid of locations in Cartesian 2D coordinate system which allows us to calculate distances in meters (not in latitude/longitude degrees). Then we'll project those coordinates back to latitude/longitude degrees to be shown on Folium map. So let's create functions to convert between WGS84 spherical coordinate system (latitude/longitude degrees) and UTM Cartesian coordinate system (X/Y coordinates in meters).

**Hexagonal grid of cells:**

we offset every other row, and adjust vertical row spacing so that every cell center is equally distant from all its neighbours.

## Visualize City center location and candidate neighbourhood centers:



Number of Candidate Neighbourhood Center = **364**

	Address	Latitude	Longitude	X	Y	Distance from center
0	93-46 210th Pl, Queens Village, NY 11428	40.716520	-73.751049	-5.820665E+06	9.836718E+06	5992.495307
1	211-30 90th Ct, Jamaica, NY 11428	40.720054	-73.750895	-5.820065E+06	9.836718E+06	5840.376700
2	89-28 213th St, Queens Village, NY 11427	40.723589	-73.750740	-5.819465E+06	9.836718E+06	5747.173218
3	214-46 Whitehall Terrace, Jamaica, NY 11427	40.727124	-73.750586	-5.818865E+06	9.836718E+06	5715.767665
4	218-17 Grand Central Pkwy, Jamaica, NY 11427	40.730659	-73.750431	-5.818265E+06	9.836718E+06	5747.173218
5	220-24 Hartland Ave, Jamaica, NY 11427	40.734194	-73.750276	-5.817665E+06	9.836718E+06	5840.376700
6	220-72 77th Ave, Flushing, NY 11364	40.737730	-73.750122	-5.817065E+06	9.836718E+06	5992.495307
7	204-12 100th Ave, Jamaica, NY 11423	40.711321	-73.755302	-5.821565E+06	9.837237E+06	5855.766389
8	93-52 Francis Lewis Blvd, Jamaica, NY 11428	40.714855	-73.755149	-5.820965E+06	9.837237E+06	5604.462508
9	90-21 208th St, Jamaica, NY 11428	40.718389	-73.754995	-5.820365E+06	9.837237E+06	5408.326913

## Foursquare:

### Use of Foursquare API to get info on restaurants in each neighbourhood

We are interested in venues in 'food' category, but only those that are proper restaurants - coffee shops, pizza places, bakeries etc. are not direct competitors so we don't care about those. So we have include in our list only venues that have 'restaurant' in category name, and we have made sure to detect and

include all the subcategories of specific 'Indian restaurant' category, as we need info on Indian restaurants in the neighbourhood.

**Food category:** - '4d4b7105d754a06374d81259'

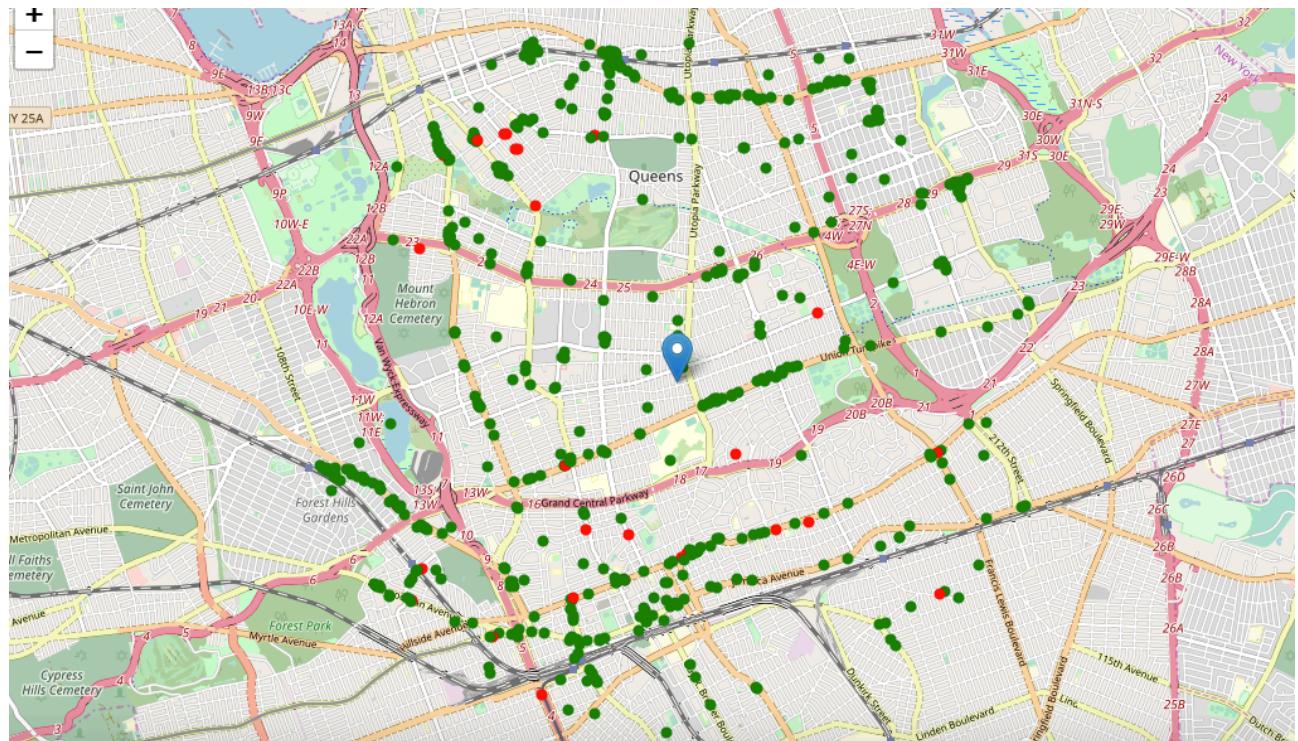
### Indian restaurant categories :-

```
[ '4bf58dd8d48988d10f941735', '54135bf5e4b08f3d2429dfe5', '54135bf5e4b08f3d2429dff3',  
  '54135bf5e4b08f3d2429dff5', '54135bf5e4b08f3d2429dfe2', '54135bf5e4b08f3d2429dff2',  
  '54135bf5e4b08f3d2429dfe1', '54135bf5e4b08f3d2429dfe3', '54135bf5e4b08f3d2429dfe8',  
  '54135bf5e4b08f3d2429dfe9', '54135bf5e4b08f3d2429dfe6', '54135bf5e4b08f3d2429dfdf',  
  '54135bf5e4b08f3d2429dfe4', '54135bf5e4b08f3d2429dfe7', '54135bf5e4b08f3d2429dfea',  
  '54135bf5e4b08f3d2429dfeb', '54135bf5e4b08f3d2429dfed', '54135bf5e4b08f3d2429dfee',  
  '54135bf5e4b08f3d2429dff4', '54135bf5e4b08f3d2429dfe0', '54135bf5e4b08f3d2429dfdd',  
  '54135bf5e4b08f3d2429dff6', '54135bf5e4b08f3d2429dfef', '54135bf5e4b08f3d2429dff0',  
  '54135bf5e4b08f3d2429dff1', '54135bf5e4b08f3d2429dfde', '54135bf5e4b08f3d2429dfec' ]
```

### Foursquare Exploration Result Summary:

Total number of restaurants	: 592
Total number of Indian restaurants	: 28
Percentage of Indian restaurants	: 4.73%
Average no. of restaurants in neighbourhood	: 2.782967032967033

Let's now see all the collected restaurants in our area of interest on map, and let's also show Indian restaurants in different colour



So now we have all the restaurants in area within few kilo meters from Queens,Ny, and we know which ones are Indian restaurants! We also know which restaurants exactly are in vicinity of every neighbourhood candidate center.

This concludes the data gathering phase - we're now ready to use this data for analysis to produce the report on optimal locations for a new Indian restaurant!

## **Methodology**

In this project we will put our efforts on exploring the areas of Queens,Ny that have

1. Low restaurant density,
2. Less number of Indian restaurants.
3. Our Analysis area Limit will be ~6km around city center

**Step 1:** In first step we have collected the required data: location and type (category) of every restaurant within 6km from New York center (Queens). We have also identified Indian restaurants (according to Foursquare categorisation).

**Step 2:** In second step our analysis will be calculation and exploration of 'restaurant density' across different areas of Queens - we will use heat-maps to identify a few promising areas close to center with low number of restaurants in general (and no Indian restaurants in vicinity) and focus our attention on those areas.

**Step 3:** 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 / neighbourhoods / addresses which should be a starting point for final 'street level' exploration and search for optimal venue location by stakeholders

## Analysis

Some basic explanatory data analysis is performed and derive some additional info from our raw data. First let's count the number of restaurants in every candidate area:

	Address	Latitude	Longitude	X	Y	Distance from center	Restaurants in area
0	93-46 210th Pl, Queens Village, NY 11428	40.716520	-73.751049	-5.820665E+06	9.836718E+06	5992.495307	2
1	211-30 90th Ct, Jamaica, NY 11428	40.720054	-73.750895	-5.820065E+06	9.836718E+06	5840.376700	1
2	89-28 213th St, Queens Village, NY 11427	40.723589	-73.750740	-5.819465E+06	9.836718E+06	5747.173218	1
3	214-46 Whitehall Terrace, Jamaica, NY 11427	40.727124	-73.750586	-5.818865E+06	9.836718E+06	5715.767665	0
4	218-17 Grand Central Pkwy, Jamaica, NY 11427	40.730659	-73.750431	-5.818265E+06	9.836718E+06	5747.173218	0
5	220-24 Hartland Ave, Jamaica, NY 11427	40.734194	-73.750276	-5.817665E+06	9.836718E+06	5840.376700	2
6	220-72 77th Ave, Flushing, NY 11364	40.737730	-73.750122	-5.817065E+06	9.836718E+06	5992.495307	2
7	204-12 100th Ave, Jamaica, NY 11423	40.711321	-73.755302	-5.821565E+06	9.837237E+06	5855.766389	1
8	93-52 Francis Lewis Blvd, Jamaica, NY 11428	40.714855	-73.755149	-5.820965E+06	9.837237E+06	5604.462508	1
9	90-21 208th St, Jamaica, NY 11428	40.718389	-73.754995	-5.820365E+06	9.837237E+06	5408.326913	1

Average number of restaurants in every area with radius=300m: 2.782967032967033

### Calculate Distance :

Calculate the distance to nearest Indian restaurant from every area candidate center (not only those within 300m - we want distance to closest one, regardless of how distant it is).

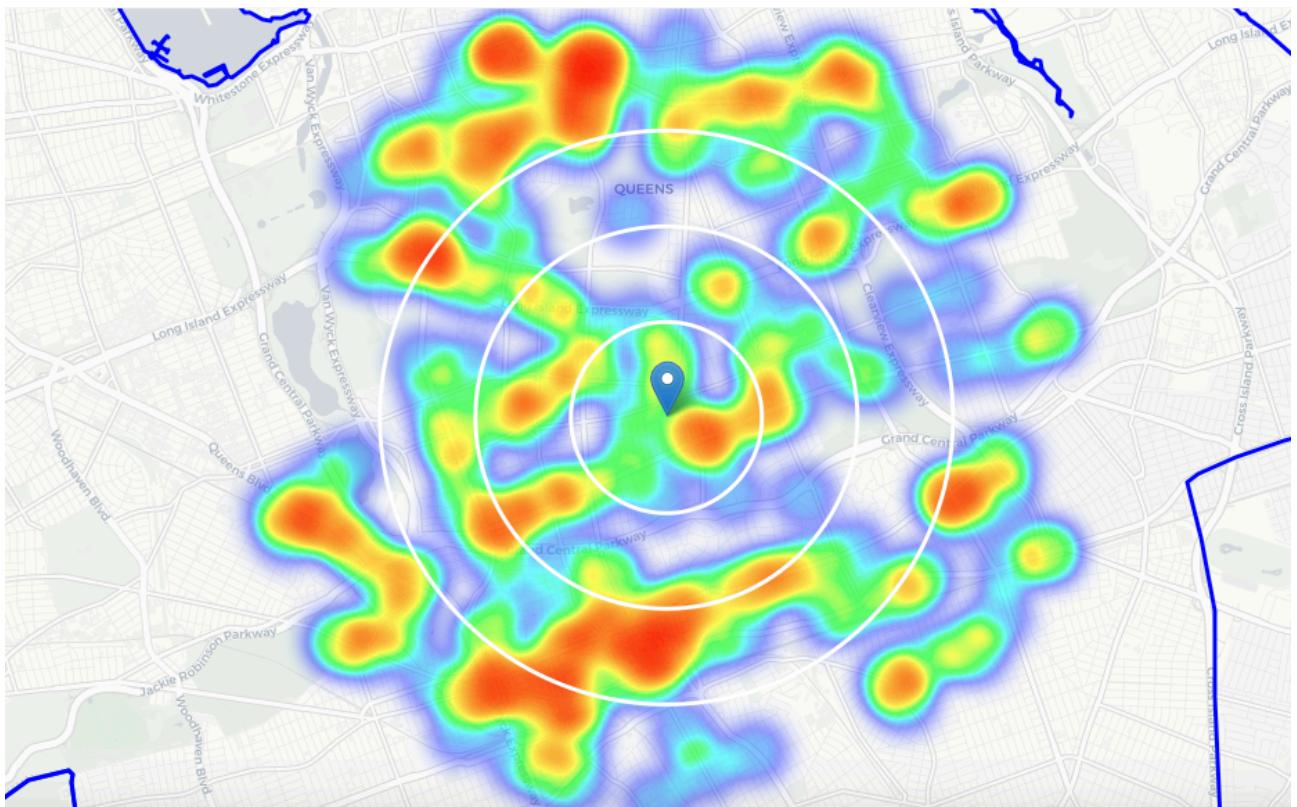
	Address	Latitude	Longitude	X	Y	Distance from center	Restaurants in area	Distance to Indian restaurant
0	93-46 210th Pl, Queens Village, NY 11428	40.716520	-73.751049	-5.820665E+06	9.836718E+06	5992.495307	2	1316.586918
1	211-30 90th Ct, Jamaica, NY 11428	40.720054	-73.750895	-5.820065E+06	9.836718E+06	5840.376700	1	1092.513353
2	89-28 213th St, Queens Village, NY 11427	40.723589	-73.750740	-5.819465E+06	9.836718E+06	5747.173218	1	1172.079238
3	214-46 Whitehall Terrace, Jamaica, NY 11427	40.727124	-73.750586	-5.818865E+06	9.836718E+06	5715.767665	0	1507.963545
4	218-17 Grand Central Pkwy, Jamaica, NY 11427	40.730659	-73.750431	-5.818265E+06	9.836718E+06	5747.173218	0	1973.357131

**Average between closest Indian restaurant & area centers:** 1337.73

Average Indian restaurant can be found within ~1500m from every area center candidate. That's fairly close

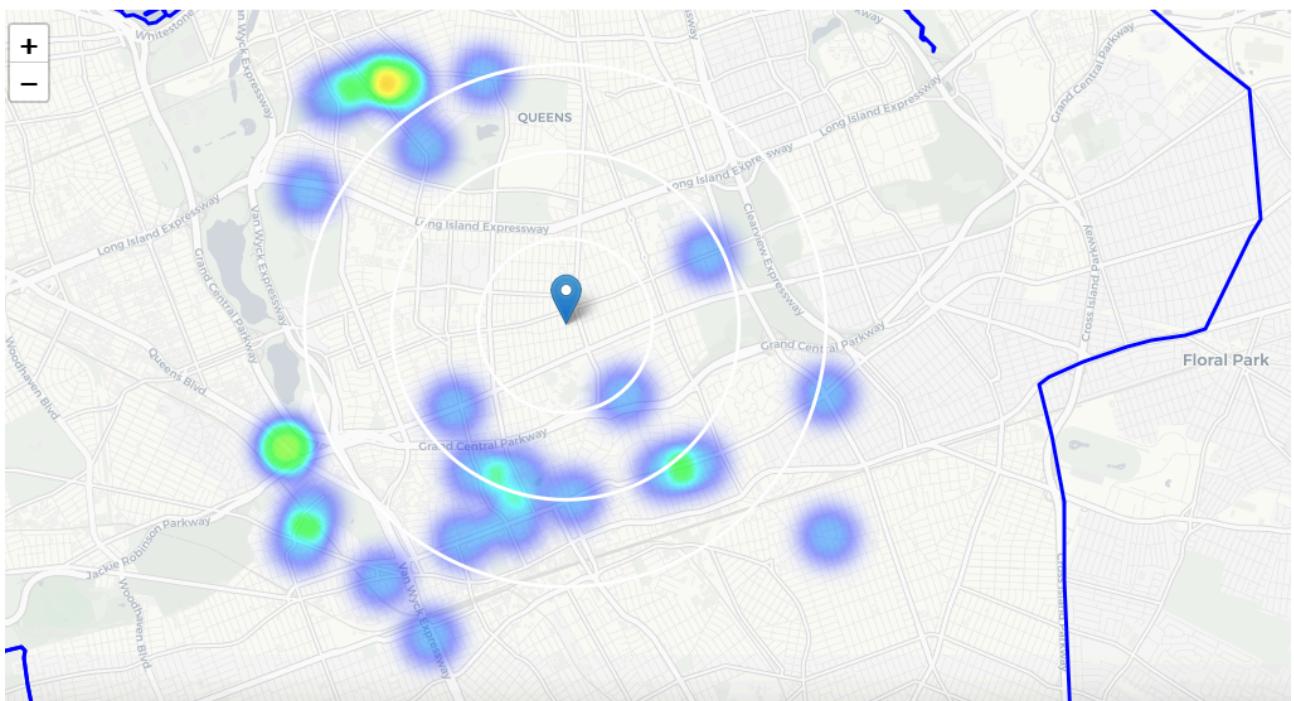
### Visualization of Restaurants:

Heat-map / density of restaurants with Borders of New York boroughs on our map and a few circles indicating distance of 1km, 2km and 3km from Queens.



A few pockets of low restaurant density closest to city center can be found in east, north-east and east from Queens.

## Visualization of Indian Restaurants: Heat-map showing the density of Indian restaurants only.



above map is not so 'hot' (Indian restaurants represent a subset of ~4.73% of all restaurants in Queens) but it also indicates higher density of existing Indian restaurants directly north and west from Queens, with closest pockets of low Indian restaurant density positioned east, south-east and south from city center.

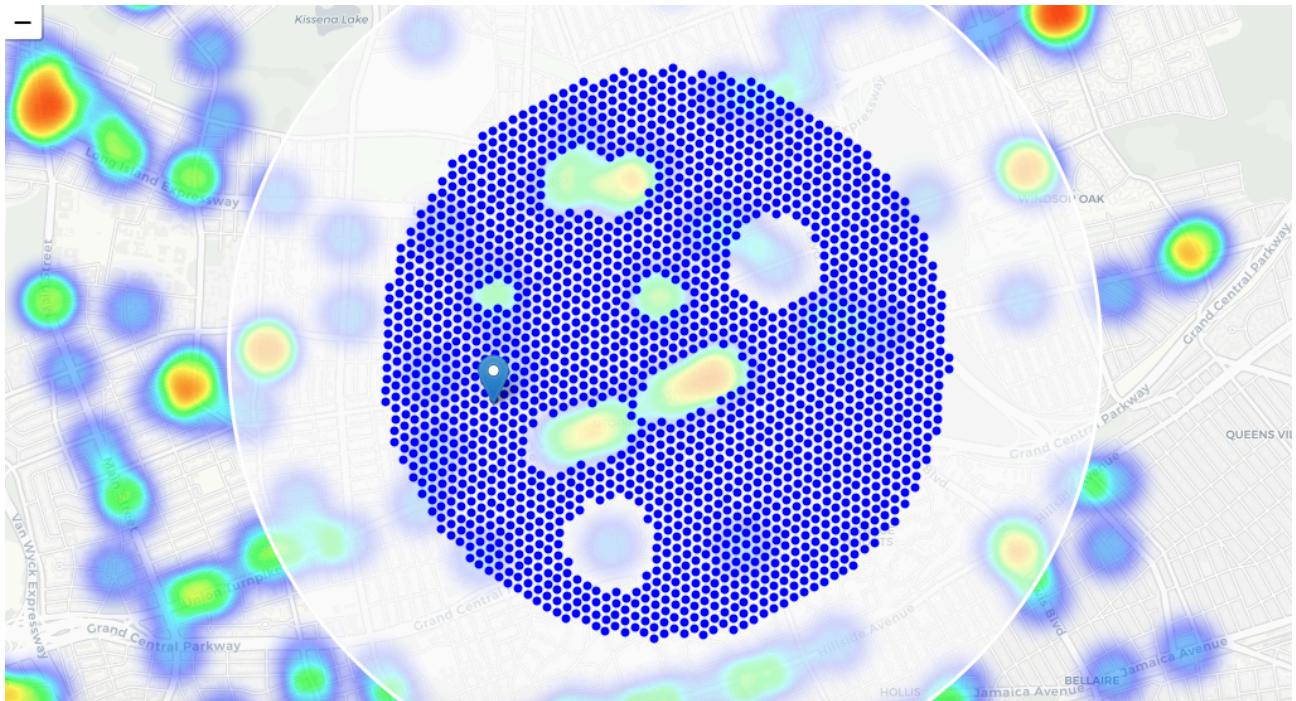
Based on this we will now focus our analysis on areas south-east, north-east and east from Queens center - we will move the center of our area of interest and reduce it's size to have a radius of 2.5km. This places our location candidates mostly in boroughs Jamaica and Auburndale.

Another potentially interesting borough is Jamaica with large low restaurant density north-east from city center, however this borough is less interesting to stakeholders as it's mostly residential and less popular with tourists.

## Justification for Further Analysis:

Selection of Jamaica and Auburndale for further analysis is based on three main aspects -

1. Popular with tourists,
2. Alternative and bohemian but booming and trendy,
3. Relatively close to city center for well connected;



## Candidate neighbourhood centers generated = 2261

Now let's calculate two most important things for each location candidate:

1. Number of restaurants in vicinity (Radius of 250 meters)
2. Distance to closest Indian restaurant

	Latitude	Longitude	X	Y	Restaurants nearby	Distance to Indian restaurant
0	40.730110	-73.763758	-5.818415E+06	9.838433E+06	0	1628.460512
1	40.730699	-73.763733	-5.818315E+06	9.838433E+06	0	1721.021313
2	40.726886	-73.764569	-5.818965E+06	9.838520E+06	0	1192.247396
3	40.727475	-73.764544	-5.818865E+06	9.838520E+06	0	1273.375268
4	40.728064	-73.764518	-5.818765E+06	9.838520E+06	0	1357.024425
5	40.728653	-73.764493	-5.818665E+06	9.838520E+06	0	1442.756393
6	40.729243	-73.764467	-5.818565E+06	9.838520E+06	0	1530.221137
7	40.729832	-73.764442	-5.818465E+06	9.838520E+06	0	1619.137872
8	40.730421	-73.764416	-5.818365E+06	9.838520E+06	0	1679.969561
9	40.731010	-73.764391	-5.818265E+06	9.838520E+06	0	1633.603268

### **Filter those locations:**

we are interested only in locations with

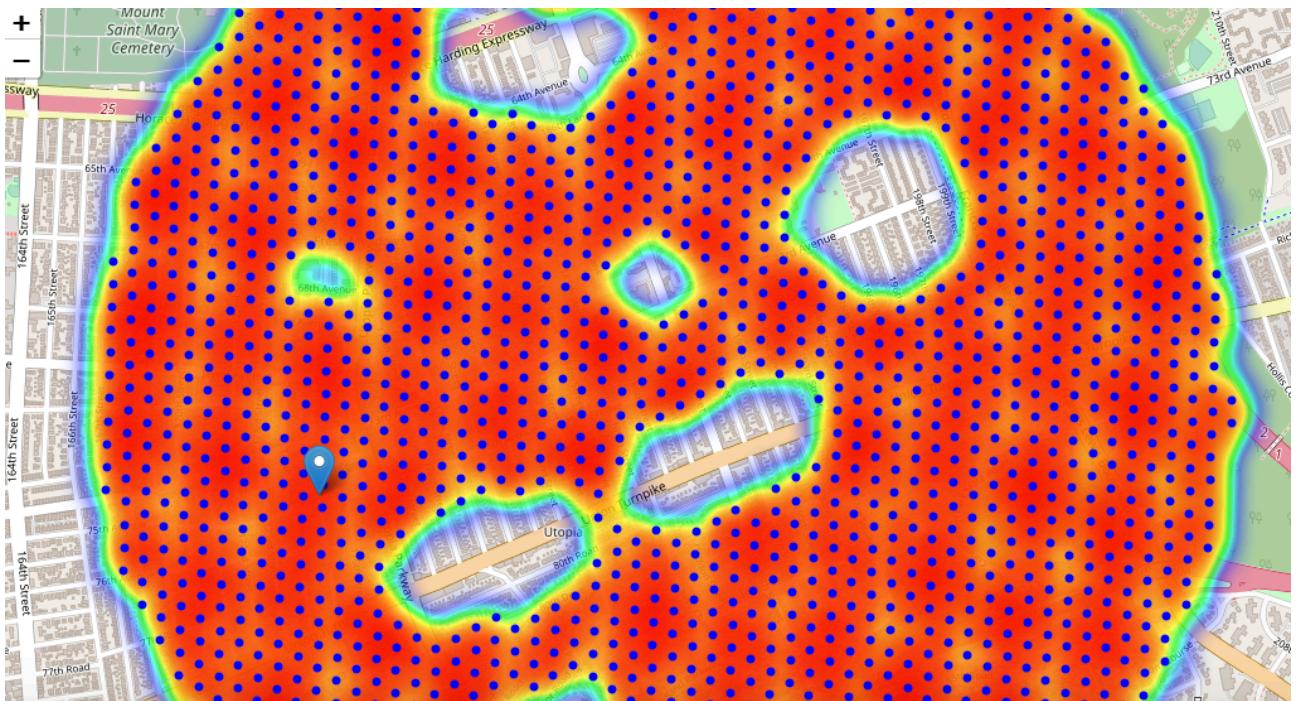
1. No more than two restaurants in radius of 250 meters,
2. No Indian restaurants in radius of 400 meters.

### **Results of Filtration:**

1. Locations with no more than two restaurants nearby : 2095
2. Locations with no Indian restaurants within 400m : 2146
3. Locations with both conditions met : 1980

We now have a bunch of locations fairly close to Queens, and we know that each of those locations has no more than two restaurants in radius of 250m, and no Indian restaurant closer than 400m. Any of those locations is a potential candidate for a new Indian restaurant, at least based on nearby competition.

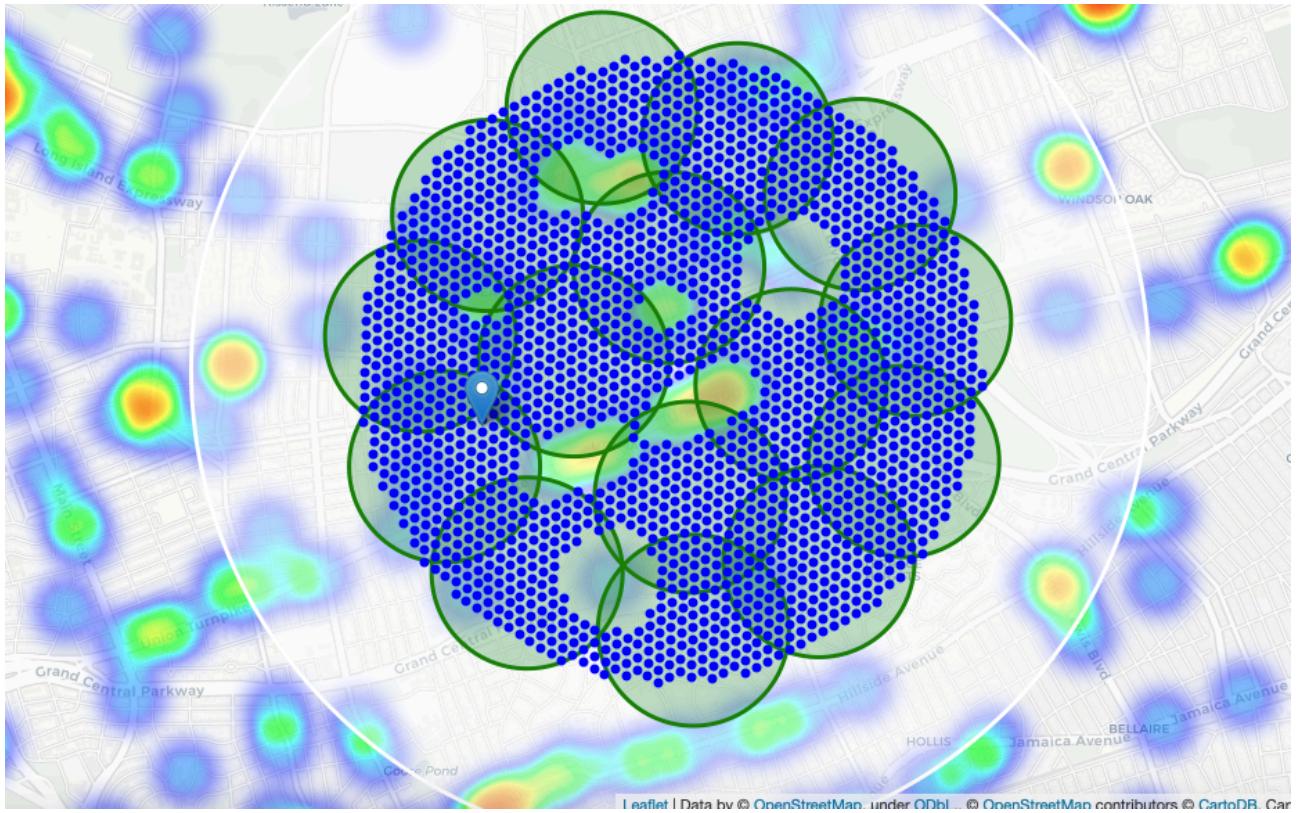
### **Visualization of good locations in a form of heat-map:**



What we have now is a clear indication of zones with low number of restaurants in vicinity, and no Indian restaurants at all nearby.

## Clustering:

Cluster these locations to create centers of zones containing good locations. Those zones, their centers and addresses will be the final result of our analysis.



Our clusters represent groupings of most of the candidate locations and cluster centers are placed nicely in the middle of the zones 'rich' with location candidates.

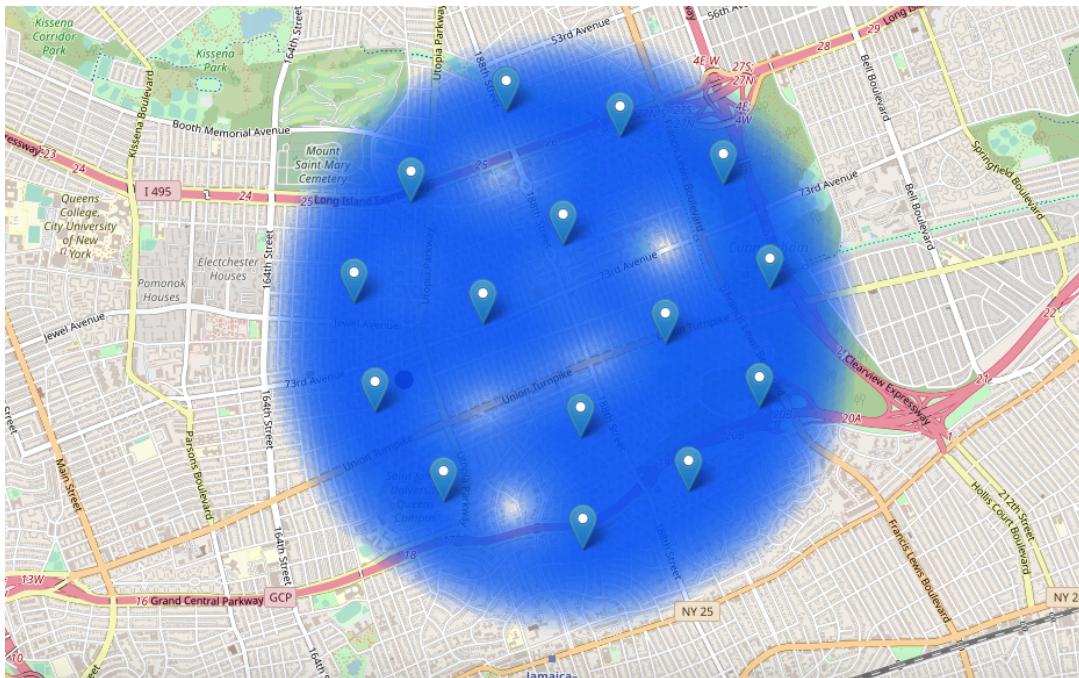
Addresses of those cluster centers will be a good starting point for exploring the neighbourhoods to find the best possible location based on neighbourhood specifics.

**Finally:** Use reverse geocode for those candidate area centers to get the addresses which can be presented to stakeholders.

=====  
Addresses of centers of areas recommended for further analysis  
=====

56-37 188th St, Fresh Meadows, NY 11365	=> 2.7km from Queens
198-15 Grand Central Pkwy, Hollis, NY 11423	=> 3.4km from Queens
184-80 Aberdeen Rd, Jamaica, NY 11432	=> 1.8km from Queens
1955 Peck Ave, Fresh Meadows, NY 11365	=> 3.1km from Queens
69-83 181st St, Flushing, NY 11365	=> 0.9km from Queens
Clearview Expy, Oakland Gardens, NY 11364	=> 3.6km from Queens
85-80 Palermo St, Jamaica, NY 11423	=> 2.9km from Queens
84-53 Avon St, Jamaica, NY 11432	=> 2.4km from Queens
76-09 172nd St, Flushing, NY 11366	=> 0.4km from Queens
188-35 69th Ave, Fresh Meadows, NY 11365	=> 2.0km from Queens
75-68 195th St, Flushing, NY 11366	=> 2.5km from Queens
170-20 69th Ave, Flushing, NY 11365	=> 0.9km from Queens
Belson Hall, Jamaica, NY 11439	=> 1.2km from Queens
196-10 Union Tpke, Fresh Meadows, NY 11366	=> 3.6km from Queens
174-02 Horace Harding Expy, Flushing, NY 11365	=> 1.7km from Queens

This concludes our analysis. We have created 15 addresses representing centers of zones containing locations with low number of restaurants and no Indian restaurants nearby, all zones being fairly close to city center (all less than 4km from Queens, and about half of those less than 2km from Queens). Although zones are shown on map with a radius of ~500 meters (green circles), their shape is actually very irregular and their centers/addresses should be considered only as a starting point for exploring area neighbourhoods in search for potential restaurant locations. Most of the zones are located in Jamaica and Fresh Meadows , which we have identified as interesting due to being popular with tourists, fairly close to city center and well connected by public transport.



## **Results and Discussion:**

Through our analysis we get a great number of restaurants in Queens (~592 in our initial area of interest which was 12x12km around Queens), there are also pockets of low restaurant density fairly close to city center. Highest concentration of restaurants was detected north and west from Queens, so we focused our attention to areas south, south-east and east, corresponding to Jamaica, Meadows and south-east corner.

We found another interesting borough Jamaica which is south-east from Queens and it is largely a residential area. So our attention was focused on Jamaica and Meadows which offer a combination of -

1. Popularity among tourists,
2. Closeness to city center,
3. Strong socio-economic dynamics
4. Number of pockets of low restaurant density.

After directing our attention to this more narrow area of interest (covering approx. 5x5km southeast from Queens) we first created a dense grid of location candidates (spaced 100m apart); those locations were then filtered so that those with more than two restaurants in radius of 250m and those with an Indian restaurant closer than 400m were removed.

Those location candidates were then clustered to create zones of interest which contain greatest number of location candidates. Addresses of centers of those zones were also generated using reverse geocoding to be used as markers/starting points for more detailed local analysis based on other factors.

Result of all this is 15 zones containing largest number of potential new restaurant locations based on number of and distance to existing venues - both restaurants in general and Indian restaurants particularly. This, of course, does not imply that those zones are actually optimal locations for a new restaurant! Purpose of this analysis was to only provide info on areas close to Queens center but not crowded with existing restaurants (particularly Indian) - it is entirely possible that there is a very good reason for small number of restaurants in any of those areas, reasons which would make them unsuitable for a new restaurant regardless of lack of competition in the area. Recommended zones should therefore be considered only as a starting point for more detailed analysis which could eventually result in location which has not only no nearby competition but also other factors taken into account and all other relevant conditions met

# **Conclusion:**

Objective of this project was to assist stakeholders in identifying areas close to center of Queens,Ny,USA with low number of restaurants (particularly Indian restaurants) for opening a new Indian restaurant.

Using Foursquare data we calculated density of restaurants to identify boroughs that justify the further analysis of two major tourist locations (jamaica and Meadows, Auburndale), and then generated extensive collection of locations which satisfy some basic requirements regarding existing nearby restaurants.

Clustering of those locations was then performed in order to create major zones of interest (containing greatest number of potential locations) and addresses of those zone centers were created to be used as starting points for final exploration by stakeholders.

Final decision on location of restaurant will be made by stakeholders based on specific characteristics of neighbourhoods and locations in every recommended zone.

Additional factors: Stakeholder can consider additional factors like

1. Attractiveness of each location (proximity to park or water)
2. Levels of noise / proximity to major roads
3. Real estate availability
4. Real estate Prices
5. Social and economic dynamics of every neighbourhood etc.