Capstone Project: The Battle of Neighbourhoods
Indian Restaurants in Berlin, Germany

Author: Kaushik Chakrabarti

1. Introduction

(1A) Background

The city of Berlin in Germany is one of the important city in Germany. Being the capital city of Germany it attracts a lot of people both for official as well as tourism visits. The population of Berlin has been estimated to be around 3.5 million. Berlin also host tourists from all over the world visiting the city for tourism. Berlin also attracts many foreign people from all over the world for both professional purpose as well as for studies. Thus along with local people many foreign nationals used to stay in Berlin for their work or study.

As Berlin attracts many people from all over the world specially the Indians so Indian food is an important aspects for the people from India. Apart from Indians many foreign nationals also like to have Indian food. As Berlin is a place of many foreign nationals thus restaurants serving Indian Food will be an interesting aspect of the study.

(1B) Business Problem

As discussed in the Background section that Berlin host many people from within Germany as well as from different parts of the world for both professional purpose as well as Tourism thus restaurants in Berlin is an important aspect to study. Out of various cuisines the Indian cuisine seemed to an interesting aspect of study as apart from Indian nationals many Foreign nationals also like to have Indian cuisine.

In this project we will try to address the following problem,

- Optimal location for opening an Indian restaurant in the city of Berlin

- Berlin being the capital city of Germany there are many restaurants all over Berlin thus we will try to identify locations in the Berlin city that are not crowded with the restaurants
- We are also interested in the locations which do not have any Indian restaurants in the vicinity
- If the above mentioned criteria gets satisfied we will also prefer the location should be as close to the city centre as opening a restaurant far away from city centre will not attract any customers and this is not good for the business

(1C) Target Audience

The main target audience for the project is anyone or stakeholder who wanted to open an Indian restaurant in the city of Berlin. This project will help the future stakeholder to analyse the present conditions that are,

- Locations of present restaurants
- Total number of restaurants
- Locations of Indian restaurants
- Total number of Indian restaurants

Based on this data and the analysed data from which he can understand,

- Location for opening an Indian restaurant
- Distance of that location from the city centre

These data will help the stakeholder to invest their money in opening an Indian restaurant in the city of Berlin.

2. Details about the Data

(2.A) Data Description

Based on our problem definition the main factors that will affect our decision are

- Number of existing restaurants in the neighbourhood
- Number of Indian restaurants in the neighbourhood
- Distance to the Indian restaurants
- Distance of the neighbourhood from the city centre

In order to address our problem we have decided to use regularly spaced grid of locations around the central location of the city of Berlin to define our neighbourhoods. For this we will be performing the following steps,

- We will identify a well known location at the central part of the city of Berlin (‘Alexanderplatz’)
- Then using the “Geolocator” function we will find the co-ordinates or Latitude and Longitude of the central part of Berlin city
- After that we will be creating centres of interesting areas for our problem using user defined algorithm
- With the help of the algorithm we will be having a set of co-ordinates for all the centres of interest
- Then we will be using “Reverse Geolocator” function in order to obtain the address for the interesting centres
- After that we will be using Foursquare API (https://developer.foursquare.com/docs/resources/categories) in order to identify the Indian restaurants having category ID of “4bf58dd8d48988d10f941735”

- After that we will be using all the categories of Indian restaurants from Foursquare API and find out the locations of the Indian restaurants
- We will be using heat map to identify the suitable locations for opening an Indian restaurants and also we will keep in mind the distance between the location and the city centre defined earlier
- Based upon our heat map we will identify few locations for the stakeholder which have low density of Indian restaurants and also they are close to the central part of Berlin
- Then based upon the Borough we will cluster our result on the basis of locations for the stakeholder to choose and invest their funds to open a business

(2B) <u>Details about the Neighbourhood Data Procurement</u>

As discussed in the problem description as well as in the Data Description we started our project by identifying a rental location in the city of Berlin named **"Alexanderplatz"**. We then used the Geolocator package in order to obtain the exact co-ordinate (Latitude and Longitude) of our selected location. After that we will be using the obtained geographical co-ordinates to obtain the 2D cartesian x and y values using pyproj package. The 2D Cartesian co-ordinate will help us to obtain the distance between the place of interest and central location of Berlin which normally is impossible using the Geographical co-ordinates. We created hexagonal grid of cells having offset in every row and we have also adjusted the vertical row spacing so that every cell centre is equidistant from all its neighbours. The idea of this is to have a grid area of candidates who are equally spaced, centred around our Berlin city centre and they are within around 6 Km from "Alexanderplatz". The neighbourhoods obtained by us will be defined as circular areas having a radius of 300 m thus mathematically our interesting neighbourhoods centres are 600 m apart from each other.

After we create the hexagonal grid of cells representing our interested neighbourhood around the initial Berlin city centre within its 6 Km range we obtained the geographical co-ordinates of all the central position of the interested neighbourhood locations along with total number of neighbourhood locations and their respective 2D cartesian co-ordinate for distance calculation.

We then used the geographical co-ordinates of the neighbourhoods centroid in folium map in order to show the extent of our neighbourhood in an around “Alexanderplatz”. Figure: 1 below shows the folium map of Berlin centred around “Alexanderplatz” with all the interesting neighbourhoods marked in circular rings.

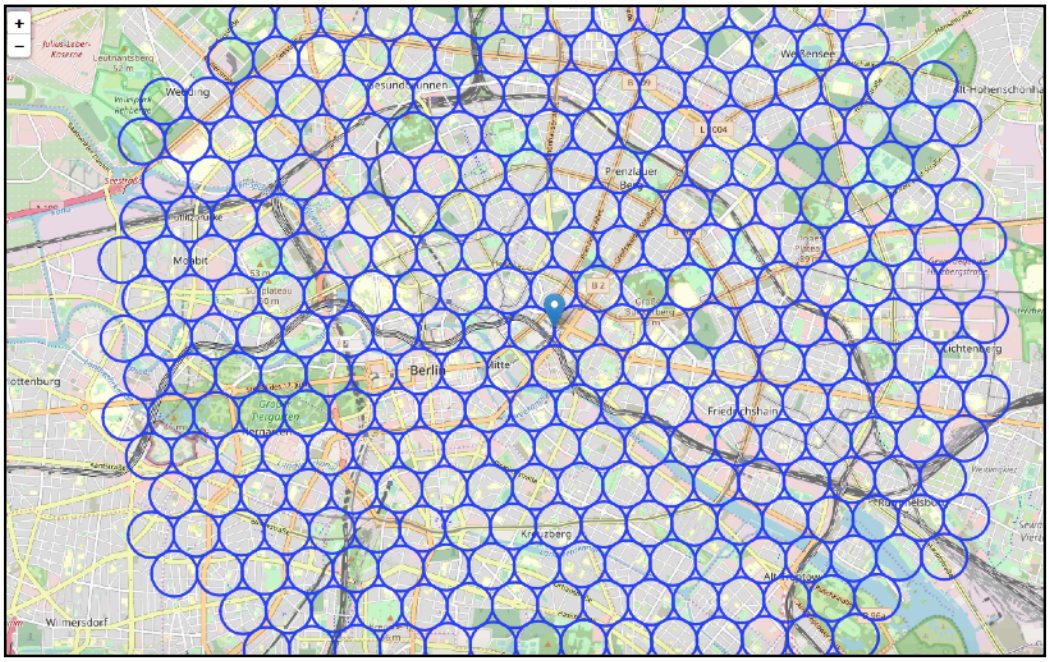

Figure: 1

After we receive the geographical co-ordinates for the centroids of interesting neighbourhoods we will use reverse Geocode function in order to get the address corresponding to the centroid’s geographical co-ordinates. After that we will be creating a Data frame showing the Geographical address for the centroids along with their respective

geographical co-ordinates and 2D cartesian co-ordinates. Figure: 2 below shows the first 10 entries of the data frame as discussed above.

	Address	Latitude	Longitude	X	Y	Distance from center
0	(ehemalige südliche Start- und Landebahn, Gart...	52.470217	13.388849	390559.94364	5.814560e+06	5992.495307
1	(ehemalige südliche Start- und Landebahn, Temp...	52.470337	13.397678	391159.94364	5.814560e+06	5840.376700
2	(ehemalige südliche Start- und Landebahn, Temp...	52.470456	13.406508	391759.94364	5.814560e+06	5747.173218
3	(Hundeauslauf, ehemalige südliche Start- und L...	52.470574	13.415337	392359.94364	5.814560e+06	5715.767665
4	(22, Warthestraße, Schillerkiez, Neukölln, Ber...	52.470692	13.424167	392959.94364	5.814560e+06	5747.173218
5	(14, Ilsenhof, Rollbergsiedlung, Neukölln, Ber...	52.470810	13.432996	393559.94364	5.814560e+06	5840.376700
6	(Mode Ateș, 215, Karl-Marx-Straße, Richardkiez...	52.470926	13.441826	394159.94364	5.814560e+06	5992.495307
7	(36, Hessenring, Gartenstadt Neu-Tempelhof, Te...	52.474705	13.375433	389659.94364	5.815079e+06	5855.766389
8	(127, Kleineweg, Gartenstadt Neu-Tempelhof, Te...	52.474826	13.384263	390259.94364	5.815079e+06	5604.462508
9	(ehemalige nördliche Start- und Landebahn, Tem...	52.474946	13.393094	390859.94364	5.815079e+06	5408.326913

Figure: 2

(2C) Details about Data procurement from Foursquare API client

Once we have the our areas or neighbourhoods of interest we will then try to access the number of total restaurants as well as number of Indian restaurants in an around the Berlin city centre covering our areas or neighbourhoods of interest. We will be using Foursquare API client for the main food category along with Indian restaurants category and sub-categories for all various Indian restaurants in our Foursquare API client call function. We will limit our search to 100 hits with a radius of 500. We will find all the restaurants in an around our location of interest and also we will find all the Indian restaurants in an around our locations of interest as this is an important aspect of our project.

We will then project the total number of restaurants, total number of Indian restaurants and average number of restaurants in the neighbourhoods along with their address and 2D cartesian co-ordinates for the distance calculation. This data will be helpful to us in order to further analyse our data. After getting the full details about the total number of restaurants and total number of Indian restaurants in an around the central location of Berlin we will

project the locations in the form of folium map of Berlin so that visually it will be easier to distinguish Indian restaurants from other restaurants. Figure: 3 shows the map of Berlin pointing all the Indian and other restaurants in an around the central location of Berlin "Alexanderplatz".

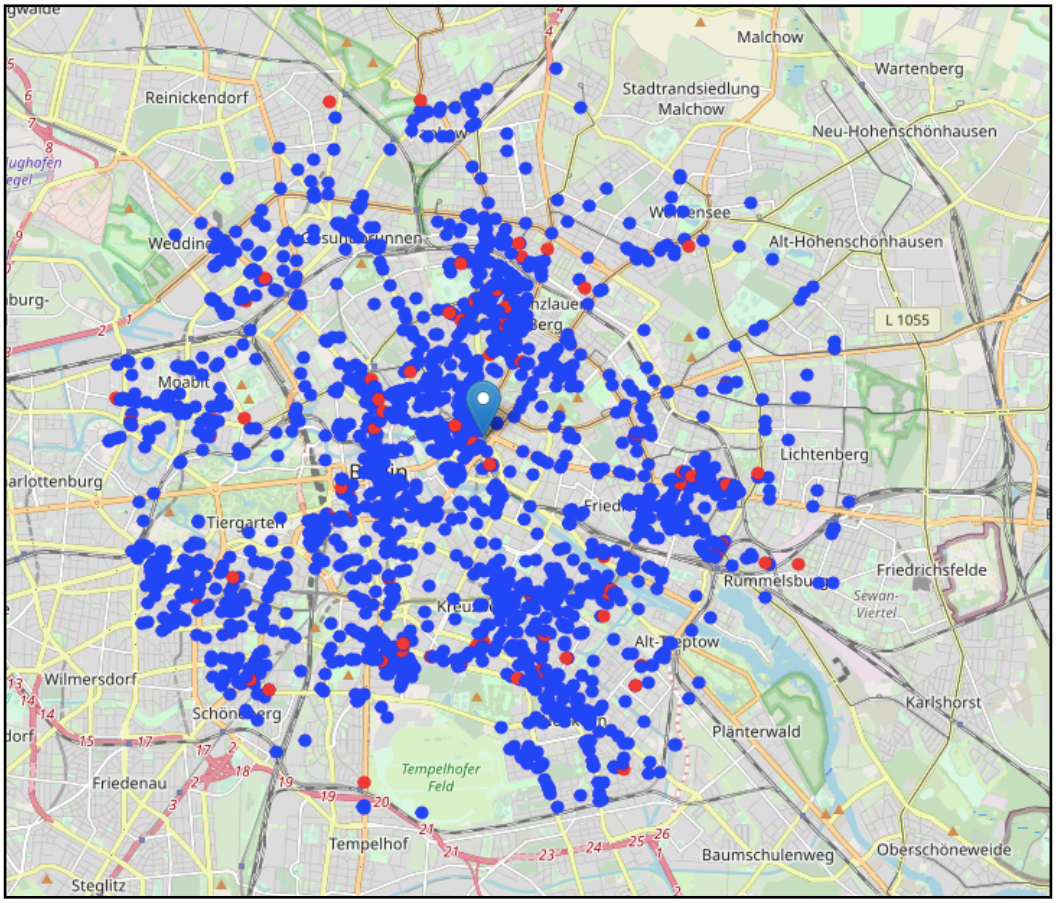

Figure: 3

This concludes our data gathering process and with the gathered data we will start to perform our analysis which well lead to identify optimal or interesting locations for opening an Indian restaurant.

3. Methodology

As we have discussed at the beginning of the project that we will be directing our efforts in order to detect areas of Berlin having low density of restaurants especially those areas where the number of Indian restaurants are low as the main aim of our project is to direct the

stakeholders helping them to identify a suitable location for opening an Indian restaurants. In this project we will limit our analysis area to around 6 Km from the previously identified central part of Berlin “Alexanderplatz”.

In the first part of our project we have already collected the required data on the interested locations or neighbourhoods along with their address within 6 Km of the Berlin city centre “Alexanderplatz”. We have also identified the total number of restaurants and total number of Indian restaurants in an around the central part of Berlin city.

In the second part of our project we will direct our energy in order to analyse the data collected by us. We will use the data collected by by in order to identify few promising areas close to the Berlin city centre with low number of restaurants in general and no Indian restaurants. We will be using heat maps to identify the area of interest and then focus our main interest on those identified areas.

In the final part of the project we will mainly focus on previously obtained areas or neighbourhoods of interest and within those ares we will create cluster of locations that will meet some of the basic requirements for the project. Established discussion with the stakeholders we will take into considerations the locations

- Having no more than two restaurants in radius of 250 m
- Locations having no Indian restaurants within radial distance of 400 m

We will create a map of all such locations. We will also create clusters using k-means clustering of the locations in order to identify general neighbourhoods or addresses which we could suggest our potential stakeholder as a venue location for opening an Indian restaurant.

4. Analysis

Before we proceed into our Analysis part we wanted to extract some additional information from our raw data. As we have discussed in earlier module that we have divided our whole neighbourhoods into circles of radius 300 m and also we have got some information about the number of Indian restaurants and total number of restaurants in an around the central location of Berlin. Now we wanted to have an information about the number of restaurants in each neighbourhood area or in each area of interest. This means we wanted to find out the total number of restaurants within a radius of 300 m. Figure: 4 shows the data frame for total number of restaurants in each interesting neighbourhood regions.

```
Average number of restaurants in every area with radius=300m: 5.010989010989011
[21]:
```

	Address	Latitude	Longitude	X	Y	Distance from center	Restaurants in area
0	(ehemalige südliche Start- und Landebahn, Gart...	52.470217	13.388849	390559.94364	5.814560e+06	5992.495307	2
1	(ehemalige südliche Start- und Landebahn, Temp...	52.470337	13.397678	391159.94364	5.814560e+06	5840.376700	0
2	(ehemalige südliche Start- und Landebahn, Temp...	52.470456	13.406508	391759.94364	5.814560e+06	5747.173218	0
3	(Hundeauslauf, ehemalige südliche Start- und L...	52.470574	13.415337	392359.94364	5.814560e+06	5715.767665	0
4	(22, Warthestraße, Schillerkiez, Neukölln, Ber...	52.470692	13.424167	392959.94364	5.814560e+06	5747.173218	0
5	(14, Ilsenhof, Rollbergsiedlung, Neukölln, Ber...	52.470810	13.432996	393559.94364	5.814560e+06	5840.376700	7
6	(Mode Ateş, 215, Karl-Marx-Straße, Richardkiez...	52.470926	13.441826	394159.94364	5.814560e+06	5992.495307	6
7	(36, Hessenring, Gartenstadt Neu-Tempelhof, Te...	52.474705	13.375433	389659.94364	5.815079e+06	5855.766389	0
8	(127, Kleineweg, Gartenstadt Neu-Tempelhof, Te...	52.474826	13.384263	390259.94364	5.815079e+06	5604.462508	0
9	(ehemalige nördliche Start- und Landebahn, Tem...	52.474946	13.393094	390859.94364	5.815079e+06	5408.326913	0

Figure: 4

After that we need to find the average distance between the centroid of interesting area and the closest Indian restaurant. This will help us to carefully select our region of interest. In order to get more insight to the data we have to create a heat map for the Berlin with centering “Alexanderplatz” along with the Borough names of Berlin on the heat map. The borough names were downloaded or obtained from https://raw.githubusercontent.com/m-

hoerz/berlin-shapes/master/berliner-bezirke.geojson. We will also show circles of 1 Km, 2 Km and 3 Km from Berlin city centre on the heat map. The heat map is depicted in Figure: 5.

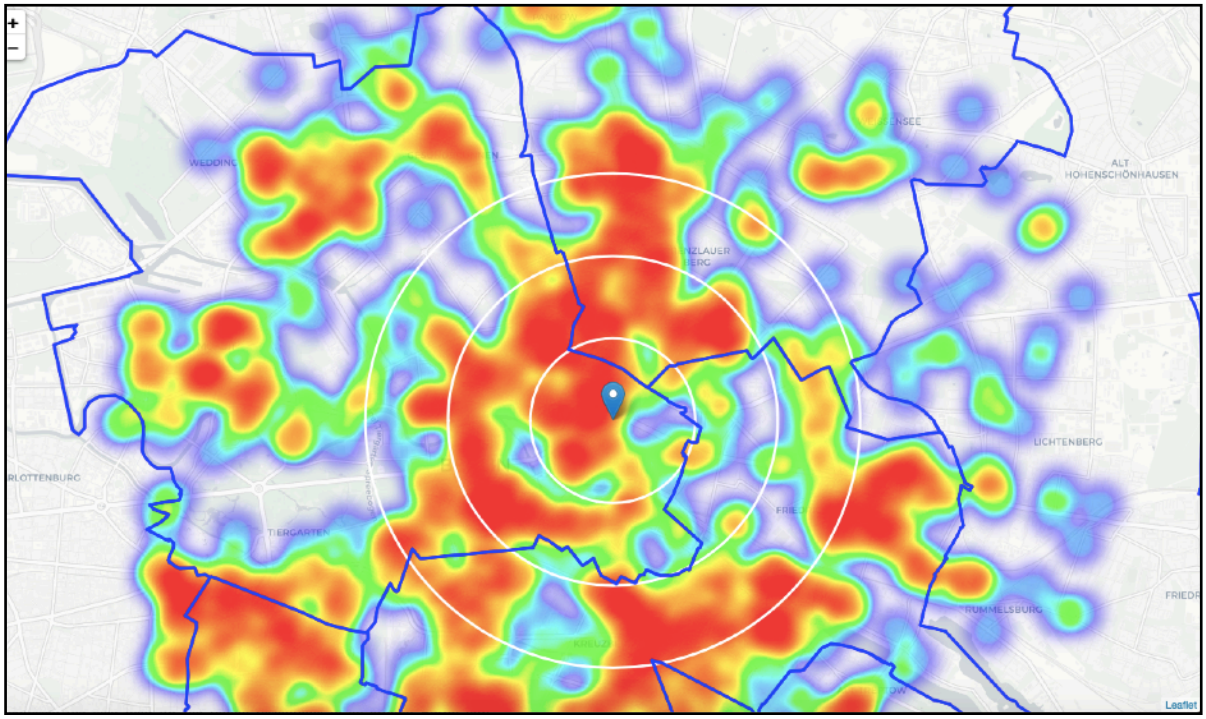

Figure: 5

From the heat map depicted in Figure: 5 we observed that we have a relatively low density of restaurant in the south, south-east and east from the central location of Berlin, "Alexanderplatz". We also wanted to have a proper idea about the Indian restaurants thus in Figure: 6 we represent the heat map for the Indian restaurants in an around Berlin city centre. From this heat map we observed that the map is not that much hot as Indian restaurants only contributes to around 4% of total number of restaurants in and around Berlin's central location.

The heat map obtained in Figure: 6 shows that we have relatively higher density of Indian restaurants in the northern and western part of "Alexanderplatz". We also observed a relatively low density of Indian restaurants in the east, south-east and south of

“Alexanderplatz”. Thus from the above analysis we could direct our primary focus towards the south-east, south, south-west and east of “Alexanderplatz”.

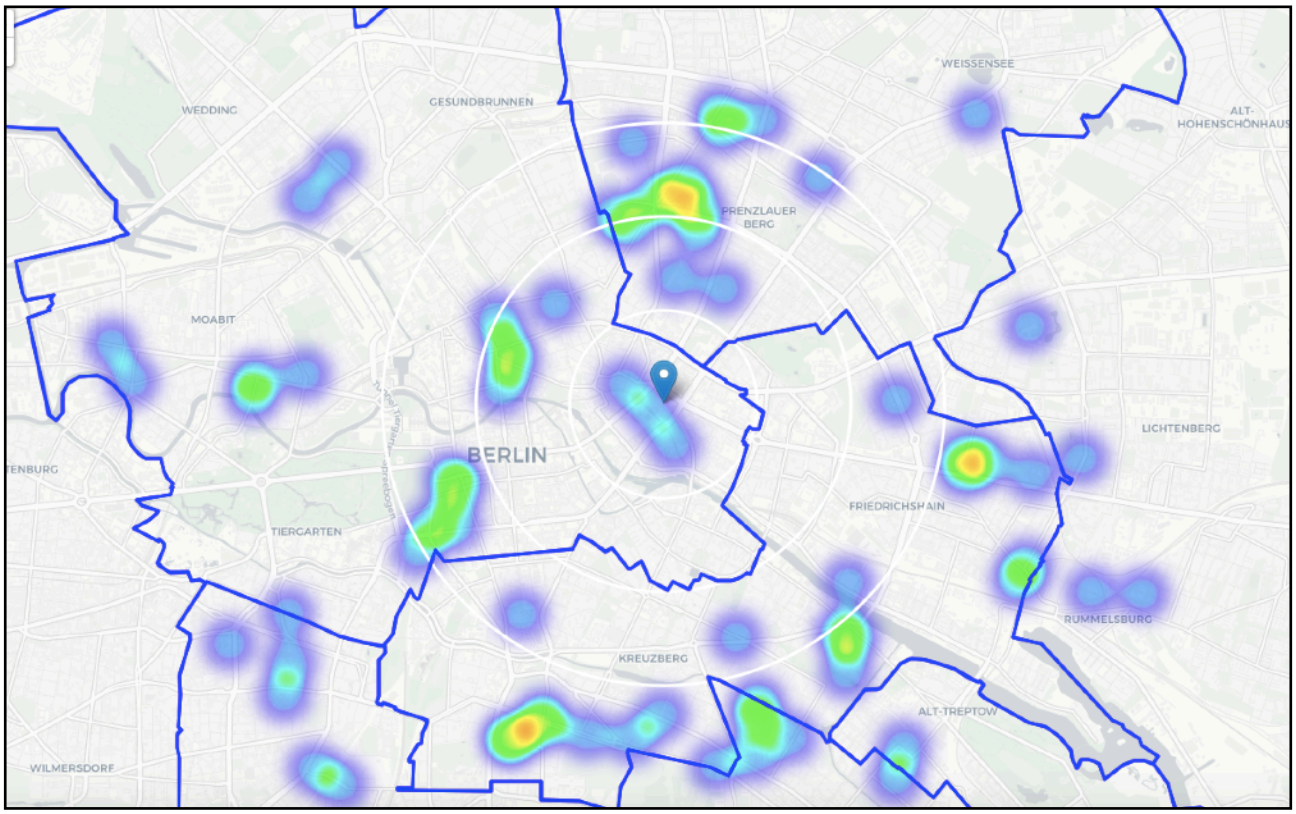

Figure: 6

Bases on our earlier analysis as we are interested in south-east, south-west, south and east of “Alexanderplatz” we will be moving the centre of the area of our interest and reduce the radius size to 2.5 Km instead of 6 Km assumed earlier in our project. When we shift or reduce the radius size to 2.5 Km from “Alexanderplatz” the primary locations or area of interest mostly fall under the boroughs of “Kreuzberg” and “Friedrichshain”. We can mainly focus our interest in these two boroughs as according to wikipedia and other sources these places are important for tourist attraction and opening a restaurant in these location will attract more people than anywhere else considering the business point of view.

As mentioned earlier from different sources like travel guides, wikipedia and other travel websites these two places “Kreuzberg” and “Friedrichshain” attract more tourist owing to its

beauty, and culture which is quite rich and interesting. The neighbourhoods of Berlin are popular not only for tourists but also for people staying in Berlin. Thus due to their popularity among the tourists and residents along with its location being close to the city centre having well connected transport facilities justify our selection of these boroughs for further analysis in our project.

Let us now re-define a new and more narrow region of interest having low restaurant count in parts of “Kreuzberg” and “Friedrichshain” closest to Berlin city centre “Alexanderplatz”. Figure:7 shows our new area of interest on the map of Berlin depicting all the pockets of low restaurant density in “Kreuzberg” and “Friedrichshain” closest to Berlin centre.

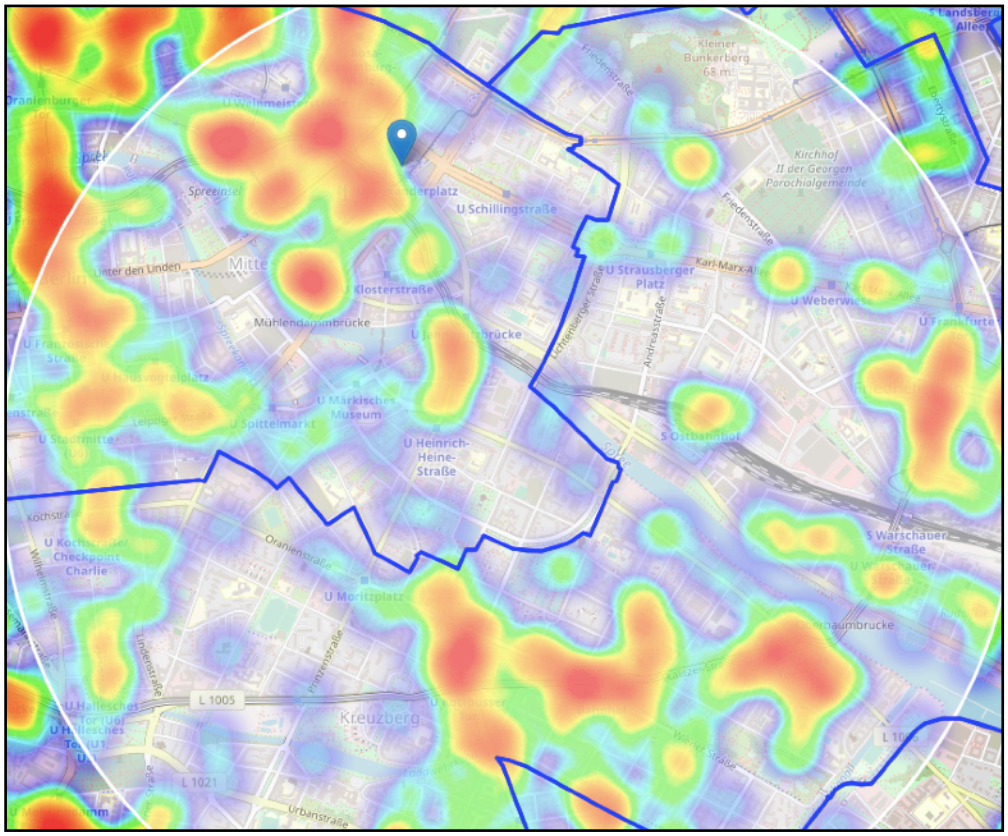

Figure: 7

Let us also create a new more dense grid of location candidates restricted to our new region of interest. In this we will be making our location candidates 100 m apart and this created new geographical co-ordinates for our renewed area of interest. After that we will

concentrate on finding the number of restaurants in the vicinity of our new are of interest along with distance to the closest Indian restaurant. After that we will try to find

- How many locations has no more than two restaurants in the locality
- How many locations has no Indian restaurants within 400 m of radius
- How many locations satisfy both the above criteria

Figure: 8 shows the details in the form of a map.

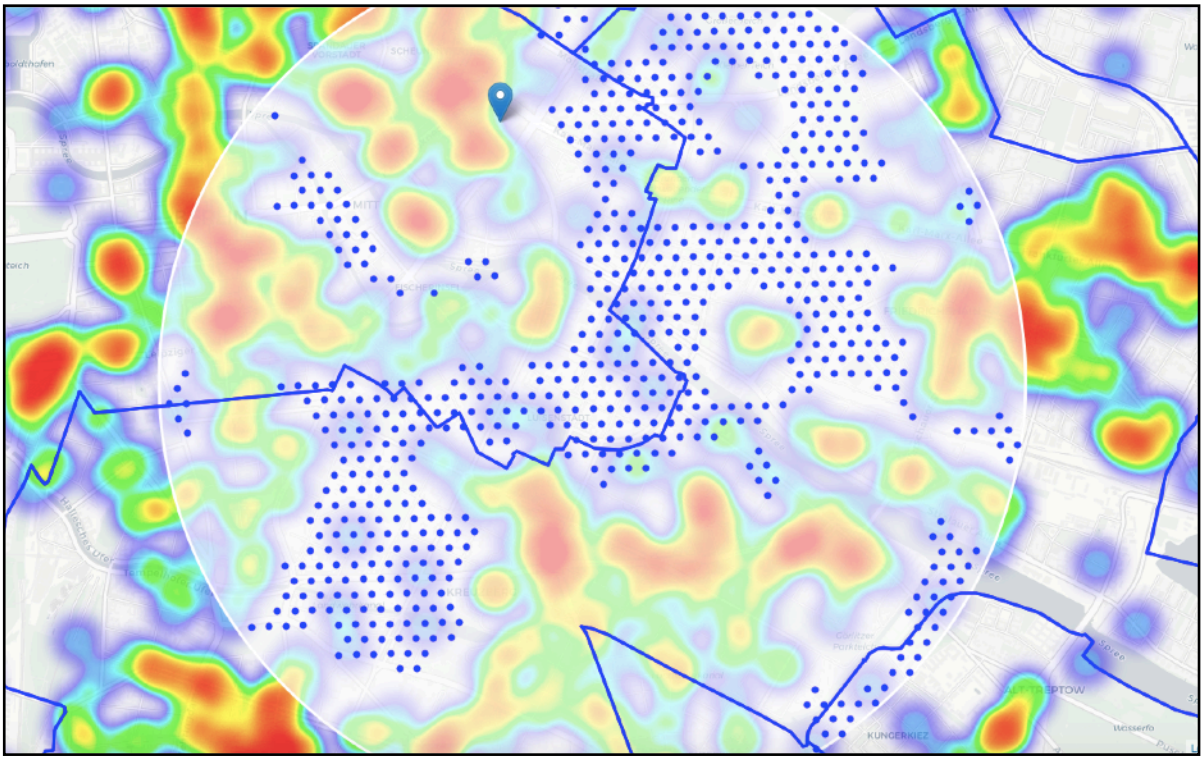

Figure: 8

Now that we have a few locations from the above analysis which are fairly close to the Berlin city centre “Alexanderplatz” mostly in “Kreuzberg”, “Friedrichshain” and south-east corner of “Mitte” boroughs satisfying the criteria set by us thus any of these locations can be a potential candidate for opening a new Indian restaurant, at-least based on nearby competition. Figure: 9 shows the interesting or final locations in the form of heat map.

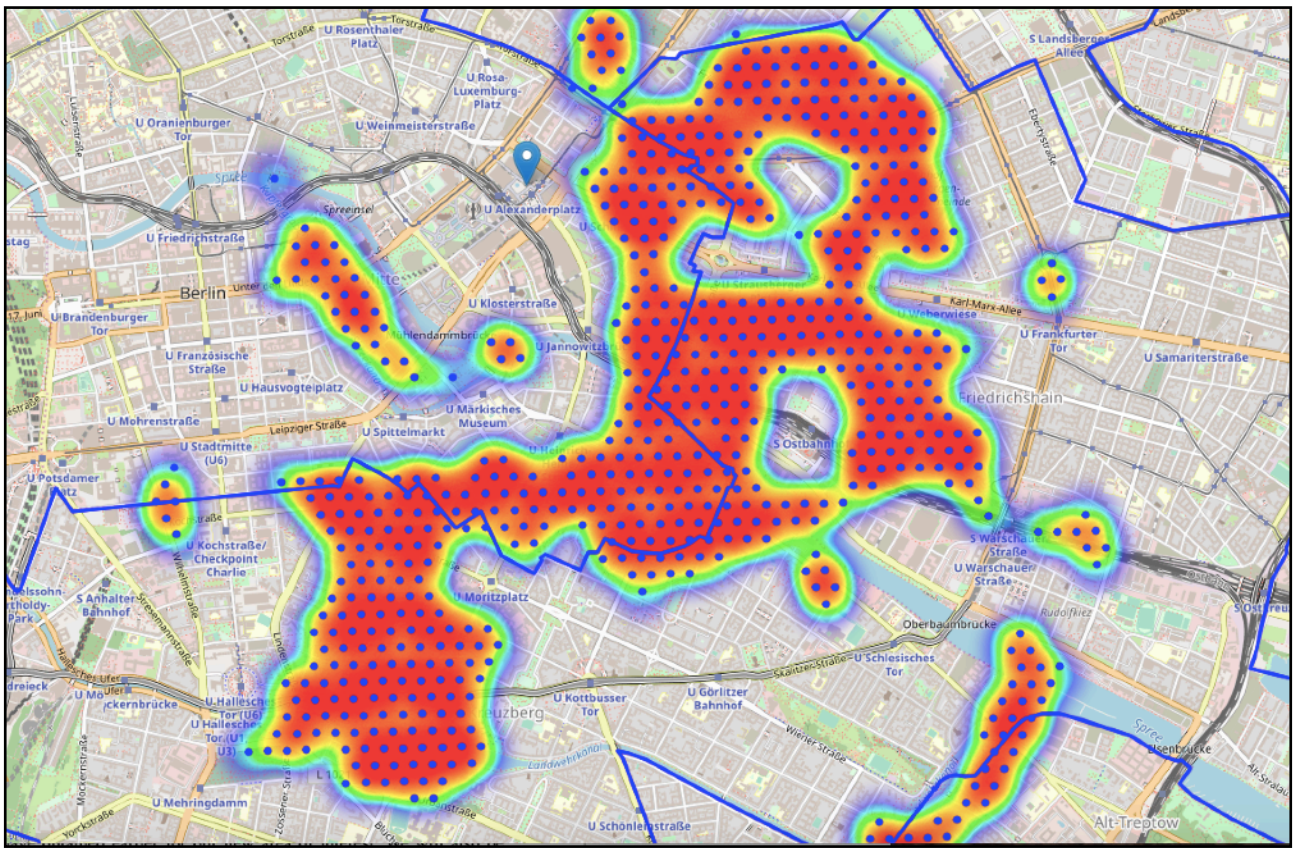

Figure: 9

Based upon our earlier analysis and from the heat map now we are in a position to create clusters of the the locations so that we can form centres of zones having preferred locations. These zones along with their centres and addresses will be the final result of our analysis.

After forming the clusters let us form the map of Berlin showing the clusters that we created. After that let us also form a magnified view of “Kreuzberg” and “Friedrichshain” to explore the neighbourhoods properly. Figures: 10, 11 and 12 shows the maps discussed above for better understanding.

Figure: 10

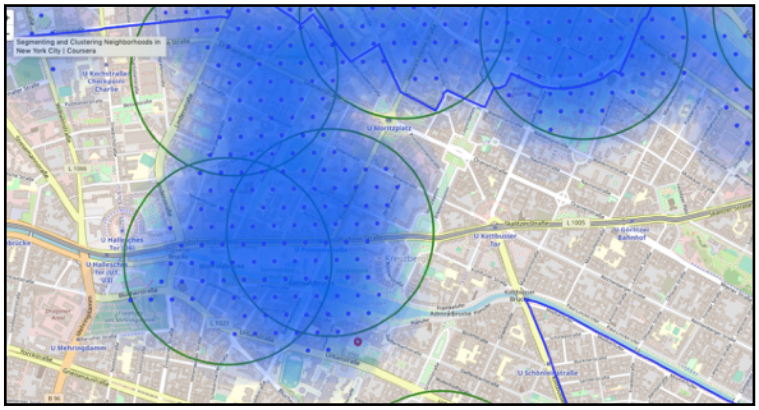

Figure: 11

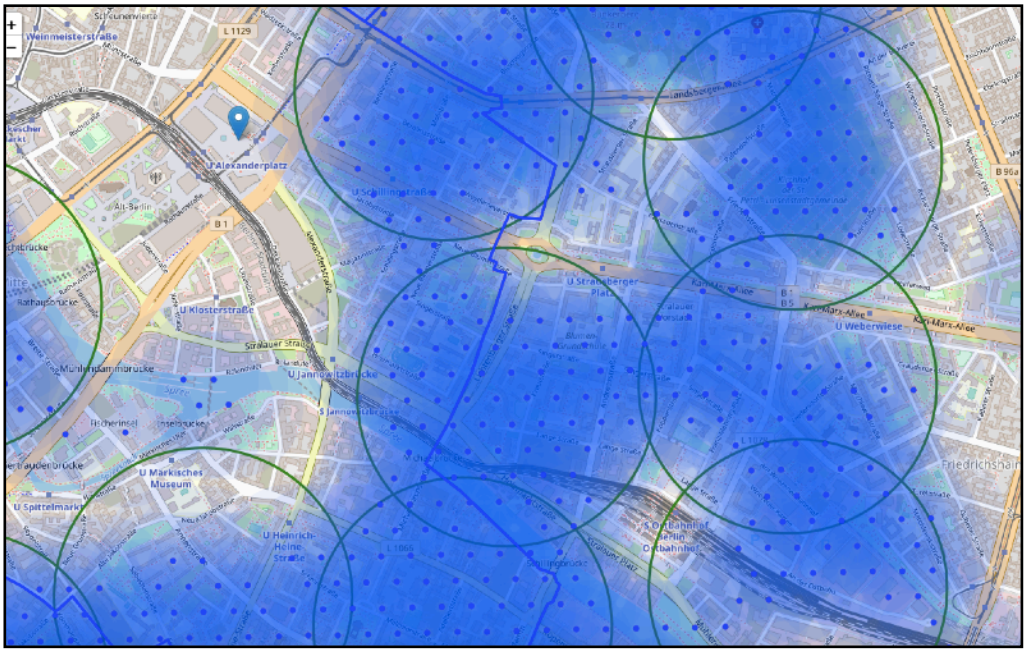

Figure: 12

After this we will use reverse geolocation in order to obtain the address of the geographical co-ordinates which we have obtained earlier for our new area of interest. We will also be calculating their distance from the Berlin city centre “Alexanderplatz”. Figure: 13 shows the final results of our analysis.

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

Themengarten Bachlauf, Fhain, Friedrichshain-Kreuzberg, Berlin, 10249, Deutschland
=> 1.5km from Alexanderplatz
Landwehrkanal, Carl-Herz-Ufer, Kreuzberg, Friedrichshain-Kreuzberg, Berlin, 10961, Deutschland
=> 2.9km from Alexanderplatz
26, Krautstraße, Fhain, Friedrichshain-Kreuzberg, Berlin, 10243, Deutschland
=> 1.2km from Alexanderplatz
5, Stralauer Allee, Fhain, Friedrichshain-Kreuzberg, Berlin, 10245, Deutschland
=> 3.5km from Alexanderplatz
87, Dresdener Straße, Mitte, Berlin, 10179, Deutschland                                                =
1.5km from Alexanderplatz
ehem. Postfuhramt, Melchiorstraße, Mitte, Berlin, 10179, Deutschland
=> 1.8km from Alexanderplatz
28, Straße der Pariser Kommune, Fhain, Friedrichshain-Kreuzberg, Berlin, 10243, Deutschland
=> 2.0km from Alexanderplatz
Kupfergraben, Spandauer Straße, Spandauer Vorstadt, Mitte, Berlin, 10178, Deutschland
=> 1.1km from Alexanderplatz
Deutsche Post, Karl-Kunger-Straße, Kungerkiez, Alt-Treptow, Treptow-Köpenick, Berlin, 12435, Deutschland
=> 3.9km from Alexanderplatz
Senatsverwaltung für Gesundheit und Soziales, 106, Oranienstraße, Kreuzberg, Friedrichshain-Kreuzberg, Ber
lin, 10969, Deutschland                                          => 2.0km from Alexanderplatz
Hellweg, An der Ostbahn, Fhain, Friedrichshain-Kreuzberg, Berlin, 10243, Deutschland
=> 2.4km from Alexanderplatz
12, Mollstraße, Mitte, Berlin, 10178, Deutschland                                                  => 0.7km
from Alexanderplatz
Tierpark Neukölln, 82, Hasenheide, Kreuzberg, Friedrichshain-Kreuzberg, Berlin, 10967, Deutschland
=> 3.8km from Alexanderplatz
4A, Diestelmeyerstraße, Fhain, Friedrichshain-Kreuzberg, Berlin, 10249, Deutschland
=> 1.9km from Alexanderplatz
Prinzeneck, Prinzenstraße, Kreuzberg, Friedrichshain-Kreuzberg, Berlin, 10969, Deutschland
=> 2.6km from Alexanderplatz
```

Figure: 13

5. Results and Discussion

Our primary analysis showed that there are plenty of restaurants in Berlin (approximatelt 2000) when we have created our initial area of interest to be12 Km x 12 Km around central location of Berlin, “Alexanderplatz”. From our analysis we could also observe some pockets of low restaurants density in an around the city centre. We have observed a relatively higher concentration of restaurants in the north and west of Berlin city centre whereas a low density were observed in the south, south-east and east of the city centre. On the basis of this result we directed our primary focus on “Kreuzberg” and “Friedrichshain” boroughs as they are quite popular not only among the tourists but also among the local people and these boroughs are quite close to the city centre, “Alexanderplatz”. Thus these boroughs can

become our primary interest while opening an Indian restaurants and they have low density of restaurants but they are important from socio-economic point of view.

Thus the selection of these boroughs will lead us to reduce the earlier 12 Km x 12 Km area around city centre to 6 Km x 6 Km of dense grid location candidates having a spacings of 100 m between them.

Then we filtered the locations on the basis of ,

- locations having no more than two restaurants in the locality
- locations having no Indian restaurants within 400 m of radius
- locations satisfying the above two criteria

The locations obtained were then clustered to create zones of interest containing good number of location candidates. The addresses of these zones were also obtained by reverse geolocator along with their distance from “Alexanderplatz”

Result of all these zones containing largest number of potential new restaurants locations based on number of and distance to existing locations (having both general restaurants particularly the Indian restaurants). This, of course, does not imply that those zones are actually optimal locations for a new restaurant. The primary purpose of this analysis was to only provide some information on areas close to Berlin centre but not crowded with existing restaurants (particularly Indian restaurants). 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.

6. Conclusion

Purpose of this project was to identify Berlin areas close to central location “Alexanderplatz” 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. By calculating restaurant density distribution from Foursquare data we have first identified general boroughs that justify further analysis,“Kreuzberg” and “Friedrichshain”. We 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 centres were created to be used as starting points for final exploration by stakeholders.

Final decision on optimal restaurant location will be made by stakeholders based on specific characteristics of neighbourhoods 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 neighbourhood etc.