

# **Final Project Report(Capestone/Coursera)**

**Sanjit Bhagat**

**April 28,2020**

## **Table of contents**

- Business Problem
- Data
- Methodology
- Analysis
- Result and Discussion
- Conclusion

## **Background**

Jaipur is the capital of India's Rajasthan state. At the centre of its stately street grid (notable in India) stands the opulent, colonnaded City Palace complex. With gardens, courtyards and museums, part of it is still a royal residence. As a capital city it has a good opportunity for an individual/business to open the restaurant/ coffee shop/ cafe near the city centre.

## **Business Problem**

This project is focusing on, finding the places where a \*restaurant/ coffee shop/ cafe\* can be opened. The project will lead to the destinations where there are \*few or no restaurants\* having the \*optimal distance\* from the \*city centre\*. This report will be targeted to stakeholders interested in opening a restaurant/ coffee shop/ cafe in \*Jaipur, India\*. AS it is a good opportunity to establish a food chain in the capital city. The optimal location will give public advantage as there are few or no restaurants nearby. The city centre is selected because it has a good amount of tourist places, shopping centres, colleges that will help to gather the audience.

## Data

Gathering the data is the most important factor of data science. More the data more the consistency in prediction. We will acquire the pin-codes of Jaipur city from the following <https://finkode.com/rj/jaipur.html> URL. This contains the postal addresses of all the locations in Jaipur almost 600 postal locations. We will need only locations that are near to the city centre as the outside the city its still a rural area. So we will filter locations according to our need with specific tools. The foursquare will help to find the public places regarding to the all coordinates we have gathered. The following methodology will be used We will require the process

- Get all the picodes of Jaiupr city
- Clean the data and will aquire almost all picodes near the location
- Get the coordinates of pincodes using geopy
- Use Foursquare to explore the needed venues with respect to all the locations we have
- Apply k-means to get the clusters having less venues
- Use Folium to plot the map wherever required

Let's first get the data from <https://finkode.com/rj/jaipur.html> , this contains all the pincodes with the concerned area name

```
|=====
| Post Office |District |Pincode
|0 |A C Jobner S.O |Jaipur |303329
|1 |Ajmer Road S.O |Jaipur |302006
|2 |Akhepura B.O |Jaipur |302013
|3 |Akoda B.O |Jaipur |303338
|4 |Akodiya B.O |Jaipur |303901
|=====
----
```

Here the data frame contains Post Office, District and \*Picode of Jaipur City\*  
----(589, 4)----

As, we can see there are 589 postal areas in jaipur city

## Methodology

In this project we will direct our efforts on detecting areas of Jaipur that have low restaurant density. We will limit our analysis to area ~20km around city center. In first step we have collected the required data: Area pincodes of jaipur city, Now further we will limit the distance to 20km from the city center. Then we will collect the venues and other details using \*FourSquare\* 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. Now, we will fetch the coordinates of the Jaipur City and its areas

----

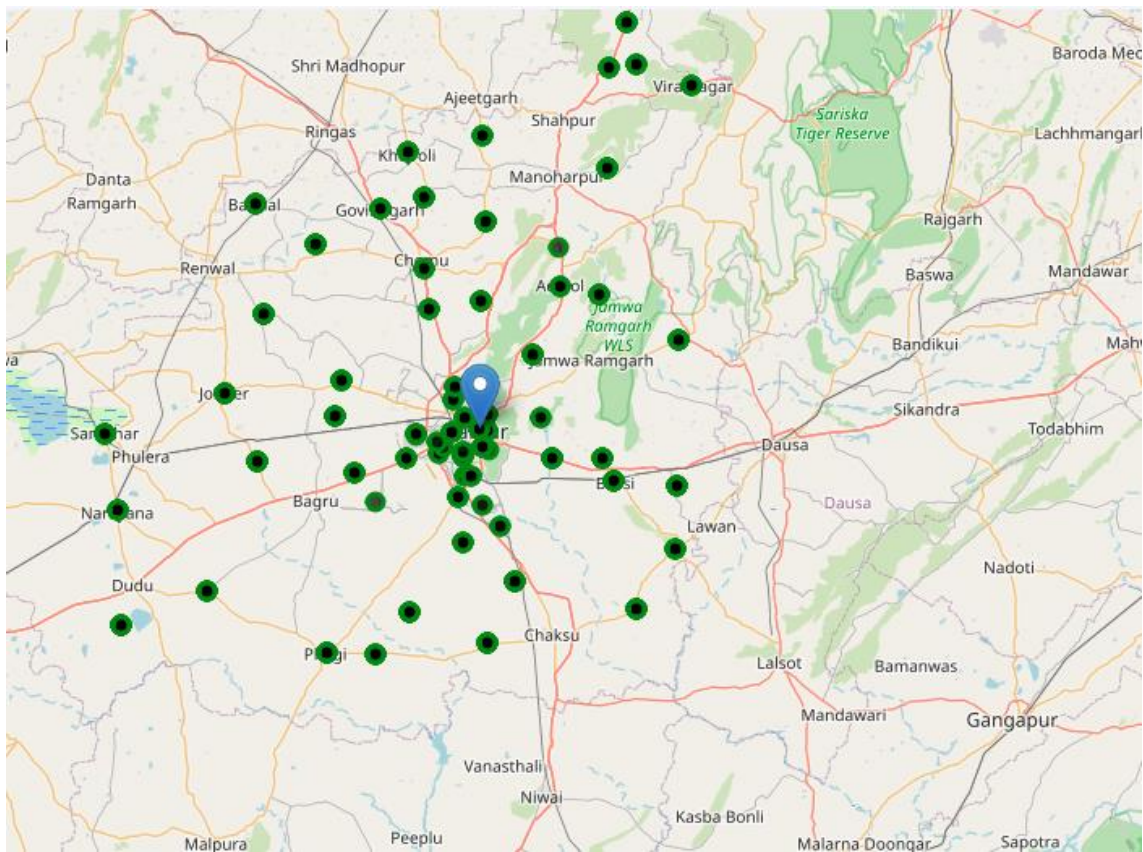
**The geograpical coordinate of Jaipur are 26.916194, 75.820349.**

----

## Now Data cleaning and processing

We will perform the operations like add, remove, sort, group to clean our data and visualize it

First let's have a look how it looks like on the map using \***folium**\*



As observed we can see we have 75 unique areas in all over jaipur

## Data Analysis

perform the analysis ,we will limit the distance to ~20kms and collect data from foursquare for the given limits.

After, we have our location candidates, let's use Foursquare API to get info on restaurants in each neighborhood.

We're interested in venues in 'food' category

Category ID corresponding to Food is taken from Foursquare web site

(<https://developer.foursquare.com/docs/resources/categories>) 1. we will get the venues nearby the areas we have filtered

Food(4d4b7105d754a06374d81259) is used as the category

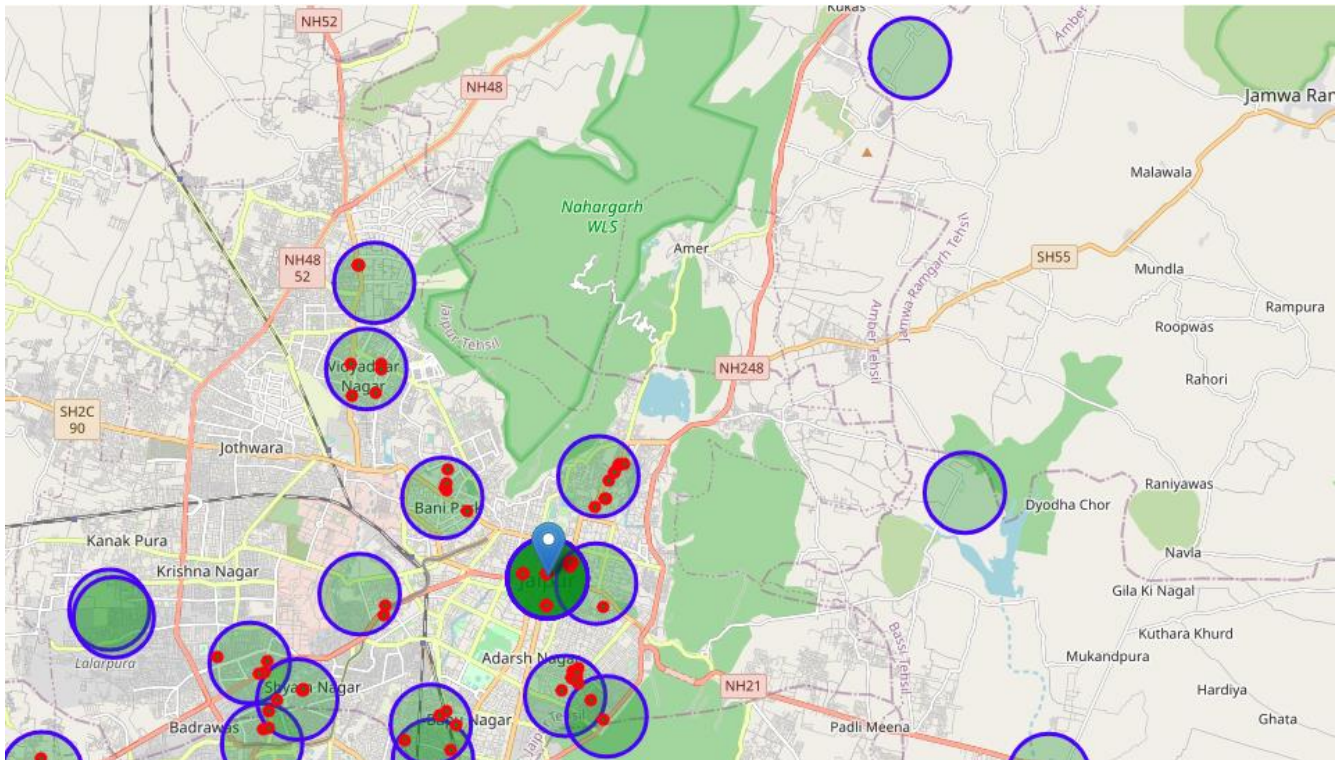
After all these process, We have the venues their coordinates distance which is less then 20kms etc..

```
|=====
| |Area |Latitude |Longitude |Pincode |No_of_restaurants
|0 |Ajmer Road S.O |26.912666 |75.773900 |302006 |2
|1 |Akhepura B.O |26.981239 |75.777517 |302013 |3
|2 |Amba Bari S.O |26.962243 |75.775541 |302039 |5
|3 |Amer Clark Hotel S.O |26.846508 |75.794000 |302018 |11
|4 |Amer Road S.O |26.938614 |75.832950 |302002 |9
|5 |Arjun Lal Sethi Nagar S.O |26.890235 |75.824723 |302004 |8
|2 |Amba Bari S.O |26.962243 |75.775541 |302039 |5
|24 |Shyam Nagar S.O (Jaipur) |26.889447 |75.758470 |302019 |5
|8 |Bajaj Nagar S.O |26.876235 |75.792026 |302015 |5
|18 |Jagatpura S.O |26.845418 |75.806878 |302017 |5
|10 |Bani Park S.O |26.933842 |75.794351 |302016 |5
|13 |Bhopawa B.O |26.916194 |75.820349 |303807 |4
|16 |Heerapura S.O |26.897467 |75.746716 |302021 |4
|23 |Nabhawala B.O |26.916194 |75.820349 |303507 |4
|17 |High Court S.O (Jaipur) |26.883897 |75.791502 |302005 |4
```

|25 |Udaipuria B.O |26.916194 |75.820349 |303509 |4  
 |11 |Bhalloji B.O |26.916194 |75.820349 |303110 |4  
 |9 |Bandha Jamwaramgarh B.O |26.916194 |75.820349 |303109 |4  
 |6 |Ashok Nagar S.O (Jaipur) |26.916194 |75.820349 |302001 |4  
 |1 |Akhepura B.O |26.981239 |75.777517 |302013 |3  
 |14 |Bilwa B.O |26.769897 |75.854512 |302022 |3  
 |15 |Haldiyan Ka Rasta S.O |26.914957 |75.832379 |302003 |3  
 |12 |Bhambhoria B.O |26.873418 |75.695346 |302026 |3  
 |21 |Mansarovar S.O |26.879649 |75.749776 |302020 |3  
 |20 |Kalyanpura B.O |26.885701 |75.834936 |303503 |2  
 |22 |NIMS University S.O |27.192247 |75.954757 |303121 |2  
 |0 |Ajmer Road S.O |26.912666 |75.773900 |302006 |2  
 |19 |Kalwara B.O |26.813043 |75.784938 |302029 |1  
 |7 |Awania B.O |26.850938 |75.608524 |303007 |1

We can observe there is maximum 11 no of restaurants in a location

Lets look into visual form with the pincodes in 1km radius and restaurants in them



Looks good, and we can see some locations don't have even a single restaurants..

We will use 7 segments to cluster the data and map a graph to the centers of the k mean  
Those locations will be Prime Location\* and can be used for further analysis

----

### **Clusters: [1 0 0 5 6 4 6]**

31

[[26.95910812 75.78246964]

[26.89563142 75.73561435]

[26.77217387 75.82393455]

[26.90401993 75.93393922]

[27.0306745 75.9100265 ]

[26.85302038 75.79386894]

[26.9126671 75.82470823]]

----

### **Addresses of centers of areas recommended for further analysis**

1. Shiv Nagar, Vidyadhar Nagar, Jaipur Municipal Corporation, Jaipur Tehsil, Jaipur, Rajasthan, 302039, India
2. Tagore Nagar, Badrawas, Jaipur Municipal Corporation, Jaipur Tehsil, Jaipur, Rajasthan, 302021, India
3. Sanganer Tehsil, Jaipur, Rajasthan, 302022, India
4. TRatanpura, Bassi Tehsil, Jaipur, Rajasthan, 302031, India
5. Khoramina, Kukas, Amber Tehsil, Jaipur, Rajasthan, 302028, India
6. Durgapura, Bapu Nagar, Jaipur Municipal Corporation, Sanganer Tehsil, Jaipur, Rajasthan, 302015, India
7. Moti Doongri Road, Fateh Tibba, Adarsh Nagar, Jaipur Municipal Corporation, Jaipur Tehsil, Jaipur, Rajasthan, 302004, India

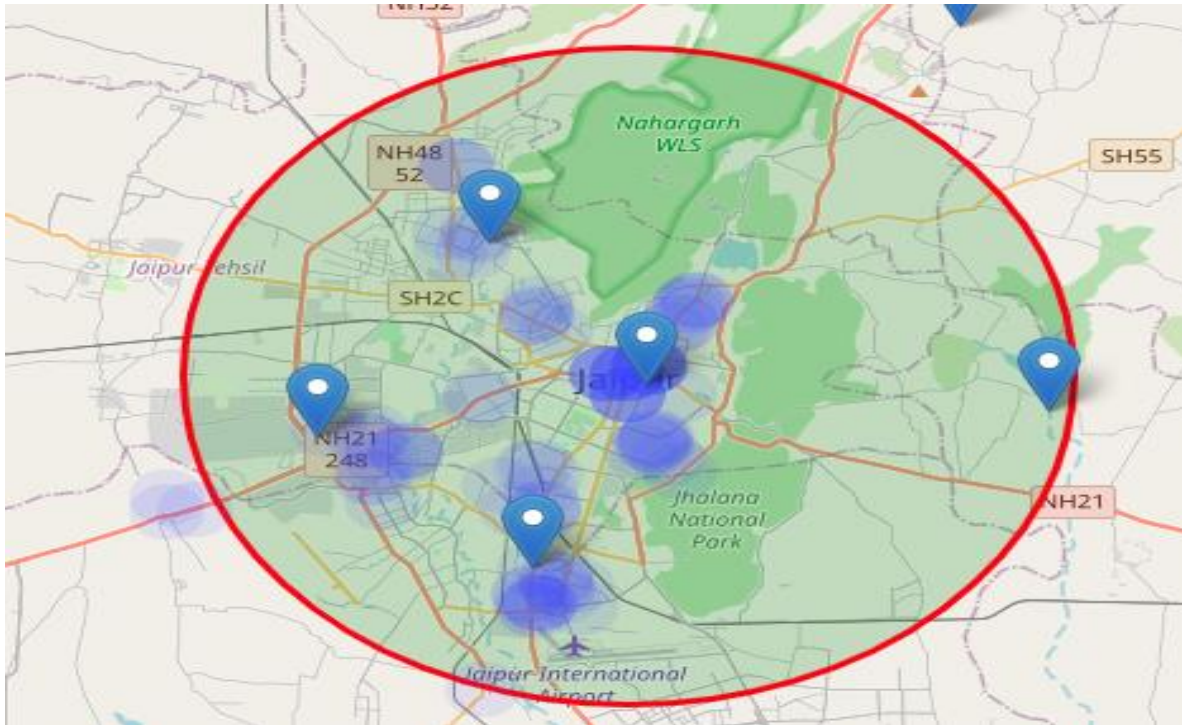
----

This concludes our analysis. We have created 15 addresses representing centers of zones containing locations with low number of restaurants. These seven locations have less than 7 restaurants in the corresponding area, so it's good to observe these sites. Although zones are shown on map with a radius of ~1000 meters (green circles), their shape is actually very irregular and their centers/addresses should be considered only as a starting point for exploring area neighborhoods in search for potential restaurant locations. Most



of the zones are located in Jaipur, which we have identified as interesting due to being popular with tourists, fairly close to city center and well connected by public transport.

Let's Visualize the final map



Here the big circle denotes the area of 12 km radius from the center, the poups shows the **Prime Locations** for further analysis having less then 7 restaurnats/cafe/coffeeshops

## Results and Discussion

Our analysis shows that there are good no of restaurants in the range of 20kms, so we can say its a good opputunity for the stakeholder to establish new business.Further we observed that some locations have more than 10 restaurants , means they are in proccessing stage.We found 26 locations where the no restaurants or less then 12 restaurants 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 7 zones containing largest number of potential new restaurant

locations based on number of and distance to existing venues. 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 Berlin center but not crowded with existing 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.

## **Conclusion**

Purpose of this project was to identify jaipur areas close to center with low number of restaurants in order to aid stakeholders in narrowing down the search for optimal location for a new restaurant. By calculating restaurant density distribution from Foursquare data we have first identified general boroughs that justify further analysis, 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 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 neighbourhood etc.