

|  |
| --- |
| Foursquare Battle of Delhi, India  2020 |
|  |
| May 27  COMPANY NAME  Authored by: KT |

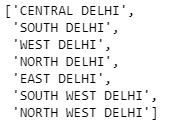
# Business Problem

|  |  |
| --- | --- |
| Leverage Foursquare locations data to explore venues in Delhi, India.  Find insights of most popular venues in sub-parts of Delhi. | |
| *“Delhi's 2020 population is now estimated at 30,290,936.”* | |
| It covers a large area, totaling around 1,484 square kilometers. The city has a population density of 20,412 people per square km which is one of the highest in the world.  Using tools learnt in coursera’s [IBM Data Science Professional Certificate](https://www.coursera.org/professional-certificates/ibm-data-science?) will help explore the same. Data On searching over the internet, an informative blogpost mentions all the “PostalCodes” along with “Districts” of Delhi.  [*https://www.whatsuplife.in/delhi/blog/zip-pin-postal-code-pincodes-delhi*](https://www.whatsuplife.in/delhi/blog/zip-pin-postal-code-pincodes-delhi)  Below are a couple of screenshots of the data from this webpage that were extracted: |

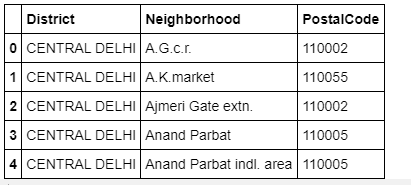


This blogpost webpage was scrapped to get –

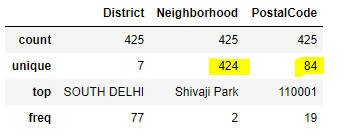
1. Districts in Delhi



1. Corresponding Neighborhoods and PostalCodes

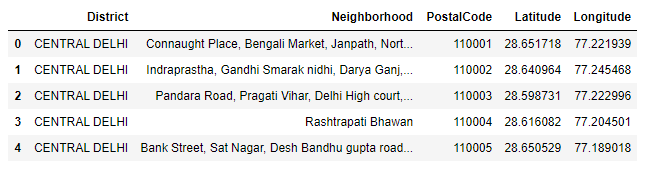


As you can notice in the data above, same postal codes have multiple neighborhoods, which are contributing to multiple rows.

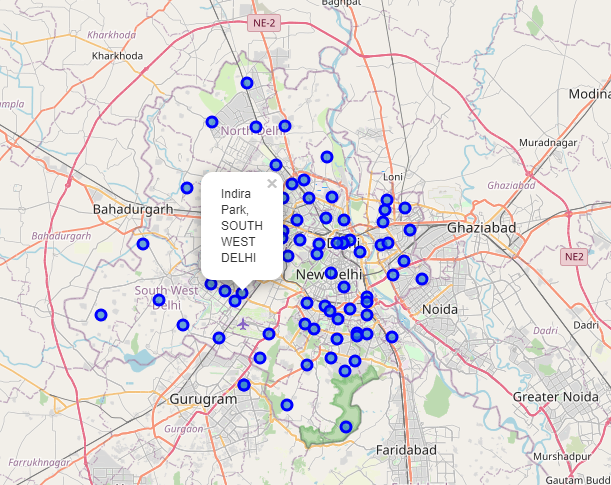


Therefore, merge the neighborhoods that have same postal code and get corresponding latitude and longitude of the same.

Here it is how it looks as the final data set that was used to explore Delhi.



Neighborhoods and Postal Codes on Delhi Map

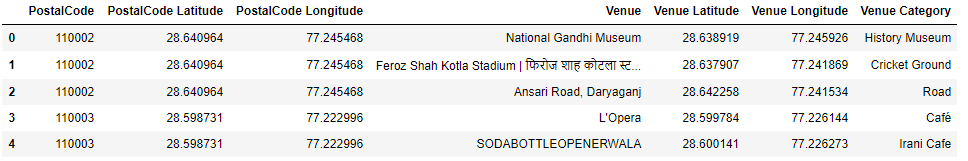


# Exploration

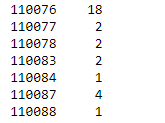
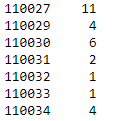
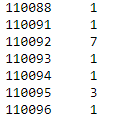
Used foursquare venues explore API - [*https://api.foursquare.com/v2/venues/explore*](https://api.foursquare.com/v2/venues/explore)to fetch the top 100 venues corresponding to each postal code.

Number of venues fetched for 84 Postal Codes were 409 which on an average is less than five venues per postal code

Here is a snapshot of Delhi venues –

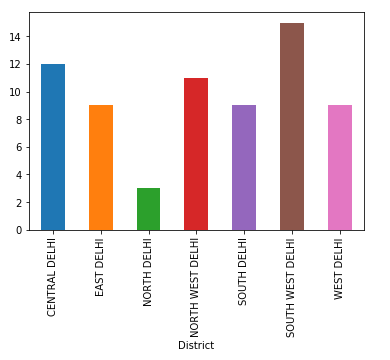


Snapshots of venues categories per postal code as returned by the Foursquare API –

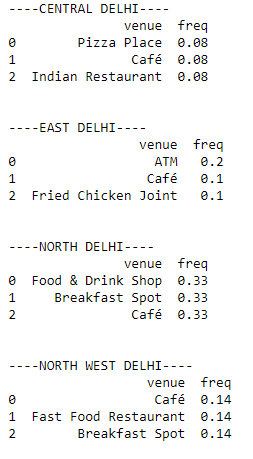
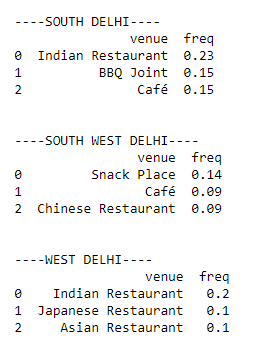
There are 99 unique categories for 84 postal codes from the foursquare data.

Since number of venues per postal code is very small, let us group venue categories by Districts.

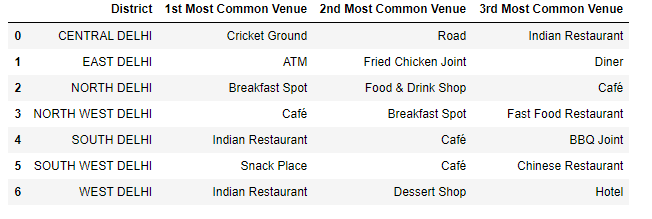


From the above graph, we can see venue categories have a good distribution over districts of Delhi.

Using one-hot encoding, top three venue categories were calculated for each district.

Printing in a dataframe



As you can see from above analysis - Central Delhi, North Delhi and North West Delhi top three venues have the same frequency, therefore the results are ambiguous. This proves the Foursquare API does not have enough "good" data for Delhi to perform analysis.

# Conclusion

Using foursquare API data, I tried to explore different postal codes of Delhi, but the data was very limited, therefore I switched to districts and analyzed the venues available. During analysis of top venues in district, I found out the data is ambiguous also which can be seen from Central Delhi, North Delhi and North West Delhi top three venues have the same frequency, therefore the results are ambiguous.

## Lesson learnt

This project is a very good reference and start point for me to further pursue exploration of Delhi, India from other data sources like google places API, and may be some local places providers.