DATA SCIENCE ON BANGALORE CITY
THE BATTLE OF NEIGHBOURHOODS – A REPORT

UMANATH VINAYAK   Umanath.vinayak@gmail.com

TABLE OF CONTENTS

INTRODUCTION

The city of Bengaluru (Bangalore) is India's own Silicon Valley. It is one of the fastest growing cities in the world, with a population of more than 8M as of 2011. The total area of the city ~2200 sq. Km with population still increasing phenomenally. Out of 8M, nearly 5M are Male and 3M are female population. By one estimate, the population Growth has been 47.18% when compared with the population senses taken by the year 2001.

This makes the city vibrant, and most attractive market in India for any business person/ entrepreneur to put his best foot to expand & grow his business.

But, the most important question haunts the Investor, "where and what business should be best bet"?

Well, this report exactly does that.
This Report tries to capture the significant Neighborhoods in the Metropolitan city, analysis the already present infrastructure/ venues and describes the best location for an entrepreneur to determine a spot to kick start his business.

City Overview Map
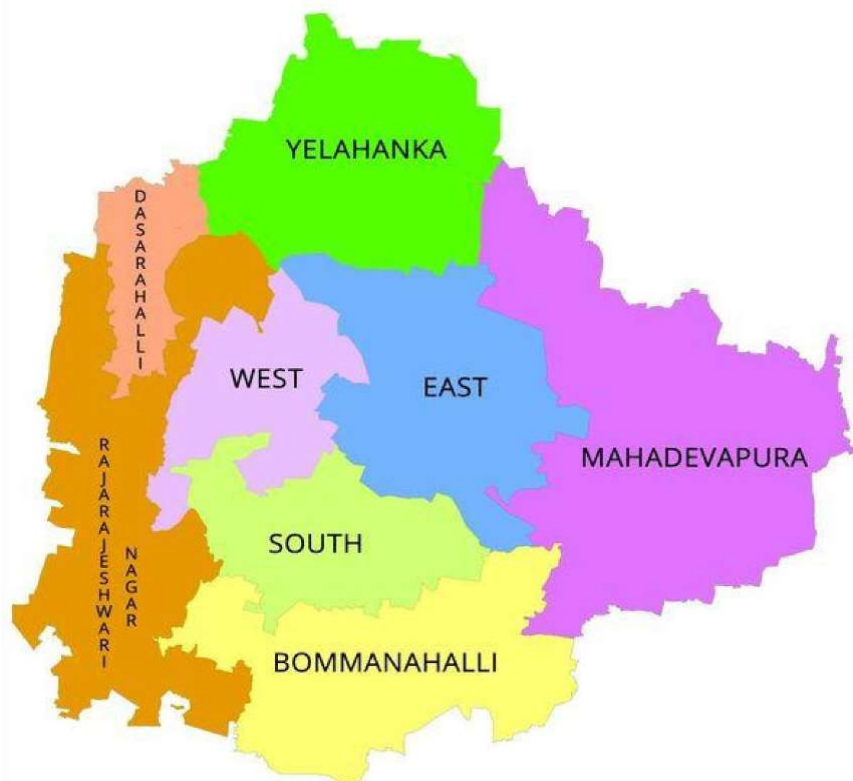
DATA COLLECTION

For this analysis, below said data is required:
- ➤ City's Neighborhood and their population
- ➤ City's Co-ordinates
- ➤ City's current infrastructure (i.e) Venues

City's Neighborhood data

The City's Neighborhood data was obtained from India's free to use datacenter website - https://data.gov.in/.

The dataset consists of Bangalore's Zone and ward wise – Area and its population – details.

City Name	Zone Name	Ward Name	Area (in sq km)	Total_Population_in_k
Bengaluru	Dasarahalli	Bagalakunte	4.3	65113
Bengaluru	Dasarahalli	Bagalakunte	5.7	57062
Bengaluru	Dasarahalli	Bagalakunte	1.9	47004
Bengaluru	Dasarahalli	Chokkasandra	3.8	59289
Bengaluru	Dasarahalli	Chokkasandra	1.0	24181
Bengaluru	Dasarahalli	Chokkasandra	1.5	33521
Bengaluru	Dasarahalli	Mallasandra	1.3	41482
Bengaluru	Dasarahalli	Mallasandra	2.1	35228
Bengaluru	Dasarahalli	Mallasandra	1.6	34299
Bengaluru	Dasarahalli	Peenya Industrial Area	5.5	57814
Bengaluru	Dasarahalli	Peenya Industrial Area	1.1	33866

City's Neighborhood Co-ordinates
The City's Neighborhood Co-ordinates were obtained using Python Software library – Geocode.
The co-ordinates can be obtained just by providing inputs, such as the ward Name or Zone Name to the Geocode.

City's current Venue details
The venue details are obtained using **Foursquare API.**

Using each Neighborhood's population detail and its already existing venues, we can understand existing infra.

This enables us to view the opportunities available and predict what Type/Category of venue is best to open as new one.

METHODOLOGY

Data Visualization

The City's Neighborhood data on a plot shows the Total population of each zone.
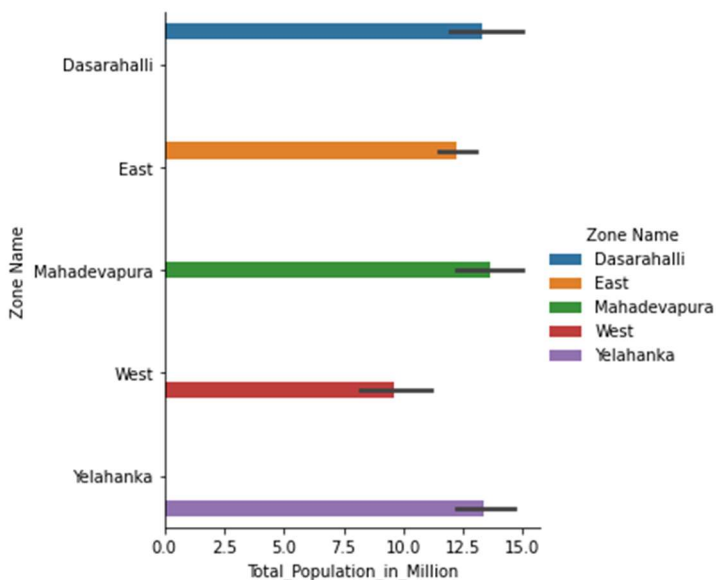

We can see that Zones - 'Dasarahalli', 'Mahdevapura' & Yelahanka' are heavily populated.

Let us see the top 5 most populated neighborhoods in all of bangalore

	City Name	Zone Name	Ward Name	Area (in sq km)
36	Bengaluru	Mahadevapura	Horamavu	23.0
64	Bengaluru	Yelahanka	Thanisandra	16.8
0	Bengaluru	Dasarahalli	Bagalakunte	11.9
8	Bengaluru	East	C.V. Raman Nagar	14.2
56	Bengaluru	Yelahanka	Atturu	11.4

	Total_Population_in_Million	Latitude	Longitude
36	17.37	13.027331	77.660151
64	17.30	13.054713	77.633926
0	16.92	13.059800	77.499883
8	16.71	12.985648	77.664977
56	15.56	13.107470	77.565685
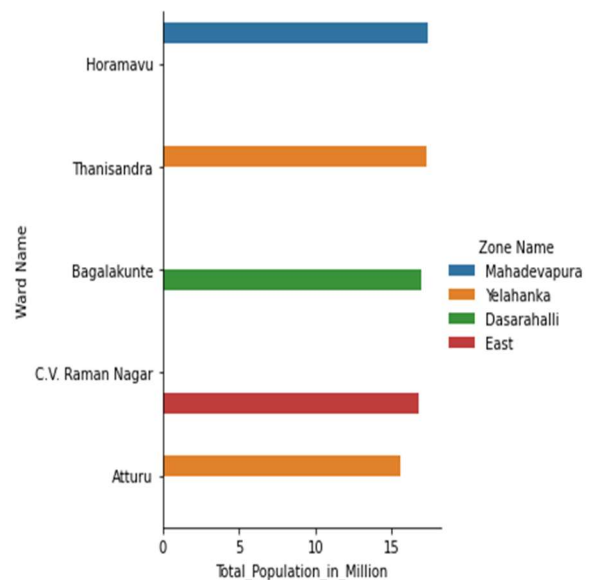
Using folium, a map was plotted to show how the different neighborhoods are spread all across Bangalore.
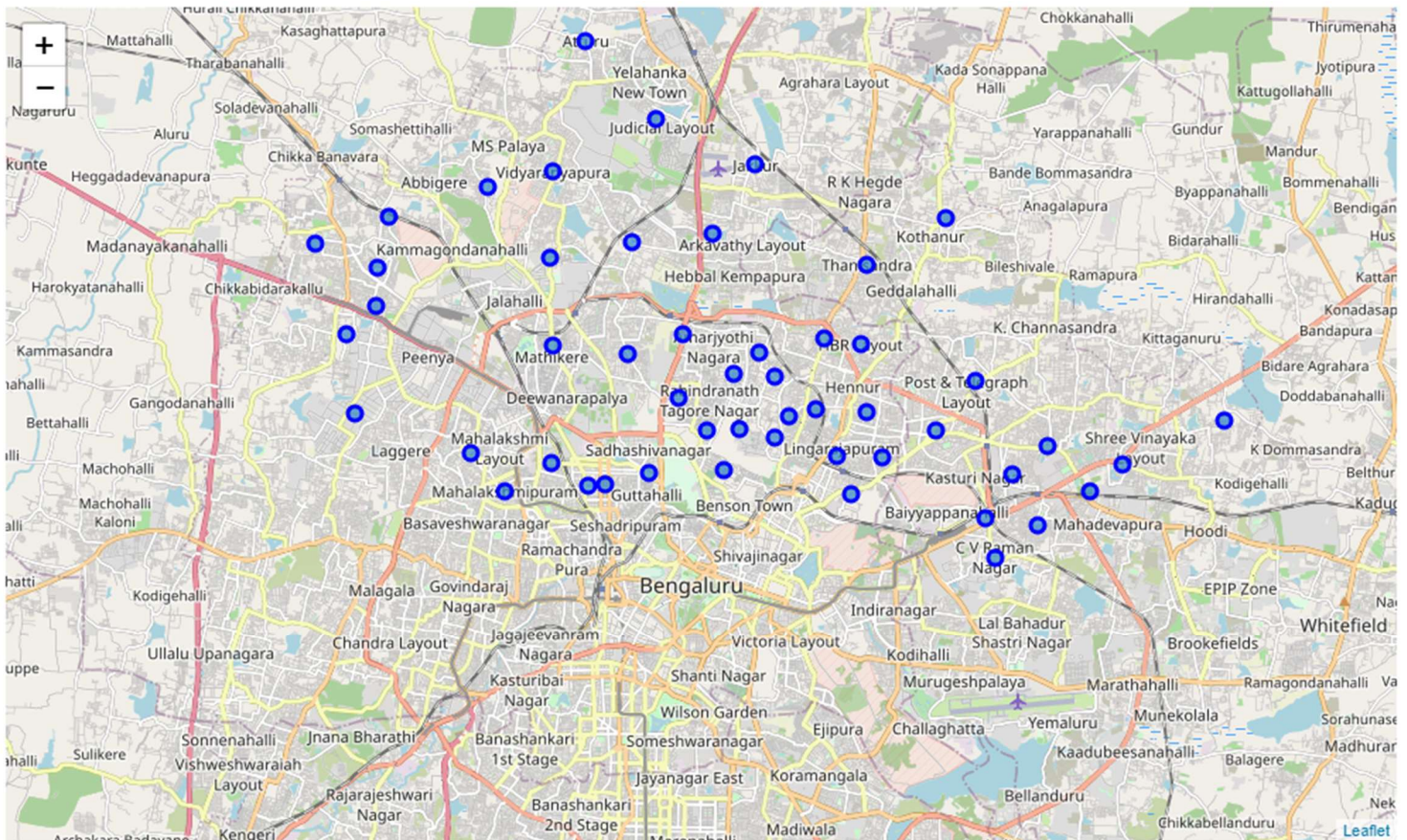
Feature Engineering was done on the venues obtained from Foursquare API using Unsupervised learning model. To do so, venue categories were converted/normalized by 'one-hot-encoding' method.

This was then loaded into data-table containing the 'Zone Name', 'Ward Name', 'Longitude', 'Latitude', 'Area', 'Population'.

Having this as the base data table 'K-means clustering' model was modelled to depict the clusters of venues.
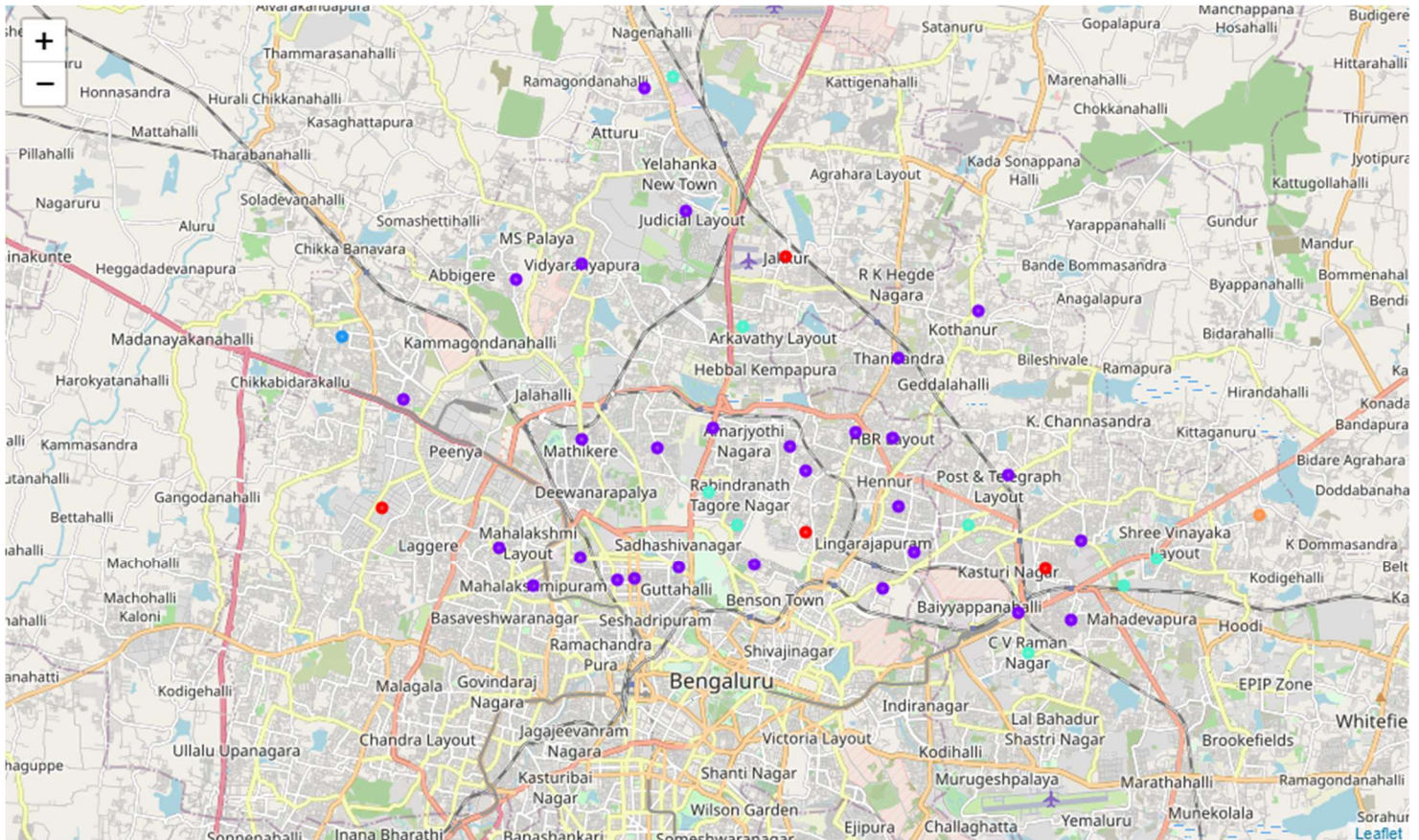
Cluster-1 – Utilities Cluster

	Ward Name	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
3	Peenya Industrial Area	13.019567	77.509589	0.0	ATM	Dry Cleaner	Restaurant	Cosmetics Shop	Dessert Shop	Diner	Donut Shop	Electronics Store	Fast Food Restaurant	Flea Market
24	Muneshwara Nagar	13.014081	77.611474	0.0	ATM	Food Truck	Dessert Shop	Diner	Donut Shop	Dry Cleaner	Electronics Store	Fast Food Restaurant	Flea Market	Flower Shop
40	Vijnanapura	13.005338	77.669363	0.0	ATM	Food Truck	Dessert Shop	Diner	Donut Shop	Dry Cleaner	Electronics Store	Fast Food Restaurant	Flea Market	Flower Shop
60	Jakkuru	13.078474	77.606894	0.0	ATM	Food Truck	Dessert Shop	Diner	Donut Shop	Dry Cleaner	Electronics Store	Fast Food Restaurant	Flea Market	Flower Shop

Cluster-2 – Hotel Industry

	Ward Name	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
5	T Dasarahalli	13.045141	77.514789	1.0	Karnataka Restaurant	Resort	Electronics Store	Indian Restaurant	Indian Sweet Shop	Metro Station	Pizza Place	ATM	Snack Place
7	Benniganahalli	12.994919	77.662603	1.0	Clothing Store	Indian Restaurant	Fast Food Restaurant	Department Store	Coffee Shop	Snack Place	Pharmacy	Convenience Store	New American Restaurant
12	HBR Layout	13.035870	77.632360	1.0	North Indian Restaurant	Coffee Shop	Café	Road	Women's Store	Flower Shop	Diner	Donut Shop	Dry Cleaner
13	Hebbala	13.038325	77.589063	1.0	Pharmacy	Indian Restaurant	Bakery	Café	Food Truck	Diner	Donut Shop	Dry Cleaner	Electronics Store
16	Kacharkanahalli	13.019958	77.633948	1.0	Indian Restaurant	South Indian Restaurant	Fast Food Restaurant	Burger Joint	Korean Restaurant	Coffee Shop	Snack Place	Electronics Store	Chinese Restaurant
18	Kammanahalli	13.009346	77.637709	1.0	Lake	Bar	Women's Store	French Restaurant	Diner	Donut Shop	Dry Cleaner	Electronics Store	Fast Food Restaurant
19	Kaval Bairasandra	13.028397	77.611513	1.0	Department Store	Fast Food Restaurant	Bakery	Women's Store	Food Truck	Diner	Donut Shop	Dry Cleaner	Electronics Store
23	Maruthi Seva Nagar	13.000561	77.630008	1.0	Fast Food Restaurant	Clothing Store	Department Store	Multiplex	Coffee Shop	Lounge	Shopping Mall	Café	Bus Station
25	Nagavara	13.037465	77.623602	1.0	Vegetarian / Vegan Restaurant	Indian Restaurant	Andhra Restaurant	Diner	Pizza Place	Bus Station	Women's Store	Food Court	Donut Shop
28	Ramaswamy Palya	13.006307	77.599353	1.0	Electronics Store	Café	Business Service	Women's Store	Department Store	Diner	Donut Shop	Dry Cleaner	Fast Food Restaurant
31	SanJayanagar	13.033560	77.575898	1.0	Park	Smoke Shop	Indian Restaurant	Fast Food Restaurant	Juice Bar	Athletics & Sports	French Restaurant	Donut Shop	Dry Cleaner
32	Vishwanath Nagenahalli	13.034010	77.607704	1.0	ATM	Fast Food Restaurant	Business Service	Department Store	Food Truck	Diner	Donut Shop	Dry Cleaner	Electronics Store
33	A Narayanapura	12.993080	77.675359	1.0	Electronics Store	ATM	Indian Restaurant	Food Truck	Dessert Shop	Diner	Donut Shop	Dry Cleaner	Fast Food Restaurant
36	Horamavu	13.027331	77.660151	1.0	Indian Restaurant	Athletics & Sports	Dry Cleaner	Chinese Restaurant	Women's Store	Food Truck	Diner	Donut Shop	Electronics Store
37	Hudi	13.065763	77.653033	1.0	Pizza Place	Sandwich Place	Gym / Fitness Center	Pub	Bakery	Badminton Court	Convenience Store	North Indian Restaurant	Italian Restaurant

RESULTS & DISCUSSION

Our Analysis shows that there are many Restaurants & eateries, grouped in clusters, are mostly clustered to the west and Central Bengaluru. Also, utility works such as Dry cleaners and Electronic stores are distributed at too much distance. This give the stakeholder to focus on the place/neighborhood he/she can spot to start his business.

CONCLUSION

The Purpose of this project was to identify the ideal locations for business inverstor to start a business.
So, using the Foursquare API was used to get the density of stores/venues that are already available in the Neighborhood.
Then clustering of vensues in each neighborhood was performed.

By looking at the clusters of venues by density, the stakeholders can deduce the type of business they would like to start in a neighborhood.