

Bengaluru Real Estate Data Analysis

A Capstone Project for Applied Data Science Capstone – IBM/Coursera

Done by: Shravan Bharadwaj

1. Introduction /Business Problem

1.1. Introduction

Bengaluru is one of the major metropolitan cities of India, often dubbed as the ‘Silicon City of India’. It is the capital of Karnataka and is one of the major hubs of the IT industry. It is home to more than 1 Crore i.e. 10 million people. Bengaluru is divided into 198 neighborhoods or wards.

The Real Estate of Bengaluru is developing rapidly and it is quite expensive to buy a plot of land. Thus, it is very important to know the information regarding any neighborhood while investing in any real estate.

1.2. Problem

In this project, Data analysis of Bengaluru Real Estate Prices will be done using data science methods. Data like prices, availability of water, proximity towards Metro, Hospital, Market etc. will be done by making use of data from various sources. The data will be clustered using a machine learning algorithm and the outcome will be presented using a dashboard application.

Targeted Audience:

The targeted audience is any individual or organization who intends to invest in the real estate of Bengaluru. This project enables them to get necessary insights from the analysis of data like Price Rates, Water Source, Proximity towards important venues like metro, hospital, grocery store etc. which are important for day-to-day life activities.

The investor would also get to know the important insights like Average prices of a land in Bengaluru. With the analysis of clustering, the investor would get to know which neighborhoods have their advantages and disadvantages while buying any plot of land in any area.

By the application of dashboard (which will be presented in the program), the investor can select the neighborhood of his desire and the important details like Price, Water Source, Population Residing and Proximity towards venues will be presented. Other important maps will also be presented.

2. Data

2.1. Factors of Influence

Based on the problem stated above, following are the factors which influence in buying any plot of land in Bengaluru:

1. Neighborhood Site Price Data in Rupees per Square Feet for prices estimation.
2. Kaveri River Water Availability based on the supply of water in number of days in a week in each neighborhood.
Note: River Kaveri is the major river which supplies water to Bengaluru. Other supplies of water would be small lakes and groundwater. Here Kaveri water availability is used in the analysis.
3. The question of how close are important venues like metro, hospital, grocery store, clothing store, bank etc. to the neighborhood, is a very important factor which comes in play while buying any plot of land.

2.2. Data Sources

Following are the data sources for this project:

1. Real Estate price data from websites like Makaan, MagicBrics, 99 Acres and Housing.com.
2. Kaveri River water availability based on the supply of water in number of days in a week from the website of Bangalore Water Supply and Sewage Board.
3. Bengaluru ward data from BBMP for names of the wards, latitudes and longitudes, population residing.
4. Venue data from Foursquare API to get the minimum distance of venues like metro, hospital, clothing store etc. from a given ward/neighborhood.

2.3 Data description:

1. House prices are estimated in Rupees per square feet in India. If a plot of land has 5000 Rs/ Sq. Ft. and its estimated area is 2400 Sq. Ft., then the total cost of land would be 2400 times 5000 which yields 12 Million rupees.
2. Water Supply of Kaveri River is based on the supply of number of days in a week. The Water board supplies the water for about 1-2 Hours on those particular days in Bengaluru. Water is stored in tanks and used for domestic purposes.
3. The wards data of Bengaluru contains the names of 198 wards/ neighborhoods. It has the names of the wards, coordinates of the wards and the population residing in the ward.
4. The Foursquare API will be used for getting the minimum distances of venues like metro, hospital etc. which will be used for clustering.
5. Bengaluru Geojson file is used for plotting choropleth maps. It contains the geographic details of all the wards in Bengaluru.

3. Methodology

3.1. Data Importing:

1. The data of house prices is imported by different websites stated above by copying it to excel. Data is not scraped from the website because the pages contain multiple pages within a page and all the data could not be scraped. Also, the names of areas are not in conjunction with the BBMP neighborhood data.
2. The Data of Kaveri water availability is imported from the website of BWSSB by copying it to excel. Data is not scraped because the data could not be scraped from the website.
3. The Bengaluru ward data is imported as a CSV file. It has all the data like Latitude, Longitudes, Population, Borough name and Ward name.
4. Venue data are obtained from Foursquare API for getting the distance of venues located near the ward.
5. Bengaluru ward Geojson file is imported as a Geojson file from BBMP.

3.2. Data Cleaning:

1. All 198 wards are matched with the Site prices manually in Excel Sheet.
2. The 198 wards all have their latitude, longitude and population.
3. The data of Kaveri data has timings related to the availability of water in a week. Some areas get the water alternate days in a week, three days in a week, two days in week, one day in a week and some do not get Kaveri water. Depending on the no of days Kaveri water is available, the number of days is assigned to a ward/neighborhood.
4. All the data is combined and converted to a CSV file. The CSV file is read in the Jupyter Notebook for processing.

	Borough	Neighborhood	Latitude	Longitude	Population	Price	Kaveri
0	Yelahanka	Chowdeswari Ward	13.121710	77.580420	19626	3644	2
1	Yelahanka	Atturu	13.102805	77.560036	24020	3295	2
2	Yelahanka	Yelahanka Satellite Town	13.090986	77.583920	25782	3900	2
3	K.R. Puram	Vijnanapura	13.006063	77.669560	35087	5840	0
4	K.R. Puram	Basavanapura	13.016847	77.715454	22012	4000	0

Figure 1. Display of the Initial Data frame

5. By using GeoPy library of Python, the map of Bengaluru can be obtained with the all the 198 neighborhoods.

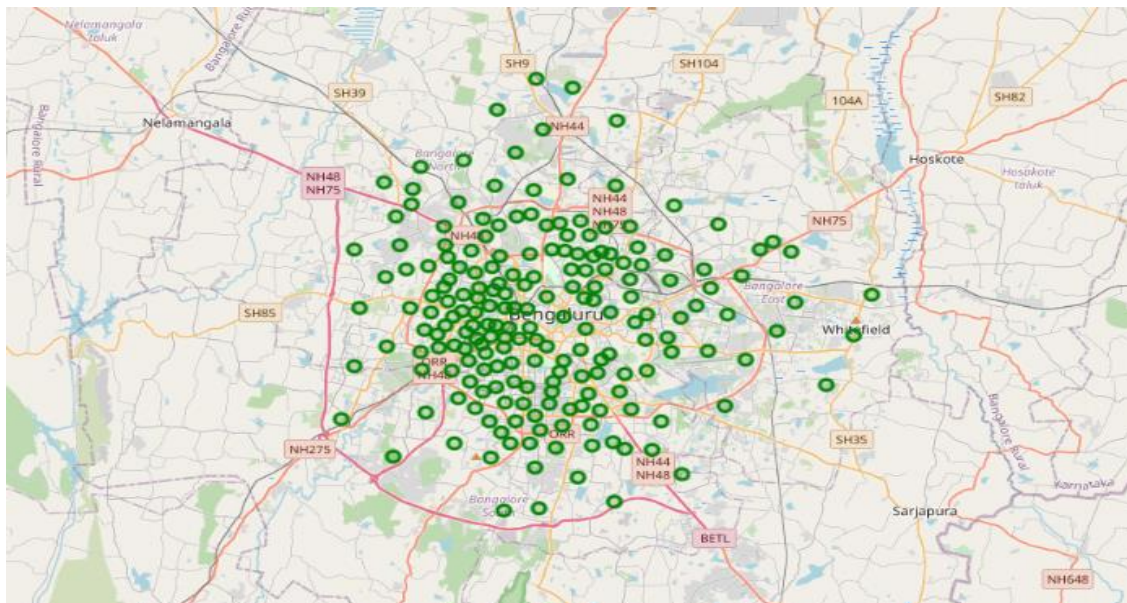


Figure 2. Map of Bengaluru with 198 wards using Folium.

3.3. Venue Extraction:

1. The next process is to extract the venues nearby each neighborhood by using Foursquare API.
2. Nine Venues of different types like Metro, Hospital, Clothing Store, Grocery Store, Bank, Sports Center, Schools, Colleges /University, Movie theaters will be searched.
3. According to a specific order of categories, the minimum distance from the neighborhood location will be extracted.

For example, in Grocery Store, there may be Super Market, Grocery Store, Fruit Market, Farmers Market etc. different categories. Our preferred order will be Super Market, Grocery Store.

4. The minimum distance will be used for clustering analysis.

Unnamed: 0		Neighborhood	Latitude	Longitude	Price	Metro	Hospital	Grocery	School	University	Movie	Clothing	Sports	Bank	Kaveri
0	0	Chowdeswari Ward	13.121710	77.580420	3844.0	5000	1842	3133	525	1007	3427	2817	2408	1272	2
1	1	Atturu	13.102805	77.580038	3295.0	5000	1380	1441	1501	1942	2853	1788	2110	1879	2
2	2	Yelahanka Satellite Town	13.090988	77.583920	3900.0	5000	758	1250	975	807	1271	337	794	524	2
3	3	Vijnanapura	13.008063	77.669580	5840.0	1855	823	744	910	930	1568	727	854	929	0
4	4	Basavanapura	13.016847	77.715454	4000.0	5000	949	682	1065	1132	3026	2189	760	789	0

Figure 3. Minimum Distance Data Frame after extracting from Foursquare API

3.4. Clustering:

1. Clustering will be done through k-means clustering algorithm.
2. The Data's of Price, Proximity, Water Source will be used to cluster the data.
3. All values will be normalized to get good result.
4. Elbow method will be used to determine the optimum number of clusters.
5. All clusters will be analyzed and proper inferences to be made.

3.5. Dashboard:

1. A Jupyter dashboard will be presented where it will contain a choropleth map of price distribution in different neighborhoods, a choropleth map of water availability based on the days in a week, a clustered map obtained due to clustering.
2. There will also be a dropdown menu of neighborhoods, where the end user can select a neighborhood he wants and see the data regarding price, coordinates, venues close by, water source availability and population in the neighborhood.

4. Analysis

4.1 Clustering:

1. Clustering is done with the help of K-Means Clustering method. The data of prices, water availability and proximity will be normalized and then used for clustering.
2. The elbow method will be used to determine the optimum k-value. It is a graph of k-values versus distortions.
3. The optimum k-value at the elbow will be the number of clusters for this project. This number was found to be 6.

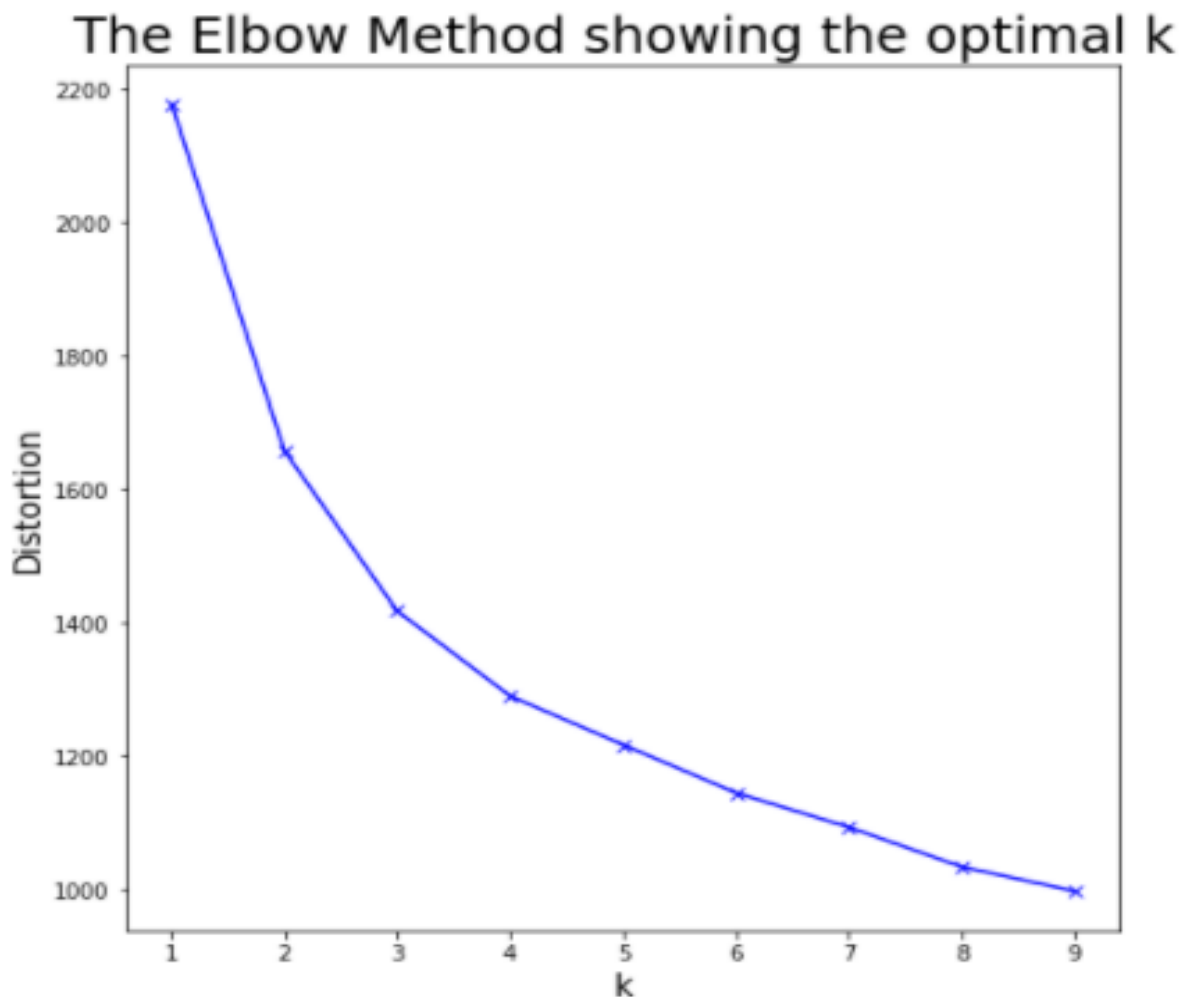


Figure 4. Elbow Method Analysis.

4. The clustered neighborhoods will be mapped using Folium Library.

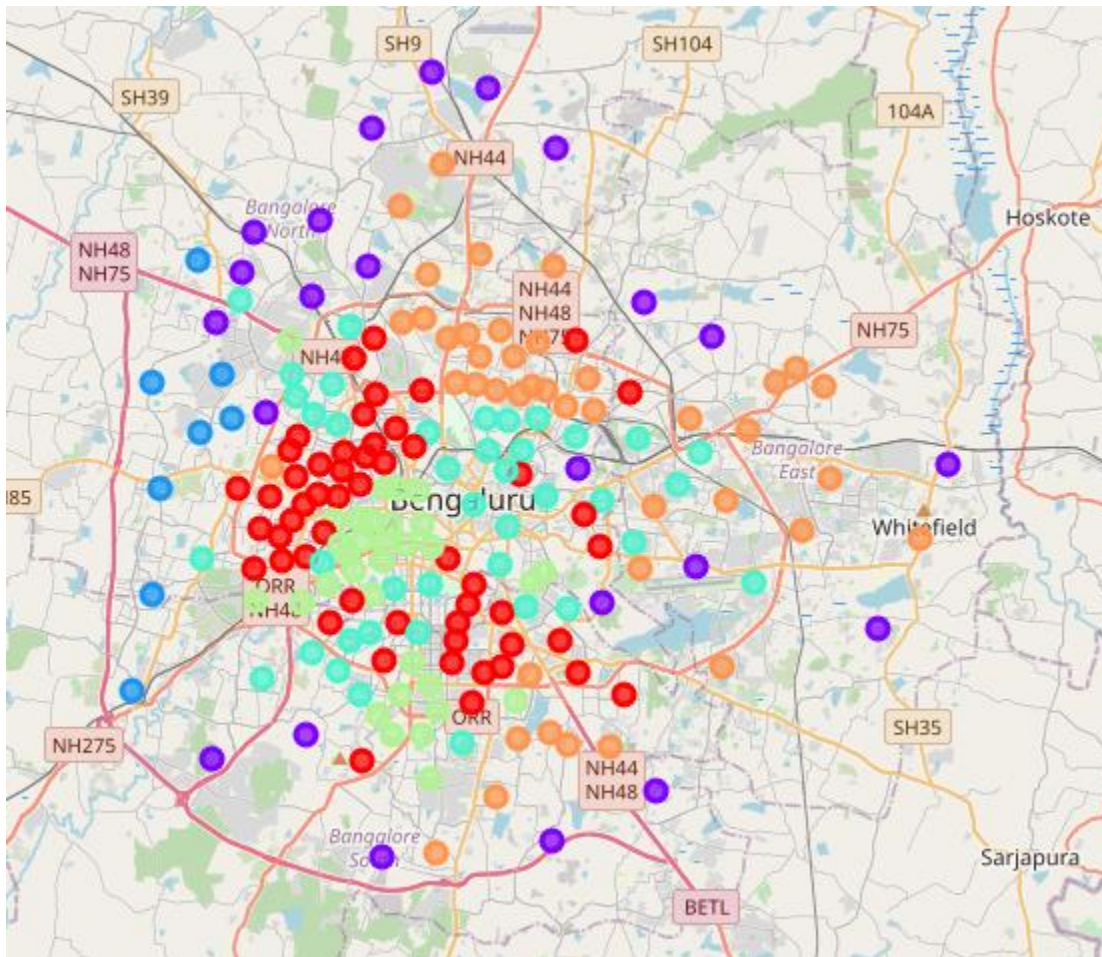


Figure 5. Clustered Map of Wards after K-Means Clustering Analysis.

5. Results and Discussion

5.1 Histogram of prices and their frequency: A histogram chart can be plotted for prices to find the frequency of prices and its occurrence over 198 neighborhoods. Results show that Prices that range between 7000-8000 and 10000-10000 Rupees per square feet had the highest occurrence.

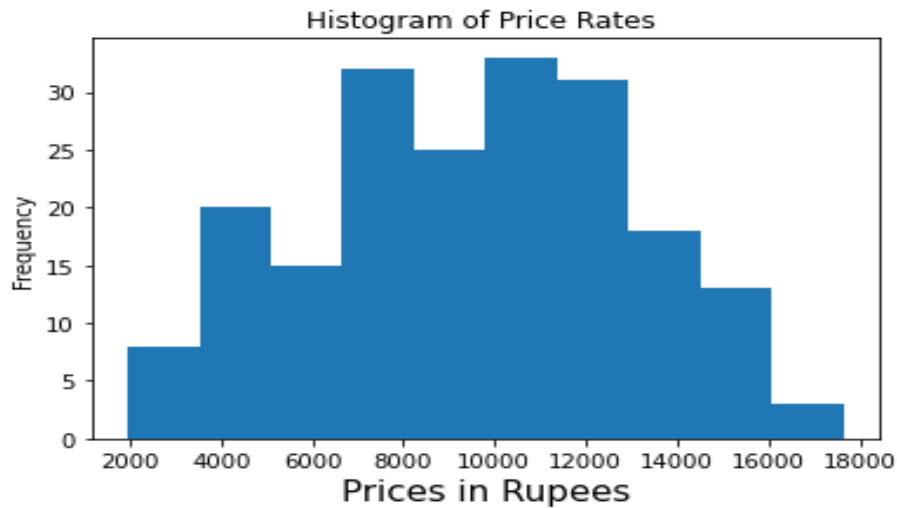


Figure 6. Histogram of Price v/s Frequency.

5.2 Box Chart of Average Mean Price: A box chart can be used to visualize the mean price per square feet in Bengaluru. The mean Price was found to be close to Rs.10000/ Sq. Ft.

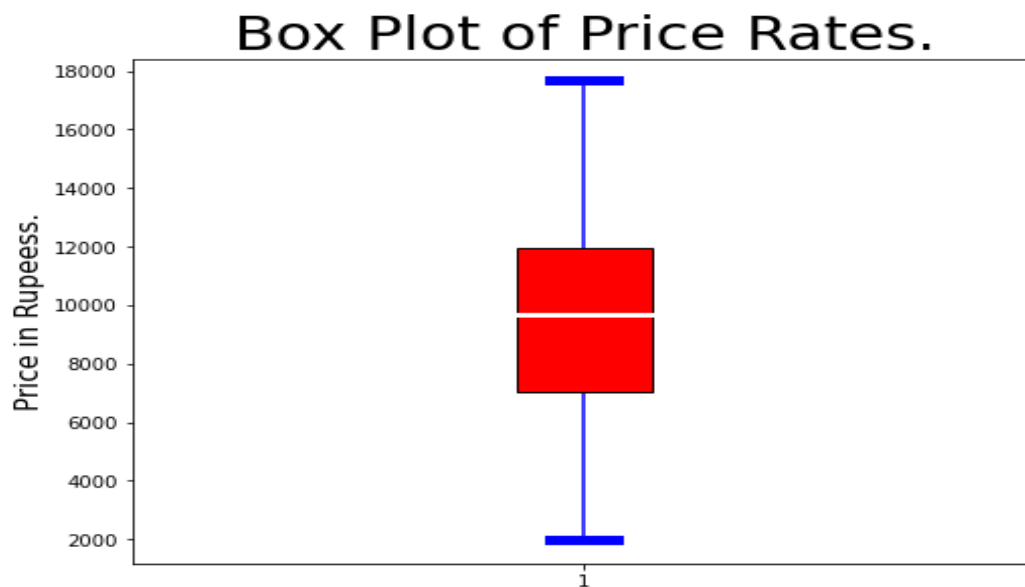


Figure 7. Box Plot to find the median Price over 198 wards.

5.3 Choropleth Maps of Price Distribution and Kaveri Water Availability in different wards:

A choropleth map is a good way to visualize the price distribution and Kaveri water Facility in all the 198 wards. Results show that the heart of the city has high price and as we move away, the price becomes lower.

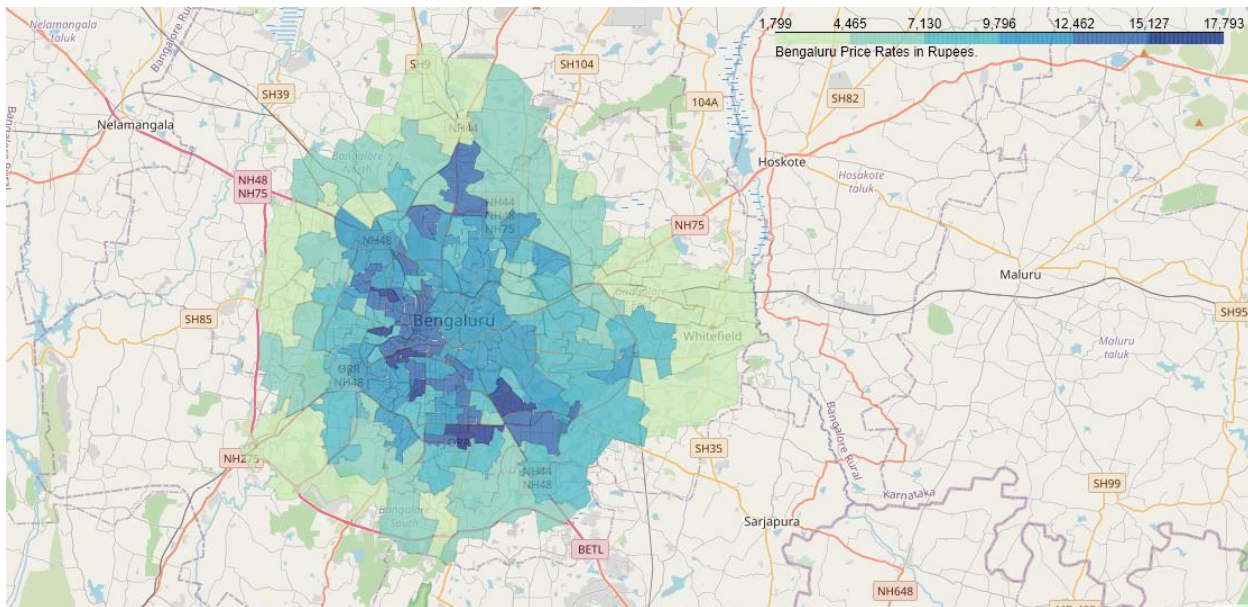


Figure 8. Choropleth Map of Price Distribution.

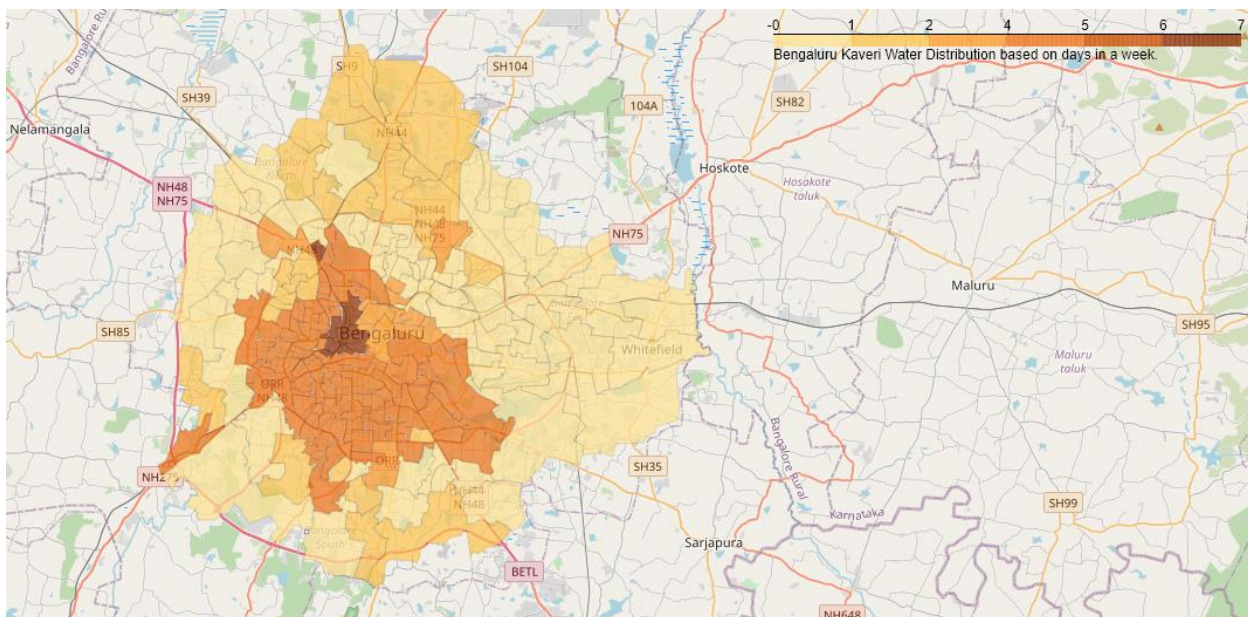


Figure 9. Choropleth Map of Kaveri Water Availability in a Week on the basis of days.

5.4 Dashboard:

A dashboard is a good way to present the end user to display the detail of a neighborhood of his selection. The dashboard made in this project contains a dropdown menu of neighborhoods, where the end user can select a neighborhood he wants and see the data regarding price, coordinates, venues close by, water source availability and population in the neighborhood. It also displays the two-choropleth map and the clustered map.

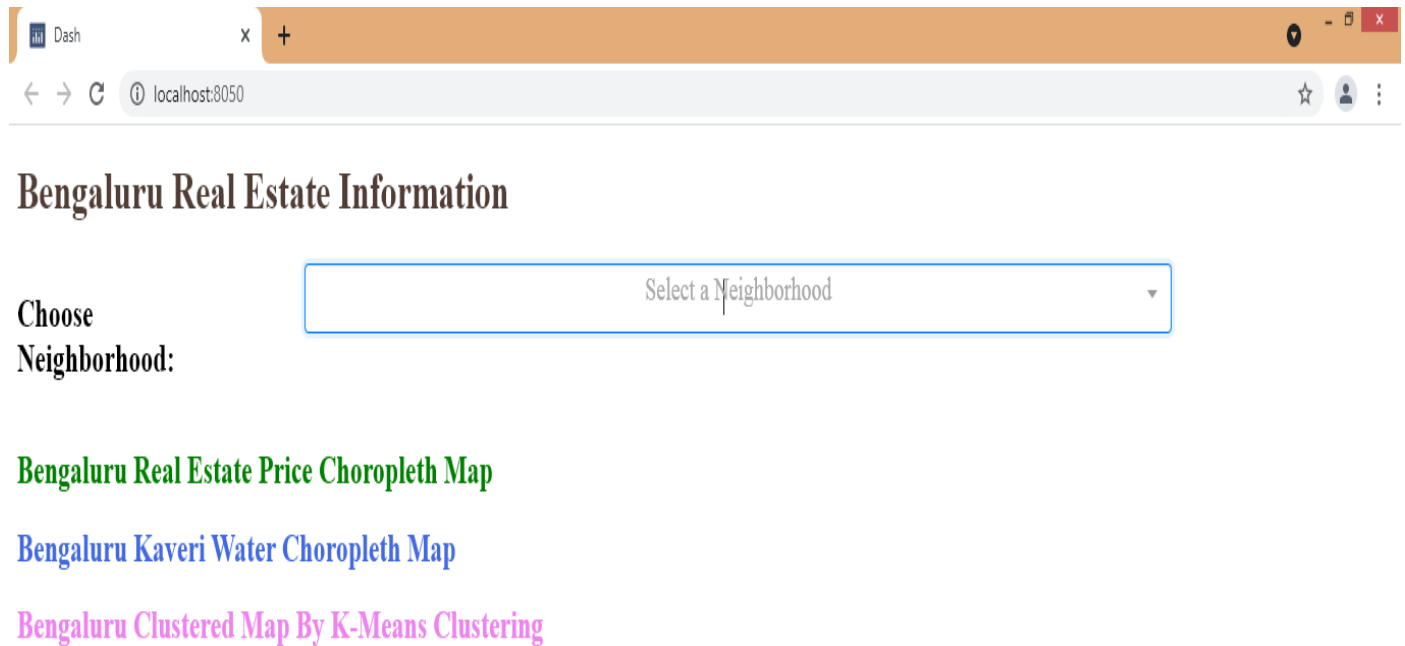


Figure 10. Initial Layout of the Dash Application.

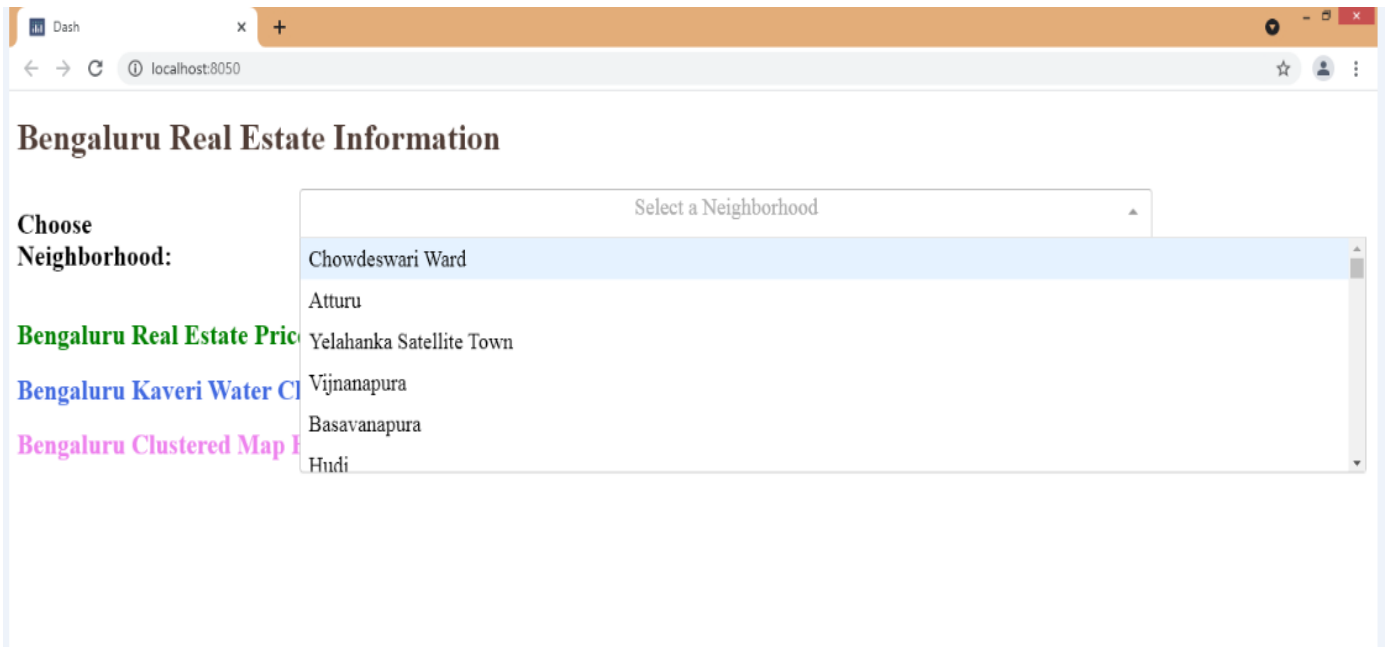
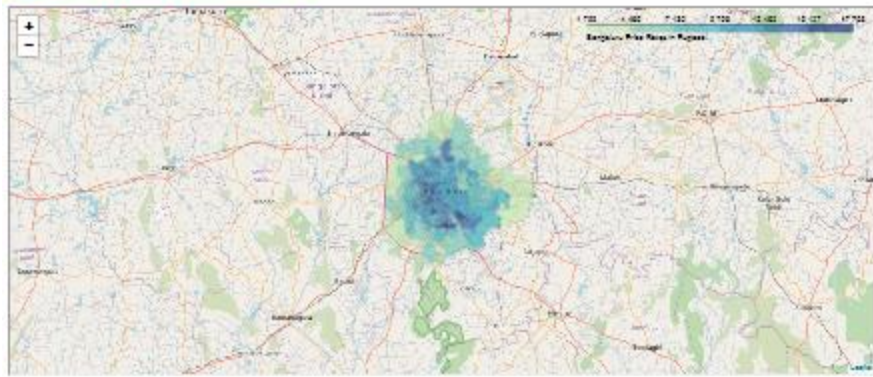


Figure 11. User Selecting the Neighborhood Dropdown Menu

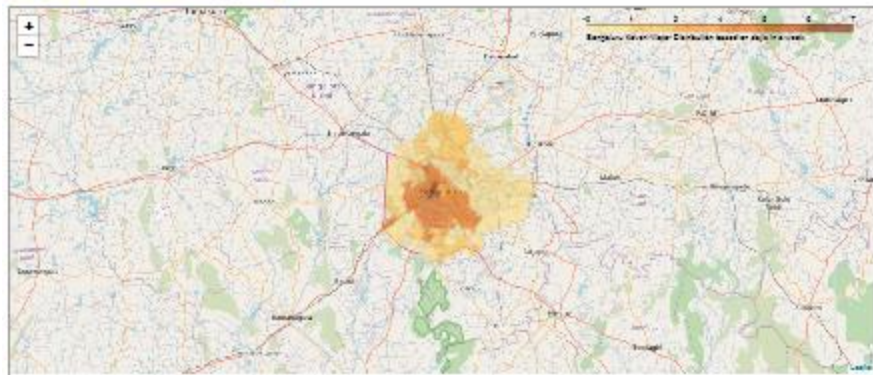
Neighborhood	Price in Rs per SQ.FT.	Metro Closeby	Hospital Closeby	Grocery Closeby	Schools Closeby	Movie Theaters Closeby	Clothing Store Closeby	University Closeby	Sports Center Closeby	Bank Closeby	Kaveri Water Availability (No of days in a week)	Population
Katriguppe	11233.65	Deepanjali Nagar Metro Station (Metro Station) Proximity: 2825 mts	Diavista Multi Speciality Hospital (Hospital) Proximity: 647 mts	Nilgiri's (Supermarket) Proximity: 3599 mts	Jumbo Kids (Nursery School) Proximity: 701 mts	Screen 2, Gopalan cinemas (Multiplex) Proximity: 1112 mts	INMARK (Clothing Store) Proximity: 506 mts	Indian Institute Of Foreign Languages (University) Proximity: 1031 mts	Revive Fitness Hub (Gym / Fitness Center) Proximity: 526 mts	Andra Bank ATM (Bank) Proximity: 195 mts	4	35736

Figure 12. Table displaying all the necessary details about the venue in a table.

Bengaluru Real Estate Price Choropleth Map



Bengaluru Kaveri Water Choropleth Map



Bengaluru Clustered Map By K-Means Clustering

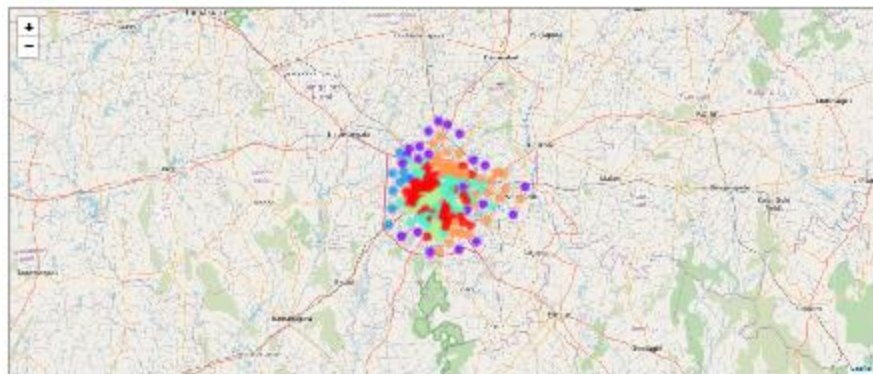


Figure 13. Other Results Displayed.

5.5 Analysis of Clusters:

The Analysis of clusters is done by looking at the group of clusters and seeing what's common among the clusters. The following are the details obtained from analysis of clusters:

Cluster 1:

	Neighborhood	Latitude	Longitude	Price	Metro	Hospital	Grocery	School	University	Movie	Clothing	Sports	Bank	Kaveri	Labels
24	Yeshwanthpura	13.026029	77.553856	10678.00	2730	584	1753	1168	1264	593	587	342	538	7	0
28	Kottegepalya	12.982456	77.514090	8478.55	2839	371	1559	854	924	1874	970	2066	274	4	0
42	Shakthi Ganapathi Nagar	12.995048	77.531586	12111.20	2748	219	935	843	1056	1610	840	278	939	2	0
43	Shankar Matt	12.999985	77.534410	13151.24	3097	543	1312	789	1432	1107	920	350	906	2	0
44	Aramane Nagara	13.015349	77.577020	10875.00	2766	1274	1363	779	1121	2384	1934	726	954	4	0
45	Mattikere	13.032590	77.561035	7480.00	3120	252	845	860	346	1631	749	202	199	4	0
46	Malleswaram	13.014073	77.561676	12470.00	2734	588	566	412	703	747	795	588	647	4	0
48	Kadu Malleshwar Ward	13.002385	77.568490	9568.00	1270	384	1492	126	584	1032	855	291	165	4	0
49	Subramanya Nagar	13.007040	77.557106	10546.00	2312	308	416	186	761	387	275	494	178	4	0
50	Gayithri Nagar	12.997437	77.560520	9986.00	1323	526	1511	559	994	1164	1290	1028	800	4	0
67	HBR Layout	13.032021	77.630130	12230.44	5106	412	1551	43	791	1513	796	400	516	3	0
68	Banasavadi	13.014699	77.648926	11254.60	2378	174	934	402	662	2291	618	257	314	2	0
85	Bharathi Nagar	12.987015	77.611040	10540.00	1363	739	963	458	788	477	532	1038	404	3	0
89	Domlur	12.963362	77.638040	10372.00	1660	309	334	383	323	499	480	237	246	4	0
94	Jogupalya	12.973724	77.632590	11974.00	724	429	1339	243	1390	388	639	739	170	4	0
98	Dattatreya Temple	12.996555	77.574135	13570.00	689	459	2378	580	583	546	57	605	379	7	0
102	Basaveshwara Nagar	12.990808	77.542130	13111.25	1876	255	621	64	423	1330	400	562	501	4	0
103	Kamakshipalya	12.986797	77.533875	6671.00	1827	409	379	732	394	2087	301	293	476	4	0

```
The size of dataframe is : (53, 15)
The Average Price of Cluster 1 is: 11380.3
Average Distance Metro: 1906.4
Average Distance Hospital: 441.5
Average Distance Grocery Store: 1065.9
Average Distance School: 508.2
Average Distance University: 941.4
Average Distance Movie: 1320.5
Average Distance Clothing: 848.1
Average Distance Sports Center: 785.8
Average Distance Bank: 448.0
```

- The average Price of Cluster 1 is Rs 11380.3.
- These wards have Average metro proximity of 2 Kms.
- Hospitals, Schools, Sports Center, Clothing Store, Grocery Store, University and Banks have average proximity of less than 1 kms.
- Movie Theaters have an average proximity of 1.3 kms.
- These areas get Kaveri water 4 days in a week.
- Cluster 1 has 53 Neighborhoods

Cluster 2:

	Neighborhood	Latitude	Longitude	Price	Metro	Hospital	Grocery	School	University	Movie	Clothing	Sports	Bank	Kaveri	Labels
0	Chowdeswari Ward	13.121710	77.580420	3644.00	5000	1842	3133	525	1007	3427	2617	2408	1272	2	1
1	Atturu	13.102805	77.560036	3295.00	5000	1380	1441	1501	1942	2653	1788	2110	1879	2	1
9	HAL Airport	12.956536	77.671500	8783.00	4308	734	1619	1237	1853	2403	1602	633	1328	0	1
10	Dodda Bommasandra	13.056595	77.558690	7500.00	3292	824	745	1177	1777	1420	2279	1541	873	2	1
11	Kuvempu Nagar	13.072080	77.541870	6956.28	3657	1501	1474	2066	1283	393	1885	1437	1284	2	1
12	Jakkuru	13.096251	77.623314	7118.00	5000	1900	2337	1647	2586	2992	3019	2410	1856	2	1
21	Jalahalli	13.046444	77.539340	3931.30	943	1119	1476	1358	1863	2395	2273	1261	1035	1	1
27	Laggere	13.007688	77.523900	4951.20	3672	670	2565	1524	2688	2281	2174	293	1025	1	1
29	Shettihalli	13.067997	77.519410	4875.63	3092	697	744	3000	1738	3074	3017	1200	1220	1	1
30	Mallasandra	13.054437	77.515130	4258.26	1619	1299	2234	2753	1853	1595	1517	1989	853	1	1
33	Chokkasandra	13.038051	77.506310	3989.90	1329	899	2623	2813	1364	1309	1236	993	1184	1	1
78	Sarvagna Nagar	12.989170	77.631040	7015.00	1466	1404	2843	1282	1982	1375	1399	1391	979	1	1
90	Agaram	12.944263	77.639046	11545.00	3785	1792	1874	1586	1723	2524	1367	891	778	4	1
164	Varthuru	12.935373	77.733610	2146.89	5000	2360	743	910	3127	3110	1289	2480	936	0	1
167	Kadugodi	12.990687	77.757910	2152.00	5000	1378	1011	1069	1152	2500	2321	895	795	0	1
177	Singasandra	12.881166	77.657650	7445.00	5000	1456	1381	1356	1090	2115	1708	737	1062	1	1
179	Uttarahalli	12.900265	77.537080	6912.11	2693	838	1113	1482	2846	3009	1017	1332	975	2	1
181	Begur	12.865000	77.622070	5362.67	5000	1926	1221	2112	2863	3125	2882	1369	1004	2	1
184	Anjanapura	12.859588	77.563286	5149.44	2801	900	1707	2706	2864	3519	3048	1750	1169	0	1
185	Kempegowda Ward	13.116188	77.599720	3876.00	5000	733	2487	1250	697	2466	2501	1404	1420	2	1
194	Ramamurthy Nagar	13.033613	77.676540	6031.00	4652	1453	1419	1330	3408	1616	2209	2286	888	0	1
195	Horamavu	13.044561	77.653275	5258.89	5483	974	1420	1008	1901	2873	2498	1469	931	0	1
197	Hemmigepura	12.891903	77.505010	2890.00	4352	1473	2875	1328	1292	2886	2156	2683	1479	1	1

```
The size of dataframe is : (23, 15)
The Average Price of Cluster 2 is: 5438.55
Average Distance Metro: 3788.87
Average Distance Hospital: 1284.87
Average Distance Grocery Store: 1760.22
Average Distance School: 1609.57
Average Distance University: 1952.13
Average Distance Movie: 2393.91
Average Distance Clothing: 2078.35
Average Distance Sports Center: 1520.09
Average Distance Bank: 1140.22
```

- The average Price of Cluster 2 is Rs. 5438.55
- These wards have Average Metro Proximity of 3.8 Kms.
- Banks and Hospitals have an average proximity of 1.2 Kms.
- Other venues have average proximity between 1.6 - 2.3 Kms.
- These areas get Kaveri water in a week for less than 2 days in a week.
- Cluster 2 has 23 Neighborhoods.

Cluster 3:

	Neighborhood	Latitude	Longitude	Price	Metro	Hospital	Grocery	School	University	Movie	Clothing	Sports	Bank	Kaveri	Labels
17	Ullalu	12.946716	77.48462	5523.45	4162	2907	2445	2023	923	842	3238	2801	491	2	2
18	Kengeri	12.915069	77.47752	3381.67	4680	2532	1495	3000	435	202	3000	1754	427	4	2
19	Dodda Bidarakallu	13.018076	77.48454	4235.00	4420	3000	3000	3000	1818	2251	3525	3000	2632	1	2
20	Herohalli	12.982807	77.48716	3896.55	5606	845	2566	2638	777	3225	2288	2382	867	1	2
31	Bagalakunte	13.058892	77.50028	4696.30	2856	1279	3039	3000	922	2825	2718	1608	739	1	2
34	Peenya Industrial Area	13.020902	77.50861	5180.11	2398	2499	3206	3065	1463	2404	2425	1955	14	1	2
35	Rajagopal Nagar	13.006200	77.51225	6234.14	3846	1925	3327	2763	1479	3495	2055	1227	757	1	2
36	Hegganahalli	13.001005	77.50098	3863.00	4756	2489	2585	3043	792	2034	1577	2577	383	1	2

```
The size of dataframe is : (8, 15)
The Average Price of Cluster 3 is: 4626.3
Average Distance Metro: 4090.5
Average Distance Hospital: 2184.5
Average Distance Grocery Store: 2707.9
Average Distance School: 2816.5
Average Distance University: 1076.1
Average Distance Movie: 2159.8
Average Distance Clothing: 2603.2
Average Distance Sports Center: 2163.0
Average Distance Bank: 788.8
```

- The average Price of Cluster 3 is Rs. 4626.3.
- These wards have Metro proximity greater than 4.0 Kms.
- Most Facilities are situated far away by more than 2.0 Kms.
- Only Banks and University are within 1 Kms.
- These areas get Kaveri water in a week for less than 2 days in a week.
- These neighborhoods are far from metro and city.

Cluster 4:

	Neighborhood	Latitude	Longitude	Price	Metro	Hospital	Grocery	School	University	Movie	Clothing	Sports	Bank	Kaveri	Labels
7	A Narayanapura	12.994474	77.672585	5100.00	1507	610	1904	1090	536	843	950	899	647	0	3
22	Rajarajeshwari Nagar	12.918757	77.522140	8796.56	1192	821	1673	757	1875	1192	250	206	497	1	3
23	J P Park	13.036629	77.552420	8247.55	2106	474	1507	1325	587	1445	1430	560	354	1	3
26	Lakshmi Devi Nagar	13.020603	77.532470	10233.00	2193	1176	2951	892	2074	1807	1811	916	29	1	3
32	T Dasarahalli	13.045358	77.514730	3566.67	693	326	3177	2154	1578	663	563	1609	640	1	3
38	Nandini Layout	13.013450	77.534140	10023.65	2988	485	2576	263	2284	1634	1912	295	257	2	3
39	Marappana Palya	13.017480	77.546240	10285.00	2897	874	1454	318	1062	790	804	695	484	2	3
40	Nagapura	13.004127	77.548134	9856.20	2858	217	1274	405	1305	453	784	314	230	2	3
41	Mahalakshimpuram	13.007350	77.540040	13414.89	3730	737	1955	157	1786	701	1317	424	302	2	3
47	Rajamahall Guttahalli	13.001372	77.579060	11456.85	1434	376	2537	741	1676	1112	713	560	335	4	3
63	Sagayarapuram	13.005577	77.616870	11912.00	3438	239	2651	306	1373	1584	925	334	251	1	3
64	S K Garden	13.005327	77.607056	11370.00	2944	478	2119	915	1384	1126	1195	301	317	1	3
65	Pulikeshinagar	12.995343	77.611855	12201.00	2247	51	3342	240	243	396	390	111	635	1	3
73	Maruthi Seva Nagar	12.999921	77.630080	7092.00	2582	1432	2536	1075	1602	322	348	285	316	0	3
75	Benniganahalli	12.999501	77.651700	7793.00	878	665	2474	1210	712	1508	1101	1096	1050	1	3
76	C V Raman Nagar	12.983950	77.665200	7579.00	1473	1218	1262	234	1037	922	734	273	984	1	3
79	Hoysala Nagar	12.978457	77.638790	7205.00	27	111	1432	74	716	526	593	228	95	1	3
80	Jeevanbhima Nagar	12.964652	77.650030	7603.00	1955	599	878	426	752	658	795	406	457	1	3
81	Sampangiram Nagar	12.976794	77.595375	10351.00	449	703	1844	692	728	1285	1326	487	281	3	3
82	Ramaswamy Palya	13.005976	77.599075	10319.00	2845	562	1929	915	1167	1386	1744	602	594	3	3

```

The size of dataframe is : (40, 15)
The Average Price of Cluster 4 is: 9760.9
Average Distance Metro: 2011.0
Average Distance Hospital: 614.0
Average Distance Grocery Store: 2519.7
Average Distance School: 655.1
Average Distance University: 1088.8
Average Distance Movie: 1009.9
Average Distance Clothing: 852.8
Average Distance Sports Center: 538.2
Average Distance Bank: 433.6

```

- The average Price of Cluster 4 is Rs. 9760.9.
- These wards have average Metro proximity of 2.0 Kms.
- All Facilities are having an average proximity of 1.0 Kms.
- These areas get Kaveri water in a week for less than 2 days in a week.
- These neighborhoods are within the heart of the city.
- Cluster 4 has 40 Neighborhoods.

Cluster 5:

	Neighborhood	Latitude	Longitude	Price	Metro	Hospital	Grocery	School	University	Movie	Clothing	Sports	Bank	Kaveri	Labels
25	HMT Ward	13.031905	77.531710	11563.25	950	919	3291	470	1614	1749	1858	1793	787	3	4
91	Vannarpet	12.954209	77.618940	11854.00	2093	689	2023	186	2032	1503	1989	1488	421	4	4
92	Nilasandra	12.951159	77.614650	10504.00	2442	266	2585	124	1650	1679	1619	1307	913	4	4
95	Chickpete	12.970688	77.577095	14851.00	479	438	3321	634	1100	367	178	1390	178	7	4
96	Cottonpete	12.970359	77.566284	13646.00	615	578	2428	768	1748	696	951	1623	433	7	4
97	Binnipete	12.971799	77.559100	13503.00	517	737	1838	755	1943	728	1746	1788	1216	7	4
99	Gandhinagar	12.981565	77.575670	14840.00	756	706	2727	934	1409	1146	944	1677	196	7	4
100	Subhash Nagar	12.982667	77.568630	13973.00	790	1092	1954	1150	1191	596	1021	1808	685	7	4
101	Okalipuram	12.982783	77.563820	14084.00	796	692	1444	692	1344	625	1250	1918	1093	7	4
117	Nayandahalli	12.945141	77.519960	8542.10	1146	1422	2664	935	1856	1820	1054	948	1072	4	4
118	Kempapura Agrahara	12.972313	77.555480	8956.20	377	736	1631	600	2277	447	2142	1458	1540	4	4
119	Vijayanagar	12.972093	77.547040	15346.00	310	369	1639	734	879	526	2377	945	633	4	4
124	Gali Anjenaya Temple ward	12.950159	77.544655	10654.20	870	1040	3152	854	1024	1322	2349	472	773	4	4
125	Deepanjali Nagar	12.944540	77.536230	11203.00	677	102	3187	473	524	2424	1884	1129	822	4	4
126	Padarayanapura	12.963702	77.548810	14630.00	1234	934	2527	1219	2554	816	1441	1082	755	4	4
127	Jagajivanaramnagar	12.960545	77.551940	14721.00	1689	1406	2869	1489	2177	329	1916	960	367	4	4
128	Rayapuram	12.964988	77.556380	14850.00	1190	1335	2443	1411	1793	600	1878	964	1021	4	4
129	Chalavadipalya	12.964580	77.564445	14100.00	1159	560	2824	536	2567	314	1211	1898	931	4	4
130	K R Market	12.963968	77.571800	14502.00	470	336	3357	264	2017	858	607	1719	506	4	4
131	Chamrajapet	12.958184	77.563774	15370.00	1243	749	3423	479	2491	876	775	1281	847	4	4
132	Azad Nagar	12.954717	77.554020	15163.00	1878	1406	3531	988	2054	381	1611	1756	422	4	4
134	Dharmaraya Swamy Temple	12.963133	77.580510	12996.00	665	648	2191	677	1465	1176	747	1423	488	4	4
140	Hanumanth Nagar	12.947122	77.559880	11465.35	1474	812	3618	667	2199	1396	1182	900	320	4	4

```

The size of dataframe is : (32, 15)
The Average Price of Cluster 5 is: 12233.6
Average Distance Metro: 1068.7
Average Distance Hospital: 689.2
Average Distance Grocery Store: 2607.9
Average Distance School: 685.9
Average Distance University: 1633.6
Average Distance Movie: 1182.2
Average Distance Clothing: 1319.9
Average Distance Sports Center: 1725.3
Average Distance Bank: 662.2

```

- The average Price of Cluster 5 is Rs. 12233.6.
- These wards have average Metro proximity of 1.0 Kms.
- All Facilities are having an average proximity of 1.2 Kms.
- These areas get Kaveri water in a week for more than 4 days in a week.
- These neighborhoods are within the heart of the city.
- Cluster 5 has 32 Neighborhoods.

Cluster 6:

	Neighborhood	Latitude	Longitude	Price	Metro	Hospital	Grocery	School	University	Movie	Clothing	Sports	Bank	Kaveri	Labels
2	Yelahanka Satellite Town	13.090986	77.583920	3900.00	5000	758	1250	975	807	1271	337	794	524	2	5
3	Vijnanapura	13.006063	77.669560	5840.00	1655	623	744	910	930	1568	727	854	929	0	5
4	Basavanapura	13.016847	77.715454	4000.00	5000	949	682	1065	1132	3026	2189	760	789	0	5
5	Hudi	13.022375	77.705490	3484.78	5883	271	570	636	153	2958	1879	195	327	0	5
6	Devasandra	13.001798	77.689125	3986.00	3368	1009	1518	806	1337	835	698	1002	945	0	5
8	Vijnana Nagar	12.978493	77.681770	4500.00	3140	434	766	598	1220	1303	1002	520	216	0	5
13	Thanisandra	13.056525	77.622410	6896.89	5000	1222	2104	1098	885	1316	1322	1420	690	2	5
14	Byatarayanapura	13.061070	77.597374	12672.56	5000	480	514	564	2179	500	631	737	649	2	5
15	Kodigehalli	13.053845	77.579090	7118.70	5183	1102	635	795	1954	726	1451	329	960	2	5
16	Vidyaranyapura	13.077092	77.569450	6235.22	5000	1463	1590	1253	1336	3385	1188	1027	514	1	5
37	Vrisabhavathi Nagar	12.989773	77.525880	9863.45	2456	1031	1207	879	1280	2451	981	989	1072	2	5
51	Radhakrishna Temple Ward	13.038305	77.570570	12481.00	4040	401	662	245	1032	2844	560	522	143	1	5
52	Sanjaya Nagar	13.039301	77.577800	14135.89	4819	409	936	494	1723	513	911	351	299	1	5
53	Ganga Nagar	13.032555	77.586360	10002.00	4897	474	1256	738	873	1270	764	884	166	1	5
54	Hebbala	13.034055	77.593020	9140.22	5337	379	1449	424	548	3862	72	938	589	1	5
55	Vishwanath Nagenahalli	13.035174	77.603740	10886.00	6051	1233	1372	224	174	2736	1108	348	1079	1	5
56	Manorayanapalya	13.026743	77.597305	9263.00	4881	238	519	528	479	3802	588	342	256	1	5
57	Gangenahalli	13.017775	77.588810	9223.00	3536	823	1406	530	1297	2660	449	321	417	1	5
58	Jayachamarajendra Nagar	13.017172	77.595980	9610.00	3967	182	835	673	1302	2651	159	1682	331	1	5
59	Kushal Nagar	13.016274	77.615020	11799.00	4401	438	1757	445	1302	2523	633	1404	1452	1	5
60	Kaval Bairasandra	13.026351	77.608570	10970.00	5241	662	941	730	955	2883	689	1047	456	1	5

```
The size of dataframe is : (42, 15)
The Average Price of Cluster 6 is: 7600.3
Average Distance Metro: 4367.9
Average Distance Hospital: 686.9
Average Distance Grocery Store: 1158.6
Average Distance School: 623.5
Average Distance University: 1127.4
Average Distance Movie: 1880.4
Average Distance Clothing: 922.1
Average Distance Sports Center: 898.9
Average Distance Bank: 640.5
```

- The average Price of Cluster 6 is Rs. 7600.3
- These wards have Metro with average proximity of 4.0 Kms.
- Facilities like Hospitals schools, clothing store, bank, sports center, university have average proximity of 1.2 Kms.
- These areas get Kaveri water in a week for less than 2-0 days in a week.
- These neighborhoods are far from metro and city.
- Cluster 6 has 42 Neighborhoods.

6. Conclusions

In this project, The Real Estate Data analysis of Bengaluru was made by utilizing data from various sources. Initial raw data was cleaned and venues details were extracted by using Foursquare API. Then the data was clustered by using K-Means Clustering Method. Six groups of clusters were obtained. The average price per square feet was near 10000 Rupees. Choropleth Maps of Price Distribution and Kaveri Water Availability was Mapped. Finally, a Dashboard was created where the end user can select a neighborhood and the details of the neighborhood is shown and important maps displayed.

The Clustered Data can be grouped as:

1. Cluster 1:

Average Price: Rs. 11000	Rate: 4.0/5
Kaveri Water in a week: 4 days	Rate: 4.0/5
Excellent Facilities Nearby.	Rate: 4.5/5
Far from Metro	Rate: 3.5/5

2. Cluster 2:

Average Price: Rs. 5400	Rate: 2.5/5
Kaveri Water in a week: 2-1 days	Rate: 2.0/5
Facilities are quite far.	Rate: 3.0/5
Very Far from Metro.	Rate: 1.0/5

3. Cluster 3:

Average Price: Rs. 4600	Rate: 2.0/5
Kaveri Water in a week: 2-0 days	Rate: 2.0/5
Facilities are quite far away	Rate: 3.0/5
Very Far from Metro.	Rate: 1.0/5

4. Cluster 4:

Average Price: Rs. 9700	Rate: 3.75/5
Kaveri Water in a week: 3-1 days	Rate: 3.0/5
Excellent Facilities Nearby.	Rate: 2.5/5
Metro is quite far.	Rate: 3.0/5

5. Cluster 5:

Average Price: Rs. 12200	Rate: 4.0/5
Kaveri Water in a week: 4 days	Rate: 4.0/5
Excellent Facilities Nearby.	Rate: 4.5/5
Very Near to Metro.	Rate: 4.5/5

6. Cluster 6:

Average Price: Rs. 7600	Rate: 3.5/5
Kaveri Water in a week: 2-0 days	Rate: 2.2/5
Very Good Facilities Nearby.	Rate: 4.0/5
Very Far from Metro.	Rate: 1.0/5

Ratings Note:

1. Price: 5-high to 0-low
2. Water: 5-Excellent supply to 0-No Supply
3. Facilities: 5- Very Near to 0- Very Far
4. Metro: 5- Very Near to 0- Very Far

- Cluster 1 has is a very good facilities and good water source supply in a week. The only draw back is that the price is quite high. These neighborhoods are mostly in the western and southern part of the city.
- Cluster 2 has facilities situated quite far and water supply of less than 2 days in a week. The advantage of these cluster is that price is quite affordable. These neighborhoods are situated in the outskirts of the city.
- Cluster 3 has facilities situated quite far and water supply of less than 2 days in a week. The advantage of these cluster is that price is quite affordable. These neighborhoods are situated in the outskirts of western part of the city.
- Cluster 4 good facilities nearby and good source of water supply of 3-1 days in a week. The price is the average price of cities in Bengaluru. These neighborhoods are distributed all over the city i.e. within the heart of the city.
- Cluster 5 has excellent facilities and excellent water source. The price is the highest among all clusters. These neighborhoods are in the heart of the city.

- Cluster 6 has good facilities but situated far from metro. The water source is average with less than 2 weeks supply within a week. The price is also quite affordable. These neighborhoods are situated in the north-eastern, eastern and south eastern part of the city.

7. References

1. Land Prices: <https://www.99acres.com/property-rates-and-price-trends-in-bangalore>.
2. Land prices: <https://www.makaan.com/price-trends/property-rates-for-buy-in-bangalore>
3. Land Prices:
https://housing.com/price-trends/property-rates-for-buy-in-bangalore_karnataka-P38f9yfbk7p3m2h1f
4. Land prices: <https://www.magicbricks.com/property-for-sale-rent-in-Bangalore/residential-real-estate-Bangalore>
5. Bengaluru Wards Geojson:
https://raw.githubusercontent.com/datameet/Municipal_Spatial_Data/master/Bangalore/BBMP.GeoJSON
6. Bengaluru Kaveri Water Supply: https://www.bwssb.gov.in/watter_supply.php
7. Venues Data: Foursquare API
8. Bengaluru wards data: <https://opencity.in/cc2b165a-d44a-49d7-b077-4bc6123ca6ab>