Bangalore_Capstone_project

Coursera_capstone M5(week5)

- Anil Kumar Naidu
16th May 2020

Why Bengaluru for investments?

- As per trends, City will be fastest growing with GDP of 8.5%
- Silicon city of India
- Huge talent force
- Growth will be definite

Problem:

1.) It is difficult to choose given so many parameters to consider while investing

2.) If some area is missed, there will be huge future loss of profit.

- The problem is further diversified to **Major possible investment strategies**:

1.) Residential investments

2.) Fast growth and return investments

3.) Small businesses

Data Description:

- Data required for this method would be Residential and commercial real-estate rates along with all the localities of Bangalore.

- Data collected by webscrapping: https://bengaluru.citizenmatters.in/615-real-estate-rates-615

How Data can be used to solve this problem?

- Data covers complete Bengaluru

- Residential and Commercial real-estate rates split up correctly

- Using the data, venues near the localities can be found

- By training K-Means algorithm, most accurate results can be obtained

Visualizing Data
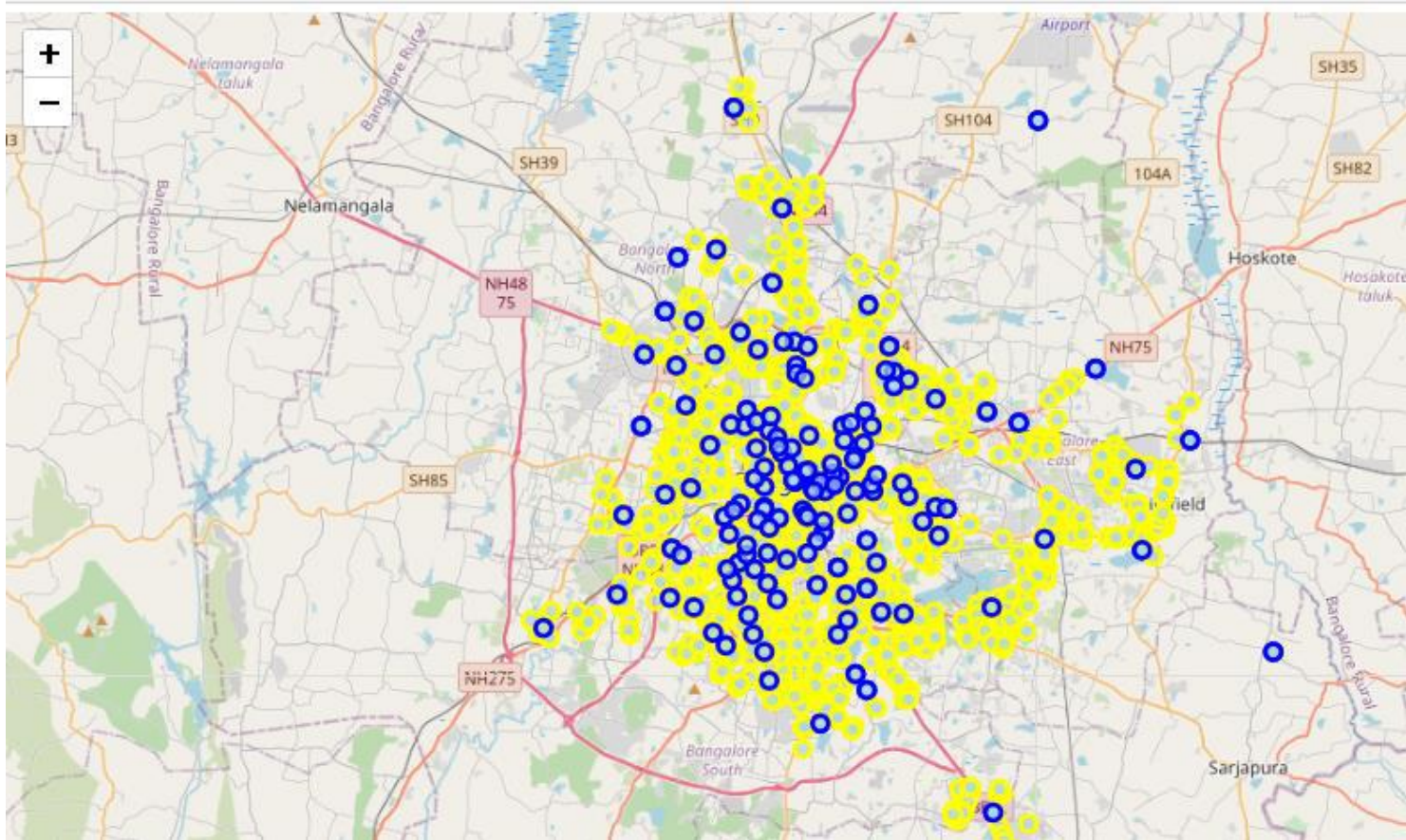
Blue markers show locality,
While yellow maked show the
Venues near the localities.

Why K-Means Classification?

- Best suited for large datasets

- Fast results

- Simple to implement

- Accurate results

- Possibility of changing Kclusters

- Kclusters =5 showed best classification

Classification output

- Figure shows various clusters:
- The proper spread of clusters shows the wide coverage of data and algorithm.
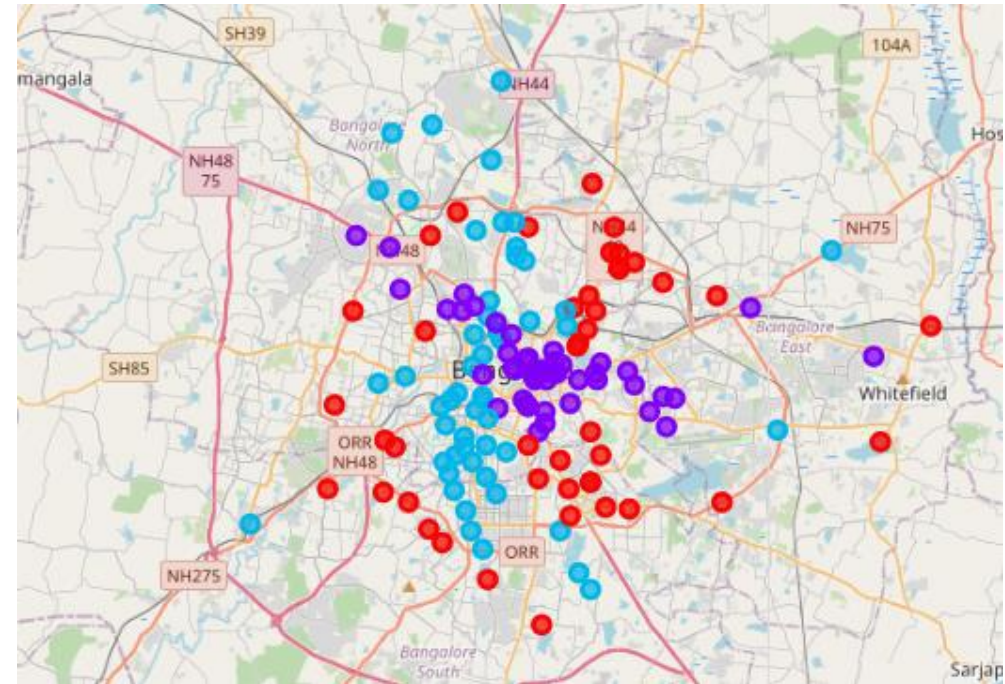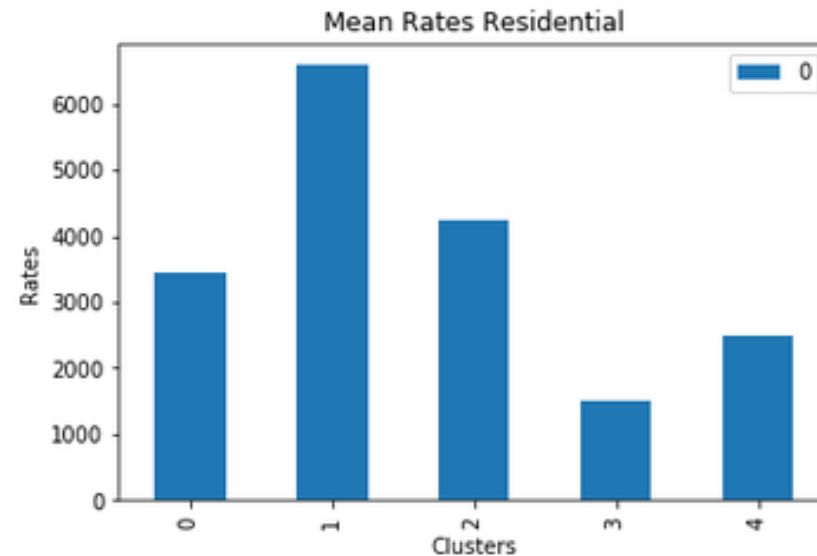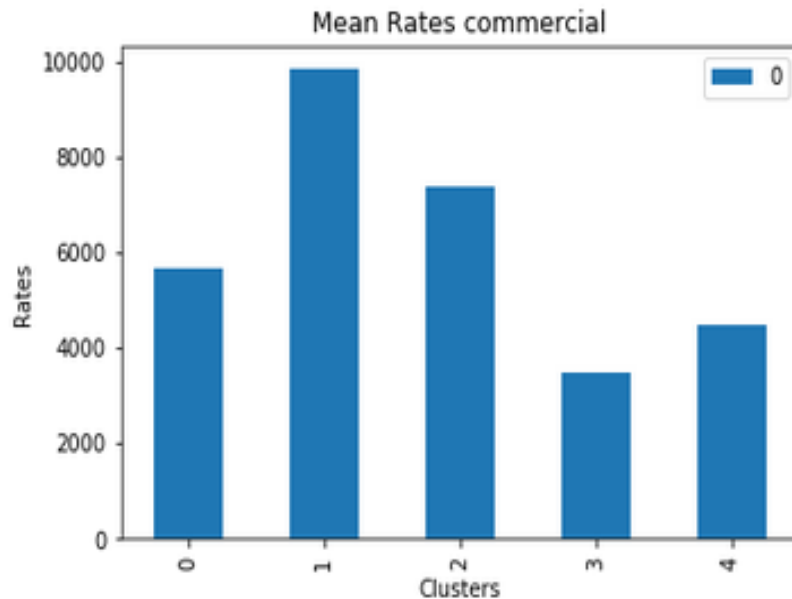
Merging classification results with Rates

	Area	Commercial_rate
Cluster		
0	37	37
1	36	36
2	46	46
3	1	1
4	1	1

Mean Rates commercial

	Area	Residential_rate
Cluster		
0	32	32
1	29	29
2	34	34
3	1	1
4	1	1

	0
0	3437.500000
1	6579.310345
2	4247.058824
3	1500.000000
4	2500.000000

Mean Rates Residential

Figures to the right show the Cluster average rates for both commercial and residential purposes. Notice Cluster 1 is the costliest in both cases. This data also shows the distribution of clusters in the Commercial and Residential areas

Results

- As per the analysis of cluster data and the Residential/commercial real estate rates, the following observations were made:

- C**luster 1** has many eateries, all round facilities.

- **Cluster 3** The residential and commercial rates are cheap and locality also is fine

- **Cluster 0** An average City area, this means, its a perfect combination of Residential and commercial space.

- **Cluster 4** Facilities are less and rates are more

- **Cluster 2** The rates are below average and adequate facilities.

Discussions:

- On a whole, if a person wants to start with small business, Cluster 3 locality is preferred. Cluster 2 is slightly costly but has adequate facilities than Cluster 3,this can be second preferred option for starting small businesses.

- If a person wants to start with Fast moving business that requires more visitors, Cluster 1 or Cluster 0 is most preferred. Cluster 1 has many eateries and businesses, this means it is a Commercial Business District. Starting Pubs or restaurants there would be advantageous

- Cluster 4 is not a preferred locality for Residential purpose as facilities are less and Rates are also more.

- Cluster 2 is most preferred for Residential Purpose followed by cluster 3 and cluster 0.

Thank you.