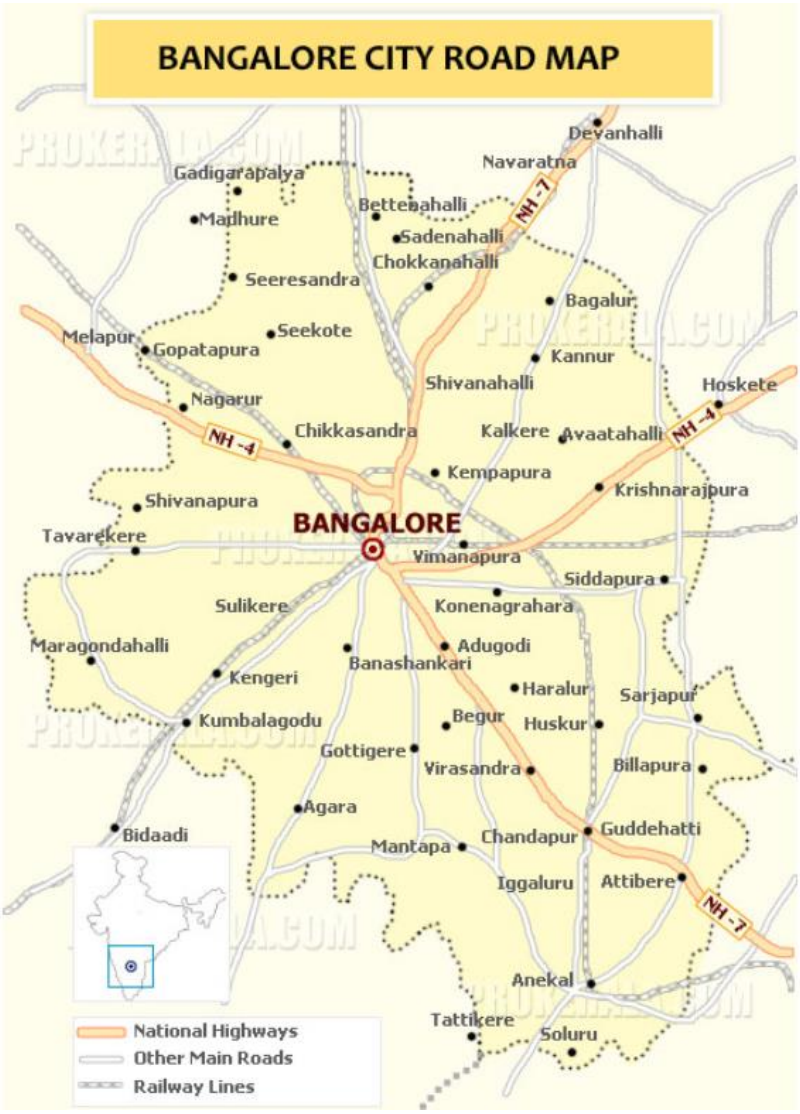
The Battle of Neighborhoods – Bangalore City.



Applied Data Science Capstone by IBM on Coursera.

Author:

Saideep R Naik.

1. INTRODUCTION: BUSINESS PROBLEM

This project deals with the major venue categories in the neighborhoods of **Bangalore, India**. This project would specifically help Business personal plan to start new Restaurants, Hotels, etc. in Bangalore, Karnataka, India.

The **Foursquare API** is used to access the venues in the neighborhoods. Since, it returns less venues in the neighborhoods, we would be analyzing areas for which countable number of venues are obtained. Then they are clustered based on their venues using Data Science Techniques. Here the **k-means clustering algorithm** is used to achieve the task. The optimal number of clusters can be obtained using **silhouette score** metrics.

Folium visualization library can be used to visualize the clusters superimposed on the map of Bangalore city. These clusters can be analyzed to help small scale business owners select a suitable location for their need such as Hotels, Shopping Malls, Restaurants or even specifically Indian restaurants or Coffee shops.

The major **Target Audience** would be small-scale business owners and stake holders planning to start their business at a location in Bangalore. This project would help them find the optimal location based on the category of their business such as,

- What is the best location to start a new hotel in Bangalore with restaurants around?
- Which area is best suitable for opening a Shopping Mall in Bangalore?

2. Data Requirements:

Bangalore has multiple neighborhoods. The Kaggle website has a dataset which has the list of locations in Bangalore along with their Latitude and Longitude in degree format. There is a total of 352 neighborhoods as shown in Fig below,

1. https://www.kaggle.com/rmenon1998/bangalore-neighborhoods

```
bang_data.head()
```

	Location	Latitude	Longitude
0	Agram	45.813177	15.977048
1	Amruthahalli	13.066513	77.596624
2	Attur	11.663711	78.533551
3	Banaswadi	13.014162	77.651854
4	Bellandur	58.235358	26.683116

Next the details of venues in each neighborhood namely **Venue, Venue Latitude, Venue Longitude,Venue Category** data needs to be obtained. Here, Foursquare API is used to obtain this data.

2. http://foursquare.com/

A total of 776 venues data have been obtained from Foursquare. The resultant venues dataset is used for the analysis process.

```
print(bang_venues.shape)
bang_venues.head()
```

```
(776, 7)
```

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Agram	45.813177	15.977048	Amélie	45.813842	15.979011	Dessert Shop
1	Agram	45.813177	15.977048	Trg bana Josipa Jelačića	45.813032	15.976868	Plaza
2	Agram	45.813177	15.977048	Vinodol	45.811666	15.975643	Mediterranean Restaurant
3	Agram	45.813177	15.977048	Tržnica Dolac	45.814070	15.977261	Farmers Market
4	Agram	45.813177	15.977048	Cremme Zagreb	45.814987	15.976296	Dessert Shop