2020 | By: Rajitha Shetty

Coursera - IBM Data Science

[New Shopping Mall In The City Of Mangalore]

CAPSTONE PROJECT

Opening a New Shopping Mall in Mangalore

Introduction

Shopping malls are becoming one of the most popular places of visit in every city. People like to visit shopping malls for different reasons. Shopping malls are becoming popular for different reasons like they have one stop shopping for different items like branded cloths, accessories, electronic, etc. And in Malls you will get wide verity of choices in each of these items. In city like Mangalore where climate is very hot; Malls provide very good shopping experience as Malls are air-conditioned. At Malls shoppers can hop between shops looking for right items. Malls also give good opportunity for window shopping. There are specialty stores, such as candle stores, bath and beauty stores and music and video stores. Finally, the mall is a wonderful entertainment center. Not only do people go there to shop, but they also got there for relaxation.

Background

Mangalore know as port city of Karnataka located in the southern part of India. It has a beautiful cost line. Mangalore is known for its swaying coconut palms, beautiful beaches and temple architecture. It is also an important industrial, commercial, **educational** and **healthcare** hub.

It is one of the fastest growing non metro (Indian city of population less than 1 million) in South India. Mangalore is the 2nd largest business centre in Karnataka. Around 75% of India's coffee, timber and cashew nuts exports are handled by the New Mangalore Port. Mangalore is one among the 5 cities in the country to have both a Major Port and an International Airport.

Mangalore has some of the tallest buildings in South India, with many more under construction. Mangalore International Airport is among the 2 International Airports in Karnataka. Mangalore is selected as one of the hundred Indian cities to be developed as a **Smart City**.

Business Problem

The objective of this project is - "Identifying a suitable location to build a Shopping Mall in City of Mangalore". Currently Mangalore has only 3 malls. And out of them 1 is comparatively small also very old. People at Mangalore like Banded cloths accessories. And list of multinational bands in these malls are limited. So there is huge potential to open up Banded outlets. And eating out is also catching up. Mangalorians like movies too; and multiplex at Mangalore play movies of different language English(Hollywood), Hindhi, Kannada, Tamil, Telugu,Tulu. document to the original contained in your current template.

This project will try to come up with a solution by using the data science methodologies and machine learning techniques like clustering; the best suited location for a new Shopping mall that is close to crowded areas of the city that has essentially high residential areas. And it should be easily accessible/ well connected as most people depend on public transport.

This project aims to provide solutions(answer) to the business question: **If a property developer is looking to open a new shopping mall in Mangalore, where would you recommend that they should open it?**

Business Stakeholders

This project is useful to **property developers** and also to the **investors** who are interested in investing in properties at Mangalore. According to *U.N Projections* population of Mangalore will grow from current **700,000** to around **1,000,000** in 2035. This steady growth in population will contribute to better business opportunities and Shopping Malls are one of the best investment options as more and more **people**(customers) like to visit malls as they are the one stop destination for Shopping, eating out, entertainment, socializing etc. This also gives better opportunity for the **Shopkeepers** to open up banded stores and ultimately **Government** will get more revenue and community too gets benefited.

Data

Based on our business problem ie. – "Identifying a suitable location to build a Shopping Mall in City of Mangalore." we need following data.

1. We need a list of neighborhoods in Mangalore to identify locations with Residential buildings, Hotels, Colleges, parks, transportation etc.
2. We need the latitude and longitude coordination of these neighborhoods for different analysis and visualization.
3. Analyze the obtained data to identify the best locations for coming up with new Shopping Mall.

Data Source

Following are the data sources needed for the analysis

1. **Google Map** to get coordinates of Mangalore City
2. **Wikipedia** to collect information about different neighborhoods (wards) in Mangalore.
3. Using **Foursquare API's** to find out the best location for the mall. Also analyze the different neighborhoods to find out the number of Residential building, Hotels, arks etc.
4. Will be using **Geopy Nominatim** to get the latitude and longitude of different locations/Wards.
5. **Pandas** will be used to clean up the data As we may have to add/remove/fix coordinates of different locations/wards. Data collection and cleaning will be a important activity in this project as there may not be enough information available from one single source.

Methodology

We can begin our process by finding out the coordinates of Mangalore City using Google Map. We will be using it to generate a folium map to visualize different results. Next we will look into Wikipedia page about "Mangalore city corporation" which has a list of all the wards(Neighbourhoods) in Mangalore. Using "Beautiful Soup" Python package we will perform web scraping activity on this Wikipedia page and generate a list of wards.

We need to perform some data clean up activity as the list contains not only the name of the Wards but also other listed information from the wiki page. And in the wards list we will find many duplicates as wards may be divided as south and north and these duplicate entries need to be removed. After removing the duplicate we will Geopy Nominatim to get the latitude and longitude of different neighbourhoods/Wards. And Nominatim may not return latitude and longitude of all the wards and we will be finding out the latitude and longitude manually and update the list. We will use Pandas extensively for this process. So all the activity of Add/Remove/Update of Neighbourhood information is performed using the pandas library. And at the end of this process we will get a list of all Wards/Neighbourhoods with latitude and longitude values.

Next by using Google search we will gather information about the current Malls at Mangalore. And we will convert this information to pandas dataframe which will be used at the time of analysis to show the geo-locations of these Mall in the clusters created by us.

Now instead of considering all kinds of venues in our neighbourhood; we will look into few categories of venues which make more sense in identifying the best locations for shopping mall. We can consider following the categories that can be used to collect information about venues in our Neighbourhood using Foursquare API's. These categories sound more relevant than using all the categories.

- Residence
- Travel & Transport
- College & University
- Outdoors & Recreation
- Professional & Other Places

Next we will find out the neighborhood venues of all these categories. One by one and then we will merge all these data to form a new dataset that we will be using for analysis.

Analysis

Next we can Perform One hot Encoding on the venues collected. And Group neighbourhood by taking the mean of the frequency of occurrence of each category. we can find out **top 10** venues in each neighbourhood. And sort this information based on Neighbourhood.

Next we have to cluster these neighbourhoods using **KMeans**. In this project neighbourhoods are clustered into **5 clusters**. Once clustered; we will use Folium to display these clusters on Map. Here we will be displaying the map of Mangalore with 5 clusters that are created. This will help us in visualizing the data.

We can next look into data available for individual clusters and analyze it. This list of individual clusters contains the top ten most common venues found in them. This data can help us in predicting the most suitable Neighbour hood for the new Shopping Mall.

Result

- Cluster 0: Neighborhoods in cluster Zero has colleges and gyms followed by recreational venues and Residential area. And it also has many new housing colonies are being designed.

- Cluster 1: Neighborhoods in cluster One has huge number of residential places is located; It also has lots of Hotels. Custer 1 neighborhoods are well connected and also have some of the schools and colleges.

- Cluster 2: Neighborhoods in cluster Two has offices and most of the recreational activities; has Beaches, River front, Sports centers, it also has few collages and residential places and new housing colonies are coming up.

Discussion

In **Cluster 0**, we see all the required inputs like Residential area, colleges and recreational areas. There are projects on new housing colonies also being designed in **Cluster 0**. **Cluster 1** would have been a better choice, but it already has **3** (all) shopping malls within a 10km radius. So, I am suggesting **Cluster 0** as a more viable option for coming up with a Shopping Mall.

Cluster 2 can also be considered as most of the recreational activities are found in this area. So lot of people will be visiting these neighborhood. It also has some Housing development activities going on. And it is also a growing part of Mangalore and current Mall are comparatively far away from **cluster 2** so it is also a ideal spot to come up with new Shopping Mall.

4
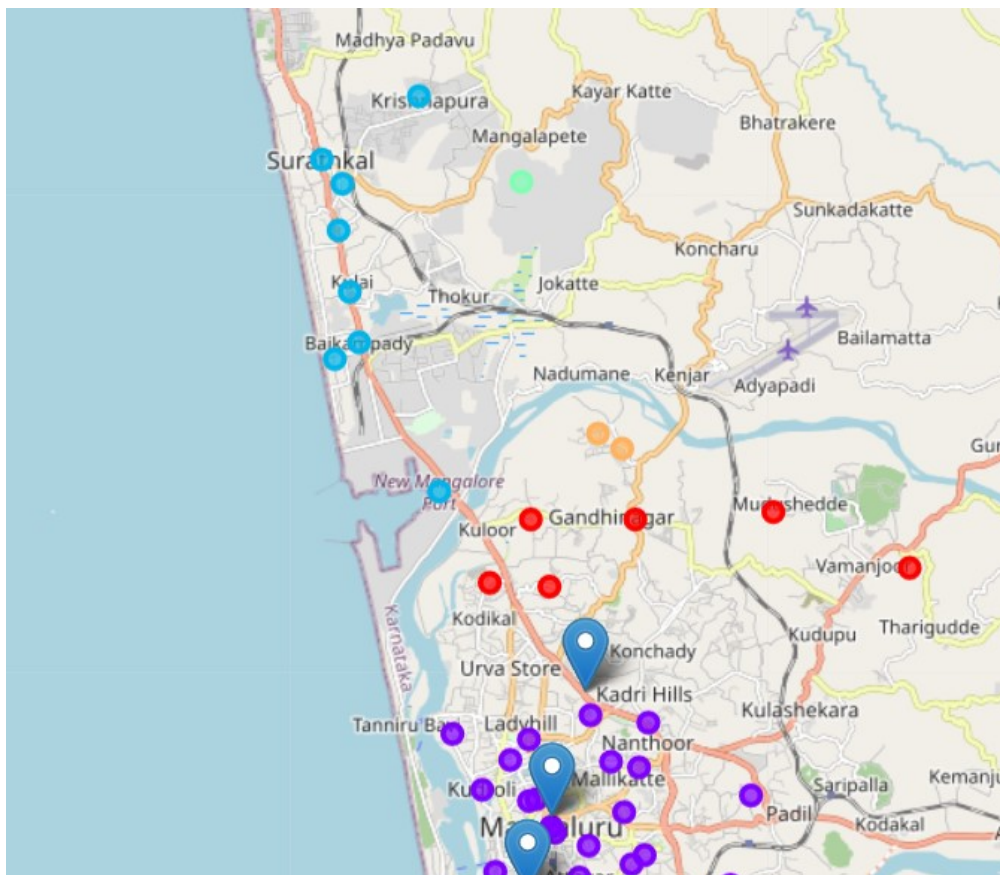
Limitations

In this project only the population of the city has been given the highest priority. So venues where selected based on the factors that where people oriented like residence, Travel (hotel), Collages, Outdoors & Recreation and Professional & Other Places. No other aspect is considered in identifying the suitable location. And as Mangalore city is growing rapidly the data used for this analysis may not be considered as sufficient. But the methodology used can be easily adopted for a large set of data with more relevant categories.

Conclusion

For a shopping mall to become successful, it needs people. So, a good place would be somewhere in a residential area, with commercial activity and colleges within the vicinity. Even Recreational activities nearby could have a good contribution of people walking in to the mall. Not all of them would be buyers, but this could create a lot of walk-ins.

Purpose of this project was to identify a suitable location to come up with a new shopping mall. And after the analysis we can say neighbourhoods in **Cluster 0** are best suited for this project,

We can consider **Cluster 2** as it has a combination of colleges, some industries and new housing projects are coming up and it is a growing part of Mangalore. And it is popular for recreational activities and beaches.