

Optimal location for opening a new cafe alongside the Bangalore Metro

Muhammed Salih

27 April 2020

1. Introduction

1.1. Background

Bangalore is one of India's biggest cities with a population of over 10 million. It is often called the 'Silicon Valley' of India as it is the hub for India's IT companies.

The Bangalore Metro Rail Corporation (BMRC) operates the Namma Metro (translated as Our Metro in English) which serves as a key player in the city's public transit. The Namma Metro which started operations in 2011 currently has 40 stations and sees a daily ridership in the range of four hundred and fifty thousand. Due to its high ridership, the metro stations and its vicinity have become centres of commercial activity, especially for food and beverage outlets.

1.2. Problem

This project tries to suggest the suitable locations near metro stations for opening new cafes for a client who is planning to expand its operations in the city over the next few years.

1.3. Interest

Our client who owns a cafe business with more than 50 cafes across the country currently, is executing this study as part of its plans to expand its operations in the city of Bangalore.

2. Data Description

2.1. Data Sources

In order to do the analysis, we need the location data of each metro station which can be readily acquired from Wikipedia. Once the location data is ready, the Foursquare API can be used to get a detailed overview of the venues that are currently operational in the vicinity of each metro station. So to summarise, the required data for this project will be acquired from the following sources.

Table 1 - Data Sources

Item	Source
Geographical locations of the metro stations	Wikipedia Pages
Venue data for each stations	Foursquare API

2.2. Data Cleaning and Formatting

Since our dataset will primarily contain the venue data from Foursquare API, missing values or incorrect data will not be a big concern. The data will be stored as a pandas dataframe in Python and appropriate Python packages will be used for EDA, visualisation, and modelling.

	Venue Name	Category	Latitude	Longitude
0	Hotel Fishland	Seafood Restaurant	12.975569	77.578592
1	Udupi Sri Krishna Bhavan	Indian Restaurant	12.971563	77.574158
2	SGS Donne Biryani	Indian Restaurant	12.970325	77.572648
3	Sangam Sweets	Dessert Shop	12.976924	77.577891
4	Sapna Book House	Bookstore	12.976355	77.578461

Fig 1 - Sample dataframe with the venues near Majestic metro Station

2.3. Approach towards using the data

The Foursquare API can give a detailed insight into the composition of the type of businesses that are currently operating near the metro stations. By using proper transformations like one-hot encoding and other operations, the relative frequency for each business type (say coffee shops, cafes, restaurants, cinema etc.) can be found out for the metro stations.

As far as the aim of the current project is concerned, a neighborhood currently overcrowded with cafes might not offer a good prospect when it comes to opening a new cafe. And at the same time, a neighborhood that has very few food and beverage outlets currently might represent a lower potential for the success of such a venture as well. So our best bet for opening a new cafe would be neighborhoods that already have an active set of food and beverages businesses running currently, but the number of cafes are not very high.

For our analysis, the metro stations would be clustered using the k-means clustering algorithm based on the neighborhood venue makeup as obtained from the Foursquare API data. Once the clustering is done, the characteristic features of the clusters (like the top most occurring type of venue) will be used to finalise the clusters that are good candidates for opening a new cafe for our client.