Coursera Capstone

Set up Café in Bangalore

IBM DATA SCIENCE CERTIFICATION

Amitabha Kanjilal
September 2019

Introduction

In the modern days that we live in, cafés have become very popular, and is a great place to hang out with friends or family. Any major city would usually have a bunch of cafés around, and many investors still look out for business opportunities for opening brand-new cafés in town. However, the location is a vital factor for getting profits out of it.

Business Problem

The objective of this capstone project is to analyse and select the best locations in the city of Bangalore, India to open a new café. Using data science methodology and machine learning techniques like clustering, this project aims to provide solutions to answer the business question: In the city of Bangalore, if an entrepreneur is looking to set up a new café, where would you recommend that they open it?

Target Audience

Anyone who is interested in setting up a café in a big city can use this approach to determine the best location in the city to make his/her business a success.

Data Collection

We need the following data to solve the problem –

- List of neighbourhoods in Bangalore

- Latitude and Longitude for those neighbourhoods

- Venues, especially cafés in 10 km radius of these co-ordinates

We are extracting the neighbourhood data by web-scraping the Wikipedia link for localities in Bangalore. After we get the neighbourhood names, we are searching for the co-ordinates of those neighbourhood in google and again scraping for the latitude and longitude values. Once the data has been fetched, it would look like –

	Neighborhood	Latitude	Longitude
1	Domlur	12.9610	77.6387
2	Indiranagar	12.9784	77.6408
3	Jeevanbheemanagar	12.9642	77.6581
4	Malleswaram	13.0055	77.5692
6	Sadashivanagar	13.0068	77.5813
7	Seshadripuram	12.9889	77.5740
8	Shivajinagar	12.9857	77.6057
9	Ulsoor	12.9817	77.6284
10	Vasanth Nagar	12.9920	77.5943
11	Bellandur	12.9304	77.6784

Once we have the co-ordinates, we will loop through the data frame and are use the foursquare API to explore the neighbourhood and fetch top 100 venues within a 10 km radius of each neighbourhood. After we fetch all the venues, the data set would look like this –

`blr_venues.head(10)`

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Domlur	12.961	77.6387	Lavonne	12.963909	77.638579	Café
1	Domlur	12.961	77.6387	Barbeque Nation	12.962684	77.641599	BBQ Joint
2	Domlur	12.961	77.6387	Smoke House Deli	12.965584	77.641498	Deli / Bodega
3	Domlur	12.961	77.6387	Anand Sweets	12.960166	77.645168	Indian Restaurant
4	Domlur	12.961	77.6387	League of Extraordinary Gamers	12.967099	77.636919	Gaming Cafe
5	Domlur	12.961	77.6387	Starbucks	12.965649	77.641718	Coffee Shop
6	Domlur	12.961	77.6387	Bodycraft	12.968497	77.641289	Spa
7	Domlur	12.961	77.6387	Puma Social Club	12.967254	77.641212	Nightclub
8	Domlur	12.961	77.6387	Big Pitcher	12.960101	77.646946	Brewery
9	Domlur	12.961	77.6387	Murphy's	12.953659	77.639397	Irish Pub