

Coursera - Capstone Project

IBM Applied Data Science Capstone

By: Shivali Sur

Food Business In Pune, Maharashtra

Introduction:

Pune city is one of the major business hubs in India. The city has a major concentration of IT companies spread across the city. Its a city where a lot of people come from different parts of India to start their careers. Also, the city has a very good education system and is home to multiple renowned educational institutes. As a result, the city is a home to a diverse people coming from different cultures, languages, locations and age groups. The city witnesses a large number of businesses open and close down every year. Few grow to new heights while others, which are not so successful, close down. The project is aimed at analysing the data available in the public domain to answer important questions while starting a business.

Business Problem:

As a budding entrepreneur, I am passionate about Ice Creams and would like to setup an Ice Cream Parlor. I would like to understand the business patterns across the Pune City and narrow down a few neighbourhoods favourable for setuping up an Ice Cream Parlor.

Target Audience Of This Project

Target audience are entrepreneurs who are looking to setup a business in Pune City. The project helps then analyse existing data and make informed decisions to make the most out of their investment.

Data

Primary Data:

To get started, I will be using the business and location data from FourSquare.

Additional Data

In addition to the Primary Data, we can use the following data to narrow down on the neighbourhoods and make informed decisions: - Spending trends across the city on various services - Demographic data detailing the diverse nature of residents, for instance, students, professionals, senior citizens etc. - Area Cost: Rent and real estate costs, which can provide insight in estimating the initial investments to be made in starting the business.

Sources of data and methods to extract them

- https://en.wikipedia.org/wiki/Category:Neighbourhoods_in_Pune
- <https://api.foursquare.com/>
- <https://geocode.arcgis.com>

Methodology

Scraping Data: Scrape Neighbourhood data for Pune City from Wikipedia. This will provide information about the various Neighbourhoods in the city.

Extract Location Data: Use Geocode API to extract the location data (Latitude and Longitude data) about these Neighbourhoods and create a map.

FourSquare API: Use data from FourSquare API to get details about various neighbourhoods, existing businesses and categorize these business.

Cluster: Divide neighbourhood into multiple clusters. This will help to broadly group multiple neighbourhoods based on neighbourhood similarity.

Result

Neighbourhoods in Pune are divided into multiple clusters which can then be analysed. Based on the top 4 most common categories are: - Indian Restaurants - Snack Place - Ice Cream Shop - Fast Food Restaurant

Since we are in India, its safe to say Indian Restaurants are in abundance and hence its safe to exclude it from the list. Based on the above data, it is evident that people like to visit Snack Places and Ice Cream Shops in the city. With the above information, we can now narrow down on the neighbourhoods where an Ice Cream Shop is not available. We can safely say people from these neighbourhoods travel to other neighbourhoods to enjoy Ice Creams due to non availability of joints in their respective neighbourhoods. These neighbourhoods can prove to be good places to open up an Ice Cream shop as people can opt to purchase from a local nearby store which will not require them to travel to a different neighbourhood for the same reason.

Limitation and Suggestion For Future Research

The above conclusions are based on available data from ForeSquare. However, the results can be refined by analysing the given data in conjunction with the following: - Spending trends across the city on various services - Demographic data detailing the diverse nature of residents, for instance, students, professionals, senior citizens etc. - Area Cost: Rent and real estate costs, which can provide insight in estimating the initial investments to be made in starting the business.

This will help narrow down the neighbourhoods and help make a much more informed decision in terms of expenses, spending trends of residents, their demographics etc.