OPENING A BAKERY SHOP IN MUMBAI- INDIA

IBM APPLIED DATA SCIENCE CAPSTONE PROJECT

JUNE, 2020

SREYA SARKAR

Introduction

Bakeries are a popular type of foodservice establishment, and they allow us to express our culinary creativity while also serving a unique market. The business for selling baked products has been there for thousands of years. However, it has also evolved over the years due to the fame and desire that the art of baking has received. While grocery stores and supermarkets are keen on selling pre-sliced breads, cakes, and pastries, many still prefer to get their kind of flour baked from a small artisanal bakery. This may be because of taste or age-old household tradition. The recent move of commercialization of baked goods has increased the quality characteristic in the trade of baking. Therefore, the passion of baking is now catching up into the small   and   big   bakery   business.

People in India are fueling their passion for food through their open food business ventures. Along with the startup wave hitting the country, and a major disruption seen by the Restaurant Industry has encouraged a lot of people to start up their restaurant venture, especially women. Professionals and home cooks alike are now fuelling their passion for food and to get into the food business. Baking is one such passion that has been catching up lately, with many big and small bakery businesses cropping up. Opening a bakery presents many unique challenges that are different from other types of businesses.

The profit potential for a bakery is good, but starting a food business is never easy. A lot of things have to be taken care of. Opening a bakery presents many unique challenges that are different from other types of businesses. The total approximate cost of opening a bakery in India can be figured but the cost of equipment and location can lead to considerable variance in the approximate cost.

One of the first steps to how to start a bakery business in India is therefore to decide a good location. The location plays a huge role in its success. Although a great location may not guarantee success, a bad location will almost always guarantee failure.

- Look out for the competitors in the area. **Lesser the competition**, easier is the sales.
- The best location combines visibility, affordability and lease terms you can live with.

Business Problem

The objective of this Capstone project is to analyze and select the best location to opening a Bakery Business in the India's Financial capital, Mumbai. Using the data Science methodology and machine learning techniques like clustering, the aim is to provide answer to the question:

If a Restaurateur is planning to open a new Bakery business in Mumbai, where would you recommend it?

Target Audience

To solve this problem, data scientist team led by myself has been engaged by a Restaurateur. The objective is to locate and recommend to the management which neighborhood will be the best choice to start off their first bakery shop. The management also expects to understand the rationale of the recommendations in the report.

Success Criteria

The success criteria of this project will be a good recommendation of the neighborhood choice in Mumbai to the Restaurateur based on 2 key factors; lack of Bakery Shops available (less competition) and higher number of residences presented (higher demand).

Data

We would need the following data to analyze the opportunity:

- List of neighborhoods in Mumbai, India. This defines the scope of the project which is confined to the city of Mumbai.
- Latitude and Longitude coordinates of those neighborhoods. This will help us to plot the locations on the map and also get the venue data.
- Venue data, particularly Venue equals to Bakery. We will use this data set to perform Clustering.

Data Sources

We will use the Wikipedia page ("https://en.wikipedia.org/wiki/Category:Suburbs_of_Mumbai") which contains list of neighborhoods in Mumbai which is close to 42. With the help of different Python Packages and using the scraping techniques, we will extract the data from the Wikipedia page. Post that, we will fetch the latitude and longitude coordinates of the neighborhoods using the geospatial data.

We will then use the Foursquare API to get the venue data for those neighborhoods. This API will provide us with many categories of venue data for those neighborhoods. As we are interested in Bakery, we will further filter out Bakery from the venue data and proceed with the analysis. This project will make use of few data science techniques like Web Scraping, Working with Foursquare API, data cleaning, data wrangling, Machine learning algorithm like K-Means clustering and Map visualization.

Going forward we will explain the methodology i.e. the steps taken in this project.