Using Data Science to identify the optimal location for a new Italian restaurant in Sydney

INTRODUCTION

In this scenario, we are going to look at the case study of an Italian restaurateur. Italian cuisine is one of the most popular cuisines in the world and for a prospective restaurateur looking to open a new Italian restaurant, it is important to find the right location. Location is single handedly the most important factor for a restaurant that can affect its success.

Business Problem

The objective of this capstone project is to explore the city of Melbourne, Australia and find the most appropriate location for opening a new Italian restaurant. We will be using a variety of data science and analytical methods such as clustering. We will aim to answer the following question : What is the best location to open an Italian restaurant in Melbourne?

Target audience

Any restaurateur who wants to find the best location to open an Italian restaurant in Melbourne. This approach can be replicated in any city in the world and modified for any other cuisine restaurant easily.

Data

To solve this problem, we will need the following data:

- List of neighborhoods in Melbourne. This defines the scope of the project, as confined to the city of Melbourne, Australia
- Latitude and Longitude co-ordinates of these neighborhoods. We will need these to plot the map and also to get the venue data
- Venue data, related to restaurants, particularly Italian restaurants. We will use this data to perform clustering and other analysis on the neighborhoods

Sources and extraction of Data

- We will use the data from the Australian government datasets located at Data.com.au which is a geojson file containing the administrative boundaries of Sydney. We can extract the district names from there as well as any map boundaries if required

- Getting latitude and longitude data of these neighborhoods via geocoder package
- Using foursquare API to get details of venues in these neighborhood