

Predicting the favourable location to open a restaurant in New York City

Chandni Mandaviya

November 27, 2020

1. Introduction

1.1 Background

This project aims to locate a favourable neighborhood to open restaurant in New York city. Favourable neighborhood will not only do profit to the surrounding business such as shopping malls or residence areas but will do a major profit to the restaurant since it can drive more customers to it.

1.2 Problem

The objective of this business problem is to analyse and select the best locations in the New York city to open restuarants. Using the data science methodology and machine learning techniques like clustering this project aims to provide solutions to the business problem: if a contractor is trying to start their own business like opening a restaurant , where would you recommend that they setup their office?

1.3 Interest

The project is particularly useful to property developers and investors looking to open or invest in new restuarants in the New York.

2. Data Acquisition

2.1 Data

To solve the problem, we will be needing the list of neighbourhoods in New York city. This defines the scope of the project.

Latitude and Longitude co-ordinates of those neighborhoods. This is needed to plot the map and also to get the venue data particularly related to the restuarants.

This data will be useful to do the clustering on the neighborhoods.

2.2 Data Source

The list of the neighborhoods of the new York city can be found out from https://en.wikipedia.org/wiki/Neighborhoods_in_New_York_City . In order to get this data, we will be using web scrapping techniques to extract the data with the help of python beautifulSoap packages. Python has a Geocoder package which will be useful in order to get the latitude and longitude of the data. After, we will be using Foursquare API to get the venue for those neighborhoods.

Hence, in this project will make use of data science skills from webscrapping, working with API(FourSquare), data cleaning, data wrangling, machine learning model(K-means , visualization (Folium package).