

Capstone Project - The Battle of Neighborhoods

Italian Restaurant in New York

Introduction

New York City (NYC), often simply called New York, is the most populous city in the United States. With an estimated 2019 population of 8,336,817 distributed over about 302.6 square miles (784 km²), New York City is also the most densely populated major city in the United States.

The restaurant industry is essential to New York City's social and economic fabric. From small, family-owned restaurants and food carts to four-star world-famous establishments, the City abounds with restaurants that offer cuisines from every corner of the globe. Since March 2020, the restaurant industry has been hit very hard by the COVID-19 pandemic. Mandatory closures, stay-at-home and social distancing orders, the onset of a severe economic recession, and travel restrictions have resulted in unprecedented upheaval for the industry. As part of this final project we need to explore the best locations for an Italian restaurant throughout the city of New York for the owner such that the restaurant can have great success and consistent profit despite the pandemic situation.

However, as with any business, opening a new restaurant will require serious considerations specifically due to pandemic situation. Therefore, our project will attempt to answer the questions "Where should the investor open an Italian Restaurant?" and "Where should I go if I want great Italian food?"

Business Problem

The objective of this Capstone project is to analyze and select the best locations in the city of New York to open a new Italian restaurant. Using Data Science methodology and instruments such as Data Analysis and Visualization, this project aims to provide solutions to answer the business question:

- Where in the city of New York, should the investor open an Italian Restaurant?
- Which areas have large number of Italian Restaurant Market?
- Which all areas have a smaller number of Restaurants?
- Which is the best place to stay if I prefer Italian Cuisine?
- What places are having best restaurant in New York?

Data

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

- New York City data containing the neighborhoods and boroughs.
- Latitude and longitude coordinates of those neighborhoods. This is required to plot the map and get the venue data.
- Venue data, particularly data related to restaurants. We are going to use this data to perform further analysis of the neighborhoods.

Data Source and methods to extract them

New York City data containing the neighborhoods and boroughs will be obtained from the open data source: https://cocl.us/new_york_dataset.

Geographical coordinates of the neighborhoods (latitude and longitude) using [Python Geocoder package](#).

Finally, we will use [Foursquare](#) API to get the venue data for the neighborhoods defined at the previous step.

This project will require using of many data science skills, from web scrapping (open source dataset), working with API (Foursquare), data cleaning, data wrangling, to map visualization (Folium). In the next Methodology section, we will discuss and describe any exploratory data analysis that we did, any inferential statistical testing that we performed, and what machine learning techniques were used.