

## **Problem Statement**

Eating out is big. Pre-COVID, people loved going out to eat and post-pandemic, I would assume the trend will continue. Chefs are being treated like rock stars. So, I wanted to take a dive into a problem that relates to the food industry in the biggest city in US i.e., New York City. New York is the most culturally diverse city in the United States with immigrants from all countries across the globe. A quick Wikipedia search suggests that Italian Americans comprise of the largest ethnic group in New York clocking in at 8.2% of the population. Italian cuisine is amongst most well-known and popular cuisines in the world. My project aims at combining the above categories - New York City and Italian cuisine. Combining these two aspects, my project will provide suggestions as to where one should open an Italian restaurant and why. It also tells a customer where they can find a good Italian restaurant to eat at.

The project aims to answer the following 4 questions

1. Most Italian restaurants by neighborhood or borough - This tells us if borough is saturated or if there is an opportunity to open a place with less competition
2. List of the most popular Italian chain restaurants - This is aimed at an entrepreneur who wants to operate a franchise instead of an independent restaurant
3. Highest rated Italian restaurants - This gives a list of restaurants which are the highest rated by borough. This is aimed at a casual consumer or someone who wants to open a restaurant and avoid strong competition
4. List restaurants by tips - This is aimed again at both a consumer and a would be restaurateur. A higher tip typically suggests a more upscale or high end restaurant and vice-versa. This suggests a customer about a restaurant based on their lunch/ budget. It also tells a restaurateur about the type of restaurants in the neighborhood based on which they can take decisions.

## **Data Set Used**

I used data from the Foursquare application using their API to access the data set. I also pulled location co-ordinate data from the following dataset - [https://cocl.us/new\\_york\\_dataset](https://cocl.us/new_york_dataset). The four square API provided me with restaurant information which is gathered via user check ins.

The location data is the form a json file which contains the following details - Borough, Neighborhood, Latitude and Longitude.

The data pulled from the Four Square API contains the following details - Italian restaurants by neighborhood, Name of the restaurant, Borough, Neighborhood, ID, Likes, Rating and Tips. The four square API will use location data from the location dataset and select restaurant details.