## **Restaurant Location Recommender – New York City**

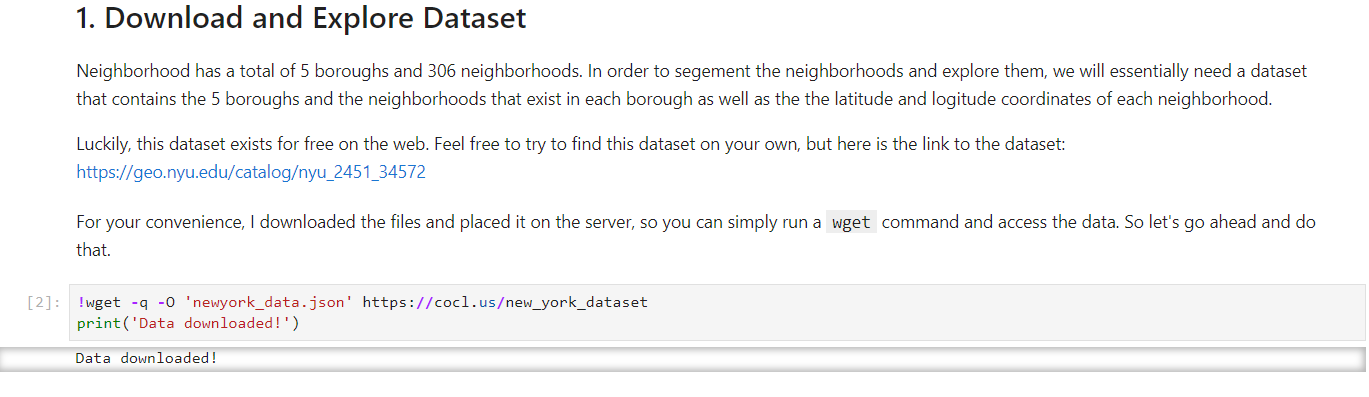
**Introduction/Business Problem：**

**This capstone project is meant to build a restaurant location recommender in New York City. NYC is one of the world’s most prosperous and diverse metropolis. There are a huge number of restaurants in NYC already, recorded at 26,697 in 2017, which is up from 25,305 in 2014 [1]. Therefore, it is indisputable that the competition in the NYC restaurant business is fierce. With the number of restaurants still increasing, it can be anticipated that the competition will become even stronger in the future. Therefore, for a prospective new restaurant owner who dreams of opening a restaurant in NYC, it is critical for them to understand where to open their restaurant based on the desired restaurant type (e.g. price range, cuisine, etc.) they would like to open, to maximize their probability of survival and profitability in the sea of restaurants in NYC. This project is targeted for anyone who is interested in opening a restaurant of any type in NYC. It should offer them a choropleth map of the profitability of opening their desired type of restaurant in all the neighborhoods in NYC.**

**[1] Restaurant Statistics, NYC** Department of Health, 2017

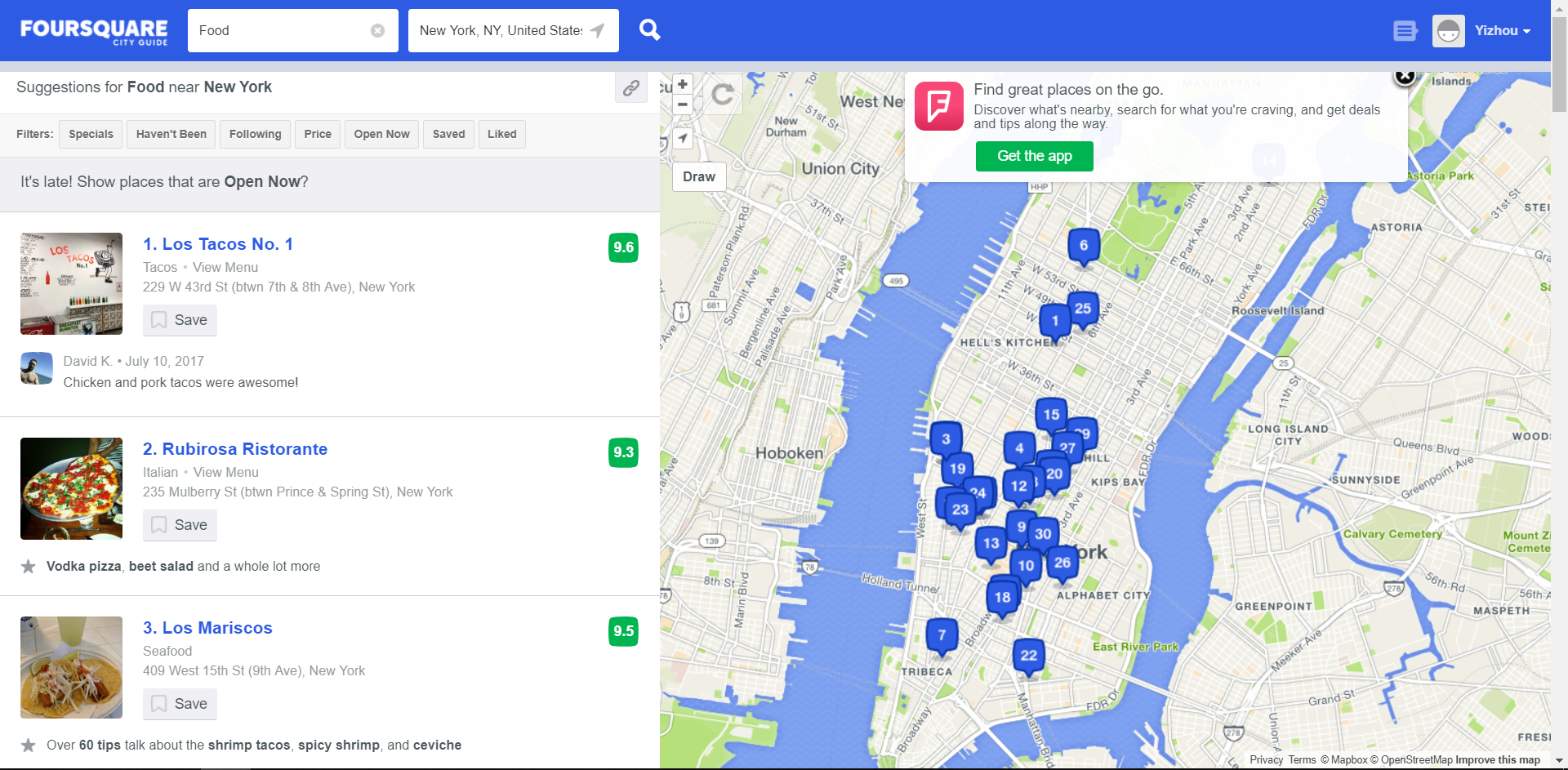
**Data：**

The first dataset needed for this project will be the json data of NYC, which includes 5 boroughs and 306 neighborhoods. This data is essential for obtaining a complete set of location data, including boroughs, neighborhoods, longitude, latitude, etc. It is also the basis of the choropleth map we are trying to create as a final product of the project. This dataset has been used in the ungraded lab of Week 3, so it will be processed similarly in this project.





Besides, there is of course Foursquare location data, which includes basically all the restaurants in NYC based on their exact location (longitude, latitude). Foursquare location data provides venues in which we can access details including ID, name, location, URL, contact information, ratings, tips and other statistics. These information provides the backbones of our restaurant locator. An example of Foursquare location data is provided below.





Foursquare location data of restaurants will provide insights into the competitiveness of a certain type of restaurants (price range, cuisine, etc.) in each neighborhood, taking into account the ratings of restaurants. For example, if a neighborhood already has a couple of Arabic restaurants, but they are all given low ratings, then the rating information will help level up the potential profitability of opening a new Arabic restaurant in that neighborhood, with an assumption that the new restaurant owner has what it takes to open a highly rated Arabic restaurant.

Finally, there is the demographics data from Data USA, namely Race and Ethnicity data of NYC. Residents of different races would prefer different cuisines, and the demographics of NYC, combined with the number of restaurants available for each cuisine, would largely affect the survivability of a restaurant with a certain cuisine. Below is a bar chart based on the Race and Ethnicity dataset of NYC.

A picture containing screenshot

Description automatically generated