

Problem Statement & Background

Setting and Running up a restaurant requires a significant investment. It is important that, the investment is made after proper research so that chances of failure are the lowest and chances of success are the highest.

There are multiple parameters to look at while setting up a restaurant. Before an investment can be made it is important to define which cuisine and which area. If the investor has constraint for one parameter then data can be used to identify which is the best option for the second parameter. So let say, If I want to open a Mexican joint as I have expertise in the same. I would like to open in an area where there are people ready to eat outside but not many Mexican joints are available.

The idea here is to understand locality based on available data and comparing the same with similar localities in the New York. Once comparison is done then idea is to find the locality with opportunity. Locality with opportunity is nothing but it matches with other localities but have least number of Mexican joints compared to similar localities. That locality should be ideal for us to look for an place for a Mexican joint.

Data Source

We will need 2 sets of data.

1. Details of New York localities
2. Details of restaurants in each locality

1st set of data will be used from https://geo.nyu.edu/catalog/nyu_2451_34572. This has details on all neighbourhoods of New York

2nd set of data will be pulled from FourSquare APIs. We will user explore function in the library to get mmost common venue categories in each neighbourhood.