Introduction and Business Problem

Introduction:

- The city of Hoboken, NJ is relatively small at ~1 square mile but it is packed with restaurants, night life and amazing people. For people that are new to Hoboken, despite its small geographic size, it can be daunting to figure out what restaurants are worth going to and where they are. For people that used to live in Hoboken or are visiting Hoboken, how do you know what the best places are to get something to eat?

Business Problem:

- For this project, I am going to put on my entrepreneur hat and create a simple guide on where to eat based on Foursquare ratings, price, category and geographic location data for restaurants in Hoboken. I will then cluster these restaurants based on their similarities so that a user can easily determine what type cuisine they are looking for and in what price range.

Data Required

For this assignment, I will be utilizing the Foursquare API to pull the following location data on restaurants in Hoboken, NJ:

- Venue Name
- Venue ID
- Venue Location
- Venue Category
- Rating
- Price
- Count of Likes
- Phrases

Data Acquisition Approach

To acquire the data mentioned above, I will need to do the following:

- Get geolocator lat and long coordinates for Hoboken, NJ
- Get a list of all venues in Hoboken
- Get venue IDs for each venue in Hoboken
- Pull venue name, location, category, rating, price, count of likes and phrases for each venue

Data Usage Approach - Clustering

I will take the gathered data (see above in Data Acquisition Approach and Data Required sections) and will create a k-means clustering algorithm that groups restaurants into 4-5 clusters so that people looking to eat in Hoboken can easily see which restaurants will cater to what their needs.