Introduction

The purpose of this project is to apply data science techniques in order to determine what type of restaurant venue and which particular location appears to be the most appealing to create within the Boston metropolitan area.

This will be accomplished by data scraping census and demographic data to first explore neighborhoods in the area, then later using Foursquare API data to see the density of specific restaurants in these neighborhoods, and to see how popular these restaurants are.

Business Problem

Suppose we want to open a restaurant somewhere in the Boston metropolitan area, and need to determine what would likely be the most successful venture. What type of restaurant would be the most successful in the Boston area? Which neighborhood would be the best location for the determined type of business?

Description

There are a myriad of factors to consider when looking to open a restaurant: Population, demographics, competition, just to name a few. By using existing data, we hope to look at these factors and use what we learn to better determine where an ideal location may be.

Census data can be used to first break down the city of Boston into its various boroughs and, on a smaller scale, neighborhoods. From there, we should be able to see the populations of these neighborhoods. Population is the first factor to look at, as a densely populated neighborhood, or one that attracts a large number of people, correlates to how successful a restaurant will be; more people means more customers.

Census data also provides demographics of these neighborhoods, which is helpful in determining what type of restaurant makes the most sense. Does a particular neighborhood have a large population of younger people? Is there a particular ethnicity that is more prevalent in certain neighborhoods? These data points are crucial in figuring out location and venue type, as specific venues will perform better if tailored to a neighborhood's chief demographic.

Foursquare data will be important once we have the population and demographics of each neighborhood, as we can see what the correlations are between specific groups of people and types of restaurants. Additionally, we can look at the density of certain restaurants as well as their popularity; this will help us avoid picking a location where competition is already high, while still having an idea if existing venues are popular enough to warrant a new restaurant.