Boston Neighborhood Analysis

**John Hennessey**

**May 22, 2020**

**Introduction**

This study will inspect and compare types of venues in different neighborhoods throughout Boston. It will result in venue data clustered by neighborhood. Graphs will display where a certain cuisine is most popular. The top venues in each neighborhood will also be considered and displayed.

**Problem**

*How does a person decide where to live when they move to a new city?* For this project I will assume that types of venues indicate types of culture and that a neighborhood with a person’s favorite type of venues will be the best place for them to live. To solve this problem, I will analyze and group venue data by neighborhood and present each neighborhood’s top cuisine choices.

**Discussion**

When someone decides to move to Boston, how will they know where to live? Most people who move to a new city attempt to live close to their job. Besides this requirement though, many who do not have a familiarity with the different boroughs and neighborhoods find themselves lost. After moving to a certain neighborhood, many people end up spending most of their time in a different neighborhood as it appeals to their interests and most importantly, it contains their favorite venues.

## Data

Unfortunately, I could not find all the necessary data in one place. I had to create an excel file containing the different neighborhoods of Boston and their coordinates which I found on Google. These coordinates were linked to zip codes, however some neighborhoods had multiple zip codes and some zip codes contained multiple neighborhoods. In some cases, I had to estimate the coordinates for a given neighborhood. I estimated coordinates for these neighborhoods which, when mapped, were close enough for the sake of visualization. Using this excel file, I input the coordinates for each neighborhood into Foursquare to attain different venues in each neighborhood. The Foursquare Data combined with the Neighborhood Data offers insight into the venue options in each neighborhood. Using all of this information, I am able to determine the best place for an individual to live based on their venue preferences.

Ultimately, I simplified this study to accept the following user inputs: five most important venues to have in the neighborhood including one “top priority” venue. Hopefully, a neighborhood exists that will have at least one instance of each of these five venues and many instances of the “top priority” venue. In order to determine if a neighborhood satisfies the “top priority” venue portion, I will check where the venue is ranked in each neighborhood. If it falls in the top 3 the requirement is satisfied (score = 1), top 5 (score = .8), top 10 (score = .6), top 20 (score = .4), outside of the top 20 (score = 0). Combining these scores with the proportion of the five venues found in the neighborhood and dividing this number by 2 will give a neighborhood ranking between 0-1. These rankings will make it easier to visualize how suited each neighborhood is to a person’s lifestyle.

## Conclusion

It is often difficult to determine the best neighborhood to live in when moving to a new city. Venue data when organized by neighborhood, grants insights which can be helpful when making such a decision. A person’s venue interests can be prioritized and compared to the venues available in each neighborhood in order to make a data driven decision.