Boston Neighborhood Analysis

**John Hennessey**

**May 22, 2020**

**I.I 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.

**I.II. 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.

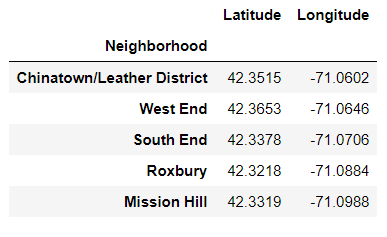
**I.III. Discussion**

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.

## II. Data Acquisition and Cleaning

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 and venue prevalence estimation. The first few rows of the cleaned data frame are shown below (Table 1):

Table 1. Boston Neighborhood Location Data Frame



Using this excel file, I input the coordinates for each neighborhood into Foursquare to attain different venues within 1000m from the center of each neighborhood. The first few rows can be seen below (Table 2). The purpose was to get venues that are close enough to be easily patronized by a person who lives in a certain neighborhood. The venue does not technically need to be in the neighborhood, as long as it is within 1000m, it is relevant. This distance of 1000m was chosen because it tended to be slightly larger than the smaller neighborhoods, but still allowed for larger neighborhoods to gather most of their relevant venues. A map of the Boston Neighborhoods and 1000m radius around each one is shown below (Map 1):

Map 1. Boston Neighborhoods with 1000m radii

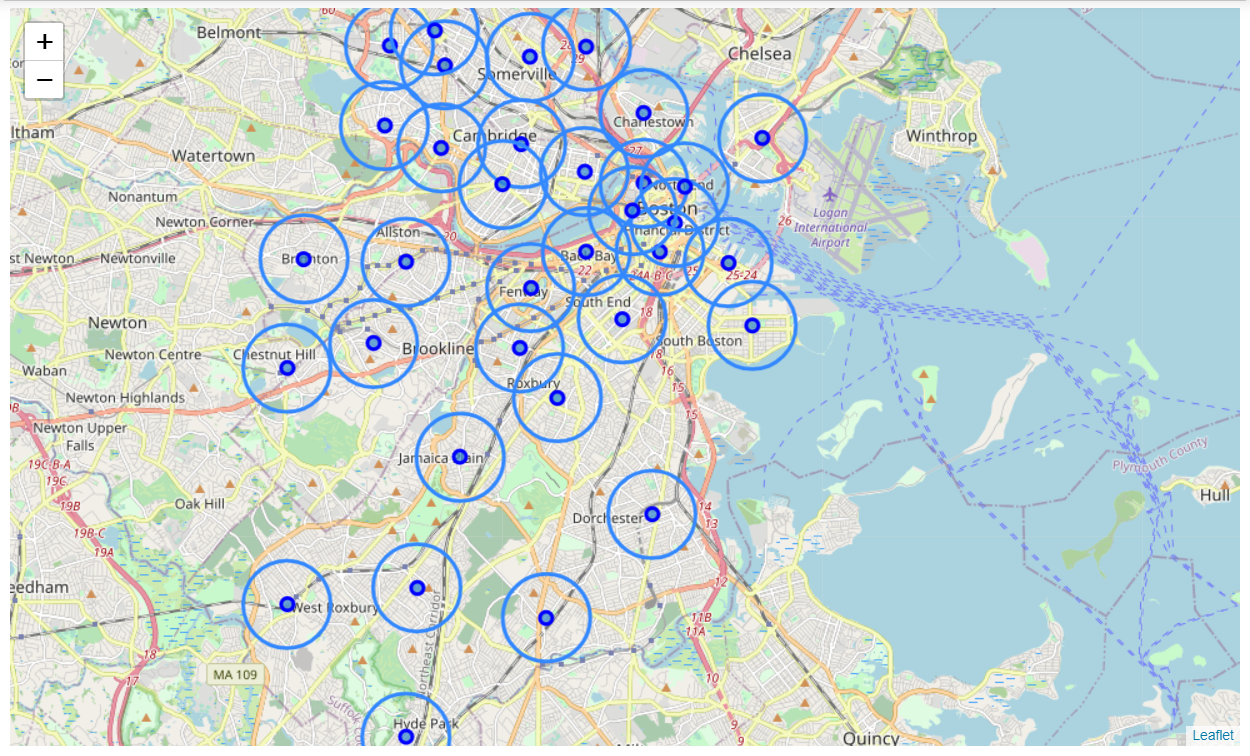
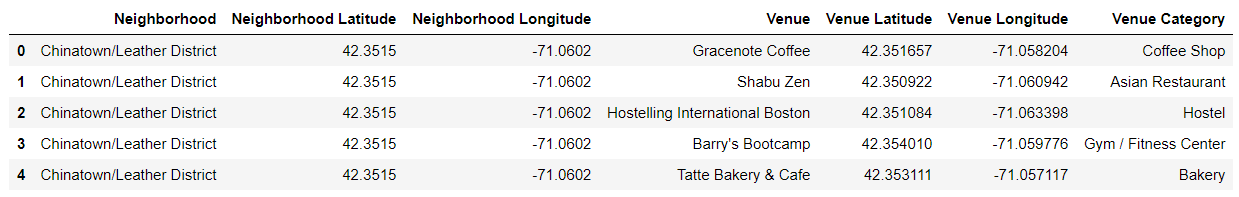


Table 2. Venues Organized by Neighborhood



The Foursquare Data combined with the Neighborhood Data offers insight into the venue options in each neighborhood. Using this information in combination with an individual’s venue preferences, I can determine the best place for an individual to live using a model.

## III. Exploratory Data Analysis

The first data exploring phase required learning about venue types. I downloaded an excel file with all of the venue types and how many of each type were found. I did the same for number of venues per neighborhood to gain more insight into the general amount of opportunity available in each neighborhood. Both of these excel files can be viewed in the Appendix (Table A and Table B respectively). The most popular venues through out Boston are Pizza Place, Park, Coffee Shop, and Italian Restaurant occurring 96, 88, 86, and 80 times respectively. Foursquare capped my venues per neighborhood at 100 and many neighborhoods hit this threshold. North Cambridge, Roxbury, and West Cambridge had the least amount of venues with 27, 27, and 26 respectively while neighborhoods close to and including Downtown all hit the 100 venue threshold.

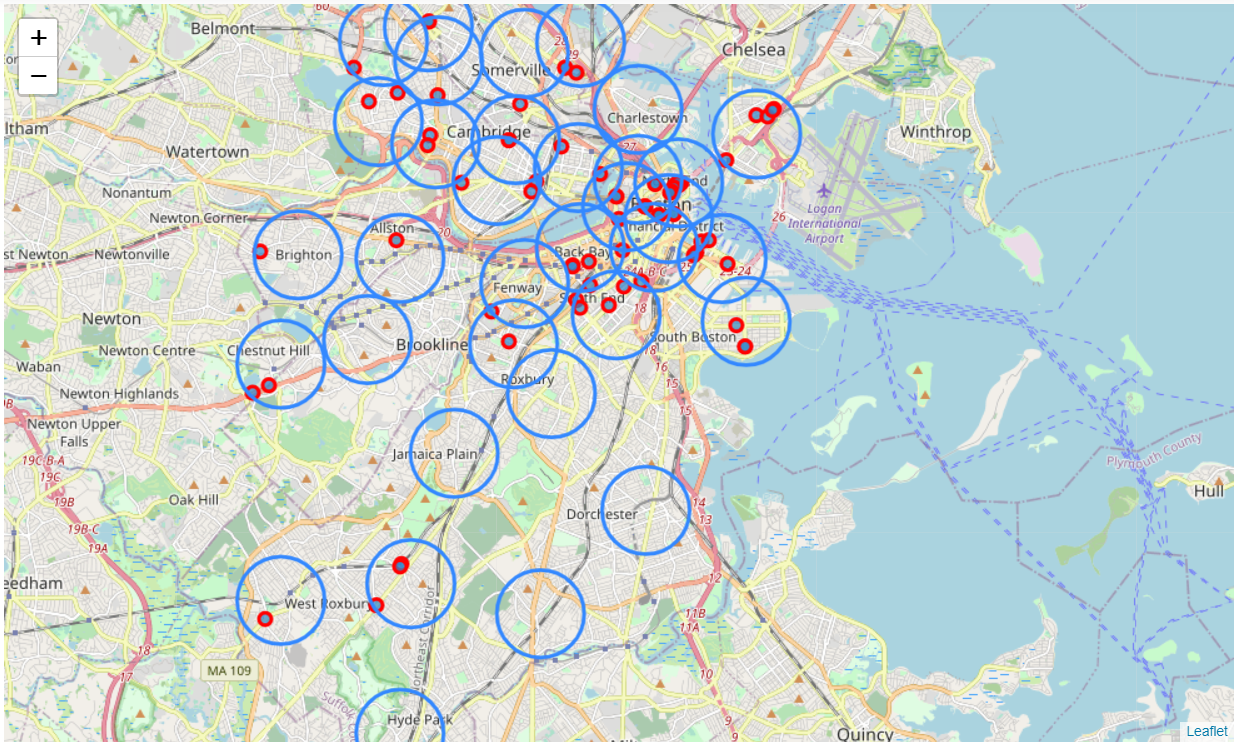
Venues in each neighborhood were organized by type. One hot encoding was performed and the top venues for each neighborhood were placed in a table. A sample is below (Table 3):

Table 3. Sample of Common Venues per Neighborhood



I also put together a map of all Italian Restaurants in Boston with neighborhood radii shown (Map 2). As you will see in the Predictive Modeling section, I perform an example analysis were Italian Restaurant is the top venue category criteria. Most of the Italian Restaurants are near Downtown or the North End (the well-known Italian Neighborhood in Boston). A map of all Italian Restaurants in Boston is shown on the next page:

Map 2. Italian Restaurants in Boston with Neighborhood Radii



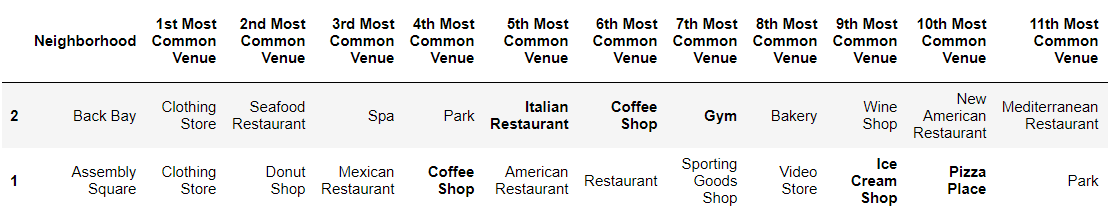
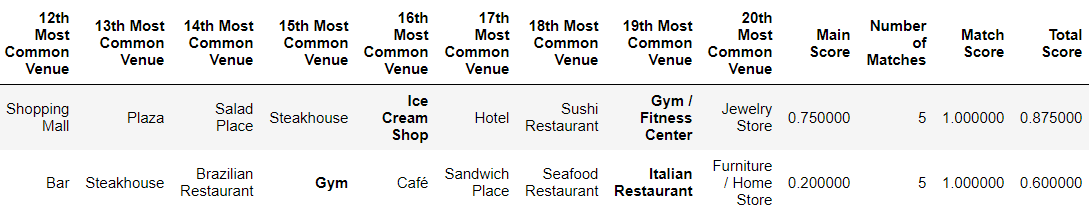
These visualizations give some insight into the differences between neighborhoods in Boston. More analysis is required (and detailed in the next section) to gain actionable and clear results showing the best neighborhoods for a person given their venue category criterion.

**IV. Predictive Modeling**

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. Functions were created to score each neighborhood. These rankings will make it easier to visualize how suited each neighborhood is to a person’s lifestyle.

I asked a friend to give me his top 5 venues which were 'Italian', 'Pizza', 'Ice Cream', 'Gym', and 'Coffee' with his top choice being ‘Italian’. Venue Types that match his criteria are 'Italian Restaurant', 'Pizza Place', 'Ice Cream Shop', 'Gym', 'Gym / Fitness Center', 'Coffee Shop'. Back Bay and Assembly Square both had all 5 venue types in their top 20, but neither had ‘Italian Restaurant’ in their top 3 (Table 4). A split table can be seen below:

Table 4. Neighborhoods with All 5 Venue Category Matches

However, the South End actually scored better than these neighborhoods because it has the top criterion ‘Italian Restaurant’ as its most prevalent venue category. A split table of the top scored neighborhoods for this criterion can be seen on the next page (Table 5) (all neighborhoods were scored and can be seen in the Appendix, Table C):

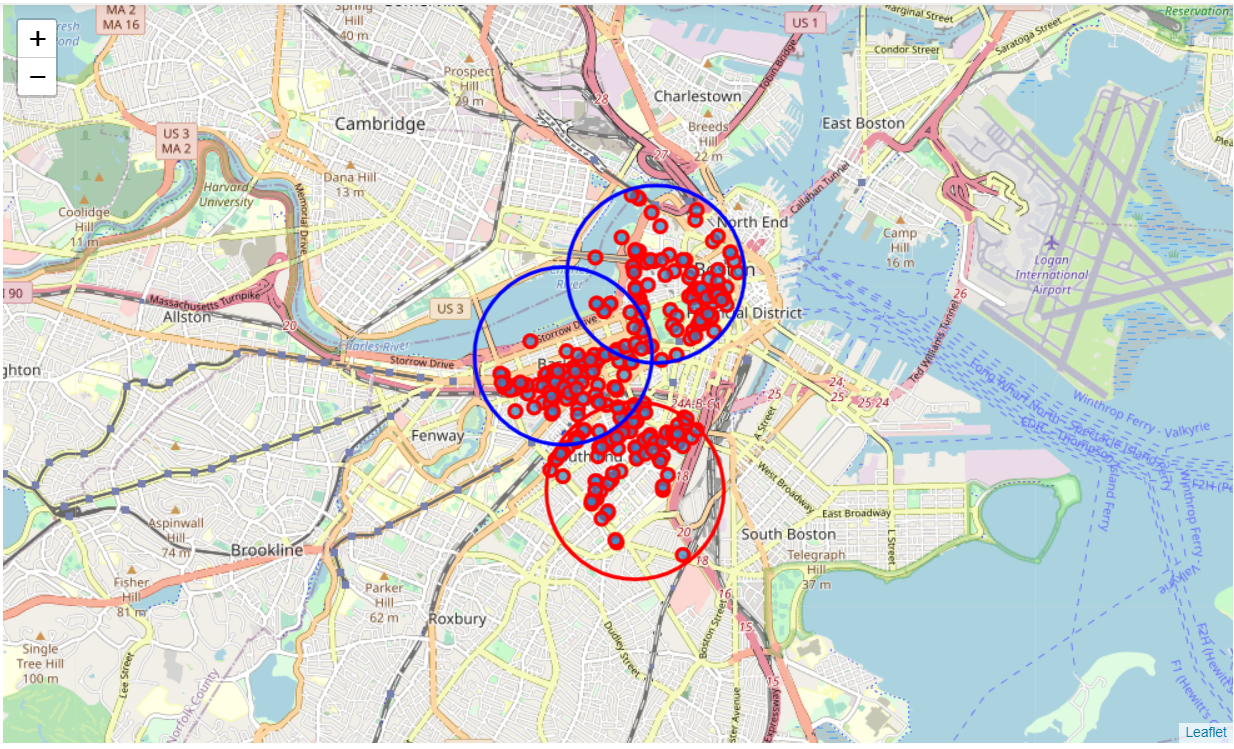
Table 5. Top Neighborhoods Scored with Top Venues Matched





I put together a map to visualize these neighborhoods with all the venues shown below (Map 3). The venues inside the red circle are within range of the top venue (South End). Check out the notebook for a more interactive version of the map. Neighborhood names appear when you click on the circles and venues are labelled as well.

Map 3. Best Neighborhoods with All Venues



It is interesting to note that these neighborhoods are all close enough to have overlapping venues. This make sense especially if key venues (such as Italian Restaurants) appear at the overlaps but that is out of the scope of this project. I did run this model a few times with different criterion and found that, more often than not, top neighborhoods were clustered close to each other. This could indicate that neighborhoods influence the cultures and venues of nearby neighborhoods. This is an interesting potential insight while out of the scope of this project, it could influence future projects that focus on neighborhood interactions and correlations.

## V. Conclusion

It is often difficult to determine the best neighborhood to live in when moving to a new city on your own. Venue data when organized by neighborhood, grants insights which can be helpful when making such a decision. Initial visualizations give more insight into neighborhood trends, but a more refined model is required to gain actionable insights into which neighborhood a person should move to. 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. In this example, with this person’s criterion, this person should look at moving to the South End. If they have other requirements such as distance to work that might affect their decision, they could look at the other top neighborhoods recommended to them such as the Back Bay and Beacon Hill. Regardless, this model helps a person to focus their apartment search on a few neighborhoods and prevents them from getting overwhelmed by the plethora of options.

**VI. Appendix**

Table A. Venues per Category

|  |  |
| --- | --- |
| **Venue Category** | **Count** |
| Pizza Place | 96 |
| Park | 88 |
| Coffee Shop | 86 |
| Italian Restaurant | 80 |
| Bakery | 68 |
| American Restaurant | 62 |
| Café | 59 |
| Seafood Restaurant | 50 |
| Mexican Restaurant | 48 |
| Sandwich Place | 48 |
| Hotel | 47 |
| Donut Shop | 46 |
| Bar | 39 |
| Ice Cream Shop | 38 |
| Grocery Store | 37 |
| Gym | 36 |
| Chinese Restaurant | 33 |
| Liquor Store | 29 |
| New American Restaurant | 29 |
| Sushi Restaurant | 28 |
| Gym / Fitness Center | 27 |
| Gastropub | 25 |
| Pub | 25 |
| Yoga Studio | 25 |
| Clothing Store | 24 |
| Pharmacy | 23 |
| Convenience Store | 22 |
| Greek Restaurant | 22 |
| Thai Restaurant | 22 |
| Bookstore | 21 |
| Historic Site | 21 |
| Restaurant | 21 |
| Japanese Restaurant | 20 |
| French Restaurant | 19 |
| Plaza | 19 |
| Burger Joint | 18 |
| Deli / Bodega | 18 |
| Mediterranean Restaurant | 18 |
| Spa | 18 |
| Steakhouse | 18 |
| Vegetarian / Vegan Restaurant | 17 |
| Asian Restaurant | 16 |
| Gourmet Shop | 16 |
| Indian Restaurant | 16 |
| Tea Room | 16 |
| Breakfast Spot | 15 |
| Vietnamese Restaurant | 15 |
| Art Gallery | 14 |
| Art Museum | 14 |
| Bank | 14 |
| Salad Place | 14 |
| Wine Shop | 14 |
| Diner | 13 |
| Gift Shop | 13 |
| Theater | 13 |
| Trail | 13 |
| Baseball Field | 12 |
| Brewery | 12 |
| Furniture / Home Store | 12 |
| Lounge | 12 |
| Market | 12 |
| Rental Car Location | 12 |
| Burrito Place | 11 |
| Cycle Studio | 11 |
| History Museum | 11 |
| Music Venue | 11 |
| Pet Store | 11 |
| Salon / Barbershop | 11 |
| Wine Bar | 11 |
| Farmers Market | 10 |
| Science Museum | 10 |
| BBQ Joint | 9 |
| Cocktail Bar | 9 |
| Fried Chicken Joint | 9 |
| Middle Eastern Restaurant | 9 |
| Scenic Lookout | 9 |
| Skating Rink | 9 |
| Athletics & Sports | 8 |
| Dive Bar | 8 |
| Falafel Restaurant | 8 |
| Garden | 8 |
| Korean Restaurant | 8 |
| Latin American Restaurant | 8 |
| Ramen Restaurant | 8 |
| Boat or Ferry | 7 |
| Boutique | 7 |
| Comedy Club | 7 |
| Cosmetics Shop | 7 |
| Department Store | 7 |
| Discount Store | 7 |
| Hotel Bar | 7 |
| Jewelry Store | 7 |
| Juice Bar | 7 |
| Playground | 7 |
| Record Shop | 7 |
| Rock Club | 7 |
| Sporting Goods Shop | 7 |
| Beer Garden | 6 |
| Brazilian Restaurant | 6 |
| Caribbean Restaurant | 6 |
| Dessert Shop | 6 |
| Harbor / Marina | 6 |
| Metro Station | 6 |
| Shopping Mall | 6 |
| Sports Bar | 6 |
| Taco Place | 6 |
| Bubble Tea Shop | 5 |
| Comfort Food Restaurant | 5 |
| Dog Run | 5 |
| Food & Drink Shop | 5 |
| Food Truck | 5 |
| Library | 5 |
| Movie Theater | 5 |
| Outdoor Sculpture | 5 |
| Southern / Soul Food Restaurant | 5 |
| Thrift / Vintage Store | 5 |
| Toy / Game Store | 5 |
| Video Store | 5 |
| Aquarium | 4 |
| Belgian Restaurant | 4 |
| Boxing Gym | 4 |
| Circus | 4 |
| Ethiopian Restaurant | 4 |
| Fast Food Restaurant | 4 |
| Fish Market | 4 |
| Food Court | 4 |
| Gas Station | 4 |
| Hobby Shop | 4 |
| Hockey Arena | 4 |
| Indie Movie Theater | 4 |
| Lake | 4 |
| Massage Studio | 4 |
| Miscellaneous Shop | 4 |
| Mobile Phone Shop | 4 |
| Museum | 4 |
| Nightclub | 4 |
| Portuguese Restaurant | 4 |
| Ski Chalet | 4 |
| Smoke Shop | 4 |
| Tapas Restaurant | 4 |
| Accessories Store | 3 |
| Arts & Crafts Store | 3 |
| Bagel Shop | 3 |
| Beach | 3 |
| Big Box Store | 3 |
| Bistro | 3 |
| Chocolate Shop | 3 |
| Comic Shop | 3 |
| Concert Hall | 3 |
| Creperie | 3 |
| Cuban Restaurant | 3 |
| Dumpling Restaurant | 3 |
| Electronics Store | 3 |
| Event Space | 3 |
| Flower Shop | 3 |
| Frozen Yogurt Shop | 3 |
| General Entertainment | 3 |
| Hardware Store | 3 |
| Hotpot Restaurant | 3 |
| Irish Pub | 3 |
| Kids Store | 3 |
| Lingerie Store | 3 |
| Men's Store | 3 |
| Nail Salon | 3 |
| Neighborhood | 3 |
| Noodle House | 3 |
| Opera House | 3 |
| Organic Grocery | 3 |
| Pastry Shop | 3 |
| Performing Arts Venue | 3 |
| Peruvian Restaurant | 3 |
| Poke Place | 3 |
| Pool | 3 |
| Shipping Store | 3 |
| Speakeasy | 3 |
| Tennis Court | 3 |
| Track | 3 |
| Train Station | 3 |
| Whisky Bar | 3 |
| ATM | 2 |
| African Restaurant | 2 |
| Airport | 2 |
| Airport Terminal | 2 |
| Antique Shop | 2 |
| Arepa Restaurant | 2 |
| Baby Store | 2 |
| Beer Store | 2 |
| Bowling Alley | 2 |
| Bus Station | 2 |
| Bus Stop | 2 |
| Business Service | 2 |
| Cajun / Creole Restaurant | 2 |
| Cheese Shop | 2 |
| Church | 2 |
| Cupcake Shop | 2 |
| Dance Studio | 2 |
| Doctor's Office | 2 |
| Food | 2 |
| Home Service | 2 |
| Hot Dog Joint | 2 |
| IT Services | 2 |
| Knitting Store | 2 |
| Laundromat | 2 |
| National Park | 2 |
| Optical Shop | 2 |
| Other Repair Shop | 2 |
| Pet Service | 2 |
| Planetarium | 2 |
| Platform | 2 |
| River | 2 |
| Soba Restaurant | 2 |
| Social Club | 2 |
| South American Restaurant | 2 |
| Sri Lankan Restaurant | 2 |
| Storage Facility | 2 |
| Supermarket | 2 |
| Taiwanese Restaurant | 2 |
| Tibetan Restaurant | 2 |
| Tour Provider | 2 |
| Turkish Restaurant | 2 |
| Wings Joint | 2 |
| Women's Store | 2 |
| Afghan Restaurant | 1 |
| Airport Service | 1 |
| Arcade | 1 |
| Auditorium | 1 |
| Auto Garage | 1 |
| Automotive Shop | 1 |
| Baseball Stadium | 1 |
| Board Shop | 1 |
| Burmese Restaurant | 1 |
| Bus Line | 1 |
| Cambodian Restaurant | 1 |
| Candy Store | 1 |
| Cantonese Restaurant | 1 |
| Casino | 1 |
| College Gym | 1 |
| College Quad | 1 |
| Community Center | 1 |
| Construction & Landscaping | 1 |
| Dim Sum Restaurant | 1 |
| Dry Cleaner | 1 |
| Event Service | 1 |
| Eye Doctor | 1 |
| Field | 1 |
| Food Service | 1 |
| Forest | 1 |
| Garden Center | 1 |
| Gay Bar | 1 |
| German Restaurant | 1 |
| Hostel | 1 |
| Intersection | 1 |
| Jazz Club | 1 |
| Jewish Restaurant | 1 |
| Karaoke Bar | 1 |
| Laundry Service | 1 |
| Martial Arts Dojo | 1 |
| Meze Restaurant | 1 |
| Monument / Landmark | 1 |
| Moroccan Restaurant | 1 |
| Motorcycle Shop | 1 |
| Moving Target | 1 |
| Multiplex | 1 |
| Music Store | 1 |
| Office | 1 |
| Other Great Outdoors | 1 |
| Paper / Office Supplies Store | 1 |
| Pedestrian Plaza | 1 |
| Pilates Studio | 1 |
| Poutine Place | 1 |
| Racetrack | 1 |
| Rental Service | 1 |
| Resort | 1 |
| Shanghai Restaurant | 1 |
| Shoe Store | 1 |
| Shopping Plaza | 1 |
| Skate Park | 1 |
| Smoothie Shop | 1 |
| Soccer Field | 1 |
| Souvenir Shop | 1 |
| Spanish Restaurant | 1 |
| Stadium | 1 |
| Supplement Shop | 1 |
| Szechuan Restaurant | 1 |
| Tattoo Parlor | 1 |
| Taxi | 1 |
| Theme Park Ride / Attraction | 1 |
| Tiki Bar | 1 |
| Tourist Information Center | 1 |
| Tunnel | 1 |
| Video Game Store | 1 |
| Warehouse Store | 1 |

Table B. Venues per Neighborhood

|  |  |
| --- | --- |
| **Neighborhood** | **Count** |
| Allston | 100 |
| Assembly Square | 100 |
| Back Bay | 100 |
| Beacon Hill | 100 |
| Central Square | 100 |
| Chinatown/Leather District | 100 |
| Davis Square | 100 |
| Downtown | 100 |
| East Cambridge | 100 |
| Fenway/Kenmore | 100 |
| Harvard Square | 100 |
| Inman Square | 100 |
| North End | 100 |
| Porter Square | 100 |
| Seaport | 100 |
| South End | 100 |
| West End | 100 |
| East Boston | 89 |
| Brighton | 77 |
| Winter Hill | 69 |
| South Boston | 66 |
| Jamaica Plain | 60 |
| Chestnut Hill | 58 |
| Mission Hill | 57 |
| Roslindale | 53 |
| Brookline Village | 51 |
| West Roxbury | 45 |
| Dorchester | 39 |
| Hyde Park | 38 |
| Charlestown | 37 |
| Mattapan | 32 |
| North Cambridge | 27 |
| Roxbury | 27 |
| West Cambridge | 26 |

Table C1. Final Results – All Neighborhoods Scored (First half)



Table C2. Final Results – All Neighborhoods Scored (Second half)

