# Where Should We Build Brooklyn's Next Donut Shop?

**Introduction**

A nationally known doughnut chain owns several popular doughnut shops throughout the United States. Serving the freshest, most colorful, and most delicious doughnuts, people line up before the shops open to get their favorite tasty treat right out of the fryer. They would love to build their first east coast shop in Brooklyn, New York, which is the owner's hometown, but they are unsure of which neighborhood would be best. The doughnut chain would like to open in a thriving neighborhood, but it does not want to get lost in a sea of other doughnut shops that would be competing for their business. They have decided data science should be used to select the right Brooklyn neighborhood for their newest shop that would attract new customers and ultimately yield great profits. If this project is successful, the national chain can continue to use this analysis branch out into new areas.

Brooklyn is New York City's most populous borough with a population of over 2.6 million residents. It is known for its music, museums, dining, shopping, and nightlife. Its residents are diverse, with 36.5% being immigrants. The median income is $56,015, and 81.6% have a high school education or higher. Using Census business data, there are currently 59 commercial bakeries in Brooklyn that sell donuts. The population per each doughnut shop is 44,663 people. Each of these bakeries also have an average revenue of about 2.3 million dollars. This is good news for our doughnut chain, who could potentially have over 40,000 new customers and sales over 2 million!

**Data**

To determine which borough is best for the new doughnut shop, the following data are used:

- **2014 New York City Neighborhood Names Points File** was used to locate Brooklyn and all of its neighborhoods for mapping purposes. This file contains that latitude and longitude for each neighborhood. This data set will be reduced to only those neighborhoods in the borough of Brooklyn. Using the folium library, these data will be mapped.

- **Foursquare API** data was used to determine the most common venue in each neighborhood. The most common venues were used to determine which neighborhoods do not have doughnut shops, which will guide the national doughnut chain to the neighborhood ripest for development.

- **Census demographic and business data** were used to find out more about the neighborhoods than just the types of venues present to aid in selecting which neighborhood may have customers for the new donut shop.

Once the Brooklyn data were mapped, they were linked with the Foursquare data to search for nearby venues. The number of unique venues was calculated and summed over each neighborhood. The top ten most common venues were calculated for each neighborhood. Using this information, neighborhoods were clustered using the k-means algorithm. Since selecting the right neighborhood for a new doughnut shop is the basis of this analysis, the focus was doughnut shops and how many are located in each neighborhood.
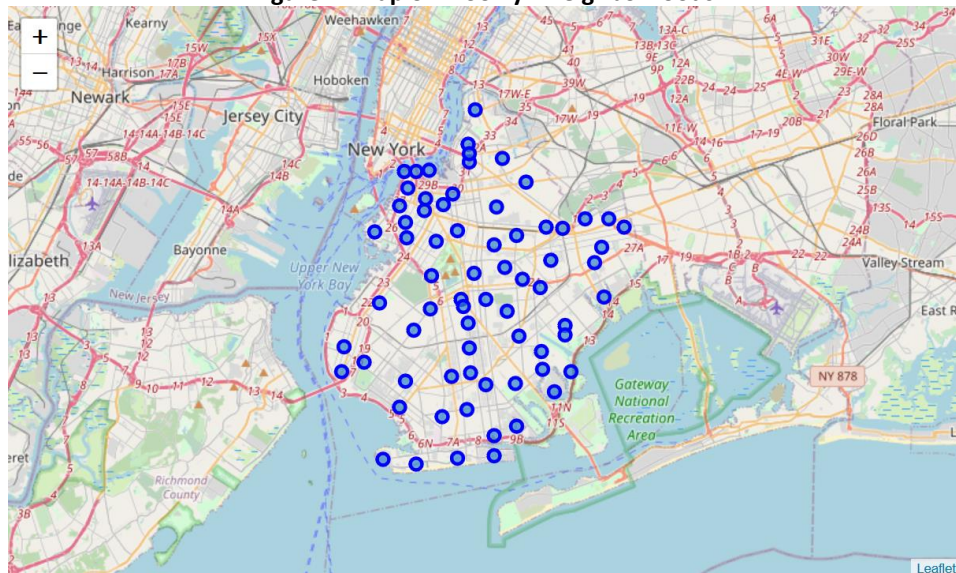
**Methodology**

Using the 2014 New York City Neighborhood Names Points file, the data was restricted to neighborhoods in Brooklyn, NY as shown below in Table 1.

**Table 1. Geospatial Info by Brooklyn Neighborhood**

|   | Borough | Neighborhood | Latitude | Longitude |
|---|---------|--------------|----------|-----------|
| 0 | Brooklyn | Bay Ridge | 40.625801 | -74.030621 |
| 1 | Brooklyn | Bensonhurst | 40.611009 | -73.995180 |
| 2 | Brooklyn | Sunset Park | 40.645103 | -74.010316 |
| 3 | Brooklyn | Greenpoint | 40.730201 | -73.954241 |
| 4 | Brooklyn | Gravesend | 40.595260 | -73.973471 |

Using the Folium library, the 70 neighborhoods were mapped:

**Figure 1. Map of Brooklyn Neighborhoods**



The Foursquare API was used to fetch the top 100 venues for each neighborhood. Using the first five rows of Table 2 below as an example, the Bay Ridge neighborhood has a spa, bagel shop, pizza place, breakfast spot, and bookstore among its venues.

**Table 2. Venues in Each Brooklyn Neighborhood with Their Coordinates**

|   | Neighborhood | Neighborhood Latitude | Neighborhood Longitude | Venue | Venue Latitude | Venue Longitude | Venue Category |
|---|--------------|----------------------|-----------------------|-------|----------------|-----------------|----------------|
| 0 | Bay Ridge | 40.625801 | -74.030621 | Pilo Arts Day Spa and Salon | 40.624748 | -74.030591 | Spa |
| 1 | Bay Ridge | 40.625801 | -74.030621 | Bagel Boy | 40.627896 | -74.029335 | Bagel Shop |
| 2 | Bay Ridge | 40.625801 | -74.030621 | Leo's Casa Calamari | 40.624200 | -74.030931 | Pizza Place |
| 3 | Bay Ridge | 40.625801 | -74.030621 | Pegasus Cafe | 40.623168 | -74.031186 | Breakfast Spot |
| 4 | Bay Ridge | 40.625801 | -74.030621 | The Bookmark Shoppe | 40.624577 | -74.030562 | Bookstore |

Using this data, Foursquare pulled 280 unique venue categories, one of which is a Donut Shop. The number of venues for each neighborhood were counted using one-hot encoding. The mean frequency of each venue type was calculated to determine the top 10 most common venues for each neighborhood.
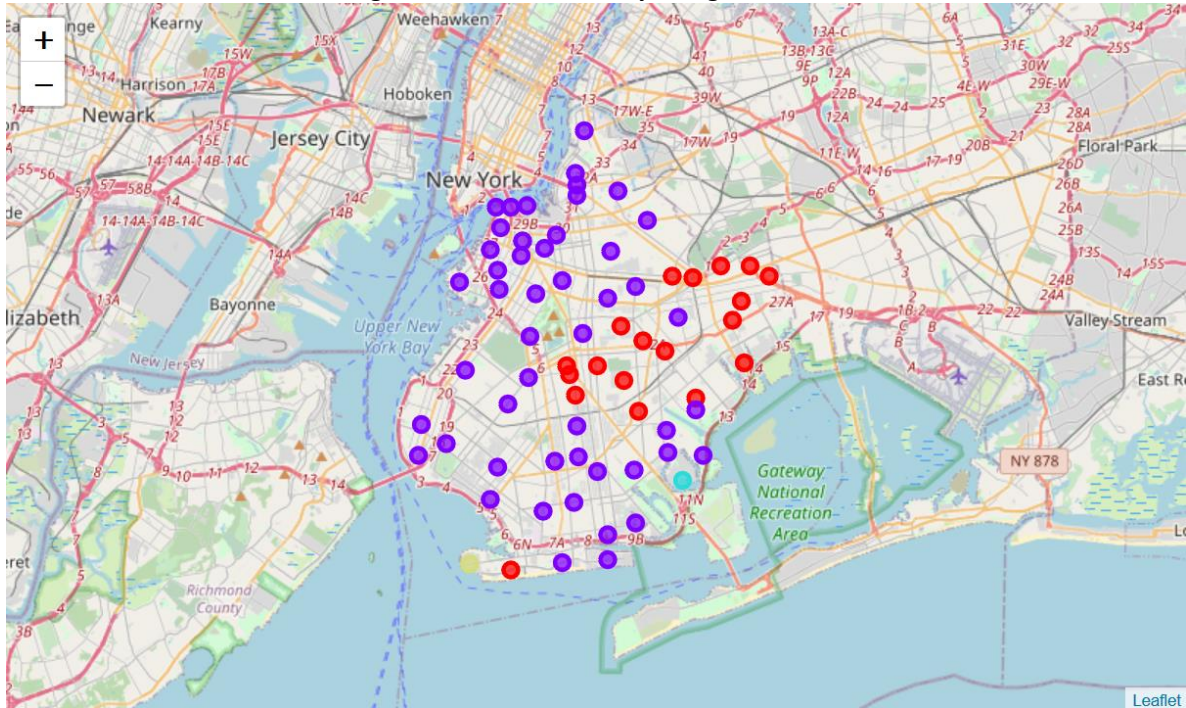
**Table 3. Top 10 Most Common Venues by Brooklyn Neighborhood**

| | Neighborhood | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue | 6th Most Common Venue | 7th Most Common Venue | 8th Most Common Venue | 9th Most Common Venue | 10th Most Common Venue |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | Bath Beach | Chinese Restaurant | Pharmacy | Fast Food Restaurant | Bubble Tea Shop | Gas Station | Sushi Restaurant | Italian Restaurant | Cantonese Restaurant | Park | Liquor Store |
| 1 | Bay Ridge | Spa | Italian Restaurant | Pizza Place | Greek Restaurant | American Restaurant | Bar | Pharmacy | Ice Cream Shop | Playground | Mediterranean Restaurant |
| 2 | Bedford Stuyvesant | Bar | Coffee Shop | Café | Pizza Place | Boutique | New American Restaurant | Gift Shop | Gourmet Shop | Juice Bar | Cocktail Bar |
| 3 | Bensonhurst | Ice Cream Shop | Donut Shop | Chinese Restaurant | Grocery Store | Sushi Restaurant | Italian Restaurant | Flower Shop | Furniture / Home Store | Liquor Store | Spa |
| 4 | Bergen Beach | Harbor / Marina | Donut Shop | Baseball Field | Playground | Athletics & Sports | Women's Store | Factory | Falafel Restaurant | Farm | Farmers Market |

Table 3 shows that in at least two neighborhoods, Bensonhurst and Bergen Beach, that Donut Shops are somewhat popular as the second most common venue.

Next, the k-means algorithm was used to cluster the 70 neighborhoods into four clusters.
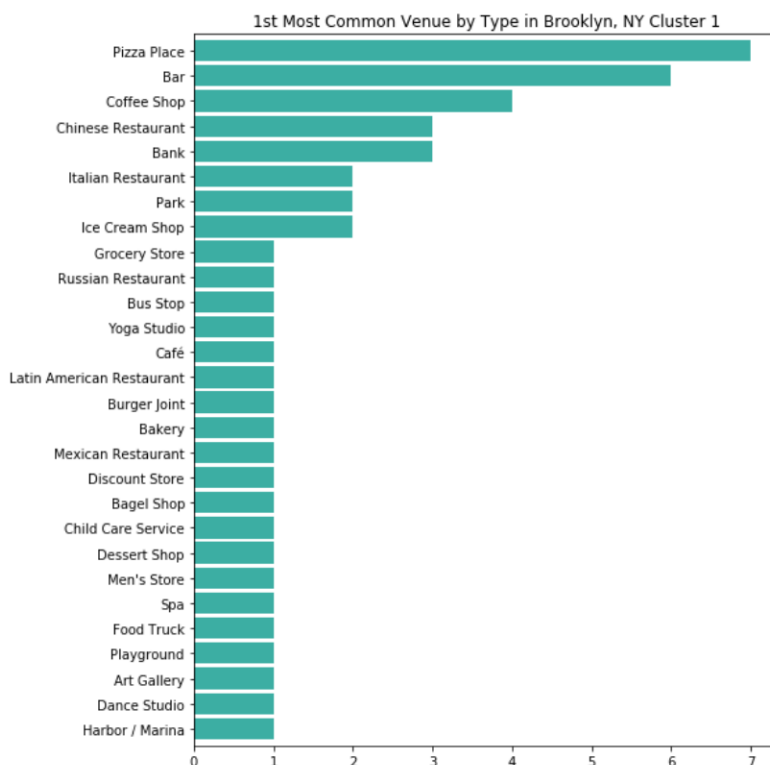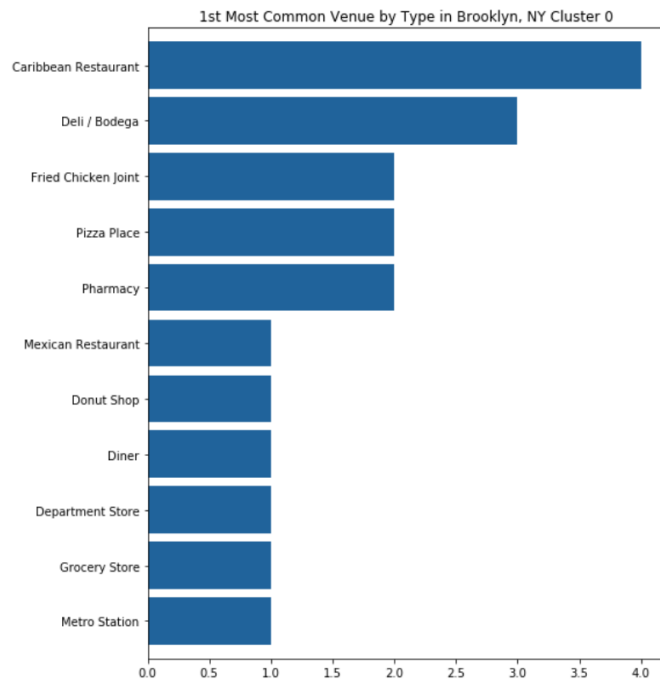
**Figure 2. Clustered Brooklyn Neighborhoods**

**Results**

Now that the neighborhoods have been clustered, the focus now becomes to look at type of venues. If a neighborhood has too many doughnut shops, it would not be a great location to select because there are more competitors nearby.

**Cluster 0** has 19 neighborhoods. This set of first most common venues are varied, where Caribbean restaurants are most common in four of the neighborhoods. A doughnut shop is the first most common venue for just one neighborhood.



1st Most Common Venue by Type in Brooklyn, NY Cluster 0



1st Most Common Venue by Type in Brooklyn, NY Cluster 1

**Cluster 1** has an even more diverse set of first common venues than Cluster 0. Pizza places were most frequent, appearing first in seven of the 49 neighborhoods in this cluster. It's important to note that doughnut shops were not first most common for any of the neighborhoods, which may be good news for the national doughnut chain looking for the perfect location.

Clusters 2 and 3 just had one neighborhood each. The first most common venue in the Mill Island neighborhood (Cluster 2) was a pool, and the first most common venue in the Sea Gate neighborhood (Cluster 3) was a spa.

Although it's important to know the most common venues, the location of all doughnut ia also important. Table 4 shows the number of doughnut shops and all venues for each neighborhood. From this table, we see that at most, some neighborhoods (City Line, Homecrest, and Manhattan Terrace) already have 3 doughnut shops. From a business perspective, it's better to build the shop where there's no similar business to compete against. As a result, City Line, Homecrest, and Manhattan Terrace are not good neighborhoods for our chain to open a new shop.

**Table 4. Doughnut Shops & Venues by Neighborhood**

|  | Neighborhood | Donut Shop | All Venues |
|---|---|---|---|
| 34 | Gerritsen Beach | 0 | 20 |
| 24 | East Flatbush | 0 | 12 |
| 25 | East New York | 0 | 15 |
| 59 | Rugby | 0 | 17 |
| 57 | Red Hook | 0 | 49 |
| ... | ... | ... | ... |
| 3 | Bensonhurst | 2 | 33 |
| 31 | Fort Hamilton | 2 | 69 |
| 43 | Manhattan Terrace | 3 | 26 |
| 14 | City Line | 3 | 27 |
| 39 | Homecrest | 3 | 35 |

**Table 5. Top Neighborhoods Without Doughnut Shops**

|  | Neighborhood | Donut Shop | All Venues |
|---|---|---|---|
| 65 | Vinegar Hill | 0 | 28 |
| 68 | Windsor Terrace | 0 | 29 |
| 67 | Williamsburg | 0 | 34 |
| 64 | Sunset Park | 0 | 37 |
| 55 | Prospect Lefferts Gardens | 0 | 46 |
| 57 | Red Hook | 0 | 49 |
| 22 | Dumbo | 0 | 53 |
| 32 | Fulton Ferry | 0 | 56 |
| 53 | Park Slope | 0 | 56 |
| 30 | Fort Greene | 0 | 59 |
| 54 | Prospect Heights | 0 | 60 |
| 11 | Bushwick | 0 | 72 |
| 1 | Bay Ridge | 0 | 86 |
| 5 | Boerum Hill | 0 | 89 |
| 16 | Cobble Hill | 0 | 94 |
| 21 | Downtown | 0 | 96 |
| 9 | Brooklyn Heights | 0 | 100 |
| 13 | Carroll Gardens | 0 | 100 |
| 62 | South Side | 0 | 100 |
| 49 | North Side | 0 | 100 |

Next the data is restricted to only those neighborhoods without a donut shop and then sorted by the number of all venues. Table 5 shows the top 20 neighborhoods with the most venues. Brooklyn Heights, Carroll Gardens, South Side, and North Side all have 100 venues but no doughnut shops. This list of neighborhoods is the starting point of choosing where to build Brooklyn's next doughnut shop.

**Discussion**

This analysis was needed to determine where to open a new branch of a doughnut chain. Combining geospatial and Foursquare data, the borough of Brooklyn was clustered into four areas using the k-means algorithm. The number of venues in each neighborhood as well as the top 10 most common venues in each neighborhood was also determined. To narrow down the list of possible areas to open, neighborhoods with currently operating doughnut shops were not considered.

Now that a list of the top 20 neighborhoods with the most venues has been prepared, the right neighborhood must be selected. A national chain would not want to open in a mostly residential neighborhood with few venues but rather a more active area to draw in more customers. It would be best to open the new doughnut shop in a neighborhood with an already thriving venue scene that is accustomed to foot traffic and customers. The list was narrowed down to the top four neighborhoods with the most venues, which are Brooklyn Heights, Carroll Gardens, South Side, and North Side. All were found to have 100 venues and located in Cluster 1.

Table 6 shows some quick demographic statistics of each neighborhood using US Census data. Brooklyn Heights has the second smallest population but the largest median income and percentage of residents with a Bachelor degree or higher. The North Side neighborhood has the largest population and the second smallest median income and percentage of residents with a Bachelor degree or higher.

**Table 6. Demographic Stats for the Neighborhoods with the Most Venues**

| Neighborhood | Population | Median Income | Percent Bachelor Degree or Higher |
|---|---|---|---|
| Brooklyn Heights | 62,823 | $124,996 | 74.1% |
| Carroll Gardens | 37,845 | $100,213 | 62.9% |
| South Side | 88,349 | $36,404 | 27.2% |
| North Side | 102,624 | $71,225 | 49.6% |

Using the analysis and demographic data, it is recommended to open the newest doughnut shop in the Brooklyn Heights neighborhood of Brooklyn. It has the highest median income, which means that people will have more disposable income to spend on simple pleasures, which include doughnuts. It also has a large population, larger than the average of all Brooklyn neighborhoods (44,663). Looking at the maps produced in this analysis, its location is directly across from the borough of Manhattan, which could potentially bring in additional customers outside of the neighborhood, since it's only a subway or ferry ride away.

**Conclusion**

This analysis was conducted to find the Brooklyn neighborhood that would be the best to open a new branch of a national doughnut shop chain. Picking the borough of Brooklyn was the first step, but with 70 neighborhoods, picking the best neighborhood was a more challenging task. Using data science tools, geospatial, and Foursquare venue data, that list of 70 neighborhoods were reduced to four: Brooklyn Heights, Carroll Gardens, South Side, and North Side. Each of these neighborhoods have a diverse set of venues that attract a wide variety of customers.

After reviewing demographic and location data, the choice was clear. The wealthy and conveniently located neighborhood of Brooklyn Heights will be the newest location of the most delicious doughnut shop in Brooklyn.