

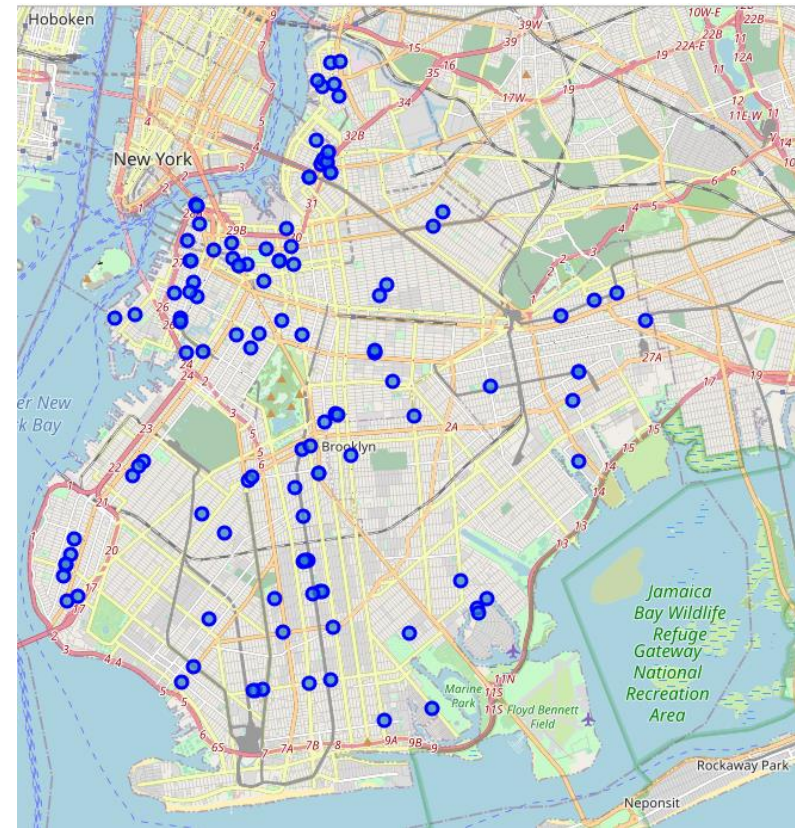
Best Neighborhood in Brooklyn to open a Pizzeria - Capstone

KEEGAN PIERI



Introduction

- Brooklyn is known for having some of the best pizza in the world. With this reputation there can feel like an oversaturation of Pizza options in certain neighborhoods. As shown by the map, there are 123 venues listed as Pizza Places across Brooklyn.
- However, certain neighborhoods may have a high demand for Pizza yet few to no options nearby.
- Understanding which neighborhoods these are would be quite advantageous for someone looking to open a Pizzeria.



Data Description

Sources:

1. Dataset containing each of the 306 neighborhoods across all 5 boroughs in New York City.
This dataset can be found here: https://geo.nyu.edu/catalog/nyu_2451_34572
2. Foursquare API, was used to gather information about the venues in Brooklyn.

Retrieval & Cleaning:

- Venue and neighborhood data was cleaned & merged, as shown below :

	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	Georgian Dream Cafe and Bakery	40.625586	-74.030196	Caucasian Restaurant
4	Bay Ridge	40.625801	-74.030621	Pegasus Cafe	40.623168	-74.031186	Breakfast Spot

Methodology

Part 1: Data Analysis and Manipulation

- The first step was to normalize the venue data by taking the mean of frequency of occurrence of each category.
- The second step was to sort the top 10 venues in each Neighborhood by their frequency, as shown below:

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
Bath Beach	Chinese Restaurant	Bubble Tea Shop	Pizza Place	Cantonese Restaurant	Gas Station	Donut Shop	Fast Food Restaurant	Italian Restaurant	Pharmacy	Peruvian Restaurant
Bay Ridge	Italian Restaurant	Spa	Pizza Place	Bar	American Restaurant	Greek Restaurant	Bagel Shop	Pharmacy	Sandwich Place	Hookah Bar
Bedford Stuyvesant	Coffee Shop	Bar	Café	Deli / Bodega	Pizza Place	Gift Shop	Bagel Shop	Juice Bar	Gourmet Shop	Cocktail Bar
Bensonhurst	Sushi Restaurant	Italian Restaurant	Chinese Restaurant	Ice Cream Shop	Bakery	Donut Shop	Cha Chaan Teng	Noodle House	Sporting Goods Shop	Coffee Shop
Bergen Beach	Harbor / Marina	Baseball Field	Playground	Park	Athletics & Sports	Event Service	Event Space	Factory	Falafel Restaurant	Farm

Methodology

Part 2: Machine Learning with KMeans Clustering

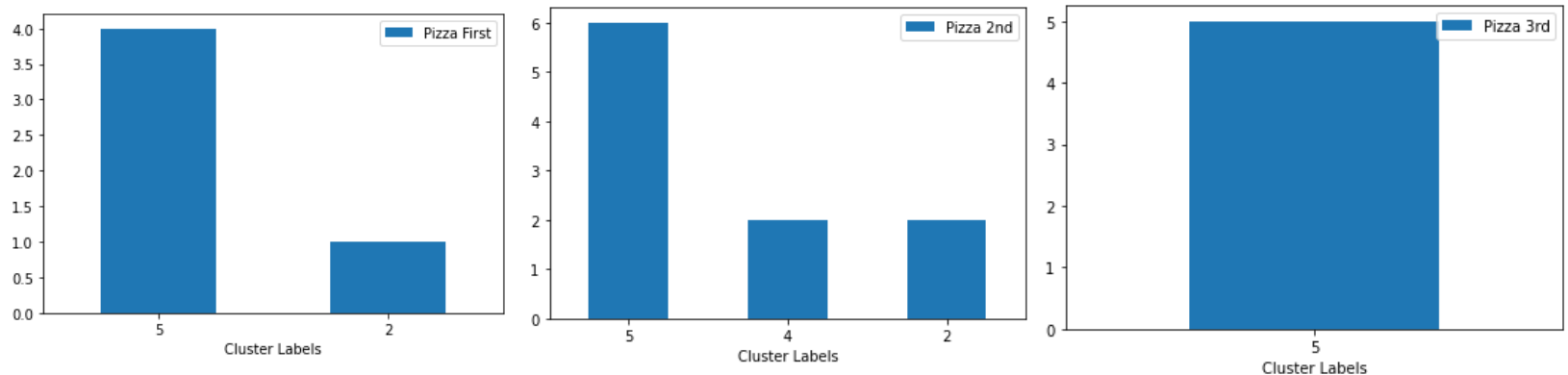
- To determine which type of neighborhood has a high demand for Pizzerias I grouped the neighborhoods into like clusters; allowing me to determine which cluster had the highest preference for Pizzerias.
- This was done using K-Means clustering, with a K value of 7. This sorted each neighborhood into one of 7 clusters and added a column to identify the cluster, as shown to the right:

Borough	Neighborhood	Latitude	Longitude	Cluster Labels
Brooklyn	Bay Ridge	40.625801	-74.030621	5
Brooklyn	Bensonhurst	40.611009	-73.995180	5
Brooklyn	Sunset Park	40.645103	-74.010316	4
Brooklyn	Greenpoint	40.730201	-73.954241	5
Brooklyn	Gravesend	40.595260	-73.973471	5

Results

Part 1: Determine High Demand Pizza Cluster

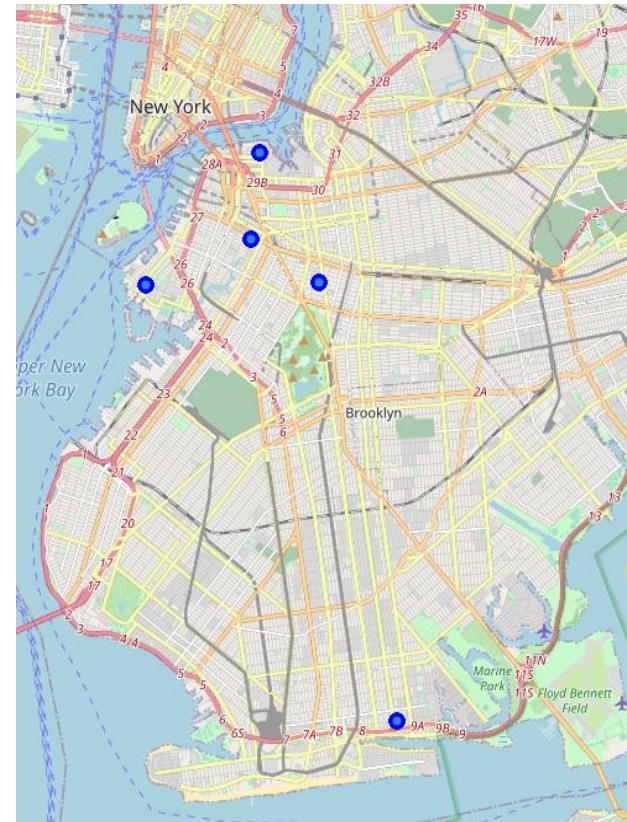
- To determine which of the 7 clusters represented neighborhoods which have a high demand for Pizza, I totaled the number of 1st, 2nd, and 3rd Most Common venues that were “Pizza Place” for each cluster.
- Clearly, Cluster #5 has the highest demand as shown in the charts below:



Results

Part 2: Identify Low Supply Pizza Neighborhoods

- Since we know that Cluster 3 has the highest demand for Pizza, the next step was to look at the neighborhoods to determine where supply might be low
- To do this I removed all Neighborhoods who had a Pizza Place as a Top ten most common venue. Additionally, I removed neighborhoods which had potentially similar venues (Italian Restaurants & Deli/Bodegas).
- **Result:** Sheepshead Bay, Prospect Heights, Red Hook, Boerum Hill, & Vinegar Hill.



Discussion

- This analysis indicates that there are 5 neighborhoods without Pizzerias in Brooklyn that are part of the cluster, as determined by KMeans, where demand for Pizza tends to be high. Given this result it seems these neighborhoods would be promising locations to open a Pizzeria.
- Additional quantitative analysis could be conducted to help determine the suitability of Brooklyn neighborhoods for a new Pizzeria. Different machine learning methods could be used to determine like clusters. Additionally, new factors such as income, population, and ethnicity could be introduced to improve the accuracy of the clustering and determine the appropriate location for a new Pizzeria.
- Additionally, it is important to consider qualitative factors in addition to the quantitative analysis. For instance, the history and success of Pizzerias in these 6 neighborhoods should be examined. Locals in each neighborhood should also be consulted to opine on whether they feel a new Pizzeria would be popular in their neighborhood.

Conclusion

- This analysis indicates that an individual looking to open a Pizzeria is likely to have success in the following neighborhoods: Sheepshead Bay, Prospect Heights, Red Hook, Boerum Hill, & Vinegar Hill.
- An individual should use the beforementioned additional quantitative and qualitative assessments to determine which of these 6 neighborhoods is most suitable for them to open a new Pizzeria in Brooklyn!