Battle of the Neighborhoods

Preferred Borough for Pizza Lovers

Aman Joshi

18-May-20

Contents

1. Introduction

1.1 Background

A lot of people visit NYC for various reasons. NYC is a hub for bankers, IT professionals, tourists etc. It is multi cultural with people from all over the world residing/visiting it. It also has variety of cuisines to satisfy the taste buds of its multi ethnic groups. NYC is divided into 5 Borough namely: Bronx, Manhattan, Brooklyn, Queens and Staten Island. If a tourist who is a big pizza lover wants to visit New York City, he/she would love to reside in that Borough from where the pizza places are easily accessible.

1.2 Problem

This research problem is for such a tourist. A tourist wants to visit New York for a trip. He is a Pizza lover and hence would love to be live in the Borough where pizza is easily accessible. Hence we need to find such a Borough which has maximum Pizza Places and also with high density of Pizza Places. The recommendation should be easy to understand and interpret.

1.3 Stakeholders

This research will benefit all the pizza lovers who will be travelling to New York. This will help them choose the Borough to stay which will give access to Pizza Places options with ease.

2. Data acquisition and cleaning

2.1 Data sources

For this problem I need data of the following:

1) Amount of Pizza places in various boroughs

2) Data of Borough and Neighborhoods with Lat Lang data

Geographical Data will be taken by Foursquare API and using the link https://geo.nyu.edu/catalog/nyu_2451_34572, https://cocl.us/new_york_dataset

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585
5	Bronx	Kingsbridge	40.881687	-73.902818
6	Manhattan	Marble Hill	40.876551	-73.910660
7	Bronx	Woodlawn	40.898273	-73.867315
8	Bronx	Norwood	40.877224	-73.879391
9	Bronx	Williamsbridge	40.881039	-73.857446
10	Bronx	Baychester	40.866858	-73.835798

Borough data with Latitude and Longitude

2.2 Data Cleaning

Check is any borough data is Not Available then those rows need to be deleted from the dataset.

2.3 Feature Selection

- For the borough data we 305 rows and the features selected were Neighborhood, Latitude and Longitude.
- For the Pizza Places in a borough the features selected were Name, Address, Latitude and Longitude. For different boroughs the size of data set varied.

3. Methodology

3.1 Finding the Borough with highest Pizza Places

Firstly the criteria will be to find the Borough with highest pizza places. Higher number of pizza places mean that the demand of pizza is that borough is high and also more competition generally means better quality. For this purpose Foursquare API was used to with Pizza places id as the parameter alongwith the borough names.

3.2 Finding the Borough with highest density of Pizza Places

For a tourist only the higher number of pizza places won't satisfy his/her need, he also will look for the ease to access and hence they should be closely located too in case he wants to have taste of pizzas at multiple pizza places. Hence I chose to find the density of the pizza places. For this I found out the mean coordinate with respect to all the pizza places and then the mean distance of the pizza places from the mean coordinate.

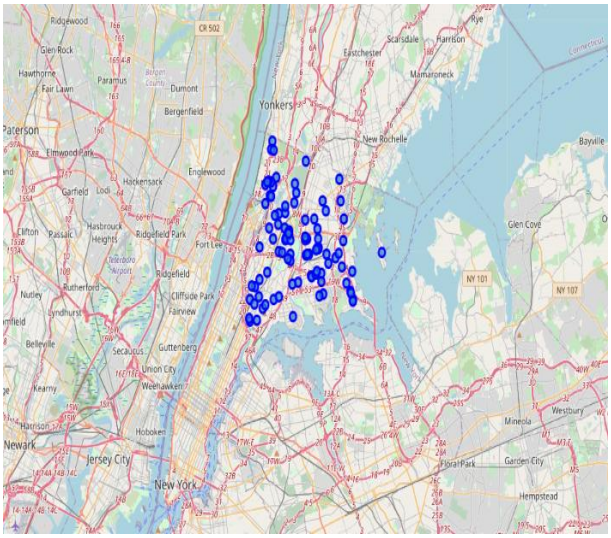
4. Results

4.1 For the number of pizza places

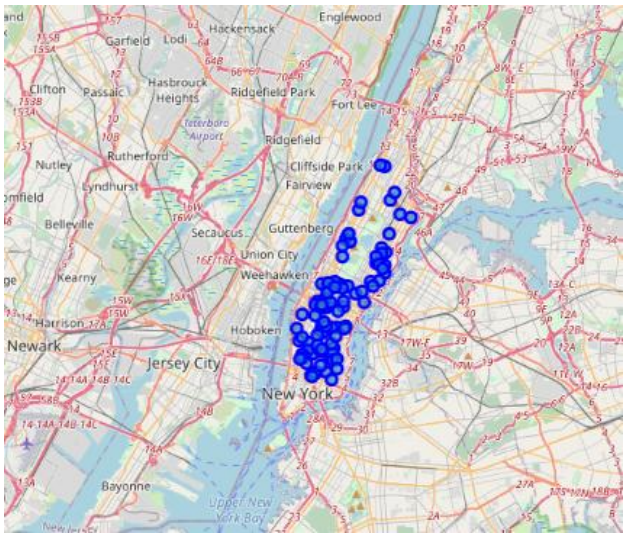
```
Total number of pizza places in Bronx =  158
Showing Top 100
Total number of pizza places in Manhattan =  252
Showing Top 100
Total number of pizza places in Brooklyn =  232
Showing Top 100
Total number of pizza places in Queens =  188
Showing Top 100
Total number of pizza places in Staten Island =  140
Showing Top 100
```

This clearly shows that Manhattan has most number of Pizza Places, followed by Brooklyn and Bronx. Hence the ranking as per the number of pizza places is:

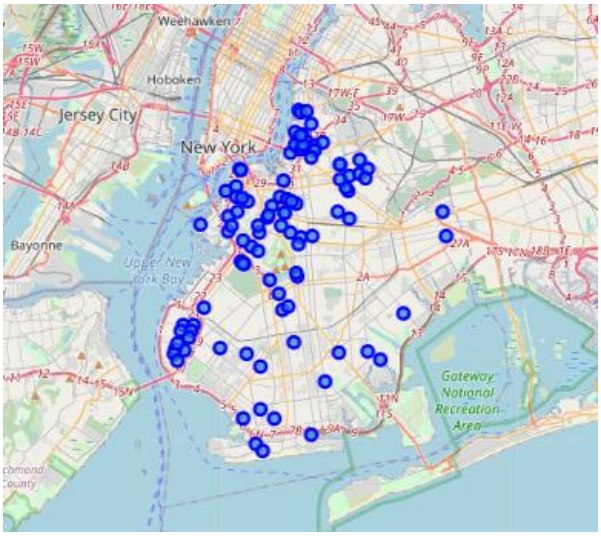
1. Manhattan
2. Brooklyn
3. Bronx
4. Queens
5. Staten Island



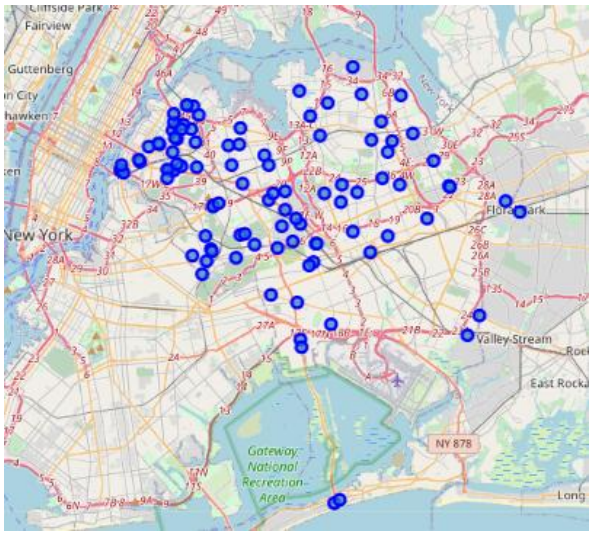
Bronx



Manhattan



Brooklyn



Queens

Staten Island

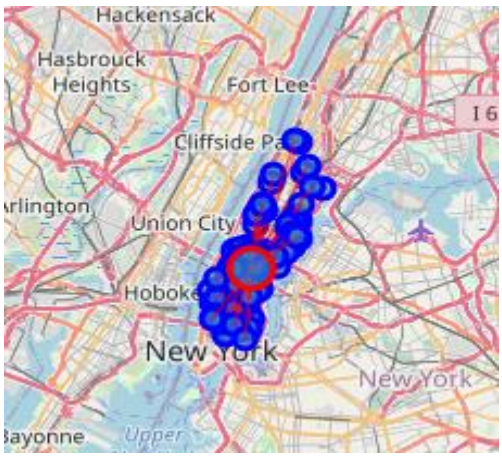
4.2 Highest density of pizza places

```
Bronx
Mean Distance from Mean coordinates
0.03498701585170117
Manhattan
Mean Distance from Mean coordinates
0.025780434583878624
Brooklyn
Mean Distance from Mean coordinates
0.0455394088214408
Queens
Mean Distance from Mean coordinates
0.06189604575236819
Staten Island
Mean Distance from Mean coordinates
0.05600907549251621
```

This clearly shows that Manhattan has the highest density of Pizza Places, followed by Bronx and Brooklyn. Hence the ranking as per the number of pizza places is:

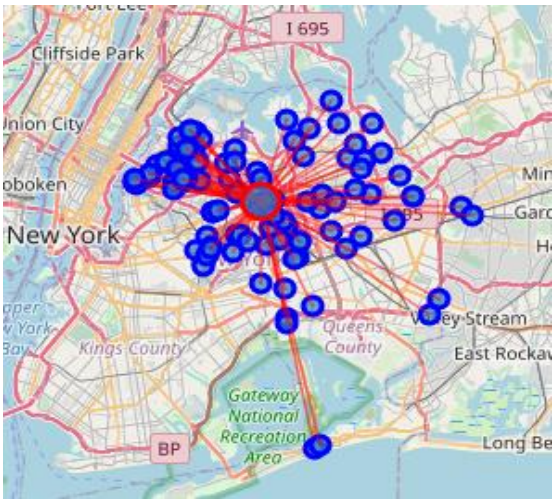
1. Manhattan
2. Bronx
3. Brooklyn
4. Queens
5. Staten Island

Bronx



Manhattan

Brooklyn



Queens

Staten Island

5. Discussion

We can see that there are 2 places in Queens which are quite far away and hence are treated as outliers. So I removed them from the dataset of Queens and again ran the Mean distance check

```
Queens
Mean Distance from Mean coordinates after removing outliers
0.05999405348415403
```

6. Conclusion

We can see that it has reduced but it's still huge and hence won't affect the findings, so the person travelling to NYC and wants to locate himself/herself as per the affinity to Pizza Places and Pizzas, he/she should prefer **Manhattan** because it has most number of the Pizza places and also the density of the location of pizza places is also maximum.

7. Future Scope

This study can help anyone travelling to NYC and wants to live in such a Borough which has easy access to Pizza places as compared to other Borough. Though there are some limitations in this study. It doesn't take into account the ratings of the pizza places in the Borough which can be an important feature if you are thinking of visiting a pizza place. Also the preference of pizza type such as thin crush, cheese burst etc can also be considered.