

IBM Data Science Professional Certificate Capstone Project

This is a capstone project for "IBM Data Science Professional Certificate" as a part of course "Applied Data Science Capstone". The capstone project is meant to define and solve a business problem leveraging Foursquare data, among other publicly available geo-data.

Business Problem

Our customer would like to open an ice cream shop in New York City and needs us -the data consultant- to provide metrics and suggestions on where the best neighborhood is for such a shop.

Data Requirement

Data that is used for segment one is [Demographic and Housing Profiles by Borough](#). The data is sorted as an excel file which needs to be refined and prepared.

Data that is used for segment two is [NYPD Shooting Incident Data](#). The data is sorted as different formats including csv.

Data for segment three is [New York Geo Data](#) as a json file.

Segment four data is acquired through Foursquare as a json file that includes information about venues near specific locations. This section requires a Foursquare account and a customized URL which is provided in the relevant section.

Methodology

We define four metrics and break the project into four correspondent segments as follows:

- 1- Segment One:** Rank boroughs in terms of total and youth population
- 2- Segment Two:** Rank boroughs in terms of safety using shooting data as the base
- 3- Segment Three:** Rank neighborhoods in terms of access by distance from a benchmark
- 4- Segment Four:** Rank neighborhoods in terms of existing ice cream shops parsing Foursquare data

Segments one and two are used to determine the best borough. Segments three and four will refine the neighborhood data to find the best neighborhood.

Results

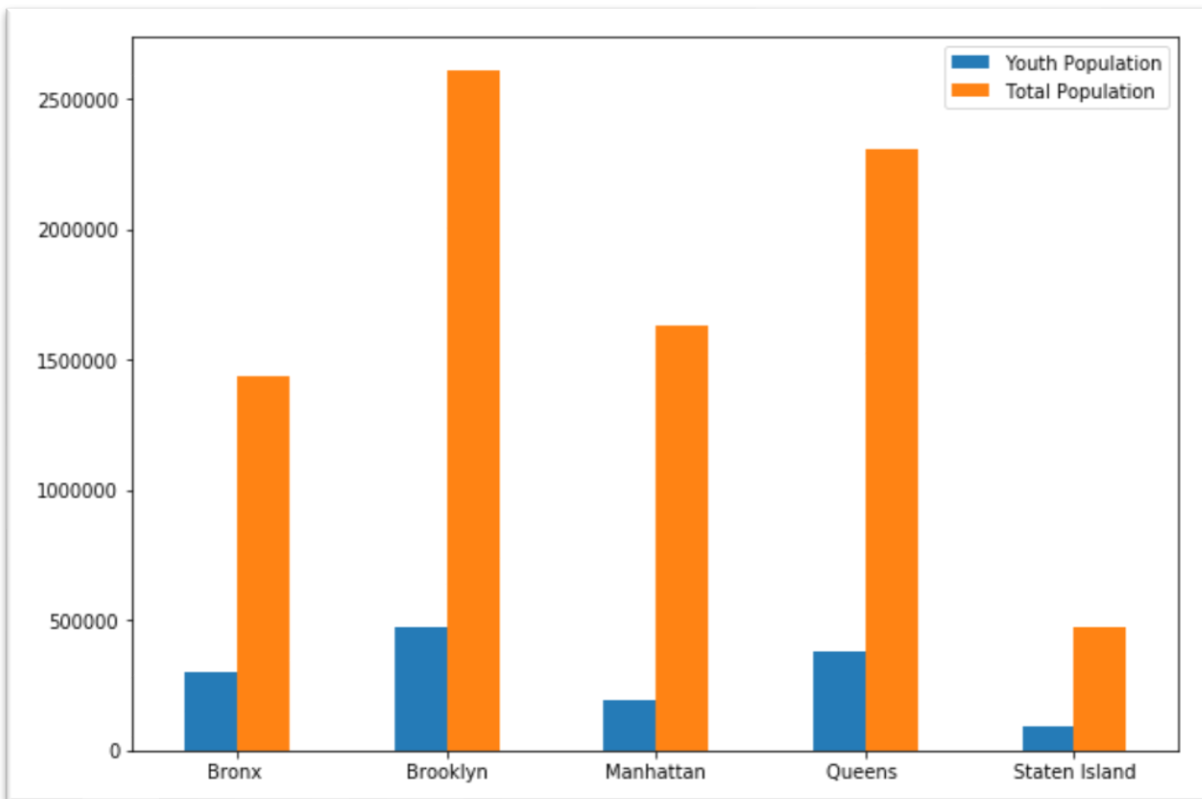
Segment One

Youth Population is defined as ages 5 to 19, therefore we need to sum three relevant rows: 5 to 9 years, 10 to 14 years, 15 to 19 years.

	BRONX	BROOKLYN	MANHATTAN	QUEENS	STATEN ISLAND
0	302283	471044	194937	376278	89522

Plot total and youth population together.

	YOUTH POPULATION	TOTAL POPULATION
BROOKLYN	471044	2606852
QUEENS	376278	2310011
BRONX	302283	1436785
MANHATTAN	194937	1634989
STATEN ISLAND	89522	473324

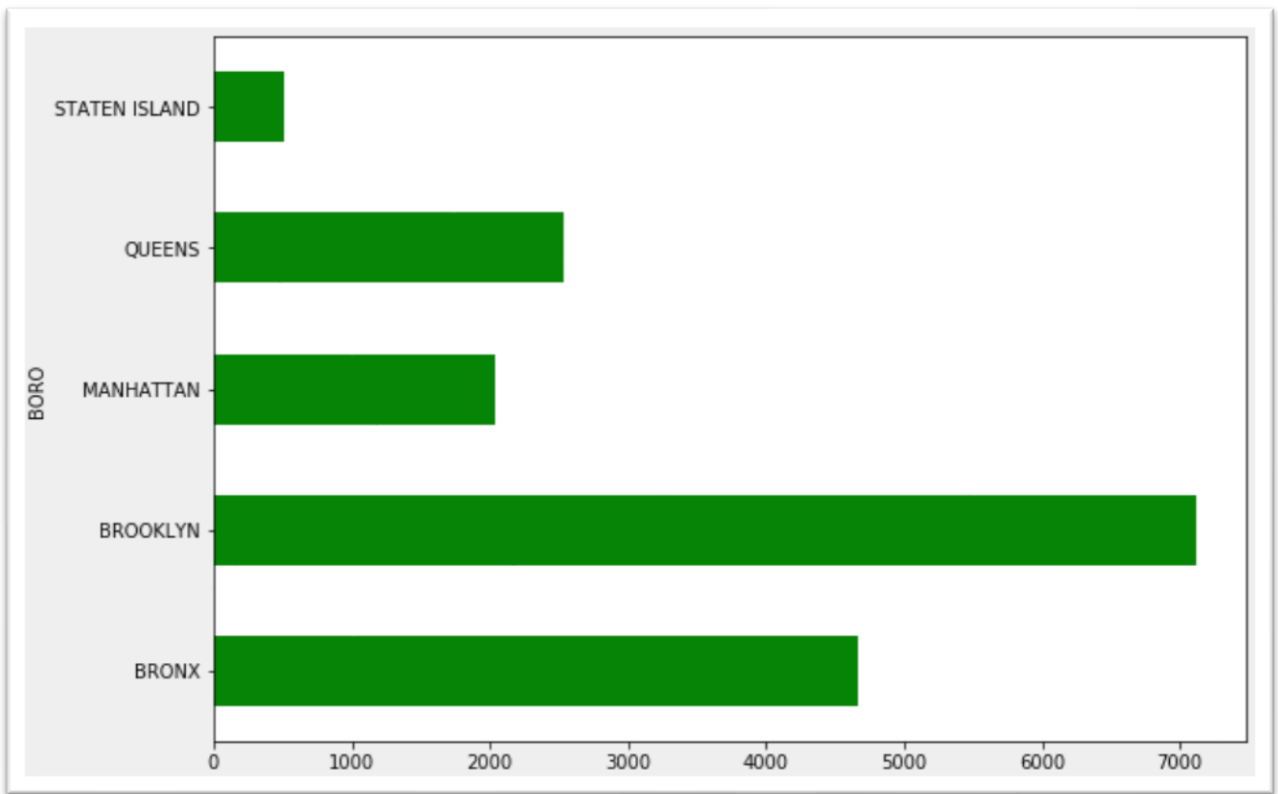


It appears that Brooklyn and then Queens have the highest total and youth population both.

Segment Two

Plot Shooting Incidents.

BOROUGH	INCIDENTS
STATEN ISLAND	505
MANHATTAN	2040
QUEENS	2532
BRONX	4662
BROOKLYN	7123



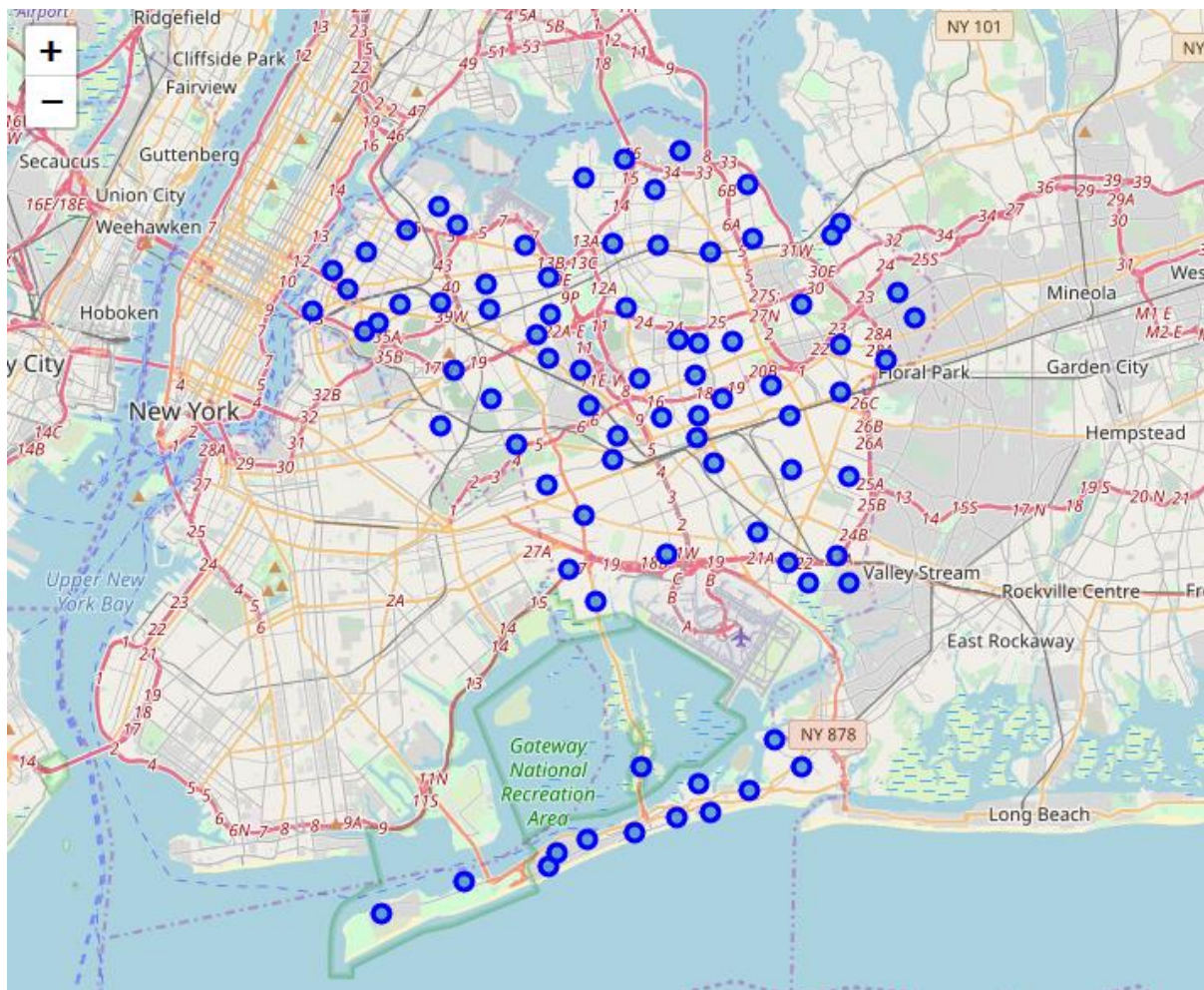
It appears that Queens has significantly less incidents than Brooklyn. Our borough pick is therefore Queens.

Segment Three

Queens Neighborhoods Data

	BOROUGH	NEIGHBORHOOD	LATITUDE	LONGITUDE
0	Queens	Astoria	40.768509	-73.915654
1	Queens	Woodside	40.746349	-73.901842
2	Queens	Jackson Heights	40.751981	-73.882821
3	Queens	Elmhurst	40.744049	-73.881656
4	Queens	Howard Beach	40.654225	-73.838138

Queens Neighborhoods Map



Manhattan Latitude/Longitude

Since Manhattan is not the selected borough, it is selected as the benchmark for this segment as being closest to Manhattan brings a higher chance of tourist and business foot traffic in addition to resident population.

5 Closest Queens Neighborhoods to Manhattan

RANK	BOROUGH	NEIGHBORHOOD	LATITUDE	LONGITUDE	DISTANCETOMAN(KM)
0	Queens	Queensbridge	40.756091	-73.945631	4.044033
1	Queens	Ravenswood	40.761705	-73.931575	4.411265
2	Queens	Long Island City	40.750217	-73.939202	4.939070
3	Queens	Hunters Point	40.743414	-73.953868	5.172607
4	Queens	Astoria	40.768509	-73.915654	5.435539

Segment Four

Foursquare API

Customized URL:

'https://api.foursquare.com/v2/venues/explore?&client_id=X3ITNPE3G1ZYYVMCCKNZNX1ILW5SJRI0FUR3SV2M0A1TWRRW3&client_secret=LFPUDM4AT4BELBVRLS5RYREYM342GTOV0KP1SZDDQFR1W42H&v=20180605&ll=40.7896239,-73.9598939&radius=500&limit=100'

Get Nearby Venues for the Top 5 Neighborhoods

	NEIGHBORHOOD	NEIGHBORHOOD_LATITUDE	NEIGHBORHOOD_LONGITUDE	VENUE	VENUE_LATITUDE	VENUE_LONGITUDE	VENUE_CATEGORY
0	Queensbridge	40.756091	-73.945631	Queensboro Bridge Pedestrian & Bike Path	40.755201	-73.950445	Scenic Lookout
1	Queensbridge	40.756091	-73.945631	Ravel Hotel	40.753989	-73.949244	Hotel
2	Queensbridge	40.756091	-73.945631	Queensbridge Park	40.756701	-73.948653	Park
3	Queensbridge	40.756091	-73.945631	Boulevard Restaurant	40.755648	-73.941834	Spanish Restaurant
4	Queensbridge	40.756091	-73.945631	SUBWAY	40.755557	-73.941276	Sandwich Place

Check Existing Ice Cream Shops in the Top 5 Neighborhood

	NEIGHBORHOOD	NEIGHBORHOOD_LATITUDE	NEIGHBORHOOD_LONGITUDE	VENUE	VENUE_LATITUDE	VENUE_LONGITUDE	VENUE_CATEGORY
50	Long Island City	40.750217	-73.939202	Ice & Vice Pop-Up	40.747730	-73.941025	Ice Cream Shop
195	Astoria	40.768509	-73.915654	Ample Hills Creamery	40.765562	-73.919012	Ice Cream Shop
216	Astoria	40.768509	-73.915654	Cold Stone Creamery	40.765696	-73.913778	Ice Cream Shop

Neighborhood Queensbridge has 0 ice cream shop(s).
 Neighborhood Ravenswood has 0 ice cream shop(s).
 Neighborhood Long Island City has 1 ice cream shop(s).
 Neighborhood Hunters Point has 0 ice cream shop(s).
 Neighborhood Astoria has 2 ice cream shop(s).

It appears that our top neighborhood does not have any competition and is closest to Manhattan.

Conclusion

Therefore, our suggested neighborhood would be "Queensbridge".