

IBM DATA SCIENCE PROFESSIONAL CERTIFICATE CAPSTONE PROJECT

Photo by Bleff (compilation), for proper author attribution see source images / [CC BY-SA 3.0](#)



CONTENTS

- **Business Problem**
- **Data Requirement**
- **Methodology**
- **Results**
- **Conclusions**

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.

[Photo](#) by Evan Kirby / Public Domain



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.

[Photo](#) by Nick Amoscato / [CC BY 2.0](#)



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.

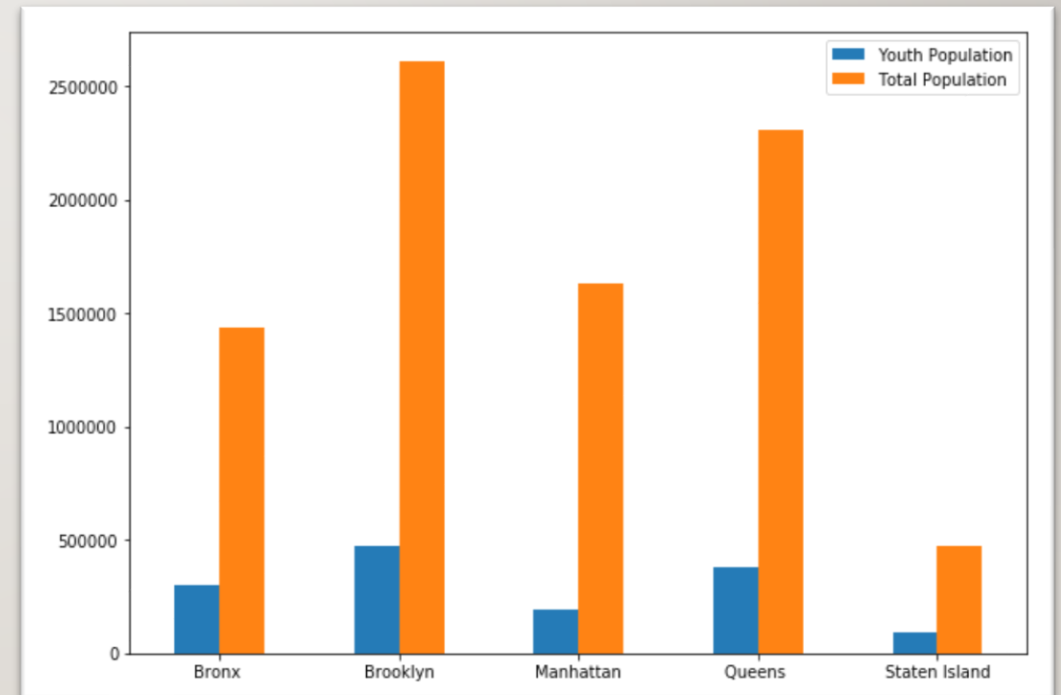
Photo by Dave Lindblom / [CC BY 2.0](#)



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.

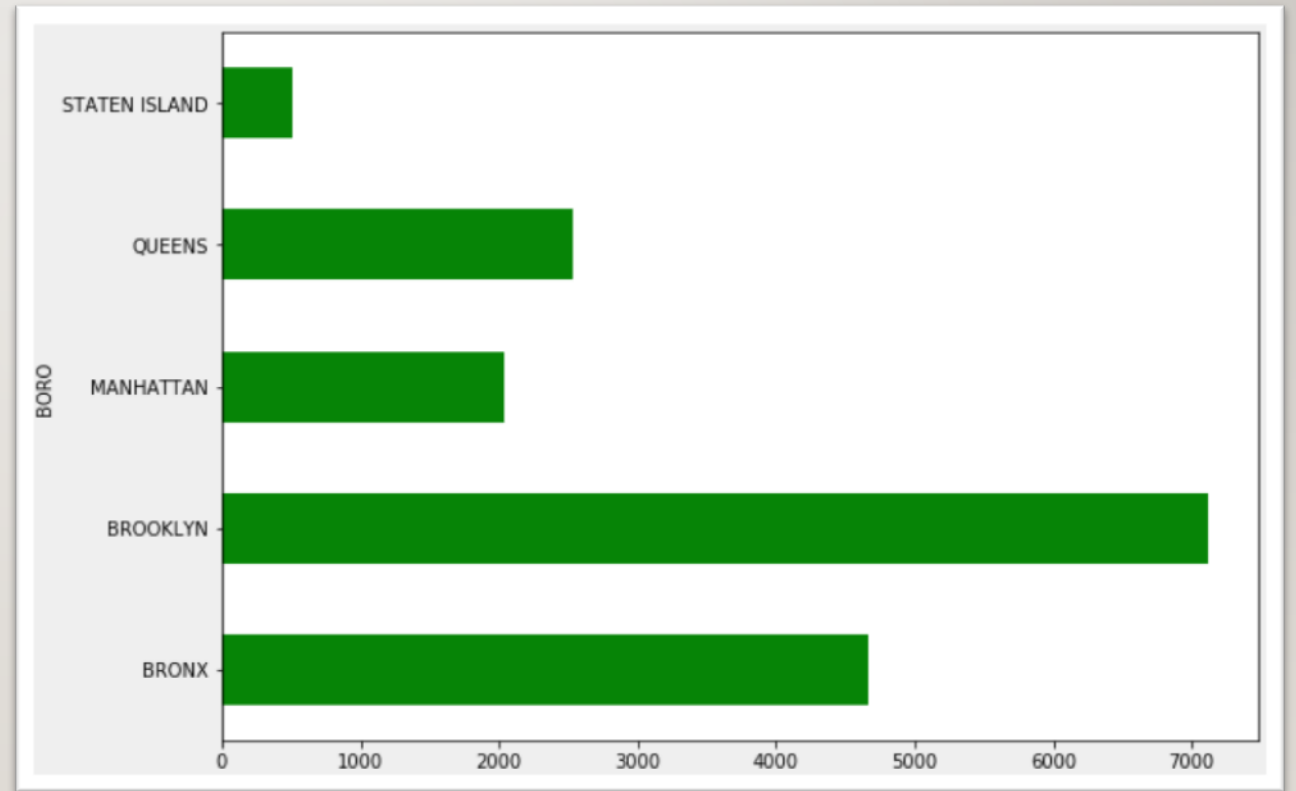
	YOUTH POPULATION	TOTAL POPULATION
BROOKLYN	471044	2606852
QUEENS	376278	2310011
BRONX	302283	1436785
MANHATTAN	194937	1634989
STATEN ISLAND	89522	473324



RESULTS

- **Segment Two**
- Plot Shooting Incidents.
- It appears that Queens has significantly less incidents than Brooklyn. Our borough pick is therefore Queens.

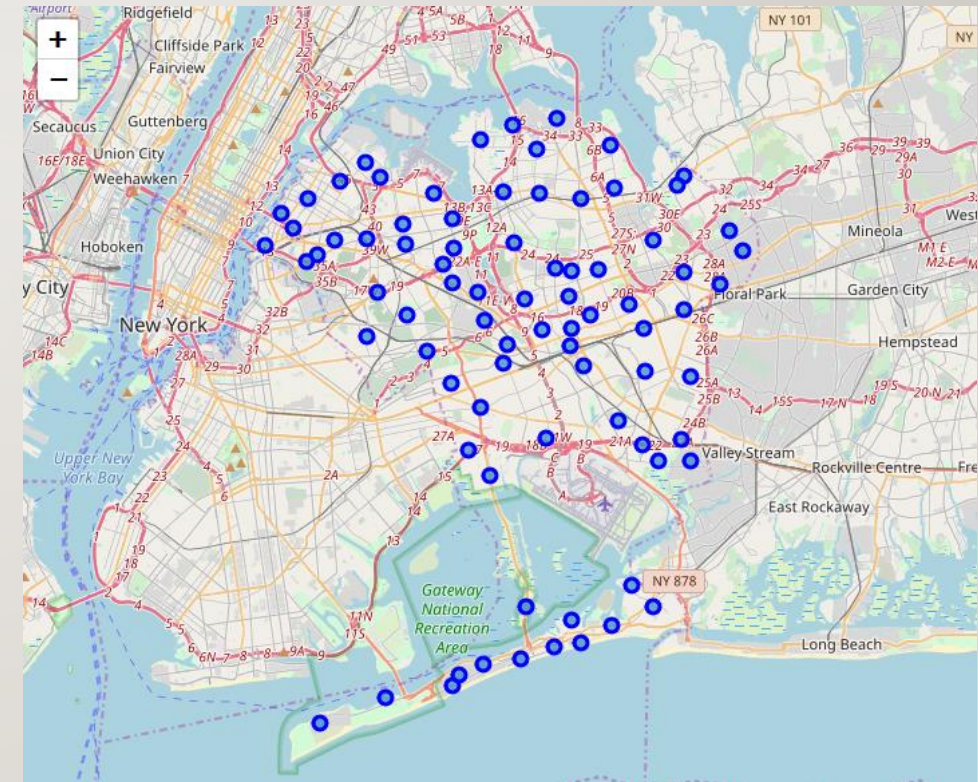
BOROUGH	INCIDENTS
STATEN ISLAND	505
MANHATTAN	2040
QUEENS	2532
BRONX	4662
BROOKLYN	7123



RESULTS

- **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



RESULTS

- **Segment Three**
- **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

RESULTS

- **Segment Four**
- Get Nearby Venues for the Top 5 Neighborhoods from Foursquare

	NEIGHBOR HOOD	NEIGHBO RHOOD_L ATITUDE	NEIGHBOR HOOD_LO NGITUDE	VENUE	VENUE_L ATITUDE	VENUE_LO NGITUDE	VENUE_CA TEGORY
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

RESULTS

- **Segment Four**
- Check Existing Ice Cream Shops in the Top 5 Neighborhood
- It appears that our top neighborhood does not have any competition and is closest to Manhattan.

	NEIGHBO RHOOD	NEIGHBO RHOOD_L ATITUDE	NEIGHBOR HOOD_LO NGITUDE	VENUE	VENUE_L ATITUDE	VENUE_LO NGITUDE	VENUE_CATEG ORY
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

CONCLUSION

- Therefore, our suggested neighborhood would be:

"Queensbridge"