



Battle of the Neighborhoods

Finding a Family-friendly
Neighborhood in New York City

Introduction

Background

- Families need data to help them make decisions regarding where to buy a home, including school quality, safety, location convenience and recreation. This project develops a way for families to make informed decisions when making a move to New York City.

Problem

- This project aims to find the New York City Borough with the lowest crime rate and the best schools for Science, Technology, Engineering and Math (STEM) education, and then find the neighborhoods with the best recreation and convenience options for a family with school-aged children.

Interest

- Any families with school-aged children relocating to New York City would find this information valuable.

Data

Data Sources

The New York City data for this effort is from various data sources.

Crime Data for each New York City Borough

- Data from Kaggle was used for Crime data for New York City from 2013-2015.
 - It includes attributes such as Borough Name, Date of Occurrence, Date of Report and Level of Offense (felony, misdemeanor, violation).
- Source: <https://www.kaggle.com/adamschroeder/crimes-new-york-city>

Education data for each New York City Borough

- SAT scores for New York City schools for the 2014-2015 school year were provided.
 - This data consists of Borough, School Name, Average SAT Math Score, Average SAT Reading Score and Average SAT Writing Score. The SAT Math Score will be used to find the best Borough for STEM education.
- Source: <https://www.kaggle.com/nycopendata/high-schools#scores.csv>

Data

Data Sources (continued)

Location and Convenience Data

- Data from foursquare.com was then used to find venue and recreation (parks) data for New York City.
- This data consists of Borough, Neighborhood, Latitude, Longitude, Venue, Venue Category. Venue Category includes Parks, but also includes Restaurants and Grocery Stores.

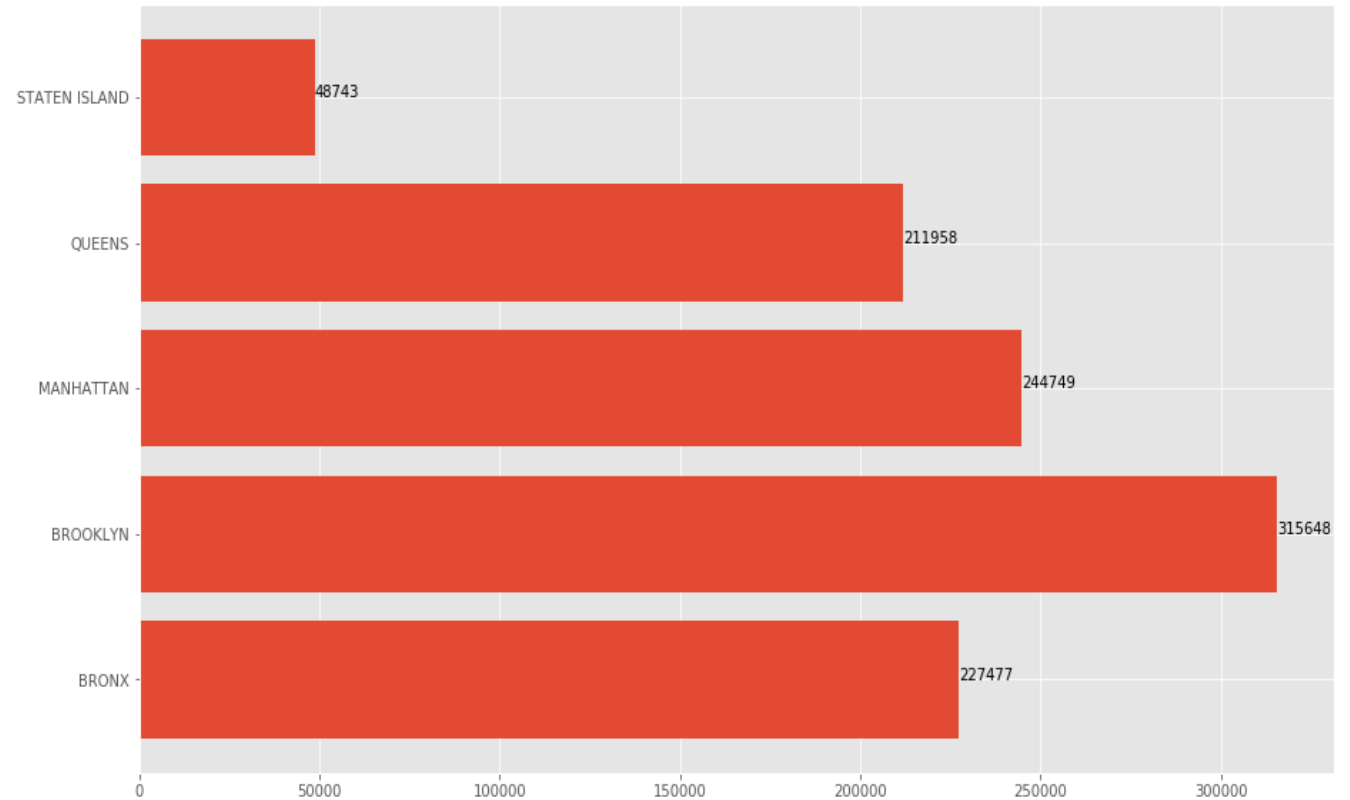
New York City Map data

- Data from a New York University site was used to find data to develop the maps for this project
- https://geo.nyu.edu/catalog/nyu_2451_34572

Methodology

Crime Data

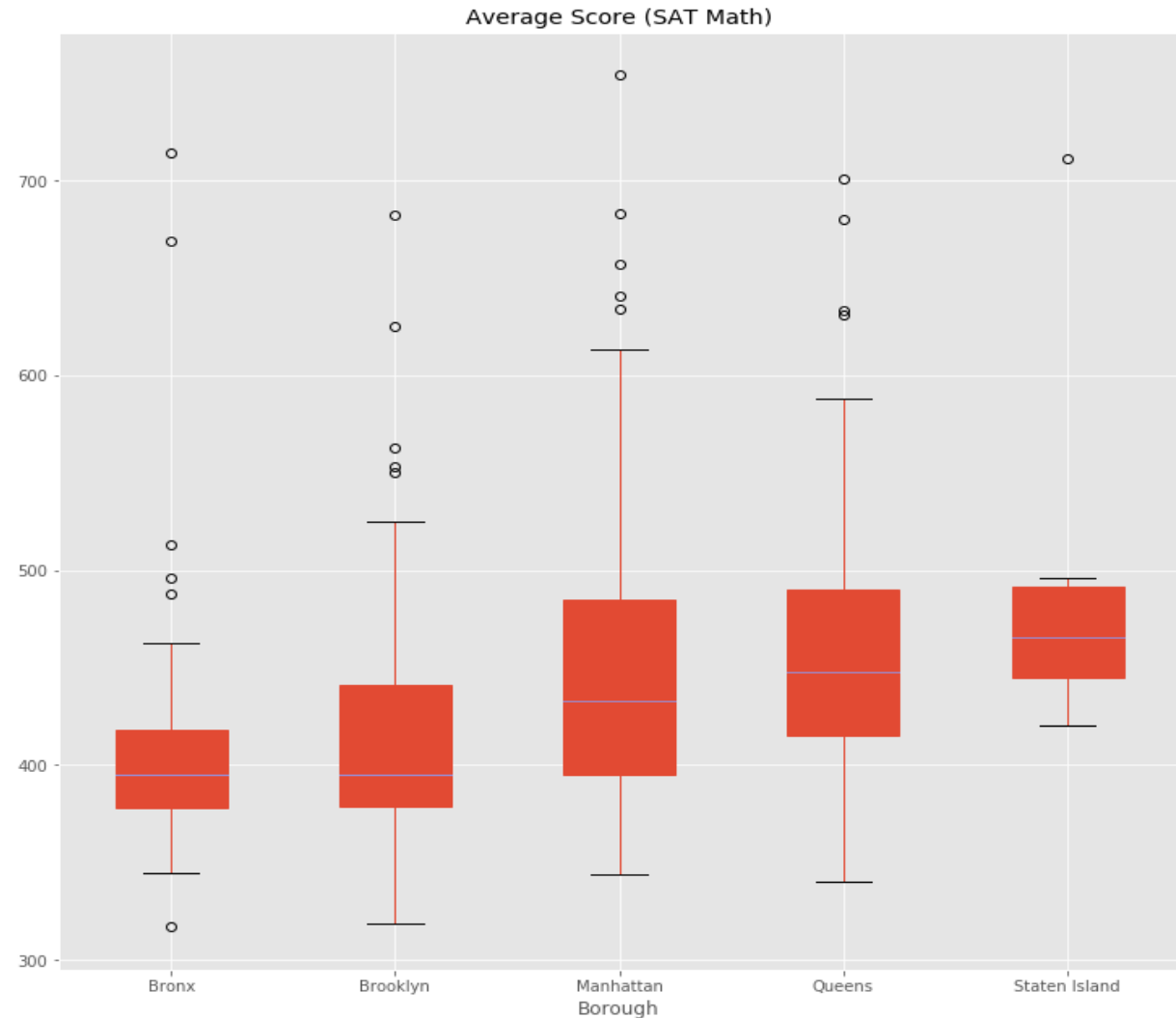
- The crime data was analyzed to determine the Borough with the lowest number of crimes in aggregate. The number of crimes for the data set were compiled and a bar chart was developed. Over the time period studied (2013-2015), Staten Island had the lowest number of crimes, with close to 50,000 compared to over 200,000 for the other Boroughs.



Methodology

Education Data

- The education data was also examined to determine the Borough with the best SAT Math test scores. A box plot was developed using the data from each Borough. It was determined that Staten Island schools had the best scores overall and should be strongly considered for relocation. The inner quartile range (25%-75%) and median were higher for Staten Island, although the maximum scores were higher in other Boroughs. In Staten Island, a parent would be guaranteed solid results.

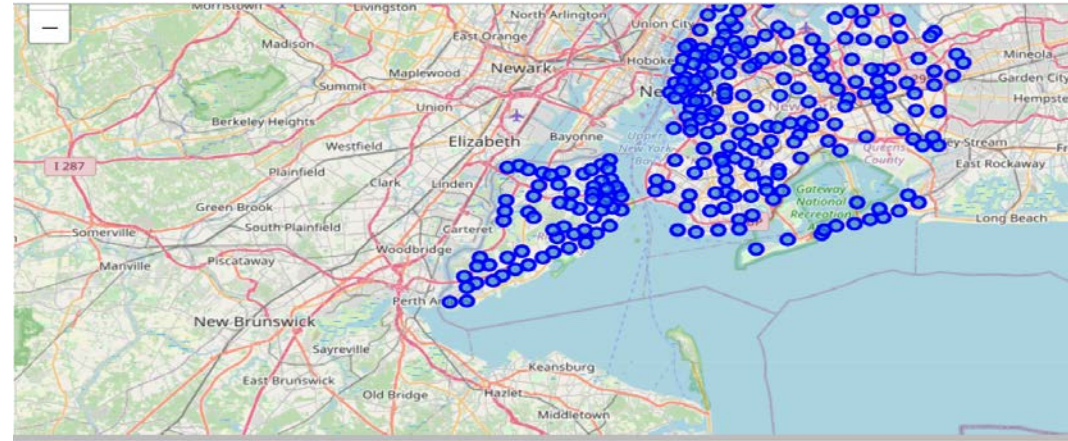


Methodology

Maps

- New York City consists of five Boroughs: Brooklyn, Bronx, Manhattan, Queens and Staten Island.
- Staten island is in the southwest portion of New York City.

New York City



Staten Island



Methodology

- Staten Island is a large island with 63 neighborhoods. K-Means clustering was performed as some initial analysis to determine the neighborhoods with the largest number of venues in proximity to a potential home to make it a convenient location for a family.
- The Foursquare API was used to obtain neighborhood venue data. One Hot encoding was performed on the venue data to allow it to be better consumed by a Machine Learning algorithm.
- In addition to K-Means clustering to help focus the neighborhoods and understand the venue distribution. Data for those neighborhoods was used to perform further analysis. A join of the “Top 10” neighborhoods for venue count table with the neighborhoods that have parks table was performed. Three neighborhoods would be ideal for a family with school-aged children: St. George, Stapleton and Grasmere.

Results

- The Top 10 neighborhoods for venue count are shown in the table below. For a family, these neighborhoods would be extremely convenient.
- The neighborhoods on Staten Island with Parks are shown in the table below. Ten of the twenty-six neighborhoods on Staten Island have parks.

Top 10 Neighborhoods by Venue Count - Staten Island

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Bulls Head	45	45	45	45	45	45
1	Eltingville	42	42	42	42	42	42
2	West Brighton	38	38	38	38	38	38
3	St. George	36	36	36	36	36	36
4	Charleston	31	31	31	31	31	31
5	Stapleton	31	31	31	31	31	31
6	Rosebank	29	29	29	29	29	29
7	Dongan Hills	24	24	24	24	24	24
8	Grant City	24	24	24	24	24	24
9	Grasmere	24	24	24	24	24	24

Neighborhoods with Parks – Staten Island

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Bloomfield	2	2	2	2	2	2
1	Clifton	1	1	1	1	1	1
2	Concord	1	1	1	1	1	1
3	Grasmere	1	1	1	1	1	1
4	New Brighton	2	2	2	2	2	2
5	Randall Manor	1	1	1	1	1	1
6	St. George	2	2	2	2	2	2
7	Stapleton	2	2	2	2	2	2
8	Todt Hill	1	1	1	1	1	1
9	Tompkinsville	1	1	1	1	1	1
10	Travis	1	1	1	1	1	1

Results

- Three neighborhoods were found to be ideal for a family with school-aged children based on crime, education, number of venues and number of parks: St. George, Stapleton and Grasmere.

Bloomfield	False
Clifton	False
Concord	False
Grasmere	True
New Brighton	False
Randall Manor	False
St. George	True
Stapleton	True
Todt Hill	False
Tompkinsville	False
Travis	False

Results

- Through the K-Means clustering analysis it was determined that St. George was close to Sporting Goods Stores, Clothing Stores, Farmers Market, Bus Stations and Scenic Lookout. There was plenty to keep a family busy. However, St. George has many bars, which may not be the best for a family atmosphere. Meanwhile, Grasmere has bus stop, restaurants and grocery stores. Stapleton has restaurants, discount stores and a bank. Given that Grasmere has restaurants and grocery stores, and also has a park, this would likely be the first choice for a family with school-aged children.

	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	St. George	Clothing Store	Sporting Goods Shop	Bar	Italian Restaurant	Monument / Landmark	Burger Joint	Bus Station	Scenic Lookout	Farmers Market	Donut Shop
2	Stapleton	Mexican Restaurant	Pizza Place	Sandwich Place	Restaurant	Discount Store	Bank	Coffee Shop	Spanish Restaurant	Fast Food Restaurant	Motorcycle Shop
3	Rosebank	Mexican Restaurant	Grocery Store	Italian Restaurant	Pharmacy	Cosmetics Shop	Discount Store	Donut Shop	Eastern European Restaurant	Restaurant	Sandwich Place
4	West Brighton	Coffee Shop	Breakfast Spot	Bus Stop	Italian Restaurant	Music Store	Bank	Bar	Pharmacy	Diner	Café
6	Todt Hill	Trail	Park	Yoga Studio	Electronics Store	Food	Flower Shop	Fish & Chips Shop	Financial or Legal Service	Filipino Restaurant	Fast Food Restaurant
8	Port Richmond	Bar	Bus Stop	Rental Car Location	Donut Shop	Basketball Court	Event Space	Food	Flower Shop	Fish & Chips Shop	Financial or Legal Service
26	Graniteville	Food Truck	Bus Stop	Sandwich Place	Grocery Store	Electronics Store	Flower Shop	Fish & Chips Shop	Financial or Legal Service	Filipino Restaurant	Fast Food Restaurant
27	Arlington	Bus Stop	American Restaurant	Home Service	Coffee Shop	Boat or Ferry	Deli / Bodega	Hookah Bar	Food	Ice Cream Shop	Fish & Chips Shop
29	Grasmere	Bus Stop	Bagel Shop	Ice Cream Shop	Nail Salon	Restaurant	Grocery Store	Park	Bank	Bakery	Business Service
32	Midland Beach	Deli / Bodega	Bus Stop	Pet Store	Beach	Basketball Court	Dessert Shop	Bookstore	Restaurant	Liquor Store	Electronics Store

Discussion

- The objective of this exercise was to enable families to make an informed decision when moving to New York City.
- The potential home buyers would be aware of the crime in the area and be able to make a purchase with safety in mind. The home buyers would also be able to understand the quality of the STEM education in a New York City Borough based on the SAT math scores at the high schools in the area. Lastly, the home buyers would be able to consider the convenience of the area based on the general number of venues (including restaurants and grocery stores) and recreation based on the parks in the area.
- Neighborhoods in Staten Island are the most suitable for a family with school-aged children. Within Staten Island, the neighborhoods of St. George, Stapleton and Grasmere should be considered. Given that Grasmere has parks, restaurants, grocery stores and a number of other family-friendly venues, it would be ideal.



Conclusion

- This project helps a family relocating to New York City an understanding of the safety of the Boroughs and the quality of the schools in each Borough.
- This study provides the ability to make a neighborhood selection based on the convenience of venues such as grocery stores and restaurants as well as recreational access through parks.
- In the future, this project could more closely examine the school rankings by neighborhood and include non-STEM scores in the analysis.

