Final Report

The Best Neighborhood in New York City

CONTENT

01

INTRODUCTION

1.1 Background Information

Moving to a new city or country is always intimidating, and people may be unsure of their choice of neighborhood to live in. There is a plethora of decision variables one has to consider when choosing the perfect neighborhood. It could be the place interests, food locations, crime rates, housing prices etc.

The purpose of this report is to allow people to make smarter and more informed decisions when selecting the neighborhood that is best fit with their interests, budgets and concerns.
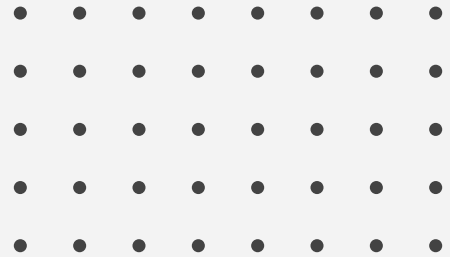
1.2 Location

Our focal point will be **The Big Apple, New York City.** New York City is one the most ethnically diverse, commercially driven and most attractive urban centre in the country. It is not a shock that people would wish to shift to this bustling city.

1.2 Target Audience

People migrating to the state of New York City, and are unsure of the neighborhood to buy a property at. These would mainly be people who wish to save as much as they can, but still live wish to enjoy a good lifestyle and ensure their security in their new homes.

02
DATA

Neighborhood Coordinates:

In order to segment and explore the neighborhoods, we can extract the data of New York City neighborhoods from the NYU Spatial Data Repository. Here, we can obtain:

1. Borough
2. Neighborhood
3. Latitude
4. Longitude

Crime Data:

Another major deciding factor for many residents are crimes rates within the area. Thus, I have retrieved the NYPD complaints dataset from NYC OpenData. This is a very large dataset, consisting of 108058 rows, thus we reduced this to be able to work with the data easier. This will give us a basic overview of crimes in boroughs in overall. We will only need the columns:

1. BORO_NM
2. Latitude
3. Longitude

Venues and Places of Interest:

This can be obtained from the FourSquare API. We will be obtaining:

1. Neighborhood
2. Neighborhood Latitude
3. Neighborhood Longitude
4. Venue
5. Venue Latitude
6. Venue Longitude
7. Venue   Category

Subsidized Housing Prices

The prices of properties is a major factor in one's decision in moving into a neighborhood. Thus, I extracted this data from CoreData.nyc to facilitate in our analysis. This data will be cleaned and obtain only:

1. boro_name
2. accessed_value
3. res_unit

03
METHODOLOGY

3.1 Geocoders

This project required the latitude and longitude coordinates of New York City and it's multiple Boroughs in our data analysis and visualizations. The Geocoders Nomatim function allowed for ease and accurate retrieval of these necessary coordinates.

3.2 Folium

All clustering and bubble map visualization was done with the help of Folium to generate a map using OpenStreetMap technology. With the add circle markers function, I was able to represent clusters in differing colours, as well as different circle radius to represent the magnitude of the data.

3.3 Bar Chart Sub-Plot

We made use of a 2 way Bar Chart Sub-Plots to compare the analysis for our Crime Data and Housing Data. This allowed us to see our data clraly in one plane, which aids in the comparison we needed to conduct.

3.4 FourSquare API

One major decision variable in choosing the best neighborhood was the venue categories in its vicinity. The FourSquare API enabled us to retrieve this dataset matching the neighborhood data we already have.

3.5 One-Hot Coding

One-Hot Coding is a process which converts data into dummy variables, assigned a 1 or 0. This was a necessary step in listing the top 5 most common venues in each neighborhoods, as well as our clustering process.

3.6 K-Means Clustering

We needed to find similarities within the neighborhood datset, and not base our recommendation only based on the largest amount of venues within the vicinity. Thus, to form these clusters, we trained and made use of the K-Means Algorithm.
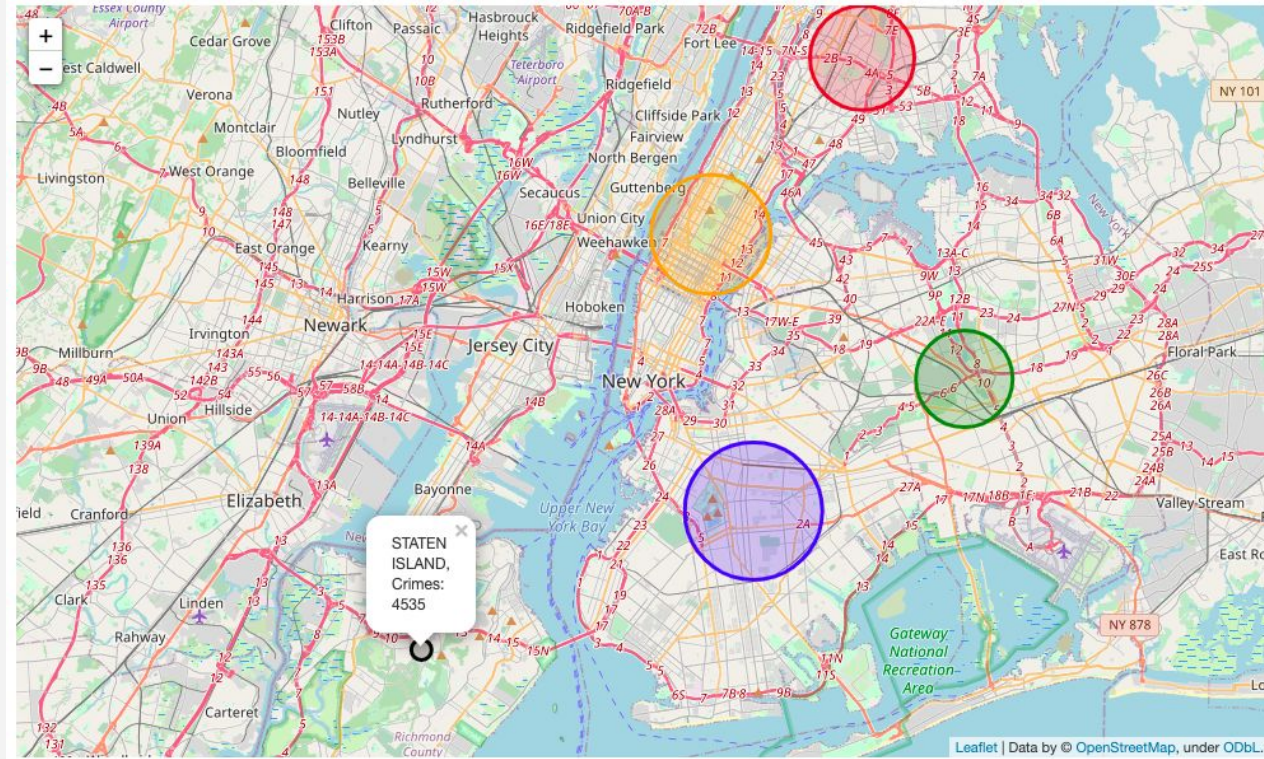
3.7 Top Venue Categories in each Cluster

The variety and category of venues is also an important aspect in our analysis in choosing the best venue. Thus, we analysed the characteristics of these clusters according to the kinds of categories it has.
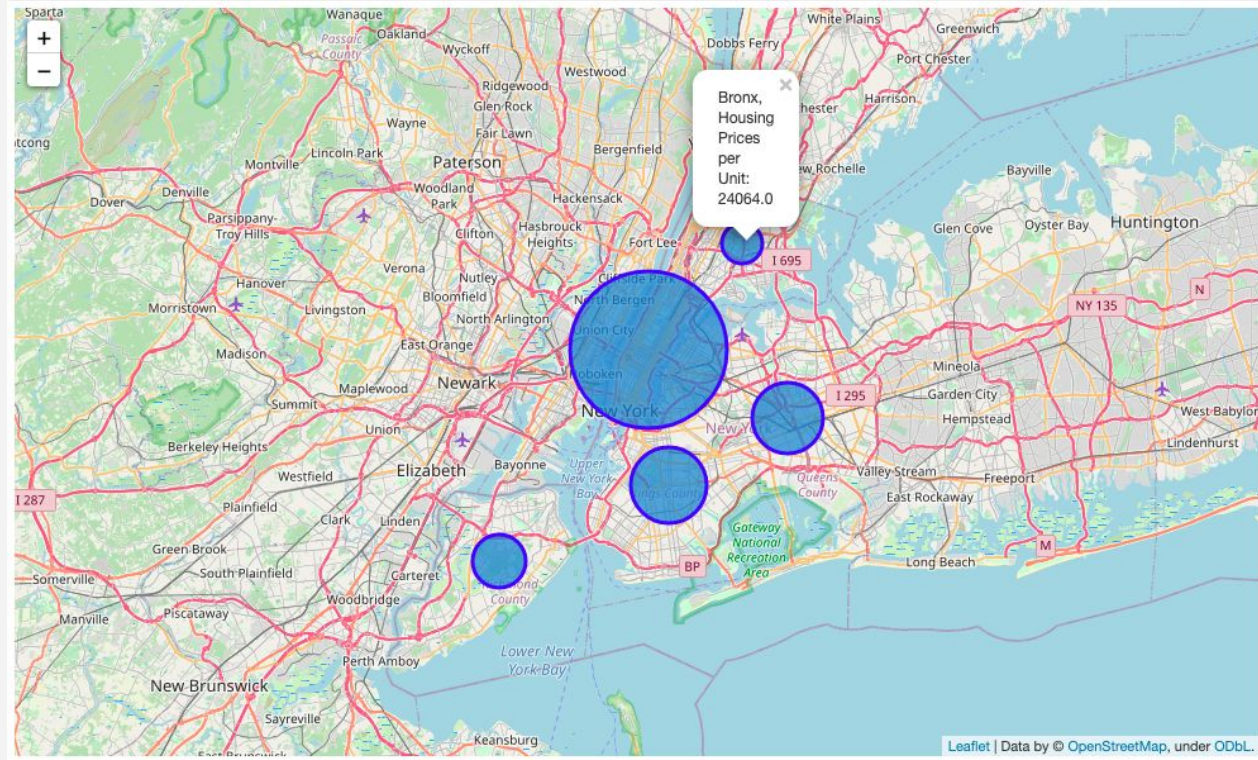
04
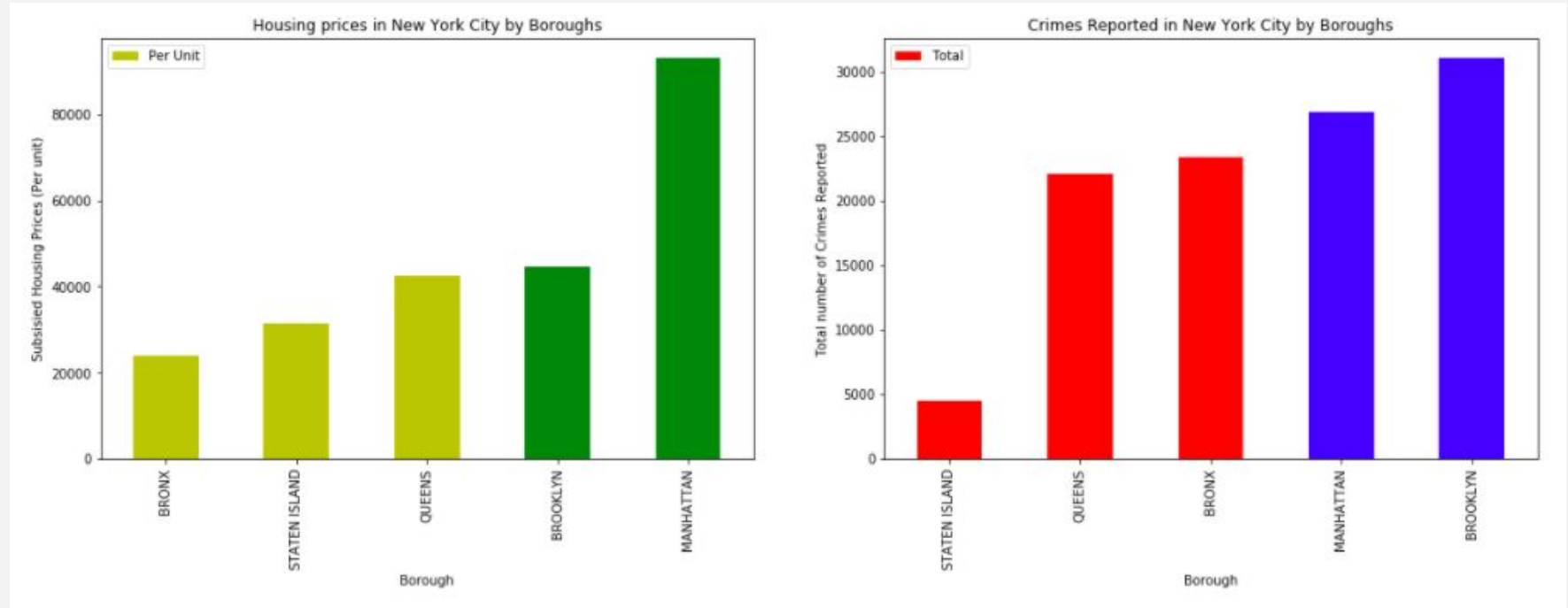## RESULTS

Crime Numbers in all 5 Boroughs : Figure 4.1

Subsidized Housing Prices in all 5 Boroughs : Figure 4.2
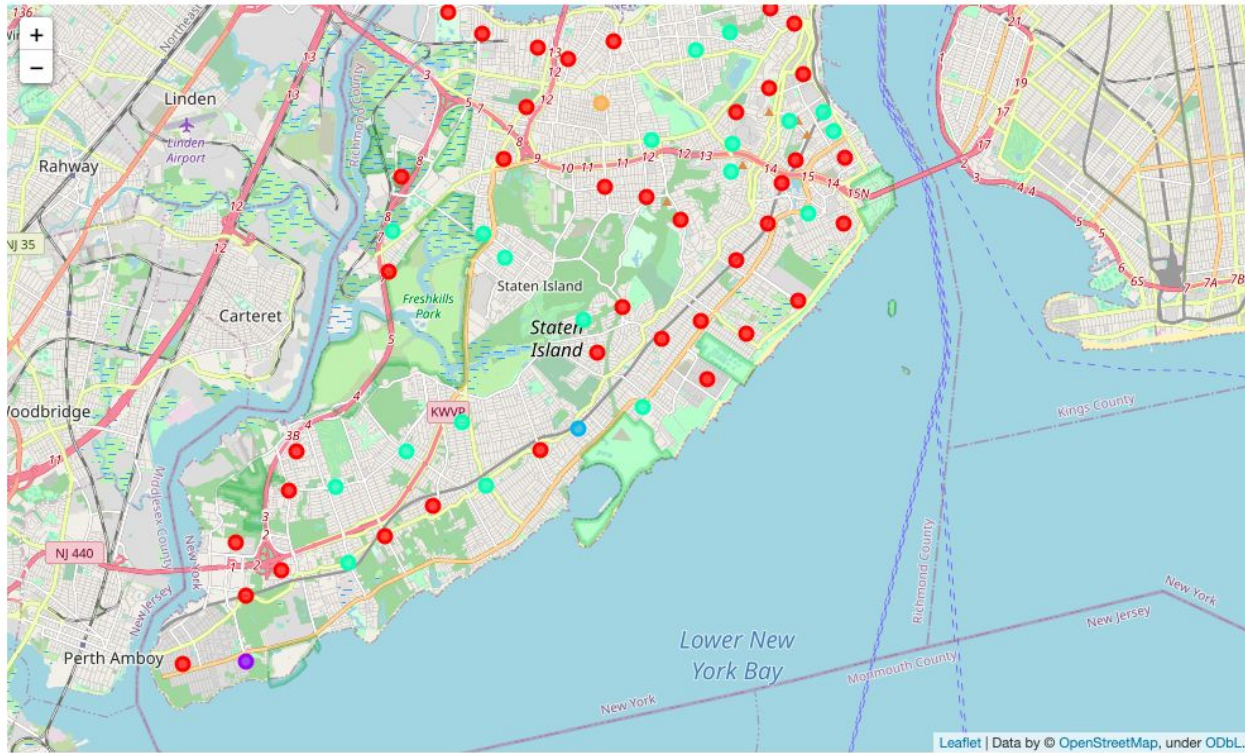
Comparison of the Crime Numbers and Housing Prices : Figure 4.3
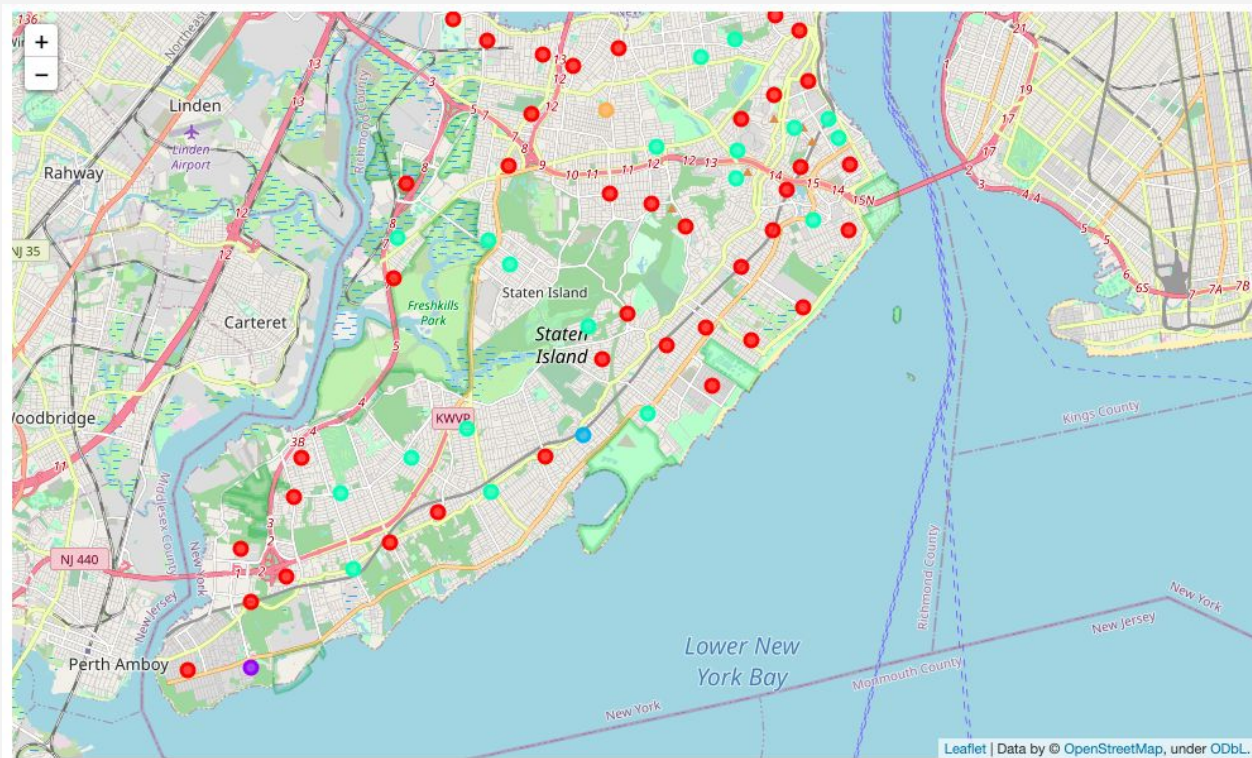
Staten Island Clusters : Figure 4.4

Cluster 0

	Total Venues
Pizza Place	35
Deli / Bodega	32
Bus Stop	31
Italian Restaurant	31
Chinese Restaurant	13
Bar	12
Sandwich Place	12
Bagel Shop	12
Cosmetics Shop	11
Donut Shop	11

Cluster 3

	Total Venues
Pizza Place	19
Bus Stop	12
Grocery Store	11
Pharmacy	10
Bagel Shop	10
Coffee Shop	10
Italian Restaurant	9
Sandwich Place	8
Bank	8
Bar	7

Bronx Clusters : Figure 4.5

	Total Venues
Pizza Place	35
Deli / Bodega	32
Bus Stop	31
Italian Restaurant	31
Chinese Restaurant	13
Bar	12
Sandwich Place	12
Bagel Shop	12
Cosmetics Shop	11
Donut Shop	11

	Total Venues
Pizza Place	19
Bus Stop	12
Grocery Store	11
Pharmacy	10
Bagel Shop	10
Coffee Shop	10
Italian Restaurant	9
Sandwich Place	8
Bank	8
Bar	7

05
DISCUSSION

Cut down Boroughs by looking at the Crime and Housing Prices results:

We cut down the Boroughs by looking at the Boroughs that returned the least values for the its crime numbers (Figure 4.1) and Housing Prices (Figure 4.2).

Our target audience moving into the state would likely prefer to live in a neighborhood that had the lowest crime rates and housing prices.

As seen in the Bar Graph Sub-Plot (Figure 4.3) theres common pattern and the 3 lowest ranking Boroughs are similar: **Staten Island, Queens and Bronx.**

This clearly affirms that these are the top 3 best boroughs in terms of safety and housing prices are the ones stated above.

To further deep dive into our analysis, we should consider the different priorities of different groups of migrants.

Some may place more importance on lower prices, and others may be greater safety. Thus, we will divide our recommendations for 2 groups of people:

I. **Staten Island:** Safety Conscious
II. **Bronx:** Price Conscious
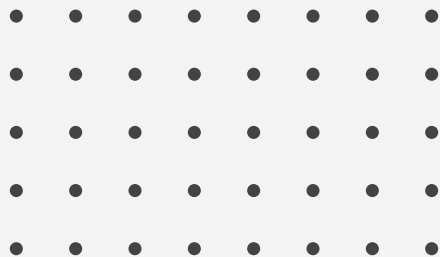
Lastly, we need to look at is Venues within the cluster of neighborhoods

We don't know the preference of individual people or families migrating, whether they wish to have more parks in the vicinity or food places.
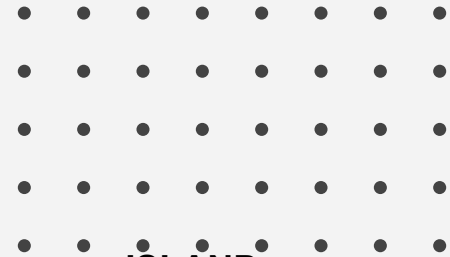
With this, we will judge the neighborhoods in terms of <u>the most venues within the clusters</u>, and <u>next the variety of venue categories</u> in those chosen clusters.

This will ensure that the migrants will have a relatively good lifestyle around the neighborhood and will cater to a variety of their needs.

FOR STATEN ISLAND:

In Figure 4.4 , we can observe that Cluster 0 is mainly a Food District, while Cluster 3, although has a lot of food places, also has a greater variety of venues such as Grocery stores, Pharmacy and Banks.

This will be greatly convenient when moving into a new area. Thus, the best neighborhoods in Staten Island are from Cluster 3, which are:

1. **West Brighton**
2. **Etingville**
3. **New Springville**

FOR   THE   BRONX:

In Figure 4.5 , it is clear that the variety of Venues in both clusters is <u>extremely similar</u> and will both provide a relatively good lifestyle for the migrants.

Thus, we judged the best neighborhood using <u>the most venue categories</u> within the vicinity, which is Cluster 2.

The top 3 neighborhoods are:

1. **Belmont**
2. **Fordham**
3. **Kingsbridge**

06
CONCLUSION

The neighbourhoods **West Brighton, Etingville** and **New Springville** of Staten Island, as well as **Belmont, Fordman** and **Kingsbridge** of the Bronx are the best neighborhoods in New York City.

These boroughs cater to a variety of people and their concerns, whether they are money conscious or wary of their safety. The neighborhoods further caters to people's interest and preferences, in terms venue categories. This will give them greater flexibility in the new area they wish to live in.

THANKS!

Final Report: <u>Click Here</u>
Notebook Coded: <u>Click Here</u>