THE BATTLE OF NEIGHBORHOODS

Applied Data Science – Capstone Project

Stefano I.

INTRODUCTION

Background

NYC is the most populous city in United States, often described as culture, financial, technology, education, tourism capital of Westrern world. Its liguistically most diverse city in the world with millions of foreign--born population of any city in the world. New york City is a place for migrating people or individuals who arrives to this place with dream of starting there lives away from there comforting families. For any individual or a family to move to a new place will be a hard decision, and moving to a metropolitan like New York City will be even more intimidating. This projects vision is to make it even more easier for individuals to take decision of which borough to choose to live in the city.

Business Problem

The objective of this Capstone Project is to analyse and choose the safest borough in the New York City based on the total crimes. This will help the people to buy/rent a home who are newly arriving to NYC. Exploring the neighbourhood and select the best among the five boroughs Brooklyn, Queens, Manhattan, The Bronx and Staten Island it has.

Who will use it

The target audience for this problem will be all the individual or the families moving to this new place to make a decision of which location is safe and will be suitable for there preferences.

DATA

NYPD Crime Data

The data for this project has been acquired from three different sources. First one is for the crime data of NYC I have used the link below.

https://data.cityofnewyork.us/Public-Safety/NYPD-Complaint-Data-Historic/qgea-i56i

The data is downloaded and the uploaded on Jupyter with the name NYPD_Crime_Data.csv. The data frame information is attained and it looks like the below Fig(a).

NYC Boroughs

Data related to boroughs and its population has been scrapped from the link below:

https://www.citypopulation.de/en/usa/newyorkcity/

Neighborhood

Data related to boroughs and its population.

https://geo.nyu.edu/catalog/nyu_2451_34572

METHODOLOGY

▸ Exploratory Data Analysis

1. Statistical Summary

```
NYC_Crime_Table.describe()
```

	Felony	Misdemeanor	Violation	Population-2019	Total
count	5.000000	5.000000	5.000000	5.000000e+00	5.000000
mean	30463.600000	50995.000000	14933.600000	1.667363e+06	96392.200000
std	15643.156037	23927.115883	6375.194334	8.098120e+05	45560.768087
min	5059.000000	10727.000000	4217.000000	4.761430e+05	20003.000000
25%	30356.000000	49857.000000	15862.000000	1.418207e+06	97201.000000
50%	31369.000000	57102.000000	15975.000000	1.628706e+06	104825.000000
75%	38903.000000	66785.000000	17367.000000	2.253858e+06	121550.000000
max	46631.000000	70504.000000	21247.000000	2.559903e+06	138382.000000

2. Merging the Data

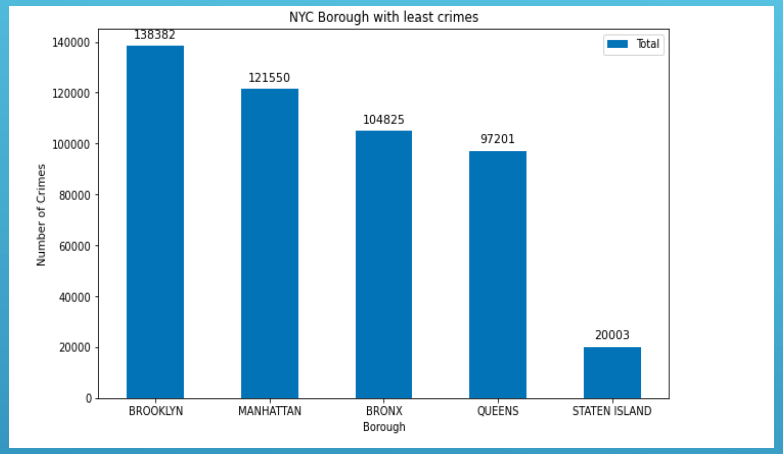
	Borough	Felony	Misdemeanor	Violation	Status	Population-2019	Total
1	BROOKLYN	46631	70504	21247	Borough	2559903	138382
2	MANHATTAN	38903	66785	15862	Borough	1628706	121550
0	BRONX	30356	57102	17367	Borough	1418207	104825
3	QUEENS	31369	49857	15975	Borough	2253858	97201
4	STATEN ISLAND	5059	10727	4217	Borough	476143	20003

3. Finding Highest Crime Rate

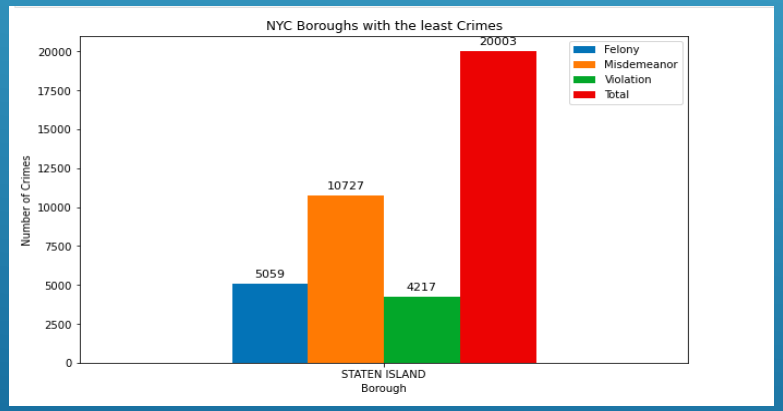
```
NYC_Crime_Table.sort_values(['Total'], ascending = False, axis = 0, inplace = True )
NYC_Crime_Table
```

	Borough	Felony	Misdemeanor	Violation	Status	Population-2019	Total
1	BROOKLYN	46631	70504	21247	Borough	2559903	138382
2	MANHATTAN	38903	66785	15862	Borough	1628706	121550
0	BRONX	30356	57102	17367	Borough	1418207	104825
3	QUEENS	31369	49857	15975	Borough	2253858	97201
4	STATEN ISLAND	5059	10727	4217	Borough	476143	20003

4. Analyzing Safety within Boroughs



5. Analyzing Crimes



SEGMENTATION & CLUSTERING (PART I)

Borough

The data included 5 boroughs and 306 neighborhoods.

```
[10]: neighborhoods.head()
```

[10]:

	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

```
[11]: print('The dataframe has {} boroughs and {} neighborhoods.'.format(
          len(neighborhoods['Borough'].unique()),
          neighborhoods.shape[0]
      )
)
```

```
The dataframe has 5 boroughs and 306 neighborhoods.
```

Foursquare API

I set the radius at 500 and limited venues to 100.

SEGMENTATION & CLUSTERING (PART II)

One Hot Encoding

Analyzing each Neighborhood of Staten Island

	Neighborhood	Accessories Store	African Restaurant	American Restaurant	Arcade	Art Gallery	Art Museum	Arts & Crafts Store	Asian Restaurant	Athletics & Sports	...	Tourist Information Center	Toy / Game Store	Trail
0	Annadale	0.0	0.0	0.100000	0.000000	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0
1	Arden Heights	0.0	0.0	0.000000	0.000000	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0
2	Arlington	0.0	0.0	0.000000	0.000000	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0
3	Arrochar	0.0	0.0	0.000000	0.000000	0.0	0.0	0.0	0.0	0.043478	...	0.0	0.0	0.0
4	Bay Terrace	0.0	0.0	0.000000	0.000000	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
58	Travis	0.0	0.0	0.000000	0.000000	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0
59	West Brighton	0.0	0.0	0.026316	0.000000	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0
60	Westerleigh	0.0	0.0	0.000000	0.333333	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0
61	Willowbrook	0.0	0.0	0.000000	0.000000	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0
62	Woodrow	0.0	0.0	0.000000	0.000000	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0

63 rows × 181 columns

Clusters of Neighborhoods

K-means clustering using supervised machine learning.

	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	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 Com V
0	Staten Island	St. George	40.644982	-74.079353	2	Clothing Store	Sporting Goods Shop	Italian Restaurant	Park	Bar	Pizza Place	Outlet Mall	Bas Sta
1	Staten Island	New Brighton	40.640615	-74.087017	0	Deli / Bodega	Bus Stop	Park	Discount Store	Playground	Flower Shop	Bowling Alley	Fi Resta
2	Staten Island	Stapleton	40.626928	-74.077902	2	Pizza Place	Discount Store	Sandwich Place	Bank	Restaurant	Spanish Restaurant	Fast Food Restaurant	Skate
3	Staten Island	Rosebank	40.615305	-74.069805	2	Pharmacy	Italian Restaurant	Grocery Store	Breakfast Spot	Beach	Pizza Place	Deli / Bodega	Cosm
4	Staten Island	West Brighton	40.631879	-74.107182	2	Coffee Shop	Pharmacy	Bank	Italian Restaurant	Music Store	Breakfast Spot	Bar	Bus

RESULTS

- **The result of the study deepend the audience knowledge about the neighborhood situation in NYC, helping them to eventually relocate to the safer borough of the city.**

- **The venues overview help the reader evaluatethe most suitable choice.**

DISCUSSION

Cluster 2

- In the clusters formed after the data is explored a individual can look for a neighborhood with good public transportation, food places we can see that Clusters 2 has bus stops, restaurants as the most common venues.

	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
26	Graniteville	Boat or Ferry	Grocery Store	Filipino Restaurant	Gas Station	Furniture / Home Store	French Restaurant	Food Truck	Food & Drink Shop	Food	Flower Shop
53	Howland Hook	Boat or Ferry	German Restaurant	Gas Station	Furniture / Home Store	French Restaurant	Food Truck	Food & Drink Shop	Food	Flower Shop	Fish & Chips Shop

Cluster 3

- If a person is looking for a neighborhood with stores and restaurants in a close proximity then the neighborhoods in the cluster 3 is suitable for them. For a family looking for a neighborhood the Cluster 3 is more suitable as it shows parks, beach, grocery stores and Gyms.

	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 M Comn Ve
0	St. George	Clothing Store	Sporting Goods Shop	Italian Restaurant	Park	Bar	Pizza Place	Outlet Mall	Baseball Stadium	Bus Stop	Bus Sta
2	Stapleton	Pizza Place	Discount Store	Sandwich Place	Bank	Restaurant	Spanish Restaurant	Fast Food Restaurant	Skate Park	New American Restaurant	Optical S
3	Rosebank	Pharmacy	Italian Restaurant	Grocery Store	Breakfast Spot	Beach	Pizza Place	Deli / Bodega	Cosmetics Shop	Ice Cream Shop	Eas Europ Restau
4	West Brighton	Coffee Shop	Pharmacy	Bank	Italian Restaurant	Music Store	Breakfast Spot	Bar	Bus Stop	Café	Sandw Pl
7	South Beach	Pier	Beach	Athletics & Sports	Deli / Bodega	Diner	Discount Store	Furniture / Home Store	French Restaurant	Food Truck	Food & D St
8	Port Richmond	Rental Car Location	Martial Arts Dojo	Donut Shop	Pizza Place	Dim Sum Restaurant	Fast Food Restaurant	French Restaurant	Food Truck	Food & Drink Shop	F
11	Castleton Corners	Bank	Pizza Place	Japanese Restaurant	Sandwich Place	Go Kart Track	Grocery Store	Mini Golf	Tattoo Parlor	Bagel Shop	
12	New Springville	Chinese Restaurant	Coffee Shop	Mobile Phone Shop	Bagel Shop	Accessories Store	Donut Shop	Sandwich Place	Restaurant	Pizza Place	Pharm

CONCLUSION

- This project gives better understanding to the neighborhoods in terms of the crimes occurring in the borough.
- The most common venues in that neighborhood.
- Safety is considered as a primary concern of everyone and has shortlisted to the safest borough in New York city.
- The most common places in the neighborhoods to present the different clusters to choose according to once preferences.