ANALYSIS OF BUSINESS SITES IN BROOKLYN NEIGHBORHOODS

INTRODUCCION

BACKGROUND AND BUSINESS ANALYSIS

IN THIS CASE STUDY
A CLEARER VISION WILL BE GIVEN TO ENTER A NEW BUSINESS IN THE BROOKLYN DISTRICT.

TO ACHIEVE THE OBJECTIVE WE WILL DO AN ANALYSIS
WHERE ALL THE COMMERCIAL PREMISES OF ALL THE NEIGHBORHOODS OF THE BROOKLYN DISTRICT WILL BE EXPLORED

WE WILL EXPLORE THEIR GEOGRAPHIC LOCATIONS, THEIR TYPES OF BUSINESSES, THE NAMES OF THE BUSINESSES AND THE NUMBER OF COMMERCIAL PLACES BY CATEGORY.

WE WILL ALSO ANALYZE THE MOST COMMON COMMERCIAL PREMISES, AND THE LESS COMMON ONES.

WITH THIS ANALYSIS
INTERESTED PERSONS OR COMPANIES MAY HAVE A CLEARER VISION TO ENTER A NEW BUSINESS IN THE BROOKLYN DISTRICT.

DATA ACQUISITION AND CLEANING

DATA SOURCE

TO ACHIEVE THE OBJECTIVE OF ANALYZING ALL THE COMMERCIAL PREMISES OF ALL THE NEIGHBORHOODS OF THE

BROOKLYN DISTRICT, WE BEGAN BY DOWNLOADING THE DATA OF THE CITY OF NEW YORK WITH ITS GEOGRAPHICAL

LOCATIONS FROM THE FOLLOWING LINK: DATOS NEW YORK

WITH THE HELP OF THE FOURSQUARE API, WE EXPLORE ALL THE BROOKLYN NEIGHBORHOODS WITH THEIR

COMMERCIAL PLACES AND CATEGORIES

DATA CLEANING

HAVING OBTAINED THE DATA FROM THE INDICATED SOURCES, WE MOVE ON TO THE NEXT PROCESS OF TRANSFORMING THE DATA INTO A PANDAS DATA FRAME TO SEGMENT AND GROUP THE DATA FOR THE NEIGHBORHOODS IN THE BROOKLYN BOROUGH.

AFTER CREATING THE NEW PANAS DATA FRAME WITH THE BROOKLYN DATA, AND OBTAINING THE BROOKLYN LATITUDE AND LONGITUDE VALUES, THEN THE BROOKLYN MAP IS CREATED.

WITH THE HELP OF THE FOURSQUARE API, WE SCANNED ALL THE BROOKLYN NEIGHBORHOODS WITH THEIR BUSINESS LOCATIONS AND CATEGORIES TO GENERATE THE DATA FRAME, AN IMAGE OF WHICH CAN BE SEEN IN THE FOLLOWING FIGURE 1.

IN THAT DATA FRAME, WE HAVE ALL THE NEIGHBORHOODS IN BROOKLYN

WITH THE 10 MOST COMMON BUSINESS LOCATIONS FOR EACH NEIGHBORHOOD.

AT THE END OF THE CASE STUDY, WE WILL APPLY MACHINE LEARNING TECHNIQUES TO ANALYZE ALL THE INFORMATION BY CLUSTERS, AND

TO GET RESPONSES TO THE FOLLOWING REQUESTS:

MOST COMMON BUSINESS TYPES

LESS COMMON BUSINESS TYPES

MOST COMMON BUSINESS NAMES

LESS COMMON BUSINESS NAMES

HOW MANY MORE COMMON BUSINESS PREMISES EXIST

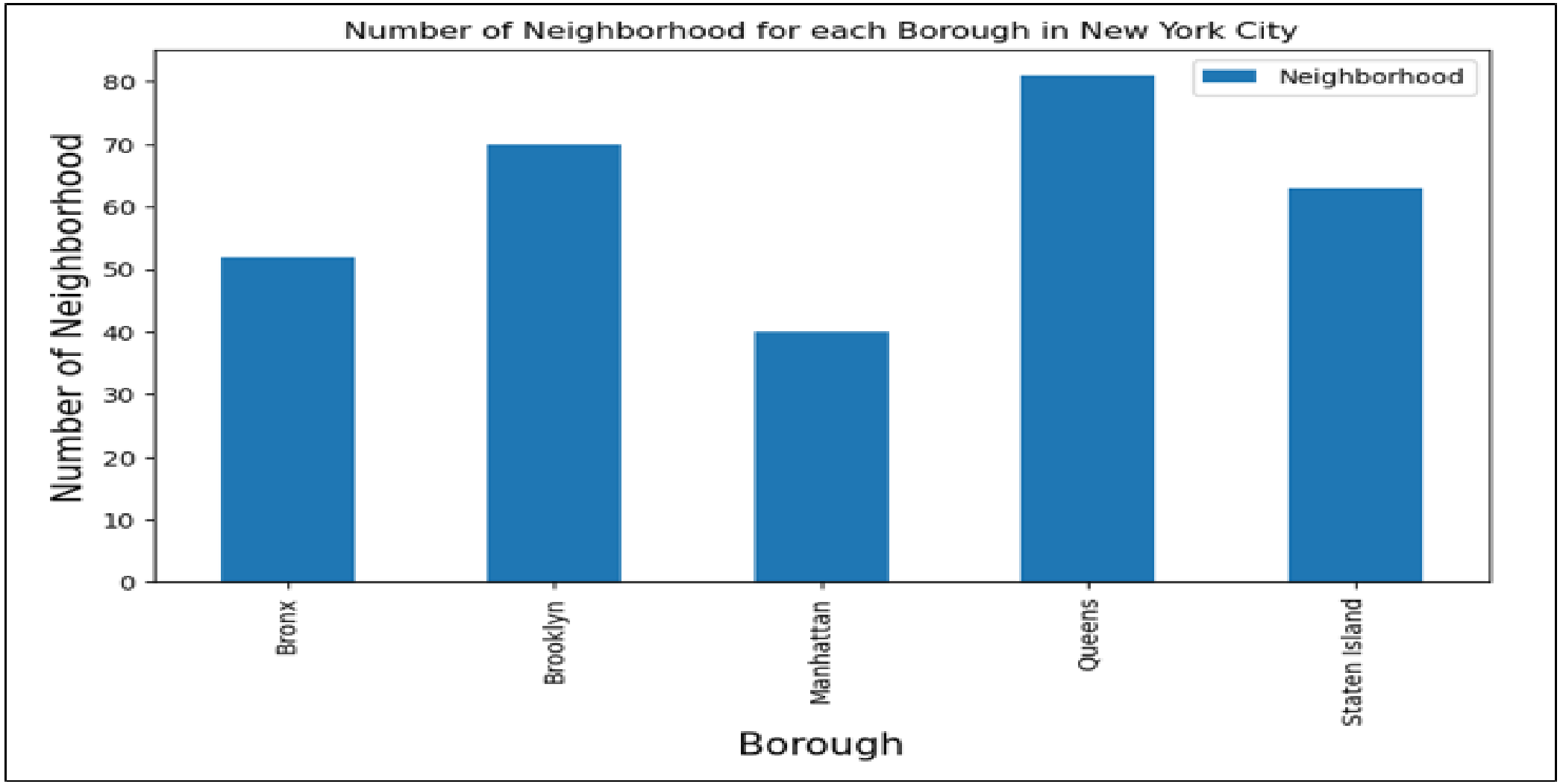
HOW MANY LESS COMMON BUSINESS PREMISES EXIST

NEIGHBORHOODS WITH THE 10 MOST COMMERCIAL PLACES

	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	Bath Beach	Bubble Tea Shop	Pharmacy	Italian Restaurant	Gas Station	Chinese Restaurant	Donut Shop	Dessert Shop	Fast Food Restaurant	Pizza Place	Cantonese Restaurant
1	Bay Ridge	Italian Restaurant	Spa	Pizza Place	Bar	American Restaurant	Greek Restaurant	Hookah Bar	Thai Restaurant	Bagel Shop	Pharmacy
2	Bedford Stuyvesant	Coffee Shop	Café	Pizza Place	Bar	Discount Store	Tiki Bar	Fried Chicken Joint	New American Restaurant	Boutique	Gift Shop
3	Bensonhurst	Chinese Restaurant	Italian Restaurant	Donut Shop	Sushi Restaurant	Park	Pizza Place	Ice Cream Shop	Smoke Shop	Bar	Noodle House
4	Bergen Beach	Harbor / Marina	Baseball Field	Hockey Field	Athletics & Sports	Playground	Park	Food	Flower Shop	Food & Drink Shop	Food Court
5	Boerum Hill	Coffee Shop	Dance Studio	Bar	Furniture / Home Store	Spa	Sandwich Place	Bakery	Arts & Crafts Store	French Restaurant	Yoga Studio
6	Borough Park	Bank	Pizza Place	Fast Food Restaurant	Café	Pharmacy	Eastern European Restaurant	Hotel	American Restaurant	Grocery Store	Deli / Bodega
7	Brighton Beach	Eastern European Restaurant	Restaurant	Russian Restaurant	Beach	Sushi Restaurant	Mobile Phone Shop	Pharmacy	Gourmet Shop	Playground	Korean Restaurant

Exploratory Data Analysis

We pass the data of the neighborhoods of each district to a dataframe, and make a representative graph.



Predictive Modeling

APPLY MACHINE LEARNING - Cluster Neighborhoods

Run K-MEANS to group the 70 neighborhood into 10 groups.

```
# set number of clusters
kclusters = 10
Brooklyn_grouped_clustering = Brooklyn_grouped.drop('Neighborhood', 1)
Brooklyn_grouped_clustering
```

```
# run k-means clustering
kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(Brooklyn_grouped_clustering)

# check cluster labels generated for each row in the dataframe
kmeans.labels_[0:10]
```

```
array([7, 7, 7, 7, 6, 7, 1, 7, 7, 7], dtype=int32)
```

We examine the Clusters

Now, we can examine each cluster and determine the categories of discriminating commercial places that distinguish each cluster.

Cluster 1

```
Clust_Brooklyn2.loc[Clust_Brooklyn2['Cluster Labels'] == 0, Clust_Brooklyn2.columns[[1] + list(range(5, Cl
ust_Brooklyn2.shape[1]))]]
```

	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
26	East New York	Deli / Bodega	Salon / Barbershop	Fried Chicken Joint	Child Care Service	Plaza	Fast Food Restaurant	Pizza Place	Event Service	Convenien Store
43	Ocean Hill	Deli / Bodega	Bus Stop	Supermarket	Southern / Soul Food Restaurant	Grocery Store	Park	Martial Arts School	Seafood Restaurant	Chinese Restaurant
67	Highland Park	Deli / Bodega	Moving Target	Grocery Store	Tennis Court	Park	Liquor Store	Spanish Restaurant	Caribbean Restaurant	Pizza Place

Cluster 9

```
Clust_Brooklyn2.loc[Clust_Brooklyn2['Cluster Labels'] == 8, Clust_Brooklyn2.columns[[1] + list(range(5, Cl
ust_Brooklyn2.shape[1]))]]
```

	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 Comm Venue
39	Sea Gate	Beach	Bus Station	Spa	American Restaurant	Lighthouse	Sports Club	Home Service	Dog Run	Bus Line	Flea Market

Cluster 10

```
Clust_Brooklyn2.loc[Clust_Brooklyn2['Cluster Labels'] == 9, Clust_Brooklyn2.columns[[1] + list(range(5, Cl
ust_Brooklyn2.shape[1]))]]
```

	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 Mo Commo Venue
68	Madison	Bagel Shop	Pilates Studio	Deli / Bodega	Dessert Shop	Candy Store	Italian Restaurant	Spa	Hobby Shop	Pizza Place	Restaura

CONCLUSIONS

Upon completion of the exploratory study and cluster analysis of all commercial locations in all neighborhoods in the brooklyn district, we can conclude the following:

In cluster 1 the most common commercial places are warehouses

In group 2, they are pizzerias

In group 3, they are car garage

In group 4, are the Caribbean Restaurants

In cluster 5, they are the pools

In group 6, they are the Grocery Store and Burger

In group 7, they are the Playground, Harbor / Marina, Athletics and Sports, Baseball Field and Hockey Field

In group 8, are the restaurants, Italian restaurants, Chinese, Latin, Russian and American restaurants

In cluster 9, they are the bus stations and Spa

In group 10 are the Bagel Shop