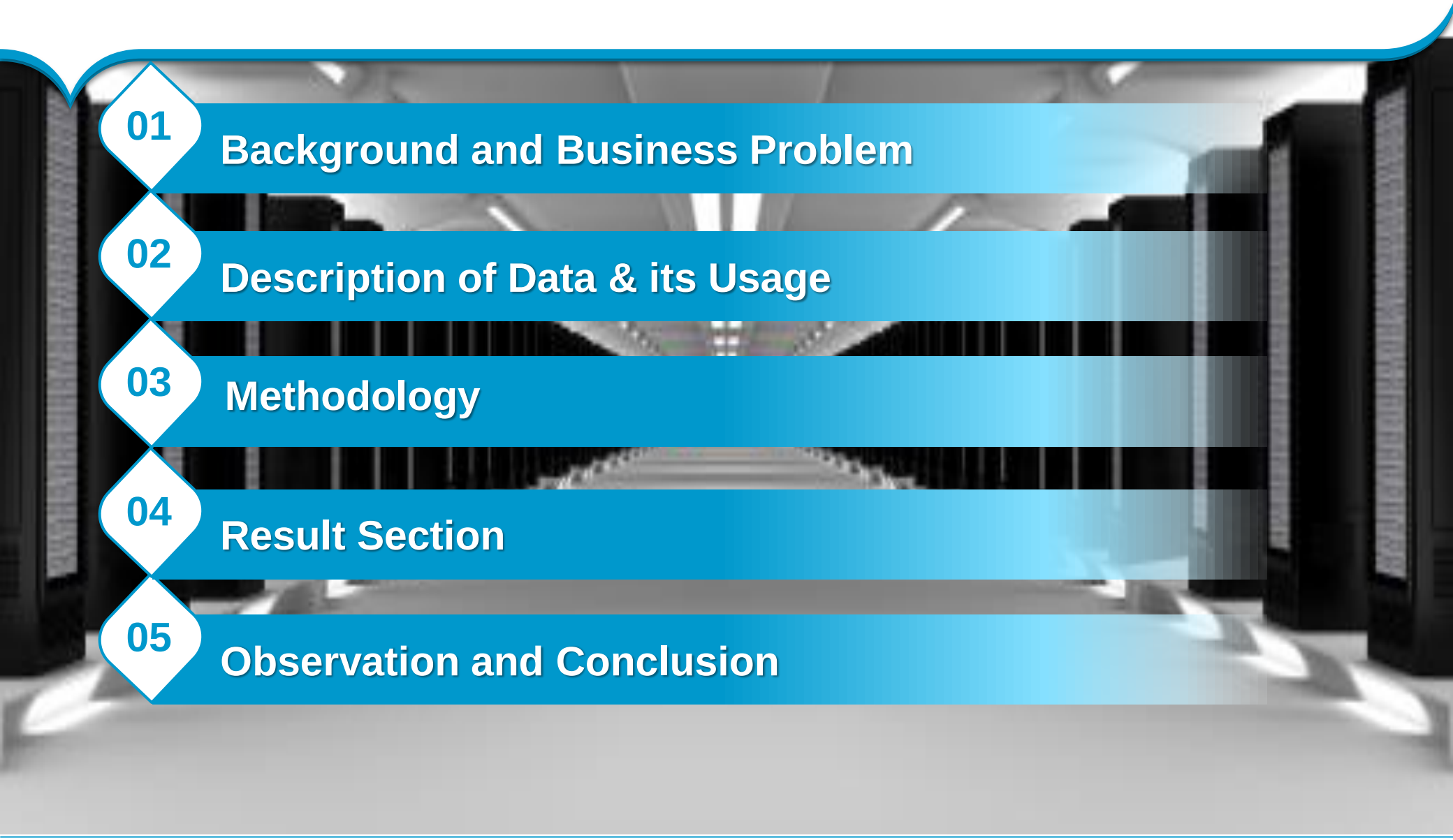
Data Science Project - Pharmacy @ Brooklyn

- Bhushan Gadhari

Contents



Background and Business Problem

Background:

This is a Data Science Project – "The Battle of Neighborhood – Most Pharmacies around Brooklyn"

Business Problem and Targeted Audience:

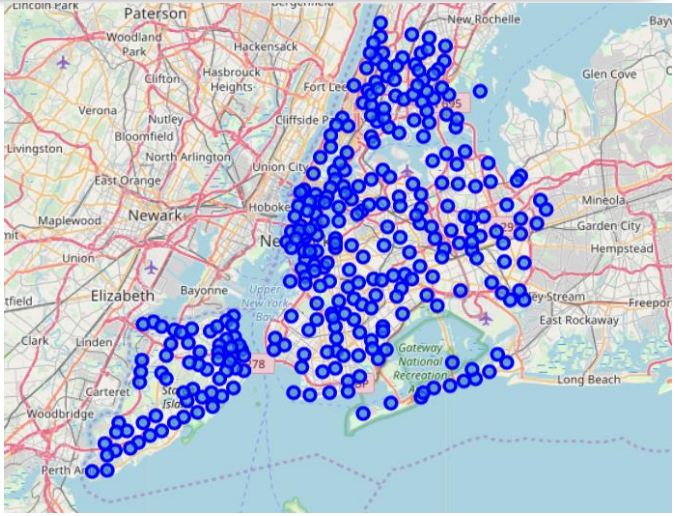
We have a client – Team of Doctor's and Medical Representatives who have been given an assignment to go to New York, Brooklyn and have been asked to set up private Hospitals around the same. They need to find an area that will yield maximum output and they can treat the patients efficiently.

Now this effectively means that we need to find out locations that have most Pharmacies and these can be ideal place to set up Hospitals

This project will be useful for the targeted Audience as we will to find out the locations around Brooklyn where we have maximum "Pharmacies" or they have "Pharmacy" as most common venue

Description of Data and its Usage

- We will use New York Data set that is readily available on the web.
- In this Data set we have 5 boroughs and 306 neighborhoods
- Foursquare Method is used to get details
- Features key of the Data set has all the relevant information
- This data set is then transformed to Pandas Data frame
- First, we plot New York map with neighborhoods superimposed on Top (Visualize using Folium Maps)



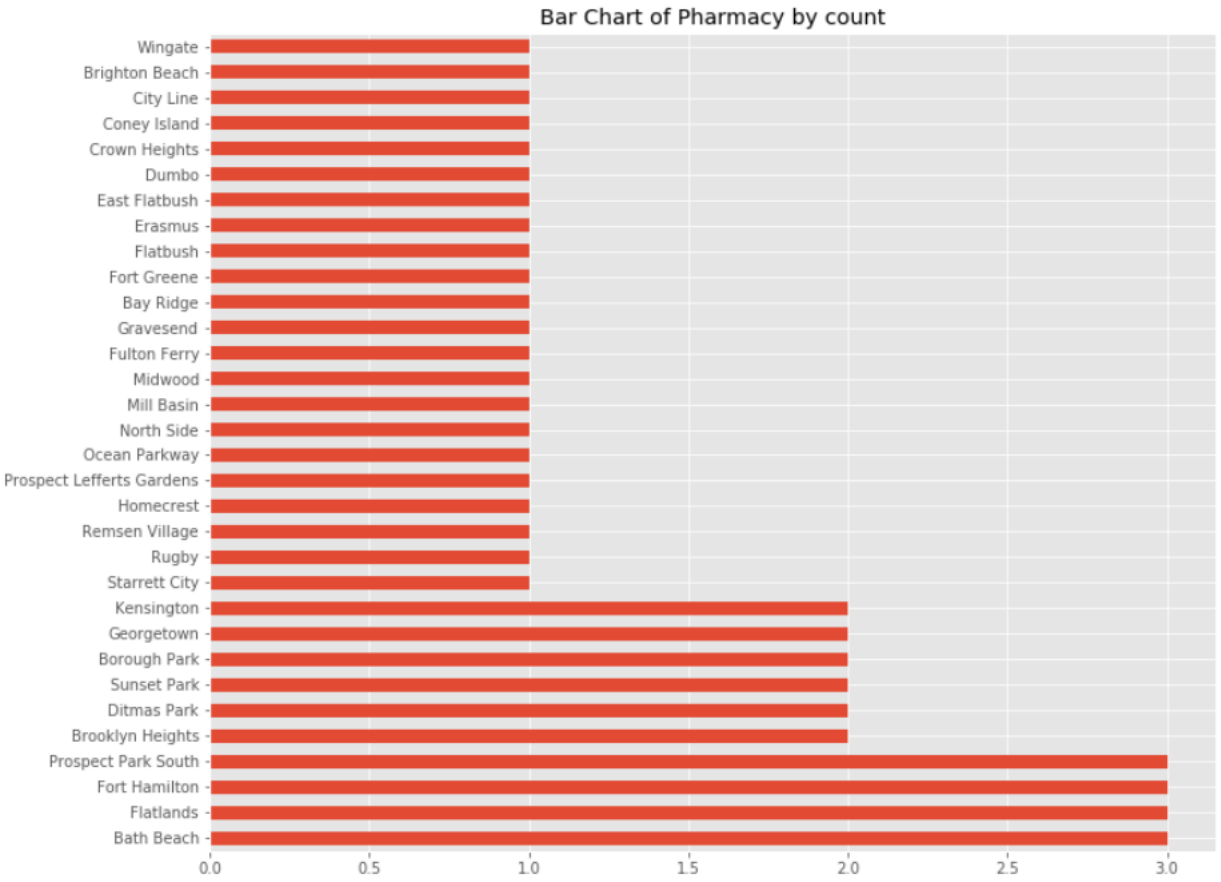
Our Interest area here is borough “Brooklyn”

- Then we plot a map of Brooklyn with its neighborhood superimposed on Top
- Then we create a new Data frame with Brooklyn data
- We then explore Brooklyn’s first neighborhood called “Bay Ridge” and its top 100 venues
- We get data from Items key using Foursquare Location data
- We repeat this to get data for all Brooklyn Neighborhoods and assess Venues separately
- We then refine Brooklyn neighborhood data to show the locations that Pharmacy as most common location as well as locations that have maximum number of Pharmacies

Methodology

- Geopy packages to plot Brooklyn maps. Folium maps created.
- We use Json file and its fields to extract relevant data and convert that to dataframe
- Foursquare credentials used to get venues and other insights
- One hot encoding used to convert the categorial data into integer values
- Bar charts used to assess the locations with highest number of Pharmacies (Matplotlib package used for plotting)
- Finally we used the Most Common Function to assess the areas that have Pharmacy as its top common locations and another dataset to get maximum number of Pharmacies among venues

Result Section (1)



- Bar Chart showing Pharmacy count by location
- Since this study is by count, it may not necessarily be the most common location
- Bath Beach, Flatlands, Fort Hamilton and Prospect Part South location in the premises of Brooklyn have most Pharmacy Locations

Results Section (2)

Finally - Concluding the Neighborhoods that have Pharmacy as most common Venue

```
print ("\n""The Neighborhoods of Brooklyn that have Pharmacy as most common Venues among top 3 are:""\n")
print(df1.Neighborhood[0])
print(df1.Neighborhood[29])
print(df2.Neighborhood[31])
print(df2.Neighborhood[33])
print(df2.Neighborhood[56])
print(df3.Neighborhood[20])
```

```
The Neighborhoods of Brooklyn that have Pharmacy as most common Venues among top 3 are:

Bath Beach
Flatlands
Fort Hamilton
Georgetown
Prospect Park South
Ditmas Park
```

- Neighborhoods showing Pharmacy as common location among top 3
- Bath Beach, Flatlands, Fort Hamilton, Georgetown, Prospect Park South and Ditmas Park are the location with most common venue as Pharmacy

- Split of the common locations as 1st, 2nd and 3rd most common location
- Data frames showing each of these categories

Find out which locations have Pharmacy as the most common location

```
df1 = neighborhoods_venues_sorted[neighborhoods_venues_sorted['1st Most Common Venue'].str.contains("Pharmacy")]
df1
```

	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	Pharmacy	Chinese Restaurant	Sushi Restaurant	Bubble Tea Shop	Pizza Place	Donut Shop	Fast Food Restaurant	Kids Store	Italian Restaurant	Cantonese Restaurant
29	Flatlands	Pharmacy	Caribbean Restaurant	Fast Food Restaurant	Deli / Bodega	Fried Chicken Joint	Discount Store	Bar	Video Store	Nightclub	Electronics Store

Find out which locations have Pharmacy as the 2nd most common location

```
df2 = neighborhoods_venues_sorted[neighborhoods_venues_sorted['2nd Most Common Venue'].str.contains("Pharmacy")]
df2
```

	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
31	Fort Hamilton	Deli / Bodega	Pharmacy	Chinese Restaurant	Sandwich Place	Bank	Bagel Shop	Diner	Cocktail Bar	Pub	Japanese Restaurant
33	Georgetown	Bank	Pharmacy	Breakfast Spot	Donut Shop	Pizza Place	Shopping Mall	Burger Joint	Shipping Store	Sandwich Place	Clothing Store
56	Prospect Park South	Caribbean Restaurant	Pharmacy	Fast Food Restaurant	Grocery Store	Pizza Place	Donut Shop	Mobile Phone Shop	Latin American Restaurant	Fried Chicken Joint	Mexican Restaurant

Find out which locations have Pharmacy as the 3rd most common location

```
df3 = neighborhoods_venues_sorted[neighborhoods_venues_sorted['3rd Most Common Venue'].str.contains("Pharmacy")]
df3
```

	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
20	Ditmas Park	Deli / Bodega	Pizza Place	Pharmacy	Kids Store	Caribbean Restaurant	Chinese Restaurant	Donut Shop	Discount Store	Women's Store	Japanese Restaurant

Observation and Conclusion

- We carried out analysis with help of Data frame containing Brooklyn's Pharmacy location
- Each of the neighborhoods that had "Pharmacy" were shown, then we plotted Bar chart showing count of "Pharmacy" locations
- This helped to show high number of such location
- This concluded one area of analysis

- Another analysis we did was around Pharmacy as most common location
- Though this may not be count-wise, we still got an idea on such locations in terms of availability frequency
- Hence if you see we got different results from bot exercises

- **Conclusive results for Doctors and Medical Representatives to set up Hospitals – Below are the locations with most common venue as Pharmacy**

- **Bath Beach**
- **Flatlands**
- **Fort Hamilton**
- **Georgetown**
- **Prospect Park South**
- **Ditmas Park**