

Capstone Project

EDA Analysis on Airbnb Booking Data Set

Presentation By:

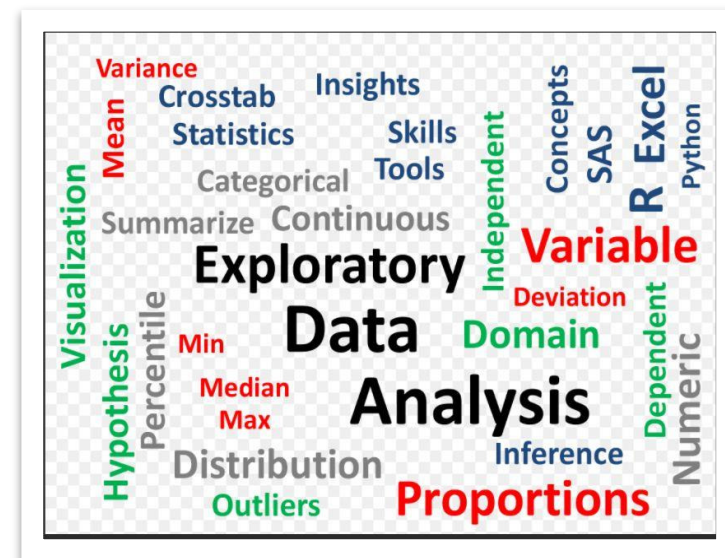
Anumalla Rajashekar

OBJECTIVE:

To perform Exploratory Data Analysis on a given dataset

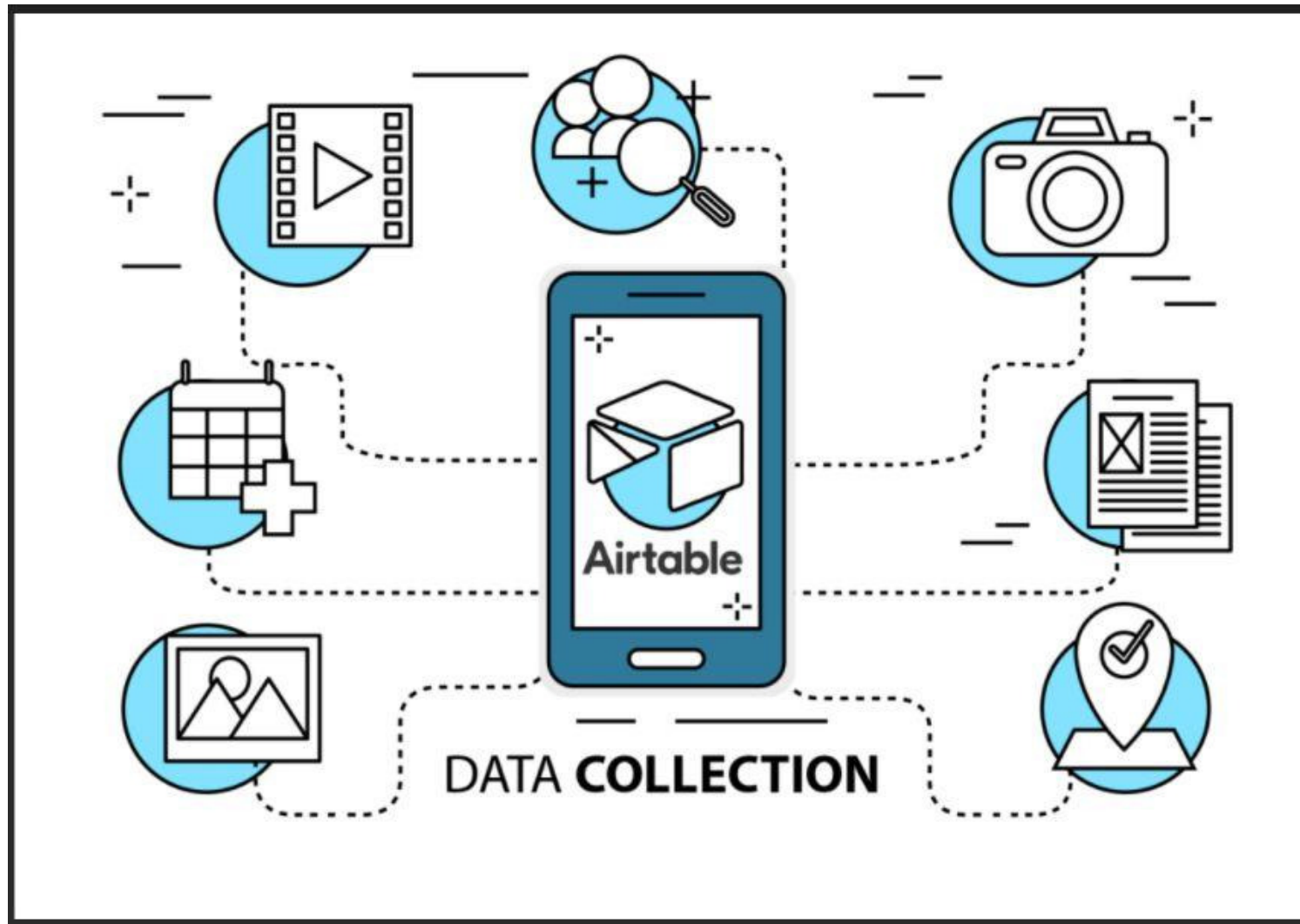
Exploratory Data Analysis (EDA)

EDA is the process of investigating the dataset to discover patterns, and anomalies (outliers), and form hypotheses based on our understanding of the dataset



EDA

- DATA COLLECTION
- DATA CLEANING
- HYPOTHESIS
- DATA VISUALIZATION



DATA CLEANING

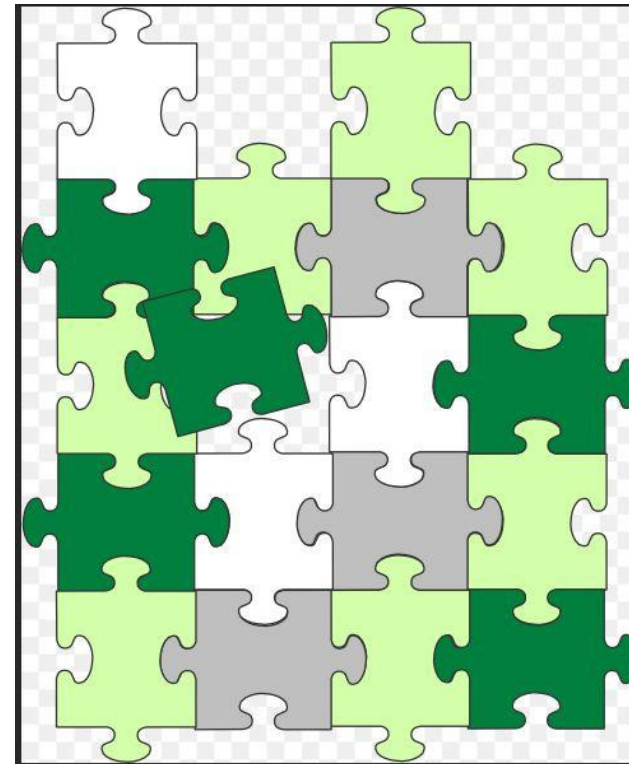


DEALING WITH MISSING DATA VALUES



Detection Of Missing Data

id	0
name	16
host_id	0
host_name	21
neighbourhood_group	0
neighbourhood	0
latitude	0
longitude	0
room_type	0
price	0
minimum_nights	0
number_of_reviews	0
last_review	10052
reviews_per_month	10052
calculated_host_listings_count	0
availability_365	0
dtype: int64	



Dropping of unnecessary features



Handling Of Missing Data



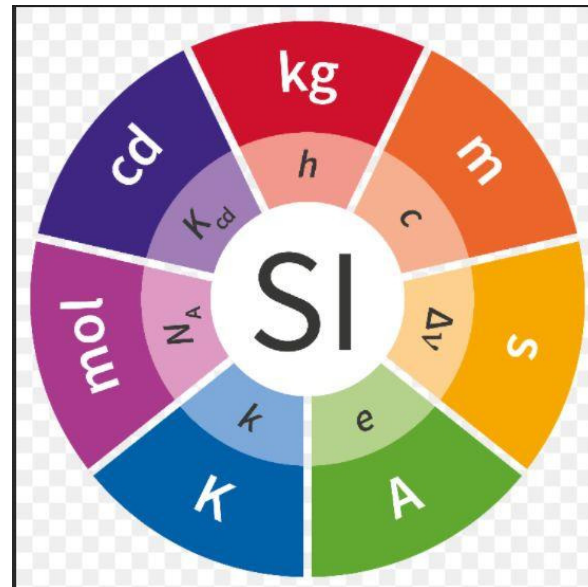
```
host_id                                0
neighbourhood_group                    0
neighbourhood                          0
latitude                              0
longitude                              0
room_type                              0
price                                  0
minimum_nights                         0
number_of_reviews                      0
reviews_per_month                      0
calculated_host_listings_count        0
availability_365                       0
dtype: int64
```

FINDING DUPLICATE RECORDS

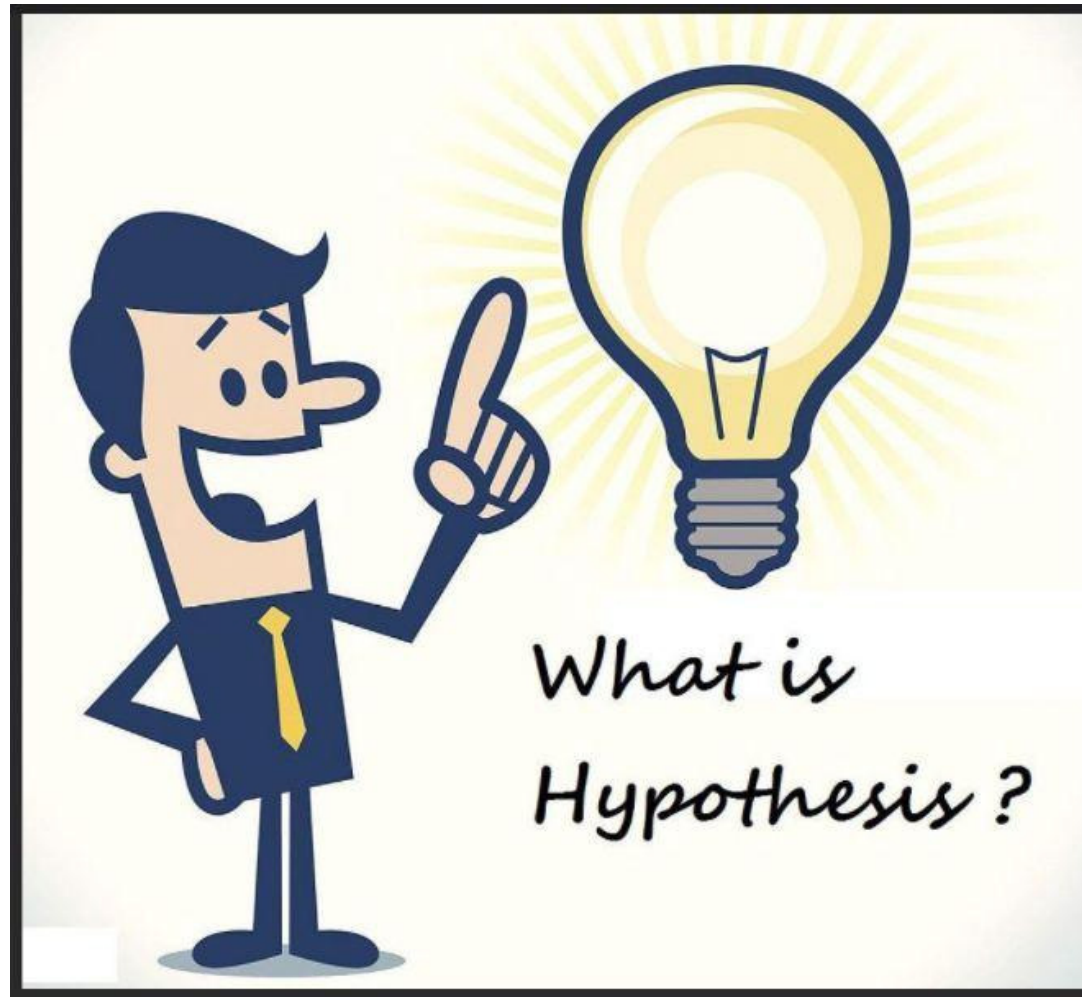


Data Formatting

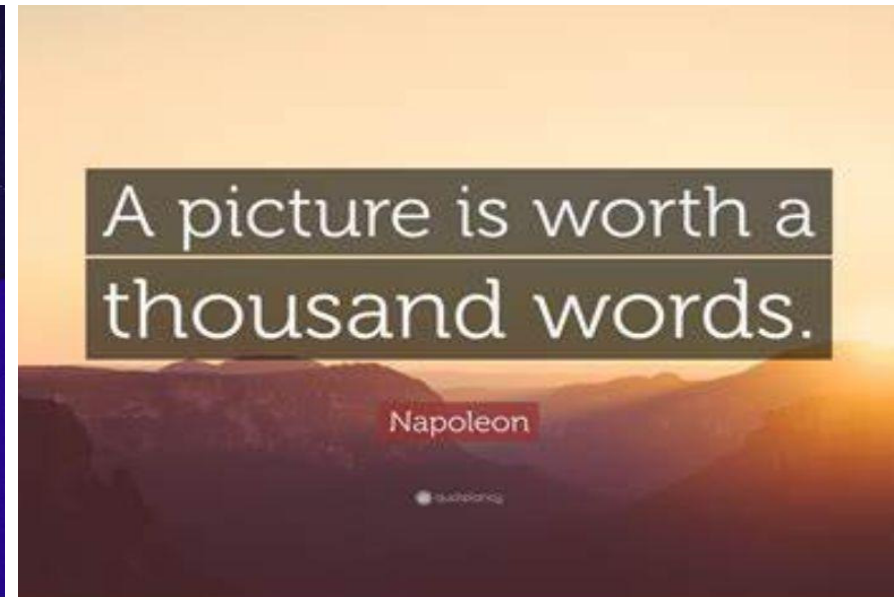
- Units
- Symbols
- AlphaNumeric



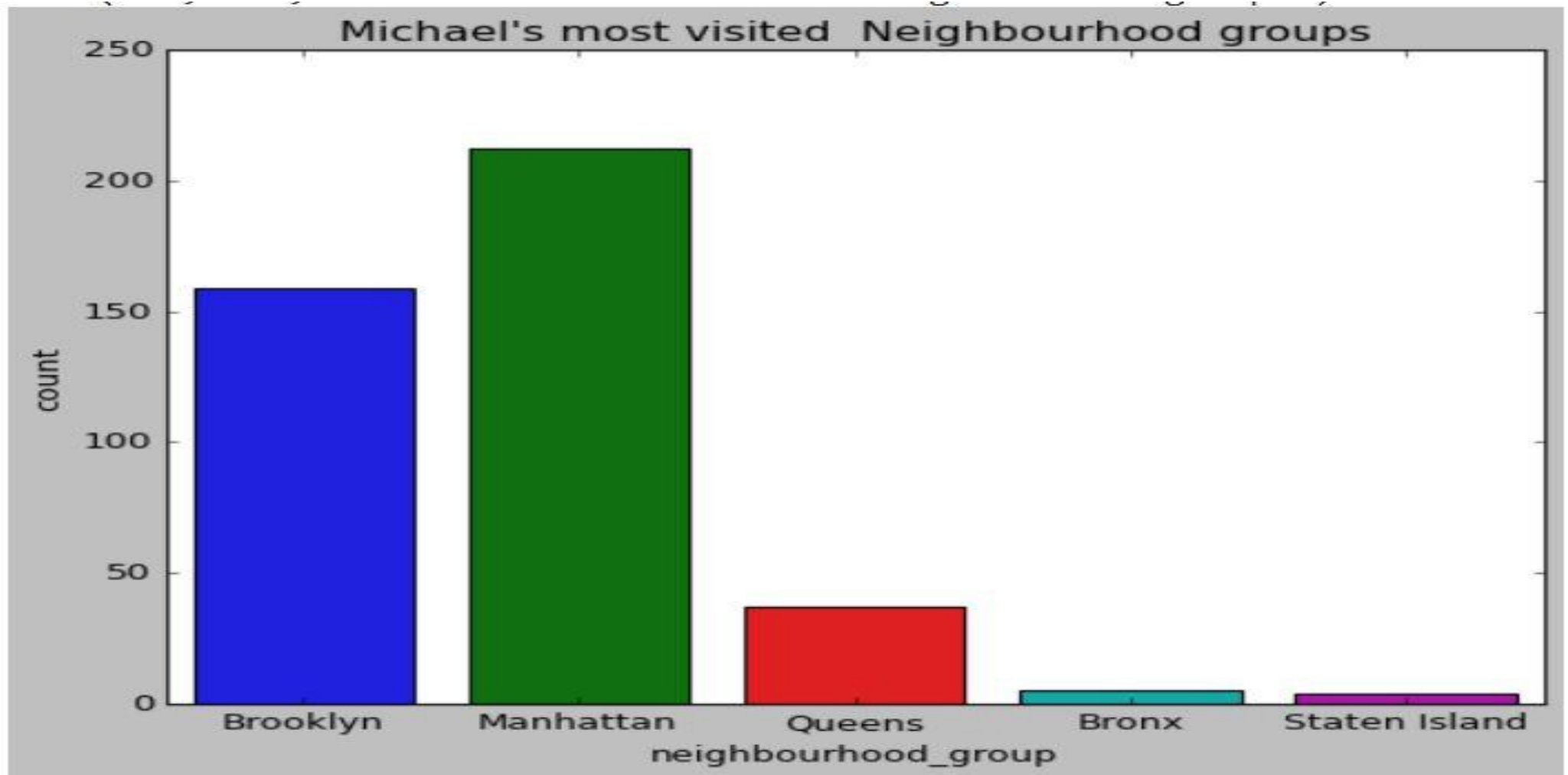
€ EU: Euro	\$ US: Dollar	¥ China: Yuan	¥ Japan: Yen
DM Germany: Deutsche Mark	Fr France: Franc	£ UK: Pound	£ Italy: Lira
R\$ Brazil: Real	C\$ Canada: Dollar	₽ Russia: Rouble	₹ India: Rupee
₧ Spain: Peseta	A\$ Australia: Dollar	Mex\$ Mexico: Peso	₩ South Korea: Won
f Netherlands: Guilder	TL Turkey: Lira	Rp Indonesia: Rupiah	SFr Switzerland: Franc



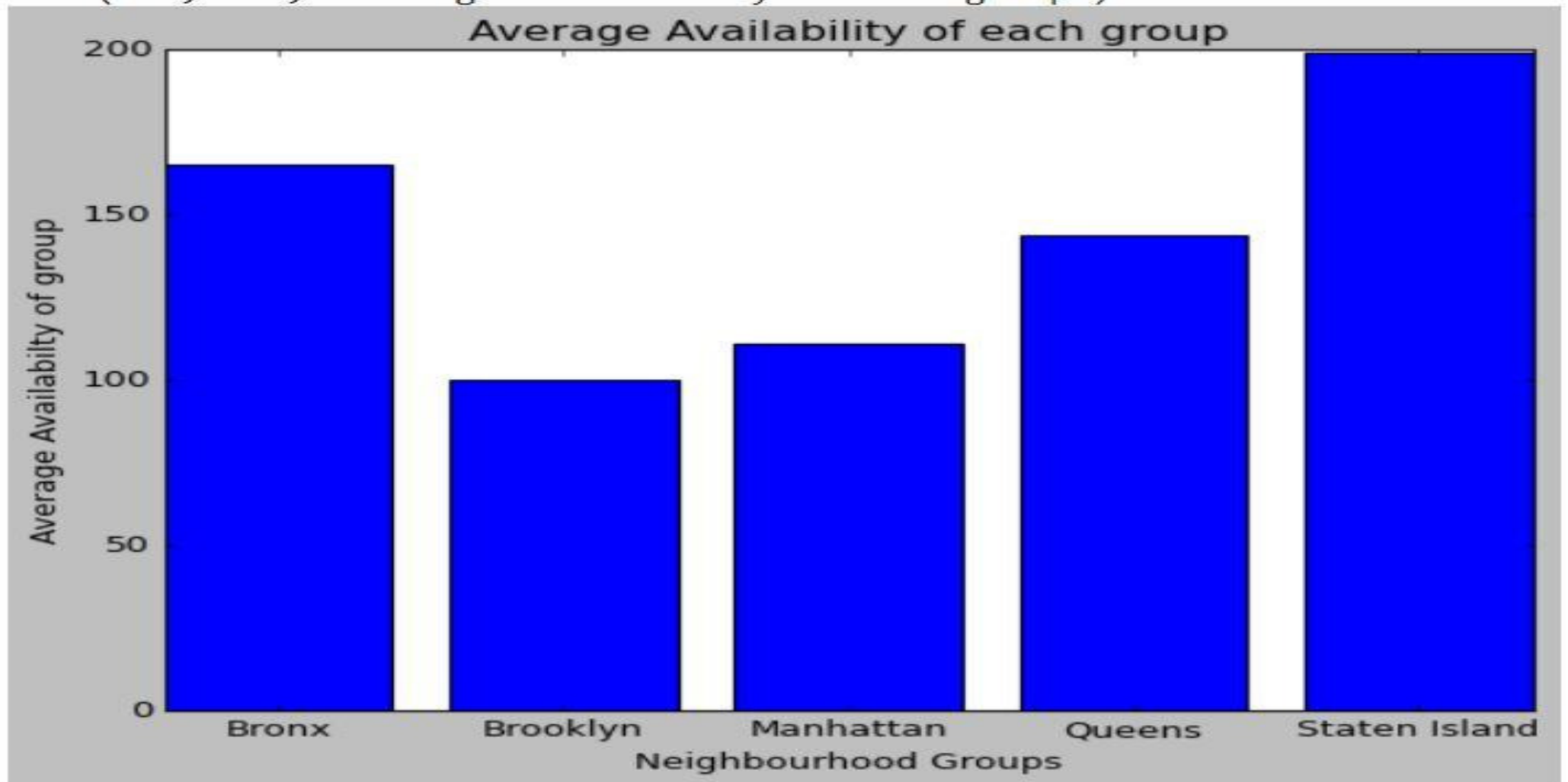
DATA VISUALIZATION



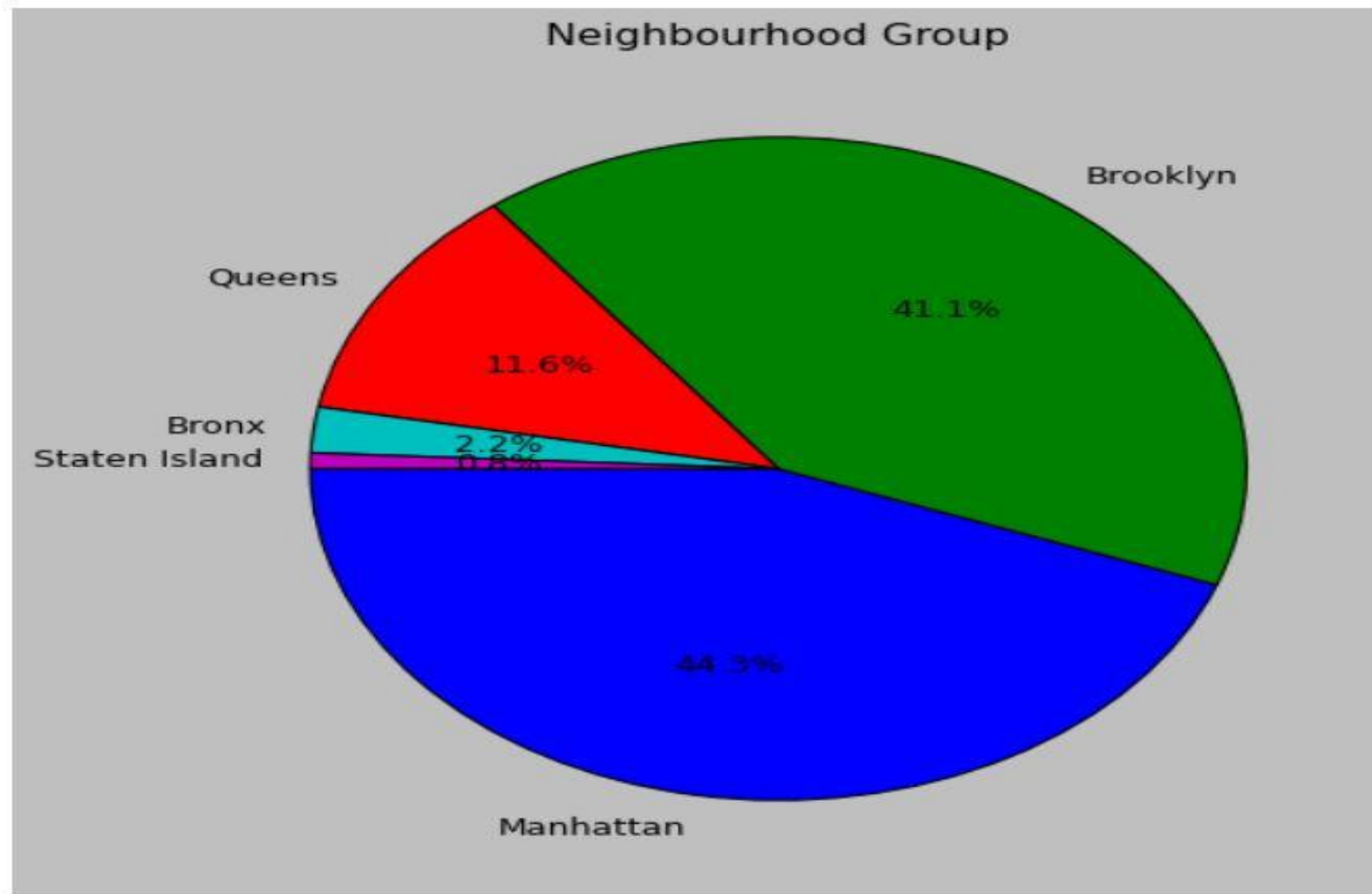
1. Analysis over Host and Areas:



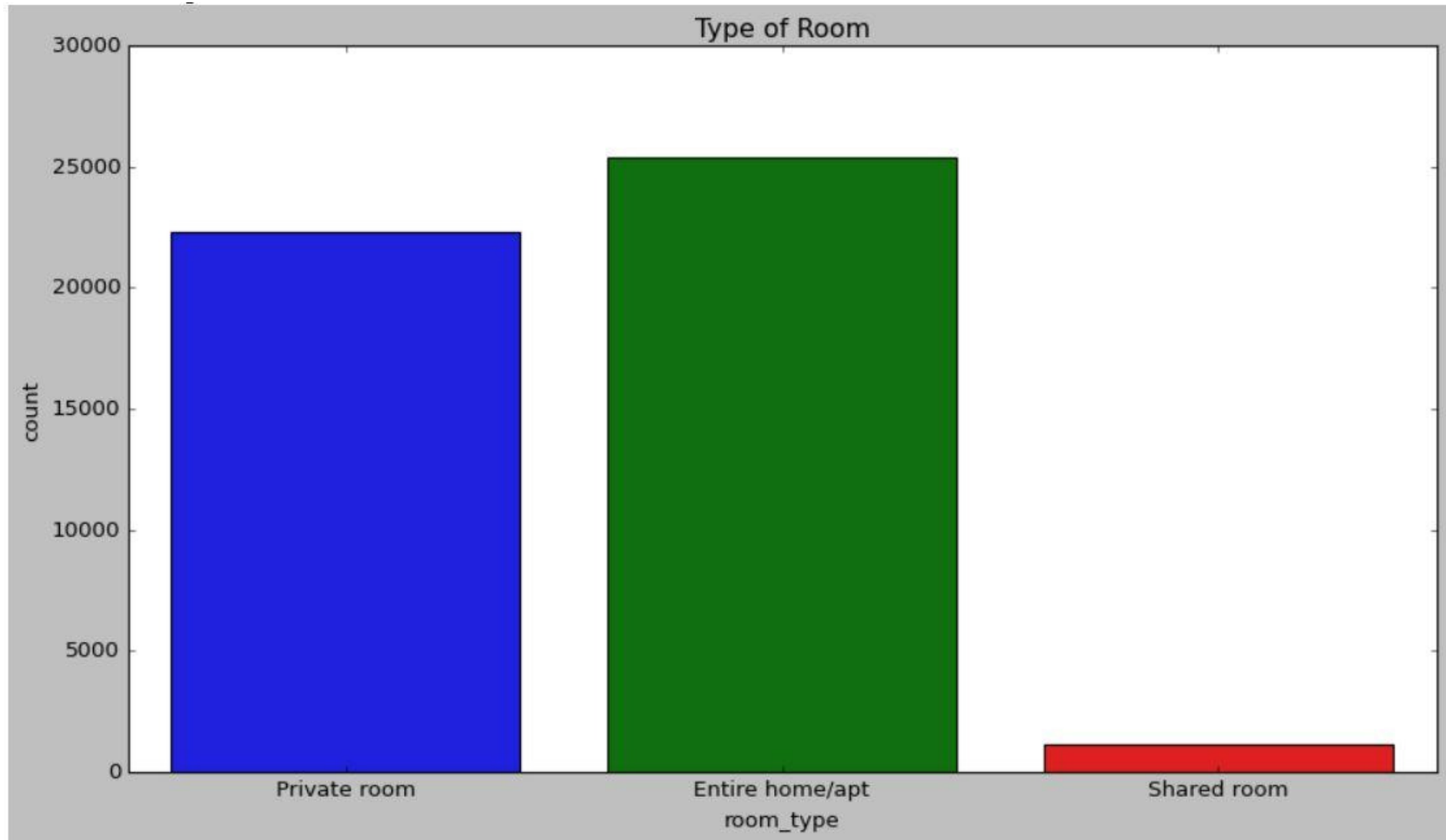
2. Neighbourhood Group vs Average Availability:



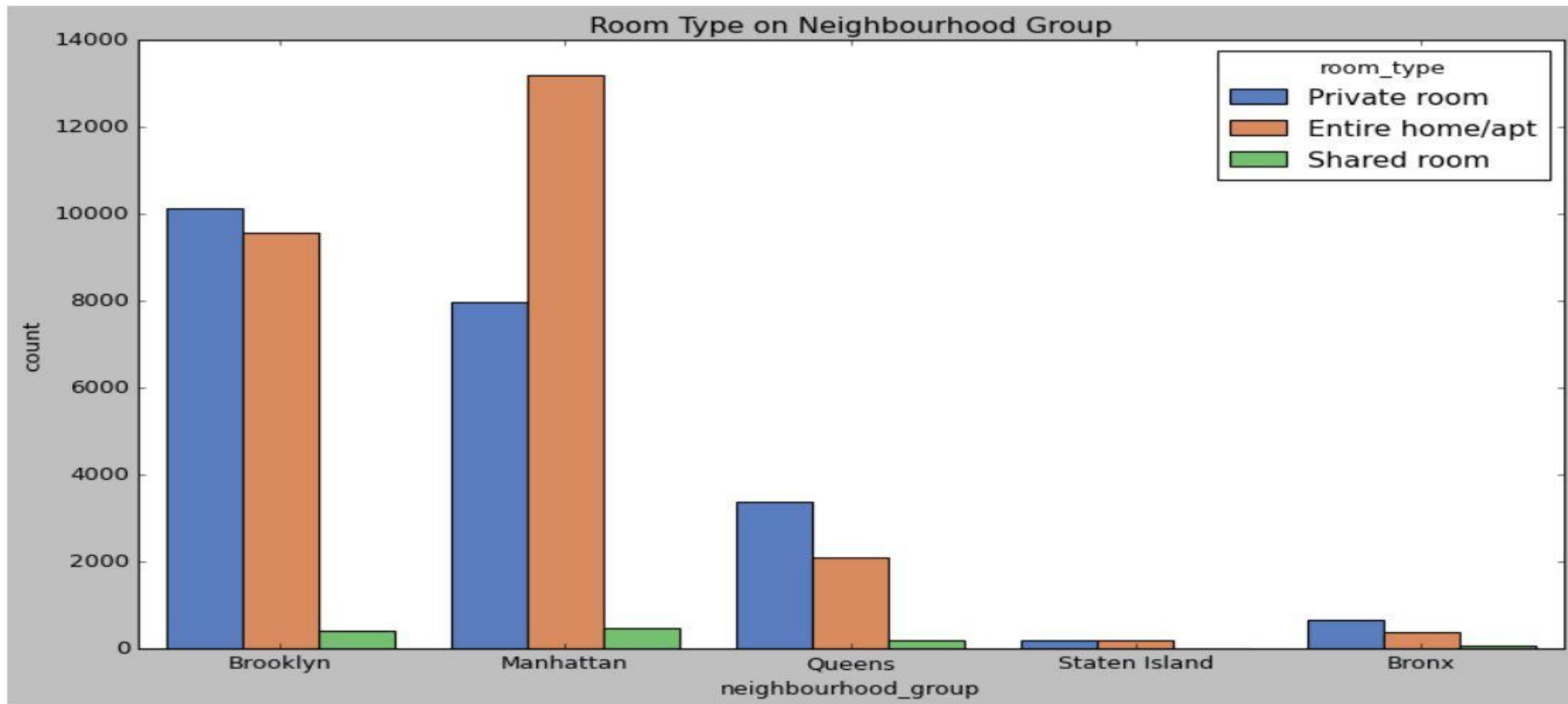
3.Count of each neighbourhood group



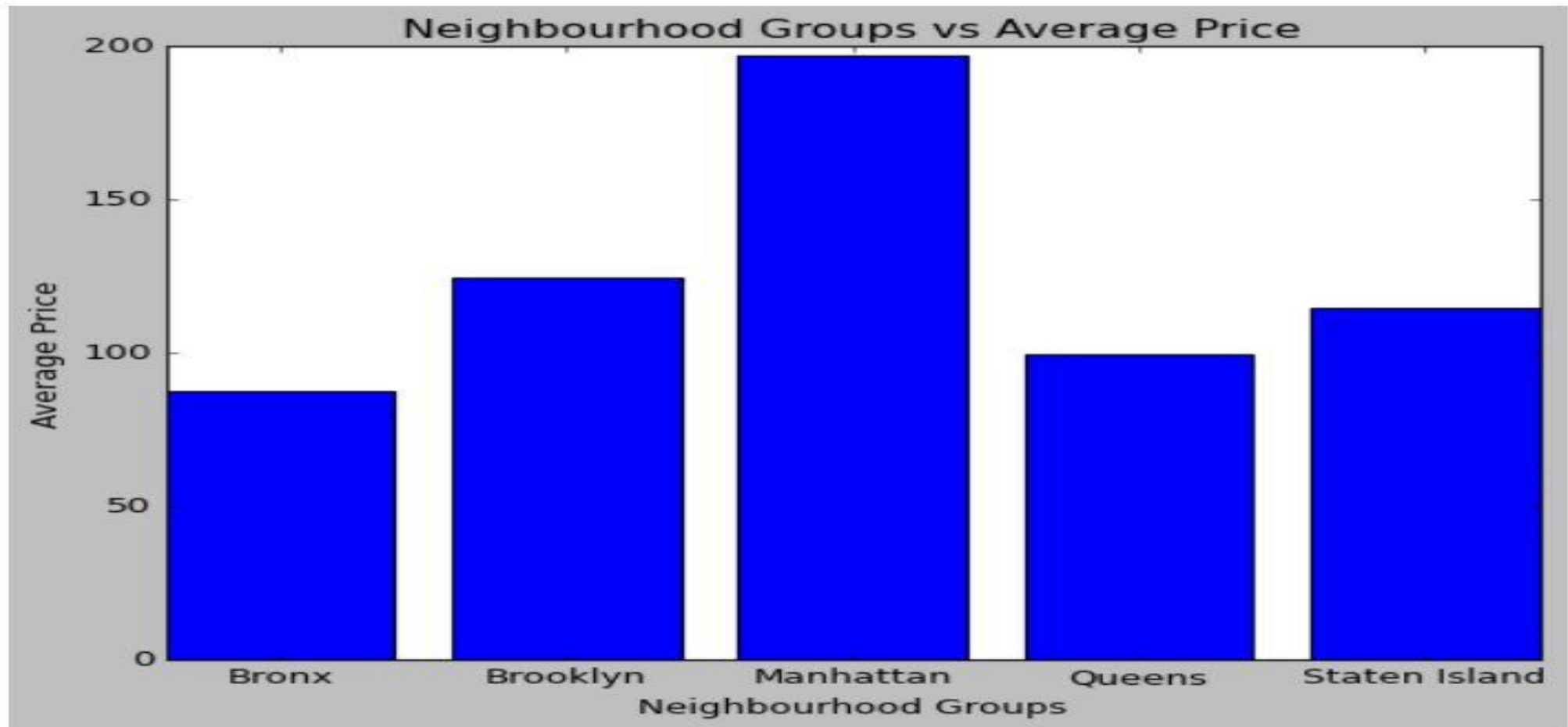
4.Room Type



5.Room Type on Neighbourhood Groups:



6.Price on Neighbourhood Groups



CONCLUSION

We believe that we have conclude our objective on this particular dataset.

Overview of tasks performed:

- ☐ Data Cleaning
- ☐ Hypotheses
- ☐ Data Visualization to support hypotheses

Thank You

