

# Capstone Project - 1

## Airbnb Exploratory Data Analysis

### Team Members

Archana Mehta

Shivansh Singh

# About Airbnb Company



- Airbnb is an online marketplace that connects people who want to rent out their homes with people who are looking for accommodations in that locale.
- It currently covers more than 100,000 cities and 220 countries worldwide.
- This company is located in USA.

# Journey Through Airbnb Dataset

1. Defining Objectives
2. Key Insights From Dataset (EDA)
3. Our Recommendations
4. Appendix:
  - Data sources
  - Data methodology

# Objective

- Key Insights about different Hosts and Neighborhood Group
- How Pricing varies with Room types and Neighborhood Group
- Explore the difference in traffic among different areas
- How reviews varies with different property types
- Which type of property customers prefer
- Most Popular places and room type and its reasons
- Is there any other crucial factor for hosts except pricing

# Data Summary

Airbnb dataset 2018 and 2019

Features = 16 and Observations= 48895

## Some Important Features

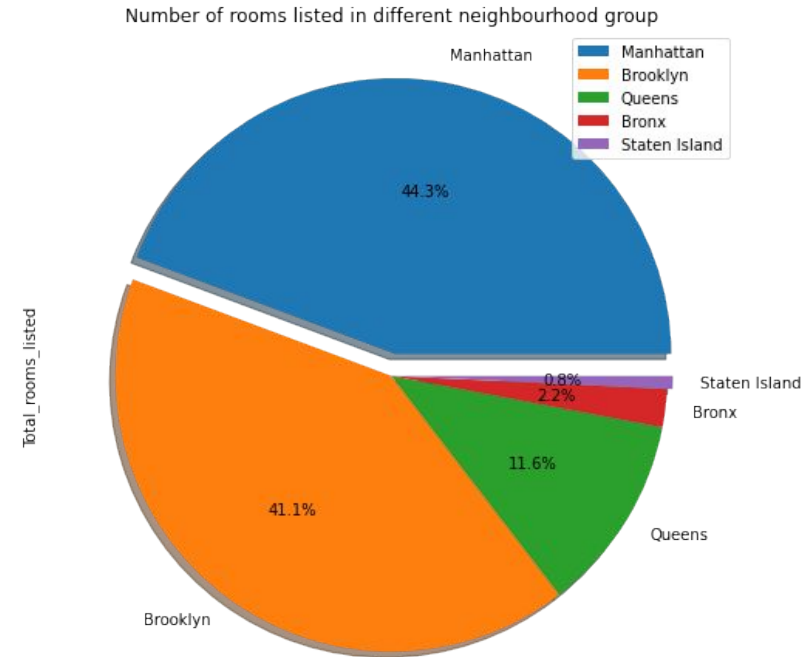
1. **Host\_id** - Unique id given to every host by Airbnb
2. **Price** - Price of the listed room
3. **Room\_type** - type of listed room( shared room, entire home, private room)
4. **Neighbourhood\_group** - Manhattan, Brooklyn, Queens, Staten Island, Bronx
5. **Minimun\_nights** - minimum number of nights the property has to booked for
6. **Number of reviews** - Total number of reviews for particular property

# Data preparation and cleaning

- Data was cleaned for missing values and outliers
- Visualized relation between different parameters using EDA
- Data is also filtered for rows having price = 0

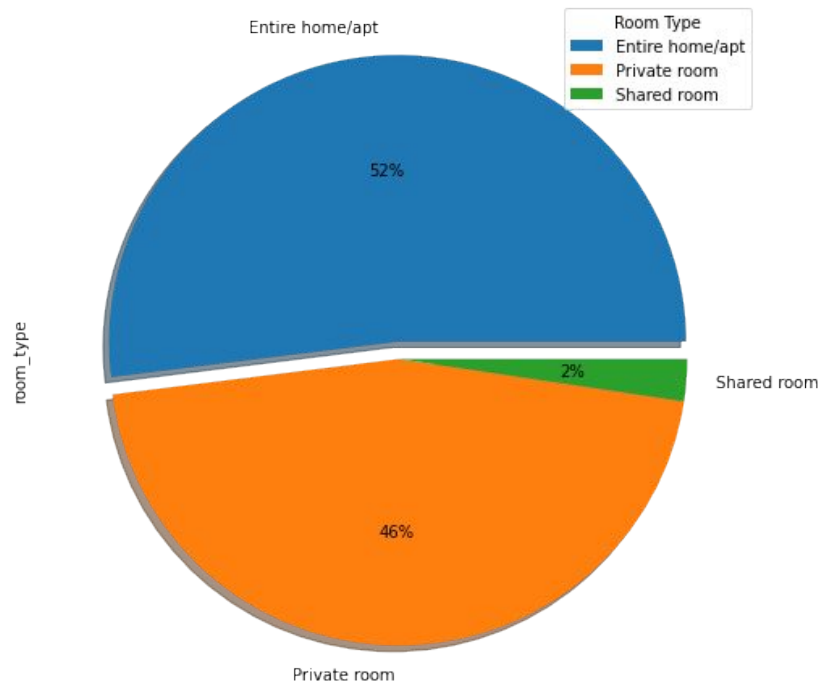
# Rooms on Different Neighborhood Group

- Manhattan and Brooklyn together makes more than 85% total property listing.
- Host either wants to list their property near **Manhattan** or **Brooklyn** Neighbourhood Group.
- May be Hosts are getting the good price for their properties in these two near locations.



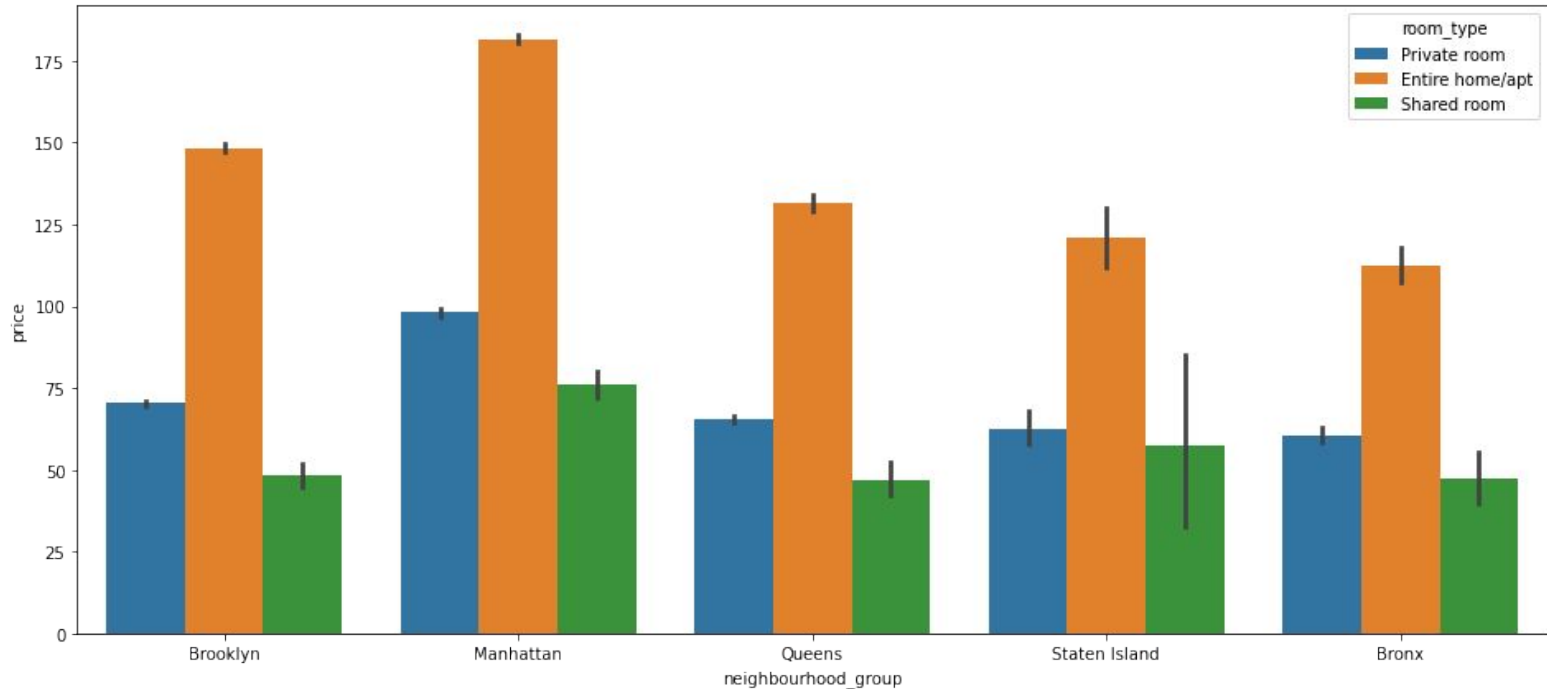
# Total Rooms Type Listings

- 52% Entire home/apt is available for rent.
- While 46% Private Room is available for rent.
- Most of the Host wants to list their entire **home/apt** or **Private Room**.





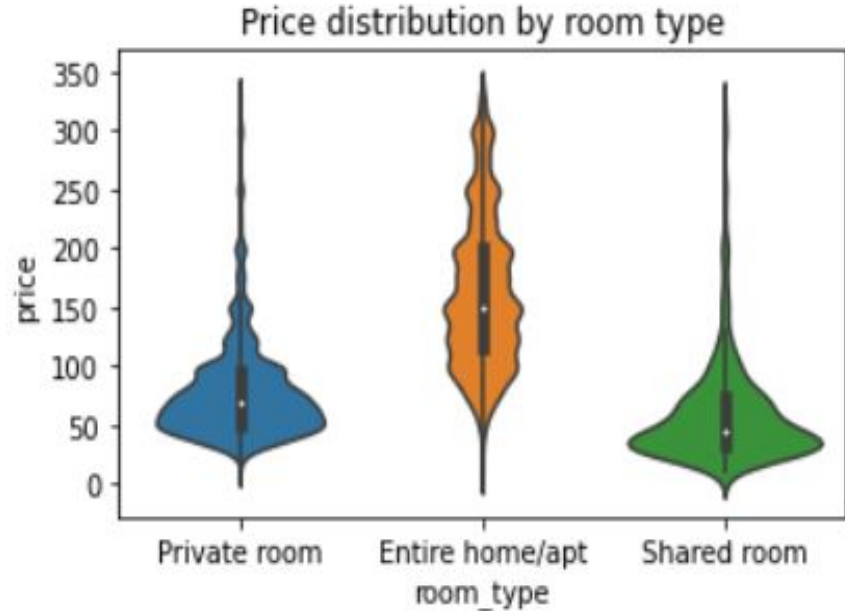
# Average Pricing For Room Types



- Entire Home/apt is the most expensive property in Manhattan and Brooklyn among all the property types and Neighborhood Groups.
- Entire home/apt have \$ **181.644693** avg pricing in Manhattan while same property type stands for \$ **148.218656** in Brooklyn.
- That is the reason why Manhattan and Brooklyn (**Approx 85%**) have the most no. of properties.
- While properties in Neighborhood Groups Bronx are the cheaper one on an average.

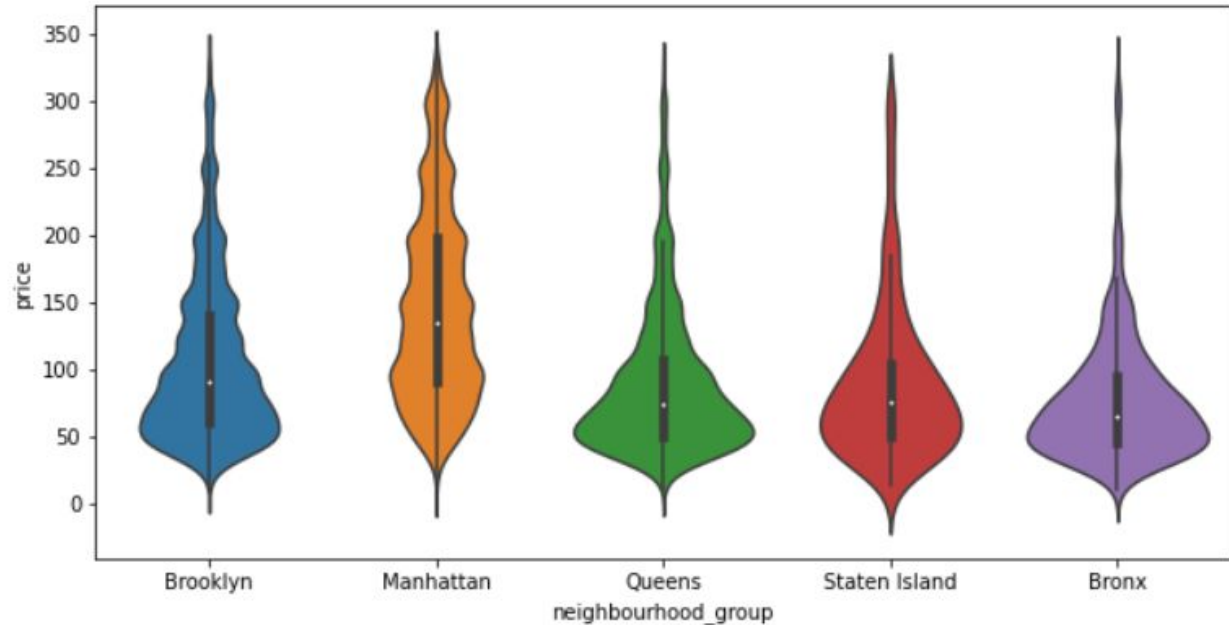
# Price Distribution by room type

- Price range varies from different room type, as expected shared room have high price range.
- Private rooms and shared rooms price range are more centered around their mean.



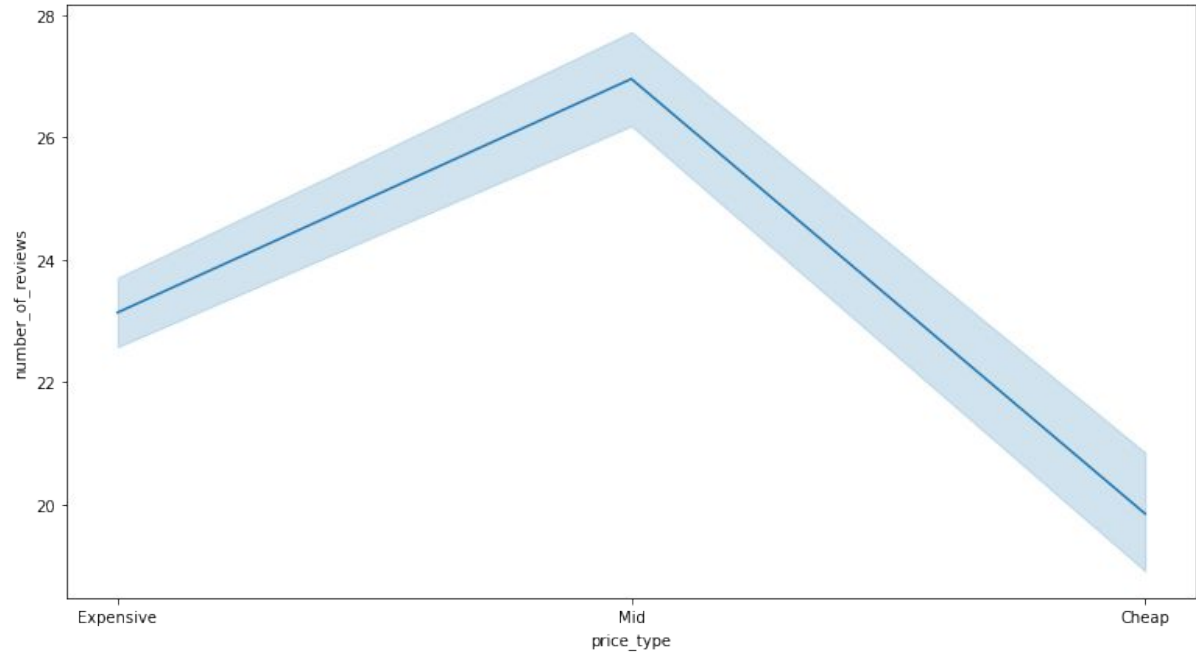
# Price Distribution in Neighborhood Group

- Most premium properties are in Manhattan and Brooklyn.
- Staten Island, Bronx and Queens have cheaper properties compare to Manhattan and Brooklyn.

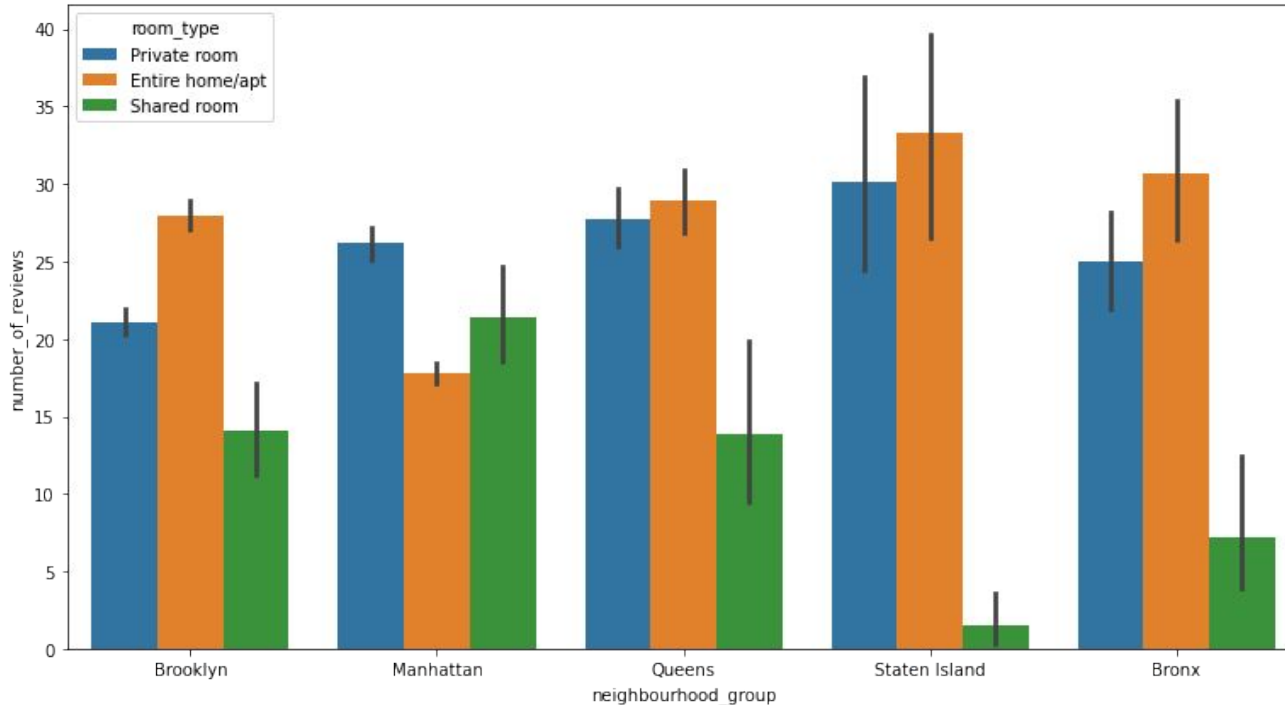


# Relation Between Reviews and Pricing

- Customer prefers Mid Range property that's why it has most no. reviews.
- Cheap property has lowest no. of reviews. We guess, customer don't prefer to review Cheap property (even for booking also.)

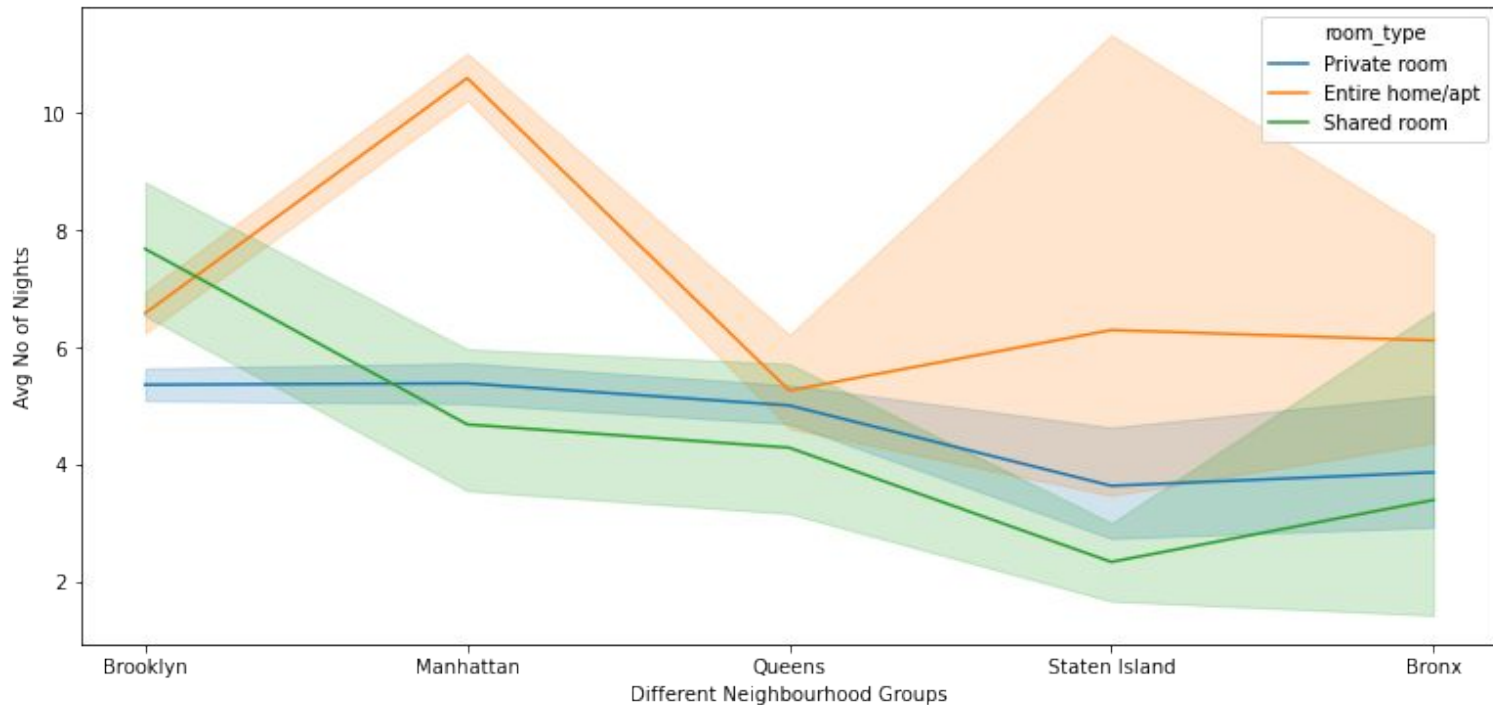


# Type of Properties Having Higher Reviews



- **Entire Room/apt** have the highest no. of reviews in all the neighborhood group while Private Room stands second.
- **Staten Island's Entire Room/apt and private rooms** stands first because it got **most no. of reviews**.
- In Staten Island, Average price for Entire home/apt is about \$ **121.089286** while, for a private room, It is about \$ **62.3** which is in the segment of lowest pricing.
- This may be the crucial factor for this property type to get more bookings.

# Average Property Booking Duration





- Average booking duration for Entire home/apt in Manhattan is higher than any other property type in any Neighborhood Group.
- If a customer wants to book a room for less than 3 days then he/she should go for Shared Room of Staten Island Group.

# Recommendation

1. Pricing and Booking rate is good for **Entire home/apt** in Manhattan and Brooklyn So, Increasing the properties in these 2 near locations will be beneficial.
2. Increase the property listing in the price range 50 \$ to 100 \$.
3. Company should give priority to Entire Home/Apt and Private Room as most of the customers prefer these 2 property type.
4. Minimum number of nights should be decreased (especially for **Entire Home/Apt**).
5. Reviews plays an important role for the host to get more bookings (and revenue), so company should do something so that more customers can review the properties.

# Appendix – Data sources

A. Here is a snapshot of our data dictionary.

- Property information such as property ID, name, minimum nights, price, ratings and location
- Customer information as host ID

B. We used the following data sources:

- Past booking data 2018-2019

## Appendix – Data methodology

B. We conducted a deep analysis of Airbnb NYC dataset. The process included:

- Data Cleaning through Python
- Used box plots to remove outliers
- Used heat maps to visualize correlation between the variables.
- Data Visualization