Customer Preferences while Booking Airbnb Listings

Agenda

- ❑ Objective
- ❑ Background
- ❑ Key findings
- ❑ Recommendations
- ❑ Appendix:
  - ➢ Data sources
 - ➢ Data methodology
  - ➢ Data model assumptions

Objective

❑ To understand the insights based on the types of properties, neighborhoods, and relationships between variables.

❑ Improve our shared understanding of the most popular localities in New York.

❑ Improve our shared understanding of the categorization of our customers based on their preferences(neighborhoods, pricing ranges, kinds of properties, etc.).

Background

- For the past few months, Airbnb has seen a major decline in revenue due to covid.

- Now that the restrictions have started lifting and people have started to travel more, bookings may increase.

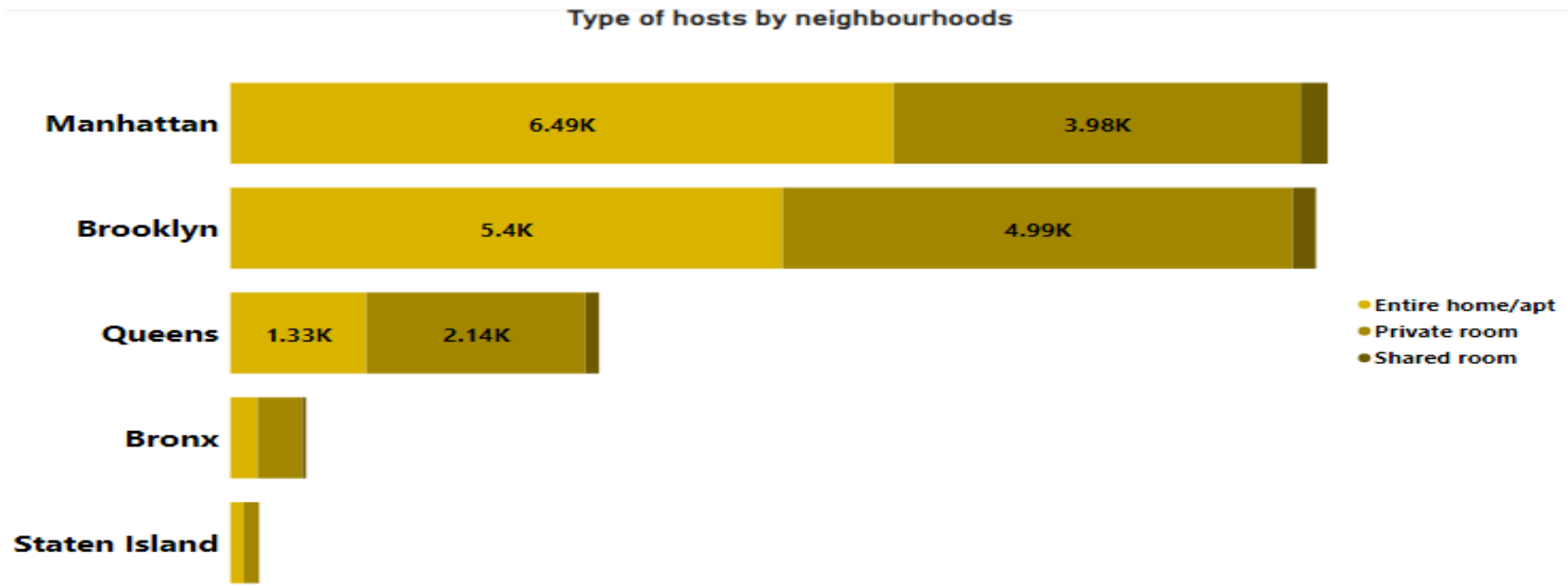- Airbnb wants to make sure that it is fully prepared for this change.

Entire home/Apartment is preferred more

- ❑ We grouped the data according to room type and found that Entire Home/Apartment is the most booked property with **52.45%** of total bookings.

- ❑ Private Rooms are next with a whopping **45%**.

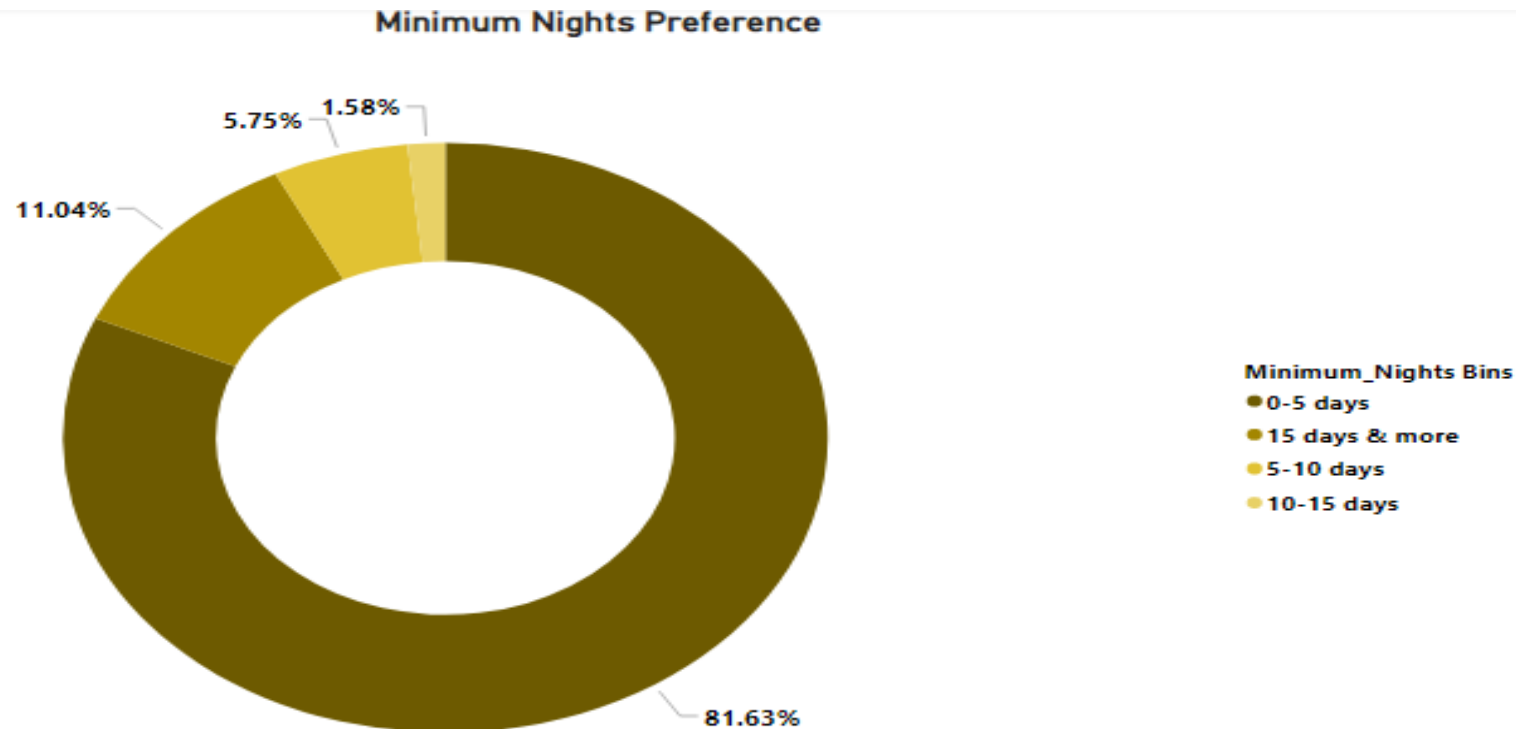- ❑ Shared Rooms have the lowest share with a minute of **2.56%**.

Most popular properties

2.56%

45%

52.45%

Room Type
- Entire home/apt
- Private room
- Shared room
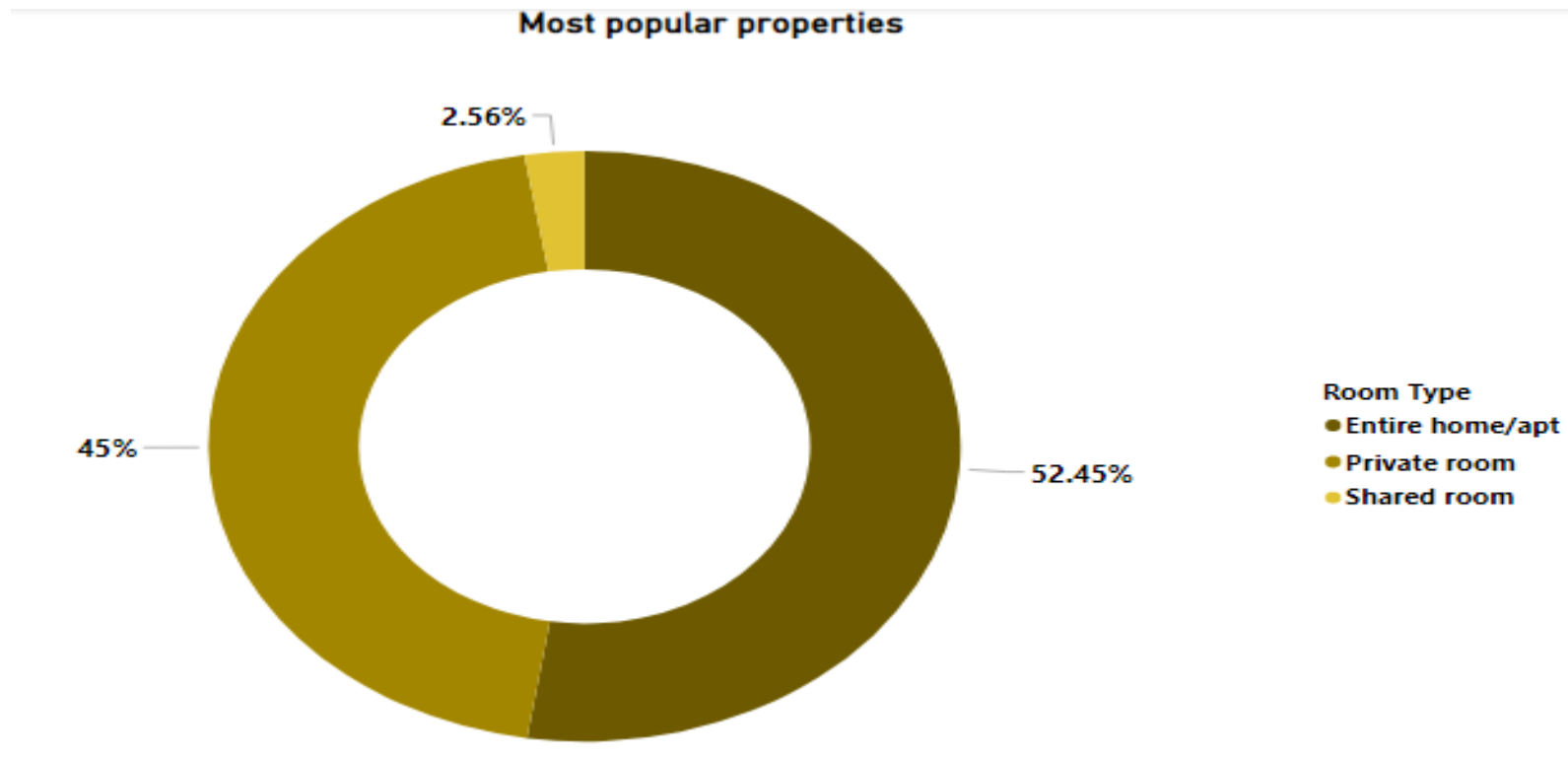
Entire Home/apt and Private room have dominating numbers

❑ We grouped the room type according to the neighborhood and found that Brooklyn and Manhattan have almost **90%** as Entire/home apt and Private rooms.

❑ Same goes with Queens, the Bronx and Staten Island have more than **90%** as Entire/home apt and Private rooms with negligible properties as Shared rooms.

❑ Shared rooms should be prioritized in order to increase the presence.

Type of hosts by neighbourhoods

Neighbourhood	Entire home/apt	Private room
Manhattan	6.49K	3.98K
Brooklyn	5.4K	4.99K
Queens	1.33K	2.14K
Bronx		
Staten Island		

● Entire home/apt
● Private room
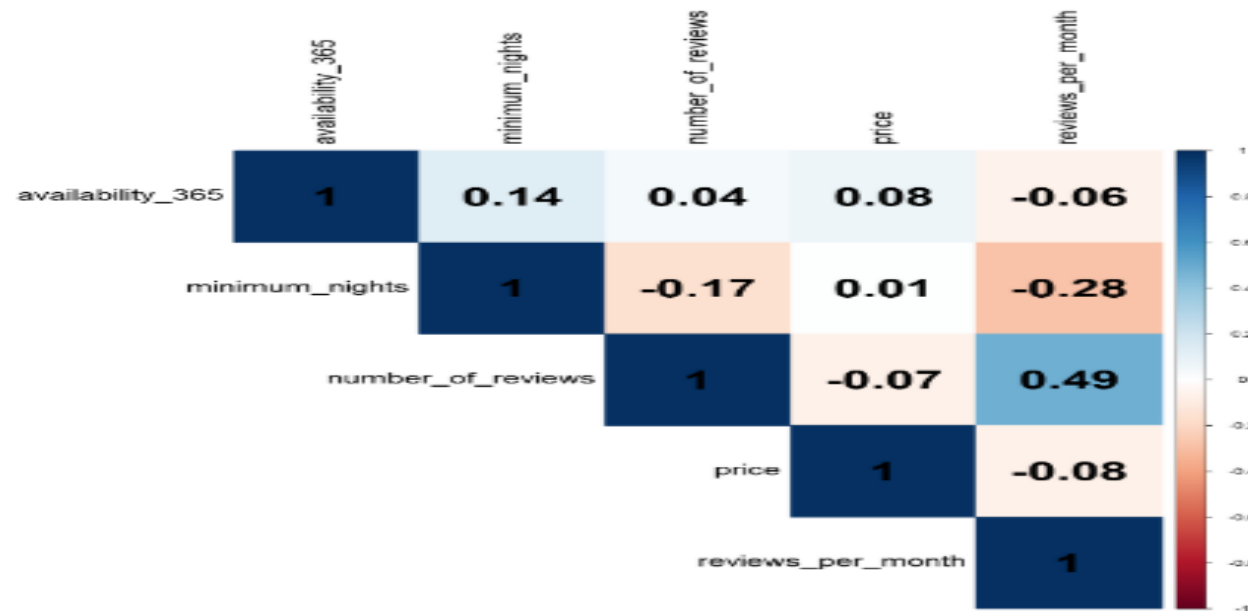● Shared room

Customers prefer shorter stays

- ❑ We created different bins of price ranges (0-5 days, 5-10 days, etc.) and counted them.

- ❑ Majority of customers preferred to stay for **0 to 5 days with 81.63%**. Surprisingly the second was **15 days or more with 11.04%**.

- ❑ Therefore our customers prefer a stays **lesser than a week**.

Minimum Nights Preference

Minimum_Nights Bins
- ● 0-5 days
- ● 15 days & more
- ● 5-10 days
- ● 10-15 days

5.75%   1.58%

11.04%

81.63%

Correlation between variables affecting bookings

❑ **Availability_365** has a positive correlation with almost all the other variables.

❑ **Minimum_nights** has a negative relation with **number_of_reviews** implying that as the minimum night increases there are a lesser number of reviews (fewer bookings).

❑ **Minimum_nights** also has a negative relation with **price** implying that as the price increases there are a lesser number of reviews (fewer bookings).

Assumption: Number of reviews means the number of bookings for a listing/property.

Recommendations

- ❑ Focus on Hosts which are providing stays **shorter than a week or a maximum of 4 weeks**.

- ❑ Acquire hosts whose prices are in the range of **0-500** dollars.

- ❑ Acquire more **Entire Home/apt** and **Private room** hosts than shared rooms as they are in demand.

Appendix - Data Sources:

❑ Here is a snapshot of our data dictionary:

➢ Property demographics such as neighborhood, neighborhood group, latitude, and longitude.

➢ Room information such as type, minimum nights, and room price.

➢ Customer likeliness attributes such as the number of reviews, last review, etc.

❑ The team used the following data sources:

➢ Airbnb listings data for New York City.

➢ Customer feedback data for the listings.

Appendix - Data Methodology

❑ We conducted a thorough analysis of the Airbnb data. The process included:

➤ Cleaning and transforming the data set using the Power BI tool.

➤ Exploratory data analysis using the PowerBI tool.

➤ The **Methodology document file** has been included separately with a detailed Data Cleaning, Data Preparation, and Data Visualization process.

Appendix - Data Assumptions

❑ There were certain assumptions made due to data unavailability:

➢ We assumed that only the listings which had been reviewed were actually booked.

➢ If the availability is zero for any listing then it is not available for booking.

➢ If the prices for listings were zero then they were not a genuine listing/property.