

NYC AIRBNB Case Study IIIT-B

Methodology Document

For this case study, we have used Jupiter Notebook for initial analysis of the data and Tableau for further analysis and visualisation.

```
1 # Importing necessary Libraries
2 import warnings
3 warnings.filterwarnings("ignore")
4 import numpy as np
5 import pandas as pd
6 import matplotlib.pyplot as plt
7 %matplotlib inline
8 import seaborn as sns
```

```
1 # importing the Data
2 df = pd.read_csv("BNB.csv")
3 df.head()
```

	id	name	host_id	host_name	neighbourhood_group	neighbourhood	latitude	longitude	room_type	price	mir
0	2539	Clean & quiet apt home by the park	2787	John	Brooklyn	Kensington	40.64749	-73.97237	Private room	149	1
1	2595	Skyliit Midtown Castle	2845	Jennifer	Manhattan	Midtown	40.75362	-73.98377	Entire home/apt	225	1
2	3647	THE VILLAGE OF HARLEM....NEW YORK I	4632	Elisabeth	Manhattan	Harlem	40.80902	-73.94190	Private room	150	3
3	3831	Cozy Entire Floor of Brownstone	4869	LisaRoxanne	Brooklyn	Clinton Hill	40.68514	-73.95976	Entire home/apt	89	1
4	5022	Entire Apt. Spacious Studio/Loft by central park	7192	Laura	Manhattan	East Harlem	40.79851	-73.94399	Entire home/apt	80	10

```
1 # Calculating the missing values in the dataset
2 df.isnull().sum()
```

```
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
```

```

1 # There are certain columns that are insignificant to the dataset and our analysis hence we will drop t
2 df.drop(['id', 'name', 'last_review'], axis = 1, inplace = True)

1 # Now reviews_per_month contains more missing values which should be replaced with 0 respectively
2 df.fillna({'reviews_per_month':0},inplace=True)

1 df.reviews_per_month.isnull().sum()

0

1 # Checking the unique values of other columns'
2 df.room_type.unique()

array(['Private room', 'Entire home/apt', 'Shared room'], dtype=object)

1 len(df.room_type.unique())

3

1 df.neighbourhood_group.unique()

array(['Brooklyn', 'Manhattan', 'Queens', 'Staten Island', 'Bronx'],
      dtype=object)

1 len(df.neighbourhood_group.unique())

5

1 len(df.neighbourhood.unique())

221

```

Step 2: Data Wrangling

- The dataset was checked for duplicate rows, and none were found.
- Null values were found in the columns name, host-name, last review, and review-per-month.
- The column “name” was dropped because the number of missing values was small and dropping it would not have a significant

impact on the analysis. Other irrelevant columns like “last review” and “ID” were also dropped.

- The formatting of the dataset was checked.
- Outliers were identified and reviewed.

Step 3: Data Visualisation and Analysis

We created visualizations in Tableau to analyze various attributes that could potentially increase Airbnb's revenue such as:

- Which type of hosts to acquire more and where?
- The categorization of customers based on their preferences.
- What are the Neighborhoods they need to target?
- What is the pricing ranges preferred by customers?
- The various kinds of properties that exist w.r.t. customer preferences.
- Adjustments in the existing properties to make it more customer oriented.
- What are the most popular localities and properties in New York currently?
- How to get unpopular properties more traction?

Created a Calculated Field for Earnings

Earnings

×

[Price]*[Number Of Reviews]

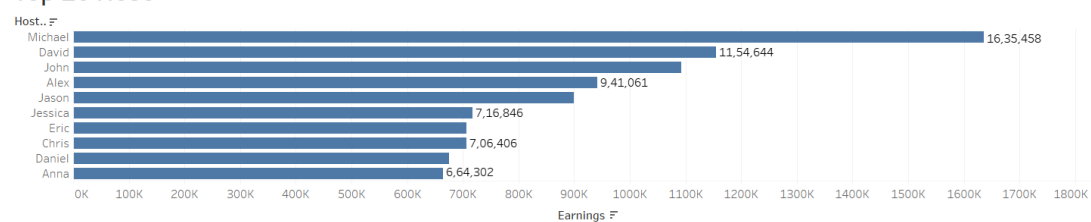
The calculation is valid.

2 Dependencies ▾

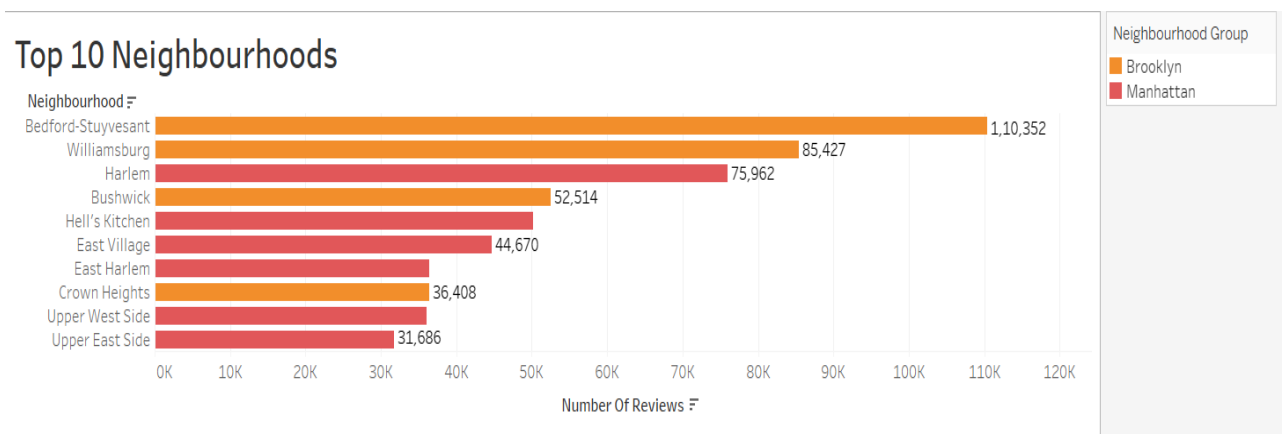
Apply

OK

Top 10 Host

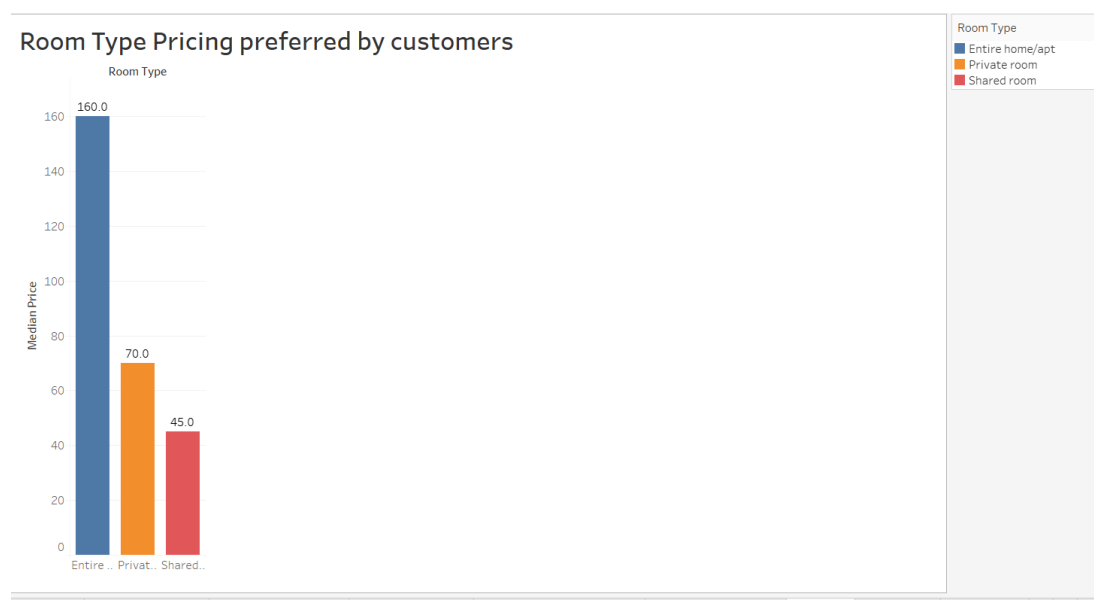


The graph shows the top 10 hosts who are earning more. We observed the host Michael has the highest number of properties and earnings.



These Neighborhoods are the most targeted by customers. It could be mostly visited as they are near a tourist spot or the facilities are best in this locality.

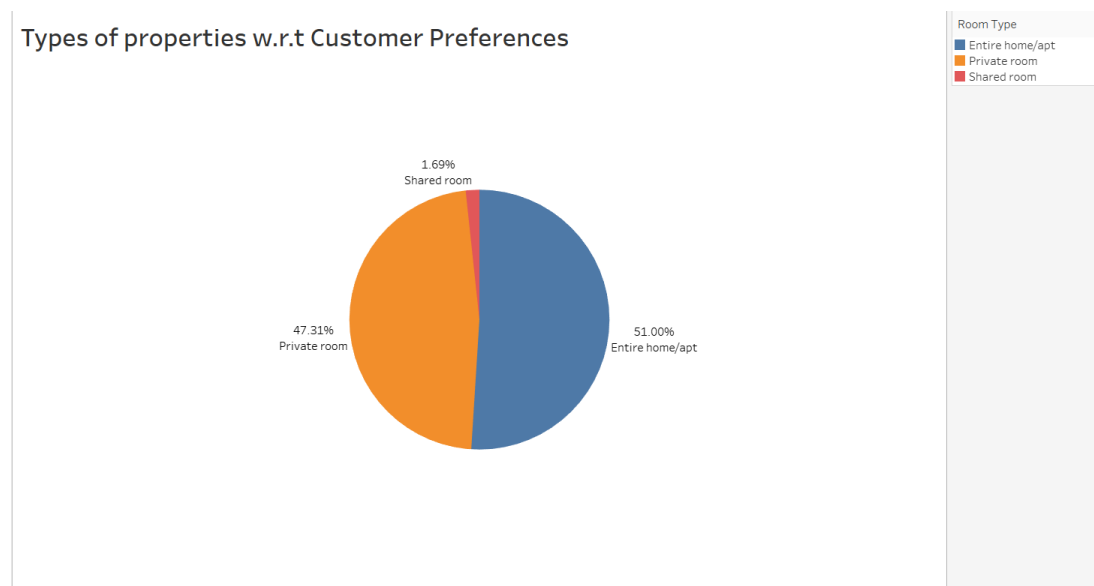
Room Type Pricing preferred by Customers



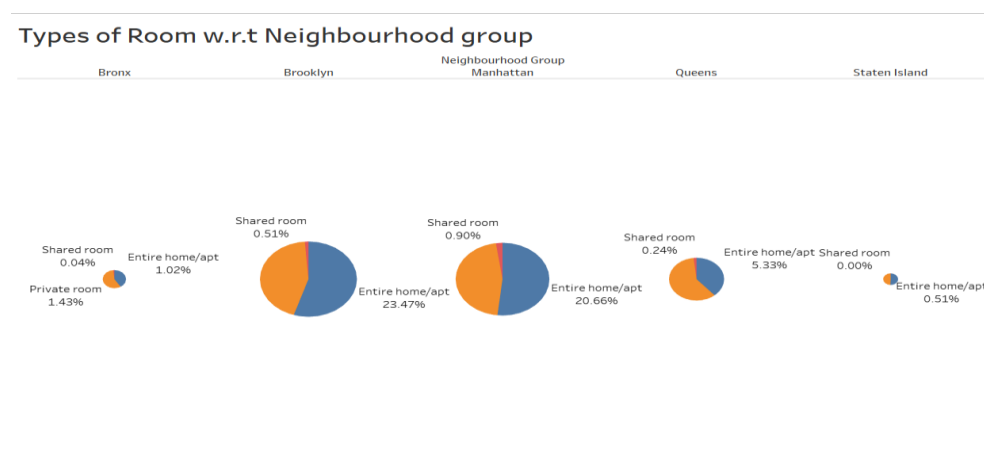
We observed, the customers prefer entire home the most. The pricing range is 160 for entire home, then private room 70 and least is shared room which is 45.

Types of properties preferred by customers

Most of the customers prefer entire home (51%) and then private room (47.31%). So, AirBnB should increase these type of rooms more and provide discounts to garner more customers. On the other hand, Shared rooms are preferred the least (1.69%) by the customers.

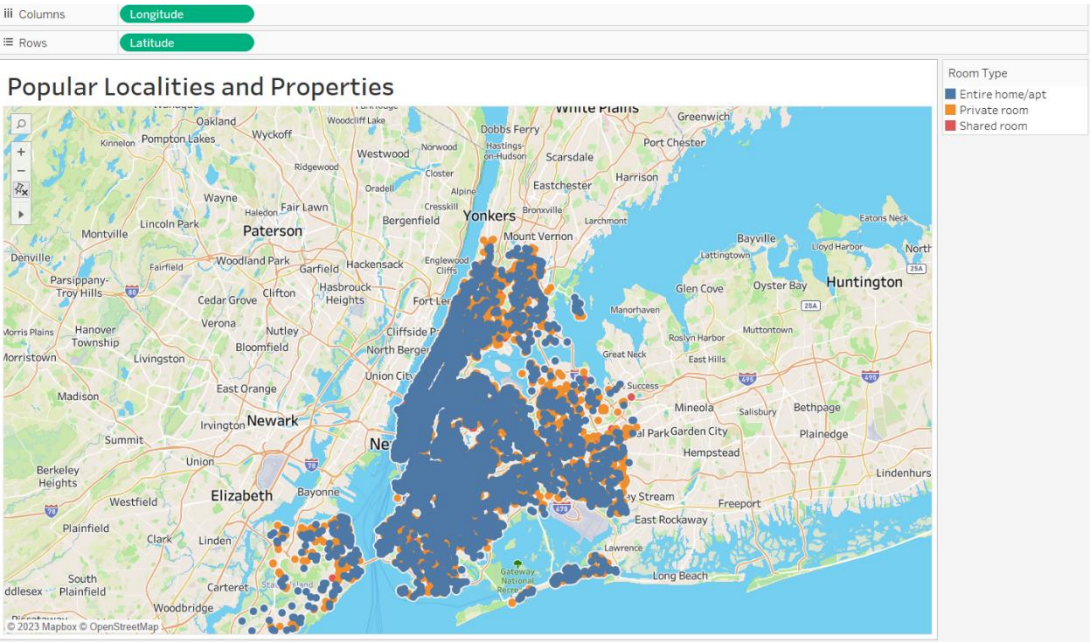


Type of room type preferred by customer's w.r.t Neighborhood group:

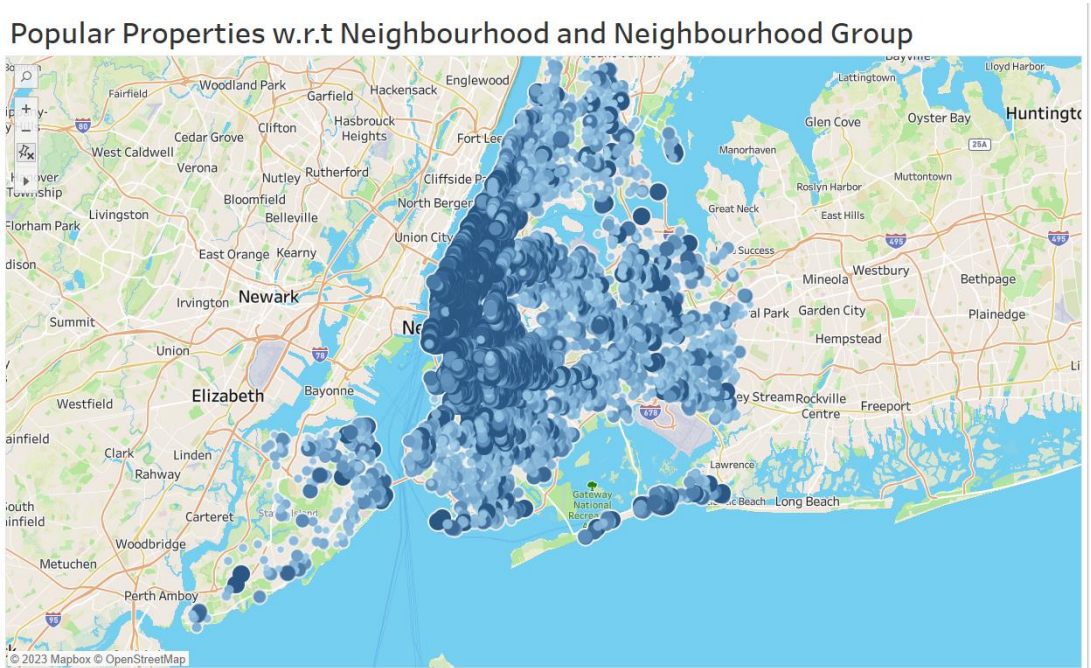


The graph above shows that in Manhattan, Brooklyn, Staten Island entire home is preferred by customers whereas in Bronx and Queens private rooms are preferred.

Popular Localities

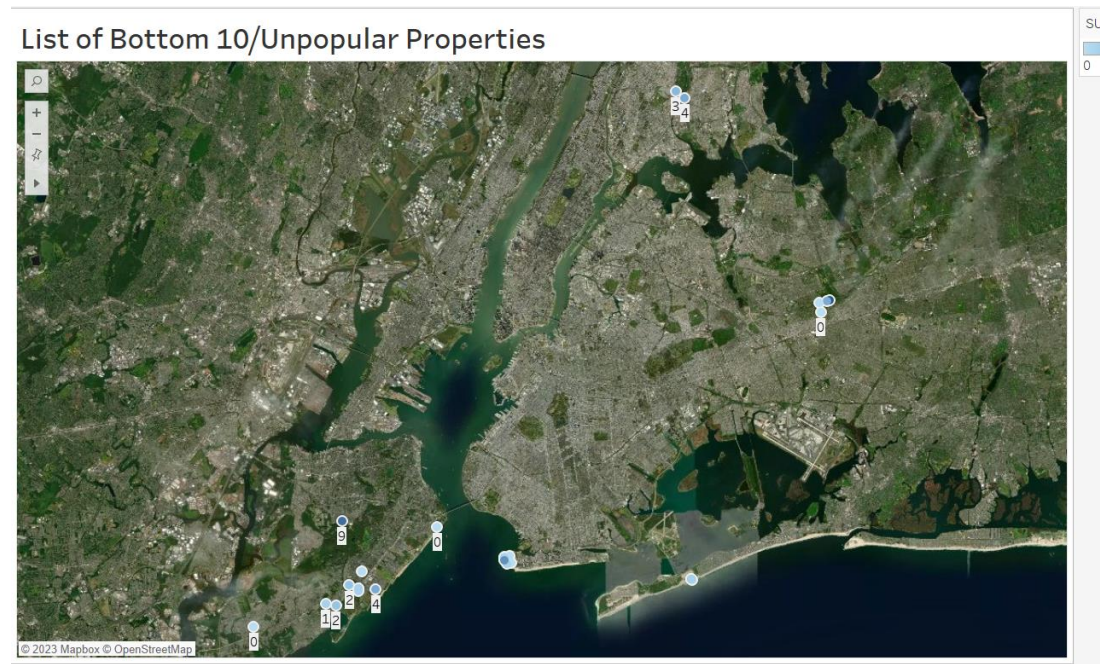


This map shows that Manhattan, Brooklyn, are popular than Queens, Bronx and Staten islands for entire home. They are depicted in the map in the darker shade



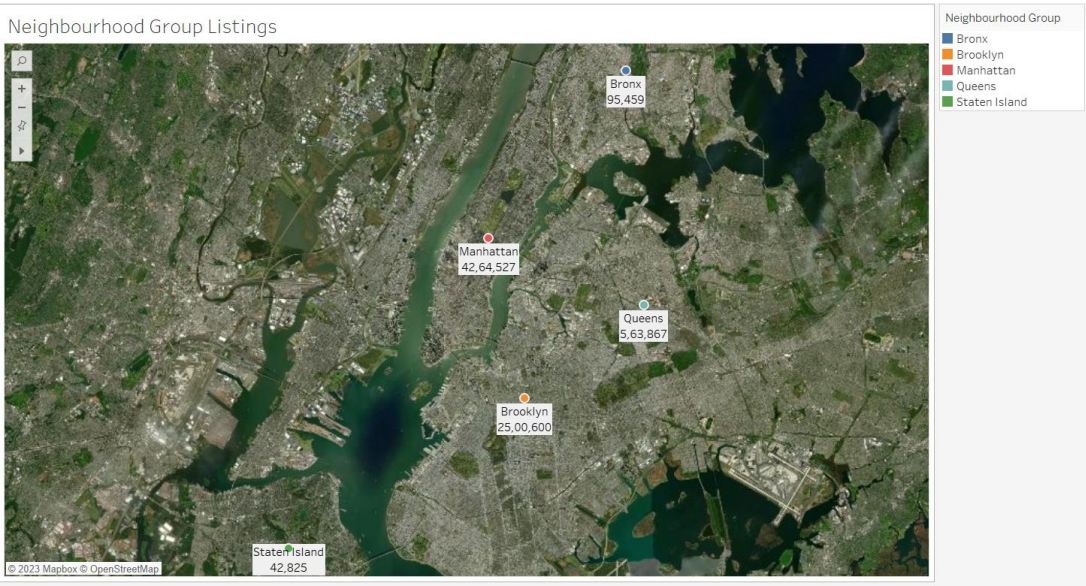
According to the map Manhattan, Brooklyn, Queens are popular than Bronx and Staten islands for entire home. They are again shown in the darker shade in the map.

List of Bottom 10 Properties



The map shows the list of the bottom 10 properties. Two possible reasons could be that these properties are located far off on the outskirts or these do not have any tourist spots.

Neighbourhood Group Listings



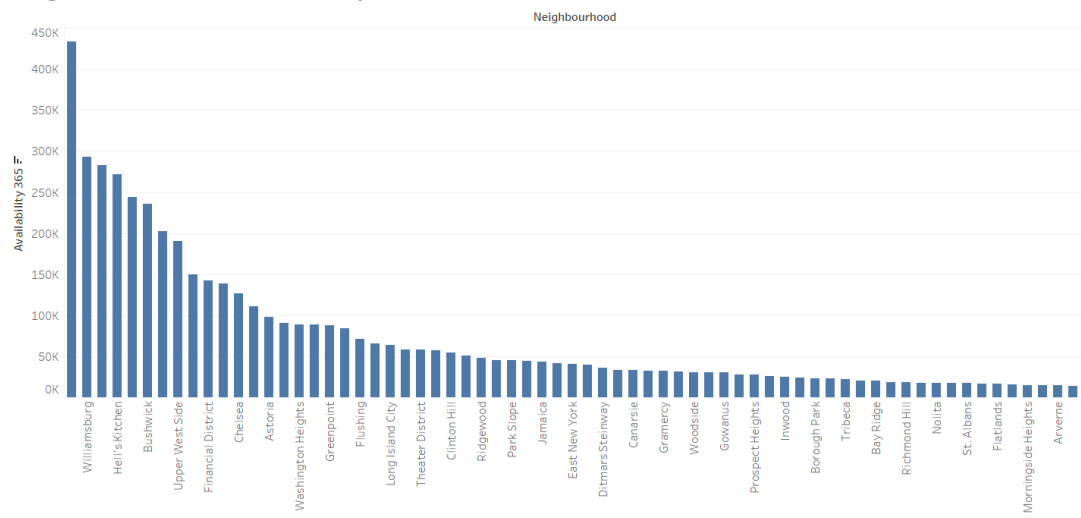
The map shows that Manhattan is more expensive and caters to high-end tourists due to its many tourist attractions. Brooklyn, on the other hand, is more affordable and preferred by locals and budget-conscious tourists.

Average Price of Neighbourhood Group



The average price of listed properties in Manhattan is around 196.9, which is highest among all neighbourhoods. Average price for Brooklyn is second highest i.e. 124.4. Bronx appears to be an affordable neighbourhood as the average price is almost half than Manhattan's average price

Neighbourhood and Availability



The graph shows that the availability of Bedford is highest and its price is on the lower side. It is a good choice for customers. After Bedford, Harlem follows the same trend. Chelsea's availability low but it is costly. On the other hand, William's price is high and has average availability.

Recommendations:

- Airbnb can increase bookings by promoting shared rooms with targeted discounts. They can also acquire new properties, such as private rooms in Manhattan and Brooklyn, and entire homes in Bronx and Queens. These areas have a higher number of bookings, and weekly or bi-weekly rentals can also be acquired to help customers who are stranded in New York City for quarantine purposes.
- The prices of private rooms in Manhattan and Brooklyn, and entire homes in Bronx and Queens are also favorable. Brooklyn has an average price of \$113, which is lower than Manhattan's average price. Therefore, Airbnb can consider expanding to Brooklyn.
- Overall, increasing acquisitions and new properties in coastal regions, as well as promoting shared rooms with discounts, can help Airbnb increase customer bookings.
- Here are some specific examples of how Airbnb can promote shared rooms with targeted discounts:
 - They can offer discounts for bookings made during the off-season.
 - They can offer discounts for bookings made by groups of travelers.
 - They can offer discounts for bookings made by members of the military or other organizations.
 - They can partner with businesses or organizations to offer discounts to their employees or members.

By promoting shared rooms with targeted discounts, Airbnb can attract more travelers and increase bookings.

- All cities except Manhattan and Brooklyn need to change their marketing strategy to increase bookings.
- Most customers are willing to spend between \$40 and \$160 on accommodation. Therefore, these cities should try new marketing tactics to attract customers, such as offering discounts and promotions.

- Unpopular cities need to change their strategy to increase revenue. This could involve creating tourism attractions, improving infrastructure, or starting recreation and sports clubs.
- They could also try to increase the purchasing power of their customers by offering financial assistance or other incentives.
- More coastal acquisitions and new construction could also lead to increased bookings from customers.