EXECUTIVE PG PROGRAMME IN DATA SCIENCE
IIIT BANGALORE

Airbnb, NYC
Case Study

A presentation for Head of Acquisitions and Operations &
Head of User Experience

BY : SHREYAS DUBEY AND KEERTHI ADEP

AGENDA

- **OBJECTIVES**
- **BACKGROUND**
- **KEY FINDINGS**
- **RECOMMENDATIONS**
- **APPENDIX :**
  - **DATA SOURCES**
  - **DATA METHODOLOGY**
 - **DATA MODEL ASSUMPTIONS**

OBJECTIVES

- **Improve our shared understanding about impact of Minimum Nights and Availability on Airbnb bookings.**

- **Applying the understanding we gained from the previous step and find the weak links which can give us losses**

- **Convey important insights which can eventually maximize chances of more bookings.**
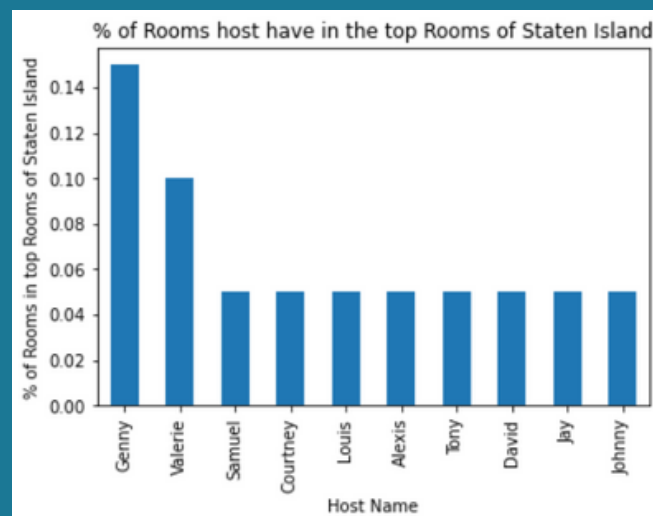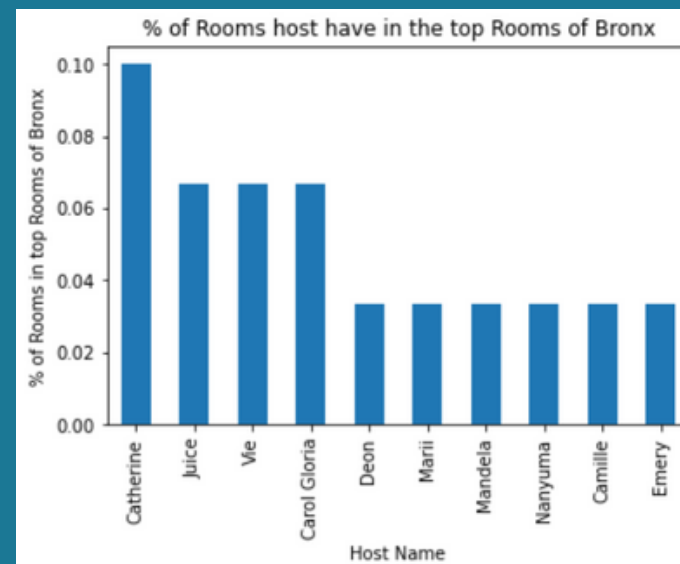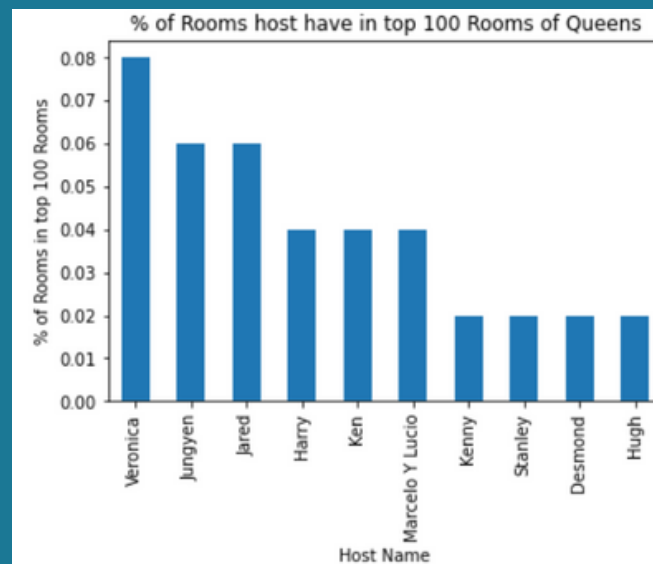
BACKGROUND

- **Customer prefer the rooms with high availability**

- **Customer prefer the rooms with low number of minimum nights for the room.**

- **If number of reviews is more, it means customers are prefering to book such rooms.**

Reduce minimum nights and increase availability in underperforming rooms of respective neighborhood
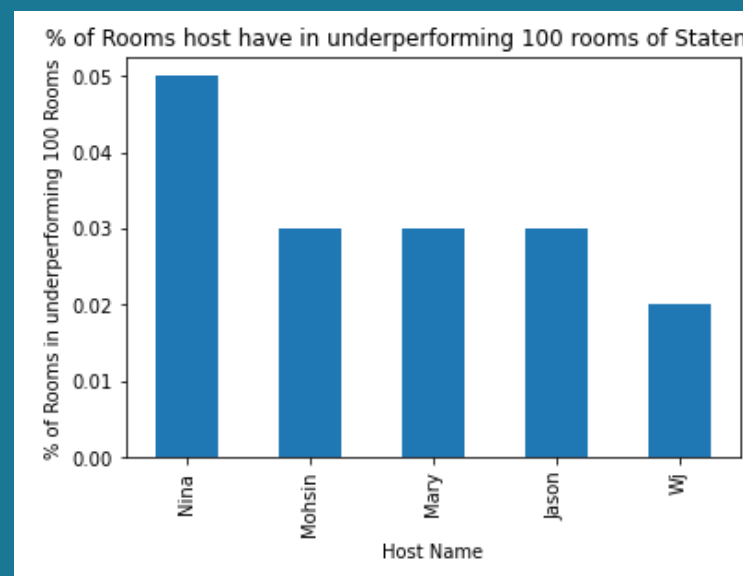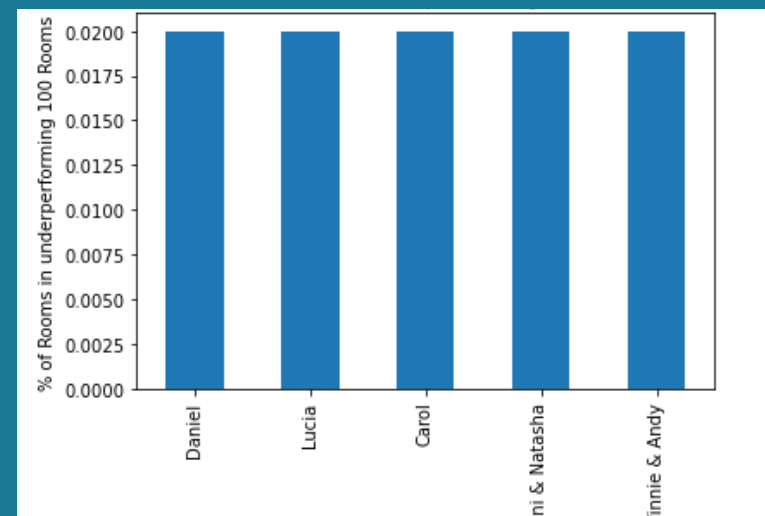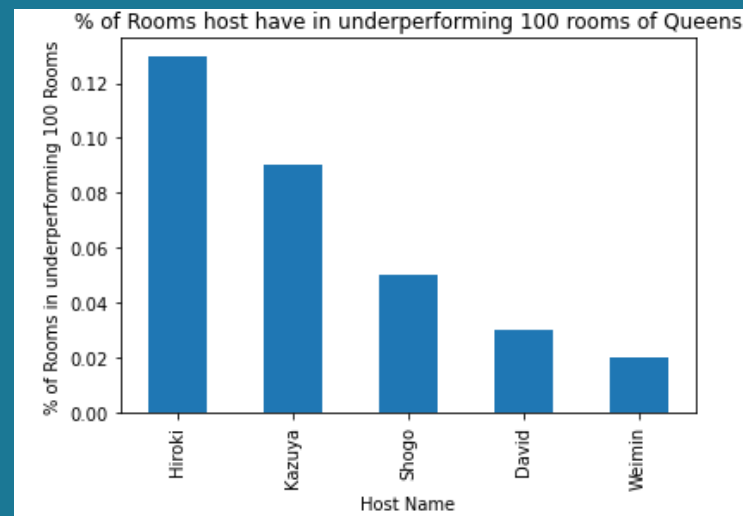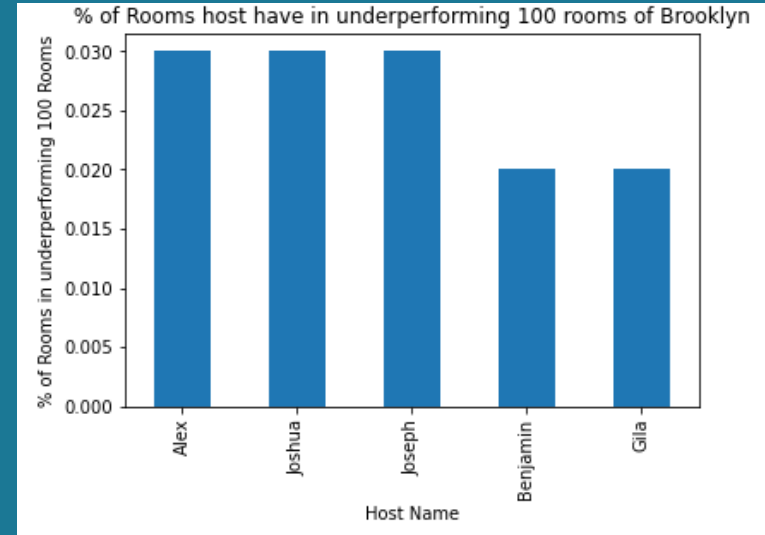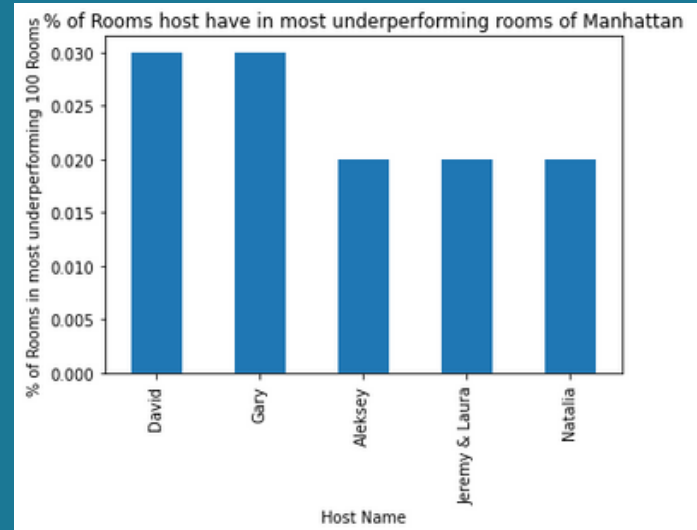
- This plot shows the percentage of rooms a host have in the 100 most underperforming rooms of each neighborhood.

- **David, Alex, Hiroki, Daniel and Nina** are the hosts of most number of underperforming rooms of **Manhattan, Brooklyn, Queens, Bronx and Staten Island** respectively.

- *Recommendation to Head of Acquisitions : We should speak to hosts of underperforming rooms to decrease the minimum nights of these rooms and increase availability.*

- *Recommendation to Head of UX : We should give Ratings to the hosts and hotels so that our customer can choose best rooms.*

Own More properties from the best hosts of the neighborhood.

% of Rooms host have in top 100 Rooms of Manhattan

% of Rooms host have in top 100 Rooms of Brooklyn

% of Rooms host have in top 100 Rooms of Queens

% of Rooms host have in the top Rooms of Bronx

% of Rooms host have in the top Rooms of Staten Island
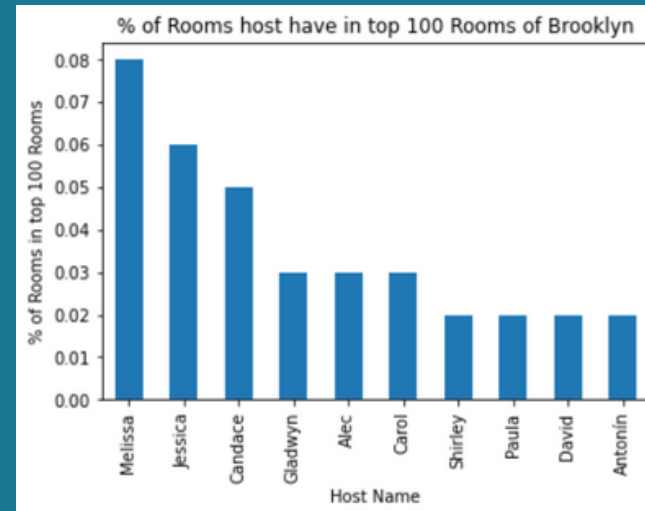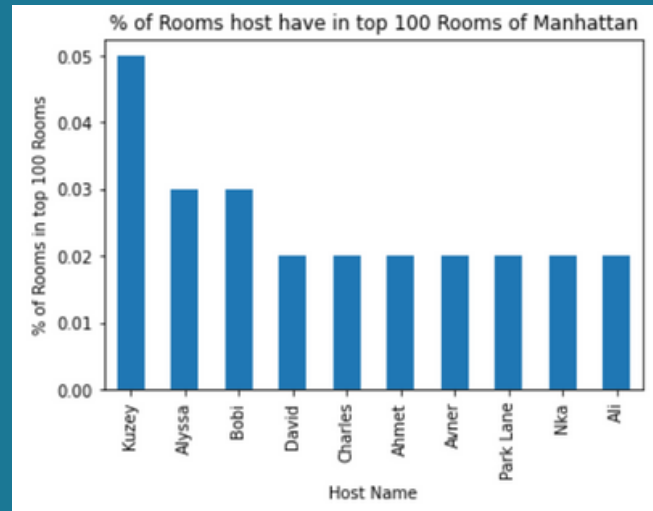
- This plot shows the percentage of rooms a host have in the best 100 rooms of each neighborhood.

- *Kuzey, Melissa, Veronica, Catherine and Genny* are the top hosts of *Manhattan, Brooklyn, Queens, Bronx and Staten Island* respectively as they have highest % of rooms in top 100 list.

- *Recommendation to Head of Acquisitions : Book more rooms from the best hosts of respective neighborhoods.*

- *Recommendation to Head of UX : We should arrange the rooms in the order of high availability, low minimum nights and high number of reviews in the website so that best rooms are visible on the top.*

Own More of Low and Medium priced private Rooms and Entire Apartments
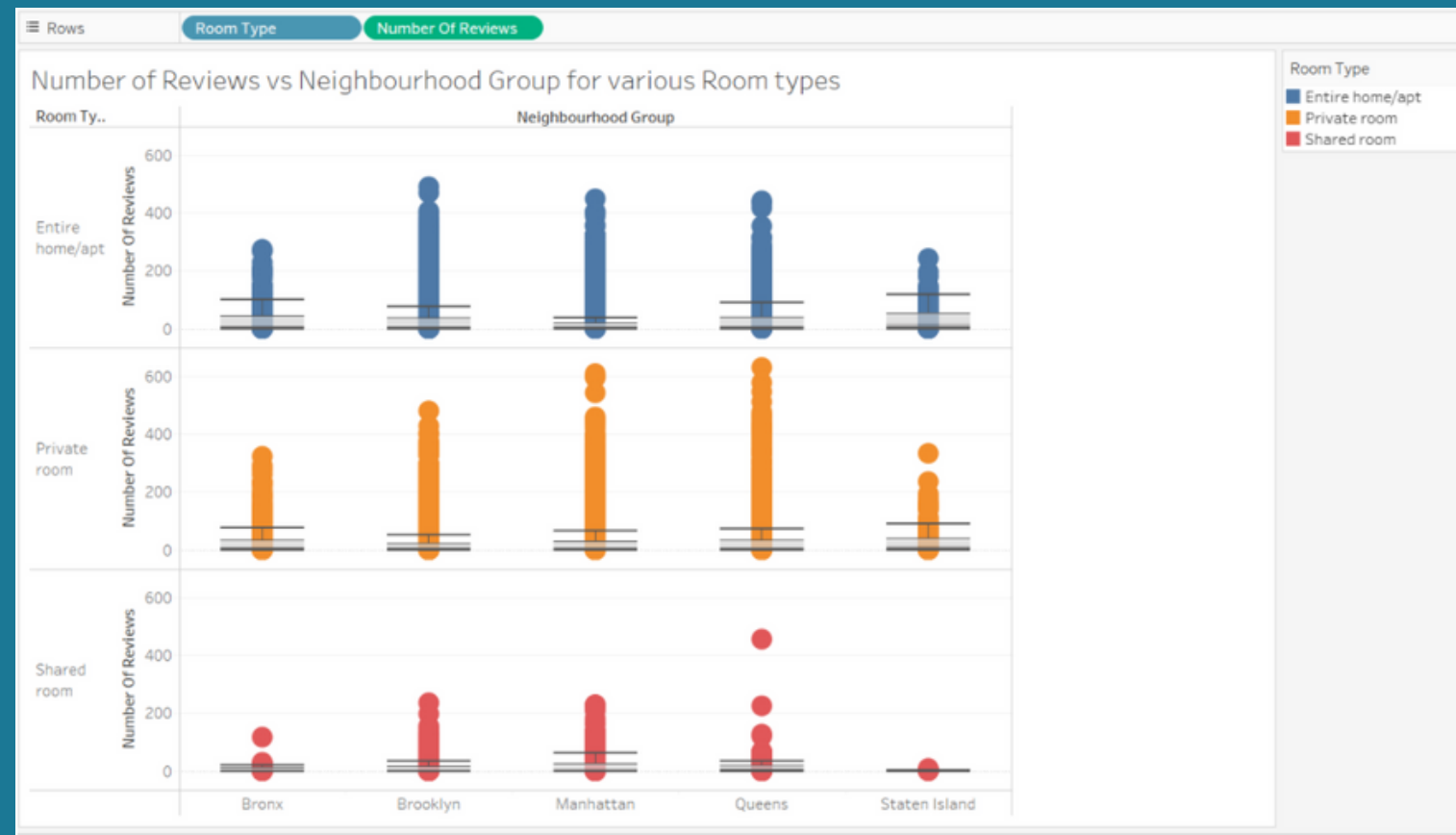
- We can observe that private rooms have the highest preference and entire houses have second highest preference shared rooms have the lowest preference.

- From the scatter plot chart we can see that most people prefer **low and medium priced private and entire houses** so we can look for such properties for ownership.
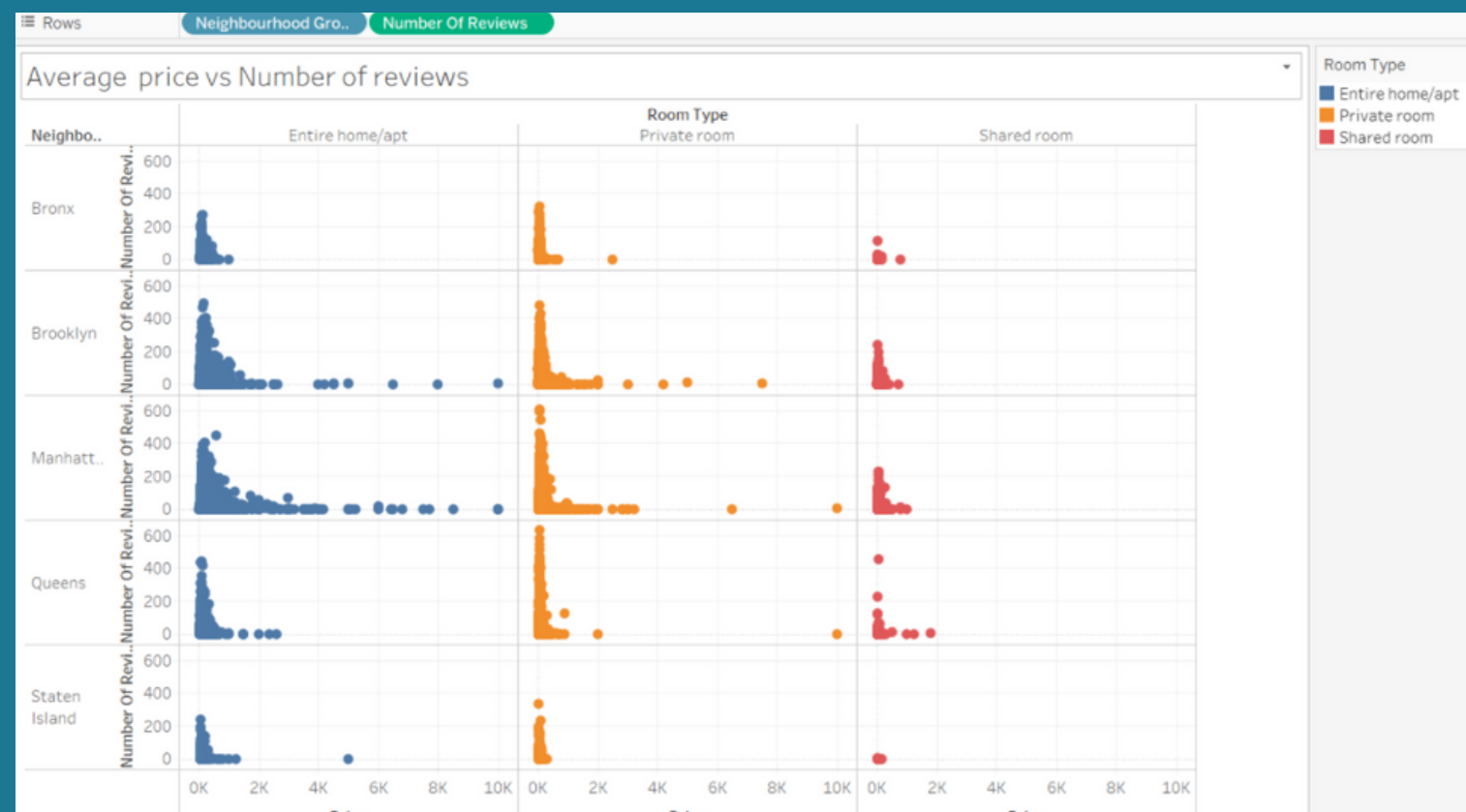
- *Recommendation to Head of Acquisitions : We can own more of low and medium priced private and entire apartment type rooms in all the neighborhood.*

- *Recommendation to Head of UX : Run more ad-campaigns for shared rooms as its full potential needs to be exploited.*

Recommendation to Head of Acquisitions

- We should speak to hosts of underperforming rooms to decrease the minimum nights of these rooms and increase availability.

- Book more rooms from the best host of respective neighbourhood.

- We can own more of low and medium priced private and entire apartment type rooms in all the neighborhood.

- we can speak to the hosts about the unoccupied shared rooms and provide them as private rooms for a small price bumps.

Recommendation to Head of UX

- We should give Ratings to the hosts and hotels so that our customer can choose best rooms.

- We should arrange the rooms in the order of high availability, low minimum nights and high number of reviews in the website so that best rooms are visible on the top.

- Run more ad-campaigns for shared rooms as its full potential needs to be exploited.

APPENDIX - DATA SOURCES:

SNIPPET OF DATA DICTIONARY

Column	Description
id	listing ID
name	name of the listing
host_id	host ID
host_name	name of the host
neighbourhood_group	location
neighbourhood	area
latitude	latitude coordinates
longitude	longitude coordinates
room_type	listing space type
price	
minimum_nights	amount of nights minimum
number_of_reviews	number of reviews
last_review	latest review
reviews_per_month	number of reviews per month
calculated_host_listings_count	amount of listing per host
availability_365	number of days when listing is available for booking

Dataset Description

This dataset contains information about different Airbnb listings along with their hosts, locations, prices and other attributes.

APPENDIX - DATA METHODOLOGY

For Methodology Document Click Here

(Internet is required to open this File)

Code To find underperforming hosts in Manhattan

```python
# underperforming Entire home in Manhattan
manhatten_under=df[((df.neighbourhood_group == 'Manhattan'))]
m1=manhatten_under.sort_values(by=['availability_365','minimum_nights','number_of_reviews'],ascending = [True, False, True])
manhatten_under=m1.head(100)
```

```python
# 5 most underperforming hosts
plt.xlabel('Host Name')
plt.ylabel('% of Rooms in most underperforming 100 Rooms')
plt.title('% of Rooms host have in most underperforming rooms of Manhattan')
manhatten_under['host_name'].value_counts(normalize=True).head().plot(kind='bar')
plt.show()
```

Code To find underperforming hosts in Manhattan

```python
# top 100 rooms in Manhattan
manhatten_best=df[((df.neighbourhood_group == 'Manhattan'))]
m1=manhatten_best.sort_values(by=['availability_365','minimum_nights','number_of_reviews'],ascending = [False, True, False])
manhatten=m1.head(100)
```

```python
# So this is the preference of Manhattan for price range
plt.xlabel('Host Name')
plt.ylabel('% of Rooms in top 100 Rooms')
plt.title('% of Rooms host have in top 100 Rooms of Manhattan')
manhattan['host_name'].value_counts(normalize=True).head(10).plot(kind='bar')
```

Code To find Relation between number of reviews vs price vs roomtype and neighbourhood are made on Tableau using room type as color tag.

APPENDIX - DATA ASSUMPTIONS

We have made a general assumption that more reviews on a listing translate as more traction/more number of bookings on that listing, the reasoning being that more reviews usually mean that more people have booked and lived in that property.