Methodology

Pre-case study

* Went through the data and understood the variables and relationships.
* Created a trail-layout of case study to figure out tools and methodology to be applied.
* Used excel pivots at initial level to play around the available dataset.
* Understood the range of values for continuous variables and no. of unique in categorical.

Python:

Imported necessary libraries and file.

Graphical user interface, text, application

Description automatically generated

Checked the import of data going through few initial rows.

Table

Description automatically generated

The dataset has below displayed rows and columns.

A picture containing table

Description automatically generated

Table

Description automatically generated

Checked the null presence in data which can possibly deviate the analysis.

Table

Description automatically generated

Dropped the variables ID (host\_id exists which gives unique count), Host name (not necessary for analysis) and last review (not necessary to find our observations in the case study).

Table

Description automatically generated with medium confidence

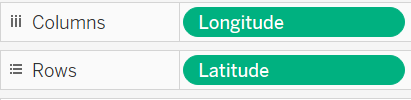
Replaced the null values with 0, meant no reviews received for the month.

A picture containing text

Description automatically generated

AIRBNB Upgrad PPT 1

1. Location Distribution:



* We took Longitude to Columns and Latitude to Rows we got scatter plot.
* Added Neighbourhood Groups in colours we got exact distribution of location
* We took Neighbourhood to Labels and Neighbourhood Groups to Details, whenever we hover across the plot we can see we will see Neighbourhood, Neighbourhood Groups, Longitude & latitude in details.

1. Top 10 Host By Reviews:

Graphical user interface, text, application

Description automatically generated

* We took Longitude to columns and Latitude to rows we got scatter plot
* We filter it by Host id where we took top 10 in terms of Reviews By Months In Counts.

1. Neighbourhood Group to Target:

Graphical user interface, text, application, chat or text message

Description automatically generated

* We took Neighbourhood Groups to columns and Number Of Reviews By Sum to Rows.
* We Got Bar Graphs we can tell how many reviews each Neighbourhood Groups gets.
* Added Neighbourhood Groups into colours so we can get different colours for different Neighbourhoods Groups.
* Added No of Reviews to Labels to get exact Sum of No of Reviews.

1. Neighbourhood Group vs Room Type vs Availability 365:

Graphical user interface, text, application, chat or text message

Description automatically generated

* We took Neighbourhood Groups & Room Type to columns and Availability 365 By Counts in Rows.
* We Sorted Neighbourhood Groups by Nested, sort order Descending, Field name Availability 365 & Aggregation in Count.

1. Customer Preference

Graphical user interface, text, application, chat or text message

Description automatically generated

* We took Room Type to columns and Price By CNT in Rows.
* Added Room Type to Colour & Price By CNT to Lable

1. Avg Price Vs Neighbourhood Group:

Graphical user interface, text, application, chat or text message

Description automatically generated

* We took Neighbourhood Groups to columns and Price By Avg in Rows.
* We Sorted Neighbourhood Groups by Nested, sort order Descending, Field name Price & Aggregation in Avg.
* To get avg price on bar we put Avg price to Labels.

1. Rooms Availability:

Graphical user interface, text, application, chat or text message

Description automatically generated

* We took Availability 365 to columns and Availability 365 By CNT in Rows.
* We fitler out 0 days to get exact Availability of Rooms.

1. Neighbourhood Group Vs Avg Availability 365

Graphical user interface, text, application, chat or text message

Description automatically generated

* We took Neighbourhood Group to columns and Availability 365 By Avg in Rows.
* Added Neighbourhood Group to Colours.

1. Rooms For Minimum Nights

Graphical user interface, text, application, chat or text message

Description automatically generated

* We took Minimum Nights to columns and Minimum Nights By CNT in Rows.

1. Neighborhood Group vs Avg Minimum Night Price

Graphical user interface, text, application, chat or text message

Description automatically generated

* We took Neighbourhood Groups to columns and Price By AVG in Rows.
* Added Neighbourhood Groups to Colours & Avg Price to Label.

1. Review For Neighbourhood Groups

Graphical user interface, text, application, chat or text message

Description automatically generated

* We took Neighbourhood Groups to columns and Number of Reviews By AVG in Rows.
* We filtered out 0 to get exact no of reviews.

1. Last Review Vs Price

Graphical user interface, text, application, chat or text message

Description automatically generated

* We took Last Review by YEAR to columns and Price By AVG in Rows.
* Added Price by AVG to Labels.

AIRBNB Upgrad PPT 2

1. Neighbourhood Group To Target

Graphical user interface, text, application, chat or text message

Description automatically generated

* We took Neighbourhood Group to columns and Number of Reviews by Sum in Rows.
* Added Neighbourhood Group to colours and Number of Reviews by Sum to Label.

1. Avg Price vs Neighbourhood Group

Text, application

Description automatically generated

* We took Neighbourhood Group to columns and Price by AVG in Rows.
* Added Neighbourhood Group to colours and Price by AVG to Label.

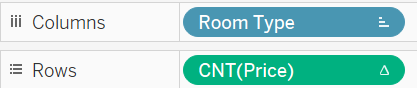
1. Neighborhood Group vs Room Type vs Avg Price

Graphical user interface, application

Description automatically generated

* We took Room Type & Neighbourhood Group to columns and Price by AVG in Rows.
* Added Neighbourhood Group to colours and Price by AVG to Label.

1. Customer Preference



* We took Room Type to columns and Price by CNT in Rows.
* Added Room Type to colours and Price by CNT to Label.

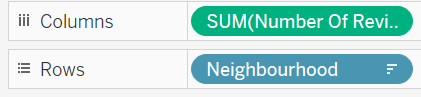
1. Avg Price/Minimum Nights vs Room Type

Graphical user interface, application, chat or text message

Description automatically generated

* We took Room Type & Measure Names to columns and Measure Values in Rows.
* Added Measure Names to Colours.
* Kept only Avg Price & Avg Minimum Nights in Measure Values.

1. Popular Neighbourhood As per Reviews/ Not Popular Neighbourhood As per Reviews



* We took Number of Reviews by Sum to columns and Neighbourhood in Rows.
* Added Neighbourhood Group to Colours to get which Neighbourhood belongs to which Neighbourhood Group.
* Top we got most popular neighbourhood & below we got Not popular Neighbourhood.

1. Popular Neighbourhood As per Availability

Graphical user interface, text, application, chat or text message

Description automatically generated

* We took Availability 365 in CNT to columns and Neighbourhood in Rows.
* Added Neighbourhood Group to Colours to get which Neighbourhood belongs to which Neighbourhood Group.

1. Host with Most Number of Reviews across all Neighbourhood Groups

Graphical user interface, application, chat or text message

Description automatically generated

* We took Host Name & Neighbourhood Group to columns and Number of Reviews by CNT in Rows.
* Added Neighbourhood Group to Colours to get which Neighbourhood Group Host operates.

1. Top Host in Manhattan, Brooklyn, Bronx, Queens & Staten Island

Graphical user interface, application, chat or text message

Description automatically generated

Graphical user interface, text, application

Description automatically generated

* We took Host Name & Neighbourhood Group to columns and Number of Reviews by CNT in Rows.
* To get top Host in all Neighbourhood Group we added neighbourhood Group to filter, and we filter it as per neighbourhood Group that we want to see.

Key observations :

* Brooklyn & Manhattan are need to target as per number of reviews.
* Queens, Bronx and Staten Island received a smaller number of reviews.
* Bronx Avg. price is not even half of Manhattan, per night
* Staten island provides room services exactly double the avg. days which Brooklyn provides
* The Manhattan’s average price for all room types is higher than the overall average, whereas other neighbourhood groups are lesser than overall average.
* Brooklyn provides the cheaper shared rooms and comparatively less average prize for entire home/apt.
* Staten island provides the cheaper private rooms.
* Entire Home/Apt has the highest Avg Price and Avg Min. Nights among other room types (211/8.5)