
Detailed Project Report (DPR)

Airbnb Data Analysis

Document Version Control

Date Issue	Version	Description	Author
09/03/2023	1	Initial DPR - V 1.0	Abdul Jaweed
25/06/2023	2	Initial DPR - V 2.0	Abdul Jaweed Aalmeen Khan Safa Ashraf

Objective

The objective of the project is to build an dashboard that will help to understand the airbnb data and its pattern.

Architecture

RAW Dataset

Exploratory
Data Analysis

Data
Preprocessing

Dashboard

About the Project

Airbnb is an American company that facilitates an online marketplace for lodging, primarily homestays for vacation rentals, and tourism activities. It basically connecting travelers with local hosts who want to rent out their homes with people who are looking for accommodations in that locality. On the other hand, this platform enables host to list their available space and earn extra income in the form of rent and it also enables travelers to book unique homestays from local hosts, saving them money and giving them a chance to interact with locals.

Dataset Information

The dataset consists of 18,724 rows and 20 columns namely- room_id, survey_id, host_id, room_type, country, city, borough, neighborhood, reviews, overall_satisfaction, accommodates, bedrooms, bathrooms, price, minstay, name, last_modified, latitude, longitude and location.

Attribute Information

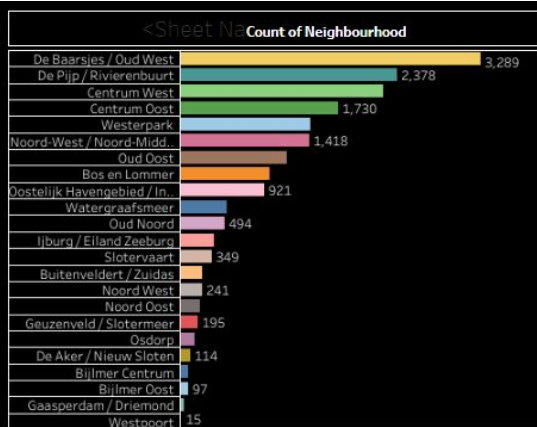
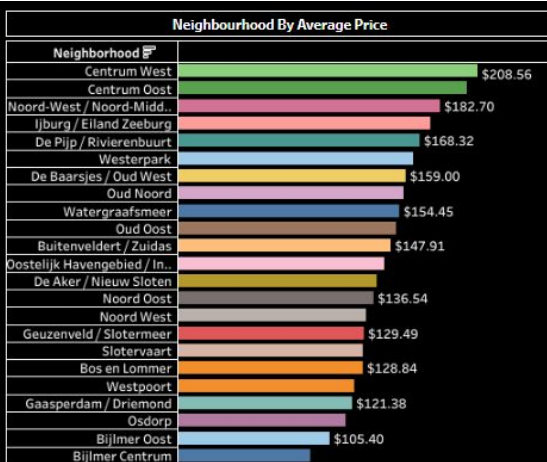
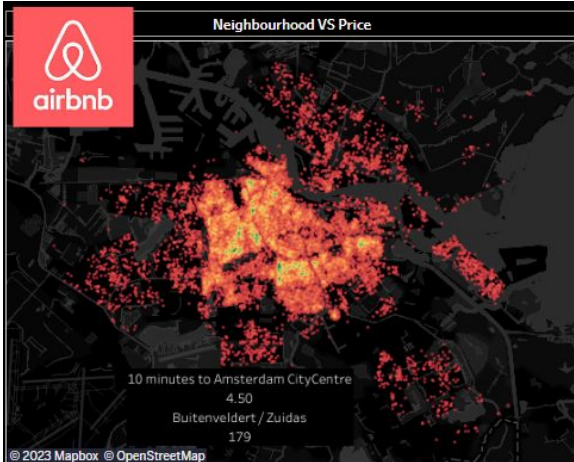
- Room ID: There are 18,724 unique room IDs in the dataset.
 - Survey ID: Only one value is provided for the survey ID, so it will not be used in our analysis.
 - Host IDs: There are a lot of duplicate values.
 - Room Type: There are three room types: shared room, private room, and entire apartment.
 - Country: The country is not mentioned in the provided information.
 - City: Only one city, Amsterdam, is given.
 - Borough: No values are given for the borough attribute.
-

Attribute Information

- Neighbourhood: The available neighbourhoods are De Baarsjes/Oud West, De Pijp/Rivierenbuurt, Centrum West, Centrum Oost, Westerpark, Noord-West/Noord-Midd., Oud Oost, Bos en Lommer, Oostelijk Havengebied/In..., Watergraafsmere Oud Noord, IJburg/Eiland Zeeburg, Slotervaart, Buitenveldert/Zuidas Noord West, Noord Oost, Geuzenveld/Slotermeer, Osdorp, De Aker/Nieuw Sloten, Bijlmer Centrum, Bijlmer Oost, Gaasperdam/Driemond, and Westpoort.
-

Attribute Information

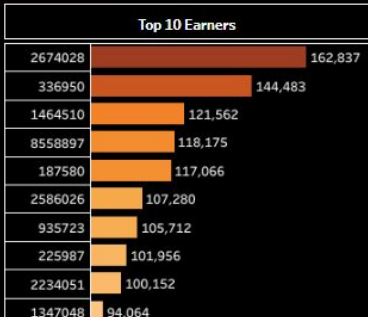
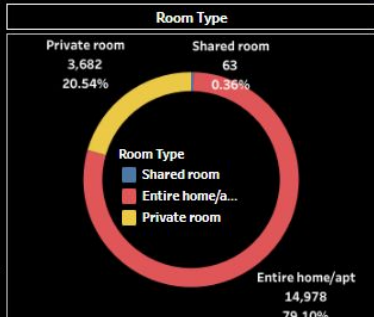
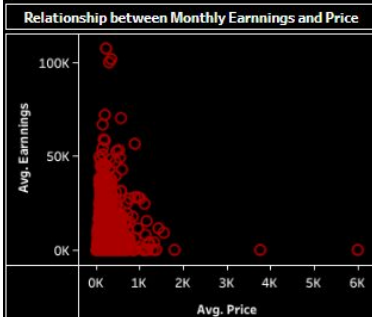
- Reviews: The review ratings range from 0 to 532.
 - Overall Satisfaction: The overall satisfaction ratings range from 0 to 5.
 - Accommodates: The number of people the rooms can accommodate ranges from 1 to 17.
 - Bedrooms: The number of bedrooms in the rooms ranges from 0 to 10.
 - Bathrooms: No values are given for the bathroom attribute, so it is not considered for analysis.
 - Price: The price range for the rooms is from 12 to 6000 euros.
-



Total Hosts
15,943

Total no. Neighbourhood
23

- Neighbourhood**
- Bijlmer Centrum
 - Bijlmer Oost
 - Bos en Lommer
 - Buitenveldert / Zu...
 - Centrum West
 - Centrum Oost
 - De Aker / Nieuw S...
 - De Baarsjes / Oud ...
 - De Pijp / Rivieren...
 - Gaasperdam / Dri...
 - Geuzenveld / Slot...
 - IJburg / Eiland Zee...
 - Noord Oost
 - Noord West
 - Noord-West / Noo...
 - Oostelijk Haveng...
 - Osdorp
 - Oud Noord
 - Oud Oost
 - Slotervaart
 - Watergraafsmeer
- Room Type**
- (All)
 - Entire home/apt
 - Private room
 - Shared room



Q & A

Q1) What is the size of your data ?

The size of the data in terms of KB is 38.0, with 18723 rows and 20 columns names are encrypted for security reasons

Q2) What are the data type ?

The columns consisted of both integer and float values.

Q & A

Q3) What were the libraries that you used Python ?

I used Pandas, Matplotlib, Seaborn, Plotly and Streamlit.

Q4) What's the complete flow you followed in this Project ?

We had a full-fledged data pipeline which would extract the data from the S3 bucket, which had come from Big Data pipeline created by the client side.

Q & A

Q5) Where did you get the data ?

The data was provided by iNeuron.

<https://drive.google.com/drive/folders/165Pjmf9W9PGy0rZjHEA22LW0Lt3Y-Q8>

THANK YOU
