AirBNB Stay Analysis

[1]Dr. Siddique Ibrahim S P
Assistant Professor,
School of Computer Science and Engineering,
VIT-AP University,
Amaravati, Andhra Pradesh, India
siddique.ibrahim@vitap.ac.in

[2]Anisha Kumar
School of Computer Science and Engineering,
VIT-AP University,
Amaravati, Andhra Pradesh, India
anisha.21bce7239@vitapstudent.ac.in

[3]Jai Gaurav
School of Computer Science and Engineering,
VIT-AP University,
Amaravati, Andhra Pradesh, India
gaurav.21bce7193@vitapstudent.ac.in

[4]Shreya Shamej
School of Computer Science and Engineering,
VIT-AP University,
Amaravati, Andhra Pradesh, India
shreya.21bce8432@vitapstudent.ac.in

[5]Sanal P K
School of Computer Science and Engineering,
VIT-AP University,
Amaravati, Andhra Pradesh, India
sanal.21bce8592@vitapstudent.ac.in

[6]Dayita Halder
School of Computer Science and Engineering,
VIT-AP University,
Amaravati, Andhra Pradesh, India
dayita.21bce8042@vitapstudent.ac.in

Abstract - This market analysis focuses on Airbnb, a prominent online hospitality marketplace that has revolutionized the way people find accommodation and lodging. Airbnb's unique platform connects hosts offering their properties and guests seeking short-term rentals, creating a vast global network of diverse and unique accommodations. The analysis begins by providing an overview of Airbnb's history, growth trajectory, and its impact on the traditional hospitality sector. Utilizing various market research methodologies and data sources, the study examines Airbnb's market share, geographical expansion, and user demographics. It evaluates the factors that attract hosts to list their properties on the platform, as well as the motivations that drive guests to choose Airbnb over traditional hotels. Ultimately, this market analysis aims to provide valuable insights into Airbnb's position within the hospitality industry, identifying its key success factors and potential areas for improvement. It serves as a valuable resource for industry stakeholders, investors, and policymakers seeking to understand the dynamics of the sharing economy and its impact on the traditional lodging sector. As Airbnb continues to redefine the travel and accommodation landscape, an in-depth analysis of its market presence proves crucial for making informed decisions and predicting its future trajectory.

Keywords— Airbnb's market share analysis, future trajectory.

I. INTRODUCTION

AirBNB is a online marketplace that connects people who want to rent out their homes with people who are looking for accommodations, people rent out their homes, apartments, or rooms to travelers on a short-term basis ,and even unique accommodation like tree house. It enables hosts to list their properties, and travelers can search for accommodations in various locations around the world. And here we are going to find out the most convenient preference by using filters like geographical area , price , rating etc

And this filters play a crucial role in ensuring a positive and efficient user experience on Airbnb, leading to increased bookings, guest satisfaction, and long-term success for both guests and hosts on the platform.
The importance of filtering in our project Filtering preferences in Airbnb offers several advantages to both hosts and guests:

Customization: Guests can easily find accommodations that meet their specific preferences and needs, such as price range, location, amenities, property type, and more. This allows them to tailor their search and find the perfect place to stay.
Time-saving: With filtering options, guests can quickly narrow down their choices, reducing the time spent searching for suitable accommodations.
In Airbnb's filtering process, data mining and data warehousing play critical roles in analyzing and processing large volumes of data to improve the accuracy and relevance of search results for users. Here's how these technologies are used:

Data mining and data warehousing are integral to Airbnb's filtering process, ensuring a seamless and personalized user experience by refining search results, detecting fraud, optimizing prices, and making data-driven business decisions.

II. LITERATURE REVIEW

Personalization: Filtering options allow guests to tailor their search to specific preferences, ensuring they find accommodations that align with their unique needs and preferences. This personalized experience enhances guest satisfaction and increases the likelihood of bookings.

Efficient Search: With numerous listings available on Airbnb, filtering streamlines the search process. Guests can quickly narrow down their options based on criteria like location, price, amenities, and property type, saving time and effort.

Enhanced User Experience: By providing relevant and customized results, filtering enhances the overall user experience on the platform. Guests are more likely to have a positive impression of Airbnb when they can easily find accommodations that meet their expectations

Better Matching: Filtering enables better matching between guests and hosts, as guests can find accommodations that match their desired criteria, and hosts can attract guests who are genuinely interested in what they offer.

Targeted Marketing: Airbnb can use the data from filtering preferences to understand user preferences and behavior, allowing them to provide targeted recommendations and marketing efforts, further enhancing user engagement.

Long-Term Success: By offering a personalized and efficient experience, filtering contributes to guest loyalty and retention. Satisfied guests are more likely to return to Airbnb for future bookings and recommend the platform to others.

Data mining and data warehousing role in our project:-
Data Storage and Integration: Airbnb collects vast amounts of data from multiple sources, including user profiles, property listings, booking transactions, reviews, and more. Data warehousing allows all this data to be stored in a single, integrated repository, making it easier to access and analyze.

Performance and Scalability: Data warehousing systems are designed for high-performance data processing, allowing Airbnb to handle large datasets and complex queries efficiently. As the platform grows, data warehousing ensures scalability without compromising performance.

Real-time Data Availability: Data warehousing enables Airbnb to have up-to-date and real-time access to data. This timely information is essential for delivering accurate and relevant search results to users based on their current preferences and demands.

Data Cleansing and Quality: With data warehousing, Airbnb can implement data cleansing and quality assurance processes. This helps to ensure that the data used for filtering is accurate, reliable, and free from errors or inconsistencies.

Historical Analysis: Data warehousing maintains historical data over time, allowing Airbnb to conduct in-depth analysis of past trends, user behaviors, and market dynamics. This historical analysis helps in understanding patterns and making informed decisions for future improvements.

Improved Filtering Algorithms: Data warehousing supports the development of complex and efficient filtering algorithms. By accessing integrated data, Airbnb can fine-tune algorithms to provide users with more accurate and personalized search results

We referred the following bibliographies
https://airbnbla.humspace.ucla.edu/bibliography/
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8561316/
https://www.blinkist.com/en/books/the-airbnb-story-en
https://www.airbnb.co.in/resources/hosting-homes

As of now output of user view is very sophisticated which can be integrated into a web application in the future.

III. PROPOSED METHOD

The main idea of this project is to develop a filtering system and analyze the stays available in New York City, for the AirBNB dataset that we have obtained from Kaggle.

Given below is the Google Colab link for the project :
https://colab.research.google.com/drive/10Vk3JkbzXKRa7ffKdbyoEkd7PLqiijXm?usp=sharing
In this particular section of the report, we shall be discussing both these aspects one after the other. Before that, we shall discuss about the python modules used for the project.

A. Modules & Softwares used:
a) **Pandas** library is a data analysis and manipulation tool available in Python. We used it for both our filter system and EDA.
b) **Numpy** library is a python library primarily made for working with arrays. This was used for EDA.
c) **Matplotlib** library is a library used for plotting graphs and this was also used for EDA.
d) Additionally, **PowerBI** was used for EDA.

B. Filter System
In this model, we have developed a 4-stage / 4-layer filter for the dataset that we have. The filters modify the dataset on these 4 grounds: Location, Room Type, Price & Number of Reviews.

The first filter is a *location filter* that reduces the existing dataset into a smaller one on the basis of the location the user wishes to reside at for their vacation. The filtering is done on the basis of the *'neighbourhood_group'* attribute. The user can simply enter the location that they wish to reside in and the dataset gets filtered accordingly.

eg : If we search neighborhood group as Manhattan we will find all the listings of Manhattan.

The second filter filters the dataset further on the ground of what *room type* the user wishes to stay in. Once the user enters the required room type, they can view all such rooms available in the selected location.

eg : Here we have searched the room_type as Private room, so we find all the corresponding listings under this category.

The third filter reduces the dataset on the basis of the *budget* of the user. This is the *price filter* that generates a dataset according to the upper and lower budget bounds of the user. In this way, the user can obtain a much more sophisticated view of the options available to them according to their preferences of stay

eg : Here we have searched for the listings having price ranging from 70-100.

Last but not the least, we have a filter that reorders the dataset in a descending order on the basis of the number of reviews a stay has got. This is the *review filter*. The reason why we have taken the number of reviews as a ground of categorization is because higher the number of reviews, the better the stay is likely to be. The user definitely wants to have the best options available for them at the top of their screen and hence we decided the sorting to be in a descending fashion, thus generating the best possible filtering system.

eg : To sort an existing Airbnb dataset by the number of views, we extracted the "number of views" column from the dataset, sorted it based on the "number of views" column in descending order. The sorted dataset will now display listings with the highest number of views at the top, allowing you to identify the most viewed properties on AirBNB.

The finally filtered dataset contains information about location (city or neighborhood), lodging type (e.g., entire home, private room), price, number of reviews and a brief description of each listing (name of listing). This comprehensive dataset is ideal for conducting detailed analysis, visualizations. This data is efficiently stored in a pandas data frame object that can easily be adapted to multiple output streams such as a desktop app or a web application. In it's current implementation this information is made visible on the command prompt in a tabular format with re-sizeable columns and variable row counts per page, dynamically wrapping the text data to ensure a clean and comprehensible output.

Given below is the flowchart representation of the entire filter system.
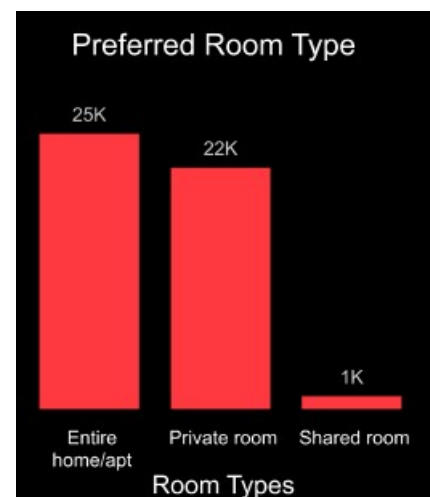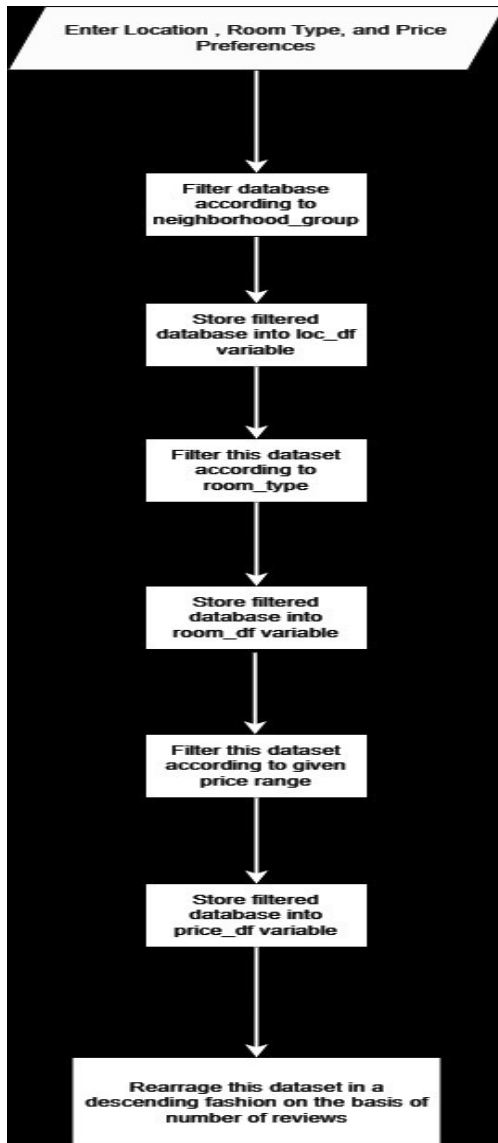
C. Stay Analysis (Exploratory Data Analysis)

For analyzing price trends, room types and availability at different locations, etc. we have performed exploratory analysis on our dataset. Using PowerBI and Python, we have generated a few of the given plots for the same purpose. Note that there might be some repetitive plots, however they are there for visualizing the same point in different ways and through different perspectives in order to make utmost use of all available tools and also to learn more about the platforms we have used in this research paper.
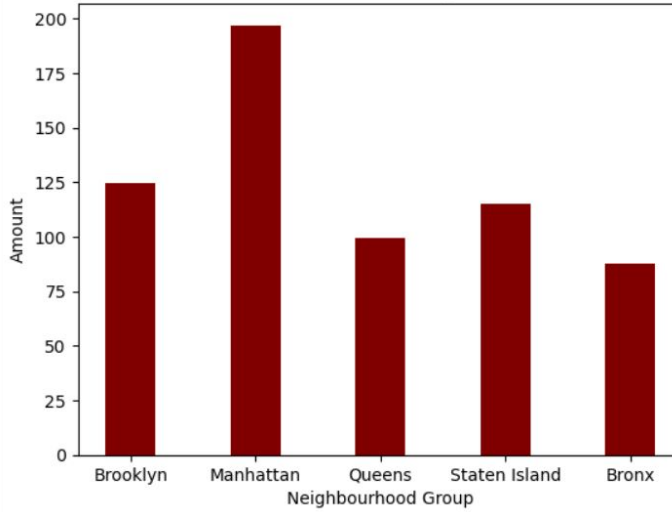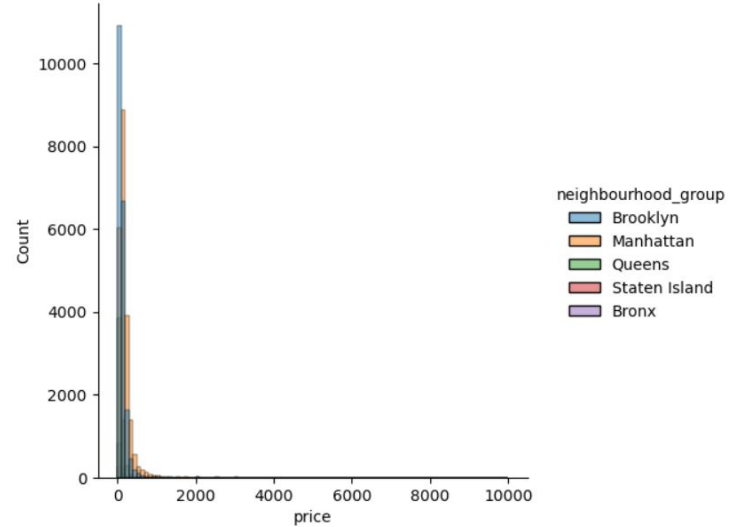
Distribution of Average Availability

- 2.00 (0.28%)
- 100.13 (13.84%)
- 198.65 (27.47%)
- 111.85 (15.47%)
- 165.41 (22.87%)
- 144.22 (19.94%)
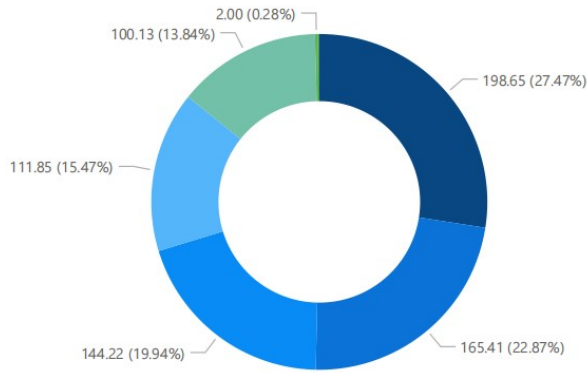
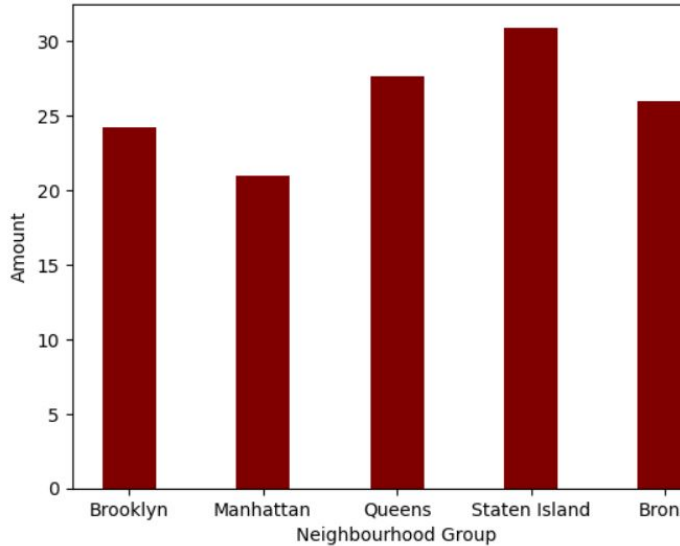## Distribution of Average Host Listings Count
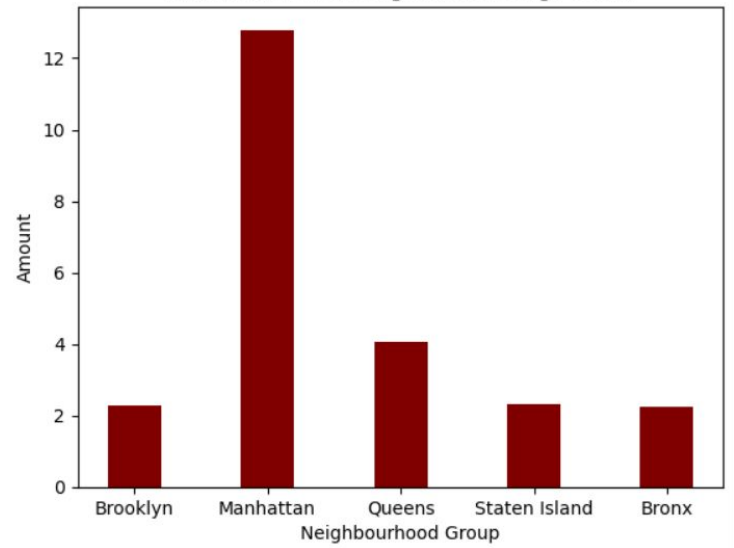
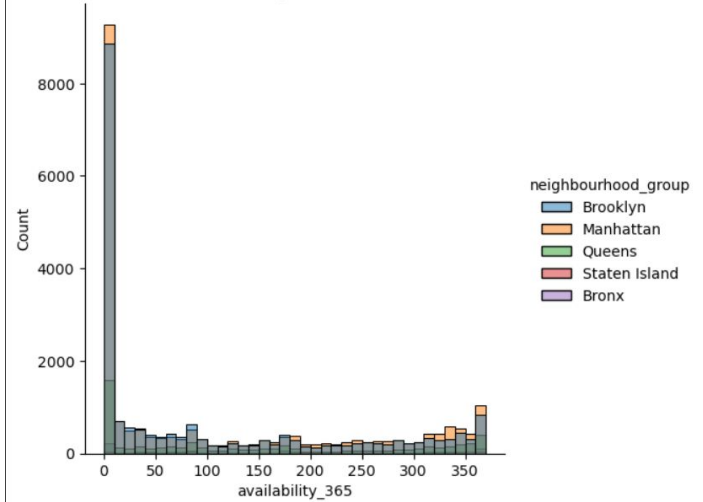## Distribution of Average Price

Price Distribution

Distribution of Average Number of Reviews

Availability Distribution

IV. RESULT AND DISCUSSION

In an era defined by shared economies and innovative travel experiences, lodging and homestay rentals has revolutionized the way people explore new destinations and find unique accommodations. Lodging and homestay rentals has sparked a wealth of data, offering an opportunity to delve deep into the world of hospitality trends, guest preferences, and property dynamics. In our Airbnb stay analysis, we sought to build a predictive model that could accurately forecast guest satisfaction based on various listing attributes.

The review data within the filtered dataset holds a pivotal role in uncovering the true essence of each lodging experience. It offers a tangible measure of the lodgings' popularity and the level of guest engagement. By quantifying the number of reviews for each listing, we gain insights into the lodgings' appeal and the satisfaction level of past guests. Taking into consideration the assumption that a higher number of reviews per stay corresponds to a more superior lodging experience compared to other listings, our model has demonstrated remarkable accuracy in providing tailored user recommendations.

We currently have a sophisticated user interface that, when implemented on a web application, could significantly enhance the user experience by making it much easier to navigate and interact with the system. When applying Exploratory Data Analysis (EDA) specifically to the review filter, this well-structured format proves particularly beneficial. It enables analysts to interact with the review-related insights effectively, allowing for comprehensive exploration and extraction of patterns, trends, and potential outliers within the review data.

V. CONCLUSION AND FUTURE SCOPE

A. Conclusion
In this paper, we analysed data on AirBNB stays in New York City in the year 2019 to find the popular areas and booked places which is based on their reviews which is reviewed by the people who have the overall experience about the stay and place. We used a variety of visualization tools like PowerBI, Jupyter Notebook, and Python etc., which help us understand the data and to illustrate our findings. By using these tools, we found the popular areas for AirBNB stays in New York City are Manhattan, Brooklyn, and Queens. These areas attract more tourists, and they offer a wide range of AirBNB listings, from private rooms to entire apartments. To get more information from the dataset, we created a map of New York City that shows the locations of AirBNB listings. We also created a bar chart that shows the distribution of average review scores, average host listings count, average price, average minimum nights, and preferred room type for AirBNB listings in New York City. And created univariates like average price in USD, minimum nights, and total availability.

B. Future Scope
In the field of future thinking, interesting ways are emerging that extend and support the framework of AirBNB's filtering and analytics services. These deals have the potential to change and improve the user experience, leading to a new era of optimization and personalization in travel and accommodation.

- Interactive map visualization and graphs takes centre stage by making visual comparisons that guide users in understanding location in relation to geography, topics and trends, transportation. The dynamic map helps to improve the assessment of the potential destinations region, giving travellers a clear view to evaluate their options
- Visual improvements and innovations can be implemented as features on the integrated AirBNB website or mobile app. By seamlessly integrating these capabilities into its existing platform, AirBNB can create a unified and user-friendly experience that meets the diverse needs and preferences of travellers.

VI. REFERENCES

1. Manuel J. Sánchez-Franco, & Maria Elena Aramendia-Muneta. (2023). Why do guests stay at Airbnb versus hotels? An empirical analysis of necessary and sufficient conditions. https://www.sciencedirect.com/science/article/pii/S2444569X23000768#:~:text=Both%20types%20of%20accommodation%20share%20similar%20themes%20that%20contribute%20to,interactions%20between%20staff%20and%20guests.
2. Valentin Toader & 3 others. (2021). Analysis of price determinants in the case of AirBNB listings. https://www.tandfonline.com/doi/full/10.1080/1331677X.2021.1962380
3. Czeslaw Adamaik. (2019). Current state and development of Airbnb accommodation offer in 167 countries. https://www.tandfonline.com/doi/full/10.1080/13683500.2019.1696758
4. M. Bharati. (2010). Data mining techniques and applications. https://www.researchgate.net/publication/49616224_Data_mining_techniques_and_applications
5. G. Sivabharathi & Dr. K. Chitra. (2020). Data mining techniques to perform market analysis. https://www.researchgate.net/publication/354877317_DATA_MINING_TECHNIQUES_TO_PERFORM_MARKET_ANALYSIS