

Data Visualization On Airbnb Dataset Using Tableau

Syed Ibrahim Hussain¹, Mahad Ansari², Mr.Kareem Basha³

^{1,2}UG Scholar, Department of AI&DS, Methodist College Of Engineering and Technology, Hyderabad, India

³Assistant Professor of CSE, Methodist College Of Engineering and Technology, Hyderabad, India

Abstract- Today, the brand new digital economy has revolutionized the way people travel and find a place to stay through platforms like Airbnb. For this purpose, this project is dedicated to delving into Airbnb data with the help of efficient data visualization methods to reveal significant patterns and insights. The study, by analyzing various factors such as pricing location types of rooms, availability, and customer reviews, is aiming at finding the answer to how different variables affect listing performance and user preferences. Through the use of visualization software, not only are complex datasets opened up in a simple and interactive visual manner like charts, graphs, and maps but it also becomes much easier to recognize the patterns and associations. Besides hosts, guests, and platform developers, the project also showcases the great potential of data visualization in enabling them to make better decisions. In short, this work illustrates the tremendous impact of visual storytelling in turning huge datasets into simpler ones and at the same time, providing useful insights in a real-life situation!!!

Keywords- Airbnb Analytics, Data Visualization, Sharing Economy, Hospitality Industry, Exploratory Data Analysis (EDA), Customer Reviews, Pricing Analysis, Geographic Visualization, Business Intelligence, Data-Driven Decision Making, Tourism Analytics, Interactive Dashboards.

I. INTRODUCTION

Amidst the rise of the internet as a platform of commerce, the internet also allows for new markets and marketplace systems to emerge because of the ease with which it propagates information. One such service is, for instance Airbnb which is an online marketplace where travelers may book rooms or entire houses for the duration of their trip.

As the number of users grows, so will the available data which can be exploited to find trends. Using Data Visualisation One way of understanding these vast amounts of raw data is through data visualisation.

Utilising visualisation techniques, raw data can be represented in a meaningful way that is not difficult to interpret. Graphs, charts and maps all provide a means by which we may be able to easily discern trends, patterns or interrelationships.

This project is to explore data given by Airbnb listing by using various visualization methods to make known a number of critical issues such as price location room type and availability. It is hoped that the project will conduce to the choice for both hosts and custom.

II. LITERATURE SURVEY

Many research works have been made to examine Airbnb data by employing data visualization and analysis methods. These researches mostly target the understanding of users' behavior, determining pricing trends, spotting location patterns, and studying the performance of the platform overall. In a paper by Jain et al. (2024), the authors discussed how data visualization assists in tracing user journeys on Airbnb. They found that studying users' behavior, like looking at listings, reading reviews, making booking decisions, etc. can cause enhanced customer satisfaction and operational efficiency of the platform. They also got to know that users count more on properties with higher ratings and detailed reviews, which is a clear testament to the role of data insights in decision-making.

One more paper that examined the New York City Airbnb listings data from 2011 to 2023 used some visualization methods, including heat maps, graphs, and word clouds, to uncover the patterns of pricing, types of rooms, and distribution of the neighbour hoods. This research uncovered that location and kind of room are the main factors that determine price, and that visualization tools present the data in a very user-friendly way.

Based on a deep research article review of more than 1000 papers, it was found that most of Airbnb studies are geared towards development of pricing tools, building customer trust, and the creation of values. Besides that, the review also touches on the increasing application of data analysis and visualization techniques for mining large datasets and guiding decision-making. Besides what I just said, scholars have used techniques like sentiment analysis and natural language processing on Airbnb reviews. They have documented that how positive reviews and ratings are why for bookings while characteristics like the location amenities, and the host's talkiveness go a long way in determining the guest's final satisfaction level. Besides this, the academics employed geospatial studies to find out the spread of Airbnb accommodations in different cities. Their results depicted that many the listings are located in renowned tourist areas and downtowns, which is a clear indication of how much location affects demand and pricing.

In short, the surveyed publications show that data visualization forms an essential element for making sense of Airbnb's data base. It is a suitable means for spotting tendencies, grasping consumers' tastes, and encouraging more informed decision-making. Even so, several papers do highlight problems like data integrity problems in the publicly available datasets, which call for thorough data validation.

III. MATERIALS AND METHODS

The Dataset we used, which is based on New York City. It covers data about hosts pricing location, room types, as well as customer reviews.

This dataset is one of the most popular datasets for carrying out analysis about pricing trends, demand patterns and host performance in the short-term rental market.

This dataset has a number of different features that will help in analysis of factors such as booking behavior, geographical distribution, and economic factors related to Airbnb listings.

Feature	Data Type	Description
id	Integer	Unique identifier for each listing
host_id	Integer	Unique ID of the host
host_name	String	Name of the host
Neighbourhood_group	Categorical	Area (Manhattan, Brooklyn, etc.)
neighbourhood	String	Specific locality

latitude	Float	Latitude coordinate
longitude	Float	Longitude coordinate
room_type	Categorical	Entire home, private room, shared room
price	Float	Price per night
minimum_nights	Integer	Minimum stay required
number_of_reviews	Integer	Total number of reviews
reviews_per_month	Float	Average monthly reviews
availability_365	Integer	Number of available days per year

Software Used: Tableau

Tableau is a very useful and comprehensive data visualization and business intelligence tool which is used to analyze and visualize data in an interactive and easy-to-explore way. It allows the users to make dashboards, charts and maps without any programming skills.

In this project, Tableau was the tool of choice to mine Airbnb data and to come up with interesting findings through many visualization techniques such as bar charts maps heat maps and trend analysis. Thanks to Tableau's drag-and-drop interface, data analysis was very quick and it was easy to notice different patterns related to prices, bookings and geographical distribution.

Some of Tableau's distinguishing factors are:

- Interactive dashboards
- Real-time data visualization
- Geographic mapping
- Easy data connectivity

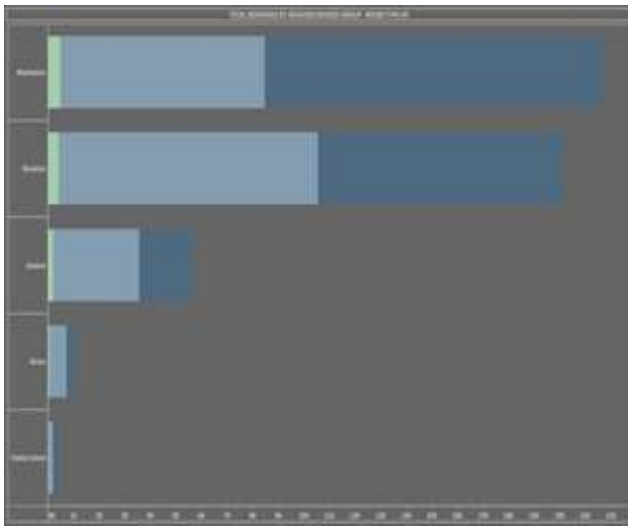
IV. DATA VISUALIZATION

A full suite of data visualization tools was developed with Tableau to study the Airbnb NYC dataset.

These visualizations uncover various elements like pricing, temporal changes in reservations, the spatial distribution of Airbnb listings, and the performance of hosts. Every visualization narrates a story by focusing on a particular part of the Airbnb platform and aid the readers to identify factors influencing the market.

Total Bookings by Neighbourhood Group Here, the total number of bookings per each NYC borough is displayed by a bar chart.

Manhattan together with Brooklyn have very high booking figures from their widespread popularity among tourists. Still, Staten Island and Bronx are Much lesser for bookings which may be due to either reduced demand or a limited number of listings.



[Figure 1: Total Bookings by Neighbourhood Group — Bar Chart]

Average Reviews per Month by Room Type and Neighbourhood Group



[Figure 2: Average Reviews per Month — Heatmap]

The heatmap above displays average monthly review numbers for various room types (Whole home/apartment, Private room, Shared room) in different neighbourhood groups. Entire homes and private rooms in Manhattan and Brooklyn are more reviewed which suggests that these areas have high demand and occupancy. Shared rooms generally have less reviews in all areas.

Top 10 Hosts by Total Reviews The horizontal bar chart above illustrates the leading 10 hosts with total reviews received. Some hosts quite literally outclass the others with review numbers, suggesting that they may have multiple listings or provide better services leading to higher customer satisfaction. More so, it points out Really a handful of very active hosts are dominating the platform.

The horizontal bar chart above illustrates the leading 10 hosts with total reviews received. Some hosts quite literally outclass the others with review numbers, suggesting that they may have multiple listings or provide better services leading to higher customer satisfaction. More so, it points out In reality a handful of very active hosts are dominating the platform.



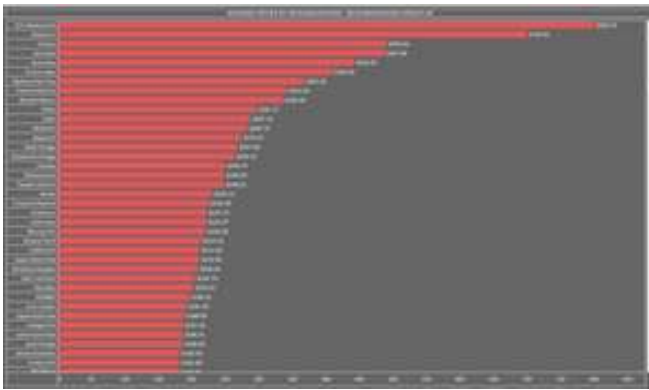
[Figure 3: Top 10 Hosts by Total Reviews — Bar Chart]

The average price allocation between different neighbourhood groups is depicted in the treemap above. Manhattan is the top dog in average price, with Brooklyn coming second. The Bronx and Staten Island are on the low end of the price spectrum. One can see here differences in economic and geographic conditions reflected through price points in the region. The average price allocation between different neighbourhood groups is depicted in the treemap above.

Average Prices by Neighbourhood This plot shows the breakdown of average price per neighbourhood. High-end places like Fort Wadsworth, Tribeca and Midtown show much

higher price levels, while the rest of the city is mostly at a medium to low price range. Such differences are a reflection of localized pricing strategies within the city.

This plot shows the breakdown of average price per neighbourhood. High-end places such as Fort Wadsworth, Tribeca and Midtown show much higher price levels, while the rest of the city is mostly at a medium to low price range. Such differences are a reflection of localized pricing strategies within the city.



[Figure 4: Average Price by Neighbourhood Group — Treemap]

Geographical Distribution of Listings



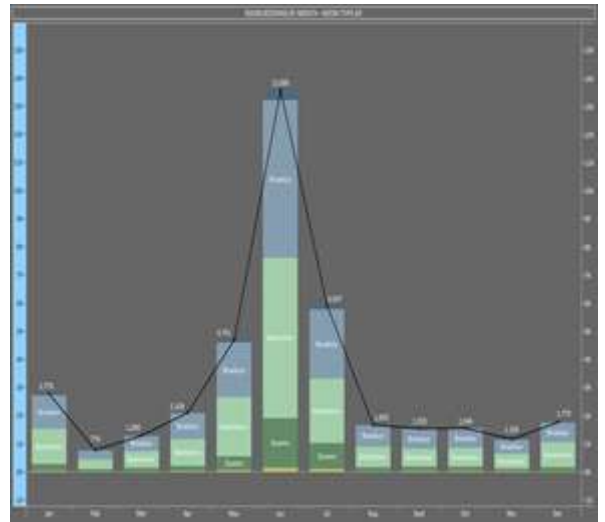
[Figure 5: Geographical Distribution Map]

The map plotted above shows the locations of Airbnb listings per their latitude and longitude coordinates. Manhattan and

Brooklyn are the two most densely packed areas of listings, thereby confirming these two boroughs as the main hubs of interest. Other areas are largely scattered with only a few listings. Other areas are largely scattered with only a few listings

Room Booking by Month

A combination of bars and lines chart displays the monthly booking trend. The pattern of bookings throughout the year is quite obvious, the highest booking seasons are found at the middle of the year (June and July), whereas bookings are lower in the months before and after. This can be interpreted as more people are traveling during the summer seasons.



[Figure 8: Room Booking by Month — Combined Chart]

V. CONCLUSION

This research verifies the power of visualizing data through Tableau to both assess Airbnb listings data and obtain insightful findings. The project, by simultaneously covering pricing, spatial location, types of rooms, and hosts' performance, thereby identifies major features of New York City short-term rental market.

The work finds out that Manhattan and Brooklyn are the top two boroughs about the number of Airbnb listings and level of booking activity. The price pattern points out premium neighborhoods tend to charge much higher prices whereas other areas provide more economical alternatives. The room type distribution indicates entire apartments and private rooms

are leading the popularity charts versus shared rooms, which aligns well with consumers' desire for privacy and comfort.

Also, the analysis points out that a handful of hosts account for a major portion of reviews that suggests we are dealing with host activity allocation being rather skewed. Data on booking seasonality demonstrate that demand is more than double during peak summer time periods, thereby indicating tourism cycles as the most likely factors. In general, adopting interactive dashboards not only makes easier comprehension of intricate datasets but also encourages evidence-based decision-making by hosts, customers, and policymakers. This analysis highlights the crucial role that visualization tools like Tableau play in converting unprocessed data into practical knowledge in the field of business analytics.

Acknowledgment

Syed Ibrahim Hussain would like to thank miss Sana Mateen, for her direction and support with Data Visualization Course and along this project work. Syed Ibrahim Hussain is also thankful to the colleagues at AI&DS Department for their help and cooperation. Thanks to Kaggle for dataset and Tableau public for visualization tools to carry out this project.

REFERENCES

1. Inside Airbnb. (2019). Airbnb New York City Dataset.
2. Source from <http://insideairbnb.com>
3. Airbnb Official Website. Source is from <https://www.airbnb.com>
4. Tableau Software. (2024). Tableau Documentation. Source is <https://www.tableau.com>
5. Kaggle. (2019). Airbnb NYC Dataset. Source is from <https://www.kaggle.com>.
6. Beyond the hype: Big data concepts, methods, and analytics. International Journal of Information Management
7. Few, S. (2006). Information Dashboard Design: The Effective Visual Communication of Data. O'Reilly Media.
8. Munzner, T. (2014). Visualization Analysis and Design. CRC Press.