Article:

Analytical solution for Short-term rental listings on Airbnb in Seattle.

-OHM SAI BHARGAV ALURI (000942201)

GitHub Airbnb

Table of Contents:

Selecting Dataset:

When selecting a dataset to create interactive dashboard, I opted the Airbnb Seattle dataset (2016-2017) from Kaggle. I found this dataset to be a great choice because it gives comprehensive details about property listings, including host information, property types, and booking calendars. It contains valuable data on availability, pricing, and booking status, which offers deep insights into the Seattle rental market. The one-year scope of the dataset helps me capture seasonal trends and market fluctuations. Using this dataset, I can build a dashboard in Power BI that gives meaningful trends for informed decision-making.

Reason For Choosing the Dataset:

- **Valuable Insights**: The dataset provides a complete view of Seattle's Airbnb market, including listings, pricing, and booking trends.
- **Real-world dynamics**: The data reflects **real-world dynamics** in a major city's short-term rental market, offering valuable Findings for Executives and policymakers.
- **Variety Of Metrics**: It contains information on **property types, locations, reviews, and availability**, enabling in-depth analysis.
- **Visualization Potential**: The dataset is well-Structured for building an **interactive dashboard in Power BI** to Investigate patterns and trends.
- **Open Source**: Accessible through **Kaggle**, making it a well-documented and reliable source for analysis.

Dataset Link:

Listings: final_calender_b

Bookings:final_listing

Introduction:

Airbnb was Founded in 2008; Airbnb has revolutionized travel by connecting guests with extraordinary lodging experiences. In Seattle, a vibrant and diverse city, Airbnb offers a wide range of accommodations—from downtown apartments to homes in beautiful neighbourhoods. I am analysing Seattle's Airbnb data to provide valuable insights into pricing trends, booking patterns, and neighbourhood dynamics, helping Executives and Hosts make more informed decisions.

Objective:

To Provide executives with a snapshot of the overall performance of Airbnb listings in Seattle, highlighting key metrics.

Initial Goals:

- ***Analyse Booking Trends:*** To identify trends in the Seattle homestay market
- **Visualize Listings**: Create an interactive map to display the locations of Airbnb listings in Seattle.
- **Understand Host Effectiveness**: To Understand the impact of host characteristics like response rate, on listing performance and customer satisfaction.

After analysing the dataset and identifying the best options, I Identified the key expectations which can achieve from the solution. These insights will help optimize decision-making and provide a clear View of Airbnb trends in Seattle.

Key Expectations from Solution/Final Goals:

1. **Pricing Insights:**
    - Identify average nightly rates for various property and room types.
    - Analyse how pricing affects booking rates.
2. **Booking Insights:**
    - Examine booking trends throughout the year and identify peak times.
    - Analyse customer preferences for different property types.
3. **Geographical Distribution:**
    - Determine which neighbourhoods have the highest concentration of listings and bookings.
    - Compare listings and bookings with average prices in those areas.
4. **Cancellation Rate:**
    - Identify cancellation rates throughout the year.
5. **Recognizing the top hosts of the year.**

Connecting The Data:

To build interactive dashboard, I used two primary datasets: the **Listings CSV** and the **Booking CSV**. The Listings CSV contains detailed information like property IDs, prices, amenities, and host details, while the Booking CSV contains booking statuses, calendar availability, and listing IDs.

Power BI was the main tool I used to integrate and connect these datasets. By Using Power BI's powerful visualization capabilities, I was able to create an interactive dashboard that Gives insights into pricing trends, booking patterns, and geographical distributions in Seattle's Airbnb market.

Initial Kips/Metrics Identified:

Total no of Listings, Total no of Bookings, Cancellation Rate, Average Nightly Prices Across Different Room Types, Overall Average Response Rate for All Hosts, Booking Trends Over Time by Hosts

Final Kips/metrics considered:

After considering the initial set of Kips. I have changed the measures according to the expectations. identified key visualization that helps the objective .

1. **Total no of Listings**: The Total number of active Airbnb properties listed in Seattle.

   Visualization: card

2. **Total no of Bookings**: The total no of bookings made across all Airbnb listings within the dataset.

   Visualization: card

3. **Cancellation Rate**: The percentage of bookings that were cancelled over the given period.

   Visualization: card

4. **Average Nightly Prices Across Different Room Types and property Types**: The average price per night, categorized by room type (e.g., entire home, private room, shared room). And another with property type.

   Visualization: bar chart

   X-axis: property type

   Y-axis: average Night price

5. **Overall Average Response Rate for All Hosts**: The average speed at which hosts respond to inquiries from guests.

   Visualization: Gauge chart

6. **Booking and Cancellation trends over the year:** A timeline analysis showing how Airbnb bookings, cancellations fluctuate throughout the year.

   Visualization: Clustered Bar chart

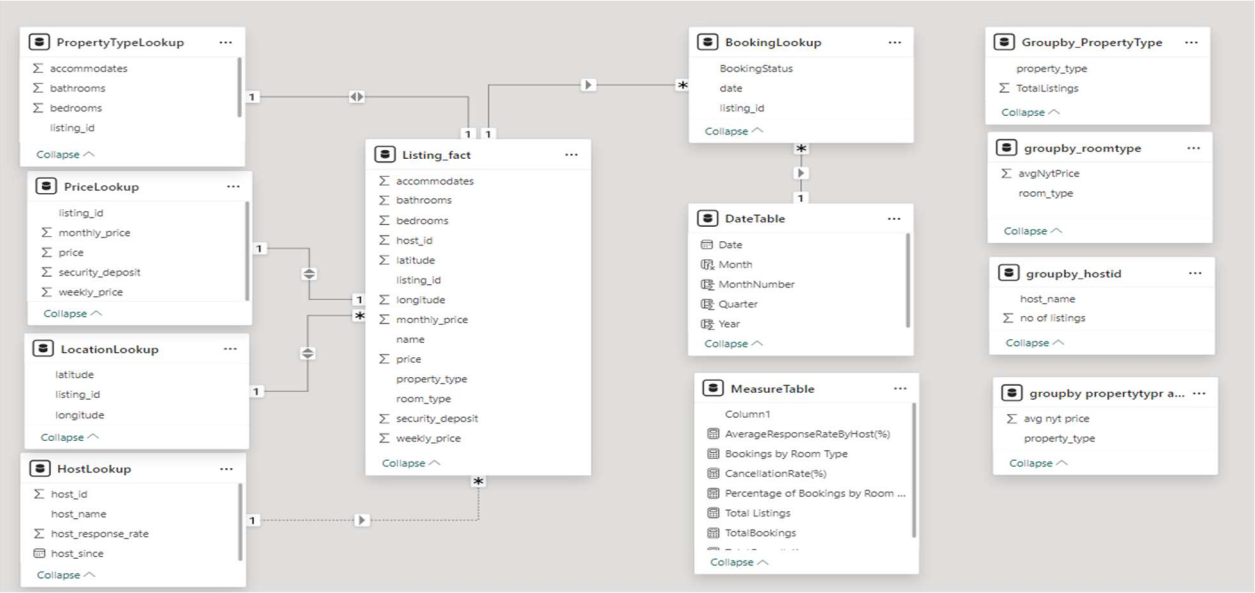
   X axis: months of the year

   Y axis: total no of bookings, total no of cancellations.

7. **Booking By Room type and property type**: shows the percentage of bookings by both the room type and property type.
   Visualization: pie chart

8. **Geographical distribution of listings and bookings**: it shows the geographical aspects of bookings and trends.
   Visualization: map

Data Model:



Data model includes lookup tables for PropertyType, Price, Location, Host, and Booking, all linked to the Listing_Fact table via Listing ID and Host ID, with a Date table connected to the Booking table.

Measure Table: I created measures like AvgResponseRate, BookingsByRoomType, CancellationRate, TotalListings, and TotalBookings for visualizations.

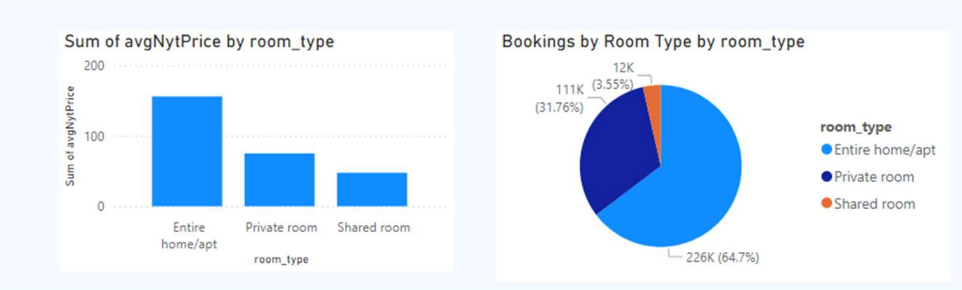
GroupBy Tables: I built four GroupBy tables:

1. PropertyType and Listings
2. RoomType and Avg Nightly Price
3. HostName and Number of Listings
4. Avg Nightly Price and PropertyType

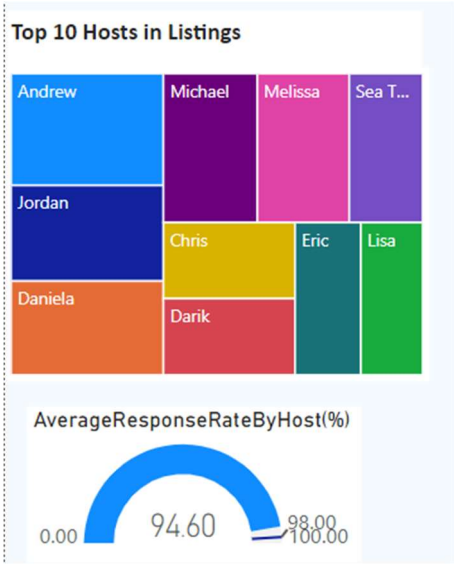
These were used to create the dashboard insights.

Results:

The Total Listings KPI indicated 1 million listings in Seattle. The Total Bookings KPI showed there were 524K bookings, showing an occupancy rate of 52%. The Cancellation Rate for that year was 33.32%. These KPIs were used to monitor monthly trends, booking types, and booking locations, providing insights into overall booking patterns.

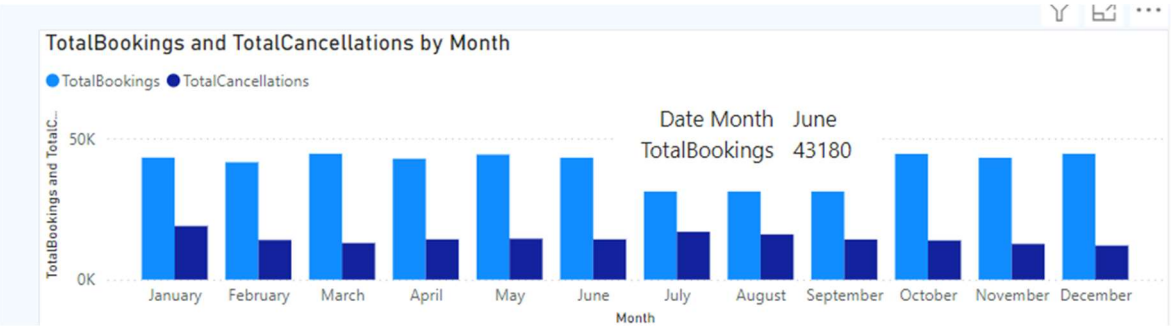


When comparing the average nightly price by room type with the total number of bookings, some interesting insights Identified. The average price for an entire home is $155.85, while the lowest is for a shared room at $47.55. Irrespective of the higher price, 64.57% of bookings are made for entire homes, most likely because families prefer them over shared or private rooms. Only 31.76% of bookings are made for private rooms, and very few opt for shared rooms, indicating people are reluctant to share with strangers. Entire home bookings also have a higher cancellation rate, as expected.

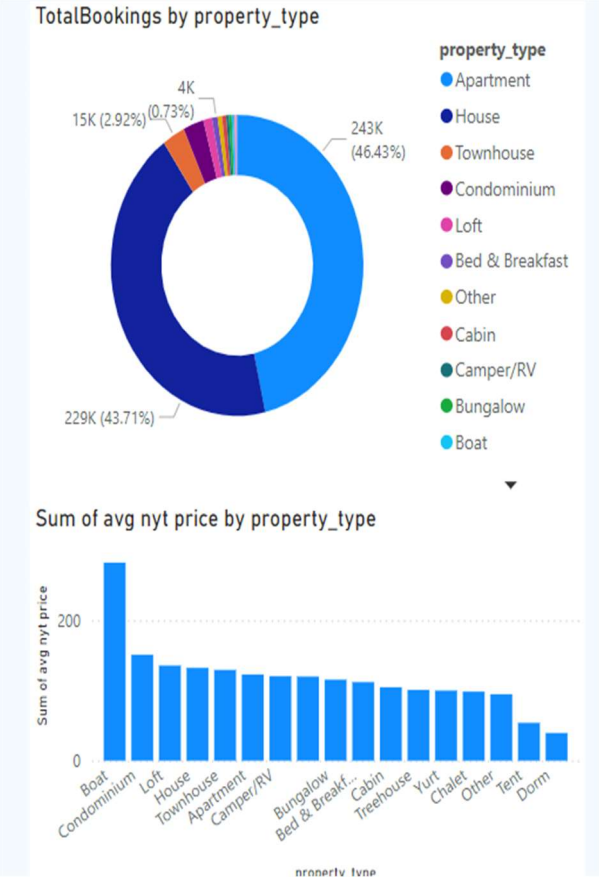


By Analysing the total number of listings and bookings, I identified the top 10 hosts in Seattle who gives the best service. Understanding the types of rooms and properties they list helps to understand customer preferences and trends. They listed the Entire homes more than the private rooms. Additionally, the average response rate was calculated at 94.60%, which needs improvement to 98% to ensure a smoother experience for guests.

https://github.com/OHMALURI/DATASCIENCE

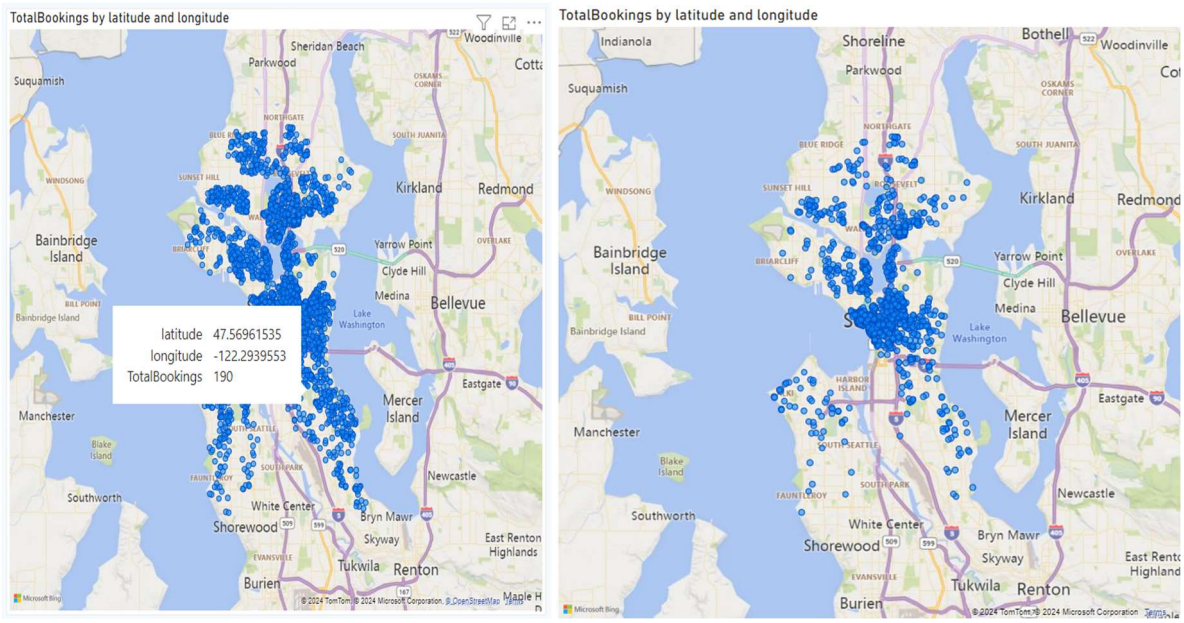


The clustered bar graph analysing the monthly distribution of bookings and cancellations shows that bookings remain steady throughout the year, Which is around 42-44K. Since Seattle is more of a business hub rather than a tourist destination, bookings stay relatively constant. The highest number of bookings occurred in June, with 43,180 due to summer vacations, while the lowest cancellations happened in December, likely because of the Christmas season. Overall, this shows consistent booking patterns, regardless of the time of year.

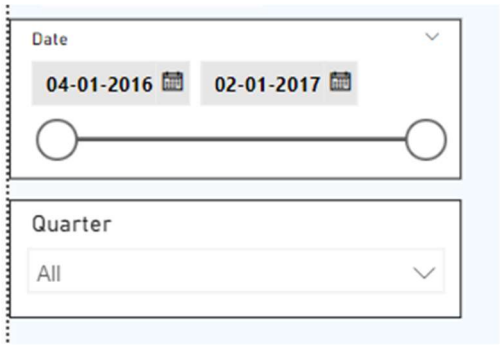


To get a detailed overview of the property types that people prefer, I created a pie chart showing bookings by property type and compared it with the average nightly price. The results shows that most bookings are for apartments (46.43%) and houses (43.71%) due to their affordability and higher availability. The least bookings were for boats, with only 1,113 bookings. Alternative accommodations like boats, cabins, lofts account for just 12% of total bookings. To create interest in these less popular options, promotions and discounts could be offered.

https://github.com/OHMALURI/DATASCIENCE



To unserstand the geographical distribution of bookings, maps were plotted for listings and bookings. The map on the left shows the listing distribution , while the one on the right shows booking distribution . This reveals that bookings are more in the central part of Seattle (the heart of the city), with more scattered bookings in areas outside the main city. By utilizing other metrics like property type, cancellation rate, and booking rate, we can gain deeper insights. The most demanding areas are the Bay Area and the city center.



A date slicer was used to analyse data for specific time periods, with quarters added to understand trends each quarter. The last quarter showed the highest number of bookings and the fewer cancellations, resulting in a higher profit margin. In contrast, the first quarter had lower bookings and was relatively slow.

https://github.com/OHMALURI/DATASCIENCE

Conclusion:

In this analysis, I used the Seattle Airbnb dataset from 2016-2017 and Power BI to create an interactive dashboard. The dashboard gave detailed visualizations and KPIs to understand booking trends, cancellation rates, and property preferences. With measures like average nightly price, and response rate, key insights were derived about customer mindset and host performance.

In conclusion, the analysis of the Airbnb Seattle dataset gave valuable insights into booking trends, customer preferences, and host performance. Key metrics such as cancellation rate, and average nightly prices helped to identify patterns in guest behaviour. Entire homes are the most popular, especially among families, managed to get 64.57% of bookings, while shared rooms are least considered. Top hosts offering entire homes improved in bookings but need to improve their response rate for a better Customer satisfaction. Geographically, bookings are concentrated at Seattle's city centre and Bay Area. The booking trends remain consistent throughout the year due to Seattle's status as a business hub, with the highest demand in June and lower cancellations in December. Lastly, giving promotions for less popular property types like boats and cabins could create more interest.

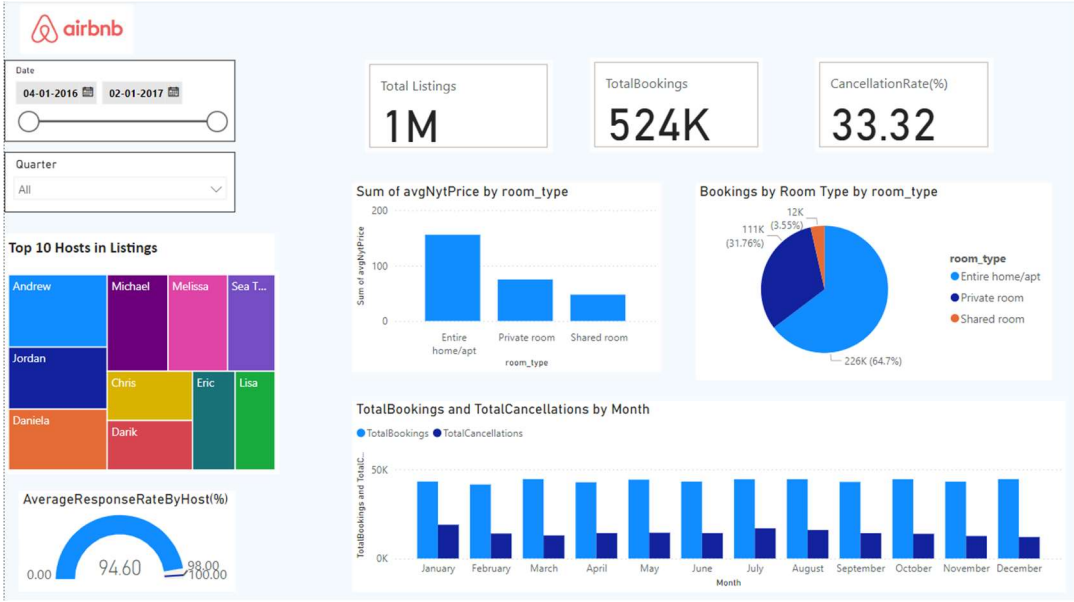
References:

Information regarding Airbnb: Airbnb

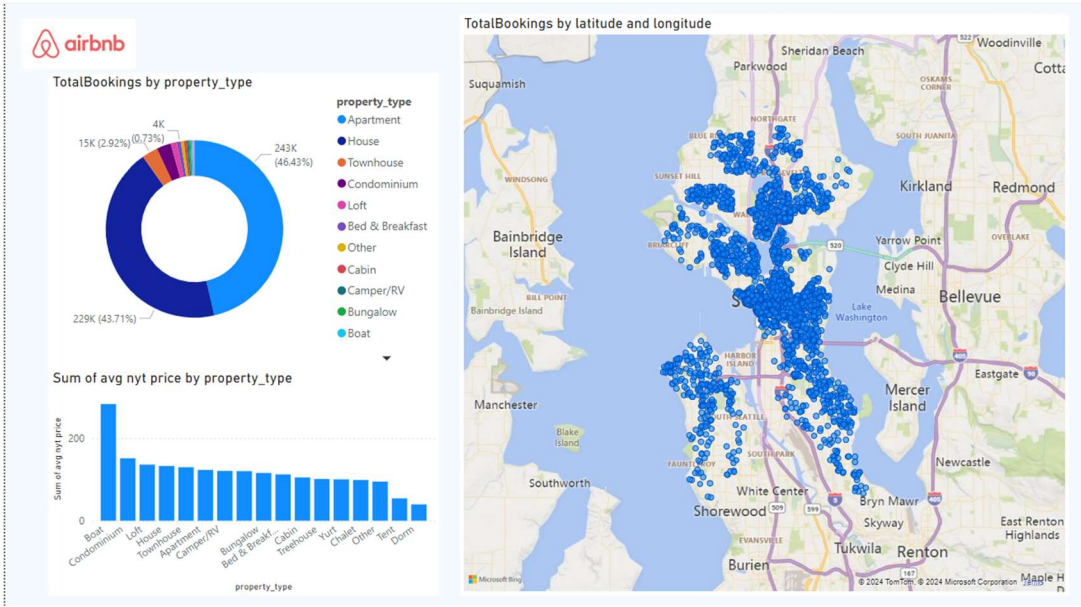
Dataset from Kaggle: Dataset_Reference

Appendix 1:

Final Dashboard:



This is the first page of the dashboard, providing an interactive report on listings, bookings, cancellations, and top hosts.



This is the second page of the dashboard, offering an analysis of property types and the geographical distribution of listings.

https://github.com/OHMALURI/DATASCIENCE