Airbnb Market Insights

for

Strategic Decision-Making

(An Albany, NY Case Study)

Mrunal Kiran

A20580436

Data Warehousing (ITMD-526)

1.Introduction and Problem Statement

1.1 Introduction

Albany, as the capital of New York State, attracts diverse visitors due to government-related travel, tourism, and local events. This influx creates a unique and vibrant Airbnb market, characterized by varying guest preferences, competitive pricing strategies, and distinct neighborhood performances. Understanding this market's intricacies is vital for hosts aiming to optimize their listings and for policymakers seeking to foster sustainable tourism growth.

1.2 Problem Statement

The primary aim of this analysis is to identify and evaluate key trends, guest engagement patterns, and listing performance within Albany's Airbnb market. By exploring these elements comprehensively, the analysis seeks to provide actionable insights that empower hosts and policymakers to enhance their strategies, maximize revenue opportunities, and better align with guest expectations and preferences.

2. Data Sources and Collection Methodology

2.1 Data Sources

The datasets used in this analysis were sourced from Kaggle, a public data repository. Specifically, Airbnb datasets covering Albany, NY, were utilized, including detailed listings, calendar availability, and guest reviews. Each dataset provided comprehensive data with more than 10,000 records and contained more than eight distinct variables, structured clearly into three primary tables: listings, calendar, and reviews.

2.2 Collection Methodology

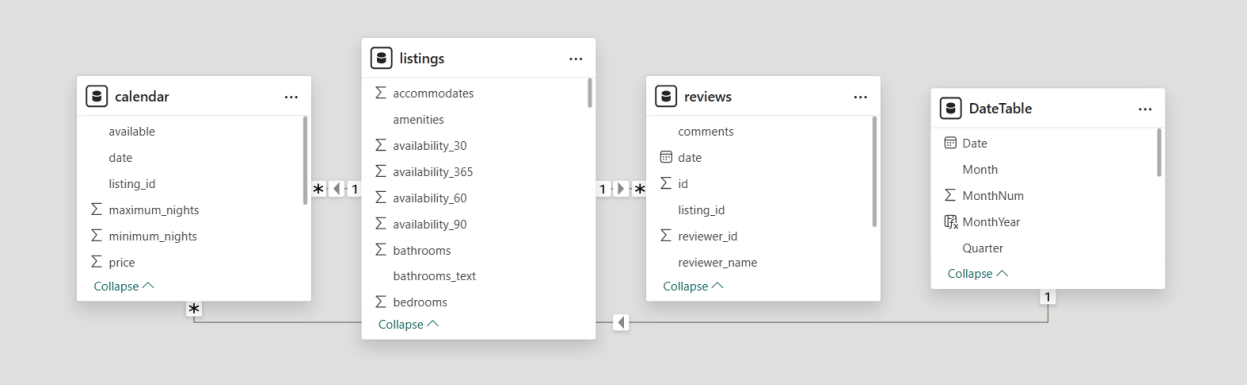
Data collection involved direct extraction of CSV files from Kaggle. Subsequent processing steps included data cleaning to remove columns with significant null values, converting price data from text to numeric formats, and transforming date fields from textual representations to standardized date formats. This meticulous data preparation ensured consistency, reliability, and suitability for detailed analytical exploration.

3. Data Model Design and Implementation

A semantic data model was developed in Power BI, establishing clear entity relationships:

- Listings: Central entity containing details about individual Airbnb listings.
- Calendar: Linked to listings by listing ID, providing pricing and availability details.
- Reviews: Linked to listings by listing ID, capturing guest engagement.
- Date Table: Custom-created to facilitate detailed temporal analysis.

An Entity-Relationship Diagram (ERD) illustrates the one-to-many relationships clearly.



3.1 Calculated Measures

Calculated measures and columns were introduced to enhance analytical capabilities:

3.1.1 Date Table

DateTable =

ADDCOLUMNS (

CALENDAR (MIN(calendar[date]), MAX(calendar[date])),

"Year", YEAR([Date]),

"Month", FORMAT([Date], "MMMM"),

"MonthNum", MONTH([Date]),

"Quarter", "Q" & FORMAT([Date], "Q")

)

This table enables detailed time-based analyses and drill-down capabilities.

3.1.2 Average Price

```
AvgPrice = AVERAGE(calendar[price])
```

This measure calculates the overall average listing price, providing insight into pricing strategies.

3.1.3 Total Reviews

```
TotalReviews = COUNT(reviews[id])
```

This measure quantifies guest engagement across listings

3.1.4 Reviews Per Listing

```
ReviewsPerListing =
DIVIDE(
    COUNT(reviews[id]),
    DISTINCTCOUNT(reviews[listing_id])
)
```

It indicates the average number of reviews per listing, reflecting guest satisfaction and listing popularity.

3.1.5 Total Listings

```
TotalListings = DISTINCTCOUNT(listings[id])
```

This measure provides the total number of unique listings, crucial for understanding market size and competition.

3.1.6 Listings Over Time

```
ListingsOverTime = DISTINCTCOUNT(calendar[listing_id])
```

Tracks the availability and variation in listings over specified periods.

3.1.7 Peak review Count

```
PeakReviewCount =
CALCULATE(
    MAXX(
        VALUES(DateTable[Month]),
        [TotalReviews]
    )
)
```

Identifies the highest review count in a month, highlighting peak engagement periods.

3.1.8 Most Active Month

```
MostActiveMonth =
VAR SummaryTable =
    ADDCOLUMNS(
        SUMMARIZE(DateTable, DateTable[Month]),
        "Total", [TotalReviews]
    )
VAR TopMonth =
    TOPN(1, SummaryTable, [Total], DESC)
RETURN
    MAXX(TopMonth, DateTable[Month])
```

Identifies the month with the highest guest activity, helping target marketing and operational decisions.

4. Visualization Approach

Power BI was selected due to its robust modeling, intuitive interface, dynamic filtering capabilities, extensive documentation, and ability to support custom visualizations such as geographic maps beyond default options. Its versatility and ease of use made it ideal for the analysis requirements.

5.Dashboards Construction

Three interactive dashboards were constructed to support a clear narrative arc:

- **Executive Summary:**
  - Presents high-level KPIs—total listings, average price, total reviews, and reviews per listing.
  - Interactive filters allow focused analysis by neighborhood and room type.
  - Provides an initial understanding of Albany's Airbnb market.
- **Guest Engagement and Listing Performance:**
  - Reveals critical relationships through bar and scatter charts.
  - Highlights correlations between listing prices, review counts, and room types.
  - Demonstrates dynamics of guest preferences and factors influencing listing success.
- **Geographical Analysis:**
  - Identifies key geographical insights using a custom map visual.
  - Highlights neighborhood-specific guest activity.
  - Specifically identifies the Sixth Ward as a hotspot for targeted strategic decisions.

5.1 Business Insights

- High guest preference for Entire homes/apartments, suggesting targeted promotions for these property types.
- Neighborhood analysis identifies areas like the Sixth Ward as hotspots, providing hosts and marketers precise locations to prioritize.
- Correlation between higher pricing and higher reviews indicates guests' willingness to pay premium prices for higher-rated properties, guiding pricing strategies.

6. Challenges Encountered and Solutions Implemented

Initial challenges included inconsistent data formats, especially pricing and dates, resolved effectively using Power Query transformations. The limited temporal range posed a challenge for

trend analysis, addressed by adjusting the temporal scope to daily trends within the available September 2024 data.

7. Conclusion

This analysis provides clear, actionable insights into Albany's Airbnb market. Hosts and policymakers can leverage these insights to enhance listings, optimize pricing strategies, and focus marketing efforts effectively. Continued application of these strategies, along with future enhancements such as predictive analytics, will support sustained growth and competitive advantage in the dynamic Airbnb marketplace of Albany.

8. Future Enhancements

Potential future improvements include integrating external data sources, such as event calendars, to better understand demand fluctuations. Advanced predictive analytics could also be employed to enhance forecasting and strategic planning.