

Airbnb Data Analysis

Team Members:

- Syed Samiullah Ghorl (Team Leader)
- Muhammad Arshad Mansoor
- Muhammad Maaz Maqsood
- Azka Salman Farsi
- Muhammad Omer Jamil

Introduction

We conducted an analysis of Airbnb data in Lisbon, Portugal to identify popular neighbourhoods and the top 5 neighbourhoods with the highest average listing price. The analysis also included exploring the distribution of room types and their average prices, as well as plotting the availability of Airbnb listings over time.

Data

We used the following files for the analysis:

- ❖ listings.csv: This file contains Airbnb listing data, such as the price, number of bedrooms, reviews per month, and room type.
- ❖ neighbourhoods.csv: This file contains neighbourhood data, such as the name of the neighbourhood and the borough it belongs to.
- ❖ calendar.csv: This file contains availability data for Airbnb listings.

Cleaning the Data

To prepare the data for analysis, we cleaned it using the following steps:

- ❖ Removed duplicates from the listings data
- ❖ Removed listings with missing data for the price, reviews_per_month, and bedrooms columns
- ❖ Converted the price column to float and added the euro (\$) sign to the prices
- ❖ Converted the last_review column to datetime format
- ❖ Replaced values in the room_type column to make them more readable
- ❖ Merged the listings and neighbourhoods data and dropped duplicates from the neighbourhood column.

We added try/except blocks to handle various types of errors that could occur during data loading, cleaning, analysis, and plotting.

Analyzing the Data

We used the following functions to analyze the cleaned data:

- ❖ `get_popular_neighborhoods`: This function identifies the top 5 most popular neighbourhoods in Lisbon by grouping the data by neighbourhood group and sorting by popularity.
- ❖ `get_avg_price_neighborhoods`: This function identifies the top 5 neighbourhoods with the highest average listing price in Lisbon by grouping the data by a neighborhood group and sorting by price.
- ❖ `get_room_type_distribution`: This function explores the distribution of room types in Lisbon and calculates their average prices.
- ❖ `plot_availability_calendar`: This function plots the availability of Airbnb listings over time by grouping the data by month and calculating the mean availability.

We used the matplotlib library to create the availability over time plot. This plot was created to explore the trend of Airbnb listings availability over time and to identify the high and low season months in Lisbon.

Results

After running the analysis, we found the following results:

- ❖ The top 5 most popular neighbourhoods in Lisbon are:
 - Lisboa
 - Cascais
 - Sintra
 - Mafra
 - Lourinha.

Lisboa has the highest number of listings with **12082**.

- ❖ The top 5 neighbourhoods with the highest average listing price in Lisbon are:
 - Sintra (\$153.48)
 - Cascais (\$147.82)
 - Azambuja (\$141.71)
 - Mafra (\$132.90)
 - Arruda Dos Vinhos (\$130.75).
 - The average prices of different room types in Lisbon are:
 - Entire (\$149.12)

- Hotel (\$155.06)
 - Private (\$65.36)
 - Shared (\$35.20).
- The availability of Airbnb listings in Lisbon has a seasonal trend, with the lowest availability occurring in the summer months from June to August.

Conclusion

This analysis provides insights into the popularity and pricing of Airbnb listings in Lisbon, Portugal. The results can be used by Airbnb hosts and guests to make informed decisions about pricing and location. Based on the analysis, it is recommended that hosts in Lisbon adjust their prices and availability based on the seasonal trends to maximize their bookings and profits.