

Project Premise and Objectives for Achievement

This project aims to better understand Airbnb pricing dynamics in popular European cities through a thorough analysis of various factors shaping rental rates. The dataset driving this exploration offers a nuanced perspective on Airbnb listings, encompassing attributes such as room types, cleanliness and satisfaction ratings, person capacity, distance from the city center, host superhost status, business indicators, and geographical coordinates. The goal is to capture a comprehensive understanding of Airbnb pricing trends and their determinants.

Source

The dataset is sourced from a diverse range of Airbnb listings across several major European cities. Airbnb, a globally recognized platform for short-term rentals, offers a vast and varied dataset comprising listings from various hosts and regions.

Data Contents

Each entry in the dataset represents a specific Airbnb listing, providing information on essential variables such as room type, pricing, cleanliness, satisfaction ratings, and location-related attributes. These variables contribute to a holistic understanding of the factors influencing Airbnb prices across European cities.

Data Limitations and Considerations

Data limitations should be acknowledged in this project. The dataset, centered around Airbnb pricing in European cities, may lack certain critical variables that influence outcomes, such as local events or specific host characteristics. Additionally, the dataset's temporal span may limit its ability to capture current market dynamics.

Questions to Consider

- How do Airbnb prices vary across different European cities?
- What are the key factors contributing to guest satisfaction in Airbnb listings?
- Does the distance from the city center influence both the pricing and overall guest satisfaction of Airbnb listings, and is there a relationship between these two factors?
- What is the relationship between a host's superhost status and both the pricing and cleanliness ratings of their listings?
- Are there regional patterns or clusters where both pricing and overall guest satisfaction align, indicating a geographical relationship?
- Is there a correlation between the attraction index (proximity to attractions) or distance to city center and overall guest satisfaction, and does this relationship vary across cities?

Data Profile					
Variable	Description	Time Variant/ Invariant	Structured/ Unstructured	Quantitative/ Qualitative	Nominal /Ordinal/ Discrete/ Continuous
City	The name of the city.	Invariant	Structured	Qualitative	Nominal
Time of Week	Weekend or weekday day.	Variant	Structured	Qualitative	Nominal
Room Type	Type of room.	Invariant	Structured	Qualitative	Nominal
Room Shared	Indicator whether the room is shared (True/False).	Invariant	Structured	Qualitative	Nominal
Room Private	Indicator whether the room is private (True/False).	Invariant	Structured	Qualitative	Nominal
Person Capacity	The capacity of the room in terms of the number of people.	Invariant	Structured	Quantitative	Continuous
Multiple Rooms	Indicator whether multiple rooms are available (True/False).	Invariant	Structured	Qualitative	Nominal
Business Listing	Indicator whether it's a business listing (True/False).	Invariant	Structured	Qualitative	Nominal
Number of Bedrooms	The number of bedrooms in the listing.	Invariant	Structured	Quantitative	Discrete
Total Price	The total price of the listing.	Invariant	Structured	Quantitative	Continuous
Cleanliness Rating	The cleanliness rating given by guests.	Invariant	Structured	Quantitative	Continuous
Overall Guest Satisfaction	The overall satisfaction rating given by guests.	Invariant	Structured	Quantitative	Continuous
Distance from City Centre	The distance of the listing from the city center.	Invariant	Structured	Quantitative	Continuous
Distance from Nearest Metro Station	The distance of the listing from the nearest metro station.	Invariant	Structured	Quantitative	Continuous
Attraction Index	The attraction index of the city.	Invariant	Structured	Quantitative	Continuous
Normalized Attraction Index	Normalized version of the attraction index.	Invariant	Structured	Quantitative	Continuous
Restaurant Index	The restaurant index of the city.	Invariant	Structured	Quantitative	Continuous
Normalized Restaurant Index	Normalized version of the restaurant index.	Invariant	Structured	Quantitative	Continuous
Longitude	The longitude coordinates of the listing.	Invariant	Structured	Quantitative	Continuous
Latitude	The latitude coordinates of the listing.	Invariant	Structured	Quantitative	Continuous