

Proposal: NYC Airbnb

April 05, 2024

Group No. – 6

Team Members

Mahima Bhasin

Het Shah

Shiva Talari

Simran Panda

Swapnil Mungi

Bharath Kalyan Nanganuri

Manohar Murthy Raju Mantena

1. Introduction

Airbnb has revolutionized the travel and hospitality industries since its 2008 founding. It has turned ordinary individuals into hospitality providers and disrupted traditional travel by generating billions of dollars in revenue and offering over six million listings worldwide.

Our study specifically attempts to pinpoint important Airbnb characteristics with high price predictability. We plan to conduct numerous hypothesis tests and univariate and multivariate regression analyses to identify variables that have the most influence on Airbnb's pricing mechanism.

Research Question: The research aims to investigate the causal effect of various factors (X) on the pricing strategy (Y) of Airbnb listings. It specifically aims to comprehend how variables like location, home type, amenities, and host attributes affect Airbnb hosts' pricing choices.

Motivation: It is important to investigate how these elements affect pricing strategy because of a number of reasons. Initially, it can offer valuable perspectives about Airbnb hosts' pricing practices, which is crucial for both hosts and guests to make knowledgeable choices. Secondly, it has the potential to facilitate policymakers' understanding of the sharing economy and its impact on traditional industries such as hotels. Finally, it can contribute to the broader knowledge on pricing strategy and consumer behavior in the context of the sharing economy.

Methodology: To evaluate the Airbnb data, the study will use a variety of econometric techniques, including multiple regression analysis and causal inference methods. It will use a dataset containing information on Airbnb listings, including pricing, location, property type, amenities, host characteristics, and guest reviews. To confirm the validity of the findings, the study will control for robustness checks and control for potential confounding variables.

Additional Information: The research intends to add to the expanding collection of knowledge regarding pricing strategy and the sharing

economy. Researchers and professionals in the domains of economics, business, and politics might benefit greatly from the study's comprehension of the elements that affect pricing decisions on Airbnb.

2. Methodology

Research question(s)

- **How does Airbnb listing prices depend on rating of listing, location of listing, amenities included in the listing, and host response time?**

We seek to ascertain the variables influencing Airbnb listing prices with this research topic. Four independent variables (i.e., the listing's rating, location, amenities, and host response time) and one dependent variable (i.e. the listing price) make up the five variables.

The location and facilities offered by the listing are categorical factors, whereas the listing price, rating, and host response time are quantitative criteria.

This study question is crucial because it helps us comprehend the Airbnb pricing structure, which in turn helps us understand the sharing economy's larger dynamics and how websites like Airbnb are transforming the conventional hotel sector.

Our theory is that higher listing costs will result from a listing with a high rating, a longer list of amenities provided, a prompt host response time, and a listing in a convenient and safe place.

Equation :

$$\text{Listings_Price} = a + b \cdot \text{amenities} + c \cdot \text{neighbourhood} + d \cdot \text{review_scores_rating} + e \cdot \text{host_response_time} + f$$

- **Endogeneity Issues**

1. There is a potential for bias in the reviews and ratings provided by individuals, which could lead to lower ratings for properties that are of high quality.
2. This bias may affect the perceived value of these properties, potentially influencing their prices.

3. Neighborhood features strongly impact Airbnb listing prices, with proximity to attractions, amenities, and transportation hubs, as well as safety and accessibility, commanding higher prices.
 4. Demand, supply, and seasonal fluctuations also play key roles, posing challenges for pricing models due to potential omitted-variable bias.
- **How can we address the Endogeneity issues?**
 1. Include Relevant Variables: Include additional relevant variables in the regression model that capture neighborhood characteristics, such as proximity to attractions, amenities, transportation hubs, safety measures, and accessibility.
 2. Use Instrumental Variables (IV): If there are endogeneity concerns with neighborhood features, consider using instrumental variables to estimate the relationship between neighborhood characteristics and listing prices.

3. Data

The `airbnb_data` dataset comes from Inside Airbnb (<http://insideairbnb.com/get-the-data/>), an open platform that provides data on Airbnb listings in different locations around the world.

The objective of Inside Airbnb is to empower communities by providing them with information and data regarding the impact of Airbnb on residential areas. Their mission is to enable communities to make informed decisions and have control over the practice of renting homes to tourists, with the ultimate goal of achieving a vision where data is used to shape this industry. The dataset is aggregated through Airbnb's public information on their website and contains information on over 42,000 listings in New York City as of March 6, 2023.

In `airbnb_data`, the observations (rows) are different Airbnb listings in NYC and the attributes (columns) are various variables that describe the listing.

It describes the host and who they are, ratings on the host and place of stay, etc. Some variables in the dataset that we chose to look into: price, neighbourhood_group, amenities, review_scores_rating

For analysis, we plan to use linear regression to model the relationship between price and various factors, such as neighborhood and property type.

While the dataset provides valuable insights, it has some limitations. For example, the dataset only includes listings from a specific time period and may not represent the current market conditions. We will address these limitations by conducting sensitivity analyses and comparing our results with external data sources.

The dataset will be made available to others for research purposes. Researchers interested in accessing the dataset can request access through Airbnb's data sharing platform.

Data Summary

| Name | airbnb_data |
|------------------------|-------------|
| Number of rows | 42931 |
| Number of columns | 75 |
| Column type frequency: | |
| character | 26 |
| Date | 5 |
| Logical | 7 |
| numeric | 37 |
| Group variables | None |

4. Conclusion and Potential findings

In conclusion, our project proposal aims to investigate the determinants of Airbnb listing prices by examining factors such as listing ratings, location, amenities, host response time, and the characteristics of listings and hosts.

Our preliminary analyses suggest that these variables significantly influence pricing decisions on Airbnb.

We have observed that listings with higher ratings, comprehensive amenities, quick host response times, and favorable locations tend to command higher prices. This indicates that the quality of the listing and its perceived value play crucial roles in determining its price.

Additionally, the appeal of listings and hosts, as reflected in their ratings and descriptions, also impacts pricing, with more attractive listings and reputable hosts being able to set higher prices.

Our findings underscore the importance of location, amenities, host characteristics, and seasonal variations in shaping Airbnb prices.

Future research will delve deeper into these aspects to provide a more nuanced understanding of their impact on pricing.

Moreover, our study has implications for policymakers in terms of regulating short-term rentals and understanding their impact on traditional hospitality industries.

Overall, this project proposal sets the stage for a comprehensive examination of the factors influencing Airbnb listing prices. By understanding these determinants, hosts can optimize their offerings to attract guests and maximize revenue, while guests can make more informed decisions when selecting accommodations. Our research aims to contribute to the broader literature on pricing strategies and consumer behavior in the context of the sharing economy.