Airbnb Booking Analysis

Ankit Sharma

Dsn Naveen Kumar

Pankaj Kumar

Abstract:-

Since 2008, guests and hosts have used Airbnb to expand on Since 2008, guests and hosts have used Airbnb to expand on traveling possibilities and present a more unique, personalized way of experiencing the world. Today, Airbnb became one of a kind service that is used and recognized by the whole world. Data analysis on millions of listings provided through Airbnb is a crucial factor for the company. These millions of listings generate a lot of data - data that can be analyzed and used for security, business decisions, understanding of customers' and providers' (hosts) behavior and performance on the platform, guiding marketing initiatives, implementation of innovative additional services and much more.traveling possibilities and present a more unique, personalized way of experiencing the world. Today, Airbnb became one of a kind service that is used and recognized by the whole world. Data analysis on millions of listings provided through Airbnb is a crucial factor for the company.

Our experiment will help analyse the data generated by millions of listings by Airbnb Booking data with the help of python programming, MS-Excel tools, and Tableau.

These findings could be used for security, business decisions, understanding of customers, marketing and research analysis

Keywords: Python, Tableau, dynamic pricing, reviews

1.Problem Statement

Data provided by Airbnb bookings gives us statistics about booking pattern of users, the bookings they make frequently, users with maximum booking and stay.

- Destinations:- Top destinations are Bronx, Staten Island, Manhattan, Queens, Brooklyn.
- Customer Reviews:- Customer reviews about their booking and stay experience.

The main objective is to build a analysis model which could help in contributing towards realisation of data for future purposes like predicting guest habits and help in marketing purposes as well.

Variables are as follows:-

Host name:- Name of the host

Host_id :- Booking ID

Neighbourhood :- Booking city

Room_type :- Category of the room chosen for booking

Minimum_nights :- Minimum number of nights the guest stayed at a Hotel room

Number_of_reviews:- Number of reviews given for a particular Hotel and room

Calculated_host_listing :- number of host/guest per hotel per city per hotel.

2. Introduction

Airbnb booking analysis uses data from hotel bookings available from hotels and website of Airbnb. This project basically deals with behaviour of the guests and can help us to analyse data of their likings about place, room type etc.

We will be able to analyse that by raw data available to us.

3. How host and places were picked

Places with highest number of bookings were picked and out of that guests with highest number of bookings and reviews were picked for further study.

4. Steps Involved:

- **Exploratory Data Analysis**
  After loading the dataset we performed this method by comparing our target variable that is Host name and number of bookings. Type with other independent variables. This process helped us figuring out various aspects and relationships among the target and the independent variables. It gave us a better idea of which feature behaves in which manner compared to the target variable.

- **Null values Treatment**
  Our dataset contains a large number of null values which might tend to disturb our accuracy hence we dropped them at the beginning of our project inorder to get a better result.

- **Encoding of categorical columns**
  We used One Hot Encoding to produce binary integers of 0 and 1 to encode our categorical features because categorical features that are in string format cannot be understood by the machine and needs to be converted to numerical format.

- **Standardization of features**
  Our main motive through this step was to scale our data into a uniform format that would allow us to utilize the data in a better way while performing fitting and applying different algorithms to it.

  The basic goal was to enforce a level of consistency or uniformity to certain practices or operations within the selected environment.

- **Using different tools**
  For analysing data we used different data interpreting tools like:-
  1. MS-Excel
  2. Tableau
- **Representation of data :-**
  After analysing data we represented data into meaningful manner by the help of charts like pie charts and graphs which are quite easy to understand.

5. Conclusion:-

That's it! We reached the end of our exercise.

Starting with loading excel file to google collab by mounting google drive and then creating it into columns further and sorted useful data out of data available to us, then with the help of

python codes we converted data into Pictorial graphs and then represented the same with Tableau and MS-Excel.

In all of our datasets modelling accuracy is >75%.

So the accuracy of our best model is >75% which can be said to be good for this large dataset. This performance could be due to various reasons like: no proper pattern of data, too much data, not enough relevant features.