

# **Exploratory Data Analysis**

**Beyond the Hotel: A Deep Dive into Airbnb in Asheville, NC  
- Trends, Prices, and More**

**Final Project Proposal**

**Fall 2023**

**Group 12**

**Lakshmi Sai Ram Marupudi**

**Niharika Ganji**

**Yerramsetti Dharma Teja**

**Sravya Vujjini**

**Dataset used :** <http://insideairbnb.com/get-the-data/>

**Dataset Title:** Airbnb Listings in Asheville, NC

**Description:** The Airbnb Listings Dataset is a comprehensive compilation of records representing individual property listings on Airbnb. Each entry in the dataset corresponds to a unique Airbnb listing, including details ranging from basic listing attributes to host-related information and customer review metrics.

### **Contents:**

- **Basic Listing Information:** Includes unique identifiers for each listing (**id**, **listing\_url**), descriptive details such as the name (**name**), and a brief description (**description**). It also features URLs to images of the property (**picture\_url**).
- **Host Details:** Contains information about the listing's host, such as the host's ID (**host\_id**), profile URL (**host\_url**), name (**host\_name**), the date they joined Airbnb (**host\_since**), location (**host\_location**), and other personal insights provided by the host (**host\_about**). It also includes host response metrics like response time (**host\_response\_time**), response rate (**host\_response\_rate**), and acceptance rate (**host\_acceptance\_rate**), along with indicators of whether they are a superhost (**host\_is\_superhost**) and if their identity is verified (**host\_identity\_verified**).
- **Geographic Information:** Offers geographical details of the listing, including the neighborhood (**neighbourhood**), cleansed neighborhood name (**neighbourhood\_cleansed**), and precise location coordinates (**latitude**, **longitude**).
- **Property Features:** Details about the property type (**property\_type**), room type (**room\_type**), the number of guests it can accommodate (**accommodates**), the number of bathrooms (**bathrooms\_text**), bedrooms (**bedrooms**), and beds (**beds**), along with a list of amenities (**amenities**).
- **Pricing and Availability:** Information on the price (**price**), minimum and maximum nights for stays (**minimum\_nights**, **maximum\_nights**), and availability

over various periods (**availability\_30**, **availability\_60**, **availability\_90**, **availability\_365**).

- **Review Scores:** Comprehensive review data, including the number of **reviews** (**number\_of\_reviews**, **number\_of\_reviews\_ltm**, **number\_of\_reviews\_l30d**) and detailed review scores for various aspects like overall rating, accuracy, cleanliness, check-in, communication, location, and value (**review\_scores\_rating**, **review\_scores\_accuracy**, **review\_scores\_cleanliness**, etc.).

- **Booking Details:** Contains booking-related features such as whether the property is instant bookable (**instant\_bookable**), various calculated counts of the host's listings (**calculated\_host\_listings\_count**, etc.), and reviews per month (**reviews\_per\_month**).

## **Dataset Size:**

The original dataset has listings across various cities around the world. We selected the city Asheville, North Carolina because of our team's affiliation with North Carolina (our mini project's team name). Also, the dataset (Asheville, NC) includes 3288 observations including numerous quantitative variables (meeting the requirement of number of observations and quantitative variables), each with detailed attributes as listed above.

## **Research Questions :**

1. After going through the attributes present in the dataset, our initial thoughts were around predicting the price of a listing based on its location, size (number of bedrooms and bathrooms) and the time of the year.
2. We'd also like to explore the possibility of answering what factors contribute most to a guest's perceived value of a listing(review score) as this can assist hosts in focusing on the factors that matter most to guests, giving them a scope to improvise their businesses.