**Dataset: "Airbnb Listings"**

**Airbnb dataset** from Kaggle or Inside Airbnb (<http://insideairbnb.com/>), which contains real-world listings, including prices, locations, ratings, and availability.

**Data Cleaning Tasks:**

1. **Handling Missing Data**
   * Identify columns with missing values.
   * Fill missing values appropriately (e.g., replace missing prices with the median, missing reviews with 0).
   * Drop irrelevant columns.
2. **Fixing Inconsistent Data**
   * Standardize city names and ensure proper casing.
   * Convert price columns to numeric values (removing $ signs).
   * Correct inconsistent date formats.
3. **Removing Duplicates**
   * Find and remove duplicate listings based on host ID and listing name.
4. **Outlier Detection**
   * Detect and handle outliers in the price column using IQR or z-score.
   * Identify properties with unrealistic availability (e.g., more than 365 days available).

**Visualization Tasks:**

1. **Price Distribution**
   * Plot a histogram of listing prices to understand price ranges.
2. **Room Type Comparison**
   * Use a bar chart to show the count of different room types (Entire home, Private room, Shared room).
3. **Location Analysis**
   * Create a scatter plot or heatmap of listings on a city map using latitude & longitude.
4. **Availability vs. Price**
   * Use a scatter plot to check the relationship between availability and price.
5. **Top Hosts**
   * Find the top 10 hosts with the most listings and visualize using a bar chart.
6. **Ratings Analysis**
   * Plot a boxplot of ratings to check how guest ratings are distributed.