Airbnb Listings Analysis

Key Findings

1. **Price Distribution**
   - The price distribution is highly right-skewed.
   - The majority of listings are priced below **$300**, while a small number of luxury properties exceed **$10,000**.
   - Median price: **$169**, Mean price: **$577.88**.

2. **Location Insights**
   - The **Near North Side**, **West Town**, and **Near West Side** neighborhoods have the highest number of listings.
   - These areas are centrally located and cater primarily to tourists and business travelers.
   - Prices in these neighborhoods are significantly higher than in suburban or rural areas.

3. **Room Type Analysis**
   - Most listings (**approximately 76%**) are **Entire homes/apartments**, while **22%** are **Private rooms**.
   - Indicates that property owners and professional hosts dominate the market.

4. **Pricing by Room Type**
   - **Hotel rooms** show extreme price outliers (over $20,000), possibly due to data errors or luxury hotel listings.
   - **Entire homes/apartments** average around **$288**, while **private rooms** average **$555**.
   - Indicates inconsistency, suggesting a need for data cleaning or improved room type classification.

5. **Reviews and Ratings**
   - Weak negative correlation (-0.05) between **price** and **number of reviews**, implying that higher-priced listings tend to have fewer reviews.
   - Budget-friendly options are reviewed more frequently, indicating higher occupancy or guest turnover.

6. **Geographical Distribution**
   - The majority of listings are concentrated in central urban areas, especially **downtown Chicago**.
   - Suburban and rural regions have fewer listings and lower prices.

7. **Seasonality in Bookings**
   - Review activity peaks around **May–June**, suggesting higher tourism and booking rates during the summer season.

8. **Data Quality Observations**
   - The "neighborhood group" column contains missing or undefined values (recorded as "Unknown").
   - Certain listings show unrealistic price or availability values, requiring further cleaning.

b. Supporting Visualizations
- Histogram of Price Distribution
- Bar Chart of Top 10 Neighborhoods by Listing Count
- Pie Chart of Room Type Distribution
- Boxplot showing Price Outliers
- Interactive Map of Listing Locations (Mapbox Visualization)
- Line Chart of Reviews Over Time
- Correlation Heatmap and Pair Plot

These visualizations demonstrate patterns in price, location density, room type dominance, and review behavior.

2. Business Recommendations

1. **Pricing Strategy Optimization**
   o Implement **dynamic pricing models** that adjust rates based on seasonality, demand, and location.
   o Filter out extreme outliers to avoid misleading averages in price algorithms.

2. **Host Performance Improvement**
   o Encourage hosts to maintain **competitive pricing** for better visibility and more frequent bookings.
   o Provide guidance on **amenities and presentation** that correlate with higher review counts and ratings.

3. **Location-Based Expansion**
   o Airbnb can focus marketing efforts on **underrepresented neighborhoods** with lower listing density but potential for growth.
   o Promote suburban or near-city areas for travelers seeking quieter or affordable stays.

4. **Data Quality Enhancements**
   o Standardize room type classifications to prevent mislabeling of hotel rooms and private stays.
   o Improve data collection accuracy for neighborhood grouping and availability fields.

5. **Customer Experience Enhancement**
   o Encourage hosts to respond quickly to reviews and maintain high cleanliness and communication scores.
   o Offer incentives or "Superhost" status for consistent positive feedback and frequent bookings.

6. **Seasonal Campaigns**
   o Leverage peak months (May–June) by offering **discount promotions** or **special events** to maximize occupancy.
   o Use predictive models to forecast demand and optimize listing visibility during off-peak months.