

WHERE SHOULD INVESTORS INVEST? A DATA-DRIVEN ANALYSIS OF AIRBNB CITY PERFORMANCE



Intro

Dataset - [Kaggle](#) : Inside AirBnB - USA

The project analyzes Airbnb performance from an investor's perspective.

Like which cities offer strong profitability and consistent demand.



- Goals:**
- Identify high-price cities
 - Measure demand through occupancy
 - Evaluate guest sentiment
 - Determine revenue performance
 - Recommend top markets for investment



Dataset & Wrangling Process

1. Data cleaning

- Converted all pricing fields to numeric values**
- Removed the “\$” symbol and “,” from price columns.
 - Converted text-based price fields into numeric types.



+ city labels to enable cross-city comparison

Added a new column indicating the city name before stacking multiple CSV files.

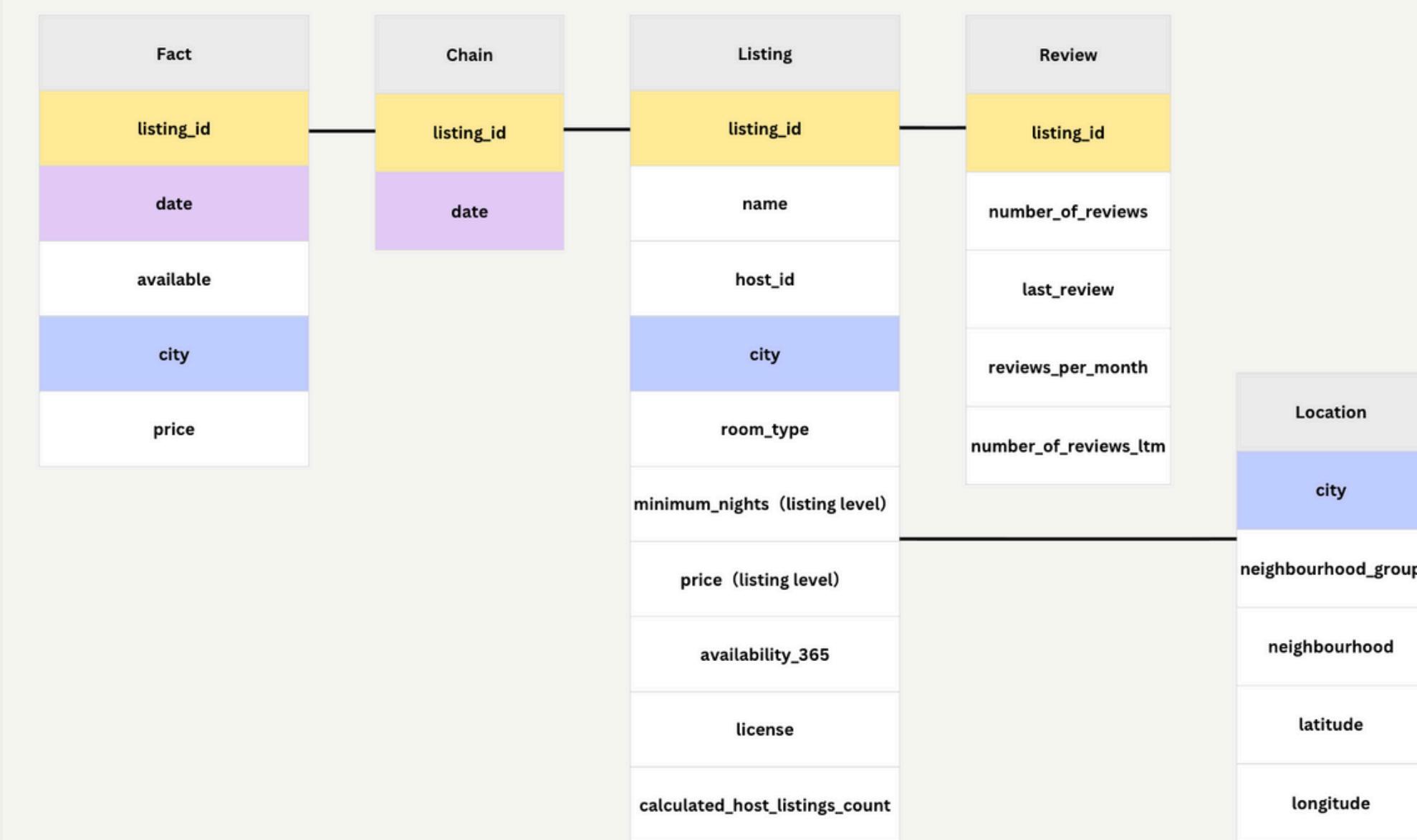
2. Variable Choosing

Avg price
City
Available
Review_per_month



Data Overview

1. After data cleaning - Schema



2. how many columns in this table?

listing.
csv

listing listing_id name host_id room_type minimum_nights price availability_365 license calculated_host_listings_count city

review listing_id number_of_reviews last_review reviews_per_month number_of_reviews_ltm

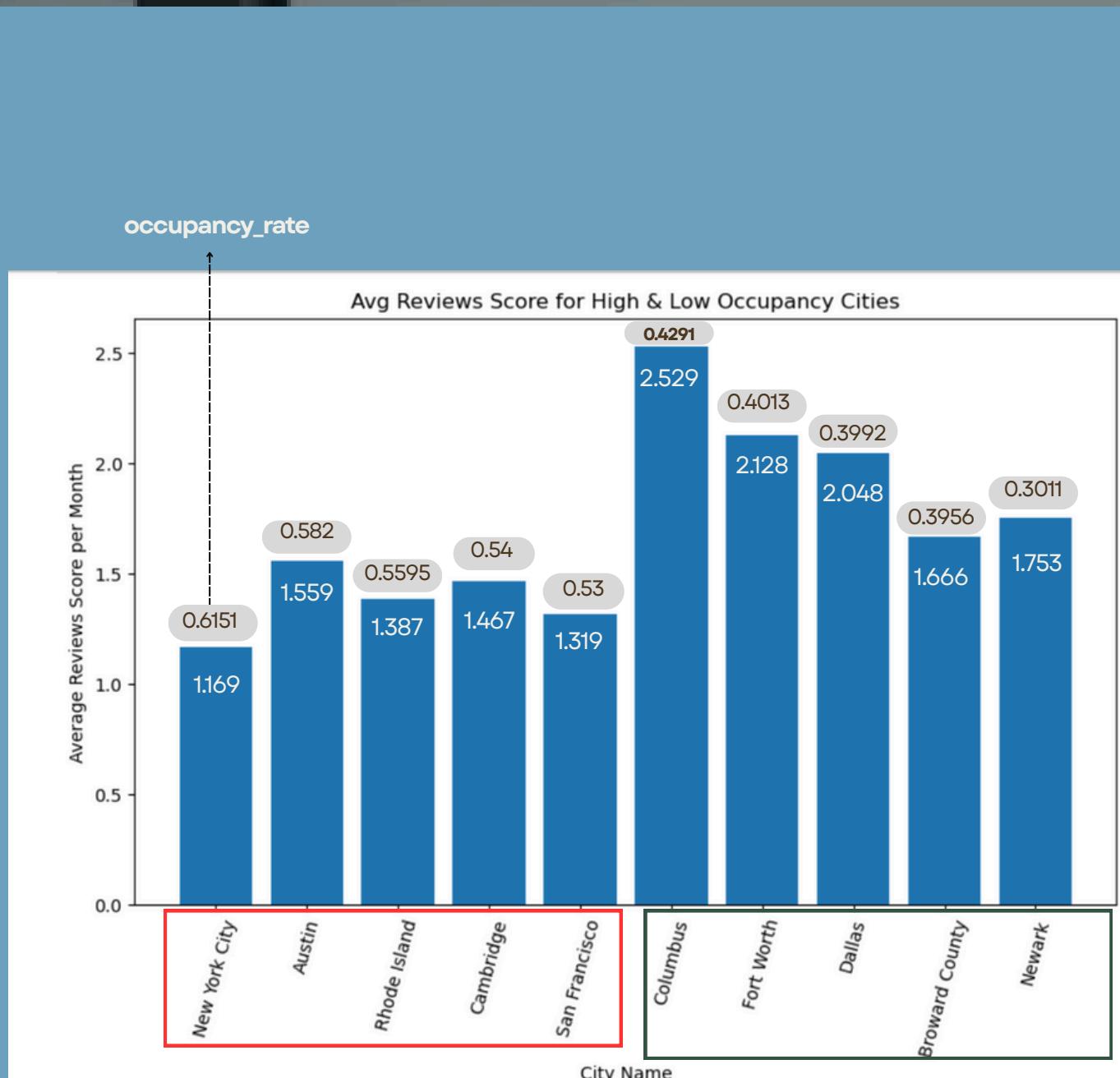
location city neighbourhood_group neighbourhood latitude longitude

calendar
.csv

fact

listing_id date available price city

@MandyChiu



DO GUESTS RATE HIGHLY IN CITIES WITH STRONG DEMAND?

Key Finding

There is no strong relationship between occupancy rate and guest satisfaction.

Supporting Evidence

- **New York City** and **Austin** show the highest occupancy rates (0.6151 and 0.582), yet their average review scores are relatively low (1.169 and 1.559).
- In contrast, **Columbus**, despite a lower occupancy rate of 0.429, achieves the highest average rating (2.529) among the cities analyzed.

Investor Takeaway

High occupancy increases exposure, not review quality. As a result, high-demand markets may generate strong bookings but only moderate or lower average ratings.

Academic research

Mattila (2003) finds that occupancy percentage is not a significant predictor of guest satisfaction, suggesting that demand does not automatically translate into better experiences.

Industry insight

According to AirDNA (2023), higher occupancy exposes listings to a more diverse guest base, increasing review volume and rating variability.

TOP VS BOTTOM REVENUE CITIES: IDENTIFYING HIGH-RETURN MARKETS

key finding

Revenue is highly concentrated in a small number of large, high-demand markets.

Top Revenue Markets	Bottom Revenue Markets
These markets combine high ADRs, strong demand, and large supply scale, resulting in significantly higher total revenue.	Smaller markets generate limited revenue due to lower scale and narrower demand, even when prices are moderate.

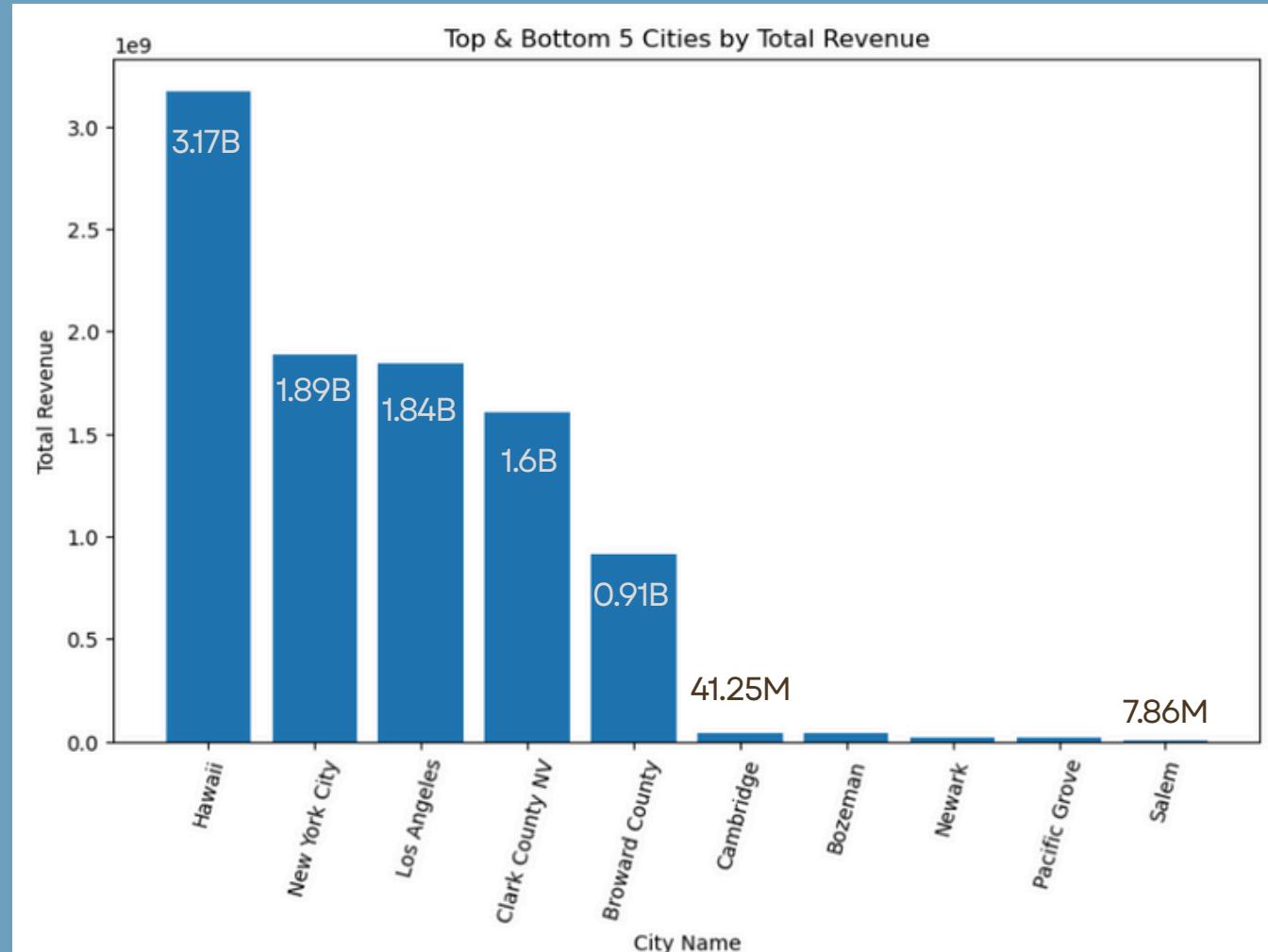
Industry Validation

AirDNA (2023) identifies Hawaii, NYC, Los Angeles, Las Vegas, and South Florida as the largest U.S. short-term rental markets, driven by tourism intensity, event-based demand, and international travel flows.

Business Inside

Revenue = Price × Demand × Supply Scale

High-revenue markets succeed not because of price alone, but due to the interaction of premium pricing, consistent demand, and large listing volume.



NEXT STEPS & LIMITATION

Next steps

- Incorporate regulatory data to adjust revenue and occupancy estimates by city.
- Segment listings by property type and price tier to refine investment insights.
- Analyze seasonality and events to capture demand volatility and peak revenue periods.
- Expand platform coverage by comparing Airbnb with hotels or other rental platforms.



Limitation

Regulatory differences

across cities may affect listings, occupancy, and revenue (e.g., stricter rules in NYC).

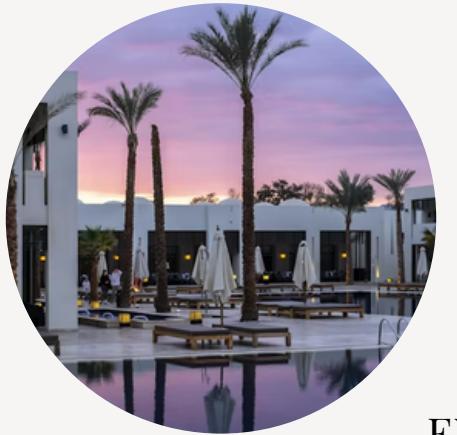
Host behavior varies

by market; pricing strategies and participation levels are not uniform.

platform bias

Airbnb data does not represent the entire lodging market or event-driven supply

CONCLUSION



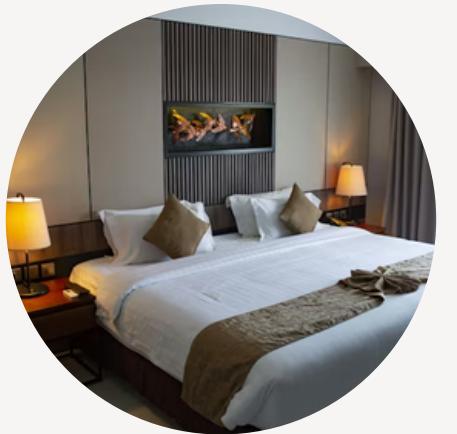
EX: HAWAII

Pricing varies widely, with tourism-driven markets commanding higher prices.

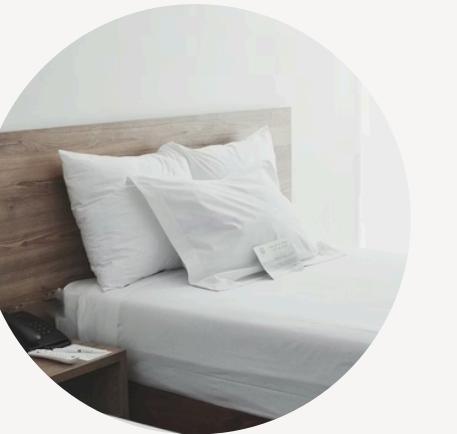


EX: NEW YORK

Demand is concentrated in large metropolitan areas with consistently high occupancy.



High occupancy does not guarantee higher ratings due to review variability.



1. Revenue patterns align closely with industry benchmarks.
2. Top markets combine high prices, strong demand, and large supply scale.

