

San Diego Airbnb Data Visualization

Group12

Cao Marin

Parekh Prashil

Ahmad Younus

Fang Xiaotian

Fuhrman Sean

Objective and application

- Analyze and visualize the Airbnb dataset of San Diego to find the trends that impact price, occupancy, etc.
- These findings will help explore how elements like neighborhood popularity, property features, proximity to the ocean affect the price and overall quality of stays.
- Analyze how pricing varies over time and identify trends that impact guest satisfaction and host performance.



Dataset Overview

- Source: Inside Airbnb San Diego
- Datasets contain 75 columns including Airbnb info, host response time, neighbourhood, # of bathroom/bedroom, ratings, bookings, etc.



<https://insideairbnb.com>

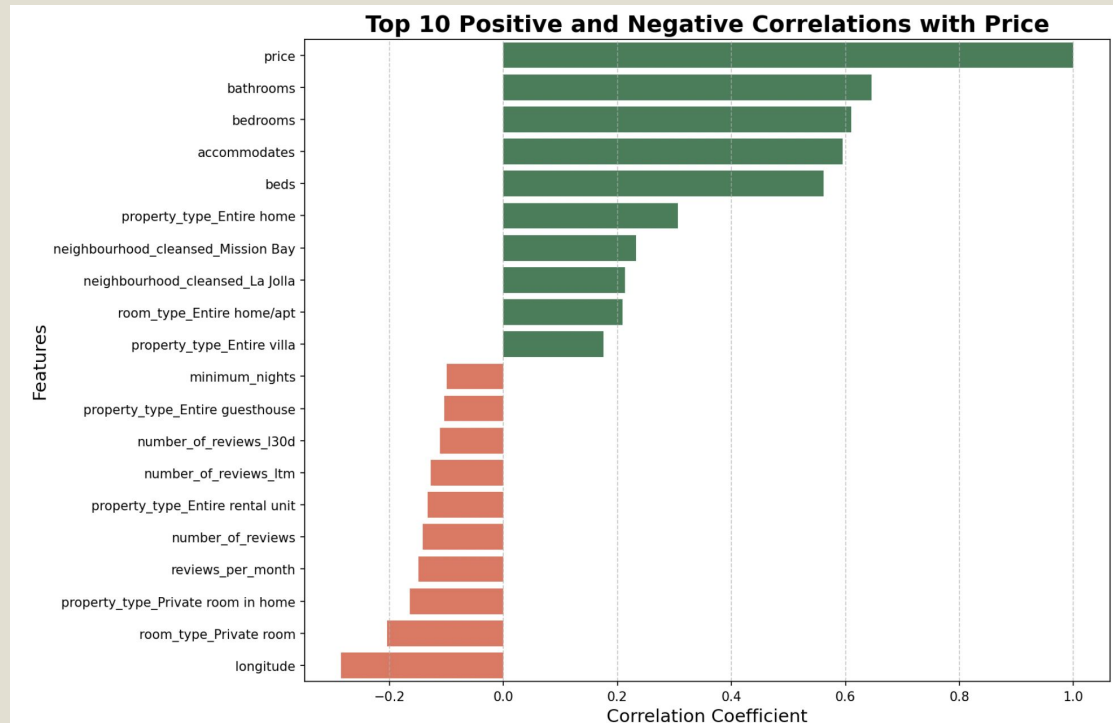
Methodology

- Data extraction and cleaning
To ensure data accuracy and reliability, missing values and inconsistencies were removed and the data was formatted appropriately for analysis.
- Analysis and finding trends
Conducted a comprehensive analysis of key trends affecting pricing, reviews, and occupancy rates. This process focus on the neighborhood level while also examining attributes of the property, such as the environment, locations, and design of the houses.
- Visualization
Developed detailed visualizations, including charts and diagrams, to illustrate pricing trends and their relationships with external elements. These visualizations will show how those factors mentioned above contribute to pricing variations and overall guest satisfaction.

Data Visualization

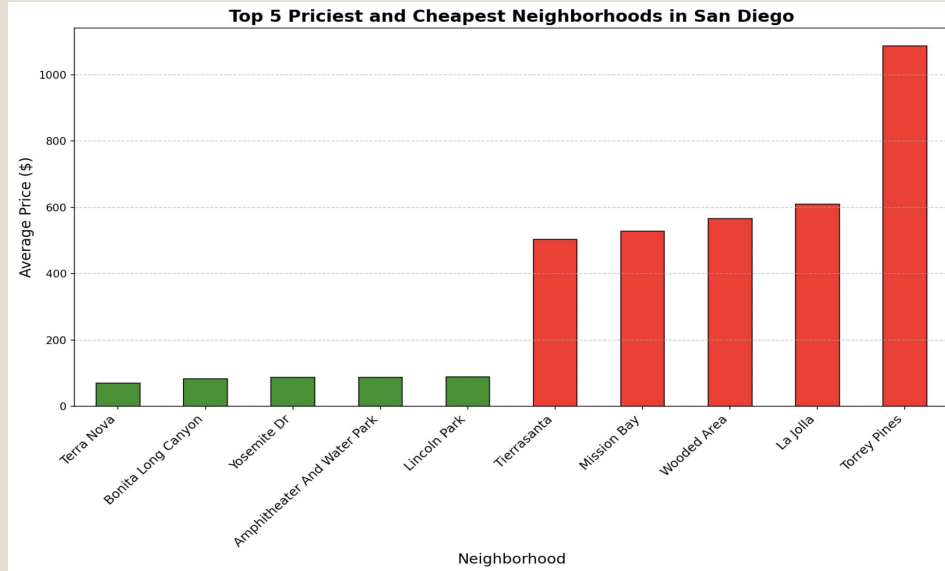


What factors are driving Airbnb prices

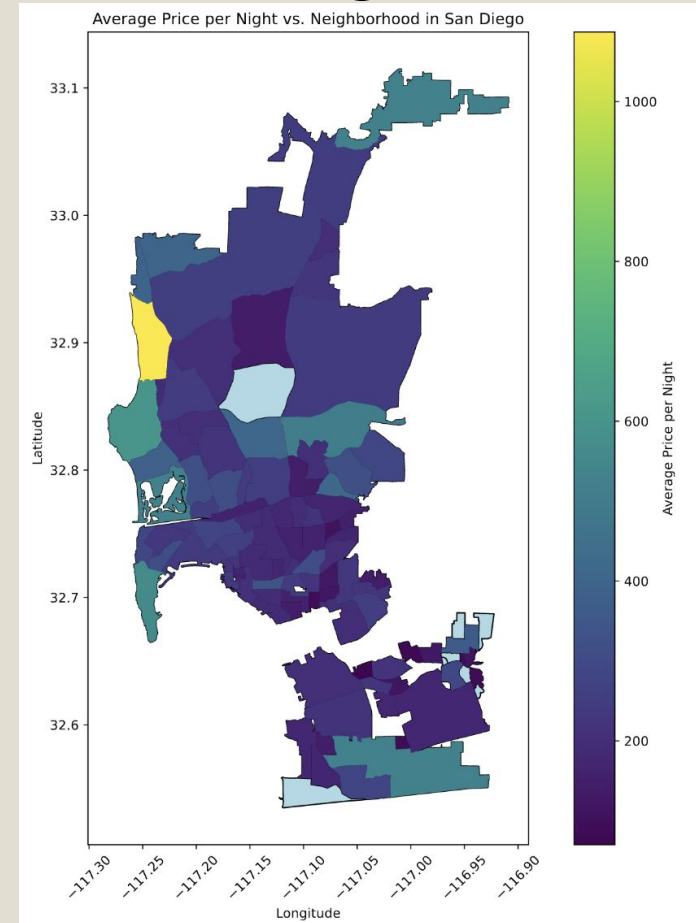


- Positive correlations → price dependency increase with more positive coefficients.
- Negative correlations → price dependency increase with more negative coefficients.
- Factors which are important for driving the price : Bathrooms, bedrooms, accommodates, beds, longitude, beds, private rooms.

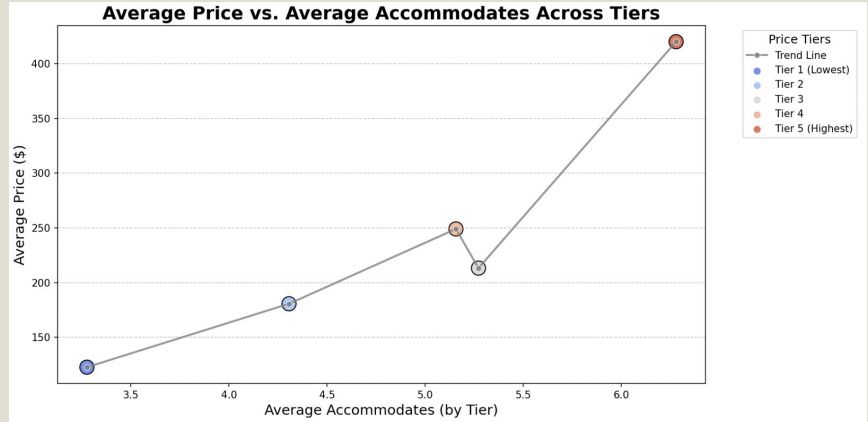
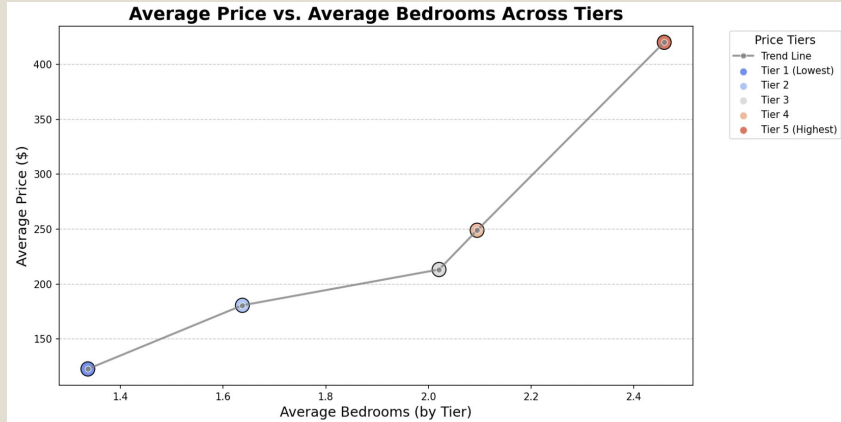
Price vs. Neighborhood analysis in San Diego



- We found most expensive neighborhoods are near the coast: La Jolla, Torrey Pines
- Also some expensive neighborhoods inland → could be due to bigger houses



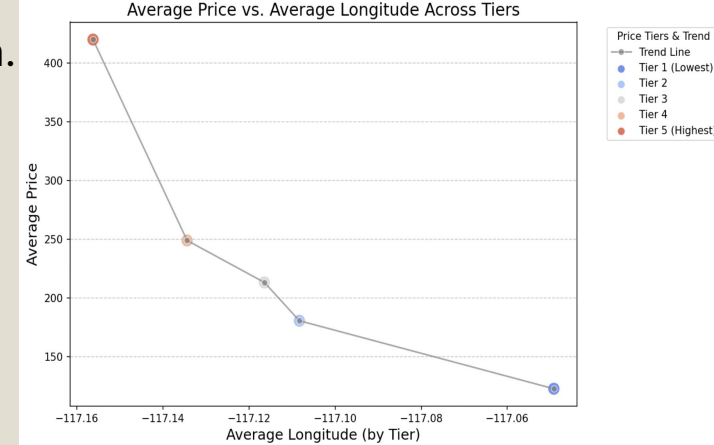
Correlation between size of the house and price



- San Diego neighbourhoods divided into 5 tiers based on the airbnb pricing.
 - Tier 1 → Cheapest, Tier 5 → priciest.
 - Tier 1: about 1.35 average bedrooms
 - Tier 1 accommodate about 3 people.
 - Tier 5 accommodate about 6 people.
- Tier 5: about 2.4 average bedrooms

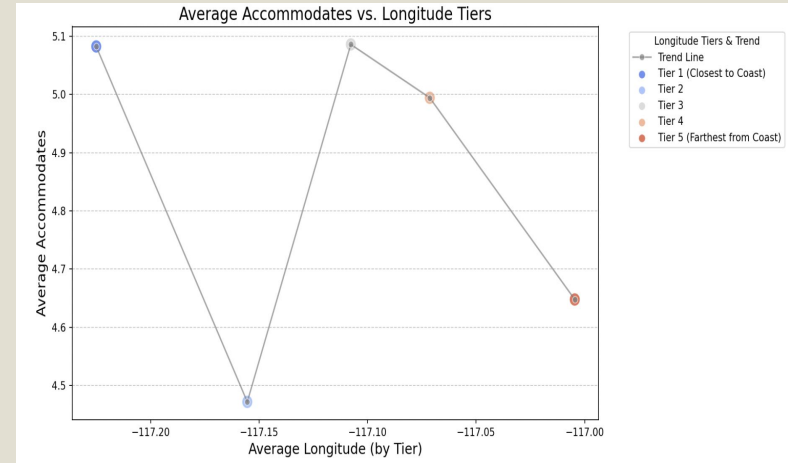
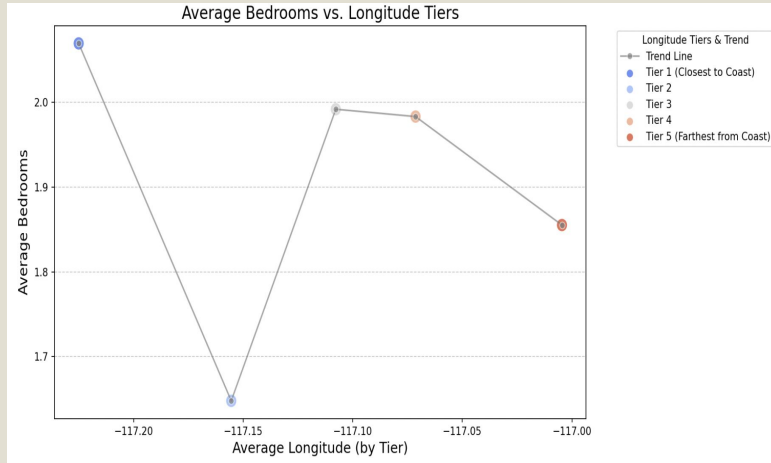
Correlation between size of the house and price

- In all the previous graphs, there is a huge gap in pricing from Tier 4 → Tier 5 (From 200\$ - 250\$ to 400\$).
- Not a lot of area increase from Tier 4 → Tier 5 though.
 - Bedrooms : 2.1 → 2.4
 - Accomodations : 5.25 → 6.25



- Tier 5 airbnbs have a slightly bigger house than Tier 4, but can give you a much appreciated airbnb price. (About 1.6 times !!).
- Tier 5 neighbourhoods : Bay Park, Del Carro, Del Mar.
- But, is this higher price because of ideal accommodations and ideal size of the house or because of its proximity to the ocean ?

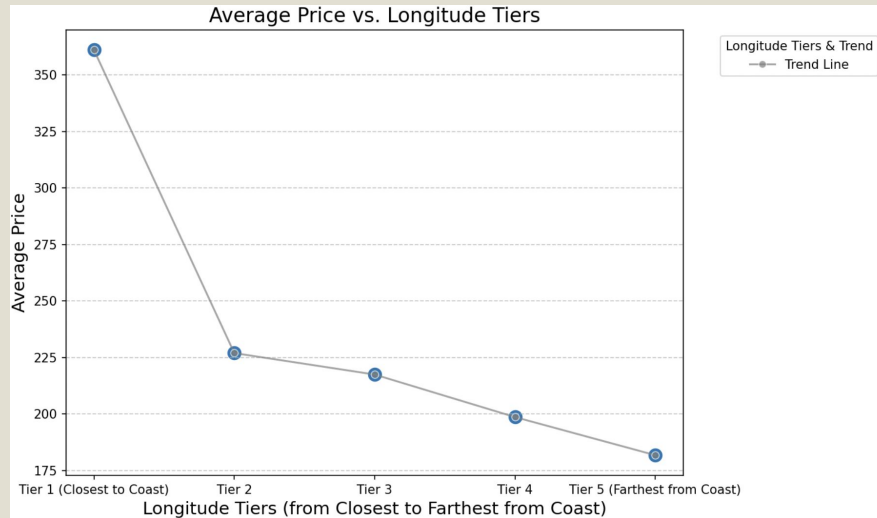
Correlation between size of the house and proximity to the ocean



- San Diego neighbourhoods divided into 5 tiers based on the longitude.
- Tier 1 → Closest to the coast, Tier 5 → Farthest from the coast.
- Tier 1: about 2.2 average bedrooms Tier 5: about 1.85 average bedrooms.
- Tier 1 airbnbs can accommodate about 5.1 people.
- Tier 5 airbnbs can accommodate about 4.6 people.
- Tier 1 airbnbs are closer to the ocean and have bigger houses, but not true for other tiers.

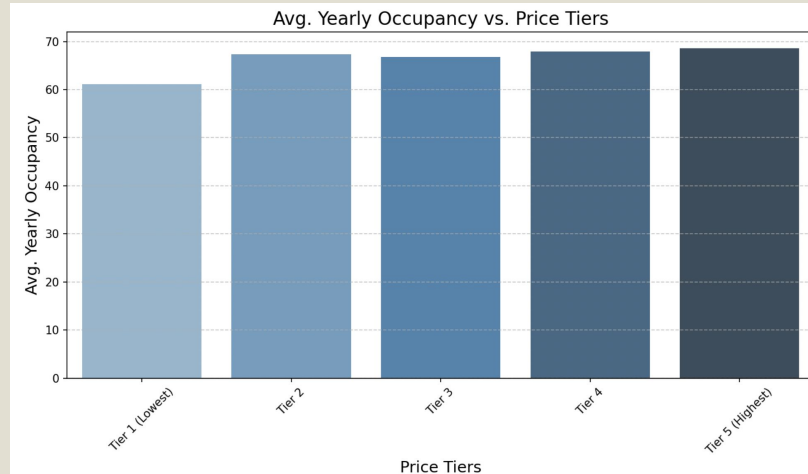
Correlation between size of the house and proximity to the ocean

- Interesting tier is Tier 2.
- From the previous graphs : Tier 2 has the least amount of bedrooms and accommodations.
- Smallest house → Tier 2, but still airbnbs charge more for Tier 2 neighbourhoods compared to Tier 3,4,5 even though they have bigger house.
- Our theory : Although smaller house, it is still close to the ocean.



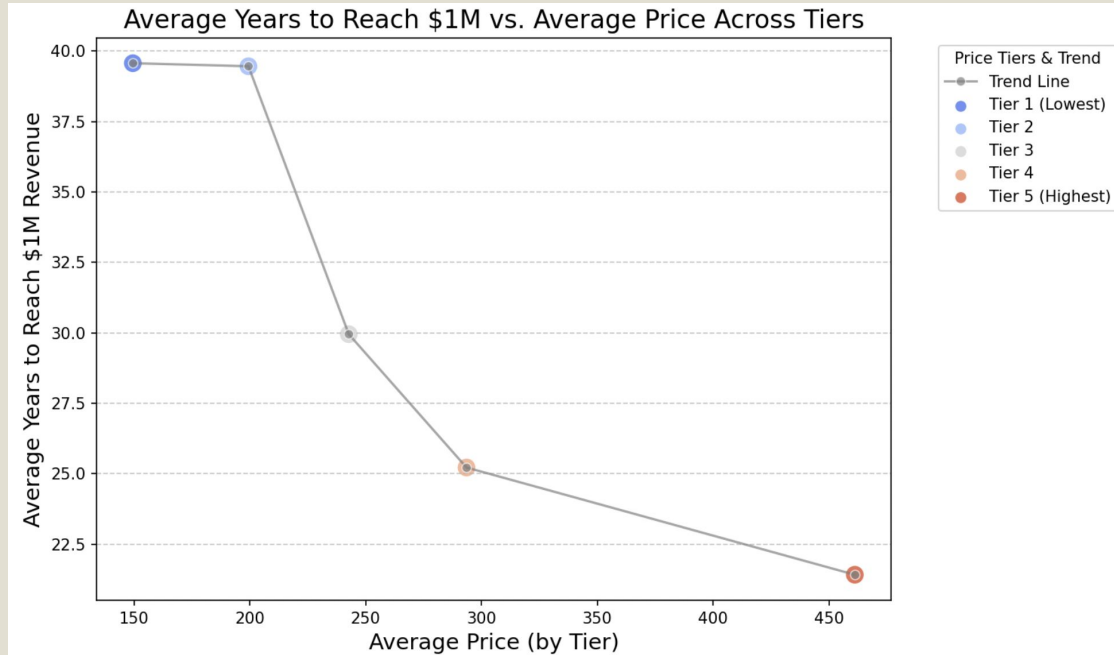
Occupancy of Airbnbs in San Diego

- San Diego neighbourhoods divided into 5 tiers based on the airbnb pricing.
- Tier 1 → Cheapest neighbourhoods, Tier 5 → priciest neighbourhoods.
- All tiers have their average occupancy between 60 - 70%.
- Tier 5 which is the priciest has the highest occupancy rate → High price of airbnbs does not lead to low occupancy for Tier 5 neighbourhoods.
- $\text{Occupancy} = (30 - \text{availability}_{30}) * 1200 / 360$



Years to reach 1M\$ revenue - for entire home as airbnb

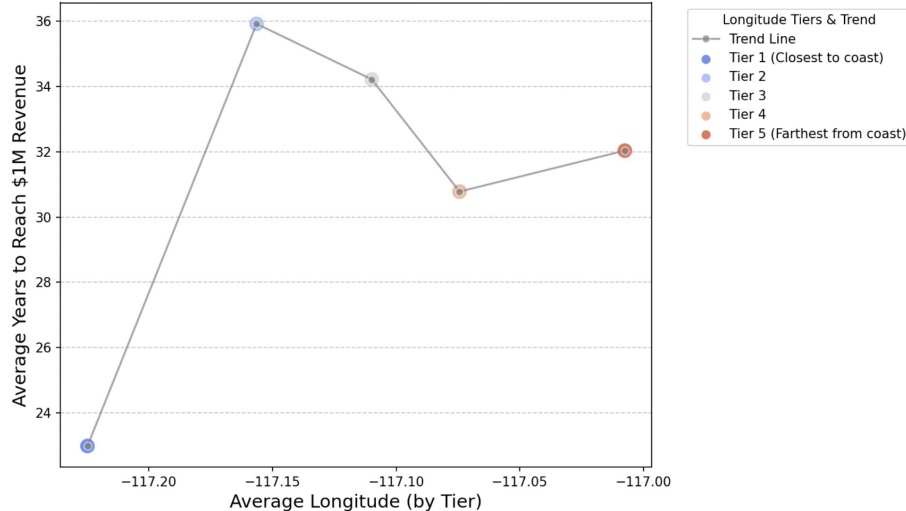
- San Diego neighbourhoods divided into 5 tiers based on the airbnb pricing.
- Tier 1 → Cheapest neighbourhoods, Tier 5 → priciest neighbourhoods.
- Tier 1 airbnbs almost take double the time to reach 1M\$ compared to Tier 5 airbnbs.
- The gap for average years to make 1M\$ revenue decreases as we go to higher tiers.



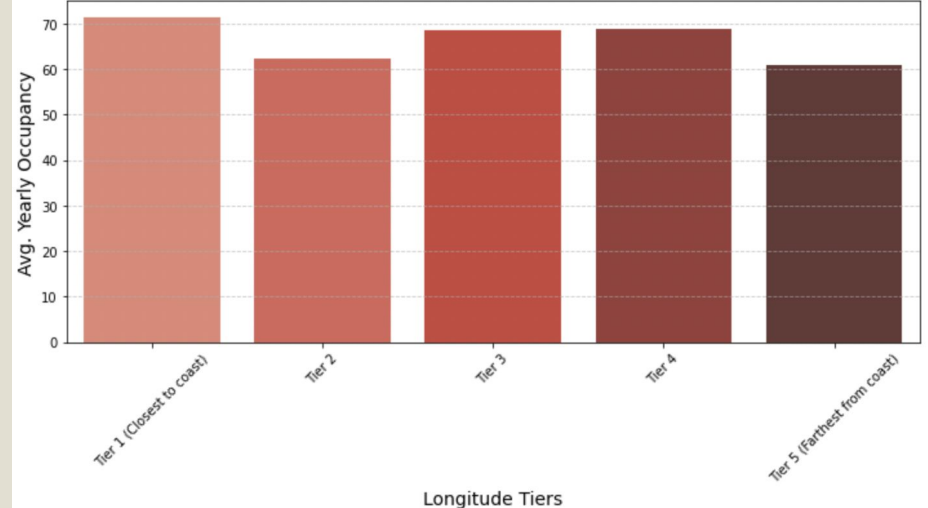
Years to reach 1M\$ revenue - for entire home as airbnb

- San Diego neighbourhoods divided into 5 tiers based on the longitude.
- Tier 1 → Closest to the coast, Tier 5 → Farthest from the coast.
- Tier 2 airbnbs, that had the smallest house from the previous house vs longitude analysis, also takes in maximum amount of time to make 1M\$ in revenue.
- This is because of lower occupancies in tier 2 airbnbs.

Average Years to Reach \$1M vs. Average Longitude Across Tiers



Avg. Yearly Occupancy vs. Longitude Tiers



Thank You