Data 301 Group 38A Final Report

**Our repository can be found here:** [**https://github.com/ubco-W2022T2-data301/project-group-group38**](https://github.com/ubco-W2022T2-data301/project-group-group38) **, along with relevant code and outputs for you to study!**

**Introduction**

Our project focused on the Airbnb listings for the state of Hawaii from this dataset <http://data.insideairbnb.com/united-states/hi/hawaii/2023-03-16/data/listings.csv.gz> . Both members of our group are new to data science and were unsure of what area we wanted to focus our project on. After looking through the datasets provided by Dr. Firas Moosvi in the canvas course shell, a group member expressed an interest in analyzing the listings of Airbnb rentals in Hawaii as he has spent time vacationing on the island and plans to continue visiting. After analyzing the dataset we quickly realized that there were many intriguing aspects and the group decided we would focus our attention on this dataset.

**Question 1:** **Which Island provides the most value based on price and quality of stay?**

To begin my analysis, I needed to find out what the different types of listings were on this dataset so that I could come up with average prices that reflected the type of room that was being listed. I quickly realized that while there are 4 room types in the dataset, two of them (Private rooms and Entire homes and apartments) make up most of the listings.

Chart, bar chart

Description automatically generated

Once I realized this, it became clear that focusing my analysis on these two room types would be most efficient and provide enough information to answer my research question. I also concluded that any data regarding price or quality of stay should be split by these room types to get an accurate picture of the value of each listing type.

Here are the visualizations I created to demonstrate the average prices of both room types on each island:

Chart, bar chart

Description automatically generated

Chart, bar chart

Description automatically generated

Once I had analyzed the price information, I needed a measure of the quality of stay and I discovered that the dataset included information about the reviews of each listing. I decided that using these review scores would be an effective way to uncover the value of each listing type and came up with average review scores.

Here are my figures outlining the average review scores for each island and room type:

Chart, bar chart

Description automatically generated

Chart, bar chart

Description automatically generated

Rather interestingly, the review score data was very similar and suggested to me that most visitors enjoyed their stay regardless of the room type or island, and that perhaps price is the most important factor.

Once this analysis was complete, I created the following report:

**Maui:**

**Private Room**

Average Price: **$606.71**

Average Review Score: **4.76/5**

**Entire home/apt:**

Average Price: **$753.44**

Average Review: **4.82/5**

### ****Honolulu****

**Private Room**

Average Price: **$756.31**

Average Review Score: **4.71/5**

**Entire Home/Apt**

Average Price: **$346.82**

Average Review Score: **4.69/5**

### ****Kauai****

**Private Room**

Average Price: **$485.07**

Average Review Score: **4.78/5**

**Entire Home/Apt**

Average Price: **$474.73**

Average Review Score: **4.75/5**

### ****Hawaii****

**Private Room**

Average Price: **$508.79**

Average Review Score: **4.79/5**

**Entire Home/Apt**

Average Price: **$432.23**

Average Review Score: **4.77/5**

This report was used to answer my research question: **Which Island provides the most value based on price and quality of stay?**

**Answer:** Based on the findings provided above, there is answer for this question, but it is broken down by room type.

Private Room: Kauai

Entire Home/Apt: Honolulu

**Question 2: What are the correlations between bedroom, price, and review score?**

I wanted to see what conclusions I could derive from review score, bedroom, and price data.

Chart, waterfall chart

Description automatically generated

Firstly, I made a heatmap of average review score per island. It is important to note that they are all very high scores, but the smallest of differences could be the most meaningful impacts for a traveller looking for the optimal experience.

Chart, scatter chart

Description automatically generated

Then, I did average review score by room type using this scatter plot. Entire homes were much more commonly listed, as you can tell from analysis done in Question 1, which may have skewed the data for having higher average review scores.

Chart, box and whisker chart

Description automatically generated

Between the islands, it seems as though Maui gets the most amount of high reviews.

Now that we have an understanding between for review scores, I will move on to price and bedroom correlations.

Chart, histogram

Description automatically generated

It can be drawn from this chart that there is a steady increase in price, as one would expect, from an increasing number of bedrooms. It is interesting to note that a 8 bedroom rental property is where one would get significantly more value than a 10 bedroom place. It is also interesting to note that there is a significant price hike when going up to 9 bedrooms from 8 bedrooms, in more than **double** the price.

For travellers in big groups, noting this difference could prove beneficial.

How many places are there by each bedroom count?

Chart, histogram

Description automatically generated

As noted above, 8-bedroom listings are where travellers could get the most value. However, it is unfortunate that they are only available for basically half of the year, and there is so few of them. If travellers are looking to travel in big groups, they must be on the hunt for these few 8-bedroom deals that could be available!

However, they will have to settle with something less than a 5.0 review. As noted above, there are only 4 listings that have 5.0 reviews. Thus, it is not reasonable to expect to get the best of the best having 8 bedrooms.

Chart

Description automatically generated

**Findings**

**What is the average review score on each of the islands?** - I found that the average review score was very similar, all very high. This tells me: - Slight differences in review scores could reveal big differences for the ultimate travelling experience. Due to Hawaii's beautiful nature, reviewers may be skewed to leave better reviews anyhow. Small differences, and 5.0 rated houses, could make all the difference in terms of elevating one's travelling experience.

**Do the number of bedrooms have an influence on average review score, and how much does it affect price?** - It was found that number of bedrooms does not influence the review score, or price very much. Once again, travellers may be skewed towards giving higher reviews, because Hawaii is so beautiful.

**Are there any bargains for a certain number of bedrooms?** - It was found that 8 bedroom places offer significantly more value than any other bedroom size. Although far and few between, as visualized above, finding these places can provide great savings for travellers in large groups.

**Which islands have the best location scores?** - It was found that although only minor changes, "The Big Island" Hawaii, Hawaii had the best average review scores. For travellers looking for the ultimate location experience, The Big Island is where you want to go, offering the highest quality Airbnb's to stay at.

**Conclusion**

In conclusion, there were a lot of interesting data points drawn from this dataset that could provide of significant use for travellers. Which islands, how many bedrooms, quality of airbnbs, location/review scores, price, etc... all provide useful information for finding the best deals and best bargains for travelling to such a beautiful place. It was a lot of fun working with this dataset throughout the term!