Visualizing Airbnb Data: Insights for Travelers and Hosts

Alicia Kaski
501075806
Dr. Sanaz Mohammad Jafari
April 15, 2023

Ryerson University

Contents

Introduction

The COVID-19 pandemic has had a significant impact on the travel industry, with businesses like Airbnb experiencing a sharp decline in demand in early 2020. As many countries implemented travel restrictions to prevent the spread of the virus, the number of bookings on Airbnb plummeted. In response, Airbnb was forced to lay off 25% of its workers (Forbes, 2020) and to implement new policies, such as offering refunds for cancellations and enforcing new cleaning protocols to stay afloat.

However, as travel restrictions have started to ease, the demand for short-term rentals has begun to bounce back. Many people are now choosing to book accommodations on Airbnb as an alternative to traditional hotels, as they offer more space and privacy. Additionally, remote work has become more prevalent, and people are taking advantage of the opportunity to work from different locations, leading to an increase in long-term stays on Airbnb.

The goal of this study was to create informative data visualizations that could be incorporated into Airbnb's website to provide valuable travel insights to both travelers and hosts. Doing so could lead to a higher number of bookings on Airbnb, could lead to better outcomes for both guests and hosts, and could help to rebuild the Airbnb business post-pandemic.

More specifically, I propose that Airbnb add a Travelers Insight Dashboard on each city's homepage to provide travelers general but vital information that they could use to book the Airbnb that would best suit their needs. Including this information on each city's Airbnb homepage would reduce the need for travelers to search for these answers on other webpages which would improve customer experience. The data visualizations in the dashboard could answer questions that travelers often have prior to booking a city, such as:

Research Question 1: How do Airbnb **prices** vary by **weekday**? Is the relationship between day and price the same for all property types?

Research Question 2: How do Airbnb **prices** vary by **month**? Is the relationship the same for all property types?

Research Question 3: How do Airbnb **prices** vary by **location**? Is the relationship the same for all property types?

Additionally, I propose that Airbnb develop a Host Insight Dashboard on their website that would only be visible to Airbnb hosts in a particular city. This dashboard would provide hosts with 1) information about Airbnbs in the same area, helping them make data-driven decisions to improve their listings and stay competitive, and 2) specific feedback about Airbnbs which hosts could use to monitor and improve their guest experience. The dashboard could answer questions that hosts often have when making business decisions about their Airbnb listings:

Research Question 4: Which **amenities** have the most impact on Airbnb price?

Research Question 5: What information can Airbnb **reviews** and **ratings** give to hosts?

The tool we used to perform this study was Python, and the packages we used included plotly, matplotlib, seaborn, NLTK, and Wordcloud. The code can be found at: https://github.com/alakaski/DS8007_Data_Visualization_Project.git.

Data Description

Publicly Available Dataset:

https://www.kaggle.com/datasets/airbnb/seattle?resource=download

The Seattle Airbnb Open Dataset contains data from 2016 about 3818 Airbnb listings in Seattle.

The dataset includes 92 columns describing each listing (location, ratings, reviews, host information, property characteristics, amenities, etc.).

It contained 3 files: Listings.csv, Calendar.csv, and Reviews.csv.

Listings Data

The listings.csv file contains 92 columns and 3818 rows. Each row represents one of the 3818 Airbnb listings in Seattle in 2016, and the columns contain information about each listing.

We cleaned the dataset to only include 23 of the most relevant columns to this project. The columns can be grouped into three categories:

- Property information (Listing ID, Name, Property Type, Room Type', Number of Bathrooms, Number of bedrooms, Number of beds, Bed Type, Amenities Included, Number of People Accommodated, Cancellation Policy Type, Price per Night)
- Location information (Neighbourhood, Latitude, Longitude)
- Rating scores (Overall Rating, Accuracy, Cleanliness, Check-in, Communication, Location, Value)

Table 1: Listings Dataset Column Data Types

Column #	Column Name	Data Type
0	id	Numerical
1	name	Categorical
2	neighbourhood_cleansed	Categorical
3	neighbourhood_group_cleansed	Categorical
4	latitude	Numerical
5	longitude	Numerical
6	accommodates	Numerical
7	property_type	Categorical
8	room_type	Categorical
9	bathrooms	Numerical
10	bedrooms	Numerical
11	beds	Numerical
12	bed_type	Categorical
13	amenities	Categorical
14	cancellation_policy	Categorical
15	review_scores_rating	Numerical

16	review_scores_accuracy	Numerical
17	review_scores_cleanliness	Numerical
18	review_scores_checkin	Numerical
19	review_scores_communication	Numerical
20	review_scores_location	Numerical
21	review_scores_value	Numerical
22	price	Numerical

Calendar Data

The calendar.csv file contains 4 columns and 1393570 rows. Each row represents one day in 2016 for a specific Airbnb listing, and the columns contain the availability and price of each listing on each day. We kept all columns in this dataset.

Table 2: Calendar Dataset Column Data Types

Column #	Column Name	Data Type
0	listing_id	Numerical
1	date	Date
2	available	Categorical
3	price	Numerical

Reviews Data

The reviews.csv file contains 84849 rows and 6 columns. Each row contains information about one text review for a specific listing. We kept all columns in this dataset.

Table 3: Reviews Dataset Column Data Types

Column #	Column Name	Data Type
0	listing_id	Numerical
1	id	Numerical
2	date	Date
3	reviewer_id	Numerical
4	reviewer_name	Categorical
5	comments	Text

Exploratory Data Analysis

Exploratory Data Analysis is a critical step in any data analysis project. It involves examining and visualizing the data to uncover patterns, relationships, and potential outliers. Through EDA, we can gain insights into the data, identify potential issues, and inform our approach to further analysis. In this section, we explore the Seattle Airbnb dataset through various visualizations and statistical methods to gain a better understanding of the data and its characteristics.

Listings Data

Numerical Attributes

Below, Tables 4-5 and Figures 1-2 provide descriptive statistics of the numerical columns of the Listings.csv dataset.

Table 4: Descriptive Statistics for Numerical Columns in the Listings Dataset (Part 1)

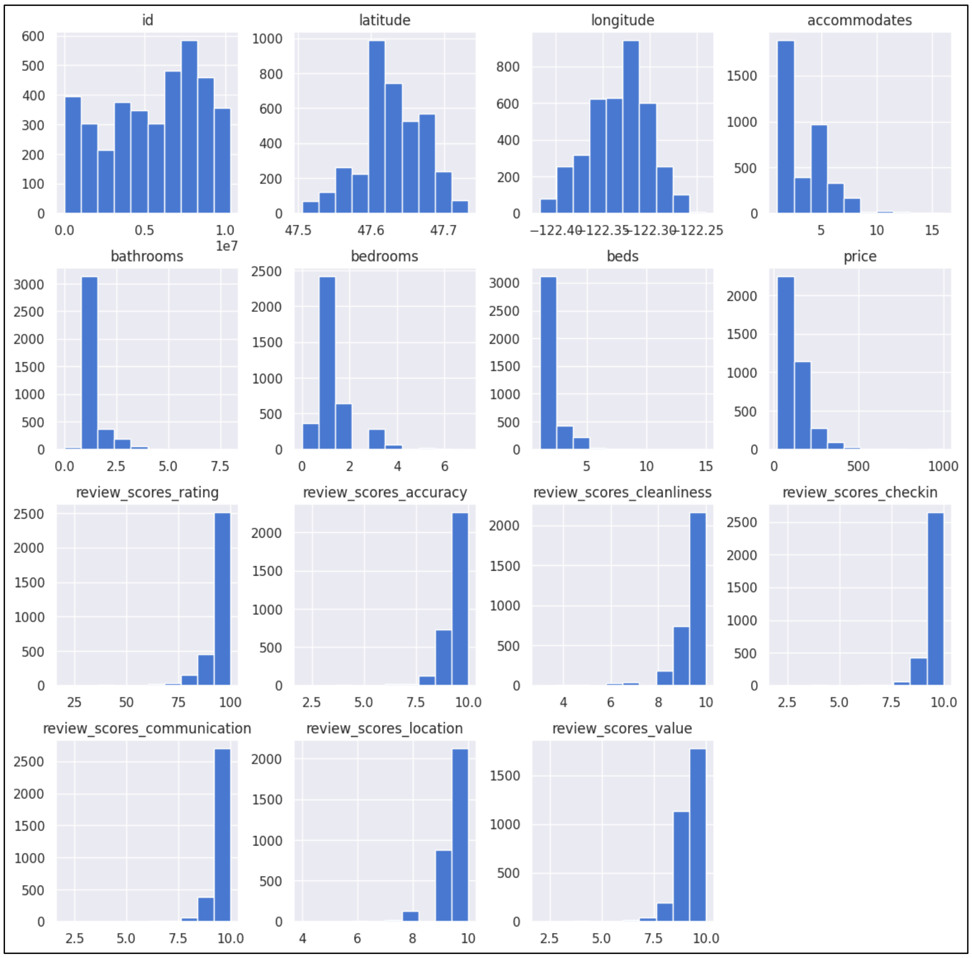
	id	latitude	longitude	accommodates	bathrooms	bedrooms	beds	price
count	3818	3818	3818	3818	3802	3812	3817	3818
mean	5550111	47.63	-122.33	3.35	1.26	1.31	1.74	127.98
std	2962660	0.04	0.03	1.98	0.59	0.88	1.14	90.25
min	3335	47.51	-122.42	1	0	0	1	20
25%	3258256	47.61	-122.35	2	1	1	1	75
50%	6118245	47.62	-122.33	3	1	1	1	100
75%	8035127	47.66	-122.31	4	1	2	2	150
max	10340165	47.73	-122.24	16	8	7	15	1000
variance	8.78E+12	0.00	0.00	3.91	0.35	0.78	1.30	8145.07
skewness	-0.31	-0.25	-0.21	1.70	2.91	1.54	2.48	3.11
kurtosis	-1.10	-0.13	-0.40	4.38	13.87	3.74	10.91	16.62

Table 5: Descriptive Statistics for Numerical Columns in the Listings Dataset (Part 2)

	review_scores_rating	review_scores_accuracy	review_scores_cleanliness	review_scores_checkin	review_scores_communication	review_scores_location	review_scores_value
count	3171	3160	3165	3160	3167	3163	3162
mean	94.54	9.64	9.56	9.79	9.81	9.61	9.45
std	6.61	0.70	0.80	0.60	0.57	0.63	0.75

min	20	2	3	2	2	4	2
25%	93	9	9	10	10	9	9
50%	96	10	10	10	10	10	10
75%	99	10	10	10	10	10	10
max	100	10	10	10	10	10	10
variance	43.64	0.49	0.64	0.35	0.32	0.40	0.56
skewness	-2.92	-3.20	-2.55	-4.97	-5.63	-1.91	-2.07
kurtosis	15.53	18.33	9.19	39.56	51.84	5.62	8.71

Figure 1: Histograms of Numerical Column Distributions



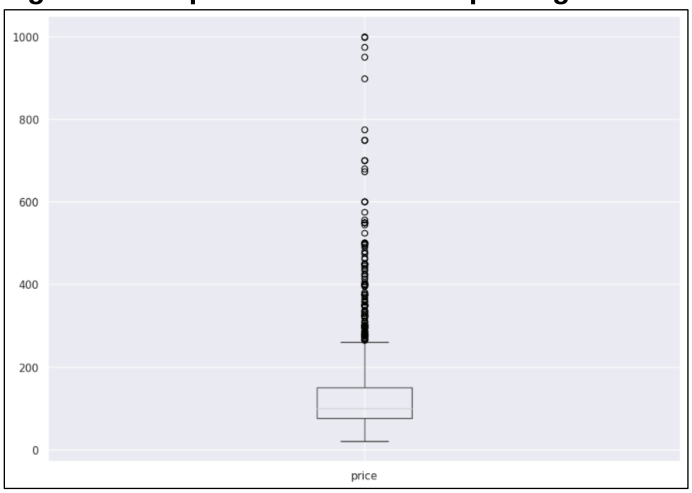
As seen in Table 4 and Table 5, some of the columns have null values, however, because there were so few null values, we kept all of the rows and dealt with the null values differently for specific tasks.

Given the statistics in Table 4 and the histograms in Figure 1, we can see that Listing IDs are evenly distributed between 3335 and 10340165. Latitude (skew = -0.25, kurt = -0.13) and Longitude (skew = -0.21, kurt = -0.40) are normally distributed around the mean coordinate (47.60, -122.33).

The columns showing the Number of People Accommodated, Number of Bathrooms, Number of Bedrooms, and Number of Beds are positively skewed with skewness values of approximately +2.00, indicating that their distributions have a greater number of smaller values. The columns all also have high kurtosis values of over +3.00, indicating that their distributions are more peaked than normal. The average number of people accommodated per Airbnb is 3.35, and the average number of bathrooms, bedrooms, and beds are close to 1.00.

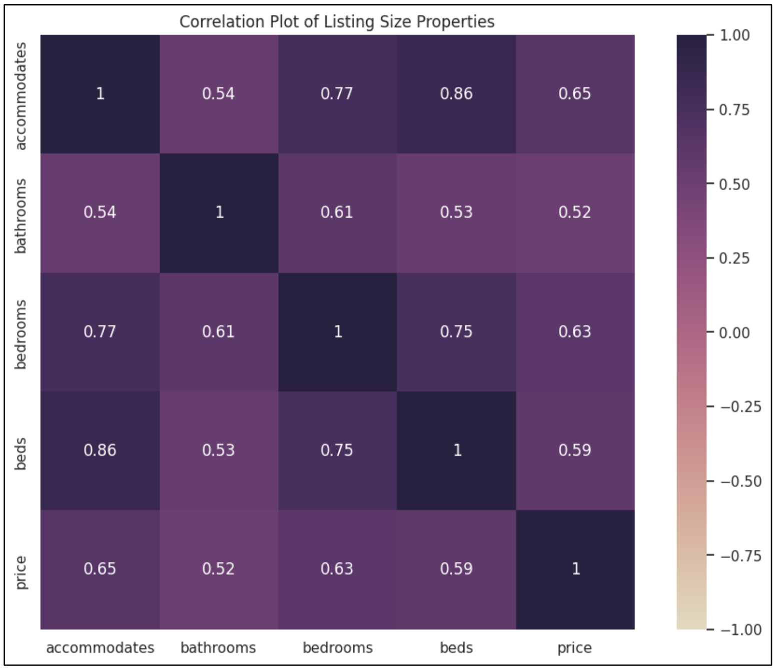
By contrast, Table 5 shows that the columns showing review scores ratings are negatively skewed with skewness values of less than -2.00, and have high kurtosis values of over +5.00. This is because the variation of ratings that travelers give is quite small, with a mean rating of over 9.00 for all rating categories, and a standard deviation of approximately 0.60. In other words, few travelers tend to give low ratings.

Figure 2: Boxplot of Airbnb Price per Night in Listings.csv Dataset



Given the importance of the Price column, we decided to take an in-depth look the distribution of its values in Figure 2. The mean price per night for an Airbnb in Seattle is $127.98, and the median price per night is $100.00. The boxplot shows us that the Price column is positively skewed (skew = 3.11) and that the distribution has high kurtosis (kurt = 16.62). The position of the top whisker is at $262.50/night, which was calculated as 1.5 * IQR (IQR = Q3 - Q1). As a result, 246 listings out of the 3818 total listings are considered outliers because their price is higher than the top whisker value. We removed the outliers before creating the data visualizations so that the visualizations would not be sensitive to the outliers and so that the visualizations showed data that was most representative of the most common Airbnbs that travelers would like to know about.

Figure 3: Heatmap of Correlations between Numerical Attributes Related to Listing Size



In Figure 3, we can see that all of the numerical attributes related to listing size (Accommodates, Bathrooms, Bedrooms, Beds) are positively correlated with each other (r^2 values > +0.50). This is unsurprising, given the nature of the information.

Figure 4: Heatmap of Correlations between Numerical Attributes about Airbnb Ratings

In Figure 4, we can see that the columns related to Airbnb ratings are all positively correlated with each other, however the strengths of the relationships do not appear to be large. This may be due to the fact that the data has a small range, with most ratings being between 9-10. This means that there is limited variability in the data, which can lead to lower correlation scores, even if there is a strong relationship between the variables as one would expect.

Categorical Attributes

In the below plots, we explore the categorical variables in the listings.csv dataset.

Figure 5: Bar Plot of Neighborhood Values

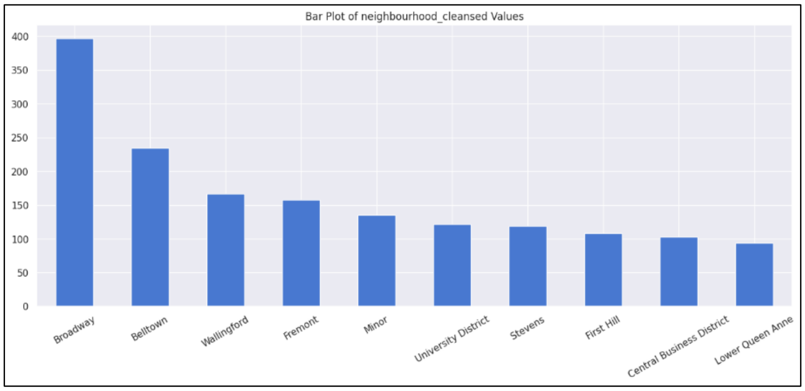


Figure 5 shows that the most common neighborhood for Airbnbs in Seattle is Broadway, and the least common neighborhood is Lower Queen Anne.

Figure 6: Bar Plot of Property Type Values

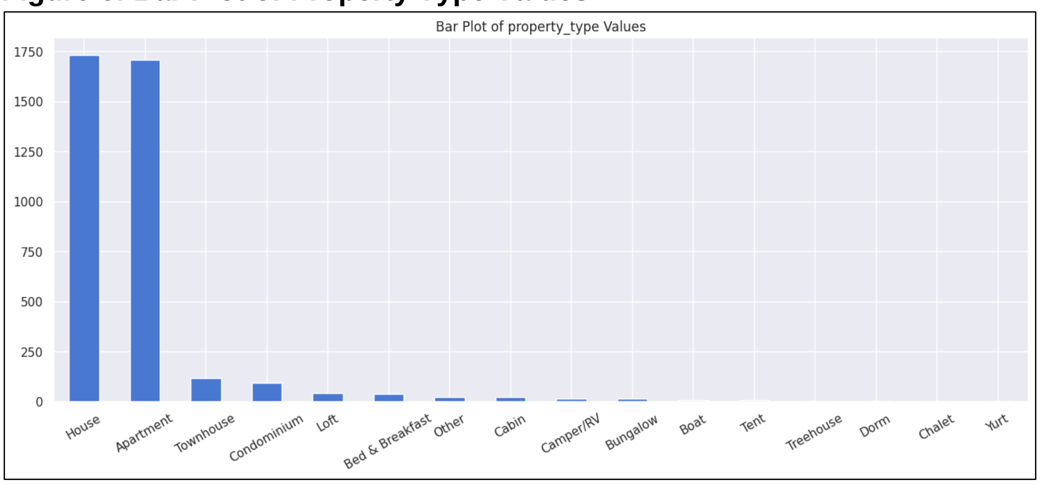


Figure 6 shows that the most common property types for Airbnbs in Seattle are houses & apartments, by a large margin, as there are over 1500 of each type of property. The next most common property type is a townhouse, however the frequency of this value is less than 250.

Figure 7: Bar Plot of Room Type Values

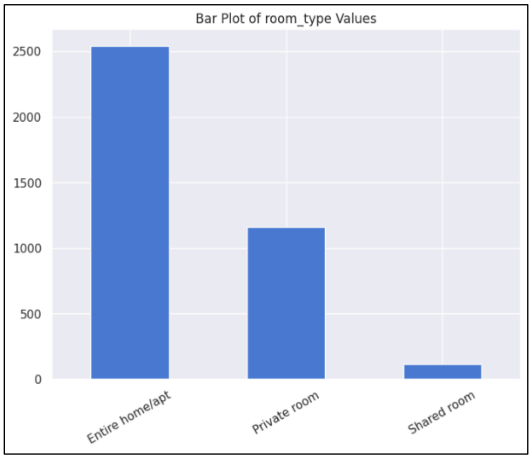


Figure 7 shows that the most common room type is the entire home or apartment, and the next most common type is a private room. The least common room type is a shared room.

Figure 8: Bar Plot of Bed Type Values

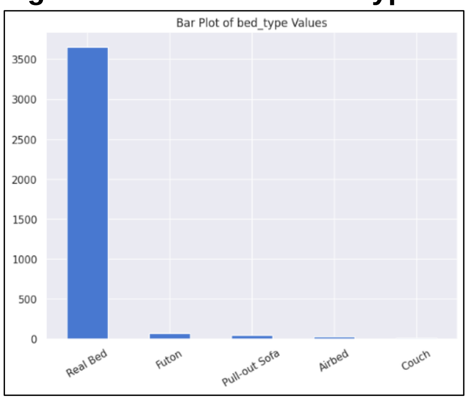


Figure 8 shows that the most common bed type by far is a real bed, as opposed to a futon or pull-out sofa.

Figure 9: Bar Plot of Cancellation Policy Values

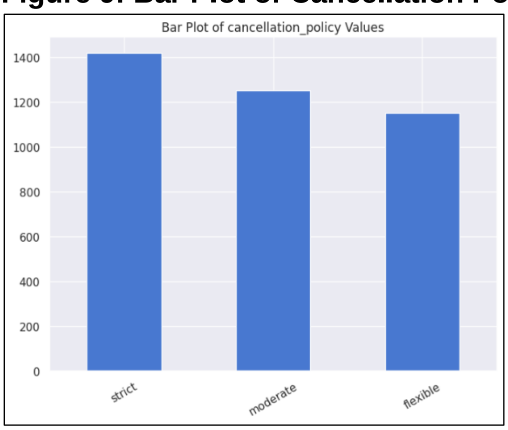


Figure 9 shows that Airbnbs are split relatively evenly across cancellation policy types (strict, moderate, or flexible), however most cancellation policies are strict.

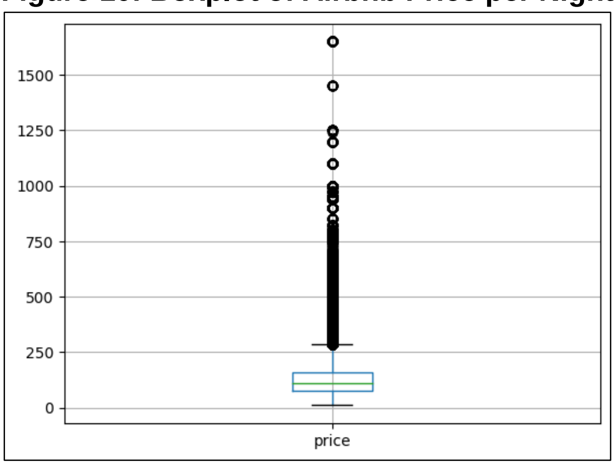
Calendar Data

For the calendar.csv dataset, we only used the ID, Date, and Price columns. The IDs have the same distribution as the listings.csv dataset, and the Date column shows every day in 2016 for every listing. The Price column differs slightly from the Price column in the listings.csv dataset because this dataset has specific listing prices per night, while the listings.csv dataset has the average price per night per listing over the whole year. Once again, as seen in Table 6 and Figure 10, the Price column was positively skewed and contained outliers. The outliers were removed prior to creating data visualizations.

Table 6: Descriptive Statistics of Price per Night for the Calendar.csv dataset

	Price
count	934542
mean	137.94
std	105.06
min	10.00
25%	75.00
50%	109.00
75%	160.00
max	1650.00

Figure 10: Boxplot of Airbnb Price per Night in Calendar.csv Dataset



Reviews Data

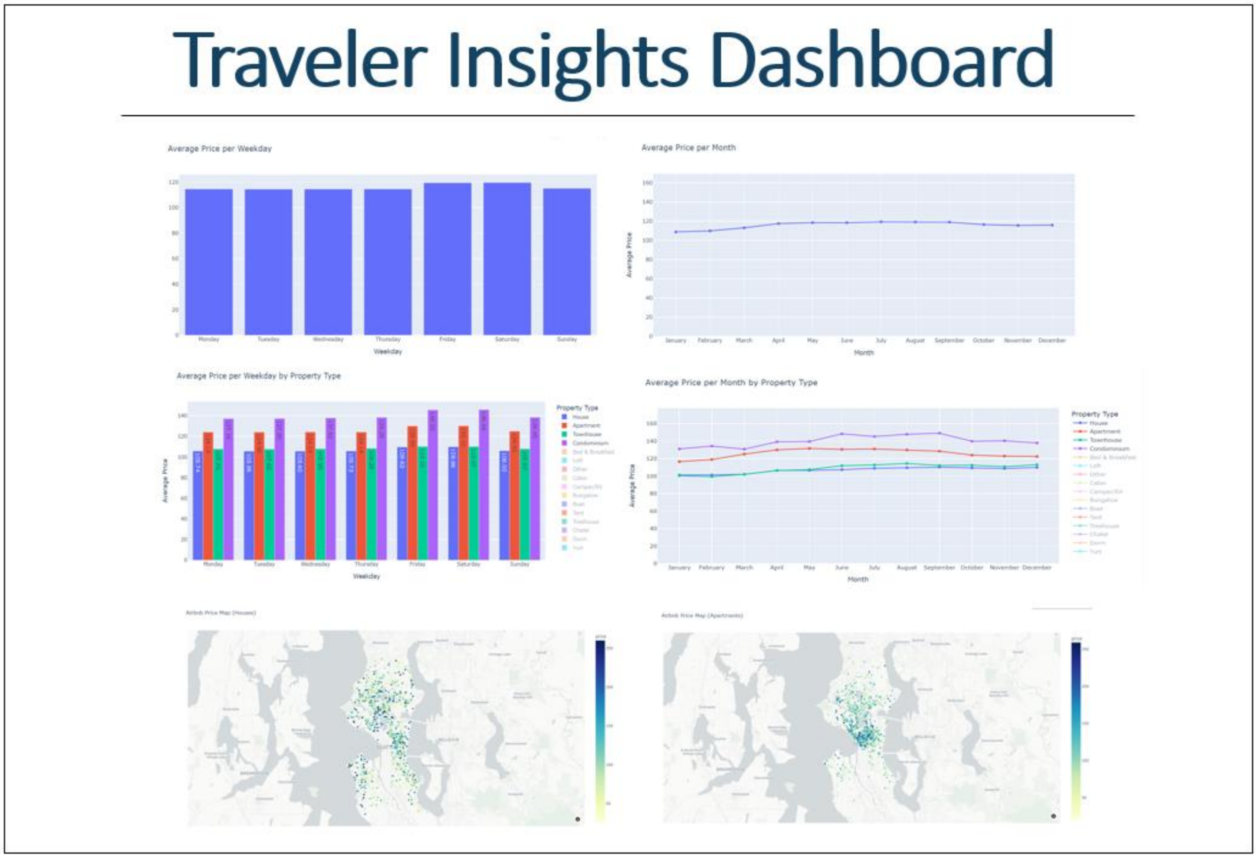
For the reviews.csv dataset, we only used the Listing ID and Comments columns. The Listing IDs had the same distribution as the listings.csv dataset, and the Comments column contained reviews of certain listings on certain days in 2016.

Results

Traveler Insights

For this project, we had five research questions to investigate using data visualizations. Three of the questions were relevant to travelers. To answer those questions, we created 6 data visualizations and combined the visualizations into one dashboard called the Traveler Insights Dashboard. The three questions answered in the Traveler Insights Dashboard are below.

Figure 11: Traveler Insights Dashboard

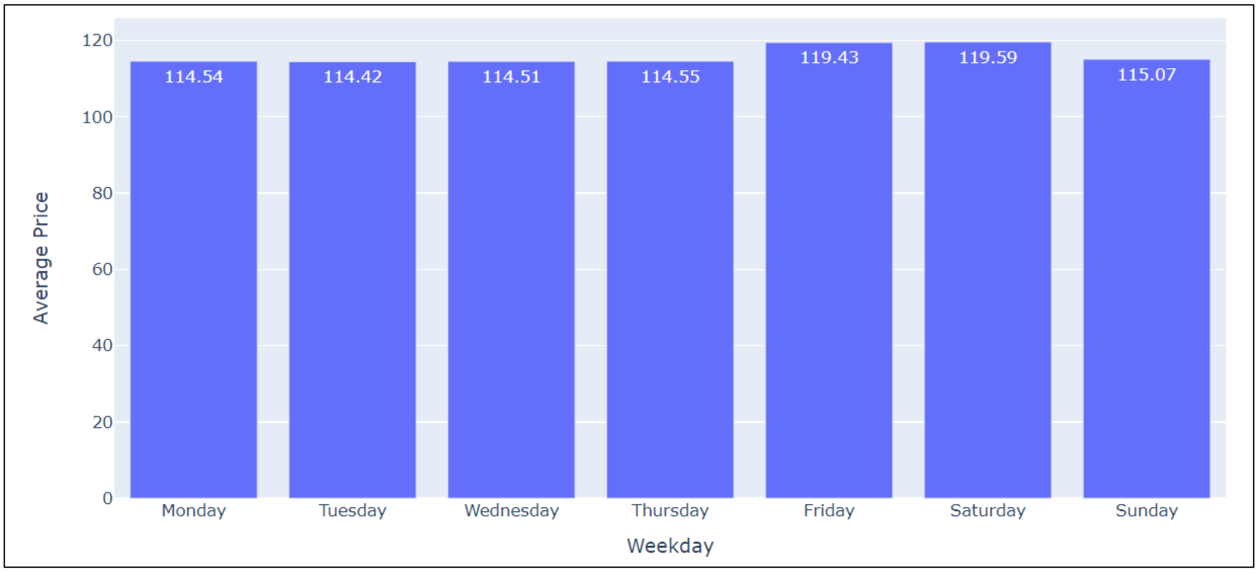


Research Question 1

How do Airbnb **prices** vary by **weekday**? Is the relationship between day and price the same for all property types?

We answered this question by creating **two interactive bar plots** using plotly.graph_objects. We decided to use bar plots to visualize this data because bar plots are easy to understand and interpret, making them a popular choice for visualizing categorical data. Moreover, they can be used to compare the frequencies of different categories and show trends over time. We also decided to make one of the bar plots interactive because interactivity allows users to explore and interact with data, giving them greater control over the visualization and enabling them to delve deeper into the data.

Figure 12: Average Airbnb Price per Night per Weekday



Looking at Figure 12, we can see that the average Airbnb price per night on weekdays in Seattle regardless of property type is $114/night, and the average price on weekends (Friday & Saturday night) is $119/night. So, although there is a difference in prices per night between weekdays and weekends, the difference is much smaller than was expected.

Figure 13: Average Airbnb Price per Night per Weekday by Property Type

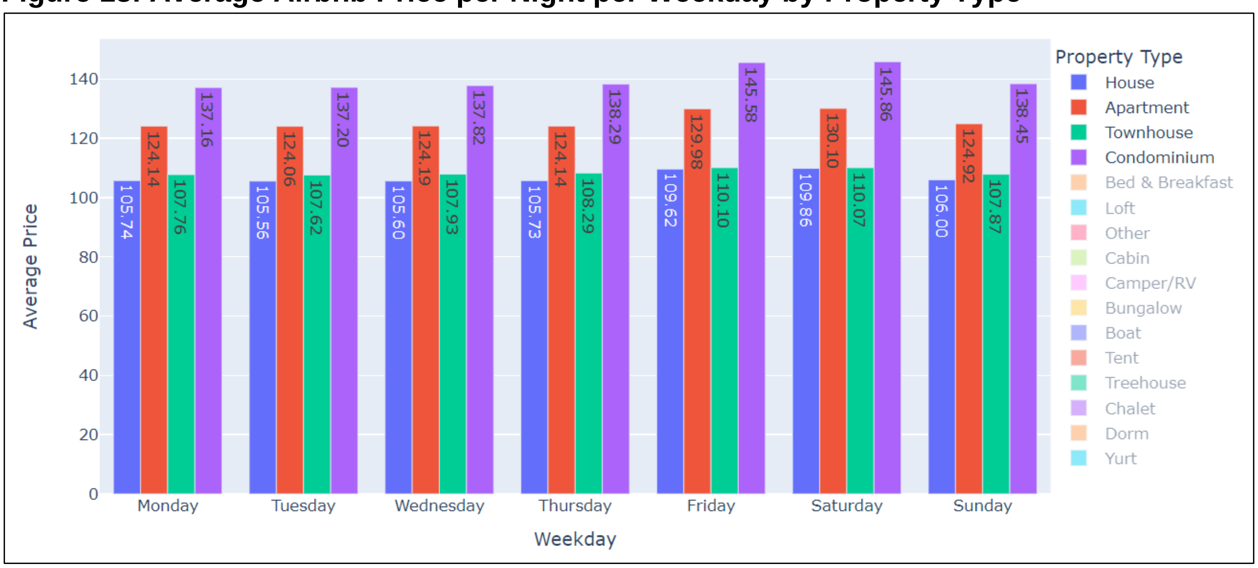


Figure 13 is an interactive plot that travelers can use to see prices per weekdays for the property types they select using the legend. In this picture, the top 4 most common property types are toggled, but travelers could select anywhere between 1-16 property types.

This plot not only allows travelers to look at differences in prices for days of the week within property types (ex. apartments in red cost $124/night on weekdays and $130/night on weekends), but also allow travelers to compare prices between property types (ex. houses in blue cost $105/night on weekdays and townhouses in green cost $108/night on weekdays).

This graph shows that the difference in Airbnb price between weekdays and weekends is relatively small for all property types (<$10 difference between weekdays and weekends). This can further be seen in Table 7 below.

Table 7: Summary of Findings from Figure 13

Property Type	Average Price on Weekdays	Average Price on Weekends	Price Difference
Houses	$105	$110	$5
Apartments	$124	$130	$6
Townhouses	$107	$110	$3
Condos	$137	$145	$8

In summary, Airbnb prices are slightly higher on weekends (Friday nights and Saturday nights) than on weekdays, and the relationship between weekday and price price is the same for all property types.

Research Question 2

How do Airbnb **prices** vary by **month**? Is this the same for all property types?

We answered this question by creating **two interactive line graphs** using plotly.graph_objects. We chose to visualize this data with line graphs because line graphs are ideal for showing trends or changes over time, as they can highlight the direction and magnitude of changes in a variable. This choice also allowed us to differentiate the graphs from the bar plots of the previous question.

Figure 14: Average Airbnb Price per Night per Month

Figure 14 shows the average Airbnb price per month in the year 2016. Travelers can hover over datapoints to see more detailed prices per month. What we can see in this graph is that the price per night is relatively stable over the year, which is different than what was expected. We expected to see a larger difference between summer months and winter months, but the graph shows that Airbnbs cost an average of $120/night in Spring & Summer months, and $110/night in Fall & Winter. Therefore, month does not have a large impact on the price of Airbnbs in Seattle.

Figure 15: Average Airbnb Price per Night per Month by Property Type

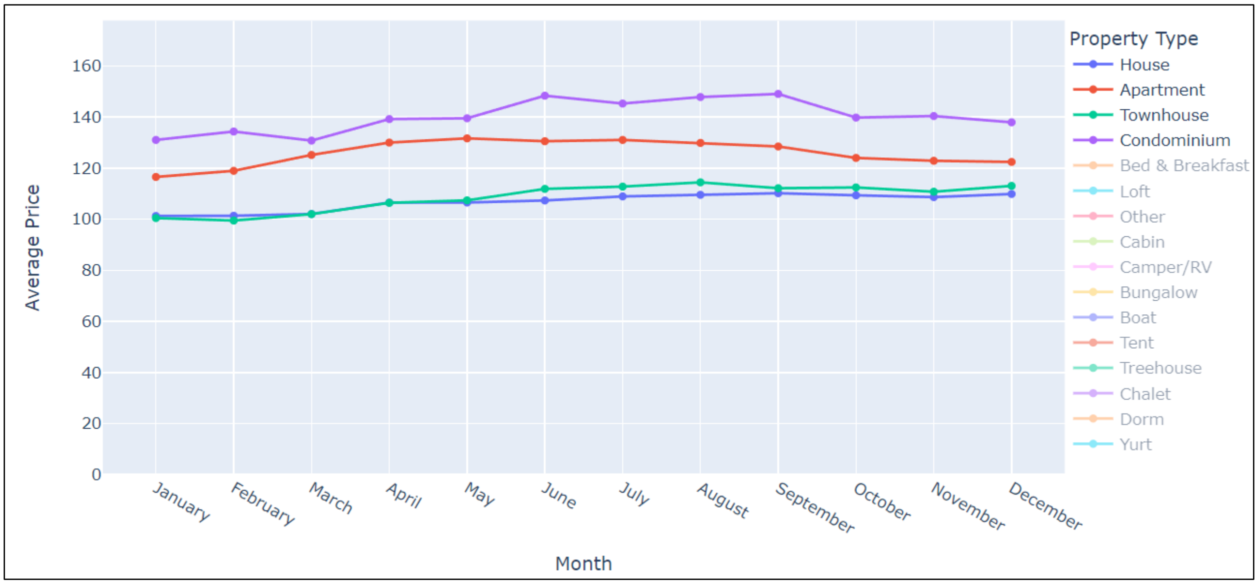


Figure 15 shows the average Airbnb price per night per month for each property type. Travelers can select which property types they would like to display on the graph by selecting one of the options in the legend on the right.

Interestingly, this graph shows that condos (purple) & apartments (red) show more variation in prices over the year than houses (blue) and townhouses (green). In other words, condos and apartments have bigger differences in prices between Summer and Winter months than houses & townhouses. House & townhouse prices are more stable throughout the year.

In summary, Airbnb prices vary slightly by month, with higher prices in summer months than in winter months. Interestingly, the relationship between month and price is more pronounced for certain property types (i.e., condominiums and apartments) than others (i.e., houses and townhouses) whose prices are more stable over the year.

Research Question 3

How do Airbnb **prices** vary by **location**? Is this the same for all property types?

We answered this question by creating **two interactive scatterplot maps** using plotly.express. We chose to visualize this data on a map because it is the most intuitive way of visualizing geospatial data. Additionally, interactive scatterplot maps allow users to explore data in a more interactive and dynamic way than traditional scatterplots, as users can zoom in and out, pan the map, and click on individual data points for more information.

Figure 16: Price per Night by Location (Houses)

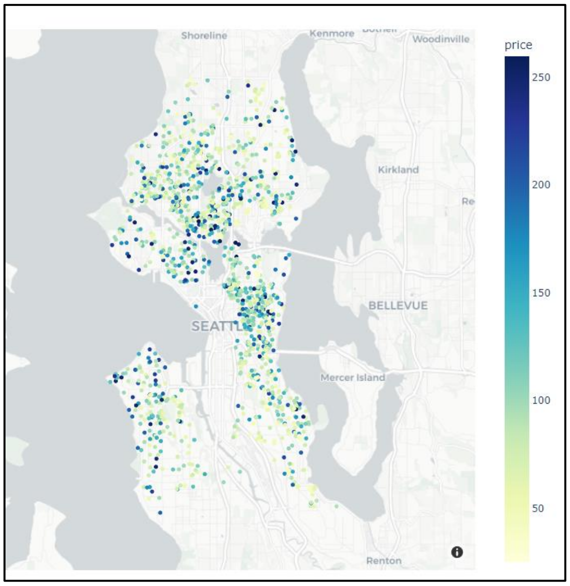


Figure 17: Price per Night by Location (Apartments)

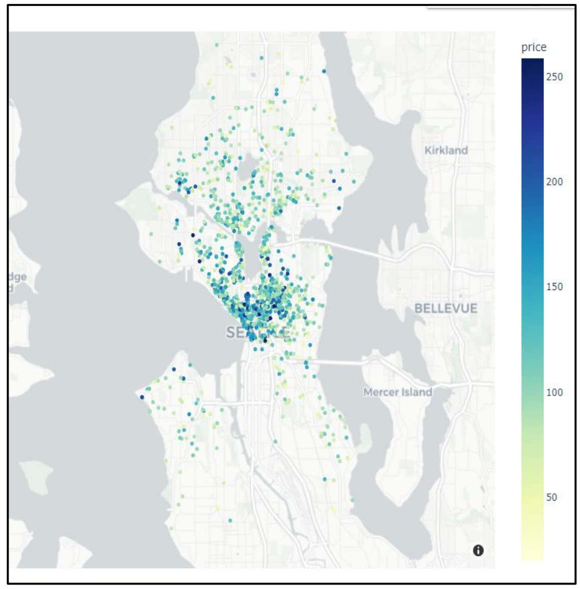


Figure 16 shows the prices of Airbnb house listings in Seattle. Figure 17 shows the prices of Airbnb apartment listings in Seattle. The circle markers represent the Airbnbs and the circle colours represent the price per night on a scale from yellow to blue.

Looking at the two figures side by side, one can notice several differences between them. On the left, it appears that houses are scattered all around Seattle, with few houses in the middle of downtown. Also, it appears that house prices are highly variable everywhere in Seattle (i.e. you can find cheap and expensive houses in every part of town), but the houses on the outskirts of Seattle are slightly cheaper than closer downtown.

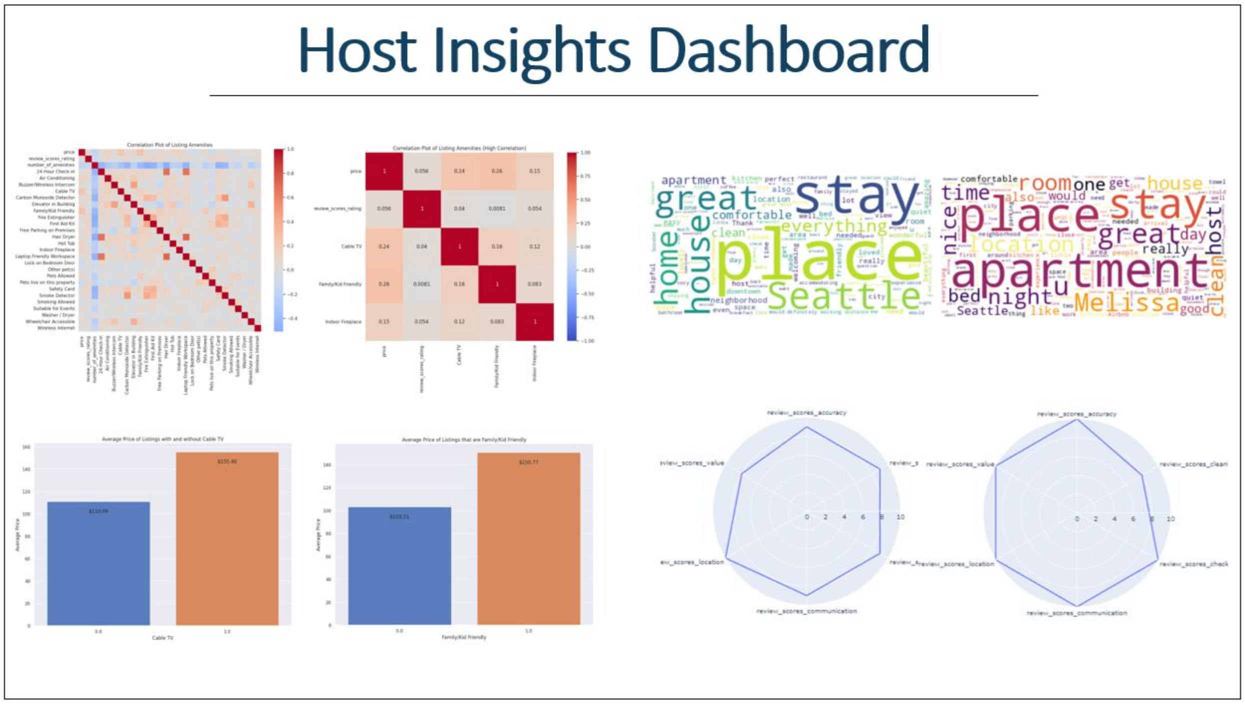
On the other hand, the figure on the right show that apartments are much more concentrated in downtown Seattle and the apartments that are downtown are more expensive than apartments in surrounding areas.

In summary, Airbnb prices do vary by location, and the relationship between price and location differs by property type.

Host Insights

Two of our research questions were relevant to hosts, and we created a Host Insights Dashboard made of 6 data visualizations to answer those questions.

Figure 18: Host Insights Dashboard



Research Question 4

Which **amenities** have the most impact on Airbnb price?

We answered this question by creating **heatmaps** and **bar plots** using seaborn. Heatmaps allowed us to show the correlations between Airbnb amenities, price, and ratings all in one plot. Heatmaps are excellent ways to show correlations because they use color to represent the strength and direction of correlations, which makes it easy to interpret and understand the relationship between variables. They also allow for quick identification of which variables are positively correlated, negatively correlated, or have no correlation.

After the heatmaps, we took a closer look at the variables which were most highly correlated with price by plotting bar graphs showing the price of Airbnbs with and without those amenities. We chose to use bar graphs to show this information because they are intuitive visualizations to emphasize the difference in magnitude between two variables.

Figure 19: Heatmap of Correlations between Airbnb Amenities, Prices, and Ratings

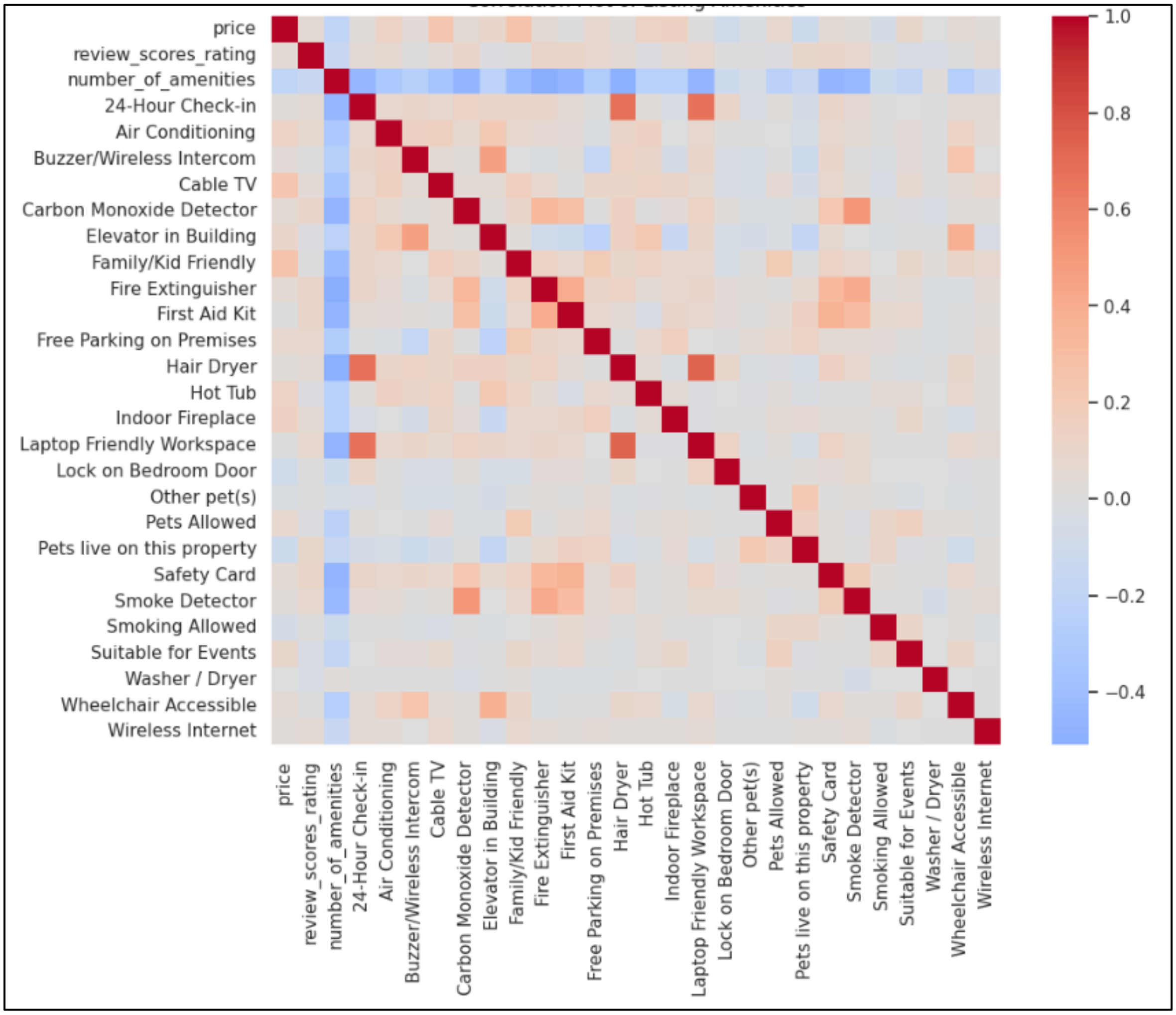


Figure 19 shows the correlations between Airbnb amenities, prices per night, and ratings in Seattle. Using this visualization, hosts could see which amenities other hosts in their area are including in their listings and see if the presence of a certain amenity is correlated with an Airbnb's price or rating.

In this heatmap, it appears that hosts that offer carbon monoxide detectors also offer smoke detectors. Similarly, hosts that offer fire extinguishers often also offer first-aid kits. Interestingly, it appears that hosts that offer 24-hour check-in often also offer hair dryers and laptop-friendly workspaces. In summary, certain groups of amenities tend to show up together at Airbnbs across Seattle.

Additionally, this heatmap shows that few amenities have high correlations with Airbnb price or rating. However, we can take a closer look at the highest correlated amenities with price in the following visualization.

One thing to note is that the amenities were one-hot encoded to create this graph. Correlations between variables can be low for one-hot encoded variables because one-hot encoding involves creating new binary columns to represent the different categories of a categorical variable. These binary columns have a value of 1 if the category is present and 0 otherwise, and therefore there is a limited range of values. This limited range of values means that even if there is a strong relationship between the two categorical variables, the resulting correlation coefficient may be low.

Figure 20: Top 3 Amenities Correlated with Airbnb Price

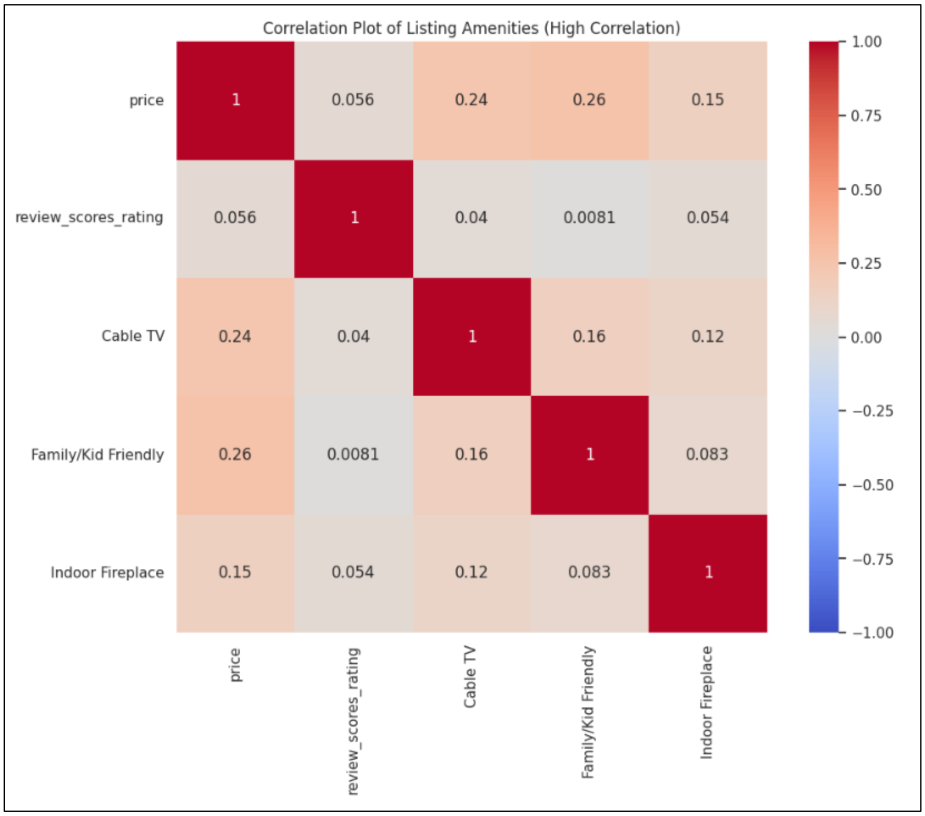


Figure 20 shows that the top 3 amenities that are positively correlated with Airbnb price are Cable TV (r=0.24), Indoor Fireplaces (r=0.15), and being Family/Kid Friendly (0.26). We can take a closer look at the relationship between these variables and price in the next figure.

Figure 21: Average Price of Airbnbs with and without Cable TV

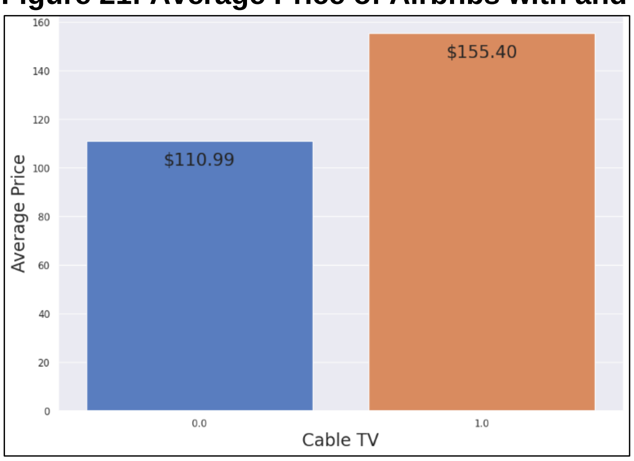


Figure 22: Average Price of Airbnbs that Are and Are Not Family/Kid Friendly

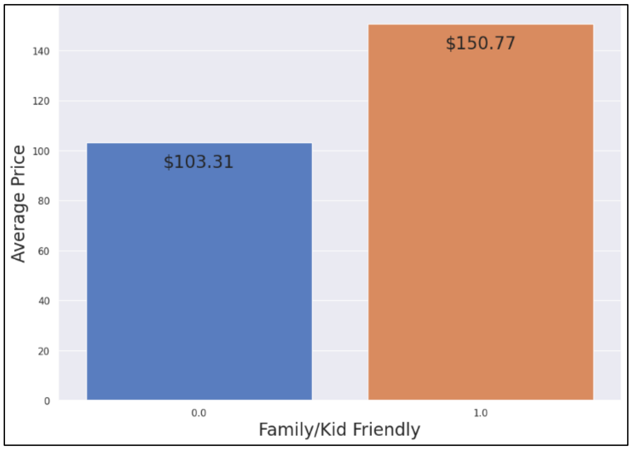
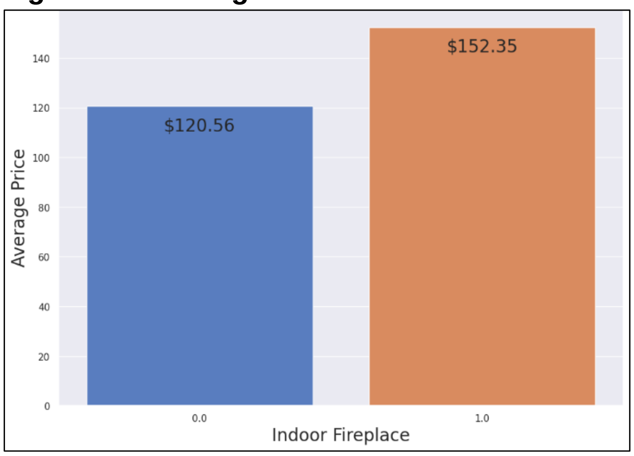


Figure 23: Average Price of Airbnbs with and without Indoor Fireplaces



Figures 21-23 show us that there can be a large difference in Airbnb prices per night depending on what amenities they have. For instance, there is a $45 difference per night between Airbnbs that offer Cable TV and do not offer Cable TV. Similarly, there is a $47 difference per night between Airbnbs that are and are not Family/Kid Friendly. Finally, there is a $32 difference per night between Airbnbs that have and do not have Indoor Fireplaces.

In summary, the amenities that have the highest impact on Airbnb price are Cable TV, Indoor Fireplaces, and being Family/Kid Friendly.

Research Question 5

What valuable information can Airbnb **reviews** and **ratings** give to hosts?

We answered this question by creating **word clouds** using NLTK and wordcloud, and **radar plots** using plotly.express.

We believe that Word clouds could be used to summarize the positive and negative commented reviews of all Airbnbs in Seattle to provide hosts with general trends in reviews that they could learn from. Word clouds are great visualizations because they are simple and easy to understand, making them accessible to a wide range of audiences. They also highlight the most frequent words or phrases in a text, which could help hosts identify important themes or trends.

Radar plots could be used to summarize rating categories for each host's Airbnb. Radar plots are great ways to visualize multidimensional data, where each dimension represents a different variable. This makes it easy to compare multiple variables at once and identify patterns and relationships between them. In this case, each variable would be a different category of rating.

Figure 24: Most Frequent Words in Positive Reviews of Airbnbs in Seattle

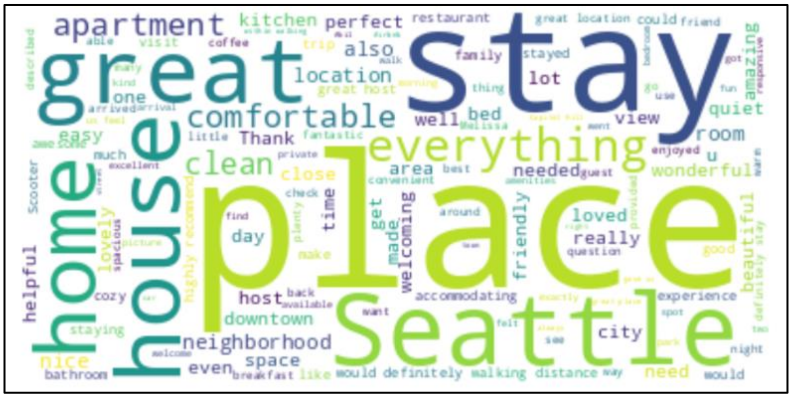


Figure 25: Most Frequent Words in Negative Reviews of Airbnbs in Seattle

Figures 24 shows the most frequent words in reviews for the top-rated Airbnbs in Seattle, and Figure 25 shows the most frequent words in reviews for the lowest-rated Airbnbs in Seattle.

At first glance, it appears that the two visualizations show similar words. This is because regardless of positive or negative sentiment, travelers make comments on the same aspects of Airbnbs. If you look more closely, you see information that could be valuable to hosts. For instance, in the negative word cloud, one can see the words "parking", "towels", and "responsiveness". This suggests that in the lowest-rated Airbnbs, travelers had negative sentiments about the parking situation, the towels, and the responsiveness of hosts. In the positive word cloud, you can see words like "comfortable", "helpful", "clean". These words give insight into things that are important to travelers in Seattle.

Therefore, word cloud summaries of positive and negative reviews of Airbnbs in Seattle can provide hosts with valuable topics and themes that are important to travelers to consider when making decisions about their listings.

Figure 26: Radar Plot of Rating Categories for Airbnb ID 4061051

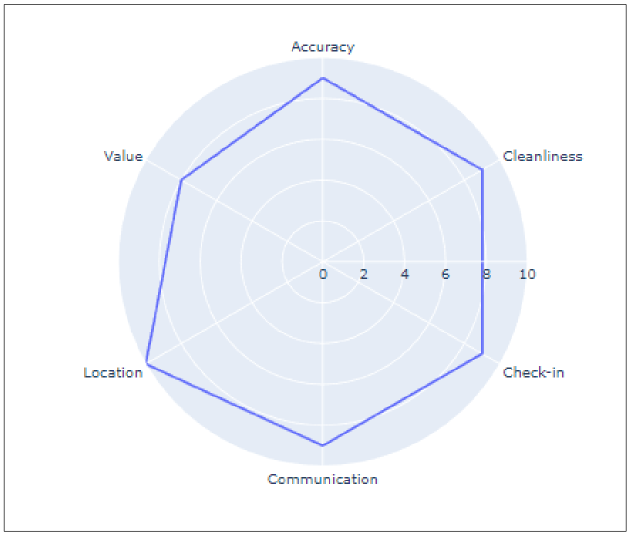
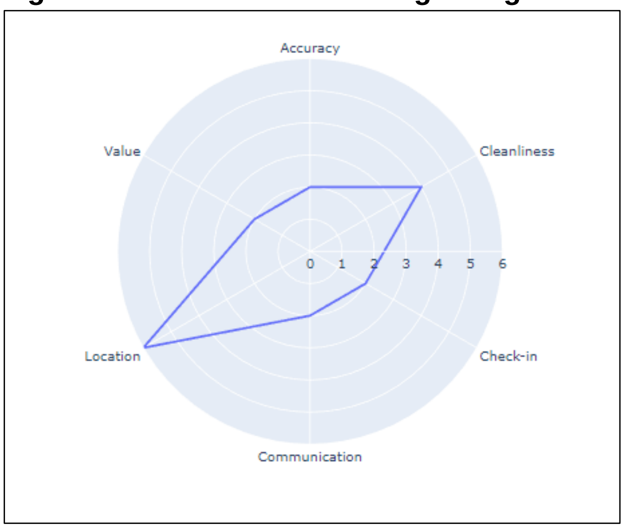


Figure 27: Radar Plot of Rating Categories for Airbnb ID 9183838



Figures 26-27 show radar plots of rating categories for specific Airbnbs. The first plot shows that the host is doing very well keeping their guests happy, but they could improve on the Accuracy of the listing's description and on their communication with their guests. The second plot shows that the host is doing quite poorly on everything except the Location rating, and therefore has much room for improvement in the other categories.

Therefore, radar plots can provide useful visual summaries of Airbnb ratings which hosts can use to track their guests' satisfaction and improve their Airbnbs.

Discussion

In this study, we found 5 insights about Airbnbs in Seattle:

Insight 1:

This study showed that Airbnb prices in Seattle are slightly higher on weekends (Friday nights and Saturday nights) than on weekdays, and this relationship between weekday and price is the same for all property types.

Providing this information on the Traveler Insights Dashboard on the front page of Airbnb's Seattle homepage would allow travelers to make decisions about the dates of their trip. For instance, knowing that Airbnb prices do not increase too much on the weekends could entice them to visit Seattle on weekends and not have to take too much time off from work just to be able to visit on weekdays which they might assume would be much cheaper.

Choosing the right weekdays to travel to Seattle depending on travelers' budgets could result in greater customer satisfaction and increased sales on the Airbnb website.

Insight 2:

This study showed that Airbnb prices in Seattle vary slightly by month, with higher prices in summer months than in winter months. However, the relationship between month and price is more pronounced for certain property types (i.e., condominiums and apartments) than others (i.e., houses and townhouses) whose prices are more stable over the year.

Showing this information on Airbnb's Seattle homepage would inform travelers about trends in Airbnb prices over the year. They could use this information to decide whether they should take seasons into account when planning their trip and if so, which season would best suit their budget. For instance, if travelers know that they want to stay in a house, they don't need to worry about which month they visit Seattle because house prices are stable over the year. On the other hand, if they want to stay in an apartment, it may be in their best interest to visit Seattle in the winter when the prices are generally lower than in the summer.

This could lead to travelers making better-informed decisions about which time of year they would like to visit Seattle, and in turn lead to higher customer satisfaction and increased Airbnb bookings.

Insight 3:

This study showed that Airbnb prices in Seattle vary by location, and the relationship between price and location differs by property type. For instance, Airbnb houses are scattered all around Seattle and their prices are highly variable all around the city except for the very outskirts of town where the prices are generally low. On the other hand, Airbnb apartments are much more concentrated in downtown Seattle and the apartments that are downtown are more expensive than apartments in surrounding areas.

Travelers could use this information to inform their decisions about which area of Seattle they want to book their Airbnb in, which could lead to better matches between guests and Airbnbs. This could lead to higher traveler satisfaction and increased Airbnb sales. .

Insight 4:

This study showed that few amenities have high correlations with Airbnb price or rating in Seattle, however the amenities that have the highest impact on Airbnb price are Cable TV, Indoor Fireplaces, and being Family/Kid Friendly.

Providing this information to hosts on the Host Insights Dashboard would allow them to make data-driven decisions about which amenities to include at their Airbnb that would provide the best monetary return. For instance, knowing the information above, an Airbnb host could reasonably expect the value of their Airbnb listing to increase if they were to provide cable TV, install an indoor fireplace, and create a family/kid-friendly environment. This could help hosts stay competitive in their area and make better business decisions.

Insight 5:

This study showed that word cloud summaries of positive and negative reviews of all Airbnbs in Seattle could provide hosts with valuable topics and themes that are important to travelers to consider when making decisions about their listings.

For instance, some of the words in the word clouds included ‘parking’, ‘towels’, ‘clean’, and ‘responsiveness’. This information could tell hosts what general things to pay particular attention to at their Airbnb to increase traveler satisfaction. Airbnb could also provide hosts with word clouds that were specific to each host’s Airbnb to provide more tailored feedback to hosts.

We also found that radar plots could provide useful visual summaries of Airbnb ratings. Hosts could use these visualizations to track their guests’ satisfaction and improve their Airbnb listings accordingly.

Conclusion

In conclusion, showing data visualizations about Airbnb statistics for travelers on each city's Airbnb homepage could help improve the Airbnb business in several ways. Firstly, it could provide valuable information to potential guests about the characteristics of the local Airbnb market, such as prices by weekday, month, and location. This information can help travelers make more informed decisions about when and where to stay, which can lead to increased bookings and higher customer satisfaction. Additionally, these data visualizations can help showcase the value of the Airbnb platform to potential hosts, particularly in cities where there is high demand for accommodations.

Secondly, data visualizations could provide hosts with valuable information about which amenities are offered in their location and which amenities have the highest impact on Airbnb price and on Airbnb ratings. This could help hosts make data-driven decisions about their properties. Moreover, data visualizations could help highlight areas of improvement for Airbnb hosts and the company as a whole. For example, if data shows that certain cities have consistently low ratings for cleanliness or communication, Airbnb can provide targeted resources to hosts in those areas to improve these issues.

Overall, data visualizations can be a powerful tool for improving the Airbnb business by providing valuable insights to both guests and hosts, and by driving increased engagement and revenue on the platform.

Future Work

Future work in this area could focus on adding more data visualizations to the Traveler Insights Dashboard and the Host Insights Dashboard that focus on other important topics. This could include data visualizations about data trends in Airbnb availability, new Airbnb properties, occupancy rates, revenue, and other metrics. It would also be beneficial to expand these dashboards to cover all cities with Airbnbs, not solely Seattle.