Airbnb in Hawaii Data Visualization

By Xiaolin Liu

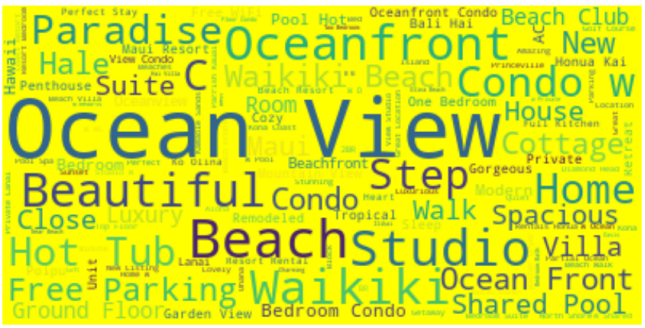

Figure 1: woldcloud generator

Figure one shows when loading the Hawaii Airbnb listing information, what are the most important keywords generated. The bigger the keyword front is, the more weight it has when generating the words. For example, the most relevant or important word would be “Ocean View”.

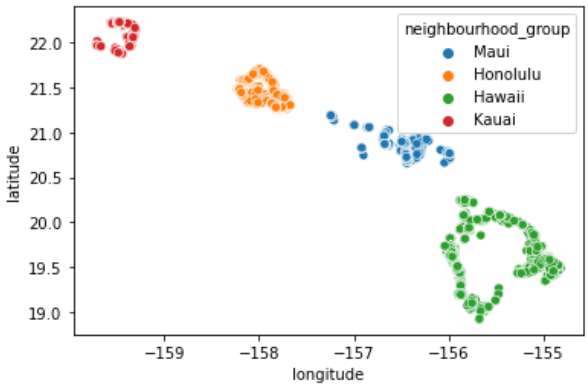

Figure 2: scatter plot of the Airbnb locations

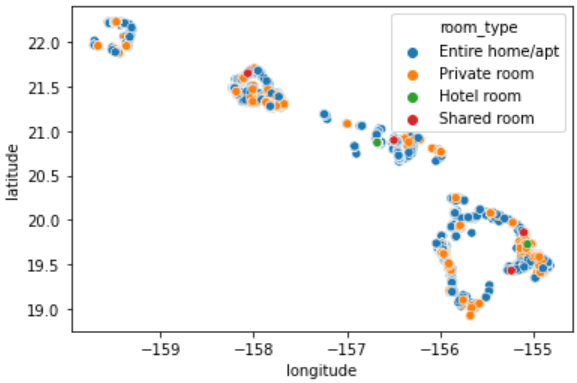

Figure 3: scatter plot of the room types

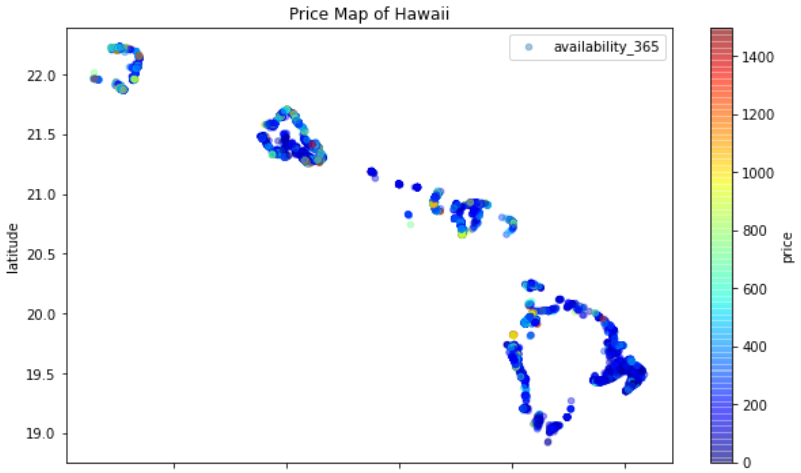

Figure 4: price map of hawaii

Figures 2,3,4 are the scatter plots to show different neighborhoods, room types and locations, and the prices and their locations. For figure 2, the colors are distinguished for the islands; For figure 3, the different colors correspond to the room types with their locations; For figure 4, the different colors are representing the corresponding prices. With the information contained, it is convenient to locate Airbnb in Hawaii and know their information. Since for figures 2 and 3, the x and y axis represent the longitude and latitude, the points correspond to the real location in the map.

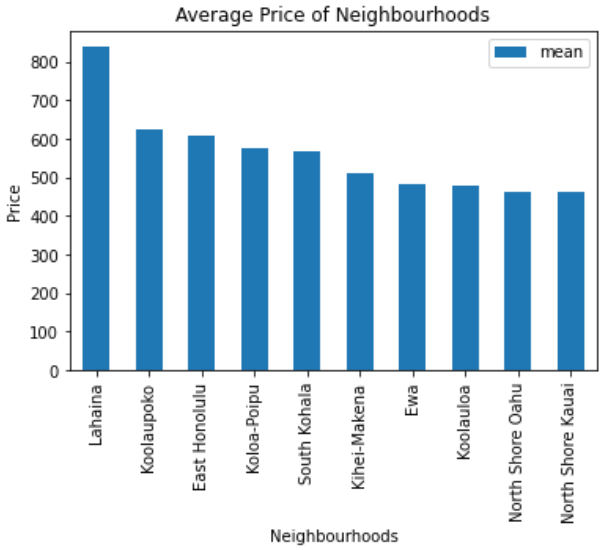

Figure 5: Average price of neighborhoods bar plot

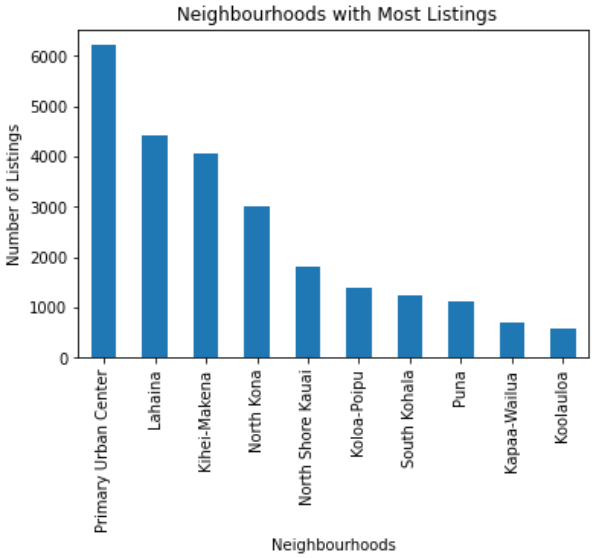

Figure 6: Neighborhoods with Most Listings

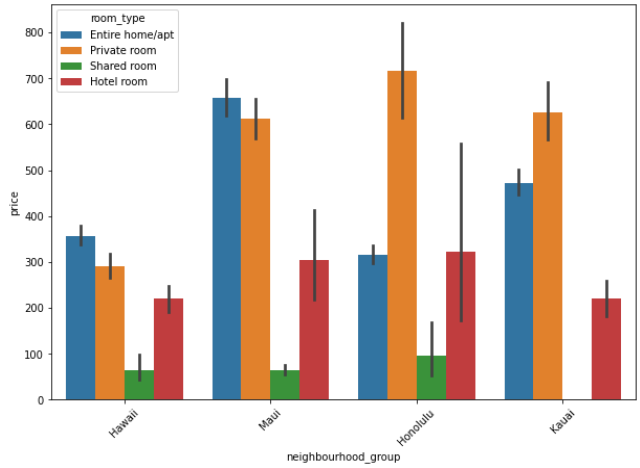

Figure 7: barplot of price and room type

Figure 5 focuses on the average prices of the listings located in Hawaii. The X-axis listed the name of the neighborhoods. Y-axis showed the average price in a certain neighborhood.
Figure 6 focuses on the neighborhood with the most number of listings. The x-axis shows the name of the neighborhoods and the y-axis has the corresponding number of listings.
Figure 7 shows the prices in the four neighborhoods, but different room types were distinguished by using different colors of bars. The black lines on the graph were shown as error

bars, which is the confidence interval of the variables. Indicated how the data are spreading around the mean value.

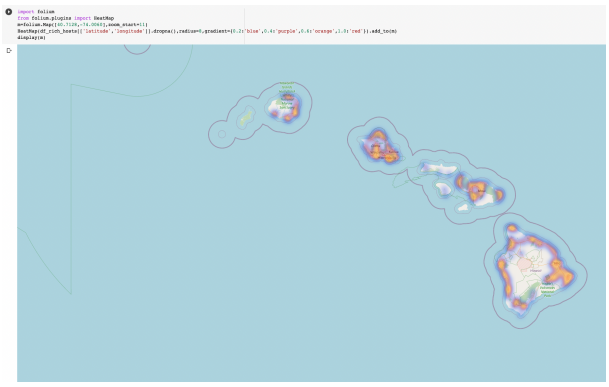

Figure 8: Folium Heatmap

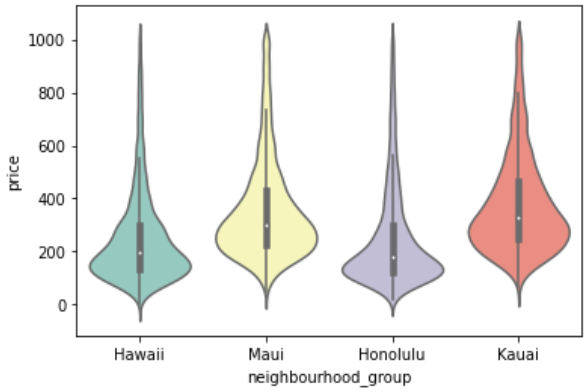

Figure 9: Violin plot of hawaii Airbnb prices

Figure 8 is an interactive folium heat map showing the location of the Airbnb. Heatmap is a holistic way to observe the topography corresponding to the location of the Airbnb locations; it can be seen that most of the locations are along the coastline. Figure 9 violin plot has the neighborhood on the x-axis and the prices on the y-axis, the who dot is the median, the thick bar are the interquartile range [Q1(lower)-25%-Q2(middle)-25%-Q3(middle)], and the wider section of the violin is representing a higher probability and the skinnier section for lower probability.

Importance: I have prepared more graphs in my code but only selected nine figures to present. Because I would like to see if all of the graphs will convey the same message. And if the information that has been illustrated matches each other. It's also good to see the data visualized in different shapes and formats. The landscape ones can show the physical location vividly but are limited to showcase the amount of data. Whereas the violin plot can show the probability with the amount since it is a combination of density graph and box plot but it cannot show the corresponding location on the single graph itself. An interactive heatmap allows the users to zoom in and out and view more detailed information but it is hard to tell the prices and the number of listings. So with all of the above-generated plots, all of the information can match each other and connect. This gives us a holistic view of the Airbnb listings in Hawaii, the ultimate paradise.

Data and method: The data I used for this assignment came from the open-source Airbnb Hawaii dataset. You may also download it directly from my Github repository. All of the graphs were generated using python. Depending on the type of the graph, different python packages were installed. A more detailed python package usage could be found on the repository as well.

Github Link:

https://colab.research.google.com/drive/1e2tukikHvGiEflpdJMqF3WbnotKxB9Ko?usp=sharing