Melbourne Airbnb Marketing Analysis

Zhi Li

Introduction:

The dataset for this project is the Melbourne Airbnb Dataset and it was obtained from the Kaggle dataset (https://www.kaggle.com/tylerx/melbourne-airbnb-open-data). This dataset describes the listing activity of homestays in Melbourne, VIC, Australia, which was compiled on 07 Dec 2018. The following Airbnb activity is included in this Melbourne Airbnb dataset :

- Listings: detailed listings data including full descriptions and average review score.
- Calendar: detailed calendar data for listings, including listing id and the price and availability for that day.
- Reviews, detailed review data for listings including unique id for each reviewer and detailed comments.
- Listings-Summary: summary information and metrics for listings (good for visualizations).
- Reviews-Summary: summary Review data and Listing ID (to facilitate time-based analytics and visualizations linked to a listing).
- Neighborhoods: neighborhood list for geo filter. Sourced from a city or open source GIS files.

I chose this particular dataset because it has abundant info of the vocational rentals in Melbourne, including numerical data, categorical data, text data, and geographical data. This will give me a great opportunity to investigate the complex dynamics of vocational rental marketing in Melbourne. In addition, this diverse datatype would enable me to have a better visualization to understand the data. Since this report is aimed to analyze the factors affecting the nightly price of vocational rentals in Melbourne. I would like to know how factors like location, number of guests, property type, availability, etc can affect the marking price of homestay in Melbourne.

Summary of Data :
After cleaning the dataset (removing some outliers), the greater Melbourne area had :

- 22,895 listings
- Across 59 neighborhoods
- Hosted by 14954 hosts

1. Distribution of the Nightly Price

At first, I need to know the distribution of the nightly price in the greater Melbourne area.

Fig 1. Histogram of Airbnb Nightly Price in Melbourne

From the histogram, we can infer that the median nightly price is around 110 AU dollars. Most of the homestays are below 200 AU dollars per night. However, there are still homestays expansive rates and some of them even above 3500 AU dollars per night.

2. Geographical Distribution of the listings

Then, I would like to know how the vocational rentals distributed in the greater Melbourne area. The above interactive Bubble plot showed that the listings of vocational rentals are not evenly distributed across Melbourne. It shows that Melbourne Central Business District (CBD) has the highest density in terms of listings. When I zoomed in the CBD area, I found that Southern Cross Station area and Melbourne Central Station area have the highest densities. This may due to the proximity to the main transit stations can attract more customers for the sake of convenience.
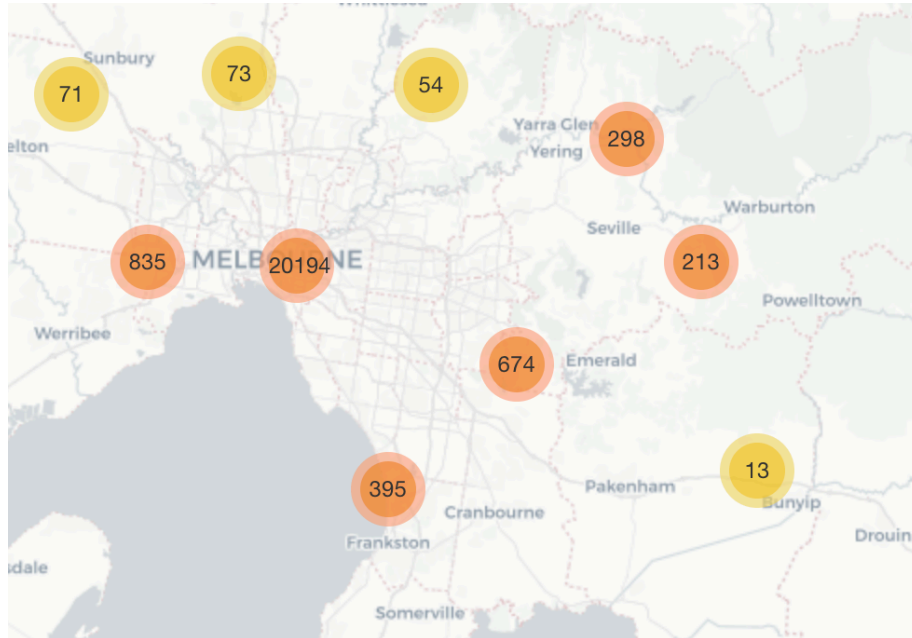
Fig. 2 Interactive Bubble Map of listings in Melbourne

3. Number of Listings by the Property and Room Types

On the Airbnb website, a host needs to specify the type of the property, e.g., apartment or house, and which part he/she is renting out, e.g., entirely or one room. These are the property type and room type of a listing. Let's see what are the types of the listings :

- Two-thirds of the listings are renting out entirely, around one third are private rooms.
- Most of the listings are apartments and houses.
- 75% of the apartments are renting out entirely while 60% of the houses are renting out as private rooms.
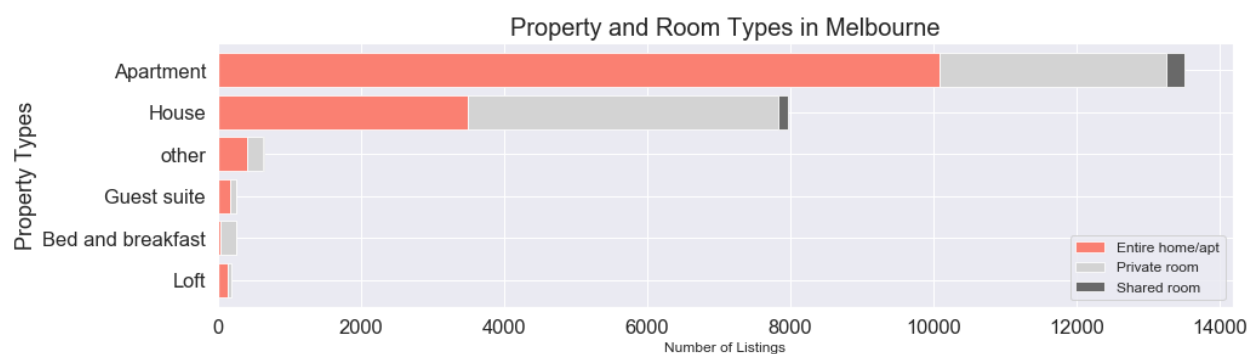

Fig. 3 Number of Listings by the Property and Room Types

4. Investigate What Numerical Factors are highly correlated with Nightly Price.

After knowing the distribution of nightly price, we would like to know what numerical factors can have a great influence on the nightly price. Fig. 4 is a heatmap that described the correlation relationship between different numerical factors with the nightly price.
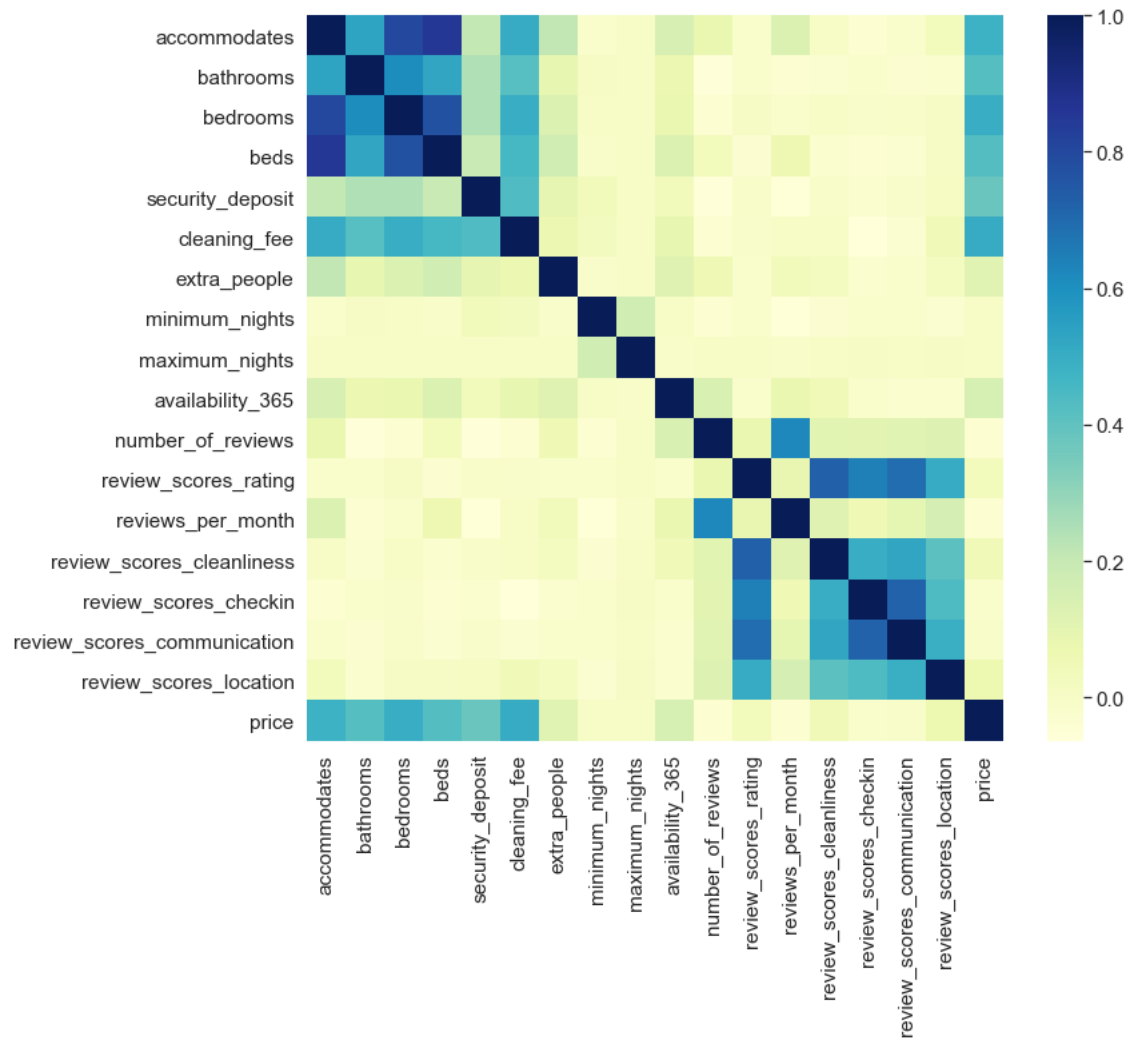
Fig. 4 Heatmap of correlation matrix between Numerical Factors with Nightly Price

From the correlation heatmap, we can infer that most of the numerical factors are positively correlated with the nightly price. Especially, the factors like accommodates, bathrooms, bedrooms, beds, security deposit are cleaning fee are highly positively correlated with nightly price. While accommodates, bedrooms are beds highly correlated with each other. Factors like minimum nights, maximum nights, review scores (cleanliness, check-in, communication, location and rating) almost have zero impact on nightly price.

5. Is there a clear pattern between the availability of properties in a year and the nightly price?

The price is usually affected by supply and demand. We can also infer that the availability of a property in a year has some influence on the nightly price from the correlation heatmap. So I'd like to know is there a clear pattern(relationship) between the availability and the nightly price. To better understand their relationship, I plotted the nightly price against the availability. The Fig

5. shows that the points are evenly distributed. Therefore, we cannot get a clear pattern (relationship) between the nightly price and the availability in 365 days.

Fig 5. Interactive Scatter Plot of Nightly Price Vs. Availability in 365 Days

6. Is there a clear pattern between the nightly price and cleaning fee?
From the heatmap, we can know that the cleaning fee is highly correlated with the nightly price. So I'd like to know is there a clear pattern(relationship) between the cleaning fee and the nightly price. To better understand their relationship, I plotted the nightly price against the cleaning fee. We can see a clear trend between these two variables in the scatter plot. Fig 6. shows that as the cleaning fee increased, the nightly price increased. Most of the cleaning fee is below 200 dollars.
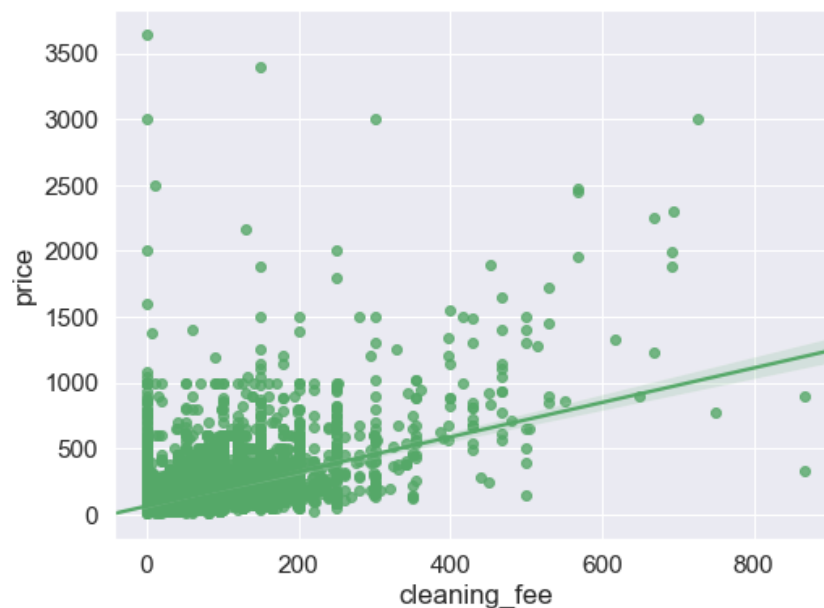

Fig 6. Scatter plot of Nightly Price Vs. Cleaning Fee
7. Does the Property Type Influence the Nightly Price.

Property Type is a categorical variable, I cannot simply use the scatter plot to investigate its relationship with the price. So I grouped the prosperities by their types and calculated the median nightly price for each group. The treemapping indicated that the property type can influence the nightly price. For instance, property types like Castle, Boutique Hotel, Boat have a high median nightly price. While the property types like Campsite and Minsu have a low median nightly price(below 50 AU dollars).
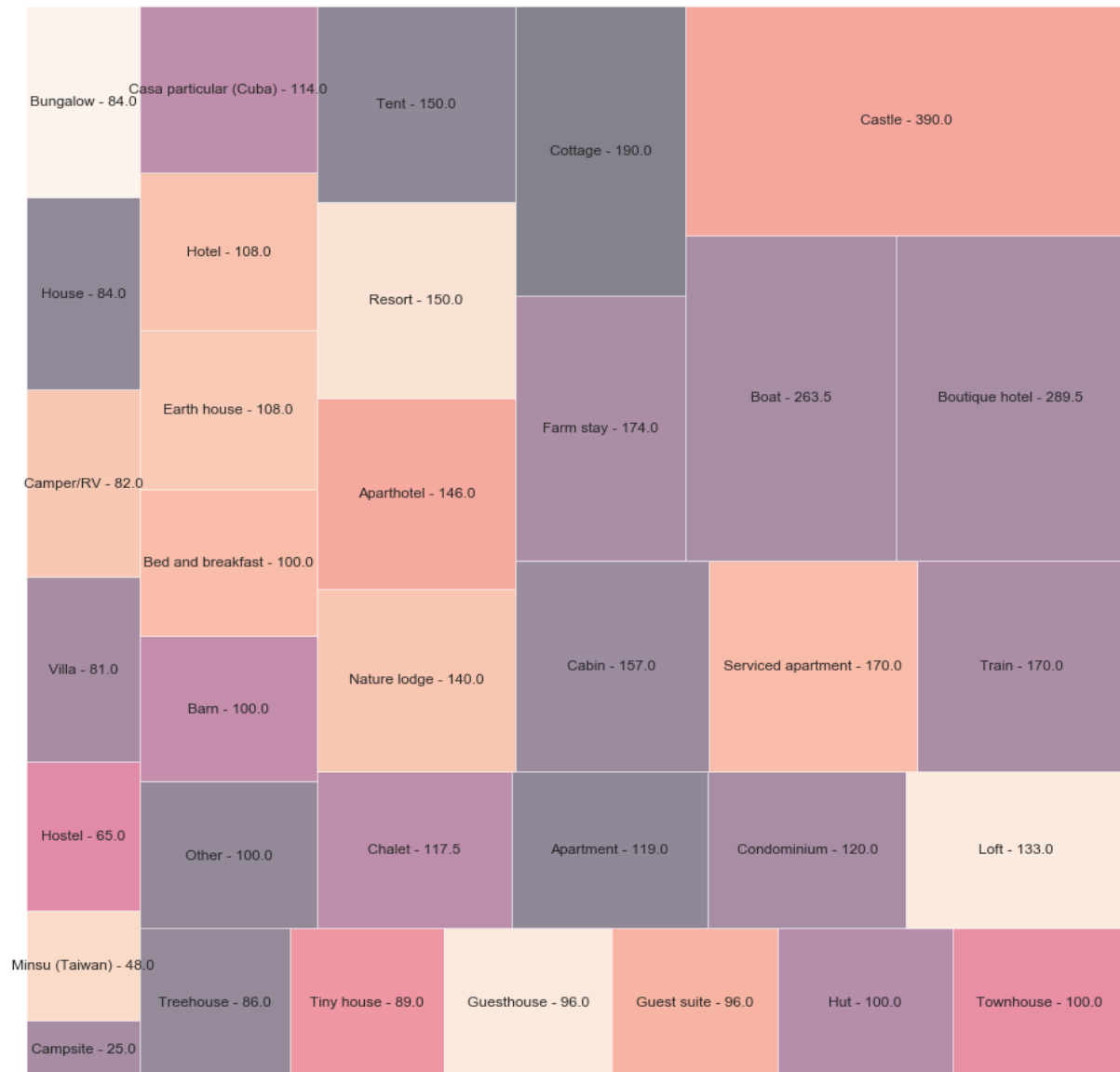

Fig 7. Tree Map of Median Nightly Price of Each Property Type

8. Does Cancellation Policy Affect the Daily Rental Price?
From the Fig.8, we can infer that the cancellation policy can influence the daily rental price. For instance, the cancellation policy of super strict 30 days has the highest daily median rental price. Though flexible cancelation policy has a relatively low median daily rental price, it has a
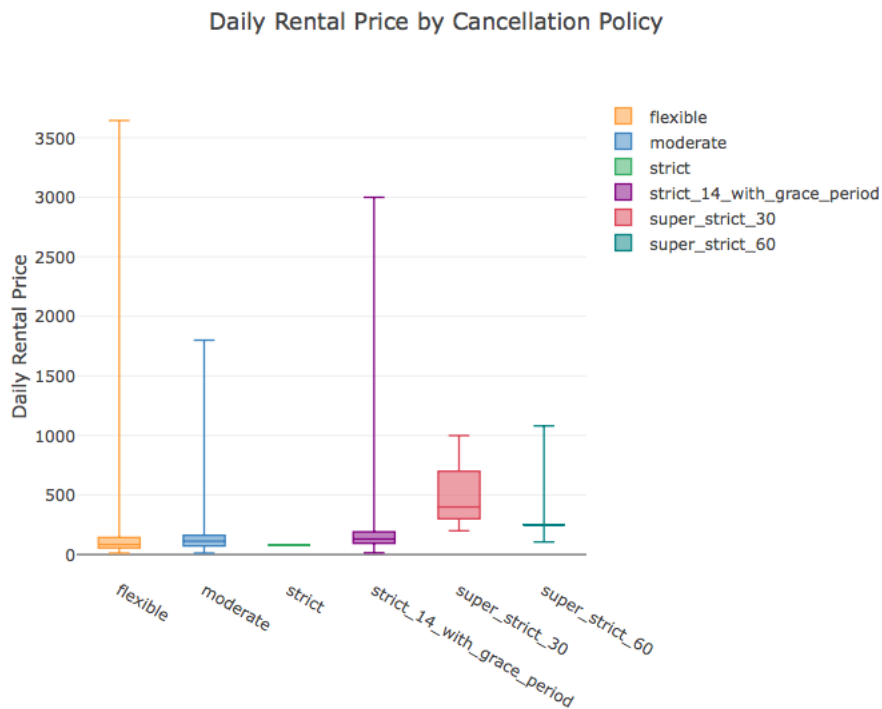
lot of variances than other groups.

Fig. 8 Boxplot of Daily Rental Price by Cancellation Policy.

9. Does the room type and the number of bathrooms affect he Daily Rental Price?

From the below stack area graph, we can infer that both the number of bathrooms and room type can affect the daily rental price. As the number of bathrooms increased, the daily rental price increased. When the number of bathrooms is fixed, the entire home/apt has the highest daily rental price while the shared room has the lowest daily rental price.
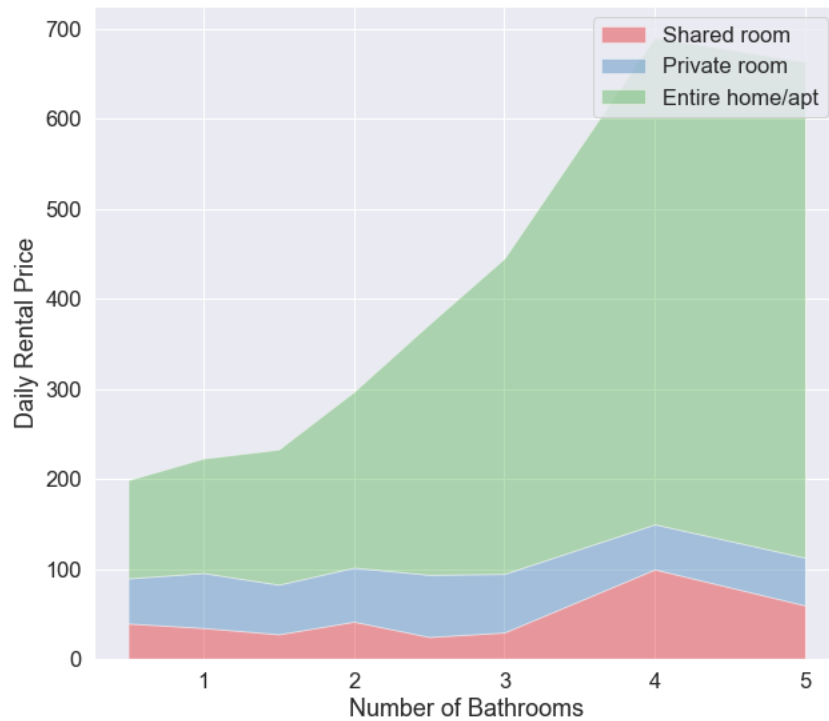
Fig.9 Stack Area Graph of Daily Rental Price vs. Number of Bathrooms by Room Types.

10. Does the neighborhood Affect the Daily Rental Price?

The below choropleth map indicated that neighborhood can has an impact on the daily rental price. When I zoom in, I found the following neighborhood had a relative high daily median rental price :

- Willamtown
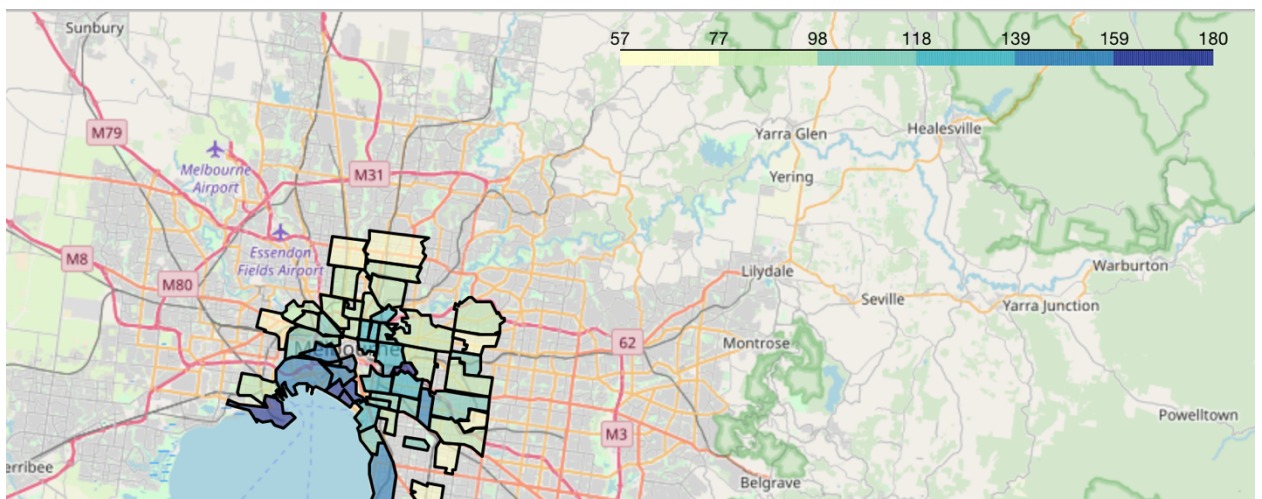- Middle Park
- Albert Park
- Burnley

Fig. 10 Interactive Choropleth Map of Median Price for Each Neighborhood

11. The Geographical difference between the property with highest daily rental price and the prosperity with the lowest rental price.

From the connection map below, we can infer there is a geographical difference between the property with highest daily rental price and the prosperity with the lowest rental price.
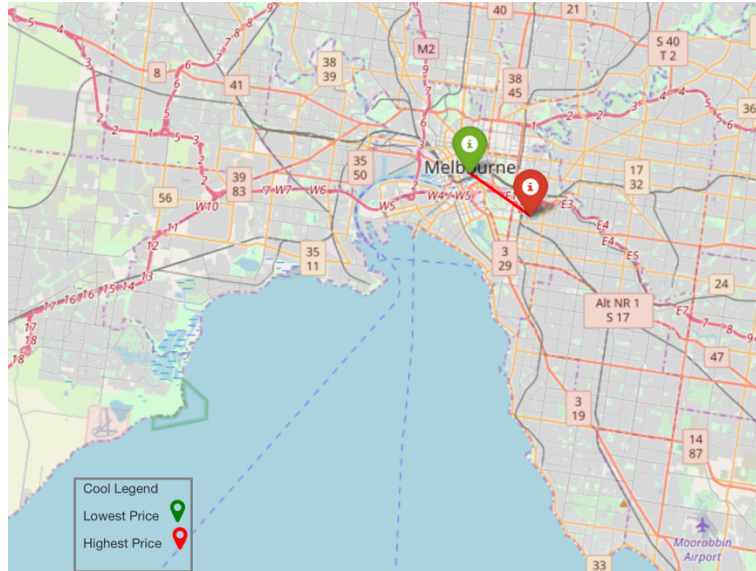

Fig. 11 Interactive Connection Map between the Property with the Highest Daily Rental Price and the Prosperity with the Lowest.

Story Line :

Travel as an industry has been evolving dramatically. In recent years, sharing economies have becomes more popular, giving rise to companies like Airbnb that utilize sharing economies to create an online marketplace for the exchange of private products and services. Nowadays, vacation rentals are booming in Melbourne, Australia. Melbourne was 6th on the list of top ten cities for users globally in 2016 and has been one of the top cities for listings globally since then. More and more people are using Airbnb services or they become Airbnb hosts recently. It's of great importance to know what is the proper a proper daily rental price in Melbourne.

From the above data summary and analysis, we can infer that the factors like accommodates, bathrooms, bedrooms, beds, security deposit are cleaning fee, property type, room type, and the neighborhood of the property are highly correlated with the daily rental price in Melbourne. The following visualization will give us a basic idea about the daily rental price in a given situation.

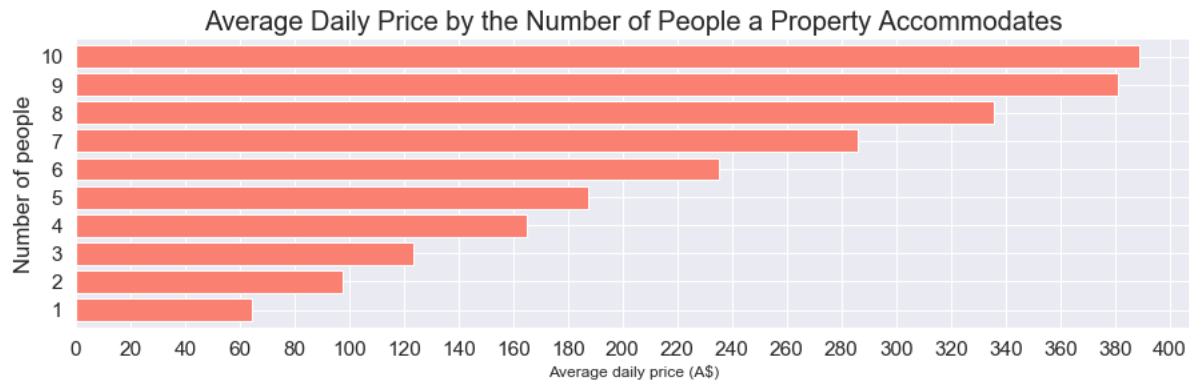1. **Daily Rental Price by Number of Guests**

Fig. 12 Average Daily Price by the Number of People a Property Accommodates

We know that accommodates are highly correlated with the nightly price, we want to further what extent that they influence the price. The prosperities are grouped by the number of people they can accommodate and then the average nightly price for each group are calculated.

From Fig 12. we can infer as following :

- $ 65 for 1 person,
- $ 100 for 2 people,
- $ 125 for 3 people,
- $ 170 for 4 people,
- $ 190 for 5 people,
- $ 235 for 6 people,
- $ 286 for 7 people,
- $ 335 for 8 people,
- $ 381 for 9 people,
- $ 388 for 10 people.

2. Daily Rental Price by Property Type and Room Type

The prosperities are grouped by the property type and room type and then the average nightly price for each group are calculated. From Fig 13. we can infer the daily rental price as following :

- $220 per night for an entire villa,
- $199 per night for an entire house,
- $136 per night for an entire apartment,
- $ 145 per night for an entire Condominium,

and

- $100 per night for a private room in an 'bed and breakfast',
- $60 per night for a private room in an Townhouse or a villa,
- $35 per night for a shared room in an apartment.

Note that we only consider a combination of property type and room type that has more than 100 listings.
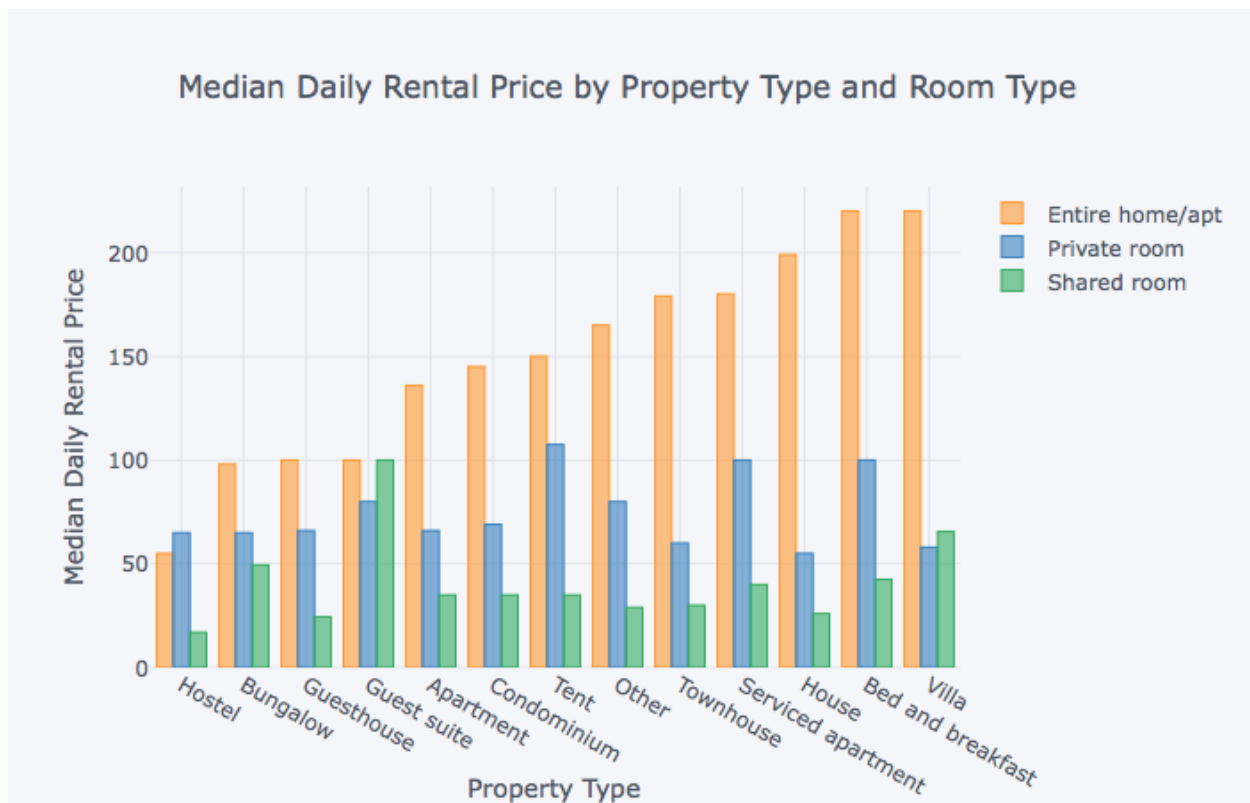
Fig. 13 Median Daily Price by Property Type and Room Type
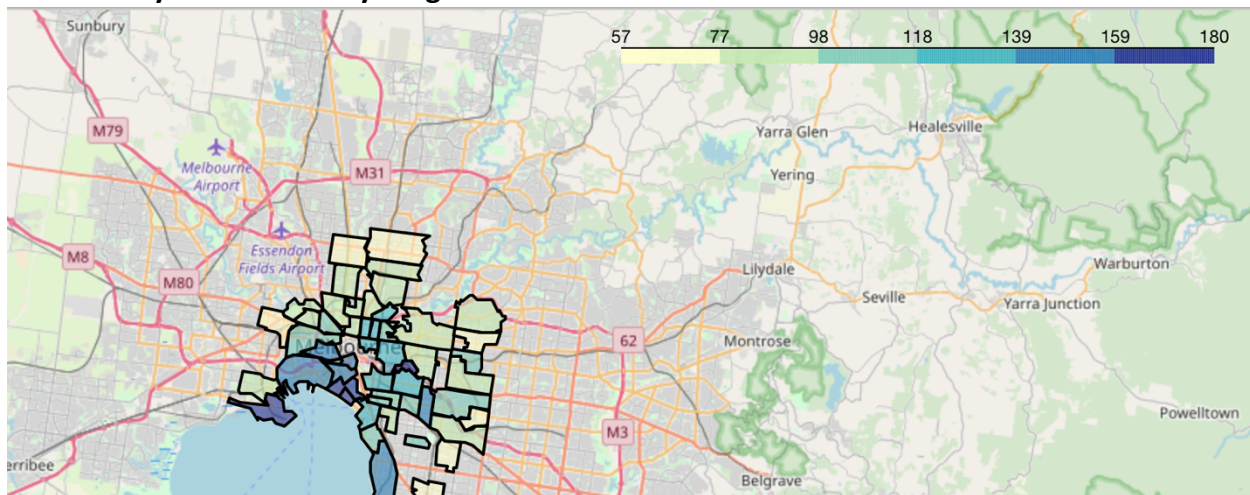
3. Daily Rental Price by Neighborhoods

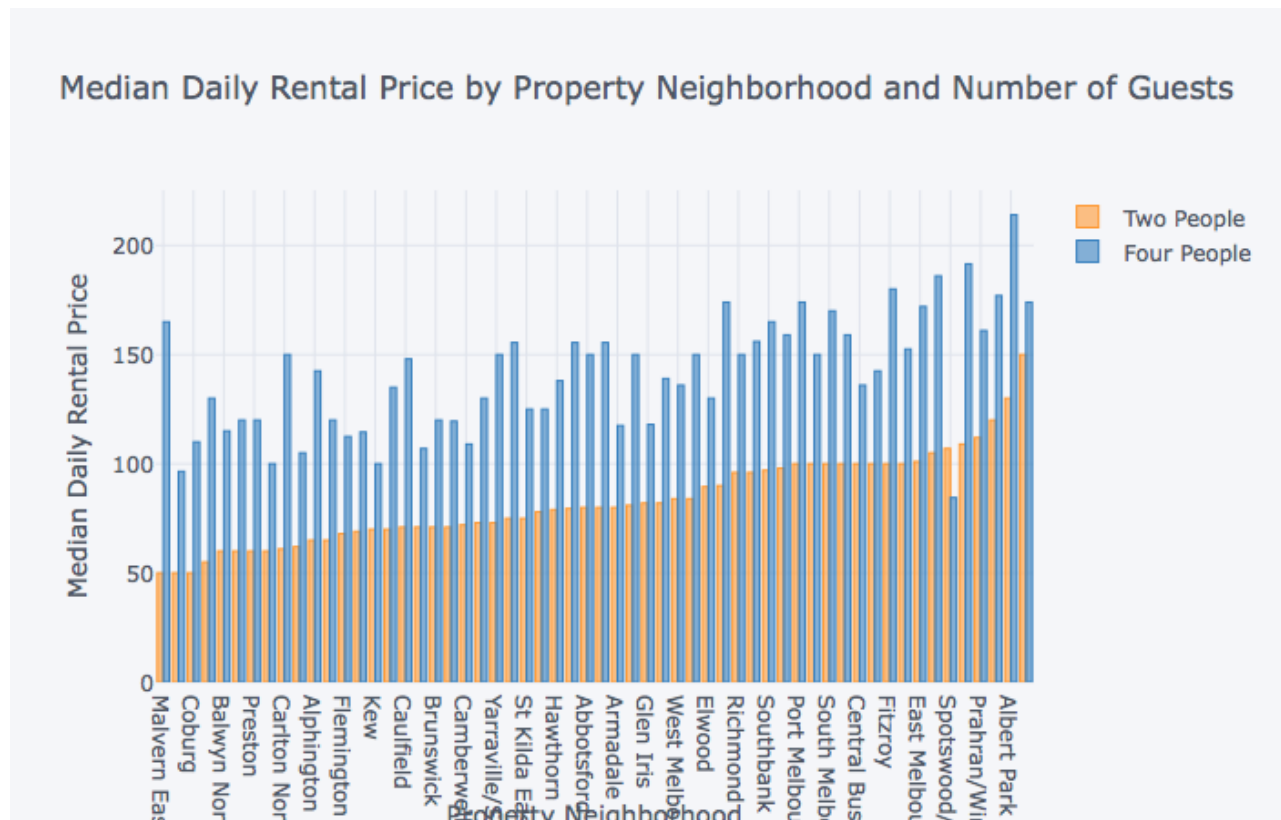Fig. 14 Choropleth Map of Median Price for Each Neighborhood

Fig. 15 Median Daily Price by Neighborhood and Number of Guests

The choropleth map indicated that the neighborhood can have a great impact on the daily rental price. The following neighborhood has a relatively high daily median rental price:
- Williamtown
- Middle Park
- Albert Park
- Burnley

Next, I'd like to know the median price of the different neighborhood and the number of people it accommodates.
- The neighborhood of Malvern East has the lowest daily rental price for accommodating two people, which is 50 AU dollars.
- The neighborhood of Albert Park has the highest daily rental price for accommodating four people, which is 214 AU dollars.

Results & Conclusion :

From the above analysis, we can infer that the factors like accommodates, bathrooms, bedrooms, beds, security deposit are cleaning fee, property type, room type, and the neighborhood of the property are highly correlated with the daily rental price in Melbourne. Customers and hosts can be based on these factors to evaluate the proper marking daily rental price in Melbourne. For instance, people have to pay more in neighborhoods like Williamstown, Middle Park, Albert Park, and Burnley. The property types of castle, villa, house also will charge more than other property

types. These marking price report will help customers to know how much they need to pay for a short-term vocational rental.

Link to your GitHub page with this analysis :
https://github.com/zhlli1/Melbourne-Airbnb-Merketing-Analysis

Citations:
https://towardsdatascience.com/visualizing-data-at-the-zip-code-level-with-folium-d07ac983db20
https://python-visualization.github.io/folium/quickstart.html#Choropleth-maps
https://medium.com/@bobhaffner/creating-a-legend-for-a-folium-map-c1e0ffc34373
https://github.com/sajal2692/Scalable-Walkability-Analysis-of-Melbourne
https://www.kaggle.com/tylerx/melbourne-airbnb-open-data

Appendix Containing All Code: