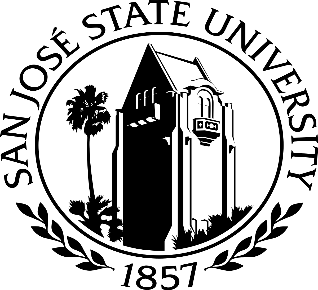
**A Project Report on**

**Analysis of Paris Airbnb Dataset**

**INFO 215: Information Visualization**

**San Jose State University**

**Master of Science in Data Analytics**



**Group Members**

Aarathi Nadathur - 013704405

Divya Puraswani - 013755391

Sai Chaitanya Tolem - 013008788

Sameer Rajput - 013802841

Shubh Johri – 013750672

Contents

[Introduction and Motivation of Study 3](#_Toc7301627)

[Description of Dataset 3](#_Toc7301628)

[Design Principles/Perceptual Properties 3](#_Toc7301629)

[Aim 4](#_Toc7301630)

[Where to List 4](#_Toc7301631)

[Which factors to consider while listing 6](#_Toc7301632)

[When to list the property 9](#_Toc7301633)

[Challenges 10](#_Toc7301634)

[Conclusions 10](#_Toc7301635)

[References 11](#_Toc7301636)

## Introduction and Motivation of Study

Paris being one of the most beautiful cities in the world attracts millions of tourists every year. Since, exquisite hotels in Paris are very expensive, it is a good opportunity for independent property owners to rent their place on Airbnb, if they want to earn extra income. Deciding the factors to look at for listing is a challenging task for the host. Therefore, in this study, we performed in-depth analysis of Airbnb data for Paris through visualizations using Python and Tableau, which would help Airbnb to increase its revenue.

## Description of Dataset

The dataset is sourced from website http://insideairbnb.com/get-the-data.html and data was compiled on 05 February 2019 (Inside Airbnb).

We used the listings.csv dataset which consists of information and metrics for listings in Paris and reviews.csv consists of review data and listing ID for the years 2009 – 2019.

## Design Principles/Perceptual Properties

The goal in our project is to translate abstract information into visual representations that can be easily, efficiently and accurately decoded. To achieve this goal, the visualizations focus on the following visual properties:

* **Color**

Color can be expressed in many ways and is used to attract the viewers’ attention. In figure 4, hue of blue was used to depict correlation. Most of the graphs have the bars in blue and we haven’t used the combination of red and green as almost 10% of men and 1% of females suffer from colorblindness (Few, February 2008).

* **Form**

Length is a very good pre-attentive attribute for quantitative information. We have used bar charts for most of the figures as they are easier for interpretation and comparison. Also, the length of the graph makes the analysis easier.

* Other pre-attentive visual properties like intensity, length and the width of the figure are also focused on the graphs below.
* We used Tufte's Design Principles to ensure high graphical integrity by showing that the labels are clear and detailed. So, all the figures have a title, labels for x, y and an interactive legend.

## Aim

Increase in the revenue of Airbnb depends on the revenue earned by the hosts as Airbnb charges 3%, only once the reservation is completed (Airbnb Financials). Hence, the month when the host lists, the location where he lists, and the type of property he lists matter!

Therefore, to achieve the desired results, we narrowed down our analysis to one neighborhood based on the average and median price of that neighborhood and the competition in that neighborhood. Then, we followed it with analyzing the highly correlated features in that neighborhood and selecting the time to list the property.

## Where to List

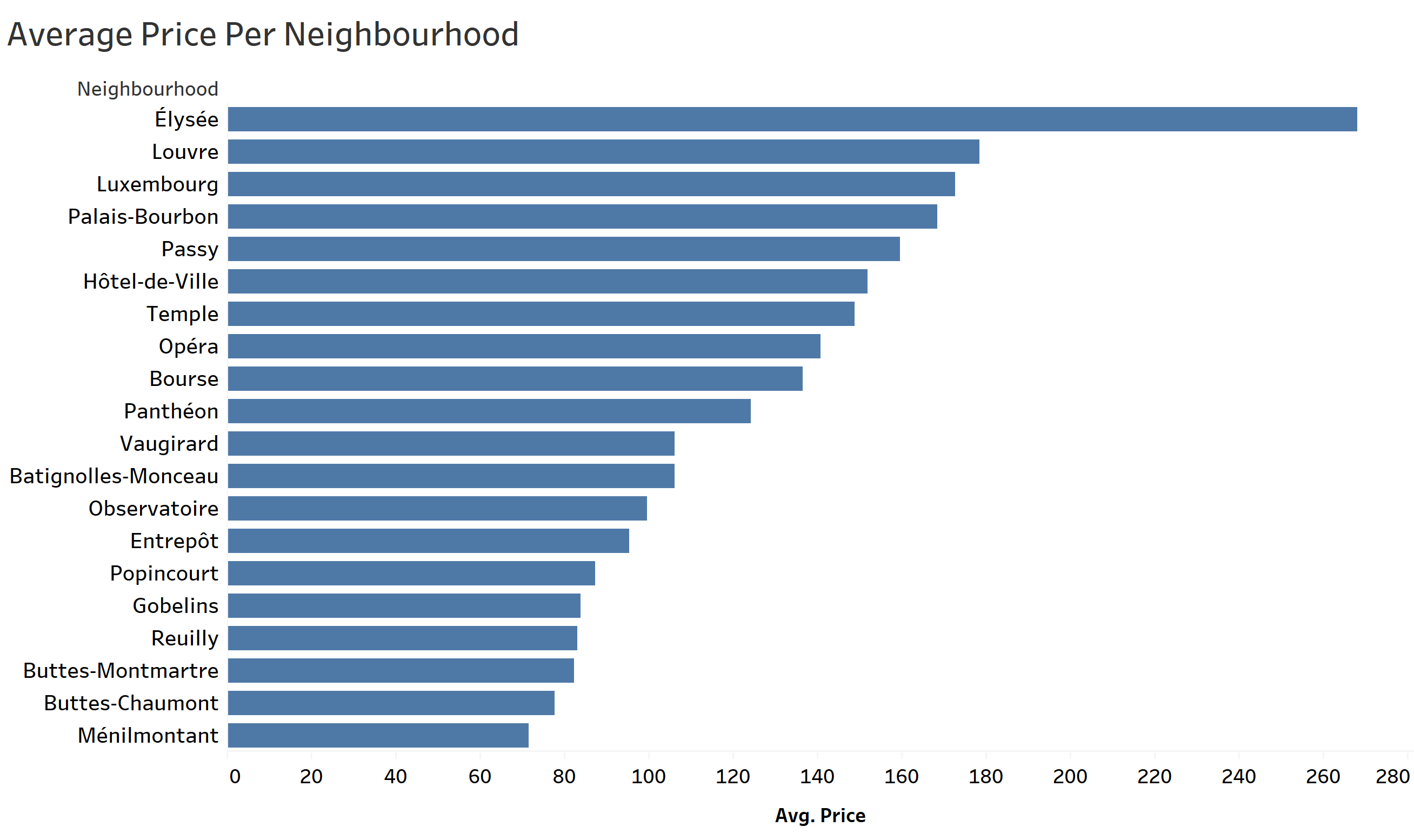


Figure 1

In Figure 1, we see that the average price in the Elysee neighborhood is the highest. Let’s plot the median price of all neighborhood to get a further insight as to what is the distribution of the price Elysee.

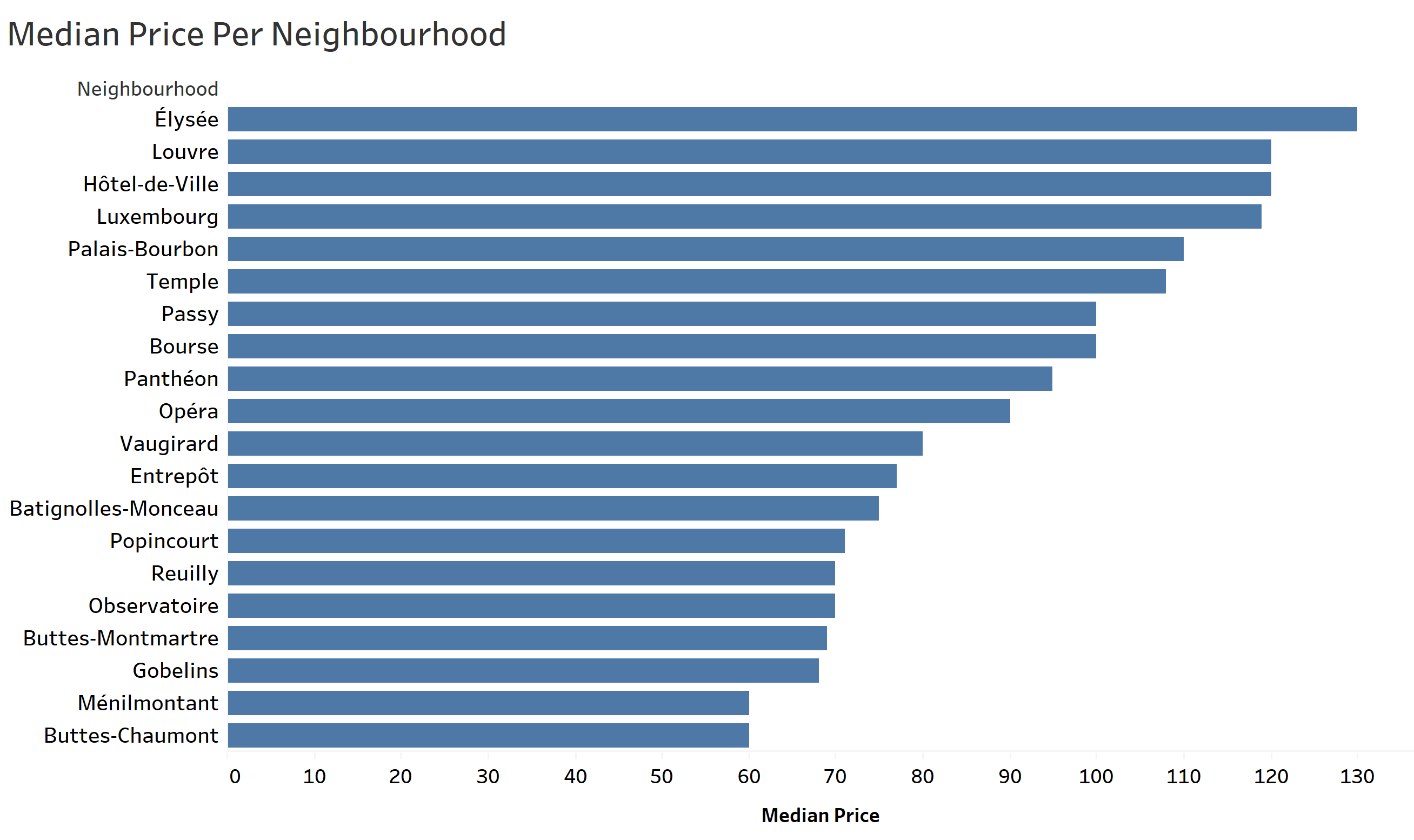


Figure 2

From Figure 1 and Figure 2, we see that half of the listings in Elysee are priced lower than $130, while the average price of the listings in Elysee is $263. From that, we can say that half of the listings in Elysee are significantly priced higher. Hence, the host may choose to list in the Elysee neighborhood, but at the same time the host would want to check how the competition is in that area.

To know about the competitiveness of the neighborhood we can see the number of listings in that neighborhood.

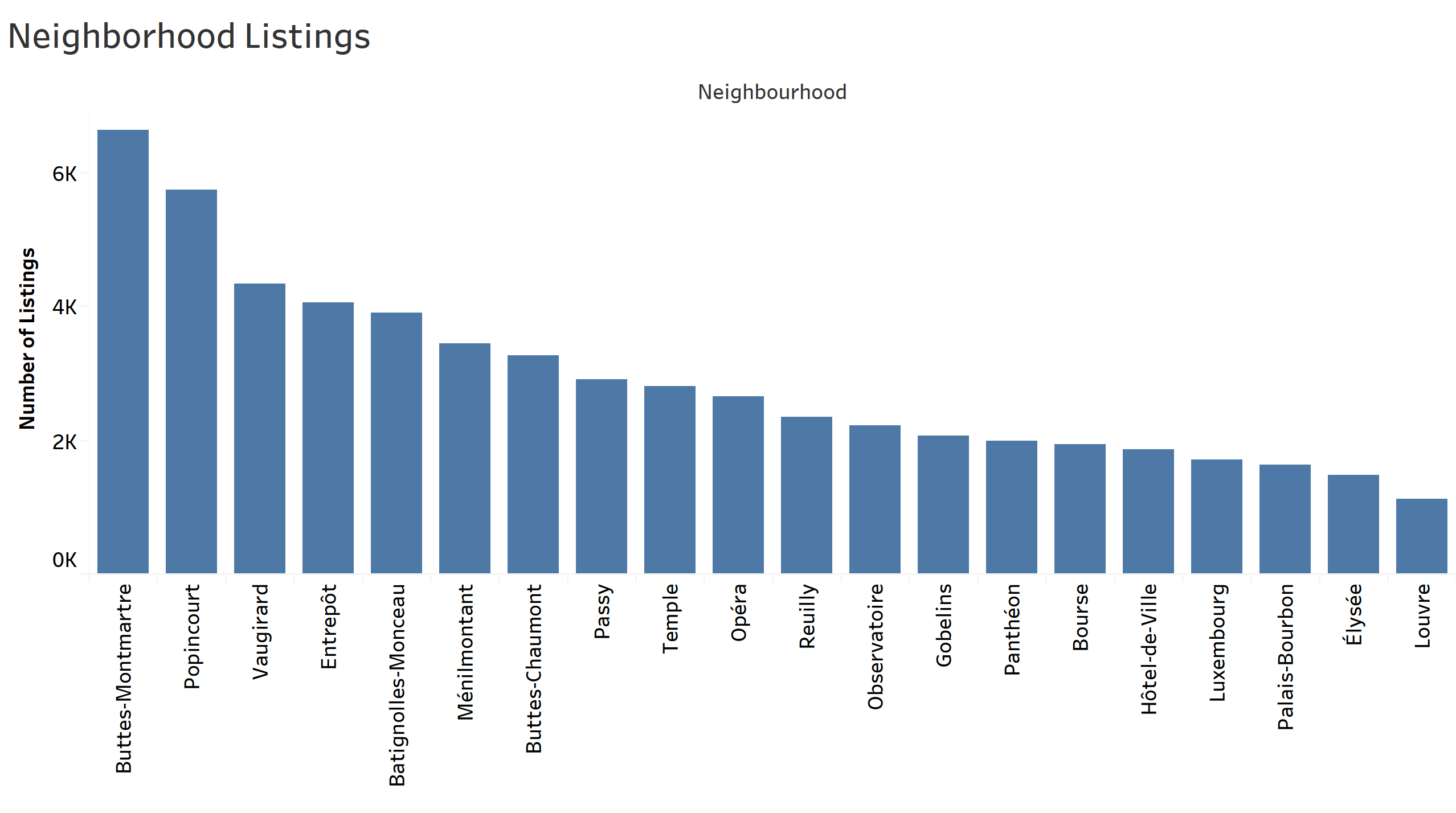


Figure 3

Here we observe that the number of listings in Elysee are quite low as compared to the other neighborhoods, which means that the competition in that area is less. Therefore, from the above three graphs, we infer that in Elysee, the host can exploit the market as the average price in that market is high while the competition is not as much.

Additionally, Elysee has several attractions nearby- Eiffel Tower is about 10 minutes away by Subway. Châtelet-les-Halles is about 15 min by Subway. It also has two airports nearby namely, Charles de Gaulle airport and Orly airport, with easy public transit (Airbnb - Elysee).

After selecting the right neighborhood, the host should consider the factors which are to be included in the listing.

## Which factors to consider while listing

To increase the revenue, the host should consider the factors which affect price the most. For that, we will see how the factors are correlated with price.

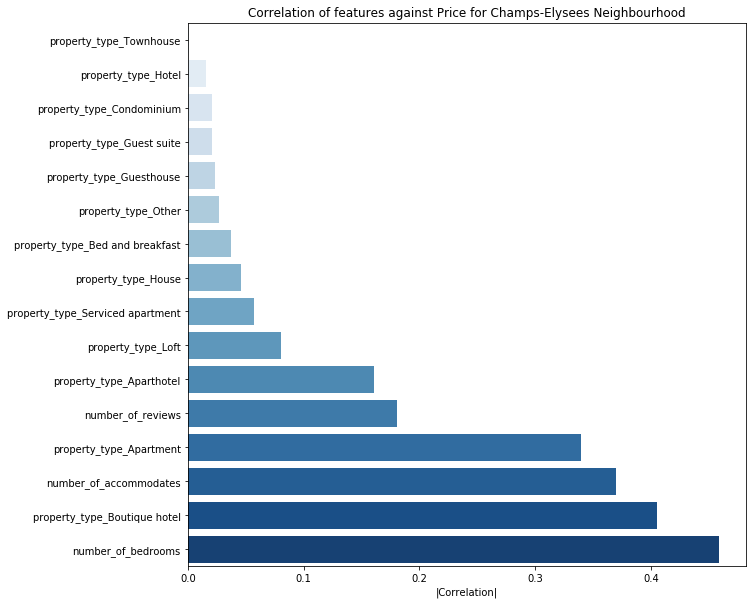


Figure 4

Based on correlation results from Figure 4, we see that number of bedrooms, property type and number of accommodates are the highly correlated features with price. Hence, after selecting the neighborhood, the host can optimize the price by selecting the right features.

Therefore, we will now look at these highly correlated features in the selected neighborhood (Elysee).

* Number of Accommodates and Number of Bedrooms

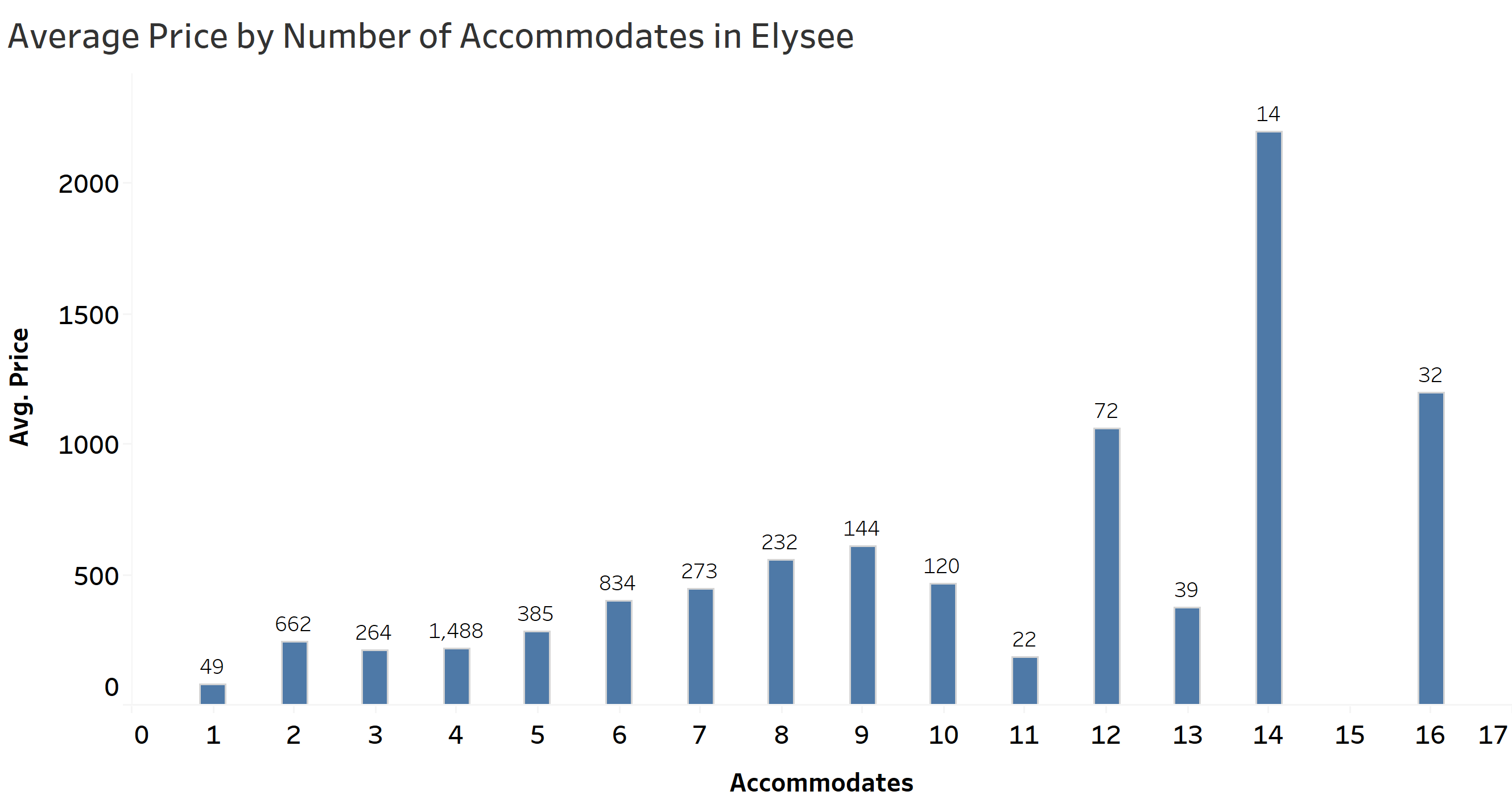


Figure 5

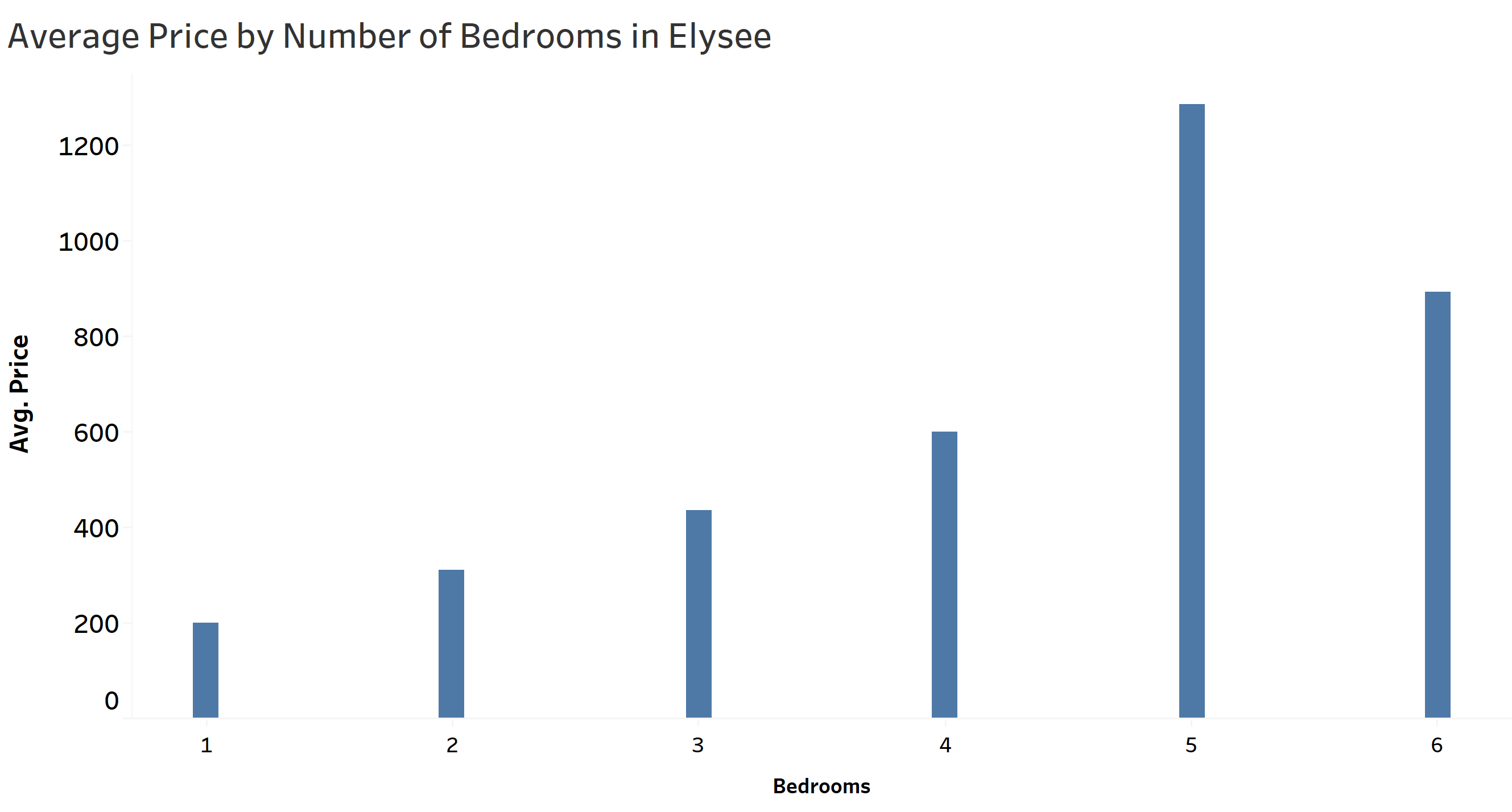


Figure 6

Figure 5 represents the average price per number of accommodates in the Elysee neighborhood with the listings at the top of each bar and figure 6 represents the average price of listings per number of bedrooms in Elysee. Based on these graphs, we see that the average price of the listing is varying after the 10 accommodates and the number of listings is less as compared to that of lower number of accommodates. Therefore, the host should allow for varying number of accommodates from 11 to 14 as he can charge a higher price and the competition is also low for those number of accommodates as the number of listings are low and can keep 5 bedrooms to be able to charge a higher price.

* Property Type

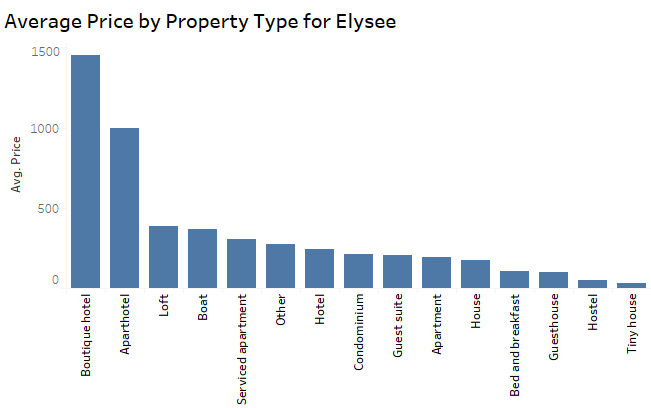


Figure 7

According to Figure 7, in Elysee, Apartment hotels and Boutique hotels are priced the highest. To decide which property type to list, the host can check the demand for these property types in Elysee. Since, we do not have the number of bookings explicitly in the data, we can use the number of reviews as a proxy to get an idea of the demand of the property type in Elysee.

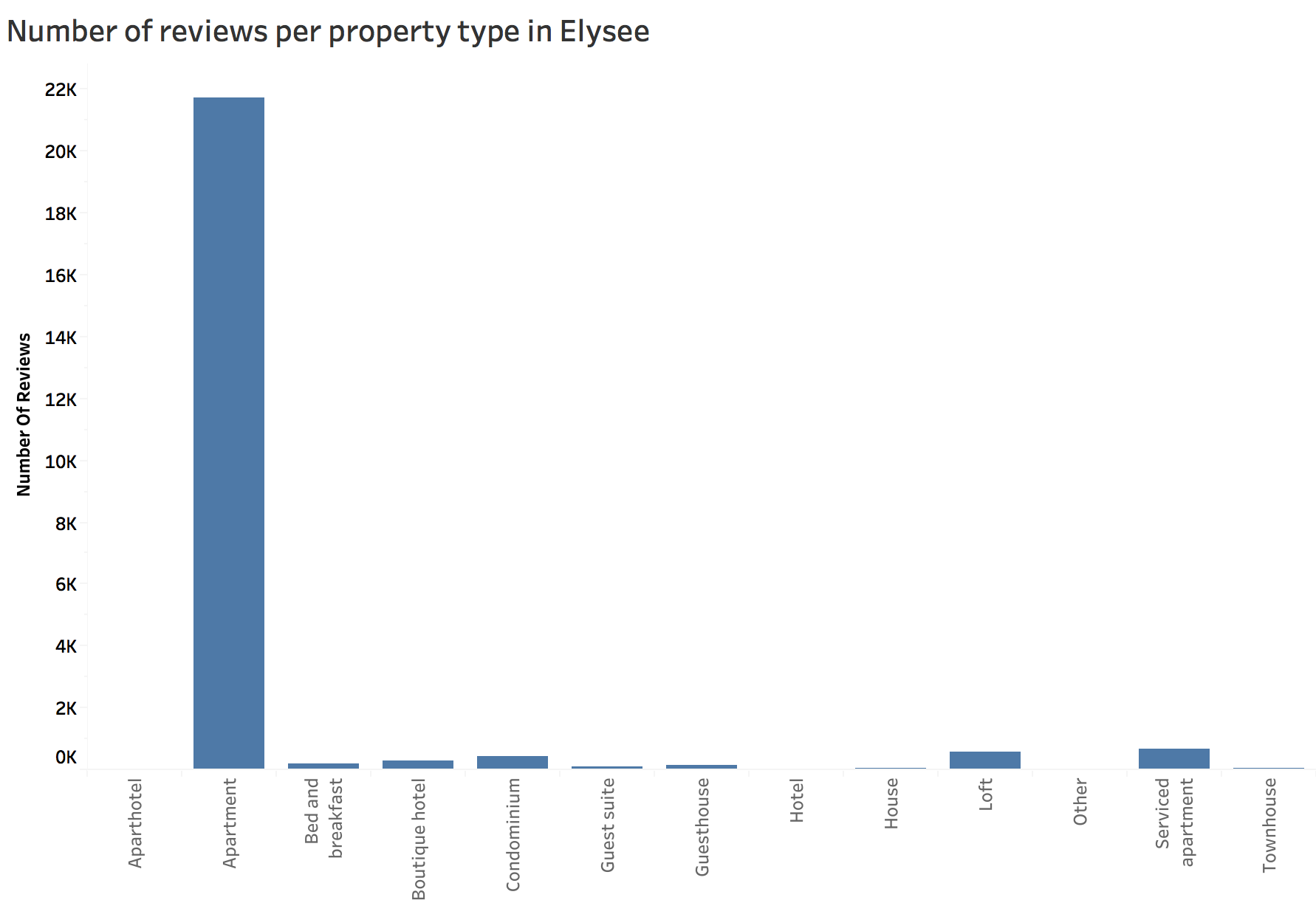


Figure 8

Figure 8 shows that the Apartments are the highest in demand and figure 7 shows low average price for them. Hence, the host can charge a comparatively higher price based on these two factors.

After selecting the property type, let us examine when to list this property on Airbnb in Elysee.

## When to list the property

To determine when to list the property in Elysee, we will consider the number of reviews in each month as a proxy for the demand of listings in that month. Hence, increased reviews on a listing would mean an increased demand for the listing. Let us have a look at the demand for the listings in Elysee monthly.

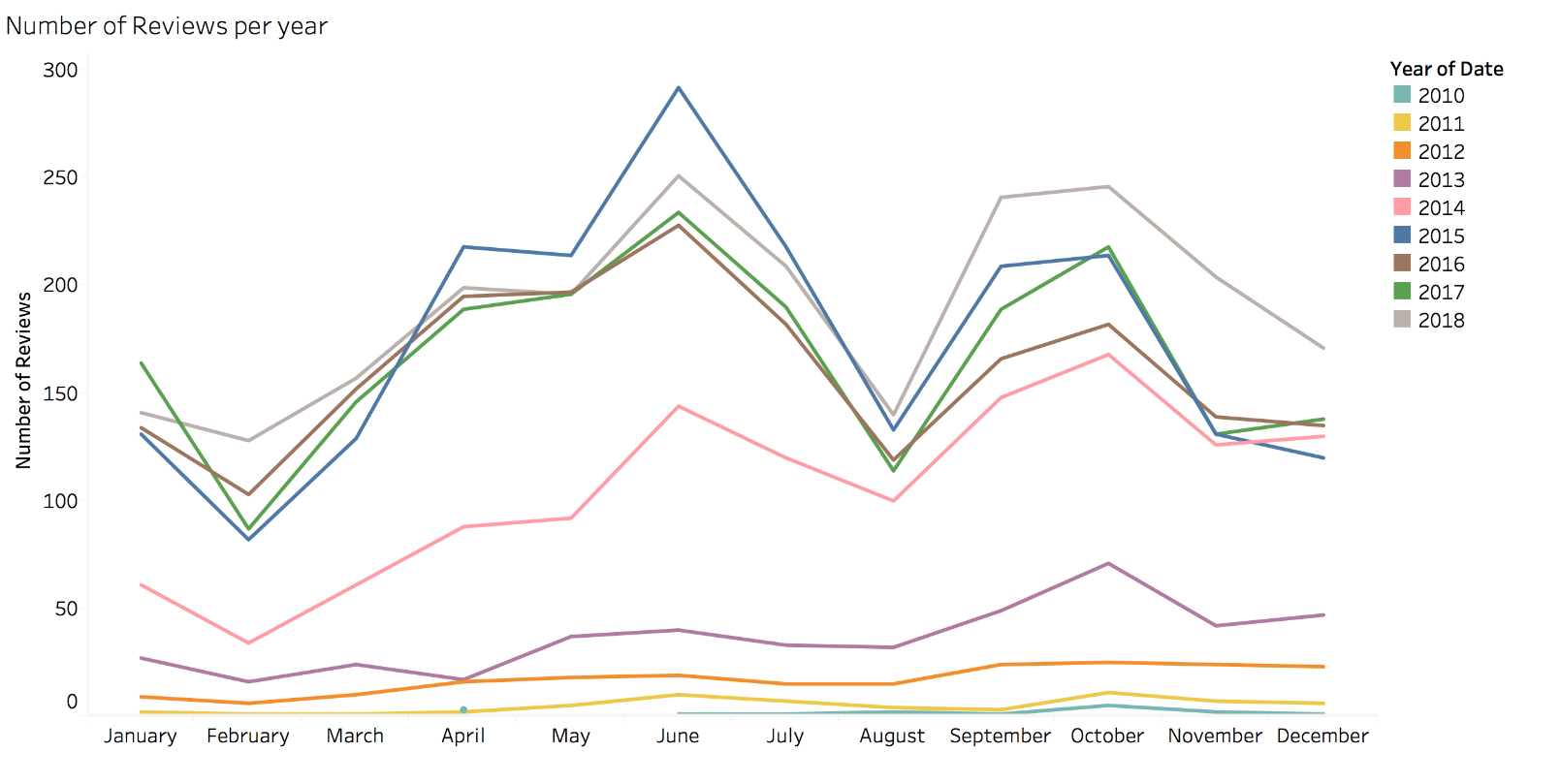


Figure 9

The number of comments shows a similar pattern in each month of recent years, particularly after 2014. We can say that the demand for Elysee listings are highest during the month of June. Looking at the graph, the host in Elysee should list his place in June.

Also, deciding to list in the month of June can be profitable based on the number of events that take place in Elysee including the Elysées Film Festival and Tour De France (Paris Insider Guide).

Listing in August might not result in favor of the host as the demand for listings during that month is quite low. The dip in the demand and the price for Airbnb listings during August is mainly because many Parisians leave Paris during that time and many restaurants are also closed during this time as the owners are out of the city, which makes it a less attractive time of the year (Culture Trip)

## Challenges

Assumptions had to be made for analysis. Since, we did not have the previous year bookings data, we used “number of reviews” feature as a proxy for demand of listings. Also, to obtain number of reviews for each month, we had to combine two data sets.

## Conclusions

In this study, we worked towards solving the problem every host faces while listing a property, i.e. to determine which neighborhood to list in, which property type to list, how many accommodates to allow and when to list that property. We solved this business problem by choosing the neighborhood by analyzing the average and median price and considering the competitiveness in that neighborhood. We found Elysee to be a neighborhood where a host could list. Then, we obtained the factors which were highly correlated with price for considering the features to be included in the listing. We found number of accommodates, number of bedrooms, and property type to be the most relevant features to be included in a listing in Elysee. Lastly, we answered the question of when to list by looking at the demand of the listings monthly and found June to be that month.

Overall, our analysis conveys the ways a host can maximize his or her revenue which in turn would increase the revenue of Airbnb.

# References

*Airbnb - Elysee*. (n.d.). Retrieved from https://www.airbnb.com/locations/paris/champs-elysees

*Airbnb Financials*. (n.d.). Retrieved from Airbnb: https://www.airbnb.com/b/financials

Culture Trip. (n.d.). *When to Go to Paris: A Month-By-Month Guide*. Retrieved from https://theculturetrip.com/europe/france/paris/articles/when-to-go-to-paris-a-month-by-month-guide/

Few, S. (February 2008). Practical Rules for Using Color in Charts. *Perceptual Edge - Visual Business Intelligence Newsletter*.

*Inside Airbnb*. (n.d.). Retrieved from http://insideairbnb.com/get-the-data.html

*Paris Insider Guide*. (n.d.). Retrieved from https://www.parisinsidersguide.com/paris-events.html