

Airbnb Listings in Paris*

Irene Huynh

March 5, 2024

Introduction

In this case study, I look at Airbnb listings in Paris, France as at 12 December 2023 and conduct an exploratory data analysis. The dataset is from Inside Airbnb (Cox 2023).

Distribution and properties of individual variables results

Prices

Figure 1 shows the distribution of prices. Since there are outliers, I will consider the log scale.

Figure 2 displays the distribution of prices using the log scale for prices more than \$1,000. I will focus on prices that are less than \$1,000 as more data appears to lie there.

Figure 3 reveals that most properties have a nightly price less than \$250. There is some bunching of prices here. Similar to Telling Stories with Data Chapter 11 (Alexander, n.d.), it might be that this is happening around numbers ending in zero or nine. Thus, Figure 4 focuses on prices between \$90 and \$210, and the bins are smaller. Figure 4 reveals that properties with a nightly price between \$90 and \$115 is the most common.

*Code and data are available at: https://github.com/IreneeeH/Airbnb_Listings_in_Paris.git

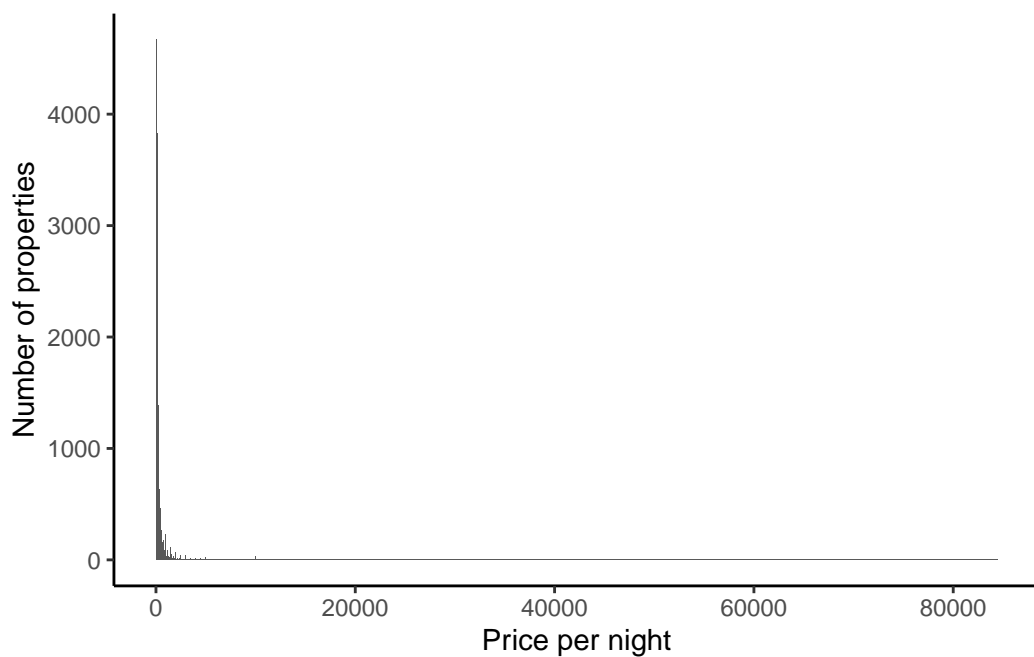


Figure 1: Distribution of prices for Paris Airbnb rentals in December 2023

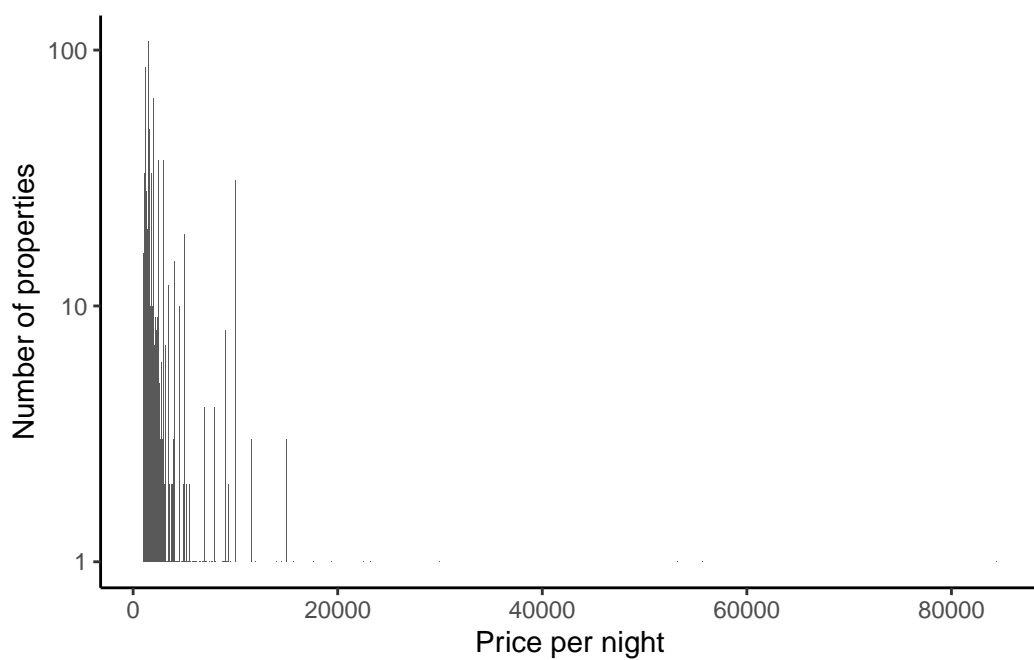


Figure 2: Distribution of prices using the log scale for prices more than \$1,000 for Paris Airbnb rentals in December 2023

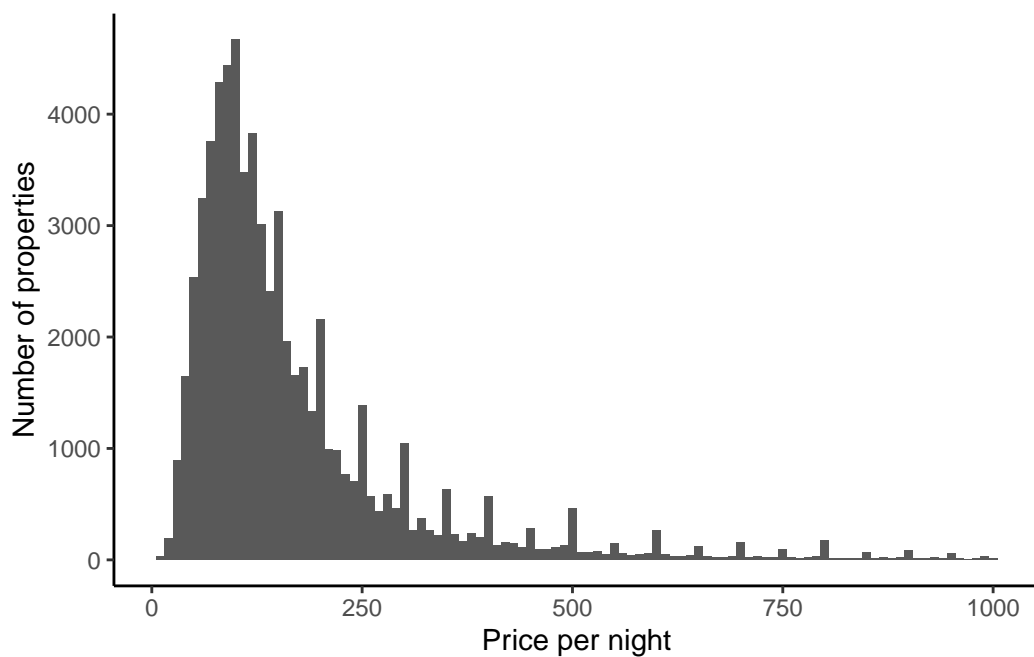


Figure 3: Distribution of prices less than \$1,000 for Paris Airbnb rentals in December 2023

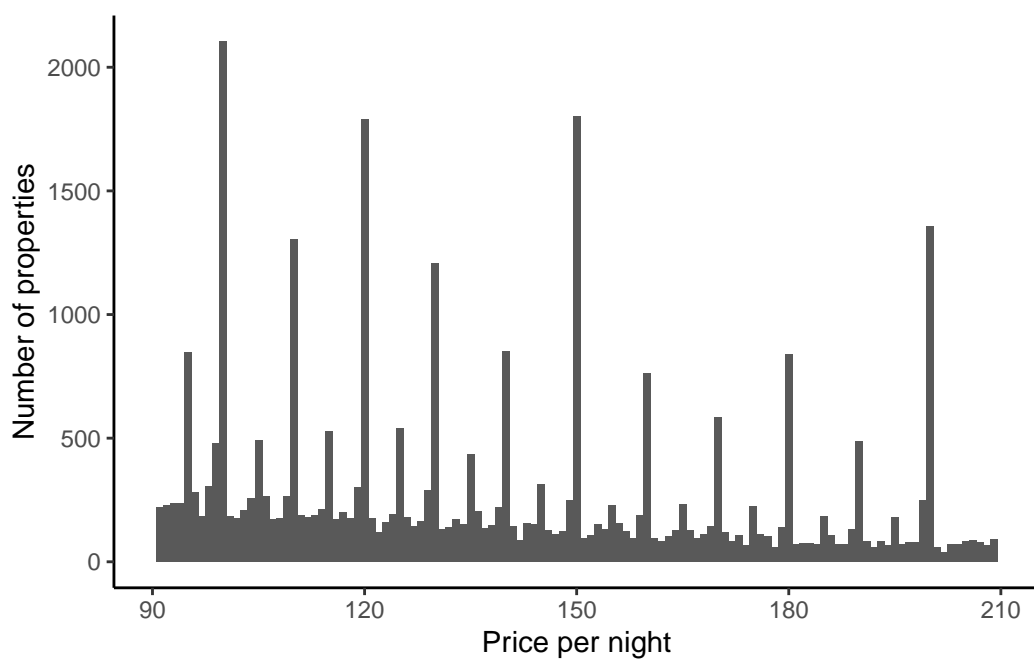


Figure 4: Distribution of prices between \$90 and \$210 for Paris Airbnb rentals in December 2023

Reviews

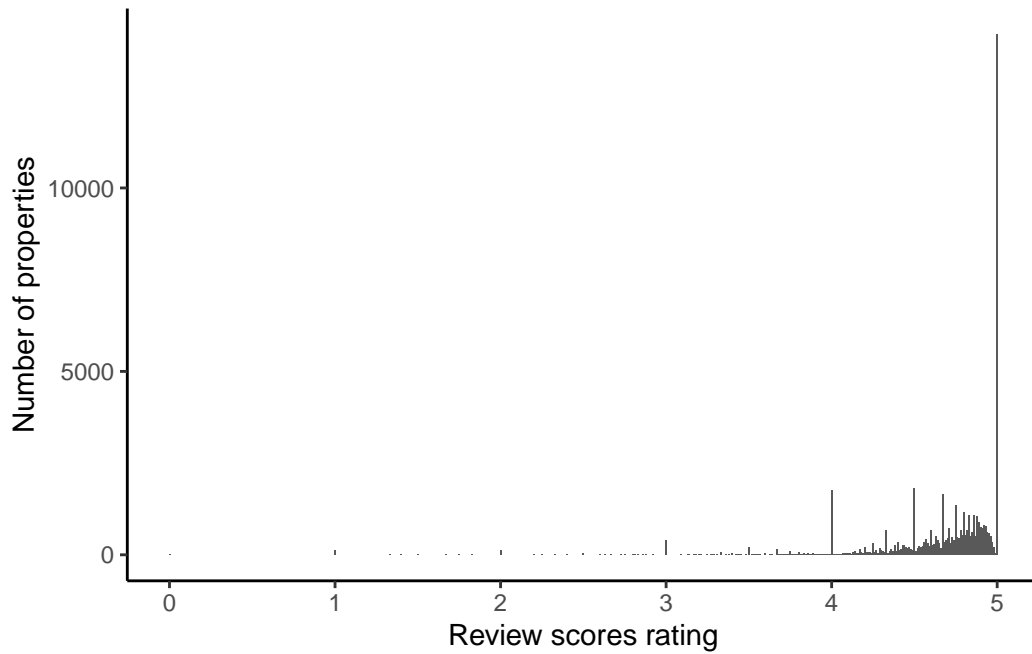


Figure 5: Distribution of review scores rating for Paris Airbnb rentals in December 2023

Figure 5 demonstrates that although guests can give ratings from one to five, it is clear that it is effectively a binary: in almost all cases, the rating is either five stars or not five stars.

Figure 6 is similar to Figure 5; however, it focuses on reviews that are not missing and the main review score.

Figure 7 illustrates the relationship between the properties with NA response time and review score. There are a lot that have an overall review of 5.

Figure 8 reveals that there are a lot of missing values in the data.

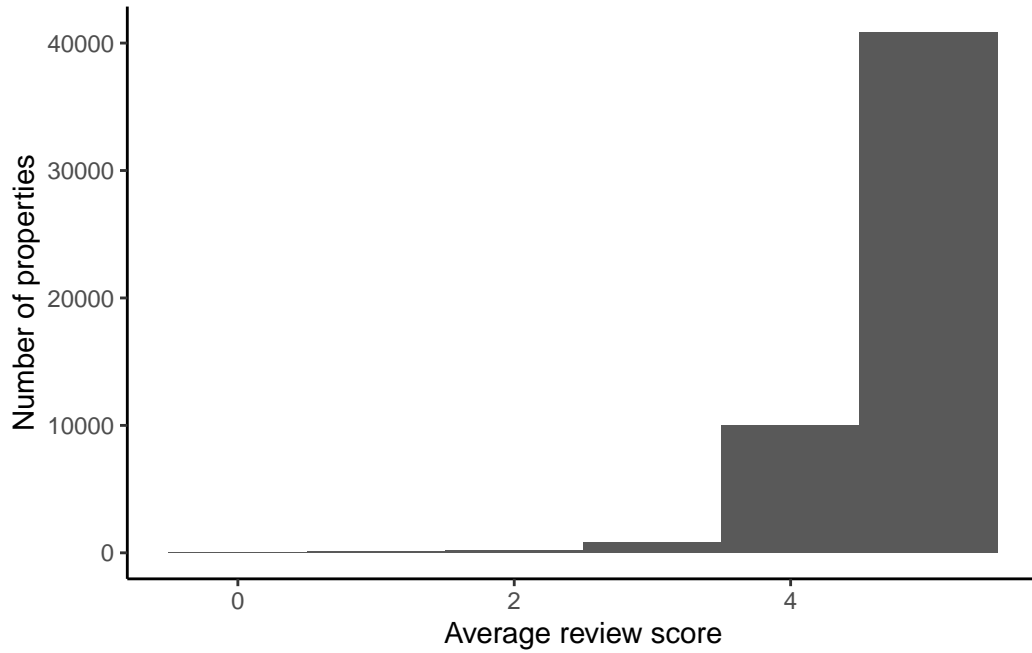


Figure 6: Average distribution of review scores rating for Paris Airbnb rentals in December 2023

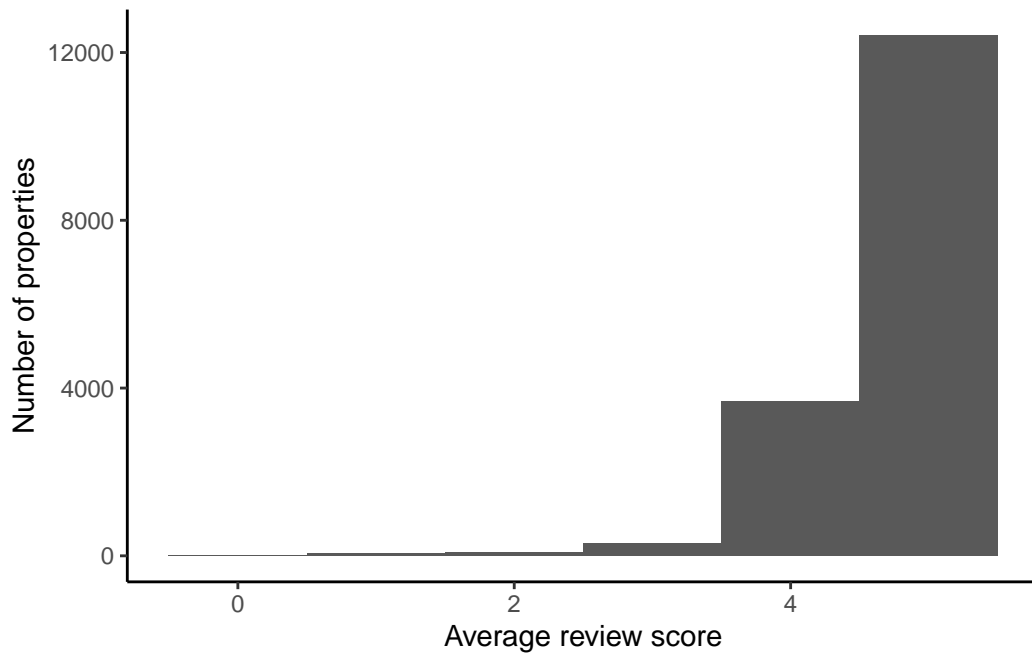


Figure 7: Average distribution of review scores for properties with NA response time for Paris Airbnb rentals in December 2023

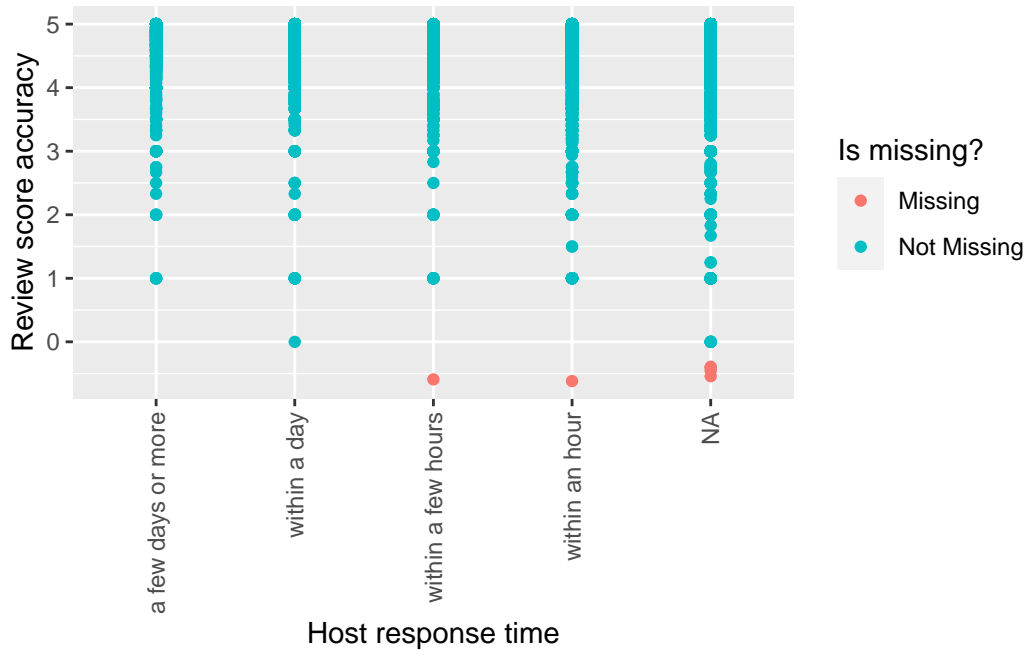


Figure 8: Missing values in Paris Airbnb data, by host response time

Hosts

From Figure 9, we can see that most hosts have only one property, and a smaller number of hosts have somewhere in the 5-100 properties range. This is expected as hosts do not generally have more than 100 properties.

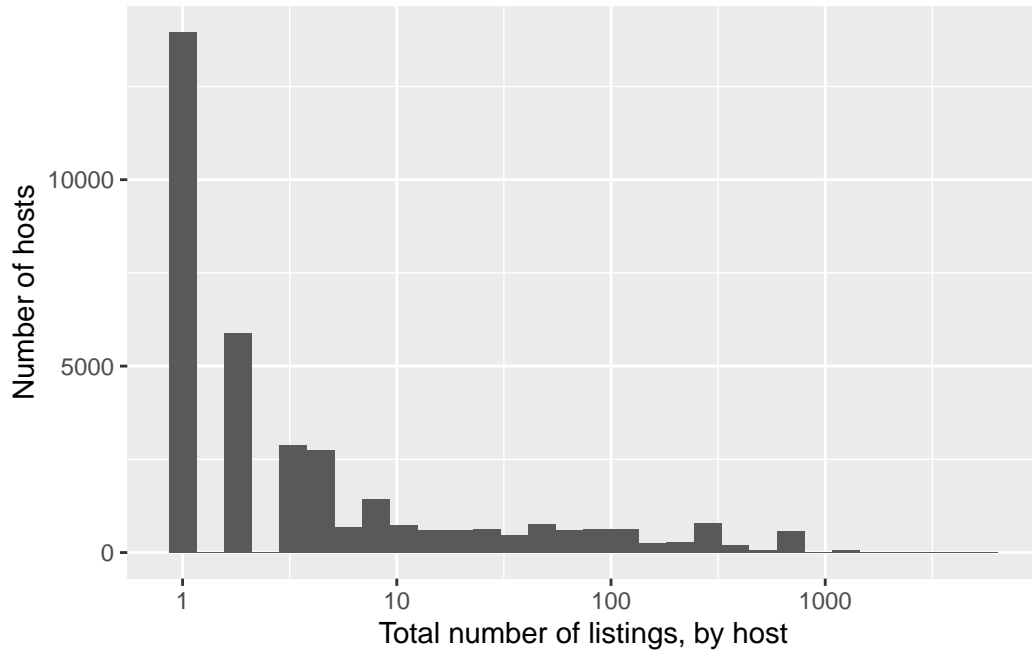


Figure 9: Distribution of the number of properties a host has on Airbnb for Paris Airbnb rentals in December 2023

Relationships between variables

Figure 10 shows that superhosts receive consistently high ratings, and listings that are below \$250 have a range mostly between three to four stars and some one to two star ratings. For listings that are over \$500, most guests rated them four to five stars. This is expected as higher priced listings are often more luxurious and provide more ammentiies. Additionally, there are fewer ratings for properties over \$500, which is expected as individuals would rather rent a cheaper property.

As Telling Stories with Data Chapter 11 mentions, one of the aspects that may make someone a superhost is how quickly they respond to inquiries (Alexander, n.d.). I explore this in Table 1.

Table 1 demonstrates that each of the levels is positively associated with the probability of being a superhost. However, having a host that responds within an hour is associated with individuals that are superhosts in our dataset.

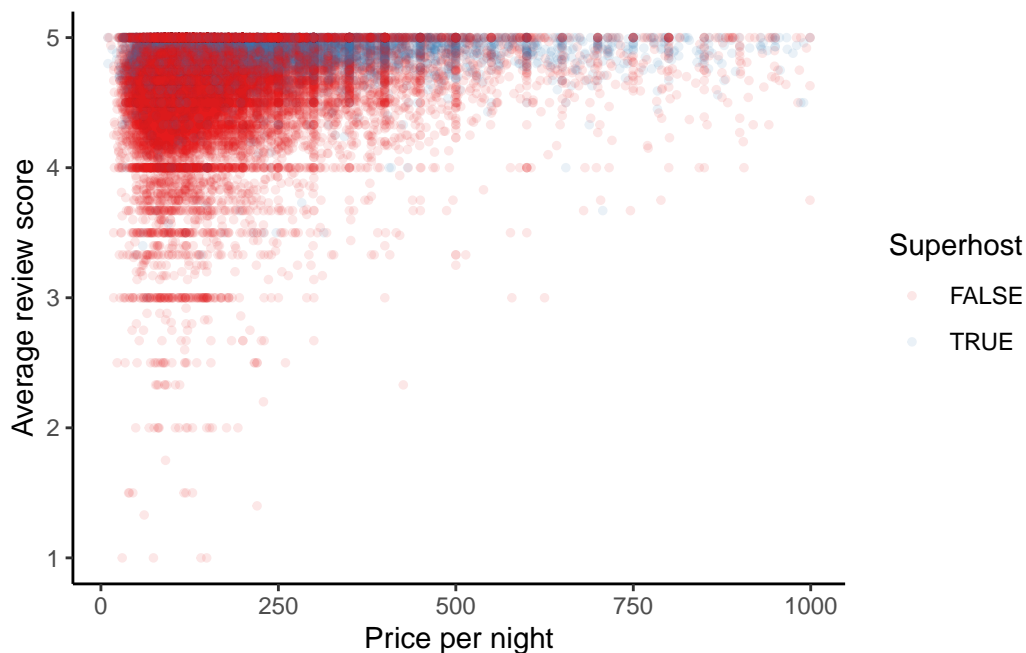


Figure 10: Relationship between price and review and whether a host is a superhost, for Paris Airbnb rentals in December 2023

Table 1: Explaining whether a host is a superhost based on their response time

	(1)
(Intercept)	−18.384 (0.377)
host_response_timewithin a day	2.283 (0.210)
host_response_timewithin a few hours	3.015 (0.209)
host_response_timewithin an hour	3.190 (0.208)
review_scores_rating	3.021 (0.065)
Num.Obs.	35 445
AIC	37 601.0
BIC	37 643.4
Log.Lik.	−18 795.504
F	674.466
RMSE	0.43

References

- Alexander, Rohan. n.d. *Telling Stories - 11 Exploratory Data Analysis: Prime Ministers of the United Kingdom*. <https://tellingstorieswithdata.com/11-eda.html#airbnb-listings-in-london-england>.
- Cox, Murray. 2023. *Inside Airbnb—Toronto Data*. <http://insideairbnb.com/get-the-data.html>.