Airbnb listings in Paris, France*

Xu Qi

March 4, 2024

Table of contents

1 Introduction

This paper is a exploratory data analysis of Airbnb listings in Paris, France, as at 12 December 2023.

2 Data and Processing

The dataset is from Inside Airbnb (Cox 2021). Data was collected and analyzed using the statistical programming software R (R Core Team 2023), with additional support packages including `tidyverse` (Wickham et al. 2019), `ggplot2` (Wickham 2016), `dplyr` (Wickham et al. 2023), `janitor` (Firke 2023), `knitr` (Xie 2014) , `naniar` (Tierney and Cook 2023) , `patchwork` (Pedersen 2024) , `arrow` (Richardson et al. 2024), `here` (Müller and Bryan 2020) and `modelsummary` (Arel-Bundock 2022).

---

*Code and data from this analysis are available at: https://github.com/xuqi2002/toronto_library_visit

Table 1

	(1)
(Intercept)	−16.262
	(0.481)
host_response_timewithin a day	2.019
	(0.211)
host_response_timewithin a few hours	2.695
	(0.210)
host_response_timewithin an hour	2.972
	(0.209)
review_scores_rating	2.624
	(0.089)
Num.Obs.	22 047
AIC	24 165.0
BIC	24 205.0
Log.Lik.	−12 077.507
RMSE	0.43

3 Results

3.1 Distribution and properties of individual variables

3.2 Relationships between variables

References

Arel-Bundock, Vincent. 2022. "modelsummary: Data and Model Summaries in R." *Journal of Statistical Software* 103 (1): 1–23. https://doi.org/10.18637/jss.v103.i01.

Cox, Murray. 2021. "Get the Data." *Insideairbnb.com.* http://insideairbnb.com/get-the-data.html.

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data.* https://CRAN.R-project.org/package=janitor.

Müller, Kirill, and Jennifer Bryan. 2020. *Here: A Simpler Way to Find Your Files.* https://CRAN.R-project.org/package=here.

Pedersen, Thomas Lin. 2024. *Patchwork: The Composer of Plots.* https://CRAN.R-project.org/package=patchwork.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.
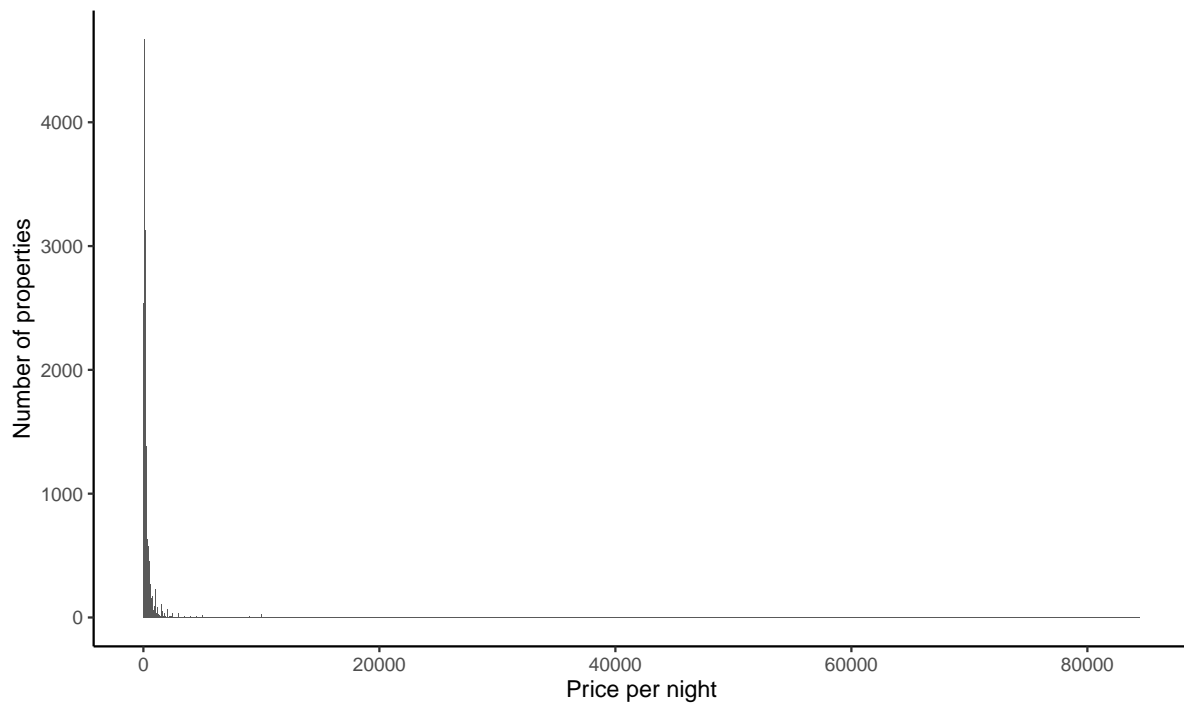
Figure 1.1 Distribution of prices

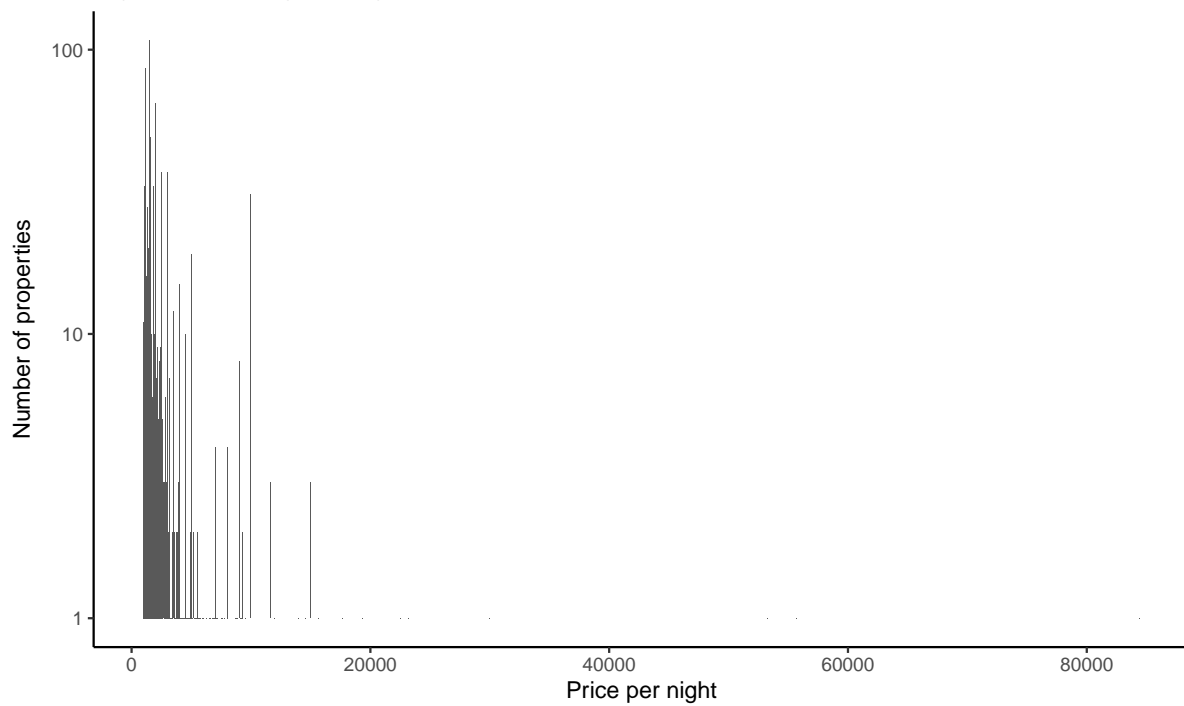Figure 1.2 Using the log scale for prices more than $1,000

Figure 1: Distribution of prices of London Airbnb rentals in March 2023

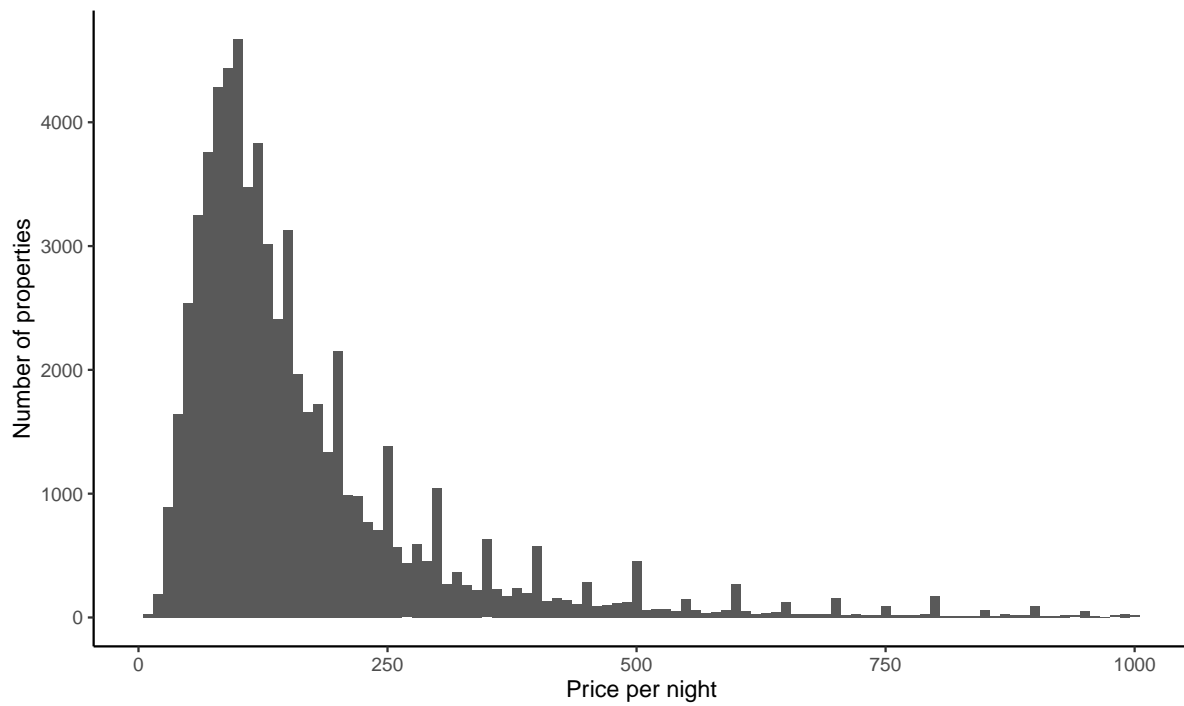Figure2.1 Prices less than $1,000 suggest some bunching

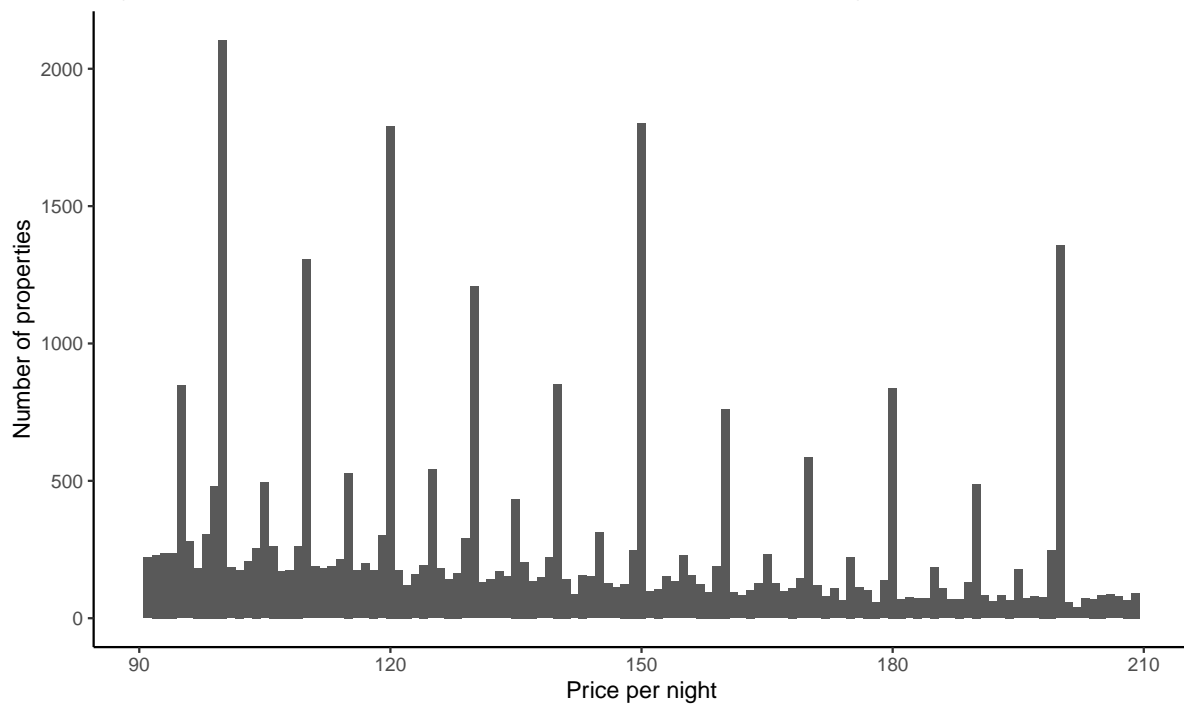Figure 2.2 Prices between $90 and $210 illustrate the bunching more clearly

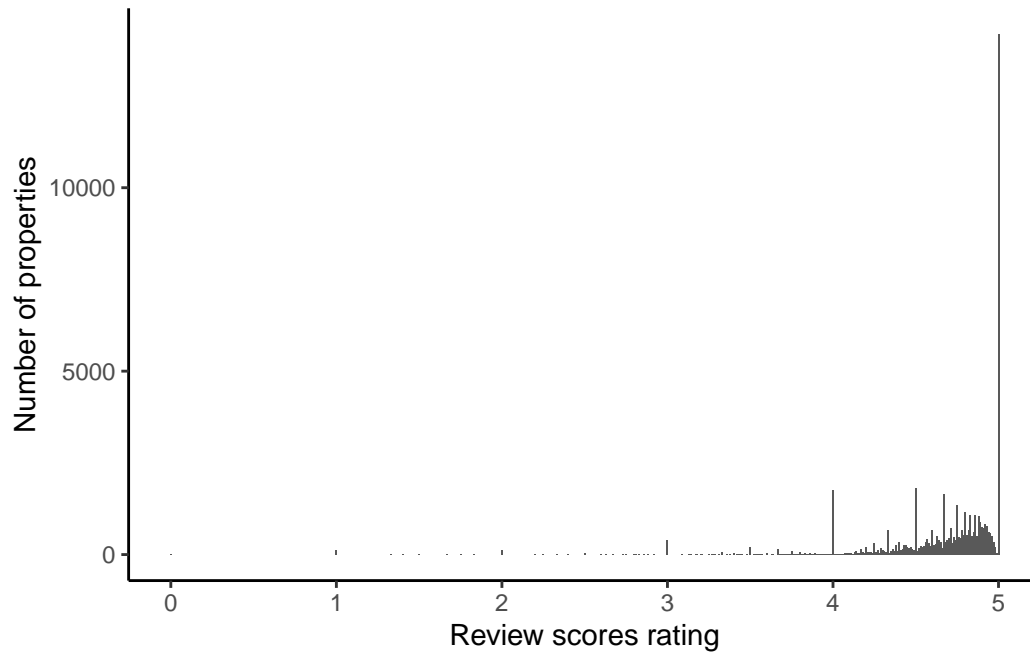Figure 2: Distribution of prices for Airbnb listings in London in March 2023

4

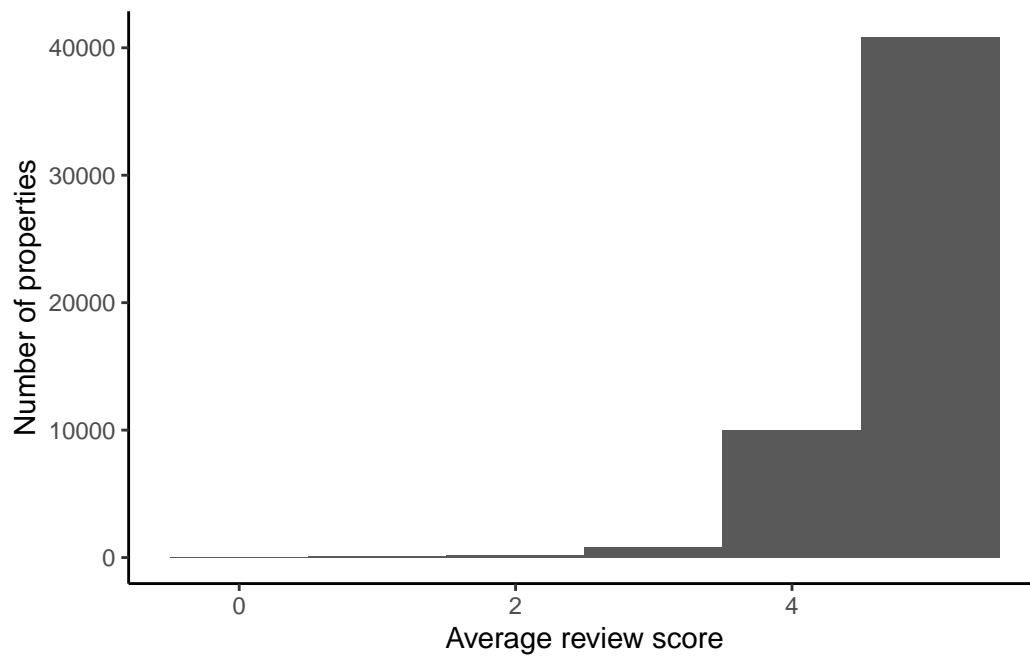Figure 3: Distribution of review scores rating for Paris Airbnb rentals in December 2023

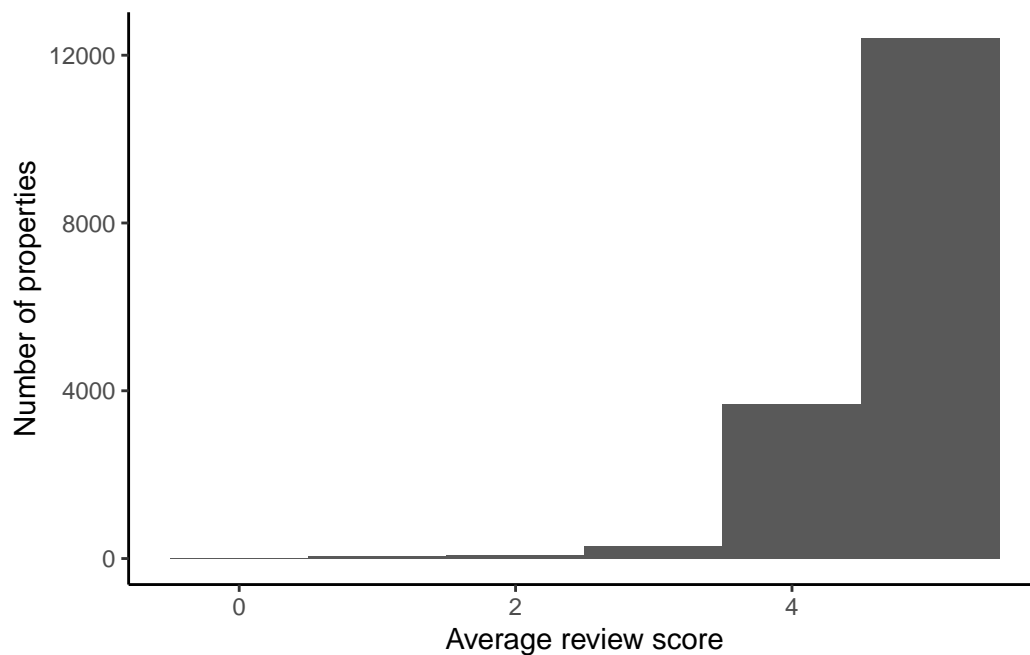Figure 4: Distribution of review scores for Paris Airbnb rentals in December 2023

Figure 5: Distribution of review scores for properties with NA response time, for Paris Airbnb rentals in December 2023
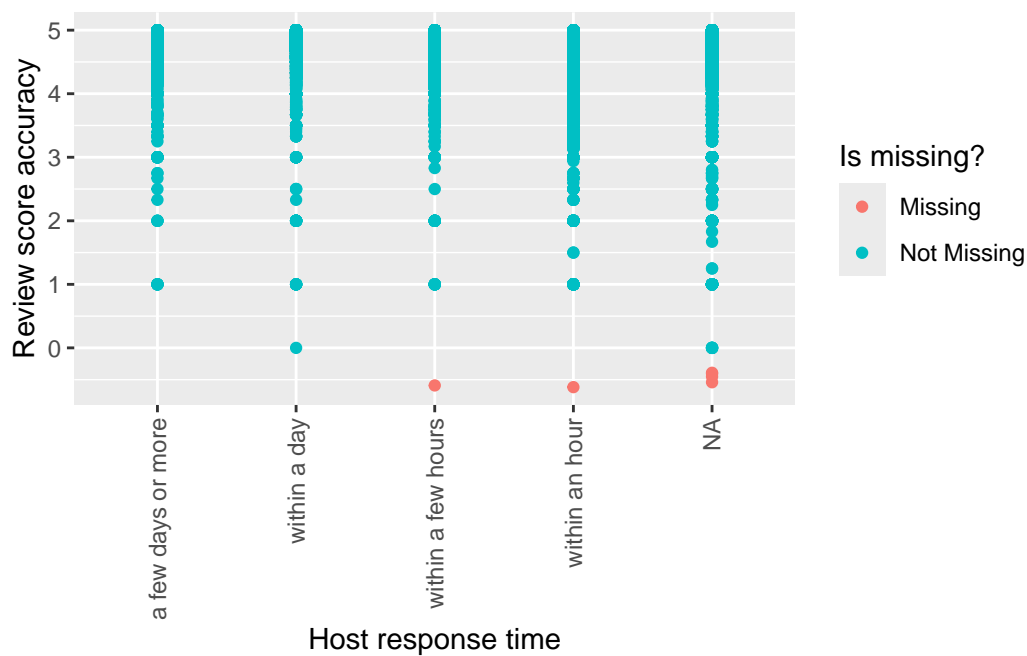

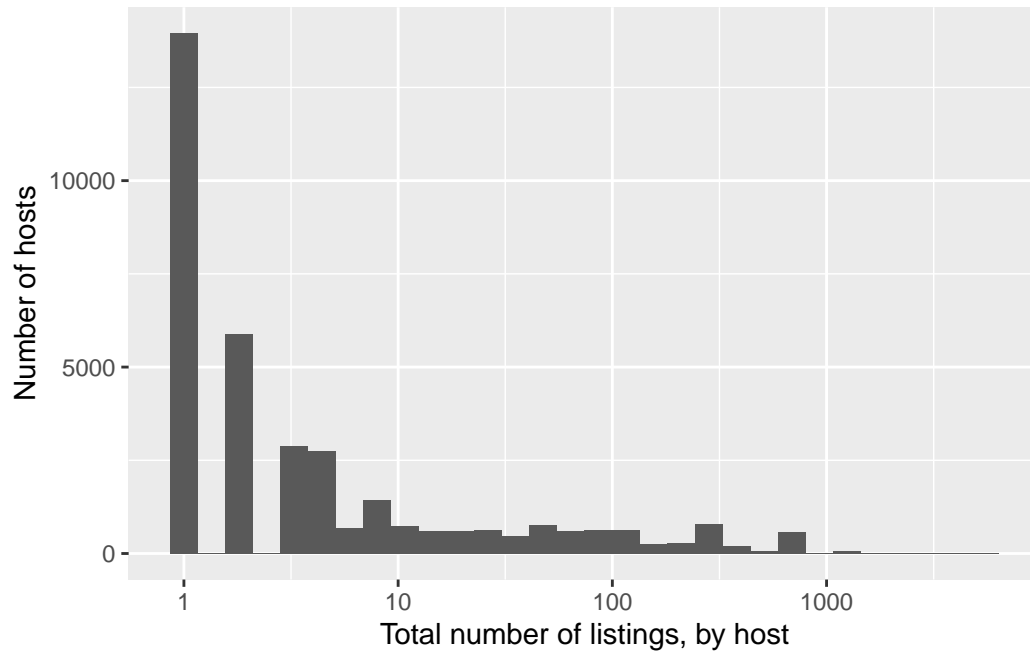Figure 6: Missing values in Paris Airbnb data, by host response time

Figure 7: Distribution of the number of properties a host has on Airbnb, for Paris Airbnb rentals in December 2023
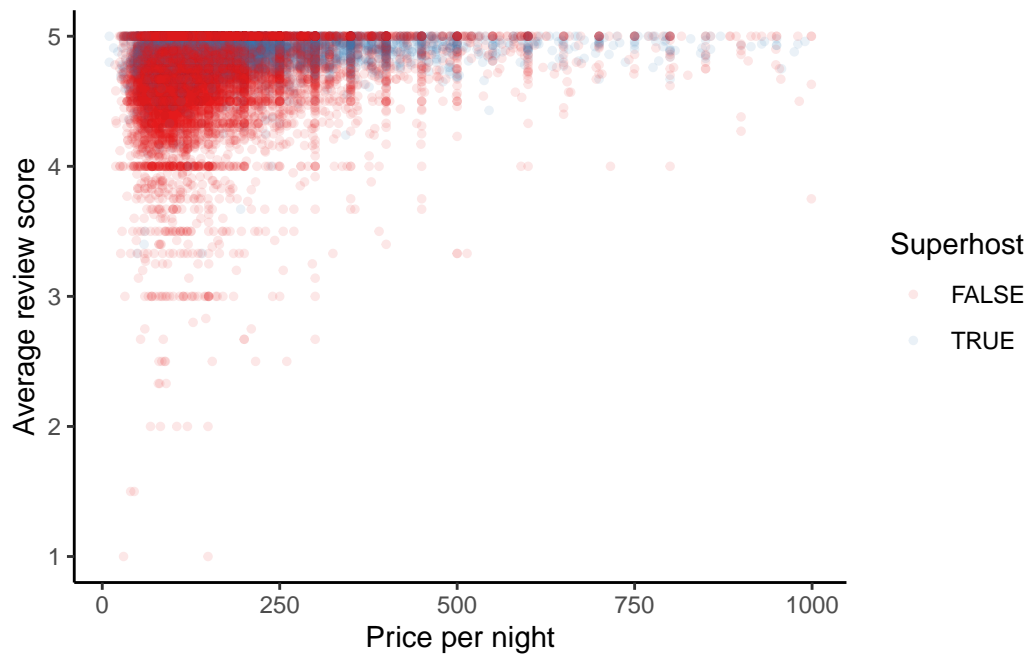

Figure 8: Relationship between price and review and whether a host is a superhost, for Paris Airbnb rentals in December 2023

Richardson, Neal, Ian Cook, Nic Crane, Dewey Dunnington, Romain François, Jonathan Keane, Dragoș Moldovan-Grünfeld, Jeroen Ooms, Jacob Wujciak-Jens, and Apache Arrow. 2024. *Arrow: Integration to 'Apache' 'Arrow'*. https://github.com/apache/arrow/.

Tierney, Nicholas, and Dianne Cook. 2023. "Expanding Tidy Data Principles to Facilitate Missing Data Exploration, Visualization and Assessment of Imputations." *Journal of Statistical Software* 105 (7): 1–31. https://doi.org/10.18637/jss.v105.i07.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2023. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Xie, Yihui. 2014. *Knitr: A Comprehensive Tool for Reproducible Research in R*. Edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC. http://www.crcpress.com/product/isbn/9781466561595.