How far can you go for love on Airbnb? - The effect of Valentine's Day on Airbnb listings in Europe

Anouk Bor, Eva Bos, Bi Xuan Guo, Mandana Khabbazi, Indi Wieggers

October 17, 2022

Introduction

In order to answer our research question, we used data from InsideAirbnb. We collected listing data from three major European cities where we suspected that Valentine's Day is a big thing here, namely Rome, Paris and Madrid. After extracting the datasets for each city, we cleaned these data. In this stage, redundant columns were removed and variables related to the analysis were added. Lastly, the data sets from the individual cities were merged into one complete dataset.

This report provides a description of all variables used in the analysis, including relevant summary statistics. Furthermore, the report includes several plots visualizing the key variables from the analysis. The main section of the report contains the three analyses that answer the research question, where the results are discussed. Ultimately, a conclusion on the effect of Valentine's Day on Airbnb listings is drawn based on these analyses.

The following questions will be answered:

1. What is the influence of Valentine's Day on the price of Airbnb listings?
2. What is the influence of Valentine's Day on the number of booked Airbnb listings?
3. Compared to 'normal days', are relatively more Airbnb accommodations with one/two beds booked during Valentine's Day?

Variable Descriptives

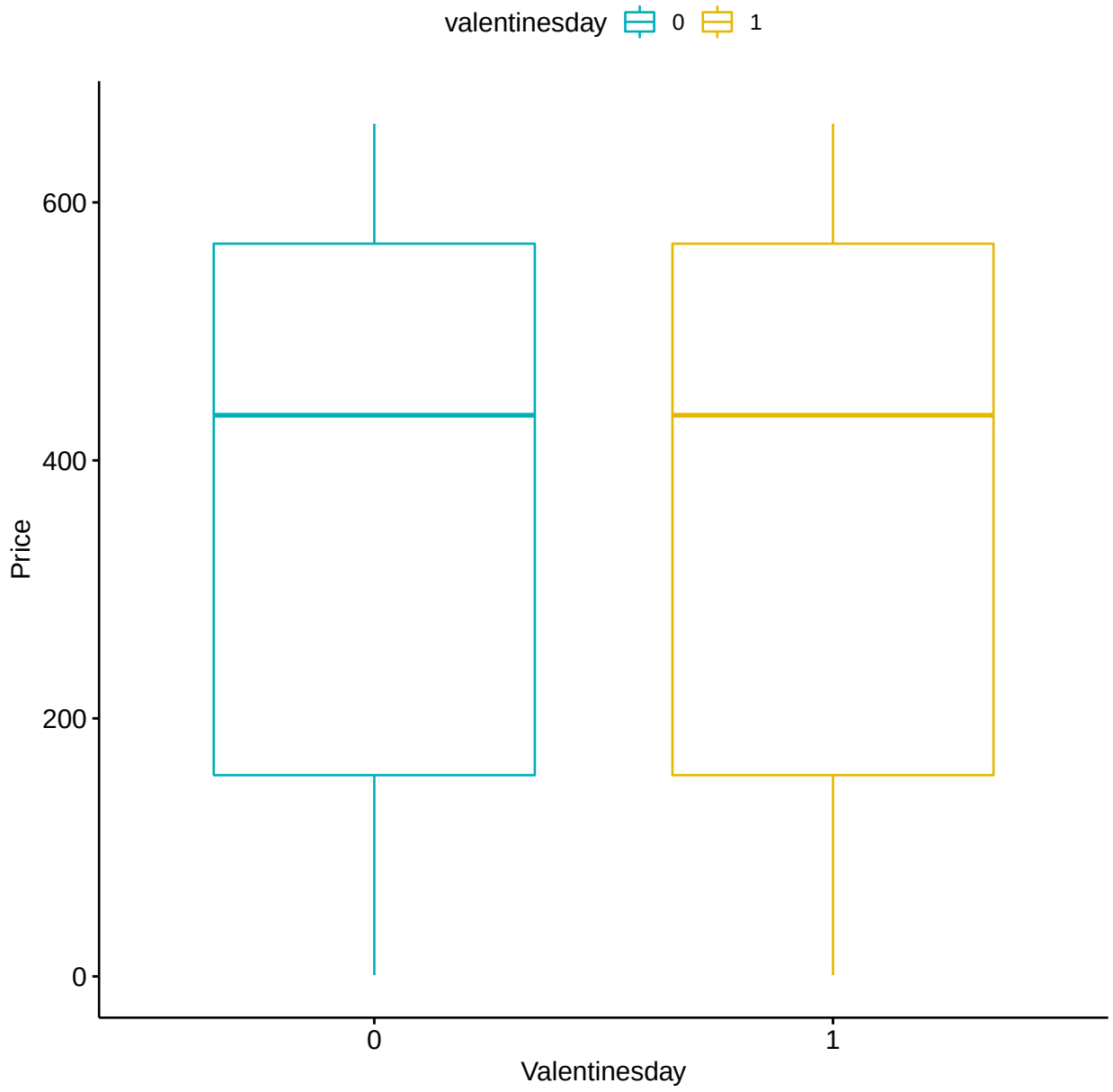
Prices

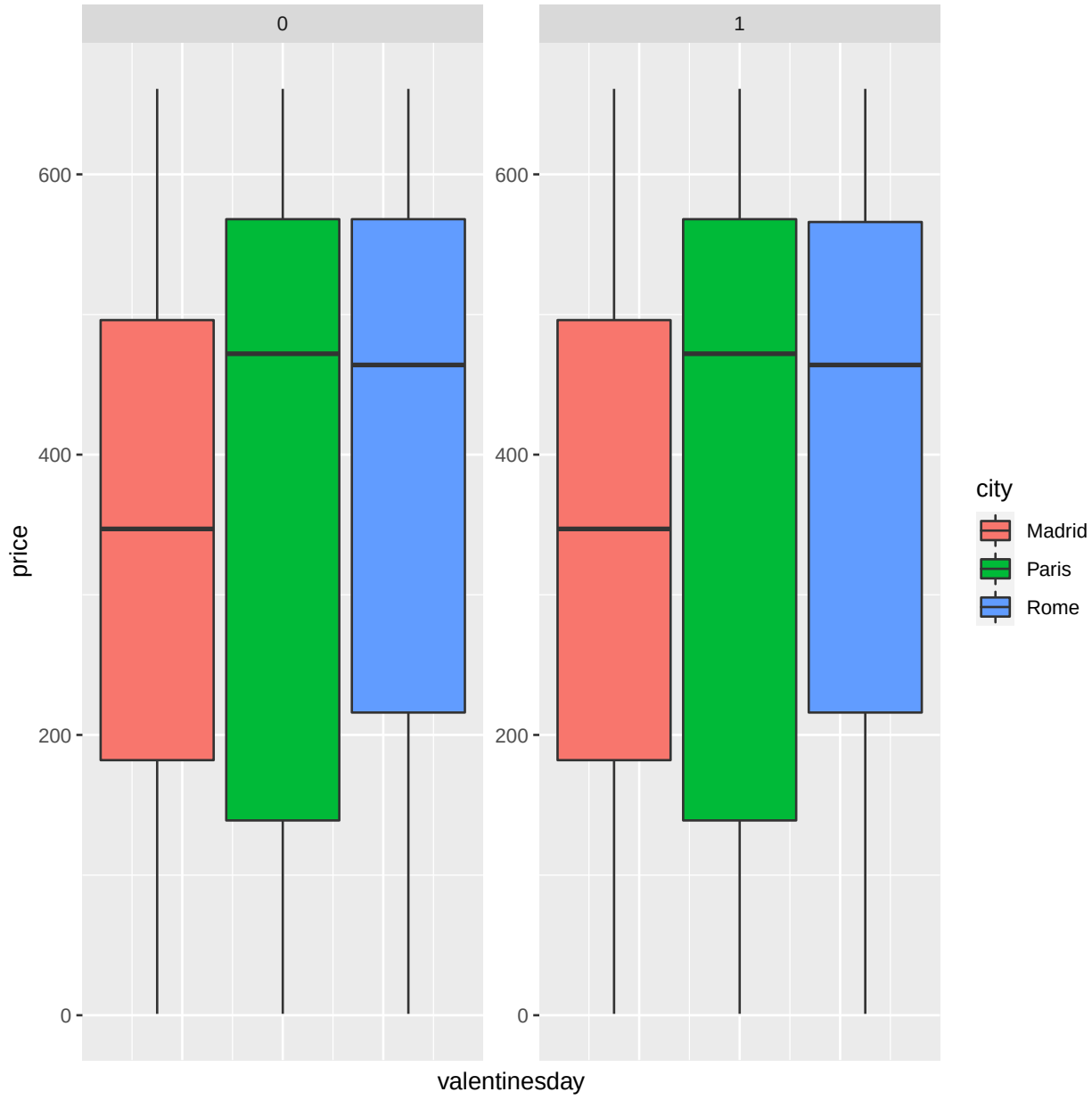
In order to analyze the influence of Valentine's Day on the prices of Airbnb accomodations, we used the metric variable 'price'. Furthermore, we created a dummy variable 'Valentinesday', which takes the value of 1 on Valentine's Day and 0 on any other day. The average price of an Airbnb listing from the full data is 379.5378929. The mean of the 'valentinesday' dummy variable is 0.0359785.

Table 1: Frequencies dummy variable 'valentinesday'

X	Var1	Freq
1	0	1437705
2	1	53657

We compare the price on Valentine's Day with the price on days two weeks before and after the celebration. The outcome of this comparison can be seen in the boxplot below. According to the generated boxplot, we can conclude that the difference between the averages price of accomodations on Valentine's Day and 'normal' days is very small. Furthermore, comparing the cities, we can conclude that the average price for Airbnb accomodations is highest in Paris. The average price is lowest in Madrid. This is verified by using the lapply function of the price per cities.





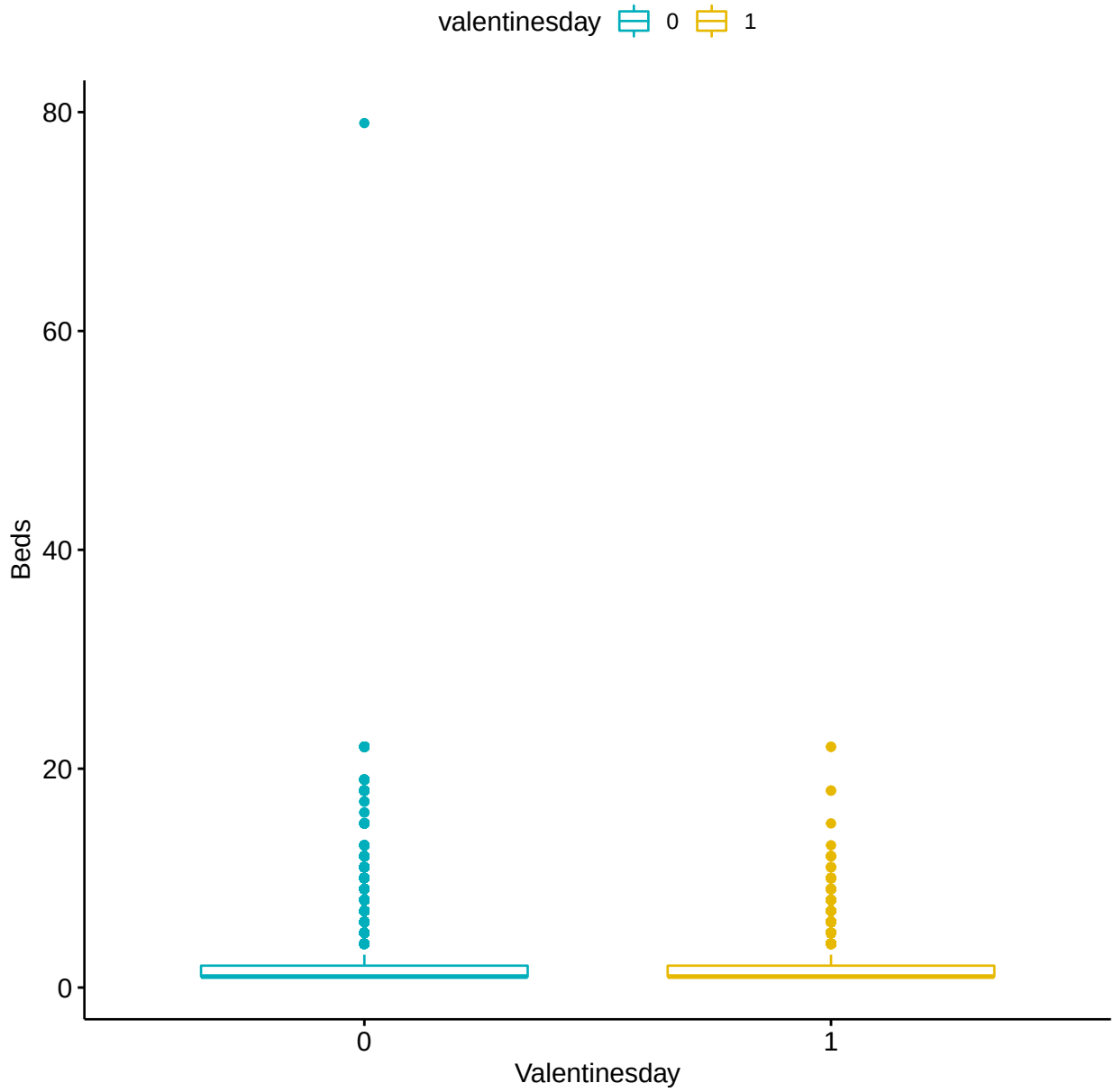
Bookings

To analyse the impact of Valentine's Day on bookings, we created a dummy variable for booked accommodations. The dummy variable booked consists of two categories; an accommodation that has not been booked (0) and an accommodation that has been booked (1). Looking at the mean 0.6977649 in the dataset, we conclude that Airbnbs are more booked than not booked in the chosen period of this dataset (31 January - 28 February). This conclusion can be verified by printing out the table and the histogram for the variable 'booked'.

A total of 53657 Airbnbs were booked in Rome, Madrid and Paris on Valentine's Day. If we take all other data from our dataset, an average of 53826 accommodations were booked for the cities. Therefore, we can conclude that there are not more accommodations booked via Airbnb during Valentine's Day, but even slightly less.

Beds

The variable beds used in this analysis consist of two categories; zero or more than two beds (0) and one or two beds (1). Looking at the mean, 0.8399752, we can conclude that there are more than average Airbnbs with one or two beds in the dataset. The table below shows that on Valentine's Day, on average, more Airbnbs with one or two beds are booked than Airbnbs with more than two beds. This is the same for a day in February when it is not Valentine's Day. This can be checked from the boxplot below. We do need to take into account that there are many outliers.



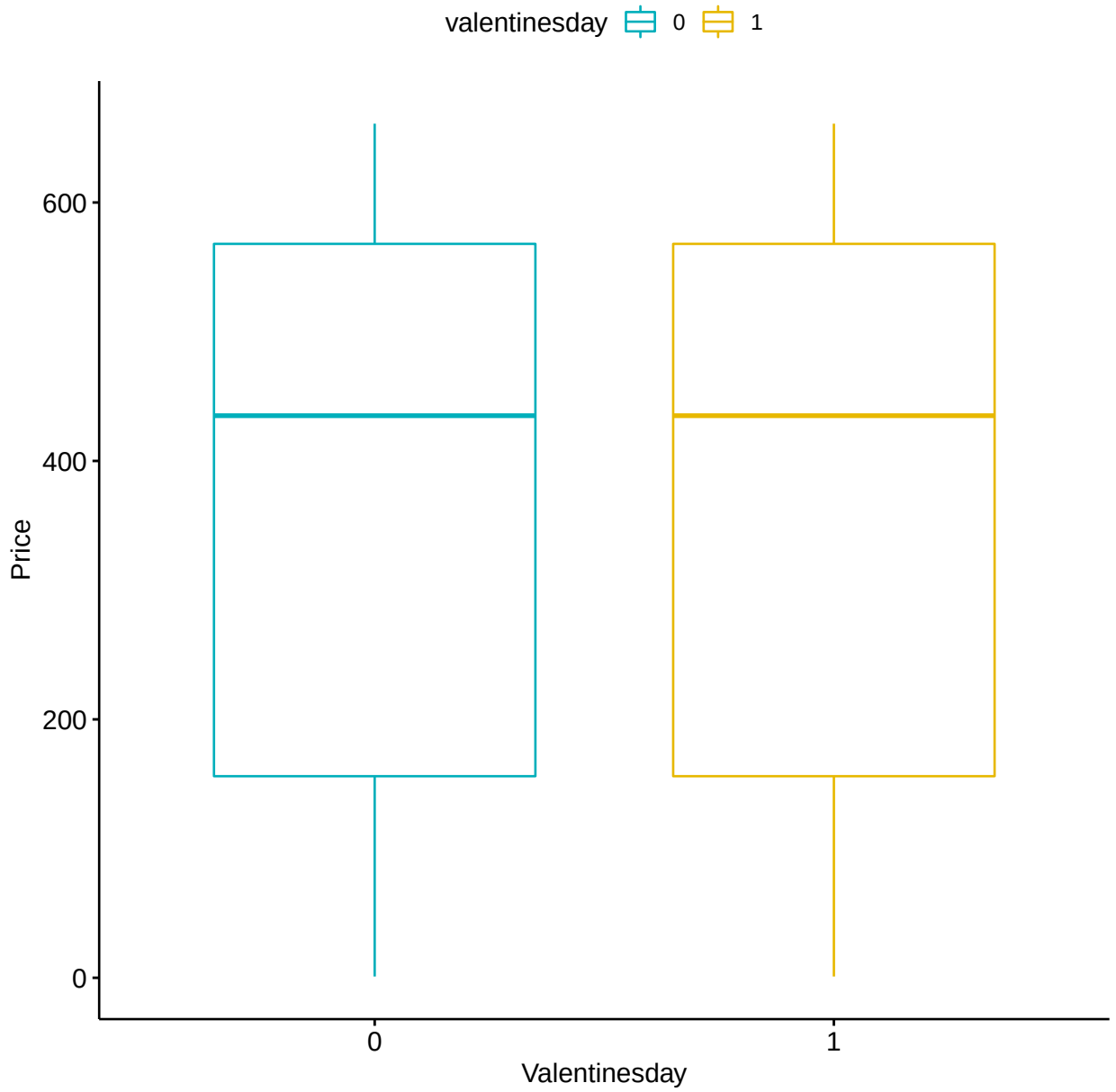
Checking assumptions

Before the analyses are conducted to answer the research questions, some assumptions are first checked.

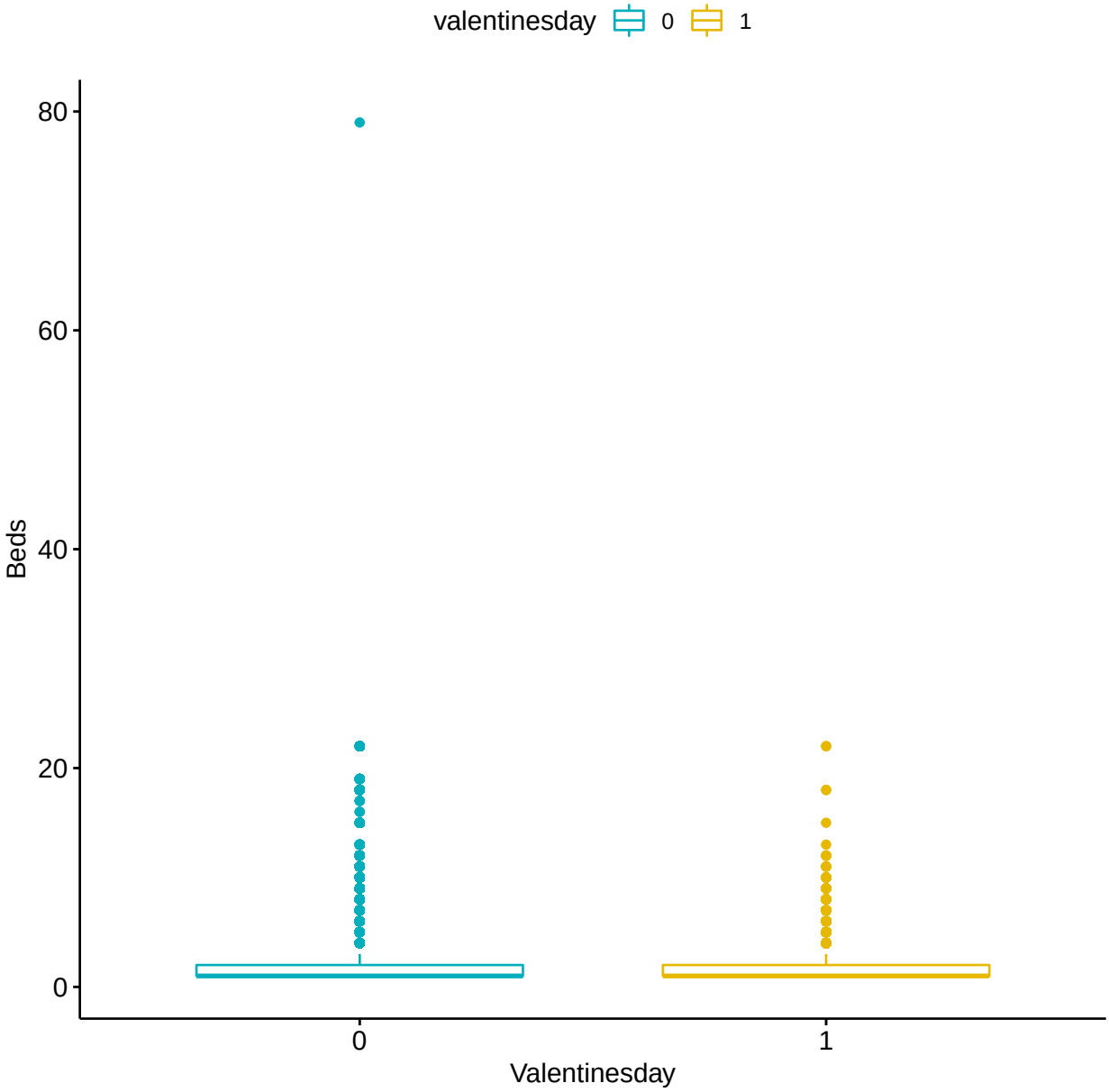
Price analysis, Bookings analysis and Bed analysis

Before the logistic regression can be performed, it is necessary to check for normality. A Shapiro-Wilk test is used to see if the data are normally distributed. The whole dataset is too large for the normality test, so a random sample of 5000 observations is generated. The data are normally distributed because $p = 0.9555875 > 0.05$, which means that a logistic regression can be performed.

The boxplot below shows that there are no extreme outliers for the price variable, so the data can be used to conduct a t-test.



However, if we look at the boxplot of the number of beds for the Airbnbs, we see many outliers and this should be taken into account when drawing conclusions based on this data.



Analysis & Results

Price analysis

This analysis will answer the following question: "What is the influence of Valentine's Day on the price of Airbnb listings?". In order to answer this question, a t-test will be performed.

First, we will look at the t-test per listing for all the cities studied. This t-test is not statistically significant, as p=0.8746806 $>$ 0.05. Additionally, the t-test results for the different cities Madrid, p = 0.9271568, Paris, p = 0.9160878, and Rome, p = 0.9726884, are all $>$ 0.05. Therefore, with a significance of 0.05, H0 cannot be rejected for these three cities in Europe.

Bookings analysis

This analysis will answer the following question: "What is the influence of Valentine's Day on the number of booked Airbnb listings?". A logistic regression will be performed.

First, we perform a logistic regression for all cities combined. The logistic regression gave us a P-value $<$ 0.01. Therefore, with a significance of 0.05, the H0 can be rejected. Therefore, a significant relation between Valentine's Day and the number of bookings has been found. When looking at the exponents of the logistic regression, we can conclude that on Valentine's Day, the odds of a listing being booked, increases with 1.025085.

Table 2: Effect of Valentine's Day on number of bookings of Airbnb listings

	Number of bookings
	Total
Valentine's Day	1.025
	(1.010)
Constant	2.307**
	(1.002)
Observations	1,491,362
Log Likelihood	−913,823.100
Akaike Inf. Crit.	1,827,650.000
Note:	*p<0.1; **p<0.05; ***p<0.01

As for the different cities, we find small differences for this relation. On Valentine's Day, the odds of booked Airnbnb's in Madrid decrease by a value of 0.9961. While in Rome, these odds increase by a value of 1.005468 and in Paris by a value of 1.006446. From the exponents, we can conclude that the odds of an Airbnb being booked on Valentine's Day, is the greatest in Paris.

Lastly, a model fit is performed with a log likelihood ratio test to check whether the model is significant and H0 can be rejected. The log likelihood ratio test gives us a P-value of 0.009886. With a significance of 0.05, H0 can be rejected. Therefore, the model has found to be significant.

Beds analysis

This analysis will answer the following question: "Compared to 'normal days', are relatively more Airbnb accommodations with one/two beds booked during Valentine's Day?". In order to analyse this question, a logistic regression will be performed.

First, we will perform a logistic regression for all cities combined. The p-value of the logistic regression is $>$ 0.05, therefore H0 cannot be rejected. This means that Valentine's Day has no significant effect on the amount of beds of booked Airbnb listings. As a result, we are not able to interpret the exponents of the regression.

Table 3: Effect of Valentine's Day on number of beds of Airbnb listings

	One or two beds
	Total
Valentine's Day	0.999
	(1.002)
Constant	2.316**
	(1.000)
Observations	1,040,620
Log Likelihood	−432,413.500
Akaike Inf. Crit.	864,831.100
Note:	*p<0.1; **p<0.05; ***p<0.01

Finally, a model fit is performed with a log likelihood ratio test to check whether the model is significant.The log likelihood ratio test gives us a value of 0.7905802. With a significance of 0.05, H0 cannot be rejected. Therefore, the model is not found to be significant.

Table 4: Effect of Valentine's Day on number of beds of Airbnb listings per city

	One or two beds		
	Madrid	Paris	Rome
Valentine's Day	1.000	0.998	1.000
	(1.006)	(1.002)	(1.006)
Constant	2.313**	2.387**	2.001**
	(1.001)	(1.000)	(1.001)
Observations	127,199	759,576	153,845
Log Likelihood	−53,329.210	−250,485.900	−99,134.800
Akaike Inf. Crit.	106,662.400	500,975.700	198,273.600
Note:		*p<0.1; **p<0.05; ***p<0.01	

Conclusion

Answering our research question, no significant relation has been found between Valentine's day and the prices of Airbnb listings. This is the case both for the accommodations in Paris, Rome and Madrid together and for the cities separately. Secondly, a significant relation has been found between Valentine's Day and the changes of an Airbnb listing being booked. However, although the changes of a listing being booked in Paris and Rome increase on Valentine's Day, the chance of a booking decreases in Madrid on this day. Despite Paris having the smallest population of the three cities, it is here that the greatest impact can be seen. Lastly, no significant relation has been found between Valentine's Day and the number of beds in the Airbnbs booked.