Project: Question 5

Gabriel Berardi

1. Exploratory Analysis
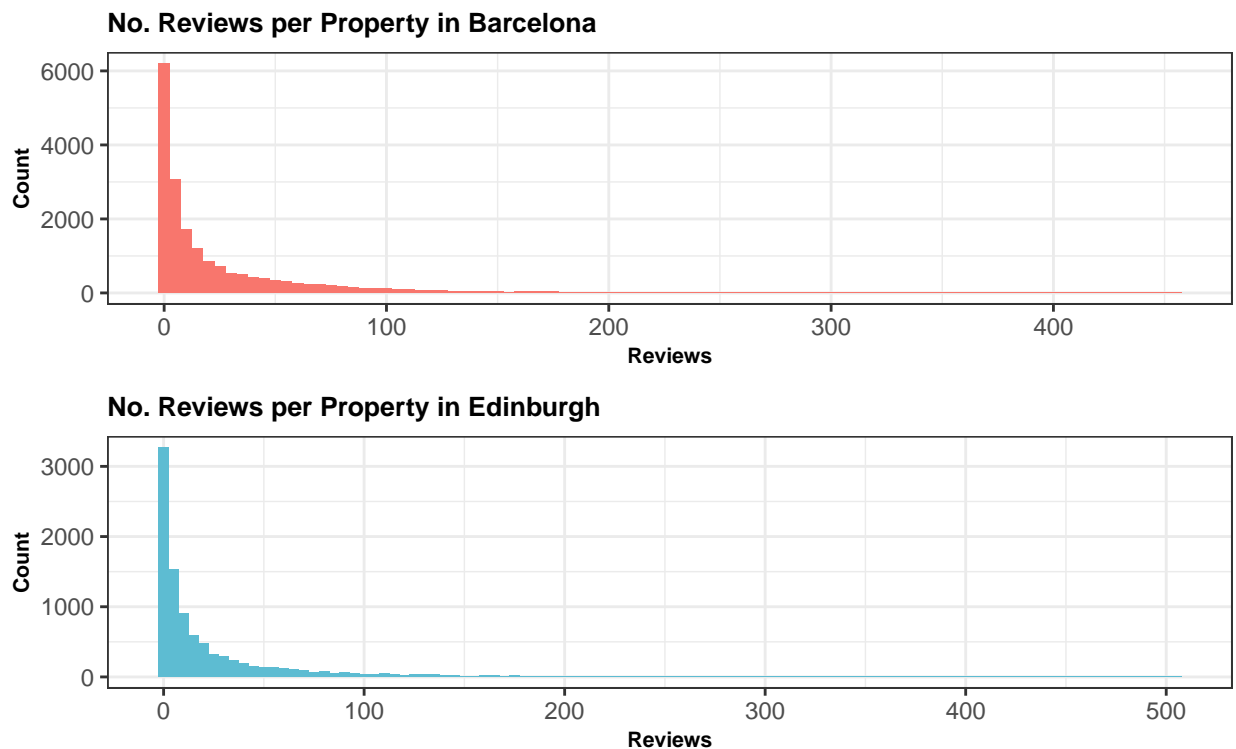
(a) Number of Reviews for each Property

Let's begin this short exploratory analysis by looking at a five-number summary of the number of reviews for each property in Barcelona and Edinburgh:
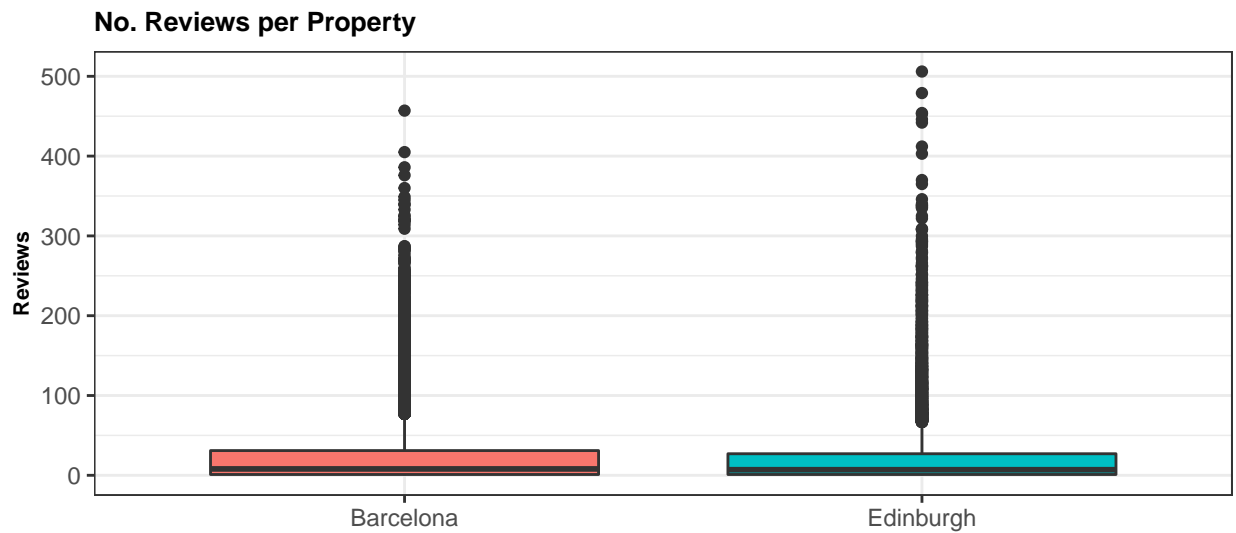
Table 1: Reviews per Property

City	Minimum	LQ	Median	Mean	UQ	Maximum
Barcelona	0	1	8	24.7125	31	457
Edinburgh	0	1	7	23.6505	27	506

As we can see from the large distance between the Median and the Mean, the data seems to be highly skewed. Since the Median lies to the left of the Mean, this data seems to be skewed to the right.

Let's look at a histogram for the number of reviews Barcelona and Edinburgh received for each property to visualize this more clearly:

No. Reviews per Property in Barcelona

No. Reviews per Property in Edinburgh

From these two histograms, we now clearly see the skewness to the right for the number of reviews per property in both cities. There seem to be quite a few very high number of reviews, which are barely visible in a histogram. We can see these outliers more clearly by plotting a boxplot:

No. Reviews per Property

Apart from this, the distribution of the number of reviews per property seems to be quite similar for both cities, but it is hard to see on this plot. Let's remove the outliers and look at the boxplot once more:

No. Reviews per Property
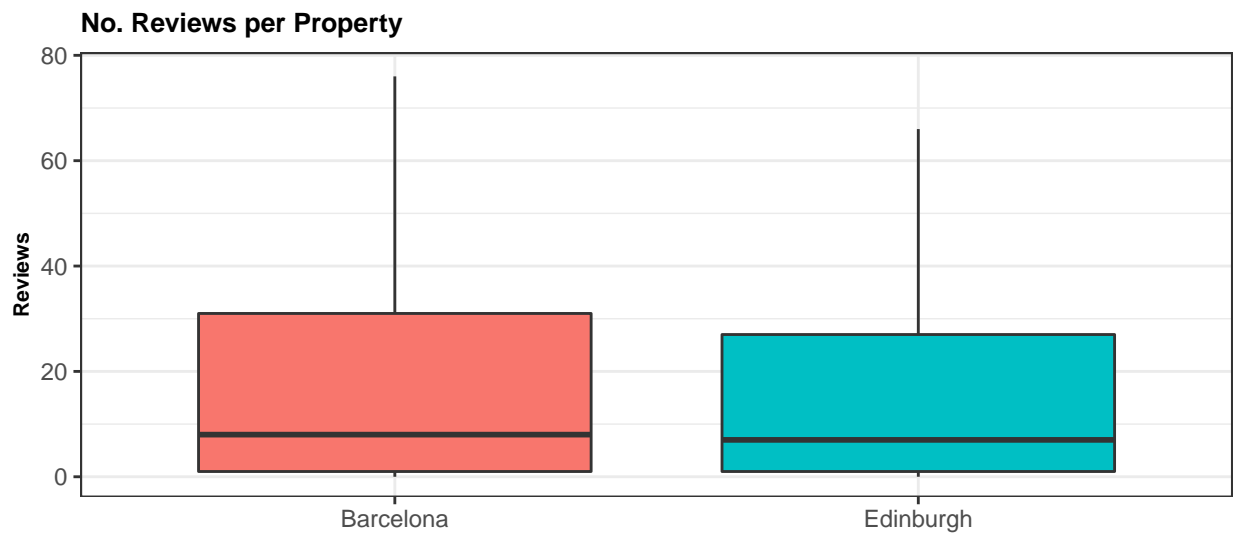
Now we can see that the number of reviews per property is distributed quite similarly in Barcelona and Edinburgh.

(b) The Room Type in each Property

Next, let's examine the type of rooms that are offered in each property in Barcelona and Edinburgh.
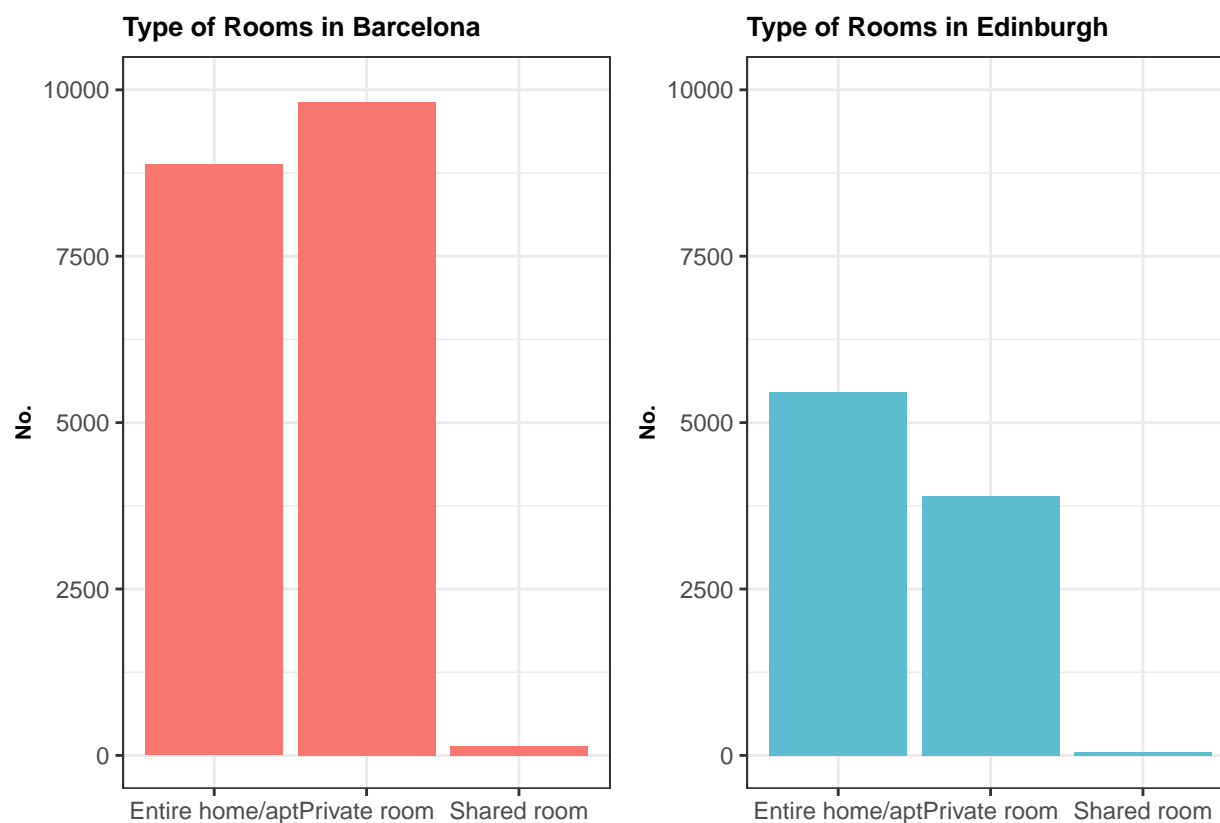
As we can see, there are in total many more offered rooms in Barcelona (n = 18.838) than in Edinburgh (n = 9.402), which makes sense since Barcelona has a population of over 1,600,000, while Edinburgh only has around 520,000 inhabitants. In terms of the distribution, we see that Barcelona has more private rooms than entire apartments, while Edinburgh has more entire apartments than private rooms. Both Barcelona and Edinburgh only have very few shared rooms:

Table 2: Room Types in Barcelona and Edinburgh

	Barcelona	Edinburgh
Entire home/apt	8877	5459
Private room	9821	3897
Shared room	140	46

Let's examine these room types a bit more and look at the Median price as well as the Median number of people that can be accommodated for each of the different room types in Barcelona and Edinburgh:

Table 3: Median Price and Capacity per Room Type

City	Room Type	Median Price	Median Capacity
Barcelona	Entire home/apt	137	4
Barcelona	Private room	48	2
Barcelona	Shared room	33	2
Edinburgh	Entire home/apt	129	4
Edinburgh	Private room	57	2
Edinburgh	Shared room	33	2

Median Price per Room Type


Median Capacity per Room Type

From Table 3 and the first barplot above we can see that the median price for different room types differ. As we would have expected, entire apartments are the most expensive, and private rooms are more expensive than shared rooms. The median prices are quite similar in Barcelona and Edinburgh.

However: As outlined in the PDF "Assessed_Project_Data_Description", there may be some values for price that were recorded per month, rather than per night, possibly resulting in a distortion of the data. This will be further investigated in Question 6.

The second barplot visualizes the median capacity per room type. Unsurprisingly, the median capacity for entire apartments is the highest, while the median number of people that can accommodated in private and shared rooms is the same. Again, the median capacity is the same in Barcelona and Edinburgh.
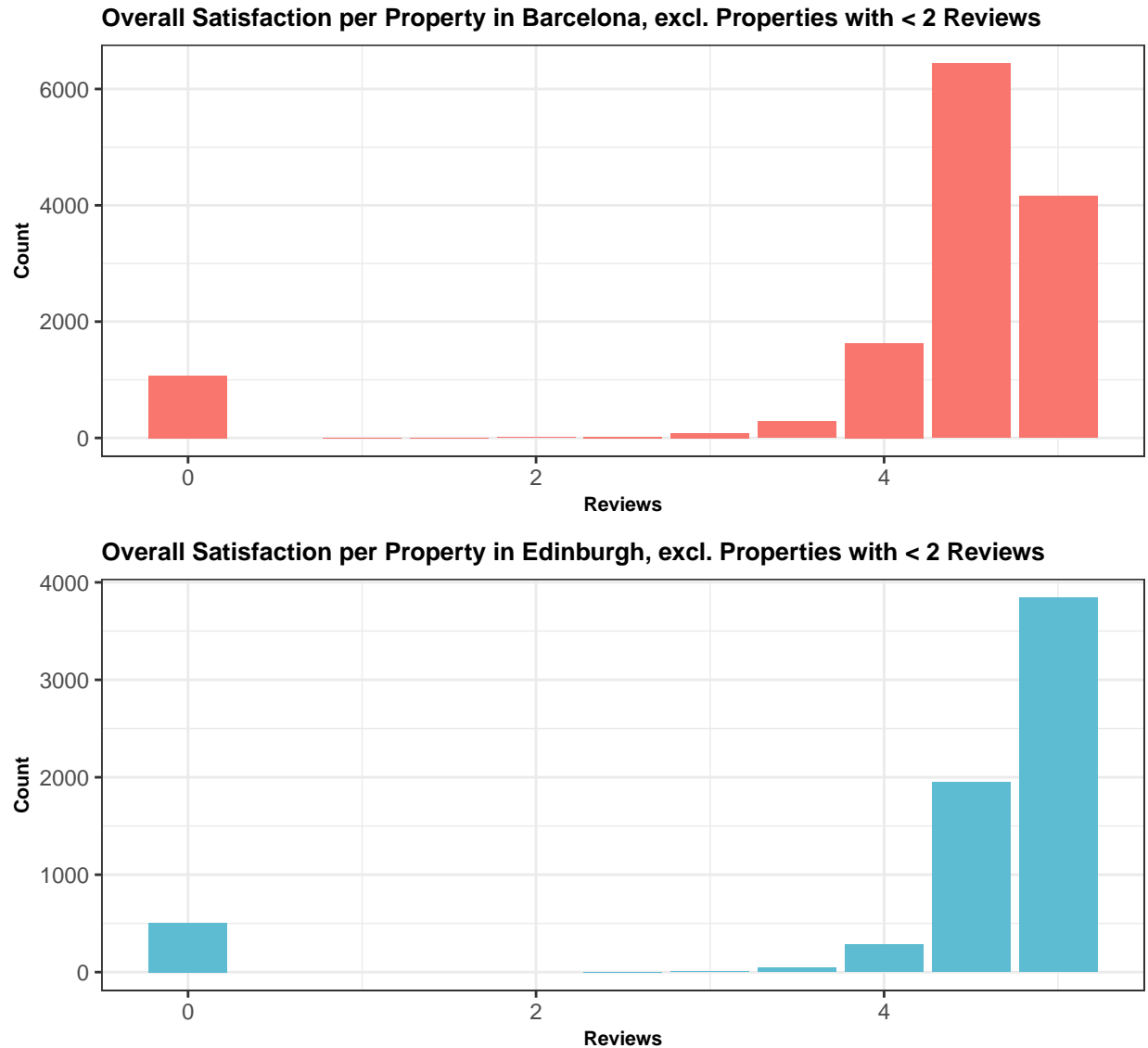
(c) Overall Satisfaction

To conclude this short exploratory data analysis, let's take a look at the overall satisfaction in Barcelona and Edinburgh:

Overall Satisfaction per Property in Barcelona

Overall Satisfaction per Property in Edinburgh
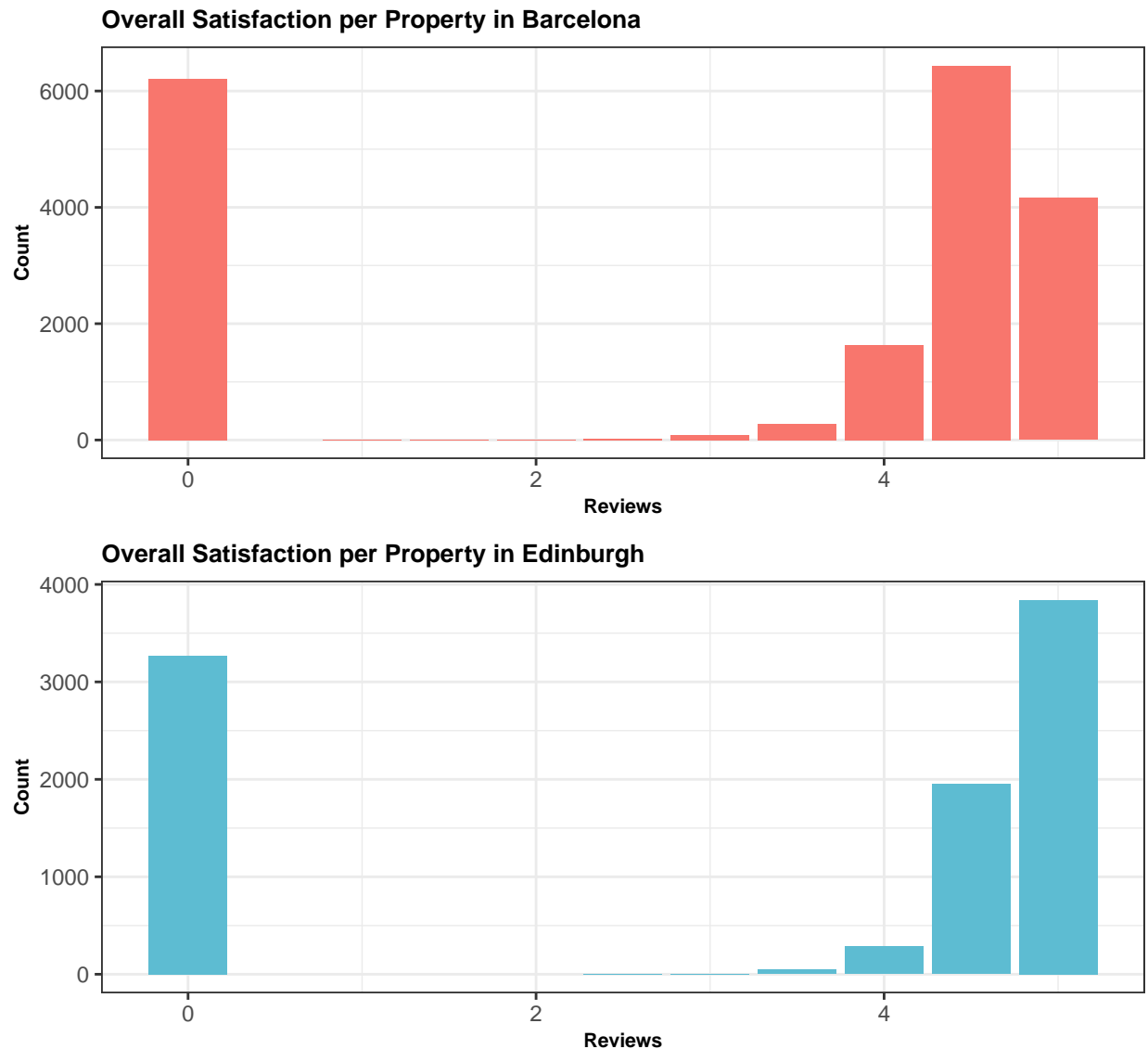
As we can see from the two barplots above, the reviews in both Barcelona and Glasgow seem to be quite extreme, with most ratings being either very bad (0.0) or really good (4.5 or 5.0).

However, we need to consider that listings with an overall satisfaction of 0.0 probably have very few, or even no reviews at all. Moreover, if a listing only has one review, the overall satisfaction does not carry a lot of information. Let's therefore filter all listings that have less than two reviews and reproduce the barplots above once more:

Overall Satisfaction per Property in Barcelona, excl. Properties with < 2 Reviews

Overall Satisfaction per Property in Edinburgh, excl. Properties with < 2 Reviews

As we have anticipated, when looking at properties with at least two reviews, a lot fewer listings have an overall satisfaction of 0.0. By excluding properties without any review, it furthermore looks like Edinburgh might have a slightly higher overall satisfaction. Let's have a closer look:

Table 4: Overall Satisfaction, excl. Properties without Reviews

City	Minimum	LQ	Median	Mean	UQ	Maximum
Barcelona	0	4.0	4.5	3.792	5	5
Edinburgh	0	4.5	5.0	3.972	5	5

Table 4 shows that the overall satisfaction for AirBnb listings seems to be better in Edinburgh, compared to Barcelona, since both the Mean and the Median are higher.

2. Comments

Summarizing the exploratory analysis, we can conclude the following points about the AirBnb listings for Barcelona and Edinburgh:

- There are more data for Barcelona (18.838 listings) than for Edinburgh (9.402 listings).

- The number of reviews per property is highly right-skewed and the distribution is quite similar for both cities. There are no missing values in this column.

- There are three room types in both cities: whole apartments, private rooms and shared rooms. Barcelona has more private rooms than entire apartments, while Edinburgh has more entire apartments than private rooms. Both Barcelona and Edinburgh only have very few shared rooms. There are no missing values in this column. Furthermore, we have seen that both the price and capacity for each room type is quite similar in both cities.

- Regarding the overall satisfaction, we have seen that most reviews were either very low (0.0) or very high (4.5 or 5.0), with only very few reviews between that. There are no missing values in this column, but many properties had no reviews at all, resulting in many misleading values of 0.0, which we have excluded above. The overall satisfaction for AirBnb listings seems to be slightly higher in Edinburgh, compared to Barcelona.