Student Hustle Data Analysis

Dprep Team 3

2022-03-21

Mean prices

First, we compared the prices of a listing, consisting of one room, by calculating the means of the price per listing in the top 10 student cities in Europe. The results showed that in Dublin, on average hosts ask the highest price for their listing with an average of 68.55 euro' s per night, while in Manchester it is on average the cheapest to rent a room with an average of 32.70 euro's per night. In the plot below, the mean prices are visually represented per city.

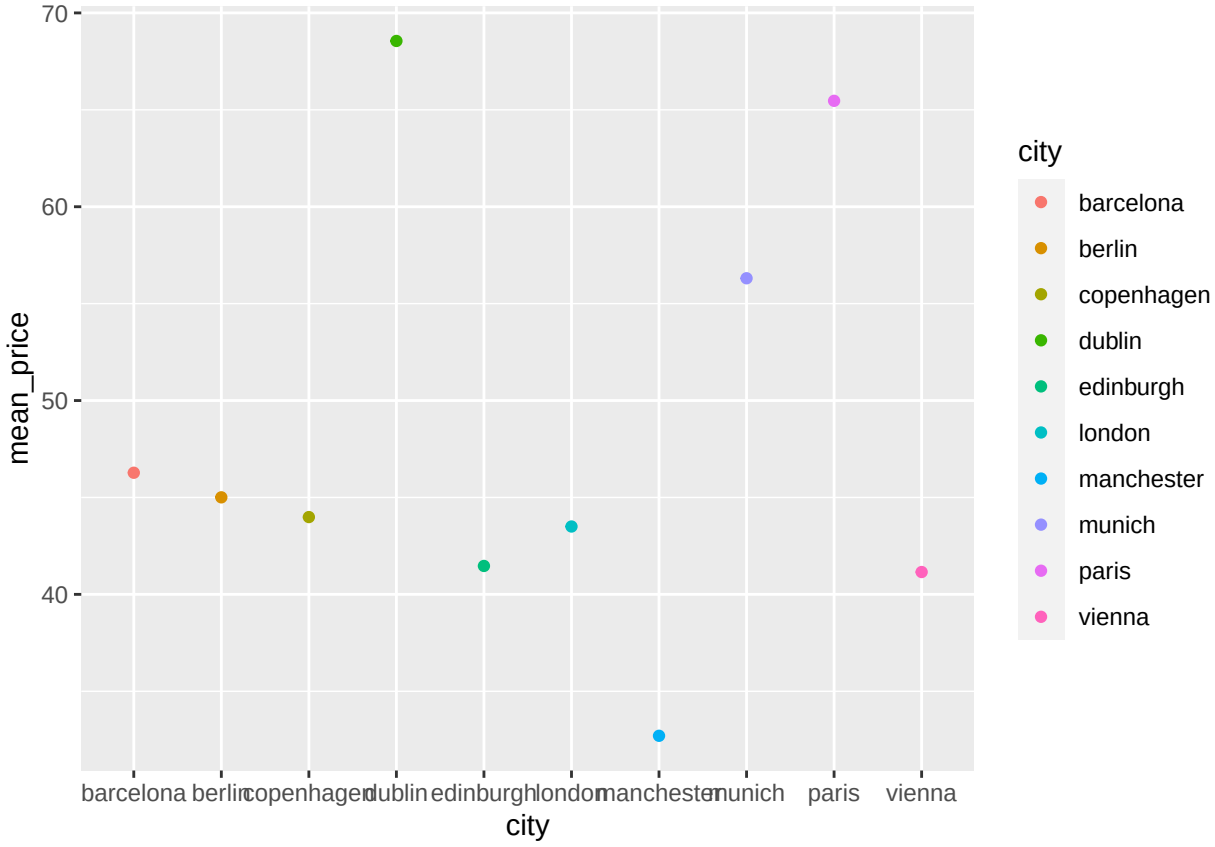

Figure 1: Plot 1: mean prices per city

Mean reviews

Second, we compared the mean review scores per city. We found that customers are most satisfied in Edinburgh, and relatively less satisfied in Barcelona. However, the average review score per listing in Barcelona is still 4.72, indicating that customers are generally very satisfied in all the student cities (see Plot 2).

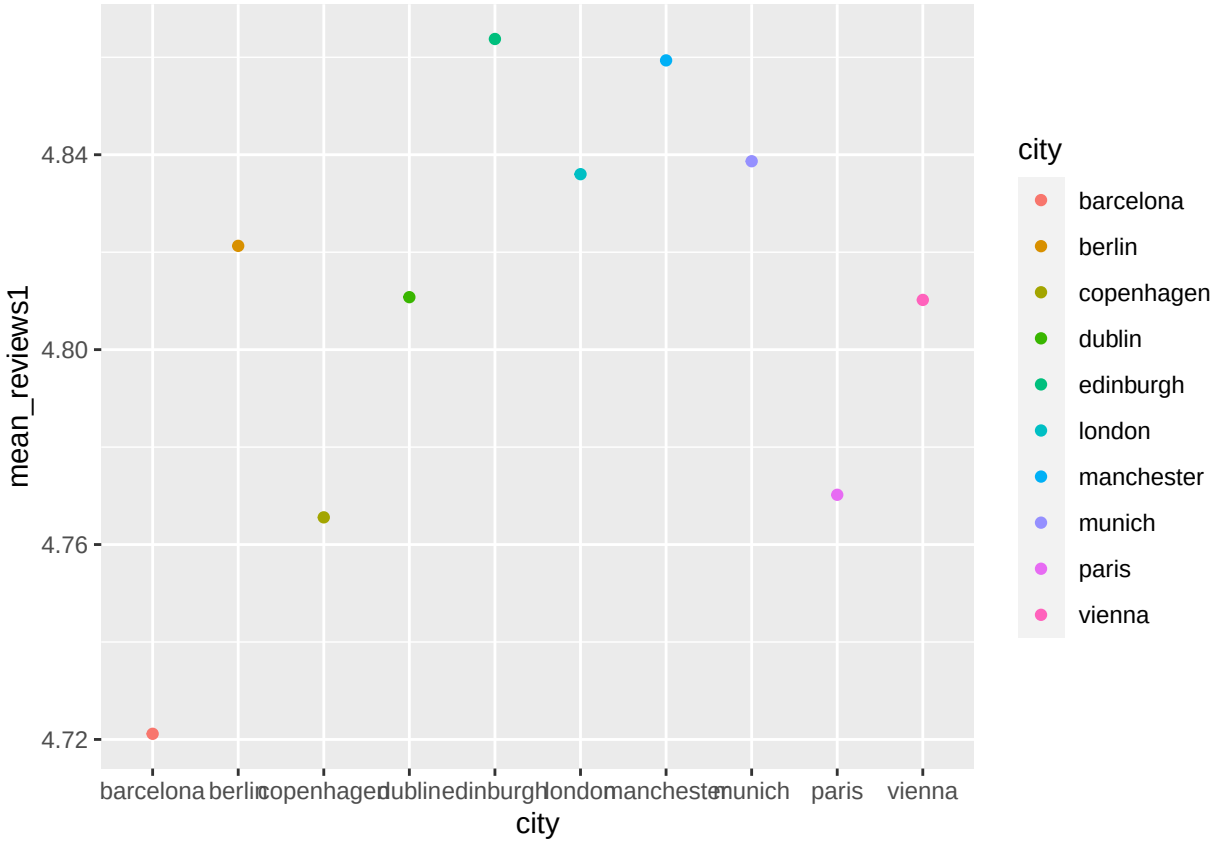

Figure 2: Plot 2: mean reviews per city

Mean availability

Next, the main availability per year per city were analyzed. In plot 3, it can be seen that in Vienna, one listings are on average most available throughout the year. One listings in Dublin, however, have the lowest availability throughout the year.

Relationship between price and review score rating

After checking those plots, it is shown that Manchester and Edinburgh have one of the highest review scores but also the lowest average price per night. Therefore, we conducted two linear regressions to investigate whether the review scores of a listing are significantly influenced by these factors. A linear regression is used, because the results of the linear regressions provide useful insights and explains the relationship between the dependent variable and one or more independent variables. In other words, it is possible to find out whether reviews are influenced price and availability, and what the direction and magnitude of this possible effect might be.

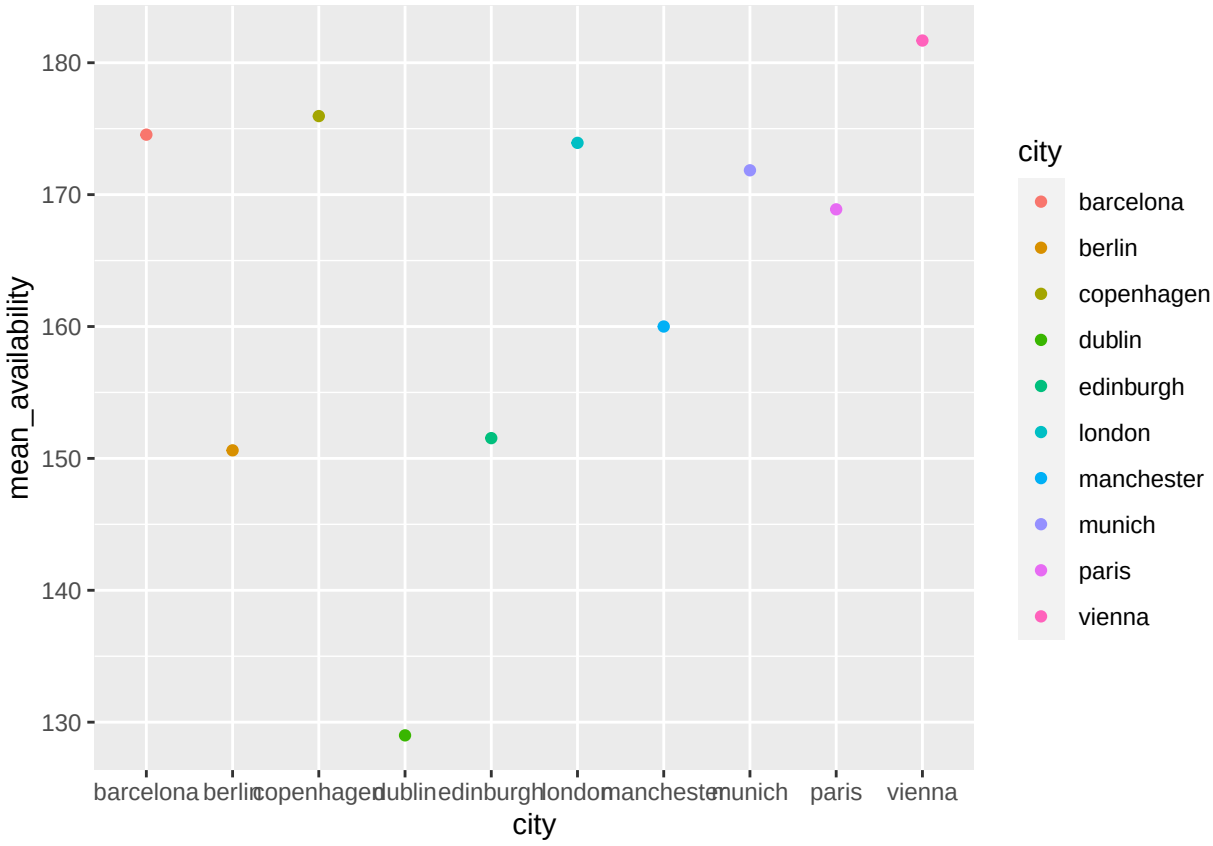

Figure 3: Plot 3: mean availability per city

To investigate whether there is a relationship between the price per listing and the review scores per listing, a linear regression was performed. As the p-value (p = 0.305) is higher than 0.05, there is no significant relationship between price and review score per listing. The results can be found in the table below

x

Title: Price Linear Regression Results
==========
Dependent variable:

Review Scores
---
availability_365 -0.0002***
(0.00002)
t = -7.120
p = 0.000
Constant 4.836***
(0.005)
t = 1,041.460
p = 0.000

Observations 3,485
R2 0.014
Adjusted R2 0.014
Residual Std. Error 0.161 (df = 3483)
F Statistic 50.688*** (df = 1; 3483) (p = 0.000)
==========
Note: *p<0.1;* ***p<0.05;*** p<0.01

Relationship between price and availability

Furthermore, when looking at the availability per listing, we found a significant relationship between availability and review scores. This means that if the availability per listing increases, the review score per listing significantly decreases.

x

Title: Availability Linear Regression Results
==========
Dependent variable:

Review Scores
---
availability_365 -0.0002***
(0.00002)
t = -7.120
p = 0.000
Constant 4.836***
(0.005)
t = 1,041.460
p = 0.000

x

Observations 3,485
R2 0.014
Adjusted R2 0.014
Residual Std. Error 0.161 (df = 3483)
F Statistic 50.688*** (df = 1; 3483) (p = 0.000)
==========
Note: *p<0.1;* ***p<0.05;*** p<0.01

Multiple regression with price, availability, short_stay and city

In order to find out which city is most attractive to rent out your room, several analyses were done. It is expected that reviews are influenced by the variables 'price', 'availability' and 'short stays'. Therefore, these variables are further analyzed.

A linear regression was performed to examine the relationship between different independent variables 'price', 'availability' and 'short stays' and the dependent variable 'review score'. The output of the linear regression can be found in table below.

x

Title: Multiple Regression Results
==========
Dependent variable:
-
Review Scores

price 0.0003***
(0.0001)
t = 4.696
p = 0.00001
availability_365 -0.0002***
(0.00002)
t = -7.242
p = 0.000
short_stay 0.042
(0.029)
t = 1.447
p = 0.149
cityberlin 0.097***
(0.010)
t = 9.355
p = 0.000
citycopenhagen -0.076**
(0.030)
t = -2.548
p = 0.011
citydublin 0.075***
(0.015)
t = 5.003
p = 0.00000

x
cityedinburgh 0.138***
(0.013)
t = 10.983
p = 0.000
citylondon 0.113***
(0.008)
t = 13.369
p = 0.000
citymanchester 0.138***
(0.016)
t = 8.373
p = 0.000
citymunich 0.114***
(0.013)
t = 8.498
p = 0.000
cityparis 0.041***
(0.010)
t = 3.950
p = 0.0001
cityvienna 0.091***
(0.017)
t = 5.296
p = 0.00000
Constant 4.735***
(0.009)
t = 550.360
p = 0.000
Observations 3,485
R2 0.091
Adjusted R2 0.088
Residual Std. Error 0.155 (df = 3472)
F Statistic 28.937*** (df = 12; 3472) (p = 0.000)
Note: *p<0.1;* ***p<0.05;*** p<0.01

The results of the linear regression showed that the variables price and availability have a significant influence on the review of the listing. The regression coefficient of price shows that price has a positive influence on reviews per listing, indicating that if the price increases the review score per listing also significantly increases. This makes sense as generally, the higher the price, the higher the 'quality' of the room. Therefore, if the price for a listing increases, the higher the review score as this generally reflects the quality of the stay.

Regarding availability, the regression coefficient tells us that if the availability per listing increases, the reviews per listing significantly decreases. A reason for this could be that if a room is booked more often because it is more available, its furniture is used more often, which might negatively influence its condition. This might have a negative influence on the review per listing.

Furthermore, the results showed that short stay does not have a significant impact on review scores. Therefore, no conclusions can be made. Additionally, all the student cities significantly influences the review scores. The relations described above are visually represented in the plots in Figure 1.

R-squared worthwhile to mention is that it can be stated that the model poorly explained the variation

as the coefficient of determination is close to 0 (Adjusted R2 = .088). This means that the independent variables explain only 8.8% of the variance in review score per listing.

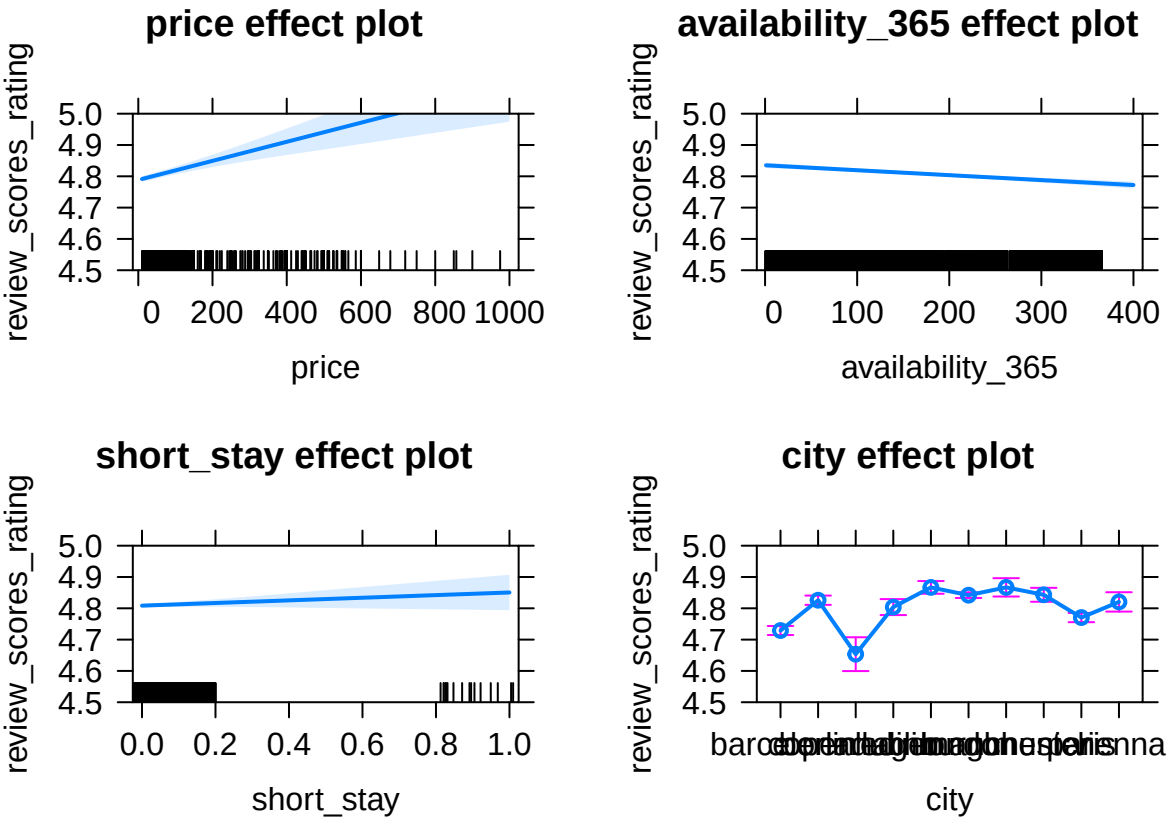

Figure 4: Figure 1: Multiple linear regression

Top 10

To find out which city is most attractive, further analysis and calculations are needed. First the mean price and mean availability are needed, which can be found in the overview below.

city	mean_price	mean_reviews1	mean_availability
barcelona	46.27527	4.721161	174.5462
berlin	45.00683	4.821298	150.6105
copenhagen	43.98841	4.765580	175.9565
dublin	68.55245	4.810769	129.0000
edinburgh	41.46637	4.863755	151.5371
london	43.50190	4.836011	173.9254
manchester	32.70435	4.859364	160.0000
munich	56.30851	4.838670	171.8457
paris	65.46387	4.770210	168.8858
vienna	41.15306	4.810204	181.6837

Together with the coefficients in Table 1, the reviews per student city can be calculated. This results in the following formula: Intercept + (price * mean price per city) + (availability * mean availability per city) +

coefficient short stay + coefficient per city The intercept coefficient in Table 1 represents Barcelona. Based on the above mentioned formula, the review score per student city can be calculated. The results per city can be found in Table 3. The table shows that Edinburgh has the highest score. Based on this score, it can be stated that Edinburgh is the most attractive student city to rent out your room.

	score
edinburgh	4.903
manchester	4.899
munich	4.881
london	4.875
berlin	4.863
dublin	4.852
vienna	4.852
paris	4.811
barcelona	4.763
copenhagen	4.687

Conclusion

To conclude, based on the data of Airbnb, it can be stated that it is most attractive to rent out your private room on Airbnb in Edinburgh. The second most attractive European city to rent out your private room is Manchester and Munich ended up third.