



# What effects price for an Airbnb?

Zackary Johnston and Eesha Jagdhane



## Summary/Background

Two former school/roommates decided to place an air mattress into their living room in San Francisco and decided to make it a bed and breakfast for customers. They started this to offset the cost of living in San Francisco. After years of expansion, planning, and growth Airbnb was formed and has flourished into one of the largest brokers for vacation rentals in the world.

## Hypothesis

Our hypothesis is that minimum nights, number of reviews, reviews per month, and availability do influence price.

Null: The effect on price from the independent variables (minimum nights, number reviews, reviews per month, and availability) is equal to 0

Alternative: The effect on price from the independent variables (minimum nights, number reviews, reviews per month, and availability) is not equal to 0

## Methods

We used the data from 18905 different listed vacation rental homes in Amsterdam, North Holland, The Netherlands from insideairbnb.com. We used price as our dependent variable and then from the data we took minimum number of nights, number of reviews, reviews per month, and availability as our independent variables. From this equation we were able to perform multiple linear regression and have a line of best fit to help predict the correlation between the variables. We were also able to use the provided maps from Airbnb and use the locations of either the entire homes/apartments or private rooms to see where higher price areas are.

## Data/Results

Listings data excel format from The Netherlands, Amsterdam

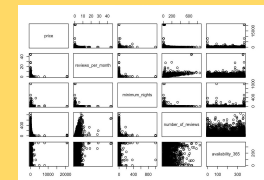
Multiple linear regression

```
## Multiple linear regression
lmfit <- lm(price ~ reviews_per_month + minimum_nights + number_of_reviews + availability_365, data = listings)
summary(lmfit)
```

Correlation Matrix

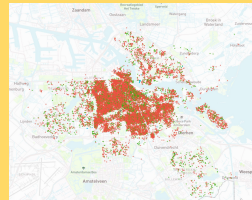


Correlation Scatter plot



Map data

Red = entire homes/apartments  
Green = private room



Correlation values, VIF values, and Tolerance Values

	price	reviews_per_month	minimum_nights	number_of_reviews	availability_365
price	1.000	-0.034	0.010	-0.028	0.087
reviews_per_month	-0.034	1.000	-0.029	0.733	0.305
minimum_nights	0.010	-0.029	1.000	-0.014	0.050
number_of_reviews	-0.028	0.733	-0.014	1.000	0.291
availability_365	0.087	0.305	0.050	0.291	1.000

VIF VALUES				
reviews_per_month	minimum_nights	availability_365	number_of_reviews	
2.209760	1.004624	1.118583	2.186622	
TOLERANCE VALUES				
reviews_per_month	minimum_nights	availability_365	number_of_reviews	
0.4525379	0.9953969	0.8939886	0.4573264	

## Conclusion

From our map data that was provided from Airbnb we can see that on average entire homes/apartments are more expensive in central Amsterdam within the city compared to private rooms. Entire homes/apartments average about 168 euros a night and for a private room it is 109 euros a night.

There is no multicollinearity, and this can be confirmed from the VIFs from the data translation. None of the independent variables have a VIF above 5 therefore there is no bias.

If we translate the listing data into R and create visualizations, we can conclude that the overall model is statistically significant as the p-value is less than 0.05 which can be confirmed from the F-statistics provided. So, we reject our null hypothesis as our independent variables significantly influence price. There is a correlation and effect on price from our independent variables.

## References

<http://insideairbnb.com/get-the-data.html>  
<https://www.airbnb.com/>