



SEVESTER



01

# AIRBNB PRICES IN BERLIN

Capstone 4





02

# Table of Contents

Presentation Outline

Introduction  
Problem  
Context  
Solutions





# Target Audience:

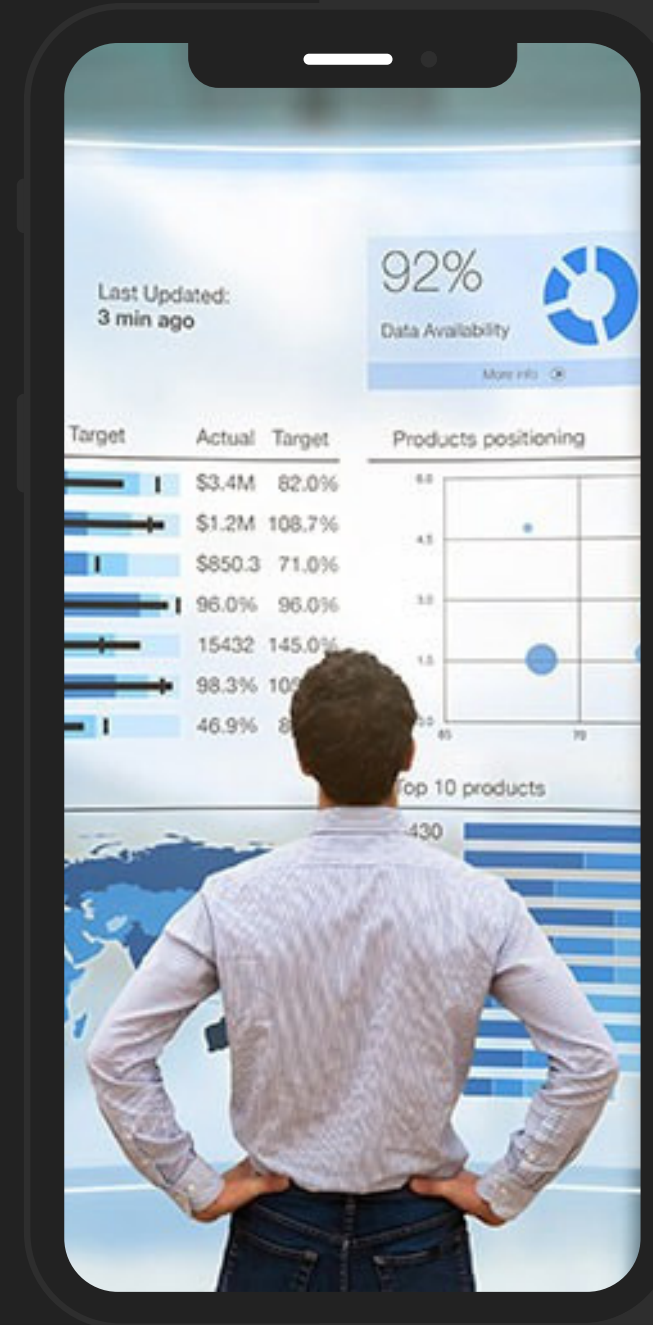
- House Owner - Host  
(New/Existing)
- Shareholders





04

# Who Am I?



Data Analyst for Airbnb





05

# Problem

- Surging growth rate
- Increasing supply/demand
- Pricing?

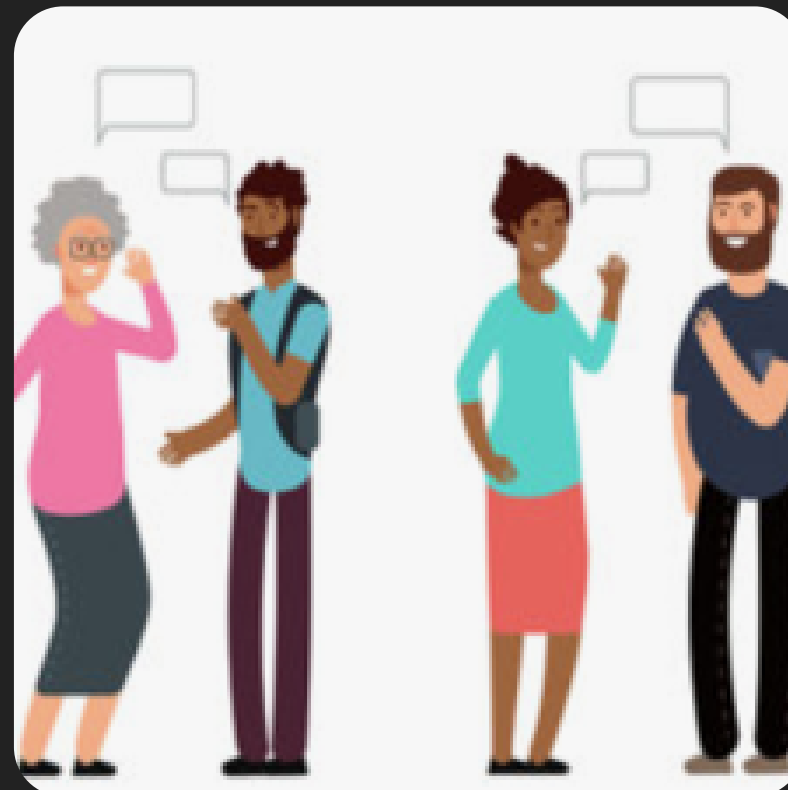




06



Convenience



# Why **Airbnb** instead of **Hotel**?

- Cheaper
- Better quality
- Convenient location
- Amenities
- Human Interaction





07

61%

Choose Airbnb Instead  
of budget or Mid -  
range hotel.



08

26%

Stay in Airbnb led them  
to increase their length  
of trip



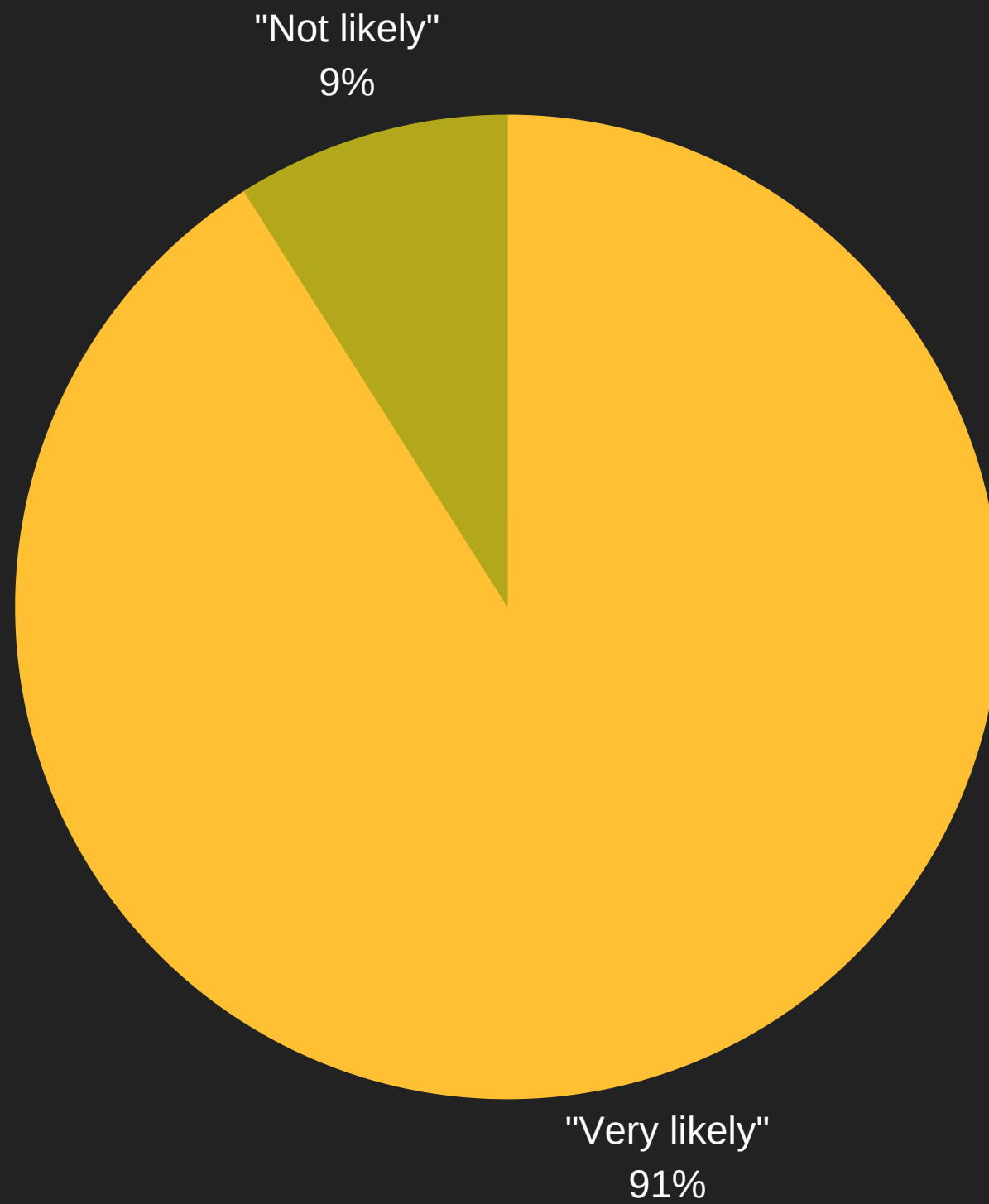




## Fun Fact Statistics

AIRBNB BERLIN REPORT

09



91% are very likely to recommend Airbnb to others





10

# How can I help?

Providing them with an informed insights  
derived from Machine learning





# Solutions

Predict price of Airbnb room using list of features data from Kaggle

## No overpriced room

Using Machine learning to predict prices that are best suited according to the features of their room. Thus, not overpriced and people can afford it better

## Potential new host

Able to put up their home/rooms up for rental on Airbnb website with confidence as they now know what their home true value is.

## Customer satisfaction

More consumer able to afford rooms that are more suited to their taste. Therefore, will greatly increases their trip's satisfaction. Thus more will be willing to use Airbnb again.

## Generating more revenue

As more rooms are rented out, this ultimately generating more revenue for the company



12

# Dataset: Kaggle



Bedrooms



Security\_deposit



Accommodates



Extra\_people



Bathrooms



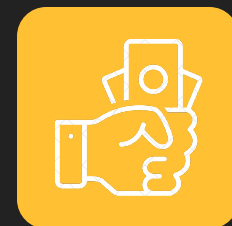
Minimum\_nights



Price



Distance\_in\_km



Cleaning\_fee



Size



13



Feature Engineering







14



## About this space

The very generous loft apartment with 43m<sup>2</sup> is situated in a typical Berlin old building. You'll find a bathroom with shower, a fully equiped Kitchen with a dining area and a large living area with 1 doublebed and a sofa bed. The apartment was completely renovated in November 2016 and newly furnished.

Scene district Berlin Friedrichshain, quiet side street, many cafes, restaurants, clubs in the vicinity. Approximately 1 km to Warschauer Straße.

### The space

We have made a lot of effort to make the apartment as attractive as possible. The old brickwork has been extensively restored. The complete furniture and the complete equipment were bought in November 16. Many pictures from famous Berlin photographers. Loving details.

### Other things to note

Please no smoking and no parties or events in the apartment or the building.

Check-out until 12:00 (if there are no other guests the same day maybe longer).

Check-in is at any time after 15:00 h

Scene district Berlin Friedrichshain, quiet side street, many cafes, restaurants, clubs in the vicinity. Approximately 1 km to Warschauer Straße.

[Show more](#) >

**Lower price.** Your dates are \$45 less than the avg. nightly rate of the last 60 days.



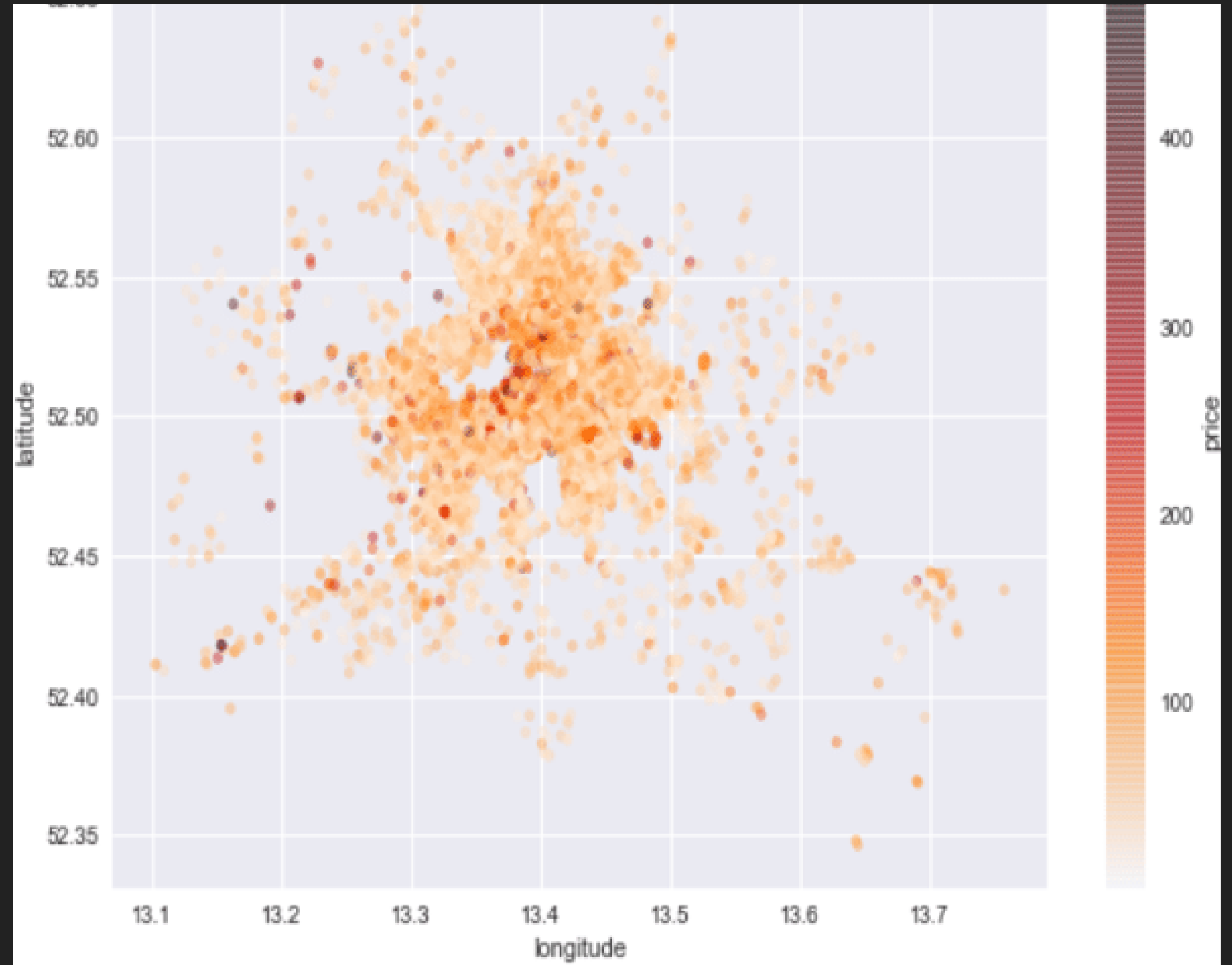
# Feature Engineering





# EDA

Price Differences on map

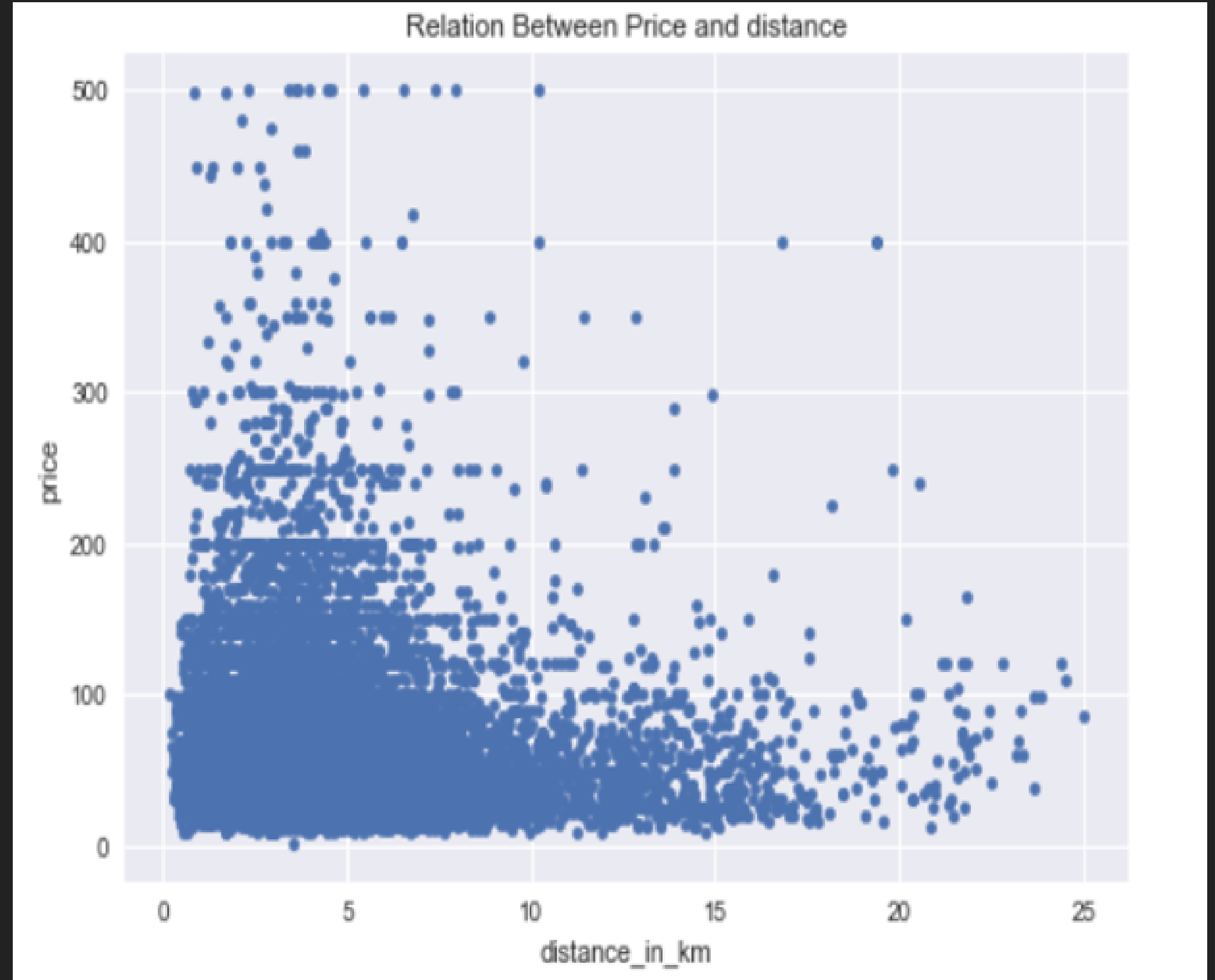




# EDA

Price differences base on  
distance

16

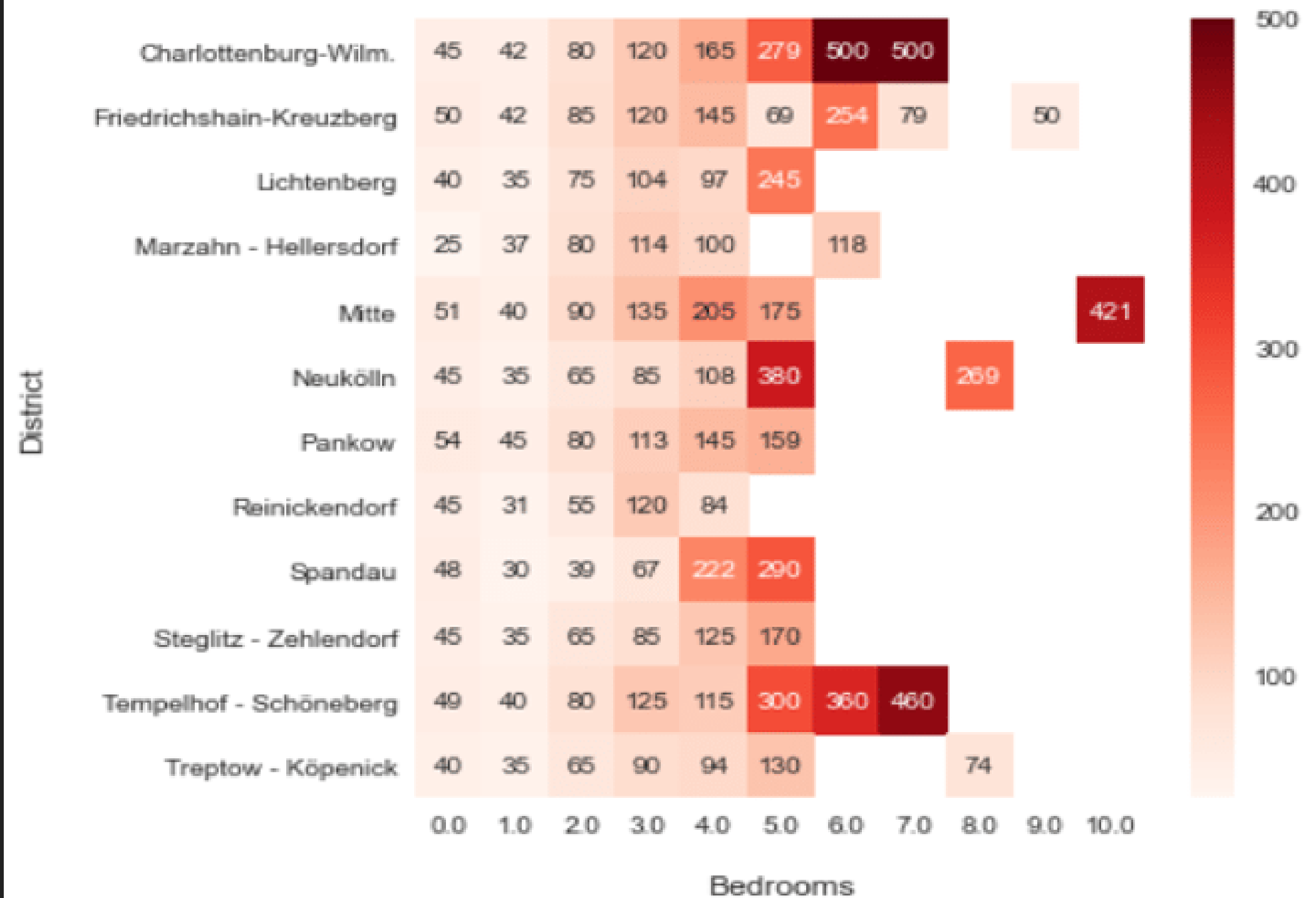




# EDA

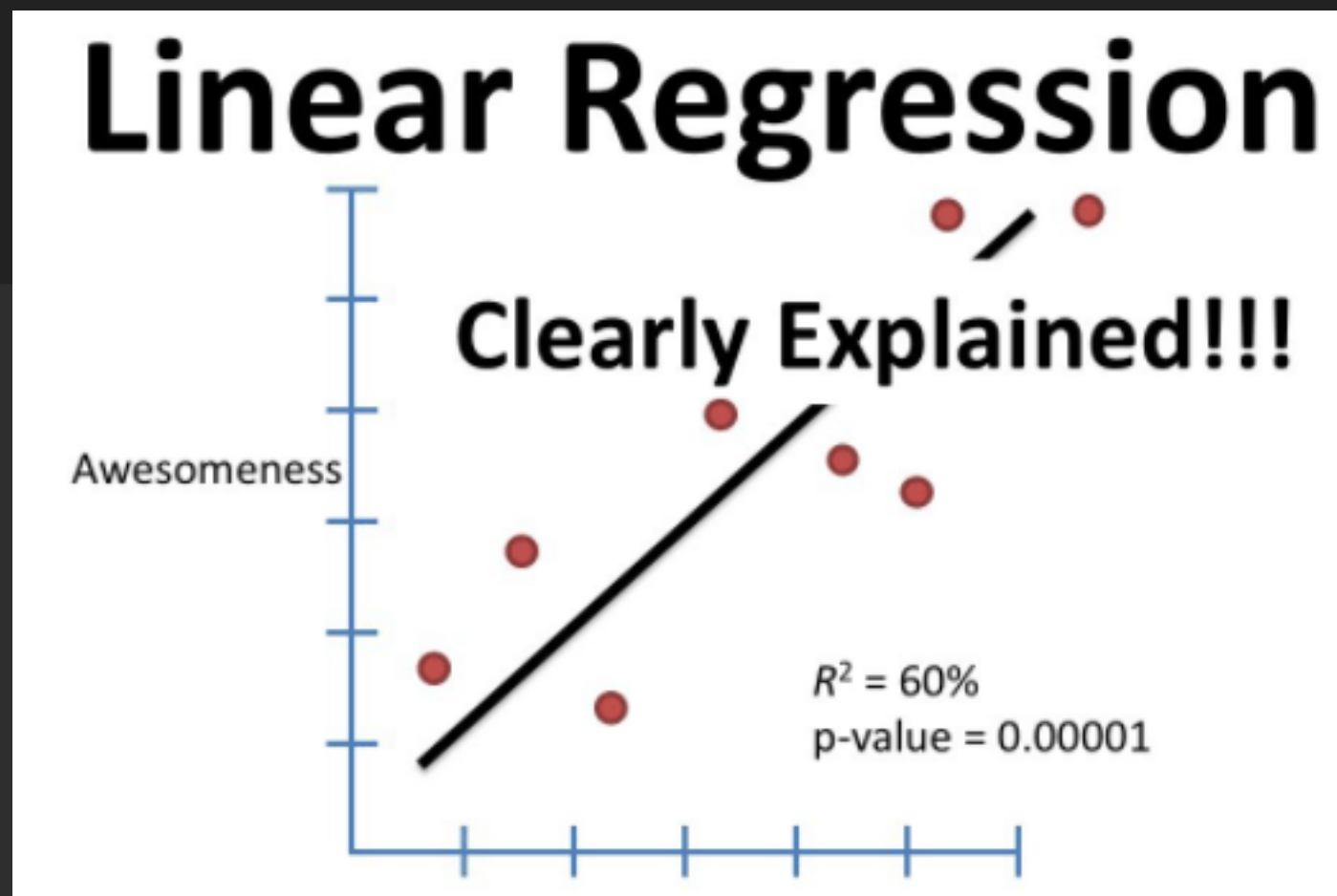
Price by neighbourhood  
and number of bedrooms

Heatmap: Prices by Neighbourhood and Number of Bedrooms

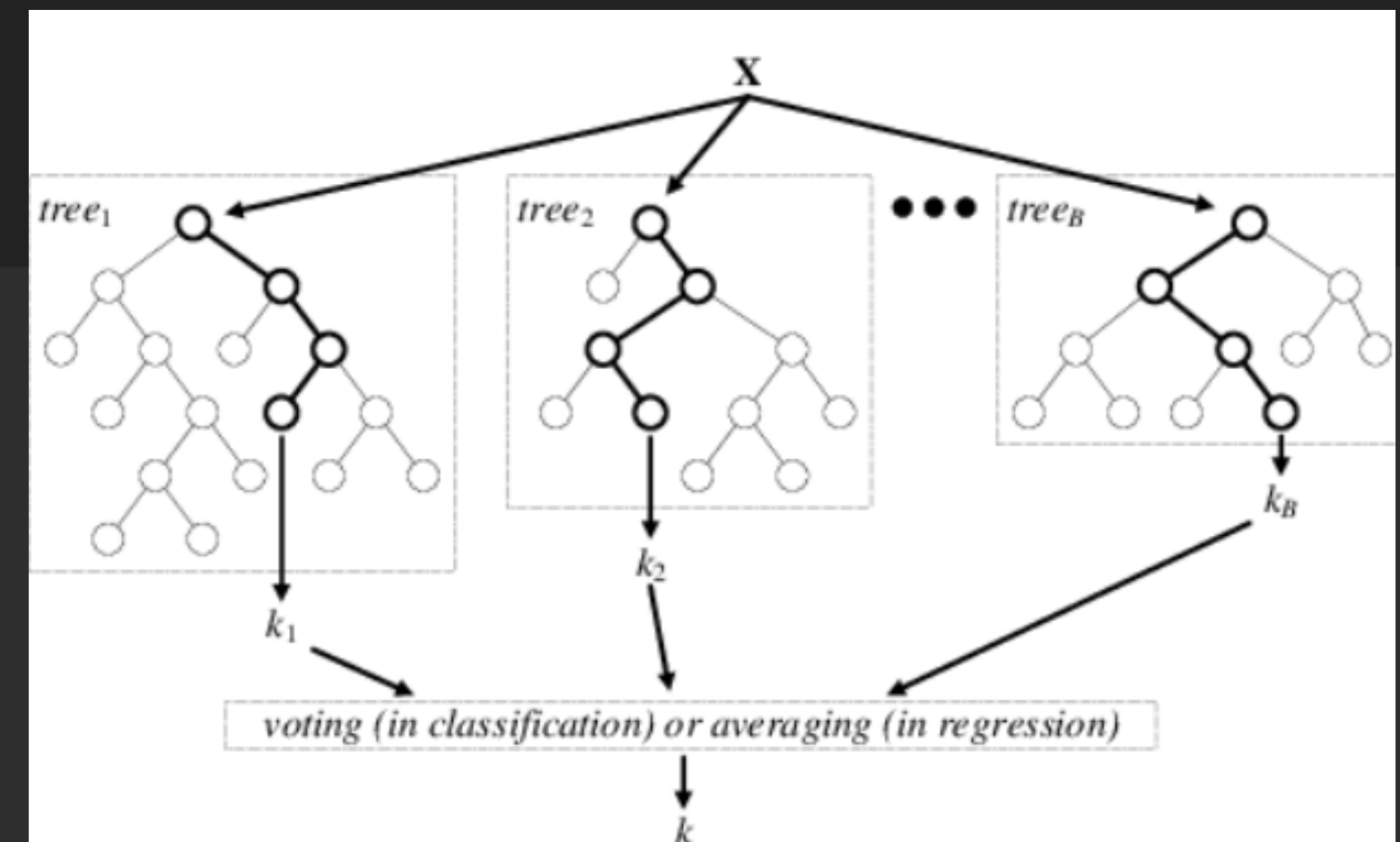


# d What are the model I am using?

10



Linear regression



Random Forest Regressor





# Model 1: Linear Regression

Result:

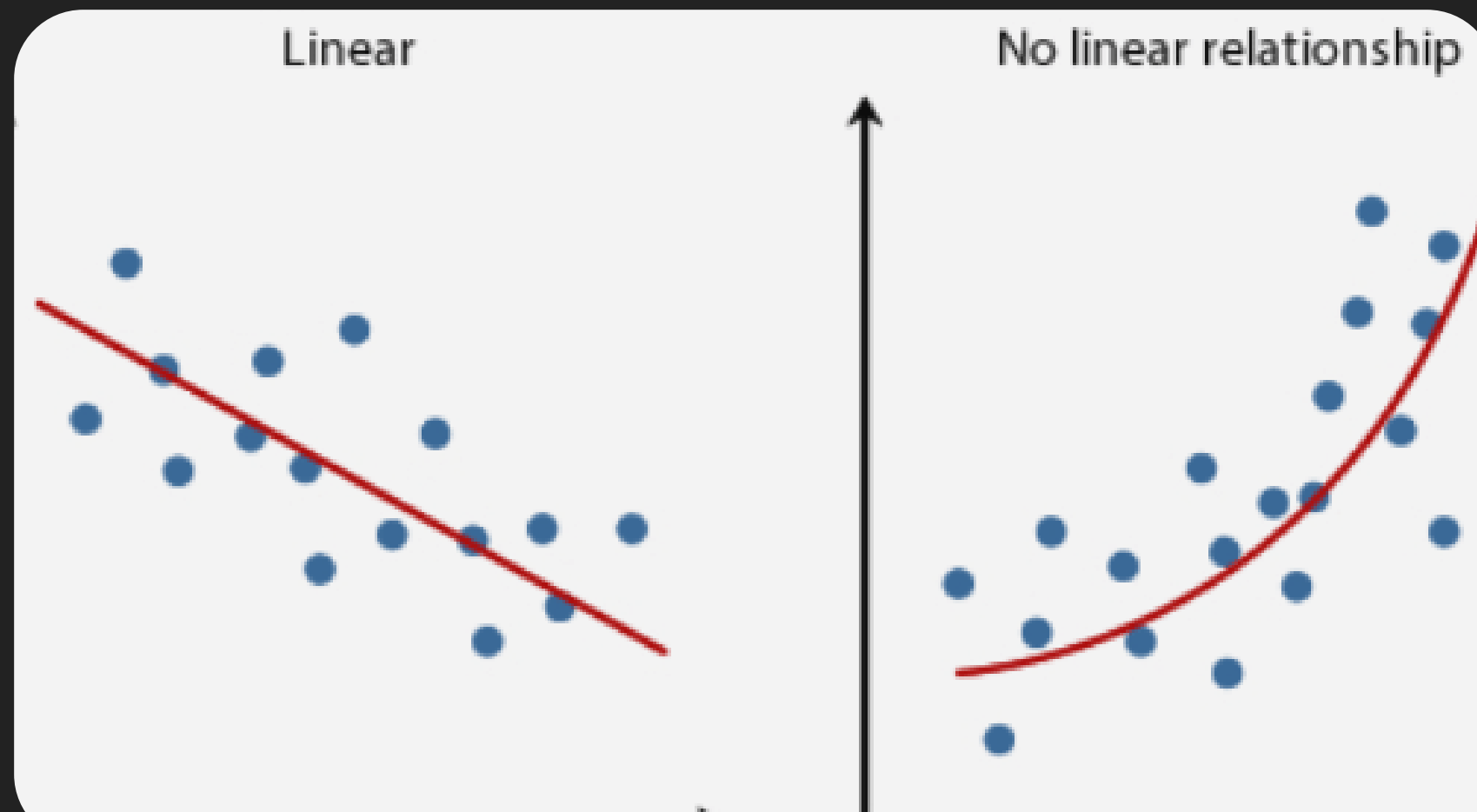
Coefficient of determination:

0.29

Score:

Regression Model training score: 0.46

Regression Model test score: 0.29



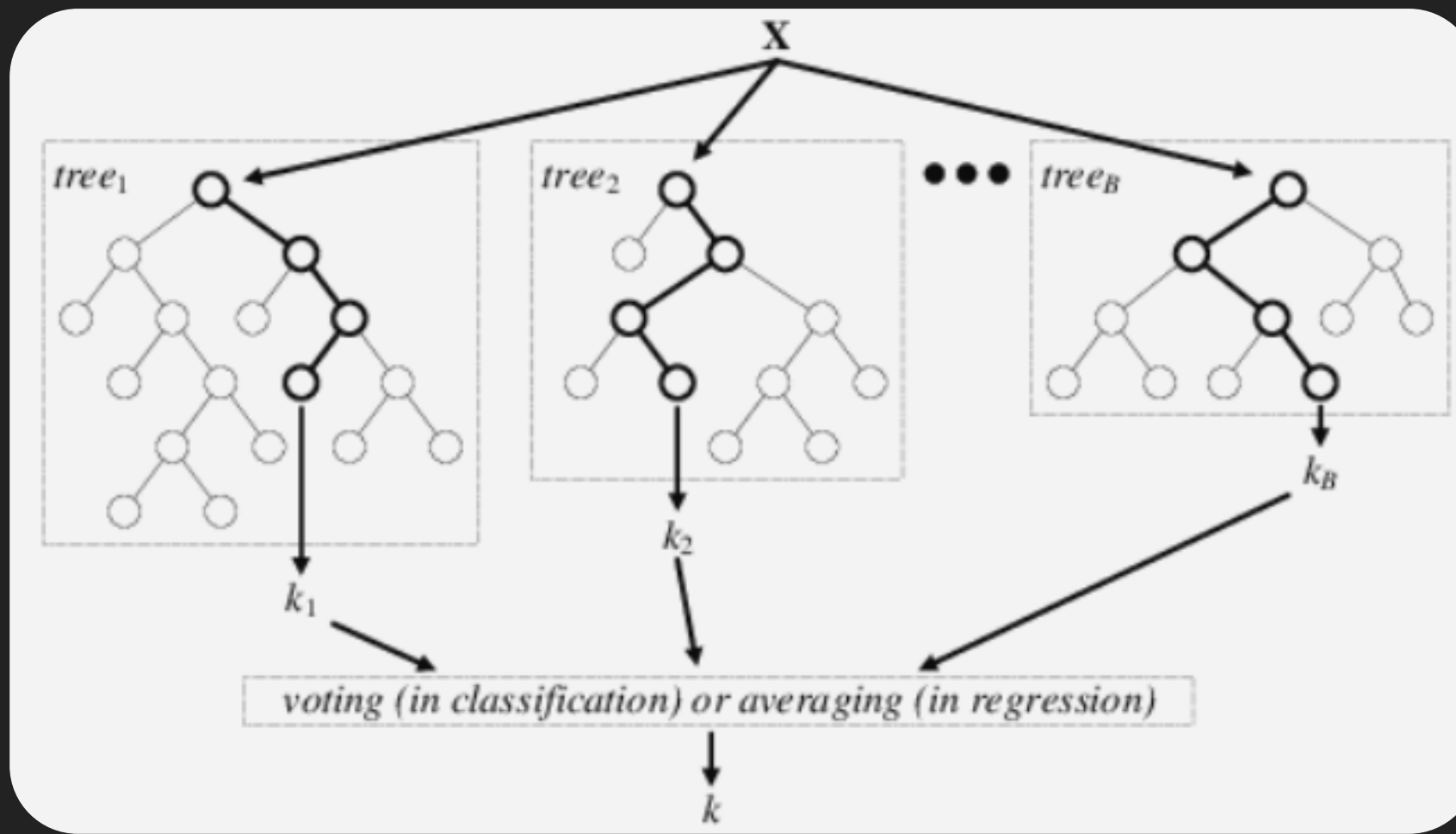


# Model 2: Random Forest Regressor

Result:

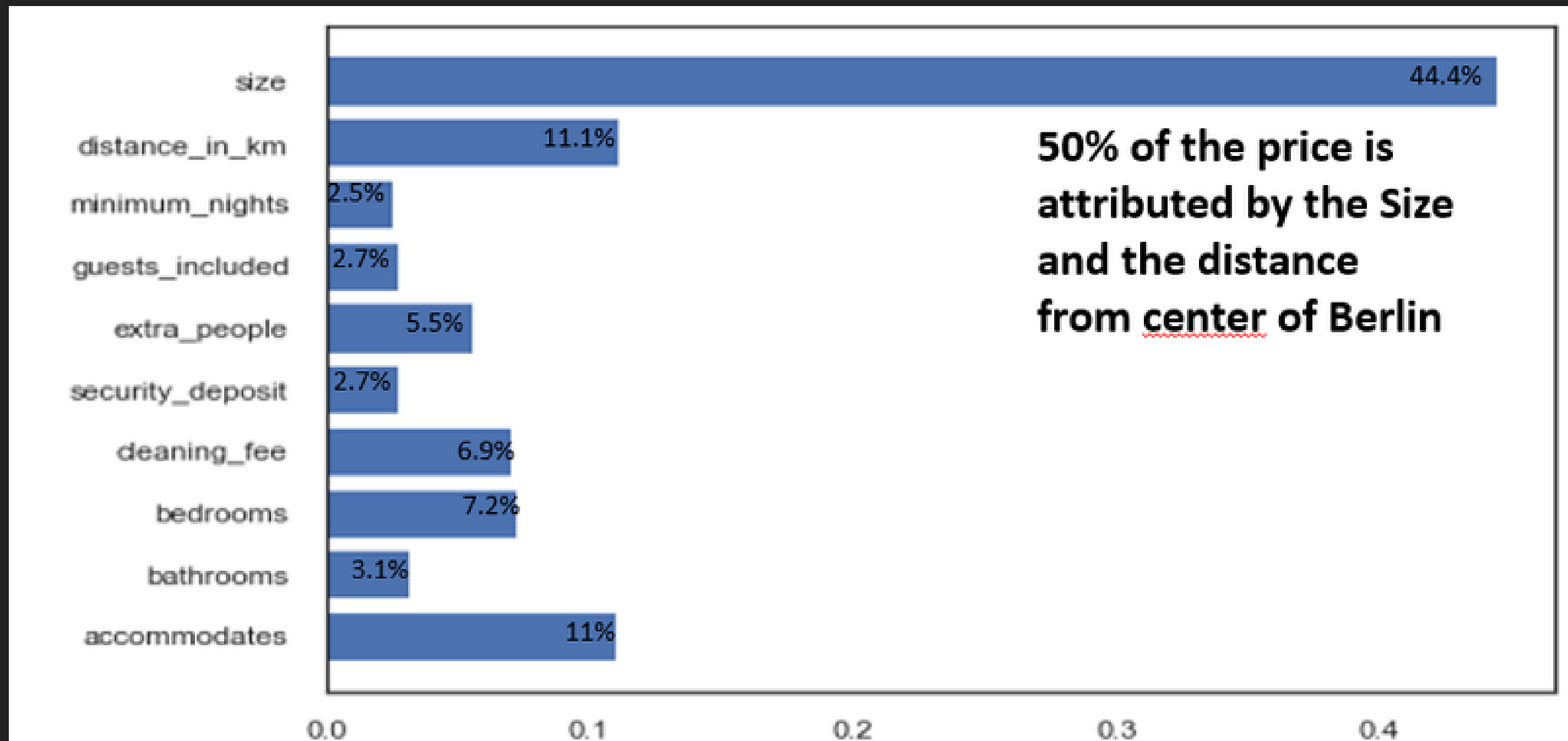
Mean absolute Percentage  
Error  
(MAPE):

Accuracy : 90.94%





# Features importance:





# Interesting Insights



21

## Location

Surprisingly, prices are affected a lot more by the size of the room than location from central of city.

## Bathroom

Weirdly, bathroom does not affect price as much. As I feel bathroom, is for basic personal hygiene. Probably affect more on longer-term guests or with bigger group of guest.

## Seasonality

High season lodging for Airbnb in Berlin is between May - July and September - October

## Super host

Host with this status generally have higher occupancy rate. Almost twice as much as other normal host.



# The Next Steps

22

Use data for Airbnb in other country to compare prices, location, etc

Take into consideration review from consumers

Do more other models to compare the results







23

# Thank You





# References:

- [https://www.kaggle.com/brittabetendorf/berlin-airbnb-data?select=listings\\_summary.csv](https://www.kaggle.com/brittabetendorf/berlin-airbnb-data?select=listings_summary.csv)
- <https://towardsdatascience.com/random-forest-in-python-24d0893d51c0>
- <https://www.ryerson.ca/news-events/news/2016/10/why-tourists-choose-airbnb-over-hotels/#:~:text=Tourists%20are%20mostly%20motivated%20to,convenient%20location%2C%20and%20household%20amenities.&text=61%20percent%20chos>  
e%20Airbnb%20as,the%20length%20of%20their%20trip
- <https://towardsdatascience.com/random-forest-ca80e56224c1>
- <https://mljar.com/blog/feature-importance-in-random-forest/#:~:text=Random%20Forest%20Built%2Din%20Feature,from%20the%20Random%20Forest%20structure.&text=The%20average%20over%20all%20tree>  
s,measure%20of%20the%20feature%20importance.