



# Ironhack Data Analytics

## Final Project

### Carles Campdepadrós Martín



# Agenda

- Project Overview
- Data Selection & Preparation
- EDA
- Hypothesis Testing
- ML Models & App Testing
- Final Insights & Conclusions
- Challenges & Learnings
- Next Steps





# PROJECT OVERVIEW

## PROBLEM



## ACTUAL SOLUTIONS

# OBJECTIVES

?

## WHAT

- Price and Client Satisfaction prediction App

?

## WHY

- Pricing tools are paid versions or subjective
- No client satisfaction tools

?

## HOW

- Hypothesis Testing to check differences among factors
- Machine Learning Models to predict price and client satisfaction

Search or enter website name

# Airbnb Price Prediction

First Name

Last Name

Phone Number

Email

Accept Terms and Conditions

Search or enter website name

# Airbnb Price Prediction

User inputs

Estimated Price

98.35 €

per night

The Clean Up Man Cleaning

Commercial Janitorial Cleaning Services

As a leading provider of cleaning services, we take pride in offering the best professional service.

Services We Offer

Sed ut perspiciatis unde omnis iste natus error sit voluptatem accusamus et iusto odio dignissimos.

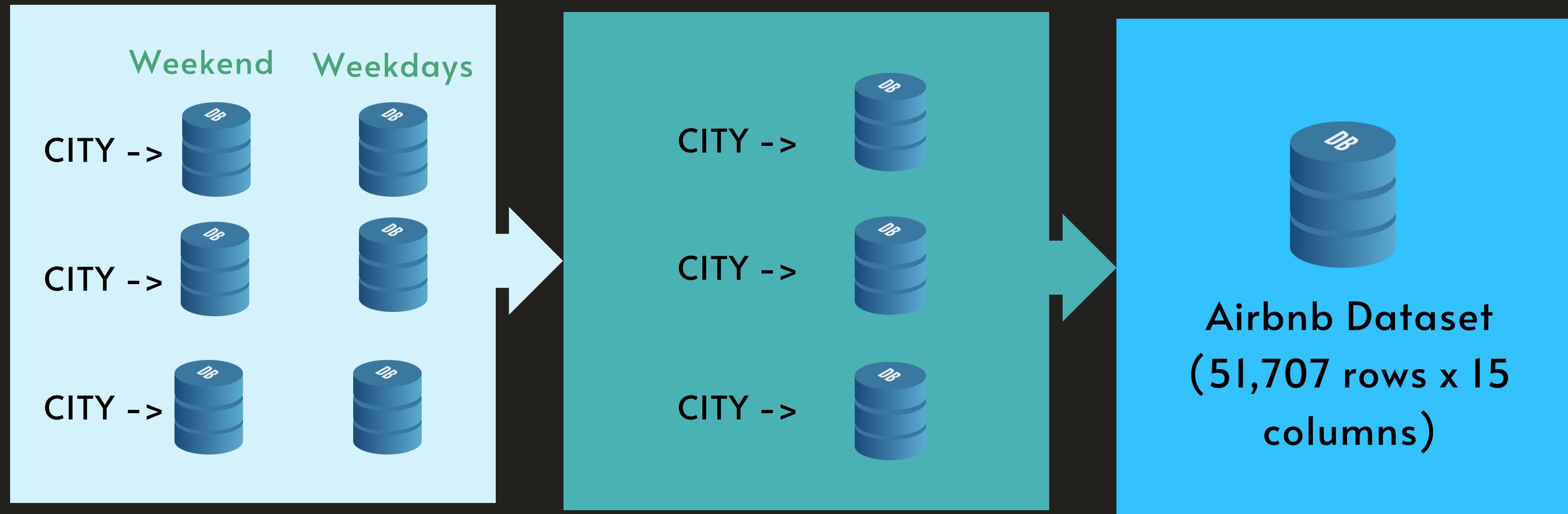
# DATA SELECTION & PREPARATION

kaggle™

	Weekdays.csv	Weekends.csv
Amsterdam	2	2
Athens	2	2
Barcelona	2	2
Berlin	2	2
Budapest	2	2
Lisbon	2	2
London	2	2
Paris	2	2
Rome	2	2
Vienna	2	2



# DATA PREPARATION

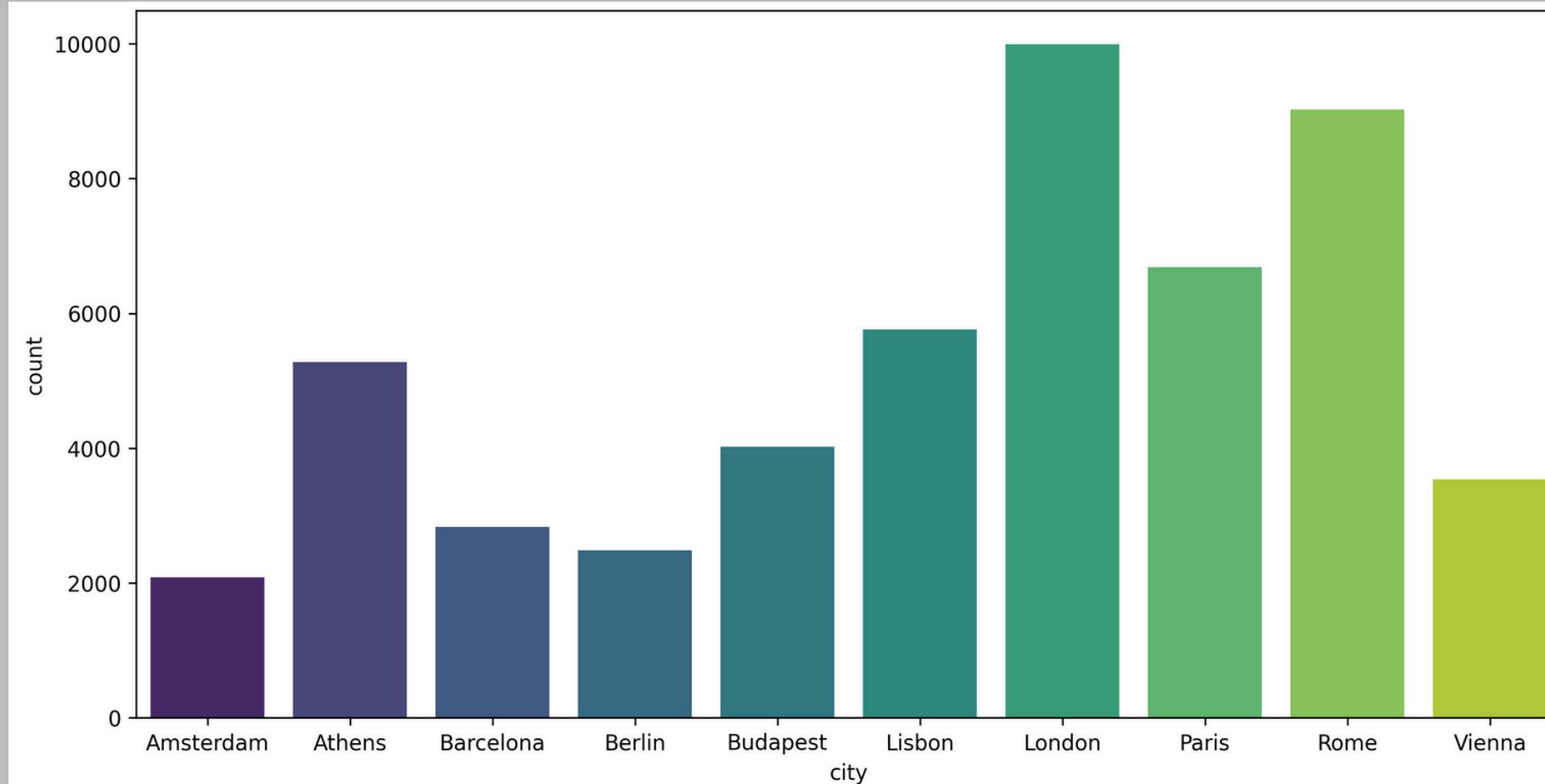




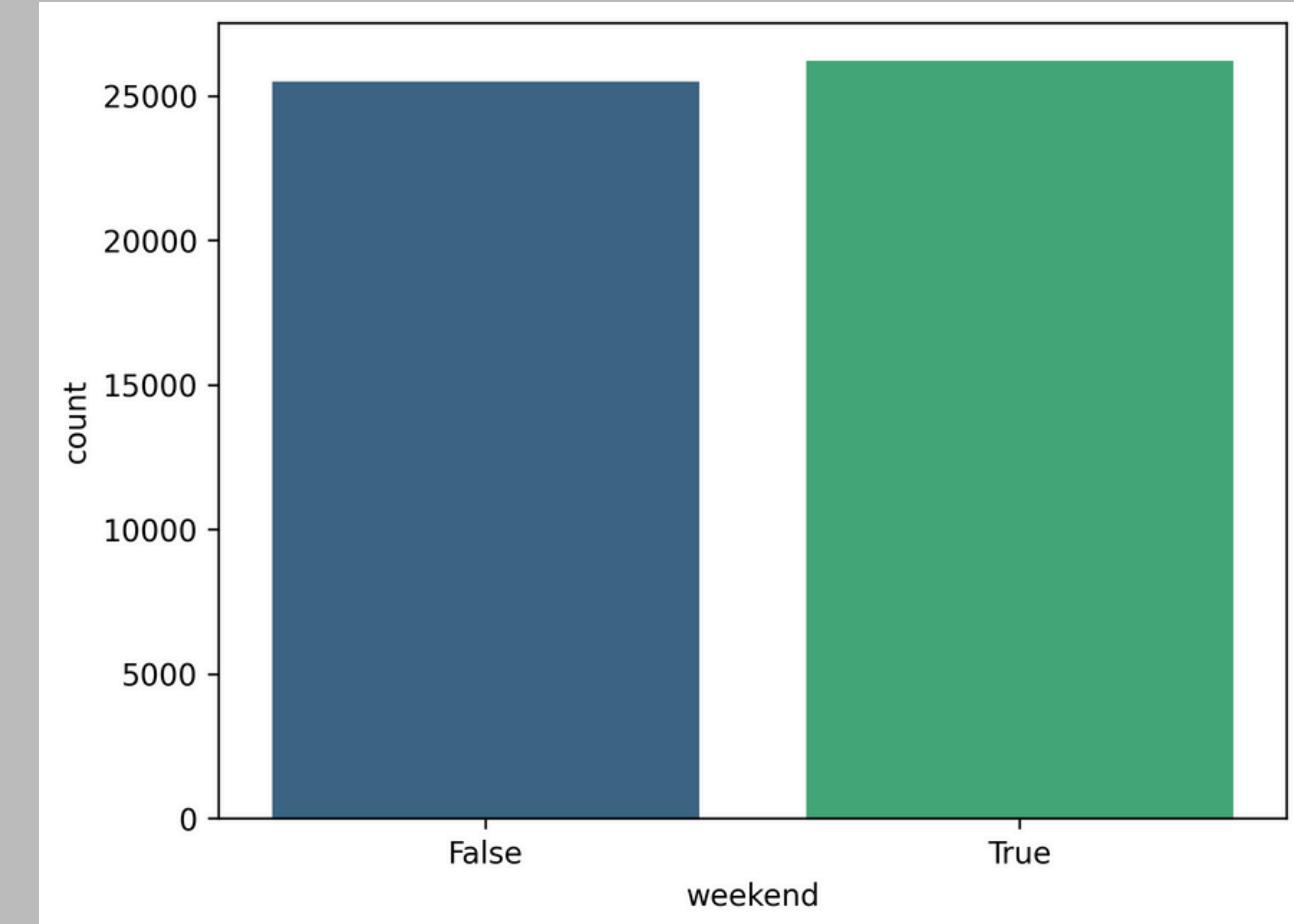
IRON  
HACK

EDA

# MAIN VARIABLES UNIVARIATE ANALYSIS



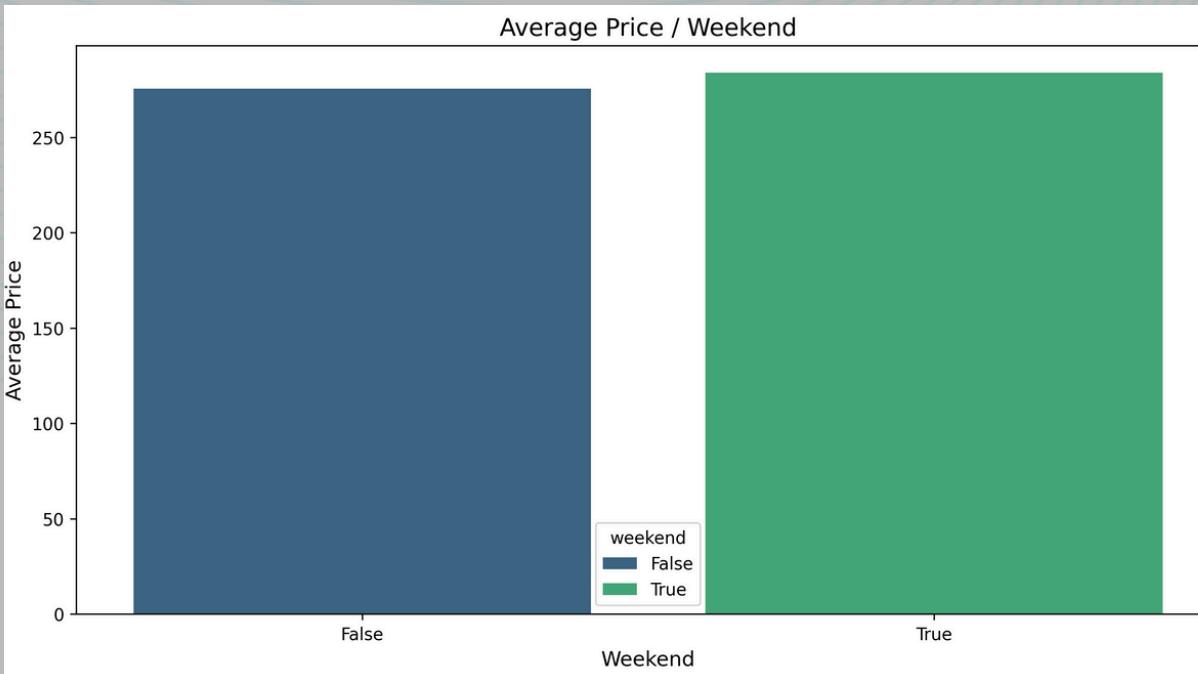
City



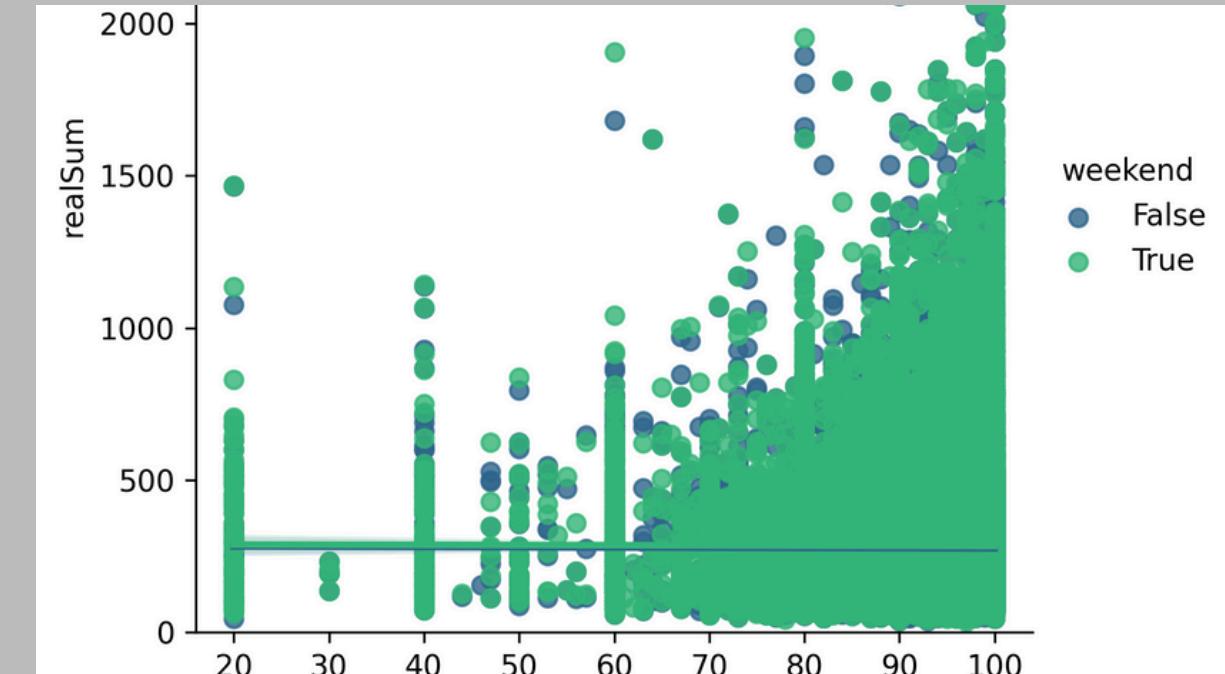
Weekend



# Superhost

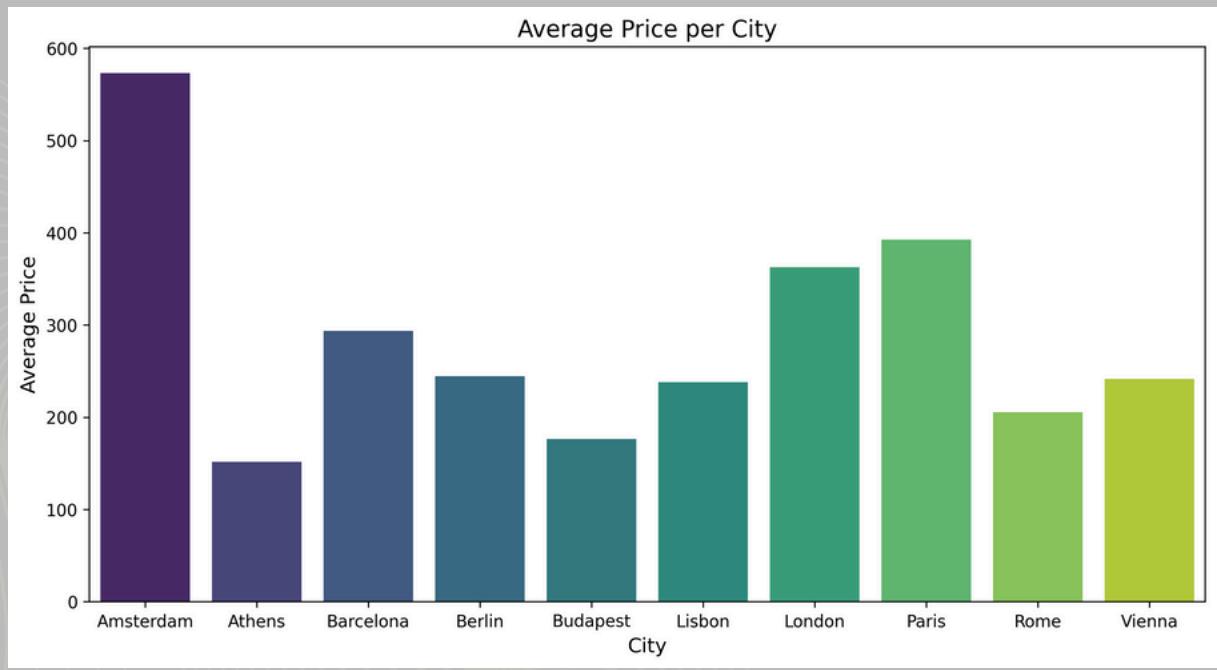


# Weekend

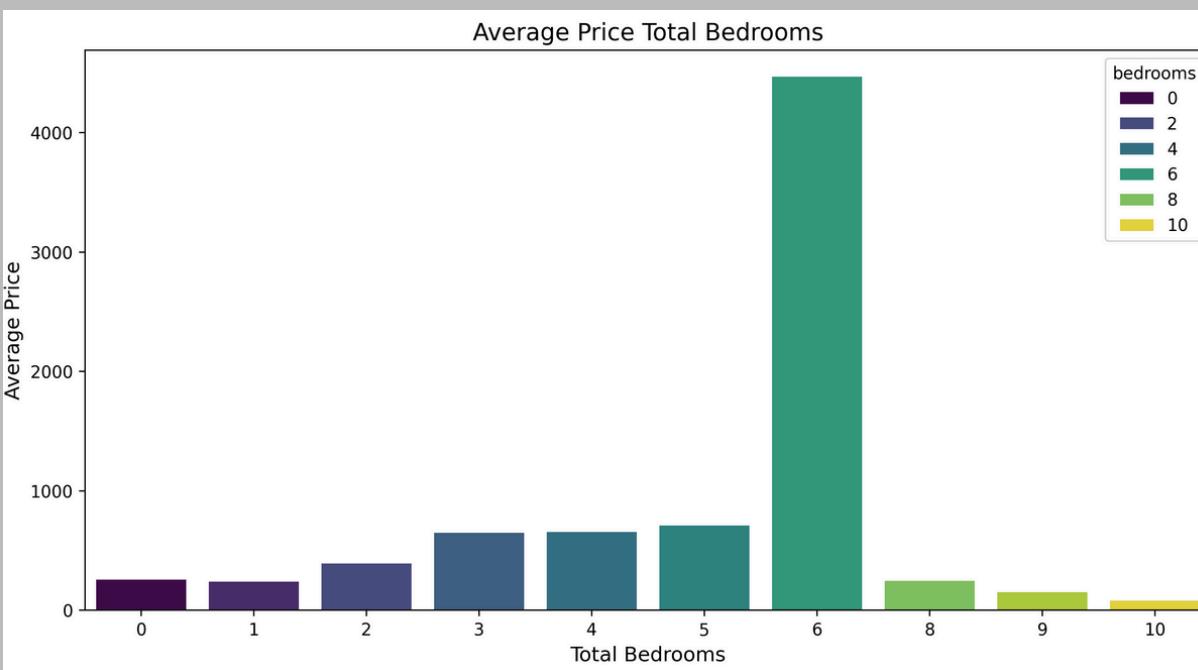


# Guest Satisfaction

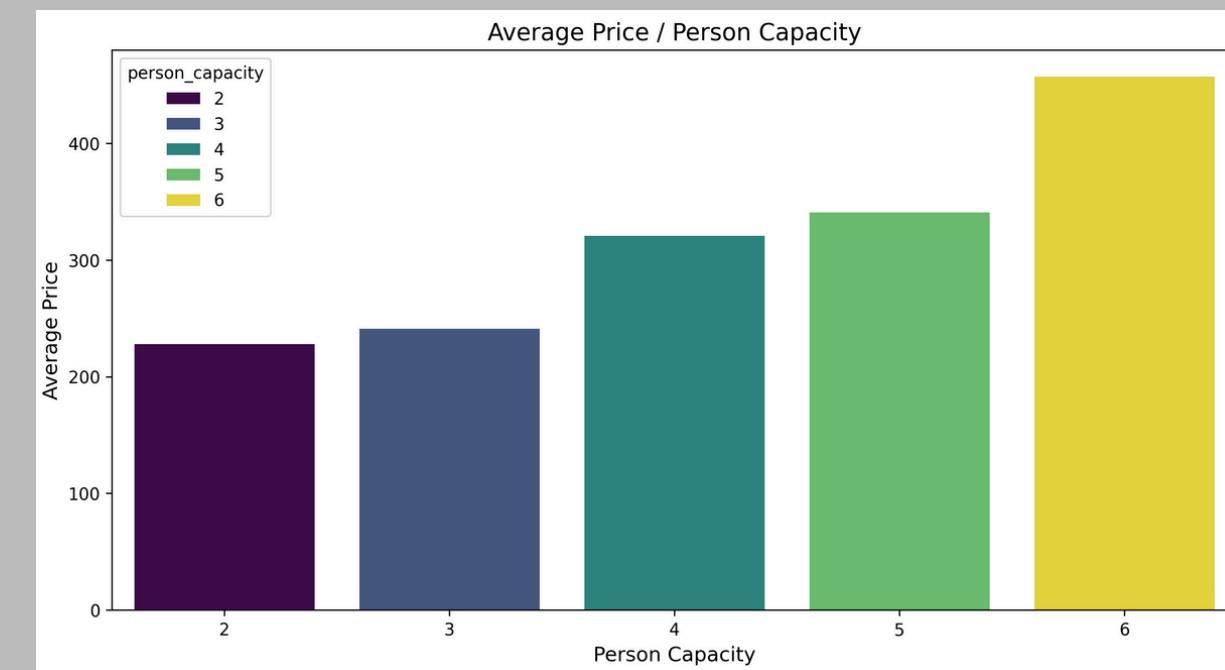
## PRICE BIVARIATE



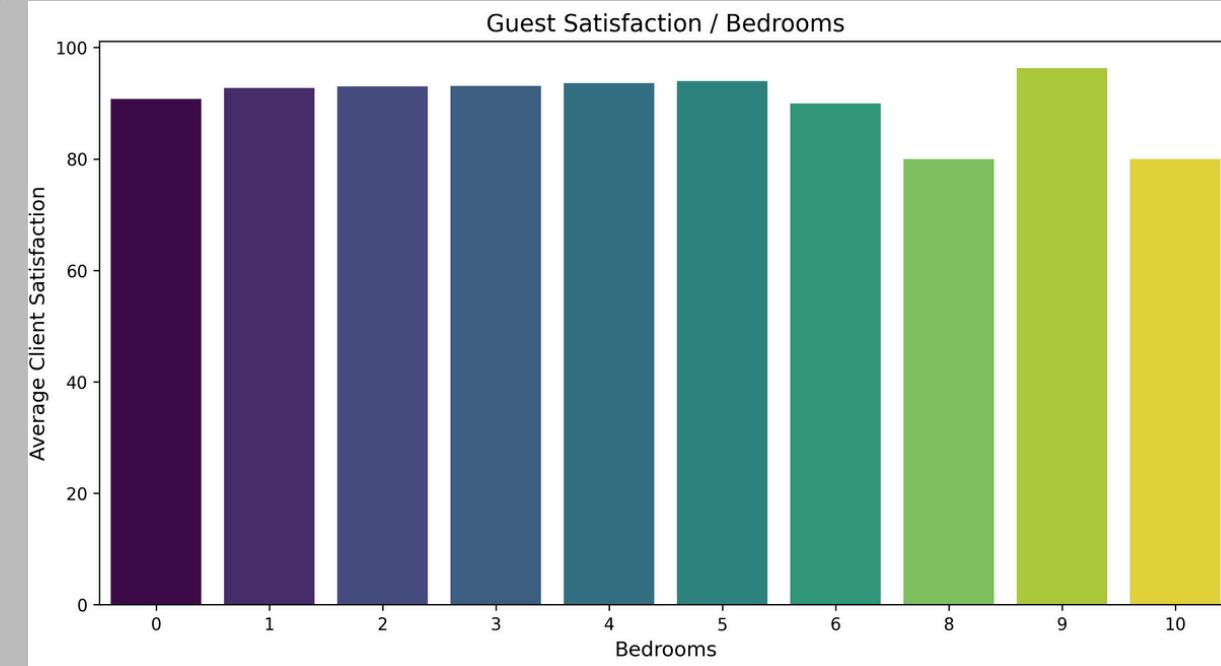
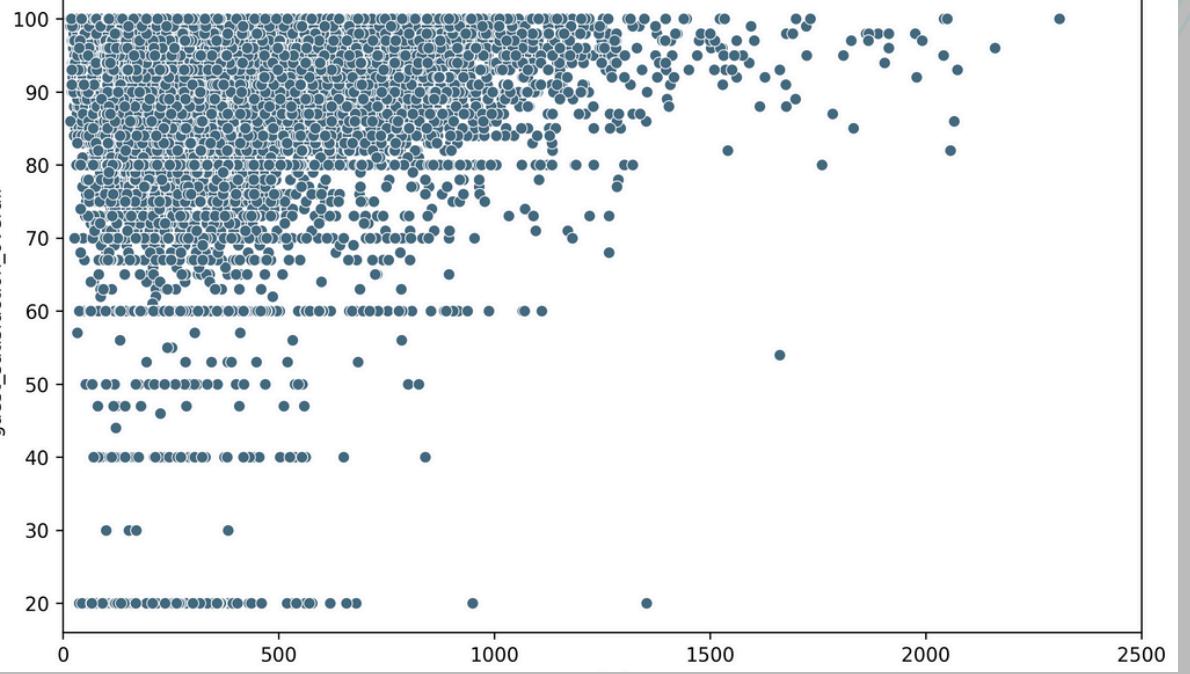
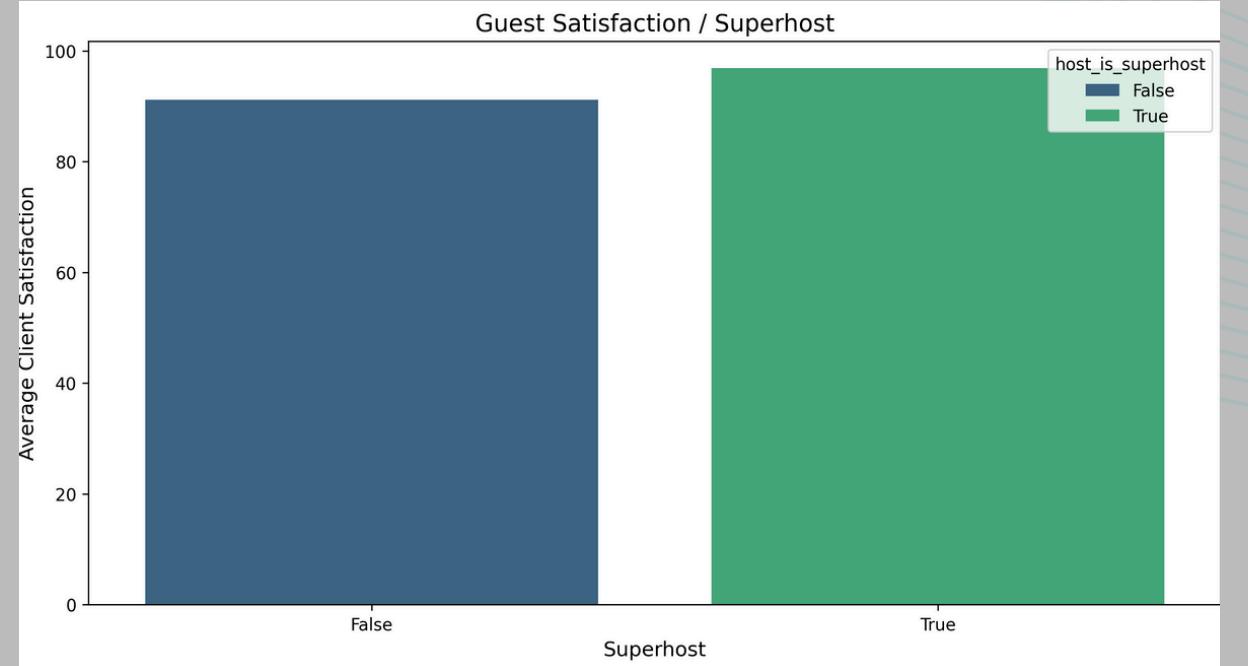
# City



# Bedrooms



# Capacity

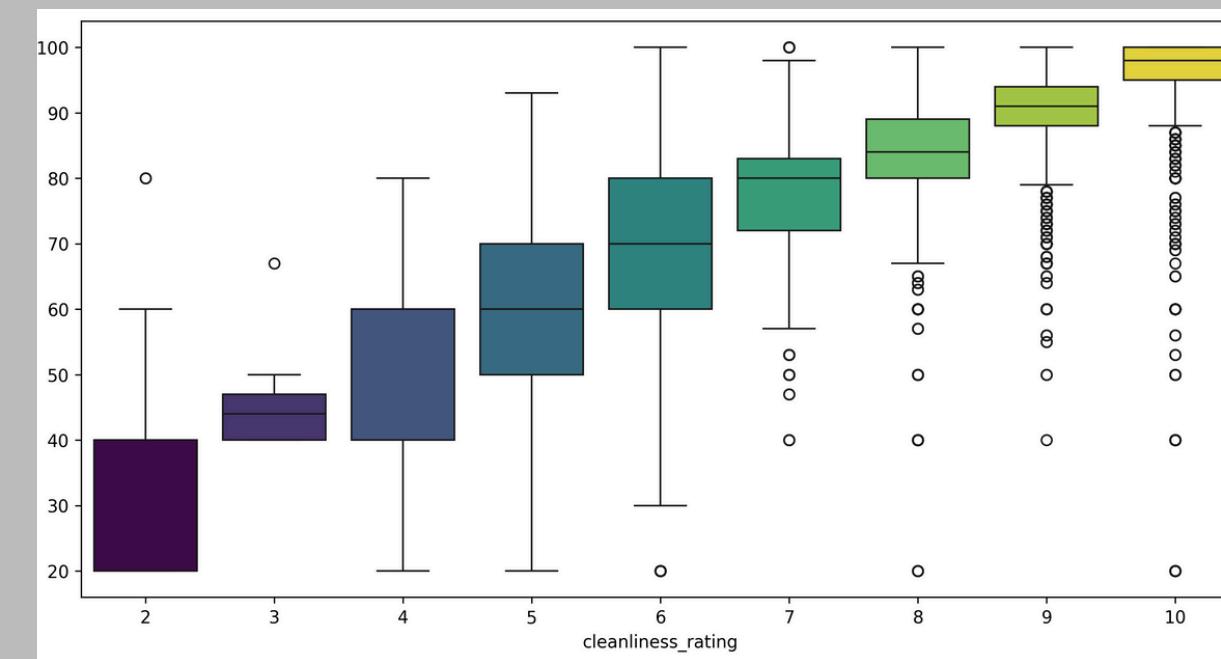
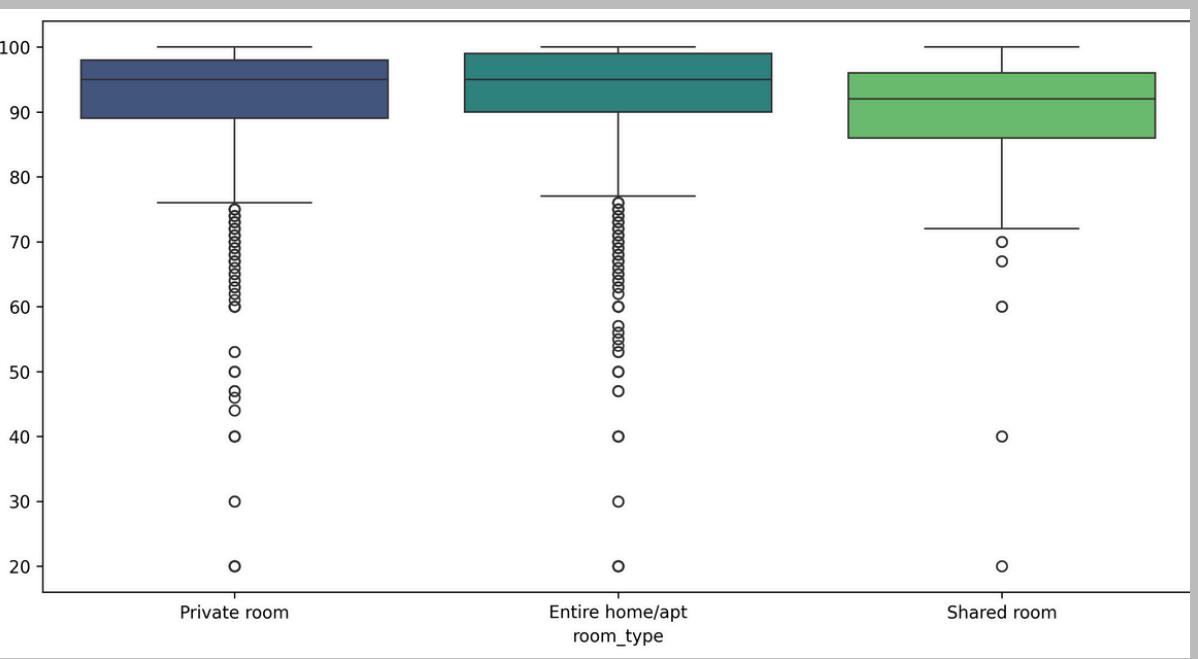
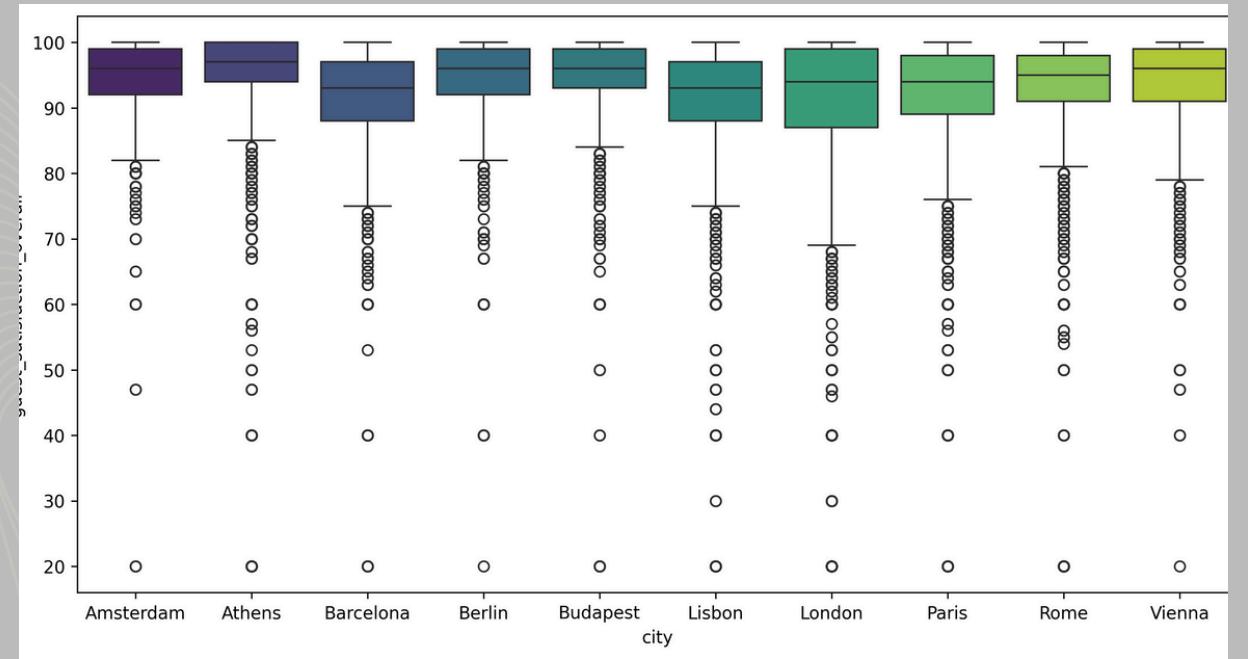


## Superhost

## Attraction Index

## Bedrooms

### C. SATISFACTION BIVARIATE

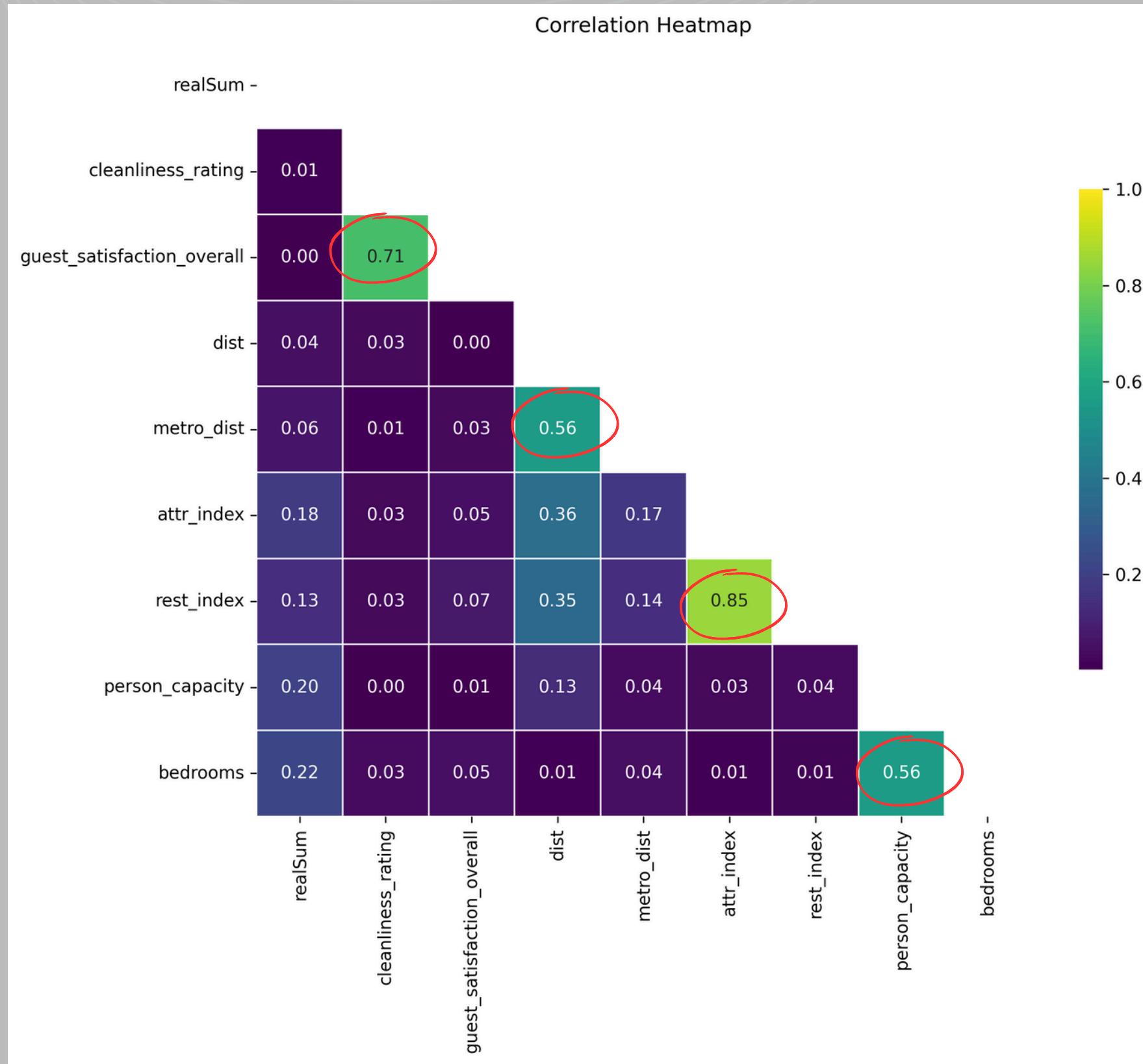


## City

## Room Type

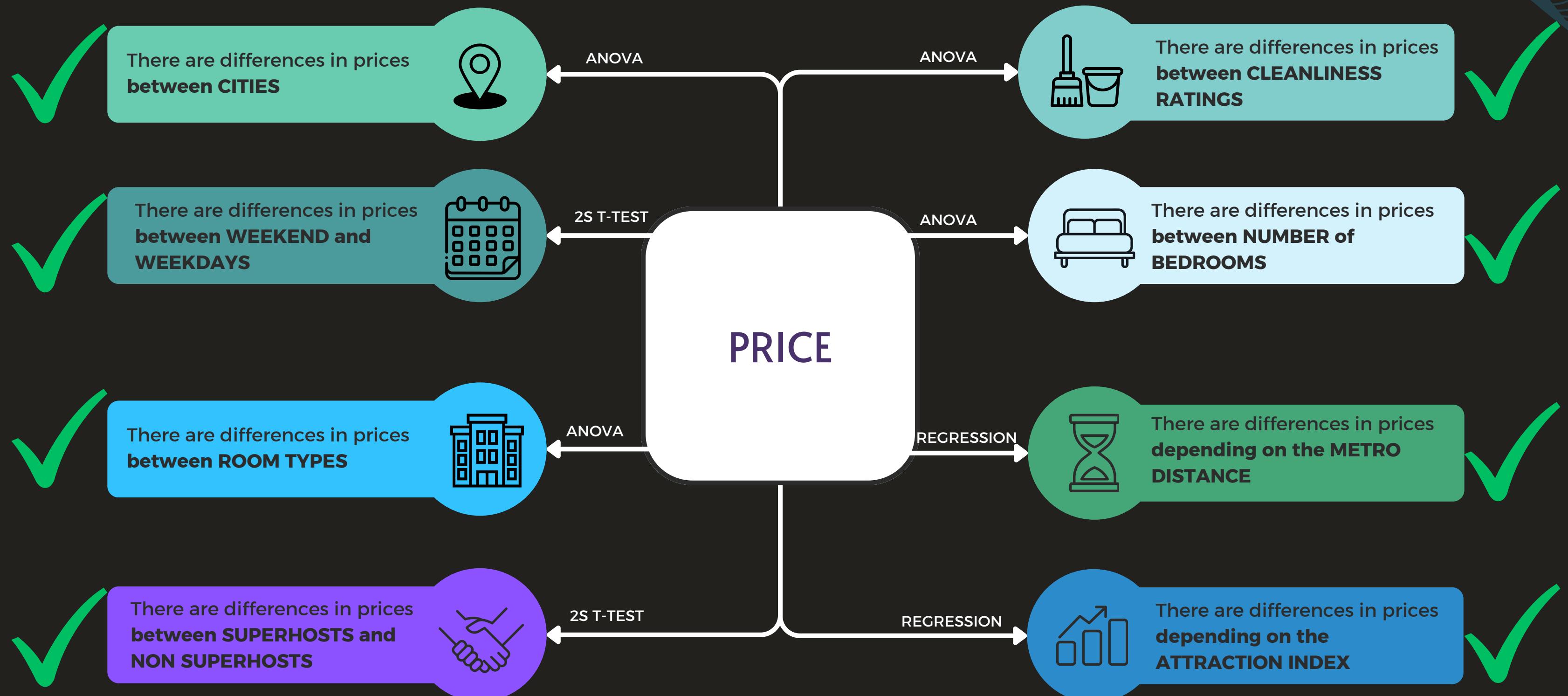
## Cleanliness Rate

# CORRELATION MATRIX



# HYPOTHESIS TESTING

# HYPOTHESIS TESTING - PRICE



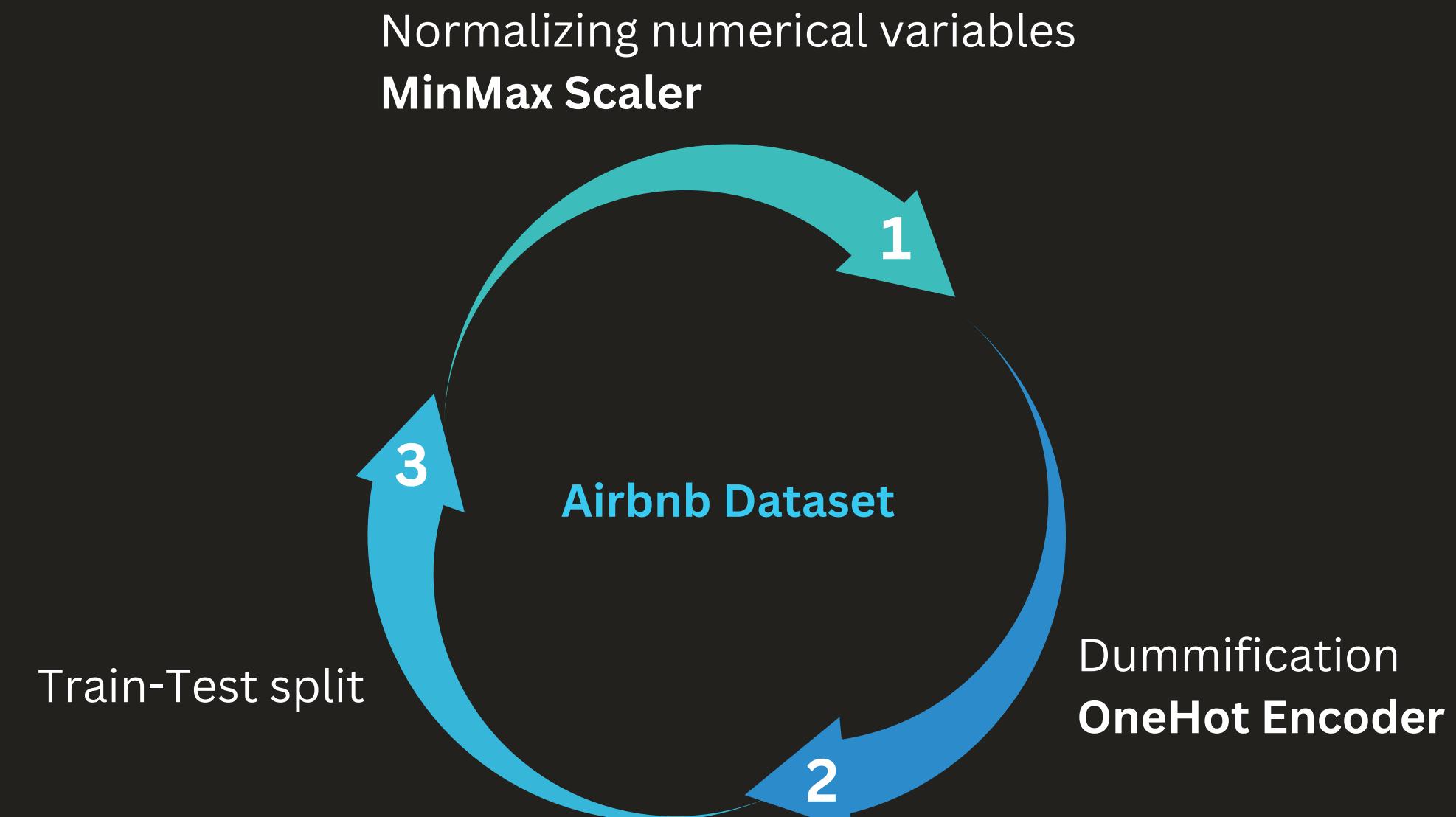
# HYPOTHESIS TESTING - CLIENT SATISFACTION



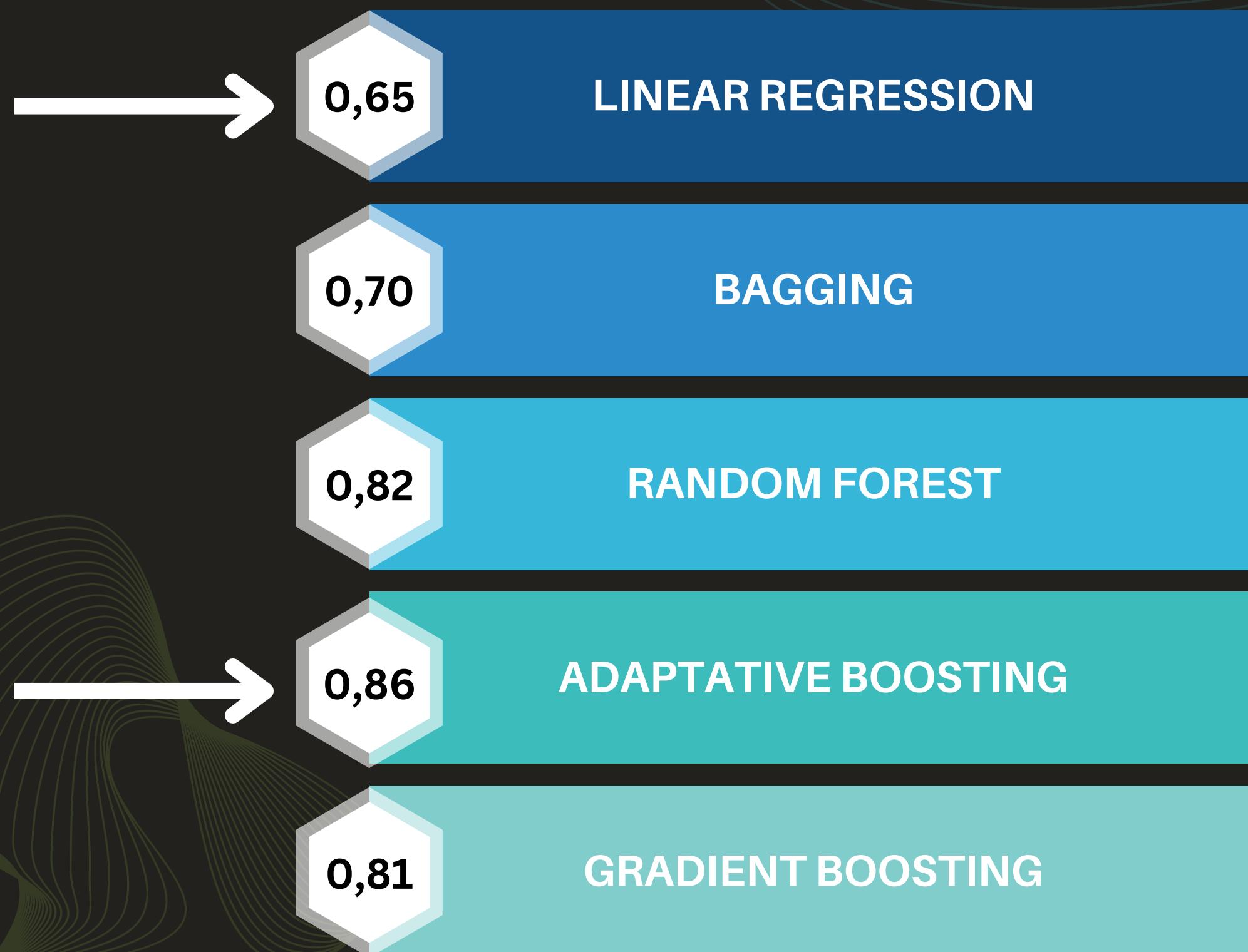
# ML MODELS & APP TESTING

A photograph of a modern, multi-story apartment building with a light-colored facade and dark horizontal cladding. A green tree is visible in the foreground.

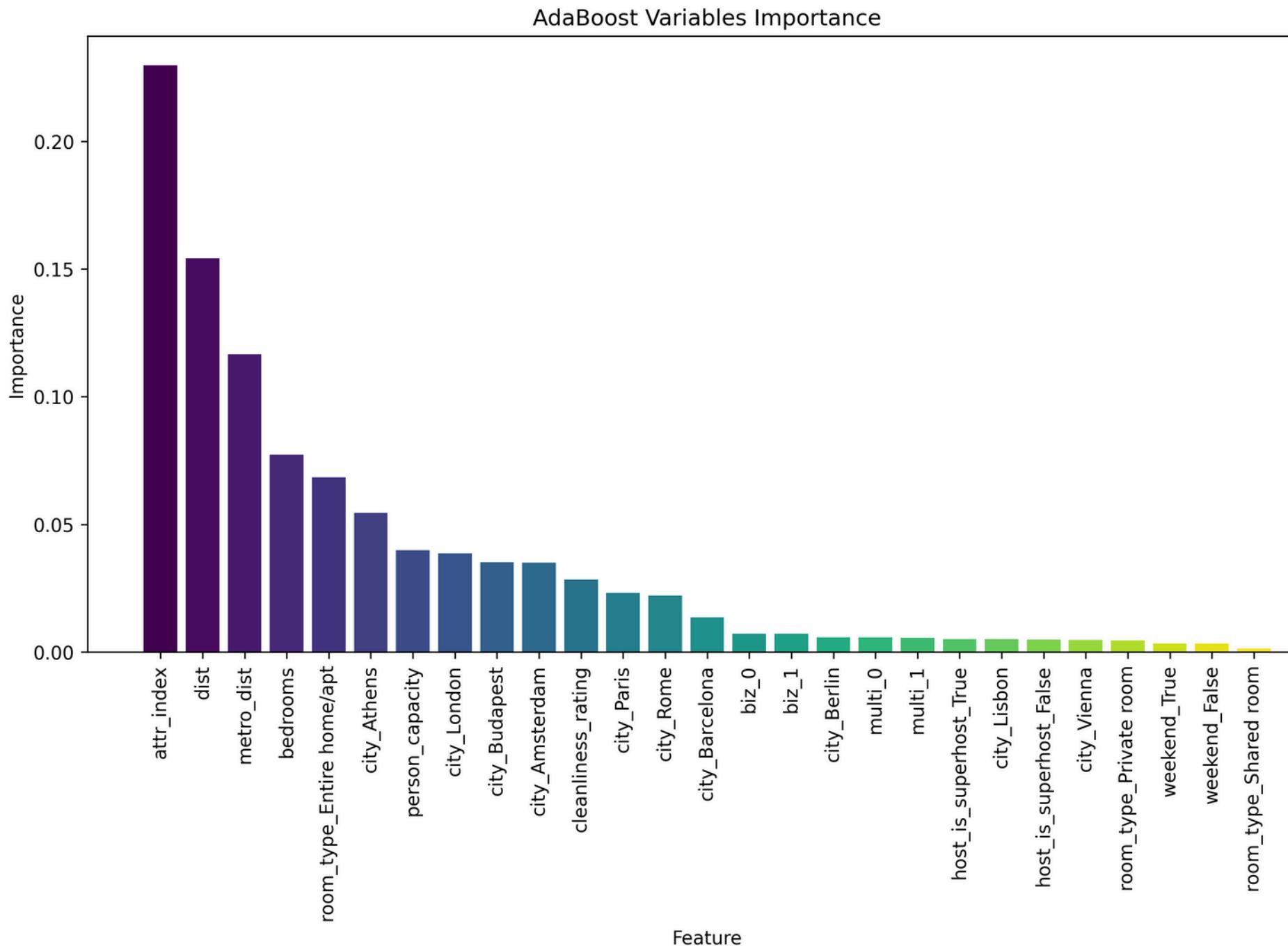
# FEATURE ENGINEERING AND SELECTION



# ML MODELS- $\log(\text{PRICE})$



# ADAPTATIVE BOOSTING ANALYSIS - log(PRICE)



## HYPERPARAMETER TUNING

Grid Search (R2) = **0,86**

Random Search (R2) = **0,86**

## OVERFITTING

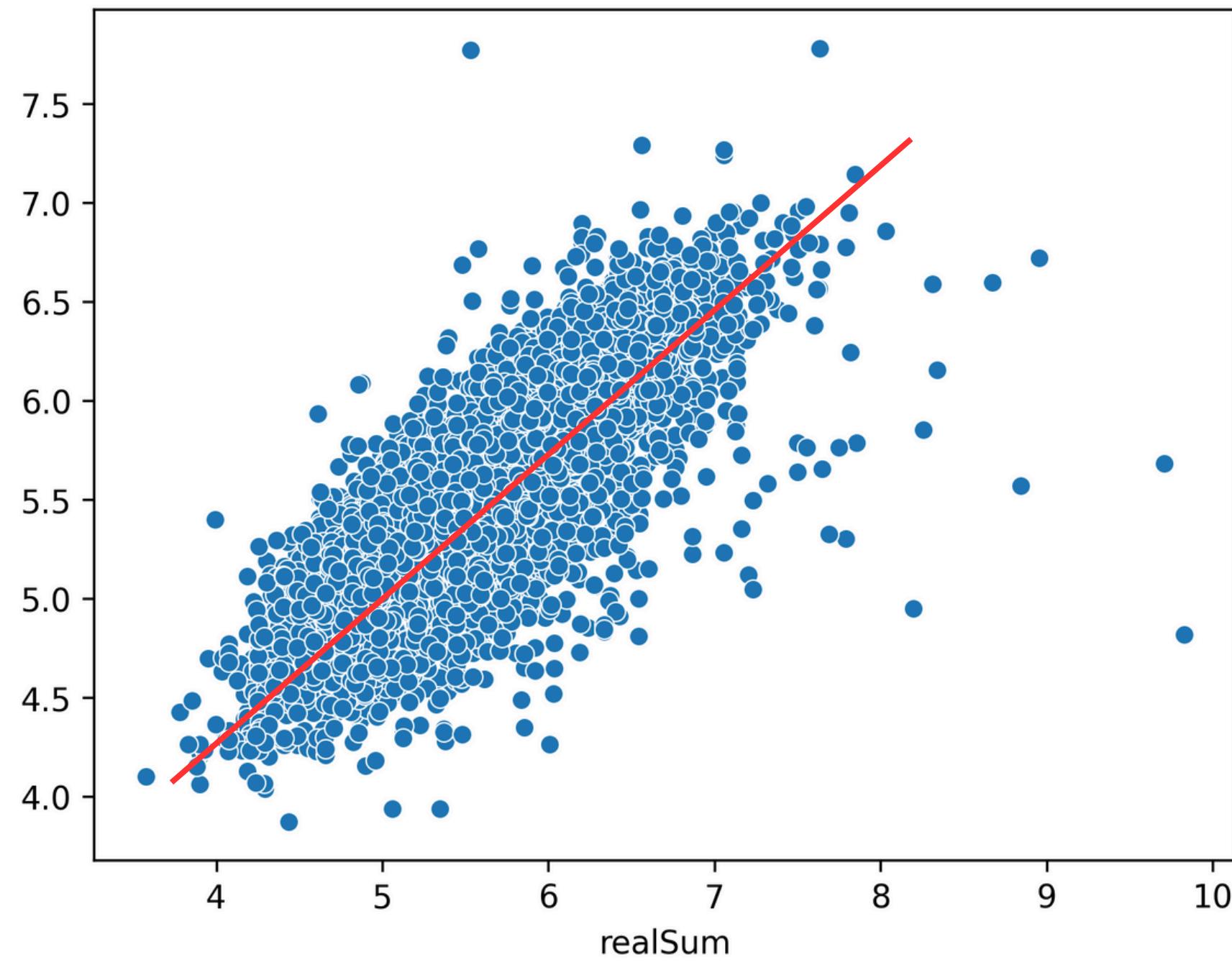
Test Set (R2) = **0,86**

Train Set (R2) = **0,99**



OVERFITTED

# LINEAR REGRESSION ANALYSIS - log(PRICE)



**OVERFITTING**

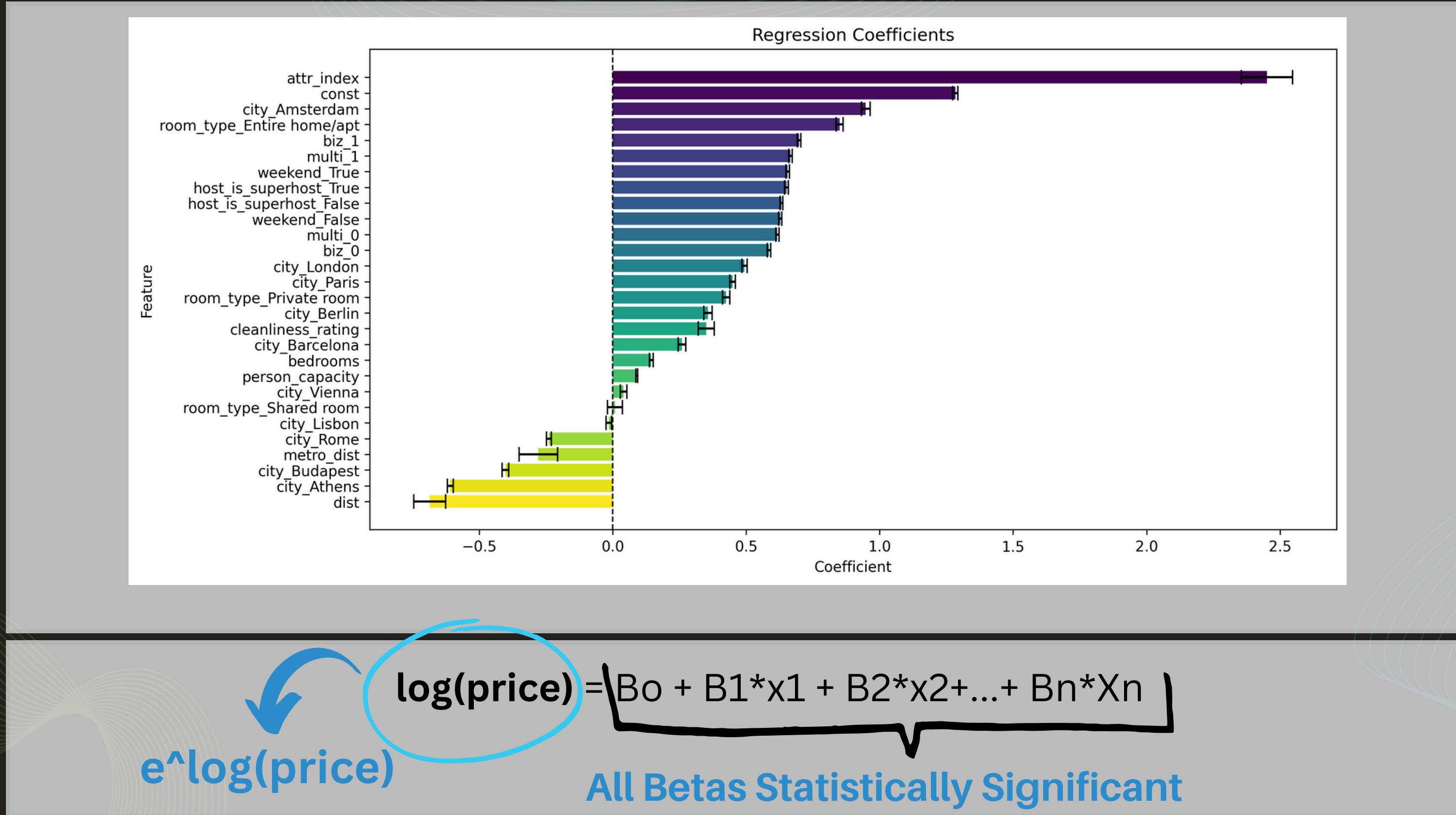
Test Set ( $R^2$ ) = 0,65

Train Set ( $R^2$ ) = 0,66

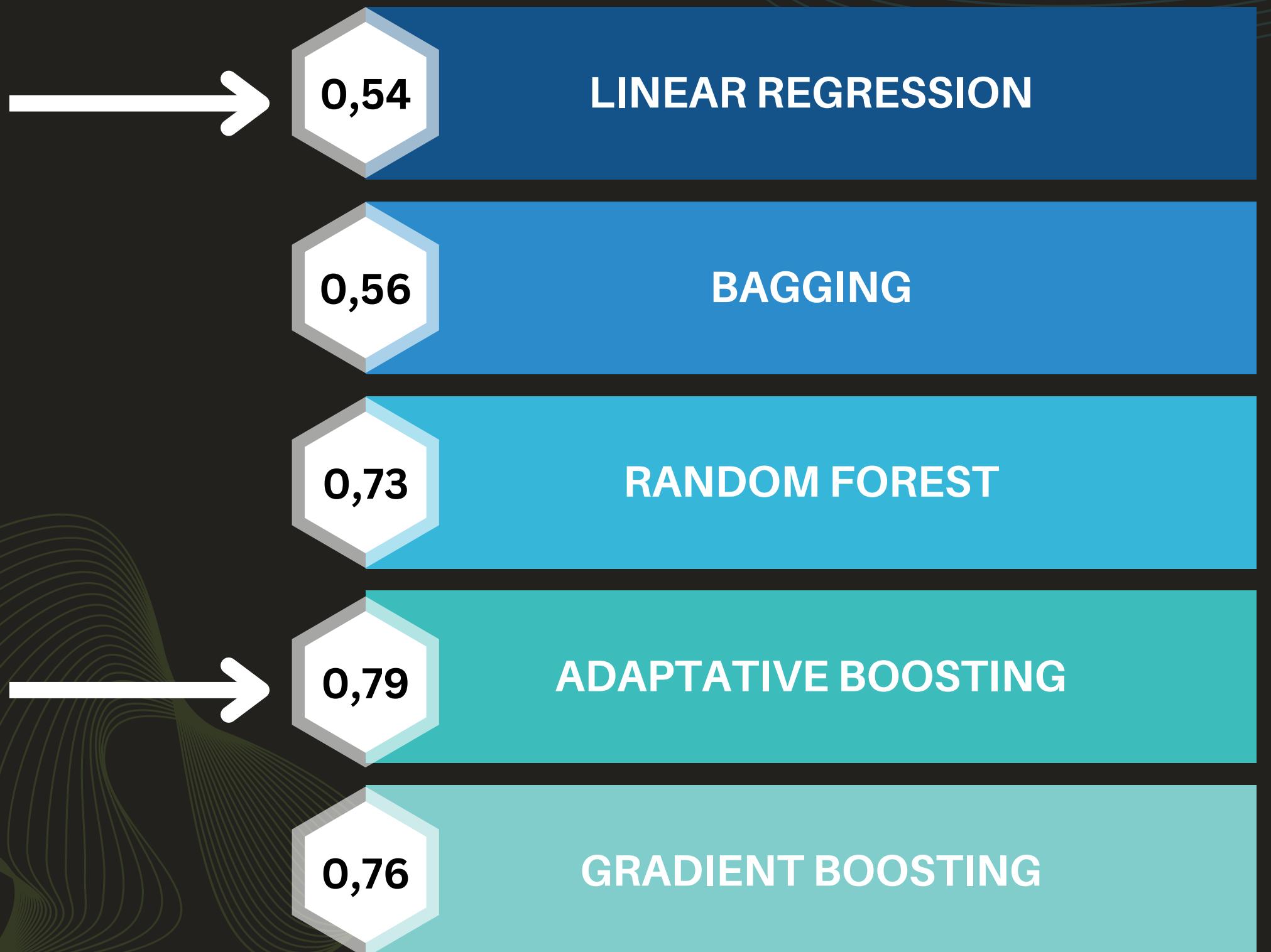


**NOT OVERFITTED**

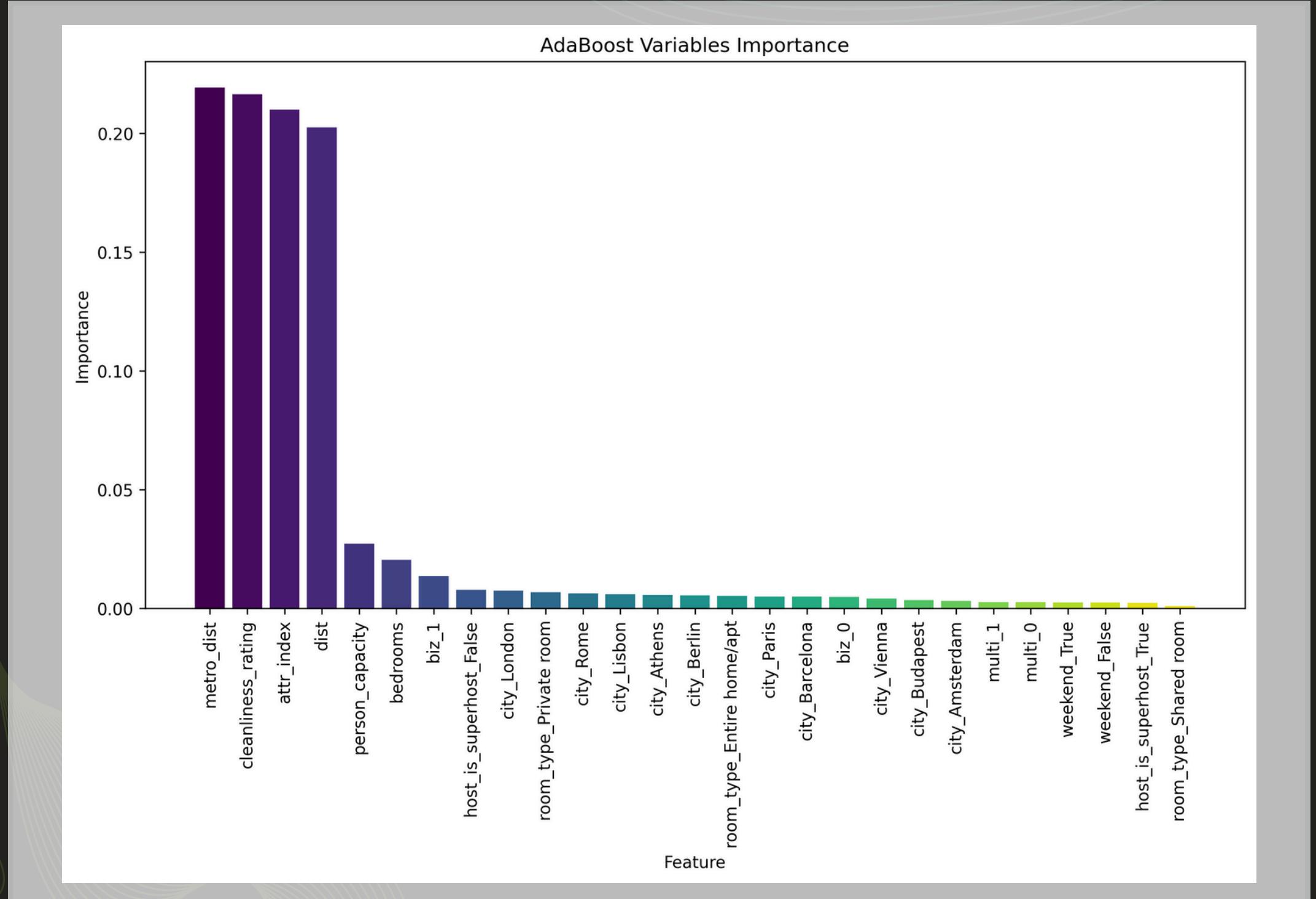
# LINEAR REGRESSION ANALYSIS - log(PRICE)



# ML MODELS- CLIENT SATISFACTION



# ADAPTATIVE BOOSTING ANALYSIS - CLIENT SATISFACTION



## HYPERPARAMETER TUNING

Grid Search (R2) = **0,80**

Random Search (R2) = **0,79**

## OVERFITTING

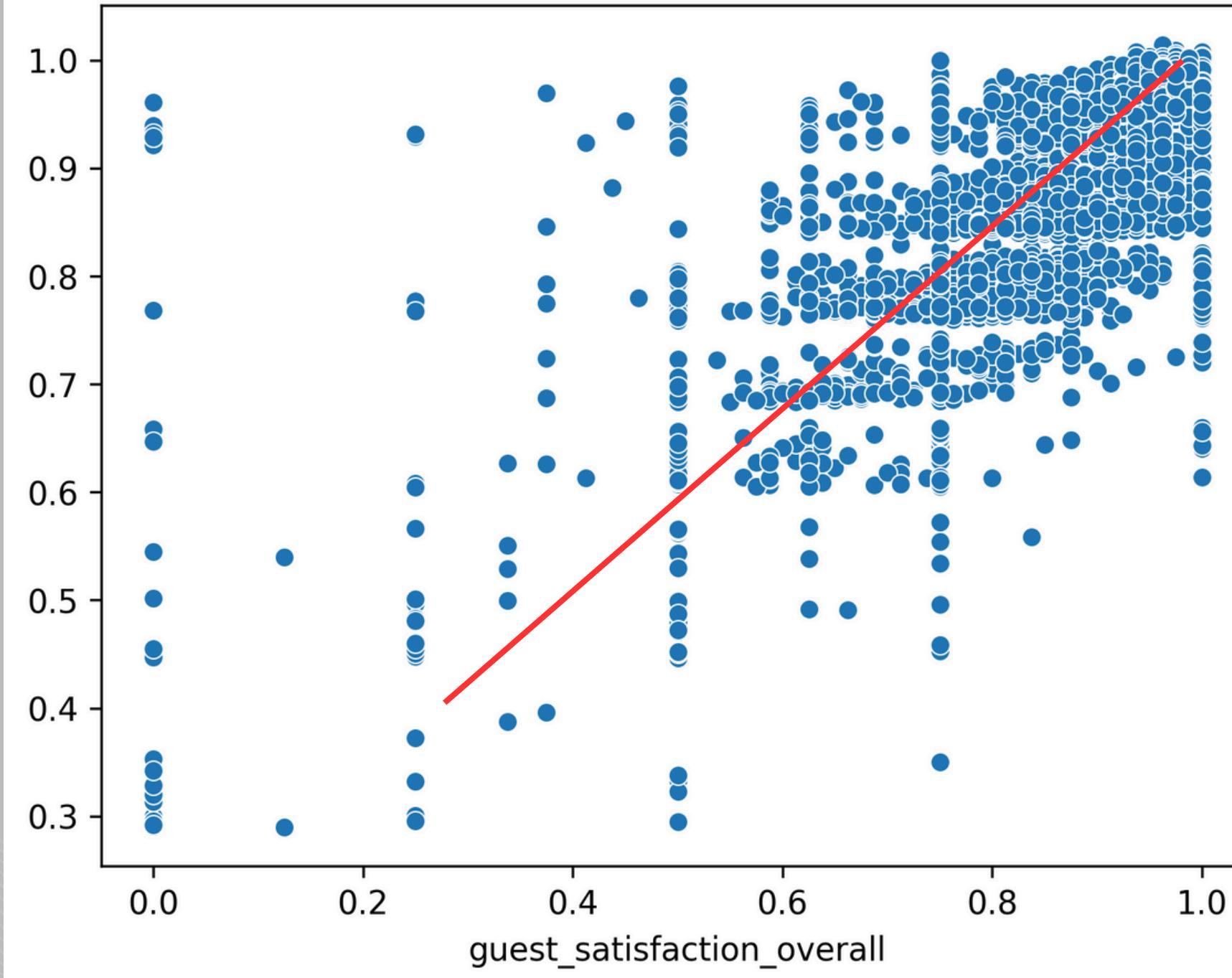
Test Set (R2) = **0,79**

Train Set (R2) = **0,97**



OVERFITTED

# LINEAR REGRESSION ANALYSIS - CLIENT SATISFACTION



OVERFITTING

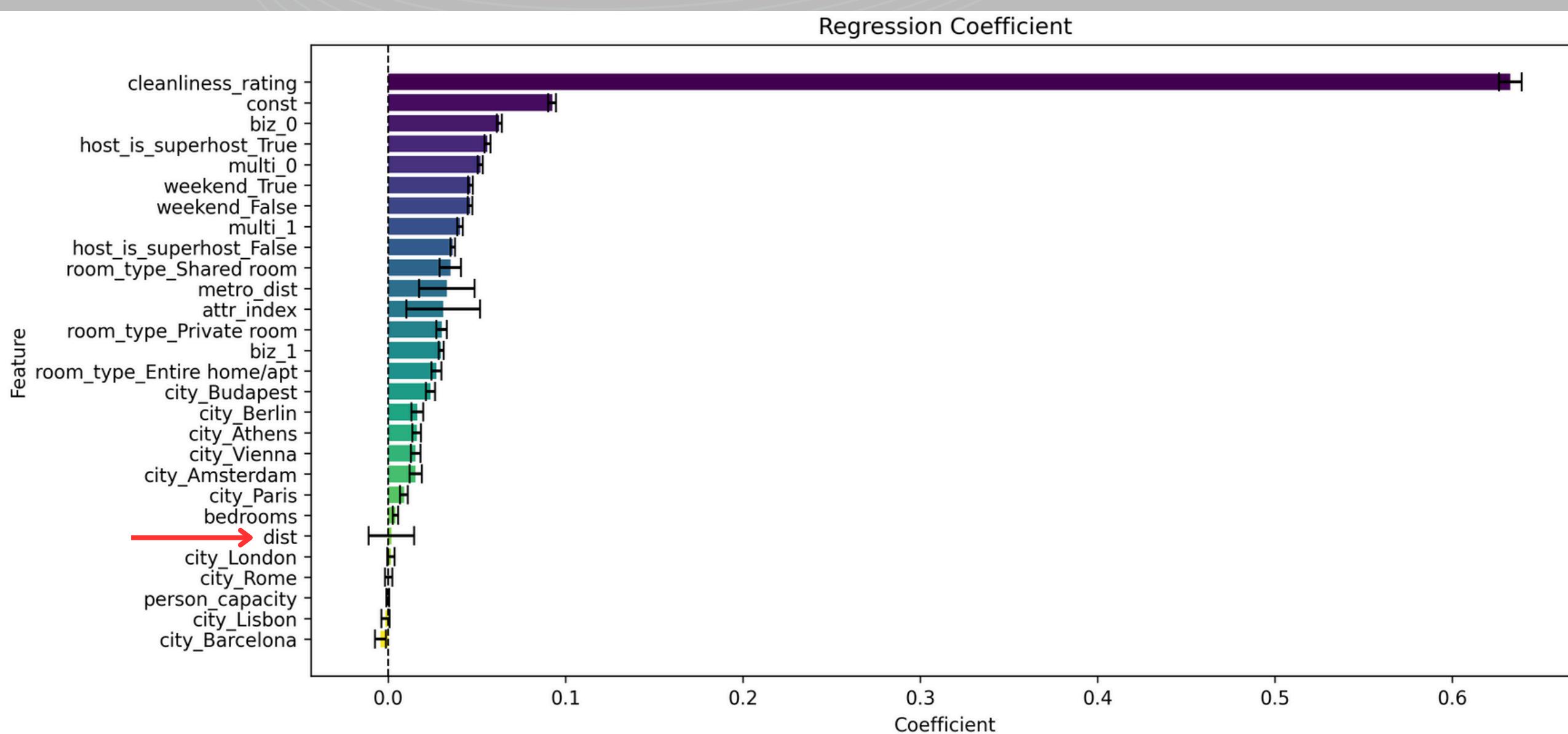
Test Set ( $R^2$ ) = 0,54

Train Set ( $R^2$ ) = 0,54



NOT OVERFITTED

# LINEAR REGRESSION ANALYSIS – CLIENT SATISFACTION



client satisfaction =  $B_0 + B_1*x_1 + B_2*x_2+...+ B_n*X_n$

All Betas NOT Statistically Significant → delete variables



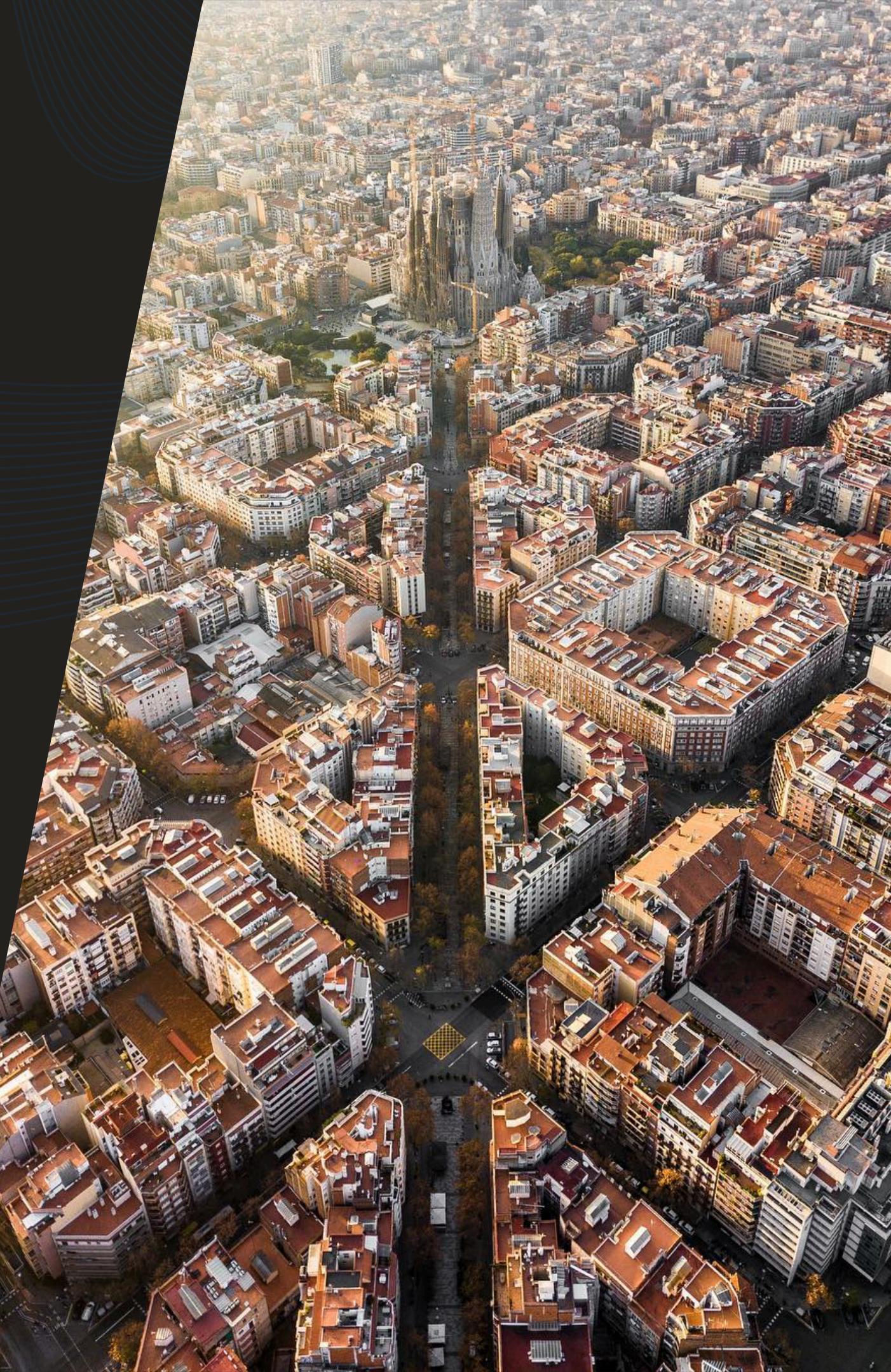
# APPs

<https://priceairbnbreg.streamlit.app>

<https://finalprojectironhacks.s3-us-west-2.amazonaws.com/finalprojectironhacks.streamlit.app>



# Streamlit



# FINAL INSIGHTS & CONCLUSIONS

# CROSS VALIDATION - AIRBNB PRICING

ATHENS

Enter the listing details

City: Athens

Room Type: Entire home/apt

Person Capacity: 2

Cleanliness Rating: 5.00

## Airbnb Price Prediction

Estimated Price

108.97 €

per night

Podrías ganar **96 €** en Airbnb con tu alojamiento

1 noche · 96 €/noche  
Descubre cómo [calculamos los ingresos](#)

Atenas · Alojamiento entero · 2 dormitorios

BARCELONA

Enter the listing details

City: Barcelona

Room Type: Entire home/apt

Person Capacity: 2

Cleanliness Rating: 5.00

## Airbnb Price Prediction

Estimated Price

261.11 €

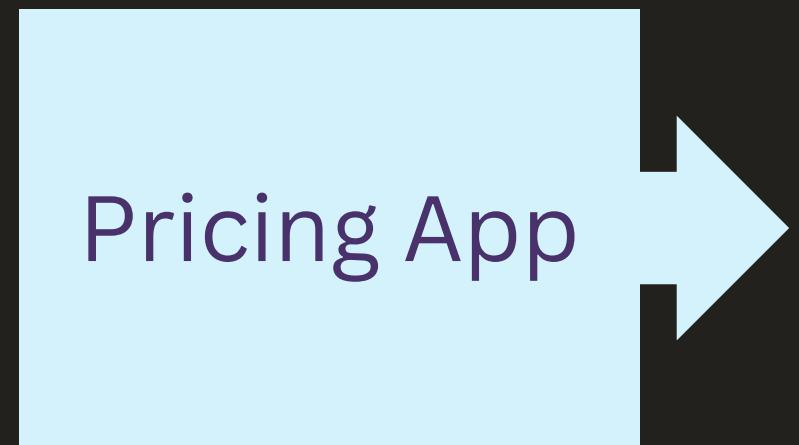
per night

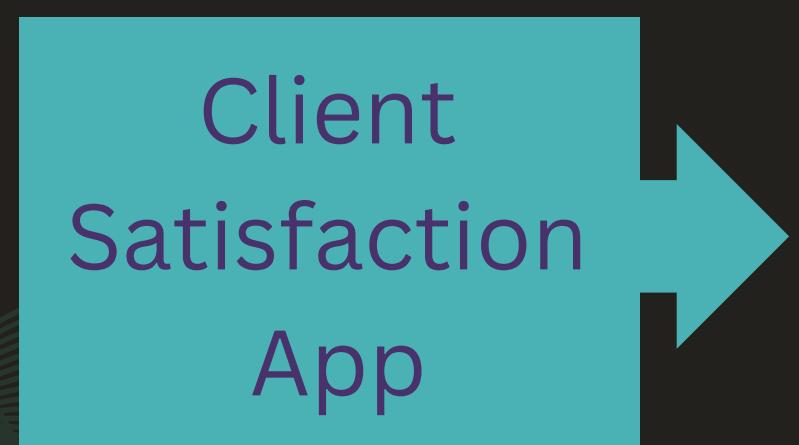
Podrías ganar **216 €** en Airbnb con tu alojamiento

1 noche · 216 €/noche  
Descubre cómo [calculamos los ingresos](#)

Barcelona · Alojamiento entero · 2 dormitorios

# CONCLUSIONS

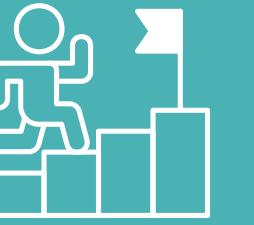


- High competition
  - Value added is the cost
  - Dynamic trends difficult to capture
- 
- 
- Client Satisfaction App
- Market Gap
  - Value added for sensitivity analysis
  - More static factors (easier to measure)



# CHALLENGES & LEARNINGS

# CHALLENGES & LEARNINGS



## Challenges

- Adding OneHot Encoder and MinMax Scaler to streamlit
- Dealing with logs
- Overfitting



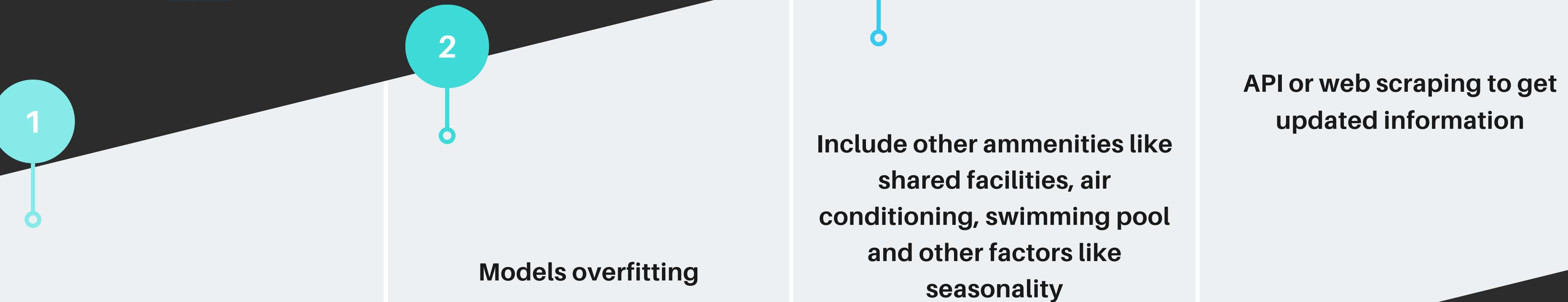
## Learnings

- Client satisfaction can be measured and predict it
- The importance of pricing factors is more spread out
- Streamlit Apps



# NEXT STEPS

# NEXT STEPS



Reach out to companies for advertising opportunities and client information

Models overfitting

Include other ammenities like shared facilities, air conditioning, swimming pool and other factors like seasonality

API or web scraping to get updated information



Thank You!  
Final Project  
Carles Campdepadrós Martín