

# Predicting & Comparing House Prices in Istanbul and London

## DSA210

Yunus Emre Gök - 32028



# Project Overview

---

## Objective:

This project aims to compare residential property prices in London and Istanbul by analyzing key influencing factors such as size and number of bathrooms.

## Approach:

A combination of data preprocessing, exploratory data analysis (EDA), and machine learning was used to uncover trends and build predictive models.



# Data Sources & Preprocessing

## Datasets:

- [London Dataset – Kaggle](#)
- [Istanbul Dataset – Kaggle](#)

## Preprocessing :

- Standardized column names
- Converted Turkish Lira to GBP (1 GBP = 20.38 TL)
- Cleaned missing or invalid entries
- Unified units (e.g., square meters)





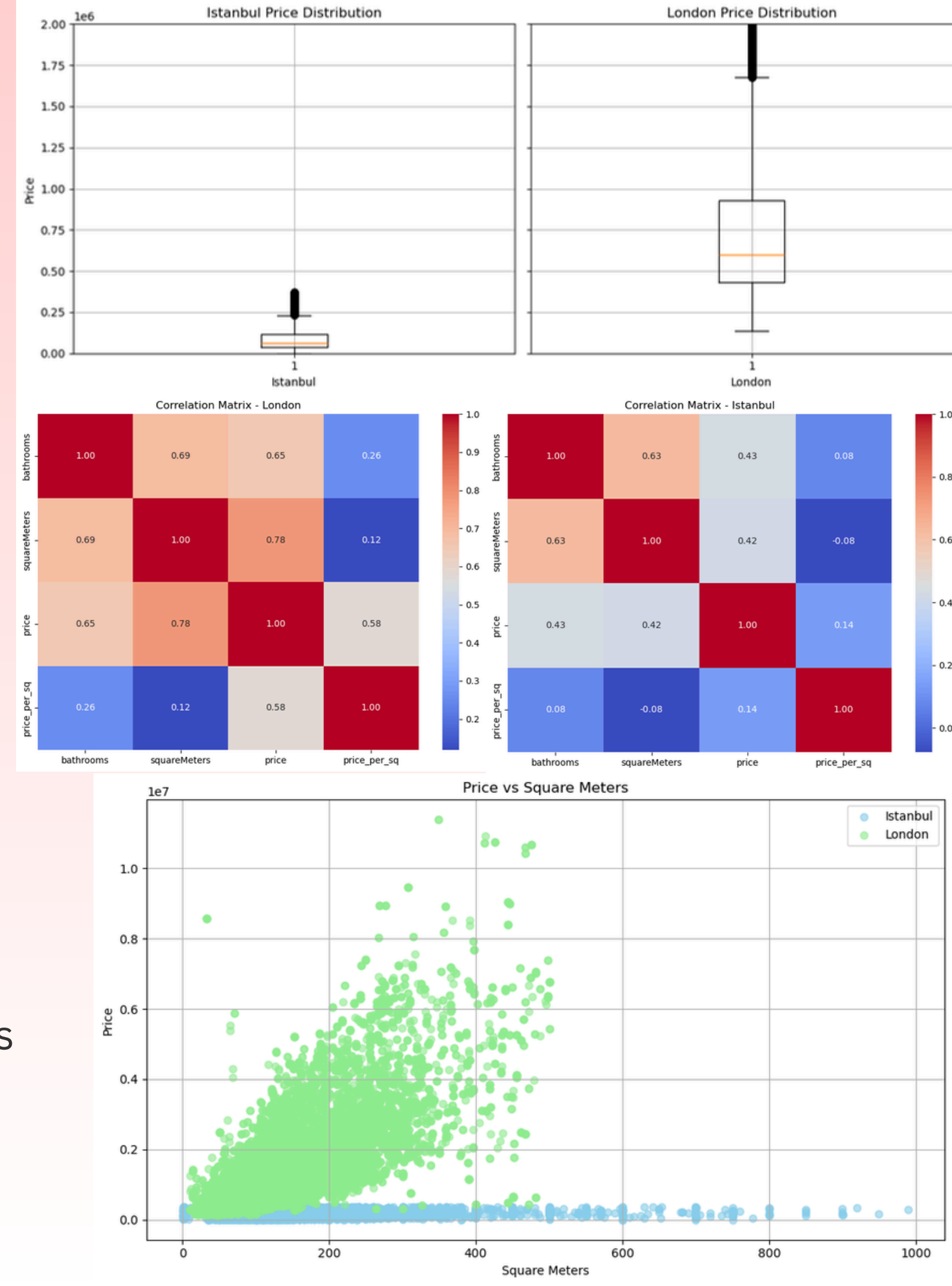
# Exploratory Data Analysis (EDA)

## Data Visualizations:

- Multiple plot types were used (histograms, boxplots, scatterplots)
- A correlation matrix was created to explore feature relationships

## Preprocessing :

- London house prices are significantly higher than Istanbul's (in GBP). And price is correlated with square meters(0.78) and bathrooms (0.69)
- Istanbul homes tend to be larger in square meters, but pricing is less predictable. This suggests location-based and unobserved factors may dominate in Istanbul



# My Hypothesis'

**T-statistic:** -139.741

**P-value:** 0.0000

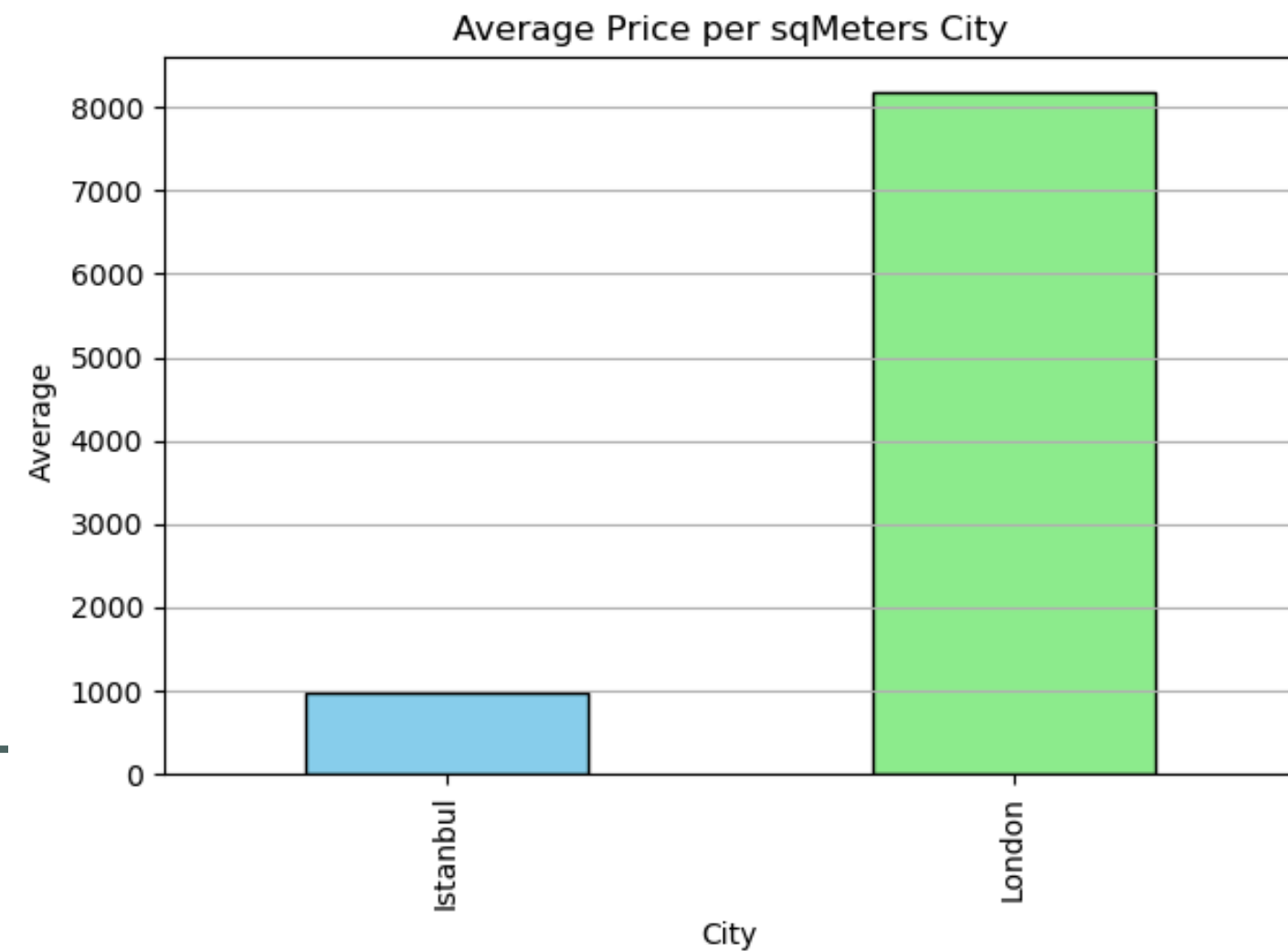
**H0 rejected:** The average price per square meter in Istanbul and London is statistically different .

## Null Hypothesis H0:

There is no statistically significant difference in the average price per square meter between Istanbul and London

## Alternative Hypothesis H1:

The average price per square meter in Istanbul and London is statistically different .



# My Hypothesis'

**T-statistic:** 67.223

**One-sided p-value:** 0.0000

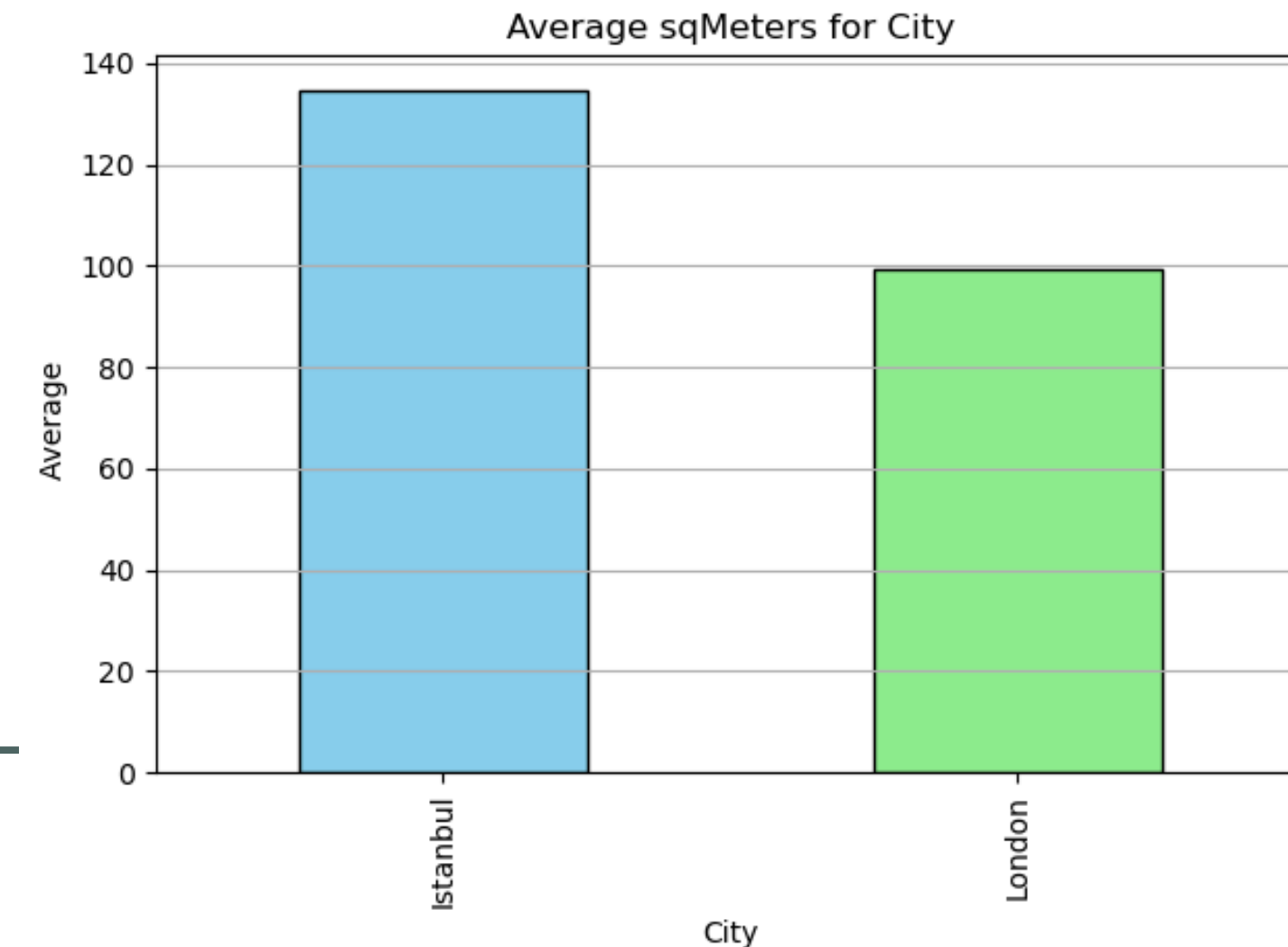
**H0 rejected:** Houses in Istanbul are significantly larger than London.

## Null Hypothesis H0:

We cannot conclude that houses in Istanbul are larger than London.

## Alternative Hypothesis H1:

Houses in Istanbul are significantly larger than London.



# My Hypothesis'

**Pearson Correlation Coefficient:** 0.777

**P-value:** 0.0000

**H0 rejected:** There is a statistically significant relationship between price and square meters in London.

## Null Hypothesis H0:

There is no statistically significant relationship between price and square meters in London.

## Alternative Hypothesis H1:

There is a statistically significant relationship between price and square meters in London.



# My Hypothesis'

**Pearson Correlation Coefficient:** 0.416

**P-value:** 0.0000

**H0 rejected:** There is a statistically significant relationship between price and square meters in Istanbul.

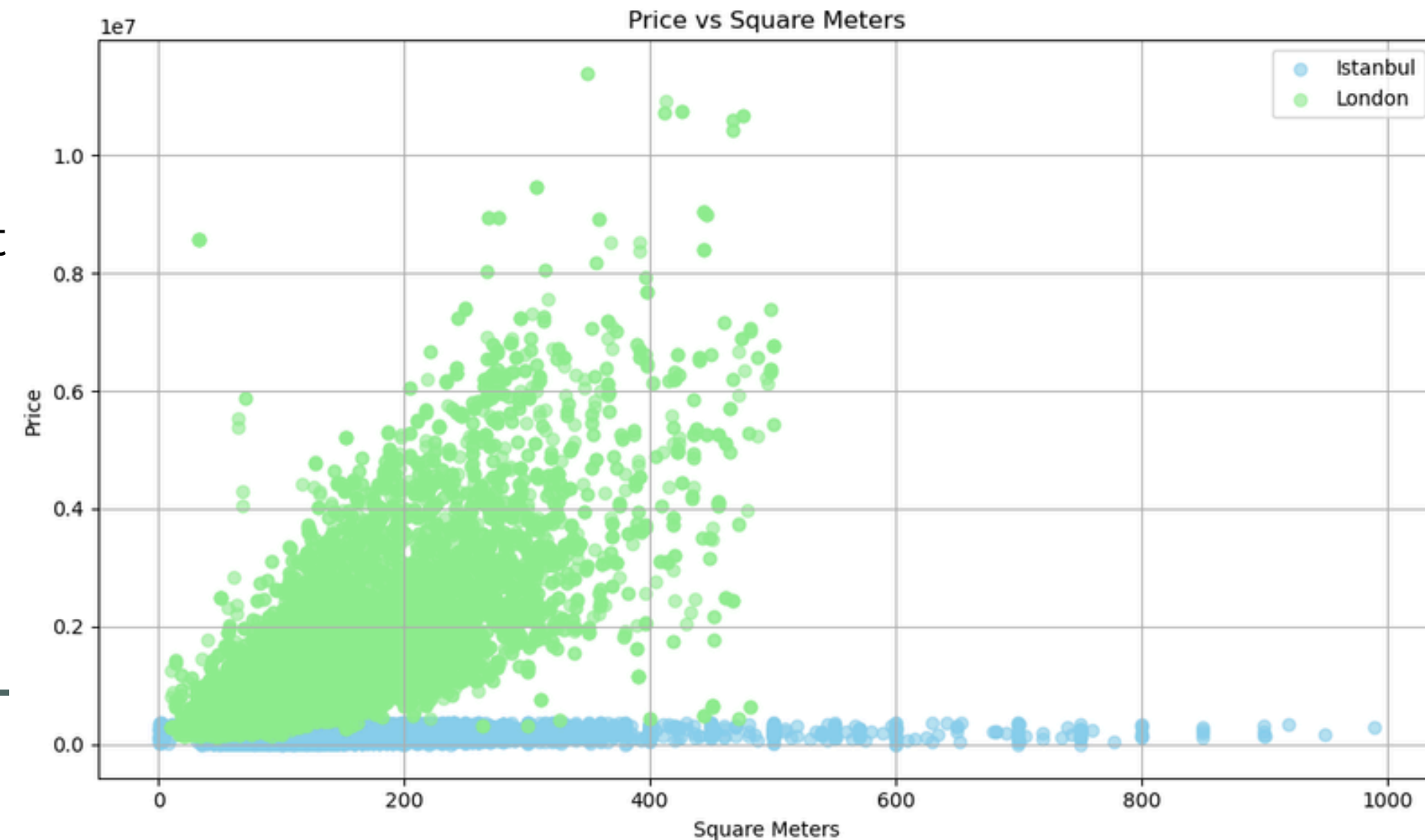
However with checking graph and low pearson core. coefficient we cannot sa there is a significant relationship.

## Null Hypothesis H0:

There is no statistically significant relationship between price and square meters in Istanbul.

## Alternative Hypothesis H1:

There is a statistically significant relationship between price and square meters in Istanbul.





# Machine Learning Part

---



## Goal:

To predict house prices in Istanbul and London using square meters and number of bathrooms.

## Models Used

- Linear Regression
- Random Forest Regressor
- XGBoost Regressor

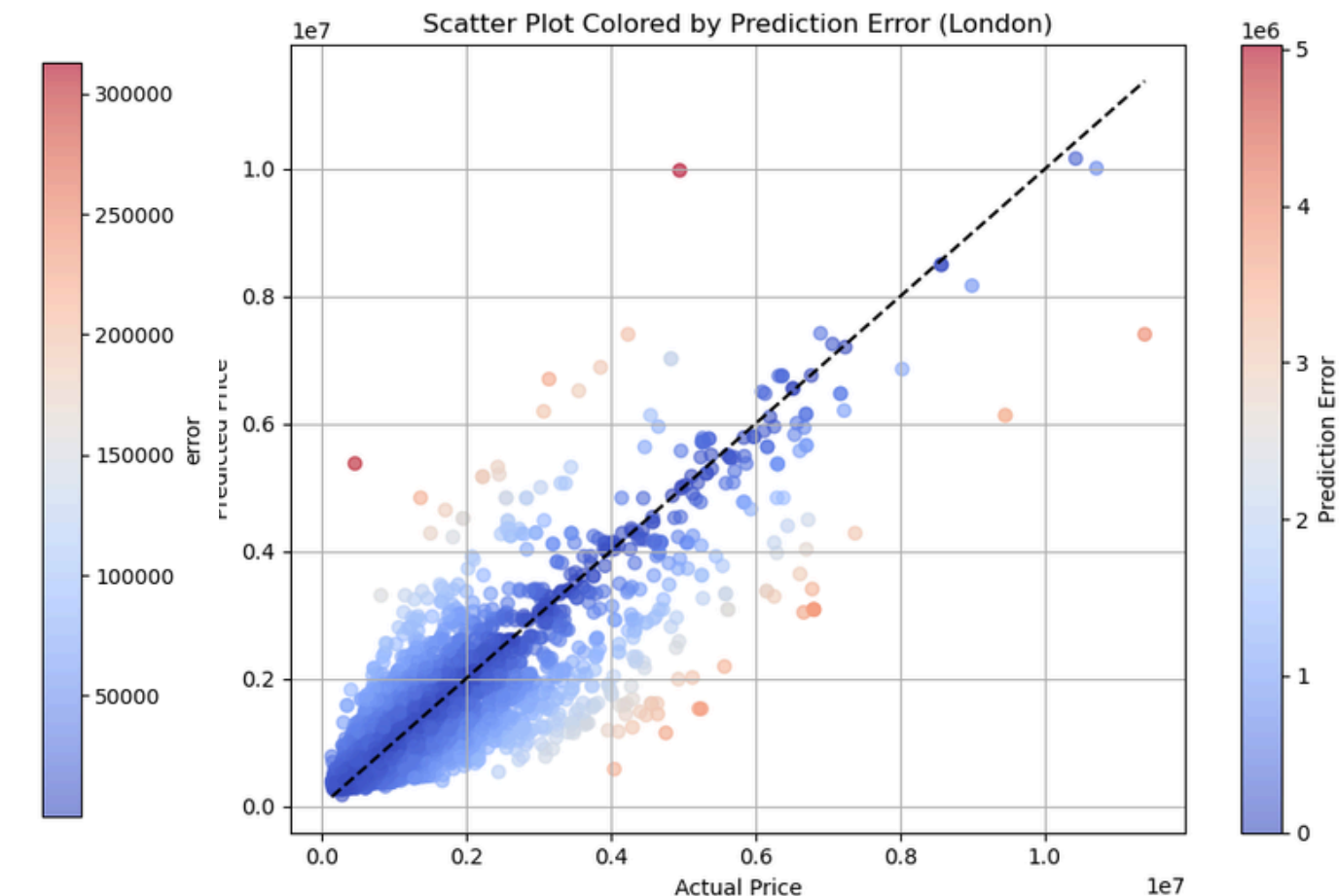
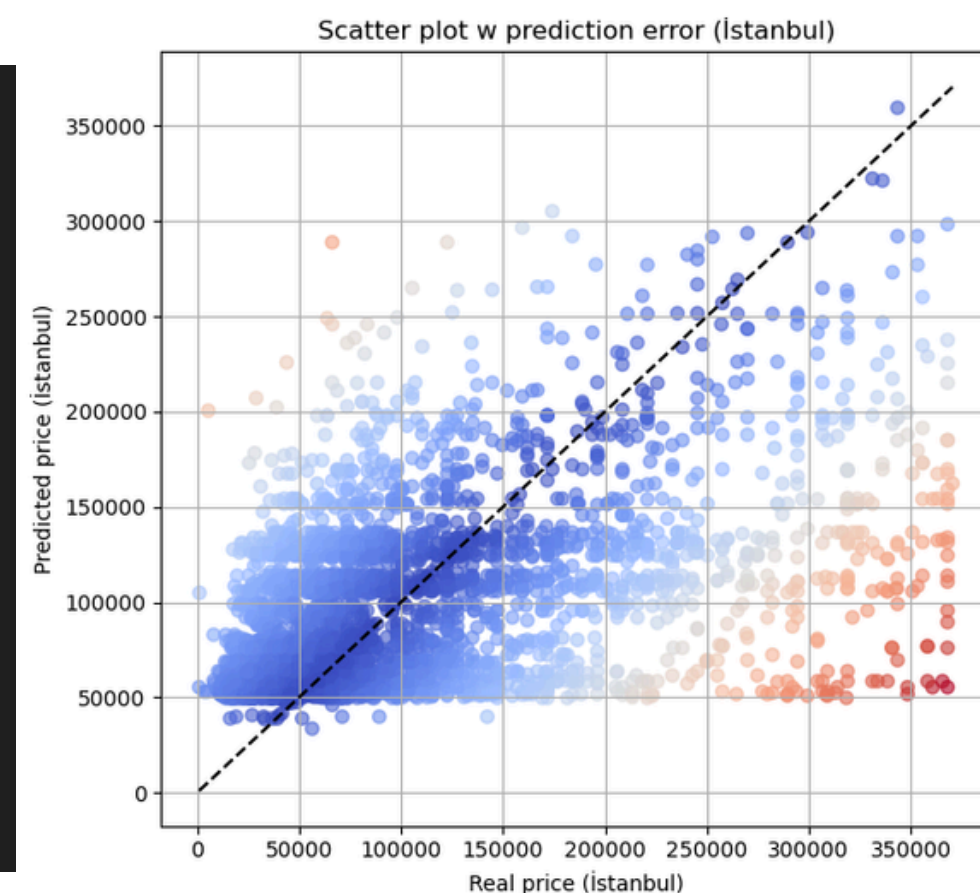
## Evaluation Metrics

- $R^2$  Score
- MAE (Mean Absolute Error)
- RMSE (Root Mean Squared Error)
- MAPE (Mean Absolute Percentage Error)

# Machine Learning Part



	City	Model	R2 Mean
0	Istanbul	Linear Regression	0.185134
1	Istanbul	Random Forest	0.196348
2	Istanbul	XGBoost	0.207405
3	London	Linear Regression	0.625061
4	London	Random Forest	0.723139
5	London	XGBoost	0.670211



- Istanbul Feature Importances: squareMeters: 0.8467 bathrooms: 0.1533
- London Feature Importances: squareMeters: 0.9121 bathrooms: 0.0879
- Models perform well in London due to strong correlation and cleaner data. Istanbul shows weak results, which may be resulted from location-based and unobserved factors

# Conclusion

---



## Key Takeaways:

- London housing prices are significantly higher than Istanbul's.
- Price per square meter is statistically different between the two cities ( $H_0$  rejected).
- Houses in Istanbul are larger, but have weaker correlation with price.
- Square meters is the most important feature in both cities.

## Machine Learning Summary:

- Models predict London house prices not bad ( $R^2 > 0.70$ ).
- Istanbul models perform poorly may due to missing location-specific data and market irregularities.
- Tree-based models (Random Forest, XGBoost) showed better performance than linear models.

A modern, two-story house with large glass windows and doors is visible in the background. In the foreground, there is a swimming pool with blue water. To the right of the pool, there is a lounge area with a large umbrella and some furniture. In the bottom foreground, there is a dark surface, possibly a patio, with several lit candles and some white fabric or pillows. The sky is a mix of blue and purple, suggesting dusk or dawn. The text "Thank you" is overlaid on a dark green rectangular background in the center of the image.

# Thank you