HOUSE PRICE PREDICTION

Athens, Greece

DIMITRIS FARANTOS
ANDREAS KATSAROS

Intro

We were tasked by real estate agency to produce a model for house price prediction

The reason for that is to have a general perception of the market and act as a metric for those who want to sell/buy houses in the area of Athens

Data Info

We pumped our data from Kaggle.

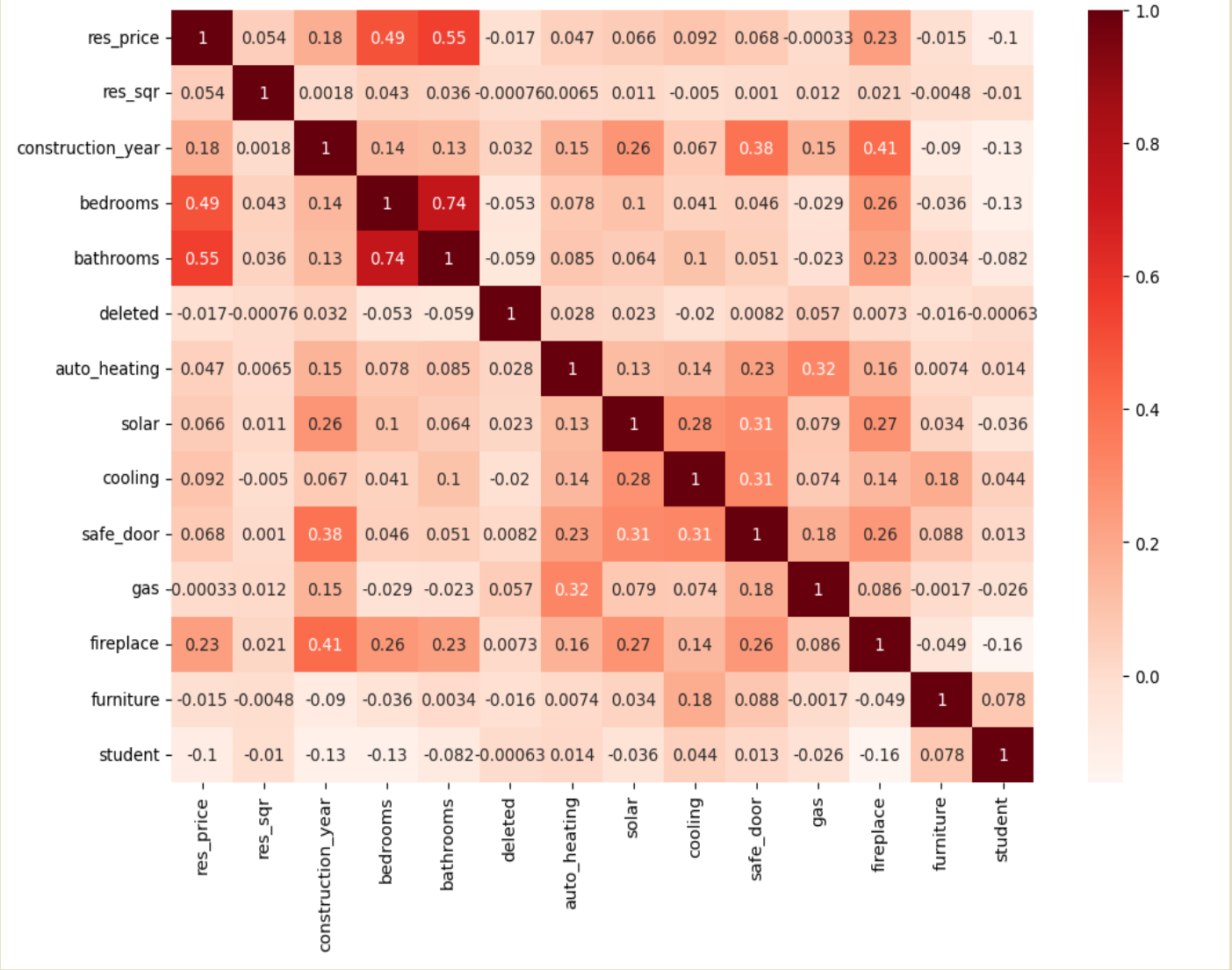
Includes house listings as publicly advertised in Greece.

This dataset is a shuffled 10% of the complete dataset of listings from 10/03/2022 to 14/06/2022.

Even though, the dataset, includes observations from both Athens and Thessaloniki areas, only the Athens ones are kept.

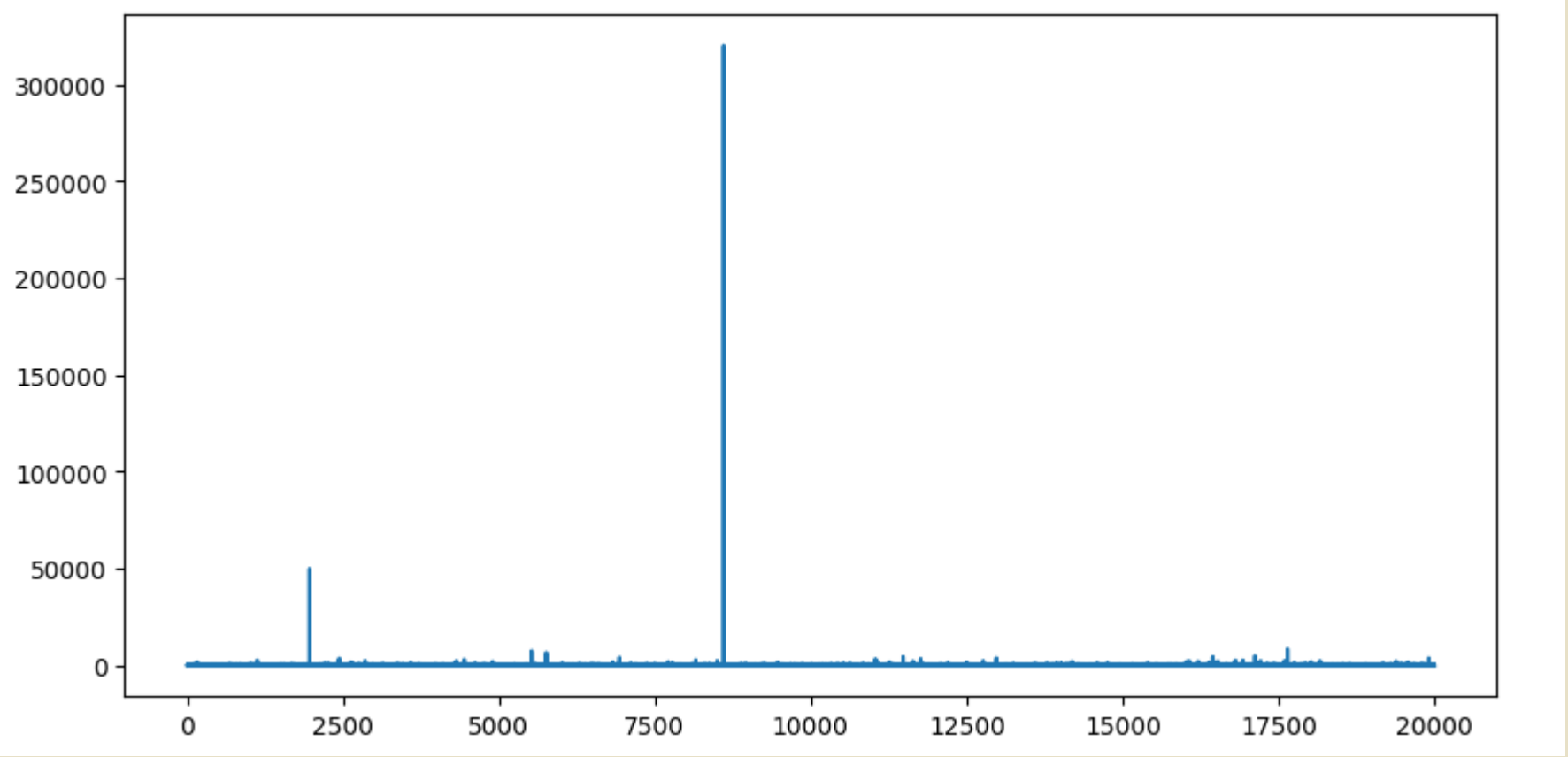
Dataset Graphs

Heatmap showing how variables affect each other



Dataset Graphs

Square Meters
Outliers

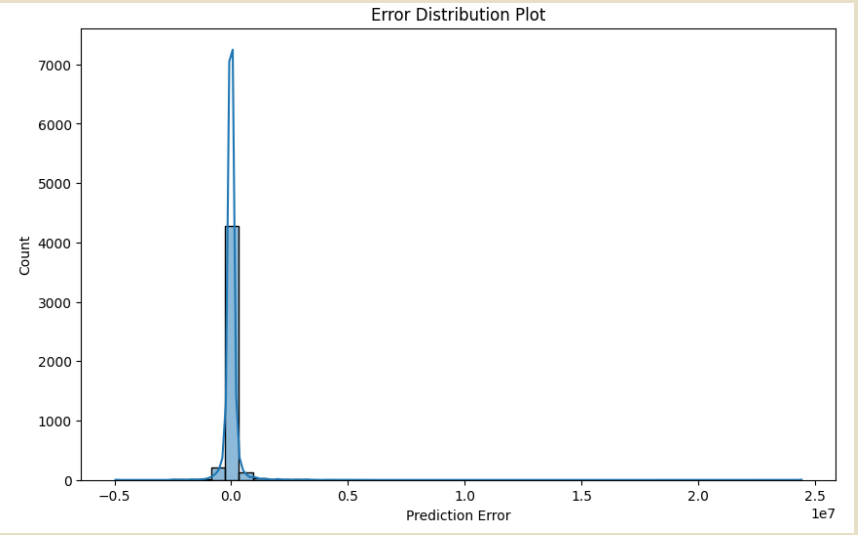


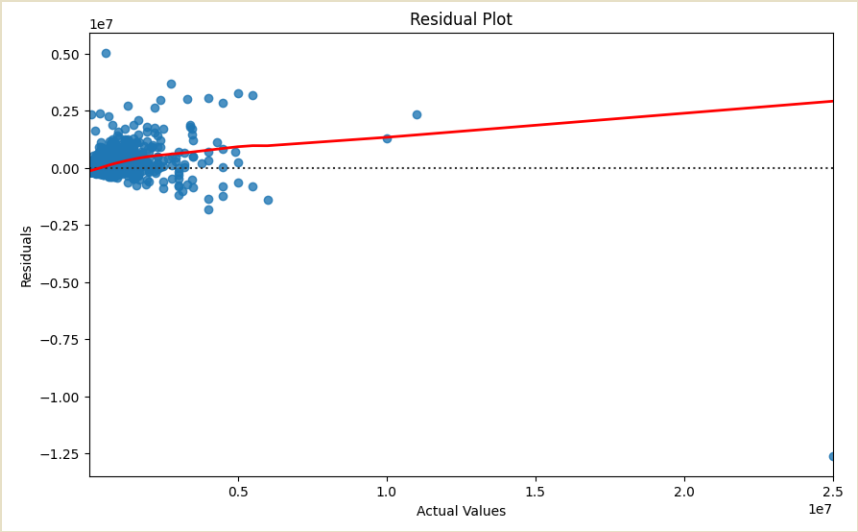
Method

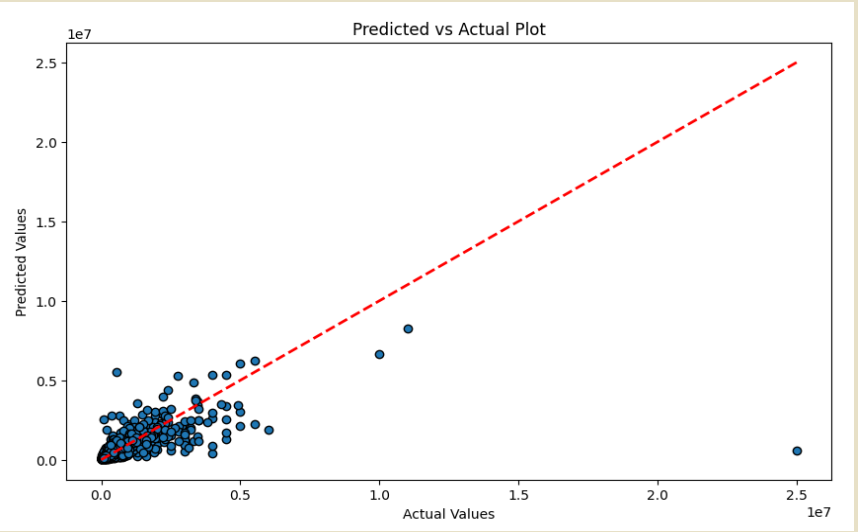
- One-Hot-Encoding on categorical data, which resulted in about 800 columns
- After that we used a KNN Imputer to handle some remaining missing values
- Then we used a standard scaler to scale our data
- And finally… we used a SelectKBest selector to find a the most valuable features
- The model we used was random forest… But the resulted metrics were not very valuable.

Results

R^2: 0.4940
MSE: 210148262844.4351
RMSE: 458419.3090
MAE: 117835.6899







Conclusion

As the results were not satisfying , the whole progress needs to be reevaluated !

Maybe with a different approach we could end up with better results