

Teruel City (TC)

Döşoğlu, Neşe*

#Summary

This paper presents a detailed analysis of house price prediction using a linear regression model based on the Teruel City dataset. The study investigates the impact of various factors on house prices, such as the number of bedrooms, bathrooms, and property area. The researchers have uncovered valuable insights into the housing market by examining these variables and applying the regression model. Traditionally, real estate valuation has relied on the expertise of professional appraisers, whose opinions may be influenced by various parties involved in a transaction. However, the emergence of automated prediction systems has revolutionized the industry by providing more objective and accurate property valuations. These systems leverage computer algorithms and machine learning techniques to analyze vast amounts of data, including historical sales and market trends, to generate reliable predictions. Such systems have become increasingly popular among stakeholders, including homebuyers, sellers, lenders, and real estate agents. The paper begins with a comprehensive literature review, highlighting previous studies that have employed machine learning models for house price prediction. These models have proven effective in analyzing large datasets and identifying patterns that may not be immediately apparent to humans. The reviewed studies encompass a range of methodologies, including deep learning models, regression algorithms, and neural network models, all of which have contributed to advancing the field of house price prediction. The Teruel City dataset, obtained from the Idealista website, is the foundation for this study. It comprises 58 observations and 14 variables that capture important attributes of houses in the Teruel city center, such as the type of house (apartment or individual), whether it is a duplex or contains an attic, location, number of bedrooms, property area, floor level, presence of a lift or garage, gas heating system, whether it is newly constructed or remodeled, presence of a box room, number of bathrooms, and the price in thousands of euros. To better understand the dataset, the researchers conduct exploratory data analysis. This includes visualizations, such as bar and scatter plots, to reveal patterns and relationships between variables. For instance, a bar plot showcases the prevalence of houses with and without a lift, providing insights into the significance of lift availability. Scatter plots are employed to examine the relationship between the number of bathrooms and house prices, which is positively correlated. The paper then delves into predictor variable analysis, focusing on the key factors influencing house prices. Multiple linear regression estimates the coefficients associated with each predictor variable. The results highlight the importance of the number of bathrooms, which significantly positively

*20080386, [Github Repo](#)

affects house prices. In contrast, the number of bedrooms harms costs, meaning that an increase in the number of bedrooms leads to decreased prices. The area of the property is also identified as a variable with a positive but relatively minor influence on house prices. To further validate the regression model, the researchers employ it to make predictions on new observations. By fitting the model and estimating the coefficients, accurate predictions can be generated, providing valuable insights for stakeholders in the real estate market. A line plot compares the predicted prices with the actual costs of houses in the Teruel City dataset, demonstrating the model's accuracy in estimating house prices. In conclusion, this paper contributes to house price prediction by showcasing the significance of specific variables and the potential of machine learning techniques in enhancing property valuation efficiency. The findings underscore the importance of considering attributes such as the number of bedrooms, bathrooms, and property area when predicting house prices. Automated prediction systems offer objective and accurate valuations, empowering stakeholders to make informed decisions. Overall, this study highlights the value of data analysis and machine learning in the real estate industry and provides a foundation for further research and advancements in house price prediction.