Housing Size and Home Values

Introduction

Housing prices affect buyers, sellers, and renters. One basic feature of a home is its size, which is commonly measured by the number of rooms. This study asks the following question: Does the average number of rooms predict the median home value?

Data methods

To answer this question, I used the Boston Housing Dataset, originally collected by Harrison and Rubinfeld (1978). The dataset contains information on 506 residential properties. After filtering unrealistic values, including homes with median values below $5,000 and homes with fewer than three rooms, the final sample consists of 504 observations. The variables used are RM, the average number of rooms, and MEDV, the median home value measured in thousands of dollars.

The distribution of the average number of rooms is roughly normally distributed. Median home values vary substantially across observations. Homes with more rooms tend to have higher median values.

Figure 1 shows the relationship between home size and median home value.

Figure 2 shows the distribution of the average number of rooms.

Statistical methods

To answer the main research question, I used a simple linear regression model to predict median home value using the average number of rooms.

$$\mathbf{MEDV} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1}(\mathbf{RM}) + \boldsymbol{\epsilon}$$

Figure 3 shows the residuals from the regression model.

Results

The linear regression produced the following results. The slope coefficient is 9.0393 and is statistically significant ($p < 0.001$). This model finds that each additional room is

associated with an increase of approximately $9,039 in median home value. The R-squared value is 0.482, meaning that the number of rooms explains about 48 percent of the variation in home prices.

Conclusion

Homes with more rooms tend to have higher median values. Using data from the Boston Housing Dataset and a simple regression model, I find a strong positive relationship between housing size and home value. While this model does not capture all factors that affect housing prices, it shows that size is an important determinant of value.

Figures:

Figure:1

Figure:2

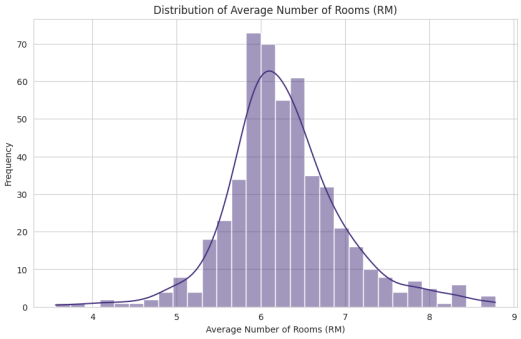
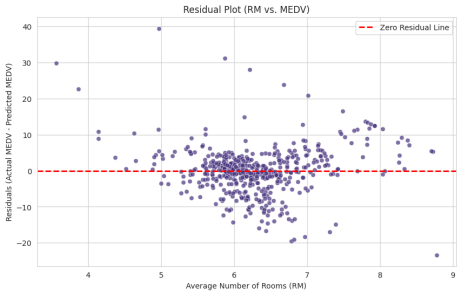


Figure:3



Sources

- **Data Source: Boston Housing Dataset** (Harrison & Rubinfeld, 1978).
- **Statistical Software:** Python (via Anaconda distribution) with the pandas, scikit-learn, and statsmodelslibraries.

- **AI Assistance:** Google Gemini (used for code generation, editing, and structure).