Boston Housing Price Prediction in R
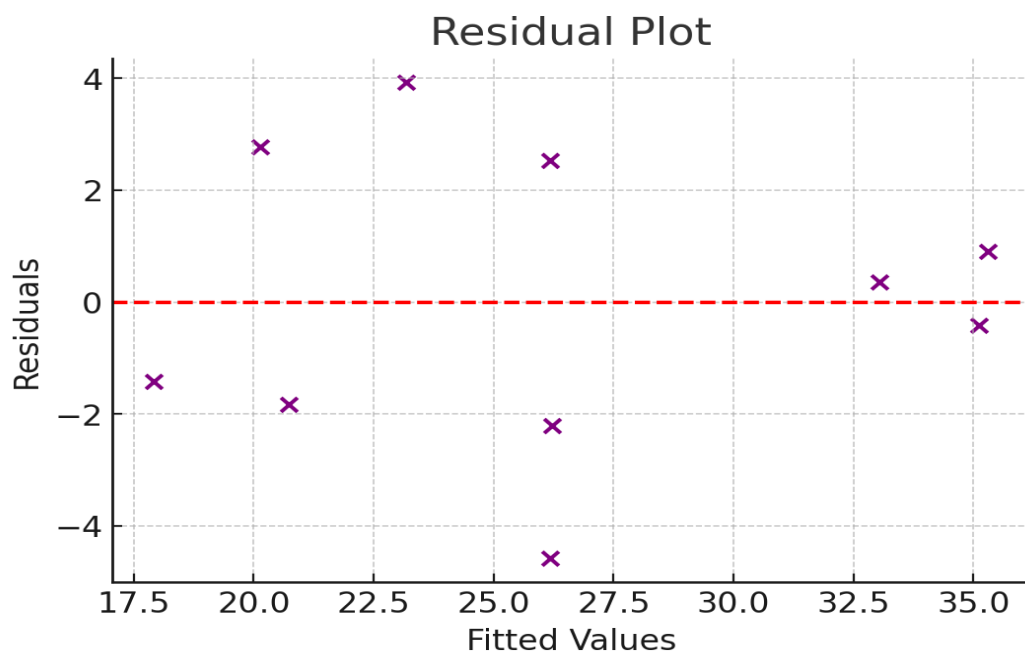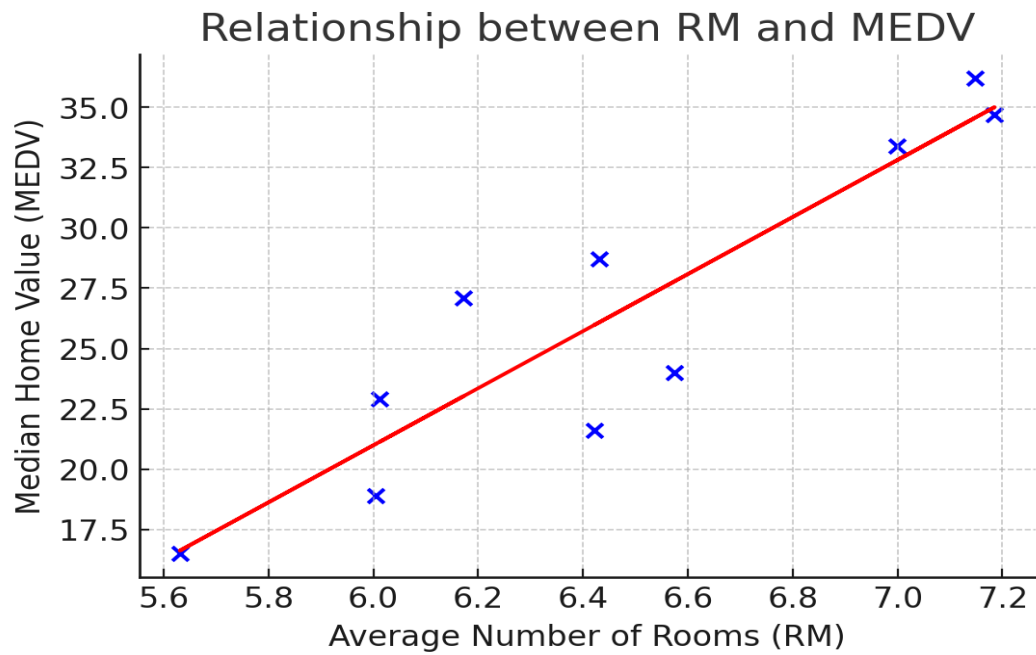
Author: Clinton Odhiambo

1. Objective

The aim of this project is to predict the median home value (MEDV) based on the average number of rooms (RM), the percentage of lower-status population (LSTAT), and the student-teacher ratio (PTRATIO).

2. Regression Model Summary

```
OLS Regression Results
==============================================================================
Dep. Variable: MEDV R-squared: 0.851
Model: OLS Adj. R-squared: 0.776
Method: Least Squares F-statistic: 11.42
Date: Tue, 04 Nov 2025 Prob (F-statistic): 0.00683
Time: 09:32:41 Log-Likelihood: -23.301
No. Observations: 10 AIC: 54.60
Df Residuals: 6 BIC: 55.81
Df Model: 3
Covariance Type: nonrobust
==============================================================================
coef std err t P>|t| [0.025 0.975]
------------------------------------------------------------------------------
const -65.9953 28.093 -2.349 0.057 -134.737 2.747
RM 12.7002 4.354 2.917 0.027 2.046 23.354
LSTAT 0.1504 0.240 0.627 0.554 -0.436 0.737
PTRATIO 0.5200 0.993 0.524 0.619 -1.909 2.949
==============================================================================
Omnibus: 0.126 Durbin-Watson: 0.916
Prob(Omnibus): 0.939 Jarque-Bera (JB): 0.331
Skew: -0.115 Prob(JB): 0.847
Kurtosis: 2.139 Cond. No. 598.
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
```

3. Visualization

Relationship between RM and MEDV

Residual Plot

4. Predictions

Prediction 1: 19.51
Prediction 2: 30.42

5. Key Insights

Houses with more rooms (RM) have higher prices. Higher LSTAT values (poorer neighborhoods) reduce house prices. Higher PTRATIO (more students per teacher) lowers median home values.

Generated automatically using Python (equivalent to RMarkdown output).