📊 Simple vs Multiple Linear Regression – Comparison

This document compares Simple Linear Regression and Multiple Linear Regression using the California Housing dataset.  
The goal is to predict house prices, first using only one feature (Median Income), and then using all available features.

# 🔍 Side-by-Side Metrics Comparison

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Metric | Simple Linear Regression | Multiple Linear Regression |  |  |
| Features Used | 1 (MedInc) | 8 (All features) |  |  |
| Coefficient | 0.4193 | Multiple coefficients |  |  |
| Intercept | 0.4446 | Intercept from full model |  |  |
| MSE | 0.7091 | 0.5559 |  |  |
| RMSE | 0.8418 | 0.7456 |  |  |
| MAE | - | 0.5332 |  |  |
| R² Score | 0.4589 | 0.5758 |  |  |

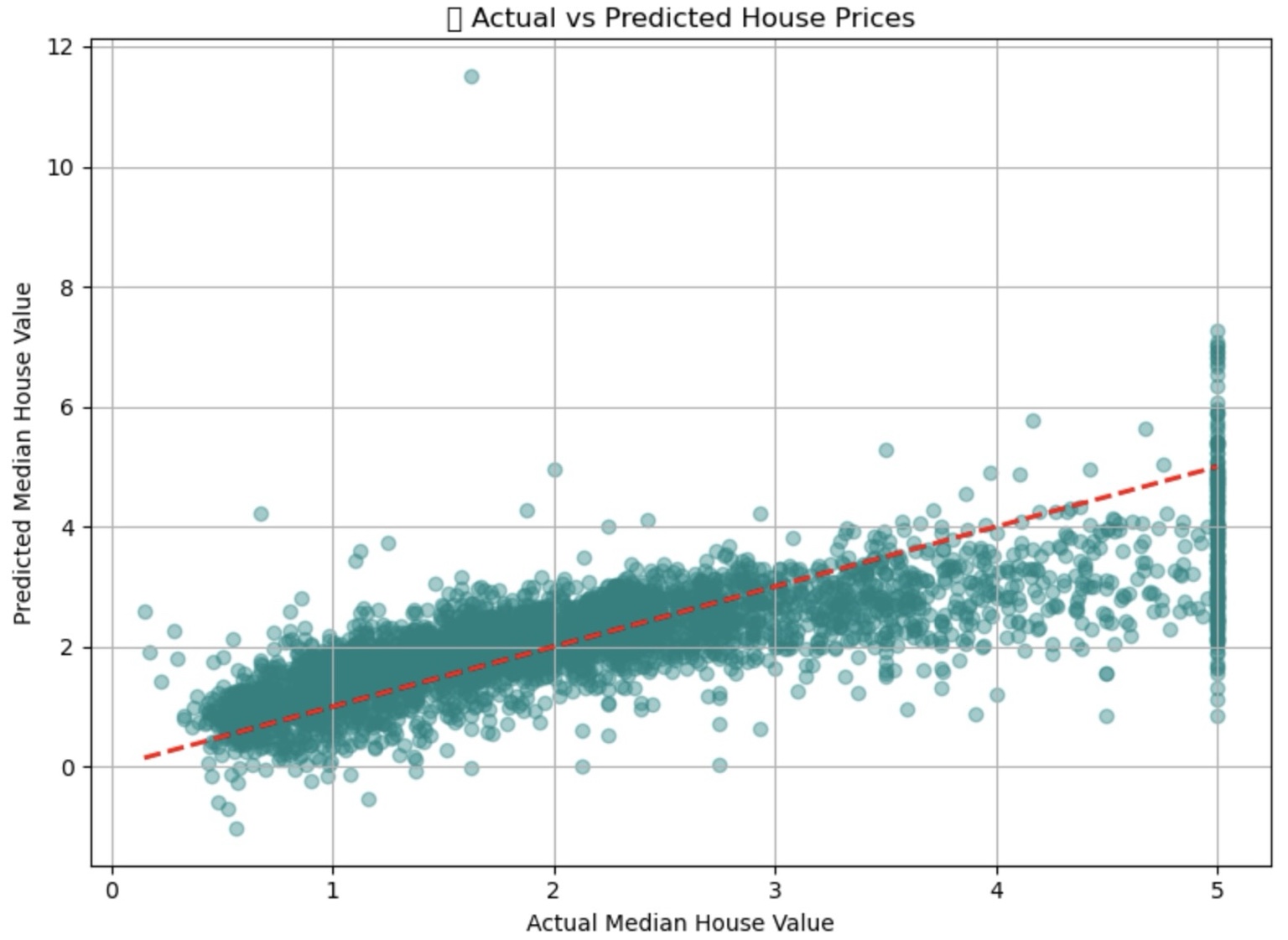
From the comparison, we see that Multiple Linear Regression:  
- Has a lower error (MSE, RMSE) than Simple Linear Regression  
- Explains more of the variance (R² = 0.5758 vs 0.4589)  
- Is more reliable due to having more features involved

# 📉 Visual Comparison

## Simple Linear Regression: Income vs House Price



## Multiple Linear Regression: Actual vs Predicted House Prices



# 📌 Final Insight

Simple Linear Regression gives us a quick way to understand one relationship, but is often too limited for accurate predictions.  
Multiple Linear Regression is more powerful and performs better when we have a variety of relevant features, as seen in the California Housing example.