📈 Simple Linear Regression – Income vs House Price

This simple linear regression model predicts house prices using only one feature: Median Income.  
The data used comes from the California Housing dataset. We fit a straight line to model how income affects house prices.

# 📊 Model Performance

- 📈 Coefficient: 0.4193   
 Every $10,000 increase in income is associated with a $41,930 increase in predicted house value.  
  
- 📉 Intercept: 0.4446   
 When income is 0, the predicted house value is approximately $44,460.  
  
- 📉 Mean Squared Error (MSE): 0.7091   
 Represents the average squared difference between predicted and actual prices.  
  
- 📈 R² Score: 0.4589   
 The model explains about 46% of the variance in house prices — indicating a moderate linear relationship.

# 📉 Visual Interpretation

The scatter plot shows actual house prices versus median income with a red prediction line fitted by the model.

## ✅ Positive Correlation

The red line slopes upward, showing that house prices tend to increase as median income increases. The coefficient quantifies this relationship.

## ⚠️ High Variability Around the Line

Although there's a trend, actual prices vary significantly at each income level. This shows that income alone is not a strong predictor of house value.

## ❌ House Price Clipping

There’s a visible ceiling at $500,000 (value = 5), where many actual values are capped. This limits the model’s accuracy at higher price ranges.

# 📌 Conclusion

This simple regression model shows a clear but limited relationship between income and house price.  
To improve predictions, multiple features should be used — as seen earlier in our Multiple Linear Regression model.



Figure: Simple Linear Regression plot – Predicting House Price from Median Income.