

# Multiple Linear Regression

Supervised Learning

Daniel E. Acuna

Associate Professor, University of Colorado Boulder



# Contents of This Video

In this video, we will cover:

- Extension from simple to multiple linear regression
- The multiple regression equation and interpretation
- Key concept: interpreting coefficients “holding others fixed”
- Benefits of multiple predictors for prediction accuracy
- Evaluating feature importance in ML context
- Real-world applications and model selection



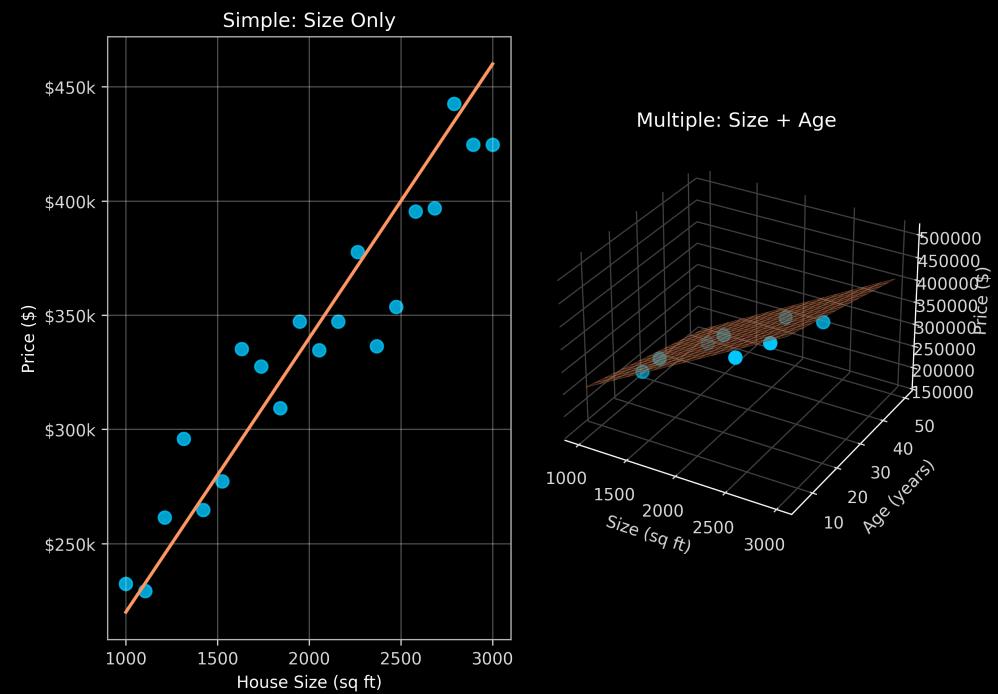
# From Simple to Multiple Predictors

## Simple Linear Regression

- One predictor variable
- $$Y = \beta_0 + \beta_1 X + \varepsilon$$
- Example: House price from size only

## Multiple Linear Regression

- Multiple predictor variables
- $$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon$$
- Example: House price from size, location, age



# The Multiple Regression Equation

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon$$

## Components:

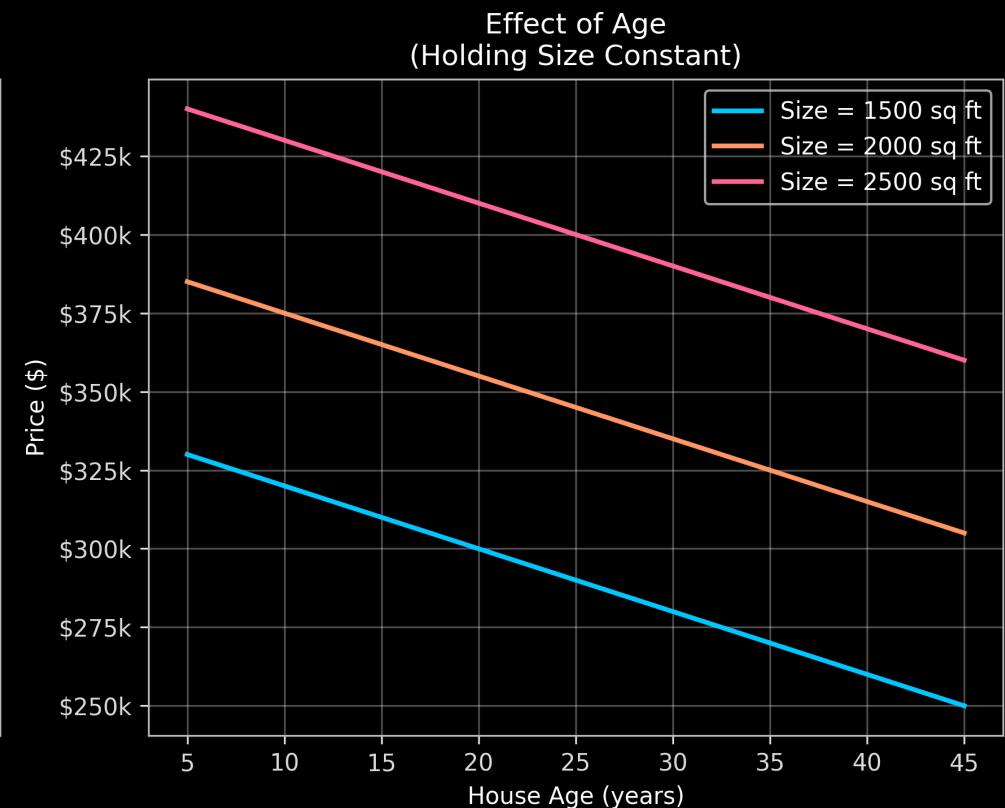
- $Y$ : Target variable (outcome we want to predict)
- $X_1, X_2, \dots, X_p$ : Predictor variables (features)
- $\beta_0$ : Intercept,  $\beta_1, \beta_2, \dots, \beta_p$ : Coefficients for each predictor
- $\varepsilon$ : Error term



# Key Concept: “Holding Others Fixed”

**Critical interpretation principle:** Each coefficient represents the effect of its variable **holding all other variables constant**

$\beta_1$  = Change in Y for 1-unit increase in  $X_1$  when  $X_2, X_3, \dots$  are fixed



# Example: House Price Prediction Model

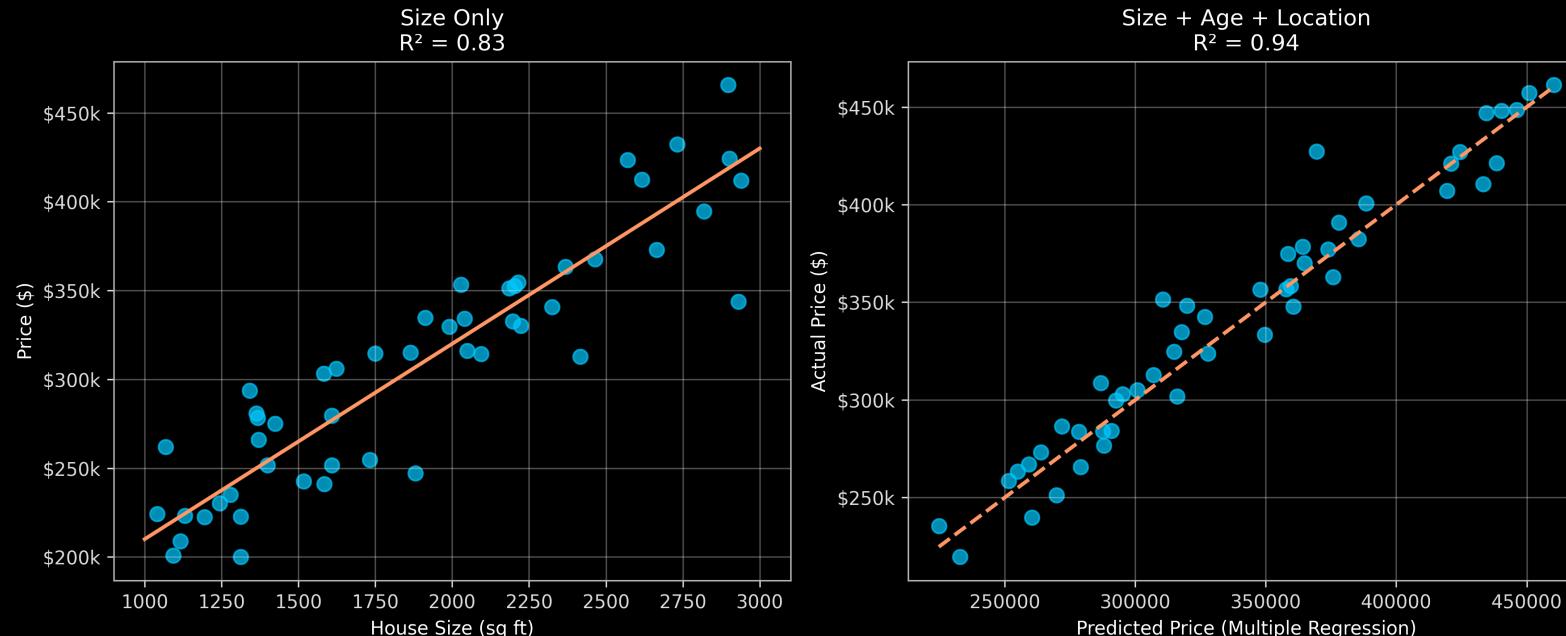
$$\text{Price} = \beta_0 + \beta_1(\text{Size}) + \beta_2(\text{Age}) + \beta_3(\text{Location}) + \varepsilon$$

**Coefficient interpretation (“holding others fixed”):**

- $\beta_1 = \$110$ : Each sq ft  $\rightarrow +\$110$  price (same age/location)
- $\beta_2 = -\$2000$ : Each year older  $\rightarrow -\$2000$  price (same size/location)
- $\beta_3 = \$50000$ : Premium location  $\rightarrow +\$50000$  (same size/age)



# Why Use Multiple Predictors? Improved Prediction



**Including more relevant features usually improves predictive accuracy**

- Single predictor may miss important relationships
- Multiple predictors can explain more variation in Y
- Example: House size alone vs. Size + Age + Location

# Why Use Multiple Predictors? Unique Contributions

Multiple regression helps understand each predictor's *unique contribution*

**Problem with single predictors:**

- Predictors can be correlated with each other
- Simple models might confuse their effects

**Solution with multiple regression:**

- Controls for other predictors
- Isolates each variable's true effect

**Example:** House price model with size and bedrooms

- Both are correlated (bigger houses → more bedrooms)
- Multiple regression separates their individual effects



# Evaluating Feature Importance

**How do we know which predictors are most useful?**

**Coefficient magnitude:** Larger absolute values suggest stronger effects

**Performance metrics:**

- R-squared improvement when adding each feature
- Cross-validation performance with/without each predictor
- Feature ablation: remove feature and measure performance drop



# Fitting Multiple Regression Models

**Same method as simple regression: Least Squares**

- Minimize sum of squared residuals:  $\sum_{i=1}^n (y_i - \hat{y}_i)^2$
- Find optimal values for  $\beta_0, \beta_1, \dots, \beta_p$
- Software handles computation automatically

**Output:** Coefficients, R-squared, predictions, residuals



# Real-World Examples

**Multiple regression is widely used across domains:**

- **Real Estate:** House prices using size, age, location
- **Healthcare:** Patient outcomes using treatment, dosage, age
- **Marketing:** Sales using advertising channels, seasonality, pricing
- **Transportation:** Fuel efficiency using engine size, weight
- **Environment:** Temperature using location, altitude, time

**Key insight:** Most outcomes depend on multiple factors



# Caution: Model Selection and Overfitting

## Adding more predictors:

-  Generally improves fit on training data ( $R^2$  increases)
-  Risk of overfitting with too many predictors

## Guidelines for model selection:

- Use domain knowledge to select meaningful predictors
- Validate performance on unseen data
- Balance complexity with interpretability



# What We've Covered

In this video, we've learned:

- Extension from simple to multiple linear regression
- Key interpretation: coefficients “holding others fixed”
- Benefits: improved prediction accuracy and unique contributions
- ML approaches to evaluating feature importance
- Real-world applications and model selection considerations

**Multiple regression: powerful framework for complex relationships**

