

Statistics and Data Analysis

Unit 04 – Lecture 02: Simple Linear Regression (OLS)

Tofik Ali

School of Computer Science, UPES Dehradun

February 14, 2026

<https://github.com/tali7c/Statistics-and-Data-Analysis>

Overview
o

Model
oo

Fit and Diagnostics
o

Exercises
oooooo

Demo
oo

Summary
oo

Quick Links

Overview

Model

Fit and Diagnostics

Exercises

Demo

Summary

Agenda

1 Overview

2 Model

3 Fit and Diagnostics

4 Exercises

5 Demo

6 Summary

Learning Outcomes

- Write the simple linear regression model

Learning Outcomes

- Write the simple linear regression model
- Interpret slope and intercept in context

Learning Outcomes

- Write the simple linear regression model
- Interpret slope and intercept in context
- Compute a prediction and a residual

Learning Outcomes

- Write the simple linear regression model
- Interpret slope and intercept in context
- Compute a prediction and a residual
- Explain R-squared (intuition)

Model: Key Points

- $y = b_0 + b_1 x + \text{error}$

Model: Key Points

- $y = b_0 + b_1 x + \text{error}$
- Slope: expected change in y for 1-unit increase in x

Model: Key Points

- $y = b_0 + b_1 x + \text{error}$
- Slope: expected change in y for 1-unit increase in x
- Intercept: predicted y at $x=0$ (interpret carefully)

Model: Key Formula

$$y = \beta_0 + \beta_1 x + \epsilon$$

Fit and Diagnostics: Key Points

- Look at residual plots for patterns

Fit and Diagnostics: Key Points

- Look at residual plots for patterns
- Outliers can dominate the fitted line

Fit and Diagnostics: Key Points

- Look at residual plots for patterns
- Outliers can dominate the fitted line
- High R^2 does not guarantee a good model

Exercise 1: Prediction

Model: $y\hat{=}10+2x$. Predict y when $x=7$.

Solution 1

- $\hat{y} = 24$

Exercise 2: Residual

If actual $y=20$ at $x=7$, compute residual.

Solution 2

- $20 - 24 = -4$

Exercise 3: Interpret slope

Slope is 5 thousand INR per extra room. Interpret.

Solution 3

- Each extra room increases predicted price by 5k INR (on average).

Mini Demo (Python)

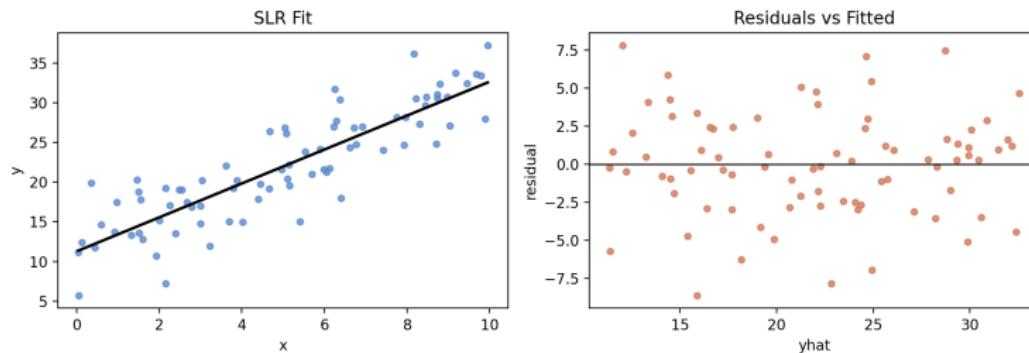
Run from the lecture folder:

```
python demo/demo.py
```

Outputs:

- images/demo.png
- data/results.txt

Demo Output (Example)



Summary

- Key definitions and the main formula.

Summary

- Key definitions and the main formula.
- How to interpret results in context.

Summary

- Key definitions and the main formula.
- How to interpret results in context.
- How the demo connects to the theory.

Exit Question

Why do we check residual plots even if R^2 is high?