Statistics and Data Analysis

Unit 04 – Lecture 01: Correlation and Regression: Concepts

Tofik Ali

School of Computer Science, UPES Dehradun

February 17, 2026

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

Quick Links

Overview Concepts Causation Warning Exercises Demo Summary

Agenda

Learning Outcomes

- Differentiate correlation and regression

Learning Outcomes

- Differentiate correlation and regression
- Explain why correlation does not imply causation

Learning Outcomes

- Differentiate correlation and regression
- Explain why correlation does not imply causation
- Interpret a scatter plot (trend, outliers)

Learning Outcomes

- Differentiate correlation and regression
- Explain why correlation does not imply causation
- Interpret a scatter plot (trend, outliers)
- Define residual and why residuals matter

Concepts: Key Points

- Correlation measures linear association

Concepts: Key Points

- Correlation measures linear association
- Regression models Y as a function of X

Concepts: Key Points

- Correlation measures linear association
- Regression models Y as a function of X
- Regression has roles: predictors vs response

Causation Warning: Key Points

- Confounding can create misleading correlation

Causation Warning: Key Points

- Confounding can create misleading correlation
- Reverse causality is possible

Causation Warning: Key Points

- Confounding can create misleading correlation
- Reverse causality is possible
- Causal claims need design or strong assumptions

Exercise 1: Pick response variable

Predict house price using size and location. What is the response variable?

Solution 1

- House price is the response (Y).

Exercise 2: Interpret r

If r=0.7 between study hours and score, what does it mean?

Solution 2

- Strong positive linear association.
- Not proof of causation.

Exercise 3: Residual sign

If y=74 and yhat=80, what is residual?

Solution 3

- Residual = y - yhat = -6 (over-prediction).

Mini Demo (Python)

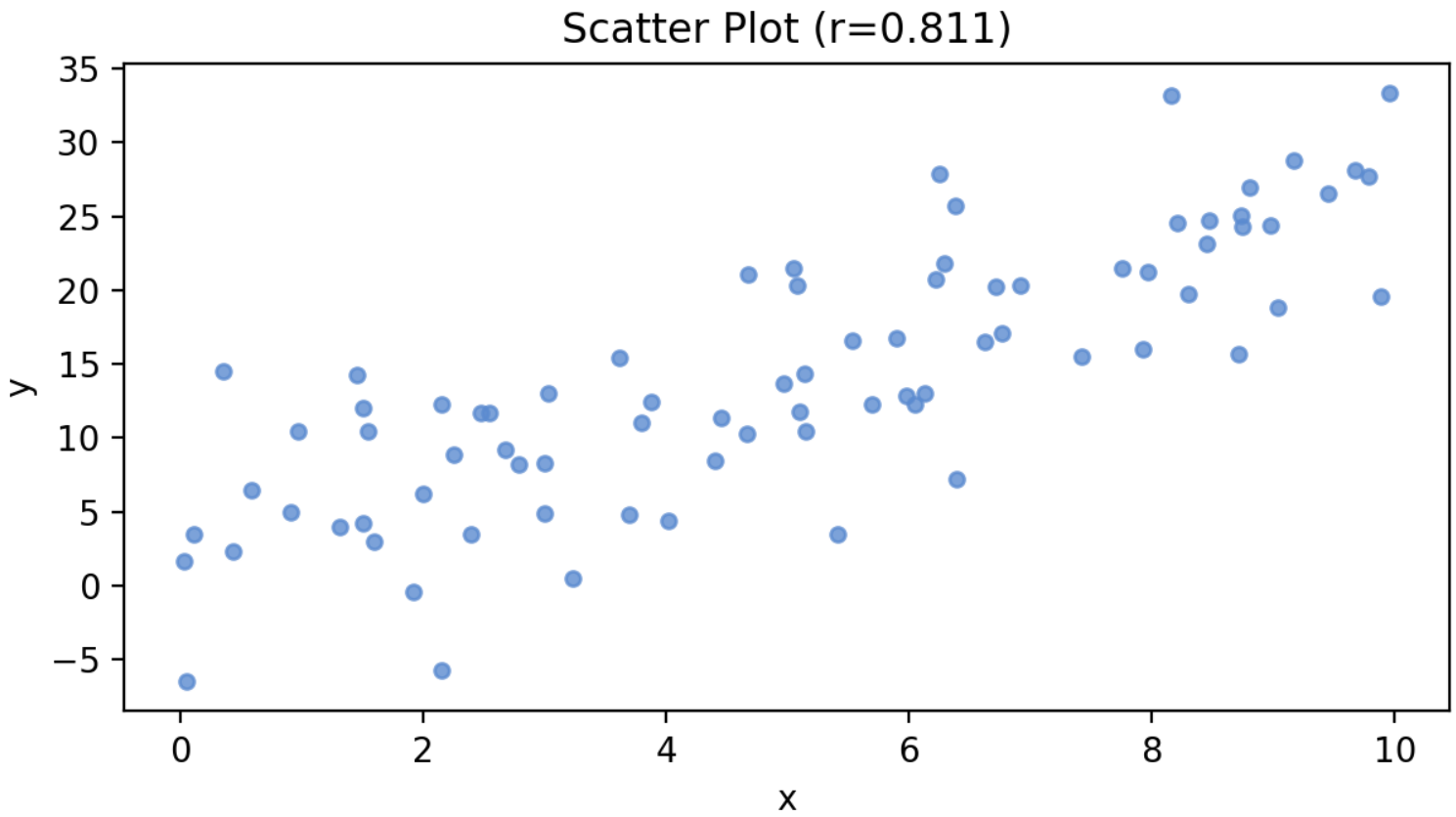
Run from the lecture folder:

```
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

Give one example of a confounder that can create a misleading correlation.