Introduction to the Course

Data Analysis for Public Policy

Sergio I. García-Ríos

Fall 2026

PA 311N: Undergraduate Quantitative Methods
LBJ School of Public Affairs

General information

Teaching team and weekly rhythm

- Professor: Sergio I. García-Ríos
- Teaching assistant: Sarah Traore
- Monday: lecture and statistical reasoning
- Wednesday: guided lab in R
- Friday: review, troubleshooting, and **On Your Own** support

Office hours: Sergio by appointment; Sarah Wednesday 12:30–2:00 and Friday 11:00–12:30.

Required materials

- *OpenIntro Statistics*, 4th edition: `https://www.openintro.org/book/os/`
  - Freely available online; an inexpensive print edition is optional.
- R, RStudio Desktop, and Quarto
- A laptop for guided labs
- Optional calculator

Course website

Undergraduate Quantitative Methods

Data Analysis for Public Policy

`https://garciarios.github.io/lbj_ug_quant/`

Course structure

Learning sequence

1. Data, measurement, and exploratory analysis
2. Probability and distributions
3. Sampling, uncertainty, and foundations for inference
4. Hypothesis testing and comparisons
5. Correlation and regression
6. Communicating quantitative findings

How the course works

Every topic connects three elements:

1. Statistical reasoning: What question can the method answer?
2. R programming: How do we carry out and check the analysis?
3. Public-policy application: What does the result mean in practice?

Methods begin with substantive questions and real data, not disconnected formulas.

Assessment

Component	Weight
Participation and in-class work	5%
Problem sets	25%
Guided labs	25%
Midterm	10%
Project proposal	5%
Individual reproducible report	20%
Policy poster and presentation	10%

Labs and problem sets

- Labs are substantial and cumulative; most coding is completed together.
- Each lab ends with a shorter independent **On Your Own** section.
- Problem sets use selected exercises from *OpenIntro Statistics*.
- Coding complexity is secondary to reasoning, interpretation, and communication.

Individual applied project

Objective: Develop an answerable policy question and use real data to answer it.

- Short project proposal
- Reproducible individual analysis in Quarto
- Policy poster and presentation during the final class meeting

The project must demonstrate your own coding and interpretation.

Midterm

There will be one midterm on Monday, November 2.

- Covers Units 1–4
- Focuses on reasoning, interpretation, and application
- One student-prepared double-sided reference sheet is permitted

Responsible work with data and AI

- Document your data sources and analytical decisions.
- Protect privacy and avoid overstating what the evidence can show.
- AI may support learning only within assignment-specific rules.
- You remain responsible for checking code, results, citations, and interpretations.
- If you are unsure whether something is allowed, ask before submitting.

Tips for success

Tips for success

- Complete readings before the relevant unit and revisit them afterward.
- Participate actively in lectures and labs.
- Ask questions early—in class, during office hours, or by email.
- Start problem sets and the project early.
- Do not let one unit pass with unanswered questions; the course is cumulative.

Before Wednesday

- Visit the course website and read the syllabus.
- Download *OpenIntro Statistics*.
- Install R, RStudio Desktop, and Quarto.
- Bring your laptop to Lab 0.

Why good data analysis and presentation matter

The Challenger launch decision

In 1986, the *Challenger* space shuttle exploded moments after liftoff.
The failure of O-rings in the solid-fuel rocket boosters was blamed for the explosion.

Could this failure have been foreseen?

The evidence engineers emphasized

Morton Thiokol engineers focused on flights with O-ring damage and worried about launching below 53°F.

Flight	Temperature (°F)
2	70
41b	57
41c	63
41d	70
51c	53
61a	79
61c	58

What is missing from this presentation of the data?

Selection changes the story

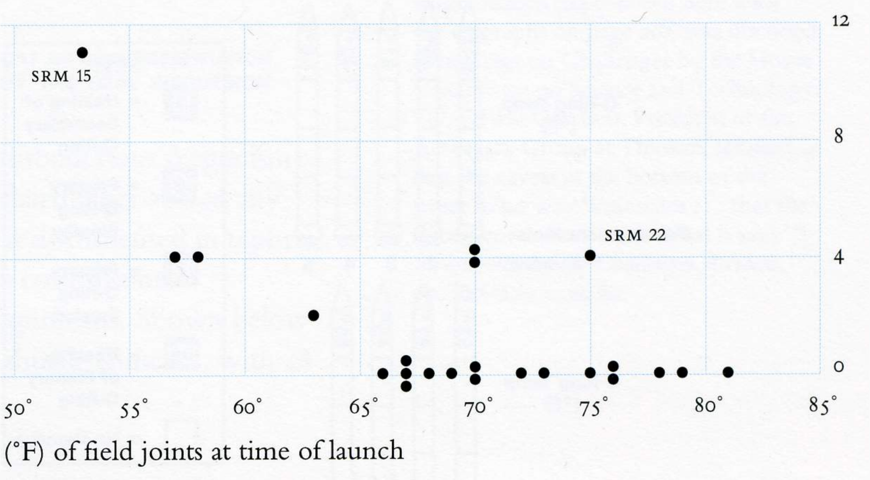
Engineers considered failures, but not the successful launches.

That is selection on the dependent variable: the failures are meaningful only when compared with the successes.

Two immediate improvements:

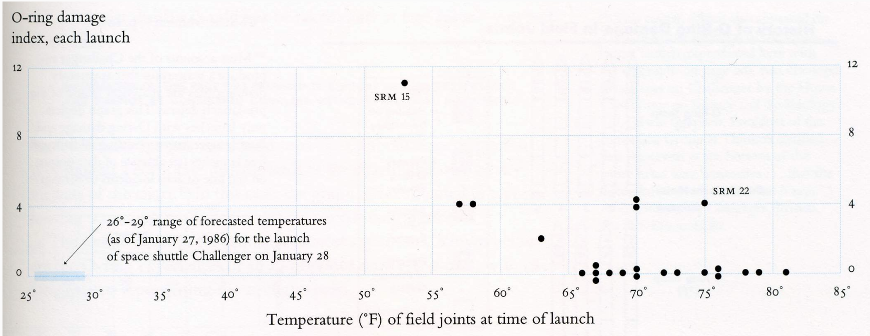
- Include all launches.
- Order the evidence by the explanatory variable—temperature—rather than flight number.

A graph reveals the relationship



A scatterplot makes the temperature–damage relationship easier to see than a table.

The launch was outside prior experience



The forecast launch temperature was only 26–29°F.

From a pattern to a decision quantity

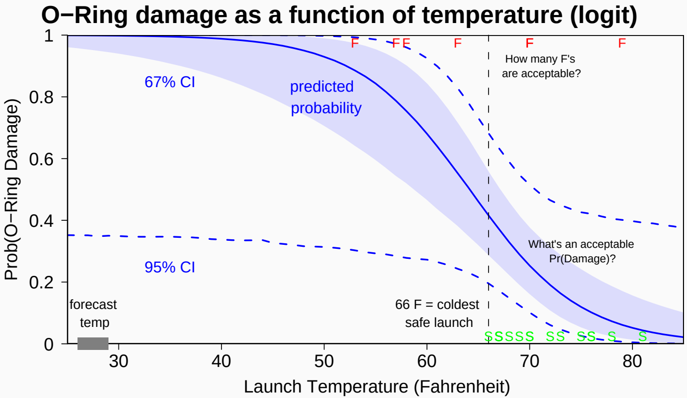
Imagine you are making the launch recommendation. The first question is:

What is the probability of O-ring damage at 29°F?

The scatterplot suggests the risk is high, but decision-makers need a more precise and communicable answer.

That requires a *model*—and a clear way to present its predictions and uncertainty.

Present predictions, data, and uncertainty



A useful graphic connects the observed data to model predictions and makes assumptions visible.

The course in one example

The *Challenger* case is not just about choosing a statistical method.

- Ask the right question.
- Use all relevant data.
- Choose an appropriate comparison and model.
- Communicate the result for a consequential decision.