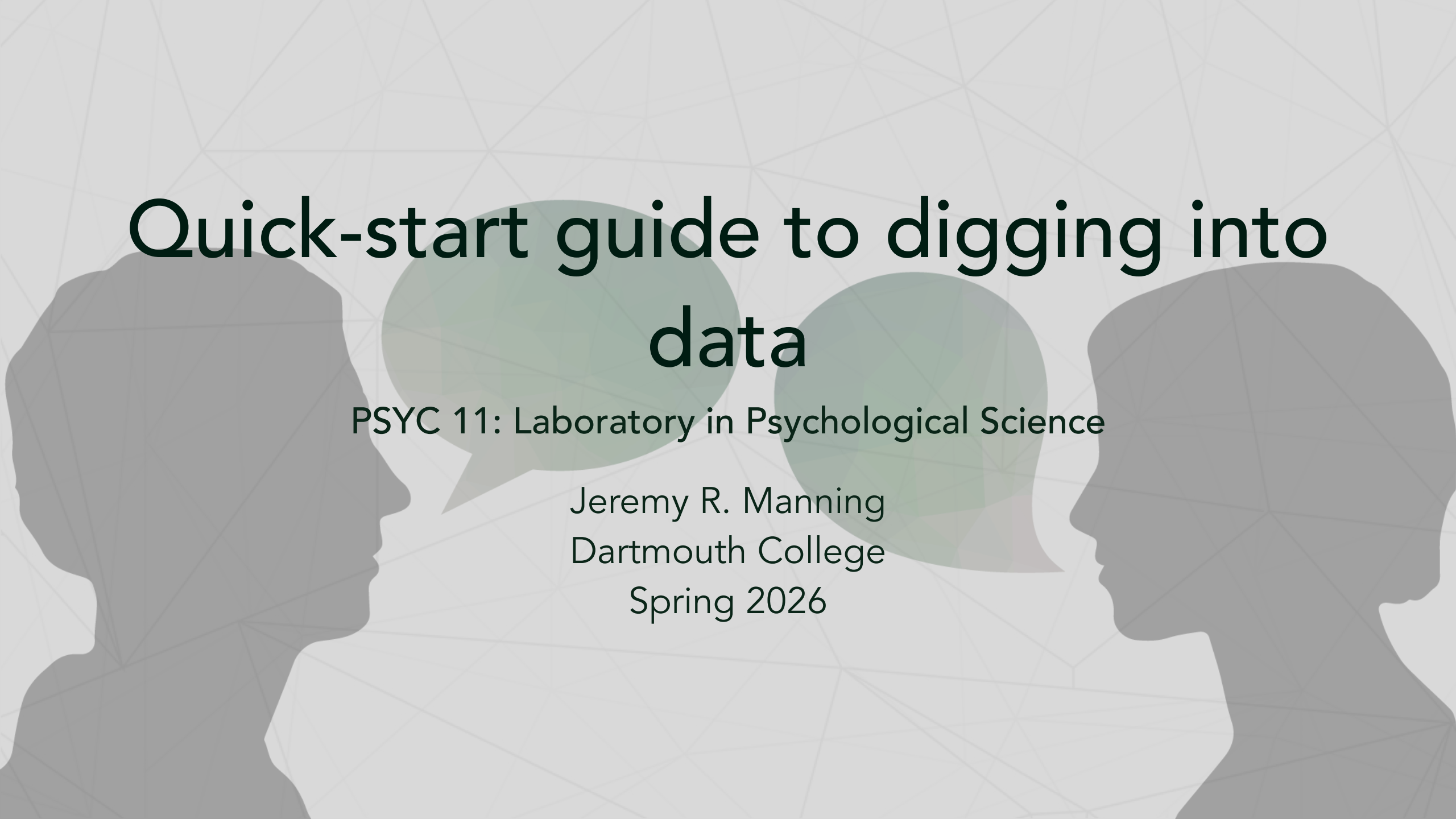

Quick-start guide to digging into data

PSYC 11: Laboratory in Psychological Science

Jeremy R. Manning
Dartmouth College
Spring 2026

Levels of exploration

From Simple to Advanced

- **Level 1:** Look at the raw data (tables, spreadsheets)
- **Level 2:** Visualize the data (histograms, scatter plots, etc.)
- **Level 3:** Run descriptive and inferential statistics
- **Level 4:** Build models and test predictions
- **Level 5:** Compare to other datasets or collect new data

Practical tip: Start with the basics

Your First 10 Minutes with Any Dataset

- How big is it? (rows x columns)
- What are the column names? What do they mean?
- Are there missing values? How are they coded?
- Pick 2-3 columns and make a quick plot

Common pitfalls

Things That Trip People Up

- **Jumping to stats too fast:** always visualize first
- **Ignoring missing data:** blanks, NaNs, and -999s can break your analysis
- **Assuming you know what a column means:** always check the documentation or codebook
- **Not saving your work:** keep a running log of what you tried and what you found

Discussion: What would you do first?

Breakout Groups (8 min)

- You just received a mystery dataset with 10 columns and 1,000 rows
- No documentation -- just data
- What are your **first 5 steps** to figure out what's going on?
- Compare strategies with another group

When you're stuck

Debugging Your Analysis

- **Confused by a column?** Look at unique values, min/max, and the most common entries
- **Plot looks weird?** Check for outliers or data entry errors
- **Stats don't make sense?** Go back to the plot -- does the visual match?
- **Still stuck?** Ask a TA, check Slack, or try a completely different approach

Wrapping up data sleuthing

What to Include in Your Report

- Describe the dataset: what is it, where did it come from, what does it contain?
- Show your key visualizations
- Answer (or explain why you can't answer) the 5 questions
- Reflect: what was surprising? What would you do differently?

Let's dig in!

Today's Goal

- Continue exploring your sleuthing dataset
- Focus on generating clear visualizations and answering the 5 questions
- Wrap up and prepare for the group discussion on Friday