



# Data wrangling: birthday matching lab

PSYC 11: Laboratory in Psychological Science  
March 29, 2024

Jeremy Manning  
[jeremy@dartmouth.edu](mailto:jeremy@dartmouth.edu)

# Posing your questions

- How many students share the same birth month and day?
- How many share the same birth month?
- How many share the same birth day (excluding month)?

# Posing your questions

- Non-scientific framing: guess
- Simple framing: just count and report!
- More sophisticated: try to **explain** what you observed (stats!)

# Statistical tests

- To actually carry out whatever tests or analyses you decide on, you need to **wrangle** your data

# Data wrangling

- Data wrangling means organizing or transforming your data into a format that is more convenient for you to work with

# What do we have?

Birthday survey (S22) ☆ 📁 ☁			
File Edit View Insert Format Data Tools Extensions Help			
↶ ↷ 🖨 📌 100% ▾ \$ % .0 .00 123 ▾ Default (Ari... ▾ 10 ▾ B I			
A1	fx	Timestamp	
	A	B	C
1	Timestamp	What is your birth MONTH?	What is your birth DAY?
2	3/28/2022 13:11:17	September	6
3	3/28/2022 13:11:18	March	27
4	3/28/2022 13:12:25	December	18
5	3/28/2022 13:12:33	October	28
6	3/28/2022 13:12:37	April	26
7	3/28/2022 13:12:39	April	19
8	3/28/2022 13:12:45	March	28
9	3/28/2022 13:12:48	July	1
10	3/28/2022 13:12:50	June	7
11	3/28/2022 13:12:51	March	14
12	3/28/2022 13:12:51	June	6
13	3/28/2022 13:12:52	January	4
14	3/28/2022 13:12:56	March	6
15	3/28/2022 13:13:02	June	21
16	3/28/2022 13:13:12	December	4
17	3/28/2022 13:13:20	April	21
18	3/28/2022 13:13:20	August	7

# Discuss (with your group)

- Are there any **challenges** to analyzing the data in its current form?
- What data format do you **want**?
- How can you “wrangle” the dataset into a more convenient format? (Try it!)

# Example analysis of the birthday data





# Next week: “pitch session” lab

- Goal: learn about how to motivate an idea or question
- Each group will come up with an idea to pitch to the class (Monday)
- You’ll present your pitches and evaluate each other (Wednesday)
- Then we’ll discuss which strategies were effective

