

## Session 5: Data Tidying — Pen-and-Paper Pair Exercise

PSY 410 | Data Science for Psychology

Name: \_\_\_\_\_ Date: \_\_\_\_\_

*No laptop today? No problem. This handout lets you practice the same skills on paper. Work with a partner who has a laptop and compare your work at the end.*

---

### The data: survey\_wide

This dataset has BDI (Beck Depression Inventory) item scores for 3 participants in wide format:

participant	bdi_1	bdi_2	bdi_3	bdi_4
1	2	1	3	2
2	1	0	2	1
3	3	2	2	1

Each column (bdi\_1 through bdi\_4) is a separate BDI item score.

---

### The task (same as the slide exercise)

Which participant has the highest mean BDI score?

**Hint from the slides:** You'll need `pivot_longer()` followed by `group_by()` + `summarize()`.

---

participant	item	response
-------------	------	----------

---

### Your pen-and-paper version

**Step 1: Pivot to long format by hand.** Rewrite the data so each row is one participant-item combination. Fill in the table:

---

participant	item	response
1	bdi_1	2
1	bdi_2	
1	bdi_3	
1	bdi_4	
2	bdi_1	
2	bdi_2	
2	bdi_3	
2	bdi_4	
3	bdi_1	
3	bdi_2	
3	bdi_3	
3	bdi_4	

---

How many rows does the long version have? \_\_\_\_\_ (vs. 3 rows in wide)

**Step 2: Group and summarize.** Calculate the mean response for each participant:

---

participant	sum of responses	n items	mean_bdi (sum / n)
1			
2			
3			

---

Which participant has the highest mean BDI score? \_\_\_\_\_

**Step 3: Write the code.** Fill in the blanks:

```
survey_wide |>
  pivot_longer(
    cols = starts_with("_____"),
    names_to = "_____",
    values_to = "_____"
  ) |>
```

```
group_by(_____) |>  
summarize(mean_bdi = _____(response))
```

---

### **Check your work**

Compare your long table and mean calculations with your partner's screen. Do your answers match?