Session 12: Data Types Grab Bag — 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: `attention_data`

You have survey data with a 1–7 attention check item where the correct answer is 4:

participant_id	attention_check
1	4
2	3
3	4
4	7
5	NA
6	4

The task (same as the slide exercise)

1. Create a new column `passed` that is `TRUE` if they answered 4, `FALSE` otherwise
2. Create a column `status` with three values: "Passed", "Failed", or "No response" (for NA)
3. What proportion of participants passed?

Your pen-and-paper version

Step 1: Create the `passed` column by hand. For each row, evaluate `attention_check == 4`:

participant_id	attention_check	passed
1	4	

participant_id	attention_check	passed
2	3	
3	4	
4	7	
5	NA	
6	4	

Tricky question: What does `NA == 4` return? (Not `FALSE`!) _____

Step 2: Create the `status` column. This requires `case_when()` logic — check conditions in order. Fill in:

participant_id	attention_check	status
1	4	
2	3	
3	4	
4	7	
5	NA	
6	4	

Why do we need to check for `is.na()` first in `case_when()`?

Your answer: _____

Step 3: Calculate the proportion who passed. Using `mean(x == 4, na.rm = TRUE)`:

- How many non-NA values are there? _____
- How many of those equal 4? _____
- Proportion passed = _____ / _____ = _____

Step 4: Write the code. Fill in the blanks:

```
attention_data |>
  mutate(
    passed = attention_check _____ 4,
    status = case_when(
      ______(attention_check) ~ "No response",
      attention_check == 4    ~ "_________",
      .default                = "_________"
    )
  )
```

```r
# Proportion who passed
attention_data |>
  summarize(prop_passed = ______(attention_check == 4, na.rm = _____))
```

Check your work

Compare your filled-in tables and code with your partner's screen. Do your answers match?