Session 13: Strings, Factors & Text — 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: `messy_survey`

You have messy survey responses:

id	gender	comment
1	" FEMALE"	"Great study!"
2	"male"	"too long"
3	"Female"	" Very interesting "
4	"MALE"	"CONFUSING INSTRUCTIONS"
5	"non-binary"	"I enjoyed this"

Notice the extra spaces and inconsistent capitalization.

The task (same as the slide exercise)

1. Clean `gender` to lowercase with no extra spaces
2. Clean `comment` to title case with no extra spaces
3. Create a logical column `is_negative` that is TRUE if the comment contains "long" or "confusing" (case-insensitive)
4. Filter to only negative comments

Your pen-and-paper version

Step 1: Clean `gender` by hand. Apply `str_trim()` then `str_to_lower()` to each value:

id	original gender	after `str_trim()`	after `str_to_lower()`
1	” FEMALE”		
2	“male”		
3	“Female”		
4	“MALE”		
5	“non-binary”		

Step 2: Clean `comment` by hand. Apply `str_trim()` then `str_to_title()`:

id	original comment	after `str_trim()`	after `str_to_title()`
1	“Great study!”		
2	“too long”		
3	” Very interesting ”		
4	“CONFUSING INSTRUCTIONS”		
5	“I enjoyed this”		

Step 3: Detect negative comments. Check if each original comment contains “long” or “confusing” (ignoring case):

id	comment	contains “long”?	contains “confusing”?	is_negative
1	“Great study!”			
2	“too long”			
3	” Very interesting ”			
4	“CONFUSING INSTRUCTIONS”			
5	“I enjoyed this”			

Which rows survive `filter(is_negative)`? Rows _____ and _____

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

```
messy_survey |>
  mutate(
    gender_clean = str_to_______(str_______(gender)),
    comment_clean = str_to_______(str_______(comment)),
```

```
    is_negative = str_detect(
      comment,
      regex("_____|__________", ignore_case = _____)
    )
  ) |>
  filter(__________)
```

Check your work

Compare your cleaned values and code with your partner's screen. Do your answers match?