Socio-Informatics 348
Practical 5

Submission Instructions

- Submit your completed practical as `studentnumber.qmd` on SocSciLearn.
- Submissions are checked for completeness, not correctness.
- At least 80% of exercises must be attempted to receive 1% towards AF assessment.
- Attendance of at least one practical session per week is required to earn the 1% for that week's practical.

Deadline

Friday 19 September, 17:00 (submit on SocSciLearn)

Exercises

Section 1: `stringr` package

1. Compare and contrast the results of `paste0()` with `str_c()` for the following inputs:

```
str_c("hi ", NA)
str_c(letters[1:2], letters[1:3])
```

2. What is the difference between `paste()` and `paste0()`? How can you recreate the equivalent of `paste()` with `str_c()`?

3. Convert the following expressions from `str_c()` to `str_glue()` or vice versa:

```
str_c("The price of ", food, " is ", price)
str_glue("I'm {age} years old and live in {country}")
str_c("\\section{", section_num, "}")
```

4. Take a vector of sentences (e.g., from `stringr::sentences`). Extract the first word, last word, and number of characters in each sentence. Hint: Use `word()` and `str_length()`.

5. Compare the results of using `str_to_lower()`, `str_to_upper()`, and `str_to_title()` on the same text. In what contexts might each be most useful?

Section 2: babynames dataset from the `babynames` package

6. Using the `babynames` dataset (from the `babynames` package), use `str_length()` and `str_sub()` to extract the middle letter from each baby name. What will you do if the string has an even number of characters?

7. Write code to answer the following question with the `babynames` dataset: Are there any major trends in the length of baby names over time? What about the popularity of first and last letters?

8. Create a new variable from `babynames` that contains only the first three letters of each name. What proportion of names share the same first three letters?