Socio-Informatics 348
Practical 8

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 10 October, 17:00 (submit on SocSciLearn)

Chapters Covered:

- R4DS: Chapter 25

Exercises

1. Practice turning the following code snippets into functions. Think about what each function does. What would you call it? How many arguments does it need?

```
mean(is.na(x))
mean(is.na(y))
mean(is.na(z))

x / sum(x, na.rm = TRUE)
y / sum(y, na.rm = TRUE)
z / sum(z, na.rm = TRUE)

round(x / sum(x, na.rm = TRUE) * 100, 1)
round(y / sum(y, na.rm = TRUE) * 100, 1)
round(z / sum(z, na.rm = TRUE) * 100, 1)
```

2. Given a vector of birthdates, write a function to compute the age in years.

3. Write `both_na()`, a summary function that takes two vectors of the same length and returns the number of positions that have an `NA` in both vectors.

4. Using the datasets from `nycflights13`, write a function that:

    1. Finds all flights that were cancelled (i.e. `is.na(arr_time)`) or delayed by more than an hour:

        ```
        flights |> filter_severe()
        ```

    2. Finds all flights that were cancelled or delayed by more than a user-supplied number of hours:

        ```
        flights |> filter_severe(hours = 2)
        ```

5. Scrape the table of "Largest cities in the world by population" from the Wikipedia page and create a data frame in R.