Module 3: R

Submodule 3: Data in R

Expected length: 0.5 day

Guiding question: What does data look like in R?

Concepts: tibble, string, substring, factor, relevel, datetime, missing values

Description: This lesson familiarizes students with the most common data types in R, to prepare them to manipulate those types in data sets.

Instructor Preparation: Run all code to ensure it generates the same output as described in the lesson.

| Materials and resources | Learning objectives |
| --- | --- |
| 03-data-in-r\_deck.html | 1. Understanding vectors and types  2. Familiarity with tibbles  3. Ability to manipulate strings in R  4. Ability to manipulate factors in R  5. Ability to manipulate dates and times in R  6. Awareness of how missing values are handled in R. |

| Length | Lesson content | Guidelines, tips, and tricks |
| --- | --- | --- |
| 30min | Atomic types & vectors  (Wickham and Grolemund, 2017 Chapter 20) |  |
| 15min | Tibbles  (Wickham and Grolemund, 2017 Chapter 10) |  |
| 45min | Strings  (Wickham and Grolemund, 2017, Chapter 14) | For regex, encourage students to use an online regex checker and follow along. |
| 15min | Factors  Wickham and Grolemund, 2017, Chapter 15 |  |
| 30min | Dates and times  Wickham and Grolemund, 2017, Chapter 16 |  |
| 15min | Missing values  Wickham and Grolemund, 2017 Chapter 5 |  |
| 60min | Formative Exercises  1. Make a tibble where the vectors do not have equal length. What happens?  2. In the following tibble, extract variable:    ```{r}  mytibble <- tibble(  A = 1:10,  B = A \* 2)  ```  3. Try using functions paste() and paste0(). Compare them to str\_c. How do they work differently?  4. Look up function str\_trim() and demonstrate application.  5. Match the sequence "'\ with regex  6. Match words that start with x with regex.  7. Match words that are 3 letters long with regex.  8. Match words that only contain consonants with regex.  9. What does ^.\*$ match?  10. What happens if you parse a string with invalid dates? | Allow 30 minutes of independent work time.  Take up for 30 minutes, with RStudio open to demonstrate.  Ask students for errors they encountered. |