

Data transformation

2021-09-02

Loading packages and data

Loading packages, `tidyverse` loads `dplyr` and `tidyr`. In this session we will use the cycling data set from the `exscidata` package.

```
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.0 --

## v ggplot2 3.3.2      v purrr   0.3.4
## v tibble  3.0.3      v dplyr  1.0.2
## v tidyr   1.1.2      v stringr 1.4.0
## v readr   1.4.0      v forcats 0.5.0

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

library(exscidata)
data("cyclingstudy")
```

To see get an overview of the data, use `glimps()`.

```
# glimpse(cyclingstudy)
```

Filter observation by their value using `filter()`

Order the data frame using `arrange()`

Pick variables (columns) from the data frame using `select()`

Create new variables in the data set using `mutate()`

Summarise data using `summarise()` (possibly over grouping through `group_by()`)

Pipe multiple operations together