

Tibbles and Tidy Data

Prof. Eric A. Suess

October 7, 2019

Tidy Data

Today we are starting the Section II **Wrangle**.

This section introduces different ways to access data and discusses the idea of *Tidy Data*.

- ▶ tibbles - a better data.frame, in the tidyverse, data_frame
- ▶ importing data - a better read.csv, in the tidyverse, read_csv
- ▶ tidy data - **variables down columns, observations accross rows**
- ▶ relational data - merging and joining different data_frames
- ▶ strings and factors
- ▶ dates and times

tibbles or tribles

- ▶ The *tibble* package is part of the tidyverse.
- ▶ The `tibble()` function allows you to create a `data_frame` in R, listing data in vectors and it puts them in columns.
- ▶ The `tribble()` functions allows you to create a `data_frame` in R, listing data in columns directly.

read_csv

Comma separated value files are very common.

The `read_csv()` function can be used much like the `tribble()` function. Read data from a .csv file directly into a tibble.

Parsing a vector of ...

- ▶ numbers
- ▶ strings, characters
- ▶ factors
- ▶ dates, times

Reading and writing files to your harddrive

- ▶ `read_csv()`
- ▶ `write_csv()`
- ▶ `read_rds()`
- ▶ `write_rds()`