library(tidyr)

library(tidyverse)

# The weather data we used in the last lesson was already almost tidy when we imported it.

# Please find a lot less tidy version in the resources for this lesson and

# go through the necessary steps to clean it.

weather <- read.csv("009 weather-untidy.csv")

A screenshot of a computer

Description automatically generated

tidy\_data <- weather %>%

gather(day, value, d1:d31, na.rm = TRUE) %>%

mutate(day = sub("^d","",day)) %>%

mutate(day = as.numeric(day)) %>%

arrange(month, day)

A screenshot of a table

Description automatically generated

# Note: read the documentation on the parse\_number() function from the readr package,

# and try to use it when cleaning the data.

?parse\_number

# Play around with the tb data, too;

# try to reproduce what we did in the lesson - you will need to do a little bit more tidying

# than we did together.

# Hint: if you are confused about what to do with the gender-age variables,

# try reading the documentation on the str\_replace() function from the stringr package.

tb <- read.csv("009 tb-untidy.csv")

A screenshot of a data sheet

Description automatically generated

tb\_clean <- tb %>%

select(-mu, -fu, -m04, -m514, -f04, -f514) %>%

replace\_na(list(m04 = 0, m514 = 0, m014 = 0, m1524 = 0, m2534 = 0, m3544 = 0, m4554 = 0, m5564 = 0, m65 = 0, f04 = 0, f514 = 0, f014 = 0, f1524 = 0, f2534 = 0, f3544 = 0, f4554 = 0, f5564 = 0, f65 = 0))

A screenshot of a computer

Description automatically generated

tb\_tidy1 <- tb\_clean %>%

gather(m014:f65, key = "column", value = "cases", na.rm = T) %>%

arrange(country)

A screenshot of a table

Description automatically generated

tb\_tydy2 <- tb\_tidy1 %>%

separate(column, into = c("sex","age"), sep=1)

tb\_tydy2$age <- str\_replace\_all(tb\_tydy2$age,"0","0-")

tb\_tydy2$age <- str\_replace\_all(tb\_tydy2$age,"15","15-")

tb\_tydy2$age <- str\_replace\_all(tb\_tydy2$age,"25","25-")

tb\_tydy2$age <- str\_replace\_all(tb\_tydy2$age,"35","35-")

tb\_tydy2$age <- str\_replace\_all(tb\_tydy2$age,"45","45-")

tb\_tydy2$age <- str\_replace\_all(tb\_tydy2$age,"55","55-")

tb\_tydy2$age <- str\_replace\_all(tb\_tydy2$age,"65","65-100")

tb\_tydy2 <- tb\_tydy2 %>%

separate(age, into = c("age.low", "age.high"))

A screenshot of a computer

Description automatically generated

tb\_united <- tb\_tydy2 %>% unite("age.new", c("age.low","age.high"))

tb\_united

A table of numbers and letters

Description automatically generated