PP4RS | R Module

Slot 3 - Additional Exercises

Dora Simon

06.09.2018

Slot 3

Exercise

Create a csv-file called "simon-family.csv" using the shell that looks like this: (hint: use `touch` and `cat`)

```
cat("
    These are the Simons!
    lastname, firstname, gender, age
    Simon, Dora, female, 28
    Simon, Adam, male, 26
    Simon, Gergely, male, 59
    Csillag, Dora, female,")
```

```
#In shell:
#cat > simon-family.csv
#These are the Simons!
#lastname, firstname, gender, age
#Simon, Dora, female, 28
#Simon, Adam, male, 26
#Simon, Gergely, male, 59
#Csillag, Dora, female,
#Ctrl+d
```

Exercise

Read in simon-family.csv using the `readr` package. Strip the first useless row while reading it in.

```
library(readr)
read_csv("simon-family.csv", skip=1)
```

```
## Parsed with column specification:
## cols(
##   lastname = col_character(),
##   firstname = col_character(),
##   gender = col_character(),
##   age = col_integer()
## )

## # A tibble: 4 x 4
##   lastname firstname gender   age
##   <chr>    <chr>     <chr>  <int>
## 1 Simon    Dora      female    28
## 2 Simon    Adam      male      26
## 3 Simon    Gergely   male      59
## 4 Csillag  Dora      female    NA
```

Exercise Dates

Calculate how many days Dora is older than Adam. Their birthdays are given by this vector: `dates <- c("14.08.90", "21.07.92")`

```
dates <- c("14.08.90", "21.07.92")

better_dates<-as.Date(dates, format="%d.%m.%y")

better_dates[2]-better_dates[1]
```

```
## Time difference of 707 days
```

Exercise Dates Excel

If you are curious about that Excel start date thing, try it yourself!

- Save an Excel file on Windows with the two dates 1899-12-30 and 1904-01-01 in Excel formatting
- Repeat the same with an Excel file on Mac
- Read in both files with read_excel, inspect the current format and bring it to R format

In case you do this, let Dora know! She uses Linux and cannot verify this.