data_cleaning

Fan Gao (Scott)

2024-01-17

Contents

This is my first (Table @ref(tab:kable)); this is my first (Figure @ref(fig:lineplot)), and this is my second (Figure @ref(fig:Histogram)).

Importing packages

```
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 4.2.3
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(readr)
library(tidyverse)
```

```
## Warning: package 'stringr' was built under R version 4.2.3
```

```
## -- Attaching core tidyverse packages ----------------------- tidyverse 2.0.0 --
## v forcats   1.0.0      v stringr   1.5.1
## v ggplot2   3.4.4      v tibble    3.2.1
## v lubridate 1.9.2      v tidyr     1.3.0
## v purrr     1.0.2
```

```
## -- Conflicts ----------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(ggplot2)
library(stats)
library(knitr)
library(kableExtra)
```

```
## Warning in !is.null(rmarkdown::metadata$output) && rmarkdown::metadata$output
## %in% : 'length(x) = 2 > 1' in coercion to 'logical(1)'
```

```
##
## Attaching package: 'kableExtra'
##
## The following object is masked from 'package:dplyr':
##
##     group_rows
```

Counting the number of datas inside BIDS

```
folder_path <- "bids_dataset/Binocular_rivalry"
files <- list.files(folder_path)
print(length(files))
```

```
## [1] 108
```

Reading in 54 participants' data

```
# creating a loop to extra all of the participants' information
participant_data <- c()
for(id in 1:58){
  for(session_id in 1:2){
    file_name <- sprintf("sub-%02d_task-rivalry_run-%02d_events", id, session_id)
    file_path <- file.path(folder_path, paste0(file_name, ".tsv"))
# debugging
#    cat("Checking file:", file_path, "\n")
    if (file.exists(file_path)){
      participant_session_data <- read_tsv(file_path, show_col_types = FALSE)
# adding participant id and session id
      participant_session_data$subject <- id
      participant_session_data$session <- session_id
# append all of participants information into a single data
      participant_data <- append(participant_data, list(participant_session_data))

    }
    #debugging
    else {
      cat("File not found:", file_path, "\n")
      if (id %in% c(43, 53, 54, 56)) {
        cat("Skipping missing participant:", id, "\n")
      }
      next
  }
}}
```

```
## File not found: bids_dataset/Binocular_rivalry/sub-43_task-rivalry_run-01_events.tsv
## Skipping missing participant: 43
## File not found: bids_dataset/Binocular_rivalry/sub-43_task-rivalry_run-02_events.tsv
## Skipping missing participant: 43
## File not found: bids_dataset/Binocular_rivalry/sub-53_task-rivalry_run-01_events.tsv
## Skipping missing participant: 53
## File not found: bids_dataset/Binocular_rivalry/sub-53_task-rivalry_run-02_events.tsv
## Skipping missing participant: 53
```