

Data exploratory analysis Antonia Chroni for St. Jude Children's Research Hospital

Contents

1	Information about this notebook	3
2	Set up	3
3	Directories and paths to file Inputs/Outputs	3
	Read metadata file 4.1 Color palette for plotting	3
5	Number of samples by Sex	4
6	Session Info	6
## ##	The following object is masked _byGlobalEnv:	
##	root_dir	

PI: Stanislav Zakharenko Project: mouse-test-dataset

Task: Training 1 - Demo Project Lead(s): NA

Department: Developmental Neurobiology DNB Bioinformatics Core Analysis Team:

Lead Analyst(s): Antonia Chroni, PhD

Group Lead: Cody A. Ramirez, PhD

 ${\bf Contact~E\text{-}mail:~antonia.chroni@stjude.org}$

DNB Bioinformatics Core Pipeline: Training 1 - Demo

Date started: 10/10/2024Date completed: 10/10/2024

Report generated: 20:41:24 CDT 10/21/2024

1 Information about this notebook

This is an exploratory analysis of the data in the project. We are investigating the number of samples per Sex. This is a demo designed for the purposes of the Best practices training in the context of the Automation and Reproducibility in Computational Biology course.

2 Set up

```
suppressPackageStartupMessages({
  library(tidyverse)
})
```

3 Directories and paths to file Inputs/Outputs

```
attach(params)
## The following object is masked _by_ .GlobalEnv:
##
##
       root_dir
demo_dir <- file.path(root_dir, "courses", "Automation-Reproducibility-compbio", "demo")</pre>
analysis_dir <- file.path(demo_dir, "analyses", "data-exploratory-analysis")</pre>
# We will first read in metadata file as we need to define sample name
metadata_file <- file.path(metadata_dir, "project_metadata.tsv") # metadata input file</pre>
palette_file <- file.path(demo_dir, "figures", "palettes", "binary_color_palette.tsv")</pre>
# File path to `plots` directory
plots dir <- file.path(analysis dir, "plots")</pre>
if (!dir.exists(plots_dir)) {
  dir.create(plots dir)}
figures_plots_dir <- file.path(plots_dir, "figures")</pre>
if (!dir.exists(figures_plots_dir)) {
  dir.create(figures_plots_dir)}
source(paste0(analysis_dir, "/util/function-create-barplot.R"))
source(pasteO(demo_dir, "/figures/scripts/theme_plot.R"))
```

4 Read metadata file

```
# Read metadata
df <- read.csv(metadata_file, sep = '\t', header = TRUE)</pre>
```

4.1 Color palette for plotting

```
# Define and order palette
palette <- palette_df$hex_codes
names(palette) <- palette_df$color_names</pre>
```

$5 \quad \text{Number of samples by Sex} \\$

Table 1: Summary of samples per Sex

Sex	n
F	3
M	5

Segen of Sex

```
pdf(file = fname, width = 6, height = 5)
print(p)
dev.off()
```

```
## pdf
## 2
```

6 Session Info

```
## R version 4.4.0 (2024-04-24)
## Platform: x86_64-pc-linux-gnu
## Running under: Red Hat Enterprise Linux 8.8 (Ootpa)
##
## Matrix products: default
          /usr/lib64/libblas.so.3.8.0
## BLAS:
## LAPACK: /usr/lib64/liblapack.so.3.8.0
## locale:
                                   LC_NUMERIC=C
##
  [1] LC_CTYPE=en_US.UTF-8
  [3] LC_TIME=en_US.UTF-8
                                   LC_COLLATE=en_US.UTF-8
   [5] LC_MONETARY=en_US.UTF-8
                                   LC_MESSAGES=en_US.UTF-8
   [7] LC_PAPER=en_US.UTF-8
                                   LC_NAME=C
##
  [9] LC_ADDRESS=C
                                   LC_TELEPHONE=C
## [11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C
##
## time zone: America/Chicago
## tzcode source: system (glibc)
## attached base packages:
## [1] grid
                 stats
                           graphics grDevices utils
                                                         datasets methods
## [8] base
## other attached packages:
  [1] ggthemes_5.1.0
                        lubridate_1.9.3 forcats_1.0.0
##
                                                         stringr_1.5.1
  [5] dplyr_1.1.4
                        purrr_1.0.2
                                        readr_2.1.5
                                                         tidyr_1.3.1
## [9] tibble_3.2.1
                        ggplot2_3.5.1
                                        tidyverse_2.0.0 yaml_2.3.10
## loaded via a namespace (and not attached):
## [1] sass_0.4.9
                          utf8_1.2.4
                                            generics_0.1.3
                                                               stringi_1.8.4
## [5] hms_1.1.3
                          digest_0.6.37
                                            magrittr_2.0.3
                                                               evaluate_0.24.0
## [9] timechange_0.3.0 fastmap_1.2.0
                                            jsonlite_1.8.8
                                                               tinytex_0.52
## [13] fansi_1.0.6
                          scales_1.3.0
                                            jquerylib_0.1.4
                                                               cli_3.6.3
## [17] rlang 1.1.4
                          crayon_1.5.3
                                            bit64 4.0.5
                                                               munsell_0.5.1
## [21] withr_3.0.1
                          cachem_1.1.0
                                            tools_4.4.0
                                                               parallel_4.4.0
## [25] tzdb_0.4.0
                          colorspace_2.1-1
                                            vctrs_0.6.5
                                                               R6_2.5.1
## [29] mime_0.12
                          lifecycle_1.0.4
                                            bit_4.0.5
                                                               vroom_1.6.5
## [33] pkgconfig_2.0.3
                          pillar_1.9.0
                                            bslib_0.8.0
                                                               gtable_0.3.5
## [37] glue 1.7.0
                          xfun 0.47
                                            tidyselect 1.2.1 highr 0.11
                          farver_2.1.2
                                            htmltools_0.5.8.1 rmarkdown_2.28
## [41] knitr_1.48
## [45] labeling_0.4.3
                          compiler_4.4.0
```