Draft Quarto document

Barbaros Eroglu

targets::tar\_config\_set(store = here::here("\_targets"))  
library(tidyverse)

── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
✔ dplyr 1.1.3 ✔ readr 2.1.4  
✔ forcats 1.0.0 ✔ stringr 1.5.0  
✔ ggplot2 3.4.4 ✔ tibble 3.2.1  
✔ lubridate 1.9.3 ✔ tidyr 1.3.0  
✔ purrr 1.0.2   
── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
✖ dplyr::filter() masks stats::filter()  
✖ dplyr::lag() masks stats::lag()  
ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

library(targets)  
library(tidymodels)

── Attaching packages ────────────────────────────────────── tidymodels 1.1.1 ──  
✔ broom 1.0.5 ✔ rsample 1.2.0  
✔ dials 1.2.0 ✔ tune 1.1.2  
✔ infer 1.0.5 ✔ workflows 1.1.3  
✔ modeldata 1.2.0 ✔ workflowsets 1.0.1  
✔ parsnip 1.1.1 ✔ yardstick 1.2.0  
✔ recipes 1.0.8   
── Conflicts ───────────────────────────────────────── tidymodels\_conflicts() ──  
✖ scales::discard() masks purrr::discard()  
✖ dplyr::filter() masks stats::filter()  
✖ recipes::fixed() masks stringr::fixed()  
✖ dplyr::lag() masks stats::lag()  
✖ yardstick::spec() masks readr::spec()  
✖ recipes::step() masks stats::step()  
• Use suppressPackageStartupMessages() to eliminate package startup messages

source(here::here("R/functions.R"))  
lipidomics <- tar\_read(lipidomics)

## Results

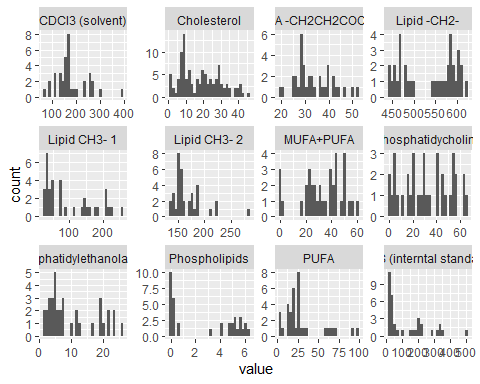
tar\_read(df\_stats\_by\_metabolite) %>%  
 mutate(MeanSD = glue::glue("{value\_mean} ({value\_sd})")) %>%  
 select(Metabolite = metabolite, `Mean SD` = MeanSD) %>%  
 knitr::kable(caption = "Descriptive Statistics of the Metabolites")

Descriptive Statistics of the Metabolites

| Metabolite | Mean SD |
| --- | --- |
| CDCl3 (solvent) | 180 (67) |
| Cholesterol | 18.6 (11.4) |
| FA -CH2CH2COO- | 33.6 (7.8) |
| Lipid -CH2- | 536.6 (61.9) |
| Lipid CH3- 1 | 98.3 (73.8) |
| Lipid CH3- 2 | 168.2 (29.2) |
| MUFA+PUFA | 32.9 (16.1) |
| PUFA | 30 (24.1) |
| Phosphatidycholine | 31.7 (20.5) |
| Phosphatidylethanolamine | 10 (7.6) |
| Phospholipids | 2.7 (2.6) |
| TMS (interntal standard) | 123 (130.4) |

tar\_read(fig\_metabolite\_distribution)

`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



## Building the model

## Running multiple models

model\_estimates <- lipidomics %>%  
 split\_by\_metabolite() %>%   
 map(generate\_model\_results) %>%   
 list\_rbind() %>%   
 filter(str\_detect(term, "metabolite\_"))

Warning: glm.fit: algorithm did not converge

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

model\_estimates

# A tibble: 12 × 5  
 term estimate std.error statistic p.value  
 <chr> <dbl> <dbl> <dbl> <dbl>  
 1 metabolite\_cd\_cl\_3\_solvent 8.70e- 2 0.865 -2.82 0.00475  
 2 metabolite\_cholesterol 2.97e+ 0 0.458 2.38 0.0175   
 3 metabolite\_fa\_ch\_2\_ch\_2\_coo 1.52e+ 0 0.387 1.09 0.276   
 4 metabolite\_lipid\_ch\_2 2.59e- 3 3.14 -1.90 0.0578   
 5 metabolite\_lipid\_ch\_3\_1 4.45e+ 1 1.41 2.70 0.00697  
 6 metabolite\_lipid\_ch\_3\_2 8.85e- 1 0.361 -0.339 0.734   
 7 metabolite\_mufa\_pufa 4.56e- 1 0.449 -1.75 0.0798   
 8 metabolite\_phosphatidycholine 1.28e-120 116628. -0.00237 0.998   
 9 metabolite\_phosphatidylethanolamine 2.69e+ 1 1.32 2.49 0.0129   
10 metabolite\_phospholipids 2.39e- 19 68964. -0.000622 1.00   
11 metabolite\_pufa 3.27e+ 0 0.560 2.11 0.0345   
12 metabolite\_tms\_interntal\_standard 5.62e- 2 0.990 -2.91 0.00363

lipidomics %>%  
 mutate(term = metabolite) %>%  
 column\_values\_to\_snake\_case(term) %>%  
 mutate(term = str\_c("metabolite\_", term)) %>%  
 distinct(term, metabolite) %>%  
 right\_join(model\_estimates, by = "term")

# A tibble: 12 × 6  
 term metabolite estimate std.error statistic p.value  
 <chr> <chr> <dbl> <dbl> <dbl> <dbl>  
 1 metabolite\_tms\_interntal\_st… TMS (inte… 5.62e- 2 9.90e-1 -2.91 0.00363  
 2 metabolite\_cholesterol Cholester… 2.97e+ 0 4.58e-1 2.38 0.0175   
 3 metabolite\_lipid\_ch\_3\_1 Lipid CH3… 4.45e+ 1 1.41e+0 2.70 0.00697  
 4 metabolite\_lipid\_ch\_3\_2 Lipid CH3… 8.85e- 1 3.61e-1 -0.339 0.734   
 5 metabolite\_lipid\_ch\_2 Lipid -CH… 2.59e- 3 3.14e+0 -1.90 0.0578   
 6 metabolite\_fa\_ch\_2\_ch\_2\_coo FA -CH2CH… 1.52e+ 0 3.87e-1 1.09 0.276   
 7 metabolite\_pufa PUFA 3.27e+ 0 5.60e-1 2.11 0.0345   
 8 metabolite\_phosphatidyletha… Phosphati… 2.69e+ 1 1.32e+0 2.49 0.0129   
 9 metabolite\_phosphatidycholi… Phosphati… 1.28e-120 1.17e+5 -0.00237 0.998   
10 metabolite\_phospholipids Phospholi… 2.39e- 19 6.90e+4 -0.000622 1.00   
11 metabolite\_mufa\_pufa MUFA+PUFA 4.56e- 1 4.49e-1 -1.75 0.0798   
12 metabolite\_cd\_cl\_3\_solvent CDCl3 (so… 8.70e- 2 8.65e-1 -2.82 0.00475

lipidomics

# A tibble: 504 × 6  
 code gender age class metabolite value  
 <chr> <chr> <dbl> <chr> <chr> <dbl>  
 1 ERI109 M 25 CT TMS (interntal standard) 208.   
 2 ERI109 M 25 CT Cholesterol 19.8   
 3 ERI109 M 25 CT Lipid CH3- 1 44.1   
 4 ERI109 M 25 CT Lipid CH3- 2 147.   
 5 ERI109 M 25 CT Cholesterol 27.2   
 6 ERI109 M 25 CT Lipid -CH2- 587.   
 7 ERI109 M 25 CT FA -CH2CH2COO- 31.6   
 8 ERI109 M 25 CT PUFA 29.0   
 9 ERI109 M 25 CT Phosphatidylethanolamine 6.78  
10 ERI109 M 25 CT Phosphatidycholine 41.7   
# ℹ 494 more rows

## Figure of model estimates

model\_estimates <- tar\_read(df\_model\_estimates)

tar\_read(fig\_model\_estimates)

