PEC1 Análisis de datos ómicos

Rodrigo Laourou

2024-11-06

Seleccionar dataset

Para seleccionar el dataset, accedemos al repositorio de metaboData dentro de GitHub (https://github.com/nutrimetabolomics/metaboData). Una vez dentro, obtenemos el enlace del dataset (https://raw.githubusercontent.com/nutrimetabolomics/metaboData/refs/heads/main/Datasets/2024-Cachexia/human_cachexia.csv) y lo cargamos desde R.

```
# Leemos el csv desde la url de Github
cachexia <- read.csv("https://raw.githubusercontent.com/nutrimetabolomics/metaboData/refs/heads/main/Da
```

Análisis del dataset

Antes de crear el contenedor apropiado, vamos a analizar nuestro dataset.

```
str(cachexia)
```

```
## 'data.frame':    77 obs. of  65 variables:
##  $ Patient.ID                 : chr  "PIF_178" "PIF_087" "PIF_090" "NETL_005_V1" ...
##  $ Muscle.loss                : chr  "cachexic" "cachexic" "cachexic" "cachexic" ...
##  $ X1.6.Anhydro.beta.D.glucose: num  40.9 62.2 270.4 154.5 22.2 ...
##  $ X1.Methylnicotinamide      : num  65.4 340.4 64.7 53 73.7 ...
##  $ X2.Aminobutyrate           : num  18.7 24.3 12.2 172.4 15.6 ...
##  $ X2.Hydroxyisobutyrate      : num  26.1 41.7 65.4 74.4 83.9 ...
##  $ X2.Oxoglutarate            : num  71.5 67.4 23.8 1199.9 33.1 ...
##  $ X3.Aminoisobutyrate        : num  1480.3 116.8 14.3 555.6 29.7 ...
##  $ X3.Hydroxybutyrate         : num  56.83 43.82 5.64 175.91 76.71 ...
##  $ X3.Hydroxyisovalerate      : num  10.1 79.8 23.3 25 69.4 ...
##  $ X3.Indoxylsulfate          : num  567 369 665 412 166 ...
##  $ X4.Hydroxyphenylacetate    : num  120.3 432.7 292.9 214.9 97.5 ...
##  $ Acetate                    : num  126.5 212.7 314.2 37.3 407.5 ...
##  $ Acetone                    : num  9.49 11.82 4.44 206.44 44.26 ...
##  $ Adipate                    : num  38.1 327 131.6 144 15 ...
##  $ Alanine                    : num  314 871 464 590 1119 ...
##  $ Asparagine                 : num  159.2 157.6 89.1 273.1 42.5 ...
##  $ Betaine                    : num  110 245 117 279 392 ...
##  $ Carnitine                  : num  265.1 120.3 25 200.3 84.8 ...
##  $ Citrate                    : num  3714 2618 863 13630 854 ...
##  $ Creatine                   : num  196.4 212.7 221.4 85.6 105.6 ...
##  $ Creatinine                 : num  16482 15835 24588 20952 6768 ...
```

```
##  $ Dimethylamine             : num  633 608 735 1064 242 ...
##  $ Ethanolamine              : num  645 488 407 821 365 ...
##  $ Formate                   : num  441 252 250 469 114 ...
##  $ Fucose                    : num  337 198.3 186.8 407.5 26.1 ...
##  $ Fumarate                  : num  7.69 18.92 7.1 96.54 19.69 ...
##  $ Glucose                   : num  395 8691 1353 863 6836 ...
##  $ Glutamine                 : num  871 602 302 1686 433 ...
##  $ Glycine                   : num  2039 1108 620 5064 395 ...
##  $ Glycolate                 : num  685.4 652 141.2 70.8 26.6 ...
##  $ Guanidoacetate            : num  154 110 183 103 53 ...
##  $ Hippurate                 : num  4582 1737 4316 757 1153 ...
##  $ Histidine                 : num  925 846 284 1043 327 ...
##  $ Hypoxanthine              : num  97.5 82.3 114.4 223.6 66.7 ...
##  $ Isoleucine                : num  5.58 8.17 9.3 37.71 40.04 ...
##  $ Lactate                   : num  107 369 750 369 3641 ...
##  $ Leucine                   : num  42.1 77.5 31.5 103.5 101.5 ...
##  $ Lysine                    : num  146.9 284.3 97.5 290 122.7 ...
##  $ Methylamine               : num  52.5 23.6 18.7 48.9 27.9 ...
##  $ Methylguanidine           : num  9.97 7.69 4.66 141.17 5.31 ...
##  $ N.N.Dimethylglycine       : num  23.3 87.4 24.5 40 46.1 ...
##  $ O.Acetylcarnitine         : num  52.98 50.4 5.58 254.68 45.6 ...
##  $ Pantothenate              : num  25.8 186.8 145.5 42.5 74.4 ...
##  $ Pyroglutamate             : num  437 437 713 567 185 ...
##  $ Pyruvate                  : num  21.1 37 29.4 64.1 12.3 ...
##  $ Quinolinate               : num  165.7 73 192.5 86.5 38.1 ...
##  $ Serine                    : num  284 392 296 1249 206 ...
##  $ Succinate                 : num  154.5 244.7 142.6 144 68.7 ...
##  $ Sucrose                   : num  45.1 459.4 160.8 111 75.2 ...
##  $ Tartrate                  : num  97.51 32.79 16.28 837.15 4.53 ...
##  $ Taurine                   : num  1920 1261 4273 1525 469 ...
##  $ Threonine                 : num  184.9 198.3 110 376.1 64.1 ...
##  $ Trigonelline              : num  943.9 208.5 192.5 992.3 86.5 ...
##  $ Trimethylamine.N.oxide    : num  2122 639 1153 1451 172 ...
##  $ Tryptophan                : num  259.8 83.1 82.3 235.1 103.5 ...
##  $ Tyrosine                  : num  290 167.3 60.3 323.8 142.6 ...
##  $ Uracil                    : num  111 47 31.5 30.6 44.3 ...
##  $ Valine                    : num  86.5 110 59.1 102.5 160.8 ...
##  $ Xylose                    : num  72.2 192.5 2164.6 125.2 186.8 ...
##  $ cis.Aconitate             : num  237 334 330 1863 101 ...
##  $ myo.Inositol              : num  135.6 376.1 86.5 247.2 750 ...
##  $ trans.Aconitate           : num  51.9 217 58.6 75.9 98.5 ...
##  $ pi.Methylhistidine        : num  157.6 308 145.5 249.6 84.8 ...
##  $ tau.Methylhistidine       : num  160.8 130.3 83.9 254.7 79.8 ...
```

Por lo que podemos observar, este está compuesto de 77 variables. Las dos primeras nos sirven para identificar tanto el paciente (Patient.ID) como si están afectados por caquexia o no (Muscle.loss). El resto de mediciones corresponden con una serie de metabolitos. Midiendo metabolitos en pacientes cachéxicos vs. controles, se pueden identificar cambios metabólicos que indican cómo la enfermedad afecta el metabolismo, identificar biomarcadaores o ver la respuesta al tratamiento a nivel metabolómico. Aunque no se indica, estas mediciones son realizadas en humanos.

Creación del contenedor *SummarizedExperiment*

Una vez hemos cargado los datos, debemos crear el contenedor *SummarizedExperiment*. Para ello, usaremos la librería con el mismo nombre.

```
# Cargamos la librería
library(SummarizedExperiment)
```

```
## Warning: package 'matrixStats' was built under R version 4.4.1
```

```
## Warning: package 'GenomicRanges' was built under R version 4.4.1
```

```
## Warning: package 'S4Vectors' was built under R version 4.4.1
```

```
## Warning: package 'IRanges' was built under R version 4.4.1
```

```
# Creamos la matriz de conteos (datos de expresión de metabolitos)
counts <- as.matrix(cachexia[, -(1:2)])  # Creamos una matriz con todas las variables medidas
rownames(counts) <- cachexia$Patient.ID  # Los nombres de filas son los IDs de los pacientes
counts <- t(counts) # Transponemos la matriz

# Creamos colData (metadatos de columnas) con la información de los pacientes
colData <- DataFrame(Muscle.loss = cachexia$Muscle.loss, row.names = cachexia$Patient.ID)

# Creamos el objeto SummarizedExperiment
se <- SummarizedExperiment(
    assays = list(counts = counts),
    colData = colData)

# Comprobamos el objeto SummarizedExperiment
se
```

```
## class: SummarizedExperiment
## dim: 63 77
## metadata(0):
## assays(1): counts
## rownames(63): X1.6.Anhydro.beta.D.glucose X1.Methylnicotinamide ...
##   pi.Methylhistidine tau.Methylhistidine
## rowData names(0):
## colnames(77): PIF_178 PIF_087 ... NETL_003_V1 NETL_003_V2
## colData names(1): Muscle.loss
```

Podemos comprobar que nuestro objeto se ha creado de la forma correcta.

```
head(se@colData)
```

```
## DataFrame with 6 rows and 1 column
##             Muscle.loss
##             <character>
## PIF_178        cachexic
## PIF_087        cachexic
```

```
## PIF_090        cachexic
## NETL_005_V1    cachexic
## PIF_115        cachexic
## PIF_110        cachexic
```

```
head(se@NAMES)
```

```
## [1] "X1.6.Anhydro.beta.D.glucose" "X1.Methylnicotinamide"
## [3] "X2.Aminobutyrate"            "X2.Hydroxyisobutyrate"
## [5] "X2.Oxoglutarate"             "X3.Aminoisobutyrate"
```

```
head(assay(se))
```

```
##                            PIF_178 PIF_087 PIF_090 NETL_005_V1 PIF_115 PIF_110
## X1.6.Anhydro.beta.D.glucose   40.85   62.18  270.43      154.47   22.20  212.72
## X1.Methylnicotinamide         65.37  340.36   64.72       52.98   73.70   31.82
## X2.Aminobutyrate              18.73   24.29   12.18      172.43   15.64   18.36
## X2.Hydroxyisobutyrate         26.05   41.68   65.37       74.44   83.93   80.64
## X2.Oxoglutarate               71.52   67.36   23.81     1199.91   33.12   47.94
## X3.Aminoisobutyrate         1480.30  116.75   14.30      555.57   29.67   17.46
##                            NETL_019_V1 NETCR_014_V1 NETCR_014_V2 PIF_154
## X1.6.Anhydro.beta.D.glucose      151.41        31.50        51.42  117.92
## X1.Methylnicotinamide             36.60         6.82        30.27   52.46
## X2.Aminobutyrate                   8.67         4.18         7.54   19.49
## X2.Hydroxyisobutyrate             42.52        12.94        34.81   72.24
## X2.Oxoglutarate                  223.63        25.03        80.64   73.70
## X3.Aminoisobutyrate               56.26         8.67        17.99   57.97
##                            NETL_022_V1 NETL_022_V2 NETL_008_V1 PIF_146 PIF_119
## X1.6.Anhydro.beta.D.glucose       20.70      127.74       59.74   89.12   23.57
## X1.Methylnicotinamide            221.41      177.68       50.91   32.79    6.89
## X2.Aminobutyrate                  15.18       12.68        6.82   10.38    2.12
## X2.Hydroxyisobutyrate             28.79       15.03       46.06   32.14    7.85
## X2.Oxoglutarate                  357.81       68.03      111.05   32.46    8.33
## X3.Aminoisobutyrate               93.69      105.64        8.08   43.38    2.97
##                            PIF_099 PIF_162 PIF_160 PIF_113 PIF_143
## X1.6.Anhydro.beta.D.glucose   41.26  589.93  112.17  167.34  183.09
## X1.Methylnicotinamide          8.67   21.98   25.28   19.89   90.92
## X2.Aminobutyrate               2.56   15.18   15.49   13.46    8.94
## X2.Hydroxyisobutyrate          7.85   46.06   47.94   31.19   64.07
## X2.Oxoglutarate                6.89   32.79   28.79   47.94   20.49
## X3.Aminoisobutyrate            6.36   31.82   16.12   79.04   18.73
##                            NETCR_007_V1 NETCR_007_V2 PIF_137 PIF_100
## X1.6.Anhydro.beta.D.glucose       208.51        34.81  333.62   32.46
## X1.Methylnicotinamide              53.52        95.58   35.87    9.68
## X2.Aminobutyrate                    5.26        23.57    7.92    3.90
## X2.Hydroxyisobutyrate              47.94        68.03   54.60   11.02
## X2.Oxoglutarate                   212.72       287.15   20.49  170.72
## X3.Aminoisobutyrate                50.40       104.58   63.43    2.97
##                            NETL_004_V1 PIF_094 PIF_132 PIF_163 NETCR_003_V1
## X1.6.Anhydro.beta.D.glucose        4.71   68.72  214.86  304.90        37.71
## X1.Methylnicotinamide             11.13   13.87  127.74   25.79        10.80
## X2.Aminobutyrate                  43.38   12.18   31.50   27.11         5.00
## X2.Hydroxyisobutyrate             30.88   25.03   33.78   40.45         8.25
```

```
## X2.Oxoglutarate                 104.58   28.22   88.23   70.81        11.70
## X3.Aminoisobutyrate              54.05   72.97   64.07  126.47         8.41
##                              NETL_028_V1 NETL_028_V2 NETCR_013_V1 NETL_020_V1
## X1.6.Anhydro.beta.D.glucose        45.60       34.12       107.77       13.33
## X1.Methylnicotinamide             473.43       92.76        16.61       50.91
## X2.Aminobutyrate                   16.28        8.25        26.84        2.92
## X2.Hydroxyisobutyrate              63.43       16.61        32.46       40.85
## X2.Oxoglutarate                   221.41       55.15        62.80       46.99
## X3.Aminoisobutyrate                15.49        3.39        29.67       22.42
##                              NETL_020_V2 PIF_192 NETCR_012_V1 NETCR_012_V2
## X1.6.Anhydro.beta.D.glucose        27.94  141.17        14.01       244.69
## X1.Methylnicotinamide              80.64   68.03        46.06       116.75
## X2.Aminobutyrate                   15.80   40.85        29.08        40.04
## X2.Hydroxyisobutyrate              64.72   12.81        24.53        61.56
## X2.Oxoglutarate                    88.23   26.05        64.07       174.16
## X3.Aminoisobutyrate                11.70   21.76        13.07        53.52
##                               PIF_089 NETCR_002_V1 PIF_179 PIF_114 NETCR_006_V1
## X1.6.Anhydro.beta.D.glucose  123.97       141.17   35.16  685.40       278.66
## X1.Methylnicotinamide         81.45        28.50   26.58   36.23        40.45
## X2.Aminobutyrate              55.15        20.29    5.21   32.46        55.15
## X2.Hydroxyisobutyrate         70.81        14.30   30.27   85.63        51.42
## X2.Oxoglutarate               92.76        97.51    7.39   25.03        74.44
## X3.Aminoisobutyrate          561.16         8.41    8.41  184.93       354.25
##                              PIF_141 NETCR_025_V1 NETCR_025_V2 NETCR_016_V1
## X1.6.Anhydro.beta.D.glucose   15.80        29.96        16.95       292.95
## X1.Methylnicotinamide         23.57        96.54       114.43        57.97
## X2.Aminobutyrate              17.99         6.55         2.53       167.34
## X2.Hydroxyisobutyrate         37.34        65.37        77.48        82.27
## X2.Oxoglutarate               21.33      1053.63      2465.13       468.72
## X3.Aminoisobutyrate           26.84        14.15        19.49        53.52
##                              PIF_116 PIF_191 PIF_164 NETL_013_V1 PIF_188 PIF_195
## X1.6.Anhydro.beta.D.glucose   29.67   18.92  127.74       34.81   65.37   15.18
## X1.Methylnicotinamide         70.11   24.53 1032.77       12.30   24.05   94.63
## X2.Aminobutyrate               5.58    3.29    8.58        5.87    4.71   11.36
## X2.Hydroxyisobutyrate         18.73   10.49   66.02       15.18   15.80    8.17
## X2.Oxoglutarate                5.53    9.68   38.09       16.78    7.24    5.64
## X3.Aminoisobutyrate            2.61   26.84   66.69       11.25    3.13    5.99
##                              NETCR_015_V1 PIF_102 NETL_010_V1 NETL_010_V2
## X1.6.Anhydro.beta.D.glucose        70.81   25.28       34.47       18.54
## X1.Methylnicotinamide              75.94  101.49       12.81        8.41
## X2.Aminobutyrate                   22.65    8.33        3.78        3.78
## X2.Hydroxyisobutyrate              60.95   59.15        8.33        4.85
## X2.Oxoglutarate                   230.44   88.23       14.30        8.08
## X3.Aminoisobutyrate                53.52   22.65       24.29       22.87
##                              NETL_001_V1 NETCR_015_V2 NETCR_005_V1 PIF_111
## X1.6.Anhydro.beta.D.glucose        37.34        33.78        22.42  146.94
## X1.Methylnicotinamide              55.15        53.52        55.15   10.07
## X2.Aminobutyrate                    7.39        18.17        20.70    6.30
## X2.Hydroxyisobutyrate              36.23        46.53        38.47   27.94
## X2.Oxoglutarate                    75.94        81.45       164.02   24.05
## X3.Aminoisobutyrate                 9.87        44.70       206.44   14.88
##                              PIF_171 NETCR_008_V1 NETCR_008_V2 NETL_017_V1
## X1.6.Anhydro.beta.D.glucose   64.07        32.46       113.30       22.20
## X1.Methylnicotinamide          6.42        14.01        43.38       20.70
```

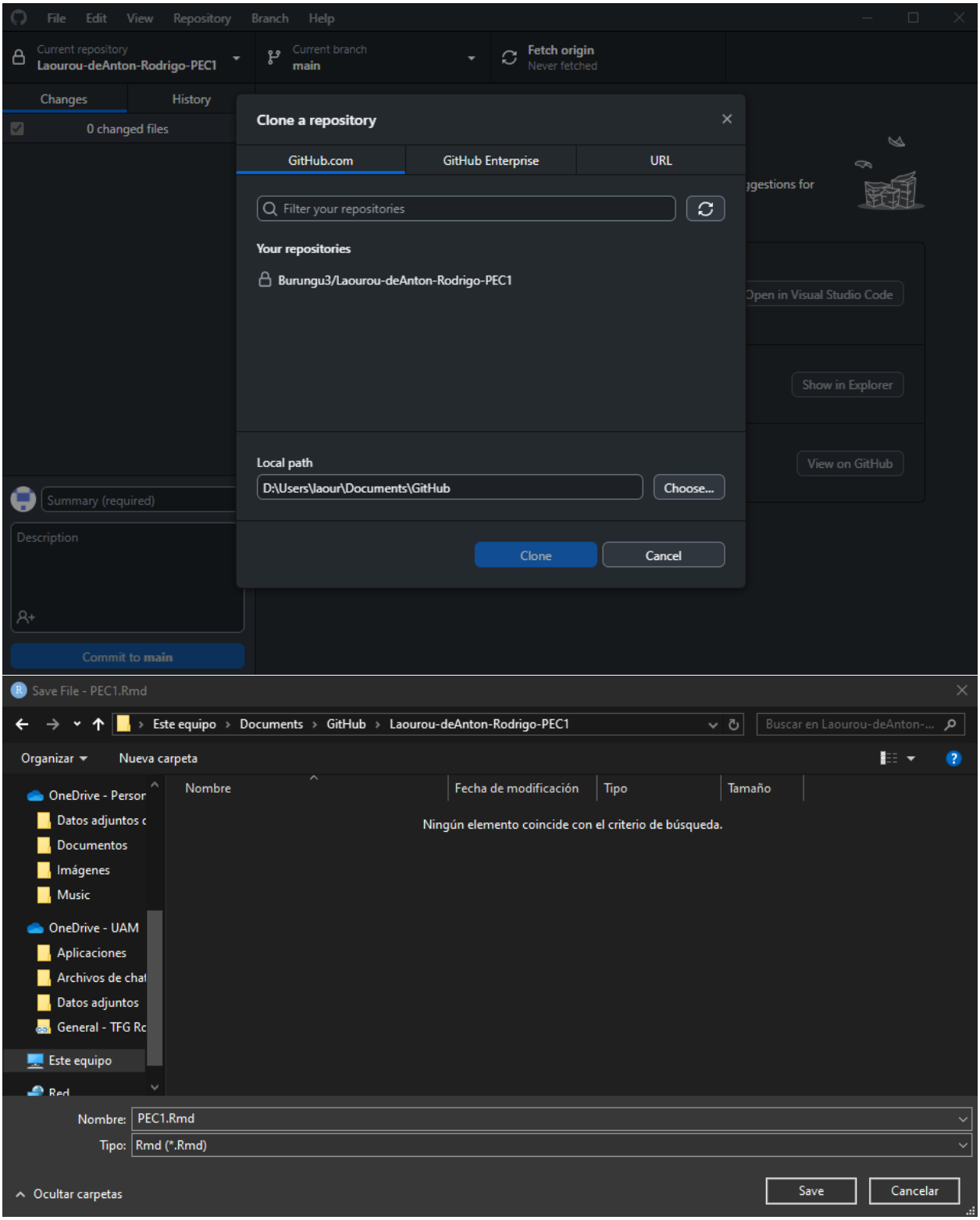
```
## X2.Aminobutyrate              28.79         2.97         4.66        7.85
## X2.Hydroxyisobutyrate         18.92         5.16        27.11       19.69
## X2.Oxoglutarate               85.63         8.08        22.42       38.47
## X3.Aminoisobutyrate           31.82         5.99        27.11        9.30
##                            NETL_017_V2 NETL_002_V1 NETL_002_V2 PIF_190
## X1.6.Anhydro.beta.D.glucose       46.53      192.48      528.48   28.79
## X1.Methylnicotinamide              9.78      108.85      225.88    9.21
## X2.Aminobutyrate                   3.10        7.77       13.46    5.53
## X2.Hydroxyisobutyrate              9.30       46.06       93.69   17.64
## X2.Oxoglutarate                   10.59       55.15      230.44   14.44
## X3.Aminoisobutyrate               13.20        7.03       10.80   15.49
##                            NETCR_009_V1 NETCR_009_V2 NETL_007_V1 PIF_112
## X1.6.Anhydro.beta.D.glucose       181.27        47.47       15.96   22.87
## X1.Methylnicotinamide              48.42         7.69       16.12   10.38
## X2.Aminobutyrate                    8.94         4.06        1.93    1.28
## X2.Hydroxyisobutyrate              51.94         9.30       15.80    5.58
## X2.Oxoglutarate                   982.40        65.37       25.28    8.50
## X3.Aminoisobutyrate               198.34        50.40       13.46   13.74
##                            NETCR_019_V2 NETL_012_V1 NETL_012_V2 NETL_003_V1
## X1.6.Anhydro.beta.D.glucose        35.16       16.95        9.39       37.71
## X1.Methylnicotinamide              52.46       15.80       14.01       18.17
## X2.Aminobutyrate                   13.87       10.49        5.16       26.05
## X2.Hydroxyisobutyrate              44.26       22.42       23.57       15.03
## X2.Oxoglutarate                    99.48       62.80       46.99       23.34
## X3.Aminoisobutyrate               208.51       10.91       13.33       33.45
##                            NETL_003_V2
## X1.6.Anhydro.beta.D.glucose       38.47
## X1.Methylnicotinamide             12.55
## X2.Aminobutyrate                  15.03
## X2.Hydroxyisobutyrate             12.55
## X2.Oxoglutarate                   22.20
## X3.Aminoisobutyrate               21.33
```

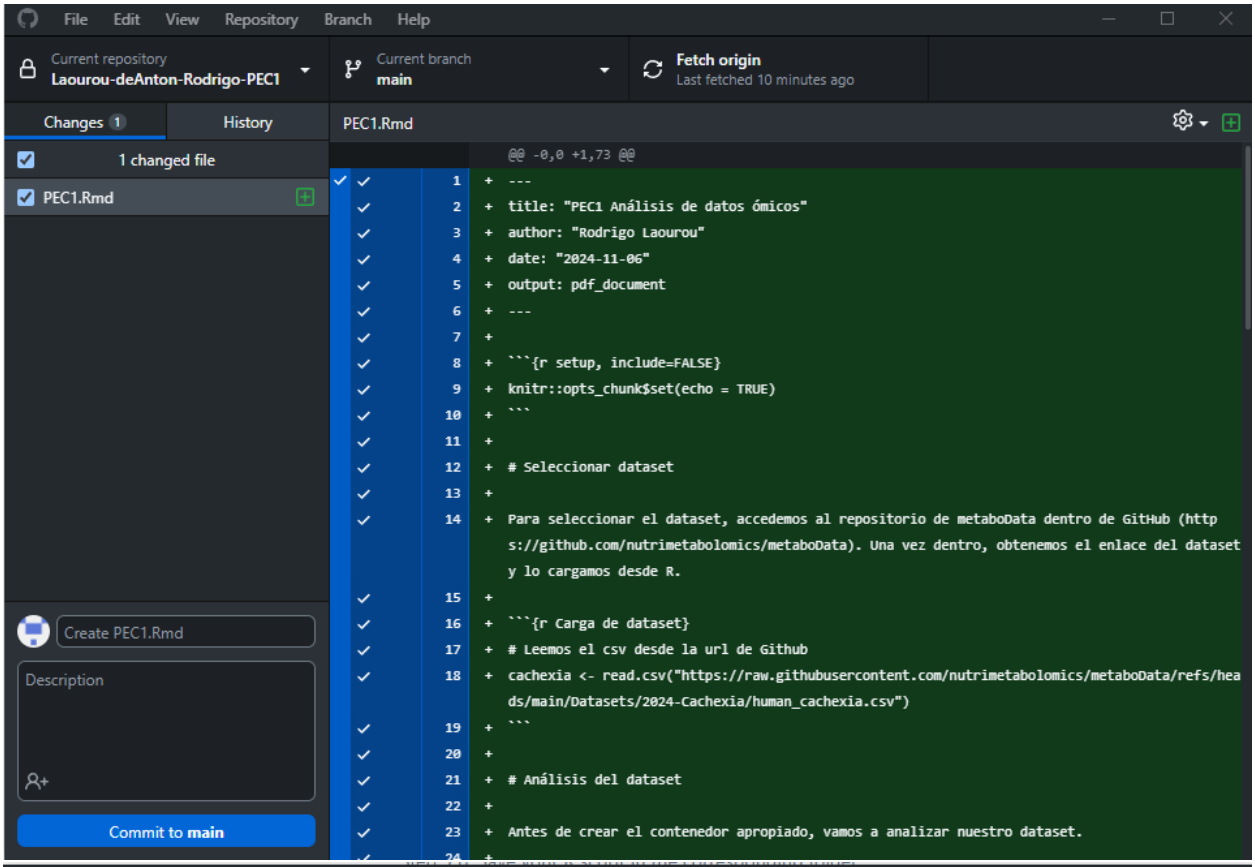
Por último, podemos exportar este objeto a un archivo de texto.

```
# Guardamos el objeto SummarizedExperiment en un archivo .rda
save(se, file = "cachexia_summarized_experiment.rda")
```

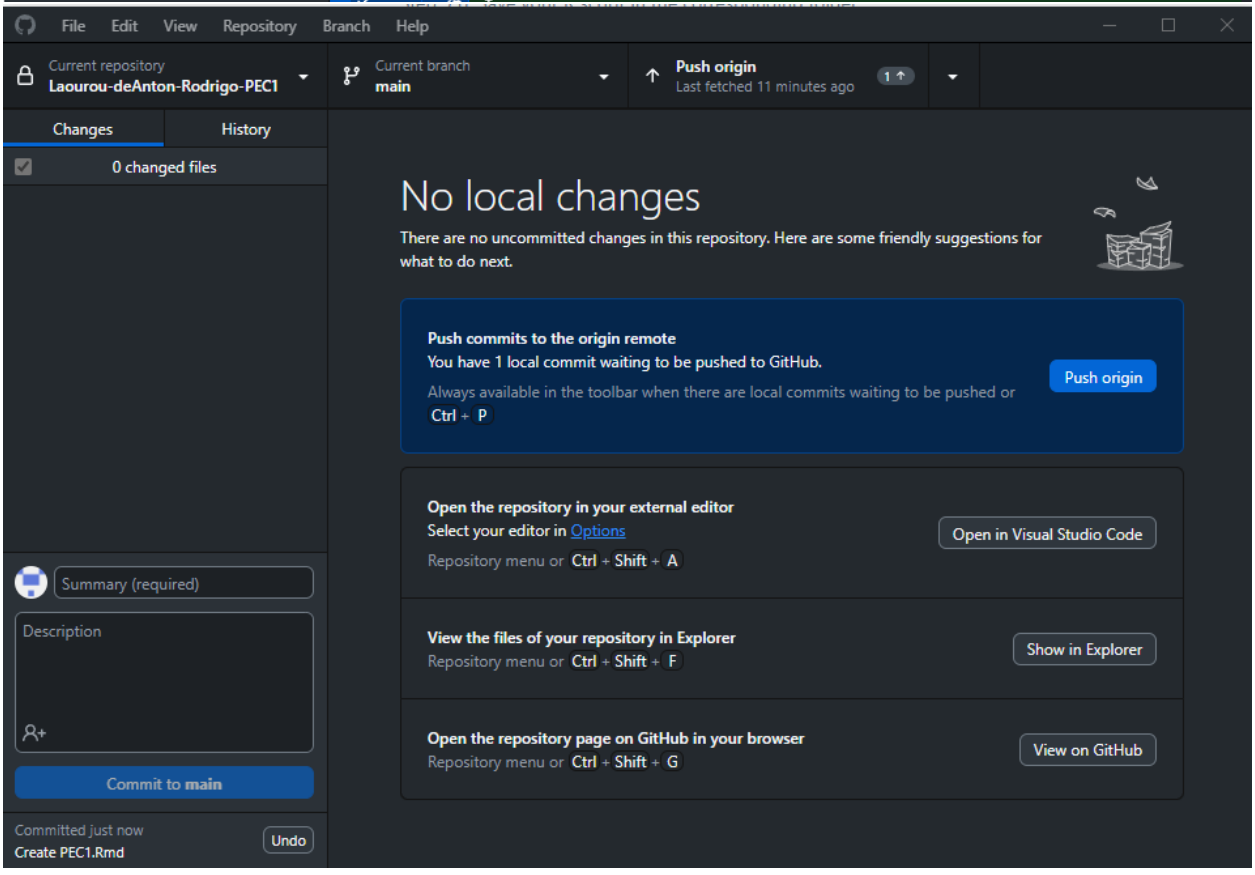
Reposición de los datos a GitHub

Para poder subir este informe a GitHub junto con el resto de datos, instalamos la aplicación de GitHub de escritorio y creamos una carpeta con nuestro repositorio. Una vez hecho esto, guardamos los diferentes archivos como por ejemplo este informe dentro del repositorio.


File Edit View Repository Branch Help
Current repository
Laourou-deAnton-Rodrigo-PEC1
Current branch
main
Fetch origin
Never fetched
Changes
History
0 changed files
Clone a repository
GitHub.com
GitHub Enterprise
URL
Filter your repositories
Your repositories
Burungu3/Laourou-deAnton-Rodrigo-PEC1
Local path
D:\Users\laour\Documents\GitHub
Choose...
Clone
Cancel
Open in Visual Studio Code
Show in Explorer
View on GitHub
Summary (required)
Description
Commit to main
Save File - PEC1.Rmd
Este equipo › Documents › GitHub › Laourou-deAnton-Rodrigo-PEC1
Buscar en Laourou-deAnton-...
Organizar
Nueva carpeta
Nombre
Fecha de modificación
Tipo
Tamaño
Ningún elemento coincide con el criterio de búsqueda.
OneDrive - Persor
Datos adjuntos c
Documentos
Imágenes
Music
OneDrive - UAM
Aplicaciones
Archivos de chat
Datos adjuntos
General - TFG Rc
Este equipo
Red
Nombre: PEC1.Rmd
Tipo: Rmd (*.Rmd)
Ocultar carpetas
Save
Cancelar


File Edit View Repository Branch Help
Current repository
Laourou-deAnton-Rodrigo-PEC1
Current branch
main
Fetch origin
Last fetched 10 minutes ago
Changes 1
History
1 changed file
PEC1.Rmd
PEC1.Rmd
@@ -0,0 +1,73 @@
---
title: "PEC1 Análisis de datos ómicos"
author: "Rodrigo Laourou"
date: "2024-11-06"
output: pdf_document

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```
Seleccionar dataset
Para seleccionar el dataset, accedemos al repositorio de metaboData dentro de GitHub (https://github.com/nutrimetabolomics/metaboData). Una vez dentro, obtenemos el enlace del dataset y lo cargamos desde R.
```{r Carga de dataset}
# Leemos el csv desde la url de Github
cachexia <- read.csv("https://raw.githubusercontent.com/nutrimetabolomics/metaboData/refs/heads/main/Datasets/2024-Cachexia/human_cachexia.csv")
```
Análisis del dataset
Antes de crear el contenedor apropiado, vamos a analizar nuestro dataset.
Create PEC1.Rmd
Description
Commit to main


File Edit View Repository Branch Help
Current repository
Laourou-deAnton-Rodrigo-PEC1
Current branch
main
Push origin
Last fetched 11 minutes ago
1
Changes
History
0 changed files
No local changes
There are no uncommitted changes in this repository. Here are some friendly suggestions for what to do next.
Push commits to the origin remote
You have 1 local commit waiting to be pushed to GitHub.
Always available in the toolbar when there are local commits waiting to be pushed or Ctrl + P
Push origin
Open the repository in your external editor
Select your editor in Options
Repository menu or Ctrl + Shift + A
Open in Visual Studio Code
View the files of your repository in Explorer
Repository menu or Ctrl + Shift + F
Show in Explorer
Open the repository page on GitHub in your browser
Repository menu or Ctrl + Shift + G
View on GitHub
Summary (required)
Description
Commit to main
Committed just now
Create PEC1.Rmd
Undo