

PEC1_script

Gabriel Regueira Huguet

2025-03-20

1) **Seleccionar el dataset:** Hem seleccionat el dataset del repositori Github proporcionat a l'enunciat de la PEC1 anomenat 2024-Cachexia

```
library(SummarizedExperiment) #Carregeum el paquet
```

```
## Cargando paquete requerido: MatrixGenerics
```

```
## Cargando paquete requerido: matrixStats
```

```
##
```

```
## Adjuntando el paquete: 'MatrixGenerics'
```

```
## The following objects are masked from 'package:matrixStats':
```

```
##
```

```
## colAlls, colAnyNAs, colAnys, colAvgsPerRowSet, colCollapse,  
## colCounts, colCummaxs, colCummins, colCumprods, colCumsums,  
## colDiffs, colIQRDiffs, colIQRs, colLogSumExps, colMadDiffs,  
## colMads, colMaxs, colMeans2, colMedians, colMins, colOrderStats,  
## colProds, colQuantiles, colRanges, colRanks, colSdDiffs, colSds,  
## colSums2, colTabulates, colVarDiffs, colVars, colWeightedMads,  
## colWeightedMeans, colWeightedMedians, colWeightedSds,  
## colWeightedVars, rowAlls, rowAnyNAs, rowAnys, rowAvgsPerColSet,  
## rowCollapse, rowCounts, rowCummaxs, rowCummins, rowCumprods,  
## rowCumsums, rowDiffs, rowIQRDiffs, rowIQRs, rowLogSumExps,  
## rowMadDiffs, rowMads, rowMaxs, rowMeans2, rowMedians, rowMins,  
## rowOrderStats, rowProds, rowQuantiles, rowRanges, rowRanks,  
## rowSdDiffs, rowSds, rowSums2, rowTabulates, rowVarDiffs, rowVars,  
## rowWeightedMads, rowWeightedMeans, rowWeightedMedians,  
## rowWeightedSds, rowWeightedVars
```

```
## Cargando paquete requerido: GenomicRanges
```

```
## Cargando paquete requerido: stats4
```

```
## Cargando paquete requerido: BiocGenerics
```

```
##
```

```
## Adjuntando el paquete: 'BiocGenerics'
```

```

## The following objects are masked from 'package:stats':
##
##     IQR, mad, sd, var, xtabs

## The following objects are masked from 'package:base':
##
##     anyDuplicated, aperm, append, as.data.frame, basename, cbind,
##     colnames, dirname, do.call, duplicated, eval, evalq, Filter, Find,
##     get, grep, grepl, intersect, is.unsorted, lapply, Map, mapply,
##     match, mget, order, paste, pmax, pmax.int, pmin, pmin.int,
##     Position, rank, rbind, Reduce, rownames, sapply, saveRDS, setdiff,
##     table, tapply, union, unique, unsplit, which.max, which.min

## Cargando paquete requerido: S4Vectors

##
## Adjuntando el paquete: 'S4Vectors'

## The following object is masked from 'package:utils':
##
##     findMatches

## The following objects are masked from 'package:base':
##
##     expand.grid, I, unname

## Cargando paquete requerido: IRanges

##
## Adjuntando el paquete: 'IRanges'

## The following object is masked from 'package:grDevices':
##
##     windows

## Cargando paquete requerido: GenomeInfoDb

## Cargando paquete requerido: Biobase

## Welcome to Bioconductor
##
##     Vignettes contain introductory material; view with
##     'browseVignettes()'. To cite Bioconductor, see
##     'citation("Biobase")', and for packages 'citation("pkgname)".

##
## Adjuntando el paquete: 'Biobase'

## The following object is masked from 'package:MatrixGenerics':
##
##     rowMedians

```

```
## The following objects are masked from 'package:matrixStats':  
##  
## anyMissing, rowMedians
```

```
library(readr) #Per llegir arxius csv
```

```
## Warning: package 'readr' was built under R version 4.4.3
```

```
library(readxl) #Per llegir arxius exzel
```

```
## Warning: package 'readxl' was built under R version 4.4.3
```

```
library(dplyr)
```

```
##  
## Adjuntando el paquete: 'dplyr'
```

```
## The following object is masked from 'package:Biobase':  
##  
## combine
```

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

```
## The following object is masked from 'package:GenomeInfoDb':  
##  
## intersect
```

```
## The following objects are masked from 'package:IRanges':  
##  
## collapse, desc, intersect, setdiff, slice, union
```

```
## The following objects are masked from 'package:S4Vectors':  
##  
## first, intersect, rename, setdiff, setequal, union
```

```
## The following objects are masked from 'package:BiocGenerics':  
##  
## combine, intersect, setdiff, union
```

```
## The following object is masked from 'package:matrixStats':  
##  
## count
```

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

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

Hem decidit utilitzar les dades 2024-Cachexia i li hem fet una modificació al document per tal que R ho llegeixi bé, ja que al principi al haver-ho dobles comillas R ho llegia tot com si fos una sola columna:

```
data_info <- read.csv("human_cachexia_clean.csv", stringsAsFactors = FALSE)
```

```
View(data_info)
```

2) Crear SummarizedExperiment:

Un cop tenim les dades accessibles procedim a crear el *SummarizedExperiment*:

```
#Extraïem les metadades de les mostres (colData):
col_data <- DataFrame(
  Patient.ID = data_info$Patient.ID,
  Muscle.loss = data_info$Muscle.loss
)

#Extreiem la matriu d'expressió (assay), treiem les dos primeres columnes (Patient.ID i Muscle.loss)
metabolite_data <- t(as.matrix(data_info[, -(1:2)])) #Trasposem la matriu per a que les files siguin el

#Assignem els noms de les mostres (files)
colnames(metabolite_data) <- data_info$Patient.ID
rownames(col_data) <- data_info$Patient.ID

#Creem l'objecte SummarizedExperiment
SumExp <- SummarizedExperiment(
  assays = list(metabolites = metabolite_data),
  colData = col_data
)
```

```
SumExp
```

```
## class: SummarizedExperiment
## dim: 65 77
## metadata(0):
## assays(1): metabolites
## rownames(65): X1 X6.Anhydro.beta.D.glucose ... pi.Methylhistidine
## tau.Methylhistidine
## rowData names(0):
## colnames(77): PIF_178 PIF_087 ... NETL_003_V1 NETL_003_V2
## colData names(2): Patient.ID Muscle.loss
```

65 files = metabolitos

77 columnas = pacientes

```
colData(SumExp)
```

```
## DataFrame with 77 rows and 2 columns
##           Patient.ID Muscle.loss
##           <character> <character>
## PIF_178      PIF_178      cachexic
## PIF_087      PIF_087      cachexic
## PIF_090      PIF_090      cachexic
## NETL_005_V1  NETL_005_V1  cachexic
## PIF_115      PIF_115      cachexic
## ...          ...          ...
## NETCR_019_V2 NETCR_019_V2  control
## NETL_012_V1  NETL_012_V1  control
## NETL_012_V2  NETL_012_V2  control
## NETL_003_V1  NETL_003_V1  control
## NETL_003_V2  NETL_003_V2  control
```

```
dim(SumExp)
```

```
## [1] 65 77
```

```
names(SumExp)
```

```
## [1] "X1" "X6.Anhydro.beta.D.glucose"
## [3] "X1.Methylnicotinamide" "X2.Aminobutyrate"
## [5] "X2.Hydroxyisobutyrate" "X2.Oxoglutarate"
## [7] "X3.Aminoisobutyrate" "X3.Hydroxybutyrate"
## [9] "X3.Hydroxyisovalerate" "X3.Indoxylsulfate"
## [11] "X4.Hydroxyphenylacetate" "Acetate"
## [13] "Acetone" "Adipate"
## [15] "Alanine" "Asparagine"
## [17] "Betaine" "Carnitine"
## [19] "Citrate" "Creatine"
## [21] "Creatinine" "Dimethylamine"
## [23] "Ethanalamine" "Formate"
## [25] "Fucose" "Fumarate"
## [27] "Glucose" "Glutamine"
## [29] "Glycine" "Glycolate"
## [31] "Guanidoacetate" "Hippurate"
## [33] "Histidine" "Hypoxanthine"
## [35] "Isoleucine" "Lactate"
## [37] "Leucine" "Lysine"
## [39] "Methylamine" "Methylguanidine"
## [41] "N" "N.Dimethylglycine"
## [43] "O.Acetylcarnitine" "Pantothenate"
## [45] "Pyroglutamate" "Pyruvate"
## [47] "Quinolate" "Serine"
## [49] "Succinate" "Sucrose"
## [51] "Tartrate" "Taurine"
## [53] "Threonine" "Trigonelline"
## [55] "Trimethylamine.N.oxide" "Tryptophan"
## [57] "Tyrosine" "Uracil"
## [59] "Valine" "Xylose"
## [61] "cis.Aconitate" "myo.Inositol"
## [63] "trans.Aconitate" "pi.Methylhistidine"
## [65] "tau.Methylhistidine"
```

```
rownames(SumExp)[1:5]
```

```
## [1] "X1" "X6.Anhydro.beta.D.glucose"
## [3] "X1.Methylnicotinamide" "X2.Aminobutyrate"
## [5] "X2.Hydroxyisobutyrate"
```

*****FALTEN MES ANALAISIS

Análisis exploratorio

Tenim un *SummarizedExperiment* que consta de : - 77 mostres (pacients) - 65 metabolits (variables) - Metadades amb *Pacient.ID* i *Muscle.loss*

```
summary(assay(SumExp)) #Resumen estadístico por metabolito
```

```
##      PIF_178      PIF_087      PIF_090      NETL_005_V1
## Min.   : 5.58   Min.   : 7.69   Min.   : 4.44   Min.   : 25.03
## 1st Qu.: 52.72   1st Qu.: 78.66   1st Qu.: 31.50   1st Qu.: 102.51
## Median : 154.47   Median : 208.51   Median : 141.17   Median : 247.15
## Mean   : 699.86   Mean   : 708.30   Mean   : 771.79   Mean   : 1021.28
## 3rd Qu.: 416.24   3rd Qu.: 412.10   3rd Qu.: 308.03   3rd Qu.: 673.71
## Max.   :16481.60   Max.   :15835.35   Max.   :24587.66   Max.   :20952.22
## NA's   :2        NA's   :2        NA's   :2        NA's   :2
##      PIF_115      PIF_110      NETL_019_V1      NETCR_014_V1
## Min.   : 4.53   Min.   : 5.05   Min.   : 2.10   Min.   : 1.73
## 1st Qu.: 44.26   1st Qu.: 35.34   1st Qu.: 26.73   1st Qu.: 7.14
## Median : 84.77   Median : 113.30   Median : 91.84   Median : 18.17
## Mean   : 441.22   Mean   : 537.48   Mean   : 400.85   Mean   : 82.77
## 3rd Qu.: 196.62   3rd Qu.: 325.58   3rd Qu.: 223.63   3rd Qu.: 52.52
## Max.   :6836.29   Max.   :15677.78   Max.   :8022.46   Max.   :2208.35
## NA's   :2        NA's   :2        NA's   :2        NA's   :2
##      NETCR_014_V2      PIF_154      NETL_022_V1      NETL_022_V2
## Min.   : 2.41   Min.   : 9.12   Min.   : 4.26   Min.   : 7.17
## 1st Qu.: 14.63   1st Qu.: 43.82   1st Qu.: 31.98   1st Qu.: 38.77
## Median : 39.65   Median : 117.92   Median : 83.93   Median : 127.74
## Mean   : 207.80   Mean   : 478.07   Mean   : 367.52   Mean   : 650.75
## 3rd Qu.: 102.00   3rd Qu.: 405.50   3rd Qu.: 182.20   3rd Qu.: 283.05
## Max.   :6634.24   Max.   :8690.62   Max.   :8433.78   Max.   :19341.34
## NA's   :2        NA's   :2        NA's   :2        NA's   :2
##      NETL_008_V1      PIF_146      PIF_119      PIF_099
## Min.   : 6.05   Min.   : 3.49   Min.   : 1.48   Min.   : 2.230
## 1st Qu.: 36.30   1st Qu.: 28.09   1st Qu.: 5.17   1st Qu.: 5.965
## Median : 83.93   Median : 71.52   Median : 17.46   Median : 18.360
## Mean   : 484.70   Mean   : 355.17   Mean   : 53.48   Mean   : 56.669
## 3rd Qu.: 218.11   3rd Qu.: 152.94   3rd Qu.: 40.65   3rd Qu.: 37.155
## Max.   :15677.78   Max.   :12209.87   Max.   :1480.30   Max.   :1635.980
## NA's   :2        NA's   :2        NA's   :2        NA's   :2
##      PIF_162      PIF_160      PIF_113      PIF_143
## Min.   : 3.29   Min.   : 3.39   Min.   : 2.92   Min.   : 3.10
## 1st Qu.: 17.91   1st Qu.: 26.06   1st Qu.: 18.82   1st Qu.: 39.26
## Median : 64.07   Median : 78.26   Median : 74.44   Median : 82.27
## Mean   : 318.71   Mean   : 424.13   Mean   : 356.55   Mean   : 461.16
## 3rd Qu.: 164.90   3rd Qu.: 193.47   3rd Qu.: 176.13   3rd Qu.: 296.43
```

| | | | |
|--------------------|-----------------|-----------------|-----------------|
| ## Max. :9701.15 | Max. :10198.54 | Max. :6974.39 | Max. :11158.98 |
| ## NA's :2 | NA's :2 | NA's :2 | NA's :2 |
| ## NETCR_007_V1 | NETCR_007_V2 | PIF_137 | PIF_100 |
| ## Min. : 4.85 | Min. : 5.26 | Min. : 4.35 | Min. : 1.55 |
| ## 1st Qu.: 28.80 | 1st Qu.: 44.97 | 1st Qu.: 39.41 | 1st Qu.: 8.85 |
| ## Median : 64.72 | Median : 98.49 | Median : 75.94 | Median : 17.81 |
| ## Mean : 460.75 | Mean : 645.12 | Mean : 546.23 | Mean : 153.92 |
| ## 3rd Qu.: 210.62 | 3rd Qu.: 397.55 | 3rd Qu.: 267.15 | 3rd Qu.: 53.80 |
| ## Max. :9798.65 | Max. :14328.42 | Max. :13359.73 | Max. :5943.18 |
| ## NA's :2 | NA's :2 | NA's :2 | NA's :2 |
| ## NETL_004_V1 | PIF_094 | PIF_132 | PIF_163 |
| ## Min. : 4.71 | Min. : 4.57 | Min. : 6.42 | Min. : 2.41 |
| ## 1st Qu.: 16.20 | 1st Qu.: 25.03 | 1st Qu.: 69.42 | 1st Qu.: 31.00 |
| ## Median : 31.19 | Median : 72.97 | Median : 196.37 | Median : 97.51 |
| ## Mean : 183.79 | Mean : 350.55 | Mean : 1237.54 | Mean : 516.61 |
| ## 3rd Qu.: 108.31 | 3rd Qu.: 186.53 | 3rd Qu.: 641.11 | 3rd Qu.: 330.37 |
| ## Max. :4865.87 | Max. :8349.86 | Max. :33860.35 | Max. :11271.13 |
| ## NA's :2 | NA's :2 | NA's :2 | NA's :2 |
| ## NETCR_003_V1 | NETL_028_V1 | NETL_028_V2 | NETCR_013_V1 |
| ## Min. : 0.790 | Min. : 10.07 | Min. : 1.82 | Min. : 2.69 |
| ## 1st Qu.: 6.425 | 1st Qu.: 46.06 | 1st Qu.: 13.33 | 1st Qu.: 31.82 |
| ## Median : 15.180 | Median : 115.58 | Median : 45.15 | Median : 70.81 |
| ## Mean : 62.813 | Mean : 738.89 | Mean : 199.61 | Mean : 376.69 |
| ## 3rd Qu.: 29.370 | 3rd Qu.: 336.99 | 3rd Qu.: 119.20 | 3rd Qu.: 267.74 |
| ## Max. :1737.150 | Max. :21590.31 | Max. :4188.09 | Max. :11731.12 |
| ## NA's :2 | NA's :2 | NA's :2 | NA's :2 |
| ## NETL_020_V1 | NETL_020_V2 | PIF_192 | NETCR_012_V1 |
| ## Min. : 2.32 | Min. : 3.19 | Min. : 2.08 | Min. : 2.01 |
| ## 1st Qu.: 14.82 | 1st Qu.: 28.64 | 1st Qu.: 20.19 | 1st Qu.: 12.94 |
| ## Median : 37.34 | Median : 61.56 | Median : 45.60 | Median : 24.05 |
| ## Mean : 227.97 | Mean : 327.88 | Mean : 191.82 | Mean : 148.51 |
| ## 3rd Qu.: 104.17 | 3rd Qu.: 153.72 | 3rd Qu.: 110.89 | 3rd Qu.: 64.39 |
| ## Max. :5431.66 | Max. :8349.86 | Max. :5014.05 | Max. :4315.64 |
| ## NA's :2 | NA's :2 | NA's :2 | NA's :2 |
| ## NETCR_012_V2 | PIF_089 | NETCR_002_V1 | PIF_179 |
| ## Min. : 5.53 | Min. : 4.01 | Min. : 3.67 | Min. : 2.18 |
| ## 1st Qu.: 42.32 | 1st Qu.: 49.70 | 1st Qu.: 21.02 | 1st Qu.: 14.88 |
| ## Median : 101.49 | Median : 116.75 | Median : 62.80 | Median : 50.91 |
| ## Mean : 496.29 | Mean : 581.79 | Mean : 270.28 | Mean : 198.65 |
| ## 3rd Qu.: 290.56 | 3rd Qu.: 330.69 | 3rd Qu.: 177.69 | 3rd Qu.: 121.56 |
| ## Max. :13359.73 | Max. :16481.60 | Max. :7631.20 | Max. :3533.34 |
| ## NA's :2 | NA's :2 | NA's :2 | NA's :2 |
| ## PIF_114 | NETCR_006_V1 | PIF_141 | NETCR_025_V1 |
| ## Min. : 5.47 | Min. : 7.32 | Min. : 1.95 | Min. : 4.01 |
| ## 1st Qu.: 32.62 | 1st Qu.: 50.91 | 1st Qu.: 21.66 | 1st Qu.: 36.88 |
| ## Median : 98.49 | Median : 119.10 | Median : 48.42 | Median : 94.63 |
| ## Mean : 502.98 | Mean : 697.47 | Mean : 279.24 | Mean : 579.72 |
| ## 3rd Qu.: 234.00 | 3rd Qu.: 404.56 | 3rd Qu.: 144.90 | 3rd Qu.: 242.27 |
| ## Max. :12332.58 | Max. :19930.37 | Max. :7115.28 | Max. :14764.78 |
| ## NA's :2 | NA's :2 | NA's :2 | NA's :2 |
| ## NETCR_025_V2 | NETCR_016_V1 | PIF_116 | PIF_191 |
| ## Min. : 2.53 | Min. : 6.62 | Min. : 1.120 | Min. : 0.90 |
| ## 1st Qu.: 61.26 | 1st Qu.: 45.40 | 1st Qu.: 7.885 | 1st Qu.: 9.68 |
| ## Median : 120.30 | Median : 127.74 | Median : 27.390 | Median : 21.98 |

| | | | |
|--------------------|-----------------|-----------------|-----------------|
| ## Mean : 745.91 | Mean : 525.02 | Mean : 143.280 | Mean : 72.36 |
| ## 3rd Qu.: 337.24 | 3rd Qu.: 497.73 | 3rd Qu.: 68.400 | 3rd Qu.: 44.70 |
| ## Max. :22247.84 | Max. :14328.42 | Max. :2864.070 | Max. :1702.75 |
| ## NA's :2 | NA's :2 | NA's :2 | NA's :2 |
| ## PIF_164 | NETL_013_V1 | PIF_188 | PIF_195 |
| ## Min. : 6.89 | Min. : 1.21 | Min. : 1.28 | Min. : 1.51 |
| ## 1st Qu.: 47.94 | 1st Qu.: 7.15 | 1st Qu.: 7.03 | 1st Qu.: 7.30 |
| ## Median : 121.51 | Median : 15.18 | Median : 18.92 | Median : 18.54 |
| ## Mean : 639.13 | Mean : 76.81 | Mean : 71.90 | Mean : 170.47 |
| ## 3rd Qu.: 306.53 | 3rd Qu.: 42.85 | 3rd Qu.: 44.26 | 3rd Qu.: 62.80 |
| ## Max. :15063.05 | Max. :2392.27 | Max. :2489.91 | Max. :4817.45 |
| ## NA's :2 | NA's :2 | NA's :2 | NA's :2 |
| ## NETCR_015_V1 | PIF_102 | NETL_010_V1 | NETL_010_V2 |
| ## Min. : 6.17 | Min. : 7.10 | Min. : 1.36 | Min. : 0.79 |
| ## 1st Qu.: 27.12 | 1st Qu.: 35.34 | 1st Qu.: 7.58 | 1st Qu.: 5.56 |
| ## Median : 106.70 | Median : 101.49 | Median : 14.30 | Median : 11.25 |
| ## Mean : 396.24 | Mean : 343.37 | Mean : 64.50 | Mean : 54.02 |
| ## 3rd Qu.: 250.44 | 3rd Qu.: 231.62 | 3rd Qu.: 35.17 | 3rd Qu.: 24.31 |
| ## Max. :9996.60 | Max. :7480.09 | Max. :1480.30 | Max. :1064.22 |
| ## NA's :2 | NA's :2 | NA's :2 | NA's :2 |
| ## NETL_001_V1 | NETCR_015_V2 | NETCR_005_V1 | PIF_111 |
| ## Min. : 1.97 | Min. : 4.39 | Min. : 4.10 | Min. : 1.77 |
| ## 1st Qu.: 20.19 | 1st Qu.: 29.52 | 1st Qu.: 26.57 | 1st Qu.: 15.34 |
| ## Median : 54.05 | Median : 87.36 | Median : 49.40 | Median : 35.87 |
| ## Mean : 289.17 | Mean : 347.33 | Mean : 361.04 | Mean : 137.42 |
| ## 3rd Qu.: 115.00 | 3rd Qu.: 234.28 | 3rd Qu.: 202.39 | 3rd Qu.: 79.56 |
| ## Max. :6974.39 | Max. :8266.78 | Max. :11849.01 | Max. :3827.63 |
| ## NA's :2 | NA's :2 | NA's :2 | NA's :2 |
| ## PIF_171 | NETCR_008_V1 | NETCR_008_V2 | NETL_017_V1 |
| ## Min. : 4.31 | Min. : 1.23 | Min. : 1.14 | Min. : 2.05 |
| ## 1st Qu.: 21.45 | 1st Qu.: 4.00 | 1st Qu.: 15.93 | 1st Qu.: 10.48 |
| ## Median : 62.18 | Median : 13.46 | Median : 46.06 | Median : 23.57 |
| ## Mean : 357.12 | Mean : 42.80 | Mean : 316.91 | Mean : 159.57 |
| ## 3rd Qu.: 177.72 | 3rd Qu.: 28.08 | 3rd Qu.: 107.86 | 3rd Qu.: 56.26 |
| ## Max. :10614.75 | Max. :1339.43 | Max. :7785.36 | Max. :5115.34 |
| ## NA's :2 | NA's :2 | NA's :2 | NA's :2 |
| ## NETL_017_V2 | NETL_002_V1 | NETL_002_V2 | PIF_190 |
| ## Min. : 1.55 | Min. : 3.29 | Min. : 6.23 | Min. : 3.03 |
| ## 1st Qu.: 5.78 | 1st Qu.: 22.43 | 1st Qu.: 50.41 | 1st Qu.: 9.30 |
| ## Median : 15.96 | Median : 49.90 | Median : 100.48 | Median : 24.05 |
| ## Mean : 63.52 | Mean : 240.74 | Mean : 467.35 | Mean : 97.24 |
| ## 3rd Qu.: 34.65 | 3rd Qu.: 125.86 | 3rd Qu.: 284.31 | 3rd Qu.: 63.47 |
| ## Max. :1571.84 | Max. :6768.26 | Max. :13359.73 | Max. :2121.76 |
| ## NA's :2 | NA's :2 | NA's :2 | NA's :2 |
| ## NETCR_009_V1 | NETCR_009_V2 | NETL_007_V1 | PIF_112 |
| ## Min. : 3.10 | Min. : 0.920 | Min. : 1.21 | Min. : 1.230 |
| ## 1st Qu.: 45.45 | 1st Qu.: 6.795 | 1st Qu.: 10.54 | 1st Qu.: 6.145 |
| ## Median : 152.93 | Median : 17.990 | Median : 26.05 | Median : 17.460 |
| ## Mean : 511.55 | Mean : 110.017 | Mean : 118.81 | Mean : 56.986 |
| ## 3rd Qu.: 323.05 | 3rd Qu.: 60.700 | 3rd Qu.: 60.37 | 3rd Qu.: 30.725 |
| ## Max. :13493.99 | Max. :2298.470 | Max. :3165.29 | Max. :1002.250 |
| ## NA's :2 | NA's :2 | NA's :2 | NA's :2 |
| ## NETCR_019_V2 | NETL_012_V1 | NETL_012_V2 | NETL_003_V1 |
| ## Min. : 3.67 | Min. : 1.84 | Min. : 2.69 | Min. : 2.51 |


```
## 1st Qu.: 22.12 1st Qu.: 10.70 1st Qu.: 9.30 1st Qu.: 14.88
## Median : 58.56 Median : 21.33 Median : 24.05 Median : 34.12
## Mean : 342.26 Mean : 142.84 Mean : 147.55 Mean : 159.46
## 3rd Qu.: 180.37 3rd Qu.: 63.44 3rd Qu.: 59.20 3rd Qu.: 90.47
## Max. :10097.06 Max. :3789.54 Max. :3498.19 Max. :3498.19
## NA's :2 NA's :2 NA's :2 NA's :2
## NETL_003_V2
## Min. : 1.62
## 1st Qu.: 12.55
## Median : 24.29
## Mean : 121.70
## 3rd Qu.: 80.70
## Max. :2864.07
## NA's :2
```

```
sum(is.na(assay(SumExp))) #Observamos que hay 154 NA
```

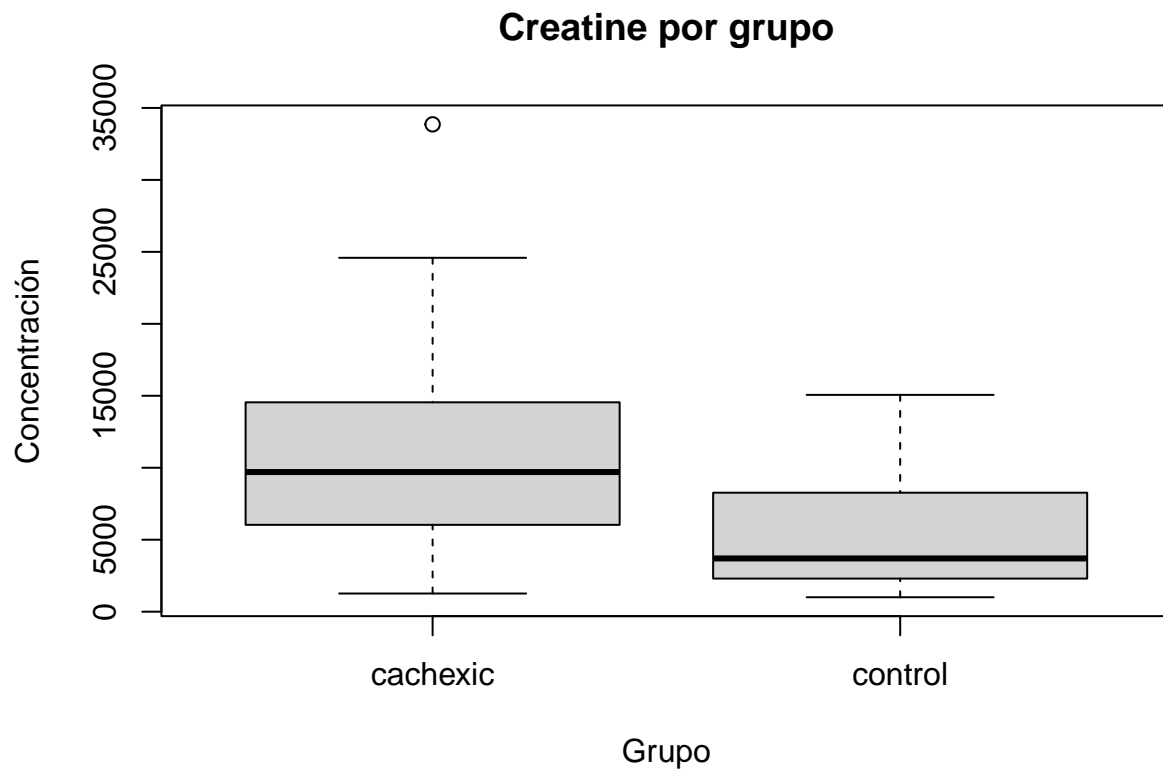
```
## [1] 154
```

```
SumExp <- SumExp[complete.cases(assay(SumExp)), ] #Eliminamos NA
```

```
sum(is.na(assay(SumExp)))
```

```
## [1] 0
```

```
boxplot(assay(SumExp)["Creatine", ] ~ colData(SumExp)$Muscle.loss,
        main="Creatine por grupo",
        ylab = "Concentración", xlab = "Grupo")
```



Por ejemplo, aquí podemos observar que las concentraciones de creatina en pacientes con Cachexic son mayores que el grupo de pacientes control.

PROVA PCA

```
#Trasponemos para que la fila sea los pacientes
df_metabolitos <- t(assay(SumExp))#añadimos la variable grupo como una columna
View(df_metabolitos)
```

```
pca <- prcomp(df_metabolitos, scale. = TRUE)
```

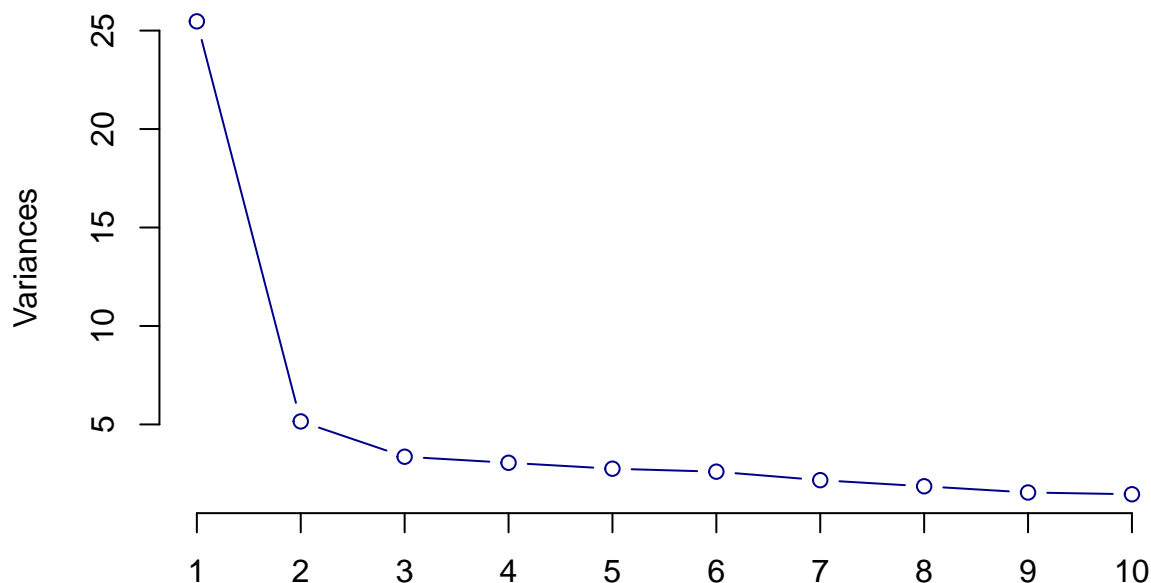
```
summary(pca)
```

```
## Importance of components:
##              PC1    PC2    PC3    PC4    PC5    PC6    PC7
## Standard deviation  5.0467 2.2701 1.83311 1.74728 1.65906 1.6130 1.47304
## Proportion of Variance 0.4043 0.0818 0.05334 0.04846 0.04369 0.0413 0.03444
## Cumulative Proportion 0.4043 0.4861 0.53941 0.58787 0.63156 0.6729 0.70730
##              PC8    PC9    PC10   PC11   PC12   PC13   PC14
## Standard deviation  1.36403 1.24275 1.20650 1.1584 1.05503 1.03620 0.9914
## Proportion of Variance 0.02953 0.02451 0.02311 0.0213 0.01767 0.01704 0.0156
## Cumulative Proportion 0.73683 0.76135 0.78445 0.8057 0.82342 0.84046 0.8561
##              PC15   PC16   PC17   PC18   PC19   PC20   PC21
```

| | | | | | | | |
|---------------------------|---------|---------|---------|---------|---------|---------|---------|
| ## Standard deviation | 0.96773 | 0.89551 | 0.86788 | 0.83041 | 0.8133 | 0.73918 | 0.72112 |
| ## Proportion of Variance | 0.01487 | 0.01273 | 0.01196 | 0.01095 | 0.0105 | 0.00867 | 0.00825 |
| ## Cumulative Proportion | 0.87093 | 0.88366 | 0.89562 | 0.90656 | 0.9171 | 0.92573 | 0.93399 |
| ## | PC22 | PC23 | PC24 | PC25 | PC26 | PC27 | PC28 |
| ## Standard deviation | 0.71053 | 0.64606 | 0.63389 | 0.5830 | 0.5442 | 0.50539 | 0.48743 |
| ## Proportion of Variance | 0.00801 | 0.00663 | 0.00638 | 0.0054 | 0.0047 | 0.00405 | 0.00377 |
| ## Cumulative Proportion | 0.94200 | 0.94863 | 0.95500 | 0.9604 | 0.9651 | 0.96916 | 0.97293 |
| ## | PC29 | PC30 | PC31 | PC32 | PC33 | PC34 | PC35 |
| ## Standard deviation | 0.42674 | 0.42427 | 0.41483 | 0.38653 | 0.35092 | 0.32424 | 0.31646 |
| ## Proportion of Variance | 0.00289 | 0.00286 | 0.00273 | 0.00237 | 0.00195 | 0.00167 | 0.00159 |
| ## Cumulative Proportion | 0.97582 | 0.97867 | 0.98141 | 0.98378 | 0.98573 | 0.98740 | 0.98899 |
| ## | PC36 | PC37 | PC38 | PC39 | PC40 | PC41 | PC42 |
| ## Standard deviation | 0.2867 | 0.28435 | 0.26060 | 0.25353 | 0.24800 | 0.21896 | 0.19537 |
| ## Proportion of Variance | 0.0013 | 0.00128 | 0.00108 | 0.00102 | 0.00098 | 0.00076 | 0.00061 |
| ## Cumulative Proportion | 0.9903 | 0.99158 | 0.99266 | 0.99368 | 0.99465 | 0.99541 | 0.99602 |
| ## | PC43 | PC44 | PC45 | PC46 | PC47 | PC48 | PC49 |
| ## Standard deviation | 0.18914 | 0.1767 | 0.16864 | 0.1580 | 0.15287 | 0.1380 | 0.13101 |
| ## Proportion of Variance | 0.00057 | 0.0005 | 0.00045 | 0.0004 | 0.00037 | 0.0003 | 0.00027 |
| ## Cumulative Proportion | 0.99659 | 0.9971 | 0.99753 | 0.9979 | 0.99830 | 0.9986 | 0.99888 |
| ## | PC50 | PC51 | PC52 | PC53 | PC54 | PC55 | PC56 |
| ## Standard deviation | 0.10759 | 0.10374 | 0.09853 | 0.08760 | 0.08258 | 0.08049 | 0.06927 |
| ## Proportion of Variance | 0.00018 | 0.00017 | 0.00015 | 0.00012 | 0.00011 | 0.00010 | 0.00008 |
| ## Cumulative Proportion | 0.99906 | 0.99923 | 0.99939 | 0.99951 | 0.99962 | 0.99972 | 0.99979 |
| ## | PC57 | PC58 | PC59 | PC60 | PC61 | PC62 | PC63 |
| ## Standard deviation | 0.05937 | 0.05673 | 0.05088 | 0.04001 | 0.02972 | 0.02789 | 0.01876 |
| ## Proportion of Variance | 0.00006 | 0.00005 | 0.00004 | 0.00003 | 0.00001 | 0.00001 | 0.00001 |
| ## Cumulative Proportion | 0.99985 | 0.99990 | 0.99994 | 0.99997 | 0.99998 | 0.99999 | 1.00000 |

```
plot(pca, type = "l", main = "Scree plot", col = "darkblue")
```

Scree plot



```
summary(pca)
```

```
## Importance of components:
##               PC1      PC2      PC3      PC4      PC5      PC6      PC7
## Standard deviation  5.0467  2.2701  1.83311  1.74728  1.65906  1.6130  1.47304
## Proportion of Variance 0.4043  0.0818  0.05334  0.04846  0.04369  0.0413  0.03444
## Cumulative Proportion 0.4043  0.4861  0.53941  0.58787  0.63156  0.6729  0.70730
##               PC8      PC9      PC10     PC11     PC12     PC13     PC14
## Standard deviation  1.36403  1.24275  1.20650  1.1584  1.05503  1.03620  0.9914
## Proportion of Variance 0.02953  0.02451  0.02311  0.0213  0.01767  0.01704  0.0156
## Cumulative Proportion 0.73683  0.76135  0.78445  0.8057  0.82342  0.84046  0.8561
##               PC15     PC16     PC17     PC18     PC19     PC20     PC21
## Standard deviation  0.96773  0.89551  0.86788  0.83041  0.8133  0.73918  0.72112
## Proportion of Variance 0.01487  0.01273  0.01196  0.01095  0.0105  0.00867  0.00825
## Cumulative Proportion 0.87093  0.88366  0.89562  0.90656  0.9171  0.92573  0.93399
##               PC22     PC23     PC24     PC25     PC26     PC27     PC28
## Standard deviation  0.71053  0.64606  0.63389  0.5830  0.5442  0.50539  0.48743
## Proportion of Variance 0.00801  0.00663  0.00638  0.0054  0.0047  0.00405  0.00377
## Cumulative Proportion 0.94200  0.94863  0.95500  0.9604  0.9651  0.96916  0.97293
##               PC29     PC30     PC31     PC32     PC33     PC34     PC35
## Standard deviation  0.42674  0.42427  0.41483  0.38653  0.35092  0.32424  0.31646
## Proportion of Variance 0.00289  0.00286  0.00273  0.00237  0.00195  0.00167  0.00159
## Cumulative Proportion 0.97582  0.97867  0.98141  0.98378  0.98573  0.98740  0.98899
##               PC36     PC37     PC38     PC39     PC40     PC41     PC42
## Standard deviation  0.2867  0.28435  0.26060  0.25353  0.24800  0.21896  0.19537
```

| | | | | | | | | |
|----|------------------------|---------|---------|---------|---------|---------|---------|---------|
| ## | Proportion of Variance | 0.0013 | 0.00128 | 0.00108 | 0.00102 | 0.00098 | 0.00076 | 0.00061 |
| ## | Cumulative Proportion | 0.9903 | 0.99158 | 0.99266 | 0.99368 | 0.99465 | 0.99541 | 0.99602 |
| ## | | PC43 | PC44 | PC45 | PC46 | PC47 | PC48 | PC49 |
| ## | Standard deviation | 0.18914 | 0.1767 | 0.16864 | 0.1580 | 0.15287 | 0.1380 | 0.13101 |
| ## | Proportion of Variance | 0.00057 | 0.0005 | 0.00045 | 0.0004 | 0.00037 | 0.0003 | 0.00027 |
| ## | Cumulative Proportion | 0.99659 | 0.9971 | 0.99753 | 0.9979 | 0.99830 | 0.9986 | 0.99888 |
| ## | | PC50 | PC51 | PC52 | PC53 | PC54 | PC55 | PC56 |
| ## | Standard deviation | 0.10759 | 0.10374 | 0.09853 | 0.08760 | 0.08258 | 0.08049 | 0.06927 |
| ## | Proportion of Variance | 0.00018 | 0.00017 | 0.00015 | 0.00012 | 0.00011 | 0.00010 | 0.00008 |
| ## | Cumulative Proportion | 0.99906 | 0.99923 | 0.99939 | 0.99951 | 0.99962 | 0.99972 | 0.99979 |
| ## | | PC57 | PC58 | PC59 | PC60 | PC61 | PC62 | PC63 |
| ## | Standard deviation | 0.05937 | 0.05673 | 0.05088 | 0.04001 | 0.02972 | 0.02789 | 0.01876 |
| ## | Proportion of Variance | 0.00006 | 0.00005 | 0.00004 | 0.00003 | 0.00001 | 0.00001 | 0.00001 |
| ## | Cumulative Proportion | 0.99985 | 0.99990 | 0.99994 | 0.99997 | 0.99998 | 0.99999 | 1.00000 |

```
## Welcome to POMA!  
## Version 1.16.0  
## POMAShiny app: https://github.com/pcastellanoescuder/POMAShiny
```

```
library(ggtext)
```

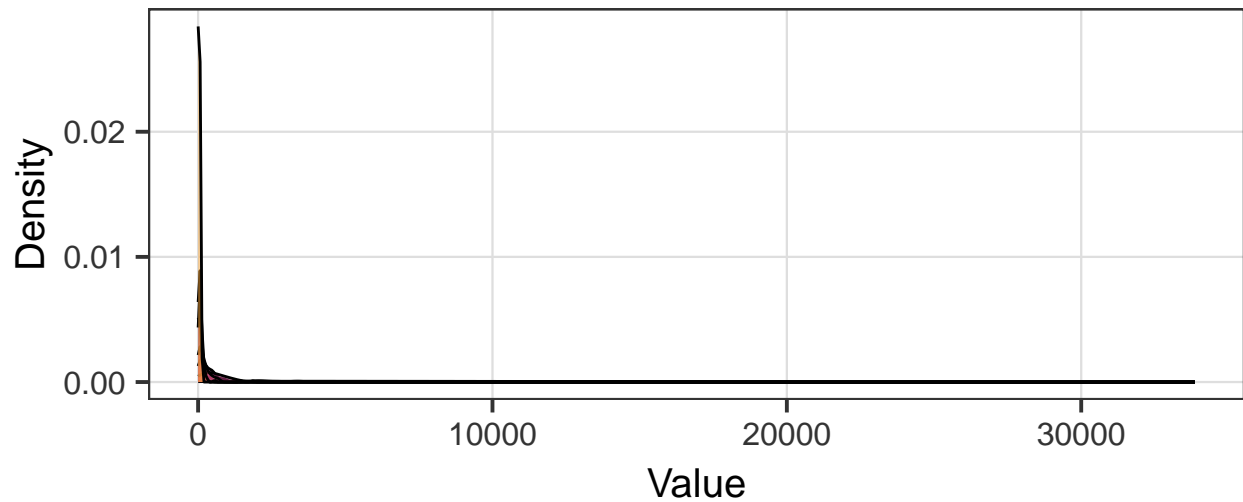
```
library(POMA)
```

```
class(norm_data)
```

```
# Seleccionamos 10 metabolitos al azar
```

```
PomaDensity(SumExp, x = "features", feature_name = selected_features) +
  ggtitle("Distribución de 10 metabolitos seleccionados")
```

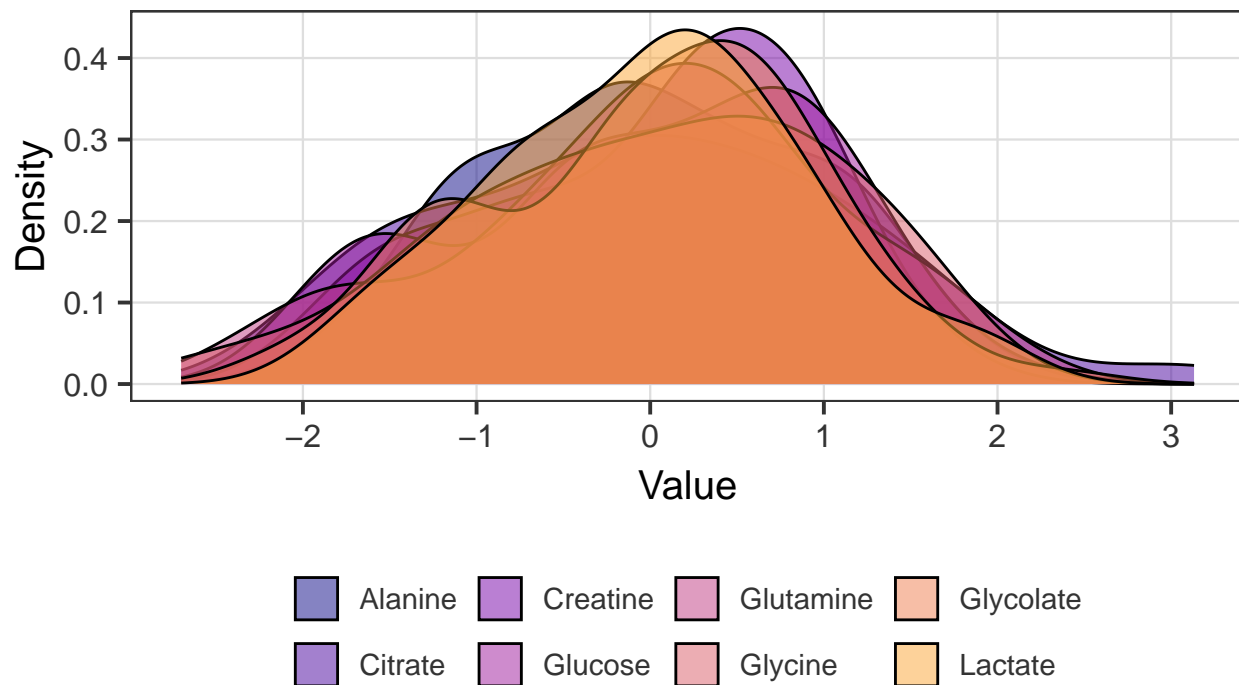
Distribución de 10 metabolitos seleccionados



```
PomaDensity(norm_data, x = "features", feature_name = selected_features) +  
  ggtitle("Distribución de 10 metabolitos seleccionados")
```

```
## Glucos, Piruvate not found
```

Distribución de 10 metabolitos seleccionados



Amb aquests gràfics veiem com afecta normalitzar les dades abans de fer l'anàlisi estadístic propi.

```
norm_data
```

```
## class: SummarizedExperiment
## dim: 63 77
## metadata(0):
## assays(1): ''
## rownames(63): X1 X6.Anhydro.beta.D.glucose ... myo.Inositol
##   trans.Aconitate
## rowData names(0):
## colnames(77): PIF_178 PIF_087 ... NETL_003_V1 NETL_003_V2
## colData names(2): Patient.ID Muscle.loss
```

```
args(POMA::PomaOutliers)
```

```
## function (data, method = "euclidean", type = "median", outcome = NULL,
##   coef = 2, labels = FALSE)
## NULL
```

```
packageVersion("POMA")
```

```
## [1] '1.16.0'
```

```

if (!requireNamespace("BiocManager", quietly = TRUE)) {
  install.packages("BiocManager")
}

# Este paso asegura que estás usando la versión buena de Bioconductor
BiocManager::install(version = "3.20")

## Bioconductor version 3.20 (BiocManager 1.30.25), R 4.4.2 (2024-10-31 ucrt)

## Installation paths not writeable, unable to update packages
## path: C:/Program Files/R/R-4.4.2/library
## packages:
## class, cluster, foreign, KernSmooth, MASS, Matrix, nlme, nnet, rpart,
## spatial, survival

## Old packages: 'cli', 'jsonlite', 'rlang'

# Ahora sí: instalar la versión correcta y limpia de POMA
BiocManager::install("POMA", force = TRUE, ask = FALSE)

## Bioconductor version 3.20 (BiocManager 1.30.25), R 4.4.2 (2024-10-31 ucrt)

## Installing package(s) 'POMA'

## Warning: package 'POMA' is in use and will not be installed

## Installation paths not writeable, unable to update packages
## path: C:/Program Files/R/R-4.4.2/library
## packages:
## class, cluster, foreign, KernSmooth, MASS, Matrix, nlme, nnet, rpart,
## spatial, survival

## Old packages: 'cli', 'jsonlite', 'rlang'

## package 'cli' successfully unpacked and MD5 sums checked

## Warning: cannot remove prior installation of package 'cli'

## Warning in file.copy(savedcopy, lib, recursive = TRUE): problema al copiar
## C:\Users\gabri\AppData\Local\R\win-library\4.4\00LOCK\cli\libs\x64\cli.dll a
## C:\Users\gabri\AppData\Local\R\win-library\4.4\cli\libs\x64\cli.dll: Permission
## denied

## Warning: restored 'cli'

## package 'jsonlite' successfully unpacked and MD5 sums checked

## Warning: cannot remove prior installation of package 'jsonlite'

```



```

## Warning in file.copy(savedcopy, lib, recursive = TRUE): problema al copiar
## C:\Users\gabri\AppData\Local\R\win-library\4.4\00LOCK\jsonlite\libs\x64\jsonlite.dll
## a
## C:\Users\gabri\AppData\Local\R\win-library\4.4\jsonlite\libs\x64\jsonlite.dll:
## Permission denied

## Warning: restored 'jsonlite'

## package 'rlang' successfully unpacked and MD5 sums checked

## Warning: cannot remove prior installation of package 'rlang'

## Warning in file.copy(savedcopy, lib, recursive = TRUE): problema al copiar
## C:\Users\gabri\AppData\Local\R\win-library\4.4\00LOCK\rlang\libs\x64\rlang.dll
## a C:\Users\gabri\AppData\Local\R\win-library\4.4\rlang\libs\x64\rlang.dll:
## Permission denied

## Warning: restored 'rlang'

##
## The downloaded binary packages are in
## C:\Users\gabri\AppData\Local\Temp\Rtmpqi1tNS\downloaded_packages

packageVersion("POMA")

## [1] '1.16.0'

args(POMA::PomaOutliers)

## function (data, method = "euclidean", type = "median", outcome = NULL,
##      coef = 2, labels = FALSE)
## NULL

```