

# Sanchez-Martin-Maria-PEC1

María Sánchez Martín

2024-11-04

## Contents

PEC 1 . . . . .	1
Selección del dataset . . . . .	1
Creación del contenedor <code>SummarizedExperiment</code> . . . . .	2
Exploración del dataset . . . . .	4
Exploración de perfil en GitHub . . . . .	17

## PEC 1

### Selección del dataset

He seleccionado el *dataset de metabolómica* del repositorio de github. Lo he clonado en la carpeta destinada a la PEC1 de esta asignatura de acuerdo con el siguiente tutorial:

- <https://github.com/nutrimetabolomics/metaboData.git>

Procedimiento:

1. Encima de la lista de archivos, haz clic en `<>` **Código**.
2. Copia la dirección URL del repositorio.
3. Abrir el terminal.
4. Cambia el directorio de trabajo actual a la ubicación en donde quieres clonar el directorio.
5. Escribe `git clone` y pegue la dirección URL que ha copiado antes.
6. Presione **Enter** para crear el clon local.

Una vez tenemos el repositorio clonado, procedemos a visualizar los archivos que contiene y seleccionar el que más nos interese.

Tras leer los archivos `README.html` y `Data_Catalog.xlsx`, he decidido seleccionar el dataset `2024-Cachexia`. De acuerdo con la información dada, Cachexia es un síndrome metabólico asociado a otras enfermedades y se caracteriza por la pérdida de masa muscular con o sin pérdida de masa grasa.

Seleccionamos y abrimos los archivos de uso:

```
data <- read.csv("metaboData/Datasets/2024-Cachexia/human_cachexia.csv", row.names = 1) # La primera co
```

Únicamente poseemos un archivo que contiene tanto los datos como los metadatos de expresión.

## Creación del contenedor SummarizedExperiment

Información sobre el paquete de Bioconductor SummarizedExperiment:

<https://bioconductor.org/packages/release/bioc/html/SummarizedExperiment.html>

<https://bioconductor.org/packages/release/bioc/vignettes/SummarizedExperiment/inst/doc/SummarizedExperiment.html>

[https://lcolladotor.github.io/rnaseq\\_LCG-UNAM\\_2021/objetos-de-bioconductor-para-datos-de-expresi%C3%B3n.html](https://lcolladotor.github.io/rnaseq_LCG-UNAM_2021/objetos-de-bioconductor-para-datos-de-expresi%C3%B3n.html)

De acuerdo con el uso del paquete, hemos de tener los datos de expresión y los metadatos separados.

```
metadata <- as.data.frame(data$Muscle.loss) # Metadatos, la primera columna contiene información sobre
expressiondata <- t(as.matrix(data[,-1])) # Convertimos los datos de expresión en matriz
rownames(metadata) <- rownames(data)
colnames(metadata) <- 'Muscle.loss'
colnames(expressiondata) <- rownames(data)
```

!!!! Importante!!! Las columnas de los datos de expresión deben coincidir con las filas de los metadatos, por eso hallamos la matriz traspuesta, así coincidirán (t())

Aplicamos la función de interés SummarizedExperiment().

```
library(SummarizedExperiment)
```

```
## 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, 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

se <- SummarizedExperiment(
  assays = SimpleList(counts = expressiondata),
  colData = metadata
)
```

## Exploración del dataset

Información general del dataset que contiene los datos de expresión:

```
expr_data <- data[, -1]
summary(expr_data)
```

```
## X1.6.Anhydro.beta.D.glucose X1.Methylnicotinamide X2.Aminobutyrate
## Min. : 4.71 Min. : 6.42 Min. : 1.28
## 1st Qu.: 28.79 1st Qu.: 15.80 1st Qu.: 5.26
## Median : 45.60 Median : 36.60 Median : 10.49
## Mean : 105.63 Mean : 71.57 Mean : 18.16
## 3rd Qu.: 141.17 3rd Qu.: 73.70 3rd Qu.: 19.49
## Max. : 685.40 Max. : 1032.77 Max. : 172.43
## X2.Hydroxyisobutyrate X2.Oxoglutarate X3.Aminoisobutyrate X3.Hydroxybutyrate
## Min. : 4.85 Min. : 5.53 Min. : 2.61 Min. : 1.70
## 1st Qu.: 15.80 1st Qu.: 22.42 1st Qu.: 11.70 1st Qu.: 5.99
## Median : 32.46 Median : 55.15 Median : 22.65 Median : 11.70
## Mean : 37.25 Mean : 145.09 Mean : 76.76 Mean : 21.72
## 3rd Qu.: 54.60 3rd Qu.: 92.76 3rd Qu.: 56.26 3rd Qu.: 29.96
## Max. : 93.69 Max. : 2465.13 Max. : 1480.30 Max. : 175.91
## X3.Hydroxyisovalerate X3.Indoxylsulfate X4.Hydroxyphenylacetate
## Min. : 0.92 Min. : 27.66 Min. : 15.49
## 1st Qu.: 5.26 1st Qu.: 82.27 1st Qu.: 41.68
## Median : 12.55 Median : 144.03 Median : 70.11
## Mean : 21.65 Mean : 218.88 Mean : 112.02
## 3rd Qu.: 30.27 3rd Qu.: 333.62 3rd Qu.: 145.47
## Max. : 164.02 Max. : 1043.15 Max. : 796.32
## Acetate Acetone Adipate Alanine
```

##	Min. : 3.49	Min. : 2.29	Min. : 1.55	Min. : 16.78
##	1st Qu.: 16.28	1st Qu.: 4.95	1st Qu.: 6.11	1st Qu.: 78.26
##	Median : 39.65	Median : 7.10	Median : 10.18	Median : 194.42
##	Mean : 66.14	Mean : 11.43	Mean : 24.76	Mean : 273.56
##	3rd Qu.: 86.49	3rd Qu.: 10.49	3rd Qu.: 19.11	3rd Qu.: 399.41
##	Max. : 411.58	Max. : 206.44	Max. : 327.01	Max. : 1312.91
##	Asparagine	Betaine	Carnitine	Citrate
##	Min. : 6.69	Min. : 2.29	Min. : 2.18	Min. : 59.74
##	1st Qu.: 20.49	1st Qu.: 28.79	1st Qu.: 14.44	1st Qu.: 788.40
##	Median : 42.10	Median : 64.72	Median : 23.81	Median : 1790.05
##	Mean : 62.28	Mean : 90.32	Mean : 52.09	Mean : 2235.35
##	3rd Qu.: 89.12	3rd Qu.: 127.74	3rd Qu.: 60.95	3rd Qu.: 3071.74
##	Max. : 273.14	Max. : 391.51	Max. : 487.85	Max. : 13629.61
##	Creatine	Creatinine	Dimethylamine	Ethanolamine
##	Min. : 2.75	Min. : 1002	Min. : 41.26	Min. : 16.12
##	1st Qu.: 17.64	1st Qu.: 3498	1st Qu.: 142.59	1st Qu.: 86.49
##	Median : 44.26	Median : 7631	Median : 304.90	Median : 204.38
##	Mean : 126.83	Mean : 8734	Mean : 358.17	Mean : 276.26
##	3rd Qu.: 117.92	3rd Qu.: 12333	3rd Qu.: 454.86	3rd Qu.: 407.48
##	Max. : 1863.11	Max. : 33860	Max. : 1556.20	Max. : 1436.55
##	Formate	Fucose	Fumarate	Glucose
##	Min. : 6.42	Min. : 5.70	Min. : 0.79	Min. : 26.84
##	1st Qu.: 53.52	1st Qu.: 29.37	1st Qu.: 2.23	1st Qu.: 80.64
##	Median : 95.58	Median : 61.56	Median : 4.10	Median : 210.61
##	Mean : 147.40	Mean : 88.67	Mean : 8.44	Mean : 559.85
##	3rd Qu.: 167.34	3rd Qu.: 123.97	3rd Qu.: 7.85	3rd Qu.: 407.48
##	Max. : 1480.30	Max. : 407.48	Max. : 96.54	Max. : 8690.62
##	Glutamine	Glycine	Glycolate	Guanidoacetate
##	Min. : 23.34	Min. : 38.09	Min. : 5.42	Min. : 7.03
##	1st Qu.: 113.30	1st Qu.: 262.43	1st Qu.: 50.91	1st Qu.: 33.78
##	Median : 225.88	Median : 528.48	Median : 130.32	Median : 64.72
##	Mean : 306.87	Mean : 880.72	Mean : 187.99	Mean : 86.37
##	3rd Qu.: 445.86	3rd Qu.: 1096.63	3rd Qu.: 267.74	3rd Qu.: 108.85
##	Max. : 1685.81	Max. : 5064.45	Max. : 720.54	Max. : 561.16
##	Hippurate	Histidine	Hypoxanthine	Isoleucine
##	Min. : 92.76	Min. : 14.15	Min. : 3.78	Min. : 1.790
##	1st Qu.: 492.75	1st Qu.: 66.69	1st Qu.: 20.70	1st Qu.: 3.900
##	Median : 1224.15	Median : 174.16	Median : 40.04	Median : 7.170
##	Mean : 2286.84	Mean : 292.64	Mean : 61.10	Mean : 8.709
##	3rd Qu.: 2921.93	3rd Qu.: 419.89	3rd Qu.: 83.93	3rd Qu.: 11.250
##	Max. : 19341.34	Max. : 1863.11	Max. : 265.07	Max. : 40.040
##	Lactate	Leucine	Lysine	Methylamine
##	Min. : 7.32	Min. : 2.51	Min. : 10.49	Min. : 1.51
##	1st Qu.: 35.52	1st Qu.: 9.12	1st Qu.: 30.27	1st Qu.: 5.26
##	Median : 81.45	Median : 19.11	Median : 69.41	Median : 14.73
##	Mean : 158.46	Mean : 24.36	Mean : 108.79	Mean : 17.38
##	3rd Qu.: 139.77	3rd Qu.: 31.19	3rd Qu.: 121.51	3rd Qu.: 24.05
##	Max. : 3640.95	Max. : 103.54	Max. : 788.40	Max. : 52.46
##	Methylguanidine	N.N.Dimethylglycine	O.Acetylcarnitine	Pantothenate
##	Min. : 1.70	Min. : 0.79	Min. : 1.23	Min. : 2.59
##	1st Qu.: 4.26	1st Qu.: 7.03	1st Qu.: 3.94	1st Qu.: 11.13
##	Median : 7.85	Median : 21.98	Median : 11.47	Median : 22.65
##	Mean : 15.32	Mean : 26.35	Mean : 19.73	Mean : 44.88
##	3rd Qu.: 19.30	3rd Qu.: 40.04	3rd Qu.: 20.91	3rd Qu.: 41.26

```

## Max. :141.17 Max. :120.30 Max. :254.68 Max. :692.29
## Pyroglutamate Pyruvate Quinolinate Serine
## Min. : 21.33 Min. : 0.90 Min. : 5.21 Min. : 16.12
## 1st Qu.: 68.72 1st Qu.: 4.85 1st Qu.: 26.58 1st Qu.: 83.10
## Median : 157.59 Median : 13.46 Median : 51.42 Median : 142.59
## Mean : 211.45 Mean : 21.29 Mean : 66.44 Mean : 197.69
## 3rd Qu.: 301.87 3rd Qu.: 29.08 3rd Qu.: 87.36 3rd Qu.: 270.43
## Max. :1064.22 Max. :184.93 Max. :259.82 Max. :1248.88
## Succinate Sucrose Tartrate Taurine
## Min. : 1.72 Min. : 6.49 Min. : 2.20 Min. : 17.81
## 1st Qu.: 8.58 1st Qu.: 19.30 1st Qu.: 6.89 1st Qu.: 99.48
## Median : 30.88 Median : 40.85 Median : 12.94 Median : 249.64
## Mean : 60.23 Mean : 113.23 Mean : 40.00 Mean : 525.12
## 3rd Qu.: 74.44 3rd Qu.: 94.63 3rd Qu.: 25.79 3rd Qu.: 665.14
## Max. :589.93 Max. :2079.74 Max. :837.15 Max. :4272.69
## Threonine Trigonelline Trimethylamine.N.oxide Tryptophan
## Min. : 8.25 Min. : 10.07 Min. : 55.7 Min. : 8.67
## 1st Qu.: 31.82 1st Qu.: 53.52 1st Qu.: 175.9 1st Qu.: 21.33
## Median : 64.07 Median : 114.43 Median : 383.8 Median : 46.99
## Mean : 95.36 Mean : 270.44 Mean : 652.2 Mean : 66.24
## 3rd Qu.:137.00 3rd Qu.: 340.36 3rd Qu.: 735.1 3rd Qu.: 96.54
## Max. :450.34 Max. :2252.96 Max. :5486.2 Max. :259.82
## Tyrosine Uracil Valine Xylose
## Min. : 4.22 Min. : 3.10 Min. : 4.10 Min. : 10.07
## 1st Qu.: 23.57 1st Qu.: 11.94 1st Qu.: 12.18 1st Qu.: 29.96
## Median : 60.34 Median : 27.39 Median : 33.12 Median : 50.40
## Mean : 81.76 Mean : 35.56 Mean : 35.67 Mean : 100.93
## 3rd Qu.:113.30 3rd Qu.: 44.26 3rd Qu.: 50.40 3rd Qu.: 89.12
## Max. :539.15 Max. :179.47 Max. :160.77 Max. :2164.62
## cis.Aconitrate myo.Inositol trans.Aconitrate pi.Methylhistidine
## Min. : 12.94 Min. : 11.59 Min. : 4.90 Min. : 11.36
## 1st Qu.: 36.23 1st Qu.: 30.27 1st Qu.: 12.43 1st Qu.: 67.36
## Median : 129.02 Median : 78.26 Median : 26.84 Median : 162.39
## Mean : 204.22 Mean :135.40 Mean : 40.63 Mean : 370.29
## 3rd Qu.: 254.68 3rd Qu.:167.34 3rd Qu.: 57.40 3rd Qu.: 387.61
## Max. :1863.11 Max. :854.06 Max. :217.02 Max. :2697.28
## tau.Methylhistidine
## Min. : 8.00
## 1st Qu.: 27.39
## Median : 68.72
## Mean : 89.69
## 3rd Qu.:130.32
## Max. :317.35

```

```

# Información general
se

```

```

## class: SummarizedExperiment
## dim: 63 77
## metadata():
## assays(1): counts
## rownames(63): X1.6.Anhydro.beta.D.glucose X1.Methylnicotinamide ...
## pi.Methylhistidine tau.Methylhistidine
## rowData names():

```

```
## colnames(77): PIF_178 PIF_087 ... NETL_003_V1 NETL_003_V2
## colData names(1): Muscle.loss
```

```
# Dimensiones
dim(se)
```

```
## [1] 63 77
```

```
# Identificadores de muestras y características
dimnames(se)
```

```
## [[1]]
## [1] "X1.6.Anhydro.beta.D.glucose" "X1.Methylnicotinamide"
## [3] "X2.Aminobutyrate" "X2.Hydroxyisobutyrate"
## [5] "X2.Oxoglutarate" "X3.Aminoisobutyrate"
## [7] "X3.Hydroxybutyrate" "X3.Hydroxyisovalerate"
## [9] "X3.Indoxylsulfate" "X4.Hydroxyphenylacetate"
## [11] "Acetate" "Acetone"
## [13] "Adipate" "Alanine"
## [15] "Asparagine" "Betaine"
## [17] "Carnitine" "Citrate"
## [19] "Creatine" "Creatinine"
## [21] "Dimethylamine" "Ethanolamine"
## [23] "Formate" "Fucose"
## [25] "Fumarate" "Glucose"
## [27] "Glutamine" "Glycine"
## [29] "Glycolate" "Guanidoacetate"
## [31] "Hippurate" "Histidine"
## [33] "Hypoxanthine" "Isoleucine"
## [35] "Lactate" "Leucine"
## [37] "Lysine" "Methylamine"
## [39] "Methylguanidine" "N.N.Dimethylglycine"
## [41] "O.Acetylcarnitine" "Pantothenate"
## [43] "Pyroglutamate" "Pyruvate"
## [45] "Quinolate" "Serine"
## [47] "Succinate" "Sucrose"
## [49] "Tartrate" "Taurine"
## [51] "Threonine" "Trigonelline"
## [53] "Trimethylamine.N.oxide" "Tryptophan"
## [55] "Tyrosine" "Uracil"
## [57] "Valine" "Xylose"
## [59] "cis.Aconitate" "myo.Inositol"
## [61] "trans.Aconitate" "pi.Methylhistidine"
## [63] "tau.Methylhistidine"
##
## [[2]]
## [1] "PIF_178" "PIF_087" "PIF_090" "NETL_005_V1" "PIF_115"
## [6] "PIF_110" "NETL_019_V1" "NETCR_014_V1" "NETCR_014_V2" "PIF_154"
## [11] "NETL_022_V1" "NETL_022_V2" "NETL_008_V1" "PIF_146" "PIF_119"
## [16] "PIF_099" "PIF_162" "PIF_160" "PIF_113" "PIF_143"
## [21] "NETCR_007_V1" "NETCR_007_V2" "PIF_137" "PIF_100" "NETL_004_V1"
## [26] "PIF_094" "PIF_132" "PIF_163" "NETCR_003_V1" "NETL_028_V1"
## [31] "NETL_028_V2" "NETCR_013_V1" "NETL_020_V1" "NETL_020_V2" "PIF_192"
```

```
## [36] "NETCR_012_V1" "NETCR_012_V2" "PIF_089"      "NETCR_002_V1" "PIF_179"
## [41] "PIF_114"      "NETCR_006_V1" "PIF_141"      "NETCR_025_V1" "NETCR_025_V2"
## [46] "NETCR_016_V1" "PIF_116"      "PIF_191"      "PIF_164"      "NETL_013_V1"
## [51] "PIF_188"      "PIF_195"      "NETCR_015_V1" "PIF_102"      "NETL_010_V1"
## [56] "NETL_010_V2"  "NETL_001_V1"  "NETCR_015_V2" "NETCR_005_V1" "PIF_111"
## [61] "PIF_171"      "NETCR_008_V1" "NETCR_008_V2" "NETL_017_V1"  "NETL_017_V2"
## [66] "NETL_002_V1"  "NETL_002_V2"  "PIF_190"      "NETCR_009_V1" "NETCR_009_V2"
## [71] "NETL_007_V1"  "PIF_112"      "NETCR_019_V2" "NETL_012_V1"  "NETL_012_V2"
## [76] "NETL_003_V1"  "NETL_003_V2"
```

```
# Metadatos
colData(se)
```

```
## DataFrame with 77 rows and 1 column
##           Muscle.loss
##           <character>
## PIF_178      cachexic
## PIF_087      cachexic
## PIF_090      cachexic
## NETL_005_V1   cachexic
## PIF_115      cachexic
## ...          ...
## NETCR_019_V2  control
## NETL_012_V1   control
## NETL_012_V2   control
## NETL_003_V1   control
## NETL_003_V2   control
```

```
metadata(se)
```

```
## list()
```

```
# Datos de expresión
data_analysis <- assay(se, "counts")
head(data_analysis)
```

```
##           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
```

Vamos a realizar un ANOVA sobre nuestros datos para determinar si existen diferencias significativas en las características estudiadas entre el grupo control y el grupo enfermo. Vamos a emplear el dataset original, en lugar del contenedor generado, ya que disponemos de todos los datos.

```
table(data$Muscle.loss)
```

```
##
## cachexic control
##      47      30
```

```
ANOVA_result <- list()

# Realizar ANOVA para cada característica
for (col in names(data)[-1]) {
  anova_mod <- aov(data[[col]] ~ Muscle.loss, data = data)
  ANOVA_result[[col]] <- summary(anova_mod)
}

# p-valor < 0.05 = Valores significativos, es decir, que difieren entre ambos grupos de datos
significativos <- list()

# Filtrar por p-value < 0.05
for (p in names(ANOVA_result)) {
  # Extraer el valor p del resumen del ANOVA
  p_value <- ANOVA_result[[p]][[1]][["Pr(>F)"]][1]

  # Verificar si el p-value es menor a 0.05
  if (!is.na(p_value) && p_value < 0.05) {
    significativos[[p]] <- ANOVA_result[[p]]
  }
}

# Tamaño
significativos
```

```
## $X2.Aminobutyrate
##              Df Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss  1   3662    3662   5.058 0.0274 *
## Residuals   75  54293     724
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $X2.Hydroxyisobutyrate
##              Df Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss  1   4324    4324   8.253 0.00529 **
## Residuals   75  39295     524
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $X3.Hydroxybutyrate
```

```

##           Df Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1    6865     6865  11.37 0.00119 **
## Residuals   75   45300      604
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $X3.Hydroxyisovalerate
##           Df Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1    4283     4283   7.468 0.00783 **
## Residuals   75   43012      573
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $X3.Indoxylsulfate
##           Df Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1  258360  258360   7.211 0.00892 **
## Residuals   75  2687195   35829
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Acetate
##           Df Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1   45831   45831   7.974 0.00607 **
## Residuals   75  431041     5747
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Adipate
##           Df Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1   12212   12212   5.059 0.0274 *
## Residuals   75  181046     2414
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Alanine
##           Df Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1  661101  661101  11.38 0.00118 **
## Residuals   75  4358038    58107
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Asparagine
##           Df Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1   20725   20725   7.753 0.00679 **
## Residuals   75  200493     2673
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Betaine
##           Df Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1   58007   58007   9.416 0.00299 **
## Residuals   75  462035     6160
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

##
## $Citrate
##           Df      Sum Sq  Mean Sq F value Pr(>F)
## Muscle.loss 1  28435309 28435309   6.496 0.0129 *
## Residuals  75 328309888  4377465
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Creatinine
##           Df      Sum Sq  Mean Sq F value  Pr(>F)
## Muscle.loss 1 4.768e+08 476841076   13.19 0.000513 ***
## Residuals  75 2.712e+09  36161105
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Dimethylamine
##           Df  Sum Sq Mean Sq F value  Pr(>F)
## Muscle.loss 1 1098238 1098238   13.5 0.000446 ***
## Residuals  75 6103134   81375
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Ethanolamine
##           Df  Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1  307788  307788   5.118 0.0266 *
## Residuals  75 4510111   60135
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Formate
##           Df  Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1  194575  194575   5.912 0.0174 *
## Residuals  75 2468219   32910
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Fucose
##           Df Sum Sq Mean Sq F value  Pr(>F)
## Muscle.loss 1  47918   47918   8.048 0.00585 **
## Residuals  75 446523   5954
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Glucose
##           Df      Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1   8623964 8623964   4.702 0.0333 *
## Residuals  75 137569910 1834265
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Glutamine
##           Df  Sum Sq Mean Sq F value  Pr(>F)
## Muscle.loss 1   862145  862145   11.6 0.00106 **
## Residuals  75 5572369   74298

```

```

## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Glycine
##           Df    Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1  4293675 4293675   5.012 0.0281 *
## Residuals   75 64251986  856693
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Hippurate
##           Df    Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1 41834878 41834878   5.37 0.0232 *
## Residuals   75 584307994  7790773
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Histidine
##           Df    Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1  618344  618344   6.804  0.011 *
## Residuals   75 6815737   90876
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Leucine
##           Df Sum Sq Mean Sq F value  Pr(>F)
## Muscle.loss 1   5740    5740   14.62 0.00027 ***
## Residuals   75  29442    393
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Methylamine
##           Df Sum Sq Mean Sq F value  Pr(>F)
## Muscle.loss 1   1779   1778.9   10.31 0.00195 **
## Residuals   75  12937   172.5
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $N.N.Dimethylglycine
##           Df Sum Sq Mean Sq F value  Pr(>F)
## Muscle.loss 1   7993    7993   15.87 0.000155 ***
## Residuals   75  37766    504
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Pyroglutamate
##           Df Sum Sq Mean Sq F value  Pr(>F)
## Muscle.loss 1 417715 417715   13.31 0.000485 ***
## Residuals   75 2353312  31377
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Pyruvate
##           Df Sum Sq Mean Sq F value Pr(>F)

```

```

## Muscle.loss 1 3744 3744 5.91 0.0174 *
## Residuals 75 47517 634
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Quinolate
## Df Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1 36137 36137 16.5 0.000119 ***
## Residuals 75 164259 2190
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Serine
## Df Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1 279597 279597 8.947 0.00376 **
## Residuals 75 2343683 31249
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Succinate
## Df Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1 45401 45401 6.733 0.0114 *
## Residuals 75 505708 6743
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Taurine
## Df Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1 2057455 2057455 4.755 0.0323 *
## Residuals 75 32452028 432694
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Threonine
## Df Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1 63128 63128 9.222 0.00329 **
## Residuals 75 513408 6845
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Trigonelline
## Df Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1 959869 959869 6.495 0.0129 *
## Residuals 75 11084041 147787
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Trimethylamine.N.oxide
## Df Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1 3412207 3412207 4.298 0.0416 *
## Residuals 75 59544091 793921
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##

```

```

## $Tryptophan
##           Df Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1  29286   29286   10.37 0.0019 **
## Residuals   75 211896    2825
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Tyrosine
##           Df Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1  43480   43480    6.748 0.0113 *
## Residuals   75 483223    6443
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Valine
##           Df Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1  11860   11860   16.12 0.000139 ***
## Residuals   75  55167    736
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $cis.Aconitate
##           Df Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1 621994  621994    8.873 0.0039 **
## Residuals   75 5257591   70101
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $myo.Inositol
##           Df Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1  260168  260168   10.04 0.00222 **
## Residuals   75 1943123   25908
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $trans.Aconitate
##           Df Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1   8079    8079    5.464 0.0221 *
## Residuals   75 110901    1479
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $tau.Methylhistidine
##           Df Sum Sq Mean Sq F value Pr(>F)
## Muscle.loss 1  30808   30808    5.468 0.022 *
## Residuals   75 422598    5635
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

De las 63 características que se han estudiado, de acuerdo con nuestro ANOVA, un total de 40 son significativas, es decir que difieren entre grupos. A partir de aquí, podríamos dividir el dataset y extraer los datos significativos.



```
features_sig <- names(significativos)
new_data <- data[, c("Muscle.loss", features_sig)]
```

## Exploración de perfil en GitHub

Ya que no tenía cuenta en GitHub me he creado una. He creado un repositorio con el nombre Sanchez-Martin-Maria-PEC1 con la idea de subir los archivos que se especifican en el enunciado de la PEC1.

Guardar el objeto contenedor con los datos y metadatos:

```
save(se, file = "Sanchez-Martin-Maria-PEC1-SummarizedExperiment.Rda")
```

He ido generando los archivos necesarios de acuerdo a lo que se exige en el enunciado de la PEC1:

- Informe: Sanchez-Martin-Maria-PEC1.Rmd, Sanchez-Martin-Maria-PEC1.pdf, Sanchez-Martin-Maria-PEC1.html.
- El código R para la exploración de los datos: Sanchez-Martin-Maria-PEC1-RScript.R
- Los metadatos acerca del dataset: Sanchez-Martin-Maria-PEC1-README.Rmd
- El objeto contenedor con los datos y los metadatos en formato binario: Sanchez-Martin-Maria-PEC1-SummarizedExperiment.Rda
- Los datos en formato texto: human\_cachexia.csv

Enlace GitHub: <https://github.com/mariasanchez13/Sanchez-Martin-Maria-PEC1.git>