

# Análisis de datos ómicos-Primera prueba de evaluación continua.

Francisco Javier García Barrera

2025-04-02

## Table of Contents

<b>Resumen</b> .....	1
<b>Objetivos</b> .....	1
<b>Métodos</b> .....	1
<b>Resultados</b> .....	2
<b>Discusión</b> .....	7
<b>Conclusión</b> .....	7
<b>Referencias</b> .....	7

## Resumen

En este trabajo se ha realizado un objeto de clase SummarizedExperiment y un análisis exploratorio de los datos extraídos del repositorio GitHub “metaboData”. Dichos datos son originarios del artículo científico “Metabotypes of response to bariatric surgery independent of the magnitude of weight loss”, en el cual se hace un estudio de la evolución de ciertos parámetros metabólicos tras el sometimiento de pacientes hombres y mujeres de distinta edad a una cirugía bariátrica.

Se observa que, tras un periodo de unos 5 meses, hay una disminución de los niveles de ciertos parámetros tales como el peso, triglicéridos, colesterol total y glucosa en sangre.

## Objetivos

El objetivo de este trabajo es la creación de un objeto de clase SummarizedExperiments y un análisis exploratorio de los datos extraídos del repositorio de GitHub metaboData.

## Métodos

Los datos a analizar se encuentran en el archivo Metabotypingpaper en el repositorio metaboData de GitHub, dichos datos pertenecen a un artículo científico llamado “Metabotypes of response to bariatric surgery independent of the magnitude of

weight loss” (<https://doi.org/10.1371/journal.pone.0198214>). En dicho trabajo se realizó un estudio de parámetros metabólicos, además de medición del peso, de pacientes de ambos géneros y distintas edades que se sometieron a una cirugía bariátrica para el tratamiento de la obesidad mórbida en un periodo de 5-6 meses.

Para extraer los datos a analizar, se vinculó un proyecto de R con GitHub y a continuación se clonaron del repositorio donde se encontraban guardándose después en el archivo del proyecto de esta PEC, para esto último se usó el comando “git clone” en el terminal de R seguido de la dirección url del repositorio (<https://github.com/nutrimetabolomics/metaboData.git>).

Una vez obtenidos los datos del repositorio, se cargaron los datos y metadatos deseados (DataValues y DataInfo del archivo Metabotypingpaper) y, tras observar la composición y el contenido de estos, se crea un objeto clase SummarizedExperiments, para ello se usó la función y paquete con el mismo nombre.

Después se realizó un análisis exploratorio. Debido a la gran cantidad de datos existente y el espacio limitado que se dispone para este trabajo, se hace análisis exploratorio de parte de los datos: género, tipo de cirugía, peso y niveles de triglicéridos, colesterol total y glucosa (los tres últimos en distintos momentos). Se crearon gráficas de líneas donde se comparaban las evoluciones de los parámetros a lo largo del tiempo según el género de los pacientes y el tipo de cirugía. Para la creación de las gráficas, se hizo la media de cada variable, de esta forma se amortiguan los valores extremos.

Para la realización de este trabajo y los análisis se utilizó el software RStudio y los paquetes “readr” y “SummarizedExperiment”.

Por último, todos los análisis, resultados e informe desarrollados y obtenidos se suben a un repositorio creado específicamente para la recepción de este trabajo y posterior evaluación (<https://github.com/fragarbar/Garcia-Barrera-Francisco-Javier-PEC1.git>).

## Resultados

```
## New names:
## Rows: 39 Columns: 696
## — Column specification
## _____ Delimiter:
## ", " chr
## (2): SURGERY, GENDER dbl (693): ...1, SUBJECTS, AGE, Group, MEDDM_T0,
## MEDCOL_T0, MEDINF_T0, MEDHT... lgl (1): X
## i Use `spec()` to retrieve the full column specification for this
data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this
message.
## New names:
```

```
## Rows: 695 Columns: 4
## — Column specification
## _____ Delimiter:
## ", " chr
## (4): ...1, VarName, varTpe, Description
## i Use `spec()` to retrieve the full column specification for this
data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this
message.
## • `` -> `...1`

## # A tibble: 6 × 696
##   ...1 SUBJECTS SURGERY AGE GENDER Group MEDDM_T0 MEDCOL_T0
MEDINF_T0
##   <dbl>   <dbl> <chr>   <dbl> <chr>   <dbl>   <dbl>   <dbl>
<dbl>
## 1     1       1 1 by pass  27 F       1       0       0
0
## 2     2       2 2 by pass  19 F       2       0       0
0
## 3     3       3 3 by pass  42 F       1       0       0
0
## 4     4       4 4 by pass  37 F       2       0       0
0
## 5     5       5 5 tubular  42 F       1       0       0
0
## 6     6       6 6 by pass  24 F       2       0       0
0
## # i 687 more variables: MEDHTA_T0 <dbl>, GLU_T0 <dbl>, INS_T0 <dbl>,
## #   HOMA_T0 <dbl>, HBA1C_T0 <dbl>, HBA1C.mmol.mol_T0 <dbl>, PESO_T0
<dbl>,
## #   bmi_T0 <dbl>, CC_T0 <dbl>, CINT_T0 <dbl>, CAD_T0 <dbl>, TAD_T0
<dbl>,
## #   TAS_T0 <dbl>, TG_T0 <dbl>, COL_T0 <dbl>, LDL_T0 <dbl>, HDL_T0
<dbl>,
## #   VLDL_T0 <dbl>, PCR_T0 <dbl>, LEP_T0 <dbl>, ADIPO_T0 <dbl>, GOT_T0
<dbl>,
## #   GPT_T0 <dbl>, GGT_T0 <dbl>, URICO_T0 <dbl>, CREAT_T0 <dbl>,
UREA_T0 <dbl>,
## #   HIERRO_T0 <dbl>, TRANSF_T0 <dbl>, FERR_T0 <dbl>, Ile_T0 <dbl>, ...
```

Podemos observar parte de los datos a analizar, las variables se encuentran en las columnas y cada sujeto de estudio se encuentran en filas. También podemos observar la cantidad de variables que hay en el fichero (696),m dicha cantidad de datos se debe en parte a que se tomaron muestras en distintos momentos a lo largo de los 6 meses.

```
## # A tibble: 6 × 4
##   ...1 VarName varTpe Description
##   <chr> <chr>   <chr>   <chr>
```

```
## 1 SUBJECTS SUBJECTS integer dataDesc
## 2 SURGERY SURGERY character dataDesc
## 3 AGE AGE integer dataDesc
## 4 GENDER GENDER character dataDesc
## 5 Group Group integer dataDesc
## 6 MEDDM_T0 MEDDM_T0 integer dataDesc
```

Al igual que ocurre con los datos, podemos observar parte de los metadatos, en el que podemos ver las variables que hay en el fichero de datos y una pequeña descripción de cada variable.

Lo siguiente que se muestra son los resultados obtenidos de la creación del objeto clase SummarizedExperiment.

```
## class: SummarizedExperiment
## dim: 39 695
## metadata():
## assays(1): counts
## rownames: NULL
## rowData names(0):
## colnames(695): SUBJECTS SURGERY ... SM.C24.0_T5 SM.C24.1_T5
## colData names(4): ...1 VarName varTpe Description
```

Se puede observar el resultado del análisis en el que se describe:

- La clase: SummarizedExperiment
- Las dimensiones: 39 filas y 695 columnas
- Metadatos: no hay datos
- Ensayos: counts o conteos
- Nombre de las filas: no se han definido nombre de filas
- Datos de las filas: no hay nombre en los datos de las filas
- Nombre de las columnas: hay nombres definidos que van desde Subjects hasta SM.C24.1\_T5
- Datos de columnas: hay 4 variables y sus nombres son ...1, Varname, Vartype y Description

Una vez obtenido SummarizedExperiments y se extrajesen los datos mencionados en el apartado Métodos, se muestra un resumen estadístico de las variables que se analizaron.

##	GENDER	SURGERY	PESO_T0	PESO_T2
##	Length:39	Length:39	Min. : 84.0	Min. : 75.7
##	Class :character	Class :character	1st Qu.:119.5	1st Qu.:110.0
##	Mode :character	Mode :character	Median :135.0	Median :120.0
##			Mean :140.0	Mean :124.5
##			3rd Qu.:155.0	3rd Qu.:140.0
##			Max. :200.0	Max. :176.0
##				NA's :2

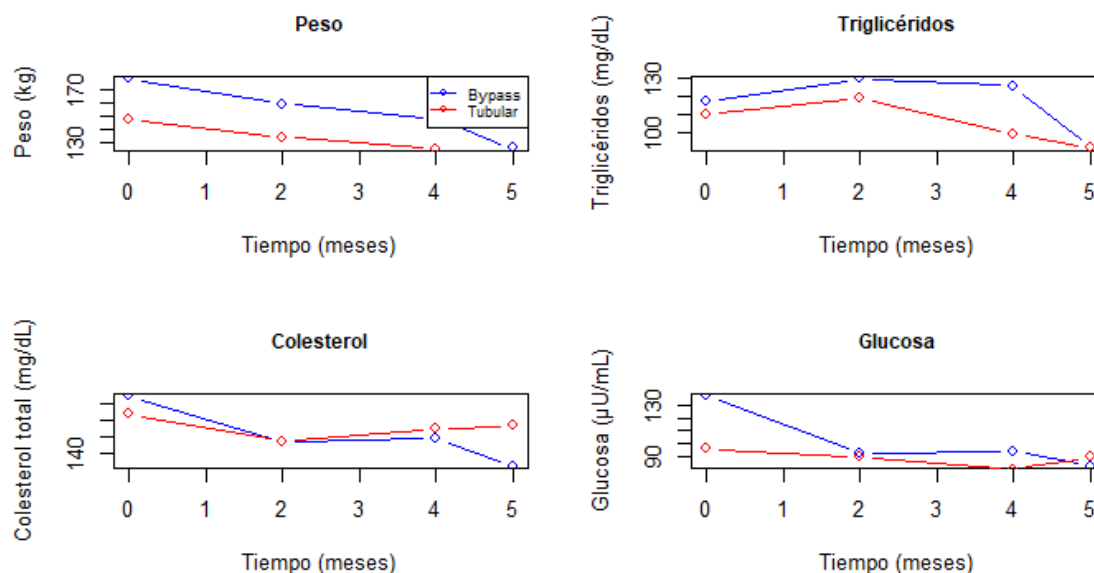
##	PESO_T4	PESO_T5	TG_T0	TG_T2
##	Min. : 74.50	Min. : 64.00	Min. : 30.0	Min. : 52.0
##	1st Qu.: 99.83	1st Qu.: 88.42	1st Qu.: 84.0	1st Qu.:100.0
##	Median :109.50	Median :100.00	Median :118.0	Median :124.0
##	Mean :113.99	Mean :100.41	Mean :129.1	Mean :121.9
##	3rd Qu.:126.75	3rd Qu.:112.25	3rd Qu.:151.0	3rd Qu.:136.5
##	Max. :157.00	Max. :142.00	Max. :268.0	Max. :215.0
##	NA's :1	NA's :7		NA's :3
##	TG_T4	TG_T5	COL_T0	COL_T2
##	Min. : 39.00	Min. : 37.00	Min. :102.0	Min. : 93.0
##	1st Qu.: 80.25	1st Qu.: 72.75	1st Qu.:180.5	1st Qu.:136.5
##	Median :110.50	Median : 87.00	Median :207.0	Median :163.0
##	Mean :109.82	Mean : 95.56	Mean :206.6	Mean :163.8
##	3rd Qu.:136.75	3rd Qu.:110.50	3rd Qu.:236.0	3rd Qu.:181.2
##	Max. :223.00	Max. :227.00	Max. :295.0	Max. :248.0
##	NA's :1	NA's :7		NA's :3
##	COL_T4	COL_T5	GLU_T0	GLU_T2
##	Min. :107.0	Min. : 91.0	Min. : 71.0	Min. : 66.00
##	1st Qu.:143.8	1st Qu.:132.2	1st Qu.: 91.0	1st Qu.: 82.50
##	Median :166.5	Median :170.0	Median :103.0	Median : 89.00
##	Mean :170.1	Mean :169.7	Mean :106.5	Mean : 89.17
##	3rd Qu.:191.2	3rd Qu.:205.5	3rd Qu.:109.5	3rd Qu.: 96.25
##	Max. :265.0	Max. :257.0	Max. :263.0	Max. :117.00
##	NA's :1	NA's :7		NA's :3
##	GLU_T4	GLU_T5		
##	Min. : 38.00	Min. : 67.00		
##	1st Qu.: 78.25	1st Qu.: 74.75		
##	Median : 83.50	Median : 84.00		
##	Mean : 83.76	Mean : 84.31		
##	3rd Qu.: 91.75	3rd Qu.: 94.25		
##	Max. :108.00	Max. :105.00		
##	NA's :1	NA's :7		

Como se puede observar en el resumen, hay 19 variables en los datos que se van a analizar (aunque la mayoría son las mismas variables repetidas en diferentes etapas del estudio).

Gender y Surgery son variables de tipo carácter ya que contienen información del género y tipo de cirugía del paciente. Se puede apreciar que, según los valores mínimos y máximos, hay ciertas variables que tienen rangos amplios, puede que sean outliers, es debido a esto por lo que se usaron las medias para amortiguar este tipo de datos al realizar las gráficas. Por último, mencionar que hay ciertas variables que contienen valores “NA’S”, son valores vacíos que se tuvieron en cuenta al realizar las medias de cada variable para su posterior uso en las gráficas.

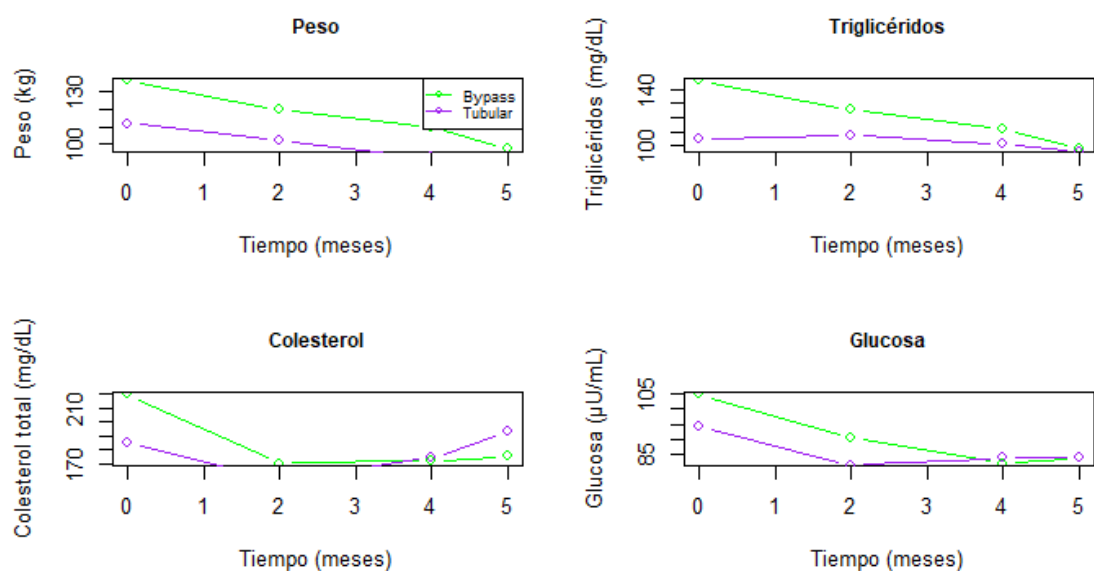
Lo siguiente que se muestra es la evolución en el tiempo de las variables según el tipo de cirugía y agrupadas por género.

## Evolución de niveles promedio de parámetros en hombres tras cirugía bariátrica



Se puede observar en las cuatro gráficas que, en términos generales, los parámetros de estudio van disminuyendo los niveles a lo largo del tiempo al igual que el peso de los pacientes. Los pacientes a los que se le hicieron el by pass empezaron de media con más peso y niveles más altos que aquellos a los que se les hicieron la cirugía tubular, aunque al final del estudio todos los pacientes llegaron a tener pesos y niveles similares, a excepción del colesterol total.

## Evolución de niveles promedio de parámetros en mujeres tras cirugía bariátrica



En estas gráficas se pueden ver patrones parecidos a las observadas en los gráficos de hombres, en términos generales los niveles de los parámetros y el peso de las pacientes descienden con el paso del tiempo, las pacientes con cirugía by pass comenzaron con mayor peso y niveles de media que las de cirugía tubular aunque al final acabaron con pesos y parámetros similares. Hay que destacar el hecho de que, en el caso del colesterol total, las pacientes con cirugía tubular acabaron teniendo mayores niveles de colesterol que cuando se comenzó el estudio.

## Discusión

En este trabajo se ha realizado un objeto de clase SummarizedExperiment, esta es una extensión de otra clase lalamda ExpressionSet. Las diferencias entre SummarizedExperiment y ExpressionSet son que, aunque en ambos se manejan gran cantidad de datos, SummarizedExperiment permite asociar rangos genómicos, es más flexible con la información de las filas y columnas y permite el manejo de varias matrices; mientras que ExpressionSet no es tan flexible con la información de los datos y sólo permite el manejo de una matriz.

Tanto en las gráficas como lo mencionado en el resumen estadístico, se observa una reducción de los niveles de todos los parámetros que se han analizado, dando como resultado también una pérdida de peso. Esto tiene sentido si se conoce un poco en qué consiste las cirugías realizadas, en ambas hay una reducción del tamaño del estómago por lo que se reduce la cantidad de comida que se ingiere, además, en el by pass se conecta parte del estómago con parte del intestino delgado disminuyendo también la absorción de nutrientes por evitar parte del intestino.

Es lógico pensar que, al disminuir la cantidad de alimento ingestado y/o la disminución de nutrientes, vayan bajando los niveles de azúcares y grasas en sangre, se está generando una condición de déficit calórico en la cual el cuerpo usa las reservas que posee para compensar el déficit y eso provoca la reducción de peso en última instancia. Esto además, acabaría también generando un mayor control de glucosa en sangre y puede mejorar la situación de pacientes diabéticos, sobretodo los tipo II.

## Conclusión

Se ha realizado un estudio con una temática de gran interés por la situación occidental actual de parte de la población con obesidad mórbida en la cual se puede observar los efectos de estas dos cirugías en la que se puede concluir que los pacientes que se someten a dichos tratamientos mejoran sus condiciones metabólicas.

## Referencias

Alexsanchezpla. Repositorio metaboData. GitHub.  
<https://github.com/nutrimetabolomics/metaboData.git>.

García-Barrera-Francisco-Javier. Repositorio GitHub.

<https://github.com/fragarbar/Garcia-Barrera-Francisco-Javier-PEC1.git>.

Palau-Rodriguez M, Tulipani S, Marco-Ramell A, Miñarro A, Jáuregui O, et al. (2018) Metabotypes of response to bariatric surgery independent of the magnitude of weight loss. PLOS ONE 13(6): e0198214. <https://doi.org/10.1371/journal.pone.0198214>.

Morgan M, Obenchain V, Hester J, Pagès H. (2023). SummarizedExperiment for Coordinating Experimental Assays, Samples, and Regions of Interest. Bioconductor.org.

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

2025. The SummarizedExperiment class. The Bioconductor project.

<https://carpentries-incubator.github.io/bioc-project/09-summarizedexperiment.html>.

ExpressionSet and SummarizedExperiment. STHDA.

<https://www.sthda.com/english/wiki/expressionset-and-summarizedexperiment>.

La gastrectomía tubular, cirugía bariátrica para tratar la obesidad. Ilahy.

<https://www.ilahy.es/cirugia/gastrectomia-tubular-obesidad.html>.

El bypass gástrico, cirugía bariátrica para tratar la obesidad. Ilahy.

<https://www.ilahy.es/cirugia/bypass-gastrico-obesidad.html>.

Gráfico de líneas en R. RCoder. <https://r-coder.com/grafico-lineas-r/>.

Márgenes en R. RCharts. <https://r-charts.com/es/r-base/margenes/>.

Función par. GitHub. <https://fhernanb.github.io/Graficos-con-R/par.html>.

Añadir textos en R con las funciones text y mtext. RCharts. <https://r-charts.com/es/r-base/textos/>.

Save: Save R Objects. RDocumentation.

<https://www.rdocumentation.org/packages/base/versions/3.6.2/topics/save>.