

PEC 1.

1. Descarga de los datos.

He seleccionado el dataset GastricCancer_NMR contenido en 2023-CIMBTutorial, procedente del repositorio

<https://github.com/nutrimetabolomics/metaboData/>, y lo he importado a R.

2. Creación del contenedor.

Para crear el contenedor con los datos he tenido que instalar y cargar el paquete SummarizedExperiment de BiocManager, y he cargado también el paquete readxl para abrir mi fichero en R mediante la función read_excel. Tras esto, he definido los metadatos, los datos de expresión y los datos row o variables, que serían contenidos en mi contenedor de datos.

3. Exploración de los datos.

En primer lugar, he realizado un resumen de los datos mediante la función summary, para tener una visión general de la variabilidad estadística de los metabolitos en las diferentes muestras. Tras esto, he instalado y cargado el paquete POMA de BiocManager, y para llevar a cabo este análisis fue necesario realizar la sustitución de los valores faltantes por valores estimados realizando una imputación. Una vez tuvimos nuestro contenedor de datos imputado, realicé una normalización y los gráficos previo y posterior a esta normalización. Las funciones que he utilizado para realizar los gráficos en nuestro análisis de los datos fueron PomaBoxplots y PomaDensity.

4. Reposición de los datos en GitHub.

Se puede acceder a mi trabajo sobre el análisis del dataset GastricCancer_NMR en GitHub haciendo click en el siguiente enlace: https://github.com/MrlSara/Moron_El-Khalifi_Sara