

Informe de Análisis de Datos de Fosforilación

PEC 1- Inés Agustí Sunyer

1. Introducción

Este informe presenta un análisis de un conjunto de datos de fosforilación de proteínas, evaluando la diferencia en las intensidades de fosforilación entre dos grupos: MSS (Muestra de Control) y PD (Muestra de Enfermedad).

A continuación, se describen los pasos realizados en el análisis, junto con los resultados obtenidos.

2. Descripción del Dataset

El conjunto de datos incluye un total de 1438 proteínas medidas en 10 muestras. Los datos se organizaron en un objeto `SummarizedExperiment`, el cual permite realizar análisis más eficientes en datos ómicos. La tabla a continuación resume las dimensiones del objeto:

- Número de Proteínas (Filas): 1438
- Número de Muestras (Columnas): 10
- Características de las Proteínas: 6 (incluyendo modificaciones de secuencia, accesiones, descripciones, puntajes, clase y fosforilación)

Los datos de intensidades de fosforilación para cada muestra son los siguientes:

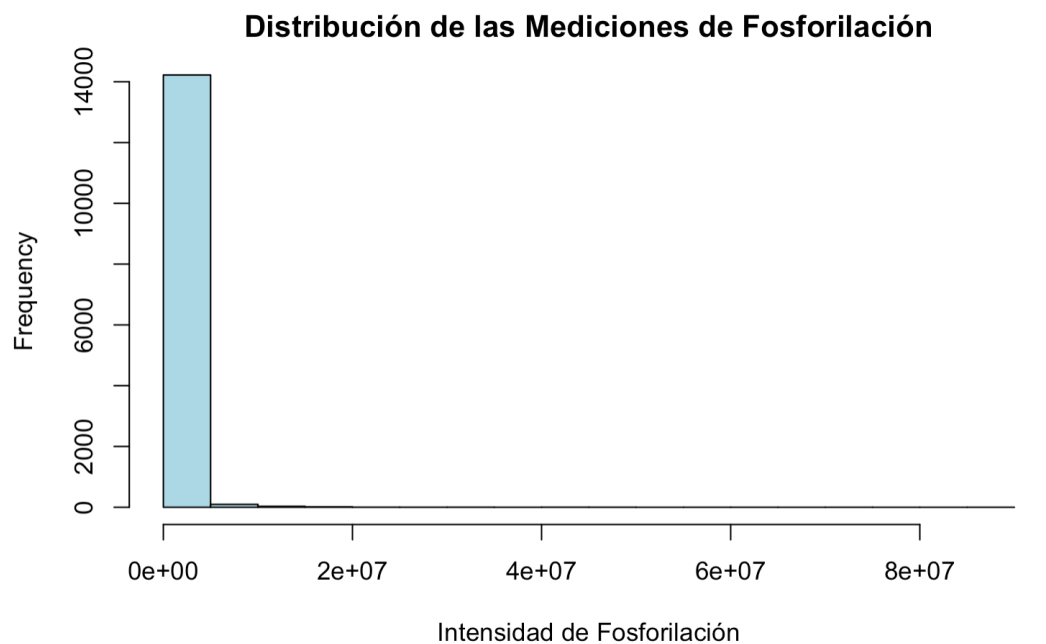
| Muestra | Intensidad |
|----------|------------|
| M1_1_MSS | 24.29438 |
| M1_2_MSS | 44475.964 |
| M5_1_MSS | 0.000 |
| M5_2_MSS | 6269.141 |
| M42_1_PD | 0.000 |
| M42_2_PD | 0.000 |
| M43_1_PD | 772.9056 |
| M43_2_PD | 2136.746 |
| M64_1_PD | 1820.724 |
| M64_2_PD | 1727.9098 |

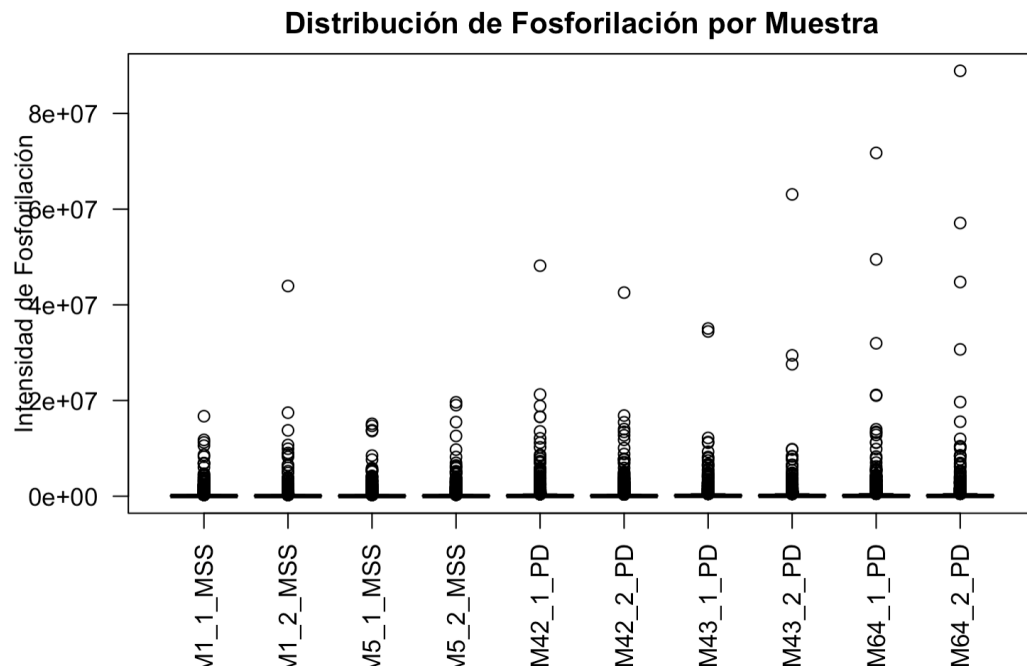
3. Análisis de Datos

Exploración Inicial

Se realizó un análisis descriptivo de los datos, calculando estadísticas resumidas como el mínimo, cuartiles, mediana, media y máximo para las intensidades de cada grupo.

- Grupo MSS:
 - Mínimo: 0
 - 1er Cuartil: 5963
 - Mediana: 29314
 - Media: 244256
 - 3er Cuartil: 125089
 - Máximo: 20378233
- Grupo PD:
 - Mínimo: 0
 - 1er Cuartil: 17355
 - Mediana: 61530
 - Media: 397537
 - 3er Cuartil: 213708
 - Máximo: 47035198





Comprobación de Valores Faltantes

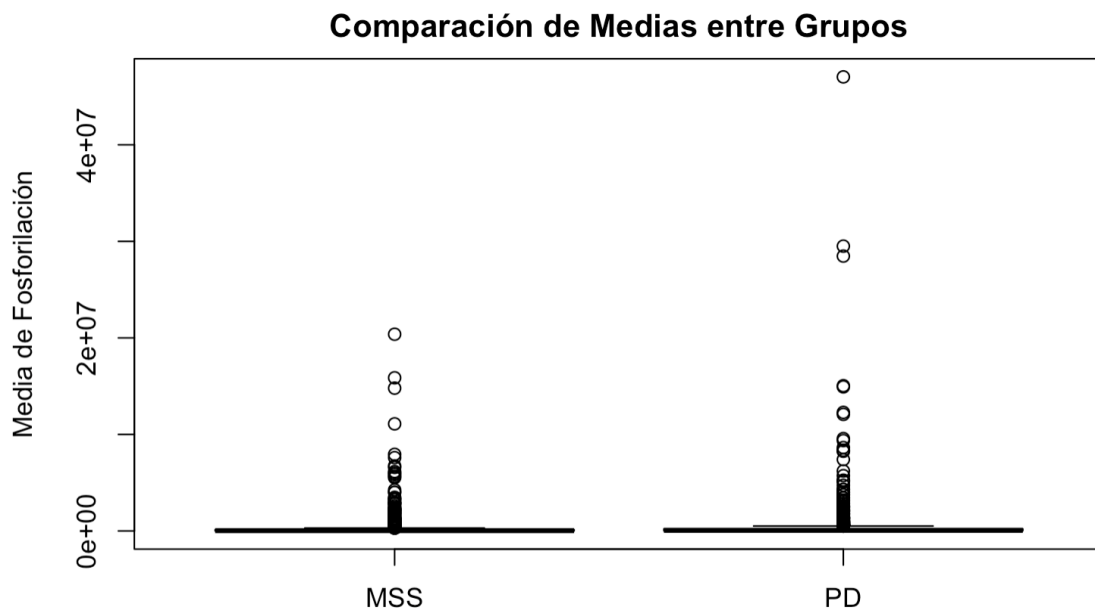
Se verificó la existencia de valores faltantes en el dataset, obteniendo un total de 0 valores NA, lo que indica que no hay datos ausentes en el conjunto analizado.

Comparación de Grupos

Se realizó una prueba t de Welch para comparar las medias de las intensidades de fosforilación entre los grupos MSS y PD.

- Resultados de la prueba t:
 - Estadístico t: -2.6129
 - Grados de libertad: 2205.1
 - p-valor: 0.009038
 - Intervalo de confianza del 95% para la diferencia de medias: [-268319.27, -38241.01]
 - Media del grupo MSS: 244256.5
 - Media del grupo PD: 397536.6

Estos resultados sugieren una diferencia estadísticamente significativa en las medias de los dos grupos, con un p-valor menor a 0.01.



4. Conclusiones

El análisis revela diferencias significativas en las intensidades de fosforilación entre los grupos MSS y PD. Estos hallazgos pueden tener implicaciones importantes para entender los mecanismos de fosforilación en el contexto de la enfermedad analizada.

Dirección URL repositorio Github: <https://github.com/inesagusti/AGUSTI-SUNYER-INES-PEC1>