Ha llegado una de las prácticas más comunes.

Comparar medidas (var. cuantitativas) por factores. Es muy común en prácticas de investigaciones como por ejemplo casos control e investigación.

Siempre que queramos comprar una medida (var. cuantitativa) por grupos de estudio estos serán los pasos que tienes que seguir.

Aquí tienes bases de datos por si quieres practicar aunque te recomiendo practicar primero con la tabla de datos “espalda.xlsx”

[>> Repositorio de datasets muy muy interesantes](https://archive.ics.uci.edu/ml/datasets.html)

# Abre la tabla de datos “espalda”

Si no quieres pensar en base de datos vamos a seguir trabajando con la tabla de datos espalda.

Más que nada para seguir un orden en la aplicación de la metodología ☺

Abre la tabla de datos “espalda.xlsx” y crea una nueva variable llamada diff\_ODI. Calcula esta variable de esta forma:

* diff\_ODI = ODI\_mes0 – ODI\_mes1

Esta variable será reflejo de la mejora de los pacientes.

Recuerda que ODI es sinónimo de “lo mal” que está el paciente por el dolor de espalda. Y con diff\_ODI calculamos la mejora de este paciente.

Construye una nueva variable llamada LC\_2. Esta variable la puedes obtener a partir de la variable LC. Recuerda que los pacientes con LC son los correspondientes al 1 o un 2..

Deberías juntar los niveles 1 y 2 de LC en un solo grupo llamado 1. Entonces la LC\_2 tendría:

* 0 = NO LC 🡪 que correspondes al 0 de la variable LC
* 1 = SI LC 🡪 que corresponden al juntar los grupos 1 y 2 de la variable LC

¿Se ve?

## Escogemos la medida o variable dependiente

Vamos a seleccionar una medida que nos interese estudiar.

En esta base de datos la medida crítica está claro que es el índice de discapacidad:

* ODI mes 0
* ODI mes 1
* Diff\_ODI

Yo te recomiendo que lo hagas con la diff\_ODI, aunque puedes trabajar con el ODI\_mes1 (para cambiar)

## Escogemos los factores de estudio o variables independientes

* Queremos ver las diferencias entre el factor grupo (investigación o control).
* El factor NHD (número de hernias discales)
* Otra variable interesante podría ser si tienen LC o no. Es decir el factor LC\_2 que has creado.

Escoge dos de estos tres.

## Describe DIFF\_ODI vs Grupo

|  |
| --- |
| Copia:   * El resumen numérico por factor * El boxplot o histograma por factor * El diagrama de error por factor   Resumen Numérico variable diff\_odi por LC    A priori hay diferencias en diff\_odi entre los LC 0 y LC 1 pues los IConfianza no se traslapan  Resumen Numerico diff\_Oddi por grupo control. Como se observa también hay aparente diferencias entre el Grupo con tratamiento convencional (0) y tratamiento experimental (1)    BoxPlot para ver gráficamente las distribuciones de diff\_oddi por grupo y por LC y Hernia Discal    Por Hernia discales no parece haber diferencias entre los grupos respecto a mejoría posiblemente sin hernia discal puede que sea ligeramente diferente pero igual se solapa    Aquí si se observa claramente que hay diferencias claras en mejoría entre el grupo con tratamiento convencional (0) y el grupo con tratamiento novedoso(1)    No observo claras diferencias en mejoría por los LC 0 y LC 1. A nivel de medias si hay diferencias pero la dispersión que muestra LC1 hace dudar .    En efecto, claramente el diagrama de Medias o error muestra que para el factor de Hernias discales no hay grandes diferencias en la mejoria    Tal como mostro el box plot de mejoría por Tipo de tratamiento, el diagrama de medias muestra claramente estas diferencias.  En efecto el diagrama de medias no muestra grandes diferencias en mejoría cuando se agrupo por LC    MTcars, Millas por galon agrupado por gear. Histograma de densidad |
|  |

## ¿Qué conclusiones obtienes?

Identifica las diferencias entre los dos grupos (control e investigación).

* Fíjate en el IC de la media
* En el diagrama de error
* En el boxplot y las cajas

|  |
| --- |
| Lista las conclusiones que has podido observar con los índices  Mis conclusiones están arriba en cada grafica y resumen numerico |

## Describe DIFF\_ODI vs Grupo vs NHD o LC\_2

Ahora podemos describir la interacción de los dos factores:

* Grupo vs NHD
* O puedes hacerlo con Grupo vs LC\_2

El que prefieras. De esta manera vemos cómo curamos los pacientes si tienes ciática o bien tienen hernias discales.

|  |
| --- |
| Copia:   * El resumen numérico por los dos factor (si puedes) * El boxplot por los dos factores * El diagrama de error por los dos factor   Box Plot Mejoria agrupada por Grupo Tratamiento y No de Hernias Discales    En definitiva, el grupo que mejor se comporta es el Grupo 1 (tratamiento avanzado). Dentro de este lo que tienen 1 o 2 Hernias presentan grandes mejorías por encima de los que no tienen Hernias Discales.  En el grupo 0 (tratamiento convencional) Es evidente que mejoran menos que el grupo experimental, sin embargo, dentro de la poca mejoría vemos que igual los que tiene al menos una (1) hernia mejora mas que los que no tienen.  El error diagram muestra que en definitiva los grupos tienen diferencias. Solo en el grupo 1 (tratamiento avanzado) parece que cuando que el paciente con una (1) o mas hernias parecen mejorar mas que los que no tienen ninguna hernia.  No podemos apreciar en el grupo 0(tratamiento convencional) parece que entre los pacientes con 0,1 y 2 hernias no muestran diferencias.  Box Plot Mejoria (diff\_oddi) por Grupo tratamiento (0 Convencional y 1 Innovador ) y Tipo de Lumbociatica    Vemos que para el grupo de tratamiento innovador (1) hay evidente mejorías frente al grupo de tratamiento convencional(0). En este grupo innovador es mucho mas evidente la mejoría entre los pacientes con algún tipo de LumboCiatica que los que no tienen.  En el grupo conencional al parecer mejorn lo de LCiatica Derecha mas que los otros. Esto hay que seguir investigando puede ser que sean muy pocos datos.  Diagrama de Error    En efecto el Diagrama de error muestra que el grupo experimental y dentro de este los que tienen algún tipo de Lumbociatica presentan grandes diferencias con los que no tienen |

## ¿Qué conclusiones obtienes?

Identifica las diferencias por la interacción de los dos factores

* Fíjate en el IC de la media por los dos factores
* En el diagrama de error por los dos factores
* En el boxplot y las cajas

Tómate tu tiempo para entender los gráficos. Son sencillos pero a veces cuesta observar los dos factores a la vez.

|  |
| --- |
| Lista las conclusiones que has podido observar con los índices |

## Repite este procedimiento para los factores que creas importantes:

Recuerda:

* Identifica la medida que quieres analizar
* Identifica unos o dos factores de estudios que pueden influir en la medida
* Aplica las herramientas descriptivas:
  + Resumen numéricos por factores
  + Boxplot por factores
  + Diagramas de error por factores

¡Con este proceso tienes más que ganada la partida! Sólo te queda el paso de la demostración con un valor de probabilidad con la estadística inferencial ☺

¡Bien!

¡Lo puedes implementar en cualquier contexto!