

Desafío - Estimando curvas normales

- Para realizar este desafío debes haber estudiado previamente todo el material disponibilizado correspondiente a la unidad.
- Una vez terminado el desafío, comprime la carpeta que contiene el desarrollo de los requerimientos solicitados y sube el `.zip` en el LMS.
- Desarrollo desafío:
 - El desafío se debe desarrollar de manera Individual.

1. Importe un csv

- Importe el set de datos creados por usted al final de la sesión pasada.

2. Importe su conjunto de funciones

- Genere un archivo `ancilliary_funcs.py` para guardar las funciones creadas en la sesión pasada e impórtelo.
- Utilice las funciones creadas para inspeccionar la cantidad de casos perdidos de la muestra importada.

3. Estime el puntaje Z de una variable

Estime el puntaje Z la variables que posea la menor cantidad de datos perdidos (o que no posea).

- Genere una nueva columna en la tabla de datos que contenga el puntaje Z de dicha variable. Recuerde que para calcular el puntaje z de una observación se utiliza la fórmula: $z_i = (x_i - \bar{x}) / \sigma(x)$
- Para ordenar los valores pueden utilizar la función `sort_values` del `DataFrame` analizado. Lean la documentación asociada al argumento `by`.
- Asegúrense de eliminar las observaciones que tengan valores perdidos con `dropna`. Lean la documentación asociada al argumento `subset`.
- Selecciones las primeras y últimas 10 observaciones de la base reordenada. ¿Qué características en común se encuentran?

4. Estime el puntaje z de las regiones

Utilizando el resultado anterior, visualice su nueva variable estandarizada (es decir, luego de calcular el puntaje Z).

- Para ello realice un `dotplot` que muestre el valor promedio de su variable por cada región.
- Agregue además una línea que muestre el promedio total de la variable (es decir, antes de agrupar).

5. Curvas de Densidad de Probabilidad

- Genere una función que permita visualizar las curvas de Densidad de Probabilidad para las variables `imf_pop`, `gle_cgdpc` y `wdi_exph`.
- La función debe contener como argumentos:
 - `df`: Tabla de datos 1.
 - `variable`: Variable a contrastar.