

Desafío - Estimando curvas de densidad de probabilidad

- Para realizar este desafío debes haber estudiado previamente todo el material disponible 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: individual.

Habilidades a evaluar

- Conocer las convenciones y principios rectores de la visualización de gráficos.
- Describir la Ley de los Grandes Números y Teorema del Límite Central y su importancia en la inferencia estadística.

Descripción

La empresa en la cual usted trabaja quedó muy satisfecha con su análisis anterior del Índice de Desarrollo Humano (IDH) de diversos países, el cual se elaboró para el Ministerio de Relaciones Exteriores a fin de poder contar con la data necesaria para poder comenzar a trabajar en relaciones diplomáticas más estrechas con diversos países. Por consiguiente, esta vez le solicitan que, a partir del trabajo hecho en el Desafío precedente, usted pueda seguir profundizando en la base de datos *Quality of Government*, de la Universidad de Gotemburgo.

Requerimientos

A continuación, revisaremos los requerimientos y acciones que el Ministerio de RREE te pide realizar.

1. **Importar un csv (2 puntos):** Importe el set de datos creados por usted al final de la sesión pasada.
2. **Importar su conjunto de funciones (2 puntos).** Para ello deberá:
 - Generar un archivo `ancillary_funcs.py` para guardar las funciones creadas en la sesión pasada e importe.

- Utilizar las funciones creadas para inspeccionar la cantidad de casos perdidos de la muestra importada.
3. **Estimar el puntaje Z de una variable (2 puntos):** Se debe utilizar la que posea la menor cantidad de datos perdidos (o que no posea).
- Generar 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. **Estimar el puntaje z de las regiones (2 puntos).** 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. **Generar curvas de Densidad de Probabilidad (2 puntos).**
- Cree 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.