

Pauta Tarea 3

Objetivo: Se espera que los estudiantes demuestren el aprendizaje de las herramientas básicas sobre visualización de datos y análisis de la forma de una distribución utilizando el software R (aprendidas en las sesiones 6 y 7 del Laboratorio), sabiendo interpretar y reflexionar críticamente sobre los resultados obtenidos.

Instrucciones:

1. Creación de un Proyecto de R
 - 1.1. Crear un .Rproject llamado “tarea03ED” y establecer una estructura de carpetas siguiendo el formato IPO, esto es, una carpeta de *input*, *procesamiento* y *output*.
 - 1.2. Guardar en la carpeta *input* la base de datos adjuntada por el equipo docente en U-Cursos; crear en la carpeta *procesamiento* un script llamado “01proc-data” y otro script llamado “02analisis”.
2. En el script de “01proc-data”:
 - 2.1. Instalar y cargar librerías correspondientes vía *sintaxis*.
 - 2.2. Cargar la base de datos ELSOC_W06_v1.0_R.RData vía *sintaxis*.
 - 2.3. Visualizar los nombres de todas las variables de base de datos vía *sintaxis*.
 - 2.4. Seleccionar y renombrar las variables *sexo* (m0_sexo), *ingreso* (m13) y *nivel_educativo* (m01), y almacenarlas en una base de datos creada vía *sintaxis* y llamarla *datos_proc*.
 - 2.5. Recodificar las variables *sexo*, *ingreso* y *nivel_educativo* vía *sintaxis*, dejando como valores perdidos las observaciones o categorías correspondientes. Además, recodificar la variable *sexo* vía *sintaxis* dejando sus valores respectivos como “Hombre” o “Mujer”. Tenga en especial consideración para este paso el libro de códigos de la base de datos.
 - 2.6. Identificar y remover valores perdidos (NA 's) vía *sintaxis*.
 - 2.7. Generar una nueva variable *tramo_ingresos* vía *sintaxis*, con los tramos: menor o igual a \$340.000, entre \$340.001 y \$680.000, entre \$680.001 y \$920.000, entre \$920.001 y \$1.220.001, y mayores a \$1.220.001.
 - 2.8. Guardar base de datos procesada en formato RData (.RData) y exportar esta base procesada a la carpeta *output* de su proyecto, ambos vía *sintaxis*.
 - 2.9. Limpiar el entorno de trabajo vía *sintaxis*.
3. En el script de “02analisis”:
 - 3.1. Instalar y cargar librerías correspondientes vía *sintaxis*.
 - 3.2. Cargar base de datos procesada alojada en carpeta *output* de su proyecto, vía *sintaxis*.

- 3.3. Generar un gráfico de barras para la variable sexo y otro para tramo_ingresos, vía syntaxis.
- 3.4. Obtener estadísticos descriptivos para las variables ingreso y nivel_educativo.
- 3.5. Generar un gráfico de cajas para la variable nivel_educativo, vía syntaxis.
- 3.6. Generar un gráfico tipo histograma para la variable ingreso y otro para nivel_educativo, vía syntaxis.
- 3.7. Generar un gráfico tipo histograma con densidad para la variable ingreso y otro para la variable nivel_educativo, y contrastar en ambos con la distribución normal, vía syntaxis.
- 3.8. Limpiar el entorno de trabajo vía syntaxis.
4. Reportar e interpretar los gráficos y resultados obtenidos en los puntos anteriores (3.3, 3.4, 3.5, 3.6 y 3.7) en la misma syntaxis usando `#`. En el caso de los gráficos tipo histograma y densidad interpretar la forma de la distribución (asimetría y curtosis), y comparar distribución normal con la empírica.

Entregue como productos de esta tarea:

- Un archivo .zip o comprimido del proyecto, el cual contenga los elementos solicitados: .Rproject, carpetas de Input (con la base de datos original), Procesamiento (con los respectivos códigos de procesamiento y análisis) y de Output (con la base de datos procesada).

Aspectos formales:

- Fecha y formato de entrega: jueves 06 de julio hasta las 23:59 vía módulo Tareas en plataforma U-Cursos.
- Entregas atrasadas hasta las 23:59 del domingo 9 de julio, pero tendrán 5 décimas de descuento sobre la nota final por cada día de atraso.
- Luego del domingo 9 de junio no se aceptarán más trabajos y quienes no entreguen serán evaluados con nota 1.0.
- En caso de detectarse plagio, la tarea será evaluada con nota 1.0. Esta falta puede resultar causal de reprobación de la asignatura.

En caso de tener dudas, no dude en contactar a sus ayudantes respectivos, o bien, vía foro U-Cursos al equipo docente de la asignatura.