**Taller - Practica en software R.**

**Lineamientos generales**

* Debe resolver el taller asociándose con otra persona o solo, entendiendo que la calificación será la misma para ambos estudiantes.
* El archivo del reporte final debe ser un archivo pdf.
* Se debe realizar en R y adjuntar el código correspondiente. No mezclar el código con el reporte final.
* El objetivo principal de este trabajo es la claridad lógica y la interpretación de los resultados. El informe no necesita ser extenso. Recuerde ser minimalista escribiendo el reporte. Se deben incluir solo aquellos gráficos, tablas, y fórmulas que son relevantes para la discusión.
* Numerar figuras y tablas, proporcionarles a un tamaño adecuado que no distorsione la información que estas contienen.
* Además del contenido, serán parte integral de la calificación: la redacción, la ortografía, la calidad de las gráficas, tablas y fórmulas, y en general, la presentación y estructura general del reporte final.
* Cualquier evidencia de plagio o copia se castigará severamente tal y como el reglamento de la Universidad lo estipula.

**IMPORTANTE:** Los resultados de estos análisis no consideran el factor de expansión por lo que lo obtenido no debe generalizarse a todo el país.

El uso de estos factores no es complicado, eventualmente se reduce al cálculo de promedios ponderados. Sin embargo, aprender como manejarlos hace parte de cursos posteriores de estadística, en particular los cursos de Análisis de encuestas por muestreo probabilístico.

**Instrucciones.**

# Descargue las bases de datos correspondientes a la Encuesta Nacional de Lectura (ENLEC) 2017 dispuesta en la página del DANE en el enlace: http://microdatos.dane.gov.co/index.php/catalog/550/get\_microdata

# Allí encontrarán 6 bases de datos, de las cuales solo se deben descargar las siguientes:

* Form\_lectura
* Admin\_Personas
* Admin\_vivienda

# Se descargarán archivos “.zip”, extraiga y quédese únicamente con los archivos tipo “.txt”. Guárdelos en la misma carpeta donde se está guardando el script de R y ejecute el código “00 LecturaBDD.R”

# El código junta las bases descargadas y selecciona únicamente las variables que se van a trabajar en el taller. Es una buena idea entenderlo.

# En el archivo “Dicccionario.xlsx” encontrará el código de cada pregunta, la etiqueta de la pregunta, su descripción y los posibles valores de respuesta.

# Utilice el software R y lo visto en clase para responder las preguntas que se mostraran más adelante.

# Se debe entregar el Script en R y el taller nombrados de la siguiente manera:

* NúmeroIdentificaciónEstudiante\_PrimerNombrePrimerApellido\_Script\_taller
* NúmeroIdentificaciónEstudiante\_PrimerNombrePrimerApellido\_Solucion\_taller

# Por ejemplo: 1234566\_SergioMartinez\_Script\_taller, 1234566\_SergioMartinez\_Solucion\_taller.

**Nota: Debe comentar todos los resultados.**

# Preguntas.

**Sobre la base de datos consolidada “baseConsolidadaENLEC.xlsx” que se genera al final del script “00 LecturaBDD.R”**

1. Construya la variable “RangosEdad”, que clasifica las edades de acuerdo con los siguientes grupos de edad.

* Menos de 18 años.
* De 18 a 24 años.
* De 25 a 34 años.
* De 35 a 44 años.
* De 45 a 54 años.
* 55 años o más.

1. Describa los encuestados en términos de clase de zona (Cabecera y centros poblados y rural disperso), sexo, rangos de edad, nivel educativo más alto alcanzado, estado civil y su cultura, pueblo o rasgos físicos.
2. La brecha existente entre el sector urbano y rural colombiano, conlleva a insuficiencias en la oferta educativa en zonas rurales y precariedad de las condiciones educativas a sí mismo restricción al acceso de materiales de lectura. Identifique el total de encuestados a los que les gusta leer o que les lean, adicional las principales razones por las que los encuestados no leen ya sea en soporte digital o soporte impreso. Analice y comente los resultados a nivel general y por clase de zona (Cabecera y centros poblados y rural disperso).
3. Es de interés los beneficios y la importancia de descubrir la lectura desde la niñez, así como la posible asociación entre las prácticas de lectura durante la infancia con el nivel educativo más alto alcanzado y el gusto por la lectura. Realice un análisis descriptivo como insumo para dicho propósito.
4. Analice las relaciones entre las horas promedio a la semana en que se leen redes sociales, periódicos impresos, artículos o documentos académicos impresos y digitales y páginas web con el estrato socioeconómico del encuestado.
5. La inclusión de las tecnologías de información y comunicación ha modificado las estrategias de aprendizaje en el estilo de leer, quizás por facilitar el acceso a los materiales de lectura e incluirlos en nuestras vidas cotidianas, sin embargo, a pesar de la cercanía existe la inquietud de si leemos poco, mal y por obligación. Identifique la cantidad promedio de libros impresos o físicos, digitales o electrónicos o audiolibros leídos, la cantidad promedio de libros leídos por trabajo o estudio y por entretenimiento o gusto. Establezca si existen diferencias estadísticamente significativas entre:
6. La cantidad de libros impresos o físicos y la cantidad de libros digitales o electrónicos o audiolibros leídos, concluya.
7. La cantidad de libros leídos por entretenimiento o gusto y la cantidad de libros leídos por trabajo o estudio, concluya.
8. La cantidad de libros impresos o físicos leídos y la cantidad de libros leídos por trabajo o estudio, concluya.
9. La cantidad de libros digitales o electrónicos o audiolibros leídos y la cantidad de libros leídos por entretenimiento o gusto, concluya.
10. Plantee una manera de resumir en un solo indicador la frecuencia de lectura de libros impresos y otro indicador para la frecuencia de lectura de libros digitales. Luego, analice los resultados con las variables que considere más pertinentes.