

Universidad de San Carlos de Guatemala

Facultad de Ingeniería

Escuela de Ciencias y Sistemas

Seminario de Sistemas 2 Sección A

Ing. William Estuardo Escobar Argueta

Aux. Glen Abra-ham Calel Robledo



Práctica 2

OBJETIVOS

- El estudiante aprenda y utilice el lenguaje de programación R para el análisis de datos.
- El estudiante pueda realizar la lectura de diferentes archivos de entrada en R.
- El estudiante pueda realizar gráficas de datos que permitan su análisis.
- El estudiante utilice plataformas para el aprendizaje del lenguaje de programación R y pueda poner en práctica lo aprendido.

DESCRIPCIÓN

La práctica tiene como objetivo principal que el estudiante implemente una solución como analista de datos utilizando un lenguaje de programación dedicado para el análisis de datos.

Debido a las múltiples implementaciones realizadas, usted ha ganado fama como analista de datos, por tal razón varias empresas desean contratarle para que pueda proporcionar soluciones a sus necesidades. Para brindarles una solución óptima se requiere que analice los mismos en R, un lenguaje de programación para computación estadística y gráficas.

IMPLEMENTACIÓN SUGERIDA

1. Utilice una distribución de Linux, Windows o Mac, como usted prefiera.
2. Instale la línea de comandos e intérprete de R.
3. Instale un IDE de su elección, por ejemplo RStudio.
4. Realice scripts individuales para cada uno de los análisis solicitados.
5. Generar una imagen por cada reporte.

REPORTES

1. **Archivo sales:**

- a. Gráfica de pie que muestre el total de ganancias de los países de Centroamérica a partir de las ventas de ropa.
- b. Gráfica de barras a su elección, que compare el número de artículos de las ventas en línea contra las ventas que no son en línea.
- c. Gráfica a su elección que muestre el año con más ordenes de prioridad M.

2. **Archivo cardio_train:**

- a. Genere el histograma y polígono de frecuencia (en la misma gráfica) del peso.

3. **Archivo suicide:**

- a. Tabla de frecuencias de los rangos de los suicidios en Guatemala (todas las edades)
- b. Gráfica de barras de la frecuencia absoluta
- c. Polígono de frecuencias.
- d. Diagrama de frecuencia acumulada.

4. **Archivo fifa:**

- a. Gráfica de barras de la potencia de disparo promedio, máximo y mínimo de los jugadores con nacionalidad alemana.
- b. Gráfica que muestre al club con la mayor agresividad promedio y al club con la menor agresividad promedio.
- c. Gráfica que muestre al club con el mayor regate promedio y al club con el menor regate promedio.

5. **Archivo regresion.**

- a. Análisis exploratorio de datos para comprobar si existe relación entre estrés, depresión y ansiedad.
- b. De existir relación, realice una regresión lineal entre los datos y demuestre que el modelo produce datos válidos.

RESTRICCIONES

- Utilizar R para el análisis de los datos.
- Las tablas de frecuencia deben tener las columnas variable, frecuencia absoluta, frecuencia acumulada, frecuencia relativa y frecuencia relativa acumulada.
- Las gráficas deben ser de los tipos solicitados y entendibles, con los nombres correspondientes en cada eje.

- Las gráficas deben mostrar los resultados numéricos, puede ser por medio de leyendas o encima de las gráficas.
- Para cada numeración se debe entregar scripts individuales (5 archivos).
- La calificación se realizará sobre lo que se envía.
- Se prohíbe modificar código dentro de la calificación de la práctica.

CONSIDERACIONES

- La entrega es individual.
- Fecha de entrega: entregar scripts y documentación en la plataforma UEDI identificado como **[SS2]Practica2_carnet.zip** el día domingo 28 de junio.
- Entregas tarde se califican con un 50% de penalización.
- Copias detectadas tendrán nota 0 y reporte a la escuela.