

Universidad de San Carlos de Guatemala

Facultad de Ingeniería

Escuela de Ciencias y Sistemas

Seminario de Sistemas 2 Sección A

Ing. Luis Alberto Vettorazzi España

Aux. Luis Gilberto Barrios Vásquez



Práctica 2

R

Objetivos

- Conocer y utilizar el lenguaje de programación R para el análisis de datos.
- Realizar la lectura de diferentes archivos de entrada en R.
- Realizar gráficas de datos que permitan su análisis

Descripción

Después de las múltiples implementaciones realizadas, usted ha ganado fama como analista de datos, por lo que varias empresas desean proporcionarle parte de sus datos almacenados, que han recopilado para propósitos varios. Para esto 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 o Windows, como usted prefiera.
2. Instale la línea de comandos e intérprete de R.
3. Instale un IDE de su elección, como RStudio.
4. Realice scripts individuales para cada uno de los análisis solicitados.

Reportes

1. Archivo de Ventas:
 - a. Tabla de frecuencias acerca de las veces que se han realizado ventas a los países de Centro América (Guatemala, El Salvador, Costa Rica, Honduras, Nicaragua, Panamá y Belice).
 - b. Gráfico de barras de la frecuencia absoluta del reporte 1.
 - c. Gráfico de frecuencia acumulada del reporte 1.
2. Archivo del desempeño de los CPU:
 - a. Grafica de barras del rendimiento promedio, rendimiento máximo y rendimiento mínimo de los CPU fabricados por *amdahl*

- b. Gráfica que muestre la empresa que fabrica los CPU con el mayor rendimiento promedio, y la empresa que fabrica los CPU con el menor rendimiento promedio. (Incluya todos sus modelos en el cálculo)
3. Deceso por sexo y edad.
 - a. Tabla de frecuencias de los rangos de los decesos (todas y únicamente las edades)
 - b. Histograma de la frecuencia absoluta
 - c. Polígono de frecuencias
 - d. Diagrama de frecuencia acumulada
4. Presión Sanguínea y Edad.
 - a. Análisis exploratorio de datos para comprobar si existe relación entre la presión sanguínea y la edad de los pacientes.
 - b. De existir relación, realice una regresión lineal entre los datos y demuestre que el modelo produce datos válidos.
5. Pobreza y desempleo.
 - a. Análisis exploratorio de datos para comprobar si existe relación entre la pobreza, el desempleo y la cantidad de asesinatos.
 - 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.
- Entregas tarde se califican con un 50% de penalización.
- Fecha de entrega: entregar scripts y documentación con screens de los reportes en la publicación de Google Classroom identificado como **[SS2]Practica2_carne.zip** el día 26/04/2019
- Fecha de calificación: Sábado 27/04/2019