

Sala de Entrenamiento 2: descripción de los Datos

Análisis Estadístico de Datos con R

Semana 2

Diplomados Online

Febrero 2019

En esta sala de entrenamiento pongamos en práctica lo aprendido en la clase.

- 1) Escriba el programa de asimetría y utilícelo para calcular el coeficiente de asimetría de los cuatro exámenes parciales en *resultados.xlsx*, los cuales, corresponden a las notas obtenidas por dos grupos de estudiantes de 119 (identificando su género), para dos semestres consecutivos (identificados como 1 y 2), en dos materias, a saber, arquitectura del computador(*arquit*) y programación (prog) calificados del 1 al 100. ¿Qué puede decir sobre estos datos?

- 2) Pearson ha dado una fórmula aproximada para la asimetría, que es más fácil que la fórmula exacta:

$$\text{Asimetría} \approx \frac{3(\text{media} - \text{mediana})}{\text{desviación Estandar}}$$

Escribe un programa para calcular esto y aplíquelo a los datos en *resultados.xlsx*. ¿Es una aproximación razonable?

- 3) Es necesario estimar el número de búferes¹ de mensajes que se utilizan en la memoria principal del sistema informático de Roantree Retailers Ltd. Para ello se corren 20 programas y se encontró que los números de búferes de mensajes en uso eran,

141, 146, 157, 151, 152, 140, 142, 156, 150, 140, 139, 135, 143, 146, 146, 152, 140, 136, 149, 148

Calcular el número medio de búferes utilizados. ¿Cuál es la desviación estándar? ¿Diría que estos datos están sesgados?

- 4) Utilicemos los datos de *resultados.xlsx*, para:

- i) Desarrollar un diagrama de caja de todos los temas en el mismo gráfico.
- ii) Obtener un gráfico de tallo y una hoja de cada tema en *resultados.xlsx*. ¿Están emergiendo patrones?

- 5) Para los 50 estudiantes que se muestras a continuación,

Femenino: 57, 59, 78, 79, 60, 65, 68, 71, 75, 48, 51, 55, 56, 41, 43, 44, 75, 78, 80, 81, 83, 83, 85

¹Instrumento digital reservado para el almacenamiento temporal de información digital, mientras que está esperando ser procesada

Masculino: 48, 49, 49, 30, 30, 31, 32, 35, 37, 41, 86, 42, 51, 53, 56, 42, 44, 50, 51, 65, 67, 51, 56, 58, 64, 64, 75

Use R para:

- (a) Formar el gráfico de tallo y la hoja para cada género, y discutir las ventajas de esta representación en comparación con el histograma tradicional;
- (b) Construir un diagrama de caja para cada género y discutir los hallazgos.