

Recuperativa

June 23, 2024

Fecha de entrega. 30 de junio. Debe enviar su prueba al correo rodrigo.barrera@upla.cl

1 Conjunto de datos tips

Este conjunto de datos contiene las siguientes variables:

- **total_bill:** La cuenta total del cliente.
- **tip:** La propina dada por el cliente.
- **sex:** Sexo del cliente (Male/Female).
- **smoker:** Si el cliente es fumador o no (Yes/No).
- **day:** Día de la semana (Thur, Fri, Sat, Sun).
- **time:** Momento del día (Lunch/Dinner).
- **size:** Tamaño del grupo del cliente.

2 Instrucciones:

1. **Calcule las medidas de tendencia central** para las variables `total_bill`, `tip`, y `size`. Esto incluye:
 - La **media**, que es el promedio de los valores.
 - La **mediana**, el valor medio que divide los valores en dos mitades.
 - La **moda**, el valor que aparece con más frecuencia.
2. **Determine las medidas de dispersión** para las mismas variables, que ayudan a entender la variabilidad de los datos:
 - El **rango**
 - La **desviación estándar**
3. **Analice la distribución de las variables categóricas** (`sex`, `smoker`, `day`, `time`), calculando la **frecuencia** y **proporción** de cada categoría.
4. **Cree gráficos** para visualizar los datos:
 - **Histogramas** para `total_bill`, `tip`, y `size`, para visualizar la distribución de estas variables.
 - **Diagramas de caja (boxplots)** para las mismas variables, para ilustrar la mediana, los cuartiles y los valores atípicos.
 - **Gráficos de barras** para las variables categóricas, para mostrar la frecuencia o proporción de cada categoría.

5. Realice un análisis de correlación entre las variables adecuadas, calculando:

- La **matriz de correlación** que muestra la relación entre cada par de variables.
- Un **mapa de calor (heatmap)** de la matriz de correlación para una visualización intuitiva de las relaciones.

```
[ ]: # Este código le puede ser de ayuda.  
import seaborn as sns  
import pandas as pd  
import numpy as np  
import matplotlib.pyplot as plt  
  
# Cargar el conjunto de datos 'tips'  
df = sns.load_dataset('tips')
```