

February 13, 2024

1 Actividad 6

1.1 Juan Carlos Solórzano Castanedo

1.2 Análítica de Datos

1.3 12 de Febrero del 2024

```
[2]: #importar las librerías
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

Matplotlib is building the font cache; this may take a moment.

```
[10]: #Generación de datos de muestra para una variable continua
np.random.seed(42)
datos = np.random.normal(loc=100, scale=15, size=1000) #distribución normal,
↳ 1000 datos con media 100 y desviación estandar 15

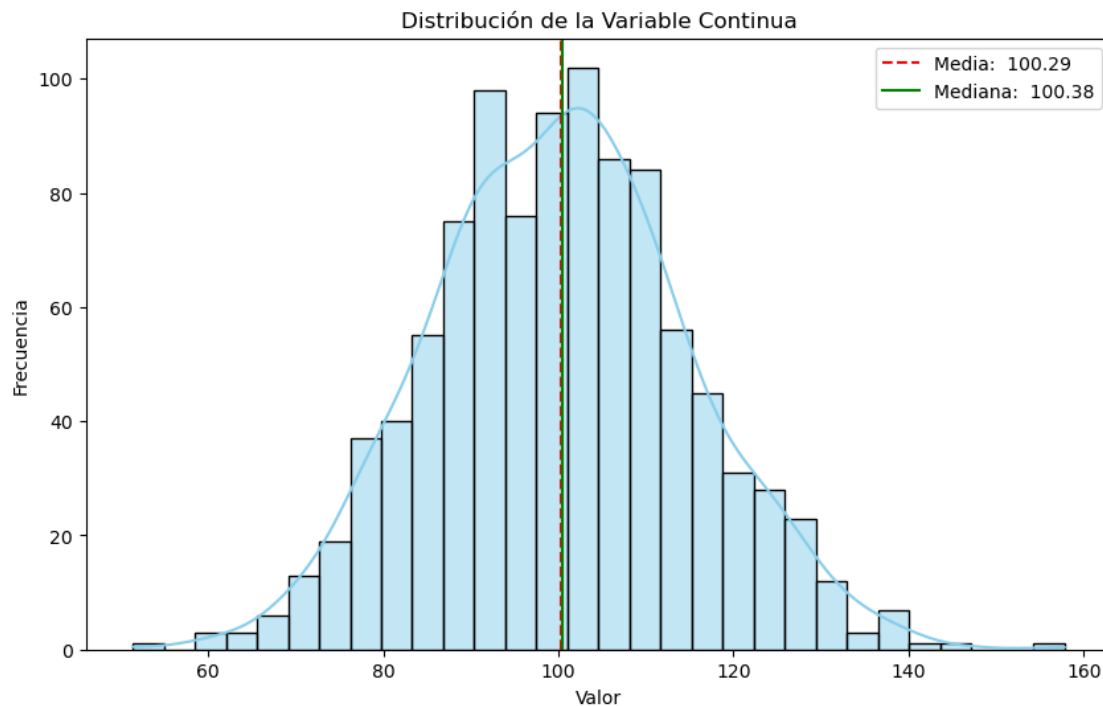
#Generación de dataframe de pandas para facilitar el análisis
df = pd.DataFrame(datos, columns=['Variable_Continua'])
```

```
[21]: #Análisis descriptivo
media = df['Variable_Continua'].mean()
mediana = df['Variable_Continua'].median()
varianza = df['Variable_Continua'].var()
desviación_estandar = df['Variable_Continua'].std()

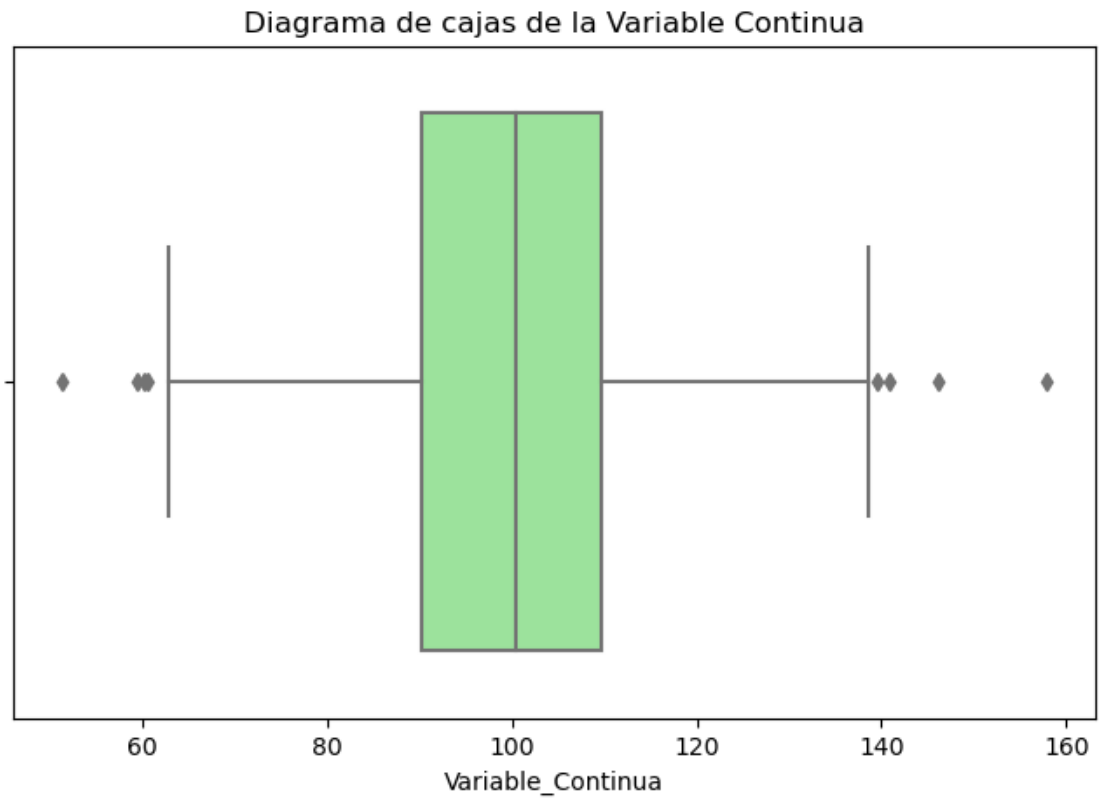
print(f"Media: {media: .2f}")
print(f"Mediana: {mediana: .2f}")
print(f"Varianza: {varianza: .2f}")
print(f"Desviación Estándar: {desviación_estandar: .2f}")
```

```
Media: 100.29
Mediana: 100.38
Varianza: 215.74
Desviación Estándar: 14.69
```

```
[29]: # Visualización - Histograma
plt.figure(figsize=(10, 6))
sns.histplot(df['Variable_Continua'], kde=True, bins=30, color='skyblue')
plt.title('Distribución de la Variable Continua')
plt.xlabel('Valor')
plt.ylabel('Frecuencia')
plt.axvline(media, color='r', linestyle='--', label=f'Media: {media: .2f}')
plt.axvline(media, color='g', linestyle='-', label=f'Mediana: {media: .2f}')
plt.legend()
plt.show()
```



```
[33]: # Visualización - Diagrama de cajas y bigotes
plt.figure(figsize=(8,5))
sns.boxplot(data=df, x="Variable_Continua", color='lightgreen')
plt.title('Diagrama de cajas de la Variable Continua')
plt.show()
```



```
[35]: # Guardad DataFrame en CSV  
df.to_csv('datos_variable_continua.csv', index=False)
```

```
[ ]:
```