

```
# Cargar librerías
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# Cargar datos
df = pd.read_csv('https://github.com/ET771/arte_analitica/raw/refs/heads/main/Social%20Media')
df.head()
```



|   | Platform | Owner     | Primary Usage              | Country          | Daily Time Spent (min) | Verified Account | Date Joined |
|---|----------|-----------|----------------------------|------------------|------------------------|------------------|-------------|
| 0 | WhatsApp | Meta      | Messaging                  | Switzerland      | 113.94                 | Yes              | 2019-03-03  |
| 1 | WeChat   | Tencent   | Messaging and social media | Madagascar       | 49.63                  | Yes              | 2023-09-21  |
| 2 | Snapchat | Snap Inc. | Multimedia messaging       | Pitcairn Islands | 29.01                  | Yes              | 2020-12-13  |

```
# Dimensiones y tipos de datos
print("Filas y columnas:", df.shape)
df.info()

# Comprobación de valores faltantes
print("\nValores faltantes por columna:")
print(df.isnull().sum())
```



```
Filas y columnas: (10000, 7)
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10000 entries, 0 to 9999
Data columns (total 7 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   Platform                             10000 non-null  object
1   Owner                                10000 non-null  object
2   Primary Usage                         10000 non-null  object
3   Country                              10000 non-null  object
4   Daily Time Spent (min)                10000 non-null  float64
5   Verified Account                      10000 non-null  object
6   Date Joined                           10000 non-null  object
dtypes: float64(1), object(6)
memory usage: 547.0+ KB
```

```
Valores faltantes por columna:
Platform      0
Owner         0
Primary Usage 0
Country       0
```

```
Daily Time Spent (min)    0
Verified Account          0
Date Joined               0
dtype: int64
```

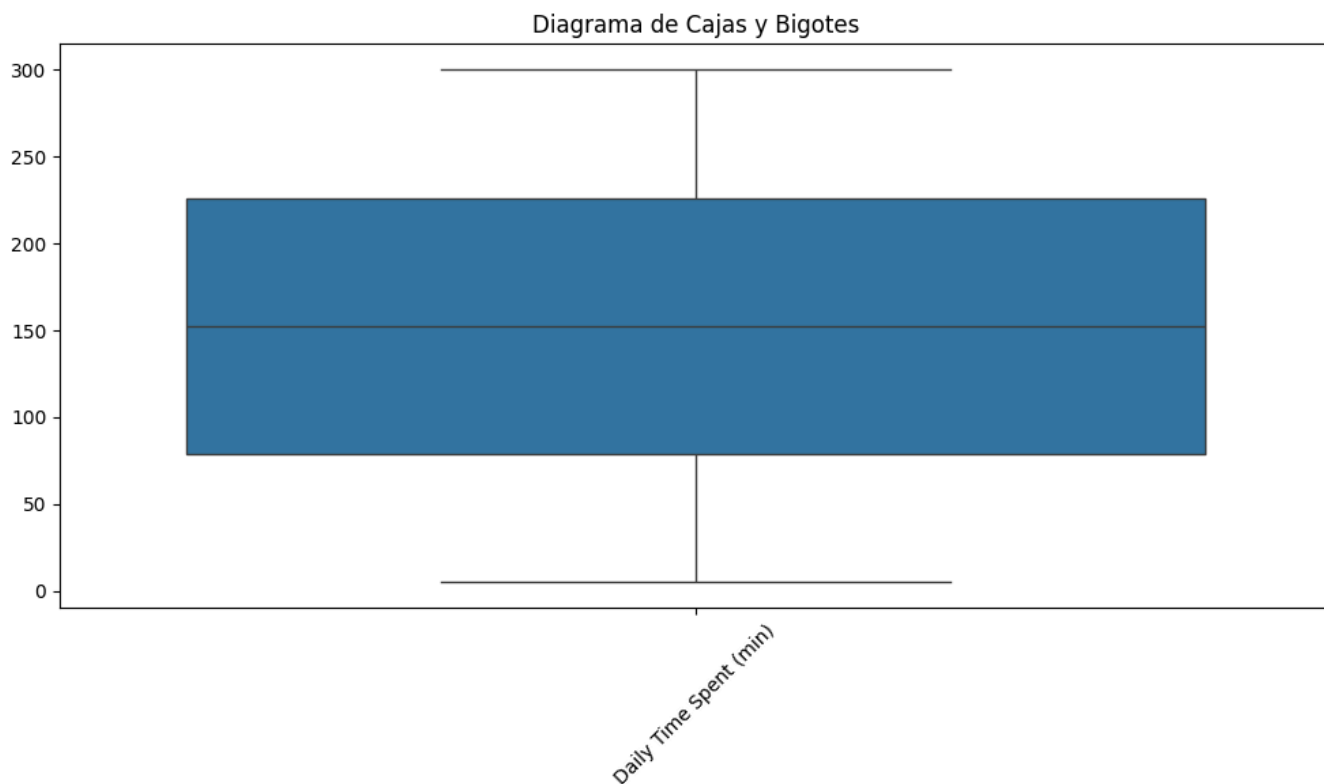
```
# Rango mínimo y máximo por variable numérica
for col in df.select_dtypes(include='number').columns:
    print(f"{col} -> Mín: {df[col].min()} | Máx: {df[col].max()}")
```

```
⇒ Daily Time Spent (min) -> Mín: 5.02 | Máx: 300.0
```

```
# Media, mediana y desviación estándar
for col in df.select_dtypes(include='number').columns:
    print(f"\n{col}")
    print("Media:", df[col].mean())
    print("Mediana:", df[col].median())
    print("Desviación estándar:", df[col].std())
```

```
⇒
Daily Time Spent (min)
Media: 152.21114500000002
Mediana: 152.735
Desviación estándar: 85.14274960911294
```

```
# Diagrama para identificar outliers y distribución
plt.figure(figsize=(10, 6))
sns.boxplot(data=df.select_dtypes(include='number'))
plt.title("Diagrama de Cajas y Bigotes")
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```

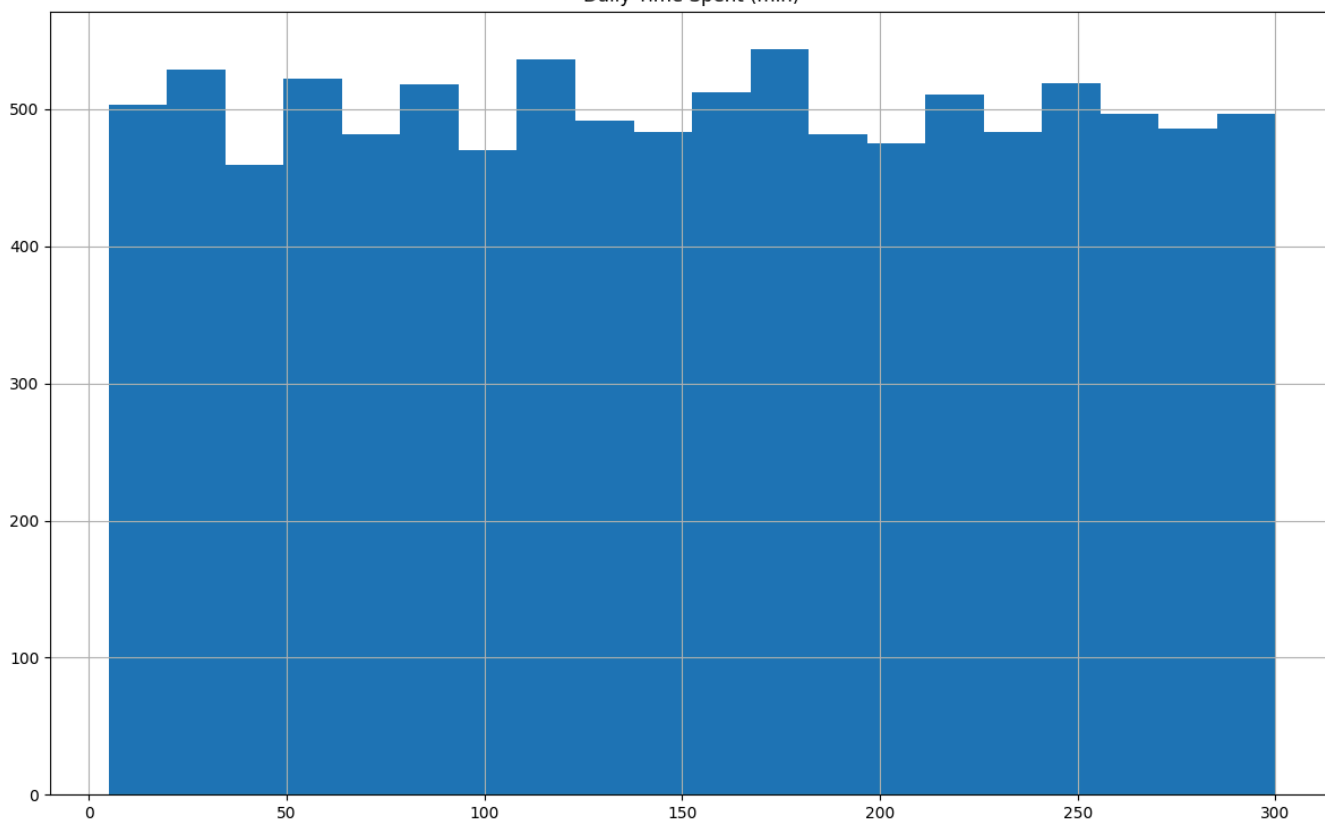


```
# Histograma para distribución de frecuencias
df.select_dtypes(include='number').hist(figsize=(12, 8), bins=20)
plt.suptitle("Histogramas de Variables Numéricas")
plt.tight_layout()
plt.show()
```



Histogramas de Variables Numéricas

Daily Time Spent (min)

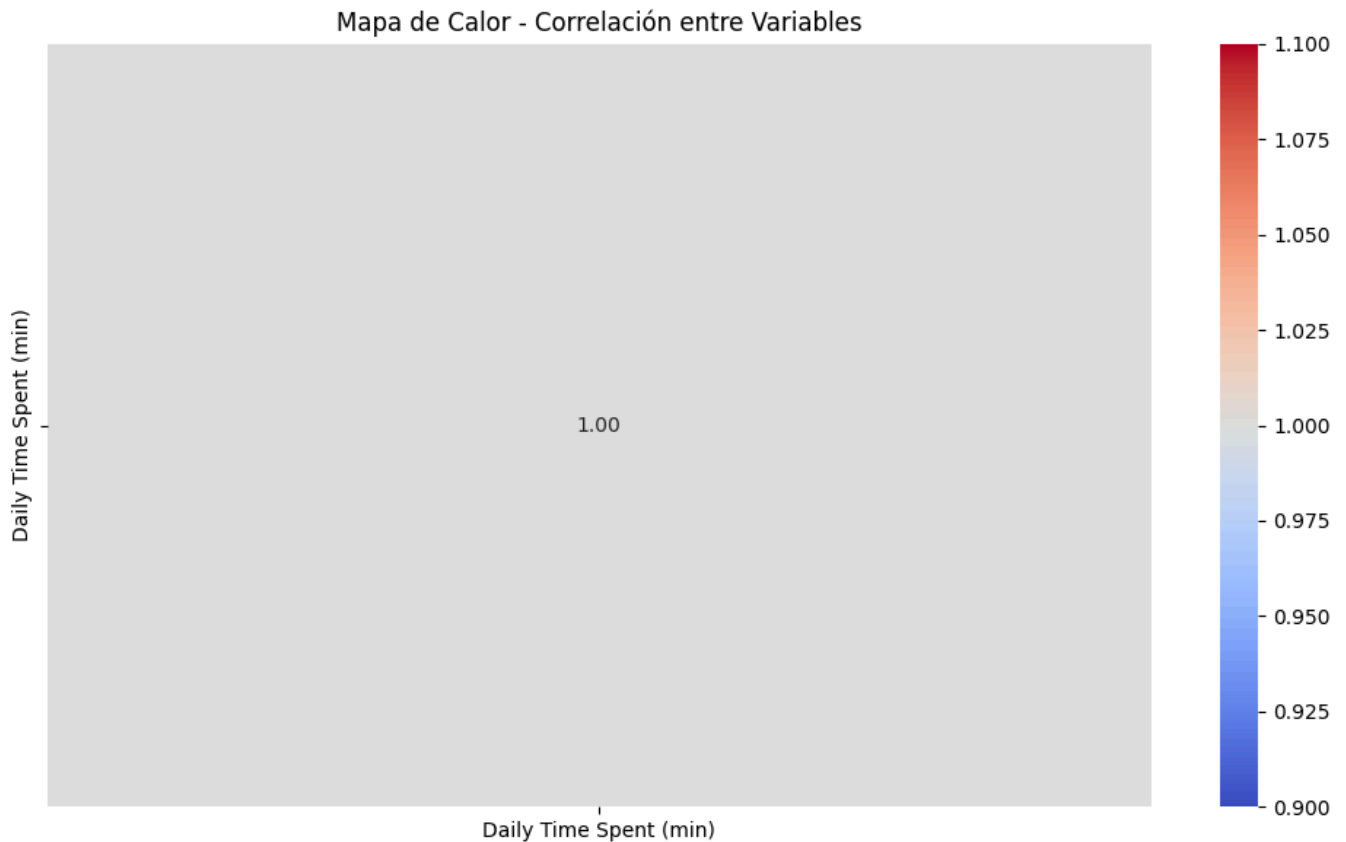


```
# Cálculo de correlaciones entre variables numéricas
# Selecting only numeric columns for correlation calculation
numeric_df = df.select_dtypes(include='number')
correlation_matrix = numeric_df.corr()
print(correlation_matrix)
```



Daily Time Spent (min)      Daily Time Spent (min)  
1.0

```
# Representación gráfica de las correlaciones
plt.figure(figsize=(10, 6))
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', fmt=".2f")
plt.title("Mapa de Calor - Correlación entre Variables")
plt.tight_layout()
plt.show()
```



Double-click (or enter) to edit

Conclusión El análisis de correlación revela que el uso diario de Twitter está débil pero positivamente correlacionado con el de LinkedIn, especialmente en días laborales. Esta relación sugiere que los picos de actividad en Twitter pueden coincidir con eventos informativos o laborales, como noticias económicas, lanzamientos corporativos o conferencias, que también impulsan la participación en LinkedIn asumiendo esto, el comportamiento profesional e

informativo digital podría estar más entrelazado de lo que se asume, mostrando que ciertas plataformas actúan como reflejos indirectos del entorno laboral diario.

Double-click (or enter) to edit