import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

In [3]:

*# Load the dataset*

data = pd.read\_csv("/content/E-commerce Website Logs.csv")

In [4]:

*# Display basic information about the dataset*

print("Dataset shape:", data.shape)

Dataset shape: (97752, 15)

In [5]:

print("\nColumns:", data.columns)

Columns: Index(['accessed\_date', 'duration\_(secs)', 'network\_protocol', 'ip', 'bytes',

'accessed\_Ffom', 'age', 'gender', 'country', 'membership', 'language',

'sales', 'returned', 'returned\_amount', 'pay\_method'],

dtype='object')

In [6]:

print("\nData types:\n", data.dtypes)

Data types:

accessed\_date object

duration\_(secs) float64

network\_protocol object

ip object

bytes float64

accessed\_Ffom object

age object

gender object

country object

membership object

language object

sales float64

returned object

returned\_amount float64

pay\_method object

dtype: object

*# Plot visualizations*

*# Distribution of sales*

plt.figure(figsize=(10, 6))

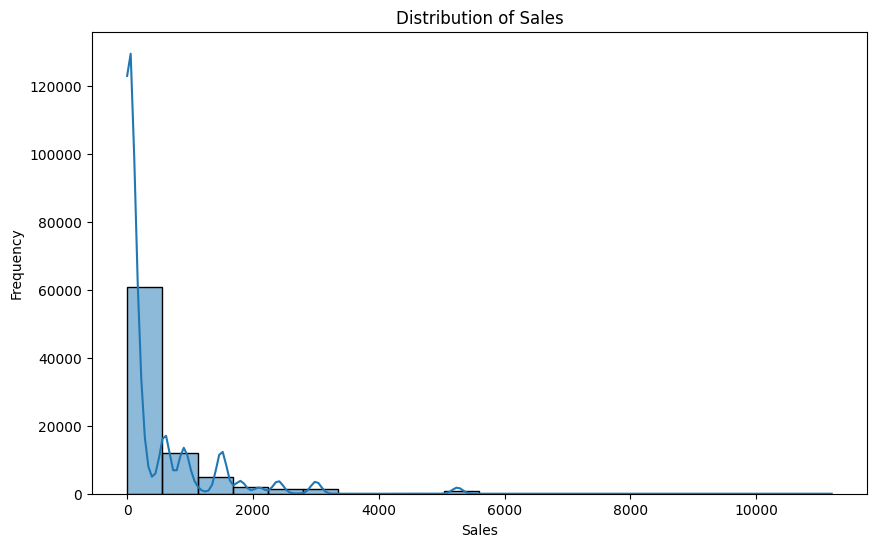
sns.histplot(data['sales'], bins=20, kde=True)

plt.title('Distribution of Sales')

plt.xlabel('Sales')

plt.ylabel('Frequency')

plt.show()



*# Sales vs. Age*

plt.figure(figsize=(10, 6))

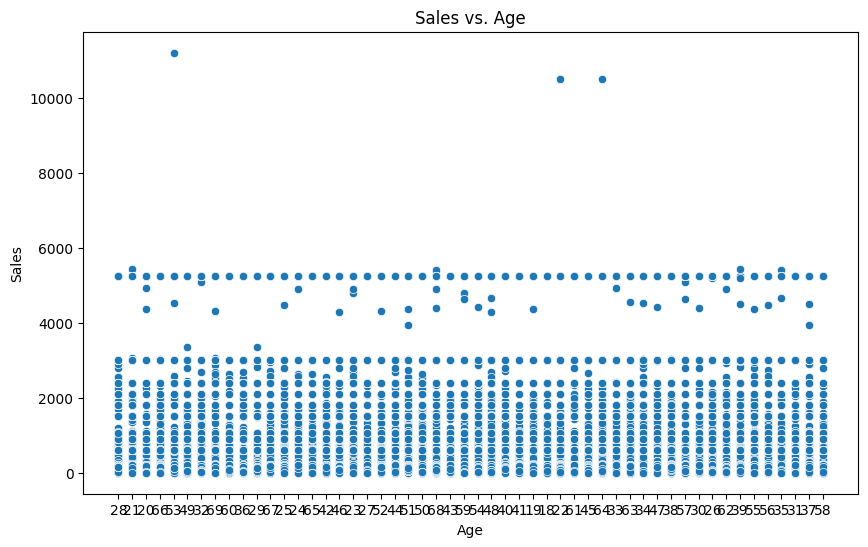
sns.scatterplot(x='age', y='sales', data=data)

plt.title('Sales vs. Age')

plt.xlabel('Age')

plt.ylabel('Sales')

plt.show()



*# Sales by Country*

plt.figure(figsize=(12, 8))

sns.boxplot(x='country', y='sales', data=data)

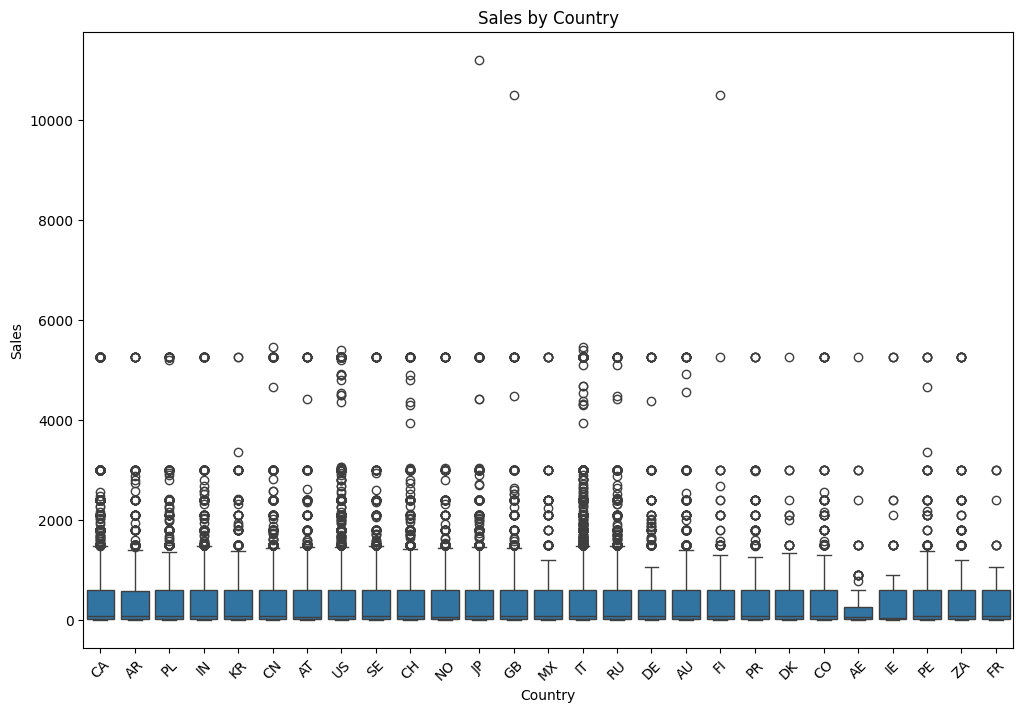
plt.title('Sales by Country')

plt.xlabel('Country')

plt.ylabel('Sales')

plt.xticks(rotation=45)

plt.show()

  
*# Sales by Language*

plt.figure(figsize=(12, 8))

sns.boxplot(x='language', y='sales', data=data)

plt.title('Sales by Language')

plt.xlabel('Language')

plt.ylabel('Sales')

plt.xticks(rotation=45)

plt.show()

