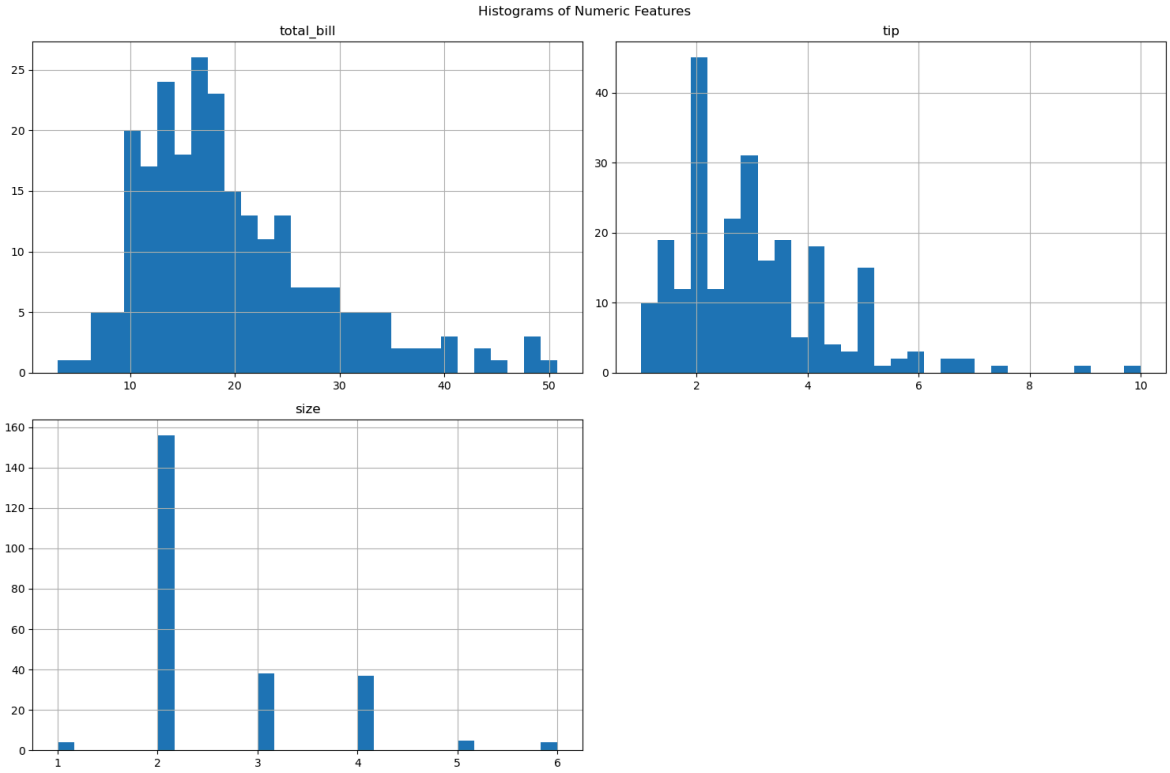
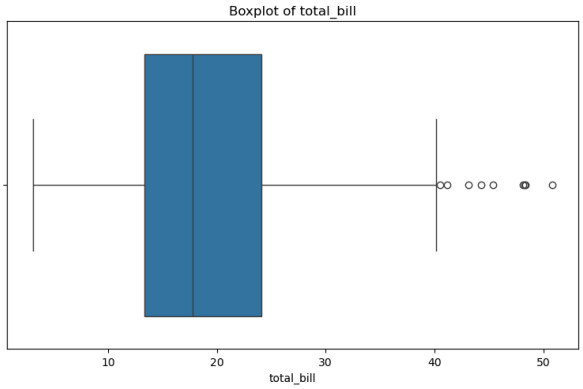
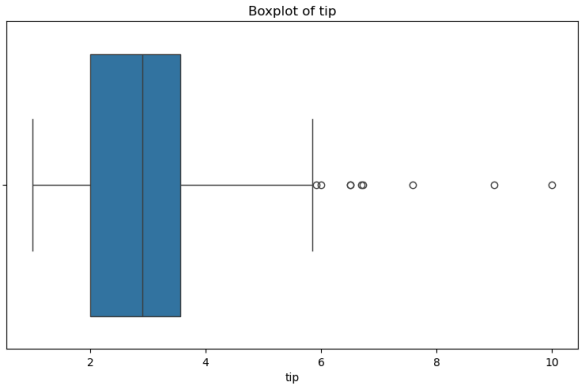
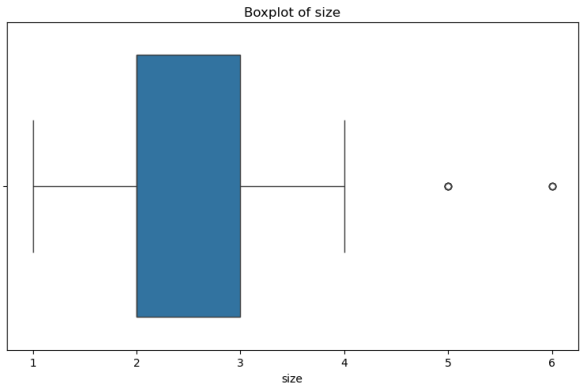
In [5]:
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.express as px

# Load dataset - Replace 'your_dataset.csv' with the actual path to your CSV fil
# For example: df = pd.read_csv('data/my_actual_dataset.csv')
# Or use a sample dataset from seaborn for testing
df = sns.load_dataset('tips')  # Using a sample dataset from seaborn

# Summary statistics
print("Summary Statistics:")
print(df.describe(include='all'))
print("\nMissing Values:")
print(df.isnull().sum())

# Histograms for numerical columns
df.hist(figsize=(15,10), bins=30)
plt.suptitle('Histograms of Numeric Features')
plt.tight_layout()
plt.show()

# Boxplots for numeric columns
plt.figure(figsize=(15,10))
for i, col in enumerate(df.select_dtypes(include='number').columns):
    plt.subplot(2, len(df.select_dtypes(include='number').columns)//2 + 1, i+1)
    sns.boxplot(data=df, x=col)
    plt.title(f'Boxplot of {col}')
plt.tight_layout()
plt.show()

# Correlation matrix heatmap
plt.figure(figsize=(10,8))
sns.heatmap(df.corr(), annot=True, cmap='coolwarm')
plt.title('Correlation Matrix')
plt.show()

# Example scatter plot using Plotly
fig = px.scatter(df, x=df.columns[0], y=df.columns[1], color=df.columns[-1])
fig.show()
```

```
Summary Statistics:
        total_bill         tip    sex smoker  day    time        size
count   244.000000  244.000000    244    244  244     244  244.000000
unique         NaN         NaN      2      2    4       2         NaN
top            NaN         NaN   Male     No  Sat  Dinner         NaN
freq           NaN         NaN    157    151   87     176         NaN
mean     19.785943    2.998279    NaN    NaN  NaN     NaN    2.569672
std       8.902412    1.383638    NaN    NaN  NaN     NaN    0.951100
min       3.070000    1.000000    NaN    NaN  NaN     NaN    1.000000
25%      13.347500    2.000000    NaN    NaN  NaN     NaN    2.000000
50%      17.795000    2.900000    NaN    NaN  NaN     NaN    2.000000
75%      24.127500    3.562500    NaN    NaN  NaN     NaN    3.000000
max      50.810000   10.000000    NaN    NaN  NaN     NaN    6.000000

Missing Values:
total_bill    0
tip           0
sex           0
smoker        0
day           0
time          0
size          0
dtype: int64
```




Boxplot of total_bill
10
20
30
40
50
total_bill


Boxplot of tip
2
4
6
8
10
tip


Boxplot of size
1
2
3
4
5
6
size