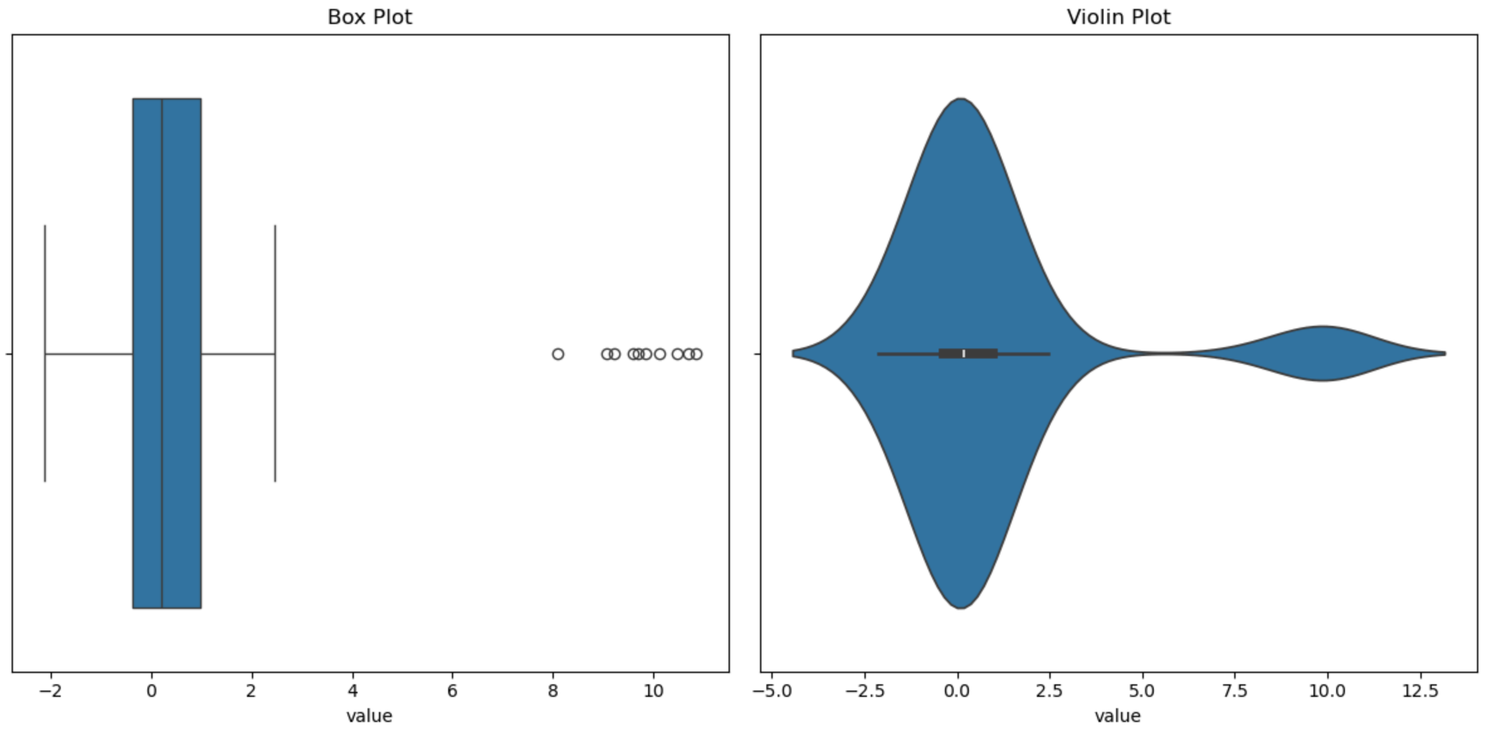
In [1]:
```
#24/05/2025

import numpy as np

#sample data (simple list os data points)
data = [10,12,14,13,100,87,5,15,16,14,13,11]

#step 1: Calculate Q1 (25th percentile) and Q3 (75th percentile)
Q1 = np.percentile(data, 25)
Q3 = np.percentile(data, 75)

#step 2: Calculate IQR (Interquartile Range)
IQR = Q3 - Q1

#step 3: Determine the lower and upper bounds for outliers
lower_bound = Q1 - 1.5 * IQR
upper_bound = Q3 + 1.5 * IQR

#step 4: Find the outliers (values outside the lower and upper bounds)
outliers = [x for x in data if x < lower_bound or x > upper_bound]

#output the results
print(f"Q1 (25th percentile): {Q1}")
print(f"Q3 (75th percentile): {Q3}")
print(f"IQR: {IQR}")
print(f"Lower Bound: {lower_bound}")
print(f"Upper Bound: {upper_bound}")
print(f"Outliers: {outliers}")
```

```
Q1 (25th percentile): 11.75
Q3 (75th percentile): 15.25
IQR: 3.5
Lower Bound: 6.5
Upper Bound: 20.5
Outliers: [100, 87, 5]
```

In [6]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

#create a sample dataset with some outliers
np.random.seed(10)
data = pd.DataFrame({
    'value':np.concatenate([np.random.normal(0, 1, 100), np.random.normal(10, 1, 10)])
})

#calculate
Q1 = data['value'].quantile(0.25)
Q3 = data['value'].quantile(0.75)
IQR = Q3 - Q1

#Define outlier thresholds
lower_bound = Q1 - 1.5 * IQR
upper_bound = Q3 + 1.5 * IQR

#Identify outliers
outliers = data[(data['value']<lower_bound) | (data['value']<upper_bound)]
#print outliers
print(f"Outliers based on Box Plot criteria:/n{outliers}")

#create a box
plt.figure(figsize=(12,6))

#Box plot
plt.subplot(1,2,1)
sns.boxplot(x=data['value'])
plt.title('Box Plot')

#violin plot with box plot inside for comparison
plt.subplot(1,2,2)
sns.violinplot(x=data['value'])
plt.title('Violin Plot')

plt.tight_layout()
plt.show()
```

```
Outliers based on Box Plot criteria:/n       value
0    1.331587
1    0.715279
2   -1.545400
3   -0.008384
4    0.621336
..        ...
95   0.918269
96  -0.482093
97   0.089588
98   0.826999
99  -1.954512

[100 rows x 1 columns]
```


Box Plot
Violin Plot
−2
0
2
4
6
8
10
value
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0
12.5
value

In [ ]: