TASK - 4

Create visualizations to understand the distribution of variables, identify outliers, and check for correlations between variables.

1. **Imports and Setup**

```
In [1]:  ▶ import numpy as np
           import pandas as pd
           import seaborn as sns
           import matplotlib.pyplot as plt
```

2. **Generate a Sample Dataset**

```
In [2]:  ▶ np.random.seed(42)
           data = {
               'Variable_A': np.random.normal(50, 15, 100),
               'Variable_B': np.random.exponential(2, 100),
               'Variable_C': np.random.uniform(0, 100, 100),
               'Variable_D': np.random.normal(30, 5, 100)
           }
           df = pd.DataFrame(data)
```
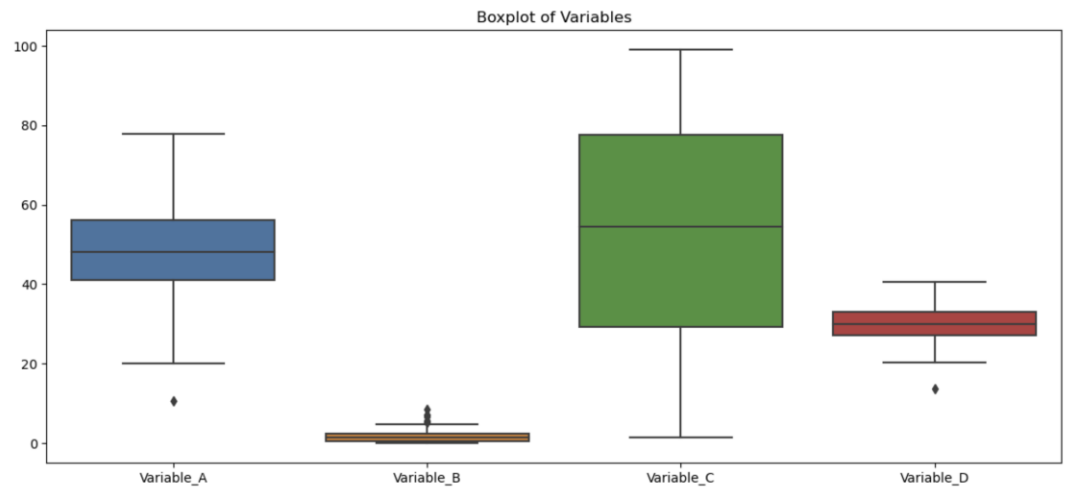
3. Create Visualizations
   a) Histograms

```
In [3]:  ▶ plt.figure(figsize=(14, 6))
           for i, column in enumerate(df.columns, 1):
               plt.subplot(2, 2, i)
               sns.histplot(df[column], kde=True)
               plt.title(f'Histogram of {column}')
           plt.tight_layout()
           plt.show()
```
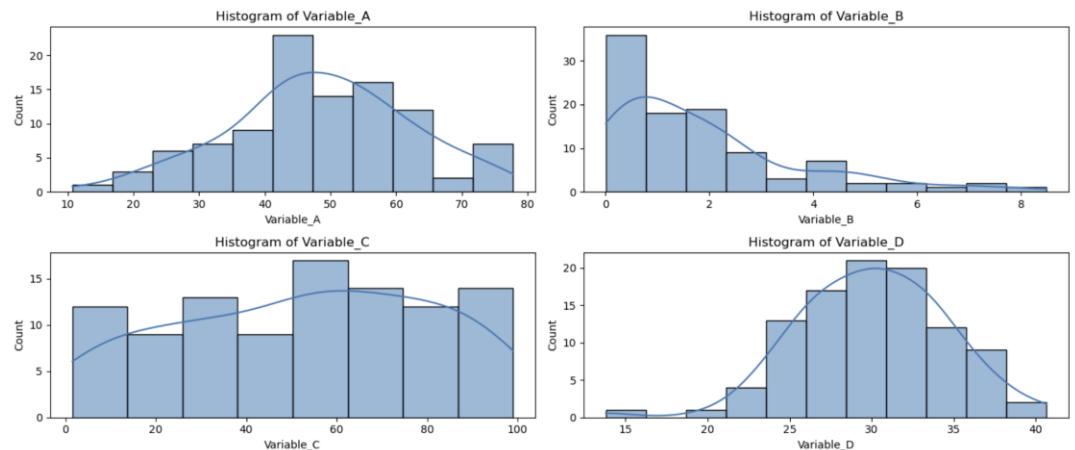
b) Boxplots

```
In [4]:  ▶ plt.figure(figsize=(14, 6))
          sns.boxplot(data=df)
          plt.title('Boxplot of Variables')
          plt.show()
```


Boxplot of Variables

c) Pair Plot (Scatter Plot Matrix)

```
sns.pairplot(df)
plt.suptitle('Pair Plot of Variables', y=1.02)
plt.show()
```

```
C:\Users\dell\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWarning: The figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```

Pair Plot of Variables