# What is Seaborn?

Seaborn is a Python data visualization library based on Matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.
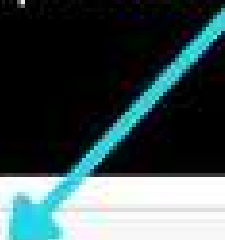
## Installation

```
pip install seaborn
```

# Basic Plotting with Seaborn

First, let's import the necessary **libraries** and load a sample **dataset**.

```python
import seaborn as sns
import matplotlib.pyplot as plt

# Load a sample dataset
tips = sns.load_dataset('tips')
```
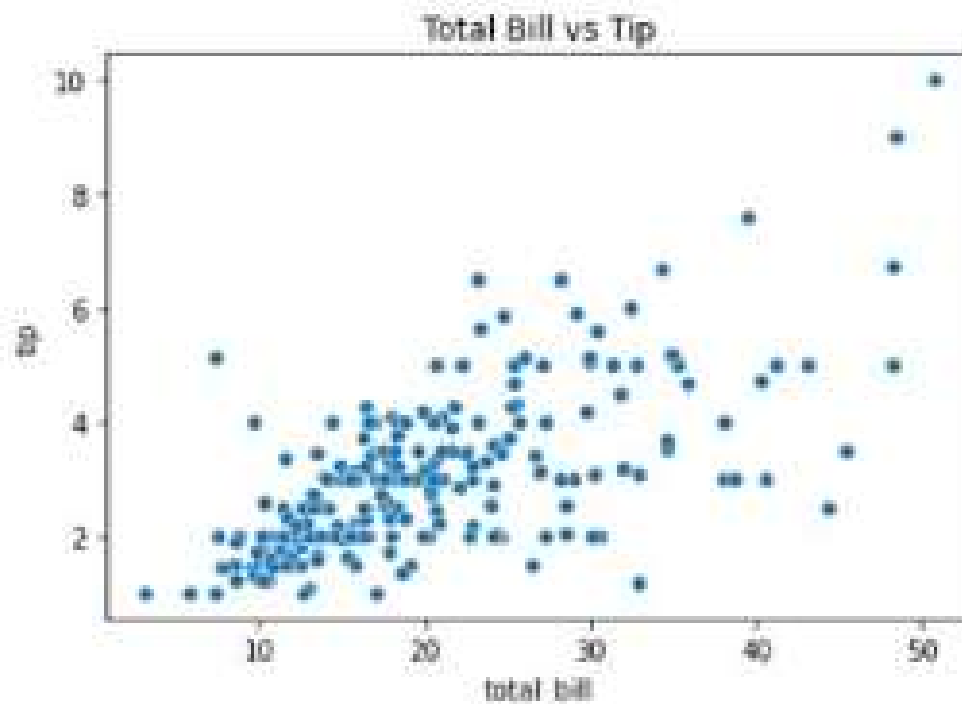
In [2]:
```python
tips.head(3)
```

Out[2]:

|   | total_bill | tip | sex | smoker | day | time | size |
|---|------------|------|--------|--------|-----|--------|------|
| 0 | 16.99 | 1.01 | Female | No | Sun | Dinner | 2 |
| 1 | 10.34 | 1.66 | Male | No | Sun | Dinner | 3 |
| 2 | 21.01 | 3.50 | Male | No | Sun | Dinner | 3 |

# Scatter Plot

A scatter plot displays values for two variables as a collection of points.
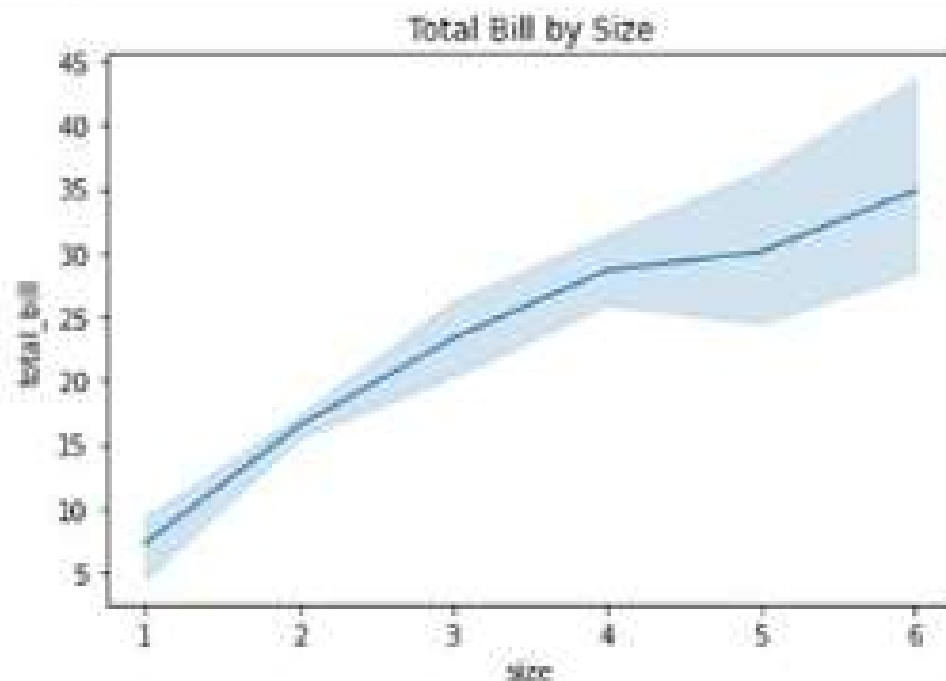
```
In [3]:   1  # Scatter plot
          2  sns.scatterplot(data=tips, x='total_bill', y='tip')
          3  plt.title('Total Bill vs Tip')
          4  plt.show()
```



Total Bill vs Tip

# Line Plot

A line plot displays information as a series of data points connected by straight line segments.

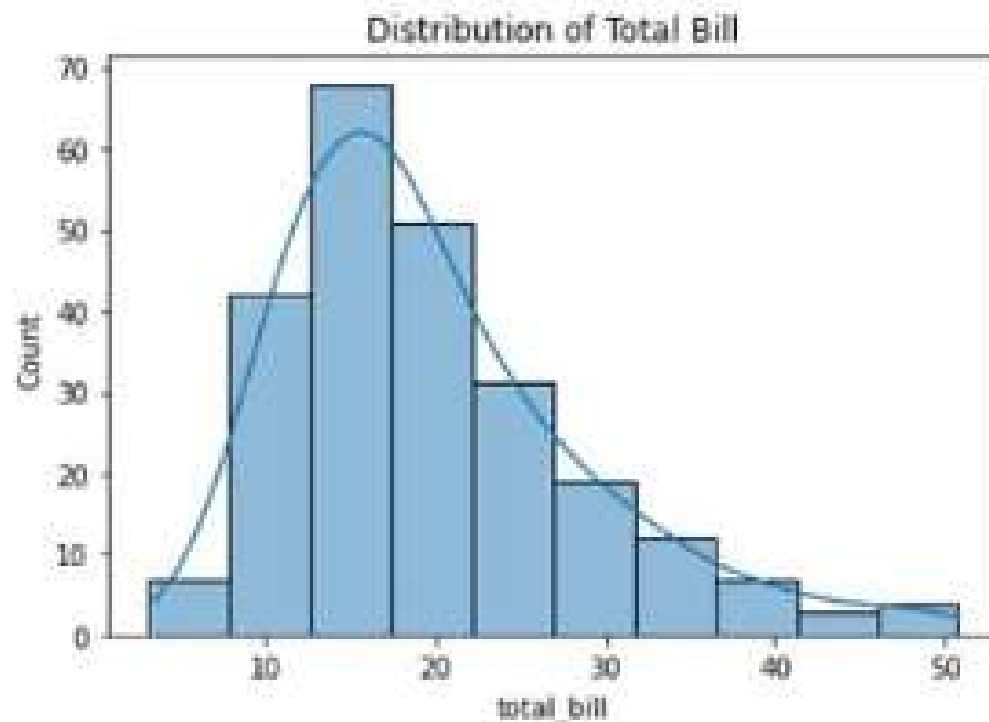```
In [4]:   1  # Line plot
          2  sns.lineplot(data=tips, x='size', y='total_bill')
          3  plt.title('Total Bill by Size')
          4  plt.show()
```



Total Bill by Size

# Histogram

*A histogram is an accurate representation of the distribution of numerical data.*
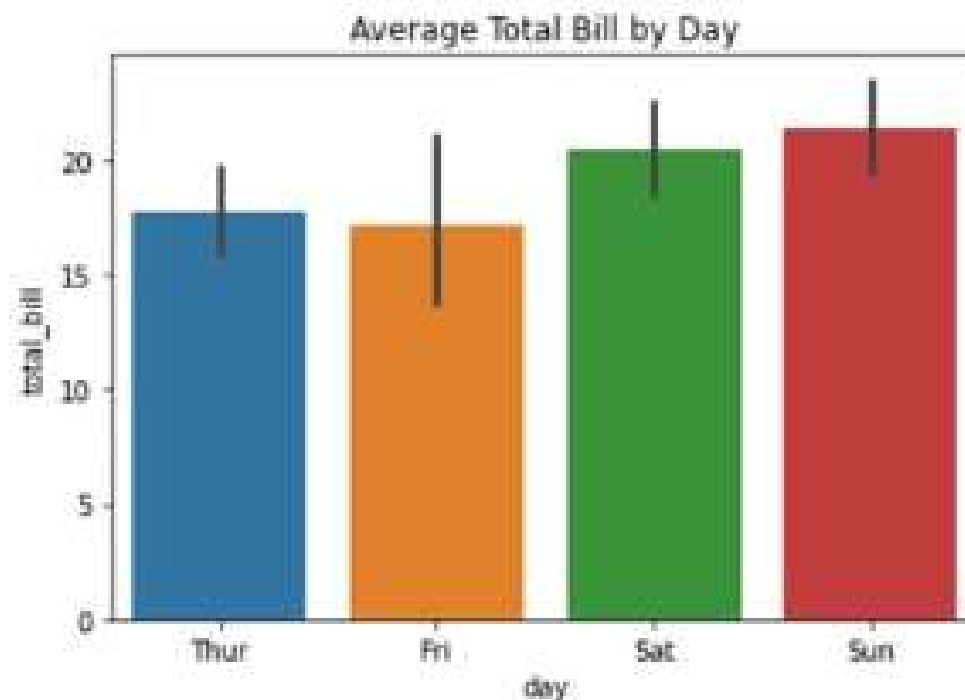
```
In [5]:   1  # Histogram
          2  sns.histplot(data=tips, x='total_bill', bins=10, kde=True)
          3  plt.title('Distribution of Total Bill')
          4  plt.show()
```



Distribution of Total Bill

# Bar Plot

*A bar plot is a way of summarizing a set of categorical data.*

```
In [6]:   1  # Bar plot
          2  sns.barplot(data=tips, x='day', y='total_bill')
          3  plt.title('Average Total Bill by Day')
          4  plt.show()
```



Average Total Bill by Day

# *List of important Seaborn functions and plots*

- **sns.scatterplot()**: Creates a scatter plot for two continuous variables.
- **sns.lineplot()**: Creates a line plot to show trends over time or a sequential variable.
- **sns.histplot()**: Plots a histogram to show the distribution of a single variable.
- **sns.barplot()**: Plots a bar plot to show the relationship between a categorical variable and a continuous variable.
- **sns.boxplot()**: Creates a box plot to show the distribution of a continuous variable and detect outliers.
- **sns.violinplot()**: Combines a box plot and a kernel density plot to show the distribution of a continuous variable.
- **sns.pairplot()**: Creates a matrix of scatter plots to show relationships between multiple variables.
- **sns.heatmap()**: Plots a heatmap to visualize matrix-like data, such as a correlation matrix.
- **sns.distplot()**: (Deprecated in favor of sns.histplot and sns.kdeplot) Combines a histogram and kernel density estimate plot.
- **sns.kdeplot()**: Plots a kernel density estimate to show the distribution of a single variable.
- **sns.jointplot()**: Plots a scatter plot with histograms for the x and y axes to show the relationship between two variables.
- **sns.lmplot()**: Creates a scatter plot with a linear regression model fit.
- **sns.catplot()**: Creates a categorical plot, combining several types like strip, swarm, and box plots.