SEABORN

**What is Seaborn?**

**Seaborn** is a Python **data visualization library** built on top of **Matplotlib**. It provides a **high-level interface** for creating visually appealing and informative statistical graphics.

**🔹 Why Use Seaborn?**

✔ **Simplifies complex visualizations** with fewer lines of code.  
✔ **Built-in support for Pandas DataFrames**, making it easy to visualize tabular data.  
✔ **Beautiful and customizable themes** for better aesthetics.  
✔ **Integration with Matplotlib**, allowing further customization.

**🔹 Installing Seaborn**

If you don’t have it installed, use:

pip install seaborn

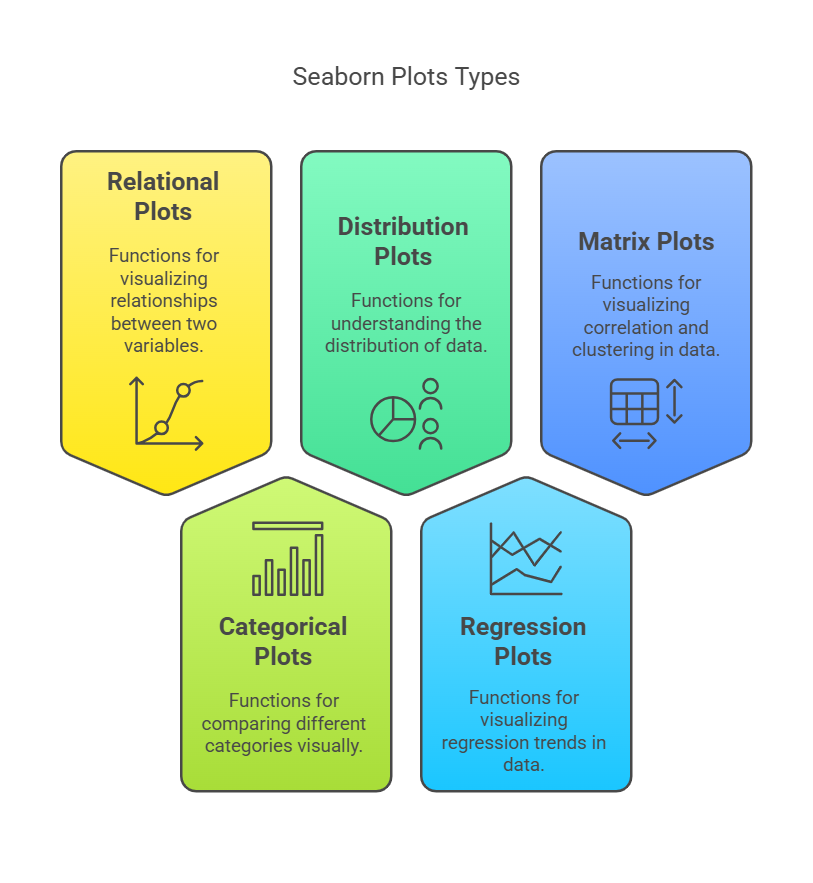
or in Jupyter Notebook:

!pip install seaborn

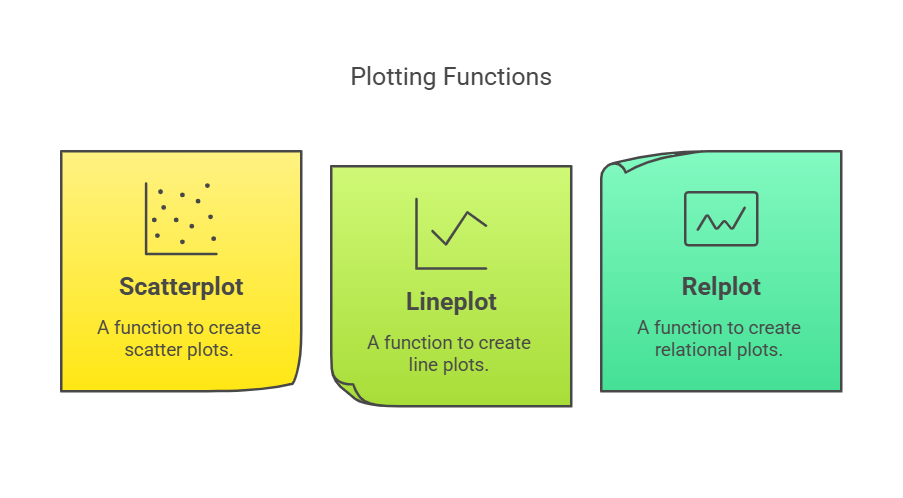
**🔹 Importing Seaborn**

import seaborn as sns

import matplotlib.pyplot as plt



# Relational Plots



**1. Scatter Plot (sns.scatterplot)**

* **Purpose**: Shows the relationship between two continuous variables, with each point representing a pair of values.
* **Analysis Type**: Bivariate
* **Types of Variables:**
  + **Independent Variable (X-axis)**: Continuous (e.g., height, temperature, income)
  + **Dependent Variable (Y-axis)**: Continuous (e.g., weight, sales, expenses)
* **Examples:**
  + **Height vs. Weight** → Analyzing body mass trends
  + **Marketing Spend vs. Revenue** → Understanding the impact of advertising
  + **Temperature vs. Ice Cream Sales** → Warmer weather increases sales

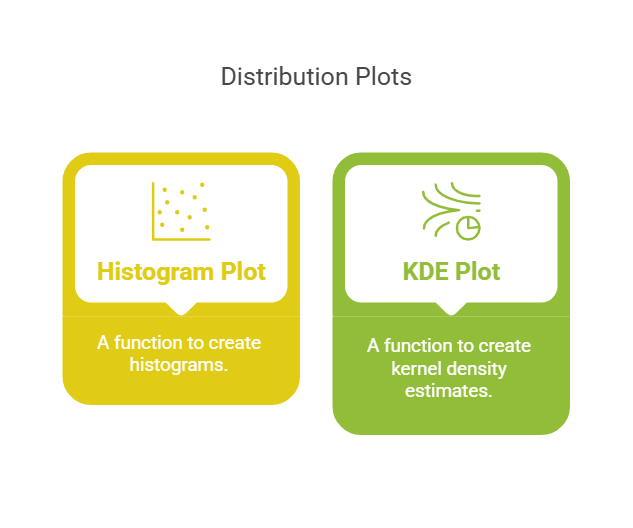
**2. Line Plot (sns.lineplot)**

* **Purpose**: Illustrates trends over time or ordered categories by connecting data points with a line.
* **Analysis Type**: Bivariate
* **Types of Variables:**
  + **Independent Variable (X-axis)**:
    - **Continuous** (e.g., time, temperature, stock prices)
    - **Ordered Categorical** (e.g., months of the year, age groups)
  + **Dependent Variable (Y-axis)**: Continuous (e.g., revenue, website traffic, sensor readings)
* **Examples:**
  + **Stock Prices Over Time** → Tracking fluctuations
  + **Website Traffic Over Days** → Analyzing daily user visits
  + **Monthly Rainfall Over a Year** → Observing seasonal trends

**3. Relational Plot (sns.relplot)**

* **Purpose**: A figure-level function for creating scatter plots or line plots with additional options for facets and hue.
* **Usage**: Useful for visualizing multiple relationships by splitting data into subplots based on categorical variables.
* **Examples:**
  + **Sales vs. Advertisement Spend, separated by Region**
  + **Temperature vs. Energy Consumption, colored by City**

# Distribution Plots



**1. Histogram (sns.histplot)**

* **Purpose**: Displays the distribution of a single continuous variable by dividing the data into bins and showing the frequency of data points in each bin.
* **Analysis Type**: Univariate
* **Examples:**
  + **Distribution of Exam Scores**
  + **Income Distribution in a Country**

**2. Kernel Density Estimate (KDE) Plot (sns.kdeplot)**

* **Purpose**: Estimates the probability density function of a continuous variable, providing a smooth curve that shows the distribution of data.
* **Analysis Type**: Univariate
* **Examples:**
  + **Salary Distribution in a Company**
  + **Height Distribution of Athletes**

**3. Histogram with KDE (sns.histplot, kde=True)**

* **Purpose**: Combines a histogram and KDE plot to visualize the distribution and density of a continuous variable.
* **Examples:**
  + **Distribution of Car Prices**

**4. Distribution Plot (sns.displot)**

* **Purpose**: A figure-level function for drawing distribution plots onto a FacetGrid.
* **Examples:**
  + **Income Distribution by Region**

# Matrix Plots

**1. Heatmap (sns.heatmap)**

* **Purpose**: Displays data in a matrix format with color-coding, often used for visualizing correlation matrices or complex datasets.
* **Analysis Type**: Bivariate or Multivariate
* **Examples:**
  + **Correlation Matrix in a Machine Learning Dataset**
  + **Monthly Sales of Products by Region**

# Categorical Plots



**1. Box Plot (sns.boxplot)**

* **Purpose**: Displays the distribution of a continuous variable through quartiles, highlighting the median, interquartile range, and outliers.
* **Analysis Type**: Bivariate
* **Examples:**
  + **Salary Distribution by Job Title**
  + **Test Scores Across Different Schools**

**2. Violin Plot (sns.violinplot)**

* **Purpose**: Combines a box plot with a KDE plot to show the distribution and probability density of a continuous variable across different categories.
* **Analysis Type**: Bivariate
* **Examples:**
  + **Customer Spending Across Different Age Groups**

**3. Bar Plot (sns.barplot)**

* **Purpose**: Represents the average value of a continuous variable for different categories.
* **Analysis Type**: Bivariate
* **Examples:**
  + **Average Monthly Salary by Industry**

**4. Count Plot (sns.countplot)**

* **Purpose**: Shows the count of observations in each categorical bin.
* **Analysis Type**: Univariate
* **Examples:**
  + **Number of Students in Different Majors**

**5. Swarm Plot (sns.swarmplot)**

* **Purpose**: Shows all data points for a categorical variable, avoiding overlap.
* **Analysis Type**: Bivariate
* **Examples:**
  + **Annual Income by Gender**

**6. Strip Plot (sns.stripplot)**

* **Purpose**: Displays individual data points for a categorical variable, with optional jitter to prevent overlap.
* **Examples:**
  + **Reaction Times by Experimental Group**

**7. Categorical Plot (sns.catplot)**

* **Purpose**: A figure-level function for creating various categorical plots (e.g., strip, box, violin) with faceting capabilities.