Seaborn

What is Seaborn?

Seaborn is a Python data visualization library based on Matplotlib.

It provides:

- Beautiful default styles
- High-level functions for **statistical plots**
- Easy integration with **Pandas DataFrames**

Installation and Import

```
pip install seaborn
import seaborn as sns
import matplotlib.pyplot as plt
import pandas as pd
```

Seaborn also comes with **built-in datasets** (like *tips*, *iris*, *titanic*):

```
data = sns.load_dataset("tips")
print(data.head())
```

---

1. Basic Scatter Plot

```
sns.scatterplot(x="total_bill", y="tip", data=data)
plt.title("Scatter Plot: Total Bill vs Tip")
plt.show()
```

---

2. Line Plot

```
sns.lineplot(x="size", y="tip", data=data)
plt.title("Line Plot Example")
plt.show()
```

You can also use a built-in dataset:

```
flights = sns.load_dataset("flights")
sns.lineplot(x="year", y="passengers", data=flights)
plt.title("Number of Passengers Over Years")
plt.show()
```

---

3. Bar Plot

```
sns.barplot(x="day", y="total_bill", data=data)
plt.title("Average Bill per Day")
plt.show()
```

You can customize:

```
sns.barplot(x="day", y="total_bill", data=data, palette="coolwarm", estimator=sum)
plt.title("Total Bill by Day")
plt.show()
```

4. Count Plot (Category Frequency)

```
sns.countplot(x="day", data=data)
plt.title("Count of Records per Day")
plt.show()
```

5. Box Plot

Shows data distribution with median and outliers.

```
sns.boxplot(x="day", y="total_bill", data=data)
plt.title("Box Plot of Total Bill by Day")
plt.show()
```

Multiple category boxplot:

```
sns.boxplot(x="day", y="total_bill", hue="sex", data=data)
plt.title("Box Plot by Gender and Day")
plt.show()
```

6. Violin Plot

Combination of **boxplot** and **density plot**:

```
sns.violinplot(x="day", y="total_bill", data=data, palette="muted")
plt.title("Violin Plot Example")
plt.show()
```

7. Pair Plot

Compares **all numeric variables** in a dataset.

```
iris = sns.load_dataset("iris")
sns.pairplot(iris)
plt.show()
```

With category coloring:

```
sns.pairplot(iris, hue="species", palette="husl")
plt.show()
```

8. Heatmap

Used for showing **correlation** or matrix data.

```
corr = data.corr()
sns.heatmap(corr, annot=True, cmap="coolwarm")
plt.title("Correlation Heatmap")
plt.show()
```

9. Histogram & KDE Plot

Histogram:

```
sns.histplot(data["total_bill"], bins=10, kde=True)
```

```python
plt.title("Distribution of Total Bill")
plt.show()
```

KDE (Kernel Density Estimation):

```python
sns.kdeplot(data["tip"], fill=True, color="green")
plt.title("KDE Plot of Tips")
plt.show()
```

---

10. Joint Plot (2D relationship + distribution)

```python
sns.jointplot(x="total_bill", y="tip", data=data, kind="scatter")
plt.show()
```

Other types:

```python
kind='hex', 'reg', 'kde', 'hist'
sns.jointplot(x="total_bill", y="tip", data=data, kind="reg")
plt.show()
```

---

11. Swarm Plot

Shows individual data points with jittering.

```python
sns.swarmplot(x="day", y="total_bill", data=data)
plt.title("Swarm Plot Example")
plt.show()
```

---

12. Regression Plot

```python
sns.regplot(x="total_bill", y="tip", data=data)
plt.title("Regression Line Example")
plt.show()
```

---

13. Styling and Themes

Seaborn has built-in **themes** for better visuals:

```python
sns.set_style("whitegrid")  # Options: darkgrid, whitegrid, dark, white, ticks
sns.set_palette("coolwarm")
```

Example:

```python
sns.set_style("whitegrid")
sns.barplot(x="day", y="total_bill", data=data)
plt.title("Bar Plot with Whitegrid Style")
plt.show()
```

---

14. Complete Example

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

sns.set_theme(style="whitegrid")
```

```
data = sns.load_dataset("tips")

plt.figure(figsize=(8,5))
sns.scatterplot(x="total_bill", y="tip", hue="day", size="size", data=data, palette="viridis",
alpha=0.8)
plt.title("Tips vs Total Bill by Day and Party Size")
plt.show()
```

Common Seaborn Datasets

Dataset Name	Description
tips	Restaurant bills and tips
iris	Flower measurements
flights	Monthly airline passengers
titanic	Titanic passenger details
penguins	Penguin species data

You can load any dataset like:

sns.load_dataset("titanic").head()