DATA VISUALIZATION IN PYTHON

Data visualization is the **graphical representation of** information and data. By using visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data.

Excel Spreadsheet

Data visualization in Python is commonly achieved using libraries such as Matplotlib, **Seaborn,** Plotly, and Bokeh etc. These libraries offer a wide range of functionalities for creating various types of visualizations. Here's an overview of some popular libraries and how they are used for data visualization in Python:

**Matplotlib**: Matplotlib is one of the most widely used plotting libraries in Python. It provides a MATLAB-like interface for creating static, interactive, and animated visualizations. Matplotlib can be used to create

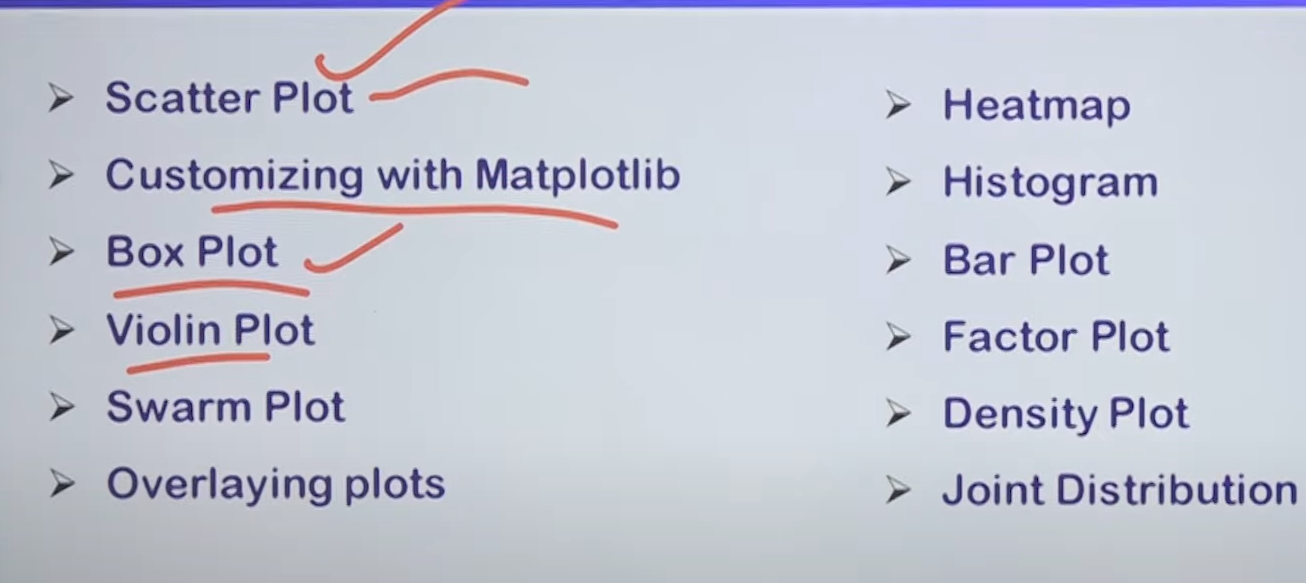
1. Line Plot
2. Scatter Plot
3. Bar Plot
4. Histogram
5. Pie Chart
6. Box Plot
7. ETC

Drawback : **too Much of code**

**Seaborn**: Seaborn is built on top of Matplotlib and provides a high-level interface for creating attractive and informative statistical graphics. It simplifies the process of creating complex visualizations like heatmaps, violin plots, pair plots, and more.

1. Bar Plot: Represents categorical data with rectangular bars, suitable for comparing quantities across categories.
2. Line Plot: Displays data points connected by straight lines, useful for visualizing trends over time.
3. Scatter Plot: Shows the relationship between two numerical variables, ideal for identifying correlations.
4. Histogram: Illustrates the distribution of a single numerical variable, useful for understanding data frequency.
5. Box Plot: Summarizes the distribution of a numerical variable, providing information about central tendency and variability.
6. Violin Plot: Combines the features of a box plot and a kernel density plot, revealing the data distribution.
7. Heatmap: Visualizes data in a matrix format using color gradients, helpful for exploring correlations in large datasets.
8. Pair Plot: Shows pairwise relationships between variables in a dataset, useful for initial exploratory data analysis.
9. Joint Plot: Displays the relationship between two variables along with their individual distributions.
10. Facet Grid: Divides a dataset into subsets and creates a grid of plots based on those subsets.
11. KDE Plot: Estimates the probability density function of a continuous variable, useful for visualizing smooth distributions.
12. Regplot: Shows the relationship between two variables along with a regression line.
13. Count Plot: Displays the count of observations in each category of a categorical variable.

Seaborn is often preferred over Matplotlib for data visualization due to its high-level interface, which allows users to create complex statistical graphics with minimal code. It offers attractive defaults, specialized plots, and seamless integration with Pandas, making it particularly well-suited for exploratory data analysis and visualizing relationships and distributions in datasets. While Matplotlib provides more flexibility and control over plot customization, Seaborn excels in simplifying the process of creating informative and visually appealing statistical visualizations.

****