

Class 3 - Seaborn

1. What Is Seaborn and How Is It Different from Matplotlib?

Seaborn is a high-level Python data visualization library built on top of Matplotlib. It is specifically designed for statistical data visualization and works seamlessly with Pandas DataFrames.

Key Differences from Matplotlib:

- Seaborn provides more attractive and informative visualizations by default
- It requires less code to create complex plots
- It focuses more on statistical relationships in data
- Matplotlib offers greater control and customization at a lower level

2. What Is a Statistical Plot?

A statistical plot is a type of visualization that represents statistical properties of data such as distribution, central tendency, variability, and relationships between variables.

Examples of Statistical Information Displayed:

- Mean and median
- Variance and spread
- Correlation between variables
- Frequency distributions

Statistical plots help in understanding the underlying structure of the data.

3. Difference Between Distplot / Histplot and Boxplot

- **Distplot / Histplot:**
Used to visualize the distribution of a single numerical variable. It shows how frequently values occur within specific ranges and helps identify skewness and data spread.
- **Boxplot:**
Used to summarize data using quartiles. It highlights the median, interquartile range, and potential outliers.

Key Difference:

Distplots focus on frequency distribution, while boxplots emphasize data summary and outlier detection.

4. What Is a Heatmap and Where Is It Used?

A heatmap is a graphical representation of data where values are represented by colors. It is commonly used to visualize correlations or patterns within a dataset.

Common Applications:

- Correlation analysis
- Feature relationship analysis
- Visualizing confusion matrices
- Identifying patterns in large datasets

Heatmaps allow quick identification of strong and weak relationships.

5. Importance of Seaborn for Data Exploration in AI Projects

Seaborn is highly useful during exploratory data analysis in AI projects as it simplifies the visualization of complex statistical relationships.

Why Seaborn Is Important in AI:

- Helps detect patterns, trends, and correlations
- Identifies outliers and anomalies
- Assists in feature selection
- Enhances understanding of data distributions

Effective data exploration using Seaborn leads to better preprocessing decisions and improved AI model performance.