

Data Visualization – Beginner Learning Notes

Bar Plot

Used to compare **average (mean)** values across categories.

Examples:

- Compare mean age by gender
- Average sales per product

Best when X is categorical and Y is numeric.

Box Plot

Shows **distribution** of data including median, quartiles, and outliers.

Examples:

- Age distribution of survivors vs non-survivors

Best for comparing medians and detecting outliers.

Line Plot

Shows **trend over time** or sequence.

Examples:

- Monthly sales trend
- Website traffic per day

Best for time-series data.

Scatter Plot

Shows **relationship** between two numeric variables.

Examples:

- Age vs fare
- Study hours vs marks

Best to detect correlation or patterns.

Histogram

Shows **frequency distribution** of a numeric variable.

Examples:

- Age distribution
- Salary distribution

Best to understand data spread and skewness.

Count Plot

Shows **number of observations** in each category.

Examples:

- Count of males vs females
- Survivors vs non-survivors

Best for categorical data counts.

Subplot

Used to draw **multiple plots in one figure**.

Examples:

- Line + Bar + Scatter together

Best for comparison and dashboards.