Week 4 – Data Visualization with Matplotlib and Seaborn
Data visualization is the process of representing data in a graphical form to
identify trends, patterns, and outliers. It uses visual elements like charts,
graphs, and maps to make data easier to understand. Matplotlib is Python's core
plotting library, offering full control over plot elements but requiring more code.
It supports essential plot types: Line Plots for trends over time, Bar Charts for
comparing categories, Histograms for showing the distribution of numeric values,
and Scatter Plots for exploring relationships between two variables. Customization
in Matplotlib includes setting titles, axis labels, legends, colors, line styles,
markers, and figure sizes. Seaborn builds on Matplotlib, providing simpler syntax
and more attractive, statistically meaningful visualizations. It includes: Count
Plots for showing the frequency of categories, Box Plots for visualizing spread and
detecting outliers, Violin Plots for combining distribution and density
information, Heatmaps for showing correlations or matrix-style data, and Pair Plots
for visualizing pairwise relationships across multiple variables, making it
especially useful for Exploratory Data Analysis (EDA). In real projects, Matplotlib
is often chosen for precise control and fundamental plotting, while Seaborn is
preferred for rapid creation of polished, data-rich visuals. Combining the two
provides both flexibility and presentation quality, ideal for building dashboards
and telling clear data stories.