Semester: 7th   **SUB CODE: CIE 405P**

1. Describing, Viewing, and Manipulating Data

- Load, explore, and manipulate datasets (e.g., handling missing values, filtering).
- **Dataset:** Titanic, Iris (Pandas, NumPy).

2. To plot the Probability Distribution Curves

- Plot different probability distributions (Normal, Binomial).
- **Dataset:** Iris, Tips (Matplotlib, Seaborn).

3. To perform Chi-Square Test on various datasets

- Conduct chi-square tests for independence between categorical variables.
- **Dataset:** Titanic, Mall Customer Segmentation (SciPy).

4. To use Python as a programming tool for the analysis of data structures

- Manipulate lists, arrays, dictionaries, and data frames.
- **Dataset:** Iris (NumPy, Pandas).

5. To perform various operations such as data storage, analysis and visualization

- Load, analyze, and visualize datasets with various plots.
- **Dataset:** COVID-19, Car Sales (Pandas, Matplotlib, Seaborn).

6. To perform descriptive statistics analysis and data visualization

- Perform summary statistics and visualize data with histograms, box plots.
- **Dataset:** Iris, Wine (Pandas, Matplotlib).

7. To perform Principal Component Analysis on datasets.

- Apply PCA for dimensionality reduction.
- **Dataset:** Breast Cancer, Wine (Scikit-learn, Pandas).

8. To perform linear regression on datasets

- Implement linear regression to model relationships.
- **Dataset:** Boston Housing, Advertising (Scikit-learn).

9. To perform Data Aggregation and Group Wise Operations

- Group data and perform aggregations (sum, mean).
- **Dataset:** Titanic, Superstore Sales (Pandas).