



## DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING - DATA SCIENCE

### Lesson Plan & Work-done Diary for AY: 2026-27, Odd Semester

| Course with Code: EDA Lab 1BAIL307A |                      |                                                                                                                                                                                                                                                                                                                                          |                            | Faculty: Mr.Chandan A      |                | Semester & Section:3 'A'   |                          |
|-------------------------------------|----------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------|----------------------------|----------------|----------------------------|--------------------------|
| Class No.                           | Date planned (DD/MM) | Topics to be covered                                                                                                                                                                                                                                                                                                                     | TLP Planned                | Date of Conduction (DD/MM) | Topics Covered | TLP Executed               | Remarks if any deviation |
| 1.                                  |                      | Load a real-world dataset such as Student Performance, Employee Information, or Sales Data using Pandas in Jupyter Notebook. Explore the dataset by displaying the first few records, identifying the number of rows and columns, examining data types, and generating a summary of the dataset to understand its structure and contents | Projector and marker board |                            |                | Projector and marker board |                          |
| 2.                                  |                      | Extract numerical attributes from a dataset and perform basic statistical analysis using NumPy and Pandas. Compute measures such as mean, median, mode, standard deviation, minimum, and maximum values, and interpret the results to understand the distribution and characteristics of the data.                                       | Projector and white board  |                            |                | Projector and marker board |                          |
| 3.                                  |                      | Analyze the quality of a dataset by identifying missing values, counting null entries in each attribute, and calculating the percentage of missing data. Generate a simple data quality report highlighting attributes that require cleaning before further analysis.                                                                    | Projector and marker board |                            |                | Projector and marker board |                          |





## DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING - DATA SCIENCE

|    |  |                                                                                                                                                                                                                                                                                    |                            |  |  |                            |  |
|----|--|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------|--|--|----------------------------|--|
| 4. |  | Perform data cleaning by handling missing values using suitable techniques such as replacement with mean, median, or mode values and removal of incomplete records. Compare the dataset before and after cleaning and analyze the impact of the cleaning process.                  | Projector and marker board |  |  | Projector and marker board |  |
| 5. |  | Detect and remove duplicate records from a dataset using Pandas functions. Compare the number of records before and after duplicate removal and discuss the importance of maintaining data consistency in analytical applications.                                                 | Projector and marker board |  |  | Projector and marker board |  |
| 6. |  | Perform data filtering, sorting, and selection operations on a dataset to answer business-oriented questions such as identifying high-performing students, top-selling products, or employees with the highest salaries. Present the extracted information in a meaningful format. | Projector and marker board |  |  | Projector and marker board |  |
| 7. |  | Apply grouping and aggregation techniques using Pandas to generate summaries from a dataset. Compute category-wise counts, averages, totals, and other statistical measures to derive meaningful insights from grouped data.                                                       | Projector and marker board |  |  | Projector and marker board |  |





## DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING - DATA SCIENCE

|     |  |                                                                                                                                                                                                                                                                                         |                            |  |  |                            |  |
|-----|--|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------|--|--|----------------------------|--|
| 8.  |  | Conduct exploratory statistical analysis on a dataset by generating descriptive statistics and identifying patterns, variations, and unusual observations. Interpret the results to understand the overall behavior of the dataset and support data-driven decision-making.             | Projector and marker board |  |  | Projector and marker board |  |
| 9.  |  | Create bar charts and pie charts using Matplotlib or Seaborn to visualize categorical data from a real world dataset. Analyze the visualizations and draw conclusions regarding trends, distributions, and comparisons among different categories.                                      | Projector and marker board |  |  | Projector and marker board |  |
| 10. |  | Develop line plots and scatter plots to visualize trends and relationships among variables in a dataset. Analyze the generated plots to identify growth patterns, variations, correlations, and potential business insights.                                                            | Projector and marker board |  |  | Projector and marker board |  |
| 11. |  | Perform correlation analysis on a dataset by computing the correlation matrix and visualizing it using a heat map. Identify strongly and weakly related attributes and discuss the significance of correlation analysis in exploratory data analysis and machine learning applications. | Projector and marker board |  |  | Projector and marker board |  |



## DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING - DATA SCIENCE

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                    |                            |  |  |                            |  |
|-----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------|--|--|----------------------------|--|
| 12. | Carry out a comprehensive Exploratory Data Analysis (EDA) micro-project on a real-world dataset from domains such as education, healthcare, finance, retail, sports, or social media. The analysis shall include dataset understanding, data quality assessment, data cleaning, statistical analysis, visualization, correlation analysis, and presentation of key findings and recommendations through a Jupyter Notebook report. | Projector and marker board |  |  | Projector and marker board |  |
|-----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------|--|--|----------------------------|--|