**Lesson Plan**

**Name of Faculty : Er Chandna Jain, Guest Faculty of CSE**

**Discipline : Computer Science and Engineering**

**Semester : 5th (odd)**

**Subject : Foundations of Data Science (PC/CDS/51-T)**

**Lesson Plan Duration : August to December-2025**

**Work Load (Lecture/Practical) per week (in hours): 03**

| Week | Theory | Topic Covered Date and Remarks |
| --- | --- | --- |
| Lecture Day | Topic (Including Assignment/Test) | Date | HOD | Director-Principal |
| 1st | 1 | Evolution of Data Science |  |  |  |
| 2 | Introduction to Data Science – Types of Data |  |  |  |
| 3 | Data Science Vs Big Data |  |  |  |
| 2nd | 4 | Concept of Big Data |  |  |  |
| 5 | Concept of Data Warehousing |  |  |  |
| 6 | Introduction to Data Mining |  |  |  |
| 3rd | 7 | Role of Data Scientist |  |  |  |
| 8 | Data Science Life Cycle |  |  |  |
| 9 | Data Science Roles – Data Science Project Stages – Data Science Applications in Various Fields – Data Security Issues |  |  |  |
| 4th | 10 | thinking in a structured way to solve data science problem statements |  |  |  |
| 11 |  |  |  |  |
| 12 | Need of Data Pre-processing |  |  |  |
| 5th | 13 | Pre-processing of data and data collection |  |  |  |
| 14 | Data Pre-Processing Overview – Data Cleaning – Data Integration and Transformation – Data Reduction – Data Discretization |  |  |  |
| 15 | Data Storage |  |  |  |
| 6th | 16 | and management |  |  |  |
| 17 | Data preparation with Sandbox for analytics. Introduction to Data Analytics/Concept of Data Analytics Types of Data Analytics |  |  |  |
| 18 | Descriptive Statistics |  |  |  |
| 7th | 19 | Mean |  |  |  |
| 20 | Standard Deviation |  |  |  |
| 21 | Skewness |  |  |  |
| 8th | 22 | and Kurtosis |  |  |  |
| 23 | Box Plots |  |  |  |
| 24 | Pivot Table |  |  |  |
| 9th | 25 | Heat Map |  |  |  |
| 26 | Correlation Statistics |  |  |  |
| 27 | ANOVA |  |  |  |
| 10th | 28 | Exploratory Data Analytics |  |  |  |
| 29 | Confidence (statistical) intervals; variances and correlations |  |  |  |
| 30 | Simple and Linear Regression – Visual Model Evaluation – Residual Plot – Distribution Plot – Polynomial Regression and Pipelines – Residual Plot – Distribution Plot – Polynomial Regression and Pipelines – In- sample Evaluation Measures – Prediction and Decision Making |  |  |  |
| 11th | 31 | Metrics for Out-of-Sample Evaluation Error – Cross Validation – Overfitting – Under fitting and Model Selection – Ridge Regression Prediction – Grid Search Testing Multiple Parameters Data handling /Data wrangling using Python Definition |  |  |  |
| 32 | Types of visualization |  |  |  |
| 33 | data visualization |  |  |  |
| 12th | 34 | Data types |  |  |  |
| 35 |  Data encoding  |  |  |  |
| 36 | mapping variables  |  |  |  |
| 13th | 37 | Conventional data visualization tools |  |  |  |
| 38 | Techniques for visual data representations |  |  |  |
| 39 | Types of data visualization |  |  |  |