SYLLABUS

ADVANCED DATA ANALYSIS USING PYTHON-SEM 5

UNIT-I: INTRODUCTION

Introduction to Data Science Stages in a Data Science Project Security Issues. Evolution of Data Science Data Science Roles -Applications of Data Science in various fields - Data

UNIT-II: DATA COLLECTION AND PRE-PROCESSING

Data Collection Strategies - Data Pre-Processing Overview - Data Cleaning - Data Integration and Transformation Data Reduction Data Discretization.

UNIT-III: EXPLORATORY DATA ANALYTICS

Descriptive Statistics Mean, Standard Deviation, Skewness and Kurtosis - Box Plots-Pivot Table - Heat Map - Correlation Statistics - ANOVA.

UNIT-IV: MODEL DEVELOPMENT

Simple and Multiple Regression - Model Evaluation using Visualization - Residual Plot - Distribution Plot - Polynomial Regression and Pipelines - Measures for In-sample Evaluation -Prediction and Decision Making.

UNIT-V: MODEL EVALUATION

Generalization Error Out-of-Sample Evaluation Metrics Cross Validation -Overfitting - Under Fitting and Model Selection - Prediction by using Ridge Regression - Testing Multiple Parameters by using Grid Search.