CO3	Implement various visualisation techniques on different data types	Ap	P	Modelling Assignments
CO4	Create prediction models using supervised techniques	Ap	P	Modelling Assignments/ / Case studies
CO5	Assess the similarity among the data using unsupervised techniques.	Ap	P	Modelling Assignments/ / Case studies
CO5	Gain insights on advanced data pre-processing techniques	U	C	Instructor-created exams / Quiz

* - Remember (R), Understand (U), Apply (Ap), Analyse (An), Evaluate (E), Create (C)

- Factual Knowledge(F) Conceptual Knowledge (C) Procedural Knowledge (P) Metacognitive Knowledge (M)

Detailed Syllabus:

Module	Unit	Content	Hrs	Marks (70)
I	**Introduction to Data Science**		**10**	**10**
	1	Introduction to Data: Types of Data – Structured Data, Semi-Structured Data, Unstructured Data and Data Streams, Statistical Data Types - Quantitative Data (Ratio and Interval Scale) and Qualitative Data (Nominal and ordinal)	2	
	2	Basic Methods of Data Analysis- Descriptive Data Analysis, Diagnostic Data Analysis or Exploratory Data Analysis, Inferential Data Analysis and Predictive Analysis.	1	
	3	Inferential Statistics: Statistical Inference, Population and Sample, Statistical Modeling, Probability Distributions – Normal, Uniform	3	
	4	Introduction to Data Science: Big Data and Data Science , Data Science Process	2	
	5	Applications of Data Science , Issues and challenges in Data Science	2	

II	**Exploratory Data Analysis**		**14**	**10**
	6	Exploratory Data Analysis (EDA) : Types of EDA - Univariate non-graphical, Multivariate non- graphical, Univariate graphical, and, Multivariate graphical. Methods of EDA – Descriptive Statistics and Data Visualization	5	
	7	Descriptive Statistics - Measures of Central Tendencies, Dispersion, Skewness and Kurtosis.	5	
	8	Data Visualization - Histograms , Box plots , Quantile-Quantile plots Scatter plots , Heat map, Bubble chart , Bar chart, Distribution plot , Pair plot , Line graph , Pie chart, Area chart	4	
III	**Data Preparation for Analysis**		**6**	**15**
	9	Data Cleaning: Handling Missing and Noisy Data, Removing outliers	2	
	10	Data Integration	1	
	11	Data Transformation: Standardization, Normalization	2	
	12	Data Reduction: Dimensionality Reduction - Principal Component Analysis	1	
1V	**Introduction to Machine Learning**		**15**	**15**
	13	Machine Learning Algorithms : Supervised Learning– Classification, Regression, Unsupervised Learning – Clustering, Dimensionality Reduction , Reinforcement Learning	3	
	14	Test /Train Split, Model Training, Bias and Variance, Overfitting and Underfitting	3	
	15	Evaluation	2	
	16	Linear Regression	1	
	17	k-Nearest Neighbors (k-NN)	1	
	18	k-means Clustering	1	
	19	Naive Bayes	1	
	20	Application of Naive Bayes - Spam Filtering	1	
	21	Singular Value Decomposition	1	
	22	Applications of Supervised, Unsupervised and Reinforcement	1	

		Learning		
V	**Hands-on Data Structures:** **Practical Applications, Case Study and Course Project**		**30**	**20**
	1	Implementation of the concepts or the algorithms learned [Binary Classification, Linear Regression, k-NN, k-means clustering, Spam Filtering]	15	
	2	Case study: Perform exploratory data analysis on a real world dataset using Python. Using appropriate Python packages parse, clean and visualize the data .	5	
	3	Capstone/Course Project: Perform an end-to-end project of the data science process.		

Mapping of COs with PSOs and POs :

	PSO1	PSO2	PSO3	PSO4	PSO 5	PSO 6	PO1	PO2	PO3	PO4	PO5	PO6
CO 1	3	-	-	-	-	1						
CO 2	1	-	2	-	-	-						
CO 3	-	-	2	-	-	-						
CO 4	-	2	3	3	-	1						
CO 5	-	2	3	3	-	1						
CO 6	-	-	-	-	-	2						

Correlation Levels:

Level	**Correlation**
-	Nil