SILVER OAK UNIVERSITY
School of Technology, Design And Computer Application
Silver Oak College of Computer Application
Master of Science Data Science
Course Name: Data Analytics Using Tools
Course Code: DSM5012C
Semester: 2nd

Prerequisite: Fundamental knowledge of data management and basic statistical concepts, along with prior experience in using spreadsheets, elementary data visualization techniques, and familiarity with analytical reasoning and structured problem-solving approaches.

Course Objective: To develop students' competencies in data analytics by enabling them to systematically acquire, process, explore, visualize, and interpret data; apply appropriate statistical and analytical techniques; design and implement predictive and prescriptive analytics workflows; and generate actionable insights to support informed, data-driven decision-making in both academic and professional environments.

Teaching Scheme:

Teaching Scheme				
L	T	P	Contact Hours	Credit
4	0	0	4	4

Content:

Unit No.	Course Contents	Teaching Hours	Weightage %
1	**Analytics Foundations** Concept of data analytics and analytical thinking, evolution and scope of data analytics in industry, role of data analytics in organizational decision making, data analytics versus data science, overview of machine learning concepts, overview of artificial intelligence concepts, types of analytics including descriptive and diagnostic analytics, types of analytics including predictive and prescriptive analytics, structured data concepts, semi-structured and unstructured data concepts, data sources including databases and transactional systems, data sources including APIs, IoT data, and log data, CRISP-DM methodology and analytics lifecycle, data warehousing concepts, data lakes, data mining fundamentals, ethics, privacy, and data governance.	14	25
2	**Data Processing and Preparation** Analytics workflow and analytics environment overview, data gathering and data ingestion concepts, data formats used in data analytics, data parsing techniques, data transformation techniques, scalability issues in data processing, real-time data processing challenges, data cleaning fundamentals, consistency checking	14	25

	methods, handling heterogeneous data, identification and treatment of missing data, data normalization, scaling, and segmentation concepts, exploratory data analysis including univariate, bivariate, and multivariate analysis, descriptive statistics, comparative statistics, and correlation analysis concepts.			
3	**Data Analytics using EXCEL** Role of spreadsheet tools in data analytics, installation and understanding of the Data Analytics Pack in EXCEL, importing data into EXCEL, data formats supported in EXCEL, data preprocessing techniques using EXCEL, data cleaning and transformation using EXCEL functions, data visualization principles in EXCEL, chart selection and graphical representation in EXCEL, measures of central tendency and dispersion using EXCEL, correlation analysis using EXCEL, regression analysis using EXCEL, statistical analysis tools and functions in EXCEL, time series analysis concepts, application of Moving Average and Exponential Moving Average	14	25	
4	**Data Analytics using KNIME and WEKA Tools** Introduction to open-source analytics tools, installation and environment overview of KNIME, data importing and supported formats in KNIME, data preprocessing and transformation using KNIME workflows, data visualization techniques in KNIME, correlation and regression analysis using KNIME, statistical method support in KNIME, implementation of classification models using KNIME, implementation of cluster-based models using KNIME, installation and environment overview of WEKA, data importing and preprocessing using WEKA, data visualization features in WEKA, classification and clustering demonstrations using WEKA, association rule mining concepts and demonstration using WEKA	14	25	

Course Outcome:

Sr.No.	CO-Statement	Unit
CO-1	Describe fundamental concepts of data analytics, types of analytics, data types, data sources, CRISP-DM methodology, and ethical governance principles.	1
CO-2	Execute data gathering, processing, cleaning, transformation, and exploratory data analysis techniques for handling missing and heterogeneous datasets.	2
CO-3	Examine datasets using statistical measures, correlation, regression, visualization, and time-series techniques in EXCEL to derive meaningful insights.	3
CO-4	Design and utilize analytics workflows using KNIME and WEKA for classification, clustering, and association mining in industry-oriented case studies.	4

Teaching & Learning Methodology:

1. Lecture-Based Learning
2. Hands-On Labs and Practical Sessions
3. Project-Based Learning
4. Group Activities and Collaborative Learning
5. Active Learning Strategies

Major Equipment:

1. Computers/Laptops with internet connectivity
2. Projector and display system for demonstrations

Books Recommended:

1. *Data Science for Business by Foster Provost and Tom Fawcett*
2. *Practical Statistics for Data Scientists by Peter Bruce, Andrew Bruce, and Peter Gedeck*

List of Open Source Software/learning website:

1. Jupyter Notebook / Google Colab
2. Kaggle (datasets and analytics notebooks)
3. RStudio (statistical analysis and visualization)
4. GitHub (project repositories and version control)
5. Coursera / edX / Udemy (online analytics learning platforms)
6. OpenRefine (data cleaning and transformation)

CO-PO-PSO MATRIX:

Co.No	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PSO1	PSO2
CO-1	2			2	2	3		3	2
CO-2				2	2			3	3
CO-3	3	3	2	3	2	3	3	3	3
CO-4	3	3	3	3	3	3	3	3	3