Mehran University of Engineering and Technology, Jamshoro
Department of Software Engineering

Title of Subject	:	**Applied Data Science (SE615)**
Discipline	:	Software Engineering (2nd Semester)
Effective	:	24MESE & onwards
Pre-requisite	:	--
Assessment	:	Theory**:** 10% Sessional, 30% Mid, 60% Final
Credit Hours	:	3 + 0 **Marks:** 100
Minimum Contact Hours:		42

Objectives of course:

- To integrate diverse datasets, apply applied analytics, and leverage data visualization techniques to derive valuable insights for complex business decisions.
- To critically analyze intricate problems and design innovative data-driven solutions with a focus on efficiency and effectiveness.
- To Cultivate a deep understanding of how to synergize data tools with domain expertise to devise contextually appropriate solutions.
- To provide hands-on exposure to complex data sets and real-world challenges, honing their skills in data manipulation, exploratory analysis, and hypothesis testing.

Course outline:

Introduction to Applied Data Science
Overview of applied data science concepts and applications.
Understanding the role of data-driven decision-making in complex business scenarios.

Data Integration and Preprocessing
Techniques for integrating and preprocessing diverse datasets.
Handling missing data and data imputation methods.
Data normalization, transformation, and feature engineering.

Exploratory Data Analysis and Visualization
Advanced data visualization techniques for gaining insights.
Exploratory data analysis to understand data distributions and patterns.
Interactive data visualization using tools like Tableau and D3.js.

Advanced Statistical Analysis
Hypothesis testing and statistical inference.
Multivariate statistical techniques for analyzing complex relationships.
Time series analysis and forecasting.

Machine Learning for Complex Problem Solving
Deep dive into advanced machine learning algorithms such as neural networks, support vector machines, and ensemble methods.
Model selection, hyperparameter tuning, and performance evaluation.
Addressing challenges in real-world data sets, such as imbalanced data and overfitting.

Big Data Technologies and Scalability
Introduction to big data concepts and distributed computing frameworks (e.g., Hadoop, Spark).
Scalable data processing and analysis using cloud platforms and parallel computing.

Ethical Considerations in Data Science

Exploring the ethical implications of data handling and decision-making.
Ensuring data privacy and security in data-driven projects.

Data Technology and Tools Showcase
Introducing a range of data technologies and tools suited for specific problem domains.
Hands-on sessions to use these tools in real-world data science projects.

BOOKS RECOMMENDED
1. David Mertz, Cleaning Data for Effective Data Science: Doing the other 80% of the work with Python, R, and command-line tools, Packt Publishing, Latest Edition.
2. Thomas Niel, Essential Math for Data Science: Take Control of Your Data with Fundamental Linear Algebra, Probability, and Statistics, O'Reilly Media, Latest Edition.

Approval:

Board of Studies:	**Resolution No. 2.3**	**Dated: 21-07-2023**
Board of Faculty:	**Resolution No. 21.9**	**Dated: 07-12-2023**
AS&RB	**Resolution No.**	
Academic Council:	**Resolution No.**	