

Semester: 1				
Programme : M.Sc. Data Science				
Course: Mathematical Foundation of Data Science and Data Visualization				
Paper code: MDTS4112			Credits: 6	
Hours/week : 8 (4Th + 3Pr)				
Category: Core/MDC/SEC/VAC: Core				
Theory / Practical / Composite: Composite				
No of Modules : 2				
Course Outcome for module 1:				
1. Remember the fundamental definitions of vectors, norms, matrix properties (rank, trace, determinant), graph terminology, and logical connectives.				
2. Understand the conceptual framework of vector spaces, the principles of linear transformations, and the structural properties of various graph types, including Eulerian and Hamiltonian graphs.				
3. Apply numerical methods for solving linear systems (Gaussian elimination, Jacobi), graph algorithms (Dijkstra, BFS, DFS), and graph coloring techniques to solve practical problems.				
4. Analyze linear algebra techniques for data representation and dimensionality reduction alongside the computational properties of shortest path and minimum spanning tree algorithms.				
5. Evaluate the validity of logical expressions using normal forms (CNF, DNF) and the effectiveness of different matrix decompositions, such as Singular Value Decomposition (SVD), for specific data science applications.				
6. Create integrated mathematical frameworks by synthesizing advanced linear algebra paradigms with discrete structures and logical reasoning to model complex data science challenges				
Prerequisites: <i>Basic knowledge about any prior course</i>				
SYLLABUS				
UNIT/Module	CONTENT	HOURS or NUMBER OF CLASSES	CO Mapping	COGNITIVE LEVEL
Module 1 / Unit 1	Vectors, norms, Vector spaces, bases, orthogonality, Gram-Schmidt orthogonalization.	6	CO1, CO2	K1, K2
Module 1 / Unit 2	Review of matrices, rank, trace, determinant, Linear transformations.	5	CO1, CO2	K1, K2
Module 1 / Unit 3	Solving linear systems: Gaussian elimination, LU, Cholesky, Jacobi, Gauss-Seidel; Projections, least squares	7	CO3	K3

Module 1 / Unit 4	Eigenvalues, eigenvectors, Spectral Decomposition; quadratic forms; Singular value decomposition.	8	CO4, CO5, CO6	K4, K5, K6
Module 2/ Unit 1	Graph Theory: Definition of Graph, Graph Terminology, Finite and Infinite graphs. Directed and undirected graphs, Degree, Isolated vertex, Pendant vertex. Null graphs. Walks: Paths and circuits. Connected and disconnected graphs, Euler's graphs, Hamiltonian paths and circuits. Planar Graph, Isomorphic graph, Graph Coloring Problem	8	CO1, CO2, CO3	K1, K2, K3
Module 2/ Unit 2	Graph Algorithms: Graph representation, Floyd's shortest Path algorithm, Trees, Minimal Spanning Tree using Kruskal Algorithm and Prim's Algorithm, Dijkstra's Algorithm, Breadth First Search (BFS), Depth First Search (DFS), Connected components. Bellman Ford Algorithm	14	CO3, CO4	K3, K4
Module 2/ Unit 3	Propositional Logic: Logical Connectives, Well-formed Formulas, Tautologies, Equivalences, Normal Forms.	4	CO1, CO5, CO6	K1, K5, K6
Data Visualisation:	<ul style="list-style-type: none"> • Introduction to R and RStudio, Importing datasets in R (CSV, Excel), Basic data handling and data frames • Univariate and Bivariate Visualisation using ggplot2: Bar plots, pie charts, Histograms, Boxplots, Scatter plots and line plots, Correlation and line of best fit • Visualisation of time series data, Introduction to animation using ganimate, Animated scatter plots and trend visualisation • Basics of R Markdown and Shiny, Creating simple interactive dashboards, Displaying plots and summaries in dashboards 			

References

1. Gilbert Strang; Linear Algebra and its Applications; Academic Press; Second Edition.
2. Hadley G. (2002): Linear Algebra. Narosa Publishing House (Reprint).

3. Mapa S. K. (2016): Higher Algebra: Abstract and Linear. Levant Books.		
4. C.L. Liu, D.P., Elements of Discrete mathematics, 2nd Edition, Tata McGraw Hill, 1985		
5. Kenneth Rosen, Discrete Mathematics and Its Applications, Sixth Edition, McGraw Hill 2006		
6. M. O. Albertson and J. P. Hutchinson, Discrete Mathematics with Algorithms, John Wiley Publication, 1988		
7. Wickham, H. ggplot2: Elegant Graphics for Data Analysis. Springer.		
8. Dalgaard, P. Introductory Statistics with R. Springer.		
9. Wilkinson, L. The Grammar of Graphics. Springer.		
10. Tabular Data Analysis with R and Tidyverse: Environmental Health. by Jean-Yves Sgro and Kristen Malecki : Click Here		
Evaluation	Theory CIA: 10 End Sem Exam: 25+25 Total: 60	Practical Continuous Assessment: 40
Paper Structure for Theory Semester Exam Module :	Short questions: 5 marks each	Long Questions: 10 Marks each
Module I	1 out of 2	2 out of 3
Module II	1 out of 2	2 out of 3

Course outcomes (COs) and Cognitive Level Mapping

COs	CO Description	Cognitive levels
CO1	Remember the fundamental definitions of vectors, norms, matrix properties (rank, trace, determinant), graph terminology, and logical connectives.	K1
CO2	Understand the conceptual framework of vector spaces, the principles of linear transformations, and the structural properties of various graph types, including Eulerian and Hamiltonian graphs.	K2
CO3	Apply numerical methods for solving linear systems (Gaussian elimination, Jacobi), graph algorithms (Dijkstra, BFS, DFS), and graph coloring techniques to solve practical problems.	K3
CO4	Analyze linear algebra techniques for data representation and dimensionality reduction alongside the computational properties of shortest path and minimum spanning tree algorithms.	K4
CO5	Evaluate the validity of logical expressions using normal forms (CNF, DNF) and the effectiveness of different matrix decompositions, such as Singular Value Decomposition (SVD), for specific data science applications.	K5
CO6	Create integrated mathematical frameworks by synthesizing advanced linear algebra paradigms with discrete structures and logical reasoning to model complex data science challenges	K6