DATA ANALYTICS

Programme Structure

Course Code	Course Title	Lecture (L) Hours Per Week	Tutorial (T) Hours Per Week	Practical (P) Hours Per Week	Total Credits
MTH2151	Optimization Techniques	3	-	-	3
MTH2251	Statistics	3	-	-	3
MTH2351	Data Mining	3	-	-	3
MTH2451	Database Management System	3	-	-	3
MTH2551	Introduction to Financial Modelling	3	-	-	3
MTH2651	Statistical Quality Control	3	-	-	3
	TOTAL				**18**

DATA ANALYTICS

Syllabus - Semester First

OPTIMIZATION TECHNIQUES

Course Code: MTH2151   **Credit Units: 03**

Course Objective:
The problems in optimization are the most common applications of mathematics. The main aim of this course is to present different methods of solving optimization problems in the areas of linear programming.

Course Contents:
Module I: Introduction to OR
Historical Development of OR, OR models and Advantages, Methodology of OR, Advantages of OR, Features of OR solution, Applications and Scope of OR

Module II: Linear Programming Problems (LPP)
Definition of LPP, General Structure of Linear Programming, Formulation of LPP, Advantages and Limitations of Linear Programming, Graphical Solutions of Linear Programming Problems. Simplex Method, Degeneracy, Duality.

Module III: Transportation Problems
Introduction to Transportation Model, Degeneracy in TP, Solution Techniques of TP, Different Methods for Obtaining Initial Basic Feasible Solutions viz. Matrix Minima Method, Row Minima Method, Column Minima Methods, Vogel's Approximation Method, Techniques for Obtaining Optimal Basic Feasible Solution – Stepping Stone and MODI Method. **Assignment Problems:** Definition, Types of Assignment Problems, Hungarian Method for AP.

Module IV: Game Theory
Concept of Game problem. Rectangular games. Pure strategy and Mixed strategy. Saddle point and its existence. Optimal strategy and value of the game. Algebraic method, Graphical method and Dominance method of solving Rectangular games. Inter-relation between the theory of Games and L.P.P

Examination Scheme:

Components	A	CT	S/V/Q	HA	EE
Weightage (%)	5	10	8	7	70

CT: Class Test, HA: Home Assignment, S/V/Q: Seminar/Viva/Quiz, EE: End Semester Examination; Att: Attendance

References:
- Operations Research, S D Sharma, KNRN Publication
- Operations Research, P.K. Gupta and D. S. Hira, Sultan Chand Ltd.
- Introduction to Operations Research, Hamdy A Taha, PHI Limited, New Delhi.
- Operations Research, J K Sharma, Macmillan Publication

Syllabus - Semester Second

STATISTICS

Course Code: MTH2251   **Credit Units: 03**

Course Objective:
Statistics plays an important role in data analytics. The main aim of this course is to help the students to read, classify and then interpret the data given to them and draw conclusions.

Module I: Data and Representation
Introduction to Statistics, Collection of data, classification and tabulation of data, Types of data: Primary data, Secondary data, Presentation of data Diagrammatic and Graphical Representation: Histogram, frequency curve, frequency polygon, Ogive curves, stem and leaf chart.

Module II: Measures of Central Tendency and Dispersion
Arithmetic Mean (A.M.) Definition, Mode, Median, Partition Values : Quartiles, Deciles and Percentiles, Box Plot, Percentile ranks. Means of transformed data, Geometric Mean (G.M.) Definition, Harmonic Mean (H.M.), Weighted Mean : Weighted A.M., G.M. and H.M.
Range, Mean deviation Mean square deviation, Variance and Standard Deviation, Combined variance (derivation for 2 groups), Combined standard deviation.

Module III: Correlation and Regression
Bivariate normal distribution, types, importance, methods of measuring correlation-scatter diagram, Karl Pearson's Coefficient of Correlation and Spearman's rank Correlation. Regression lines, Difference between regression and correlation, uses of Regression.

Module IV : Sampling theory and tests of significance
Methods of sampling (Description only): Simple random sampling with and without replacement (SRSWR and SRWOR) stratified random sampling, systematic sampling.
Tests of significance – z, t, chi-square and F.

Examination Scheme:

Components	CD	CT1	SA	A	EE
Weightage	5	10	10	5	70

CD= Class Discussion, CT 1= Class Test, SA= Short Assignments, A= Attendance. EE= External Examination

References:
- Mood, A. M., Graybill, F. A. And Boes, D.C. : Introduction to the Theory of Statistics, McGraw Hill.
- Biswas and Srivastava – A textbook, mathematical Statistics, Ist Edition, Narosa Publishing House, New Delhi.
- Gupta, S.C. and V. K. Kapoor – Mathemathical Statistics, Sultan Chand and sons.
- Hogg,R.V. and Craig, A.T: Introduction to Mathematical Statistics, McMillan.
- S. C. Gupta – Fundamentals of Statistics, Himalaya Publishing House.

DATA MINING

Course Code: MTH2351   **Credit Units: 03**

Module 1: DATA WAREHOUSING
Data warehousing Components –Building a Data warehouse –- Mapping the Data Warehouse to a Multiprocessor Architecture – DBMS Schemas for Decision Support – Data Extraction, Cleanup, and Transformation Tools –Metadata.

Module 2: BUSINESS ANALYSIS
Reporting and Query tools and Applications – Tool Categories – The Need for Applications – Cognos Impromptu – Online Analytical Processing (OLAP) – Need – Multidimensional Data Model – OLAP Guidelines – Multidimensional versus Multirelational OLAP – Categories of Tools – OLAP Tools and the Internet.

Module 3 : DATA MINING , CLUSTERING AND APPLICATIONS AND TRENDS IN DATA MINING
Introduction – Data – Types of Data – Data Mining Functionalities – Interestingness of Patterns – Classification of Data Mining Systems – Data Mining Task Primitives – Integration of a Data Mining System with a Data Warehouse – Issues –Data Preprocessing, Cluster Analysis - Types of Data – Categorization of Major Clustering Methods – Kmeans – Partitioning Methods – Hierarchical Methods - Density-Based Methods –Grid Based Methods – Model-Based Clustering Methods – Clustering High Dimensional Data - Constraint – Based Cluster Analysis – Outlier Analysis – Data Mining Applications.

Module 4 : ASSOCIATION RULE MINING AND CLASSIFICATION
Mining Frequent Patterns, Associations and Correlations – Mining Methods – Mining Various Kinds of Association Rules – Correlation Analysis – Constraint Based Association Mining – Classification and Prediction - Basic Concepts - Decision Tree Induction - Bayesian Classification – Rule Based Classification – Classification by Backpropagation – Support Vector Machines – Associative Classification – Lazy
Learners – Other Classification Methods – Prediction

Examination Scheme:

Components	A	CT	S/V/Q	HA	EE
Weightage (%)	5	10	8	7	70

CT: Class Test, HA: Home Assignment, S/V/Q: Seminar/Viva/Quiz, EE: End Semester Examination; Att: Attendance

TEXT BOOKS:
1. Alex Berson and Stephen J. Smith, " Data Warehousing, Data Mining & OLAP", Tata McGraw – Hill Edition, Tenth Reprint 2007.
2. Jiawei Han and Micheline Kamber, "Data Mining Concepts and Techniques", Second Edition, Elsevier, 2007.

REFERENCES:

1. Pang-Ning Tan, Michael Steinbach and Vipin Kumar, " Introduction To Data Mining", Person Education, 2007.
2. K.P. Soman, Shyam Diwakar and V. Ajay ", Insight into Data mining Theory and Practice", Easter Economy Edition, Prentice Hall of India, 2006.
3. G. K. Gupta, " Introduction to Data Mining with Case Studies", Easter Economy Edition, Prentice Hall of India, 2006.
4. Daniel T.Larose, "Data Mining Methods and Models", Wile-Interscience, 2006.

DATABASE MANAGEMENT SYSTEM

Course Code: MTH2451   **Credit Units: 03**

Module-I
INTRODUCTION: Introduction; An example; Characteristics of Database approach; Actors on the screen; Workers behind the scene; Advantages of using DBMS approach; A brief history of database applications; when not to use a DBMS. Data models, schemas and instances; Three-schema architecture and data independence; Database languages and interfaces; The database system environment; Centralized and client-server architectures; Classificaton of Database Management systems.

Module-II
ENTITY-RELATIONSHIP MODEL: Using High-Level Conceptual Data Models for Database Design; An Example Database Application; Entity Types, Entity Sets, Attributes and Keys; Relationship types, Relationship Sets, Roles and Structural Constraints; Weak Entity Types; Refining the ER Design; ER Diagrams, Naming Conventions and Design Issues; Relationship types of degree higher than two.

Module-III
RELATIONAL MODEL AND RELATIONAL ALGEBRA: Relational Model Concepts; Relational Model Constraints and Relational Database Schemas; Update Operations, Transactions and dealing with constraint violations; Unary Relational Operations: SELECT and PROJECT; Relational Algebra Operations from Set Theory; Binary Relational Operations : JOIN and DIVISION; Additional Relational Operations; Examples of Queries in Relational Algebra; Relational Database Design Using ER- to-Relational Mapping.

Module-IV
TRANSACTION MANAGEMENT: The ACID Properties; Transactions and Schedules; Concurrent Execution of Transactions; Lock - Based Concurrency Control; Performance of locking; Transaction support in SQL; Introduction to crash recovery; 2PL, Serializability and Recoverability; Lock Management; Introduction to ARIES; The log; Other recovery-related structures; The write-ahead log protocol; Checkpointing; Recovering from a System Crash; Media Recovery; Other approaches and interaction with concurrency control.

Examination Scheme:

Components	A	CT	S/V/Q	HA	EE
Weightage (%)	5	10	8	7	70

CT: Class Test, HA: Home Assignment, S/V/Q: Seminar/Viva/Quiz, EE: End Semester Examination; Att: Attendance

TEXT BOOKS:
1. **Fundamentals of Database Systems** – Elmasri and Navathe, 5th Edition, Addison-Wesley, 2007
2. **Database Management Systems** – Raghu Ramakrishnan and Johannes Gehrke – 3rd Edition, McGraw-Hill, 2003.

REFERENCE BOOKS:
1. **Data Base System Concepts** – Silberschatz, Korth and Sudharshan, 5th Edition, Mc-GrawHill, 2006.
2. **An Introduction to Database Systems** – C.J. Date, A. Kannan, S. Swamynatham, 8th Edition, Pearson Education, 2006.

INTRODUCTION TO FINANCIAL MODELLING

Course Code: MTH2551   **Credit Units: 03**

Prerequisites: Introduction to Probability Theory

Basic notions %G–%@ Cash flow, present value of a cash flow, securities, fixed income securities, types of markets.

Forward and futures contracts, options, properties of stock option prices, trading strategies involving options, option pricing using Binomial trees, Black %G–%@ Scholes model, Black %G–%@ Scholes formula, Risk-Neutral measure, Delta %G–%@ hedging, options on stock indices, currency options.

Examination Scheme:

Components	A	CT	S/V/Q	HA	EE
Weightage (%)	5	10	8	7	70

CT: Class Test, HA: Home Assignment, S/V/Q: Seminar/Viva/Quiz, EE: End Semester Examination; Att: Attendance

Texts / References

- D.G. Luenberger, Investment Science, Oxford University Press, Oxford, 1998.

- J.C. Hull, Options, Futures and Other Derivatives, 4th ed., Prentice-Hall, New York, 2000.

- J.C. Cox and M. Rubinstein, Options Market, Englewood Cliffs, N.J.: Prentice Hall, 1985.

- C.P Jones, Investments, Analysis and Measurement, 5th ed., John Wiley and Sons, New York, 1996.

STATISTICAL QUALITY CONTROL

Course Code: MTH2651   **Credit Units: 03**

Total quality control in an industry. Quality planning, quality conformance, quality ad-herence. Quality assurance and quality management functions.

Control charts and allied techniques. Concept of quality and meaning of control. Concept of inevitability of variation-chance and assign-able causes. Pattern of variation. Principles of rational sub-grouping.

Different types of control charts. Concept of process capability and its comparison with design specifications, CUSUM charts.

Acceptance sampling. Sampling inspection versus 100 percent inspection. Basic concepts of attributes and variables inspection. OC curve, Single, double, multiple and sequential sampling plans, Management and organisation of quality control.

Examination Scheme:

Components	A	CT	S/V/Q	HA	EE
Weightage (%)	5	10	8	7	70

CT: Class Test, HA: Home Assignment, S/V/Q: Seminar/Viva/Quiz, EE: End Semester Examination; Att: Attendance

Texts / References :

- A.J. Duncan, Quality Control and Industrial Statistics, 5th ed., Richard D. Irwin, 1986.

- E.L. Grant and R. Levenworth, Statistical Quality Control, 6th ed., McGraw-Hill, 1988.

- J.M. Juran and F. M. Grayna, Quality Planning and Analysis, Tata McGraw-Hill, 1970.

- D.C. Montgomery, Introduction to Statistical Quality Control, Wiley, 1985.

- T.P. Ryan, Statistical Methods for Quality Improvement, Wiley, New York, 2000.