GURU NANAK COLLEGE
(AUTONOMOUS)

GURU NANAK SALAI, VELACHERY, CHENNAI – 600042
(Re-accredited at 'A-Grade' by NAAC) Affiliated to University of Madras

Bachelor of Science – B.Sc. (Data Analytics)

(SEMESTER PATTERN WITH CHOICE BASED CREDIT SYSTEM)

Syllabus

(For the candidates admitted in the Academic year 2021-22 and thereafter)

PREAMBLE

This course presents an introduction into the concepts of data analysis, the role of a Data Analyst, and the tools that are used to perform functions. One will gain an understanding of the data ecosystem and the fundamentals of data analysis, such as data gathering or data mining.

Data Analytics is the most demanded professions of the decade, and the demand for data scientists who can analyze data and communicate results to inform data driven decisions has never been greater. This Professional program do not require any prior knowledge of computer science or programming languages required. This program is integrated with SAS India PVT Limited. SAS provides Hands on Training with the following modules

1. Programming for Analytics.
2. Big Data Management
3. Applied Data Mining and Machine Learning
4. Applied Statistics and Time Series Modelling
5. Optimization and Text Analytics

Also, the course will provide students with the latest job-ready tools and skills, including open source tools and libraries, Python, databases, SQL, data visualization, data analysis, statistical analysis, predictive modeling, and machine learning algorithms. One can learn data science through hands-on practice in the SAS Cloud using real data science tools and data sets. In addition, every student can earn digital badges and International Certifications from SAS recognizing your proficiency in data science.

This course will help the students to differentiate between the roles of a Data Analyst, Data Scientist, and Data Engineer. One can learn the responsibilities of a Data Analyst and exactly what data analysis entails. The course also helps the students to summarize the data ecosystem, such as databases and data warehouses. Big Data platforms such as Hadoop, Hive, and Spark are also covered in the Big Data Management. Students are also provided with a real time project.

LEARNING OUTCOME BASED CURRICULUM FRAMEWORK
From the Academic Year 2021- 22 and there after

Vision

Train the students on technologies of Data Analytics to analyze data, create interactive dashboards, and publish insights of the data. Students can begin a rewarding carrier as a Junior Data Analyst after completing the course.

Mission

Strive for good graduate education that will prepare students for careers as competent professionals in industry and academia, capable of taking on new challenges in an ever-changing world. Students will understand the patterns and trends and find the significance in solving complex problems. Encourage the transmission of technologies to aid socioeconomic growth.

PROGRAMME OUTCOMES
B.Sc. DATA ANALYTICS

The following points are the expected outcomes of the three-year B.Sc. Data Analytics Programme of Guru Nanak College:

PO1 : Manifest proficiency with Statistical analysis of data.
PO2 : Emerge the ability to construct and appraise Databased Models.
PO3 : Accomplish statistical scrutiny with competent statistical software like SAS.
PO4 : Exhibit dexterity in Data Management.
PO5 : Apply Data Science concepts and methods to solve problems in real-world situations and to transmit these solutions efficiently.

PROGRAMME SPECIFIC OUTCOMES
B.Sc. DATA ANALYTICS

PSO1: Expertise in planning, organizing, directing, and leading large-scale data analytics initiatives and businesses.
PSO2: Perform interpretive and probabilistic reasoning approaches, to fit complicated structures usingspecialized statistical tools (e.g., SAS).

COURSE STRUCTURE 2021-24 Batch onwards
B. SC. DATA ANALYTICS

Semester	Part	Course Component	Subject Name	Credits	Hours	Internal	External	Total
Semester - I	I	Language	Language - I	3	6	50	50	100
	II	English	English - I	3	4	50	50	100
	III	Core Paper-I	Programming in C++	4	6	50	50	100
	III	Core Paper-II	Practical – Programming for Analytics I	4	4	50	50	100
	III	Allied-I	Linear Algebra	5	6	50	50	100
	IV	Non Major Elective-I /Basic/ Advance Tamil	Practical – Multimedia/Basic Tamil - I / Advanced Tamil -I	2	2		100	100
	IV	Soft Skills-I	Introduction to Study Skills	3	2		100	100
Total Credits: 24 / Total Hours per week:30								
Semester - II	I	Language	Language - II	3	6	50	50	100
	II	English	English - II	3	4	50	50	100
	III	Core Paper-III	Operating Systems and Linux	4	5	50	50	100
	III	Core Paper-IV	Practical – Programming for Analytics II	4	5	50	50	100
	III	Allied-II	Statistics I	5	6	50	50	100
	IV	Non Major Elective-II /Basic/ Advance Tamil	Practical –E-Commerce Lab/Basic Tamil – II/ Advanced Tamil - II	2	2		100	100
	IV	Soft Skills-II	Life Skills	3	2		100	100
Total Credits: 24 / Total Hours per week:30								
Semester - III	III	Core Paper-V	Big Data Technologies	4	6	50	50	100
	III	Core Paper-VI	Database Management System	4	6	50	50	100
	III	Core Paper-VII	R Programming	4	6	50	50	100
	III	Core Paper-VIII	Practical: Big Data Management	4	4	50	50	100
	III	Allied-III	Operations Research	5	6	50	50	100
	IV	Soft Skills-III	Job-Oriented Skills	3	2		100	100
Total Credits: 24 / Total Hours per week:30								
Semester - IV	III	Core Paper-IX	Data Mining and Data Warehousing	4	5	50	50	100
	III	Core Paper-X	Python Programming	3	5	50	50	100
	III	Core Paper-XI	Data Structures and Algorithm	4	6	50	50	100
	III	Core Paper-XII	Practical : Applied Data Mining and Machine Learning	4	4	50	50	100
	III	Allied-IV	Statistics - II	5	6	50	50	100
	IV	Soft Skills-IV	Quantitative Aptitude	3	2		100	100
	IV	EVS	EVS	2	2		100	100
Total Credits: 25 / Total Hours per week:30								

COURSE STRUCTURE 2021-24 Batch onwards
B. SC. DATA ANALYTICS

Semester	Part	Course Component	Subject Name	Credits	Hours	Internal	External	Total
Semester - V	III	Core Paper XIII	Machine Learning Basics	4	6	50	50	100
	III	Core Paper XIV	Software Engineering	4	6	50	50	100
	III	Core Paper XV	Practical: Applied Statistics and Time Series Modelling	4	6	50	50	100
	III	Core Paper XVI	Cyber Security	4	4	50	50	100
	III	Elective-I IDE	Cloud Computing	5	6	50	50	100
	IV	Value Education	Value Education	2	2		100	100
	V	Internship	Internship	2				
Total Credits: 25/ Total Hours per week: 30								
Semester - VI	III	Core Paper-XVII	Web Design and Development	4	6	50	50	100
	III	Core Paper-XVIII	Natural Language Processing	4	6	50	50	100
	III	Core Paper-XIX	Practical – Optimization and Text Analytics	4	6	50	50	100
	III	Core Paper-XX	Mini Project	4	6	50	50	100
	III	Elective-II	Business Analytics	5	6	50	50	100
	V	Extension Activity	Participation in NSS/NCC/ROTRACT etc.,	1		-	-	-
Total Credits: 22 / Total Hours per week:30								
Grand Total Credits: 144 / Total Hours: 180								

ANNEXURE - I

Course Component	Subject Name
Elective - I	1. Cloud computing 2. Digital Logic Fundamentals
Elective - II	1. Business Analytics 2. Introduction to IoT 3. Introduction to Deep Learning

CORE I - PROGRAMMING IN C++

SUBJECTCODE:	THEORY	MARKS: 100
SEMESTER: I	CREDITS: 4	NO. OF HOURS : 90

COURSE FRAMEWORK:
This course introduces the Object Oriented Programming concepts in C++.

COURSE OUTCOME:
On completion of the course the students will be able

1. To describe the procedural and object-oriented paradigm with concepts of streams, classes, functions, data, and objects
2. To understand dynamic memory management techniques using pointers, constructors, destructors, etc.
3. To apply the concept of function overloading, operator overloading, virtual function and polymorphism
4. To classify inheritance with the understanding of early and late binding, usage of exception handling, and generic programming
5. To demonstrate the use of various OOPs concepts with the help of programs

UNIT I:
Principles of Object-Oriented Programming: A Look at Procedure Oriented Programming – Object- Oriented Programming Paradigm – Basic Concepts of Object- Oriented Programming – Benefits of OOP – Applications of OOP. Tokens, Expressions and Control Structures: Introduction – Tokens – Keywords - Identifiers and Constants – Basic Data Types – User Defined Data Types – Derived Data Types – Symbolic Constants – Type Compatibility – Declaration of Variables – Dynamic Initialization of Variables – Reference Variables. **(18 Hours)**

UNIT II:
Operators in C++ - Scope Resolution Operator – Member Dereferencing Operators – Memory Management Operators – Manipulators – Type Cast Operator – Expressions and their Types – Special Assignment Expressions – Implicit Conversions – Operator Overloading – Operator Precedence – Control Structures. Functionsin C++: Introduction – The main function – Function Prototyping – Call by Reference – Inline Functions – Default Arguments – Function Overloading– Math Library Functions. **(18 Hours)**

UNIT III:
Classes and Objects: Introduction – C Structures Revisited – Specifying a Class – Defining Member Functions – A C++ Program with Class – Making an Outside Function Inline – Arrays within a Class – Memory Allocation for Objects – Static Data Members – Static Member Functions – Arrays of Objects - Objects as Function Arguments – Friendly Functions – Returning Objects. Constructors and Destructors: Introduction – Constructor – Parameterized Constructors – Multiple Constructors in a Class – Constructors with Default Arguments – Dynamic Initialization of Objects – Copy Constructors – Constructing Two-dimensional Arrays – constant Objects – Destructors. **(18 Hours)**

UNIT IV:

Operator Overloading and Type Conversions: Introduction – Defining Operator Overloading – Overloading Unary Operators – Overloading Binary Operators – Overloading Binary Operators Using Friends – Manipulation of Strings Using Operators – Rules for Overloading Operators – Type Conversions. Inheritance: Extending Classes: Introduction – Defining Derived Classes – Single Inheritance – Making a Private Member Inheritable – Multilevel Inheritance – Multiple Inheritance – Hierarchical Inheritance – Hybrid Inheritance – Constructors in Derived Classes – Member Classes: Nesting of Classes. **(18 Hours)**

UNIT V:

Pointers, Virtual Functions and Polymorphism: Introduction – Pointers – Pointers to Objects – this Pointer Polymorphism - Pointers to Derived Classes – Virtual Functions – Pure Virtual Functions. Managing Console I/O Operations: Introduction – C++ Streams – C++ Stream Classes – Unformatted I/O Operations – Formatted Console I/O Operations – Managing Output with Manipulators. Working with Files: Introduction – Classes for File Stream Operations – Opening and Closing a File – Detecting end-of-file – More about Open(): Files Modes – File Pointers and Their manipulations – Sequential Input and Output Operations – Updating a File – Error Handling During File Operations. **(18 Hours)**

PRESCRIBED BOOKS:

1. E. Balagurusamy,1995, Object Oriented Programming with C++, Tata McGraw Hill Publishing Company Ltd.

REFERENCE BOOKS:

1. Robert Lafore, Object Oriented Programming in Microsoft C++, Galgotia publication.
2. H.Schildt, C++,1998,The Complete Reference-1998-TMH Edition, 1998

WEBSITES:

1. http://www.cprogramming.com/algorithms-and-data-structures.html

QUESTION PAPER PATTERN:

Section	Question Component	Numbers	Marks	Total
A	**Definition/Principle** Answer any 10 out of 12 questions (each in 50 words)	1-12	3	**30**
B	**Short Answer** Answer any 5 out of 7 questions (each in 300 words)	13-19	6	**30**
C	**Essay** Answer any 4 out of 6 questions (each in 600 words)	20-25	10	**40**
TOTAL MARKS				**100**

DISTRIBUTION OF QUESTIONS:

Section	Units	No. of Questions	
		Theory	**Problems**
A	Unit – 1	2	
	Unit – 2	3	
	Unit – 3	3	
	Unit – 4	2	
	Unit – 5	2	
B	Unit – 1	1	
	Unit – 2	1	
	Unit – 3	2	
	Unit – 4	2	
	Unit – 5	1	
C	Unit – 1	2	
	Unit – 2	1	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	

CORE II PRACTICAL - PROGRAMMING FOR ANALYTICS I

SUBJECT CODE:	PRACTICAL	MARKS: 100
SEMESTER: I	CREDITS: 4	NO. OF HOURS : 60

COURSE FRAMEWORK:

This course is for users who want to learn how to write SAS programs to access, explore, prepare, and analyze data. It is the entry point to learning SAS programming for data science, machine learning, and artificial intelligence. In addition, the course focuses on data manipulation techniques using the SAS DATA step and procedures to access, transform, and summarize data.

COURSE OUTCOME:

On completion of the course the students will be able

1. To Demonstrate Access SAS, Microsoft Excel
2. To explore and validate data
3. To analyze and report on data
4. To understand and create custom formats
5. To implement restructure tables

EXERCISES:

- Access SAS, Microsoft Excel, and text data.
- Explore and validate data.
- Prepare data by sub-setting rows and computing new columns.
- Analyze and report on data.
- Create an accumulating column and process data in groups.
- Manipulate data with functions.
- Convert column type.
- Create custom formats.
- Concatenate and merge tables.
- Process repetitive code.
- Restructure tables.

REQUIRED TEXTS

1. SAS Programming Essentials, SAS Institute, Cary.
2. SAS Data Manipulation Techniques, SAS Institute, Cary.

REFERENCE TEXTS

1. Step-by-Step Programming with Base SAS, SAS Institute, Cary.
2. The Little SAS Book: A Primer, Sixth Edition, SAS Institute, Cary.

SAS INTERNAL REFERENCE

LAB1; C1 – PG1V2, C2 – PG2V2; Exam Weightage – 50:50.

ALLIED - I

LINEAR ALGEBRA

SUBJECTCODE:	THEORY	MARKS: 100
SEMESTER: I	CREDITS: 5	NO. OF HOURS : 90

COURSE FRAMEWORK:

This course will cover the analysis and implementation of algorithms used to solve linear algebra problems in practice.

COURSE OUTCOME:

1. Compute the eigen values and eigen vectors. Apply Cayley Hamilton theorem
2. Solve the linear transformations
3. Solve Algebraic equations numerically by Gauss seidel and Gauss Jordan methods
4. Find the inverse of the matrix using Gauss Elimination method.
5. Evaluate the singular value decompostions

UNIT I:

Matrix and Basic properties of matrix & vectors: Matrix, scalar multiplication, linear transformation, transpose, conjugate, rank, determinant, Inner and outer products, matrix multiplication rule and various algorithms, matrix inverse, square matrix, identity matrix, triangular matrix, idea about sparse and dense matrix, UNIT vectors, symmetric matrix, Hermitian, skew-Hermitian and Unitary matrices. **(18 Hours)**

UNIT II:

Special matrices and Vector Space: Matrix factorization concept/LU decomposition, Gaussian/Gauss-Jordan elimination, solving Ax=b linear system of equation, vector space, subspaces, basis, span, dimension of subspace, orthogonality, orthonormality, linear least square, Eigenvalues, eigenvectors, and diagonalization. **(18 Hours)**

UNIT III:

Linear Transformations - Definition and example of linear transformation, Null space, range, rank and nullity of linear transformation, matrix representation of a linear transformation, dual space, dual basis, double dual, composition of linear transformation and matrix multiplication. **(18 Hours)**

UNIT IV:

Transformation Diagonalization: Diagonalizability, matrix Limits and Markov Chains and the Caley- Hamilton Theorem. **(18 Hours)**

UNIT V: Numerical Linear Algebra: Regularization, Principal Component Analysis, Singular-Value Decomposition, Latent Semantic Analysis. **Case Studies**: Recommender Systems, Page Ranking. **(18 Hours)**

PRESCRIBED BOOKS:

1. Stephen H. Friedberg, Arnold J. Insel, Lawrence E. Spence, Linear Algebra, 4th Ed., PrenticeHall ofIndia Pvt. Ltd., New Delhi, 2004.
2. David C. Lay, Linear Algebra and its Applications, 3rd Ed., Pearson Education Asia, Indian Reprint,2007.

REFERENCE BOOKS:

1. S. Lang, Introduction to Linear Algebra, 2nd Ed., Springer, 2005.
2. Gilbert Strang, Linear Algebra and its Applications, Thomson, 2007.
3. A.I. Kostrikin, Introduction to Algebra, Springer Verlag, 1984.
4. Richard Bronson, Theory and Problems of Matrix Operations, Tata McGraw Hill, 1989.

QUESTION PAPER PATTERN:

Section	Question Component	Numbers	Marks	Total
Section A	**Definition/Principle** Answer any 10 out of12 questions	1 – 12	3	**30**
Section B	**Short Answer** Answer any 5 out of 7 questions	13–19	6	**30**
Section C	**Essay** Answer any 4 out of 6 questions	20– 25	10	**40**
			TOTAL	**100**

DISTRIBUTION OF QUESTIONS:

Sections	Units	No. of Questions	
		Theory	**Problems**
Section A	Unit – 1	2	1
	Unit – 2	1	2
	Unit – 3	1	1
	Unit – 4	1	1
	Unit – 5	1	1
Section B	Unit – 1	1	1
	Unit – 2	1	1
	Unit – 3		1
	Unit – 4		1
	Unit – 5		1
Section C	Unit – 1		1
	Unit – 2		1
	Unit – 3	1	
	Unit – 4	1	1
	Unit -5		1

NME – PRACTICAL – MULTIMEDIA LAB

SUBJECT CODE :	**PRACTICAL**	**MARKS : 100**
SEMESTER : I	**CREDITS : 2**	**NO. OF HOURS : 30**

COURSE FRAMEWORK:

- Autonomy and initiative in knowledge acquisition and integration of multimedia technologies.
- Concept, development and implementation of new multimedia systems andapplications based in emergent technologies.

COURSE OUTCOME:

On completion of the course the students will be able

1. To demonstrate selection tools
2. To apply separation of background object and combining images
3. To implementation of Transform tools
4. To handle different filters
5. To implementation of design visiting card

GIMP

1. Implementation of different Selection Tool.
2. Applying different View Options.
3. Implementation of Transforming and sizing.
4. Images-adding, Deleting and Moving.
5. Layers-Implementation of Paint Tool.
6. Implementation of Transform Tool.
7. Implementation of different Filters.
8. Implementation of different Color Tools

CORE III OPERATING SYSTEMS AND LINUX

SUBJECT CODE :	THEORY	MARKS: 100
SEMESTER: II	CREDITS: 4	NO. OF HOURS : 90

COURSE FRAMEWORK:

At the end of this course students will have in-depth knowledge in core concepts of operating system andwork with UNIX.

COURSE OUTCOME:

On completion of the course the students will be able

1. To understand the basics of operating systems like kernel, shell, types, and views of operating systems
2. To describe the various CPU scheduling algorithms and remove deadlocks.
3. To explain various memory management techniques and concepts of thrashing
4. To recognize file system interface, protection, and security mechanisms.
5. To explain the various features of OS like Unix.

UNIT-I:

Introduction: What Operating Systems Do? – Operating System Structure – Operating system operations - Storage Management -Protection and Security - Computing Environments– System Structures: Operating System Services - Operating System Structure. **(18 Hours)**

UNIT-II:

Process Concept: Process Concept – Process Scheduling- Operations on Processes - Process Scheduling: Basic Concepts-Scheduling Criteria – Scheduling Algorithms-Deadlocks: System model – Deadlock Characterization-Methods of handling Deadlocks-Deadlock Prevention – Dead Lock Avoidance-Deadlock Detection - Recovery from Deadlock. **(18 Hours)**

UNIT-III:

Memory Management Strategies: Background – Swapping (Standard Swapping alone) - Contiguous Memory Allocation – Segmentation (Basic Method alone) – Paging (Basic Method alone). **(18 Hours)**

UNIT-IV:

Introduction to Unix – Unix Components – Commands in Unix – Command Substitution – File and File Organization: Unix files – categories of files- file system – directory commands File related commands. File Attributes and Permission – Standard I/O-Redirection, pipes and filters – Sample database file – Handling Columns and Fields**.** **(18 Hours)**

UNIT-V:

Shell Programming: Shell variables – positional parameters – branching control structures – loop control structures - Regular Expressions –The Grep family – The process- Parent and child process- types of process – foreground and background process- internal and external commands. **(18 Hours)**

PRESCRIBED BOOKS:

1. Abraham Silberschatz Peter B. Galvin, G. Gagne, “Operating System Concepts”, Ninth Edition, International Student Version. John Wiley & Sons (Asia) Pvt. Ltd, 2014.
2. M.G. Venkateshmurthy,”Introduction to Unix and shell programming “, Pearson Education India, New Delhi, 2009.

REFERENCE BOOKS:

1. Operating systems - Internals and Design Principles, W. Stallings, 6th Edition.

QUESTION PAPER PATTERN:

Section	Question Component	Numbers	Marks	Total
A	**Definition/Principle** Answer any 10 out of 12 questions (each in 50 words)	1-12	3	**30**
B	**Short Answer** Answer any 5 out of 7 questions (each in 300 words)	13-19	6	**30**
C	**Essay** Answer any 4 out of 6 questions (each in 600 words)	20-25	10	**40**
TOTAL MARKS				**100**

DISTRIBUTION OF QUESTIONS:

Section	Units	No. of Questions	
		Theory	**Problems**
A	Unit – 1	3	
	Unit – 2	3	
	Unit – 3	2	
	Unit – 4	2	
	Unit – 5	2	
B	Unit – 1	2	
	Unit – 2	2	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	
C	Unit – 1	2	
	Unit – 2	1	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	

CORE IV PRACTICAL - PROGRAMMING FOR ANALYTICS II

SUBJECT CODE :	PRACTICAL	MARKS: 100
SEMESTER: II	CREDITS: 4	NO. OF HOURS : 60

COURSE FRAMEWORK:

The first part of the course focuses on using the SAS macro facility to design, write, and debug macro programs, with an emphasis on understanding how programs that contain macro code areprocessed. In the second part, the course focuses on how to process SAS data using Structured Query Language (SQL).

COURSE OUTCOME:

On completion of the course the students will be able

1. To understand and perform text substitution in SAS code
2. To identify and use macro variables and macro functions
3. To write self-modifying and data-driven programs
4. To create data-driven macro variables using query
5. To implement DBMS data with SAS/Access technology

EXERCISES:

- Perform text substitution in SAS code.
- Use macro variables and macro functions.
- Automate and customize the production of SAS code.
- Conditionally or iteratively construct SAS code.
- Write self-modifying, data-driven programs.
- Query and subset data.
- Summarize and present data.
- Combine tables using joins and set operators.
- Create and modify tables and views.
- Create data-driven macro variables using a query.
- Access DBMS data with SAS/ACCESS technology.

COURSE MATERIALS REQUIRED TEXTS

1. SAS Macro Language Essentials, SAS Institute, Cary.
2. SAS SQL Essentials, SAS Institute, Cary.

REFERENCE TEXTS

1. SAS Macro Programming Made Easy, Third Edition, SAS Institute, Cary.
2. Carpenter's Complete Guide to the SAS Macro Language, Third Edition, SAS Institute, Cary.
3. PROC SQL: Beyond the Basics Using SAS®, Third Edition, SAS Institute, Cary

SAS INTERNAL REFERENCE

LAB2; C1 – MC1V2, C2 – SQ1M6; Exam Weightage – 50:50

ALLIED - II
STATISTICS - I

SUBJECT CODE :	**THEORY**	**MARKS: 100**
SEMESTER: II	**CREDITS: 5**	**NO. OF HOURS : 90**

COURSE FRAMEWORK:
To develop the student's ability to deal with numerical and quantitative issues in business

COURSE OUTCOME:
1. Compute univariate and bivariate frequency distribution with samples of size not proceeding 200
2. Create the Diagrammatic and graphical representation of data using Simple bar diagram, Multiple bar diagram, sub-divided bar diagram, Deviation bar diagram, Histogram and Pie diagram, Cumulative frequency curve and Lorenz curves.
3. Compute various measures of location, dispersion, moments, skewness and kurtosis
4. Compare two variables using correlation and regression lines.
5. Discuss the test of significance based on t, chi-square and F distributions with respect to mean and variance.

UNIT I:
Introduction to Statistics: Primary and Secondary data – Nominal, Ordinal, Ratio, and Interval scale (with examples) - Graphical Representation of data – Bar-charts, Pie-diagrams, Histograms, Frequency polygon, Ogives. **(18 Hours)**

UNIT II:
Measures of central tendency: properties – merits and demerits – weighted means– graphical location of median, quartiles, deciles, percentiles, and mode – relation between arithmetic mean, geometric mean and harmonic mean. **(18 Hours)**

UNIT III
Measures of dispersion: – characteristics – Coefficient of dispersion – Coefficient of variation – Moments – Relation between moments about mean in terms of moments about point – Pearson's coefficients. **(18 Hours)**

UNIT IV:
Skewness and Kurtosis – Pearson's coefficient of skewness – Bowley's coefficient of Skewness – coefficient of Skewness based upon moments. **Curve fitting** – Principle of least squares – Fitting of straight line, parabola, exponential and power curve. **(18 Hours)**

UNIT V:
Correlation and Regression: Simple correlation – Karl Pearson's coefficient of correlation –Rank correlation –Simple Regression – lines of regression – properties of regression coefficient –Multiple and Partial correlation coefficient in three variables. **Hypothesis Testing:** Estimation andHypothesis testing, t-test, chi-square test, ANOVA. **(18 Hours)**

PRESCRIBED BOOKS:
1. Agarwal.B.L (1996): Basic Statistics, 3/e, New Age International (P) Ltd.,
2. S.P.Gupta (2014), Statistical Methods, Sultan Chand & sons

REFERENCE BOOKS:

1. Gupta,S.C. and Kapoor, V.K.(2000): Fundamentals of Mathematical Statistics, 10/e, SultanChand and Sons
2. Sanjay Arora & Bansilal (2002): New Mathematical statistics, Meerat Publications, New Delhi
3. Hooda.R.P.(2003): Statistics for Business and Economics, 3/e, Mac Millan.

QUESTION PAPER PATTERN

Section	Question Component	Numbers	Marks	Total
Section A	**Definition / Principles** Answer any **10** out of 12 questions	1 – 12	3	**30**
Section B	**Short Answer** Answer any **5** out of 7 questions	13–19	6	**30**
Section C	**Essay** Answer any 4 out of 6 questions	20– 25	10	**40**
TOTAL				**100**

DISTRIBUTION OF QUESTIONS:

Sections	Units	No. of Questions	
		Theory	**Problems**
Section A	Unit – 1	2	1
	Unit – 2	1	2
	Unit – 3	1	1
	Unit – 4	1	1
	Unit – 5	1	1
Section B	Unit – 1	1	1
	Unit – 2	1	1
	Unit – 3		1
	Unit – 4		1
	Unit – 5		1
Section C	Unit – 1		1
	Unit – 2		1
	Unit – 3		1
	Unit – 4		2
	Unit -5		1

NME – PRACTICAL – E-COMMERCE LAB

SUBJECT CODE:	**PRACTICAL**	**MARKS: 100**
SEMESTER: II	**CREDITS: 2**	**NO. OF HOURS : 30**

COURSE FRAMEWORK:

- Learn how to design, develop and implement ecommerce web applications.
- Demonstrate how businesses sell products and services on the Web

COURSE OUTCOME:

On completion of the course the students will be able

1. To understand basic HTML tags
2. To create table tag
3. To familiarize with Hyperlink and Images
4. To implementation of Frames
5. To design Forms and Controls

1. Implements basic HTML tags
2. Creation of List
3. Working with Hyperlink, Images
4. Implementation of Table tag
5. Implementation of FRAMES
6. Working with forms and controls

CORE V BIG DATA TECHNOLOGIES

SUBJECT CODE:	THEORY	MARKS: 100
SEMESTER: III	CREDITS: 4	NO. OF HOURS : 90

COURSE FRAMEWORK:

- To understand Analytical Process Model.
- To learn Predictive Analytics and Descriptive Analytics.
- To learn Social Network Analytics.

COURSE OUTCOME:

On completion of the course the students will be able

1. To understand Big data basic nomenclature and analytics
2. To familiarize Predictive analytics and Evaluate Predictive Models
3. To understand Descriptive Analytics and Survival Analysis
4. To acquire knowledge of Social Network Analytics
5. To elucidate model design and documentation

UNIT – I : Basic nomenclature - Analytics process model - Analytics model requirements - Types of data sources – Sampling - types of data elements - Visual Data Exploration and Exploratory Statistical Analysis - Missing Values - Outlier Detection and Treatment - Standardizing Data – Categorization - weights of evidence coding - Variable selection - Segmentation. **(18 Hours)**

UNIT –II: Predictive Analytics: Target Definition - Linear Regression - Logistic Regression - Decision Trees - Neural Networks - Support Vector machines - Ensemble Methods - Multiclass Classification Techniques - Evaluating Predictive Models. **(18 Hours)**

UNIT – III: Descriptive Analytics: Association Rules - Sequence Rules - Segmentation. Survival Analysis: Survival Analysis Measurements - Parametric Survival Analysis-Proportional Hazards Regression – Extensions of Survival Analysis Models – Evaluating Survival Analysis Models. **(18 Hours)**

UNIT – IV : Social Network Analytics: Social Network Definitions - Social Network Metrics - Social Network Learning -Relational Neighbour Classifier - Probabilistic Relational Neighbour Classifier -Relational logistic Regression - Collective Inferencing. **(18 Hours)**

UNIT – V: Analytics: Backtesting Analytical Models - Benchmarking - Data Quality - Software – Privacy - Model Design and Documentation - Corporate Governance. Example applications:Credit Risk Modeling - Fraud Detection. **(18 Hours)**

PRESCRIBED BOOKS:

1.Bart Baesens, 2014, Analytics in a Big Data World: The Essential Guide to Data Science and Itsapplications, Wiley India Private Limited.

REFERENCES BOOKS:

1. Michael Minelli, Michele Chambers, 2013, Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses, Wiley CIO
2. Stephan Kudyba, 2014, Big Data, Mining and Analytics: Components of Strategic Decision Making, CRC Press.

QUESTION PAPER PATTERN:

Section	Question Component	Numbers	Marks	Total
A	**Definition/Principle** Answer any 10 out of 12 questions (each in 50 words)	1-12	3	**30**
B	**Short Answer** Answer any 5 out of 7 questions (each in 300 words)	13-19	6	**30**
C	**Essay** Answer any 4 out of 6 questions (each in 600 words)	20-25	10	**40**
TOTAL MARKS				**100**

DISTRIBUTION OF QUESTIONS:

Section	Units	No. of Questions	
		Theory	Problems
A	Unit – 1	3	
	Unit – 2	3	
	Unit – 3	2	
	Unit – 4	2	
	Unit – 5	2	
B	Unit – 1	2	
	Unit – 2	2	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	
C	Unit – 1	2	
	Unit – 2	1	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	

CORE VI DATABASE MANAGEMENT SYSTEMS

SUBJECT CODE:	THEORY	MARKS: 100
SEMESTER: III	CREDITS: 4	NO. OF HOURS : 90

COURSE FRAMEWORK:

On completion of this course student will be able to understand and the fundamental concepts of Database Management Systems and work with SQL

COURSE OUTCOME:

On completion of the course the students will be able

1. To understand database Concepts and Architecture, Query languages
2. To design simple database system and demonstrate basic tasks of modeling and designing
3. To execute SQL basic queries and complex SQL Retrieval queries
4. To understand query Transaction Processing
5. To familiarize recovery management and database security issues

UNIT I:

Databases and Database Users: Introduction – Characteristics of Database Approach – Actors of the Scene- Workers behind the scene – Advantages of using the DBMS Approach – Database System Concepts and Architecture : Three Schema Architecture and Data Independence – DBMS Languages. **(18 Hours)**

UNIT II:

Data Modelling Using Entity Relationship Model : Entities, Attributes, Entity types, Entity Sets, Keys and Value Sets, Relationship types, Relationship sets, roles and structural constraints, Weak Entity types – ER Diagrams, Naming Conventions, Design Issues - Notation for ER Diagrams – Alternative Notations for ER Diagrams – Enhanced Entity-Relationship Model : Specialization and Generalization (Basic concepts alone with example). **(18 Hours)**

UNIT III:

Basics of Functional Dependencies and Normalization for Relational Databases: Informal Design Guidelines for Relation Schemas – Functional Dependencies -Normal Forms based on General Definitions of Second and Third Normal form-Boyce Code Normal form. **(18 Hours)**

UNIT IV:

Basic SQL - SQL Data Definition and Data Types – Specifying Constraints in SQL- Basic Retrieval Queries in SQL - INSERT, DELETE, and UPDATE Statements in SQL - More SQL: MoreComplex SQL Retrieval Queries – Views (Virtual Tables) in SQL – Schema Change Statements in SQL. **(18 Hours)**

UNIT V:

Introduction to Transaction Processing Concepts and Theory: Desirable Properties of Transactions (ACID Properties). Database Recovery Techniques: Recovery Concepts- No-Undo / Redo Recovery based on Deferred Update – Recovery Techniques based on immediate update - Shadow Paging – Database Backup and Recovery from Catastrophic Failures - Database Security: Introduction to Database Security Issues - Discretionary Access Control Based on Granting and Revoking Privileges - Challenges to Maintaining Database Security. **(18 Hours)**

PRESCRIBED BOOKS:

1. Ramez Elmasri & Shamkant B. Navathe, Fundamentals of Database systems, 7th Edition , Pearson2016.

REFERENCE BOOKS:

1. H. F. Korth & A. Silverschatz, Database Concepts, Tata McGraw Hill, New Delhi
2. C. J. Date, Database Systems, Prentice Hall of India, New Delhi.
3. Ivan Bayross, SQL, PL/SQL, The programming language of Oracle.

QUESTION PAPER PATTERN:

Section	Question Component	Numbers	Marks	Total
A	**Definition/Principle** Answer any 10 out of 12 questions (each in 50 words)	1-12	3	**30**
B	**Short Answer** Answer any 5 out of 7 questions (each in 300 words)	13-19	6	**30**
C	**Essay** Answer any 4 out of 6 questions (each in 600 words)	20-25	10	**40**
TOTAL MARKS				**100**

DISTRIBUTIONS OF QUESTIONS:

Section	Units	No. of Questions	
		Theory	**Problems**
A	Unit – 1	3	
	Unit – 2	3	
	Unit – 3	2	
	Unit – 4	2	
	Unit – 5	2	
B	Unit – 1	2	
	Unit – 2	2	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	
C	Unit – 1	2	
	Unit – 2	1	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	

CORE VII R PROGRAMMING

SUBJECT CODE:	THEORY	MARKS: 100
SEMESTER: III	CREDITS: 4	NO. OF HOURS : 90

COURSE FRAMEWORK:

R is the powerful Tool for statistical analysis and mostly used for implementing various concepts of Data Science and Analytics. Student will be highly benefited bylearning such open source Tools.

COURSE OUTCOME:

On completion of the course the students will be able

1. To acquire basic knowledge of R data structures
2. To demonstrate Filtering and Matrix operations
3. To illustrate List and List functions
4. To understand Data Frames and functions
5. To illustrate Factors and Tables and understand R Programming Structures

UNIT I:

Interactive Mode - Batch Mode - Introduction to Functions: Variable Scope, Default Arguments - R Data Structures - Scalars, Vectors, Arrays, and Matrices - Adding and Deleting Vector Elements, Obtaining the Length of a Vector - Matrices and Arrays as Vectors, Common Vector Operations - Vector Arithmetic and Logical Operations - Vector Indexing - Using all() and any(). **(18 Hours)**

UNIT II:

Vectorized Operations: Vector In, Vector Out - NA and NULL Values – Filtering: Generating Filtering Indices - Filtering with the subset() Function- Matrices and Arrays : Creating Matrices - General Matrix Operations: Performing Linear Algebra Operations on Matrices- Matrix Indexing - Filtering on Matrices - Applying Functions to Matrix Rows and Columns: Using the apply() Function- Finding Outliers. **(18 Hours)**

UNIT III:

Adding and Deleting Matrix Rows and Columns: More on the Vector/Matrix Distinction - Naming Matrix Rows and Columns - Higher-Dimensional Arrays – List: Creating Lists. - General List Operations - List Indexing, Adding and Deleting List Elements, Getting the Size of a List, Accessing List Components and Values - Applying Functions to Lists: Using the lapply() andsapply() Functions, Recursive Lists. **(18 Hours)**

UNIT IV:

Data Frames - Creating Data Frames, Accessing Data Frames - Other Matrix-Like Operations Extracting Subdata Frames - More on Treatment of NA Values - Using the rbind() and cbind() Functions - Applying apply(), Merging Data Frames - Applying Functions to Data Frames: Using lapply() and sapply() on Data Frames. **(18 Hours)**

UNIT V:

Factors and Tables - Factors and Levels, Common Functions Used with Factors: The tapply() and Split() and By() Function- Working with Tables - Matrix/Array-Like Operations on Tables- aggregate() and cut() function - R Programming Structures - Control Statements, Arithmeticand Boolean Operators and Values, Default Values for Arguments - Return Values- Functions Are Objects- Accessing the Keyboard and Monitor - Reading and Writing Files. **(18 Hours)**

PRESCRIBED BOOKS:

1. THE ART OF R PROGRAMMING - 2011 A Tour of Statistical, Software Design by NormanMatlof - No Starch Press
2. Statistical Analysis with R for Dummies, Joseph Schmuller, John Wiley Publication - 2017.

QUESTION PAPER PATTERN:

Section	Question Component	Numbers	Marks	Total
A	**Definition/Principle** Answer any 10 out of 12 questions (each in 50 words)	1-12	3	**30**
B	**Short Answer** Answer any 5 out of 7 questions (each in 300 words)	13-19	6	**30**
C	**Essay** Answer any 4 out of 6 questions (each in 600 words)	20-25	10	**40**
TOTAL MARKS				**100**

DISTRIBUTION OF QUESTIONS:

Section	Units	No. of Questions	
		Theory	Problems
A	Unit – 1	3	
	Unit – 2	3	
	Unit – 3	2	
	Unit – 4	2	
	Unit – 5	2	
B	Unit – 1	2	
	Unit – 2	2	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	
C	Unit – 1	2	
	Unit – 2	1	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	

CORE VIII PRACTICAL - BIG DATA MANAGEMENT

SUBJECT CODE :	PRACTICAL	MARKS: 100
SEMESTER: III	CREDITS: 4	NO. OF HOURS : 60

COURSE FRAMEWORK:

This course trains the students to manage big data. This course teaches how to use SAS programming methods to read, write, and manipulate Hadoop data. Base SAS methods are used for reading and writing raw data with the DATA step and managing the Hadoop file system and executing Pig code from SAS via the HADOOP procedure. The course discusses the SAS/ACCESS Interface to Hadoop methods that allow LIBNAME access and SQL pass-through techniques to read and write Hadoop Hive table structures. In the final part of the course, we cover what is data integration, and in this context how do we register for sources and targets; create and work with jobs; and work with transformations.

COURSE OUTCOME:

On completion of the course the students will be able

1. To understand read and write Hadoop files
2. To create and use SQL procedure
3. To demonstrate use base SAS procedures with Hadoop
4. To familiarize create jobs and the functions of job editor in DI studio
5. To implement various transformations in DI Studio

EXERCISES:

- Read and write Hadoop files with the FILENAME statement.
- Execute and use Hadoop commands with the HADOOP procedure.
- Invoke the execution of Pig programs in Hadoop within a SAS program.
- Access Hadoop distributions using the LIBNAME statement and the SQL pass - through facility.
- Create and use SQL procedure pass-through queries.
- Use Base SAS procedures with Hadoop.
- Modify DS2 programs to execute in-database in Hadoop.
- Use data in Hadoop as disk storage for SAS Viya in-memory tables.
- register source data and target tables in DI Studio
- create jobs and explore the functionality of the job editor in DI Studio
- work with many of the various transformations in DI Studio

COURSE MATERIALS

REQUIRED TEXT

1. Introduction to SAS and Hadoop, SAS Institute, Cary.
2. Data Integration Studio: Essentials, SAS Institute, Cary.

REFERENCE TEXT

1. SAS and Hadoop Technology Overview, SAS Institute, Cary.

SAS INTERNAL REFERENCE

LAB3; C1 – DIAHSM, C2 –DI148; Exam Weightage – 50:50

ALLIED III OPERATIONS RESEARCH

SUBJECT CODE:	THEORY	MARKS: 100
SEMESTER: III	CREDITS: 5	NO. OF HOURS : 90

COURSE FRAMEWORK:

The objective of this course is to enable the student to understand and analyse managerial problems to equip him to use the resources such as capitals, materials, productions, controlling, directing, staffing, and machines more effectively.

COURSE OUTCOME:

1. Identify and develop operation research models from the verbal description of the real system.
2. Formulate the Linear Programming Problem and Evaluate the LPP using Graphical Method, Simplex method and Big-M method
3. Develop a report that describes the model and solving transportation, assignment problems using different techniques.
4. Use techniques to plan, schedule and control project techniques.
5. Use basic ideas of Queuing models-elements of queuing system-characteristics of queuing model, Elementary ideas of different models

UNIT I: (18 Hours)

Basics of Operation Research: Definition of O.R. – Characteristics of O.R. - Necessity of O.R. in Industry-Scope of O.R. – Operation Research and Decision Making. **Linear Programming:** Introduction – Requirements for a Linear Programming Problem – Assumptions in Linear ProgrammingModels -Applications of Linear Programming Method-Areas of Application of Linear Programming- Formulation of Linear Programming Problems.

UNIT II: (18 Hours)

Linear Programming: Graphical Method of Solutions – Some Exceptional Cases. **Simplex Method:** The Simplex Method (Technique or Algorithm) **- Artificial Variables Techniques**: Big M- Method - Two Phase method.

UNIT III: (18 Hours)

The Transportation Model: Introduction to the Model – Assumptions in the Transportation Model - Definitions of the Transportation Model - Matrix Terminology - **Formulation and Solution of Transportation Models:** Find Basic feasible solution by North West Corner Method, Least Cost Method, Vogel's Approximation Method.

UNIT IV: (18 Hours)

Variants in Transportation Problems: Unbalanced Transportation Problem, Maximization Problem.- **The Assignment Model :** Definition of the Assignment Model – Mathematical Representation of the Assignment Model – Solution of the Assignment Models – Formulation and Solution of the Assignment models - **Variations of the Assignment Problem**:Unbalanced Assignment Problem, Maximization Problem.

UNIT V: (18 Hours)

Sequencing Models and Related Problems: Sequencing Problems- Assumptions in Sequencing Problems- Processing n jobs through One Machine, Processing n jobs through Two Machine, Processing n jobs through Three Machine, Processing Two jobs through m Machine, Processing n jobs through m Machine. Problem related to Sequencing (Routing Problems in Networks)- Minimal Path Problems (Shortest Acyclic route Models).

(Exercise on the topics mentioned must be strictly as in Prescribed Text Book)

PRESCRIBED BOOKS:

1. Prem Kumar Gupta & D. S. Hira, Operations Research, 7th Revised Edition, S. Chand & Company Ltd.,2014.

REFERENCE BOOKS:

1. Kandiswarup, P. K. Gupta, Man Mohan, Operations Research, S. Chand & Sons Education Publications, New Delhi, 12th Revised edition.
2. S. Dharani Venkata Krishnan, Operations Research Principles and Problems, Keerthi publishing house PVT Ltd.

QUESTION PAPER PATTERN:

Section	Question Component	Numbers	Marks	Total
Section A	**Definition / Principles** Answer any **10** out of 12questions	1 – 12	3	**30**
Section B	**Short Answer** Answer any **5** out of 7questions	13–19	6	**30**
Section C	**Essay** Answer any 4 out of 6questions	20– 25	10	**40**
TOTAL				**100**

DISTRIBUTION OF QUESTIONS:

Sections	Units	No. of Questions	
		Theory	**Problems**
Section A	Unit – 1	2	1
	Unit – 2	2	1
	Unit – 3	1	1
	Unit – 4	1	1
	Unit – 5	1	1
Section B	Unit – 1	2	
	Unit – 2	1	1
	Unit – 3		1
	Unit – 4		1
	Unit – 5		1
Section C	Unit – 1	1	
	Unit – 2		1
	Unit – 3		1
	Unit – 4		2
	Unit -5		1

DEPARTMENT OF ENGLISH

UG Part IV SOFT SKILLS
SECOND YEAR

SUBJECT CODE :	**THEORY**	**MARKS: 100**
SEMESTER: III	**CREDITS: 3**	**NO. OF HOURS : 30**

THIRD SEMESTER: Job-oriented Skills

COURSE FRAMEWORK:

- To prepare the students to be job-ready.
- To help learners use English Language appropriately to the role or situation.
- To develop confidence in them to face Interviews.
- To train them to prepare their own CV/Resume

Different kinds of Interviews

Letter of Application and CV

Technical Writing - Circulars,

Memos, Agenda and Minutes

Group Discussion

Review - 1. Books 2. Films

BOOKS FOR REFERENCE:

1. Harishankar, Bharathi. ed. Essentials of Spoken and Presentation Skills. University of Madras.
2. John, Seely. 1998. The Oxford Guide to writing and speaking. Oxford U P, 1998, Delhi.
3. The Princeton Language Institute and Lanny Laskowski.2001. 10 days to more confident PublicSpeaking. Warner Books.
4. http://jobsearch.about.com/cs/curriculumvitae.html.//
5. http://www.cvtips.com//

QUESTION PAPER PATTERN

UG - SOFT SKILLS

TIME – 3 HRS **MAXIMUM MARKS – 50**

PART – A (5X2=10)

Answer any FIVE from the questions given below from Q.No.1 to Q.No.7 (5 out of 7)

PART – B (4X5=20)

Answer any FOUR from the questions given below from Q.No.8 to Q.No.13 (4 out of 6)

PART – C (2X10=20)

Answer TWO questions only choosing one each from Q.No.14 &Q.No.15 (Internal Choice)

CORE IX DATA MINING AND DATA WAREHOUSING

SUBJECT CODE :	THEORY	MARKS: 100
SEMESTER: IV	CREDITS: 4	NO. OF HOURS : 75

COURSE FRAMEWORK:

- To Understand Data pre-processing and Data Generalization
- To study Data warehouse and OLAP Technology
- To Understand Mining Object, Multimedia, Text and Web Data

COURSE OUTCOME:

On completion of the course the students will be able

1. To understand basic data mining concepts
2. To acquire knowledge on statistical perspectives data mining algorithms
3. To familiarize clustering and association rules
4. To understand and evaluate decision support systems
5. To illustrate data warehouse and design technology

UNIT-I : (15 Hours)

Introduction: Basics of Data Mining – Data Mining Versus Knowledge Discovery in Database – Data Mining Issues – Data Mining Metrics – Social Implications of Data Mining – Data Mining from a Database Perspective. Related Concepts: Database/OLTP Systems – Fuzzy Sets and Fuzzy Logic – Information Retrieval – Decision Support Systems – Dimensional Modelling – OLAP – Web Search Engines.

UNIT-II: (15 Hours)

Data Mining Techniques: Introduction – A Statistical Perspective on Data Mining – Similarity Measures – Decision Trees – Neural Networks – Genetic Algorithms – Classification: Introduction – Statistical Based Algorithms – Distance Based Algorithms – Decision Tree Based Algorithms – Neural Network Based Algorithms – Rule Based Algorithms.

UNIT-III: (15 Hours)

Clustering: Introduction – Similarity and Distance Measures – Hierarchical Algorithms – Partitional Algorithms – Clustering Large Database. Association Rules: Introduction – Large Item sets – Basic Algorithms – Parallel and Distributed Algorithms –Comparing Approaches – Incremental Rules – Advanced Association Rule Techniques – Measuring the Quality of Rule Techniques – Measuring the Quality of Rules.

UNIT-IV: (15 Hours)

Evolution of Decision Support Systems: The Evolution-Problems with naturally Evolving Architecture- The development life cycle- Patterns of Hardware Utilization-setting the stage for Re- engineering- The Data warehouse Environment: the structure of the Data warehouse –Granularity- Exploration and Data Mining – Living sample database – structuring data in the Data warehouse-Auditingand the Data warehouse-Incorrect Data in the Data warehouse.

UNIT-V: (15 Hours)

The Data Warehouse and Design: Beginning with Operational Data – Data/Process Models and the Architected Environment- Data warehouse and Data Models - Meta data - Cyclicity of Data- Complexity of Transformation and Integration - Data warehouse and Technology: Managing LargeAmounts of Data - Interfaces to Many Technologies - Compaction of Data - Other Technological features.

PRESCRIBED BOOKS:

1. Data Mining Introductory and Advanced Topics, Margaret H.Dunham, Pearson Education [LPE] First Impression, 2006.
2. Building the Data Warehouse, W.H. Inmon, Wiley Publishing, Inc. Fourth Edition, 2005.

REFERENCE BOOKS:

1. Data Warehousing in the Real World, Sam Anahory, Dennis Murray, Pearson Education [LPE], Thirteenth Indian Reprint, 2005.
2. Data Mining Concepts and Techniques, Jiawei Han, Micheline Kamber, Jian Pei, Morgan-Kaufmann series, Third Edition, 2012.

QUESTION PAPER PATTERN:

Section	Question Component	Numbers	Marks	Total
A	**Definition/Principle** Answer any 10 out of 12 questions (each in 50 words)	1-12	3	**30**
B	**Short Answer** Answer any 5 out of 7 questions (each in 300 words)	13-19	6	**30**
C	**Essay** Answer any 4 out of 6 questions (each in 600 words)	20-25	10	**40**
TOTAL MARKS				**100**

DISTRIBUTION OF QUESTIONS:

Section	Units	No. of Questions	
		Theory	Problems
A	Unit – 1	3	
	Unit – 2	3	
	Unit – 3	2	
	Unit – 4	2	
	Unit – 5	2	
B	Unit – 1	2	
	Unit – 2	2	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	
C	Unit – 1	2	
	Unit – 2	1	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	

CORE X PYTHON PROGRAMMING

SUBJECT CODE:	THEORY	MARKS: 100
SEMESTER: IV	CREDITS: 3	NO. OF HOURS : 75

COURSE FRAMEWORK:

- To understand the basic concepts of Python programming and Data Analysis.
- At the end of this course, student will be able to analyze data using Python.

COURSE OUTCOME:

On completion of the course the students will be able

1. To understand python programming basics and data types
2. To demonstrate code control statements, function and modules
3. To evaluate the concept of testing and debugging
4. To determine how to use exception handling in Python applications for error handling
5. To demonstrate object oriented programming through real time entities and implement GUI programs

UNIT I:

An Introduction to Python programming: Introduction to Python – How to use IDLE to develop programs – How to write your first programs – Basic coding skills – How to work with data types and variables – How to work with numeric data – How to work with string data – How to use five of the Python functions. **(15 Hours)**

UNIT II:

How to code control statements: How to code Boolean expressions - How to code the selection structure – How to use the iteration structure – How to define and use functions and modules – More skills for defining and using functions – How to create and use Modules – How to use standard modules – How to plan the functions of a program. **(15 Hours)**

UNIT III:

How to test and debug a program: An Introduction to testing and debugging – Four techniques for testing and debugging – How to use IDLE debugger – How to work with lists and tuples – Basic skills for working with lists – How to work with a list of lists – More skills for working with lists – Howto work with tuples. **(15 Hours)**

UNIT IV:

How to work with file I/O: Introduction – How to use text files – How to use CSV files – Howto handle Exceptions – Introduction – How to handle multiple exceptions – Other concepts and skills – How to work with numbers and with strings. **(15 Hours)**

UNIT V:

Object-Oriented Programming: How to define and use your own classes – An Introduction to Classes and Objects – How to define a class – How work with object composition – How to work with Encapsulation – How to work with Inheritance – How to work with a database – How to build a GUI program. **(15 Hours)**

PRESCRIBED BOOKS:

1.Murach's Python Programming, Michael Urban, Joel Murach, Mike Murach & Associates, Inc.2016.

REFERENCE BOOKS:

1. Python All-in-one dummies, John Shovic, Alan Simpson, John Wiley & Sons, Inc. 2019.

QUESTION PAPER PATTERN:

Section	Question Component	Numbers	Marks	Total
A	**Definition/Principle** Answer any 10 out of 12 questions (each in 50 words)	1-12	3	**30**
B	**Short Answer** Answer any 5 out of 7 questions (each in 300 words)	13-19	6	**30**
C	**Essay** Answer any 4 out of 6 questions (each in 600 words)	20-25	10	**40**
TOTAL MARKS				**100**

DISTRIBUTION OF QUESTIONS:

Section	Units	No. of Questions	
		Theory	**Problems**
A	Unit – 1	3	
	Unit – 2	3	
	Unit – 3	2	
	Unit – 4	2	
	Unit – 5	2	
B	Unit – 1	2	
	Unit – 2	2	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	
C	Unit – 1	2	
	Unit – 2	1	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	

CORE XI DATA STRUCTURES AND ALGORITHM

SUBJECT CODE:	THEORY	MARKS: 100
SEMESTER: IV	CREDITS: 4	NO. OF HOURS : 90

COURSE FRAMEWORK:

To understand and implement basic concepts of data structures andalgorithms.

COURSE OUTCOME:

On completion of the course the students will be able

1. To understand basic concepts of data structures and algorithm analysis
2. To implementation of Array, Linked List and Queue
3. To illustrate binary tree and implementation
4. To demonstrate Sorting and Searching techniques
5. To implementation of Graph traversal

UNIT I:

Data Structures and Algorithms – A Philosophy of Data Structures – Abstract Data Types and Data Structures – Design Patterns – Problems, Algorithms and Programs – Mathematical Preliminaries – Algorithm Analysis – Asymptotic Analysis. **(18 Hours)**

UNIT II:

Fundamental Data Structures – Lists, Array-Based List Implementation –Linked Lists – Comparison of List Implementations – Element Implementations - Doubly Linked Lists – Stacks – Array-Based Stacks – Linked Stacks – Comparison of Array-Based and Linked Stacks – ImplementingRecursion – Queues – Array – Based Queues – Linked Queues – Comparison of Array-Based and Linked Queues - Dictionaries. **(18 Hours)**

UNIT III:

Binary Trees – Definitions and Properties – Binary Tree Traversals – Binary Tree Node Implementations – Binary Search Trees – Heaps and Priority Queues – Huffman Coding Trees – Non- Binary Trees – General Tree Definitions and Terminology – Implementations – K-ary Trees. **(18 Hours)**

UNIT IV:

Sorting and Searching **–** Internal Sorting – Sorting Terminology and Notation – Three $\Theta(n^2)$ Sorting Algorithms – Shell sort – Merge sort – Quick sort – Heap sort – Searching – Searching Unsorted and Sorted Arrays – Hashing – Indexing. **(18 Hours)**

.

UNIT V:

Graphs – Terminology and Representations – Graph Implementations – Graph Traversals – Shortest Paths Problems – Minimum-Cost Spanning Trees – Balanced Trees – The AVL Tree – SpatialData Structures – Patterns of Algorithms – Dynamic Programming – The Knapsack Problem – All- Pairs shortest Paths. **(18 Hours)**

PRESCRIBED BOOKS:

1. Clifford A. Shaffer, “Data Structures and Algorithm Analysis”, Third Edition, Dover Publications, 2012.

REFERENCE BOOKS:

1. Ellis Horowiz, Sartaj Sahni and Sanguthevar Rajasekaran, Galgotia Publications, 2001, "Fundamentals of Computer Algorithms ".
2. A.V.Aho, J.D. Ullman, J.E. Hopcraft, 1983, Data Structures and Algorithms, Addison Wesley,Boston.

QUESTION PAPER PATTERN:

Section	Question Component	Numbers	Marks	Total
A	**Definition/Principle** Answer any 10 out of 12 questions (each in 50 words)	1-12	3	**30**
B	**Short Answer** Answer any 5 out of 7 questions (each in 300 words)	13-19	6	**30**
C	**Essay** Answer any 4 out of 6 questions (each in 600 words)	20-25	10	**40**
TOTAL MARKS				**100**

DISTRIBUTION OF QUESTIONS:

Section	Units	No. of Questions	
		Theory	**Problems**
A	Unit – 1	3	
	Unit – 2	3	
	Unit – 3	2	
	Unit – 4	2	
	Unit – 5	2	
B	Unit – 1	2	
	Unit – 2	2	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	
C	Unit – 1	2	
	Unit – 2	1	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	

CORE XII

PRACTICAL – APPLIED DATA MINING AND MACHINE LEARNING

SUBJECT CODE:	PRACTICAL	MARKS: 100
SEMESTER: IV	CREDITS: 4	NO. OF HOURS : 60

COURSE FRAMEWORK:

To train the students to apply data mining and machine learning. The first part of this course is designed for students who want to perform statistical analyses. The second part of the course covers the skills that are required to assemble analysis flow diagrams using a set of tools for both pattern discovery (segmentation, association, andsequence analyses) and predictive modeling (decision tree, regression, and neural network models).

COURSE OUTCOME:

On completion of the course the students will be able

1. To understand analysis of variance
2. To demonstrate potential outliers in multiple regression
3. To understand predictive models like decision trees and regression models
4. To familiarize complex models
5. To apply association and sequence discovery techniques to transaction data

EXERCISES:

- Perform analysis of variance.
- Perform linear regression and assess the assumptions.
- Use diagnostic statistics to identify potential outliers in multiple regression.
- Use chi-square statistics to detect associations among categorical variables.
- Fit a multiple logistic regression model.
- Use diagnostic statistics to identify potential outliers in multiple regression.
- Use chi-square statistics to detect associations among categorical variables.
- Fit a multiple logistic regression model.
- Build and understand predictive models such as decision trees and regression models.
- Compare and explain complex models.
- Apply association and sequence discovery to transaction data.

COURSE MATERIALSREQUIRED TEXT

1. SAS Enterprise Guide: ANOVA, Regression, and Logistic Regression, SAS Institute, Cary.
2. Applied Analytics Using SAS Enterprise Miner, SAS Institute, Cary.

REFERENCE TEXT

1. Rawlings, J. O. 1988.Applied Regression Analysis: A Research Tool. Pacific Grove, CA:Wadsworth & Brooks.
2. Hosmer, D.W. and Lemeshow, S. 2000.Applied Logistic Regression 2ndEdition. NewYork: John Wiley & Sons.
3. Beck, A. 1997. “Herb Edelstein discusses the usefulness of data mining” DS Star. Vol.1:No. 2

SAS Internal Reference

LAB4; C1 – EGBS82, C2 – AAEM51; Exam Weightage – 50:50

ALLIED - IV

STATISTICS - II

SUBJECT CODE :	THEORY	MARKS: 100
SEMESTER: IV	CREDITS: 5	NO. OF HOURS : 90

COURSE FRAMEWORK

On completion of this course student will be able to use the mathematical knowledge in optimal use of resources. To understand basics of Probability and Sampling theory.

COURSE OUTCOME

On completion of the course the students will be

1. Convert real-world problems into probability models. Discuss the concepts of Probability, conditional probability and Bayes' theorem and its applications.
2. Discuss Probability distribution functions.
3. Understand Mathematical expectation, Identify the different types of probability distributions ,use them to solve real life problems
4. To know about Probability Distributions of Functions of Random Variables
5. To study the concepts of the various distributions.

UNIT I

Basic Probability - Random Experiments - Sample Spaces Events - The Concept of Probability -The Axioms of Probability - Assignment of Probabilities -Conditional Probability – Baye's Theorem and its applications. (**18 hours)**

UNIT II

Random Variables - Random Variables - Discrete Probability Distributions - Distribution Functions for Random Variables - Distribution Functions for Discrete Random Variables-Continuous Random Variables **(18 hours)**

UNIT III

Mathematical Expectation - Definition of Mathematical Expectation - Functions of Random Variables - Theorems on Expectation - Variance and Standard Deviation - Theorems on Variance - Standardized Random Variables. **(18 hours)**

UNIT IV

Probability Distributions - Graphical Interpretations Joint Distributions Independent Random Variables - Change of Variables - Probability Distributions of Functions of Random Variables – Convolutions – Conditional Distributions Applications of Geometric Probability. **(18 hours)**

UNIT V

Special Probability Distributions - Binomial Distribution - Normal Distribution - Poisson Distribution – Geometric distribution and Exponential distribution.

(18 hours)

REFERENCE BOOKS:

1. Statistical Methods - S.P.Gupta , Sultan Chand & Sons 45th Edition(2017)
2. "Fundamentals of Mathematical Statistics" - S. C Gupta and V. K. Kapoor, 11th edition, S.Chand and Sons.

(Problems - 60% and Theory - 40%)

QUESTION PAPER PATTERN:

Section	Question Component	Numbers	Marks	Total
A	Answer any 10 out of 12 questions	1 – 12	3	**30**
B	Answer any 5 out of 7 questions	13–19	6	**30**
C	Answer any 4 out of 6 questions	20– 25	10	**40**
TOTAL				**100**

DISTRIBUTION OF QUESTIONS:

Section	Distribution of Questions:
A	At least two and not more than three questions from each unit.
B	At least one and not more than two questions from each unit.
C	At least one question from each unit.

SOFT SKILLS IV - QUANTITATIVE APTITUDE

SUBJECT CODE :	Theory	MARKS: 100
SEMESTER: IV	CREDITS: 3	NO. OF HOURS : 30

COURSE FRAMEWORK:
To impart quantitative aptitude skills in students.

COURSE OUTCOME:
On completion of the course the students will be able
1. To understand the concepts of HCF and LCM
2. To calculate Profit and Loss
3. To familiarize the concepts of Time, Work and Distance
4. To calculate Simple and Compound Interest
5. To find area, volume and surface area of different shapes

UNIT I:
Divisibility – HCF and LCM – Decimal Fractions – Square roots and Cube Roots – Logarithms – Antilogarithms.

UNIT II:
Averages – Percentage – Profit and Loss - Ratio and Proposition – Partnership – Alligation and mixture.

UNIT III:
Time and work – Pipes and Cistern – Time and Distance – Boats and Streams.

UNIT IV:
Simple Interest – Compound Interest – Stocks and Shares – True Discount – Banker's discount.

UNIT V:
Area – Volume and surface Areas – Heights and Distances – Data Interpretation-Tabulation – Bar Graphs – Pie Charts – Line Graphs.

BOOKS FOR REFERENCE:

1. R.S. Aggarwal, Objective Arithmetic, S. Chand & Company, New Delhi, 2005.
2. Govind Prasad Singh and Rakesh Kumar, Text Book of Quickest Mathematics (for allCompetitive Examinations), Kiran Prakashan, 2012
3. R.S. Aggarwal, Quantitative Aptitude, S. Chand & Company, New Delhi, 2012

QUESTION PAPER PATTERN

UG - SOFT SKILLS

TIME – 3 HRS **MAXIMUM MARKS –** 100

PART – A (20X5=100)

Answer any TWENTY from the questions given from Q.No.1 to Q.No.25 (20 out of 25)

EVS - ENVIRONMENTAL STUDIES

SUBJECT CODE:	THEORY	MARKS : 100
SEMESTER : IV	CREDITS : 2	NO. OF HOURS : 30

COURSE FRAMEWORK:

- This course introduces the concepts of Environmental Studies

COURSE OUTCOME:

On completion of the course the students will be able

1. To know the importance of environmental studies and methods of conservation of natural resources.
2. To describe the structure and function of an ecosystem.
3. To identify the values and conservation of bio-diversity.
4. To explain the causes, effects and control measures of various types of pollutions. 5.To select the appropriate methods for waste management.

UNIT I:

The Multidisciplinary nature of environmental studies Definition; Scope and importance, Need for public awareness. **(6hrs)**

UNIT II:

Natural Resources: Renewable and non-renewable resources: Natural resourcesand associated problems.

a) Forest resources: Use and Over-exploitation, deforestation, case studies. Timber extraction, mining, dams and their effects on forests and tribal people.
b) Water resources: Use and over-utilization of surface and ground water, floods, drought, conflicts over water, dams benefits and problems.
c) Mineral resources: Use and exploitation, environmental effects of extracting and using mineral resources, case studies.
d) Food resources: World food problems, changes caused by agriculture and overgrazing, effects of modern agriculture, fertilizer-pesticide problems, water logging, salinity, case studies.
e) Energy resources: Growing energy needs, renewable and non-renewable energy sources, use of alternate energy sources, Case studies.
f) Land resources: Land as a resource, land degradation, man induced landslides, soil erosion and desertification. - Role of an individual in conservation of natural resources. - Equitable use of resources for sustainable lifestyles. **(6hrs)**

UNIT III:

Ecosystems - Concept of an ecosystem. - Structure and function of an ecosystem.

- Producers, consumers and decomposers. - Energy flow in the ecosystem.

-Ecological succession. - Food chains, food webs and ecological pyramids. - Introduction, types, characteristic features, structure and function of the following ecosystem: -

a. Forest ecosystem
b. Grassland ecosystem
c. Desert ecosystem
d. Aquatic ecosystems (ponds, streams, lakes, rivers, oceans, estuaries) **(6hrs)**

UNIT IV:

Biodiversity and its Conservation

- Introduction-Definition: genetic, species and ecosystem diversity.
- Bio geographical classification of India.
- Value of biodiversity: consumptive use, productive use, social, ethical, aesthetic and option values.
- Biodiversity at global, National and local levels.
- India as a mega-diversity nation.
- Hot-spots of biodiversity.
- Threats to biodiversity: habitual loss, poaching of wildlife, man-wildlife conflicts.
- Endangered and endemic species of India.
- Conservation of biodiversity: In-situ and Ex-situ conservation of biodiversity.

UNIT V:

Environmental Pollution: Definition - Causes, effects and control measures of: -

a. Air pollution
b. Water pollution
c. Soil pollution
d. Marine pollution
e. Noise pollution
f. Thermal pollution
g. Nuclear hazards

- Solid waste Management: Causes, effects and control measures of urban and industrialwastes. - Role of an individual in prevention of pollution. - Pollution case studies.
- Disaster management: floods, earthquake, cyclone and landslides. (6hrs)

PRESCRIBED BOOKS:

Agarwal, K.C. 2001 Environmental Biology, Nidi Publ. Ltd. Bikaner.

REFERENCE BOOKS:

1. Cunningham, W.P.Cooper, T.H. Gorhani, E & Hepworth,M.T.2001,

2. Environmental Encyclopedia, Jaico Publ. House, Mumbai, 1196p.

QUESTION PAPER PATTERN:

Section	Question Component	Numbers	Marks	Total
A	**Definition/Principle** Answer any 5 out of 8 questions (each in 250 words)	1-8	8	**40**
B	**Short Answer** Answer any 3 out of 6 questions (each in 600 words)	9-14	20	**60**
TOTAL MARKS				**100**

DISTRIBUTION OF QUESTIONS:

Section	Units	No. of Questions	
		Theory	Problems
A	Unit – 1	2	
	Unit – 2	2	
	Unit – 3	2	
	Unit – 4	1	
	Unit – 5	1	
B	Unit – 1	1	
	Unit – 2	1	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	2	

CORE XIII MACHINE LEARNING BASICS

SUBJECT CODE :	THEORY	MARKS: 100
SEMESTER: V	CREDITS: 4	NO. OF HOURS : 90

COURSE FRAMEWORK:

- Understand the concept of Machine Learning
- Implement Machine language

COURSE OUTCOME:

On completion of the course the students will be able

1. To understand the Machine learning types and its main challenges
2. To Perform Data cleaning and Data classification
3. To describe Training models
4. To explain Supervised machine learning techniques
5. To learn Unsupervised learning techniques

UNIT-I : (18 Hours)

The Machine Learning Landscape : Introduction - Types of Machine Learning Systems - Supervised/Unsupervised Learning, Batch and Online Learning, Instance-Based Versus Model-BasedLearning - Main Challenges of Machine Learning - Non-representative Training Data, Poor-Quality Data - Underfitting the Training Data - Testing and Validating - Frame the Problem - Select a Performance Measure - Check the Assumptions - Create a Test Set - Visualizing Geographical Data- correlation - Prepare the Data for Machine Learning Algorithms.

UNIT-II: (18 Hours)

Data Cleaning - Handling Text and Categorical Attributes- Feature Scaling- Transformation Pipelines- Select and Train a Model- Training and Evaluating on the Training Set- Fine-Tune Your Model- Grid Search- Randomized Search. Classification: MNIST- Training a Binary Classifier- Performance Measures- Measuring Accuracy Using Cross-Validation- Confusion Matrix- Precision and Recall – Multiclass, multi-label and multi-output classification.

UNIT-III: (18 Hours)

Training Models: Linear Regression- The Normal Equation- Computational Complexity- Gradient and batch gradient Descent- Stochastic Gradient Descent- Mini-batch Gradient Descent- Polynomial Regression- Learning Curves- Ridge Regression- Lasso Regression- Logistic Regression-Estimating Probabilities- Training and Cost Function- Decision Boundaries.

UNIT-IV: (18 Hours)

Decision Trees: Training and Visualizing a Decision Tree- Making Predictions- Estimating Class Probabilities- CART Training Algorithm- Computational Complexity – EnsembleLearning and Random Forests – Voting classifiers – Bagging and Pasting - Random Patches andRandom Subspaces – Random Forests – Boosting – Adaboost – Gradient Boosting–Stacking.

UNIT-V: (18 Hours)

Dimensionality Reduction – Main Approaches for Dimensionality Reduction – PCA -Unsupervised Learning Techniques- Clustering - K-Means - Limits of K-Means - Using clusteringfor image segmentation, Preprocessing, Semi supervised Learning- DBSCAN.

PRESCRIBED BOOKS:

1. Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow by Aurélien Géron, 2019Aurélien Géron. Published by O'Reilly Media, Inc.

REFERENCE BOOKS:

1. Machine Learning Step-by-Step Guide to Implement Machine Learning Algorithms with Python ByRudolph Russell-Copyright 2018.
2. Introduction to Machine Learning Second Edition by Ethem Alpaydın, MIT Press - Cambridge,Massachusetts-London.

QUESTION PAPER PATTERN:

Section	Question Component	Numbers	Marks	Total
A	**Definition/Principle** Answer any 10 out of 12questions (each in 50 words)	1-12	3	**30**
B	**Short Answer** Answer any 5 out of 7 questions (each in 300 words)	13-19	6	**30**
C	**Essay** Answer any 4 out of 6 questions (each in 600 words)	20-25	10	**40**
TOTAL MARKS				**100**

DISTRIBUTION OF QUESTIONS:

Section	Units	No. of Questions	
		Theory	**Problems**
A	Unit – 1	3	
	Unit – 2	3	
	Unit – 3	2	
	Unit – 4	2	
	Unit – 5	2	
B	Unit – 1	2	
	Unit – 2	2	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	
C	Unit – 1	2	
	Unit – 2	1	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	

CORE XIV SOFTWARE ENGINEERING

SUBJECT CODE:	THEORY	MARKS: 100
SEMESTER: V	CREDITS: 4	NO. OF HOURS : 90

COURSE FRAMEWORK:

To understand the concepts of software engineering. To study the software testing and evolution.

COURSE OUTCOME:

On completion of the Course the students will be able

1. To study Professional software development
2. To understand the concepts of Requirements Engineering
3. To learn Architectural design and its implementation
4. To describe Software testing
5. To understand Software evolution processes

UNIT I:

Introduction: Professional Software Development – Software Engineering Ethics – **Software Processes:** Software Process Models-Process Activities-Coping with change. **(18 Hours)**

UNIT II:

Requirements Engineering: Introduction – Functional and Non-Functional requirements- Requirements Engineering Processes- Requirements elicitation – Requirement Specification – Requirements Validation - Requirement Change- **System Modeling :** Context Models – Interaction Models – Structural Models- Behavioural Models. **(18 Hours)**

UNIT III:

Architectural Design: Architectural Design Decisions – Architectural Views – Application Architectures - **Design and Implementation:** Object oriented design using the UML – Design patterns – Implementation Issues. **(18 Hours)**

UNIT IV:

Software Testing: Development testing – Test Driven Development – Release Testing – User Testing – **Quality Management:** Software Quality – Software Standards – Reviews and Inspections. **(18 Hours)**

UNIT V:

Software Evolution: Evolution Processes – Legacy Systems - Software Maintenance - **Agile Software Development:** Agile Methods. **(18 Hours)**

PRESCRIBED BOOKS:

1. Ian Sommerville, Software Engineering, 10th Edition, Pearson Education Limited, 2016.

REFERENCE BOOKS:

1. Roger S. Pressman, Software Engineering, 5th Edition, McGrawHill, 2001.

QUESTION PAPER PATTERN:

Section	Question Component	Numbers	Marks	Total
A	**Definition/Principle** Answer any 10 out of 12 questions (each in 50 words)	1-12	3	**30**
B	**Short Answer** Answer any 5 out of 7 questions (each in 300 words)	13-19	6	**30**
C	**Essay** Answer any 4 out of 6 questions (each in 600 words)	20-25	10	**40**
TOTAL MARKS				**100**

DISTRIBUTION OF QUESTIONS:

Section	Units	No. of Questions	
		Theory	**Problems**
A	Unit – 1	3	
	Unit – 2	3	
	Unit – 3	2	
	Unit – 4	2	
	Unit – 5	2	
B	Unit – 1	2	
	Unit – 2	2	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	
C	Unit – 1	2	
	Unit – 2	1	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	

CORE XV

PRACTICAL – APPLIED STATISTICS AND TIME SERIES MODELLING

SUBJECT CODE :	**PRACTICAL**	**MARKS: 100**
SEMESTER: V	**CREDITS: 4**	**NO. OF HOURS : 90**

COURSE FRAMEWORK:
This course trains the students to apply statistics and time series modelling. The first part of the course introduces students to build predictive models in an interactive, exploratory way using in-memory analytics. The second part of the course discusses the fundamentals of modeling time series data – this section focuses on the applied use of the three main model types used to analyze univariate time series: exponential smoothing, autoregressive integrated moving average with exogenous variables (ARIMAX), and unobserved components (UCM).

COURSE OUTCOME:
On completion of the course the students will be able
1. To understand statistical analysis of data
2. To Perform decision tree modelling
3. To demonstrate stratified model fitting
4. To familiarize time series data
5. To identify relative strength and weaknesses of the model types

EXERCISES:
- Perform Statistical analysis of data of any size.
- Create segments, or clusters, of input variables.
- Perform decision tree modelling.
- Perform stratified model fitting.
- Compare models.
- Generate score code.
- Create time series data.
- Accommodate trend, as well as seasonal and event-related variation, in time series models.
- Diagnose, fit, and interpret exponential smoothing models, ARIMAX models, and UCMs.
- Identify relative strengths and weaknesses of the three model types.

COURSE MATERIALS
REQUIRED TEXT
1. SAS Visual Statistics in SAS Viya: Interactive Model Building, SAS Institute, Cary.
2. Time Series Modelling Essentials, SAS Institute, Cary.

REFERENCE TEXT
1. Myers, R. H. 1990.Classical and Modern Regression with Applications, Second Edition.
   Boston: Duxbury Press.
2. Yaffee, Robert A., and Monnie McGee. 2000. Introduction to Time Series Analysis and Forecasting with Applications of SAS and SPSS. Academic Press: San Diego, CA.

SAS INTERNAL REFERENCE
LAB5; C1 – SVSO35, C2 – STSM51; Exam Weightage – 50:50.

CORE XVI CYBER SECURITY

SUBJECT CODE	THEORY	MARKS: 100
SEMESTER: V	CREDITS: 4	NO. OF HOURS : 60

COURSE FRAMEWORK:

- To Understand Cyber Security Fundamentals
- To study the Attacker Techniques and Motivations
- To Understand Defense and Analysis Techniques

COURSE OUTCOME:

On completion of the course the students will be able

1. To learn the basics of Cyber security concepts
2. To understand Hackers tracks and Detection techniques
3. To study Exploration and Web exploit tools
4. To demonstrate Malicious code and Spyware
5. To demonstrate about Memory forensics and Memory analysis frameworks

UNIT-I : (12 Hours)

Cyber Security Fundamentals: Network and Security Concepts – Information Assurance Fundaments – Basic Cryptography – Symmetric Encryption – Public key encryption – DNS –Firewalls – Virtualization – Radio-frequency Identification – Microsoft windows Security Principles – Windows Tokens – Window Messaging – Windows Program Execution.

UNIT-II: (12 Hours)

Attacker Techniques and Motivations: How Hackers Cover their Tracks – Types of Proxies - Tunneling Techniques – HTTP, ICMP – Detection and Prevention - Fraud Techniques – Phishing, Smishing, Vishing and Mobile Malicious Code - Threat Infrastructure.

UNIT-III: (12 Hours)

Exploitation – Techniques to Gain a Foothold – Shellcode – Integer Overflow Vulnerabilities – Stack Based Buffer Overflows – Format String Vulnerabilities – SQL Injection – Malicious PDF Files – Reducing the Risks of Malicious PDF Files – Race Conditions – Web Exploit Tools – DoS Conditions – Brute Force and Dictionary Attacks.

UNIT-IV: (12 Hours)

Malicious Code – Self-Replication Malicious Code - Evading Detection and Elevating Privileges – Persistent Software Techniques – Rootkits – Spyware- Attacks against Privileged User Accounts and Escalation of Privileges – Token Kidnapping – Virtual Machine Detection – Stealing Information and Exploitation – DLL-Injection – Browser Helper Objects.

UNIT-V: (12 Hours)

Defense and Analysis Techniques – Memory Forensics – Why Memory Forensics is Important – Capabilities of Memory Forensics – Memory Analysis Frameworks – Dumping Physical Memory – Installing and Using Volatility – Finding Hidden Processes – Volatility Analyst Pack – HoneyPots – Malicious Code Naming-Intrusion Detection Systems.

PRESCRIBED BOOKS:

1. Cyber Security Essentials, James Graham, Richard Howard, Ryan Olson, CRC, 2011

REFERENCE BOOKS:

1. Cyber Security Fundamentals, Chuck Easttom, Pearson, Second Edition.

QUESTION PAPER PATTERN:

Section	Question Component	Numbers	Marks	Total
A	**Definition/Principle** Answer any 10 out of 12 questions (each in 50 words)	1-12	3	**30**
B	**Short Answer** Answer any 5 out of 7 questions (each in 300 words)	13-19	6	**30**
C	**Essay** Answer any 4 out of 6 questions (each in 600 words)	20-25	10	**40**
TOTAL MARKS				**100**

DISTRIBUTION OF QUESTIONS:

Section	Units	No. of Questions	
		Theory	Problems
A	Unit – 1	3	
	Unit – 2	3	
	Unit – 3	2	
	Unit – 4	2	
	Unit – 5	2	
B	Unit – 1	2	
	Unit – 2	2	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	
C	Unit – 1	2	
	Unit – 2	1	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	

Elective - I
CLOUD COMPUTING

SUBJECT CODE:	**THEORY**	**MARKS:100**
SEMESTER: V	**CREDITS: 5**	**NO. OF HOURS : 90**

COURSE FRAMEWORK:

- Understand the concept of cloud computing
- Work in Amazon Cloud Platform

COURSE OUTCOME:

On completion of the course the students will be able

1. To learn Amazon Web Services (AWS) cloud computing
2. To create an AWS account
3. To understand the concept of EC2 Virtual machine
4. To deploy a simple Web application with AWS elastic beanstalk
5. To learn how to store data in the cloud

UNIT-I: (18 Hours)

Amazon Web Services-cloud computing- Hosting a web shop-Implementing a highly available system-Innovative and fast-growing platform- Services solve common problems-Enabling automation-Flexible capacity (scalability)- Built for failure (reliability)- Reducing time to market- Global infrastructure-Free Tier-Pay-per-use opportunities.

UNIT-II: (18 Hours)

Creating an AWS Account-Signing In-Creating a key pair-Create a billing alarm to keep track of your AWS bill- simple example: WordPress- Creating and exploring your infrastructure- Resource groups-Load balancer-MySQL database- Network file System.

UNIT-III: (18 Hours)

Using virtual machines: EC2 - Monitoring and debugging a virtual machine-Changing the size of a virtual machine- Starting a virtual machine in another data center. Programming your infrastructure - Infrastructure as Code- Installing and configuring the CLI- Controlling virtual machines with SDK.

UNIT-IV: (18 Hours)

Creating a virtual machine and run a deployment script on startup with AWS Cloud Formation- Deploying a simple web application with AWS Elastic Beanstalk-Components of AWS Elastic Beanstalk- Securing your system: IAM, security groups, and VPC- Securing your AWS account.

UNIT-V: (18 Hours)

STORING DATA IN THE CLOUD: Storing your objects: S3 and Glacier- Amazon S3- Backing up your data on S3 with AWS CLI-Using S3 for static web hosting-Elastic Block Store (EBS) - Sharing data volumes between machines - EFS-Creating a file system-Using Cloud Formation to describe a file system.

PRESCRIBED BOOKS:

1. Amazon Web Services in Action, Second Edition, MICHAEL WITTIG ANDREAS WITTIG FOREWORD BY BEN WHALEY MANNING©2019 by Manning Publications.

REFERENCE BOOKS:

1. Beginning Serverless Computing Developing with Amazon Web Services, Microsoft Azure, andGoogle Cloud Copyright © 2018 by Maddie Stigler A Press.
2. Amazon Web Services™ For Dummies® Published by: John Wiley & Sons, Inc., 2013, New Jersey.

QUESTION PAPER PATTERN:

Section	Question Component	Numbers	Marks	Total
A	**Definition/Principle** Answer any 10 out of 12 questions (each in 50 words)	1-12	3	**30**
B	**Short Answer** Answer any 5 out of 7 questions (each in 300 words)	13-19	6	**30**
C	**Essay** Answer any 4 out of 6 questions (each in 600 words)	20-25	10	**40**
TOTAL MARKS				**100**

DISTRIBUTION OF QUESTIONS:

Section	Units	No. of Questions	
		Theory	Problems
A	Unit – 1	3	
	Unit – 2	3	
	Unit – 3	2	
	Unit – 4	2	
	Unit – 5	2	
B	Unit – 1	2	
	Unit – 2	2	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	
C	Unit – 1	2	
	Unit – 2	1	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	

VALUE EDUCATION

SUBJECT CODE :	THEORY	MARKS: 100
SEMESTER: V	CREDITS: 2	NO. OF HOURS : 30

COURSE FRAMEWORK:

Values are socially accepted norms to evaluate objects, persons, and situations that form part and parcel of sociality. A value system is a set of consistent values and measures. Knowledge of the values are inculcated through education. It contributes in forming true human being, who are able to face life and make it meaningful. There are different kinds of values like, ethical or moral values, doctrinal or ideological values, social values and aesthetic values. Values can be defined as broad preferences concerning appropriate courses of action or outcomes. Assuch, values reflect a person's sense of right and wrong or what "ought" to be. There are representative values like, "Equal rights for all", "Excellence deserves admiration". "People should be treated with respect and dignity". Values tend to influence attitudes and behavior and help to solve common human problems. Values are related to the norms of a culture.

COURSE OUTCOME:

On completion of the course the students will be able

1. To inculcate the value system in their real life scenarios.
2. To implement the role of culture and civilization, roles and responsibilities in the society.
3. To effectively follow Salient values for life such as forgiveness, ability to sacrifice, self - esteem, teamwork and creative thinking.
4. To reflect the human rights, social values and welfare of the citizen.
5. To consider the relation between values and personal behavior affecting the achievementof a sustainable future.

UNIT I: (6HRS)

Value education-its purpose and significance in the present world – Value system – The role of culture and civilization-Holistic living – Balancing the outer and inner – Body, Mind and Intellectual level- Duties and responsibilities.

UNIT II : (6HRS)

Salient values for life- Truth, commitment, honesty and integrity, forgiveness and love, empathy and ability to sacrifice, care, UNITY , and inclusiveness, Self esteem and self confidence, punctuality – Time, task and resource management – Problem solving and decision making skills- Interpersonal and Intra personal relationship – Team work – Positive and creative thinking.

UNIT III: (6HRS)

Human Rights – Universal Declaration of Human Rights – Human Rights violations – National Integration – Peace and non-violence – Dr. A P J Kalam's ten points for enlightened citizenship – Social Values and Welfare of the citizen – The role of media in value building.

UNIT IV: (6HRS)

Environment and Ecological balance – interdependence of all beings – living and non-living. The binding of man and nature – Environment conservation and enrichment.

UNIT V: (6HRS)

Social Evils – Corruption, Cybercrime, Terrorism – Alcoholism, Drug addiction – Dowry – Domestic violence – untouchability – female infanticide – atrocities against women-How to tackle them.

BOOKS FOR REFERENCE:

1. M.G.Chitakra: Education and Human Values, A.P.H.Publishing Corporation, New Delhi, 2003
2. Chakravarthy, S.K.: Values and ethics for Organizations: Theory and Practice, Oxford UniversityPress, New Delhi, 1999.
3. Satchidananda, M.K.: Ethics, Education, Indian Unity and Culture, Ajantha Publications, Delhi,1991.
4. Das, M.S. & Gupta, V.K.: Social Values among Young adults: A changing Scenario, M.D.Publications, New Delhi, 1995.

QUESTION PAPER PATTERN:

Section	Question Component	Numbers	Marks	Total
A	**Essay** Answer any 5 out of10 questions **(each in 1200 words)**	1-10	20	**100**

DISTRIBUTION OF QUESTIONS:

Section	Units	No. of Questions	
		Theory	**Problems**
A	Unit – 1	2	
	Unit – 2	2	
	Unit – 3	2	
	Unit – 4	2	
	Unit – 5	2	

CORE XVII WEB DESIGN AND DEVELOPMENT

SUBJECT CODE :	THEORY	MARKS: 100
SEMESTER: VI	CREDITS: 4	NO. OF HOURS : 90

COURSE FRAMEWORK:

- To understand Web based programming and scripting languages.
- To learn the basic web concepts and to create rich internet applications that use most recent client-side programming technologies.
- To learn the basics of CSS, Java Script and PHP.

COURSE OUTCOME:

On completion of the course the students will be able

1. To acquire the basics of Web designing
2. To understand the concepts of cascading style sheets
3. To demonstrate colors and backgrounds
4. To explain Java Scripts and its anatomy
5. To describe the basics of PHP and its scripts

UNIT I:

Getting started in web design – Gearing up for web design – The Internet Versus the Web – Serving up your Information – The Anatomy of a Web Page – A Multitude of Devices – Sticking with the Standards – Progressive Enhancement – The Need for speed. **(18 Hours)**

UNIT II:

CSS for presentation – Introducing Cascading Style Sheets – The benefits of CSS – How Style Sheets Work – The Big Concepts – CSS units of measurement – Moving Forward with CSS - Formatting Text – Basic Font Properties – Advanced Typography with CSS3 – Changing Text Color –A Few More Selector Types – Text Line Adjustments. **(18 Hours)**

UNIT III:

Colors and Backgrounds – Specifying Color Values – Foreground Color – Background Color – Clipping the Background – Background Images – Thinking inside the Box – Specifying Box Dimensions – Padding – Borders - Margins **–** Floating and Positioning – Normal Flow – Floating – Fancy text wrap with CSS Shapes. **(18 Hours)**

UNIT IV:

JavaScript: What is JavaScript, Adding JavaScript to a page, The Anatomy of a Script, The Browser Object – Events – Putting it all together – Learning more about JavaScript – Using JavaScript – Meet the DOM – Ploy**f**ills – JavaScript Libraries – Big Finish**.** **(18 Hours)**

UNIT V:

PHP – The Basics – Examining the Structure of a PHP Script – Looking at PHP Syntax –Writing PHP Code – Displaying Content in a Webpage – Using PHP Variables – Using PHP Constants – Understanding Data Types – Using Arrays – Using Dates and Times – Understanding PHP Error Messages. **(18 Hours)**

PRESCRIBED BOOKS:

1. Learning Web Design, A Beginner's Guide to HTML, CSS, JavaScript and Web graphics, JenniferNiederst Robbins, Fifth Edition, O'Reilly Pub., 2018.(UNIT I-IV)
2. PHP, MySQL, JavaScript & HTML All-in-one for Dummies, Steven Suehring, Janet Valade, JohnWiley & Sons publication, 2013. (UNIT V)

REFERENCE BOOKS:

Laura Lemay, Rafe Colburn , Jennifer Kyrnin, "Mastering HTML, CSS & Javascript Web Publishing",2016.

QUESTION PAPER PATTERN:

Section	Question Component	Numbers	Marks	Total
A	**Definition/Principle** Answer any 10 out of 12 questions (each in 50 words)	1-12	3	**30**
B	**Short Answer** Answer any 5 out of 7 questions (each in 300 words)	13-19	6	**30**
C	**Essay** Answer any 4 out of 6 questions (each in 600 words)	20-25	10	**40**
TOTAL MARKS				**100**

DISTRIBUTION OF QUESTIONS:

Section	Units	No. of Questions	
		Theory	Problems
A	Unit – 1	3	
	Unit – 2	3	
	Unit – 3	2	
	Unit – 4	2	
	Unit – 5	2	
B	Unit – 1	2	
	Unit – 2	2	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	
C	Unit – 1	2	
	Unit – 2	1	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	

CORE XVIII NATURAL LANGUAGE PROCESSING

SUBJECT CODE :	THEORY	MARKS: 100
SEMESTER: VI	CREDITS: 4	NO. OF HOURS : 90

COURSE FRAMEWORK:

- To Understand Linguistic Resources for NLP
- To study the Linguistic studies of speech
- To Understand Syntactic Sphere

COURSE OUTCOME:

On completion of the course the students will be able

1. To study the concepts of Corpus and its anatomy
2. To understand Corpus annotation and Speech processing
3. To explain about Morphology sphere and Finite state machines
4. To learn Syntax sphere and Syntactic constituents
5. To understand the Syntactic formation and Recursive transition networks

UNIT-I:

The concept of a corpus – Corpus taxonomy – Written versus spoken – The historical point of view – The language of corpora - Who collects and distributes corpora? – The Gutenberg project – The Linguistic Data Consortium - The Lifecycle of a Corpus – Needs analysis – Design ofScenarios to collect data for the Corpus – Collection of the Corpus – Transcription. **(18 Hours)**

UNIT-II:

Corpus Annotation – Corpus Documentation – Statistical analysis of data – The Use of Corpora in NLP – Examples of Existing Corpora – The Sphere of Speech – Linguistic Studies of Speech – Phonetics – Phonology – Speech Processing – Automatic Speech Recognition – Speech Synthesis. **(18 Hours)**

UNIT-III:

Morphology Sphere – Elements of Morphology – Morphological typology – Morphology of English – Parts of speech – Terms, Collocations and Colligations – Automatic Morphological Analysis – Stemming – Regular Expressions for Morphological Analysis – Informal Introduction to finite-state machines – Two-level Morphology and FST – Part-of-Speech Tagging. **(18 Hours)**

UNIT-IV:

Syntax Sphere – Basic Syntactic Concepts – Delimitation of the field of Syntax – the Concept of Grammaticality – Syntactic Constituents – Syntactic typology of topology and agreement – Syntactic ambiguity – Syntactic specificities of spontaneous oral Language - Elements of Formal Syntax – Syntax Trees and Rewrite Rules – Languages and Formal Grammars – Hierarchy of Languages – Feature structures and Unification – Definite Clause Grammar. **(18 Hours)**

UNIT-V:

Syntactic Formalisms – X-bar – Head-driven Phrase Structure Grammar – Lexicalized Tree-adjoining Grammar – Automatic Parsing – Finite-State Automata – Recursive Transition Networks – Top-down approach – Bottom-up approach – Mixed approach – Tabular Parsing – Probabilistic Parsing – Neural Networks – Parsing algorithms for Unification-based grammars – Robust Parsing Approaches – Generation Algorithms.

(18 Hours)

PRESCRIBED BOOKS:

Natural Language Processing and Computational Linguistics 1, Speech, Morphology and Syntax, Mohamed Zakaria Kurdi, Wiley publications, 2016.

BOOKS FOR REFERENCE:

1. Natural Language Processing with python, Steven Bird, Ewan Lein, Edward Loper, 2009, OReilly Media.

QUESTION PAPER PATTERN:

Section	Question Component	Numbers	Marks	Total
A	**Definition/Principle** Answer any 10 out of 12 questions (each in 50 words)	1-12	3	**30**
B	**Short Answer** Answer any 5 out of 7 questions (each in 300 words)	13-19	6	**30**
C	**Essay** Answer any 4 out of 6 questions (each in 600 words)	20-25	10	**40**
TOTAL MARKS				**100**

DISTRIBUTION OF QUESTIONS:

Section	Units	No. of Questions	
		Theory	**Problems**
A	Unit – 1	3	
	Unit – 2	3	
	Unit – 3	2	
	Unit – 4	2	
	Unit – 5	2	
B	Unit – 1	2	
	Unit – 2	2	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	
C	Unit – 1	2	
	Unit – 2	1	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	

CORE XIX
PRACTICAL: OPTIMIZATION AND TEXT ANALYTICS

SUBJECT CODE:	**PRACTICAL**	**MARKS: 100**
SEMESTER: VI	**CREDITS: 4**	**NO. OF HOURS : 90**

COURSE FRAMEWORK:

The first part of the course focuses on formulating and solving mathematical optimization models in SAS. The course covers linear, integer, mixed integer, and nonlinear programming problems, with an emphasis on model formulation and construction. The second part of the course describes the functionality of SAS Text Miner to uncover underlying themes or concepts contained in large document collections, automatically group documents into topical clusters, classify documents into predefined categories, and integrate text data with structured data to enrich predictive modelling endeavors.

COURSE OUTCOME:

On completion of the course the students will be able

1. To understand and to formulate linear programming problems in SAS
2. To solve nonlinear programming problems using SAS
3. To identify topics or concepts that appear in a document collection
4. To implement user-influenced topic tables from scratch
5. To use derived topic tables

EXERCISES:

- Formulate and solve linear programming problems in SAS
- Solve integer and mixed integer programming problems using SAS
- Solve nonlinear programming problems using SAS
- Read documents from a variety of sources (web pages, flat files, data elements in a relationaldatabase, spreadsheet cells, and so on) into SAS tables.
- Process textual data for text mining.
- Convert unstructured text-based character data into structured numeric data.
- Identify topics or concepts that appear in a document collection.
- Create user-influenced topic tables from scratch or by modifying machine-generated topics orconcepts using domain knowledge.
- Use derived topic tables or pre-existing user-influenced topic tables (or both) to enhanceinformation retrieval and document classification.

COURSE MATERIALS

REQUIRED TEXT

1. Building and Solving Optimization Models with SAS/OR, SAS Institute, Cary.
2. Text Analytics Using SAS® Text Miner, SAS Institute, Cary.

REFERENCE TEXT

1. Chvátal, V. 1980.Linear Programming. New York: W.H. Freeman and Company.
2. Bracken, J. and G.P. McCormick. 1968.Selected Applications of Nonlinear Programming. New York: Wiley.
3. Albright, Russell. 2004.Taming Text with the SVD . SAS Institute White Paper.
4. Jurafsky, Daniel, and James H. Martin. 2000. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Upper SaddleRiver, New Jersey: Prentice Hall.

SAS INTERNAL REFERENCE

LAB6; C1 – OROP41, C2 – DMTX5; Exam Weightage – 50:50

CORE XX MINI PROJECT

SUBJECT CODE :	**PRACTICAL**	**MARKS: 100**
SEMESTER: VI	**CREDITS: 4**	**NO. OF HOURS : 90**

A mini-project should be done by the students based on concepts they have already learnt in the first two years of the programme. It may be based on database concepts, object oriented concepts, image processing, data Science, optimization tools, Big Data, etc.

COURSE FRAMEWORK:

Working on Mini project is to get used to the larger project, which will be handled in Industry. The project work constitutes an important component of the B.Sc. (DATA ANALYTICS) programme and it is to be carried out with due care and should be executed with seriousness by the students. The objective of this mini project is to help thestudent develop the ability to apply theoretical and practical tools/ techniques to solve real life problems related to industry, academic institutions and research laboratories.

COURSE OUTCOME:

On completion of the course the students will be able

1. To understand real life problem and finding solutions
2. To perform requirement analysis and identify design methodologies
3. To apply advanced programming techniques
4. To design technical report using various visualization tools and evaluation metrics
5. To implement final application

Guidelines: A student is expected to devote about 3 months in planning, analysing, designingand implementing the project. The initiation of project should be with the project proposal that isto be treated as an assignment:

Mini-project evaluation: The evaluation of the mini-project will be based on the project reports submitted by the student, a presentation and a demonstration.

ELECTIVE –II BUSINESS ANALYTICS

SUBJECT CODE:	THEORY	MARKS: 100
SEMESTER: VI	CREDITS: 5	NO. OF HOURS : 90

COURSE FRAMEWORK:

- Understand how business can be improved using business intelligence.
- Work in different business intelligence tools

COURSE OUTCOME:

On completion of the course the students will be able

1. To introduce Software models and to explore Data
2. To find the Relationships and Numerical variables
3. To explain about Probability distribution and elements of Decision analysis
4. To understand Multistage decision problems and Sampling terminology
5. To learn Regression analysis and Linear relationships

UNIT-I: (18 Hours)

Introduction - The Methods-The Software- Modeling and Models: Graphical Models, Algebraic Models, Spreadsheet Models, A Seven-Step Modeling Process. Exploring Data: Describing the Distribution of a Single Variable: Data Sets, Variables, and Observations, Types of Data, Descriptive Measures for Categorical and Numerical variables, Charts for Numerical Variables, Time Series Data, Outliers and Missing Values.

UNIT-II: (18 Hours)

Finding Relationships among Variables: Relationships among Categorical Variables and a Numerical Variable, Scatterplots, Correlation and Covariance, Pivot Tables, Probability and Decision making under Uncertainty: Probability Essentials, Rule of Complements-Addition Rule- Conditional Probability and the Multiplication Rule.

UNIT-III: (18 Hours)

Probability Distribution of a Single Random Variable- Conditional Mean and Variance-Normal, Binomial, Poisson, and Exponential Distributions - Decision Making under Uncertainty-Elements of Decision Analysis-Payoff Tables- Possible Decision Criteria- Expected Monetary Value(EMV)- Decision Trees- The Precision Tree Add-In-Bayes' Rule.

UNIT-IV: (18 Hours)

Multistage Decision Problems and the Value of Information- Sampling and Sampling Distributions- Sampling Terminology- Methods for Selecting Random Samples- Introduction to Estimation- Confidence Interval Estimation - Hypothesis Testing.

UNIT-V: (18 Hours)

Regression Analysis: Estimating Relationships- Linear versus Nonlinear Relationships- Outliers- Correlations: Indicators of Linear Relationships- Simple Linear Regression- Multiple Regression.

PRESCRIBED BOOKS:

1. Business Analytics: Data Analysis and Decision Making, Fifth Edition, S. Christian Albright andWayne L. Winston Cengage Learning, US.

REFERENCE BOOKS:

1. Business Analysis For Dummies®Published by: John Wiley & Sons, Inc.111 River Street Hoboken, NJ07030-5774, www.wiley.com.

QUESTION PAPER PATTERN:

Section	Question Component	Numbers	Marks	Total
A	**Definition/Principle** Answer any 10 out of 12 questions (each in 50 words)	1-12	3	**30**
B	**Short Answer** Answer any 5 out of 7 questions (each in 300 words)	13-19	6	**30**
C	**Essay** Answer any 4 out of 6 questions (each in 600 words)	20-25	10	**40**
TOTAL MARKS				**100**

DISTRIBUTION OF QUESTIONS:

Section	Units	No. of Questions	
		Theory	Problems
A	Unit – 1	3	
	Unit – 2	3	
	Unit – 3	2	
	Unit – 4	2	
	Unit – 5	2	
B	Unit – 1	2	
	Unit – 2	2	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	
C	Unit – 1	2	
	Unit – 2	1	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	

REFER ANNEXURE - I

ELECTIVE-I

Annexure I

Elective - I
CLOUD COMPUTING

SUBJECT CODE :	**THEORY**	**MARKS:100**
SEMESTER: V	**CREDITS: 5**	**NO. OF HOURS : 90**

COURSE FRAMEWORK:

- Understand the concept of cloud computing
- Work in Amazon Cloud Platform

COURSE OUTCOME:

On completion of the course the students will be able

1. To learn Amazon Web Services (AWS) cloud computing
2. To create an AWS account
3. To understand the concept of EC2 Virtual machine
4. To deploy a simple Web application with AWS elastic beanstalk
5. To learn how to store data in the cloud

UNIT-I: (18 Hours)

Amazon Web Services-cloud computing- Hosting a web shop-Implementing a highly available system-Innovative and fast-growing platform- Services solve common problems-Enabling automation-Flexible capacity (scalability)- Built for failure (reliability)- Reducing time to market- Global infrastructure-Free Tier-Pay-per-use opportunities.

UNIT-II: (18 Hours)

Creating an AWS Account-Signing In-Creating a key pair-Create a billing alarm to keep track of your AWS bill- simple example: WordPress- Creating and exploring your infrastructure- Resource groups-Load balancer-MySQL database- Network file System.

UNIT-III: (18 Hours)

Using virtual machines: EC2 - Monitoring and debugging a virtual machine-Changing the size of a virtual machine- Starting a virtual machine in another data center. Programming yourinfrastructure - Infrastructure as Code- Installing and configuring the CLI- Controlling virtual machines with SDK.

UNIT-IV: (18 Hours)

Creating a virtual machine and run a deployment script on startup with AWS Cloud Formation- Deploying a simple web application with AWS Elastic Beanstalk-Components of AWS Elastic Beanstalk- Securing your system: IAM, security groups, and VPC- Securing your AWS account.

UNIT-V: (18 Hours)

STORING DATA IN THE CLOUD: Storing your objects: S3 and Glacier-Amazon S3-Backing up your data on S3 with AWS CLI-Using S3 for static web hosting-Elastic Block Store (EBS) - Sharing data volumes between machines - EFS-Creating a file system-Using Cloud Formation todescribe a file system.

PRESCRIBED BOOKS:

1. Amazon Web Services in Action, Second Edition, MICHAEL WITTIG ANDREAS WITTIGFOREWORD BY BEN WHALEY MANNING©2019 by Manning Publications.

REFERENCE BOOKS:

1. Beginning Serverless Computing Developing with Amazon Web Services, Microsoft Azure, andGoogle Cloud Copyright © 2018 by Maddie Stigler A Press.
2. Amazon Web Services™ For Dummies® Published by: John Wiley & Sons, Inc., 2013, New Jersey.

QUESTION PAPER PATTERN:

Section	Question Component	Numbers	Marks	Total
A	**Definition/Principle** Answer any 10 out of 12 questions (each in 50 words)	1-12	3	**30**
B	**Short Answer** Answer any 5 out of 7 questions (each in 300 words)	13-19	6	**30**
C	**Essay** Answer any 4 out of 6 questions (each in 600 words)	20-25	10	**40**
TOTAL MARKS				**100**

DISTRIBUTION OF QUESTIONS:

Section	Units	No. of Questions	
		Theory	Problems
A	Unit – 1	3	
	Unit – 2	3	
	Unit – 3	2	
	Unit – 4	2	
	Unit – 5	2	
B	Unit – 1	2	
	Unit – 2	2	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	
C	Unit – 1	2	
	Unit – 2	1	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	

ELECTIVE I
DIGITAL LOGIC FUNDAMENTALS

SUBJECT CODE :	THEORY	MARKS: 100
SEMESTER: V	CREDITS: 5	NO. OF HOURS : 90

COURSE FRAMEWORK:
This course introduces Logic gates, Boolean algebra and combinationalcircuits

COURSE OUTCOME:
On completion of the course the students will be able
1. To understand number systems and number conversions.
2. To gain knowledge about the combinational and sequential logic circuits.
3. To procure knowledge Karnaugh Map
4. To enumerate combinational logic circuits
5. To understand sequential logic

UNIT I: (18Hours)
Binary number Systems - Code conversion: Binary to decimal conversion, Decimal to binary conversion, Hexadecimal number system, Hexadecimal to decimal conversion, Decimal to hexadecimal conversion, Hexadecimal to binary conversion, Binary to hexadecimal, Octal number system, Octal to decimal conversion, Decimal to Octal conversion, Octal to binary conversion, Binaryto octal conversion.

UNIT II: (18 Hours)
Logic Gates : Introduction, Analog and digital signal – basic logic gates, NOT, OR, AND – Logic circuits and logic expressions – Sum of Products (SOP) – Product of Sums (POS) – NAND andNOR gates – EX-OR and EX-NOR gates - Boolean Algebra: Laws of Boolean Algebra, De Morgan'sTheorems, NAND as Universal Gate, NOR as Universal Gate.

UNIT III: (18 Hours)
Simplification of Boolean functions: Karnaugh Map (k-map), Min terms and Max terms –Relation between k-map and truth table – 2-variable k-map using min terms - 3-variable k-map using min terms - 4-variable k-map using min terms – Don't care conditions.

UNIT IV: (18 Hours)
Combinational Logic: Half Adder and Full Adder - Half Subtractor and Full Subtractor –Multiplexer – De-multiplexer.

UNIT V: (18 Hours)
Sequential Logic: RS Clocked RS, D, JK, Master Slave JK, T Flip-Flops – Shift Registers –Types of Shift Registers.

PRESCRIBED BOOKS:

1. V. Vijayendran, 2004, Digital Fundamentals, S. Viswanathan (Printers & Publishers) Pvt. Ltd.
2. M. Morris Mano, 2014, 4th Edition, Digital Logic and Computer Design, Prentice-Hall of India Pvt. Ltd.

REFERENCE BOOKS

1. W. Stallings, Computer Organization and Architecture Designing for Performance, 8th Edition,Prentice Hall of India, 2009
2. A. J. Dos Reis, Assembly Language and Computer Architecture using C++ and JAVA, CourseTechnology, 2004

QUESTION PAPER PATTERN:

Section	Question Component	Numbers	Marks	Total
A	**Definition/Principle** Answer any 10 out of 12 questions (each in 50 words)	1-12	3	**30**
B	**Short Answer** Answer any 5 out of 7 questions (each in 300 words)	13-19	6	**30**
C	**Essay** Answer any 4 out of 6 questions (each in 600 words)	20-25	10	**40**
TOTAL MARKS				**100**

DISTRIBUTION OF QUESTIONS:

Section	Units	No. of Questions	
		Theory	**Problems**
A	Unit – 1	3	
	Unit – 2	3	
	Unit – 3	2	
	Unit – 4	2	
	Unit – 5	2	
B	Unit – 1	2	
	Unit – 2	2	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	
C	Unit – 1	2	
	Unit – 2	1	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	

ELECTIVE-II

ELECTIVE –II BUSINESS ANALYTICS

SUBJECT CODE :	THEORY	MARKS: 100
SEMESTER: VI	CREDITS: 5	NO. OF HOURS : 90

COURSE FRAMEWORK:

- Understand how business can be improved using business intelligence.
- Work in different business intelligence tools

COURSE OUTCOME:

On completion of the course the students will be able

1. To introduce Software models and to explore Data
2. To find the Relationships and Numerical variables
3. To explain about Probability distribution and elements of Decision analysis
4. To understand Multistage decision problems and Sampling terminology
5. To learn Regression analysis and Linear relationships

UNIT-I: (18 Hours)

Introduction - The Methods-The Software- Modeling and Models: Graphical Models, Algebraic Models, Spreadsheet Models, A Seven-Step Modeling Process. Exploring Data:Describing the Distribution of a Single Variable: Data Sets, Variables, and Observations, Types of Data, Descriptive Measures for Categorical and Numerical variables, Charts for Numerical Variables, Time Series Data, Outliers and Missing Values.

UNIT-II: (18 Hours)

Finding Relationships among Variables: Relationships among Categorical Variables anda Numerical Variable, Scatterplots, Correlation and Covariance, Pivot Tables, Probability andDecision making under Uncertainty: Probability Essentials, Rule of Complements-Addition Rule-Conditional Probability and the Multiplication Rule.

UNIT-III: (18 Hours)

Probability Distribution of a Single Random Variable- Conditional Mean and Variance-Normal, Binomial, Poisson, and Exponential Distributions - Decision Making under Uncertainty- Elements of Decision Analysis-Payoff Tables-Possible Decision Criteria- Expected MonetaryValue (EMV)- Decision Trees- The Precision Tree Add-In-Bayes' Rule.

UNIT-IV: (18 Hours)

Multistage Decision Problems and the Value of Information- Sampling and Sampling Distributions- Sampling Terminology- Methods for Selecting Random Samples- Introduction to Estimation- Confidence Interval Estimation - Hypothesis Testing.

UNIT-V: (18 Hours)

Regression Analysis: Estimating Relationships- Linear versus Nonlinear Relationships- Outliers- Correlations: Indicators of Linear Relationships- Simple Linear Regression- MultipleRegression.

PRESCRIBED BOOKS:

1. Business Analytics: Data Analysis and Decision Making, Fifth Edition, S. Christian Albright andWayne L. Winston Cengage Learning, US.

REFERENCE BOOKS:

1. Business Analysis For Dummies®Published by: John Wiley & Sons, Inc.111 River Street Hoboken,NJ 07030-5774, www.wiley.com.

QUESTION PAPER PATTERN:

Section	Question Component	Numbers	Marks	Total
A	**Definition/Principle** Answer any 10 out of 12 questions (each in 50 words)	1-12	3	**30**
B	**Short Answer** Answer any 5 out of 7 questions (each in 300 words)	13-19	6	**30**
C	**Essay** Answer any 4 out of 6 questions (each in 600 words)	20-25	10	**40**
TOTAL MARKS				**100**

DISTRIBUTION OF QUESTIONS:

Section	Units	No. of Questions	
		Theory	Problems
A	Unit – 1	3	
	Unit – 2	3	
	Unit – 3	2	
	Unit – 4	2	
	Unit – 5	2	
B	Unit – 1	2	
	Unit – 2	2	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	
C	Unit – 1	2	
	Unit – 2	1	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	

ELECTIVE II INTRODUCTION TO IoT

SUBJECT CODE :	THEORY	MARKS: 100
SEMESTER: VI	CREDITS: 5	NO. OF HOURS : 90

COURSE FRAMEWORK:

This course introduces the concept of Internet of Things

COURSE OUTCOME:

On completion of the course the students will be able

1. To understand about Internet of Things and its design principles
2. To familiarize prototyping and embedded devices
3. To enumerate prototyping the physical design
4. To demonstrate techniques for writing embedded code
5. To understand about manufacturing techniques

UNIT I: (18 Hours)

The Internet of Things: An Overview: The Flavour of the Internet of Things, The "Internet" of "Things", The Technology of the Internet of Things, Enchanted Objects, Who is Making the Internet of Things? **Design Principles for Connected Devices:** Calm and Ambient Technology, Magic as Metaphor, Privacy, Keeping Secrets, Whose Data Is It Anyway? WebThinking for Connected Devices, Small Pieces, Loosely Joined, First-Class Citizens On The Internet, Graceful Degradation, Affordances. **Internet Principles: I**nternet Communications: An Overview, IP, TCP, The IP Protocol Suite (TCP/IP), UDP, IP Addresses, DNS, Static IP Address Assignment, Dynamic IP Address Assignment, IPv6, MAC Addresses, TCP and UDP Ports, An Example: HTTP Ports, Other Common Ports, Application Layer Protocols, HTTP, **HTTPS: Encrypted HTTP, Other Application Layer Protocols.**

UNIT II: (18 Hours)

Thinking About Prototyping: Sketching, Familiarity, Costs versus Ease of Prototyping, Prototypes and Production, Changing Embedded Platform, Physical Prototypes and Mass Personalisation, Climbing into the Cloud, Open Source versus Closed Source, Why Closed? Why Open? Mixing Open and Closed Source, Closed Source for Mass Market Projects, Tapping into the Community. **Prototyping Embedded Devices:** Electronics, Sensors, Actuators, Scaling Up the Electronics, Embedded Computing Basics, Microcontrollers, System-on-Chips, Choosing Your Platform, Arduino, Developing on the Arduino, Some Notes on the Hardware, Openness, RaspberryPi, Cases and Extension Boards, Developing on the Raspberry Pi, Some Notes on the Hardware, Openness.

UNIT III: (18 Hours)

Prototyping the Physical Design: Preparation, Sketch, Iterate, and Explore, Nondigital Methods, Laser Cutting, Choosing a Laser Cutter, Software, Hinges and Joints, 3D Printing, Types of 3D Printing, Software, CNC Milling, Repurposing/Recycling.

Prototyping Online Components: Getting Started with an API, Mashing Up APIs, Scraping,Legalities, Writing a New API, Clockodillo, Security, Implementing the API, Using Curl to Test, Going Further, Real-Time Reactions, Polling, Comet, Other Protocols, MQ Telemetry Transport, Extensible Messaging and Presence Protocol, Constrained Application Protocol.

UNIT IV: **(18 Hours)**

Techniques for Writing Embedded Code: Memory Management, Types of Memory, Making the Most of Your RAM, Performance and Battery Life, Libraries, Debugging **Business Models:** A Short History of Business Models, Space and Time, From Craft to Mass Production, The Long Tail of the Internet, Learning from History, The Business Model Canvas, Who Is the Business Model For? Models, Make Thing, Sell Thing, Subscriptions, Customisation, Be a Key Resource, Provide Infrastructure: Sensor Networks, Take a Percentage, Funding an Internet of Things Startup, Hobby Projects and Open Source, Venture Capital, Government Funding, Crowdfunding, Lean Startups.

UNIT V: **(18 Hours)**

Moving to Manufacture: What Are You Producing? Designing Kits, Designing Printed circuit boards, Software Choices, The Design Process, Manufacturing Printed Circuit Boards, Etching Boards, Milling Boards. Assembly, Testing, Mass-Producing the Case and Other Fixtures, Certification, Costs, Scaling Up Software, Deployment, Correctness and Maintainability, Security, Performance, User Community. **Ethics:** Characterizing the Internet of Things, Privacy, Control, Disrupting Control, Crowdsourcing, Environment, Physical Thing, Electronics, Internet Service, Solutions, The Internet of Things as Part of the Solution, Cautious Optimism, The Open Internet of Things Definition.

PRESCRIBED BOOKS:

1. Adrian McEwen, Hakim Cassimally, Designing the Internet of Things, WILEY First Edition 2014
2. Raj Kamal, Internet of Things – Architecture and Design, McGraw Hill, 1st Edition, 2017

REFERENCE BOOKS:

1. Cuno Pfister, Getting Started with the Internet of Things, O'Reilly, 6th Edition, 2018

QUESTION PAPER PATTERN:

Section	Question Component	Numbers	Marks	Total
A	**Definition/Principle** Answer any 10 out of 12 questions (each in 50 words)	1-12	3	**30**
B	**Short Answer** Answer any 5 out of 7 questions (each in 300 words)	13-19	6	**30**
C	**Essay** Answer any 4 out of 6 questions (each in 600 words)	20-25	10	**40**
TOTAL MARKS				**100**

DISTRIBUTION OF QUESTIONS:

Section	Units	No. of Questions	
		Theory	Problems
A	Unit – 1	3	
	Unit – 2	3	
	Unit – 3	2	
	Unit – 4	2	
	Unit – 5	2	
B	Unit – 1	2	
	Unit – 2	2	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	
C	Unit – 1	2	
	Unit – 2	1	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	

ELECTIVE II
INTRODUCTION TO DEEP LEARNING

SUBJECT CODE :	THEORY	MARKS: 100
SEMESTER: VI	CREDITS: 5	NO. OF HOURS : 90

COURSE FRAMEWORK:
This course introduces the basic concepts of Deep Learning

COURSE OUTCOME:
On completion of the course the students will be able
1. To understand the concepts about neural network
2. To Implement neural networks
3. To familiarize about embedding and representation
4. To implement a part-of-speech tagger
5. To understand memory augmented neural networks

UNIT I: **(18 Hours)**
The Neural Network- Building Intelligent Machines – The Limits of Traditional ComputerPrograms – The Mechanics of Machine Learning – The Neuron – Expression Linear Perceptrons as Neurons – Feed-Forward Neural Networks – Linear Neurons and their Limitations – Training Feed-Forward Neural Networks – The Fast-Food Problem – The Test sets, Validation sets and Overfitting.

UNIT II: **(18 Hours)**
Implementing Neural Networks in TensorFlow – What is TensorFlow – How Does Tensorflow Compare to Alternatives? – Creating and manipulating Tensor Flow Variables –TensorFlow Operations – Beyond Gradient Descent – The Challenges with Gradient Descent –Model Identifiability – Flat Regions in the Error Surface.

UNIT III: **(18 Hours)**
Embedding and Representation Learning – Learning Lower-Dimensional Representations – Principal Component Analysis – Motivating the Autoencoder in Tensorflow – Sparsity in Autoencoders – Implementing the Skip-Gram Architecture – Models for Sequence Analysis – Analyzing Variable-Length Inputs.

UNIT IV: **(18 Hours)**
Implementing a Part-of-Speech Tagger – Dependency Parsing and SyntaxNet – Beam Search and Global Normalization – Recurrent Neural Networks – The Challenges with Vanishing Gradients – TensorFlow Primitives for RNN Models – Implementing a Sentiment Analysis Model –Solving Seq2seq Tasks with Recurrent Neural Networks.

UNIT V: **(18 Hours)**
Memory Augmented Neural Networks – Neural Turing Machines – Attention –Based Memory Access – NTM Memory Addressing Mechanisms – Differentiable Neural Computers – DNC memory reuse – Temporal Linking of DNC Writes – The DNC Controller Network – Visualizing the DNC in Action – Implementing the DNC in Tensorflow.

PRESCRIBED BOOKS:

1. Fundamentals of Deep Learning, Designing Next-Generation Machine Intelligence Algorithms, byNikhil Buduma and Nicholas Lacascio, 2017 Published by O'Reilly Media, Inc.,

REFERENCE BOOKS:

1. Bengio, Yoshua, Ian Goodfellow, Aaron Courville, Deep learning, MIT press, 2016.
2. Raúl Rojas, Neural Networks: A Systematic Introduction, 1996, 2nd edition

QUESTION PAPER PATTERN:

Section	Question Component	Numbers	Marks	Total
A	**Definition/Principle** Answer any 10 out of 12 questions (each in 50 words)	1-12	3	**30**
B	**Short Answer** Answer any 5 out of 7 questions (each in 300 words)	13-19	6	**30**
C	**Essay** Answer any 4 out of 6 questions (each in 600 words)	20-25	10	**40**
TOTAL MARKS				**100**

DISTRIBUTION OF QUESTIONS:

Section	Units	No. of Questions	
		Theory	Problems
A	Unit – 1	3	
	Unit – 2	3	
	Unit – 3	2	
	Unit – 4	2	
	Unit – 5	2	
B	Unit – 1	2	
	Unit – 2	2	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	
C	Unit – 1	2	
	Unit – 2	1	
	Unit – 3	1	
	Unit – 4	1	
	Unit – 5	1	