RASHTRASANT TUKADOJI MAHARAJ NAGPUR UNIVERSITY, NAGPUR
THREE YEAR BACHELOR OF SCIENCE (B. Sc.) DATA SCIENCE DEGREE COURSE
(C B S)

B. Sc. (Data Science)		**Semester - III**	
Course Name : Programming in JAVA (Paper – I)		Course Code: (3DST01)	
Periods per week (1 Period is 60 minutes)		4	
Credits		4	
		Hours	Marks
Evaluation Scheme	Theory Examination	3	80
	Internal Assessment	--	20
		Total Marks	100

Learning Objectives:
This course aims to provide an in-depth understanding of object oriented programming: abstract data types, encapsulation, inheritance and polymorphism Fundamental features of an object oriented language like Java: object classes and interfaces, exceptions and libraries of object collections

Learning Outcomes:
After completion of this course successfully the students will be able to:

1. To learn why Java is useful for the design of desktop and web applications.
2. To learn how to implement object-oriented designs with Java.
3. To identify Java language components and how they work together in applications.
4. To design and program stand-alone Java applications.
5. To learn how to design a graphical user interface (GUI) with Java Swing.
6. To learn how to extend Java classes with inheritance and dynamic binding.
7. To learn how to use exception handling in Java applications.
8. To understand how to design applications with threads in Java.
9. To learn how to read and write files in Java.

B.Sc. (Data Science)
Semester – III
Paper – I
Programming in Java (3DST01)

UNIT – I

Introduction - Java history, Java features, How Java differs from C and C++, Java and internet, Java and world wide web, Java environment. Simple Java programs, Java program structure, Java tokens, Java virtual machine, Command line arguments. **Variables, Data Types and Simple I/O** –Variables, Data Types, Scope of variables, Symbolic constants, Type casting, Standards default values, Getting Simple User Input. **Operators in java**-Introduction, Arithmetic operators, Relational operators, Logical operators, Assignment operators, Increment and decrement operators, Conditional operators, Bitwise operators, Special operators, Mathematical functions. **Decision Making and Branching** –If…Else statement, Nesting of If...Else statement, the switch statement, The ? : Operators.

UNIT-II

Decision Making and Looping – Introduction, The while statement, the do statement, for statement, Jumps in loops, Labeled loops. **Arrays, Strings** – One Dimensional Array, Two Dimensional Array, Strings. **Classes, Objects and Methods** -- Introduction, Defining a class, Methods declaration, Creating objects, Accessing class members, Constructors, Method overloading, Static members, **Inheritance**: Extending a class, Overriding methods, Final variables and methods, Final classes, Finalizer methods, Abstract methods and classes, Visibility Controls. **Interfaces:**– Introduction, Defining interfaces, Implementing interfaces, Accessing interface variables.

UNIT - III

Packages:– Introduction, Java API Packages, Using system packages, Naming conventions, Creating packages, Accessing a package, Using a package. **Introduction to Thread** - Creating threads, Life cycle of thread. **Managing Errors and Exceptions** – Introduction, Types of errors, Exceptions, Syntax of exceptions handling code, catch statements, using finally statements, throwing our own exceptions. **Graphics Programming** – Introduction, The graphics class, Lines and rectangles, Circles and ellipses, Drawing arcs, Drawing polygons.

UNIT-IV

Applet Programming – Introduction, How applet differ from application, Preparing to write applet, Building applet code, Applet life cycle, Creating an executable applet, Designing a web page, Applet tag, Adding applet to HTML file, Running the applet, Passing parameters to applet, Displaying numerical values, Getting input from the user.

Managing Input / Output Files in JAVA – Introduction, Concepts of streams, Streams classes, Bytes streams classes, Character streams classes, Using the file classes, Input / Output exception, Creation of files, Reading/Writing character, Reading/Writing bytes.

Text Book:

1. E. Balagurusamy, Programming with Java, McGraw-Hill.
2. Java Programming for the absolute beginner, PHI, Joseph P. Russel
3. Schildt, The Complete Reference Java 2, McGraw-Hill.
4. Rashmi Kanta Das, Core Java for Beginners, Vikas Publishing.

RASHTRASANT TUKADOJI MAHARAJ NAGPUR UNIVERSITY, NAGPUR
THREE YEAR BACHELOR OF SCIENCE (B. Sc.) DATA SCIENCE DEGREE COURSE
(C B S)

B. Sc. (Data Science)		**Semester - III**	
Course Name : Database Management Systems (Paper – II)		Course Code: (3DST02)	
Periods per week (1 Period is 60 minutes)		4	
Credits		4	
		Hours	Marks
Evaluation Scheme	Theory Examination	3	80
	Internal Assessment	--	20
		Total Marks	100

Learning Objectives:
The objective of the course is to present an introduction to database management systems, with an emphasis on how to organize, maintain and retrieve - efficiently, and effectively - information from a DBMS. It also provides students with theoretical knowledge and practical skills in the use of databases and database management systems in information technology applications.

Learning Outcomes:
After completion of this course successfully the students will be able to:

1. Understand the basic concepts and the applications of database systems.
2. Master the basics of SQL and construct queries using SQL.
3. Understand the relational database design principles.
4. Familiar with database storage structures and access techniques.
5. Use the PL/SQL code constructs of IF-THEN-ELSE and LOOP types as well as syntax and command functions.
6. Solve Database problems using Oracle 9i SQL and PL/SQL. This will include the use of Procedures, Functions, Packages, and Triggers

B.Sc. (Data Science)
Semester – III
Paper – II
Database Management Systems (3DST02)

UNIT - I:
CODD'S Rules, Oracle Database Objects, Sub Languages of SQL, Data types, Operators. DDL Statement: Creating Tables, Deriving Table from existing table, Altering, Dropping Tables. Integrity Constraints, Specifying Names for the Constraints, Viewing Integrity Constraints, Adding and Dropping Constraints. DML Statements: SELECT statement, Insert, Update, Delete, Working with Sequences and Synonyms. Built-in functions: Arithmetic, Date, Character, Conversion, Single row, Aggregate, Decode. Joins, Set Operators and Sub queries. DCL and TCL Statements: Grant, Revoke, Commit, Rollback and Savepoints.

UNIT - II:
VIEWS: Creating Views, Dropping Views, Inserting, Updating and Deleting Data using Views, Types of Views. PL/SQL Programming: PL/SQL Data Types, Identifiers, Operators and Expressions, Iterative Statements, Conditional Statements, emphasis on Problems

UNIT - III:
Exception Handling: Predefined Exceptions, User defined Exceptions. Cursors: Declaring Cursors, Opening and Retrieving Records, Closing cursors. Attributes of Explicit and Implicit Cursors, Parameter Passing in Cursors. Procedures: Create and Drop Procedure, Creating Procedures with Parameters, Calling Procedures, Granting the EXECUTE Permission Problems on Exception Handling, Cursors and Procedures.

UNIT – IV:
Function: Creating and Dropping Function, Purity Levels in Functions, Executing Functions. Triggers: Create Triggers, Type of Triggers, Creating BEFORE and AFTER Triggers, INSTEAD-OF Triggers, Trigger Predicates, Inserting, Updating and Deleting Triggers, Enabling, Disabling and Dropping Triggers. Problems on Functions and Triggers

Text Books:

1. Understanding ORACLE By Ivan Bayross [BPB Publication]
2. Gupta, Database Management Systems, McGraw-Hill.
3. Alexis Leon, Mathews Leon, Database Management System, Leaon Vikas.
4. Mike Mcgrath, SQL in Easy Steps, McGraw-Hill.
5. Silberschatz, Korth, Sudarshan, Database System concepts, McGraw-Hill.

RASHTRASANT TUKADOJI MAHARAJ NAGPUR UNIVERSITY, NAGPUR
THREE YEAR BACHELOR OF SCIENCE (B. Sc.) DATA SCIENCE DEGREE COURSE
(C B S)

B. Sc. (Data Science)		**Semester - III**	
Course Name : Statistical Inference (Paper – III)		Course Code: (3DST03)	
Periods per week (1 Period is 60 minutes)		4	
Credits		4	
		Hours	Marks
Evaluation Scheme	Theory Examination	3	80
	Internal Assessment	--	20
		Total Marks	100

Learning Objectives:
This course introduces students to the basic theory behind the development and assessment of statistical analysis techniques in the areas of point and interval estimation, as well as hypothesis testing. The student can perform point estimation, hypothesis testing and interval estimation under a large variety of discrete and continuous probability models. Further, the student can evaluate the properties of these estimators and tests, for both finite sample sizes and asymptotically as the sample size tends to infinity.

Learning Outcomes:
After completion of this course successfully the students will be able to:

1. To study distributions for categorical data.
2. To describe and make statistical inference for contingency tables.
3. Understand problem of statistical inference, problem of point estimation
4. Properties of point estimator such Consistency, Unbiasedness, Sufficiency
5. Obtain minimum variance unbiased estimator
6. Quantify information in statistic using Fisher Information
7. Construct minimal sufficient statistic and minimal sufficient statistic for exponential family.
8. Understand problem of statistical inference, problem of Interval estimation

B.Sc. (Data Science)
Semester – III
Paper – III
Statistical Inference (3DST03)

UNIT I:
Important concepts in Probability: Definition of Probability—Classical and relative frequency approach to Probability. Richard Von Mises, Cramer and Kolmogorov's approaches to Probability, merits and demerits of these approaches. Only general ideas to be given. Random Experiment: Trial, sample point and sample space, definition of an event, operation of events, mutually exclusive and exhaustive events. Discrete sample space, properties of Probability based on axiomatic approach.

UNIT II:
Conditional Probability, independence of events, Bayes' theorem and its applications. Chebyshev's inequality and applications with problems.

UNIT III:
Random variables: Definition of discrete random variables, Probability mass function, idea of continuous random variable, Probability density function, illustrations of random variables and its properties, expectation of a random variable and its properties.

UNIT IV:
Moments, measures of location, dispersion, skewness and kurtosis, Probability generating function (if it exists), moment generating function, their properties and uses.

Text Books:
1. Bhat B.R, Srivenkataramana T And RaoMadhava K.S. (1997): Statistics: A Begineer's Text,Vol II, New Age International (P) Ltd
2. Edward P.J, Ford J.S And Lin(1974): Probability for Statistical Decision making, Prentice Hall
3. Goon A. M,Gupta M. K, Das Gupta,B (1999): Fundamentals of Statistics, Vol II World Press, Calcutta.
4. Mood A.M,Graybill F.A And Boes D.C (1974): Introduction to the Theory of Statistics, McGraw Hill.
5. Freund J. E: Mathematical Statistics (Prentice Hall India)
6. P. L. Meyer: Introductory Probability and Statistical Applications, (Oxford and IBH)

RASHTRASANT TUKADOJI MAHARAJ NAGPUR UNIVERSITY, NAGPUR
THREE YEAR BACHELOR OF SCIENCE (B. Sc.) DATA SCIENCE DEGREE COURSE
(C B S)

B. Sc. (Data Science)		**Semester - III**	
Course Name : Data Communication and Network (Paper – IV)		Course Code: (3DST04)	
Periods per week (1 Period is 60 minutes)		4	
Credits		4	
		Hours	Marks
Evaluation Scheme	Theory Examination	3	80
	Internal Assessment	--	20
		Total Marks	100

Learning Objectives:
The objective of the course understand the basics of data communication, networking, internet and their importance, Analyze the services and features of various protocol layers in data networks. Analyze TCP/IP and their protocols. Differentiate wired and wireless computer networks, Student will be able to understand various types of transmission media, network devices; and parameters of evaluation of performance for each media and device.

Learning Outcomes:
After completion of this course successfully the students will be able to:

1. Understand and Contrast the concept of Signals, OSI & TCP/IP reference models and discuss the functionalities of each layer in these models.
2. Discuss and Analyze flow control and error control mechanisms and apply them using standard data link layer protocols
3. Analyze and apply various routing algorithms to find shortest paths for packet delivery.
4. Explain the details of Transport Layer Protocols (UDP, TCP) and suggest appropriate protocol in reliable/unreliable communication.
5. Understand the basic aspects of packet-based protocol design and implementation..

B.Sc. (Data Science)
Semester –III
Paper – IV
Data Communication and Network (3DST04)

UNIT - I :
Data Communication Data Transmission- Concept and Terminology, Analog & Digital Data Transmission, Transmission Impairment, Transmission Media. Data Encoding- Digital Data, Analog Data, Digital Signal, Analog Signal. Digital Data Communication: Asynchronous and Synchronous transmission, Error detection technique, interfacing. Data Link Control: Line configurations, Flow control, Error control, Data link control protocols.

UNIT - II :
Multiplexing-Frequency division multiplexing, Synchronous Time Division Multiplexing. : Circuit Switching: Communication Networks, Circuit switching, Single Node network, Digital switching concept, Control Signaling. Packet Switching: Packet switching principles, Virtual circuits and diagrams, Routing, Traffic control, X.25. LAN & MAN- LAN, MAN technology.

UNIT - III :
Bus, Tree, Star and Hybrid Topologies, Optical fiber Bus, Ring Topology, Medium Access Control Protocols, LAN/MAN standards. Internetworking: Principles of Internetworking, Bridges, Routers, Repeaters, Gateways, Connection Oriented Internetworking, Connectionless Internetworking, Connectionless Internetwork Protocol, Router-level protocol. Communication Architecture Protocols & Architecture: Protocols,

UNIT - IV :
The Layers Approach, OSI Model, TCP/IP protocol suite, System Network Architecture.
Transport Protocols- Transport services, Protocol Mechanism, Network services, ISO Transport Standards, TCP, UDP, TCP and UDP Packet format, Lightweight Transport Protocol.

Reference Books:

1. William Stalling, Data and Computer Communication, PHI Publication.
2. Forouzan, Data Communication and Networks, Tata McGraw Hill.
3. Godbole, Data Communication and Network, TMH
4. Tanenbum, Computer Networks , ,PHI Publication.
5. Comer Internetworking with TCP/IP Vol-1, PHI Publication

RASHTRASANT TUKADOJI MAHARAJ NAGPUR UNIVERSITY, NAGPUR
THREE YEAR BACHELOR OF SCIENCE (B. Sc.) DATA SCIENCE DEGREE COURSE
(C B S)

B. Sc. (Data Science)		**Semester - III**	
Course Name : Design and Analysis of Algorithms (Paper – V)		Course Code: (3DST05)	
Periods per week (1 Period is 60 minutes)		4	
Credits		4	
		Hours	Marks
Evaluation Scheme	Theory Examination	3	80
	Internal Assessment	--	20
		Total Marks	100

Learning Objectives:
The course introduces the basics of computational complexity analysis and various algorithm design paradigms. The goal is to provide students with solid foundations to deal with a wide variety of computational problems, and to provide a thorough knowledge of the most common algorithms and data structures. Able to Compare between different data structures and pick an appropriate data structure for a design situation.

Learning Outcomes:
After completion of this course successfully the students will be able to:

1. to apply knowledge of computing and mathematics to algorithm design;
2. to analyze a problem and identify the computing requirements appropriate for its solution;
3. to design, implement, and evaluate an algorithm to meet desired needs; and
4. to apply mathematical foundations, algorithmic principles, and computer science theory to the modeling and design of computer
   - Based systems in a way that demonstrates comprehension of the trade-offs involved in design choices.
   - An ability to apply design and development principles in the construction of software systems of varying complexity.
   - An ability to function effectively as a member of a team in order to accomplish a common goal.
   - Recognition of the need for and an ability to engage in continuing professional development.
   - An ability to use current techniques, skills, and tools necessary for computing practice.

Semester – III
Paper – V
Design and Analysis of Algorithms (3DST05)

Unit – I:
Algorithm, analysis, Time complexity and space complexity, O-notation, Omega notation and Theta notation, Heaps and Heap sort, Sets and disjoint set, union and find algorithms. Sorting in linear time. Tower of Hanoi, Heaps and Heap sort, **Divide and Conquer** General Strategy, Binary Search, Quick Sort, Merge Sort.

Unit – II:
Greedy Method: General Strategy, Knapsack problem Job sequencing with Deadlines, Optimal merge patterns, Minimal Spanning Trees, Dijkstra's algorithm. Optimal storage on tapes, Fractional Knapsack problem, Job Sequencing

Unit – III:
Linear Programming Various definitions, statements of basic theorems and properties, Advantages, Limitations and Application areas of Linear Programming, **Linear Programming** – The Graphical method – Graphical Solution methods of Linear Programming problem, Problems, Phase II of the Simplex Method, Primal and Dual Simplex Method, Big –M method.

Unit- IV:
Transportation Problem and its solution, Assignment Problem and its solutions by Hungarian Method. Sequential model and related Problems Processing n jobs through A) 1 machine and B) 2 machines, **Queuing Theory** Characteristics of Queuing Models, Transient and Steady states of the System, Model – I [(M/M/1) : (FCFS / ∞ /∞)], Model II – Generalization of Model, [(M/M/1) : (FCFS / ∞ / ∞)] (BirthDeath Process).

Text Books

1. Bressard, "Fundamental of Algorithm." PHI
2. Introduction to Operation Research : A Computer Oriented Algorithm Approach By Filet B. E.
3. Fundamentals of Queuing Theory By Gross D. and Ilaris C.M.
4 Horowitz/Sahani, "Fundamentals of computer Algorithms", Galgotia.
5 Magnifying Data Structures, Arpita Gopal : PHI Publications.
6 Introduction to Operation Research By Hiller F. and Lieberman G. J., TMH,8th Ed.
7 Operations Research By Kanti swarup, Gupta P.K. and ManMohan, S.Chand And Sons, 15th Ed.

RASHTRASANT TUKADOJI MAHARAJ NAGPUR UNIVERSITY, NAGPUR
THREE YEAR BACHELOR OF SCIENCE (B. Sc.) DATA SCIENCE DEGREE COURSE
(C B S)

B. Sc. (Data Science)		**Semester - III**	
Course Name : Object-Oriented Programming in Java Lab		Course Code: (3DSP01)	
Periods per week (1 Period is 60 minutes)		4	
Credits		2	
		Hours	Marks
Evaluation Scheme	Practical Examination	6-8*	50
		Total Marks	50

Practical's based on Java Programming and implementing the concept of Design and Analysis of Algorithms

RASHTRASANT TUKADOJI MAHARAJ NAGPUR UNIVERSITY, NAGPUR
THREE YEAR BACHELOR OF SCIENCE (B. Sc.) DATA SCIENCE DEGREE COURSE
(C B S)

B. Sc. (Data Science)		**Semester - III**	
Course Name : Database Management Systems Lab		Course Code: (3DSP02)	
Periods per week (1 Period is 60 minutes)		4	
Credits		2	
		Hours	Marks
Evaluation Scheme	Practical Examination	6-8*	50
		Total Marks	50

Practicals based on Database Management Systems Lab

RASHTRASANT TUKADOJI MAHARAJ NAGPUR UNIVERSITY, NAGPUR
THREE YEAR BACHELOR OF SCIENCE (B. Sc.) DATA SCIENCE DEGREE COURSE
(C B S)

B. Sc. (Data Science)		**Semester - IV**	
Course Name : Python Programming (Paper – I)		Course Code: (4DST01)	
Periods per week (1 Period is 60 minutes)		4	
Credits		4	
		Hours	Marks
Evaluation Scheme	Theory Examination	3	80
	Internal Assessment	--	20
		Total Marks	100

Learning Objectives:
This course is a foundation for the development of more advanced mathematical concepts, mainly emphasis on the use appropriate set, function, or relation models for analysis of practical examples and interpretation of the associated operations and terminology in context. To formulate problems precisely, solve the problems, apply formal proof techniques, and explain their reasoning clearly.

Learning Outcomes:

After completion of this course successfully the students will be able to:

1. Interpret the fundamental Python syntax and semantics and be fluent in the use of Python control flow statements.
2. Express proficiency in the handling of strings and functions.
3. Determine the methods to create and manipulate Python programs by utilizing the data structures like lists, dictionaries, tuples and sets.
4. Identify the commonly used operations involving file systems and regular expressions.
5. Apply Mathematical thinking, Mathematical proofs, and algorithmic thinking, and be able to apply them in problem solving.

B.Sc. (Data Science)
Semester-IV
Paper – I
Python Programming (4DST01)

Unit- I:

Introduction to Python Language: Overview, Features of Python, Execution of a Python Program, Innards of Python, Frozen Binaries, Python Interpreter, Comparison of Python with C and Java, Installing Python, Writing & Executing, IDLE Data Types,

Variables And Other Basic Elements: Comments, Docstrings, Data types-Numeric, Compound, Boolean, Dictionary, Sets, Mapping, Basic Elements of Python, **Variables Input and Output Operations:** Input Function, Output Statements, Command Line Arguments **Control Statements:** Control Statements- Loop Statement, The else Suite, break Statement, continue Statement, pass Statement, assert Statement, return Statement.

Unit- II:

Functions: Defining & Calling a Function, Returning Results, Returning Multiple Values, Built-in Functions, Parameters and Arguments, Recursive Functions, Anonymous or Lambda **Functions Operators:** Arithmetic operators, Assignment operators, Unary minus operator, Relational operators, Logical operators, Bitwise operators, Membership operators, Identity operators, Precedence of Operators, Associativity of Operators

Arrays: Creating Arrays, Indexing and Slicing, Basic Array Operations, Arrays Processing, Mathematical Operations on Array, Aliasing Arrays, Slicing and Indexing in NumPy Arrays, Basic Slicing. Advanced Indexing. Dimensions of Arrays, Attributes of an Array.

Strings: Creating Strings, Functions of Strings, Working with Strings, Length of a String, Indexing & Slicing, Repeating & Concatenation of Strings, Checking Membership, Comparing Strings, Removing Spaces, Finding Substrings, Counting Substrings, Strings are Immutable, Splitting and Joining Strings, Changing Case, Checking Starting and Ending of a String, Sorting & Searching in the Strings, Formatting the Strings, Working with Characters

Unit- III:

Lists and Tuples: Lists, List Functions and Methods, List Operations,

Tuples Dictionaries: Creating a Dictionary, Operators in Dictionary, Dictionary Methods, Using for Loop with Dictionaries, Operations on Dictionaries, Ordered Dictionaries

Regular Expressions: What is a Regular Expression? Sequence Characters in Regular Expressions, Quantifiers in Regular Expressions, Special Characters in Regular Expressions, Introduction to NumPy: Understanding Data Types in Python, The Basics of NumPy Arrays, Computation on NumPy Arrays: Universal Functions, Aggregations: Min, Max, and Everything In Between, Computation on Arrays: Broadcasting, Comparisons, Masks, and Boolean Logic, Fancy Indexing, Sorting Arrays, Structured Data: NumPy's Structured Arrays

Unit- IV:

Data Manipulation with Pandas: Introducing Pandas Objects, Data Indexing and Selection, Operating on Data in Pandas, Handling Missing Data, Hierarchical Indexing, Combining Datasets: Concat and Append, Combining Datasets: Merge and Join, Aggregation and Grouping, Pivot Tables, Vectorized String Operations, Working with Time Series. High-Performance Pandas:

eval() and query( ) Visualization with Matplotlib: Simple Line Plots, Simple Scatter Plots, Visualizing Errors, Density and Contour Plots, Histograms, Binnings, and Density, Customizing Plot Legends, Customizing Colorbars, Multiple Subplots, Text and Annotation, Customizing Ticks, Customizing Matplotlib: Configurations and Stylesheets, Three-Dimensional Plotting in Matplotlib, Geographic Data with Basemap, Visualization with Seaborn

Text Book

1. Programming through Python M. T. Savaliya, R.K Maurya, G.M Magar Staredu Solutions 1 st 2018
2. Python Data Science Handbook Jake VanderPlas O'Reilly Media 1st 2016
3. Let Us Python Y. Kanetkar, BPB 1 st 2019
4. Programming in Python 3 Mark Summerfield Pearson Education 2nd 2018
5. Learning Python Lutz M O'ReillyShroff 5th 2013
6. Beginning Python Magnus Lie Hetland Apress 2 nd 2009
7. Star Python Star Certification Star Certification 1 st 2018

RASHTRASANT TUKADOJI MAHARAJ NAGPUR UNIVERSITY, NAGPUR
THREE YEAR BACHELOR OF SCIENCE (B. Sc.) DATA SCIENCE DEGREE COURSE
(C B S)

B. Sc. (Data Science)		**Semester - IV**	
Course Name : Cloud Computing (Paper – II)		Course Code: (4DST02)	
Periods per week (1 Period is 60 minutes)		4	
Credits		4	
		Hours	Marks
Evaluation Scheme	Theory Examination	3	80
	Internal Assessment	--	20
		Total Marks	100

Learning Objectives:

The course will introduce domain and cover the topics of cloud infrastructures, virtualization, software defined networks and storage, cloud storage, and programming models. As an introduction, the course will discuss the benefits and challenges of the cloud, as well as service models, service level agreements (SLAs), security, example cloud service providers and use cases. Choose among various cloud technologies for implementing applications, Implement different types of Virtualization technologies and Service Oriented Architecture systems.

Learning Outcomes:

After completion of this course successfully the students will be able to:

1. Understand the fundamental principles of distributed computing.
2. Understand how the distributed computing environments known as Grids can be built from lower level services.
3. Understand the importance of virtualization in distributed computing and how this has enabled the development of Cloud Computing.
4. Analyze the performance of Cloud Computing.
5. Understand the concept of Cloud Security.
6. Learn the Concept of Cloud Infrastructure Model

B.Sc. (Data Science)
Semester-IV
Paper – 2
Cloud Computing (4DST02)

Unit- I:
Overview of Computing Paradigm : Recent trends in Computing: Grid Computing, Cluster Computing, Distributed Computing, Utility Computing, Cloud Computing, Introduction to Cloud Computing : Introduction to Cloud Computing, History of Cloud Computing, Cloud service providers, Benefits and limitations of Cloud Computing

Unit- II:
Cloud Computing Architecture : Comparison with traditional computing architecture (client/server), Services provided at various levels, Service ModelsInfrastructure as a Service(IaaS), Platform as a Service(PaaS), Software as a Service(SaaS), How Cloud Computing Works, Deployment Models- Public cloud, Private cloud, Hybrid cloud, Community cloud, Case study of NIST architecture. SE Case Studies : Case study of Service model using Google App Engine, Microsoft Azure, Amazon EC2,Eucalyptus

Unit- III:
Service Management in Cloud Computing : Service Level Agreements(SLAs), Billing & Accounting, Comparing Scaling Hardware: Traditional vs. Cloud, Economics of scaling.

Unit- IV:
Cloud Security : Infrastructure Security- Network level security, Host level security, Application level security, Data security and Storage- Data privacy and security Issues, Jurisdictional issues raised by Data location, Authentication in cloud computing.

Text Books:

1.Cloud Computing Bible, Barrie Sosinsky, Wiley-India, 2010
2. Cloud Computing: Principles and Paradigms, Editors: Rajkumar Buyya, James Broberg, Andrzej M. Goscinski, Wile, 2011 3.
3.Cloud Computing: Principles, Systems and Applications, Editors: Nikos Antonopoulos, Lee Gillam, Springer, 2012
4. Cloud Security: A Comprehensive Guide to Secure Cloud Computing, Ronald L. Krutz, Russell Dean Vines, Wiley-India, 2010
5. Gautam Shroff, Enterprise Cloud Computing Technology Architecture Applications, Adobe Reader ebooks available from eBooks.com,2010

RASHTRASANT TUKADOJI MAHARAJ NAGPUR UNIVERSITY, NAGPUR
THREE YEAR BACHELOR OF SCIENCE (B. Sc.) DATA SCIENCE DEGREE COURSE
(C B S)

B. Sc. (Data Science)		**Semester - IV**	
Course Name : Data Warehousing & Mining (Paper – III)		Course Code: (4DST03)	
Periods per week (1 Period is 60 minutes)		4	
Credits		4	
		Hours	Marks
Evaluation Scheme	Theory Examination	3	80
	Internal Assessment	--	20
		Total Marks	100

Learning Objectives:

To extract knowledge from data repository for data analysis, frequent pattern, classification and Prediction, This course will introduce the concepts of data ware house and data mining, which gives a complete description about the principles, used, architectures, applications, design and implementation of data mining and data ware housing concepts. Apply data mining techniques and methods to large data sets.

Learning Outcomes:

After completion of this course successfully the students will be able to:

1. Identify the scope and necessity of Data Mining & Warehousing for the society
2. Describe the designing of Data Warehousing so that it can be able to solve the root problems.
3. To understand various tools of Data Mining and their techniques to solve the real time problems.
4. To develop ability to design various algorithms based on data mining tools.
5. To develop further interest in research and design of new Data Mining techniques

B.Sc. (Data Science)
Semester-IV
Paper – III
Data Warehousing & Mining (4DST03)

UNIT I:
Introduction to Data Warehousing: Evolution of decision support systems, Failure of past decision support system, Operational v/s decision support systems, Data warehousing lifecycle, Architecture, Building blocks, Components of DW, Data Marts and Metadata

UNIT II:
Data Preprocessing: Why preprocess the data?, Descriptive data summarization, Data cleaning, Data integration and transformation, Data reduction, Data Discretization and Concept Hierarchy Generation.

UNIT III:
OLAP Analytical Processing: OLAP in Data warehouse, Demand for online analytical processing, need for multidimensional analysis, limitations of other analysis methods, OLAP definitions and rules, OLAP characteristics, major features and functions. OLAP models- ROLAP, MOLAP, HOLAP, Differentiation, Data cubes and operations on cubes. Introduction of Data Mining: Motivation, Importance, Data Mining functionalities, KDD and Data Mining, Data Mining v/s Query tools, Interesting patterns,

UNIT IV:
Architecture, Classification of Data Mining systems, Major issues from Data warehousing and Data Mining, Applications of Data Mining. Mining Frequent Patterns and Association: Basic Concepts: Market Basket analysis, motivating example, Frequent Item sets, Closed Item sets and Association rules, Frequent Pattern Mining Efficient and Scalable Frequent Item set. Mining Methods: Apriori Algorithm, Generating Association rules from Frequent Item sets, mining various kinds of association rules.

Text Books:

1. Data Mining (Concepts and Techniques) - Han and Kamber
2. Data Mining and Business Intelligence - Shinde and Chandrashekhar, Dreamtech Press

RASHTRASANT TUKADOJI MAHARAJ NAGPUR UNIVERSITY, NAGPUR
THREE YEAR BACHELOR OF SCIENCE (B. Sc.) DATA SCIENCE DEGREE COURSE
(C B S)

B. Sc. (Data Science)		**Semester - IV**	
Course Name : Business Research Methods (Paper – IV)		Course Code: (4DST04)	
Periods per week (1 Period is 60 minutes)		4	
Credits		4	
		Hours	Marks
Evaluation Scheme	Theory Examination	3	80
	Internal Assessment	--	20
		Total Marks	100

Learning Objectives:
The primary objective of this course is to develop a research orientation among the scholars and to acquaint them with fundamentals of research methods. Specifically, the course aims at introducing them to the basic concepts used in research and to scientific social research methods and their approach. It includes discussions on sampling techniques, research designs and techniques of analysis.

Learning Outcomes:

After completion of this course successfully the students will be able to:

1. Have an understanding of various kinds of research, objectives of doing research, research process research designs and sampling.
2. Be able to formulate research problem and develop a sufficiently coherent research design.
3. Have basic knowledge on qualitative, quantitative as well as measurement & scaling techniques.
4. Have a basic awareness of data analysis, including descriptive & inferential measures.
5. Be able to write & develop independent thinking for critically analyzing research reports.

B.Sc. (Data Science)
Semester-IV
Paper – IV
Business Research Methods (4DST04)

Unit-I:
Theory Building and Research Proposal - Meaning and nature of Theory, Nature of proposition, Scientific Method, Verifying Theory, Inductive and Deductive Reasoning, Nature of business problem, importance of problem definition, the process of problem definition, research questions and research objectives, research proposal, anticipating outcomes, Literature Review

Unit-II:
Sampling Design - Census and Sample Survey, Implication of Sample Design, Steps in Sample Design, Sampling Procedure, Different Types of Sampling Designs, How to select a random sample, Random sample from infinite universe, Sample size calculation.

Unit-III:
Measurement and Scaling Techniques - Measurement in research, Measurement Scales, Source of errors in measurement, Test for sound measurement, Techniques for developing measurement tools, Scaling - Meaning, Scale classification bases, Important Scaling Techniques, Scale construction techniques. **Methods of Data Collection** - Collection of Primary Data (Observation, Interview, Questionnaire, Schedules), Other methods of data collection, Collection of Secondary Data (Reliability, Suitability and Adequacy), Selection of appropriate method for data collection, Constructing Questionnaire/e-questionnaire/Schedule.

Unit-IV:
Interpretation and Report Writing - Meaning and Techniques of Interpretation, Significance and Steps of Report Writing, Layout of research reports, types of reports, Mechanics and Precautions for writing a research report.

Text Books:

1. Panneerselvam, "Research Methodology", Prentice Hall India
2. Donald Cooper and Pamela Schindler, "Business Research Methods ", Tata McGraw Hill
3. Krishnaswami & Rangantham, "Methodology of Research ", Himalaya Publishing House
4. Alan Bryman and Emma Bell, "Business Research Methods ", Oxford Publication

RASHTRASANT TUKADOJI MAHARAJ NAGPUR UNIVERSITY, NAGPUR
THREE YEAR BACHELOR OF SCIENCE (B. Sc.) DATA SCIENCE DEGREE COURSE
(C B S)

B. Sc. (Data Science)		**Semester - IV**	
Course Name : Time Series Analysis (Paper – V)		Course Code: (4DST05)	
Periods per week (1 Period is 60 minutes)		4	
Credits		4	
		Hours	Marks
Evaluation Scheme	Theory Examination	3	80
	Internal Assessment	--	20
		Total Marks	100

Learning Objectives:

The aim of this course is to provide students with the essential expertise to handle modern time series techniques. Idea is to introduce students to comprehensive set of tools and techniques for analyzing various forms of univariate and multivariate time series and for understanding the current literature in applied time series. After the course students will also be able to appreciate and apply key concepts of estimation and forecasting in a time series context. Endeavor will be to provide simple examples that illustrate how the theoretical results are used and applied in Practice

Learning Outcomes:

After completion of this course successfully the students will be able to:

1. To introduce discrete and continuous time stochastic processes
2. Forecast the trend pattern exhibited by the given data by using various methods
3. Run and interpret time series models and regression models for time series
4. time series data, its applications to various fields and components of time series

B.Sc. (Data Science)
Semester-IV
Paper – V
Time Series Analysis (4DST05)

Unit-I:
Basics of Time series : A model Building strategy, Time series and Stochastic process, stationarity, Auto correlation, meaning and definition – causes of auto correlation - consequence of autocorrelation – test for auto – correlation. Study of Time Series model and their properties using correlogram, ACF and PACF. Yule walker equations

Unit-II:
Time Series Models : White noise Process, Random walk, MA, AR, ARMA and ARIMA models, Box- Jenkins's Methodology fitting of AR(1), AR(2), MA(1), MA(2) and ARIMA(1,1) process. Unit root hypothesis, Co-integration, Dicky Fuller test unit root test, augmented Dickey – Fuller test.

Unit-III:
Non-linear time series models :ARCH and GARCH Process, order identification, estimation and diagnostic tests and forecasting. Study of ARCH (1) properties. GARCH (Conception only) process for modelling volatility.

Unit- IV:
Multivariate Liner Time series : Introduction, Cross covariance and correlation matrices, testing of zero cross correlation and model representation. Basic idea of Stationary vector Autoregressive Time Series with order one: Model Structure, Granger Causality, stationarity condition, Estimation, Model checking.

Text Book:

1. Box, G. E. P. and Jenkins, G. M. (1976). Time Series Analysis – Forecasting and Control, Holden – day, San Francisco.
2. Chatfield, C. (2003) Analysis of Time Series, An Introduction, CRC Press.
3. Ruey S. Tsay (2005). Analysis of Financial Time Series, Second Ed. Wiley & Sons.
4. Ruey S. Tsay (2014). Multivariate Time series Analysis: with R and Financial Application, Wiley & Sons.
5. Introduction to Statistical Time Series : W.A. Fuller

RASHTRASANT TUKADOJI MAHARAJ NAGPUR UNIVERSITY, NAGPUR
THREE YEAR BACHELOR OF SCIENCE (B. Sc.) DATA SCIENCE DEGREE COURSE
(C B S)

B. Sc. (Data Science)		**Semester - IV**	
Course Name : Python Programming Lab		Course Code: (4DSP01)	
Periods per week (1 Period is 60 minutes)		4	
Credits		2	
		Hours	Marks
Evaluation Scheme	Practical Examination	6-8*	50
		Total Marks	50

Practical's based on Python Programming Lab and implementing concepts of Business Research Methods using Python Programming

RASHTRASANT TUKADOJI MAHARAJ NAGPUR UNIVERSITY, NAGPUR
THREE YEAR BACHELOR OF SCIENCE (B. Sc.) DATA SCIENCE DEGREE COURSE
(C B S)

B. Sc. (Data Science)		**Semester - III**	
Course Name : Data Mining Lab		Course Code: (4DSP02)	
Periods per week (1 Period is 60 minutes)		4	
Credits		2	
		Hours	Marks
Evaluation Scheme	Practical Examination	6-8*	50
		Total Marks	50

Practical's based on Data Mining Lab and Time Series Analysis