Rayat Shikshan Sanstha's

Sadguru Gadage Maharaj College, Karad.

(An Autonomous College)

Undergraduate Programme

B. Sc. II Data Science

Syllabi of the course

(To be implemented w.e.f. from June 2025)

1) General Objectives of the Course: -

1. To create graduates with sound knowledge of Data Science, who can contribute towards recent advances in technology.
2. To provide advanced and in-depth knowledge of data science and specialization in one or two subjects of the new era of technology.
3. To prepare Graduates who will achieve peer-recognition, as an individual or in a team, through demonstration of good analytical, design, programming, and implementation skills.
4. To enable students, pursue a professional career in Data Science in related industry, business and research.
5. To impact industry knowledge and practical skills of current trends in IT field to the students.
6. To develop the ability among students to formulate, analyze and solve real life problems faced in Computer Science industry. To produce computer science professionals who can be directly employed or start his/her own work as
    - Data Scientist.
    - Business Analyst.
    - Data Analytics Manager.
    - Data Architect.
    - Data Administrator.
    - Business Intelligence Manager.
    - Entrepreneur in Computer Science industry.
7. To Develop designing, analyzing and critical thinking skill among students.

2) Duration:

The course shall be a full-time course.
The course shall be for **three** years, consisting of **six semesters**.

3) Fee Structure:

• **Course Fee:** as prescribed by the Institute.

4) Eligibility for Admission:

• As per Rule (2) for graduates of this Institute.

5) Medium of instruction:

-English

6) OUTCOMES:

After completing this courses student shall be expert in following things:

1. Avail yourself of Current trends in IT Industries and new Technologies.
2. Apply knowledge of programming platforms in Data Science and AI in real life.
3. Students should avail detailed knowledge of Data Science, Artificial Intelligence, Machine Learning, and Big Data etc.

4. Demonstrate their ability of advanced programming to design and develop innovative applications.
5. Access, evaluate, understand, and compare digital information from various sources and apply it for scientific knowledge acquisition as well as scientific data analysis and presentation.
6. Critically evaluate, analyze, and comprehend a scientific problem. Think creatively, experiment and research into innovation and creatively design scientific solutions to problems.
7. Exemplify generate a solution independently, check and validate it and modify if necessary.
8. Translate project plans, use management skills, and lead a team for planning and execution of a task.
9. Can start his own business or start up.

7) SCOPE:

After Successful completion of three years Degree in Data Science, we observed that the students have the ample opportunities in diversified areas such as:

1. Data Scientist
2. Data Analytics
3. Machine learning Engineer
4. Business Intelligence Analyst
5. Data Engineer.
6. Research Scientist.
7. Data Consultant

B.Sc. Data Science Part-II Semester- III
Computer Science Paper-I
Course Code-BDST24-301
Course Title: Object Oriented Programming Using C++
Total Hours: 30 hrs (38 lectures of 48 min)
Credits: 02
Total Marks: 40+10=50

Course outcomes:

The student should-

- Understand basic concepts of object oriented programming.
- Able to use various control structures to improve programming logic.
- Design classes and objects.
- Able to use constructor and destructor.
- Utilize the OOP techniques like operator overloading, inheritance, and polymorphism.

UNIT 1: Object Oriented Concepts (7)

- Difference between POP and OOP
- Concepts of OOP- Class, Object, Data abstraction, Encapsulation, Inheritance, Polymorphism, Data Binding, Exception handling.
- Basics of C++
- Terminology-Tokens, Operators, Keywords, Identifiers, constants.
- Basic data types -Structure of C++ program. Input and output streams.

UNIT 2: Operators in C++ (10)

- Dynamic Memory allocation (New and Delete), this pointer. Dynamic initialization of variable, reference variables.
- Control structures- Branching and looping statements.
- Features of OOP: Classes and objects-Definitions, class declaration, Member function
- Access modifiers: private, public and protected, Defining Member functions, static data members.
- Array of objects, passing object as parameter, inline function, reference arguments
- Friend function and friend class.

UNIT 3: Constructors and Destructor (6)

- Constructors- Definition, types- Default constructor, Copy constructor, Parameterized constructor. Destructors.
- Operator overloading-Definition overloading unary and binary perators.
- Overloading operators using friend function. Rules for overloading operator.

UNIT 4: OOPs in C++ (7)

- Inheritance- Defining base and derived class.
- Types of Inheritance–Single, multiple, multilevel, hierarchical, hybrid.
- Polymorphism-Definition. Types of polymorphism. Virtual function.

Reference Books:

1. Object oriented programming By E. Balagurusamy.
2. C++ Programming–By D. Ravichandran
3. Let Us C++By Yashawant Kanetkar.
4. Object Oriented Programming in C++-Dr. G.T. Thampi, Dr. S. S. Mantha
5. Mastering C++- By Venugopal.

Computer Science Paper-IV
Course Code-BDST24-302
Course Title: Data Science Using Python
Total Hours: 30 hrs (38 lectures of 48 min)
Credits: 02
Total Marks: 40+10=50

UNIT 1: Introduction to Data Science (8)

Data Science: Introduction, Data Analysis Sequence, Data Acquisition Pipeline, Report Structure

Files and Working with Text Data: Types of Files, Creating and Reading Text Data, File Methods to Read and Write Data, Reading and Writing Binary Files, The Pickle Module, Reading and Writing CSV Files, Python os and os.path Modules.

Working with Text Data: JSON and XML in Python

Unit – II: Data Processing (7)

Working with Text Data: Processing HTML Files, Processing Texts in Natural Languages

Regular Expression Operations: Using Special Characters, Regular Expression Methods, Named Groups in Python Regular Expressions, Regular Expression with glob Module

Unit – III: Working with Databases and Numeric Data (8)

Working with Databases: Setting Up a MySQL Database, Using a MySQL Database: Command Line, Using a MySQL Database, Taming Document Stores: MongoDB

Working with Tabular Numeric Data (Numpy with Python): NumPy Arrays Creation Using array() Function, Array Attributes, NumPy Arrays Creation with Initial Placeholder Content, Integer Indexing, Array Indexing, Boolean Array Indexing, Slicing and Iterating in Arrays, Basic Arithmetic Operations on NumPy Arrays, Mathematical Functions in NumPy, Changing the Shape of an Array, Stacking and Splitting of Arrays, Broad casting in Arrays.

Unit – IV: Data Series and Frames (7)

Working with Data Series and Frames: Pandas Data Structures, Reshaping Data, Handling Missing Data, Combining Data, Ordering and Describing Data, Transforming Data, Taming Pandas File I/O, **Plotting:** Basic Plotting with PyPlot, Getting to Know Other Plot Types, Mastering Embellishments, Plotting with Pandas

Course Code-BDSP24-303
Course Title: Practical III
Credits: 02
Total Marks: 50
Lab on OOP and Python

Practical List:

1. Programs based on branching and looping statements.
2. Programs based on constructor and destructor.
3. Programs based on Friend Class
4. Programs based on friend function.
5. Programs based on function overloading concept
6. Programs based on operator overloading concept
7. Programs based on member functions.
8. Programs based on use of constructor and destructor
9. Programs based on inheritance.
10. Programs based on polymorphism

Practical in Python

1. Write programs to parse text files, CSV, HTML, XML and JSON documents and extract relevant data. After retrieving data check any anomalies in the data, missing values etc.
2. Write programs for reading and writing binary files
3. Write programs for searching, splitting, and replacing strings based on pattern matching using regular expressions
4. Design a relational database for a small application and populate the database. Using SQL do the CRUD (create, read, update and delete) operations.
5. Create a Python MongoDB client using the Python module pymongo. Using a collection object practice functions for inserting, searching, removing, updating, replacing, and aggregating documents, as well as for creating indexes
6. Write programs to create numpy arrays of different shapes and from different sources,

reshape and slice arrays, add array indexes, and apply arithmetic, logic, and aggregation functions to some or all array elements

7. Write programs to use the pandas data structures: Frames and series as storage containers and for a variety of data-wrangling operations, such as:

- Single-level and hierarchical indexing
- Handling missing data
- Arithmetic and Boolean operations on entire columns and tables
- Database-type operations (such as merging and aggregation)
- Plotting individual columns and whole tables
- Reading data from files and writing data to files

Statistics

Class: B.Sc. II DS (CBCS) Curriculum with effect from June 2025

DS Syllabus

Preamble:

This syllabus is framed to give sound knowledge with understanding of Statistics to undergraduate students at first year of three years of B.Sc. (DS) degree course.

Students learn Statistics as a separate subject from B.Sc. II (DS) .The goal of the syllabus is to make the study of Statistics popular, interesting and encouraging to the students for the higher studies.

General Objectives of the Program:

- Provide students with learning experiences that develop broad knowledge and understanding of key concepts of Statistics
- Develop abilities in students to design and develop innovative solutions for benefits of society, leadership, teamwork and lifelong learning..
- To develop scientific attitude among the student so as to make students curious and open minded.
- Provides Students with skills that enable them to get employment in industries or pursue higher studies.

General outcomes of Program:

1. The students will graduate with proficiency in the subject of their choice.
2. The students will be eligible to continue higher studies in the subject.
3. The students will be eligible to appear the examination for jobs in government organizations.
4. The students will be eligible to apply for jobs with a minimum requirement of B.Sc. (DS)

Program Specific Outcomes:

The studies will acquire;

1. Knowledge about the Testing of hypothesis.
2. Knowledge about the univariate, bivariate data analysis.
3. Knowledge about the advanced excel,Minitab ,SPSS
4. Knowledge of different methods of estimation about inference of parameter of standard discrete and continuous probability distribution.

B.Sc. Part II (DS)

1. **Title** : Statistics
2. **Year of implementation**: The syllabus will be implemented from June 2025 onwards
3. **Duration** : The course shall be a fulltime
4. **Pattern:** Semester examination
5. **Medium of Instruction** : English
6. **Structure of Course** :

B.Sc.-II (DS) : Semester –III

Subject Code	Title of theory Paper	Credit	Lectures /Practical per week	TH/PR	SEE		CCE		Total Theory/ Practical Marks
					Max	Min	Max	Min	
BDST24-304	Testing of Hypothesis	02	2	TH	40	16	10	4	50
BDST24-305	Basic Excel	02	2	TH	40	16	10	4	50
BDSP24-306	Practical -III	02	4	PR	50	20	---	---	50

B.Sc.-II(DS) : Semester –IV

Subject Code	Title of theory Paper	Credit	Lectures /Practical per week	TH/PR	SEE		CCE		Total Theory/ Practical Marks
					Max	Min	Max	Min	
BDST24-404	Advanced Excel	02	2	TH	40	16	10	4	50
BDST24-405	Continuous Probability distributions	02	2	TH	40	16	10	4	50
BDSP24-406	Practical - IV	02	4	PR	50	20	---	---	50

Titles of Papers of B.Sc. course:

B.Sc.II (DS) Semester-III :

Theory:

BDST24- 304: Testing of Hypothesis
BDST24- 305: Basic Excel

Practical (Semester):

BDSP24-306: Practical - III

B.Sc.II (DS) Semester-IV

Theory:

BDST24- 404: Advanced Excel
BDST24- 405: Continuous Probability distributions

Practical (Semester):

BDSP24-406: Practical - IV

B. Sc. II(DS) (Semester-III)

BDST24- 304: Testing of Hypothesis

Theory- 30 Hours Credits:02

Objectives:

The main objectives of this course are:

1. To understand the concept of Hypothesis Testing.
2. Applying the appropriate small sample tests and large sample tests in various situations.

Syllabus Contents

Unit-1 **15hrs.**

2.1: Testing of Hypothesis - I:

Notion of Population, Sample, Parameter, Statistic, Sampling distribution of Statistics $\bar{X}$ and S^2 when sample is drawn from normal distribution (statement only).Hypothesis, Simple and composite hypothesis, Null and alternative hypothesis, type I and type II errors, Critical region, level of significance, p-value.One and two tailed test, power of test.And real life examples.

Unit-2 **15hrs.**

2.2: **Testing of Hypothesis - II:** General procedure of testing of hypothesis.

Small Sample Tests:

t- test : Test for means: i) $H_0: \mu = \mu_0$,ii) $H_0: \mu_1 = \mu_2$, (where $\sigma_1^2 = \sigma_2^2$), iii) Paired t-test
χ^2 – test : Test for population variance $H_0: \sigma^2 = \sigma_0^2$ (Mean Known and unknown)
F – test : Test for equality of two population variances $H_0: \sigma_1^2 = \sigma_2^2$

Large Sample Tests:

A) Tests for means: i) Testing of population mean; $H_0: \mu = \mu_0$
ii) Testing equality of population means; $H_0: \mu_1 = \mu_2$

B) Tests for Proportion: i) Testing of population Proportion; $H_0: P = P_0$
ii) Testing equality of population Proportion; $H_0: P_1 = P_2$

C) Test for population correlation by using Fisher's Z- transformation:
i)Testing of population correlation; $H_0: \rho = \rho_0$

ii) Testing equality of population correlations;H_0: $\rho_1 = \rho_2$

D) χ^2– tests for :i)Goodness of fit of given probability distribution and

ii)Test for independence of attributes when data is in the form of:

a) m x n contingency table

b) 2 x 2 contingency table, Yate's correction for continuity.

And real life examples.

Reference Books:

1. Fundamental of mathematical statistics: Gupta &Kapoor
2. Statistical Methods: S.P.Gupta, Sultan Chand and Sons Publishers
3. Business Statistics: S.L. Agarwal , Kalyani Publishers
4. Introduction to Statistics Methods : C.B. Gupta and Vijay Gupta, Vikas Publication
5. Business Statistics :G.V Kumbhojkar, PhadkePrakashan

BDST24- 305: Basic Excel

Theory- 30 Hours Credits: 02

Objectives:

The main objectives of this course are:

i) To understand Ms-Excel

ii) To understand various data import methods in excel

iii) To understand data manipulation in excel

iv) To create visualizations and plots using excel

Unit1 (1credit)

1.1 IntroductiontoMs-Excel

1.2 Using of functions and formulae in excel to calculate Measuresofcentraltendency-ArithmeticMean,Geometricmean,Harmonicmean, Median, Quartiles Deciles, Percentiles, Mode.

1.3 Using of functions and formulae in excel to calculate , Measures ofdispersion -Range, Quartile deviation, Absolute measure, Relative measure, Interquartilerange,Meandeviation,Variance,StandardDeviation,Coefficientofvariatio.

Unit2 (1credit)

2.1 Using of functions and formulae in excel to calculate ,Correlation,Karl-Pearson Correlation coeff(r),Spearman's correlationcoeff.(R), Regreession.

2.2 Graphicalrepresentation**,**Sub–dividedbardiagram,multiplebardiagram,subdividedbar diagram, scatter diagram, frequency curve, frequency Polygon,

2.3 Using of functions and formulae in excel to calculate , Fittingofdistributions**,**Binomial,Poisson,NegativeBinomial,Hypergeometric, Exponential &Normal Also test the goodness of fit and model sampling.

Reference Books:

i) Kore B.G-. MS-EXCEL for Data Analysis, Nirali Publication (2024)
ii) Paul McFedries -Microsoft Excel Formulas and functions ,
iii) Excel 2007 Bible – by John Walkenbach
iv) Handbook Microsoft Excel Till you understand – Department of Statistics

BDSP24-306: Practical - III

Practical - 60 Hours Credits:02

Objectives:

The main objectives of this course are:

1. Test various hypothesis about parameters of specified distribution for given data.

2. Fit and test the goodness of fit of specified distribution for given data.

3. Efficiently manipulate the data in MS-Excel using use basic tools like find and replace, sort and filter.

4. Visualize the data using MS-Excel.

5. Develop a deep understanding of Excel formulas and functions in MS-Excel, enabling them to perform mathematical, trigonometric, statistical, date and time operations.

List of Practical's:

1) Basic functions Using Ms-excel
2) Measures of central tendency
3) Measures of Dispersion
4) Correlation and regression
5) Graphical representation
6) Fitting of Distributions
7) Large sample tests for means.
8) Large sample tests for proportions.
9) Tests based on Chi square distribution.(Test for population variance, Test for goodness of fit.)Tests for independence.
10) Tests based on t distribution ($\mu = \mu_0$, $\mu_1 = \mu_2$; paired t test)
11) Tests based on F distribution. ($\sigma_1^2 = \sigma_2^2$)
12)Error handling and formula Auditing in Excel

B.Sc. Data Science Part-II Semester- III
Open Elective Paper-III
Course Code-OEDST24-301
Course Title: Computer Organization (30)

UNIT I: MEMORY ORGANISATION: (07)

Introduction, Characteristics of memory systems ,Main memory design, Memory hierarchy, Cache memory, Memory mapping, Virtual Memory, Memory management concepts (paging and segmentation).

UNIT II: CONTROL UNIT: (07)

Introduction, Hardware control-design methods, Microprogrammed control unit, Microinstruction addressing and architecture of typical micro programmed control unit.

UNIT III I/O ORGANISATION: (08)

Peripheral devices, Input output Interface, IO mapped IO, Memory mapped IO, Asynchronous data transfer, Modes of transfer, Priority Interrupts, Direct memory access, Input output Processor, serial communication.

UNIT IV CPU ORGANISATION: (08)

Introduction, General register organization, Stack Organization, Instruction formats, Addressing modes Program Control, Arithmetic and Logic Unit (One bit and multiple bit), Bit processor.

Reference Books

Computer Organization - J.P. Hays TMH

B.Sc. Data Science Part-II Semester- III
SEC Paper-I
Course Code-SECDSP24-301
Course Title: **Power BI for Data Analysis and Visualization**

Credits: 02

Total Marks: 50

Unit 1: Power BI Essentials – Data Preparation & Modeling

Introduction to Business Intelligence (BI): Understanding BI concepts and their applications, Power BI ecosystem overview: Desktop, Service, and Mobile

Getting Started with Power BI Desktop User interface walkthrough ,Creating and saving projects ,Connecting to data sources (Excel, CSV, Web)

Data Transformation using Power Query: Basic cleaning: renaming, removing columns/rows, data types, Using "Add Column from Example", Append and Merge Queries, Combine files from a folder, Introduction to M Query editor and advanced transformations, Parameters and templates for reusable queries

- **Data Modeling in Power BI**: Understanding table relationships ,Creating and editing relationships, Model enhancements: formatting, hiding, categorizing data ,What-If parameters for simulation and scenarios

Introduction to DAX (Data Analysis Expressions):Creating calculated columns and tables, Common DAX functions: SUM, AVERAGE, COUNT, IF ,Navigation and logical functions ,Creating basic measures and calculated fields ,Intro to time intelligence functions (YTD, MTD, Previous Year)

Unit 2: Data Visualization, Power BI Service & Advanced DAX

- **Report Creation & Visualizations:** Building visuals: bar, line, pie, card, matrix, map ,Interactive elements: slicers, filters, drill-down, bookmarks ,Custom visuals and storytelling features ,Visualizing tabular, categorical, trend, and goal-tracking data
- **Using Power BI Service (Cloud Platform)**Publishing reports from Desktop to Power BI Service, Creating and sharing dashboards ,Power BI Q&A (natural language queries),Setting up App Workspaces for collaboration, Subscriptions and alerts, Export and embed reports, Excel and Power BI integration
- **Data Refresh and Scheduling**: Manual vs automatic refresh, Installing data gateway, Scheduling refreshes for cloud datasets
- **Power BI Mobile**: Overview of the Power BI Mobile App, Designing mobile-optimized reports, Viewing and interacting with mobile dashboards
- **Advanced DAX Functions**: Logical and conditional functions (IF, SWITCH) , Time intelligence deep dive (TOTALYTD, SAMEPERIODLASTYEAR) , Table functions (FILTER, SUMMARIZE) , X-Functions and A-Functions (e.g., CALCULATE, ALL) ,DAX as a query language (EVALUATE, RETURN)

B.Sc. Data Science Part-II Semester- III
Vocational Skill Course Paper-I
Course Code-VSCDSP24-301
Course Title: **Computer Hardware**

Credits: 02

Total Marks: 50

Practical-I (Hardware Components Identification):

Identify and handling of Internal components in the PC Cabinets like SMPS and its connection to Motherboard and various devices, Motherboard, CPU, Chipset, Slots, Memory modules, memory slots, Hard Disc Drives, CDROM/DVD/Blue-Ray Disc, etc.

Practical-II (System Integration):

Assembly of PC using various parts, Interconnection between devices, cable polarities and connections, SMPS installation and power connection. Various types of Add-on Cards, Motherboard slot and their application.

Practical-III (OS Installation I) :

Multiple HDD installation and creation single large volume out of it,

Practical-IV (OS Installation II) :

Installation of Operating System like Windows8, Windows 2008

Practical-V (OS and other software Installation) :

Installation of Operating System Linux flavors like Ubantu, SUSE, RedHat, Introduction to VMware Virtualization etc.

Practical-VI to X (Troubleshooting) :

Troubleshooting of various hardware problems like SMPS failure, Display not there, missing OS or re-installation of user software or system software. CDROM, DVD lens cleaning or replacement, CMOS setup, Battery replacement on motherboard in case BIOS is not retaining correct values. Driver software downloading and installation, Antivirus Software installation, scanning for viruses, removing .tmp files from WINDOWS machine, etc

B.Sc. Data Science Part-II Semester- III

AEC Paper-I

Course Code-English

B.Sc. Data Science Part-II Semester- III

Co-curriculum Course

Course Code

SEMESTER- IV

Course Code-BDST24-401
Course Title: Data Structure and Algorithm
Total Hours: 30 hrs (38 lectures of 48 min)
Credits: 02
Total Marks: 40+10=50

Unit-1 : Concepts of Data structure, Array (6)

Concept of Data, Data Object, Types of Data- Atomic Data, Non-atomic Data

Concept of Data Structure Abstract data type (ADT), Array Definition, Array Operations,

Applications of Array(Polynomial evaluation and addition of two polynomials), Multi-dimensional arrays.

Unit-2 : Algorithm Analysis (7)

Space complexity, time complexity

Asymptotic notation (Big O, Omega Ω, Theta θ)

Searching algorithms- Linear search, binary search and their algorithms

Sorting algorithm-Bubble Sort, insertion sort, selection sort, quick sort and their algorithms.

Unit-3 : Stack and Queue (8)

Stack, Concept of Stack, Operations on Stack-push, pop, peek

Array implementation of Stack, Linked List implementation of Stack Applications of Stack-Recursion, Infix, Prefix, Postfix, conversion from Infix to Prefix and Infix to Postfix Queue,

Concepts of queue, Operations on Queue-Insert, Delete, peek, Array implementation of queue,

Linked List Implementation of Queue, Types of Queue-Linear, Circular and Priority, Applications of Queue

Unit-4 : Linked List and Tree (9)

Linked List, Concept of Linked List, Memory representation of Linked List

Operations on Linked List (Insertion, Deletion, Display and Search)

Types of Linked List: Singly, Doubly Linked List & Circular Linked List Tree- Concept of Tree,

Tree terminology (root, child, parent, sibling, descendent, ancestor, leaf/external node, branch node/internal node, degree, edge, path, level, depth, height of node, height of tree, forest)

Binary Tree- definition , types (Full/Proper / Plane, Complete, Perfect, Skewed, Balanced)

Binary search tree- Operations on BST – Create, Insert, Search, Delete, traversals (Preorder, Inorder, Post order)

Reference Books :

1. Data structure through C++- Yashwant Kanitkar (BPB Publications)

2. Principles of Data structures using C++ - Vinu V. Das(New Age International Publication)

3. Data Structures with C- SEYMOUR LIPSCHUTZ(Tata McGraw-Hill)

4. Data structures, Algorithms and Applications in C++, S. Sahni, University Press

(India) Pvt. Ltd, 2nd edition, Universities Press Orient Longman Pvt. Ltd.

Course Code-BDST24-402
Course Title: Advanced SQL
Total Hours: 30 hrs (38 lectures of 48 min)
Credits: 02
Total Marks: 40+10=50

Unit I Structured Query Language: (7)

Writing Basic SQL Select Statements, Restricting and Sorling Data, Single-Row Functions, Joins (Displaying Data from Multiple Tables),Aggregating Data using Group Functions, Subqueries, Manipulating Data, Creating and Managing Tables, Including Constraints, Creating Views, Creating other Database Objects (Sequences, Indexes and Synomyms)

Unit II Advanced SQL: (8)

Controlling user Access, using SET operators, Data Time Functions, Enhancements to Group by clause (cube, Rollup and Grouping), Advanced Subqueries (Multiple column subqueries, Subqueries in FROM clause, Scalar and correlated subqueries), WITH Clause, Hierarchical retrieval.

Unit III PLSQL: (7)

Introduction, Overview and benefits of PL/SQL, Subprograms, types of PL/SQL blocks, Simple Anonymous Block, Identifiers, types of identifiers, Declarative Section, variables,

Scalar Data Types, The % Type attribute, bind variables, sequences in PL/SQL expressions, Executable statements, PL/SQL block syntax, comment the code.

Control Structures: Conditional processing using IF statements and CASE statements, Loop Statement, while loop statement, for loop statement, the continue statement

Unit IV Stored Procedures: (7)

Create a Modularized and Layered Subprogram Design, the PL/SQL Execution Environment, differences between Anonymous Blocks and Subprograms, Create, Call, and Remove Stored Procedures, Implement Procedures Parameters and Parameters Modes, View Procedure Information, Stored Functions.

Dynamic SQL:

The Execution Flow of SQL, Declare Cursor Variables, Dynamically Executing a PL/SQL Block, Configure Native Dynamic SQL to Compile PL/SQL Code, invoke DBMS_SQL Package.

Reference Books:

Murach's Oracle SQL and PLSQL by Joel Murach, Murach and Associates.

Oracle Database 11g PL/SQL Programming Workbook, ISBN:9780070702264,
By: Michael McLaughlin, John Harper, Tata McGraw-Hill.

Oracle PL/SQL Programming, Fifth Edition By Steven Feuerstein, Bill Pribyl

Oracle 11g : SQL Reference Oracle press

Oracle 11g : PL/SQL Reference Oracle Press.

Expert Oracle PL/SQL, By : Ron Hardman, Michael McLaughlin, Tata McGraw-Hill

Oracle database 11g : hands on SQL/PL SQL by Satish Asnani (PHI) EEE edition

Course Code-BDSP24-403
Course Title: Practical IV
Credits: 02
Total Marks: 50
Lab on Data structure & Algorithm and Advanced SQL

Practical List:

1. Write a C++ programs to implement recursive i) Linear search ii) Binary search

2. Write a C++ programs to implement
   i) Bubble sort ii) Selection sort
   iii) quick sort iv) insertion sort
3. Write a C++ programs to implement the following using an array.
   a) Stack ADT b) Queue ADT

4. Write a C++ programs to implement list ADT to perform following operations:
   a) Insert an element into a list. b) Delete an element from list

c) Search for a key element in list d) Count number of nodes in list

5. Write C++ programs to implement the following using a singly linked list.
   a) Stack ADT b) Queue ADT

6. Write a C++ program to perform the following operations:
   a) Insert an element into a binary search tree.
   b) Delete an element from a binary search tree.
   c) Search for a key element in a binary search tree.

7. Write C++ programs for implementing the following sorting methods:
   a) Insertion Sort b) Bubble Sort
   c) Selection Sort d) Quick Sort

1. Create Tables with Appropriate Constraints
   Concepts: DDL, Primary key, Foreign key, NOT NULL, CHECK
   Program: Create Employee and Department tables using proper constraint.

2. Write SELECT Query with Filtering and Sorting
   Concepts: SELECT, WHERE, ORDER BY
   Program: Retrieve employee names and salaries where salary > 50,000, sorted by name.

3. INNER JOIN Between Tables
   Concepts: JOINs
   Program: Display employee names with their corresponding department names using an inner join.

4. Aggregate Functions with GROUP BY
   Concepts: SUM, AVG, COUNT, GROUP BY
   Program: Display average salary and total employees in each department.

5. PL/SQL Block with IF and LOOP Statements
   Concepts: PL/SQL, Control Structures
   Program: Write a PL/SQL block to check if a salary is above 60,000, and loop through numbers 1 to 5.

6. Stored Procedure with IN Parameter Concepts: Stored Procedure, Parameters
Program: Create a stored procedure to show employee details when given an employee ID.

7. Create a View for High Salary Employees Concepts: Views
   Program: Create a view to display employees earning more than 70,000.

8. Cursor: Fetch and display employee details and department using an explicit cursor.

9. Trigger: Prevent deletion of employees who are assigned to active projects.

B. Sc. II(DS) (Semester-IV)

BDST24- 404: Advanced Excel

Theory- 30 Hours Credits:02

Objectives:

The main objectives of this course are:

1) To optimize the use of MS-Excel for powerful data analysis
2) To apply correct data visualization technique to gain optimal presentation of data
3) To apply enhanced features of MS-Excel

Unit 1 (15Hrs)

Workbook and worksheets-Navigation with keyboard, Tabs and ribbons, protecting excel workbook and worksheet,importingandexportingdata,co-authoring;DataandFormatting-AddingData, CutCopyPaste,Datafill,DataMovement,CellFormatting,ConditionalFormatting,Cell Operations , Reusable Lists , Data Validation , Sorting And Filtering ,Tables, Record macro freeze panes , cell references formulas, Audit formula, define Name, protect workbook and worksheet

Unit2 (15Hrs)

Understanding formulas operators in formula named ranges, calculations, functions in formula, relative and absolute address inreferencing cells outside the worksheet and workbook, functions - logical, summarizing, text , lookup, reference, data and time, math functions; error handling, formula auditing, Lookup functions (with H lookup and V lookup)
Charts types and uses, Chart depiction–column, line, pie, bar, bubble, histogram Analysis- Pivot Table, Pivot Charts, Data tools ,subtotal and what if analysis.

Reference Books:

i) Kore B.G-. MS-EXCEL for Data Analysis, Nirali Publication (2024)
ii) Paul McFedries -Microsoft Excel Formulas and functions ,
iii) Excel 2007 Bible – by John Walkenbach
iv) Handbook Microsoft Excel Till you understand – Department of Statistics

B. Sc. II(DS) (Semester-IV)

BDST24- 405: Continuous Probability Distributions

Theory- 30 Hours Credits:02

Objectives:

The main objectives of this course are:

i) Bivariate discrete distributions with real life situations.
ii) Continuous random variable and find the various measures, probabilities using its probability distribution.
iii) Some standard continuous probability distributions with real life situations.
iv) The relations among the different distributions.

Syllabus Contents

Unit-1 **15 hrs.**

Continuous Univariate Distribution:

Definition of the continuous sample space with illustration, Definition of continuous random variable (r.v.),probability density function (p.d.f.), cumulative distribution function (c.d.f.) and its properties.
Expectation of r.v. expectation of function of r.v., mean, median, mode, quartiles, variance, harmonic mean, raw and central moments, skewness and kurtosis, examples
Momentsgeneratingfunction(m.g.f.):definitionandproperties(i)Standardization property $M_x(0) = 1$, (ii) Effect of change of origin and scale , (iii) Uniqueness property of m.g.f., (statement only). Generation of raw and central moments.
Cumulativegeneratingfunction(c.g.f.):definition,relationbetweencumulantsand central moments (up to order four), Examples.

Unit-2 **15 hrs.**

Continuous Bivariate Distribution:

Definition of bivariate continuous random variable (X,Y), joint p.d.f., c.d.f. with properties , marginal and conditional distribution , independence of random variables , evaluation of probabilities of various regions bounded by straight lines. Expectation of function of r.v.s means, variances, covariance, correlation coefficient , conditional expectation , regression as conditional expectation if it is linear function of other variable and conditional variance, without proof of i) $E(X \pm Y) = E(X) \pm E(Y)$, ii)$E[E(X/Y) = E(X)$.
If X and Y are independent r.v.s. then (i) $E(XY)=E(X)E(Y)$, (ii) $M_{x+y}(t)= M_x(t)$ $M_y(t)$. Examples.

Continuous Uniform, Exponential and Normal Distribution:

Uniform distribution: Definition of Uniform distribution over (a,b), c.d.f.,

m.g.f., mean, variance, moments and real life examples.
Exponential distribution : p.d.f. (one parameter), c.d.f., m.g.f., c.g.f., mean, variance,c.v.,moments,Cumulants,Median,Quartiles,lackofmemoryproperty, real life examples.
Normal distribution : normal distribution with parameters μ & σ^2, Definition of standard normal distribution, properties of normal curve, m.g.f., c.g.f., mean, variance, median,
mode,meandeviation,moments,cumulants,measuresofskewness&kurtosis, real life examples.

Reference Books:

1. Parimal Mukhopadhyaya: An Introduction to the Theory of Probability. World Scientific Publishing.
2. Hogg R.V. and Criag A.T.: Introduction to Mathematical Statistics (Third edition), Macmillan Publishing, New York.
3. Gupta S.C.& Kapoor V. K.: Fundamentals of Mathematical Statistics. Sultan Chand & sons, New Delhi.
4. Gupta S.C. & Kapoor V. K.: Applied Statistics. Sultan Chand & sons, New Delhi.
5. Walpole R.E. & Mayer R. H.: Probability & Statistics.(Chapter4,5,6,8,10) MacMillan
6. Goon A.M., Gupta M. K. and Dasgupta B: Fundamentals of Statistics Vol. I and Vol. II World Press, Calcutta

BDSP24-406 : Practical - IV

Practical - 60 Hours Credits:02

Objectives:

The main objectives of this course are:

i) Know applications of some standard bivariate discrete probability distribution.

ii) Understand how to obtain random sample from various probability distributions.

iii) Know applications of some standard continuous probability distributions.

List of Practical's:

1. Data Management using Excel
2. Formatting Techniques, Macros, and cell reference
3. Working with Functions and lookup Operations
4. Data visualization and Analytics tools in excel

5. Workbook Sketching of Continuous Probability Distributions.
6. Fitting of Uniform Distributions
7. Fitting of Exponential Distributions
8. Fitting of Normal Distributions
9. Model Sampling of uniform and Exponential distributions.
10. Model Sampling of Normal Distributions.

B.Sc. Data Science Part-II Semester- IV
Open Elective Paper-IV
Course Code-OEDST24-401

Course Title: Computer Networking (30)

UNIT-1: Physical and Data link Layer (6)

Physical layer: -Digital-to-analog conversion: concept, Amplitude Shift Keying, Frequency Shift Keying, Analog-to-digital coversion:- Pulse Code Modulation (PCM),Delta Modulation (DM), Data link layer , Design issues, Framing, error detection and correction, Protocols:- Sliding window protocol: one bit sliding window protocol, protocol using go back, protocol using selective repeat.

UNIT-2: Network and Transport layer (8)

Network layer: - Design issue, Concept of routing, Routing algorithm (shortest path, Flooding, distance vector) , Congestion control algorithms (Leaking bucket, Token Bucket) , Transport layer: Services: Connection oriented and connection less services , Transport Layer Primitives: listen, connect, send, receive, disconnect , Protocols: TCP, UDP.

UNIT – III Session and Presentation layer (9)

Session layer: Services: dialog management, synchronization, activity management, exception handling Remote procedure calls
Presentation layer: Services: Translation, compression, encryption Cryptography: concept, symmetric key & asymmetric key cryptography

UNIT – IV Application layer and network security (7)

Application layer: Function Domain name system (DNS),Hypertext Transfer Protocol (HTTP),

Simple Mail Transfer Protocol (SMTP) ,Telnet, File Transfer Protocol (FTP) Network security:

Security concept and services Message Authentication, Digital Signatures, Entity authentication

B.Sc. Data Science Part-II Semester- IV

SEC Paper-II

Course Code-SECDSP24-401

Course Title: Web Technology

Credits: 02

Total Marks: 50

1 Create HTML page to add basic tags

2 Write an HTML code to illustrate the usage of the following: • Ordered List • Unordered List • Definition List.

3 Write HTML page to add image and 2 paragraph,and insert a table.

4 Use tag and Div tag and design page

5 Programs based on marquee, hyperlink, image maps

6 Program based on frame tags

7 Write an HTML code to demonstrate the usage of inline CSS. C3

8 Write an HTML code to demonstrate the usage of internal CSS.

9 Write an HTML code to demonstrate the usage of external CSS.

10 Programs based on CSS, cross browser testing

11 Programs based on creating forms, inputting values

12 Programs based on drop down and list box, text area, password

13 Program based on action buttons, radio, checkbox

B.Sc. Data Science Part-II Semester- IV

AEC Paper-II

Course Code-English

B.Sc. Data Science Part-II Semester- IV

CEP

Course Code

Chairman,
Board of Studies in Data Science