DEPARTMENT OF COMPUTER SCIENCE

BSc. (Hons.) Computer Science -DSC

Category I

DISCIPLINE SPECIFIC CORE COURSE – 4: Object Oriented Programming

CREDIT DISTRIBUTION, ELIGIBILITY AND PRE-REQUISITES OF THE COURSE

Course title & Code	Credits	Credit distribution of the course			Eligibility criteria	Pre-requisite of the course (if any)
		Lecture	Tutorial	Practical/ Practice		
DSC04 Object Oriented Programming with C++	4	3	0	1	Class XII pass with Mathematics	Nil

Learning Objectives

This course is designed to introduce programming concepts using C++ to students. The course aims to develop structured as well as object-oriented programming skills using C++ programming language. The course also aims to achieve competence amongst its students to develop correct and efficient C++ programs to solve problems spanning multiple domains.

Learning outcomes

On successful completion of the course, students will be able to:

- Write simple programs using built-in data types of C++.
- Implement arrays and user defined functions in C++.
- Write programs using dynamic memory allocation, handling external files, interrupts and exceptions.
- Solve problems spanning multiple domains using suitable programming constructs in C++.
- Solve problems spanning multiple domains using the concepts of object oriented programming in C++.

SYLLABUS OF DSC-4

UNIT – I (3 Hours)

Introduction to C++: Overview of Procedural and Object-Oriented Programming, Using main() function, Header Files, Compiling and Executing Simple Programs in C++

UNIT – II (12 Hours)

Programming Fundamentals: Data types, Variables, Operators, Expressions, Arrays, Keywords, Decision making constructs, Iteration, Type Casting, Input-output statements, Functions, Command Line Arguments/Parameters

UNIT – III (15 Hours)

Object Oriented Programming: Concepts of Abstraction, Encapsulation. Creating Classes and objects, Modifiers and Access Control, Constructors, Destructors, Implementation of Inheritance and Polymorphism, Template functions and classes

UNIT – IV (9 Hours)

Pointers and References: Static and dynamic memory allocation, Pointer and Reference Variables, Implementing Runtime polymorphism using pointers and references

UNIT – V (6 Hours)

Exception and File Handling: Using try, catch, throw, throws and finally; Nested try, creating user defined exceptions, File I/O Basics, File Operations

Practical component (if any) -30 Hours

1. Write a program to compute the sum of the first n terms of the following series:

   $sum = 1 - \frac{1}{2^2} + \frac{1}{3^3} -$ . . . .

   The number of terms n is to be taken from the user through the command line. If the command line argument is not found then prompt the user to enter the value of n.
2. Write a program to remove the duplicates from an array.
3. Write a program that prints a table indicating the number of occurrences of each alphabet in the text entered as command line arguments.
4. Write a menu driven program to perform string manipulation (without using inbuilt string functions):
   a. Show address of each character in string
   b. Concatenate two strings.
   c. Compare two strings
   d. Calculate length of the string (use pointers)
   e. Convert all lowercase characters to uppercase
   f. Reverse the string
   g. Insert a string in another string at a user specified position
5. Write a program to merge two ordered arrays to get a single ordered array.

6. Write a program to search a given element in a set of N numbers using Binary search

   (i) with recursion (ii) without recursion.

7. Write a program to calculate GCD of two numbers (i) with recursion (ii) without recursion.

8. Create a Matrix class. Write a menu-driven program to perform following Matrix operations (exceptions should be thrown by the functions if matrices passed to them are incompatible and handled by the main() function):

   a. Sum

   b. Product

   c. Transpose

9. Define a class Person having name as a data member. Inherit two classes Student and Employee from Person. Student has additional attributes as course, marks and year and Employee has department and salary. Write display() method in all the three classes to display the corresponding attributes. Provide the necessary methods to show runtime polymorphism.

10. Create a Triangle class. Add exception handling statements to ensure the following conditions: all sides are greater than 0 and sum of any two sides are greater than the third side. The class should also have overloaded functions for calculating the area of a right angled triangle as well as using Heron's formula to calculate the area of any type of triangle.

11. Create a class Student containing fields for Roll No., Name, Class, Year and Total Marks. Write a program to store 5 objects of Student class in a file. Retrieve these records from the file and display them.

12. Copy the contents of one text file to another file, after removing all whitespaces.

Essential/recommended readings

1. Stephen Prata, *C++ Primer Plus*, 6th Edition, Pearson India, 2015.
2. E Balaguruswamy, *Object Oriented Programming with C++*, 8th edition, McGraw-Hill Education, 2020.
3. D.S. Malik, *C++ Programming: From Problem Analysis to Program Design*, 6th edition, Cengage Learning, 2013.

Suggestive readings

(i) Schildt, H. *C++: The Complete Reference*, 4th edition, McGraw Hill, 2003

(ii) Forouzan, A. B., Gilberg, R. F. *Computer Science: A Structured Approach using C++,* 2nd edition, Cengage Learning, 2010

Note: Examination scheme and mode shall be as prescribed by the Examination Branch, University of Delhi, from time to time.

DISCIPLINE SPECIFIC CORE COURSE – 5: Discrete Mathematical Structures

Credit distribution, Eligibility and Pre-requisites of the Course

Course title & Code	Credits	Credit distribution of the course			Eligibility criteria	Pre-requisite of the course (if any)
		Lecture	Tutorial	Practical/ Practice		
DSC 05 Discrete Mathematical Structures	**4**	**3**	**0**	**1**	**Class XII pass with Mathematics**	**Nil**

Learning Objectives

This course is designed as a foundational course to make students learn about the mathematical constructs that are used in Computer Science such as Boolean algebra, sets, relations, functions, principles of counting, and recurrences. In this course, the knowledge of mathematical notation, ideas and concepts learnt at the pre-college levels is extended to orient the students towards mathematical thinking required in Computer Science.

Learning outcomes

On successful completion of the course, students will be able to:

- Relate mathematical concepts and terminology to examples in the domain of Computer Science.
- Model real world problems using various mathematical constructs.
- Use different proofing techniques; construct simple mathematical proofs using logical arguments.
- Formulate mathematical claims and construct counterexamples.

SYLLABUS OF DSC- 5
UNIT – I (06 Hours)
Sets, Functions, Sequences and Summations, Relations: Sets: Set Operations, Computer Representation of Sets, Countable and Uncountable Set, Principle of Inclusion and Exclusion, Multisets; Functions: One-to-one and Onto Functions, Inverse Functions and Compositions of

Functions, Graphs of Functions Sequences and Summations: Sequences, Special Integer Sequences, Summations; Relations: Properties of Binary Relations, Equivalence relations and Partitions, Partial Ordering Relations and Lattices.

UNIT – II (09 Hours)

Logic and Proofs: Propositional Logic, Propositional Equivalences, Use of first-order logic to express natural language predicates, Quantifiers, Nested Quantifiers, Rules of Inference, Introduction to Proofs, Proof Methods and Strategies, Mathematical Induction.

UNIT – III (09 Hours)

Number Theory: Division and Integers, Primes and Greatest Common Divisors, Representation of Integers, Algorithms for Integer Operations, Modular Exponentiation, Applications of Number Theory.

UNIT – IV (06 Hours)

Combinatorics/Counting: The Pigeonhole Principle, Permutations and Combinations, Binomial Coefficients, Generalized Permutations and Combinations, Generating Permutations and Combinations.

UNIT – V (09 Hours)

Graphs and Trees: Graphs: Basic Terminology, Multigraphs and Weighted Graphs, Paths and Circuits, Eulerian Paths and Circuits, Hamiltonian paths and Circuits, Shortest Paths, Spanning Trees, Graph Isomorphism, Planar Graphs; Trees: Trees, Rooted Trees, Path Lengths in Rooted Trees.

UNIT – VI (06 Hours)

Recurrence: Recurrence Relations, Generating Functions, Linear Recurrence Relations with Constant Coefficients and their solution.

Practical component (if any) – 30 Hours

1. Create a class SET. Create member functions to perform the following SET operations:
    1) is member: check whether an element belongs to the set or not and return value as true/false.
    2) powerset: list all the elements of the power set of a set .
    3) subset: Check whether one set is a subset of the other or not.
    4) union and Intersection of two Sets.
    5) complement: Assume Universal Set as per the input elements from the user.
    6) set Difference and Symmetric Difference between two sets.
    7) cartesian Product of Sets.

    Write a menu driven program to perform the above functions on an instance of the SET class.

2. Create a class RELATION, use Matrix notation to represent a relation. Include member functions to check if the relation is Reflexive, Symmetric, Anti-symmetric, Transitive. Using these functions check whether the given relation is: Equivalence or Partial Order relation or None

3. Write a Program that generates all the permutations of a given set of digits, with or without repetition.
4. For any number n, write a program to list all the solutions of the equation $x_1 + x_2 + x_3 + ...+ x_n = C$, where C is a constant (C<=10) and $x_1, x_2, x_3, ..., x_n$ are nonnegative integers, using brute force strategy.
5. Write a Program to evaluate a polynomial function. (For example store $f(x) = 4n^2 + 2n + 9$ in an array and for a given value of n, say n = 5, compute the value of f(n)).
6. Write a Program to check if a given graph is a complete graph. Represent the graph using the Adjacency Matrix representation.
7. Write a Program to check if a given graph is a complete graph. Represent the graph using the Adjacency List representation.
8. Write a Program to accept a directed graph G and compute the in-degree and out-degree of each vertex.

Essential/recommended readings

1. Liu, C. L., Mohapatra, D. P. *Elements of Discrete Mathematics: A Computer Oriented Approach*, 4th edition, Tata McGraw Hill, 2017.
2. Rosen, K. H.. *Discrete Mathematics and Its Applications*, 8th edition, McGraw Hill, 2018.

Suggestive readings

(i) Cormen, T. H., Leiserson, C. E., Rivest, R. L., Stein C. *Introduction to Algorithms*, 4th edition, Prentice Hall of India. 2022.
(ii) Trembley, J. P., Manohar, R. *Discrete Mathematical Structures with Application to Computer Science*, Tata McGraw Hill, 1997.
(iii) Albertson, M. O. and Hutchinson, J. P. *Discrete Mathematics with Algorithms*, John Wiley and Sons, 1988.

DISCIPLINE SPECIFIC CORE COURSE – 6: Probability for Computing

Credit distribution, Eligibility and Pre-requisites of the Course

Course title & Code	Credits	Credit distribution of the course			Eligibility criteria	Pre-requisite of the course (if any)
		Lecture	Tutorial	Practical/ Practice		
DSC06 Probability for computing	**4**	**3**	**0**	**1**	Class XII pass with Mathematics	**Nil**

Learning Objectives

This course introduces the students to the fundamental concepts and topics of probability and statistics, whose knowledge is important in other computer science courses. The course aims to build the foundation for some of the core courses in later semesters.

Learning outcomes

After successful completion of this course, the student will be able to:

- Use probability theory to evaluate the probability of real-world events.
- Describe discrete and continuous probability distribution functions and generate random numbers from the given distributions.
- Find the distance between two probability distributions
- Define and quantify the information contained in the data.
- Perform data analysis in a probabilistic framework.
- Visualize and model the given problem using mathematical concepts covered in the course.

SYLLABUS OF DSC-6 UNIT-I (09 Hours)

Basic Probability: Introduction to the notion of probability, Random experiment, Sample space and Events, Probability defined on events, Algebra of events. Conditional probabilities, independent events, Bayes' theorem.

UNIT-II (12 Hours)

Random Variables: Introduction to Random Variables, Probability mass/density functions, Cumulative distribution functions. Discrete Random Variables (Bernoulli, Binomial, Poisson, Multinomial and Geometric). Continuous Random Variables (Uniform, Exponential and Normal). Expectation of a Random Variable, Expectation of Function of a Random Variable and Variance. Markov inequality, Chebyshev's inequality, Central Limit Theorem, Weak and Strong Laws of Large Numbers.

UNIT-III (09 Hours)

Joint Distributions: Jointly distributed Random Variables, Joint distribution functions, Independent Random Variables, Covariance of Random Variables, Correlation Coefficients,

Conditional Expectation.

UNIT-IV (15 Hours)

Markov Chain and Information Theory: Introduction to Stochastic Processes, Chapman–Kolmogorov equations, Classification of states, Limiting and Stationary Probabilities. Random Number Generation, Pseudo Random Numbers, Inverse Transformation Method, Rejection Method, Uncertainty, Information and Entropy, Mutual Information, KL Divergence.

Practical component (if any) – 30 Hours

The goal of this lab is to develop data interpretation skills. Following exercises are designed to enable students to understand data characteristics either by visualization or by interpreting computed measures. All the exercises are to be completed using MS Excel functions and graphs. At the end of each exercise, the student should be able to draw a conclusion and state in a concise manner. Teachers are expected to guide students to obtain real data available through the internet for the following exercises.

1. Plotting and fitting of Binomial distribution and graphical representation of probabilities.
2. Plotting and fitting of Multinomial distribution and graphical representation of probabilities.
3. Plotting and fitting of Poisson distribution and graphical representation of probabilities.
4. Plotting and fitting of Geometric distribution and graphical representation of probabilities.
5. Plotting and fitting of Uniform distribution and graphical representation of probabilities.
6. Plotting and fitting of Exponential distribution and graphical representation of probabilities.
7. Plotting and fitting of Normal distribution and graphical representation of probabilities.
8. Calculation of cumulative distribution functions for Exponential and Normal distribution.
9. Given data from two distributions, find the distance between the distributions.
10. Application problems based on the Binomial distribution.
11. Application problems based on the Poisson distribution.
12. Application problems based on the Normal distribution.
13. Presentation of bivariate data through scatter-plot diagrams and calculations of covariance.
14. Calculation of Karl Pearson's correlation coefficients.
15. To find the correlation coefficient for a bivariate frequency distribution.
16. Generating Random numbers from discrete (Bernoulli, Binomial, Poisson) distributions.

17. Generating Random numbers from continuous (Uniform, Normal) distributions.
18. Find the entropy from the given data set.

Essential/recommended readings

1. Ross **S**heldon M. *Introduction to Probability Models*, 12th Edition, Elsevier, 2019.
2. Trivedi, K. S. *Probability and Statistics with Reliability, Queuing and Computer Science Applications*, 2nd edition, Wiley, 2015.
3. Deisenroth, Marc Peter, Faisal A. Aldo and Ong Cheng Soon, *Mathematics for Machine Learning*, 1st edition, Cambridge University Press, 2020.
4. Ian F. Blake, *An Introduction to Applied Probability*, John Wiley.

Suggestive readings

(i) Johnson James L., *Probability and Statistics for Computer Science*, 6th edition, Wiley, 2004.

(ii) Forsyth David, *Probability and Statistics for Computer Science*, 1st edition, Springer, 2019.

(iii) Freund J.E., *Mathematical Statistics with Applications*, 8th edition, Pearson Education, 2013.

(iv) Devore Jay L., *Probability and Statistics for Engineering and the Sciences*, 9th edition, Cengage Learning, 2020.

BSc. (Physical Sciences/ Mathematical Sciences) with Computer Science as one of the Core Disciplines

Category II

DISCIPLINE SPECIFIC CORE COURSE (DSC-2): Data Structures using C++

CREDIT DISTRIBUTION, ELIGIBILITY AND PRE-REQUISITES OF THE COURSE

Course title & Code	Credits	Credit distribution of the course			Eligibility criteria	Pre-requisite of the course (if any)
		Lecture	Tutorial	Practical/ Practice		
DSC02: Data Structures using C++	**4**	**3**	**0**	**1**	**Class XII pass with Mathematics**	Nil

Learning Objectives

The course aims at developing the ability to use basic data structures like arrays, stacks, queues, lists, trees to solve problems. C++ is chosen as the language to understand implementation of these data structures.

Learning outcomes

On successful completion of the course, students will be able to:

- Compare two functions for their rates of growth.
- Understand abstract specification of data-structures and their implementation.
- Compute time and space complexity of operations on a data-structure.
- Identify the appropriate data structure(s) for a given application and understand the trade-offs involved in terms of time and space complexity.
- Apply recursive techniques to solve problems.

SYLLABUS OF DSC-2 UNIT – I (06 Hours)

Growth of Functions, Recurrence Relations. Functions used in analysis, asymptotic notations, asymptotic analysis, solving recurrences using recursion tree, Master Theorem.

UNIT – II (12 Hours)

Arrays, Linked Lists, Stacks, Queues, Deques. Arrays: array operations, applications, sorting, two-dimensional arrays, dynamic allocation of arrays; Linked Lists: singly linked lists, doubly linked lists, circularly linked lists, Stacks: stack as an ADT, implementing stacks using arrays, implementing stacks using linked lists, applications of stacks; Queues:

queue as an ADT, implementing queues using arrays, implementing queues using linked lists, double-ended queue as an ADT. Time complexity analysis of operations on all data structures.

UNIT – III (06 Hours)

Sorting: Insertion Sort, Count Sort and their complexity analysis.

UNIT – IV (03 Hours)

Recursion: Recursive functions, linear recursion, binary recursion.

UNIT – V (06 Hours)

Trees, Binary Trees. Trees: definition and properties, binary trees: definition and properties, traversal of binary trees and their time complexity analysis.

UNIT – VI (09 Hours)

Binary Search Trees, Balanced Search Trees: Binary Search Trees: insert, delete (by copying), search operations, time complexity analysis of these operations; Balanced Search Trees and (2,4) Trees: motivation and introduction.

UNIT – VII (03 Hours)

Binary Heap, Priority Queue: Binary Heaps: motivation and introduction, application of heaps - Priority Queues.

Practical component (if any) – 30 Hours

1. Perform matrix addition and multiplication.
2. Implement following recursive functions:
    a. Factorial of a number
    b. N^{th} fibonacci number
    c. Power function: x^y
3. Implement singly linked lists.
4. Implement doubly linked lists.
5. Implement circular linked lists.
6. Implement stack data structure and its operations using arrays.
7. Implement stack data structure and its operations using linked lists.
8. Convert Prefix expression to Infix and Postfix expressions, and evaluate.
9. Implement queue data structure and its operations using arrays.
10. Implement queue data structure and its operations using linked lists.
11. Implement Binary Trees and its traversals.

Essential/recommended readings

1. Goodrich, M., Tamassia, R., & Mount, D., *Data Structures and Algorithms Analysis in C++,* 2nd edition. Wiley, 2011.
2. Cormen, T.H., Leiserson, C.E., Rivest, R. L., Stein C., *Introduction to Algorithms*, 3rd edition, Prentice Hall of India, 2010.
3. Drozdek, A., *Data Structures and Algorithms in C++*, 4th edition, Cengage Learning, 2012.

Suggestive readings

(i) Sahni, S. *Data Structures, Algorithms and applications in C++*. 2nd Edition. Universities Press, 2011.

(ii)Tanenbaum, A. M., Augenstein, M. J., & Langsam Y., *Data Structures Using C and C++*. 2nd edition. Prentice Hall of India, 2009.

Note: Examination scheme and mode shall be as prescribed by the Examination Branch, University of Delhi, from time to time

BA (Prog.) with Computer Science as Major

Category III

DISCIPLINE SPECIFIC CORE COURSE (DSC-2): Data Structures

CREDIT DISTRIBUTION, ELIGIBILITY AND PRE-REQUISITES OF THE COURSE

Course title & Code	Credits	Credit distribution of the course			Eligibility criteria	Pre-requisite of the course (if any)
		Lecture	Tutorial	Practical/ Practice		
DSC02: Data Structures	**4**	**3**	**0**	**1**	**Class XII pass with Mathematics**	Nil

Learning Objectives

The course aims at developing the ability to define, differentiate, implement the basic data structures like arrays, stacks, queues, lists, trees and use them to solve problems. C++ is chosen as the language to understand implementation of these data structures.

Learning outcomes

On successful completion of the course, students will be able to:

- Understand abstract specification of data-structures.
- Implement data structures as ADT..
- Identify the appropriate data structure(s) for a given application.
- Apply recursive techniques to solve problems.

SYLLABUS OF DSC-2

UNIT – I (15 Hours)

Arrays, Linked Lists, Stacks, Queues, Deques: Arrays: array operations, applications, sorting, two-dimensional arrays, dynamic allocation of arrays; Linked Lists: singly linked lists, doubly

linked lists, circularly linked lists, Stacks: stack as an ADT, implementing stacks using arrays, implementing stacks using linked lists, applications of stacks; Queues: queue as an ADT, implementing queues using arrays, implementing queues using linked lists, double-ended queue as an ADT.

UNIT – II (06 Hours)

Searching and Sorting: Linear Search, Binary Search, Insertion Sort, Count Sort.

UNIT – III (09 Hours)

Recursion: Recursive functions, linear recursion, binary recursion.

UNIT – IV (06 Hours)

Trees, Binary Trees: Trees: definition and properties, binary trees: definition and properties, traversal of binary trees.

UNIT – V(09 Hours)

Binary Search Trees: insert, delete (by copying), search operations.

Practical component (if any) – 30 Hours

1. Perform matrix addition and multiplication.
2. Implement following recursive functions:

   Factorial of a number

   N^{th} fibonacci number

   Power function: x^y
3. Implement singly linked lists.
4. Implement doubly linked lists.
5. Implement circular linked lists.
6. Implement stack data structure and its operations using arrays.
7. Implement stack data structure and its operations using linked lists.
8. Convert Prefix expression to Infix and Postfix expressions, and evaluate.
9. Implement queue data structure and its operations using arrays.
10. Implement queue data structure and its operations using linked lists.
11. Implement Binary Trees and its traversals.

Essential/recommended readings

1. Goodrich, M.T., Tamassia, R., & Mount, D. *Data Structures and Algorithms Analysis in C++,* 2nd edition, Wiley, 2011.
2. Cormen, T.H., Leiserson, C.E., Rivest, R. L. Stein C. *Introduction to Algorithms*, 4th edition, Prentice Hall of India, 2022.
3. Drozdek, A. *Data Structures and Algorithms in C++*, 4th edition, Cengage Learning, 2012.

Suggestive readings

(i) Sahni, S., *Data Structures, Algorithms and applications in C++*, 2nd edition, Universities Press, 2011.

(ii) Langsam Y., Augenstein, M. J., & Tanenbaum, A. M. *Data Structures Using C and C++*, Pearson, 2009.

Note: Examination scheme and mode shall be as prescribed by the Examination Branch, University of Delhi, from time to time.

DISCIPLINE SPECIFIC CORE COURSE – A2 : DATA INTERPRETATION AND VISUALIZATION USING PYTHON

CREDIT DISTRIBUTION, ELIGIBILITY AND PRE-REQUISITES OF THE COURSE

Course title & Code	Credits	Credit distribution of the course			Eligibility criteria	Pre-requisite of the course (if any)
		Lecture	Tutorial	Practical/ Practice		
A2: Data Interpretation and Visualization using Python	**4**	**3**	**0**	**1**	**Class XII pass with Mathematics**	knowledge of Python

Learning Objectives

This course is designed to introduce the students to the application of Python to get a deterministic view of data and interpret results..

Learning outcomes

On successful completion of the course, students will be able to:

- Interpret Data
- Obtain a deterministic view of data
- Perform data handling using Numpy arrays
- Load, clean, transform, merge and reshape data using Pandas
- Visualize data using Pandas and matplot libraries

SYLLABUS OF A2

UNIT – I (06 Hours)

Introduction to basic statistics and analysis: Fundamentals of Data Analysis, Statistical foundations for Data Analysis, Types of data, Descriptive Statistics, Correlation and covariance, Linear Regression, Statistical Hypothesis Generation and Testing, Python Libraries: NumPy, Pandas, Matplotlib

UNIT – II (09 Hours)

Array manipulation using Numpy: Numpy array: Creating Numpy arrays; various data types of Numpy arrays, indexing and slicing, swapping axes, transposing arrays, data processing using Numpy arrays

UNIT – III (12 Hours)

Data Manipulation using Pandas: Data Structures in Pandas: Series, DataFrame, Index objects, Loading data into Pandas data frame, Working with Data Frames: Arithmetics, Statistics, Binning, Indexing, Reindexing, Filtering, Handling missing data, Hierarchical indexing, Data wrangling: Data cleaning, transforming, merging and reshaping

UNIT – IV (12 Hours)

Plotting and Visualization: Using Matplotlib to plot data: figures, subplots, markings, color and line styles, labels and legends, plotting functions in Pandas: Line, bar, Scatter plots, histograms, stacked bars, Heatmap

UNIT-V (06 Hours)

Data Aggregation and Group operations: Group by Mechanics, Data aggregation, General split-apply-combine, Pivot tables and cross tabulation.

Practical component (if any) – 30 Hours

Use a dataset of your choice from Open Data Portal (https:// data.gov.in/, UCI repository) or load from scikit, seaborn library for the following exercises to practice the concepts learnt.

1. Load a Pandas dataframe with a selected dataset. Identify and count the missing values in a dataframe. Clean the data after removing noise as follows
   a. Drop duplicate rows.
   b. Detect the outliers and remove the rows having outliers
   c. Identify the most correlated positively correlated attributes and negatively correlated attributes
2. Import iris data using sklearn library or (Download IRIS data from: https://archive.ics.uci.edu/ml/datasets/iris or import it from sklearn.datasets)

i. Compute mean, mode, median, standard deviation, confidence interval and standard error for each feature
ii. Compute correlation coefficients between each pair of features and plot heatmap
iii. Find covariance between length of sepal and petal
iv. Build contingency table for class feature

3. Load Titanic data from sklearn library, plot the following with proper legend and axis labels:

a. Plot bar chart to show the frequency of survivors and non-survivors for male and female passengers separately
b. Draw a scatter plot for any two selected features
c. Compare density distribution for features age and passenger fare
d. Use a pair plot to show pairwise bivariate distribution

4. Using Titanic dataset, do the following
a. Find total number of passengers with age less than 30
b. Find total fare paid by passengers of first class
c. Compare number of survivors of each passenger class

5. Download any dataset and do the following
a. Count number of categorical and numeric features
b. Remove one correlated attribute (if any)
c. Display five-number summary of each attribute and show it visually

Essential/recommended readings

1. McKinney W. *Python for Data Analysis: Data Wrangling with Pandas, NumPy and IPython*, 2nd edition, O'Reilly Media, 2018.
2. Molin S. *Hands-On Data Analysis with Pandas*, Packt Publishing, 2019.
3. Gupta S.C., Kapoor V.K. *Fundamentals of Mathematical Statistics*,12th edition, Sultan Chand & Sons, 2020.

Suggestive readings

(i) Chen D. Y. *Pandas for Everyone: Python Data Analysis*, 1st edition, Pearson Education, 2018.

(ii) Miller J.D. *Statistics for Data Science,* Packt Publishing Limited, 2017.

Note: Examination scheme and mode shall be as prescribed by the Examination Branch, University of Delhi, from time to time.

BA (Prog.) with Computer Science as Non-Major

Category III

DISCIPLINE SPECIFIC CORE COURSE (DSC-2): Data Structures

CREDIT DISTRIBUTION, ELIGIBILITY AND PRE-REQUISITES OF THE COURSE

Course title & Code	Credits	Credit distribution of the course			Eligibility criteria	Pre-requisite of the course (if any)
		Lecture	Tutorial	Practical/ Practice		
DSC02: Data Structures	**4**	**3**	**0**	**1**	**Class XII pass with Mathematics**	Knowledge of C++

Learning Objectives

The course aims at developing the ability to define, differentiate, implement the basic data structures like arrays, stacks, queues, lists, trees and use them to solve problems. C++ is chosen as the language to understand implementation of these data structures.

Learning outcomes

On successful completion of the course, students will be able to:

- Understand abstract specification of data-structures.
- Implement data structures as ADT..
- Identify the appropriate data structure(s) for a given application.
- Apply recursive techniques to solve problems.

SYLLABUS OF DSC-2

UNIT – I (15 Hours)

Arrays, Linked Lists, Stacks, Queues, Deques: Arrays: array operations, applications, sorting, two-dimensional arrays, dynamic allocation of arrays; Linked Lists: singly linked lists, doubly linked lists, circularly linked lists, Stacks: stack as an ADT, implementing stacks using arrays, implementing stacks using linked lists, applications of stacks; Queues: queue as an ADT, implementing queues using arrays, implementing queues using linked lists, double-ended queue as an ADT.

UNIT – II (06 Hours)

Searching and Sorting: Linear Search, Binary Search, Insertion Sort, Count Sort.

UNIT – III (09 Hours)

Recursion: Recursive functions, linear recursion, binary recursion.

UNIT – IV (06 Hours)

Trees, Binary Trees: Trees: definition and properties, binary trees: definition and properties, traversal of binary trees.

UNIT – V (09 Hours)

Binary Search Trees: insert, delete (by copying), search operations.

Practical component (if any) – 30 Hours

1. Perform matrix addition and multiplication.
2. Implement following recursive functions:
    i. Factorial of a number
    ii. N^{th} fibonacci number
    iii. Power function: x^y
3. Implement singly linked lists.
4. Implement doubly linked lists.
5. Implement circular linked lists.
6. Implement stack data structure and its operations using arrays.
7. Implement stack data structure and its operations using linked lists.
8. Convert Prefix expression to Infix and Postfix expressions, and evaluate.
9. Implement queue data structure and its operations using arrays.
10. Implement queue data structure and its operations using linked lists.
11. Implement Binary Trees and its traversals.

Essential/recommended readings

1. Goodrich, M.T., Tamassia, R., & Mount, D. *Data Structures and Algorithms Analysis in C++,* 2nd edition, Wiley, 2011.
2. Cormen, T.H., Leiserson, C.E., Rivest, R. L. Stein C. *Introduction to Algorithms*, 4th edition, Prentice Hall of India, 2022.
3. Drozdek, A. *Data Structures and Algorithms in C++*, 4th edition, Cengage Learning, 2012.

Suggestive readings

(i) Sahni, S., *Data Structures, Algorithms and applications in C++*, 2nd edition, Universities Press, 2011.

(ii) Langsam Y., Augenstein, M. J., & Tanenbaum, A. M. *Data Structures Using C and C++*, Pearson, 2009.

Note: Examination scheme and mode shall be as prescribed by the Examination Branch, University of Delhi, from time to time.

Common Pool of Generic Electives (GE) Courses
Offered by Department of Computer Sciences
Category-IV

GENERIC ELECTIVES (GE-2a): Data Analysis and Visualization

Credit distribution, Eligibility and Pre-requisites of the Course

Course title & Code	Credits	Credit distribution of the course			Eligibility criteria	Pre- requisite of the course
		Lecture	Tutorial	Practical/ Practice		
GE2a Data Analysis and Visualization using Python	**4**	**3**	**0**	**1**	**Class XII pass with Mathematics**	knowledge of Python

Learning Objectives

This course is designed to introduce the students to real-world data analysis problems, their analysis and interpretation of results in the field of exploratory data science using Python.

Learning outcomes

On successful completion of the course, students will be able to:

- Apply descriptive statistics to obtain a deterministic view of data
- Apply basic and advanced level statistical function on data
- Perform data handling using Numpy arrays
- Do data cleaning and transformation before extracting useful information
- Visualize data for ease of understanding the revealed information

SYLLABUS OF GE-2a

UNIT – I & II (09 Hours)

Introduction to basic statistics and analysis: Fundamentals of Data Analysis, Statistical foundations for Data Analysis, Types of data, Descriptive Statistics, Python Libraries: NumPy, Pandas, Matplotlib

Array manipulation using NumPy: NumPy array: Creating NumPy arrays, various data types of NumPy arrays

UNIT – I & II (09 Hours)

Introduction to basic statistics and analysis: contd..

Correlation and covariance, Linear Regression, Statistical Hypothesis Generation and Testing

Unit 2 Array manipulation using Numpy: contd..

Indexing and slicing, swapping axes, transposing arrays, data processing using Numpy arrays

UNIT – III (15 Hours)

Data Manipulation using Pandas: Data Structures in Pandas: Series, Data Frame, Index objects, loading data into Panda's data frame, Working with Data Frames: Arithmetics, Statistics, Binning, Indexing, Reindexing, Filtering, Handling missing data, Hierarchical indexing, Data wrangling: Data cleaning, transforming, merging and reshaping

UNIT – IV (12 Hours)

Plotting and Visualization: Using Matplotlib to plot data: figures, subplots, markings, color and line styles, labels and legends, Plotting functions in Pandas: Lines, bar, Scatter plots, histograms, stacked bars, Heatmap

Practical component (if any) – 30 Hours

Use data set of your choice from Open Data Portal (https:// data.gov.in/, UCI repository) or load from scikit, seaborn library for the following exercises to practice the concepts learnt.

1. Load a Pandas data frame with a selected dataset. Identify and count the missing values in a data frame. Clean the data after removing noise as follows
   a. Drop duplicate rows.
   b. Detect the outliers and remove the rows having outliers
   c. Identify the most correlated positively correlated attributes and negatively correlated attributes
2. Import iris data using sklearn library or (Download IRIS data from: https://archive.ics.uci.edu/ml/datasets/iris or import it from sklearn.datasets)
   a. Compute mean, mode, median, standard deviation, confidence interval and standard error for each feature
   b. Compute correlation coefficients between each pair of features and plot heatmap
   c. Find covariance between length of sepal and petal
   d. Build contingency table for class feature
3. Load Titanic data from sklearn library , plot the following with proper legend and axis labels:
   a. Plot bar chart to show the frequency of survivors and non-survivors for male and female passengers separately
   b. Draw a scatter plot for any two selected features
   c. Compare density distribution for features age and passenger fare

d. Use a pair plot to show pairwise bivariate distribution

4. Using Titanic dataset, do the following
   a. Find total number of passengers with age less than 30
   b. Find total fare paid by passengers of first class
   c. Compare number of survivors of each passenger class

Project students are encouraged to work on a good dataset in consultation with their faculty and apply the concepts learned in the course.

Essential/recommended readings

1. McKinney W. *Python for Data Analysis: Data Wrangling with Pandas*, *NumPy and IPython*. 2nd edition, O'Reilly Media, 2018.
2. Molin S. *Hands-On Data Analysis with Pandas*, Packt Publishing, 2019.
3. Gupta S.C., Kapoor V.K., *Fundamentals of Mathematical Statistics*, Sultan Chand & Sons, 2020.

Suggestive readings

(i) Chen D. Y, *Pandas for Everyone: Python Data Analysis*, Pearson, 2018.

(ii) Miller J.D. *Statistics for Data Science,* Packt Publishing, 2017.

GENERIC ELECTIVES (GE-2b): Data Analysis and Visualization using Spreadsheet

Credit distribution, Eligibility and Pre-requisites of the Course

Course title & Code	Credits	Credit distribution of the course			Eligibility criteria	Pre-requisite of the course
		Lecture	Tutorial	Practical/ Practice		
GE2b **Data Analysis and Visualization using Spreadsheet**	**4**	**3**	**0**	**1**	**Class XII pass with Mathematics**	**Nil**

Learning Objectives

This course is designed to inculcate statistical thinking about data to the students who have studied Mathematics up to Class 10^{th} ONLY. It gently introduces the students to basic statistics, and builds gradually to cover advanced functionalities for data analysis in spreadsheets. The objective is to enhance the knowledge of statistics and enable students to make sense of data by analyzing and visualizing it using spreadsheets, interpreting the results and gaining insights.

Learning outcomes

On successful completion of the course, students will be able to:

- Analyze and visualize data using spreadsheets
- Apply basic and advanced level statistical functions in spreadsheets
- Gain practical, hands-on experience of data analysis using spreadsheet

SYLLABUS OF GE-2b

UNIT – I (09 Hours)

Introduction to Basic Statistics

Fundamentals of Data Analysis, Statistical foundations for Data Analysis, Types of data, Descriptive Statistics, Correlation and covariance, Linear Regression.

UNIT – II (15 Hours)

Data Handling

Spreadsheet concepts, managing worksheets, formatting cells, entering data, Handling operators in formula, Cell referencing and naming of cells and cell ranges, Sorting, Multilayer sorting, Data validation, Find and Replace, Paste special, Filter and advanced filter, Formatting as table, Pivot tables, Formulae vs functions, Cell formulae vs Array formulae.

Mathematical functions, Statistical functions, Logical functions, Date and Time functions, Lookup and reference: Hlookup, and Vlookup, Index and Match functions, Text functions.

What–if-analysis: Goal-seek, Data tables, Scenario manager.

UNIT – III (12 Hours)

Data Analysis

Explore a data model: its content, and its structure, using the Power Pivot add-in. Learning DAX formula language. Create calculated fields and calculated measure for each cell, filter context for calculation, and explore several advanced DAX functions.

Cube formulas to retrieve data from data model.

UNIT – IV (09 Hours)

Data Visualization

Different types of charts including Pivot charts: Column, Line, Pie, Bar, Scatter charts. Fine tuning of charts: Chart Elements, Chart Styles, Chart Filters, Box Plot.

Practical component (if any) – 30 Hours

1. In a meeting of a marketing department of an organization it has been decided that price of selling an item is fixed at Rs. 40. It was resolved to increases the selling of more items and getting the profit of Rs. 50000/-. Use Goal Seek of find out how many items you will have to sell to meet your profit figure.
2. Create worksheet related to crop production of various crops in Indian stats in last five years (wheat, rice, pulses, soya-bean, and cane-sugar etc).
    i) Make a bar chart
    ii) Make a pie chart
    iii) Make a box plot
3. Study and perform the various DAX functions to analyse the data.
4. Create workbook related to sales of Business Company having various product in last four quarters for 10 sales persons.
    i) Make a line graph to show the growth/decline in the sales
    ii) Show the graph of each sales person sales
    iii) Find the two sales persons done in last 2 quarters
    iv) Find the sales persons consistent in last four quarters
    v) Find the most popular product of the company and the current popular product of company.
5. Create a Pivot-table showing the Customer Names who placed orders with GSS during 2019-2022. For each customer, also show the total number of orders, Total Sales, and Total Profit. Add a Slicer or a Filter that can be used to show the information specifically for each Customer Segment. Use information from the Pivot- table to answer the following questions (Hint: Filter and sort the data in the Pivot- table to locate the answer):
    i) Which small business customer had the highest sales?
    ii) Which corporate customer placed the greatest number of orders in 2019-2022? How many orders were placed by the corporate customer?
    iii) Which consumer customer was the most profitable one?
    iv) What is the sales figure of the least profitable home office customer?
6. Consider the following worksheet: (enter 5 records)

FULL NAME	GRADE 1/2/3	BASIC SALARY	HRA	PF	GROSS	NET	VA	VA>HRA

HRA is calculated as follows:

Grade	HRA (% of basic)
1	40%
2	35%
3	30%

PF is 8% for all grades

VA is 15000, 10000, 7000 for Grades 1, 2 and 3.

Gross=Basic + HRA+VA

Net=Gross - PF

i) Find max, min and average salary of employees in respective Grade.
ii) Count no. of people where VA>HRA
iii) Find out most frequently occurring grade.
iv) Extract records where employee name starts with "A" has HRA>10000
v) Print Grade wise report of all employees with subtotals of net salary and also grand totals.
vi) Use subtotal command.
vii) Extract records where Grade is 1 or 2 and salary is between 10000 and 20000 both inclusive.

7. Create workbook related to sales of Business Company having various product in last ten quarters for 20 sales persons. Perform the following on workbook:
   i) Create and modify a Pivot-table
   ii) Apply Pivot-table styles and formatting
   iii) Filter a Pivot-table
   iv) Insert a slicer to filter a Pivot-table
   v) Create a Pivot Chart

8. Create a PivotTable showing Total Sales breakdown by Region, Product Category, and Product Sub-Category. Use information from the PivotTable to answer the following questions:
   i) What was the Total Sales figure included in this data set?
   ii) Which Product Category had the highest sales?

iii) Which Region had the lowest sales?
iv) What was the Total Sales of Appliances in Delhi?

9. You are required to prepare a payroll statement in the given format making maximum use of cell referencing facility:

Code	Name	Category	Is HRA to be Paid	Basic	DP	DA	HRA	TA	CCA	Gross
1			Y							
2			N							
	Total									

Required:

- Basic salary (Allow any Basic salary in the range of Rs.10000-35000)
- DP is 50% of Basic Salary.
- DA (as a Percentage of Basic + DP) is more than 35000 then 40% of basic else 30% of basic.
- HRA is to be paid @ 40% of (Basic plus DP) to those whom HRA payable is yes.
- TA is to be paid @ Rs. 800 PM if Basic Salary is Less than Rs.12000, otherwise the TA is Rs. 1000 PM)
- CCA is to be paid @ Rs. 300 PM if Basic Salary is less than Rs.12000/- otherwise the CCA is Rs. 500 PM)
- Gross salary is the sum of Salary and all other allowances
- Deduction: a) GPF 10% of (Basic +DP) subject to a minimum of Rs.2000/- b) IT 10% of Gross Salary
- Net salary is Gross salary minus total deductions.

10. Consider the following worksheet for APS 1st year students:

S.No.	Name	Physics	Chem	Bio	Maths	CS	Total	%	Grade

1									
2									
3									
4									
5									

The value of Grade is calculated as follows:

If % >=90	Grade A

If % >=80 & <90	Grade B
If % >=70 & <80	Grade C
If % >=60 & <70	Grade D

Otherwise, students will be declared fail.

i) Calculate Grade using if function
ii) Sort the data according to total marks
iii) Apply filter to display the marks of the students having more than 65% marks.
iv) Enter the S.No. of a student and find out the Grade of the student using VLOOKUP.
v) Extract all records where name
   a) Begins with “A”
   b) Contains “A”
   c) Ends with “A”

Essential/recommended readings

1. Gupta, S.P., *Elementary Statistical Methods*, Sultan Chand and Sons, New Delhi, 2017.

2. Goldmeier, J., *Advanced Excel Essentials*, Apress, 2014.
3. Slager, D., *Essential Excel 2016: A Step-by-Step Guide*, Apress, 2016.
4. Valerie M. Sue and Matthew T. Griffin, *Data Visualization and Presentation with Microsoft Office*, SAGE, 2016.
5. Schmuller, J., *Statistical Analysis with Excel for Dummies*, 4th edition., Wiley India Pvt Ltd., 2020.

Note: Examination scheme and mode shall be as prescribed by the Examination Branch, University of Delhi, from time to time.