

## 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	Pass in Class XII	Pass in Class XII and 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 (5 Weeks)

**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 (2 Weeks)**

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

#### **UNIT – III (3 Weeks)**

**Recursion:** Recursive functions, linear recursion, binary recursion.

#### **UNIT – IV (2 Week)**

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

#### **UNIT – V(3 Weeks)**

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

#### **Practical component (if any) -**

1. Perform matrix addition and multiplication.
2. Implement following recursive functions:
  - Factorial of a number
  - $N^{\text{th}}$  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++*, 2<sup>nd</sup> edition, Wiley, 2011.
2. Cormen, T.H., Leiserson, C.E., Rivest, R. L. Stein C. *Introduction to Algorithms*, 4<sup>th</sup> edition, Prentice Hall of India, 2022.
3. Drozdek, A. *Data Structures and Algorithms in C++*, 4<sup>th</sup> edition, Cengage Learning, 2012.

#### **Suggestive readings**

- (i) Sahni, S., *Data Structures, Algorithms and applications in C++*, 2<sup>nd</sup> 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	Pass in Class XII	Pass in Class XII, and 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 matplotlib libraries

## SYLLABUS OF A2

### UNIT – I (2 Weeks)

**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 (3 Weeks)**

**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 (4 Weeks)**

**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 (4 Weeks)**

**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 (2 Weeks)**

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

### **Practical component (if any)**

Use a dataset of your choice from Open Data Portal ([https:// data.gov.in/](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*, 2<sup>nd</sup> 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*, 12<sup>th</sup> edition, Sultan Chand & Sons, 2020.

### Suggestive readings

- (i) Chen D. Y. *Pandas for Everyone: Python Data Analysis*, 1<sup>st</sup> 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.