

SEMESTER - I
STATISTICS FOR DATA
SCIENCE – I

1. Course Description

Programme: M.SC. DATA SCIENCE
Course Code: P24/CDS/DSC/102
Course Type: DSC
No. of credits: 3

Max. Hours: 45
Hours per week: 3 hrs.
Max. Marks: 100

2. Course Objectives:

1. To enable the students to understand the fundamentals of statistics to apply descriptive measures and probability for data analysis.
2. To help the students understand the importance of testing an hypothesis and to teach them how to draw conclusions.

3. Course Outcomes:

On completion of the course the student will be able to:

CO1: To examine the distribution of data and also able to measure it. To understand different concepts of probability and also to apply them to real life situations.

CO2: To make use of different concepts of random variables in understanding the scope of different distributions and to apply appropriate discrete and continuous distributions to any given situation.

CO3: To analyse and apply the concepts of hypothesis testing for large and small samples

DEPARTMENT OF COMPUTER SCIENCE
ST. FRANCIS COLLEGE FOR WOMEN, HYDERABAD

PROFESSOR
Department of Computer Science & Engineering
Osmania University
Hyderabad-500 007
P. V. S. S. S.

4. Course Content:**MODULE I: Descriptive statistics and probability** (10 Hrs)

Review of Measures of Central tendency and Dispersion. Moments, Measures of skewness and Kurtosis, Theorems on probability, conditional probability, and independent events, Addition theorem and Multiplication theorem of probability, Bayes' theorem and its applications.

MODULE- II: Random Variables And Probability Distributions (20 Hrs)

Definition of random variable, discrete and continuous random variables, functions of random variables, probability mass function and probability density function with illustrations. Distribution function and its properties, Notion of bivariate random variable, joint, marginal and conditional distributions, Independence of random variables. Mathematical Expectation of a random variable, Theorems on Expectations (only statements).

Discrete and Continuous distributions : Bernoulli, Binomial, Poisson, and Normal Distributions - definition, properties and applications with numerical problems

MODULE-III: Testing of Hypothesis (15 Hrs)

Concepts of population, parameter, random sample, statistic, sampling distribution and standard error. Exact sampling distributions- statement and properties of χ^2 , t and F distributions. Concepts of statistical hypothesis, Null and Alternative hypothesis, Critical region, two types of errors, Level of significance and Power of a test, One and two tailed tests, procedure for hypothesis testing, Large sample tests for means and proportions, t-test for means and Chi-square test for goodness of fit.

5. References:

1. Goon A.M., Gupta M.K. and Dasgupta B. (2002): Fundamentals of Statistics, Vol. I & II, 8th Edn. The World Press, Kolkata.
2. Miller, Irwin and Miller, Marylees (2006): John E. Freund's Mathematical Statistics with Applications, (7th Edn.), Pearson Education, Asia.
3. V.K. Kapoor and S.C. Gupta (2014): Fundamentals of Mathematical Statistics, Sultan Chand & Sons, New Delhi
4. Mood, A.M. Graybill, F.A. and Boes, D.C. (2007): Introduction to the Theory of Statistics, 3rd Edn., (Reprint), Tata McGraw-Hill Pub. Co. Ltd.
5. Hoel P.G: Introduction to mathematical Statistics, Asia Publishing house.

DEPARTMENT OF COMPUTER SCIENCE
ST. FRANCIS COLLEGE FOR WOMEN, HYDERABAD

PROFESSOR
Department of Computer Science & Engineering
University of Osmania
College of Engineering (A)
Osmania University
Hyderabad-500 007

6. Syllabus Focus

a) Relevance to Local, Regional, National and Global Development Needs

Local /Regional/National /Global Development Needs	Relevance
Global	<p>Descriptive statistics, probability theory and random variables have diverse applications across various fields including data analysis, decision making, gaming theory, risk assessment, modelling uncertainty, and predictive modelling.</p> <p>probability distributions and hypothesis testing are integral components of statistical analysis with broad applications across diverse fields globally, playing a crucial role in decision-making, research, and problem-solving.</p>

b) Components on Skill Development/Entrepreneurship Development/Employability

SD/ED/EMP	Syllabus Content	Description of Activity
Employability	Testing of Hypothesis	<p>The demand for professionals with expertise in statistical analysis, hypothesis testing, and data-driven decision-making continues to grow across various industries which include Data Scientist/Analyst, Research Analyst, Quality control analyst, Business/Marketing Analyst etc.</p>

7. Pedagogy

S. No	Student Centric Methods Adopted	Type / Description of Activity
1.	MCQ test	Experiential learning
2.	Assignment	Experiential learning
3.	Case Study	Problem Solving

8. Course Assessment Plan**a) Weightage of Marks in Continuous Internal Assessments and End Semester Examination**

CO	Continuous Internal Assessments CIA -40%	End Semester Examination-60%
CO1	CIA-1- Written Exam	Written Exam
CO2	CIA-1- Written Exam	
CO3	CIA-2 written test/Assignment	
CO4	CIA-2 MCQ test	

b) Model Question Paper- End Semester Exam

MODEL QUESTION PAPER

STATISTICS FOR DATA SCIENCE – I
THEORYCourse Code: P24/CDS/DSC/102
Credits: 3Max. Marks: 60
Time: 2 1/2Hrs.

SECTION –A

I. Answer the following

3X12=36

1. Define Central and Non central moments. Establish the relationship between central moments in terms of raw moments.

(OR)

2. (i) State and prove Baye's theorem?

(ii) From 6 positive and 8 negative numbers, 4 numbers are chosen at random without replacement and multiplied, what is the probability that the product is positive.

3. Let X be a continuous random variable with distribution:

$$f(x) = kx^2, \text{ if } 0 \leq x \leq 1$$

(i) Evaluate k (ii) find $P\left(\frac{1}{4} \leq x \leq \frac{3}{4}\right)$ (iii) Find $P\left(x \leq \frac{2}{3}\right)$

(OR)

4. a) What are the characteristics of normal distribution

b) The G.P.A of university students is assumed to be Normal with mean 2.85 and variance 0.1225. Find the percentage of students whose G.P.A is

(i) more than 3.5 (ii) less than 2.55 . (TABLE

VALUE: $P(0 < Z < 1.86) = 0.9686$; $P(0 < Z < 0.86) = 0.3051$)

5. (i) Explain the large sample test procedure for testing the mean.

(ii) A geneticist is interested in the proportions of males and females in a population who have a certain minor blood disorder. In a random sample of 100 males, 31 are found to be afflicted, whereas only 24 of 100 females tested have the disorder. Can we conclude at the 0.01 level of significance that the proportion of men in the population afflicted with this blood disorder is significantly greater than the proportion of women afflicted?

(OR)

6. A marketing agency gives the following information about age group and the liking for a particular model of bike which a company plans to introduce.

	Age group		
	Below20	20 – 39	40 - 59
Liked	125	420	60
Unliked	75	220	100

On the basis of the above data, can it be concluded that model appeal is independent of the age group. (Table value = 42.786).

SECTION –B

II. Answer any FOUR.

4x6=24


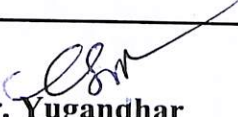
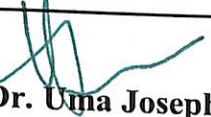
7. Explain the concept of skewness.
8. If A and B are independent events then show that (i) \bar{A} and B (ii) A and \bar{B} are also independent.
9. Define Distribution function? State the properties?
10. Two dice are rolled. Find the expectation of sum of the points on the two dice.
11. If X is a poisson variate such that $3P(X = 4) = \frac{1}{2} P(X = 2) = P(X = 0)$,
Find (i) Mean of X (ii) $P(X = 2)$.
12. A sample of 1000 tyres is taken from a lot. The mean life of the tyres is found to be 39,350 km with a standard deviation of 3260 km. Could the sample come from a population with mean life of 40,000km?

c) Question Paper Blueprint


Modules	Hours Allotted in the Syllabus	COs Addressed	Section A (No. of Questions)	Total Marks	Section B (No. of Questions)	Total Marks
1	10	CO-1	2	12	2	3
2	20	CO-2	2	12	2	3
3	15	CO-3	2	12	2	3

9. CO-PO Mapping

CO	PO	Cognitive Level	Class room sessions (Hrs.)
1	2	Apply/Understand	10
2	2	Remember/Apply	20
3	1	Analyze/Apply	15

Prepared by	Checked & verified by	Approved by
 G. Priyanka Teaching Faculty	 Dr. Yugandhar Head, Department of Statistics	 Dr. Uma Joseph, Principal

DEPARTMENT OF COMPUTER SCIENCE
 ST. FRANCIS COLLEGE FOR WOMEN, HYDERABAD

PROFESSOR
 Department of Computer Science & Engineering
 Osmania University
 Hyderabad-500 007.


SEMESTER II
DATA ANALYSIS USING R
PRACTICAL

1.Course Description**Programme: M.Sc. Data Science****Max. Hours: 60****Course Code: P24/CDS/DSC/102/P****Hours per week: 4****Course Type: DSC****Max. Marks: 50****No. of credits: 2****2.Course Objective**

- To explore various Data analysis and Data visualization methods using R.

3.Course Outcomes**CO1:** To be able to perform Exploratory data analysis.**CO2:** Applying visualization tools and libraries such as ggplot2 in R to visualize the data.

4. Practical Sessions

Installation of library-ggplot2.

1. Data Analyze/Visualize with ggplot2
 - a. Creating ggplot2
 - b. Facets
 - c. Geometric Objects
 - d. Statistical Transformations
 - e. Position Adjustments
 - f. Coordinate Systems
 - g. The Layered Grammar of Graphics
2. Data Transformation with dplyr
 - a. Filter Rows with filter ()
 - b. Arrange Rows with arrange ()
 - c. Select Columns with select ()
 - d. Add New Variables with mutate ().
 - e. Grouped Summaries and Grouped Mutates
3. Exploratory Data Analysis
 - a. Variation
 - b. Missing Values
 - c. Covariation
4. Working on Tibbles with Tibble, Data Import with readr, Tidy Data with tidyr
 - a. Creating Tibbles, Tibbles Versus data. Frame
 - b. Data Import with readr: Parsing a Vector-Numbers, Strings
 - c. Tidy Data with tidyr: Tidy Data, Spreading and Gathering, Separating and Pull, Missing Values, Nontidy data.
5. Computation of Relational Data with dplyr, Strings with stringr.

MODEL QUESTION PAPER
DATA ANALYSIS USING R
PRACTICAL




COURSE CODE: P24/CDS/DSC/102/P

Max.Marks: 50

Exam Duration: 2 hrs

Answer any one of the following.

1. Using R perform exploratory data analysis methods on any dataset.
2. Create the following Data Transformation with dplyr.
 - a) Filter Rows with filter ().
 - b) Arrange Rows with arrange ()
 - c) Select Columns with select ()
 - d) Add New Variables with mutate ().
3. Implement statistical transformations on any data set and visualize using ggplot2.
4. Show the Grouped Summaries and Grouped Mutates of any dataset.
5. Show the computation relational data with dplyr.

Prepared by	Checked & Verified by	Approved by
 Ms.Khalida Tabassum Teaching Faculty	 Ms. D. Sowjanya Head of the Department	 Dr.Uma Joseph Principal

SEMESTER I
DATA STRUCTURES
THEORY

1. Course Description

Programme: M.Sc. Data Science
Course Code: P24/CDS/DSC/104
Course Type: DSC
No. of credits: 3

Max. Hours:45
Hours per week:3
Max. Marks: 100

2.CourseObjective:

- To understand the importance of structuring the data for easy access and storage.
- To design and implement and develop various data structure algorithms.
- To acquire skills in using generic principles for data representation and manipulation with a view for efficiency, maintainability and code reuse.

3.Course Outcomes:

- CO1:** To evaluate and analyze the complexity of given algorithms to apply for Problems solving like sorting, searching.
- CO2:** To understand the concepts of stacks and queues.
- CO3:** To implement operations like insertion, deletion, traversing mechanism on linear and non- linear data structures like lists and trees.

DEPARTMENT OF COMPUTER SCIENCE
ST.FRANCIS COLLEGE FOR WOMEN, HYDERABAD

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Shaikfa University
Hyderabad-500 007.

4. Course Content**MODULE I: INTRODUCTION TO DATA STRUCTURES (15 Hrs.)**

Introduction to data structures. Sorting and Searching- Sorting Techniques- Bubble sort, mergesort, quicksort, heapsort, searching Techniques- linear search, binary search.

MODULE II: STACKS AND QUEUES AND LINKED LISTS (15 Hrs.)

Stacks-Introduction, Stacks, stack operations, stack Implementation.

Queues- Introduction, Queues-Basic Concept, Queue Operations, Queue Implementation
Circular Queues, Double-Ended Queues.

Linked Lists- Introduction, Linked List-Basic Concept, Linked List Implementation

MODULE III: LINKED LISTS, TREES, GRAPHS (15 Hrs.)

Trees- Introduction, Basic Concept, Binary Tree, Binary Tree Representation, Binary Tree Traversal, Binary Search Tree.

Graphs- Basic Concept, Graph Terminology, Graph Implementation, Shortest path Algorithm, Graph Traversal.

5. References:

1. Balaguruswamy, Data Structures using Java. McGraw Hill Education (India) Private Limited, 2019.
2. Data Structures, by G.S. Baluja, Dhanpat Rai & Co. (P) Limited, First Edition
3. Data Structures & Algorithms in Java, by Robert Lafore, Pearson Education, Second Edition.
4. Data Structures & Algorithms in Java, by Michael T. Goodrich. Robert Tamassia, Wiley; 5th edition, Fifth edition
5. Langsam, Augenstein, Tenenbaum, "Data Structures using Java"

6. Syllabus Focus

a) Relevance to Local, Regional, National and Global Development Needs

Local /Regional/National /Global Development Needs	Relevance
Global Development	This course helps students learn different Algorithms in data Structures. It provides a means to manage large amounts of data efficiently. And efficient data structures are key to designing efficient algorithms

b) Components on Skill Development/Entrepreneurship Development/Employability

SD/ED/EMP	Syllabus Content	Description of Activity
SD	Modules 1 and 2	Presentation. You can build real-world applications and projects with Java enhancing your employability.
EMP	Modules 3	Java developers use data structures to store and organize data. Good knowledge of data structures in Java can help programmers write efficient and optimized Java programs.

7. Pedagogy

S. No	Student Centric Methods Adopted	Type / Description of Activity
1.	Group Discussion	Participative Learning Group Discussion
2.	Interactive Classroom games	Experiential Learning
3.	Case studies	Problem solving

8. Course Assessment Plan

a) Weightage of Marks in Continuous Internal Assessments and End Semester Examination

CO	Continuous Internal Assessments CIA-40%	End Semester Examination-60%
CO1	CIA-1 written exam	Written Exam
CO2	CIA-1 Assignment	
CO3	CIA-2 Presentation	

b) Model Question Paper- End Semester Exam

DATA STRUCTURES
MODEL QUESTION PAPER
THEORY

Course Code: P24/CDS/DSC/104
Credits: 3

Max Marks: 60
Time: 2 ½ Hrs.

SECTION – A

I Answer the following

3 x 12=36

1. a) Write a program to implement selection sort.
b) Trace the following list using selection sort 23 12 35 19 41 38

OR

2. Explain the merge sort with an example.
3. Explain the stack operations in detail. Give an example.

OR

4. Write the procedures to insert and delete data using a single linked list.
5. Explain binary trees. List and explain the various tree traversals with an example

OR

6. Explain different graph search methods with an example.

SECTION - B

II. Answer any Four

4 x 6 = 24

7. Write a program to sort the elements using bubble sort.
8. Explain the merge sort with an example.
9. Write short notes on the applications of circular queues.
10. Write the node structure for a single linked list.
11. Explain binary tree operations. Give an example for In-order tree traversal.
12. Explain BFS with an example.

c) Question Paper Blueprint

Modules	Hours Allotted in the Syllabus	COs Addressed	Section A (No. of Questions)	Total Marks	Section B (No. of Questions)	Total Marks
1	15	CO-1	2	12	2	12
2	15	CO-2	2	12	2	6
3	15	CO-3	2	12	2	6

9. CO-PO Mapping

CO	PO	Cognitive Level	Classroom sessions(hrs.)
1	1	2	15
2	2	3,4	15
3	2	5	15

DATA STRUCTURES PRACTICAL

Programme: M.Sc. Data Science
Course Code: P24/CDS/DSC/104/P
Course Type: DSC
No. of credits: 2

Max. Hours: 60
Hours per week:4
Max. Marks: 50

Course Objective:

- To introduce the fundamental concepts of data structures through JAVA language
- To apply the concepts of data structures.

Course Outcomes:

CO1: To understand searching and sorting techniques using coding exercise

CO2: Apply the principles of stacks, queues and linked lists

CO3: Develop appropriate tree and graphs search algorithms

Practical Sessions

1. Program to implement linear search.
2. Program to implement Binary search.
3. Program to implement Bubble sort, Selection sort and Insertion sort.
4. Program to implement Merge sort.
5. Program to implement Quick sort.
6. Program to implement Heap sort.
7. Program to implement Stack using arrays.
8. Program to implement Queue using arrays.
9. Program to implement Singly linked list.
10. Program to implement Linked list using stack.
11. Program to implement Linked list using queue.
12. Program to implement Doubly linked list.
13. Program to implement Binary tree traversals.
14. Program to implement Breadth First Search.
15. Program to implement Breadth First Search.

**DATA STRUCTURES
MODEL QUESTION PAPER
PRACTICAL**




Course Code: P24/CDS/DSC/104/P
Credits: 2

Max. Marks: 50
Max. Time: 2 Hrs.

I. Answer any ONE of the following

2 x 25 = 50

1. Write a Java Program to implement Merge sort.
2. Write a Java Program to insert and delete an element from a Stack using linked lists.
3. Write a Java Program to implement Binary tree traversals.

Prepared by	Checked & verified by	Approved by
 Ms. B. Jyothi PG Coordinator	 Ms. D. Sowjanya Head of Department	 Dr. Uma Joseph Principal

SEMESTER – I
FUNDAMENTALS OF DATA SCIENCE
THEORY

1. Course Description

Programme:	M.Sc. Data Science	Max. Hours: 45
Course Code:	P24/CDS/DSC/101	Hours per week: 3
Course Type:	DISCIPLINE SPECIFIC CORE	Max. Marks: 100
No. of credits:	3	

2. Course Objectives

- Learn to explore and analyze mega data from various sources.
- To optimize business processes or for decision support.

3. Course Outcomes

On completion of the course the student will be able to:

CO1: *Illustrate* and *explain* the various stages of data science. (Cognitive level – 2)

CO2: *Analyze & Use* of NoSQL databases. (Cognitive level – 3,4)

CO3: *Apply* NoSQL development tools on different types of NoSQL Databases. (Cognitive level - 5)

DEPARTMENT OF COMPUTER SCIENCE
ST. FRANCIS COLLEGE FOR WOMEN, HYDERABAD

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.
P.V. Srinivas

4. Course Content

MODULE I: INTRODUCTION TO DATA SCIENCE

(15 Hrs.)

Benefits and uses, Facets of data, The data science process, Applications of data science. Overview of the data science process- Retrieving data, Cleansing, integrating, and transforming data, Exploratory data analysis, Build the models.

MODULE II: GRAPH AND XML DATABASES

(15 Hrs.)

Graph Databases – Graphs and graph structures, graph data structures, the property graph model, storing property graphs in relational tables, advanced graph models. XML Databases – XML concepts, XML query languages, storing XML in relational databases, native XML storage. Key value storage – map reduce, Document databases – JSON, REST

MODULE III: OBJECT AND DISTRIBUTED DATABASES

(15 Hrs.)

Column stores – column wise storage, column striping. Object Databases – Object orientation, object relational mapping, object mapping APIs, object databases Distributed Databases – scaling horizontally, distribution transparency, failure in distributed systems. Data fragmentation – properties, types of fragmentation, fragmentation approaches, data allocation.

5. References

1. Davy Cielen, Arno D. B. Meysman and Mohamed Ali, Manning.
2. Advanced Data Management: For SQL, NoSQL, Cloud and Distributed.
3. Databases (De Gruyter Textbook) by Lena Wiese, 29 October 2015.
4. The Art of Data Science by Roger D Peng and Elizabeth Matsui.
5. NoSQL Distilled by Pramod J. Sadalage Martin Fowler, Addison-Wesley, 2013.
6. Next Generation Databases NoSQL, NewSQL, and Big Data by Guy Harrison, APress 2015.
7. NoSQL for Mere Mortals by Dan Sullivan, Addison-Wesley, 2015.

6. Syllabus Focus

a) Relevance to Local, Regional, National and Global Development Needs

Local/Regional/National /Global Development Needs	Relevance
Global Development	Fundamentals of data science are essential for advancing global development goals by enabling evidence-based decision-making, predictive modeling, resource optimization, monitoring and evaluation, equity promotion, and capacity building. Organizations can accelerate progress towards achieving sustainable development outcomes and improving the well-being of communities worldwide.

b) Components on Skill Development/Entrepreneurship Development/Employability

SD/ED/EMP	Syllabus Content	Description of Activity
SD	Modules 1	Learn benefits, different facets of data, and overview of data science processes.
EMP	Modules 2 and 3	Represent data using NoSQL formats.

7. Pedagogy

S. No	Student Centric Methods Adopted	Type / Description of Activity
1.	Participative	Seminars
2.	Experimental	Quiz
3.	Problem solving	Troubleshoot (debug) code

8. Course Assessment Plan

a) Weightage of Marks in Continuous Internal Assessments and End Semester Examination

CO	Continuous Internal Assessments CIA -40%	End Semester Examination-60%
CO1	CIA 1 – Written	Written Exam
CO2	CIA 2 – Assignment	
CO3	CIA 2 – Lab	

DEPARTMENT OF COMPUTER SCIENCE
ST. FRANCIS COLLEGE FOR WOMEN, HYDERABAD

PROFESSOR
Department of Computer Science & Engineering
Osmania University
Hyderabad-500 007.

b) **Model Question Paper- End Semester Exam**

FUNDAMENTALS OF DATA SCIENCE

MODEL QUESTION PAPER

THEORY

Course Code: P24/CDS/DSC/101
Credits: 3

Max Marks: 60
Time: 2 ½ Hrs.

I. Answer the following

3 x 12 = 36

1. Explain the main categories of data in data science in detail.

OR

2. Explain in detail the steps followed in the data science process with the help of a diagram.

3. Explain how data is represented in graph structures. Explain with an example the graph models.

OR

4. Explain the XML format of the database with key value storage in detail.

5. Explain the need for data fragmentation. Explain the properties for fragmentation and the various approaches of fragmentation.

OR

6. Explain the techniques of distributing the databases in detail.

SECTION – B

II. Answer any FOUR

4 x 6 = 24

7. Explain applications of data science in detail.

8. Explain in detail data cleaning, integrating and transforming of data with the help of diagrams.

9. Explain how to document databases using JSON.

10. Explain advanced graph models in detail.

11. Explain how to store data column wise and need for striping it with an example.

12. Explain distribution transparency in detail.

DEPARTMENT OF COMPUTER SCIENCE
ST. FRANCIS COLLEGE FOR WOMEN, HYDERABAD

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

c) Question Paper Blueprint

Modules	Hours Allotted in the Syllabus	COs Addressed	Section A (No. of Questions)	Total Marks	Section B (No. of Questions)	Total Marks
I	15	1	2	12	2	12
II	15	2	2	12	2	6
III	15	3	2	12	2	6

NOTE: From

Section-A any 3 questions can be answered (INTERNAL CHOICE).

Section-B any 4 questions can be answered. (EXTERNAL CHOICE)

9. CO-PO Mapping

CO	PO	Cognitive Level	Classroom sessions(hrs)
1	1	2	15
2	2	3,4	15
3	2	5	15

DEPARTMENT OF COMPUTER SCIENCE
ST. FRANCIS COLLEGE FOR WOMEN, HYDERABAD

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

FUNDAMENTALS OF DATA SCIENCE

PRACTICAL

Course Description

Programme: M.Sc. Data Science
Course Code: P24/CDS/DSC/101/P
Course Type: DISCIPLINE SPECIFIC CORE
No. of credits: 2

Max. Hours: 60
Hours per week: 4
Max. Marks: 100

Course Objectives

- To explore and analyze mega data from various sources.

Course Outcomes

On completion of the course the student will be able to:

CO1: *Analyze & Use* of NoSQL databases.
(Cognitive level – 3,4)

CO2: *Apply* NoSQL development tools on different types of NoSQL Databases. (Cognitive level - 5)

Practical Sessions

Suggested Exercises:

Neo4j

1. Query to create Nodes
2. Query to create Relationships between nodes:
3. Query to return nodes and relationships:
4. Query to return nodes and relationships depending on condition.

BaseX

5. Query to store Employee Data in XML Document
6. Query to display employee details who belong to the IT department.
7. Query to count and display the number of employees who work in the marketing department.
8. Query to display empname, ethnicity, age, and annual salary of employees whose age is more than 40 years.
9. Query to display employee data in descending order of their age.
10. Query to display employee data depending on their city (use group by).

MonetDB

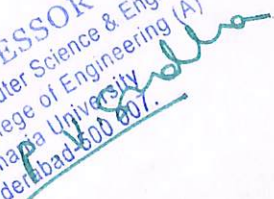
DEPARTMENT OF COMPUTER SCIENCE
ST. FRANCIS COLLEGE FOR WOMEN, HYDERABAD

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

11. Query to create a table
12. Query to insert data into the table and retrieve all data.
13. Query to retrieve data depending on condition.
14. Query to retrieve data using aggregate functions.
15. Query to perform group-by and order-by.

DEPARTMENT OF COMPUTER SCIENCE
ST. FRANCIS COLLEGE FOR WOMEN, HYDERABAD

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.



MODEL QUESTION PAPER

FUNDAMENTALS OF DATA SCIENCE PRACTICAL

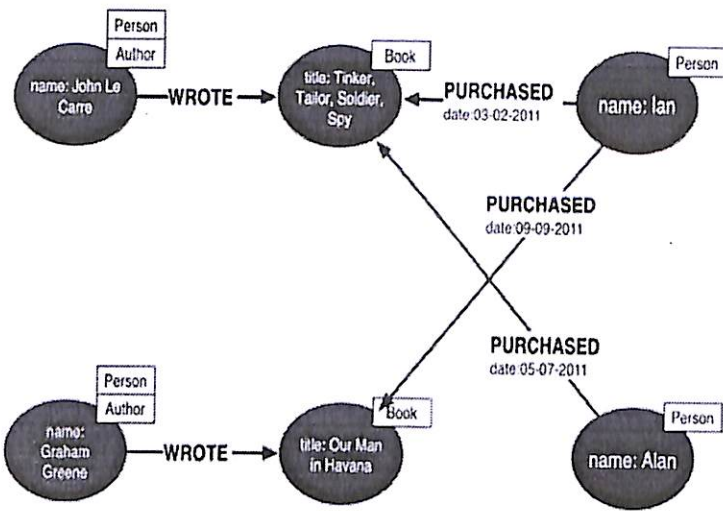
Course Code: P24/CDS/DSC/101/P

Time: 2Hrs
Max Marks:50

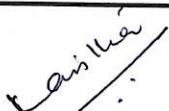


Answer any two:

Rank	Name	Age	Wealth (\$bil)	Wealth per year
1	Bernard Arnault & family	74	211	2.85
2	Elon Musk	51	212	4.16
3	Jeff Bezos	59	213	3.61
4	Larry Ellison	78	214	2.74
5	Warren Buffett	92	215	2.34
6	Bill Gates	67	216	3.22
7	Michael Bloomberg	81	217	2.68
8	Carlos Slim & family	83	218	2.63
9	Mukesh Ambani	65	219	3.37
10	Steve Ballmer	67	220	3.28

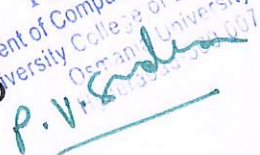
1. Create the above structure using Monetdb.
 - Insert minimum of 10 records.
 - Display the details of people whose age >70.
 - Display the details of people whose rank <5.
 - Display the sum of the wealth.
2. Create the above structure using BaseX.
 - Insert minimum of 10 records.
 - Display the details of people in ascending order of their age.
 - Display the details of people whose wealth per year is >3.00.
 - Count the number of people whose age is >80.



3. Create the above graph and perform the following queries.
- Display all names of the persons in the graph.
 - Display the details of the person who purchased books.
 - Display name of the author who wrote a book titled spy.

Prepared by	Checked & verified by	Approved by
 Ms. B. Kavitha Teaching Faculty	 Ms. D. Sowjanya Head of the Department	 Dr. Uma Joseph Principal

DEPARTMENT OF COMPUTER SCIENCE
 ST. FRANCIS COLLEGE FOR WOMEN, HYDERABAD

PROFESSOR
 Department of Computer Science & Engineering
 Osmania University
 Hyderabad 500 007.


SEMESTER - I
PYTHON PROGRAMMING
THEORY

1. Course Description

Programme: M.Sc. Data Science
Course Code: P24/CDS/DSC/103
Course Type: DISCIPLINE SPECIFIC CORE
No. of credits: 3

Max. Hours:45
Hours per week: 3
Max. Marks: 100

2. Course Objectives

- To understand the basic theory underlying Python programming and control structures.
- To learn and explore lists, tuples, dictionaries, files and exceptions.
- To learn and explore OOPS concepts and GUI programming.

3. Course Outcomes

On completion of the course the student will be able to:

CO1: *Illustrate* and *explain* control structures and how python is different from other programming languages. (Cognitive levels – 2)

CO2: *Apply* the usage of lists, tuples, dictionaries and analyze the files and exceptions.
(Cognitive level – 3,4)

CO3: *Illustrate* and *Design* object-oriented concepts, the graphical interfaces using tkinter module.
(Cognitive levels – 2,5,6)

DEPARTMENT OF COMPUTER SCIENCE
ST. FRANCIS COLLEGE FOR WOMEN, HYDERABAD

PROFESSOR
Department of Computer Science & Engineering
Osmania University
Hyderabad-500 007.
P.V. Sridhar

4. Course Content

I:MODULE INTRODUCTION TO PYTHON PROGRAMMING (15 Hrs.)

How a Program Works, Using Python, Why Python, Input, Processing, and Output, Displaying Output with the Print Function, Comments, Variables, Reading Input from the Keyboard, Performing Calculations (Operators. Type conversions, Expressions), More about Data Output, Indentation. Decision Structures and Boolean Logic: if, if-else, if-elif-else Statements, Nested Decision Structures, Comparing Strings, Logical Operators, Boolean Variables. Repetition Structures: Introduction, while loop, for loop, Calculating a Running Total, Input Validation Loops, Nested Loops. Functions, Modules.

MODULE II: LISTS AND TUPLES, DICTIONARIES AND SETS, FILE AND EXCEPTIONS (15 Hrs.)

Lists and Tuples: Sequences, Introduction to Lists, List slicing, Finding Items in Lists with the in Operator, List Methods and Useful Built-in Functions, Copying Lists, Processing Lists, Two Dimensional Lists, Tuples. Strings: Basic String Operations, String Slicing, Testing, Searching, and Manipulating Strings. Dictionaries and Sets: Dictionaries, Sets, Serializing Objects. Recursion: Introduction, Problem Solving with Recursion, Examples of Recursive Algorithms. File and Exceptions: Introduction to File Input and Output, Using Loops to Process Files, Processing Records, Exceptions.

MODULE III: OOPS CONCEPTS AND GUI PROGRAMMING (15 Hrs.)

OOPs Concept: Introduction to OOP, Classes and objects, Inheritance Method overloading and method overriding, Abstract method and Abstract class, Interfaces in python, Abstract class VS Interfaces, constructor, instance methods, class methods, static methods. GUI Programming: Graphical User Interfaces, Using the tkinter Module, display text with Label Widgets, Organizing Widgets with Frames, Button Widgets and Info Dialog Boxes, Getting Input with Entry Widget, Using Labels as Output Fields, Radio Buttons, Check Buttons.

5. References

1. Text Book: Tony Gaddis, Starting Out with Python (4th edition).
2. Kenneth A. Lambert, Fundamentals of Python
3. Clinton W. Brownley, Foundations for Analytics with Python
4. James Payne, Beginning Python using Python 2.6 and Python 3
5. Charles Dierach, Introduction to Computer Science using Python
6. Paul Gries, Practical Programming: An Introduction to Computer Science using Python 3 P

DEPARTMENT OF COMPUTER SCIENCE
ST. FRANCIS COLLEGE FOR WOMEN, HYDERABAD

PROFESSOR
Department of Computer Science & Engineering
Osmania University
Hyderabad-500 007.
H. V. S. Reddy

6. Syllabus Focus

a) Relevance to Local, Regional, National and Global Development Needs

Local /Regional/National /Global Development Needs	Relevance
Global Development	Python programming is versatile and has extensive libraries, data science capabilities, web development, scripting, automation, cross-platform compatibility, open-source nature, ease of learning.

b) Components on Skill Development/Entrepreneurship Development/Employability

SD/ED/EMP	Syllabus Content	Description of Activity
SD	Modules 1 and 2	Applying control structures, functions, lists and dictionaries on a given problem.
EMP	Modules 3	Testing CRUD operations

DEPARTMENT OF COMPUTER SCIENCE
ST. FRANCIS COLLEGE FOR WOMEN, HYDERABAD

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Hyderabad-500 007.

A. V. S. Reddy

7. Pedagogy

S. No	Student Centric Methods Adopted	Type / Description of Activity
1.	Participative	Seminars
2.	Experimental	Quiz
3.	Problem solving	Implementing OOPS concepts and GUI programming.

8. Course Assessment Plan**a) Weightage of Marks in Continuous Internal Assessments and End Semester Examination**

CO	Continuous Internal Assessments CIA -40%	End Semester Examination-60%
CO1	CIA 1 – Written	Written Exam
CO2	CIA 2 – Assignment	
CO3	CIA 2 – Lab	

DEPARTMENT OF COMPUTER SCIENCE
ST. FRANCIS COLLEGE FOR WOMEN, HYDERABAD

PROFESSOR
Department of Computer Science & Engineering
Osmania University
Hyderabad-500 007.

b) Model Question Paper- End Semester Exam**PYTHON PROGRAMMING****MODEL QUESTION PAPER
THEORY****Course Code: P24/CDS/DSC/103**
Credits: 3**Max Marks: 60**
Time: 2 ½ Hrs.**SECTION – A****I: Answer any Three****3 x 12 = 36**

1. Explain in detail about Decision structures and Repetition structures.
OR
2. Describe in detail functions and modules.
3. Define lists and tuples. Explain the various operations of lists and tuples with examples.
OR
4. Explain in detail about files and exceptions.
5. Discuss about Inheritance Method overloading and method overriding with an example.
OR
6. Explain in detail about GUI programming.

II. Answer any Four:**4 x 6 = 24**

7. Discuss about operators.
8. Write a short note on type conversions.
9. Explain in detail about dictionaries.
10. Write a short note on recursion.
11. Discuss classes and objects.
12. Differentiate between abstract class and interfaces.

c) Question Paper Blueprint

Modules	Hours Allotted in the Syllabus	COs Addressed	Section A (No. of Questions)	Total Marks	Section B (No. of Questions)	Total Marks
I	15	1	2	12	2	12
II	15	2	2	12	2	6
III	15	3	2	12	2	6

NOTE: From

Section-A any 3 questions can be answered (INTERNAL CHOICE).

Section-B any 4 questions can be answered. (EXTERNAL CHOICE)

9. CO-PO Mapping

CO	PO	Cognitive Level	Classroom sessions(hrs)
1	1	2	15
2	2	3,4	15
3	2	5,6	15

SEMESTER III
PYTHON PROGRAMMING
PRACTICAL

1.Course Description

Programme: M.Sc. Data Science
Course Code: P24/CDS/DSC/103/P
Course Type: DSC
No. of credits: 2

Max. Hours: 60
Hours per week: 4
Max. Marks: 50

2.Course Objective

To learn and develop modular programming and explore advanced concepts in python.

3.Course Outcomes

CO1: Exploring control structures, functions, oops concepts.
CO2: Applying CRUD operations by effectively connecting with databases.

4. Practical Sessions

1. Write a program that displays the following information: Your name, Full address, Mobile number, College name, Course subjects.
2. Write a program to find the largest three integers using if-else and conditional operator.
3. Write a program that asks the user to enter a series of positive numbers (The user should enter a negative number to signal the end of the series) and the program should display the numbers in order and their sum.
4. Write a program to find the product and sum of two matrices [A]m_xp and [B]p_xr using Numpy
5. Write recursive and non-recursive functions for the following:
 - a. To find GCD of two integers.
 - b. To find the factorial of positive integer
 - c. To print Fibonacci Sequence up to given number n
6. Write a program to display two random numbers that are to be added, such as: 247 + 129, the program should allow the student to enter the answer. If the answer is correct, a message of congratulations should be displayed. If the answer is incorrect, a message showing the correct answer should be displayed.
7. Write a function to demonstrate variable length arguments.
8. WAP to Demonstrate about Fundamental Data types (sequential and non-sequential) in Python

DEPARTMENT OF COMPUTER SCIENCE
ST. FRANCIS COLLEGE FOR WOMEN, HYDERABAD

PROFESSOR
Department of Computer Science & Engineering
St. Francis College for Women
Manama University
Hyderabad-500 007

P.V. Susha

Programming using type function.

9. Write a program to create a file, write the content and display the contents of the file.
10. In a program, write a function that accepts two arguments: a list and a number n. The function displays all of the numbers in the list that are greater than the number n.
11. Write a program with a function that accepts a string as an argument and returns the no. of vowels that the string contains. Another function to return the number of consonants.
12. Write a program that opens a specified text file and then displays a list of all the unique words found in the file. (Store each word as an element of a set.)
13. Write a program to analyze the contents of two text files using set operations.
14. Write a program to implement the inheritance and dynamic polymorphism.
15. Write a GUI program that converts Celsius temperatures to Fahrenheit temperatures.
16. Write a GUI program that displays you details when a button is clicked.

DEPARTMENT OF COMPUTER SCIENCE
ST. FRANCIS COLLEGE FOR WOMEN, HYDERABAD

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 008

MODEL QUESTION PAPER

PYTHON PROGRAMMING
PRACTICAL




COURSE CODE: P24/CDS/DSC/103/P

Max.Marks: 50

Exam Duration: 2 Hrs.

Answer any one of the following.

1. Write recursive and non-recursive functions for the following: a. To find GCD of two integers. b. To find the factorial of positive integer c. To print Fibonacci Sequence up to given number n
2. Write a function to demonstrate variable length arguments.
3. In a program, write a function that accepts two arguments: a list and a number n. The function displays all of the numbers in the list that are greater than the number n.
4. Write a program with a function that accepts a string as an argument and returns the no. of vowels that the string contains. Another function to return the number of consonants.
5. Write a program to implement the inheritance and dynamic polymorphism.

Prepared by	Checked & verified by	Approved by
 Ms. Khalid Tabassum Teaching Faculty	 Ms. D. Sowjanya Head of the Department	 Dr. Uma Joseph Principal

DEPARTMENT OF COMPUTER SCIENCE
ST. FRANCIS COLLEGE FOR WOMEN, HYDERABAD

PROFESSOR
Department of Computer Science & Engineering
University of Osmania
College of Engineering (A)
Hyderabad-500 007

