Second Year B.Tech. Program in Computer Science and Engineering (Data Science) - Semester III

Scheme for Second Year B.Tech. Program in Computer Science & Engineering (Data Science): Semester III

Sr	Course Code	Course	Teaching Scheme				Semester End Examination (A)						Continuous Assessment (B)					Aggregate (A+B)	Credits earned	
			Theory (hrs.)	Practical (hrs.)	Tutorial (hrs.)	Credits	Duration (Hrs)	Theory	Oral	Pract	Oral & Pract	End Sem Exam Total	Term Test 1 (TT1)	Term Test 2 (TT2)	Avg (TT1 & TT2)	Term Work Total	CA Total			
1	DJ19DSC301	Mathematics for Intelligent Systems	3	--	--	3	3	75	--	--	--	75	25	25	25	--	25	100	3	4
	DJ19DSL301	Mathematics for Intelligent Systems - Tutorial	--	--	1	1	--	--	--	--	--	--	--	--	--	25	25	25	1	
2	DJ19DSC302	Data Structures and Algorithms	3	--	--	3	3	75	--	--	--	75	25	25	25	--	25	100	3	4
	DJ19DSL302	Data Structures and Algorithms Laboratory	--	2	--	1	2	--	--	--	25	25	--	--	--	25	25	50	1	
3	DJ19DSC303	Foundations of Data Analysis	3	--	--	3	3	75	--	--	--	75	25	25	25	--	25	100	3	4
	DJ19DSL303	Foundations of Data Analysis Laboratory	--	2	--	1	2	--	--	--	25	25	--	--	--	25	25	50	1	
4	DJ19DSC304	Database Management Systems	3	--	--	3	3	75	--	--	--	75	25	25	25	--	25	100	3	4
	DJ19DSL304	Database Management Systems Laboratory	--	2	--	1	--	--	--	--	--	--	--	--	--	25	25	25	1	
5	DJ19DSC305	Statistics for Data Science	3	--	--	3	3	75	--	--	--	75	25	25	25	--	25	100	3	4
	DJ19DSL305	Statistics for Data Science Laboratory	--	2	--	1	--	--	--	--	--	--	--	--	--	25	25	25	1	
6	DJ19DSL306	Programming with Python Laboratory	--	2	--	1	2	--	--	--	25	25	--	--	--	25	25	50	1	1
7	DJ19A2	Innovative Product Development-I	--	2	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
8	DJ19A3	Constitution of India	1	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
		Total	**16**	**12**	**1**	**21**	**21**	**375**	**0**	**0**	**75**	**450**	**125**	**125**	**125**	**150**	**275**	**725**	**21**	**21**

Program: Second Year B. Tech. in Computer Science and Engineering (Data Science)	**Semester: III**
Course: Mathematics for Intelligent Systems	**Course Code:** DJ19DSC301
Course: Mathematics for Intelligent Systems Tutorial	**Course Code:** DJ19DST301

Teaching Scheme (Hours / week)				Evaluation Scheme						
				Semester End Examination Marks (A)			Continuous Assessment Marks (B)			Total marks (A+ B)
Lectures	Practical	Tutorial	Total Credits	Theory			Term Test 1	Term Test 2	Avg.	
				75			25	25	25	100
				Laboratory Examination			Term work		Total Term work	
3	--	1	4	Oral	Practical	Oral & Practical	Laboratory Work	Tutorial / Mini project / presentation/ Journal		25
				--	--	--	15	10	25	

Pre-requisites: Concepts of basic matrices, partial derivatives and basic probability.

Objective:

To build the strong foundation in learners of mathematics needed for building concepts of machine learning.

Outcomes: On completion of the course, learner will be able to:

1. Analyze probability of random variable and probability distributions
2. Demonstrate knowledge of linear algebra
3. Apply concepts of matrix theory
4. Demonstrate concepts of calculus
5. Analyze different optimization techniques

Detailed Syllabus: (unit wise)		
Unit	**Description**	**Duration**
1.	**Probability, Random variables and Probability distributions** Probability: Conditional probability, mutually and pair wise independent events, Bayes' theorem Random variables: discrete random variable, probability mass function, discrete distribution function, continuous random variable, probability density function, continuous distribution function, mathematical expectation, moment generating function, two-dimensional random variable and its joint probability mass and density function, marginal distribution function, conditional distribution functions, covariance, joint moments Probability distributions: discrete probability distribution: binomial distribution, poisson distribution, hypergeometric distribution, continuous probability distribution: uniform	**10**

	distribution, exponential distribution, normal distribution, beta distribution, gamma distribution, central limit theorem	
2.	**Linear algebra** vectors in n-dimensional vector space, properties, dot product, cross product, norm and distance, vector spaces over real field, properties of vector spaces over real field, subspaces, linear independence and dependence of vectors, span of vectors, basis of a vector space, dimension of a vector space, Cauchy Schwarz inequality, linear transformation, Norms and spaces, orthogonal compliments and projection operator, Kernel Hilbert spaces	**08**
3.	**Matrix theory** Characteristic equation, Eigen values and Eigen vectors, properties of Eigen values and Eigen vectors, Cayley-Hamilton theorem, examples based on verification of Cayley Hamilton theorem, similarity of matrices, diagonalization of matrices, functions of square matrix, derogatory and non-derogatory matrices, least squared and minimum normed solutions	**08**
4.	**Calculus** Gradient, directional derivatives, Jacobian, hessian, convex sets, convex functions and its properties	**04**
5	**Optimization** Unconstrained and Constrained optimization, convergence Unconstrained optimization techniques: newton's method, quasi newton method Constrained optimization techniques: gradient descent, stochastic gradient descent, penalty function method, lagrange multiplier method, Karush–Kuhn–Tucker method, simplex method, penalty and duality, dual simplex method, downhill simplex method	**12**

List of tutorials: (any eight)

1. To solve numerical on discrete probability distributions
2. To solve numerical on continuous probability distributions
3. To solve numerical on vector spaces (basis and dimension)
4. To solve numerical on cauchy-schwarz inequality and linear transformation
5. To solve numerical on diagonalizability using eigenvalues and eigenvectors
6. To solve numerical on minimal polynomial and functions of a matrix
7. To solve numerical on calculus
8. To solve numerical on Gradient descent and Lagrange's multiplier method
9. To solve numerical on KKT method
10. To solve numerical on all forms of simplex method

Any other tutorial based on syllabus may be included, which would help the learner to understand topic/concept.

Books Recommended:

Text Books:

1. Higher Engineering Mathematics, Dr. B. S. Grewal, Khanna Publication, 44th Edition, Publication Year 1965
2. Mathematical Methods in Science and Engineering, Kanti B. Datta, Cengage Learning India, 1st Edition, Publication Year 2011
3. Operations Research - An Introduction, Hamdy A. Taha, Pearson, 10th Edition, Publication Year 2010
4. Operations Research, Kanti Swarup, P. K. Gupta, Mohan Man; S Chand Publication, 2020 Edition, Publication Year 2005

Reference Books:

1. Analysis for Applied Mathematics, W. Cheney, New York: Springer Science+Business Media, 1st edition, Publication Year 2001
2. Linear Algebra Done Right, S. Axler, (Third Edition), Springer International Publishing, Publication Year 2015
3. Numerical Optimization, J. Nocedal and S. J. Wright, New York: Springer Science+Business Media, 2nd Edition, Publication Year 2006
4. A First Look at Rigorous Probability Theory (Second Edition), J. S. Rosenthal, Singapore: World Scientific Publishing, 2nd Edition, Publication Year 2006
5. Linear Algebra‖ Schaum‘s outline series, Seymour Lipschutz and Marc Lipson, Mc-Graw Hill Publication, 4th Edition, Publication Year 2009
6. Advanced Engineering Mathematics, Erwin Kreysizg, John Wiley & Sons, Inc, 10th Edition, Publication Year 2000

Evaluation Scheme:

Semester End Examination (A):

Theory:

1. Question paper will be based on the entire syllabus summing up to 75 marks.
2. Total duration allotted for writing the paper is 3 hrs.

Continuous Assessment (B):

Theory:

1. Two term tests of 25 marks each will be conducted during the semester out of which; one will be a compulsory term test (on minimum 02 Modules) and the other can either be a term test or an assignment on live problems or a course project.
2. Total duration allotted for writing each of the paper is 1 hr.
3. Average of the marks scored in both the two tests will be considered for final grading.

Laboratory: (Term work)

Term work shall consist of minimum 8 Tutorials.

The distribution of marks for term work shall be as follows:

i. Laboratory work (Performance of Experiments): 15 Marks
ii. Journal documentation (Write-up and/or Assignments): 10 marks

The final certification and acceptance of term work will be subject to satisfactory performance of laboratory work and upon fulfilling minimum passing criteria in the term work.

Prepared by Checked by Head of the Department Principal

Second Year B.Tech. Program in Computer Science and Engineering (Data Science) Semester III

Program: Second Year B. Tech. in Computer Science and Engineering (Data Science)	**Semester: III**
Course: Data Structures and Algorithms	**Course Code:** DJ19DSC302
Course: Data Structures and Algorithms Laboratory	**Course Code:** DJ19DSL302

Teaching Scheme (Hours / week)				Evaluation Scheme						
				Semester End Examination Marks (A)			Continuous Assessment Marks (B)			Total marks (A+ B)
Lectures	Practical	Tutorial	Total Credits	Theory			Term Test 1	Term Test 2	Avg.	
				75			25	25	25	100
				Laboratory Examination			Term work		Total Term work	
3	2	--	4	Oral	Practical	Oral & Practical	Laboratory Work	Tutorial / Mini project / presentation/ Journal		50
				--	--	25	15	10	25	

Prerequisite: C Programming

Objectives:

The course intends to introduce and familiarize students with data structures, their use in solving real time complex problems and implementation of these data structures. The course also aims to provide mathematical approach for analyzing algorithms using asymptotic notation and for measuring efficiency of algorithms. Finally, the course intends to make students learn various sorting and searching techniques and choose efficient one based on their efficiency.

Outcomes: On completion of the course, learner will be able to:

1. Implement various operations like searching, insertion, deletion, traversal, etc. on various data structures.
2. Choose appropriate (efficient) sorting, searching and hashing technique for given problem and implement it.
3. Choose appropriate (efficient) data structure and algorithm and apply them to solve specified problems.
4. Evaluate and analyze the efficiency of algorithms based on time and space complexity.
5. Formulate new solutions for given problems or improve existing one for better efficiency and optimization.

Detailed Syllabus: (unit wise)		
Unit	**Description**	**Duration**
1	**Review and Introduction** Review: Pointers, Structures, Function, Recursion Introduction to Data Structures: Need of Data Structures, Types of Data Structures, Abstract Data Type (ADT) Introduction to Algorithms and Analysis: Need of writing algorithm, SDLC (System Development Life Cycle) and role of algorithm, Asymptotic Notation (Big-Oh, Big Omega, Theta Notations), Order of Growth function, Complexity Analysis Techniques, Few examples of analysis of algorithms (like Fibonacci, prefix average, etc.)	**06**
2	**Linked List** Basic concept of Linked List, Comparison of sequential (array-based) and linked organizations, Dynamic Memory Management, ADT of Linked list, Singly Linked list, Doubly Linked list, Circular linked list, Various basic and advanced operations on linked list (Insertion, Deletion, Merge, Traversal, copy, reverse, etc.) and their analysis, Applications of linked list.	**06**
3	**Stack and Queue** Stacks: Introduction to Stack, Stack as an ADT, Stack ADT implementation using array and Linked List with respective analysis and comparison, Applications of Stack: Expression Conversion (Infix to prefix and postfix) and Evaluation (Postfix expression evaluation), Parenthesis Correctness, etc. Queues: Introduction to Queue, Queue as an ADT, Queue ADT implementation using array and Linked List with respective analysis and comparison, Linear Queue, Circular queue, Priority Queue: Heap based implementation, DEQueue, Applications of Queue.	**08**
4	**Trees** Introduction to Trees, Basic Terminology, Types of Trees, Binary tree representation, Traversal of binary tree, Expression Tree, Binary Search Tree, Operations on binary search tree and their analysis, AVL tree, Applications of tree.	**08**
5	**Graphs** Representation of Graph, Types of Graph, Breadth-First Search (BFS), Depth–First Search (DFS), Minimum Spanning Tree: Prim's & Kruskal's Algorithm, Applications of graphs.	**06**
6	**Sorting and Searching Techniques** Sorting: Bubble Sort, Selection Sort, Insertion Sort, Quick Sort, Merge Sort, Heap Sort, Radix sort. Analysis of Sorting Techniques Searching: Linear search, Binary search, Hashing techniques and Collision resolution techniques, Linear hashing, Hashing with chaining, Separate Chaining, Open Addressing, Rehashing, Analysis of Searching Techniques.	**08**

Suggested list of Laboratory Experiments: (At Least 12)

Note: Students are required to complete 12 experiments. At least one experiments is mandatory from each topic.

Any other experiment based on syllabus may be included, which would help the learner to understand topic/concept.

Recursion

- Implementation of recursive algorithms to solve various fundamental problems like: addition of elements in an array, reversing an array, adding all digits of a given numeral, prefix average, factorial of a given number, finding Fibonacci sequence etc.

Sorting and Searching

- Implementation of Insertion sort, Selection sort menu driven program
- Implementation of Quick Sort.
- Implementation of Merge Sort
- Implementation of Heap Sort
- Implementation of Binary search
- Implementation of hashing functions with different collision resolution techniques.

Linked List

- Implementation of Linked Lists menu driven program.
- Implementation of different operations on linked list – copy, concatenate, split, reverse, count no. of nodes etc.
- Implementation of polynomials operations (addition, subtraction) using Linked List.

Stack and Queue

- Implementation of Infix to Postfix Transformation and its evaluation program.
- Implementation of Infix to Prefix Transformation and its evaluation program.
- Implementation of double ended queue menu driven program.
- Implementation of queue menu driven program.
- Implementation of Circular queue menu driven program.
- Implementation of Priority queue program using array.
- Implementations of Linked Lists menu driven program (stack and queue).
- Implementations of Double ended queue using Linked Lists.
- Implementation of Priority Queue program using Heap

Trees

- Implementation of BT (Binary Tree) program
- Implementation of BST program.
- Implementation of various operations on tree like – copying tree, mirroring a tree, counting the number of nodes in the tree, counting only leaf nodes in the tree.
- Implementation of construction of expression tree using postfix expression

Graphs

- Implementation of Graph menu driven program (DFS & BFS).

Books Recommended:

Text Books:

1. R. F. Gilberg and B. A. Forouzan, “Data Structures – A Pseudocode Approach with C”, 2nd Edition, Cengage Learning, 2005.
2. Ellis Horowitz, Sartaj Sahni and Susan Anderson-Freed, “Fundamentals of Data Structures in C”, 2nd Edition, W. H. Freeman and Company 2008.

Reference Books:

1. Mark A. Weiss, "Data Structures and Algorithm Analysis in C", 4th Edition, Pearson, 2014.
2. M. T. Goodrich, R. Tamassia, D. Mount, "Data Structures and Algorithms in C++", Wiley, 2004.
3. Tenenbaum, Langsam, Augenstein, "Data Structures using C", Pearson, 2004.
4. Aho, Hopcroft, Ullman, "Data Structures and Algorithms", Addison-Wesley, 2010.
5. Reema Thareja, "Data Structures using C", Oxford, 2017.

Evaluation Scheme:

Semester End Examination (A):

Theory:

1. Question paper will be based on the entire syllabus summing up to 75 marks.
2. Total duration allotted for writing the paper is 3 hrs.

Continuous Assessment (B):

Theory:

1. Two term tests of 25 marks each will be conducted during the semester out of which; one will be a compulsory term test (on minimum 02 Modules) and the other can either be a term test or an assignment on live problems or a course project.
2. Total duration allotted for writing each of the paper is 1 hr.
3. Average of the marks scored in both the two tests will be considered for final grading.

Laboratory: (Term work)

Laboratory work will be based on **DJ19DSL302** with minimum 12 experiments to be incorporated.

The distribution of marks for term work shall be as follows:

I. Laboratory work (Performance of Experiments): 15 Marks

II. Journal documentation (Write-up and/or Assignments): 10 marks

The final certification and acceptance of term work will be subject to satisfactory performance of laboratory work and upon fulfilling minimum passing criteria in the term work.

Prepared by Checked by Head of the Department Principal

Second Year B.Tech. Program in Computer Science and Engineering (Data Science) Semester III

Program: Second Year B. Tech. in Computer Science and Engineering (Data Science)	**Semester: III**
Course: Foundations of Data Analysis	**Course Code:** DJ19DSC303
Course: Foundations of Data Analysis Laboratory	**Course Code:** DJ19DSL303

Teaching Scheme (Hours / week)				**Evaluation Scheme**						
				Semester End Examination Marks (A)			**Continuous Assessment Marks (B)**			**Total marks (A+ B)**
Lectures	**Practical**	**Tutorial**	**Total Credits**	**Theory**			**Term Test 1**	**Term Test 2**	**Avg.**	
				75			**25**	**25**	**25**	**100**
				Laboratory Examination			**Term work**		**Total Term work**	
3	**2**	**--**	**4**	**Oral**	**Practical**	**Oral & Practical**	**Laboratory Work**	**Tutorial / Mini project / presentation/ Journal**		**50**
				--	**--**	**25**	**15**	**10**	**25**	

Prerequisite: Basic Mathematics

Objectives:

To develop skills of data analysis techniques for data modelling.

Outcomes: On completion of the course, learner will be able:

1. Apply visualization techniques to understand Data.
2. Apply ETL and perform OLAP operation.
3. Perform various techniques to improve quality of data.
4. Apply appropriate feature engineering technique to prepare data for modelling.
5. Apply Sampling techniques to sample data for modelling.

Detailed Syllabus: (unit wise)		
Unit	**Description**	**Duration**
1.	**Data:** Data objects and attributes: nominal, binary, ordinal, numeric, discrete, continuous; Characteristics of data sets: dimensionality, sparsity, resolution Types of data sets: record data, data matrix, graph-based data, sequential data, sequence data, time series data, spatial data. **Data visualization:** Temporal: Scatter plots, Time Series sequences, Line graphs; Hierarchal: Tree diagrams, Ring charts; Network: Matrix charts, Node-link diagrams, Word clouds, Alluvial diagrams; Multidimensional: Pie chart, Venn diagrams, Stacked bar graph, Histograms; Geospatial: Flow map, Density map, Heat maps.	**06**
2.	**ETL Process and OLAP:** Major steps in ETL process, Data extraction: Techniques, Data transformation: Basic tasks, Major transformation types, Data Loading: Applying Data, OLTP Vs OLAP, OLAP definition,	**08**

	Dimensional Analysis, Hypercubes, OLAP operations: Drill down, Roll up, Slice, Dice and Rotation, OLAP models: MOLAP, ROLAP.	
3.	**Data Preprocessing** Data Quality: measurement error, data collection error, noise, artifacts, precision, bias, accuracy, outliers, missing values, inconsistent values, duplicate values Data Cleaning: handling missing values and noisy data Data Transformation: smoothing, attribute construction, aggregation, normalization; Data Discretization: binning, histogram analysis, cluster Outlier detection: types of outliers, challenges, statistical method, proximity-based method, clustering-based method	**10**
4.	**Feature Engineering** Curse of Dimensionality, Feature selection: Univariate methods (Pearson correlation, F score, Chi-square, Signal to noise ratio) and Multivariate methods (Forward selection, backward selection and stepwise selection), Feature extraction: principal component analysis	**10**
5.	**Elementary Sampling Theory:** Census and Sampling Survey, Steps in Sampling Design, Criteria of selecting a good sample procedure, Characteristics of a good sample design, Types of sample design: Non Probability and Probability sampling, Complex Random Sampling Design: Symmetric Sampling, Stratified Sampling, Cluster Sampling, Area Sampling, Sequential Sampling and Multi-stage Sampling.	**08**

Suggested list of Laboratory Experiments: (At least 8 experiments)

Visualization experiments can be performed using Tableau and Data Preprocessing experiments can be performed using Python/R.

1. Create new measures on a given dataset and visualize them using a bar graph.
2. Perform time series aggregation, apply filters on a given dataset, create line and area charts.
3. Apply maps, scatter plots on a given dataset and create a dashboard.
4. Perform joins, blends and create dual axis chart.
5. Perform table calculations, bins, distributions and create Heat maps.
6. Create an interactive data story.
7. Perform Exploratory Data Analysis on a given dataset.
8. Perform Data cleaning on a given dataset.
9. Perform necessary Data Transformation on a given dataset.
10. Perform correlation analysis on a given dataset.
11. Perform dimensionality reduction using PCA.

Any other experiment based on syllabus may be included, which would help learners to understand the topic/concept.

Books Recommended:

Text Books:

1. Data Preparation for Machine Learning, Jason Browniee, ebook by Machine Learning Mastry.
2. Best Practices in Data Cleaning: A Complete Guide to Everything you Need to Do Before and After Collecting Your Data, Jason Osborne, Sage Publication, 2012.
3. Bad Data Handbook: Cleaning Up the Data so you can get back to work, Ethan McCallum, O'Reilly, 2012
4. Feature Engineering and Selection: A practical Approach for Predictive Models, Max Kuhn and Keijell Johnson, CRC Press, 2020

Reference Books:

1. The Big Book of Dashboards: Visualizing your Data using Readl-World Business Scenarios, Jeffrey Shaffer, Steve Wexier, Andy Cotgreave, Wiley 2017
2. Research Methodology-Methods and Techniques, C. R. Kothari, New Age International, Second Edition.
3. Fundamentals of mathematical statistics, S.C.Gupta and V.K.Kapoor, second edition, Sultan Chand Publisher
4. Data Warehousing Fundamentals: A Comprehensive Guide for IT Professionals, Paulraj Ponniah, Wiley, Second Edition
5. Practical Tableau, Rayan Sleeper, O'Reilly 2018
6. Data Mining Concepts and Techniques, Han, Kamber, Morgan Kaufmann 3nd Edition.
7. Python for Data Analysis, Wes McKinney, O'Reilly, Second Edition, 2018

Evaluation Scheme:

Semester End Examination (A):

Theory:

1. Question paper will be based on the entire syllabus summing up to 75 marks.
2. Total duration allotted for writing the paper is 3 hrs.

Continuous Assessment (B):

Theory:

1. Two term tests of 25 marks each will be conducted during the semester out of which; one will be a compulsory term test (on minimum 02 Modules) and the other can either be a term test or an assignment on live problems or a course project.
2. Total duration allotted for writing each of the paper is 1 hr.
3. Average of the marks scored in both the two tests will be considered for final grading.

Laboratory: (Term work)

Laboratory work will be based on **DJ19DSL303** with minimum 8 experiments to be incorporated.

The distribution of marks for term work shall be as follows:

i. Laboratory work (Performance of Experiments): 15 Marks
ii. Journal documentation (Write-up and/or Assignments): 10 marks

The final certification and acceptance of term work will be subject to satisfactory performance of laboratory work and upon fulfilling minimum passing criteria in the term work.

Prepared by Checked by Head of the Department Principal

Second Year B.Tech. Program in Computer Science and Engineering (Data Science) Semester III

Program: Second Year B. Tech. in Computer Science and Engineering (Data Science)	**Semester: III**
Course: Database Management System	**Course Code:** DJ19DSC304
Course: Database Management System Laboratory	**Course Code:** DJ19DSL304

Teaching Scheme (Hours / week)				Evaluation Scheme						
				Semester End Examination Marks (A)			Continuous Assessment Marks (B)			Total marks (A+ B)
Lectures	Practical	Tutorial	Total Credits	Theory			Term Test 1	Term Test 2	Avg.	
				75			25	25	25	100
				Laboratory Examination			Term work		Total Term work	
3	2	--	4	Oral	Practical	Oral & Practical	Laboratory Work	Tutorial / Mini project / presentation/ Journal		25
				--	--	--	15	10	25	

Prerequisite: Computer Basics

Objectives:

The course intends to introduce the students to the management of database systems, with an emphasis on how to design, organize, maintain and retrieve information efficiently and effectively from a database.

Outcomes: On completion of the course, learner will be able:

1. Design an optimized database.
2. Create and populate a relational database and retrieve information from the database by formulating SQL queries
3. Explain the concepts of transaction, concurrency and recovery.
4. Apply indexing mechanisms for efficient retrieval of information from database.

Detailed Syllabus: (unit wise)

Unit	Description	Duration
1.	**Introduction Database Concepts:** Introduction, Characteristics of databases, File system v/s Database system, Users of Database system, Data Independence, DBMS system architecture, Database Administrator	03
2.	**Entity–Relationship Data Model** The Entity-Relationship (ER) Model: Entity types: Weak and strong entity sets, Entity sets, Types of Attributes, Keys, Relationship constraints: Cardinality and Participation Extended Entity-Relationship (EER) Model: Generalization, Specialization and Aggregation	08
3.	**Relational Model and Relational Algebra** Introduction to the Relational Model, relational schema and concept of keys, Mapping the ER and EER Model to the Relational Model Relational Algebra – unary and set operations, Relational Algebra Queries.	08

4.	**Structured Query Language (SQL)** Overview of SQL, Data Definition Commands, Data Manipulation commands, Data Control commands, Transaction Control Commands. Integrity constraints - key constraints, Domain Constraints, Referential integrity, check constraints, set and string operations, aggregate function, group by clause, having Clause Views in SQL, joins, Nested and complex queries	**09**
5	**Relational–Database Design** Pitfalls in Relational-Database designs, Concept of normalization, Functional Dependencies, First Normal Form, 2NF, 3NF, BCNF **Transactions Management and Concurrency** Transaction concept, Transaction states, ACID properties, Concurrent Executions, Serializability – Conflict and View, Concurrency Control: Lock-based, Timestamp-based protocols. **Recovery System:** Introduction to Recovery system	**10**
6.	**Indexing Mechanism** Hashing techniques, Types of Indexes: Single Level Ordered Indexes, Multilevel Indexes, Overview of BTrees and B+ Trees.	**04**

Suggested List of Experiments for Database Management System Laboratory (DJ19DSL304): (At Least Ten)

1. To draw an ER diagram for a problem statement.
2. To implement Basic SQL commands.
3. To access & modify Data using SQL.
4. To implement Joins and Views.
5. To implement Subqueries.
6. To implement Integrity Constraints.
7. To implement triggers.
8. To implement procedures, functions and cursors.
9. To simulate ARIES recovery algorithm.
10. To demonstrate export-import commands.
11. To implement B-trees/B+ trees.

Any other experiment based on syllabus may be included, which would help the learner to understand topic/concept.

Books Recommended:

Text books:

1. Database System Concepts, Korth,Silberchatz,Sudarshan, 6th Edition, McGraw – Hill
2. Fundamentals of Database Systems, Elmasri and Navathe, 5th Edition, Pearson education.
3. Database Systems Design, Implementation and Management, Peter Rob and Carlos Coronel, Thomson Learning, 5th Edition.
4. Distributed Database System, Chhanda Ray, Pearson Education India.
5. Database Management Systems, G. K. Gupta, McGraw – Hill.

Reference Books:

1. SQL and PL/SQL for Oracle 10g, Black Book, Dr. P.S. Deshpande, Dreamtech Press.
2. Introduction to Database Management‖, Gillenson, Paulraj Ponniah, Wiley Publication.
3. Database Management Systems, Raghu Ramkrishnan and Johannes Gehrke,TMH.

4. Principles of Distributed Database, M. Tamer Ozsu, Patrick Valduriez, PearsonEducation India.

Evaluation Scheme:

Semester End Examination (A):

Theory:

1. Question paper will be based on the entire syllabus summing up to 75 marks.
2. Total duration allotted for writing the paper is 3 hrs.

Continuous Assessment (B):

Theory:

1. Two term tests of 25 marks each will be conducted during the semester out of which; one will be a compulsory term test (on minimum 02 Modules) and the other can either be a term test or an assignment on live problems or a course project.
2. Total duration allotted for writing each of the paper is 1 hr.
3. Average of the marks scored in both the two tests will be considered for final grading.

Laboratory: (Term work)

Laboratory work will be based on **DJ19DSL304** with minimum 10 experiments to be incorporated.

The distribution of marks for term work shall be as follows:

i Laboratory work (Performance of Experiments): 15 Marks

ii Journal documentation (Write-up and/or Assignments): 10 marks

The final certification and acceptance of term work will be subject to satisfactory performance of laboratory work and upon fulfilling minimum passing criteria in the term work.

Prepared by Checked by Head of the Department Principal

Second Year B.Tech. Program in Computer Science and Engineering (Data Science) Semester III

Program: Second Year B. Tech. in Computer Science and Engineering (Data Science)	**Semester: III**
Course: Statistics for Data Science	**Course Code:** DJ19DSC305
Course: Statistics for Data Science Laboratory	**Course Code:** DJ19DSL305

Teaching Scheme (Hours / week)				Evaluation Scheme						
				Semester End Examination Marks (A)			Continuous Assessment Marks (B)			Total marks (A+ B)
Lectures	Practical	Tutorial	Total Credits	Theory			Term Test 1	Term Test 2	Avg.	
				75			25	25	25	100
				Laboratory Examination			Term work			
3	2	--	4	Oral	Practical	Oral & Practical	Laboratory Work	Tutorial / Mini project / presentation/ Journal	Total Term work	25
				--	--	--	15	10	25	

Prerequisite: Probability, Probability distribution

Objectives:
To build the strong foundation in statistics which can be applied to analyze data and make predictions.

Outcomes: On completion of the course, learner will be able to:
1. Interpret data using descriptive statistics
2. Demonstrate sampling distributions and estimate statistical parameters
3. Develop hypothesis based on data and perform testing using various statistical techniques.
4. Perform analysis of variance on data
5. Examine relations between data

Detailed Syllabus: (unit wise)

Unit	Description	Duration
1.	**Introduction to Statistics** Types of statistics, population vs sample Measures of Central Tendency: arithmetic mean, properties, weighted mean, properties, median, mode, grouped and ungrouped data, empirical relation between the mean, median and mode, geometric mean, harmonic mean, relation between arithmetic, geometric and harmonic mean, outlier Measures of dispersion: range, quartile deviation, mean deviation, standard deviation, properties, variance, root mean square deviation, empirical relations between measures of dispersion, absolute and relative dispersion, coefficient of variation, moments, Pearson's β and γ coefficients, skewness, kurtosis, population parameters and sample statistics, histogram, frequency polygon Measures of position: quartiles, interquartile range, semi interquartile range, percentiles, percentile rank, 10–90 percentile range, box and whisker plot	08

2.	**Sampling distribution and Estimation:** Sampling distribution: Central limit theorem, population distribution, chi-square distribution, Z - distribution, student's t-distribution, F-Distribution Statistical Estimation: Characteristics of estimators, consistency, unbiasedness, unbiased estimates, efficient estimates, sufficient estimators, point estimates, interval estimates, determination of sample size for estimating mean and proportions, estimates of population parameters, probable error	**07**
3.	**Hypothesis Testing for data driven decision making:** Hypothesis testing: Test of significance, null and alternative hypothesis, type I and type II error, factors affecting Type II error, probability of Type II error, power of test, p Value, critical region, level of significance Confidence interval: Population mean, difference between two population means, population proportion, difference between two population proportions, variance, ratio of variances of two populations Goodness of fit test using Kolmogorov-Smirnov test and Anderson Darling test Tests using z-statistics: difference between sample proportion and population proportion, difference between two sample proportion, difference between sample mean and population mean with known σ and unknown σ, difference between two sample means, one tailed and two tailed tests Test using t-statistics: difference between sample mean and population mean, difference between two independent sample means, difference between means from the same group Test using F-statistics: equality of population variance Test using chi-square statistics: test of independence, goodness of fit	**12**
4.	**Analysis of Variance (ANOVA) for data analysis** Sample size calculation, one way ANOVA, POST-HOC Analysis (Tukey's Test), randomized block design, two way ANOVA	**07**
5	**Examining Relationship** Correlation: Scatter plot, covariance, Karl Pearson's coefficient of correlation, hypothesis test for correlation, correlation vs causation, extreme data values, limits of correlation coefficient, Rank correlation, Spearman's rank correlation coefficient, Repeated ranks, partial and multi correlation Regression: linear regression analysis, lines of regression, regression coefficients, scatter plot with regression lines, hypothesis test for regression, multiple regression, coefficient of determination, residuals, collinearity, influential observations	**08**

List of Laboratory Experiments: (any eight using Python)

1. To perform descriptive statistics on data.
2. To visualize descriptive statistics on data.
3. To prove central limit theorem.
4. To study sampling distributions and their parameters
5. To perform statistical estimation tests on data
6. To calculate confidence interval for different parameters
7. To perform goodness of fit using Kolmogorov-Smirnov test and Anderson Darling test
8. To perform hypothesis test using Z statistics
9. To perform hypothesis test using t statistics
10. To perform hypothesis test using F statistics
11. To perform hypothesis test using Chi square
12. To perform ANOVA on given data
13. To perform correlation on given data
14. To perform regression on given data regression and evaluate the model

Any other experiment based on syllabus may be included, which would help the learner to understand topic/concept.

Books Recommended:

Text books:

1. An Introduction to Statistics with Python, Thomas Hasalwanter, Springer, 2016
2. Think Stats: Probability and Statistics for Programmers, Allen B. Downey, Green Tea Press, 2011
3. Testing Statistical Hypotheses, E. L. Lehmann, Joseph P. Romano, Springer, 2008, third edition.
4. Statistical Methods, S. P. Gupta, Sultan Chand, 2014, forty third edition

Reference Books:

1. Practical Statistics for data scientists 50+ Essential Concepts Using R and Python, Peter Bruce, Andrew Bruce, Peter Gedeck, Orelly, second edition, 2020
2. Statistics, Freedman, David, Robert Pisani, Roger Pervis, W. W. Norton, 2007
3. Fundamentals of mathematical statistics, S. C. Gupta, V. K. Kapoor, Sultan Chand, 2002, tenth edition

Evaluation Scheme:

Semester End Examination (A):

Theory:

1. Question paper will be based on the entire syllabus summing up to 75 marks.
2. Total duration allotted for writing the paper is 3 hrs.

Continuous Assessment (B):

Theory:

1. Two term tests of 25 marks each will be conducted during the semester out of which; one will be a compulsory term test (on minimum 02 Modules) and the other can either be a term test or an assignment on live problems or a course project.
2. Total duration allotted for writing each of the paper is 1 hr.
3. Average of the marks scored in both the two tests will be considered for final grading.

Laboratory: (Term work)

Term work shall consist of minimum 8 experiments.

The distribution of marks for term work shall be as follows:

i. Laboratory work (Performance of Experiments): 15 Marks
ii. Journal documentation (Write-up and/or Assignments): 10 marks

The final certification and acceptance of term work will be subject to satisfactory performance of laboratory work and upon fulfilling minimum passing criteria in the term work.

Prepared by Checked by Head of the Department Principal

Second Year B.Tech. Program in Computer Science and Engineering (Data Science) Semester III

Program: Second Year B. Tech. in Computer Science and Engineering (Data Science)	**Semester: III**
Course: Programming with Python Laboratory	**Course Code:** DJ19DSL306

Teaching Scheme (Hours / week)				**Evaluation Scheme**						
				Semester End Examination Marks (A)			**Continuous Assessment Marks (B)**			**Total marks (A+ B)**
Lectures	**Practical**	**Tutorial**	**Total Credits**	**Theory**			**Term Test 1**	**Term Test 2**	**Avg.**	
				--			--	--	--	--
				Laboratory Examination			**Term work**		**Total Term work**	
-	4	--	2	**Oral**	**Practical**	**Oral & Practical**	**Laboratory Work**	**Tutorial / Mini project / presentation/ Journal**		**75**
				--	--	25	25	25	50	

Prerequisite: Programming Fundamentals

Objectives:

1. To learn the basic and OOP concepts of Python.
2. To study various advance python concept like inheritance, exception handling, modules etc.
3. Learn to develop GUI based standalone and web application.

Outcomes: On completion of the course, learner will be able:

1. Demonstrate basic data types and data structures in python.
2. Demonstrate the concepts of Object-oriented programming.
3. Apply file, directory handling and text processing concepts in python.
4. Apply database connectivity, client-server communication using python. 5.
5. Apply various advance modules of Python for data analysis.

Detailed Syllabus: (unit wise)		
Unit	**Description**	**Duration**
1.	**Python basics:** Data types in python, Operators in python, Input and Output, Control statement, Arrays in python, String and Character in python, Functions, List and Tuples, Dictionaries	**10**
2.	**Introduction to OOP** Classes, Objects, Constructor, Methods, Abstraction, Inheritance, Magic Methods, Exception Handling	**12**
3.	**Advanced Python** Building Modules, Packages: Python Collections Module, Opening and Reading Files and Folders (Python OS Module, Python Datetime Module, Python Math and Random Modules Text Processing & Regular expression in python	**12**
4.	**Python Integration Primer**	**10**

	Graphical User interface using Tkinter: Form designing Networking in Python: Client Server socket programming, Python database connectivity using SQL lite	
5	**Python advance Modules** Numpy: Working with Numpy, Constructing Numpy arrays, Printing arrays, Arithmetic Operations on matrix's, numpy zeros() Matplotlib: Matplotlib-Installation & Sample code, Bar Chart Pandas: Data Processing, Pandas-Data structure, Pandas-Series data, Data Frames	**12**

Suggested List of Laboratory Experiments:

1. Exploring basics of python like data types (strings, list, array, dictionaries, set, tuples) and control statements.
2. Demonstrate the concept of Abstraction in Python.
3. Demonstrate the concept of Inheritance
4. Demonstrate exception handling.
5. Python program to explore different types of Modules
6. Exploring Files and directories
    a) Python program to append data to existing file and then display the entire file.
    b) Python program to count number of lines, words and characters in a file.
    c) Python program to display file available in current directory
6. Make use of RE module to do text processing.
7. Creating GUI with python containing widgets such as labels, textbox, radio, checkboxes and custom dialog boxes.
8. Program to demonstrate CRUD (create, read, update and delete) operations on database (SQLite/ MySQL) using python.
9. Creation of simple socket for basic information exchange between server and client.
10. Make use of advance modules of Python like Matplotlib, Numpy, Pandas

Any other experiment based on syllabus may be included, which would help the learner to understand topic/concept.

Books Recommended:

Text Books:

1. Learn Python the Hard Way, Zed Shaw's Hard Way Series, third edition
2. Python Projects, Laura Cassell, Alan Gauld, wrox publication

Digital resources:

1. The Python Tutorial", http://docs.python.org/release/3.0.1/tutorial/
2. http://spoken-tutorial.org
3. www.staredusolutions.org

Evaluation Scheme:

Practical and Oral(A):

Oral & Practical examination will be based on the practical's performed during laboratory sessions.

1. *Implementation:15 Marks*
2. *Oral:10 Marks*

Total :25 Marks

Continuous Assessment (B):

Term Work:

Laboratory work will be based on syllabus with minimum 10 experiments to be incorporated. Experiments should be completed by students in the given time duration.

The final certification and acceptance of term work will be subject to satisfactory performance of laboratory work and upon fulfilling minimum passing criteria in the term work.

1. *Laboratory work (Performance of Experiments): 15 Marks*
2. *Mini Project:10 Marks*

   Total: 25 Marks

Prepared by Checked by Head of the Department Principal

Second Year B.Tech. Program in Computer Science and Engineering (Data Science) Semester III

<table>
<tr><td colspan="8">Program: Common for All Programs</td><td colspan="3">Semester: III & IV Combined</td></tr>
<tr><td colspan="8">Course: Innovative Product Development-I</td><td colspan="3">Course Code: DJ19A2</td></tr>
<tr><td colspan="4" rowspan="2">Teaching Scheme (Hours/week)</td><td colspan="7">Evaluation Scheme</td></tr>
<tr><td colspan="4">Semester End Examination Marks (A)</td><td colspan="3">Continuous Assessment Marks (B)</td><td rowspan="2">Total marks (A+ B)</td></tr>
<tr><td rowspan="3">Lectures</td><td rowspan="3">Practical</td><td rowspan="3">Tutorial</td><td rowspan="3">Total Credits</td><td colspan="4">Theory</td><td>Term Test 1</td><td>Term Test 2</td><td>Avg.</td></tr>
<tr><td colspan="4">--</td><td>--</td><td>--</td><td>--</td><td>--</td></tr>
<tr><td colspan="4">Laboratory Examination</td><td colspan="2">Semester review</td><td rowspan="3">Total</td><td rowspan="3">100</td></tr>
<tr><td rowspan="2">--</td><td rowspan="2">2</td><td rowspan="2">--</td><td rowspan="2">--</td><td>Oral</td><td>Practical</td><td>Oral & Practical</td><td>Review 1</td><td>Review 2</td></tr>
<tr><td>--</td><td>--</td><td>--</td><td>50</td><td>50</td><td>100</td></tr>
</table>

Objectives:

1. To acquaint the students with the process of identifying the need (considering a societal requirement) and ensuring that a solution is found out to address the same by designing and developing an innovative product.
2. To familiarize the students with the process of designing and developing a product, while they work as part of a team.
3. To acquaint the students with the process of applying basic engineering fundamentals, so as to attempt at the design and development of a successful value added product.
4. To inculcate the basic concepts of entrepreneurship and the process of self-learning and research required to conceptualize and create a successful product.

Outcome:

Learner will be able to:

1. Identify the requirement for a product based on societal/research needs.
2. Apply knowledge and skills required to solve a societal need by conceptualising a product, especially while working in a team.
3. Use standard norms of engineering concepts/practices in the design and development of an innovative product.
4. Draw proper inferences through theoretical/ experimental/simulations and analyse the impact of the proposed method of design and development of the product.
5. Develop interpersonal skills, while working as a member of theteam or as theleader.
6. Demonstrate capabilities of self-learning as part of the team, leading to life-long learning, which could eventually prepare themselves to be successful entrepreneurs.
7. Demonstrate product/project management principles during the design and development work and also excel in written (Technical paper preparation) as well as oral communication.

Guidelines for the proposed product design and development:

- Students shall form a team of 3 to 4 students (max allowed: 5-6 in extraordinary cases, subject to the approval of the department review committee and the Head of the department).
- Students should carry out a survey and identify the need, which shall be converted into conceptualisation of a product, in consultation with the faculty supervisor/head of department/internal committee of faculty members.
- Students in the team shall understand the effective need for product development and accordingly select the best possible design in consultation with the faculty supervisor.
- Students shall convert the best design solution into a working model, using various components drawn from their domain as well as related interdisciplinary areas.
- Faculty supervisor may provide inputs to students during the entire span of the activity, spread over 2 semesters, wherein the main focus shall be on self-learning.
- A record in the form of an activity log-book is to be prepared by each team, wherein the team can record weekly progress of work. The guide/supervisor should verify the recorded notes/comments and approve the same on a weekly basis.
- The design solution is to be validated with proper justification and the report is to be compiled in a standard format and submitted to the department. Efforts are to be made by the students to try and publish a technical paper, either in the institute journal, "Techno Focus: Journal for Budding Engineers" or at a suitable publication, approved by the department research committee/ Head of the department.
- The focus should be on self-learning, capability to design and innovate new products as well as on developing the ability to address societal problems. Advancement of entrepreneurial capabilities and quality development of the students through the year long course should ensure that the design and development of a product of appropriate level and quality is carried out, spread over two semesters, i.e. during the semesters III and IV.

Guidelines for Assessment of the work:

- The review/ progress monitoring committee shall be constituted by the Head of the Department. The progress of design and development of the product is to be evaluated on a continuous basis, holding a minimum of two reviews in each semester.
- In the continuous assessment, focus shall also be on each individual student's contribution tothe team activity, their understanding and involvement as well as responses to the questionsbeing raised at all points in time.
- Distribution of marks individually for the both reviews as well as for the first review during the subsequent semester shall be as given below:
    - Marks awarded by the supervisor based on log-book : 20
    - Marks awarded by review committee : 20
    - Quality of the write-up : 10

In the last review of the semester IV, the marks will be awarded as follows.

- Marks awarded by the supervisor (Considering technical paper writing) : 30
- Marks awarded by the review committee : 20

Note- A Candidate needs to secure a minimum of 50% marks to be declared to have completed the audit course.

Review/progress monitoring committee may consider the following points during the assessment.

- In the semester III, the entire design proposal shall be ready, including components/system selection as well as the cost analysis. Two reviews will be conducted based on the presentation given by the student'steam.
    - First shall be for finalisation of the product selected.
    - Second shall be on finalisation of the proposed design of the product.
- In the semester IV, the expected work shall be procurement of components/systems, building of the working prototype, testing and validation of the results based on work completed in semester III.
    - First review is based on readiness of building the working prototype.
    - Second review shall be based on a presentation as well as the demonstration of the working model, during the last month of semester IV. This review will also look at the readiness of the proposed technical paper presentation of the team.

The overall work done by the team shall be assessed based on the following criteria;

1. Quality of survey/ need identification of the product.
2. Clarity of Problem definition (design and development) based on need.
3. Innovativeness in the proposed design.
4. Feasibility of the proposed design and selection of the best solution.
5. Cost effectiveness of the product.
6. Societal impact of the product.
7. Functioning of the working model as per stated requirements.
8. Effective use of standard engineering norms.
9. Contribution of each individual as a member or the team leader.
10. Clarity on the write-up and the technical paper prepared.

- The semester reviews (III and IV) may be based on relevant points listed above, as applicable.

Guidelines for Assessment of Semester Reviews:

- The write-up should be prepared as per the guidelines given by the department.
- The design and the development of the product shall be assessed through a presentation and demonstration of the working model by the student team to a panel of Internal and External Examiners, preferably from industry or any research organisations having an experience of more than five years, approved by the Head of the Institution. The presence of the external examiner is desirable only for the 2nd presentation in semester IV.Students are compulsorily required to present the outline of the technical paper prepared by them during the final review in semester IV.

Prepared by Checked by Head of the Department Principal

Second Year B.Tech. Program in Computer Science and Engineering (Data Science) Semester III

Program: Common for All Programs	**Semester** : III
Course : Constitution of India	**Course Code** : DJ19A3

Teaching Scheme (Hours / week)				**Evaluation Scheme**					
				Semester End Examination Marks (A)			**Continuous Assessment Marks (B)**			**Total marks (A+ B)**
Lectures	**Practical**	**Tutorial**	**Total Credits**	**Theory**			**TT1**	**TT2**	**Avg.**	
				-			-	-	-	-
				Laboratory Examination			**Term Work**			
01	-	-	-	**Oral**	**Practical**	**Oral & Practical**	-			-
				-	-	-				

Objectives:

1. To provide basic information about Indian constitution.
2. To identify individual role and ethical responsibility towards society.
3. To understand human rights and its implications.

Outcomes: On completion of the course, learner will be able to

1. Have general knowledge and legal literacy and thereby to take up competitive examinations.
2. Understand state and central policies, fundamental duties.
3. Understand Electoral Process, special provisions.
4. Understand powers and functions of Municipalities, Panchayat's and Co- operative Societies,
5. Understand Engineering ethics and responsibilities of Engineers
6. Understand Engineering Integrity & Reliability

Detailed Syllabus (unit wise)

Unit	**Description**	**Duration**
1	**Introduction to the Constitution of India** The Making of the Constitution and Salient features of the Constitution. Preamble to the Indian Constitution Fundamental Rights & its limitations.	2
2	**Directive Principles of State Policy:** Relevance of Directive Principles State Policy Fundamental Duties. Union Executives – President, Prime Minister Parliament Supreme Court of India.	3
3	**State Executives:** Governor, Chief Minister, State Legislature High Court of State. Electoral Process in India, Amendment Procedures, 42nd, 44th, 74th, 76th, 86th & 91st Amendments.	3

4	**Special Provision**s: For SC & ST Special Provision for Women, Children & Backward Classes, Emergency Provisions. **Human Rights:** Meaning and Definitions, Legislation Specific Themes in Human Rights- Working of National Human Rights Commission in India Powers and functions of Municipalities, Panchyats and Co – Operative Societies.	3
5	**Scope & Aims of Engineering Ethics:** Responsibility of Engineers Impediments to Responsibility. Risks, Safety and liability of Engineers, Honesty, Integrity & Reliability in Engineering	3

Books Recommended :

Text books:

1. Durga Das Basu: "Introduction to the Constitution on India", (Students Edn.) Prentice –Hall EEE, 19th / 20th Edn., 2001
2. Charles E. Haries, Michael S Pritchard and Michael J. Robins "Engineering Ethics" Thompson Asia, 2003-08-05.

Reference Books:

1. M.V.Pylee, "An Introduction to Constitution of India", Vikas Publishing, 2002.
2. M.Govindarajan, S.Natarajan, V.S.Senthilkumar, "Engineering Ethics", Prentice – Hall of India Pvt. Ltd. New Delhi, 2004
3. Brij Kishore Sharma," Introduction to the Constitution of India", PHI Learning Pvt. Ltd., New Delhi, 2011.
4. Latest Publications of Indian Institute of Human Rights, New Delhi

Web Resources :

1. www.nptel.ac.in
2. www.hnlu.ac.in
3. www.nspe.org
4. www.preservearticles.com

Prepared by Checked by Head of the Department Principal