LEARNING OUTCOME BASED CURRICULUM FRAMEWORK

[LOCF]

Sanskar Sarjan Education Society's

DTSS COLLEGE OF COMMERCE

[**AUTONOMOUS**]

PROGRAMME CODE: BDS0021

Bachelor of Science in Data Science

[B. Sc. In Data Science]

w. e. f. 2021-22

PROGRAMME STRUCTURE

1. **Title of the Program** : B.Sc. in Data Science
2. **Programme Code** : BDS0021
3. **Introduction of the Program**

   The B.Sc. Data Science is designed to provide in-depth knowledge of data techniques, and their applications for effective decision making in improving processes in various sectors. B.Sc. Data Science is a three-year integrated undergraduate inter-disciplinary course spread over six semesters. The program curriculum has been designed with feedback from industry and academia. The main attraction of the program is live projects in the third year of the course where students will explore their knowledge while working on real-world organizational problems.
4. **Program Objectives**
   a. To explain idea of data analysis techniques and quantitative modeling for the solution of real world business problems.
   b. To report findings of analysis and effectively present them using data visualization techniques.
   c. To demonstrate knowledge of statistical data analysis techniques utilized in business decision making.
   d. To provide insights about the roles of a Data Scientist, such as a developer, an analyst, a statistical expert etc.
   e. To understand techniques and tools for transformation of data, Data Mining, Data formats, Machine Learning Algorithms, Data Visualization and Optimization.
5. **System : CHOICE BASED CREDIT SYSTEM [CBCS]**
6. **Duration of the Program** : 3 years
7. **Total No of Semesters** : 6
8. **Eligibility for Admission :** An applicant must have a minimum 45% marks aggregate (open) and 40% marks aggregate (for reserved category) (with Mathematics)
9. **Intake capacity :** 60
10. **Total Credits :** 132 [Theory: 64, Practical:68]
11. **Types of Courses:**

Course Type	Total Courses (Sem I to VI)
a. Core Courses	16
b. Elective Courses	3 to be selected out of 6
c. Ability Enhancement Courses	04
d. Skill Based Courses	06
e. Multi-disciplinary / inter-disciplinary courses	03
f. Practicals	27
g. Hands on Projects	02

12. Fee Structure :

13. Teacher's Qualification: Post Graduate in Information Technology/Computer Science or Equivalent one from Recognized University, NET /SET or Ph.D.

14. Per week Work-load of the Teachers :

a. **Theory :** 80 periods per week

b. **Practical :** 45 periods per week

c. **Total :** 125 periods per week

15. Total Courses:

Semester	Category of Course	No of Courses	Credits Allotted
I	A. Core Courses (Theory)	3	6
	B. Elective Courses (Theory)	-	-
	C. Skill Based (Theory)	1	2
	D. Ability Enhancement (Theory)	1	2
	E. Inter-Disciplinary (Theory)	1	2
	F. Practical	5	10
II	A. Core Courses (Theory)	3	6
	B. Elective Courses (Theory)	-	-
	C. Skill Based (Theory)	1	2

	D. Ability Enhancement (Theory)	1	2
	E. Inter-Disciplinary (Theory)	1	2
	F. Practical	5	10
III	A. Core Courses (Theory)	3	6
	B. Elective Courses (Theory)	1	2
	C. Skill Based (Theory)	1	2
	D. Ability Enhancement (Theory)	1	2
	E. Inter-Disciplinary (Theory)	-	-
	F. Practical	5	10
IV	A. Core Courses (Theory)	3	6
	B. Elective Courses (Theory)	-	-
	C. Skill Based (Theory)	1	2
	D. Ability Enhancement (Theory)	1	2
	E. Inter-Disciplinary (Theory)	1	2
	F. Practical	5	10
V	A. Core Courses (Theory)	2	4
	B. Elective Courses (Theory)	1	2
	C. Skill/Ability Enhancement (Theory)	1	2
	D. Inter-Disciplinary (Theory)	-	-
	E. Practical	4	8
	F. Hands on Project	1	6
VI	A. Core Courses (Theory)	2	4
	B. Elective Courses (Theory)	1	2
	C. Skill/Ability Enhancement (Theory)	1	2
	D. Inter-Disciplinary (Theory)	-	-
	E. Practical	4	8
	F. Hands on Project	1	6
	Theory Credits : **64 Practical** Credits : **68**	**Total Credits**	**132**

16. Evaluation Pattern

a. Total Marks : 5000

Sem-I	Sem-II	Sem-III	Sem-IV	Sem-V	Sem-VI	Total
850	850	850	850	800	800	5000

b. Passing Criteria : 40% in theory as well as in practical.

c. **Marking Scheme** 60:40

d. **Mode of Evaluation of Answer-book :** Offline

e. Paper Pattern

[A]	**Evaluation Scheme for Theory Courses**	**Marks**
	I. Continuous Internal Assessment (C.I.A.) 1 C.I.A.-I: Test – 20 Marks of 40 mins. duration 2 C.I.A.-II: Assignment/Presentation/Project -15 Marks Active Participation/Attendance - 05 Marks	**40 Marks**
	II. Semester End Examination (SEE)	**60 Marks**
Q. 1	Objective/Short Answer (Covering the whole Syllabus)	10 Marks
Q. 2	Answer any two Descriptive	10 Marks
Q. 3	Answer any two Descriptive	10 Marks
Q. 4	Answer any two Descriptive	10 Marks
Q. 5	Answer any two Descriptive	10 Marks
Q. 6	Answer any two Descriptive	10 Marks
[B]	**Evaluation Scheme for Practical Courses**	**50 Marks**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

17. Programme Outcome

After completion of the course the participants will be able to –

PO1: Use proper techniques for understanding and presenting data.

PO2: Learn how to use optimization techniques and software tools for data analysis.

PO3: Sharpen their analytical skills and develop the ability of analyzing data properly.

PO4: Develop strong analytical and problem-solving skills.

Semester - I				
Course Code	**Course Type**	**Course Title**	**Credits**	**Marks**
BDSCC101	Core Course	Discrete Mathematical Structure	2	100
BDSCC102	Core Course	Algebra	2	100
BDSCC103	Core Course	Operating Systems	2	100
BDSSB104	Skill Based Course	Intro. to software fundamentals	2	100
BDSAE105	Ability Enhancement Course	Communication Skills	2	100
BDSID106	Inter-disciplinary Course	Green Computing	2	100
BDSCCP101	Core Course Practical	Discrete Mathematical Structure Practical	2	50
BDSCCP102	Core Course Practical	Introduction to SPSS Practical	2	50
BDSCCP103	Core Course Practical	Operating Systems Practical	2	50
BDSSBP104	Skill Based Course Practical	Intro. to software fundamentals Practical	2	50
BDSAEP105	Ability Enhancement Course Practical	Communication Skills Practical	2	50
BDSIDP106	Inter-disciplinary Course Practical	-	-	-
		Total Credits	**22**	**850**

Semester - II				
Course Code	**Course Type**	**Course Title**	**Credits**	**Marks**
BDSCC201	Core Course	Probability	2	100
BDSCC202	Core Course	Linux	2	100
BDSCC203	Core Course	Algorithms and Data structures	2	100
BDSSB204	Skill Based Course	Introduction to python	2	100
BDSAE205	Ability Enhancement Course	Introduction to Statistics	2	100
BDSID206	Inter-disciplinary Course	Environmental Science	2	100
BDSCCP201	Core Course Practical	Probability Practical	2	50
BDSCCP202	Core Course Practical	Linux Practical	2	50
BDSCCP203	Core Course Practical	Algorithms and Data structures Practical	2	50
BDSSBP204	Skill Based Course Practical	Introduction to python Practical	2	50
BDSAEP205	Ability Enhancement Course Practical	Introduction to Statistics Practical	2	50
BDSIDP206	Inter-disciplinary Course Practical	-	-	-
		Total Credits	**22**	**850**

Semester - III				
Course Code	**Course Type**	**Course Title**	**Credits**	**Marks**
BDSCC301	Core Course	Database systems	2	100
BDSCC302	Core Course	Sampling Methods	2	100
BDSCC303	Core Course	Data visualization	2	100
BDSSB304	Skill Based Course	Data Analysis in Excel	2	100
BDSAE305	Ability Enhancement Course	Introduction to R/R studio	2	100
BDSEL306 BDSEL307	Elective Course	1. Data Warehouse 2. Optimization Techniques	2	100
BDSCCP301	Core Course Practical	Database systems Practical	2	50
BDSCCP302	Core Course Practical	Sampling Methods Practical	-	-
BDSCCP303	Core Course Practical	Data visualization Practical	2	50
BDSSBP304	Skill Based Course Practical	Data Analysis in Excel Practical	2	50
BDSAEP305	Ability Enhancement Course Practical	Introduction to R/R studio Practical	2	50
BDSELP306 BDSELP307	Elective Course Practical	1. Data Warehouse 2. Optimization Techniques	2	50
		Total Credits	**22**	**850**

Semester - IV				
Course Code	**Course Type**	**Course Title**	**Credits**	**Marks**
BDSCC401	Core Course	Introduction to Data Science	2	100
BDSCC402	Core Course	Artificial Intelligence	2	100
BDSCC403	Core Course	Statistical Consulting	2	100
BDSSB404	Skill Based	Internet of Things	2	100
BDSAE405	Ability Enhancement	Statistical computing (R)	2	100
BDSID406	Inter-disciplinary	Cyber Law	2	100
BDSCCP401	Core Course Practical	Introduction to Data Science Practical	2	50
BDSCCP402	Core Course Practical	Artificial Intelligence Practical	2	50
BDSCCP403	Core Course Practical	Statistical Consulting Practical	2	50
BDSSBP404	Skill Based Practical	Internet of Things Practical	2	50
BDSAEP405	Ability Enhancement Practical	Statistical computing (R) Practical	2	50
BDSIDP406	Inter-disciplinary Practical	-	-	-
		Total Credits	**22**	**850**

Semester - V				
Course Code	**Course Type**	**Course Title**	**Credits**	**Marks**
BDSCC501	Core Course	Soft Computing	2	100
BDSCC502	Core Course	Statistical Modelling in Python	2	100
BDSSB503	Skill Based Course	Advances techniques in Data Science	2	100
BDSAE504	Ability Enhancement Course	Capstone Project 1 (R)	3	100
BDSEL505 BDSEL506	Elective Course	1. Cloud Computing 2. Big Data	2	100
BDSCCP501	Core Course Practical	Soft Computing Practical	2	50
BDSCCP502	Core Course Practical	Statistical Modelling in Python	2	50
BDSSBP503	Skill Based Course Practical	Advances techniques in Data Science	2	50
BDSAEP504	Ability Enhancement Course Practical	Capstone Project 1 (R)	3	100
BDSELP505 BDSELP506	Elective Course Practical	1. Cloud Computing Practical 2. Big Data Practical	2	50
		Total Credits	**22**	**800**

Semester - VI				
Course Code	**Course Type**	**Course Title**	**Credits**	**Marks**
BDSCC601	Core Course	Data mining	2	100
BDSCC602	Core Course	Business Intelligence	2	100
BDSSB603	Skill Based Course	Machine Learning in Python	2	100
BDSAE604	Ability Enhancement Course	Capstone Project 2 (python)	3	100
BDSEL605 BDSEL606	Elective Course	1. Data Compression 2. SQA	2	100
BDSCCP601	Core Course Practical	Data Mining Practical	2	50
BDSCCP602	Core Course Practical	Business Intelligence Practical	2	50
BDSSBP603	Skill Based Course Practical	Machine Learning in Python Practical	2	50
BDSAEP604	Ability Enhancement Course Practical	Capstone Project 2 (Python) Practical	3	100
BDSELP605 BDSELP606	Elective Course Practical	1. Data Compression Practical 2. SQA Practical	2	50
		Total Credits	**22**	**800**

Bachelor of Science in Data Science

[**B. Sc. In Data Science**]

Semester - I

COURSE STRUCTURE

1. **Title of the Course :** Discrete mathematics
2. **Semester** : I
3. **Course Code: For Theory :** BITCC101
   For Practical: BITCCP101
4. **Course Objective:**

1. Mathematical reasoning: Students are expected to use mathematical reasoning in order to read, comprehend, and construct mathematical arguments. Students will learn basic concepts of mathematical logic and proof.

2. Combinatorial analysis: Students will count or enumerate objects and perform combinatorial analysis.

3. Discrete structures: Students will learn the basic concepts of sets, permutations, relations, graphs, trees and finite state machines. Students will represent discrete objects and relationships using abstract mathematical structures.

4. Algorithmic thinking: Students will verify whether an algorithm works well and perform analysis in terms of memory and time.

5. Applications and modeling: Discrete mathematics has been used in numerous applications. Students will formulate and model problems with the concepts and techniques of discrete mathematics.

5.Category of Course : Core Course

6.Total Hours: 60

7.Total Credits: 04 Credits (02 Credits for Theory & 02 Credits for Practical)

8.Evaluation Pattern:

- **Total Marks** : 150 Marks (10 Point Grading)
- **Passing Criteria** : 40 % (4 Grade Points)
- **Marking Scheme** : 60:40:50 Pattern
  - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
  - 40 Marks - Internal Assessment (Passing = 16 Marks)
  - 50 Marks - Practical Assessment (Passing = 20 Marks)
- **Mode of Evaluation of Answer-books** : Online/Offline

9. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BIT10	Discrete Mathematics	5	3	2	2	4

Module	Detailed Content	Hours
1	**Introduction:** Variables, the Language of Sets, the Language of Relations and Function Set Theory: Definitions and the Element Method of Proof, Properties of Sets, Venn diagram, Cartesian product, the principle of inclusion exclusion, the principle of inclusion exclusion using Venn diagram Disproof, Algebraic Proofs, Boolean Algebras, and Russell's Paradox and the Halting Problem. **The Logic of Compound Statements**: Logical Form and Logical Equivalence, Conditional Statements, Valid and Invalid Arguments	12
2	**Quantified Statements**: Predicates and Quantified Statements, Statements with Multiple Quantifiers, Arguments with Quantified Statements **Elementary Number Theory and Methods of Proof**: Introduction to Direct Proofs, Rational Numbers, Divisibility, Division into Cases and the Quotient-Remainder Theorem, Floor and Ceiling, Indirect Argument: Contradiction and Contraposition, Two Classical Theorems, Applications in algorithms.	12
3	**Sequences, Mathematical Induction, and Recursion:** Sequences, Mathematical Induction, Strong Mathematical Induction and the Well- Ordering Principle for the Integers, Correctness of algorithms, defining sequences recursively, solving recurrence relations by iteration, Second order linear homogenous recurrence relations with constant coefficients. General Recursive definitions and structural induction. **Functions**: Functions Defined on General Sets, One-to-One and Onto, Inverse Functions, Composition of Functions, Cardinality with Applications to Computability	12
4	**Relations**: Relations on Sets, Reflexivity, Symmetry, and Transitivity, Equivalence Relations, Partial Order Relations **Graphs and Trees**: Definitions and Basic Properties, Trails, Paths, and Circuits, Matrix Representations of Graphs, Isomorphism's of Graphs, Trees, Rooted Trees, Isomorphism's of Graphs, Spanning trees and shortest paths.	12
5	**Counting and Probability**: Introduction, Possibility Trees and the Multiplication Rule, Possibility Trees and the **Multiplication Rule, Counting Elements of Disjoint Sets**: The Addition Rule, The Pigeonhole Principle, Counting	12

	Subsets of a Set: Combinations, r- Combinations with Repetition Allowed, Probability Axioms and, Conditional Probability, Bayes 'Formula, and Independent Events.	
	Total	60

Sr. No.	List of Practical
1	**Set Theory** a. Inclusion Exclusion principle. b. Power Sets c. Mathematical Induction
2	**Functions and Algorithms** a. Recursively defined functions b. Cardinality c. Polynomial evaluation d. Greatest Common Divisor.
3	**Counting** a. Sum rule principle b. Product rule principle c. Factorial d. Binomial coefficients e. Permutations f. Permutations with repetitions g. Combinations h. Combinations with repetitions i. Ordered partitions j. Unordered partitions
4	**Probability Theory** a. Sample space and events b. Finite probability spaces c. Equiprobable spaces

	d. Addition Principle e. Conditional Probability f. Multiplication theorem for conditional probability g. Independent events h. Repeated trials with two outcomes
5	**Graph Theory** a. Paths and connectivity b. Minimum spanning tree c. Isomorphism
6	**Directed Graphs** a. Adjacency matrix b. Path matrix
7	**Recurrence relations** a. Linear homogeneous recurrence relations with constant coefficients b. Solving linear homogeneous recurrence relations with constant coefficients c. Solving general homogeneous linear recurrence relations
8	**Algebraic Systems** a. Properties of operations b. Roots of polynomials
9	**Properties of integers** a. Division algorithm b. Primes c. Euclidean algorithm d. Fundamental theorem of arithmetic e. Congruence relation f. Linear congruence equation 8. Algebraic Systems a. Properties of operations b. Roots of polynomials
10	**Algebraic Systems** a. Properties of operations

	b. Roots of polynomials 9. Boolean Algebra a. Basic definitions in Boolean Algebra b. Boolean algebra as lattices

10.Evaluation Pattern:

- **Total Marks** : 150 Marks (10 Point Grading)
- **Passing Criteria** : 40 % (4 Grade Points)
- **Marking Scheme** : 60:40:50 Pattern
  - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
  - 40 Marks - Internal Assessment (Passing = 16 Marks)
  - 50 Marks - Practical Assessment (Passing = 20 Marks)
- **Mode of Evaluation of Answer-books** : Online/Offline

11.Paper Pattern:

a. **Internal Assessment:**

- Assessment consists of a class test of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
- Students have to submit assignments after completion of each module which will carry 15 marks and 5 marks are for attendance.

b. **Semester End Theory Examination :**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

12. Course Outcome:

Students will be able to:

CO1. Write an argument using logical notation and determine if the argument is or is not valid.

CO2. Demonstrate the ability to write and evaluate a proof or outline the basic structure of and give examples of each proof technique described.

CO3. Understand the basic principles of sets and operations in sets.

CO4. Prove basic set equalities.

CO5. Apply counting principles to determine probabilities.

CO6. Determine when a function is 1-1 and "onto".

CO7. Demonstrate different traversal methods for trees and graphs.

CO8. Model problems in Computer Science using graphs and trees

13. References:

1. Discrete Mathematics with Applications BY Sussana S. Epp Cengage Learning 4th 2010

2. Discrete Mathematics, Schaum's Outlines Series BY Seymour Lipschutz, Marc Lipson Tata MCGraw Hill 2007

3. Discrete Mathematics and its Applications BY Kenneth H. Rosen Tata MCGraw Hill

4. Discrete mathematical structures BY B Kolman RC Busby, S Ross PHI

5. Discrete structures BY Liu Tata MCGraw Hill

COURSE STRUCTURE

1. **Title of the Course:** Algebra
2. **Semester**: I
3. **Course Code: For Theory:** BDSCC102
   For Practical: BDSCCP102
4. **Course Objective:**
   The course is aimed to develop the basic Mathematical skills of learners that are imperative for effective understanding of data science subjects. The topics introduced will serve as basic tools for specialized studies in many fields of engineering and technology.
   a. Determinants & Matrices: To provide knowledge of determinants & matrices which is applied for solving system of linear equations and useful in various fields of technology.
   b. Complex numbers: This course enables the learner to learn the concept of imaginary numbers and gives awareness about algebra of complex numbers which helps in understanding of area of subjects like electrical circuits and complex analysis etc.
   c. Vector & Vector Space: This course enables the learners to understand the concept of Vector & Vector Space and its applications.
   d. This course will also enable the learners to understand the concept of fast Fourier and its usability in the field of Data Science.
   e. This course will also enable the learners to understand the concept of probability and statistics and its usability in the field of Data Science.
5. **Category of Course**: Core
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSCC102	Algebra	5	3	2	2	4

Module	Detailed Content	Hours
1	**Determinants:** The properties of determinants, Permutations and cofactors, Cramer's rule, Inverses and Volumes. **Introduction to vectors:** Vectors and Linear combinations, Lengths and Dot products, Matrices. **Solving Linear equations:**	12

	Vectors and Linear equations, the idea of elimination, Elimination using matrices, Rules for matrix operations, Inverse matrices, Elimination = Factorization: A=LU, Transpose and permutations.	
2	**Vector spaces and subspaces:** Spaces of vectors, the nullspace of A: solving AX = 0 and Rx = 0, The complete solution to AX = b, Independence, Basis and Dimension, Dimensions of the four subspaces. **Orthogonality:** Orthogonality of the four subspaces, Projections, Least square approximations, Orthonormal bases and Gram-Schmidt. **Eigenvalues and Eigenvectors:** Introduction to Eigenvalues, Diagonalizing a Matrix, Systems of Differential equations, Symmetric Matrices, Positive Definite Matrices.	12
3	**The singular value decomposition (SVD):** Image processing by Linear algebra, Bases and matrices in the SVD, Principal component analysis (PCA by SVD), The geometry of the SVD. **Linear Transformation:** The idea of linear transformation, The matrix of a linear transformation, The search for a good basis.	12
4	**Complex vectors and matrices:** Complex Numbers, Hermitian and Unitary matrices, The Fast Fourier. **Applications:** Graphs and networks, Matrices in engineering, Markov matrices, population and economics, Linear programming, Fourier series: Linear algebra for functions, Computer graphics, Linear algebra for cryptography.	12
5	**Numerical linear algebra:** Gaussian elimination in practice, Norms and condition numbers, Iterative methods and preconditioners. **Linear algebra in probability and statistics:** Mean, variance, and probability, Covariance matrices and joint probabilities, Multivariate gaussian and weighted least squares.	12
	Total	60

Sr. No.	**List of Practical**
1.	Using SPSS, execute the basic commands like Importing from Excel, Characteristics of Variables, Adding Value Labels.
2.	Using SPSS, write a program to understand basic commands on Grouping Data, Transforming Variables, Selecting a Subset, Producing summary stat.

3.	Using SPSS, write a program for understanding Frequencies, Percentages, Averages, Measures of spread.
4.	Using SPSS, execute the basic commands for producing Bar Charts, Cluster Bar Charts.
5.	Using SPSS, execute the basic commands for producing Histograms, Pie Charts.
6.	Using SPSS, execute the basic commands for producing Boxplots, Scatter Diagrams.
7.	Using SPSS, execute the basic commands for producing Tables, Two Way Tables.
8.	Using SPSS, execute the basic commands Interpreting Output, Drawing Conclusions Exporting to Word and PDF.

9. **Evaluation Pattern:**
   a. **Total Marks:** 150 Marks (10 Point Grading)
   b. **Passing Criteria**: 40 % (4 Grade Points)
   c. **Marking Scheme:** 60:40:50 Pattern
      - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
      - 40 Marks - Internal Assessment (Passing = 16 Marks)
      - 50 Marks - Practical Assessment (Passing = 20 Marks)
   d. **Mode of Evaluation of Answer-books:** Online/Offline

10. **Paper Pattern:**
   a. **Internal Assessment:**
      - Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
      - Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

   b. **Semester End Theory Examination:**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each.

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11. **Course Outcome:**

On successful completion of this course, the Learner should be able to:

CO1: Apply the knowledge of determinants and matrices to solve the problems in field of Image processing, Computer Graphics, Network Security etc.

CO2: Ability to interpret the mathematical results in physical or practical terms for complex numbers.

CO3: Solve and analyse the vector & Application in related field of engineering.

CO4: Solve and analyse fast Fourier in fields like Image processing.

CO5: Solve and Analyse problems in basic probability and statistics.

12. **References:**

1. A text book of Applied Mathematics by P.N.Wartikar and J.N.Wartikar, Vol – I and – II , 9th Edition, Pune Vidyarthi Graha, 2010.
2. Higher Engineering Mathematics by Dr. B. S. Grewal, 42nd Edition, Khanna Publication, 2017.
3. Advanced Engineering Mathematics by Erwin Kreyszig, 9th Edition, Wiley Eastern Limited.
4. A Textbook of Matrices by Shanti Narayan & P K Mittal, S. Chand publication, 1953.
5. Elementary Linear Algebra Application by Howard Anton and Christ Rorres, 11th edition, Wiley.

COURSE DETAILS

1) **Title of the Course:** Operating System
2) **Course Code: For Theory : BDSCC103**
   For Practical: BDSCCP103
3) **Course Objective:**

1**.** To understand the basic Operating System concepts and founding the services and advantages of it.

2. Importance of virtualization and cloud computing in today's IT industries.

3. Learning the features of different Operating System like Linux, Windows, and Android etc.

4. Understand how Operating system manage the File and Directory system.

5. Gaining the knowledge about Scheduling algorithm.

4) **Category of Course :** Core Course

5) **Semester :** I

6) **Total Hours:** 60 lectures

7) **Total Credits:** 04 Credits (02 Credits for Theory & 02 Credits for Practical)

8) **Modules:-**

Course Code	**Course Name**	**Teaching Scheme (Hours /Week)**		**Credits Assigned**		
		Theory	**Practical / Tutorial**	**Theory**	**Practical / Tutorial**	**Total**
BDSCC103	Operating System	5	3	2	2	4

Unit	**Details**	**Lectures**
I	**Introduction:** What is an operating system? History of operating system, computer hardware, different operating systems, operating system concepts, system calls, operating system structure. **Processes and Threads:** Processes, threads, inter process communication, scheduling, IPC	**12**
II	**Memory Management:** No memory abstraction, memory abstraction: address spaces, virtual memory, page replacement algorithms, design issues for paging systems, implementation issues, segmentation. **File Systems:** Files, directories, file system implementation, file-system management and optimization, MS-DOS file system, UNIX V7 file system, CD ROM file system	**12**

III	**Input-Output:** Principles of I/O hardware, Principles of I/O software, I/O software layers, disks, clocks, user interfaces: keyboard, mouse, monitor, thin clients, power management, **Deadlocks:** Resources, introduction to deadlocks, the ostrich algorithm, deadlock detection and recovery, deadlock avoidance, deadlock prevention, issues	**12**
IV	**Virtualization and Cloud:** History, requirements for virtualization, type 1 and 2 hypervisors, techniques for efficient virtualization, hypervisor microkernels, memory virtualization, I/O virtualization, Virtual appliances, virtual machines on multicore CPUs, Clouds. **Multiple Processor Systems** Multiprocessors, multicomputer, distributed systems.	**12**
V	**Case Study on LINUX and ANDROID:** History of Unix and Linux, Linux Overview, Processes in Linux, Memory management in Linux, I/O in Linux, Linux file system, security in Linux. Android **Case Study on Windows:** History of windows through Windows 10, programming windows, system structure, processes and threads in windows, memory management, caching in windows, I/O in windows, Windows NT file	**12**
	Total	**60**

Practical List :-

List of Practical	
1.	Installation of virtual machine software.
2.	Installation of Linux operating system (RedHat / Ubuntu) on virtual machine.
3.	Installation of Windows operating system on virtial machine.
4.	**Linux commands: Working with Directories:**
a.	pwd, cd, absolute and relative paths, ls, mkdir, rmdir,
b.	file, touch, rm, cp. mv, rename, head, tail, cat, tac, more, less, strings, chmod
5.	**Linux commands: Working with files:**
a.	ps, top, kill, pkill, bg, fg,
b.	grep, locate, find, locate.
c.	date, cal, uptime, w, whoami, finger, uname, man, df, du, free, whereis, which.
d.	Compression: tar, gzip.
6.	**Windows (DOS) Commands – 1**
a.	Date, time, prompt, md, cd, rd, path.
b.	Chkdsk, copy, xcopy, format, fidsk, cls, defrag, del, move.
7.	**Windows (DOS) Commands – 2**
a.	Diskcomp, diskcopy, diskpart, doskey, echo
b.	Edit, fc, find, rename, set, type, ver
8.	**Working with Windows Desktop and utilities**
a.	Notepad
b.	Wordpad
c.	Paint
d.	Taskbar
e.	Adjusting display resolution
f.	Using the browsers
g.	Configuring simple networking
h.	Creating users and shares
9.	**Working with Linux Desktop and utilities**
a.	The vi editor.
b.	Graphics
c.	Terminal

d.	Adjusting display resolution
e.	Using the browsers
f.	Configuring simple networking
g.	Creating users and shares
10.	**Installing utility software on Linux and Windows**

9) Evaluation Pattern:

a. **Total Marks** : 150 Marks (10 Point Grading)

b. **Passing Criteria** : 40 % (4 Grade Points)

c. **Marking Scheme** : 60:40:50 Pattern

- 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
- 40 Marks - Internal Assessment (Passing = 16 Marks)
- 50 Marks - Practical Assessment (Passing = 20 Marks)

d. **Mode of Evaluation of Answer-books** : Online/Offline

10) Paper Pattern:

a. **Internal Assessment:**

- Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
- Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

b. **Semester End Theory Examination :**

Question No.	**Description**	**Marks**
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11) Course Outcome:

Students will be able to:

CO1:- Illustrate the fundamentals of Operating System and its features.

CO2:- Explain the different types and services provided by an Operating System.

CO3:- Comprehend the concepts of Virtualization and Cloud computing.

CO4:-Describe the different scheduling algorithm.

CO5:-Discuss the properties of different Operating System like Linux Windows, Android etc.

CO 6:-Understand and Execute the Linux Commands in brief.

12) References:

1. Modern Operating Systems Andrew S. Tanenbaum,

Herbert Bos Pearson 4th 2014

2. Operating Systems –Internals and Design Principles Willaim Stallings Pearson 8th 2009

3. Operating System Concepts Abraham Silberschatz, Peter B. Galvineg GagneWiley 8th

4. Operating Systems Godbole and Kahate McGrawHill 3rd

COURSE STRUCTURE

1. **Title of the Course :** Introduction to software fundamentals
2. **Semester** : I
3. **Course Code: For Theory :** BDSSB104
   For Practical: BDSSBP104
4. **Course Objective:**
   a. To learn the fundamental programming concepts and methodologies which are essential to building good C programs.
   b. To practice the fundamental programming methodologies in the C programming language via laboratory experiences. Microsoft Visual Studio is the programming environment that will used.
   c. To code, document, test, and implement a well-structured, robust computer program using the C programming language.
   d. Students will be able to develop logics which will help them to create programs, applications in C. Also by learning the basic programming constructs they can easily switch over to any other language in future
   e. To write reusable modules (collections of functions).
   f. The course is designed to provide complete knowledge of C language.
5. **Category of Course** : Skill Based
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSSB 104	Introduction to software fundamentals	5	3	2	2	4

Module	Detailed Content	Hours
1	**Introduction:** Types of Programming languages, History, features and application. Simple program logic, program development cycle, pseudocode statements and flowchart symbols, sentinel value to end a program, programming and user environments, evolution of programming models., desirable program characteristics. **Fundamentals:** Structure of a program. Compilation and Execution of a Program, Character Set, identifiers and keywords, data types, constants, variables and arrays,	12

	declarations, expressions, statements, Variable definition, symbolic constant.	
2	**Operators and Expressions:** Arithmetic operators, unary operators, relational and logical operators, assignment operators, assignment operators, the conditional operator, library functions. **Data Input and output:** Single character input and output, entering input data, scanf function, printf function, gets and puts functions, interactive programming.	12
3	**Conditional Statements and Loops:** Decision Making Within A Program, Conditions, Relational Operators, Logical Connectives, If Statement, If-Else Statement, Loops: While Loop, Do While, For Loop. Nested Loops, Infinite Loops, Switch Statement. **Functions:** Overview, defining a function, accessing a function, passing arguments to a function, specifying argument data types, function prototypes, recursion, modular programming and functions, standard library of c functions, prototype of a function: foo1lal parameter list, return type, function call, block structure, passing arguments to a function: call by reference, call by value.	12
4	**Program structure:** Storage classes, automatic variables, external variables, static variables, multifile programs, more library functions, **Preprocessor:** Features, #define and #include, Directives and Macros **Arrays:** Definition, processing, passing arrays to functions, multidimensional arrays, arrays and string.	12
5	**Pointers** : Fundamentals, Declarations, Pointer address operators, Pointer type declaration, Pointer assignment, Pointer initialization, Pointer Arithmetic, Functions and Pointers, Arrays and Pointers, Pointer Arrays, Passing functions to other functions. **Structures and Unions:** Structure variables, Initialization, Structure assignment, Nested structures, structures and functions, structures and arrays, Arrays of structures, structures containing arrays Union, structure and pointers.	12
	Total	60

Sr. No.	List of Practical
1	Write a program to find the addition, subtraction, multiplication and division of two numbers
2	Write a program to swap two numbers without using third variable.
3	Write a program to find the area of rectangle, square and circle.
4	Write a program to check whether the number is even or odd.
5	Write a program to find the factorial of a number.
6	Write a program to check whether the entered number is prime or not.
7	Write a program to find the sum of squares of digits of a number.
8	Write a programs to print the Fibonacci series.
9	Write a program to find whether a given number is palindrome or not.
10	Write a program to find the factorial of a number using recursive function..
11	Write a program to find the largest value that is stored in the array.
12	Write a program to demonstrate the use of pointers.
13	Programs on structures.

9. **Evaluation Pattern:**
   a. **Total Marks** : 150 Marks (10 Point Grading)
   b. **Passing Criteria** : 40 % (4 Grade Points)
   c. **Marking Scheme** : 60:40:50 Pattern
      - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
      - 40 Marks - Internal Assessment (Passing = 16 Marks)
      - 50 Marks - Practical Assessment (Passing = 20 Marks)
   d. **Mode of Evaluation of Answer-books** : Online/Offline

10. **Paper Pattern:**
   a. **Internal Assessment:**
      - Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
      - Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

   b. **Semester End Theory Examination :**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10

3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11. Course Outcome:

Upon successful completion of this course, students should be able to develop application

CO1: To describe the advantages of a high level language like C/C++, the programming process and the compilation process

CO2: To describe and use software tools in the programming process (IDE)

CO3: To apply good programming principles to the design and implementation of C programs

CO4: To design, implement, debug and test programs using the fundamental elements of C.

CO5: 5: To demonstrate an understanding of primitive data types, values, operators and expressions in C

12. References:

1. Programming with C Byron Gottfried Tata McGRAW Hill 2nd 1996
2. Programming Logic and Design Joyce Farell Cengage Learning 8th 2014
3. "C" Programming" Brian W. Kernighan and Denis M. Ritchie. PHI 2nd
4. Let us C Yashwant P. Kanetkar, BPB publication
5. C for beginners Madhusudan Mothe X-Team Series 1st 2008
6. 21st Century C Ben Klemens OReilly 1st 2012

COURSE DETAILS

1) **Title of the Course:** Communication Skill
2) **Course Code: For Theory : BDSAE105**
   For Practical: BDSAEP105

3) **Course Objective:**

1. Understand how they use their energy to work effectively.

2 .Learn how to manage themselves better, especially when facing work situations which cause them stress.

3. Be more aware of the impact they have on other people.

4. Be more skillful at understanding how and why other people behave and react as they do.

4) **Category of Course :** Ability Enhancement Course

5) **Semester :** I

6) **Total Hours:** 60 lectures

7) **Total Credits:** 04 Credits (02 Credits for Theory & 02 Credits for Practical)

8) **Modules:-**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical / Tutorial	Theory	Practical / Tutorial	Total
BITAE205	Communication Skill	5	3	2	2	4

Unit	Details	Lectures
I	**The Seven Cs of Effective Communication:** Completeness, Conciseness, Consideration, Concreteness, Clarity, Courtesy, Correctness **Understanding Business Communication:** Nature and Scope of Communication, Non-verbal Communication, Cross-cultural communication, Technology-enabled Business Communication	**12**

II	**Writing Business Messages and Documents:** Business writing, Business Correspondence, Instructions Business Reports and Proposals, Career building and Resume writing. **Developing Oral Communication Skills for Business:** Effective Listening, Business Presentations and Public Speaking, Conversations, Interviews	**12**
III	**Developing Oral Communication Skills for Business:** Meetings and Conferences, Group Discussions and Team Presentations, Team Briefing, **Understanding Specific Communication Needs:** Communication across Functional Areas	**12**
IV	**Understanding Specific Communication Needs:** Corporate Communication, Persuasive Strategies in Business Communication, Ethics in Business Communication, Business Communication Aids	**12**
V	**Presentation Process:** Planning the presentations, executing the presentations, Impressing the audience by performing, Planning stage: Brainstorming, mind maps / concept maps, executing stage: chunking theory, creating outlines, Use of templates. Adding graphics to your presentation: Visual communication, Impress stage: use of font, colour, layout, Importance of practice and performance.	**12**
	Total	**60**

Practical List :-

List of Practical Questions:	
1.	Communication Origami, Guessing Game, Guessing the emotion
2.	Body Language, Follow All Instructions, Effective Feedback Skills
3.	The Name Game, Square Talk (Effective Communication), Room 101 (Influential and persuasive skills)
4.	Back to Back Communication, Paper Shapes (Importance of two-way communication), Memory Test(Presentation Skills)
5.	Exercises on Communication Principles
6.	Exercises on communication icebreakers
7.	Communication exercises
	For the following practicals, Microsoft Office, Open Office, Libre Office or any other software suite can be used.

8.	Use of word processing tools for communication
9.	Use of spreadsheet tools for communication
10.	Use of presentation tools for communication

9) Evaluation Pattern:

a. **Total Marks** : 150 Marks (10 Point Grading)

b. **Passing Criteria** : 40 % (4 Grade Points)

c. **Marking Scheme** : 60:40:50 Pattern

- 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
- 40 Marks - Internal Assessment (Passing = 16 Marks)
- 50 Marks - Practical Assessment (Passing = 20 Marks)

d. **Mode of Evaluation of Answer-books** : Online/Offline

10) Paper Pattern:

a. **Internal Assessment:**

- Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
- Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

b. **Semester End Theory Examination :**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total

2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11) Course Outcome:

After studying this course the students would gain enough knowledge on:

1. Students will be able to understand and apply knowledge of human communication and language processes as they occur across various contexts, e.g., interpersonal, intrapersonal, small group, organizational, media, gender, family, intercultural communication, technologically mediated communication, etc. from multiple perspectives.
2. Students will be able to understand and evaluate key theoretical approaches used in the interdisciplinary field of communication. I.e., students will be able to explain major theoretical frameworks, constructs, and concepts for the study of communication and language, summarize the work of central thinkers associated with particular approaches, and begin to evaluate the strengths and weaknesses of their approaches.
3. Students will be able to understand the research methods associated with the study of human communication, and apply at least one of those approaches to the analysis and evaluation of human communication.
4. Students will be able to find, use, and evaluate primary academic writing associated with the communication discipline.
5. Students will develop knowledge, skills, and judgment around human communication that facilitate their ability to work collaboratively with others. Such skills could include communication competencies such as managing conflict, understanding small group processes, active listening, appropriate self-disclosure, etc.
6. Students will be able to communicate effectively orally and in writing.

12) References:

1. Business Communication Edited by Meenakshi Raman and Prakash Singh Oxford University Press Second.

2. Professional Communication Aruna Koneru Tata McGraw Hill

3. Strategies for improving your business communication Prof. M. S. Rao Shroff publishers and distributors 2016.

4. Business Communication Dr. Rishipal and Dr. Jyoti Sheoran SPD 2014.

5. Communication Skills Dr. Nageshwar Rao Dr. Rajendra P. Das Himalaya Publishing House.

COURSE DETAILS

1.Title of the Course : Green Computing

2.Course Code: For Theory : BDSID106

For Project:- BDSIDP106

3.Course Objective:

1. To understand how to reduce the use of hazardous materials, maximize energy efficiency during the product life time.
2. Importance of recycling, biodegradability of defunct products and factory waste. Changing the way of work with GREEN in mind

4. Category of Course : Inter-disciplinary

5. Semester : I

6. Total Hours: 60 lectures

7. Total Credits: 04 Credits (02 Credits for Theory & 02 Credits for Practical/Project)

8. Modules:-

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSID106	Green Computing	5	3	2	2	4

Unit	Details	Lectures
I	**Overview and Issues:** Problems: Toxins, Power Consumption, Equipment Disposal, Company's Carbon Footprint: Measuring, Details, reasons to bother, Plan for the Future, Cost Savings: Hardware, Power. **Initiatives and Standards:** Global Initiatives: United Nations, Basel Action Network, Basel	**12**
II	**Minimizing Power Usage:** Power Problems, Monitoring Power Usage, Servers, Low-Cost Options, Reducing Power Use, Data De-Duplication, Virtualization, Management, Bigger Drives, Involving the Utility Company, Low-Power Computers, PCs, Linux, Components, Servers, Computer Settings, Storage, Monitors, Power Supplies, Wireless Devices, Software.	**12**

	Cooling: Cooling Costs, Power Cost, Causes of Cost, Calculating Cooling Needs, Reducing Cooling Costs, Economizers, On-Demand Cooling, HP's Solution, Optimizing Airflow, Hot Aisle/Cold Aisle, Raised Floors, Cable Management, Vapour Seal, Prevent Recirculation of Equipment Exhaust, Supply Air Directly to Heat Sources, Fans, Humidity, Adding Cooling, Fluid Considerations, System Design,	
III	**Changing the Way of Work:** Old Behaviours, starting at the Top, Process Reengineering with Green in Mind, Analysing the Global Impact of Local Actions, Steps: Water, Recycling, Energy, Pollutants, Teleworkers and Outsourcing, Telecommuting, Outsourcing, how to Outsource. **Going Paperless:** Paper Problems, The Environment, Costs: Paper and Office, Practicality, Storage, Destruction, Going Paperless, Organizational Realities, Changing Over, Paperless Billing, Handheld Computers vs. the Clipboard, Unified Communications, Intranets, What to Include, Building an Intranet, Microsoft Office SharePoint Server 2007, Electronic Data Interchange (EDI), Nuts and Bolts, Value Added	**12**
IV	**Recycling:** Problems, China, Africa, Materials, Means of Disposal, Recycling, Refurbishing, Make the Decision, Life Cycle, from beginning to end, Life, Cost, Green Design, Recycling Companies, Finding the Best One, Checklist, Certifications, Hard Drive Recycling, Consequences, cleaning a Hard Drive, Pros and cons of each method, CDs and DVDs, good and bad about CD and DVDs disposal, Change the mind-set, David vs. America Online **Hardware Considerations:** Certification Programs, EPEAT, RoHS, Energy Star, Computers, Monitors, Printers, Scanners, All-in-Ones, Thin Clients, Servers, Blade Servers, Consolidation, Products, Hardware Considerations, Planned Obsolescence, Packaging, Toxins, Other Factors, Remote Desktop, Using Remote Desktop, Establishing a Connection, In Practice	**12**
V	**Greening Your Information Systems:** Initial Improvement Calculations, Selecting Metrics, Tracking Progress, Change Business Processes, Customer Interaction, Paper Reduction, Green Supply Chain, Improve Technology Infrastructure, Reduce PCs and Servers, Shared Services, Hardware Costs, Cooling. **Staying Green:** Organizational Check-ups, Chief Green Officer, Evolution, Sell the CEO, SMART Goals, Equipment Check-ups, Gather Data, Tracking the data, Baseline Data, Benchmarking, Analyse Data, Conduct Audits, Certifications, Benefits, Realities, Helpful Organizations.	**12**
	Total	**60**

Practical/Project List:

Project and Viva Voce	
1.	**A project should be done based on the objectives of Green Computing. A report of minimum 50 pages should be prepared. The report should have a font size of 12, Times new roman and 1.5 line spacing. The headings should have font size 14. The report should be hard bound.**
2.	**The project can be done individually or a group of two students.**
3.	**The students will have to present the project during the examination.**
4.	**A certified copy of the project report is essential to appear for the examination.**

9. Evaluation Pattern:

a. **Total Marks** : 150 Marks (10 Point Grading)

b. **Passing Criteria** : 40 % (4 Grade Points)

c. **Marking Scheme** : 60:40:50 Pattern

- 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
- 40 Marks - Internal Assessment (Passing = 16 Marks)
- 50 Marks - Project Assessment (Passing = 20 Marks)

d. **Mode of Evaluation of Answer-books** : Online/Offline

10. Paper Pattern:

a. **Internal Assessment:**

- Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
- Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

b. **Semester End Theory Examination :**

Question No.	**Description**	**Marks**
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

c. **Semester End Project Submission**: Total Marks:-50

11. Course Outcome :

CO1:- Practice of environmentally sustainable production practices, energy efficient computers.

CO2:- Understand the importance of energy efficiency, power consumption and other way is making green software to thrive the industry and make innovatory products.

CO3:- Comprehend the concepts of Recycling like water recycling.

12. Books and References:

1. Green IT Toby Velte, Anthony Velte, Robert Elsenpeter McGrawHill 2008

2. Green Data Center: Steps for the Journey AlvinGalea, Michael Schaefer, Mike Ebbers Shroff Publishers and Distributers 2011

3. Green Computing and Green IT Best Practice Jason Harris Emereo

4. Green Computing Tools and Techniques for Saving Energy, Money
And Resources Bud E. Smith CRC Press 2014

Bachelor of Science in Data Science

[**B. Sc. In Data Science**]

Semester - II

COURSE STRUCTURE

1. **Title of the Course :** Probability
2. **Semester** : II
3. **Course Code: For Theory :** BDSCC201
   For Practical: BDSCCP201
4. **Course Objective:**

   1. To make students to use measure-theoretic and analytical techniques for understanding probability concept.
   2. Use basic counting techniques (multiplication rule, combinations, and permutations) to compute probability and odds.
   3. Work with continuous random variables. In particular, know the properties of uniform, normal and exponential distributions.
   4. Compute the covariance and correlation between jointly distributed variables

5. **Category of Course**: Core Course
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Evaluation Pattern:**
   - **Total Marks** : 150 Marks (10 Point Grading)
   - **Passing Criteria** : 40 % (4 Grade Points)
   - **Marking Scheme** : 60:40:50 Pattern
     - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
     - 40 Marks - Internal Assessment (Passing = 16 Marks)
     - 50 Marks - Practical Assessment (Passing = 20 Marks)
   - **Mode of Evaluation of Answer-books** : Online/Offline

9. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSCC201	Probability	5	3	2	2	4

Module	Detailed Content	Hours
1	**Combinatorial Analysis** Introduction The Basic Principle of Counting Permutations Combinations Multinomial Coefficients **Axioms of Probability**	12

	Introduction Sample space and Events Axioms of probability Sample spaces having equally likely outcomes	
2	**Conditional Probability and Independence** Introduction Conditional Probability Bayes' Formula Independent Events **Random Variables** Random Variables Discrete Random Variables Expected Value Expectation of a function of a Random Variable Variance The Bernoulli and Binomial Random Variables The Poisson Random Variable Other Discrete Probability Distributions Expected values of Sums of Random Variables Properties of Cumulative Distribution Function	12
3	**Continuous Random Variables** Introduction Expectation and Variance of Continuous Random Variables The Uniform Random Variable Normal Random Variable Exponential Random Variable Other Continuous Distributions The Distribution of a Function of a Random Variable	12
4	**Jointly Distributed Random Variables** Joint Distribution Functions Independent Random Variables Sums of Independent Random Variables Conditional Distributions Joint probability distributions of Functions of Random Variables	12
5	**Properties of Expectation** Introduction Expectation of sums of random variables Moments of the number of events that occur Covariance, Variance of Sums, and Correlations Conditional Expectation Conditional Expectation and Prediction Moment generating Function Additional Properties of Normal Random Variables	12

	General Definition of expectation **Limit Theorem** Introduction Chebyshev's Inequality and weak law of large numbers The Central Limit Theorem The strong law of large numbers Other Inequalities	
	Total	60

Sr. No.	**List of Practical**
1	**Probability Theory** Sample space and events Finite probability space
2	**Probability Theory** Equiprobable space Addition Principle
3	**Probability Theory** Conditional Probability Multiplication theorem for conditional probability
4	**Probability Theory** Independent events Repeated trials with two outcomes
5	**Program to generate random variables**
6	**Program to fit binomial distribution.**
7	**Program to fit Poisson distribution.**
8	**Program for Uniform distribution.**
9	**Program for Bernoulli distribution.**
10	**Program for Negative binomial distribution.**

10. Evaluation Pattern:

- **Total Marks** : 150 Marks (10 Point Grading)
- **Passing Criteria** : 40 % (4 Grade Points)
- **Marking Scheme** : 60:40:50 Pattern
  - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
  - 40 Marks - Internal Assessment (Passing = 16 Marks)
  - 50 Marks - Practical Assessment (Passing = 20 Marks)
- **Mode of Evaluation of Answer-books** : Online/Offline

11. Paper Pattern:

a. **Internal Assessment:**

- Assessment consists of a class test of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
- Students have to submit assignments after completion of each module which will carry 15 marks and 5 marks are for attendance.

b. **Semester End Theory Examination :**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

12. Course Outcome:

Students will be able to:
CO1: Understand measure and measurable functions
CO2: Analyze probability concepts using measure-theoretic approach
CO3: Identify applications of different limit theorems in statistical problems
CO4: Apply Radon-Nikodym theorem in conditional probability

13. References:

1. Billingsley, P. (2008) Probability and Measure, Second edition, John Wiley
2. Bhat, B.R. (2018) Modern Probability Theory, Second edition, Wiley Eastern
3. Rohatgi, V.K. and Salah, A.K.E. (2011) an Introduction to Probability and Statistics, John Wiley & Sons.
4. Durrett, Rick. Probability: Theory and Examples. 4th ed. Cambridge University Press, 2010
5. A first course in probability by Sheldon Ross (Pearson Publication)

COURSE STRUCTURE

1. **Title of the Course :** Linux
2. **Semester** : II
3. **Course Code: For Theory:** BDSCC202
   For Practical: BDSCCP202
4. **Course Objective:**
   a. To impart knowledge and skills on various practical and theoretical aspects of Linux operating system (OS) basics and Linux OS based server configuration, management.
   b. This course introduces various tools and techniques commonly used by Linux programmers, system administrators and end users to achieve their day to day work in Linux environment.
   c. It is designed for computer students who have limited or no previous exposure to Linux.
5. **Category of Course** : Core Course
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSCC202	Linux	5	3	2	2	4

Module	Detailed Content	Hours
1	Introduction Distributions, Linux kernel vs distribution. Importance of Linux in software ecosystem: web servers, supercomputers, mobile, servers. Linux Structure Linux Architecture, Filesystem basics, The boot process, init scripts, runlevels, shutdown process, Very basic introductions to Linux processes, Packaging methods: rpm/deb, Graphical Vs Command line.	12
2	Graphical Desktop Session Management, Basic Desktop Operations, Network Management, Installing and Updating	12

	Software, Text editors: gedit, vi, vim, emacs, Graphics editors, Multimedia applications. Command Line Command line mode options, Shells, Basic Commands, General Purpose Utilities, Installing Software, User management, Environment variables, Command aliases.	
3	Linux Documentation man pages, GNU info, help command, More documentation sources File Operations Filesystem, Filesystem architecture, File types, File attributes, Working with files, Backup, compression	12
4	Security Understanding Linux Security, Uses of root, sudo command, working with passwords, Bypassing user authentication, Understanding ssh Networking Basic introduction to Networking, Network protocols: http, ftp etc., IP address, DNS, Browsers, Transferring files. ssh, telnet, ping, traceroute, route, hostname, networking GUI	12
5	Basic Shell Scripting Features and capabilities, Syntax, Constructs, Modifying files, File manipulation utilities, Dealing with large files and Text, String manipulation, Boolean expressions, File tests, Case, Debugging, Regular expressions.	12
	Total	60

Sr. No.	**List of Practical**
1	Install your choice of Linux distribution e.g. Ubuntu, Fedora, Debian
2	Customize desktop environment by changing different default options like changing default background, themes, screensavers, Networking: Get the current networking configuration for your desktop.
3	Installing and Removing Software: Install gcc package. Verify that it runs, and then remove it.
4	Basic Commands: cd, mkdir, cat,more,less,head,tail,ls,date, cal, rmdir, mv,rm,cp Demonstration of chmod command
5	Advance Commands Demonstration of chmod command,Chgrp,chown etc
6	Use environment:create user , delete user, login status,
7	Linux Editors: vim/emacs Create, modify, search, navigate a file in editor.
8	Linux Security: Use of sudo to change user privileges to root, Create a new user and add it to sudo configuration file, Set password for new user
9	Network: Get IP address of your machine using ifconfig, If IP is not set, then assign an IP address according to your network settings, Get hostname of

	your machine, Use ping to check the network connectivity to remote machines.
10	Write a Shell Script For Cheking Even/Odd numbers Using && Operator.

9. **Evaluation Pattern:**
   a. **Total Marks** : 150 Marks (10 Point Grading)
   b. **Passing Criteria** : 40 % (4 Grade Points)
   c. **Marking Scheme** : 60:40:50 Pattern
      - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
      - 40 Marks - Internal Assessment (Passing = 16 Marks)
      - 50 Marks - Practical Assessment (Passing = 20 Marks)
   d. **Mode of Evaluation of Answer-books** : Online/Offline

10. **Paper Pattern:**
   a. **Internal Assessment:**
      - Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
      - Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

   b. **Semester End Theory Examination :**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

   c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total

2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11. **Course Outcome:**

Upon successful completion of this course, learner should be able to:

CO1: Upon completion of this course, students should have a good working knowledge of Linux, from both a graphical and command line perspective, allowing them to easily use any Linux distribution.

CO2: This course shall help student to learn advanced subjects in computer science practically.

CO3: Student shall be able to progress as a Developer or Linux System Administrator using the acquired skill set.

12. **References:**

1. Unix Concepts and Applications by Sumitabha Das.
2. Additional References:
   1) Linux kernel Home: http://kernel.org
   2) Open Source Initiative: https://opensource.org/
   3) The Linux Foundation: http://www.linuxfoundation.org/

COURSE STRUCTURE

1. **Title of the Course :** Algorithms and Data structures
2. **Semester** : II
3. **Course Code: For Theory:** BDSCC203
   For Practical: BDSCCP203
4. **Course Objective:**
   a. To impart the basic concepts of data structures and algorithms
   b. To understand concepts about searching and sorting techniques
   c. To Understand basic concepts about stacks,queues,lists,trees and graphs
   d. To understanding about writing algorithms and step by step approach in solving problems with the help of fundamental data structures
5. **Category of Course** : Core Subject
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BITCC303	Data structure	5	3	2	2	4

Module	Detailed Content	Hours
1	**Introduction:** Data and Information, Data Structure, Classification of Data Structures, Primitive Data Types, Abstract Data Types, Data structure vs. File Organization, Operations on Data Structure ,Algorithm, Importance of Algorithm Analysis, Complexity of an Algorithm, Asymptotic Analysis and Notations, Big O Notation, Big Omega Notation, Big Theta Notation, Rate of Growth and Big O Notation. **Array**: Introduction, One Dimensional Array, Memory Representation of One Dimensional Array, Traversing, Insertion, Deletion, Searching, Sorting, Merging of Arrays, Multidimensional Arrays, Memory Representation of Two Dimensional Arrays, General Multidimensional Arrays, Sparse	12

	Arrays, Sparse Matrix, Memory Representation of Special kind of Matrices, Advantages and Limitations of Arrays.	
2	**Linked List:** Linked List, One-way Linked List, Traversal of Linked List, Searching, Memory Allocation and De-allocation, Insertion in Linked List, Deletion from Linked List, Copying a List into Other List, Merging Two Linked Lists, Splitting a List into Two Lists, Reversing One way linked List, Circular Linked List, Applications of Circular Linked List, Two way Linked List, Traversing a Two way Linked List, Searching in a Two way linked List, Insertion of an element in Two way Linked List, Deleting a node from Two way Linked List, Header Linked List, Applications of the Linked list, Representation of Polynomials, Storage of Sparse Arrays, Implementing other Data Structures.	12
3	**Stack:** Introduction, Operations on the Stack Memory Representation of Stack, Array Representation of Stack, Applications of Stack, Evaluation of Arithmetic Expression, Matching Parenthesis, infix and postfix operations, Recursion. **Queue:** Introduction, Queue, Operations on the Queue, Memory Representation of Queue, Array representation of queue, Linked List Representation of Queue, Circular Queue, Some special kinds of queues, Deque, Priority Queue, Application of Priority Queue, Applications of Queues	12
4	**Sorting and Searching Techniques**: Bubble, Selection, Insertion, Merge Sort. Searching: Sequential, Binary, Indexed Sequential Searches, Binary Search. **Tree:** Tree, Binary Tree, Properties of Binary Tree, Memory Representation of Binary Tree, Operations Performed on Binary Tree, Reconstruction of Binary Tree from its Traversals, Huffman Algorithm, Binary Search Tree, Operations on Binary Search Tree, Heap, Memory Representation of Heap, Operation on Heap, Heap Sort. **Advanced Tree Structures:** Red Black Tree, Operations Performed on Red Black Tree, AVL Tree, Operations performed on AVL Tree, 2- 3 Tree, B-Tree.	12
5	**Hashing Techniques:** Hash function, Address calculation techniques, Common hashing functions Collision resolution, Linear probing, Quadratic, Double hashing, Bucket hashing, Deletion and rehashing **Graph:** Introduction, Graph, Graph Terminology, Memory Representation of Graph, Adjacency Matrix Representation of Graph, Adjacency List or Linked Representation of Graph, Operations Performed on Graph, Graph Traversal, Applications	12

	of the Graph, Reachability, Shortest Path Problems, Spanning Trees.	
	Total	60

Sr. No.	List of Practical
1	**Implement the following:** a. Write a program to store the elements in 1-D array and perform the operations like searching, sorting and reversing the elements. [Menu Driven] b. Read the two arrays from the user and merge them and display the elements in sorted order.[Menu Driven] c. Write a program to perform the Matrix addition, Multiplication and Transpose Operation. [Menu Driven]
2	**Implement the following for Linked List:** a. Write a program to create a single linked list and display the node elements in reverse order. b. Write a program to search the elements in the linked list and display the same c. Write a program to create double linked list and sort the elements in the linked list.
3	**Implement the following for Stack:** a. Write a program to implement the concept of Stack with Push, Pop, Display and Exit operations. b. Write a program to convert an infix expression to postfix and prefix conversion.
4	**Implement the following for Queue:** a. Write a program to implement the concept of Queue with Insert, Delete, Display and Exit operations. b. Write a program to implement the concept of Circular Queue c. Write a program to implement the concept of Deque.
5	**Implement the following sorting techniques**: a. Write a program to implement bubble sort. b. Write a program to implement selection sort. c. Write a program to implement insertion sort.
6	**Implement the following data structure techniques:** a. Write a program to implement merge sort. b. Write a program to search the element using sequential search. c. Write a program to search the element using binary search.
7	**Implement the following data structure techniques:** a. Write a program to create the tree and display the elements. b. Write a program to construct the binary tree. c. Write a program for inorder, postorder and preorder traversal of tree

8	**Implement the following data structure techniques:** a. Write a program to insert the element into maximum heap. b. Write a program to insert the element into minimum heap
9	**Implement the following data structure techniques:** a. Write a program to implement the collision technique. b. Write a program to implement the concept of linear probing
10	**Implement the following data structure techniques:** a. Write a program to generate the adjacency matrix. b. Write a program for shortest path diagram.

9. **Evaluation Pattern:**
   a. **Total Marks** : 150 Marks (10 Point Grading)
   b. **Passing Criteria** : 40 % (4 Grade Points)
   c. **Marking Scheme** : 60:40:50 Pattern
      - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
      - 40 Marks - Internal Assessment (Passing = 16 Marks)
      - 50 Marks - Practical Assessment (Passing = 20 Marks)
   d. **Mode of Evaluation of Answer-books** : Online/Offline

10. **Paper Pattern:**
   a. **Internal Assessment:**
      - Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
      - Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

   b. **Semester End Theory Examination :**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11. Course Outcome:

Upon successful completion of this course, students should be able to:

CO1: Ability to analyse algorithms and algorithm correctness

CO2: Ability to summarize searching and sorting techniques

CO3: Ability to describe stack, queue and linked list operation.

CO4: Ability to have knowledge of tree and graphs concepts.

12. References:

1. A Simplified Approach to Data Structures Lalit Goyal, Vishal Goyal, Pawan Kumar SPD 1st 2014
2. An Introduction to Data Structure with Applications Jean – Paul Tremblay and Paul Sorenson Tata MacGraw Hill 2nd 2007
3. Data Structure and Algorithm Maria Rukadikar SPD 1st 2017
4. Schaum's Outlines Data structure Seymour Lipschutz Tata McGraw Hill 2nd 2005
5. Data structure – A Pseudo code Approach with C AM Tanenbaum, Y Langsam and MJ Augustein Prentice Hall India 2nd 2006
6. Data structure and Algorithm Analysis in C Weiss, Mark Allen Addison Wesley 1st 2006

COURSE STRUCTURE

1. **Title of the Course :** Introduction to python
2. **Semester** : II
3. **Course Code: For Theory: BDSSB204**
   For Practical: BDSSBP204
4. **Course Objective:**

The learning objectives of this course are:

- To understand why Python is a useful scripting language for developers.
- To learn how to design and program Python applications.
- To learn how to use lists, tuples, and dictionaries in Python programs.
- To learn how to identify Python object types.

5. **Category of Course** : Skill Based
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSSB204	Introduction to python	5	3	2	2	4

Module	Detailed Content	Hours
1	**Introduction:** The Python Programming Language, History, features, Installing Python, Running Python program, Debugging : Syntax Errors, Runtime Errors, Semantic Errors, Experimental Debugging, Formal and Natural Languages, The Difference Between Brackets, Braces, and Parentheses, **Variables and Expressions** Values and Types, Variables, Variable Names and Keywords, Type conversion, Operators and Operands, Expressions, Interactive Mode and Script Mode, Order of Operations. **Conditional Statements:** if, if-else, nested if –else	12

	Looping: for, while, nested loops **Control statements:** Terminating loops, skipping specific conditions.	
2	**Functions:** Function Calls, Type Conversion Functions, Math Functions, Composition, Adding New Functions, Definitions and Uses, Flow of Execution, Parameters and Arguments, Variables and Parameters Are Local, Stack Diagrams, Fruitful Functions and Void Functions, Why Functions? Importing with from, Return Values, Incremental Development, Composition, Boolean Functions, More Recursion, Leap of Faith, Checking Types **Strings:** A String Is a Sequence, Traversal with a for Loop, String Slices, Strings Are Immutable, Searching, Looping and Counting, String Methods, The in Operator, String Comparison, String Operations.	12
3	**Lists:** Values and Accessing Elements, Lists are mutable, traversing a List, Deleting elements from List, Built-in List Operators, Concatenation, Repetition, In Operator, Built-in List functions and methods **Tuples and Dictionaries:** Tuples, Accessing values in Tuples, Tuple Assignment, Tuples as return values, Variable-length argument tuples, Basic tuples operations, Concatenation, Repetition, in Operator, Iteration, Built-in Tuple Functions Creating a Dictionary, Accessing Values in a dictionary, Updating Dictionary, Deleting Elements from Dictionary, Properties of Dictionary keys, Operations in Dictionary, Built-In Dictionary Functions, Built-in Dictionary Methods **Files:** Text Files, The File Object Attributes, Directories, **Exceptions:** Built-in Exceptions, Handling Exceptions, Exception with Arguments, User-defined Exceptions	12
4	**Regular Expressions –** Concept of regular expression, various types of regular expressions, using match function. **Classes and Objects:** Overview of OOP (Object Oriented Programming), Class Definition, Creating Objects, Instances as Arguments, Instances as return values, Built-in Class Attributes,	12

	Inheritance, Method Overriding, Data Encapsulation, Data Hiding **Multithreaded Programming:** Thread Module, creating a thread, synchronizing threads, multithreaded priority queue **Modules:** Importing module, Creating and exploring modules, Math module, Random module, Time module	
5	**Creating the GUI Form and Adding Widgets:** **Widgets:** Button, Canvas, Checkbutton, Entry, Frame, Label, Listbox, Menubutton, Menu, Message, Radiobutton, Scale, Scrollbar, text,Toplevel, Spinbox, PanedWindow, LabelFrame, tkMessagebox.Handling Standard attributes and Properties of Widgets. **Layout Management:** Designing GUI applications with proper Layout Management features. **Look and Feel Customization**: Enhancing Look and Feel of GUI using different appearances of widgets. **Storing Data in Our MySQL Database via Our GUI :** Connecting to a MySQL database from Python, Configuring the MySQL connection, Designing the Python GUI database, Using the INSERT command, Using the UPDATE command, Using the DELETE command, Storing and retrieving data from MySQL database.	12
	Total	60

Sr. No.	**List of Practical**
1	**Write the program for the following:** a.Create a program that asks the user to enter their name and their age. Print out a message addressed to them that tells them the year that they will turn 100 years old. b. Enter the number from the user and depending on whether the number is even or odd, print out an appropriate message to the user. c. Write a program to generate the Fibonacci series. d. Write a function that reverses the user defined value. e. Write a function to check the input value is Armstrong and also write the function for Palindrome. f. Write a recursive function to print the factorial for a given number.

2	**Write the program for the following:** a.Write a function that takes a character (i.e. a string of length 1) and returns True if it is a vowel, False otherwise. b. Define a function that computes the *length* of a given list or string. c. Define a *procedure* histogram() that takes a list of integers and prints a histogram to the screen. For example, histogram([4, 9, 7]) should print the following: **** ********* *******
3	**Write the program for the following:** a.A *pangram* is a sentence that contains all the letters of the English alphabet at least once, for example: *The quick brown fox jumps over the lazy dog*. Your task here is to write a function to check a sentence to see if it is a pangram or not. b. Take a list, say for example this one: a = [1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89] and write a program that prints out all the elements of the list that are less than 5.
4	**Write the program for the following:** a.Write a program that takes two lists and returns True if they have at least one common member. b. Write a Python program to print a specified list after removing the 0th, 2nd, 4th and 5th elements. c. Write a Python program to clone or copy a list
5	**Write the program for the following:** a.Write a Python script to sort (ascending and descending) a dictionary by value. b. Write a Python script to concatenate following dictionaries to create a new one. Sample Dictionary : dic1={1:10, 2:20} dic2={3:30, 4:40} dic3={5:50,6:60} Expected Result : {1: 10, 2: 20, 3: 30, 4: 40, 5: 50, 6: 60} c. Write a Python program to sum all the items in a dictionary.
6	**Write the program for the following:** a.Write a Python program to read an entire text file. b. Write a Python program to append text to a file and display the text. c. Write a Python program to read last n lines of a file.
7	**Write the program for the following:**

	a.Design a class that store the information of student and display the same b. Implement the concept of inheritance using python c. Create a class called Numbers, which has a single class attribute called MULTIPLIER, and a constructor which takes the parameters x and y (these should all be numbers). i. Write a method called add which returns the sum of the attributes x and y. ii. Write a class method called multiply, which takes a single number parameter a and returns the product of a and MULTIPLIER. iii. Write a static method called subtract, which takes two number parameters, b and c, and returns b - c. iv. Write a method called value which returns a tuple containing the values of x and y. Make this method into a property, and write a setter and a deleter for manipulating the values of x and y.
8	**Write the program for the following:** Open a new file in IDLE ("New Window" in the "File" menu) and save it as geometry.py in the directory where you keep the files you create for this course.Then copy the functions you wrote for calculating volumes and areas in the "Control Flow and Functions" exercise into this file and save it. Now open a new file and save it in the same directory. You should now be able to import your own module like this: import geometry Try and add print dir(geometry) to the file and run it. Now write a function pointyShapeVolume(x, y, squareBase) that calculates the volume of a square pyramid if squareBase is True and of a right circular cone if squareBase is False. x is the length of an edge on a square if squareBase is True and the radius of a circle when squareBase is False. y is the height of the object.First use squareBase to distinguish the cases. Use the circleArea and squareArea from the geometry module to calculate the base areas. b. Write a program to implement exception handling.
9	**Write the program for the following:** a.Try to configure the widget with various options like: bg="red", family="times", size=18 b. Try to change the widget type and configuration options to experiment with other widget types like Message, Button, Entry, Checkbutton, Radiobutton, Scale etc.
10	**Design the database applications for the following:** Design a simple database application that stores the records and retrieve the same. b. Design a database application to search the specified record from the database. c. Design a database application to that allows the user to add, delete and modify

	the records

9. **Evaluation Pattern:**
   a. **Total Marks** : 150 Marks (10 Point Grading)
   b. **Passing Criteria** : 40 % (4 Grade Points)
   c. **Marking Scheme** : 60:40:50 Pattern
      - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
      - 40 Marks - Internal Assessment (Passing = 16 Marks)
      - 50 Marks - Practical Assessment (Passing = 20 Marks)
   d. **Mode of Evaluation of Answer-books** : Online/Offline

10. **Paper Pattern:**
   a. **Internal Assessment:**
      - Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
      - Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

   b. **Semester End Theory Examination :**

Question No.	**Description**	**Marks**
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

   c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11. **Course Outcome:**

Upon successful completion of this course, Learner should be able to:

CO1: Explain basic principles of Python programming language

CO2: Implement object oriented concepts.

CO3: Implement database and GUI applications.

12. **References:**

1.Think Python Allen Downey O'Reilly 1st 2012

2. An Introduction to Computer Science using Python 3 Jason Montojo, Jennifer Campbell, Paul Gries SPD 1st 2014.

3. Python GUI Programming Cookbook Burkhard A. Meier Packt 2015

4.Fundaments of Database System by Ramez Elmasri and Shamkant B. Navathe, 7th Edition, Pearson Education India, 2010

5. Object-oriented Programming in Python Michael H. Goldwasser, David Letscher Pearson Prentice Hall 1st 2008.

COURSE STRUCTURE

1. **Title of the Course:** Statistics
2. **Semester**: II
3. **Course Code: For Theory:** BDSAE205
   For Practical: BDSAEP205
4. **Course Objective:**
   This course aims
   a. To equip the learners with a working knowledge of statistics and modelling in the presence of uncertainties.
   b. To understand the concept of hypothesis and significance tests.
   c. To help the students to develop an intuition and an interest for random phenomena.
   d. To introduce both theoretical issues and applications that may be useful in real life.
5. **Category of Course**: Ability Enhancement
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSAE205	Statistics	5	3	2	2	4

Module	Detailed Content	Hours
1	**Descriptive Statistics:** Populations, Samples and Processes, Pictorial and tabular methods in Descriptive Statistics, Measures of Location, Measures of variability. **Point Estimation:** Introduction, General concepts of point estimation, Methods for point estimation.	12
2	**Statistical Intervals Based on a single sample:** Introduction, Basic properties of Confidence intervals, Large-sample confidence intervals for a Population mean and proportion, Intervals based on normal population distribution, Confidence intervals for the variance and standard deviation of a normal population. **Tests of hypothesis based on a single sample:** Hypothesis and test procedures, Tests about population mean, Tests concerning a population proportion, P-values, Selecting a test.	12

3	**Inferences based on Two Samples:** Z-test for difference between two population means, T-tests, Analysis for paired data, Inference concerning a difference between population proportions, Inferences concerning two population variances.	12
4	**The Analysis of Variance:** One-way ANOVA, Multiple Comparisons, More in One-way ANOVA. **Multifactor Analysis of variance:** Two-way ANOVA, Three-way ANOVA, 2^p factorial experiments.	12
5	**Distribution-free procedures:** The Wilcoxon Signed-rank test, The Wilcoxon Rank-sum test, Non parametric Confidence intervals, Non parametric ANOVA **Simple Linear Regression and Correlation:** The simple linear regression model, Estimating model parameters, Inferences about model parameters, Prediction of future values, Correlation.	12
	Total	60

Sr. No.	**List of Practical**
1.	Execute the basic commands, array, list and frames.
2.	Create a Matrix and Perform the operations addition, inverse, transpose and multiplication operations.
3.	Execute the statistical functions: mean, median, mode, quartiles, range, inter quartile range histogram.
4.	Calculate the standard deviation, variance, co-variance.
5.	Execute the statistical functions: perform Z test.
6.	Perform t test.
7.	Perform the hypothetical testing.
8.	Perform the Chi-squared Test.
9.	Perform the correlation.
10.	Perform the Linear Regression.

9. **Evaluation Pattern:**
   a. **Total Marks:** 150 Marks (10 Point Grading)
   b. **Passing Criteria**: 40 % (4 Grade Points)
   c. **Marking Scheme:** 60:40:50 Pattern
      - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
      - 40 Marks - Internal Assessment (Passing = 16 Marks)
      - 50 Marks - Practical Assessment (Passing = 20 Marks)

d. **Mode of Evaluation of Answer-books:** Online/Offline

10. Paper Pattern:

a. **Internal Assessment:**

- Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
- Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

b. **Semester End Theory Examination:**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each.

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11. Course Outcome:

On successful completion of this course, the Learner should be able to:

CO1: Distinguish between quantitative and categorical data.

CO2: Apply different statistical measures on data.

CO3: Identify, formulate and solve problems on Statistics and Hypothesis.

CO4: Use Correlation and Regression and their fundamental applications.

12. References:

1. Fundamental of Mathematical Statistics by S.C. Gupta & V.K. Kapoor, 11th Revised Edition, Sultan Chand and Sons, 2011.
2. Mathematical Statistics by J.N. Kapur & H.C. Saxena, 12th Revised Edition, S. Chand, 2005.
3. Introduction to Probability & Statistics by J.Susan Milton & Jesse C. Arnold, 4th Edition, Tata McGraw Hill, 2007.
4. R for Everyone: Advanced Analytics and Graphics by Jared P. Lander, 2nd Edition, O'Reilly, 2017.

COURSE STRUCTURE

1. **Title of the Course :** Environmental Science
2. **Semester** : II
3. **Course Code: For Theory:** BDSID206
4. **Course Objective:**
   a. To make students aware about environment and various issues related to it.
   b. The course will provide brief introduction of various topic as pollution, sustainable development, environment and economic etc.
   c. Developing an attitude of concern for the environment.
   d. Motivating public to participate in environment protection and environment improvement.
   e. Acquiring skills to help the concerned individuals in identifying and solving environmental problems.
   f. Striving to attain harmony with Nature.
5. **Category of Course**: Inter-disciplinary Course
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSID 206	Environmental Science	5	-	2	-	2

Module	Detailed Content	Hours
1	**Introduction to environmental studies:** Multidisciplinary nature of environmental studies; Scope and importance; Concept of sustainability and sustainable development. **Ecosystems :** What is an ecosystem? Structure and function of ecosystem; Energy flow in an ecosystem: food chains, food webs and ecological	12

	succession. Case studies of the following ecosystems: a) Forest ecosystem b) Grassland ecosystem c) Desert ecosystem d) Aquatic ecosystems (ponds, streams, lakes, rivers, oceans, estuaries)	
2	**Natural Resources : Renewable and Non-renewable Resources:** a) Land resources and land-use change; Land degradation, soil erosion and desertification. b) Deforestation: Causes and impacts due to mining, dam building on environment, forests, biodiversity and tribal populations. c) Water: Use and over-exploitation of surface and ground water, floods, droughts, conflicts over water (international & interstate). d) Energy resources: Renewable and non renewable energy sources, use of alternate energy sources, growing energy needs, case studies. **Biodiversity and Conservation :** a) Levels of biological diversity : genetic, species and ecosystem diversity; Biogeographic zones of India; Biodiversity patterns and global biodiversity hot spots b) India as a mega biodiversity nation; Endangered and endemic species of India c) Threats to biodiversity: Habitat loss, poaching of wildlife, man-wildlife conflicts, biological invasions; Conservation of biodiversity: In-situ and Ex-situ conservation of biodiversity. d) Ecosystem and biodiversity services: Ecological, economic, social, ethical, aesthetic and Informational value.	12
3	**Environmental Pollution :** a) Environmental pollution : types, causes, effects and controls; Air, water, soil and noise pollution b) Nuclear hazards and human health risks c) Solid waste management: Control measures of urban and industrial waste. d) Pollution case studies.	12
4	**Environmental Policies & Practices:** a) Climate change, global warming, ozone layer depletion, acid rain	12

	and impacts on human communities and agriculture b) Environment Laws: Environment Protection Act; Air (Prevention & Control of Pollution) Act; Water (Prevention and control of Pollution) Act; Wildlife Protection Act; Forest Conservation Act. International agreements: Montreal and Kyoto protocols and Convention on Biological Diversity (CBD). c) Nature reserves, tribal populations and rights, and human wildlife conflicts in Indian context.	
5	**Human Communities and the Environment :** a) Human population growth: Impacts on environment, human health and welfare. b) Resettlement and rehabilitation of project affected persons; case studies. c) Disaster management: floods, earthquake, cyclones and landslides. d) Environmental movements: Chipko, Silent valley, Bishnois of Rajasthan. e) Environmental ethics: Role of Indian and other religions and cultures in environmental conservation. f) Environmental communication and public awareness, case studies (e.g., CNG vehicles in Delhi). **Field work :** a) Visit to an area to document environmental assets: river/ forest/ flora/fauna, etc. b) Visit to a local polluted site-Urban/Rural/Industrial/Agricultural. c) Study of common plants, insects, birds and basic principles of identification. d) Study of simple ecosystems---pond, river, Delhi Ridge, etc.	12
	Total	60

9. **Evaluation Pattern:**
   a. **Total Marks** : 10
   b. 0 Marks (10 Point Grading)
   c. **Passing Criteria** : 40 % (4 Grade Points)
   d. **Marking Scheme** : 60:40:50 Pattern
      - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
      - 40 Marks - Internal Assessment (Passing = 16 Marks)

e. **Mode of Evaluation of Answer-books** : Online/Offline

10. Paper Pattern:

a. **Internal Assessment:**

- Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
- Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

b. **Semester End Theory Examination :**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

11. Course Outcome:

Upon successful completion of this course, students should be able to:

CO1: Understand the eco-system and need to protect it.

CO2: Understand various danger to environment and how to protect it.

CO3: Appreciate the ethical, cross-cultural, and historical context of environmental issues and the links between human and natural systems.

CO4: Understand core concepts and methods from ecological and physical sciences and their application in environmental problem-solving.

CO5: Reflect critically on their roles, responsibilities, and identities as citizens, consumers and environmental actors in a complex, interconnected world.

12. References:

1. This Fissured Land: An Ecological History of India by Gadgil, M., & Guha, R. Univ. of California Press 1993.
2. Principles of Conservation Biology by Groom, Martha J., Gary K. Meffe, and Carl Ronald Carroll.Sunderland: Sinauer Associates, 2006.
3. Fundamentals of Ecology by Odum, E.P., Odum, H.T. & Andrews, J. Philadelphia:

Saunders 1971.
4. Environmental and Pollution Science by Pepper, I.L., Gerba, C.P. & Brusseau, M.L. Academic Press 2011.
5. Environment by Raven, P.H., Hassenzahl, D.M. & Berg, L.R. 2012. 8th edition. John Wiley & Sons.
6. Ecology, Environmental Science and Conservation by Singh, J.S., Singh, S.P. and Gupta, S.R. S. Chand Publishing, New Delhi. 2014.

Bachelor of Science in Data Science

[**B. Sc. In Data Science**]

Semester - III

COURSE STRUCTURE

1. **Title of the Course :** Database Systems
2. **Semester** : III
3. **Course Code: For Theory: BDSCC301**
   For Practical: BDSCCP301
4. **Course Objective:**
   a. To present an introduction to database management systems, with an emphasis on how to organize, maintain and retrieve - efficiently, and effectively - information from a DBMS.
   b. To understand the different issues involved in the design and implementation of a database system.
   c. To study the physical and logical database designs, database modelling, relational, hierarchical, and network models.
   d. To understand and use data manipulation language to query, update, and manage a database.
   e. To develop an understanding of essential DBMS concepts such as: database security, integrity and concurrency.
   f. To design and build a simple database system and demonstrate competence with the fundamental tasks involved with modelling, designing, and implementing a DBMS.
   g. Develop efficient **PL/SQL** programs to access **Oracle** databases.
   h. Design modular applications using packages, procedures and functions.
5. **Category of Course** : Core Course
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSCC 301	Database System	5	3	2	2	4

Module	Detailed Content	Hours
1	**Introduction to Databases**:-What is database system, purpose of database system, view of data, relational databases, database architecture and different types of databases.	12

	Data Models: - The importance of data models, Basic building blocks, Business rules, The evolution of data models, Degrees of data abstraction.	
2	**Database design, ER Diagram and unified modelling language:-** Database design and ER model: overview, ER model, Constraints, ER design, ER issues, weak entity sets, Codd's rules, Relational schemas, Introduction to UML **Relational database model: -** Logical view of data, keys, integrity rules. **Relational Database design: -** features of good relational database design, atomic domain and Normalization (1NF, 2NF, 3NF, BCNF).	12
3	**Relational Algebra and Calculus:-** **Relational algebra:** introduction, Selection and projection, set operations, renaming, Joins, Division, syntax, semantics. Operators, grouping and ungrouping, relational comparison. **Calculus:** Tuple relational calculus, Domain relational Calculus, calculus vs. algebra.	12
4	**Constraints, Views and SQL** **Constraints**:-What are constraints, types, Integrity constraints. **Views:-** Introduction to views, data independence, security, updates on views, comparison between tables and views **SQL:**-data definition, aggregate function, Null Values, nested sub queries, Joined relations. Triggers.	12
5	**Transaction management and Concurrency control:** What is transaction, ACID properties, serializability and concurrency control, Lock based concurrency control (2PL, Deadlocks), Time stamping methods, optimistic methods and Database recovery management. **PL-SQL :** Beginning with PL/SQL, identifies and keywords, Operators Expressions, Sequences, Control structures, Cursors and Transactions, Collections and composite data types, procedures and functions, Exception handling, packages, with clause and hierarchical retrieval, triggers	12
	Total	60

Sr. No.	**List of Practical**
1	Design a Database and create required tables. For e.g. Bank, College Database
2	Apply the constraints like Primary Key, Foreign key, NOT NULL to the tables.
3	Write a SQL statement for implementing ALTER,UPDATE and DELETE

4	Write the queries to implement the joins.
5	Write the query for implementing the following functions: MAX(),MIN(),AVG(),COUNT()
6	Write the query to implement the concept of Integrity constrains
7	Write the query to create the views.
8	Perform the queries for triggers.
9	Perform the following operation for demonstrating the insertion , updation and deletion using the referential integrity constraints
10	Write the query for creating the users and their role.
11	Creation of Synonyms, Sequence, Indexes, WITH Clause, Hierarchical retrieval.
12	Study of PL/ SQL block.
13	Creation of Procedures & Implementation of Functions.
14	Write a PL/SQL block that handles types of Cursors, Cursor Variables , attributes & loops.
15	Implementation of Triggers – Row level and Statement level triggers.

9. **Evaluation Pattern:**
   a. **Total Marks** : 150 Marks (10 Point Grading)
   b. **Passing Criteria** : 40 % (4 Grade Points)
   c. **Marking Scheme** : 60:40:50 Pattern
      - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
      - 40 Marks - Internal Assessment (Passing = 16 Marks)
      - 50 Marks - Practical Assessment (Passing = 20 Marks)
   d. **Mode of Evaluation of Answer-books** : Online/Offline

10. **Paper Pattern:**
   a. **Internal Assessment:**
      - Assessment consists of a class test of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
      - Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

   b. **Semester End Theory Examination :**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10

5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11. Course Outcome:

Upon successful completion of this course, students should be able to:

CO1: Describe the fundamental elements of relational database management systems. Improve the database design by normalization.

CO2: Explain the basic concepts of relational data model, entity-relationship model, relational database design, relational algebra and SQL.

CO3: Design ER-models to represent simple database application scenarios

CO4: Convert the ER-model to relational tables, populate relational database and formulate SQL queries on data.

CO5: Develop an understanding of the differences between OODBMS, ORDBMS and RDBMS and the practical implications of each approach.

CO6: Analyse, design and develop a real database application using DBMS.

CO7: Explore the use of Procedures, Functions, Packages, and Triggers.

CO8: Examine the characteristics of PL/SQL and how it is used to extend and automate SQL to administer the Oracle database.

12. References:

1. Database System and Concepts By Abraham Silberschatz and Henry Korth and S. Sudarshan , 6th Edition, McGraw-Hill, 2011
2. Database System- Design, Implementation and Management by Peter Rob and Carlos Coronel , 7th Edition , Cengage Learning , 2007
3. Database Management Systems by Raghu Ramakrishnan and Johannes Gehrke, 3rd Edition, McGraw Hill, 2003
4. Fundaments of Database System by Ramez Elmasri and Shamkant B. Navathe, 7th Edition, Pearson Education India, 2010
5. Murach's Oracle SQL and PLSQL by Joel Murach, Murach and Associates.
6. Oracle Database 11g PL/SQL Programming Workbook, ISBN :9780070702264, By : Michael McLaughlin, John Harper, Tata McGrawHill. "IT Savvy: What Top Executives Must Know to Go from Pain to Gain" by Peter Weill. Harvard Business Press, 2009.
7. Oracle PL/SQL Programming, Fifth Edition By Steven Feuerstein, Bill Pribyl.

COURSE DETAILS

1) **Title of the Course: Sampling Methods**

2) **Course Code: For Theory : BDSCC302**
For Practical: BDSCCP302

3) **Course Objective:**

1) Distinguish the two foci of sampling theory.

2) Identify the circumstances the make sampling unnecessary and the reason they re rare.

3) Identify the relation between the desired sample, the obtained sample, the sampling frame, and sample quality.

4) Define and distinguish probability and nonprobability sampling.

5) Define the major types of probability sampling method and indicate when each is preferred.

6) Explain when nonprobability sampling methods may be preferred.

7) Describe the concept of sampling error and explain how its size is affected by the number of cases sampled, the heterogeneity of the population, and the fraction of population included in the sample.

4) **Category of Course :** Core Course

5) **Semester :** III

6) **Total Hours:** 60 lectures

7) **Total Credits:** 04 Credits (02 Credits for Theory & 02 Credits for Practical)

8) **Modules:-**

Course Code	**Course Name**	**Teaching Scheme (Hours /Week)**		**Credits Assigned**		
		Theory	**Practical /Tutorial**	**Theory**	**Practical /Tutorial**	**Total**
BDSCC302	Sampling Methods	5	3	2	2	4

Unit	Details	Lectures
I	1) Introduction to Sampling • Introduction • A brief history of sampling • Sampling concepts • Guidelines of good sampling 2) Defining and framing the population • Defining population units • Setting population boundaries • Obtaining a list • Problems with lists • Coping with omission • Coping with ineligibles • Coping with duplications • Coping with clustering • Framing population without a list	**12**

II	3) Drawing the sample and executing the research • Drawing the sample • Simple random sample • Systematic sampling • Physical sampling • Executing the research • Controlling non-response bias • Calculating response rates 4) Setting sample size • Sampling error illustrated • Sample size based on confidence intervals • Sample size based on hypothesis testing power • Sample size based on value of information • Informal methods of setting sample size	**12**
III	5) Stratified Sampling • When should stratified sampling be used? • Other uses of stratified sampling • How to draw a stratified sample 6) Cluster sampling • When are cluster samples appropriate? • Increased sample variability as a result of clustering • Optimum cluster size • Defining clusters • How to draw a cluster sample?	**12**

IV	7) Estimating population characteristics from samples • Weighing sample data • Using models to guide sampling and estimation • Using models to cope with violations of probability sampling assumptions • Conclusions about the use of models • Measuring the uncertainty of estimates from complex or nonprobability samples 8) Sampling in special context • Sampling for online research • Sampling visitors to a place • Sampling rare populations • Sampling organizational populations • Sampling groups such as influence groups or elites • Panel sampling • Sampling in international context • Big data and survey sampling • Incorporating smartphones, social media and technological changes	**12**
V	9) Evaluating Samples • The sample report • How good must the sample be?	**12**
	Total	**60**

Practical List :-

10 practical's covering the entire syllabus must be performed. The detailed list of practical will be circulated later in the official workshop.

9) Evaluation Pattern:

a. **Total Marks** : 150 Marks (10 Point Grading)
b. **Passing Criteria** : 40 % (4 Grade Points)
c. **Marking Scheme** : 60:40:50 Pattern
   - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)

- 40 Marks - Internal Assessment (Passing = 16 Marks)
- 50 Marks - Practical Assessment (Passing = 20 Marks)

d. **Mode of Evaluation of Answer-books** : Online/Offline

10) Paper Pattern:

a. **Internal Assessment:**

- Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
- Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

b. **Semester End Theory Examination :**

Question No.	**Description**	**Marks**
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11) Course Outcome:

By the completion of this course, learners will be able to:

CO1) define principal concepts about sampling.

CO2) Explains the advantages of sampling.

CO3) Lists the stages of sampling process.

CO 4) Categorizes and defines the sampling methods.

CO 5) apply the Simple Random Sampling (SRS) method.

CO 6) Expresses sample select process on SRS.

CO 7) apply the Systematic Sampling (SS) and Cluster Sampling (CS) methods.

CO 8) Expresses sample select process on SS and CS methods

CO 9) Relates SS to the other sampling method.

CO 10) Estimates the parameters using SS method.

CO 11) Applies SS method to strata.

CO12) Estimates the parameters for equal and unequal cluster sizes.

12) References:

1) Survey Sampling Methods Paperback by Laishram Ladusingh (Author)

2) “Categorical Data Analysis” by A Agresti

3) “Regression Analysis – Theory, Methods and Applications” by A A Sen and M Srivastava

4) “Regression Analysis – Theory, Methods and Applications” by A A Sen and M Srivastava

5) “Nonparametric Statistical Methods” by M Hollander and D A Wolfe

COURSE DETAILS

1) **Title of the Course: Data Visualization**
2) **Course Code: For Theory : BDSCC303**

For Practical: BDSCCP303

3) **Course Objective:**

1) to explore sources
2) to tell stories
3) to predict sales volumes
4) to identify areas that need attention or improvement
5) to understand what factors influence customers' behaviour
6) to know which products to place where
7) to discover how to increase revenues or reduce expenses
8) spreadsheets are hard to visualize
9) patterns and trends can be spotted quickly and easily
10) saves time and energy

4) **Category of Course :** Core Course

5) **Semester :** III

6) **Total Hours:** 60 lectures

7) **Total Credits:** 04 Credits (02 Credits for Theory & 02 Credits for Practical)

8) **Modules:-**

Course Code	**Course Name**	**Teaching Scheme (Hours /Week)**		**Credits Assigned**		
		Theory	**Practical / Tutorial**	**Theory**	**Practical / Tutorial**	**Total**
BDSCC303	Data visualization	5	3	2	2	4

Unit	Details	Lectures
I	Creating visual analytics with tableau desktop: The shortcomings of traditional information analysis. The Business case for visual analysis. Tableau's desktop tools. Introducing the Tableau desktop workspace. Connecting to your data**:** How to connect to your data, Connecting to desktop sources. Understanding the connect page. Connecting to a database Connecting to cloud services ,Connecting to desktop sources Understanding the connect page Connecting to a database Connecting to cloud services problems.	**12**
II	Building your first virtualization: Fast and easy analysis via Show Me. The analytics pane. Creating calculations to enhance data: What is aggression? What are calculated fields and table calculations?	**12**
III	Using maps to improve insights New map features Creating a standard map view. Developing an Ad Hoc analysis environment. Data discovery as a creative process. Providing self service Ad Hoc analysis with parameters	**12**
IV	Tips, tricks and timesavers. Saving time and improving formatting Customizing shapes, colours, fonts, and images. Advanced chart types Bringing it all together with dashboards .How dashboards facilitate analysis and understanding. How tableau improves the dashboard-building process. The wrong way to build a dashboard. The right way to build a dashboard. Building your first advanced dashboard. Sharing your dashboard with tableau reader. Using the tableau performance recorder to improve load speed. Sharing dashboards with tableau online or tableau server	**12**

V	Designing for mobile. The physics of mobile consumption Security considerations for mobile consumption. Offline access Typical mobile usage patterns. Design best practices for mobile consumption. A tablet dashboard example. Mobile authoring and editing. A note on project elastic. Conveying your findings with stories. Turning analysis into insight .Building a story. Formatting story points. Sharing your story point deck. Managing tableau server. Managing published dashboards in tableau server. Navigating tableau server. Organizing reports for consumption. Options for securing reports.Improve efficiency with the data server. Consuming information in the tableau server. Authorizing and editing reports via server. What is required to author reports on the web? Saving and exporting via the web-table environment. Sharing connections, data models, and data extracts. Embedding tableau reports securely on the web.	**12**
	Total	**60**

Practical List :-

10 practical's covering the entire syllabus must be performed. The detailed list of practical will be circulated later in the official workshop.

9) Evaluation Pattern:

a. **Total Marks** : 150 Marks (10 Point Grading)
b. **Passing Criteria** : 40 % (4 Grade Points)
c. **Marking Scheme** : 60:40:50 Pattern
   - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
   - 40 Marks - Internal Assessment (Passing = 16 Marks)
   - 50 Marks - Practical Assessment (Passing = 20 Marks)
d. **Mode of Evaluation of Answer-books** : Online/Offline

10) Paper Pattern:

a. **Internal Assessment:**

- Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
- Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

b. **Semester End Theory Examination :**

Question No.	**Description**	**Marks**
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11) Course Outcome:

By the completion of this course, learners will be able to:

CO1: Design and create data visualizations.

CO2: Conduct exploratory data analysis using visualization.

CO3: Craft visual presentations of data for effective communication.

CO4: Use knowledge of perception and cognition to evaluate visualization design alternatives.

CO5: Design and evaluate color palettes for visualization based on principles of perception.

CO6: Apply data transformations such as aggregation and filtering for visualization.

CO7: Identify opportunities for application of data visualization in various domains.

CO8: Critique existing visualizations based on data visualization theory and principles.

12) References:

1) “Visual Thinking for Design” by Colin Ware.

2) “Semiology of Graphics: Diagrams, Networks, Maps” by Jacques Bertin.

3) “Data Visualization: A Handbook for Data Driven Design” by Andy Kirk.

4) “Show Me the Numbers: Designing Tables and Graphs to Enlighten, Second Edition” by Stephen Few.

5) Tableau your data by Dan Murray, Wiley publication

COURSE STRUCTURE

1. **Title of the Course :** Data Analysis in Excel
2. **Semester** : III
3. **Course Code: For Theory:** BDSSB304
   For Practical: BDSSBP304
4. **Course Objective:**
   a. Add analysis capabilities to Excel spreadsheets and provide foundation to learn about more advanced data analytics with Excel.
   b. To give a working knowledge of Excel with the aim of getting to use it for more advance topics in Business Statistics.
   c. To transform raw data into information that is useful for decision-making.
   d. To learn Descriptive and Inferential Statistics with the help of simple but practical examples.
5. **Category of Course**: Skill Based Course
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSSB 304	Data Analysis in Excel	5	3	2	2	4

Module	Detailed Content	Hours
1	**About variables and values** • Variables and Values • Scales of measurement • Charting numeric variables in excel • Understanding frequency distributions **How values cluster together**	12

	• Calculating the mean • Calculating the median • Calculating the mode • From central tendency to variability	
2	**Variability: How values disperse** • Measuring variability with the range • The concept of standard deviation • Calculating the standard deviation and variance • Bias in estimates and degrees of freedom • Excel's variability functions **Correlation** • Understanding correlation • Using correlation • Using TREND( ) for multiple regression **Charting Statistics** • Characteristics of Excel charts • Histogram charts • Box and Whisker plots	12
3	**Using Excel with the Normal distribution** • About the normal distribution • Excel functions for the normal distribution • Confidence intervals and the normal distribution • The central limit theorem **Telling the truth with statistics** • A context of inferential statistics • Problems with Excel's documentation • The f-test two-sample variance • Reproducibility **Testing difference between means: the basics**	12

	• Testing means: the rationale • Using t-test instead of z-test	
4	**Testing difference between means: further issues** •Using Excel's T.DIST() and T.INV() functions to test hypothesis • Using the T.TEST() function • Using data analysis add-in t-tests **Testing differences between means: The Analysis of Variance** • Why not t-test? • The logic of ANOVA • Using Excel's F worksheet functions • Unequal group sizes • Multiple comparisons procedures	12
5	**Statistical Power** • Controlling the risk • The statistical power of t-tests • The non-centrality parameter in the f-distribution • Calculating the power of f-test **Multiple regression analysis and effect coding: the basics** • Multiple regression and ANOVA • Multiple regression and proportion of variance • Assigning effect coding in excel • Using Excel's regression tool with unequal group sizes • Effect coding, regression, and factorial designs in Excel • Using TREND( ) to replace squared semi-partial correlations	12
	Total	60

List of Practical: 10 practicals covering the entire syllabus must be performed.

9. **Evaluation Pattern:**
   a. **Total Marks** : 150 Marks (10 Point Grading)
   b. **Passing Criteria** : 40 % (4 Grade Points)
   c. **Marking Scheme** : 60:40:50 Pattern
      - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
      - 40 Marks - Internal Assessment (Passing = 16 Marks)
      - 50 Marks - Practical Assessment (Passing = 20 Marks)
   d. **Mode of Evaluation of Answer-books** : Online/Offline

10. **Paper Pattern:**
   a. **Internal Assessment:**
      - Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
      - Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

   b. **Semester End Theory Examination :**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

   c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11. **Course Outcome:**

Upon successful completion of this course, students should be able to:

CO1: Learn about Excel tables and what is their advantage over regular ranges.

CO2: Use formulas to aggregate the data as an alternative to pivot tables for more flexible reporting layouts.
CO3: Summarize data in flexible ways, enabling quick exploration of data and producing valuable insights from the accumulated data.
CO4: Learn about clustering, regression and correlations.
CO5: Use statistical power like Z-test, F-test, T-test and ANNOVA.

12. **Book:** Statistical Analysis: Microsoft excels 2013 by Conrad Carlberg, Pearson publication India.

13. **References:**

1. Excel Data Analysis: Modeling and Simulation by Hector Guerrero 2010th Edition.
2. Microsoft Excel 2019 Data Analysis and Business Modeling by Wayne Winston Microsoft Press,U.S , 6th Edition.

COURSE STRUCTURE

1. **Title of the Course:** Introduction to R and R Studio
2. **Semester**: III
3. **Course Code: For Theory:** BDSAE305
   For Practical: BDSAEP305
4. **Course Objective:**
   a. In this course learner will learn how to program in R and how to use R for effective data analysis.
   b. In this course learner will learn how to install and configure software necessary for a statistical programming environment and describe generic programming language concepts as they are implemented in a high-level statistical language.
   c. The course covers practical issues in computing which includes programming in R, reading data into R, accessing R packages, writing R functions, debugging, profiling R code, and organizing and commenting R code.
5. **Category of Course**: Ability Enhancement
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSAE305	Introduction to R and R Studio	5	3	2	2	4

Module	Detailed Content	Hours
1	**Introduction and Preliminaries:** The R environment, related software and documentation, R and statistics, R and Windows system, Using R interactively, an introductory session, getting help with functions and features, R commands, case sensitivity, recall and correction of previous commands, executing commands from or diverting output to a file, Data permanency and removing objects. **Simple manipulation, numbers and vectors:** Vectors and assignment, vector arithmetic, generating regular sequences, Logical vectors, missing values, character vectors, Index vectors, selecting and modifying subsets of a data set, Other types of objects	12
2	**Objects, their modes and attributes:**	12

	Intrinsic attributes, mode and length, Changing the length of an object, getting and setting attributes, The class of an object. **Ordered and Unordered factors:** The function tapply() and ragged arrays, Ordered factors.	
3	**Arrays and matrices:** Arrays, Array indexing, subsections of an array, Index matrices, The array() function, outer product of 2 arrays, generalized transpose of an array, Matrix facilities, forming partitioned matrices, cbind() and rbind(), The concatenation function, c(), with arrays, Frequency tables from factors. **Lists and data frames:** Lists, constructing and modifying lists, Data frames.	12
4	**Reading data from files:** The read.table() function, The scan() function, Accessing built-in datasets, Editing data. **Probability distribution:** R as a set of statistical tables, Examining the distribution of a set of data. **Grouping, loops and conditional execution:** Grouped expression, Control statements.	12
5	**Writing your own functions:** Simple examples, defining new binary operators, named arguments and defaults, the '…' argument, Assignments within functions, more advanced examples, Scope, Customizing the environment, Classes, generic functions and object orientation. **Graphical procedures:** High-level plotting commands, Low-level plotting commands, interacting with graphics, Using graphic parameters, Graphic parameters list, Device drivers, dynamic graphics.	12
	Total	60

Sr. No.	**List of Practical**
1.	Using R execute the basic commands.
2.	Using R, write a program to understand basic commands on various vector operations.
3.	Using R, write a program for understanding modes and attributes of objects.
4.	Using R, write a program using tapply() function, ragged array, ordered factor.
5.	Using R, create a Matrix using R and Perform the operations addition, inverse, transpose and multiplication operations.
6.	Using R execute the basic commands of list and data frame.
7.	Using R, write a program to read data from a file through various functions.
8.	Using R, write a program to create statistical table and examining the set of data.

9.	Using R, write a program to create a customised environment, class and custom function.
10.	Using R, write a program to understand various graphic plotting commands.

9. **Evaluation Pattern:**
   a. **Total Marks:** 150 Marks (10 Point Grading)
   b. **Passing Criteria**: 40 % (4 Grade Points)
   c. **Marking Scheme:** 60:40:50 Pattern
      - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
      - 40 Marks - Internal Assessment (Passing = 16 Marks)
      - 50 Marks - Practical Assessment (Passing = 20 Marks)
   d. **Mode of Evaluation of Answer-books:** Online/Offline

10. **Paper Pattern:**
   a. **Internal Assessment:**
      - Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
      - Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

   b. **Semester End Theory Examination:**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each.

   c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11. **Course Outcome:**

On successful completion of this course, the Learner should be able to:
CO1: Understand basic concepts such as data type and index and use them in their work.
CO2: Demonstrate use of basic functions.

CO3: Conceptualize and create loops to solve different types of problems.
CO4: Create their own customized functions.
CO5: Construct tables and figures for descriptive statistics.
CO6: Learn to understand new data sets and functions.

12. References:

1. An introduction to R by W.N. Venables, D.M. Smith and the R core team, 2021.
2. Hands on Programming with R: Write Your Own Functions and Simulations by Garrett Gorlemund, 1st Edition, O'Reilly, 2017.
3. R for Everyone: Advanced Analytics and Graphics by Jared P. Lander, 2nd Edition, O'Reilly, 2017.

COURSE STRUCTURE

1. **Title of the Course :** Data Warehouse
2. **Semester** : III
3. **Course Code: For Theory:** BDSEL306
   For Practical: BDSELP306
4. **Course Objective:**
   a. To understand data warehouse concepts, architecture, business analysis and tools
   b. To understand data pre-processing and data visualization techniques
   c. To study algorithms for finding hidden and interesting patterns in data
5. **Category of Course** : Elective
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSEL306	Data Warehouse	5	3	2	2	4

Module	Detailed Content	Hours
1	**Introduction to Data Warehousing:** Introduction, Necessity, Framework of the datawarehouse, options, developing datawarehouses, end points. **Data Warehousing Design Consideration and Dimensional Modeling:** Defining Dimensional Model, Granularity of Facts, Additivity of Facts, Functional dependency of the Data, Helper Tables, Implementation manyto-many relationships between fact and dimensional modelling.	12
2	**An Introduction to Oracle Warehouse Builder:** Installation of the database and OWB, About hardware and operating systems, Installing Oracle database software, Configuring the listener, Creating the database, Installing the OWB standalone software, OWB components and architecture, Configuring the repository and workspaces. **Defining and Importing Source Data Structures:** An overview of Warehouse Builder Design Center, Importing/defining source metadata, Creating a project, Creating a module, Creating an Oracle Database module,	12

	Creating a SQL Server database module, Importing source metadata from a database, Defining source metadata manually with the Data Object Editor, Importing source metadata from files	
3	**Designing the Target Structure:** Data warehouse design, Dimensional design, Cube and dimensions, Implementation of a dimensional model in a database, Relational implementation (star schema), Multidimensional implementation (OLAP),Designing the ACME data warehouse, Identifying the dimensions, Designing the cube, Data warehouse design in OWB, Creating a target user and module, Create a target user, Create a target module, OWB design objects. **Creating the Target Structure in OWB:** Creating dimensions in OWB, The Time dimension, Creating a Time dimension with the Time Dimension Wizard, The Product dimension, Product Attributes (attribute type),Product Levels, Product Hierarchy (highest to lowest),Creating the Product dimension with the New Dimension Wizard, The Store dimension, Store Attributes (attribute type), data type and size, and (Identifier),Store Levels, Store Hierarchy (highest to lowest),Creating the Store dimension with the New Dimension Wizard, Creating a cube in OWB, Creating a cube with the wizard, Using the Data Object Editor	12
4	**Extract, Transform, and Load Basics:** ETL, Manual ETL processes, Staging, To stage or not to stage, Configuration of a staging area, Mappings and operators in OWB, The canvas layout, OWB operators, Source and target operators, Data flow operators, Pre/post-processing operators. **Designing and building an ETL mapping:** Designing our staging area, Designing the staging area contents, Building the staging area table with the Data Object Editor, Designing our mapping, Review of the Mapping Editor, Creating a mapping.	12
5	**ETL: Transformations and Other Operators:** STORE mapping, Adding source and target operators, Adding Transformation Operators, Using a Key Lookup operator, Creating an external table, Creating and loading a lookup table, Retrieving the key to use for a Lookup Operator, Adding a Key Lookup operator, PRODUCT mapping, SALES cube mapping, Dimension attributes in the cube, Measures and other attributes in the cube, Mapping values to cube attributes, Mapping measures' values to a cube, Mapping PRODUCT and STORE dimension values to the cube, Mapping DATE_DIM values to the cube, Features and benefits of OWB. **Validating, Generating, Deploying, and Executing Objects:** Validating, Validating in the Design Center, Validating from the editors, Validating in the Data Object Editor, Validating in the Mapping, Editor, Generating, Generating in the Design Center, Generating from the editors, Generating in the Data Object Editor, Generating in the Mapping Editor, Deploying, The Control Center Service, Deploying in the Design Center	12

	and Data Object Editor, The Control Center Manager, The Control Center Manager window overview, Deploying in the Control Center ,Manager, Executing, Deploying and executing remaining objects, Deployment Order, Execution order	
	Total	60

Sr. No.	List of Practical
1	Importing the source data structures in Oracle
2	Design the target data structure using Oracle
3	Create the target structure in OWB (Oracle Web Builder)
4	Designed and build the ETL mapping 5. Perform the ETL process and transform it to data marts.
5	Perform the ETL process and transform it to data marts. 6. Create the cube and process it in OWB.
6	Create the cube and process it in OWB.
7	Generate the different types of reports in using Oracle.
8	Perform the deployment of Warehouse
9	Create the Pivot table and Pivot chart using some existing data or create the new data.
10	Import the cube in access and create Pivot table and chart.

9. **Evaluation Pattern:**
   a. **Total Marks** : 150 Marks (10 Point Grading)
   b. **Passing Criteria** : 40 % (4 Grade Points)
   c. **Marking Scheme** : 60:40:50 Pattern
      - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
      - 40 Marks - Internal Assessment (Passing = 16 Marks)
      - 50 Marks - Practical Assessment (Passing = 20 Marks)
   d. **Mode of Evaluation of Answer-books** : Online/Offline

10. **Paper Pattern:**
   a. **Internal Assessment:**
      - Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
      - Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

   b. **Semester End Theory Examination :**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11. Course Outcome:

Upon successful completion of this course, students should be able to:

CO1: Understand Data Warehouse fundamentals.

CO2: Design data warehouse with dimensional modelling and apply OLAP operations.

CO3: Identify appropriate data mining algorithms to solve real world problems

CO4: Design a Data warehouse system and perform business analysis with OLAP tools.

12. References:

1. "Data Mining", Ian H. Witten, Eibe Frank and Mark A. Hall, 3rd Edition
2. Introduction to Data Mining by Pang-Ning Tan, Michael Steinbach and Vipin Kumar.
3. "Data Mining Methods", R. Chattamvelli, 2nd Edition.
4. Data Warehousing, Data Mining & OLAP by Alex Berson and Stephen J.Smith Tata McGraw – Hill Edition, 35th Reprint 2016

COURSE STRUCTURE

1. **Title of the Course:** Optimization Techniques
2. **Semester**: III
3. **Course Code: For Theory:** BDSEL307
   For Practical: BDSELP307
4. **Course Objective:**
   a. Introduction to optimization techniques using both linear and non-linear programming.
   b. The focus of the course is on convex optimization, though some techniques will be covered for non-convex function optimization too.
   c. After an adequate introduction to linear algebra and probability theory, learners will learn to frame engineering minima maxima problems in the framework of optimization problems.
5. **Category of Course**: Elective
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSEL307	Optimization Techniques	5	3	2	2	4

Module	Detailed Content	Hours
1	**Mathematical preliminaries:** Linear algebra and matrices, Vector space, eigen analysis, Elements of probability theory, Elementary multivariable calculus.	12
2	**Linear Programming:** Introduction to linear programming model, Simplex method, Duality, Karmarkar's method.	12
3	**Unconstrained optimization:** One-dimensional search methods, Gradient-based methods, Conjugate direction and quasi-Newton methods.	12
4	**Constrained Optimization:** Lagrange theorem, FONC, SONC, and SOSC conditions.	12

5	**Non-linear problems:** Non-linear constrained optimization models, KKT conditions, Projection methods.	12
	Total	60

Sr. No.	**List of Practical**
1.	Perform Matrix operations.
2.	Differentiation of a vector and matrix.
3.	Integration of a vector and matrix.
4.	Perform Simplex algorithm.
5.	Implementation of Newton's method.
6.	Implementation of Secant method.
7.	Implementation of Lagrange multiplier method.
8.	Implementation of KKT theorem.
9.	Implementation of BFGS method

9. **Evaluation Pattern:**
   a. **Total Marks:** 150 Marks (10 Point Grading)
   b. **Passing Criteria**: 40 % (4 Grade Points)
   c. **Marking Scheme:** 60:40:50 Pattern
      - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
      - 40 Marks - Internal Assessment (Passing = 16 Marks)
      - 50 Marks - Practical Assessment (Passing = 20 Marks)
   d. **Mode of Evaluation of Answer-books:** Online/Offline

10. **Paper Pattern:**
   a. **Internal Assessment:**
      - Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
      - Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

   b. **Semester End Theory Examination:**

Question No.	**Description**	**Marks**
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10

4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each.

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11. Course Outcome:

On successful completion of this course, the Learner should be able to:

CO1: Model engineering minima/maxima problems as optimization problems.

CO2: Use various tools/software to implement optimization algorithms.

12. References:

1. An introduction to Optimization by Edwin P K Chong, Stainslaw Zak, 4th Edition, Wiley, 2017.
2. Nonlinear Programming by Dimitri Bertsekas, 2nd Edition, Athena Scientific, 1999.

Bachelor of Science in Data Science

[**B. Sc. In Data Science**]

Semester - IV

COURSE STRUCTURE

1. **Title of the Course :** Introduction to Data Science
2. **Semester** : IV
3. **Course Code: For Theory:** BDSCC401
   For Practical: BDSCCP401
4. **Course Objective:**

   a. To explain idea of data analysis techniques and quantitative modeling for the solution of real world business problems.
   b. To report findings of analysis and effectively present them using data visualization techniques.
   c. To demonstrate knowledge of statistical data analysis techniques utilized in business decision making.
   d. To provide insights about the roles of a Data Scientist, such as a developer, an analyst, a statistical expert etc.
   e. To understand techniques and tools for transformation of data, Data Mining, Data formats, Machine Learning Algorithms, Data Visualization and Optimization.

5. **Category of Course**: Core Course
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSCC401	Data Science	5	3	2	2	4

Module	Detailed Content	Hours
1	**Data Science Technology Stack:** Rapid Information Factory Ecosystem, Data Science Storage Tools, Data Lake, Data Vault, Data Warehouse Bus Matrix, Data Science Processing Tools ,Spark, Mesos, Akka , Cassandra, Kafka, Elastic Search, R ,Scala, Python, MQTT, The Future **Layered Framework:** Definition of Data Science Framework, Cross- Industry Standard Process for Data Mining (CRISP-DM), Homogeneous Ontology for Recursive Uniform Schema, The Top	12

	Layers of a Layered Framework, Layered Framework for High-Level Data Science and Engineering. **Business Layer:** Business Layer, Engineering a Practical Business Layer. **Utility Layer:** Basic Utility Design, Engineering a Practical Utility Layer.	
2	**Three Management Layers:** Operational Management Layer, Processing-Stream Definition and Management, Audit, Balance, and Control Layer, Balance, Control, Yoke Solution, Cause-and-Effect, Analysis System, Functional Layer, Data Science Process. **Retrieve Superstep :** Data Lakes, Data Swamps, Training the Trainer Model, Understanding the Business Dynamics of the Data Lake, Actionable Business Knowledge from Data Lakes, Engineering a Practical Retrieve Superstep, Connecting to Other Data Sources.	12
3	**Assess Superstep:** Assess Superstep, Errors, Analysis of Data, Practical Actions, Engineering a Practical Assess Superstep.	12
4	**Process Superstep :** Data Vault, Time-Person-Object-Location-Event Data Vault, Data Science Process, Data Science. **Transform Superstep :** Transform Superstep, Building a Data Warehouse, Transforming with Data Science, Hypothesis Testing, Overfitting and Underfitting, Precision-Recall, Cross-Validation Test.	12
5	**Transform Superstep:** Univariate Analysis, Bivariate Analysis, Multivariate Analysis, Linear Regression, Logistic Regression, Clustering Techniques, ANOVA, Principal Component Analysis (PCA), Decision Trees, Support Vector Machines, Networks, Clusters, and Grids, Data Mining, Pattern Recognition, Machine Learning, Bagging Data,Random Forests, Computer Vision (CV) , Natural Language Processing (NLP), Neural Networks, TensorFlow. **Organize and Report Supersteps :** Organize Superstep, Report Superstep, Graphics, Pictures, Showing the Difference	12
	Total	60

Sr. No.	**List of Practical**
1	Creating Data Model using Cassandra.
2	Conversion from different formats to HOURS format. a. Text delimited csv format. b. XML c. JSON

	d. MySQL Database e. Picture (JPEG) f. Video g. Audio
3	Utilities and Auditing
4	Retrieving Data
5	Assessing Data
6	Processing Data
7	Transforming Data
8	Organizing Data
9	Generating Reports
10	Data Visualization with Power BI

9. **Evaluation Pattern:**
   a. **Total Marks** : 150 Marks (10 Point Grading)
   b. **Passing Criteria** : 40 % (4 Grade Points)
   c. **Marking Scheme** : 60:40:50 Pattern
      - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
      - 40 Marks - Internal Assessment (Passing = 16 Marks)
      - 50 Marks - Practical Assessment (Passing = 20 Marks)
   d. **Mode of Evaluation of Answer-books** : Online/Offline

10. **Paper Pattern:**
   a. **Internal Assessment:**
      - Assessment consists of a class test of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
      - Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

   b. **Semester End Theory Examination :**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10

3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11. Course Outcome:

Upon successful completion of this course, students should be able to:

CO1: Develop relevant programming abilities.

CO2: Demonstrate proficiency with statistical analysis of data.

CO3: Develop the ability to build and assess data-based models.

CO4: Execute statistical analyses with professional statistical software.

CO5: Apply data science concepts and methods to solve problems in real-world contexts and will communicate these solutions effectively

CO6: Formulate simple algorithms to solve problems, and can code them in a high-level language appropriate for data science work (e.g., Python, SQL, R, Java).

CO7: Integrate data from disparate sources, can transform data from one format to another, and can program data management in relational databases.

12. References:

1. Practical Data Scienceby Andreas François Vermeulen , APress, 2018.
2. Principles of Data Science by Sinan Ozdemir, PACKT, 2016.
3. Data Science from Scratch by Joel Grus, O'Reilly, 2015.
4. Data Science from Scratch first Principle in python by Joel Grus, Shroff Publishers, 2017.
5. Experimental Design in Data science with Least Resources by N C Das, Shroff Publishers, 2018.

COURSE STRUCTURE

1. **Title of the Course :** Artificial Intelligence
2. **Semester** : IV
3. **Course Code: For Theory:** BDSCC402
   For Practical: BDSCCP402
4. **Course Objective:**
   a. To explore the applied branches of artificial intelligence
   b. To enable the student to understand applications of artificial intelligence
   c. To enable the student to solve the problem aligned with derived branches of artificial intelligence
5. **Category of Course** : Core Course
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSCC402	Artificial Intelligence	5	3	2	2	4

Module	Detailed Content	Hours
1	**Introduction:** What is artificial intelligence? Foundations of AI, history, the state of art AI today. **Intelligent agents**: Agents and environment, good behavior, nature of environment, the structure of agent.	12
2	**Solving problems by searching:** Problem solving agents, examples problem, searching for solutions, Uninformed search, informed search strategies, heuristic functions **Beyond classical search:** Local search algorithms, searching with non-deterministic action, searching with partial observations Online search agents and unknown environments	12

3	**Adversarial search:** Games, optimal decision in games, alpha-beta pruning, Stochastic games, partially observable games, state of the art game programs **Logical agents:** Knowledge based agents, the Wumpus world Logic, propositional logic, propositional theorem proving Effective propositional model checking, agents based on propositional logic	12
4	**First order logic::** Syntax and semantics, using first order logic, Knowledge engineering in first order logic. **Inference in first order logic:** Proposition vs first order, unification and lifting, forward and backward chaining, resolution	12
5	**Planning:** Definition of classical planning, algorithms for planning as state space search,Planning graphs, other classical planning approaches, analysis of planning approaches, Time, schedules and resources, hierarchical planning, Planning and acting in nondeterministic domains, multi agent planning. **Knowledge representation:** Categories and objects, events, mental events and objects, reasoning systems for categories, reasoning with default information, internet shopping world.	12
	Total	60

List of Practical	
1.	a.Write a program to implement depth first search algorithm.
	b,Write a program to implement breadth first search algorithm
2.	a.Write a program to simulate 4-Queen / N-Queen problem.
	b.Write a program to solve tower of Hanoi problem.
3.	a.Write a program to implement alpha beta search. b.Write a program for Hill climbing problem.
4.	a. Write a program to implement A* algorithm b. Write a program to implement AO* algorithm.
5.	a. Write a program to solve water jug problem. b. Design the simulation of tic – tac – toe game using min-max algorithm.
6.	a. Write a program to solve Missionaries and Cannibals problem. b. Design an application to simulate number puzzle problem.
7.	a. Write a program to shuffle Deck of cards. b. Solve traveling salesman problem using artificial intelligence technique.
8.	a. Solve the block of World problem. b. Solve constraint satisfaction problem
9.	a. Derive the expressions based on Associative law b. Derive the expressions based on Distributive law

10.	a. Write a program to derive the predicate. (for e.g.: Sachin is batsman , batsman is cricketer) - >Sachin is Cricketer. b. Write a program which contains three predicates: male, female, parent. Make rules for following family relations: father, mother, grandfather,grandmother, brother, sister, uncle, aunt, nephew and niece, cousin. Question: i. Draw Family Tree. ii. Define: Clauses, Facts, Predicates and Rules with conjunction and disjunction

The practicals can be implemented in C / C++ / Java/ Python / R /Prolog / LISP or any other language.

9. Evaluation Pattern:

a. **Total Marks** : 150 Marks (10 Point Grading)

b. **Passing Criteria** : 40 % (4 Grade Points)

c. **Marking Scheme** : 60:40:50 Pattern

- 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
- 40 Marks - Internal Assessment (Passing = 16 Marks)
- 50 Marks - Practical Assessment (Passing = 20 Marks)

d. **Mode of Evaluation of Answer-books** : Online/Offline

10. Paper Pattern:

a. **Internal Assessment:**

- Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
- Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

b. **Semester End Theory Examination :**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11. Course Outcome:

Upon successful completion of this course, students should be able to:

CO1: Be able to use probability and concept of fuzzy sets for solving AI based problems

CO2: Be able to understand the fundamentals concepts of expert system and its applications.

CO3: Be able to understand the applications of Machine Learning. The learner can also apply fuzzy system for solving problems.

CO4: A student can use knowledge representation techniques in natural language processing.

CO5: Student will be able to apply to understand the applications of genetic algorithms in different problems related to artificial intelligence.

12. References:

1. Artificial Intelligence by Saroj Kaushik, 1st, 2019
2. Artificial Intelligence: A Modern Approach by A. Russel, Peter Norvig, 1st, 2019
3. Artificial Intelligence by Elaine Rich,Kevin Knight,Shivashankar B. Nair, 3nd Edition.2019

COURSE DETAILS

1) **Title of the Course: Statistical Consulting**

2) **Course Code: For Theory : BDSCC403**
For Practical: BDSCCP403

3) **Course Objective:**

The goal of this course is to help you develop the skills required of a statistical consultant by completing a substantive consulting project for a client under faculty supervision. Communication and consulting management skills will be emphasized in addition to formal data analysis and problem solving.

4) **Category of Course :** Core Course

5) **Semester :** IV

6) **Total Hours:** 60 lectures

7) **Total Credits:** 04 Credits (02 Credits for Theory & 02 Credits for Practical)

8) **Modules:-**

Course Code	**Course Name**	**Teaching Scheme (Hours /Week)**		**Credits Assigned**		
		Theory	**Practical / Tutorial**	**Theory**	**Practical / Tutorial**	**Total**
BDSCC403	Statistical Consulting	5	3	2	2	4

Unit	Details	Lectures
I	1) Introduction to statistical consulting • History of scientific method • The development of statistics • An overview of statistical consulting • Statistical consulting environments 2) Communication • Verbal interaction • Other aspects of verbal interaction • How to write reports • Basic guidelines for writing • How to make effective presentations • The importance of quality graphics	**12**
II	3) Methodological aspects • Data collection • Data processing • Statistical issues • Statistical methods used in consulting • Standard methods • General methods • Design of experiments • Statistical software 4) A consultation project from A to Z • Prior information • Financial issues • Session I: the first meeting • Documentation • Project analysis • Session II: presenting the results • The final report • Postscript	**12**

III	5) Introduction to the case studies • Presentation format for the case studies • Case study details 6) Case study from group I • Job promotion discrimination • The case of the lost mail • A device to reduce engine emissions • Reverse psychology	**12**
IV	7) Case study from Group II • The flick tail study • Does it have good taste? • Expenditures in NY Municipalities • Measuring quality time 8) Case study from group III • A tale of two thieves • Plastic explosives detection • A market research study • Sales of orthopedic equipment	**12**
V	9) Additional case studies • Improving teaching • Random sampling? • Left or right? • Making horse sense • The tall redhead • Bentley's revenge • Wear what you like? • An AIDS study	**12**
	Total	**60**

Practical List :-

10 practicals' covering the entire syllabus must be performed. The detailed list of practical will be circulated later in the official workshop.

9) Evaluation Pattern:

a. **Total Marks** : 150 Marks (10 Point Grading)

b. **Passing Criteria** : 40 % (4 Grade Points)

c. **Marking Scheme** : 60:40:50 Pattern

- 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
- 40 Marks - Internal Assessment (Passing = 16 Marks)
- 50 Marks - Practical Assessment (Passing = 20 Marks)

d. **Mode of Evaluation of Answer-books** : Online/Offline

10) Paper Pattern:

a. **Internal Assessment:**

- Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
- Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

b. **Semester End Theory Examination :**

Question No.	**Description**	**Marks**
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11) Course Outcome:

By the completion of this course, learners will be able to:

CO1: Students will learn a systematic approach to **statistical consulting** and

CO2: , how to communicate about **statistics** with non-mathematical audiences,

CO3: develop the ability to apply appropriate **statistical** techniques to research questions.

CO4. Be able to identify subject matter questions and translate into statistical language. Ask good questions and restate answers to ensure understanding.

CO5. Know ethical guidelines specific to statisticians.

CO6. Be able to verbally explain statistical concepts.

CO7. Know how to choose a statistical method that is appropriate for answering the client's question.

12) References:

1. Boen JR, Zahn DA (1982) The Human Side Statistical Consulting , Lifetime Learning Publications.

2. Chatfield C (1988) Problem solving: A statistician's guide , Chapman & Hall/CRC Press.

3. J Derr (2000) Statistical Consulting: A Guide to Effective Communication , Duxbury Press.

COURSE STRUCTURE

1. **Title of the Course:** Internet of Things
2. **Semester**: IV
3. **Course Code: For Theory:** BDSSB404
   For Practical: BDSSBP404
4. **Course Objective:**
   a. The aim of this course is to make students aware about 'Internet of Things'-IOT, which is an emerging technology through which all the manual process is to be converted in to system operated process and also integrates with the business.
   b. Learners will understand the concepts of Internet of Things and can able to build IoT applications.
5. **Category of Course**: Skill Based
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSSB404	Internet of Things	5	3	2	2	4

Module	Detailed Content	Hours
1	**The Internet of Things:** An Overview: Flavour of the Internet of Things, the "Internet" of "Things", The Technology of the Internet of Things, Enchanted Objects, who is Making the Internet of Things? **Design Principles for Connected Devices:** Calm and Ambient Technology, Magic as Metaphor, Privacy, Keeping Secrets, Whose Data Is It Anyway? Web Thinking for Connected Devices, Small Pieces, Loosely Joined, First-Class Citizens on The Internet, Graceful Degradation, Affordances. **Internet Principles:** Internet Communications: An Overview, IP, TCP, The IP Protocol Suite (TCP/IP), UDP, IP Addresses, DNS, Static IP Address Assignment, Dynamic IP Address Assignment, IPv6, MAC Addresses, TCP and UDP Ports, An Example: HTTP Ports, Other Common Ports, Application Layer Protocols, HTTP. **HTTPS:** Encrypted HTTP, Other Application Layer Protocols.	12
2	**Thinking About Prototyping:** Sketching, Familiarity, Costs versus Ease of Prototyping, Prototypes and Production,	12

	Changing Embedded Platform, Physical Prototypes and Mass Personalisation, climbing into the Cloud, Open Source versus Closed Source, Why Closed? Why Open? Mixing Open and Closed Source, Closed Source for Mass Market Projects, Tapping into the Community. **Prototyping Embedded Devices:** Electronics, Sensors, Actuators, Scaling Up the Electronics, Embedded Computing Basics, Microcontrollers, System-on-Chips, Choosing Your Platform, Arduino, developing on the Arduino, Some Notes on the Hardware, Openness, Raspberry Pi, Cases and Extension Boards, Developing on the Raspberry Pi, Some Notes on the Hardware, Openness.	
3	**Prototyping the Physical Design:** Preparation, Sketch, Iterate, and Explore, Nondigital Methods, Laser Cutting, Choosing a Laser Cutter, Software, Hinges and Joints, 3D Printing, Types of 3D Printing, Software, CNC Milling, Repurposing/Recycling. **Prototyping Online Components:** Getting Started with an API, Mashing Up APIs, Scraping, Legalities, writing a New API, Clockodillo, Security, Implementing the API, Using Curl to Test, Going Further, Real-Time Reactions, Polling, Comet, Other Protocols, MQ Telemetry Transport, Extensible Messaging and Presence Protocol, Constrained Application Protocol.	12
4	**Techniques for Writing Embedded Code:** Memory Management, Types of Memory, Making the Most of Your RAM, Performance and Battery Life, Libraries, Debugging **Business Models:** A Short History of Business Models, Space and Time, From Craft to Mass Production, The Long Tail of the Internet, Learning from History, The Business Model Canvas, Who Is the Business Model For? Models, Make Thing, Sell Thing, Subscriptions, Customisation, Be a Key Resource, Provide Infrastructure: Sensor Networks, Take a Percentage, Funding an Internet of Things Start-up, Hobby Projects and Open Source, Venture Capital, Government Funding, Crowdfunding, Lean Start-ups.	12
5	**Moving to Manufacture:** What Are You Producing? Designing Kits, Designing Printed circuit boards, Software Choices, The Design Process, Manufacturing Printed Circuit Boards, Etching Boards, Milling Boards. Assembly, Testing, Mass-Producing the Case and Other Fixtures, Certification, Costs, Scaling Up Software, Deployment, Correctness and Maintainability, Security, Performance, User Community.	12

	Ethics: Characterizing the Internet of Things, Privacy, Control, Disrupting Control, Crowdsourcing, Environment, Physical Thing, Electronics, Internet Service, Solutions, The Internet of Things as Part of the Solution, Cautious Optimism, The Open Internet of Things Definition.	
	Total	60

Sr. No.	List of Practical
1.	Starting Raspbian OS, Familiarising with Raspberry Pi Components and interface, Connecting to ethernet, Monitor, USB.
2.	Displaying different LED patterns with Raspberry Pi.
3.	Displaying Time over 4-Digit 7-Segment Display using Raspberry Pi.
4.	Raspberry Pi Based Oscilloscope.
5.	Controlling Raspberry Pi with WhatsApp.
6.	Setting up Wireless Access Point using Raspberry Pi.
7.	Fingerprint Sensor interfacing with Raspberry Pi.
8.	Raspberry Pi GPS Module Interfacing.
9.	IoT based Web Controlled Home Automation using Raspberry Pi.
10.	Visitor Monitoring with Raspberry Pi and Pi Camera.
11.	Interfacing Raspberry Pi with RFID.
12.	Building Google Assistant with Raspberry Pi.
13.	Installing Windows 10 IoT Core on Raspberry Pi.

9. **Evaluation Pattern:**
   a. **Total Marks:** 150 Marks (10 Point Grading)
   b. **Passing Criteria**: 40 % (4 Grade Points)
   c. **Marking Scheme:** 60:40:50 Pattern
      - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
      - 40 Marks - Internal Assessment (Passing = 16 Marks)
      - 50 Marks - Practical Assessment (Passing = 20 Marks)
   d. **Mode of Evaluation of Answer-books:** Online/Offline

10. **Paper Pattern:**
   a. **Internal Assessment:**
      - Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
      - Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

b. **Semester End Theory Examination:**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each.

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11. Course Outcome:

On successful completion of this course, the Learner should be able to:

CO1: Understand the concepts of Internet of Things.

CO2: Analyse basic protocols in wireless sensor network.

CO3: Design IoT applications in different domain and be able to analyse their performance.

CO4: Implement basic IoT applications on embedded platform**.**

12. References:

1. Designing the Internet of Things by Adrian McEwen, Hakim Cassimally, 1st Edition, WILEY, 2014.
2. Internet of Things – Architecture and Design by Raj Kamal, 1st Edition, McGraw Hill, 2017.
3. Getting Started with the Internet of Things by Cuno Pfister, 6th Edition, O'Reilly, 2018.
4. Getting Started with Raspberry Pi by Matt Richardson and Shawn Wallace, 3rd Edition, SPD, 2016.

COURSE STRUCTURE

1. **Title of the Course :** Statistical computing (R)
2. **Semester** : IV
3. **Course Code: For Theory:** BDSAE405
   For Practical: BDSAEP405
4. **Course Objective:**
   a. Computationally intensive statistical methods are a key component to modern data analysis methods.
   b. Student will be able to use statistical software to implement both traditional and state-of-the-art methods in computational statistics as well as recognize situations where these methods are required.
   c. Students should understand the structure of a research study that produces data appropriate for an independent-measures t hypothesis test.
   d. Students should understand the structure of a research study that produces data appropriate for a repeated-measures t hypothesis test.
   e. Processing raw data into formatted data.
5. **Category of Course** : Ability Enhancement
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSAE405	Statistical computing (R)	5	3	2	2	4

Module	Detailed Content	Hours
1	**Drawing statistical conclusions** **S**tatistical inference and study design Measuring uncertainty in randomized experiment Measuring uncertainty in observational studies **Inference using t-distribution** One-sample t-test and the paired t-test A t-ratio for two-sample inference	12

	Inferences in a two-treatment randomized experiment **A closer look at assumptions** Robustness of the two-sample t-tools Resistance of the two-sample t-tools Practical strategies for the two-sample problem Transformations of data	
2	**Alternatives to the t-tools** The rank-sum test Other alternatives for two independent samples Alternatives for paired data **Comparisons among several samples** Comparing any two of the several means The one-way analysis of variance f-test More applications of the extra sums of squares f-test Robustness and model checking **Linear combinations and multiple comparison of means** Inferences about linear combinations of group means Simultaneous inferences Some multiple comparison procedures	12
3	**Simple linear regression: a model for the mean** The simple linear regression model Least squares regression estimation Inferential tools **A closer look at assumptions for simple linear regression** Robustness of least squares inferences Graphical tools for model assessment Interpretation after log transformations Assessment of the fit using analysis of variance	12
4	**Multiple regression** Regression coefficients Specially constructed explanatory variables A strategy for data analysis Graphical methods for data exploration and presentation **Inferential tools for multiple regression** Inferences about regression coefficients Extra sum of squares f-tests **Model checking and refinement** Residual plots A strategy for dealing with influential observation Case-influence statistics Refining the model	12
5	**Strategies for variable selection** Specific issues relating to many explanatory variables	12

	Sequential variable-selection techniques Model selection among all subsets Posterior beliefs about different models Analysis of sex discrimination data **Exploratory tools for summarizing multivariate responses** Linear combination of variables Principal component analysis Canonical correlations analysis Introduction to other multivariate tools **Comparisons of proportion of odds** Inferences for the difference of two proportions Inferences about the ratio of two odds Inference from retrospective studies	
	Total	60

Sr. No.	**List of Practical**
1	Program for Linear regression.
2	Program for Polynomial Regression.
3	Program for multiple linear regression
4	Program for non-linear regression.
5	Import the data from Excel / .CSV and perform the hypothetical testing.

9. **Evaluation Pattern:**
   a. **Total Marks** : 150 Marks (10 Point Grading)
   b. **Passing Criteria** : 40 % (4 Grade Points)
   c. **Marking Scheme** : 60:40:50 Pattern
      - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
      - 40 Marks - Internal Assessment (Passing = 16 Marks)
      - 50 Marks - Practical Assessment (Passing = 20 Marks)
   d. **Mode of Evaluation of Answer-books** : Online/Offline

10. **Paper Pattern:**
   a. **Internal Assessment:**
      - Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
      - Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

b. **Semester End Theory Examination :**

Question No.	**Description**	**Marks**
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11. Course Outcome:

Upon successful completion of this course, students should be able to:

CO1: Students will be able to enter, manipulate and plot data and run basic statistical analyses in R.

CO2: Students will be able to implement estimators for non-standard statistical problems in R.

CO3: Describe the challenges involved in handling Big Data and the strategies used to address these challenges

CO4: Design and implement simulation studies to test and compare statistical methods

12. References:

1. Book: the statistical sleuth: A course in methods of data analysis
2. Maria Rizzo, Statistical Computing with R, Chapman and Hall / CRC (2008).
3. Zuur, A.F., Ieno E.N., & Meesters, E.H.W.G., A Beginner's Guide to R, Springer (2009).

COURSE STRUCTURE

1. **Title of the Course :** Cyber Law
2. **Semester** : IV
3. **Course Code: For Theory:** BDSID406
4. **Course Objective:**

   a. The objective of this course is to enable learner to understand, explore, and acquire a critical understanding of Intellectual Property Rights and Cyber Law.
   b. How to Protect Intellectual Properties and legal procedures about them.
   c. Develop competencies for dealing with frauds and deceptions (confidence tricks, scams) and other cyber crimes for example, Child Pornography etc. that are taking place via the internet.

5. **Category of Course**: Inter-disciplinary Course
6. **Total Hours**: 60
7. **Total Credits**: 02 Credits (02 Credits for Theory)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSEL 406	Cyber Law	5	-	2	-	2

Module	Detailed Content	Hours
1	**Basic Principles and Acquisition of Intellectual Property Rights:** Philosophical Aspects of Intellectual Property Laws, Basic Principles of Patent Law, Patent Application procedure, Drafting of a Patent Specification, Understanding Copyright Law, Basic Principles of Trade Mark, Basic Principles of Design Rights, International Background of Intellectual Property, Unfair Competition.	12
2	**Information Technology Related Intellectual Property Rights:** Computer Software and Intellectual Property-Objective, Copyright Protection, Reproducing, Defences,	12

	Patent Protection. **Database and Data Protection**: Objective, Need for Protection, UK Data Protection Act, 1998, US Safe Harbor Principle, Enforcement. **Protection of Semi-conductor Chips**: Objectives Justification of protection, Criteria, Subject-matter of Protection, WIPO Treaty, TRIPs, SCPA. **Domain Name Protection:** Objectives, domain name and Intellectual Property, Registration of domain names, disputes under Intellectual Property Rights, Jurisdictional Issues, and International Perspective.	
3	**Ownership and Enforcement of Intellectual Property :** Patents-Objectives, Rights, Assignments, Defences in case of Infringement Copyright-Objectives, Rights, Transfer of Copyright, work of employment Infringement, Defences for infringement Trademarks-Objectives, Rights, Protection of good will, Infringement, Passing off, Defences. Designs-Objectives, Rights, Assignments, Infringements, Defences of Design Infringement.	12
4	**Basic Concepts of Technology and Law:** Understanding the Technology of Internet, Scope of Cyber Laws, Cyber Jurisprudence. **Law of Digital Contracts:** The Essence of Digital Contracts, The System of Digital Signatures, The Role and Function of Certifying Authorities, The Science of Cryptography. **Intellectual Property Issues in Cyber Space:** Domain Names and Related issues, Copyright in the Digital Media, Patents in the Cyber World. **Rights of Netizens and E-Governance :** Privacy and Freedom Issues in the Cyber World, E-Governance, Cyber Crimes and Cyber Laws	12
5	**Information Technology Act 2000 :** Information Technology Act-2000-1 (Sec 1 to 13), Information Technology Act-2000-2 (Sec 14 to 42 and Certifying authority Rules), Information	12

	Technology Act-2000-3 (Sec 43 to 45 and Sec 65 to 78), Information Technology Act-2000-4(Sec 46 to Sec 64 and CRAT Rules), Information Technology Act-2000-5 (Sec 79 to 90), Information Technology Act- 2000-6 (Sec 91-94) Amendments in 2008. **International Scenario in Cyber Laws :** Data Protection Laws in EU and USA, Child Abuse Protection Laws in EU and USA, Cyber Laws - the Malaysian Approach.	
	Total	60

---

9. **Evaluation Pattern:**
   a. **Total Marks** : 150 Marks (10 Point Grading)
   b. **Passing Criteria** : 40 % (4 Grade Points)
   c. **Marking Scheme** : 60:40:50 Pattern
      - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
      - 40 Marks - Internal Assessment (Passing = 16 Marks)
      - 50 Marks - Practical Assessment (Passing = 20 Marks)
   d. **Mode of Evaluation of Answer-books** : Online/Offline

10. **Paper Pattern:**
   a. **Internal Assessment:**
      - Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
      - Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

   b. **Semester End Theory Examination :**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10

5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

11. Course Outcome:

Upon successful completion of this course, students should be able to:

CO1: Make learner conversant with the social and intellectual property rights and their issues emerging from cyberspace.

CO2: Explore the legal and policy developments in various countries to regulate cyberspace.

CO3: Develop the understanding of relationship between commerce and cyberspace.

CO4: Give learners in depth knowledge of Information Technology Act and Legal Framework of Right to Privacy, Data Security and Data Protection.

CO5: Make study on various case studies on real time crimes.

12. References:

1. "IT Governance: How Top Performers Manage IT Decision Rights for Superior Results" by Peter Weill, Jeanne Ross. Hardcover – 1 June 2004.
2. "Enterprise Architecture As Strategy: Creating a Foundation for Business Execution" by Jeanne W. Ross. Kindle Edition.
3. "IT Savvy: What Top Executives Must Know to Go from Pain to Gain" by Peter Weill. Harvard Business Press, 2009.
4. "How To Register Your Own Copyright" by Marx Warda, Sphinx Publishing
5. "Managing Intellectual Property: The Strategic Importance" by V. V. Sopale, 2nd edition (PHI).

Bachelor of Science in Data Science

[**B. Sc. In Data Science**]

Semester - V

COURSE STRUCTURE

1. **Title of the Course:** Soft Computing
2. **Semester**: V
3. **Course Code: For Theory:** BDSCC501
   For Practical: BDSCCP501
4. **Course Objective:**
   This course aims
   a. To introduce learners to soft computing concepts and techniques and foster their abilities in designing and implementing soft computing-based solutions for real-world and engineering problems.
   b. To introduce learners to fuzzy systems, fuzzy logic and its applications.
   c. To explain the learners about Artificial Neural Networks and various categories of ANN.
   d. To explain the learners about Genetic Algorithm and various categories of it.
5. **Category of Course**: Core
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSCC501	Soft Computing	5	3	2	2	4

Module	Detailed Content	Hours
1	**Introduction:** What is Soft Computing? Difference between Hard and Soft computing, Requirement of Soft computing, Major Areas of Soft Computing, Applications of Soft Computing. **Introduction to Fuzzy Systems:** Fuzzy Set theory, Fuzzy versus Crisp set, Fuzzy Relation, Fuzzification, Minmax Composition, Defuzzification Method, Fuzzy Logic.	12
2	**Fuzzy Systems:** Fuzzy Rule based systems, Predicate logic, Fuzzy Decision Making, Fuzzy Control Systems, Fuzzy Classification. **Fuzzy Backpropagation Networks:** LR type Fuzzy numbers, Fuzzy Neuron, Fuzzy BP Architecture, Learning in Fuzzy BP, Application of Fuzzy BP Networks.	12
3	**Neural Networks:** What is Neural Network, Learning rules and various activation functions, Single layer Perceptrons, Back	12

	Propagation networks, Architecture of Backpropagation (BP) Networks, Backpropagation Learning, Variation of Standard Back propagation Neural Network, Introduction to Associative Memory, Adaptive Resonance theory and Self Organizing Map, Recent Applications.	
4	**Genetic Algorithm:** History of Genetic Algorithms (GA), Working Principle, Various Encoding methods, Fitness function, GA Operators- Reproduction, Crossover, Mutation, Convergence of GA, Bit wise operation in GA, Multilevel Optimization.	12
5	**GA based Backpropagation Networks:** GA based Weight Determination, K - factor determination in Columns. **Hybrid Systems:** Sequential Hybrid Systems, Auxiliary Hybrid Systems, Embedded Hybrid Systems, Neuro-Fuzzy Hybrid Systems, Neuro-Genetic Hybrid Systems, Fuzzy-Genetic Hybrid Systems.	12
	Total	60

Sr. No.	**List of Practical**
1.	Create a perceptron with appropriate no. of inputs and outputs. Train it using fixed increment learning algorithm until no change in weights is required. Output the final weights.
2.	Create a simple ADALINE network with appropriate no. of input and output nodes. Train it using delta learning rule until no change in weights is required. Output the final weights.
3.	Train the autocorrelator by given patterns: A1=(-1,1,-1,1), A2=(1,1,1,-1), A3=(-1, -1, - 1, 1). Test it using patterns: Ax=(-1,1,-1,1), Ay=(1,1,1,1), Az=(-1,-1,-1,-1).
4.	Train the hetrocorrelator using multiple training encoding strategy for given patterns: A1=(000111001), B1=(010000111), A2=(111001110) B2=(100000001), A3=(110110101) B3(101001010). Test it using pattern A2.
5.	Implement Union, Intersection, Complement and Difference operations on fuzzy sets. Also create fuzzy relation by Cartesian product of any two fuzzy sets and perform maxmin composition on any two fuzzy relations.
6.	Solve Greg Viot's fuzzy cruise controller using MATLAB Fuzzy logic toolbox.
7.	Solve Air Conditioner Controller using MATLAB Fuzzy logic toolbox.
8.	Implement TSP using GA.

9. **Evaluation Pattern:**
   a. **Total Marks:** 150 Marks (10 Point Grading)
   b. **Passing Criteria**: 40 % (4 Grade Points)
   c. **Marking Scheme:** 60:40:50 Pattern
      - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
      - 40 Marks - Internal Assessment (Passing = 16 Marks)

- 50 Marks - Practical Assessment (Passing = 20 Marks)

d. **Mode of Evaluation of Answer-books:** Online/Offline

10. Paper Pattern:

a. **Internal Assessment:**

- Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
- Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

b. **Semester End Theory Examination:**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each.

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11. Course Outcome:

On successful completion of this course, the Learner should be able to:

CO1: Understand soft computing techniques and their role in problem solving.

CO2: Conceptualize and parameterize various problems to be solved through basic soft computing techniques.

CO3: Analyse and integrate various soft computing techniques in order to solve problems effectively and efficiently.

12. References:

1. Neural Networks, Fuzzy Logic & Genetic Algorithm: Synthesis and Applications by S. Rajasekaran & G. A. Vijayalakshmi Pai, Phi, 2003.
2. Soft Computing: Methodologies and Applications by Hoffmann, F., Koeppen, M., Klawonn, F. & Roy, R., Springer, 2005.
3. Principles of Soft Computing by S. N. Sivanandam & S.N. Deepa, Wiley, 2007.
4. Genetic Algorithms by David E. Goldberg, Pearson Education India, 2006.

5. Soft Computing and Its Applications by Rafik Aziz, O. Aliev, R. R. Aliev, World Scientific, 2001.
6. Artificial Neural Networks by B. Yagnanarayana, PHI, 2009.
7. Neural Networks and Learning Machines by Simon O. Haykin, 3rd Edition, Prentice Hall, 2009.

COURSE STRUCTURE

1. **Title of the Course :** Statistical modelling in Python
2. **Semester** : V
3. **Course Code: For Theory:** BDSCC502
   For Practical: BDSCCP502
4. **Course Objective:**
   a. Analyze data using descriptive statistics and graphical tools
   b. Fit a probability distribution to data (estimate distribution parameters)
   c. Express various risk measures as statistical tests
   d. Determine quintile measures of various risk metrics
5. **Category of Course** : Core Course
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSC C502	Statistical modelling in Python	5	3	2	2	4

Module	Detailed Content	Hours
1	**IPython: Beyond normal python** Shell or notebook? Help and documentation in IPython Keyboard shortcut in the IPython shell IPython magic commands Input and output history IPython and shell commands Shell-related magic commands Errors and debugging	12

	Profiling and timing code More IPython resources	
2	**Introduction to NumPy** Understanding data types in Python The basics of NumPy arrays Computation on NumPy arrays: Universal functions Aggregations: Min, max and everything in between Computation on arrays: broadcasting Comparisons, Masks and Boolean logic Fancy indexing Sorting arrays Structured data: NumPy's structured arrays	12
3	**Data manipulation with Pandas** Installing and using Pandas Introducing Pandas objects Data indexing and selection Operating on data in Pandas Handling missing Data Hierarchical indexing Combining datasets: concat and append Combining datasets: merge and join Aggregation and grouping Pivot tables Vectorized string operations Working with time series High performance Pandas: eval() and query()	12
4	Visualization with Matplotlib General Matplotlib tips Two interfaces for the price of one	12

	Simple line plots Simple scatter plots Visualizing errors Density and contour plots Histogram, binning and density Customizing plot legends Customizing colorbars Multiple subplots Text and annotation Customizing ticks Customizing Matplotlib: configurations and stylesheets Three-dimensional plotting in Matplotlib Geographic data with Basemap Visualization with seaborn Further resources	
5	**Machine learning** What is machine learning? Introducing Scikit-Learn Hyperparameters and model validation Feature engineering Naïve Bayes classification Linear regression Support vector machines Decision trees and random forests Principal component analysis K-means clustering	12
	Total	60

List of Practical:

10 practical covering the entire syllabus must be performed. The detailed list of practical will be circulated later in the official workshop.

9. **Evaluation Pattern:**
   a. **Total Marks** : 150 Marks (10 Point Grading)
   b. **Passing Criteria** : 40 % (4 Grade Points)
   c. **Marking Scheme** : 60:40:50 Pattern
      - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
      - 40 Marks - Internal Assessment (Passing = 16 Marks)
      - 50 Marks - Practical Assessment (Passing = 20 Marks)
   d. **Mode of Evaluation of Answer-books** : Online/Offline

10. **Paper Pattern:**
   a. **Internal Assessment:**
      - Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
      - Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

   b. **Semester End Theory Examination :**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

   c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11. **Course Outcome:**

Upon successful completion of this course, students should be able to:

CO1: Ability to derive the distributional results needed for statistical inference.

CO2: Ability to conduct appropriate hypothesis tests for comparing two or more means and for regression.

CO3: Demonstrate understanding that hypothesis tests, regression and analysis of variance can be seen as part of the same statistical theory of linear models.

References:

1. Python Data science handbook by Jake Vanderplas, O'Reilly publication

COURSE STRUCTURE

1. **Title of the Course :** Advances techniques in data science
2. **Semester** : V

Course Code: For Theory: BDSSB503

3. **For Practical:** BDSSBP503
4. **Course Objective:**

   a. Demonstrate knowledge of statistical data analysis techniques utilized in business decision making.
   b. Practice problem analysis and decision-making.
   c. Identify patterns, trends, correlations, and causal relationships in big databases.
   d. Use concepts and methods of mathematical disciplines relevant to data analytics and statistical modelling.

5. **Category of Course**: Core Course
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSSB 503	Advances techniques in data science	5	3	2	2	4

Module	Detailed Content	Hours
1	**1) Statistical learning** • What is statistical learning? • Assessing model accuracy **2) Linear regression** • Simple linear regression • Multiple linear regression • Other considerations in the regression model	12

	• The marketing plan • Comparison of linear regression with K-nearest neighbours	
2	**3) Classification** • An overview of classification • Why not linear regression? • Logistic regression • Linear discriminant analysis • A comparison of classification methods **4) Resampling methods** • Cross-validation • The bootstrap	12
3	**5) Linear model selection and regularization** • Subset selection • Shrinkage methods • Dimension reduction methods • Considerations in high dimensions 6) Tree based methods • The basics of decision trees • Bagging, random forests, boosting	12
4	**7) Support vector machines** • Maximal margin classifier • Support vector classifier • Support vector machines • SVMs with more than two classes • Relationship to logistic regression	12
5	**8) Unsupervised learning** • The challenge of unsupervised learning • Principle component analysis • Clustering methods	12

	9) Project	
	Total	60

List of Practical: 10 practicals covering the entire syllabus must be performed.

Evaluation Pattern:

a. **Total Marks** : 150 Marks (10 Point Grading)
b. **Passing Criteria** : 40 % (4 Grade Points)
c. **Marking Scheme** : 60:40:50 Pattern
   - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
   - 40 Marks - Internal Assessment (Passing = 16 Marks)
   - 50 Marks - Practical Assessment (Passing = 20 Marks)
d. **Mode of Evaluation of Answer-books** : Online/Offline

9. **Paper Pattern:**

a. **Internal Assessment:**
   - Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
   - Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

b. **Semester End Theory Examination :**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

10. **Course Outcome:**

Upon successful completion of this course, students should be able to:

CO1: Utilize statistical concepts of data analysis, data collection, modeling, and inference.

CO2: Use and adapt statistical software packages and scalable computing infrastructure to formulate problems, identify and gather relevant existing data, and analyze the data to provide insights.

CO3: Apply computing theory, languages, and algorithms, as well as mathematical and statistical models, and the principles of optimization to appropriately formulate and use data analyses.

CO4: Execute statistical analyses with professional statistical software.

CO5: Compare the performance of multiple methods and models, recognize the connections between how the data were collected and the scope of conclusions from the resulting analysis, and articulate the limitations and abuses of formal inference and modeling.

CO6: Understand the linear regression model, classification and sampling methods.

11. **References:**

1. Introduction to statistical learning with applications in R, Springer publication
2. Practical Statistics for Data Scientists by Peter Bruce, Paperback – 6 June 2017

COURSE STRUCTURE

1. **Title of the Course :** Capstone Project 1 (R)
2. **Semester** : V
3. **Course Code: For Theory:** BDSAE504
   For Practical: BDSAEP504
4. **Course Objective:**
   a. Students will demonstrate an ability to handle a problem in data science from the point of problem definition through delivery of a solution. In doing so, they will demonstrate proficiency in collecting and processing real-world data, in designing the best methods to solve the problem, in implementing a solution, and quantifying the robustness and accuracy of their model.
   b. Students will demonstrate competence in presenting material by delivering two presentations: a proposal on how to approach the problem and their final solution.
   c. Students will learn how to work in small teams with at least one other student on their project.
   d. Students will write a report on their project for evaluation by the instructor(s) in consultation with the project advisors. The report will be structured as a typical research paper, and hence will include three main sections:
      a. Motivation, problem definition, and existing approaches
      b. Proposed solution and details of implementation
      c. Results, conclusion, and directions for future work
5. **Category of Course** : Ability Enhancement Course
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSAE504	Capstone Project 1 (R)			3	3	6

Course Description

The purpose of the Capstone Project is for the students to apply theoretical knowledge acquired during the Data Science program to a project involving actual data in a realistic setting. During the project, students engage in the entire process of solving a real-world data science project, from collecting and processing actual data to applying suitable and appropriate analytic methods to the problem. Both the problem statements for the project assignments and the datasets originate from real-world domains similar to those that students might typically encounter within industry, government, non-governmental organizations (NGOs), or academic research.

Illustrative project examples

A large insurance company has an anonymized dataset of worker compensation claims. The insurance claims dataset incorporates claimant demographics, claims payments, etc. A team comprised of capstone students, advised by the instructor in conjunction with a technical coach from the company, employ the dataset to develop and implement an analytic solution to reduce workplace injuries using software tools studied in previous courses.

Description of Project Requirements

- Demonstrate ability to carry out a data science project from end to end.
- Demonstrate proficiency in preparation and walk through of a presentation.
- Demonstrate ability to carry out a literature search and summarize the state of the art.
- Demonstrate ability to translate the project objects into a realistic work plan that draws on multiple people.
- Demonstrate ability to design and implement required software using tools such as R, MatLab, Torch, and traditional programming languages such as C, C++, Java.
- Demonstrate ability to professionally present the project plan and results.

9. **Evaluation Pattern:**
   a. **Total Marks** : 200 Marks (10 Point Grading)
   b. **Passing Criteria** : 40 % (4 Grade Points)
   c. **Marking Scheme** : 60:40 Pattern
   d. **Mode of Evaluation of Answer-books** : Online/Offline
   e.

10. **Paper Pattern:**

Semester End Practical Examination:

Exam Duration (in Hours)	Project Demonstration	Project Documentation	Total
2 Hours 30 min per batch	100 Marks	100 Marks	200 Marks

COURSE STRUCTURE

1. **Title of the Course :** Cloud Computing
2. **Semester** : V
3. **Course Code: For Theory:** BDSEL505
   For Practical: BDSELP505
4. **Course Objective:**
   a. To understand data warehouse concepts, architecture, business analysis and tools
   b. To learn how to use Cloud Services.
   c. To implement Virtualization.
   d. To implement Task Scheduling algorithms.
   e. Apply Map-Reduce concept to applications. To build Private Cloud
5. **Category of Course** : Elective
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSEL505	Cloud Computing	5	3	2	2	4

Module	Detailed Content	Hours
1	**Introduction to Cloud Computing:** Introduction, Historical developments, Building Cloud Computing Environments, **Principles of Parallel and Distributed Computing:** Eras of Computing, Parallel v/s distributed computing, Elements of Parallel Computing, Elements of distributed computing, Technologies for distributed computing. **Virtualization**: Introduction, Characteristics of virtualized environments, Taxonomy of virtualization techniques, Virtualization and cloud computing, Pros and cons of virtualization, Technology examples. Logical Network Perimeter, Virtual Server, Cloud Storage Device, Cloud usage monitor, Resource replication, Ready-made environment.	12
2	**Cloud Computing Architecture:** Introduction, Fundamental concepts and models, Roles and boundaries, Cloud Characteristics, Cloud Delivery models, Cloud Deployment models, Economics of the cloud, Open challenges. **Fundamental Cloud Security:** Basics, Threat agents, Cloud	12

	security threats, additional considerations. **Industrial Platforms and New Developments:** Amazon Web Services, Google App Engine, Microsoft Azure.	
3	**Specialized Cloud Mechanisms:** Automated Scaling listener, Load Balancer, SLA monitor, Pay-per-use monitor, Audit monitor, fail over system, Hypervisor, Resource Centre, Multi device broker, State Management Database. **Cloud Management Mechanisms:** Remote administration system, Resource Management System, SLA Management System, Billing Management System, **Cloud Security Mechanisms:** Encryption, Hashing, Digital Signature, Public Key Infrastructure (PKI), Identity and Access Management (IAM), Single 12 11 Sign-On (SSO), Cloud-Based Security Groups, Hardened Virtual Server Images	12
4	**Fundamental Cloud Architectures:** Workload Distribution Architecture, Resource Pooling Architecture, Dynamic Scalability Architecture, Elastic Resource Capacity Architecture, Service Load Balancing Architecture, Cloud Bursting Architecture, Elastic Disk Provisioning Architecture, Redundant Storage Architecture. **Advanced Cloud Architectures:** Hypervisor Clustering Architecture, Load Balanced Virtual Server Instances Architecture, Non-Disruptive Service Relocation Architecture, Zero Downtime Architecture, Cloud Balancing Architecture, Resource Reservation Architecture, Dynamic Failure Detection and Recovery Architecture, Bare-Metal Provisioning Architecture, Rapid Provisioning Architecture, Storage Workload Management Architecture	12
5	**Cloud Delivery Model Considerations:** Cloud Delivery Models: The Cloud Provider Perspective, Cloud Delivery Models: The Cloud Consumer Perspective, **Cost Metrics and Pricing Models:** Business Cost Metrics, Cloud Usage Cost Metrics, Cost Management Considerations, **Service Quality Metrics and SLAs:** Service Quality Metrics, SLA Guidelines	12
	Total	60

Sr. No.	**List of Practical**
1	Write a program for implementing Client Server communication model using TCP. a. A client server based program using TCP to find if the number entered is prime b. A client server TCP based chatting application
2	Write a program for implementing Client Server communication model using UDP a. A client server based program using UDP to find if the number entered is even or odd. b. A client server based program using UDP to find the factorial of the

	entered number. c. A program to implement simple calculator operations like addition, subtraction, multiplication and division. d. A program that finds the square, square root, cube and cube root of the entered number
3	A multicast Socket example
4	Write a program to show the object communication using RMI a. A RMI based application program to display current date and time. b. A RMI based application program that converts digits to words, e.g. 123 will be converted to one two three.
5	Show the implementation of web services. a. Implementing "Big" Web Service. b. Implementing Web Service that connects to MySQL database.
6	Implement Xen virtualization and manage with Xen Center
7	Implement virtualization using VMWare ESXi Server and managing with vCenter
8	Implement Windows Hyper V virtualization
9	Develop application for Microsoft Azure.
10	Develop application for Google App Engine

9. **Evaluation Pattern:**
   a. **Total Marks** : 150 Marks (10 Point Grading)
   b. **Passing Criteria** : 40 % (4 Grade Points)
   c. **Marking Scheme** : 60:40:50 Pattern
      - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
      - 40 Marks - Internal Assessment (Passing = 16 Marks)
      - 50 Marks - Practical Assessment (Passing = 20 Marks)
   d. **Mode of Evaluation of Answer-books** : Online/Offline

10. **Paper Pattern:**
   a. **Internal Assessment:**
      - Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
      - Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

   b. **Semester End Theory Examination :**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10

2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11. Course Outcome:

Upon successful completion of this course, students should be able to:

CO1: Analyze the Cloud computing setup with its vulnerabilities and applications using different architectures.

CO2: Design different workflows according to requirements and apply map reduce programming model.

CO3: Apply and design suitable Virtualization concept, Cloud Resource Management and design scheduling algorithms.

CO4: Create combinatorial auctions for cloud resources and design scheduling algorithms for computing clouds.

C05: Assess cloud Storage systems and Cloud security, the risks involved, its impact and develop cloud application

12. References:

1. Mastering Cloud Computing Foundations and Applications Programming by Rajkumar Buyya,Christian Vecchiola, S. Thamarai Selvi, 2013
2. Cloud Computing Concepts, Technology & Architecture by Thomas Erl, Zaigham Mahmood,and Ricardo Puttini, 2013
3. Distributed and Cloud Computing, From Parallel Processing to the Internet of Things by Kai Hwang, Jack Dongarra, Geoffrey Fox, 2012

COURSE STRUCTURE

1. **Title of the Course :** Big Data
2. **Semester** : V
3. **Course Code: For Theory:** BDSEL506
   For Practical: BDSELP506
4. **Course Objective:**
   a. To provide an overview of an exciting growing field of big data analytics.
   b. To introduce the tools required to manage and analyze big data like Hadoop, NoSql MapReduce.
   c. To teach the fundamental techniques and principles in achieving big data analytics with scalability and streaming capability.
   d. To enable students to have skills that will help them to solve complex realworld problems in for decision support.
5. **Category of Course** : Elective
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSEL506	Big Data	5	3	2	2	4

Module	Detailed Content	Hours
1	**Introduction to Big Data,** Characteristics of Data, and Big Data Evolution of Big Data, Definition of Big Data, Challenges with big data, Why Big data? Data Warehouse environment, Traditional Business Intelligence versus Big Data. State of Practice in Analytics, Key roles for New Big Data Ecosystems, Examples of big Data Analytics. Big Data Analytics, Introduction to big data analytics, Classification of Analytics, Challenges of Big Data, Importance of Big Data, Big Data Technologies, Data Science, Responsibilities, Soft state eventual consistency. Data Analytics Life Cycle	12
2	**Analytical Theory and Methods:** Clustering and Associated Algorithms, Association Rules, Apriori Algorithm, Candidate Rules, Applications of Association Rules, Validation and Testing, Diagnostics, Regression, Linear Regression, Logistic Regression, Additional Regression Models.	12

3	**Analytical Theory and Methods:** Classification, Decision Trees, Naïve Bayes, Diagnostics of Classifiers, Additional Classification Methods, Time Series Analysis, Box Jenkins methodology, ARIMA Model, Additional methods. Text Analysis, Steps, Text Analysis Example, Collecting Raw Text, Representing Text, Term Frequency-Inverse Document Frequency (TFIDF), Categorizing Documents by Topics, Determining Sentiments	12
4	**Data Product**, Building Data Products at Scale with Hadoop, Data Science Pipeline and Hadoop Ecosystem, Operating System for Big Data, Concepts, Hadoop Architecture, Working with Distributed file system, Working with Distributed Computation, Framework for Python and Hadoop Streaming, Hadoop Streaming, MapReduce with Python, Advanced MapReduce. In-Memory Computing with Spark, Spark Basics, Interactive Spark with PySpark, Writing Spark Applications,	12
5	**Distributed Analysis and Patterns**, Computing with Keys, Design Patterns, Last-Mile Analytics, Data Mining and Warehousing, Structured Data Queries with Hive, HBase, Data Ingestion, Importing Relational data with Sqoop, Injesting stream data with flume. Analytics with higher level APIs, Pig, Spark's higher level APIs	12
	Total	60

Sr. No.	List of Practical
1-10	10 Practical based on above syllabus, covering entire syllabus

9. **Evaluation Pattern:**
   a. **Total Marks** : 150 Marks (10 Point Grading)
   b. **Passing Criteria** : 40 % (4 Grade Points)
   c. **Marking Scheme** : 60:40:50 Pattern
      - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
      - 40 Marks - Internal Assessment (Passing = 16 Marks)
      - 50 Marks - Practical Assessment (Passing = 20 Marks)
   d. **Mode of Evaluation of Answer-books** : Online/Offline

10. **Paper Pattern:**
   a. **Internal Assessment:**
      - Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
      - Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

b. **Semester End Theory Examination :**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11. Course Outcome:

Upon successful completion of this course, students should be able to:

CO1: Understand the key issues in big data management and its associated applications in intelligent business and scientific computing.

CO2: Acquire fundamental enabling techniques and scalable algorithms like Hadoop, Map Reduce and NO SQL in big data analytics

CO3: Interpret business models and scientific computing paradigms, and apply software tools for big data analytics

CO4: Achieve adequate perspectives of big data analytics in various applications like recommender systems, social media applications etc.

12. References:

1. Data Analytics with Hadoop An Introduction for Data Scientists by Benjamin Bengfort and Jenny Kim, 2016
2. Big Data and Hadoop by V.K Jain, 2018

Bachelor of Science in Data Science

[**B. Sc. In Data Science**]

Semester - VI

COURSE STRUCTURE

1. **Title of the Course :** Data Mining
2. **Semester** : VI
3. **Course Code: For Theory:** BDSCC601
   For Practical: BDSCCP601
4. **Course Objective:**
   a. To identify the scope and essentiality of Data Mining.
   b. Students will be train on fundamentals of these techniques by exploring various data mining algorithms.
   c. This course will introduce students the various techniques and methods used for data mining
5. **Category of Course** : Core Course
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSCC601	Data Mining	5	3	2	2	4

Module	Detailed Content	Hours
1	**Data Mining: W**hat is data mining?, Statistical limits on data mining, MapReduce and the new software stack,Distributed file systems. **MapReduce :** Algorithms using MapReduce**,** The communication Cost model**,** Complexity theory for MapReduce	12
2	**Finding similar items:** Application of near neighbor search, Shingling of documents, Similarity-preserving summaries of sets Locality- sensitive hashing for documents, Distance measures, The theory of Locality-sensitive functions, LSH families for other distance measures, Application of locality-sensitive hashing, Methods for high degrees of similarity, **Mining Data streams:** The stream data model, Sampling data in a stream, Filtering streams, Counting distinct elements in a stream, Estimating moments, Counting ones in a window, Decaying windows,	12
3	**Link Analysis:** PageRank, Efficient computation of PageRank, Topic sensitive PageRank, Link spam, Hubs and authorities,	12

	Frequent Itemsets: The market-based model, Market baskets and the A-priory algorithm, Handling larger datasets in the main memory, Limited-pass algorithm, Counting frequent items in a stream,	
4	**Clustering:** Introduction to clustering techniques, Hierarchical clustering, K-means algorithms, The CURE algorithm, Clustering in non-Euclidean spaces, Clustering for streams and parallelism, **Advertising on the Web:** Issues on the On-Line advertising, On-Line algorithms, The matching problem, The Adwords problem, Adwords implementation,	12
5	**Recommendation systems:** A model for recommendation systems, Content-based recommendations, Collaborative filtering, Dimensionality reduction, The Netflix challenge, **Mining social network graphs:** Social networks as graphs, Clustering of social-network graphs, Direct discovery of communities, Simrank, Counting triangles, Neighborhood properties of graphs, **Dimensionality reduction:** Eigenvalues and Eigenvectors, Principal-component analysis, Singular-value decomposition, CUR decomposition, **Large-scale machine learning:** The machine-learning model, Perceptrons, Support vector machines, Learning from nearest neighbors, Comparison of learning methods,	12
	Total	60

Sr. No.	List of Practical
	List of Practical: 10 practicals covering the entire syllabus must be performed. The detailed list of practical will be circulated later in the official workshop

9. Evaluation Pattern:

a. **Total Marks** : 150 Marks (10 Point Grading)

b. **Passing Criteria** : 40 % (4 Grade Points)

c. **Marking Scheme** : 60:40:50 Pattern

- 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
- 40 Marks - Internal Assessment (Passing = 16 Marks)
- 50 Marks - Practical Assessment (Passing = 20 Marks)

d. **Mode of Evaluation of Answer-books** : Online/Offline

10. Paper Pattern:

a. **Internal Assessment:**

- Assessment consists of a class tests of 20 marks. The class test is to be conducted

when approx. 40% syllabus is completed. Test will be of one hour.

- Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

b. **Semester End Theory Examination :**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11. Course Outcome:

Upon successful completion of this course, students should be able to:

CO1: Understand Data Mining Principles

CO2: Identify appropriate data mining algorithms to solve real world problems

CO3: Compare and evaluate different data mining techniques like classification, prediction, clustering and association rule mining

CO4: Describe complex data types with respect to spatial and web mining. 6. Benefit the user experiences towards research and innovation.

12. References:

1. "Data Mining", Ian H. Witten, Eibe Frank and Mark A. Hall, 3rd Edition
2. Introduction to Data Mining by Pang-Ning Tan, Michael Steinbach and Vipin Kumar.
3. "Data Mining Methods", R. Chattamvelli, 2nd Edition.

COURSE STRUCTURE

1. **Title of the Course :** Business Intelligence
2. **Semester** : VI
3. **Course Code: For Theory:** BDSCC602

   For Practical: BDSCCP602
4. **Course Objective:**
   a. Data extraction: Investigate data to establish new relationships and patterns
   b. Predictive Analytic and Predictive Modelling: Analyse the correlation between different variables
   c. Logistic Regression: Analyze the possibility of default and generate customer records
   d. Problem analysis: Understand and explore problems in business
   e. Data interpretation: Use tools such as Excel and open source to interpret data
   f. Problem-solving: Use analytics to solve business problems

5. **Category of Course** : Core Course
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSCC602	Business Intelligence	5	3	2	2	4

Module	Detailed Content	Hours
1	**Business intelligence:** Effective and timely decisions, Data, information and knowledge, The role of mathematical models, Business intelligence architectures, Ethics and business intelligence **Decision support systems**: Definition of system, Representation of the decision-making process, Evolution of information	12

	systems, Definition of decision support system, Development of a decision support system	
2	**Mathematical models for decision making:** Structure of mathematical models, Development of a model, Classes of models **Data mining**: Definition of data mining, Representation of input data , Data mining process, Analysis methodologies Data preparation: Data validation, Data transformation, Data reduction	12
3	**Classification:** Classification problems, Evaluation of classification models, Bayesian methods, Logistic regression, Neural networks, Support vector machines **Clustering:** Clustering methods, Partition methods, Hierarchical methods, Evaluation of clustering models	12
4	**Business intelligence applications:** Marketing models: Relational marketing, Sales force management, And Logistic and production models: Supply chain optimization, Optimization models for logistics planning, Revenue management systems. **Data envelopment analysis:** Efficiency measures, Efficient frontier, The CCR model, Identification of good operating practices	12
5	**Knowledge Management:** Introduction to Knowledge Management, Organizational Learning and Transformation, Knowledge Management Activities, Approaches to Knowledge Management, Information Technology (IT) In Knowledge Management, Knowledge Management Systems Implementation, Roles of People in Knowledge Management **Artificial Intelligence and Expert Systems:** Concepts and Definitions of Artificial Intelligence, Artificial Intelligence Versus Natural Intelligence, Basic Concepts of Expert Systems, Applications of Expert Systems, Structure of Expert Systems, Knowledge Engineering, Development of Expert Systems	12
	Total	60

Sr. No.	List of Practical
1	Import the legacy data from different sources such as (Excel, SqlServer, Oracle etc.) and load in the target system. (You can download sample database such as Adventureworks, Northwind, foodmart etc.)
2	Perform the Extraction Transformation and Loading (ETL) process to construct the database in the Sqlserver.
3	a. Create the Data staging area for the selected database. b. Create the cube with suitable dimension and fact tables based on ROLAP, MOLAP and HOLAP model.
4	A.Create the ETL map and setup the schedule for execution. b. Execute the MDX queries to extract the data from the data warehouse.
5	a. Import the data warehouse data in Microsoft Excel and create the Pivot table and Pivot Chart. b. Import the cube in Microsoft Excel and create the Pivot table and Pivot Chart to perform data analysis
6	Apply the what – if Analysis for data visualization. Design and generate necessary reports based on the data warehouse data.
7	Perform the data classification using classification algorithm.
8	Perform the data clustering using clustering algorithm
9	Perform the Linear regression on the given data warehouse data.
10	Perform the logistic regression on the given data warehouse data.

9. Evaluation Pattern:

a. **Total Marks** : 150 Marks (10 Point Grading)

b. **Passing Criteria** : 40 % (4 Grade Points)

c. **Marking Scheme** : 60:40:50 Pattern

- 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
- 40 Marks - Internal Assessment (Passing = 16 Marks)
- 50 Marks - Practical Assessment (Passing = 20 Marks)

d. **Mode of Evaluation of Answer-books** : Online/Offline

10. **Paper Pattern:**

a. **Internal Assessment:**

- Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
- Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

b. **Semester End Theory Examination :**

Question No.	**Description**	**Marks**
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11. **Course Outcome:**

Upon successful completion of this course, students should be able to:

CO1: Describe the concepts and components of Business Intelligence (BI).

CO2: Critically evaluate use of BI for supporting decision making in an organisation.

CO3: Understand and use the technologies and tools that make up BI (e.g. Data warehousing, Data reporting and use of online analytical processing (OLAP)).

CO4: Understand and design the technological architecture that underpins BI systems.

CO5: Plan the implementation of a BI system.

12. **References:**

1. Business Intelligence Data Mining and Optimization for Decision Making (Carlo Vercellis) Wiley 1st 2009
2. Decision support and Business Intelligence Systems (Efraim Turban, Ramesh Sharda, Dursun Delen) Pearson 9th 2011
3. Fundamentals of Business Intelligence (Grossmann W, Rinderle-Ma)

COURSE STRUCTURE

1. **Title of the Course :** Machine Learning in Python
2. **Semester** : VI
3. **Course Code: For Theory :** BDSSB603
   For Practical: BDSSBP603
4. **Course Objective:**
   a. To The objective of this course is to introduce machine learning fundamentals to students.
   b. This course provides introductory concepts of various machine learning techniques to students which will help to build foundation for further understanding.
   c. This course also aims to provide details of various steps involved in machine learning pipeline such as data collection, pre- processing, feature engineering etc.
   d. This course also introduce popular tools used in the area of machine learning
5. **Category of Course** : Skill Based
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSSB503	Machine Learning in Python	5	3	2	2	4

Module	Detailed Content	Hours
1	**Giving computers the ability to learn from Data:** Building intelligent machines to transform data into knowledge The three different types of machine learning Introduction to the basic terminology and notations A roadmap to building machine learning systems Using Python for machine learning **Training simple machine learning algorithms for classification:**Artificial neurons- a brief glimpse into the early history of machine learning Implementation of perceptron learning algorithm in python Adaptive linear neurons and the convergence of learning	12
2	A **tour of machine learning classifiers using scikit-learn:** Choosing a classification algorithmFirst steps with scikit learn – training a perceptron Modelling class probabilities via logistic regression Maximum margin classification with support vector machines Solving non-linear problems using a kernel SVM, Decision tree learning	12

	K-nearest neighbors- a lazy learning algorithm **Building good training sets – Data preprocessing:** Dealing with missing data, Handling categorical data Partitioning a dataset into separate training and test sets Bringing features onto the same scale, Selecting meaningful features, Assessing feature importance with random forests	
3	**Compressing Data via dimensionality reduction:** Unsupervised dimensionality reduction via principal component analysis Supervised data compression via linear discriminant analysis Using kernel principal component analysis for nonlinear mappings **Learning best practices for model evaluation and hyper parameter tuning:**Streamlining workflows with pipelines Using k-fold cross-validation to assess model performance Debugging algorithms with learning and validation curves Fine-tuning machine learning models via grid search Looking at different performance evaluation metrics Dealing with class imbalance	12
4	**Combining different models for ensemble learning:** Learning with ensembles Combining classifiers via majority vote Bagging- building an ensemble of classifiers from bootstrap samplesLeverage weak learners via adaptive boosting **Applying machine learning to sentiment analysis:** Prepare the IMDB movie review data for text processing Introducing the bag of words model Training a logistic regression model for document classification Working with bigger data- online algorithms and out of core learning Topic modelling with Latent Dirichlet Allocation **Predicting continuous target variables with regression analysis:** Introducing linear regression Exploring the housing dataset Implementing an ordinary least square linear regression model Fitting a robust regression model using RANSAC Evaluating the performance of linear regression models Using regularization methods for regression Turning a linear regression model into a curve- polynomial regression Dealing with nonlinear relationship using random forests	12
5	**Working with Unlabeled data – clustering analysis:** Grouping objects by similarity using k-means Organizing clusters as a hierarchical tree Locating regions of high density via DBSCAN **Implementing a multilayer artificial neural network from scratch:** Modelling, complex functions with artificial neural networks**,** Classifying handwritten digits**,**Training an artificial neural network.About the convergence in neural networks**,** A few words about the neural network implementation **Parallelizing neural network training with TensorFlow:**	12

	TensorFlow and training performance, Training neural network efficiently with high-level TensorFlow APIs, Choosing activation functions for multilayer networks	
	Total	60

Sr. No.	List of Practical
	List of Practical: 10 practicals covering the entire syllabus must be performed. The detailed list of practical will be circulated later in the official workshop

9. Evaluation Pattern:

a. **Total Marks** : 150 Marks (10 Point Grading)

b. **Passing Criteria** : 40 % (4 Grade Points)

c. **Marking Scheme** : 60:40:50 Pattern

- 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
- 40 Marks - Internal Assessment (Passing = 16 Marks)
- 50 Marks - Practical Assessment (Passing = 20 Marks)

d. **Mode of Evaluation of Answer-books** : Online/Offline

10. Paper Pattern:

a. **Internal Assessment:**

- Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
- Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

b. **Semester End Theory Examination :**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11. **Course Outcome:**

Upon successful completion of this course, students should be able to develop application

CO1: Understand the various processes involve in machine learning.
CO2: Perform data cleaning and pre-processing
CO3: Decide and classify the problem as classification, prediction or clustering
CO4: Train and test machine learning algorithms

12. **References:**

1. Understanding machine learning: From theory to algorithms, by Shalev-Shwartz, Shai, and Shai Ben-David, 2014.
2. Practical machine learning tools and techniques, by Ian H., et al. 2016

COURSE STRUCTURE

1. **Title of the Course :** Capstone Project 1 (R)
2. **Semester** : VI
3. **Course Code: For Theory:** BDSAE604
   For Practical: BDSAEP604
4. **Course Objective:**
   a. Students will demonstrate an ability to handle a problem in data science from the point of problem definition through delivery of a solution. In doing so, they will demonstrate proficiency in collecting and processing real-world data, in designing the best methods to solve the problem, in implementing a solution, and quantifying the robustness and accuracy of their model.
   b. Students will demonstrate competence in presenting material by delivering two presentations: a proposal on how to approach the problem and their final solution.
   c. Students will learn how to work in small teams with at least one other student on their project.
   d. Students will write a report on their project for evaluation by the instructor(s) in consultation with the project advisors. The report will be structured as a typical research paper, and hence will include three main sections:
      a. Motivation, problem definition, and existing approaches
      b. Proposed solution and details of implementation
      c. Results, conclusion, and directions for future work
5. **Category of Course** : Ability Enhancement Course
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSA E604	Capstone Project 2 (python)			3	3	6

Course Description

The purpose of the Capstone Project is for the students to apply theoretical knowledge acquired during the Data Science program to a project involving actual data in a realistic setting. During the project, students engage in the entire process of solving a real-world data science project, from collecting and processing actual data to applying suitable and appropriate analytic methods to the problem. Both the problem statements for the project assignments and the datasets originate from real-world domains similar to those that students might typically encounter within industry, government, non-governmental organizations (NGOs), or academic research.

Illustrative project examples

A large insurance company has an anonymized dataset of worker compensation claims. The insurance claims dataset incorporates claimant demographics, claims payments, etc. A team comprised of capstone students, advised by the instructor in conjunction with a technical coach from the company, employ the dataset to develop and implement an analytic solution to reduce workplace injuries using software tools studied in previous courses.

Description of Project Requirements

- Demonstrate ability to carry out a data science project from end to end.
- Demonstrate proficiency in preparation and walk through of a presentation.
- Demonstrate ability to carry out a literature search and summarize the state of the art.
- Demonstrate ability to translate the project objects into a realistic work plan that draws on multiple people.
- Demonstrate ability to design and implement required software using tools such as Python
- Demonstrate ability to professionally present the project plan and results.

9. **Evaluation Pattern:**
   a. **Total Marks** : 200 Marks (10 Point Grading)
   b. **Passing Criteria** : 40 % (4 Grade Points)
   c. **Marking Scheme** : 60:40 Pattern
   d. **Mode of Evaluation of Answer-books** : Online/Offline
   e.

10. **Paper Pattern:**

Semester End Practical Examination:

Exam Duration (in Hours)	Project Demonstration	Project Documentation	Total
2 Hours 30 min per batch	100 Marks	100 Marks	200 Marks

COURSE STRUCTURE

1. **Title of the Course:** Data Compression
2. **Semester**: VI
3. **Course Code: For Theory:** BDSEL605
   For Practical: BDSELP605
4. **Course Objective:**
   a. To introduce learners to basic applications, concepts, and techniques of Data Compression.
   b. To develop skills for using recent data compression software to solve practical problems in a variety of disciplines.
   c. To gain experience doing independent study and research.
5. **Category of Course**: Elective
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSEL605	Data Compression	5	3	2	2	4

Module	Detailed Content	Hours
1	**Compression Techniques:** Lossless Compression, Lossy Compression, Measures of Performance. **Mathematical Preliminaries for Lossless Compression Models:** Physical Models, Probability Models, Markov Models, Composite Source Model, Coding Uniquely Decodable Codes, Prefix Codes, Algorithmic Information Theory, Minimum Description Length principle Strings.	12
2	**Huffman Coding:** The Shannon Fano Coding, The Huffman Coding Algorithm, Minimum Variance Huffman Codes, Adaptive Huffman Coding, Application of Huffman Coding.	12
3	**Arithmetic Coding:** Overview, Introduction, coding a Sequence, Generating a Binary Code, Comparison of Huffman and Arithmetic Coding, **Quantization:**	12

	The Quantization Problem, Scalar Quantization, Vector Quantization, Discrete Cosine Transform.	
4	**Dictionary Techniques:** Overview, Introduction, Static Dictionary, Adaptive Dictionary (LZ77, LZ78), LZW.	12
5	**Context-Based Compression:** Overview, Introduction, Prediction with Partial Match (ppm), Dynamic Markov Compression. **Lossless Image Compression:** Overview, Introduction, CALIC, JPEG-LS, Multi resolution Approaches, Facsimile Encoding.	12
	Total	60

Sr. No.	List of Practical
1.	Write a Program to check whether the given code is prefix or not.
2.	Write a program to determine whether the set of given codes is uniquely decodable or not.
3.	Write a program to implement Shannon-Fano Compression Algorithm.
4.	Write a program to implement Huffman Coding Compression Algorithm.
5.	Write a program to implement Arithmetic Coding Compression Algorithm.
6.	Write a program to compress and decompress the given input string.
7.	Write a program to implement LZ77 Compression Algorithm.
8.	Write a program to implement LZ77 Decompression Algorithm.
9.	Write a program to implement LZ78 Compression Algorithm.
10.	Write a program to implement LZ78 Decompression Algorithm.
11.	Write a program to implement LZW Compression Algorithm.
12.	Write a program to implement LZW Decompression Algorithm.

9. **Evaluation Pattern:**
   a. **Total Marks:** 150 Marks (10 Point Grading)
   b. **Passing Criteria**: 40 % (4 Grade Points)
   c. **Marking Scheme:** 60:40:50 Pattern
      - 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
      - 40 Marks - Internal Assessment (Passing = 16 Marks)
      - 50 Marks - Practical Assessment (Passing = 20 Marks)
   d. **Mode of Evaluation of Answer-books:** Online/Offline

10. **Paper Pattern:**
   a. **Internal Assessment:**
      - Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.

- Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

b. **Semester End Theory Examination:**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10
2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each.

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11. Course Outcome:

On successful completion of this course, the Learner should be able to:

CO1: Understand importance of data compression.
CO2: Develop a reasonably sophisticated data compression application.
CO3: Select methods and techniques appropriate for the task.
CO4: Develop the methods and tools for the given task.

12. References:

1. Introduction to Data Compression by Khalid Sayood, 5th Edition, Morgan Kaufmann Publishers, 2017.
2. Data Compression: The Complete Reference by David Saloman, 4th Edition, Springer, 2006.
3. The Data Compression Book by Mark Nelson and Jean-Loup Gaily, 2nd Edition, John Wiley & Sons, 1995.

COURSE STRUCTURE

1. **Title of the Course :** Software Quality Assurance
2. **Semester** : VI
3. **Course Code: For Theory:** BDSEL606
   For Practical: BDSELP606
4. **Course Objective:**

   a. Create and apply a software quality assurance plan for all software projects.
   b. Create and manage a software quality assurance team
   c. Create and maintain appropriate metrics to measure and maintain quality
   d. Introduce basic concepts of software testing.
   e. To understand white box, block box, object oriented and web based testing.
   f. To know in details automation testing and tools used for automation testing.
   g. To understand the importance of software quality and assurance software systems development.

5. **Category of Course**: Elective Course
6. **Total Hours**: 60
7. **Total Credits**: 04 Credits (02 Credits for Theory & 02 Credits for Practical)
8. **Modules:**

Course Code	Course Name	Teaching Scheme (Hours /Week)		Credits Assigned		
		Theory	Practical/ Tutorial	Theory	Practical/ Tutorial	Total
BDSEL 506	Software Quality Assurance	5	3	2	2	4

Module	Detailed Content	Hours
1	**Introduction to Quality:** Historical Perspective of Quality, What is Quality? (Is it a fact or perception?), Definitions of Quality, Core Components of Quality, Quality View, Financial Aspect of Quality, Customers, Suppliers and Processes, Total Quality Management (TQM), Quality Principles of Total Quality Management, Quality Management Through Statistical Process Control, Quality Management Through Cultural Changes, Continual (Continuous) Improvement Cycle, Quality in Different Areas, Benchmarking and Metrics, Problem Solving Techniques, Problem Solving Software	12

	Tools. **Software Quality:** Introduction, Constraints of Software Product Quality Assessment, Customer is a King, Quality and Productivity Relationship, Requirements of a Product, Organisation Culture, Characteristics of Software, Software Development Process, Types of Products, Schemes of Criticality Definitions, Problematic Areas of Software Development Life Cycle, Software Quality Management, Why Software Has Defects? Processes Related to Software Quality, Quality Management System Structure, Pillars of Quality Management System, Important Aspects of Quality Management.	
2	**Fundamentals of testing:** Introduction, Necessity of testing, What is testing? Fundamental test process, The psychology of testing, Historical Perspective of Testing, Definitions of Testing, Approaches to Testing, Testing During Development Life Cycle, Requirement Traceability Matrix, Essentials of Software Testing, Workbench, Important Features of Testing Process, Misconceptions About Testing, Principles of Software Testing, Salient Features of Good Testing, Test Policy, Test Strategy or Test Approach, Test Planning, Testing Process and Number of Defects Found in Testing, Test Team Efficiency, Mutation Testing, Challenges in Testing, Test Team Approach, Process Problems Faced by Testing, Cost Aspect of Testing, Establishing Testing Policy, Methods, Structured Approach to Testing, Categories of Defect, Defect, Error, or Mistake in Software, Developing Test Strategy, Developing Testing Methodologies (Test Plan), Testing Process, Attitude Towards Testing (Common People Issues), Test Methodologies/Approaches, People Challenges in Software Testing, Raising Management Awareness for Testing, Skills Required by Tester, Testing throughout the software life cycle, Software development models, Test levels, Test types, the targets of testing, Maintenance testing	12
3	**Unit Testing:** Boundary Value Testing: Normal Boundary Value Testing, Robust Boundary Value Testing, Worst-Case Boundary Value Testing, Special Value Testing, Examples, Random Testing, Guidelines for Boundary Value Testing. **Equivalence Class Testing:** Equivalence Classes, Traditional Equivalence Class Testing, Improved Equivalence Class Testing, Edge Testing, Guidelines and Observations. **Decision Table–Based Testing:** Decision Tables, Decision Table Techniques, Cause-and-Effect Graphing, Guidelines and Observations.	12

	Path Testing: Program Graphs, DD-Paths, Test Coverage Metrics, Basis Path Testing, Guidelines and Observations. **Data Flow Testing:** Define/Use Testing, Slice-Based Testing, Program Slicing Tools.	
4	**Software Verification and Validation:** Introduction, Verification, Verification Workbench, Methods of Verification, Types of reviews on the basis od Stage Phase, Entities involved in verification, Reviews in testing lifecycle, Coverage in Verification, Concerns of Verification, Validation, Validation Workbench, Levels of Validation, Coverage in Validation, Acceptance Testing, Management of Verification and Validation, Software development verification and validation activities. **V-test Model:** Introduction, V-model for software, Testing during Proposal stage, Testing during requirement stage, Testing during test planning phase, Testing during design phase, Testing during coding, VV Model, Critical Roles and Responsibilities.	12
5	**Levels of Testing:** Introduction, Proposal Testing, Requirement Testing, Design Testing, Code Review, Unit Testing, Module Testing, Integration Testing, Big-Bang Testing, Sandwich Testing, Critical Path First, Sub System Testing, System Testing, Testing Stages. **Special Tests:** Introduction, GUI testing, Compatibility Testing, Security Testing, Performance Testing, Volume Testing, Stress Testing, Recovery Testing, Installation Testing, Requirement Testing, Regression Testing, Error Handling Testing, Manual Support Testing, Intersystem Testing, Control Testing, Smoke Testing, Adhoc Testing, Parallel Testing, Execution Testing, Operations Testing, Compliance Testing, Usability Testing, Decision Table Testing, Documentation Testing, Training testing, Rapid Testing, Control flow graph, Generating tests on the basis of Combinatorial Designs, State Graph, Risk Associated with New Technologies, Process maturity level of Technology, Testing Adequacy of Control in New technology usage, Object Oriented Application Testing, Testing of Internal Controls, COTS Testing, Client Server Testing, Web Application Testing, Mobile Application Testing, eBusiness, eCommerce Testing, Agile Development Testing, Data Warehousing Testing.	12
	Total	60

Sr. No.	List of Practical
1	Setting up a company that sells testing services to software houses.
2	Write a SRS
3	Black Box Testing – Equivalence Partitioning and Boundary value Analysis
4	Black Box Testing – Equivalence Partitioning and Boundary value Analysis
5	Black Box Testing –: Decision table and Cause Effect Graphing
6	Branch – Decision – Condition Coverage
7	State Transition Testing
8	Data Flow Testing
9	Structured Testing – Loop Coverage, Call coverage and Path Coverage.
10	Test Automation using Selenium IDE

9. Evaluation Pattern:

a. **Total Marks** : 150 Marks (10 Point Grading)

b. **Passing Criteria** : 40 % (4 Grade Points)

c. **Marking Scheme** : 60:40:50 Pattern

- 60 Marks - Written/Semester End Exam (Passing = 24 Marks)
- 40 Marks - Internal Assessment (Passing = 16 Marks)
- 50 Marks - Practical Assessment (Passing = 20 Marks)

d. **Mode of Evaluation of Answer-books** : Online/Offline

10. Paper Pattern:

a. **Internal Assessment:**

- Assessment consists of a class tests of 20 marks. The class test is to be conducted when approx. 40% syllabus is completed. Test will be of one hour.
- Students have to submit assignment after completion of each module which will carry 15 marks and 5 marks are for attendance.

b. **Semester End Theory Examination :**

Question No.	Description	Marks
1	Objectives or Short Answers (Covering All Modules)	10

2	Answer any two Questions (Descriptive based on module 1)	10
3	Answer any two Questions (Descriptive based on module 2)	10
4	Answer any two Questions (Descriptive based on module 3)	10
5	Answer any two Questions (Descriptive based on module 4)	10
6	Answer any two Questions (Descriptive based on module 5)	10

Note: Q.2 to Q.6 will include total 4 sub questions having 5 marks each

c. **Semester End Practical Examination:**

Exam Duration (in Hours)	Practical + Oral	Journal	Total
2 Hours 30 min per batch	45 Marks	05 Marks	50 Marks

11. Course Outcome:

Upon successful completion of this course, students should be able to:

CO1: Describe fundamental concepts of software quality assurance.

CO2: Explore test planning and its management.

CO3: Understand fundamental concepts of software automation.

CO4: Apply Selenium automation tool for testing web based application.

CO5: Demonstrate the quality management, assurance, and quality standard to software system.

CO6: Demonstrate Software Quality Tools and analyze their effectiveness.

12. References:

1. Software Testing and Continuous Quality Improvement By William E. Lewis, 3rd Edition, CRC Press, 2016.
2. Software Testing: Principles, Techniques and Tools by M. G. Limaye, TMH , 2017.
3. Foundation of Software Testing by Dorothy Graham, Erik van Veenendaal, Isabel Evans, Rex Black, 3rd Edition, Cengage Learning, 2003.
4. Software Testing: A Craftsman's Approach by Paul C. Jorgenson, 4th Edition, CRC Press, 2017.