Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

(Affiliated to Savitribai Phule Pune University, Pune)

Choice Based Credit System (CBCS)

Proposal to Introduce New Academic Programme

In

B.Sc. (Data Science)

Implemented from

Academic Year 2022 - 23

1. Introduction/ Prologue :

Artificial Intelligence and Data Science Programme prepare students with the skills to perform intelligent data analysis which is a key component in numerous real-world applications. During the past ten years, data science has emerged as one of the most high-growth, dynamic, and lucrative careers in technology. This course aims at providing not only the core technologies such as artificial intelligence, data mining and data modelling but also gives intensive inputs in the areas of machine learning and big data analytics. By the end of this course, the students will gain cross-disciplinary skills across fields such as statistics, computer science, machine learning, and logic, data scientists and will have career opportunities in healthcare, business, eCommerce, social networking companies, climatology, biotechnology, genetics, and other important areas. The major focus of this programme is to equip students with statistical, mathematical reasoning, machine learning, knowledge discovery, and visualization skills.

2. Objective of the Programme:

The Broad goals and objectives of the BSc(Data Science)are as follows :

1.To prepare students to develop strong analytical thinking skills and problem solving abilities that benefits graduate personally and professionally

2. While learning Data Science it is necessary to be aware with the skills of Artificial Intelligence to perform intelligent data analysis which is a key component in numerous real-world applications.

3. This course aims at providing not only the core technologies such as artificial intelligence, data mining and data modelling but also gives intensive inputs in the areas of machine learning and big data analytics and by the end of this course, the students will gain cross-disciplinary skills across fields such as statistics, computer science, machine learning, and logic, data scientists and will have career opportunities in healthcare, business, eCommerce, social networking companies, climatology, biotechnology, genetics, and other important areas.

4. The major focus of this programme is to equip students with statistical, mathematical reasoning, machine learning, knowledge discovery, and data visualization skills.

5. To achieve and demonstrate knowledge of statistical analysis techniques utilized in business decision making .

6. To explore,sort and analyze mega data from various sources in order to take advantage of Data Science applications and reach conclusions to optimize business processes for decision support.

7. To prepare necessary knowledge base for research and development in Computer Science

3. Programme Outcomes:

1. Data Science graduates will be able to design, and develop intelligent business applications to solve various industrial problems.
2. Students can use the latest tools and open source technologies to recommend the required solutions.
3. Students can figure out how to evaluate the ethical, legitimate, proficient and social standards of engineering knowledge and practices.
4. These graduates can also exhibit their domain knowledge in data handling, knowledge extraction, mobile and distributed application development, intelligence web/ecommerce development, database administration, computer hardware, networking, education and training and decision support systems using AI and Data Science tools and techniques.
5. Students will be able to analyze a given dataset and derive insights which provide value to the business and society at large
6. The syllabus also develops requisite professional skills and problem solving abilities for pursuing a career in domains like Healthcare, Business, Finance.
7. B.Sc. (Data Science) graduates can go for higher study in programmes like M.Sc. in Data Science and Big Data and Analytics to further add a professional touch to their knowledge and become ready for the corporate world.
8. Full Time Industry Project – Internship gives hands on experience in solving a real world problem.

4. Program Structure:

- The Program is of a Three Years (Six semesters) Full Time Degree Program.
- The program shall be based on credit system comprising 148 credit points.
- Theory Courses offered shall be of 2 to 3 credits and practical courses of 2 credits each.
- For Theory Course, one credit is equivalent to one clock hour direct teaching in a week and for Practical Course, one credit is equivalent to one and half hours of laboratory work in a week.

5. Eligibility for Admission:

- Any candidate who has passed the XII standard Examination in Science stream from, Maharashtra State Board of Secondary and Higher Secondary Education or equivalent Board of Examination with Mathematics, is eligible for admission to the First Year of this program.

OR

Passed Three Year Diploma Course approved by the DTE, Maharashtra State or Equivalent authority.

- Admission will be given on the basis of entrance examination conducted by the college. Entrance examination consist (a) MCQ Paper of 70 marks and 30 marks for GD-PI.

- Admission is given on the basis of merit. Merit list is prepared by applying the crieteria as :- 50 % weightage is given for marks obtained in 12th Examination and 50 % marks from Entrance Examination.
- Intake : 60

6. **Medium of Instruction**: English.

7. **Award of Credits:**

- Course having 2 credits shall be evaluated out of 50 (35:15) marks and student should secure at least 20 marks (40%)in continuous assessment as well as term end exam to earn full credits of that course.
- Course having 3 credits shall be evaluated out of 75 (50:25) marks and student should secure at least 30 marks (40%) in continuous assessment as well as term end exam to earn full credits of that course.
- GPA shall be calculated based on the marks obtained in the respective subject provided that student should have obtained credits for that course.
- Evaluation Pattern:
    - Each course carrying 50/75 marks shall be evaluated with Continuous Assessment (CA) and University Evaluation (UE) mechanism.
    - For Continuous assessment weight of 30% marks is given and 70% evulation is carried out by external examination.
    - To pass in a course, a student has to secure minimum 40 % marks separately for CA and external examiniation.

8. **ATKT Rules:**

- Minimum number of credits required to take admission to Second Year of B.Sc. Data Science .: 34
- Minimum number of credits required to take admission to Third Year of B.Sc. Data Science.: 72

9. **Completion of Degree Program:**

- A student who earns 148 credits, shall be considered to have completed the requirements of the B.Sc. Data Science degree program and CGPA will be calculated for such student.

10**. Examination Patterns:**

- Continuous Internal Evaluation is periodically conducted by the respective subject teacher.
- Semester End Examination will be conducted as per standard norms of autonomy for end semester examination.

11**. Credit system :**

a. Definitions and Keywords:

1. **Credit Based Semester System**: The degree programme and total credits of the same are divided into semesters. To complete the degree the students, need to complete the prescribed options of credits during all the semesters of an academic programme.
2. **Choice-Based Credit System**: The choice-based credit system provides choices for students to select courses from the range of available programmes.
3. **Academic Programmes**: A range of learning experiences offered to students formally for one to four years leading to a certificate, diploma, or degree. All the specified degrees were notified by the University Grants Commission (UGC) in the Gazette of India on 05 July 2014.
4. **Semester**: Semester consists of 15 to 18 weeks of academic work equivalent to 90 actual teaching days. The odd semester of the academic programme is usually scheduled from June to December and even semesters from January to May.
5. **Academic Year**: The consecutive (One odd and One even) semesters constitute one academic year.
6. **Course:** A course (Paper) is a unit of credits in a formal degree programme.
7. **Credit:** A unit by which the course work is measured. It determines the number of hours of instructions required per week. One credit is equivalent to one hour of teaching or two hours of practical work/fieldwork per week.
8. **Credit Point:** It is a product of grade points and the number of credits of a course.
9. **Grade Point:** It is a numerical weight prescribed to each letter grade on a 10-point scale.
10. **Letter Grade:** Letter grade is an index of the performance of students in a said course. Grades are denoted by letters O, A+, A, B+, B, C, D, and F.
11. **Semester Grade Point Average (SGPA)**: It measures the performance of work done in a semester. It is a ratio of total credit points secured by a student in various courses registered in a semester and the total course credits taken during that semester. It shall be expressed up to two decimal places.
12. **Cumulative Grade Point Average (CGPA):** It measures the overall cumulative performance of overall semesters. The CGPA is the ratio of total credit points secured by a student in various courses in all semesters and the sum of the total credits of all courses in all the semesters. It is expressed up to two decimal places.

13. **Certificate (Marksheet**): Based on the grades earned, a Marksheet will be issued to all the registered students after every semester. The grade certificate will display the course details (code, title, number of credits, grade secured) and SGPA of that semester and CGPA earned till that semester.
14. **Transcript**: The degree transcript will be issued as per the demand of students.

b. Types of Courses:

1. **Discipline-Specific Core Courses (DSCC):** These are the courses compulsorily studied by a student as a core requirement to complete a programme in a said discipline.
2. **Elective Course (DSEC):** Elective course is a course that can be chosen from a group of papers.

   It may be:

   - Supportive to the discipline of study
   - Provides an expanded scope
   - Enables exposure to some other discipline/domain
   - Nurtures student's proficiency/skill.

   Elective Courses may be Discipline Specific Elective Courses (DSEC), Skill Enhancement Courses (SEC) and Generic Elective (GE), or Open Elective.

   a. **Discipline-Specific Elective Course (DSEC):** Elective courses offered by the main discipline/subject of study are referred to as Discipline Specific Elective Courses. The institute may offer discipline-related elective courses of interdisciplinary nature (to be offered by the main discipline/subject of study).
   b. **Generic Elective (GE) Courses:** An elective course chosen from an unrelated discipline/subject, to seek exposure is called a Generic Elective.
   c. **Ability Enhancement Core Courses (AECC):** The Ability Enhancement (AE) Courses may be of two kinds: Ability Enhancement Compulsory Courses (AECC) and Skill Enhancement Courses (SEC). "AECC" courses are the courses based upon the content that leads to knowledge enhancement; i. Environmental Awareness and ii. English/MIL Communication. SEC courses are value-based and/or skill-based aim to provide hands-on training, competencies, skills, etc.

c. Letter Grades and Grade Points

Letter Grade	Meaning	Percentage of Marks	Grade Points
O	Outstanding	90 and Above	10
A+	Excellent	75-89	9
A	Very Good	60-74	8
B+	Good	55-59	7
B	Above Average	50-54	6
C	Average	45-49	5
D	Pass	40-44	4
F	Fail	Less than 40	0
Ab	Absent	Absent	0
-	-	Fx (Detained, Repeat the course)	0
-	-	IC (Incomplete Course - Absent for examination but continue for the course)	0

d. CGPA and Grade

Sr. No.	CGPA (Credit Courses)	Grade
1.	9.50 and above	O (Outstanding)
2.	8.25 to 9.49	A+ (Excellent)
3.	6.75 to 8.24	A (Very Good)
4.	5.75 to 6.74	B+ (Good)
5.	5.25 to 5.74	B (Above Avarage)
6.	4.75 to 5.24	C (Avarage)
7.	4.00 to 4.74	D (Pass)
8.	Less than 4.00	F (Fail)
Sr. No.	CGPA (Non Credit Courses)	Remark
1.		S(Satisfactory)
2.		U(Unsatisfactory)

* The statutory requirement for eligibility to enter as an assistant professor in colleges and universities in the disciplines of arts, science, commerce, etc., is a minimum average mark of 50% and 55% in relevant postgraduate degrees respectively for reserved and general categories. Hence, the cut-off marks for grade B shall not be less than 50% and for grade B+, it should not be less than 55% in CBCS System.

- A student obtaining Grade F shall be considered a fail and will be required to reappear in the examination.
- For non-credit courses 'Satisfactory' or "Unsatisfactory' shall be indicated instead of the letter grade and this will not be counted for the computation of SGPA/CGPA. The students with unsatisfactory remark will be required to reappear in the examination.

e. Computation of SGPA and CGPA

As per the UGC recommendations the following procedure to compute the Semester Grade Point Average (SGPA) and Cumulative Grade Point Average (CGPA):

1. The SGPA is the ratio of the sum of the product of the number of credits with the grade points scored by a student in all the courses taken by a student and the sum of the number of credits of all the courses undergone by a student,

i.e SGPA (Si) = $\Sigma(C_i \times G_i) / \Sigma C_i$

where C_i is the number of credits of the ith course and G_i is the grade point scored by the student in the ith course.

2. The CGPA is also calculated in the same manner taking into account all the courses undergone by a student over all the semesters of a programme,

i.e. CGPA = $\Sigma(C_i \times S_i) / \Sigma C_i$

where S_i is the SGPA of the ith semester and C_i is the total number of credits in that semester.

iii. The SGPA and CGPA shall be rounded off to 2 decimal points and reported in the transcripts.

f. Illustration of Computation of SGPA and CGPA (B. Sc. Data Science)

Class	Subject	Course	Credit	Letter Grade	Grade Point	Credit Point (Grade points X Credits)
B.Sc (D.S)-I Semester I	DSCC - 01 T	Basic Programming in C	02	A	8	8 X 2 =16
	DSCC – 02 T	Database Management System	02	A	8	8 X 2 =16
	DSCC – 03 T	Elements of Information Technology	02	A	8	8 X 2 =16
	DSCC – 04 T	Introduction to R programming	02	B+	7	7 X 2 =14
	DSCC – 05 T	Introduction to Data	02	B+	7	7 X 2 =14
	DSCC – 06 T	Probability and Basic Statistics	02	B+	7	7 X 2 =14
	DSCC – 07 T	Matrix Computations	02	A	8	8 X 2 =16
	DSCC – 08 T	Discrete Mathematics	02	A	8	8 X 2 =16
	DSCC – 09 P	Lab Course on 101	02	A	8	8 X 2 =16
	DSCC – 10 P	Lab Course on 102	02	B+	7	7 X 2 =14
	DSCC – 11 P	Lab Course on 104,106,107,108	02	B+	7	7 X 2 =14
Total			22			166

Method for Computation of SGPA and CGPA

Thus, SGPA =166/22= 7.5

Illustration for CGPA

Class	Semester	Total Credits	SGPA
B.Sc (D.S)-I	I	22	7.5
B.Sc (D.S)-I	II	22	8.5
B.Sc (D.S)-II	III	26	6.9
B.Sc (D.S)-II	IV	26	9.0
B.Sc (D.S)-III	V	22	8.6
B.Sc (D.S)-III	VI	22	8.0
Total		140	64.4

Thus CGPA = 22 x 7.5+22 x 8.5+26 x 6.9+26 x 9.0+22 x 8.6+22 x 8.0/ 140 = 8.07

(Very Good)

12**. Details of credit system:**

The B.Sc. (Data Science) programme is for 3 academic years and 6 semesters. The minimum total number of credits required for each programme is 140 plus 08 additional credits.

Bachelor of Science (Data Science): Structure

Class	Semester	Subjects	Courses	DSCC	AECC	DSEC	SEC**	GE	Project**	Total Credit
B.Sc (D.S)-I	I	1	08T + 03P +	11		-				22
	II	1	08T + 03P +	11		-				22
B.Sc (D.S)-II	III	1	03T + 03P + 02 AECC+ 02 DSEC+ 01 SEC	06	A-Foreign Language B- Critical thinking / Scientific Temper	02 T	01 T			26
	IV	1	03T+ 03P + 02 AECC+ 02 DSEC+ 01 SEC	06	A- Foreign Language B- Environmental Awareness	02 T	01 T			26
B.Sc (D.S)-III	V	1	03T + 02P + 02 DSEC 01 SEC + 01 GE	05		01 T 01 P	01 T	01 T		22
	VI	1	03T + 02 P + 02 DSEC + 01 SEC 01 GE	05		01 T 01P	01 T	01 T		22
Total										140

13**. Credit Distribution**

Class	Semester	Subjects	Courses	DSCC	AECC	DSEC	SEC	GE*	Total Credits
First Year	I	11	08T +03 P	22	00	00	00	00	22
First Year	I	11	08 T +03 P	22	00	00	00	00	22
Second Year	III	11	06 T +03 P	14	04	06	02	00	26
Second Year	IV	11	06 T +03 P	14	04	06	02	00	26
Third Year	V	09	06 T +03 P	13	00	04	03	02	22
Third Year	VI	09	06 T +03 P	13	00	05	02	02	22
Total									140

Note:

- Each theory credit is equivalent to 15 clock hours of teaching.
- The duration of each theory semester is 15-18 weeks in which at least 12-week classroom teaching.

Marking Scheme under Choice Based Credit System for B.Sc (Data Science) program.

First Year –B.Sc.Data Science - Semester -I

Class	Subject	Course Code	Course Name	Credit	Internal Evaluation	External Evaluation	Total Maximum Marks
First Year. Semester I	DSCC-1T	BSC-DS 101 T	Basic Programming in C	02	15	35	50
	DSCC–2T	BSC-DS 102 T	Database Management System	02	15	35	50
	DSCC–3T	BSC-DS 103 T	Elements of Information Technology	02	15	35	50
	DSCC–4T	BSC-DS 104 T	Introduction to R programming	02	15	35	50
	DSCC–5T	BSC-DS 105 T	Introduction to Data	02	15	35	50
	DSCC–6T	BSC-DS 106 T	Probability and Basic Statistics	02	15	35	50
	DSCC–7T	BSC-DS 107 T	Matrix Computations	02	15	35	50
	DSCC–8T	BSC-DS 108 T	Discrete Mathematics	02	15	35	50
	DSCC–9 P	BSC-DS 109 P	Lab Course on 101	02	15	35	50
	DSCC–10 P	BSC-DS 110 P	Lab Course on 102	02	15	35	50
	DSCC– 11 P	BSC-DS111 P	Lab Course on 104,106, 107,108	02	15	35	50
				22	165	385	550

First Year –B.Sc.Data Science - Semester -II

Class	Subject	Course Code	Course Name	Credit	Internal Evaluation	External Evaluation	Total Maximum Marks
First Year. Semester II	DSCC - 12 T	BSC-DS 201 T	Advanced C Programming	02	15	35	50
	DSCC – 13 T	BSC-DS 202 T	Advance Database Concepts	02	15	35	50
	DSCC – 14 T	BSC-DS 203 T	Fundamentals of Data Science	02	15	35	50
	DSCC – 15 T	BSC-DS 204 T	Discrete Probability Distribution	02	15	35	50
	DSCC – 16 T	BSC-DS 205 T	Continous Probability Distribution	02	15	35	50
	DSCC – 17 T	BSC-DS 206 T	Linear Algebra	02	15	35	50
	DSCC – 18 T	BSC-DS 207 T	Operation Research	02	15	35	50
	DSCC – 19 T	BSC-DS 208 T	Numerical Methods	02	15	35	50
	DSCC – 20 P	BSC-DS 209 P	Lab Course on 201	02	15	35	50
	DSCC – 21 P	BSC-DS 210 P	Lab Course on 202	02	15	35	50
	DSCC – 22 P	BSC-DS 211 P	Lab Course on 204, 205, 206,207,208	02	15	35	50
				22	165	385	550

Second Year –B.Sc.Data Science - Semester -III

Class	Subject	Course Code	Course Name	Credit	Internal Evaluation	External Evaluation	Total Maximum Marks
Second Year. Semester III	DSCC –23T	BSC-DS 301T	Data Structures	03	25	50	75
	DSCC– 24T	BSC-DS 302T	Introduction to Artificial Intelligence	03	25	50	75
	DSCC– 25T	BSC-DS 303T	Linear Algebra in Data Science	02	15	35	50
	DSCC–26 P	BSC-DS 304P	Lab Course on 301	02	15	35	50
	DSCC–27 P	BSC-DS 305P	Lab Course on 302 and 307 (DSEC-01) and 308(DSEC-02)	02	15	35	50
	DSCC–28 P	BSC-DS 306P	Lab Course on 303 and 309(SEC-01)	02	15	35	50
	DSEC-01 T	BSC-DS 307T	Introduction to Regression Analysis / Open Course	03	25	50	75
	DSEC-02 T	BSC-DS 308T	Introduction to Time Series / Open Course	03	25	50	75
	AECC-01	AECC-01	MIL:Foreign Language	02	15	35	50
	AECC-02	AECC-02	Critical Thinking / Scientific Temper	02	15	35	50
	SEC-01 T	BSC-DS 309 T	Introduction to Analog and Digital Electronics	02	15	35	50
				26	205	445	650

Second Year –B.Sc.Data Science - Semester -IV

Class	Subject	Course Code	Course Name	Credit	Internal Evaluation	External Evaluation	Total Maximum Marks
Second Year. Semester IV	DSCC –29 T	BSC-DS 401 T	Object Oriented Concepts Using Java Programming	03	25	50	75
	DSCC –30 T	BSC-DS 402 T	NoSQL Databases	03	25	50	75
	DSCC –31 T	BSC-DS 403 T	Artificial Intelligence in Data Science	02	15	35	50
	DSCC –32 P	BSC-DS 404 P	Lab course on 401	02	15	35	50
	DSCC –33 P	BSC-DS 405 P	Lab course on 402	02	15	35	50
	DSCC –34 P	BSC-DS 406 P	Lab course on 403 , 407 T (DSEC-03) and 408 T (DSEC-04)	02	15	35	50
	DSEC-03	BSC-DS 407 T	Statistical Inference / Open Course	03	25	50	75
	DSEC-04	BSC-DS 408 T	Baysian Inference / Open Course	03	25	50	75
	AECC-03	AECC-03	MIL:Foreign Language	02	15	35	50
	AECC-04	AECC-04	Critical Thinking / Scientific Temper	02	15	35	50
	SEC-02	BSC-DS 409 T	Data Communication and Networking	02	15	35	50
				26	205	445	650

Third Year –B.Sc. Data Science - Semester -V

Class	Subject	Course Code	Course Name	Credit	Internal Evaluation	External Evaluation	Total Maximum Marks
Third Year. Semester V	DSCC – 35T	BSC-DS 501T	Python Programming	03	25	50	75
	DSCC – 36T	BSC-DS 502 T	Data Acquisition and Visualization	03	25	50	75
	DSCC – 37T	BSC-DS 503 T	Machine Learning	03	25	50	75
	DSCC –38P	BSC-DS 504 P	Lab Course on 501	02	15	35	50
	DSCC – 39P	BSC-DS 505 P	Lab Course on 502	02	15	35	50
	DSEC-05	BSC-DS 506 T	Programming in SCALA / Open Course	02	15	35	50
	DSEC-06	BSC-DS 507 P	Lab Course on 506 (DSEC-05)	02	15	35	50
	SEC-03	BSC-DS 508 T	Operating System Concepts	03	25	50	75
	GE-01	BSC-DS 509 T	Consumer Behaviour and Marketing Research	02	15	35	50
				22	175	375	550

Third Year –B.Sc.Data Science - Semester –VI

Class	Subject	Course Code	Course Name	Credit	Internal Evaluation	External Evaluation	Total Maximum Marks
Third Year. Semester VI	DSCC – 40T	BSC-DS 601 T	Big Data Security	03	25	50	75
	DSCC – 41T	BSC-DS 602 T	Social Web Analytics	03	25	50	75
	DSCC – 42T	BSC-DS 603 T	Real Time Analysis	03	25	50	75
	DSCC – 43P	BSC-DS 604 P	Lab Course on 601	02	15	35	50
	DSCC –44P	BSC-DS 605 P	Lab Course on 602	02	15	35	50
	DSEC-07	BSC-DS 606 T	Azure for Analytics / Open Course	03	25	50	75
	DSEC-08	BSC-DS 607 P	Lab Course on 606 (DSEC-08)	02	15	35	50
	SEC-04	BSC-DS 608 T	Data Mining and Data Warehousing	02	15	35	50
	GE-02	BSC-DS 609 T	Professional ethics	02	15	35	50
				22	175	375	550

14. Distribution of Internal and External Marks

Class	Semester	Credit	Internal Evaluation	External Evaluation	Total Maximum Marks
First Year B.Sc. DS	I	22	165	385	550
First Year B.Sc. DS	II	22	165	385	550
Second Year B.Sc. DS	III	26	205	445	650
Second Year B.Sc. DS	IV	26	205	445	650
Third Year B.Sc. DS	V	22	175	375	550
Third Year B.Sc. DS	VI	22	175	375	550
	06	140	1090	2410	3500

Additional grade-based credits for all UG Programmes

Sr. No.	Title	Credits	Remark
1.	Democracy, Election and Governance: Semester I	02	Compulsory
2.	Physical Education: Semester II	02	Compulsory
3.	Completion of skill-based certificate programme organized by any department of the college	02	Compulsory
4.	SWAYAM certificate course	02	Optional
5.	Participation in NSS Winter Camp	02	Optional
6.	'C' Certificate in NCC	02	Optional
7.	Selection and participation in RDC parade at New Delhi	02	Optional
8.	Representation at State/ National level Co-curricular Activities	02	Optional
9.	Representation at State/ National level Extra-curricular Activities	02	Optional
10.	Winning Medal/ Prize at International/National level Co-curricular/ Extracurricular activities	02	Optional
11.	Prize in Curricular/ Extracurricular/ Cultural Activities at college level	01	Optional
12.	Active participation in Excursion tours/Study tours and Experiential Learning Activities	01	Optional
13.	Survey Report	02	Optional
14.	Book Review on a book suggested by the Academic Council	02	Optional

15. Programme Structure and Course Titles: (All academic years)

Sr. No.	Class	Semester	Course Code	Course Title	Credits
1	BSC-DS-I	I	BSC-DS 101 T	Basic Programming in C	02
2	BSC-DS-I	I	BSC-DS 102 T	Database Management System	02
3	BSC-DS-I	I	BSC-DS 103 T	Elements of Information Technology	02
4	BSC-DS-I	I	BSC-DS 104 T	Introduction to R programming	02
5	BSC-DS-I	I	BSC-DS 105 T	Introduction to Data	02
6	BSC-DS-I	I	BSC-DS 106 T	Probability and Basic Statistics	02
7	BSC-DS-I	I	BSC-DS 107 T	Matrix Computations	02
8	BSC-DS-I	I	BSC-DS 108 T	Discrete Mathematics	02
9	BSC-DS-I	I	BSC-DS 109 P	Lab Course on 101	02
10	BSC-DS-I	I	BSC-DS 110 P	Lab Course on 102	02
11	BSC-DS-I	I	BSC-DS 111 P	Lab Course on 104,106, 107,108	02
12	BSC-DS-I	II	BSC-DS 201 T	Advanced C Programming	02
13	BSC-DS-I	II	BSC-DS 202 T	Advance Database Concepts	02
14	BSC-DS-I	II	BSC-DS 203 T	Fundamentals of Data Science	02
15	BSC-DS-I	II	BSC-DS 204 T	Discrete Probability Distribution	02
16	BSC-DS-I	II	BSC-DS 205 T	Continous Probability Distribution	02
17	BSC-DS-I	II	BSC-DS 206 T	Linear Algebra	02
18	BSC-DS-I	II	BSC-DS 207 T	Operation Research	02
19	BSC-DS-I	II	BSC-DS 208 T	Numerical Methods	02
20	BSC-DS-I	II	BSC-DS 209 P	Lab Course on 201	02
21	BSC-DS-I	II	BSC-DS 210 P	Lab Course on 202	02

22	BSC-DS-I	II	BSC-DS 211 P	Lab Course on 204, 205, 206,207,208	02
23	BSC-DS-I I	III	BSC-DS 301 T	Data Structures	03
24	BSC-DS-I I	III	BSC-DS 302 T	Introduction to Artificial Intelligence	03
25	BSC-DS-I I	III	BSC-DS 303 T	Linear Algebra in Data Science	02
26	BSC-DS-I I	III	BSC-DS 304 P	Lab Course on 301	02
27	BSC-DS-I I	III	BSC-DS 305 P	Lab Course on 302 and 307 and 308	02
28	BSC-DS-I I	III	BSC-DS 306 P	Lab Course on 303 and 309	02
29	BSC-DS-I I	III	BSC-DS 307 T	Introduction to Regression Analysis / Open Course DSEC-01	03
30	BSC-DS-I I	III	BSC-DS 308 T	Introduction to Time Series / Open Course DSEC-02	03
31	BSC-DS-I I	III	AECC-01	MIL:Foreign Language	02
32	BSC-DS-I I	III	AECC-02	Critical Thinking / Scientific Temper	02
33	BSC-DS-I I	III	BSC-DS 309 T	Introduction to Analog and Digital Electronics SEC-01	02
34	BSC-DS-I I	IV	BSC-DS 401 T	Object Oriented Concepts Using Java Programming	03
35	BSC-DS-I I	IV	BSC-DS 402 T	NoSQL Databases	03
36	BSC-DS-I I	IV	BSC-DS 403 T	Artificial Intelligence in Data Science	02
37	BSC-DS-I I	IV	BSC-DS 404 P	Lab course on 401	02
38	BSC-DS-I I	IV	BSC-DS 405 P	Lab course on 402	02
39	BSC-DS-I I	IV	BSC-DS 406 P	Lab course on 403,407 and 408	02
40	BSC-DS-I I	IV	BSC-DS 407 T	Statistical Inference / Open Course DSEC-03	03
41	BSC-DS-I I	IV	BSC-DS 408 T	Baysian Inference / Open Course	03

				DSEC-04	
42	BSC-DS-I I	IV	AECC-03	MIL:Foreign Language	02
43	BSC-DS-I I	IV	AECC-04	Critical Thinking / Scientific Temper	02
44	BSC-DS-I I	IV	BSC-DS 409 T	Data Communication and Networking SEC-02	02
45	BSC-DS-III	V	BSC-DS 501 T	Python Programming	03
46	BSC-DS-III	V	BSC-DS 502 T	Data Acquisition and Visualization	03
47	BSC-DS-III	V	BSC-DS 503 T	Machine Learning	03
48	BSC-DS-III	V	BSC-DS 504 P	Lab Course on 501	02
49	BSC-DS-III	V	BSC-DS 505 P	Lab Course on 502	02
50	BSC-DS-III	V	BSC-DS 506 T	Programming in SCALA / Open Course DSEC-05	02
51	BSC-DS-III	V	BSC-DS 507 P	Lab Course on 506	02
52	BSC-DS-III	V	BSC-DS 508 T	Operating System Concepts SEC-03	03
53	BSC-DS-III	V	BSC-DS 509 T	Consumer Behaviour and Marketing Research GE-01	02
54	BSC-DS-III	VI	BSC-DS 601 T	Big Data Security	03
55	BSC-DS-III	VI	BSC-DS 602 T	Social Web Analytics	03
56	BSC-DS-III	VI	BSC-DS 603 T	Real Time Analysis	03
57	BSC-DS-III	VI	BSC-DS 604 P	Lab Course on 601	02
58	BSC-DS-III	VI	BSC-DS 605 P	Lab Course on 602	02
59	BSC-DS-III	VI	BSC-DS 606 T	Azure for Analytics / Open Course DSEC-09	03
60	BSC-DS-III	VI	BSC-DS 607 P	Lab Course on 606	02
61	BSC-DS-III	VI	BSC-DS 608 T	Data Mining and Data Warehousing SEC-05	02
62	BSC-DS-III	VI	BSC-DS 609 T	Professional ethics GE-03	02

16.Detail Syllabus:

Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

Syllabus of F.Y B.Sc. Data Science Semester I

Under

Faculty of Science

Semester – I	**Paper – I**
Course Code: BSC-DS 101 T	**Title of the Course: Basic Programming in C**
Credits: 02	**Total Lectures: 33 Hrs.**

Course Outcomes (COs):

1. To learn the fundamental concepts of computers and problem solving techniques.
2. To think logically and demonstrate the functional behaviour of various real world problems.
3. Analyse a computational problem and develop an algorithm/flowchart to find its solution
4. Develop readable C programs with branching and looping statements, which uses Arithmetic,
   Logical, Relational or Bitwise operators.
5. Divide a given computational problem into a number of modules and develop a readable multi-function C program by using recursion if required, to find the solution to the computational problem

Detailed Syllabus:

Unit I: Programming Principles (07)

1.1 Introduction: Types of Programming languages, History, features and application
1.2 Introduction to Algorithms: Definition & Characteristics of algorithm
1.3 Pseudo code and Flowchart.

- pseudocode statements and flowchart symbols

1.4 Steps in Problem Solving
1.5 Problem Solving Strategies
Top down design.
1.6 Introduction to Programming

- Program and Programming
- Programming Languages
- Program development cycle
- Types of software's
  - compiler, Interpreter, Loader and Linker

1.7 Fundamentals in C
History of 'C' , Features of C, A Simple C Program ,Program execution phases Character set, Identifier, Keywords.

Constants, Integer Constants, Real Constants, Character Constants, String Constants, , Backslash character constants.
Variables Rules for naming Variables Declaration of Variable, Assigning Values to Variables, Initialization.

Unit II: Data Types, Operators and Expressions (06)

2.1 Data types

- Basic data types
- Enumerated types
- Type casting
- Declarations and Expressions
- Expressions, Precedence and Associatively, Expression Evaluation, Type conversions.
- Modifiers, Const Qualifier
- Dealing with each data types, Memory representation of each type

2.2 Types

Arithmetic operators

- Increment and decrement operators
- Relational operators
- Logical operators
- The bitwise operators
- The assignment operators
- The conditional operator
- The size of operator
- The comma operator
- Type casting operator
- Precedence and order of evaluation

Unit III: Input-Output Library Functions (05)

3.1 Unformatted I-O Functions

- Single Character Input-Output
- String Input-Output

3.2 Formatted I-O Functions

- printf() with Width Specifier
- scanf() with Width Specifier

Unit IV: Control statements (05)

4.1 Introduction

4.2 Types of Statements :

- Expression Statements, Compound Statements.
- Selection Statements:
    - If, if…else, switch ,
- Iterative Statements :
    - for loop, while loop, do-while loop,
- Jump Statements :Goto, continue, break, return, Exit()

Unit V: Function (04)

5.1 Introduction

- Definition, need of using functions, Advantages of using functions

5.2 Function Prototype

- Declaration, calling a function, Defining a function

5.3 Return statement

5.4 Types of functions

- Recursion, Nested functions, main() function, Library Function

5.5 Local and global variables

Unit VI: Array (06)

6.1 Introduction

- Definition, Declaration of array, Need, Boundary Checking

6.2 One Dimensional arrays

- Initialization, accessing element of 1D arrays, Reading and displaying elements

6.3 Two dimensional arrays

- Declaration of 2D arrays, Initialization of 2D arrays, Accessing element of 2D arrays , Reading and displaying elements.

6.4 Memory representation of array [Row Major, Column Major]

6.5 Multidimensional array

Suggested Readings:

1. R.G.Dromey, "How to Solve it by Computer", Pearson Education, India, 2008.
2. "C" Programming" Brian W. Kernighan and Denis M. Ritchie.
PHI 2nd Edition
3. Let us C Yashwant P. Kanetkar,
BPB publication
4. 21st Century C Ben Klemens OReilly 1st 2012
5. E. Balaguruswamy, "Programming in ANSI C", ISBN: 9781259004612, Tata Mc-Graw Hill Publishing Co Ltd.-New Delhi

Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

Syllabus of F.Y B.Sc. Data Science Semester I

Under

Faculty of Science

Semester – I	**Paper – II**
Course Code: BSC-DS 102 T	**Title of the Course: Database Management Systems**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

1. To understand the different issues involved in the design and implementation of a database system.
2. To study the physical and logical database designs and understand, database modelling.
3. To understand and learn Structured Query language and data manipulation language.
4. To develop an understanding of essential DBMS concepts.

Detailed Syllabus:

Unit I: File Organization (03)
1.1 Introduction –Basic concept of File, File system, File operations
1.2 Physical / logical files
1.3 Record organization (fixed, variable length)
1.4 Types of file organization (heap, sorted, indexed, hashed)
Unit II: Introduction of DBMS (07)
2.1 Overview- Data, information, database, DBMS, field, record
2.2 File system Vs. DBMS
2.3 Component of database system :Database Architechture
2.4 Describing & storing data (Data models - relational, hierarchical, network),
2.5 Levels of abstraction
2.6 Data independence
2.7 Structure of DBMS
2.8 Users of DBMS
2.9 Advantages of DBMS
Unit III: Conceptual Database Design (E-R model) (09)
3.1 Overview of Database design
3.2 ER data model- E-R diagram (entities, attributes, entity sets, relations, relationship sets)
3.3 Additional constraints
(key constraints, participation constraints, strong entities, weak entities)
3.4 Additional features of database design: aggregation, generalization, specialization
3.5 Codd's rules, Relational Schemas.
3.6 Relational database model: Logical view of data, keys, integrity rules.

3.7 Relational Database design
3.7.1 features of good relational database design
3.7.2 atomic domain and Normalization (1NF, 2NF, 3NF, BCNF).

Unit IV: Relational Algebra and Calculus (06)
4.1 Relational algebra:
4.1.1 Introduction
4.1.2 Selection and projection, set operations, renaming, Joins, Division, syntax, semantics. Operators, grouping and ungrouping, relational comparison.
4.2 Calculus: Tuple relational calculus, Domain relational Calculus, calculus vs algebra, computational capabilities.

Unit- V Constraints and SQL (05)
5.1 What is constraints, types of constrains, Integrity constraints
5.2 SQL: data definition, aggregate function, Null Values, nested sub queries, Joined relations.

Suggested Readings:
1. Henry F. Korth, Abraham Silberschatz, S. Sudarshan Database System Concepts, ISBN: 9780071289597, Tata McGraw-Hill Education
2. Korry Douglas, PostgreSQL, ISBN: 9780672327568
3. John Worsley, Joshua Drake Practical PostgreSQL (B/CD), ISBN: 9788173663925 Shroff/O'reilly
4. Joshua D. Drake, John C Worsley Practical PostgreSQL, O'Reilly

Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

Syllabus of F.Y B.Sc. Data Science Semester I

Under

Faculty of Science

Semester – I	**Paper – III**
Course Code: BSC-DS 103 T	**Title of the Course: Elements of Information Technology**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

1. Identify various components of a computer system.
2. Understand the working of computer system.
3. Understand the working principle of different I/O devices.
4. Aware about the different emerging technologies of data communication.

Detailed Syllabus:

UNIT I Introduction (7)

1.1 Fundamentals of Computer: History of Computer, Generation of Computer
1.2 Classification of Computers
1.2.1 Basic Anatomy of Computer System
1.3 Basic Concept of Assembly language, high level language, compiler, interpreter and assembler.
1.4 Human Computer Interface
1.4.1 Types of software
1.4.2 Operating system as user interface
1.4.3 utility courses.

UNIT II Computer Organization and Architecture (6)

2.1 Architecture of Computer
2.1.1 C.P.U. Basic : registers, system bus
2.1.2 main memory unit, cache memory
2.1.3 Primary, secondary, auxiliary memory, RAM: DRAM, SRAM, ROM: PROM, EPROM, E2PROM, cache memory, hard disks, optical disks
2.1.4 Inside a computer, SMPS, Motherboard, Ports and Interfaces, expansion cards, ribbon cables.

UNIT III Peripheral Devices (7)

3.1 Input Devices (with connections and practical demo)

3.1.1 Types and Working: keyboard, mouse, joystick, scanner, OCR, OMR, bar code reader, web camera

3.2 Output Devices (with connections and practical demo)

3.2.1 Types and Working: monitor, printer, plotter. etc

UNIT IV. **Overview of Networks and Emerging Technologies:** (10)

4.1 Introduction to Networking

4.1.1 Basic of Computer networks

4.2 Types of Networks : LAN, MAN, WAN

4.2.1 Network Topology

4.3 Concept of Internet

4.3.1 Applications of Internet

4.3.2 WWW and Web Browsers : Web Browsing software

4.3.3 Connecting to Internet; ISP

4.3.4 Internet and Intranet

4.3.5 Search Engines; URL; Domain name, Bluetooth

References :

1. Norton Peter, “Introduction to computers”, TMH, 4th Ed., 2006.

2. Simon Haykins, “Communication System”, John Wiley & Sons, 2006.

3. B. Basaraj, “Digital Fundamentals”, Vikas Publications, 1999.

4. V. Rajaraman, “Introduction to Information Technology”, PHI, 2006.

5. V. Rajaraman, “Fundamentals of Computers”, PHI, 5th Ed., 2006.

6. David Anfinson and Ken Quamme, “IT Essentials PC Hardware and Software Component on Guide”, Pearson, 3rd Ed., 2008

Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

Syllabus of F.Y B.Sc. Data Science Semester I

Under

Faculty of Science

Semester – I	**Paper – IV**
Course Code: BSC-DS 104 T	**Title of the Course: Introduction to R programming**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

1. Student will have sufficient computational skill through R- programming software.
2. Student will understand the difference in data visualization using R.
3. Student will understand the difference in output of analysis using R.
4. The overall logical thinking as a base of data science will be improved.
5. Student will have skill of result interpretation.

Detailed Syllabus:

Unit-I		Data Visualization using R	15 L
	1.1	Fundamentals of R Creating a vector using scan function, creating a data frame using edit command, Importing data from MS- Excel file. Using read.table command, saving the R-output in a file using MS-Excel, concept of Rscript file, Graphics using R: (a) High level plotting functions (b) Low level plotting functions (c) Interactive graphic functions.	
	1.2	Diagrams: Simple bar diagram, Subdivided bar diagram, Multiple bar diagram, Pie diagram, stem and leaf diagram. Graphs: Boxplot for one and more than one variables, rod or spike plot, histogram for raw data with prob =TRUE option and for both equal and unequal class intervals, frequency polygon, ogive curves, empirical distribution function, Saving the diagram and graphs using R	
Unit-II		Basic Statistics and Probability using R	15 L
	2.1	Use of R commands to compute measures of Central Tendency, dispersion, skewness and kurtosis Computations of following measures for all types of data. a) Central tendency mean, mode, median, quartiles, deciles, percentiles, geometric mean and harmonic mean. b) Dispersion: variance, standard deviation, coefficient of variation, mean deviation. c) Skewness: Bowley's coefficient and Karl Pearson's coefficient of skewness.	

	2.2	Probability distributions: Simulation from distributions, computations of probabilities, cumulative probabilities, quantiles and drawing random sample using d, p, q, r functions and graphs of pmf/ pdf by varying parameters for following distributions: Binomial, Poisson, Hypergeometric , normal, exponential, gamma, uniform, Fitting of Poisson and normal distribution, testing normality of data by Shapiro-Wilks test.	

Suggested Readings Books:

1. Crawley, M.J. (2006). Statistics – An introduction using R. John Wiley London.
2. Purohit, S.G., Deshmukh, S.R. and Gore, S.D., (2015). Statistics using R. Alpha Science International.
3. Verzani, J., (2018). Using R for introductory statistics. CRC press.
4. Schumacker, R.E., (2014). Learning statistics using R. Sage Publications

Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

Syllabus of F.Y B.Sc. Data Science Semester I

Under

Faculty of Science

Semester – I	**Paper –V**
Course Code: BSC-DS 105 T	**Title of the Course: Introduction to Data**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

1. Understanding basic concept of data
2. Understanding how data is acquired ,the possible sources of error
3. Understand how to identify and overcome errors in data
4. Understand Data represenation

Detailed Syllabus:

UNIT I Introduction (10)

1.1 Build intuition about data and data science
Students are exposed to about 12 to 15 different data like an image, a sound pattern, the night sky in visual and x ray and radio views, patient health data etc. They are asked various questions that they must answer by looking / studying in whatever way at the data."

UNIT II Applications of Data Science Concepts (15)

"Given 12 to 15 different data objects like (complex) images, sounds etc., the student is challenged in various ways. Some example challenges are:

1. Think about how the data has been acquired.
2. The possible sources of error in the data.
3. Ideas to either overcome the errors or work around them,
4. Represent the data on paper in various ways: pick a most convenient one.
5. Summarise the data using paper in the best possible way.
6. Consider the effect of merging two data objects, what can now be asked that was not possible earlier.

Students are expected to complete weekly lab assignments and several mini-projects throughout the course."
"Many are available on the Internet. Some sources are:
1. UC Irvine ML Repository

2. The Azure Open Datasets from Microsoft
3. The DeepAl Datasets

UNIT III Understanding Different kinds of Data **(5)**

a. images of varying parameters like resolution, colours, sizes etc.
b. sound pieces with variations as for images,
c. hand written text samples, text samples with different font designs,
d. tweets and Facebook posts that can be collected through their APIs.
e. measurement of some artifact that is scientifically or engineering wise interesting (e.g. fuel consumption of a car) ... can, and should also be acquired by the instructor and used"

References:

Textbooks " No prescribed textbook. Some resources on the Internet are to be used . Some courses listed below use some programming language (R or Python) for support. The decision of the use of the programming language is left to the instructor .One possible Book can be:

1. Modern Data Science with R, 2nd edition,Benjamin S. Bauner, Daniel T.Kaplan and Nicholas J. Horton, Chapman and Hall/ CRC;ISBN-10:0367191490 ISBN-13:978-0367191498

Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

Syllabus of F.Y B.Sc. Data Science Semester I

Under

Faculty of Science

Semester – I	**Paper –VI**
Course Code: BSC-DS 106 T	**Title of the Course: Probability and Basic Statistics**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

1. Along with regular teaching and learning, student will groom to design posters and power point presentations.
2. Student will improve the process of logical thinking.
3. Students will be aware of the variety of fields in which Statistics is used widely.
4. Student will also gain the knowledge of computational tools.

Detailed Syllabus:

Unit-I		Statistics	15 L
	1.1	Definition of Statistics, applications of Statistics, Types of data: a) Primary data and Secondary data. b) Data collection methods (register, questionnaire, interview method) c) Categorical data, directional data, Binary data, time series data, Panel data and Cross sectional data. d) Image, Voice, Audio, Animated images, Text, Video data	
	1.2	Measures of Central Tendency: Concept and Definition of Central Tendency, Characteristics of good measures of Central Tendency Types of central Tendency, Arithmetic Mean(A.M): Definition of Mean, formulae for ungrouped data Properties of A.M , Trimmed AM, Weighted AM, Median: Definition of Median, Formulae for ungrouped Graphical representation, Partition values Quartiles, Deciles, Percentiles, Quantiles, Mode: Definition of Mode, formulae for ungrouped, Graphical Representation. Empirical relation between mean, median and mode.	
	1.3	Concept and Definition of dispersion, Characteristics of good measures of Dispersion, Types of Dispersion, Absolute and relative measures of dispersions, Range: Definition, range and coefficient of range, mean deviation: definition, coefficient of mean deviation, minimal property of MD, variance and Standard deviation: definition, formula, combined variance. Minimal property of variance (Mean square deviation, coefficient of quartile deviation and coefficient of mean deviation, coefficient of variation(C.V)	
Unit-II		Probability	15 L

	2.1	Basics of Probability, Counting Principles-Additive principle, multiplicative principle, counting Rules-Permutations and combinations. Rules and relationship between permutations and combinations (without proof), concept of deterministic and non-deterministic models (Random experiments)	
	2.2	Definitions of sample space and types of sample space: (i) Sample space, (ii) Types of sample space: finite, countably infinite and uncountable. Real life examples. Definitions of Event and types of event: Event and concept of occurrence of an event, (i) Elementary event, (ii) complement of an event, (iii) certain event, (iv) impossible event, (v) Relative complement event, Mutually exclusive events or Disjoint events (for two and three events), mutually, Exhaustive events (for two and three events), mutually exclusive and exhaustive events, partition of sample space. Algebra of events including De Morgan's rules and its representation in set theory notation.	
	2.3	Classical definition of probability and its limitations. Equiprobable and non-equiprobable sample space, classical definition of probability, Addition theorem on probability, limitations of classical definition. Situations where classical definition of probability is applicable. Axiomatic approach of probability-Axioms of Probability, Situations where axiomatic approach of probability is applicable, Addition theorem on probability and its generalization. Various results on Probability, Boole's inequality. Numerical examples and problems. Definition of conditional probability of an event. Results on conditional probability, Definition of independence of two events P (A ∩ B) = P (A) * P (B), Pairwise independence and mutual independence for three events, Multiplication theorem, P (A ∩ B) = P (B)*P (A|B). Generalization to P (A ∩ B ∩ C).	

Suggested Readings Books:

1. Agarwal, B. L. (2003). Programmed Statistics, Second Edition, New Age International Publishers, New Delhi.
2. Gupta, S. C. and Kapoor, V. K. (1983). Fundamentals of Mathematical Statistics, Eleventh Edition, Sultan Chand and Sons Publishers, New Delhi.
3. Mood, A.M. Graybill, F.A. and Boes, D.C. (2007):Introduction to the Theory of Statistics, 3rd Edn. (Reprint), Tata McGraw-Hill Pub. Co. Ltd.
4. Sarma, K. V. S. (2001). Statistics Made it Simple: Do it yourself on PC. Prentice Hall of India, New Delhi.
5.Snedecor G. W. and Cochran W. G.(1989). Statistical Methods, Eighth Ed. East-West Press.

Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

Syllabus of F.Y B.Sc. Data Science Semester I

Under

Faculty of Science

Semester – I	**Paper –VII**
Course Code: BSC-DS 107 T	**Title of the Course: Matrix Computations**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcome:

1.Use matrix computations in theory and practice to solve linear systems of equations.
2. Use sophisticated scientific computing and visualization environments to solve application problems involving matrix computation algorithms.
3. Analyze numerical algorithms, and understand the relationships between the computational effort and the accuracy of these algorithms.

Detailed Syllabus:

Unit I : Linear systems (06)

All variants of Gaussian elimination and LU factorization, Cholesky factorization.

Unit II: Linear least-squares problem (08)

Normal equations, rotators and reflectors, QR factorization via rotators,

reflectors and Gram Schmidt orthonormalization, QR method for linear least-squares problems, rank

deficient least-squares problems.

Unit III: Singular value decomposition (SVD) (08)

numerical rank determination via SVD, solution of least squares

problems, Moore- Penrose inverse, low rank approximations via SVD, Principal Component Analysis,

applications to data mining and image recognition.

Unit IV: Eigenvalue Decomposition (08)

Power, inverse power and Rayleigh quotient iterations, Schur's decomposition,

unitary similarity transformation of Hermitian matrices to tridiagonal form, QR algorithm, implementation of

explicit QR algorithm for Hermitian matrices.

Suggested Readings:

1. L. N. Trefethen and David Bau, Numerical Linear Algebra, SIAM, Philadelphia, 1997.
2. D. S. Watkins, Fundamentals of Matrix Computation, 2nd Edition, Wiley, 2002.
3. L. Elden Matrix Methods in Data Mining and Pattern Recognition, SIAM, Philadelphia, 2007.

Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

Syllabus of F.Y B.Sc. Data Science Semester I

Under

Faculty of Science

Semester – I	**Paper –VIII**
Course Code: BSC-DS 108 T	**Title of the Course: Discrete Mathematics**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes:

1. Ability to apply mathematical logic to solve problems.
2. Understand sets, relations, functions and discrete structures.
3. Able to model and solve real world problems using graphs and trees.

Detailed Syllabus:

Unit I: Sets (07)

Set Theory - sets and classes, relations and functions, recursive definitions, posets, Zorn - s lemma, cardinal and ordinal numbers;

Unit II: Logic (07)

Logic - propositional and predicate calculus, well-formed formulas, tautologies, equivalence, normal forms, theory of inference.

Unit III: Combinatorics (08)

Combinatorics - permutation and combinations, partitions, pigeonhole principle, inclusion-exclusion principle, generating functions, recurrence relations.

Unit IV: Graph Theory (08)

Graph Theory - graphs and digraphs, Eulerian cycle and Hamiltonian cycle, adjacency and incidence matrices, vertex colouring, planarity, trees.

Suggested Readings:

1. J.P. Tremblay and R. Manohar, Discrete Mathematical Structures with

Applications to Computer Science, Tata McGraw Hill, New Delhi, 2001.

C. L. Liu, Elements of Discrete Mathematics, 2nd Edn., Tata McGraw-Hill, 2000.

2. K. H. Rosen, Discrete Mathematics & its Applications, 6th Edn., Tata McGraw-Hill, 2007.

V. K. Balakrishnan, Introductory Discrete Mathematics, Dover, 1996.

3. J. L. Hein, Discrete Structures, Logic, and Computability, 3rd Edn., Jones and Bartlett, 2010.

4. N. Deo, Graph Theory, Prentice Hall of India, 1974.

Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

Syllabus of F.Y B.Sc. Data Science Semester I

Under

Faculty of Science

Semester – I	**Paper –IX**
Course Code: BSC-DS 109 P	**Title of the Course: Lab Course on 101**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

Practical Implementation of the topic covered in theory includes :
1. Understand the logic for a given problem, Write the algorithm of a given problem and draw a flow chart of a given problem.
2. Recognize and understand the syntax and construction of C programming code.
3. Know the steps involved in compiling, linking and debugging C code.
4. Learn the methods of iteration or looping and branching.
5. Make use of different data-structures like arrays
6. Understand function declaration and definition and use of user defined functions.

Detailed Syllabus:

Assignments :Sample assignments are given . Students are expected to complete atleast 5 assignments per topic as per theory syllabus in addition to list given below

Assignment 1.
a)Write a C program to find sum and average of three numbers.
b)Write a C program to find the sum of individual digits of a given positive integer.
c) a) Write a C program to find the roots of a quadratic equation..
Assignment 2.
a)Write a C program to generate prime numbers between 1 to n.
b)Write a C program to Check whether given number is Armstrong Number or Not.
c) Write a C program to evaluate algebraic expression (ax+b)/(ax-b).
Assignment 3.
a)Write a C program to check whether given number is perfect number or Not.
b)Write a C program to check whether given number is armstrong number or not.
Assignment 4.
a) Write a C program to generate the first n terms of the Fibonacci sequence
b)Write a C program perform arithmetic operations using switch statement.
Assignment 5
a)Write a C program to find factorial of a given integer using non-recursive function.
b)Write a C program to find factorial of a given integer using recursive function.
a)Write C program to find GCD of two integers by using recursive function.
Assignment 6

a)Write a C program to find both the largest and smallest number in a list of integers.
b) Write a C Program to Sort the Array in an Ascending Order.
c) Write a C Program to find whether given matrix is symmetric or not.

Assignment 7

a) Write a C program to perform addition of two matrices.
b)Write a C program that uses functions to perform Multiplication of Two Matrices.

Assignment 8

a)Write a C program to use function to insert a number in to given main array at a given position.
b) Write a C program that uses functions to delete n numbers from a given position in a given array.

Assignment 9

a)Write a C program using user defined functions to determine whether the given number is palindrome or not.
b)Write a C program using user defined functions to determine whether the given number is armstrong or not.

Assignment 10

a) Write a C program to perform addition of diagonal elements of an 2 D array.
b) Write a C program to pass a 1 D array to a function. using user defined function calculate the sum and average of the array.

Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

Syllabus of F.Y B.Sc. Data Science Semester I

Under

Faculty of Science

Semester – I	**Paper –X**
Course Code: BSC-DS 110 P	**Title of the Course: Lab Course on 102**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

Practical Implementation of the topic covered in theory includes :
1. Prepare E-R Diagram for the given problem statement
2. Formulate appropriate SQL DDL Queries
3. Formulate appropriate SQL DML Queries

Detailed Syllabus:

Assignments :Sample assignments are given . Students are expected to complete atleast 5 assignments per topic as per theory syllabus in addition to list given below

Assignment

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 Intergrity constrains

7) Perform the following operation for demonstrating the insertion , updation and deletion using the referential integrity constraints

8) Nested Queries Assignment

9) Write the query for creating the users and their role.

Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

Syllabus of F.Y B.Sc. Data Science Semester I

Under

Faculty of Science

Semester – I	**Paper –XI**
Course Code: BSC-DS 111 P	**Title of the Course: Lab Course on 104,106,107,108**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

1. On completion of all the practical's student will have sufficient computational skill through software and programming.
2. Student will understand the difference in data visualization by R.
3. Student will understand the difference in output of analysis by R.
4. The overall logical thinking as a base of data science will be improved.
5. Student will have skill of result interpretation.

Detailed Syllabus:

1) Introduction of different R- Commands and help I R.
2) Practical on import of data in R from various platforms.
3) Identification of different data types in given big data.
4) Present the given data using Graphs, Diagrams and Charts .
5) Measures of central tendencies based on raw data .
6) Practical based on partition values.
7) Measures of central tendencies based Continuous frequency distribution.
8) Measures of dispersion based on different data types .
9) Measures of absolute and relative measures of dispersion.
10) Verification of different rules of counting principles .
11) Drawing of different events using vein- diagram .
12) Computation of different probabilities .
13) Conditional probabilities .
14) Practical based on probability mass function and cumulative distribution function .
15) Practical based on mathematical expectation.

SEM II

Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

Syllabus of F.Y B.Sc. Data Science Semester II

Under

Faculty of Science

Semester – II	**Paper –I**
Course Code: BSC-DS 201 T	**Title of the Course: Advanced C programming**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

1. Use of Preprocessor Directive.
2. Implementation of Pointers.
3. Learn to dynamically allocate memory.
4. Efficient use of file handling

Detailed Syllabus:

UNIT I: Array and Function **(04)**

1.1 1D array and function

1.1.1 Passing 1D array to function :

- Passing individual array elements to a function
- Passing individual array elements address to a function
- Passing whole 1 D array to a function

1.2 2 D array and function

1.2.1Passing 2D array to function :

- Passing individual array elements to a function
- Passing individual array elements address to a function
- Passing whole 2D array to a function

UNIT II :Introduction to C Preprocessor **(04)**

2.1 Definition of Preprocessor
2.2 Types of Preprocessors
2.3 Macro substitution directives
2.4 Macros versus function
2.5 File inclusion directives
2.6 Conditional compilation processors
2.7 Predefined macros
2.8 Preprocessor Operator

Unit III: Pointers **(08)**

3.1 Introduction

- Definition and declaration
- Initialization

3.2 Indirection operator and Address of operator
3.3 Types of Pointers
3.4 Pointer arithmetic
3.5 Dynamic memory allocation
3.6 Arrays and pointers
3.7 Pointer to array
3.8 Array of pointers

3.9 Function and pointers

- Call by value and call by refernce
- Passing pointer to function
- Returning pointer from function,
- Function pointer

3.10 Pointers & const:

- Constant pointer
- Pointer to a constant

Unit IV: Strings **(7)**

4.1 Introduction

- Definition
- Declaration
- Initialization

4.2 Importance of terminating NULL character
4.3 Strings & pointers
4.4 String and Function

- User Defined
- Standard library function strlen(), strcpy(), strcat(),strcmp() etc

4.5 Command line arguments – argc and argv

Unit V: Structures and Union **(4)**

5.1 Introduction

- Definition
- Declaration

5.2 Variables initialization
5.3 Accessing fields and structure operations
5.4 Nested structures
5.5 Array of structure variables
5.6 Structure and function
5.7 pointer and structure

- Declaration
- Initialization
- Accessing members using pointers

5.8 Introduction to union

- Definition Declaration
- Initialization

5.9 Differentiate between Union and structure

5.10 Accessing fields and structure operations
5.11 Nested structures and unions

Unit VI: File Handling **(3)**

6.1 Introduction

- Definition
- Types of files

6.2 Concept of streams
6.3 Operations on files:

- Text files
- Binary files
- Random access file

6.4 Library functions for file handling – fopen, fclose, fgetc, fseek, fgets, fputc etc,feof,rewind etc

Suggested Readings:

1. R.G.Dromey, "How to Solve it by Computer", Pearson Education, India, 2008.
2. "C" Programming" Brian W. Kernighan and Denis M. Ritchie.
PHI 2nd Edition
3. Let us C Yashwant P. Kanetkar,
BPB publication
4. 21st Century C Ben Klemens OReilly 1st 2012
5. E. Balaguruswamy, "Programming in ANSI C", ISBN: 9781259004612, Tata Mc-Graw Hill Publishing Co Ltd.-New Delhi

Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

Syllabus of F.Y B.Sc. Data Science Semester II

Under

Faculty of Science

Semester – II	**Paper –II**
Course Code: BSC-DS 202 T	**Title of the Course: Advanced Database Management System**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

1. Formulate SQL queries using advanced SQL features.
2. Perform Database operations using PL/PostgreSQL
3. Compare and contrast different concurrency control and recovery techniques.
4. Apply mechanisms for database security.
5. Analyze various database system architectures.

Detailed Syllabus:

Unit I: Relational Database Design (8)

1.1. PL/Postgresql: Language Structure

1.2. Controlling the Program Flow, Conditional Statements, Loops

1.3. Views

1.4. Functions and procedures

1.5. Handling Errors and Exceptions

1.6. Cursors

1.7. Triggers

Unit II: Transaction Concepts (6)

2.1 Transaction, Properties of Transaction, States of Transactions

2.2 Concurrent Execution of Transactions and Conflicting Operations

2.3 Schedules, Types of Schedules, Concept of Serializability, Precedence Graph for Serializability

Unit III: Concurrency Control (8)

3.1 Ensuring Serializability by Locks, Different Lock Modes

3.2 2PL And Its Variations

3.3 Multiple Granularity Locking Protocol

3.4 Basic Timestamp Method for Concurrency, Thomas Write Rule

3.5 Locks with Multiple Granularity, Dynamic Database Concurrency (Phantom Problem)

3.6 Timestamps versus Locking

3.7 Optimistic Concurrency Control Algorithm, Multi Version Concurrency Control

3.8 Deadlock Handling Methods –

3.8.1 Detection And Recovery (Wait For Graph).

3.8.2 Prevention Algorithms (Wound-Wait, Wait-Die)

3.8.3 Deadlock Recovery Techniques (Selection of Victim, Starvation, Rollback)

Unit IV: Crash Recovery (4)

4.1 Transaction Failure Classification

4.2 Recovery Concepts

4.3 Checkpoints

4.4 Recovery with Concurrent Transactions (Rollback, Checkpoints, Commit)

4.5 Log Base Recovery Techniques (Deferred and Immediate Update)

4.6 Buffer Management

4.7 Database Backup and Recovery from Catastrophic Failures

4.8 Shadow Paging

Unit V: Database Security and Database System Architecture (4)

5.1 Introduction to Database Security Concepts

5.2 Methods for Database Security

5.3 Centralized and Client – Server Architectures

5.4 Server System Architectures

5.5 Introduction to Parallel Systems

5.6 Introduction to Distributed Systems

5.7 Introduction to Object Based Databases

5.8 Introduction to Web based databases

Suggested Readings:

1. Database System Concepts – Abraham Silberschatz, Henry F. Korth, S. Sudarshan, 6th editionMcGraw-Hill

2. Fundamentals of Database Systems- Ramez Elmasri, Shamkant B. Navathe, 6th edition– Pearson.

3. Database Management Systems -Raghu Ramakrishnan, Johannes Gehrke, 3rd edition, Tata McGraw Hill

4. Introduction to Database Management System- Bipin Desai, 3rd edition, Galgotia Publication.

5. An Introduction to Database Systems - C.J. Date, 7 th edition, Addison-Wesley

6. Practical PostgreSQL- Joshua D. Drake, John C Worsley, O'Reilly Publications

Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

Syllabus of F.Y B.Sc. Data Science Semester II

Under

Faculty of Science

Semester – II	**Paper –III**
Course Code: BSC-DS 203 T	**Title of the Course: Fundamentals of Data Science**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

1. Apply data visualisation in big-data analytics
2. Utilise EDA techniques
3. Apply data pre-processing techniques
4. Understand the basics in R programming in terms of constructs, control statements, string functions
5. Understand the use of R for Big Data analytics
6. Able to appreciate and apply the R programming from a statistical perspective

Detailed Syllabus:

UNIT I Introduction (8)
1.1 Big Data and Data Science - Big Data Analytics,
1.2 Business intelligence vs Big data
1.3 big data frameworks
1.4 Current landscape of analytics
1.5 data visualisation techniques, visualisation software
1.6 data repositories- analyst perspective
1.7 Business drivers for analytics
1.8 Typical analytical architecture
1.9 Drivers of Big data analytics
1.10 Role of data scientist in Big data ecosystem
1.11 Applications of Big data analytics.

UNIT II Exploratory Data Analysi (2)

2.1 Exploratory Data Analysis (EDA)
2.2 Statistical measures
2.3 Basic tools (plots, graphs and summary statistics) of EDA

UNIT III Basic Statistical Inference (4)

3.1 Developing Initial Hypotheses
3.2 Identifying Potential Data Sources
3.3 EDA case study

3.4 Testing hypotheses on means
3.5 Proportions and variances

UNIT IV Data Analytics Lifecycle (7)
4.1 Need of Data analytic lifecycle
4.2 Key roles for successful analytic projects
4.3 Phases of Data analytic lifecycle: Discovery, Data Preparation, Model Planning, Model Building, Communicating Results, Operationalization

UNIT V Data Analytics using SPARK (9)
5.1 Introduction to Spark
5.2 Apache Spark Architecture
5.3 Components of Spark, Spark RDDs
5.4 RDD Operations: Transformation & Actions
5.5 Spark SQL Library – DataFrames, Leverage Hive for Spark
5.6 Machine learning with Spark and Realtime Screening:

- Machine Learning using Spark ML
- Illustrate ML Algorithm using PySpark,
- Into to Kafka for Spark Streaming
- Apache Spark Streaming Features
- Spark Streaming Workflow
- Streaming Context and Dstreams

Recommended Books:
1.Doing Data Science, Straight Talk From The Frontline, Cathy O'Neil and Rachel Schutt, O'Reilly (2014).
2.Big Data and Business Analytics, Jay Liebowitz, CRC press (2013)
3.Learning R Richard Cotton O'Reilly

Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

Syllabus of F.Y B.Sc. Data Science Semester II

Under

Faculty of Science

Semester – II	**Paper –IV**
Course Code: BSC-DS 204 T	**Title of the Course: Discrete Probability Distribution**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

1. Understand the need of standard probability distribution as a model.
2. Understand real life situations for the use of these models.
3. Learn interrelation among the different probability distributions.
4. Understand the concept of discrete random variable taking countably finite and countably infinite value
5. Learn Poisson distribution as the limiting form of the binomial distribution
6. Understand Geometric and Negative Binomial as waiting time distribution.

Detailed Syllabus:

Unit-I		**Univariate Probability Distributions-Based on Finite Support**	15 L
	1.1	Univariate probability mass function (p.m.f.)-Concept and definition of a random variable. Types of random variable. Concept and definition of a discrete random variable, Probability mass function (p.m.f) and cumulative distribution function (c.d.f), F(.) of discrete random variable, properties of c.d.f., graphical representation of p.m.f. and c.d.f., Mode and median of discrete probability distribution, Numerical examples and problems.	
	1.2	**Discrete uniform (DU) distribution:** Real life situations, definition of DU distribution, plot of p.m.f., cumulative distribution function (CDF), plot of CDF. **Bernoulli Distribution**: Concept of Bernoulli trials, genesis of p.m.f. of Bernoulli distribution, definition of Bernoulli distribution with parameter p, Notation, real life situations, plot of probability mass function, cumulative distribution function (CDF), plot of CDF, mean and variance	
	1.3	**Binomial Distribution**: Definition of binomial distribution with parameters n and p, Notations, distribution of number of successes in n independent Bernoulli trial as a Binomial distribution with parameters n and p, conditions for the applications of binomial distribution, an illustration of use of	

		binomial distribution in SRSWR, real life situations, distribution of $n - X$ if X has $B\ (n, p)$, mean and variance **Hypergeometric Distribution**: Failure of assumptions of binomial distribution in SRSWR, genesis of p.m,f. of hypergeometric distribution with parameters $N, M\ and\ n$, difference between hypergeometric and binomial distribution, conditions for the applications of hypergeometric distribution, real life situations, mean and variance,	
Unit-II		**Univariate Probability Distributions-Based on Countable Support**	15 L
	2.1	**Poisson Distribution**: Poisson distribution as a model for the situations where chances of occurrence of an event in a short time interval is with high probability, real life situations, definition of Poisson distribution, Notation, mean and variance, MGF, additive property, generalization of additive property, conditional distribution of X_1 given $X_1 + X_2 = n$, Poison distribution as a limiting form of binomial distribution, fitting of Poisson distribution, mode of Poisson distribution	
	2.2	**Geometric distribution**: Genesis of p.m.f. of geometric distribution (for both forms), definition of Geometric distribution, Notation, Geometric distribution as waiting time distribution, mean, variance, mode, relation between mean and variance, , Lack of memory property and its interpretation, real life applications.	

Suggested Readings Books:

1. Gupta, S.C. and Kapoor, V. K. (1983). Fundamentals of Mathematical Statistics, Eighth Edition, Sultan Chand and Sons Publishers, NewDelhi.
2. Agarwal B. L. (2003). Programmed Statistics, second edition, New Age International Publishers, New Delhi.
3. Hoel P. G. (1971). Introduction to Mathematical Statistics, John Wiley and Sons, New York.
4. Hogg R.V. and Craig R.G. (1989). Introduction to Mathematical Statistics, MacMillan Publishing Co., New York.
5. Mood, A. M. and Graybill, F. A. and Boes D.C. (1974). Introduction to the Theory of Statistics, Ed. 3, McGraw Hill Book Company.

Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

Syllabus of F.Y B.Sc. Data Science Semester II

Under

Faculty of Science

Semester – I	**Paper –V**
Course Code: BSC-DS 205 T	**Title of the Course: Continuous Probability Distribution**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

1. On completion of this course students will get in depth knowledge of general concepts of univariate probability distributions.
2. Students will get sound knowledge of a probability distributions with lot of applications in industrial field.
3. Students will get the complete idea about probability distributions that helps the comparison and to draw the inference.
4. Practice of learning of different concepts in this course using R coding will improve their logical thinking.

Detailed Syllabus:

Unit-I		General Concepts in Continuous Univariate Distributions (15)
	1.1	Continuous sample space: Definition, illustrations. Continuous random variable: Definition, probability density function (p.d.f.), cumulative distribution function (c.d.f.), properties of c.d.f., probabilities of events related to random variable, Plot p.d.f. and c.d.f. of some non-standard probability distributions.
	1.2	Expectation of continuous r.v., expectation of function of r.v., E [g(X)], mean, variance, geometric mean, harmonic mean, median, mode, Partition values: Quartiles (Q_1, Q_2, Q_3), Deciles, Percentiles. Raw and central moments, skewness, kurtosis, mean deviation about constant, about mean, about median, about mode.
	1.3	Uniform or Rectangular Distribution- Probability density function (p.d.f.) $f(x) = \begin{cases} \frac{1}{(b-a)}, & a \le x \le b \\ 0, & Otherwise \end{cases}$

		Notation: X→ U [a, b] p. d. f., sketch of p. d. f. & c. d. f., mean, median, mode, variance, standard deviation, C.V., symmetry, quartiles, deciles and percentiles. Model sampling from Uniform distribution. Examples and problems
Unit-II		Standard Continuous Probability Distributions (15)
	2.1	Exponential Distribution: Probability density function (p. d. f.) $f(x) = \begin{cases} \alpha e^{-\alpha x}, & x \geq 0, \alpha > 0 \\ 0, & Otherwise \end{cases}$ Notation: X~ Exp(α) Nature of density curve, interpretation of α as a interarrival rate of customer's joining the queue and $\frac{1}{\alpha}$ as mean, mean, median, mode, variance, standard deviation, C.V., lack of memory or forgetfulness property, quartiles, deciles, percentiles, mean deviation about mean, Additive property, Distribution of min (X, Y) and max (X, Y) with X, Y i.i.d. exponential random variables.
	2.2	Normal or Gaussian Distribution: Probability density function (p. d. f.) $f(x) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}, & -\infty < X, \mu < \infty, \sigma > 0 \\ 0, \ Otherwise \end{cases}$ Notation: X~ N (μ, σ^2) p. d. f. curve, identification of scale and location parameters, nature of probability curve, mean, median, mode, variance, standard deviation, C.V, skewness, kurtosis, quartiles, deciles, percentiles, points of inflexion of probability curve, mean deviation, additive property, probability distribution of : i) ($\frac{x-\mu}{\sigma}$) Standard Normal Variable (SNV), ii) aX+b, iii) aX+bY+c where X and Y are independent normal variates. Probability distribution of $\bar{X}$, the mean of n i.i.d. N(μ,σ^2) r. v s., computations of normal probabilities using R. Central Limit Theorem (CLT) for r.v.s. with finite variance (statement only), its illustration for Poisson and Binomial distributions. Box-Muller transformation and normal probability plot. Plot p.d.f. and c.d.f. of Exponential distribution using R code. R code for quartiles, deciles and percentiles. Draw random sample of size n from Normal distribution using R code.

Suggested Readings Books:

1. Fundamentals of Mathematical Statistics, by Gupta and V.K. Kapoor.
2. Continuous Univariate Distributions – 1, by Norman L. Johnson and Samuel Kotz

Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

Syllabus of F.Y B.Sc. Data Science Semester II

Under

Faculty of Science

Semester – II	**Paper –VI**
Course Code: BSC-DS 206T	**Title of the Course: Linear Algebra**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcome:

1. Solve systems of linear equations and interpret their results.
2. Demonstrate an understanding of vector spaces and subspaces.
3. Demonstrate an understanding of eigenvalues and eigenvectors.

Detailed Syllabus:

Unit I: Euclidean Vector Spaces [08 Hours]

1.1 Vectors in 2-Space, 3-Space, and n-Space

1.2 Norm, Dot Product, and Distance in R^n

1.3 Orthogonality

1.4 The Geometry of Linear Systems

1.5 Cross Product

Unit II: General Vector Spaces [08 Hours]

2.1 Real Vector Spaces

2.2 Subspaces

2.3 Linear Independence

2.4 Coordinates and Basis

2.5 Dimension

2.6 Change of Basis

Unit III: Eigenvalues and Eigenvectors [04 Hours]

3.1 Eigenvalues and Eigenvectors

3.2 Diagonalization

Unit-IV: General Linear Transformations [10 Hours]

4.1 General Linear Transformations

4.2 Compositions and Inverse Transformations

4.3 Isomorphism

4.4 Matrices for General Linear Transformations

Suggested Readings:

1. Howard Anton, Chris Rorres, Elementary Linear Algebra, Application Version, Wiley, 11th edition, 2013.

Unit I: Chapter 3

Unit II: Chapter 4

Unit III: Chapter 5

Unit IV: Chapter 8

2. K. Hoffman and R. Kunze, Linear Algebra, Prentice Hall of India, 2nd edition (2014), New Delhi

3. S. Lang, Introduction to Linear Algebra, 2nd edition,1986, Springer-Verlag, New York, Inc.

4. S. H. Friedberg, A. J. Insel and L. E. Spence, Linear Algebra, 4th ed., Prentice Hall, 2003.

5. G. Strang, Linear Algebra and its Applications, Cengage Learning, 15th Re-print edition, 2014.

Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

Syllabus of F.Y B.Sc. Data Science Semester II

Under

Faculty of Science

Semester – II	**Paper –VII**
Course Code: BSC-DS 207 T	**Title of the Course: Operation Research**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes:

1. Identify and develop operational research models from the verbal description of the real system.
2. Understand the mathematical tools that are needed to solve optimisation problems.
3. Use mathematical software to solve the proposed models.

Detailed Syllabus:

1. **Modeling with Linear Programming:** [5 lectures]

Two variable LP Model, Graphical LP solution, Selected LP Applications, Graphical Sensitivity analysis.

2. **The Simplex Method:** [8 lectures]

LP Model in equation form, Transition from graphical to algebraic solutions, the simplex method, Artificial starting solutions.

3. **Duality:** [6 lectures]

Definition of the dual problem, Primal dual relationship.

4. **Transportation Model:** **[**8Lectures]

Definition of the Transportation model. The Transportation algorithm.

5. **The Assignment Model**: [03 Lectures]

The Hungarian method, Simplex explanation of the Hungarian method.

Suggested Readings:

1. Hamdy A. Taha, Operation Research (Eighth Edition, 2009), Prentice Hall of India Pvt. Ltd, New Delhi.
2. Frederick S. Hillier, Gerald J. Lieberman, Introduction to Operation Research (Eighth Edition) Tata McGraw Hill.
3. J K Sharma, Operations Research (Theory and Applications, second edition, 2006), Macmilan India Ltd.
4. Hira and Gupta, Operation Research

Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

Syllabus of F.Y B.Sc. Data Science Semester II

Under

Faculty of Science

Semester – II	**Paper –VIII**
Course Code: BSC-DS 208 T	**Title of the Course: Numerical Methods**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes(COs):

1. Demonstrate understanding of common numerical methods and how they are used to obtain approximate solutions to otherwise intractable mathematical problems.
2. Apply numerical methods to obtain approximate solutions to mathematical problems.
3. Derive numerical methods for various mathematical operations and tasks, such as interpolation, differentiation, integration, the solution of linear and nonlinear equations, and the solution of differential equations.

Detailed Syllabus:

Unit I : Algebraic and Transcendental Equation [04 Hours]

1.1 Introduction to Errors
1.2 False Position Method
1.3 Newton-Raphson Method

Unit II: Calculus of Finite Differences and Interpolation [10 Hours]

2.1 Differences
2.2 Forward Differences
2.3 Backward Differences
2.4 Central Differences
2.5 Other Differences (δ, μ operators)
2.6 Properties of Operators
2.7 Relation between Operators
2.8 Newton's Gregory Formula for Forward Interpolation
2.9 Newton's Gregory Formula for Backward Interpolation
2.10 Lagrange's Interpolation Formula
2.11 Divided Difference
2.12 Newton's Divided Difference Formula

Unit III: Numerical Integration [08 Hours]

3.1 General Quadrature Formula
3.2 Trapezoidal Rule
3.3 Simpson's one-Third Rule
3.4 Simpson's Three-Eight Rule

Unit IV: Numerical Solution of Ordinary Differential Equation [08 Hours]

4.1 Euler's Method
4.2 Euler's Modified Method
4.3 Runge-Kutta Methods

Suggested Readings:

1. A textbook of Computer Based Numerical and Statistical Techniques, by A. K. Jaiswal and Anju Khandelwal. New Age International Publishers.

   Unit 1: Chapter 2: Sec. 2.1, 2.5, 2.7

   Unit 2: Chapter 3: Sec. 3.1, 3.2, 3.4, 3.5, Chapter 4: Sec. 4.1, 4.2, 4.3,

   Chapter 5: Sec. 5.1, 5.2, 5.4, 5.5

   Unit 3: Chapter 6: Sec. 6.1, 6.3, 6.4, 6.5, 6.6, 6.7

   Unit 4: Chapter 7: Sec. 7.1, 7.4, 7.5, 7.6

2. S.S. Sastry; Introductory Methods of Numerical Analysis, 3rd edition, Prentice Hall of India, 1999.
3. H.C. Saxena; Finite differences and Numerical Analysis, S. Chand and Company.
4. K.E. Atkinson; An Introduction to Numerical Analysis, Wiley Publications.

Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

Syllabus of F.Y B.Sc. Data Science Semester II

Under

Faculty of Science

Semester – II	**Paper –IX**
Course Code: BSC-DS 209 P	**Title of the Course: Lab Course on 201**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

1. Implement Programs with pointers and arrays, perform pointer arithmetic, and
2. Efficient use of pre-processor.
3. Write programs that perform operations using derived data types.
4. Handling of data through files. (text and binary)

Detailed Syllabus:

Assignments:

1. Write the Program to implement macros for example: -define constant and array size
2) Write the Program to
a. find maximum of two integers
b. check whether a number is positive, negative or Zero
c. check given number is even or odd
3) Write the Program to illustrate the use of #pragma
4) Write a program to Interchange values of two numbers using pointers
5) Write a program to display the elements of an array containing n integers in reverse order using pointer
6) Write a program to reverse the elements of an array containing n integers using pointer
7) Write a program to multiply two numbers using function pointer
8) Write a Program to accept an array and print the same using double pointer
9) Write a program to calculate average of array of n numbers. Pass the array to a function and use pointers
10) Write a program to find the number of vowels, consonants, digits and white space in a string.
11) Write a program to accept a word and a string. Remove / delete the given word from a string. Example: - if word is= "Hello" and the String is "Hello All Well Come" The output is:- "All Well Come"
12) Write a program that accepts names of n cities and write functions for the following: a) Search for a city b) Display the longest names
13) Write a program which accepts a sentence from the user and replaces all lower case letters by uppercase letters. 4) Write a program to find the First Capital Letter in a String.

write a function iscap() to find the first capital letter. 5) Write a program to remove all other characters in a string except alphabets.
14) Write a program to compare two strings. If they are not equal display their length and if equal concatenate them
15) Write a program to pass two strings to user defined function and copy one string to another using pointer
16) Write a program to reverse string, without using another string variable.

Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

Syllabus of F.Y B.Sc. Data Science Semester II

Under

Faculty of Science

Semester – II	**Paper -X**
Course Code: BSC-DS 210 P	**Title of the Course: Lab Course on 202**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

a. Formulate SQL queries using advanced features
b. Write stored procedures, cursors and triggers using PL/PostgreSQL.
c. Design a database using database normalization technique

Detailed Syllabus:

***Topics that are not included in theory must be covered with the relevant assignments**

Assignment 1: Designing a Database using normalization theory for

the given application/database design

Assignment 2: Simple and Nested Queries

Assignment 3: Views Creation

Assignment 4: Stored Functions

1) A Simple Stored Function

2) A Stored Function that returns

3) A Stored Function recursive

Assignment 5: Cursors

1)Simple Cursor

2) Parameterize Cursor

Assignment 6: Error and Exception handling

1) Simple Exception- Raise Debug Level Messages

2) Simple Exception- Raise Notice Level Messages

3) Simple Exception- Raise Exception Level Messages

Assignment 7: Triggers

1) Before Triggers (insert, update, delete)

2) After Triggers (insert, update, delete)

Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

Syllabus of F.Y B.Sc. Data Science Semester II

Under

Faculty of Science

Semester – II	**Paper –XI**
Course Code: BSC-DS 211 P	**Title of the Course: Lab Course on 204,205,206,207,208**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

a. On completion of all the practical's student will have sufficient computational skill through software and programming.
b. Student will understand the difference in data visualization by MS-Excel and R.
c. Student will understand the difference in output of analysis by MS-Excel and R.
d. The overall logical thinking as a base of data science will be improved.
e. Student will have skill of result interpretation.

Detailed Syllabus: List of Practical's

1) Practical based on Plotting of pmf and cdf of on Bernoulli binomial distribution
2) Practical based on Plotting of pmf and cdf of on Hypergeometric distribution
3) Practical based on Plotting of pmf and cdf of on Poisson and Geometric distribution
4) Practical based on Plotting of pmf and cdf of on uniform, exponential & normal distribution
5) Practical based on computation of probabilities of Bernoulli binomial distribution
6) Practical based on computation of probabilities of Hypergeometric distribution
7) Practical based on computation of probabilities of Poisson and Geometric distribution
8) Practical based on computation of probabilities of uniform, exponential & normal distribution
9) Fitting of Binomial distribution and computation of expected frequencies
10) Fitting of Poisson distribution and computation of expected frequencies
11) Fitting of Geometric distribution and computation of expected frequencies
12) Applications of Binomial & hypergeometric distributions
13) Applications of Poisson & geometric distributions
14) Model sampling from Poisson and Binomial distributions
15) Model sampling from Normal, Exponential and Uniform distribution

17**. Question paper Pattern:**

Question Paper Format (For 35 and 50 marks)

Q-1 M.C.Q.	05 Marks
Q-2 Answer (any 05 out of 07)	15 Marks (03 Each)
Q-3 Answer (any 03 out of 05)	15 Marks (05 Each)

For 50 MARKS include Q4

Q-4 Answer (any 03 out of 05)	15 Marks (05 Each)