



KANNUR UNIVERSITY

POST GRADUATE DIPLOMA IN DATA SCIENCE & ANALYTICS

(Credit Based Semester System)

Regulations, Curricula, Syllabus and Scheme of Evaluation

(With Effect from 2020 admission)

1. Introduction

Data Science and Artificial Intelligence are two prominent domains that are going to rule the entire world in the future decades. Currently these domains are being used in many real life applications like Business, Commerce & Banking, GIS, Health, Climate change, Automobiles, Crime analysis etc. The key ideas of Artificial Intelligence and Data Science are mainly used for identifying and making prediction on unseen information hidden in the enormous data available in the real world application. Using these techniques, one can identify and analyse those relevant data and make prediction of the future of data. Prediction can be done by analysing in the form of patterns hidden in the original data and further it can be utilized for the benefits of the society by applying various mathematical and Statistical tools that are available. Since one has to deal with huge amount of data and information, ordinary database applications cannot deal them easily. The significance of Big Data analysis and Machine Learning comes here. The Deep Learning mechanism as part of Machine Learning provide the capability of identifying the unseen information that normally one cannot identify/retrieve through the conventional SQL query processing. Data mining and data ware housing techniques are the important mechanism for storing and managing mass amount of data that comes in different styles and format which replicates the real world scenario and provide rich set of tools and techniques for extracting the most relevant information for making prediction of the data with the help of Artificial Intelligence that conventional database management system cannot do.

Future job market needs people who posses' technical knowledge and programming skills in dealing these techniques and many conventional programming techniques are being getting obsolete with the advent of these new methods. The conventional courses are presently dealing with foundation courses and specialized mainly in the standard programming concepts. A few institutions may be offering one or two electives during their course of study without dealing with practical aspects of the domains that will not be sufficient for them to fetch a job in the current industry.

We, Kannur University planning to offer a specialized Data Science course by the joint venture of three departments like Information Technology, Mathematical Science and Statistical Science in a single umbrella to build up practically talented man power capable of handling solution to such complex real world problems. In present scenario, there are two types of professionals to

be evolved for handling the Data Science related problems such as Data Analysts and Data Scientist. The practical aspects of analysing and interpreting the data is done by the Data Analyst and Data Scientist who carry out build mathematical and statistical models for dealing complex data by the Data Analyst. We have confident that by the joint venture of these three departments will achieve this target without any ambiguity. Another important highlight of this course is that we give more stress on the development of practical skills to the students in the related domain and do not give much more importance of the conventional examination systems followed in the University. The Course is offered in the pattern of 60-40. The 60% of evaluation is in continuous assessment and 40% of external evaluation pattern. Another important aspects of the course is that Information Technology, Mathematics and Statistics departments are situated in a single building neighbours to each other will make the smooth and fruitful running of the course.

Initially Department of Information Technology will initiate a one-year Post-Graduate diploma course in the coming academic year (2020-2021) with inputs from MCA/ M.Sc. Computer Science/M.Sc. Information Technology/ M.Sc. Bioinformatics/M. Tech./M.Sc. Mathematics/M.Sc. Statistics/ M.Sc. Applied Mathematics/ MSc. Physics/M.Sc. Electronics/M.Sc. Geology/M.Sc. Geography/ M.Sc. Biostatistics/M.Sc. Applied Statistics/M.A Economics/ B.Tech-MBA etc. After establishing and stabilizing this PG program the University can think of starting a full-fledged PG program for Data Science.

1.1 Objective of the Course

The Post Graduate Diploma in Data Science & Analytics is a one year program offered by School of Information Science & Technology in collaboration with department of Mathematical Science and Statistical Science is an excellent blend of knowledge and practice in the field of Data Science and its industrial applications. The program is targeted for creating qualified Data Science professionals. The Program also offers 10 weeks of industry oriented hands-on real – life analytical projects for students to get equipped with strong analytical and programming background which makes them highly competitive and employable on completion of the program.

1.2 Outcome of the Course

On completion of the course, the participants will learn the concept of Data Analytics using open source statistical tools like R, Python, Hadoop, Tableau and some very good visualization tools and techniques. They will be capable of implementing industry oriented Data Analytics Project.

1.3 Duration of the Course

Duration of the Post Graduate Diploma in Data Science shall be one (01) year- fulltime course divided into 2semesters. Each semester should have 18 weeks. Maximum time limit for completing the course is 4 semesters (two years)

1.4 Number of Seats:

This course has intake of **25** seats per semester and filling of seats as according to the rotation matrix maintained time-to-time by the Department of Information Technology/Kannur University norms. The rotation matrix of the seats to the course will be announced at the time of notification of the program.

1.5 Course Structure

This course contains total seven modules in first semester and two modules in second semester followed by 300 hours real time project work using any of the topics studied to earn the Diploma. All these components are mandatory for the completion of the course. The course comprises 30 Hours (5 × 6 Hours) per week comprising 18 weeks of teaching and learning activities.

1.6 Eligibility

MCA/ M.Sc. Computer Science/M.Sc. Information Technology/ M.Sc. Bioinformatics/M. Tech.(computer Science, Electrical & Electronics)/M.Sc. Mathematics/M.Sc. Statistics/ M.Sc. Applied Mathematics/ MSc. Physics/M.Sc. Electronics/M.Sc. Geology/M.Sc. Geography/ M.Sc. Biostatistics/M.Sc. Applied Statistics/M.A Economics/ B.Tech-MBA of this University or any other University / Institution, recognized by this University as equivalent thereto, with a minimum aggregate of 55% marks or equivalent grade.(For SEBC and Physically Challenged candidates the aggregate marks required is 50%. For SC and ST, a minimum pass in the degree examination is sufficient)

1.7 Selection Criteria of the candidates

The selection to the course shall be based on common admission test conducted by Kannur University. The test will be duration of two hours comprising 100 × 4 marks multiple choice questions from Computer Science, Mathematics, and Statistics and Aptitude and Mental ability in under graduate level. The pattern of question paper shall be as follows:

Sl.	Subjects	No. of Questions
1	Mathematics	25× 4 = 100
2	Statistics	25× 4 = 100
3	Computer Science	25× 4 = 100
4	Aptitude & Mental ability	25× 4 = 100
	Total	100× 4 = 400
	Each Right answer will be awarded 4 Marks	
	Each Wrong answer will be awarded -1 Mark	

1.8 Course Fee Structure

Sl.	Fee Details	Amount in Rs./-
1	Registration fee (Application Fee)	1,000/-
	For SC/ST	500/-

2	Admission Fee	555/-
3	Tuition Fee(Per Semester)	33,000/-
4	Library Fee	325/-
5	Student Welfare fund	360/-
6	Special fee	125/-
7	Caution Deposit (Refundable)	5,000/-
8	Laboratory Fee(Per Semester)	7,000/-
9	Student Affiliation Fee	440/-
10	Sports Affiliation Fee	220/-
11	University Union Fee	110/-
12	University Development Fund	60/-

1.9 Placement

This program is scheduled under the industrial collaboration and the experts from different industry have agreed to handle different sessions to the course to build thorough practical knowledge to the students and provide placement assistance to students who successfully qualify the course with the mandate required for the industry.

2.0 Programme Structure

2.1 Attendance

The minimum attendance required for each course shall be 75% of the total number of classes conducted for each semester. Those who secure the minimum attendance in a semester alone will be allowed to register for the End Semester Examination. Condonation of attendance to a maximum of 10 days in a semester subject to a maximum of two spells within a programme will be granted by the Vice-Chancellor. Benefit of condonation of attendance will be granted to the students on health grounds, for participating in University Union activities, meetings of the university bodies and participation in extracurricular activities on production of genuine supporting documents with the recommendation of the Head of the Department concerned. A student who is not eligible for condonation shall repeat the course with the subsequent batch.

2.2 Credits: One credit of the course is defined as a minimum of one hour lecture or a minimum of 2 hours lab/tutorial per week for 18 weeks in a Semester. The minimum credits required to complete Post Graduate Diploma in Data Science & Analytics (PGDDSA) programme is 42.

2.2 Seminar

Each student should select a relevant topic and prepare a seminar report, under the guidance of a faculty member. Students should prepare an abstract of the topic and distribute it to every faculty member at least two weeks ahead of the seminar. Presentation shall be for a minimum of 30 minutes duration. Presentation and seminar report will be evaluated by a group of at least two faculty members (Mark distribution:50% for report and 50% for presentation and discussion).

2.3 Assignments

Each student shall be required to submit a minimum of three assignments for each course. The details such as number of assignments, mark distribution and the weightage for each assignment will be announced by the faculty in charge of the course at the beginning of the semester.

2.4 Tests

A minimum of two class tests will be conducted for each course. The details such as number of tests, mark distribution and weightage for each test will be announced by the faculty in charge of the course at the beginning of the semester.

2.5 Seminar / Viva-voce / Case studies / Lab assignments

The faculty in charge of the course shall design the evaluation pattern based on one or more of these components and will be announced at the beginning of the semester.

2.6 Evaluation

Evaluation of the students shall be done by the faculty member who handles the course on the basis of continuous evaluation and End Semester Examination. The proportion of the distribution of marks including CE (Continuous Evaluation) and ESE (End Semester Examination) shall be 60-40.

2.6.1 Continuous Evaluation (CE): Continuous Evaluation (CE) of a course shall be based on periodic written tests, assignments, and Seminar / Viva-voce / Case studies/Project work/Attending workshops/Participating and presenting papers in Conferences/Publishing articles in Journals/Proceeding etc in respect of each courses

Components of Continuous Evaluation (Theory)

Sl.	Component	Marks
1	Seminar	10
2	Case studies / Project(individual)	15
3	Assignments	10
4	Presenting papers/Lectures (Outside the Department)/Group Discussion or work	10
5	Publishing Papers in Conference/ Book Chapters/ Publishing Articles in Approved Journals	15
	Total	60

NB: for Journal Publication 10 marks and Conference Proceeding/Book Chapters 05 Marks

Components of Continuous Evaluation (Practical)

Sl.	Component	Marks
1	Record Work/Lab Assignments	10
2	Implementing the experiment in the Lab	35
3	Viva-voce	15
	Total	60

2.6.2 Evaluation of Practical courses

The details regarding the CE as well as ESE for each practical course will be specified as part of the syllabus for the course.

2.7 End-Semester Evaluation (ESE)

All odd semester examinations will be conducted by the Head of the Department and even semester examinations will be conducted by the Controller of Examination, Kannur University. To conduct the end-semester examination, the Head of the department shall submit a confidential panel of examiners not less than ten experts from the outside of the Kannur University duly approved by the department council for the approval of Vice Chancellor. All teachers who engage classes on the course except industrial experts will be the members of the Board of Examiners (BoE) with Head of the Department as the chairman of the BoE. All the faculty in-charge of the course shall prepare and submit three (03) unique set of question papers for their course in odd semester well in advance to the Head of the department for the conduct of End Semester Examination of the respective batch. The Head of the department shall conduct a scrutiny meeting of the above question papers submitted by the concerned faculty by inviting at least two external experts from the list approved by the Vice Chancellor. The even semester examination question papers for the elective courses shall be set by the Controller of Examination by selecting the external experts of the question paper setters approved by the Vice Chancellor. The Head of the department shall submit the detailed syllabus, model question papers of the elective courses offered in the even semester to the Controller of Examination along with the panel of experts duly approved by the Vice Chancellor for setting the question papers to those electives soon after the commencement of the course.

2.7.1 Pattern of question papers and evaluation criteria for (ESE)

Pattern of Questions: Questions shall be set to assess knowledge acquired, standard application of knowledge, application of knowledge in new situations, critical evaluation of knowledge and the ability to synthesize knowledge. The duration of the examination is two (02) hours only.

Question paper for end semester theory examination shall consist of two parts

Part A

(Short essay type)

Answer all Questions.

Each question carries four marks.

- 1.
- 2
- 3
- 4
- 5

(5 × 4 = 20 Marks)

Part B

(Essay type)

Answer any two questions.

Each question carries 10 Marks.

- 6.
- 7.
- 8.

(2 × 10 = 20 Marks)

2.8 Project

A project work has to be undertaken by all students enrolled in the program. The project can be software related or establishing mathematical/statistical models evolved for the development of data science & analytics following software development lifecycle or an R&D related project. The hours allotted for project work may be clustered into a single slot so that students can do their work at a centre or location for a continuous period of time. The Major project work should be carried out in the Department /Institution or in an Industry / R & D organization of national repute. Project work shall be carried out under the supervision of a Teacher. If the project is carried out in an Industry / R & D organization outside the campus, then a co-guide shall be selected from the concerned organization. If the project work is of interdisciplinary nature, a co-guide shall be taken from the other department concerned. Every student should do the Project individually and no grouping is allowed. The candidates are required to get the synopsis and the guide approved by the department before the commencement of the project. A co-guide should be a postgraduate in CS/Application/IT/Mathematics/Statistics or allied subject or a person of eminence in the area in which student has chosen the project. At the end of the semester the candidate shall submit the Project report (two bound copies and one soft copy) duly approved by the guide and co-guide for End Semester Evaluation. The project report shall be prepared according to the guidelines appended along with this regulations/Guidelines approved by the University.

2.8.1 Evaluation of Project:

- i. A Departmental committee duly constituted by the Head of the Department will review the project periodically.
- ii. **Continuous Assessment of project work:** There shall be three internal presentations before the committee (Minimum two members, including the guide). The assessment is based on presentation, interim report and viva voce. The total mark for CA shall be divided among the three presentations in the ratio **20%:30%:50%**. Each internal presentation shall be evaluated based on the following components:

	COMPONENTS	% OF MARKS
i	Understanding of the problem / concepts	25%
ii	Adhering to methodology	20%
iii	Quality of presentation and demonstration (Demonstration is optional)	15%
iv	Quantum of work / effort	30%
v	Organization and content of Project report	10%

- iii. End Semester Assessment of Project: A board of two examiners appointed by the University shall conduct ESE evaluation. The evaluation shall be based on the report, presentation of the work, demonstration of the work and a detailed viva voce based on the work carried out. A candidate will not be permitted to attend the Project evaluation

without duly certified project reports. Also, a project will be evaluated only if the candidate attends the ESE presentation and Viva voce on the scheduled date and time. A board shall evaluate a maximum of eight (08) candidates in a day. The End Semester evaluation shall consist of the following components:

	COMPONENTS	% OF MARKS
i	Understanding of the problem/requirements/ concepts related to the project	15
ii	Adhering to methodology (Software engineering phases or research methodology) and the candidates understanding of the components of methodology	15
iii	Quality of Modeling of the problem and solution/ database design / form design / reports / testing (For research projects - relevance /novelty of the work(s)/ use of data/ proposal of new models /analysis of algorithms/ comparison and analysis of results /findings)	20
iv	Quality of presentation / demonstration	15
v	Quantum of work / effort - assessed through the content of report, presentation and viva	25
vi	Organization and content of report	10

- iv. A student shall pass in the Project course if she/he secures a separate minimum of 50 % for the external and 50% for ESE and CA put together.
- v. If a candidate fails in the evaluation of Project, he/she has to repeat the project course along with the next batch and undergo both CA and ESE. Unlike theory/practical courses, the CA mark will not be retained.
- vi. There shall be no improvement chance for the marks obtained in the Project course.

2.8.2 Guideline for preparing project Report

- i. Arrangement of contents:

The sequence in which the project report material should be arranged and bound should be as follows:

1. Cover Page & Title Page
2. Plagiarism Report
3. Bonafide Certificate
4. Abstract
5. Table of Contents
6. List of Tables
7. List of Figures
8. List of Symbols, Abbreviations and Nomenclature
9. Chapters

The chapters may be broadly divided into 3 parts (i) Introductory chapter, (ii) Chapters developing the main theme of the project work (iii) implementation details (if any) and Conclusion. The main text will be divided in to several chapters and each chapter may be further divided into several divisions and sub-divisions. Each chapter should be given an appropriate title.

Tables and figures in a chapter should be placed in the immediate vicinity of the reference where they are cited.

Footnotes should be used sparingly. They should be typed single space and placed directly underneath in the very same page, which refers to the material they annotate.

10. Appendices
11. References

The tables and figures shall be introduced at appropriate places.

- ii. Page Dimension and binding specifications:
The dimension of the project report should be in A4 size. The project report should be bound using flexible cover of the thick white art paper. The cover should be printed in black letters and the text for printing should be identical.
- iii. All the project report submitted by the students should be **plagiarism checked** using **Turnitin software** and the plagiarism report generated by the software should be verified and signed by the Head of the Department or person-in charge of the Project Coordinator.

3.0 Grading

An alphabetical grading system shall be adopted for the assessment of student's performance in a course. The grade is based on a ten-point scale. The following table gives the range of marks, grade points and the alphabetical grade.

Range of marks %	Grade points	Alphabetical grade
90–100	9	A+
80–89	8	A
70–79	7	B+
60–69	6	B
50–59	5	C
<50	0	F

A minimum of grade point 5 (Grade C) is needed for the successful completion of a course.

Performance of a student at the end of each semester is indicated by the Grade Point Average (GPA) and is calculated by taking the weighted average of grade points of the courses successfully completed. Following formula is used for the calculation. The average will be rounded off to two decimal places.

$$\text{GPA} = \frac{\text{Sum of (grade points in a course multiplied by its credit)}}{\text{Sum of credits of courses}}$$

The overall performance of a student is indicated by the Cumulative Grade Point Average (CGPA) and is calculated using the same formula given above.

Empirical formula for calculating the percentage of marks will be $\text{CGPA} \times 10 + 5$. Based on CGPA overall letter grade of the student shall be in the following way.

CGPA	Overall letter grade
8.5 and above	A+
7.5 and above but less than 8.5	A
6.5 and above but less than 7.5	B+
5.5 and above but less than 6.5	B
4.5 and above but less than 5.5	C

Conversion of Grades into classification

Classification	Overall letter grade
First Class with distinction	A+ and A
First Class	B+ and B
Second Class	C

3.1 Grade Card

The Controller of Examination, Kannur University is the authority to issue the semester wise grade card and consolidated grade statement and certificates on completion of the programme based on the authenticated documents submitted by the Head of the Department after the approval of the department council on end of each semester.

4.0 Supplementary Examinations for Failed Candidates

1. Candidates who have failed (F grade) in the semester examinations (except project work) can appear for the failed papers for the particular semester along with regular students. However, the Continuous Evaluation (CE) marks shall remain the same. Two such supplementary chances will be given for each semester within two years.
2. In the event of failure in Project Work the candidate shall re-register for project work, re do the project work and resubmit the project report a fresh for evaluation. The Continuous Evaluation marks shall be freshly allotted in this case.

Appearance for continuous Evaluation and End Semester Evaluation are compulsory and no grade shall be awarded to a candidate if he/she is absent for CE/ESE or both.

A student who fails to complete the programme / semester can repeat the full programme /semester once, if the department council permits so.

There shall be no provision for improvement of CE or ESE.

5.0 Department Council

This program is a joint venture of three departments such as Information Technology, Mathematical science and Statistical science with an industrial collaboration. So the conventional structure of Department council is not sufficient for dealing the smooth conduct of this course. Hence the department council for the Post Graduate Diploma in Data Science & Analytics may be reconstituted as follows:

1. **Chairman:** The Head of the Department of Information Technology.
2. **Members:** All faculty members who engages classes for this course.
3. The Head of the Department of Mathematical Science/ Senior Faculty nominated by the Head of the department concerned
4. The Head of the Department of Statistical Science/Senior Faculty nominated by the Head of the department concerned.
5. One /two experts from industry/Academician nominated by the Vice Chancellor as per the recommendation of the Department Council for a period of one year or duration of the course

6.0 Industrial Collaboration

This program is intended to make 100% industrial collaboration and the experts from the industry will handle different sessions for the course. The course comprised 10 modules including the industrial related project. Therefore each course, the expert from that domain handles at least hundred (100 [10 modules ×10 sessions =100 sessions]) sessions for each course. The duration of a session is two (02) hours. The Head of the Department will provide the remuneration/TA/DA to the expert as per the University norms. To smooth conduct of the

program the Finance Officer of the Kannur University has to make a provision to transfer the required fund to the Head of the Department of Information Technology on request soon after the beginning of the first semester of the course.

Details of the financial assistance required for the conduct of session by the industrial/academic experts.

Sl.	Description	Amount	Remarks
1	Remuneration/TA/DA for the experts from industry/academia	Rs. 7,00,000	(per Session Rs.4000/-)
2	Expenses for conducting odd semester examinations.	Rs. 35,000	
	Total	Rs.7,35,000	

NB: Economic class flight fare is eligible for the experts coming from other state/country with the prior approval from the Vice Chancellor of the Kannur University.

7.0 Grievance Redressal Mechanism

Committees will be constituted at the Department and University levels to investigate the written complaints regarding continuous Evaluation (CE). Department Level Committee (DLC) will consist of the Department Council and a student nominee of the Department Students' Union from the concerned faculty.

University Level Committee (ULC) will consist of the Pro-Vice-Chancellor (Chairman and Convener), the Convener of the Curriculum Committee (vice-chairman), the Head of the Department concerned and a nominee of the Students' Union. Department Level Committee will be presided over by the HOD and the University Level Committee by the Pro-Vice Chancellor. Department Level Committee will have initial jurisdiction over complaints against CE and University Level Committee will hear appeals against Department level decisions. Complaints will have to be submitted to the Department concerned within two weeks of publication of results of CE and disposed of within two weeks of receipt of complaint. Appeals to university Level Committee should be made within one month of the decision taken by the Department level committee and disposed within two months of the receipt of the complaint.

Complaints unsolved by the University Level Grievance committee shall be placed before the Vice Chancellor.

COURSE STRUCTURE

Semester	Theory	Practical
Semester I	5	2
Semester II	2 electives	1 (project)

Semester I

Subject Code	Subject	Instructional Hrs./week			Marks			Credit
		L	P	T	CE	ESE	Tot	
PGDSC01	Mathematical and Statistical Methods for Data Science Using R	3	0	0	60	40	100	4
MCSC02	Introduction to Data Science and Algorithm Design	3	0	0	60	40	100	4
MCSC03	Machine Learning for Data Science and Introduction to Python	4	0	0	60	40	100	4
MCSC04	Business Analytics and Prediction	4	0	0	60	40	100	3
MCSC05	Advanced Computational Methods for Data Science	4	0	0	60	40	100	3
MCSC06	Lab –I	0	6	2	60	40	100	3
MCSC07	Lab–II	0	6	3	60	40	100	3
Total		18	12	5	420	280	700	24

Semester II

Subject Code	Subject	Instructional Hrs./week			Marks			Credit
		L	P	T	CE	ES	Tot	
MCSC01	Elective I	4	0	0	60	40	100	4
MCSC02	Elective II	4	0	0	60	40	100	4
MCSC03	Project		12	5	120	80	200	10
Total		8	12	5	240	160	400	18

ELECTIVES

Subject Code	Subject
Elective1	Advanced Machine Learning
	Natural Language Processing
Elective2	Big Data Analytics
	Data Warehousing
Elective3	Deep Learning
	Time Frequency Analysis
Elective4	Artificial Intelligence
	Computer Vision

Course Contents:-

Subject I: Mathematical and Statistical Methods for Data Science Using R

1. Fundamentals of R
2. Basic Linear algebra for data science
3. Linear programming problem
4. Basic statistical concepts and important sampling methods.
5. Data cleaning and visualization using R
6. Measures of Central Tendency, dispersion, skewness and kurtosis
7. Pearson correlation coefficient, rank correlation, intra-class correlation
8. Basics of probability and random variable
9. Probability distributions-Binomial, Poisson, Uniform, Normal, Beta and Gamma and multinomial and multivariate normal distributions.
10. Statistical inference: basics of estimation, testing (parametric and non-parametric) and confidence interval estimation

11. Maximum likelihood estimation and Bayesian estimation.
12. Basic concepts of Markov chain

Subject II: Introduction to Data Science and Algorithm Design

1. Basics of Data and data science
2. Data Science Pipeline – Data Wrangling, Exploratory Analysis, Modeling
3. Data structure: Linear and non- linear data structure
4. Important problem types, algorithm design, decisions prior to designing
5. Basic technique for design of efficient algorithm
6. Brute Force approach
7. Divide-and-Conquer approach
8. Branch-and-Bound technique
9. Greedy method
10. Dynamic Programming
11. Backtracking.
12. Importance of algorithm analysis, time and space complexity
13. Techniques to analyze algorithm

Subject III: Machine Learning for Data science and Introduction to Python

1. Introduction to python programming
2. Basics of Machine Learning
3. Supervised Machine Learning – K-NN, Naïve Bayes, Decision tree, SVM
4. Unsupervised Machine Learning – K means, Apriori algorithms
5. Computation with Python – NumPy, SciPy
6. Data Manipulation in Python- Pandas
7. Understanding Data Frame
8. Data Visualization in Python - matplotlib
9. Introduction to Scikit – Machine learning
10. Web Scraping in Python – BeautifulSoup
11. Introduction to Hadoop/Map Reduce

Subject IV: Business Analytics and Prediction

1. Introduction to analytics
2. Simple and multiple linear regressions
3. Regression diagnostics
4. Logistic regression
5. Time series analysis and forecasting
6. Autoregressive and moving average models
7. Graphics programming: charts, graphs, animations, and techniques for visualization of high dimensional data
8. Presentation and visualization of data for effective communication
9. TABLEAU: Visual Analytics | Mapping | Calculations | Dashboard and Stories
10. POWER BI: Intro to Power BI | Visualization with BI | Data Analysis Expressions

Subject V: Advanced Computational Methods for Data Science

1. Basics of statistical simulation
2. Monte Carlo simulation
3. Basic concepts of Bayesian computing
4. MCMC methods, MH algorithm and Gibbs sampling
5. Multivariate data visualization and analysis
6. Dimension reduction methods (PCA, CCA and FA)
7. Expectation Maximization (EM) algorithm
8. Gaussian Mixture Models (GMM)
9. Mixture Regression Models
10. Hidden Markov Models (HMM)