

Semester- 1

Data Science and Management			
Course Code	MCS102	CIE Marks	50
Teaching Hours/Week (L:P:SDA)	4:0:2	SEE Marks	50
Total Hours of Pedagogy	50	Total Marks	100
Credits	03	Exam Hours	03
Course Learning objectives:			
<ul style="list-style-type: none">• Explain the foundational concepts of data science, including its history, significance, and the data science process.• Apply statistical methods and data analysis techniques to interpret and draw insights from complex datasets.• Implement various machine learning algorithms and assess their performance using appropriate evaluation metrics in real-world scenarios.• Utilize data visualization tools and techniques to effectively communicate findings and insights to diverse audiences.			
Module-1			
Module 1: Chapter 1. Introduction: What Is Data Science?, Big Data and Data Science Hype Getting Past the Hype Why Now? Datafication, The Current Landscape (with a Little History),Data Science Jobs, A Data Science Profile, Chapter 2. Statistical Inference, Exploratory Data Analysis, and the Data Science Process, Statistical Thinking in the Age of Big ,Data, Statistical Inference, Populations and Samples, Populations and Samples of Big Data, Big Data Can Mean Big Assumptions, Modeling, Exploratory Data Analysis, Philosophy of Exploratory Data Analysis,, Exercise: EDA, The Data Science Process, A Data Scientist's Role in This Process.			
Module-2			
Module 2: . Algorithms, Machine Learning Algorithms, Three Basic Algorithms, Linear Regression,k-Nearest Neighbors (k-NN),k-means, Naive Bayes, Bayes Law, A Spam Filter for Individual Words,A Spam Filter That Combines Words: Naive Bayes Exercise: Basic Machine Learning Algorithms, Spam Filters, Naive Bayes, and Wrangling, Logistic Regression ,Thought Experiments Classifiers, Logistic Regression Case Study, Estimating α and β ,Newton's Method, Stochastic Gradient Descent, Implementation, Evaluation, Sample R Code.			
Module-3			
Module 3: Time Stamps and Financial Modeling,Kyle Teague and GetGlue,Timestamps,Exploratory Data Analysis (EDA),Metrics and New Variables or Features,Financial Modeling,In-Sample, Out-of-Sample, and Causality,Preparing Financial Data,Log Returns			

7. 2

Module-4

Module 4:Extracting Meaning from Data,Background: Data Science Competitions,Background: Crowdsourcing,The Kaggle Model,A Single Contestant,Their Customers, Data Visualization and Fraud Detection,Data Visualization History,A Sample of Data Visualization Projects,Mark's Data Visualization ,Data Science and Risk.About Square.The Risk Challenge.The Trouble with Performance Estimation, Model Building Tips.

Module-5

Module 5:Data Engineering: MapReduce, Pregel, and Hadoop, MapReduce, Word Frequency Problem,Enter MapReduce, Other Examples of MapReduce, Pregel, Thought Experiment, On Being a Data Scientist, Data Abundance Versus Data Scarcity, Designing Models,, Economic Interlude: Hadoop, A Brief Introduction to Hadoop, Cloudera, Next-Generation Data Scientists, Hubris, and Ethics.

Assessment Details (both CIE and SEE)

The weightage of Continuous Internal Evaluation (CIE) is 50% and for Semester End Exam (SEE) is 50%. The minimum passing mark for the CIE is 50% of the maximum marks. Minimum passing marks in SEE is 40% of the maximum marks of SEE. A student shall be deemed to have satisfied the academic requirements and earned the credits allotted to each subject/ course if the student secures not less than 50% (50 marks out of 100) in the sum total of the CIE (Continuous Internal Evaluation) and SEE (Semester End Examination) taken together.

Continuous Internal Evaluation:

1. Two Unit Tests each of **25 Marks**
2. Two assignments each of **25 Marks** or **one Skill Development Activity of 50 marks**

to attain the COs and POs

The sum of two tests, two assignments/skill Development Activities, will be **scaled down to 50 marks**

CIE methods /question paper is designed to attain the different levels of Bloom's taxonomy as per the outcome defined for the course.

Semester-End Examination:

1. The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to 50.
2. The question paper will have ten full questions carrying equal marks.
3. Each full question is for 20 marks. There will be two full questions (with a maximum of four sub-questions) from each module.
4. Each full question will have a sub-question covering all the topics under a module.
5. The students will have to answer five full questions, selecting one full question from each module

7

Suggested Learning Resources:**Books****Textbooks:**

1. " Doing Data Science- Straight talk from the Frontline" , Rachel Schutt & Cathy O'Neil 2014,Orielly publication
2. "Data Science from Scratch: First Principles with Python" by Joel Grus, 2nd Edition (2019)

Reference Books:

3. "The Elements of Statistical Learning" by Trevor Hastie, Robert Tshigami, and Jerome Friedman, 2nd Edition (2009)

Web links and Video Lectures (e-Resources):

- <https://www.coursera.org/specializations/jhu-data-science>
- <https://www.kaggle.com/learn/data-science>
- <https://www.edx.org/professional-certificate/harvardx-data-science>
- <https://www.youtube.com/playlist?list=PL4cUxeGkcC9g1s4L6G8p8Fq5XK6Pq7b1k>

Skill Development Activities Suggested

- The students, with the help of the course teacher can take up relevant technical activities which will enhance their skill. The prepared report shall be evaluated for CIE marks.

Course outcome (Course Skill Set)

At the end of the course the student will be able to :

Sl. No.	Description	Blooms Level
CO1	Explore the foundational concepts of data science, history, significance, and process.	L3
CO2	Apply statistical methods and data analysis techniques to interpret and draw insights from complex datasets.	L3

72

Program Outcome of this course

Sl. No.	Description	POs
1	Demonstrate the ability to independently conduct research and development work to address practical engineering problems.	PO1
2	Develop and deliver comprehensive technical presentations that effectively convey complex information to diverse audiences.	PO2
3	Exhibit mastery in the specialized study area, surpassing the requirements of a relevant bachelor's program.	PO3
4	Analyse engineering problems critically and apply appropriate techniques, skills, and modern tools to develop innovative solutions.	PO4

Mapping of COS and POs

	PO1	PO2	PO3	PO4	PO5	PO6
CO1	x			x		x
CO2			x		x	
CO3		x				
CO4	x					x

Request to Circulate Updated Syllabus for Data Science and Management (CSE MTech Stream)_Final

1 message

Dept of CSE,VTU Belagavi <chair.cse@vtu.ac.in>

Thu, Jan 16, 2025 at 1:19 PM

To: registrar@vtu.ac.in

Cc: sbhvtuso@gmail.com, sbhvtuso2022@gmail.com, sbhalbhavi@vtu.ac.in, "sld@vtu.ac.in" <sldeshpande@gmail.com>, sld@vtu.ac.in

Dear Sir,

I would like to inform you that the syllabus for the course **Data Science and Management (MCS102)** has been updated based on the feedback from faculty members. This course is common to all branches of Computer Science and Engineering (CSE).

Action Required:


1. Please circulate the updated syllabus to all concerned parties.
2. Ensure that the updated syllabus is uploaded to the official website.

Please find the attached in word and pdf format.

Thanking you

Regards,

Dr S L Deshpande
Chairperson
Department of Computer Science and Engineering,
Visvesvaraya Technological University, Belagavi-590018
Phone:0831-2498251

**2 attachments** **Data Science and Management_Sem1_MTech_AllBranchesCommon.pdf**
159K **Data Science and Management_Sem1_MTech_AllBranchesCommon.doc**
69K