

## SCHEME OF COURSE WORK

### Course Details:

<b>Course Title</b>	: FUNDAMENTALS OF DATA SCIENCE								
<b>Course Code</b>	: 22CD1101	<b>L</b>	<b>T</b>	<b>P</b>	<b>C</b>	3	0	0	3
<b>Program:</b>	: B.Tech.								
<b>Specialization:</b>	: Data Science								
<b>Semester</b>	: 3rd Semester								
<b>Prerequisites</b>	:								
<b>Courses to which it is a prerequisite</b>	:								

### Course Outcomes (COs):

1	Describe the significance of data science and understand the Data Science process.
2	Explain how data is collected, managed and stored for data science.
3	Build, and prepare data for use with a variety of statistical methods and models.
4	Analyze Data using various Visualization techniques.
5	Choose contemporary models, such as machine learning, AI, techniques to solve practical problems.

### Program Outcomes (POs):

A graduate of data Science Specialization will be able to

1	Graduates will be able to apply the knowledge of Mathematics, Science, Engineering Fundamentals, Principles of Computer Science and Engineering and Data Science to solve complex problems in different domains.
2	Graduates can identify a problem, conduct data analysis experiments to interpret data, and formulate appropriate strategies to evolve solution using the basic principles of science and engineering.
3	Graduates will have the ability to design a component or a product applying all the relevant standards and with realistic constraints, including public health, safety, culture, society and environment.
4	Graduates will be able interpret data, extract meaningful information, and assess findings as per the prevailing research methods.
5	Graduates will be able to choose and apply new tools and innovative methodologies necessary for engineering practice to solve data-driven problems.
6	Graduates will be able to analyze the impact of data analytic systems on individuals, organizations, society and design appropriate solutions related to society health and safety.
7	Graduates will have adaptive thinking and adaptability in relation to environmental context and sustainable development
8	Graduates will be able to have clear understanding of professional thinking and innovation to provide more economical and effective solutions.
9	Graduates will be able to have cross cultural competency exhibited by working as a member or in teams identifying the significance of each individual's contribution.
10	Graduates will be able to have a good working knowledge of communicating in English – communication with engineering community and society
11	Graduates will be able to have good cognitive load management skills and manage data science projects adhering to financially viable options
12	Graduates will engage themselves in independent and continuous learning in the broad context of data science and other computer Science related domains.

**Course Outcome versus Program Outcomes:**

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO-1	M	M		S	M							
CO-2	S	M	S	S	M					S		
CO-3	S	S	M	S	M					S		
CO-4	S	S		S	S							
CO-5	M	M	M	M	S	S						M

*S* - Strongly correlated, *M* - Moderately correlated, *Blank* - No correlation

<b>Assessment Methods:</b>	Assignment / Quiz / Seminar / Case Study / Mid-Test / End Exam
----------------------------	--

## Teaching-Learning and Evaluation

Week	TOPIC / CONTENTS	Course Outcomes	Sample questions	TEACHING-LEARNING STRATEGY	Assessment Method & Schedule
1	<b>Introduction To Data Science:</b> Definition, Big Data and Data Science Hype, Datafication, Data Science Profile, Meta-Definition, Data Scientist,	CO1	1. Define Data Science 2. Explain Datafication 3. Explain roles of data Scientist	Lecture / PPT /Group Discussion	Assignment-1, Test- 1 Quiz-1
2	Statistical Inference, Populations and Samples Populations and Samples of Big Data, Big Data Can Mean Big Assumptions, Modeling,	CO1	1. What is Statistical Inference 2. Define population and Sample 3. Describe Big data	Lecture / PPT	Assignment-1, Test- 1 Quiz-1
3	Philosophy of Exploratory Data Analysis, The Data Science Process, A Data Scientist's Role in this Process Case Study: RealDirect.	CO1	1. Describe in detail data science process with a neat diagram 2. Explain role of Data scientist in Data science process	Lecture / PPT	Assignment-1, Test- 1 Quiz-1
4	<b>Mathematical Preliminaries:</b> Probability, Descriptive Statistics, Correlation Analysis	CO2	1. Explain Correlation analysis and correlation coefficient 2. Explain descriptive statistics	Lecture / PPT	Assignment-1, Test- 1 Quiz-1
5	<b>Data Munging:</b> Properties of Data, Languages for Data Science, Collecting Data, Cleaning Data, Crowdsourcing.	CO2	1. Explain different categories of data 2. Write a note on languages for data science 3. Explain the methods to clean the data	Lecture / PPT	Assignment-1, Test- 1 Quiz-1
7	<b>Scores and Rankings:</b> Developing Scoring Systems, Z-scores and Normalization, Advanced Ranking Techniques	CO3	1. What is Z-scores Normalization 2. Explain Advanced Ranking Techniques	Lecture / PPT	Assignment-1, Test- 1 Quiz-1
6	<b>MID TEST-1</b>				
8	<b>Statistical Analysis:</b> Sampling from Distributions, Statistical Distributions, Statistical Significance, Permutation Tests and P-values	CO3	1. Briefly explain statistical distributions 2. Describe p-test	Lecture / PPT	Assignment-1, Test- 1 Quiz-1

9	<b>Visualizing Data:</b> Exploratory Data Analysis, Developing a Visualization Aesthetic, Chart Types, Great Visualizations	CO4	1. Explain different chart types for visualization 2. What is Exploratory Data Analysis	Lecture / PPT	Assignment-1, Test- 1 Quiz-1
10	<b>Mathematical Models:</b> Philosophies of Modeling, A Taxonomy of Models, Baseline Models, Evaluating Models, Evaluation Environment	CO4	1. Explain taxonomy of models 2. How are models developed are evaluated	Lecture / PPT	Assignment-2, Test- 2 Quiz-2
11	<b>Supervised Learning:</b> Linear Regression, Better Regression Models, Regression as Parameter Fitting, Simplifying Models through Regularization	CO5	1. Explain Linear Regression with example 2. What is Gradient Descent	Lecture / PPT	Assignment-2, Test- 2 Quiz-2
12	Classification and Logistic Regression, Issues in Logistic Classification, Naive Bayes, Decision Trees Classifiers	CO5	1. Explain Classification 2. Describe Logistic Regression with example 3. How are decision Trees constructed	Lecture / PPT	Assignment-2, Test- 2 Quiz-2
13	<b>MID TEST-2</b>				
14	<b>END EXAM</b>				