Data Science and Visualization	**3**	**0**	**2**	Probability, Statistics, Linear Algebra

Course Objective: 1. To familiarize with different types of data and its visualization.
2.To understand and practice data pre-processing and data exploration.
3. To solve real-world analytical problems in data science.

S. NO.	Course Outcomes (CO)
CO1	Ability to identify different types of data and data distributions.
CO2	Ability to understand and apply different data cleaning and data transformation techniques.
CO3	Ability to understand and implement different data visualization techniques.
CO4	Ability to understand and execute different data exploration techniques.
CO5	Ability to implement different real-world applications of data science.

S. NO.	Contents	Contact Hours
UNIT 1	Introduction to data science: Basics of Probability & Statistics (Random Variables, Bayes's Theorem, Normal distribution, Central Limit Theorem). Defining data science, Recognizing different types of data, Data distributions. Data acquisition and data storage.	**10**

UNIT 2	Data pre-processing: Missing data problem, Outlier definition. Data cleaning, Data transformation or data wrangling procedures such as merging, ordering and aggregating.	**10**
UNIT 3	Data visualization: Introduction to data visualization. Definition of Dashboard, Dashboard design and principles. Basic charts and plots, Box plots, Histogram, Graphs, Networks, Hierarchies, Reports.	**10**
UNIT 4	Exploratory Data Analysis: Data exploration for univariate data. Outlier detection techniques. Descriptive statistics (mean, standard deviation etc.) for data exploration. Correlation statistics for data exploration. Data exploration for multivariate data. Use of multivariate visualization tools such as bar charts, bar plots, heat maps, bubble charts, run charts, and scatter plots.	**12**
	TOTAL	**42**