

Reg. No. :

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

**Question Paper Code: UEG71**

B.E. / B.Tech. DEGREE EXAMINATION, NOV 2024

Professional Elective

21ADVG71-FUNDAMENTALS OF DATA SCIENCE

(Regulations 2021)

(Common to Biomedical & Mechanical Engineering Branches)

Duration: Three hours

Maximum: 100 Marks

Answer All Questions

PART – A (5 x 20= 100 Marks)

1. (a) Apply the concept of R Charts and Graphs with syntax and example program. CO2-App (20)

Or

(b) Apply the concept of R interfaces with syntax and example program. CO2-App (20)

2. (a) Define the clustering methods in detail. CO1-U (20)

Or

(b) Define the Decision tree in detail. CO1-U (20)

3. (a) Define Naïve Bayes theorem with example. CO1-U (20)

Or

(b) Define Clustering and its methods with necessary diagrams. CO1-U (20)

4. (a) Discuss the challenges and considerations involved in designing a big data architecture that can handle the processing and analysis of large volumes of data. Provide examples of different architectural models and their advantages and disadvantages. CO1-U (20)

Or

- (b) Explain the importance of data quality in big data solutions. CO1-U (20)  
Discuss the key steps involved in ensuring data quality, such as data cleansing, transformation, and validation. Provide examples of tools and techniques that can be used to ensure data quality in big data solutions
5. (a) Explain the concept of social media sentiment analysis using big data. How can businesses use sentiment analysis to improve their social media strategies? CO1-U (20)
- Or
- (b) Explain the purpose and architecture of Hadoop Distributed File System (HDFS). How does HDFS address the storage needs of big data applications? CO1-U (20)