R22 Regulation *Subject code:4E6HA*

TKR COLLEGE OF ENGINEERING AND TECHNOLOGY

(Autonomous, Accredited by NAAC with 'A+' Grade)

B.Tech VI Semester Regular/Supplementary Examinations, April 2026

BIGDATA ANALYTICS

(CSE(DS))

Maximum Marks: 60 Date:13.04.2026 Duration: 3 hours

Note: **1.This question paper contains two parts A and B.**
2. Part A is compulsory which carries 10 marks. Answer all questions in Part A.
3. Part B consists of 5 Units. Answer any one full question from each unit.
4. Each question carries 10 marks and may have a, b, c, d as sub questions.

Part-A

All the following questions carry equal marks (10X1M=10 Marks)		Marks	CO	Bloom Tx
1.a)	What is the role of analytics in decision making?	1M	1	L1
b)	List any two layers of the Big Data stack.	1M	1	L1
c)	How do queues help in load leveling?	1M	2	L2
d)	State one application of Pipes and Filters in analytics.	1M	2	L1
e)	List any two data acquisition considerations.	1M	3	L2
f)	Why is Spark faster than traditional disk-based processing systems?	1M	3	L1
g)	What is the output of a Reduce function?	1M	4	L1
h)	Why is batch processing suitable for news article analysis?	1M	4	L1
i)	What is meant by a table (relation) in SQL databases?	1M	5	L1
j)	How do document databases differ from relational databases?	1M	5	L2

Part-B

Answer All the following questions. (5X10M=50Marks)		Marks	CO	Bloom Tx
2	a) Discuss how the characteristics of Big Data create challenges for data processing.	5M	1	L6
	b) Explain the importance of data preparation in Big Data analytics.	5M	1	L2
	OR			
3	a) Differentiate between batch analytics and stream analytics.	5M	1	L4
	b) Explain how visualization helps in extracting insights from Big Data.	5M	1	L2
4	a) Explain the major components of an analytics architecture with a neat diagram.	5M	2	L2
	b) Explain MapReduce and its operations with example.	5M	2	L2
	OR			
5	a) Describe different sharding strategies used for large-scale data storage.	5M	2	L2
	b) Explain the Scheduler–Agent–Supervisor design pattern in analytics systems.	5M	2	L2

6	Describe the components and working of Big Data collection systems used for large-scale data ingestion.	10M	3	L2
	OR			
7	Explain the role of Scala in developing scalable and high-performance analytics applications.	10M	3	L2
8	Describe the architecture of Pig and explain the functions of Pig Latin, parser, optimizer, and execution engine.	10M	4	L2
	OR			
9	Write a MapReduce program to find the maximum temperature recorded for each year from NCDC data.	10M	4	L6
10	a) Explain NoSQL databases and justify their need in modern data-intensive applications. b) Describe the MVT (Model–View–Template) architecture used in Django.	5M 5M	5 5	L2 L2
	OR			
11	a) Compare key–value databases with relational databases. b) Discuss use cases of column family databases in Big Data applications.	5M 5M	5 5	L4 L6