| Question Paper Code | 13989 |
|---|---|

## M.E. / M.Tech - DEGREE EXAMINATIONS, NOV / DEC 2025
Third Semester
### Big Data Analytics
### 24PBDOE301 - BUSINESS DATA ANALYTICS
Regulations - 2024

Duration: 3 Hours                                                    Max. Marks: 100

### PART - A (MCQ) (10 × 1 = 10 Marks)
Answer ALL Questions

| | | Marks | K – Level | CO |
|---|---|---|---|---|
| 1. | Which of the following best describes Business Analytics? <br> (a) Collection of data      (b) Interpretation of data to support decision-making <br> (c) Designing marketing campaigns    (d) Accounting analysis | 1 | K1 | CO1 |
| 2. | Which of the following is not a driver for Business Analytics? <br> (a) Data availability      (b) Increased computing power <br> (c) Lack of skilled professionals    (d) Customer-centric focus | 1 | K1 | CO1 |
| 3. | In descriptive statistics, mean is also called: <br> (a) Median      (b) Average      (c) Frequency      (d) Variance | 1 | K1 | CO2 |
| 4. | Which measure indicates how far data values deviate from the mean? <br> (a) Range      (b) Variance      (c) Mode      (d) Frequency | 1 | K1 | CO2 |
| 5. | In hypothesis testing, the probability of rejecting a true null hypothesis is known as: <br> (a) Confidence level (b) P-value      (c) Type I Error    (d) Type II Error | 1 | K2 | CO3 |
| 6. | The function of HDFS in Hadoop is to: <br> (a) Process data      (b) Store data across multiple nodes <br> (c) Manage databases      (d) Analyze structured data | 1 | K1 | CO4 |
| 7. | MapReduce works based on the principle of: <br> (a) Parallel processing      (b) Sequential storage <br> (c) Centralized computation    (d) Manual scripting | 1 | K1 | CO4 |
| 8. | Hive Query Language (HQL) is similar to: <br> (a) SWL      (b) Python      (c) Java      (d) HTML | 1 | K1 | CO5 |
| 9. | MongoDB is an example of: <br> (a) Relational DB    (b) NoSQL DB    (c) Graph DB    (d) Distributed File System | 1 | K1 | CO5 |
| 10. | Which of the following tools is used for data transfer between Hadoop and RDBMS? <br> (a) Sqoop      (b) Pig      (c) Spark      (d) Drill | 1 | K1 | CO5 |

### PART - B (12 × 2 = 24 Marks)
Answer ALL Questions

| | | Marks | K – Level | CO |
|---|---|---|---|---|
| 11. | Define business analytics and mention any two of its applications. | 2 | K1 | CO1 |
| 12. | Summarize any two key skills required for a business analyst. | 2 | K2 | CO1 |
| 13. | Differentiate between qualitative and quantitative data. | 2 | K2 | CO2 |
| 14. | What is a data dashboard? Give one business use case. | 2 | K2 | CO2 |
| 15. | Define hypothesis testing. Give one real-world example. | 2 | K1 | CO3 |
| 16. | Recall the concept of conditional probability. | 2 | K1 | CO3 |
| 17. | What is hadoop distributed file system (HDFS)? | 2 | K1 | CO4 |
| 18. | State two advantages of MapReduce framework. | 2 | K2 | CO4 |
| 19. | List any two features of NoSQL databases. | 2 | K1 | CO5 |
| 20. | Differentiate between Hive and Pig. | 2 | K2 | CO5 |
| 21. | Outline the importance of data visualization in business analytics. | 2 | K2 | CO2 |
| 22. | What are the major components of the business analytics life cycle? | 2 | K1 | CO1 |

*K1 – Remember; K2 – Understand; K3 – Apply; K4 – Analyze; K5 – Evaluate; K6 – Create*                **13989**

## PART - C (6 × 11 = 66 Marks)
### Answer ALL Questions

23. a) i) A retail chain wants to improve customer satisfaction by analyzing feedback data collected through surveys and social media. Explain how the business analytics process can be applied to this scenario.    6   K2   CO1

     ii) Illustrate the type of analytics most useful in this case.    5   K2   CO1

**OR**

   b) A healthcare provider aims to reduce patient wait time. Illustrate how business analytics can help identify the bottlenecks and suggest data-driven decisions.    11   K2   CO1

24. a) Using suitable examples, demonstrate how descriptive statistics and visualization tools can help a company monitor sales performance.    11   K3   CO2

**OR**

   b) Analyze dataset (hypothetical) showing customer ratings across four product categories. Suggest two data visualization types that best communicate trends to management.    11   K3   CO2

25. a) A company wants to verify if it's new training program improved employee productivity. Formulate null and alternate hypotheses and describe how to test them statistically.    11   K2   CO3

**OR**

   b) Explain how sampling and confidence intervals help decision-making when analyzing large-scale marketing data?    11   K2   CO3

26. a) An e-commerce company handles 1 TB of clickstream data daily. Explain how Hadoop and MapReduce can be used to process and extract user behavior insights.    11   K2   CO4

**OR**

   b) Outline the role of HDFS and MapReduce in distributed big data processing. Provide a practical example of an analytical task.    11   K2   CO4

27. a) Compare and contrast Hive, Pig, and Spark in handling large datasets. Illustrate with one example each.    11   K2   CO5

**OR**

   b) A startup uses MongoDB to store user interaction data. Relate how it can be integrated with Hadoop for better analytical performance.    11   K2   CO5

28. a) Explain how data-driven decision-making supports organizational strategy with a suitable business case.    11   K2   CO1-5

**OR**

   b) A telecom company wants to predict customer churn using historical data. Outline the steps in a predictive analytics model suitable for this case.    11   K2   CO1-5