

CLASS: XII
ARTIFICIAL INTELLIGENCE (SUBJECT CODE - 843)
SESSION 2026-27
UNIT-5 Introduction to Big Data and Data Analytics

1. Big Data refers to:

- A. Small and simple datasets
- B. Extremely large and complex datasets
- C. Only numerical data
- D. Only social media data

Answer: B

2. Which of the following is NOT a source of Big Data?

- A. Transactional data
- B. Machine data
- C. Social data
- D. Handwritten notebooks

Answer: D

3. Amazon uses Big Data mainly for:

- A. Weather forecasting
- B. Product recommendations
- C. Space research
- D. Printing books

Answer: B

4. Structured data is:

- A. Unorganized
- B. Stored with a defined structure
- C. Only images and videos
- D. Difficult to search

Answer: B

5. Which of the following is an example of structured data?

- A. Audio files
- B. Videos
- C. Transaction records
- D. Social media posts

Answer: C

6. JSON files are examples of:

- A. Structured data
- B. Semi-structured data
- C. Unstructured data
- D. Multimedia data

Answer: B

7. Which type of data has no fixed format?

- A. Structured
- B. Semi-structured
- C. Unstructured

D. Tabular

Answer: C

8. Emails are generally classified as:

A. Structured

B. Semi-structured

C. Unstructured

D. Numerical

Answer: C

9. Which is an advantage of Big Data?

A. Increased security risks

B. Better customer insights

C. Data quality issues

D. High implementation cost

Answer: B

10. Big Data helps organizations in:

A. Guesswork

B. Data-driven decision making

C. Avoiding analysis

D. Reducing data collection

Answer: B

11. A major disadvantage of Big Data is:

A. Better productivity

B. Innovation

C. Privacy concerns

D. Customer satisfaction

Answer: C

12. GDPR is related to:

A. Data protection

B. Programming

C. Networking

D. Robotics

Answer: A

13. The original framework for Big Data characteristics is:

A. 2Vs

B. 3Vs

C. 5Vs

D. 7Vs

Answer: B

14. Which of the following is NOT one of the 3Vs?

A. Volume

B. Velocity

C. Variety

D. Value

Answer: D

15. Velocity refers to:

- A. Data size
- B. Data accuracy
- C. Speed of data generation
- D. Data value

Answer: C

16. Volume refers to:

- A. Amount of data
- B. Speed of data
- C. Data quality
- D. Data format

Answer: A

17. Variety means:

- A. Data quality
- B. Different forms of data
- C. Data security
- D. Data duplication

Answer: B

18. Which V focuses on data accuracy?

- A. Value
- B. Variety
- C. Veracity
- D. Velocity

Answer: C

19. The benefits obtained from Big Data are referred to as:

- A. Velocity
- B. Value
- C. Variety
- D. Variability

Answer: B

20. Variability in Big Data refers to:

- A. Storage size
- B. Changing data patterns
- C. Data security
- D. Data ownership

Answer: B

21. Which company example is used in the chapter?

- A. Netflix
- B. Amazon
- C. OnDemandDrama
- D. Flipkart

Answer: C

22. Real-time recommendations are related to:

- A. Volume
- B. Velocity
- C. Veracity

D. Value

Answer: B

23. Big Data Analytics deals with:

- A. Small datasets only
- B. Structured data only
- C. Large and diverse datasets
- D. Printed records only

Answer: C

24. Data Analytics is used to:

- A. Hide information
- B. Discover patterns and trends
- C. Delete databases
- D. Increase hardware size

Answer: B

25. Which analytics predicts future outcomes?

- A. Descriptive
- B. Diagnostic
- C. Predictive
- D. Historical

Answer: C

26. Moore's Law contributed to Big Data by:

- A. Reducing data
- B. Increasing computing power
- C. Decreasing storage
- D. Eliminating networks

Answer: B

27. Which trend generates massive user-generated content?

- A. Cloud Computing
- B. Mobile Computing
- C. Social Networking
- D. Robotics

Answer: C

28. Cloud Computing allows:

- A. Local-only storage
- B. Remote access to resources
- C. No internet usage
- D. Data deletion

Answer: B

29. The first step in Big Data Analytics is:

- A. Analyze Data
- B. Clean Data
- C. Gather Data
- D. Visualize Data

Answer: C

30. Batch processing works on:

- A. Real-time streams only
- B. Large blocks of data
- C. Images only
- D. Audio only

Answer: B

31. Stream processing helps in:

- A. Faster decision making
- B. Data deletion
- C. Reducing storage devices
- D. Printing reports

Answer: A

32. Data cleaning removes:

- A. Useful information
- B. Duplicate and irrelevant data
- C. Hardware errors
- D. Networks

Answer: B

33. Missing data may lead to:

- A. Better results
- B. Accurate insights
- C. Inaccurate insights
- D. Faster processing

Answer: C

34. Which is a Big Data Analytics tool?

- A. MS Paint
- B. Tableau
- C. Notepad
- D. Calculator

Answer: B

35. Hadoop is mainly used for:

- A. Big Data processing
- B. Gaming
- C. Graphic design
- D. Video editing

Answer: A

36. In Orange, the File widget is used to:

- A. Train models
- B. Load datasets
- C. Visualize graphs
- D. Test models

Answer: B

37. Normalization scales values to:

- A. Increase size
- B. Fixed range such as 0–1
- C. Delete values

D. Create duplicates

Answer: B

38. Which widget is used for normalization?

- A. Predict
- B. File
- C. Preprocess
- D. Scatter Plot

Answer: C

39. Missing values are handled using:

- A. Normalize
- B. Impute
- C. Predict
- D. Score

Answer: B

40. Mean replacement is a method of:

- A. Prediction
- B. Imputation
- C. Clustering
- D. Validation

Answer: B

41. Logistic Regression is mainly used for:

- A. Prediction
- B. Storage
- C. Data entry
- D. Visualization

Answer: A

42. K-Means is used for:

- A. Classification
- B. Clustering
- C. Storage
- D. Cleaning

Answer: B

43. Test and Score widget is used for:

- A. Validation
- B. Storage
- C. Collection
- D. Visualization

Answer: A

44. Cross-validation is a:

- A. Storage technique
- B. Validation method
- C. Cleaning method
- D. Collection method

Answer: B

45. A data stream is:

- A. Stored historical data only
- B. Continuous real-time flow of data
- C. Printed records
- D. Structured tables only

Answer: B

46. Mining data streams means:

- A. Deleting streams
- B. Extracting patterns from real-time data
- C. Storing data permanently
- D. Compressing files

Answer: B

47. Which can generate data streams?

- A. Sensors
- B. Satellite systems
- C. Web traffic
- D. All of these

Answer: D

48. Real-time analytics provides:

- A. Delayed insights
- B. Immediate insights
- C. No insights
- D. Historical records only

Answer: B

49. Future predictive analytics will be enhanced by:

- A. AI and Machine Learning
- B. Printers
- C. Scanners
- D. Keyboards

Answer: A

50. Quantum Computing is expected to:

- A. Reduce processing power
- B. Solve complex problems faster
- C. Eliminate Big Data
- D. Replace the internet

Answer: B

Answer Key:

1-B, 2-D, 3-B, 4-B, 5-C, 6-B, 7-C, 8-C, 9-B, 10-B,
11-C, 12-A, 13-B, 14-D, 15-C, 16-A, 17-B, 18-C, 19-B, 20-B,
21-C, 22-B, 23-C, 24-B, 25-C, 26-B, 27-C, 28-B, 29-C, 30-B,
31-A, 32-B, 33-C, 34-B, 35-A, 36-B, 37-B, 38-C, 39-B, 40-B,
41-A, 42-B, 43-A, 44-B, 45-B, 46-B, 47-D, 48-B, 49-A, 50-B.