UNIT-1

MCQs – Data Warehousing (DW) & Business Intelligence (BI)

1. Which of the following best describes the main purpose of Business Intelligence (BI)?
A) To store large volumes of data
B) To transform raw data into meaningful insights for decision-making
C) To perform transaction processing
D) To replace the need for data warehousing
Answer: B

2. What is the primary function of a Data Warehouse?
A) Real-time transaction processing
B) Long-term storage of operational data
C) Storing integrated, subject-oriented, time-variant, and non-volatile data for analysis
D) Replacing operational databases
Answer: C

3. Which of the following is NOT a defining feature of a data warehouse?
A) Subject-oriented
B) Time-variant
C) Volatile
D) Non-volatile
Answer: C

4. What does OLAP stand for?
A) Online Analytical Processing
B) Online Application Processing
C) Offline Analytical Processing
D) Onsite Analytical Performance
Answer: A

5. Which of the following best describes a Data Mart?
A) A subset of a data warehouse focused on a specific business area
B) A replacement for a data warehouse
C) A transactional database
D) An operational data store
Answer: A

6. Which of the following best describes Metadata in the context of data warehousing?
A) Data stored in the warehouse
B) Data about the data stored in the warehouse
C) Raw operational data
D) Real-time transactional records
Answer: B

7. Which of the following is an example of an OLTP system?
A) Banking transaction system
B) Sales trend analysis dashboard
C) Monthly business report generator
D) Data warehouse query tool
Answer: A

8. Which of the following is NOT part of the BI lifecycle?
A) Data collection
B) Data integration
C) Transaction logging
D) Data analysis
Answer: C

9. Which of the following best describes Dimensional Analysis in BI?
A) Analysis of transactions in a normalized database
B) Viewing data in multiple perspectives such as time, geography, and product
C) Only storing data in flat files
D) Maintaining indexes on all columns
Answer: B

10. In the context of BI, "Raw Data to Valuable Information"
means: A) Storing raw operational data directly in a warehouse
B) Processing, cleansing, and summarizing raw data for decision-making
C) Using raw logs without transformation
D) Ignoring redundant data
Answer: B

11. Which of the following is true about OLTP systems?
A) Optimized for read-heavy workloads
B) Contain historical data for analysis
C) Optimized for large numbers of small transactions
D) Designed for complex analytical queries
Answer: C

12. Which component of a data warehouse is responsible for extracting, transforming, and loading data?
A) OLAP cube
B) ETL process
C) Data mart
D) Metadata repository
Answer: B

13. Which is an advantage of using Data Warehousing for BI?
A) Supports operational transaction updates
B) Provides integrated and consistent data for analysis
C) Eliminates the need for any data governance
D) Requires no data transformation

Answer: B
14. Which of the following describes the relationship between BI and DW?
A) BI replaces DW completely
B) DW stores the data, BI analyzes it
C) BI is a form of DW
D) DW performs reporting while BI stores data
Answer: B

15. The time-variant characteristic of a data warehouse means:
A) Data is updated in real-time
B) Data is stored with historical context
C) Data is erased after a certain period
D) Data is encrypted before storage
Answer: B

16. Which of the following is a current trend in data warehousing?
A) Manual data entry
B) Cloud-based data warehouses
C) Only on-premises deployments
D) Avoiding big data integration
Answer: B

17. The "non-volatile" nature of a data warehouse means:
A) Data changes frequently
B) Data is not updated or deleted once entered
C) Data is volatile until analyzed
D) Data is stored in RAM
Answer: B

18. OLAP differs from OLTP because:
A) OLAP supports complex analytical queries, OLTP supports day-to-day transactions
B) OLAP stores operational data, OLTP stores historical data
C) OLAP is faster for transactions
D) OLAP uses normalized databases only
Answer: A

19. Which of the following is part of the basic elements of data warehousing?
A) ETL process
B) Metadata
C) Data marts
D) All of the above
Answer: D

20. In BI, "Reporting and Analyzing Data" typically involves:
A) Only storing data
B) Presenting processed data using dashboards, charts, and summaries
C) Raw data collection without processing
D) Performing transactions in real-time

Answer: B

5-Mark Descriptive Questions

1. Explain the relationship between Business Intelligence (BI) and Data Warehousing (DW) with examples.
2. Describe the defining features of a Data Warehouse and explain why each is important.
3. Compare and contrast OLAP and OLTP systems, including use cases for each. 4. Discuss the lifecycle of data from raw data to valuable information in the context of BI.
5. Explain the role of Metadata in a data warehouse and why it is critical for BI operations.
6. Describe the trends in modern data warehousing and their impact on BI systems.
7. Define dimensional analysis and explain its role in OLAP-based reporting. 8. What are Data Marts? Discuss their advantages and disadvantages compared to a full-scale data warehouse.
9. **1.** Which of the following best describes Business Intelligence (BI)?
   A) A set of tools for storing raw data
   B) Techniques for analyzing data to support decision making
   C) A process for designing operational systems
   D) A way to secure data
   Answer: B
10. **2.** In the BI process, raw data is transformed into:
    A) Metadata
    B) Valuable information
    C) Unstructured logs
    D) Temporary tables
    Answer: B
11. **3.** Which of the following is **NOT** a feature of a data warehouse? A) Subject-oriented
    B) Time-variant
    C) Non-volatile
    D) Real-time transaction processing
    Answer: D
12. **4.** The **lifecycle of data** typically begins with:
    A) Data analysis
    B) Data creation/collection
    C) Data visualization
    D) Data disposal
    Answer: B
13. **5.** A **Data Mart** can be defined as:
    A) A large enterprise-wide database
    B) A smaller, subject-specific subset of a data warehouse
    C) A temporary storage for real-time analytics
    D) A type of OLTP system
    Answer: B
14. **6.** Metadata in a data warehouse is used for:
    A) Encrypting the data

B) Storing raw transaction logs
C) Describing the structure, content, and rules of data
D) Running operational queries
Answer: C

15. **7.** Which is a **primary reason** for the need of data warehousing? A) To improve data redundancy
B) To integrate data from multiple sources for analysis
C) To replace transactional systems
D) To remove old data
Answer: B
16. **8.** In today's perspective, BI and DW are often integrated to: A) Replace all operational systems
B) Enable real-time transactional processing
C) Support strategic decision making
D) Reduce storage costs
Answer: C
17. **9.** OLAP is mainly used for:
A) High-speed transaction processing
B) Complex analytical queries and reporting
C) Data entry forms
D) Backup and recovery
Answer: B
18. **10.** Which statement is true about OLAP?
A) It handles thousands of short, simple transactions
B) It is optimized for large analytical queries
C) It is used only for operational systems
D) It stores only unstructured data
Answer: B
19. **11.** OLTP systems are primarily:
A) Analytical in nature
B) Designed for frequent updates and transactions
C) Time-variant
D) Non-volatile
Answer: B
20. **12.** The difference between OLAP and OLTP is that: A) OLAP focuses on operations, OLTP on analysis
B) OLAP focuses on analysis, OLTP on operations
C) Both are used for data entry
D) Both store historical data only
Answer: B
21. **13.** Which of the following is a dimensional model concept?
A) Entity-Relationship diagram
B) Fact tables and dimension tables
C) Normal forms
D) Trigger-based schema
Answer: B
22. **14.** Dimensional analysis is commonly used in:
A) Data entry systems

B) Data mining and OLAP queries
C) Transactional backups
D) Network monitoring
Answer: B

23. **15.** Which of the following is **NOT** a component of a typical data warehouse architecture?
A) Data sources
B) ETL processes
C) OLAP engine
D) Printer server
Answer: D
24. **16.** The term "non-volatile" in data warehouse context means:
A) Data can be modified at any time
B) Data is not erased after power loss
C) Data is stable and primarily read-only
D) Data is constantly updated in real-time
Answer: C
25. **17.** Which is a trend in modern data warehousing?
A) Manual data entry from paper forms
B) Cloud-based data warehouse solutions
C) Elimination of BI tools
D) Reducing the size of stored data to zero
Answer: B
26. **18.** In BI, dashboards are primarily used to:
A) Clean the data
B) Visually monitor KPIs and metrics
C) Replace metadata
D) Perform ETL operations
Answer: B
27. **19.** Which is an example of time-variant data?
A) Current bank account balance
B) Historical sales data for the last 5 years
C) Today's temperature
D) Live sensor feed
Answer: B
28. **20.** The relation between BI and DW can be summarized as:
A) BI is the foundation, DW is the application
B) DW stores and organizes data, BI analyzes it
C) DW performs analytics, BI stores data
D) Both are used for transactional processing
Answer: B

UNIT-2

MCQs (1 Mark Each)

Q1. Which of the following best describes *Data Mining*?

A) The process of cleaning data
B) The process of discovering patterns and knowledge from large amounts of data
C) The process of storing data in a database
D) The process of deleting irrelevant data
Answer: B

Q2. Which of the following is **NOT** a functionality of Data Mining?
A) Classification
B) Clustering
C) Summarization
D) Data Encoding
Answer: D

Q3. Which classification of Data Mining systems is based on the type of data handled?
A) Based on data models
B) Based on kinds of databases mined
C) Based on applications
D) Based on degree of user interaction
Answer: B

Q4. The *KDD Process* stands for:
A) Knowledge Discovery in Data
B) Key Data Discovery
C) Knowledge Definition and Design
D) Knowledge Data Derivation
Answer: A

Q5. Which of the following is a **Data Mining task primitive**?
A) Specifying the kind of knowledge to be mined
B) Data cleaning process
C) Database backup operation
D) Normalization of data
Answer: A

Q6. Integration of a Data Mining system with a Data Warehouse
improves: A) Data redundancy
B) Performance and scalability
C) Manual report writing
D) Disk fragmentation
Answer: B

Q7. Which is **NOT** an issue in Data Mining?
A) Mining methodology and user interaction
B) Performance issues
C) Security and social issues
D) Table formatting in spreadsheets
Answer: D

Q8. Which stage in the KDD process involves removing noise and inconsistencies?
A) Data Selection
B) Data Cleaning

C) Data Transformation
D) Pattern Evaluation
Answer: B

Q9. In classification of DM systems, systems that work without human intervention are classified as:
A) Interactive DM systems
B) Autonomous DM systems
C) Data-centric DM systems
D) Pattern-oriented DM systems
Answer: B

Q10. Which Data Mining functionality aims to assign data items to predefined categories?
A) Clustering
B) Classification
C) Association
D) Summarization
Answer: B

5-Mark Descriptive Questions

Q1. Explain the motivation for Data Mining. Discuss with examples why traditional data analysis techniques are not sufficient for large datasets.

Q2. Define Data Mining. Explain the major functionalities of Data Mining with examples.

Q3. Describe the classification of Data Mining systems based on the kinds of databases mined, kinds of knowledge mined, and techniques used.

Q4. What are Data Mining task primitives? Discuss their importance in specifying a Data Mining query.

Q5. Explain the KDD process in detail. How is Data Mining related to KDD?

Q1. Explain the *Motivation for Data Mining* with suitable examples.
Q2. Define Data Mining. Discuss its main functionalities in detail.
Q3. Describe the classification of Data Mining systems based on various criteria.
Q4. What are Data Mining task primitives? Explain each with an example.
Q5. Describe the integration of a Data Mining system with a Database or Data Warehouse. **Q6.** Explain the major issues in Data Mining.
Q7. Describe the steps of the Knowledge Discovery in Database (KDD) process.