Data Warehouse and Mining
Assignment -2

Multiple Choice Questions.

1. _____ is a subject-oriented, integrated, time-variant, nonvolatile collection of data in support of management decisions.
A. Data Mining.
B. Data Warehousing.
C. Web Mining.
D. Text Mining.

2. The data Warehouse is_____.
A. read only.
B. write only.
C. read write only.
D. none.

3. Expansion for DSS in DW is_____.
A. Decision Support system.
B. Decision Single System.
C. Data Storable System.
D. Data Support System.

4. The important aspect of the data warehouse environment is that data found within the data warehouse is_____.
A. subject-oriented.
B. time-variant.
C. integrated.
D. All of the above.

5. The time horizon in Data warehouse is usually _____.
A. 1-2 years.
B. 3-4years.
C. 5-6 years.
D. 5-10 years.

6. The data is stored, retrieved & updated in _____.

A. OLAP.

B. OLTP.

C. SMTP.

D. FTP.

7. _____describes the data contained in the data warehouse.

A. Relational data.

B. Operational data.

C. Metadata.

D. Informational data.

8. _____predicts future trends & behaviors, allowing business managers to make proactive,

knowledge-driven decisions.

A. Data warehouse.

B. Data mining.

C. Datamarts.

D. Metadata.

9. _____ is the heart of the warehouse.

A. Data mining database servers.

B. Data warehouse database servers.

C. Data mart database servers.

D. Relational data base servers.

10. _____ is the specialized data warehouse database.

A. Oracle.

B. DBZ.

C. Informix.

D. Redbrick.

11. _____defines the structure of the data held in operational databases and used by

operational applications.

A. User-level metadata.

B. Data warehouse metadata.

C. Operational metadata.

D. Data mining metadata.

12. _____ is held in the catalog of the warehouse database system.

A. Application level metadata.

B. Algorithmic level metadata.

C. Departmental level metadata.

D. Core warehouse metadata.

13. _____maps the core warehouse metadata to business concepts, familiar and useful to end users.

A. Application level metadata.

B. User level metadata.

C. Enduser level metadata.

D. Core level metadata.

14. _____consists of formal definitions, such as a COBOL layout or a database schema.

A. Classical metadata.

B. Transformation metadata.

C. Historical metadata.

D. Structural metadata.

15. _____consists of information in the enterprise that is not in classical form.

A. Mushy metadata.

B. Differential metadata.

C. Data warehouse.

D. Data mining.

16. . _____databases are owned by particular departments or business groups.

A. Informational.

B. Operational.

C. Both informational and operational.

D. Flat.

17. The star schema is composed of _____ fact table.
A. one.
B. two.
C. three.
D. four.
18. The time horizon in operational environment is _____.
A. 30-60 days.
B. 60-90 days.
C. 90-120 days.
D. 120-150 days.
19. The key used in operational environment may not have an element of_____.
A. time.
B. cost.
C. frequency.
D. quality.
20. Data can be updated in _____environment.
A. data warehouse.
B. data mining.
C. operational.
D. informational.
21. Record cannot be updated in _____.
A. OLTP
B. files
C. RDBMS
D. data warehouse
22. The source of all data warehouse data is the_____.
A. operational environment.
B. informal environment.
C. formal environment.
D. technology environment.
23. Data warehouse contains_____data that is never found

in the operational environment.

A. normalized.

B. informational.

C. summary.

D. denormalized.

24. The modern CASE tools belong to _____ category.

A. a. analysis.

B. b.Development

C. c.Coding

D. d.Delivery

25. Bill Inmon has estimated_____of the time required to build a data warehouse, is consumed in

the conversion process.

A. 10 percent.

B. 20 percent.

C. 40 percent

D. 80 percent.

26. Detail data in single fact table is otherwise known as_____.

A. monoatomic data.

B. diatomic data.

C. atomic data.

D. multiatomic data.

27. _____test is used in an online transactional processing environment.

A. MEGA.

B. MICRO.

C. MACRO.

D. ACID.

28. _____ is a good alternative to the star schema.

A. Star schema.

B. Snowflake schema.

C. Fact constellation.

D. Star-snowflake schema.

29. The biggest drawback of the level indicator in the classic star-schema is that it limits_____.

A. quantify.

B. qualify.

C. flexibility.

D. ability.

30. A data warehouse is _____.

A. updated by end users.

B. contains numerous naming conventions and formats

C. organized around important subject areas.

D. contains only current data.

31. An operational system is _____.

A. used to run the business in real time and is based on historical data.

B. used to run the business in real time and is based on current data.

C. used to support decision making and is based on current data.

D. used to support decision making and is based on historical data.

32. The generic two-level data warehouse architecture includes _____.

A. at least one data mart.

B. data that can extracted from numerous internal and external sources.

C. near real-time updates.

D. far real-time updates.

33. The active data warehouse architecture includes _____

A. at least one data mart.

B. data that can extracted from numerous internal and external sources.

C. near real-time updates.

D. all of the above.

34. Reconciled data is _____.

A. data stored in the various operational systems throughout the organization.

B. current data intended to be the single source for all decision support systems.

C. data stored in one operational system in the organization.

D. data that has been selected and formatted for end-user support applications.

35. Transient data is _____.

A. data in which changes to existing records cause the previous version of the records to be eliminated.

B. data in which changes to existing records do not cause the previous version of the records to be
eliminated.

C. data that are never altered or deleted once they have been added.

D. data that are never deleted once they have been added.

36. The extract process is _____.

A. capturing all of the data contained in various operational systems.

B. capturing a subset of the data contained in various operational systems.

C. capturing all of the data contained in various decision support systems.

D. capturing a subset of the data contained in various decision support systems.

37. Data scrubbing is _____.

A. a process to reject data from the data warehouse and to create the necessary indexes.

B. a process to load the data in the data warehouse and to create the necessary indexes.

C. a process to upgrade the quality of data after it is moved into a data warehouse.

D. a process to upgrade the quality of data before it is moved into a data warehouse

38. The load and index is _____.

A. a process to reject data from the data warehouse and to create the necessary indexes.

B. a process to load the data in the data warehouse and to create the necessary indexes.

C. a process to upgrade the quality of data after it is moved into a data warehouse.

D. a process to upgrade the quality of data before it is moved into a data warehouse.

39. Data transformation includes _____.

A. a process to change data from a detailed level to a summary level.

B. a process to change data from a summary level to a detailed level.

C. joining data from one source into various sources of data.

D. separating data from one source into various sources of data.

40. _____ is called a multifield transformation.

A. Converting data from one field into multiple fields.

B. Converting data from fields into field.

C. Converting data from double fields into multiple fields.

D. Converting data from one field to one field.

41. The type of relationship in star schema is _____.

A. many-to-many.

B. one-to-one.

C. one-to-many.

D. many-to-one.

42. Fact tables are _____.

A. completely demoralized.

B. partially demoralized.

C. completely normalized.

D. partially normalized.

43. _____ is the goal of data mining.

A. To explain some observed event or condition.

B. To confirm that data exists.

C. To analyze data for expected relationships.

D. To create a new data warehouse.

44. Business Intelligence and data warehousing is used for _____.

A. Forecasting.

B. Data Mining.

C. Analysis of large volumes of product sales data.

D. All of the above.

45. The data administration subsystem helps you perform all of the following, except_____.

A. backups and recovery.

B. query optimization.

C. security management.

D. create, change, and delete information.

46. The most common source of change data in refreshing a data warehouse is _____.

A. queryable change data.

B. cooperative change data.

C. logged change data.

D. snapshot change data.

47. _____ are responsible for running queries and reports against data warehouse tables.

A. Hardware.

B. Software.

C. End users.

D. Middle ware.

48. Query tool is meant for _____.

A. data acquisition.

B. information delivery.

C. information exchange.

D. communication.

49. Classification rules are extracted from _____.

A. root node.

B. decision tree.

C. siblings.

D. branches.

50. Dimensionality reduction reduces the data set size by removing _____.

A. relevant attributes.

B. irrelevant attributes.

C. derived attributes.

D. composite attributes.

51. _____ is a method of incremental conceptual clustering.

A. CORBA.

B. OLAP.

C. COBWEB.

D. STING.

52. Effect of one attribute value on a given class is independent of values of other attribute is called _____.

A. value independence.

B. class conditional independence.

C. conditional independence.

D. unconditional independence.

53. The main organizational justification for implementing a data warehouse is to provide _____.

A. cheaper ways of handling transportation.

B. decision support.

C. storing large volume of data.

D. access to data.

54. Multidimensional database is otherwise known as_____.

A. RDBMS

B. DBMS

C. EXTENDED RDBMS

D. EXTENDED DBMS

55. Data warehouse architecture is based on _____.
A. DBMS.
B. RDBMS.
C. Sybase.
D. SQL Server.
56. Source data from the warehouse comes from _____.
A. ODS.
B. TDS.
C. MDDB.
D. ORDBMS.
57. _____ is a data transformation process.
A. Comparison.
B. Projection.
C. Selection.
D. Filtering.
58. The technology area associated with CRM is _____.
A. specialization.
B. generalization.
C. personalization.
D. summarization.
59. SMP stands for _____.
A. Symmetric Multiprocessor.
B. Symmetric Multiprogramming.
C. Symmetric Metaprogramming.
D. Symmetric Microprogramming.
60. _____ are designed to overcome any limitations placed on the warehouse by the nature of the
relational data model.
A. Operational database.
B. Relational database.
C. Multidimensional database.
D. Data repository.

ANSWER: C