

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.

ANSWER: B

2. The data Warehouse is _____.

- A. read only.
- B. write only.
- C. read write only.
- D. none.

ANSWER: A

3. Expansion for DSS in DW is _____.

- A. Decision Support system.
- B. Decision Single System.
- C. Data Storable System.
- D. Data Support System.

ANSWER: A

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.

ANSWER: D

5. The time horizon in Data warehouse is usually _____.

- A. 1-2 years.
- B. 3-4years.

- C. 5-6 years.
- D. 5-10 years.

ANSWER: D

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

- A. OLAP.
- B. OLTP.
- C. SMTP.
- D. FTP.

ANSWER: B

7. _____describes the data contained in the data warehouse.

- A. Relational data.
- B. Operational data.
- C. Metadata.
- D. Informational data.

ANSWER: C

8. _____predicts future trends & behaviors, allowing business managers to make proactive, knowledge-driven decisions.

- A. Data warehouse.
- B. Data mining.
- C. Datamarts.
- D. Metadata.

ANSWER: B

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.

ANSWER: B

10. _____ is the specialized data warehouse database.

- A. Oracle.
- B. DBZ.
- C. Informix.
- D. Redbrick.

ANSWER: D

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.

ANSWER: C

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.

ANSWER: B

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.

ANSWER: A

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.

ANSWER: A

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.

ANSWER: A

16. _____ databases are owned by particular departments or business groups.

- A. Informational.
- B. Operational.
- C. Both informational and operational.
- D. Flat.

ANSWER: B

17. The star schema is composed of _____ fact table.

- A. one.
- B. two.
- C. three.
- D. four.

ANSWER: A

18. The time horizon in operational environment is _____.

- A. 30-60 days.
- B. 60-90 days.
- C. 90-120 days.
- D. 120-150 days.

ANSWER: B

19. The key used in operational environment may not have an element of _____.

- A. time.
- B. cost.
- C. frequency.
- D. quality.

ANSWER: A

20. Data can be updated in _____ environment.

- A. data warehouse.
- B. data mining.
- C. operational.
- D. informational.

ANSWER: C

21. Record cannot be updated in _____.

- A. OLTP
- B. files
- C. RDBMS
- D. data warehouse

ANSWER: D

22. The source of all data warehouse data is the_____.

- A. operational environment.
- B. informal environment.
- C. formal environment.
- D. technology environment.

ANSWER: A

23. Data warehouse contains_____data that is never found in the operational environment.

- A. normalized.
- B. informational.
- C. summary.
- D. denormalized.

ANSWER: C

24. The modern CASE tools belong to _____ category.

- A. a. analysis.
- B. b.Development
- C. c.Coding
- D. d.Delivery

ANSWER: A

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.

ANSWER: D

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

- A. monoatomic data.
- B. diatomic data.

- C. atomic data.
- D. multiatomic data.

ANSWER: C

27. _____ test is used in an online transactional processing environment.

- A. MEGA.
- B. MICRO.
- C. MACRO.
- D. ACID.

ANSWER: D

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

- A. Star schema.
- B. Snowflake schema.
- C. Fact constellation.
- D. Star-snowflake schema.

ANSWER: C

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.

ANSWER: C

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.

ANSWER: C

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.

ANSWER: B

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.

ANSWER: C

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.

ANSWER: D

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.

ANSWER: B

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.

ANSWER: A

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.

ANSWER: B

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.

ANSWER: D

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.

ANSWER: B

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.

ANSWER: A

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.

ANSWER: A

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.

ANSWER: C

42. Fact tables are _____.

- A. completely demoralized.
- B. partially demoralized.
- C. completely normalized.
- D. partially normalized.

ANSWER: C

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.

ANSWER: A

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.

ANSWER: D

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.

ANSWER: D

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.

ANSWER: A

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

- A. Hardware.

- B. Software.
- C. End users.
- D. Middle ware.

ANSWER: C

48. Query tool is meant for _____.

- A. data acquisition.
- B. information delivery.
- C. information exchange.
- D. communication.

ANSWER: A

49. Classification rules are extracted from _____.

- A. root node.
- B. decision tree.
- C. siblings.
- D. branches.

ANSWER: B

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

- A. relevant attributes.
- B. irrelevant attributes.
- C. derived attributes.
- D. composite attributes.

ANSWER: B

51. _____ is a method of incremental conceptual clustering.

- A. CORBA.
- B. OLAP.
- C. COBWEB.
- D. STING.

ANSWER: C

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.

ANSWER: A

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.

ANSWER: C

54. Multidimensional database is otherwise known as_____.

- A. RDBMS
- B. DBMS
- C. EXTENDED RDBMS
- D. EXTENDED DBMS

ANSWER: B

55. Data warehouse architecture is based on _____.

- A. DBMS.
- B. RDBMS.
- C. Sybase.
- D. SQL Server.

ANSWER: B

56. Source data from the warehouse comes from _____.

- A. ODS.
- B. TDS.
- C. MDDB.
- D. ORDBMS.

ANSWER: A

57. _____ is a data transformation process.

- A. Comparison.
- B. Projection.
- C. Selection.
- D. Filtering.

ANSWER: D

58. The technology area associated with CRM is _____.

- A. specialization.
- B. generalization.
- C. personalization.
- D. summarization.

ANSWER: C

59. SMP stands for _____.

- A. Symmetric Multiprocessor.
- B. Symmetric Multiprogramming.
- C. Symmetric Metaprogramming.
- D. Symmetric Microprogramming.

ANSWER: A

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

61. _____ 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

62. MDDB stands for _____.

- A. multiple data doubling.
- B. multidimensional databases.
- C. multiple double dimension.
- D. multi-dimension doubling.

ANSWER: B

63. _____ is data about data.

- A. Metadata.
- B. Microdata.

- C. Minidata.
- D. Multidata.

ANSWER: A

64. _____ is an important functional component of the metadata.

- A. Digital directory.
- B. Repository.
- C. Information directory.
- D. Data dictionary.

ANSWER: C

65. EIS stands for _____.

- A. Extended interface system.
- B. Executive interface system.
- C. Executive information system.
- D. Extendable information system.

ANSWER: C

66. _____ is data collected from natural systems.

- A. MRI scan.
- B. ODS data.
- C. Statistical data.
- D. Historical data.

ANSWER: A

67. _____ is an example of application development environments.

- A. Visual Basic.
- B. Oracle.
- C. Sybase.
- D. SQL Server.

ANSWER: A

68. The term that is not associated with data cleaning process is _____.

- A. domain consistency.
- B. deduplication.
- C. disambiguation.
- D. segmentation.

ANSWER: D

69. _____ are some popular OLAP tools.

- A. Metacube, Informix.
- B. Oracle Express, Essbase.
- C. HOLAP.
- D. MOLAP.

ANSWER: A

70. Capability of data mining is to build _____ models.

- A. retrospective.
- B. interrogative.
- C. predictive.
- D. imperative.

ANSWER: C

71. _____ is a process of determining the preference of customer's majority.

- A. Association.
- B. Preferencing.
- C. Segmentation.
- D. Classification.

ANSWER: B

72. Strategic value of data mining is _____.

- A. cost-sensitive.
- B. work-sensitive.
- C. time-sensitive.
- D. technical-sensitive.

ANSWER: C

73. _____ proposed the approach for data integration issues.

- A. Ralph Campbell.
- B. Ralph Kimball.
- C. John Raphlin.
- D. James Gosling.

ANSWER: B

74. The terms equality and roll up are associated with _____.

- A. OLAP.

- B. visualization.
- C. data mart.
- D. decision tree.

ANSWER: C

75. Exceptional reporting in data warehousing is otherwise called as _____.

- A. exception.
- B. alerts.
- C. errors.
- D. bugs.

ANSWER: B

76. _____ is a metadata repository.

- A. Prism solution directory manager.
- B. CORBA.
- C. STUNT.
- D. COBWEB.

ANSWER: A

77. _____ is an expensive process in building an expert system.

- A. Analysis.
- B. Study.
- C. Design.
- D. Information collection.

ANSWER: D

78. The full form of KDD is _____.

- A. Knowledge database.
- B. Knowledge discovery in database.
- C. Knowledge data house.
- D. Knowledge data definition.

ANSWER: B

79. The first International conference on KDD was held in the year _____.

- A. 1996.
- B. 1997.
- C. 1995.
- D. 1994.

ANSWER: C

80. Removing duplicate records is a process called _____.

- A. recovery.
- B. data cleaning.
- C. data cleansing.
- D. data pruning.

ANSWER: B

81. _____ contains information that gives users an easy-to-understand perspective of the information stored in the data warehouse.

- A. Business metadata.
- B. Technical metadata.
- C. Operational metadata.
- D. Financial metadata.

ANSWER: A

82. _____ helps to integrate, maintain and view the contents of the data warehousing system.

- A. Business directory.
- B. Information directory.
- C. Data dictionary.
- D. Database.

ANSWER: B

83. Discovery of cross-sales opportunities is called _____.

- A. segmentation.
- B. visualization.
- C. correction.
- D. association.

ANSWER: D

84. Data marts that incorporate data mining tools to extract sets of data are called _____.

- A. independent data mart.
- B. dependent data marts.
- C. intra-entry data mart.
- D. inter-entry data mart.

ANSWER: B

85. _____ can generate programs itself, enabling it to carry out new tasks.

- A. Automated system.
- B. Decision making system.
- C. Self-learning system.
- D. Productivity system.

ANSWER: D

86. The power of self-learning system lies in _____.

- A. cost.
- B. speed.
- C. accuracy.
- D. simplicity.

ANSWER: C

87. Building the informational database is done with the help of _____.

- A. transformation or propagation tools.
- B. transformation tools only.
- C. propagation tools only.
- D. extraction tools.

ANSWER: A

88. How many components are there in a data warehouse?

- A. two.
- B. three.
- C. four.
- D. five.

ANSWER: D

89. Which of the following is not a component of a data warehouse?

- A. Metadata.
- B. Current detail data.
- C. Lightly summarized data.
- D. Component Key.

ANSWER: D

90. _____ is data that is distilled from the low level of detail found at the current detailed level.

- A. Highly summarized data.
- B. Lightly summarized data.
- C. Metadata.
- D. Older detail data.

ANSWER: B

91. Highly summarized data is _____.

- A. compact and easily accessible.
- B. compact and expensive.
- C. compact and hardly accessible.
- D. compact.

ANSWER: A

92. A directory to help the DSS analyst locate the contents of the data warehouse is seen in _____.

- A. Current detail data.
- B. Lightly summarized data.
- C. Metadata.
- D. Older detail data.

ANSWER: C

93. Metadata contains atleast _____.

- A. the structure of the data.
- B. the algorithms used for summarization.
- C. the mapping from the operational environment to the data warehouse.
- D. all of the above.

ANSWER: D

94. Which of the following is not a old detail storage medium?

- A. Phot Optical Storage.
- B. RAID.
- C. Microfinche.
- D. Pen drive.

ANSWER: D

95. The data from the operational environment enter _____ of data warehouse.

- A. Current detail data.
- B. Older detail data.
- C. Lightly summarized data.
- D. Highly summarized data.

ANSWER: A

96. The data in current detail level resides till _____ event occurs.

- A. purge.

- B. summarization.
- C. archived.
- D. all of the above.

ANSWER: D

97. The dimension tables describe the _____.

- A. entities.
- B. facts.
- C. keys.
- D. units of measures.

ANSWER: B

98. The granularity of the fact is the _____ of detail at which it is recorded.

- A. transformation.
- B. summarization.
- C. level.
- D. transformation and summarization.

ANSWER: C

99. Which of the following is not a primary grain in analytical modeling?

- A. Transaction.
- B. Periodic snapshot.
- C. Accumulating snapshot.
- D. All of the above.

ANSWER: B

100. Granularity is determined by _____.

- A. number of parts to a key.
- B. granularity of those parts.
- C. both A and B.
- D. none of the above.

ANSWER: C

101. _____ of data means that the attributes within a given entity are fully dependent on the entire primary key of the entity.

- A. Additivity.
- B. Granularity.
- C. Functional dependency.
- D. Dimensionality.

ANSWER: C

102. A fact is said to be fully additive if _____.

- A. it is additive over every dimension of its dimensionality.
- B. additive over atleast one but not all of the dimensions.
- C. not additive over any dimension.
- D. None of the above.

ANSWER: A

103. A fact is said to be partially additive if _____.

- A. it is additive over every dimension of its dimensionality.
- B. additive over atleast one but not all of the dimensions.
- C. not additive over any dimension.
- D. None of the above.

ANSWER: B

104. A fact is said to be non-additive if _____.

- A. it is additive over every dimension of its dimensionality.
- B. additive over atleast one but not all of the dimensions.
- C. not additive over any dimension.
- D. None of the above.

ANSWER: C

105. Non-additive measures can often combined with additive measures to create new _____.

- A. additive measures.
- B. non-additive measures.
- C. partially additive.
- D. All of the above.

ANSWER: A

106. A fact representing cumulative sales units over a day at a store for a product is a _____.

- A. additive fact.
- B. fully additive fact.
- C. partially additive fact.
- D. non-additive fact.

ANSWER: B

107. _____ of data means that the attributes within a given entity are fully dependent on the entire primary key of the entity.

- A. Additivity.
- B. Granularity.
- C. Functional Dependency.
- D. Dependency.

ANSWER: C

108. Which of the following is the other name of Data mining?

- A. Exploratory data analysis.
- B. Data driven discovery.
- C. Deductive learning.
- D. All of the above.

ANSWER: D

109. Which of the following is a predictive model?

- A. Clustering.
- B. Regression.
- C. Summarization.
- D. Association rules.

ANSWER: B

110. Which of the following is a descriptive model?

- A. Classification.
- B. Regression.
- C. Sequence discovery.
- D. Association rules.

ANSWER: C

111. A _____ model identifies patterns or relationships.

- A. Descriptive.
- B. Predictive.
- C. Regression.
- D. Time series analysis.

ANSWER: A

112. A predictive model makes use of _____.

- A. current data.
- B. historical data.
- C. both current and historical data.

D. assumptions.

ANSWER: B

113. _____ maps data into predefined groups.

A. Regression.

B. Time series analysis

C. Prediction.

D. Classification.

ANSWER: D

114. _____ is used to map a data item to a real valued prediction variable.

A. Regression.

B. Time series analysis.

C. Prediction.

D. Classification.

ANSWER: B

115. In _____, the value of an attribute is examined as it varies over time.

A. Regression.

B. Time series analysis.

C. Sequence discovery.

D. Prediction.

ANSWER: B

116. In _____ the groups are not predefined.

A. Association rules.

B. Summarization.

C. Clustering.

D. Prediction.

ANSWER: C

117. Link Analysis is otherwise called as _____.

A. affinity analysis.

B. association rules.

C. both A & B.

D. Prediction.

ANSWER: C

118. _____ is a the input to KDD.

A. Data.

- B. Information.
- C. Query.
- D. Process.

ANSWER: A

119. The output of KDD is _____.

- A. Data.
- B. Information.
- C. Query.
- D. Useful information.

ANSWER: D

120. The KDD process consists of _____ steps.

- A. three.
- B. four.
- C. five.
- D. six.

ANSWER: C

121. Treating incorrect or missing data is called as _____.

- A. selection.
- B. preprocessing.
- C. transformation.
- D. interpretation.

ANSWER: B

122. Converting data from different sources into a common format for processing is called as _____.

- A. selection.
- B. preprocessing.
- C. transformation.
- D. interpretation.

ANSWER: C

123. Various visualization techniques are used in _____ step of KDD.

- A. selection.
- B. transformaion.
- C. data mining.
- D. interpretation.

ANSWER: D

124. Extreme values that occur infrequently are called as _____.

- A. outliers.
- B. rare values.
- C. dimensionality reduction.
- D. All of the above.

ANSWER: A

125. Box plot and scatter diagram techniques are _____.

- A. Graphical.
- B. Geometric.
- C. Icon-based.
- D. Pixel-based.

ANSWER: B

126. _____ is used to proceed from very specific knowledge to more general information.

- A. Induction.
- B. Compression.
- C. Approximation.
- D. Substitution.

ANSWER: A

127. Describing some characteristics of a set of data by a general model is viewed as

- _____
- A. Induction.
 - B. Compression.
 - C. Approximation.
 - D. Summarization.

ANSWER: B

128. _____ helps to uncover hidden information about the data.

- A. Induction.
- B. Compression.
- C. Approximation.
- D. Summarization.

ANSWER: C

129. _____ are needed to identify training data and desired results.

- A. Programmers.
- B. Designers.
- C. Users.
- D. Administrators.

ANSWER: C

130. Overfitting occurs when a model _____.

- A. does fit in future states.
- B. does not fit in future states.
- C. does fit in current state.
- D. does not fit in current state.

ANSWER: B

131. The problem of dimensionality curse involves _____.

- A. the use of some attributes may interfere with the correct completion of a data mining task.
- B. the use of some attributes may simply increase the overall complexity.
- C. some may decrease the efficiency of the algorithm.
- D. All of the above.

ANSWER: D

132. Incorrect or invalid data is known as _____.

- A. changing data.
- B. noisy data.
- C. outliers.
- D. missing data.

ANSWER: B

133. ROI is an acronym of _____.

- A. Return on Investment.
- B. Return on Information.
- C. Repetition of Information.
- D. Runtime of Instruction

ANSWER: A

134. The _____ of data could result in the disclosure of information that is deemed to be confidential.

- A. authorized use.

- B. unauthorized use.
- C. authenticated use.
- D. unauthenticated use.

ANSWER: B

135. _____ data are noisy and have many missing attribute values.

- A. Preprocessed.
- B. Cleaned.
- C. Real-world.
- D. Transformed.

ANSWER: C

136. The rise of DBMS occurred in early _____.

- A. 1950's.
- B. 1960's
- C. 1970's
- D. 1980's.

ANSWER: C

137. SQL stand for _____.

- A. Standard Query Language.
- B. Structured Query Language.
- C. Standard Quick List.
- D. Structured Query list.

ANSWER: B

138. Which of the following is not a data mining metric?

- A. Space complexity.
- B. Time complexity.
- C. ROI.
- D. All of the above.

ANSWER: D

139. Reducing the number of attributes to solve the high dimensionality problem is called as _____.

- A. dimensionality curse.
- B. dimensionality reduction.
- C. cleaning.
- D. Overfitting.

ANSWER: B

140. Data that are not of interest to the data mining task is called as _____.

- A. missing data.
- B. changing data.
- C. irrelevant data.
- D. noisy data.

ANSWER: C

141. _____ are effective tools to attack the scalability problem.

- A. Sampling.
- B. Parallelization
- C. Both A & B.
- D. None of the above.

ANSWER: C

142. Market-basket problem was formulated by _____.

- A. Agrawal et al.
- B. Steve et al.
- C. Toda et al.
- D. Simon et al.

ANSWER: A

143. Data mining helps in _____.

- A. inventory management.
- B. sales promotion strategies.
- C. marketing strategies.
- D. All of the above.

ANSWER: D

144. The proportion of transaction supporting X in T is called _____.

- A. confidence.
- B. support.
- C. support count.
- D. All of the above.

ANSWER: B

145. The absolute number of transactions supporting X in T is called _____.

- A. confidence.
- B. support.

- C. support count.
- D. None of the above.

ANSWER: C

146. The value that says that transactions in D that support X also support Y is called _____.

- A. confidence.
- B. support.
- C. support count.
- D. None of the above.

ANSWER: A

147. If T consist of 500000 transactions, 20000 transaction contain bread, 30000 transaction contain jam,10000 transaction contain both bread and jam. Then the support of bread and jam is _____.

- A. 2%
- B. 20%
- C. 3%
- D. 30%

ANSWER: A

148. 7 If T consist of 500000 transactions, 20000 transaction contain bread, 30000 transaction contain jam,10000 transaction contain both bread and jam. Then the confidence of buying bread with jam is _____.

- A. 33.33%
- B. 66.66%
- C. 45%
- D. 50%

ANSWER: D

149. The left hand side of an association rule is called _____.

- A. consequent.
- B. onset.
- C. antecedent.
- D. precedent.

ANSWER: C

150. The right hand side of an association rule is called _____.

- A. consequent.
- B. onset.
- C. antecedent.

D. precedent.

ANSWER: A

151. The data Warehouse is_____.

- A. read only.
- B. write only.
- C. read write only.
- D. none.

ANSWER: A

152. Expansion for DSS in DW is_____.

- A. Decision Support system.
- B. Decision Single System.
- C. Data Storable System.
- D. Data Support System.

ANSWER: A

153. 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.

ANSWER: D

154. The time horizon in Data warehouse is usually _____.

- A. 1-2 years.
- B. 3-4years.
- C. 5-6 years.
- D. 5-10 years.

ANSWER: D

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

- A. OLAP.
- B. OLTP.
- C. SMTP.
- D. FTP.

ANSWER: B

156. _____describes the data contained in the data warehouse.

- A. Relational data.
- B. Operational data.
- C. Metadata.
- D. Informational data.

ANSWER: C

157. _____ predicts future trends & behaviors, allowing business managers to make proactive, knowledge-driven decisions.

- A. Data warehouse.
- B. Data mining.
- C. Datamarts.
- D. Metadata.

ANSWER: B

158. _____ 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.

ANSWER: B

159. Which of the following is not a component of a data warehouse?

- A. Metadata.
- B. Current detail data.
- C. Lightly summarized data.
- D. Component Key.

ANSWER: D

160. _____ is the specialized data warehouse database.

- A. Oracle.
- B. DBZ.
- C. Informix.
- D. Redbrick.

ANSWER: D

161. _____ 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.

ANSWER: C

162. _____ 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.

ANSWER: B

163. _____ 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.

ANSWER: A

164. _____ 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.

ANSWER: A

165. _____ consists of information in the enterprise that is not in classical form.

- A. Mushy metadata.
- B. Differential metadata.
- C. Data warehouse.
- D. Data mining.

ANSWER: A

166. . _____ databases are owned by particular departments or business groups.

- A. Informational.
- B. Operational.
- C. Both informational and operational.
- D. Flat.

ANSWER: B

167. The star schema is composed of _____ fact table.

- A. one.
- B. two.
- C. three.
- D. four.

ANSWER: A

168. The time horizon in operational environment is _____.

- A. 30-60 days.
- B. 60-90 days.
- C. 90-120 days.
- D. 120-150 days.

ANSWER: B

169. The key used in operational environment may not have an element of_____.

- A. time.
- B. cost.
- C. frequency.
- D. quality.

ANSWER: A

170. Data can be updated in _____environment.

- A. data warehouse.
- B. data mining.
- C. operational.
- D. informational.

ANSWER: C

171. Record cannot be updated in _____.

- A. OLTP
- B. files
- C. RDBMS
- D. data warehouse

ANSWER: D

172. The source of all data warehouse data is the_____.

- A. operational environment.
- B. informal environment.
- C. formal environment.
- D. technology environment.

ANSWER: A

173. Data warehouse contains_____data that is never found in the operational environment.

- A. normalized.
- B. informational.
- C. summary.
- D. denormalized.

ANSWER: C

174. The modern CASE tools belong to _____ category.

- A. a. analysis.
- B. b.Development
- C. c.Coding
- D. d.Delivery

ANSWER: A

175. 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.

ANSWER: D

176. Detail data in single fact table is otherwise known as _____.

- A. monoatomic data.
- B. diatomic data.
- C. atomic data.
- D. multiatomic data.

ANSWER: C

177. _____ test is used in an online transactional processing environment.

- A. MEGA.
- B. MICRO.
- C. MACRO.
- D. ACID.

ANSWER: D

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

- A. Star schema.
- B. Snowflake schema.
- C. Fact constellation.
- D. Star-snowflake schema.

ANSWER: C

179. The biggest drawback of the level indicator in the classic star-schema is that it limits _____.

- A. quantify.
- B. qualify.
- C. flexibility.
- D. ability.

ANSWER: C

180. 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.

ANSWER: C

181. 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.

ANSWER: B

182. 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.

ANSWER: C

183. 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.

ANSWER: D

184. 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.

ANSWER: B

185. 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.

ANSWER: A

186. 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.

ANSWER: B

187. Research on mining multi-types of data is termed as _____ data.

- A. Graphics.
- B. Multimedia.
- C. Meta.

D. Digital.

ANSWER: B

188. _____ are some popular OLAP tools.

- A. Metacube, Informix.
- B. Oracle Express, Essbase.
- C. HOLAP.
- D. MOLAP.

ANSWER: A

189. 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.

ANSWER: A

190. _____ 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

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

- A. many-to-many.
- B. one-to-one.
- C. one-to-many.
- D. many-to-one.

ANSWER: C

192. Data marts that incorporate data mining tools to extract sets of data are called _____.

- A. independent data mart.
- B. dependent data marts.
- C. intra-entry data mart.
- D. inter-entry data mart.

ANSWER: B

193. _____ 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.

ANSWER: A

194. 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.

ANSWER: D

195. 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.

ANSWER: D

196. 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.

ANSWER: A

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

- A. Hardware.
- B. Software.
- C. End users.
- D. Middle ware.

ANSWER: C

198. Strategic value of data mining is _____.

- A. cost-sensitive.
- B. work-sensitive.
- C. time-sensitive.
- D. technical-sensitive.

ANSWER: C

199. Classification rules are extracted from _____.

- A. root node.
- B. decision tree.
- C. siblings.
- D. branches.

ANSWER: B

200. Capability of data mining is to build _____ models.

- A. retrospective.
- B. interrogative.
- C. predictive.

D. imperative.

ANSWER: C

201. _____ 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.

ANS:B

202. Data cleaning is

[A] Large collection of data mostly stored in a computer system.

[B] The removal of noise errors and incorrect input from a database.

[C] The systematic description of the syntactic structure of a specific database. It describes the structure of the attributes the tables and foreign key relationships.

[D] None of these.

ANS:B

203Expansion for DSS in DW is_____.

[A] Decision Support system.

[B] Decision Single System.

[C] Data Storable System.

[D] Data Support System.

ANS : A

204.The time horizon in Data warehouse is usually _____.

[A]1-2 years.

[B] 3-4years.

[C] 5-6 years.

[D] 5-10 years.

ANS : D

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

[A] Mushy metadata.

[B] Differential metadata.

[C] Data warehouse.

[D] Data mining.

ANS : A

206. The star schema is composed of _____ fact table.

- [A] one.
- [B] two.
- [C] three.
- [D] four.

ANS : A

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

- [A] Star schema.
- [B] Snowflake schema.
- [C] Fact constellation.
- [D] Star-snowflake schema.

ANS : C

208. 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.

ANS : D

209. _____ 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.

ANS : A

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

- [A] many-to-many.
- [B] one-to-one.
- [C] one-to-many.
- [D] many-to-one.

ANS : C

211. Fact tables are which of the following?

- [A] Completely demoralized.
- [B] Partially demoralized.
- [C] Completely normalized.
- [D] Partially normalized.

ANS : C

212. Classification rules are extracted from _____.

- [A] root node.
- [B] decision tree.

- [C]siblings.
- [D] branches.

ANS:B

213. Multidimensional database is otherwise known as_____.

- [A] RDBMS.
- [B] DBMS.
- [C] EXTENDED RDBMS.
- [D]EXTENDED DBMS.

ANS:B

214. Data warehouse architecture is based on _____.

- [A] DBMS.
- [B] RDBMS.
- [C]Sybase.
- [D] SQL Server.

ANS:B

215. SMP stands for _____.

- [A] Symmetric Multiprocessor.
- [B]Symmetric Multiprogramming.
- [C] Symmetric Metaprogramming.
- [D]Symmetric Microprogramming.

ANS : A

216.Which type of schema has dimensions with extended dimensions.

- [A] Snow Flake
- [B]Star
- [C]extended star
- [D]Hybrid star

ANS : A

217.Which OLAP category offers best performance.

- [A]ROLAP
- [B]DOLAP
- [C]MOLAP
- [D]AOLAP

ANS : C

218._____ is data about data.

- [A] Meta Data
- [B] Micro Data
- [C] Mini Data
- [D]Multi Data

ANS : A

219.Data Warehouse Architecture is base on _____

- [A]DBMS
- [B]RDBMS
- [C]Sybase
- [D]SQL Server

ANS:B

220. MDDDB stands for _____.

- [A] multiple data doubling.
- [B] multidimensional databases.
- [C] multiple double dimension.
- [D] multi-dimension doubling.

ANS:B

221. _____ is an example of application development environments.

- [A] Visual Basic.
- [B] Oracle.
- [C]Sybase.
- [D] SQL Server.

ANS : A

222. Discovery of cross-sales opportunities is called _____.

- [A] segmentation.
- [B] visualization.
- [C] correction.
- [D] association.

ANS : D

223. How many components are there in a data warehouse?

- [A] two.
- [B] three.
- [C] four.
- [D] five

ANS : D

224.The dimension tables describe the _____.

- [A] entities.
- [B] facts.
- [C] keys.
- [D] units of measures.

ANS:B

225.Granularity is determined by _____.

- [A] number of parts to a key.
- [B] granularity of those parts.

- [C] both A and B.
- [D] none of the above.

ANS : C

226.A _____ model identifies patterns or relationships.

- [A] Descriptive.
- [B] Predictive.
- [C] Regression.
- [D] Time series analysis.

ANS : A

227. In _____, the value of an attribute is examined as it varies over time.

- [A] Regression.
- [B] Time series analysis.
- [C] Sequence discovery.
- [D] Prediction.

ANS:B

228. _____ is a the input to KDD.

- [A]Data.
- [B] Information.
- [C] Query.
- [D] Process.

ANS : A

229.Which Approach to build a data warehouse involves individual data marts to be integrated into the enterprise data warehouse.

- [A] Bottom Up
- [B] Top Up
- [C] Vertical
- [D] Horizontal

ANS : A

230.Which data mining is appropriate for undirected data mining.

- [A] Association Rules
- [B] Statistical.
- [C] Decision Tree.
- [D] Neural Network

ANS : A

231. _____ is a data transformation process.

- [A]Comparison
- [B]Projection

- [C] Selection
- [D] Filtering

ANS : D

232. The time horizon in operational environment is _____.

- [A] 30-60 days.
- [B] 60-90 days.
- [C] 90-120 days.
- [D] 120-150 days.

ANS: B

233. The KDD Process Consists of _____ Step.

- [A] five.
- [B] two.
- [C] three.
- [D] four.

ANS : A

234. _____ 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.

ANS: B

235. _____ 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.

ANS : A

236. In web mining, _____ is used to find natural groupings of users, pages, etc.

- [A] clustering.
- [B] associations.
- [C] sequential analysis.
- [D] classification.

ANS : A

237. _____ is the specialized data warehouse database.

- [A] Oracle.
- [B] DBZ.
- [C] DBZ.
- [D] Redbrick.

ANS : D

238. Record cannot be updated in_____.

- [A] OLTP.
- [B] Files.
- [C] RDBMS.
- [D] Data Warehouse.

ANS : D

239. The actual amount of reduction at each learning step may be guided by _____.

- [A] Learningcost.
- [B] LearninLevel.
- [C] Learning Rate.
- [D] LearningTime.

ANS : C

240. In _____ algorithm each cluster is represented by the center of gravity of the cluster.

- [A] k-medoid.
- [B] k-means.
- [C] STIRR.
- [D] ROCK.

ANS:B

241. Which of the following is a clustering algorithm?

- [A] A priori.
- [B] CLARA.
- [C]Pincer-Search.
- [D] FP-growth.

ANS:B

242. Rule based classification algorithms generate _____ rule to perform the classification.

- [A] if-then.
- [B] while..
- [C]do while.
- [D]switch.

ANS : A

243. GA stands for _____.

- [A] Genetic algorithm
- [B] Gene algorithm.
- [C] General algorithm.
- [D] Geo algorithm.

ANS : A

244. Research on mining multi-types of data is termed as _____ data.

- [A] graphics.
- [B] multimedia.
- [C] meta.
- [D] digital.

ANS:B

245. Which of the following is/are the Data mining tasks?

- [A] Regression
- [B] Classification.
- [C] Clustering.
- [D] All.

ANS : D

246. .Data mining is an integral part of_____

- [A]SE
- [B]DBMS
- [C]KDD
- [D]OS

ANS : C

247. _____Maps the core warehouse metadata to business concepts, familiar and useful to end user.

- [A] Application level metadata.
- [B] User level metadata
- [C] Enduser level metadata
- [D] Core level metadata

ANS : A

248. A Data Mining is

- [A] The Actual Discovery phase of a knowledge discovery process.
- [B] The Stage of selecting the right data for a KDD process.
- [C] A Subject-Oriented integrated time variant non volatile collection of data in support of management.
- [D] None of All.

ANS : A

249. The Terms equality and roll up are associated with_____.

- [A] OLAP
- [B] Visualization
- [C] Data Mart
- [D] Decision Tree

ANS : C

250. The Form of data, having an associated time interval during which it is valid is known as

- [A] Temporal Data
 - [B] Snapshot Data
 - [C] Chunk Data
 - [D] Point in time data
- ANS : A

******* BEST OF LUCK *******