

# CS614-Data Warehouse MID TERM MCQS

Prepared by: JUNAID MALIK

## AL-JUNAID TECH INSTITUTE



[www.vulmshelp.com](http://www.vulmshelp.com)



### Language Courses Training Available

I'm providing paid courses in different languages within 3 Months, Certificate will be awarded after completion.

- HTML
- CSS
- JAVASCRIPT
- BOOTSTRAPS
- JQUERY
- PHP MYSQL
- NODES.JS
- REACT JS

### LMS Handling Services

LMS Activities Paid Task

Assignments 95% Results

Quizes 95% Results

GDB 95% Results

For CS619 Project Feel Free To Contact With Me

Ph# 0304-1659294  
Email: junaidfzal08@gmail.com

# AL-JUNAID TECH INSTITUTE

1. Online Extraction is a kind of-----data extraction.
  - Logical
  - Dimensional
  - **Physical page 132**
  - Multi valued
2. The ----- saw the advent of disk storage, or DASD( direct Access Storage Device) :
  - 1960s
  - **1970s page 13**
  - 1950s
  - 1990s
3. In context of data warehouse, normally it becomes difficult to extract data from different sources because these sources are normally.
  - **Heterogeneous page 140**
  - Homogeneous
  - Centralized
  - Baseline
4. 5Which of the following is not a task of Data Transformation?
  - Conversion
  - Summarization
  - Enrichment
  - **Full Data Refresh page 135**
5. . Which of the following is not an Orr’s Law of Data Quality”?
  - “Data that is not used cannot be corrected!”
  - “Data quality is a function of its use, not its collection!”
  - “Data will be no better than its most stringent use!”

# AL-JUNAID TECH INSTITUTE

- **“Data duplication can be harmful for the organization!”**  
page 181

6. Flat files are one of the prevalent structures used in -----  
data extraction:

- Online
- **Offline** page 134
- Incremental
- Full

7. Which of the following is NOT one of the advantages of changed data capture (CDC) technique?

- Flat files are not required
- **Limited query interface is required for data extraction** page 152
- No incremental on-line I/O required for log tape
- Extraction of changed data occurs immediately

8. The most common range partitioning is on

- Color
- **Date** page 66
- PhoneNo
- Name

9. A relation is said to be in first normal form(1NF), if it does not contain

- Single value column
- **Multi-valued column** page 43
- Derived column
- Composite column

10. In a fully normalized database, too many \_\_\_\_\_ are  
required

- Values
- **Joins** page 49
- Queries

# AL-JUNAID TECH INSTITUTE

- Conditions
11. In the data warehouse, data is collection from ----- sources:
- Homogeneous
  - **Heterogeneous** page 21
  - External
  - Internal
12. De-normalization is more like a “controlled crash” with the aim to ----- without loss of information:
- Check
  - Balance
  - Decrease
  - **Enhance** page 49
13. ----- is making all efforts to increase effectiveness and efficiency in meeting and accepted customer expectation:
- Quality assurance
  - **Quality improvement** page 183
  - Quality maintenance
  - Quality Establishment
14. ----- is the application of intelligence and experience to get common goals.
- **Wisdom** page 11
  - Education
  - Power
  - Information
15. In the data transformation, ----- is the rearrangement and simplification of individual
- Aggregation
  - **Enrichment** page 136

# AL-JUNAID TECH INSTITUTE

- Splitting joining
  - Conversion
16. Grain of a fact table means :
- **The meaning of one fact table row page 109**
  - The meaning of one dimensional table row
  - Summary of aggregates in all fact tables
  - Summary of aggregates in all dimension tables
17. Normalization ----- :
- **Reduces redundancy page 41**
  - Increases redundancy
  - Reduces joins
  - Reduces tables
18. 1 Which of the following is not an example of a typical grain :
- Individual transaction
  - Daily aggregates
  - Monthly aggregates
  - **Normalized attributes page 111**
19. Multi-dimensional databases(MDDs) typically use -----  
----- formats to store pre-summarized cube structures:
- SQL
  - **Proprietary file page 79**
  - Object oriented
  - Non-proprietary file
20. ----- provides a combination of “relational databases  
access” and “cube” data structures within a single framework:
- **HOLAP page 78**
  - DOLAP
  - MOLAP
  - ROLAP

# AL-JUNAID TECH INSTITUTE

21. Data Warehouse provides the best support for analysis while OLAP carries out the ----- task:
- Mandatory
  - Whole
  - **Analysis** page 69
  - Prediction
22. ----- involves splitting a table by columns so that a group of columns is placed into the new table and the remaining columns are placed in another new table:
- **Vertical splitting** page 56
  - Horizontal splitting
  - Adding redundant column
  - None of the given option
23. OLAP implementations are highly/completely ----- :
- Normalized
  - **Demoralized** page 69
  - Predictive
  - Additive
24. If each cell of Relation R contains a single value ( no repeating values) then it is confirmed that :
- **Relation R is in 1<sup>st</sup> Normal Form** page 43
  - Relation R is in 2<sup>nd</sup> Normal Form
  - Relation R is in 3<sup>rd</sup> Normal Form
  - Relation R is in 3<sup>rd</sup> Normal Form but not in 2<sup>nd</sup> Normal Form
25. Which kind of relationships is captured by Fact less fact table:
- **Many- to- Many** page 121
  - One-to-many
  - One-to-one
  - None of the given option
26. Which of the following is not an example of dimension:
- Product

# AL-JUNAID TECH INSTITUTE

- Date
  - Region
  - **Sales volume** page 78
27. Which people criticize Dimensional Modeling (DM) as being a data mart oriented approach?
- Those that consider ER models as Data marts
  - **Those that consider Business processes as Data marts** page 110
  - Those that consider Data marts as Data warehouse
  - Those that consider dimensional model
  - Those that consider dimensional modeling as de-normalization approach
28. In a fully normalized form:
- **To many joins are required** page 49
  - Relationships lose their significance
  - No joins are required
  - Data integrity becomes an issue
29. Which of the following is an example of Non-Additive Facts:
- Quality sold
  - Total sale in Rs.
  - **Discount in percentage** page 119
  - Count of orders in a store
30. Which of the following is not a CUBE operation?
- **ANSI SQL** page 81
  - Roll Up
  - Drill Down
  - Pivoting
31. ----- allows download of “cube” structures to a desktop platform without the need for shared or cube server:

# AL-JUNAID TECH INSTITUTE

- MPLAP
  - ROLAP
  - **DOLAP** page 78
  - HOLAP
32. ROLAP provides access to information via a relational database using:
- **ANSI standard SQL** page 78
  - Proprietary file format
  - Comma Separated Values
  - All of the given option
33. ----- is usually deployed when expression can be used to group data together in such a way that access can be targeted to a small set of partitions:
- Expression elimination
  - **Expression partitioning** page 67
  - Expression indexing
  - None of the given option
34. Taken jointly, the extract programs or naturally evolving systems formed a spider web, also known as
- Distributed Systems Architecture
  - **Legacy System Architecture** page 14
  - Online System Architecture
  - Intranet System Architecture
35. The data has to be checked , cleaned and transformed into a ---  
----- format to allow easy and fast access
- **Unified** page 20
  - Predicated
  - Qualified
  - Proactive



# AL-JUNAID TECH INSTITUTE

36. Suppose in a system A, the values of “PhoneNo” attribute were stored in “countrycode-phone-extension” format, however after transformation into data warehouse the separate columns were used for “countrycode”, “phone” and “extension”. The above scenario is an example of :

- One-to-one scalar transformation
- **One-to-many element transformation** page 144+conceptual
- Many-to-one element transformation
- Many-to-many element transformation

37. In decision support system ease of use is achieved by:

- Normalization
- **Denormalization** page no 49
- Drill up
- Drill down

38. Which of the following is one of the methods to simplify an ER model?

- Normalization
- **Denormalization** page no 103
- HOLAP
- Hybrid schema

39. In ETL process data transformation includes -----

- **Data cleansing** page 129
- Data aggregation
- Behavior checking
- Pattern recognition

40. Non-uniform use of abbreviations, units, and values refers to:

- **Syntactically dirty data** page 160
- Semantically dirty data
- Coverage anomaly
- Extraction issue

# AL-JUNAID TECH INSTITUTE

41. Suppose the size of the attribute “Computerized National Card (CNIC) no. is changed in NADRA database. This transformation refers to:

• **Format revision** page 153

- Field splitting
- Field decoding
- Calculation of derived value

42. The divide and conquer cube partitioning approach helps alleviate the ----- limitations of MOLAP implementation:

- Flexibility
- Maintainability
- Security

• **Scalability** page 85

43. identify the TRUE statement:

- DM is inherently dimensional in nature
- DM comprises of a single central fact table
- DM comprises of a set of dimensional tables

• **All of the given option** Page 103

44. ----- can be used when some columns are rarely accessed rather than other columns or when the table has wide rows or header or both:

- Horizontal splitting
- Pre-joining

• **Vertical splitting** page 56

- Derived attributes

45. Which of the following is an example of derived attributes?

• **Age** page 61

- Size
- Color
- Length

46. The online high performance transaction processing was evolved in -----:

# AL-JUNAID TECH INSTITUTE

- 1980
  - **1975** page 12
  - 1977
  - 1965
47. Cube is a logical entity containing values of a certain aggregation level at an intersection of a combination of -----  
----- :
- Facts
  - **Dimension** page 88
  - Summary tables
  - Primary and foreign key
48. Which of the following is TRUE regarding Entity relationship modeling?
- It does not really model business, but models the micro relationships among data elements.
  - ER modeling does not have “business rules,” it has “data rules
  - ER modeling helps retrieval of individual records having certain critical identifiers.
  - **All of the given option** page 102
49. -----Facilitates a mobile computing paradigamn:
- HOLAP
  - **DOLAP** page78
  - ROLAP
  - MOLAP
50. The main reason(s )for the increase in cube size may be:
- Increase in the number of dimensions
  - Increase in the cardinality of the dimensions
  - Increase in the amount of detail data
  - **All of the given options** page 87

# AL-JUNAID TECH INSTITUTE

51. Suppose the amount of data recorded in an organization is doubled in year. This increase in -----
- Linear
  - Quadratic
  - **Exponential** page 15
  - Logarithmic
52. The data in the data warehouse is ----- :
- Volatile
  - **Non-volatile** page 69
  - Static
  - Non-structured
53. ----- models the macro relationships among data elements with an overall deterministic strategy:
- **Dimensional model** page 102
  - Entity relationship model
  - Object oriented model
  - Structured model
54. ----- technique requires a separate column to specify the time and date when the last modification was occurred:
- Checkmarks
  - **Timestamps** page 150
  - Just-in-Time
  - Real Time extraction
55. Which of the de-normalization technique squeezes master table into detail?
- **Pre-joining** page 58
  - Horizontal splitting
  - Vertical splitting
  - Adding redundant column
56. De-normalization can help:
- Minimize joins

# AL-JUNAID TECH INSTITUTE

- Minimize foreign keys
  - Resolve aggregates
  - **All of the given options** page 51
57. The domain of the “gender” field in some database may be (‘F’, ‘M’) or as (“Female”, “Male”) or even as (1, 0). This is:
- Primary key problem
  - **Non primary key problem** page 163
  - Normalization problem
  - All of the given option
58. Increasing level of normalization ----- number of tables:
- **Increases** page 51
  - Decreases
  - Does not effect
  - None of the given option
59. Which of the following is not a Data Quality Validation Technique:
- Referential integrity
  - Using Data Quality Rules
  - Data Histograming
  - **Indexes** page 189
60. This technique can be used when column from one table is frequently accessed in a large scale join in conjunction with a column from another table:
- Horizontal splitting
  - Pre-joining
  - **Adding redundant column** page 58
  - Derived attributes
61. Data cleansing requires involvement of domain expert because:
- Domain expert has deep knowledge of data aggregation

# AL-JUNAID TECH INSTITUTE

- Change Data captures requires involvement of domain expert
  - **Domain knowledge is required to correct anomalies** page 158
  - Domain expert has deep knowledge of data summarization
62. Relational databases allow you to navigate the data in -----  
--- that is appropriate using the primary , foreign key structure with  
in the data model:
- Only One Direction
  - **Any Direction** page 19
  - Two Direction
  - None of these
63. History is excellent predictor of the -----:
- Past
  - Present
  - **Future** page 15
  - History
64. De- normalization is the process of selectively transforming  
normalized relations into un-normalized physical record  
specifications, with the aim to:
- Well structure the data
  - Well model the data
  - **Reduce query processing time** page 50
  - None of the given option
65. ----- gives total view of an organization:
- OLAP
  - OLTP
  - **Data Warehouse** page 16
  - Database
66. Suppose in system A, the possible values of “Gender” attribute  
were “Male”& “Female”, however in data warehouse ,the values stored

# AL-JUNAID TECH INSTITUTE

were “M” for male and “F” for female. This above scenario is an example of :

- **One-to-one scalar transformation** page 144
  - One-to-many element transformation
  - Many-to-one element transformation
  - Many-to-many element transformation
67. Enrichment is one of the basic tasks in data ----- :
- Extraction
  - **Transformation** page 138
  - Loading
  - Summarization
68. Which of the following is not a technique of De-normalization?
- Pre-joining
  - Splitting tables
  - Adding redundant columns
  - **ER modeling** page 52
69. Which of the following is an example of Additive Facts?
- **Sales Amount** page 119
  - Average
  - Discount
  - Ratios
70. Robotic libraries are needed for -----:
- Cubes
  - Data marts
  - **Data warehouse** page 131
  - Aggregates
71. Normally ROLAP is implemented using -----
- **Star schema** page 87
  - Hybrid schema
  - Pre-defined aggregate
  - All of the given options

# AL-JUNAID TECH INSTITUTE

72. The relation R will be in 2<sup>nd</sup> Normal Form if
- It is in 1NF and each cell contains single value
  - **It is in 1NF and each non key attribute is dependent upon entire primary key** page 44
  - It is in 1NF and non key attribute is dependent upon a single column of composite primary key
  - It is in 1NF and Primary key is composite

73.

n--  
----

- **Nested loop join of quadratic time complexity does not hurt the performance**

**Typical OLTP environments** page 22

- Data warehouse
- DSS
- OLAP

74. In Extract, Load, Transform(ELT) process, data transformation -----:

- **Takes place on the data warehouse server** page 147
- Takes place on a separate transformation server
- Depends on the nature of the source database
- Does not take place

75. Node of a B-Tree is stored in memory block and traversing a B-Tree involves ----- page faults:

- $O(n \log n)$
- **$O(\log n)$**  page 22
- $O(n)$
- $O(n^2)$



# AL-JUNAID TECH INSTITUTE

76. As dimensions get less detailed (e.g. , year vs. day) cubes get -----

-----  
• **Smaller** page 84

- Larger
- Partitioned
- Merged

77. Which of the following is not a technique of “ Changed Data Capture” in currently used Modern Source System?

- Timestamps
- Partitioning
- Triggers
- **Dimensional Modeling** page 150

78. The trade-offs of de-normalization is/are:

- Storage
- Performance
- Ease-of-use
- **All of the given options** page 62

79. If actual data structure does not conform to documented formats then it is called:

• **Syntactically dirty data** page 160

- Semantically dirty data
- Coverage anomaly
- Extraction issue

80. “Header size is reduced, allowing more rows per block , thus reducing I/O” .The above statement is TRUE with respect to:

• **Vertical splitting** page 56

- Horizontal splitting
- Adding redundant column
- None of the given options

# AL-JUNAID TECH INSTITUTE

-----Breaks a table into multiple tables based upon common  
--- column values

**Horizontal** page  
• **splitting** 54

- Vertical splitting
- Adding redundant column
- None of the given option

81. Which of the following is NOT an example of derived attribute?

- Age
- CGPA
- Area of rectangle
- **Height (Conceptual)**

82. Which of the following is NOT an example of derived attribute?

- Age
- CGPA
- Annual Salary

83. If a table is expected to have six columns but some or all of the records do not have six columns then it is example of:

- **Syntactically dirty data** page 160
- Semantically dirty data
- Coverage anomaly
- Extraction issue

84. MDX by Microsoft is an example of -----:

- HOLAP
- DOLAP
- ROLAP
- **None of the given options** page 79

# AL-JUNAID TECH INSTITUTE

85. The growth of master files and magnetic tapes exploded around the mid- -----
- 1950s
  - **1960s page 12**
  - 1970s
  - 1980s
86. If one or more records in a relational table do not satisfy one or more integrity constraint , then the data:
- Is syntactically dirty
  - **Is semantically dirty page 160**
  - Has Coverage anomaly
  - Has extraction issue
87. OLAP is:
- **Analytical processing page 69**
  - Transaction processing
  - Additive processing
  - Active processing
88. One of the possible issues faced by web scrapping is that:
- **Web pages may contain junk data page 141**
  - Web pages do not contain multiple facts
  - Web pages do not contain multiple dimensions
  - Web pages does not support transformation
89. Which of the following is\are example of dimension:
- **Product page 79**
  - Region
  - Data
  - None of the given
90. An OLTP system is always good at -----:
- **Evolving data page 122**
  - Keeping static data

# AL-JUNAID TECH INSTITUTE

- Tracking past data
- Maintaining historic data

91. In case of multiple sources for the same data element , we need to prioritize the source systems per element based, the process is called:

- **Ranking page 143**

- Prioritization
- Element selection
- Measurement selection

92. One feature of Change Data Capture (CDC) is that:

- It pre-calculates changed aggregates
- It loads the transformed data in real time
- It only processes the data has been changed

- **It can automate the transformation of extracted data page 149**

93. In ----- SQL generation is vastly simplified for front-end tools when the data is highly structure:

- MOLAP

- **Star Schema page 107**

- Hybrid schema
- Object oriented schema

94. Dirty data means:

- Data cannot be aggregated
- Data contains non-additive facts
- Data does not fulfill dimensional modeling rules

- **Data does not conform to proper domain definitions page 158**

# AL-JUNAID TECH INSTITUTE

95. In Context of Change Data Capture (CDC) sometimes a -----  
----- object can be used to store recently modified data:
- Buffer table
  - **Change table page 149**
  - Checkmark table
  - Change control table
96. “Sometimes during data collection complete entities are missed”. This statement is an example of :
- **Missing tuple page 161**
  - Missing attribute
  - Missing aggregates
  - Semantically dirty data
97. Table collapsing technique is applied in case of:
- **One-by-one relation or many-to –many relation page 52**
  - One-to-many relation
  - Many-to-many relation
  - None of the given option
98. Which of the following is an example of dimension?
- Product
  - Region
  - Date
  - **All of the given optionpage 78**
99. Data warehouse stores -----:
- Operational data
  - **Historical data page 24**
  - Meta data
  - Log files data
100. The business process covered by ER diagrams:
- **Do not co-exist in time and space page 109**
  - Co-exist in time and space

# AL-JUNAID TECH INSTITUTE

- Do not physically exist in real time context
  - None of the given options
101. The main goal of normalization is to eliminate -----:
- **Data redundancy page 41**
  - Data sharing
  - Data security
  - Data consistency
102. Serious ----- involves decomposing and reassembling the data:
- **Data cleansing page 168**
  - Data transformation
  - Data loading
  - Data extraction
103. In the data warehouse environment the data is -----
- **Subject-oriented page 69**
  - Time-oriented
  - Both subject and time oriented
  - Neither time-oriented nor subject-oriented
104. For large record spaces and large number of records, the run time of the clustering algorithms:
- **Prohibitive page 164**
  - Static
  - Exponential
  - Numerical
105. ----- can result in costly errors, such as, False frequency distributions and incorrect aggregates due to double counting:
- **Data duplication page 165**
  - Data reduction
  - Data anomaly
  - Data transformation

# AL-JUNAID TECH INSTITUTE

106. The degree to which values are present in the attributes that require them is known as -

-----:

• **Completeness** page 185

- Uniqueness
- Accessibility
- Consistency

107. Time complexity of Key Creation process in basic Sorted Neighborhood (BSN) Method is

-----:

- $O(n \log n)$
- $O(\log n)$
- **$O(n)$**  page 171
- $O(2n)$

108. Which of the following is an example of slowly changing dimensions?

• **Inheritance** page 124

- Aggregation
- Association
- Asset disposal

109. The ----- operator proves useful in more complex metrics applicable to the dimensions and accessibility:

• **Max** page 188

- Min
- Max and Min
- None of the given

110. In OLAP, the typical write operation is ----- :

• **Bulk insertion** page 75

- Single insertion
- Sequential insertion
- No insertion

111. The issue(s) of “ Adding redundant column” includes(s):

# AL-JUNAID TECH INSTITUTE

- Increase in table size
- Maintenance
- Loss of information
- **All of the given option page 65**

112. ----- is applicable in Profitability analysis:

- OLTP
- **Data warehouse page 36,37**
- Information System(IS)
- Management Information System(MIS)

113. The hardware (CPU) utilization in data warehouse environment is full or ----- :

- Fixed
- Partial
- **Not at all page 24**
- Slow

114. Time variant is a characteristics of data warehouse which means:

- **Data loaded in data warehouse will be time stamped page 20**
- Data can be loaded in data warehouse anytime
- Data can be loaded in data warehouse only at a particular time
- Data cannot be loaded in data warehouse with respect to time

115. In which class of aggregates AVERAGE function can be placed:

- **Algebraic page 120**
- Distributed
- Associative



# AL-JUNAID TECH INSTITUTE

- Holistic
116. Considered the following Employee table and identify the column which causes that the table is not in first normal form(1NF):  
(Emp\_ID, Emp\_Name ,Emp\_skills, Emp\_Designation)
- Emp\_ID
  - Emp\_Name
  - **Emp\_skills page 43(conceptual)**
  - Emp\_Designation
117. The application of data and information leads to -----
- Intelligence
  - Experience
  - **Knowledge page 11**
  - Power
118. 1----- segregate data into separate partitions so that queries do not need to examine all data in a table when WHERE clause filters specify only a subset of the partitions.
- Pre-joining technique
  - Collapsing table technique
  - **Horizontal splitting technique page 56**
  - Vertical splitting technique
119. -----should not be present in a relation, so that it would be in second normal form (2NF).
- **Partial dependency page 44 (conceptual)**
  - Full functional dependency
  - Multivalued dependency
  - Transitive dependency
120. 1Records referring to the same entity are represented in different formulas in the different data sets or are represented erroneously. Thus duplicate records will appear in the merged database. This problem is known as-----.

# AL-JUNAID TECH INSTITUTE

- **Merge/purge problem** page 168

- Duplication problem
- Redundant duplication problem
- Redundant problem

121. The data perspective in OLTP system is operational, while that in data warehouse

is:

- Fully normalized
- Fully de-normalized
- Fully summarized

- **Historical and detailed** page 30

122. 1Simple scalar transformation is a-----mapping from one set of values to another set of values using straightforward rules.

- **One-to-one** page 144

- One-to-many
- Many-to-many
- Many-to-one

123. -----can be created in operational systems to keep tracks of recently updated records.

- **Triggers** page 150

- Timestamps
- Partitioning
- ELT

124. 1Development of data warehouse is hard because data sources are usually-----

- Structured and homogeneous

- **Unstructured and heterogeneous** page 31

- Structured and heterogeneous
- Unstructured and homogeneous

125. In a decision support environment, the decision maker is interested in -----.

# AL-JUNAID TECH INSTITUTE

- Only limited organizational data
  - **Big picture of organizational data page 21**
  - Only sale related data
  - Only customer related data
126. Information can answer question like “what”, “who” and “when” while knowledge can answer question like-----  
--.
- Why
  - Where
  - Which
  - **How page 11**
127. OLTP implementations are fully-----.
- **Normalized page 69**
  - Denormalized
  - Predictive
  - Additive
128. Which logical data extraction has significant performance impacts on the data warehouse server?
- **Incremental Extraction page 133**
  - Online Extraction
  - Offline Extraction
  - Legacy Vs OLTP
129. Consider the following Student table and identify the column which causes that the table is not in first normal form(1NF).
- Student(Std\_ID, Std\_Name ,Std\_CGPA ,Std\_Hobbies)
- Std\_ID
  - Std\_Name
  - Std\_CGPA

# AL-JUNAID TECH INSTITUTE

- **Std\_Hobbies page 43(Conceptual)**

130. Analytical processing uses -----

- **Multi-level aggregates page 74**

- Record level aggregates
- Table level aggregates
- All of the given options

131. Which is not a class of anomalies in following?

- **Dirty anomalies page 160**

- Syntactically dirty data
- Semantically dirty data
- Coverage anomalies

132. 1----- is a system of activities that assures conformance of product to pre-established requirements.

- **Quality assurance page 183**

- Quality improvement
- Quality Maintenance
- Quality Establishment

133. Two interesting examples of quality dimensions that can make use of min operator are -----

- **Believability and appropriate amount of data page 188**

- Believability and consistency
- Believability and Redundancy
- Reliability and appropriate amount of data

134. ----- in database or data warehouse has no actual value; it only has potential value.

- **Data page 181**

- Entity
- Flat tables
- Data marts

# AL-JUNAID TECH INSTITUTE

135. In OLTP environment the selectivity is ----- and ---  
----- in data warehouse environment.
- **High, Low page 22**
  - Low, High
  - High, Fixed
  - Fixed, Low
136. Which is not a/an characteristics of data quality?
- **Reliability page 186**
  - Uniqueness
  - Accessibility
  - Consistency
137. If a product meets formally defined “requirement specifications”, yet fails to be a quality product from the customer’s perspective, this means the requirements were -----.
- **Defective page 180**
  - Unclear
  - Unrefined
  - Undefined
138. The relation R will be in 3<sup>rd</sup> Normal Form if:
- It is in 2NF each cell contains single value
  - **It is in 2NF and every non-key column is non-key transitively dependent upon its primary key. Page 46**
  - It is in 1NF and each non key attribute is dependent upon a single column of composite primary key.
  - It is in 2NF and each non key attribute is dependent upon other non-key attribute.
139. Decision support system queries deal with number of columns ---  
-----
- Having numeric values
  - In a single table

# AL-JUNAID TECH INSTITUTE

- In a single view
  - **Spanning across multiple tables page 21**
140. Normalization is used to reduce:
- **Reduces redundancy page 41**
  - Increases redundancy
  - Reduces joins
  - Reduces tables
141. The end user of data warehouse are-----
- Programmers
  - Database developers
  - Data entry operator
  - **Business executives page 18 + 19**
142. Which one are the characteristics of data warehouse queries?
- Use primary key
  - High selectivity
  - **Use multiple tables page 30**
  - Very low performance
143. Referential Integrity (RI) using a naive approach would take ----- time.
- **$O(n)$  page 160**
  - $O(1)$
  - $O(\log n)$
  - None of the given
144. Web scrapping is a process of applying ----- techniques to the web
- **Screen scrapping page 146**
  - Data scrapping
  - Text scrapping
  - Meta scrapping
145. Which is not an issue of ROLAP in the following?
- **Standard hierarchy of dimensions page 92**

# AL-JUNAID TECH INSTITUTE

- Non-standard conventions
  - Maintenance
  - Aggregation
146. One of the fundamental purpose of de-normalization is to -----  
----- a number of physical tables which ultimately reduce the  
number of joins to answer a query.
- Delete
  - **Reduce page 50**
  - Increase
  - Decrease
147. ----- is not the characteristic of data warehouse.
- Time variant
  - Subject-oriented
  - Integrated
  - **Volatile page 69**
148. Which is not a/an step of data cleansing procedure?
- **Aggregation page 168**
  - Elementizing
  - Standardizing
  - Verifying
149. 1 Instance matching between different sources is then achieved by  
a standard -----  
on identifying attribute(s), if you are very, very, very lucky.
- **Equi-join page 169**
  - Inner join
  - Outer join
  - Fuller join
150. Ad-hoc access of data warehouse means:
- That have predefined database access pattern

# AL-JUNAID TECH INSTITUTE

- **That does not have predefined database access pattern**  
**page 18**

- That could be accessed by any user
- That could not be accessed by any user

151. In OLTP environment, the size of tables is relatively-----  
---

- Large
- Fixed
- Moderate
- **Small** **page 22**

152. 1----- is a/an measure of how current or up to date the data is:

- **Timeliness** **page 185**
- Completeness
- Accessibility
- Consistency

153. 1The process of converting entity relationship model in to dimensional model of -----  
steps:

- Two
- Three
- **Four** **page 109**
- Five

154. A ----- Is defined by a group of records that have similar characteristics (“behavior”) for p% of the fields in the data set, where p is a user- defined value(usually above 90).

- **Pattern** **page 164**
- Cluster
- Entity
- Attribute



# AL-JUNAID TECH INSTITUTE

155. ----- is known as state of being only one of its kind or being without an equal or parallel.

- Completeness
- **Uniqueness page 185**
- Accessibility
- Consistency

156. Which of the following is not an example of fact?

• **Account no page 74**

- Sales quantity
- Per unit sales amount
- Sales amount

157. 1-----is the degree to which data accurately reflects the real world object that the data represents.

• **Intrinsic data quality page 181**

- Realistic data quality
- Strong data quality
- Weak data quality

158. 1Which one among the following data warehouse stores data containing long period?

- Telecommunication data warehouse
- Financial data warehouse
- Human resource data warehouse
- **Insurance data warehouse page 36**

159. A \_\_\_\_\_ dimension is a collection of random transactional codes, flags and/text attributes that are unrelated to any particular dimension. The \_\_\_\_\_ dimension is simply a structure that provides a convenient place to store the \_\_\_\_\_ attributes.

# AL-JUNAID TECH INSTITUTE

- **Junk**

- Time
- Parallel
- None of these

160. Data Warehouse is about taking / collecting data from different \_\_\_\_\_ sources.

- Harmonized
- Identical
- **Homogeneous**
- Heterogeneous

NOT CONFIRM

161. Taken jointly, the extract programs or naturally evolving systems formed a spider web, also known as

- Distributed Systems Architecture
- **Legacy Systems Architecture**
- Online Systems Architecture
- Intranet Systems Architecture

162. It is observed that every year the amount of data recorded in an organization

- **Doubles**
- Triples
- Quartiles

163. The users of data warehouse are knowledge workers in other words they are \_\_\_\_\_ in the organization.

- DWH Analyst
- **Decision maker**
- Database Administrator
- Manager

164. Node of a B-Tree is stored in memory block and traversing a B-Tree involves \_\_\_\_\_ page faults.

- $O(n \lg n)$
- **$O(\log n)$  {  $O(\log n)$  it's the real answer }**
- $O(n)$
- $O(n^2)$

165. In \_\_\_\_\_ system, the contents change with time.

# AL-JUNAID TECH INSTITUTE

- **OLTP**

- ATM
- DSS
- OLAP

166. The growth of master files and magnetic tapes exploded around the mid-\_\_\_\_\_.

- 1950s.
- **1960s.**
- 1970s.
- 1980s.

167. Relational databases allow you to navigate the data in \_\_\_\_\_ that is appropriate using the primary, foreign key structure within the data model

- Only One Direction
- **Any Direction**
- Two Direction
- None of these

168. Naturally Evolving architecture occurred when an organization had a \_\_\_\_\_ approach to handling the whole process of hardware and software architecture.

- **Relaxed**
- Good
- Not Relaxed
- None

169. \_\_\_\_\_ gives total view of an organization

- OLAP
- OLTP
- **Data Warehouse**
- Database

170. Suppose the amount of data recorded in an organization is doubled every year. This increase is

- Linear
- Quadratic
- **Exponential**

# AL-JUNAID TECH INSTITUTE

- Logarithmic

171. \_\_\_\_\_ is the application and experience to get common goals.

➤ **Wisdom**

- Power
- Information
- Education

172. Normally ROLAP is implemented using \_\_\_\_\_.

➤ **Star Schema**

- Pre-defined aggregates
- Hybrid Schema
- All of the given options

173. Identify data warehouse query from the following.

- Number of students studying CS614 course
- **Factors which can effect students' performance**

- List of students belong to Lahore city
- Total number of students that have 3.5 CGPA

174. In \_\_\_\_\_ SQL generation is vastly simplified for front-end tools when the data is highly structured.

- Hybrid Schema
- **Star Schema**
- MOLAP
- Object oriented schema

175. \_\_\_\_\_ models the Marco relationships among data elements with an overall deterministic strategy.

➤ **Dimensional model**

- Structured model
- Object oriented model
- Entity relationship model

# AL-JUNAID TECH INSTITUTE

176. The \_\_\_\_\_ task is typically performed after most other transformation and cleaning steps have taken place.

➤ Data extraction

➤ **Data duplicate elimination**

➤ Data transformation

➤ Data loading

177. \_\_\_\_\_ is known as state of being only one of its kind or being without an equal or parallel.

➤ Completeness

➤ Consistency

➤ Accessibility

➤ **Uniqueness**

178. Relational modeling techniques are used to develop \_\_\_\_\_ systems.

➤ **OLTP**

➤ MOLAP

➤ OLAP

➤ ROLAP

179. Which of the following is not CUBE operation?

➤ **ANSI SQL**

➤ Drill Down

➤ Pivoting

➤ Roll UP

180. The hybrid OLAP(HOLAP) solution is a mix of \_\_\_\_\_ architecture that supports queries against summary and transaction in an integrated.

➤ ROLAP

➤ OLTP

➤ **MOLAP and ROLAP**

➤ MOLAP

# AL-JUNAID TECH INSTITUTE

181. The response time of OLTP queries are in milliseconds to seconds. While that of data warehouse queries are in:

- Days of weeks
- Hours to days
- Milliseconds to microseconds
- **Minutes to hours**

182. \_\_\_\_\_ allows download of “cube” structures to a desktop platform without the need for shared relational or cube server.

- ROLAP
- HOLAP
- **DOLAP**
- MOLAP

183. In case of multiple sources for the same data element. We need to prioritize the source systems per element bases, the process is called:

- **Ranking**
- Measurement selection
- Element selection
- Prioritization

184. OLAP is:

- Implementation technique
- **Framework**
- None of the given options
- Physical database design

185. Which of the following is not a task of Data Transformation?

- Summarization
- Enrichment
- Conversion
- **Full data refresh**

# AL-JUNAID TECH INSTITUTE

186. Simple scalar transformation is a \_\_\_\_\_ mapping from one set of values to another set of values using straightforward rules.

- Many-to-one
- Many-to-many
- **One-to-one**
- One-to-many

187. Which of the following is not a data quality validation technique?

- **Indexes**
- Using data quality rules
- Referential integrity
- Data Histograming

188. \_\_\_\_\_ are created out from the data warehouse to service different needs of departments in the organization.

- Databases
- **Data marts**
- MIS
- Tables

189. An OLTP system is always good at \_\_\_\_\_.

- Maintaining historical data
- Keeping static data
- Tracking past data

➤ **Evolving data**

190. \_\_\_\_\_ incorporates the concept of product quality, process control, quality improvement.

- Intrinsic data quality management
- Realistic Data Quality Management
- **Total Quality Management**
- Strong Data Quality Management

# AL-JUNAID TECH INSTITUTE

191. “Sometimes during data collection complete entities are missed”. This statement is an example of:

- Missing aggregate
- **Missing tuple**
- Semantically dirty data
- Missing attribute

192. The last step of Software Development Life Cycle (SDLC) is implementation. While that of data warehouse is:

- Integration
- **Understanding requirements**
- Testing
- Analysis

193. In which class of aggregates Max function can be places?

- Algebraic
- Associative
- Holistic
- **Distributive**

194. Telecommunication data warehouse is dominated by the \_\_\_\_\_ volume of data generated at the call level.

- Incomplete
- Semi-incomplete
- **Sheer**
- Partial

195. If a table is expected to have six columns but some or all of the records do not have six columns then it is example of:

- Semantically dirty data
- **Syntactically dirty data**
- Extraction issue
- Coverage anomaly

196. In the data warehouse environment the date is \_\_\_\_\_.



# AL-JUNAID TECH INSTITUTE

- **Both subject and time oriented**
- Neither time-oriented nor subject-oriented
- Subject-oriented
- time-oriented

197. **dirty data means that:**

- Data control be aggregated
- Data contain non-additive facts
- Data does not fulfill dimensional modeling rules
- **Data does not conform to proper domain definitions**

198. **One of the main reason for the failure of DWH deployments is \_\_\_\_\_.**

- Data anomaly
- Data integrity
- **Data quality**
- Data duplication

199. **The data in the data warehouse is \_\_\_\_\_.**

- Non-structured
- Volatile
- Static

➤ **Non-volatile**

200. **Which of the following is NOT an example of a typical grain?**

- Daily aggregates
- Monthly aggregates
- Individual transaction
- **Normalized attribute**