

AL-JUNAID INSTITUTE GROUP



LMS handling service are
available

PAID

Assignment, QUIZ, GDBZ and
GRAND QUIZ SERVICE are
available

0304-1659294

VULMSHELP.COM

JUNAID MALIK

0304-1659294

AL-JUNAID INSTITUTE GROUP

CS614 GRAND QUIZ

Online Extraction is a kind of-----data extraction.

- Logical
- Dimensional

○ **Physical page 132**

- Multi valued

The ----- saw the advent of disk storage, or DASD(direct Access Storage Device) :

- 1960s
- **1970s page 13**
- 1950s
- 1990s

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

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

- Conversion
- Summarization
- Enrichment
- **Full Data Refresh page 135**

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

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

- Online
- **Offline page 134**

AL-JUNAID INSTITUTE GROUP

- Incremental
- Full

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

The most common range partitioning is on

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

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

In a fully normalized database, too many _____ are required

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

In the data warehouse, data is collection from ----- sources:

- Homogeneous
- **Heterogeneous** page 21
- External
- Internal

De-normalization is more like a “controlled crash” with the aim to ----- without loss of information:

- Check
- Balance
- Decrease
- **Enhance** page 49

----- is making all efforts to increase effectiveness and efficiency in meeting and accepted customer expectation:

- Quality assurance

AL-JUNAID INSTITUTE GROUP

- **Quality improvement** page 183

- Quality maintenance
- Quality Establishment

----- is the application of intelligence and experience to get common goals.

- **Wisdom** page 11

- Education
- Power
- Information

In the data transformation, ----- is the rearrangement and simplification of individual

- Aggregation

- **Enrichment** page 136

- Splitting joining
- Conversion

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

Normalization ----- :

- **Reduces redundancy** page 41

- Increases redundancy
- Reduces joins
- Reduces tables

Which of the following is not an example of a typical grain :

- Individual transaction
- Daily aggregates
- Monthly aggregates
- **Normalized attributes** page 111

Multi-dimensional databases(MDDs) typically use ----- formats to store pre-summarized cube structures:

- SQL
- **Proprietary file** page 79
- Object oriented
- Non-proprietary file

AL-JUNAID INSTITUTE GROUP

----- provides a combination of “relational databases access” and “cube” data structures within a single framework:

- **HOLAP** page 78
- DOLAP
- MOLAP
- ROLAP

Data Warehouse provides the best support for analysis while OLAP carries out the ----- task:

- Mandatory
- Whole
- **Analysis** page 69
- Prediction

----- 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

OLAP implementations are highly/completely ----- :

- Normalized
- **Demoralized** page 69
- Predictive
- Additive

If each cell of Relation R contains a single value (no repeating values) then it is confirmed that :

- **Relation R is in 1st Normal Form** page 43
- Relation R is in 2nd Normal Form
- Relation R is in 3rd Normal Form
- Relation R is in 3rd Normal Form but not in 2nd Normal Form

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

Which of the following is not an example of dimension:

- Product
- Date
- Region

AL-JUNAID INSTITUTE GROUP

- **Sales volume** page 78

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

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

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

Which of the following is not a CUBE operation?

- **ANSI SQL** page 81
- Roll Up
- Drill Down
- Pivoting

----- allows download of “cube” structures to a desktop platform without the need for shared or cube server:

- MPLAP
- ROLAP
- **DOLAP** page 78
- HOLAP

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

AL-JUNAID INSTITUTE GROUP

----- 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

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

The data has to be checked , cleaned and transformed into a ----- format to allow easy and fast access

- **Unified** page 20
- Predicated
- Qualified
- Proactive

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

In decision support system ease of use is achieved by:

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

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

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

In ETL process data transformation includes -----

AL-JUNAID INSTITUTE GROUP

- **Data cleansing** page 129

- Data aggregation
- Behavior checking
- Pattern recognition

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

- **Syntactically dirty data** page 160

- Semantically dirty data
- Coverage anomaly
- Extraction issue

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

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

- Flexibility
- Maintainability
- Security

- **Scalability** page 85

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

----- 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

Which of the following is an example of derived attributes?

- **Age** page 61

- Size
- Color
- Length

AL-JUNAID INSTITUTE GROUP

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

- 1980
- **1975** page 12
- 1977
- 1965

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

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

-----Facilitates a mobile computing paradigam:

- HOLAP
- **DOLAP** page78
- ROLAP
- MOLAP

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

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

- Linear
- Quadratic
- **Exponential** page 15
- Logarithmic

The data in the data warehouse is ----- :

- Volatile

AL-JUNAID INSTITUTE GROUP

- **Non-volatile** page 69

- Static
- Non-structured

----- models the macro relationships among data elements with an overall deterministic strategy:

- **Dimensional model** page102

- Entity relationship model
- Object oriented model
- Structured model

----- 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

Which of the de-normalization technique squeezes master table into detail?

- **Pre-joining** page 58

- Horizontal splitting
- Vertical splitting
- Adding redundant column

De-normalization can help:

- Minimize joins
- Minimize foreign keys
- Resolve aggregates

- **All of the given options** page 51

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

Increasing level of normalization

number of
tables:

- **Increases** page 51

- Decreases
- Does not effect

AL-JUNAID INSTITUTE GROUP

- None of the given option

Which of the following is not a Data Quality Validation Technique:

- Referential integrity
- Using Data Quality Rules
- Data Histograming
- **Indexes** page 189

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

Data cleansing requires involvement of domain expert because:

- Domain expert has deep knowledge of data aggregation
- 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

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

History is excellent predictor of the -----:

- Past
- Present
- **Future** page 15
- History

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

----- gives total view of an organization:

- OLAP
- OLTP

AL-JUNAID INSTITUTE GROUP

- **Data Warehouse** page 16

- Database

Suppose in system A, the possible values of “Gender” attribute were “Male” & “Female”, however in data warehouse, the values stored 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

Enrichment is one of the basic tasks in data ----- :

- Extraction

- **Transformation** page 138

- Loading
- Summarization

Which of the following is not a technique of De-normalization?

- Pre-joining
- Splitting tables
- Adding redundant columns

- **ER modeling** page 52

Which of the following is an example of Additive Facts?

- **Sales Amount** page 119

- Average
- Discount
- Ratios

Robotic libraries are needed for -----:

- Cubes
- Data marts

- **Data warehouse** page 131

- Aggregates

Normally ROLAP is implemented using -----

- **Star schema** page 87

- Hybrid schema
- Pre-defined aggregate
- All of the given options

The relation R will be in 2nd Normal Form if

- It is in 1NF and each cell contains single value

AL-JUNAID INSTITUTE GROUP

- **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

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

- **Typical OLTP environments** page 22

- Data warehouse
- DSS
- OLAP

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

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
- $O(n)$
- $O(n^2)$

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

- **Smaller** page 84

- Larger
- Partitioned
- Merged

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

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

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

AL-JUNAID INSTITUTE GROUP

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

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

“Header size is reduced, allowing more rows per back , 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

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

- **Horizontal splitting** page 54

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

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

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

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

- Age
- CGPA
- Annual Salary

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

AL-JUNAID INSTITUTE GROUP

- Semantically dirty data
- Coverage anomaly
- Extraction issue

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

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

The growth of master files and magnetic tapes exploded around the mid-----

- 1950s
- **1960s** **page 12**
- 1970s
- 1980s

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

OLAP is:

- **Analytical processing** **page 69**
- Transaction processing
- Additive processing
- Active processing

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

Which of the following is\are example of dimension:

- **Product** **page 79**
- Region
- Data
- None of the given

AL-JUNAID INSTITUTE GROUP

An OLTP system is always good at -----:

- **Evolving data page 122**

- Keeping static data
- Tracking past data
- Maintaining historic data

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

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**

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

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**

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

AL-JUNAID INSTITUTE GROUP

“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

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

Which of the following is an example of dimension?

- Product
- Region
- Date
- **All of the given option page 78**

Data warehouse stores -----:

- Operational data
- **Historical data page 24**
- Meta data
- Log files data

The business process covered by ER diagrams:

- **Do not co-exist in time and space page 109**

- Co-exist in time and space
- Do not physically exist in real time context
- None of the given options

The main goal of normalization is to eliminate -----:

- **Data redundancy page 41**

- Data sharing
- Data security
- Data consistency

Serious ---- involves decomposing and reassembling the data:

- **Data cleansing page 168**

- Data transformation
- Data loading
- Data extraction

AL-JUNAID INSTITUTE GROUP

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

For large record spaces and large number of records , the run time of the clustering algorithms:

- **Prohibitive** page 164

- Static
- Exponential
- Numerical

----- 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

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

-----:

- **Completeness** page 185

- Uniqueness
- Accessibility
- Consistency

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

-----:

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

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

- **Inheritance** page 124

- Aggregation
- Association
- Asset disposal

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

- **Max** page 188

AL-JUNAID INSTITUTE GROUP

- Min
- Max and Min
- None of the given

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

• **Bulk insertion page 75**

- Single insertion
- Sequential insertion
- No insertion

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

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

----- is applicable in Profitability analysis:

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

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

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

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

In which class of aggregates AVERAGE function can be placed:

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

AL-JUNAID INSTITUTE GROUP

- Holistic

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

The application of data and information leads to -----

- Intelligence
- Experience
- **Knowledge page 11**
- Power

----- 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

-----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

Records 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-----.

- **Merge/purge problem page 168**
- Duplication problem
- Redundant duplication problem
- Redundant problem

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

- Fully normalized
- Fully de-normalized
- Fully summarized

AL-JUNAID INSTITUTE GROUP

- **Historical and detailed** page 30

Simple 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

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

- **Triggers** page 150

- Timestamps
- Partitioning
- ELT

Development of data warehouse is hard because data sources are usually-----

- Structured and homogeneous
- **Unstructured and heterogeneous** page 31

- Structured and heterogeneous
- Unstructured and homogeneous

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

- Only limited organizational data
- **Big picture of organizational data** page 21

- Only sale related data
- Only customer related data

Information can answer question like “what”, “who” and “when” while knowledge can answer question like-----.

- Why
- Where
- Which

- **How** page 11

OLTP implementations are fully-----.

- **Normalized** page 69

- Denormalized
- Predictive
- Additive

Which logical data extraction has significant performance impacts on the data warehouse server?

AL-JUNAID INSTITUTE GROUP

• **Incremental Extraction** page 133

- Online Extraction
- Offline Extraction
- Legacy Vs OLTP

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

• **Std_Hobbies** page 43(Conceptual)

Analytical processing uses -----

• **Multi-level aggregates** page 74

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

Which is not a class of anomalies in following?

• **Dirty anomalies** page 160

- Syntactically dirty data
- Semantically dirty data
- Coverage anomalies

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

• **Quality assurance** page 183

- Quality improvement
- Quality Maintenance
- Quality Establishment

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

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

• **Data** page 181

AL-JUNAID INSTITUTE GROUP

- Entity
- Flat tables
- Data marts

In OLTP environment the selectivity is ----- and ----- in data warehouse environment.

- **High, Low** page 22

- Low, High
- High, Fixed
- Fixed, Low

Which is not a/an characteristics of data quality?

- **Reliability** page 186

- Uniqueness
- Accessibility
- Consistency

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

The relation R will be in 3rd Normal Form if:

- It is in 2NF each 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.

Decision support system queries deal with number of columns -----

- Having numeric values
- In a single table
- In a single view
- **Spanning across multiple tables** page 21

Normalization is used to reduce:

- **Reduces redundancy** page 41
- Increases redundancy

AL-JUNAID INSTITUTE GROUP

- Reduces joins
- Reduces tables

The end user of data ware house are-----.

- Programmers
- Database developers
- Data entry operator
- **Business executives page 18 + 19**

Which one are the characteristics of data warehouse queries?

- Use primary key
- High selectivity
- **Use multiple tables page 30**
- Very low performance

Assume a company with a multi- million row customer table i.e. n rows. Checking for Referential Integrity (RI) using a naive approach would take ----- time.

- **O(n) page 160**
- O(1)
- O(log n)
- None of the given

Web scrapping is a process of applying ----- techniques to the web

- **Screen scrapping page 146**
- Data scrapping
- Text scrapping
- Meta scrapping

Which is not an issue of ROLAP in the following?

- **Standard hierarchy of dimensions page 92**
- Non-standard conventions
- Maintenance
- Aggregation

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

----- is not the characteristic of data warehouse.

- Time variant
- Subject-oriented
- Integrated

AL-JUNAID INSTITUTE GROUP

- **Volatile** page 69

Which is not a/an step of data cleansing procedure?

- **Aggregation** page 168
- Elementizing
- Standardizing
- Verifying

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

Ad-hoc access of data warehouse means:

- That have predefined database access pattern
- **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

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

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

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

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

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

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

AL-JUNAID INSTITUTE GROUP

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

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

- Completeness
- **Uniqueness** page 185

- Accessibility
- Consistency

Which of the following is not an example of fact?

- **Account no** page 74

- Sales quantity
- Per unit sales amount
- Sales amount

-----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

Which 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

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.

- **Junk**

- Time
- Parallel
- None of these

AL-JUNAID INSTITUTE GROUP

Data Warehouse is about taking / collecting data from different _____ sources.

- Harmonized
- Identical
- **Homogeneous** NOT CONFIRM
- Heterogeneous

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

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

- **Doubles**
- Triples
- Quartiles

The users of data warehouse are knowledge workers in other words they are _____ in the organization.

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

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 }**
- $O(n)$
- $O(n^2)$

In _____ system, the contents change with time.

- **OLTP**
- ATM
- DSS
- OLAP

The growth of master files and magnetic tapes exploded around the mid- _____.

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

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

AL-JUNAID INSTITUTE GROUP

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

_____ gives total view of an organization

- OLAP
- OLTP

- **Data Warehouse**

- Database

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

- Linear
- Quadratic
- **Exponential**
- Logarithmic

Which people criticized Dimensional Modeling (DM) as being a data mart oriented approach?

Those that consider ER model as Data marts

Which of the following is not a CUBE operation?

ANSI SQL

If actual data structure does not confirm to documented formats then it is called:

Semantically dirty data

This technique can be used when a column from one table is frequently accessed in a large scale join in conjunction with

Adding redundant column

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

Vertical splitting

Gives total view of an organization.

Data Warehouse

ROLAP provides access to information via a relational database using

ANSI standard SQL

The main reason(s) for the increase in cube size may be

AL-JUNAID INSTITUTE GROUP

All of the given

If each cell of Relation R contains a single value (no repeating values) then it is confirmed that

Relation R is in 3rd Normal Form but in 2nd Normal Form

In Extract, Load, Transformation (ELT) process, you don't need to purchase extra devices to achieve parallelism because.

You already have parallel data warehouse servers

Which of the following is not a technique of "Changed Data Capture" in currently used Modern Source Systems?

Dimensional Modeling

Breaks a table into multiple tables based upon common column values

Horizontal splitting

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

Height

Grain of a fact table means

The meaning of one fact table row

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

Email Address

Which of the following is an example of Additive Facts?

Average

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

Involves splitting a table by columns so that a group of columns is placed into the new table and the remaining

Vertical splitting

Normalization.

Reduces redundancy

De-normalization affects:

Database size and query performance

AL-JUNAID INSTITUTE GROUP

MDX by Microsoft is an example of

None of the given options

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

Scalability

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

Syntactically dirty data

Allows download of “cube” structures to a desktop platform without the need for shared relational or cube server.

DOLAP

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

O (lg n)

The growth of master files and magnetic tapes exploded around the mid

1960s.

<https://vulmshep.com/>