Prepared by: Dr. Tariq Hanif
Midterm Spring-13
For more help rep at:
qirathanif@gmail.com
Cell# 0303-7300008



CS614 Data Warehousing

01. _____ is one class of decision support environment.

OLAP 30page

OLTP
Data Cleansing
ETL
2. The confusion created by data redundancy makes it difficult for companies to

Create online processing capabilities.
Work in batch processing load.
Use a distributed database.
Integrate data from different sources.
3. Effects of de-normalization on database performance are

Unpredictable pg 62

Predictable
Conventional
Unsurprising
04. OLAP is a (n) _____ of application.

Classification pg74

Amalgamation
Unification
Blending
5. DOLAP model facilitates _____ computing paradigm.

Mobile pg 78

Permanent
Rigid
Strict
6. _____ is the lowest level of detail or the atomic level of data stored in the warehouse.

Cube
Grain pg 111
Virtual Cube

Solved By Dr. Tariq Hanif

Prepared by: Dr. Tariq Hanif MHT-4th VU-Sahiwal Campus
Email at:qirathanif@gmail.com
Cell# 0303-7300008

Aggregate
7. Extract, Transform, Load (ETL) process consist of steps which are _____.
Independent and Interrelated 131

Independent or Interrelated
Dependent and Interrelated
Dependent or Interrelated
8. In _____ system, the contents change with time.

OLTP pg 20
DSS
ATM
OLAP
9. _____ is an application of intelligence and experience.

Skill
Power
Wisdom pg 11
Knowledge
10. 3NF removes even more data redundancy than 2NF but it is at the cost of

Simplicity and Performance pg 48
Complexity
Number of tables
Relations
11. Collapsing tables can be done on the _____ relationship(s)

Only One-to-One
Only Many-to-Many
Only One-to-Many
Both One-to-One and Many-to-Many pg 52
12. Transactional fact tables do not have records for events that do not occur. These are called

Not Recording Facts pg 120
Fact-less Facts
Null Facts
Empty Facts
13. Semantically "Dirty Data" class of anomalies includes which of the following:

I) Lexical Errors
II) Integrity Constraints Violation
III) Business Rule Contradiction
IV) Irregularities
V) Duplication
(I) and (II) only
(I), (II), and (III)
(II), (III), and (V) only pg 160
(I), (IV), and (V) only
14. Relational databases allow you to navigate the data in _____ that is appropriate using the primary, foreign key structure within the data model.

Solved By Dr. Tariq Hanif

Prepared by: Dr. Tariq Hanif MHT-4th VU-Sahiwal Campus
Email at:qirathanif@gmail.com
Cell# 0303-7300008

Only One Direction
Any Direction pg 19
Two Directions
Partitions
15. One major goal of horizontal splitting is

Splitting rows for exploiting parallelism pg 54
Splitting columns for exploiting parallelism
Splitting schema for exploiting parallelism
Splitting relationships for exploiting parallelism
16. MOLAP usually builds "cubes" in proprietary file format of a multi-dimensional database (MDD) or a user defined data structure, therefore _____ is not supported.

ANSI pg 78
Microsoft
Oracle
SAP
17. A company has implemented data warehouse for analytical purpose. Quantity sold is stored as a fact. This quantity sold is

Additive Fact 115
Non-Additive Fact
Associative Fact
Non-Associative Fact
18. Typically a data mart is much smaller to data warehouse and it is pretty easy to take its _____ as compare to data warehouse.

Backup pg 131
Cube
Load
Schema
19. "Change Data Capture" is one of the challenging technical issues in _____

Data Extraction pg 149
Data Loading
Data Transformation
Data Cleansing
20. Within the data warehousing domain, data _____ is applied especially when several databases are merged.

Extraction
Loading
Cleansing pg 168
Join

CS614 Data Warehousing

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

Distributed Systems Architecture
Legacy Systems Architecture pg 14

Solved By Dr. Tariq Hanif

Prepared by: Dr. Tariq Hanif MHT-4th VU-Sahiwal Campus
Email at:qirathanif@gmail.com
Cell# 0303-7300008

Online Systems Architecture
Intranet Systems Architecture
2. Suppose the amount of data recorded in an organization is doubled every year. This increase is

Linear
Quadratic
Logarithmic
Exponential pg 15
3. The most common use of range partitioning in data warehouse is on

Date pg 66
Most redundant column
Fact
Dimensions
4. OLAP is a (n) _____ of application.

Classification pg 74
Amalgamation
Unification
Blending
5. ER is a _____ design technique that seeks to remove the redundancy in data.

Logical pg 98
Physical
Data Dependent
Transaction Dependent
6. _____ is the lowest level of detail or the atomic level of data stored in the warehouse.

Cube
Grain pg 111
Virtual Cube
Aggregate
7. It is called a _____ violation, if we have null values for attributes where NOT NULL constraint exists.

Load
Transform
Constraint pg 161
Extraction
8. In the Information Age, the _____ learning organization is at a distinct disadvantage. This term means "impaired functioning."

Functional
Dysfunctional pg 181
Purposeful
Serviceable
9. In _____ system, the contents change with time.

OLTP pg 20
DSS
ATM

Solved By Dr. Tariq Hanif

Prepared by: Dr. Tariq Hanif MHT-4th VU-Sahiwal Campus
Email at:qirathanif@gmail.com
Cell# 0303-7300008

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

Doubles pg 15
Triples
Quartiles
Remains same as previous year
11. Normalization is the process of efficiently organizing data in a database by _____ a relational table into smaller tables by projection.

Composing
Joining / Merging
Combining
Decomposing pg 41
12. 3NF removes even more data redundancy than 2NF but it is at the cost of

Simplicity and Performance 48
Complexity
Number of tables
Relations
13. Which statement is true for De-Normalization?

Redundant data is a performance liability at query time, but is a performance benefit at update time.
Redundant data is a performance benefit at both query time and update time.
Redundant data is a performance liability at both query time and update time.

Redundant data is a performance benefit at query time, but is a performance liability at update time. 51
14. The goal of star schema design is to simplify _____

Logical data model
Physical data model pg 107
Conceptual data model
Semantic data model
15. Source systems for extraction are typically OLTP systems. Extraction is a very complex task due to reasons:

1. Very complex and poorly documented source system.
2. Data has to be extracted not once but many times.
3. People extracting data have limited expertise.

Which of the following option represents correct reason?
1 & 2 only pg 132
1 & 3 only
2 & 3 only
All 1, 2 and 3
16. When tables are populated for the first time, it is a full data refresh. This may be called as:

1. Block Insert pg
2. Block Slamming

Solved By Dr. Tariq Hanif

Prepared by: Dr. Tariq Hanif MHT-4th VU-Sahiwal Campus
Email at:qirathanif@gmail.com
Cell# 0303-7300008

3. Bulk Insert
4. Bulk Slamming

Which of the following option is true?
Option 1 & 3

Option 1 & 2 139

Option 1 & 4
Option 1, 2 & 3
17. The TQM philosophy of management is _____. All members of a total quality management organization strive to systematically manage the improvement of the organization through the ongoing participation of all employees in problem solving efforts across functional and hierarchical boundaries.

Customer-Oriented pg 182
Employee-Oriented
Employer-Oriented
Organization-Oriented
18. Identify the correct option. One Petabyte (PB) equals to ___

2^{52} or 10^{13} bytes
2^{50} or 10^{15} bytes pg 15
2^{50} or 10^{10} bytes
2^{48} or 10^{12} bytes
19. Pre-computed _____ can solve performance problems

Aggregates pg 111
Facts
Dimensions
Primary Keys
20. Single value attributes during recording of a transaction are _____

Dimensions pg 115
Facts
Aggregates
Constraints

CS614 Data Warehousing

1. Development of data warehouse is hard because data sources are

Unstructured & Heterogeneous 31
Structured & Heterogeneous
Unstructured & Homogeneous
Structured and Homogeneous
2. The confusion created by data redundancy makes it difficult for companies to

Create online processing capabilities.
Work in batch processing load.
Use a distributed database.
Integrate data from different sources.
3. Select the statement which is true for Insurance Data Warehouse

Solved By Dr. Tariq Hanif

Prepared by: Dr. Tariq Hanif MHT-4th VU-Sahiwal Campus
Email at:qirathanif@gmail.com
Cell# 0303-7300008

It has Long Operational Business Cycle 36

It has Long Development & Implementation Cycle
It has Short Operational Business Cycle
It has Short Development & Implementation Cycle
4. Redundancy causes anomalies which are called

Selection Anomalies

Update Anomalies 43

SQL Anomalies
Data Warehouse Anomalies
5. 3NF removes even more data redundancy than 2NF but it is at the cost of

Simplicity and Performance pg 48

Complexity
Number of tables
Relations
6. Which statement is true for De-Normalization?

Redundant data is a performance liability at query time, but is a performance benefit at update time.
Redundant data is a performance benefit at both query time and update time.
Redundant data is a performance liability at both query time and update time.

Redundant data is a performance benefit at query time, but is a performance liability at update time.51

7. Pre-join technique is used to avoid

Run time join pg 58

Compile time join
Load time join
8. OLAP is used for analytical process. For analytical processing we need

Multi-level aggregates 74

Record level access
Data level access
Row level access
9. The cube clause which is a part of SQL: 1999 is

GROUP BY CUBE (V_1, V_2 .... V_n) pg 90

SELECT BY CUBE (V_1, V_2 .... V_n)
JOIN BY CUBE (V_1, V_2 .... V_n)
None of these
10. ER is a logical design technique that seeks to remove the _____ in data.

Redundancy pg 98

Normalization
Anomalies
11. Non recording facts have a disadvantage that it has

Solved By Dr. Tariq Hanif

Prepared by: Dr. Tariq Hanif MIT-4th VU-Sahiwal Campus
Email at:qirathanif@gmail.com
Cell# 0303-7300008

Lack of Information 120

Redundant Information
Repeated Information
Normalized Information

12. Once the data has been transformed and ready to be loaded in to data warehouse, we adopt one of two prevalent _____ strategies.

Loading 139

Transformation
Quality
Indexing

13. Syntactically Dirty Data class of anomalies includes which of the following:

1. Lexical Errors
2. Integrity Constraints Violation
3. Business Rule Contradiction
4. Irregularities
5. Duplication

Option 1 and 4 pg 160

Option 2 and 3
Option 2, 3, and 5
Option 1, 4, and 5

14. Records referring to the same entity are represented in different formats in the different data sets or are represented erroneously. Thus, duplicate records will appear in the merged database. The issue is to identify and eliminate these duplicates. The problem is known as the _____ .

Merge/Purge Problem pg 168

Cleansing Problem
Transformation Problem
Data Quality Problem

15. since this form is useful for longitudinal comparisons illustrating trends of continuous improvement. Many traditional data quality metrics, such as free-of-error, completeness, and consistency take this form. This statement is about which of the following:

Simple Ratio pg 187

Min Operation
Max Operation
Weighted Average

16. To handle dimensions that require the aggregation of multiple data quality indicators, which of the following operation can be applied

Minimum or Maximum pg 188

Complex Ratio
Aggregate Average

17. Companies collect and record their own operational data, but at the same time they also use reference data obtained from _____ sources such as codes, prices etc.

None of these
Operational

Solved By Dr Tariq Hanif

Prepared by: Dr. Tariq Hanif MHT-4th VU-Sahiwal Campus
Email at:qirathanif@gmail.com
Cell# 0303-7300008

Internal
External pg 21
18. Source systems for extraction are typically OLTP systems. Extraction is a very complex task due to reasons:

1. Very complex and poorly documented source system.
2. Data has to extracted not once but many times
3. People extracting data have limited expertise

Which of the following option represents correct reason?
1 & 2 only pg 132
1 & 3 only
2 & 3 only
All 1, 2 and 3
19. _____ is about taking/collecting data from different heterogeneous sources.

Data Warehouse pg 21
Data Mart
Data Mining
20. In ROLAP access to information is provided via relational database using _____ standard SQL.

ANSI pg 78
Microsoft
Oracle
SAP

CS614 Data Warehousing

1. A typical example of the crisis in credibility in the naturally evolving architecture is the decision of CEO based on politics and personalities on receiving two different reports for the same query. We say CEO is
Very Subjective and Non-Scientific pg 14
Very Objective and Non-Scientific
Very Subjective and Scientific
Very Objective and Scientific
2. Development of data warehouse is hard because data sources are

Unstructured & Heterogeneous 31
Structured & Heterogeneous
Unstructured & Homogeneous
Structured and Homogeneous
3. Financial data warehouses have some severe drawbacks that are not found elsewhere. For example it is almost impossible to reconcile down to the rupee. This is because of many reasons. Select the statement which shows the possible reason(s).

The accounting periods may be different in different operational systems or the classifications of regions may change pg 35
The accounting periods may be different in Data Warehouse application
Data warehouse uses dynamic classifications of regions
During aggregation data warehouse neglect amount in rupees
4. Redundancy causes anomalies which are called

Solved By Dr. Tariq Hanif

Prepared by: Dr. Tariq Hanif MHT-4th VU-Sahiwal Campus
Email at:qirathanif@gmail.com
Cell# 0303-7300008

Selection Anomalies
Update Anomalies pg 43
SQL Anomalies
Data Warehouse Anomalies
5. Normalization is the process of efficiently organizing data in a database by decomposing / splitting a relational table into _____ tables by projection.

Smaller pg 41
Larger
Combined
Joined
6. One major goal of horizontal splitting is

Splitting rows for exploiting parallelism pg 54
Splitting columns for exploiting parallelism
Splitting schema for exploiting parallelism
7. The most common use of range partitioning in data warehouse is on

Date pg 66
Most redundant column
Fact
Dimensions
8. ER Model can be simplified in -------- ways

One
Two pg 103
Three
Four
9. _____ is the lowest level of detail or the atomic level of data stored in the warehouse.

Cube
Grain pg 111
Virtual Cube
Aggregate
10. A company has implemented data warehouse for analytical purpose. Quantity sold is stored as a fact. This quantity sold is

Additive Fact 119
Non-Additive Fact
11. Fact-less fact table is a fact table without numeric fact columns. It is used to capture relationship between _____

Dimensions pg 121
Attributes
Tables
Facts
12. Full and Incremental extraction techniques are types of _____

Logical Extraction pg 133
Physical Extraction
Both Logical and Physical Extraction

Solved By Dr. Tariq Hanif

Prepared by: Dr. Tariq Hanif MHT-4th VU-Sahiwal Campus
Email at:qirathanif@gmail.com
Cell# 0303-7300008

None of these

13. Rearranging the grouping of source data, delivering it to the destination database, and ensuring the quality of data are crucial to the process of loading the data warehouse. Data _____ is vitally important to the overall health of a warehouse project.

1. Cleansing
2. Cleaning
3. Scrubbing

Which of the following options is true?

Option 1 only pg 158

Option 2 only

Option 1 & 2 only

Option 1, 2 & 3

14. Syntactically Dirty Data class of anomalies includes which of the following:

6. **Lexical Errors**
7. Integrity Constraints Violation
8. Business Rule Contradiction
9. **Irregularities**
10. Duplication

Option 1 and 4 pg 160

Option 2 and 3

Option 2, 3, and 5

Option 1, 4, and 5

15. It is called a _____ violation, if we have null values for attributes where NOT NULL constraint exists.

Load

Transform

Constraint

Extraction

16. As consumers, human beings judge the quality of things during their life-time.

I	**Consciously**
II	**Subconsciously**
III	Unconsciously

Which of the following statement is true?

I Only

II Only

III Only

I & II Only pg 179

17. All data is _____ of something real.

I	**An Abstraction**
II	A Representation

Which of the following option is true?

I Only pg 180

II Only

Both I & II

Solved By Dr. Tariq Hanif

Prepared by: Dr. Tariq Hanif MHT-4th VU-Sahiwal Campus
Email at:qirathanif@gmail.com
Cell# 0303-7300008

None of I & II

18. _____queries deal with number of variables spanning across number of tables (i.e. join operations) and looking at lots of historical data.

OLTP
DBMS
DSS pg 21
None of these

19. Collapsing tables can be done on the _____ relationships
Many-to-Many
Both One-to-One and Many-to-Many pg 52
None of these
One-to-One

20. In data warehouse, a query results in retrieval of hundreds of records from very large table. The ratio of number of records retrieved to total number of record present is high and selectivity is

Low
High pg 22
Average
Can not be calculated

CS614 Data Warehousing

1. _____ is one class of decision support environment.

OLAP pg 30

OLTP

Data Cleansing

ETL

2. The confusion created by data redundancy makes it difficult for companies to

Create online processing capabilities.

Work in batch processing load.

Use a distributed database.

Integrate data from different sources.

3. Effects of de-normalization on database performance are

Unpredictable 62
Predictable

Conventional

Unsurprising

4. OLAP is a (n) _____ of application.

Solved By Dr. Tariq Hanif

Prepared by: Dr. Tariq Hanif MHT-4th VU-Sahiwal Campus
Email at:qirathanif@gmail.com
Cell# 0303-7300008

Classification pg 74

Amalgamation

Unification

Blending

5. DOLAP model facilitates _____ computing paradigm.

Mobile pg 97

Permanent

Rigid

Strict

6. _____ is the lowest level of detail or the atomic level of data stored in the warehouse.

Cube

Grain pg 111

Virtual Cube

Aggregate

7. Extract, Transform, Load (ETL) process consist of steps which are _____.

Independent and Interrelated 131

Independent or Interrelated

Dependent and Interrelated

Dependent or Interrelated

8. In _____ system, the contents change with time.

OLTP pg 20

DSS

ATM

OLAP

9. _____ is an application of intelligence and experience.

Skill

Power

Wisdom

Knowledge pg 11

Solved By Dr. Tariq Hanif

Prepared by: Dr. Tariq Hanif MHT-4th VU-Sahiwal Campus
Email at:qirathanif@gmail.com
Cell# 0303-7300008

10. 3NF removes even more data redundancy than 2NF but it is at the cost of

Simplicity and Performance pg 48

Complexity

Number of tables

Relations

11. Collapsing tables can be done on the _____ relationship(s)

Only One-to-One

Only Many-to-Many

Only One-to-Many

Both One-to-One and Many-to-Many pg 52

12. Transactional fact tables do not have records for events that do not occur. These are called

Not Recording Facts pg 120

Fact-less Facts

Null Facts

Empty Facts

13. Semantically "Dirty Data" class of anomalies includes which of the following:

I) Lexical Errors

II) Integrity Constraints Violation

III) Business Rule Contradiction

IV) Irregularities

V) Duplication

(I) and (II) only

(I), (II), and (III)

(II), (III), and (V) only 160

(I), (IV), and (V) only

14. 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

Solved By Dr. Tariq Hanif

Prepared by: Dr. Tariq Hanif MIT-4th VU-Sahiwal Campus
Email at:qirathanif@gmail.com
Cell# 0303-7300008

Any Direction pg 19

Two Direction

Partitions

15. One major goal of horizontal splitting is

Splitting rows for exploiting parallelism 54

Splitting columns for exploiting parallelism

Splitting schema for exploiting parallelism

Splitting relationships for exploiting parallelism

16. MOLAP usually builds "cubes" in proprietary file format of a multi-dimensional database (MDD) or a user defined data structure, therefore _____ is not supported.

ANSI pg 78

Microsoft

Oracle

SAP

17. A company has implemented data warehouse for analytical purpose. Quantity sold is stored as a fact. This quantity sold is

Additive Fact 113

Non-Additive Fact

Associative Fact

Non-Associative Fact

18. Typically a data mart is much smaller to data warehouse and it is pretty easy to take its _____ as compare to data warehouse.

Backup

Cube pg 131

Load

Schema

19. "Change Data Capture" is one of the challenging technical issues in _____

Data Extraction pg 149

Data Loading

Data Transformation

Data Cleansing

Solved By Dr. Tariq Hanif

Prepared by: Dr. Tariq Hanif MHT-4th VU-Sahiwal Campus
Email at:qirathanif@gmail.com
Cell# 0303-7300008

20. Within the data warehousing domain, data _____ is applied especially when several databases are merged.

Extraction

Loading

Cleansing pg 168

Join

CS614 Data Warehousing

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

Distributed Systems Architecture

Legacy Systems Architecture pg 14

Online Systems Architecture

Intranet Systems Architecture

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

Linear

Quadratic

Logarithmic

Exponential 15

3. The most common use of range partitioning in data warehouse is on

Date pg 66

Most redundant column

Fact

Dimensions

4. OLAP is a (n) _____ of application.

Classification pg 74

Amalgamation

Unification

Blending

5. ER is a _____ design technique that seeks to remove the redundancy in data.

Logical pg 98

Solved By Dr. Tariq Hanif

Prepared by: Dr. Tariq Hanif MHT-4th VU-Sahiwal Campus
Email at:qirathanif@gmail.com
Cell# 0303-7300008

Physical

Data Dependent

Transaction Dependent

6. _____ is the lowest level of detail or the atomic level of data stored in the warehouse.

Cube

Grain pg 111

Virtual Cube

Aggregate

7. It is called a _____ violation, if we have null values for attributes where NOT NULL constraint exists.

Load

Transform

Constraint pg 161

Extraction

8. In the Information Age, the _____ learning organization is at a distinct disadvantage. This term means "impaired functioning."

Functional

Dysfunctional pg 181

Purposeful

Serviceable

9. In _____ system, the contents change with time.

OLTP pg 20

DSS

ATM

OLAP

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

Doubles pg 15

Triples

Quartiles

Remains same as previous year

Solved By Dr. Tariq Hanif

Prepared by: Dr. Tariq Hanif MIT-4th VU-Sahiwal Campus
Email at:qirathanif@gmail.com
Cell# 0303-7300008

11. Normalization is the process of efficiently organizing data in a database by _____ a relational table into smaller tables by projection.

Composing

Joining / Merging

Combining

Decomposing pg 41

12. 3NF removes even more data redundancy than 2NF but it is at the cost of

Simplicity and Performance pg 48

Complexity

Number of tables

Relations

13. Which statement is true for De-Normalization?

Redundant data is a performance liability at query time, but is a performance benefit at update time.

Redundant data is a performance benefit at both query time and update time.

Redundant data is a performance liability at both query time and update time.

Redundant data is a performance benefit at query time, but is a performance liability at update time. 51

14. The goal of star schema design is to simplify _____

Logical data model

Physical data model pg 107

Conceptual data model

Semantic data model

15. Source systems for extraction are typically OLTP systems. Extraction is a very complex task due to reasons:

1. Very complex and poorly documented source system.
2. Data has to be extracted not once but many times.
3. People extracting data have limited expertise.

Which of the following option represents correct reason?

1 & 2 only pg 132

1 & 3 only

Solved By Dr. Tariq Hanif

Prepared by: Dr. Tariq Hanif MHT-4th VU-Sahiwal Campus
Email at:qirathanif@gmail.com
Cell# 0303-7300008

2 & 3 only

All 1, 2 and 3

16. When tables are populated for the first time, it is a full data refresh. This may be called as:

1. Block Insert
2. Block Slamming
3. Bulk Insert
4. Bulk Slamming

Which of the following option is true?

Option 1 & 3

Option 1 & 2

Option 1 & 4 pg 139

Option 1, 2 & 3

17. The TQM philosophy of management is _____. All members of a total quality management organization strive to systematically manage the improvement of the organization through the ongoing participation of all employees in problem solving efforts across functional and hierarchical boundaries.

Customer-Oriented pg 182

Employee-Oriented

Employer-Oriented

Organization-Oriented

18. Identify the correct option. One Petabyte (PB) equals to ___

2^{52} or 10^{13} bytes

2^{50} or 10^{15} bytes pg 15

2^{50} or 10^{10} bytes

2^{48} or 10^{12} bytes

19. Pre-computed _____ can solve performance problems

Aggregates pg 111

Facts

Dimensions

Primary Keys

20. Single value attributes during recording of a transaction are _____

Solved By Dr. Tariq Hanif

Prepared by: Dr. Tariq Hanif MHT-4th VU-Sahiwal Campus
Email at:qirathanif@gmail.com
Cell# 0303-7300008

Dimensions pg 115

Facts

Aggregates

Constraints

CS614 Data Warehousing

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

Linear
Quadratic
Logarithmic
Exponential pg 15
2. _____ is one class of decision support environment.

OLAP pg 30
OLTP
Data Cleansing
ETL
3. De-Normalization normally speeds up

Data Retrieval pg 51
Data Modification
Development Cycle
Data Replication
4. In horizontal splitting, we split a relation into multiple tables on the basis of
Common Column Values
Common Row Values
Different Index Values
Value resulted by ad-hoc query
5. The most common use of range partitioning in data warehouse is on

Date pg 66
Most redundant column
Fact
Dimensions
6. OLAP is a (n) _____ of application.

Blending
Characterization pg 74
Amalgamation
Unification
7. One of the OLAP characteristics is Multi-dimensional, which is _____ for OLAP.

Essential 76
Optional
Discretionary
Not Obligatory

Solved By Dr. Tariq Hanif

Prepared by: Dr. Tariq Hanif MHT-4th VU-Sahiwal Campus
Email at:qirathanif@gmail.com
Cell# 0303-7300008

8. Non recording facts have a disadvantage that it has

Lack of Information pg 120
Redundant Information
Repeated Information
Normalized Information
9. During ETL process of an organization, suppose you have data which can be transformed using any of the transformation method. Which of the following strategy will be your choice for least complexity?

One-to-One Scalar Transformation 144
One-to-Many Element Transformation
Many-to-Many Element Transformation
Many-to-One Element Transformation
10. All data is _____ of something real.

I An Abstraction pg 180
II A Representation
Which of the following option is true?

I Only
II Only
Both I & II
None of I & II
11. _____ is an application of information and data.

Skill
Knowledge pg 11
Intelligence
Power
12. In data warehouse, a query results in retrieval of hundreds of records from very large table. The ratio of number of records retrieved to total number of records present is high and selectivity is:

Low

High 22
Average
Non computable
13. "The environment is smart enough to develop or compute higher level aggregates using lower level or more detailed aggregates". Which of the following approach is described by the above statement?

Aggregate awareness pg 87
Cube partitioning
Indexing
MOLAP cube aggregation
14. The goal of star schema design is to simplify _____

Logical data model

Solved By Dr. Tariq Hanif

Prepared by: Dr. Tariq Hanif MHT-4th VU-Sahiwal Campus
Email at:qirathanif@gmail.com
Cell# 0303-7300008

Physical data model pg 107
Conceptual data model
Semantic data model
15. Syntactically "Dirty Data" class of anomalies includes _____

I) Lexical Errors
II) Integrity Constraints Violation
III) Business Rule Contradiction
IV) Irregularities
V) Duplication

(I) and (IV) only 160
(II) and (III) only
(II), (III), and (IV) only
(I), (IV), and (V) only
16. Experience showed that for a single pass of magnetic tape that scanned 100% of the records, only _____ of the records, sometimes were actually required.

5% pg 12
30%
50%
80%
17. Pre-computed _____ can solve performance problems

Aggregates pg 111
Facts
Dimensions
Primary Keys
18. Single value attributes during recording of a transaction are _____

Dimensions pg 115
Facts
Aggregates
Constraints
19. In full extraction, data is extracted completely from the source system. Therefore there is no need to keep track of changes to the _____

Data Source pg 133
DWH
Data Mart
Data Destination
20. Within the data warehousing domain, data _____ is applied especially when several databases are merged.

Extraction

Loading
Cleansing pg 168
Join

Solved By Dr. Tariq Hanif
Remember me in Prayers
The End

Prepared by: Dr. Tariq Hanif MHT-4th VU-Sahiwal Campus
Email at:qirathanif@gmail.com
Cell# 0303-7300008