

Paper 1:

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

Answer (select correct option)

Linear

Quadratic

Logarithmic

Exponential page 15

2. ----- is the lowest level of detail or the atomic level of data stored in the warehouse.

Answer (select correct option)

Cube

Grain page 111

Virtual cube

Aggregate

3. After implementing change data capture, the advantage we have is that, data is able to be integrated and transformed -----

Answer (select correct option)

In-flight page 152

Off-flight

Stored data

Over-flight

4. ----- If fits into memory, costs only one disk I/O access to locate a record by given key.

Answer (select correct option)

A Dense index page 223

A Sparse index

An inverted index

None of these

5. A dense index , if fits into memory , costs only ----- disk I/O access to locate a record by given key

Answer (select correct option)

One page 223

Two

$Ig(n)$

n

6. ----- if too big and does not fit into memory , will be expensive when used to find a record by given key.

Answer (select correct option)

A dense index page 223

A Sparse index

An inverted index

None of these

7. All data is ----- of something real.

1. An Abstraction

2. A Representation

Which of the following options is true?

Answer (select correct option)

1 **Only page 180**

3 only

2 only

Both 1 &2

None of 1 &2

8. In the information age, the ----- learning organization is at a distinct disadvantage. This term means “impaired functioning.”

Answer (select correct option)

Functional

Dysfunctional page 181

Purposeful

Serviceable

9. Many DW projects do not deliver to full potential because they treat data quality as a one-time undertaking as part of UAT. Here UAT stands for

Answer (select correct option)

User Acceptance testing page 193

Uninterrupted availability of testing

Universal acceptance test

Universal applied test

10. NUMA stands for -----

Answer (select correct option)

Non-uniform memory access page 206

Non-updateable memory architecture

New universal memory architecture

11. Parallelisms can ----- system performance on over-utilized systems or systems with small I/O bandwidth.

Answer (select correct option)

Reduce page 202

Enhance

Maintain

Boost

12. There are many variants of the traditional nested-loop join. If the index is built as part of the query plan and subsequently dropped, it is called

Answer (select correct option)

Naïve nested-loop joins

Index nested-loop joins

Temporary index nested-loop joins page 243

None of these

13. Data mining evolve as a mechanism to cater the limitations of ----- systems to deal massive data sets with high dimensionality, new data types, multiple heterogeneous data resources etc.

Answer (select correct option)

OLTP page 254

OLAP

DSS

DWH

14. In contrast to data mining, statistics is ----- driven.

Answer (select correct option)

Assumption

Knowledge

Discovery page 255

Database

15. ----- is the technique in which existing heterogeneous segments are reshuffled, relocated into homogeneous segments.

Answer (select correct option)

Clustering page 264

Aggregation
Segmentation
Partitioning

16. A ----- Implementation approach is generally useful for projects where the technology is mature and well understood, as well as where the business problems that must be solved are clear and well understood.

Answer (select correct option)

Top down page 283

Bottom up
Waterfall
Spiral

17. Implementing a data warehouse requires ----- integrated activities.

Answer (select correct option)

Slackly
Lethargically
Tightly page 289
Loosely

18. The Kimball's iterative data warehouse development approach drew on decades of experience to develop the -----.

Answer (select correct option)

Business definition lifecycle
Business dimensional lifecycle page 289
Data warehouse dimension
OLAP dimension

19. In DWH project, it is assured that ----- environment is similar to the production environment

Answer (select correct option)

Designing
Development page 314

Analysis
Implementation

20. Pipeline parallelism focuses on increasing throughput of task execution, NOT on ----- sub-task execution time.

Answer (select correct option)

Increasing

Decreasing page 215

Maintaining

None of these

21. The application development quality-assurance activities cannot be completed until the data is -----.

Answer (select correct option)

Stabilized page 308

Identified

Finalized

Computerized

22. One needs to slot the alternative tools into categories that allow for meaningful comparison in order to -----

Answer (select correct option)

Evaluate tools page 315

Reduce cost

None of these

23. Focusing on data warehouse delivery only often end up -----.

Answer (select correct option)

Rebuilding page 315

Success

Good stable product

None of these

24. Invest years in architecture and forgetting the primary purpose of solving business problem, results in inefficient application. This is the example of ----- mistake.

Answer (select correct option)

Extreme technology design

Extreme architecture design

None of these page 315

25. Pakistan is one of the five major -----countries in the world.

Answer (select correct option)

Cotton-growing page 330

Rice-growing

Weapon producing

26. ----- Division is cotton hub of Punjab.

Answer (select correct option)

Lahore

Faisalabad

Multan page 331

Bahawalpur

27. ----- is a systematic field sampling process that provide field specific information on pest pressure and crop injury.

Answer (select correct option)

Pest scouting page 333

Soil survey

Seed survey

Water survey

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

Answer (select correct option)

1950s.

1960s. page 12

1970s.

1980s.

29. Redundancy causes ----- anomalies

Answer (select correct option)

Update page 43

Select

Both update & select

None of these

30. Modification anomalies occur in ----- form

Answer (select correct option)

2nf page 45

3nf

BCNF

1NF

31. If we remove the modification anomalies the table comes in ----- form

Answer (select correct option)

1NF

2NF

3NF page 47

4NF

32. Data transformation services (DTS) provide a set of ---- that lets you extract, transform, and consolidate data from disparate sources into single or multiple destinations supported by DTS connectivity.

Answer (select correct option)

Tools page 372

Documentations

Guidelines

33. Data transformation services (DTS) provide a set of tools that lets you extract, transform, and consolidate data from disparate sources into ----- supported by DTS connectivity.

Answer (select correct option)

Single destinations

Multiple destinations

Single or multiple destinations page 372

34. Execution can be completed successfully or it may be stopped due to some error. In case of successful completion of execution all the transactions will be -----

Answer (select correct option)

Committed to the database page 419

Rolled back

35. Execution can be completed successfully or it may be stopped due to some error. If some error occurs, execution will be terminated abnormally and all transactions will be -----

Answer (select correct option)

Committed to the database

Rolled back page 419

36. To identify the degree of transformation required we need to perform -----.

Answer (select correct option)

Data profiling page 437

Data anomaly detection

Data cleansing

None of the given

37. To identify the ----- required we need to perform data profiling.

Answer (select correct option)

Degree of transformation page 437

Complexity

Cost

Time

38. To judge effectiveness we perform data profiling twice.

Answer (select correct option)

One before extraction and the other after extraction

One before transformation and the other after transformation page 441

One before loading and the other after loading

39. If the dates are missing we must need to consult -----.

Answer (select correct option)

Golden copy page 457

Default system data

Silver copy

None of the given

40. In ROLAP access to information is provided via relational database using ----- standard SQL.

Answer (select correct option)

ANSI page 78

Microsoft
Oracle
SAP

Paper 2:

1. Which of the following statement is true? 1 PB is

Answer (Select your correct option)

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

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

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

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

2. Normalized design is likely to perform much faster than de-normalized design for queries that probe.

Answer (Select your correct option)

Master table only may be

Detail tables only

Both

3. Partition elimination is not possible with

Answer (Select your correct option)

Round-Robin page 66

De-normalization

Normalization

4. ER is a logical design technique that seeks to remove the ---- in data.

Answer (Select your correct option)

Redundancy page 99

Normalization

Anomalies

5. Merging information is one of the major types of -----.

Answer (Select your correct option)

Transformation page 154

Extraction

Loading

None of these

6. The goal of ----- is to look at as few blocks as possible to find the matching records(s).

Answer (Select your correct option)

Indexing page 222

Partitioning

Joining

7. If every key in the data file is represented in the index file then index is

Answer (Select your correct option)

Dense index page 223

Sparse index

Inverted index

None of these

8. -----, if fits into memory, costs only one disk I/O access to locate a record by given key.

Answer (Select your correct option)

A dense index page 223

A sparse index

An inverted index

None of these

9. A dense index, if fits into memory, costs only ----- disk I/O access to locate a record by given key.

Answer (Select your correct option)

One page 223

Two

$Lg(n)$

n

10. ---- means meeting customer`s needs, not necessarily exceeding them.

Answer (Select your correct option)

Quality page 180

Marketing

DSS

OLAP

11. The purpose of the House of Quality technique Is to reduce ---- types of risk.

Answer (Select your correct option)

Two page 195

Three

Four

All

12. Parallelisms can ----- system performance on over-utilized systems or systems with small I/O bandwidth.

Answer (Select your correct option)

Reduce page 202

Enhance

Maintain

Boost

13. Data dependencies between different phases of computation introduce synchronization requirements that force ---- execution.

Answer (Select your correct option)

Sequential page 204

Parallel

Analogous

Interactive parallel

14. There are many variants of the traditional nested-loop join. When the entire table is scanned it is called

Answer (Select your correct option)

Naïve nested-loop joins page 243

Index nested-loop joins

Temporary index nested-loop joins

None of these

15. Data mining is a/an ---- approach, where browsing through data using data mining techniques may reveal something that might be of interest to the user as information that was unknown previously.

Answer (Select your correct option)

Exploratory page 249

Non-exploratory

Computer science

16. Data mining uses ----- algorithms to discover patterns and regularities in data.

Answer (Select your correct option)

Statistical page 251

Mathematical

Computational

17. As opposed to the outcome of classification, estimation deal with ---- valued outcome.

Answer (Select your correct option)

Continuous page 260

Discrete

Isolated

Distinct

18. ----- is the technique in which existing heterogeneous segments are reshuffled, relocated into homogeneous segments.

Answer (Select your correct option)

Clustering page 264

Aggregation

Segmentation

Partitioning

19. Majority of data warehouse projects fail due to the complexity of the -----.

Answer (Select your correct option)

Development process page 283

Analytical process of cube

Query complexity

Indexing complexity

20. Implementing a data warehouse requires ----- integrated activities.

Answer (Select your correct option)

Tightly page 289

Loosely

Slackly

Lethargically

21. In DWH project, it is assured that ----- environment is similar to the production environment.

Answer (Select your correct option)

Designing

Development page 314

Analysis

Implementation

22. For a DWH, the key requirement are ----- and product experience.

Answer (Select your correct option)

Tools

Industry page 320

Software

None of these

23. The application development quality-assurance activities cannot be completed until the data is -----.

Answer (Select your correct option)

Stabilized page 308

Identified

Finalized

Computerized

24. Many data warehouse project teams waste enormous amount of time searching in vain for a -----.

Answer (Select your correct option)

Silver Bullet page 315

Golden bullet

Suitable hardware

Compatible product

25. One needs to slot the alternative tools into categories that allow for meaningful comparison in order to -----.

Answer (Select your correct option)

Evaluate tools page 315

Reduce cost

None of these

26. Investing years in architecture and forgetting the primary purpose of solving business problems, results in inefficient application. This is the example of ---- mistake.

Answer (Select your correct option)

Extreme technology design

Extreme architecture design

None of these (page # 303)

27. Pakistan is one of the five major ---- countries in the world.

Answer (Select your correct option)

Cotton-growing page 330

Rice-growing

Weapon producing

28. -----division is cotton hub of Punjab.

Answer (Select your correct option)

Lahore

Faisalabad

Multan page 331

Bahawalpur

29. -----is a process which involves gathering of information about column through execution of certain queries with intention to identify erroneous records.

Answer (Select your correct option)

Data profiling page 439

Data anomaly detection

Record duplicate detection

None of these

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

Answer (Select your correct option)

Only one direction

Any direction page 19

Two directions

None of these

31. In ----- system, the contents change with time.

Answer (Select your correct option)

OLTP page 20

DSS

ATM

OLAP

32. Primary key is repeated in ----- splitting.

Answer (Select your correct option)

Horizontal

Vertical page not sure

33. Data transformation services (DTS) provide a set of tools that lets you extract, transform, and consolidate data from disparate sources into ----- supported by DTS connectivity.

Answer (Select your correct option)

Single destination

Multiple destinations

Single or multiple destinations page 372

34. If some error occurs, execution will be terminated abnormally and all transactions will be rolled back. In this case when we will access the database we will find it in the state that was before the -----.

Answer (Select your correct option)

Execution of package page 419

Creation of package

Connection of package

35. Execution can be completed successfully or it may be stopped due to some error. If some error occurs, execution will be terminated abnormally and all transactions will be ----.

Answer (Select your correct option)

Committed to the database

Rolled back page 419

36. To identify the degree of transformation required we need to perform ----.

Answer (Select your correct option)

Data profiling page 437

Data anomaly detection
Data cleansing
None of the given

37. To identify the ----- required we need to perform data profiling.

Answer (Select your correct option)

Degree of transformation page 437

Complexity
Cost
Time

38. To judge effectiveness we perform data profiling twice.

Answer (Select your correct option)

One before extraction and the other after extraction

One before transformation and the other after transformation page 441

One before loading and the other after loading

39. If the dates are missing we must need to consult -----.

Answer (Select your correct option)

Golden copy page 457

Default system date
Silver copy
None of the given

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

Answer (Select your correct option)

Doubles page 15

Triples
Quartiles
Remains same as previous year

Q1: The automated, prospective analyses offered by data mining move beyond the analysis of past events provided by respective tools typical of _____.

- OLTP
- OLAP
- **Decision Support systems**
- None of these

Q2: As opposed to the outcome of classification, estimation deal with _____ valued outcome.

- Discrete
- Isolated
- **Continuous**
- Distinct

Q3: The goal of ideal parallel execution is to completely parallelize those parts of a computation that are not constrained by data dependencies. The _____ the portion of the program that must be executed sequentially, the greater the scalability of computation.

- Larger
- **Smaller**
- Unambiguous
- Superior

Q5: The goal of _____ is to look at as few block as possible to find the matching records.

- **Indexing**
- Partitioning
- Joining
- none of these

Q6: There are many variants of the traditional nested-loop join, if there is an index is exploited, and then it is called.....

- Naïve nested loop join
- **index nested loop join**
- temporary index nested loop join
- none of these

Q7: The technique that is used to perform these feats in data mining modeling and this act of model building is something that people have been doing for long time, certainly before the _____ of computers or data mining technology.

- Access
- **Advent** <http://www.netcoach.eu.com/index.php?id=188>
- Ascent
- Avowal

Spring 2010

Question No: 1

A data warehouse may include

► **Legacy systems**

- Only internal data sources
- Privacy restrictions
- Small data mart

Question No: 2

De-Normalization normally speeds up

► **Data Retrieval (page no 51)**

- Data Modification
- Development Cycle
- Data Replication

Question No: 3

In horizontal splitting, we split a relation into multiple tables on the basis of

► **Common Column Values page no 54**

- Common Row Values
- Different Index Values
- Value resulted by ad-hoc query

Question No: 4

Multidimensional databases typically use proprietary _____ format to store pre-summarized cube structures.

► **File page no 79**

- Application
- Aggregate
- Database

Question No: 7

The key idea behind _____ is to take a big task and break it into subtasks that can be processed concurrently on a stream of data inputs in multiple, overlapping stages of execution.

► **Pipeline Parallelism page no 214**

- Overlapped Parallelism
- Massive Parallelism
- Distributed Parallelism

Question No: 8

Non uniform distribution, when the data is distributed across the processors, is called

► **Skew in Partition (P # 218)**

- ▶ Pipeline Distribution
- ▶ Distributed Distribution
- ▶ Uncontrolled Distribution

Question No: 9

The goal of ideal parallel execution is to completely parallelize those parts of a computation that are not constrained by data dependencies. The smaller the portion of the program that must be executed _____, the greater the scalability of the computation.

- ▶ None of these

_____ ▶ **Sequentially page no 204**

- ▶ in Parallel
- ▶ Distributed

Question No: 10 (Marks: 1) - Please choose one

If 'M' rows from table-A match the conditions in the query then table-B is accessed 'M' times. Suppose table-B has an index on the join column. If 'a' I/Os are required to read the data block for each scan and 'b' I/Os for each data block then the total cost of accessing table-B is _____ logical I/Os approximately.

_____ ▶ **$(a + b)M$**

- ▶ $(a - b)M$
- ▶ $(a + b + M)$
- ▶ $(a * b * M)$

Question No: 14

To measure or quantify the similarity or dissimilarity, different techniques are available. Which of the following option represent the name of available techniques?

- ▶ Pearson correlation is the only technique
- ▶ Euclidean distance is the only technique

_____ ▶ **Both Pearson correlation and Euclidean distance page no 270**

- ▶ None of these

Question No: 15

For a given data set, to get a global view in un-supervised learning we use

_____ ▶ **One-way Clustering (P# 271)**

- ▶ Bi-clustering
- ▶ Pearson correlation
- ▶ Euclidean distance

Question No: 17 (Marks: 1) - Please choose one

For a DWH project, the key requirement are _____ and product experience.

- ▶ Tools
- ▶ Industry (P# 320)
- ▶ Software
- ▶ **None of these**

Question No: 24

DSS queries do not involve a primary key

▶ **True (p # 21)**

- ▶ False

Question No: 25

_____ contributes to an under-utilization of valuable and expensive historical data, and inevitably results in a limited capability to provide decision support and analysis.

- ▶ The lack of data integration and standardization (P# 330)
- ▶ Missing Data
- ▶ Data Stored in Heterogeneous Sources

Question No: 26

DTS allows us to connect through any data source or destination that is supported by _____

▶ **OLE DB (p # 373)**

- ▶ OLAP
- ▶ OLTP
- ▶ Data Warehouse

Question No: 29

If some error occurs, execution will be terminated abnormally and all transactions will be rolled back. In this case when we will access the database we will find it in the state that was before the _____.

▶ **Execution of package (p # 419)**

- ▶ Creation of package
- ▶ Connection of package

Spring 2009

Question No: 3

Pre-computed _____ can solve performance problems

Aggregates (page # 111)

Facts

Dimensions

Question No: 5

the degree of similarity between two records, often measured by a numerical value between _____, usually depends on application characteristics.

0 and 1 (page # 169)

0 and 10

0 and 100

0 and 99

Question No: 8

which are the least appropriate join operation for Pipeline parallelism?

Hash Join

Inner Join

Outer Join

Sort-Merge Join page 217

Question No: 10

Data mining derives its name from the similarities between searching for valuable

business information in a large database, for example, finding linked products in

gigabytes of store scanner data, and mining a mountain for a _____ of valuable ore.

Furrow

Streak

Trough

Vein

Question No: 11

With data mining, the best way to accomplish this is by setting aside some of your data in a _____ to isolate it from the mining process; once the mining is

complete, the results can be tested against the isolated data to confirm the model's validity.

Cell

Disk

Folder

Vault

Question No: 12

The Kimball s iterative data warehouse development approach drew on decades

of experience to develop the _____.

Business Dimensional Lifecycle (page # 289)

Data Warehouse Dimension

Business Definition Lifecycle

OLAP Dimension

Question No: 13

we must try to find the one access tool that will handle all the needs of their users.

True

False [p # 315]

They believe their mission is to find the one access tool that will handle all the needs of all their users. Don't even try. It can't be done. One size does not fit all.

Question No: 14

for a smooth DWH implementation we must be a technologist.

True

False (page # 306)

Question No: 15

during the application specification activity, we also must give consideration to

the organization of the applications.

True (page # 307)

False

Question No: 17

the most recent attack is the _____ attack on the cotton crop during 2003-04, resulting in a loss of nearly 0.5 million bales.

Boll Worm (VIDO LECTURE # 38)

Purple Worm

Blue Worm

Cotton Worm

Question No: 18

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

Decision maker (page# 10)

Manager

Database Administrator

DWH Analyst

Question No: 19

_____ breaks a table into multiple tables based upon common column values.

Horizontal splitting (page # 46)

Vertical splitting

Question # 1

During the development phase we should follow standards for:

Naming Conventions

Calculation

Libraries

All of the given options (Page 307)

Question # 2

As per Kimball's approach the second phase of DWH life cycle is:

Project Planning

Requirement definition (Page 299)

Requirement verification

Requirement validation

Question # 3

Bill Inmon argues that requirements are well understood only after

They are documented

They extracted and verified

Data warehouse is populated (Page 285)

All schemes are defined

Question # 4

Waterfall model is appropriate when

When budget is low

When the deadline is strict
When resources are limited

Requirements are clearly defined (Page 284)

Question # 5

A bottom up implementation approach is useful when

Technology is matured and well understands (page 283)

Organization cannot implement latest technologies

The business objectives are very much clear

All requirements are well documented

Question # 6

Technical architecture design specifies the:

Project schedule (Page 300)

Minimum project completion time

Required Components

Umbrella activities

Question # 7

_____ says that as far as company goals are concerned, the entire company pursues in the same direction

Westerman (Page 285)

Bill inmon

Kimball

Bohnlein

Question # 8

As per Bill Inmon, a data warehouse, in contrast with classical applications is:

Data driven (Page 285)

Resource driven

Requirement driven

Time Sensitive

Question # 9

The _____ phase of fundamental data warehouse life cycle model includes data warehouse scheduled maintenance

Deployment

Operation (Page 288)

Enhancement
Maintenance

Question # 10

Following the coding standards during development helps to:

Minimize the future rework (Page 307)

Verify the requirements
To refine the project scope
To establish the test cases

Question # 1

As per Kimball, _____ is the main operational process

Requirement extraction
Global design

Business process (Page 285)

Scheme design

Question

In _____ phase of a fundamental data warehouse life cycle model, a working model of data warehouse is deployed for a selective set of users

Design

Prototype (Page 287)

Deployment
Operation

Question # 3

Waterfall is a/an _____ model

Iterative

Simple linear sequential (Page 284)

Object oriented
Rapid development

Question # 4

One of the drawbacks of waterfall model is that:

Customers can not review the product during development (Page 284)

It does not work when the resources are limited
It does not define the project timeline/schedule

All of the given options

Question # 5 of 10 (Total Marks: 1) Select correct option:

Spiral model is _____

Sequence of waterfall model

Risk oriented model

An iterative model

All of the given (Page 284)

Question # 6 of 10 (Total Marks: 1) Select correct option:

_____ refers to the overall process of discovering useful knowledge from data and data mining refers to a particular step in this process.

Statistics

Knowledge discovery in database (Page 249)

Clustering

Information cleansing

Question # 7 of 10 (Total Marks: 1) Select correct option:

In data mining, initially you _____ what you are looking for.

Know

Don't Know (Page 250)

May or may not know

None of the given options

Question # 8 of 10 (Total Marks: 1) Select correct option:

In context of the most fundamental data warehouse life cycle model, which of the following is NOT one of the data warehouse design activities?

End-user interviews and re-interviews

Source system cataloguing

Definition of key performance indicators

System vision development (Page 287)

Question # 9 of 10 (Total Marks: 1) Select correct option:

A top down implementation approach is useful when

Technology is matured and well understands (Page 283)

Organization cannot implement latest technologies

The business objectives are very much clear

All requirements are well documented

Question # 10 of 10 (Total Marks: 1) Select correct option:

Which of the following is NOT one of the methodologies for Data Warehouse project development?

Goal Driven

Data Driven

User Driven

System Driven (Page 283)

Question # 1 of 10 (Total Marks: 1) Select correct option:

In context of data mining definition, the term "value" means

The primary key of table

Index location of the record

Importance of hidden patterns discovered (Page 250)

Numerical or string measure assigned to an attribute

Question # 2 of 10 (Total Marks: 1) Select correct option:

Data mining is all about

Knowledge discovery in database

Finding hidden patterns in data

Finding hidden relations in data

All of the given options (Page 249)

Question # 3 of 10 (Total Marks: 1) Select correct option:

In context of clustering, the term "distance" means

Similarity/dissimilarity of records (Page 272)

The difference between the primary keys of two records

The relation of a record with corresponding record in child table

None of the given options

Question # 4 of 10 (Total Marks: 1) Select correct option:

We should follow a proper _____ cycle to implement a change even if it is smaller one.

Development -> QA -> Production (Page 314)

Production -> QA -> Development

Development -> Production -> QA

Production->Development -> QA

Question # 5 of 10 (Total Marks: 1) Select correct option:

For smooth DWH project implementation, one of the recommendations is to have:

Full time project manager assigned to the project (Page 318)

QA team ready for testing the code

Design of complete project before implementation

Components of complete project before implementation

Question # 6 of 10 (Total Marks: 1) Select correct option:

Normally it is recommended to have

Different servers for development and production environment

Same server for development and production environment

Interference while having different database environments on a single server (Page 313)

All of the given options

Question # 7 of 10 (Total Marks: 1) Select correct option:

Identify the TRUE statement:

Clustering is unsupervised learning and classification is supervised learning (Page 270)

Clustering is supervised learning and classification is unsupervised learning

Both clustering and classification are unsupervised learning

Both clustering and classification are supervised learning

Question # 8 of 10 (Total Marks: 1) Select correct option:

Mining multi-dimensional databases allow users to:

Categorize the data

Analyze the data

Summarize the data

All of the given options (Page 250)

Question # 1 of 10 (Total Marks: 1) Select correct option:

Technical architecture design supports the communication about technical requirements: I) within the team II) Upward to management III) Outward to vendors

(I) only

(II) Only

(III) Only

(I), (II) and (III) (Page 300)

Question # 2 of 10 (Total Marks: 1) Select correct option:

Improper documentation results the problem(s) like:

Maintenance issue

New developers unable to configure already existing code (Page 314)

Lot of time required for enhancing the code

All of the given options

Question # 3 of 10 (Total Marks: 1) Select correct option:

In Four Cell Quadrant Technique, The quadrant's vertical axis refers to:

Scope

Feasibility

Resources available

The potential impact or value to the business (Page 297)

Question # 4 of 10 (Total Marks: 1) Select correct option:

Which of the following activity executes parallel with all other activities in Kimball's DWH development approach?

Requirement elicitation

Project Planning

Project management (Page 289)

Deployment

Question # 5 of 10 (Total Marks: 1) Select correct option:

"What means what". The phrase refers to:

Meta data (Page 338)

External data

Transformed data

Internal representations

Question # 6 of 10 (Total Marks: 1) Select correct option:

Which of the following is NOT one of the top-10 mistakes that should be avoided during DWH development?

Not interacting directly with end user

Not being an accommodating person (Page 316)

Isolating IT support people from business users

Training the users with dummy data and considering it success

Question # 7 of 10 (Total Marks: 1) Select correct option:

In _____ phase of kimballs approach, we identify the components needed now and in future.

Requirement definition

Architectural design (Page 300)

Product development

Analytical application development

Question # 8 of 10 (Total Marks: 1) Select correct option:

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

Question # 9 of 10 (Total Marks: 1) Select correct option:

_____ technique requires a separate column to specify the time and date when the last modification was occurred.

CheckMarks

TimeStamps (Page 150)

Just-In-Time

RealTime extraction

Question # 10 of 10 (Total Marks: 1) Select correct option:

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

Syntactically dirty data (Page 160)

Semantically dirty data

Coverage anomaly

Extraction issue

Question # 1 of 10 (Total Marks: 1) Select correct option:

In Extract, Load, Transform (ELT) process, data transformation _____

Takes place on the data warehouse server

Takes place on a separate transformation server (Page 147)

Depends on the nature of the source database

Does not take place

Question # 2 of 10 (Total Marks: 1) Select correct option:

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

Question # 3 of 10 (Total Marks: 1) Select correct option:

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

Conversion

Summarization

Enrichment

Full Data Refresh (Page 135)

Question # 4 of 10 (Total Marks: 1) Select correct option:

Robotic libraries are needed for _____

Cubes

Data marts

Data warehouse (Page 131)

Aggregates

Question # 5 of 10 (Total Marks: 1) Select correct option:

Change Data Capture (CDC) can be challenging task because:

Aggregates don't change in real time

Transformation of extracted data is difficult (Page 149)

Identifying the recently modified data may be difficult

Source systems may not support extraction of changed aggregates

Question # 6 of 10 (Total Marks: 1) Select correct option:

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 (Page 143)

Prioritization

Element Selection

Measurement Selection

Question # 7 of 10 (Total Marks: 1) Select correct option:

In context of Change Data Capture (CDC), sometimes a _____ object can be used to store the recently modified data:

Buffer table

Change table (Page 149)

Checkmark table

Change control table

Question # 8 of 10 (Total Marks: 1) Select correct option:

Which of the following is NOT one of the examples of dynamic attributes?

Daily Sale

Date of Birth cnfim not (Page 342)

Air pressure

None of the given options

Question # 9 of 10 (Total Marks: 1) Select correct option:

In context of requirement definition phase of Kimball's DWH development approach, _____ is positioned as a findings review and prioritization meeting.

System analysis

Scope definition

Requirement configuration

Requirements wrap-up presentation (Page 297)

Question # 10 of 10 (Total Marks: 1) Select correct option:

Which of the following activity/activities is/are part of project planning phase in Kimball's DWH development approach?

Obtain resources

Establish the preliminary scope and justification

Assess organization's readiness for a data warehouse initiative

All of the given options (Page 290)

Question # 1 of 10 (Total Marks: 1) Select correct option:

Goal driven approach of data warehouse development was result of _____ work

Bill Inmon

Ralph Kimball

Question # 2 of 10 (Total Marks: 1) Select correct option:

In contrast to data mining, statistics is _____ driven.

Assumption

Knowledge

human

Database

(Page 255)

Question # 3 of 10 (Total Marks: 1) Select correct option:

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

This increase is _____ .

Linear

Quadratic

Exponential

Logarithmic

(Page 15)

Question # 4 of 10 (Total Marks: 1) Select correct option:

In context of requirement definition phase in Kimball's DWH development approach, activities like debriefing, documentation, and prioritization are considered to be the part of:

Requirement preplanning

Business requirements collection

Post collection

None of the given options

(Page 294)

Question # 5 of 10 (Total Marks: 1) Select correct option:

Users do not care, how advance the front end of your DWH is, what they care is that:

Tables should be properly de-normalized

Proper partitioning technique should be used

At least star or snow flake schema should be implementing

They should get information in timely manner

(Page 314)

Question # 6 of 10 (Total Marks: 1) Select correct option:

Which of the following is NOT one of the activities of "Maintenance and Growth" phase in Kimball's DWH development approach?

Education

Technical education

Program Support

Interface Deployment (Page 309)

Question # 7 of 10 (Total Marks: 1) Select correct option:

Which of the following is/are example(s) of static attributes?

Employee Name (Page 342)

Employee Date of Birth

Employee Blood Group

All of the given options

Question # 8 of 10 (Total Marks: 1) Select correct option:

Three parallel tracks in kimbal's DWH development approach converge at _____ Phase

Project planning

Business requirement definition

Deployment (Page 308)

Maintenance and growth

Question # 9 of 10 (Total Marks: 1) Select correct option:

Which of the following is not a technique of "De-Normalization" ?

Pre-joining

Splitting Tables

Adding Redundant Columns

ER Modeling (Page 52)

Question # 10 of 10 (Total Marks: 1) Select correct option:

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

Timestamps

Partitioning

Triggers

Dimensional Modeling (Page 150)

Question # 1 of 10 (Total Marks: 1) Select correct option:

Normally the term "DWH face to the business user" refers to:

Lifecycles technology track

Lifecycle Data track

Lifecycle Analytical Applications track

(Page 306)

Lifecycle Maintenance track

Question # 2 of 10 (Total Marks: 1) Select correct option:

An effective user education program includes, among other, the following guideline(s):

Understand the audience, don't overwhelm

Train after delivery of the data and analytic applications

Postpone education, if DWH not ready

All of the given options

(Page 308)

Question # 3 of 10 (Total Marks: 1) Select correct option:

In context of analytical applications track, the application development activity can begin after

Database Design is complete

(Page 307)

A subset of historical data has been loaded

The data access tools and metadata are installed

All of the given options

QuestionNo:1(Marks:1)-Pleasechooseo

ne

To identify the _____ required we need to perform data profiling

▶ **Degree of Transformation (Page437)**

▶ Complexity

▶ Cost

▶ Time

QuestionNo:2(Marks:1)-Pleasechooseo

ne

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.

▶ Operational

▶ None of these

▶ Internal

▶ **External (Page21)**

QuestionNo:3(Marks:1)-Pleasechooseo

ne

Ad-hoc access means to run such queries which are known already.

► True

► **False** (Page 18)

QuestionNo:4(Marks:1)-Pleasedchooseo

ne

_____ in agriculture extension is that pest population beyond which the benefit of spraying outweighs its cost.

► Profit Threshold Level

► **Economic Threshold Level** (Page 332)

► Medicine Threshold Level

► None of these

QuestionNo:5(Marks:1)-Pleasedchooseon

e

People that design and build the data warehouse must be capable of working across the organization at all levels

► **True** (Page 320)

► False

QuestionNo:6(Marks:1)-Pleasedchooseon

e

The _____ is only a small part in realizing the true business value buried within the mountain of data collected and stored within organizations business systems and operational databases.

► Independence on technology

► **Dependence on technology** (Page 320)

► None of these

QuestionNo:7(Marks:1)-Pleasedchooseon

e

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** (Page 19)

- Two Direction
- None of these

QuestionNo:8(Marks:1)-Pleasechooseone

DSS queries do not involve a primary key

► **True** (Page 21)

- False

QuestionNo:10(Marks:1)-Pleasechooseone

The need to synchronize data upon update is called

- Data Manipulation
- Data Replication
- **Data Coherency** (Page 12)
- Data Imitation

QuestionNo:11(Marks:1)-Pleasechooseone

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

- Distributed Systems

Architecture

► **Legacy Systems Architecture** (Page 14)

- Online Systems

Architecture

- Intranet Systems

Architecture

QuestionNo:12(Marks:1)-Pleasechooseone

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

- $O(n)$

- ▶ $O(n^2)$
- ▶ $O(n \lg n)$
- ▶ **$O(\lg n)$ (Page 22)**

Question No: 13 (Marks: 1) - Please choose one

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.**
- (Page 51)**

Question No: 14 (Marks: 1) - Please choose one

As opposed to the outcome of classification, estimation deal with _____ valued outcome.

- ▶ Discrete
- ▶ Isolated
- ▶ **Continuous (Page 260)**
- ▶ Distinct

Question No: 15 (Marks: 1) - Please choose one

The goal of _____ is to look at as few block as possible to find the matching records.

- ▶ **Indexing (Page 222)**
- ▶ Partitioning
- ▶ Joining
- ▶ none of these

Question No: 16 (Marks: 1) - Please choose one

The technique that is used to perform these feats in data mining modeling, and this act of model building is something that people have been doing for long time, certainly before the _____ of computers or data mining technology.

► **Access Advent** [click for detail](#)

► Ascent Avowal

QuestionNo:17(Marks:1)-Pleasechoose one

De-Normalization normally speeds up

► **Data Retrieval** (Page51)

► Data Modification

► Development Cycle

► Data Replication

QuestionNo:18(Marks:1)-Pleasechoose one

In horizontal splitting, we split a relation into multiple tables on the basis of

► **Common Column Values** (Page54)

► Common Row Values

► Different Index Values

► Value resulted by ad-hoc query

QuestionNo:19(Marks:1)-Pleasechoose one

For good decision making, data should be integrated across the organization to cross the LoB (Line of Business). This is to give the total view of organization from:

► Owner"s Perspective

► **Customer's Perspective** (Page 16)

► Decision Maker"s

Perspective

► Employee's Perspective

Question No: 20 (Marks: 1) - Please choose

one

The automated, prospective analyses offered by data mining move beyond the analysis of past events provided by respective tools typical of _____.

- ▶ OLTP
- ▶ OLAP
- ▶ **Decision Support systems**
- ▶ None of these

Question No : 24 (Marks: 1) - Please choose one

DSS queries do not involve a primary key

- ▶ **True (Page21)**
- ▶ False

Question No : 25 (Marks: 1) - Please choose one

_____ contributes to an under-utilization of valuable and expensive historical data, and inevitably results in a limited capability to provide decision support and analysis.

- ▶ **The lack of data integration and standardization (Page 330)**
- ▶ Missing Data
- ▶ Data Stored in Heterogeneous Sources

Question No : 26 (Marks: 1) - Please choose one

DTS allows us to connect through any data source or destination that is supported by _____

- ▶ **OLE DB (Page373)**
- ▶ OLAP
- ▶ OLTP
- ▶ Data Warehouse

Question No : 29 (Marks: 1) - Please choose one

If some error occurs, execution will be terminated abnormally and all transactions will be rolled back. In this case when we will access the database we will find it in the state that was before the _____.

- ▶ **Execution of package (Page419)**

- ▶ Creation of package
- ▶ Connection of package

Question No: 17 (Marks: 1) - Please choose one

The most recent attack is the _____ attack on the cotton crop during 2003- 04, resulting in a loss of nearly 0.5 million bales.

▶ **Boll Worm** (Page 333)

- ▶ Purple Worm
- ▶ Blue Worm
- ▶ Cotton Worm

Question No: 18 (Marks: 1) - Please choose one

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

▶ **Decision maker** (Page 18)

- ▶ Manager
- ▶ Database Administrator
- ▶ DWH Analyst

Question # 1 of 10 (Total Marks: 1)

With data mining, the best way to accomplish this is by setting aside some of your data in a vault to isolate it from the mining process; once the mining is complete, the results can be tested against the isolated data to confirm the model's _____.

▶ **Validity** [Click here for detail](#)

- ▶ Security
- ▶ Integrity
- ▶ None of above

Question # 2 of 10 (Total Marks: 1)

Classification consists of examining the properties of a newly presented observation and assigning it to a predefined _____.

- ▶ Object
- ▶ Container
- ▶ Subject
- ▶ **Class** (Page 259)

Question # 3 of 10 (Total Marks: 1)

During business hours, most _____ systems should probably not use parallel execution.

- ▶ OLAP
- ▶ DSS
- ▶ Data Mining
- ▶ **OLTP (Page 206)**

Question # 4 of 10 (Total Marks: 1)

In contrast to statistics, data mining is _____ driven.

- ▶ Assumption
- ▶ **Knowledge (Page 254)**
- ▶ Human
- ▶ Database

Question # 5 of 10 (Total Marks: 1)

As opposed to the outcome of classification, estimation deal with _____ valued outcome.

- ▶ Discrete
- ▶ Isolated
- ▶ **Continuous (Page 260)**
- ▶ Distinct

Question # 7 of 10 (Total Marks: 1)

The goal of _____ is to look at as few blocks as possible to find the matching records(s).

- ▶ **Indexing (Page 222)**
- ▶ Partitioning
- ▶ Joining
- ▶ None of above

Question # 10 of 10 (Total Marks: 1)

An optimized structure which is built primarily for retrieval, with update being only a secondary consideration is

▶ OLTP

▶ OLAP

▶ DSS

▶ **inverted Index** (Page 232)

Question # 1 of 10 (Total Marks: 1)

Which of the following is not an "Orr's Law of Data Quality"?

- ▶ Data that is not used cannot be correct!"
- ▶ 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!"**

Question # 2 of 10 (Total Marks: 1)

(Page 181)

_____, if too big and does not fit into memory, will be expensive when used to find a record by given key.

- ▶ An Inverted Index
- ▶ A Sparse Index
- ▶ **A Dense Index** (Page 223)
- ▶ None of these

Question # 3 of 10 (Total Marks: 1)

Which of the following is not an activity of Data Quality Analysis Project?

- ▶ "Define"
- ▶ Measure"
- ▶ Analyze"
- ▶ **"Compression"** (Page 194)

Question # 4 of 10 (Total Marks: 1)

The automated, prospective analyses offered by data mining move beyond the analyses of past events provided by retrospective tools typical of _____ .

- ▶ OLTP
- ▶ OLAP
- ▶ **Decision Support Systems** [Click here for detail](#)
- ▶ None of these

Question # 5 of 10 (Total Marks: 1)

Data mining, the _____ of hidden predictive information from large databases, is a powerful new technology with great potential to help companies focus on the most important information in their data warehouses.

- ▶ **Extraction** [Click here for detail](#)
- ▶ Insertion
- ▶ Inclusion
- ▶ Enclosure

Question # 6 of 10 (Total Marks: 1)

If someone told you that he had a good model to predict customer usage, the first thing you might try would be to ask him to apply his model to your customer _____, where you already knew the answer.

- ▶ **Base** [Click here for detail](#)

- ▶ Drive
- ▶ File
- ▶ Log

Question # 7 of 10 (Total Marks: 1)

Data mining uses _____ algorithms to discover patterns and regularities in data.

- ▶ Mathematical
- ▶ Computational
- ▶ **Statistical** (Page 251)
- ▶ None of these

Question # 10 of 10 (Total Marks: 1)

A data warehouse implementation without an OLAP tool is always possible.

- ▶ True
- ▶ **False** (Page 69)

Question # 1 of 10 (Total Marks: 1)

_____ modeling technique is more appropriate for data warehouses.

► entity-relationship

► **dimensional** [Click here for detail](#)

► physical

► None of the given

Question # 2 of 10 (Total Marks: 1)

The performance in a MOLAP cube comes from the $O(1)$ look-up time for the array data structure.

► **True** (Page 79)

► False

Question # 3 of 10 (Total Marks: 1)

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

► SQL

► **proprietary file** (Page 79)

► Object oriented

► Non-proprietary file

Question # 4 of 10 (Total Marks: 1)

Slice and Dice is changing the view of the data.

► True

ue

► **False** (Page 81)

Question # 5 of 10 (Total Marks: 1)

Data warehousing and on-line analytical processing (OLAP) are _____ elements of decision support system.

► Unusual

ual

► **Essential** (Page 69)

► Optional

nal

► None of the given

Question # 6 of 10 (Total Marks: 1)

Virtual cube is used to query two similar cubes by creating a third "virtual" cube by a join between two cubes.

► True

ue

► **False** (Page 86)

Question # 7 of 10 (Total Marks: 1)

Analytical processing uses _____, instead of record level access.

► **multi-level aggregates** (Page 74)

► Single-level

aggregates

- ▶ Single-level hierarchy
- ▶ None of the Given

Question # 8 of 10 (Total Marks: 1)

The divide&conquer cube partitioning approach helps alleviate the _____ limitations of MOLAP

- ▶ implementation.
- ▶ Flexibility
- ▶ Maintainability
- ▶ Security
- ▶ **Scalability (Page 85)**

Question # 9 of 10 (Total Marks: 1)

In a traditional MIS system, there is an almost linear sequence of queries.

- ▶ **True (Page 72)**
- ▶ False

Question # 10 of 10 (Total Marks: 1)

Data Warehouse provides the best support for analysis while OLAP carries out the _____ task.

- ▶ Mandatory
- ▶ Whole
- ▶ **Analysis (Page 69)**
- ▶ Prediction

Question # 1 of 10 (Total Marks: 1) Select correct option:

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

Question # 2 of 10 (Total Marks: 1) Select correct option:

In contrast to data mining, statistics is _____ driven.

Assumption (Page 255)

Knowledge
Discovery
Database

Question # 4 of 10 (Total Marks: 1) Select correct option:

Which of the following is not a technique of „De-Normalization“?

Pre-joining
Splitting Tables
Adding Redundant Columns

ER Modeling (Page 54)

Question # 5 of 10 (Total Marks: 1) Select correct option:

Which of the following is an example of Additive Facts?

Sales Amount (Page 119)

Average
Discount
Ratios

2
0

Question # 6 of 10 (Total M a r k s: 1) Select correct option:

Which of the following is not technique of "changed data capture" in currently used modern source system?

Timestamps

Partitioning

Triggers

Dimensional modeling (Page 150)

Question # 7 of 10 (Total M a r k s: 1) Select correct option:

The automated, prospective analyses offered by data mining move beyond the analyses of past events provided

by _____ tools typical of decision support systems.

Introspective

Intuitive

Reminiscent

Retrospective [Click here for detail](#)

Question # 8 of 10 (Total M a r k s: 1) Select correct option:

There are many variants of the traditional nested-loop join. When the entire table is scanned it is called

Index nested-loop join

Naive nested-loop join (Page 243)

Temporary index nested-loop join

None of these

Question # 9 of 10 (Total M a r k s: 1) Select correct option:

The relation R will be in 2nd 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 each non key attribute dependent upon a single column of composite primary key.

it is in 1NF and Primary Key is composite.

Question # 10 of 10 (Total Marks: 1) Select correct option:

Which one of the following is not data quality validation technique?

Using data quality Rules

Data Histograming

Indexes (Page 189)

Question # 1 of 10 (Total Marks: 1) Select correct option:

Which of the following is not a technique of Data Mining?

Estimation

Prediction

Clustering

Normalization (Page 259)

Question # 2 of 10 (Total Marks: 1) Select correct option:

The STAR schema used for data design is a _____ consisting of fact and dimension tables.

► Network odel

► **Relational model (Page 87)**

► Hierarchical data model

► None of the given

Question # 3 of 10 (Total Marks: 1) Select correct option:

DOLAP allows download of "cube" structures to a desktop platform with the need for shared relational or cube server.

► True

► **False (Page 78)**

Question # 4 of 10 (Total Marks: 1) Select correct

option:

There are many variants of the traditional nested-loop join, if there is an index is exploited, then it is called.....

- ▶ Naïve nested loop join
- ▶ **index nested loop join (Page 243)**
- ▶ temporary index nested loop

join

- ▶ none of these

Question # 5 of 10 (Total M a r k s: 1) Select correct

option:

Data mining is a/an _____ approach , where browsing through data using mining techniques may reveal something that might be of interest to the user as information that was unknown previously.

- ▶ Non-Exploratory
- ▶ **Exploratory (Page**

249)

- ▶ Compute Science
- ▶ none of these

Question # 6 of 10 (Total M a r k s: 1) Select correct option:

Pre-join technique is used to avoid

- ▶ **Run time join (Page 58)**
- ▶ Compile time join
- ▶ Load time join

Question # 7 of 10 (Total M a r k s: 1) Select correct option:

Cube is a _____ entity containing values of a certain fact at a certain aggregation level at an intersection of a combination of dimensions.

- ▶ **Logical (Page 88)**
- ▶ Physical
- ▶ Analytical
- ▶ None of these

Question # 8 of 10 (Total M a r k s: 1) Select correct option:

The goal of star schema design is to simplify

-
- ▶ Logical data model
 - ▶ **Physical data model (Page 107)**
 - ▶ Conceptual data model
 - ▶ None of these

Question # 9 of 10 (Total M a r k s: 1) Select correct option:

Grain is the _____ level of data stored in the warehouse.

- ▶ **Atomic (Page 111)**
- ▶ Summarized
- ▶ Aggregated
- ▶ Cube

Question # 10 of 10 (Total M a r k s: 1) Select correct option:

Transactional fact tables do not have records for events that do not occur.

These are called

- ▶ **Not Recording Facts** (Page 120)
- ▶ Fact-less Facts
- ▶ Null Facts
- ▶ None of these

Question # 1 of 10 (Total Marks: 1) Select correct option:

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** (Page 127)
- ▶ Time
- ▶ Parallel
- ▶ None of these

Question # 2 of 10 (Total Marks: 1) Select correct option:

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**
(Page 144)
- ▶ One-to-Many Element Transformation
- ▶ Many-to-Many Element Transformation
- ▶ Many-to-One Element Transformation

Question # 3 of 10 (Total Marks: 1) Select correct option:

Change Data Capture is one of the challenging technical issues in

-
- ▶ **Data Extraction** (Page 149)
 - ▶ Data Loading
 - ▶ Data

Transformation

► Data Cleansing

Question # 4 of 10 (Total M a r k s: 1) Select correct option:

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 (Page 158)**

► Option 2 only

► Option 1 & 2

only

► Option 1, 2 & 3

Question # 5 of 10 (Total M a r k s: 1) Select correct option:

When performing objective assessments, companies follow a set of principles to develop metrics specific to their needs, there is hard to have "one size fits all" approach. Which of the following statement represents the pervasive functional forms?

► **Simple Ratio, Min or Max Operation, Weighted Average (Page 186)**

► Only Complex Ratio, Min Operation, Max Operation

► Only Simple Ratio, Min or Max Operation

► Only Min or Max Operation, Weighted Average

Question # 6 of 10 (Total M a r k s: 1) Select correct option:

The input to the data warehouse can come from OLTP or transactional system but not from other third party database.

► **True (Page 19)**

► False

Question # 7 of 10 (Total M a r k s: 1) Select correct option:

Normalization effects
performance

► **True**

► False

Question # 8 of 10 (Total M a r k s: 1) Select correct option:

Collapsing tables can be done on the _____
relationships

► One-to-One

► Many-to-Many

► **Both One-to-One and Many-to-Many (Page 52)**

► None of these

Question # 9 of 10 (Total M a r k s: 1) Select correct option:

_____ breaks a table into multiple tables based upon common
column values.

► **Horizontal splitting (Page 54)**

► Vertical splitting

Question # 10 of 10 (Total M a r k s: 1) Select correct option:

If w is the window size and n is the size of data set, then the complexity of merging
phase in BSN method
is _____

► $O(n)$

► $O(w)$

► **$O(w n)$ (Page 171)**

► $O(w \log n)$

Question # 1 of 10 (Total M a r k s: 1) Select correct

option:

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

- ▶ Harmonized
- ▶ Identical
- ▶ Homogeneous
- ▶ **Heterogeneous (Page 21)**

Question # 2 of 10 (Total M a r k s: 1) Select correct

option:

_____ is class of Decision Support Environment.

- ▶ OLTP
- ▶ **OLAP (Page 30)**
- ▶ DBMS
- ▶ Network

Question # 3 of 10 (Total M a r k s: 1) Select correct

option:

For a relation to be in 4NF it must be:-

- ▶ In 2NF if and only if it is in first normal form and all nonkey attributes are fully functionally dependent on the key.
- ▶ In 3NF and every non-key column is non-transitively dependent upon its primary key.
- ▶ In 3NF and every non-key column is non-transitively dependent upon its Foreign key.
- ▶ **In 3NF and It does not have multi valued dependencies (Page 48)**

Question # 4 of 10 (Total M a r k s: 1) Select correct

option:

Horizontal splitting breaks a table into multiple tables based upon _____

- ▶ Common Row values
- ▶ Range of Data.
- ▶ Redundant data.
- ▶ **Common column values. (Page 54)**

Question # 5 of 10 (Total M a r k s: 1) Select correct option:

D-Normalization speeds up_____

▶ **Data Retrieval (Page 51)**

- ▶ Data Duplication
- ▶ Data Extraction
- ▶ Data Loading

Question # 6 of 10 (Total M a r k s: 1) Select correct option:

The most common use of range partitioning is on _____.

▶ **Date (Page 66)**

- ▶ Rows
- ▶ DSS
- ▶ None of these

Question # 7 of 10 (Total M a r k s: 1) Select correct option:

The fact table is a way of *visualizing* as an _____cube

- ▶ Rolled Up
- ▶ **Un-rolled (Page 87)**
- ▶ Rolled Down
- ▶ None of these

Question # 8 of 10 (Total M a r k s: 1) Select correct option:

High Granularity, more _____.

▶ **Details** [Click here for detail](#)

- ▶ Dimensions
- ▶ Aggregates
- ▶ Cubes

Question # 9 of 10 (Total M a r k s: 1) Select correct option:

Transactional fact tables do not have records for events that do not occur.

These are called

► **Not Recording Facts (Page 120)**

- Fact-less Facts
- Null Facts
- None of these

Question # 10 of 10 (Total M a r k s: 1) Select correct option:

Cubes not drilled down is, It is a _____.

- **Data Mart**
- DWH
- DOLAP
- None of these

Question # 1 of 10 (Total M a r k s: 1) Select correct option:

Full & Incremental Extraction are the types of _____
Extraction

► **Logical (Page 132)**

- Physical
- Both Logical & Physical
- None of Above

Question # 2 of 10 (Total M a r k s: 1) Select correct option:

Selectivity is low in _____
environment.

► **DWH (Page 22)**

- DBMS
- OLTP
- None of Above

Question # 3 of 10 (Total M a r k s: 1) Select correct option:

Normalization effects performance

- True
- **False**

Question # 4 of 10 (Total M a r k s: 1) Select correct option:

MOLAP physically builds "cubes" for direct access in a multi-dimensional database (MDD) Therefore _____ is not supported.

- ▶ One-to-One
- ▶ Facts
- ▶ **ANSI SQL (Page 78)**
- ▶ Dimensions
- ▶ None of these

Question # 5 of 10 (Total M a r k s: 1) Select correct option:

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

- ▶ **Decision maker (Page 18)**
- ▶ Manager
- ▶ Database Administrator
- ▶ DWH Analyst

Question # 6 of 10 (Total M a r k s: 1) Select correct option:

In _____ system, the contents change with time.

Select correct option:

- ▶ **OLTP (Page 20)**
- ▶ ATM
- ▶ DSS
- ▶ OLAP

Question # 7 of 10 (Total M a r k s: 1) Select correct option:

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

Select correct option:

- ▶ 1950s.
- ▶ **1960s. (Page 12)**
- ▶ 1970s.

► 1980s.

Question # 8 of 10 (Total M a r k s: 1) Select correct

option:

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

► **Relaxed (Page 14)**

► Good

► Not Relaxed

► None

Question # 9 of 10 (Total M a r k s: 1) Select correct

option:

_____ gives total view of an organization

► OLAP

► OLTP

► **Data Warehouse (Page 16)**

► Database

Question # 10 of 10 (Total M a r k s: 1) Select correct

option:

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

► Linear

► Quadratic

► **Exponential (Page 15)**

► logarithmic

Question # 1 of 10 (Total M a r k s: 1) Select correct

option:

_____ is an application of information and data.

► **Knowledge (Page 11)**

► Intelligence

► Power

► Education

Question # 2 of 10 (Total M a r k s: 1) Select correct option:

A single database, couldn't serve both operational high performance transaction processing and DSS, analytical processing, all at the same time.

▶ **True (Page 13)**

▶ False

Question # 3 of 10 (Total M a r k s: 1) Select correct option:

B-Tree is used as an index to provide access to records

▶ **Without scanning the entire table (Page 22)**

▶ By scanning the entire meta data

▶ By scanning the entire table

▶ None of these

Question # 4 of 10 (Total M a r k s: 1) Select correct option:

in agriculture extension is that pest population beyond which the benefit of spraying outweighs levels

▶ **True (Page 332)**

▶ False

