



CS614 - Data Warehousing

Solved MCQS
From Final term Papers

Feb 25, 2013

MC100401285

Moaaz.pk@gmail.com

Mc100401285@gmail.com

PSMD01

FINAL TERM EXAMINATION

Spring 2012

CS614- Data Warehousing

Question No: 1 (Marks: 1) – Please choose one

To identify the _____ required we need to perform data profiling

► **Degree of Transformation** (Page 437)

► Complexity

► Cost

► Time

Question No: 2 (Marks: 1) – Please choose one

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** (Page 21)

Question No: 3 (Marks: 1) – Please choose one

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

► True

► **False** (Page 18)

Question No: 4 (Marks: 1) – Please choose one

_____ 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

خدا کے سوا کسی سے امید مت رکھو

Muhammad Moaaz Siddiq – MCS(4th)

Moaaz.pk@gmail.com

Campus:- Institute of E-Learning & Modern Studies
(IEMS) Samundari

Question No: 5 (Marks: 1) – Please choose one

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

► **True** (Page 320)

► False

Question No: 6 (Marks: 1) – Please choose one

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

Question No: 7 (Marks: 1) – Please choose one

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

Question No: 8 (Marks: 1) – Please choose one

DSS queries do not involve a primary key

► **True** (Page 21)

► False

Question No: 9 (Marks: 1) – Please choose one

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.

► **Tools** (Page 373)

► Documentations

► Guidelines

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

The need to synchronize data upon update is called

► Data Manipulation

► Data Replication

► **Data Coherency** (Page 12)

► Data Imitation

دنیا میں سب سے مشکل کام اپنی اصلاح اور سب سے آسان کام دوسروں پر نکتہ چینی کرنا ہے

Question No: 11 (Marks: 1) – Please choose one

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

Question No: 12 (Marks: 1) – Please choose one

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

- ▶ O (n)
- ▶ O (n²)
- ▶ 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 apposed to the out come 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

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

De-Normalization normally speeds up

▶ **Data Retrieval** (Page 51)

- ▶ Data Modification
- ▶ Development Cycle
- ▶ Data Replication

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

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

▶ **Common Column Values** (Page 54)

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

Question No: 19 (Marks: 1) – Please choose 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

اللہ کا خوف سب سے بڑی دانائی ہے

FINAL TERM EXAMINATION
Spring 2010
CS614- Data Warehousing

Question No: 1 (Marks: 1) – Please choose one

A data warehouse may include

- ▶ **Legacy systems** (Page 135)
- ▶ Only internal data sources
- ▶ Privacy restrictions
- ▶ Small data mart

Question No: 2 (Marks: 1) – Please choose one

De-Normalization normally speeds up

- ▶ **Data Retrieval** (Page 51)
- ▶ Data Modification
- ▶ Development Cycle
- ▶ Data Replication

Question No: 3 (Marks: 1) – Please choose one

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

- ▶ **Common Column Values** (Page 54)
- ▶ Common Row Values
- ▶ Different Index Values
- ▶ Value resulted by ad-hoc query

Question No: 4 (Marks: 1) – Please choose one

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

- ▶ **File** (Page 79)
- ▶ Application
- ▶ Aggregate
- ▶ Database

Question No: 5 (Marks: 1) – Please choose one

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

- ▶ **One** (Page 223)
- ▶ Two
- ▶ $\lg(n)$
- ▶ n

ایماندار کو غصہ دیر سے آتا ہے اور جلدی دور ہو جاتا ہے

Question No: 6 (Marks: 1) – Please choose one

All data is _____ of something real.

I An Abstraction

II A Representation

Which of the following option is true?

▶ **I Only** (Page 180)

▶ II Only

▶ Both I & II

▶ None of I & II

Question No: 7 (Marks: 1) – Please choose one

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 214)

▶ Overlapped Parallelism

▶ Massive Parallelism

▶ Distributed Parallelism

Question No: 8 (Marks: 1) – Please choose one

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

▶ **Skew in Partition** (Page 218)

▶ Pipeline Distribution

▶ Distributed Distribution

▶ Uncontrolled Distribution

Question No: 9 (Marks: 1) – Please choose one

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 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: 11 (Marks: 1) – Please choose one

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.

▶ **Exploratory** (Page 249)

- ▶ Non-Exploratory
- ▶ Computer Science

Question No: 12 (Marks: 1) – Please choose one

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.

▶ **OLTP** (Page 254)

- ▶ OLAP
- ▶ DSS
- ▶ DWH

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

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

▶ **Clustering** (Page 264)

- ▶ Aggregation
- ▶ Segmentation
- ▶ Partitioning

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

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 270)
- ▶ None of these

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

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

▶ **One-way Clustering** (Page 271)

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

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

In DWH project, it is assured that _____ environment is similar to the production environment

- ▶ Designing
- ▶ **Development** (Page 314)
- ▶ Analysis
- ▶ Implementation

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

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

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

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

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

- ▶ Increasing
- ▶ **Decreasing** (Page 215)
- ▶ Maintaining
- ▶ None of these

Question No: 19 (Marks: 1) – Please choose one

Many data warehouse project teams waste enormous amounts of time searching in vain for a

- ▶ **Silver Bullet** (Page 315)
- ▶ Golden Bullet
- ▶ Suitable Hardware
- ▶ Compatible Product

Question No: 20 (Marks: 1) – Please choose one

Focusing on data warehouse delivery only often end up _____.

- ▶ **Rebuilding** (Page 315)
- ▶ Success
- ▶ Good Stable Product
- ▶ None of these

Question No: 21 (Marks: 1) – Please choose one

Pakistan is one of the five major _____ countries in the world.

- ▶ **Cotton-growing** (Page 330)
- ▶ Rice-growing
- ▶ Weapon Producing

Question No: 22 (Marks: 1) – Please choose one

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

- ▶ **Data profiling** (Page 439)
- ▶ Data Anomaly Detection
- ▶ Record Duplicate Detection
- ▶ None of these

Question No: 23 (Marks: 1) – Please choose one

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

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

DSS queries do not involve a primary key

- ▶ **True (Page 21)**
- ▶ 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 (Page 373)**
- ▶ OLAP
- ▶ OLTP
- ▶ Data Warehouse

Question No: 27 (Marks: 1) – Please choose one

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.

- ▶ **Tools (Page 373) rep**
- ▶ Documentations
- ▶ Guidelines

Question No: 28 (Marks: 1) – Please choose one

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 _____

- ▶ **Committed to the database (Page 419)**
- ▶ Rolled back

بری صحبت سے تنہائی بہتر ہے اور تنہائی سے نیک صحبت بہتر ہے

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** (Page 419)

- Creation of package
- Connection of package

Question No: 30 (Marks: 1) – Please choose one

To judge effectiveness we perform data profiling twice.

- 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

FINAL TERM EXAMINATION
Spring 2009
CS614- Data Warehousing

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

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

► **Doubles** (Page 15)

- Triples
- Quartiles
- Remains same as previous year

Question No: 3 (Marks: 1) - Please choose one

Pre-computed _____ can solve performance problems

► **Aggregates** (Page 11)

- Facts
- Dimensions

Question No: 4 (Marks: 1) - Please choose one

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

► **A Dense Index** (Page 223)

- A Sparse Index
- An Inverted Index
- None of These

زندگی میں کامیابی کا یہی راز ہے کہ پریشانیوں سے پریشان مت بنو

Question No: 5 (Marks: 1) - Please choose one

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: 6 (Marks: 1) - Please choose one

The purpose of the House of Quality technique is to reduce _____ types of risk.

▶ **Two** (Page 194)

▶ Three

▶ Four

▶ All

Question No: 7 (Marks: 1) - Please choose one

NUMA stands for _____

▶ **Non-uniform Memory Access** (Page 206)

▶ Non-updateable Memory Architecture

▶ New Universal Memory Architecture

Question No: 8 (Marks: 1) - Please choose one

Which is the least appropriate join operation for Pipeline parallelism?

▶ Hash Join

▶ Inner Join

▶ Outer Join

▶ **Sort-Merge Join**

Question No: 9 (Marks: 1) - Please choose one

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

▶ Naive nested-loop join

▶ Index nested-loop join

▶ **Temporary index nested-loop join** (Page 243)

▶ None of these

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

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** [Click here for detail](#)

Muhammad Moaaz Siddiq – MCS(4th)

Moaaz.pk@gmail.com

**Campus:- Institute of E-Learning & Modern Studies
(IEMS) Samundari**

Question No: 11 (Marks: 1) - Please choose one

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

[Click here for detail](#)

Question No: 12 (Marks: 1) - Please choose one

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 (Marks: 1) - Please choose one

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

- ▶ True
- ▶ **False** (Page 315)

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

For a smooth DWH implementation we must be a technologist.

- ▶ True
- ▶ **False** (Page 319)

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

During the application specification activity, we also must give consideration to the organization of the applications.

- ▶ **True** (Page 307)
- ▶ False

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

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

- ▶ **Extreme Technology Design**
- ▶ **Extreme Architecture Design**
- ▶ ~~None of these~~ (Page 315)

دنیا کی سب سے بڑی فتح نفس پر قابو رکھنا ہے

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 No: 19 (Marks: 1) - Please choose one

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

► **Horizontal splitting** (Page 54)

► Vertical splitting

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

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

► **Committed to the database** (Page 419) rep

► Rolled back

جھوٹ انسان اور ایمان دونوں کا دشمن ہے

خود کو تمھیں سے بڑھ کر کوئی اچھا مشورہ نہیں دے سکتا

عقل مند کہتا ہے میں کچھ نہیں جانتا جبکہ بے وقوف کہتا ہے کہ میں سب کچھ جانتا ہوں

Muhammad Moaaz Siddiq – MCS(4th)

Moaaz.pk@gmail.com

Campus:- Institute of E-Learning & Modern Studies
(IEMS) Samundari

Cs614 – Final Term Quizzes (Quiz No.3 & 4)

Quiz No.3 (7th June 2012)

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 # 6 of 10 (Total Marks: 1)

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 the computation.

- ▶ Larger
- ▶ **Smaller** (Page 204)
- ▶ Unambiguous
- ▶ Superior

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 # 8 of 10 (Total Marks: 1)

In nested-loop join case, if there are 'M' rows in outer table and 'N' rows in inner table, time complexity is

- ▶ $O(M \log N)$
- ▶ $O(\log MN)$
- ▶ $O(MN)$
- ▶ $O(M + N)$

Question # 9 of 10 (Total Marks: 1)

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

- ▶ **Dense Index** (Page 223)
- ▶ Sparse Index
- ▶ Inverted Index
- ▶ None

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)

جو لوگوں کے سامنے فخر کرتا ہے وہ لوگوں کی نظروں سے گر جاتا ہے

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!"** (Page 181)

Question # 2 of 10 (Total Marks: 1)

_____, 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

جو شخص ناکامیوں سے ڈر کر بھاگتا ہے کامیابی اُس سے ڈر کر بھاگتی ہے

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 # 8 of 10 (Total Marks: 1)

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

► **OLTP** (Page 254)

► OLAP

► DSS

► DWH

Question # 9 of 10 (Total Marks: 1)

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 plus 'b' I/Os for each data block then the total cost of accessing table-B is _____ logical I/Os approximately.

Select correct option:

► **(a + b)M** (Right Answer)

► (a - b)M

► (a + b + M)

► (a * b * M)

Question # 10 of 10 (Total Marks: 1)

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

► True

► **False** (Page 69)

بد صورت چہرہ بد صورت دماغ سے بہتر ہے

Muhammad Moaaz Siddiq – MCS(4th)

Moaaz.pk@gmail.com

**Campus:- Institute of E-Learning & Modern Studies
(IEMS) Samundari**

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
- ▶ **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
- ▶ **Essential** (Page 69)
- ▶ Optional
- ▶ 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
- ▶ **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

CS614 Quiz No.3 & 4 (Fall 2012)

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

If every key in the data file is represented in the index file then it is called

Dense Index (Page 223)

Sparse Index

Inverted Index

A Multi level Sparse Index

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

Question # 6 of 10 (Total Marks: 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 Marks: 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 Marks: 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 Marks: 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)

Some More MCQs

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 Marks: 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 M a r k s: 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 Marks: 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.

Cleansing
Cleaning
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 Marks: 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 Marks: 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 Marks: 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 Marks: 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 Marks: 1) Select correct option:

D-Normalization speeds up_____

- ▶ **Data Retrieval (Page 51)**
- ▶ Data Duplication
- ▶ Data Extraction
- ▶ Data Loading

Question # 6 of 10 (Total Marks: 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 Marks: 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 Marks: 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

