



## CS614 - Data Warehousing

Solved MCQS  
From Midterm Papers

Nov 30, 2012

MC100401285

Moaaz.pk@gmail.com

Mc100401285@gmail.com

PSMD01

### MIDTERM EXAMINATION Spring 2011 CS614- Data Warehousing

#### Question No: 1 ( 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** [Click here for detail](#)
- ▶ None of these

#### Question No: 2 ( 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: 3 ( 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 \_\_\_\_\_ the portion of the program that must be executed sequentially, the greater the scalability of computation.

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

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

**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.

- ▶ An inverted index
- ▶ A sparse index
- ▶ **A dense index (Page 223)**
- ▶ none of these

**Question No: 5 ( 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: 6 ( Marks: 1 ) - Please choose one**

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 No: 7 ( 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 here for detail)**
- ▶ Ascent
- ▶ Avowal

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

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

خدا کے سوا کسی سے امید مت رکھو

**Muhammad Moaaz Siddiq – MCS(4th)**

**Moaaz.pk@gmail.com**

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

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

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

**MIDTERM EXAMINATION**  
**Spring 2010**  
**CS614- Data Warehousing (Session - 6)**

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

The need to synchronize data upon update is called

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

**Question No: 2 ( 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: 3 ( 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

اللہ کا خوف سب سے بڑی دانائی ہے

**Question No: 4 ( 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 (n<sup>2</sup>)
- ▶ O (n lg n)
- ▶ **O (log n) (Page 22)**

**Question No: 5 ( 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: 6 ( Marks: 1 ) - Please choose one**

Pre-join technique is used to avoid

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

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

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

The goal of star schema design is to simplify \_\_\_\_\_

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

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

**Muhammad Moaaz Siddiq – MCS(4th)**

**Moaaz.pk@gmail.com**

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

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

Grain is the \_\_\_\_\_ level of data stored in the warehouse.

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

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

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

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

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

Change Data Capture is one of the challenging technical issues in \_\_\_\_\_

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

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

**Muhammad Moaaz Siddiq – MCS(4th)**

**Moaaz.pk@gmail.com**

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

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

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

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

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

Normalization effects performance

- ▶ **True**
- ▶ False

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

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

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

**Muhammad Moaaz Siddiq – MCS(4th)**

**Moaaz.pk@gmail.com**

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

**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**

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)**
- ▶  $O(w \log n)$

**MIDTERM EXAMINATION**  
**Spring 2010**  
**CS614- Data Warehousing**

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

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

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

**Question No: 2 (Marks: 1 ) - Please choose one**

\_\_\_\_\_ is class of Decision Support Environment.

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

دنیا کی سب سے بڑی فتح نفس پر قابور کھنا ہے

**Question No: 3 (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) rep**
- ▶ Decision Maker's Perspective
- ▶ Employee's Perspective

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

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

Horizontal splitting breaks a table into multiple tables based upon\_\_\_\_\_

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

**Question No: 6 (Marks: 1 ) - Please choose one**

D-Normalization speeds up\_\_\_\_\_

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

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

The most common use of range partitioning is on \_\_\_\_\_.

- ▶ **Date (Page 66)**
- ▶ Rows
- ▶ DSS
- ▶ None of these

جھوٹ انسان اور ایمان دونوں کا دشمن ہے

**Muhammad Moaaz Siddiq – MCS(4th)**

**Moaaz.pk@gmail.com**

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



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

The fact table is a way of *visualizing* as an \_\_\_\_\_cube

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

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

High Granularity, more \_\_\_\_\_.

- ▶ **Details** [Click here for detail](#)
- ▶ Dimensions
- ▶ Aggregates
- ▶ Cubes

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

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

Cubes not drilled down is, It is a \_\_\_\_\_.

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

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

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) rep**
- ▶ One-to-Many Element Transformation
- ▶ Many-to-Many Element Transformation
- ▶ Many-to-One Element Transformation

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 & Moderen Studies  
(IEMS) Samundari**

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

Full & Incremental Extraction are the types of \_\_\_\_\_ Extraction

▶ **Logical (Page 132)**

- ▶ Physical
- ▶ Both Logical & Physical
- ▶ None of Above

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

Selectivity is low in \_\_\_\_\_ environment.

▶ **DWH (Page 22)**

- ▶ DBMS
- ▶ OLTP
- ▶ None of Above

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

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) rep**

- ▶ Only Complex Ratio, Min Operation, Max Operation
- ▶ Only Simple Ratio, Min or Max Operation
- ▶ Only Min or Max Operation, Weighted Average

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

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

▶ **True (Page 19) rep**

- ▶ False

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

Normalization effects performance

- ▶ True
- ▶ False

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

**Muhammad Moaaz Siddiq – MCS(4th)**

**Moaaz.pk@gmail.com**

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

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

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 No: 19 (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: 20 (Marks: 1 ) - Please choose one**

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) rep**
- ▶ O (w log n)

**MIDTERM EXAMINATION**  
**Spring 2009**  
**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

**جو لوگوں کے سامنے فخر کرتا ہے وہ لوگوں کی نظروں سے گرجاتا ہے**

**Question No: 2 ( Marks: 1 ) - Please choose one**

De-Normalization normally speeds up

- ▶ **Data Retrieval (Page 51) rep**
- ▶ 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) rep**
- ▶ 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) rep**
- ▶ 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
- ▶  $\lg(n)$
- ▶  $n$

**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

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

**Muhammad Moaaz Siddiq – MCS(4th)**

**Moaaz.pk@gmail.com**

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

**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

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 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

عقل مندا اپنے عیب خود دیکھتا ہے اور بیوقوفوں کے عیب دنیا دیکھتی ہے

**Muhammad Moaaz Siddiq – MCS(4th)**

**Moaaz.pk@gmail.com**

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

**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

عقل مند آدمی اس وقت تک نہیں بولتا جب تک خاموشی نہیں ہو جاتی

**Muhammad Moaaz Siddiq – MCS(4th)**

**Moaaz.pk@gmail.com**

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

**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)**
- ▶ Documentations
- ▶ Guidelines

بہترین تجربہ وہ ہے جس سے نصیحت حاصل ہو

**Muhammad Moaaz Siddiq – MCS(4th)**

**Moaaz.pk@gmail.com**

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



**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

**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

**MIDTERM EXAMINATION**  
**Spring 2008**  
**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: 2 ( Marks: 1 ) - Please choose one**

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

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

انسان دکھ نہیں دیتے بلکہ انسانوں سے وابستہ امیدیں دکھ دیتی ہیں

**Question No: 3 ( Marks: 1 ) - Please choose one**  
Pre-computed \_\_\_\_\_ can solve performance problems

- ▶ **Aggregates (Page 111)**
- ▶ 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) rep**
- ▶ A Sparse Index
- ▶ An Inverted Index
- ▶ None of These

**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

فٹنہ انگیز سچائی سے مصلحت آمیز جھوٹ بہتر ہے

**Muhammad Moaaz Siddiq – MCS(4th)**

**Moaaz.pk@gmail.com**

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

**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](#)

**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)

خوبصورتی علم و ادب سے ہوتی ہے لباس و حسن سے نہیں

**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)**

Ref:- Extremes of Tech. Arch. Design

**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) rep**
- ▶ 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) rep**
- ▶ Vertical splitting

تم اچھا کرو زمانہ تم کو برا سمجھے یہ اس سے بہتر ہے کہ تم برا کرو اور زمانہ تم کو اچھا سمجھے

**Muhammad Moaaz Siddiq – MCS(4th)**

**Moaaz.pk@gmail.com**

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

**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)**
- ▶ Rolled back

## CS614- Data Warehousing Quiz No. 1 & 2

### Quiz No.1

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

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

Select correct option:

- ▶ **Doubles (Page 15) rep**
- ▶ Remains same as previous year
- ▶ Triples
- ▶ Quartiles

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

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

Select correct option:

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

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

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

Select correct option:

- ▶ 1950s.
- ▶ **1960s. (Page 12)**
- ▶ 1970s.
- ▶ 1980s.

**خاموشی غصے کا بہترین علاج ہے**

**Question # 4 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 # 5 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 # 6 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 # 7 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 # 8 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 # 9 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

انسان کے لئے بری صحبت سے بڑھ کر بری کوئی چیز نہیں

**Muhammad Moaaz Siddiq – MCS(4th)**

**Moaaz.pk@gmail.com**

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

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

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

- ▶ True
- ▶ **False (Page 18)**

## Quiz No.2

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

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) rep**
- ▶ None of these

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

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

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

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

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

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

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

Focusing on data warehouse delivery only often end up \_\_\_\_\_.

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

افضل انسان وہ ہے جو اپنی اصلاح کی کوشش کرتا ہے

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

Pakistan is one of the five major \_\_\_\_\_ countries in the world.

- ▶ **Cotton-growing**
- ▶ Rice-growing
- ▶ Weapon Producing

**Question # 6 of 10 (Total M a r k s: 1)**

Pre-computed \_\_\_\_\_ can solve performance problems

- ▶ **Aggregates (Page 111) rep**
- ▶ Facts
- ▶ Dimensions

**Question # 7 of 10 (Total M a r k s: 1)**

The degree of similarity between two records, often measured by a numerical value between \_\_\_\_\_, usually depends on application characteristics.

- ▶ **0 and 1 (Page 169) rep**
- ▶ 0 and 10
- ▶ 0 and 100
- ▶ 0 and 99

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

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

- ▶ **Two (Page 194) rep**
- ▶ hree
- ▶ Four
- ▶ All

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

NUMA stands for \_\_\_\_\_

- ▶ **Non-uniform Memory Access (Page 206) rep**
- ▶ Non-updateable Memory Architecture
- ▶ New Universal Memory Architecture

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

\_\_\_\_\_ modeling technique is more appropriate for data warehouses.

- ▶ entity-relationship
- ▶ **dimensional** [Click here for detail](#)
- ▶ physical
- ▶ None of the given

جھوٹ رزق کو کھا جاتا ہے

**Muhammad Moaaz Siddiq – MCS(4th)**

**Moaaz.pk@gmail.com**

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



## Quiz No.1

### Question # 1 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 # 2 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 # 3 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 # 4 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 # 5 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

اطمینان قلب چاہتے ہو تو حسد سے دور رہو

Muhammad Moaaz Siddiq – MCS(4th)

Moaaz.pk@gmail.com

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

**Question # 6 of 10 (Total M a r k s: 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)

## Quiz No.2

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

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

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

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

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

Select correct option:

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

**Question # 3 of 10 (Total M a r k s: 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 # 4 of 10 (Total M a r k s: 1)**

To identify the \_\_\_\_\_ required we need to perform data profiling

- ▶ **Degree of Transformation** (Page 437)
- ▶ Complexity
- ▶ Cost
- ▶ Time

اس سے پہلے کہ تمہیں شہوت فتنے میں ڈالے نکاح کر لو

**Muhammad Moaaz Siddiq – MCS(4th)**

**Moaaz.pk@gmail.com**

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

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

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 # 6 of 10 (Total M a r k s: 1)**

\_\_\_\_\_ 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

**Question # 7 of 10 (Total M a r k s: 1)**

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

- ▶ **True (Page 332)**
- ▶ False

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

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

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 these

ہر چیز کی ایک پہچان ہوتی ہے اور عقلمند کی پہچان غور و فکر کرنا ہے اور غور و فکر کی پہچان خاموشی ہے

**Muhammad Moaaz Siddiq – MCS(4th)**

**Moaaz.pk@gmail.com**

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

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

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

- ▶ OLTP
- ▶ OLAP
- ▶ DSS
- ▶ **Inverted Index** (Page 232)

## Quiz No.2

**Question # 2 of 10 (Total M a r k s: 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 # 3 of 10 (Total M a r k s: 1)**

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 # 4 of 10 (Total M a r k s: 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 these

وہ لوگ مبارک ہیں جو الفاظ سے نصیحت نہیں کرتے بلکہ عمل سے کرتے ہیں

**Question # 6 of 10 (Total M a r k s: 1)**

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

▶ True (Page 320)

▶ False

**Question # 7 of 10 (Total M a r k s: 1)**

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 here for detail](#)

▶ Ascent Avowal

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

De-Normalization normally speeds up

▶ Data Retrieval (Page 51)

▶ Data Modification

▶ Development Cycle

▶ Data Replication

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

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

## Quiz No.2

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

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

▶ True

▶ False (Page 69)

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

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

▶ True (Page 79)

▶ False

اپنی مرضی اور اللہ کی مرضی میں فرق کا نام غم ہے

**Muhammad Moaaz Siddiq – MCS(4th)**

**Moaaz.pk@gmail.com**

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

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

Slice and Dice is changing the view of the data.

▶ True

▶ **False** (Page 81)

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

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

▶ **True** (Page 72)

▶ False

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

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

▶ True

▶ **False** (Page 78)

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

**Question # 6 of 10 (Total M a r k s: 1)**

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

Select correct option:

▶ Network model

▶ **Relational model** (Page 87)

▶ Hierarchical data model

▶ None of the given

کامیاب و کامران زندگی یہی ہے کہ جہاں رہو جس حال میں رہو خوش رہو

**Muhammad Moaaz Siddiq – MCS(4th)**

**Moaaz.pk@gmail.com**

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