# CS-614 Data Warehousing

## Update MCQS For Quiz-3 File

## Solve By Vu Topper RM

**80 To 100% Marks**




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

**For More Help Contact What's app 03224021365**

"What means What". The phrase refers to:

**A. Meta data Page 338**

B. External data

C. Transformed data

D. Internal representations

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

A. Scope

B. Feasibility

C. Resources available

**D. The potential impact or value to the business Page 297**

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

A. Data driven

**B. System driven Page 283**

C. Resource driven

D. Requirement driven

Which of the following is the most ignored step during data warehouse development

A. Schema validation

B. The vision definition

C. The requirement verification

**D. Success criteria development**

Waterfall model is appropriate when

A. When the budget is low

B. When the deadline is strict

C. When resources are limited

**D. Requirements are clearly defined Page 284**

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

A. Tables should be properly denormalized

VU Topper RM

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

For More Help Contact What's app 03224021365

B. Proper partitioning technique should be used
C. At least star or snow flake schema should be implemented
**D. They should get information in timely manner and the way they want**

A top down implementation approach is useful when
A. Business objectives are unclear
**B. Technology is mature and well understood** **Page 283**
C. Problem to be solved are not well documented
D. Organization can not implement latest technologies

Identify the TRUE statement:
A. Both clustering and classification are supervised learning
B. Both clustering and classification are unsupervised learning
**C. Clustering is unsupervised learning and classification is supervised learning** **Page 270**
D. Clustering is supervised learning and classification is unsupervised learning

The _____ phase of fundamental data warehouse life cycle model includes data warehouse daily maintenance activities
A. Project
**B. Operation** **Page 288**
C. Deployment
D. Maintenance

Horizontally wide data means:
A. Dataset has attribute skews
B. Dataset has partitioning skews
C. Dataset has large no. of records
**D. Dataset has large no. of attributes** **Page 330**

Which of the following is/are included in the list of Top-7 key steps for smooth DWH implementation?
**A. All of the given options** **Page 318**

بری صحبت سے تنہائی بہتر ہے اور تنہائی سے نیک صحبت بہتر ہے
For More Help Contact What's app 03224021365
Vu Topper RM

B. Be a diplomat NOT a technologist
C. Assign significant resources for ETL
D. Consider handing-off project management

Implementation of a data warehouse requires _____ activities
A. Highly integrated
B. Loosely integrated
C. Tightly decoupled
**D. None of the given Page 289**

Goal driven approach of data warehouse development was result of _____ work
A. Bill Inmon
B. Westerman
C. Ralph Kimball
**D. Böhnlein and Ulbrich-vom Page 285**

In context of analytical applications track, the application development activity can begin after:
**A. Database Design is complete Page 307**
B. A subset of historical data has been loaded
C. The data access tools and metadata are installed
D. All of the given options

In _____ phase of a fundamental data warehouse life cycle model, a working model of data warehouse is deployed for a selective set of users
A. Design
B. Operation
**C. Prototype Page 287**
D. Deployment

Which of the following activity executes parallel with all other activities in Kimball's DWH development approach?
A. Deployment
B. Project planning

Vu Topper RM
بری صحبت سے تنہائی بہتر ہے اور تنہائی سے نیک صحبت بہتر ہے
For More Help Contact What's app 03224021365

**C. Project management** **Page 289**
D. Requirement elicitation

The first step of the "12-steps data warehouse implementation approach" of Shaku Atre is:

**A. Finding user needs** **Page 336**
B. Finding system scope
C. Planning system resources
D. Data acquisition and cleansing

A typical cycle of implementing the change in DWH comprises of the sequence:

A. Production -> Development -> QA
**B. Development-> QA -> Production** **Page 314**
C. Development -> Production -> QA
D. Production -> QA -> Development

Which of the following is NOT one of the three parallel tracks in Kimballs approach?

A. Lifecycle technology track
B. Lifecycle Data track
**C. Lifecycle Maintenance track** **Page 299**
D. Lifecycle Analytical Applications track

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

**A. Data driven** **Page 285**
B. Time sensitive
C. Resource driven
D. Requirement driven

Normally the input data structure (a database table) for a data mining algorithm:

A. Has limited amount of data
B. Has limited amount of relations

VU Topper RM

بری صحبت سے تنہائی بہتر ہے اور تنہائی سے نیک صحبت بہتر ہے
For More Help Contact What's app 03224021365

C. Has more number of records than attributes
**D. Has more number of attributes than records**

Normally the term "DWH face to the business user" refers to:
A. Lifecycle Data track
B. Lifecycle technology track
C. Lifecycle Maintenance track
**D. Lifecycle Analytical Applications track**

An effective user education program includes, among others, the following guideline(s):
A. Postpone education, if DWH not ready
B. Understand the audience, don't overwhelm
C. Train after delivery of the data and analytic applications
**D. All of the given options Page 308**

In _____ phase of kimballs approach, we identify the components needed now and in future.
**A. Architectural design Page 300**
B. Product development
C. Requirement definition
D. Analytical application development

Spiral model is _____
A. Risk oriented model
B. Risk oriented model
C. Sequence of waterfall model
**D. All of the given Page 284**

In context of requirement definition phase of Kimball's DWH development approach, _____ is positioned as a findings review and prioritization meeting.
A. System analysis
B. Scope definition
C. Requirement configuration

بری صحبت سے تنہائی بہتر ہے اور تنہائی سے نیک صحبت بہتر ہے
For More Help Contact What's app 03224021365
VU Topper RM

**D. Requirements wrap-up presentation Page 297**

In analytical application development phase, we follow standards for:

A. Naming conventions.
B. Standard for libraries.
C. Standard for calculations.
**D. All of the given option Page 307**

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

**A.** (I) only
**B.** (II) Only
**C.** (III) Only
**D. (I), (II) and (III) Page 300**

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

A. Source system cataloguing
**B. System vision development Page 287**
C. End-user interviews and re-interviews
D. Definition of key performance indicators

Improper documentation results the problem(s) like:

A. Maintenance issue
B. Lot of time required for enhancing the code
**C. New developers unable to configure already existing code Page 314**
D. All of the given options

Vertically wide data means:

A. Dataset has attribute skews
B. Dataset has partitioning skews

Vu Topper RM

بری صحبت سے تنہائی بہتر ہے اور تنہائی سے نیک صحبت بہتر ہے
For More Help Contact What's app 03224021365

**C. Dataset has large no. of records**
D. Dataset has large no. of attributes

In _____learning you don't know the number of clusters and no idea about their attributes.

A. Supervised learning
**B. Un-Supervised learning** **Page 269**
C. Multi Dimension modeling
D. None of the given options

Bill Inmon argues that requirements are well understood only after

A. They are documented
B. All schemes are defined
C. They extracted and verified
**D. Data warehouse is populated** **Page 285**

Waterfall is a/an _____ model

A. Iterative
B. Object Oriented
C. Rapid development
**D. Simple linear sequential** **Page 284**

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

A. Obtain resources
B. Establish the preliminary scope and justification
C. Assess organization's readiness for a data warehouse initiative
**D. All of the given options** **Page 290**

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

A. Daily Sale
B. Air pressure
**C. Date of Birth** **Page 342**
D. None of the given options

VuTopperRM

بری صحبت سے تنہائی بہتر ہے اور تنہائی سے نیک صحبت بہتر ہے
For More Help Contact What's app 03224021365

A bottom up implementation approach is useful when

A. All the requirements are well documented
B. The business objectives are very much clear
**C. Technology is mature and well understood Page 283**
D. Organization can not implement latest technologies

The phase-3 i.e. "Rollout and Maintenance" phase in "12-steps data warehouse implementation approach" of Shaku Atre comprises of _____ activity.

A. On Line Analytical Processing
B. Construct Metadata Repository
**C. Deployment & System Management**
D. None of the these

In lifecycle data track, we begin with translating the requirements into dimensional model, which then transforms into _____

A. System structure
B. Logical structure
**C. Physical structure Page 290**
D. Conceptual structure

In context of clustering, the term "distance" means:

**A. Similarity/dissimilarly of records Page 272**
B. The difference between the primary keys of two records
C. The relation of a record with corresponding record in child table
D. None of the given options

As per Kimball, _____ is the main operational process

A. Goal design
B. Schema design
**C. Business process Page 285**
D. Requirement extraction

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

**A. Employee Name Page 342**
B. Employee Blood Group

بری صحبت سے تنہائی بہتر ہے اور تنہائی سے نیک صحبت بہتر ہے
For More Help Contact What's app 03224021365
Vu Topper RM

C. Employee Date of Birth
D. All of the given options

Which of the following is NOT one of the possible pitfalls in DWH Life Cycle & Development?

A. Improper documentation
B. Not having multiple servers
C. Low priority for OLAP Cube Construction
**D. None of the given options Page 313**

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

A. Support
B. Education Program
C. Education Technical
**D. Interface Deployment Page 309**

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

A. Bill inmon
B. Kimball
C. Bohnlein
**D. Westerman Page 285**

One of the drawbacks of waterfall model is that:

A. It does not define the project timeline/schedule
B. It does not work when the resources are limited
**C. Customers can not review the product during development Page 284**
D. All of the given options

_____ index uses even less space than _____ index, but the block has to be searched, even for unsuccessful searches.

**Sparse, dense Page 224**
Dense, inverted
Sparse, inverted

VU Topper RM

بری صحبت سے تنہائی بہتر ہے اور تنہائی سے نیک صحبت بہتر ہے
For More Help Contact What's app 03224021365

Dense, sparse

In context of data mining definition, the term “value” means:
The primary key of table
The index location of the record
**Importance of hidden patterns discovered Page 250**
Numerical or string measure assigned to an attribute

If a task takes “T” time units to execute on a single data item, then execution of this task on “N” data items will take _____ time units.
**N*t**
N+t
N/t
N-t

Which of the following is NOT one of the variants of Nested-loop join?
Naïve nested-loop join
Indexed nested-loop join
**Binary index nested-loop join**
Temporary index nested-loop join

_____ index stores first value in each block in the sequential file and a pointer to the block.
Hash
Dense
**Sparse Page 224**
B-Tree

_____ refers to the overall process of discovering useful knowledge from data and data mining refers to a particular step in this process.
Statistics
Clustering
Information cleansing
**Knowledge discovery in database Page 254**

Vu Topper RM

بری صحبت سے تنہائی بہتر ہے اور تنہائی سے نیک صحبت بہتر ہے
For More Help Contact What’s app 03224021365

Which of the following tasks can NOT be parallelized?
Partitioned index scans
Creation of large indexes
Large table scans and joins
**None of the given options**

In nested-loop join case, if there are ‘M’ rows in outer table and ‘N’ rows in inner table, time complexity is
**O (MN)**
O (M + N)
O (log MN)
O (M log N)

In context of data mining definition, the term “nontrivial” means:
We can not discover information
Discovering information is a simple task
**Discovering information is a complex task**
We simply find things rather than discovery

In contrast to statistics, data mining is _____ driven.
Human
Database
Assumption
**Knowledge** **Page 254**

Parallelism can be exploited, if there is:
**All of the given options**
Sufficient I/O bandwidth
Symmetric multi-processors (SMP)
Underutilized or intermittently used CPUs

One of the preconditions to decide about operations to be parallelized is that:
Operations share same namespace
Operations share same memory location
Output of one operation becomes input of other

VU Topper RM

بری صحبت سے تنہائی بہتر ہے اور تنہائی سے نیک صحبت بہتر ہے
For More Help Contact What’s app 03224021365

**Operations can be implemented independent of each other**

In B-tree index, the lowest level index blocks are called leaf blocks, and these blocks contain:
NULL value to make the leaf terminal node
Every indexed data value and pointer to root block
Every indexed data value and pointer to next level block
**Every indexed data value and a corresponding ROWID Page 225**

Which of the following is NOT one of the parallel hardware architectures?
**Shared Memory**
Symmetric Multi Processing
Massively Parallel Processing
Non-uniform Memory Access

In context of joining tables, the join condition is specified in ____ clause
FROM
SELECT
**WHERE**
GROUP BY

Classification consists of examining the properties of a newly presented observation and assigning it to a predefined _____.
**Class Page 259**
Object
Subject
Container

An optimized structure which is built primarily for retrieval, with update being only a secondary consideration is
DSS
OLAP
OLTP
**Inverted Index Page 232**

Vu Topper RM

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

For More Help Contact What's app 03224021365

Data mining is all about:
**All of the given options** **Page 249**
Finding relationships in data
Finding hidden patterns in data
Knowledge discovery in database

Mining multi dimensional databases allow users to:
**Analyze the data**
Categorize the data
Summarize the data
All of the given options

In context of data parallelism, the work done by query processor should be:
Pipelined
Maximum
**Almost zero**
Filtered across partitions

_____ do not (typically) keep the index values in sorted order
Dense index
Sparse index
B-Tree index
**Hash based index** **Page 227**

_____ refers to “Parallel execution of single data operation across multiple partitions of data”
**Data parallelism** **Page 212**
Software parallelism
Hardware parallelism
Operational parallelism

In _____ index, the ith bit is set to "1" if the ith row of the base table has the value for the indexed column.
Join index
Cluster index

VU Topper RM

بری صحبت سے تنہائی بہتر ہے اور تنہائی سے نیک صحبت بہتر ہے
For More Help Contact What’s app 03224021365

**Bitmap index**
Inverted index

_____ is the technique in which existing heterogeneous segments are reshuffled, relocated into homogeneous segments.
**Clustering** **Page 269**
Partitioning
Aggregation
Segmentation

In context of bitmap index, the length of the bit vector is:
None of the given options
**The number of records in the base table**
The possible number of domain values in corresponding field (column)
The possible number of bitmap tables formed for corresponding field (column)

Bitmap index is appropriate for:
Clustered data
Aggregated data
High cardinality data
**Low cardinality data**

In context of data parallelism, to get a speed-up of N with N partitions, it must be ensured that:
**All of the given options**
There are enough computing resources
Work done in each partition almost same
Query-coordinator is very fast as compared to query servers

VU Topper RM

بری صحبت سے تنہائی بہتر ہے اور تنہائی سے نیک صحبت بہتر ہے
For More Help Contact What's app 03224021365

Question # 3 of 10 ( Start time: 01:04:49 PM ) Total Marks: 1
Which of the following is NOT one of the top-10 mistakes that should be avoided during DWH development?
Select correct option:
Not interacting directly with end user
**Not being an accommodating person (Page No. 316)**
Isolating IT support people from business users
Training the users with dummy data and considering it success

Question # 10 of 10 ( Start time: 09:10:52 PM ) Total Marks: 1
Which of the following is NOT one of the top-10 mistakes that should be avoided during DWH development?
Select correct option:
Not interacting directly with end user
Not being an accommodating person(CORRECT)
Isolating IT support people from business users
Training the users with dummy data and considering it success

Question # 4 of 10 ( Start time: 10:17:31 PM ) Total Marks: 1
In contrast to data mining, statistics is _____ driven.
Select correct option:
Assumption (Correct)
Knowledge
Discovery
Database

2_ _____ do not (typically) keep the index values in stored order

Hash based index

3_ if every key in the data is represented in the index file then it is called

Dense index

6_ A join is identified by multiple tables in the

____ clause. From

9_ As the number of processors increase, the speedup should also increase. thus we should have linear speedup. Which of the following is NOT the one of the barriers to achieve this linear speed-up?

Amdah'l Law not sure

Q1: in context of nested-loop join, actual number os matching rows returned as a result of the join would be _____ of the order of tables. Independent.

Vu Topper RM

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

For More Help Contact What's app 03224021365

Q3: If resources increase in proportion to increase in data size. time is constant'. The statement refers to:
Scale-Up

Q4: If every key in the data file is represented in the index file then it is called? Dense Index

Q7: The goal of_____ is to look at as few blocks as possib le to find the matching records(s). Indexing

Q8: Parallelism can be exploited, if there is. All of the given options

Q13: if every key the data file is represented in the index file then it is called. Dense Index

Q16: A join is identified by multiple tables in the_____ clause. From

Q17: Parallelism can exploited, if there is
All of the given options

Q19: As the number of processors increase, the speedup should also increase. Thus we should have linear speedup. Which of the following is NOT one of the barriers to achieve this linear speed-up? Amdahl' Law

Q22: _____ lists each term in the collection only once and then shows a list of all the documents the contain the given term.
Inverted index

Q23: "More resources means proportionally less time for given amount of data". The statement refers to:
Speed-UP

Q27: A_____ index, if fits in the memory, costs only one disk I/O access to locate a record given a key.
Dense Index

Q28: In context of nested-loop join, actual number of matching rows returned as a result of the join would be ___ of the order of tables
Independent

Q29: _____ refers to " Parallelexectution of single data operation across multiple partitions of data" Data Parallelism.

A join is identified by multiple tables in the _ FROM ___ clause

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

2_ _____ do not (typically) keep the index values in stored order

Hash based index

Vu Topper RM

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

For More Help Contact What's app 03224021365

3_ if every key in the data is represented in the index file then it is called

Dense index

6_ A join is identified by multiple tables in the

_____ clause. From

9_ As the number of processors increase, the speedup should also increase. thus we should have linear speedup. Which of the following is NOT the one of the barriers to achieve this linear speed-up?

Amdah'l Law not sure

Q1: in context of nested-loop join, actual number os matching rows returned as a result of the join would be _____ of the order of tables. Independent.

Q3: If resources increase in proportion to increase in data size. time is constant'. The statement refers to:
Scale-Up

Q4: If every key in the data file is represented in the index file then it is called? Dense Index

Q7: The goal of_____ is to look at as few blocks as possib le to find the matching records(s). Indexing

Q8: Parallelism can be exploited, if there is.
All of the given options

Q13: if every key the data file is represented in the index file then it is called. Dense Index

Q16: A join is identified by multiple tables in the_____ clause. From

Q17: Parallelism can exploited, if there is
All of the given options

Q19: As the number of processors increase, the speedup should also increase. Thus we should have linear speedup. Which of the following is NOT one of the barriers to achieve this linear speed-up? Amdahl' Law

Q22: _____ lists each term in the collection only once and then shows a list of all the documents the contain the given term.
Inverted index

Q23: "More resources means proportionally less time for given amount of data". The statement refers to:
Speed-UP

Q27: A_____ index, if fits in the memory, costs only one disk I/O access to locate a record given a key.
Dense Index

Vu Topper RM

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

For More Help Contact What's app 03224021365

Q28: In context of nested-loop join, actual number of matching rows returned as a result of the join would be ___ of the order of tables
Independent

Q29: _____ refers to " Parallelexectution of single data operation across multiple partitions of data" Data Parallelism.

A join is identified by multiple tables in the _ FROM ___ clause

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

it must be ensured that:

Customers can not review the product during development
It does not work when the resources are limited
It does not define the project timeline/schedule
**All of the given options (Correct)**

The primary key of table
The index location of the record
**Importance of hidden patterns discovered (Correct)**
Numerical or string measure assigned to an attribute

Question # 3 of 10 ( Start time: 09:28:57 PM ) Total Marks: 1
As opposed to the outcome of classification, estimation deals with _____ valued outcome. Select correct option:
Discrete (Correct)
Isolated
Continuous
Distinct

**CS614 Data Warehousing Solved MCQs from Quiz # 2 Solved and shared by Gulshan**

Data mining uses _____ algorithms to discover patterns and regularities in data. Select correct option:
Mathematical
Computational
Statistical
Non
e of
thes
e

Vu Topper RM

بری صحبت سے تنہائی بہتر ہے اور تنہائی سے نیک صحبت بہتر ہے
For More Help Contact What's app 03224021365

If every key in the data file is represented in the index file then index is Select correct option:

**Dense Index**

Sparse Index

Inverted Index

None of these

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

None of these

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**

OLAP

DSS

DWH

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

Two

Linear

Quadratic

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**

Unambiguous

Superior

_____, 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**

None of these

Vu Topper RM

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

For More Help Contact What's app 03224021365

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

Data mining uses _____ algorithms to discover patterns and regularities in data. Mathematical
Computational
**Statistical**
None of these

_____, 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**
None of these

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
None of these

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.
In Parallel
Distributed
**Sequentially**
None of these

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.
Non-Exploratory
**Exp**
**lora**
**tory**
Co
mp
ute
Sci
enc
e
none of these

To identify the _____ required we need to perform data profiling **Degree of Transformation**
Complexity
Cost
Time

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 _____
Committed to the database

Vu Topper RM

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

For More Help Contact What's app 03224021365

**Rolled back**

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**

_____ in agriculture extension is that pest population beyond which the benefit of spraying outweighs its cost. Profit Threshold Level
Economic Threshold Level
Medicine Threshold Level
**None of these**

People that design and build the data warehouse must be capable of working across the organization at all levels **True**
False
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**
None of these

1_ As per Bill Inmost, a data warehouse, in contrast with classical applications is:

Data driven _ pg285

2_ Which of the following is NOT one of the three parallel tracks in Kimball's approach? Lifecycle Maintenance track

The optimizer uses a hash join to join two tables if they are joined using an equijoin and Select correct option:

Large amount of data need to be joined (Correct) A large amount of data needs to be joined.
A large portion of the table needs to be joined.

Quiz Start Time: 04:09 PM
Time Left 20 sec(s)

Question # 5 of 10 ( Start time: 04:15:07 PM ) Total Marks: 1
In context of nested-loop join, actual number of matching rows returned as a result of the join would be _____ of the order of tables
Select correct option:

Independent (Correct)

In case of nested-loop join, Inner table is accessed ____ for each qualifying row (or touple) in outer table Select correct option:

VU Topper RM

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

For More Help Contact What's app 03224021365

One Time (Correct)

Quiz Start Time: 04:09 PM
Time Left 10 sec(s)

Question # 10 of 10 ( Start time: 04:22:32 PM ) Total Marks: 1 In contrast to data mining, statistics is _____ driven. knowledge (Correct)

1 In data mining, initially you _____ what you are looking for. Select correct option:

Know
Don't know (Answer)
May or may not know
None of the given options

Quiz Start Time: 08:50 PM
Time Left 52 sec(s)

Question # 10 of 10 ( Start time: 09:00:41 PM ) Total Marks: 1
The optimizer uses a hash join to join two tables if they are joined using an equijoin and Select correct option:

Outer table has less number of rows
Inner table has less number of rows
Cardinality of tables is equal
Large amount of data needs to be joined (Answer)

The optimizer uses a hash join to join two tables if they are joined using an equijoin and

Select correct option:

Outer table has less number of rows

Inner table has less number of rows

Cardinality of tables is equal

Large amount of data needs to be joined

"More resources means proportionally less time for given amount of data". The statement refers to:

Vu Topper RM

بری صحبت سے تنہائی بہتر ہے اور تنہائی سے نیک صحبت بہتر ہے
For More Help Contact What's app 03224021365

Select correct option:

Scale-Up

Speed-Up

Size-up

Over-utilized system

If we apply Run Length Encoding on the input “11001100”, the output will be:

Select correct option:

41#40

14#04

12#02#12#02

21#20#21#20

The possible number of bitmap tables formed for corresponding field (column) None of the given options

“More resources means proportionally less time for given amount of data”. The statement refers to:

Scale-Up

Speed-Up

Size-up

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

For More Help Contact What’s app 03224021365

Vu Topper RM

Over-utilized system

“If resources increase in proportion to increase in data size, time is constant”. The statement refers to:

Scale-Up

Speed-Up

Size-up

Over-utilized system

In context of nested-loop join. actual number of matching rows returned as a result of the join would be _____ of the order of tables.

Dependent

Independent

Superset

Subset

“If resources increase in proportion to increase in data size, time is constant”. The statement refers to:

Scale-Up

Speed-Up

Size-up

Vu Topper RM

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

For More Help Contact What’s app 03224021365

Over-utilized system

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

Dense Index

Sparse Index

Inverted Index

None

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

Indexing

ETL

Joining

Dimentional Modeling

Parallelism can be exploited, if there is.

Symmetric multi-processors (SMP)

Sufficient I/O bandwidth

Underutilized or intermittently used CPUs

All of the above

Vu Topper RM

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

For More Help Contact What's app 03224021365

as the number of processors increase, the speedup should also increase thus we should have linear speedup. Which of the following is not one of the barriers to achive this linear speedup?

Amdhal law

Startup

No interference

Skew

Bitmap index is approperiate lists each term in the collection only once and then shows a list of all the documents that contain the given term.

Inverted Index

Bitmap index

Clustor index

Join index

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

A Dense Index

A Sparse Index

Vu Topper RM

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

For More Help Contact What's app 03224021365

An Inverted Index

None of These

1_In context of nested-loop join, actual number of matching rows returned as a result of the join would be _____ of the order of tables.

Independent

3 "If resources increase in proportion to increase in data size. Time is constant'.

The statement refers to:

Scale Up

4_ If every key in the data file is represented in the index file then it is called

Dense Index

6_ In nested-loop join case, if there are 'M' rows in outer table and 'N' rows in inner table, time complexity is o(MN)

7_ The goal of _____ is to look at as few blocks as possible to find matching records(s).

Indexing

Vu Topper PM

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

For More Help Contact What's app 03224021365

9_ If we apply Run Length Encoding on the input "11001100", the output will be: 21#20#21#20

10_ Which of the following is NOT one of variants of Nested-loop join?

Binary index nested-loop join

2_ _____ do not (typically) keep the index values in stored order

Hash based index

3_ if every key in the data is represented in the index file then it is called Dense index

6_ A join is identified by multiple tables in the _____ clause.

From

9_ As the number of processors increase, the speedup should also increase. thus we should have linear speedup. Which of the following is NOT the one of the barriers to achieve this linear speed-up?

Amdah'l Law not sure

Q1: in context of nested-loop join, actual number os matching rows returned as a result of the join would be _____ of the order of tables.

Vu Topper RM

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

For More Help Contact What's app 03224021365

Independent

Q3: If resources increase in proportion to increase in data size. time is constant'.

The statement refers to:

Scale-Up

Q4: If every key in the data file is represented in the index file then it is called? Dense Index

Q7: The goal of_____ is to look at as few blocks as possib le to find the matching records(s).

Indexing

Q8: Parallelism can be exploited, if there is.

All of the given options

Q13: if every key the data file is represented in the index file then it is called. Dense Index

Q16: A join is identified by multiple tables in the_____ clause.

From

Q17: Parallelism can exploited, if there is

All of the given options

VuTopperRM

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

For More Help Contact What's app 03224021365

Q19: As the number of processors increase, the speedup should also increase. Thus we should have linear speedup. Which of the following is NOT one of the barriers to achieve this linear speed-up?

Amdahl' Law

Q22: _____ lists each term in the collection only once and then shows a list of all the documents the contain the given term.

Inverted index

Q23: "More resources means proportionally less time for given amount of data".

The statement refers to:

Speed-UP

Q27: A_____ index, if fits in the memory, costs only one disk I/O access to locate a record given a key.

Dense Index

Q28: In context of nested-loop join, actual number of matching rows returned as a result of the join would be ___ of the order of tables

Independent

Vu Topper RM

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

For More Help Contact What's app 03224021365

Q29: _____ refers to " Parallelexectution of single data operation across multiple partitions of data" Data Parallelism.

A join is identified by multiple tables in the _ FROM ___ clause

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

Question # 3 of 10 ( Start time: 01:04:49 PM ) Total Marks: 1
Which of the following is NOT one of the top-10 mistakes that should be avoided during DWH development?
Select correct option:
Not interacting directly with end user
**Not being an accommodating person (Page No. 316)**
Isolating IT support people from business users
Training the users with dummy data and considering it success

3. In _____ phase of kimball's approach, we identify the components needed now and in future.

1. Project Planning
2. **Requirements Definition (Pg 294)**
3. Technical Architecture Design
4. Maintenance & Growth

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

**Ans: Not being an accommodating person**
7.Which of the following is NOT one of the top-10 mistakes that should be avoided during DWH development?
**Ans: Not being an accommodating person**

- *2_ Which of the following is NOT one of the three parallel tracks in Kimball's approach?*
- ***Lifecycle Maintenance track***

بری صحبت سے تنہائی بہتر ہے اور تنہائی سے نیک صحبت بہتر ہے
For More Help Contact What's app 03224021365
Vu Topper RM

Vu Topper RM

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

For More Help Contact What's app 03224021365

***Clustering is unsupervised learning and classification is supervised learning _ pg 270***

- 
- •
- 1_ As per Bill Inmost, a data warehouse, in contrast with classical applications is:
- Data driven _ pg285
- 2_ Which of the following is NOT one of the three parallel tracks in Kimball's approach? Lifecycle Maintenance track

  The optimizer uses a hash join to join two tables if they are joined using an equijoin and Select correct option:
- Large amount of data need to be joined (Correct) A large amount of data needs to be joined. A large portion of the table needs to be joined.
- Quiz Start Time: 04:09 PM Time Left 20 sec(s)
- Question # 5 of 10 ( Start time: 04:15:07 PM ) Total Marks: 1
  In context of nested-loop join, actual number of matching rows returned as a result of the join would be _____ of the order of tables Select correct option:
- Independent (Correct)

  In case of nested-loop join, Inner table is accessed ____ for each qualifying row (or touple) in outer table Select correct option:
- One Time (Correct)
- Quiz Start Time: 04:09 PM Time Left 10

sec(s)

- Question # 10 of 10 ( Start time: 04:22:32 PM ) Total Marks: 1 In contrast to data mining, statistics is _____ driven.
  knowledge (Correct)

- 1 In data mining, initially you _____ what you are looking for. Select correct option:
- Know
  Don't know (Answer)
  May or may not know
  None of the given options
- Quiz Start Time: 08:50 PM Time Left 52 sec(s)
- Question # 10 of 10 ( Start time: 09:00:41 PM ) Total Marks: 1
  The optimizer uses a hash join to join two tables if they are joined using an equijoin and Select correct option:
- Outer table has less number of rows Inner table has less number of rows
  Cardinality of tables is equal

Vu Topper RM

Large amount of data needs to be joined (Answer)

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

Select correct option:

Not interacting directly with end user

**Not being an accommodating person**

Isolating IT support people from business users

Training the users with dummy data and considering it success

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

Select correct option:

Discrete (Correct)

Isolated

Continuous

Distinct

Vu Topper RM

*2_ Which of the following is NOT one of the three parallel tracks in Kimball's approach?*

***Lifecycle Maintenance track***

Vu Topper RM

**For More Help Contact What's app 03224021365**