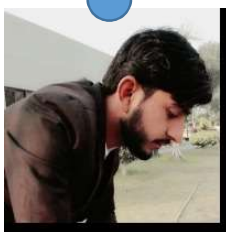
AL-JUNAID INSTITUTE GROUP

CS614-Data Warehousing

Solved MCQS for Final terms papers

Solved by JUNAID MALIK and Team



AL-Junaid Institude

Contact no: 03041659294

VU LMS HELP vulmshelp.com

LMS Handling

Subject Enrollment

1) Online classes available
HTML, CSS, Javascript
Phyton..etc
2) Solved Graded Activities
- Assignment's
- Quizze's
- GDB's

3) Sloved Quiz File's
4) Short Note's
5) Past paper's &
Current paper's

Final project CS619

1) SRS
(Software Requirement's Specification)
2) DD
(Design Document)
3) Test phase + viva
4) Viva preparation
5) Final Deliverable

Website Link
vulmshelp.com

Gmail
junaidfazal08@gmail.com

ALL answers are verified if found any mistake then Correct ACCORDINGLY

AL-JUNAID INSTITUTE GROUP

1. Classification consist of examining the properties of a newly presented observation and assigning it to a predefined_____
   - Object
   - Class
   - Container
   - Subjects
2. In context of clustering the term "distance" means
   - The relation of a record with corresponding records in child table
   - None of these
   - The difference between the primary keys of two records
   - Similarity dissimilarity of record
3. In contrast to data mining statistics is_____driven
   - Discovery
   - Knowledge
   - Database
   - Assumption
4. _____is the technique in which existing heterogeneous segments are reshuffled, relocated into homogenous segments.
   - Clustering
   - Partitioning
   - Aggregation
   - Segmentation
5. In context of data mining definition, the term "Value" means.
   - Importance of hidden pattern discovered
   - The primary key of table
   - The index location of record
   - Numerical or string measure assigned to an attributes
6. As oppose to the outcome of classification, estimation deal with____valued outcomes
   - Discrete
   - Continuous
   - Isolated
   - distinct
7. _____ incorporates the con cept of product quality, process control, quality assurance, and quality improvement.
   - **Total Quality Management**
   - Intrinsic Data Quality Management
   - Realistic Data Quality Management
   - Strong Data Quality Management

8. The extent to which data is in appropriate languages, symbols and units, and the definitions are clear is known as _____.
    - **Interpretability**
    - Uniqueness
    - Accessibility
    - Consistency
9. The degree to which values are present in the attributes that require them is known as _____.
    - **Completeness**
    - Uniqueness
    - Accessibility
    - Consistency
10. The _____ dimension represents data correctness.

- **Free-of-error**
- Completeness
- Consistency
- Correctness

11. 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 a corresponding ROWID**
    - Every indexed data value and pointer to next level block
    - Every indexed data value and pointer to root block
12. Data is the _____ on which a Data Warehouse (DWH) runs.
    - **Fuel**
    - Element
    - Component
    - Entity
13. Mining Multi dimensional databases allow users to:
    - Categorize the data
    - Summarize the data
    - Analyze the data
    - All of the given

12. In context of data parallelism to get a speed-up of N with N partitions, it must be ensured that:
    - There are enough computing resources
    - Query-coordinator is very fast as compared to query servers
    - Work done in each partition almost same
    - **All of the given options**
13. Which of the following is not an activity of Data Quality Analysis Project?
    - "Define"
    - "Measure"
    - "Analyze"
    - **"Compression"**
14. Which of the following is not a Data Quality Validation Technique?
    a. Referential Integrity
    b. Using Data Quality Rules
    c. Data Histograming
    d. Indexes
15. One of the preconditions to decide about operations to be parallelized is that
    a. Operation can be implemented independent of each other
    b. Output of one operation becomes input of other
    c. Operations share same memory location
    d. Operations share same namespace
16. _____do not (typically) keep the index values in sorted order
    a. Dense index
    b. Sparse index
    c. B-Tree Index
    d. Hash Based index
17. Parallelism can be exploited, if there is :
    - Symmetric multi processors (SMP)
    - Sufficient I/O bandwidth
    - Underutilized or intermittently used CPUs
    - **All of the given options**
18. Which of the following is NOT one of the parallel hardware architectures?
    - Symmetric Multi-Processing
    - Massively Parallel Processing
    - Non-uniform Memory Access
    - **Shared Memory**

AL-JUNAID INSTITUTE GROUP

19. Two interesting examples of quality dimensions that can make use of the min operator are _____.
    - **Believability and appropriate amount of data**
    - Believability and Consistency
    - Believability and Redundancy
    - Reliability and appropriate amount of data
20. 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
    - Start-up
    - **No Interference**
    - Skew
21. In _____index, the ith bit is set to "1" if the ith row of the base table has the value for the indexed column
    - Inverted index
    - **Bitmap index**
    - Cluster index
    - Join index
22. _____ 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
    - Cluster index
    - Join index
23. The exact formula for Speed-up is:
    - **(Time on Serial Processor) / (Time on parallel processors)**
    - (Time on Serial Processor) * (Time on parallel processors)
    - (Time on Serial Processor) + (Time on parallel processors)
    - (Time on Serial Processor) - (Time on parallel processors)
24. _____is the degree to which data accurately reflects the real world object that the data represents
    - **Intrinsic data quality**
    - Realistic data quality
    - Strong data quality
    - Weak data quality

25. Assume a company with a multi-million row customer table i.e. n rows. Checking for Referential Integrity (RI), using a smart technique with some kind of a tree data structure would require _____ time.
    - **O(log n)**
    - O(n)
    - O(1)
    - None of the given
26. Which of the following is NOT one of the variants of Nested-loop join?
    - Naive nested-loop join
    - Indexed nested-loop join
    - Temporary index nested-loop join
    - **Binary index nested-loop join**
27. "More resources means proportionally less time for given amount of data" that statement refers to
    - Scale-Up
    - **Speed-Up**
    - Size-Up
    - Over-Utilized system
28. The optimizer uses a hash join to join two tables if they are joined using an equijoin and
    - outer table has less number of rows
    - inner table has less number of rows
    - cardinality of table is equal
    - **large amount of data needs to be joined**
29. "If resources increase in proposition to increase in data size, time is constant". The statement refers to
    - **Scale-up**
    - Speed-up
    - Size-up
    - Over-utilized system
30. If a product meets formally defined "requirement specifications", yet fails to be a quality product form the customer's perspective, this means the requirements were _____.
    - **Defective**
    - Unclear
    - Unrefined
    - Undefined

AL-JUNAID INSTITUTE GROUP

31. _____is the extent to which data is regarded as true and credible.
    - **Believability**
    - Completeness
    - Accessibility
    - consistency
32. Which is not a/an data quality validation technique?
    - **Consistency integrity**
    - Referential integrity
    - Attribute domain
    - Using data quality rules
33. Which of the following is not an "Orr's law of data quality"?
    - Data that Is not used is cannot be correct
    - Data quality is 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**

34. _____ is known as state of being only one of its kind or being without an equal or parallel.
    - Completeness
    - **Uniqueness**
    - Accessibility
    - Consistency
35. Which is not a/an characteristics of data quality?
    - **Reliability**
    - Uniqueness
    - Accessibility
    - Consistency
36. If every key in the data file is represented in the index file then it is called
    - **Dense Index**
    - Sparse Index
    - Inverted Index
    - A Multi level Sparse Index
37. One of the main reasons for the failure of DWH deployment is ____
    - **Data quality**
    - Data integrity
    - Data duplication
    - Data anomaly

38. The _____ operator is conservative in that it assigns to the dimension an aggregate value no higher than the value of its weakest data quality indicator.
    - **Min**
    - Max
    - Min and Max
    - None of given
39. ____ is making all efforts to increase effectiveness in meeting and deficiency in meeting except customer expectations
    - Quality assurance
    - **Quality improvement**
    - Quality maintenance
    - Quality establishment
40. Most DWH implantations today do not use ____ enforced by the database, but as TQM method improved overall data quality and database optimizers.
    - Consistency integrity
    - **Referential integrity**
    - Attribute domain
    - Using data quality rules
41. If a task takes "T" time units to execute on a single data item, then execution of the Task on "N" data items will take _____time units.
    - **N*T**
    - N/T
    - N+T
    - N-T
42. An optimized structure which is built primarily for retrieval, with update being only a secondary consideration is
    - OLTP
    - OLAP
    - DSS
    - **Inverted Index**
43. _____ refers to "Parallel execution of single data operation across multiple partitions of data"
    - Hardware parallelism
    - Software parallelism
    - **Data parallelism**
    - Operational parallelism

44. _____ in a database or data warehouse has no actual value, it only has potential value.
    - **Data**
    - Entity
    - Flat tables
    - Data marts

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

46. A join is identified by multiple tables in the _____ clause
    - **FROM**
    - SELECT
    - GROUP BY
    - SORT BY

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

48. _____ is a/an measure of how current or up to date the data is
    - **a. Timeliness page 185**
    - b. Completeness
    - c. Accessibility
    - d. Consistency

49. In context of data parallelism, the work done by query processor should be:
    - **a. Almost zero Goggle**
    - b. Maximum
    - c. Pipelined
    - d. Filtered across partitions

50. In context of joining tables, the join condition is specified in _____clause
    - FORM
    - SELECT
    - **WHERE**
    - GROUP BY

51. A _____ index, if fits in the memory, costs only one disk I/O access to locate a record given a key.
    - **Dense**
    - Sparse
    - B-Tree
    - Hash
52. _____ index uses even less space than _____ index, but the block has to be searched, even for unsuccessful searches.
    - Dense, sparse
    - **Sparse, dense**
    - Dense, inverted
    - Sparse, inverted
53. _____ is the degree of utility and value the data has to support the enterprise processes that enable accomplishing enterprise objectives.
    - Intrinsic Data Quality
    - **Realistic Data Quality**
    - Strong Data Quality
    - Weak Data Quality
54. _____ is a system of activities that assures conformance of product to pre-established requirements.
    - **Quality assurance**
    - Quality improvement
    - Quality maintenance
    - Quality establishment
55. 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
56. In context of bitmap index, the length of the bit vector is:
    - a. The possible number of domain values in corresponding field (column)
    - **b. The number of records in the base table**
    - c. The possible number of bitmap tables formed for corresponding field (column)
    - d. None of the given options

AL-JUNAID INSTITUTE GROUP

57. The _____ operator proves useful in more complex metrics applicable to the dimensions of timeliness and accessibility.

- **Max**
- Min
- Min and Max
- None of given

58. 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)

59 Assume a company with a multi-million row customer table i.e. n rows.

Checking for Referential Integrity (RI), using a naïve approach would take _____ time.

- **O(n)**
- O(1)
- O(log n)
- None of the given

60. Bitmap index is appropriate for:

- **Low cardinality data**
- High cardinality data
- Clustered data
- Aggregated data

Question No: 1

A data warehouse may include

- ► **Legacy systems**
- ► Only internal data sources
- ► Privacy restrictions
- ► Small data mart

Question No: 2

De-Normalization normally speeds up

- ► **Data Retrieval (page no 51)**
- ► Data Modification
- ► Development Cycle

- ► Data Replication

Question No: 3

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

- ► **Common Column Values page no 54**
- ► Common Row Values
- ► Different Index Values
- ► Value resulted by ad-hoc query

Question No: 4

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

- ► **File page no 79**
- ► Application
- ► Aggregate
- ► Database

Question No: 5

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

- ► **One [page no 233]**
- ► Two
- ► lg (n)
- ► n

Question No: 6

All data is _____ of something real.

IAn Abstraction

IIA Representation

Which of the following option is true?

- ► I Only
- ► II Only
- ► **Both I & II (P# 181)**
- ► None of I & II

Question No: 7

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

- ► **Pipeline Parallelism page no 214**
- ► Overlapped Parallelism
- ► Massive Parallelism
- ► Distributed Parallelism

Question No: 8

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

- ► **Skew in Partition (P # 218)**
- ► Pipeline Distribution
- ► Distributed Distribution
- ► Uncontrolled Distribution

Question No: 9

The goal of ideal parallel execution is to completely parallelize those parts of a computation that are not constrained by data dependencies. The smaller the portion of

the program that must be executed _____, the greater the scalability of the computation.

- ► None of these
- ► **Sequentially page no 204**
- ► In Parallel
- ► Distributed

Question No: 10

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

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 no 249**
- ► Non-Exploratory
- ► Computer Science

Question No: 12

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 no 254**
- ► **OLAP**
- ► **DSS**
- ► **DWH**

Question No: 13

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

- ► **Clustering [page no 264]**
- ► **Aggregation**
- ► **Segmentation**
- ► **Partitioning**

Question No: 14

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

- ► **Pearson correlation is the only technique**
- ► **Euclidean distance is the only technique**
- ► **Both Pearson correlation and Euclidean distance [page no 270**
- ► **None of these**

Question No: 15

► One-way Clustering (P# 271)
► Bi-clustering
► Pearson correlation
► Euclidean distance

Question No: 16

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

► Designing
► Development
► Analysis
► Implementation

Question No: 17

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

► Tools
► Industry (P# 320)
► Software
► None of these

Question No: 18

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

► Increasing
► Decreasing (P# 215)
► Maintaining
► None of these

Question No: 19

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

► Silver Bullet (pg#375)
► Golden Bullet
► Suitable Hardware
► Compatible Product

Question No: 20

Focusing on data warehouse delivery only often end up _____.

► Rebuilding (pg# 315)
► Success
► Good Stable Product
► None of these

Question No: 21

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

► Cotton-growing (p # 330)
► Rice-growing
► Weapon Producing

AL-JUNAID INSTITUTE GROUP

Question No: 22

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

- **► Data profiling (P# 439)**
- **► Data Anomaly Detection**
- **► Record Duplicate Detection**
- **► None of these**

Question No: 23

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 (p # 19)**
- **► Two Direction**
- **► None of these**

Question No: 24

DSS queries do not involve a primary key

- **► True (p # 21)**
- **► False**

Question No: 25

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

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

Question No: 26

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

- **► OLE DB (p # 373)**
- **► OLAP**
- **► OLTP**
- **► Data Warehouse**

Question No: 27

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 (p #373)**
- **► Documentations**
- **► Guideline**

Question No: 28

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 (p # 419)**
- **► Rolled back**

AL-JUNAID INSTITUTE GROUP

Question No: 29

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

- **► Execution of package (p # 419)**
- **► Creation of package**
- **► Connection of package**

Question No: 30

To judge effectiveness we perform data profiling twice.

- **► One before Extraction and the other after Extraction**
- **► One before Transformation and the other after Transformation [p # 441**
- **► One before Loading and the other after Loading**

Question No: 31

During the development phase we should follow standards for:

- ➢ Naming Conventions
- ➢ Calculation
- ➢ Libraries
- ➢ **All of the given options (Page 307)**

Question No: 32

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

- ➢ Project Planning
- ➢ **Requirement definition (Page 299)**
- ➢ Requirement verification
- ➢ Requirement validation

Question No 33

Bill Inmon argues that requirements are well understood only after

- ➢ They are documented
- ➢ They extracted and verified
- ➢ **Data warehouse is populated (Page 285)**
- ➢ All schemes are defined

AL-JUNAID INSTITUTE GROUP

Question # 34

Waterfall model is appropriate when

- When budget is low
- When the deadline is strict
- When resources are limited
- **Requirements are clearly defined (Page 284)**

Question # 35

A bottom up implementation approach is useful when

- **Technology is matured and well understands (page 283)**
- Organization cannot implement latest technologies
- The business objectives are very much clear
- All requirements are well documented

Question # 36

Technical architecture design specifies the:

- **Project schedule (Page 300)**
- Minimum project completion time
- Required Components
- Umbrella activities

Question # 37

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

- **Wester man (Page 285)**
- Bill Inmon
- Kimball
- Bohnlein

Question#38

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

- **Data driven (Page 285)**
- Resource driven
- Requirement driven
- Time Sensitive

Question # 39

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

- Deployment
- **Operation (Page 288)**
- Enhancement
- Maintenance

AL-JUNAID INSTITUTE GROUP

Question # 40

Following the coding standards during development helps to:

- **Minimize the future rework (Page 307)**
- Verify the requirements
- To refine the project scope
- To establish the test cases

Question # 41

As per Kimball, _____ is the main operational process

- Requirement extraction
- Global design
- **Business process (Page 285)**
- Scheme design

Question # 42

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

- Design
- **Prototype (Page 287)**
- Deployment
- Operation

Question #43

Waterfall is a/an _____ model

- Iterative
- **Simple linear sequential (Page 284)**
- Object oriented
- Rapid development

Question # 44

One of the drawbacks of waterfall model is that:

- **Customers can not review the product during development (Page 284)**
- It does not work when the resources are limited
- It does not define the project timeline/schedule
- All of the given options

AL-JUNAID INSTITUTE GROUP

Question # 45

Spiral model is _____

- Sequence of waterfall model
- Risk oriented model
- An iterative model
- **All of the given (Page 284)**

Question # 46

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

- Statistics
- **Knowledge discovery in database (Page 249)**
- Clustering
- Information cleansing

Question # 47

Data mining is all about

- Knowledge discovery in database
- Finding hidden patterns in data
- Finding hidden relations in data
- **All of the given options (Page 249)**

Question # 48(Total Marks: 1) Select correct option:

In context of clustering, the term "distance" means

- **Similarity/dissimilarity of records (Page 272)**
- The difference between the primary keys of two records
- The relation of a record with corresponding record in child table
- None of the given options

Question # 48

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

- **Development -> QA -> Production (Page 314)**
- Production -> QA ->Development
- Development -> Production -> QA
- Production->Development -> QA

Question # 49 (Total Marks: 1) Select correct option:

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

- **Full time project manager assigned to the project (Page 318)**
- QA team ready for testing the code
- Design of complete project before implementation
- Components of complete project before implementation

AL-JUNAID INSTITUTE GROUP

Question # 50

Normally it is recommended to have

- Different servers for development and production environment
- Same server for development and production environment
- **Interference while having different database environments on a single server (Page 313)**
- All of the given options

Question # 51

Identify the TRUE statement:

- **Clustering is unsupervised learning and classification is supervised learning (Page 270)**
- Clustering is supervised learning and classification is unsupervised learning
- Both clustering and classification are unsupervised learning
- Both clustering and classification are supervised learning

Question #51

Mining multi-dimensional databases allow users to:

- Categorize the data
- **All of the given options (250)**
- Analyze the data
- Summarize the data

Question # 52

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

- only
- Only
- Only
- **(I), (II) and (III) (Page 300)**

Question #53

Improper documentation results the problem(s) like:

- Maintenance issue
- **New developers unable to configure already existing code (Page 314)**
- Lot of time required for enhancing the code
- All of the given options

AL-JUNAID INSTITUTE GROUP

Question # 54

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

- Scope
- Feasibility
- Resources available
- **The potential impact or value to the business (Page 297)**

Question # 55

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

- Requirement elicitation
- Project Planning
- **Project management (Page 289)**
- Deployment

Question # 56

"What means what". The phrase refers to:

- **Meta data (Page 338)**
- External data
- Transformed data
- Internal representations

Question # 57

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

- Not interacting directly with end user
- **Not being an accommodating person (Page 316)**
- Isolating IT support people from business users
- Training the users with dummy data and considering it success

Question # 58

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

- Requirement definition
- **Architectural design (Page 300)**
- Product development
- Analytical application development

AL-JUNAID INSTITUTE GROUP

Question # 59

Which of the following is NOT one of the advantages of changed data capture (CDC) technique?

- Flat files are not required
- **Limited query interface is required for data extraction (Page 152)**
- No incremental on-line I/O required for log tape
- Extraction of changed data occurs immediately

Question #60

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

- Check Marks
- **Time Stamps (Page 150)**
- Just-In-Time
- Real Time extraction

Question # 61

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

- **Syntactically dirty data (Page 160)**
- Semantically dirty data
- Coverage anomaly
- Extraction issue

Question #62

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

- Takes place on the data warehouse server
- **Takes place on a separate transformation server (Page 147)**
- Depends on the nature of the source database
- Does not take place

Question #63

If one or more records in a relational table do not satisfy one or more integrity constraint, then the data:

- Is syntactically dirty
- **Is semantically dirty (Page 160)**
- Has Coverage anomaly
- Has extraction issue

AL-JUNAID INSTITUTE GROUP

Question # 64

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

- Conversion
- Summarization
- Enrichment
- **Full Data Refresh (Page 135)**

Question # 65

Robotic libraries are needed for _____

- Cubes
- Data marts
- **Data warehouse (Page 131)**
- Aggregates

Question #66

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

- Aggregates don't change in real time
- **Transformation of extracted data is difficult (Page 149)**
- Identifying the recently modified data may be difficult
- Source systems may not support extraction of changed aggregates

Question # 67

In case of multiple sources for the same data element, we need to prioritize the source systems per element bases, the process is called:

- **Ranking (Page 143)**
- Prioritization
- Element Selection
- Measurement Selection

Question # 68

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

- Buffer table
- **Change table (Page 149)**
- Checkmark table
- Change control table

AL-JUNAID INSTITUTE GROUP

Question # 69

In data mining, initially you_____what you are looking for.

- Know
- **Don't Know (Page 250)**
- May or may not know
- None of the given options

Question # 70

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

- End-user interviews and re-interviews
- Source system cataloguing
- Definition of key performance indicators
- **System vision development (Page 287)**

Question # 71

A top down implementation approach is useful when

- **Technology is matured and well understands (Page 283)**
- Organization cannot implement latest technologies
- The business objectives are very much clear
- All requirements are well documented

Question # 72

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

- Goal Driven
- Data Driven
- User Driven
- **System Driven (Page 283)**

Question # 73

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

- The primary key of table
- Index location of the record
- **Importance of hidden patterns discovered (Page 250)**
- Numerical of string measure assigned to an attribute

AL-JUNAID INSTITUTE GROUP

Question # 74

Data mining is all about

- Knowledge discovery in database
- Finding hidden patterns in data
- Finding hidden relations in data
- **All of the given options (Page 249)**

Question # 75

In context of clustering, the term "distance" means

- **Similarity/dissimilarity of records (Page 272)**
- The difference between the primary keys of two records
- The relation of a record with corresponding record in child table
- None of the given options

Question # 76

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

- **Development -> QA -> Production (Page 314)**
- Production -> QA ->Development
- Development -> Production -> QA
- Production->Development -> QA

Question # 77

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

- **Full time project manager assigned to the project (Page 318)**
- QA team ready for testing the code
- Design of complete project before implementation
- Components of complete project before implementation

Question # 78

Normally it is recommended to have

- Different servers for development and production environment
- Same server for development and production environment
- **Interference while having different database environments on a single server (Page 313)**
- All of the given options

Question # 79

Identify the TRUE statement:

- **Clustering is unsupervised learning and classification is supervised learning (Page 270)**
- Clustering is supervised learning and classification is unsupervised learning
- Both clustering and classification are unsupervised learning
- Both clustering and classification are supervised learning

AL-JUNAID INSTITUTE GROUP

Question # 80

Mining multi-dimensional databases allow users to:

- Categorize the data
- Analyze the data
- Summarize the data
- **All of the given options (Page 250)**

Question # 81

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

- only
- Only
- Only
- **(I), (II) and (III) (Page 300)**

Question # 82

Improper documentation results the problem(s) like:

- Maintenance issue
- **New developers unable to configure already existing code (Page 314)**
- Lot of time required for enhancing the code
- All of the given options

Question # 83

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

- Scope
- Feasibility
- Resources available
- **The potential impact or value to the business (Page 297)**

Question # 84

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

- Requirement elicitation
- Project Planning
- **Project management (Page 289)**
- Deployment

AL-JUNAID INSTITUTE GROUP

Question # 85

"What means what". The phrase refers to:

- **Meta data (Page 338)**
- External data
- Transformed data
- Internal representations

Question # 86

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

- Not interacting directly with end user
- **Not being an accommodating person (Page 316)**
- Isolating IT support people from business users
- Training the users with dummy data and considering it success

Question # 87

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

- Requirement definition
- **Architectural design (Page 300)**
- Product development
- Analytical application development

Question # 88

Which of the following is NOT one of the advantages of changed data capture (CDC) technique?

- Flat files are not required
- **Limited query interface is required for data extraction (Page 152)**
- No incremental on-line I/O required for log tape
- Extraction of changed data occurs immediately

Question # 89

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

- Check Marks
- **Time Stamps (Page 150)**
- Just-In-Time
- Real Time extraction

AL-JUNAID INSTITUTE GROUP

Question # 90

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

- **Syntactically dirty data (Page 160)**
- Semantically dirty data
- Coverage anomaly
- Extraction issue

Question # 91

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

- Takes place on the data warehouse server
- **Takes place on a separate transformation server (Page 147)**
- Depends on the nature of the source database
- Does not take place

Question # 92

If one or more records in a relational table do not satisfy one or more integrity constraint, then the data:

- Is syntactically dirty
- **Is semantically dirty (Page 160)**
- Has Coverage anomaly
- Has extraction issue

Question # 93

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

- Conversion
- Summarization
- Enrichment
- **Full Data Refresh (Page 135)**

Question # 94

Robotic libraries are needed for _____

- Cubes
- Data marts
- **Data warehouse (Page 131)**
- Aggregates

AL-JUNAID INSTITUTE GROUP

Question #95

During the development phase we should follow standards for:

- Naming Conventions
- Calculation
- Libraries
- All of the given options (Page 307)

Question # 96

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

- Project Planning
- Requirement definition (Page 299)
- Requirement verification
- Requirement validation

Question #97

Bill Inmon argues that requirements are well understood only after

- They are documented
- They extracted and verified
- Data warehouse is populated (Page 285)
- All schemes are defined

Question # 98

Waterfall model is appropriate when

- When budget is low
- When the deadline is strict
- When resources are limited
- Requirements are clearly defined (Page 284)

Question # 99

A bottom up implementation approach is useful when

- Technology is matured and well understands (page 283)
- Organization cannot implement latest technologies
- The business objectives are very much clear
- All requirements are well documented

Question # 100

Technical architecture design specifies the:

- Project schedule (Page 300)
- Minimum project completion time
- Required Components
- Umbrella activities

AL-JUNAID INSTITUTE GROUP

Question # 101

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

- Westerman (Page 285)
- Bill inmon
- Kimball
- Bohnlein

Question #102

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

- Data driven (Page 285)
- Resource driven
- Requirement driven
- Time Sensitive

Question #103

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

- Deployment
- Operation (Page 288)
- Enhancement
- Maintenance

Question #104

Following the coding standards during development helps to:

- Minimize the future rework (Page 307)
- Verify the requirements
- To refine the project scope
- To establish the test cases

Question # 105

As per Kimball, _____ is the main operational process

- Requirement extraction
- Global design
- Business process (Page 285)
- Scheme design

AL-JUNAID INSTITUTE GROUP

Question #106

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

- Design
- Prototype (Page 287)
- Deployment
- Operation

Question #107

Waterfall is a/an _____ model

- Iterative
- Simple linear sequential (Page 284)
- Object oriented
- Rapid development

Question #108

One of the drawbacks of waterfall model is that:

- Customers can not review the product during development (Page 284)
- It does not work when the resources are limited
- It does not define the project timeline/schedule
- All of the given options

Question #109

Spiral model is _____

- Sequence of waterfall model
- Risk oriented model
- An iterative model
- All of the given (Page 284)

Question #110

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

- Statistics
- Knowledge discovery in database (Page 249)
- Clustering
- Information cleansing

Question #111

In data mining, initially you_____what you are looking for.

- Know
- Don't Know (Page 250)
- May or may not know
- None of the given options

AL-JUNAID INSTITUTE GROUP

Question #112

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

- End-user interviews and re-interviews
- Source system cataloguing
- Definition of key performance indicators
- System vision development (Page 287)

Question #113

A top down implementation approach is useful when

- Technology is matured and well understands (Page 283)
- Organization cannot implement latest technologies
- The business objectives are very much clear
- All requirements are well documented

Question #114

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

- Goal Driven
- Data Driven
- User Driven
- System Driven (Page 283)

Question #115

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

- The primary key of table
- Index location of the record
- Importance of hidden patterns discovered (Page 250)
- Numerical of string measure assigned to an attribute

Question #115

Data mining is all about

- Knowledge discovery in database
- Finding hidden patterns in data
- Finding hidden relations in data
- All of the given options (Page 249)

Question #116

In context of clustering, the term "distance" means

- Similarity/dissimilarity of records (Page 272)
- The difference between the primary keys of two records
- The relation of a record with corresponding record in child table
- None of the given options

AL-JUNAID INSTITUTE GROUP

Question #117

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

- Development -> QA -> Production (Page 314)
- Production -> QA ->Development
- Development -> Production -> QA
- Production->Development -> QA

Question # 118

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

- Full time project manager assigned to the project (Page 318)
- QA team ready for testing the code
- Design of complete project before implementation
- Components of complete project before implementation

Question #119

Normally it is recommended to have

- Different servers for development and production environment
- Same server for development and production environment
- Interference while having different database environments on a single server (Page 313)
- All of the given options

Question # 120

Identify the TRUE statement:

- Clustering is unsupervised learning and classification is supervised learning (Page 270)
- Clustering is supervised learning and classification is unsupervised learning
- Both clustering and classification are unsupervised learning
- Both clustering and classification are supervised learning

Question #121

Mining multi-dimensional databases allow users to:

- Categorize the data
- Analyze the data
- Summarize the data
- All of the given options (Page 250)

AL-JUNAID INSTITUTE GROUP

Question # 122

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

- only
- Only
- Only
- (I), (II) and (III) (Page 300)

Question #123

Improper documentation results the problem(s) like:

- Maintenance issue
- New developers unable to configure already existing code (Page 314)
- Lot of time required for enhancing the code
- All of the given options

Question #124

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

- Scope
- Feasibility
- Resources available
- The potential impact or value to the business (Page 297)

Question # 125

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

- Requirement elicitation
- Project Planning
- Project management (Page 289)
- Deployment

Question # 126

"What means what". The phrase refers to:

- Meta data (Page 338)
- External data
- Transformed data
- Internal representations

AL-JUNAID INSTITUTE GROUP

Question #127

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

- Not interacting directly with end user
- Not being an accommodating person (Page 316)
- Isolating IT support people from business users
- Training the users with dummy data and considering it success

Question # 128

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

- Requirement definition
- Architectural design (Page 300)
- Product development
- Analytical application development

Question # 129

Which of the following is NOT one of the advantages of changed data capture (CDC) technique?

- Flat files are not required
- Limited query interface is required for data extraction (Page 152)
- No incremental on-line I/O required for log tape
- Extraction of changed data occurs immediately

Question # 130

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

- Check Marks
- Time Stamps (Page 150)
- Just-In-Time
- Real Time extraction

Question # 131

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

- Syntactically dirty data (Page 160)
- Semantically dirty data
- Coverage anomaly
- Extraction issue

AL-JUNAID INSTITUTE GROUP

Question #132

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

- Takes place on the data warehouse server
- Takes place on a separate transformation server (Page 147)
- Depends on the nature of the source database
- Does not take place

Question #133

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

- Conversion
- Summarization
- Enrichment
- Full Data Refresh (Page 135)

Question # 134

Robotic libraries are needed for _____

- Cubes
- Data marts
- Data warehouse (Page 131)
- Aggregates

Question # 135

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

- Aggregates don't change in real time
- Transformation of extracted data is difficult (Page 149)
- Identifying the recently modified data may be difficult
- Source systems may not support extraction of changed aggregate

Question #136

In case of multiple sources for the same data element, we need to prioritize the source systems per element bases, the process is called:

- Ranking (Page 143)
- Prioritization
- Element Selection
- Measurement Selection

Question # 137

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

- Buffer table
- Change table (Page 149)
- Checkmark table
- Change control table

AL-JUNAID INSTITUTE GROUP

Question #138

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

- Daily Sale
- Date of Birth (confirm not) (Page 342)
- Air pressure
- None of the given options

Question # 139

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

- System analysis
- Scope definition
- Requirement configuration
- Requirements wrap-up presentation (Page 297)

Question # 140

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

- Obtain resources
- Establish the preliminary scope and justification
- Assess organization's readiness for a data warehouse initiative
- All of the given options (Page 290)

Question # 150

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

- Bill Inmon
- Ralph Kimball
- Böhnlein and Ulbrich-vom (Page 285)
- Westerman

Question # 151

In contrast to data mining, statistics is _____ driven.

- Assumption
- Knowledge
- human (Page 255)
- Database

Question #152

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

- Linear
- Quadratic
- Exponential (Page 15)
- Logarithmic

AL-JUNAID INSTITUTE GROUP

Question # 153

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

- Requirement preplanning
- Business requirements collection
- Post collection (Page 294)
- None of the given options

Question # 154

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

- Tables should be properly de-normalized
- Proper partitioning technique should be used
- At least star or snow flake schema should be implementing
- They should get information in timely manner (Page 314)

Question # 155

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

- Education
- Technical education
- Program Support
- Interface Deployment (Page 309)

Question #156

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

- Employee Name (Page 342)
- Employee Date of Birth
- Employee Blood Group
- All of the given options

Question # 157

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

- Project planning
- Business requirement definition
- Deployment (Page 308)
- Maintenance and growth

Question # 158

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

- Pre-joining
- Splitting Tables
- Adding Redundant Columns
- ER Modeling (Page 52)

AL-JUNAID INSTITUTE GROUP

Question # 160

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

- Timestamps
- Partitioning
- Triggers
- Dimensional Modeling (Page 150)

Question #161

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

- Lifecycles technology track
- Lifecycle Data track
- Lifecycle Analytical Applications track (Page 306)
- Lifecycle Maintenance track

Question # 162

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

- Understand the audience, don't overwhelm
- Train after delivery of the data and analytic applications
- Postpone education, if DWH not ready
- All of the given options (Page 308)

Question # 163

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

- Database Design is complete (Page 307)
- A subset of historical data has been loaded
- The data access tools and metadata are installed
- All of the given options

Question #164

During the development phase we should follow standards for:

- Naming Conventions
- Libraries
- Calculation
- All of the given options (page 307)

Question #165

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

- Project Planning
- Requirement definition (page 299)
- Requirement verification
- Requirement validation

AL-JUNAID INSTITUTE GROUP

Question #166

Bill Inmon argues that requirements are well understood only after

- They are documented
- They extracted and verified
- Data warehouse is populated (page 285)
- All schemes are defined

Question #167

Waterfall model is appropriate when

- When budget is low
- When the deadlines is strict
- When resources are limited
- Requirements are clearly defined (page 284)

Question #168

A bottom up implementation approach is useful when

- Technology is matured and well understands (page 283)
- Organization cannot implement latest technologies
- The business objectives are very much clear
- All requirements are well documented

Question #169

Technical architecture design specifies the:

- Project schedule (page 300)
- Minimum project completion time
- Required components
- Umbrella activities

Question #170

----- says that as far as company goals are connected, the entire company pursues in the same direction

- Westerman (page 285)
- Bill inmon
- Kimball
- Bohnlein

Question #171

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

- Data driven (page 285)
- Resource driven
- Requirement driven
- Time sensitive

AL-JUNAID INSTITUTE GROUP

Question #172

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

- Deployment
- Operation (page 288)
- Enhancement
- Maintenance

Question #173

Following the coding standards during development helps to:

- Minimize the future work (page 307)
- Verify the requirements
- To refine the project scope
- To establish the test cases

Question #174

As per Kimball, ……..is the main operational process.

- Requirements extraction
- Global design
- Business process (page 285)
- Scheme design

Question #175

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

- Design
- Prototype (page 287)
- Deployment
- Operation

Question #176

In case of multiple sources for the same data element, we need to prioritize the source systems per element bases, the process is called:

- Ranking (page 143)
- Prioritization
- Element Selection
- Measurement Selection

AL-JUNAID INSTITUTE GROUP

Question #177

Goal driven approach of data ware house development was result of work.

- Bill Inmon
- Ralph Kimball
- Bohnlein and ulbrich-vom (page 285)
- Westerman

Question #178

In contrast to data mining, statistics is driven.

- Assumption
- Knowledge
- Human (page 255)
- Database

Question #179

Which of the following is example of static attributes?

- Employee Name (page 342)
- Employee date of birth
- Employee blood group
- All of the given

Question #180

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

- Database design is complete (page 307)
- A subset of historical data has been loaded
- The data access tools and metadata are installed
- All of the given options

1. Bill Inmon argues that requirements are well understood only after
   a. They are documented
   b. They extracted and verified
   c. Data warehouse is populated page 285
   d. All the schemas are defined
2. Spiral model is _____
   a. Sequence of waterfall model page 284
   b. Risk oriented model
   c. An iterative model
   d. All of the given options

AL-JUNAID INSTITUTE GROUP

3. Waterfall is a/an_____model
    a. Iterative
    b. Simple linear sequential page 284
    c. Object oriented
    d. Rapid development
4. An effective user education program includes, among others, the followingguideline(s);
    a. Understand the audience, don't overwhelm
    b. Train after delivery of data and analytic applications
    c. Postpone education, if DWH not ready
    d. All of the given options page 308
5. One of the drawbacks of waterfall model is that:
    a. Customers can not review the product during development
    b. It does not work when the resources are limited
    c. It does not define the project timeline/schedule
    d. All of the given options page 284
6. As per Bill Inmon, a data warehouse, in contrast with classical applications is:
    a. Data driven page 285
    b. Resource driven
    c. Requirement driven
    d. Time sensitive
7. Horizontally wide data means:
    a. Dataset has large no. of attributes page 330
    b. Dataset has large no. of records
    c. Dataset has attribute skews
    d. Dataset has partitioning skews
8. 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
    d. Requirements wrap-up presentation page 297

9. In analytical application development phase, we follow standards for:
    a. Naming conventions
    b. Standard for calculations
    c. Standard for libraries
    d. All of the given options page 307

10. In lifecycle data track, we begin with translating the requirements intodimensional model, which then transforms into_____.
    a. Physical structure page 290
    b. Logical structure
    c. Conceptual structure
    d. System structure

11. Technical architecture design supports the communication about technicalrequirements:
    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

12. A top down implementation approach is useful when
    a. Technology is mature and well understood page 283
    b. Organization can not implement latest technologies
    c. Business objectives are unclear
    d. Problem to be solved are not well documented

13. Which of the following is NOT one of the possible pitfalls in DWH Life Cycle & Development?
    a. Not having multiple servers
    b. Low priority OLAP Cube Construction
    c. Improper documentation
    d. None of the given options page 312, 313, 314

14. Goal driven approach of data warehouse development was result of _____ work.
    a. Bill Inmon
    b. Ralph Kimball
    c. Bohnlein and Ulbrich-vom page 285
    d. Westerman
15. Which of the following is/are included in the list of Top-7 key steps for smooth DWH implementation?
    a. Consider handing-off project management
    b. Assign significant resources for ETL
    c. Be a diplomat NOT a technologist
    d. All of the given options page 318, 319
16. A typical cycle of implementing the change in DWH comprises of the sequence:
    a. Production -> QA -> Development
    b. Development -> QA -> Production page 314
    c. Development -> Production-> QA
    d. Production-> Development -> QA
17. 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. Prototype page 287
    c. Deployment
    d. Operation
18. In context of requirement definition phase in Kimball's DWH developmentapproach, activities like debriefing, documentation, and prioritization are considered to be the part of
    a. Requirement preplanning
    b. Business requirements collection
    c. Post collection page 294
    d. None of the given options

19. 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 analytical applications track
    d. Lifecycle maintenance track page 299
20. Normally the term "DWH face to the business user" refers to:
    a. Lifecycle technology track
    b. Lifecycle data track
    c. Lifecycle analytical applications track page 306
    d. Lifecycle maintenance track
21. 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
22. Improper documentation results the problem(s) like:
    a. Maintenance issue
    b. New developers unable to configure already existing code
    c. Lot of time required for enhancing the code
    d. All of the given options
23. A bottom up implementation approach is useful when
    a. Technology is mature and well understood
    b. Organization can not implement latest technologies
    c. The business objectives are very much clear
    d. All the requirements are well documented page 283
24. As per Kimball,_____is the main operational process
    a. Requirement extraction
    b. Goal design
    c. Business process page 285
    d. Schema design
25. Which of the following is NOT one of the top-10 mistakes that should beavoided during DWH development?
    a. Not interacting directly with end user
    b. Not being an accommodating person page 316, 317
    c. Isolating IT support people from business users
    d. Training the users with dummy data and considering it success

AL-JUNAID INSTITUTE GROUP

26. The_____phase of fundamental data warehouse life cycle modelincludes data warehouse daily maintenance activities
    a. Deployment
    b. Operation page 288
    c. Enhancement
    d. Maintenance
27. In context of the most fundamental data warehouse life cycle model, whichof the following is NOT one of the data warehouse design activities?
    a. End-user interviews and re-interviews
    b. Source system cataloguing
    c. Definition of key performance indicators
    d. System vision development page 287
28. Which of the following is NOT one of the methodologies for Data Warehouse project development?
    a. Goal Driven
    b. Data Driven
    c. User Driven
    d. System Driven page 283
29. In context of analytical applications track, the application developmentactivity can begin after:
    a. Database design is complete
    b. A subset of historical data has been loaded
    c. The data access tools and metadata are installed
    d. All of the given options page 307
30. 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
31. Users do not care, how advance the front end of your DWH is, what theycare is that:
    a. Tables should be properly Denormalized
    b. Proper partitioning technique should be used
    c. At least star or snowflake schema should be implemented
    d. They should get information in timely manner and the way they want page 314

AL-JUNAID INSTITUTE GROUP

32. Which of the following activity executes parallel with all other activities in Kimball's DWH development approach?
    a. Requirement elicitation
    b. Project planning
    c. Project management
    d. Deployment
33. Which of the following is the most ignored step during data warehousedevelopment
    a. The requirement verification
    b. The vision definition
    c. Schema validation
    d. Success criteria development
34. Which of the following is NOT one of the activities of "Maintenance and Growth phase in Kimball's DWH development approach?
    a. Education
    b. Technical Education
    c. Program Support
    d. Interface Deployment page 309
35. In_____phase of Kimballs approach, we identify the componentsneeded now and in future.
    a. Requirement definition
    b. Architectural design page 300
    c. Product development
    d. Analytical application development
36. Implementation of a data warehouse requires_____activities
    a. Highly integrated
    b. Loosely integrated
    c. Tightly decoupled
    d. None of the given page 289
37. Which of the following activity/activities is/are part of project planningphase 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

AL-JUNAID INSTITUTE GROUP

38. says that as far as company goals are concerned, the entirecompany pursues in the same direction

a. Westerman page 285

b. Bill Inmon

c. Kimball

d. Bohnlein