



In the Name of Allāh, the Most Gracious, the Most Merciful

CS614 - Data Warehousing - Midterm Papers Solved MCQ(S) (1 TO 22 Lectures)

1. If there are N records in a table, then the selectivity of primary key column is

- **1/N** PG # 22
- N/2
- lg(N)

2. _____ is one class of decision support environment.

- **OLAP** PG # 30
- OLTP
- Data Cleansing
- ETL

3. The users of data warehouse are _____

- Decision makers
- Knowledge workers
- **Both Knowledge workers and Decision makers** PG # 18

The users of data warehouse are **knowledge workers in other words they are decision makers** in the organization.

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

➤ **Decision maker** PG # 18

➤ Manager

➤ Database Administrator

➤ DWH Analyst

5. Effects of de-normalization on database performance are

➤ **Unpredictable** PG # 62

➤ Predictable

➤ Conventional

➤ Unsurprising

6. OLAP is used for analytical process. For analytical processing we need

➤ **Multi-level aggregates** PG # 74

➤ Record level access

➤ Data level access

➤ Row level access

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

➤ **File** PG # 79

➤ Application

➤ Aggregate

➤ Database

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

8. Normally Selectivity of query in OLTP system is

- **High** PG # 30
- Low
- Not measured

OLTP	DWH
Primary key used	Primary key NOT used
No concept of Primary Index	Primary index used
May use a single table	Uses multiple tables
Few rows returned	Many rows returned
High selectivity of query	Low selectivity of query
Indexing on primary key (unique)	Indexing on primary index (non-unique)

Table-4.1: Comparison of OLTP and DWH for given queries

9. Normally Selectivity of query in data warehouse is

- High
- **Low** PG # 30
- Not measured

10. De-Normalization normally speeds up

- **Data Retrieval** PG # 51
- Data Modification
- Development Cycle
- Data Replication

11. De-Normalization normally slows down

- Data Retrieval
- **Data Modification** PG # 51
- Development Cycle
- Data Replication

De-normalization usually speeds up data retrieval, but it can slow the data modification processes.

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

12. Geography is a good example of

- One-dimensional Hierarchy
- **Multidimensional Hierarchy** PG # 52
- Non-Dimensional
- Linear Hierarchy

13. Partition elimination is not possible with

- **Round-Robin** PG # 66
- De-normalization
- Normalization

14. OLAP is Analytical processing instead of Transaction processing, and OLAP is a

- Physical database design
- Implementation technique
- **Framework** PG # 69

OLAP is Analytical Processing instead of Transaction Processing. It is also NOT a physical database design or implementation technique, but a framework.

15. ER is a logical design technique that seeks to remove the _____ in data.

- **Redundancy** PG # 98
- Normalization
- Anomalies

16. ER is a _____ design technique that seeks to remove the redundancy in data.

- **Logical** PG # 98
- Physical
- Data Dependent
- Transaction Dependent

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

17. Pre-computed _____ can solve performance problems

- **Aggregates** **PG # 111**
- Facts
- Dimensions

18. Extract, Transform, Load (ETL) process consist of steps which are _____.

- **Independent and interrelated** **PG # 131**
- Independent or interrelated
- Dependent and interrelated
- Dependent or interrelated

19. In full extraction, data is extracted completely from the source system. Therefore there is no need to keep track of changes to the _____

- **Data Source** **PG # 133**
- DWH
- Data Mart

20. Data Extraction from source systems is a difficult task because source systems are _____

- Homogeneous
- **Heterogeneous** **PG # 140**
- Identical
- Standardized

21. Lexical errors fall in which type of class of anomalies

- **Syntactically Dirty Data** **PG # 160**
- Semantically Dirty Data
- Coverage Anomalies
- Missing Values Anomalies

22. The degree of similarity between two records, often measured by a numerical value between _____, usually depends On application characteristics.

- **0 and 1** PG # 169
- 0 and 10
- 0 and 100
- 0 and 99

23. As consumers, human beings judge the quality of things during their life-time.

I Consciously

II Subconsciously

III Unconsciously

Which of the following statement is true?

- I Only
- II Only
- III Only
- **I & II Only** PG # 179

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

- 1950s.
- **1960s.** PG # 12
- 1970s.
- 1980s.

” گفتگو ایسی چیز ہے۔
” جسکی وجہ سے انسان یا تو دل میں اتر جاتا ہے یا پھر دل سے اتر جاتا ہے۔“

25. 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**
- Two Direction
- None of these

PG # 19

26. In ROLAP access to information is provided via relational database using _____ standard SQL.

- **ANSI**
- Microsoft
- Oracle
- SAP

PG # 78

27. Which of the following statement is true? 1 GB is

- **2^{30} or 10^9 bytes**
- 2^{30} or 10^6 bytes
- 2^{32} or 10^9 bytes
- 2^{32} or 10^8 bytes

PG # 15

How Much Data is that?		
1 MB	2^{20} or 10^6 bytes	Small novel – 3 1/2 Disk
1 GB	2^{30} or 10^9 bytes	Paper rims that could fill the back of a pickup van
1 TB	2^{40} or 10^{12} bytes	50,000 trees chopped and converted into paper and printed
2 PB	1 PB = 2^{50} or 10^{15} bytes	Academic research libraries across the U.S.
5 EB	1 EB = 2^{60} or 10^{18} bytes	All words <u>ever</u> spoken by human beings

Table-2.1: Quantifying size of data

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

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

- $O(n)$
- $O(w)$
- **$O(w n)$**
- $O(w \log n)$

PG # 171

29. The telecommunications data warehouse is dominated by the sheer volume of data generated at the call level _____ area.

- **Subject**
- Object
- Aggregate
- Detail

PG # 35

30. "Decision making is an iterative process; which must involve the users" is a classic statement of

- **OLAP**
- ATM
- OLTP
- Data Mining

PG # 69

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

- **An Intersection**
- A Union
- A Subtraction
- A Subset

PG # 88

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

32. After implementing Change Data Capture, the advantage we have is that, data is able to be integrated and transformed _____

- **In-flight** [Click Here For Reference Detail](#) **PG # 152**
- Off-flight
- Stored Data
- Over-flight

Finally data is able to be integrated and transformed "in-flight". Once the update/transaction data has been pulled from the log tape, the DWH is free to re-sequence, reformat, convert, merge, summarize, etc.

33. All data is _____ of something real.

- I An Abstraction
- II A Representation

Which of the following option is true?

- **I Only** **PG # 180**
- II Only
- Both I & II
- None of I & II

34. _____ queries deal with number of variables spanning across number of tables (i.e. join operations) and looking at lots of historical data.

- OLTP
- DBMS
- **DSS** **PG # 21**
- ETL

35. In data warehouse, a query results in retrieval of hundreds of records from very large table. the ratio of number of records retrieved to total number of records present is high and selectivity is:

- **Low** **PG # 22**
- High
- Average
- Non computable

36. At small scale e.g. hundreds of rows, difference between performance of complexity algorithms “ $O(n \lg n)$ ”

and “ $O(n^2)$ ” is not _____.

- Noticeable
- Small
- Symmetric
- **Asymmetric**

PG # 32

37. Collapsing tables can be done on the _____ relationship(s)

- Only One-to-One
- Only Many-to-Many
- Only One-to-Many
- **Both One-to-One and Many-to-Many**

PG # 52

38. HOLAP provides a combination of relational database access and cube data structures. The goal is to get the best of both MOLAP and ROLAP i.e.

- **Scalability and high performance**
- Backward compatibility to RDBMS
- Forward compatibility to new DBMS
- Usability of new DBMS

PG # 78

اے اللہ
مشکلات میں میرا ہاتھ پکڑ لے
اور میرے سامنے ہر معاملے میں حقائق کے سارے پہلو کھول دے

39. Primary key is repeated in _____ splitting.

- Horizontal
- **Vertical**
- Cross
- Diagonal

PG # 56

40. “Change Data Capture” is one of the challenging technical issues in _____

- **Data Extraction**
- Data Loading
- Data Transformation
- Data Cleansing

PG # 150

41. Merging information is one of the major types of _____

- **Transformation**
- Extraction
- Loading
- Updation

PG # 152, 153

Data merging is part of data transformation where multiple values are summarized into single summarized value.

اپنے آپ کو بدل دو تمہاری
قسمت خود بخود بدل جائے گی

42. Which of the following is NOT an example of a “Dimension”?

- Date
- Region
- Product
- **Units sold**

PG # 74

FACTS: Quantitative values (numbers) or “measures.”

e.g., units sold, sales \$, Co, Kg etc.

DIMENSIONS: Descriptive categories.

e.g., time, geography, product etc.

DIM often organized in hierarchies representing levels of detail in the data (e.g., week, month, quarter, year, decade etc.).

43. Multi-value dependencies are removed in _____

- 1st Normal form
- 2nd Normal form
- 3rd Normal form
- **4th Normal form**

PG # 48

44. Which of the following is not one of the characteristics of a Data warehouse?

- Ad-hoc access
- Complete repository
- Historical data
- **Volatile**

PG # 18

A complete repository of historical corporate data extracted from transaction systems that is available for ad-hoc access by knowledge workers

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

45. To handle dimensions that require the aggregation of multiple data quality indicators, which of the following operations can be applied:

- **Minimum or Maximum** PG # 188
- Complex Ratio
- Aggregate Average
- Weighted Ratio

46. Full and Incremental extraction techniques are types of _____

- **Logical Extraction** PG # 133
- Physical Extraction
- Both Logical Extraction and Physical Extraction
- Semantic Extraction

The two logical data extraction types are full and incremental extraction techniques.

47. “The environment is smart enough to develop or compute higher level aggregates using lower level or more detailed aggregates”. Which of the following approach is described by the above statement?

- **Aggregate awareness** PG # 87
- Cube partitioning
- Indexing
- MOLAP cube aggregation

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

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

- Linear
- Quadratic
- Logarithmic
- **Exponential**

PG # 15

49. ER Model can be simplified in ----- ways

- One
- **Two**
- Three
- Four

PG # 103

There are actually two ways of “simplifying” the ER model i.e. (i) De-normalization and (ii) Dimensional Modeling.

50. Non recording facts have a disadvantage that it has

- **Lack of Information**
- Redundant Information
- Repeated Information
- Normalized Information

PG # 120

51. Fact-less fact table is a fact table without numeric fact columns. It is used to capture relationship between

- **Dimensions**
- Attributes
- Tables
- Facts

PG # 121

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

52. A cube is not a data warehouse, it is a _____

- **Data Mart** PG # 131
- Data Extraction
- Data Loading
- Query Manager

53. If we remove the modification anomalies the table comes in _____ form

- 1NF
- 2NF
- **3NF** PG # 47
- 4NF

54. Modification anomalies occur in _____ form

- 1NF
- **2NF** PG # 45
- 3NF
- BCNF

55. In MOLAP physically build cubes for direct access, support is not available for _____ SQL.

- **ANSI** PG # 78
- Microsoft
- Oracle
- SAP

MOLAP physically builds “cubes” for direct access - usually in the proprietary file format of a multi-dimensional database (MDD) or a user defined data structure. Therefore **ANSI SQL is not supported.**

جھوٹ انسان اور ایمان دونوں کا دشمن ہے

56. _____ is the lowest level of detail or the atomic level of data stored in the warehouse.

- Aggregate
- Cube
- **Grain** PG # 111
- Virtual Cube

57. Redundancy causes _____ anomalies

- **Update** PG # 43
- Select
- Both Update & Select
- None of these

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

- 2^{52} or 10^{13} bytes
- **2^{50} or 10^{15} bytes** PG # 15
- 2^{50} or 10^{10} bytes
- 2^{48} or 10^{12} bytes

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

- $O(n)$
- $O(n^2)$
- $O(n \lg n)$
- **$O(\log n)$** PG # 22

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

60. One major goal of horizontal splitting is

- Splitting rows for exploiting parallelism
- Splitting columns for exploiting parallelism
- Splitting schema for exploiting parallelism
- **Spreading rows for exploiting parallelism.**

PG # 46

Splitting Tables: Horizontal splitting

Breaks a table into multiple tables based upon common column values. Example: Campus specific queries.

GOAL

- **Spreading rows for exploiting parallelism.**
- Grouping data to avoid unnecessary query load in WHERE clause.

61. Fact-less fact table is a fact table without numeric fact columns. It is used to capture relationship between

- **Dimensions**
- Attributes
- Tables
- Facts

PG # 121

A Fact-less Fact Table

- "Fact-less" fact table
 - **A fact table without numeric fact columns**
 - **Captures relationships between dimensions**
 - Use a dummy fact column that always has value 1

ایک جھوٹ کو چھپانے کیلئے
سو جھوٹ بولنے پڑ جاتے ہیں

62. _____ gives total view of an organization

- OLTP
- **Data warehouse** PG # 16
- OLAP
- Data base

63. Experience showed that for a single pass magnetic tape that scanned 100% of the records, only _____ of the records, sometimes even were actually required.

- **5%** PG # 12
- 30%
- 50%
- 80%

64. In _____ system, the contents change with time.

- **OLTP** PG # 20
- DSS
- ATM
- OLAP

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

- **Doubles** PG # 15
- Triples
- Quartiles
- Remains same as previous year

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

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

- **Master table only** PG # 64
- Details tables only
- Both master and detail tables

67. Partition elimination is not possible with

- **Round-Robin** PG # 66
- De-normalization
- Normalization

For More Solved Papers + Help Material Visit: www.lmshelp.blogspot.com

پردہ عورت کے چہرے پر ہوتا ہے

اگر چہرے پر نہیں

تو سمجھ لو کہ اُس کی عقل پر ضرور ہے

Note: If you found any mistake in mcqz please mail at above mentioned email address.
And tell me your answer with references.



*Winning is not everything,
but wanting to win is
everything.....
Go Ahead.... Best Of Luck !*

JOKESMANTRA.COM

please pray for me
and I will pray for you too



THE END

Campus (AKMPO1)

ARSLAN ARSHAD

Arslan.arshad01@gmail.com