

Objective Questions Reference
BCA & BSC (IT)- SEM. IV

Sub: Data Warehousing with SQL Server 2012

(1) What is DWH ?

Ans. In computing, a data warehouse, also known as an enterprise data warehouse, is a system used for reporting and data analysis, and is considered a core component of business intelligence. DWs are central repositories of integrated data from one or more disparate sources.

(2) Give the full form of DDS.

Ans. dimensional data store

(3) what is a stage in data stage/store ?

Ans. Database **stages** represent data sources or data targets. **Data Stage** provides three types of **stages**:

(4) What is third party data ?

Ans. Third-party data is any information collected by an entity that does not have a direct relationship with the user the data is being collected on. Often times, third-party data is collected from a variety of websites and platforms and is then aggregated together by a third-party data provider such as a DMP.

(5) Full form of UID.

Ans. unique identifier

(6) Logical design entity know as in physical design.

Ans. A logical design is a conceptual, abstract design. You do not deal with the physical implementation details yet; you deal only with defining the types of information that you need. The process of logical design involves arranging data into a series of logical relationships called entities and attributes.

(7) what is table space ?

Ans. Table spaces consist of one or more containers. A container can be a directory name, a device name, or a file name. A single table space can have several containers

(8) What is Role playing dimension ?

Ans. Dimensions are often recycled for multiple purposes within the same database. This is often referred to as a "role-playing dimension"

(9) Full form of ETL.

Ans. Extraction, transformation, and loading

(10) Full form of SSIS.

Ans. SQL Server Integration Services

(11) Full form of SSDT.

Ans. SQL Server Data Tools

(12) Give the full form of DTS.

Ans. Data Transformation Services

(13) Full form of DQS.

Ans. Data Quality Services

(14) Full form of DQAF.

Ans. Data Quality Assessment Framework

(15) what is the use of SSIS Script task ?

Ans. To run the package, use the Script Task to create a new VSTA script. In most cases, scripts are converted automatically to use a supported version, when you open a SQL Server Integration Services package

(16) _____ is used to generate log messages as well as other logics related to packages and SSIS.

Ans. XML

(17) What is SQL CLR?

Ans. SQL CLR or SQLCLR (SQL Common Language Runtime) is technology for hosting of the Microsoft . NET common language runtime engine within SQL Server. The SQLCLR allows managed code to be hosted by, and run in, the Microsoft SQL Server environment.

(18) Full form of BIML.

Ans. Business Intelligence Markup Language

(19) Full form of BIDS.

Ans. Business Intelligence Development Studio

(20) What are the applications of data warehousing?

Ans. Banking Industry, Finance Industry, Consumer Goods Industry, Government and Education, Healthcare, Hospitality Industry, Insurance, Manufacturing and Distribution Industry.

(21) what is meant by non-volatile nature in data warehouse?

Ans. Data warehouse is also non-volatile means the previous data is not erased when new data is entered in it. Data is read-only and periodically refreshed. This also helps to analyze historical data and understand what & when happened. It does not require transaction process, recovery and concurrency control mechanisms.

(22) What is OLTP?

Ans. OLTP is an operational system that supports transaction-oriented applications in a 3-tier architecture. It administers the day to day transaction of an organization. The full form of OLTP is Online Transaction Processing.

(23) DW Stand for _____

Ans. Data Ware House

(24) RDBMS stands for _____

Ans. Relational Database Management System.

(25) Explain MOLAP.

Ans. MOLAP stands for Multidimensional online analytical processing. It is a type of OLAP process which utilizes a multidimensional data model so that analysis of data can be done easily. Data in MOLAP is pre-computed, pre-summarized and is stored in MOLAP.

(26) What is dimension table ?

Ans. A dimension table is a table in a star schema of a data warehouse. A dimension table stores attributes, or dimensions, that describe the objects in a fact table.

(27) Explain any one Different in Dimension table and fact table.

Ans. Fact table contains measurements, metrics, and facts about a business process while the Dimension table is a companion to the fact table which contains descriptive attributes to be used as query constraining.

(28) SSAS stands for _____

Ans. SQL Server Analytics Services.

(29) SSIS stands for _____

Ans. SQL Server Integration Services.

(30) what is real-time data Warehousing ?

Ans. A **real-time data warehouse** is one that acquires, cleanses, transforms, stores, and disseminates information in **real time**. An active **data warehouse**, on the other hand, operates in a non-**real-time** response mode with one-or-more OLTP systems.

(31) what is meta data repository ?

Ans. A **metadata repository** is a software tool that stores descriptive information about the **data** model used to **store** and **share metadata**. **Metadata repositories** combine diagrams and text, enabling **metadata** integration and change.

(32) DQS Stands for _____

Ans. Data Quality Services.

(33) What is data cleansing ?

Ans. **Data cleansing** or **data cleaning** is the process of detecting and correcting (or removing) corrupt or inaccurate records from a record set, table, or database and refers to identifying incomplete, incorrect, inaccurate or irrelevant parts of the **data** and then replacing, modifying, or deleting the dirty or coarse **data**.

(34) what is Data Matching In Data Quality ?

Ans. **data matching** process enables you to reduce **data** duplication and improve **data** accuracy in a **data** source. **Matching** analyzes the degree of duplication in all records of a single **data** source, returning weighted probabilities of a **match** between each set of records compared.

(35) What is data Quality ?

Ans. Data quality refers to the state of qualitative or quantitative pieces of information.

(36) what is a checkpoint in SSIS ?

Ans. We can configure a **CHECKPOINT** file in the **SSIS** package to log package execution information in it. If the package execution fails, **SSIS** uses the information in the **checkpoint** file to restart it from the point of failure.

(37) How do you deploy SSIS Packages ?

Ans. **SSIS packages deployed** to SQL Server are stored in the msdb database. There are three ways to **deploy** our **packages**: Create a **deployment** utility from our project. Use the DTUTIL command line tool.

(38) SSRS Stands for _____

Ans. SQL Server Reporting Services.

(39) List of different types of SSRS Reports.

Ans. Linked Reports, Clickthrough Reports, DrillDown Reports, Drillthrough Reports, Cached Reports, Ad hoc Reports

(40) Who is developed by data warehouse?

Ans. Bill Inmon

(41) Full form of OLTP.

Ans. Online Transactional Processing

(42) Full form of ODS.

Ans. operational data store

(43) What is DATA?

Ans. Data are units of information, often numeric, that are collected through observation. In a more technical sense, data are a set of values of qualitative or quantitative variables about one or more persons or objects, while a datum (single value) of a single variable.

(44) What is UML?

Ans. The Unified Modeling Language is a general-purpose, developmental, modeling language in the field of software engineering that is intended to provide a standard way to visualize the design of a system.

(45) Full form of DBA.

Ans. Database Administrator

(46) What is a Fact Table?

Ans. A factless **fact table** is a **fact table** that does not have any measures. It is essentially an intersection of dimensions (it contains nothing but dimensional keys).

(47) What is a dimension table?

Ans. A dimension is a structure that categorizes facts and measures in order to enable users to answer business questions. Commonly used dimensions are people, products, place and time.

(48) What is Server ?

Ans. A server is a computer that serves information to other computers. These computers, called clients, can connect to a server through either a local area network or a wide area network, such as the internet. A server is a vital piece of your IT infrastructure.

(49) List out of all Microsoft SQL Services.

Ans. SQL Server Browser, SQL Server Full Text Search, SQL Server Integration Services, SQL Server Reporting Services, SQL Server Analysis Services.

(50) What is Container ?

Ans. A **container** is an application, including all its dependencies, libraries and other binaries, and the configuration files needed to run it, bundled into a single package that can be moved, in total, from one computing environment to another.

(51) Full form of VSTA.

Ans. very small aperture terminal

(52) Full form of MDS.

Ans. Master Data Services Overview

(53) Full form of DQA.

Ans. Data Quality Assessment.

(54) What is Reporting Service ?

Ans. SQL Server Reporting Services is a server-based report generating software system from Microsoft. It is part of a suite of Microsoft SQL Server services, including SSAS and SSIS. Administered via a Web interface, it can be used to prepare and deliver a variety of interactive and printed reports.

(55) What is data analysis ?

Ans. Data analysis is a process of inspecting, cleansing, transforming, and modeling data with the goal of discovering useful information, informing conclusions, and supporting decision-making.

(56) Full form of BISM.

Ans. Business Intelligence Semantic Model

(57) Define Data Mining.

Ans. **Data mining** is the process of identifying interesting and useful patterns in large databases by means of automated methods.

(58) What is clustering?

Ans. **Clustering** is the process of making a group of abstract objects into classes of similar objects. Points to Remember. A **cluster** of **data** objects can be treated as one group. While doing **cluster** analysis, we first partition the set of **data** into groups based on **data** similarity and then assign the labels to the groups.

(59) Which is used to illustrate the hierarchical clustering technique?

Ans. **Hierarchical clustering** may be represented by a two-dimensional **diagram** known as a dendrogram, which **illustrates** the fusions or divisions made at each successive stage of **analysis**.

(60) What is surrogate key?

Ans. **Surrogate keys** are widely used and accepted design standard in **data warehouses**. It is sequentially generated unique number attached with each and every record in a Dimension table in any **Data Warehouse**.

(61) Give the name of in three tiered DWH architecture.

Ans. Data Warehouses usually have a **three-level (tier) architecture** that includes: Bottom **Tier** (**Data Warehouse Server**) Middle **Tier** (**OLAP Server**) **Top Tier** (**Front end Tools**).

(62) _____ is the last phase of CRISP Data Mining cycle.

Ans. Evaluation

(63) Which component of DWH architecture shows analysis in graphical format for taking decision?

Ans. Data Warehouse is used for analysis and decision making in which extensive database is required, including historical data, which operational database does not typically maintain.

(63) Dimension table normally includes _____ data.

Ans. Primary

(64) HOLAP stands for _____

Ans. Hybrid Online Analytical Processing

(65) Which are two types of hierarchical clustering?

Ans. Hierarchical clustering can be divided into two main types: agglomerative and divisive.

Agglomerative clustering: It's also known as AGNES (Agglomerative Nesting). It works in a bottom-up manner.

Divisive hierarchical clustering: It's also known as DIANA (Divise Analysis) and it works in a top-down manner.

(66) What is WEKA?

Ans. Waikato Environment for Knowledge Analysis, developed at the University of Waikato, New Zealand, is free software licensed under the GNU General Public License, and the companion software to the book "Data Mining: Practical Machine Learning Tools and Techniques".

(67) What is a stage in data store ?

Ans. A staging area, or landing zone, is an intermediate **storage** area used for **data** processing during the extract, transform and load (ETL) process. The **data** staging area sits between the **data** source(s) and the **data** target(s), which are often **data** warehouses, **data** marts, or other **data** repositories.

UNLOCK YOUR POTENTIAL