**1.What is a DML and what do they do?**

DML can be defined as a set of syntax elements that are used to manage the data in the database.

Data manipulation language (DML) statements query or manipulate data in existing schema objects. It is a computer programming language that is used to perform select,insert,update ,delete data in a database.

They enable you to:

1. **Retrieve data from one or more tables or views (SELECT)**
2. **Add new rows of data into a table or view (INSERT)**
3. **Change column values in existing rows of a table or view (UPDATE)**
4. **Remove rows from tables or views (DELETE)**

**2.Define NOT NULL constraint**

Constraints are the rules enforced on the data columns of a table .These are used to limit the type of data that can go into the table.

NOT NULL constraint enforces a column to not accept null values.This enforces a field to not accept null values.This enforces a field to always contain a value, which means that you cannot insert a new record or update a record without adding a value to this field

**3. What are different Oracle database objects?**

* TABLES
* VIEWS
* INDEXES
* SYNONYMS
* SEQUENCES
* TABLESPACES

**4. What are the different types of indexes supported by Oracle?**

The different types of indexes are:

* + **B-tree indexes**
  + **B-tree cluster indexes**
  + **Hash cluster indexes**
  + **Reverse key indexes**
  + **Bitmap indexes**

**5. What is the difference between a database, a data warehouse and a data mart?**

**Database:**

A database is an organised collection of structured information typically stored electronically in computer system

**Data Warehouse:**

A data warehouse is a large collection of business data used to help an organisation to make decisions

It includes the assortments of all sorts of data. The data is taken out only according to the customer's needs.

**Datamart:**

A datamart is a simple form of data warehouse focused on a single subject or a line of business. For instance, an organization having a different chunk of data for its different departments i.e. sales, finance, marketing etc.

**6. What is the difference between Mapping and Mapplet?**

**Mapping:**

* It is a collection of source, target and transformation.
* It is developed with different transformations and is not reusable.
* It is developed around what data to move the target and what modification is performed on it.

**Mapplet:**

* It is a collection of transformation only.
* It can be reused with other mapping and also mapplets.
* It is developed for complex calculations used in multiple mappings.

**7. What is the meaning of Enterprise Data Warehousing?**

Enterprise Data Warehousing is about organizing the data that can be created or developed at a single point of access. The data is globally accessed and viewed through a single source as the server is linked to a single source

**8. Explain sessions. Explain how batches are used to combine executions?**

* A session is a teaching set that is implemented to convert data from a source to a target. Session's manager is used to carry out a session by using the “pmcmd” command.
* Batch execution can be used to combine sessions executions either in a serial manner or in a parallel. Batches can have different sessions carrying forward in a parallel or serial manner.

**9.Explain the difference between mapping parameter and mapping variable?**

**Mapping variable:**

* Mapping variables are the values which change during the session’s execution.
* After completion, the Informatica server stores the end value of a variable and is reused when session restarts.

**Mapping parameter:**

* Mapping parameters are the values which do not change during the session execution.
* Mapping procedure explains mapping parameters and their usage. Values are allocated to these parameters before starting the session.

**10.What is Rank Transformation in Informatica?**

Rank is an active and connected transformation that performs the filtering of data based on the groups and ranks.

* Rank Transformation is a type of an active T/R which allows you to find out either top performance or bottom performers.
* Rank T/R is created with the following types of the port:
  1. Input Port (I).
  2. Output Port (O).
  3. Rank Port (R).
  4. Variable Port (V).

**11.Illustrate the differences that exist between joiner and Lookup Transformation.**

**Joiner:**

* It is not possible to override the query.
* Only the ‘=’ operator is available.
* Users can’t restrict the number of rows while reading relational tables.
* Tables are joined using Joins.

**Lookup:**

* It is possible to override the query.
* All operators are available for use.
* Users can restrict the number of rows while reading relational tables.
* It behaves as Left Outer Join while connecting with the database.

**12.What are the different ways to filter rows using Informatica transformations?**

 Ways to filter rows using Informatica transformations are as follows.

* Source Qualifier.
* Joiner.
* Filter.
* Router.

**13.State the differences between SQL Override and Lookup Override?**

**SQL Override:**

* The role of SQL Override is to limit the number of incoming rows entering the mapping pipeline.
* It doesn’t use the “Order By” clause and it should be manually entered in the query if we require it. SQL Override can provide any kind of ‘join’ by writing the query.
* It will not give a record if it finds multiple records for a single condition.

**Lookup Override:**

* It is used to limit the number of lookup rows to avoid the whole table scan by saving the lookup time and the cache it uses.
* It uses the “Order By” clause by default. Lookup Override provides only Non-Equi joins.
* It gives only one record even if it finds multiple records for a single condition.

**14.Name the different lookup caches?**

* Informatica lookups can be cached or uncached.
* Cached lookups can be either static or dynamic.
* A lookup cache can also be divided as persistent or non-persistent based on whether Informatica retains the cache even after completing session run or if it deletes it.
* Different lookup caches are.
* Static cache
* Dynamic cache
* Persistent cache
* Shared cache
* Recache

**15.What is diffrence between Co-related sub query and nested sub query?** Correlated subquery runs once for each row selected by the outer query. It contains a reference to a value from the row selected by the outer query.

Nested subquery runs only once for the entire nesting (outer) query. It does not contain any reference to the outer query row.

**16.What is the difference between clustered and a non-clustered index?**

A**clustered index** is a special type of index that reorders the way records in the table are physically stored. A **Nonclustered index**is a special type of index in which the logical order of the index does not match the physical stored order of the rows on disk.

**17.What are the types of lookup transformation?**

There are four different types of lookup transformation:

* **Relational or flat-file lookup**: It performs a lookup on relational tables.
* **Pipeline lookup**: It performs a lookup on application sources.
* **Connected or unconnected lookup**: While the connected lookup transformation receives data from the source, performs a lookup, and returns the result to the pipeline,
* **Cached or uncached lookup**: Lookup transformation can be configured to cache lookup data, or we can directly query the lookup source every time a lookup is invoked.

**18.How can we filter rows in Informatica?**

There are two ways to filter rows in Informatica, they are as follows:

* **Source Qualifier Transformation**: It filters rows while reading data from a relational data source.
* **Filter Transformation**: It filters rows within mapped data from any source. It is added close to the source to filter out the unwanted data and maximize performance

**19.What is a Sequence Generator Transformation?**

Sequence Generator Transformation generates unique values such as primary keys for a target in mapping and it functions like a sequence object in a database

Sequence generator transformation is a passive and connected transformation that generates numeric sequence values such as 1,2,3 and so on it does not affect the input rows.

Sequence generator is used to create unique primary key values and replace missing primary keys

**20.What is the difference between static and dynamic cache?**

The difference between static and dynamic cache are:

**Static** – Static cache gets created once and reused through out the session. Once the data is cached, it will not change. Static cache can be used to both connected and unconnected lookups ( ex: unconnected lookup uses static cache).

**Dynamic**– Dynamic cache gets created initially and then data kept is inserted/updated during the session .Dynamic cache can be used to only connected lookups (ex: connected lookup).

21.What is worklet?

Worklet is an object representing a set of tasks created to reuse a set of workflow logic in multiple workflows. A repository server controls the complete repository, which includes tables, charts, and various procedures, etc

A powerhouse server governs the implementation of various processes among the factors of the server’s database repository.

**22. What is Informatica PowerCenter?**

Informatica PowerCenter is an ETL/data integration tool that has a wide range of applications. This tool allows users to connect and fetch data from different heterogeneo us sources and subsequently process the same.

**23.How can we delete duplicate rows from flat files?**

We can delete duplicate rows from flat files by using GROUP BY function in an aggregator or in source qualifier in database

**24.What is the difference between DELETE and TRUNCATE command?**

**DELETE** statement is used to remove single or multiple records from exiting table.It is a DML command and this command eliminates records one by one.

**TRUNCATE** command remove the complete data from an existing table but not the table itself. It is a DDL command and this command deletes the entire data page