1. Tell me about yourself?
2. What is your project where you involve?
3. What are your roles & responsibilities?
4. What you done using information tool?
5. **What is DWH?**

**Ans**:- Data warehousing combines data from multiple, variable sources into one database which can be easily manipulated. This can be used for Analysis, Transformation and Reporting. Data warehouse is a subject oriented ,Integrated, Time varying, Non-volatile collection of data in support of the management decision making process.

**Subject Oriented** :This is used to analyze particular subject area.

**Integrated** :-It integrates the data from different sources.

**Time Varient**:-Historical data will be maintained and retrieval of data can be any period time.

**Non-Volatile:-**Once if you place the data in warehouse , it can not be altered ,deleted and update

collection of data for decision making purpose .

**What is ETL process?**

**Ans:-** ETL, which stands for extract, transform and load, is a data integration process that combines data from multiple data sources into a single, consistent data store that is loaded into a data warehouse

1. **What is surrogate key?**

**Ans**:-some time in database table we cannot make primary key from real data. In this situation, we have to add one artificial column in table which is unique and not null, and make this column as primary in table. This primary which is generated from artificial column is called as surrogate key.

1. **What is data mart and its type?**

Ans:- Data Mart is nothing but the smallest version of the DWH. Data Mart deals with single subject area. Data Mart focuses on one area hence they draw data from limited data source.

Its type:-1.depended data mart

2.independed data mart

3.hybrid data mart

1. **What is ODS?**

Ans:- An operational data store (ODS) is a central database that provides a understanding of the latest data from multiple transactional systems for operational reporting. It enables organizations to combine data in its original format from various sources into a single destination to make it available for business reporting.

1. **What is OLTP/OLAP?**

Ans:-1.**OLTP**:- OLTP is nothing but a database which actually stores the daily transactions which is are created from a one and more applications.

Data in OLTP is called as the current data. Mostly normalized data is used in OLTP system.

2.**OLAP**:- OLAP is use to store analyzing the data for decision making and planning ,designing etc.

Data is OLAP Is called as the historical data .Mostly de normalized is OLAP system.

1. **What is dimension table & its types?**

Ans:- A dimension table contains dimensions of a fact. They are joined to fact table via a foreign key. Dimension tables are de-normalized tables. Table which describes dimension involved are called as dimension table.

**Types of Dimensions**

1. SCD

2. Confirmed dimension

3. Degenerated dimension

4. Junk dimension

5. Roll playing dimension

6. Static dimension

7. Shrunken dimension

1. **What is fact table? Its types**

Ans:- Facts are nothing but measures or we can simply called it as numbers. The fact table is a central table in the data schemas. It is found in the center of a star schema or snowflake schema and surrounded by a dimension table. Facts are known as measurements or matrices.

**Types of Facts**

1. Additive Fact

2. Semi-Additive Fact

3. Non-Additive Fact

1. **Tell me schema types?**

Ans:- A database schema defines its entities and the relationship among them. Database schema is a descriptive detail of the database, which can be implemented by means of schema diagrams. All these activities are done by database designer to principle architect in order to give some ease of understanding of detailed database.

Types of schemas

1.star schema

2.snowflake schema

3.fact consolation schemas

1. **What is staging area?**

**Ans:-** After that the next step is extract, where the required data from data source layer is extracted and put into the data staging area. Data staging area is intermediate layer between data source layer and data storage layer used for processing data during the ETL process.

Basically, need staging area to hold the data and to perform data transformation, before loading the data into warehouse.

1. **What is Data cloning?**

**Ans:-** The general process of copying source data to the target database.

1. **How to verify facts & Dimension table**

**Ans**:- If a table has a composite key then it is a fact table. If a table does not have a composite key then it is a dimension table.

1. **What are the active trent used in your project?**
2. **Which validation you perform?**
3. **Latest entry in table (Query)**
4. **Difference between self-join and join**

**Ans:-** self join is nothing but join with it self and

Join is used to return only matching records.

1. **In my schema, 100 tables are there, some tables have some column name, I want to display how many tables have some column name?**

**Ans:-**select \* from information\_schema.columns where table name = “table name”

1. **What is Fact table? What is fact table in your project?**
2. Ans:- Facts are nothing but measures or we can simply called it as numbers. The fact table is a central table in the data schemas. It is found in the center of a star schema or snowflake schema and surrounded by a dimension table. Facts are known as measurements or matrices.

1. Transaction Fact Table: The transaction fact table is a basic approach to operate the businesses. These fact tables represent an event that occurs at the primary point. A line exists in the fact table for the customer or product when the transaction occurs.

1. **What is SCD1, SCD2, SCD3**

**Ans:-**dimensions that changes slowly over a period of time ,rather than changing on regular schedule. A slowly changing dimension is a dimension that stores and manages both current and historical data over time in a data warehouse.

1. **What is PK, FK, Union all?**

**Ans:-1) primary key:-** primary key is always identify the unique records in to column of the table. basically primary key is defined in numeric values.

**2)foreign keys:- f**oreign keys is column is one table that refer to the primary key in another table.null value can be allowed in foreign key column.

3)**union all**:- The operator is used to combine two or more table using select statement when both the table have same no of columns.

1. **How to update all columns at a time?**

Ans:- update table name set column name1,column name2,column name3 where condition

1. **Display 5th highest salary?**

Ans**:- 1. With fifth\_salary as**

(select \*,dense\_rank () over (order by salary desc) as new from table name)

Select \* from fifth\_salary where new=5;

2.select max(salary) from table name where salary<( select max(salary) from table name where salary<( select max(salary) from table name where salary<( select max(salary) from table name where salary<( select max(salary) from table name ))))

**27.Difference between Delete, Truncate, Drop?**

Ans:-1)delete:-delete statement use to delete the data from table row by row.by using delete statement we can’t delete the structure. We can delete the table data at one time and row by row by specifying an condition.

2)truncate:-truncate statement allow you to delete the records from a table at once.we can’t delete the structure .in truncate we cant’t delete the data row by row by specifying a condition.

3)drop:-drop statement will delete the table structure as well as table data .drop statement we can drop or delete the database.

1. **Difference between Union & Union all**

**Ans:-** 1) union:- the union operator is used to combine the result set of two or more select statement .

2)union all:-this operator is used to combine two or mor tables using select statement when tables using select statement when both the tables have same no of columns.

1. **Which process running n Unix, How to know?**

Ans:- Type the ps command to see all running process in Unix.

1. **How to Kill process in Unix?**

Ans:- Type the kill command

1. **Display 3 to 7 rows in table?**

**Ans:-** select \* from table\_name where id between 3 and 7

1. **Display duplicate records?**

**Ans:-** select column1,column2,….column n ,count(\*) as duplicate forom table name group by column1,column2,….column n having count(\*)>1;

1. **Delete Duplicate record**

**Ans:-** with duplicate as,

(select \*,dense\_rank() over (order by salary desc) as new froom table\_name)

Delete from duplicate where new >1;

1. **Testing Life cycle**
2. **Bug Life cycle?**
3. **Priority severity?**
4. **Display Even and odd number**

**Ans:-**1)**even**:- select \* from table\_name where id %2=0

2)**odd**:- select \* from table \_name where id %2=1

select \* from table \_name where id %2!=0;

1. **Display particular record in table**

Ans:- select \* from table\_name where specified the condition;

1. **Only distinct value**

Ans:- select distinct from table\_name

1. **Covert null value columns in a table**

Ans:- select isnull(isnull,”no\_data”) from table \_name

1. **Difference between view an materialize view**

**Ans:-** A view uses a query to pull data from the underlying tables. A materialized view is a table on disk that contains the result set of a query

1. **How to design test case to test ETL Job?**

**Ans:-** Analyze the requirement, understanding the data model.List out the source and target, analysis the data workflow.Deeply go through the ETL mapping to understand the ETL transformation rule.

1. **How to understand customers requirements?**

**Ans:-** Create a buyer's persona. To understand your customers' needs, who your customer is. give feedback from customers. Analyze competitors.a customer needs statement.

1. **What is operation data store & staging area?**

Ans:- An operational data store (ODS) is a central database that provides a understanding of the latest data from multiple transactional systems for operational reporting. It enables organizations to combine data in its original format from various sources into a single destination to make it available for business reporting.

After that the next step is extract, where the required data from data source layer is extracted and put into the data staging area. Data staging area is intermediate layer between data source layer and data storage layer used for processing data during the ETL process.

Basically, need staging area to hold the data and to perform data transformation, before loading the data into warehouse.