* What is ETL:

ANS:

ETL stands for Extract, Tranform, Load which is a process used to collect data from multiple system and store into one system

It consist of 3 steps.

Extraction: in this process data from multiple database, files, legacy system, web server is collected.

Transform: In this stage all data quality is check & then it is cleaned. Process for cleansing, deduplication, verification is applied. to transform it

Load: In this stage data is load to a new destination containing valuable form like data warehouse.

4

* **What is ELT:**

It is similar to ETL but it occurence of step are dissimilar to ELT here seq is Extract, load, Transform

Data from different sources is collected & load into one system then transform it is more efficient then ETL because Transformation is done after load as transform take Processing power of database to work better

* 3 tier Architecture in DE :-

ANS

The 3 tier architecture in is a design pattern for building data processing system

The 3 layers are

Application : It is responsible for business logic that includes process & services to transform data like ETL, machine learning model etc

Presentation : It is responsible for UI It enables user to interact with the system in form of deshboard, report & visualization.

Data layer : It is responsible for storing data either in raw or processed form it include data lake, warhouse or database.

* ETL tool (any 3)

1) Informatica Power Center
2) Apache Sprit
3) IBM infoSphere DataStage.