Ajay Boddu
359962

1. *What is ETL?*

   ETL stands for Extract, Transform, Load. It's a process used in data management to pull data from various sources, convert it into a usable format, and then load it into a target system like a data warehouse.

2. *How did ETL evolve?*

   ETL has grown from basic data movement tasks to more advanced tools that handle complex data transformations. This evolution was driven by the increasing need for businesses to manage and analyze large volumes of data efficiently.

3. *What is the difference between ETL and ELT?*

   The key difference is in the order of operations. ETL transforms the data before loading it into the destination system, while ELT loads the raw data first and then transforms it in the target system, usually leveraging the processing capabilities of modern databases.

4. *How does ETL work?*

   ETL works in three steps:

   - *Extract:* Data is gathered from multiple sources.

   - *Transform:* The data is cleaned and modified to fit the business needs.

   - *Load:* The transformed data is then loaded into the target system, like a data warehouse.

5. *What are four other types of data integration methods?*

   - *Data Replication:* Copies data from one system to another.

   - *Data Virtualization:* Allows access to data without moving it physically.

- *Data Federation:* Combines data from different sources into a unified view without physically merging the data.

   - *Change Data Capture (CDC):* Tracks and records changes in data in real-time.

6. *What are the benefits of ETL?*

   ETL offers several benefits, such as providing consistent and reliable data for analysis, improving data quality, ensuring regulatory compliance, and helping businesses make better-informed decisions by integrating data from various sources.

7. *Four capabilities of ETL tools:*

   - *Data Cleansing:* Ensures the quality of data by removing errors and inconsistencies.

   - *Data Transformation:* Converts and reshapes data to meet specific requirements.

   - *Data Integration:* Merges data from different sources into a cohesive format.

   - *Job Scheduling:* Automates the ETL process, ensuring data is regularly updated.