

ETL vs ELT

What's the
Difference?



What is ETL?

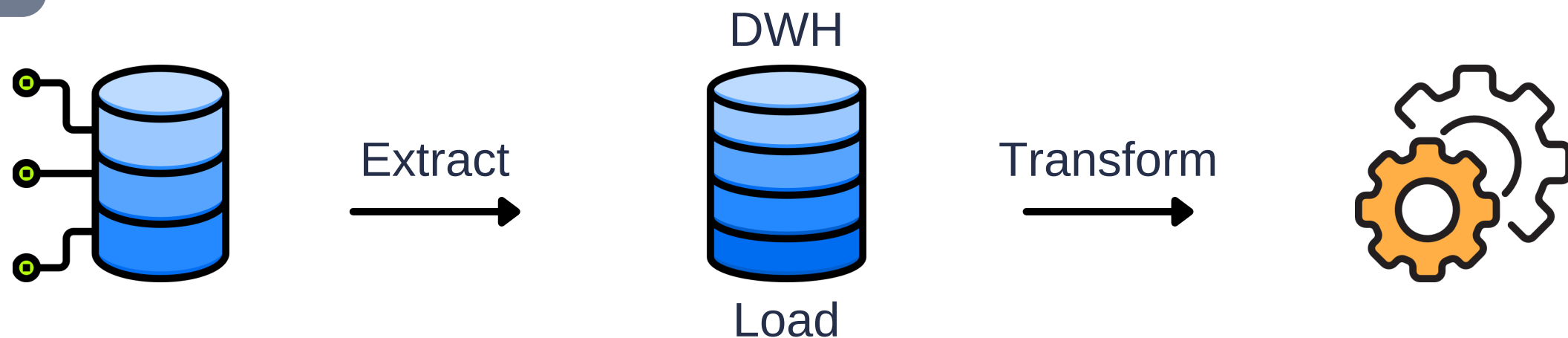
ETL



Extract, transform, and load (ETL) is a data integration methodology that extracts raw data from sources, transforms the data on a secondary processing server, and then loads the data into a target database.

What is ELT?

ELT



ELT (extract, load, and transform) loads raw data directly into a target data warehouse. Data cleansing, enrichment, and data transformation all occur inside the data warehouse itself. Raw data is stored indefinitely in the data warehouse, allowing for multiple transformations.

What is ELT?



Cloud-based data warehouses offer near-endless storage capabilities and scalable processing power. For example, platforms like **Amazon Redshift** and **Google BigQuery** make ELT pipelines possible because of their incredible processing capabilities.

ETL

Advantages

Maturity. ETL was developed first and has been in practice for more than two decades. This means that there are more engineers with experience in ETL implementations and more ETL tools in the marketplace to build data pipelines within organizations.

Compliance. ETL transforms data before it reaches its destination. When companies are subject to data privacy regulations such as GDPR, ETL allows them to remove, mask, or encrypt sensitive data before it's loaded to the data warehouse to ensure compliance.

Drawbacks

Frequent maintenance. ETL data pipelines handle both extraction and transformation. But they have to undergo refactors if analysts require different data types or if the source systems start to produce data with deviating formats and schemas.

Higher upfront cost. Defining business logic and transformations can increase the scope of a data integration project.

ELT

Advantages

High speed. ELT allows for all of the data to go into the system immediately, and from there, users can determine the exact data they need to both transform and analyze.

Lower cost. Requires a less-powerful server for data transformation and takes advantage of resources already in the warehouse. This results in cost savings and resource efficiencies.

Flexibility. Analysts no longer have to determine what insights and data types they need in advance but can perform transformations on the data as needed in the warehouse.

Drawbacks

Security gaps. Storing all the data and making it accessible to various users and applications come with security risks. Companies must take steps to ensure their target systems are secure by properly masking and encrypting data.

Increased latency. The need to continually transform data slows down the overall time it takes to perform queries/analysis.

Was this helpful?



Follow for more updates

 **ginacostag**