

3_ETL vs ELT

Tuesday, February 10, 2026 2:50 PM

What's the Difference Between ETL and ELT?

- **Extract, transform, and load (ETL)** and extract, load, and transform (ELT) are two data-processing approaches for analytics. Large organizations have several hundred (or even thousands) of data sources from all aspects of their operations—like applications, sensors, IT infrastructure, and third-party partners.
- They have to filter, sort, and clean this large data volume to make it useful for analytics and business intelligence. The ETL approach uses a set of business rules to process data from several sources before centralized integration.
- The ELT approach loads data as it is and transforms it at a later stage, depending on the use case and analytics requirements.
- The ETL process requires more definition at the beginning.
- Analytics must be involved from the start to define target data types, structures, and relationships.

Data scientists mainly use ETL to load legacy databases in the data warehouse, while ELT has become the norm today.

What are the similarities between ETL and ELT?

Both extract, transform, and load (ETL) and extract, load, and transform (ELT) are sequences of processes that prepare data for further analysis. They capture, process, and load data for analysis across three steps.

Extraction

Extraction is the first step of both ETL and ELT. This step is about collecting raw data from different sources. These could be databases, files, software as a service (SaaS) applications, Internet of Things (IoT) sensors, or application events. You can collect semi-structured, structured, or unstructured data at this stage.

Transformation

In the ETL process, transformation is the second step, while in ELT it is the third. This step focuses on changing raw data from its original structure into a format that meets the requirements of the target system where you plan to store the data for analytics. Here are some examples of transformation:

- Changing data types or formats
- Removing inconsistent or inaccurate data.
- Removing data duplication.

You apply rules and functions to clean and prepare data for analysis in the target system.

Load

In this phase, you store data into the target database. ETL processes load data as a final step, so that reporting tools can use it directly to generate actionable reports and insights. However, in ELT, you still need to transform the extracted data after loading it.

How do the ELT and ETL processes differ from each other?

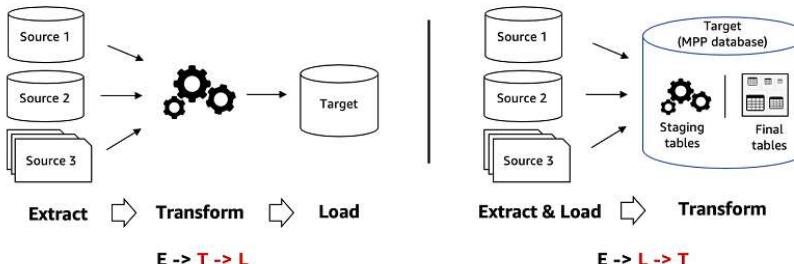
Next, we outline the processes of extract, transform, and load (ETL) and extract, load, and transform (ELT). You can also read some historical background.

ETL process

ETL has three steps:

1. You extract raw data from various sources
2. You use a secondary processing server to transform that data
3. You load that data into a target database

The transformation stage ensures compliance with the target database's structural requirements. You only move the data once it is transformed and ready.



ELT process

These are the three steps of ELT:

1. You extract raw data from various sources
2. You load it in its natural state into a data warehouse or data lake
3. You transform it as needed while in the target system

With ELT, all data cleansing, transformation, and enrichment occur within the data warehouse. You can interact with and transform the raw data as many times as needed.

History of ETL and ELT

ETL has been around since the 1970s, becoming especially popular with the rise of data warehouses. However, traditional data warehouses required custom ETL processes for each data source.

The evolution of cloud technologies changed what was possible. Companies could now store unlimited raw data at scale and analyze it later as required. ELT became the modern data integration method for efficient analytics.

Key differences: ETL vs. ELT

Extract, load, and transform (ELT) has improved extract, transform, and load (ETL) in several ways.

Transform and load location

Transformation and load occur in different locations and use distinct processes. The ETL process transforms data on a secondary processing server.

In contrast, the ELT process loads raw data directly into the target data warehouse. Once there, you can transform the data whenever you need it.

Data compatibility

ETL is best suited for structured data that you can represent in tables with rows and columns. It transforms one set of structured data into another structured format and then loads it.

In contrast, ELT handles all types of data, including unstructured data like images or documents that you can't store in tabular format. With ELT, the process loads the various data formats into the target data warehouse. From there, you can transform it further into the format you require.

Speed

ELT is faster than ETL. ETL has an additional step before it loads data into the target that is difficult to scale and slows the system down as data size increases.

In contrast, ELT loads data directly into the destination system and transforms it in parallel. It uses the processing power and parallelization that cloud data warehouses offer to deliver real-time or near real-time data transformation for analytics.

Costs

The ETL process requires analytics involvement from the start. It needs analysts to plan ahead on the reports they want to generate and define data structures and formatting. The time required for setup increases, which adds to costs. Additional server infrastructure for transformations may also cost more.

ELT has fewer systems than ETL, as all transformations occur within the target data warehouse. With fewer systems, there is less to maintain, leading to a simpler data stack and lower setup costs.

Security

When you work with personal data, you must comply with data privacy regulations. Companies must protect personally identifiable information (PII) from unauthorized access.

In ETL, developers have to build custom solutions, like masking PII to monitor and protect data.

On the other hand, ELT solutions provide many security features—like granular access control and multifactor authentication—directly within the data warehouse. You can invest more time in analytics and less time meeting data regulation requirements.

ETL vs ELT — When to Choose What

ETL (Extract → Transform → Load)

Use ETL when **you must clean, validate, or reshape data before it enters the target system**.

When ETL is the right choice

- **Source systems are messy or sensitive**
You need to enforce strict validation, cleansing, or business rules before loading.
- **Target system cannot handle heavy transformations**
Example: traditional data warehouses (SSIS → SQL Server DW).
- **You need strong control over data quality upfront**
Healthcare, finance, compliance-heavy environments.
- **Data volumes are moderate**
Transformations happen on an ETL engine (SSIS, Informatica), not on a scalable compute cluster.
- **You want a curated, clean dataset before loading**
Staging → Transform → Load into DW.

Typical tools

SSIS, Informatica, Talend, DataStage, ADF Mapping Data Flows (acts like ETL).

ELT (Extract → Load → Transform)

Use ELT when **your target system is powerful enough to transform data after loading**.

When ELT is the right choice

- **You have a modern cloud data warehouse**
Snowflake, BigQuery, Synapse, Databricks — all design for in-warehouse transformations.
- **You deal with large or semi-structured data**
JSON, Parquet, logs, clickstream, IoT.
- **You want to load raw data quickly**
“Load first, transform later” supports agility and schema evolution.
- **You want to support multiple downstream transformations**
Raw → Clean → Curated → Gold layers (medallion architecture).
- **You prefer SQL-based transformations**
dbt, Databricks SQL, Synapse SQL.

Typical tools

ADF (Copy Activity + SQL scripts), dbt, Databricks notebooks, Snowflake tasks.

The simplest rule to remember

If your warehouse is powerful → choose ELT.

If your warehouse is traditional or transformations are sensitive → choose ETL.