🎓 Video 12 Summary: ETL vs. ELT – Data Pipeline Integration Techniques

📦 Why Data Pipeline Organization Matters

As a cloud data analyst, how you **organize your data pipeline** directly affects the **efficiency** and **effectiveness** of data storage and analysis.

🔄 Two Main Integration Techniques

You'll choose between two pipeline models based on business needs:

1. ETL – Extract, Transform, Load

- **Extract**: Gather data from one or more sources.
- **Transform**: Clean and standardize the data.
- **Load**: Insert the transformed data into a target system (e.g., database, warehouse, data lake).

✅ **Best for**: Traditional systems, structured data, and when transformation must happen before storage.

2. ELT – Extract, Load, Transform

- **Extract**: Gather raw data.
- **Load**: Insert raw data directly into the destination system.
- **Transform**: Clean and format the data **after** loading.

✅ **Best for**: Cloud-based systems, large datasets, and real-time analytics.

🧸 Examples to Illustrate

🎓 ETL Example: University

- Combines data from student systems, financial records, and course catalogs.
- Enables analysis of enrollment, finances, and course performance.

🏭 ELT Example: Manufacturing Company

- Collects data from sensors, control systems, and quality checks.

- Loads data immediately for **near-real-time** decision-making.
- Transformation happens later, offering **flexibility** and **speed**.

💡 Why Choose ELT?

- **Saves time**: Especially with large or streaming datasets.
- **Scalable**: Cloud platforms handle large volumes efficiently.
- **Flexible**: You can decide how to transform data after loading.

🧠 Takeaway

Both **ETL** and **ELT** are valid approaches to building data pipelines.
Your choice depends on:

- The **timing** of transformation
- The **volume** of data
- The **tools** and **infrastructure** available