

## Video 20 – Challenges of Data Transformation

### Why Data Transformation Matters

- **Data is a raw material** for business decisions.
  - Organizations are managing **larger volumes of data** than ever before.
  - Cloud data analysts help uncover insights from this growing data flow.
- 

## Key Challenges in Data Transformation

### 1. Resource Intensity

- **Computational Power:** Large datasets require significant processing capacity.
- **Cloud Computing:** Offers scalable, cost-effective access to virtual resources.
- **Cost Planning:** Even cloud services must be budgeted carefully.

### 2. Storage Management

- **ETL vs. ELT Pipelines:** Both involve extracting, transforming/loading data into warehouses or lakehouses.
- **Unused Data:** Can lead to unnecessary storage costs.
- **Solution:** Regularly review and delete unused data.

### 3. Time and Skill Requirements

- **Skilled Personnel:** Transformation requires expertise across tools and techniques.
  - **Tool Selection:** Not all tools are equal—choose based on task and context.
- 

## Data Integrity: The Ongoing Challenge

- **Definition:** Accuracy, completeness, consistency, and trustworthiness of data.
- **Sources of Errors:**
  - Manual entry mistakes (e.g., typos)
  - Machine-generated errors (e.g., bulk reading issues)
  - Transformation errors (e.g., wrong formats, missing values, incorrect aggregations)

- **Impact:** Errors can compromise reports and dashboards, leading to poor decisions.
  - **Solution:** Implement a strong integrity management plan to minimize mistakes.
- 



## Your Role as a Cloud Data Analyst

- **Responsibility:** Extract value from complex, massive datasets.
- **Strategy:** Build a transformation plan that balances resource use and data integrity.
- **Goal:** Deliver reliable insights to support stakeholders and business growth.