

## Video 9 Summary: Introduction to Data Pipelines

### Welcome Back!

You're progressing well in your journey as a cloud data analyst!

---

### Recap: The Five Stages of the Data Journey

1. **Collect**
2. **Process**
3. **Store**
4. **Analyze**
5. **Activate**

These stages guide how data is handled from raw input to actionable insights.

---

### Focus of This Video: Data Pipelines

Data pipelines help convert raw data into a format ready for **storage** and **analysis**, supporting business needs efficiently.

---

### The Three Main Stages of a Data Pipeline

#### 1. Extract

- Data is **collected** from various sources.
- This stage involves **ingestion** of raw data.

#### 2. Transform

- Data is **cleaned, standardized, and manipulated**.
- Ensures consistency and usability for analysis.

#### 3. Load

- Transformed data is **loaded into storage systems**.

- Makes data available for further analysis and visualization.

---

### Takeaway

Understanding how **ETL (Extract, Transform, Load)** pipelines work is essential for building scalable, reliable data workflows in the cloud.