GCP CDA Module 3 - Video 4 Summary: The Data Lifecycle

Understanding how data moves through its entire life helps cloud professionals stay organized and efficient. The data lifecycle consists of six key stages: Plan, Capture, Process, Manage, Analyze, Archive, and Destroy. Each stage plays a vital role in handling data responsibly and effectively.

Plan

This stage begins before any analysis. It involves defining a business question or objective, deciding what data to collect, who is responsible for each task, and how success will be measured.

Capture

Data is collected from internal or external sources. This stage also identifies gaps in current data collection and improves them through iteration.

Process

Raw data is cleaned, transformed, compressed, and encrypted to make it usable for analysis. This ensures data is in the right format and secure.

Manage

Ensures proper data maintenance and secure storage. This is an ongoing process throughout the project for all types of data.

Analyze

Data is used to answer business questions or meet objectives. Analysts find trends, create visualizations, and make recommendations based on insights.

Archive

Data is stored for future use if needed. This stage ensures data is preserved according to organizational guidelines.

Destroy

Data that is no longer useful is securely deleted to prevent breaches and comply with privacy regulations like GDPR.

Roles in the Data Lifecycle

Each stage of the data lifecycle involves different data professionals:
- Data Analysts: Work primarily during the Analyze stage to interpret data and provide insights.
- Data Engineers: Handle the Process and Manage stages by building infrastructure and transforming data.

- Data Architects: Design the overall data management plan and structure.
- Data Scientists: Use data to create models and understand complex patterns.