**Task 08**

- ELT vs ETL (Main difference, when to use which one)

- Batch vs Streaming Pipeline (Main difference, when to use which one, demonstrate a proper use-case)

**Answer**

**ELT vs ETL**

ELT (Extract, Load, Transform) and ETL (Extract, Transform, Load) are two data integration processes that differ mainly in the sequence of steps.

In ETL, data is extracted from the source, transformed to fit the target system, and then loaded into the destination. This method suits environments where data integrity and accuracy are crucial, and the target system has limited processing power or requires complex transformations.

ELT, however, involves extracting data, loading it directly into the target system, and then performing transformations within that system. This approach leverages the computational power of modern data warehouses and data lakes, making it ideal for big data scenarios where quick loading and efficient in-system transformations are necessary.

**Batch vs. Streaming Pipeline**

Batch processing handles large data sets collected over time, processing them together. It suits tasks that can tolerate delays, such as end-of-day reporting, monthly payroll, or periodic data integration, where real-time processing is unnecessary.

Streaming processing, in contrast, deals with continuous data streams, processing data in real-time as it arrives. This is crucial for use-cases needing immediate insights or actions, such as real-time fraud detection, live analytics, and monitoring systems, where timely data processing and immediate responses are essential.

**Use-Case Demonstration**

An e-commerce platform can use an ETL batch processing pipeline for end-of-day sales reports. Data from sources like sales, inventory, and customer databases is extracted, transformed to calculate daily metrics, and loaded into a data warehouse. This ensures accurate and comprehensive reporting without needing real-time data.

For real-time order monitoring to detect fraud or ensure timely fulfillment, a streaming ELT pipeline is better. Data from customer orders is extracted and loaded into a data lake as generated, with immediate transformations to detect anomalies or trigger alerts for out-of-stock items, ensuring rapid responses.

Thus, ETL with batch processing is suitable for reporting, while ELT with streaming is ideal for real-time monitoring, optimizing efficiency for both needs.