BWT Task-07 Exercise

Submitted By: ABRAR SAEED

ELT VS ETL

ELT:

- Stands for Extract, Load and Transform
- Transformation happens after loading into the target database which is why the target database needs more computing power
- It is used in scenarios where speed of data loading is crucial which means ELT are used in big data and real time applications
- ELT provides the flexibility to transform data based on the use cases.
- ELT can be used where flexible data exploration and cloud based solutions are needed.
- **Use Case:** Real-time analytics for business intelligence.

ETL:

- Stands for Extract, Transform and Load.
- Transformation happens before loading into the target database which is why the target database needs less computing power.
- It is used in scenarios where complex transformations are required before the data can be used.
- ETL is used in scenarios when the target database lacks the computing power to efficiently handle large scale transformations.
- Used in historical data processing.
- **Use Case:** Daily Sales Reports

Batch VS Streaming Pipeline

Batch Pipeline:

- Its an approach to process and analyze large volumes of data in the form of batches which have been collected over a period
- It involves stages like data collection, storage, process and reporting.
- It is used when the analysis of large sets of data is to be carried out to find trends and patterns.
- Time insensitive complex calculations can be carried out with the help of batch ideology.

E-Commerce Sales Analysis

- Analysis of monthly sales data to understand trends can be a use case for this.

- Collect sales data from all sources.
- Store sales data in a warehouse.
- Aggregate and analyse monthly data in form of batches.
- Generate reports and make dashboards.

Streaming Pipeline:

- It is an approach to process data in real-time as the data is generated.
- It involves stages like data ingestion, real-time processing, data storage and real-time monitoring and reporting.
- It is used in scenarios of event detection.
- For the purposes of tracking and responding to user interactions in real time.

Real-Time Fraud Detection

- Detection of fraudulent activities in real-time to minimize financial loss can be a use case for this.
- Continuously ingest real-time transaction data.
- Analyze each transaction as it occurs.
- Detect fraud in real time.
- Block and notify.