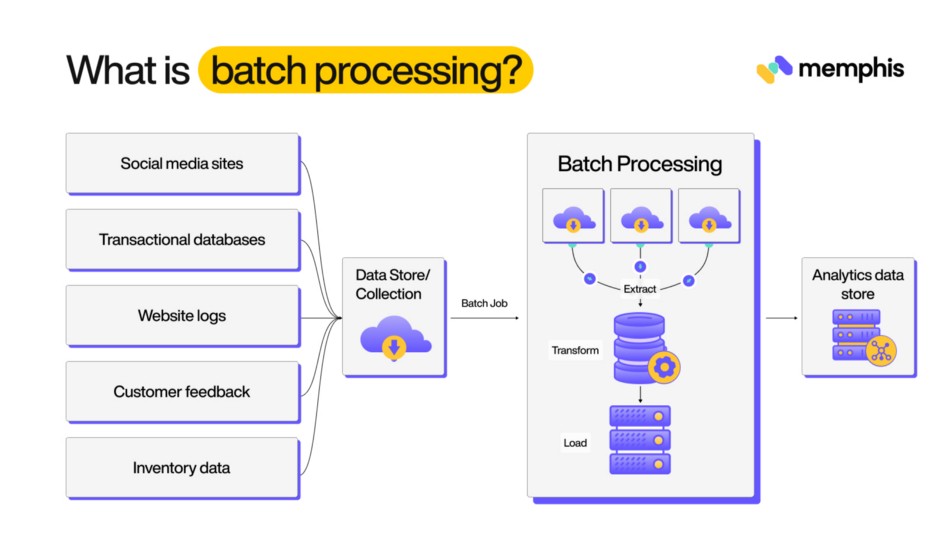
Batch Processing

* Done at the specific scheduled time
* Size of the data is known
* Automated using workflow orchestration tools like Airflow, prefect, flyte, dagster



Useful in following cases

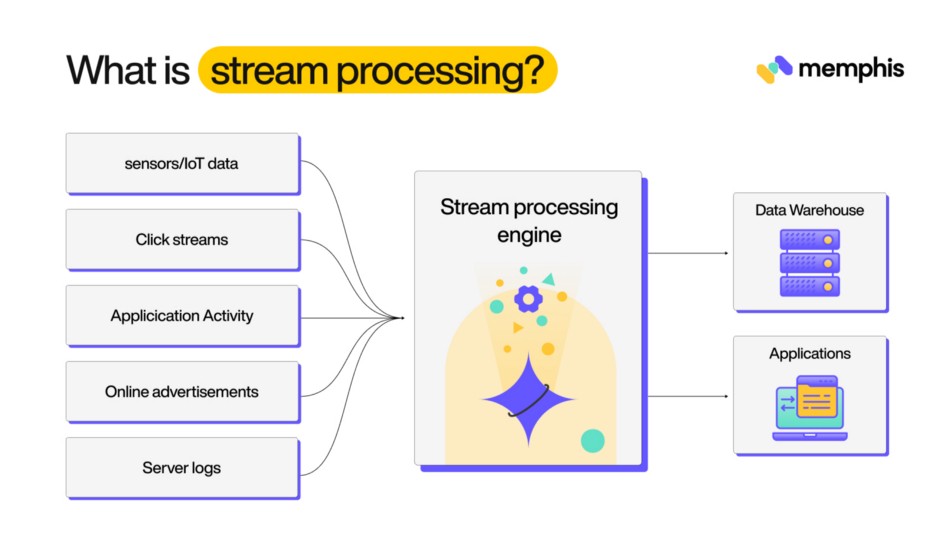
* Anomaly detection
* Customer segmentation
* Payroll systems
* Banking systems
* Billing Service

Challenges in Batch Processing

* Needs human monitoring
* Debugging in the batch process Is Difficult
* Change in data will be delayed till the next batch runs
* Dependent delays are created
* Needs more time in processing large scale data.

Stream Processing

* Extracting, processing, and delivering data in real-time
* It is a stateless operation
* It is used to generate alerts in case of any errors are detected.
* Process large volumes of data with minimum delay (low latency)



* Stream processing enables organization to analyze time-series data and identify pattern in them
* Size of the data is unknown
* Market changes ,consumer needs and business possibilities are more quick by stream processing
* Platforms
  + Spark Streaming
  + Memphisdev
  + Kafka
* Used to get immediate analytics insights.
* Fraud Detection :
* Sentiment Analysis
* Log Monitoring
* Customer Satisfaction

Challenges :

* Scalabilty is stream processing is challencing when error happens
* Data is not always consistent and durable.
* Application must be cognizant of their assumptions on ACID transactions when evaluation data streams
* 