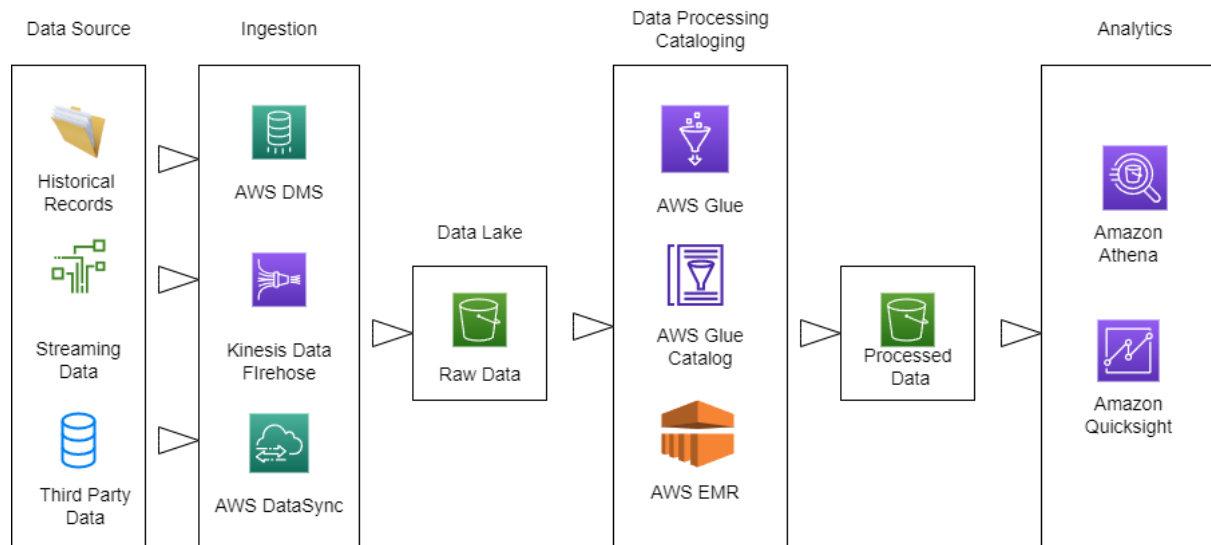# Final Capstone Project

Task: To ingest various forms of company, store them in a Data Lake such that the data can be processed later to get insights from the data.

Solution:
Data Ingestion Techniques:

- Streaming Data: AWS Kinesis Data Firehose
- Historical Records: AWS DMS
- Third Party Data: AWS Data Sync

Ingestion of the data into a S3 bucket makes it a data lake. We catalogue the data using AWS Glue Catalog, AWS Glue and AWS EMR as the company is familiar with Apache Hadoop based Software. Then for deriving insights from the data we use Amazon Athena and Amazon Quicksight.