

# Serverless Data Lake for Real-Time Financial Analytics

BATCH-13

2200031095 - N KAVYA SRI

2200030920 - CHAITANYA

2200031952 - S ABHIRAM

# Problem Statement & Objectives

## ► Problem Statement:

- ▶ Financial institutions face challenges in processing and analyzing massive transaction data in real-time.
- ▶ Existing systems are often costly, slow, and require heavy infrastructure.

## ► Objectives:

- ▶ Build a scalable, cost-effective serverless data lake.
- ▶ Enable real-time financial transaction analytics.
- ▶ Integrate machine learning for fraud detection and trend prediction.

# Background & Literature Review

- ▶ Traditional financial analytics systems rely on on-premise servers, leading to high costs and slow scalability.
- ▶ Cloud-based serverless computing offers pay-as-you-go flexibility.
- ▶ Literature highlights AWS Lambda, Glue, Athena, and SageMaker as effective tools for big data and ML in finance.
- ▶ Predictive analytics in finance improves fraud detection and forecasting accuracy.

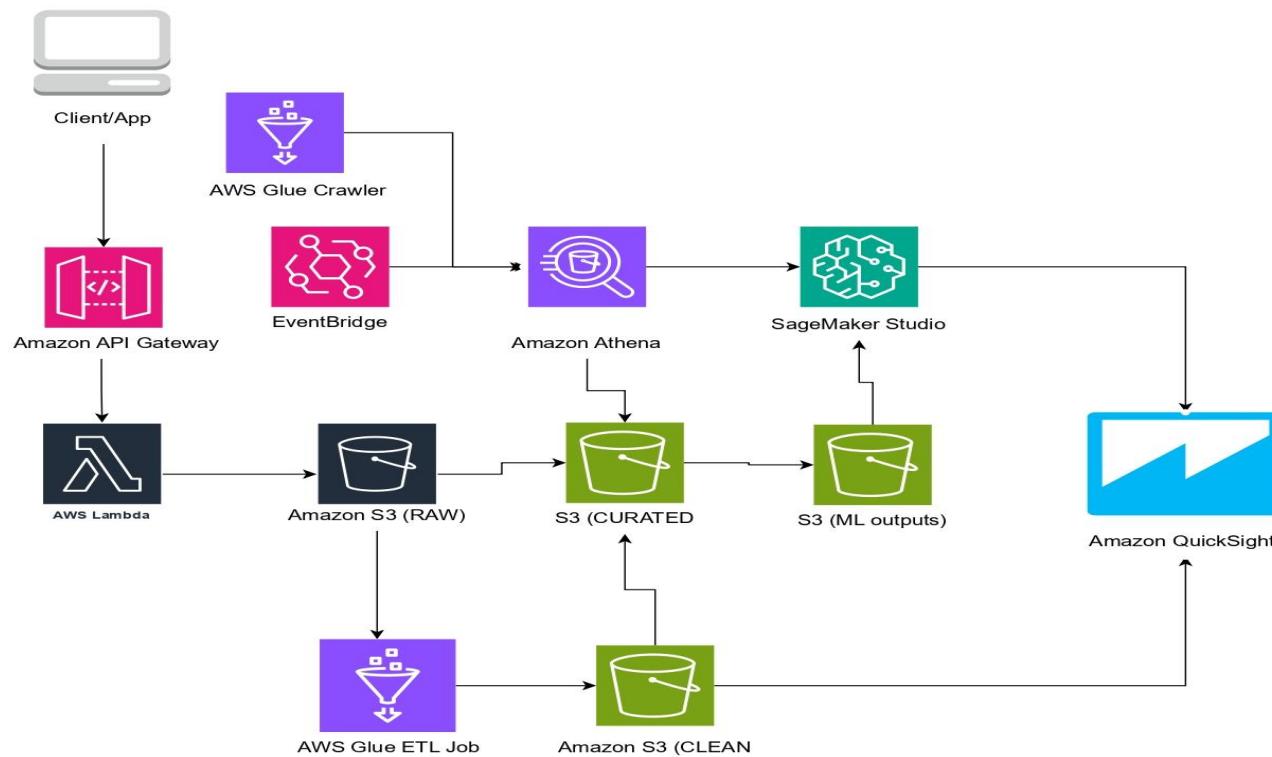
# Methodology & Planning

- ▶ 1. Requirement analysis and architecture design.
- ▶ 2. Data ingestion pipeline using AWS Lambda and Kinesis.
- ▶ 3. Storage in Amazon S3 (Data Lake).
- ▶ 4. ETL with AWS Glue.
- ▶ 5. Querying with Athena.
- ▶ 6. ML model development in SageMaker.
- ▶ 7. Testing, validation, and optimization.
- ▶ Timeline: 5 phases over ~3 months.

# Tech stack and services used:

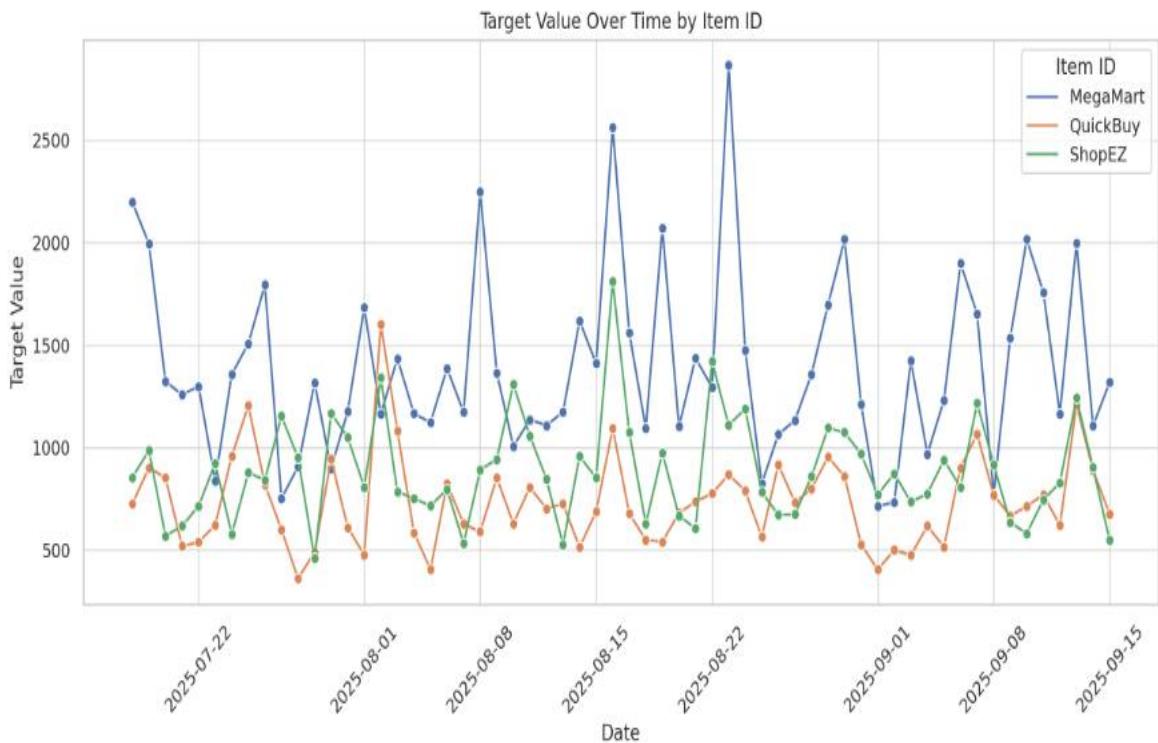
- ▶ **AWS Lambda** (event-driven processing)
- ▶ **Amazon S3** (acts as the Data Lake)
- ▶ **AWS Glue** (for cleaning and transforming data)
- ▶ **Amazon Athena** (serverless SQL querying directly on S3)
- ▶ **AWS SageMaker** (fraud detection, trend prediction models)
- ▶ **Amazon QuickSight** (for BI dashboards and insights)

# Architecture Diagram



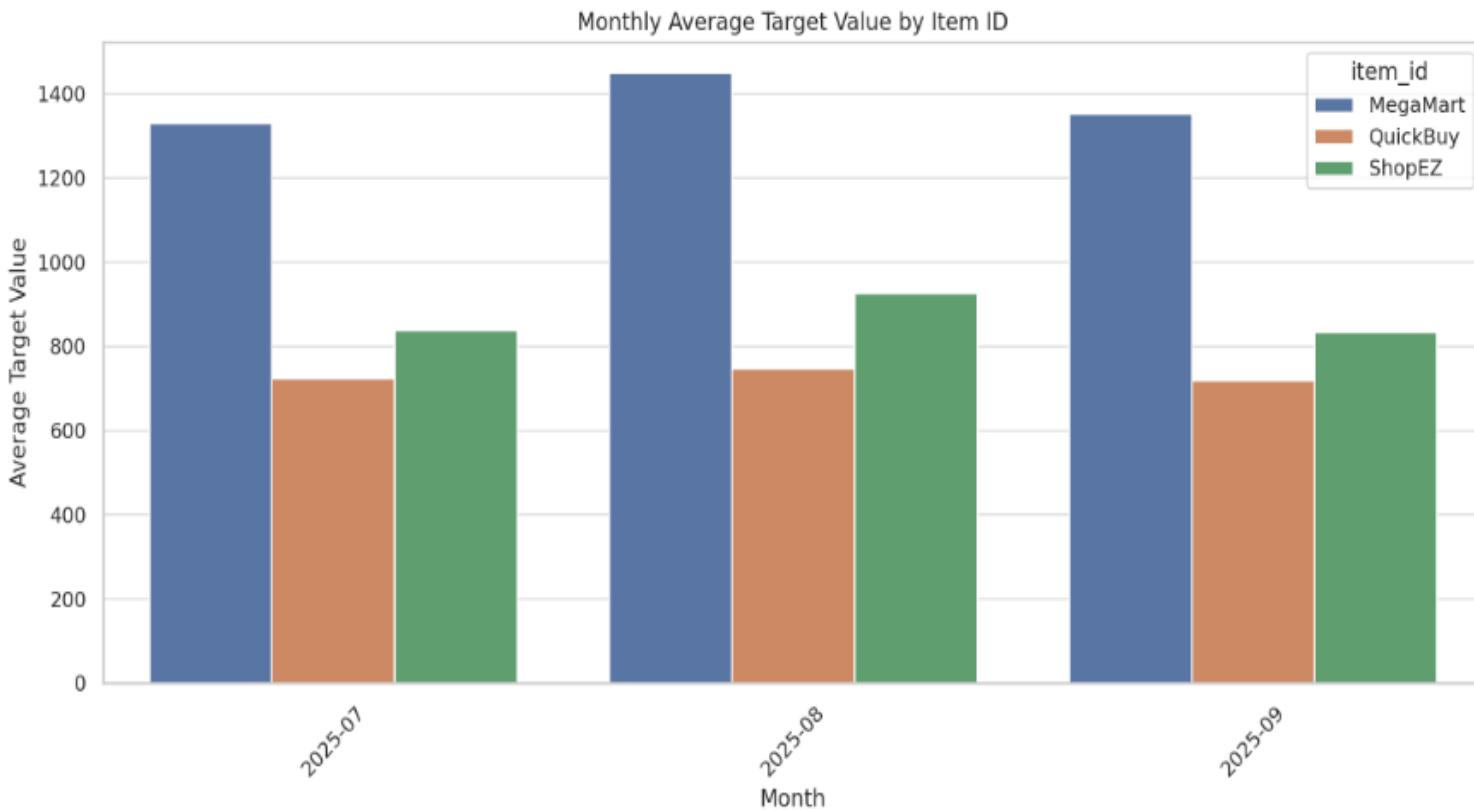
# Outputs:

[2]



...	timestamp	item_id	target_value
0	2025-07-18	MegaMart	2197.40
1	2025-07-18	QuickBuy	725.93
2	2025-07-18	ShopEZ	854.44
3	2025-07-19	MegaMart	1995.04
4	2025-07-19	QuickBuy	901.37

# Outputs:



# Stakeholder & User Need Analysis

## ► Stakeholders:

- ▶ Financial institutions
- ▶ Data analysts
- ▶ Risk management teams
- ▶ Compliance officers

## ► User Needs:

- ▶ Fast, accurate financial data insights.
- ▶ Real-time fraud detection.
- ▶ Easy-to-use dashboards.
- ▶ Secure, compliant data handling.

# Presentation & Communication

- ▶ Clear and concise presentation of objectives, methods, and results.
- ▶ Use of visual aids (architecture diagrams, flowcharts).
- ▶ Real-time demo of analytics capabilities.
- ▶ Effective Q&A handling to address stakeholder queries.



THANK  
YOU