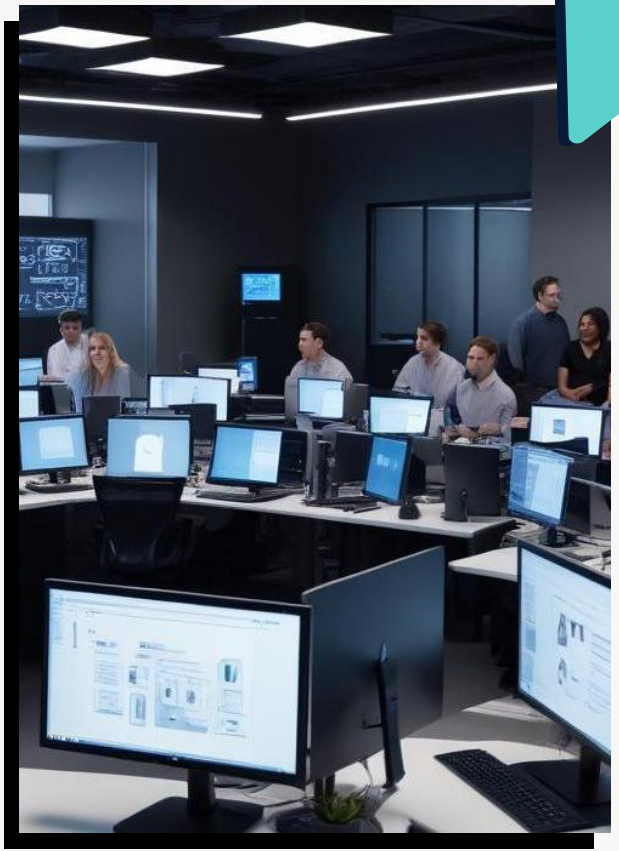
Serverless Data Lake for Real-Time Financial Analytics

Abstract

The project aims to develop a serverless data lake architecture for real-time financial analytics, enabling companies to analyze and process large volumes of financial data in a cost-effective manner. Traditional financial analytics systems often face challenges related to scalability, real-time data processing, and high infrastructure costs. By utilizing serverless technologies such as AWS Lambda, Amazon S3, and Amazon Athena, the proposed solution eliminates the need for server management while providing a highly scalable, flexible, and affordable platform for storing and processing financial data.



Key focus

- Main focus: To develop a **serverless data lake** architecture
- Purpose: Enable **real-time financial analytics**
- Goal: Process large financial datasets **efficiently and cost-effectively**

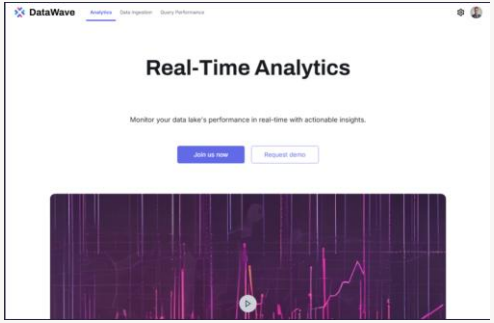
Methodologies

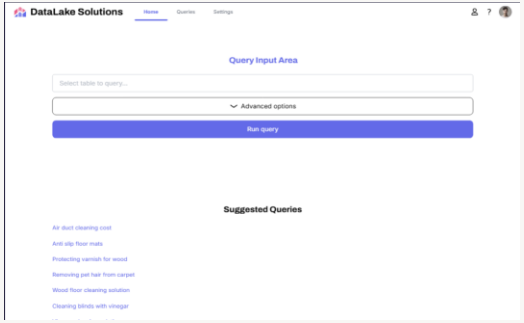
- **Serverless architecture** using event-driven data processing
- Use of **data ingestion, transformation, and real-time analysis**
- **Streamlined pipeline** for collecting and analyzing data using cloud-native tools

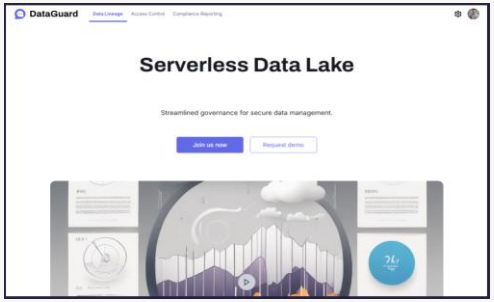
Tools & Technologies

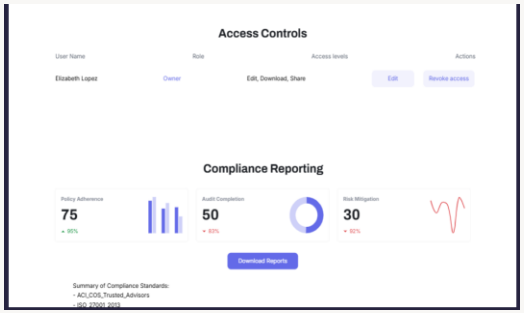
- **AWS Lambda** – for serverless compute
- **Amazon S3** – for data storage
- **Amazon Athena** – for query processing
- **Event-driven processing-**and anaytics pipeline
- **Amazon Athena**-Used for visualization

Model Design









Limitations

- Latency in querying large datasets with Athena
- Cost management for high-volume data processing
- Requires strong **governance** and **data compliae measures**

Gaps or Problems

- Integration complexity between data sources and processing layers
- Monitoring and debugging in serverless environments is limited
- Potential vendor lock-in with specific cloud services
- Data freshness challenges in event-driven pipelines

Goals

- Scalability
- Cost Efficiency
- Real-Time Analytics

Conclusion

- The project demonstrates effective use of serverless architecture.
- Real-time processing enables actionable financial insights instantly.
- Cost efficiency is achieved with minimal infrastructure overhead.
- The system supports scalability and high availability.
- Future improvements can focus on deeper automation and analytics.

Thank you!