

Quickly implement a Low-Code Data Lake for No-Code Exploratory Data Analysis

Businesses are generating data at an explosive rate, but have barely scratched the surface on how to tap into this valuable resource. IDC estimates that 175 zettabytes of data will be created, consumed, and stored annually by 2025, making effective and efficient data storage more critical than ever before. As businesses pivot towards being data-driven, business intelligence, artificial intelligence, and machine learning become competitive advantages. Using the suite of data storage and analytics options on AWS, proactive businesses are realizing that value and gaining advantage today.

Caylent's data engineering experts will work with you to build a foundational data lake and enable your teams on how to use no-code solutions for wrangling and exploring your data. Our Serverless Data Lake Catalyst will shorten your time to market from months to weeks and set you up to easily grow your data lake by scaling ingest sources through low-code solutions.

Key Activities**Discovery & Planning**

Through a discovery workshop, we'll discuss your data lake foundation and data exploration needs, then capture your requirements on data import and access patterns

Design & Implementation

Based on your requirements, we'll design the baseline infrastructure, data pipelines, and security model, build a serverless data lake in your production account and import up to 3 single schema data sources

Enablement

We'll then conduct an exploratory data analysis (EDA) lab so your team can gain hands-on experience using AWS Glue DataBrew and Lake Formation for data exploration, cataloging and governance

Engagement Details**Highlights**

- Take advantage of AWS serverless services such as AWS Lake Formation, Amazon S3, AWS Glue, and AWS Glue workflows to greatly reduce time to market through a low code data lake
- Seamlessly run no-code data analysis leveraging Amazon Athena and AWS Glue DataBrew
- Easily scale your data sources and pipelines through prebuilt IaC automation

Deliverables

- Data landscape educational workshop and requirements documentation
- Architecture diagram of data lake and data workflow designs
- MVP Serverless Data Lake
- EDA enablement lab on data wrangling and data stewardship
- All diagrams and documentation developed during the engagement including source code, scripts, templates, and other technical artifacts