LUMEN**DATA**
Modernize Data, Maximize Insights

SQL Migration Quickstart for Databricks

Our Services

01 **Cloud Modernization**
Transformation of legacy systems into cloud-based infrastructure

02 **Data Engineering and Analytics**
Extracting, analyzing, and acting on diverse data sources with advanced data tools.

03 **Data Strategy and Business Value Assessment**
Aligning data initiatives for optimal business outcomes.

QuickStart Overview

Timeline: X Days

- **INTRODUCE** the data team to Databricks analytics.

- **SET UP** Databricks and data services in the non-productive environment.

- **EXECUTE** 1 SQL Workload Migration.

- **LEVERAGE** SQL Analytics and BI tools.

QuickStart: Outcomes

Day 1   Lakehouse 101

- Data Acquisition
- Simple Data Transformations
- Organizing Data
- Security
- BI, Reporting
- Optimizing Costs & Management

Day 2-3   Up & Running

- Implementation → Powered by Infra-as-Code
- Security Config → LumenData Security Rapid Config
- Lakehouse Optimization → LumenData Lakehouse Monitor

databricks

SQL Warehouse	Data Engineering	Data Streaming	Data Sci & ML

Unity Catalog
Fine-grained governance for Data & AI

Delta Lakehouse
Data reliability & performance

Cloud Data Lake

Day 4-8   Data Pipelines

- Identify Data Sources
- Historical & Current Data
- Data Transformations
- Data Quality
- Data Set Creation, Scheduled
- Tables in the Lakehouse, Ready!

Day 8-10   BI & Analysis, Optimize, & Roadmap

- Power BI or Tableau
- DBSQL for Ad hoc Analysis
- Dashboards & Reports
- Utilization Management
- Roadmap the Future

QuickStart: Lakehouse Components

databricks

Streamed or Scheduled Ingestion	

RAW
- Data as Sources
- Retained
- Stream | Batch

CLEAN
- Enriched/ Enhanced
- Validation/ Quality
- ML Ready Data

CURATED
- DW & Modeled
- Secured
- BI/ Analytics Ready

Data Validation & Observability

Continuous Processing

Automatic Recovery

Security & Compliance

Automated Deployments

Reverse ETL & Sharing

BI & Analytics

AI/ML

Event & Alert

ID Access Mgmt	Key Management	Provisioning	Operational Readiness	Lakehouse Monitor

LUMENDATA

Deployment Journey – One Step at a Time

1

AWS or Azure Readiness
- Account with create-rights
- Terraform installed

QuickStart Data Sources
- Cloud accessible srcs. identified
- Read-only account

2

Scripted Build
- Resource / Admin Groups
- Storage, Databricks Workspace, Clusters
- Unity Catalog / Metastore/ Permissions
- LumenData Lakehouse Optimizer

Data Ingestion Pipelines
- 1 SQL workload migrated

Workflows Active
- Enable Job/Workflows, Scheduled

3

Databricks is LIVE!
- SQL Analysis Workshop
- Power BI / Tableau
- Engineering & Workflow Demos

Lakehouse Optimized
- Monitor jobs
- Understand costs
- Identify orphaned workloads

QuickStart: Deliverables

Build a net-new use case and validate cost v/s performance and usability of Databricks platform.

- TCO and performance projection report
- Established Databricks Lakehouse environment
- Data ingestion pipeline
- Platform utilization monitoring app
- Working end-to-end use case/business process deployed to non-production environment
- Results report and demonstration

Cost & Timeline: $X, X weeks
(tool subscriptions complimentary)

About LumenData

LumenData is a leading provider of Enterprise Data Management, Cloud & Analytics solutions. We help businesses navigate their data visualization and analytics anxieties and enable them to accelerate their innovation journeys. Founded in 2008, with locations in multiple countries, LumenData is privileged to serve over 100 leading companies. LumenData is SOC2 certified and has instituted extensive controls to protect client data, including adherence to GDPR and CCPA regulations.

LUMENDATA

Get in touch with us:
info@lumendata.com

Let us know what you need:
lumendata.com/contact-us