

Thank you for your interest in SQL School's Databricks Training Program.
I'm **Mr. Sai Phanindra**, your trainer and mentor for this course. With 20+ years of experience in Database Technologies, Data Engineering, Business Intelligence, and Data Warehousing, I bring extensive real-world expertise into every training session.

This course is designed to be 100% practical and job-oriented, with step-by-step hands-on learning, real-time Databricks projects, industry use cases, best practices, and implementation scenarios to help you build strong, industry-ready Databricks skills.
Let's connect: linkedin.com/in/saiphanindra/

Databricks Data Engineer Training – 100% Practical & Project-Based

This Hands-On training program is divided into **FOUR** Comprehensive Modules:
Module 1: SQL Server (MSSQL) & T-SQL
Module 2: Databricks Data Engineering – Python, Spark & PySpark
Module 3: Real-Time Project
Module 4: Databricks Data Engineer Associate Certification Guidance

Learning Outcomes

By the end of this course, you will be able to:

- Design enterprise-grade Lakehouse architectures
- Build scalable ETL pipelines using Databricks
- Implement Bronze, Silver, and Gold data layers
- Perform advanced data transformations using PySpark

- Secure data using Unity Catalog
- Deliver business-ready datasets for analytics
- Build executive dashboards in Power BI
- Gain hands-on experience through production-style Data Engineering projects



Detailed Course Content

Module 1: MSSQL & TSQL Queries, Tuning

Ch 1: SQL Database Job Roles

- ✓ Database Intro
- ✓ OLTP, DWH, OLAP
- ✓ DBMS Basics
- ✓ Data Engineer Job Roles

Ch 2: Database Intro & Installations

- ✓ SQL Server Installations
- ✓ Instance & Collations
- ✓ SSMS Tool Installation
- ✓ Connections, Authentications

Ch 3: SQL Basics V1 (Commands)

- ✓ SQL Basics (DDL, DML, etc..)
- ✓ Creating Databases, Tables
- ✓ Data Inserts (GUI, SQL)
- ✓ Basic SELECT Queries

Ch 4: SQL Basics V2 (Commands, Operators)

- ✓ DDL: Create, Alter, Drop
- ✓ DML: Insert, Update, Delete
- ✓ DQL: Select, Fetch
- ✓ SQL Operators

Ch 5: Data Types & Variables

- ✓ Integer, Character Data Types
- ✓ Decimal & Boolean Data Types
- ✓ Date and Time Data Types
- ✓ SQL_Variant Type

Ch 6: Data Imports

- ✓ Data Imports with Excel
- ✓ Data Types Detection
- ✓ OLE-DB Connections
- ✓ ORDER BY, TOP, OFFSET

Ch 7: Schemas & Batches

- ✓ Schemas, Table Grouping
- ✓ Real-world Banking Database
- ✓ 2 Part, 3 Part, 4 Part Naming
- ✓ Batch & "Go" Command

Ch 8: Constraints, Keys & RDBMS

- ✓ Null, Not Null Constraints
- ✓ Unique Key & Check
- ✓ Primary Key & Default
- ✓ Foreign Keys, DB Diagrams

Ch 9: Normal Forms & RDBMS

- ✓ Normal Forms: 1 NF, 2 NF
- ✓ 3 NF, BCNF and 4 NF
- ✓ 1:1, 1:M, M:1 Cardinality
- ✓ Self Referencing Keys, ER Models

Ch 10: Joins & Queries

- ✓ Joins & TSQL Queries
- ✓ Inner Joins, Matching Data
- ✓ Outer Joins: Left, Right, Full
- ✓ Self Joins, Conditions

Ch 11: Sub Queries

- ✓ Sub Queries
- ✓ Nested Queries
- ✓ Correlated Sub Queries
- ✓ Union, Union All

Ch 12: Group By Queries

- ✓ Group By, Having
- ✓ Cube( ) and Rollup( )
- ✓ Grouping( )
- ✓ IsNull(), Coalesce()

Ch 13: Joins with Group By, Sub Queries

- ✓ Join, Group By, Sub Queries
- ✓ IIF(), CASE Statement
- ✓ EXISTS, NOT EXISTS
- ✓ Query Execution Order

Ch 14: Views & RLS

- ✓ Views: Realtime Usage
- ✓ DML, SELECT with Views
- ✓ WITH CHECK OPTION
- ✓ RLS, Important System Views

Ch 15: Functions & Queries

- ✓ Scalar, Table Value Functions
- ✓ Variables & Parameters
- ✓ Date Time, Conversion Functions
- ✓ String, Aggregated Functions

Ch 16: Advanced SQL, Window Functions

- ✓ Window Functions (Rank)
- ✓ Row_Number, DenseRank
- ✓ Partition By, Lead & Lag
- ✓ Pivot, UnPivot, Running Totals

Ch 17: Stored Procedures - 1

- ✓ Stored Procedures & Parameters
- ✓ SELECT & Conditional DML
- ✓ IF Statement, SP Recompilation
- ✓ System Stored Procedures

Ch 18: Stored Procedures - 2

- ✓ Merge Statement (Upsert)
- ✓ Merge with OLTP & DWH
- ✓ Matched and Not Matched
- ✓ Merge & SP Recompilations

Ch 19: Triggers & Automations

- ✓ DDL & DML Triggers
- ✓ For / After Triggers
- ✓ Instead Of Triggers
- ✓ Disabling DMLs Triggers

Ch 20: Transactions & ACID

- ✓ Auto Commit & Explicit
- ✓ Commit, Rollback, ReadPast
- ✓ Checkpoint, Query Blocking
- ✓ TRY...CATCH, Error Handling

Ch 21: SQL Query Performance & Indexing

- ✓ Clustered, NonClustered Indexes
- ✓ Execution Plans, Tuning Queries
- ✓ SARGable Queries
- ✓ Query Optimization Examples

Ch 22: CTEs & Tuning

- ✓ Common Table Expression
- ✓ CTEs for Retrieval, DML
- ✓ Data Cleansing Techniques
- ✓ Duplicate Detection/Removal

Ch 23: Temp Tables & Query Techniques

- ✓ Local & Global Temp Tables
- ✓ SELECT..INTO Statement
- ✓ Cursor Basics – When to Use & Avoid
- ✓ NULLIF, TSQL Queries

Ch 24: Bonus / Advanced Module: SQL Server Internals

- ✓ Database Engine Components
- ✓ Parser, Compiler & Optimizer
- ✓ Checkpoint, SQL OS
- ✓ Memory Manager & IO Managers

Module 2: Databricks (Python, Spark, PySpark, Unity Catalog)

Ch 1: Databricks Introduction

- ✓ Cloud ETL, DWH
- ✓ Databricks Concepts
- ✓ ETL, Lakehouse & Big Data
- ✓ Databricks Workspace UI

Ch 2: Databricks Architecture

- ✓ Two Plane Architecture
- ✓ Databricks Runtime (DBR)
- ✓ Spark Compute Concepts
- ✓ Databricks APIs

Ch 3: Unity Catalog

- ✓ Unity Catalog Concepts
- ✓ Managed & External
- ✓ Lineage, Data Governance

- ✓ Genie AI with Unity Catalog

Ch 4: Databricks Operations (GUI)

- ✓ Catalog & Schema Creations
- ✓ Volumes, Files & Tables
- ✓ Catalog Explorer
- ✓ Need for ETL

Ch 5: Spark SQL

- ✓ Spark SQL Notebooks
- ✓ Catalog, Schemas
- ✓ Tables, DDL Operations
- ✓ Spark Data Types & Clones

Ch 6: Spark Queries

- ✓ Data Inserts, Sequences
- ✓ Spark Joins
- ✓ Aggregations
- ✓ Data Analytics

Ch 7: Spark SQL: Functions

- ✓ Aggregate Functions
- ✓ Math, Sort Functions
- ✓ String, DateTime Functions
- ✓ Window & Null Functions

Ch 8: Spark SQL: Time Travel

- ✓ Time Travel Concepts
- ✓ HISTORY, EXTENDED
- ✓ Versions & Time Stamp
- ✓ Time Travel Settings

Ch 9: Python Introduction, Print

- ✓ Python Intro & Versions
- ✓ Python Implementations
- ✓ Python Print(), display()
- ✓ Single, Multiline Statements

Ch 10: Python Variables

- ✓ Variable Declarations
- ✓ Variable Values
- ✓ Multi Variable Values
- ✓ Type Casting, Collections

Ch 11: Python Operators

- ✓ Arithmetic, String Operators
- ✓ Logical, Assignment Operators
- ✓ Comparison, Format Operators

- ✓ **Indexing, Bitwise Operators**

Ch 12: Python Data Types

- ✓ **Integer, Float Data Types**
- ✓ **String, List Data Types**
- ✓ **Dictionary Data Type**
- ✓ **Tuple Data Type**

Ch 13: Python Control Statements

- ✓ **If ... Else Statement**
- ✓ **ELIF & ELSE IF Statements**
- ✓ **OR, AND Concepts**
- ✓ **FOR and WHILE Loops**

Ch 14: Python Modules & DataFrames

- ✓ **Pandas, NumPy**
- ✓ **DataFrame Concepts**
- ✓ **Handling Nulls**
- ✓ **Data Cleansing Concepts**

Ch 15: Python DataFrames Transformations

- ✓ **Concat(), Indexing**
- ✓ **Merge()**
- ✓ **Join KIND, ON Clauses**
- ✓ **Merging 3 Dataframes**

Ch 16: PySpark Concepts

- ✓ **Python Vs PySpark**
- ✓ **PySpark DataFrames**
- ✓ **select(), filter()**
- ✓ **withColumn()**

Ch 17: Medallion Architecture - 1

- ✓ **File Data Sources**
- ✓ **Bronze, Silver and Gold**
- ✓ **Temp Views, Aggregations**
- ✓ **Spark Tables (Parquet)**

Ch 18: Medallion Architecture - 2

- ✓ **Azure SQL DB Connections**
- ✓ **Joins & Aggregations**
- ✓ **Gold Data Consumption**
- ✓ **Bronze → Silver → Gold → BI/Analytics**

Ch 19: Databricks Delta Lake

- ✓ **Schema Evolution**
- ✓ **Delta Table API**
- ✓ **MERGE / Upsert Operations**

- ✓ Delta Transaction Log

Ch 20: PySpark: Widgets

- ✓ Parameters, Text Widgets
- ✓ dbutils() & get() Functions
- ✓ Widget Prompts, Edits
- ✓ UI & JSON for Widgets

Ch 21: Lake Flow Jobs

- ✓ Workflows & CRON
- ✓ Job Compute, Running Tasks
- ✓ Python Script Tasks, Dependencies
- ✓ Concurrency, Branch Control

Ch 22: PySpark: Auto Loader

- ✓ Cloud Files & Checkpoints
- ✓ Data Stream, Schemas
- ✓ Schema Evaluation Modes
- ✓ File Ingestions, Rescued Column

Ch 23: LakeFlow Declarative Pipelines

- ✓ SDP: Spark Declarative Pipelines
- ✓ Delta LIVE Tables, Data Streams
- ✓ Bronze, Silver, Gold Data
- ✓ Data Quality, Materialized Views

Ch 24: Databricks Optimizations

- ✓ Lazy Evaluation, Data Skew
- ✓ Data Shuffling, Partitions
- ✓ Broadcast Joins, Vacuum
- ✓ Liquid Clusters, Optimize, Z-Order

Ch 25: Databricks Security, AI

- ✓ ACL: Access Control List
- ✓ Workspace & Catalog Security
- ✓ Schema, Volume Security
- ✓ Table & Column Security

Ch 26: Databricks with Power BI

- ✓ Access Keys (Tokens)
- ✓ Server Host, HTTP Path
- ✓ Spark Connectors
- ✓ Databricks Connectors

Ch 27: Genie AI

- ✓ Unity Catalog & Genie AI
- ✓ AI Components in Databricks
- ✓ Debugging Controls

- ✓ **Using AI for Notebook Design**

Ch 28: GitHub & Asset Bundle

- ✓ **GitHub & GIT**
- ✓ **Branches, Asset Bundle**
- ✓ **Integrating GIT**
- ✓ **Process Integration**

Ch 29: Apache Airflow

- ✓ **Apache Airflow Concepts**
- ✓ **Python Scripts, Airflow Jobs**
- ✓ **DatabricksRunNowOperator**
- ✓ **DatabricksSubmitRunOperator**

Module 3: Real-Time Projects

Project 1: HealthCare Domain

Project Requirement:

Solve 20+ Real-World Healthcare Business Requirements.

Project Workflow:

Database Design → SQL Development → Data Analysis → Performance Optimization

Project Operational Flow:

Healthcare DB → Patients → Doctors → Appointments → Treatments → Billing → Insurance → SQL Analysis → Stored Procedures → Performance Tuning → Business Reports

Business Requirements:

Monthly Hospital Revenue, Top Doctors By Patients Treated, Repeat Patients, Department Performance, Average Treatment Cost, Unpaid Bills, Insurance Claim Analysis, Patient Trends, Month-Over-Month Revenue And Top Procedures…

Project 2: E-Commerce Domain

Project Title

Retail Sales Analytics Platform using Databricks (Medallion Architecture)

Project Overview

Students will build a production-style data platform using Lakehouse Architecture (Bronze, Silver, and Gold), implementing modern ETL practices, data quality validation, incremental processing, orchestration, optimization, and reporting.

Business Scenario

A multinational retail company receives daily data from multiple operational systems. The company wants to build a centralized analytics platform to:

- Consolidate data from different business domains
- Improve data quality and consistency
- Track sales performance across regions
- Analyse customer purchasing behaviour
- Monitor inventory levels
- Generate business KPIs for decision-makers
- Deliver Power BI dashboards with near real-time insights

Source Systems

Students will work with multiple datasets including:

- Customer Master
- Product Master
- Orders
- Inventory
- Region Master

Technologies Covered

- Apache Spark
- PySpark
- Spark SQL
- Delta Lake
- Unity Catalog
- Databricks Workflows
- Git Integration

Resume Highlights

After completing this project, you can confidently showcase experience in:

- PySpark & Spark SQL
- Delta Lake
- Medallion Architecture
- ETL Pipeline Development
- Incremental Data Processing
- Data Quality & Validation
- Unity Catalog
- Workflow Automation
- Performance Optimization

Module 4: Databricks Data Engineer Certification Guidance

- ✓ **Exam Q & A, Scenarios**
- ✓ **Certification Exam Objectives**
- ✓ **Topic-Wise Revision**
- ✓ **Scenario-Based Questions**
- ✓ **Practice Assessments**
- ✓ **Mock Exams**
- ✓ **Databricks Interview Questions**

- ✓ **Real-Time Scenario Discussions**
- ✓ **Resume Project Explanation**
- ✓ **Technical Mock Interview**

Why Choose This Training?

🚀 Industry-Oriented Curriculum
💼 End-to-End Real-Time Projects & Solutions
🏗️ Medallion Architecture Implementation
📊 100% Hands-on Practical Sessions
⚡ Production-Ready ETL Pipelines
🤝 PySpark & Spark SQL from Basics to Advanced
🎯 Interview Preparation

At SQL School, our focus goes beyond completing a syllabus. You will learn through hands-on labs, production-style scenarios, end-to-end projects, assignments, interview preparation, and certification guidance designed to build practical Databricks Data Engineering skills.

BUILD REAL SKILLS. WORK ON REAL PROJECTS. BECOME JOB READY.



🚀 **Ready to Become an Expert?**
✅ **Join Our Next Batch. Reach us today!**
📞 **Call/WhatsApp: +91 9951440801**
🌐 **Course Details: https://sqlschool.com/databricks-data-engineer-training/**