

A promotional banner for Spark SQL training. The left side has a dark blue background with yellow and white text. It says 'LAUNCHING A NEW COURSE' with a megaphone icon, followed by 'Spark SQL' in large letters, and 'From SQL School Training Institute.' Below this is a 'TRAINING HIGHLIGHTS' section with two columns of icons and text. The right side features a photo of a man, Mr. Sai Phanindra Tholeti, sitting at a desk with a laptop. The SQL School logo is in the top right corner.

LAUNCHING A NEW COURSE

Spark SQL

From
SQL School Training Institute.

TRAINING HIGHLIGHTS

- Introduction to Spark SQL Architecture
- Working with DataFrames and Datasets
- SQL Queries on Big Data
- Joins, Aggregations & Advanced SQL
- Performance Optimization & Best Practices
- Spark SQL Functions & UDFs
- Integration with Hive, Parquet & JSON
- Real-time Use Cases & Hands-on Labs

Ph: +91 9030040801
(Whatsapp, Call)

Whatsapp or Call
for Enquiries

Trainer:
Mr. Sai Phanindra Tholeti

SQL School
— Training Institute —
Learn | Implement | Succeed

What is Database?

Database is a platform used to store & manage any type of data. Any amount of data. In an organized format.

What is SQL?

SQL (Structured Query Language) is a platform to connect, operate with database. We can store data, secure, report data and analyze data using SQL.

What is SparkSQL?

Spark SQL is a specialized module within the Apache Spark framework designed for processing structured and semi-structured data. Using Spark SQL, we can do Big Data Storage, Big Data Analytics and ETL Operations with Warehouse Implementations.

Where Spark SQL is Used?

Spark is used very frequently in:

- ✓ Big Data Analytics
- ✓ Data Warehousing
- ✓ ETL Pipelines
- ✓ Real-time Data Processing

Where is Spark SQL implemented?

Spark is actively implemented in Cloud Platforms like:

- ✓ Azure Databricks
- ✓ AWS EMR
- ✓ Google Dataproc
- ✓ Microsoft Fabric

Who can join this course?

Anyone. There are NO Pre-requisites for this course. We start the classes right from Basics

What about the Lab? What are the System Requirements?

Any OS is ok; 8 GB RAM; Any Processor. We shall guide you for Installation, Step by Step.

Spark SQL : Detailed Course Content

This detailed content with 100% Hands-On, Daily Assignments make you 100% Job ready !

Ch 1: Big Data & Spark SQL Introduction

- ✓ What is Big Data?
- ✓ Limitations of Traditional Databases
- ✓ Hadoop vs Spark
- ✓ Spark SQL Architecture
- ✓ Industry Use Cases

Ch 2: Apache Spark Architecture

- ✓ Driver Program
- ✓ Executors
- ✓ Cluster Manager
- ✓ DAG & Lazy Evaluation
- ✓ Spark Session

Ch 3: Environment Setup

- ✓ Spark Installation
- ✓ Databricks Setup
- ✓ Cluster Creation
- ✓ Spark UI
- ✓ Notebook Usage

Ch 3: Spark SQL Basics - 1

- ✓ Catalog Concept
- ✓ Catalog Creation
- ✓ Spark Databases
- ✓ DB Operations
- ✓ Table Creations
- ✓ Data Insertions

Ch 4: Spark SQL Basics - 2

- ✓ DDL Operations
- ✓ ALTER Operations
- ✓ Table ALTER
- ✓ Database ALTER
- ✓ Table DROP
- ✓ Database DROP

- ✓ Other DDLs

Ch 5: Spark SQL Basics

- ✓ Temporary Views
- ✓ SELECT Statements
- ✓ WHERE Clause
- ✓ DISTINCT
- ✓ ORDER BY
- ✓ LIMIT

Ch 6: Filtering & Conditional Logic

- ✓ AND/OR
- ✓ BETWEEN
- ✓ IN Operator
- ✓ CASE WHEN
- ✓ NULL Handling

Ch 7: String Functions

- ✓ CONCAT
- ✓ SUBSTRING
- ✓ UPPER/LOWER
- ✓ TRIM
- ✓ REGEXP Functions

Ch 8: Numeric Functions

- ✓ ROUND
- ✓ CEIL/FLOOR
- ✓ ABS
- ✓ POWER
- ✓ SQRT

Ch 9: Date & Time Functions

- ✓ CURRENT_DATE
- ✓ DATE_ADD
- ✓ DATEDIFF
- ✓ MONTH/YEAR
- ✓ Timestamp Formatting

Ch 10: Aggregations

- ✓ COUNT
- ✓ SUM
- ✓ AVG
- ✓ GROUP BY
- ✓ HAVING Clause

Ch 11: Joins – Part 1

- ✓ INNER JOIN
- ✓ LEFT JOIN

- ✓ RIGHT JOIN
- ✓ FULL OUTER JOIN

Ch 12: Joins – Part 2

- ✓ Broadcast Join
- ✓ Self Join
- ✓ Join Optimization
- ✓ Handling Skew Data

Ch 13: Set Operators

- ✓ UNION
- ✓ INTERSECT
- ✓ EXCEPT
- ✓ Deduplication Concepts

Ch 14: Window Functions – Part 1

- ✓ ROW_NUMBER
- ✓ RANK
- ✓ DENSE_RANK
- ✓ PARTITION BY
- ✓ LEAD
- ✓ LAG
- ✓ Running Totals
- ✓ Moving Average

Ch 15: Understanding DataFrames

- ✓ Schema & Structure
- ✓ Creating DataFrames
- ✓ Reading CSV/JSON/Parquet
- ✓ DataFrame Operations

Ch 16: Working with Nested Data

- ✓ Struct Type
- ✓ Array Type
- ✓ Map Type
- ✓ Explode Function

Ch 17: Handling Semi-Structured Data

- ✓ JSON Processing
- ✓ Schema Inference
- ✓ Nested File Queries

Ch 18: Parquet & Delta Lake

- ✓ Parquet Basics
- ✓ Delta Lake
- ✓ ACID Transactions

- ✓ Time Travel
- ✓ Managed Tables
- ✓ External Tables

Ch 19: Performance Optimization

- ✓ Caching
- ✓ Partitioning
- ✓ Repartition vs Coalesce
- ✓ Catalyst Optimizer
- ✓ Tungsten Engine
- ✓ Execution Plans

Ch 20: Spark SQL with Hive

- ✓ Hive Metastore
- ✓ Managed Tables
- ✓ External Tables

Ch 21: Spark SQL with Delta Tables

- ✓ MERGE
- ✓ UPSERTS
- ✓ SCD Concepts
- ✓ Deletes & Updates

Ch 22: Error Handling & Debugging

- ✓ Common Errors
- ✓ Log Analysis
- ✓ Performance Troubleshooting

Ch 23: Real-Time Industry Use Cases

- ✓ Banking
- ✓ Healthcare
- ✓ Telecom

Ch 24: Interview Preparation & Career Guidance

- ✓ Interview Questions
- ✓ Resume Preparation
- ✓ Career Roadmap
- ✓ Mock Interviews
- ✓ Bonus Features Included

Tools Covered

- ✓ Apache Spark
- ✓ Databricks
- ✓ Delta Lake
- ✓ Hive

All sessions are practical, step by step.

www.sqlschool.com

LAUNCHING A NEW COURSE —

Spark SQL

From
SQL School Training Institute.

TRAINING HIGHLIGHTS

-  Introduction to Spark SQL Architecture
-  Working with DataFrames and Datasets
-  SQL Queries on Big Data
-  Joins, Aggregations & Advanced SQL
-  Performance Optimization & Best Practices
-  Spark SQL Functions & UDFs
-  Integration with Hive, Parquet & JSON
-  Real-time Use Cases & Hands-on Labs



Ph: +91 9030040801
(Whatsapp, Call)



Whatsapp or Call
for Enquiries



SQL School

— Training Institute —

Learn | Implement | Succeed

Trainer:

**Mr. Sai
Phanindra Tholeti**

