

Azure Data Engineering Training

From SQL Fundamentals to Cloud Data Engineering
Build job-ready data engineering skills with SQL School's instructor-led Azure Data Engineering Training Program. Learn SQL Server TSQL, Azure Data Factory, Azure Data Lake Storage, Azure Synapse Analytics, Azure Databricks, Spark, Python, PySpark, and Microsoft Fabric through practical exercises and real-world projects.

Azure Data Engineer Training Course Content

This hands-on training program is divided into **SIX** comprehensive modules:
Module 1: SQL Server (MSSQL) & TSQL Queries, Tuning
Module 2: Azure Funda, Data Factory & Synapse
Module 3: Azure Databricks (Python, Spark, PySpark, Unity Catalog)
Module 4: Microsoft Fabric Migrations
Module 5: Realtime Projects, Certifications & Resume

Training Mode: LIVE Online Instructor-Led Training
Duration: 10 Weeks (core training; advanced practice may require additional learning hours)
Projects: Healthcare, Online Travel and Retail Data Engineering

Trainer: Mr. Sai Phanindra Tholeti (20+ Yrs of expertise)
Trainer Profile: https://www.linkedin.com/in/saiphanindra/
Book Your FREE Demo Today. Call us on +91 9951 440801, +91 9666 640801

What You Will Learn

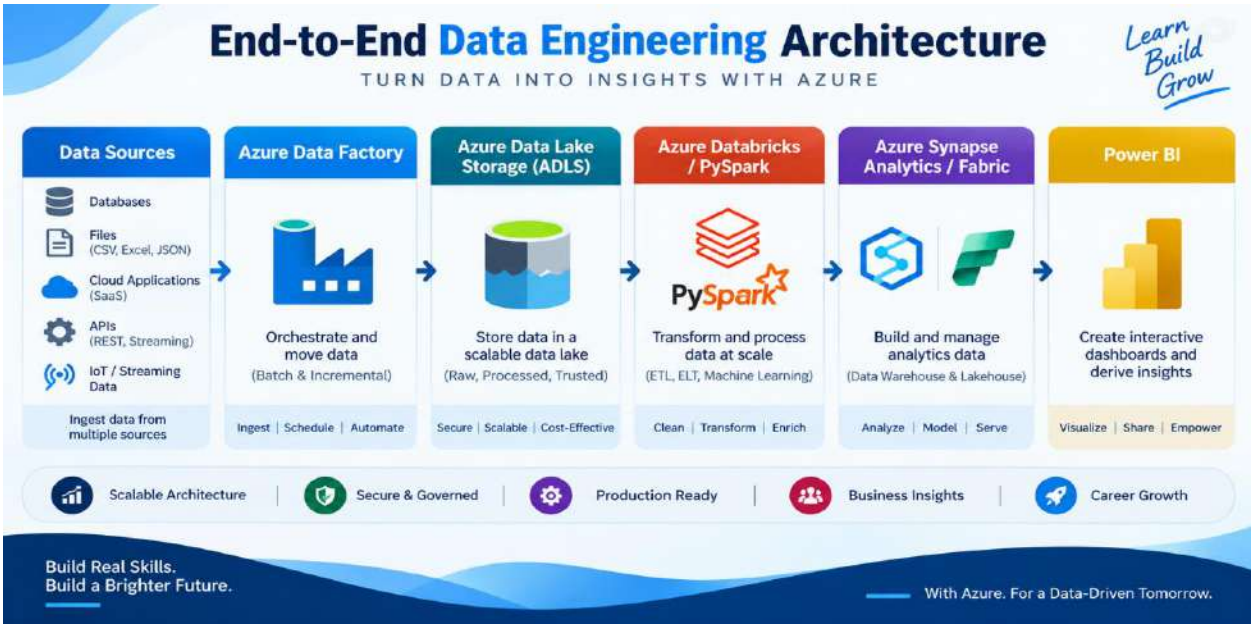
- Build SQL-based data solutions and optimize queries.
- Develop ETL and ELT pipelines using Azure Data Factory.
- Manage scalable cloud data storage using ADLS Gen2.
- Transform and process data using Azure Databricks, Spark and PySpark.
- Implement Delta Lake and Bronze, Silver and Gold architectures.
- Explore Azure Synapse Analytics and Microsoft Fabric.
- Implement pipeline security, monitoring, orchestration and CI/CD.
- Develop end-to-end projects and prepare for relevant certifications.

Prerequisites

None. The course starts with fundamentals and progresses to advanced Data Engineer concepts!

Understand the Complete Data Engineering Journey

The following architecture shows the **end-to-end data journey you will learn and implement during this program: Sources → ADF → ADLS → Databricks/ Python / PySpark → Synapse → Power BI**



🔥 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: Azure Concepts, ADF & Synapse

Ch 1: Azure Fundamentals

- ✓ Cloud Intro, Implementations
- ✓ Azure Account, Subscription
- ✓ Azure ETL & DWH Resources
- ✓ Azure Storage, IoT Resources

Ch 2: Azure Storage & ADLS

- ✓ Azure Storage Account
- ✓ Azure Data Lake Storage
- ✓ BLOB Containers
- ✓ Folders & Files

Ch 3: Azure Security Concepts

- ✓ Microsoft Entra, RBAC
- ✓ Users & Groups
- ✓ Storage Security, IAM, ACL
- ✓ Access Keys, SAS Keys

Ch 4: Azure Deployments, Azure SQL

- ✓ Azure SQL Server Deployment
- ✓ Azure SQL DB Deployment
- ✓ Azure SQL Database (OLTP)
- ✓ Firewall Rules, SSMS Connections

Ch 5: Azure SQL DB Migrations

- ✓ On-Premise SQL DB, bacpac
- ✓ Azure SQL DB Migration
- ✓ SQL Server Imports from BLOB
- ✓ Migration Verifications

Ch 6: Azure Synapse Analytics (DWH)

- ✓ Synapse Pool Architecture
- ✓ Control Node, Compute Node
- ✓ DMS (Data Movement Service)
- ✓ Partitions, Distributions & MAXDOP

Ch 7: Azure Data Factory – Pipelines

- ✓ ADF Intro & Pipeline Artifacts
- ✓ IR, Linked Services, Datasets
- ✓ Copy Data Activity
- ✓ Azure SQL DB → Synapse Loads

Ch 8: ADF: Azure Blob Data Loads

- ✓ ADF Author & Pipeline Runs
- ✓ Azure BLOB Storage Access
- ✓ Pipeline Staging & Logging
- ✓ Validations & Schedules

Ch 9: Table Incremental Loads

- ✓ SCD & Incremental Loads
- ✓ Business Keys & Data Delta
- ✓ Copy Method: Upsert
- ✓ Initial & Incremental Loads

Ch 10: On-Premises Loads, Optimizations

- ✓ Self Hosted IR Setup
- ✓ IR Keys, Security
- ✓ On-Premise Data Loads
- ✓ ADF Optimizations

Ch 11: ADF Data Flow - 1

- ✓ Data Flow Activity
- ✓ Debug Cluster, Data Preview
- ✓ Data Cleansing, Conditional Split
- ✓ Select, Sort, Union, Loops

Ch 12: ADF Data Flow - 2

- ✓ Join Transform, Broadcast
- ✓ Surrogate Keys, Derived Cols
- ✓ ETL Loads Dates, Sink Options
- ✓ Aggregated Data Loads

Ch 13: ADF Data Flow - 3

- ✓ Pivot, Data Duplication
- ✓ Group By & Pivot Keys
- ✓ Column Name Pattern
- ✓ Tuning Data Flow, Spark

Ch 14: ADF Data Flow - 4

- ✓ Get Metadata Activity
- ✓ IF Activity Expressions
- ✓ Lookup & Loops in ADF
- ✓ Data Flow Cluster

Ch 15: ADF Parameters, Security

- ✓ Linked Service Parameters
- ✓ Logins & Users, Permissions
- ✓ ADF Connection Parameters
- ✓ Dynamic Linked Services

Ch 16: SCD & ETL with Control Tables

- ✓ ADF Templates, Data Delta
- ✓ Control Tables, Watermarks
- ✓ Lookup, SP Activity For SCD
- ✓ Pipeline Parameters, SPs

Ch 17: ADF Metrics, Alerts

- ✓ ADF Insights & Metrics
- ✓ DTU, CPU & Memory Metrics
- ✓ Alerts and Notifications
- ✓ Action Groups, Tuning Options

Ch 18: ADF with DevOps & CI/CD

- ✓ ADF Git integration
- ✓ Dev / Test / Prod environments
- ✓ Parameterized deployments
- ✓ Production release workflow

Ch 19: Azure Stream Analytics

- ✓ Azure IoT Hubs
- ✓ Azure Stream Analytics
- ✓ LIVE Data Ingestions
- ✓ SAQL Queries, JSON

Ch 20: Azure Key Vaults

- ✓ Azure Key Vaults
- ✓ Azure Keys & Secrets
- ✓ Access Policies
- ✓ Secret Management

Module 3: Azure 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

Ch 30: Azure Databricks

- ✓ Azure Databricks Deployments
- ✓ Azure Databricks Regions
- ✓ Classic Deployments
- ✓ Access connector for Databricks

Ch 31: Azure Databricks with ADLS

- ✓ Creating IAM Users
- ✓ ADLS Security for Databricks
- ✓ Creating Service Principals
- ✓ Enterprise Applications
- ✓ Client ID, Secret ID
- ✓ External Tables (Delta)

Ch 32: Azure Databricks with ADF

- ✓ ADF Pipelines with Databricks
- ✓ Databricks Notebook Activities
- ✓ ADB Clusters inside ADF
- ✓ End to End Workflow

Module 4: Microsoft Fabric Migrations

Ch 1: Microsoft Fabric

- ✓ Microsoft Fabric Cloud
- ✓ OneLake Concepts
- ✓ Fabric Warehouse
- ✓ Fabric Lakehouse

Ch 2: Microsoft Fabric Warehouse

- ✓ Fabric Warehouse Creation
- ✓ Connections, Operations
- ✓ TSQL Concepts
- ✓ Azure DWH Vs Fabric Warehouse

Ch 3: Microsoft Fabric Lakehouse

- ✓ Fabric Lakehouse Creation
- ✓ Connections, Operations
- ✓ Queries, Notebooks
- ✓ ADLS Vs Fabric Lakehouse

Ch 4: Microsoft Fabric Migrations

- ✓ Azure Synapse to Warehouse Migrations
- ✓ ADF to Fabric Data Factory Migrations
- ✓ ADLS to Fabric OneLake Migrations
- ✓ End to End Workflow

Module 5: Realtime Projects for Data Engineering

Real-Time Project 1: HealthCare Domain

Project Requirement:

Solve 20+ Real-World Healthcare Business Requirements.

Project Workflow:

Database Design → SQL → 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…

Real-Time Project 2: Online Travel Data Engineering Platform

Project Objective:

Build an end-to-end Azure Data Engineering solution to process, transform, and analyze Online Travel business data from multiple sources.

Skills Gained:

- ✓ Data Ingestion & ETL Development
- ✓ Azure Data Factory Pipelines
- ✓ Data Orchestration (End to End)
- ✓ Data Lake Architecture, RBAC
- ✓ Real-Time Industry Experience
- ✓ Azure IoT, Stream Analytics
- ✓ Azure Databricks with Data Factory

Components For Project

- ✓ ADF Pipelines
- ✓ Databricks Notebooks
- ✓ Apache Spark Schemas
- ✓ Monitoring & Alerts
- ✓ CI/CD Deployment
- ✓ End to End Integrations
- ✓ IAM & Managed Identity

Real-Time Project 3: 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 ETL practices, data quality validation, incremental processing and orchestration.

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
- ✓ Generate business KPIs for decision-makers

Technologies Covered

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

Project Workflow

Sources → ADLS → Bronze → Silver → Gold → Databricks Workflows → Power BI

Certifications & Career

Exam 1: Databricks Data Engineer Associate

- ✓ Exam Details
- ✓ Exam Dumps
- ✓ Scenario-Based Questions
- ✓ Practice Assessments

Exam 2: Microsoft Databricks (DP-750)

- ✓ Exam Details
- ✓ Exam Dumps
- ✓ Scenario-Based Questions
- ✓ Practice Assessments

Resume Highlights

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

- ✓ PySpark & Spark SQL
- ✓ ETL Pipeline Development
- ✓ Data Quality & Validation
- ✓ Automation, Optimizations

Interview & Career Preparation

- End-to-End Real-Time Projects & Solutions
- 100% Hands-on Practical Sessions
- Interview Preparation
- Assignments & Practice Labs
- 100% Practical • No Unnecessary Theory
- Learn Like You're Working in a Real Company

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



🚀 **Ready to Upgrade Your Skills & Become Job-Ready?**
🎓 **Book Your FREE Demo Today!**
WhatsApp / Call: +91 9666 44 0801, +91 9666 64 0801
🌐 www.sqlschool.com

🚀 Learn Practical Skills. Build Confidence. Accelerate Your Career!