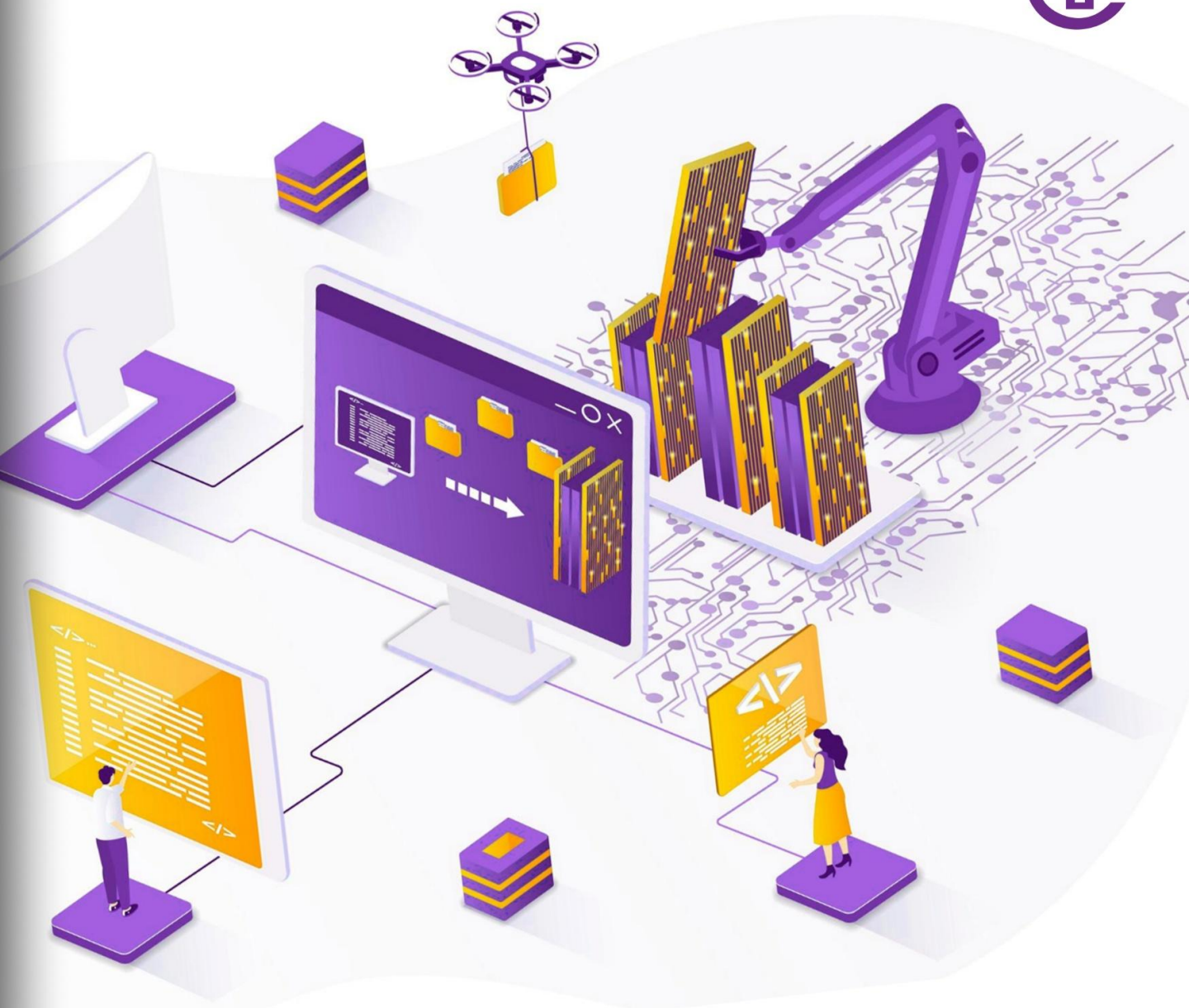




Data Engineering with Databricks

From data pipelines to Databricks certification

The Content

About Us

03

Why Data Engineering?

04

The Future of Data

05

Program Overview

06

Career Paths

07

Salary Growth

08

Certification

09

Meet Your Trainers

10

Curriculum Breakdown

11

Capstone Projects

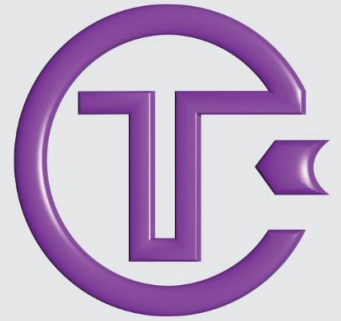
16

Career Outcomes & Industry relevance

17

About TeqCertify

TeqCertify reimagines EdTech by blending innovation with hands-on, AI-powered learning in Data Domain. We equip professionals with future-proof skills through real-world case studies and industry projects.



Empowering Organizations with **Ready-to-Deploy** Data Domain Talent.

- › Customized Training for Industry-Specific Needs
- › Real-Time Performance Reports
- › Flexible Training Solutions
- › 100% Placement (Assistance & Guarantee)

Why Data Engineering?

In the era of AI and decision intelligence, data is the new oil but without engineering, it's just crude. Data engineering is what makes it possible for companies to use data to make smart decisions every day.



Scalable Infrastructure

Learn how to build systems that manage large amounts of data smoothly and quickly.



AI-Ready Foundations

Set up data the right way so it's ready for AI tools, predictions, and automation.



High-Demand Skillset

Data engineers are needed in every industry it's a career path full of opportunities.



CEO, databricks



Ali Ghodsi

*In five or 10 years, to be the **CEO** in any of the industries, you'll need to have a Data-and-AI background.*

At the heart of this program lies the belief that modern data engineering isn't just about managing data — it's about engineering intelligence at scale.

Inspired by this vision, our curriculum empowers students to not only master tools like Spark, Delta Lake, and Databricks SQL, but to build intelligent systems that drive decision-making in the real world.

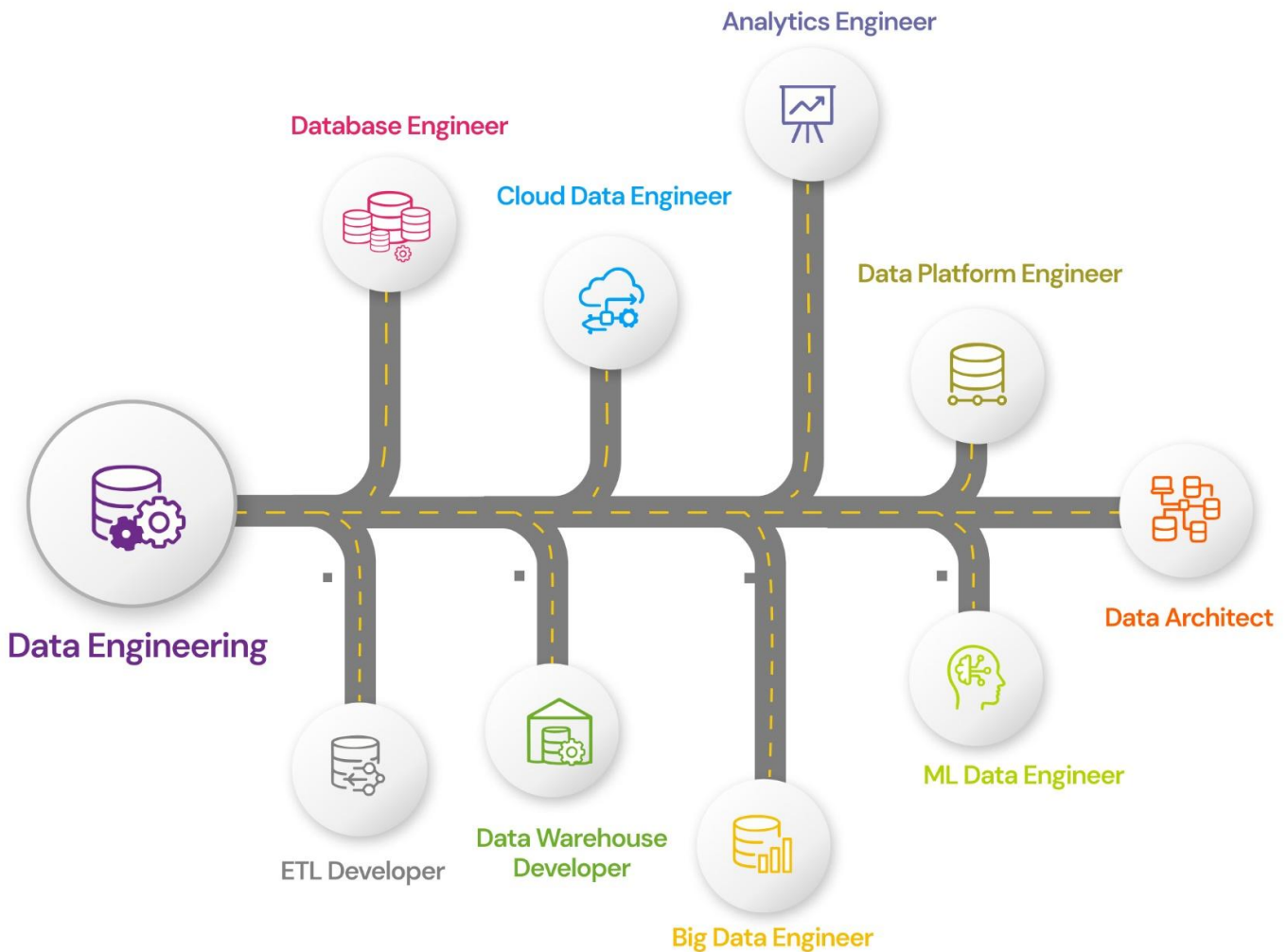
Program Overview

In today's world, data engineering goes beyond pipelines it powers AI and intelligent decision-making. This program combines practical training, real-world projects, and expert mentorship to help students gain certified skills that matter in the industry.



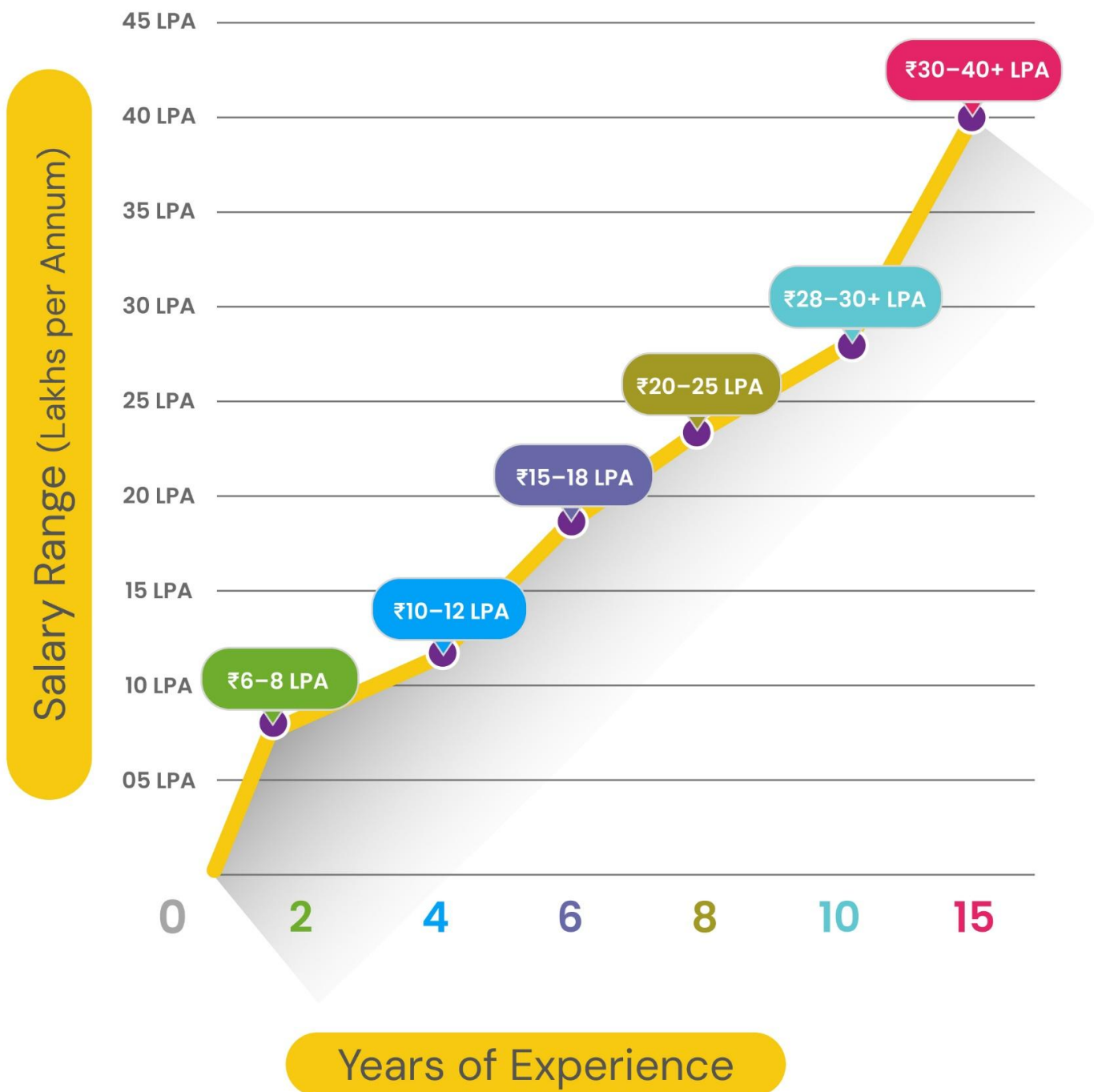
Data Engineering with Databricks is a hands-on certification program that transforms students and freshers into job-ready data engineers. It prepares learners for the **Databricks Data Engineer Associate certification** while teaching them how to build smart, scalable data systems

Career Paths

**Note:**

Power the pipelines behind modern data ecosystems. Learn to design scalable infrastructure, automate workflows, and optimize data delivery in real time. This is your gateway to high-demand backend and cloud engineering roles.

Salary Growth



Databricks Certification

The **Databricks Data Engineer Associate Certification** is an industry-recognized credential that proves your ability to build reliable, scalable data pipelines using Apache Spark, Delta Lake, and the Databricks Lakehouse Platform.



*Earning your certification isn't just proof of your skills
It's a testament to your mastery of Data Engineering.
Companies recognize it. Recruiters remember it.*

Top Mentors



Srinath Raja
Chief Data Officer

Co-founder and Chief Data Officer of GWC Data.ai, with extensive expertise in analytics and insights across banking, retail, and healthcare Industries. A dynamic leader from Tamil Nadu driving global innovation in cloud and business intelligence solutions and also a trusted partner for DOMO.



Kalpana Kudumula
Senior Data Engineer

Senior Boomi Developer | Data Engineer With her expertise spanning business consulting to data engineering, Kalpana transforms data challenges into innovative solutions. From Business Consultant at Multiplier IT Solutions to Business Analyst at Accenture, she now excels as a Senior Data Analyst and Boomi Developer, driving growth and mentoring in the data domain.



Denzina Eric
L&D Analyst

As a Learning & Development Analyst at TeqCertify, Denzina Eric brings 11+ years of experience, having empowered 1200+ aspiring leaders at Hope Foundation in a short span. With a solid background as a Business Analyst at Merstan Impex, she now leads talent creation, shaping the next generation of visionary leaders through transformative, forward-thinking programs.

Curriculum Overview

Our comprehensive 6-month, full-time Data Engineering Program equips freshers with hands-on skills through immersive, project-based learning.

The curriculum is aligned with the Databricks Data Engineer Associate certification, focusing on real-world tools like Apache Spark, Delta Lake, and MLflow to prepare students for industry-ready roles.



Foundations of Data Engineering

Develop strong core skills in SQL, Python, and cloud fundamentals to prepare for building robust data systems.

SQL for Data Engineering

- Advanced SQL queries (joins, subqueries, aggregations)
- Window functions and data transformations
- Project: Analyze sales and customer data using complex SQL queries
- Deliverable: SQL script and report with KPIs

Python Programming for Data Handling

- Python basics, control structures, functions
- Pandas and NumPy for data wrangling
- Project: Load and transform data using Python
- Deliverable: Cleaned dataset and script for preprocessing

Cloud Platforms & Data Modeling

- Intro to AWS, Azure, GCP
- OLTP vs OLAP, star/snowflake schema
- ETL concepts and Databricks workspace overview
- Project: Design a cloud-based data warehouse schema
- Deliverable: Schema diagram and data modeling rationale

Apache Spark Foundations

- Spark architecture, RDDs, DataFrames, Spark SQL
- User-defined functions (UDFs)
- Project: Transform data using PySpark
- Deliverable: Spark job script and transformation output

Core Data Engineering with Databricks

Dive into real-time data processing, Delta Lake, and scalable pipeline design using Spark and orchestration tools.

Delta Lake and Data Reliability

- ACID transactions on data lakes
- Schema enforcement and time travel
- Project: Build a Delta Lake pipeline with schema evolution
- Deliverable: Delta pipeline code and versioned data snapshots

Databricks SQL & Dashboarding

- Query structured data using Databricks SQL
- Building dashboards in Databricks
- Project: Create a live dashboard using Databricks SQL
- Deliverable: SQL queries and working dashboard

Batch & Streaming Pipelines

- Batch vs real-time data
- ETL pipeline design
- Project: Design and implement batch and streaming pipelines
- Deliverable: ETL pipeline flow diagram and working code

Workflow Orchestration

- Apache Airflow basics
- DAGs and scheduling in production pipelines
- Project: Build an Airflow DAG to automate data pipelines
- Deliverable: Working DAG and Airflow setup

Enterprise Engineering & Performance Optimization

Apply real-world engineering skills including performance tuning, governance, and scalable data architectures.

Performance Tuning in Spark

- Memory management, partitioning, caching
- Project: Optimize a slow-running Spark job
- Deliverable: Before-and-after performance comparison

MLflow for Experiment Tracking

- Track ML experiments, model versioning
- Project: Log and manage an ML pipeline using MLflow
- Deliverable: MLflow experiment dashboard

Data Governance & Quality

- Data validation, access controls, GDPR compliance
- Project: Implement data quality checks in a pipeline
- Deliverable: Data validation report and governance checklist

Lakehouse & Modern Architectures

- Data lake vs warehouse vs lakehouse
- Real-time stream processing with Kafka/Azure Event Hub
- Project: Build a mini lakehouse with real-time ingestion
- Deliverable: Architecture diagram and stream output

Capstone Projects and Career Readiness

Apply your skills in a simulated real-world project environment and get job-market ready.

Project Planning & Dataset Preparation

- Define business problem, scope, and KPIs
- Project: Choose domain (e.g., fintech, healthcare), plan project
- Deliverable: Project proposal with goals and dataset

Data Pipeline & Modeling

- Build full pipeline with transformations, storage
- Project: Implement end-to-end data pipeline using Spark/Delta
- Deliverable: Functional pipeline and documentation

Dashboard, Governance, & Reporting

- Create visuals, apply quality checks
- Project: Build dashboards and add data governance layers
- Deliverable: Presentation-ready project with audit trail

Final Project Demo & Career Prep

- Present project, interview readiness, resume prep
- Project: Present capstone to panel of mentors and peers
- Deliverable: Final presentation and job-ready portfolio

Sample Projects

Apply your knowledge to business-critical challenges inspired by real companies:



E-commerce Optimization

Built a scalable data pipeline to track customer behavior, leading to 22% improvement in targeted promotions and inventory planning.

Healthcare Data Lake

Developed a unified data lakehouse to manage patient records across hospitals, enabling faster diagnostics and predictive care models.



Streaming Analytics

Processed billions of user interactions daily with Delta Lake to recommend personalized content improving engagement by 18%.

Finance Fraud Detection

Designed real-time data streams using Apache Spark to detect transaction anomalies and reduce fraud risk by 30%.



Supply Chain Intelligence

Automated ETL pipelines to provide end-to-end supply chain visibility, reducing delays and boosting operational efficiency.

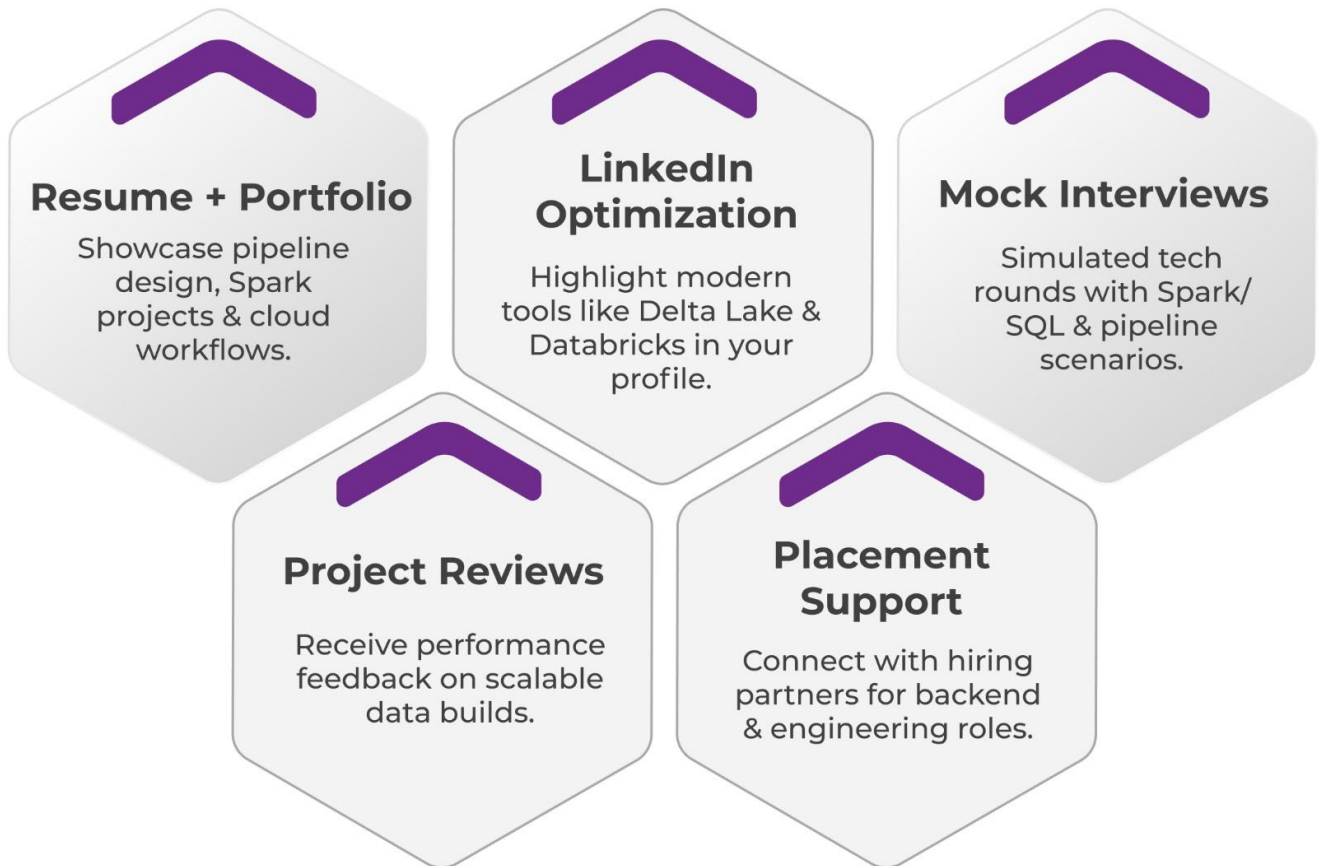
Toolstack

Build, scale, and manage production-grade data systems.



Careers

Launch your career in high-demand data roles with practical expertise.



Our Alumni have landed analyst roles in:



Fintech



Health-tech



Banking



SaaS



Ecommerce

Use Cases

Real-world data engineering simulations designed to build mastery in large-scale, high-performance data systems.



E-commerce: Clickstream Processing

Gain expertise in segmenting customers using data to drive higher ROI on targeted advertising and personalized marketing campaigns.



Healthcare: Unified Patient Records

Integrate structured and unstructured health records into a scalable lakehouse architecture for faster diagnostics.



Finance: Real-Time Fraud Detection

Stream financial transaction data using Spark and Delta Live Tables to detect anomalies instantly.



Media: Viewer Personalization

Ingest and transform petabyte-scale video interaction data to fuel personalized content delivery.



Logistics: Shipment Tracking System

Develop end-to-end batch + streaming pipelines to optimize delivery networks and reduce delays.

Certification & Job Readiness

Students build deep hands-on skills for real-world data infrastructure roles. **You'll Graduate With:**



Real-World Exposure



Fireside chats with
data engineers from
top tech firms



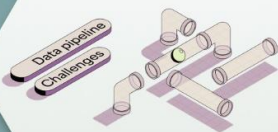
Deep dives into
Databricks production
environments



Spark/Delta Lake
hands-on labs



Architecture
teardown sessions



Peer-reviewed
pipeline challenges



Built to simulate
enterprise-grade data ops

Dream Companies Await





EdTech

Our offline locations

Hosur, Salem, Dharmapuri
Tirupattur, Bengaluru



www.teqcertify.ai