Professional Data Engineer

Course Overview

A Professional Data Engineer course—such as the widely recognized Google Cloud learning path—prepares you to design, build, and operationalize robust data pipelines. The curriculum covers core skills in data processing systems, building data pipelines, and machine learning model operations

Course Description

Professional Data Engineer program summarizes the course's purpose, target audience, and key learning outcomes. It explains what you will study and how those skills apply to real-world engineering roles.
This course provides a comprehensive foundation in designing, building, operationalizing, and securing robust data processing systems. Students will learn how to transform raw, unstructured data into high-quality, actionable insights that drive business decisions. The curriculum emphasizes scalable architecture, data quality, and automation across batch and real-time streaming environments.Who Should EnrollAspiring Data Engineers looking to build a structured foundation.Software Developers transitioning into data-focused roles.Data Analysts wanting to move from querying data to building infrastructure.IT Professionals preparing for professional cloud data engineering certifications

Learning Curriculum

Module 1

Design Data Processing Systems

Storage Systems: Choosing the optimal storage (e.g., Relational, NoSQL, or Data Warehouses) based on schema, scalability, and performance needs.Data Pipelines: Designing batch and streaming pipelines.Security and Compliance: Implementing encryption, access control, and compliance regulations

Module 2

Build and Operationalize Data Processing Systems

Extract, Transform, and Load (ETL): Processing and moving data from source to destination using tools like Apache Spark or dataflow.Data Warehousing and Lakes: Building analytics engines and structuring data for business intelligence.Orchestration and Automation: Automating workflows, scheduling jobs, and monitoring system health

Module 3

Operationalize Machine Learning Models

ML Pipeline Integration: Building repeatable and scalable pipelines for training and evaluating models.Model Deployment: Serving models in production, tracking performance, and retraining based on new data

Module 4

Ensure Solution Quality

Testing and Validation: Testing pipelines for accuracy and reliability.Monitoring and Logging: Tracking system metrics, setting up alerts, and debugging pipeline errors.

© 2026 Proasia Upskilling Sdn Bhd | www.proasia.my