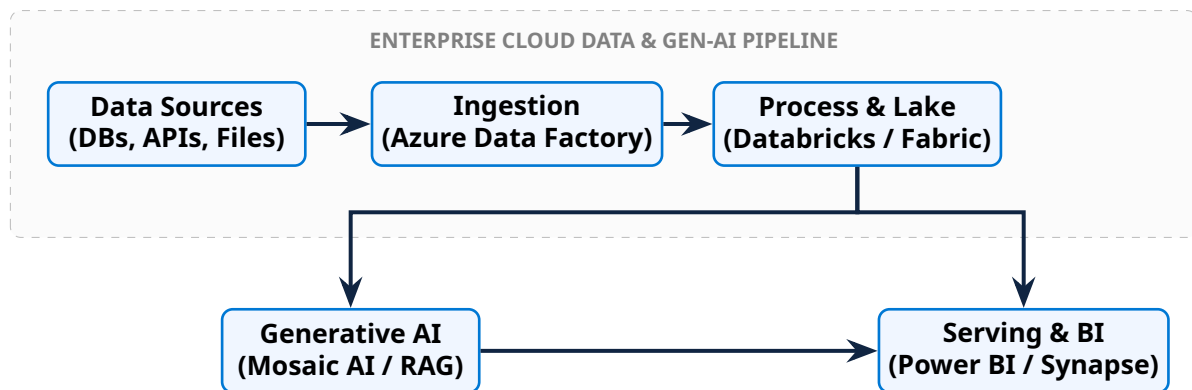


MADHU CLOUD ACADEMY

Azure Data Engineering & AI Master Program

Become Industry-Ready in Azure Data Engineering & Generative AI

Master Azure Data Engineering, Microsoft Fabric, Databricks, Power BI, and Generative AI with real-time projects, hands-on implementation, and production-level training.



Why Enroll in This Program?

This comprehensive program is designed by cloud experts to bridge the gap between academic theory and actual production environments. We focus heavily on modern data warehouse patterns, unified analytics with Microsoft Fabric, and production-level LLM integrations using Databricks and Mosaic AI.

For admissions, inquiries, and schedules, visit:

<https://madhucloudacademy.com>

Program Overview & Core Matrix

Program Highlights

- ✓ **Industry-Oriented Curriculum:** Up-to-date with current enterprise standards.
- ✓ **Real-Time Projects:** Build end-to-end working production architectures.
- ✓ **Production-Level Scenarios:** Handle performance tuning, recovery, and logging.
- ✓ **Mock Interviews & Resume Prep:** Comprehensive career counseling.
- ✓ **Placement Assurance:** Direct access to hiring opportunities.
- ✓ **Flexible Batches:** Weekend & weekday batches.
- ✓ **Hybrid Learning:** Classroom & online dynamic sessions.
- ✓ **Recorded Access:** Self-paced learning and review.
- ✓ **Integrated Path:** Unified Azure and Generative AI learning.

Who Can Join?

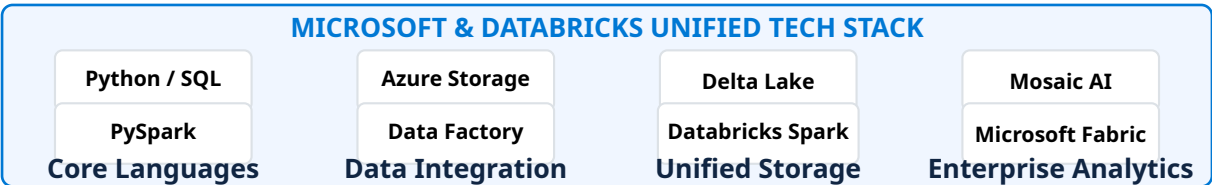
This program is strategically tailored for both technology newcomers and experienced developers aiming for a career transition:

- Freshers & College Graduates
- SQL & Database Developers
- ETL / Traditional BI Developers
- Data Analysts & Business Analysts
- Working Professionals seeking upskilling
- Career Transition Learners

Technologies Covered

We provide deep-dive practical learning across the entire modern data platform ecosystem:

Languages & Querying	Cloud & Orchestration	Data Processing	BI, AI & DevOps
Python	Azure Storage Gen2	Azure Databricks	Microsoft Fabric
SQL	Azure Data Factory (ADF)	Delta Lake	Power BI
PySpark	Azure Synapse Analytics	Mosaic AI	Azure DevOps & Git



Syllabus Modules (1 — 5)

Module 1: SQL & Database Fundamentals

Establish a solid foundation in relational database management systems and querying.

- SQL Basics & DDL/DML/DQL Commands
- Inner, Left, Right, Full, and Cross Joins
- Subqueries, Nested Queries, and Correlated Queries
- Common Table Expressions (CTEs)
- Window Functions (Row Number, Rank, Lead, Lag)
- Views, Stored Procedures, and Input/Output Parameters
- Query Optimization, Execution Plans, and Indexing
- Real-Time SQL Scenarios and Live Case Studies

Module 2: Python for Data Engineering

Master the preferred backend programming language for building highly customizable pipelines.

- Python Basics, Variables, and Built-in Data Types
- Core Functions, Scoping, and Control Loops
- File Handling (JSON, CSV, Parquet, and Text files)
- Robust Exception Handling and Try-Except blocks
- Object-Oriented Programming (OOP) Concepts
- Pandas Basics: DataFrames, Filtering, and Series
- Comprehensive Python Data Engineering Interview Q&A

Module 3: Cloud & Azure Fundamentals

Unlock core cloud architecture designs and Azure administration paradigms.

- Cloud Computing Concepts (IaaS, PaaS, SaaS)
- Navigation of the Azure Portal & CLI
- Resource Groups and Lifecycle Management
- Identity & Access Management (IAM) and RBAC
- Blob vs. File vs. ADLS Storage Options
- Virtual Networks and Cloud Security Baselines
- Monitoring, Log Analytics, and Alert Setups
- Cloud Architectural Interview Questions

Module 4: Azure Data Factory (ADF)

Orchestrate code-free ingestion pipelines with enterprise scheduling and monitoring.

- Linked Services and Dataset Definitions
- Standard Activities, Control Flow, and Dataflows
- Dynamic Pipeline Parameters and Global Variables
- Complex Incremental Loads (Watermark Tables)
- Event, Schedule, and Tumbling Window Triggers
- Custom Centralized Logging Frameworks
- CI/CD Git Integrations with Azure DevOps
- **Project:** End-to-End Dynamic Metadata Ingestion

Module 5: Azure Data Lake Storage Gen2 (ADLS Gen2)

Construct reliable cloud data reservoirs using optimized file management frameworks.

- Storage Account configurations with Hierarchical Namespace (HNS)
- Designing Enterprise Folder Structures and File Management Best Practices
- Securing data using Access Control Lists (ACLs) and Shared Access Signatures (SAS)
- Deep-dive into Delta-centric Lakehouse Topologies: Bronze, Silver, and Gold Data Layers
- Data Governance, Security Best Practices, and Lifecycle Policies

Syllabus Modules (6 — 7)

Module 6: Azure Databricks

Leverage Spark cluster computing architectures to build performant, highly scalable pipelines.

- Spark Engine Architecture and Execution Plans
- Resilient Distributed Datasets (RDDs)
- High-level PySpark DataFrames and Spark SQL
- Narrow vs. Wide Transformations and Actions
- Optimized Joins, Broadcast Joins, and Aggregations
- Medallion Architecture (Bronze, Silver, Gold)
- Structured Streaming and Real-Time Basics
- Performance Tuning, Partitioning, and Z-Ordering
- Databricks Notebook Workflows and Job Scheduling
- **Project:** End-to-End Enterprise Databricks Pipeline

Latest Databricks Enterprise Features Covered:

- **Unity Catalog:** Data Governance
- **Databricks Volumes:** File Mgmt
- **Delta Live Tables (DLT):** ETL
- **Auto Loader:** File Ingestion
- **Photon Engine:** Sub-Second Query
- **Serverless SQL Warehouses**
- **Lakehouse Federation**
- **Databricks Workflows**
- **Asset Bundles (DABs)**
- **Model Serving & Vector Search**
- **AI Functions & Mosaic AI**
- **Genie & AI Assistants**

Databricks Integrated AI Companions Explained

- **Databricks Genie:** An AI-powered, context-aware conversational business assistant designed to query complex enterprise data warehouses using plain natural language.
- **Databricks Assistant:** Built-in context-aware AI coding assistant optimized for writing, debugging, refactoring, and documenting complex PySpark, SQL, and Python code blocks.



Module 7: Delta Lake

Bridge the gap between Object Storage and Relational Databases using robust transactional layers.

- ACID Compliance and Transaction Log Secrets
- Time Travel & History Queries for Audit
- High-Performance Merge, Upserts, and Deletes
- Schema Evolution and Schema Enforcement
- Vacuum and Optimize Command Execution
- Partitioning, Pruning, and File Compaction

Syllabus Modules (8 — 12)

Module 8: Synapse Analytics

Design large-scale enterprise data warehousing systems.

- Dedicated SQL Pools (MPP Architecture)
- Serverless SQL Pools for Ad-Hoc Queries
- Spark Pools and Synapse Studio Pipelines
- Dimensional Modeling: SCD Type 1 & Type 2
- Distribution Styles (Hash, Round-Robin, Replicated)

Module 10: Power BI

Generate rich visual insights from modern data storage tiers.

- Power BI Desktop Interface & Connectivity
- Advanced Modeling, Star Schema, and Card Key-Link
- DAX Basics, Calculated Columns, and Measures
- Power Query and M-Language Transformations
- Direct Lake Mode for Sub-Second Fabric Queries
- Workspace Admin, App Deployment Pipelines

Module 9: Microsoft Fabric

Unify your analytical infrastructure under Microsoft's SaaS platform.

- OneLake Architecture: Single Data SaaS Model
- Lakehouses, Warehouses, and Shortcuts
- Notebooks and Data Factory Integration
- Real-Time Intelligence & KQL Basics
- Activators & Native Power BI Integrations
- **Project:** Fabric End-to-End Implementation

Module 11: CI/CD & Azure DevOps

Automate pipeline deployments with Enterprise git paradigms.

- Git Repos, Pull Requests, and Trunk Development
- Azure Repos, Pipelines, and Branching Strategies
- ARM Templates, Bicep, and Databricks Bundles
- Environment Automation (Dev, QA, Prod)
- Automated Testing, Approvals, and Releases

Module 12: Mosaic AI & Generative AI Integration

Integrate advanced Large Language Models into enterprise-grade analytics pipelines.

- Modern LLMs and prompt engineering fundamentals
- System configurations for Mosaic AI Gateway
- PySpark AI Functions for large scale NLP
- Databricks Vector Search and Index Syncs
- Retrieval-Augmented Generation (RAG) Architecture
- Real-Time Model Serving & Inference Endpoints
- LLM Security, Grounding, and Toxicity Controls
- **Project:** Interactive RAG QA Bot over DBs



Capstone Projects, Methodology & Placement Support

Module 13: Real-Time Production Projects

Work with industry-scale architectures on fully functional cloud deployments.

- **Project 1: Retail Data Migration Pipeline**

Ingest multi-source retail transactional systems, load them dynamically via Azure Data Factory, transform historical schemas using PySpark, and store them securely in ADLS Gen2 Delta Tables.

- **Project 2: Healthcare Analytics Pipeline**

Handle high-security patient metrics. Implement HIPAA standards and configure row/column-level access control with Unity Catalog, executing analytics via serverless Azure Databricks engines.

- **Project 3: Streaming Data Processing Pipeline**

Real-time event processing from high-velocity telemetry devices. Stream payloads into Event Hubs, apply sliding window aggregates using PySpark Structured Streaming, and write directly to Delta.

- **Project 4: Microsoft Fabric End-to-End Implementation**

Deploy a unified company dashboard from scratch. Store data directly in OneLake, execute Python-based ML training in Fabric Notebooks, and display interactive charts using Direct Lake Power BI.

- **Project 5: AI-Powered Enterprise Chatbot**

Design an enterprise support agent. Sync internal knowledge documents in Vector Search databases, construct prompt contexts via Databricks Vector indexes, and call Model Serving LLMs dynamically.

Module 14: Placement & Interviews

Complete career grooming to ensure you clear assessments.

- Strategic Resume & ATS Optimization
- Professional LinkedIn Branding & Profiles
- One-on-One Technical Mock Interviews
- Standard HR & Scenario-Based Preparations
- Dedicated ADF, Databricks, and Fabric Q&A
- Project Architectural Walkthrough Prep

Training Methodology

Ensuring deep mastery and production execution:

- Extensive Hands-on Laboratory Sessions
- Step-by-Step, zero-to-advanced breakdown
- Industry Standard Scenario Implementations
- Interactive Enterprise architectural design
- Daily coding assignments & direct reviews
- Weekly Mock Assessments with Feedback

Career Opportunities & Professional Pathways

Graduates from our academy successfully pivot into highly technical enterprise roles:

Azure Data Engineer • **Databricks Engineer** • **Microsoft Fabric Engineer** • **Data Analyst**
BI Developer • **Data Platform Engineer** • **AI Data Engineer** • **Cloud Data Engineer**

Are you ready to transform your career and become an Industry-Ready Cloud & AI Specialist?

<https://madhucloudacademy.com>