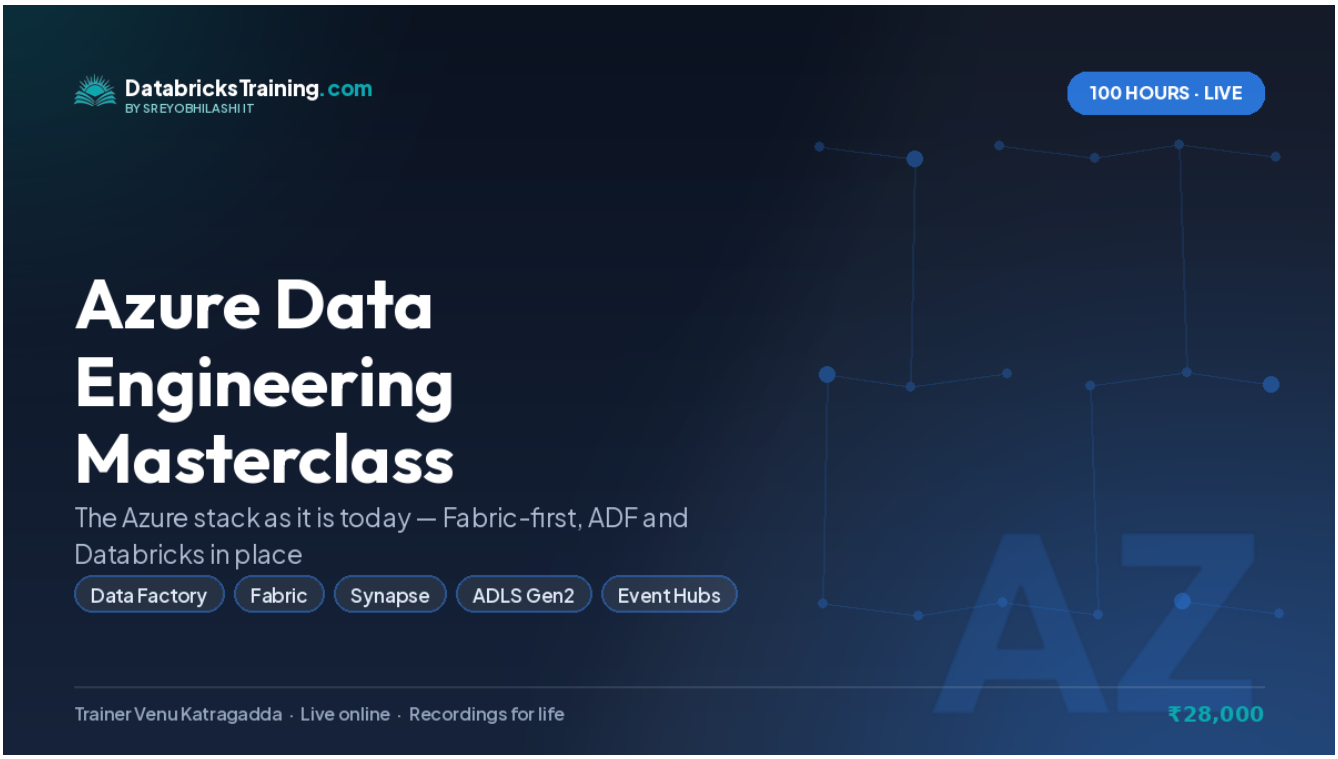

COURSE SYLLABUS

Azure Data Engineering Masterclass

The Azure stack as it is in 2026 — Fabric-first, with ADF, Databricks and Synapse in their real-world places.

100 hours	10 modules	3 projects	**₹28,000**
LIVE INSTRUCTION	STRUCTURED PATH	PORTFOLIO READY	EMI AVAILABLE

Level	Beginner to Advanced
Mode	Live, instructor-led online sessions. Every session recorded; lifetime access.
Trainer	Venu Katragadda — 14+ years in industry (L&T Infotech, Capgemini, Brillio, Manthan Systems). 1,200+ professionals trained.
Batch timings	Morning 6:00–7:30 AM IST · Morning 7:30–9:00 AM IST · Evening 7:30–9:00 PM IST · Weekend (Sat & Sun)
Certification	DP-700 Fabric Data Engineer (primary) · DP-203 concepts covered
Enquiries	+91-8500002025 / +91-9247159150 (WhatsApp) · info@databrickstraining.com

What you will be able to do

Every outcome below is demonstrated in a lab or a graded project — not just discussed in a slide.

- Build metadata-driven ADF/Fabric pipelines that scale to hundreds of tables without copy-paste
- Design an ADLS Gen2 medallion lake with Delta and Unity Catalog on Azure Databricks
- Work fluently in Microsoft Fabric: OneLake, Lakehouse, Warehouse, Notebooks, Pipelines, Direct Lake
- Stream with Event Hubs, Kafka-on-Event-Hubs, Stream Analytics and Fabric Real-Time Intelligence
- Govern with Purview/Fabric governance, managed identities, Key Vault and private endpoints
- Pass DP-700 (and understand what carried over from the retired DP-203)

Who this masterclass is for

- SSIS/Informatica developers moving to Azure
- Data engineers on the Microsoft stack
- Power BI developers going upstream
- DP-700 candidates

Prerequisites

- SQL; basic Python helpful but taught in-course
- Azure free account (₹12,000 / $200 credit) — setup walkthrough in session 1

Tools and technologies

Azure Data Factory · ADLS Gen2 · Azure Databricks · Azure Synapse Analytics · Microsoft Fabric · OneLake · Delta Lake · Event Hubs · Stream Analytics · Azure SQL · Cosmos DB · Key Vault · Microsoft Purview · Azure DevOps · Terraform · Power BI · Airflow

Curriculum — 100 hours across 10 modules

Sessions run live in small batches. Roughly 60% of each session is hands-on lab work.

MODULE 1 Azure foundations and storage — 10 hrs

- Subscriptions, resource groups, RBAC, managed identities, service principals
- ADLS Gen2: hierarchical namespace, ACLs vs RBAC, access tiers, lifecycle policies
- Networking basics: VNets, private endpoints, firewalls — the parts that break pipelines
- Key Vault for secrets; Azure CLI, Bicep and Terraform
- Cost management, budgets and tagging
- Lab: secure storage account + private endpoint + Key Vault-backed access

MODULE 2 Azure Data Factory / Fabric Data Factory — 14 hrs

- Pipelines, activities, integration runtimes (Azure, self-hosted, SSIS)
- Linked services, datasets, parameters, variables, expressions
- Copy activity at scale: staging, PolyBase, partitioned copy, DIU tuning
- Mapping Data Flows vs Databricks — when each is the right tool
- Metadata-driven framework: control tables, ForEach, dynamic datasets, watermarks
- Triggers, monitoring, alerts, retry policies, CI/CD with ADF Git integration
- Lab: build the metadata framework used in Project 1

MODULE 3 Azure Databricks and Delta Lake — 14 hrs

- Workspace, clusters, pools, cluster policies; Unity Catalog on Azure
- Mounting vs abfss:// paths; credential passthrough and managed identities
- PySpark essentials, Delta Lake operations, MERGE, time travel, OPTIMIZE
- Auto Loader from ADLS, Structured Streaming from Event Hubs
- Delta Live Tables and Workflows; Databricks + ADF orchestration patterns
- Lab: bronze→silver→gold medallion lake with quality expectations

MODULE 4 Microsoft Fabric — 14 hrs

- Fabric architecture: capacities, workspaces, OneLake, shortcuts, domains
- Lakehouse vs Warehouse vs Eventhouse — choosing correctly
- Fabric Notebooks, Spark pools, environments, V-Order
- Data pipelines, Dataflows Gen2, Copy Job, Mirroring
- Direct Lake semantic models, Power BI integration, DAX basics for engineers
- Real-Time Intelligence: Eventstreams, KQL databases, activator alerts
- Fabric CI/CD: deployment pipelines, git integration, Fabric APIs
- Lab: build Project 2 end to end in Fabric

MODULE 5 Azure Synapse Analytics — 8 hrs

- Dedicated SQL pools: distributions (hash/round-robin/replicated), partitions, resource classes
- Serverless SQL pools over the lake; OPENROWSET, external tables, cost per TB
- Synapse Spark pools and pipelines; Synapse Link for Cosmos DB / Dataverse
- Where Synapse still fits now that Fabric exists — an honest assessment
- Lab: tune a dedicated pool query with the right distribution key

MODULE 6 Streaming on Azure — 10 hrs

- Event Hubs: partitions, consumer groups, capture, throughput units, Kafka endpoint
- Kafka fundamentals and running Kafka workloads against Event Hubs
- Stream Analytics: windowing, reference data, output sinks, watermark behaviour
- Databricks Structured Streaming from Event Hubs/Kafka with exactly-once sinks
- IoT Hub overview; Fabric Eventstreams
- Lab: Project 3 real-time telemetry pipeline

MODULE 7 SQL, Cosmos DB and modelling — 8 hrs

- Azure SQL Database / Managed Instance for data engineers; elastic pools
- T-SQL for ETL: MERGE, temporal tables, partition switching
- Cosmos DB: partition keys, RU/s, change feed, analytical store
- Dimensional modelling: star schema, SCDs, surrogate keys, fact grain
- Lab: design and load a conformed star schema

MODULE 8 Orchestration with Airflow and DevOps — 8 hrs

- Airflow on Azure: Azure Data Factory Managed Airflow, AKS-hosted, or Astronomer
- DAG patterns for Azure: sensors, ADF/Databricks operators, backfills
- Azure DevOps and GitHub Actions: build/release for ADF, Databricks (DABs) and Fabric
- Testing data pipelines; Great Expectations / Fabric data quality
- Lab: full CI/CD promotion dev → test → prod

MODULE 9 Governance, security and observability — 6 hrs

- Microsoft Purview: scanning, classification, lineage, glossary
- Fabric governance: domains, endorsement, sensitivity labels, Purview integration
- Managed identity everywhere; eliminating keys from pipelines
- Log Analytics, KQL for pipeline monitoring, cost telemetry
- Lab: end-to-end lineage from source system to Power BI report

MODULE 10 Architecture, capstone and DP-700 prep — 8 hrs

- Reference architectures: Fabric-first, Databricks-first, hybrid — with trade-offs
- Migration: SSIS → ADF, on-prem SQL → Fabric, Synapse → Fabric
- Capstone presentation and review
- DP-700 exam guide walkthrough, 120 practice questions, two mocks
- Azure data engineer interview questions and resume framing

Hands-on projects

You leave with work you can demo in an interview. Each project is code-reviewed and returned with written feedback.

PROJECT 1 Metadata-driven ingestion framework

One ADF pipeline + control table ingesting 120 source tables with watermarking, schema drift handling, retries and audit logging.

PROJECT 2 Fabric lakehouse with Direct Lake

OneLake medallion lakehouse, Fabric notebooks for transformation, semantic model and a Power BI report on Direct Lake mode.

PROJECT 3 Real-time telemetry

Event Hubs → Databricks Structured Streaming → Delta → Power BI, with alerting and late-data handling.

Frequently asked questions

Is DP-203 still valid?

Microsoft retired the DP-203 Azure Data Engineer Associate exam and DP-700 (Fabric Data Engineer) is the current data engineering certification. We teach the DP-700 objectives as the certification path, and still cover ADF, Databricks and Synapse in depth because that is what production Azure estates actually run on. ~90% confident on retirement specifics — always confirm current status on Microsoft Learn before booking.

Fabric or Databricks — which should I learn?

Both, and the course covers both. Fabric is where Microsoft-centric BI-driven organisations are heading; Azure Databricks remains the default for heavy engineering and ML workloads. Module 10 gives you a decision framework you can defend in an architecture review.

Will I need a paid Azure subscription?

The free account credit covers most labs. Fabric labs use the free Fabric trial capacity. A few Synapse dedicated-pool labs cost a few dollars — we always pause/delete resources at the end of the session.

I come from SSIS. Is this too advanced?

No, it is built for that transition. Module 2 explicitly maps SSIS concepts (control flow, data flow, packages, configurations) to their ADF and Fabric equivalents.

How to enrol

- Book a free demo. Sit in on a live session with the current batch — no card, no commitment. WhatsApp +91-9247159150 or call +91-8500002025.
- Pick your batch. Weekday morning, weekday evening or weekend. All times IST.
- Pay the fee (₹28,000, instalments and EMI available) and get your workspace, material and calendar invites.
- Start building. Labs from session one; recordings posted the same day.

What is included

- 100 hours of live instruction by Venu Katragadda
- Lifetime access to all session recordings
- 3 code-reviewed portfolio projects
- Preparation for: DP-700 Fabric Data Engineer (primary) · DP-203 concepts covered
- Resume review, LinkedIn optimisation and one mock interview with written feedback
- Free re-attend of any future batch of the same masterclass
- Doubt-clearing sessions and a private student group

Refund policy in brief

When you cancel	Refund
Before the batch starts	Up to 90%
Within 5 days of start, max 2 classes attended	Up to 70%
After 7 days, or more than 5 classes attended	No refund — free transfer to a later batch

Full policy: databrickstraining.com/refund-policy/

Talk to us

Call	+91-8500002025 · +91-9247159150
WhatsApp	+91-9247159150 — usually answered within the hour
Email	info@databrickstraining.com
Website	databrickstraining.com/courses/azure-data-engineering-masterclass/
Foundation tier	databrickstraining.in — shorter 70–80 hour courses if this one is too advanced
Institute	Sreyobhilashi IT · Udyam UDYAM-TS-02-0334643 · 112, Samshraya Homes, opp. PF Office, Balanagar, Kukatpally, Hyderabad, Telangana 500037