Mastering Azure Data Explorer (ADX)

Course Description

This course provides a comprehensive, hands-on introduction to Azure Data Explorer (ADX), covering fundamentals, ingestion pipelines, advanced KQL analytics, schema design, optimization, and governance. Participants will learn to design scalable data solutions, optimize query performance, and implement secure, governed ADX environments.

Duration

3 Days (24 hours)

Prerequisites

- Basic knowledge of Azure services
- Familiarity with SQL or query languages
- Understanding of data concepts (tables, schemas, ingestion)

Learning Objectives

By the end of this course, participants will be able to:

- Understand ADX architecture and use cases
- Ingest data from multiple sources and formats
- Apply advanced KQL for analytics and anomaly detection
- Design efficient schemas and retention policies
- Optimize performance and manage ADX clusters
- Implement security, governance, and monitoring strategies

Day-wise Agenda

Day 1 – Fundamentals & Data Ingestion

- Introduction to Azure Data Explorer
    - What is ADX?
    - ADX architecture (Cluster, Database, Tables)
    - Key use cases (log analytics, telemetry, IoT, clickstream)
- Data Ingestion Concepts
    - Batch vs Streaming ingestion
    - Ingestion pipelines overview
- Ingestion Sources
    - Azure Event Hub
    - Azure IoT Hub
    - Azure Blob Storage / ADLS Gen2
    - Kafka / Logstash
- Data Formats & Mappings
    - JSON, CSV, TSV, AVRO, PARQUET
    - Ingestion mappings and schema alignment

Day 2 – Advanced Analytics & Data Modeling

- Advanced KQL Techniques
    - Joins and performance patterns
    - Time-series analysis
    - Anomaly detection functions
    - Window functions
    - Materialized views
    - User-defined functions (UDFs)

- Data Modeling & Schema Design
    - Fact vs Dimension tables
    - Hot vs Cold data strategies
    - Update policies
    - Retention and caching policies
    - Reference data handling

Day 3 – Optimization, Security & Operations

- Performance Optimization
    - Query optimization techniques
    - Shard and ingestion tuning
    - Best practices for joins
    - Using hint.strategy
    - Materialized views vs on-demand queries
    - Capacity planning and scaling clusters
- Security & Governance
    - Authentication (AAD, Managed Identity)
    - Authorization and RBAC
    - Row-level security
    - Cross-tenant and cross-cluster access
    - Private endpoints and network security
    - Auditing and compliance basics
- Monitoring & Operations
    - Monitoring ADX clusters
    - Ingestion failures and diagnostics
    - Query performance monitoring

- Cost management and optimization
- Backup and disaster recovery strategies