

Azure Data Eng: ADF, Databricks & Delta Lake

Learn to build data pipelines with Azure Data Factory and process data at scale with Databricks and Delta Lake.

5 Days

DURATION

68

TOPICS

COURSE OVERVIEW

This course covers Azure storage options and Azure Data Factory, including copy, control flow and data flow activities such as filter, lookup, join and pivot transformations, along with pipeline parameters, triggers, scheduling and CI/CD using GitHub or DevOps. It then moves to Azure Databricks, covering cluster setup, notebooks, tables, joins, aggregates and visualisations, followed by Databricks Jobs for scheduling, dependencies and alerts, and Delta Lake for data ingestion, updates, merges, constraints and versioning. It concludes with integrating Azure Data Factory and Databricks, running notebooks within pipelines, passing parameters, working with Delta Lake tables and connecting to Power BI.

WHO THIS COURSE IS FOR

Data Engineers

ETL Developers

BI Developers

Data Analysts

Database Administrators

LEARNING OBJECTIVES

- Configure Azure storage services such as ADLS Gen2 and Azure SQL Database
- Build Azure Data Factory pipelines using copy and control flow activities
- Apply data flow transformations such as join, pivot and derived column
- Schedule and trigger pipelines using event-based and scheduled triggers
- Configure and manage Azure Databricks clusters and notebooks
- Create and schedule Databricks Jobs with dependencies and alerts
- Implement Delta Lake tables including updates, merges and versioning
- Integrate Azure Data Factory pipelines with Azure Databricks notebooks

68 Topics

DAY 1 Azure Data Factory and Activity

TOPICS

- Overview of Different Types Azure Storage - Overview
- ADLS Gen2
- Azure SQL Database | Azure SQL Server
- Key Component of Azure Data Factory
- Copy Data Flow Activity in ADF
- How to Ingest Data from On-prem SQL Server To Azure Data Factory - Will show Demo but Participants may not be testing due to time constraints
- Control Flow Activity
- Get Metadata Activity
- Filter Activity
- If Activity
- Append Activity - High Level Overview
- Wait Activity - High Level Overview
- Lookup Activity
- Dataflow Transformation
- Source Transformation
- Sink Transformation

DAY 2 Azure Data Factory and Transformation

TOPICS

- Derived Column Transformation
- Select Transformation
- Join Transformation
- Pivot Transformation
- Execute Pipeline
- Store Procedure Activity
- File format Conversion using copy activity (csv to parquet) - High Level Overview
- Pipeline Parameters
- Email Notification Activity - High Level Overview
- Event Based Trigger
- Scheduling Pipeline Through Scheduling Window
- CI/CD using GitHub /DevOps

TOPICS

- Overview of Apache Spark and Databricks
- Databricks Architecture
- Azure implementation
- Getting started with Databricks
- Creating and configuring your Cluster
- Uploading data
- Creating a Table
- Connecting to a Spark data source
- Previewing your Table
- Columns and Datatypes basics
- View aggregates
- Perform Joins
- Visualisations & Data Frames

PLOTS

- Choosing Chart types
- Chart Toolbar
- Layout and styling considerations

TOPICS

- Databricks Jobs
- Creating a Job
- View Jobs and Job details
- Running your first Job
- Scheduling Jobs
- Setting Parameters
- Viewing completed jobs
- Managing Dependencies
- Setting up Alerts
- Delta Lake and Delta Tables
- Getting data into Delta Lake
- Delete, update, merge
- Constraints
- Versioning
- Concurrency - High Level Overview
- Delta Commands

TOPICS

- Describe how Azure Databricks notebooks can be run in a pipeline
- Create an Azure Data Factory linked service for Azure Databricks
- Use a Notebook activity in a pipeline
- Pass parameters to a notebook
- Create Spark catalogue tables for Delta Lake data - High Level Overview
- Use Delta Lake tables for streaming data - High Level Overview
- Azure Databrick connection With PowerBI - Overview
- Devops/Github

WHAT IS INCLUDED

Everything You Need to Succeed

- Koenig course material for every module
- Delivery by certified instructors and subject matter experts
- Hands-on exercises throughout the course
- Course completion certificate from Koenig

ABOUT KOENIG

Who We Are

Established in 1993, Koenig Solutions is a globally recognised training provider with 35+ technology partnerships, 5,000+ training programmes and a network of certified instructors and subject matter experts, supporting individuals and enterprise teams worldwide.

FLEXIBLE DELIVERY

Learn the Way That Suits You

Classroom Training

Live Online Classes (ILO)

Fly-Me-A-Trainer (FMAT)

Flexi (Self-Paced)

1-on-1 Training

Ready to Get Started?

Talk to a learning advisor about schedules, group training and customised delivery.

Visit koenig-solutions.com

✉ info@koenig-solutions.com