Extended Data Engineering on Azure with Azure Synapse, Microsoft Purview and Power BI

Course Description: This comprehensive course provides an in-depth exploration of data engineering on Azure. Participants will learn how to design, implement, and manage data engineering workflows using a range of Azure services, including Azure Synapse Analytics, Microsoft Purview, Azure Databricks, and Power BI. The course covers data storage, processing, quality management, governance, streaming, and visualization techniques, integrating practical scenarios and real-world applications.

Duration: 32 hours

Target Audience:

- Data Engineers
- Data Analysts
- Data Architects
- Database Administrators
- IT Professionals with a focus on data management
- Anyone interested in learning about data engineering on Azure

Prerequisites:

- Basic understanding of data concepts and database management
- Familiarity with cloud computing concepts
- Experience with SQL
- Basic programming knowledge (preferably Python)
- An Azure account (for hands-on exercises)

Content Coverage :

Module 1: Get Started with Data Engineering on Azure

- Lesson 1: Introduction to Data Engineering on Azure
- Lesson 2: Introduction to Azure Data Lake Storage Gen2
- Lesson 3: Introduction to Azure Synapse Analytics

Module 2: Data Quality and Governance with Microsoft Purview

- Lesson 1: Introduction to Microsoft Purview
- Lesson 2: Managing Data Quality in Microsoft Purview
- Lesson 3: Registering and Scanning Data Sources
- Lesson 4: Utilizing a Business Glossary for Data Standardization
- Lesson 5: Searching and Navigating the Data Catalog
- Lesson 6: Customizing Metadata with Enrichments and Classifications
- Lesson 7: Integrate Microsoft Purview and Azure Synapse Analytics

Module 3: Build Data Analytics Solutions Using Azure Synapse Analytics Serverless SQL Pools

- Lesson 1: Use a Serverless SQL Pool to Query Files in a Data Lake
- Lesson 2: Use a Serverless SQL Pool to Transform Data
- Lesson 3: Create a Lake Database

Module 4: Perform Data Engineering with Azure Synapse Apache Spark Pools

- Lesson 1: Analyze Data with Apache Spark in Azure Synapse Analytics
- Lesson 2: Transform Data with Apache Spark in Azure Synapse Analytics
- Lesson 3: Use Delta Lake in Azure Synapse Analytics
- Lesson 4: Bronze Layer: Raw Data Ingestion
- Lesson 5: Silver Layer: Cleaned and Enriched Data
- Lesson 6: Gold Layer: Aggregated and Business-ready Data

Module 5: Work with Data Warehouses Using Azure Synapse Analytics

- Lesson 1: Analyze Data in a Relational Data Warehouse
- Lesson 2: Load Data into a Relational Data Warehouse

Module 6: Transfer and Transform Data with Azure Synapse Analytics Pipelines

- Lesson 1: Build a Data Pipeline in Azure Synapse Analytics
- Lesson 2: Use Spark Notebooks in an Azure Synapse Pipeline
- Lesson 3: Implementing Medallion Architecture in Azure Synapse Analytics

Module 7: Work with Hybrid Transactional and Analytical Processing (HTAP) Solutions Using Azure Synapse Analytics

- Lesson 1: Plan Hybrid Transactional and Analytical Processing
- Lesson 2: Implement Azure Synapse Link with Azure Cosmos DB
- Lesson 3: Implement Azure Synapse Link for SQL

Module 8: Implement a Data Streaming Solution with Azure Stream Analytics

- Lesson 1: Get Started with Azure Stream Analytics
- Lesson 2: Ingest Streaming Data Using Azure Stream Analytics and Azure Synapse Analytics
- Lesson 3: Visualize Real-Time Data with Azure Stream Analytics and Power BI

Module 9: Data Engineering with Azure Databricks

- Lesson 1: Explore Azure Databricks
- Lesson 2: Use Apache Spark in Azure Databricks
- Lesson 3: Run Azure Databricks Notebooks in Azure Data Factory

Module 10: Visualizing Data with Power BI

- Lesson 1: Introduction to Power BI
- Lesson 2: Connecting Power BI to Azure Synapse

- Lesson 3: Creating Dashboards and Reports in Power BI
- Lesson 4: Real-Time Data Visualization with Power BI
- Lesson 5: End-to-End Data Flow: Data Ingestion to Visualization
- Lesson 6: Processing Data in Synapse and Consuming in Power BI