DP-203T00: Data Engineering on Microsoft Azure

Duration: 4 Days

Method: Instructor-Led Training (ILT) | Live Online Training

Certification: *Microsoft Certified: Azure Data Engineer Associate —*
Exam: *DP-203: Data Engineering on Microsoft Azure*

Course Description

In this course, participants will learn about data engineering as it pertains to working with batch and real-time analytical solutions using Azure data platform technologies. They will learn how to create, manage, use, and configure data services in the Azure portal. Through a blend of hands-on labs and interactive lectures, participants will learn to build and maintain secure and compliant data processing pipelines by using different tools and techniques. Participants will learn to use various Azure data services and languages to store and produce cleansed and enhanced datasets for analysis. Lastly, participants will be able to create a real-time analytical system to create real-time analytical solutions.

Target Audience

This course is intended for:

- Data Professionals
- Data Architects
- Data Engineer
- Business Intelligence Professionals

Prerequisites

To attend this course, candidates must have:

- Professional experience with data solutions
- Knowledge of cloud computing and core data concepts or completed the following courses:
  - *AZ-900T00: Azure Fundamentals*
  - *DP-900T00: Microsoft Azure Data Fundamentals*

REAL WORLD
TECHNOLOGY TRAINING & SOLUTIONS
"Training You Can Really Use"

Course Objectives

Upon successful completion of this course, attendees will be able to:

- Explore compute and storage options for data engineering workloads in Azure
- Run interactive queries using serverless SQL pools
- Perform data Exploration and Transformation in Azure Databricks
- Explore, transform, and load data into the Data Warehouse using Apache Spark
- Ingest and load Data into the Data Warehouse
- Transform Data with Azure Data Factory or Azure Synapse Pipelines
- Integrate Data from Notebooks with Azure Data Factory or Azure Synapse Pipelines
- Support Hybrid Transactional Analytical Processing (HTAP) with Azure Synapse Link
- Perform end-to-end security with Azure Synapse Analytics
- Perform real-time Stream Processing with Stream Analytics
- Create a Stream Processing Solution with Event Hubs and Azure Databricks.

Course Topics

Module 1: Explore Compute and Storage Options for Data Engineering Workloads

- Introduction to Azure Synapse Analytics
- Describe Azure Databricks
- Introduction to Azure Data Lake Storage
- Describe Delta Lake Architecture
- Work with Data Streams by Using Azure Stream Analytics

Module 2: Run Interactive Queries Using Azure Synapse Analytics Serverless SQL Pools

- Explore Azure Synapse Serverless SQL Pools Capabilities
- Query data in the lake using Azure Synapse Serverless SQL Pools
- Create Metadata Objects in Azure Synapse Serverless SQL Pools
- Secure Data and Manage Users in Azure Synapse Serverless SQL Pools

Module 3: Data Exploration and Transformation in Azure Databricks

- Describe Azure Databricks
- Read and Write Data in Azure Databricks
- Work with DataFrames in Azure Databricks
- Work with DataFrames Advanced Methods in Azure Databricks

Module 4: Explore, Transform, and Load Data into the Data Warehouse Using Apache Spark

- Understand Big Data Engineering with Apache Spark in Azure Synapse Analytics
- Ingest Data with Apache Spark Notebooks in Azure Synapse Analytics
- Transform Data with DataFrames in Apache Spark Pools in Azure Synapse Analytics
- Integrate SQL and Apache Spark Pools in Azure Synapse Analytics

Course Topics *Continued*

Module 5: Ingest and Load Data into the Data Warehouse

- Use Data Loading Best Practices in Azure Synapse Analytics
- Petabyte-Scale Ingestion with Azure Data Factory

Module 6: Transform Data with Azure Data Factory or Azure Synapse Pipelines

- Data Integration with Azure Data Factory or Azure Synapse Pipelines
- Code-Free Transformation at Scale with Azure Data Factory or Azure Synapse Pipelines

Module 7: Orchestrate Data Movement and Transformation in Azure Synapse Pipelines

- Orchestrate Data Movement and Transformation in Azure Data Factory

Module 8: End-to-End Security with Azure Synapse Analytics

- Secure a Data Warehouse in Azure Synapse Analytics
- Configure and Manage Secrets in Azure Key Vault
- Implement Compliance Controls for Sensitive Data

Module 9: Support Hybrid Transactional Analytical Processing (HTAP) with Azure Synapse Link

- Design Hybrid Transactional and Analytical Processing Using Azure Synapse Analytics
- Configure Azure Synapse Link with Azure Cosmos DB
- Query Azure Cosmos DB with Apache Spark Pools
- Query Azure Cosmos DB with Serverless SQL Pools

Module 10: Real-Time Stream Processing with Stream Analytics

- Enable Reliable Messaging for Big Data Applications Using Azure Event Hubs
- Work with Data Streams by Using Azure Stream Analytics
- Ingest Data Streams with Azure Stream Analytics

Module 11: Create a Stream Processing Solution with Event Hubs and Azure Databricks

- Process Streaming Data with Azure Databricks Structured Streaming

LABS INCLUDED