Document Generated: 07/27/2024

Learning Style: Virtual Classroom

Provider: Microsoft

Difficulty: Intermediate

Course Duration: 1 Day

Implementing a Data Analytics Solution with Azure Databricks (DP-3011)

If you enroll in this course at the listed price, you receive a Free Official Exam Voucher for the DP-3011 Exam. This course does not include Exam Voucher if enrolled within the Master Subscription, however, you can request to purchase the Official Exam Voucher separately.

About This Course:

Implementing a Data Analytics Solution with Azure Databricks course is designed for data professionals, including data engineers and data scientists, who aim to utilize Azure Databricks for advanced data analytics solutions. This course provides a comprehensive understanding of how to implement, configure, and optimize data analytics pipelines using Azure Databricks, an Apache Spark-based analytics platform.

Course Objectives:

- Introduction to Azure Databricks
- Setting Up Azure Databricks Workspaces
- Data Ingestion and Processing
- Developing Data Pipelines
- Data Exploration and Analysis
- Machine Learning and AI Integration
- Optimizing and Tuning Performance
- Security and Compliance

Audience:

- Data Engineer
- Data Scientist

Prerequisites:

None

Course Outline:

Explore Azure Databricks

- Provision an Azure Databricks workspace
- Identify core workloads and personas for Azure Databricks
- Describe key concepts of an Azure Databricks solution
- Lab: Explore Azure Databricks

Use Apache Spark in Azure Databricks

- Describe key elements of the Apache Spark architecture
- Create and configure a Spark cluster
- Describe use cases for Spark
- Use Spark to process and analyze data stored in files
- Use Spark to visualize data
- Lab: Use Spark in Azure Databricks

Use Delta Lake in Azure Databricks

- Describe core features and capabilities of Delta Lake
- Create and use Delta Lake tables in Azure Databricks
- Create Spark catalog tables for Delta Lake data
- Use Delta Lake tables for streaming data
- Lab: Use Delta Lake in Azure Databricks

Use SQL Warehouses in Azure Databricks

- Create and configure SQL Warehouses in Azure Databricks
- Create databases and tables
- Create queries and dashboards
- Lab: Use a SQL Warehouse in Azure Databricks

Run Azure Databricks Notebooks with Azure Data Factory

- 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
- Lab: Run an Azure Databricks Notebook with Azure Data Factory