

DP-3011: Implementing a Data Analytics Solution with Azure Databricks

Course Outline

1) 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

2) 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

3) 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

4) 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

5) 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