



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

Microsoft - Data & AI

- **Nível:** Intermédio
 - **Duração:** 7h
-

Sobre o curso

Learn how to harness the power of Apache Spark and powerful clusters running on the Azure Databricks platform to run large data engineering workloads in the cloud.

Objetivos

- Provision an Azure Databricks workspace.
 - Identify core workloads and personas for Azure Databricks.
 - Describe key concepts of an Azure Databricks solution.
 - 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.
 - 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.
 - Create and configure SQL Warehouses in Azure Databricks.
 - Create databases and tables.
 - Create queries and dashboards.
 - 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.
-

Programa

- Explore Azure Databricks
- Use Apache Spark in Azure Databricks
- Use Delta Lake in Azure Databricks
- Use SQL Warehouses in Azure Databricks
- Run Azure Databricks Notebooks with Azure Data Factory

Explore Azure Databricks

- Get started with Azure Databricks
- Identify Azure Databricks workloads
- Understand key concepts

Use Apache Spark in Azure Databricks

- Get to know Spark
- Create a Spark cluster
- Use Spark in notebooks
- Use Spark to work with data files
- Visualize data

Use Delta Lake in Azure Databricks

- Get Started with Delta Lake
- Create Delta Lake tables
- Create and query catalog tables

- Use Delta Lake for streaming data

Use SQL Warehouses in Azure Databricks

- Get started with SQL Warehouses
- Create databases and tables
- Create queries and dashboards

Run Azure Databricks Notebooks with Azure Data Factory

- Understand Azure Databricks notebooks and pipelines
- Create a linked service for Azure Databricks
- Use a Notebook activity in a pipeline
- Use parameters in a notebook