

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

Course Details

Course Code: DP-3011

Duration: 1 day

Notes:

- This course syllabus should be used to determine whether the course is appropriate for the students, based on their current skills and technical training needs.
- Course content, prices, and availability are subject to change without notice.
- Terms and Conditions apply

Elements of this syllabus are subject to change.

About this course

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.

This learning path helps prepare you for Exam DP-203: Data Engineering on Microsoft Azure.

Prerequisites

- None.

Academy IT Pty Ltd

Level 4, 45 Grenfell Street
ADELAIDE 5000

Email: sales@academyit.com.au

Web: www.academyit.com.au

Phone: 08 7324 9800

Brian: 0400 112 083

Explore Azure Databricks

Azure Databricks is a cloud service that provides a scalable platform for data analytics using Apache Spark.

Learning objectives

In this module, you'll learn how to:

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

Use Apache Spark in Azure Databricks

Azure Databricks is built on Apache Spark and enables data engineers and analysts to run Spark jobs to transform, analyze and visualize data at scale.

Learning objectives

In this module, you'll learn how to:

- 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.

Use Delta Lake in Azure Databricks

Delta Lake is an open source relational storage area for Spark that you can use to implement a data lakehouse architecture in Azure Databricks.

Learning objectives

In this module, you'll learn how to:

- 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.

Use SQL Warehouses in Azure Databricks

Azure Databricks provides SQL Warehouses that enable data analysts to work with data using familiar relational SQL queries.

Learning objectives

In this module, you'll learn how to:

- Create and configure SQL Warehouses in Azure Databricks.
- Create databases and tables.
- Create queries and dashboards.

Run Azure Databricks Notebooks with Azure Data Factory

Using pipelines in Azure Data Factory to run notebooks in Azure Databricks enables you to automate data engineering processes at cloud scale.

Learning objectives

In this module, you'll learn how to:

- 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.