



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

## Description

Discover how to master Apache Spark, the essential tool for processing massive data, by using the high-performance clusters of Azure Databricks. This course, “**Implementing a Data Analytics Solution with Azure Databricks (DP-3011)**”, guides you through best practices for managing and analyzing large volumes of data in the cloud, enabling you to become an expert in data engineering on cutting-edge cloud platforms.

Enhance your data science skills and open the door to new professional opportunities by mastering large-scale data processing with Spark and Databricks in the “Implementing a Data Analytics Solution with Azure Databricks (DP-3011)” training.

## Niveau

Intermédiaire

## Course Content

### Module 1 - Exploring Azure Databricks

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

### Module 2 - Using 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

### Module 3 - Using 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

#### **Module 4 - Using SQL warehouses in Azure Databricks**

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

#### **Module 5 - Running 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

#### **Documentation**

- Digital course material included

#### **Participant profiles**

- Data scientists
- Data ingénieurs

#### **Prerequisites**

- Having basic knowledge of Azure

#### **Objectives**

- Understanding Azure Databricks and its ecosystem
- Mastering Apache Spark in Azure Databricks
- Exploiting Delta Lake for advanced data management
- Integrating SQL into Azure Databricks
- Automating with Azure Data Factory

#### **Description**

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

#### **Classroom Registration Price (CHF)**

900

#### **Virtual Classroom Registration Price (CHF)**

850

#### **Duration (in Days)**

1

#### **Reference**

DP-3011