

# Implementing a Data Analytics Solution with Azure Databricks

**Course Duration: 32 Hrs.**

**Course Code: DP-3011**

## Course Overview

The DP-3011 Implementing a Data Analytics Solution with Azure Databricks Course is designed to equip professionals with skills to apply data analytics solutions using Azure Databricks, which is a powerful platform for large data analytics and artificial intelligence.

## Course Content

### Module 01:

#### Explore Azure Databricks

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

### Module 02:

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

### Module 03:

- 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.

#### Module 04:

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

#### Module 05:

- 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.

