

# Implémentation d'une solution Data Analytics avec Azure Databricks (DP-3011)

Durée: 1 jour

Public cible: Professionnel TI|Analyste de données

## À qui s'adresse cette formation ?

Ingénieurs de données, analystes de données avancés, architectes de solutions, développeurs et professionnels de l'informatique.

## Objectifs de la formation

- Explorer et utiliser l'environnement Azure Databricks (workspaces, clusters, notebooks)
- Ingérer, transformer et analyser des données à grande échelle avec Apache Spark (DataFrames, SQL, PySpark)

## Description sommaire

Apprenez à exploiter la puissance d'Apache Spark et des clusters performants sur la plateforme Azure Databricks pour exécuter des charges de travail d'ingénierie des données à grande échelle dans le cloud.

## Contenu du plan de cours

- Explorer Azure Databricks
- Utiliser Apache Spark dans Azure Databricks
- Utiliser Delta Lake dans Azure Databricks
- Utiliser des entrepôts SQL dans Azure Databricks
- Exécuter des notebooks Azure Databricks avec Azure Data Factory

## Approche et méthode pédagogique

Approche pratique et structurée combinant théorie ciblée et ateliers guidés. Les participants implémentent progressivement une solution de Data Analytics avec Azure Databricks à travers des exercices concrets inspirés de scénarios professionnels, favorisant une mise en application immé-

diate des apprentissages. Ils apprennent à traiter, transformer et analyser des données à grande échelle en exploitant les capacités de Spark et les outils collaboratifs de Databricks. Animée par un formateur certifié Microsoft, la formation privilégie l'interactivité et le développement de compétences techniques directement transférables pour concevoir des solutions analytiques performantes en contexte professionnel.

## **Prérequis**

- Notions de base sur Azure (AZ-900 ou contenu équivalent).
- Connaissances de base en Python et en visualisation de données avec matplotlib/seaborn.

## **Recommandations**

- Bases en analyse de données (nettoyage, transformation, visualisation)
- Connaissance de Python et/ou SQL
- Compréhension des concepts de traitement des données (batch, big data)
- Familiarité avec Azure et les services de données Notions de Spark ou traitement distribué (atout)