

SPARK FUNDAMENTALS



Durée

3 jours

Cette formation aborde les principes fondamentaux de Spark, un moteur open source pour le traitement de données à grande échelle qui révolutionne le monde de l'analyse et du big data.

Cette formation est l'occasion d'apprendre des leaders du secteur qui utilisent Spark, technologie reposant sur la vitesse, la facilité d'utilisation et l'analyse, et offre des opportunités et des projets pratiques pour renforcer la confiance avec l'ensemble des outils Spark.

Cette formation guidée par les instructeurs est destinée aux personnes qui souhaitent déployer Spark, un système pour le traitement de très grandes quantités de données.



Objectifs pédagogiques

À la fin de la formation, vous serez en mesure de :

- Présenter une introduction à Spark, Spark Structured APIs, Spark SQL, Spark operations, Structured Streaming concepts, Spark optimization, and Data processing

L'exercice pratique sur le notebook Databricks est inclus



Programme

Jour 1 :

Matin :

- Introduction à Spark
- LAB : introduction à Databricks
- API structurée Spark

Après-midi :

- Spark SQL avec Dataframes
- LAB : Lire et écrire des Dataframes

Jour 2 :

Matin :

- Spark Operations (Dataframe)
- LAB: Data Manipulation

Après-midi :

- Concepts de flux structurés
- Optimisation Spark

Jour 3 :

Matin :

- Traitement des données avec Azure Databricks
- LAB : Intégration avec d'autres services Azure

Après-midi :

- Suite LAB : Intégration avec les autres services Azure

 **Public cible**

Chef de projet, Data scientist, Développeurs

 **Prérequis**

Connaissances de base en programmation python/scala/SQL

 **Niveau**

Intermédiaire

 **Evaluation**

Quiz

 **Modalités d'apprentissage**

Stage pratique en présentiel, classe pratique à distance

 **Mode de délivrance**

Intra et inter

 **Moyens**

Slides
Labs

 **Certification associée**

N/A

 **Session délivrée en**

Français et Anglais

 **Langue(s) des supports**

Anglais

 **Tarifs**

Inter-entreprise : 2 200 €HT par participant
Intra-entreprise : sur demande