

Formation Spark SQL

Réf. SF33594

2 journées de formation pour maîtriser le module Spark pour le traitement des données structurées !

2 j Durée	1 500 € HT Inter / participant	4 100 € HT Intra (≤12 pers.)	0 Sessions à venir
--------------	-----------------------------------	---------------------------------	-----------------------

Objectifs pédagogiques

- ✓ Installer et configurer Spark SQL
- ✓ Requêter des jeux de données dans différents formats
- ✓ Réaliser des analyses de données en utilisant Spark SQL
- ✓ Visualiser les données et effectuer des requêtes sur les résultats

Public visé & prérequis

Public visé Administrateurs de base de données, Data analysts, Data scientists, Ingénieurs Big Data

Prérequis Une expérience dans l'utilisation des requêtes SQL, ainsi qu'en programmation.

Programme détaillé

Introduction à la formation Spark SQL

Présentation générale de la formation, objectifs et approche pédagogiques
Les différentes approches pour l'accès aux données avec Spark : Hive, bases de données...
Les fonctionnalités et l'architecture de Spark
Installer Spark et effectuer les configurations de base
Le module Spark SQL : fonctionnalités, intérêt et avantages par rapport à Hive
Principes de fonctionnement et architecture (Data Source API, DataFrameAPI, SQL Service...)

Requêtes, DataFrames et DataSets

Comprendre les DataFrames dans Spark
Le point de départ : SparkSession
Créer des DataFrames
Requêtes SQL de base
Vues temporaires
Créer des DataSets
Exécuter des agrégations, jointures et requêtes imbriquées

Requêtes sur différents types de données

Accéder aux données
Préparer les données
Schémas RDD
Requêter des fichiers JSON, Parquet, CSV...
Troubleshooting
Mise en cache In-Memory

Pourquoi choisir Sparks Formation ?

20 ans d'expertise IT, des formateurs experts et des sessions garanties contre l'annulation.

20 ans
d'expertise

1 068+
formations

11 178
apprenants formés/an

+67
Net Promoter Score

Avis des participants

Note moyenne des participants de la filière Data Engineering & Big Data : 4,9/5 — 13 avis

Arnaud M. — That was really useful for me, even though I am a PDA. The trainer was very patient with me (I needed more explanation as I was the only...

Thierry R. — Formation idéale pour appréhender la technologie nosql Bon support de formation. Bonne pédagogie

Votre formateur

Gérald B. — Formateur

Gérald intervient notamment sur la gestion de projet et les certifications associées, ainsi que sur l'IA appliquée aux métiers et sa gouvernance. Son activité de formation remonte à 2015, soit...

Ils nous font confiance

Cette année notre filière Data Engineering & Big Data a accompagné plus de 73 entreprises dont :

SOPRA STERIA CAPGEMINI TECHNOLOGY SERVICES AMADEUS SAS GROUPAMA SUPPORTS & SERVICES AUBAY FRAIKIN FRANCE LBC FRANCE
ORANGE BUSINESS SERVICES SA ACATUS INFORMATIQUE ACENSI SANA +63 autres

Financement & certifications

Cette formation est éligible aux financements OPCO et FNE-Formation. Nos équipes vous accompagnent dans le montage de votre dossier.



Réservez votre place dès maintenant

20 ans d'expertise · 1 000+ formations · NPS +67 · réponse sous 24h ouvrées

0 805 950 800

demande@sparks-formation.com