

Formation Spark SQL

Analyser des données avec le module Spark SQL

A l'issue de cette formation, les participants sont en mesure de réaliser des analyses de données avec Spark SQL. Ils maîtrisent l'installation et la configuration de l'outil, et sont en mesure de requêter et visualiser des jeux de données dans différents formats.

Durée

2 jours

Objectifs pédagogiques

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

Public

Data analysts, data scientists, ingénieurs
Big Data...

Prérequis

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

Programme de formation

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

Moyens et méthodes pédagogiques

- La formation alterne entre présentations des concepts théoriques et mises en application à travers d'ateliers et exercices pratiques.
- Les participants bénéficient des retours d'expérience terrains du formateur ou de la formatrice
- Un support de cours numérique est fourni aux stagiaires

Modalités d'évaluation

- **En amont de la session de formation**, un questionnaire d'auto-positionnement est remis aux participants, afin qu'ils situent leurs connaissances et compétences déjà acquises par rapport au thème de la formation.
- **En cours de formation**, l'évaluation se fait sous forme d'ateliers, exercices et travaux pratiques de validation, de retour d'observation et/ou de partage d'expérience.
- **En fin de session**, le formateur évalue les compétences et connaissances acquises par les apprenants grâce à un questionnaire reprenant les mêmes éléments que l'auto-positionnement, permettant ainsi une analyse détaillée de leur progression.