

Dremio Lakehouse & Data Engineering

Description

Dremio est un Data Lakehouse dit autrement une « plate-forme de données selfservice » qui accélère l'interrogation de sources de données variées.

Objectifs

À la fin de cette formation, vous serez en mesure de concevoir et modéliser vos ensemble de données, implémenter vos flux de transformation de données, publier et partager les ensembles de données et les intégrer aux outils d'analyse de données d'entreprise.

Programme

COMPRENDRE LES BASES DE DREMIO

- Concepts, Briques socles

EXPLOITER LE CLIENT

- Data Curation, SQL, API, BI

GÉRER SOURCE DE DONNÉES

- Data lakes & C3 Iceberg, Nessie, SGBDR, NoSQL

CONSTRUIRE COUCHES SÉMANTIQUES

- Conception des couches, Gouvernance des données

COMPRENDRE / OPTIMISER LE MOTEUR DE REQUÊTAGE

- Charge de travail, Passage à l'Échelle, Optimisation

BÂTIR DES ACCÉLÉRATION

- Stratégie, plan d'exécution, rafraîchissement

OPTIMISER CES REQUÊTES

- Architecture, Profil, Scénario

Public concerné

Data Ingénieur, Data Analyste, Data Scientiste

Prérequis

Connaître SQL, les systèmes de base de données. Maîtriser Python, Spark, git facilite une meilleure appropriation de la formation.

Durée

3 jours