

Durée : 21h | Formats : Présentiel · Distanciel · Blended

Résumé

Cette formation permet de comprendre les paradigmes NoSQL et de maîtriser MongoDB, de la modélisation des documents jusqu'à l'intégration applicative, en passant par les requêtes avancées, les agrégations et l'optimisation des performances.

Public cible

- Développeurs souhaitant élargir leurs compétences base de données
- Data engineers débutants
- Étudiants en informatique ou cursus data
- Toute personne ayant des bases SQL souhaitant découvrir le NoSQL

Prérequis

- Notions de bases de données relationnelles (SQL)
- Bases en Python (manipulation de dictionnaires, boucles, fonctions)
- Notions de JSON

Objectifs pédagogiques

- Comprendre les paradigmes NoSQL et savoir choisir le bon type de base selon le contexte
- Modéliser des données orientées documents avec MongoDB
- Maîtriser les opérations CRUD et les requêtes avancées
- Construire des pipelines d'agrégation pour l'analyse de données
- Optimiser les performances avec l'indexation
- Intégrer MongoDB dans une application Python

Compétences acquises

- Distinguer les familles NoSQL : document, clé-valeur, colonne, graphe
- Concevoir un schéma de documents adapté à un besoin métier
- Écrire des requêtes MongoDB (find, update, delete, opérateurs)
- Construire et optimiser des pipelines d'agrégation
- Créer et gérer des index pour améliorer les performances
- Connecter MongoDB à une application Python via PyMongo

Programme

Module 1 — Introduction au NoSQL

3h

- Limites du modèle relationnel — pourquoi le NoSQL ?
- Les 4 familles NoSQL : document (MongoDB), clé-valeur (Redis), colonne (Cassandra), graphe (Neo4j)
- Théorème CAP et compromis de cohérence
- Critères de choix d'une base NoSQL selon le contexte
- TD : analyse de cas réels — quelle base pour quel besoin ?

Module 2 — Modélisation orientée documents avec MongoDB

4h

- Structure d'un document BSON — types, imbrication, tableaux
- Embedding vs referencing — quand imbriquer, quand référencer
- Patterns de modélisation : one-to-many, many-to-many
- Comparaison avec le modèle relationnel
- TD : modélisation d'un schéma pour un cas métier concret

Module 3 — CRUD & requêtes avancées

4h

- Opérations CRUD — insertOne, find, updateMany, deleteOne
- Opérateurs de comparaison, logiques et d'array
- Projection, tri, pagination
- Requêtes sur documents imbriqués et tableaux
- TD : requêtage complet d'un jeu de données réel

Module 4 — Agrégations & pipelines

4h

- Introduction au framework d'agrégation MongoDB
- Stages essentiels : \$match, \$group, \$project, \$sort, \$lookup
- Jointures entre collections avec \$lookup
- Pipelines d'analyse de données
- TD : construction de pipelines d'agrégation sur données réelles

Module 5 — Indexation & performance

3h

- Fonctionnement des index dans MongoDB
- Types d'index : simple, composé, texte
- Analyse des requêtes avec explain()
- Anti-patterns et pièges courants
- TD : optimisation d'une collection lente — diagnostic et correction

Module 6 — Intégration Python

3h

- PyMongo — connexion, CRUD, agrégations depuis Python
- Manipulation de résultats MongoDB en Python
- Bonnes pratiques — gestion des connexions, erreurs, index
- TD : script Python complet d'interrogation et transformation de données MongoDB

Moyens pédagogiques

- Apports théoriques illustrés par des démonstrations en direct
- Exercices pratiques sur jeux de données réels
- Mise en place d'une base MongoDB et d'une intégration Python

- Support de cours remis aux participants

Évaluation

- Évaluation continue via les exercices pratiques
- Livrable final : base MongoDB modélisée, requêtée et interrogée depuis Python
- Validation des acquis par cas concret