

# FORMATION

## DATA CLOUD INGENIEUR

### Niveau :

Débutant



### Le métier :

Le data engineer est un ingénieur spécialisé dans la conception de solutions de récupération et d'exploitation de gros volumes de données ou big data.

Responsable de l'infrastructure des données, il est chargé des processus de collecte, de stockage et de modélisation des données pour que celles-ci soient faciles d'accès et prêtes à l'analyse par les data scientists et les data analysts.

### Pré-requis :

- Savoir utiliser un ordinateur
- Comprendre et parler français

Le data engineer automatise l'acquisition de données provenant de différentes sources (création des pipelines d'acquisition), les collecte et les agrège dans une data warehouse (base de données relationnelle) sur serveur ou sur le cloud (test d'intégration, mise en place et maintenance des outils d'automatisation) et gère les bases de données de stockage.

Son rôle est ensuite de nettoyer, de consolider et de structurer ces données à l'aide d'outils et d'algorithmes (création de tableaux de bord, intégration de modèles statistiques).

### Durée :

300 heures

Le data engineer doit posséder un large éventail de compétences techniques (langages de développement IA, requêtes, base de données, stockage, analyse de données, cloud computing, machine learning, deep learning, IA, DevOps..).

Pour l'industriatisation des modèles de machine learning, il doit avoir des connaissances en statistiques et mathématiques.

### Lieu :

Au centre, dans vos locaux ou en hybride

### Objectifs :

#### Analyse de Données (Data Analysis)

- Exploration de données
- Tableaux de bord et reporting
- Analyse prédictive
- Conseil en stratégie de données

#### Ingénierie des Données (Data Engineering)

- Intégration de données
- Gestion de bases de données
- Transformation des données

#### Automatisation et Optimisation

- Automatisation des processus métier
- Optimisation des opérations

#### Sécurité et Confidentialité des Données

- Protection des données
- Conformité réglementaire

### Prix :

Devis sur demande



## **Programme :**

- Introduction aux concepts de Cloud Computing et de Data Engineering
- Vue d'ensemble des services cloud (AWS / Azure / GCP)
- Configuration de l'environnement de travail
- Introduction à Python et Python pour Data Engineering
- Les principes fondamentaux de l'environnement Linux
- Les principes fondamentaux du scripting Shell
- Introduction SQL/ SQL for data engineering
- Exploration des solutions de stockage cloud et Open source
- Introduction aux écosystèmes Big Data
- Traitement des données avec Spark
- ETL dans le cloud
- Orchestration des workflows avec Airflow
- Statistiques pour data science
- Machine Learning
- Data visualization
- Principes d'Infrastructure as Code
- Conteneurisation avec Docker et Kubernetes
- Déploiement et gestion des applications avec Git / Github, FastApi, services Cloud

## **Emplois visés :**

- Data Engineer
- Data scientist
- Architecte Data

## **Rémunération :**

Un Data Engineer gagne en moyenne entre 40000€ et 50000€ par ans en France. Son salaire évolue en fonction de son niveau d'expérience et de la structure dans laquelle il exerce ses fonctions.

## **Les perspectives d'évolution :**

Après quelques années d'expérience, le Data Engineer pourra devenir Data Architecte, chief data officer.



## **Moyens pédagogiques :**

Formation délivrée en présentiel ou distanciel (classe virtuelle).

L'action de formation débutera par une présentation du formateur et de chacun des stagiaires afin de permettre une adaptation optimale de la formation.

Le formateur alterne entre méthode démonstrative, interrogative et active (via des travaux pratiques et/ou des mises en situation).

Les moyens pédagogiques mis en œuvre sont :

- Ordinateurs mis à disposition par l'organisme de formation, connexion internet libre, tableau blanc ou paperboard, vidéoprojecteur ou outils de visioconférence en cas de formation à distance.

- Logiciel de formation installé sur les postes de travail ou en ligne

- Supports de cours et exercices

En cas de formation sur site externe, le client s'assure et s'engage également à avoir toutes les ressources matérielles pédagogiques nécessaires (équipements informatiques...) au bon déroulement de l'action de formation visée conformément aux prérequis indiqués dans le programme de formation communiqué.

## **Modalités d'évaluation :**

La validation de la formation de Data Engineer s'appuie sur l'évaluation de compétences à travers :

- Des évaluations au long de votre parcours;
- La rédaction d'un dossier professionnel.

En parallèle, la formation de data engineer s'appuie sur la réalisation de projets techniques individuels ou collectifs pour s'immerger dans le monde professionnel.

## **Assiduité :**

Pour répondre aux contraintes réglementaires, chaque matin et après-midi, la signature d'une feuille d'émargement est demandée aux apprenants et formateur. Toute absence est déclarée au financeur de la formation.

La formation est dispensée de 9h à 17h avec une pause déjeuner d'une heure.