



# Formation **Data Engineer**

📍 Certification RNCP Bac +5

🕒 +5 mois de formation

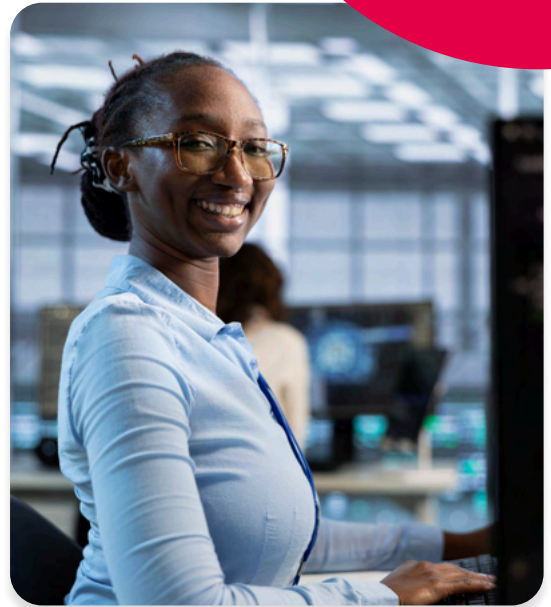
💶 100% financée

# DATA ENGINEER

## un métier émergent d'expert de la data !

L'Expert-e en ingénierie de données massives et intelligence artificielle est un ingénieur spécialisé dans la conception de solutions répondant à un besoin de stockage, d'organisation, d'accession, d'agrégation, de protection et d'analyse de données.

Ce besoin découle de la massification récente des données disponibles, phénomène désigné sous le nom de big data.



L'Expert-e en ingénierie de données massives et intelligence artificielle vient se positionner entre les administrateurs IT (technologies de l'information) qui fournissent une infrastructure IT opérationnelle et les data analysts/scientists qui vont exploiter les données.

Premier acteur du processus de traitement de la donnée, son travail intervient en amont de l'exploitation des données (data analysts/scientists), directement après que l'infrastructure technique ait été mise en œuvre par les architectes et les administrateurs.

La finalité de l'Expert-e en ingénierie de données massives et intelligence artificielle est de rendre les données faciles d'accès et prêtes à l'analyse par ses clients internes (notamment les data analysts/scientists).

### Ses principales activités :

- ✔ Définir la stratégie du projet en ingénierie de données massives et/ou IA
- ✔ Manager un projet en ingénierie de données massives et/ou IA
- ✔ Assurer la gouvernance, gestion des risques et de la conformité d'un projet en ingénierie de données massives et IA
- ✔ Concevoir et déployer des infrastructures de données
- ✔ Développer des solutions data et IA



# L'ESSENTIEL EN UN COUP D'OEIL

## 🕒 DURÉE

- 5,5 mois de formation = 805 heures de formation (35h par semaine)
- 5 mois de stage = 665 heures en entreprise (stage)

## 🎓 CERTIFICATION

Ce parcours de formation vise la certification professionnelle

**“Expert en ingénierie de données massives et intelligence artificielle”** RNCP-40167 de niveau 7, équivalent à Bac +5.



Organisme certificateur : Association SUP DE VINCI (CFA LÉONARD DE VINCI)

## ⚙️ PRÉ-REQUIS

- soit être titulaire d'un titre ou d'un diplôme de niveau 5 en informatique avec une expérience professionnelle de 2 années minimum dans le secteur des systèmes et réseaux informatiques.
- soit être titulaire d'un titre ou d'un diplôme de niveau 6 en informatique.

Cette formation s'adresse en priorité à des personnes disposant :

- de compétences en gestion de projet : identification des besoins, rédaction technique, suivi du planning, liaison client-équipes
- de compétences techniques telles que installation et gestion de services, programmation de scripts d'automatisation, conception de bases de données, maîtrise de SQL et manipulation de données pour qualification de jeux de données

Votre état d'esprit :

- Une très forte motivation, à démontrer lors de la candidature
- Curieux et créatif, avec une bonne capacité d'expression à l'oral et à l'écrit
- Prêt à travailler en équipe et à collaborer autour de projets
- La volonté de vous investir dans la durée dans un parcours de formation intense

Si vous êtes en **situation de handicap**, il est possible de mettre en place des aménagements.



# PROGRAMME DE FORMATION

## 0. INTRODUCTION À LA FORMATION

*Durant les premières semaines, découvrez votre futur métier, expérimentez la pédagogie active de Simplon en vous initiant aux principes du développement et de l'algorithme et faites connaissance avec votre promotion. Vous aurez une vision claire des compétences à acquérir et des critères d'évaluation pour l'obtention de la certification finale.*

*Durant la formation, vous appréhendez des langages de programmation comme Python, Java ou Scala, ainsi que les outils Big Data tels que Hadoop, Spark et SQL. Vous acquerez également des connaissances en architecture distribuée et en cloud computing.*

## 1. VOUS INTERVENEZ SUR DES ETL ET DES BASES DE DONNÉES ET LES FAITES ÉVOLUER

*Vous comprenez l'organisation et la modélisation des données dans une base de données, dans un Data Warehouse ou un Datamart et le fonctionnement de la mise à disposition de ces données aux utilisateurs autorisés via une API.  
Par exemple :*

- Mettre à jour la procédure d'ETL vers un Datamart
- Piloter et réaliser la mise à jour d'un contrat d'interface
- Créer une base de données Open Data à partir d'un Data Warehouse existant

## 2. VOUS INTERVENEZ SUR UNE INFRASTRUCTURE DATA ET VOUS LA FAITES ÉVOLUER

*Vous agissez sur les données stockées en ajoutant des données, par lots ou en continu, en supprimant des données ou en intégrant de nouvelles sources, mettez en place un système de surveillance de l'infrastructure, et gérez les accès à un data lake. Par exemple :*

- Intégrer de nouveaux jeux de données à un data warehouse
- Effectuer une EDA des données
- Connecter un data lake existant à de nouvelles sources de données en respectant le RGPD

## 3. VOUS CONCEVEZ ET CREZ DES ETL, BASES DE DONNÉES ET LA MISE À DISPOSITION DES DONNÉES AUX ÉQUIPES D'ANALYSE

*vous abordez les pipelines de chargement, de transformation et de stockage des données vers une plateforme cloud, mettez en place de nouveaux flux de données pour répondre à un besoin métier et optimisez les flux de transformation. Vous découvrez les fondements du Machine Learning. Par exemple :*

- Transposer des analyses et/ou visualisations dans une nouvelle plateforme cloud
- Créer des nouveaux flux de données pour une nouvelle demande métier
- Optimiser l'orchestration des flux de transformation de données sur l'ensemble de la chaîne de traitement

## 4. VOUS CONCEVEZ ET CRÉEZ UN DATA WAREHOUSE

*Vous concevez et créez un entrepôt de données cloud pour y effectuer la migration d'une solution de stockage d'origine, en cartographiant l'ensemble des sources de données et en assurant sa gestion complète, de la configuration à la sécurisation et au changement d'échelle. Par exemple :*

- Concevoir un data warehouse à partir de la cartographie des données de l'organisation
- Améliorer la qualité de données issues d'un flux de traitement existant
- Organiser et mettre en oeuvre la migration de données depuis une solution On Premise vers le Cloud
- Accompagner la réponse à un besoin Data

## 5. VOUS CONCEVEZ ET CRÉEZ UN DATA LAKE

*Vous gérez et stockez des données non structurées dans un data lake que vous concevez et mettez en place. Par exemple :*

- Concevoir et créer un data lake à l'échelle d'une organisation
- Mettre en place une démarche DataOps au sein de l'entreprise pour la gestion des flux de données
- Concevoir et mettre en place une collecte de données massives

## 6. VOUS DÉVELOPPEZ DES SOLUTIONS IA POUR L'AIDE À LA DÉCISION

*Vous travaillerez sur la dernière étape de la gestion de la donnée en étant capable de mener une analyse statistique avancée des données pour entraîner un réseau de neurones profond en s'assurant que le modèle ne comporte pas de biais et respecte le RGPD, et en proposant un tableau de bord analytique complet répondant aux besoins des équipes métier. Par exemple :*

- Réaliser une analyse statistique avancée des données
- Choisir l'algorithme adapté, entraînez des réseaux de neurones et traitez des données textuelles
- Intégrer le modèle d'IA dans un pipeline automatisé sur le Cloud
- S'assurer que l'IA ne comporte pas de biais, respectez le RGPD et justifier le ROI
- Créer un tableau de bord décisionnel intégrant des données fournies par un modèle de Deep Learning

## 7. VOUS EFFECTUEZ UN STAGE EN ENTREPRISE

*La pédagogie active promue par Simplon se traduit - entre autres - par l'implication de l'entreprise pendant la formation.*

## 8. VOUS PRÉPAREZ ET PASSEZ VOTRE TITRE PROFESSIONNEL

*Pour attester de votre maîtrise de l'ensemble des compétences visées par le titre professionnel, vous réalisez un projet que vous présentez devant un jury de professionnels.*



## Pourquoi choisir l'expérience Simplon ?

Que vous soyez en reconversion, étudiant ou déjà professionnel de la Tech, Simplon vous donne toutes les clés pour réussir :

- Des formations qui ciblent des compétences recherchées par les entreprises sur des métiers porteurs.
- Une pédagogie active qui vous prépare au monde professionnel avec des mises en situation concrètes et des projets réalisés en équipe.
- Des compétences validées par des titres professionnels et des certifications reconnus par France Compétences.
- Un accompagnement personnalisé pour valoriser votre profil et donner un nouvel élan à votre carrière.

### Pour en savoir plus

Pour découvrir nos formations, notre actualité  
ou pour toute information, rdv ici ↓

[www.simplon.co](http://www.simplon.co)