



### Contact

Christophe DEDOUCHE – Responsable de la Formation Professionnelle

[pro@ecam-strasbourg.eu](mailto:pro@ecam-strasbourg.eu)

Cédric Bobenrieth – Responsable pédagogique de la formation

[cedric.bobenrieth@ecam-strasbourg.eu](mailto:cedric.bobenrieth@ecam-strasbourg.eu)



### Présentation générale

La formation 'Data Scientist' prépare les stagiaires à la certification 'Analyser les données: intégration, exploitation, visualisation, industrialisation', qui valide les quatre domaines de compétences nécessaires à la mise en œuvre d'un projet d'analyse de données de bout en bout :

- Intégrer les données
- Exploiter les données
- Visualiser les données
- Industrialiser le traitement des données



### Blocs de compétences et capacités professionnelles

- **Bloc de compétence n°1 : Intégrer les données**
  - Capacité C1 : Connaître les enjeux de la maîtrise de la donnée pour les entreprises.
  - Capacité C2 : Connaître les principes de fonctionnement des bases de données relationnelles et être capable de programmer des requêtes SQL complexes
  - Capacité C3 : Savoir extraire et transformer des données
  - Capacité C4 : Être capable de mener un projet de développement dans un langage orienté objet de bout en bout
- **Bloc de compétence n°2 : Exploiter les données**
  - Capacité C5 : Savoir programmer une application simple en Python
  - Capacité C6 : Connaître les modèles statistiques mis en œuvre dans le Big Data afin de pouvoir choisir le modèle adapté aux besoins d'un projet. Savoir manipuler les algorithmes associés à ces modèles.
  - Capacité C7 : Connaître les principes théoriques du machine learning, de l'optimisation et de la détection d'anomalie afin de savoir choisir, développer et paramétrer un algorithme en fonction du besoin de traitement des données.
  - Capacité C8 : Savoir entraîner un modèle prédictif linéaire et explorer les données avec des algorithmes non supervisés.
- **Bloc de compétence n°3 : Visualiser les données**
  - Capacité C9 : Connaître les bases permettant de réaliser une application web.
  - Capacité C10 : Savoir manipuler un outil représentatif de la visualisation pour le big data.
  - Capacité C11 : Concevoir les fonctionnalités d'une application exploitant une plateforme d'analyse de données (assistants virtuels, chatbot, outil d'analyse de mails)
- **Bloc de compétence n°4 : Industrialiser le traitement des données**
  - Capacité C12 : Savoir mettre en œuvre une politique Informatique et Liberté pour des applications de traitement de données
  - Capacité C13 : Connaître et savoir choisir l'outil adapté à la gestion d'un projet data.
  - Capacité C14 : Savoir planifier un projet agile et organiser une équipe agile
  - Capacité C15 : Utiliser des APIs d'une plateforme représentative



### Objectifs pédagogiques de la formation

- Permettre une reconversion vers les métiers du numérique
- Permettre aux stagiaires de maîtriser la chaîne de valeur de la donnée depuis l'intégration jusqu'à sa visualisation, en passant par l'analyse
- Donner aux stagiaires les compétences leur permettant d'être opérationnels dans la mise en œuvre de projets d'analyse de données : gestion de projet, technologies logicielles, méthodologie d'analyse des données



### Modalités pratiques

- **Public** : Demandeurs d'emplois ayant une formation scientifique poussée (au minimum niveau Bac +3 avec de bonnes connaissances en mathématiques et statistiques, Bac+5 à Bac+8 en sciences, physique, biologie).
- **Prérequis** : Compétences en mathématiques, statistiques, logique, bases de données.
- **Conditions d'accès** : Recrutement en 3 étapes :
  - test d'évaluation des compétences en mathématiques, statistiques, logique, bases de données,
  - entretien de motivation professionnel
  - entretien de recrutement par l'entreprise accueillant le stagiaire durant la période en entreprise
- **Conditions de démarrage** : Nombre de participants de 8 personnes (minimum) à 15 personnes (maximum) pour l'ensemble de la formation
- **Lieu** : ECAM Strasbourg Europe, 2 Rue de Madrid, 67300 SCHILTIGHEIM
- **Nombre d'heures** : 798 heures réparties de la manière suivante :
  - 518 heures en centre de formation
  - 280 heures en Entreprise
- **Horaires** : 9h–12h30 et 13h30–17h00 (accueil à partir de 8h30)
  
- **Type de formation** : Formation en présentiel
- **Date début de formation** : 18 octobre 2021
- **Validation** : Chaque module de la formation ne peut être validé que si la présence du stagiaire est attestée sur l'ensemble des sessions de formation du module.  
La formation ne peut être validée que si l'ensemble des modules sont validés.  
L'atteinte des compétences de la formation est attestée par :
  - La présence à l'ensemble des sessions de formation
  - 1 soutenance de projet pour les modules suivants
    - 'Intégrer les données'
    - 'Analyser les données'
    - 'Visualiser les données'
  - 1 évaluation écrite pour le module 'Industrialiser le traitement des données'
  - 1 soutenance de présentation de la mission réalisée pendant la séquence en entreprise



### Moyens pédagogiques

- **Alternance de théorie, travaux pratiques et projets d'études.**
- **Accompagnement** : Un module 'Techniques de Recherche d'Emploi' (TRE) vous permet de développer vos capacités professionnelles et de vous préparer à valoriser les compétences acquises durant la formation. Un accompagnement personnalisé vous est proposé, dont les modalités seront précisées par le formateur lors de la première séance du module.



### Programme détaillé

- **Intégrer les données (147 heures)**
  - **Introduction à la data science (7 heures)**
    - Connaître les enjeux de la maîtrise de la donnée pour les entreprises.
    - Avoir des notions des spécificités du marché de l'analyse de données.
    - Connaître les spécificités des projets données en entreprise.
  - **Programmation SQL (35 heures)**
    - Connaître les principes de fonctionnement des bases de données relationnelles
    - Être capable de programmer des requêtes SQL complexes
  - **Outils pour le traitement des données (14 heures)**
    - Connaître les modèles de données en vue de leur traitement
    - Savoir préparer des données pour le traitement
    - Savoir extraire et transformer des données

- **Programmation Objet (28 heures)**
  - Être familier avec les concepts de la programmation objet
  - Savoir programmer une application dans un langage orienté objet
  - Connaître et comprendre les technologies récentes pour la création d'API Web (ex : Spring Boot)
  - Être capable de mener un projet de développement de bout en bout
  - Connaître les facteurs impactant la performance des applications
- **Bases de données avancées (FOAD, 35 heures)**
  - Algèbre relationnelle
  - Bases de données NoSQL
- **Projet « Intégration des données » (28 heures)**
- **Exploiter et analyser les données (147 heures)**
  - **Python pour la science des données (21 heures)**
    - Savoir programmer une application simple en python
    - Savoir mettre en œuvre de la programmation script, procédurale, objet, en python
    - Connaître les structures de données spécifiques au langage
  - **Statistiques pour le big data (28 heures)**
    - Connaître les modèles statistiques mis en œuvre dans le big data
    - Connaître et savoir manipuler les algorithmes associés à ces modèles
    - Savoir choisir un modèle statistique adapté aux besoins d'un projet big data
    - Être capable de mener un projet de mise en œuvre des statistiques pour le big data
  - **Apprentissage automatique (28 heures)**
    - Connaître les principes théoriques du machine learning, de l'optimisation, de la détection d'anomalie
    - Savoir choisir un algorithme en fonction du besoin de traitement des données
    - Savoir développer et paramétrer ses propres algorithmes d'analyse de données
  - **Analyse de données avancée (FOAD, 35 heures)**
    - Savoir choisir un algorithme de machine learning
    - Savoir entraîner un modèle prédictif linéaire
    - Explorer les données avec des algorithmes non supervisés
  - **Projet « Exploiter et analyser les données » (35 heures)**
- **Visualiser les données (126 heures)**
  - **Technologies Web (35 heures)**
    - Connaître les bases permettant de réaliser une application web
    - Connaître et comprendre les technologies récentes pour la création d'application Web (ex : Angular)
    - Être capable de mener un projet web de bout en bout
  - **Visualisation des données (35 heures)**
    - Connaître les particularités des outils de visualisation pour le big data
    - Savoir manipuler un outil représentatif de la visualisation pour le big data
    - Savoir mettre en œuvre une démarche de data story telling
    - Être capable de mener un projet de mise en œuvre de la visualisation
    - Être capable de convaincre avec les données et la visualisation
  - **PowerBI (21 heures)**
    - Concevoir des rapports permettant de mettre en avant les Indicateurs pertinents
  - **Projet « Visualiser les données » (35 heures)**

- **Industrialiser le traitement des données (77 heures)**
  - **Informatique et liberté (14 heures)**
    - Connaître les principes de la protection des données privées
    - Savoir mettre en œuvre une politique Informatique et Liberté pour des applications de traitement de données
  - **Gestion de projet Big Data (21 heures)**
    - Connaître les outils de gestion de projet data
    - Savoir choisir l'outil adapté selon le type de projet
  - **Gestion d'un projet informatique (FOAD, 7 heures)**
    - Savoir organiser une équipe agile
    - Planifier un projet agile
    - Choisir les outils de gestion de projet adaptés à l'agilité
  - **Modèles de données intégrés pour l'analyse de données et solutions interactives pour l'analyse de données (35 heures)**
    - Concevoir les fonctionnalités d'une application exploitant une plate-forme d'analyse de données (assistants virtuels, chatbot, outil d'analyse de mails)
    - Savoir concevoir une architecture logicielle d'analyse de données avec des plates-formes intégrées
- **Techniques de Recherche d'Emploi (14 heures)**
- **Stage en Entreprise (280 heures)**