



Contact

Responsable de la Formation Professionnelle

pro.strasbourg-europe@icam.fr

03 90 40 09 63



Accessibilité universelle

Formation accessible aux personnes en situation de handicap

Pour tout renseignement, contacter referent-handicap.strasbourg-europe@icam.fr



Présentation générale

La formation 'Data Scientist : de l'intégration à l'industrialisation des données' prépare les stagiaires à la certification validant les quatre domaines de compétences nécessaires à la mise en œuvre d'un projet d'analyse de données de bout en bout :

- Intégration des données provenant de différentes sources dans une vue unifiée pour rendre les données plus exploitables
- Exploitation et analyse des données pour obtenir des informations permettant de répondre à une problématique donnée
- Visualisation des données pour présenter des informations pertinentes
- Industrialisation du traitement des données



Suites possibles après cette formation

- **Métiers** : Reconversion dans le domaine du numérique, et plus spécifiquement dans les métiers de l'analyse de données. L'analyse de données a pour objectif de permettre aux entreprises et organisations de valoriser leur capital 'données', et de les administrer, les analyser et les visualiser : Data Scientist, Data Analyst, Data Miner.
- **Formations** : Possibilités de suivre des formations spécialisées afin d'approfondir ces connaissances dans des domaines spécifiques tel que le Report BI, le Machine Learning, le Deep Learning, le Data Mining et le Data Engineering



Blocs de compétences et capacités professionnelles

- **Bloc de compétence n°1 : Intégration des données provenant de différentes sources dans une vue unifiée pour rendre les données plus exploitables**
 - Capacité C1 : Identifier les enjeux de la maîtrise de la donnée pour les entreprises afin de déterminer les spécificités des projets
 - Capacité C2 : Mettre en œuvre les outils d'extraction et de transformation des données pour préparer leur traitement
 - Capacité C3 : Concevoir une base de données SQL et les requêtes nécessaires pour récupérer, stocker et exploiter les données



« Opération financée par la Région Grand-Est »

- **Bloc de compétence n°2 : Exploitation et analyse des données pour obtenir des informations permettant de répondre à une problématique donnée**
 - Capacité C4 : Implémenter des applications simples en Python pour réaliser des scripts d'analyse de données
 - Capacité C5 : Mener un projet de mise en œuvre des statistiques pour le big data afin d'obtenir une analyse pertinente des données
 - Capacité C6 : Mettre en œuvre les outils d'apprentissage automatique et d'intelligence artificielle en fonction du besoin de traitement des données pour répondre à une problématique précise
 -
- **Bloc de compétence n°3 : Visualisation des données pour présenter des informations pertinentes**
 - Capacité C7 : Concevoir une application Web pour présenter visuellement les données
 - Capacité C8 : Mettre en œuvre des outils de visualisation pour le Big Data pour permettre du data story telling
 - Capacité C9 : Concevoir des rapports PowerBI pour mettre en avant les indicateurs pertinents
- **Bloc de compétence n°4 : Industrialisation du traitement des données**
 - Capacité C10 : Auditer les solutions Big Data mises en place pour s'assurer du respect des principes de la loi « informatique et liberté » N78-17 du 6 janvier 1978 dans sa récente rédaction
 - Capacité C11 : Gérer un projet Big Data pour s'assurer de son bon déroulement
 - Capacité C12 : Concevoir les fonctionnalités d'une application exploitant les modèles de données intégrés et solutions interactives pour l'analyse de données afin de les intégrer à un environnement industriel

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



« Opération financée par la Région Grand-Est »



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** : Icam site de Strasbourg Europe, 2 Rue de Madrid, 67300 SCHILTIGHEIM
- **Nombre d'heures** : 1200 heures réparties de la manière suivante :
 - 600 heures en centre de formation
 - 600 heures en Entreprise
- **Horaires** : 9h–12h30 et 13h30–17h00 (accueil à partir de 8h30)
- **Type de formation** : Formation en présentiel
- **Date début de formation** : 2 Octobre 2023
- **Examen de validation des compétences** : 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 en groupe et une évaluation écrite individuelle 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 projet « Data Scientist : De l'intégration à l'industrialisation » individuelle
 - 1 soutenance de présentation de la mission réalisée pendant la séquence en entreprise
- **Validation** : Sous réserve de la réussite aux examens, attestation de compétences délivrée par un Jury composé à 50% de professionnels issus du secteur de l'informatique et Certification professionnelle RNCP (en cours d'inscription).



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.



« Opération financée par la Région Grand-Est »



Programme détaillé

- **Intégrer les données (148 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 » (29 heures)**
- **Exploiter et analyser les données (148 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



« Opération financée par la Région Grand-Est »

- **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 » (36 heures)**
- **Visualiser les données (127 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 » (36 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)**
- **Projet 'Data Scientist: De l'intégration à l'industrialisation' (70 heures)**
- **Stage en Entreprise (600 heures)**



« Opération financée par la Région Grand-Est »