

Concepteur développeur en science des données

RNCP35288 · RNCP · Publiée

L'essentiel

Niveau de qualification	Niveau 6
Code(s) NSF	326, 114g
Formacode(s)	11052, 11036, 31028, 31026
Code(s) ROME	M1403, M1810, M1805, M1803, M1806
Type d'enregistrement	Enregistrement sur demande
Date d'enregistrement	10/02/2021
Date d'échéance	10/02/2026

Certificateur(s)

JEDHA · SIRET : 83872651100044

JEDHA · SIRET : 83872651100036

<https://www.jedha.co/>

JEDHA · SIRET : 83872651100028

Partenaires habilités

11 organismes habilités, d'après la liste publiée par France compétences : 10 à former et organiser l'évaluation et 1 à former.

Organisme	SIRET	Habilitation
26 ACADEMY	83069794200027	Former et organiser l'évaluation
ALBERT GLOBAL DATA SCHOOL	90517525300046	Former et organiser l'évaluation
IFPASS SERVICES	82379935800028	Former
JEDHA	83872651100010	Former et organiser l'évaluation
JEDHA	83872651100036	Former et organiser l'évaluation
JEDHA	83872651100044	Former et organiser l'évaluation
M2I	33354415300344	Former et organiser l'évaluation
M2I SCRIBTEL	39336786700214	Former et organiser l'évaluation
NEOMA BUSINESS SCHOOL	83429535400015	Former et organiser l'évaluation
NEOMA BUSINESS SCHOOL	83429535400049	Former et organiser l'évaluation

Liste complète des 11 partenaires habilités : <https://data.aktaconseil.fr/certifications/rncp35288-concepteur-developpeur-en-science-des-donnees/partenaires>

Résumé de la certification

OBJECTIFS ET CONTEXTE

La Science des Données (ou Data Science en anglais) est devenue au cours de ces dernières années un élément indispensable pour permettre aux entreprises d'innover et de se différencier, et l'Intelligence Artificielle un outil de

performance indispensable. Les directions métiers tous secteurs confondus utilisent déjà ces dernières, et cette tendance est amenée à augmenter exponentiellement au cours des prochaines années. L'objectif visé de la certification de "Concepteur Développeur en Science des Données " est de devenir un spécialiste en science des données ayant la capacité d'avoir une vision globale des projets de Science de Données, capable de les concevoir et diriger dans leur ensemble comme d'intervenir à un point précis de ces derniers.

En effet, la certification confère au candidat les compétences pour exercer le métier : l'ensemble de ces compétences, décrites dans le référentiel, lui permet de créer des infrastructures de gestion de données robustes et adaptées, de les alimenter, de développer des algorithmes d'intelligence artificielle, de les mettre en production, mais aussi de collaborer avec les différentes équipes-métiers d'une organisation pour évaluer et adapter les besoins data. Il lui est donc possible de diriger des projets de gestion de données de bout-en-bout, d'en rendre compte, d'être force de proposition et de les adapter aux besoins de son environnement professionnel.

Le candidat certifié peut mettre en œuvre l'ensemble des compétences nécessaires pour être opérationnel rapidement.

ACTIVITÉS VISÉES

A1.

Construction d'une infrastructure de gestion de données adaptée à l'organisation

A2. Collecte de données

A3. Gestion d'entrepôts de données (Data Warehouse)

A4. Analyse exploratoire de données

A5. Visualisation et présentation de données

A6. Mise en place d'un algorithme d'apprentissage automatique

A7. Segmentation et réduction de base de données

A8. Optimisation des performances des algorithmes d'apprentissage automatique

A9. Mise en place d'un apprentissage automatique profond

A10. Optimisation des performances des algorithmes d'apprentissage automatique profond pour industrialisation

A11. Industrialisation d'algorithmes d'apprentissage automatique

A12. Production d'applications d'intelligence artificielle utilisables par toutes les équipes métier

A13. Définition d'une modélisation statistique qui permette de répondre aux problématiques des directions métiers

A14. Construction d'un système de gestion et de suivi de projet d'analyse et de gestion de données

A15. Direction de projets de gestion de données

COMPÉTENCES ATTESTÉES

Bloc n°1 - Construction et alimentation d'une infrastructure de gestion de données

C1.1 - Concevoir une architecture de

données robuste et adaptée en créant des lacs de données (Data Lake en anglais) et des entrepôts de données (Data Warehouse en anglais) afin de répondre aux besoins de stockage, d'utilisation, de sécurité et de protection de l'organisation définie par un cahier des charges

C1.2 - Intégrer la dimension de stockage et de calcul distribuée à

l'infrastructure de données via l'utilisation d'outils comme Spark ou AWS

Redshift afin de l'adapter à des besoins de gestion de données massives (Big Data en anglais)

C1.3 - Collecter des données

provenant de différentes sources (Web, Logiciels internes de type Sage / Excel ou externes de type Google Analytics) via des

bibliothèques de programmation de type Scrapy ou BeautifulSoup

dans le respect des normes de protection des données utilisateurs définies dans le RGPD pour alimenter le Data Lake afin d'affiner le résultat d'analyses futures.

C1.4 - Nettoyer et organiser les

données dans l'entrepôt de données (Data Warehouse en anglais) en écrivant des processus d'extraction, transformation et chargements (ETL en anglais) afin de rendre ces données disponibles et compréhensibles pour les autres équipes métiers.

Bloc n°2 - Analyse exploratoire, descriptive et inférentielle de données

C2.1 - Traiter des bases de données grâce à des analyses statistiques descriptives et inférentielles via des bibliothèques de programmation comme Numpy ou Pandas, pour les organiser et les nettoyer afin de les normaliser par rapport à la population étudiée.

C2.2 - Effectuer des analyses univariées et multivariées sur des bases de données structurées afin de préciser des relations entre plusieurs variables et d'établir des liens statistiques entre elles.

C2.3 - Optimiser les analyses statistiques grâce au traitement parallélisé via l'utilisation d'outils comme Spark pour accélérer le temps de calcul d'un ordinateur afin de pouvoir analyser des volumes de données massifs (Big Data)

C2.4 - Présenter le résultat d'une analyse statistique de données structurées, massives ou non, grâce à des bibliothèques de programmation comme Plotly ou Matplotlib pour synthétiser ce résultat devant un public profane afin de faciliter la prise de décisions et appuyer leurs déclinaisons opérationnelles

Bloc n°3 - Analyse prédictive de données structurées par l'intelligence artificielle

C3.1 - Traiter des données

structurées en créant un pipeline de traitement grâce à des bibliothèques de programmation comme Scikit-Learn pour encoder, normaliser et découper des données afin de les rendre interprétables par un algorithme d'apprentissage automatique (Machine Learning en anglais)

C3.2 - Effectuer des analyses prédictives sur un jeu de données structurées grâce à des algorithmes d'apprentissage automatique supervisés adaptés afin d'automatiser des tâches liées aux résultats des prédictions de ces algorithmes

C3.3 - Élaborer un algorithme d'apprentissage automatique non-supervisé pour segmenter une base de données en différents groupes homogènes ou réduire la dimension de cette dernière afin de pouvoir comprendre des observations de manière granulaire et de permettre leur visualisation

C3.4 - Évaluer la performance prédictive des algorithmes d'apprentissage automatique en déterminant l'influence des différentes variables pour pouvoir l'améliorer afin de démontrer son utilité aux directions métiers, par rapport aux processus déjà établis dans l'organisation

Bloc n°4 - Analyse prédictive de données non-structurées par l'intelligence artificielle

C4.1 - Traiter des données non-structurées (image, texte, audio) par la création de fonction de traitements via l'utilisation de bibliothèques de programmation comme TensorFlow ou Numpy pour les transformer en matrices afin de les rendre interprétables par un algorithme d'apprentissage automatique profond (Deep learning en anglais)

C4.2 - Élaborer des réseaux de neurones adaptés (classiques, convolutifs ou récurrents) en superposant des couches neuronales via des bibliothèques de programmation comme TensorFlow pour analyser des données non-structurées afin de détecter des signaux sur ces dernières

C4.3 - Créer un algorithme robuste et précis en configurant un réseau de neurones pré-entraîné profond afin de répondre à des problématiques de prédiction sur des volumes de données massifs

C4.4 - Créer des données non-structurées en élaborant des réseaux de neurones adverses afin de construire de nouvelles bases d'entraînement pour des applications d'intelligence artificielle

C4.5 - Évaluer la performance d'un algorithme d'apprentissage automatique profond en évaluant des indicateurs sur des données d'entraînement et de validation afin d'industrialiser son utilisation

Bloc n°5 - Industrialisation d'un algorithme d'apprentissage automatique et automatisation des processus de décision

C5.1 - Standardiser la construction et l'environnement informatique d'un algorithme d'apprentissage automatique grâce des outils de production comme MLflow et Docker afin de faciliter la mise en production de projets d'intelligence artificielle sur tous types de plateformes

C5.2 - Créer une interface de programmation applicative grâce à des outils comme AWS sagemaker afin de donner un accès à échelle aux prédictions des algorithmes d'apprentissage automatique à l'ensemble des équipes métiers concernées

C5.3 - Déployer une application web intégrant des algorithmes de statistiques prédictives (Machine Learning et Deep Learning) grâce à des outils comme Flask, Heroku ou AWS sagemaker pour les rendre utilisables par l'ensemble des

équipes métiers afin d'automatiser leurs processus de décision

Bloc n°6 - Direction de projets de gestion de données

C6.1 - Traduire les enjeux métiers en problématiques mathématiques/data grâce à une compréhension des besoins propres à chaque projet data afin de pouvoir répondre aux objectifs de l'organisation

C6.2 - Maîtriser les technologies les plus récentes et adaptées du marché grâce à de la veille technologique et de la pratique constante pour développer une expertise afin d'être à même de proposer aux directions métiers les solutions les plus adaptées actuellement à une problématique et l'amélioration constante des process de gestion de données déjà en place

C6.3 - Définir un cahier des charges, un retroplanning et un budget afin de défendre et détailler aux directions métier un projet data répondant aux besoins de l'organisation

C6.4 - Gérer un projet d'analyse et de gestion de données (analyse statistique descriptive, Machine Learning, Deep Learning, Big Data ou non) grâce à l'élaboration d'indicateurs adaptés et de tableaux de bords, afin de faire le suivi et le bilan de l'action, ainsi que de la déclinaison opérationnelle de ses résultats, le tout dans le respect des normes de protection des données utilisateurs définies dans le RGPD

C6.5 - Transmettre aux directions-métiers le process d'extraction d'informations et d'analyse de données en le vulgarisant afin de soutenir la mise en place d'une stratégie et d'actions futures.

C6.6 - Diriger un projet de gestion de données, allant de sa conception à la mise en place de solutions, afin de le mener jusqu'à son terme, d'être la personne clé disposant de toutes les informations sur le projet à tout moment, et d'accompagner d'autres services de l'organisation dans l'ensemble des activités relatives à celui-ci

Blocs de compétences

RNCP35288BC01 — Construction et alimentation d'une infrastructure de gestion de données

LISTE DE COMPÉTENCES

C1.1 - Concevoir une architecture de données robuste et adaptée en créant des lacs de données (Data Lake en anglais) et des entrepôts de données (Data Warehouse en anglais) afin de répondre aux besoins de stockage, d'utilisation, de sécurité et de protection de l'organisation définie par un cahier des charges

C1.2 - Intégrer la dimension de stockage et de calcul distribuée à l'infrastructure de données via l'utilisation d'outils comme Spark ou AWS Redshift afin de l'adapter à des besoins de gestion de données massives (Big Data en anglais)

C1.3 - Collecter des données provenant

de différentes sources (Web, Logiciels internes de type Sage / Excel

ou externes de type Google Analytics) via des bibliothèques de

programmation de type Scrapy ou BeautifulSoup

dans le respect des normes de protection des données utilisateurs définies dans le RGPD pour alimenter le Data Lake afin d'affiner le résultat d'analyses futures.

C1.4 - Nettoyer et organiser les données dans l'entrepôt de données (Data Warehouse en anglais) en écrivant des processus d'extraction, transformation et chargements (ETL en anglais) afin de rendre ces données disponibles et compréhensibles pour les autres équipes métiers.

MODALITÉS D'ÉVALUATION

Type d'évaluation : Une étude de cas sur des données réelles

Thème d'évaluation : Construction d'une infrastructure Cloud accueillant des données Big Data

RNCP35288BC02 — Analyse exploratoire, descriptive et inférentielle de données

LISTE DE COMPÉTENCES

C2.1 - Traiter des bases de données grâce à des analyses statistiques descriptives et inférentielles via des bibliothèques de programmation comme Numpy ou Pandas, pour les organiser et les nettoyer afin de les normaliser par rapport à la population étudiée.

C2.2 - Effectuer des analyses univariées et multivariées sur des bases de données structurées afin de préciser des relations entre plusieurs variables et d'établir des liens statistiques entre elles.

C2.3 - Optimiser les analyses statistiques grâce au traitement parallélisé via l'utilisation d'outils comme Spark pour accélérer le temps de calcul d'un ordinateur afin de pouvoir analyser des volumes de données massifs (Big Data)

C2.4 - Présenter le résultat d'une analyse statistique de données structurées, massives ou non, grâce à des bibliothèques de programmation comme Plotly ou Matplotlib pour synthétiser ce résultat devant un public profane afin de faciliter la

prise de décisions et appuyer leurs déclinaisons opérationnelles.

MODALITÉS D'ÉVALUATION

Type d'évaluation : Deux études de cas sur des données réelles

Thème d'évaluation :

- Gestion de valeurs manquantes et aberrantes d'une base de données non-massives puis analyse pour déterminer et présenter des tendances par le biais de graphiques.
- Analyse d'une base de données massives déstructurées adaptée à une problématique définie.

RNCP35288BC03 — Analyse prédictive de données structurées par l'intelligence artificielle

LISTE DE COMPÉTENCES

C3.1 - Traiter des données structurées en créant un pipeline de traitement grâce à des bibliothèques de programmation comme Scikit-Learn pour encoder, normaliser et découper des données afin de les rendre interprétables par un algorithme d'apprentissage automatique (Machine Learning en anglais)

C3.2 - Effectuer des analyses prédictives sur un jeu de données structurées grâce à des algorithmes d'apprentissage automatique supervisés adaptés afin d'automatiser des tâches liées aux résultats des prédictions de ces algorithmes

C3.3 - Élaborer un algorithme d'apprentissage automatique non-supervisé pour segmenter une base de données en différents groupes homogènes ou réduire la dimension de cette dernière afin de pouvoir comprendre des observations de manière granulaire et de permettre leur visualisation.

C3.4 - Évaluer la performance prédictive des algorithmes d'apprentissage automatique en déterminant l'influence des différentes variables pour pouvoir l'améliorer afin de démontrer son utilité aux directions métiers, par rapport aux processus déjà établis dans l'organisation

MODALITÉS D'ÉVALUATION

Type d'évaluation : trois études de cas pratiques tirées de cas réels

Thème d'évaluation

- Optimisation des processus par le biais d'algorithmes d'apprentissage supervisés
- Optimisation d'algorithmes d'apprentissage automatique supervisé sur des bases de données déséquilibrées
- Localisation de zones de densité géographique par l'élaboration d'algorithmes d'apprentissage automatique non-supervisé

RNCP35288BC04 — Analyse prédictive de données non-structurées par l'intelligence artificielle

LISTE DE COMPÉTENCES

C4.1 - Traiter des données non-structurées (image, texte, audio) par la création de fonction de traitements via l'utilisation de bibliothèques de programmation comme TensorFlow ou Numpy pour les transformer en matrices afin de les rendre interprétables par un algorithme d'apprentissage automatique profond (Deep learning en anglais)

C4.2 - Élaborer des réseaux de neurones adaptés (classiques, convolutifs ou récurrents) en superposant des couches neuronales via des bibliothèques de programmation comme TensorFlow pour analyser des données non-structurées afin de détecter des signaux sur ces dernières

C4.3 - Créer un algorithme robuste et précis en configurant un réseau de neurones pré-entraîné profond afin de répondre à des problématiques de prédiction sur des volumes de données massifs

C4.4 - Créer des données non-structurées en élaborant des réseaux de neurones adverses afin de construire de nouvelles bases d'entraînement pour des applications d'intelligence artificielle

C4.5 - Évaluer la performance d'un

algorithme d'apprentissage automatique profond en évaluant des indicateurs sur des données d'entraînement et de validation afin d'industrialiser son utilisation

MODALITÉS D'ÉVALUATION

Type d'évaluation : une étude de cas pratique sur des données non-structurées

Thème d'évaluation : Analyse de sentiment, par l'élaboration d'un algorithme permettant de déterminer le sentiment d'un utilisateur à l'égard d'un produit

RNCP35288BC05 — Industrialisation d'un algorithme d'apprentissage automatique et automatisation des processus de décision

LISTE DE COMPÉTENCES

C5.1 - Standardiser la construction et l'environnement informatique d'un algorithme d'apprentissage automatique grâce des outils de production comme MLflow et Docker afin de faciliter la mise en production de projets d'intelligence

artificielle sur tous types de plateformes

C5.2 - Créer une interface de programmation applicative grâce à des outils comme AWS Sagemaker afin de donner un accès à l'échelle aux prédictions des algorithmes d'apprentissage automatique à l'ensemble des équipes métiers concernées

C5.3 - Déployer une application web intégrant des algorithmes de statistiques prédictives (Machine Learning et Deep Learning) grâce à des outils comme Flask, Heroku ou AWS Sagemaker pour les rendre utilisables par l'ensemble des équipes métiers afin d'automatiser leurs processus de décision

MODALITÉS D'ÉVALUATION

Type d'évaluation : Étude de cas pratique sur le déploiement d'un algorithme d'apprentissage automatique

Thème d'évaluation : construction et mise en production d'une application web d'intelligence artificielle

RNCP35288BC06 — Direction de projets de gestion de données

LISTE DE COMPÉTENCES

C6.1 - Traduire les enjeux métiers en

problématiques mathématiques/data grâce à une compréhension des besoins propres à chaque projet data afin de pouvoir répondre aux objectifs de l'organisation

C6.2 - Maîtriser les technologies les plus récentes et adaptées du marché grâce à de la veille technologique et de la pratique constante pour développer une expertise afin d'être à même de proposer aux directions métiers les solutions les plus adaptées actuellement à une problématique et l'amélioration constante des processus de gestion de données déjà en place

C6.3 - Définir un cahier des charges, un retroplanning et un budget afin de défendre et détailler aux directions métier un projet data répondant aux besoins de l'organisation

C6.4 - Gérer un projet d'analyse et de gestion de données (analyse statistique descriptive, Machine Learning, Deep Learning, Big Data ou non) grâce à l'élaboration d'indicateurs adaptés et de tableaux de bords, afin de faire le suivi et le bilan de l'action, ainsi que de la déclinaison opérationnelle de ses résultats, le tout dans le respect des normes de protection des données utilisateurs définies dans le RGPD

C6.5 - Transmettre aux

directions-métiers le processus d'extraction d'informations et d'analyse de données en le vulgarisant afin de soutenir la mise en place d'une stratégie et d'actions futures.

C6.6 - Diriger un projet de gestion de données, allant de sa conception à la mise en place de solutions, afin de le mener jusqu'à son terme, d'être la personne clé disposant de toutes les informations sur le projet à tout moment, et d'accompagner d'autres services de l'organisation dans l'ensemble des activités relatives à celui-ci

MODALITÉS D'ÉVALUATION

Type d'évaluation : projet data conçu de A à Z.

Thème d'évaluation : libre. Les apprenants peuvent préparer le projet data de leur choix. Celui-ci peut être personnel, développé par le candidat dans le cadre de son activité professionnelle, ou défini par une entreprise partenaire. Il fera l'objet d'une soutenance orale de 10 minutes suivie de 5 à 10 minutes de questions.

Secteur d'activité et type d'emploi

SECTEURS D'ACTIVITÉS

Depuis l'émergence de l'Intelligence Artificielle et des Big Data, les Concepteurs Développeurs en Science des Données sont recrutés dans de différents et nombreux domaines (commercial, finance, santé, recherche, administration...), à des postes pouvant différer par leur niveau hiérarchique et leur niveau de technicité, et par une grande variété d'organisations : dans l'industrie, par de grandes entreprises comme de jeunes start-up, dans le service public et l'administration, par des institutions scientifiques, des organisations médicales ou paramédicales et même des structures associatives. En effet, leur travail d'analyse peut facilement se retrouver à la base de l'élaboration de la stratégie générale d'une organisation.

Tous les secteurs d'activités utilisant de la donnée sont ainsi concernés et quel que soit la taille de l'entreprise. Le "Concepteur Développeur en Science des Données" pourra exercer son activité :

- Dans une entreprise de Services Numériques (ESN ou ex SSII) réalisant des prestations de science des données ou d'intelligence artificielle, en régie ou au forfait
- Dans l'ensemble des structures utilisatrices d'Intelligence Artificielle, possédant un service dédié à l'analyse de données et aux développements de ces IA, du secteur privé ou public, entreprises ou association (ex : Banque, Assurance, startups spécialisées dans l'Intelligence Artificielle, certains établissements de recherche ou de santé, etc)
- En tant qu'indépendant pour des missions de gestion de données, de conceptions ou mise en production d'algorithmes, ou de pilotage de projet Data.
- Dans des organismes de formation spécialisés dans l'Intelligence Artificielle.

TYPE D'EMPLOIS ACCESSIBLES

- Data Scientist
- Data Engineer
- Machine Learning Engineer
- Développeur Full-Stack Python (ou Software Engineer)
- Data Analyst (ou Data Miner)
- Business Analyst
- Consultant Intelligence Artificielle (ou Consultant Data)
- Product Manager Data (ou Product Owner ou chef de produit Data)
- Chef de Projet Data (ou Project Manager Data)
- Ingénieur d'études / Chargé d'études
- Ingénieur en Business Intelligence
- Directeur / Directrice de l'Organisation et des Systèmes d'Information -DOSI
- Attaché / Attachée d'études statistiques

RÉGLEMENTATIONS D'ACTIVITÉ

L'ensemble du processus de conception et de développement d'algorithmes d'apprentissage automatique doit se dérouler dans le respect des réglementations en vigueur quant à l'utilisation de données personnelles (RGPD) et des critères définissant l'utilisation responsable des données.

Le règlement général sur la protection des données (RGPD, ou encore GDPR, de l'anglais General Data Protection Regulation) est le règlement de l'Union Européenne qui constitue le texte de référence en matière de protection des données à caractère personnel : il renforce la protection de ces données pour les individus, responsabilise les acteurs du traitement de ces données et augmente le pouvoir des autorités de régulation au sein de l'Union européenne. Il est entré en application dans l'ensemble des 28 États membres de l'Union européenne à compter du 25 mai 2018. Il met en place des principes nouveaux, dont les plus importants sont :

- Un consentement explicite et positif doit être donné par l'utilisateur. Il dispose également d'un droit à l'effacement de ses données personnelles dans les meilleurs délais et peut demander à recevoir l'ensemble des données personnelles détenues par l'entreprise le concernant.
- Le principe de protection des données dès la conception (Privacy by Design) impose aux organisations de prendre en compte les exigences relatives à la protection des données personnelles dès la conception des produits, services et systèmes exploitant ces données.
- La minimisation des données impose à l'organisation de ne collecter, traiter et stocker que les données nécessaires en vue de l'exécution de ses tâches.
- Les organisations publiques, traitant des données très sensibles ou de plus de 250 personnes ont l'obligation de nommer un délégué à la protection des données pour contrôler le respect du règlement, de conseiller le responsable des traitements sur son application et de faire office de point de contact avec l'autorité de contrôle, de répondre aux sollicitations de personnes qui souhaitent exercer leurs droits.
- L'ensemble des dispositions du RGPD s'appliquent aux organisations établies en dehors de l'UE qui traitent les données relatives aux activités des organisations de l'UE ou appartenant à des résidents de l'UE

L'utilisation responsable des données implique également de manière plus large un certain nombre de critères qui préviennent la création d'intelligences artificielles intrusives, pouvant porter atteinte à l'ordre public, inciter à la haine ou de manière plus générale promouvant ou permettant toute activité interdite par la loi ou allant à l'encontre de l'éthique. L'éthique des données est une problématique en perpétuelle évolution, mobilisant nombre de groupes de

travail publics comme indépendants. Si le cadre actuel est défini par le RGPD, ces acteurs et groupes de travail essayent de créer des cadres pour faire en sorte que les évolutions permises en science des données bénéficient directement à l'intérêt général et pas uniquement à des acteurs privés. Ces cadres incluent par exemple le site d'open source de bases de données Kaggle ou la plateforme de travail Data for Good.

Marché du travail

Indicateurs au niveau des familles professionnelles (FAP) des métiers visés, agrégés au niveau national — France Travail / DARES (open data).

Tension de recrutement (DARES 2024)	Élevée (4/5) Moyenne pondérée par l'emploi ; les familles ci-dessous s'échelonnent de 3 à 5 sur 5.
Facteurs principaux	Manque de main-d'œuvre disponible, Lien formation-emploi
Salaire médian (mensuel)	2 897 €
Diplôme le plus fréquent	Diplôme de niveau bac+5 ou plus (doctorat, master ou équivalent) en communication et information (41 % des actifs en emploi)
Familles professionnelles (FAP)	L4X82 — Techniciens et agents de maîtrise chargés d'études socio-économiques · tension 4/5 (Élevée) L5X93 — Chargés d'études socio-économiques · tension 3/5 (Modérée) M1X80 — Techniciens d'étude et de développement en informatique · tension 4/5 (Élevée) M1X81 — Techniciens de production, d'exploitation, d'installation, et de maintenance, support et services aux utilisateurs en informatique · tension 4/5 (Élevée) M2X90 — Ingénieurs et cadres d'étude, recherche et développement en informatique et télécom · tension 5/5 (Très élevée) M2X91 — Chefs de projet et directeurs de service informatique · tension 4/5 (Élevée) M2X92 — Responsables et cadres de la production, de l'exploitation et de la maintenance informatique et télécom · tension 3/5 (Modérée) M2X93 — Experts et consultants en systèmes d'information · tension 5/5 (Très élevée)
Projets de recrutement (BMO 2026)	46 226
Jugés difficiles à recruter	49 %
Emplois saisonniers	2 %
Familles professionnelles (FAP)	L4X82 — Techniciens et agents de maîtrise chargés d'études socio-économiques L5X93 — Chargés d'études socio-économiques M1X80 — Techniciens d'étude et de développement en informatique M1X81 — Techniciens de production, d'exploitation, d'installation, et de maintenance, support et services aux utilisateurs en informatique M2X90 — Ingénieurs et cadres d'étude, recherche et développement en informatique et télécom M2X91 — Chefs de projet et directeurs de service informatique M2X92 — Responsables et cadres de la production, de l'exploitation et de la maintenance informatique et télécom M2X93 — Experts et consultants en systèmes d'information

Voies d'accès

MODALITÉS D'ACCÈS

Formation initiale, Formation continue, Apprentissage, Contrat de professionnalisation, VAE

PRÉREQUIS À L'ENTRÉE EN FORMATION

- Un niveau 5

dans un domaine scientifique et/ou économique. À l'international, un niveau équivalent est requis. Ce niveau sera justifié par le CV de l'apprenant.

- La réussite d'un test d'admission "coding challenge", validant des compétences nécessaires scientifiques en mathématiques appliquées ainsi qu'en programmation python.

- La réussite d'un entretien préalable qui jugera:

- o De la plus-value de l'acquisition des compétences décrites dans le référentiel "Concepteur Développeur en Science des Données " pour le projet professionnel

- o De la pertinence et de la faisabilité de ce projet professionnel

- o De la capacité du candidat à suivre la Formation dans les conditions définies au préalable (possibilité de suivre la formation en présentiel ou distanciel)

- Une expérience professionnelle adaptée sera valorisée pour l'admission dans la formation.

- Une

pratique courante de l'anglais (niveau B2 en compréhension & expression écrite et B1 de compréhension & expression orale du référentiel européen CECRL) est requise.

- Dans certains cas, la réussite aux tests d'admission et à l'entretien pourront se substituer à un diplôme de niveau 5.

COMPOSITION DU JURY — FORMATION INITIALE, FORMATION CONTINUE, APPRENTISSAGE, CONTRAT DE PROFESSIONNALISATION

Le jury de certification est composé de :

- 1 représentant de Jedha, qui préside le jury : le président, le directeur des contenus, le directeur des programmes ou leur représentant désigné

- 1 membre externe, préférentiellement choisi parmi les membres externes du conseil pédagogique

- 1 à 2 autres membre(s) extérieur(s), spécialistes ou recruteurs spécialisé dans en science des données

Les membres du jury ont une expérience significative en Sciences des Données, suffisante pour pouvoir évaluer les apprenants sur la présentation de leur travail. Tous les membres du jury ont une expérience technique significative et peuvent venir de différents secteurs d'activités (Finance, Banque, Marketing etc.).

COMPOSITION DU JURY — VAE

Le dossier VAE est étudié directement par :

- 1 représentant de Jedha, qui préside le jury : le président, le directeur des contenus, le directeur des programmes ou leur représentant désigné

- 1 membre externe, choisi parmi les membres externes du conseil pédagogique

- 1 à 2 autres membre(s) extérieur(s), spécialistes ou recruteurs spécialisés dans les sciences des données

Financement en alternance (NPEC)

Niveau de prise en charge d'un contrat d'apprentissage par les OPCO, fixé par France compétences par branche professionnelle (CPNE).

Prise en charge usuelle	8 195 €
Selon la branche	7 000 € – 9 480 €
Prise en charge	Branches concernées
8 195 €	Valeur par défaut · 192 branches
9 480 €	CPNE Casinos, CPNE autoroutes, CPNE de l'édition de livres, de l'édition phonographique et de l'édition de musique, CPNE des Télécoms, CPNE des industries de santé (CPNEIS)
7 923 €	CPNE FP Branche des magasins du Bricolage, CPNEFP Commerce succursaliste de la Chaussure, CPNEFP Commerce à distance, CPNEFP Grands Magasins/ Magasins Populaires, CPNEFP des commerces et services de l'audiovisuel, de l'électronique et de l'équipement ménager
8 400 €	CPNEFP de l'habillement, mercerie, chaussure et jouet
8 255 €	CPNEFP Professions de la Photographie
8 134 €	CPNE FP Commerce de Détail de l'Horlogerie Bijouterie

7 967 €	CPNEFP de la branche des entreprises du bureau et du numérique (commerces et services)
7 124 €	CPNEFP de la blanchisserie, laverie, location de linge, nettoyage à sec, pressing et teinturerie
7 000 €	CPNEFP de la métallurgie

Base légale

Date de décision	10/02/2021
Durée de l'enregistrement	5 ans
Date d'échéance de l'enregistrement	10/02/2026

Statistiques

Année	Certifiés	Dont VAE	Insertion 6 mois	Métier 2 ans
2020	60	0	78 %	—
2019	60	0	86 %	—

Pour plus d'informations

LIEN VERS LE DESCRIPTIF

<https://www.jedha.co/page/certificat-concepteur-developpeur>

Fiche officielle (France Compétences)

<https://www.francecompetences.fr/recherche/rncp/35288/>

Entrées et sorties de formation

Entrées en formation	1 235
Parcours menés à terme	1 187
Taux d'achèvement	96 %
Période	02/2022 – 09/2026

Entrées et sorties — Caisse des Dépôts (open data), tous organismes confondus.

Demande CPF — formations engagées

Dossiers CPF engagés	1 408
Montant total engagé	3 282 962 €
Coût moyen par dossier	2 332 €
Durée moyenne	173 h
Période	02/2021 – 02/2026
France compétences	75 %
France Travail	12 %
Reste à charge titulaire	6 %
Régions	4 %

Entreprise	2 %
État	0 %
OPCO	0 %
Compte d'engagement citoyen	0 %

Formations engagées via le CPF — Caisse des Dépôts (open data), tous organismes confondus.