

Data Analyst

RNCP37429 · RNCP · Publiée

L'essentiel

Niveau de qualification	Niveau 6
Code(s) NSF	326
Formacode(s)	31025
Code(s) ROME	M1805
Type d'enregistrement	Enregistrement sur demande
Date d'enregistrement	27/03/2023
Date d'échéance	27/03/2026

Certificateur(s)

SIMPLON.CO · SIRET : 79279132900016

WILD CODE SCHOOL · SIRET : 79492606300247

INNOV'EDUC · SIRET : 79492606300155

Résumé de la certification

OBJECTIFS ET CONTEXTE

Les compétences de data Analyst sont parmi les plus demandées sur le marché. L'OPIIEC (Observatoire des métiers du numérique) indique une augmentation de + 51,7% d'offres d'emploi sur un an (Janv.2022, agrégation des sites d'offres d'emplois, données du ministère du travail).

Le volume croissant des données disponibles, leur variété sans cesse élargie, et la rapidité avec laquelle elles sont créées et sont susceptibles de générer elles-mêmes de la donnée, ont fait basculer nombre d'entreprises dans l'ère du big data. Aujourd'hui, la donnée est en effet partout, permettant aux entreprises d'accroître le suivi de leur performance via des indicateurs toujours plus nombreux, d'analyser des comportements clients et par là même d'identifier des opportunités de marché. Dans ce contexte, les entreprises ont besoin de mobiliser des compétences spécifiques pour collecter les données, les stocker, les modéliser ou encore les protéger. Qu'il s'agisse d'informations nominatives, postales ou bancaires sur leurs prestataires, leurs clients, leurs partenaires, leurs salariés, les entreprises ont toujours amassé des données. Elles les ont toujours utilisées comme données de gestion : pour la paie, les facturations, les recouvrements, etc. Mais aujourd'hui, avec l'afflux massif de données, la multiplication de solutions de collecte, de stockage et les possibilités accentuées de traitement en temps réel, les entreprises se voient offrir de nouvelles opportunités. Une fois les données collectées et stockées, tout l'enjeu pour les entreprises est de tirer profit de ces données, c'est-à-dire d'en extraire leur valeur pour développer de nouveaux services, de nouveaux produits ou de nouvelles opportunités marché. Le traitement et l'analyse des données se situent au cœur des enjeux du big data. Deux métiers sont emblématiques de ce travail : celui du data scientist et du data analyst.

La certification « Data analyst » s'inscrit donc dans un contexte où la gestion (collecte, traitement, diffusion) des données constitue un enjeu majeur pour toutes les entreprises et tous les acteurs amenés à gérer et utiliser des données clients ou des données pouvant avoir un caractère stratégique de par l'enjeu qu'elles représentent pour des entreprises sur différents sujets stratégiques et économiques. Elle s'inscrit également dans un contexte particulier où les datas sont omniprésentes et où l'Europe a été amenée à définir et à renforcer les cadres de protection des données afin qu'ils soient plus à même de protéger les personnes et les organisations et qu'ils soient plus cohérents dans l'union en exigeant une application rigoureuse des règles. Le nombre d'offres d'emploi de data analyst publiées est d'ailleurs en constante augmentation.

Cette certification de niveau intermédiaire, est en phase avec les besoins du monde professionnel, axée sur un métier

(data analyst) en pleine évolution, essentiel à la chaîne des professionnels de la donnée et complémentaire à des métiers d'un niveau supérieur (data scientist, data ingénieur, data architecte).

Son objectif est de permettre :

- aux entreprises de la branche de bénéficier de compétences nouvelles et pointues, et de pouvoir recruter des analystes de la donnée opérationnels, bénéficiant des derniers apports technologiques et ayant suivi une formation répondant concrètement aux besoins spécifiques et contextualisés du secteur.
- aux titulaires de pouvoir collecter, exploiter, modéliser, sécuriser tous types de données et les traiter en fonction des enjeux des organisations tout en assurant la mise en place de process et de modes de fonctionnement conformes au règlement général de protection des données RGPD tel que défini par le législateur et à l'ensemble des réglementations applicables, afin qu'ils soient capables d'intervenir immédiatement sur des problématiques de collectes et d'analyses de données en mettant à disposition des organisations des outils adaptés et performants s'inspirant de ce qui se fait de mieux actuellement, y compris en termes de structure des modèles et des outils de présentation et de partage de données.

Tous les secteurs sont concernés, tous les métiers et toutes les tailles d'entreprise. Les sources de données sont très variées : la géolocalisation, les objets connectés, l'utilisation d'Internet ou de programmes, sources média, réseaux sociaux, sites web, etc. À ces nombreuses sources vient s'ajouter l'open data qui désigne l'ensemble des données en accès libre appartenant au domaine public.

ACTIVITÉS VISÉES

Le data analyst transforme les données brutes en informations pertinentes permettant la prise de décision. Pour cela, il est amené à pratiquer les activités suivantes :

- Il identifie et analyse des besoins d'entreprises, d'organisations et d'utilisateurs concernant des utilisations de données.
- Il définit et propose une stratégie de prise de décision par les données suivant les besoins métier
- Il traduit des datas afin de les rendre utilisables et exploitables en fonction des besoins et objectifs d'utilisateurs.
- Il analyse des bases de données, effectue des traitements et des analyses sur des données existantes. Il réalise des tableaux de bord (dashboards) d'organisations.
- Il conçoit et structure les étapes de collecte et de traitement de données pour les optimiser au mieux.
- Il effectue des requêtes et il modifie les bases de données grâce au langage SQL.
- Il réalise des agrégations, des jointures et des sous-requêtes sur des bases de données relationnelles SQL.
- Il automatise des collectes de contenu sur des pages web (scraping HTML), et effectue des requêtes grâce aux API REST.
- Il prend en compte les enjeux du RGPD et les responsabilités de chacun liées à la collecte, au traitement et à l'utilisation de données.
- Il sélectionne, paramètre, utilise des logiciels de rédaction et d'exécution de code
- Il participe activement à la gestion d'un projet "data", en utilisant les méthodologies modernes (agile, scrum, etc...)
- Il utilise différents algorithmes, notamment en langage Python (variables, boucles, itérateur, conditions, fonctions), se sert des structures de données pour effectuer des analyses.
- Il effectue des calculs matriciels (NumPy), utilise les DataFrames (Pandas) pour la création, l'import, la manipulation, le retraitement de données et de valeurs.
- Il interprète des données, utilise des plateformes de data science (de type KNIME, dataiku), les statistiques descriptives pour analyser des datas.
- Il explique des résultats d'analyses de datas, différencie les corrélations et les causalités en développant son sens critique par rapport à la science.
- Il réalise des Dashboards (tableaux de bords) avec des logiciels adaptés et définit des cartographies (Folium) de collecte et de compilation de données.
- Il utilise des tableurs (de type Excel) et réalise des tableaux croisés dynamiques afin de permettre à des utilisateurs d'effectuer des requêtes sur des datas spécifiques.
- Il accompagne des utilisateurs et des équipes dans l'utilisation d'outils, dans des recherches, des interprétations de données brutes et retraitées.
- Il analyse des utilisations d'outils et datas retraitées par des utilisateurs afin de proposer des évolutions et améliorations des outils.
- Il identifie des possibilités d'amélioration des recherches et exploitations de datas par des organisations et des utilisateurs.
- Il relaie l'information au sein de son équipe en utilisant une communication efficace et adaptée.

- Il réalise une veille permanente sur la date, se forme aux nouveaux outils et fait évoluer sa pratique.
- Il conseille les utilisateurs sur la protection des données, sur l'éthique des algorithmes, sur la prise en compte du handicap et la responsabilité numérique de chacun. Il/elle présente à l'oral et à l'écrit de manière claire, concise et sans ambiguïté, les informations.

COMPÉTENCES ATTESTÉES

- Identifier les possibilités d'utilisation des données, en fonction des besoins métier, en étant force de proposition dans l'exploration, l'évaluation de la qualité et l'interprétation de ces données
- Définir une stratégie de prise de décision par les données suivant les besoins métier
- Modéliser des bases de données relationnelles
- Réaliser des requêtes avancées répondant à des besoins métier complexes
- Automatiser des collectes de données
- Mettre en place une interface standard de partage automatique de données entre différentes applications et langages
- Contrôler les modalités de collecte et d'utilisation de données et mesurer les enjeux du RGPD
- Effectuer des choix méthodologiques pour l'automatisation des traitements et les documenter avec clarté et concision
- Utiliser les outils et méthodes modernes : méthodes agiles afin de permettre le travail en équipe, outils de suivi de projets, logiciel adapté à la rédaction de code
- Manipuler des structures de données et utiliser l'algorithmie afin de traduire en script des besoins de traitements de données
- Appliquer les bonnes pratiques de la programmation afin d'avoir un code organisé, réutilisable et partageable
- Utiliser les tableaux de données afin de faciliter l'import, la manipulation et la fusion de données
- Nettoyer les données, retraiter les valeurs aberrantes (outliers) et les valeurs manquantes
- Utiliser les expressions régulières (RegEx) pour traiter les valeurs textuelles et permettre une anonymisation des données personnelles dans le cadre du RGPD
- Utiliser les statistiques descriptives afin de modéliser les données
- Maîtriser le process d'apprentissage automatique (Machine Learning)
- Modéliser des régressions afin de définir des modèles de prévisions, et de trouver des tendances futures
- Modéliser des classifications et interpréter les métriques associées afin de catégoriser automatiquement des informations
- Traiter automatiquement le langage naturel (NLP) à partir de texte brut
- Contrôler et documenter les biais d'un modèle et des données d'entraînement afin d'estimer les risques éthiques
- Identifier et prioriser les informations à rendre accessibles et à présenter visuellement
- Utiliser les visualisations descriptives afin de représenter graphiquement des données statistiques et des informations modélisées
- Manipuler la Dataviz interactive et dynamique
- Réaliser de la cartographie
- Utiliser un Tableur afin de proposer des croisements de variables pour obtenir des informations recherchées
- Réaliser des tableaux de bord avec des outils de Business Intelligence afin d'intégrer et de croiser des informations utiles à des approches stratégiques de problématiques
- Prendre en compte les handicaps visuels afin de produire des graphiques lisibles par tous
- Présenter à l'oral et à l'écrit de manière claire, concise et sans ambiguïté les informations

Blocs de compétences

RNCP37429BC01 — Collecte des données : exploration et requêtage des différents types de bases de données, et récupération des données

LISTE DE COMPÉTENCES

Identifier les possibilités d'utilisation des données, en fonction des besoins métier, en étant force de proposition dans l'exploration, l'évaluation de la qualité et l'interprétation de ces données

Définir une stratégie de prise de décision par les données suivant les besoins métier

Modéliser des bases de données relationnelles (notamment SQL) afin de répondre avec rigueur aux besoins des utilisateurs

Réaliser des requêtes avancées répondant à des besoins métier complexes : agrégations, jointures, vues et sous-

requêtes en s'assurant de l'intégrité des données

Automatiser des collectes de données afin d'exploiter les contenus et les informations récoltées sur des pages web ("web scraping") en s'assurant du respect de la réglementation en vigueur

Mettre en place une interface standard de partage automatique de données entre différentes applications et langages (par exemple avec API REST)

Contrôler les modalités de collecte et d'utilisation de données et mesurer les enjeux du RGPD afin d'estimer les risques et les responsabilités de chacun

MODALITÉS D'ÉVALUATION

Mise en situation professionnelle : Mise en place d'une collecte de données

Cadre de mise en œuvre et de déroulement de l'évaluation : Les candidats sont évalués individuellement lors de la réalisation d'une mise en situation au cours de laquelle ils doivent effectuer des recherches à partir de questions proposées, et automatiser la récupération des données depuis une base de données.

RNCP37429BC02 — Automatisation du traitement des données : nettoyage, complétion, correction, uniformisation

LISTE DE COMPÉTENCES

Effectuer des choix méthodologiques pour l'automatisation des traitements et les documenter avec clarté et concision

Utiliser les outils et méthodes modernes : méthodes agiles afin de permettre le travail en équipe, outils de suivi de projets, logiciel adapté à la rédaction de code

Manipuler des structures de données (chaînes de caractères, listes, dictionnaires et tuples) et utiliser l'algorithmie (variables, boucles, itérateur, conditions, fonctions) afin de traduire en script des besoins de traitements de données

Appliquer les bonnes pratiques de la programmation afin d'avoir un code organisé, réutilisable et partageable dans un cadre professionnel

Utiliser les tableaux de données (notamment les DataFrames avec Python et Pandas) afin de faciliter l'import, la manipulation et la fusion de données

Nettoyer les données, retraiter les valeurs aberrantes (outliers) et les valeurs manquantes afin d'éviter les impacts sur l'exploitation et le traitement des bases de données

Utiliser les expressions régulières (RegEx) pour traiter les valeurs textuelles et permettre une anonymisation des données personnelles dans le cadre du RGPD

MODALITÉS D'ÉVALUATION

Mise en situation professionnelle : Structuration d'outils / programmes de traitement de données selon des objectifs

Cadre de mise en œuvre et de déroulement de l'évaluation : Les candidats doivent, à partir d'objectifs définis de traitement de données, structurer des outils et utiliser des algorithmes afin de manipuler des tableaux de données, et automatiser le nettoyage notamment des valeurs manquantes et aberrantes.

Évaluation individuelle des candidats par un examinateur

RNCP37429BC03 — Modélisation des données structurées : identification des corrélations existantes et utilisation du Machine Learning pour établir des prévisions

LISTE DE COMPÉTENCES

Utiliser les statistiques descriptives afin de modéliser les données et en faire émerger des informations pertinentes

Maîtriser le process d'apprentissage automatique (Machine Learning) afin de permettre à des algorithmes d'apprendre automatiquement à partir de données : syntaxe, découpage jeu d'entraînement et de validation, entraînement, prédiction, mesure

Modéliser des régressions et interpréter les métriques associées afin de définir des modèles de prévisions, et de trouver des tendances futures pour des valeurs numériques

Modéliser des classifications et interpréter les métriques associées afin de catégoriser automatiquement des informations

Traiter automatiquement le langage naturel (NLP) à partir de texte brut afin d'en tirer de la valeur en fonction de classification (Analyse de sentiments)

Contrôler et documenter les biais d'un modèle et des données d'entraînement afin d'estimer les risques éthiques de ce modèle, notamment dans le cas d'usages de données personnelles

Communiquer et vulgariser le fonctionnement interne d'un algorithme d'apprentissage automatique afin d'éviter le phénomène de "boîte noire"

MODALITÉS D'ÉVALUATION

Mise en situation professionnelle : Modélisation de données structurées et détermination de prévisions

Descriptif de l'évaluation : Les candidats doivent présenter individuellement une proposition de modélisation de données structurées permettant de décrire les données de manière simple, d'en tirer des tendances, de prévoir des valeurs futures, et d'en interpréter les résultats.

RNCP37429BC04 — Visualisation des données : valorisation et interprétation des données pertinentes, et mise en forme dans un tableau de bord

LISTE DE COMPÉTENCES

Identifier et prioriser, en fonction du besoin métier, les informations à rendre accessibles et à présenter visuellement, afin de structurer des représentations graphiques de tableaux bord

Utiliser les visualisations descriptives, notamment les nuages de points, les boîtes à moustache et les histogrammes afin de représenter graphiquement des données statistiques et des informations modélisées à destination d'analyste de données

Manipuler la Dataviz interactive et dynamique (par exemple avec Plotly ou Bokeh) afin de réaliser différents types de représentations graphiques et visuelles de données à destination d'utilisateurs opérationnels

Réaliser de la cartographie (notamment avec Folium) afin de représenter des informations géographiques

Utiliser un Tableur, et notamment les tableaux croisés dynamiques afin de proposer des croisements de variables pour obtenir des informations recherchées

Réaliser des tableaux de bord avec des outils de Business Intelligence (par exemple PowerBI ou Tableau) afin d'intégrer et de croiser des informations utiles à des approches stratégiques de problématiques

Prendre en compte les handicaps visuels afin de produire des graphiques lisibles par tous

Présenter à l'oral et à l'écrit de manière claire, concise et sans ambiguïté les informations

MODALITÉS D'ÉVALUATION

Mise en situation professionnelle : Présentation d'un Tableau de bord réalisé à partir de besoins d'un client

Cadre de mise en œuvre et de déroulement de l'évaluation : Les candidats doivent présenter individuellement un tableau de bord réalisé à partir d'un cahier des charges détaillant, à partir de bases de données existantes, des objectifs et des attentes en termes de présentations graphiques et visuelles d'informations et de données croisées.

Secteur d'activité et type d'emploi

SECTEURS D'ACTIVITÉS

Le data analyst exerce son activité dans tous les secteurs d'activités. Il peut aussi bien être recruté par des industries, des grandes entreprises, des entreprises du secteur financier, que par des commerces ou même des organisations médicales ou paramédicales.

TYPE D'EMPLOIS ACCESSIBLES

- data analyst
- consultant/analyst BI (Business Intelligence)
- business analyst

RÉGLEMENTATIONS D'ACTIVITÉ

Le métier de data analyst n'est pas réglementé. Toutefois l'ensemble du processus de traitement des données 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.

Dans certaines structures, le Data Analyst a aussi en charge la politique de données. Il doit veiller à ce que celle-ci respecte les réglementations en vigueur, en particulier le RGPD et la loi Libertés et Informatique. Il peut aussi travailler en collaboration avec des experts en cybersécurité pour garantir la sécurité et l'intégrité des data.

Le data analyst devra également veiller aux principes de l'accessibilité numérique. S'il travaille au sein d'un établissement public, il devra se conformer aux critères définis dans les référentiels concernant les institutions publiques.

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)	Très élevée (5/5) Moyenne pondérée par l'emploi ; les familles ci-dessous s'échelonnent de 4 à 5 sur 5.
Facteurs principaux	Manque de main-d'œuvre disponible, Lien formation-emploi
Salaire médian (mensuel)	2 864 €
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)	M1X80 — Techniciens d'étude et de développement en informatique · tension 4/5 (Élevée) M2X90 — Ingénieurs et cadres d'étude, recherche et développement en informatique et télécom · tension 5/5 (Très élevée)
Projets de recrutement (BMO 2026)	27 458
Jugés difficiles à recruter	53 %
Emplois saisonniers	1 %
Familles professionnelles (FAP)	M1X80 — Techniciens d'étude et de développement en informatique M2X90 — Ingénieurs et cadres d'étude, recherche et développement en informatique et télécom

Voies d'accès

PRÉREQUIS À L'ENTRÉE EN FORMATION

Aucun diplôme n'est exigé. Il faut être majeur et disposer de certains prérequis techniques :

- être à l'aise dans l'utilisation des outils informatiques en général, des tableurs en particulier (feuilles de calcul comme Microsoft Excel ou Google Sheet),
- valider des connaissances de base en programmation, mathématiques et statistiques.

En complément, une forte motivation et un projet professionnel solide constituent des critères de sélection prioritaires.

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

Le jury de certification est composé de 2 personnes :

1.

Un professionnel externe- Data analyst/consultant BI. Président du jury: il apporte une expertise en BI ainsi que la vision pratique et technique du métier, il apporte également son expertise en analyse de données

2. Un professionnel externe- Responsable data analyst. Membre du jury:

il fournit une perspective sur la gestion et les attentes en matière de compétences et de performance des Data Analysts au sein d'une organisation.

Afin d'accompagner le jury dans sa mission,

les documents suivants leurs seront remis:

- 1.règlement d'examen,
- 2.grilles de critères détaillées
3. Les attentes et des objectifs précis.

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	9 475 €
Selon la branche	6 840 € – 12 000 €
Prise en charge	Branches concernées

9 475 €	Valeur par défaut · 178 branches
12 000 €	CPNE Entreprises des activités du déchet et de la propreté urbaine, CPNE autoroutes, CPNE de l'édition de livres, de l'édition phonographique et de l'édition de musique, CPNEF de la Publicité, CPNEF de la presse et des agences de presse
9 480 €	CPNE Casinos, CPNE de la branche de la Distribution Directe, CPNEF de l'Exploitation cinématographique et de la Distribution de films, CPNEF de l'audiovisuel
11 055 €	CPNEFP de l'enseignement privé indépendant, CPNEFP de la Restauration rapide
8 990 €	CPNEF branche retraite complémentaire et prévoyance, CPNEFP Mutualité
7 923 €	CPNEFP des commerces et services de l'audiovisuel, de l'électronique et de l'équipement ménager, CPNEFP du Négoce de l'ameublement
11 845 €	CPNE des Télécoms
11 844 €	CPNE Travail Temporaire
10 500 €	CPNEFP des industries des tuiles et briques
10 000 €	CPNE des exploitations frigorifiques
9 880 €	CPNEFP de la Mutualité Sociale Agricole
9 027 €	CPNEFP Grands Magasins/ Magasins Populaires
9 000 €	CPNEFP de la branche des prestataires de services du secteur tertiaire
8 500 €	CPNE des industries de santé (CPNEIS)
8 285 €	CPNEFP Commerce à distance
8 057 €	CPNEFP de la branche des entreprises du bureau et du numérique (commerces et services)
8 036 €	CPNE FP Branche des magasins du Bricolage
8 029 €	CPNEFP des jardineries et graineteries
8 000 €	CPNE chimie
7 000 €	CPNEFP de la métallurgie
6 840 €	CPNEFP de la Branche des Caisses régionales de Crédit Agricole et autres organismes

Base légale

Date de décision	27/03/2023
Durée de l'enregistrement	3 ans
Date d'échéance de l'enregistrement	27/03/2026
Date de dernière délivrance possible	27/03/2030

Statistiques

Année	Certifiés	Dont VAE	Insertion 6 mois	Métier 2 ans
2021	31	0	100 %	—
2020	19	0	100 %	100 %

Liens avec d'autres certifications

NOUVELLE(S) CERTIFICATION(S)
RNCP42652

Pour plus d'informations

LIEN VERS LE DESCRIPTIF
<https://www.wildcodeschool.com/fr-FR/formations/formation-data-analyst>

Fiche officielle (France Compétences)
<https://www.francecompetences.fr/recherche/rncp/37429/>

Entrées et sorties de formation

Entrées en formation	643
Parcours menés à terme	344
Taux d'achèvement	53 %
Période	06/2023 – 07/2026

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

Demande CPF — formations engagées

Dossiers CPF engagés	677
Montant total engagé	2 285 537 €
Coût moyen par dossier	3 376 €
Durée moyenne	412 h
Période	05/2023 – 03/2026
France compétences	57 %
France Travail	19 %
Régions	17 %
Entreprise	3 %
Reste à charge titulaire	3 %
OPCO	1 %
Compte d'engagement citoyen	0 %

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