

Syllabus du programme Data Scientist

Un programme *Social Builder* et *VIVADATA* (Paris)

Le programme se compose de 4 grands modules de formation répartis sur 4 mois.

| | |
|--------------------------------------|------------|
| MODULE 1 - LEADERSHIP | 44 heures |
| MODULE 2 - CULTURE NUMÉRIQUE | 35 heures |
| MODULE 3 - INSERTION PROFESSIONNELLE | 29 heures |
| MODULE 4 - FORMATION DATA SCIENCE | 400 heures |

| | |
|--------------------------------------|------------|
| Durée de la formation : | 4 mois |
| Nombre d'heures total de formation : | 508 heures |

- Cette formation bénéficie d'un **tarif préférentiel** pour les demandeuses d'emploi de 4500€ au lieu de 7500€. (Ce prix inclut l'accompagnement Social Builder).
- Certaines places pourront bénéficier de **bourses** pour les profils qui ne peuvent pas payer la formation.

DESCRIPTION DE LA FORMATION

Le dispositif s'articule autour de 2 grands piliers :

Un pilier d'accompagnement et de formation autour de l'acculturation au numérique et du développement des compétences personnelles et professionnelles favorisant le retour à l'emploi rapide et durable (Dispensé par *Social Builder*).

Un pilier formation technique pour faire monter en compétences les apprenantes sur les métiers liés à l'intelligence artificielle et la data science (Dispensé par *VIVADATA*).

En deux mots, cette formation vous apprend à programmer en Python / à collecter, traiter, gérer et visualiser de la donnée / à déployer des réseaux de neurones et des algorithmes de machine learning pour résoudre des problèmes. A partir de nombreux exercices pratiques, elle vous permet d'acquérir les bonnes pratiques de développement et de mise en production d'un projet data. Vous trouverez davantage d'informations sur le curriculum [ici](#).

A l'issue de la formation, vous pourrez vous lancer dans une aventure entrepreneuriale, travailler sur des projets IA en tant que freelance ou postuler comme data analyst/scientist junior dans les nombreuses entreprises qui recrutent.

MODULE 1 - LEADERSHIP - 44h

Objectifs du module :

- *Clarifier ses priorités et prochains pas*
- *Comprendre et développer son leadership personnel*
- *Mobiliser des outils d'affirmation et de promotion de soi*
- *Comprendre son propre fonctionnement et celui des autres*
- *Comprendre la puissance des réseaux et créer son influence*

Séminaire Leadership et trajectoire professionnelle numérique - 7h

1 journée, animée par un.e coach professionnel.le, consacrée à un travail approfondi destiné à clarifier sa trajectoire, comprendre et développer son leadership personnel, renforcer l'affirmation de soi et partager avec le groupe.

Coaching collectif - 21h

10 séances de 2h seront prévues sur la durée du programme pour se former aux méthodes et outils de structuration d'un projet professionnel et capacité à résoudre ses difficultés d'insertion et développer des compétences essentielles dans toute carrière professionnelle telles que l'assertivité, le travail d'équipe ou la communication, en demi-groupes de 10 apprenantes.

Mentorat numérique individuel - 10h

Les mentor.e.s professionnel.le.s seront choisi.e.s pour former 2h/mois les apprenantes sur les enjeux du numérique, les métiers sur lesquelles elles peuvent se positionner et les éventuelles compétences à développer. Une séance de présentation au méthodes du mentorat sera organisée par Social Builder, à la fois pour les mentor.e.s et les mentorées.

Atelier de sensibilisation aux enjeux de l'égalité professionnelle et de la mixité en entreprise - 3h

Atelier de 3h, dispensé par une experte des questions d'égalité et de mixité en entreprise, destiné à sensibiliser les apprenantes et les différent.e.s intervenant.e.s du programme (mentor.e.s et coaches) à la cause soutenue par Social Builder et promouvoir une remise en question des pratiques et comportements adoptés dans le monde professionnel.

E-learning Social Builder Academy - 3h

Un accès à une formation de 3h en ligne sur Social Builder Academy, avec des cours pratiques et accessibles sur le leadership. Social Builder Academy est une plateforme en ligne qui offre des cours pratiques et accessibles sur des compétences numériques et la création d'entreprise. 🖱️ Le cours [ici](#)

MODULE 2 - CULTURE NUMÉRIQUE - 35h

Objectifs du module :

- *Découvrir de nouvelles méthodes de gestion de projet*
- *Comprendre les notions de lean startup*
- *Avoir les éléments clés pour pitcher son projet ou son parcours*
- *Appréhender les notions clés d'une présence sur le web réussie*
- *Découvrir des outils simples et gratuits au service de ses projets*
- *Comprendre les bases du Marketing Digital*
- *Élaborer une stratégie actionnable*
- *Savoir utiliser les outils du Marketing Digital*

Ateliers sur les nouvelles compétences numériques - 10h

Ateliers mensuels de 2h proposés par des professionnel.le.s du numérique (ex: identité numérique, cybersécurité, webmarketing, etc...) sur des compétences techniques favorisant une acculturation aux nouvelles compétences et le retour à l'emploi dans ce secteur en mutation constante.

E-learning Social Builder Academy - 8h

Un accès à une formation de 7h en ligne sur Social Builder Academy, avec des cours pratiques et accessibles sur le marketing digital, le web et l'entrepreneuriat. 🖱️ Le cours [ici](#)

Séminaires thématiques - 17h

Être agile est de plus en plus essentiel et demandé par les entreprises. Un.e coach agile proposera une introduction aux méthodes agiles et des jeux afin d'acquérir des outils et de les utiliser en situation professionnelle. Nous vous proposerons également plusieurs afterwork thématiques liés à la transformation numérique et aux métiers d'avenir.

MODULE 3 - INSERTION PROFESSIONNELLE - 29h

Objectifs du module :

- *Gérer son identité numérique*

- *Travailler son CV et sa lettre de motivation*
- *Se préparer aux entretiens de recrutement*
- *Apprendre à valoriser ses compétences*

Meet-ups communautés Social Builder et VIVADATA - 15h

Des rencontres avec l'ensemble de la communauté Social Builder et VIVADATA une fois par mois pour renforcer son réseau professionnel : coachs, mentor.e.s, apprenantes, alumnis, partenaires de structures variées du secteur du numérique telles que des startups, des entreprises de conseil digital ou encore de grands groupes Tech.

Ateliers de préparation aux entretiens d'embauche - 6h

2 ateliers de 3h de préparation aux entretiens d'embauche, dispensés par des professionnel.le.s du recrutement du secteur du numérique.

Jobdating - 4h

Entretiens avec des entreprises recruteuses faisant intervenir l'ensemble des partenaires impliqués dans le programme, afin de faciliter l'insertion des apprenantes du programme et de permettre aux entreprises de mettre en avant leurs besoins en recrutement.

Ateliers insertion professionnelle - 4h

Des ateliers de 2h sont animés par des professionnels autour de l'insertion professionnelle. Ces ateliers ont lieu au minimum une fois par mois.

MODULE 4 - FORMATION DATA SCIENCE - 400h

Objectifs du module :

- *Acquérir de solides fondamentaux en programmation informatique en maîtrisant le langage Python, le plus utilisé en data science.*
- *Collecter des données depuis des sites et applications web et effectuer des requêtes dans des bases de données.*
- *Manipuler et visualiser des données pour effectuer des analyses exploratoires*
- *Développer une connaissance opérationnelle des modèles et outils de machine learning, natural language processing et deep learning*
- *Mettre en production dans le cloud des programmes d'apprentissage machine*

Présélection et e-learning - (10h)

Il est demandé à chaque candidate d'effectuer la track « Learn Python » sur codecademy.com pour permettre aux apprenantes futures de s'initier au langage informatique et les conforter dans leur choix de formation.

Travail préparatoire - (40h)

Le travail préparatoire doit être réalisé chez soi avant le début de la formation par les apprenantes sélectionnées. Il est accessible sur une plateforme en ligne dédiée et permet de comprendre les notions de base avant de commencer la formation. Cette introduction facilitera le processus d'apprentissage par la suite.

Programmation en Python - 40h

Ce premier module vise à poser de solides bases en programmation informatique. À la fin du module, les apprenantes sauront :

- écrire des programmes en Python, le langage de référence en data science, utilisé par les leaders de l'industrie ;
- manipuler les concepts fondamentaux de la programmation : variables, boucles, conditions, méthodes, objets et classes ;
- travailler en ligne de commande, versionner sous git, lire et écrire dans des fichiers, gérer les erreurs et les exceptions..

Data Collection & Engineering - 40h

Lors de ce deuxième module, les apprenantes découvriront les meilleures stratégies pour récupérer la donnée partout où elle se trouve, notamment grâce au scrapping de pages web et à l'utilisation d'APIs. Elles s'initieront également aux requêtes dans des bases de données relationnelles au moyen du langage SQL et de la librairie Python SQLAlchemy. Enfin, elles apprendront à paralléliser le traitement de grands volumes de données avec PySpark.

Data Analysis - 40h

Le troisième module permet d'entrer dans le vif du sujet en manipulant des données au moyen des librairies NumPy et Pandas. L'intuition des apprenantes est également développée grâce à la visualisation de données sous Matplotlib et Seaborn. Enfin, ce module est l'occasion de travailler avec des notebooks Jupyter et de s'initier à l'analyse exploratoire de données dans le cadre d'un projet concret.

Machine Learning - 120h

Ce quatrième module est le coeur du réacteur : les apprenantes découvriront le monde de l'apprentissage machine avec Scikit-Learn, la librairie de référence made in France. Elles travailleront en

apprentissage supervisé et non supervisé, développeront leurs compétences en feature engineering, apprendront à gérer les comportements d'overfitting et à sélectionner les modèles pertinents, mettront en place des pipelines pour structurer leurs projets de machine learning. Elles s'attaqueront à des projets concrets (recommender system et customer churn) pour adopter les meilleures pratiques du moment.

Natural Language Processing - 40h

Le cinquième module explorera la communication homme-machine au moyen d'outils de traitement du langage naturel (NLTK, SpaCy et Gensim). Les apprenantes découvriront le preprocessing de texte, l'analyse de sentiments et le topic modeling dans le cadre d'un projet NLP concret.

Deep Learning - 120h

Ce sixième et dernier module vise à enrichir les techniques d'apprentissage machine en explorant les réseaux de neurones artificiels (fully connected et convolutionnels). Les apprenantes seront formées à l'utilisation de TensorFlow, le célèbre framework de Google, et de la librairie Keras. Elles apprendront également à conteneuriser leur code et à envoyer leurs calculs sur des serveurs GPU dans le cloud avec Paperspace et AWS.