

Data Science Immersive France

English & French Syllabi

 **GENERAL ASSEMBLY**



Overview

General Assembly's Data Science Immersive (DSI) is a transformative course that prepares students to break into data careers.

Designed to get you hired, this 12-week data science bootcamp features best-in-class instruction, career coaching, and professional connections to top employers. Join us on campus or online in our remote classroom.

As a graduate, you'll leave poised to succeed in a variety of data science and advanced analytics roles, creating predictive models that drive decision-making and strategy throughout organisations of all kinds.

Throughout this expert-designed programme, you'll:

- Collect, extract, query, clean, and aggregate data for analysis.
- Gather, store, and organise data using SQL and Git.
- Perform visual and statistical analysis on data using Python and its associated libraries and tools.
- Craft and share compelling narratives through data visualisation.
- Build and implement appropriate machine learning models and algorithms to evaluate data science problems spanning finance, public policy, and more.
- Compile clear stakeholder reports to communicate the nuances of your analyses.
- Apply question, modeling, and validation problem-solving processes to data sets from various industries to provide insight into real-world problems and solutions.
- Prepare for the world of work, compiling a professional-grade portfolio of solo, group, and client projects.

To assess your skills, you will need to:

- Attend all class sessions throughout the course.
- Complete and present all projects to your instructor. A minimum score of 66% must be reached to pass each project.

Upon successfully meeting all of the course requirements, you will receive a certificate of completion.



Prerequisites

This is an intermediate-level course with some prerequisites:

- Students should arrive with a mathematical foundation and familiarity with Python and programming fundamentals. Some students have a technical background, such as a degree in mathematics or computer science or work experience in research or analysis. Other students engage in self-learning to build a foundation ahead of class.
- Take the Admission task provided by the Admission Team.
- Have access to a laptop (PC or Mac). It must be no more than four-years-old and able to run the most recent operating system, updated prior to starting the course.
- If remote, a webcam, headphones, and good access to the internet are required.

Note: The course also recommends a good level of English proficiency (for non-native speakers, a CEFR level of B2 or a TOEFL score of 90 is recommended). Check this [self-assessment grid](#). Our [Admissions team](#) can discuss your background and learning goals to advise if this course is a good fit for you.



What to Expect

Pre-Course Learning Paths

Set yourself up for success in this Immersive with up to 12 hours of preparatory lessons covering essential concepts in Python programming and applied math for data science. Designed to introduce you to foundational data skills and context, Data Science Fundamentals is a self-paced online learning path that will help you hit the ground running on day one of class.

The In-Class Experience

Engage in full-time, project-based learning that's designed to inspire a lifetime of discovery. As a DSI student, you'll:

- Explore new concepts and tools through expert-led lectures and discussions.
- Complete hands-on programming and modeling exercises to reinforce newly learned skills.
- Develop fluency in industry-essential topics and techniques via independent, pair, and group labs and real-world client work.
- Receive individualised feedback and support from your expert instructional team.
- Build out a professional portfolio to showcase your job-ready data science skills to potential employers and collaborators.

Dedicated Career Coaching

As an Immersive student, you'll receive dedicated support from career coaches who will help you set goals, build connections with employers, and stay on track with your job search. Throughout the course, you'll:

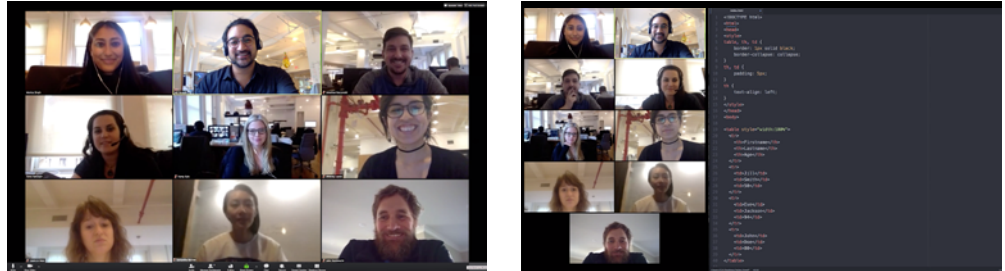
- Get an inside look at the industry through day-in-the-life talks, hiring panel discussions, guest speakers, and more.
- Cultivate a competitive candidate mindset, learning to assess your skill set against job descriptions, track progress, and recognize opportunities.
- Develop your professional brand: Polish your online and in-person presence.
- Practise technical challenges and whiteboarding skills to set yourself apart in interviews.
- Tap into an exclusive global network of experts, influencers, and peers, plus learn strategies for leveraging your existing connections, in person and online.
- Become an active contributor to the data science community.

After graduation, you'll also gain access to resources to help fuel a lifetime of learning. Dive into new topics or continue honing your data science skills with discounts on a suite of tools, passes and packages to premier events, and more. You can also apply tuition discounts to future GA courses, classes, and workshops, both on campus and online.



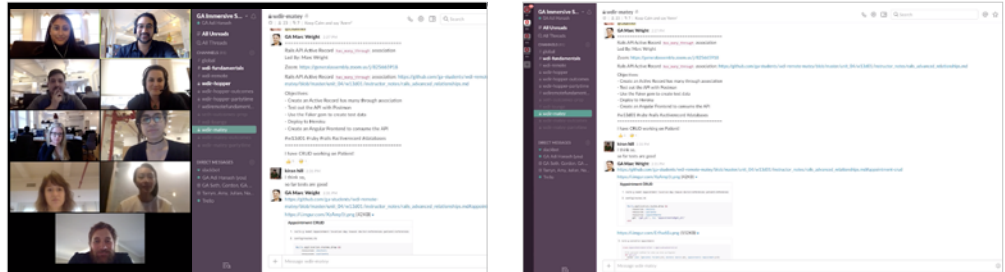
Inside the Remote Classroom

Engage in live, online, project-based learning that's designed to transform your career — from anywhere. Powered by Zoom, the major video conferencing tool, and the collaborative messaging platform Slack, our interactive classroom mirrors what you'll encounter as part of the modern workforce.



Instructor-led courses with Zoom.

You'll be able to not only share your screen and present your work to classmates but also collaborate in small groups via breakout rooms. Additionally, you can revisit recordings of class sessions in the future.



Chat and share with Slack.

Beyond class hours, get guidance, feedback, technical assistance, and more during frequent one-on-one check-ins and office hours.



What You'll Learn

Course Duration: 480 Hours

Pre-Work

Data Science Fundamentals

Dive into a series of self-paced lessons on the essentials of Python programming and applied math for data science before the course begins.

- Explore fundamental Python programming concepts, including variables, lists, loops, dictionaries, and data sets.
- Leverage programming tools like GitHub and the command line interface to manage data science projects.
- Practise solving coding challenges similar to the questions used in task-based data science interviews.
- Write and run Python functions using multiple arguments.
- Discover how key math concepts like statistical significance and probability distribution are applied throughout data science.

Unit 1 Fundamentals

Get acquainted with essential data science tools and techniques, working in a programming environment to gather, organise, and share projects and data with Git and UNIX.

- Demonstrate familiarity with introductory programming concepts using Python and NumPy to navigate data sources and collections.
- Utilise UNIX commands to navigate file systems and modify files.
- Learn to track changes and iterations using Git version control from your terminal.
- Define and apply descriptive statistical fundamentals to sample data sets.
- Practice plotting and visualising data using Python libraries like Matplotlib and Seaborn.

Project: Apply NumPy and Python programming skills to answer questions based on a clean data set.

Unit 2 Exploratory Data Analysis

Perform exploratory data analysis. Generate visual and statistical analyses, using Python and its associated libraries and tools to approach problems in fields like finance, marketing, and public policy.

- Design an experimental study with a well-thought-out problem statement and data framework.
- Use Pandas to read, clean, parse, and plot data, extracting and rearranging data through indexing, grouping, and JOINing.



Unit 2 Exploratory Data Analysis (Cont.)

- Review statistical testing concepts (p values, confidence intervals, lambda functions, correlation/causation) with SciPy and StatsModels.
- Learn to scrape website data using popular scraping tools.
- Explore bootstrapping, resampling and building inferences about your data.

Project: Leverage Pandas to apply advanced NumPy and Python skills, cleaning, analysing, and testing data from multiple messy data sets.

Unit 3 Classical Statistical Modeling

Explore effective study design and model evaluation and optimisation, implementing linear and logistic regression and classification models. Collect and connect external data to add nuance to your models using web scraping and APIs.

- Use scikit-learn and StatsModels to run linear and logistic regression models and learn to evaluate model fit.
- Begin to look at classification models by implementing the k-nearest neighbors (kNN) algorithm.
- Articulate the bias-variance trade-off as you practise evaluating classical statistical models.
- Use feature selection to deepen your knowledge of study design and model evaluation.
- Learn to apply optimisation and regularisation for fitting and tuning models.
- Dive into the math and theory behind how gradient descent helps to optimise loss functions for machine learning models.

Project: Explore, clean, and model data based on a provided data set, outlining your strategy and explaining your results.

Unit 4 Machine Learning Models

Build machine learning models. Explore the differences between supervised and unsupervised learning via clustering, natural language processing, and neural networks.

- Define clustering and its advantages and disadvantages as compared to classification models.
- Build and evaluate ensemble models using decision trees, random forests, bagging, and boosting.
- Get acquainted with natural language processing (NLP) through sentiment analysis of scraped website data.



Unit 4 Machine Learning Models (Cont.)

- Learn how Naive Bayes can simplify the process of analysing data for supervised learning algorithms.
- Explore the history and use of Hadoop, as well as the advantages and disadvantages of using parallel or distributed systems to store, access, and analyse big data.
- Understand how Hive interacts with Hadoop and discover Spark's advantages through big data case studies.
- Analyse and model time series data using the ARIMA model.

Project: Students will scrape and model their own data using multiple methods, outlining their approach and evaluating any risks or limitations.

Unit 5 Advanced Topics and Trends

Dive deeper into recommender systems, neural networks, and computer vision models, implementing what you've learned to productise models.

- Compare and contrast different types of neural networks and demonstrate how they are fit with back propagation.
- Build and apply basic recommender systems in order to predict on sample user data.
- Work with career coaches to create and polish your professional portfolio.
- Practise with data science case studies to prepare for job interviews.

Final Project: Choose a data set to explore and model, providing a detailed notebook of your technical approach and a public presentation on your findings.



Frequently Asked Questions

Why is this course relevant today?

It's ranked among the top fields in LinkedIn's [Emerging Jobs Report](#) for three years running and, according to the latest numbers, data scientist roles are experiencing 37% annual growth. Across industries from information technology and software to financial services and higher education, organisations are moving quickly to build robust in-house teams of data scientists and advanced analysts, and there's not enough talent to go around.

For those with quantitative backgrounds or existing analytical skills, building out this specialised skill set can spell ample opportunity to secure a high-paying job in data science, advanced analytics, business intelligence, and more.

What are the professional backgrounds of data science students?

DSI students come from all walks of life but share one common mission: They are passionate about launching a career in data science or advanced analytics. We see career-changers from diverse professional backgrounds, including engineers and recent STEM graduates, mid-career marketing and financial analysts, and business strategists, as well as more those from more far-flung fields like sales and the law.

What skills are assessed by the end of the programme?

In order to earn a certificate, you will be evaluated based on your ability to do the following:

- Collect, extract, query, clean, and aggregate data for analysis.
- Gather, store, and organise data using SQL and Git.
- Perform visual and statistical analysis on data using Python and its associated libraries and tools.
- Craft and share compelling narratives through data visualisation.
- Build and implement appropriate machine learning models and algorithms to evaluate data science problems spanning finance, public policy, and more.
- Compile clear stakeholder reports to communicate the nuances of your analyses.
- Apply question, modeling, and validation problem-solving processes to data sets from various industries to provide insight into real-world problems and solutions.

What does my tuition cover?

Here are just some of the benefits full-time students can expect at GA:

- Expert instruction in the skills you need to successfully transition into a data science career. If remote, you'll be given free access to Zoom, the videoconferencing application for this instructor-led course.
- Access to self-paced pre-work to explore data science fundamentals and prepare to hit the ground running on day one of class.
- Robust coursework, including expert-vetted lesson decks, project toolkits, and more. Refresh and refine your knowledge throughout your professional journey as needed.
- A professional-grade portfolio to showcase your ability to solve real-world data problems to potential employers and collaborators.



- Individual feedback and guidance from instructors and TAs during office hours. Stay motivated and make the most of your experience with the help of GA's dedicated team.
- Access to the collaborative chat tool Slack to communicate with your instructors, peers, and alumni throughout the course.
- Dedicated career services to help you navigate your personal job search experience, from technical challenges, to salary negotiation, and more.
- Technical interview prep, including resume reviews, mock interviews, and whiteboarding practice.
- Exclusive access to alumni discounts, networking events, and career workshops.
- A GA course certificate to showcase your new skill set on LinkedIn.
- Connections with a professional network of instructors and peers that lasts well beyond the course. The global GA community can help you navigate and succeed in the field.

Will I receive a certificate?

Upon successfully meeting all of the course requirements, you will receive a Data Science Immersive certificate. To pass, you will need to:

- Attend all class sessions.
- Complete and present all projects to your instructor. A minimum score of 66% must be reached to pass each project.

Thousands of GA alumni use their course certificate to demonstrate skills to potential employers — including our 19K+ hiring partners — along with their LinkedIn networks. GA's data programmes are well-regarded by many top employers, who contribute to our curriculum and partner with us to train their own teams.

What projects will I work on during this course?

For your capstone project, you'll tackle a real-world data problem from end to end. Develop a pitch and problem statement, source and collect relevant data, conduct an exploratory data analysis, and build a predictive model. You'll document and share your findings through a presentation, technical report, and non-technical summary.

Throughout this Immersive, you'll also compile a portfolio of projects designed to reinforce what you've learned in each unit. Gain hands-on experience with statistical and machine learning models, Python programming tools, recommender systems, neural networks, and more.

How does this course relate to GA's other data programmes?

This Data Science Immersive course is meant for students who are committed to changing careers and provides the most direct pathway to data science and other advanced analytics roles. For professionals who already work with data and want to perform more complex analysis involving computation, we also offer an evening and weekend [Data Science](#) course. If you're searching for a more entry-level course, [Data Analytics](#) teaches beginners how to perform rigorous analysis with Excel, SQL, and Tableau.

Which format should I take this course in — on campus or online?

It's up to you! Our remote courses offer a learning experience that mirrors GA's on-campus



offerings but allow you to learn from the comfort of home. If you don't live near a GA campus, have a busy travel schedule, or just want to save yourself the commute, a remote course could be a good option for you, if available in your market. You'll still get access to the expert instruction, learning resources, and support network that GA is known for.

If you prefer to learn alongside your peers and can make it to campus, our in-person courses allow you to take advantage of our beautiful classrooms and workspaces. Our [Admissions team](#) can advise you on the best format for your personal circumstances and learning style.

What do career services look like for remote students?

The same as our on-campus experience! We approach our remote Outcomes programming with the same philosophy, promise to, and expectations of our students. Get an insider's look at the industry through virtual sessions like day-in-the-life talks, hiring panel discussions, guest speakers, and more. You'll work individually with your career coach to understand your local job market, find opportunities, and connect with the local data community. Our Admissions team can provide more details on the dedicated support you'll receive on the path to landing a data science role.

In what language is this programme taught?

Our courses will be delivered in two ways:

- Fully in English
- In French, with English materials

Please check with the Admissions team to confirm which is true of your course of interest and what is relevant for your situation. In both cases, if you are not native speaker, a CEFR level of B2 or a TOEFL score of 90 is recommended. (Review this [self-assessment grid in English](#).)

What are my financing options?

We offer a variety of different [financing options](#), including instalment plans, so you can focus on what counts — your education. Speak with our [Admissions team](#) to learn more and find the best fit for you and your finances.

Our programme is currently not eligible for governmental funds in France such as CPF, however, we are working to become eligible in the future.

The "Compte personnel de formation" (i.e., "personal training account") was created in order to provide training opportunities throughout an employee's professional life in France, encouraging everyone to undertake training or study throughout their working life until retirement age.

The CPF is made available to employees working in the private sector, as well as job-seekers and youth who are entering the world of work (at 15- or 16-years-old) exclusively in France.

However, do not hesitate to contact your company to find out if it can utilise its personal funds to help you finance training, or contact your National Job Centre (Pôle Emploi) advisor if you are a job-seeker. Indeed, many structures in France such as your region, Pôle Emploi, or AGEFIPH can help you finance your training. Your job-seeker advisor at Pôle Emploi will be able to direct you toward these different systems or offer you Individual Training Assistance (AIF).



General Assembly is registered on the [DataDock](#) database and therefore meets the quality requirements set by French law.



Take the Next Step

Have questions about our Data Science Immersive course? Our [Admissions team](#) is here to help and can advise on if this programme is right for you and your learning goals. You can also:

- Attend an info session [online](#) or at your [local campus](#).
- Explore your [financing options](#).
- [Apply](#) to enrol in the course.*
- Email us directly at paris-inscription@generalassemb.ly or call at +33 1.88.24.42.51.

Course Enrolment Availability

Enrolments are available up to 14 days before the course starts.

Course Dates & Prices

GA delivers about four sessions per year. To learn about our courses dates and prices, please explore [this page](#) or contact our Admission team.

** Course modality options vary by location, pending market availability. Please contact our Admissions team to discuss what version is available in your location.*

Accessibility

GA provides access to its courses for people with disabilities. Please contact our Admissions team to discuss your needs and learn how we could support you throughout your learning journey.



Présentation de la formation

La formation Data Science Immersive de General Assembly est un accélérateur de carrière. En trois mois intensifs, formez-vous à la data science et soyez prêt à vous lancer sur le marché de l'emploi !

Diplômé, vous aurez les compétences nécessaires pour réussir en data science et créer des modèles prédictifs qui pourront orienter la prise de décision et la stratégie. En effet, le data scientist est un spécialiste de la donnée. Il est chargé d'analyser la data que l'entreprise récupère par divers canaux et la restitue sous forme de prospective, conseils, amélioration produits, ou performance de l'entreprise. Il part généralement d'une problématique dont l'entreprise a besoin de résoudre grâce à la modélisation statistique, détermine les données pertinentes à exploiter, puis les analyse afin de resituer les résultats et propose des solutions.

Chez General Assembly, la formation mélange des séances théoriques et pratiques, animées par un expert métier. Vous réalisez pas à pas, des projets concrets afin de constituer votre portfolio professionnel.

À la fin de cette formation, vous serez capable de :

- Collecter, extraire, interroger, nettoyer, et agréger des données pour analyse
- Effectuer une analyse visuelle et statistique des données à l'aide d'UNIX, de Git, de SQL, de Python, et de leurs bibliothèques et outils associés
- Créer et implémenter les modèles d'apprentissage automatique et les algorithmes appropriés pour évaluer des problèmes de data science couvrant la finance, les politiques publiques, etc.
- Créer et partager des récits convaincants grâce à la visualisation de données
- Compiler des rapports clairs des partenaires pour communiquer les nuances de vos analyses
- Identifier les problèmes de big data et expliquer comment les systèmes distribués et les technologies informatiques parallèles permettent de résoudre ces difficultés
- Appliquer des processus de résolution de problèmes en matière d'interrogation, de modélisation, et de validation à des ensembles de données provenant de divers secteurs pour donner un aperçu de problèmes et de solutions tirés du monde réel



Pour valider la formation, il faut :

- Être assidu aux journées de cours
- Réaliser et soutenir les projets face au formateur. Pour attester de la maîtrise de l'ensemble des compétences visées, vous devez obtenir un score de min. 66% pour chacun des projets de la formation.

Le respect de ces deux critères donne accès à la validation de la formation et à son certificat.

Lancez-vous dans une nouvelle carrière !

Pré-requis

Il s'agit d'une formation de niveau intermédiaire qui nécessite des pré-requis :

- Un niveau Bac S en mathématiques et des notions de programmation en Python sont requis. Certains étudiants ont par exemple un diplôme en mathématiques, en informatique, ou encore une expérience professionnelle dans la data, d'autres, sont autodidactes
- Valider le test d'entrée fourni, par notre équipe d'admission
- Avoir un ordinateur portable PC ou Mac (le système d'exploitation doit être le plus récent - moins de 4 ans - ou mis à jour avant la formation)
- Si la formation se déroule à distance, une connexion à Internet, webcam et des écouteurs sont nécessaires.

Lorsque la formation est délivrée en anglais, il est recommandé de :

- Disposer d'un niveau B2 en anglais (CECRL) ou un score au TOEFL d'au moins 90. En effet, en fonction de votre choix et des disponibilités par notre campus, la formation peut être entièrement en anglais ou en français avec des supports pédagogiques en anglais. Vous pouvez vous auto-évaluer à partir de ces grilles ([description en français](#) / [description en anglais](#))

N'hésitez pas à contacter notre [équipe d'admission](#) (ou écrivez directement à paris-inscription@generalassemb.ly) pour faire le point sur votre parcours et vos objectifs afin de vérifier si ce bootcamp est adapté à vos besoins.



Déroulé de la formation

Vous participez à une formation très pratique, basée sur des projets concrets et conçue pour couvrir tous les concepts et outils de base. Notre expert vous apporte toutes les compétences nécessaires pour réaliser, pas à pas, vos projets et mobiliser vos compétences. Vous construisez ainsi votre portfolio professionnel et faites avancer votre carrière.

Pre-work : Préparez votre formation

Pour préparer la formation, vous avez accès au “pre-work,” des cours en ligne sur notre plateforme, entièrement dédiés à la data science :

- Les concepts essentiels de la programmation Python
- Des mathématiques appliqués y compris les statistiques
- Une introduction aux outils de développement et la visualisation des données

Ce travail préparatoire, conçu pour vous initier aux compétences et au contexte des données fondamentales, est d’une durée d’environ 12h entièrement en anglais. Il est à compléter avant le début de la formation. Il vous aidera à partir du bon pied dès le premier jour de cours.

L'expérience d'apprentissage

Devenir, vous aussi, data scientist : suivez, pendant trois mois à temps plein, une formation complète basée sur des projets concrets, et conçus pour inspirer une vie toute entière de découvertes.

Que vous suiviez la formation à distance ou entièrement à distance, en tant qu'étudiant chez GA :

- Vous découvrez de nouveaux concepts et outils à travers des conférences et des discussions menées par des experts
- Vous réalisez des exercices pratiques d'analyse et de programmation pour renforcer les compétences nouvellement acquises
- Vous développez la maîtrise des sujets et des techniques essentiels au secteur lors des travaux pratiques, en binôme et en groupe, et en travaillant sur des projets professionnalisants
- Vous recevez un soutien personnalisé de la part de votre équipe pédagogique d'experts
- Vous créez un portfolio professionnel pour présenter vos compétences prêtes à l'emploi en data science à des employeurs et collaborateurs potentiels

Coaching carrière spécialisé

En tant qu'étudiant dans un programme d'immersion, vous recevrez un soutien spécialisé par des coachs de carrière qui vous aideront à fixer des objectifs, à établir une feuille de route vers le succès, et à rester sur la bonne voie dans votre recherche d'emploi. Tout au long de la formation, en ligne ou sur notre campus, vous :



- Assistez à des interventions d'experts externes qui partagent avec vous leur réalité professionnelle et participez à des ateliers pour vous préparer à l'emploi dans un nouveau secteur
- Apprenez à identifier des opportunités d'emploi et valoriser vos compétences en fonction des offres sélectionnées
- Développez une stratégie de recherche d'emploi en suivant vos progrès grâce à un tableau de bord
- Développez votre personal branding en compilant un portfolio convaincant, en améliorant votre présence en ligne et en face à face, en vous entraînant à résoudre des défis techniques, et à pitcher vos compétences pour vous démarquer lors des entretiens
- Bénéficiez d'un réseau mondial exclusif d'experts, d'influenceurs, et d'anciens élèves, afin de partager des stratégies pour tirer parti de vos connexions existantes, en présentiel et en ligne
- Devenez contributeur actif de la communauté data science en partageant votre travail par le biais de discussions en ligne et de rencontres en personne

Après l'obtention du certificat, vous aurez également accès à des ressources pour vous aider à parfaire votre apprentissage tout au long de votre carrière. Plongez dans de nouveaux sujets ou continuez à perfectionner vos compétences en data science grâce à nos offres promotionnelles sur une suite d'outils, des pass et des forfaits pour des événements de premier plan, et plus encore.

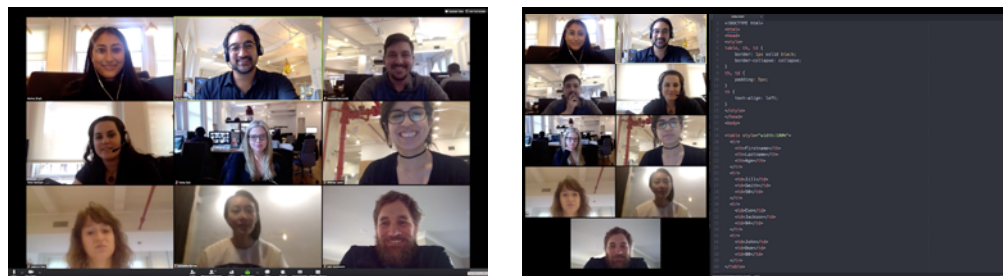
Vous pouvez aussi demander des promotions sur les frais d'inscription aux futures autres formations GA, sur les workshops et les ateliers de GA, que ce soit sur le campus et en ligne.

Si vous suivez la formation entièrement en ligne

Vous vivrez l'expérience campus, depuis chez vous!

Comment? Grâce aux sessions en classe virtuelle, avec tous les participants. Notre expert animera la formation via [Zoom](#), notre outil de visioconférence.

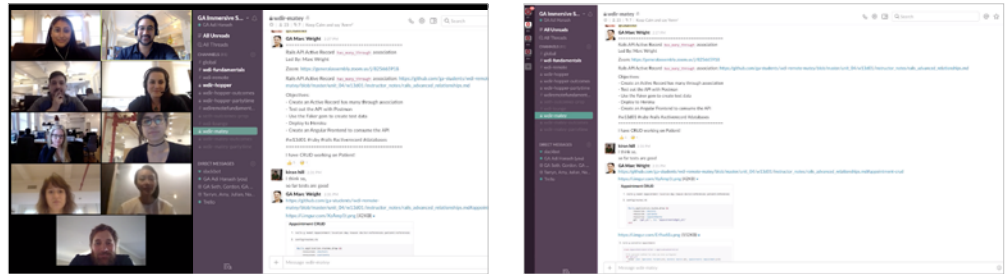
Vous assisterez en temps réel à la formation et pourrez collaborer avec tous les autres participants en classe entière ou en groupe. Ces sessions sont interactives. Vous pourrez produire, interagir, partager votre écran, et présenter vos travaux, travailler en petit groupe via nos classes virtuelles et le chat intégré, pour suivre la formation, comme si vous y étiez. Vous pouvez d'ailleurs visionner les vidéos en replay pour voir, ou revoir, le cours.



Les classes virtuelles en groupe



Grâce à [Slack](#), une application de messagerie téléchargeable sur votre ordinateur et votre mobile, vous pourrez chatter pendant ou en dehors du cours avec votre formateur ou les participants. Cet espace d'échange et de collaboration permettra de partager des documents, discuter, débattre ou même faire un point sur votre avancement.



Le chat collaboratif

En ligne, comme sur nos campus, vous participez à une formation très pratique, basée sur des projets professionnalisants et conçue pour couvrir tous les concepts et outils de base. Notre expert vous apporte toutes les compétences nécessaires pour réaliser pas à pas, votre projet et mobiliser vos compétences. Vous construisez ainsi votre portfolio et faites avancer votre carrière.



Ce que vous apprendrez

Durée de la formation : 480h

Travail préparatoire (pre-work)

Les fondamentaux en data science

Accédez à notre plateforme de cours en ligne et préparez-vous pour la formation. Apprenez les bases de Python appliquées à la data science et aux mathématiques.

- Découvrir les concepts de base de la programmation Python et les types de données, y compris les variables, les listes, les dictionnaires, les boucles, et les fonctions
- Créer des fonctions qui acceptent plusieurs arguments et renvoient plusieurs valeurs
- Comprendre l'objectif des itérateurs et des compréhensions de liste dans les workflows de data science du monde réel
- Décrire l'utilisation et le but des DataFrames et comment ils peuvent être utilisés pour manipuler des données avec Pandas
- Appréhender les visualisations de graphiques avec Matplotlib et Seaborn
- Se familiariser avec les statistiques descriptives et inférentielles et avec la façon de les calculer
- Calculer les combinaisons et permutations
- Découvrir les tests de signification et calculer les intervalles de confiance
- Calculer l'algèbre linéaire et les équations de régression

Module 1 Fondamentaux

Se familiariser avec les outils et techniques essentiels de la science des données (data science) en travaillant dans un environnement de programmation pour collecter, organiser, et partager des projets et des données avec Git et UNIX.

- Maîtriser les concepts de programmation en Python et NumPy et parcourir les sources de données et les collections
- Utiliser les commandes UNIX pour parcourir les systèmes de fichiers et modifier les fichiers
- Suivre les modifications et les itérations à l'aide du gestionnaire de version Git depuis votre terminal
- Définir et appliquer les fondamentaux de la statistique descriptive aux échantillons d'ensembles de données
- Appliquer le tracé graphique et visualiser des données à l'aide de bibliothèques Python telles que Matplotlib et Seaborn

Projet : Nettoyez un ensemble de données en appliquant vos compétences de programmation avec NumPy et Python.



Module 2 **Analyse exploratoire des données**

Générer des analyses visuelles et statistiques en utilisant Python, ses bibliothèques, et les outils associés pour aborder des problèmes dans des domaines tels que la finance, le marketing, et les politiques publiques.

- Concevoir une étude expérimentale avec une problématique bien pensée et un cadre de données
- Lire, nettoyer, analyser, et tracer des données avec Pandas
- Extraire et ré-organiser les données via l'indexation, le regroupement, et la fonction JOIN
- Passer en revue les concepts de tests statistiques (valeurs p, intervalles de confiance, fonctions lambda, corrélation/causalité) avec SciPy et StatsModels
- Récupérer les données du site web à l'aide d'outils de récupération populaires
- Appliquer le bootstrapping, le rééchantillonnage, et la construction d'inférences sur vos données

Projet : Tirez parti de Pandas pour appliquer les compétences avancées de NumPy et Python. Puis, nettoyez, analysez, et testez les données de plusieurs ensembles de données désordonnés.

Module 3 **Modélisation statistique classique**

Concevoir efficacement des études et évaluer des modèles pour les optimiser par la mise en œuvre de modèles de classification et de régression linéaires et logistiques. Collecter et connecter des données externes pour ajouter des nuances à vos modèles à l'aide du web scraping et des API.

- Exécuter des modèles de régression linéaire et logistique grâce à scikit-learn et StatsModels et évaluer l'ajustement du modèle
- Examiner les modèles de classification en implémentant l'algorithme k-plus proches voisins (kNN)
- Gérer le compromis biais-variance tout en vous entraînant à évaluer des modèles statistiques classiques
- Utiliser la sélection de fonctionnalités pour approfondir vos connaissances sur la conception d'études et l'évaluation de modèles
- Appliquer l'optimisation et la régularisation pour les modèles d'ajustement et de réglage
- Plonger dans les mathématiques appliquées et la théorie expliquant la façon dont la descente de gradient aide à optimiser les fonctions de perte pour les modèles d'apprentissage automatique

Projet : Explorez, nettoyez, et modélisez les données en vous basant sur un ensemble de données proposé. Puis, décrivez votre stratégie, en expliquant vos résultats.



Module 4 Modèles d'apprentissage automatique

Créer des modèles d'apprentissage automatique en explorant les différences entre l'apprentissage supervisé et non supervisé via le clustering, le traitement du langage naturel, et les réseaux de neurones.

- Définir le clustering, ses avantages et inconvénients par rapport aux modèles de classification
- Créer et évaluer des modèles d'ensemble à l'aide d'arbres décisionnels, de forêts aléatoires, du bagging, et du boosting
- Se familiariser avec le traitement du langage naturel (NLP) grâce à l'analyse des sentiments à partir des données récupérées de site web
- Découvrir comment Naive Bayes permet de simplifier le processus d'analyse des données avec les algorithmes d'apprentissage supervisé
- Découvrir l'histoire et l'utilisation de Hadoop, ainsi que les avantages et inconvénients de l'utilisation de systèmes parallèles ou distribués pour stocker et analyser les mégadonnées
- Comprendre l'interaction entre Hive et Hadoop
- Identifier les avantages de Spark à travers des études de cas de big data
- Analyser et modéliser les données d'une série temporelle à l'aide du modèle ARIMA

Projet : Récupérez et modélisez vos propres données à l'aide de plusieurs méthodes, en décrivant votre approche, et en évaluant les risques ou les limites.

Module 5 Sujets complexes et tendances

Plonger au cœur des systèmes de recommandation, des réseaux de neurones, et les modèles de vision par ordinateur, en mettant en œuvre ce qui a été appris pour produire des modèles.

- Comparer les différents types de réseaux de neurones et montrer comment ils sont adaptés à la propagation arrière
- Créer et appliquer des systèmes de recommandation de base afin de prédire des échantillons de données utilisateurs
- S'entraîner avec des études de cas en data science pour vous préparer aux entretiens d'embauche

Projet final : Choisissez un ensemble de données à explorer et à modéliser, en fournissant des notes détaillées de votre approche technique. Puis, construisez une présentation visuelle pour soutenir vos résultats à l'oral.



Questions fréquemment posées

Pourquoi cette formation en data science est-elle pertinente aujourd'hui ?

Les individus et les organisations de tous horizons exploitent les données pour relever les grands défis d'aujourd'hui, allant des politiques publiques et de la robotique, aux sites de rencontres et au eCommerce. En conséquence, les organisations évoluent rapidement pour créer des équipes de gestion des données internes fiables et prendre de l'avance dans ce secteur en évolution. Et il n'y a pas assez de talents pour combler tous les postes vacants.

Selon [Forbes](#), « Plus de 150 zettaoctets (150 billions de gigaoctets) de données devront être analysés d'ici 2025. » Comblez l'écart et ouvrez les portes d'un monde d'opportunités.

Quel type de communauté vais-je rencontrer lors de cette formation de trois mois ?

Notre formation Data Science Immersive attire des étudiants qui sont aussi passionnés que vous par le développement et le lancement d'une nouvelle carrière dans le domaine de la data science. Ils sont généralement issus du design, du product management, de la data, et de nombreux autres domaines. L'expérience chez General Assembly peut créer des liens, des collaborations et des opportunités professionnels durables qui vous accompagneront tout au long de votre carrière.

Quelles sont les compétences évaluées ?

À la fin de notre formation Data Science Immersive, vous serez en mesure de :

- Collecter, extraire, interroger, nettoyer, et agréger des données pour analyse
- Effectuer une analyse visuelle et statistique des données à l'aide d'UNIX, de Git, de SQL, de Python, et de leurs bibliothèques et outils associés
- Créer et implémenter les modèles d'apprentissage automatique et les algorithmes appropriés pour évaluer des problèmes de data science couvrant la finance, les politiques publiques, etc.
- Créer et partager des récits convaincants grâce à la visualisation de données
- Compiler des rapports clairs des partenaires pour communiquer les nuances de vos analyses
- Identifier les problèmes de big data et expliquer comment les systèmes distribués et les technologies informatiques parallèles permettent de résoudre ces difficultés
- Appliquer des processus de résolution de problèmes en matière d'interrogation, de modélisation, et de validation à des ensembles de données provenant de divers secteurs pour donner un aperçu de problèmes et de solutions tirés du monde réel

Comment se déroule la formation et que couvrent les frais d'inscription ?

Suivre une formation chez General Assembly, c'est accéder à de nombreux services.

- Une formation dirigée par des experts, sur les compétences, les méthodes, et les meilleures pratiques appliquées à la data science en campus ou entièrement à distance (en classe virtuelle)



- Un accès à notre plateforme de cours en ligne myGA pour effectuer le travail préparatoire pre-work afin d'actualiser et d'affiner vos connaissances tout au long de votre parcours professionnel dès que nécessaire
- Un accompagnement personnalisé par notre équipe pédagogique (expert et équipe de student experience) pendant les heures d'accueil
- Un accès à notre outil de communication collaboratif, Slack, avant, pendant, et après la formation pour échanger avec notre communauté d'étudiants, d'experts, et d'alumni, et retrouver ainsi de nombreux conseils, évènements, ou opportunités même après votre formation
- Une attestation et un certificat GA pour mettre en valeur vos nouvelles compétences sur LinkedIn
- Un accompagnement préparatoire à la recherche d'emploi : CV, entretiens d'embauche, simulations d'interview...
- L'accompagnement post-formation par un coach de carrière pour vous orienter lors de votre recherche d'emploi : conseils, soutien, défis techniques, négociation salariale, et bien plus encore
- Un accès exclusif aux avantages en tant qu'alumni : promotions, invitations aux événements de réseautage, et aux ateliers de développement professionnel pour booster votre carrière

Quelles sont les conditions d'admission et pré-requis?

Il s'agit d'une formation de niveau intermédiaire qui nécessite des pré-requis :

- Un niveau Bac S en mathématiques et des notions de programmation en Python sont requis. Certains étudiants ont par exemple un diplôme en mathématiques, en informatique, ou encore une expérience professionnelle dans la data. D'autres, sont autodidactes.
- Valider le test d'entrée fourni par notre équipe d'admission
- Avoir un ordinateur portable PC ou Mac (le système d'exploitation doit être le plus récent - moins de 4 ans - ou mis à jour avant la formation)
- Si la formation se déroule à distance, une connexion à Internet, webcam et des écouteurs sont nécessaires.

Lorsque la formation est délivrée en anglais, il est recommandé de :

- Disposer d'un niveau B2 en anglais (CECRL) ou un score au TOEFL d'au moins 90. En effet, en fonction de votre choix et des disponibilités par notre campus, la formation peut être entièrement en anglais ou en français avec des supports pédagogiques en anglais. Vous pouvez vous auto-évaluer à partir de ces grilles ([description en français](#) / [description en anglais](#))

N'hésitez pas à contacter notre [équipe d'admission](#) (ou écrivez directement à paris-inscription@generalassemb.ly) pour faire le point sur votre parcours et vos objectifs afin de vérifier si cette formation est adaptée à vos besoins.

Comment valider la formation et obtenir le certificat ?

À la fin de la formation, si vous remplissez les conditions de validation de la formation, vous obtenez une attestation de réussite et son certificat. Pour valider la formation, il vous faut :

- Être assidu aux journées de cours



- Réaliser et soutenir vos projets face au formateur. Pour attester de la maîtrise de l'ensemble des compétences visées, vous devez obtenir un score de min. 66% pour chacun des projets de la formation.

Le respect de ces deux critères donne accès à la validation de la formation et à son certificat.

Des milliers d'anciens étudiants de GA utilisent leur certification pour démontrer leurs compétences aux employeurs et à leurs réseaux sur LinkedIn. La formation Data Science Immersive est appréciée et reconnue par de nombreuses grandes entreprises qui participent à la conception de notre programme et/ou l'utilisent afin de former leurs propres équipes.

Quel est le lien entre la formation Data Science Immersive et les autres formations en data de General Assembly ?

Cette formation Data Science Immersive est un bootcamp de trois mois intensif qui vous prépare au métier de data scientist ou tout autre fonction en data analyse. Elle permet par exemple de préparer une reconversion professionnelle.

Pour les professionnels qui souhaitent approfondir et perfectionner leur compétences en sciences de la données, nous proposons la formation courte [Data Science](#). Si vous souhaitez simplement une introduction aux compétences de base en analyse de donnée, nous vous proposons la formation [Data Analytics](#). Vous apprendrez à analyser la donnée avec Excel, SQL et Tableau.

Dans quel format puis-je suivre ce cours : sur le campus ou en ligne ?

Cela dépend de vous ! Nos formations à distance offrent une expérience d'apprentissage qui reflète l'offre de GA sur le campus, tout en vous permettant d'apprendre chez vous, confortablement. Si vous ne vivez pas à proximité d'un campus GA et que vous avez un agenda chargé, ou si vous souhaitez simplement vous épargner le trajet, la formation à distance pourrait être la bonne solution pour vous. Vous aurez toujours accès aux conseils d'experts, aux ressources d'apprentissage, et au réseau de soutien qui font la réputation de GA.

Si vous préférez apprendre aux côtés de vos pairs et avez la possibilité de vous rendre sur le campus, nos formations en présentiel vous permettront de profiter du confort de nos campus.

Notre équipe admission (à contacter via ce [formulaire](#) ou écrivez directement à paris-inscription@generalassemb.ly) peut vous conseiller sur le meilleur format en fonction de votre situation personnelle et de votre mode d'apprentissage.

En quoi consiste le service carrière pour les étudiants à distance ?

Ce sont les mêmes engagements que pour nos étudiants sur le campus ! Découvrez l'univers de la data science via des conversations quotidiennes, des débats d'experts sur le recrutement, des conférenciers spécialisés, etc. Vous travaillez individuellement avec votre coach carrière pour comprendre votre marché du travail local, trouver en emploi, et créer des liens avec la communauté de data scientists de votre région.

Notre équipe d'admission pourra vous en dire davantage sur l'accompagnement dont vous bénéficierez pour décrocher un poste en tant que data scientist.



Dans quelle langue a lieu la formation?

Nous proposons la formation sous les deux formes suivantes :

- Entièrement en anglais
- En français, sur la base de supports pédagogiques en anglais

Dans les deux cas, nous vous recommandons d'avoir le niveau suivant en anglais :

- Un niveau B2 en anglais (CECRL) ou un score au TOEFL d'au moins 90. En effet, en fonction de votre choix, la formation peut être entièrement en anglais ou en français avec des supports pédagogiques en anglais. Vous pouvez vous auto-évaluer à partir de ces grilles ([description en français](#) / [description en anglais](#))

Quelles sont mes options de financement ?

Nous proposons différentes [options de financement](#), y compris des programmes de paiements échelonnés, afin que vous puissiez vous concentrer sur ce qui compte : votre formation.

Nos formations ne sont pour le moment pas finançables par le CPF, mais nous travaillons activement à cela. N'hésitez pas cependant à contacter votre entreprise pour savoir s'ils peuvent mobiliser leurs fonds personnels pour vous aider à financer la formation, ou contacter votre conseiller Pôle Emploi si vous êtes demandeur d'emploi.

En effet, de nombreuses structures telles que votre région, Pôle Emploi ou l'AGEFIPH peuvent vous aider à financer votre projet de formation. Votre conseiller Pôle Emploi pourra vous orienter vers ces différents dispositifs, ou vous proposer une Aide Individuelle à la Formation (AIF).

Discutez avec notre [équipe d'admission](#) pour en savoir plus et trouver la solution la mieux adaptée à vous et vos finances.



General Assembly est [Datadocké](#) et répond aux critères qualités énoncés par le décret qualité.





Passez à la prochaine étape

Vous avez des questions sur notre formation “Data Science Immersive” ? Notre équipe d’admission est là pour vous aider à déterminer si cette formation est faite pour vous et adaptée à vos objectifs.

Vous pouvez aussi :

- Assister à une séance d’information [en ligne](#) ou sur [votre campus local](#)
- Découvrir vos [options de financement](#)
- [Envoyer votre candidature](#) pour vous inscrire au cursus*
- Nous contacter par email à paris-inscription@generalassemb.ly ou par téléphone au +33 1.88.24.42.51.

Délai d’accès de la formation

Les inscriptions peuvent avoir lieu jusqu’à 14 jours avant le début de la formation. Contactez-nous via ce [formulaire](#).

Dates et prix de la formation

Nous proposons régulièrement des sessions de formation : en moyenne, 4 sessions de formation par an. Pour obtenir les dates et le tarif, consultez directement notre [site web](#) ou contactez notre équipe Admission.

** Les options de modalités de formation varient selon le lieu géographique, dans l’attente de la disponibilité du marché. N’hésitez pas à contacter notre équipe d’admission pour savoir quelles modalités sont disponibles dans votre région.*

Accessibilité aux personnes en situation de handicap

General Assembly soutient la diversité, l’équité et l’inclusion des personnes en situation de handicap. Pour cela, nos formations sont disponibles sous différentes modalités (à distance ou en présentiel). Notre équipe peut vous proposer une assistance spécifique sur demande et étude de vos besoins.

Contactez notre équipe Admission afin de faire le point sur vos besoins et discuter avec notre représentant au handicap sur la meilleure façon de vous accompagner lors de votre expérience d’apprentissage.