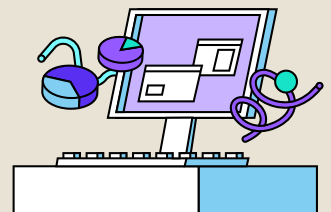




SYLLABUS MLOps



DISCOVER THE DIFFERENT JOBS IN DATA

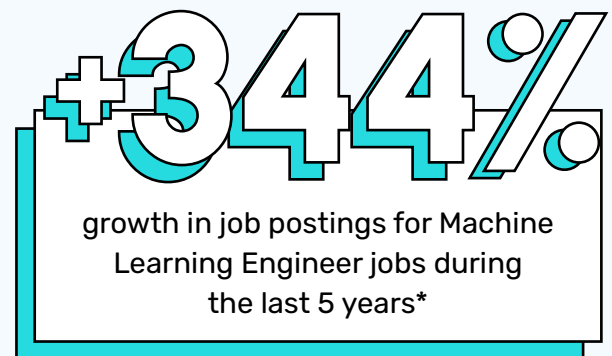
OUR STORY

DataScientest is the leading training organization in Europe for Data Science. We offer business-oriented training courses for upskilling and reskilling to professionals and individuals alike.

TRAINING DESIGNED WITH AND FOR COMPANIES

Data has become the main resource to exploit for companies. It allows them to **guide their decision-making** and to predict the behaviour of their environment. Therefore, staying **competitive** means exploiting data to its fullest potential. Thus, given that data is everywhere, it is essential to **increase workforce capabilities** on these new skills, tools and technologies to ensure stable development.

This is why DataScientest has been training the Data teams of more than **70 Fortune 500 groups** as well as many other SMEs and Startups for over **6 years**. Since 2021, more than **5500 professionals** have been trained with us.



*Source - Indeed Data

A CURRICULUM ADAPTED TO INDIVIDUALS

Having been designed for our **partner companies**, our training courses are professionally oriented, as close as possible to the **needs of the market**. Thus, in order to face the shortage of Data profiles and bridge this gap, we have decided to open our training courses to **individuals**. Under this impulse, DataScientest is constantly developing thanks to **new training courses** (Data Manager) or **new specialized courses** (Deep Learning).



"DataScientest is a rich, personalized learning experience that responds to the real problems of the company."

Samir Ait Idir
Chief Data Officer & Head of Data Intelligence @Orange Bank

DATASCIENTEST KEY FIGURES

+70

**FORTUNE 500 COMPANIES
AS CLIENTS**

98%

**COMPLETION
RATE**

94%

**SATISFACTION
RATE**

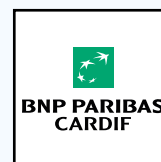
+5500

ALUMNI

2000h

**OF EXCLUSIVE
CONTENT**

THEY TRUST US





OUR PARTNERS

DataScientest has developed partnerships with academic institutions on the one hand and with software publishers on the other. The content of our training courses on Business Intelligence or Cloud tools is henceforth certified by the companies that own them. Here are our different partnerships:



AMAZON WEB SERVICES
Editor partner

Today, DataScientest enjoys the exclusive status of **Amazon Training Partner**. We are therefore authorized by Amazon to train teams on the products and services of the American company. Thanks to this partnership, we have established **several training courses** that prepare for **official AWS certifications**. The registration fee for the official exam is included in the price of the course.

In some of our courses, you will be prepared to take the fundamental **AWS certification: Cloud Practitioner**.



MICROSOFT - POWER BI
Editor partner

DataScientest is a **Microsoft Learning Partner**, which means that we can train you for **official Microsoft certifications**. These certifications attest to a level of expertise in Azure, the set of cloud computing products and services, and in **Power BI**, Microsoft's business intelligence tool. The registration fee for the official exam is included in the price of the course.

In each training courses, DataScientest prepares you to pass official Microsoft certifications. Additionally, with the **Data Management** course, DataScientest allows you to prepare for the **PL-900** certification, **Microsoft Certified Power Platform Fundamentals**.



HEC DATA MINDS
Academic partner

Since October 2020, DataScientest has been providing **HEC Paris** students with access to its platform to master coding and analytical problem-solving through data science.



QUALIOPI
Quality criteria

This label recognises that the processes implemented by the training organization meet the quality requirements demanded by the State. **All DataScientest training courses are Qualiopi certified**.

OUR TEAM

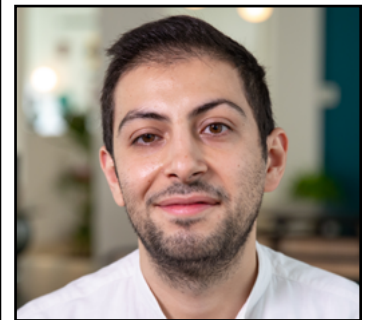
OUR TEACHERS

DataScientest's teaching team is made up of internal professors. They are dedicated to teaching and research for our various courses and expert curricula.

All our professors are always there for you. They have created and continue to update all our courses. They will accompany you by video-conference and on our platform throughout your training.

With their excellent academic background and varied professional experience, they are the Data Science experts who will enable you to join the Data team in the sector that interests you most (Banking & Insurance, Industry, Finance, Medical and many others).

Charles S.
Head of Academic
8 years of expérience



A graduate of the École Polytechnique, Charles has been involved in the development of the courses from day one. He is specialized in programming, Machine Learning and Deep Learning.

Raphael K.
Pedagogical Director
8 years of expérience



Holder of the ISF master's degree specializing in Statistical Learning and Data Science from the University of Paris-Dauphine, Raphaël designed the Data Analyst and Data Management courses thanks to his knowledge in programming, dataviz and Machine Learning.

OUR TEAM

Gaspard G.
Data Scientist training leader
2 years experience



With a research Master's degree in Mathematics and a Master's degree in Data Science from the Ecole Normale Supérieure and the Université Paris-Dauphine, Gaspard rapidly joined DataScientest. Specialized in Computer Vision and NLP, he participated in the development of our expert courses and supervised many Deep Learning projects. He is currently in charge of the Data Scientist training.

Laurene B.
MLOps expert course leader
3 years experience



Holder of a Master's degree in IASD from the University Paris-Dauphine PSL (in partnership with Les Mines and Normale Sup), Laurene has been a member of the DataScientest team for over a year. In charge of developing the MLOps curriculum, she is specialized in Machine Learning and Deep Learning.



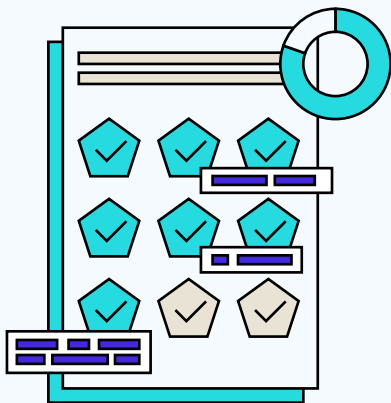
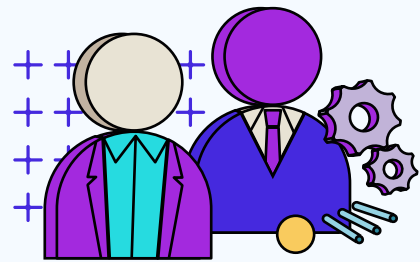
During your training, you will progress on our platform. However, **you will never be alone**, Daniel is there for you. All our teachers and Program Managers take turns to **answer all your questions live via Slack** and on a dedicated support platform. With support every day of the week, from 9am to 5pm, you are guided throughout your course.

OUR PEDAGOGY

HYBRID FORMAT

DataScientest offers 100% distance learning in a hybrid format :

During approximately 20% of the duration, you are accompanied by your teacher and your cohort during **coaching sessions via videoconference**. The rest of the time, you work on our teaching platform and are **guided** via Slack by our **Data Scientists**. This format gives you the flexibility to organize yourself. We also monitor your progress and accompany you to ensure the **successful completion** of your training.



MASTERCLASS

Each training period on the platform (called a Sprint) is accompanied by one or more **masterclasses by videoconference**. The objectives and teaching formats of each masterclass change according to the sprint: correction of concrete use cases, lectures on specific topics, competitions between cohorts, etc.

EXAMS

At DataScientest, there are no automated MCQs. **Each paper is reviewed and corrected** by hand by our teachers. They go over all your difficulties with you.

BACKBONE PROJECT

From the beginning of your training, you will carry out a concrete project with the aim to deploy it. It requires an investment of about 60 hours of work throughout the training.



You will work in pairs or in triads on a project that you will select from our regularly updated **project catalogue**. Our subjects are inspired by the work we do with companies.

This is a crucial step in your career path that will make you fully **operational**. You will be working on un-cleaned data sets, and will be expected to produce work of professional quality.

Reviews are done at regular intervals by your teacher to guide you, and **coach you**. This allows you to move efficiently from theory to practice and to ensure that you master the skills required on the different modules.

It is also a project that is **highly valued by companies**. It confirms your skills and knowledge acquired at the end of the **Data Analyst** course. You will then be able to justify your skills with a **successful Data Science project** during your interviews



DataScientest is a rich, personalized course, perfectly adapted to today's corporate data science topics.

Xavier Bocher
Head of Credit Risk Internal Models & Operational Research @Groupe Crédit Agricole

OUR PEDAGOGY

PREREQUISITES

To start a MLOps training, the minimum requirement is a first Data Scientist training from a recognized training organization, or 2 years of professional experience as a Data Scientist in a company.

CONTINUOUS

Only one format is available for the MLOps training :



CONTINUOUS FORMAT

Flexible, it is the appropriate format
for the practice of a profession

DURATION 4 months

RYTHME 8-10h / week

CURRICULUM

MLOps

1 ADVANCED PROGRAMMING



Webscraping

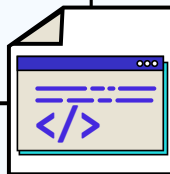


Linus & Bash Programming



Git/Github

Unit testing



2 DATAOPS - ISOLATION

FastAPI



Securing APIs

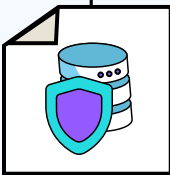
Docker



Flask



Bootstraps



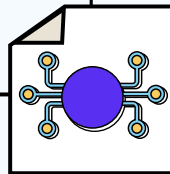
3 DATAOPS - ORCHESTRATION



Kubernetes



AirFlow



4 MODELOPS

MLFlow

Data acculturation



ADVANCED PROGRAMMING - duration 26h

Web Scraping

- Introduction to Web Scraping
- Introduction to web languages
- Web Scraping with BeautifulSoup library
- Web Scraping of Google searches

Linux System and Bash Script

- Presentation of Linux System
- Handling and use of a terminal
- Implementation of Bash scripts

Git and Github

- Introduction to the Git control system
 - Setting of a Git account
 - Discovery of the branch principle and the systems that result from it
- Discovering collaborative work on Git and Github
- Management of Git projects with the repository system
- File and project sharing with the push function
- Updating your local repository with the pull
- Participation in the improvement of public projects
- Introduction to workflows and their automation with Github Actions

Unit Testing

- Setting up unit tests with Pytest
- Introduction to integration tests and their functions
- Presentation of the benefits of testing :
 - Time saving
 - Legibility
 - Quality
 - Code improvement

Skills acquired

- Master the Linux operating system
- Learn to use a terminal
- Create and manage Bash executables
- Master the versioning tools
- Cooperate and save projects with Git and Github
- Automate workflows with Github Actions
- Understand how workflows work in general and CI/CD in particular
- Be able to manage unit tests
- Verify the functioning of independent code units during development
- Master the BeautifulSoup library for Web Scraping

DATA OPS - ISOLATION - duration 32h

API

- Introduction to APIs and discovery of microservices architectures
- Presentation of the different HTTP methods and their functions
- Use of the FastAPI and Flask libraries to develop APIs
- Documentation of an API with the OpenAPI specification
- Error and performance management of an API
- Implementing security in APIs with Flask and FastAPI

Docker

- Presentation of containerization and its usefulness in relation to virtualization
- Introduction to the functioning of Docker
- Handling images and containers
- Communication with containers
- Data persistence thanks to volumes
- Creating a Docker image via Dockerfile
- Sharing images on the Dockerhub
- Use of docker-compose

Securing APIs

- Introduction to API security and API Keys
- Mastery of HTTP Basic authentication
- Introduction to JSON Web Token and HTTPS

Skills acquired

- Understanding APIs
- Learn how to create an API with Flask and FastAPI
- Requesting an HTTP API
- Understand virtualization
- Master containerization and container orchestration techniques and tools

DATA OPS - ORCHESTRATION - duration 20h

Kubernetes

- Deploy and manage containers
- Orchestrate multiple services and manage scalability

Airflow

- Discovery of Airflow concepts
- Presentation of the principles of orchestration and their utility
- Directed Acyclic Graphs (DAG)
- Operators
- Task management through specific Operators
- Monitoring of DAGs via the Airflow graphical interface

Skills acquired

- Understand the principles of orchestration
- Automate specific tasks with Apache Airflow
- Manage scalability with Kubernetes



MODELOPS - duration 12h

MLFlow

- Lifecycle of a Machine Learning project
- Data Preparation
- Model training
- Model deploying
- Model serving
- Presentation of MLFlow architecture
 - MLFlow Tracking
 - MLFlow Projects
 - MLFlow Models
 - MLFlow Registry

Data Acculturation (optional)

- GDPR and ethics in data
- The data professions
- Managing a data team
- Definitions of the most used terms in data

Skills acquired

- Master the entire life cycle of a Machine Learning project
- Master the MLFlow architecture

ALUMNI TESTIMONIES



Adrien MOREAU

Data scientist at Siderlog Conseil

I found the training very interesting. In the first instance, I think it is necessary to mention the content, very rich and accessible. The 75% autonomy formula is also a format that I found particularly beneficial. By being supervised and «pushed» to pass the certifications to reach specific milestones, it maintains a rhythm.



Karina CASTILLO

Data Expert Consultant at Expertime



I followed the Data Scientist bootcamp, and I must say that even if it is demanding, the teachers are very friendly and always listening, in addition to being very professional. I would also like to underline the fact that they are always up to date with the technologies and topics related to the data business in general. The training combines perfectly both theory and practice throughout each step of the way.

One of their strong points is that they understand very well the needs of companies and it shows in the content of their training. Finally, once I finished the training, the career department accompanied me throughout the year by giving me advice for interviews, my CV, salary ranges and even proposed me as a candidate for a job offer which means I was able to land a job as a full-time Data Expert Consultant. I fully recommend DataScientest's training courses.



Cyrille Chemama



I went through the Data Scientist training in bootcamp mode in March 2021 and despite my amateurism in the field of computer science and programming languages, the curriculum is very educational and covers all the concepts necessary for the position of data scientist: from learning Python to data acquisition, especially via web scraping, to the most advanced skills in machine learning, deep learning and reinforcement learning.

Datascientest's learning platform allows you to work at your own pace and even in case of difficulty, the support and hotline are always available to help you out. Glad I signed up!



Emeline SILVESTRE



Datascientest allows me today to realize my dream, to become a Data Scientist!

The continuous training model allows you to progress at your own pace and whenever I wanted. The courses available on the online platform are very educational. You gain in skill little by little on complex concepts. We have access to many resources to help us, a forum, an instant messaging system and the possibility of making regular calls with the supervisors. We also have a class with whom we share, in particular, the project. In addition to the training itself, many events are organized, from meditation sessions to CV workshops, as well as data workshops, which are very enriching.

For me, with their learning method, this is the best way to combine fun and effectiveness. You learn thanks to concrete examples, through experiences and mistakes.

TO GO FURTHER

If you wish to reinforce your skills, DataScientest has developed various expert courses as well as certifications from publishers such as Microsoft or AWS to enable you to deepen your knowledge and improve your skills in data!

DEEP LEARNING

Are you interested in how Siri works? How autonomous cars are developed? This course is for you! Through two specialization tracks, you will learn all about language processing with Natural Language Processing (NLP) technologies, or image processing with Computer Vision.

SOFTWARE CERTIFICATION

You can specialize by preparing for an AWS technical certification or by training on Azure :

ARCHITECTING ON AWS

If you want to know how to make architectural decisions on the AWS cloud, this is the certification for you! Earn the **«AWS Certified Solutions Architect Associate»** status.

DEVELOPING ON AWS

If you are looking for the ability to develop and implement cloud-based applications, look no further! This training is for you. Upon completion, you will be certified as an **«AWS Certified Developer Associate»**.

AZURE DATA ENGINEER

If you want to master the implementation of Azure solutions in an organization, this training is for you! At the end of the training, pass the official Microsoft certification and become «**Azure Data Engineer Associate**».

AZURE ADMINISTRATOR

If you are interested in gaining skills and knowledge about managing storage and virtual networks on Azure, this training is what you are looking for. Get trained and earn the official «**Microsoft Certified Azure Administrator Associate**» certification.

POWER BI

Do you want to be able to provide a complete analysis of a dataset and improve your dashboard creation skills? This training is for you! Learn to master Power BI and earn your official Microsoft certification by becoming a «**Power BI Data Analyst Associate**».



**YOU WANT TO BECOME A
MACHINE LEARNING ENGINEER ?**

