

# **DIPLÔME D'UNIVERSITÉ SORBONNE DATA ANALYTICS (FORMATION INITIALE)**

---

Sorbonne Data Analytics est un diplôme universitaire de niveau Master, géré par la Mention "Econométrie-Statistiques" de l'Ecole d'Economie de la Sorbonne. Son objet est de former aux Sciences des Données (DataScience). **Sa particularité est d'être essentiellement en distanciel, mais avec un tutorat individualisé et renforcé.**

Sorbonne Data Analytics permet donc d'acquérir, de manière souple, un solide complément de formation en sciences des données. Il s'adresse à des personnes souhaitant se réorienter ou se renforcer dans le domaine de la data, mais qui ne peuvent le faire dans le cadre d'une formation classique nécessitant un présentiel important.

Le langage de programmation utilisé est le langage Python. La formation reprend le triptyque classique d'une formation en Sciences des Données, à savoir : *Mathématiques-Informatique-Statistiques*. Cependant parce que la formation se déroule principalement en distanciel (avec un tutorat individualisé et renforcé), l'apprentissage par les participants de Sorbonne Data Analytics des outils et méthodes des sciences des données se fera via des projets pratiques dans chaque enseignement. Par ailleurs, un mémoire de fin d'étude dont le sujet est validé par la direction pédagogique de Sorbonne Data Analytics, doit être soutenu. Ce qui renforce encore plus l'acquisition de compétences. Ce mémoire peut être un rapport de stage si le participant souhaite en faire un, ou un sujet portant sur une problématique de l'entreprise dans laquelle il/elle travaille. De plus, les participants à Sorbonne Data Analytics doivent obligatoirement réussir trois parmi les cinq certifications Microsoft Azure suivantes : DP900, DP100, DP203, AI100, AZ104. Enfin, chaque participant disposera d'un accès illimité à une machine virtuelle sur laquelle seront installés tous les logiciels nécessaires.

## **Infos pratiques**

**Composante :** Formation continue Panthéon-Sorbonne (FCPS)

**Campus :** Centre Broca (FCPS & IREST)

**Campus :** Maison des Sciences Économiques

### En savoir plus :

## Communiqué de Presse du Partenariat avec Microsoft France

[https://economie.panthéonsorbonne.fr/sites/default/files/inline-files/20210208\\_CP\\_SDA\\_Parteneriat\\_MicrosoftFrance\\_EES\\_FR.pdf](https://economie.panthéonsorbonne.fr/sites/default/files/inline-files/20210208_CP_SDA_Parteneriat_MicrosoftFrance_EES_FR.pdf)

## Emploi du temps de la session d'Avril

<https://calendar.google.com/calendar/u/0?cid=c29yYm9ubmVkYXRhYW5hbHI0aWNzQGdtYWlsLmNvbQ>

## Présentation

Sorbonne Data Analytics est un diplôme universitaire de niveau Master, géré par la Mention "Econométrie-Statistiques" de l'Ecole d'Economie de la Sorbonne. Son objet est de former aux Sciences des Données (DataScience). **Sa particularité est d'être essentiellement en distanciel, mais avec un tutorat individualisé et renforcé.**

Sorbonne Data Analytics permet donc d'acquérir, de manière souple, un solide complément de formation en sciences des données. Il s'adresse à des personnes souhaitant se réorienter ou se renforcer dans le domaine de la data, mais qui ne peuvent le faire dans le cadre d'une formation classique nécessitant un présentiel important.

Le langage de programmation utilisé est le langage Python. La formation reprend le triptyque classique d'une formation en Sciences des Données, à savoir : *Mathématiques-Informatique-Statistiques*. Cependant parce que la formation se déroule principalement en distanciel (avec un tutorat individualisé et renforcé), l'apprentissage par les participants de Sorbonne Data Analytics des outils et méthodes des sciences des données se fera via des projets pratiques dans chaque enseignement. Par ailleurs, un mémoire de fin d'étude dont le sujet est validé par la direction pédagogique de Sorbonne Data Analytics, doit être soutenu. Ce qui renforce encore plus l'acquisition de compétences. Ce mémoire peut être un rapport de stage si le participant souhaite en faire un, ou un sujet portant sur une problématique de l'entreprise dans laquelle il/elle travaille. De plus, les participants à Sorbonne Data Analytics doivent obligatoirement réussir trois parmi les cinq certifications Microsoft Azure suivantes : DP900,

DP100, DP203, AI100, AZ104. Enfin, chaque participant disposera d'un accès illimité à une machine virtuelle sur laquelle seront installés tous les logiciels nécessaires.

## Objectifs

Former aux outils et méthodes des sciences des données

## Savoir faire et compétences

- \* Programmer en python
- \* Traiter des problématiques liées aux sciences des données
- \* Traiter des problématiques liées à l'ingénierie des données

## Les + de la formation

- \* Solide formation en sciences des données
- \* Triptyque "Mathématiques-Informatique-Statistiques"
- \* Acquisition de connaissances via plusieurs projets pratiques
- \* Mémoire de fin d'étude portant sur un sujet pratique
- \* Accompagnement pour passer cinq Certifications Microsoft : Bases de données sur Azure (DP900), Solutions Data Science sur Azure (DP100), Solutions Data Engineering sur Azure (DP203), Intelligence Artificielle sur Azure (AI100), Administrateur Azure (AZ104)

## Organisation

### Contrôle des connaissances

- \* Une unité d'enseignement est acquise par le suivi de toutes les matières incluses dans cette unité d'enseignement et la réussite aux évaluations soumises aux étudiants en contrôle continu.
- \* Une matière (dans une unité d'enseignement) est validée si l'étudiant y a obtenu la note de 10/20.
- \* Une unité d'enseignement est validée si l'étudiant y a obtenu la moyenne de 10/20. Les matières dans une unité d'enseignement se compensent.
- \* Un semestre est validé si l'étudiant y a obtenu la moyenne de 10/20. Les unités d'enseignement dans un semestre se compensent, hors « mémoire-projet général et stage ».
- \* Pour le calcul de la moyenne, il s'applique conformément à la maquette.

### Aménagements particuliers

Il est possible pour le participant de ne suivre qu'une partie de la formation. Le participant obtiendra alors une Certification Sorbonne dans le domaine considéré, et non tout le diplôme. Trois certificats sont disponibles. Le coût de chacune est de 800 euro si le participant est inscrit sous le régime de la formation initiale et résidant en France. Ce coût est de 1540 euro sinon.

- \* Certification "Sorbonne Machine Learning" : requiert la réussite à Machine Learning 1, Machine Learning 2 et à AI100
- \* Certification "Sorbonne Deep Learning" : requiert la réussite à Deep Learning et à AI100
- \* Certification "Sorbonne Solutions Cloud" : requiert la réussite à Distributed and Large Scale Computing, à DP203 et à AZ104

## Stages

**Stage :** Facultatif

**Durée du stage :** 4 mois minimum

**Stage à l'étranger :** Facultatif

**Durée du stage à l'étranger :** 4 mois minimum

Un mémoire est obligatoire dans le cadre de Sorbonne Data Analytics. Cependant ce mémoire peut prendre la forme d'un rapport de stage si le participant au programme souhaite en faire un.

## Membres de l'équipe pédagogique

- \* Responsables : [✉ Marc-Arthur DIAYE](#), [✉ Rania KAFFEL](#)
- \* Responsable Adjoint : [✉ Aryan RAZAGHI](#)
- \* Community Manager : [✉ Lina TALEB](#)
- \* Contact : [✉ sorbonnedataanalytics@univ-parisl.fr](mailto:sorbonnedataanalytics@univ-parisl.fr)

## Admission

### Conditions d'accès

Les candidats doivent justifier :

-soit d'un diplôme national conférant le grade de licence dans un domaine compatible avec celui du diplôme universitaire "Sorbonne Data Analytics";

-soit d'une des validations prévues aux articles L 613-3 (validation d'études supérieures accomplies en France ou à l'étranger), L 613-4 (validation des acquis de l'expérience) et L 613-5 (validation d'acquis) du code de l'éducation.

La procédure de sélection des candidats est effectuée par le jury d'admission sur dossier, sur tests et sur entretien. L'admission est prononcée par l'université sur proposition du jury d'admission.

## Modalités d'inscription

Les inscriptions se dérouleront à partir du lundi 26 avril sur [✉ l'application eCandidat de l'université Paris 1](#) (Voir dans l'offre de formation, la rubrique "Candidatures Economie : diplôme d'université", puis la rubrique "Diplôme d'université de niveau M1").

L'inscription administrative est annuelle et s'effectue avant le début de chaque session de formation.

Par ailleurs, Sorbonne Data Analytics est en grande partie en ligne, mais avec un suivi en parti en présentiel. De ce fait, les étudiants résidant à l'étranger ne peuvent s'y inscrire, sauf s'il existe une convention de prestation spécifique avec une institution de formation dans le pays de résidence, pour leur suivi.

Enfin, le redoublement de la formation peut être accordé par décision du jury, à titre exceptionnel, aux étudiants qui n'auraient pas validé l'intégralité des enseignements. Les étudiants devront alors suivre une nouvelle fois les matières dans lesquelles ils n'auraient pas obtenu au moins 10/20. Les frais de scolarité sont cependant entièrement dus pour la nouvelle année scolaire.

## Tarifs

- \* **1000 Euro en formation initiale (résidence France)**
- \* 4600 Euro en formation continue ou en formation initiale avec résidence à l'étranger
- \* 4 bourses sur critères sociaux sont offertes par an

## Pré-requis nécessaires

- \* Notions d'algèbre matricielle
- \* Statistique inférentielle de base

- \* Probabilité de base

### **Pré-requis recommandés**

- \* Des notions de programmation quel que soit le langage sont un plus
- \* Des notions en bases de données relationnelles sont un plus

# Programme

## Organisation

Le diplôme d'université "Sorbonne Data Analytics" est organisé sous la forme d'un parcours type de formation formant un ensemble cohérent d'unités d'enseignement et organisant une progression pédagogique adaptée, au regard des finalités du diplôme. En l'espèce, les sciences des données.

La préparation s'effectue sur onze mois (dont quatre à cinq mois de préparation du mémoire ou stage).

Une session de formation supplémentaire peut être, en fonction de la demande, organisée pendant l'année scolaire. La formation comprend des enseignements théoriques, méthodologiques et appliqués. L'offre pédagogique est composée de séquences essentiellement "à distance". Certaines séquences de coaching pouvant toutefois être en présentiel physique. Les techniques pédagogiques utilisées pour les séquences sont aussi diversifiées et dynamiques que possible : coaching, soutenance, tutoriel, vidéo, étude de cas pratiques, exercices, assistance, etc.

Chaque matière est organisée sous la forme d'une séquence d'enseignement. Chaque séquence est constituée de trois plages destinées respectivement aux "Cours", "Exercices" et "Evaluation". Les plages "Cours" et "Exercices" sont exclusivement en ligne (via une plateforme dédiée) et sont soumises à un compte à rebours. La plage "Evaluation" est aussi soumise à un compte à rebours. La formation comprend également une base de données documentaire, accessible aux participants visant à compléter le parcours.

### \* SEMESTRE 1

-Introduction à Python et Projet Python

-Analytics & Visualisation des données

-Datamining & Textmining

-Statistique 1

-Statistique 2 (Séries temporelles)

-Economie digitale et valorisation des données

-CRM-Analytics

-Gouvernance des données (GDPR, Solutions de gouvernance : Collibra, DataGalaxy)

- **Projet Minecraft (sous Python)**, en partenariat avec Microsoft

## \* SEMESTRE 2

-Machine Learning 1

-Machine Learning 2

-Deep Learning

-Distributed and Large Scale Computing (écosystèmes Hadoop, Cloudera CDP, Azure,...)

-Cinq Certifications Microsoft Azure : Bases de données sur Azure (DP900), Solutions Data Science sur Azure (DP100), Solutions Data Engineering sur Azure (DP203), Intelligence Artificielle sur Azure (AI100), Administrateur Azure (AZ104)

-Mémoire

## Diplôme d'université Sorbonne Data Analytics (Formation initiale)

### Semestre 1

---

#### UE 1 : Introduction à Python pour la data-science

Introduction à Python	5h
Projet Python	15h

#### UE 2 : Analyse de données

Analytics & Visualisation des données	20h
Datamining & Textmining	20h

#### UE 3 : Analyse statistique

Statistique 1 (SVM, regression régularisée, forêts,...)	20h
Statistique 2 (Séries temporelles)	10h

#### UE 4 : Eco-système Data Science

CRM-Analytics	10h
Economie digitale et valorisation des données	5h
Gouvernance des données (GDPR, Solutions de gouvernance : Co	5h

### Semestre 2

---

#### UE 1 : Machine Learning

Deep Learning	15h
Machine Learning 1	20h
Machine Learning 2	20h

#### UE 2 : Apprentissage et Systèmes

##### répartis

Certification Azure Data Scientist	10h
Distributed and Large Scale Computing	10h

#### UE 3 : Mémoire

Mémoire - Projet général et Stage	2h
Tutorat	6h