



EXECUTIVE EDUCATION

EXECUTIVE MASTER

STATISTIQUE & BIG DATA



Le programme dédié aux futurs Data Scientists souhaitant traiter les problématiques d'entreprise liées au big data

L'essentiel

Cette formation est adossée au CEREMADE, laboratoire de mathématiques appliquées qui héberge en son sein des enseignants-chercheurs mondialement reconnus dans leur domaine. Elle se veut le pendant naturel du parcours Mathématiques, Apprentissage et Sciences Humaines créé en 2014 au sein du Master 2 Mathématiques appliquées.

Au-delà de l'acquisition classique du bagage scientifique du Data Scientist, cette formation met l'accent sur la rigueur indispensable de traitement des données et sur la présentation des méthodologies statistiques récentes. Ces dernières permettent de produire les outils d'analyse mathématique, d'estimation statistique et de prédiction adaptées, qui se déclinent en algorithmes novateurs programmés dans des langages informatiques performants.

Leur maîtrise offrira aux Data Scientists de cette formation une visibilité et une compétence rare et précieuse sur le marché du travail.

Il est possible d'intégrer directement le module 2 ou 3 sous réserve d'avoir réussi un test en ligne.

Objectifs

Modéliser, analyser et traiter les données

- Développer une compétence de haut niveau en mathématiques appliquées qui couvre toutes les problématiques de nature statistique rencontrées en entreprise.
- Illustrer l'utilisation des techniques idoines pour le traitement des problématiques actuelles et futures en marketing, actuariat et finance.
- Maîtriser les langages de programmation pour le traitement des données massives.
- Maîtriser l'implémentation d'algorithmes complexes et le traitement parallèle et distribué des données massives.

Modalités pratiques

Durée

300 heures sur 17 mois

Rythme

2 journées d'intégration consécutives, 3h de cours chaque mardi en fin de journée de 18h à 21h + 1 jeudi complet par mois (de 08h30 à 18h30)

Lieu

PariSanté Campus (Paris 15e)

Programme

La formation comprend un volume horaire de 300 heures de cours à l'exclusion du temps consacré aux évaluations. L'enseignement repose sur 32 cours répartis en 3 modules de niveau progressif. Il est possible de rentrer directement en module 2 (un test de prérequis est organisé). Certains documents de cours sont en anglais.

Module 1 : Maîtriser les bases de la statistique

Ce module constitue une remise à niveau en statistique pour pouvoir se familiariser avec « la culture de la donnée ». Il aborde le concept de modélisation aléatoire et les techniques élémentaires de la statistique. Le cadre du

Public

La formation a été conçue spécifiquement pour des professionnels en activité disposant d'un diplôme de niveau Bac+3 en mathématiques ou de niveau admis comme équivalent. Ainsi, la qualité de l'expérience professionnelle peut être prise en considération pour accepter des candidats ne disposant pas d'un diplôme Bac+3.

Admission

La sélection se fait sur dossier, qui comprendra un CV, une lettre de motivation et une lettre de recommandation (facultatif).

Les candidats retenus à l'issue de l'examen des dossiers seront invités à passer un test de mathématiques en ligne. La validation du test de mathématiques est nécessaire pour être admis.

L'Executive Master Statistique & Big Data délivre un Diplôme d'Université (DU) niveau Bac+5 validé par le conseil d'administration de l'Université Paris Dauphine-PSL.



L'Executive Master Statistique & Big Data permet d'obtenir un bloc de compétences du Diplôme Mathématiques et applications, éligible au CPF.

modèle linéaire gaussien permet de formuler de manière élémentaire de nombreuses situations concrètes et illustre la résolution de problématiques. Une mise en pratique informatique à l'aide du logiciel R est déployée.

- Remise à niveau en probabilités
- Estimation classique, tests, régions de confiance
- Modèle linéaire gaussien
- Mise en pratique statistique sous le logiciel R

NB : Les titulaires d'un diplôme de niveau bac + 4 sanctionnant un bon niveau en modélisation aléatoire, en probabilité et en statistique pourront être dispensés de ce premier module.

Module 2 : Traiter les données par la statistique avancée

Il est possible d'intégrer directement le module 2, sous réserve d'avoir réussi le test d'entrée.

Ce module décline les outils mathématiques et statistiques qui permettent d'envisager le traitement transversal de nombreuses problématiques. L'objectif est de présenter un très large panorama des techniques modernes à la fois sur les plans méthodologiques et pratiques et d'être en capacité d'effectuer des choix de modèles probants et de les appliquer de manière pertinente selon les résultats et objectifs attendus.

- Choix de modèles
- Modèles linéaires généralisés
- Régression non-paramétrique
- Analyse de données

- Bases de données sous SQL
- Valeurs extrêmes
- Approche bayésienne
- Séries temporelles
- Données manquantes et données censurées
- R avancé
- Python

NB : Les auditeurs dispensés du Module 1 pourront se voir offrir une formation accélérée au logiciel R sur proposition des directeurs de la formation.

Module 3 : Développer l'apprentissage et le big data

Il est possible d'intégrer directement le module 3, sous réserve d'avoir réussi le test d'entrée.

Ce module est essentiellement dévolu à la thématique du big data qui recouvre toutes les problématiques pour lesquelles les volumes des données sont tels que celles-ci ne peuvent être gérées par les approches classiques. Les techniques idoines sont celles de l'apprentissage statistique qui se situe à l'interface de l'informatique et de la statistique.

- Optimisation
- Apprentissage supervisé
- Graphes

- Deep learning
- Pratique du Deep Learning
- Classification non-supervisée
- Ethique et protection des données
- Calcul parallèle
- Cloud computing
- Visualisation des données
- Mini-projets
- Mise en pratique marketing
- Mise en pratique actuariat
- Case studies in Finance

Modalités pédagogiques

Les cours en présentiel et en distanciel sont illustrés par des jeux de données artificielles et réelles avec une mise en œuvre directe sur machine.

- Les intervenants s'assureront que les auditeurs disposent d'ordinateur personnel et des logiciels nécessaires mis à jour
- Les auditeurs auront accès au cluster informatique du CEREMADE pour les cours qui le nécessitent
- Des études de cas feront l'objet de travaux individuels ou collectifs en inter-session
- La diversité des pratiques professionnelles et le

- développement de la recherche en statistique imposent que les auditeurs assurent un travail d'analyse et de synthèse sur la base d'une bibliographie préparée par les responsables de la formation, ainsi que des recherches documentaires
- Ponctuellement, des conférences viennent compléter la formation, afin que des professionnels présentent leur pratique de la science des données en entreprise
- Plusieurs projets rythment la formation, afin de s'exercer à l'implémentation et à l'interprétation des techniques rencontrées

Modalités d'évaluation

La validation des acquis se fait tout au long du programme :

- Module 1 : devoirs sur table
- Modules 2 et 3 : travaux pratiques et mise en application sur données réelles, projets individuels et de groupes



Contact

Laurence LELIÈVRE

✉ contact.stat-big-data@dauphine.psl.eu
06 26 71 77 67

Site de la formation

[executive-education.dauphine.psl.eu/formations/executive-master-diplome-universite/
statistique-big-data](http://executive-education.dauphine.psl.eu/formations/executive-master-diplome-universite/statistique-big-data)



UNIVERSITÉ PARIS DAUPHINE-PSL
Place Maréchal de Lattre de Tassigny
75116 Paris

executive-education.dauphine.psl.eu



[linkedin.com/company/formation-continue](https://www.linkedin.com/company/formation-continue)