

DU ANALYSE DES DONNEES

RÉSUMÉ DE LA FORMATION

Type de diplôme : Diplôme d'université

Domaine(s) ministériel(s) : Sciences, technologies, santé

PLUS D'INFOS

Durée : 1 an

Niveau d'étude : BAC +2

Public concerné

* Enseignement à distance

Formation à distance : Obligatoire

Nature de la formation : Diplôme d'université

Présentation

La diffusion de plus en plus large de logiciels d'Analyse des données performants permet aujourd'hui de mettre en œuvre par un simple clic les procédures les plus sophistiquées. L'utilisateur peut-il considérer ces procédures comme une boîte noire dont il ferait usage comme d'un téléphone portable ? Nous ne le pensons pas. Nous croyons au contraire que la connaissance des fondements de ces méthodes est nécessaire pour juger de la pertinence de leur application, interpréter correctement les sorties et être capable de suivre les évolutions qui ne manqueront pas de se produire dans une discipline en renouvellement constant. L'objectif poursuivi n'est pas d'accroître les compétences académiques de l'étudiant, mais de lui permettre de mettre en œuvre des méthodes en toute connaissance de cause.

Qu'est-ce qu'un Diplôme Universitaire (DU) ?

Après la validation du mémoire, l'étudiant obtient un Diplôme Universitaire (DU). C'est un diplôme délivré par une université française, contrairement à la licence, au master et au doctorat, qui sont des diplômes nationaux, c'est-à-dire des "diplômes d'état" délivrés au nom du ministère.

Le DU correspond à un domaine restreint, à vocation temporaire ou professionnelle. Chaque université est donc habilitée pour des DU spécifiques. Le mode d'accès, la durée des études, le mode d'évaluation peuvent être très différents selon le but professionnalisant du diplôme.

Objectifs

* Apporter un complément de formation en analyse des données aux étudiants et aux professionnels issus de différentes disciplines (Mathématiques, Informatique, Biologie, Physique, Économie, Gestion, Sciences Humaines, etc).

* Au-delà d'une application mécanique des logiciels existants, permettre une compréhension profonde des méthodes sous-jacentes, en réduisant cependant l'investissement mathématique au minimum nécessaire.

Contenu de la formation

Durée de la formation

Du 1^{er} novembre au 31 octobre de l'année universitaire.

Deux lectures du cours possibles

On peut soit parcourir les chapitres de façon séquentielle. Dans l'ensemble, l'ordre séquentiel correspond à une difficulté croissante.

Ou bien, on peut suivre un cheminement personnel, en sautant certains chapitres, à condition de respecter les relations de dépendance existantes entre certains chapitres.

Dans les deux cas, il est conseillé de lire dans un premier temps les théorèmes sans les démonstrations et de passer directement à l'application du chapitre. Dans un deuxième temps, on pourra revenir sur les démonstrations pour acquérir une compréhension plus approfondie, et contrôler cette compréhension en travaillant les exemples traités et en faisant les exercices proposés.

Par contre, il est important qu'à travers les applications, l'étudiant accroisse ses compétences d'utilisateur. On a choisi délibérément de petits fichiers de données faciles à superviser, qui pourraient être traités "à la main". Bien entendu, l'Analyse des données ne prend vraiment tout son intérêt que sur des fichiers plus volumineux.

Ici, il s'agit simplement de montrer la mise en œuvre des méthodes dans des situations simples. Dans chaque étude de cas, la démarche proposée n'a rien d'impératif, et nous ne prétendons pas qu'elle soit la plus efficiente. Elle peut se prêter à de multiples variantes.

En complément

Des travaux pratiques (TP) en lien avec le cours sont proposés. Nous incitons fortement les étudiants à réaliser les exercices de TP qui seront régulièrement fournis (avec les corrigés) pour acquérir la maîtrise du logiciel SAS (fourni par la formation). Ces TP ne sont pas évalués.

* Chapitre 0 : Introduction

Présentation du cours et comment le travailler.

* Chapitre 1 : Notions élémentaires en Analyse des données

Rappels de quelques outils d'algèbre linéaire, d'analyse et de statistiques : moyenne, variance, écart-type, matrice de covariance, inerties, etc. Exemples.

* Chapitre 2 : Description bidimensionnelles et mesure de liaison entre variables

1 Mesure de liaison : Lorsque des juges ont exprimé leurs préférences sous forme de classements, comment déterminer le degré d'accord entre eux ? S'il est jugé suffisant, comment dégager une opinion commune ? Exemples.

2 Etude des tableaux de contingence à double entrée. Exemples.

* **Chapitre 3 : Méthodes de classification**

1 Classification par centres mobiles. Exemples.

2 Classification hiérarchique ascendante : Méthodes du saut minimal, du diamètre maximal, de la distance moyenne et de la règle de Ward. Exemples.

* **Chapitre 4 : Analyse en composante principale (ACP)**

Présentation de la méthode dans ses divers aspects. Exemples.

* **Chapitre 5 : Analyse de variance**

1 Analyse de variance à un facteur

2 Analyse de variance à deux facteurs répétés ou non.

3 Exemples d'illustration.

* **Chapitre 6 : Analyse Discriminante (AD)**

Etude de la méthode et interprétation. Exemple.

* **Chapitre 7 : Analyse factorielle des correspondances (AFC)**

Etude de la méthode dans ses divers aspects. Interprétations. Exemple.

* **Chapitre 8 : Analyse factorielle des correspondances multiples (AFCm)**

Présentation de la méthode et lien avec l'AFC. Interprétations et exemples.

* **Chapitre 9 : Régression linéaire multiple**

Le modèle linéaire standard. Estimation des paramètres par les moindres carrés ordinaires. Théorème de Gauss-Markov. Préviation. Significativité.

Contrôle des connaissances

D'une part l'étudiant doit envoyer selon le calendrier prévu les exercices du cours par courriers postal ou électronique (4 vagues d'exercices).

D'autre part, il doit préparer **un mémoire de 30 à 50 pages en lien avec le cours** sur un sujet et des données choisis par l'étudiant. Ce choix, pour être valable, doit être validé par l'enseignant en charge du DU.

Conditions d'accès

* **Diplôme exigé** : Avoir validé le L2 des licences de Mathématiques, Informatique, Physique-Chimie, Biologie, Economie-Gestion, AES, Géographie, etc.

* **Sur demande** : possibilité de validation des études, des expériences professionnelles et des acquis personnels. Renseignements auprès du SFC, Service Formation Continue.

1) Envoyer avant le 20 octobre de l'année universitaire en cours (en format .pdf).

- * Le [formulaire de pré-inscription que vous pouvez télécharger](#)
- * Votre CV
- * Une lettre de motivation
- * L'ensemble étant envoyé à [M.Saïd HAMADENE](#)

2) Vous recevrez une réponse par e-mail. La capacité d'accueil est limitée à 20 étudiants.

3) L'inscription définitive s'effectuera selon des modalités qui vous seront indiquées dans la réponse. Elle peut s'effectuer à distance. Il vous sera notamment demandé une photocopie de vos diplômes, 2 photos d'identité et le versement des droits d'inscription.

- * [Fiche d'autorisation de transfert](#)
- * [Transfert de dossier](#)

Pour toute demande de dossier d'inscription, [contacter la scolarité](#).

- * Formation initiale

Les tarifs pour la formation initiale seront communiqués en juillet 2019 via ce lien :

- * [Renseignements sur le site de l'Université](#)
- * Formation continue / Reprise d'études

Le Service Formation Continue (SFC) est en charge de l'orientation des publics souhaitant reprendre leurs études et de leur accompagnement dans les démarches de recherche de financements associées.

- * [Les tarifs pour les publics en formation continue et en reprise d'études](#)

Public cible

Ce cours s'adresse à des étudiants qui ont déjà une certaine familiarité avec les statistiques, les probabilités et l'algèbre linéaire. Celles qu'on peut acquérir dans les premiers cycles d'Economie, de Gestion, de Sciences de la nature et de la vie, de Physique et bien sûr de Mathématiques.

Le SFC, Service de Formation Continue universitaire, vous accueille, vous oriente et vous accompagne dans votre projet de reprise d'études.

- * [Financer sa formation ?](#)
- * [Modalités d'inscription en formation continue](#)
- * [Renseignements sur la validation des acquis](#)

Insertion professionnelle

L'analyse des données est née de l'expansion considérable qu'a connue ces quarante dernières années un chapitre de la Statistique qu'on appelait « Statistique descriptive ». Elle fait maintenant partie des connaissances de base de l'ingénieur, du gestionnaire, de l'économiste. Parmi les

innombrables applications, citons dans le domaine industriel : la fiabilité des matériels, le contrôle de qualité, l'analyse des résultats de mesure et leur planification, la prévision ; dans le domaine de l'économie et des sciences humaines : les modèles économétriques, les sondages, les enquêtes d'opinion, les études de marchés.

Contacts

Contact administratif

Hamadene Saïd

Said.hamadene@univ-lemans.fr

UFR Sciences et Techniques - Scolarité

sco-sciences@univ-lemans.fr

Florence Rousseau

florence.rousseau@univ-lemans.fr

Bougard Brigitte