

### Note d'information sur les cours STAT2360-STAT2370<sup>1</sup> Séminaires en Data Management<sup>2</sup>

### QUELS SONT LES BUTS ET PARTICULARITÉS DES SÉMINAIRES EN DATA MANAGEMENT

Organisée en collaboration avec la société SAS, ces deux séminaires vous proposent une formation à la gestion informatique et au traitement statistique de grandes bases de données à l'aide du système SAS. Cette formation répond à un besoin important des entreprises qui recherchent des personnes capables de gérer de grandes bases de données mais aussi d'en analyser le contenu pour répondre aux questions stratégiques de la société à l'aide de méthodes statistiques adaptées.

Le système SAS a été choisi comme base de cette formation car il est, actuellement, le seul outil informatique capable de gérer de grandes bases de données, faire de la statistique de haut niveau et proposant un outil intégré de data mining. Il est largement répandu dans les grandes entreprises de Belgique, d'Europe et du Monde. Une formation au data management en SAS vous permettra donc d'être directement opérationnel dans les sociétés qui l'utilisent.

## EN QUOI CONSISTE LES SÉMINAIRES EN DATA MANAGEMENT : STAT2360 ET STAT2370 ?

L'objectif de ces deux séminaires est de vous amener à maîtriser les principaux outils de data management du logiciel SAS. Au terme de ceux-ci vous aurez une connaissance approfondie de la structure du système SAS, serez capable de programmer en SAS Base, d'écrire des macros et de mettre en œuvre vos connaissances sur des grandes bases de données.

Attention, ces cours sont plus exigeants que d'autres cours du Master en Statistique/Biostatistique mais vous délivrent, en plus d'un diplôme universitaire, une reconnaissance de votre formation SAS qui sera très appréciée dans votre CV et directement valorisable sur le marché du travail.

Ces deux séminaires traiteront des sujets suivants :

- Création et manipulation de Data Sets SAS : importation, exportation, définition de formats, fusion de tables, manipulation, création et transformation de variables.
- Préparation de tableaux résumés des données, production de rapports sous différents formats (texte, html...).
- Traitement de bases de données sous SAS, écriture de requêtes vers d'autres systèmes de base de données (Oracle...). Utilisation du langage SQL avec SAS.
- Développement de Macros SAS: description du macro langage, création et utilisation de variables macro, programmation, debugging.
- Programmation avancée et efficace avec SAS: stockage et manipulation efficace de grands Data Sets, optimisation de la mémoire et de l'espace de stockage, techniques de programmation efficace, ...
- Pratique des outils sur des études de cas avancées

<sup>&</sup>lt;sup>1</sup> STAT2360 : Data Management I : programmation de base en SAS [15 heures, 6 crédits] STAT2370 : Data Management II : programmation avancée en SAS [35 heures, 6 crédits]

<sup>&</sup>lt;sup>2</sup> Cette formation a pu être mise en place grâce à un partenariat de type "initiative pédagogique " entre l'UCL et la filiale belge de la société SAS Institute au travers de son programme SAS-Academic.

Deux séminaires sont organisés:

- STAT2360 (15h-6C) SAS base programming
- STAT2370 (35h-6C) SAS advanced programming (SQL, Macros, efficiency techniques)

Le séminaire STAT2360 organisé au premier quadrimestre est constitué de :

- Une demi-journée d'introduction à SAS et d'explication sur le cours
- Deux journées de 6 heures de cours Cours SAS base partie I et II présentée en français par un formateur de SAS Belgique en salle informatique (UCL).
- La préparation individuelle du certificat via un tutorial informatisé avec coaching et contrôle régulier des connaissances par C. Bugli et un assistant STAT.
- Une séance de « tricks and tips » de 3 heures donnée par un formateur de SAS en fin de préparation.
- Un prétest de simulation du passage du certificat à l'Institut de Statistique.
- Le passage du certificat « SAS base programming » tenant lieu d'examen (payant et financé par SAS). Cet examen aura lieu dans une salle informatique de l'UCL, dans les locaux de SAS (Tervuren) ou dans un centre d'examen indépendant de SAS (Bruxelles) en fonction du nombre d'étudiants inscrits.

Le séminaire STAT2370 organisé au second quadrimestre est constitué de :

- Quatre journées de 6 heures de cours donnés en français par un formateur de SAS Belgique en salle informatique (UCL) au début du second quadrimestre (SQL Processing using the SAS System (1j), SAS Programming III (2j), Macros (1j)).
- Deux demi-journées, l'une de questions/réponses et l'autre de présentation de l'étude de cas.
- Une étude de cas à réaliser avec encadrement (questions-réponses) par C. Bugli.
- Une demi-journée de préparation à l'examen ou au certificat « SAS avanced programming » et de présentation de business cases avec SAS.
- Un examen consistant en la réalisation en salle informatique d'un programme SAS utilisant l'ensemble des techniques de data management étudiées. Cet examen a lieu chez SAS (Tervuren).

# QUELS SONT LES CONDITIONS D'ACCÈS AUX SÉMINAIRES EN DATA MANAGEMENT ?

Le séminaire STAT2360 est ouvert, par ordre de priorité :

- à tous les étudiants en Master et doctorants STAT / ACTU de la LSBA,
- à tous les étudiants en Master et doctorants de l'Académie Louvain (sur dossier),
- à tous les étudiants en Master et doctorants d'autres universités de Belgique(sur dossier) qui ont inscrit ce cours à leur programme académique.

Les étudiants doivent de plus faire preuve d'une bonne connaissance passive de l'anglais. L'accès gratuit au passage du certificat n'est donné qu'aux étudiants dont ce cours est inscrit à leur programme académique et ayant démontré leur capacité à le réussir (participation active aux séances de préparation et réussite du prétest avec 65% minimum). Le passage du certificat est financé une seule fois par SAS. Un étudiant qui échoue au certificat pourra soit repasser le certificat à ses frais ou passer un examen proposé par le titulaire du cours (ne donnant pas droit au certificat mais seulement à une cote UCL) ou obtenir une cote établie sur base de ses résultats aux autotests et au prétest de simulation.

La participation au séminaire STAT2370 est de plus soumise à la condition d'avoir réussi leur certificat « SAS base programming » (cours STAT2360).

En fonction des disponibilités en termes de places, les séminaires STAT2360 et STAT2370 pourront être ouverts par ordre de priorité à tous les chercheurs et membres du personnel UCL, de l'Académie Louvain et d'autres universités belges. Néanmoins, le passage du certificat ne sera pas financé par SAS. Les personnes ayant suivi le cours mais dont le passage du certificat n'est pas financé par SAS pourront passer le certificat à leurs frais (coût d'inscription réduit au prix coutant) lors des sessions organisées à l'UCL. De plus, la possibilité de passer la certification « SAS Advanced Programming » sera également offerte à un cout d'inscription réduit (prix coutant) lors des sessions organisées à l'UCL.

Note: Les étudiants inscrits en certificat STAT et ACTU pourront ajouter les séminaires STAT2360 et STAT2370 à leur programme de certificat, moyennant un cout d'inscription supplémentaire (info sur demande).

### COMMENT SERONT ÉVALUÉS LES COURS STAT2360 ET STAT2370 ?

La cote attribuée au cours STAT2360 sera basée sur :

• Votre résultat au certificat «SAS base programming»

La cote attribuée au cours STAT2370 sera basée sur :

• L'appréciation délivrée par SAS suite à la réalisation de votre case-study

#### **COORDINATEURS DES COURS**

**Titulaire**: Catherine Legrand, <u>catherine.legrand@uclouvain.be</u> (010/478798) **Encadrement de la préparation du certificat et du case study:** C. Bugli, <u>celine.bugli@uclouvain.be</u> (010 47 86 49) et C. Timmermans, <u>catherine.timmermans@uclouvain.be</u>