

Introduction

Présentation de l'ouvrage

Cet ouvrage s'adresse à tous les utilisateurs débutants ou de niveau intermédiaire du langage SAS, qu'ils soient en formation dans les départements d'économie, de gestion, de mathématiques appliquées, de statistique, de biologie, de STID, de MIAGE, d'IAE des universités, dans les écoles de commerce, d'ingénieurs, ou qu'ils évoluent déjà dans le monde professionnel.

Le système SAS est très vaste : on pourrait presque dire qu'il existe autant d'utilisations de SAS que d'utilisateurs. Il existe cependant un socle commun de connaissances mobilisées quotidiennement par tous ces utilisateurs. C'est ce socle que se propose d'explorer cet ouvrage.

L'ouvrage vise à présenter les connaissances fondamentales et transversales nécessaires à la programmation SAS, plus particulièrement à la création, la modification, la manipulation et la gestion des tables de données. Il s'intéresse également aux procédures d'exploration et de première exploitation des données. Enfin, il aborde la production de graphiques, de documents de présentation en format HTML, RTF ou PDF, ainsi que le langage macro propre au système SAS.

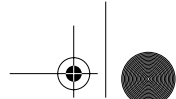
Pour les utilisateurs qui ont pour finalité la manipulation et la transformation de données, puis la production de rapports présentant ces dernières, cet ouvrage détaille les outils qu'ils doivent mobiliser.

Pour les personnes qui utilisent en plus les outils économétriques et statistiques, nous couvrons dans cet ouvrage les phases en amont et en aval de l'analyse qu'elles doivent mener au moyen du système SAS.

Aucun travail d'analyse ou de construction de rapport n'est possible sur des données non préparées. Les outils que nous aborderons au cours des six premiers chapitres vous aideront à mettre en forme ces données et à les explorer.

Il s'agit d'une phase essentielle de votre travail puisque les données, dans leur forme brute, n'ont jamais la forme des données sur lesquelles vous pouvez directement appliquer vos outils économétriques et statistiques ou sur la base desquelles vous pouvez immédiatement rédiger un rapport.

Les données brutes que l'on vous confiera devront dans un premier temps être transformées en une table SAS parfaitement fidèle à vos données originales. C'est précisément l'objet des chapitres 1 et 2. Nous examinons dans le chapitre 3 les outils à votre disposition pour créer, à partir des variables présentes dans une table brute et au moyen de transformations et de combinaisons, les variables nécessaires à votre travail.



2 ♦ SAS

Dans de très nombreux cas, il est aussi nécessaire de mobiliser plusieurs sources de données et donc de combiner des tables entre elles. C'est ce que nous voyons dans le chapitre 4. Ainsi, ce n'est pas une seule mais plusieurs tables que vous manipulerez, dans des versions plus ou moins évoluées. Il est donc nécessaire de les organiser et de les gérer de façon à travailler le plus efficacement possible. Ce point essentiel est traité dans le chapitre 5.

Nous examinons dans le chapitre 6 un ensemble de procédures à même de vous permettre d'explorer, de connaître et de comprendre vos données. On ne peut en effet analyser ou présenter des données que si on les connaît bien. Ces outils d'exploration des tables vont servir aussi dans la phase de retransmission de vos résultats puisque, même si votre finalité première est l'analyse économétrique ou statistique, vous aurez très souvent besoin de présenter au moyen de tableaux les données qui servent de base à votre analyse.

Nous présentons dans le chapitre 7 les procédures graphiques qui peuvent, ici encore, être utiles lors des phases d'exploration et de présentation de vos résultats.

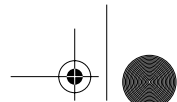
La phase de présentation des résultats est réalisée en aval de l'analyse. C'est une phase importante puisque votre travail d'analyse ne rencontrera jamais l'attention qu'il mérite s'il est livré dans une forme brute. Le document doit être attrayant par sa forme et indiquer de façon claire les principaux résultats que vous avez obtenus.

Les entreprises construisent des bases de données de plus en plus importantes et c'est le rôle du statisticien de tirer de ces données des informations essentielles au développement de son entreprise. Ce rôle, de par la position qu'occupe le statisticien, est particulier. Les analyses que peut fournir le statisticien sont destinées non pas à d'autres statisticiens mais à des personnes qui devront, sur la base de ces analyses, prendre des décisions.

Une part essentielle du travail du statisticien en entreprise est donc d'expliquer au moyen d'un langage non technique l'essentiel de ses résultats. Bien sûr, il ne s'agit pas de privilégier le fond au détriment de la forme : les analyses que vous fournirez devront être techniquement indiscutables mais vous n'obtiendrez pas l'attention que mérite votre analyse si leur forme n'est pas attrayante.

La retransmission de vos résultats et la rédaction de rapports mobilisent à nouveau les procédures présentées dans le chapitre 6 ainsi que les procédures graphiques explorées dans le chapitre 7. Nous examinons dans le chapitre 8 l'*Output Delivery System* (ODS) livré avec SAS. Au moyen de cet outil offert par SAS, vous avez en effet la possibilité de produire en quelques instructions des documents attrayants au format HTML, RTF (*Rich Text Format*), que vous pouvez coller dans votre traitement de texte, ou PDF. Ce dernier format peut être utilisé si vous souhaitez effectuer une présentation au moyen d'un vidéoprojecteur.

Le chapitre 9 consacré au langage macro conclut cet ouvrage. En effet, pour percevoir l'intérêt du langage macro, il faut avoir une certaine expérience de SAS et être conscient que votre travail comprend des phases répétitives ou qu'il suffit de modifier quelques instructions pour qu'un programme dédié au traitement d'une table spécifique puisse aussi analyser une autre table. Après une large introduction au langage macro de SAS, nous présentons un exemple mobilisant les connaissances que vous aurez acquises au cours des huit premiers chapitres.



Progresser dans votre connaissance de SAS

Cet ouvrage ne saurait prétendre à l'exhaustivité. Cependant, il aborde les problèmes les plus courants et vous donne les outils qui vous permettront de les résoudre.

Ainsi, le chapitre 7, consacré aux graphiques, explore principalement deux types de graphiques : les histogrammes et les courbes. À cette occasion, sont présentés les éléments de langage communs à toutes les procédures graphiques. En effet, une fois que vous maîtriserez ces deux outils ainsi que le langage propre à l'environnement graphique, vous pourrez aisément exploiter les possibilités des autres procédures graphiques offertes par SAS.

Chaque chapitre commence par expliquer les concepts de base avant d'aborder des notions plus complexes. Ainsi, dans l'exposé d'une procédure, il y aura peu de rapport entre les résultats simples que nous proposerons dans un premier temps et les derniers résultats. Ces derniers résultats ne peuvent cependant pas être compris sans les premières parties dans lesquelles les principaux concepts sont exposés avant d'être mis en pratique.

C'est pour cette même raison que le chapitre consacré au langage macro est présenté en dernier : il mobilise en effet l'ensemble des connaissances abordées dans les huit premiers chapitres.

Nous avons fait le choix de passer sous silence les interfaces proposées par SAS qui permettent en quelques clics de réaliser un graphique, d'importer des données, de créer des tableaux, etc. Ces interfaces, certes conviviales, écrivent le programme à votre place mais ne vous dispensent absolument pas de connaître les outils de la programmation SAS. En effet, ces interfaces ne peuvent mobiliser 100 % des outils de programmation offerts par les procédures qu'elles émulent. Si votre demande s'éloigne des cas standard et si vous n'avez aucune connaissance des outils de la programmation SAS, vous ne pourrez pas comprendre le programme créé par votre interface et le modifier de manière qu'il produise le résultat souhaité.

Les utilisateurs avancés de SAS qui prétendent ne pas coder ajoutent généralement que c'est parce qu'ils ont beaucoup codé...

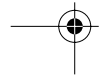
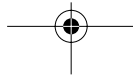
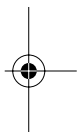
Comment travailler avec cet ouvrage

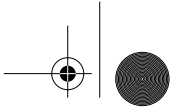
Le site compagnon de cet ouvrage est le complément indispensable à votre travail d'apprentissage de SAS. Vous retrouverez ce site à l'adresse suivante :

www.sas-sr.com

Ce site vous offre :

- tous les programmes proposés dans l'ouvrage ;
- toutes les sorties générées par ces programmes, y compris celles non reproduites dans cet ouvrage ;
- tous les fichiers nécessaires à la reproduction des exemples ;
- tous les fichiers nécessaires aux exercices proposés ;
- les solutions aux exercices proposés ;
- des compléments à cet ouvrage ;





4 ♦ SAS

- l'ensemble des liens vers des sites Internet proposés dans cet ouvrage ;
- des outils de recherche, simples d'utilisation, qui permettent de retrouver des exemples de programmes qui emploient des commandes et des options spécifiques.

Tout au long des chapitres, des exercices vous sont proposés. N'attendez pas d'avoir lu l'intégralité d'un chapitre pour les effectuer : ils sont indispensables à la compréhension des notions abordées et constituent généralement des passerelles entre les différents concepts.

Ne consultez les solutions qu'au terme d'une réelle réflexion ! C'est en commettant des erreurs que l'on progresse.

À partir du chapitre 5, les exercices sont moins nombreux. Vous êtes invité à reproduire l'intégralité des nombreux exemples de programmation proposés afin de bien comprendre l'impact de chacune des commandes et options utilisées. Vous avez la possibilité d'en retirer certaines ou de les remplacer par d'autres. Les exemples qui concluent les différentes sections des chapitres 6, 7, 8, et 9 mobilisent l'ensemble des concepts évoqués auparavant. Si vous ne deviez reproduire que ces exemples, le seul enseignement que vous risquez d'en tirer, c'est le constat d'une identité entre ce que vous aurez sur votre écran et ce que vous décrit l'ouvrage.

Si vous ne comprenez pas les concepts, vous ne pourrez rien apprendre d'un programme offrant plus de cinq instructions. Vous maîtriserez les concepts lorsque, avant même l'exécution du programme, vous aurez une idée précise de la structure du résultat produit.

La certification SAS

Cet ouvrage couvre l'intégralité du programme SAS 9 Certification Base Programming. Il va même au-delà en approfondissant certains domaines et en explorant des domaines non couverts par la certification comme SAS/GRAPH et le langage macro. Les sections qui font partie du programme de la certification SAS sont indiquées au moyen d'un astérisque dans le sommaire. Pour plus de détails sur le programme de cette certification, vous pouvez consulter le site suivant :

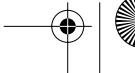
support.sas.com/certify/creds/testbp9.html

Cette certification est reconnue internationalement et constitue un réel plus dans un curriculum vitae. Nous vous encourageons vivement à la passer dans l'un des centres de test PROMETRIC. Si votre formation est partenaire de SAS ACADEMIC, vous pourrez bénéficier de coûts réduits. L'épreuve, d'une durée de deux heures, prend la forme d'un QCM de 70 questions en anglais sur ordinateur. Pour chaque question, 4 réponses sont proposées. Si vous avez au moins 46 bonnes réponses, vous obtenez la certification. Bien entendu, pendant le temps de l'épreuve, vous ne disposez pas de SAS.

Le centre de formation SAS, de son côté, propose une offre associant une journée de préparation à la certification, un manuel d'entraînement et l'examen.

www.sas.com/france/services/training/certification.html

Attention, j'attire votre attention sur le fait que, pour réussir cette certification, vous devez connaître parfaitement le système SAS. Vous connaîtrez le système SAS lorsque vous pourrez dire, en examinant différents programmes mais sans les exécuter, ce que vous allez obtenir ou pourquoi ils ne peuvent pas fonctionner. Cette connaissance ne



s'acquiert qu'au moyen d'une pratique régulière et approfondie des outils évoqués dans le programme de la certification.

La compréhension du système SAS ne suffit pas. Cette phase de compréhension, premier stade de votre apprentissage, vous permet uniquement de comprendre pourquoi un programme ne fonctionne pas lorsque vous l'exécutez.

Le site compagnon de cet ouvrage donne des exemples de questions qui pourraient vous être posées dans le cadre de cette certification¹.

Aller plus loin

Nous indiquerons très souvent des liens Internet vers des documents devant vous permettre d'approfondir vos connaissances dans un domaine particulier. Cependant, malgré le soin apporté et le temps passé à la rédaction de cet ouvrage, il est possible que vous ne trouviez pas de réponse à la question que vous vous posez. Dans ce cas, vous devrez de vous-même trouver d'autres sources.

Votre première source de renseignements est bien entendu l'aide SAS installée sur votre ordinateur, que vous activez avec la touche F1 de votre clavier. Vous pouvez aussi ouvrir le menu Aide, puis sélectionnez Introduction au système SAS. Vous naviguez dans cette aide de trois manières :

- **Par l'onglet Sommaire.** Vous pouvez consulter alors l'aide SAS comme vous consultez un livre.
- **Par l'onglet Index.** Vous devez alors indiquer l'instruction au sujet de laquelle vous recherchez des renseignements.
- **Par l'onglet Rechercher.** Indiquez plusieurs mots clés et explorez les pages de l'aide qui contiennent ces mots.

Nous ferons souvent référence à des articles de l'aide SAS, soit en partant de l'onglet Sommaire, soit en indiquant l'entrée à consulter dans l'Index. En vous invitant à consulter tel ou tel article, nous souhaitons vous aider là aussi dans votre apprentissage de l'aide SAS dans laquelle il pourrait vous sembler difficile de naviguer. Dans cet ouvrage, lorsque nous ferons référence à une entrée de l'aide, merci de respecter minuscules et majuscules pour retrouver l'article cité.

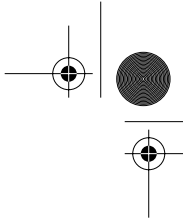
Si l'aide SAS n'est pas installée sur votre ordinateur, vous pouvez aussi la consulter sur le site support de SAS : support.sas.com/documentation/

À partir de cette page, vous pouvez accéder à l'aide de :

- SAS 8.2 : v8doc.sas.com/sashtml/
- SAS 9.1.3 : support.sas.com/onlinedoc/913/docMainpage.jsp
- SAS 9.2 : support.sas.com/cdlsearch?ct=80000

L'organisation de l'aide en ligne de SAS 9.2 diffère de l'aide installée sur votre ordinateur. L'aide SAS 8.2 de par son organisation reste intéressante à utiliser lorsque vous consultez la page d'aide consacrée à une instruction de procédure, vous avez la possibilité de revenir au chapitre auquel appartient la procédure, contrairement aux aides de SAS 9.x.

1. L'exercice 3.1 proposé dans la section 3.1.3. est tout à fait typique des questions posées à cet examen.



6 ♦ SAS

Vous avez aussi accès à un ensemble de documentations techniques qui ne sont pas comprises dans l'aide. SAS France propose des documentations en français à partir de son propre site support :

www.sas.com/france/services/support/

SAS : une communauté active d'utilisateurs

Si l'aide SAS doit toujours être votre référence lorsque vous recherchez une information sur une procédure ou une commande, les exemples qu'elle vous propose sont généralement simples et ne vous permettent pas toujours d'apprécier la richesse et les possibilités offertes par une procédure particulière. Il est toujours intéressant de s'éloigner de l'aide SAS pour regarder ce que d'autres utilisateurs peuvent faire avec SAS.

Les utilisateurs de SAS forment une communauté très active et il existe, essentiellement basés aux États-Unis, de très nombreux groupes d'utilisateurs au premier rang desquels se trouve SAS Global Forum (anciennement SUGI – *SAS User Group International*) : www.support.sas.com/events/sasglobalforum/

Ce groupe d'utilisateurs organise chaque année une conférence internationale au cours de laquelle des articles d'utilisateurs illustrant les possibilités de telle ou telle procédure sont exposés avant d'être mis à la disposition des utilisateurs SAS du monde entier.

Vous pouvez aisément consulter ces articles sur le site suivant : www.lexjansen.com/
Ce dernier site vous permet aussi de consulter les articles rédigés dans le cadre des groupes régionaux basés aux États-Unis¹ :

Enfin, à partir de ce site, vous avez aussi la possibilité de consulter les articles présentés au cours des conférences de deux groupes d'utilisateurs de l'industrie pharmaceutique :

- PharmaSUG (*The Pharmaceutical Industry SAS Users Group*) : www.pharmasug.org/
- PhUSE (*Pharmaceutical Users Software Exchange*) basé en Europe : www.phuse.eu/

Les forums de discussion constituent aussi une source de renseignements appréciable. Un forum très actif et existant depuis 1996 est hébergé par l'université de Géorgie :

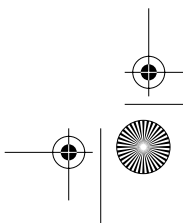
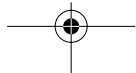
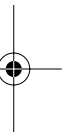
- Forum SAS-L : www.listserv.uga.edu/archives/sas-l.html

SAS Institute héberge aussi son propre forum : support.sas.com/forums/index.jspa

Attention, cependant, il s'agit là d'un forum d'utilisateurs qui ne saurait se substituer au support technique offert par SAS et que chaque client SAS (sauf les clients ACADEMIC) peut saisir en cas de problème. Le support SAS peut être contacté à cette adresse : support.sas.com/techsup/

À l'heure actuelle, il existe peu de forums à accès gratuit en langue française. Vous pouvez cependant entrer en contact moyennant un abonnement annuel avec d'autres utilisateurs français, participer au Club utilisateurs annuel de deux jours, à la journée SAS Forum Tech, et accéder à l'extranet utilisateurs ainsi qu'à son forum sur le site SAS FORUM France : www.sas.com/france/sasforum/

1. SESUG, MWSUG, NESUG, SCSUG, WUSS, et PNWSUG.



Vous trouverez aussi sur le forum developpez.net une section consacrée à SAS : www.developpez.net/forums/forumdisplay.php?f=886

Vous pouvez aussi consulter le très intéressant blog de Véronique Bourcier consacré à SAS : www.sasreference.fr

Divers groupes d'utilisateurs sont aussi présents sur Internet. Vous pouvez par exemple consulter les archives et contacter les utilisateurs du groupe Google <http://groups.google.fr/group/comp.soft-sys.sas/topics>

SAS Institute soutient l'activité des groupes d'utilisateurs et prend en considération les expériences des utilisateurs du système SAS. L'existence du SASware BALLOT est une expression de cet intérêt. Chaque année, SAS Institute demande aux utilisateurs quels ajouts ils souhaiteraient. Pour plus de détails sur le SASware BALLOT, vous pouvez consulter le lien support.sas.com/community/ballot/

Les différentes versions de SAS

Vous devez vous demander si cet ouvrage est adapté à la version de SAS que vous avez à disposition. En effet, différentes versions de SAS existent ou ont existé. SAS Institute distingue les versions principales (SAS 6.12, SAS 8.2 et SAS 9.2) des versions intermédiaires.

Le passage d'une version à une autre est toujours synonyme d'ajouts importants. Les ajouts qui concernent l'objet de cet ouvrage sont peu nombreux et seront signalés. ODS est cependant un cas particulier. ODS a été introduit avec SAS 8.0 en 1999. De très nombreuses modifications et améliorations ont été apportées depuis et certains programmes, notamment dans la partie consacrée à ODS PDF, peuvent ne pas fonctionner avec des versions antérieures à SAS 9.2.

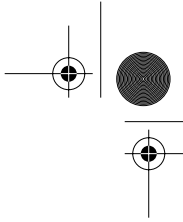
Vous trouverez sur le site compagnon de cet ouvrage des versions des programmes proposés dans le chapitre 8 compatibles avec SAS 9.1.3. Nous tentons aussi, dans la mesure du possible, de proposer des versions compatibles avec SAS 8.2.

Le système SAS est conçu de telle manière qu'un programme qui fonctionne dans une version donnée fonctionnera aussi dans les versions ultérieures de SAS. Certaines anciennes fonctions, remplacées depuis par des fonctions plus puissantes, même si elles ne sont plus documentées, continuent et continueront à pouvoir être utilisées. Le passage d'une version à une autre n'est donc pas synonyme d'un complet réapprentissage de SAS : ce que vous apprenez aujourd'hui sera encore valable demain.

SAS est loué par votre employeur, votre école ou votre université : il n'y a pas de coût supplémentaire demandé au client SAS lorsqu'une nouvelle version apparaît. Seules les contraintes propres aux clients¹ peuvent expliquer le maintien d'une version relativement ancienne de SAS. Vous comprenez donc que différentes versions de SAS peuvent subsister à un moment donné du temps : on peut ainsi considérer qu'actuellement, trois versions de SAS coexistent : SAS 8.2, SAS 9.1.3 et SAS 9.2.

SAS 9.2 a commencé à être déployé et distribué au printemps 2008. Le déploiement complet de SAS 9.2 doit se terminer courant 2009 avec la livraison des outils d'analyse propres aux métadonnées. Lorsque SAS 9.2 sera totalement déployé, SAS 9.1.3. disparaîtra.

1. Notamment en matière de tests suite à la migration des données entre différentes versions.



8 ♦ SAS

SAS Windows, Unix, OpenVMS, z/OS

Cet ouvrage traitera exclusivement de SAS Windows. En effet, si vous débutez sous SAS, il y a de très fortes chances que la version de SAS à votre disposition soit une version Windows. Si, par la suite, vous êtes appelé à travailler dans un environnement non Windows, vous découvrirez que les différences entre votre SAS Windows et SAS dans un autre environnement résident dans :

- la façon de débiter une session SAS ;
- l'adressage des fichiers externes à SAS ;
- l'existence de quelques options spécifiques à votre environnement.

Les programmes écrits au moyen de SAS Windows fonctionneront parfaitement sur d'autres systèmes d'exploitation. Il n'y a quasiment aucune différence entre SAS Windows et les versions UNIX, OpenVMS et z/OS de SAS.

Disposer du système SAS

Si vous êtes actuellement dans une université ou une école, il se peut que la formation que vous suivez soit inscrite au programme SAS ACADEMIC¹. L'inscription à ce programme permet à votre formation de bénéficier d'un certain nombre de licences SAS à un coût tout à fait intéressant et autorise l'installation sur votre ordinateur personnel de SAS pendant la durée de votre formation initiale (y compris pendant la durée de votre stage) contre une inscription au Club ACADEMIC de SAS et la signature d'un contrat de licence. Pour tous renseignements sur SAS ACADEMIC, vous pouvez consulter le site suivant : www.sas.com/france/academic/

Si vous ne pouvez bénéficier de l'offre SAS ACADEMIC, vous avez la possibilité de découvrir et de travailler sur une version limitée de SAS (SAS Learning Edition 4.1), suffisante pour l'apprentissage de SAS proposé par cet ouvrage – le site support.sas.com/learn/le/ fournit toutes les informations dont vous pouvez avoir besoin.

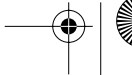
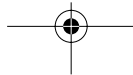
Cette version allégée de SAS contient les modules suivants : Base SAS (V9.1.3 SP4), SAS/STAT, SAS/GRAPH, SAS/QC, SAS/ETS et SAS entreprise guide 4.1². Cette version restreint l'affichage et les manipulations aux 1 500 premières lignes de votre table³. Le nombre de variables n'est pas limité.

Ce programme a une durée de vie limitée (au moment où est rédigée cette introduction, la version 4.1 de SAS Learning Edition est programmée pour ne plus fonctionner après le 31 décembre 2011) et n'est pas gratuit. Le coût est tout à fait raisonnable (60 \$ sans la documentation en anglais si vous êtes étudiant, 160 € avec la documentation dans les autres cas). Cette version de SAS n'est pas compatible avec Microsoft Windows® Vista™.

1. Environ 100 établissements et 200 filières sont actuellement partenaire SAS au travers du programme SAS ACADEMIC.

2. Certaines procédures propres aux modules cités ne sont pas contenues dans cette version.

3. Toutes les tables utilisées dans cet ouvrage contiennent moins de 1 500 observations.



Conventions typographiques

Les programmes (instructions, noms de variable, noms de tables) que vous pourrez écrire dans SAS peuvent employer minuscules et majuscules. Seules les valeurs prises par les modalités des variables alphanumériques doivent respecter la casse. Nous avons cependant employé dans cet ouvrage les majuscules et les minuscules de façon à vous aider au mieux dans votre compréhension des nombreux programmes proposés.

Tous les mots clés propres au langage SAS apparaissent en majuscules dans les programmes et dans le corps du texte. Les mots qui apparaissent en gras dans les programmes sont les instructions qui débutent et qui terminent une étape (ici, DATA et RUN). À l'intérieur du programme, les instructions propres au langage SAS apparaissent en majuscules (SET, IF, THEN, DO, RANUNI, etc.). Les noms des tables, des variables, des FORMAT et INFORMAT qui apparaissent dans le programme seront systématiquement écrits en minuscules. Nous indiquons par l'emploi des minuscules que le programme fonctionnera tout aussi bien si vous utilisez d'autres noms pour les tables ou les variables.

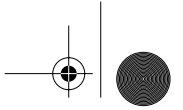
```
DATA test2;  
  SET test;  
  FORMAT date ddmmyy.;  
  ATTRIB x FORMAT=8.4  
         LABEL="Variable analysée";  
  IF x<15 THEN DO;  
    x7=RANUNI(234);  
    x9=x2+x12;  
  END;  
RUN;
```

Dans le corps du texte, seront systématiquement repris en majuscules les noms des commandes, fonctions, options, mais aussi le nom des variables, des FORMAT, INFORMAT, etc.

Les programmes seront indentés afin de vous permettre de mieux suivre et de mieux comprendre leur finalité. Il y a toujours au moins un niveau d'indentation, parfois plusieurs. Dans le cas du programme présenté ici, pour la commande ATTRIB, nous indentons une seconde fois pour signifier au lecteur que l'instruction LABEL est un élément de la commande ATTRIB. Nous indentons une seconde fois après l'instruction IF pour indiquer au lecteur que les deux instructions permettant de définir les variables X7 et X9 ne s'appliquent pas à toutes les observations (cette construction est en fait conditionnée au respect de la condition énoncée par l'instruction IF). L'instruction CARDS et les données qui suivent cette instruction ne seront pas indentées.

Il n'est pas forcément nécessaire d'organiser les programmes de cette manière mais au fur et à mesure que les programmes que vous rédigerez s'allongeront, vous constaterez qu'une telle organisation facilite grandement la compréhension.

À l'exception de certains programmes très courts, les programmes proposés dans l'ouvrage sont numérotés, ce qui vous permet de les retrouver facilement sur le site compagnon de cet ouvrage.



Un peu d'histoire

Le logiciel SAS est le fruit du travail d'un groupe de chercheurs de la North Carolina State University débuté en 1966. En 1967, James H. Goodnight, actuel CEO de SAS Institute, rejoint le projet et travaille à la stabilité du logiciel et à la création des procédures. Ce logiciel a connu à partir de 1971 un grand succès aussi bien auprès de la communauté universitaire qu'auprès de l'industrie pharmaceutique et de l'agriculture, en raison notamment de sa capacité à gérer les données manquantes.

En 1976, l'entreprise SAS Institute est créée notamment par James H. Goodnight et John P. Sall. SAS s'est depuis imposé comme le leader mondial de l'informatique décisionnelle. Il est aujourd'hui présent dans 96 des 100 premières entreprises du classement Fortune Global 500® et dans près de 44 000 sites clients répartis dans 107 pays. SAS Institute a réalisé un chiffre d'affaires de 2,15 milliards de dollars en 2007 dont 21 % a été réinvesti dans la recherche.

SAS signifiait à l'origine *Statistical Analysis System*. La signification de l'acronyme a depuis été perdue, marquant ainsi le fait que SAS est beaucoup plus de nos jours qu'un simple logiciel de traitement statistique. Le logiciel SAS en tant que tel n'existe plus ; il faudrait plutôt parler d'un système SAS, ou d'un progiciel SAS, composé de logiciels SAS intégrés.

Les modules du système SAS explorés

Le système SAS est composé de différents modules ayant tous une fonctionnalité propre. Le module Base SAS constitue le cœur de votre système SAS – il est nécessaire au fonctionnement des autres produits SAS.

Il permet de plus l'exécution des étapes DATA, du langage macro, des outils offerts par ODS (*Output Delivery System*), ainsi que les analyses statistiques simples, les procédures utilitaires comme PROC CONTENTS, PROC PRINT ou PROC SORT. Nous consacrerons l'essentiel de cet ouvrage à l'exploration des fonctionnalités de Base SAS. Le langage SAS est un langage dit de quatrième génération (L4G). Il s'agit, comme nous allons le découvrir dans cet ouvrage, d'un langage peu technique, très proche de la syntaxe naturelle.

SAS/GRAPH est le second module que nous explorerons dans cet ouvrage. Cette application permet de créer des graphiques et, depuis SAS 9.2 et l'introduction d'ODS GRAPHICS, d'intégrer automatiquement des graphiques prédéterminés dans les sorties des procédures d'analyses.

Une présentation succincte d'autres modules de SAS¹

SAS/ACCESS permet d'accéder aux données produites par d'autres logiciels, de les utiliser dans SAS et même de créer à partir de SAS des fichiers dans des formats externes. SAS/ACCESS propose différents modules qui permettent une intégration à différents systèmes de gestion de base de données (SGBD) du marché et à certains progiciels de gestion intégrés (PGI ou ERP – *Enterprise Resource Planning*). Voici une liste d'exemples non exhaustive : SAS/ACCESS Interface to SAP BW, to Teradata, to DB2, to MySQL, to ODBC, to OLE DB, to ORACLE, to PC Files (Microsoft Excel, Microsoft Access, Lotus 1-2-3, DBF), to PeopleSoft, etc.

1. Merci à Grégoire de Lassence pour l'aide apportée à la rédaction de cet encadré.



SAS AppDev Studio est l'outil de développement de SAS permettant notamment de :

- créer des interfaces clientes ;
- créer des servlets, des applets, pour construire des applications clientes sur les technologies Web ;
- créer des portlets pour compléter le portail ;
- créer des plug-in à ajouter aux clients Java ;
- développer des applications analytiques et de pilotage en client léger, Java ou Windows ;
- exploiter au mieux la puissance du serveur SAS dans des applications de pilotage sur mesure ;

SAS AppDev Studio constitue un environnement de développement complet et autonome pour concevoir tous types d'applications.

SAS/CONNECT permet de connecter des ordinateurs sur lesquels est installé SAS. Vous partagerez au moyen de cette application vos tables de données ; des programmes écrits sur un ordinateur pourront être exécutés sur un second.

SAS Data Integration Studio est une interface cliente Java de la plate-forme décisionnelle SAS. Elle fait partie des packages « SAS Data Integration Server » et « SAS Enterprise Data Integration Server ». Plus de détails peuvent être obtenus à l'adresse www.sas.com/france/software/technologies/dw.html.

SAS Enterprise Guide permet de manipuler des données, de faire des analyses statistiques et des rapports. Cette application développée en .NET est donc une application uniquement Windows. Enterprise Guide génère du code SAS, qui est exécuté par un moteur SAS base. SAS Enterprise Guide ne peut être utilisé sans une connexion à un socle Base SAS.

SAS Enterprise Miner est le module de Data Mining de SAS. C'est une « usine » à modèle permettant l'industrialisation de l'exploitation intensive de gigantesques bases de données.

SAS/ETS (*Econometrics, Times Series*) est le module qui comprend les outils pour l'économétrie des séries temporelles et la prévision (ARIMA, VARMAX, TIMESERIES...).

SAS High Performance Forecasting est une application d'aide à la production de prévisions.

SAS/IML (*Interactive Matrix Language*) permet d'utiliser un langage matriciel.

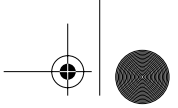
SAS/INSIGHT est un outil de visualisation des données.

Ce module est remplacé depuis SAS 9.2 par **SAS Stat Studio** (www.support.sas.com/rnd/app/studio/studio.html).

SAS/Integration Technologies permet de partager des ressources et d'intégrer SAS dans vos diverses autres applications.

SAS/IntrNet permet l'implémentation de vos applications SAS sur Internet.

SAS/IRP (*Inventory Replenishment Planning*) est un outil de gestion des stocks.



12 ♦ SAS

SAS NLS (*National Language Support*) n'est pas un module en tant que tel mais une application qui permet à SAS de fonctionner au mieux sur des systèmes non américains.

SAS OLAP Server est un SGBD¹ multidimensionnel.

SAS Open Metadata Architecture est une architecture permettant la mise en place, le déploiement et l'administration de la plate-forme SAS.

SAS/OR (*Operational Research*) regroupe des fonctions avancées de recherche opérationnelle comme la programmation (linéaire ou non) et la résolution d'optimisation de graphe.

SAS/SHARE permet l'accès de plusieurs utilisateurs à une même table de données.

SAS/STAT offre les outils de l'analyse statistique comme l'analyse de la variance et la plupart des formes de régression.

Vous pouvez accéder à des descriptifs complets des produits et des solutions proposés par le système SAS sur le site de SAS Institute :

www.sas.com/products/

1. Système de gestion de bases de données.

