# SPSS Magazine

UNE ÉDITION SPSS BELUX ANNÉE 1 NR 1 - MARS 2005
INFORMATION POUR PROFESSIONELS EN RECHERCHE, ANALYSE ET RAPPORTAGE

- L'atelier de data mining Clementine 9.0: un Retour sur investissement encore plus rapide à mesurer
- Le text mining: un trajet critique pour une meilleure analyse commerciale
- Conseils & Astuces: combiner des données
- Témoignage Client › De Boeck Optimiser les processus de production avec LexiQuest de SPSS
- Que pouvez-vous attendre de la dernière version SPSS 13.0 ?
- Formations SPSS

# L'atelier de data mining Clementine 9.0: un Retour sur investissement encore plus rapide à mesurer

**Clementine est déjà connu sur le marché comme l'outil de data mining par excellence, procurant un ROI mesurable dans un temps record. Clementine 9.0 continue dans cette voie. L'efficacité des modèles permet aux utilisateurs de tirer le maximum de leur base de données de données, tout en restant concentré sur leurs objectifs métiers.**

### Toujours plus intégré à vos bases de données

Avec cette dernière version de Clementine, vous êtes à même de mieux réaliser votre préparation de données, la construction de modèle, puis son intégration directement au sein de votre base de données, sans échange lourd de données.
SPSS capitalise sur la capacité de Clementine à exploiter la puissance de manipulation de données des SGDB, et l'étend à la modélisation et au scoring directement dans les bases de données. Cette version renforce l'intégration avec IBM® DB2® Data Warehouse Enterprise Edition et Oracle® Database 10g en permettant la construction et le scoring de modèles directement dans ces bases de données. Elle supporte la modélisation avec IBM® DB2 Intelligent Miner et Oracle Data Mining. Ainsi, Clementine offre un rapport résultat-temps très court en exécutant de façon unique l'ensemble du processus de data mining, de la préparation des données jusqu'à son intégration et son utilisation dans la base de données.

### Intégration de Model Manager™

SPSS propose avec Clementine 9 un nouveau module applicatif optionnel qui assure la gestion d'un référentiel collaboratif des flux, modèles et autres objets créés par une équipe d'analystes. Clementine intègre désormais, avec Model Manager, un moyen rapide pour organiser, retrouver et réutiliser les modèles prédictifs dans d'autres projets de data mining.
De plus, il est sécurisé et il est possible d'indiquer les groupes ayant accès à telle ou telle information. Ce système de gestion implique bien sûr aussi une fonction de recherche très détaillée.

### Le Text mining : à utiliser directement comme extension

L'interface de cette nouvelle version vous permet un accès direct à des fonctionnalités de text mining. Vous pourrez donc aussi directement structurer des données de texte non structurées et enrichir ainsi vos données provenant par exemple de comptes-rendus d'entretien, emails, comptes-rendus de visites et de toute autre source pertinente de textes non structurées. Une société de télécommunications a – en intégrant les comptes-rendus d'entretien du centre d'appel dans le processus d'analyse – amélioré ses prévisions d'attrition clients de 20%. En approchant ces clients de la manière appropriée, on réalise une économie annuelle de quelques millions d'euros.

### Arbres de décision interactifs

Clementine 9.0 vous aide à tester différents modèles et à conclure rapidement sur le plus performant. La technique d'arbre de décision a été élargie avec 3 nouveaux algorithmes : Quest, CHAID et CHAID exhaustif. Les arbres de décision disposent d'un mode interactif pour permettre aux analystes de choisir à chaque niveau de l'arbre la meilleure segmentation sur recommandation de l'algorithme, ou d'utiliser leurs connaissances métiers pour personnaliser un arbre et affiner sa conception. Si vous souhaitez évaluer différentes opportunités durant le développement, la réalisation de graphiques vous donne immédiatement une idée claire des résultats. Si les résultats sont tels qu'ils peuvent être déployés, il vous suffit d'insérer ce modèle directement dans votre processus de data mining. Clementine 9 simplifie également le processus de data mining en fournissant des outils de segmentation automatique pour construire, tester et vérifier les modèles.

### L'utilisation de Clementine dans votre organisation

Souhaitez-vous savoir ce que Clementine peut vous apporter ? Retournez alors **le bulletin réponse** pour obtenir de plus amples renseignements. Si vous souhaitez travailler vous-mêmes avec ce système, vous pourrez alors assister à une **formation d'une journée 'Faire connaissance avec Clementine'** (vous trouverez les dates dans le schéma de formation).

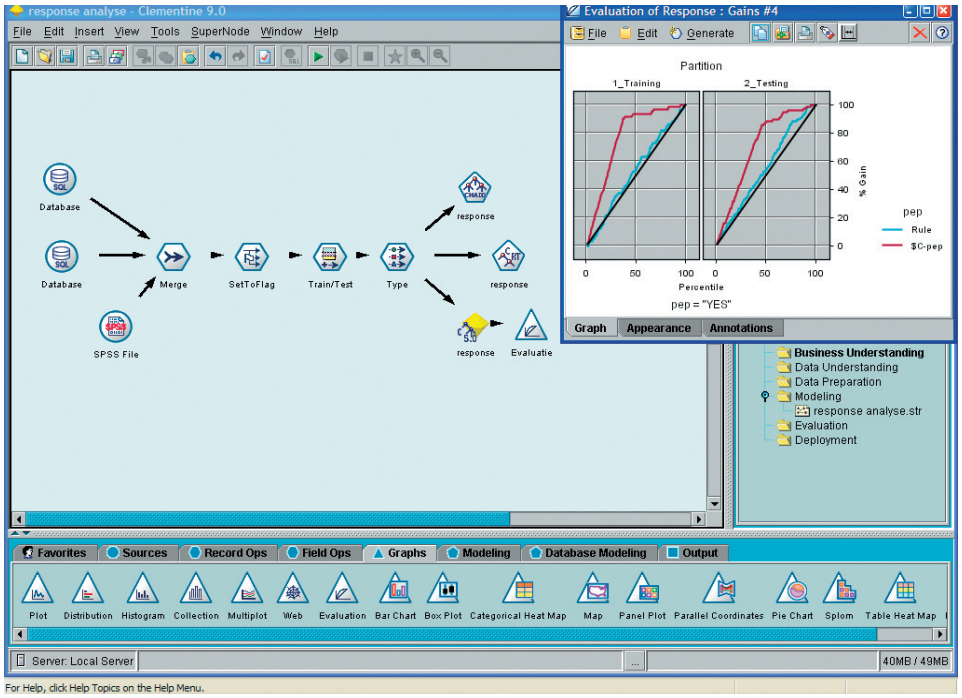



*Illustration: Le graphique d' évaluation vous montre la bonne performance de prévisions de réactions du modèle data mining ($C-pep) dans le présent cas. L' axe inférieur montre l' ensemble de la population, l' axe vertical (gain) montre le nombre total des répondants. La ligne noire est la ligne dite ' 50/50' : si vous devez spéculer sur le quota de réponse de votre public, vous trouveriez en approchant les 50% du public, aussi environ 50% de répondants. Le modèle data mining est à même, dans le jeu d' essai, de trouver environ 88% de tous les répondants dans 51% du public. Il n' est donc pas nécessaire d' approcher 37% de votre public, comme vous auriez dû le faire normalement, pour arriver à 88% de répondants. En fonction des nombres, cela représente une économie sérieuse. De plus, vous voyez dans le graphique la ligne bleue, avec le label ' règle' : c' est la règle commerciale actuelle (basée sur l' intuition) que cette organisation avait employée jusqu' à présent pour la sélection de son groupe cible. En la comparant maintenant au modèle de data mining, vous constaterez que cette intuition fonctionne à peine mieux que la spéculation...*



# Préface

Coordination
Dafne Brouwer

Réalisation
EVV Reclamestudio

Ont collaboré à ce numéro:
Pierre-Yves Thomas (De Boeck)
Jos den Ronden (SPSS)
Wiepke van der Wal (SPSS)
Sophie Van Malder (SPSS)

*Le "SPSS Magazine" est distribué gratuitement via SPSS Belux. Les articles peuvent être réutilisés après accord écrit de la rédaction. Pour toute information complémentaire quant à SPSS et ses produits et solutions n' hésitez pas à contacter notre département ventes. En cas de questions ou problèmes techniques lors de l' utilisation du logiciel notre HelpDesk se tient à votre disposition.*

*SPSS est une marque déposée et les produits annexes de SPSS sont une marque déposée de SPSS Inc. Les autres produits et applications sont la propriété industrielle des organisations respectives.*

SPSS Belux
Buro & Design Center B38
Heizel Esplanade Heysel
1020 Bruxelles
Tel. : +32 2 474 03 60
Fax : +32 2 474 03 63
Email : **belgium@spss.com**
Internet : **www.spss.com/be**

# Le text mining: un trajet critique pour une meilleure analyse commerciale

**La quête de l'analyse complète des activités et des clients d'une entreprise a entraîné des changements considérables en matière d'initiatives du stockage des données.**

Dans leur aspiration à une analyse complète des activités d'entreprise, les organisations ont minutieusement converti leurs informations en matière de transactions, d'opérations et de compétitivité en des « données structurées ». Ceci leurs permet dès lors d'enregistrer ces données afin de commencer l'analyse, grâce aux outils d'intelligence commerciale tels les analyses prédictives et OLAP, et ensuite de rapporter les résultats. L'enregistrement et l'analyse de données structurées ne représentent cependant qu'un seul aspect de la quête d'une meilleure analyse commerciale. Un aspect de l'intelligence commerciale souvent inexploité, est l'information textuelle ou l'information « non structurée ».

### L'analyse du texte est précieuse

D'importantes informations pour l'entreprise sont dissimulées dans les emails, les mémos d'avancement, les rapports de visite, les comptes-rendus régularisés et dans de nombreux autres documents divers. Les informations contenues dans les textes forment une source inexploitée d'intelligence commerciale. Ces informations offrent souvent le savoir de base et les connaissances nécessaires afin de prendre des décisions commerciales dans des conditions spécifiques. En associant les informations extraites aux données structurées, l'analyse du texte peut fournir aux organisations le contexte nécessaire afin de prendre les décisions adéquates.
L'analyse de texte ne figurait pas dans le système d'intelligence commerciale car, jusqu'à il y a peu, il était difficile d'automatiser l'analyse de texte. Aujourd'hui, les technologies de Text Mining permettent d'exploiter les données textuelles de manière plus efficace afin d'effectuer de meilleures analyses. Des compilations de matériel textuel sont dès lors analysées à l'aide de techniques de Text Mining, ce qui permet ainsi de révéler des informations essentielles, des projets, des thèmes cachés, des relations et des tendances.

### Introduction au Text Mining

Outre l'approche manuelle – où l'on examine au peigne fin les textes pour y trouver des informations pertinentes – il existe d'autres possibilités de text mining, bien plus efficaces et effectives. L'une d'elle est la solution automatisée basée sur des statistiques et des réseaux neuraux. Grâce aux techniques informatiques, des projets réels peuvent passer plus rapidement au scanneur et être détectés. Malheureusement la précision de telles solutions est relativement faible. La plupart des systèmes basés sur des statistiques comptent simplement le nombre de présences de termes ou notions et calculent leur analogie statistique avec des termes ou notions reliés, de sorte à produire de nombreux résultats non pertinents (dits aussi « bruissement ») et à ne pas révéler certains résultats qui auraient dû être trouvés (connus sous le terme « silence »).

Une autre solution de Text Mining est la lexicométrie. Cette forme de text mining offre la rapidité et la rentabilité de systèmes orientés sur des statistiques. De plus, elle offre une précision bien plus importante avec une intervention humaine moindre. Les techniques de text mining basées sur la lexicométrie résultent du champ d'examen connu comme Natural Language Processing (NLP), aussi connu sous le nom de lexicométrie informatique.

Pour montrer la différence entre les approches orientées sur les statistiques d'un côté et les approches de lexicométrie, donnons un exemple. Un résultat orienté sur les statistiques comprend probablement une ligne disant que si un texte traite de "management commercial", le mot "manager" sera très vraisemblablement également utilisé dans ce texte. Si les deux termes sont utilisés, le système orienté sur les statistiques sélectionnera le document. Ce qui implique que les textes passés au scanneur devront être suffisamment longs pour servir de source mais rien ne garantit que le mode attendu de l'expression du projet, figure dans le texte. Un utilisateur peut ainsi ne pas remarquer un article intéressant sur le management, si l'auteur a employé le terme de « leader » au lieu de « manager ». Une approche lexicométrique compare les termes tels que "leader", "manager" et même "mngr", s'ils sont employés dans des contextes similaires. De plus, la lexicométrie est à même d'analyser le ton du texte et la compréhension linguistique de la solution met fin à l'ambiguïté propre au texte.

### L'association Text Mining - Data Mining

L'un de l'association data mining et text mining, le text mining prend le dessus quant au pouvoir prédictif. Les techniques prédictives disponibles du data mining, peuvent être appliquées à des concepts issus de données textuelles. Ces concepts sont alors incorporés aux données structurées existantes et l'ensemble de la base de données est évaluée à l'aide du data mining.

L'association text mining et data mining peut être appliquée à toute une gamme de questions commerciales dans de nombreux secteurs d'activité. Par exemple, une société de télécommunication utilisera cette technique d'association afin d'endiguer le problème de l'attrition de la clientèle. Dans ces sociétés, les données sur l'organisation, le client, la démographie, l'utilisation de produits ou d'autres données structurées, seront traitées par un outil de data mining afin de vérifier si un client a tendance à partir. De plus, comme les clients s'expriment souvent via des Call Centers ou via des enquêtes enquêtes, le text mining est la solution pour analyser les réflexions / impressions des clients afin d'identifier les idées principales – mauvais service, trop cher, accès insuffisant – par la suite, ces idées seraient structurées en nouveaux champs et prêts à l'analyse. Le data mining sera alors utilisé pour réunir, classer, segmenter, profiler et prévoir les données afin de calculer les futures réactions du clients (attrition, fidélisation,...).

### Exploitation d'une source riche en informations

Le texte est une source inexploitée, riche en informations et un élément essentiel à l'intelligence commerciale. En incorporant le text mining et le data mining dans leurs systèmes d'intelligence commerciale, les organisations sont en mesure de mieux comprendre leur secteur d'activités et leurs clients.

### Renseignements complémentaires?

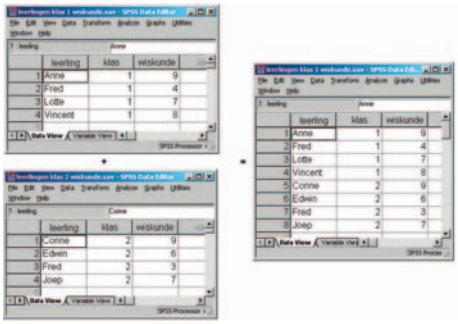
Si vous souhaitez de plus amples renseignements quant au Text Mining ou quant à l'implication de notre solution au sein de votre organisation, n'hésitez en aucun cas à nous retourner le **bulletin réponse au +32 2 474 03 63 .**

# **Conseils & Astuces:** combiner des données (1)

Quelques conseils et astuces quant à la combinaison de données avec le logiciel SPSS.

## Ajout de « cases » (variables égales)

Situation (à gauche: fichier comprenant les données actuelles - à droite: le fichier souhaité).



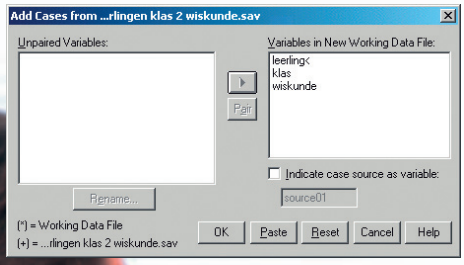
Commentaire:

- Les deux fichiers comprennent exactement les mêmes variables.

Procédure:

- Ouvrir l'un des fichiers, par exemple *élèves classe 1 mathématiques.sav.*
- Choisir: *Data | Merge Files / Add Cases.*
- Sélectionner le second fichier: *élèves classe 2 mathématiques.sav.*

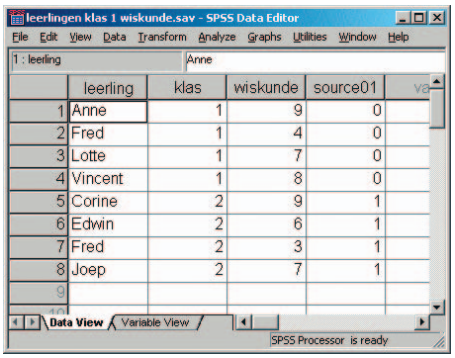
Vous voyez la boîte de dialogue suivante:




Commentaire:

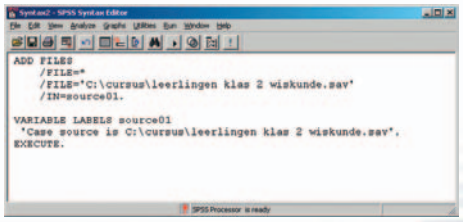
- Sous *Unpaired Variables* se trouvent les variables présentes dans l'un des fichiers, mais pas dans l'autre. Comme les variables sont identiques dans nos deux fichiers, la boîte de dialogue est vide.
- Sous *Variables* in *New Working Data* se trouvent des variables qui apparaîtront dans le fichier combiné.
- Avec l'option *Indicate case source as svariable* vous pouvez décider de créer une variable indiquant de quel fichier le case provient. Nous cochons cette option.

En confirmant par *OK* on obtient le fichier suivant:




Remarque: les « cases » du fichier de travail (le fichier ouvert, en l'occurrence *élèves classe 1 mathématiques.sav*) obtiennent la valeur 0 - *source01*, les « cases » du fichier Add Cases obtiennent la valeur 1.

En syntaxe on obtient le résultat suivant:



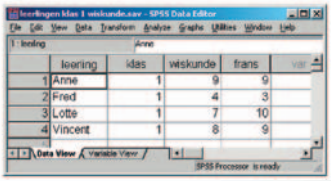

Commentaire:

- Avec la commande partielle /*File* chacun des fichiers est mentionné. L'astérisque * réfère au fichier de travail (le fichier ouvert élèves classe 1 mathématiques.sav).
- Avec la commande partielle /IN le nom de variable donné indique la source.
- Avec la commande VARIABLES LABELS, source01 est clairement définie.

L'option Syntaxe présente l'avantage important de pouvoir combiner plusieurs fichiers en une seule fois, et ce en référant chacun des fichiers à la commande partielle /FILE. Admettons par exemple, que vous vouliez combiner 5 fichiers, il faudrait le réaliser en 4 fois via les boîtes de dialogue, alors que syntaxe le réalise en une seule fois.

## Ajout de variables (« cases » égaux!)

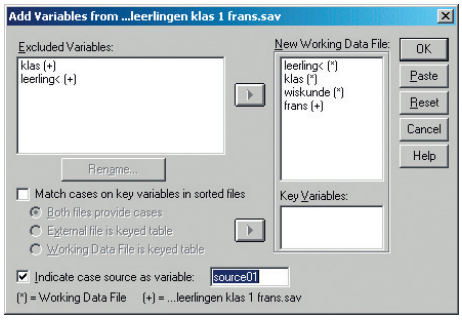
Situation (ci-dessus: les deux fichiers disponibles; ci-dessous: le fichier souhaité).



A toutes fins utiles nous remarquons que les deux fichiers comprennent exactement les mêmes « cases », dans exactement le même ordre. Nous pouvons donc combiner les fichiers en les mettant simplement l'un contre l'autre.

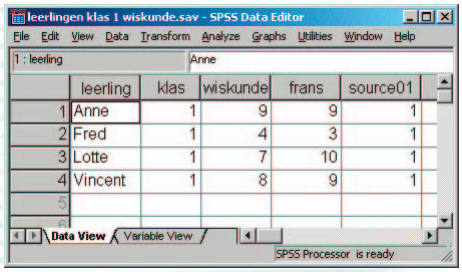
Procédure:
- Ouvrir l'un des deux fichiers, par exemple *élèves classe 1 mathématiques.sav.*
- Choisir: *Data | Merge Files | Add Variables.*
- Sélection du second fichier, ici: *élèves classe 1 français.sav.*

Vous aurez alors la boîte de dialogue suivante:



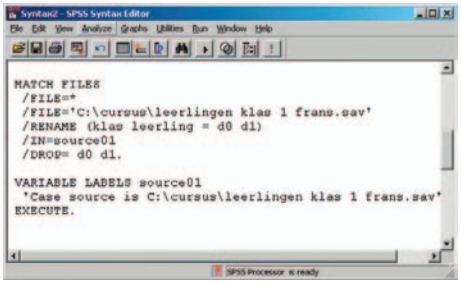

Commentaire:
- Sous *Excluded Variables* se trouvent des variables qui ne seront pas intégrées dans le nouveau fichier. Dans ce cas, ce sont les variables *classe* et *élève*, provenant de *élèves classe 1 français.sav*, témoin le + (en bas de la boîte de dialogue - vous verrez que le + réfère à ce fichier; l'astérisque * réfère au fichier de travail ou au fichier ouvert).
- Sous *New working Data File* figurent les variables qui apparaîtront dans le fichier combiné. Vous verrez également, après chaque variable de quel fichier la variable provient.
- Ici aussi vous pouvez décider à l'aide de l'option *Indicate case source as variable* de créer une variable qui vous indique de quel fichier le « case » vient. Nous cochons cette option.

En confirmant par *OK*, on obtient le fichier suivant:



Remarque: chaque « case » a la valeur 1 sur *source01*. Il indiqu ainsi que chaque « case » est également présent dans le fichier externe *élèves classe 1 français.sav.*

En syntaxe on obtient le résultat suivant:

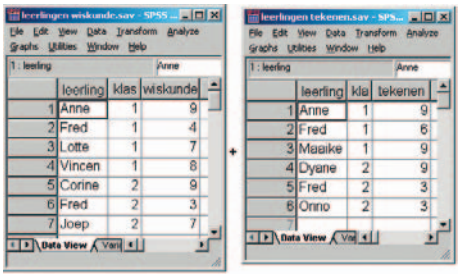



Commentaire:
- Les commandes partielles /*FILE* et /*IN* ont la même fonction que celle décrite ci-dessus à *ADD FILES*.
- A /*RENAME* les variables *classe et élève classe 1 français.sav* sont renommées respectivement *d0* et *d1*.
- Au moyen de /*DROP*, *d0* et *d1* (donc *classe* et *élève classe 1 français.sav*) sont éliminés. L'option Syntaxe présente ici aussi l'avantage, par rapport au menu, de pouvoir combiner plusieurs fichiers de données en 1 seule fois.

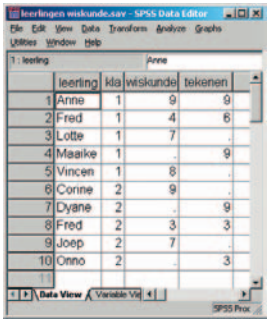
Remarque: dans les deux fichiers les « cases » sont identiques et dans le même ordre. Cette situation ne se présentera certes pas souvent. Il arrivera plus souvent que les « cases » ne sont pas identiques dans les deux fichiers, voir la suite.

## Ajout de variables (« cases » et/ou variables inégaux)

Situation: (ci-dessus: les deux fichiers disponibles - ci-dessous : le fichier souhaité).



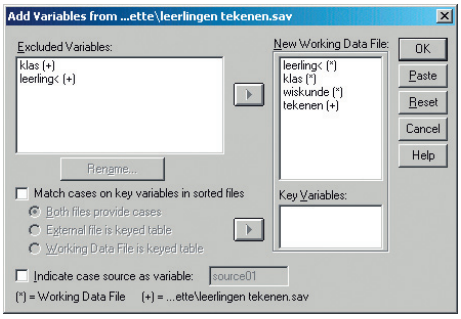

=




Remarques:
- Les fichiers comprennent des « cases » et des variables différents.
- Les fichiers sont classés sur « classe » et sous classes « *élève* ».
- Les classes 1 et 2 comprennent un *Fred*. Il est impossible de mettre les fichiers simplement côte à côte comme dans le cas précédent. Il faut dès lors combiner les deux fichiers en fonction d'une clé: *mathématiques* de *Anne* devra être apposé à *dessin* de *Anne*. Ou bien: Fred de la classe 1 dans l'un des fichiers doit être combiné avec Fred de classe 1 de l'autre fichier. On constate donc que la clé consiste en la combinaison de *classe* et *élève*.

Procédure:
- Sélection des deux fichiers sur *classe*, et en sous classe « *élève* ».
- Ouvrir l'un des deux fichiers, par exemple *élèves mathématiques.sav.*
- Choisir: *Data | Merge Files | Add Variables.*
- Sélection du second fichier, ici: *élèves dessin.sav.*

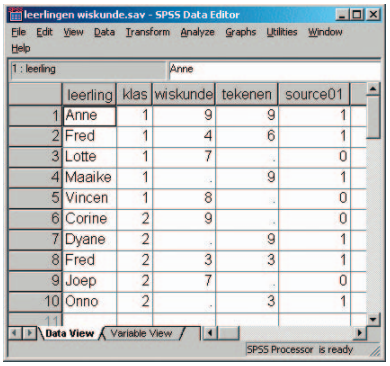
Vous verrez alors la boîte de dialogue suivante:




Commentaire:
- Sous *Excluded Variables* figurent *classe* et *élèves du fichier élèves dessin.sav*, dont témoigne le +.
- Sous New Working *Data File* sont indiquées les variables figurant dans le fichier combiné. Vous verrez après chaque variable son origine.
- Ici aussi vous pouvez décider de créer une variable avec l'option *Indicate case source as variable comme variable* qui indique de quel fichier le « case » vient. Nous activons cette option.

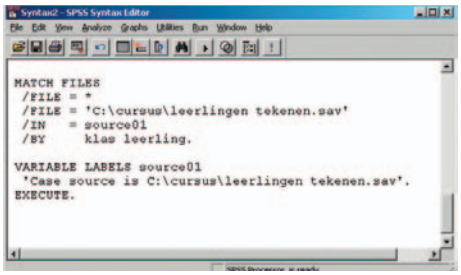
Il importe maintenant d'indiquer que classe et élève forment la clé:
- Cocher l'option *Match cases on key variables in sorted files.*
- Sélectionner *classe* sous *Excluded Variables*: et cliquer sur *Key Variables.*
- Sélectionner *élève* sous *Excluded Variables* et cliquer sur *Key Variables.*

En confirmant par *OK*, on obtient le fichier suivant:




Commentaire:
- Le nouveau fichier comprend tous les « cases » et toutes les variables.
- Lorsque des données sont inconnues (par exemple Lotte n'a pas de note pour *dessin*), figurera en system missings.

La syntaxe peut être utilisée ici aussi.




## Conclusion

Nous avons précédemment examiné trois cas pour combiner des fichiers:
- Ajouter des « cases », en cas de variables égales.
- Ajouter des variables, en cas de « cases » égaux.
- Ajouter à la fois des variables et « cases ».

## Renseignements :

Pour de plus amples renseignements, n'hésitez pas à nous contacter au +32 2 474 03 60 ou via belgium@spss.com.

# **Témoignage Client** › De Boeck
# Optimiser les processus de production avec LexiQuest de SPSS

### La situation

Acteur réputé sur le marché de l'édition avec 220 collaborateurs et un chiffre d'affaires de 32 millions d'euros, le groupe De Boeck compte six maisons d'éditions spécialisées:

- De Boeck Education (éditions scolaires, universitaires et scientifiques)
- Uitgeverij De Boeck (éditions scolaires, universitaires et scientifiques)
- Larcier (publications juridiques)
- Estem (publications médicales)
- Duculot (publications universitaires et scientifiques)
- Lexalis (publications professionelles sur Internet)



Le groupe intègre également une filiale d'édition informatique, qui développe et commercialise des solutions de gestion documentaire et de diffusion sur Internet auprès de maisons d'édition internes ou auprès d'importants clients tels le Parlement Européen, la Cour de Justice et l'Office des Publications Officielles des Communautés Européennes.

### La problématique

En 2002, la filiale d'édition juridique Larcier décide de mettre progressivement en ligne l'ensemble de ses revues via un portail. L'objectif était de proposer aux clients un accès Internet aux revues de leur choix grâce à un abonnement. «Dès la mise en ligne de la première revue, s'est posé le problème d'homogénéiser la taxinomie et ses catégories, pour offrir une arborescence unique d'accès à des articles liés dans des centaines de revues. La complexité était que la taxinomie qui décrit le contenu d'une revue était, sur papier, différente d'une année de publication à l'autre. Il est essentiel pour un client recherchant un sujet spécifique, de trouver tous les articles pertinents – ainsi la seule solution est de lui offrir une taxinomie unique aux revues,» explique Pierre-Yves Thomas, Manager de la filiale DBit. «Fédérer plusieurs revues via cette même arborescence de recherche complexifiait encore le problème.»

De plus, chaque maison d'édition utilisait en interne un processus de workflow manuel pour dispatcher tout article entrante vers ses rédacteurs – un processus assez lent. «Pour optimiser davantage cette chaîne de production, nous souhaitions automatiser la catégorisation à la volée des documents en vue de les orienter rapidement vers les bonnes cibles.»

### La solution

De Boeck constate alors que ces deux problématiques - offrir un accès Web cohérent aux revues juridiques et optimiser le processus interne de production - convergent vers un besoin commun: obtenir une solution de catégorisation automatique des documents. Chargé du projet, DBit effectue une rapide étude de marché quant aux solutions disponibles. Les contraintes pour une telle solution étaient d'avoir la capacité de travailler avec le néerlandais en plus du français, de l'anglais et de l'allemand, et d'offrir une architecture technique suffisamment ouverte (qui supporte le standard XML et fonctionne en configuration client/serveur) pour s'intégrer facilement à la chaîne de production existante. «Seuls deux logiciels répondaient à la contrainte linguistique, dont la solution LexiQuest Categorize de SPSS. Nous l'avons retenue in fine pour son prix, son ouverture, sa facilité de mise en œuvre et pour la réactivité de son éditeur», précise Pierre-Yves Thomas. «Nous avons également bénéficié d'une assistance professionnelle de la société.»

«LexiQuest optimise et fiabilise notre chaîne de production et nous aide à mettre en place un service à valeur ajoutée d'accès en ligne à nos revues.»

**Pierre-Yves Thomas, Manager de la filiale IT DBit du groupe De Boeck**

### Mise en œuvre

Fin mars 2003, le groupe De Boeck acquiert LexiQuest Categorize pour l'ensemble de ses filiales et charge DBit de la mise en œuvre, ce qui inclut l'élaboration des taxonomies, le paramétrage du produit et son déploiement. Aujourd'hui, après une phase de refonte du processus de workflow, le projet d'intégration de Categorize à la chaîne de production est en phase de finalisation. DBit a presque complété la création de la taxinomie (qui va compter à peu près 2000 catégories) pour fédérer l'accès en ligne aux revues (soit 250.000 documents). LexiQuest Categorize est utilisé pour placer, d'une façon automatique, chaque document dans une ou plusieurs de ces 2000 catégories. Le bénéfice attendu de cette approche est d'offrir aux clients une réponse de haute qualité en terme de recherche documentaire. En sélectionnant une catégorie, un client sera capable de trouver en même temps tous les documents liés dans cette catégorie.

### L'intégration

«Les développeurs de DBit, chargés de la mise en œuvre du logiciel, apprécient beaucoup la flexibilité du produit, sa facilité d'intégration et la qualité de sa documentation,» poursuit Pierre-Yves Thomas. «Non seulement le logiciel répond bien à nos besoins internes, mais nous avons aussi constaté qu'il pouvait bien répondre aux besoins de nos clients, qui utilisent notre logiciel de Knowledge Management (Gestion de la connaissance) Strad@ Office. Aussi nous avons décidé, en septembre 2003, de passer un accord contractuel avec SPSS pour intégrer à Strad@ Office, LexiQuest Categorize ainsi que la solution LexiQuest Mine. Le point fort de LexiQuest Mine est que ce logiciel aide à identifier la taxinomie la plus pertinente. Acquise plus tôt, cette solution aurait pu accélérer les phases d'élaboration des taxinomies de nos deux projets, et donc réduire leurs coûts.» L'emploi de LexiQuest Mine va bénéficier directement aux clients de Strad@ Office. Sa nouvelle version, qui est enrichie de LexiQuest Categorize et LexiQuest Mine, suscite beaucoup d'intérêt depuis sa commercialisation officielle fin septembre 2003. A ses fonctions de capture, de recherche et de diffusion de l'information s'ajoute maintenant la possibilité pour chaque client de définir et d'utiliser ses propres taxinomies et catégories.

Et Pierre-Yves Thomas conclut: «Avec l'aide de SPSS, il est devenu possible d'offrir aux clients un accès en ligne rapide et efficace aux revues, pour répondre à leurs besoins de recherche. Nous avons également réduit le temps qu'il faut pour dispatcher les documents nouveaux vers les rédacteurs ciblés. En outre, les clients de Strad@ Office peuvent maintenant définir leurs propres taxinomies et catégoriser de façon automatique leur contenu.»

### Intéressé?

Pour de plus amples renseignements, n'hésitez pas à nous contacter au +32 2 474 03 60 ou via belgium@spss.com.

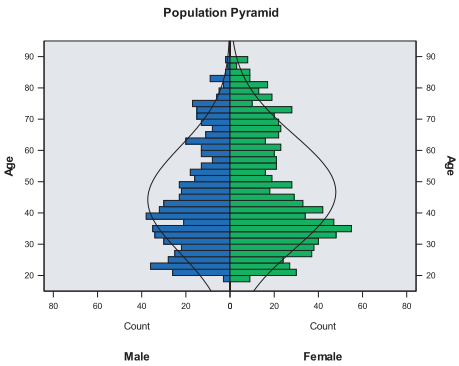
# Que pouvez-vous attendre de la dernière version SPSS 13.0 ?

**SPSS 13.0 ... la dernière version de SPSS, vous offre encore plus que la version 12.0: une amélioration de comptes-rendus/rapports, une identification plus facile des groupes, des analyses prédictives plus étendues et de meilleures possibilités de gestion des données et sorties.**

*Arbre de décision dans SPSS Classification Trees*

### Amélioration des comptes-rendus/rapports

Le moteur graphique nouveau dans la version précédente, a été amélioré et les comptes-rendus sont ainsi mieux présentés. Différents graphiques ont été ajoutés et pour leur présentation, les fonctionnalités ont été élargies – créant ainsi une flexibilité plus grande. De plus, vous pouvez directement exporter des graphiques, des tableaux et des cartes vers MS PowerPoint. Ces modifications ont été apportées au module de base. Mais le module SPSS Base n'est pas le seul à profiter des améliorations du moteur graphique, SPSS Tables bénéficie également d'améliorations.




*Nouveaux moyens graphiques dans SPSS 13.0*

### Davantage d'analyses prédictives

Un nombre d'algorithmes prédictifs a été ajouté à SPSS Complex Samples. Ce module vous permettra donc de mieux composer votre sondage: vous ne faites pas de sondage à l'aveuglette, mais vous considérez la composition de la population. Les algorithmes ajoutés sont les suivants: GLM pour la régression, ANOVA, ANCOVA et la régression logistique pour effectuer des résultats binaires et multi nominaux.

### Optimiser la gestion de données

SPSS Base 13.0 comprend quelques nouveaux éléments permettant une économie de temps et une amélioration de la précision. L'amélioration se trouve notamment au niveau de questions libres, pour la réponse à ces questions, il est maintenant possible d'indiquer davantage de caractères grâce aux 'chaînes de caractères étendues'. En outre, il comprend en plus un Assistant Heure/Date facilitant ainsi le travail avec les heures et dates. Cet Assistant vous permet par exemple de calculer la date de lancement d'un produit et de voir depuis combien de temps les organisations utilisent le produit. Il est également plus facile de calculer l'âge des personnes interrogées grâce à leur date de naissance. Avec auto-recode vous pouvez directement convertir une chaîne de caractères variables en une chaîne de caractères numériques. SPSS peut conserver les codes de catégorie préparés automatiquement, ce qui implique, en cas de modifications dans les catégories, que les codes seront maintenus. Pour agréger des fichiers, vous pouvez alors transférer le résultat directement vers la base de données. Vous disposez de plus d'une nouvelle option 'Division de fichiers', qui vous permet de fractionner une base de données au niveau des lignes et colonnes. Ce qui peut être utile si vous avez par exemple beaucoup de variables et si vous envisagez de manipuler les données en créant ainsi de nouvelles variables à la fin du fichier.

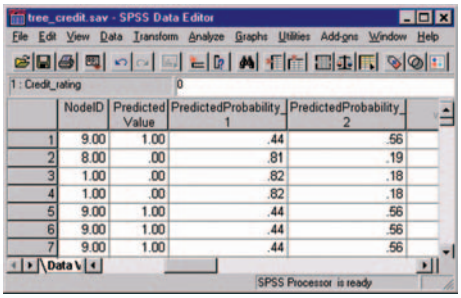
### Nouveau module: SPSS Classification Trees

Ce module complémentaire est entièrement intégré dans le module de base et vous permet d'identifier des groupes, de découvrir des liens entre les groupes et finalement de prévoir le comportement de ces groupes. Les arbres de décision sont visuels, les résultats seront donc évidents pour tout le monde – même pour des personnes non techniques qui veulent uniquement voir le résultat. SPSS Classification Trees comprend quatre algorithmes différents: CHAID, CHAID approfondi, CRT et QUEST. Ces différents algorithmes vous permettront de voir lequel correspond le mieux à vos données.

### Intéressé?

Si vous désirez savoir plus sur les nouveautés de SPSS 13.0 ou si vous êtes intéressé par le nouveau module, n'hésitez pas à nous renvoyer le bulletin réponse ci-joint.






*Avec SPSS Classification Trees, les résultats sont directement intégrés dans votre base de données*

# Trainingschedule SPSS training till July 2005

| | days | April | May | June | July |
|---|---|---|---|---|---|
| **Product training (€ 450,- a day)** | | | | | |
| Introduction to SPSS | 2 | 11-12 | 18-19 | 9-10 | 14-15 |
| SPSS Advanced, Further Data Manipulation | 2 | 21-22 | | | |
| Update to SPSS 13 | 1 | | 23 | | |
| Presenting Data With SPSS Tables, Introduction | 1 | 15 | | | 20 |
| Presenting Data With SPSS Tables, Advanced | 1 | | | 22 | |
| Clementine | 1 | 5 | | | |
| Introduction to Clementine | 2 | | 9-10 | | 4-5 |
| Programming with SPSS Scripts | 1 | | | 17 | |
| **Analytical training (€ 450,- a day)** | | | | | |
| Introduction to Statistics | 2 | 27-28 | | 28-29 | |
| Advanced Techniques: Regression | 2 | | 26-27 | | |
| Introduction to Answer Tree | 1 | | 30 | | |
| **Workshops (€ 1175,-)** | | | | | |
| Data Mining (Crisp-DM) | 2 | | | 6-7 | |



# FORMATIONS

## Apprenez les bases de SPSS en suivant notre formation 'Introduction to SPSS'

Vous êtes un nouvel utilisateur ? Vous débutez et souhaitez acquérir plus d'expérience, travailler de manière plus efficace et productive ? Nous vous proposons notre formation « Introduction to SPSS ».

Les sessions en français auront lieu les 11-12 Avril, 18-19 Mai, 9-10 Juin, 14-15 Juillet.

## Vous désirez approfondir vos connaissances ?

Vous utilisez déjà SPSS et vous souhaitez approfondir vos connaissances, n'hésitez pas à suivre le cours « SPSS Advanced, Further Data Manipulation ».

Les sessions en française se dérouleront les 21-22 Avril.

### Désirez-vous vous inscrire à l'une de nos formations ?

N'hésitez pas à nous retourner le formulaire ci-joint. Le prix de nos formations publiques s'élève à 450,00 € par jour par personne.