

IBM SPSS Conjoint 19



Note: Before using this information and the product it supports, read the general information under Notices sur p. 42.

This document contains proprietary information of SPSS Inc, an IBM Company. It is provided under a license agreement and is protected by copyright law. The information contained in this publication does not include any product warranties, and any statements provided in this manual should not be interpreted as such.

When you send information to IBM or SPSS, you grant IBM and SPSS a nonexclusive right to use or distribute the information in any way it believes appropriate without incurring any obligation to you.

© Copyright SPSS Inc. 1989, 2010.

Préface

IBM® SPSS® Statistics est un système complet d'analyse de données. Le module complémentaire facultatif Conjoint fournit les techniques d'analyse supplémentaires décrites dans ce manuel. Le module complémentaire Conjoint doit être utilisé avec le système central SPSS Statistics auquel il est entièrement intégré.

A propos de SPSS Inc., an IBM Company

SPSS Inc., an IBM Company, est un des leaders dans le domaine des solutions logicielles d'analyse prédictive. Le portfolio complet des produits de la société — Data collection, Statistics, Modeling et Deployment — capture les opinions et les attitudes du public, prédit les résultats des interactions futures des clients, et agit ensuite sur ces données en intégrant les analyses dans les processus commerciaux. Les solutions SPSS Inc. répondent aux objectifs commerciaux interdépendants d'une organisation dans sa totalité en se concentrant sur la convergence des analyses, de l'architecture informatique et des processus commerciaux. Des clients issus du milieu des affaires, du milieu gouvernemental ou du milieu académique, dans le monde entier, font confiance à la technologie SPSS Inc., et la considère comme un atout pour attirer et retenir leurs clients, ou encore augmenter leur nombre, tout en réduisant les fraudes et les risques. SPSS Inc. a été acheté par IBM en octobre 2009. Pour plus d'informations, visitez le site <http://www.spss.com>.

Support technique

Un support technique est disponible pour les clients du service de maintenance. Les clients peuvent contacter l'assistance technique pour obtenir de l'aide concernant l'utilisation des produits SPSS Inc. ou l'installation dans l'un des environnements matériels pris en charge. Pour contacter l'assistance technique, consultez le site Web SPSS Inc. à l'adresse <http://support.spss.com>, ou recherchez votre représentant local à la page <http://support.spss.com/default.asp?refpage=contactus.asp> Votre nom, celui de votre société, ainsi que votre contrat d'assistance vous seront demandés.

Service clients

Si vous avez des questions concernant votre envoi ou votre compte, contactez votre bureau local, dont les coordonnées figurent sur le site Web à l'adresse : <http://www.spss.com/worldwide>. Veuillez préparer et conserver votre numéro de série à portée de main pour l'identification.

Séminaires de formation

SPSS Inc. propose des séminaires de formation, publics et sur site. Tous les séminaires font appel à des ateliers de travaux pratiques. Ces séminaires seront proposés régulièrement dans les grandes villes. Pour plus d'informations sur ces séminaires, contactez votre bureau local dont les coordonnées sont indiquées sur le site Web à l'adresse : <http://www.spss.com/worldwide>.

Documents supplémentaires

Les ouvrages *SPSS Statistics : Guide to Data Analysis*, *SPSS Statistics : Statistical Procedures Companion*, et *SPSS Statistics : Advanced Statistical Procedures Companion*, écrits par Marija Norušis et publiés par Prentice Hall, sont suggérés comme documentation supplémentaire. Ces publications présentent les procédures statistiques des modules SPSS Statistics Base, Advanced Statistics et Regression. Que vous soyez novice dans les analyses de données ou prêt à utiliser des applications plus avancées, ces ouvrages vous aideront à exploiter au mieux les fonctionnalités offertes par IBM® SPSS® Statistics. Pour obtenir des informations supplémentaires y compris le contenu des publications et des extraits de chapitres, visitez le site web de l'auteur : <http://www.norusis.com>

Contenu

| | | |
|---|---|-----------|
| 1 | <i>Introduction à l'analyse conjointe</i> | 1 |
| | L'approche Profil entier | 2 |
| | Un tableau orthogonal | 2 |
| | Les stimuli expérimentaux | 2 |
| | Collecte et analyse des données | 3 |
| Partie I: Guide de l'utilisateur | | |
| 2 | <i>Génération d'un plan orthogonal</i> | 5 |
| | Définition des valeurs d'un plan orthogonal | 7 |
| | Options du plan orthogonal | 8 |
| | Fonctionnalités supplémentaires de la commande ORTHOPLAN | 8 |
| 3 | <i>Affichage d'un plan</i> | 9 |
| | Afficher plan orthogonal : Titres | 10 |
| | Fonctionnalités supplémentaires de la commande PLANCARDS | 10 |
| 4 | <i>Exécution d'une analyse conjointe</i> | 11 |
| | Configuration requise | 11 |
| | Spécification du fichier de plan et du fichier de données | 12 |
| | Spécification de la manière dont les données ont été enregistrées | 12 |
| | Sous-commandes facultatives | 13 |

Partie II: Exemples

5 *Utilisation de l'analyse conjointe pour le modèle Préférence de la shampouineuse* **17**

| | |
|---|----|
| Génération d'un plan orthogonal | 18 |
| Création de stimuli expérimentaux : Affichage du plan | 21 |
| Exécution de l'analyse | 23 |
| Scores des outils | 25 |
| Coefficients | 26 |
| Importance relative | 26 |
| Corrélations | 27 |
| Inversements | 28 |
| Exécution de simulations | 28 |
| Probabilités des préférences des simulations | 29 |

Annexes

A *Fichiers d'exemple* **31**

B *Notices* **42**

Bibliographie **44**

Index **46**

Introduction à l'analyse conjointe

L'analyse conjointe est un outil d'études de marché pour le développement d'un design de produit effectif. En utilisant l'analyse conjointe, le chercheur peut répondre aux questions suivantes : Quelle est l'importance de tel ou tel attribut du produit pour le consommateur ? Quelle est l'importance de certains niveaux d'attributs de produits par rapport à d'autres dans l'esprit du consommateur ? Quelles sont les parts de marché des produits leaders de la concurrence par rapport à nos produits existants ou proposés ?

L'avantage de l'analyse conjointe est qu'elle demande au répondant de faire des choix similaires à ceux que ferait le consommateur en comparant des fonctionnalités les unes aux autres.

Par exemple, supposons que vous souhaitez réserver un billet d'avion. Vous avez le choix entre un siège étroit et un siège spacieux. Si cela était la seule considération, votre choix serait clair. Vous préféreriez certainement le siège spacieux. Supposez maintenant que vous avez le choix du prix du billet : 183 ou 652 €. En vous basant uniquement sur le prix, vous choisiriez le plus bas. Enfin, supposons que vous pouvez prendre soit un vol direct de deux heures ou soit un vol avec escale d'une durée de cinq heures. La plupart des gens choisiraient le vol direct.

L'inconvénient d'une approche comme celle-ci réside dans le fait que les alternatives de choix sont présentées avec un seul attribut à la fois. L'analyse conjointe présente des propositions d'alternatives de choix entre les produits, définies par des ensembles d'attributs. Ceci est illustré par le choix suivant : Préférez-vous un vol avec escale, dans un siège étroit pour 183 € ou un vol direct dans un siège spacieux pour 652 € ? En considérant le confort, le prix et la durée comme des attributs appropriés, voici les huit produits potentiels :

| Produit | Confort | Price | Durée |
|---------|----------|-------|----------|
| 1 | étroit | \$225 | 2 heures |
| 2 | étroit | \$225 | 5 heures |
| 3 | étroit | \$800 | 2 heures |
| 4 | étroit | \$800 | 5 heures |
| 5 | spacieux | \$225 | 2 heures |
| 6 | spacieux | \$225 | 5 heures |
| 7 | spacieux | \$800 | 2 heures |
| 8 | spacieux | \$800 | 5 heures |

Etant donné les alternatives ci-dessus, le produit 4 est sans doute le moins intéressant et le produit 5 celui qui a la préférence des répondants. Les préférences des répondants pour les autres offres de produit sont déduites implicitement à partir de ce qui semble important aux yeux du répondant.

En utilisant l'analyse conjointe, vous pouvez déterminer à la fois l'importance relative de chaque attribut ainsi que les niveaux des attributs qui semblent avoir la préférence des répondants. Si le produit ayant la préférence des répondants n'est pas envisageable pour certaines raisons

comme le coût, vous avez déjà connaissance de la seconde alternative préférée par les répondants. Si vous disposez d'autres informations sur les répondants, comme des données démographiques, vous pourrez identifier les segments de marché pour lesquels des produits distincts pourront être destinés. Par exemple, un homme d'affaires et un étudiant qui voyagent peuvent avoir des préférences différentes auxquelles pourront répondre des offres de produit distinctes.

L'approche Profil entier

La procédure conjointe utilise l'**approche Profil entier** (aussi connue sous le nom Concept entier) où les répondants classent, trient ou notent un ensemble de profils ou de cartes selon leurs préférences. Chaque profil décrit un produit ou un service complet et consiste en une combinaison différente de niveaux de facteurs pour tous les facteurs (attributs) d'intérêt.

Un tableau orthogonal

Un problème potentiel survient avec l'approche Profil entier et devient clair lorsque beaucoup de facteurs sont concernés et que chacun de ces facteurs comportent plusieurs niveaux. Le nombre total de profils résultant de toutes les combinaisons possibles de niveaux devient trop important pour que les répondants puissent les classer ou les noter de façon cohérente. Pour résoudre ce problème, l'approche du profil entier utilise ce qu'on appelle un **modèle factoriel fractionnel** qui présente une fraction appropriée de toutes les combinaisons possibles de niveaux de facteurs. L'ensemble qui en résulte, appelé un **tableau orthogonal** est conçu pour saisir les effets principaux de chaque niveau de facteur. On suppose que les interactions entre les niveaux d'un facteur avec les niveaux d'un autre facteur sont négligeables.

La procédure de génération de plan orthogonal est utilisée pour générer un tableau orthogonal et est habituellement le point de départ pour une analyse conjointe. Cela vous permet également de générer des combinaisons de niveaux de facteurs, appelées **observations traitées**, qui sont classées par les sujets, mais pas utilisées pour la construction du modèle de préférence. A la place, elles sont utilisées pour vérifier la validité du modèle.

Les stimuli expérimentaux

Chaque ensemble de niveaux de facteurs est un plan orthogonal représentant une version différente du produit étudié et doit être présenté aux sujets sous la forme d'un profil de produit individuel. Ceci permet au répondant de se concentrer uniquement sur le produit unique actuellement évalué. Les stimuli doivent être normalisés en s'assurant que les profils ont tous une apparence physique similaire, sauf pour les différentes combinaisons de fonctionnalités.

La création de profils de produits est facilitée par la procédure Afficher plan orthogonal. Elle prend un plan généré par la procédure de génération de plan orthogonal ou entré par l'utilisateur et produit un ensemble de profils de produits dans un format prêt à utiliser.

Collecte et analyse des données

Puisque il y a souvent un grand nombre de différences entre les préférences des différents sujets, la plupart des analyses conjointes portent sur un sujet unique. Pour généraliser les résultats, un échantillon aléatoire de sujet est sélectionné dans la population cible afin que les résultats de groupe puissent être examinés.

La taille de l'échantillon dans les études conjointes varie grandement. Dans un rapport (Cattin et Wittink, 1982), les auteurs précisent que la taille d'un échantillon dans des études commerciales conjointes est comprise entre 100 et 1 000, 300 à 550 étant la plage la plus courante. Dans une autre étude (Akaah et Korgaonkar, 1988), il est dit que les échantillons plus petits (moins de 100) sont les plus courants. Comme toujours, la taille de l'échantillon doit être suffisamment grande pour assurer la fiabilité des résultats.

Une fois l'échantillon choisi, le chercheur distribue l'ensemble des profils ou cartes à chaque répondant. La procédure conjointe permet trois méthodes d'enregistrement des données. Dans la première méthode, on demande aux sujets d'assigner un score à chaque profil, selon leurs préférences. Ce type de méthode est courant lorsqu'une échelle Likert est utilisée ou lorsque l'on demande aux sujets d'assigner un nombre entre 1 et 100 pour indiquer leurs préférences. Dans la seconde méthode, on demande aux sujets d'affecter un rang à chaque profil, de 1 jusqu'au nombre total de profils. Dans la troisième méthode, on demande aux sujets de trier les profils en termes de préférences. Avec cette dernière méthode, le chercheur enregistre le numéro des profils dans l'ordre donné par chaque sujet.

L'analyse des données est réalisée avec la procédure conjointe (disponible uniquement via la syntaxe de commande) et résulte en une note utilitaire appelée **valeur de partie** pour chaque niveau de facteur. Ces notes outils, semblables à des coefficients de régression, fournissent une mesure quantitative de préférence pour chaque niveau de facteur, les valeurs les plus grandes correspondant aux préférences les plus marquées. Les valeurs de partie sont exprimées dans une unité commune, leur permettant ainsi d'être ajoutées pour donner l'outil total, ou préférence générale, pour toute combinaison de niveaux de facteurs. Les valeurs de partie constituent ensuite un modèle de prévision de la préférence de tout profil de produit, y compris les profils appelés **observations de simulation** qui n'étaient pas présentés lors de l'expérience.

Les informations obtenues via une analyse conjointe peuvent être appliquées à une large variété de questions d'études de marché. Elles peuvent être utilisées pour étudier des domaines tels que le design des produits, les parts de marché, la publicité stratégique, les analyses coûts-bénéfices et la segmentation du marché.

Bien que ce manuel mette l'accent sur les applications d'étude de marché, l'analyse conjointe peut également s'avérer utile dans la plupart des applications scientifiques ou professionnelles où la mesure de la perception ou du jugement des gens s'avère importante.

Partie I: Guide de l'utilisateur

Génération d'un plan orthogonal

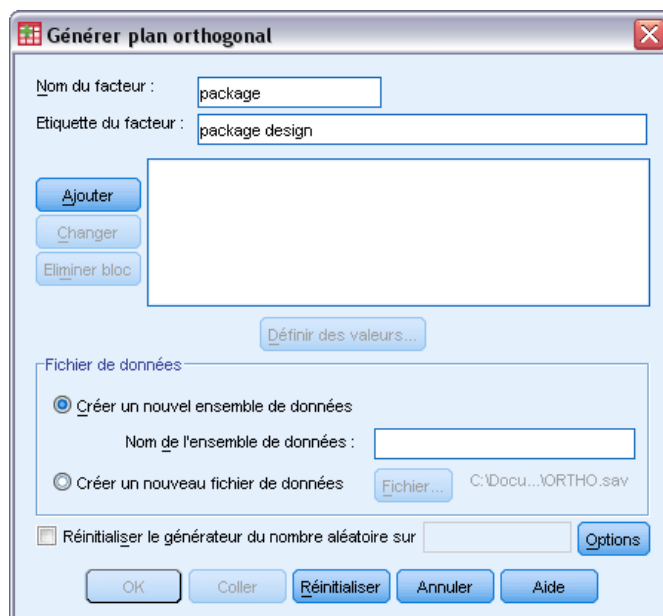
Générer plan orthogonal génère un fichier de données comprenant un plan orthogonal avec des effets principaux, qui permet de tester les statistiques de plusieurs critères sans tester chaque combinaison de niveaux de facteur. Ce plan peut être affiché au moyen de la procédure d'affichage en plan orthogonal et le fichier de données peut être utilisé par d'autres procédures, telles que la procédure conjointe.

Exemple : Une nouvelle compagnie aérienne à bas coûts souhaiterait connaître l'importance relative de divers facteurs qui composent son offre de produits auprès de clients potentiels. Le prix est clairement un facteur primaire, mais quelle est l'importance d'autres facteurs comme la taille des sièges, le nombre d'escaliers ou le fait que des boissons ou repas soient offerts ? Une enquête demandant aux répondants de classer les profils de produits représentant toutes les combinaisons de facteurs possibles n'est pas envisageable étant donné le grand nombre de profils. La procédure de génération de plan orthogonal permet de créer un ensemble de profils de produits réduit, assez petit pour être inclus dans une enquête mais assez grand pour mesurer l'importance relative de chaque facteur.

Pour générer un plan orthogonal

- ▶ A partir des menus, sélectionnez :
Données > Plan orthogonal > Générer...

Figure 2-1
Boîte de dialogue Générer plan orthogonal



- ▶ Définissez au moins un critère. Entrez un nom dans la zone de texte Nom du facteur. Le nom des facteurs peut être un nom de variable valide, à l'exception de *status_* ou *card_*. Vous pouvez également affecter une étiquette de facteur facultative.
- ▶ Cliquez sur Ajouter pour ajouter le nom du facteur et une étiquette facultative. Pour supprimer un facteur, sélectionnez-le dans la liste et cliquez sur Supprimer. Pour modifier une étiquette ou le nom d'un facteur, sélectionnez-le dans la liste, modifiez le nom ou l'étiquette puis cliquez sur Modifier.
- ▶ Définissez les valeurs de chaque critère en sélectionnant les critères et en cliquant sur Définir les valeurs.

Fichier de données. Vous permet de contrôler la destination du plan orthogonal. Vous pouvez enregistrer le plan dans un nouvel ensemble de données de la session en cours ou dans un fichier de données externe.

- **Créer un nouvel ensemble de données.** Crée un nouvel ensemble de données dans la session en cours contenant les facteurs et les observations générés par le plan.
- **Créer un nouveau fichier de données.** Crée un fichier de données externe contenant les facteurs et les observations générés par le plan. Par défaut, ce fichier est nommé *ortho.sav* et enregistré dans le répertoire actuel. Cliquez sur Fichier pour définir un nom et une destination de fichier différents.

Réinitialiser le générateur du nombre aléatoire sur. Réinitialise le générateur de nombre aléatoire sur la valeur spécifiée. Le générateur peut être une valeur entière comprise entre 0 et 2 000 000 000. Au sein d'une session, un générateur différent est utilisé à chaque génération d'un ensemble de nombres aléatoires, produisant ainsi des résultats différents. Pour dupliquer des nombres aléatoires, vous devez définir la valeur du générateur avant de générer votre premier plan et réinitialiser le générateur sur cette valeur à chaque génération de votre plan.

Sinon, vous pouvez :

- Cliquez sur Options pour définir le nombre minimum d'observations dans le plan orthogonal et pour sélectionner les observations traitées.

Définition des valeurs d'un plan orthogonal

Figure 2-2

Boîte de dialogue Générer plan orthogonal : Définir les valeurs

| | Valeur | Etiquette |
|----|--------|-----------|
| 1: | 1 | A* |
| 2: | 2 | B* |
| 3: | 3 | C* |
| 4: | | |
| 5: | | |
| 6: | | |
| 7: | | |
| 8: | | |
| 9: | | |

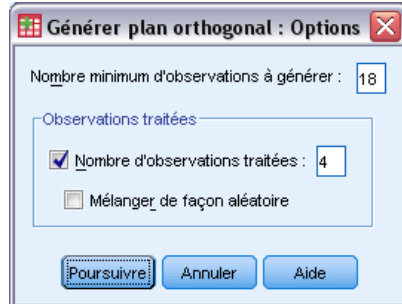
Vous devez affecter des valeurs à chaque niveau des critères sélectionnés. Le nom du facteur sera affiché après Valeurs et étiquettes pour.

Saisissez chaque valeur du facteur. Vous pouvez choisir d'attribuer des étiquettes descriptives aux valeurs. Si vous n'affectez aucune étiquette aux valeurs, les étiquettes correspondant aux valeurs sont automatiquement affectées (c'est-à-dire, l'étiquette 1 est affectée à la valeur 1, l'étiquette 3 est affectée à la valeur 3, etc.).

Remplissage auto. Permet de renseigner automatiquement les zones Valeur avec des valeurs consécutives commençant par 1. Saisissez la valeur maximale et cliquez sur Remplir pour renseigner les valeurs.

Options du plan orthogonal

Figure 2-3
Boîte de dialogue Générer plan orthogonal : Options



Nombre minimum d'observations à générer. Indique le nombre minimum d'observations relatives au plan. Sélectionnez un entier positif inférieur ou égal au nombre total d'observations issues de toutes les combinaisons possibles des niveaux de facteur. Si vous n'indiquez pas explicitement le nombre minimum d'observations à générer pour le plan orthogonal, ce nombre est automatiquement généré. Si la procédure Orthoplan ne peut pas générer au moins le nombre minimal de profils requis, elle génère le plus grand nombre possible correspondant aux critères et aux niveaux spécifiés. Remarquez que le plan ne comporte pas nécessairement le nombre exact d'observations spécifiées ; il comporte plutôt le plus petit nombre d'observations possible dans le plan orthogonal en utilisant cette valeur comme valeur minimale.

Observations traitées. Vous pouvez définir les observations traitées qui sont évaluées par sujets, mais qui ne pas incluses dans l'analyse conjointe.

- **Nombre d'observations traitées.** Crée des observations traitées en plus des observations relatives au plan standard. Les observations traitées sont évaluées par sujets, mais ne sont pas utilisées lorsque la procédure Conjointe évalue les utilitaires. Vous pouvez indiquer un entier positif inférieur ou égal au nombre total d'observations issues de toutes les combinaisons possibles des niveaux de facteur. Les observations traitées sont générées à partir d'un autre plan aléatoire et non à partir du plan expérimental avec effets principaux. Les observations traitées ne dupliquent pas les profils expérimentaux ou ne se dupliquent pas entre elles. Par défaut, aucune observation traitée n'est produite.
- **Mélanger de façon aléatoire.** Mélange de façon aléatoire les observations traitées aux observations expérimentales. Lorsque cette option est désélectionnée, les observations traitées apparaissent séparément, après les observations expérimentales.

Fonctionnalités supplémentaires de la commande ORTHOPLAN

Le langage de syntaxe de commande vous permet aussi de :

- Ajouter le plan orthogonal à l'ensemble de données actif au lieu d'en créer un nouveau.
- Spécifier les observations de simulation avant de générer le plan orthogonal, plutôt que de le faire après.

Pour obtenir des renseignements complets sur la syntaxe, reportez-vous au manuel *Command Syntax Reference*.

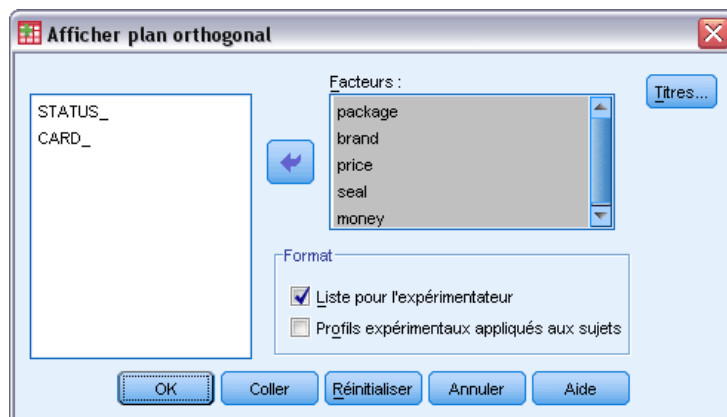
Affichage d'un plan

La procédure d'affichage en plan orthogonal permet d'imprimer un plan expérimental. Vous pouvez imprimer le plan sous forme de liste en mode brouillon ou en profils que vous pouvez appliquer aux sujets dans une étude commune. Cette procédure permet d'afficher les plans créés avec la procédure de génération de plan orthogonal ou avec n'importe quel plan affiché dans un ensemble de données actif.

Pour afficher un plan orthogonal

- ▶ A partir des menus, sélectionnez :
Données > Plan orthogonal > Afficher...

Figure 3-1
Boîte de dialogue Afficher plan orthogonal



- ▶ Placez un ou plusieurs critères dans la liste Facteurs.
- ▶ Sélectionnez un format pour l'affichage des profils dans le résultat.

Format : Vous pouvez choisir parmi les options de format suivantes :

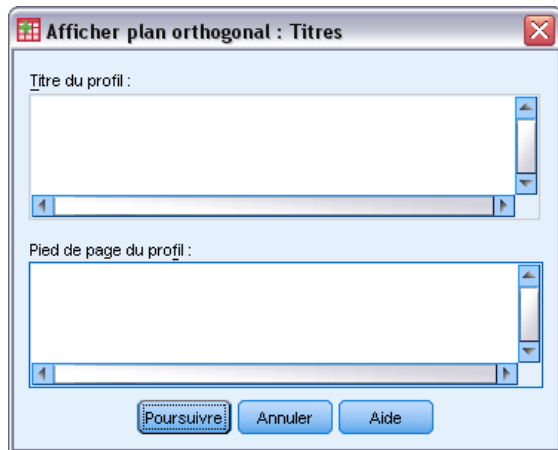
- **Liste pour l'expérimentateur :** Affiche le plan au format brouillon qui différencie les profils traités des profils expérimentaux et répertorie les profils de simulation séparément en suivant les profils traités et expérimentaux.
- **Profils expérimentaux appliqués aux sujets :** Génère des profils pouvant être appliqués aux sujets. Ce format ne différencie pas les profils traités et ne produit aucun profil de simulation.

Sinon, vous pouvez :

- Cliquez sur Titres pour définir des en-têtes et des pieds de page pour les profils.

Afficher plan orthogonal : Titres

Figure 3-2
Boîte de dialogue Afficher plan orthogonal : Titres



Titre de protocole : Saisissez jusqu'à 80 caractères pour le titre du profil. Les titres apparaissent en haut du résultat si vous avez sélectionné Liste pour l'expérimentateur et en haut de chaque nouveau profil si vous avez sélectionné Profils appliqués aux sujets dans la boîte de dialogue principale. Pour Profils expérimentaux appliqués aux sujets, si la séquence de caractères spéciaux)CARD est spécifiée quelque part dans le titre, la procédure la remplacera par le numéro séquentiel du profil. La séquence des caractères n'est pas traduite pour Liste pour l'expérimentateur.

Bas de page de protocole : Saisissez jusqu'à 80 caractères pour le pied de page du profil. Les pieds de page apparaissent en bas du résultat si vous avez sélectionné Liste pour l'expérimentateur et en bas de chaque nouveau profil si vous avez sélectionné Profils appliqués aux sujets dans la boîte de dialogue principale. Pour Profils expérimentaux appliqués aux sujets, si la séquence de caractères spéciaux)CARD est spécifiée quelque part dans le bas de page, la procédure la remplacera avec le numéro séquentiel du profil. La séquence des caractères n'est pas traduite pour Liste pour l'expérimentateur.

Fonctionnalités supplémentaires de la commande PLANCARDS

Le langage de syntaxe de commande vous permet aussi de :

- Ecrire des profils expérimentaux appliqués aux sujets dans un fichier externe (en utilisant la sous-commande `OUTFILE`).

Pour obtenir des renseignements complets sur la syntaxe, reportez-vous au manuel *Command Syntax Reference*.

Exécution d'une analyse conjointe

Une interface graphique utilisateur pour la procédure conjointe n'est pas encore disponible. Pour obtenir une analyse conjointe, vous devez entrer la syntaxe de commande pour la commande `CONJOINT` dans la fenêtre de syntaxe, puis l'exécuter.

- Pour obtenir un exemple d'une syntaxe de commande pour la commande `CONJOINT` dans le contexte d'une analyse conjointe complète, y compris la création et l'affichage d'un plan orthogonal, reportez-vous au [le chapitre 5](#).
- Pour plus d'informations sur la syntaxe de commande à propos de la commande `CONJOINT`, consultez *Command Syntax Reference*.

Pour exécuter une commande à partir d'une fenêtre de syntaxe

A partir des menus, sélectionnez :

Fichier > Nouveau > Syntaxe...

Cette opération ouvre une fenêtre de syntaxe.

- ▶ Entrez la syntaxe de commande pour la commande `CONJOINT`.
- ▶ Sélectionnez la commande dans la fenêtre de syntaxe et cliquez sur le bouton Exécuter (le triangle pointant vers la droite) dans la barre d'outils de l'Editeur de syntaxe.

Reportez-vous au manuel *Guide de l'utilisateur du système central* pour obtenir plus d'informations sur l'exécution de commandes dans les fenêtres de syntaxe.

Configuration requise

La procédure conjointe requiert deux fichiers (un fichier de données et un fichier de plan) et la spécification de la manière dont les données ont été enregistrées (par exemple, chaque point de donnée correspond à un score de préférence compris entre 1 et 100). Le fichier de plan consiste en un ensemble de profils de produits à noter par les sujets et doit être généré à l'aide de la procédure [Générer plan orthogonal](#). Le fichier de données contient les scores de préférence ou le classement de ces profils collectés chez les sujets. Les fichiers de plan et de données sont respectivement spécifiés à l'aide des sous-commandes `PLAN` et `DATA`. La méthode d'enregistrement des données est spécifiée à l'aide des sous-commandes `SEQUENCE`, `RANK` ou `SCORE`. La syntaxe de commande suivante illustre la spécification minimale :

```
CONJOINT PLAN='CPLAN.SAV' /DATA='RUGRANKS.SAV'  
/SEQUENCE=PREF1 TO PREF22.
```

Spécification du fichier de plan et du fichier de données

La commande `CONJOINT` fournit plusieurs options pour la spécification du fichier de plan et du fichier de données.

- Vous pouvez spécifier de manière explicite les noms de fichier des deux fichiers. Par exemple :

```
CONJOINT PLAN='CPLAN.SAV' /DATA='RUGRANKS.SAV'
```

- Si seul un fichier de plan ou un fichier de données est spécifié, la commande `CONJOINT` lit le fichier spécifié et utilise l'ensemble de données actif pour le second. Par exemple, si vous spécifiez un fichier de données mais que vous omettez de spécifier un fichier de plan (vous devez spécifier un fichier au moins sur les deux), l'ensemble de données actif est utilisé comme plan, comme illustré dans l'exemple qui suit :

```
CONJOINT DATA='RUGRANKS.SAV'
```

- Vous pouvez utiliser le symbole astérisque (*) pour indiquer un ensemble de données au lieu d'un nom de fichier, comme illustré dans l'exemple qui suit :

```
CONJOINT PLAN='CPLAN.SAV' /DATA=*
```

L'ensemble de données actif est utilisé comme donnée de préférence. Notez que vous ne pouvez pas utiliser l'astérisque (*) à la fois pour le fichier de données et le fichier de plan.

Spécification de la manière dont les données ont été enregistrées

Vous devez spécifier la manière dont les données de préférence ont été enregistrées. Les données peuvent être enregistrées de trois manières différentes : de manière séquentielle, sous forme de classement ou sous forme de scores de préférence. Ces trois méthodes sont indiquées par les sous-commandes `SEQUENCE`, `RANK` et `SCORE`. Vous ne devez spécifier qu'une seule de ces sous-commandes uniquement au sein d'une commande `CONJOINT`.

Sous-commande SEQUENCE

La sous-commande `SEQUENCE` indique que les données ont été enregistrées de manière séquentielle de manière à ce que chaque point de données dans le fichier de données soit un numéro de profil, en commençant par le profil préféré et en terminant par le moins intéressant. C'est l'ordre dans lequel les données sont enregistrées si l'on demande au sujet de classer les profils du plus intéressant au moins intéressant. Le chercheur enregistre le premier numéro de profil, le second, et ainsi de suite.

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'  
/SEQUENCE=PREF1 TO PREF22.
```

- La variable `PREF1` contient le numéro de profil du profil préféré parmi les 22 profils du plan orthogonal. La variable `PREF22` contient le numéro de profil du profil le moins intéressant dans le plan.

Sous-commande RANK

La sous-commande RANK indique que chaque point de données correspond à un classement, commençant par le classement du profil 1, puis celui du profil 2, et ainsi de suite. C'est la manière selon laquelle les données sont enregistrées lorsque l'on demande au sujet d'attribuer un numéro de classement à chaque profil, compris entre 1 et n , où n correspond au nombre de profils. Plus le numéro de classement est petit, plus grande est la préférence.

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'  
/RANK=RANK1 TO RANK22.
```

- La variable *RANK1* contient le numéro de classement du profil 1 sur les 22 profils du plan orthogonal. La variable *RANK22* contient le numéro de classement du profil 22.

Sous-commande SCORE

La sous-commande SCORE indique que chaque point de données correspond à un score de préférence attribué à chacun des profils, commençant par le score du profil 1, puis celui du profil 2, et ainsi de suite. Ce type de données peut être généré, par exemple, en demandant aux sujets d'attribuer un nombre compris entre 1 et 100 au profil, indiquant leur appréciation de ce dernier. Plus le score est important, plus grande est la préférence.

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'  
/SCORE=SCORE1 TO SCORE22.
```

- La variable *SCORE1* contient le score du profil 1 et la variable *SCORE22* contient le score du profil 22.

Sous-commandes facultatives

La commande CONJOINT offre un certain nombre de sous-commandes facultatives surpassant le contrôle et les fonctionnalités requises.

Sous-commande SUBJECT

La sous-commande SUBJECT vous permet de spécifier une variable du fichier du données à utiliser en tant qu'identificateur de sujet. Si vous ne spécifiez aucune variable de sujet, la commande CONJOINT assume que l'ensemble des observations d'un même fichier de données proviennent d'un même sujet. L'exemple suivant spécifie que la variable *ID* provenant du fichier *rugranks.sav* doit être utilisée en tant qu'identificateur du sujet.

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'  
/SCORE=SCORE1 TO SCORE22 /SUBJECT=ID.
```

Sous-commande FACTORS

La sous-commande FACTORS vous permet de spécifier le modèle décrivant la relation attendue entre les facteurs et les classements ou scores. Si vous ne spécifiez aucun modèle de facteur, la commande CONJOINT suppose un modèle discret. Vous pouvez spécifier l'un des quatre modèles suivants :

DISCRETE. Le modèle DISCRETE indique que les niveaux de facteurs sont catégoriques et qu'aucune hypothèse n'est faite à propos des relations entre le facteur et les scores ou classements. Il s'agit de la valeur par défaut.

LINEAR. Le modèle LINEAR indique une relation linéaire attendue entre le facteur et les scores ou classements. Vous pouvez spécifier la direction attendue de la relation linéaire avec les mots clés MORE et LESS. MORE indique qu'on s'attend à ce que des niveaux plus élevés d'un facteur soient préférés, alors que LESS indique qu'on s'attend à ce que des niveaux plus bas d'un facteur soient préférés. La spécification des mots clés MORE ou LESS n'affecte en rien l'estimation des outils. Ces mots-clés sont utilisés simplement pour identifier les sujets dont les estimations ne correspondent pas à la direction attendue.

IDEAL. Le modèle IDEAL indique une relation quadratique attendue entre les scores ou classements et le facteur. Il est supposé qu'il existe un niveau idéal pour le facteur et plus la distance vers ce point idéal est grande (dans n'importe quelle direction) moins la préférence est élevée. Les facteurs décrits dans ce modèle doivent présenter au moins trois niveaux.

ANTIIDEAL. Le modèle ANTIIDEAL indique une relation quadratique attendue entre les scores ou classements et le facteur. Il est supposé qu'il existe un niveau idéal pour le facteur et plus la distance vers ce point idéal est grande (dans n'importe quelle direction) plus la préférence est élevée. Les facteurs décrits dans ce modèle doivent présenter au moins trois niveaux.

La syntaxe de commande suivante fournit un exemple utilisant la sous-commande FACTORS :

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'
/RANK=RANK1 TO RANK22 /SUBJECT=ID
/FACTORS=PACKAGE BRAND (DISCRETE) PRICE (LINEAR LESS)
SEAL (LINEAR MORE) MONEY (LINEAR MORE).
```

- Notez que le *conditionnement* et la *marque* sont tous deux modélisés en tant que valeurs discrètes.

Sous-commande PRINT

La sous-commande PRINT vous permet de contrôler le contenu du tableau de résultats. Par exemple, si vous disposez d'un grand nombre de sujets, vous pouvez choisir de limiter le résultat au résumé des résultats, en omettant les détails pour chaque sujet, comme illustré dans l'exemple suivant :

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'
/RANK=RANK1 TO RANK22 /SUBJECT=ID
/PRINT=SUMMARYONLY.
```

Vous pouvez également choisir d'inclure l'analyse des données expérimentales, les résultats des observations de simulation dans le fichier de plan, les deux à la fois ou aucun des deux dans le résultat. Les observations de simulation ne sont pas notées par les sujets mais représentent

des profils de produits intéressants. La procédure conjointe utilise l'analyse des données expérimentales pour effectuer des prévisions sur la préférence relative correspondant à chaque profil de simulation. Dans l'exemple qui suit, le résultat détaillé de chaque sujet est supprimé et le résultat est limité aux résultats des simulations :

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'  
/RANK=RANK1 TO RANK22 /SUBJECT=ID  
/PRINT=SIMULATION SUMMARYONLY.
```

Sous-commande PLOT

La sous-commande PLOT contrôle l'insertion des diagrammes dans le résultat. De la même manière que pour le tableau de résultats (sous-commande PRINT), vous pouvez décider de limiter le résultat au résumé des résultats ou d'inclure les résultats de chaque sujet. Par défaut, aucun diagramme n'est généré. Dans l'exemple qui suit, le résultat inclut les diagrammes disponibles :

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'  
/RANK=RANK1 TO RANK22 /SUBJECT=ID  
/PLOT=ALL.
```

Sous-commande UTILITY

La sous-commande UTILITY écrit un fichier de données au format IBM® SPSS® Statistics contenant des informations détaillées pour chaque sujet. Elle inclut les outils des facteurs DISCRETE, la pente et les fonctions quadratiques des facteurs LINEAR, IDEAL et ANTIIDEAL, la constante de régression et les estimations des scores de préférence. Ces valeurs peuvent ensuite être utilisées dans d'autres analyses ou pour la création d'autres diagrammes dans d'autres procédures. L'exemple qui suit crée un fichier utilitaire portant le nom *rugutil.sav* :

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'  
/RANK=RANK1 TO RANK22 /SUBJECT=ID  
/UTILITY='RUGUTIL.SAV'.
```

Partie II: Exemples

Utilisation de l'analyse conjointe pour le modèle Préférence de la shampoineuse

Dans un exemple courant d'analyse conjointe (Green et Wind, 1973), une société intéressée par la commercialisation d'une nouvelle shampoineuse souhaite examiner l'influence de cinq critères sur la préférence du consommateur : la conception du conditionnement, la marque, le prix, une étiquette *Economique* et une garantie « satisfait ou remboursé ». Il existe trois niveaux de critère pour la conception du conditionnement, suivant l'emplacement de l'applicateur, trois marques (*K2R*, *Glory* et *Bissell*), trois niveaux de prix et deux niveaux (non ou oui) pour chacun des deux derniers critères. Le tableau suivant indique les variables utilisées dans l'étude sur la shampoineuse, ainsi que les valeurs et étiquettes correspondantes.

Table 5-1
Variables dans l'étude sur la shampoineuse

| Nom de variable | l'étiquette Variable | Etiquette de valeur |
|------------------------|---------------------------------|------------------------|
| <i>conditionnement</i> | conception du conditionnement | A*, B*, C* |
| <i>marque</i> | nom de marque | K2R, Glory, Bissell |
| <i>prix</i> | prix | \$1.19, \$1.39, \$1.59 |
| <i>étiquette</i> | Etiquette Economique | non, oui |
| <i>argent</i> | garantie Satisfait ou remboursé | non, oui |

Il peut y avoir d'autres facteurs et niveaux de facteurs pour caractériser les shampoineuses, mais ceux-ci sont les seuls qui intéressent la direction. Il s'agit d'un point important dans les analyses conjointes. Vous voulez choisir uniquement les facteurs (variables indépendantes) dont vous pensez qu'ils influencent le plus les préférences du sujet (la variable dépendante). A l'aide de l'analyse conjointe, vous allez développer un modèle des préférences clients en fonction de ces cinq facteurs.

Cet exemple utilise les informations des fichiers de données suivants : *carpet_prefs.sav* contient les données collectées dans les sujets, *carpet_plan.sav* contient les profils des produits ayant été étudiés et *conjoint.sps* contient la syntaxe de commande nécessaire à l'exécution de l'analyse. [Pour plus d'informations, reportez-vous à la section Fichiers d'exemple dans l'annexe A dans IBM SPSS Conjoint 19.](#)

Génération d'un plan orthogonal

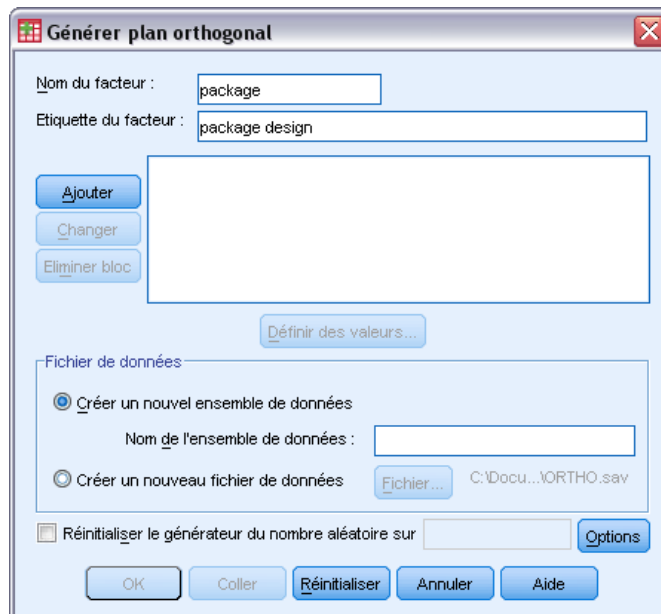
La première étape d'une analyse conjointe est de créer les combinaisons de niveaux de facteurs qui sont présentés aux sujets comme les profils des produits. Puisque même un petit nombre de facteurs et quelques niveaux pour chaque facteur génèrent un nombre ingérable de profils de produits potentiels, vous devez générer un sous-ensemble représentatif connu sous le nom de **tableau orthogonal**.

La procédure de génération de plan orthogonal crée un tableau orthogonal—également nommé **plan orthogonal**—et stocke les informations dans un fichier de données. Contrairement aux procédures, un fichier de travail n'est pas requis avant l'exécution de la procédure de génération de plan orthogonal. Si vous n'avez pas de fichier de travail, vous pouvez en créer un en générant des noms de variables, des étiquettes de variables et des étiquettes de valeurs à partir des options que vous sélectionnez dans les boîtes de dialogue. Si vous avez déjà un fichier de travail, vous pouvez le remplacer ou enregistrer le plan orthogonal comme fichier de données distinct.

Pour générer un plan orthogonal :

- ▶ A partir des menus, sélectionnez :
Données > Plan orthogonal > Générer...

Figure 5-1
Boîte de dialogue Générer plan orthogonal



- ▶ Entrez conditionnement dans la zone de texte Nom facteur et conception du conditionnement dans la zone de texte Etiquette facteur.
- ▶ Cliquez sur Ajouter.

Vous créez ainsi un élément avec l'étiquette *conditionnement 'conception du conditionnement'* (?). Sélectionnez cet élément.

- Cliquez sur Définir les valeurs.

Figure 5-2

Boîte de dialogue Générer plan orthogonal : Définir les valeurs

- Saisissez les valeurs 1, 2, et 3 pour représenter la conception du conditionnement A*, B*, et C*. Entrez également les étiquettes A*, B* et C*.
- Cliquez sur Poursuivre.

Vous allez maintenant répéter ce processus pour les facteurs restants *marque*, *prix*, *étiquette* et *argent*. Utilisez les valeurs et les étiquettes du tableau suivant, qui comprennent les valeurs que vous avez déjà entrées pour *conditionnement*.

| Nom du facteur | Etiquette du facteur | Valeurs | Etiquettes |
|-----------------|---------------------------------|------------------|------------------------|
| conditionnement | conception du conditionnement | 1, 2, 3 | A*, B*, C* |
| marque | nom de marque | 1, 2, 3 | K2R, Glory, Bissell |
| prix | prix | 1.19, 1.39, 1.59 | \$1.19, \$1.39, \$1.59 |
| étiquette | Etiquette Economique | 1, 2 | non, oui |
| argent | garantie Satisfait ou remboursé | 1, 2 | non, oui |

Une fois que vous avez terminé de spécifier les facteurs :

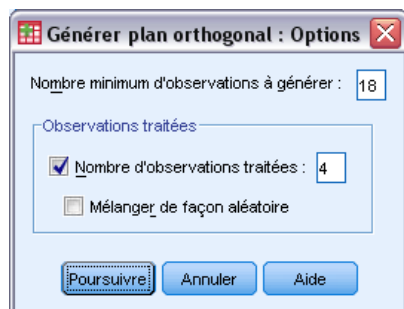
- Dans le groupe Fichiers de données, laissez la valeur par défaut de Créer un nouvel ensemble de données et entrez un nouveau nom d'ensemble de données. Le plan généré sera sauvegardé dans un nouvel ensemble de données dans la session en cours avec un nom spécifié.
- Sélectionnez Réinitialiser le générateur du nombre aléatoire sur et entrez la valeur 2000000.

La génération d'un plan orthogonal requiert un ensemble de nombres aléatoires. Pour dupliquer un plan, dans ce cas, le plan utilisé pour l'étude de cas, vous devez définir la valeur du générateur avant de générer votre plan et réinitialiser le générateur sur cette valeur à chaque génération de votre plan. Le plan utilisé pour cette étude de cas a été généré avec une valeur de générateur de 2000000.

- Cliquez sur Options.

Figure 5-3

Boîte de dialogue Générer plan orthogonal : Options



- Dans la zone de texte Nombre minimum d'observations à générer, tapez 18.

Par défaut, le nombre minimum d'observations nécessaires à un tableau orthogonal est généré. La procédure détermine le nombre d'observations devant être administrées pour permettre une estimation des outils. Vous pouvez également indiquer un nombre minimum d'observations à générer, comme vous venez de le faire. Vous souhaitez peut-être faire ceci car le nombre par défaut d'observations minimum est trop petit pour être utile ou parce que vous envisagez un plan expérimental qui requiert un certain nombre d'observations minimum.

- Sélectionnez Nombre d'observations traitées et tapez 4.

Les observations traitées sont évaluées par sujets, mais ne sont pas utilisées par l'analyse conjointe pour évaluer les outils. Elles sont utilisées pour vérifier la validité des outils estimés. Les observations traitées sont générées à partir d'un autre plan aléatoire et non à partir du plan expérimental orthogonal.

- Cliquez sur Continuer dans la boîte de dialogue Générer plan orthogonal : Options.
- Cliquez sur OK dans la boîte de dialogue Générer plan orthogonal.

Figure 5-4
Plan orthogonal pour l'exemple de la shampoineuse

| | package | brand | price | seal | money | STATUS_ | CARD |
|----|---------|---------|--------|------|-------|---------|------|
| 14 | A* | Glory | \$1,19 | oui | non | Design | 14 |
| 15 | B* | K2R | \$1,39 | oui | oui | Design | 15 |
| 16 | A* | K2R | \$1,19 | non | non | Design | 16 |
| 17 | A* | Bissell | \$1,59 | non | oui | Design | 17 |
| 18 | B* | Bissell | \$1,19 | non | non | Design | 18 |
| 19 | A* | Bissell | \$1,59 | oui | non | Holdout | 19 |
| 20 | C* | K2R | \$1,19 | oui | non | Holdout | 20 |
| 21 | A* | Glory | \$1,59 | non | non | Holdout | 21 |
| 22 | A* | Bissell | \$1,19 | non | non | Holdout | 22 |

Le plan orthogonal est affiché dans l'éditeur de données et pour optimiser son affichage, affichez les étiquettes de valeurs plutôt que les valeurs de données actuelles. Pour ce faire, choisissez *Étiquettes de valeurs* dans le menu *Afficher*.

Les variables dans le fichier de données sont les facteurs utilisés pour spécifier le plan. Chaque observation représente un profil de produit du plan. Notez que deux variables supplémentaires, *CARD_* et *STATUS_*, apparaissent dans le fichier de données. *CARD_* attribue un numéro séquentiel à chaque profil utilisé pour identifier le profil. *STATUS_* indique si un profil fait partie du plan expérimental (les 18 premières observations), une observation traitée (les 4 dernières observations) ou une observation de simulation (qui sera abordée dans une rubrique ultérieure de cette étude de cas).

Le plan orthogonal est une saisie requise pour l'analyse des données. Vous allez donc enregistrer votre plan dans un fichier de données. Dans un souci de commodité, le plan actuel a été enregistré dans *carpet_plan.sav* (les plans orthogonaux sont également appelés simplement des **plans**).

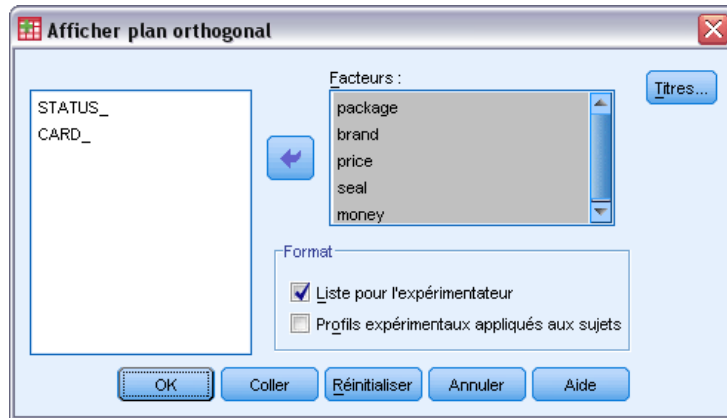
Création de stimuli expérimentaux : Affichage du plan

Une fois que vous avez créé un plan orthogonal, vous allez l'utiliser pour créer les profils des produits devant être évalués par les sujets. Vous pouvez obtenir une liste des profils dans un seul tableau ou afficher chaque profil dans un tableau distinct.

Pour afficher un plan orthogonal :

- A partir des menus, sélectionnez :
Données > Plan orthogonal > Afficher...

Figure 5-5
Boîte de dialogue Afficher plan orthogonal



- Sélectionnez *conditionnement*, *marque*, *prix*, *étiquette* et *argent* comme facteurs.

Les informations contenues dans les variables *STATUS_* et *CARD_* sont automatiquement incluses dans les résultats, il n'y a donc pas besoin de les sélectionner.

- Sélectionnez Liste pour l'expérimentateur dans le groupe Format. Cette opération affiche tout le plan orthogonal dans un simple tableau.
- Cliquez sur OK.

Figure 5-6
Affichage d'un plan orthogonal : Présentation en tableau simple

| | ID de carte | package design | brand name | price | Good Housekeeping seal | money-back guarantee |
|-----------------|-------------|----------------|------------|--------|------------------------|----------------------|
| 1 | 1 | A* | Glory | \$1,39 | oui | non |
| 2 | 2 | B* | K2R | \$1,19 | non | non |
| 3 | 3 | B* | Glory | \$1,39 | non | oui |
| 4 | 4 | C* | Glory | \$1,59 | non | non |
| 5 | 5 | C* | Bissell | \$1,39 | non | non |
| 6 | 6 | A* | Bissell | \$1,39 | non | non |
| 7 | 7 | B* | Bissell | \$1,59 | oui | non |
| 8 | 8 | A* | K2R | \$1,59 | non | oui |
| 9 | 9 | C* | K2R | \$1,39 | non | non |
| 10 | 10 | C* | Glory | \$1,19 | non | oui |
| 11 | 11 | C* | K2R | \$1,59 | oui | non |
| 12 | 12 | B* | Glory | \$1,59 | non | non |
| 13 | 13 | C* | Bissell | \$1,19 | oui | oui |
| 14 | 14 | A* | Glory | \$1,19 | oui | non |
| 15 | 15 | B* | K2R | \$1,39 | oui | oui |
| 16 | 16 | A* | K2R | \$1,19 | non | non |
| 17 | 17 | A* | Bissell | \$1,59 | non | oui |
| 18 | 18 | B* | Bissell | \$1,19 | non | non |
| 19 ^a | 19 | A* | Bissell | \$1,59 | oui | non |
| 20 ^a | 20 | C* | K2R | \$1,19 | oui | non |
| 21 ^a | 21 | A* | Glory | \$1,59 | non | non |
| 22 ^a | 22 | A* | Bissell | \$1,19 | non | non |

a. Élément exclu

Les résultats ressemblent à un plan orthogonal comme indiqué dans l'éditeur de données, une ligne pour chaque profil avec les facteurs comme colonnes. Notez cependant que les en-têtes de colonnes sont les étiquettes de variables et non les noms de variables que vous affichez dans Data editor. Notez également que les observations traitées sont identifiées par une note de bas de page. Ceci est intéressant pour les expérimentateurs, mais vous ne voulez pas que les sujets puissent connaître les observations qui ont été traitées.

En fonction de la manière dont vous créez et distribuez vos profils de produits finaux, vous souhaitez peut-être enregistrer ce tableau en tant que fichier HTML, Word/RTF, Excel ou PowerPoint. Ceci est facile à accomplir, il suffit de sélectionner le tableau dans l'Editeur de résultats, de cliquer sur le bouton droit de la souris et de sélectionner Exporter. De plus, si vous utilisez la version exportée pour créer les profils du produit final, assurez-vous de modifier les notes de bas de page pour les observations traitées.

La génération d'un tableau distinct pour chaque profil de produit est peut-être plus adaptée aux besoins de votre enquête. Ce choix permet de faciliter l'export vers PowerPoint car chaque tableau (profil de produit) est placé sur une diapositive PowerPoint distincte.

Pour afficher chaque profil dans un tableau distinct :

- ▶ Cliquez sur le bouton Rappeler boîte de dialogue et sélectionnez Afficher plan orthogonal.
- ▶ Désélectionnez Liste pour l'expérimentateur et sélectionnez Profils expérimentaux appliqués aux sujets.
- ▶ Cliquez sur OK.

Figure 5-7

Affichage d'un plan orthogonal : Présentation multitableau

Numéro de profil 1

| ID de carte | package design | brand name | price | Good Housekee ping seal | money-back guarantee |
|-------------|----------------|------------|--------|-------------------------|----------------------|
| 1 | A* | Glory | \$1,39 | oui | non |

Numéro de profil 2

| ID de carte | package design | brand name | price | Good Housekee ping seal | money-back guarantee |
|-------------|----------------|------------|--------|-------------------------|----------------------|
| 2 | B* | K2R | \$1,19 | non | non |

Les informations pour chaque profil de produit sont affichées dans un tableau distinct. En outre, il est impossible de distinguer les observations traitées du reste des observations, il n'y a donc aucun problème à supprimer des identificateurs d'observations traitées, à la différence d'une présentation en tableau simple.

Exécution de l'analyse

Vous avez généré un plan orthogonal et appris comment afficher les profils de produits associés. Vous pouvez maintenant apprendre à exécuter une analyse conjointe.

Figure 5-8
Données de préférence pour l'exemple de la shampoineuse

| | ID | PREF1 | PREF2 | PREF3 | PREF4 | PREF5 | PREF6 | PREF7 | PREF8 |
|----|----|-------|-------|-------|-------|-------|-------|-------|-------|
| 1 | 1 | 13 | 15 | 1 | 20 | 14 | 7 | 11 | 19 |
| 2 | 2 | 15 | 7 | 18 | 2 | 12 | 3 | 11 | 20 |
| 3 | 3 | 2 | 18 | 14 | 16 | 22 | 13 | 20 | 10 |
| 4 | 4 | 13 | 10 | 20 | 14 | 2 | 18 | 16 | 22 |
| 5 | 5 | 13 | 18 | 2 | 10 | 20 | 15 | 9 | 5 |
| 6 | 6 | 15 | 2 | 3 | 12 | 18 | 7 | 20 | 10 |
| 7 | 7 | 13 | 7 | 15 | 18 | 2 | 3 | 10 | 20 |
| 8 | 8 | 15 | 7 | 13 | 4 | 6 | 16 | 8 | 22 |
| 9 | 9 | 20 | 9 | 10 | 11 | 4 | 5 | 13 | 15 |
| 10 | 10 | 8 | 21 | 19 | 17 | 4 | 11 | 12 | 7 |

Les données de préférence fournies par les sujets sont stockées dans *carpet_prefs.sav*. Ces données se composent de réponses de 10 sujets, chacun étant identifié par une valeur unique de la variable *ID*. On a demandé aux sujets de classer les 22 profils de produits du préféré au moins intéressant. Les variables allant de *PREF1* à *PREF22* contiennent les ID des profils de produits associés, c'est-à-dire les ID de carte de *carpet_plan.sav*. Le sujet 1, par exemple, a préféré le profil 13 aux autres, donc *PREF1* prend la valeur 13.

L'analyse des données est une tâche qui requiert l'utilisation de syntaxe de commande, plus particulièrement la commande *CONJOINT*. La syntaxe de commande nécessaire a été fournie dans le fichier *conjoint.sps*.

```
CONJOINT PLAN='file specification'
  /DATA='file specification'
  /SEQUENCE=PREF1 TO PREF22
  /SUBJECT=ID
  /FACTORS=PACKAGE BRAND (DISCRETE)
  PRICE (LINEAR LESS)
  SEAL (LINEAR MORE) MONEY (LINEAR MORE)
  /PRINT=SUMMARYONLY.
```

- La sous-commande *PLAN* indique le fichier contenant le plan orthogonal ; dans cet exemple il s'agit de *carpet_plan.sav*.
- La sous-commande *DATA* indique le fichier contenant les données de préférence—dans cet exemple, il s'agit de *carpet_prefs.sav*. Si vous choisissez les données de préférence comme fichier de travail, vous pouvez remplacer la spécification de fichier par un astérisque (*), sans les parenthèses.
- La sous-commande *SEQUENCE* spécifie que chaque point de données des données de préférence est un numéro de profil, en commençant par le profil préféré et en terminant par le moins intéressant.
- La sous-commande *SUBJECT* spécifie que la variable *ID* identifie les sujets.
- La sous-commande *FACTORS* spécifie un modèle décrivant la relation attendue entre les données de préférence et les niveaux de facteur. Les facteurs spécifiés réfèrent à des variables définies dans le fichier de plan nommé dans la sous-commande *PLAN*.
- Le mot-clé *DISCRETE* est utilisé lorsque les niveaux de facteur sont qualitatifs et qu'aucune hypothèse n'est faite sur la relation entre les niveaux et les données. C'est le cas pour les facteurs *conditionnement* et *marque* qui représentent la conception du conditionnement et le nom de marque respectivement. *DISCRETE* est considéré si un facteur n'est pas étiqueté

de l'une des quatre alternatives (DISCRETE, LINEAR, IDEAL, ANTIIDEAL) ou n'est pas inclus dans la sous-commande FACTORS.

- Le mot-clé LINEAR, utilisé pour les facteurs restants, indique que les données doivent être liées linéairement au facteur. Par exemple, la préférence doit habituellement être liée linéairement au prix. Vous pouvez également spécifier les modèles quadratiques (qui ne sont pas utilisés dans cet exemple) avec les mots-clés IDEAL et ANTIIDEAL.
- Les mots-clés MORE et LESS suivant LINEAR indique une direction attendue de cette relation. Puisque nous attendons que les préférences soient plus élevées pour des prix plus bas, le mot-clé LESS est utilisé pour *prix*. Cependant, nous attendons des préférences élevées pour une étiquette d'approbation *Economique* ou une garantie Satisfait ou remboursé, ainsi le mot-clé MORE est utilisé pour *étiquette* et *argent* (souvenez-vous que les niveaux de ces deux facteurs étaient définis à 1 pour *non* et 2 pour *oui*).

Spécifier MORE ou LESS ne change pas les signes des coefficients ou n'affecte pas les estimations des outils. Ces mots-clés sont utilisés simplement pour identifier les sujets dont les estimations ne correspondent pas à la direction attendue. De même, le choix IDEAL à la place d'ANTIIDEAL, et vice versa, n'affecte pas les coefficients ou outils.

- La sous-commande PRINT spécifie que les résultats contiennent des informations pour le groupe de sujets uniquement en tant qu'ensemble (mot-clé SUMMARYONLY). Les informations pour chaque sujet séparément sont supprimées.

Essayez d'exécuter cette syntaxe de commande. Vérifiez que vous avez inclus des chemins d'accès valides aux fichiers *carpet_prefs.sav* et *carpet_plan.sav*. Pour obtenir une description complète de toutes les options, reportez-vous à la commande CONJOINT dans *Command Syntax Reference*.

Scores des outils

Figure 5-9
Scores des outils

| | | Estimation d'utilité | Std. Erreur |
|-------------|---------|-------------------------|-------------|
| package | A* | -2,233 | ,192 |
| | B* | 1,867 | ,192 |
| | C* | ,367 | ,192 |
| brand | K2R | ,367 | ,192 |
| | Glory | -,350 | ,192 |
| | Bissell | -,017 | ,192 |
| price | 1 | -7,703 | 1,154 |
| | 2 | -8,811 | 1,320 |
| seal | Non | 2,000 | ,287 |
| | Oui | 4,000 | ,575 |
| money | Non | 1,250 | ,287 |
| | Oui | 2,500 | ,575 |
| (Constante) | | 12,870 | 1,282 |

Ce tableau montre les scores des outils (valeur de partie) ainsi que leurs erreurs standard pour chaque niveau de facteur. Les valeurs d'outils supérieures indiquent une préférence élevée. Conformément à ce qui était attendu, il existe une relation inverse entre le prix et l'utilitaire, avec des prix supérieurs pour un utilitaire inférieur (des valeurs négatives importantes indiquent un

utilitaire inférieur). La présence d'une étiquette d'approbation ou d'une garantie Satisfait ou remboursé correspond à un utilitaire supérieur, comme prévu.

Puisque tous les outils sont exprimés dans une même unité, ils peuvent être additionnés pour donner l'**utilitaire total** d'une combinaison. Par exemple, l'utilitaire total d'une shampoineuse avec la conception du conditionnement B^* , la marque $K2R$, le prix $0,95$ € et aucune étiquette d'approbation ou de garantie Satisfait ou remboursé est :

$$\begin{aligned} & \text{utilitaire}(\text{conditionnement } B^*) + \text{utilitaire}(K2R) + \text{utilitaire}(0,95 \text{ €}) \\ & + \text{utilitaire}(\text{aucune étiquette}) + \text{utilitaire}(\text{aucune garantie Satisfait ou remboursé}) + \text{constante} \end{aligned}$$

ou

$$1.867 + 0.367 + (-6.595) + 2.000 + 1.250 + 12.870 = 11.759$$

Si la shampoineuse avait une conception du conditionnement C^* , une marque $Bissell$, un prix $0,95$ €, une étiquette d'approbation et une garantie Satisfait ou remboursé, l'utilitaire total serait :

$$0.367 + (-0.017) + (-8.811) + 4.000 + 2.500 + 12.870 = 10.909$$

Coefficients

Figure 5-10
Coefficients

| | Coefficient |
|-------|-------------|
| | Estimation |
| price | -5,542 |
| seal | 2,000 |
| money | 1,250 |

Ce tableau affiche les coefficients de régression linéaire pour les facteurs spécifiés en tant que LINEAR (pour les modèles IDEAL et ANTIIDEAL, il y aura également un terme quadratique). L'utilitaire pour un niveau de facteur particulier est déterminé en multipliant le niveau par le coefficient. Par exemple, l'utilitaire estimé pour un prix de $0,95$ € a été répertorié en tant que $-6,595$ dans le tableau des outils. Il s'agit simplement de la valeur du niveau de prix, $0,95$ multiplié par le coefficient du prix, $-5,542$.

Importance relative

La plage des valeurs d'utilitaire (plus élevé à moins élevé) pour chaque facteur fournit une mesure de l'importance du facteur dans la préférence générale. Les facteurs avec des plages d'outils supérieures tiennent un rôle plus important que ceux avec des plages plus petites.

Figure 5-11
Valeurs d'importance

| | |
|---------|--------|
| package | 35,635 |
| brand | 14,911 |
| price | 29,410 |
| seal | 11,172 |
| money | 8,872 |

Score d'importance moyen

Ce tableau fournit une mesure de l'importance relative de chaque facteur, appelée score ou valeur d'**importance**. Les valeurs sont calculées en prenant la plage d'utilitaire de chaque facteur séparément et en la divisant par la somme des plages d'utilitaires pour tous les facteurs. Les valeurs représentent donc des pourcentages et ont la particularité de produire une somme de 100 lorsqu'elles sont toutes additionnées. Ces calculs, il faut le noter, sont effectués séparément pour chaque sujet et les résultats sont ensuite considérés par leur moyenne pour tous les sujets.

Notez que les outils généraux ou récapitulatifs et les coefficients de régression des plans orthogonaux sont les mêmes avec ou sans une sous-commande SUBJECT, les importances sont en général différentes. Pour des résultats récapitulatifs sans la sous-commande SUBJECT, les importances peuvent être calculées directement à partir des outils récapitulatifs, de la même manière que pour des sujets individuels. Cependant, lorsqu'une sous-commande SUBJECT est utilisée, les importances pour les sujets individuels sont considérées par leur moyenne et les moyennes des importances ne correspondent pas en général à celles calculées à l'aide des outils récapitulatifs.

Les résultats indiquent que la conception du conditionnement est celle qui, parmi toutes les préférences, a le plus d'influence. Ceci signifie qu'il y a une différence importante de préférence entre les profils de produits avec le conditionnement le plus attractif et ceux avec le conditionnement le moins attractif. Les résultats indiquent également que la garantie Satisfait ou remboursé tient le rôle le moins important dans la détermination de la préférence générale. Le prix joue un rôle important mais pas autant que la conception du conditionnement. Ceci est peut-être dû au fait que la plage des prix n'est pas assez étendue.

Corrélations

Figure 5-12
Coefficients de corrélation

| | Valeur | Sig. |
|---|--------|------|
| r de Pearson | ,982 | ,000 |
| Tau de Kendall | ,892 | ,000 |
| Tau de Kendall pour les éléments exclus | ,667 | ,087 |

Ce tableau affiche deux statistiques, R de Pearson et tau de Kendall, qui fournissent des mesures de la corrélation entre les préférences observées et celles estimées.

Le tableau affiche également le tau de Kendall uniquement pour les profils traités. Souvenez-vous que les profils traités (quatre dans l'exemple précédent) ont été évalués par sujets mais ne sont pas utilisés par la procédure Conjointe pour les outils estimés. En fait, la procédure

conjointe calcule les corrélations entre les rangs observés et ceux estimés pour ces profils pour vérifier la validité des outils.

Dans de nombreuses analyses conjointes, le nombre de paramètres est proche du nombre de profils évalués, ce qui augmente de manière artificielle la corrélation entre les résultats observés et ceux estimés. Dans ces observations, les corrélations des profils traités peuvent donner une meilleure indication de l'ajustement d'un modèle. Gardez à l'esprit cependant que les éléments traités produisent des coefficients de corrélation plus bas.

Inversements

En spécifiant des modèles `LINEAR` pour *prix*, *étiquette* et *argent*, nous choisissons une direction attendue (`LESS` ou `MORE`) pour la relation linéaire entre la valeur de la variable la préférence pour cette valeur. La procédure conjointe garde une trace du nombre de sujets dont les préférences ont indiqué le contraire de la relation attendue—par exemple, une préférence supérieure pour des prix élevés ou une préférence inférieure pour une garantie « satisfait ou remboursé ». Ces observations sont appelées des **inversements**.

Figure 5-13

Nombre d'inversements par facteur et sujet

| | | | |
|---------|---------|----------|---|
| Facteur | price | | 3 |
| | money | | 2 |
| | seal | | 2 |
| | brand | | 0 |
| | package | | 0 |
| Sujet | 1 | Sujet 1 | 1 |
| | 2 | Sujet 2 | 2 |
| | 3 | Sujet 3 | 0 |
| | 4 | Sujet 4 | 0 |
| | 5 | Sujet 5 | 0 |
| | 6 | Sujet 6 | 1 |
| | 7 | Sujet 7 | 0 |
| | 8 | Sujet 8 | 0 |
| | 9 | Sujet 9 | 1 |
| | 10 | Sujet 10 | 2 |

Ce tableau affiche le nombre d'inversements pour chaque facteur et pour chaque sujet. Par exemple, trois sujets indiquent un inversement pour *prix*. En fait, ils préfèrent les profils de produits avec des prix élevés.

Exécution de simulations

Le véritable avantage de l'analyse conjointe est de pouvoir prévoir la préférence des profils de produits qui n'ont pas été évalués par les sujets. Ces observations sont appelées des **observations de simulation**. Les observations de simulation sont incluses dans le plan avec les profils du plan orthogonal et les profils traités.

La manière la plus simple d'entrer des observations de simulation est d'utiliser Data Editor, en utilisant les étiquettes de valeurs créées lorsque vous générez le plan expérimental.

Pour entrer une observation de simulation dans le fichier de plan :

- ▶ Dans une nouvelle ligne de la fenêtre de l'éditeur de données, sélectionnez une cellule puis la valeur souhaitée dans la liste (les étiquettes de valeurs peuvent être affichées en sélectionnant *Étiquettes de valeurs* dans le menu *Afficher*). Répétez cette opération pour toutes les variables (facteurs).
- ▶ Sélectionnez *Simulation* pour la valeur de la variable *STATUS_*.
- ▶ Entrez une valeur entière à utiliser comme identificateur de la valeur *CARD_*. Les observations de simulation devraient être numérotées séparément des autres observations.

Figure 5-14

Données de la shampoineuse incluant des observations de simulation

| package | brand | price | seal | money | STATUS_ | CARD_ |
|---------|---------|--------|------|-------|------------|-------|
| A* | Bissell | \$1.59 | Non | Oui | Design | 17 |
| B* | Bissell | \$1.19 | Non | Non | Design | 18 |
| A* | Bissell | \$1.59 | Oui | Non | Holdout | 19 |
| C* | K2R | \$1.19 | Oui | Non | Holdout | 20 |
| A* | Glory | \$1.59 | Non | Non | Holdout | 21 |
| A* | Bissell | \$1.19 | Non | Non | Holdout | 22 |
| C* | K2R | \$1.19 | Non | Non | Simulation | 1 |
| B* | Glory | \$1.19 | Oui | Oui | Simulation | 2 |

Ce schéma illustre une partie du fichier de plan pour l'étude sur la shampoineuse avec deux observations de simulation ajoutées. Dans un souci de commodité, ces dernières ont été incluses dans *carpet_plan.sav*.

L'analyse des observations de simulation est accomplie avec la même syntaxe de commande que celle utilisée précédemment, c'est-à-dire la syntaxe du fichier *conjoint.sps*. En fait, si vous avez exécuté la syntaxe décrite précédemment, vous avez remarqué que les résultats incluent également les résultats pour les observations de simulation car elles sont incluses dans *carpet_plan.sav*.

Vous pouvez choisir d'exécuter les simulations avec votre analyse initiale, comme c'est le cas ici, ou d'exécuter les simulations plus tard simplement en incluant des observations de simulation dans votre fichier de plan et d'exécuter à nouveau *CONJOINT*. Pour obtenir plus d'informations, reportez-vous à la commande *CONJOINT* dans *Command Syntax Reference*.

Probabilités des préférences des simulations

Figure 5-15

Résultats de simulation

| Numéro de carte | ID | Nombre maximal d'utilités ^a | Bradley-Terry-Luce | Logit |
|-----------------|----|--|--------------------|-------|
| 1 | 1 | 30,0% | 43,1% | 30,9% |
| 2 | 2 | 70,0% | 56,9% | 69,1% |

Ce tableau donne les probabilités prévues de choisir chacune des observations de simulation comme étant l'observation préférée, en fonction de trois modèles différents de probabilité de choix. Le **modèle utilitaire maximum** détermine la probabilité en divisant le nombre de répondants dont on prévoit qu'ils choisiront le profil par le nombre total de répondants. Pour chaque répondant, le choix prévu est simplement le profil avec le plus grand utilitaire total. Le

modèle BTL (Bradley-Terry-Luce) détermine la probabilité comme le rapport de l'utilitaire du profil par rapport à celui de tous les profils de simulation, considérée comme moyenne entre tous les répondants. Le **modèle logit** est semblable à BTL mais utilise le log népérien des outils à la place des outils. Parmi les 10 sujets de cette étude, les trois modèles ont indiqué que le profil de simulation 2 était préféré.

Fichiers d'exemple

Les fichiers d'exemple installés avec le produit figurent dans le sous-répertoire *Echantillons* du répertoire d'installation. Il existe un dossier distinct au sein du sous-répertoire *Echantillons* pour chacune des langues suivantes : Anglais, Français, Allemand, Italien, Japonais, Coréen, Polonais, Russe, Chinois simplifié, Espagnol et Chinois traditionnel.

Seuls quelques fichiers d'exemples sont disponibles dans toutes les langues. Si un fichier d'exemple n'est pas disponible dans une langue, le dossier de langue contient la version anglaise du fichier d'exemple.

Descriptions

Voici de brèves descriptions des fichiers d'exemple utilisés dans divers exemples à travers la documentation.

- **accidents.sav.** Ce fichier de données d'hypothèse concerne une société d'assurance qui étudie les facteurs de risque liés à l'âge et au sexe dans les accidents de la route survenant dans une région donnée. Chaque observation correspond à une classification croisée de la catégorie d'âge et du sexe.
- **adl.sav.** Ce fichier de données d'hypothèse concerne les mesures entreprises pour identifier les avantages d'un type de thérapie proposé aux patients qui ont subi une attaque cardiaque. Les médecins ont assigné de manière aléatoire les patients du sexe féminin ayant subi une attaque cardiaque à un groupe parmi deux groupes possibles. Le premier groupe a fait l'objet de la thérapie standard tandis que le second a bénéficié en plus d'une thérapie émotionnelle. Trois mois après les traitements, les capacités de chaque patient à effectuer les tâches ordinaires de la vie quotidienne ont été notées en tant que variables ordinales.
- **advert.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend un détaillant pour examiner la relation existant entre l'argent dépensé dans la publicité et les ventes résultantes. Pour ce faire, il collecte les chiffres des ventes passées et les coûts associés à la publicité.
- **aflatoxin.sav.** Ce fichier de données d'hypothèse concerne le test de l'aflatoxine dans des récoltes de maïs. La concentration de ce poison varie largement d'une récolte à l'autre et au sein de chaque récolte. Un processeur de grain a reçu 16 échantillons issus de 8 récoltes de maïs et a mesuré les niveaux d'aflatoxine en parties par milliard (PPB).
- **aflatoxin20.sav.** Ce fichier de données contient les mesures d'aflatoxine de chacun des 16 échantillons des récoltes 4 et 8 du fichier de données *aflatoxin.sav*.
- **anorectic.sav.** En cherchant à développer une symptomatologie standardisée du comportement anorexique/boulimique, des chercheurs (Van der Ham, Meulman, Van Strien, et Van Engeland, 1997) ont examiné 55 adolescents souffrant de troubles alimentaires. Chaque patient a été

observé quatre fois sur une période de quatre années, soit un total de 220 observations. A chaque observation, les patients ont été notés pour chacun des 16 symptômes. En raison de l'absence de scores de symptôme pour le patient 71/visite 2, le patient 76/visite 2 et le patient 47/visite 3, le nombre d'observations valides est de 217.

- **autoaccidents.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend un analyste en assurances pour modéliser le nombre d'accidents de la route par conducteur tout en prenant en compte l'âge et le sexe du conducteur. Chaque observation représente un conducteur distinct et enregistre son sexe, son âge et le nombre d'accidents de la route au cours des cinq dernières années.
- **band.sav.** Ce fichier de données contient les chiffres de ventes hebdomadaires hypothétiques de CD musicaux d'un groupe. Les données relatives à trois variables explicatives possibles sont également incluses.
- **bankloan.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend une banque pour réduire le taux de défaut de paiement. Il contient des informations financières et démographiques sur 850 clients existants et éventuels. Les premières 700 observations concernent des clients auxquels des prêts ont été octroyés. Les 150 dernières observations correspondent aux clients éventuels que la banque doit classer comme bons ou mauvais risques de crédit.
- **bankloan_binning.sav.** Ce fichier de données d'hypothèse concerne des informations financières et démographiques sur 5 000 clients existants.
- **behavior.sav.** Dans un exemple classique (Price et Bouffard, 1974), on a demandé à 52 étudiants de noter les combinaisons établies à partir de 15 situations et de 15 comportements sur une échelle de 0 à 9, où 0 = « extrêmement approprié » et 9 = « extrêmement inapproprié ». En effectuant la moyenne des résultats de l'ensemble des individus, on constate une certaine différence entre les valeurs.
- **behavior_ini.sav.** Ce fichier de données contient la configuration initiale d'une solution bidimensionnelle pour *behavior.sav*.
- **brakes.sav.** Ce fichier de données d'hypothèse concerne le contrôle qualité effectué dans une usine qui fabrique des freins à disque pour des voitures haut de gamme. Le fichier de données contient les mesures de diamètre de 16 disques de 8 machines de production. Le diamètre cible des freins est de 322 millimètres.
- **breakfast.sav.** Au cours d'une étude classique (Green et Rao, 1972), on a demandé à 21 étudiants en MBA (Master of Business Administration) de l'école de Wharton et à leurs conjoints de classer 15 aliments du petit-déjeuner selon leurs préférences, de 1= « aliment préféré » à 15= « aliment le moins apprécié ». Leurs préférences ont été enregistrées dans six scénarios différents, allant de « Préférence générale » à « En-cas avec boisson uniquement ».
- **breakfast-overall.sav.** Ce fichier de données contient les préférences de petit-déjeuner du premier scénario uniquement, « Préférence générale ».
- **broadband_1.sav.** Ce fichier de données d'hypothèse concerne le nombre d'abonnés, par région, à un service haut débit. Le fichier de données contient le nombre d'abonnés mensuels de 85 régions sur une période de quatre ans.
- **broadband_2.sav.** Ce fichier de données est identique au fichier *broadband_1.sav* mais contient les données relatives à trois mois supplémentaires.

- **car_insurance_claims.sav.** Il s'agit d'un ensemble de données présenté et analysé ailleurs (McCullagh et Nelder, 1989) qui concerne des actions en indemnisation pour des voitures. Le montant d'action en indemnisation moyen peut être modélisé comme présentant une distribution gamma, à l'aide d'une fonction de lien inverse pour associer la moyenne de la variable dépendante à une combinaison linéaire de l'âge de l'assuré, du type de véhicule et de l'âge du véhicule. Le nombre d'actions entreprises peut être utilisé comme pondération de positionnement.
- **car_sales.sav.** Ce fichier de données contient des estimations de ventes hypothétiques, des barèmes de prix et des spécifications physiques concernant divers modèles et marques de véhicule. Les barèmes de prix et les spécifications physiques proviennent tour à tour de *edmunds.com* et des sites des constructeurs.
- **car_sales_uprepared.sav.** Il s'agit d'une version modifiée de *car_sales.sav* qui n'inclut aucune version transformée des champs.
- **carpet.sav.** Dans un exemple courant (Green et Wind, 1973), une société intéressée par la commercialisation d'un nouveau nettoyeur de tapis souhaite examiner l'influence de cinq critères sur la préférence du consommateur : la conception du conditionnement, la marque, le prix, une étiquette *Economique* et une garantie satisfait ou remboursé. Il existe trois niveaux de critère pour la conception du conditionnement, suivant l'emplacement de l'appliqueur, trois marques (*K2R*, *Glory* et *Bissell*), trois niveaux de prix et deux niveaux (non ou oui) pour chacun des deux derniers critères. Dix consommateurs classent 22 profils définis par ces critères. La variable *Préférence* indique le classement des rangs moyens de chaque profil. Un rang faible correspond à une préférence élevée. Cette variable reflète une mesure globale de préférence pour chaque profil.
- **carpet_prefs.sav.** Ce fichier de données repose sur le même exemple que celui décrit pour *carpet.sav*, mais contient les classements réels issus de chacun des 10 clients. On a demandé aux consommateurs de classer les 22 profils de produits, du préféré au moins intéressant. Les variables *PREF1* à *PREF22* contiennent les identificateurs des profils associés, tels qu'ils sont définis dans *carpet_plan.sav*.
- **catalog.sav.** Ce fichier de données contient des chiffres de ventes mensuelles hypothétiques relatifs à trois produits vendus par une entreprise de vente par correspondance. Les données relatives à cinq variables explicatives possibles sont également incluses.
- **catalog_seasfac.sav.** Ce fichier de données est identique à *catalog.sav* mais contient en plus un ensemble de facteurs saisonniers calculés à partir de la procédure de désaisonnalisation, ainsi que les variables de date correspondantes.
- **cellular.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend un opérateur téléphonique pour réduire les taux de désabonnement. Des scores de propension au désabonnement sont attribués aux comptes, de 0 à 100. Les comptes ayant une note égale ou supérieure à 50 sont susceptibles de changer de fournisseur.
- **ceramics.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend un fabricant pour déterminer si un nouvel alliage haute qualité résiste mieux à la chaleur qu'un alliage standard. Chaque observation représente un test séparé de l'un des deux alliages ; le degré de chaleur auquel l'alliage ne résiste pas est enregistré.

- **cereal.sav.** Ce fichier de données d'hypothèse concerne un sondage de 880 personnes interrogées sur leurs préférences de petit-déjeuner et sur leur âge, leur sexe, leur situation familiale et leur mode de vie (actif ou non actif, selon qu'elles pratiquent une activité physique au moins deux fois par semaine). Chaque observation correspond à un répondant distinct.
- **clothing_defects.sav.** Ce fichier de données d'hypothèse concerne le processus de contrôle qualité observé dans une usine de textile. Dans chaque lot produit à l'usine, les inspecteurs prélèvent un échantillon de vêtements et comptent le nombre de vêtements qui ne sont pas acceptables.
- **coffee.sav.** Ce fichier de données concerne l'image perçue de six marques de café frappé (Kennedy, Riquier, et Sharp, 1996). Pour chacun des 23 attributs d'image de café frappé, les personnes sollicitées ont sélectionné toutes les marques décrites par l'attribut. Les six marques sont appelées AA, BB, CC, DD, EE et FF à des fins de confidentialité.
- **contacts.sav.** Ce fichier de données d'hypothèse concerne les listes de contacts d'un groupe de représentants en informatique d'entreprise. Chaque contact est classé selon le service de l'entreprise où il travaille et le classement de son entreprise. Sont également enregistrés le montant de la dernière vente effectuée, le temps passé depuis la dernière vente et la taille de l'entreprise du contact.
- **creditpromo.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprind un grand magasin pour évaluer l'efficacité d'une promotion récente de carte de crédit. A cette fin, 500 détenteurs de carte ont été sélectionnés au hasard. La moitié a reçu une publicité faisant la promotion d'un taux d'intérêt réduit sur les achats effectués dans les trois mois à venir. L'autre moitié a reçu une publicité saisonnière standard.
- **customer_dbase.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprind une société pour utiliser les informations figurant dans sa banque de données et proposer des offres spéciales aux clients susceptibles d'être intéressés. Un sous-groupe de la base de clients a été sélectionné au hasard et a reçu des offres spéciales. Les réponses des clients ont été enregistrées.
- **customer_information.sav.** Un fichier de données d'hypothèse qui contient les informations postales du client, telles que le nom et l'adresse.
- **customer_subset.sav.** Un sous-ensemble de 80 observations de *customer_dbase.sav*.
- **customers_model.sav.** Ce fichier de données d'hypothèse concerne les personnes ciblées par une campagne de marketing. Ces données incluent des informations démographiques, un récapitulatif de l'historique d'achat et indiquent si chaque personne a répondu ou non à la campagne. Chaque observation représente une personne distincte.
- **customers_new.sav.** Ce fichier de données d'hypothèse concerne les personnes constituant des cibles potentielles pour une campagne de marketing. Ces données incluent des informations démographiques et un récapitulatif de l'historique d'achat pour chaque personne. Chaque observation représente une personne distincte.
- **debate.sav.** Ce fichier de données d'hypothèse concerne des réponses appariées à une enquête donnée aux participants à un débat politique avant et après le débat. Chaque observation représente un répondant distinct.
- **debate_aggregate.sav.** Il s'agit d'un fichier de données d'hypothèse qui rassemble les réponses dans le fichier *debate.sav*. Chaque observation correspond à une classification croisée de préférence avant et après le débat.

- **demo.sav.** Ce fichier de données d'hypothèse concerne une base de données clients achetée en vue de diffuser des offres mensuelles. Les données indiquent si le client a répondu ou non à l'offre et contiennent diverses informations démographiques.
- **demo_cs_1.sav.** Ce fichier de données d'hypothèse concerne la première mesure entreprise par une société pour compiler une base de données contenant des informations d'enquête. Chaque observation correspond à une ville différente. La région, la province, le quartier et la ville sont enregistrés.
- **demo_cs_2.sav.** Ce fichier de données d'hypothèse concerne la seconde mesure entreprise par une société pour compiler une base de données contenant des informations d'enquête. Chaque observation correspond à un ménage différent issu des villes sélectionnées à la première étape. La région, la province, le quartier, la ville, la sous-division et l'identification sont enregistrés. Les informations d'échantillonnage des deux premières étapes de la conception sont également incluses.
- **demo_cs.sav.** Ce fichier de données d'hypothèse concerne des informations d'enquête collectées via une méthode complexe d'échantillonnage. Chaque observation correspond à un ménage différent et diverses informations géographiques et d'échantillonnage sont enregistrées.
- **dmdata.sav.** Ceci est un fichier de données d'hypothèse qui contient des informations démographiques et des informations concernant les achats pour une entreprise de marketing direct. *dmdata2.sav* contient les informations pour un sous-ensemble de contacts qui ont reçu un envoi d'essai, et *dmdata3.sav* contient des informations sur les contacts restants qui n'ont pas reçu l'envoi d'essai.
- **dietstudy.sav.** Ce fichier de données d'hypothèse contient les résultats d'une étude portant sur le régime de Stillman (Rickman, Mitchell, Dingman, et Dalen, 1974). Chaque observation correspond à un sujet distinct et enregistre son poids en livres avant et après le régime, ainsi que ses niveaux de triglycérides en mg/100 ml.
- **dvdplayer.sav.** Ce fichier de données d'hypothèse concerne le développement d'un nouveau lecteur DVD. A l'aide d'un prototype, l'équipe de marketing a collecté des données de groupes spécifiques. Chaque observation correspond à un utilisateur interrogé et enregistre des informations démographiques sur cet utilisateur, ainsi que ses réponses aux questions portant sur le prototype.
- **german_credit.sav.** Ce fichier de données provient de l'ensemble de données « German credit » figurant dans le référentiel Machine Learning Databases (Blake et Merz, 1998) de l'université de Californie, Irvine.
- **grocery_1month.sav.** Ce fichier de données d'hypothèse est le fichier de données *grocery_coupons.sav* dans lequel les achats hebdomadaires sont organisés par client distinct. Certaines variables qui changeaient toutes les semaines disparaissent. En outre, le montant dépensé enregistré est à présent la somme des montants dépensés au cours des quatre semaines de l'enquête.
- **grocery_coupons.sav.** Il s'agit d'un fichier de données d'hypothèse qui contient des données d'enquête collectées par une chaîne de magasins d'alimentation qui cherchent à déterminer les habitudes de consommation de ses clients. Chaque client est suivi pendant quatre semaines et chaque observation correspond à une semaine distincte. Les informations enregistrées concernent les endroits où le client effectue ses achats, la manière dont il les effectue, ainsi que les sommes dépensées en provisions au cours de cette semaine.

- **guttman.sav.** Bell (Bell, 1961) a présenté un tableau pour illustrer les groupes sociaux possibles. Guttman (Guttman, 1968) a utilisé une partie de ce tableau, dans lequel cinq variables décrivant des éléments tels que l'interaction sociale, le sentiment d'appartenance à un groupe, la proximité physique des membres et la formalité de la relation, ont été croisées avec sept groupes sociaux théoriques, dont les foules (par exemple, le public d'un match de football), l'audience (par exemple, au cinéma ou dans une salle de classe), le public (par exemple, les journaux ou la télévision), les bandes (proche d'une foule, mais qui serait caractérisée par une interaction beaucoup plus intense), les groupes primaires (intimes), les groupes secondaires (volontaires) et la communauté moderne (groupement lâche issu d'une forte proximité physique et d'un besoin de services spécialisés).
- **health_funding.sav.** Ce fichier de données d'hypothèse concerne des données sur le financement des soins de santé (montant par groupe de 100 individus), les taux de maladie (taux par groupe de 10 000 individus) et les visites chez les prestataires de soins de santé (taux par groupe de 10 000 individus). Chaque observation représente une ville différente.
- **hivassay.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend un laboratoire pharmaceutique pour développer une analyse rapide de détection d'infection HIV. L'analyse a pour résultat huit nuances de rouge, les nuances les plus marquées indiquant une plus forte probabilité d'infection. Un test en laboratoire a été effectué sur 2 000 échantillons de sang, la moitié de ces échantillons étant infectée par le virus HIV et l'autre moitié étant saine.
- **hourlywagedata.sav.** Ce fichier de données d'hypothèse concerne les salaires horaires d'infirmières occupant des postes administratifs et dans les services de soins, et affichant divers niveaux d'expérience.
- **insurance_claims.sav.** Il s'agit d'un fichier de données hypothétiques qui concerne une compagnie d'assurance souhaitant développer un modèle pour signaler des réclamations suspectes, potentiellement frauduleuses. Chaque observation correspond à une réclamation distincte.
- **insure.sav.** Ce fichier de données d'hypothèse concerne une compagnie d'assurance qui étudie les facteurs de risque indiquant si un client sera amené à déclarer un incident au cours d'un contrat d'assurance vie d'une durée de 10 ans. Chaque observation figurant dans le fichier de données représente deux contrats, l'un ayant enregistré une réclamation et l'autre non, appariés par âge et sexe.
- **judges.sav.** Ce fichier de données d'hypothèse concerne les scores attribués par des juges expérimentés (plus un juge enthousiaste) à 300 performances de gymnastique. Chaque ligne représente une performance distincte ; les juges ont examiné les mêmes performances.
- **kinship_dat.sav.** Rosenberg et Kim (Rosenberg et Kim, 1975) se sont lancés dans l'analyse de 15 termes de parenté (cousin/cousine, fille, fils, frère, grand-mère, grand-père, mère, neveu, nièce, oncle, père, petite-fille, petit-fils, sœur, tante). Ils ont demandé à quatre groupes d'étudiants (deux groupes de femmes et deux groupes d'hommes) de trier ces termes en fonction des similarités. Deux groupes (un groupe de femmes et un groupe d'hommes) ont été invités à effectuer deux tris, en basant le second sur un autre critère que le premier. Ainsi, un total de six "sources" a été obtenu. Chaque source correspond à une matrice de proximité 15×15 , dont le nombre de cellules est égal au nombre de personnes dans une source moins le nombre de fois où les objets ont été partitionnés dans cette source.
- **kinship_ini.sav.** Ce fichier de données contient une configuration initiale d'une solution tridimensionnelle pour *kinship_dat.sav*.

- **kinship_var.sav.** Ce fichier de données contient les variables indépendantes *sexe*, *génération* et *degré* (de séparation) permettant d'interpréter les dimensions d'une solution pour *kinship_dat.sav*. Elles permettent en particulier de réduire l'espace de la solution à une combinaison linéaire de ces variables.
- **marketvalues.sav.** Ce fichier de données concerne les ventes de maisons dans un nouvel ensemble à Algonquin (Illinois) au cours des années 1999–2000. Ces ventes relèvent des archives publiques.
- **nhis2000_subset.sav.** Le NHIS (National Health Interview Survey) est une enquête de grande envergure concernant la population des États-Unis. Des entretiens ont lieu avec un échantillon de ménages représentatifs de la population américaine. Des informations démographiques et des observations sur l'état de santé et le comportement sanitaire sont recueillies auprès des membres de chaque ménage. Ce fichier de données contient un sous-groupe d'informations issues de l'enquête de 2000. National Center for Health Statistics. National Health Interview Survey, 2000. Fichier de données et documentation d'usage public. ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NHIS/2000/. Accès en 2003.
- **ozone.sav.** Les données incluent 330 observations portant sur six variables météorologiques pour prévoir la concentration d'ozone à partir des variables restantes. Des chercheurs précédents (Breiman et Friedman, 1985), (Hastie et Tibshirani, 1990), ont décelé parmi ces variables des non-linéarités qui pénalisent les approches standard de la régression.
- **pain_medication.sav.** Ce fichier de données d'hypothèse contient les résultats d'un essai clinique d'un remède anti-inflammatoire traitant les douleurs de l'arthrite chronique. On cherche notamment à déterminer le temps nécessaire au médicament pour agir et les résultats qu'il permet d'obtenir par rapport à un médicament existant.
- **patient_los.sav.** Ce fichier de données d'hypothèse contient les dossiers médicaux de patients admis à l'hôpital pour suspicion d'infarctus du myocarde suspecté (ou « attaque cardiaque »). Chaque observation correspond à un patient distinct et enregistre de nombreuses variables liées à son séjour à l'hôpital.
- **patlos_sample.sav.** Ce fichier de données d'hypothèse contient les dossiers médicaux d'un échantillon de patients sous traitement thrombolytique après un infarctus du myocarde. Chaque observation correspond à un patient distinct et enregistre de nombreuses variables liées à son séjour à l'hôpital.
- **polishing.sav.** Il s'agit du fichier de données du « Nambeware Polishing Times » de la Data and Story Library. Il concerne les mesures qu'entreprend un fabricant de vaisselle en métal (Nambe Mills, Santa Fe, Nouveau-Mexique) pour planifier sa production. Chaque observation représente un article différent de la gamme de produits. Le diamètre, le temps de polissage, le prix et le type de produit sont enregistrés pour chaque article.
- **poll_cs.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend un enquêteur pour déterminer le niveau de soutien du public pour un projet de loi avant législature. Les observations correspondent à des électeurs enregistrés. Chaque observation enregistre le comté, la ville et le quartier où habite l'électeur.
- **poll_cs_sample.sav.** Ce fichier de données d'hypothèse contient un échantillon des électeurs répertoriés dans le fichier *poll_cs.sav*. L'échantillon a été prélevé selon le plan spécifié dans le fichier de plan *poll_csplan* et ce fichier de données enregistre les probabilités d'inclusion et les pondérations d'échantillon. Toutefois, ce plan faisant appel à une méthode d'échantillonnage de probabilité proportionnelle à la taille (PPS – Probability-Proportional-to-Size), il existe

également un fichier contenant les probabilités de sélection conjointes (*poll_jointprob.sav*). Les variables supplémentaires correspondant à la répartition démographique des électeurs et à leur opinion sur le projet de loi proposé ont été collectées et ajoutées au fichier de données une fois l'échantillon prélevé.

- **property_assess.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend un contrôleur au niveau du comté pour maintenir les évaluations de valeur de propriété à jour sur des ressources limitées. Les observations correspondent à des propriétés vendues dans le comté au cours de l'année précédente. Chaque observation du fichier de données enregistre la ville où se trouve la propriété, l'évaluateur ayant visité la propriété pour la dernière fois, le temps écoulé depuis cette évaluation, l'évaluation effectuée à ce moment-là et la valeur de vente de la propriété.
- **property_assess_cs.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend un contrôleur du gouvernement pour maintenir les évaluations de valeur de propriété à jour sur des ressources limitées. Les observations correspondent à des propriétés de l'état. Chaque observation du fichier de données enregistre le comté, la ville et le quartier où se trouve la propriété, le temps écoulé depuis la dernière évaluation et l'évaluation alors effectuée.
- **property_assess_cs_sample.sav.** Ce fichier de données d'hypothèse contient un échantillon des propriétés répertoriées dans le fichier *property_assess_cs.sav*. L'échantillon a été prélevé selon le plan spécifié dans le fichier de plan *property_assess_csplan* et ce fichier de données enregistre les probabilités d'inclusion et les pondérations d'échantillon. La variable supplémentaire *Valeur courante* a été collectée et ajoutée au fichier de données une fois l'échantillon prélevé.
- **recidivism.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend une agence administrative d'application de la loi pour interpréter les taux de récidive dans la juridiction. Chaque observation correspond à un récidiviste et enregistre les informations démographiques qui lui sont propres, certains détails sur le premier délit commis, ainsi que le temps écoulé jusqu'à la seconde arrestation si elle s'est produite dans les deux années suivant la première.
- **recidivism_cs_sample.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend une agence administrative d'application de la loi pour interpréter les taux de récidive dans la juridiction. Chaque observation correspond à un récidiviste libéré suite à la première arrestation en juin 2003 et enregistre les informations démographiques qui lui sont propres, certains détails sur le premier délit commis et les données relatives à la seconde arrestation, si elle a eu lieu avant fin juin 2006. Les récidivistes ont été choisis dans plusieurs départements échantillonnés conformément au plan d'échantillonnage spécifié dans *recidivism_cs.cspan*. Ce plan faisant appel à une méthode d'échantillonnage de probabilité proportionnelle à la taille (PPS - Probability proportional to size), il existe également un fichier contenant les probabilités de sélection conjointes (*recidivism_cs_jointprob.sav*).
- **rfm_transactions.sav.** Un fichier de données d'hypothèse qui contient les données de transaction d'achat, y compris la date d'achat, le/les élément(s) acheté(s) et le montant monétaire pour chaque transaction.
- **salesperformance.sav.** Ce fichier de données d'hypothèse concerne l'évaluation de deux nouveaux cours de formation en vente. Soixante employés, divisés en trois groupes, reçoivent chacun une formation standard. En outre, le groupe 2 suit une formation technique et le groupe 3 un didacticiel pratique. A l'issue du cours de formation, chaque employé est testé et

sa note enregistrée. Chaque observation du fichier de données représente un stagiaire distinct et enregistre le groupe auquel il a été assigné et la note qu'il a obtenue au test.

- **satisf.sav.** Il s'agit d'un fichier de données d'hypothèse portant sur une enquête de satisfaction effectuée par une société de vente au détail au niveau de quatre magasins. Un total de 582 clients ont été interrogés et chaque observation représente la réponse d'un seul client.
- **screws.sav.** Ce fichier de données contient des informations sur les descriptives des vis, des boulons, des écrous et des clous. (Hartigan, 1975).
- **shampoo_ph.sav.** Ce fichier de données d'hypothèse concerne le processus de contrôle qualité observé dans une usine de produits capillaires. A intervalles réguliers, six lots de sortie distincts sont mesurés et leur pH enregistré. La plage cible est 4,5–5,5.
- **ships.sav.** Il s'agit d'un ensemble de données présenté et analysé ailleurs (McCullagh et al., 1989) et concernant les dommages causés à des cargos par les vagues. Les effectifs d'incidents peuvent être modélisés comme des incidents se produisant selon un taux de Poisson en fonction du type de navire, de la période de construction et de la période de service. Les mois de service totalisés pour chaque cellule du tableau formé par la classification croisée des facteurs fournissent les valeurs d'exposition au risque.
- **site.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend une société pour choisir de nouveaux sites pour le développement de ses activités. L'entreprise a fait appel à deux consultants pour évaluer séparément les sites. Ces consultants, en plus de fournir un rapport approfondi, ont classé chaque site comme constituant une éventualité « bonne », « moyenne » ou « faible ».
- **smokers.sav.** Ce fichier de données est extrait de l'étude National Household Survey of Drug Abuse de 1998 et constitue un échantillon de probabilité des ménages américains. (<http://dx.doi.org/10.3886/ICPSR02934>) Ainsi, la première étape dans l'analyse de ce fichier doit consister à pondérer les données pour refléter les tendances de population.
- **stroke_clean.sav.** Ce fichier de données d'hypothèse concerne l'état d'une base de données médicales une fois celle-ci purgée via des procédures de l'option Validation de données.
- **stroke_invalid.sav.** Ce fichier de données d'hypothèse concerne l'état initial d'une base de données médicales et comporte plusieurs erreurs de saisie de données.
- **stroke_survival.** Ce fichier de données d'hypothèse concerne les temps de survie de patients qui quittent un programme de rééducation à la suite d'un accident ischémique et rencontrent un certain nombre de problèmes. Après l'attaque, l'occurrence d'infarctus du myocarde, d'accidents ischémiques ou hémorragiques est signalée, et le moment de l'événement enregistré. L'échantillon est tronqué à gauche car il n'inclut que les patients ayant survécu durant le programme de rééducation mis en place suite à une attaque.
- **stroke_valid.sav.** Ce fichier de données d'hypothèse concerne l'état d'une base de données médicales une fois les valeurs vérifiées via la procédure Validation de données. Elle contient encore des observations anormales potentielles.
- **survey_sample.sav.** Ce fichier de données concerne des informations d'enquête dont des données démographiques et des mesures comportementales. Il est basé sur un sous-ensemble de variables de la 1998 NORC General Social Survey, bien que certaines valeurs de données aient été modifiées et que des variables supplémentaires fictives aient été ajoutées à titre de démonstration.

- **telco.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend une société de télécommunications pour réduire les taux de désabonnement de sa base de clients. Chaque observation correspond à un client distinct et enregistre diverses informations démographiques et d'utilisation de service.
- **telco_extra.sav.** Ce fichier de données est semblable au fichier de données *telco.sav* mais les variables de permanence et de dépenses des consommateurs transformées log ont été supprimées et remplacées par des variables de dépenses des consommateurs transformées log standardisées.
- **telco_missing.sav.** Ce fichier de données est un sous-ensemble du fichier de données *telco.sav* mais certaines des valeurs de données démographiques ont été remplacées par des valeurs manquantes.
- **testmarket.sav.** Ce fichier de données d'hypothèse concerne une chaîne de fast foods et ses plans marketing visant à ajouter un nouveau plat à son menu. Trois campagnes étant possibles pour promouvoir le nouveau produit, le nouveau plat est introduit sur des sites sur plusieurs marchés sélectionnés au hasard. Une promotion différente est effectuée sur chaque site et les ventes hebdomadaires du nouveau plat sont enregistrées pour les quatre premières semaines. Chaque observation correspond à un site-semaine distinct.
- **testmarket_1month.sav.** Ce fichier de données d'hypothèse est le fichier de données *testmarket.sav* dans lequel les ventes hebdomadaires sont organisées par site distinct. Certaines variables qui changeaient toutes les semaines disparaissent. En outre, les ventes enregistrées sont à présent la somme des ventes réalisées au cours des quatre semaines de l'enquête.
- **tree_car.sav.** Ce fichier de données d'hypothèse concerne des données démographiques et de prix d'achat de véhicule.
- **tree_credit.sav.** Ce fichier de données d'hypothèse concerne des données démographiques et d'historique de prêt bancaire.
- **tree_missing_data.sav** Ce fichier de données d'hypothèse concerne des données démographiques et d'historique de prêt bancaire avec un grand nombre de valeurs manquantes.
- **tree_score_car.sav.** Ce fichier de données d'hypothèse concerne des données démographiques et de prix d'achat de véhicule.
- **tree_textdata.sav.** Ce fichier de données simples ne comporte que deux variables et vise essentiellement à indiquer l'état par défaut des variables avant affectation du niveau de mesure et des étiquettes de valeurs.
- **tv-survey.sav.** Ce fichier de données d'hypothèse concerne une enquête menée par un studio de télévision qui envisage de prolonger la diffusion d'un programme ou de l'arrêter. On a demandé à 906 personnes si elles regarderaient le programme dans diverses situations. Chaque ligne représente un répondant distinct et chaque colonne une situation distincte.
- **ulcer_recurrence.sav.** Ce fichier contient des informations partielles d'une enquête visant à comparer l'efficacité de deux thérapies de prévention de la récurrence des ulcères. Il fournit un bon exemple de données censurées par intervalle et a été présenté et analysé ailleurs (Collett, 2003).

- **ulcer_recurrence_recoded.sav.** Ce fichier réorganise les informations figurant dans le fichier *ulcer_recurrence.sav* pour que vous puissiez modéliser la probabilité d'événement pour chaque intervalle de l'enquête plutôt que la probabilité d'événement de fin d'enquête. Il a été présenté et analysé ailleurs (Collett et al., 2003).
- **verd1985.sav.** Ce fichier de données concerne une enquête (Verdegaal, 1985). Les réponses de 15 sujets à 8 variables ont été enregistrées. Les variables présentant un intérêt sont divisées en trois ensembles. Le groupe 1 comprend l'*âge* et la *situation familiale*, le groupe 2 les *animaux domestiques* et la *presse*, et le groupe 3 la *musique* et l'*habitat*. A la variable *animal domestique* est appliqué un codage nominal multiple et à *âge*, un codage ordinal ; toutes les autres variables ont un codage nominal simple.
- **virus.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend un fournisseur de services Internet pour déterminer les effets d'un virus sur ses réseaux. Il a suivi le pourcentage (approximatif) de trafic de messages électroniques infectés par un virus sur ses réseaux sur la durée, de la découverte à la circonscription de la menace.
- **wheeze_steubenville.sav.** Il s'agit d'un sous-ensemble d'une enquête longitudinale des effets de la pollution de l'air sur la santé des enfants (Ware, Dockery, Spiro III, Speizer, et Ferris Jr., 1984). Les données contiennent des mesures binaires répétées de l'état asthmatique d'enfants de la ville de Steubenville (Ohio), âgés de 7, 8, 9 et 10 ans, et indiquent si la mère fumait au cours de la première année de l'enquête.
- **workprog.sav.** Ce fichier de données d'hypothèse concerne un programme de l'administration visant à proposer de meilleurs postes aux personnes défavorisées. Un échantillon de participants potentiels au programme a ensuite été prélevé. Certains de ces participants ont été sélectionnés au hasard pour participer au programme. Chaque observation représente un participant au programme distinct.

Notices

Licensed Materials – Property of SPSS Inc., an IBM Company. © Copyright SPSS Inc. 1989, 2010.

Patent No. 7,023,453

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: SPSS INC., AN IBM COMPANY, PROVIDES THIS PUBLICATION “AS IS” WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. SPSS Inc. may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-SPSS and non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this SPSS Inc. product and use of those Web sites is at your own risk.

When you send information to IBM or SPSS, you grant IBM and SPSS a nonexclusive right to use or distribute the information in any way it believes appropriate without incurring any obligation to you.

Information concerning non-SPSS products was obtained from the suppliers of those products, their published announcements or other publicly available sources. SPSS has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-SPSS products. Questions on the capabilities of non-SPSS products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to SPSS Inc., for the purposes of developing,

using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. SPSS Inc., therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided “AS IS”, without warranty of any kind. SPSS Inc. shall not be liable for any damages arising out of your use of the sample programs.

Trademarks

IBM, the IBM logo, and [ibm.com](http://www.ibm.com) are trademarks of IBM Corporation, registered in many jurisdictions worldwide. A current list of IBM trademarks is available on the Web at <http://www.ibm.com/legal/copytrade.shtml>.

SPSS is a trademark of SPSS Inc., an IBM Company, registered in many jurisdictions worldwide.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Java and all Java-based trademarks and logos are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

This product uses WinWrap Basic, Copyright 1993-2007, Polar Engineering and Consulting, <http://www.winwrap.com>.

Other product and service names might be trademarks of IBM, SPSS, or other companies.

Adobe product screenshot(s) reprinted with permission from Adobe Systems Incorporated.

Microsoft product screenshot(s) reprinted with permission from Microsoft Corporation.



Bibliographie

- Akaah, R. P., et P. K. Korgaonkar. 1988. A conjoint investigation of the relative importance of risk relievers in direct marketing. *Journal of Advertising Research*, 28:4, .
- Bell, E. H. 1961. *Social foundations of human behavior: Introduction to the study of sociology*. New York: Harper & Row.
- Blake, C. L., et C. J. Merz. 1998. "UCI Repository of machine learning databases." Available at <http://www.ics.uci.edu/~mlearn/MLRepository.html>.
- Breiman, L., et J. H. Friedman. 1985. Estimating optimal transformations for multiple regression and correlation. *Journal of the American Statistical Association*, 80, .
- Cattin, P., et D. R. Wittink. 1982. Commercial use of conjoint analysis: A survey. *Journal of Marketing*, 46:3, .
- Collett, D. 2003. *Modelling survival data in medical research*, 2 éd. Boca Raton: Chapman & Hall/CRC.
- Green, P. E., et V. Rao. 1972. *Applied multidimensional scaling*. Hinsdale, Ill.: Dryden Press.
- Green, P. E., et Y. Wind. 1973. *Multiattribute decisions in marketing: A measurement approach*. Hinsdale, Ill.: Dryden Press.
- Guttman, L. 1968. A general nonmetric technique for finding the smallest coordinate space for configurations of points. *Psychometrika*, 33, .
- Hartigan, J. A. 1975. *Clustering algorithms*. New York: John Wiley and Sons.
- Hastie, T., et R. Tibshirani. 1990. *Generalized additive models*. Londres: Chapman and Hall.
- Kennedy, R., C. Riquier, et B. Sharp. 1996. Practical applications of correspondence analysis to categorical data in market research. *Journal of Targeting, Measurement, and Analysis for Marketing*, 5, .
- McCullagh, P., et J. A. Nelder. 1989. *Generalized Linear Models*, 2nd éd. Londres: Chapman & Hall.
- Price, R. H., et D. L. Bouffard. 1974. Behavioral appropriateness and situational constraints as dimensions of social behavior. *Journal of Personality and Social Psychology*, 30, .
- Rickman, R., N. Mitchell, J. Dingman, et J. E. Dalen. 1974. Changes in serum cholesterol during the Stillman Diet. *Journal of the American Medical Association*, 228, .
- Rosenberg, S., et M. P. Kim. 1975. The method of sorting as a data-gathering procedure in multivariate research. *Multivariate Behavioral Research*, 10, .
- Van der Ham, T., J. J. Meulman, D. C. Van Strien, et H. Van Engeland. 1997. Empirically based subgrouping of eating disorders in adolescents: A longitudinal perspective. *British Journal of Psychiatry*, 170, .
- Verdegaal, R. 1985. *Meer sets analyse voor kwalitatieve gegevens (en néerlandais)*. Leiden: Department of Data Theory, University of Leiden.

Ware, J. H., D. W. Dockery, A. Spiro III, F. E. Speizer, et B. G. Ferris Jr.. 1984. Passive smoking, gas cooking, and respiratory health of children living in six cities. *American Review of Respiratory Diseases*, 129, .

- Afficher plan orthogonal, 2, 9, 21
 -)CARD, 10
 - enregistrement des profils, 10
 - format de liste, 9
 - format de profil unique, 9
 - Pieds de page, 10
 - Titres, 10
- approche du profil entier, 2

-)CARD
 - dans Afficher plan orthogonal, 10
- card_ variable
 - dans Générer plan orthogonal, 5
- Coefficients, 26
- Coefficients de corrélation, 27

- facteurs, 2, 17–18
- Fichiers de données
 - dans Générer plan orthogonal, 5
- fichiers d'exemple
 - emplacement, 31
- format de liste
 - dans Afficher plan orthogonal, 9
- format de profil unique
 - dans Afficher plan orthogonal, 9

- Génération de nombres aléatoires
 - dans Générer plan orthogonal, 5
- Générer plan orthogonal, 2, 5, 18
 - définition des noms, des étiquettes et des valeurs des facteurs, 7
 - Fichiers de données, 5
 - Génération de nombres aléatoires, 5
 - Nombre minimum d'observations, 8
 - observations de simulation, 8
 - observations traitées, 8

- inversements, 28

- legal notices, 42

- modèle anti-idéal, 25
- Modèle BTL (Bradley-Terry-Luce), 29
- modèle discret, 24
- modèle idéal, 25
- Modèle linéaire, 25

- modèle logit, 29
- modèle utilitaire maximum, 29

- niveaux de facteurs, 2, 17–18

- observations de simulation, 3, 14, 28
 - dans Générer plan orthogonal, 8
- observations traitées, 2
 - dans Générer plan orthogonal, 8

- Pieds de page
 - dans Afficher plan orthogonal, 10
- plans orthogonaux
 - affichage, 21
 - Affichage, 9
 - génération, 5, 18
 - Nombre minimum d'observations, 8
 - observations traitées, 8

- R de Pearson, 27
- résultats de simulation
 - Modèle BTL (Bradley-Terry-Luce), 29
 - modèle logit, 29
 - modèle utilitaire maximum, 29

- scores des outils, 3, 25
- scores d'importance, 26
- status_ variable
 - dans Générer plan orthogonal, 5
- Syntaxe
 - Commande CONJOINT, 24
- Syntaxe de commande
 - Commande CONJOINT, 24

- tableau orthogonal, 2
- taille de l'échantillon, 3
- Tau de Kendall, 27
- Titres
 - dans Afficher plan orthogonal, 10
- trademarks, 43

- utilitaire total, 25

- valeurs de partie, 3
- valeurs d'importance, 26