

IBM SPSS Custom Tables 21



Remarque : Avant d'utiliser ces informations et le produit qu'elles concernent, lisez les informations générales sous Remarques sur p. 208.

Cette version s'applique à IBM® SPSS® Statistics 21 et à toutes les publications et modifications ultérieures jusqu'à mention contraire dans les nouvelles versions.

Les captures d'écran des produits Adobe sont reproduites avec l'autorisation de Adobe Systems Incorporated.

Les captures d'écran des produits Microsoft sont reproduites avec l'autorisation de Microsoft Corporation.

Matériel sous licence - Propriété d'IBM

© Copyright IBM Corporation 1989, 2012.

Droits limités pour les utilisateurs au sein d'administrations américaines : utilisation, copie ou divulgation soumise au GSA ADP Schedule Contract avec IBM Corp.

Préface

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

A propos de IBM Business Analytics

Le logiciel IBM Business Analytics offre des informations complètes, cohérentes et précises permettant aux preneurs de décision d'améliorer leurs performances professionnelles. Un portefeuille complet de solutions de [business intelligence](#), [d'analyses prédictives](#), [de performance financière et de gestion de la stratégie](#), et [d'applications analytiques](#) permet une connaissance claire et immédiate et offre des possibilités d'actions sur les performances actuelles et la capacité de prédire les résultats futurs. En combinant des solutions du secteur, des pratiques prouvées et des services professionnels, les entreprises de toute taille peuvent générer la plus grande productivité, automatiser les décisions en toute confiance et apporter de meilleurs résultats.

Dans le cadre de ce portefeuille, le logiciel IBM SPSS Predictive Analytics aide les entreprises à prédire des événements futurs et à agir de manière proactive en fonction de ces prédictions pour apporter de meilleurs résultats. Des clients dans les domaines commerciaux, gouvernementaux et académiques se servent de la technologie IBM SPSS comme d'un avantage concurrentiel pour attirer ou retenir des clients, tout en réduisant les risques liés à l'incertitude et à la fraude. En intégrant le logiciel IBM SPSS à leurs opérations quotidiennes, les entreprises peuvent effectuer des prévisions, et sont capables de diriger et d'automatiser leurs décisions afin d'atteindre leurs objectifs commerciaux et d'obtenir des avantages concurrentiels mesurables. Pour plus d'informations ou pour contacter un représentant, visitez le site <http://www.ibm.com/spss>.

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 IBM Corp. ou l'installation dans l'un des environnements matériels pris en charge. Pour contacter l'assistance technique, visitez le site IBM Corp. à l'adresse <http://www.ibm.com/support>. Votre nom, celui de votre société, ainsi que votre contrat d'assistance vous seront demandés.

Support technique pour les étudiants

Si vous êtes un étudiant qui utilise la version pour étudiant, personnel de l'éducation ou diplômé d'un produit logiciel IBM SPSS, veuillez consulter les pages [Solutions pour l'éducation](#) (<http://www.ibm.com/spss/rd/students/>) consacrées aux étudiants. Si vous êtes un étudiant utilisant une copie du logiciel IBM SPSS fournie par votre université, veuillez contacter le coordinateur des produits IBM SPSS de votre université.

Service clients

Si vous avez des questions concernant votre livraison ou votre compte, contactez votre bureau local. Veuillez préparer et conserver votre numéro de série à portée de main pour l'identification.

Séminaires de formation

IBM Corp. 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, accédez au site <http://www.ibm.com/software/analytics/spss/training>.

Contenu

1 Initiation aux tableaux personnalisés 1

Structure des tableaux et terminologie	1
Tableaux pivotants	1
Variables et niveau de mesure	2
Lignes, colonnes et cellules	3
Empilement	3
Tableau croisé	4
Emboîtement	4
Strates	5
Tableaux pour variables avec modalités partagées	5
Vecteurs de réponses multiples	6
Totaux et sous-totaux	6
Statistiques récapitulatives personnalisées pour les totaux	7
Fichier de données d'exemple	8
Construction d'un tableau	8
Ouverture du générateur de tableaux personnalisés	9
Sélection de variables en ligne et en colonne	10
Insertion de totaux et de sous-totaux	15
Récapitulation des variables d'échelle	18

2 Interface de génération de tableaux 26

Construction de tableaux	26
Pour construire un tableau	29
Empilement de variables	30
Emboîtement de variables	31
Strates	32
Affichage et masquage des noms et/ou étiquettes de variable	33
Statistiques récapitulatives	34
Modalités et totaux	41
Catégories calculées	46
Tableaux de variables avec modalités partagées (Comperimeter Tables)	48
Personnalisation du générateur de tableaux	49
Tableaux personnalisés : Onglet Options	49
Tableaux personnalisés : Onglet Titres	51
Tableaux personnalisés : Onglet Statistiques de test	52

3 Tableaux simples pour variables qualitatives 56

Une seule variable qualitative	57
Pourcentages	58
Totaux	60
Tableau croisé	61
Pourcentages dans des tableaux croisés	62
Contrôle du format d'affichage	63
Totaux marginaux	64
Tri et exclusion de modalités	65

4 Empilement, emboîtement et strates avec des variables qualitatives 70

Empilement de variables qualitatives	70
Empilement et tableau croisé	71
Emboîtement de variables qualitatives	73
Suppression d'étiquettes de variable	76
Tableau croisé emboîté	78
Strates	81
Deux variables qualitatives de strate empilées	83
Deux variables qualitatives de strate emboîtées	84

5 Totaux et sous-totaux pour les variables qualitatives 86

Total simple pour une seule variable	86
Tel écran, tel totalisé	87
Position d'affichage des totaux	88
Totaux pour tableaux emboîtés	89
Totaux des variables de strate	91
Sous-totaux	93
Tel écran, tel sous-totalisé	94
Masquage de modalités de sous-total	95
Sous-totaux d'une variable de strate	97

6 Les catégories calculées pour les variables qualitatives 99

Catégorie calculée simple	99
Masquer les catégories dans une catégorie calculée	101
Référencement des sous-totaux dans une catégorie calculée	103
Utilisation des catégories calculées pour afficher des sous-totaux non-exhaustifs	107

7 Tableaux pour variables avec modalités partagées 110

Tableau d'effectifs	110
Tableau de pourcentages	113
Contrôle des totaux et des modalités	115
Emboîtement dans des tableaux comportant des modalités partagées	117

8 Statistiques récapitulatives 119

Variable source des statistiques récapitulatives	120
Source des statistiques récapitulatives pour les variables qualitatives	120
Source des statistiques récapitulatives pour les variables d'échelle	123
Variables empilées	125
Autres statistiques récapitulatives en total pour les variables qualitatives	128
Affichage des valeurs des modalités	131

9 Récapitulation des variables d'échelle 134

Variables d'échelle empilées	134
Statistiques récapitulatives multiples	135
Effectif, N valide et valeurs manquantes	136
Différents récapitulatifs pour différentes variables	138
Récapitulatifs de groupe dans des modalités	140
Variables de regroupement multiples	141
Emboîtement de variables qualitatives dans des variables d'échelle	143

10 Statistiques de test **145**

Tests d'indépendance (Khi-deux)	145
Effets de l'emboîtement et de l'empilement sur les tests d'indépendance	148
Comparaison des moyennes de colonne	150
Effets de l'emboîtement et de l'empilement sur les tests des moyennes de colonne	155
Comparaison des proportions de colonne	157
Effets de l'emboîtement et de l'empilement sur les tests des proportions de colonne	162
Une remarque sur les pondérations et les vecteurs multiréponses	164

11 Vecteurs de réponses multiples **165**

Effectifs, réponses, pourcentages et totaux	165
Utilisation de vecteurs multiréponses avec d'autres variables	169
Variable source de statistiques et statistiques récapitulatives disponibles	171
Vecteurs de modalités multiples et réponses en double	172
Test de signification pour les vecteurs de réponses multiples	174
Tests d'indépendance avec des vecteurs de réponses multiples	175
Comparaison des moyennes de colonne avec des vecteurs de réponses multiples	178

12 Valeurs manquantes **180**

Tableaux dépourvus de valeurs manquantes	180
Inclusion de valeurs manquantes dans les tableaux	182

13 Formatage et personnalisation de tableaux **186**

Formats d'affichage des statistiques récapitulatives	186
Affichage des étiquettes pour les statistiques récapitulatives	190
Largeur des colonnes	192
Affichage d'une valeur dans les cellules vides	194
Affichage d'une valeur pour les statistiques manquantes	194

Annexes

A Fichiers d'exemple **197**

B Remarques **208**

Index **211**

Initiation aux tableaux personnalisés

De nombreuses procédures génèrent des résultats sous forme de tableaux. Toutefois, le module complémentaire Tableaux personnalisés offre des fonctionnalités spéciales conçues pour prendre en charge une grande variété de fonctions de génération de rapports personnalisés. Un grand nombre de fonctions personnalisées sont particulièrement utiles pour l'analyse d'enquêtes et la recherche en marketing.

Le présent guide suppose que vous connaissez déjà les fonctionnalités de base de IBM® SPSS® Statistics. Si vous n'êtes pas familiarisé avec les fonctionnalités de base, consultez le didacticiel d'introduction fourni avec le logiciel. Dans la barre de menus d'une fenêtre SPSS Statistics ouverte, choisissez :

Aide > Didacticiel

Structure des tableaux et terminologie

Le module complémentaire Tableaux personnalisés permet de produire une grande variété de tableaux personnalisés. Même si vous pouvez découvrir une grande partie de ses fonctionnalités simplement en manipulant l'interface de génération de tableaux, il peut s'avérer utile de connaître les bases de la structure des tableaux, ainsi que les termes employés pour décrire différents éléments structurels que vous pouvez utiliser dans un tableau.

Tableaux pivotants

Les tableaux produits via le module Tableaux personnalisés apparaissent sous la forme de **tableaux pivotants** dans la fenêtre du Viewer. Les tableaux pivotants offrent une grande flexibilité quant au formatage et à la présentation des tableaux.

Pour plus d'informations sur l'utilisation des tableaux pivotants, reportez-vous à l'aide.

- ▶ A partir des menus d'une fenêtre ouverte, sélectionnez :
Aide > Rubriques
- ▶ Dans le panneau Contenu, double-cliquez sur Système central.
- ▶ Dans le sommaire développé, double-cliquez ensuite sur Tableaux pivotants.

Variables et niveau de mesure

Dans une certaine mesure, les opérations que vous pouvez effectuer avec une variable sont limitées par le niveau de mesure défini. La procédure Tableaux personnalisés distingue deux types de variables de base, par rapport au niveau de mesure :

Qualitatives : Données ayant un nombre limité de valeurs ou de modalités distinctes (par exemple, sexe ou religion). Elles sont parfois également qualifiées de données qualitatives. Les variables qualitatives peuvent être des données chaîne (alphanumérique) ou des variables numériques qui utilisent des codes chiffrés pour représenter les modalités (par exemple, 0 = *Féminin* et 1 = *Masculin*). Les variables qualitatives peuvent en outre être divisées selon les types suivants :

- **Nominal**. Une variable peut être traitée comme étant nominale si ses valeurs représentent des modalités sans classement intrinsèque (par exemple, le service de la société dans lequel travaille un employé). La région, le code postal ou l'appartenance religieuse sont des exemples de variables nominales.
- **Ordinal**. Une variable peut être traitée comme étant ordinale si ses valeurs représentent des modalités associées à un classement intrinsèque (par exemple, des niveaux de satisfaction allant de Très mécontent à Très satisfait). Exemples de variable ordinale : des scores d'attitude représentant le degré de satisfaction ou de confiance, et des scores de classement des préférences.

Les variables définies comme étant nominales ou ordinales dans Data Editor sont traitées en tant que variables qualitatives dans la procédure Tableaux personnalisés.

Echelle. Une variable peut être traitée comme une variable d'échelle (continue) si ses valeurs représentent des modalités ordonnées avec une mesure significative, de sorte que les comparaisons de distance entre les valeurs soient adéquates. L'âge en années et le revenu en milliers de dollars sont des exemples de variable d'échelle. Ces données sont aussi appelées données quantitatives ou données continues. Les variables définies comme étant des variables d'échelle dans Data Editor sont traitées en tant que variables d'échelle dans la procédure Tableaux personnalisés.

Étiquettes de valeurs

Pour les variables qualitatives, l'aperçu affiché dans le panneau de canevas du générateur de tableaux se base sur les **étiquettes de valeur** définies. Les modalités affichées dans le tableau sont, en fait, les étiquettes de valeur définies pour cette variable. Si aucune étiquette de valeur n'a été définie pour la variable, l'aperçu affiche deux modalités génériques. Le nombre réel des modalités qui apparaîtront dans le tableau final est déterminé par le nombre d'occurrences de valeurs distinctes dans les données. L'aperçu suppose simplement que deux modalités au moins seront présentes.

En outre, certaines fonctions de génération de tableaux personnalisés ne sont pas disponibles pour les variables qualitatives pour lesquelles aucune étiquette de valeur n'a été définie.

Lignes, colonnes et cellules

Chaque dimension d'un tableau est définie par une variable unique ou une combinaison de variables. Les variables qui apparaissent de haut en bas à gauche d'un tableau sont appelées **variables de ligne**. Elles définissent les lignes d'un tableau. Les variables qui apparaissent de gauche à droite en haut d'un tableau sont appelées **variables de colonne**. Elles définissent les colonnes d'un tableau. Le corps d'un tableau est constitué de **cellules**, qui contiennent les informations de base fournies par le tableau (effectifs, sommes, moyennes, pourcentages, etc). Une cellule est formée par l'intersection d'une ligne et d'une colonne d'un tableau.

Empilement

Empiler revient à prendre des tableaux séparés et à les coller ensemble dans le même affichage. Vous pouvez, par exemple, afficher des informations sur les variables *sexe* et *tranche d'âge* dans des sections séparées du même tableau.

Figure 1-1
Variables empilées

		Effectif
Sexe	Homme	1232
	Femme	1600
Catégorie d'âge	Moins de 25	242
	25 à 34	627
	35 à 44	679
	45 à 54	481
	55 à 64	320
	65 et +	479

Bien que le terme « empilement » fasse généralement référence à un affichage vertical, vous pouvez également procéder à un empilement horizontal des variables.

Figure 1-2
Empilement horizontal

Sexe		Catégorie d'âge					
Homme	Femme	Moins de 25	25 à 34	35 à 44	45 à 54	55 à 64	65 et +
1232	1600	242	627	679	481	320	479

Tableau croisé

La technique des tableaux croisés est une technique de base qui permet d'examiner la relation entre deux variables qualitatives. Par exemple, en utilisant la variable *Tranche d'âge* comme variable de ligne et la variable *Sexe* comme variable de colonne, vous pouvez créer un tableau croisé bidimensionnel présentant le nombre d'hommes et de femmes dans chaque tranche d'âge.

Figure 1-3
Tableau croisé bidimensionnel simple

		Sexe	
		Homme	Femme
Catégorie d'âge	Moins de 25	108	134
	25 à 34	276	351
	35 à 44	309	370
	45 à 54	221	260
	55 à 64	136	184
	65 et +	178	301

Emboîtement

L'emboîtement, comme la technique des tableaux croisés, permet de présenter la relation entre deux variables qualitatives ; la seule différence est qu'une variable est emboîtée dans l'autre dans la même dimension. Vous pouvez, par exemple, emboîter la variable *Sexe* dans la variable *Tranche d'âge* dans la dimension de ligne, ce qui a pour effet de présenter le nombre d'hommes et de femmes dans chaque tranche d'âge.

Dans cet exemple, et pour les deux variables, le tableau emboîté affiche principalement les mêmes informations qu'un tableau croisé.

Figure 1-4
Variables emboîtées

Catégorie d'âge		Effectif	
		Homme	Femme
Moins de 25	Homme	108	
	Femme		134
25 à 34	Homme	276	
	Femme		351
35 à 44	Homme	309	
	Femme		370
45 à 54	Homme	221	
	Femme		260
55 à 64	Homme	136	
	Femme		184
65 et +	Homme	178	
	Femme		301

Strates

Vous pouvez utiliser des strates pour ajouter une dimension de profondeur à vos tableaux et créer ainsi des « cubes » tridimensionnels. Les strates sont en fait relativement similaires à l'emboîtement ; la différence principale réside dans le fait qu'une seule modalité de strate est visible à la fois. Par exemple, l'utilisation de la variable *Tranche d'âge* comme variable de ligne et de la variable *Sexe* comme variable de strate a pour effet de produire un tableau dans lequel les informations relatives aux hommes et aux femmes apparaissent dans différentes strates.

Figure 1-5
Variables en strates

Sexe du répondant Femme		
Catégories d'âges	Moins de 25	134
	25 à 34	351
	35 à 44	370

Sexe du répondant Homme		
Catégories d'âges	Moins de 25	108
	25 à 34	276
	35 à 44	309
	45 à 54	221
	55 à 64	136
	65 ou plus	178

Tableaux pour variables avec modalités partagées

Les enquêtes contiennent souvent de nombreuses questions caractérisées par un ensemble commun de réponses possibles. Notre échantillon d'enquête contient ainsi un certain nombre de variables relatives à la confiance accordée à divers services et institutions publics et privés, qui présentent toutes le même ensemble de modalités de réponses : 1 = *Beaucoup*, 2 = *Seulement certaines* et 3 = *Presque aucune*. Vous pouvez utiliser la technique de l'empilement pour afficher ces variables liées dans le même tableau et afficher les modalités de réponses partagées dans les colonnes du tableau.

Figure 1-6
Variables empilées avec modalités de réponses partagées dans les colonnes

	Enormém	Un peu	Pas du
Confiance dans les institutions bancaires et financières	490	1068	306
Confiance dans les multinationales	500	1078	243
Confiance dans l'éducation	511	1055	315
Confiance dans la presse	176	878	808
Confiance dans la médecine	844	864	167
Confiance dans la télévision	196	936	744

Vecteurs de réponses multiples

Les vecteurs de réponses multiples utilisent plusieurs variables pour enregistrer les réponses à des questions auxquelles le répondant peut donner plusieurs réponses. Notre échantillon d'enquête pose ainsi la question « Parmi les sources suivantes, quelles sont les plus fiables dans le domaine de l'information ? » Les répondants peuvent sélectionner une combinaison de cinq choix possibles : *Internet, télévision, radio, journaux* et *journaux télévisés*. Chacun de ces choix est stocké sous la forme d'une variable distincte dans le fichier de données ; ils forment ensemble un vecteur de réponses multiples. Via le module Tableaux personnalisés, vous pouvez définir un vecteur multiréponse sur la base de ces variables et utiliser ce vecteur dans les tableaux créés.

Figure 1-7
Vecteur multiréponses affiché dans un tableau

		Effectif	N % colonne
News sources	Obtention des informations à partir d'internet	867	41,7%
	Obtention des informations à partir de la radio	551	26,5%
	Obtention des informations à partir de la télévision	1077	51,8%
	Obtention des informations à partir des magazines	294	14,1%
	Obtention des informations à partir de la presse quotidienne	805	38,7%

Vous pouvez noter, dans cet exemple, que le total des pourcentages est supérieur à 100 %. Chaque répondant devant choisir plusieurs réponses, le nombre total de réponses peut être supérieur au nombre total de répondants.

Totaux et sous-totaux

Vous avez un contrôle étendu sur l'affichage des totaux et des sous-totaux, notamment :

- Totaux globaux des lignes et colonnes
- Totaux de groupe pour les tableaux emboîtés, empilés et en strates
- Totaux de sous-groupe

Figure 1-8
Sous-totaux, totaux de groupe et totaux de tableau

		Effectif	%
Sexe	Homme	Moins de 25	108 8,8%
		25 à 34	276 22,5%
		35 à 44	309 25,2%
		Sous-total < 45	693 56,4%
		45 à 54	221 18,0%
		55 à 64	136 11,1%
		65 et +	178 14,5%
		Sous-total 45 +	535 43,6%
		Total	1228 100,0%
		Femme	Femme
25 à 34	351 21,9%		
35 à 44	370 23,1%		
Sous-total < 45	855 53,4%		
45 à 54	260 16,3%		
55 à 64	184 11,5%		
65 et +	301 18,8%		
Sous-total 45 +	745 46,6%		
Total	1600 100,0%		
Total	Total		
		25 à 34	627 22,2%
		35 à 44	679 24,0%
		Sous-total < 45	1548 54,7%
		45 à 54	481 17,0%
		55 à 64	320 11,3%
		65 et +	479 16,9%
		Sous-total 45 +	1280 45,3%
		Total	2828 100,0%

Statistiques récapitulatives personnalisées pour les totaux

En ce qui concerne les tableaux qui contiennent des totaux et sous-totaux, vous pouvez obtenir des statistiques récapitulatives différentes des récapitulatifs affichés pour chaque modalité. Vous pouvez, par exemple, afficher les effectifs pour une variable de ligne qualitative ordinale et afficher la moyenne de la statistique de total.

Figure 1-9
Variable qualitative et statistiques récapitulatives dans la même dimension

Avez-vous confiance dans les médias	Beaucoup	Effectif	196
	Moyennement	Effectif	936
	Pas du tout	Effectif	744
	Total	Effectif	1876
		Moyenne	2

Fichier de données d'exemple

La plupart des exemples présentés ici utilisent le fichier de données *survey_sample.sav*. [Pour plus d'informations, reportez-vous à la section Fichiers d'exemple dans l'annexe A sur p. 197.](#) Ce fichier de données est une enquête fictive concernant plusieurs milliers de personnes ; il contient des informations démographiques et des réponses à une série de questions, qui s'étend des tendances politiques aux habitudes des téléspectateurs.

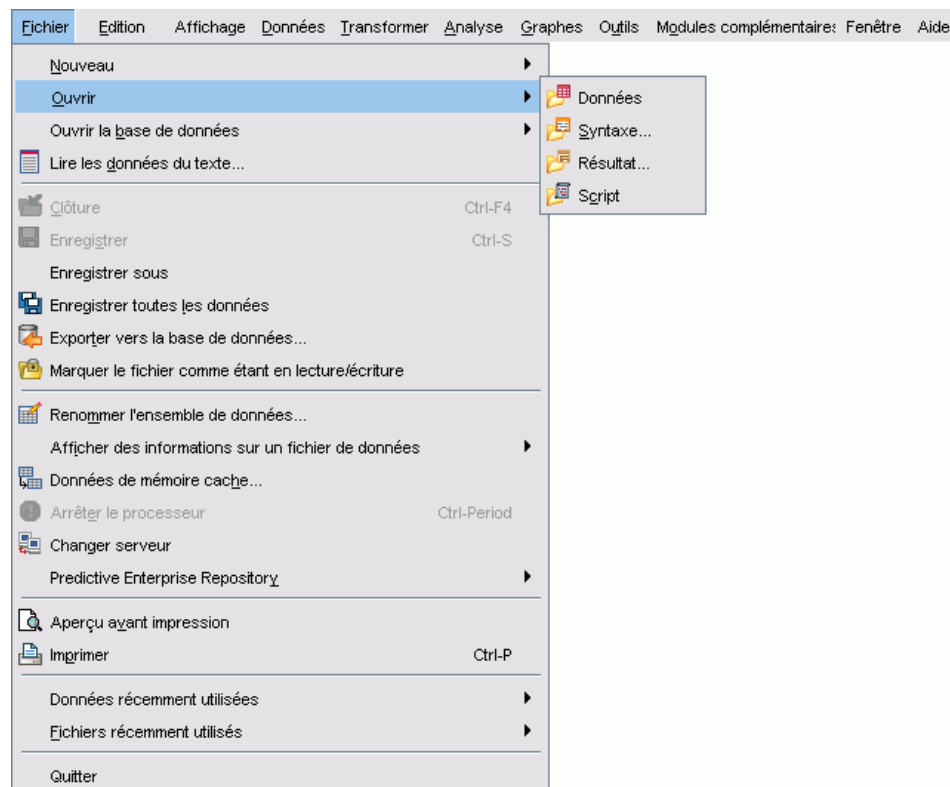
Construction d'un tableau

Pour pouvoir construire un tableau, vous devez disposer de données.

- A partir du menu, sélectionnez :

Fichier > Ouvrir > Données

Figure 1-10
Menu Fichier, Ouvrir



Vous pouvez également utiliser le bouton Ouvrir fichier de la barre d'outils.

Figure 1-11
Bouton Ouvrir fichier de la barre d'outils

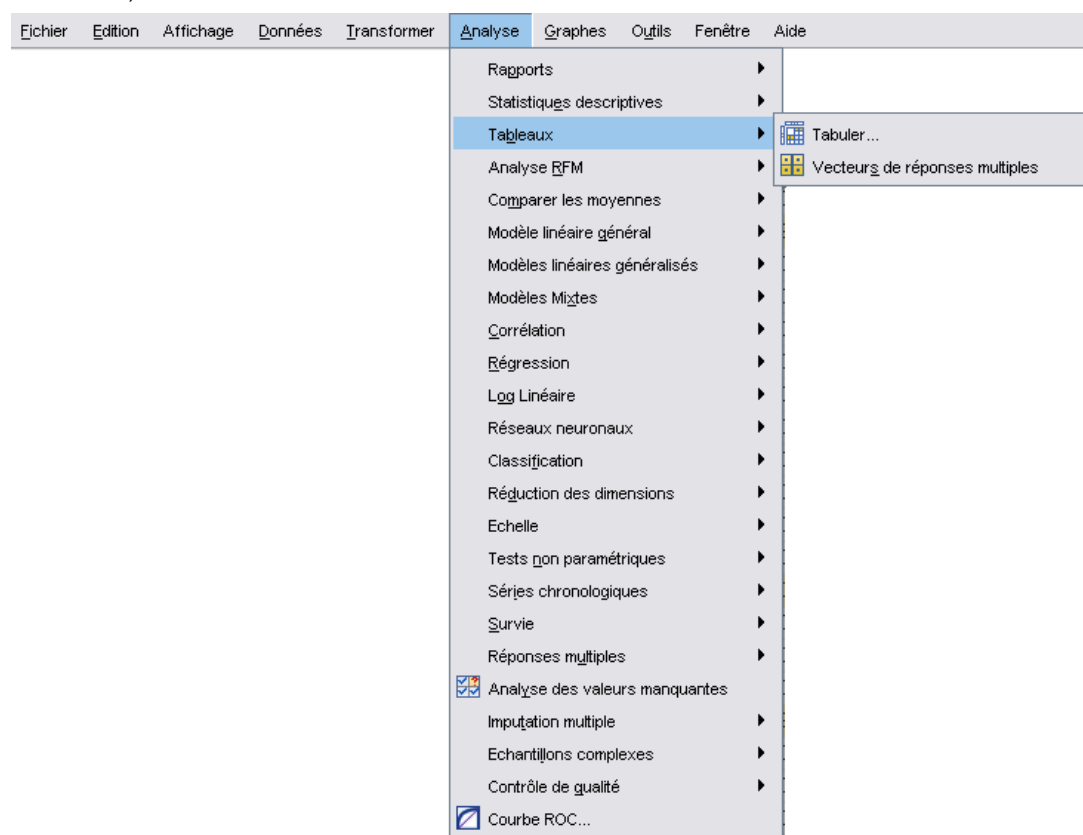


- ▶ Pour utiliser le fichier de données de cet exemple, reportez-vous à [Fichiers d'exemple](#) sur p. 197 pour plus d'informations sur l'emplacement des fichiers de données.
- ▶ Ouvrez le fichier *survey_sample.sav*.

Ouverture du générateur de tableaux personnalisés

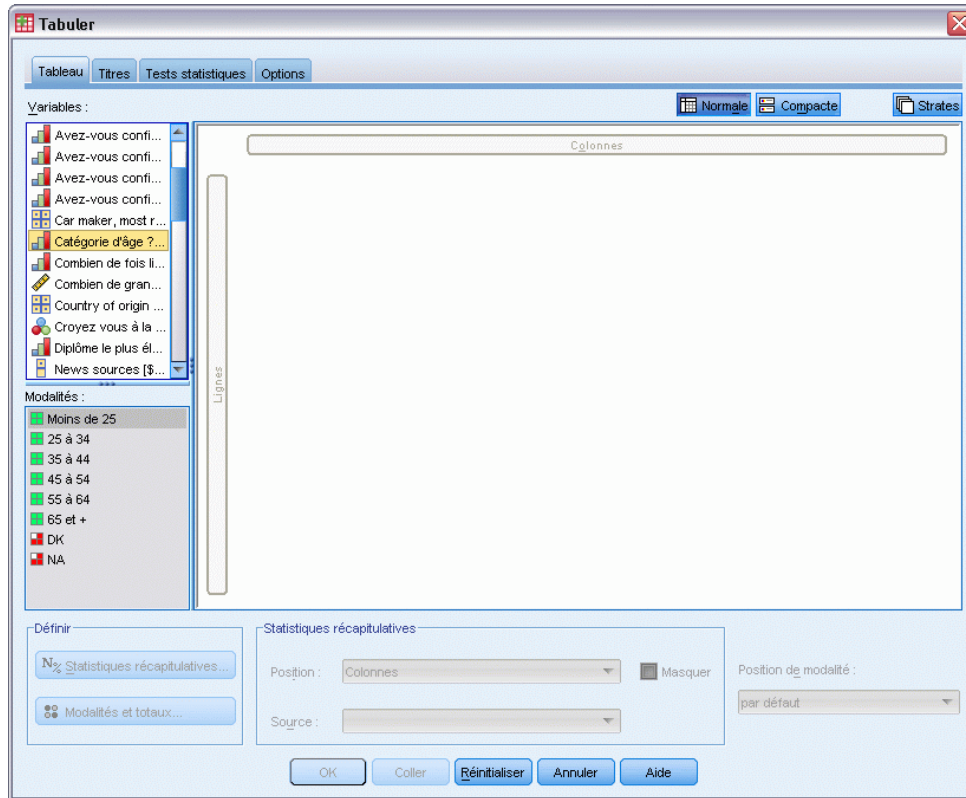
- ▶ Pour ouvrir le générateur de tableaux personnalisés, sélectionnez à partir du menu :
Analyse > Tableaux > Tableaux personnalisés...

Figure 1-12
Menu Analyse, Tableaux



Le générateur de tableaux personnalisés apparaît.

Figure 1-13
Générateur de tableaux personnalisés

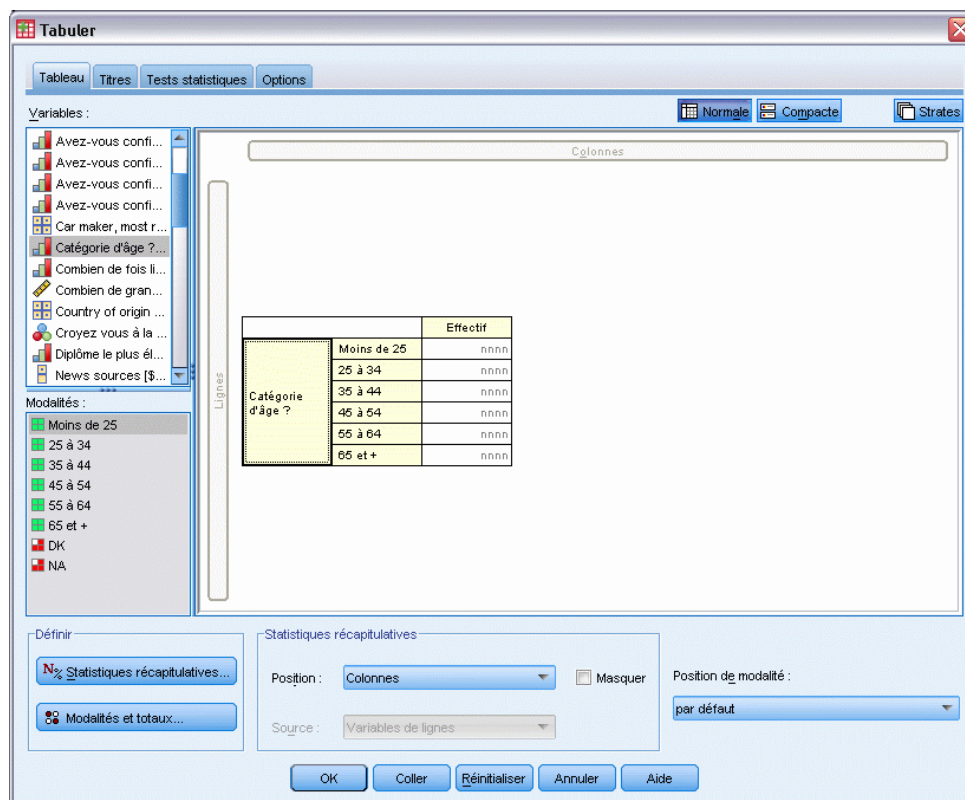


Sélection de variables en ligne et en colonne

Pour créer un tableau, procédez simplement à un « glisser-déplacer » des variables vers l’endroit où vous souhaitez qu’elles apparaissent dans le tableau.

- Sélectionnez *Catégorie d'âge* dans la liste des variables, puis appliquez-lui un « glisser-déplacer » vers la zone Lignes du panneau de canevas.

Figure 1-14
Sélection d'une variable de ligne

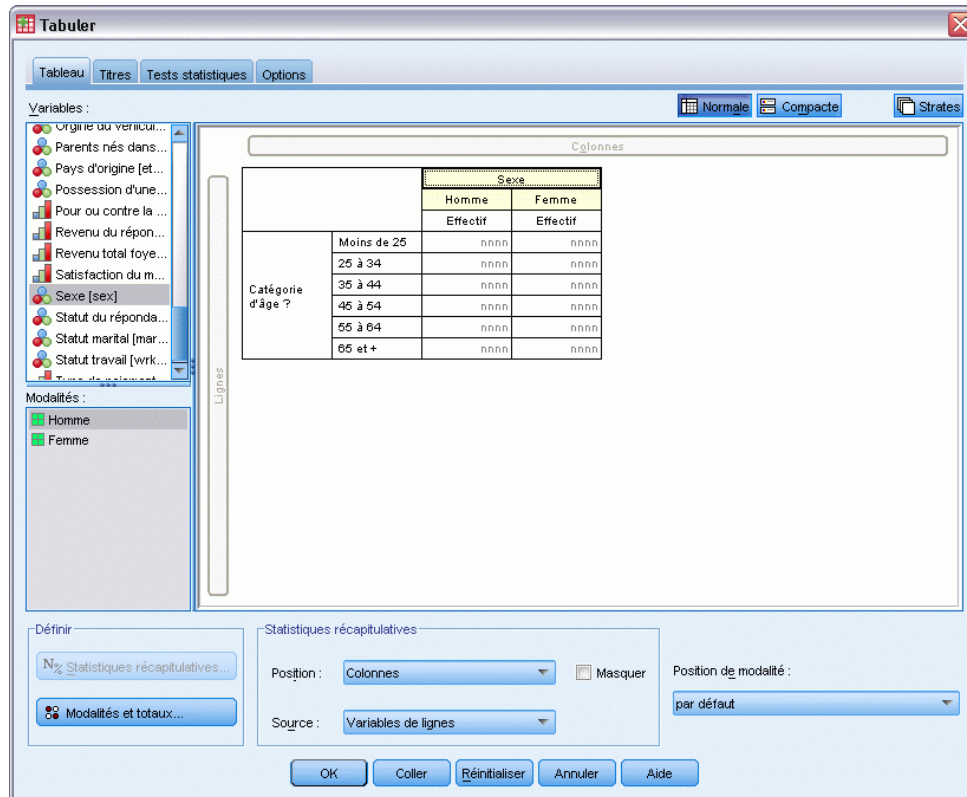


Le panneau de canevas affiche le tableau qui serait créé si cette variable de ligne unique était utilisée.

L'aperçu n'affiche pas les valeurs réelles qui apparaîtraient dans le tableau ; il affiche uniquement la présentation de base du tableau.

- Sélectionnez *sexe* dans la liste des variables, puis appliquez-lui un « glisser-déplacer » vers la zone Colonnes du panneau de canevas (pour accéder à cette variable, vous devrez peut-être parcourir la liste).

Figure 1-15
Sélection d'une variable de colonne



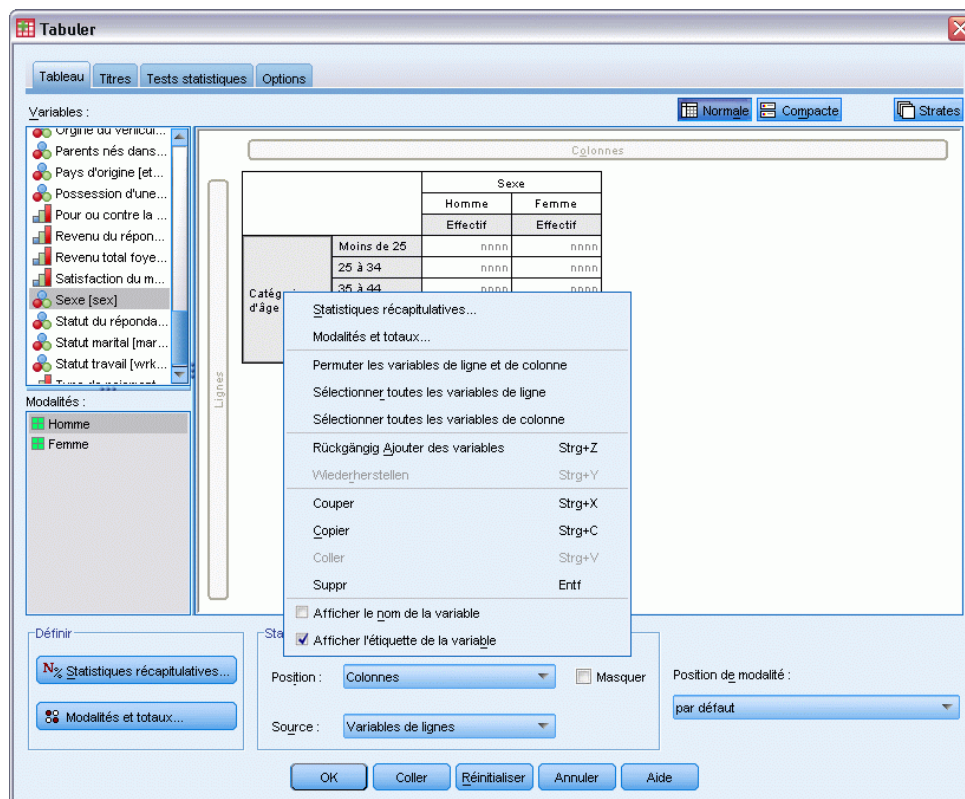
Le panneau de canevas affiche désormais un tableau croisé à deux entrées, à savoir *tranche d'âge* et *sexe*.

Par défaut, les effectifs sont affichés dans les cellules relatives aux variables qualitatives. Vous pouvez également afficher le pourcentage par ligne, par colonne et/ou le pourcentage total.

- Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Tranche d'âge* et sélectionnez *Statistiques récapitulatives* dans le menu contextuel.

Figure 1-16

Menu contextuel des variables qualitatives dans le panneau de canevas

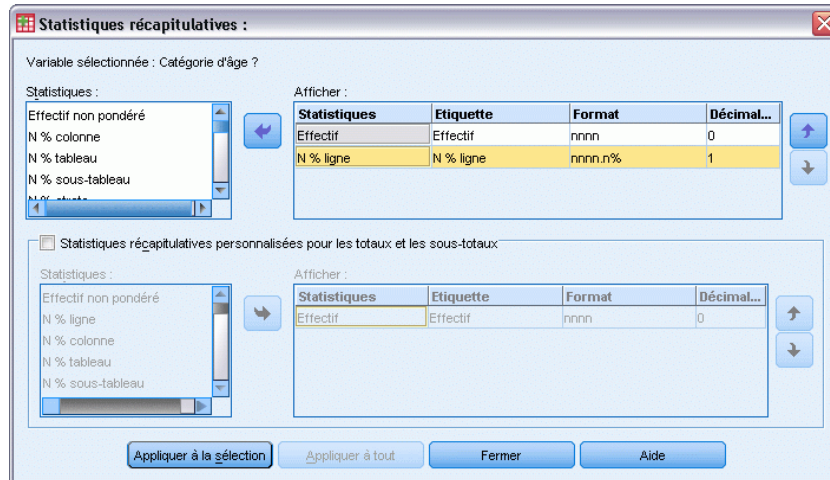


- Dans la boîte de dialogue *Statistiques récapitulatives*, sélectionnez *N % ligne* dans la liste *Statistiques* et cliquez sur le bouton fléché pour l'ajouter à la liste *Afficher*.

Les pourcentages et les effectifs des lignes apparaissent désormais dans le tableau.

Figure 1-17

Boîte de dialogue Statistiques récapitulatives pour les variables qualitatives

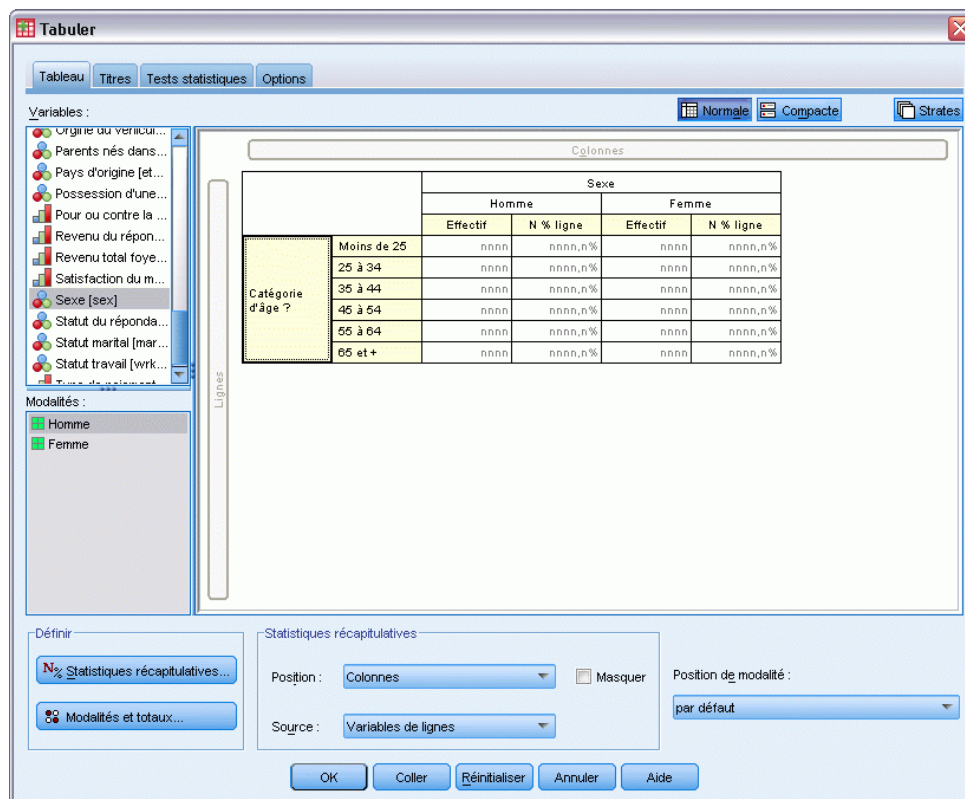


- Cliquez sur Appliquer à la sélection pour enregistrer ces paramètres et revenir au générateur de tableaux.

Le panneau de canevas reflète les modifications apportées : il affiche des colonnes à la fois pour les pourcentages de ligne et les effectifs.

Figure 1-18

Pourcentages de ligne et effectifs affichés dans le panneau de canevas



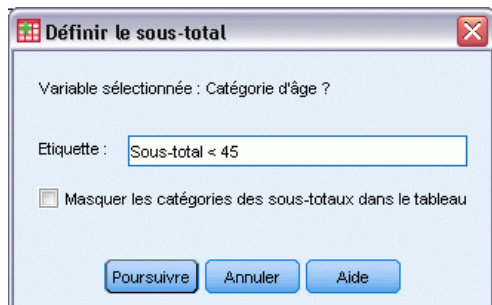
Insertion de totaux et de sous-totaux

Les totaux ne sont pas affichés par défaut dans les tableaux personnalisés, mais vous pouvez facilement ajouter des totaux et des sous-totaux à un tableau.

- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Tranche d'âge* et sélectionnez *Modalités et totaux* dans le menu contextuel.
- ▶ Dans la boîte de dialogue *Modalités et totaux*, sélectionnez 3,00 dans la liste *Valeur(s)*.
- ▶ Cliquez sur *Ajouter un sous-total*.

- Dans la boîte de dialogue Définir le sous-total, saisissez Sous-total <45 puis cliquez sur Poursuivre.

Figure 1-19
Boîte de dialogue Définir le sous-total



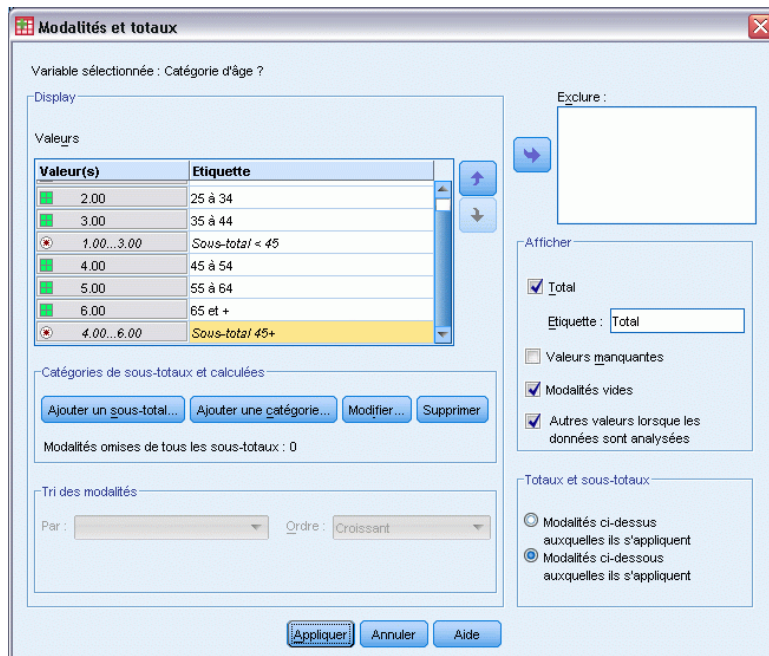
Cette action insère une ligne contenant le sous-total des trois premières tranches d'âge.

- Dans la liste Valeur(s), sélectionnez 6,00.
- Cliquez sur Ajouter un sous-total.
- Dans la boîte de dialogue Définir le sous-total, saisissez Sous-total 45+ puis cliquez sur Poursuivre.

Cette action insère une ligne contenant le sous-total des trois dernières tranches d'âge.

- Pour inclure un total global, cochez la case Total dans le groupe Présenter.

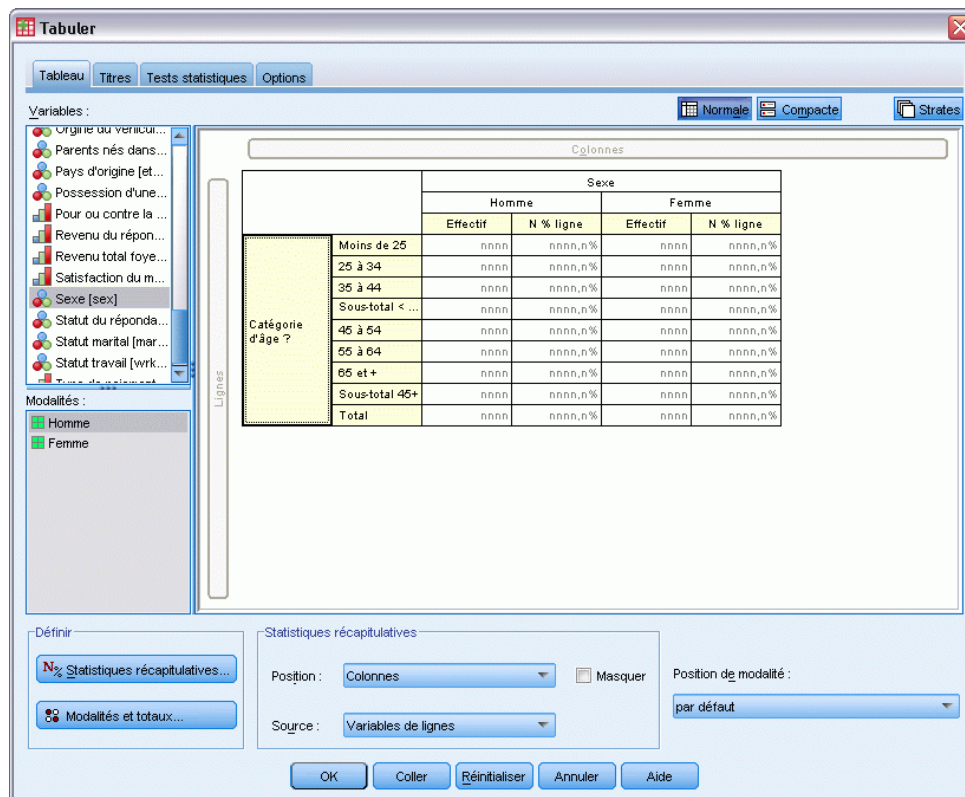
Figure 1-20
Insertion de totaux et de sous-totaux



- Cliquez ensuite sur Appliquer.

L'aperçu du panneau de canevas inclut désormais des lignes pour les deux sous-totaux et le total global.

Figure 1-21
Total et sous-totaux dans le panneau de canevas



- Cliquez sur OK pour générer ce tableau.

Le tableau apparaît dans le Viewer.

Figure 1-22
Tableau croisé avec totaux et sous-totaux

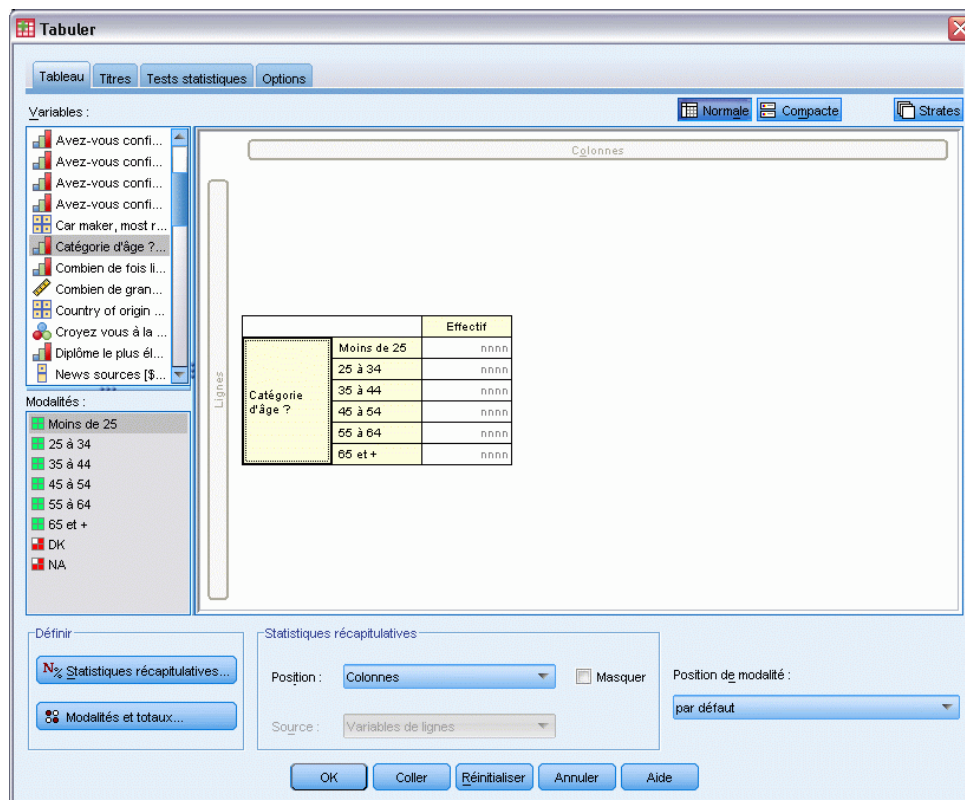
		Sexe			
		Homme		Femme	
		Effectif	N % ligne	Effectif	N % ligne
Catégorie d'âge	Moins de 25	108	44,6%	134	55,4%
	25 à 34	276	44,0%	351	56,0%
	35 à 44	309	45,5%	370	54,5%
	Sous-total < 45	693	44,8%	855	55,2%
	45 à 54	221	45,9%	260	54,1%
	55 à 64	136	42,5%	184	57,5%
	65 et +	178	37,2%	301	62,8%
	Sous-total 45 +	535	41,8%	745	58,2%
	Total	1228	43,4%	1600	56,6%

Récapitulation des variables d'échelle

Un tableau croisé simple de deux variables qualitatives affiche des effectifs ou des pourcentages dans des cellules, mais vous pouvez également afficher des récapitulatifs des variables d'échelle dans les cellules du tableau.

- ▶ Pour rouvrir le générateur de tableaux personnalisés, sélectionnez à partir du menu : Analyse > Tableaux > Tableaux personnalisés...
- ▶ Cliquez sur Restaurer pour effacer toute sélection antérieure.
- ▶ Sélectionnez *Catégorie d'âge* dans la liste des variables, puis appliquez-lui un « glisser-déplacer » vers la zone Lignes du panneau de canevas.

Figure 1-23
Sélection d'une variable de ligne



- Sélectionnez *Heures par jour devant la télé* dans la liste des variables et appliquez-lui un « glisser-déplacer » vers la droite de la variable *Catégorie d'âge* dans la dimension de ligne du tableau.

Figure 1-24

Application d'un « glisser-déplacer » à une variable d'échelle dans la dimension de ligne

The screenshot shows the Tabuler software interface. On the left, a list of variables includes 'Nombre d'heures ...'. The main area displays a pivot table with 'Catégorie d'âge ?' on the vertical axis and 'Nombre d'heures ...' on the horizontal axis. The table data is as follows:

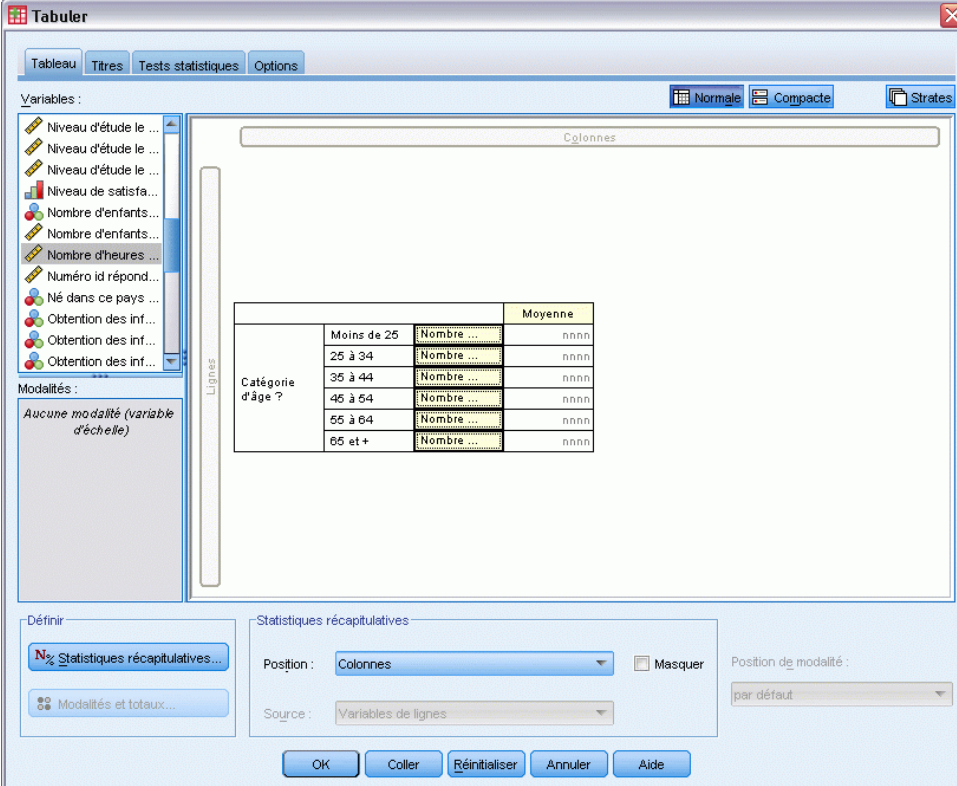
	Moins de 25	Effectif
Moins de 25		nnnn
25 à 34		nnnn
35 à 44		nnnn
45 à 54		nnnn
55 à 64		nnnn
65 et +		nnnn

At the bottom, the 'Statistiques récapitulatives' section shows 'Position' set to 'Colonnes' and 'Source' set to 'Variables de lignes'. The 'Position de modalité' is set to 'par défaut'. Buttons for 'OK', 'Coller', 'Réinitialiser', 'Annuler', and 'Aide' are visible at the bottom.

Désormais, à la place des effectifs par tranche, le tableau affiche le nombre moyen d'heures passées devant la télévision pour chaque tranche d'âge.

Figure 1-25

Variables d'échelle récapitulées dans les cellules du tableau



The screenshot shows the Tabuler software interface. The main window displays a pivot table with the following data:

			Moyenne
Catégorie d'âge ?	Moins de 25	Nombre ...	nnnn
	25 à 34	Nombre ...	nnnn
	35 à 44	Nombre ...	nnnn
	45 à 54	Nombre ...	nnnn
	55 à 64	Nombre ...	nnnn
	65 et +	Nombre ...	nnnn

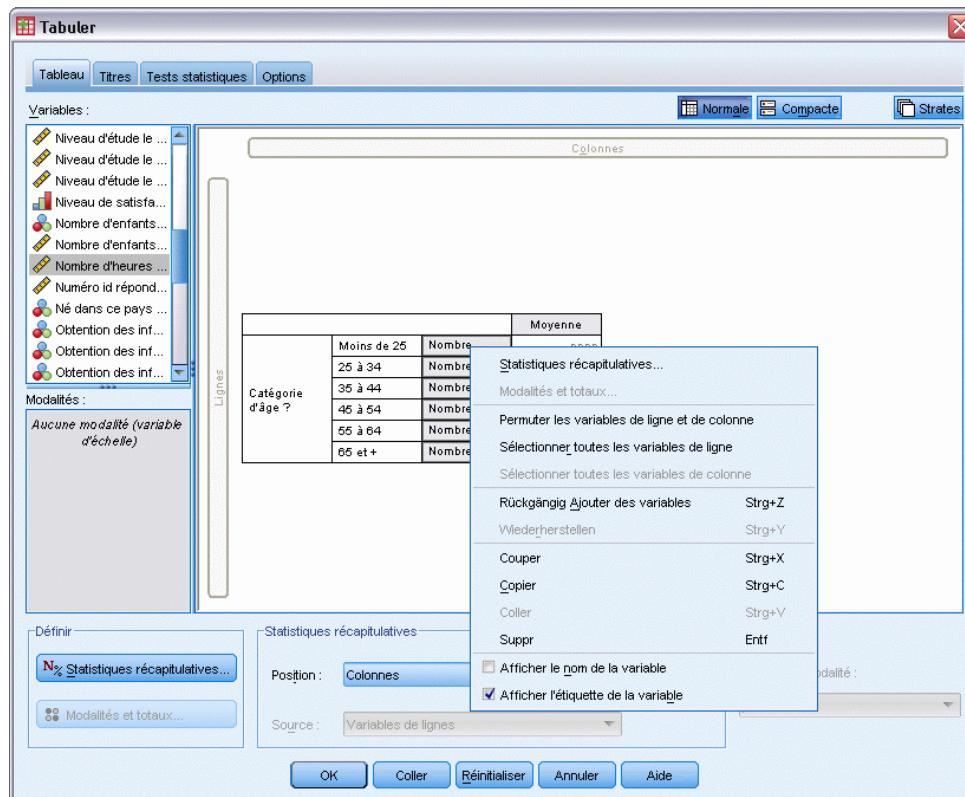
The interface includes a 'Variables' list on the left, a 'Modalités' section, and a 'Statistiques récapitulatives' section at the bottom. The 'Statistiques récapitulatives' section shows 'Position' set to 'Colonnes' and 'Source' set to 'Variables de lignes'. The 'Position de modalité' is set to 'par défaut'.

La moyenne est la statistique récapitulative par défaut pour les variables d'échelle. Vous pouvez ajouter ou modifier les statistiques récapitulatives affichées dans le tableau.

- Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur la variable d'échelle et sélectionnez Statistiques récapitulatives dans le menu contextuel.

Figure 1-26

Menu contextuel des variables d'échelle dans l'aperçu du tableau



- Dans la boîte de dialogue Statistiques récapitulatives, sélectionnez Médiane dans la liste Statistiques et cliquez sur le bouton fléché pour l'ajouter à la liste Afficher.

La moyenne et la médiane apparaissent désormais dans le tableau.

Figure 1-27

Boîte de dialogue Statistiques récapitulatives pour les variables d'échelle



- Cliquez sur Appliquer à la sélection pour enregistrer ces paramètres et revenir au générateur de tableaux.

Le panneau de canevas indique désormais que la moyenne et la médiane apparaîtront dans le tableau.

Figure 1-28

Récapitulatifs des variables d'échelle sous la forme d'une moyenne et d'une médiane dans le panneau de canevas

The screenshot shows the 'Tabuler' software interface. The main window displays a pivot table with the following data:

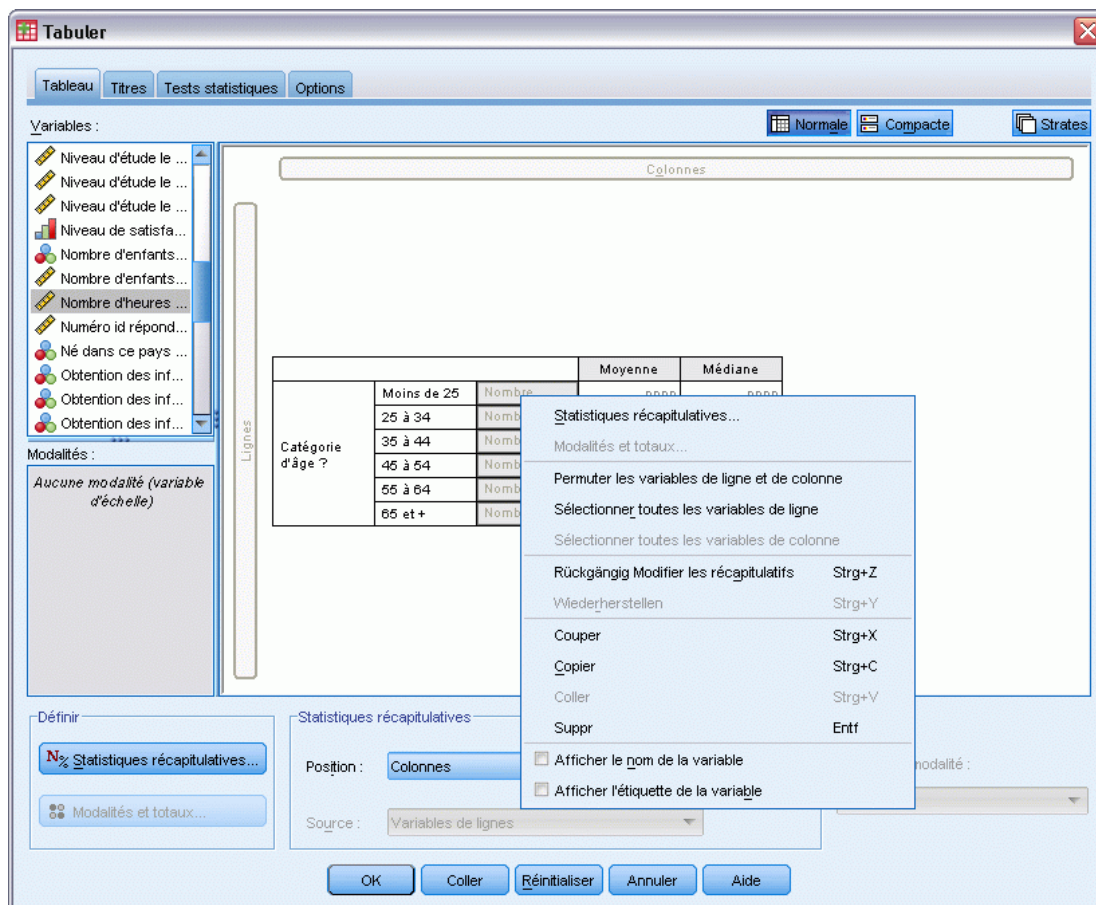
			Moyenne	Médiane
Catégorie d'âge ?	Moins de 25	Nombre ...	nnnn	nnnn
	25 à 34	Nombre ...	nnnn	nnnn
	35 à 44	Nombre ...	nnnn	nnnn
	45 à 54	Nombre ...	nnnn	nnnn
	55 à 64	Nombre ...	nnnn	nnnn
	65 et +	Nombre ...	nnnn	nnnn

The interface includes a 'Variables' list on the left, a 'Modaliés' section with 'Aucune modalité (variable d'échelle)', and a 'Statistiques récapitulatives' section at the bottom with settings for 'Position: Colonnes', 'Source: Variables de lignes', and 'Position de modalité: par défaut'. Buttons for 'OK', 'Coller', 'Réinitialiser', 'Annuler', and 'Aide' are visible at the bottom.

Avant de créer ce tableau, procédons à une petite opération de nettoyage.

- Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur la variable *Heures par jour* et désélectionnez *Afficher l'étiquette de la variables* dans le menu contextuel.

Figure 1-29
Suppression de l'affichage des étiquettes de variable



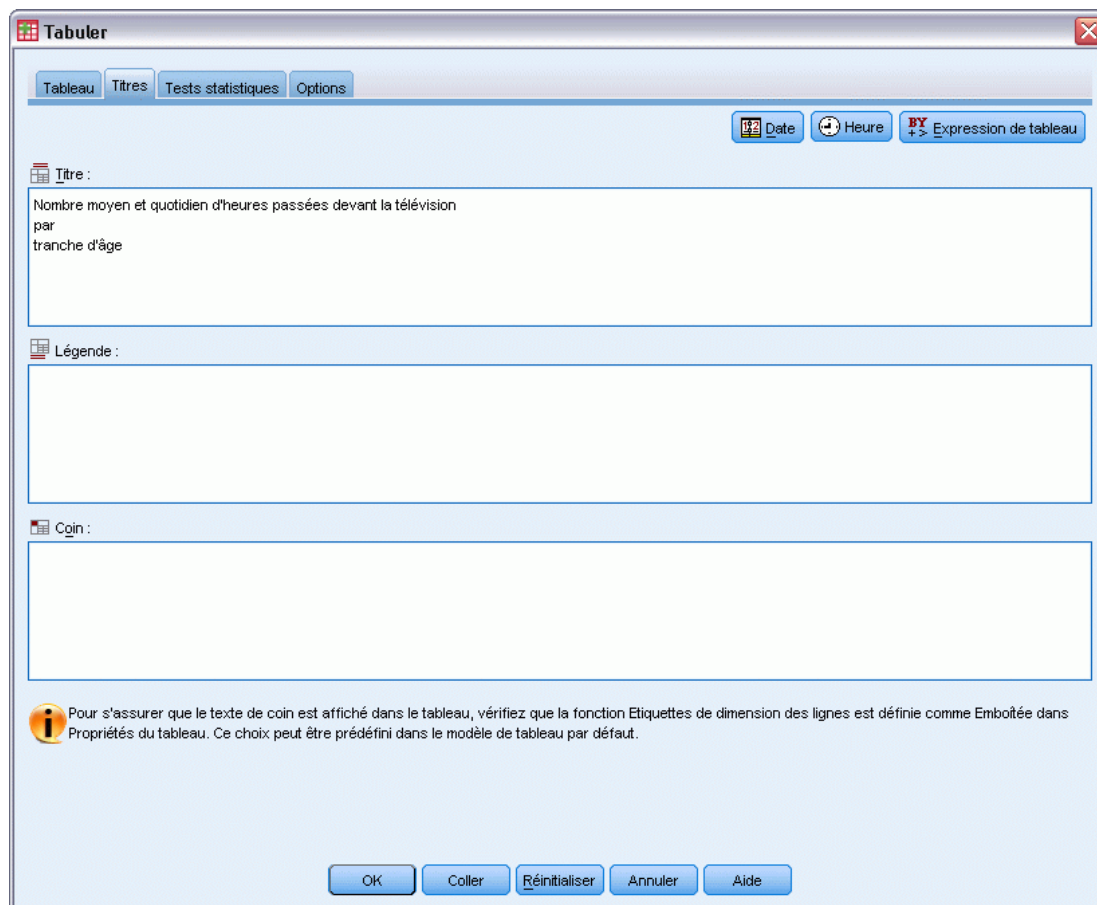
La colonne reste affichée dans l'aperçu du tableau (le texte de l'étiquette de variable est grisé), mais elle n'apparaîtra pas dans le tableau final.

- Dans le générateur de tableaux, cliquez sur l'onglet *Titres*.

- Entrez un titre descriptif pour le tableau, comme Nombre moyen et quotidien d'heures passées devant la télévision par tranche d'âge.

Figure 1-30

Boîte de dialogue *Tabuler*, onglet *Titres*



- Cliquez sur OK pour créer le tableau.

Le tableau apparaît dans la fenêtre du Viewer.

Figure 1-31

Moyenne et médiane relatives au nombre d'heures passées devant la télévision par tranche d'âge

**Nombre moyen et quotidien d'heures passées devant la télévision
par
tranche d'âge**

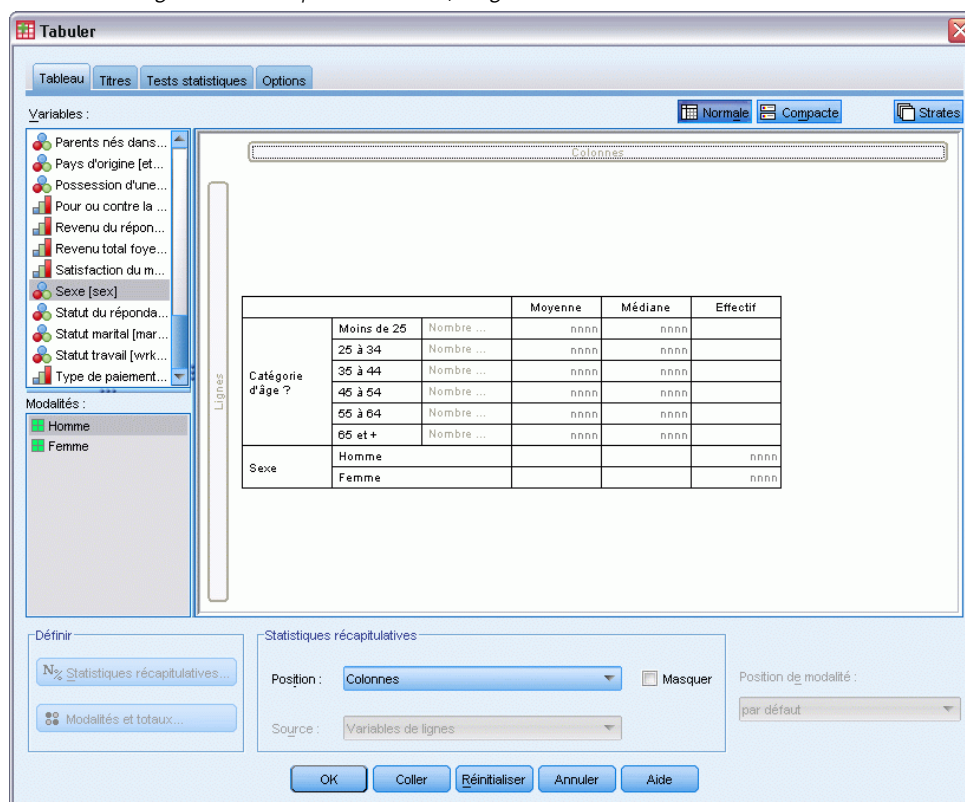
		Moyenne	Médiane
Catégorie d'âge	Moins de 25	2,85	2,00
	25 à 34	2,78	2,00
	35 à 44	2,56	2,00
	45 à 54	2,58	2,00
	55 à 64	3,02	2,50
	65 et +	3,58	3,00

Interface de génération de tableaux

L'option Tableaux personnalisés utilise une interface de génération de tableaux simple, de type « glisser-déplacer », qui vous permet d'obtenir un aperçu de votre tableau lorsque vous sélectionnez des variables et des options. Elle offre également un niveau de flexibilité introuvable dans les boîtes de dialogue classiques, comme la possibilité de modifier la taille de la fenêtre et des panneaux au sein de la fenêtre.

Construction de tableaux

Figure 2-1
Boîte de dialogue Tableaux personnalisés, onglet Tableau



L'onglet Tableau du générateur de tableaux vous permet de sélectionner les variables et les mesures récapitulatives qui apparaîtront dans vos tableaux.

Liste des variables : Les variables du fichier de données apparaissent dans le panneau supérieur gauche de la fenêtre. L'option Tableaux personnalisés opère une distinction entre deux niveaux de mesure différents pour les variables. Par conséquent, les variables sont traitées différemment selon leur niveau de mesure :

Qualitatives : Données ayant un nombre limité de valeurs ou de modalités distinctes (par exemple, sexe ou religion). Les variables qualitatives peuvent être des données chaîne (alphanumérique) ou des variables numériques qui utilisent des codes chiffrés pour représenter les modalités (par exemple, 0 = *Masculin* et 1 = *Féminin*). Elles sont parfois également qualifiées de données qualitatives. Les variables qualitatives peuvent être **nominales** ou **ordinales**

- **Nominal.** Une variable peut être traitée comme étant nominale si ses valeurs représentent des modalités sans classement intrinsèque (par exemple, le service de la société dans lequel travaille un employé). La région, le code postal ou l'appartenance religieuse sont des exemples de variables nominales.
- **Ordinal.** Une variable peut être traitée comme étant ordinale si ses valeurs représentent des modalités associées à un classement intrinsèque (par exemple, des niveaux de satisfaction allant de Très mécontent à Très satisfait). Exemples de variable ordinale : des scores d'attitude représentant le degré de satisfaction ou de confiance, et des scores de classement des préférences.

Échelle. Données mesurées sur une échelle d'intervalle ou de rapport, où les valeurs de données indiquent à la fois l'ordre des valeurs et la distance qui les sépare. Par exemple, un salaire de 58 160 € est supérieur à un salaire de 42 212 € et la distance entre les deux valeurs est de 15 948 €. Ces données sont aussi appelées données quantitatives ou données continues.

Les variables qualitatives définissent des modalités (ligne, colonnes et strates) dans le tableau et la statistique récapitulative par défaut est l'effectif (nombre d'observations dans chaque modalité). Par exemple, le tableau par défaut d'une variable qualitative telle que le sexe afficherait simplement le nombre d'hommes et le nombre de femmes.

Les variables d'échelle sont généralement récapitulées au sein des modalités des variables qualitatives et la statistique récapitulative par défaut est la moyenne. Par exemple, le tableau par défaut présentant la variable revenu au sein des modalités de la variable Sexe afficherait le revenu moyen des hommes et le revenu moyen des femmes.












Vous pouvez également récapituler les variables d'échelle seules, sans utiliser de variable qualitative pour définir des groupes. Ceci est essentiellement utile pour l'**empilement** des récapitulatifs de plusieurs variables d'échelle. [Pour plus d'informations, reportez-vous à la section Empilement de variables sur p. 30.](#)



Vecteurs de réponses multiples

L'option Tableaux personnalisés prend également en charge un type spécifique de variable nommé **vecteur multiréponses**. Les vecteurs multiréponses ne sont pas réellement des variables au sens habituel du terme. En effet, vous ne pouvez pas les voir dans l'éditeur de données et ils ne sont pas reconnus par les autres procédures. Les vecteurs multiréponses utilisent plusieurs variables pour enregistrer les réponses à des questions auxquelles le répondant peut donner plusieurs réponses. Les vecteurs multiréponses sont traités comme des variables qualitatives, et la plupart des actions appliquées à ces dernières peuvent également l'être aux vecteurs multiréponses. [Pour](#)

plus d'informations, reportez-vous à la section Vecteurs de réponses multiples dans le chapitre 11 sur p. 165.

Dans la liste des variables, une icône indique le type de chaque variable.

	Numérique	Chaîne	Date	Heure
Echelle (continue).		n/a		
Ordinal				
Nominal				

Vecteur multiréponses - modalités multiples	
Vecteur multiréponses - dichotomies multiples	

Pour modifier le niveau de mesure d'une variable dans le générateur de tableaux, cliquez avec le bouton droit de la souris sur la variable dans la liste, et sélectionnez Qualitative ou Echelle dans le menu contextuel. Vous pouvez modifier de façon permanente le niveau de mesure d'une variable dans l'affichage des variables de l'éditeur de données. Les variables définies comme étant **nominales** ou **ordinales** sont traitées comme étant qualitatives par l'option Tableaux personnalisés.

Catégories. Lorsque vous sélectionnez une variable qualitative dans la liste des variables, les modalités définies pour la variable apparaissent dans la liste Modalités. Ces modalités apparaissent également dans le panneau de canevas lorsque vous utilisez la variable dans un tableau. Si la variable ne comporte aucune modalité définie, la liste Modalités et le panneau de canevas affichent deux modalités de substitution : *Modalité 1* et *Modalité 2*.

Les modalités définies affichées dans le générateur de tableaux sont basées sur des **étiquettes de valeurs**, des étiquettes descriptives attribuées à différentes valeurs de données (par exemple, les valeurs numériques 0 et 1, avec les étiquettes de valeurs *homme* et *femme*). Vous pouvez définir des étiquettes de valeurs dans l'affichage des variables de l'éditeur de données ou via l'option Définir les propriétés de variable du menu Données de la fenêtre de l'éditeur de données.

Panneau de canevas. Pour construire un tableau, faites glisser des variables vers les lignes et colonnes du panneau de canevas. Le panneau de canevas affiche un aperçu du tableau qui sera créé. Le panneau de canevas n'affiche pas les valeurs de donnée réelles dans les cellules, mais fournit un aperçu relativement précis de la présentation du tableau final. En ce qui concerne les variables qualitatives, le tableau final peut contenir davantage de modalités que l'aperçu si le fichier de données contient des valeurs uniques pour lesquelles aucune étiquette de valeur n'a été définie.

- Le mode Normal affiche toutes les lignes et colonnes qui seront incluses dans le tableau, y compris les lignes et/ou colonnes pour les statistiques récapitulatives et les modalités des variables qualitatives.
- Le mode Compact affiche uniquement les variables qui apparaîtront dans le tableau, sans aperçu des lignes et colonnes qui composeront le tableau.

Règles élémentaires et limites relatives à la construction d'un tableau

- Pour les variables qualitatives, les statistiques récapitulatives sont basées sur la variable la plus interne dans la dimension source.
- La dimension source par défaut (ligne ou colonne) des statistiques pour les variables qualitatives est basée sur l'ordre dans lequel vous avez déplacé les variables vers le panneau de canevas. Par exemple, si vous faites glisser d'abord une variable vers la zone Lignes, la dimension de ligne est la dimension source par défaut des statistiques.
- Les variables d'échelle peuvent être récapitulées uniquement au sein des modalités de la variable la plus interne dans la dimension de ligne ou de colonne. (Vous pouvez positionner la variable d'échelle à tout niveau du tableau, mais elle est récapitulée au niveau le plus interne.)
- Les variables d'échelle ne peuvent pas être récapitulées au sein d'autres variables d'échelle. Vous pouvez empiler les récapitulatifs de plusieurs variables d'échelle ou récapituler des variables d'échelle au sein des modalités de variables qualitatives. Vous ne pouvez pas emboîter une variable d'échelle dans une autre, ou placer une variable d'échelle dans la dimension de ligne et une autre dans la dimension de colonne.
- Si une variable du fichier de travail comprend plus de 12 000 étiquettes de valeurs définies, vous ne pouvez pas utiliser le générateur de tableaux pour créer des tableaux. Si vous n'avez pas besoin d'inclure dans vos tableaux des variables dépassant cette limite, vous pouvez définir et appliquer des groupes de variables qui les excluent. Si vous devez inclure des variables comportant plus de 12 000 étiquettes de valeurs définies, utilisez la syntaxe de commande `CTABLES` pour générer les tableaux.

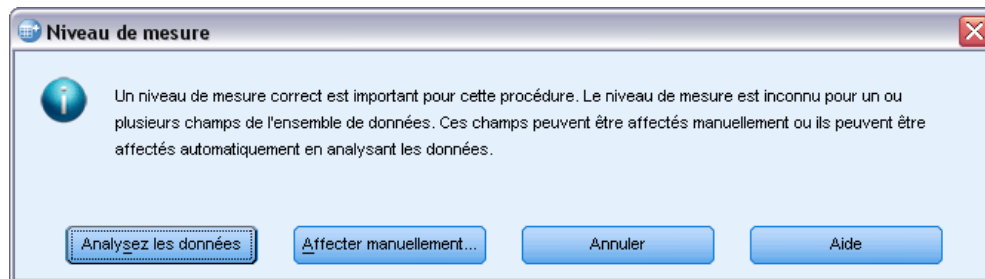
Pour construire un tableau

- ▶ A partir du menu, sélectionnez :
Analyse > Tableaux > Tabuler...
- ▶ Faites glisser des variables vers les zones Lignes et/ou Colonnes du panneau de canevas.
- ▶ Cliquez sur OK pour créer le tableau.
Pour supprimer une variable du panneau de canevas dans le générateur de tableaux :
 - ▶ Dans le panneau de canevas, sélectionnez la variable (en cliquant dessus).
 - ▶ Faites glisser la variable vers l'extérieur du panneau de canevas ou appuyez sur la touche Suppr.
- ▶ Pour modifier le niveau de mesure d'une variable :
 - ▶ Dans la liste des variables, cliquez avec le bouton droit de la souris sur la variable (vous ne pouvez effectuer cette opération que dans la liste des variables, pas dans le canevas).
 - ▶ Dans le menu contextuel, sélectionnez Qualitative ou Echelle.

Champs avec un niveau de mesure inconnu

L'alerte du niveau de mesure apparaît lorsque le niveau de mesure d'une ou plusieurs variables (champs) de l'ensemble de données est inconnu. Le niveau de mesure ayant une incidence sur le calcul des résultats de cette procédure, toutes les variables doivent avoir un niveau de mesure défini.

Figure 2-2
Alerte du niveau de mesure



- **Analysez les données.** Lit les données dans l'ensemble de données actifs et attribue le niveau de mesure par défaut à tous les champs ayant un niveau de mesure inconnu. Si l'ensemble de données est important, cette action peut prendre un certain temps.
- **Attribuer manuellement.** Ouvre une boîte de dialogue qui répertorie tous les champs ayant un niveau de mesure inconnu. Vous pouvez utiliser cette boîte de dialogue pour attribuer un niveau de mesure à ces champs. Vous pouvez également attribuer un niveau de mesure dans l'affichage des variables de l'éditeur de données.

Le niveau de mesure étant important pour cette procédure, vous ne pouvez pas accéder à la boîte de dialogue d'exécution de cette procédure avant que tous les champs n'aient des niveaux de mesure définis.

Empilement de variables

Empiler revient à prendre des tableaux séparés et à les coller ensemble dans le même affichage. Vous pouvez, par exemple, afficher des informations sur les variables *sexe* et *tranche d'âge* dans des sections séparées du même tableau.

Pour empiler des variables

- ▶ Dans la liste des variables, sélectionnez toutes les variables à empiler, puis faites-les glisser simultanément vers les lignes ou colonnes du panneau de canevas.

ou

- ▶ Faites glisser les variables séparément, en positionnant chacune d'elles au-dessus ou au-dessous des variables existantes dans les lignes, ou à droite ou à gauche des variables existantes dans les colonnes.

Figure 2-3
Variables empilées

		Effectif
Sexe	Homme	1232
	Femme	1600
Catégorie d'âge	Moins de 25	242
	25 à 34	627
	35 à 44	679
	45 à 54	481
	55 à 64	320
	65 et +	479

Pour plus d'informations, reportez-vous à la section Empilement de variables qualitatives dans le chapitre 4 sur p. 70.

Emboîtement de variables

L'emboîtement, comme la technique des tableaux croisés, permet de présenter la relation entre deux variables qualitatives ; la seule différence est qu'une variable est emboîtée dans l'autre dans la même dimension. Vous pouvez, par exemple, emboîter la variable *Sexe* dans la variable *Tranche d'âge* dans la dimension de ligne, ce qui a pour effet de présenter le nombre d'hommes et de femmes dans chaque tranche d'âge.

Vous pouvez également emboîter une variable d'échelle dans une variable qualitative. Vous pouvez, par exemple, emboîter la variable *Revenu* dans la variable *Sexe*, ce qui a pour effet de présenter des valeurs de revenu moyen (ou la médiane ou toute autre mesure récapitulative) distinctes pour les hommes et les femmes.

Pour emboîter des variables

- ▶ Faites glisser une variable qualitative vers la zone Lignes ou Colonnes du panneau de canevas.
- ▶ Faites glisser une variable qualitative ou une variable d'échelle vers la gauche ou la droite de la variable qualitative de ligne, ou au-dessus ou au-dessous de la variable qualitative de colonne.

Figure 2-4
Variables qualitatives emboîtées

		Effectif	
Catégorie d'âge	Moins de 25	Homme	108
		Femme	134
	25 à 34	Homme	276
		Femme	351
	35 à 44	Homme	309
		Femme	370
	45 à 54	Homme	221
		Femme	260
	55 à 64	Homme	136
		Femme	184
	65 et +	Homme	178
		Femme	301

Figure 2-5
Variable d'échelle imbriquée dans une variable qualitative

Nombre moyen et quotidien d'heures passées devant la télévision par tranche d'âge

		Moyenne	Médiane
Catégorie d'âge	Moins de 25	2,85	2,00
	25 à 34	2,78	2,00
	35 à 44	2,56	2,00
	45 à 54	2,58	2,00
	55 à 64	3,02	2,50
	65 et +	3,58	3,00

Remarque : Techniquement, le tableau précédent illustre l'emboîtement d'une variable qualitative dans une variable d'échelle, mais les informations fournies par ce tableau sont essentiellement les mêmes que celles issues de l'emboîtement de la variable d'échelle dans la variable qualitative, sans étiquettes redondantes pour la variable d'échelle. (Essayez l'opération inverse et vous comprendrez.)

Pour plus d'informations, reportez-vous à la section [Emboîtement de variables qualitatives dans le chapitre 4 sur p. 73](#).

Remarque : Les tableaux personnalisés ne prennent pas en charge le traitement de fichiers scindés en strates. Pour obtenir le même résultat que les fichiers scindés en strates, placez les variables de fichiers scindés dans les strates emboîtées les plus à l'extérieur du tableau.

Strates

Vous pouvez utiliser des strates pour ajouter une dimension de profondeur à vos tableaux et créer ainsi des « cubes » tridimensionnels. Les strates sont similaires à l'emboîtement ou à l'empilement ; la différence principale réside dans le fait qu'une seule modalité de strate est visible à la fois. Par exemple, l'utilisation de la variable *Tranche d'âge* comme variable de ligne et de la variable *Sexe*

comme variable de strate a pour effet de produire un tableau dans lequel les informations relatives aux hommes et aux femmes apparaissent dans différentes strates.

Pour créer des strates

- ▶ Dans l'onglet Tableau du générateur de tableaux, cliquez sur Strates pour afficher la liste Strates.
- ▶ Faites glisser les variables d'échelle ou qualitatives qui définiront les strates vers la liste Strates.

Figure 2-6
Variables en strates

Sexe du répondant Femme		
Catégories d'âges	Moins de 25	134
	25 à 34	351
	35 à 44	370

Sexe du répondant Homme		
Catégories d'âges	Moins de 25	108
	25 à 34	276
	35 à 44	309
	45 à 54	221
	55 à 64	136
	65 ou plus	178

Vous ne pouvez pas mélanger des variables d'échelle et des variables qualitatives dans la liste Strates. Toutes les variables doivent être du même type. Les vecteurs multiréponses sont traités comme des variables qualitatives pour la liste Strates. Les variables d'échelle dans les strates sont toujours empilées.

Si vous disposez de plusieurs variables qualitatives en strates, les strates peuvent être empilées ou emboîtées.

- L'affichage d'une modalité en strates équivaut à l'empilement. Une strate distincte sera affichée pour chaque modalité de chaque variable de strate. Le nombre total de strates est simplement la *somme* du nombre de modalités pour chaque variable de strate. Par exemple, si vous disposez de trois variables de strate, chacune comportant trois modalités, le tableau sera composé de neuf strates.
- L'affichage de chaque combinaison de modalités en strate équivaut à l'emboîtement ou au croisement de strates. Le nombre total de strates est le *produit* du nombre de modalités pour chaque variable de strate. Par exemple, si vous disposez de trois variables, chacune comportant trois modalités, le tableau sera composé de 27 strates.

Affichage et masquage des noms et/ou étiquettes de variable

Les options suivantes sont disponibles pour l'affichage des noms et étiquettes de variable :

- Affichage des étiquettes de variable uniquement. Pour toute variable sans étiquette de variable définie, le nom de la variable apparaît. Il s'agit du paramètre par défaut.
- Affichage des noms de variable uniquement.

- Affichage des étiquettes et des noms de variable.
- Masquage des noms ou des étiquettes de variable. Bien que la colonne/ligne qui contient l'étiquette ou le nom de variable reste affichée dans l'aperçu du tableau dans le panneau de canevas, cette colonne/ligne n'apparaîtra pas dans le tableau final.

Pour afficher ou masquer les étiquettes ou noms de variable :

- ▶ Dans l'aperçu du tableau dans le panneau de canevas, cliquez avec le bouton droit de la souris sur la variable.
- ▶ Dans le menu contextuel, sélectionnez **Afficher l'étiquette des variables** ou **Afficher le nom des variables** pour activer ou désactiver l'affichage des étiquettes ou des noms. Une coche en regard de la sélection indique que ces éléments seront affichés.

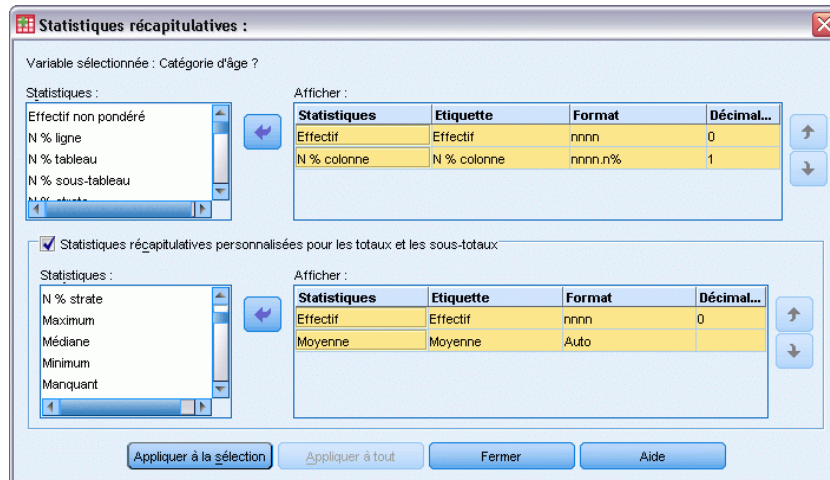
Statistiques récapitulatives

La boîte de dialogue Statistiques récapitulatives vous permet d'effectuer les opérations suivantes :

- Ajouter et supprimer des statistiques récapitulatives dans un tableau.
- Modifier les étiquettes des statistiques.
- Modifier l'ordre des statistiques.
- Modifier le format des statistiques, comme le nombre de décimales.

Figure 2-7

Boîte de dialogue des statistiques récapitulatives et des variables qualitatives



Les statistiques récapitulatives (et autres options) disponibles ici dépendent du niveau de mesure de leur variable source, comme affiché en haut de la boîte de dialogue. La source des statistiques récapitulatives (variable sur laquelle sont basées les statistiques) est déterminée par les éléments suivants :

- **Niveau de mesure.** Si un tableau (ou une section dans un tableau de type empilé) contient une variable d'échelle, les statistiques récapitulatives sont basées sur cette dernière.

- **Ordre de sélection des variables.** La dimension source par défaut (ligne ou colonne) des statistiques pour les variables qualitatives est basée sur l'ordre dans lequel vous avez glissé les variables vers le panneau de canevas. Par exemple, si vous faites glisser d'abord une variable vers la zone Lignes, la dimension de ligne est la dimension source par défaut des statistiques.
- **Emboîtement :** Pour les variables qualitatives, les statistiques récapitulatives sont basées sur la variable la plus interne dans la dimension source.

Un tableau empilé peut comporter plusieurs variables sources de statistiques récapitulatives (variable d'échelle et variable qualitative), mais chaque section du tableau comporte une seule source de statistiques récapitulatives.

Pour modifier la dimension source des statistiques récapitulatives

- ▶ Dans le groupe Statistiques récapitulatives de l'onglet Tableau, sélectionnez la dimension (ligne, colonne ou strate) dans la liste déroulante Source.

Pour contrôler les statistiques récapitulatives affichées dans un tableau

- ▶ Dans le panneau de canevas de l'onglet Tableau, sélectionnez la variable source des statistiques récapitulatives (cliquez dessus).
- ▶ Dans le groupe Définir de l'onglet Tableau, cliquez sur Statistiques récapitulatives.
ou
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur la variable source des statistiques récapitulatives et sélectionnez Statistiques récapitulatives dans le menu contextuel.
- ▶ Sélectionnez les statistiques récapitulatives à inclure dans le tableau. Pour déplacer les statistiques sélectionnées de la liste Statistiques vers la liste Afficher, vous pouvez utiliser la flèche ou faire glisser les statistiques.
- ▶ Cliquez sur les flèches Haut ou Bas pour modifier la position d'affichage de la statistique récapitulative sélectionnée.
- ▶ Dans la liste déroulante Format, sélectionnez un format d'affichage pour la statistique récapitulative sélectionnée.
- ▶ Entrez le nombre de décimales à afficher dans la cellule Décimales pour la statistique récapitulative sélectionnée.
- ▶ Cliquez sur Appliquer à la sélection pour inclure les statistiques récapitulatives sélectionnées pour les variables sélectionnées dans le panneau de canevas.
- ▶ Cliquez sur Appliquer à tout pour inclure les statistiques récapitulatives définies pour toutes les variables empilées du même type dans le panneau de canevas.

Remarque : L'option Appliquer à tout diffère de l'option Appliquer à la sélection uniquement pour les variables empilées du même type déjà présentes dans le panneau de canevas. Dans les deux cas, les statistiques récapitulatives sélectionnées sont automatiquement incluses pour tout ajout de variables empilées du même type au tableau.

Statistiques récapitulatives pour les variables qualitatives

Les statistiques de base disponibles pour les variables qualitatives sont les effectifs et les pourcentages. Vous pouvez également indiquer d'autres statistiques récapitulatives pour les totaux et sous-totaux. Ces autres statistiques récapitulatives incluent les mesures de tendance centrale (comme la moyenne et la médiane) et de dispersion (comme l'écart-type) qui peuvent convenir à certaines variables qualitatives ordinales. [Pour plus d'informations, reportez-vous à la section Autres statistiques récapitulatives en total pour les variables qualitatives sur p. 39.](#)

Effectif. Nombre d'observations dans chaque cellule du tableau ou nombre de réponses pour les vecteurs multiréponses.

Occurrence non pondérée : Nombre non pondéré d'observations dans chaque cellule du tableau.

Pourcentages en colonne : Pourcentages dans chaque colonne. La somme des pourcentages dans chaque colonne d'un sous-tableau (pour les pourcentages simples) est égale à 100 %. Les pourcentages en colonne sont généralement utiles uniquement si vous disposez d'une variable qualitative de *ligne*.

Pourcentages en ligne : Pourcentages dans chaque ligne. La somme des pourcentages dans chaque ligne d'un sous-tableau (pour les pourcentages simples) est égale à 100 %. Les pourcentages en ligne sont généralement utiles uniquement si vous disposez d'une variable qualitative de *colonne*.

Pourcentages en ligne et en colonne de strate : La somme des pourcentages en ligne ou en colonne (pour les pourcentages simples) est égale à 100 % sur l'ensemble des sous-tableaux d'un tableau emboîté. Si le tableau contient des strates, la somme des pourcentages en ligne ou en colonne est égale à 100 % sur l'ensemble des sous-tableaux emboîtés dans chaque strate.

Pourcentages en strate : Pourcentages dans chaque strate. Pour les pourcentages simples, la somme des pourcentages en cellule au sein de la strate actuellement visible est égale à 100 %. Si vous ne disposez d'aucune variable de strate, il s'agit de l'équivalent des pourcentages en tableau.

Pourcentages en tableau : Les pourcentages pour chaque cellule sont basés sur l'intégralité du tableau. Tous les pourcentages en cellule sont basés sur le même nombre total d'observations et leur somme est égale à 100 % (pour les pourcentages simples) dans l'intégralité du tableau.

Pourcentages en sous-tableau : Les pourcentages dans chaque cellule sont basés sur le sous-tableau. Tous les pourcentages en cellule du sous-tableau sont basés sur le même nombre total d'observations et leur somme est égale à 100 % dans le sous-tableau. Dans les tableaux emboîtés, la variable qui précède le niveau d'emboîtement le plus interne définit les sous-tableaux. Par exemple, dans un tableau dans lequel la variable *Situation familiale* est emboîtée dans la variable *Sexe*, elle-même emboîtée dans la variable *Tranche d'âge*, la variable *Sexe* définit les sous-tableaux.

Les pourcentages des vecteurs multiréponses peuvent être basés sur des observations, des réponses ou des occurrences. [Pour plus d'informations, reportez-vous à la section Statistiques récapitulatives pour vecteurs multiréponses sur p. 37.](#)

Tableaux empilés

Pour les calculs de pourcentage, chaque section de tableau définie par une variable d'empilement est traitée comme un tableau distinct. La somme des pourcentages en ligne de strate, en colonne de strate et en tableau est égale à 100 % (pour les pourcentages simples) dans chaque section du tableau empilé. La base de pourcentage pour différents calculs se fonde sur les observations dans chaque section du tableau empilé.

Base de pourcentage

Il est possible de calculer les pourcentages de trois façons différentes, selon le traitement des valeurs manquantes dans la base de calcul :

Pourcentage simple. Les pourcentages sont basés sur le nombre d'observations utilisées dans le tableau et leur somme est toujours égale à 100 %. Si une modalité est exclue du tableau, les observations de cette modalité sont exclues de la base. Les observations avec valeurs manquantes par défaut sont toujours exclues de la base. Les observations avec valeurs manquantes spécifiées par l'utilisateur sont exclues si des modalités manquantes spécifiées par l'utilisateur sont exclues du tableau (par défaut) ou incluses si des modalités manquantes spécifiées par l'utilisateur sont incluses dans le tableau. Tout pourcentage dont le nom ne comporte pas la mention *N valide* ou *N total* est un pourcentage simple.

Pourcentage N total. Les observations avec valeurs manquantes par défaut et valeurs manquantes spécifiées par l'utilisateur sont ajoutées à la base de pourcentage simple. La somme des pourcentages peut être inférieure à 100 %.

Pourcentage N valide. Les observations avec valeurs manquantes spécifiées par l'utilisateur sont supprimées de la base de pourcentage simple, même si des modalités manquantes spécifiées par l'utilisateur sont incluses dans le tableau.

Remarque : Les observations des modalités exclues manuellement, autres que les modalités manquantes spécifiées par l'utilisateur, sont toujours exclues de la base.

Statistiques récapitulatives pour vecteurs multiréponses

Les statistiques récapitulatives supplémentaires suivantes sont disponibles pour les vecteurs multiréponses.

% réponses strate/ligne/colonne : Pourcentage basé sur les réponses.

% réponses strate/ligne/colonne (Base : Effectif). Les réponses sont le numérateur et le total est le dénominateur.

Effectif % de strate/ligne/colonne (Base : Réponses). L'effectif est le numérateur et le total des réponses le dénominateur.

% réponses ligne/colonne de strate : Pourcentage dans les sous-tableaux. Pourcentage basé sur les réponses.

% réponses ligne/colonne de strate (Base : Effectif). Pourcentages dans les sous-tableaux. Les réponses sont le numérateur et le total est le dénominateur.

% réponses ligne/colonne de strate (Base : Réponses). Pourcentages dans les sous-tableaux. L'effectif est le numérateur et le total des réponses le dénominateur.

Réponses. Effectif des réponses.

% réponses tableau/sous-tableau : Pourcentage basé sur les réponses.

% réponses tableau/sous-tableau (Base : Effectif). Les réponses sont le numérateur et le total est le dénominateur.

% effectif de tableau/sous-tableau (Base : Réponses). L'effectif est le numérateur et le total des réponses le dénominateur.

Statistiques récapitulatives pour les variables d'échelle et autres statistiques récapitulatives en total pour les variables qualitatives

Outre les effectifs et les pourcentages disponibles pour les variables qualitatives, les statistiques récapitulatives suivantes sont disponibles pour les variables d'échelle, et sont également disponibles en tant qu'autres statistiques récapitulatives en total et sous-total pour les variables qualitatives. Ces statistiques récapitulatives ne sont pas disponibles pour les vecteurs multiréponses ou les variables chaîne (alphanumériques).

Moyenne. Moyenne arithmétique ; somme divisée par le nombre d'observations.

Médiane. Valeur au-dessus ou au-dessous de laquelle se trouve la moitié des observations ; 50e centile.

Mode. Valeur la plus fréquente. Si deux valeurs sont à égalité, la plus petite valeur est affichée.

Minimum. Plus petite valeur (valeur la moins élevée).

Maximum. Plus grande valeur (valeur la plus élevée).

Manquant : Effectif des valeurs manquantes (valeurs manquantes par défaut et valeurs manquantes spécifiées par l'utilisateur).

Centile. Vous pouvez inclure le 5e, le 25e, le 75e, le 95e et/ou le 99e centile.

Intervalle. Différence entre les valeurs minimale et maximale.

Erreur standard de la moyenne. Mesure du taux de variation de la valeur de la moyenne sur des échantillons provenant de la même distribution. Cette mesure permet de comparer approximativement la moyenne observée avec une valeur hypothétique (autrement dit, vous pouvez conclure que ces deux valeurs sont différentes si le rapport de la différence avec l'erreur standard est inférieur à -2 ou supérieur à $+2$).

Écart type. Mesure de la dispersion des valeurs autour de la moyenne. Dans le cas d'une distribution normale, 68 % des observations se situent à l'intérieur d'un écart-type de la moyenne et 95 % se situent à l'intérieur de deux écarts-types. Par exemple, si l'âge moyen est de 45 ans, avec un écart-type de 10, 95 % des observations se situent entre 25 et 65 dans une distribution normale (racine carrée de la variance).

Somme : Somme des valeurs.

Pourcentage de somme : Pourcentages basés sur des sommes. Disponible pour les lignes et colonnes (au sein de sous-tableaux), l'intégralité des lignes et colonnes (parmi les sous-tableaux), les strates, les sous-tableaux et l'intégralité des tableaux.

N total : Effectif des valeurs non manquantes, des valeurs manquantes spécifiées par l'utilisateur et des valeurs manquantes par défaut. N'inclut pas les observations des modalités exclues manuellement autres que les modalités manquantes spécifiées par l'utilisateur.

N valide : Effectif des valeurs non manquantes. N'inclut pas les observations des modalités exclues manuellement autres que les modalités manquantes spécifiées par l'utilisateur.

Variance. Mesure de la dispersion des valeurs autour de la moyenne, égale à la somme des carrés des écarts par rapport à la moyenne, divisée par le nombre d'observations moins un. La variance est mesurée en unités qui représentent le carré des unités de la variable (le carré de l'écart-type).

Tableaux empilés

Chaque section de tableau définie par une variable d'empilement est traitée comme un tableau distinct et les statistiques récapitulatives sont calculées en conséquence.

Autres statistiques récapitulatives en total pour les variables qualitatives

En ce qui concerne les tableaux de variables qualitatives qui contiennent des totaux et sous-totaux, vous pouvez obtenir des statistiques récapitulatives différentes des récapitulatifs affichés pour chaque modalité. Vous pouvez, par exemple, afficher les effectifs et les pourcentages en colonne pour une variable de ligne qualitative ordinale, et afficher la médiane de la statistique de total.

Pour créer le tableau d'une variable qualitative avec une autre statistique récapitulative en total :

- ▶ A partir du menu, sélectionnez :
Analyse > Tableaux > Tabuler...

Le générateur de tableaux s'ouvre.
- ▶ Faites glisser une variable qualitative vers la zone Lignes ou Colonnes du canevas.
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur la variable et sélectionnez Modalités et totaux dans le menu contextuel.
- ▶ Cochez la case Total, puis cliquez sur Appliquer.
- ▶ Dans le panneau de canevas, cliquez de nouveau avec le bouton droit de la souris sur la variable et sélectionnez Statistiques récapitulatives dans le menu contextuel.
- ▶ Cochez la case Personnaliser les statistiques récapitulatives pour les totaux et les sous-totaux, puis sélectionnez les autres statistiques récapitulatives de votre choix.

Par défaut, toutes les statistiques récapitulatives, y compris les autres statistiques, apparaissent dans la dimension opposée à celle contenant la variable qualitative. Par exemple, si vous disposez d'une variable de ligne qualitative, les statistiques récapitulatives définissent des colonnes dans le tableau :

Figure 2-8
Emplacement par défaut des statistiques récapitulatives

		Effectif	Moyenne
Confiance dans la télévision	Enormément	196	
	Un peu	936	
	Pas du tout	744	
	Total	1876	2,29

Pour afficher les statistiques récapitulatives dans la même dimension que la variable qualitative :

- Dans l'onglet Tableau du générateur de tableaux, dans le groupe Statistiques récapitulatives, sélectionnez la dimension dans la liste déroulante Position.

Par exemple, si la variable qualitative est affichée en lignes, sélectionnez Lignes dans la liste déroulante.

Figure 2-9
Variable qualitative et statistiques récapitulatives dans la même dimension

Avez-vous confiance dans les médias	Beaucoup	Effectif	196
	Moyennement	Effectif	936
	Pas du tout	Effectif	744
	Total	Effectif	1876
		Moyenne	2

Formats d'affichage des statistiques récapitulatives

Les options de format d'affichage suivantes sont disponibles :

nnnn : Format numérique simple.

nnnn% : Le signe % est ajouté à la fin de la valeur.

Auto : Format d'affichage de variable défini, y compris le nombre de décimales.

N=nnnn : Affiche $N=$ devant la valeur. Ceci peut être utile pour les effectifs, le N valide et le N total des tableaux dans lesquels les étiquettes des statistiques récapitulatives ne sont pas affichées.

(nnnn) : Toutes les valeurs sont mises entre parenthèses.

(nnnn)(valeur nég.) : Seules les valeurs négatives sont mises entre parenthèses.

(nnnn%) : Toutes les valeurs sont mises entre parenthèses et le signe % est ajouté à la fin des valeurs.

n,nnn.n : Format virgule. La virgule est utilisée comme séparateur de groupe et le point comme indicateur décimal, quels que soient les paramètres régionaux.

n.nnn,n : Format point. Le point est utilisé comme séparateur de groupe et la virgule comme indicateur décimal, quels que soient les paramètres régionaux.

\$n,nnn.n : Format dollar. Le signe dollar est affiché devant la valeur ; la virgule est utilisée comme séparateur de groupe et le point comme indicateur décimal, quels que soient les paramètres régionaux.

CCA, CCB, CCC, CCD, CCE : Formats monétaires personnalisés (devises). Le format actuellement défini pour chaque devise est affiché dans la liste. Ces formats sont définis dans l'onglet Devise de la boîte de dialogue Options (menu Edition, Options).

Règles générales et limites

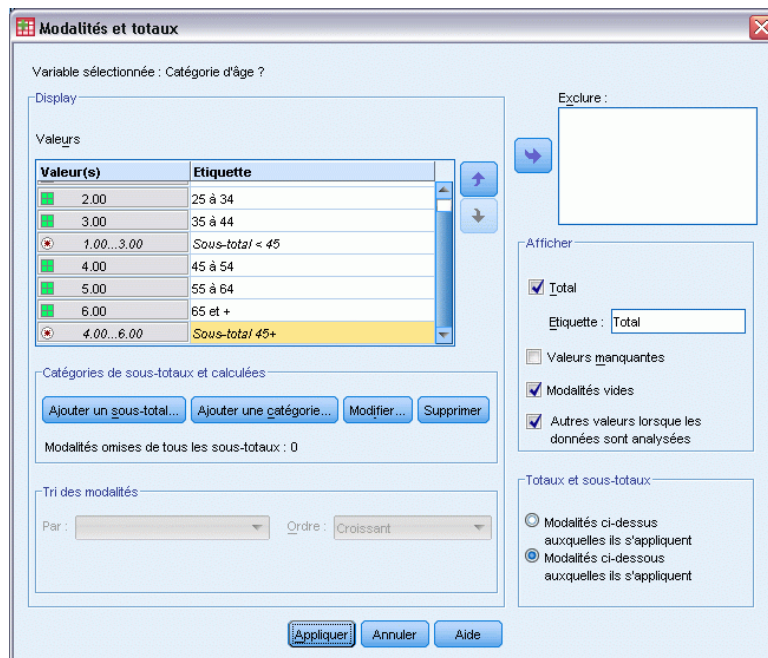
- Sauf pour le format Auto, le nombre de décimales est déterminé par le paramètre de la colonne Décimales.
- Sauf pour les formats virgule, dollar et point, l'indicateur décimal utilisé est celui défini pour les paramètres régionaux dans la zone Options régionales du Panneau de configuration Windows.
- Avec les formats virgule, dollar et point, le séparateur de groupe est soit une virgule, soit un point. Cependant, aucun format d'affichage n'est disponible lors de la création pour afficher un séparateur de groupe sur la base des paramètres régionaux en cours (définis dans la zone Options régionales du Panneau de configuration Windows).

Modalités et totaux

La boîte de dialogue Modalités et totaux vous permet d'effectuer les opérations suivantes :

- Réorganiser et exclure des modalités.
- Insérer des sous-totaux et des totaux.
- Insérer des catégories calculées
- Inclure ou exclure des modalités vides.
- Inclure ou exclure des modalités définies comme contenant des valeurs manquantes.
- Inclure ou exclure des modalités qui ne présentent aucune étiquette de valeur définie.

Figure 2-10
Boîte de dialogue Modalités et totaux



- Cette boîte de dialogue est disponible uniquement pour les variables qualitatives et les vecteurs multiréponses. Elle n'est pas disponible pour les variables d'échelle.
- Dans le cas de la sélection de plusieurs variables avec différentes modalités, vous ne pouvez pas insérer des sous-totaux, ni des catégories calculées, exclure des modalités ou réorganiser manuellement les modalités. Ceci se produit uniquement si vous sélectionnez plusieurs variables dans l'aperçu du canevas et accédez à cette boîte de dialogue pour toutes les variables sélectionnées simultanément. Vous pouvez toujours exécuter ces actions séparément pour chaque variable.
- En ce qui concerne les variables ne présentant aucune étiquette de valeur définie, vous pouvez uniquement trier les modalités et insérer des totaux.

Pour accéder à la boîte de dialogue Modalités et totaux

- Faites glisser une variable qualitative ou un vecteur multiréponses vers le panneau de canevas.

- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur la variable et sélectionnez Modalités et totaux dans le menu contextuel.

ou

- ▶ Dans le panneau de canevas, sélectionnez la variable et cliquez sur Modalités et totaux dans le groupe Définir de l'onglet Tableau.

Vous pouvez également, dans le panneau de canevas, sélectionner plusieurs variables qualitatives dans la même dimension :

- ▶ Dans panneau de canevas, cliquez sur chaque variable tout en maintenant la touche Ctrl enfoncée.

ou

- ▶ Cliquez hors de l'aperçu du tableau dans le panneau de canevas, puis cliquez et faites glisser le pointeur pour sélectionner la zone incluant les variables à sélectionner.

ou

- ▶ Cliquez avec le bouton droit de la souris sur une variable dans une dimension et choisissez Sélectionner toutes les variables de la dimension pour sélectionner toutes les variables de cette dimension.

Pour réorganiser les modalités

Pour réorganiser manuellement les modalités :

- ▶ Sélectionnez une modalité dans la liste (cliquez dessus).
- ▶ Cliquez sur la flèche Haut ou Bas pour déplacer la modalité vers le haut ou vers le bas dans la liste.

ou

- ▶ Cliquez dans la colonne Valeur(s) correspondant à la modalité, puis faites glisser cette dernière vers un autre emplacement.

Pour exclure des modalités

- ▶ Sélectionnez une modalité dans la liste (cliquez dessus).
- ▶ Cliquez sur la flèche en regard de la liste Exclure.

ou

- ▶ Cliquez dans la colonne Valeur(s) correspondant à la catégorie, puis faites glisser cette dernière hors de la liste.

Si vous excluez des modalités, toutes les modalités sans étiquette de valeur définie seront également exclues.

Pour trier les modalités

Vous pouvez trier les modalités par valeur de donnée, étiquette de valeur, effectif de cellule ou statistiques récapitulatives, dans l'ordre croissant ou décroissant.

- ▶ Dans le groupe Trier les modalités, cliquez sur la liste déroulante Par et sélectionnez le critère de tri à utiliser : valeur, étiquette, effectif ou statistique récapitulative (par exemple, la moyenne, la médiane, le mode). Les statistiques récapitulatives disponibles pour le tri dépendent des statistiques récapitulatives que vous avez choisi d'afficher dans le tableau.
- ▶ Cliquez sur la liste déroulante Tri pour sélectionner l'ordre de tri (croissant ou décroissant).
Le tri des modalités n'est pas disponible si vous avez exclu des modalités.

Sous-totaux

- ▶ Dans la liste, sélectionnez la dernière modalité de la plage à inclure dans le sous-total.
- ▶ Cliquez sur Ajouter un sous-total...
- ▶ Dans la boîte de dialogue Définir le sous-total, modifiez le texte de l'étiquette du sous-total, au besoin.
- ▶ Pour afficher uniquement un sous-total et supprimer l'affichage des modalités définissant le sous-total, sélectionnez Masquer les catégories des sous-totaux dans le tableau.
- ▶ Cliquez sur Poursuivre, pour ajouter le sous-total.

Totaux

- ▶ Cochez la case Total. Vous pouvez également modifier le texte de l'étiquette du total.
Si la variable sélectionnée est emboîtée dans une autre, des totaux seront insérés pour chaque sous-tableau.

Position d'affichage pour les totaux et sous-totaux

Les totaux et sous-totaux peuvent être affichés au-dessus ou au-dessous des modalités incluses dans chaque total.

- Si l'option En dessous est sélectionnée dans le groupe Affichage des totaux et des sous-totaux, les totaux apparaissent au-dessus de chaque sous-tableau, et toutes les modalités situées au-dessus, y compris la modalité sélectionnée (mais au-dessous des sous-totaux qui précèdent), sont incluses dans chaque sous-total.
- Si l'option Au-dessus est sélectionnée dans le groupe Affichage des totaux et des sous-totaux, les totaux apparaissent au-dessous de chaque sous-tableau, et toutes les modalités situées au-dessous, y compris la modalité sélectionnée (mais au-dessus des sous-totaux qui précèdent), sont incluses dans chaque sous-total.

Important : Vous devez sélectionner la position d'affichage des sous-totaux avant de définir ces derniers. La modification de la position d'affichage affecte tous les sous-totaux (et non pas uniquement le sous-total sélectionné) ; en outre, elle *modifie les modalités incluses dans les sous-totaux*.

Catégories calculées

Les catégories calculées peuvent être affichées depuis les statistiques récapitulatives; les totaux, les sous-totaux et/ou les constantes. [Pour plus d'informations, reportez-vous à la section Catégories calculées sur p. 46.](#)

Autres statistiques récapitulatives en total et sous-total

Vous pouvez, à l'aide de la boîte de dialogue Statistiques récapitulatives, afficher des statistiques autres que des totaux dans les zones Totaux et Sous-totaux du tableau. [Pour plus d'informations, reportez-vous à la section Statistiques récapitulatives pour les variables qualitatives sur p. 36.](#)

Remarque : Si vous sélectionnez des statistiques totales personnalisées multiples qui sont également dans le corps du tableau et que vous masquez les étiquettes des statistiques, alors les totaux seront retriés dans le même ordre que dans le corps du tableau —et puisque les étiquettes ne sont pas affichées, vous ne pouvez pas savoir ce que chaque statistique totale représente. En général, le choix de statistiques multiples et le fait de masquer les étiquettes des statistiques ne sont pas conseillés.

Totaux, sous-totaux et modalités exclues

Les observations des modalités exclues ne sont pas incluses dans le calcul des totaux.

Valeurs manquantes, modalités vides et valeurs sans étiquette de valeur

Valeurs manquantes : Cette option contrôle l'affichage des valeurs **manquantes spécifiées par l'utilisateur** ou des valeurs définies comme contenant des valeurs manquantes (par exemple, le code 99 représente la réponse "non applicable" pour la grossesse chez les hommes). Par défaut, les valeurs manquantes spécifiées par l'utilisateur sont exclues. Sélectionnez cette option pour inclure, dans les tableaux, les modalités manquantes spécifiées par l'utilisateur. Bien que la variable puisse contenir plusieurs modalités avec valeurs manquantes, l'aperçu du tableau dans le canevas affiche uniquement une modalité générique avec valeurs manquantes. Toutes les modalités manquantes spécifiées par l'utilisateur et définies seront incluses dans le tableau. Les **valeurs manquantes par défaut**(cellules vides pour les variables numériques dans l'éditeur de données) sont toujours exclues.

Modalités vides : Les modalités vides sont des modalités avec étiquettes de valeurs définies, mais qui ne comprennent aucune observation pour un tableau ou un sous-tableau donné. Par défaut, les modalités vides sont incluses dans les tableaux. Désélectionnez (décochez) cette option pour exclure les modalités manquantes du tableau.

Autres valeurs lorsque les données sont analysées : Par défaut, les valeurs de modalité du fichier de données qui ne présentent aucune étiquette de valeur définie sont automatiquement incluses dans les tableaux. Désélectionnez (décochez) cette option pour exclure du tableau les valeurs sans étiquette de valeur définie. Si vous excluez des modalités présentant des étiquettes de valeurs définies, les modalités sans étiquette de valeur définie sont également exclues.

Catégories calculées

En plus d'afficher les résultats agrégés des statistiques récapitulatives, un tableau peut afficher une ou plusieurs catégories calculées de ces résultats agrégés, à partir de valeurs constantes, de totaux et sous-totaux ou d'une combinaison de ces derniers. Les résultats sont affichés sous la forme de catégories ou de post-calculs. Une catégorie calculée agit comme une catégorie dans une variable unique, et présente les similarités et différences suivantes :

- Une catégorie calculée est positionnée comme les autres catégories.
- Une catégorie calculée utilise les mêmes statistiques que les autres catégories.
- Les catégories calculées n'ont pas d'incidence sur les sous-totaux, totaux ou les tests de signification.
- Par défaut, les valeurs des catégories calculées utilisent le même format de statistiques récapitulatives que les autres catégories. Ce format peut être supprimé et remplacé lors de la définition de la catégorie calculée.

Les catégories calculées étant similaires aux sous-totaux, elles peuvent être utilisées pour calculer le total des résultats agrégés. Toutefois, elles présentent les avantages suivants par rapport aux sous-totaux :

- Les catégories calculées peuvent être calculées à partir des résultats d'autres sous-totaux.
- Elles peuvent se chevaucher entre elles, et opérer sur les mêmes (ou une partie des mêmes) catégories.
- Les catégories calculées ne doivent pas inclure de valeurs provenant des autres catégories supérieures ou inférieures à la catégorie calculée. Cela signifie que les catégories calculées ne sont pas exhaustives.
- Les catégories calculées peuvent inclure des valeurs provenant de catégories non-adjacentes.

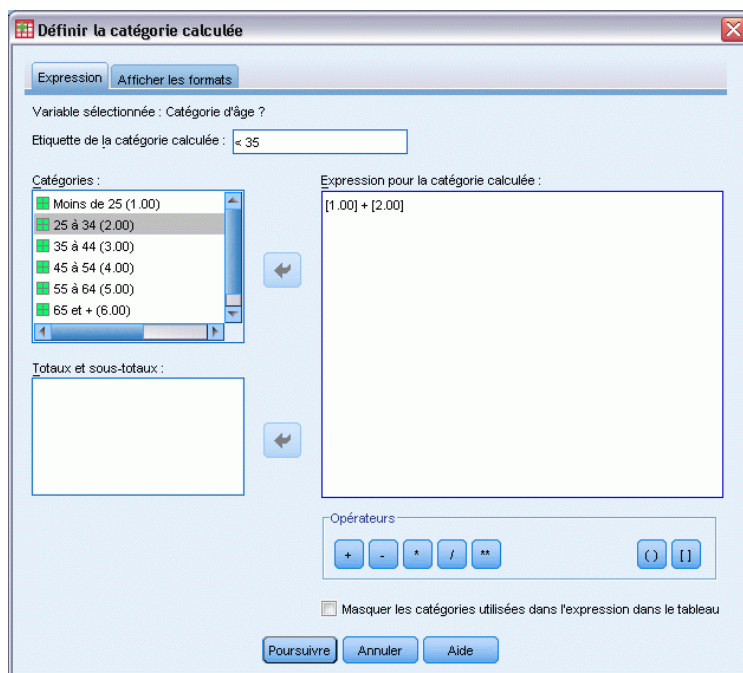
A la différence des totaux et des sous-totaux, le calcul des catégories calculées est effectué à partir des données agrégées et non à partir des données d'origine. Ainsi, il est possible que les valeurs des catégories calculées ne correspondent pas aux résultats des totaux et des sous-totaux. En outre, à cause de l'option qui permet de masquer les catégories source lors de la définition de la catégorie calculée, l'interprétation des sous-totaux peut être rendue difficile dans le tableau des résultats. Si vous utilisez des catégories calculées, nous vous recommandons de spécifier des étiquettes personnalisées pour les sous-totaux.

Pour définir une catégorie calculée

Les catégories calculées sont ajoutées dans la boîte de dialogue Modalités et totaux. Pour des informations supplémentaires sur cette boîte de dialogue, consultez la rubrique [Modalités et totaux](#) sur p. 41.

- ▶ Dans la boîte de dialogue Modalités et totaux, sélectionnez Ajouter une catégorie....

Figure 2-11
Boîte de dialogue Définir la catégorie calculée



- ▶ Dans le champ Etiquette de la catégorie calculée, spécifiez une étiquette pour la catégorie calculée. Vous pouvez faire glisser des catégories depuis la liste Modalités et inclure les étiquettes de ces catégories.
- ▶ Construisez une expression en sélectionnant des catégories et/ou des totaux ou des sous-totaux et en utilisant des opérateurs pour définir les catégories calculées. Vous pouvez aussi saisir des valeurs constantes (par exemple, 500) à inclure dans l'expression.
- ▶ Pour afficher uniquement une catégorie calculée et supprimer l'affichage des modalités la définissant, sélectionnez Masquer les catégories utilisées dans l'expression dans le tableau.
- ▶ Cliquez sur l'onglet Formats d'affichage pour modifier le format d'affichage et le nombre de décimales de la catégorie calculée. [Pour plus d'informations, reportez-vous à la section Formats d'affichage des catégories calculées sur p. 47.](#)
- ▶ Cliquez sur Poursuivre, pour ajouter la catégorie calculée.

Formats d'affichage des catégories calculées

Par défaut, une catégorie calculée utilise le même format d'affichage et le même nombre de décimales que les autres catégories de la variable. Vous pouvez les annuler dans l'onglet Formats d'affichage de la boîte de dialogue Catégorie calculée. L'onglet Formats d'affichage répertorie les statistiques récapitulatives actuelles sur lesquelles est basée la catégorie calculée en plus des formats d'affichage et du nombre de décimales de ces statistiques.

Pour chaque statistique récapitulative, vous pouvez :

- ▶ sélectionner un format d’affichage pour la statistique récapitulative dans la liste déroulante Format. Pour une liste complète des formats d’affichage, reportez-vous à la rubrique [Formats d’affichage des statistiques récapitulatives](#) sur p. 40.
- ▶ Entrez le nombre de décimales à afficher dans la cellule Décimales pour la statistique récapitulative sélectionnée.

Tableaux de variables avec modalités partagées (Comperimeter Tables)

Les enquêtes contiennent souvent de nombreuses questions caractérisées par un ensemble commun de réponses possibles. Vous pouvez utiliser la technique de l’empilement pour afficher ces variables liées dans le même tableau et afficher les modalités de réponses partagées dans les colonnes du tableau.

Pour créer un tableau pour plusieurs variables avec modalités partagées

- ▶ Faites glisser les variables qualitatives de la liste des variables vers la zone Lignes du panneau de canevas. Les variables doivent être **empilées**. [Pour plus d’informations, reportez-vous à la section Empilement de variables sur p. 30.](#)
- ▶ Dans la liste déroulante Emplacement de la modalité, sélectionnez Etiquettes de ligne dans les colonnes.

Figure 2-12

Variables empilées avec modalités de réponses partagées dans les colonnes

	Enormém	Un peu	Pas du
Confiance dans les institutions bancaires et financières	490	1068	306
Confiance dans les multinationales	500	1078	243
Confiance dans l’éducation	511	1055	315
Confiance dans la presse	176	878	808
Confiance dans la médecine	844	864	167
Confiance dans la télévision	196	936	744

[Pour plus d’informations, reportez-vous à la section Tableaux pour variables avec modalités partagées dans le chapitre 7 sur p. 110.](#)

Personnalisation du générateur de tableaux

Contrairement aux boîtes de dialogue standard, vous pouvez modifier la taille du générateur de tableaux de la même manière que vous pouvez modifier la taille d'une fenêtre classique :

- ▶ Cliquez sur un des bords ou des coins du générateur de tableaux, puis faites-le glisser pour augmenter ou diminuer sa taille.

Dans l'onglet Tableau, vous pouvez également modifier la taille de la liste des variables, de la liste Modalités et du panneau de canevas.

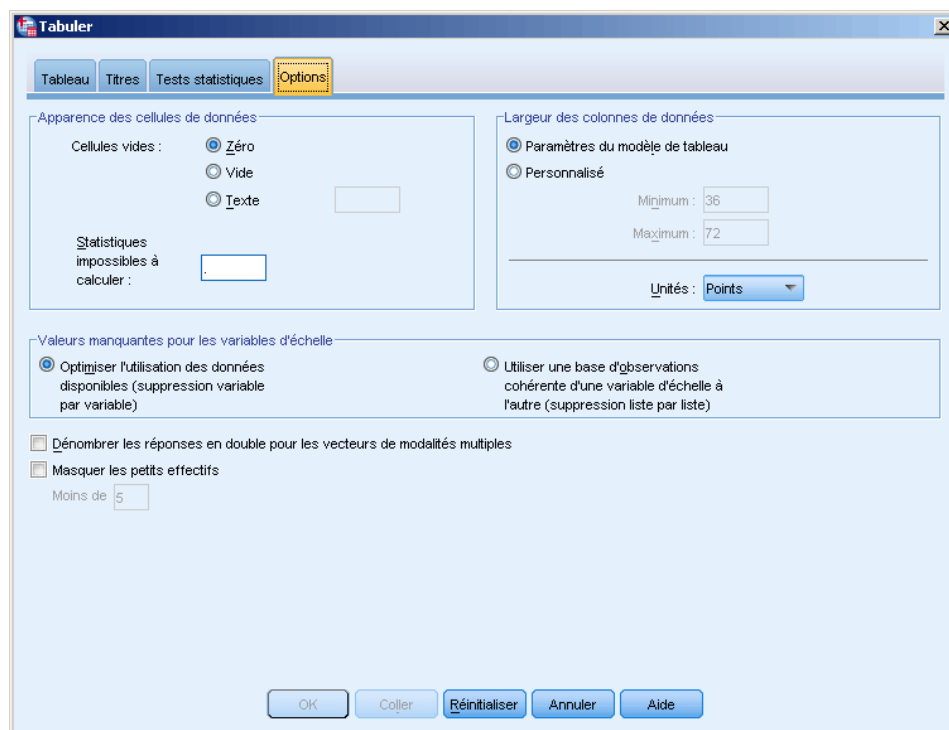
- ▶ Cliquez sur la barre horizontale située entre la liste des variables et la liste Modalités, puis faites-la glisser pour agrandir ou réduire les listes. Le fait de la déplacer vers le bas agrandit la liste des variables et réduit la liste Modalités. Le fait de la déplacer vers le haut produit l'effet inverse.
- ▶ Cliquez sur la barre verticale qui sépare la liste des variables et la liste Modalités du panneau de canevas, puis faites-la glisser afin d'élargir ou de rétrécir les listes. Le canevas est automatiquement redimensionné pour s'adapter à l'espace restant.

Tableaux personnalisés : Onglet Options

L'onglet Options vous permet d'effectuer les opérations suivantes :

- Spécifier ce qui doit apparaître dans les cellules vides et les cellules pour lesquelles aucune statistique ne peut être calculée.
- Contrôler le traitement des valeurs manquantes dans le calcul des statistiques de variable d'échelle.
- Définir la largeur minimum et/ou maximum des colonnes de données.
- Contrôler le traitement des réponses en double dans les vecteurs de modalités multiples.

Figure 2-13
Boîte de dialogue Tabuler, onglet Options



Apparence des cellules de données : Contrôle ce qui doit apparaître dans les cellules vides et les cellules pour lesquelles aucune statistique ne peut être calculée.

- **Cellules vides :** Pour les cellules du tableau qui ne contiennent aucune observation (effectif de cellule de 0), vous pouvez sélectionner l'une des trois options d'affichage : zéro, blanc ou une valeur de texte que vous indiquez. La valeur de texte peut comporter jusqu'à 255 caractères.
- **Statistiques impossibles à calculer :** Texte affiché si une statistique ne peut pas être calculée (par exemple, la moyenne pour une modalité sans observation). La valeur de texte peut comporter jusqu'à 255 caractères. La valeur par défaut est un point (.).

Largeur des colonnes de données : Contrôle la largeur maximum et minimum des colonnes de données. Ce paramètre n'affecte pas la largeur des colonnes pour les étiquettes de ligne.

- **Paramètres du modèle de tableau :** Utilise la spécification de la largeur des colonnes de données du modèle de tableau par défaut en cours. Vous pouvez créer votre propre modèle de tableau par défaut personnalisé qui sera utilisé lors de la création de tableaux, et vous pouvez contrôler, à l'aide d'un modèle de tableau, la largeur des colonnes de données et d'étiquettes de ligne.
- **Personnalisé.** Remplace les paramètres du modèle de tableau par défaut pour la largeur des colonnes de données. Spécifiez la largeur maximum et minimum des colonnes de données pour le tableau, ainsi que l'unité de mesure : points, pouces ou centimètres.

Valeurs manquantes pour les variables d'échelle : Pour les tableaux comportant au moins deux variables d'échelle, cette option contrôle le traitement des données manquantes pour les statistiques de variable d'échelle.

- **Optimiser l'utilisation des données disponibles (suppression variable par variable) :** Toutes les observations avec valeurs valides pour chaque variable d'échelle sont incluses dans les statistiques récapitulatives pour cette variable d'échelle.
- **Utiliser une base d'observations cohérente d'une variable d'échelle à l'autre (suppression liste par liste) :** Les observations avec valeurs manquantes pour toute variable d'échelle dans le tableau sont exclues des statistiques récapitulatives pour toutes les variables d'échelle dans le tableau.

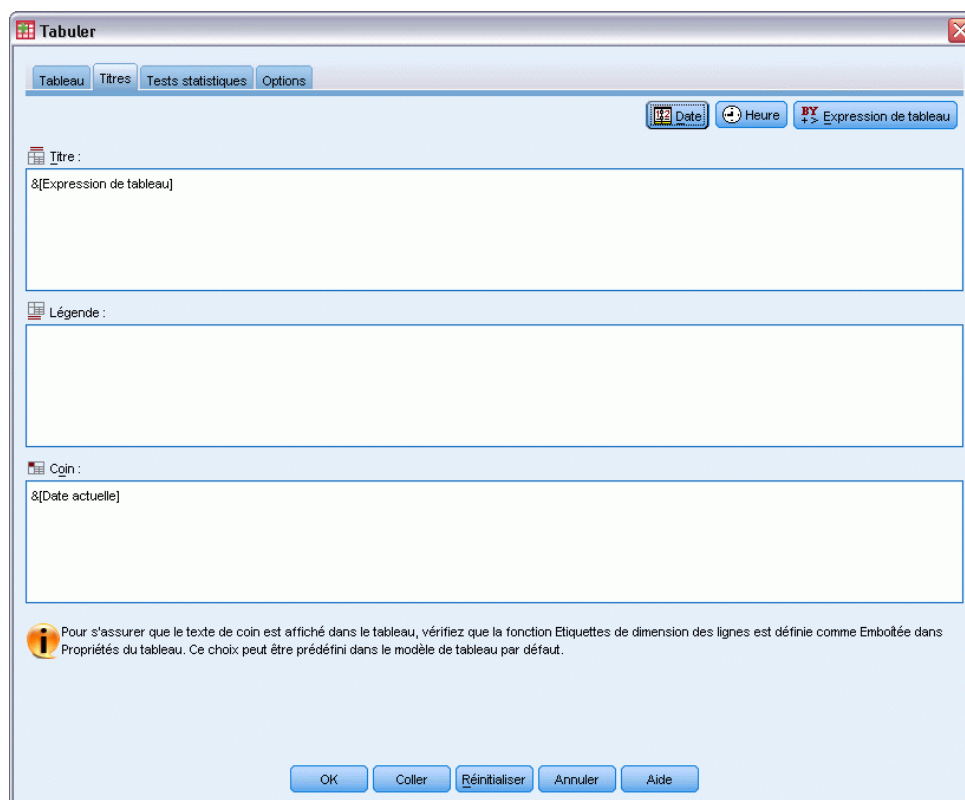
Dénombrer les réponses en double pour les vecteurs de modalités multiples : Une réponse en double est la même réponse pour au moins deux variables dans le vecteur de modalités multiples. Par défaut, les réponses en double ne sont pas comptées, mais vous pouvez parfaitement les inclure dans l'effectif (comme un vecteur de modalités multiples représentant le fabricant des trois dernières voitures achetées par un répondant).

Masquer les petits effectifs. Vous pouvez choisir de masquer les effectifs inférieurs à un entier spécifique. Les valeurs masquées seront affichées en tant que <N, où N est l'entier spécifié. L'entier spécifié doit être supérieur ou égal à 2.

Tableaux personnalisés : Onglet Titres

L'onglet Titres contrôle l'affichage des titres, légendes et étiquettes de coins.

Figure 2-14
Boîte de dialogue Tabuler, onglet Titres



Titre : Texte affiché au-dessus du tableau.

Légende : Texte affiché au-dessous du tableau et au-dessus des notes de bas de page.

Coin : Texte affiché dans le coin supérieur gauche du tableau. Le texte de coin est affiché uniquement si le tableau contient des variables en ligne et si la propriété de l'étiquette de dimension de ligne du tableau pivotant est paramétrée sur Emboîtées. Il ne s'agit *pas* du paramètre de modèle de tableau par défaut.

Vous pouvez inclure les valeurs suivantes, générées automatiquement, dans le titre tableau, la légende ou l'étiquette de coin du tableau :

Date. Année, mois et jour en cours, affichés au format basé sur le paramétrage de vos options régionales Windows.

Heure : Heure, minutes et secondes en cours, au format basé sur le paramétrage de vos options régionales Windows.

Expression de tableau : Variables utilisées et façon dont elles sont utilisées dans le tableau. Si une variable présente une étiquette de variable définie, l'étiquette est affichée. Dans le tableau généré, les symboles suivants indiquent comment les variables sont utilisées :

- + indique des variables empilées.
- > indique un emboîtement.
- **BY** indique un tableau croisé ou des strates.

Tableaux personnalisés : Onglet Statistiques de test

L'onglet Statistiques de test vous permet de demander divers tests de signification pour vos tableaux personnalisés, notamment :

- Tests du Khi-deux pour tester l'indépendance.
- Tests d'égalité des moyennes de colonnes.
- Tests d'égalité des proportions des colonnes.
- Tests de signification pour les vecteurs multiréponses et les sous-totaux. (Pour obtenir des informations sur le test de signification des vecteurs multiréponses, consultez [Test de signification pour les vecteurs de réponses multiples](#) sur p. 174.)

Ces tests ne sont pas disponibles pour les tableaux dans lesquels les étiquettes de modalité sont déplacées hors de leur dimension de tableau par défaut.

Figure 2-15
Boîte de dialogue Tabuler, onglet Test

Comparer les moyennes de colonne (tests-t) : Cette option produit des tests par paire de l'égalité des moyennes de colonnes pour les tableaux dans lesquels au moins une variable qualitative existe dans les colonnes et au moins une variable d'échelle existe dans les lignes. Le tableau doit inclure la moyenne en tant que statistique récapitulative. Vous pouvez indiquer si les valeurs p des tests sont ajustées à l'aide de la méthode Bonferroni. Vous pouvez également spécifier le niveau alpha level du test qui doit être une valeur supérieure à 0 et inférieure à 1. Enfin, alors que la variance pour le test de moyennes est toujours basée uniquement sur les catégories comparées pour les tests de multiréponses, elle peut être estimée uniquement à partir des catégories comparées ou de toutes les catégories pour les variables de catégories ordinaires.

Comparer les proportions de colonne (tests-z) : Cette option produit des tests par paire de l'égalité des proportions des colonnes pour les tableaux dans lesquels au moins une variable qualitative existe à la fois dans les lignes et les colonnes. Le tableau doit inclure les effectifs ou les pourcentages simples en colonne. Vous pouvez indiquer si les valeurs p des tests sont ajustées à l'aide de la méthode Bonferroni. Vous pouvez également indiquer le niveau alpha du test, qui doit être une valeur supérieure à 0 et inférieure à 1.

Identifiez les différences significatives. Si vous sélectionnez Comparer les moyennes de colonne ou Comparer les proportions de colonne, vous pouvez choisir la manière d'indiquer des différences significatives.

- **Dans un tableau distinct.** Les résultats des tests de signification sont affichés dans un tableau distinct. Si deux valeurs sont différentes de manière significative, la cellule correspondant à la valeur la plus élevée affiche une clé identifiant la colonne de la valeur la plus basse. En voici un exemple.

Figure 2-16

Comparaisons des proportions des colonnes dans un tableau séparé

		Statut marital				
		Marié(e) (A)	Veuf/Veuve (B)	Divorcé(e) (C)	Séparé(e) (D)	Jamais marié(e) (E)
Statut professionnel	Travail à plein temps	B		A B	B	B
	Travail à temps partiel					A B C
	Ne travaille pas temporairement					
	Chômage				. ^a	A B
	Retraité(e)	E	A C D E	E		
	Ecolier(e)					A B C
	Femme/Homme au foyer	C E	C E		C E	
Autre						

- **Dans le tableau principal à l'aide d'indices de style APA.** Le tableau principal identifie les différences significatives dans un format de style APA à l'aide d'indices. Si deux valeurs sont différentes de manière significative, elles sont affichées avec des indices *différents*. Ces indices ne sont pas des notes de bas de page. Si cette option est activée, le style de notes en bas de page du modèle de tableau actuel est annulé et les notes en bas de page seront affichées sous la forme d'exposants. En voici un exemple.

Figure 2-17

Comparaison des proportions de colonnes à l'aide d'indices de style APA

		Statut marital				
		Marrié	Veuf/veuve	Divorcé	Séparé	Célibataire
		N % colonne	N % colonne	N % colonne	N % colonne	N % colonne
Statut travail	Plein temps	57.8% _a	15.5% _b	66.1% _c	62.4% _{a,c}	59.1% _{a,c}
	Temps partiel	10.3% _a	7.1% _a	7.8% _a	9.7% _{a,b}	15.4% _b
	Ne travaille pas temporairement	1.7% _a	.7% _a	2.0% _a	1.1% _a	1.7% _a
	Chômage	1.0% _a	1.1% _a	2.2% _{a,b}	.0% ¹	4.8% _b
	Retraité	12.5% _a	53.0% _b	11.9% _a	6.5% _{a,c}	2.6% _c
	Etudiant	.7% _a	.4% _a	1.6% _a	2.2% _{a,b}	9.0% _b
	Au foyer	14.9% _a	19.4% _a	5.6% _b	14.0% _a	5.3% _b
	Autre	1.2% _a	2.8% _a	2.7% _a	4.3% _a	2.1% _a

Pour l'exemple complet sur la façon de créer et d'interpréter ces tableaux, consultez [Comparaison des proportions de colonne](#) sur p. 157.

Remarque : La tableau de style APA comprend une légende qui explique comment interpréter les indices du tableau. Si vous spécifiez une légende sur l'onglet Titres, la légende spécifiée apparaît à la place de la légende d'explication des indices.

Tests d'indépendance (Khi-deux) : Cette option produit un test d'indépendance (Khi-deux) pour les tableaux dans lesquels au moins une variable qualitative existe à la fois dans les lignes et les colonnes. Vous pouvez également indiquer le niveau alpha du test, qui doit être une valeur supérieure à 0 et inférieure à 1.

Utiliser les sous-totaux à la place des catégories de sous-totaux. Si l'option est sélectionnée, chaque sous-total remplace ses catégories pour le test de signification. Dans le cas contraire, seuls les sous-totaux pour lesquels les catégories de sous-totaux sont masquées remplacent leurs catégories pour le test.












Inclure les variables multiréponses. Si l'option est sélectionnée, les tests sont réalisés en utilisant les catégories de vecteurs multiréponses. Dans le cas contraire, les vecteurs multiréponses sont ignorés lors de la réalisation des tests.

Tableaux simples pour variables qualitatives

La plupart des tableaux que vous souhaitez créer incluent probablement au moins une **variable qualitative**. Une variable qualitative est une variable ayant un nombre limité de valeurs ou modalités distinctes (par exemple, sexe ou religion). Les variables qualitatives peuvent être **nominales** ou **ordinales**.

- **Nominal**. Une variable peut être traitée comme étant nominale si ses valeurs représentent des modalités sans classement intrinsèque (par exemple, le service de la société dans lequel travaille un employé). La région, le code postal ou l'appartenance religieuse sont des exemples de variables nominales.
- **Ordinal**. Une variable peut être traitée comme étant ordinale si ses valeurs représentent des modalités associées à un classement intrinsèque (par exemple, des niveaux de satisfaction allant de Très mécontent à Très satisfait). Exemples de variable ordinale : des scores d'attitude représentant le degré de satisfaction ou de confiance, et des scores de classement des préférences.

Dans la liste des variables, une icône indique le type de chaque variable.

	Numérique	Chaîne	Date	Heure
Echelle (continue).		n/a		
Ordinal				
Nominal				

Les tableaux personnalisés ont été optimisés pour être utilisés avec les variables qualitatives présentant des **étiquettes de valeurs** définies. [Pour plus d'informations, reportez-vous à la section Construction de tableaux dans le chapitre 2 sur p. 26.](#)

Fichier de données d'exemple

Les exemples de cette du chapitre utilisent le fichier de données *survey_sample.sav*. [Pour plus d'informations, reportez-vous à la section Fichiers d'exemple dans l'annexe A sur p. 197.](#)

Tous les exemples fournis ici affichent des étiquettes de variable dans des boîtes de dialogue ; ces étiquettes sont triées par ordre alphabétique. Les propriétés d'affichage de la liste des variables sont paramétrées dans l'onglet Général de la boîte de dialogue Options (menu Modifier, Options).

Une seule variable qualitative

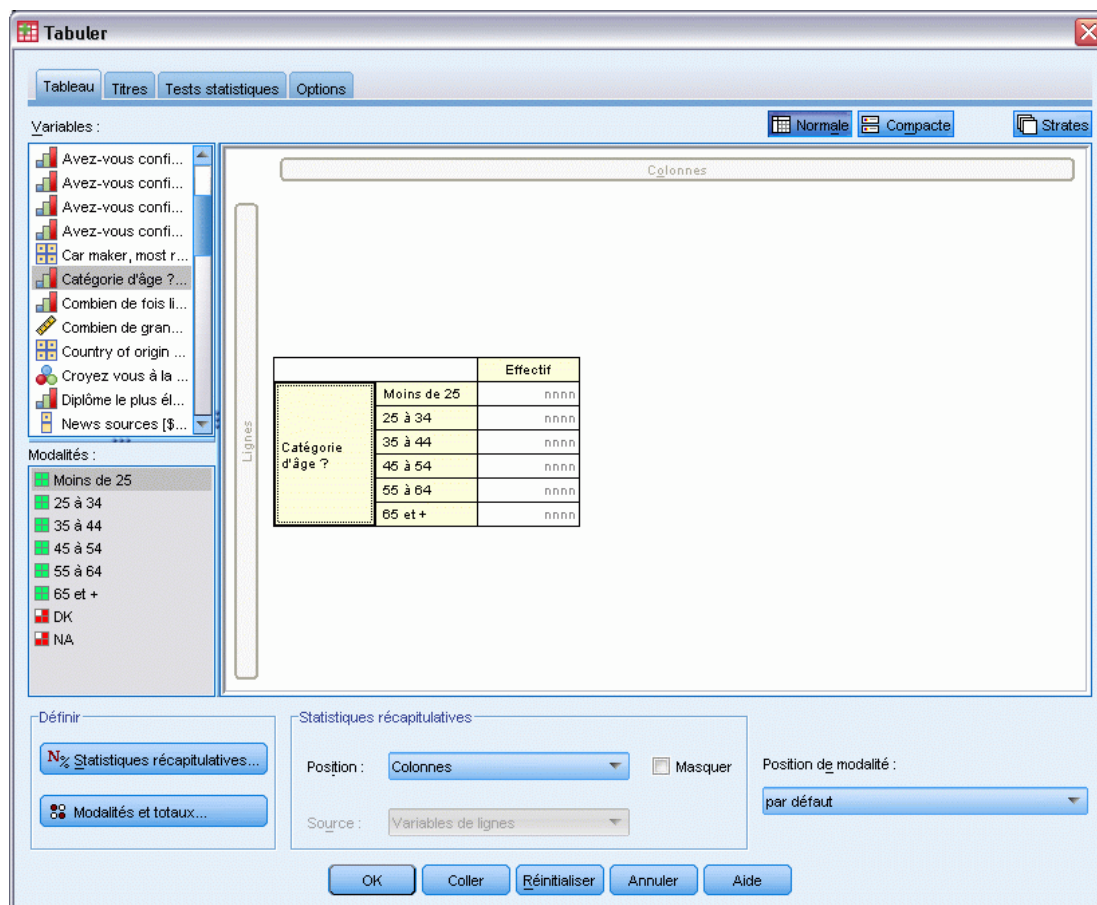
Bien qu'un tableau comportant une seule variable qualitative soit l'un des tableaux les plus simples que vous puissiez créer, il correspondra le plus souvent à ce que vous souhaitez ou à ce dont vous avez besoin.

- ▶ A partir du menu, sélectionnez :
Analyse > Tableaux > Tableaux personnalisés...
- ▶ Dans le générateur de tableaux, faites glisser la variable *Tranche d'âge* de la liste des variables vers la zone Lignes du panneau de canevas.

Un aperçu du tableau apparaît dans le panneau de canevas. L'aperçu n'affiche pas les valeurs de données réelles ; il présente des valeurs de substitution à l'emplacement où des données apparaîtront.

Figure 3-1

Variable qualitative simple en ligne dans l'aperçu du canevas du générateur de tableaux



- ▶ Cliquez sur OK pour créer le tableau.
Le tableau apparaît dans la fenêtre du Viewer.

Figure 3-2
Variable qualitative en lignes unique

		Effectif
Catégories d'âges	Moins de 25	242
	25 à 34	627
	35 à 44	679
	45 à 54	481
	55 à 64	320
	65 ou plus	479

Dans ce tableau simple, l'en-tête de colonne *Effectif* n'est pas vraiment nécessaire et vous pouvez créer le tableau sans cet en-tête.

- ▶ Ouvrez à nouveau le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Dans le groupe Statistiques récapitulatives, sélectionnez (cliquez sur) Masquer.
- ▶ Cliquez sur OK pour créer le tableau.

Figure 3-3
Une seule variable qualitative sans étiquette de colonne de statistiques récapitulatives

Catégories d'âges	Moins de 25	242
	25 à 34	627
	35 à 44	679
	45 à 54	481
	55 à 64	320
	65 ou plus	479

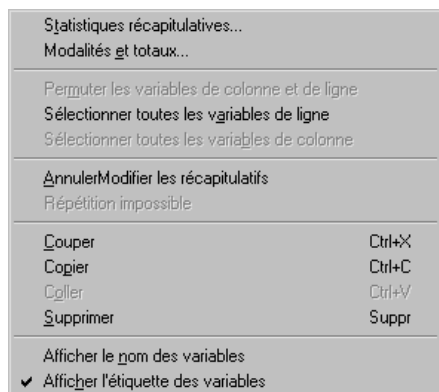
Pourcentages

Outre les effectifs, vous pouvez également afficher les pourcentages. Pour un tableau simple composé d'une seule variable qualitative, si la variable est affichée en lignes, vous souhaitez probablement examiner les pourcentages en colonne. Inversement, pour une variable affichée en colonnes, vous souhaitez probablement examiner les pourcentages en ligne.

- ▶ Ouvrez à nouveau le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Dans le groupe Statistiques récapitulatives, désélectionnez Masquer. Puisque ce tableau présentera deux colonnes, vous souhaitez afficher les étiquettes de colonne afin de savoir ce que représente chaque colonne.
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Tranche d'âge* et sélectionnez Statistiques récapitulatives dans le menu contextuel.

Figure 3-4

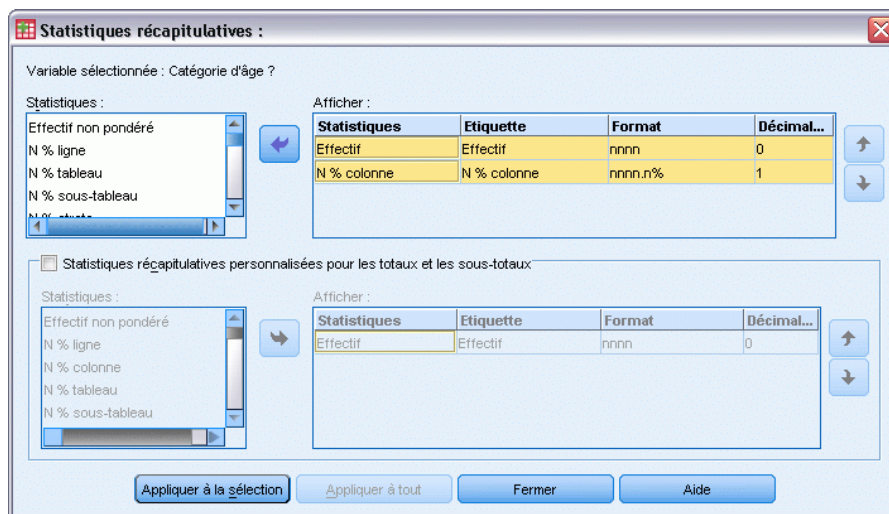
Menu contextuel (accessible via un clic sur le bouton droit de la souris) dans le panneau de canevas



- ▶ Dans la boîte de dialogue Statistiques récapitulatives, sélectionnez N % colonne dans la liste Statistiques et cliquez sur la flèche pour ajouter l'élément à la liste Afficher.
- ▶ Dans la cellule Etiquette de la liste Afficher, supprimez l'étiquette par défaut et tapez Pourcentage.

Figure 3-5

Boîte de dialogue des statistiques récapitulatives et des variables qualitatives



- ▶ Dans le générateur de tableaux, cliquez sur Appliquer à la sélection, puis sur OK pour créer le tableau.

Figure 3-6

Effectifs et pourcentages en colonne

		Effectif	Pourcentage
Catégories d'âges	Moins de 25	242	8,6%
	25 à 34	627	22,2%
	35 à 44	679	24,0%
	45 à 54	481	17,0%
	55 à 64	320	11,3%
	65 ou plus	479	16,9%

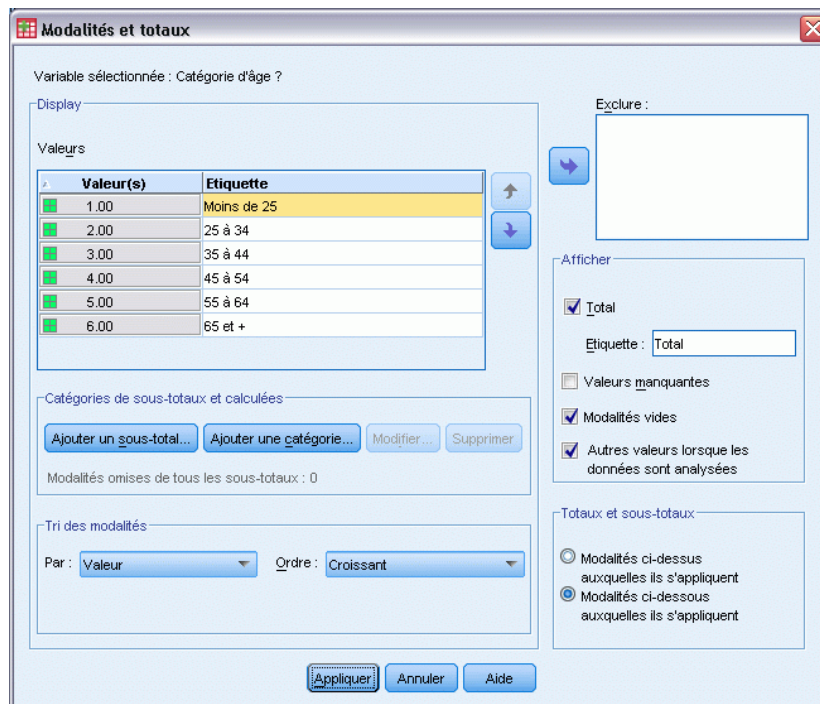
Totaux

Les totaux ne sont pas automatiquement inclus dans les tableaux personnalisés, mais vous pouvez les ajouter facilement.

- ▶ Ouvrez à nouveau le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Tranche d'âge* et sélectionnez Modalités et totaux dans le menu contextuel.
- ▶ Dans la boîte de dialogue Modalités et totaux, sélectionnez Total.

Figure 3-7

Boîte de dialogue Modalités et totaux



- ▶ Dans le générateur de tableaux, cliquez sur Appliquer, puis sur OK pour créer le tableau.

Figure 3-8

Effectifs, pourcentages en colonne et totaux

Catégories		Effectif	Pourcentage
Catégories d'âges	Moins de 25	242	8,6%
	25 à 34	627	22,2%
	35 à 44	679	24,0%
	45 à 54	481	17,0%
	55 à 64	320	11,3%
	65 ou plus	479	16,9%
Total		2828	100,0%

Pour plus d'informations, reportez-vous à la section Totaux et sous-totaux pour les variables qualitatives dans le chapitre 5 sur p. 86.

Tableau croisé

La technique des tableaux croisés est une technique de base qui permet d'examiner la relation entre deux variables qualitatives. Par exemple, en utilisant la variable *Tranche d'âge* comme variable de ligne et la variable *Sexe* comme variable de colonne, vous pouvez créer un tableau croisé bidimensionnel présentant le nombre d'hommes et de femmes dans chaque tranche d'âge.

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Cliquez sur Restaurer pour supprimer les sélections précédentes dans le générateur de tableaux.
- ▶ Dans le générateur de tableaux, faites glisser la variable *Tranche d'âge* de la liste des variables vers la zone Lignes du panneau de canevas.
- ▶ Faites glisser la variable *Sexe* de la liste des variables vers la zone Colonnes du panneau de canevas. (Vous pouvez être amené à parcourir la liste pour accéder à cette variable.)

Figure 3-9

Tableau croisé dans l'aperçu du canevas du générateur de tableaux

The screenshot shows the 'Tabuler' window with a cross-tabulation table. The table structure is as follows:

		Sexe	
		Homme	Femme
Catégorie d'âge ?	Moins de 25	Effectif	Effectif
	nnnn	nnnn	
	25 à 34	nnnn	nnnn
	35 à 44	nnnn	nnnn
	45 à 54	nnnn	nnnn
	55 à 64	nnnn	nnnn
65 et +	nnnn	nnnn	

The interface also shows the following settings:

- Variables:** Sexe [sex] is selected.
- Modalités:** Homme and Femme are checked.
- Statistiques récapitulatives:** Position: Colonnes, Source: Variables de lignes.

- ▶ Cliquez sur OK pour créer le tableau.

Figure 3-10
Tableau croisé avec les variables *Tranche d'âge* et *Sexe*

		Sexe du répondant	
		Homme	Femme
		Effectif	Effectif
Catégories d'âges	Moins de 25	108	134
	25 à 34	276	351
	35 à 44	309	370
	45 à 54	221	260
	55 à 64	136	184
	65 ou plus	178	301

Pourcentages dans des tableaux croisés

Dans un tableau croisé bidimensionnel, les pourcentages en ligne et en colonne peuvent fournir des informations utiles.

- ▶ Ouvrez à nouveau le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Sexe*.

Vous pouvez remarquer que l'option Statistiques récapitulatives est désactivée dans le menu contextuel. En effet, vous ne pouvez sélectionner l'option Statistiques récapitulatives que pour la variable la plus interne dans la dimension source des statistiques. La dimension source par défaut (ligne ou colonne) des statistiques pour les variables qualitatives est basée sur l'ordre dans lequel vous avez glissé les variables vers le panneau de canevas. Dans cet exemple, nous avons d'abord déplacé la variable *Tranche d'âge* vers la dimension de ligne ; étant donné que la dimension de ligne ne comprend pas d'autres variables, la variable *Tranche d'âge* est la variable source des statistiques. Vous pouvez modifier la dimension source des statistiques, mais ceci n'est pas nécessaire dans cet exemple. [Pour plus d'informations, reportez-vous à la section Statistiques récapitulatives dans le chapitre 2 sur p. 34.](#)

- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Tranche d'âge* et sélectionnez Statistiques récapitulatives dans le menu contextuel.
- ▶ Dans la boîte de dialogue Statistiques récapitulatives, sélectionnez N % colonne dans la liste Statistiques et cliquez sur la flèche pour ajouter l'élément à la liste Afficher.
- ▶ Dans la liste Statistiques, sélectionnez N % ligne et cliquez sur la flèche pour ajouter cet élément à la liste Afficher.
- ▶ Dans le générateur de tableaux, cliquez sur Appliquer à la sélection, puis sur OK pour créer le tableau.

Figure 3-11
Tableau croisé avec pourcentages en ligne et en colonne

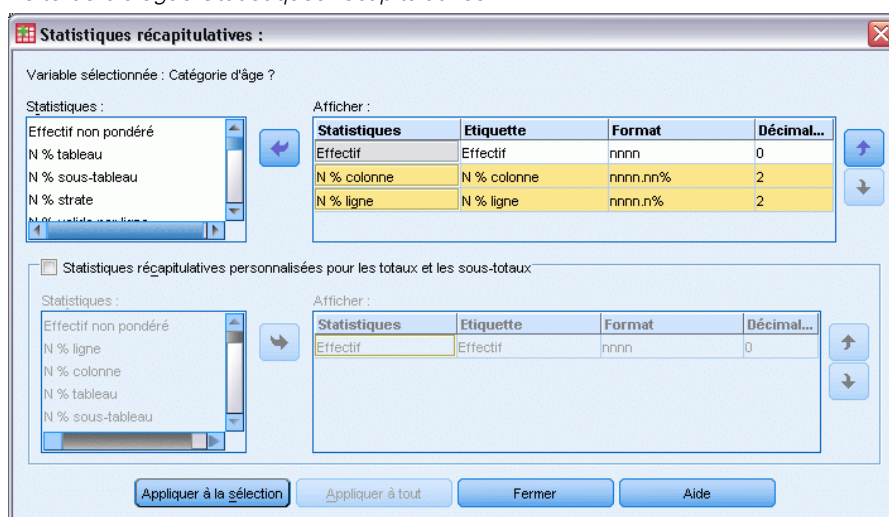
		Sexe du répondant					
		Homme			Femme		
		Effectif	% colonne	% ligne	Effectif	% colonne	% ligne
Catégories d'âges	Moins de 25	108	8,8%	44,6%	134	8,4%	55,4%
	25 à 34	276	22,5%	44,0%	351	21,9%	56,0%
	35 à 44	309	25,2%	45,5%	370	23,1%	54,5%
	45 à 54	221	18,0%	45,9%	260	16,3%	54,1%
	55 à 64	136	11,1%	42,5%	184	11,5%	57,5%
	65 ou plus	178	14,5%	37,2%	301	18,8%	62,8%

Contrôle du format d'affichage

Vous pouvez contrôler le format d'affichage, y compris le nombre de décimales affichées dans les statistiques récapitulatives. Par exemple, les pourcentages sont affichés par défaut avec une décimale et le signe %. Si vous souhaitez que les valeurs des cellules affichent deux décimales sans signe de pourcentage, procédez de la façon suivante.

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Tranche d'âge* et sélectionnez Statistiques récapitulatives dans le menu contextuel.
- ▶ Pour les deux statistiques récapitulatives de pourcentage sélectionnées (N % colonne et N % ligne), sélectionnez nnnn,n dans la liste déroulante Format et tapez 2 dans la cellule Décimales pour les deux statistiques.

Figure 3-12
Boîte de dialogue Statistiques récapitulatives



- ▶ Cliquez sur OK pour créer le tableau.

Figure 3-13
Affichage de cellule formaté pour les pourcentages en ligne et en colonne

		Sexe du répondant					
		Homme			Femme		
		Effectif	% colonne	% ligne	Effectif	% colonne	% ligne
Catégories d'âges	Moins de 25	108	8,79	44,63	134	8,38	55,37
	25 à 34	276	22,48	44,02	351	21,94	55,98
	35 à 44	309	25,16	45,51	370	23,13	54,49
	45 à 54	221	18,00	45,95	260	16,25	54,05
	55 à 64	136	11,07	42,50	184	11,50	57,50
	65 ou plus	178	14,50	37,16	301	18,81	62,84

Totaux marginaux

Dans les tableaux croisés, il est très courant d'afficher des **totaux marginaux**, c'est-à-dire des totaux pour chaque ligne et colonne. Etant donné que ces derniers ne sont pas inclus par défaut dans les tableaux personnalisés, vous devez les ajouter explicitement aux tableaux.

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Cliquez sur Restaurer pour supprimer les sélections précédentes dans le générateur de tableaux.
- ▶ Dans le générateur de tableaux, faites glisser la variable *Tranche d'âge* de la liste des variables vers la zone Lignes du panneau de canevas.
- ▶ Faites glisser la variable *Sexe* de la liste des variables vers la zone Colonnes du panneau de canevas. (Vous pouvez être amené à parcourir la liste pour accéder à cette variable.)
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Tranche d'âge* et sélectionnez Modalités et totaux dans le menu contextuel.
- ▶ Dans la boîte de dialogue Modalités et totaux, sélectionnez Total, puis cliquez sur Appliquer.
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur la variable *Sexe* et sélectionnez Modalités et totaux dans le menu contextuel.
- ▶ Dans la boîte de dialogue Modalités et totaux, sélectionnez Total, puis cliquez sur Appliquer.
- ▶ Dans le groupe Statistiques récapitulatives, sélectionnez (cliquez sur) Masquer. (Puisque vous affichez uniquement les effectifs, vous n'avez pas besoin d'identifier la « statistique » affichée dans les cellules de données du tableau.)
- ▶ Cliquez sur OK pour créer le tableau.

Figure 3-14
Tableau croisé avec totaux marginaux

		Sexe du répondant		
		Homme	Femme	Total
Catégories d'âges	Moins de 25	108	134	242
	25 à 34	276	351	627
	35 à 44	309	370	679
	45 à 54	221	260	481
	55 à 64	136	184	320
	65 ou plus	178	301	479
	Total	1228	1600	2828

Tri et exclusion de modalités

Par défaut, les modalités apparaissent dans l'ordre croissant des valeurs de donnée représentées par les étiquettes de valeurs des modalités. Par exemple, bien que les étiquettes de valeurs de *Moins de 25*, *Entre 25 et 34*, *Entre 35 et 44* etc. apparaissent pour les tranches d'âge, les valeurs de données sous-jacentes réelles sont 1, 2, 3, etc. et contrôlent l'ordre d'affichage par défaut des modalités.

Vous pouvez facilement modifier l'ordre des modalités et exclure des modalités que vous ne souhaitez pas voir apparaître dans le tableau.

Tri des modalités

Vous pouvez réorganiser manuellement ou trier des modalités dans l'ordre croissant ou décroissant pour :

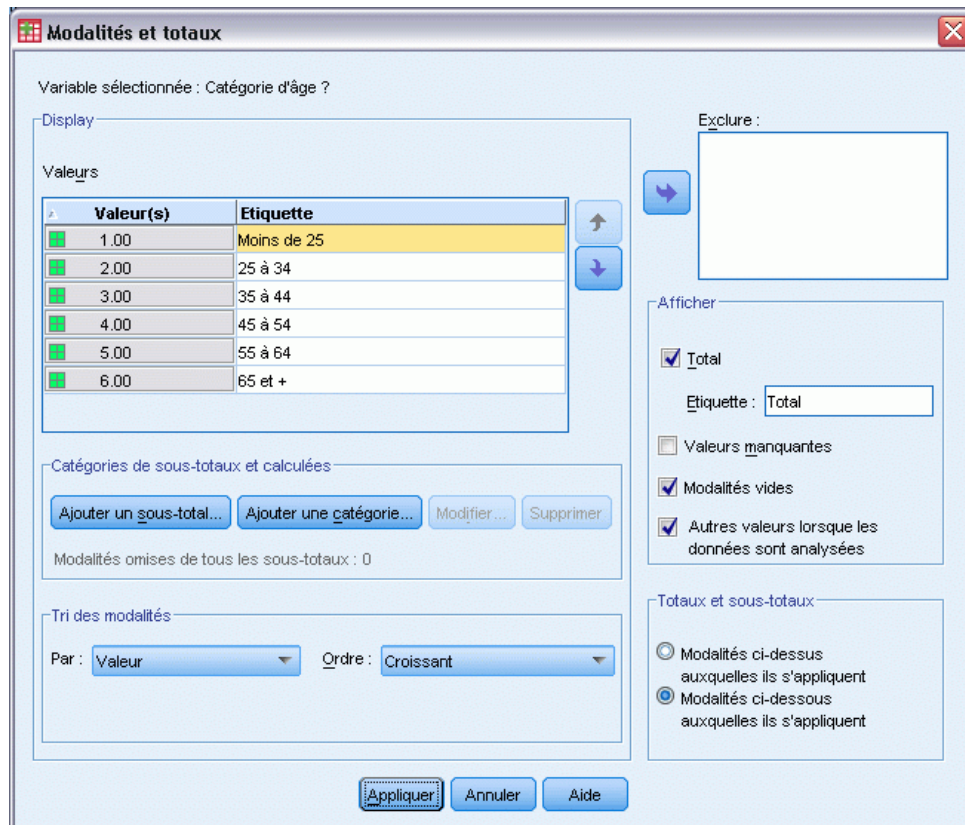
- Les valeurs de données
 - Les étiquettes de valeurs
 - Les effectifs de cellule
 - Statistiques récapitulatives : Les statistiques récapitulatives disponibles pour le tri dépendent des statistiques récapitulatives que vous avez choisi d'afficher dans le tableau.
- Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- Si la variable *Tranche d'âge* n'est pas affichée dans la zone Lignes du panneau de canevas, faites-la glisser pour l'y amener.

- Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Tranche d'âge* et sélectionnez *Modalités et totaux* dans le menu contextuel.

Les valeurs de données et les étiquettes de valeurs associées apparaissent dans l'ordre de tri actuel ; dans ce cas, il s'agit toujours de l'ordre croissant des valeurs de données.

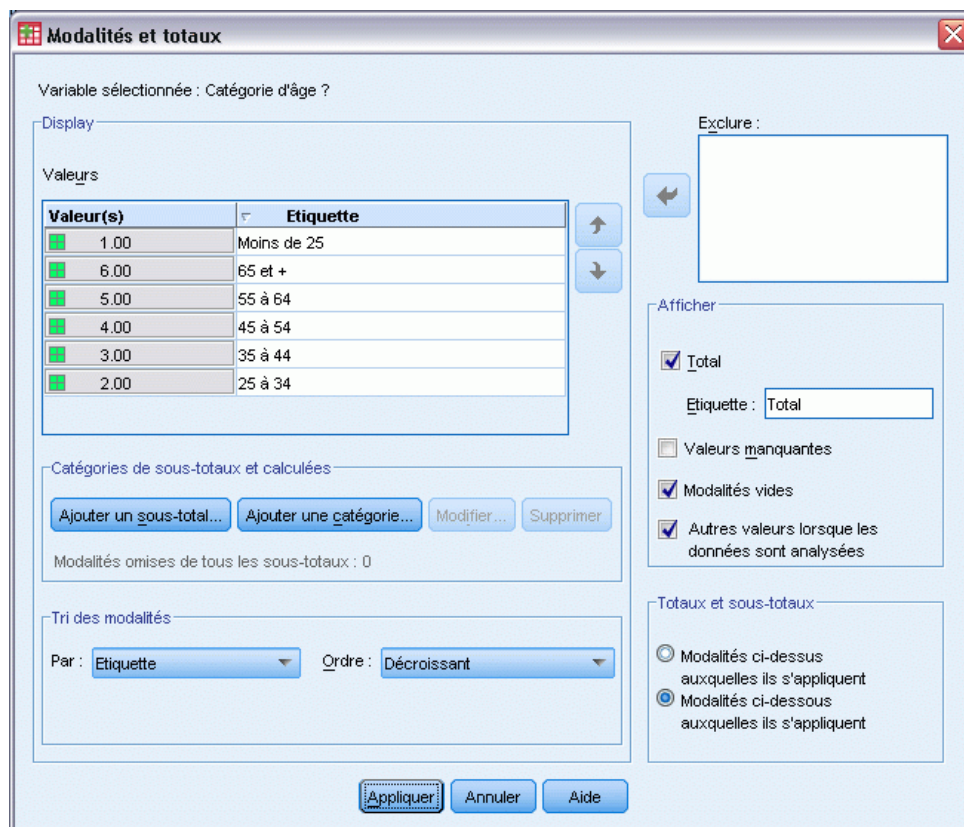
Figure 3-15

Ordre par défaut des modalités - ordre croissant selon les valeurs de données



- Dans le groupe Trier les modalités, sélectionnez *Décroissant* dans la liste déroulante *Ordre*. L'ordre de tri est désormais inversé.
- Dans la liste déroulante *Par*, sélectionnez *Etiquettes*. Les modalités sont désormais triées dans l'ordre alphabétique décroissant des étiquettes de valeurs.

Figure 3-16
Ordre de tri alphabétique décroissant

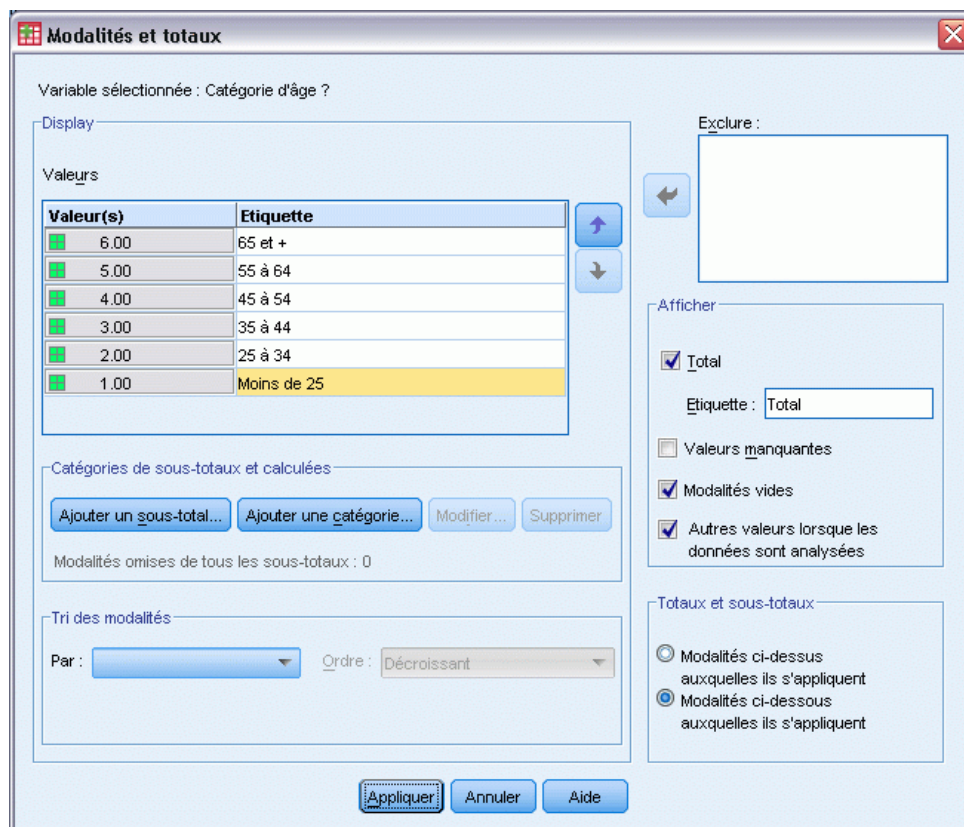


Notez que la modalité portant l'étiquette *Moins de 25* est située au sommet de la liste. Dans le tri alphabétique, les lettres arrivent après les nombres. Puisqu'il s'agit de la seule étiquette qui commence par une lettre et que la liste est triée dans l'ordre décroissant (inverse), cette modalité arrive en début de liste.

Si vous souhaitez qu'une modalité donnée apparaisse à un emplacement différent dans la liste, vous pouvez facilement la déplacer.

- ▶ Dans la liste Etiquette, cliquez sur la modalité portant l'étiquette *Moins de 25*.
- ▶ Cliquez sur la flèche vers le bas, située à droite de la liste. La modalité descend d'une ligne dans la liste.
- ▶ Continuez à cliquer sur la flèche vers le bas jusqu'à ce que la modalité se retrouve au bas de la liste.

Figure 3-17
Modalités organisées manuellement



Exclusion de modalités

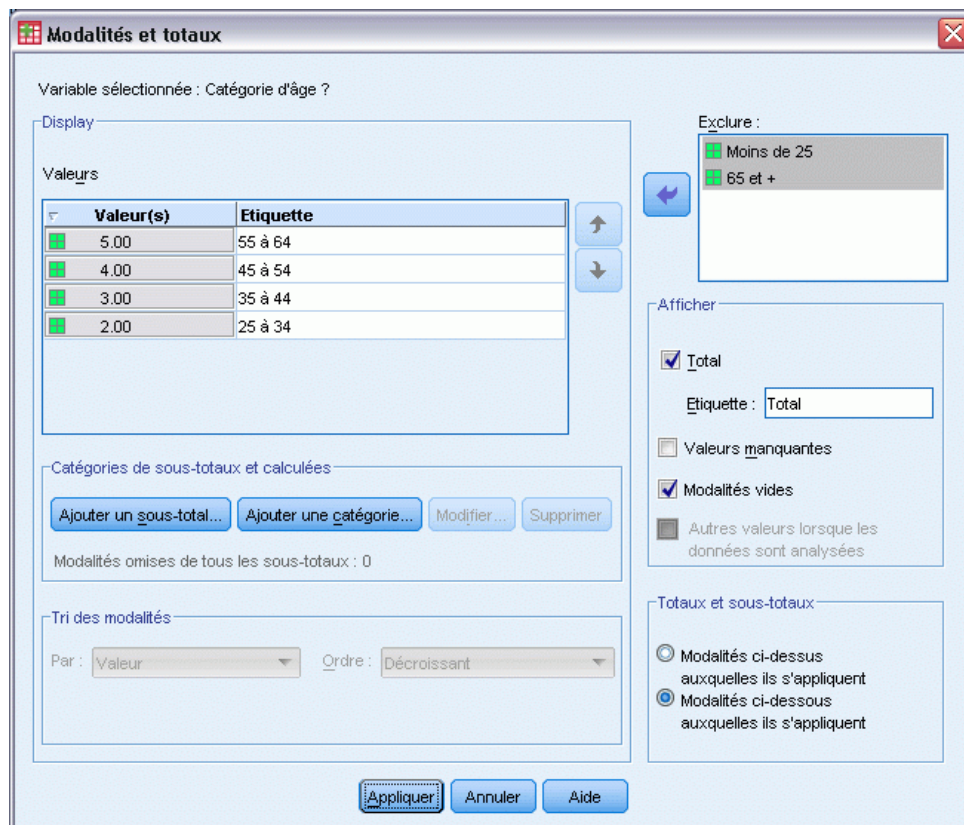
Si vous souhaitez que certaines modalités n'apparaissent pas dans le tableau, vous pouvez les exclure.

- ▶ Dans la liste Etiquette, cliquez sur la modalité portant l'étiquette *Moins de 25*.
- ▶ Cliquez sur le bouton fléché situé à gauche de la liste Exclure.
- ▶ Dans la liste Etiquette, cliquez sur la modalité portant l'étiquette *65 ans et plus*.

- Cliquez de nouveau sur le bouton fléché situé à gauche de la liste Exclure.

Les deux modalités sont déplacées de la liste Afficher vers la liste Exclure. Si vous changez d'avis, vous pouvez facilement les renvoyer vers la liste Afficher.

Figure 3-18
Modalités exclues manuellement



- Dans le générateur de tableaux, cliquez sur Appliquer, puis sur OK pour créer le tableau.

Figure 3-19
Tableau trié par étiquette de valeur (ordre décroissant), avec exclusion de certaines modalités

		Sexe du répondant		
		Homme	Femme	Total
Catégories d'âges	55 à 64	136	184	320
	45 à 54	221	260	481
	35 à 44	309	370	679
	25 à 34	276	351	627
	Total	942	1165	2107

Notez que les totaux étaient supérieurs avant l'exclusion des deux modalités. En effet, les totaux sont basés sur les modalités incluses dans le tableau. Toute modalité exclue est éliminée du calcul des totaux. [Pour plus d'informations, reportez-vous à la section Totaux et sous-totaux pour les variables qualitatives dans le chapitre 5 sur p. 86.](#)

Empilement, emboîtement et strates avec des variables qualitatives

L'empilement, l'emboîtement et les strates sont des méthodes permettant l'affichage de plusieurs variables dans le même tableau. Ce chapitre présente l'utilisation de ces techniques avec des variables qualitatives, bien qu'elles puissent également être utilisées avec des variables d'échelle.

Fichier de données d'exemple

Les exemples de ce chapitre utilisent le fichier de données *survey_sample.sav*. [Pour plus d'informations, reportez-vous à la section Fichiers d'exemple dans l'annexe A sur p. 197.](#)

Tous les exemples fournis ici affichent des étiquettes de variable dans des boîtes de dialogue ; ces étiquettes sont triées par ordre alphabétique. Les propriétés d'affichage de la liste des variables sont paramétrées dans l'onglet Général de la boîte de dialogue Options (menu Modifier, Options).

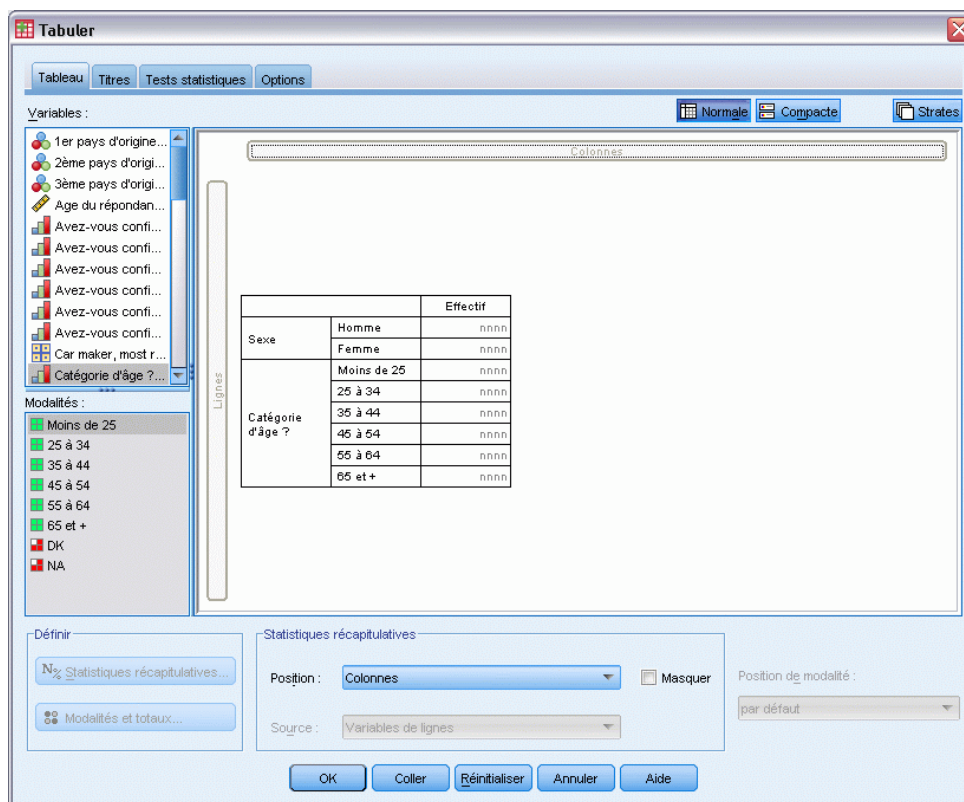
Empilement de variables qualitatives

Empiler revient à prendre des tableaux séparés et à les coller ensemble dans le même affichage. Vous pouvez, par exemple, afficher des informations sur les variables *sexe* et *tranche d'âge* dans des sections séparées du même tableau.

- ▶ A partir du menu, sélectionnez :
Analyse > Tableaux > Tabuler...
- ▶ Dans le générateur de tableaux, faites glisser la variable *Sexe* de la liste des variables vers la zone Lignes du panneau de canevas.
- ▶ Faites glisser la variable *Tranche d'âge* de la liste des variables vers la zone Lignes sous *Sexe*.

Les deux variables sont désormais empilées dans la dimension de ligne.

Figure 4-1
Variables qualitatives empilées affichées dans le panneau de canevas



- Cliquez sur OK pour créer le tableau.

Figure 4-2
Tableau des variables qualitatives empilées en lignes

		Effectif
Sexe	Homme	1232
	Femme	1600
Catégorie d'âge	Moins de 25	242
	25 à 34	627
	35 à 44	679
	45 à 54	481
	55 à 64	320
	65 et +	479

Vous pouvez également empiler de la même façon des variables en colonnes.

Empilement et tableau croisé

Un tableau empilé peut inclure d'autres variables dans d'autres dimensions. Par exemple, vous pouvez croiser deux variables empilées dans les lignes avec une troisième variable affichée dans la dimension de colonne.

- ▶ Ouvrez à nouveau le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Si les variables *Tranche d'âge* et *Sexe* ne sont pas encore empilées dans les lignes, suivez les instructions précédentes pour les empiler.
- ▶ Faites glisser la variable *Informations sur Internet* de la liste des variables vers la zone Colonnes du panneau de canevas.
- ▶ Cliquez sur OK pour créer le tableau.

Figure 4-3

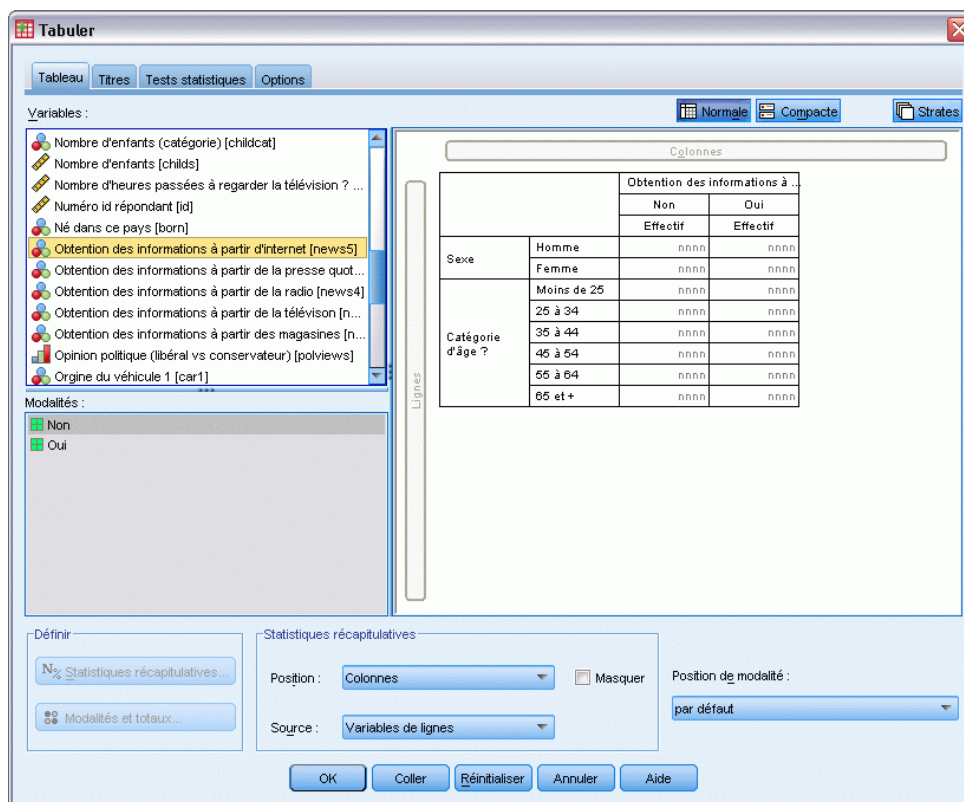
Tableau croisé de deux variables en ligne empilées avec une variable en colonne

		Obtention de l'information via internet	
		Non	Oui
		Effectif	Effectif
Sexe du répondant	Homme	873	359
	Femme	1092	508
Catégories d'âges	Moins de 25	146	96
	25 à 34	368	259
	35 à 44	435	244
	45 à 54	346	135
	55 à 64	252	68
	65 ou plus	416	63

Remarque : Plusieurs variables ayant des étiquettes qui commencent par *Informations...*, il peut s'avérer difficile de les distinguer dans la liste des variables (les étiquettes peuvent en effet être trop longues pour être affichées complètement dans la liste des variables). Deux méthodes permettent de visualiser l'intégralité de l'étiquette de variable :

- Positionnez le pointeur de la souris sur une variable dans la liste afin d'afficher l'intégralité de l'étiquette dans une note d'aide.
- Dans le panneau de canevas, faites glisser la barre verticale qui sépare la liste des variables et la liste des modalités afin d'élargir les listes.

Figure 4-4
Liste des variables élargies afin d'afficher l'intégralité des étiquettes de variable



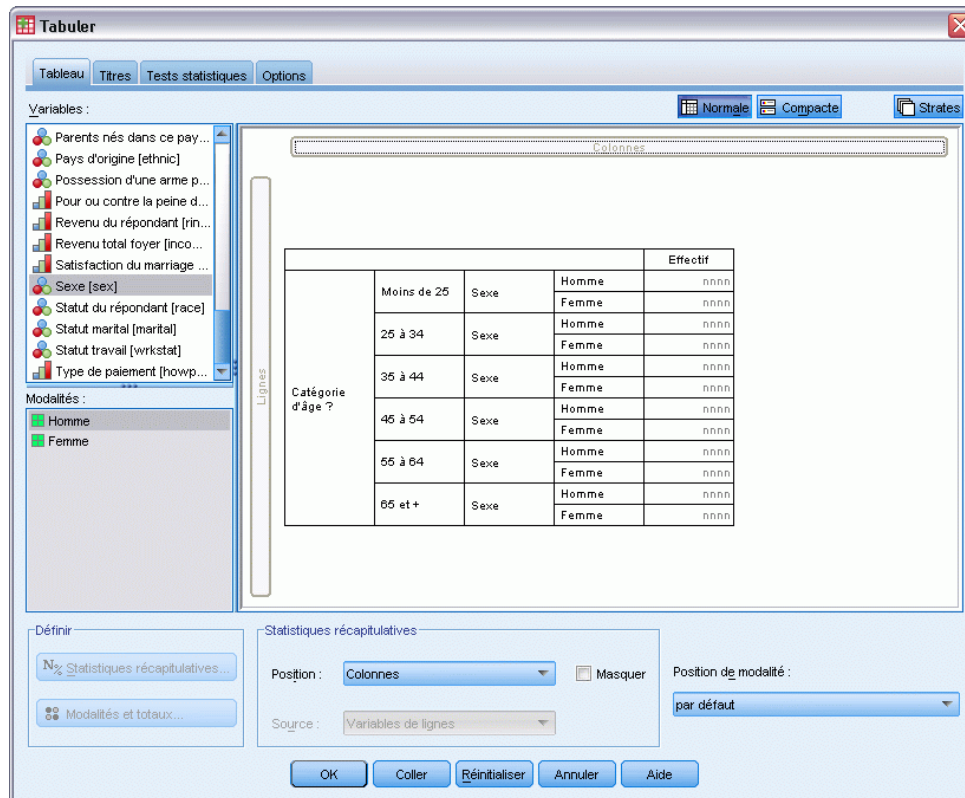
Emboîtement de variables qualitatives

L'emboîtement, comme la technique des tableaux croisés, permet de présenter la relation entre deux variables qualitatives ; la seule différence est qu'une variable est emboîtée dans l'autre dans la même dimension. Vous pouvez, par exemple, emboîter la variable *Sexe* dans la variable *Tranche d'âge* dans la dimension de ligne, ce qui a pour effet de présenter le nombre d'hommes et de femmes dans chaque tranche d'âge.

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Cliquez sur Restaurer pour supprimer les sélections précédentes dans le générateur de tableaux.
- ▶ Dans le générateur de tableaux, faites glisser la variable *Tranche d'âge* de la liste des variables vers la zone Lignes du panneau de canevas.
- ▶ Faites glisser la variable *Sexe* de la liste des variables vers la droite de la variable *Tranche d'âge* dans la zone Lignes.

Dans le panneau de canevas, l'aperçu indique désormais que le tableau emboîté sera composé d'une seule colonne d'effectifs, chaque cellule contenant le nombre d'hommes et de femmes dans chaque tranche d'âge.

Figure 4-5
Variable Sexe emboîtée dans la variable Tranche d'âge

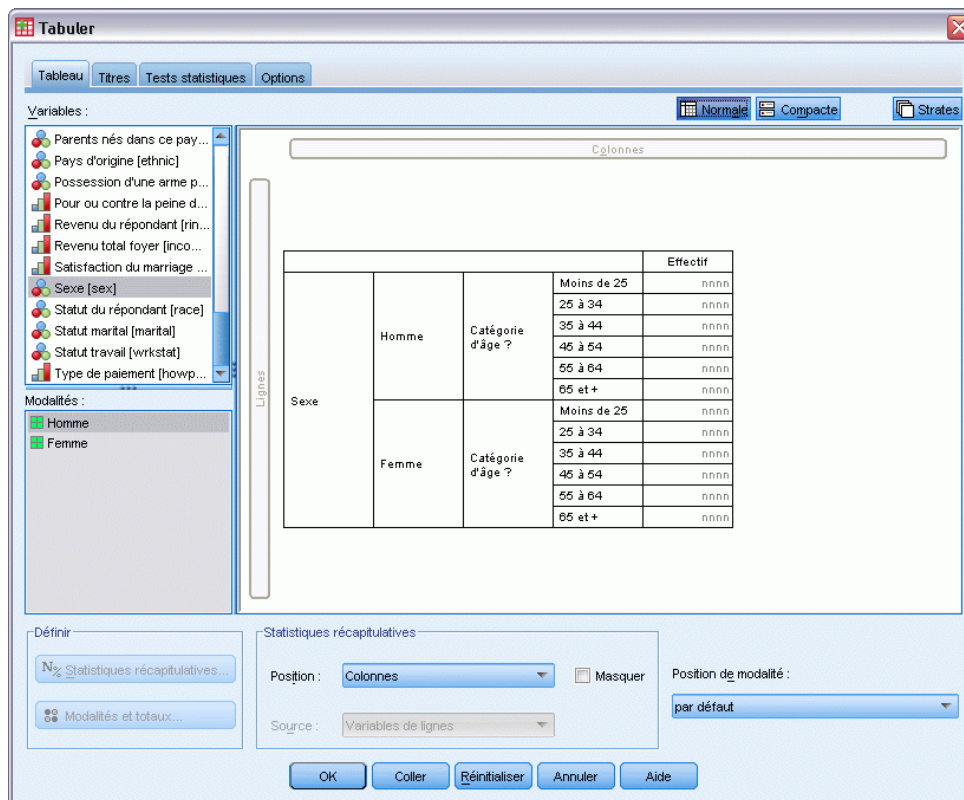


Vous pouvez remarquer que l'étiquette de variable *Sexe* apparaît plusieurs fois, pour chaque tranche d'âge. Vous pouvez minimiser ce type de répétition en plaçant la variable contenant le plus petit nombre de modalités dans le niveau d'emboîtement le plus externe.

- ▶ Dans le panneau de canevas, cliquez sur l'étiquette de variable *Sexe*.
- ▶ Faites glisser la variable le plus à gauche possible vers la zone Lignes.

Désormais, au lieu que la variable *Sexe* soit répétée six fois, la variable *Tranche d'âge* est répétée deux fois. Le tableau est alors moins confus et produira essentiellement les mêmes résultats.

Figure 4-6
Variable Tranche d'âge emboîtée dans la variable Sexe dans l'aperçu du générateur de tableaux



- Cliquez sur OK pour créer le tableau.

Figure 4-7
Tableau présentant la variable Tranche d'âge emboîtée dans la variable Sexe

				Effectif
Sexe du répondant	Homme	Catégories d'âges	Moins de 25	108
			25 à 34	276
			35 à 44	309
	Femme		45 à 54	221
			55 à 64	136
			65 ou plus	178
	Catégories d'âges	Moins de 25	134	
		25 à 34	351	
		35 à 44	370	
		45 à 54	260	
		55 à 64	184	
		65 ou plus	301	

Remarque : Les tableaux personnalisés ne prennent pas en charge le traitement de fichiers scindés en strates. Pour obtenir le même résultat que les fichiers scindés en strates, placez les variables de fichiers scindés dans les strates emboîtées les plus à l'extérieur du tableau.

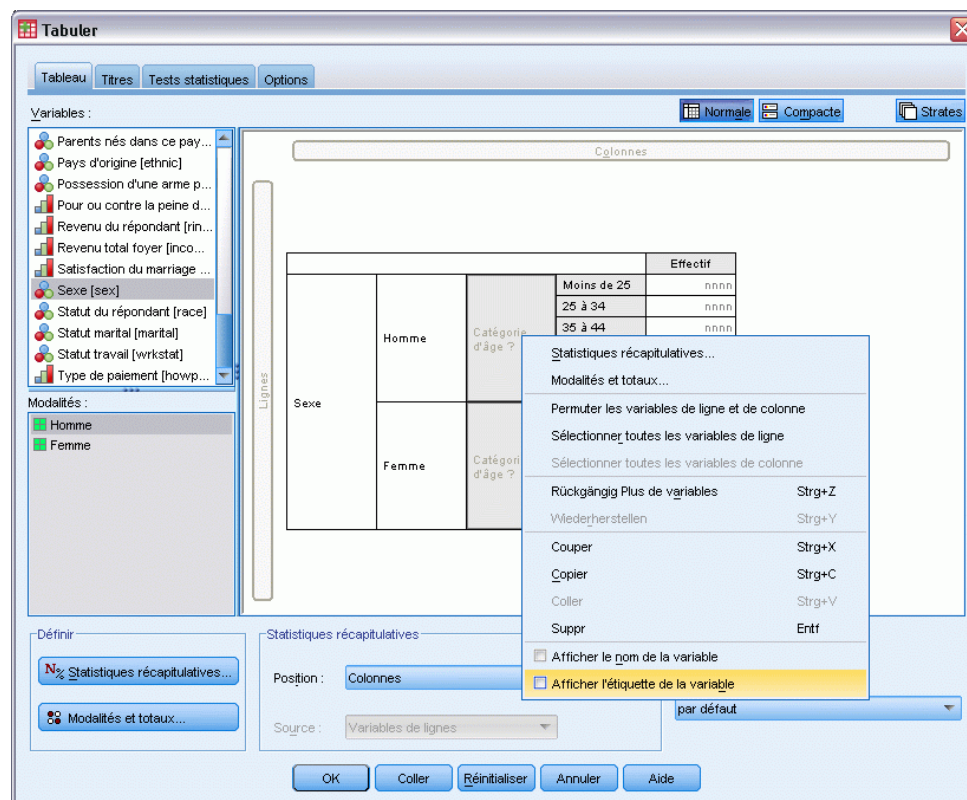
Suppression d'étiquettes de variable

Une autre solution au problème de la redondance des étiquettes de variable dans les tableaux emboîtés consiste simplement à supprimer l'affichage des noms ou étiquettes de variable. Les étiquettes de valeurs des variables *Sexe* et *Tranche d'âge* étant suffisamment descriptives sans les étiquettes de variable, nous pouvons éliminer les étiquettes pour les deux variables.

- Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur la variable *Tranche d'âge* et désélectionnez Afficher l'étiquette des variables dans le menu contextuel.
- Répétez l'opération pour la variable *Sexe*.

Figure 4-8

Suppression des étiquettes de variable via le menu contextuel dans le générateur de tableaux



Les étiquettes de variable restent affichées dans l'aperçu du tableau, mais ne seront pas incluses dans le tableau.

- Cliquez sur OK pour créer le tableau.

Figure 4-9
Tableau emboîté sans étiquettes de variable

		Effectif
Homme	Moins de 25	108
	25 à 34	276
	35 à 44	309
	45 à 54	221
	55 à 64	136
	65 ou plus	178
Femme	Moins de 25	134
	25 à 34	351
	35 à 44	370
	45 à 54	260
	55 à 64	184
	65 ou plus	301

Si vous souhaitez que les étiquettes de variable soient incluses quelque part dans le tableau, sans toutefois les afficher plusieurs fois dans le corps du tableau, vous pouvez les inclure dans le titre ou l'étiquette de coin du tableau.

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Cliquez sur l'onglet Titres.
- ▶ Cliquez n'importe où dans la zone de texte Titre.
- ▶ Cliquez sur Expression de tableau. Le texte *&[Expression de tableau]* apparaît dans la zone de texte Titre. Ceci génère un titre de tableau incluant les étiquettes des variables utilisées dans le tableau.
- ▶ Cliquez sur OK pour créer le tableau.

Figure 4-10
Étiquettes de variable dans le titre du tableau

Sexe du répondant > Catégories d'âges

		Effectif
Homme	Moins de 25	108
	25 à 34	276
	35 à 44	309
	45 à 54	221
	55 à 64	136
	65 ou plus	178
Femme	Moins de 25	134
	25 à 34	351
	35 à 44	370
	45 à 54	260
	55 à 64	184
	65 ou plus	301

Le signe “ supérieur à “ (>) dans le titre indique que la variable *Tranche d'âge* est emboîtée dans la variable *Sexe*.

Tableau croisé emboîté

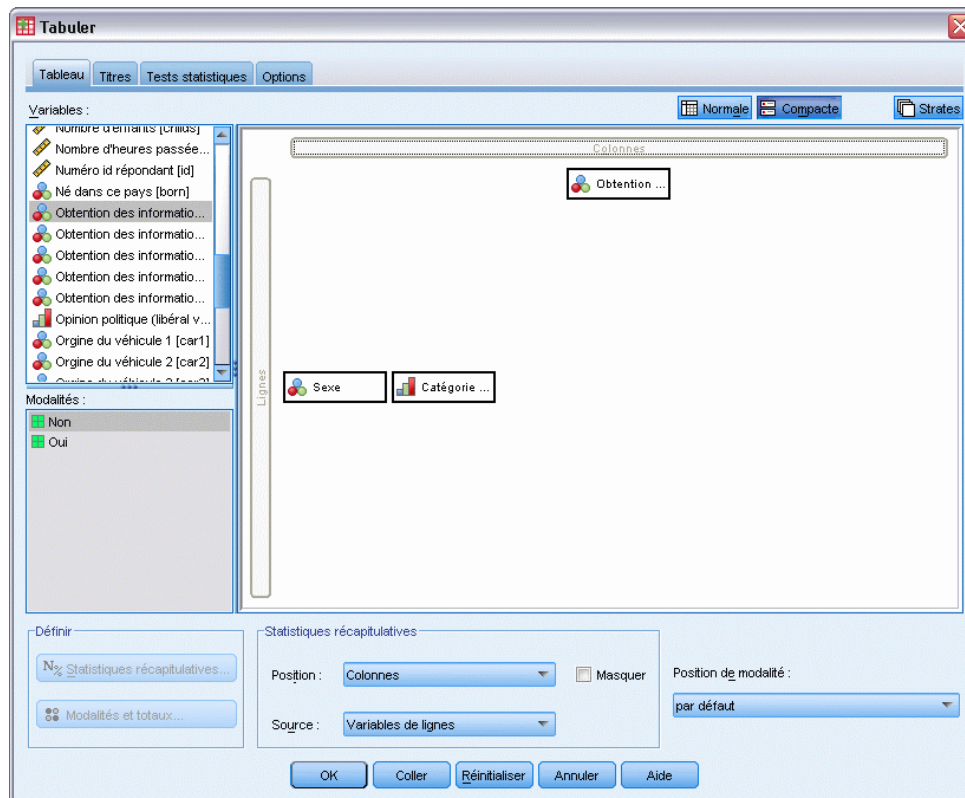
Un tableau emboîté peut contenir d'autres variables dans d'autres dimensions. Par exemple, vous pouvez emboîter la variable *Tranche d'âge* dans la variable *Sexe* dans les lignes et croiser les lignes emboîtées avec une troisième variable dans la dimension de colonne.

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Si la variable *Tranche d'âge* n'est pas encore emboîtée dans la variable *Sexe* dans les lignes, suivez les instructions précédentes pour les emboîter.
- ▶ Faites glisser la variable *Informations sur Internet* de la liste des variables vers la zone Colonnes du panneau de canevas.

Vous pouvez remarquer que le tableau est trop grand pour être affiché entièrement dans le panneau de canevas. Vous pouvez utiliser le défilement vertical ou horizontal dans le panneau de canevas pour visualiser davantage l'aperçu du tableau ou vous pouvez :

- Dans le générateur de tableaux, cliquez sur Compact pour obtenir le mode Compact. Ceci affiche uniquement les étiquettes de variable, sans aucune information sur les modalités ou les statistiques récapitulatives incluses dans le tableau.
- Augmentez la taille du générateur de tableaux en faisant glisser ses bords ou ses coins.

Figure 4-11
Mode Compact dans le panneau de canevas



- ▶ Cliquez sur OK pour créer le tableau.

Figure 4-12
Tableau croisé emboîté

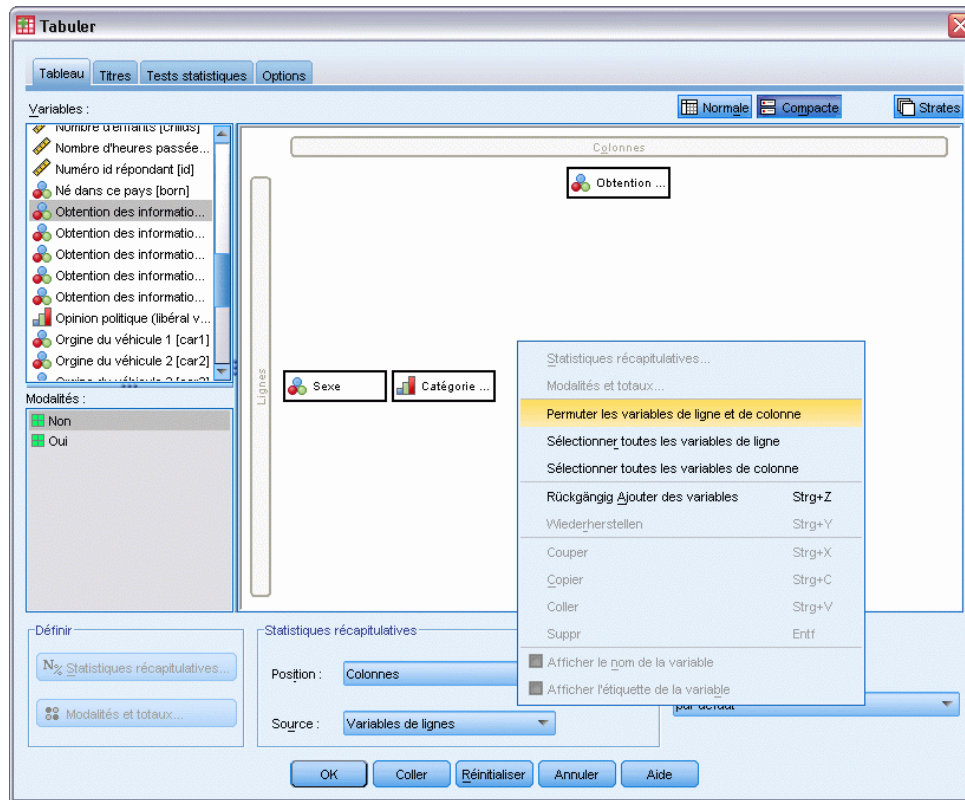
				Obtention de l'information via internet		
				Non	Oui	
				Effectif	Effectif	
Sexe du répondant	Homme	Catégories d'âges	Moins de 25	59	49	
			25 à 34	159	117	
			35 à 44	217	92	
			45 à 54	169	52	
			55 à 64	112	24	
		Femme	Catégories d'âges	65 ou plus	155	23
		Moins de 25		87	47	
		25 à 34		209	142	
		35 à 44		218	152	
		45 à 54		177	83	
			55 à 64	140	44	
			65 ou plus	261	40	

Permutation des lignes et des colonnes

Après avoir passé un certain temps à élaborer un tableau complexe, vous jugez que ce dernier est absolument parfait, si ce n'est que vous souhaitez en modifier l'orientation, c'est-à-dire placer toutes les variables en ligne dans les colonnes et vice versa. Par exemple, vous avez créé un tableau croisé emboîté avec les variables *Tranche d'âge* et *Sexe* emboîtées dans les lignes, mais vous souhaitez désormais que ces deux variables démographiques soient plutôt emboîtées dans les colonnes.

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Dans le panneau de canevas, cliquez n'importe où avec le bouton droit de la souris et sélectionnez *Permuter les variables en colonne et en ligne* dans le menu contextuel.

Figure 4-13
Permutation des variables en ligne et en colonne



Les variables en ligne et en colonne ont été permutées.

Avant de créer le tableau, procédons à quelques modifications afin de rendre l'affichage moins confus.

- ▶ Sélectionnez Masquer pour supprimer l'affichage de l'étiquette de colonne des statistiques récapitulatives.
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur la variable *Sexe* et désélectionnez Afficher l'étiquette des variables.
- ▶ Cliquez sur OK pour créer le tableau.

Figure 4-14
Tableau croisé dans lequel les variables démographiques sont emboîtées dans les colonnes

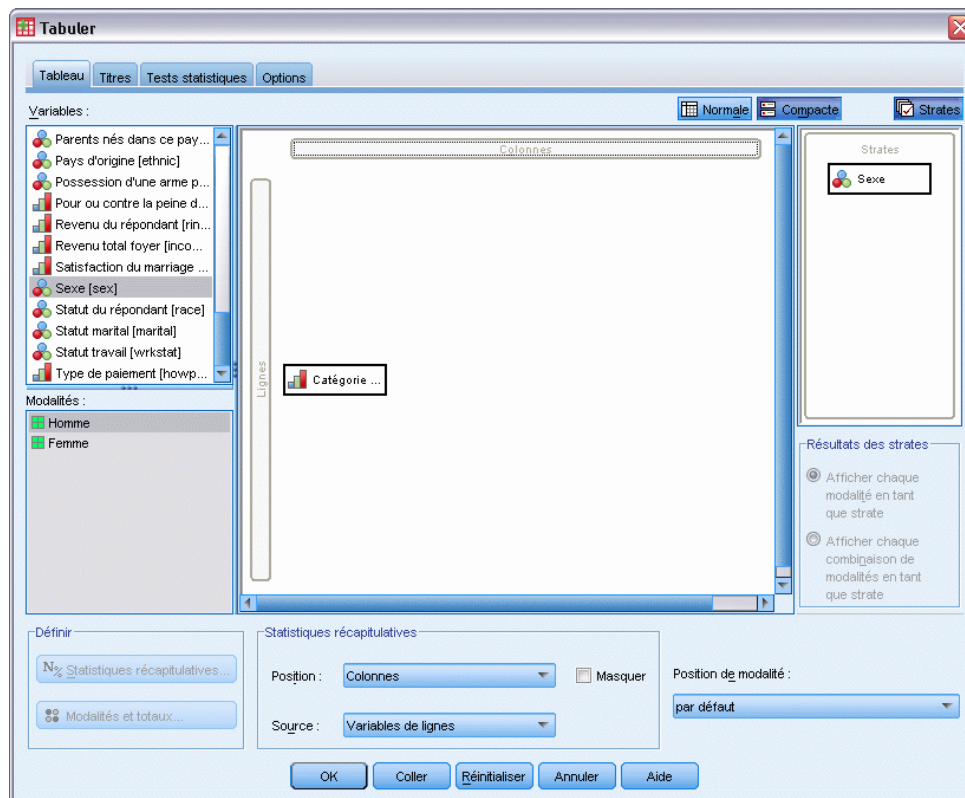
		Homme						Femme					
		Catégories d'âges						Catégories d'âges					
		Moins de	25 à 34	35 à 44	45 à 54	55 à 64	65 ou	Moins de	25 à 34	35 à 44	45 à 54	55 à 64	65 ou
Obtention de l'information via internet	Non	59	159	217	169	112	155	87	209	218	177	140	261
	Oui	49	117	92	52	24	23	47	142	152	83	44	40

Strates

Vous pouvez utiliser des strates pour ajouter une dimension de profondeur à vos tableaux et créer ainsi des « cubes » tridimensionnels. Les strates sont en fait relativement similaires à l'emboîtement ou à l'empilement ; la différence principale réside dans le fait qu'une seule modalité de strate est visible à la fois. Par exemple, l'utilisation de la variable *Tranche d'âge* comme variable de ligne et de la variable *Sexe* comme variable de strate a pour effet de produire un tableau dans lequel les informations relatives aux hommes et aux femmes apparaissent dans différentes strates.

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Cliquez sur Restaurer pour supprimer les sélections précédentes dans le générateur de tableaux.
- ▶ Dans le générateur de tableaux, faites glisser la variable *Tranche d'âge* de la liste des variables vers la zone Lignes du panneau de canevas.
- ▶ Cliquez sur Strates au sommet du générateur de tableaux pour afficher la liste Strates.
- ▶ Faites glisser la variable *Sexe* de la liste des variables vers la liste Strates.

Figure 4-15
Variable *Tranche d'âge* en lignes, variable *Sexe* en strates



A ce stade, vous pouvez remarquer que l'ajout d'une variable de strate n'a pas d'effet visible sur l'aperçu affiché dans le panneau de canevas. Les variables de strate n'affectent pas l'aperçu du panneau de canevas, sauf si la variable de strate est la variable source des statistiques et si vous modifiez les statistiques récapitulatives.

- Cliquez sur OK pour créer le tableau.

Figure 4-16

Tableau simple en strates

Sexe du répondant Homme		Effectif
Catégories d'âges	Moins de 25	108
	25 à 34	276
	35 à 44	309
	45 à 54	221
	55 à 64	136
	65 ou plus	178

Au premier abord, ce tableau ne semble pas différent d'un tableau simple comportant une seule variable qualitative. La seule différence réside dans la présence de l'étiquette *Sexe - Masculin* au sommet du tableau.

- Dans la fenêtre du Viewer, double-cliquez sur le tableau pour l'activer.
- Vous pouvez désormais remarquer que l'étiquette *Sexe - Masculin* représente en fait un choix dans la liste déroulante.
- Cliquez sur la flèche vers le bas de la liste déroulante pour afficher la liste complète des strates.

Figure 4-17

Liste des strates dans un tableau pivotant activé

Strate	Sexe du répondant Homme	
	Sexe du répondant Homme	
Catégories d'âges	Sexe du répondant Femme	
	25 à 34	276
	35 à 44	309
	45 à 54	221
	55 à 64	136
	65 ou plus	178

Pour ce tableau, un seul autre choix existe dans la liste.

- Dans la liste déroulante, sélectionnez *Sexe - Féminin*.

Figure 4-18

Tableau simple en strates dans lequel une autre strate est affichée

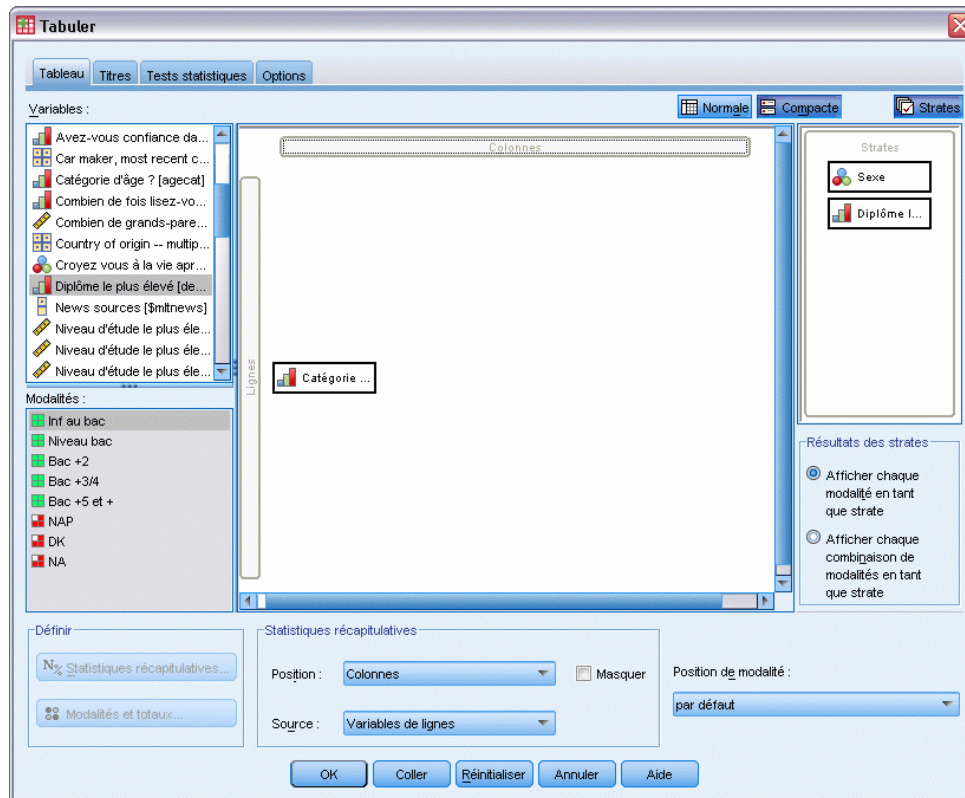
Sexe du répondant Femme		Effectif
Catégories d'âges	Moins de 25	134
	25 à 34	351
	35 à 44	370
	45 à 54	260
	55 à 64	184
	65 ou plus	301

Deux variables qualitatives de strate empilées

Si plusieurs variables qualitatives sont présentes dans les strates, vous pouvez les empiler ou les emboîter. Par défaut, les variables de strate sont empilées. (*Remarque* : Si vous disposez de variables d'échelle en strates, les variables de strate peuvent uniquement être empilées.)

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Si la variable *Tranche d'âge* ne se situe pas déjà dans les lignes et la variable *Sexe* dans les strates, suivez les instructions précédentes pour créer un tableau en strates.
- ▶ Faites glisser la variable *Diplôme le plus important* de la liste des variables vers la liste Strate, sous *Sexe*.

Figure 4-19
Variables de strate empilées dans le générateur de tableaux



Les deux boutons radio situés sous la liste Strate dans le groupe Sortie strate sont désormais activés. La sélection par défaut est Afficher chaque modalité en strate. Ceci équivaut à l'empilement.

- ▶ Cliquez sur OK pour créer le tableau.
- ▶ Dans la fenêtre du Viewer, double-cliquez sur le tableau pour l'activer.
- ▶ Cliquez sur la flèche vers le bas de la liste déroulante pour afficher la liste complète des strates.

Figure 4-20
Liste des strates empilées dans un tableau pivotant activé

Strate	Sexe du répondant Homme
	Sexe du répondant Homme
Catégorie d'âges	Sexe du répondant Femme
	Diplôme le plus élevé < Bac
	Diplôme le plus élevé Bac
	Diplôme le plus élevé Licence
	Diplôme le plus élevé Maîtrise
	Diplôme le plus élevé Bac + 5

Le tableau comporte sept strates : deux strates pour les deux modalités de la variable *Sexe* et cinq strates pour les cinq modalités de la variable *Diplôme le plus important*. En ce qui concerne les strates empilées, le nombre total de strates correspond au nombre de modalités pour les variables de strate (y compris les modalités de total et sous-total demandées pour les variables de strate).

Deux variables qualitatives de strate emboîtées

L'emboîtement de variables qualitatives de strate crée une strate distincte pour chaque combinaison de modalités de la variable de strate.

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Si vous ne l'avez déjà fait, suivez les instructions précédentes pour créer un tableau de strates empilées.
- ▶ Dans le groupe Sortie strate, sélectionnez Afficher chaque combinaison de modalités en strate. Ceci équivaut à l'emboîtement.
- ▶ Cliquez sur OK pour créer le tableau.
- ▶ Dans la fenêtre du Viewer, double-cliquez sur le tableau pour l'activer.
- ▶ Cliquez sur la flèche vers le bas de la liste déroulante pour afficher la liste complète des strates.

Figure 4-21
Liste des strates emboîtées dans un tableau pivotant activé

Strate	Sexe du répondant Homme Diplôme le plus élevé < Bac
	Sexe du répondant Homme Diplôme le plus élevé < Bac
	Sexe du répondant Homme Diplôme le plus élevé Bac
Catégorie d'âges	Sexe du répondant Homme Diplôme le plus élevé Licence
	Sexe du répondant Homme Diplôme le plus élevé Maîtrise
	Sexe du répondant Homme Diplôme le plus élevé Bac + 5
	Sexe du répondant Femme Diplôme le plus élevé < Bac
	Sexe du répondant Femme Diplôme le plus élevé Bac
	Sexe du répondant Femme Diplôme le plus élevé Licence

Le tableau comporte 10 strates (vous devez parcourir la liste pour les visualiser toutes), une pour chaque combinaison des variables *Sexe* et *Diplôme le plus important*. En ce qui concerne les strates emboîtées, le nombre total de strates correspond au *produit* du nombre de modalités pour chaque variable de strate (dans cet exemple, $5 \times 2 = 10$).

Impression de tableaux en strates

Par défaut, seule la strate actuellement visible est imprimée. Pour imprimer toutes les strates d'un tableau :

- ▶ Dans la fenêtre du Viewer, double-cliquez sur le tableau pour l'activer.
- ▶ A partir du menu Fenêtre du Viewer, sélectionnez :
Format > Propriétés du tableau
- ▶ Cliquez sur l'onglet Impression.
- ▶ Sélectionnez Imprimer toutes les strates.

Vous pouvez également enregistrer ce paramètre comme partie constituante d'un modèle de tableau, y compris le modèle de tableau par défaut.

Totaux et sous-totaux pour les variables qualitatives

Vous pouvez inclure à la fois des totaux et des sous-totaux dans les tableaux personnalisés. Vous pouvez appliquer des totaux et des sous-totaux à des variables qualitatives à tout niveau d'emboîtement, dans toute dimension (ligne, colonne et strate).

Fichier de données d'exemple

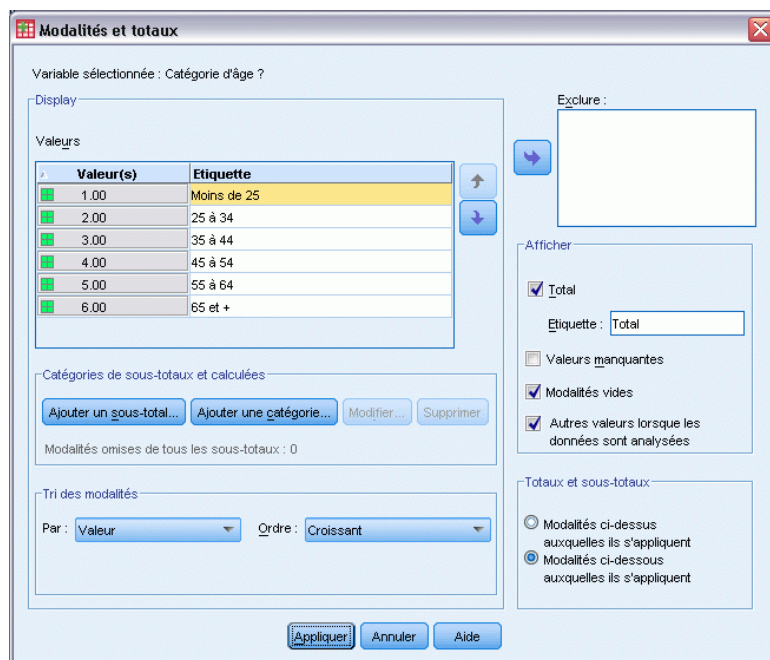
Les exemples de ce chapitre utilisent le fichier de données *survey_sample.sav*. [Pour plus d'informations, reportez-vous à la section Fichiers d'exemple dans l'annexe A sur p. 197.](#)

Tous les exemples fournis ici affichent des étiquettes de variable dans des boîtes de dialogue ; ces étiquettes sont triées par ordre alphabétique. Les propriétés d'affichage de la liste des variables sont paramétrées dans l'onglet Général de la boîte de dialogue Options (menu Modifier, Options).

Total simple pour une seule variable

- ▶ A partir du menu, sélectionnez :
Analyse > Tableaux > Tabuler...
- ▶ Dans le générateur de tableaux, faites glisser la variable *Tranche d'âge* de la liste des variables vers la zone Lignes du panneau de canevas.
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Tranche d'âge* et choisissez Statistiques récapitulatives dans le menu contextuel.
- ▶ Dans la boîte de dialogue Statistiques récapitulatives, sélectionnez N % colonne dans la liste Statistiques et cliquez sur la flèche pour ajouter l'élément à la liste Afficher.
- ▶ Dans la cellule Etiquette de la liste Afficher, supprimez l'étiquette par défaut et tapez Pourcentage.
- ▶ Cliquez sur Appliquer à la sélection.
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Tranche d'âge* et choisissez Modalités et totaux dans le menu contextuel.
- ▶ Dans la boîte de dialogue Modalités et totaux, sélectionnez Total.

Figure 5-1
Boîte de dialogue Modalités et totaux



- Dans le générateur de tableaux, cliquez sur Appliquer, puis sur OK pour créer le tableau.

Figure 5-2
Total simple pour une seule variable qualitative

		Effectif	Pourcentage
Catégories d'âges	Moins de 25	242	8,6%
	25 à 34	627	22,2%
	35 à 44	679	24,0%
	45 à 54	481	17,0%
	55 à 64	320	11,3%
	65 ou plus	479	16,9%
	Total	2828	100,0%

Tel écran, tel totalisé

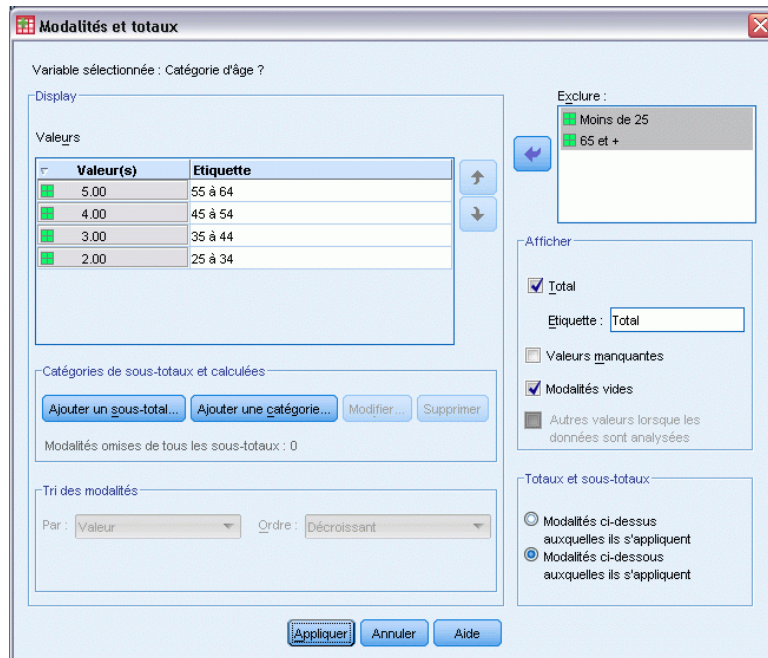
Les totaux sont basés sur les modalités affichées dans le tableau. Si vous choisissez d'exclure certaines modalités d'un tableau, les observations issues de ces modalités ne sont pas incluses dans les calculs des totaux.

- Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Tranche d'âge* et choisissez Modalités et totaux dans le menu contextuel.
- Dans la liste Etiquette, cliquez sur la modalité portant l'étiquette *Moins de 25*.
- Cliquez sur le bouton fléché situé à gauche de la liste Exclure.
- Dans la liste Etiquette, cliquez sur la modalité portant l'étiquette *65 ans et plus*.

- Cliquez de nouveau sur le bouton fléché situé à gauche de la liste Exclure.

Les deux modalités sont déplacées de la liste Afficher vers la liste Exclure.

Figure 5-3
Modalités exclues manuellement



- Dans le générateur de tableaux, cliquez sur Appliquer, puis sur OK pour créer le tableau.

Figure 5-4
Total dans un tableau dans lequel des modalités ont été exclues

	Effectif	Pourcentage
Catégories d'âges		
25 à 34	627	29,8%
35 à 44	679	32,2%
45 à 54	481	22,8%
55 à 64	320	15,2%
Total	2107	100,0%

Le total dans ce tableau est seulement de 2 107, alors qu'il est de 2 828 lorsque toutes les modalités sont incluses. Seules les modalités utilisées dans le tableau sont incluses dans le total. (Le total des pourcentages est tout de même de 100 % car tous les pourcentages sont basés sur le nombre total d'observations utilisées dans le tableau, et non sur le nombre total d'observations dans le fichier de données.)

Position d'affichage des totaux

Par défaut, les totaux sont affichés sous les modalités totalisées. Vous pouvez modifier la position d'affichage des totaux et les placer au-dessus des modalités totalisées.

- Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).

- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Tranche d'âge* et choisissez Modalités et totaux dans le menu contextuel.
- ▶ Dans le groupe Affichage des totaux et des sous-totaux, sélectionnez Modalités ci-dessus auxquelles ils s'appliquent.
- ▶ Dans le générateur de tableaux, cliquez sur Appliquer, puis sur OK pour créer le tableau.

Figure 5-5

Total affiché au-dessus des modalités totalisées

Catégories	Total	Effectif	Pourcentage
d'âges	25 à 34	627	29,8%
	35 à 44	679	32,2%
	45 à 54	481	22,8%
	55 à 64	320	15,2%

Totaux pour tableaux emboîtés

Puisque les totaux peuvent être appliqués aux variables qualitatives à tout niveau d'emboîtement, vous pouvez créer des tableaux qui contiennent des totaux de groupe à plusieurs niveaux d'emboîtement.

Totaux de groupe

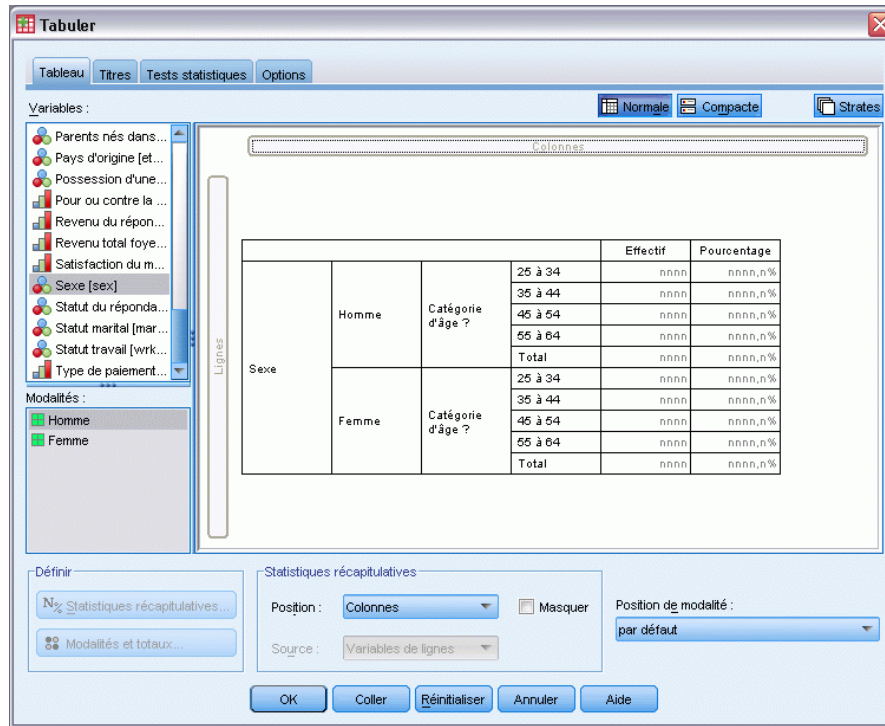
Les totaux des variables qualitatives emboîtées dans d'autres variables qualitatives représentent des totaux de groupe.

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Faites glisser la variable *Sexe* vers la gauche de la variable *Tranche d'âge* dans le panneau de canevas.
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Tranche d'âge* et choisissez Modalités et totaux dans le menu contextuel.

Avant de créer le tableau, repositionnons les totaux sous les modalités totalisées.

- ▶ Dans le groupe Affichage des totaux et des sous-totaux, sélectionnez Modalités ci-dessous auxquelles ils s'appliquent.
- ▶ Cliquez sur Appliquer pour enregistrer le paramètre et revenir au générateur de tableaux.

Figure 5-6
Variable Tranche d'âge emboîtée dans la variable Sexe dans le générateur de tableaux



- Cliquez sur OK pour créer le tableau.

Figure 5-7
Totaux de la variable Tranche d'âge au sein des modalités de la variable Sexe

				Effectif	Pourcentage
Sexe du répondant	Homme	Catégories d'âges	25 à 34	276	29,3%
			35 à 44	309	32,8%
			45 à 54	221	23,5%
			55 à 64	136	14,4%
			Total	942	100,0%
	Femme	Catégories d'âges	25 à 34	351	30,1%
			35 à 44	370	31,8%
			45 à 54	260	22,3%
			55 à 64	184	15,8%
			Total	1165	100,0%

Le tableau comporte désormais deux totaux de groupe : un pour les hommes et un pour les femmes.

Totaux généraux

Les totaux appliqués aux variables emboîtées sont toujours des totaux de groupe et non des totaux généraux. Si vous souhaitez obtenir des totaux pour l'intégralité du tableau, vous pouvez appliquer des totaux à la variable située au niveau d'emboîtement le plus externe.

- Ouvrez à nouveau le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).

- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur la variable *Sexe* et choisissez Modalités et totaux dans le menu contextuel.
- ▶ Dans la boîte de dialogue Modalités et totaux, sélectionnez Total.
- ▶ Dans le générateur de tableaux, cliquez sur Appliquer, puis sur OK pour créer le tableau.

Figure 5-8
Totaux généraux pour un tableau emboîté

				Effectif	Pourcentage
Sexe du répondant	Homme	Catégories d'âges	25 à 34	276	29,3%
			35 à 44	309	32,8%
			45 à 54	221	23,5%
			55 à 64	136	14,4%
			Total	942	100,0%
	Femme	Catégories d'âges	25 à 34	351	30,1%
			35 à 44	370	31,8%
			45 à 54	260	22,3%
			55 à 64	184	15,8%
			Total	1165	100,0%
Total	Catégories d'âges	25 à 34	627	29,8%	
		35 à 44	679	32,2%	
		45 à 54	481	22,8%	
		55 à 64	320	15,2%	
		Total	2107	100,0%	

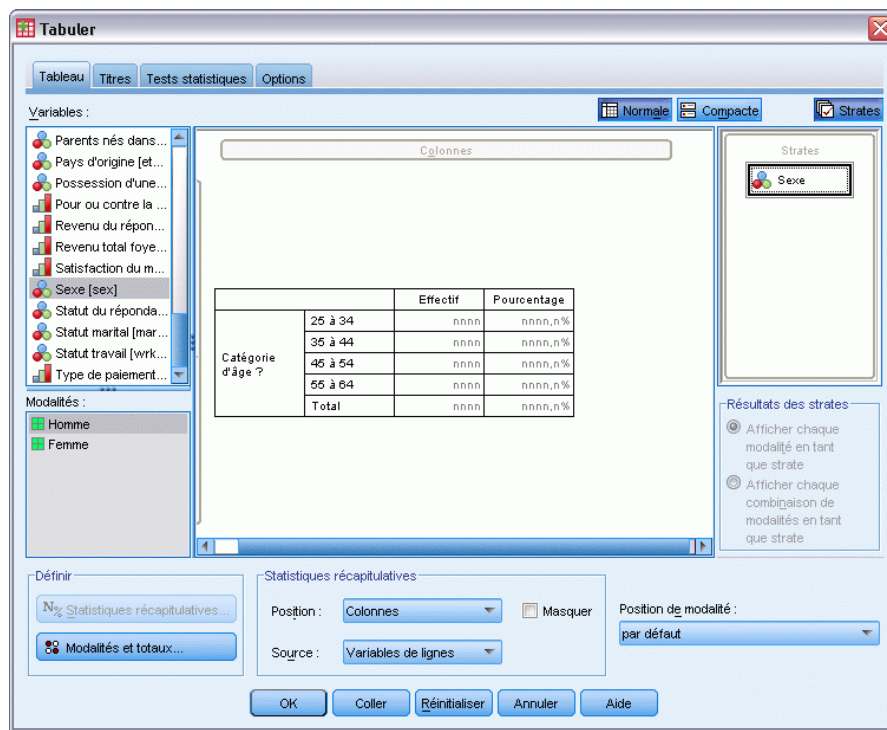
Notez que le total général n'est que de 2 107, et non de 2 828. Deux tranches (modalités) d'âge étant toujours exclues du tableau, les observations de ces modalités sont donc exclues de tous les totaux.

Totaux des variables de strate

Les totaux des variables de strate apparaissent sous la forme de strates distinctes dans le tableau.

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Dans le générateur de tableaux, cliquez sur Strates pour afficher la liste Strates.
- ▶ Faites glisser la variable *Sexe* de la zone Lignes du panneau de canevas vers la liste Strates.

Figure 5-9
Variable de strate dans le générateur de tableaux



Remarque : Puisque vous avez déjà spécifié des totaux pour la variable *Sexe*, il n'est pas nécessaire de le faire à ce stade. Le fait de déplacer la variable entre les dimensions n'affecte aucun des paramètres de cette variable.

- ▶ Cliquez sur OK pour créer le tableau.
- ▶ Dans le Viewer, double-cliquez sur le tableau pour l'activer.
- ▶ Cliquez sur la flèche vers le bas dans la liste déroulante *Strate* pour afficher une liste de toutes les strates du tableau.

Le tableau comporte trois strates : *Sexe - Masculin*, *Sexe - Féminin* et *Sexe - Total*.

Figure 5-10
Strate Total dans la liste *Strate* dans le tableau pivotant activé

Strate	Sexe du répondant		
	Sexe du répondant Homme		
	Sexe du répondant Femme		
	Sexe du répondant Total		
Catégorie d'âges	45 à 54	221	23,5%
	55 à 64	136	14,4%
	Total	942	100,0%

Position d'affichage des totaux de strate

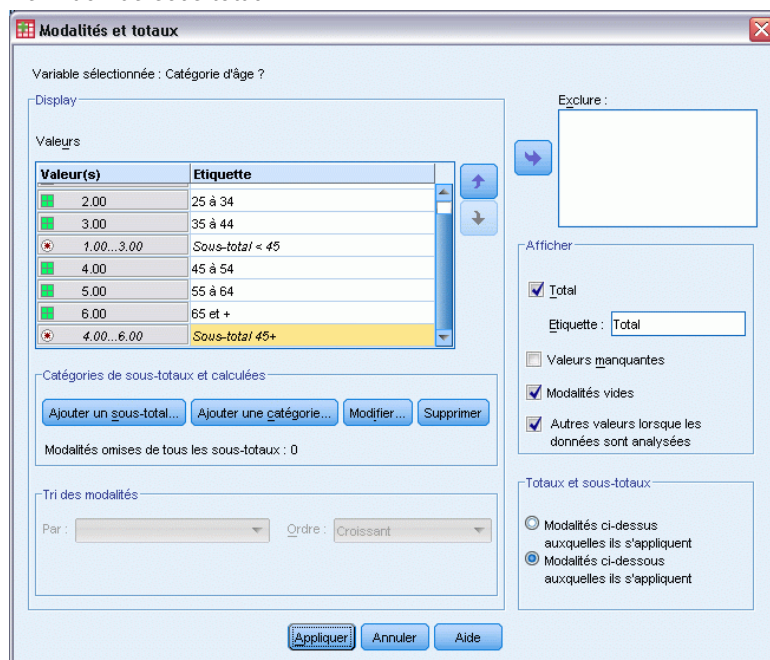
En ce qui concerne les totaux de variable de strate, la position d'affichage (au-dessus ou au-dessous) des totaux détermine la position de la strate pour les totaux. Par exemple, si vous spécifiez Modalités ci-dessus auxquelles ils s'appliquent pour un total de variable de strate, la strate du total est la première strate affichée.

Sous-totaux

Vous pouvez inclure des sous-totaux pour des sous-groupes de modalités d'une variable. Par exemple, vous pouvez inclure des sous-totaux pour les tranches (modalités) d'âge qui représentent tous les répondants âgés de moins et de plus de 45 ans dans l'échantillon d'enquête.

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Cliquez sur Restaurer pour effacer les paramètres précédents dans le générateur de tableaux.
- ▶ Dans le générateur de tableaux, faites glisser la variable *Tranche d'âge* de la liste des variables vers la zone Lignes du panneau de canevas.
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Tranche d'âge* et choisissez Modalités et totaux dans le menu contextuel.
- ▶ Sélectionnez 3.00 dans la liste Valeur(s).
- ▶ Cliquez sur Ajouter un sous-total pour afficher la boîte de dialogue Définir le sous-total.
- ▶ Saisissez Sous-total < 45 dans la zone de texte Etiquette.
- ▶ Cliquez ensuite sur Poursuivre.
Cette action insère une ligne contenant le sous-total des trois premières tranches d'âge.
- ▶ Sélectionnez 6.00 dans la liste Valeur(s).
- ▶ Cliquez sur Ajouter un sous-total pour afficher la boîte de dialogue Définir le sous-total.
- ▶ Saisissez Sous-total 45+ dans la zone de texte Etiquette.
- ▶ Cliquez ensuite sur Poursuivre.

Figure 5-11
Définition de sous-totaux



Remarque importante : Vous devez sélectionner la position d'affichage des totaux et des sous-totaux (Modalités ci-dessus auxquelles ils s'appliquent ou Modalités ci-dessous auxquelles ils s'appliquent) avant de définir un sous-total. La modification de la position d'affichage affecte tous les sous-totaux (et non pas uniquement le sous-total sélectionné) ; en outre, elle *modifie les modalités incluses dans les sous-totaux*.

- Dans le générateur de tableaux, cliquez sur Appliquer, puis sur OK pour créer le tableau.

Figure 5-12
Sous-totaux pour la variable Tranche d'âge

Catégories d'âges		Effectif
Moins de 25		242
25 à 34		627
35 à 44		679
Sous-total < 45		1548
45 à 54		481
55 à 64		320
65 ou plus		479
Sous-total 45+		1280

Tel écran, tel sous-totalisé

Tout comme les totaux, les sous-totaux sont basés sur les modalités incluses dans le tableau.

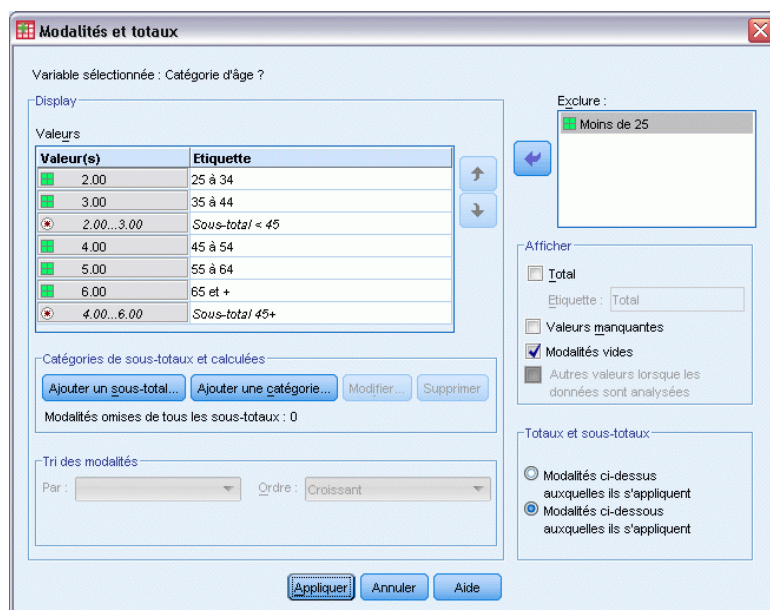
- Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).

- Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Tranche d'âge* et choisissez Modalités et totaux dans le menu contextuel.

Remarque : La valeur (non l'étiquette) affichée pour le premier sous-total est 1...3 ; elle indique que le sous-total inclut toutes les valeurs de la liste entre 1 et 3.

- Sélectionnez 1 dans la liste Valeur(s) (ou cliquez sur l'étiquette *Moins de 25*).
- Cliquez sur le bouton fléché situé à gauche de la liste Exclure.

Figure 5-13
Sous-taux lorsque des modalités sont exclues



Désormais, la première tranche (modalité) d'âge est exclue et la valeur affichée pour le premier sous-total passe à 2...3 ; elle indique que la modalité exclue ne sera pas incluse dans le sous-total car les sous-totaux sont basés sur les modalités incluses dans le tableau. L'exclusion d'une modalité entraîne son exclusion automatique des sous-totaux ; vous ne pouvez donc pas, par exemple, afficher uniquement des sous-totaux sans les modalités sur lesquelles ils se basent.

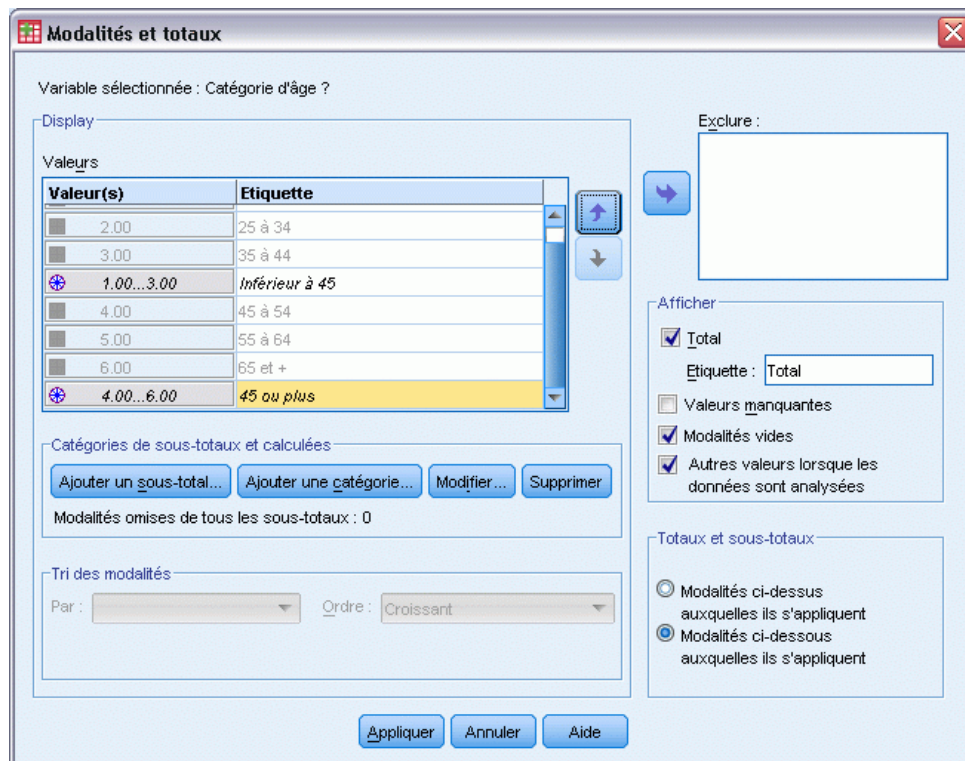
Masquage de modalités de sous-total

Vous pouvez supprimer l'affichage des modalités définissant un sous-total et afficher uniquement le sous-total, sans affecter les données sous-jacentes, en « fusionnant » efficacement les modalités.

- Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- Cliquez sur Restaurer pour effacer les paramètres précédents dans le générateur de tableaux.
- Dans le générateur de tableaux, faites glisser la variable *Tranche d'âge* de la liste des variables vers la zone Lignes du panneau de canevas.

- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Tranche d'âge* et choisissez Modalités et totaux dans le menu contextuel.
 - ▶ Sélectionnez 3.00 dans la liste Valeur(s).
 - ▶ Cliquez sur Ajouter un sous-total pour afficher la boîte de dialogue Définir le sous-total.
 - ▶ Saisissez Inférieur à 45 dans la zone de texte Etiquette.
 - ▶ Sélectionnez (cochez) Masquer les catégories des sous-totaux dans le tableau.
 - ▶ Cliquez ensuite sur Poursuivre.
- Cette action insère une ligne contenant le sous-total des trois premières tranches d'âge.
- ▶ Sélectionnez 6.00 dans la liste Valeur(s).
 - ▶ Cliquez sur Ajouter un sous-total pour afficher la boîte de dialogue Définir le sous-total.
 - ▶ Saisissez 45 ou plus dans la zone de texte Etiquette.
 - ▶ Sélectionnez (cochez) Masquer les modalités de sous-total.
 - ▶ Cliquez ensuite sur Poursuivre.
 - ▶ Pour inclure un total avec les sous-totaux, cochez la case Total dans le groupe Présenter.

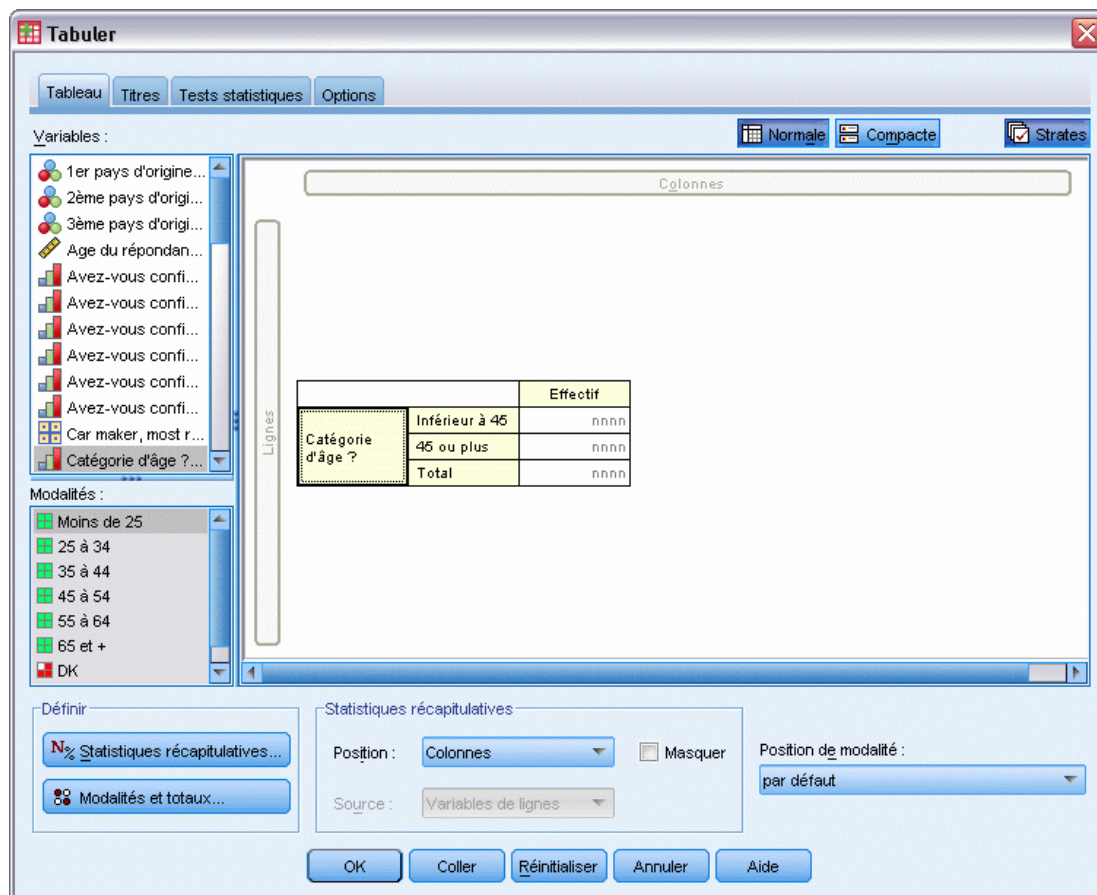
Figure 5-14
Masquage des modalités de sous-total



- ▶ Cliquez sur Appliquer.

Le canevas montre que les sous-totaux seront affichés, mais que les modalités définissant les sous-totaux seront exclues.

Figure 5-15
Canevas contenant les sous-totaux, sans les modalités de sous-total



- Cliquez sur OK pour générer ce tableau.

Figure 5-16
Tableau contenant uniquement les sous-totaux et les totaux

		Effectif
Catégories d'âges	Moins de 45	1548
	45 ou plus	1280
	Total	2828

Sous-totaux d'une variable de strate

Tout comme les totaux, les sous-totaux des variables de strate apparaissent sous la forme de strates distinctes dans le tableau. Les sous-totaux sont principalement traités comme des modalités. Chaque modalité est une strate distincte dans le tableau et l'ordre d'affichage des modalités de

strate est déterminé par l'ordre des modalités dans la boîte de dialogue Modalités et totaux, y compris la position d'affichage des modalités de sous-totaux.

Les catégories calculées pour les variables qualitatives

Vous pouvez inclure des catégories calculées dans les tableaux personnalisés. Il s'agit de nouvelles catégories calculées à partir de catégories de la même variable au niveau de l'imbrication selon n'importe quelle dimension : ligne, colonne ou strate. Par exemple, vous pouvez inclure une catégorie calculée qui montre la différence entre deux catégories.

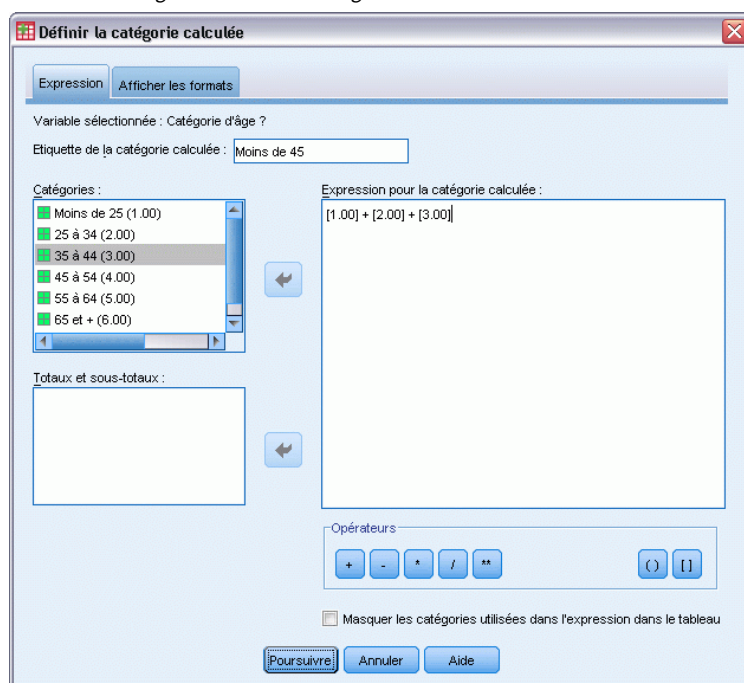
Fichier de données d'exemple

Les exemples de ce chapitre utilisent le fichier de données *survey_sample.sav*. Pour plus d'informations, reportez-vous à la section Fichiers d'exemple dans l'annexe A sur p. 197.

Catégorie calculée simple

- ▶ A partir du menu, sélectionnez :
Analyse > Tableaux > Tabuler...

Figure 6-1
Boîte de dialogue Définir la catégorie calculée



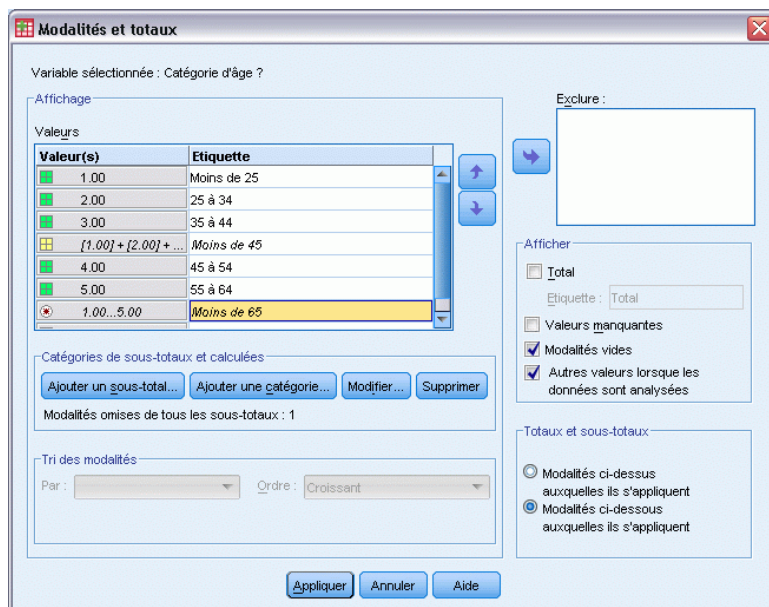
- ▶ Dans le générateur de tableaux, faites glisser la variable *Tranche d'âge* de la liste des variables vers la zone Lignes du panneau de canevas.

- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Tranche d'âge* et choisissez Modalités et totaux dans le menu contextuel.
- ▶ Sélectionnez 3.00 dans la liste Valeur(s).
- ▶ Cliquez sur Ajouter une catégorie pour afficher la boîte de dialogue Définir la catégorie calculée.
- ▶ Saisissez Inférieur à 45 dans la zone de texte Etiquette de la catégorie calculée.
- ▶ Sélectionnez Inférieur à 25 (1,00) dans la liste des catégories et cliquez sur le bouton fléché pour copier cette valeur dans la zone de texte Expression pour la catégorie calculée. [1] s'affiche dans l'expression.
- ▶ Cliquez sur le bouton plus (+) du clavier numérique de la boîte de dialogue (ou appuyez sur la touche + sur votre clavier).
- ▶ Sélectionnez 25 à 34 (2,00) dans la liste des catégories et cliquez sur le bouton fléché pour copier cette valeur dans la zone de texte Expression pour la catégorie calculée.
- ▶ Cliquez sur le bouton plus (+) du clavier numérique de la boîte de dialogue (ou appuyez sur la touche + sur votre clavier).
- ▶ Sélectionnez 35 à 44 (3,00) dans la liste des catégories et cliquez sur le bouton fléché pour copier cette valeur dans la zone de texte Expression pour la catégorie calculée.
- ▶ Cliquez ensuite sur Poursuivre.
Cette action insère une ligne contenant le sous-total des trois premières tranches d'âge.
- ▶ Sélectionnez 5.00 dans la liste Valeur(s).
- ▶ Cliquez sur Ajouter un sous-total pour afficher la boîte de dialogue Définir le sous-total.
- ▶ Saisissez Inférieur à 65 dans la zone de texte Etiquette.

- Cliquez ensuite sur Poursuivre.

Cette action insère une ligne contenant le sous-total des cinq premières catégories.

Figure 6-2
Boîte de dialogue Modalités et totaux



- Dans le générateur de tableaux, cliquez sur Appliquer, puis sur OK pour créer le tableau.

Figure 6-3
Les catégories calculées avec des sous-totaux

	Effectif
Catégorie d'âge ?	
Moins de 25	242
25 à 34	627
35 à 44	679
Moins de 45	1548
45 à 54	481
55 à 64	320
Moins de 65	2349
65 et +	479

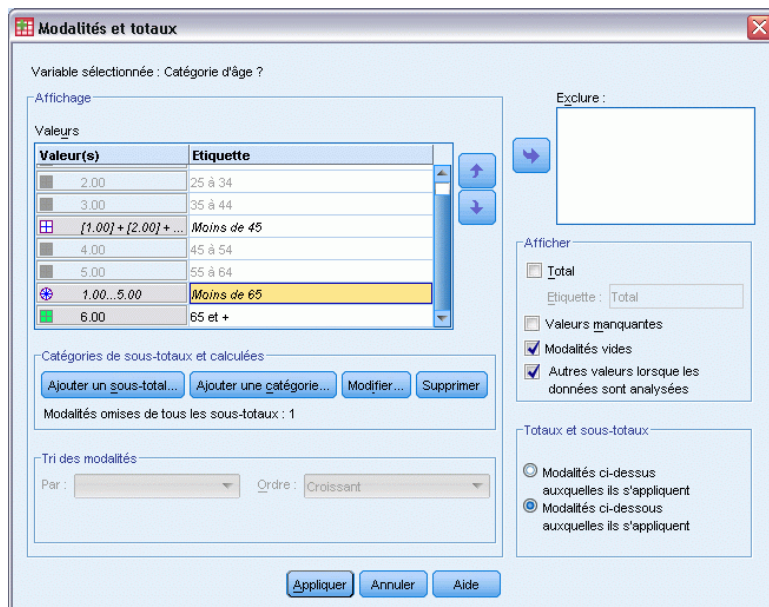
Le tableau comprend une catégorie calculée (*Inférieur à 45*) et un sous-total (*Inférieur à 65*). Le sous-total inclut les catégories comprises dans la catégorie calculée. Il n'est pas possible de créer le même tableau uniquement à l'aide des sous-totaux, car ces derniers ne peuvent pas partager les mêmes catégories.

Masquer les catégories dans une catégorie calculée

Comme pour les sous-totaux, vous avez la possibilité de supprimer l'affichage des catégories utilisées dans une expression de catégorie calculée, et de n'afficher que la catégorie calculée. L'exemple suivant est construit sur le précédent.

- ▶ A partir du menu, sélectionnez :
Analyse > Tableaux > Tabuler...
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Tranche d'âge* et choisissez Modalités et totaux dans le menu contextuel.
- ▶ Dans la liste Valeur(s), sélectionnez la catégorie calculée *Inférieur à 45*.
- ▶ Cliquez sur Edition pour afficher la boîte de dialogue Définir la catégorie calculée.
- ▶ Sélectionnez Masquer les catégories utilisées dans l'expression dans le tableau.
- ▶ Cliquez ensuite sur Poursuivre.
- ▶ Dans la liste Valeur(s), sélectionnez le sous-total *Inférieur à 65*.
- ▶ Cliquez sur Edition pour afficher la boîte de dialogue Définir le sous-total.
- ▶ Sélectionnez Masquer les catégories des sous-totaux dans le tableau.
- ▶ Cliquez ensuite sur Poursuivre.

Figure 6-4
Boîte de dialogue Modalités et totaux affichant les catégories masquées



- ▶ Dans le générateur de tableaux, cliquez sur Appliquer, puis sur OK pour créer le tableau.

Figure 6-5
Catégories calculées avec sous-totaux et catégories masquées

Catégorie d'âge ?		Effectif
Moins de 45		1548
Moins de 65		2349
65 et +		479

Comme dans l'exemple précédent, le tableau comprend une catégorie calculée et un sous-total. Toutefois, dans ce cas les catégories sont masquées pour chacun d'entre eux et seuls les totaux sont affichés.

Référencement des sous-totaux dans une catégorie calculée

Vous pouvez inclure des sous-totaux dans une expression de catégorie calculée.

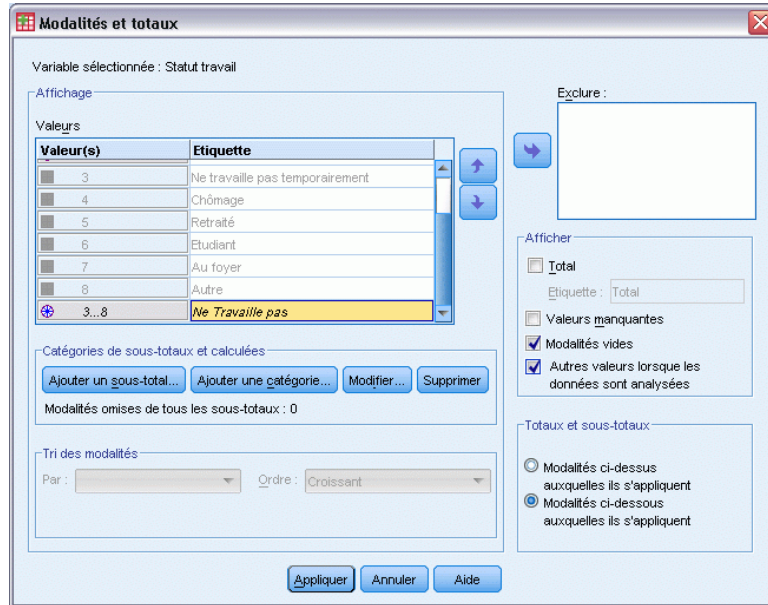
- ▶ A partir du menu, sélectionnez :
Analyse > Tableaux > Tabuler...
- ▶ Cliquez sur Restaurer pour effacer les paramètres précédents dans le générateur de tableaux.
- ▶ Dans le générateur de tableaux, faites glisser la variable *Situation d'activité* de la liste des variables vers la zone Lignes du panneau de canevas.
- ▶ Faites glisser la variable *Situation familiale* de la liste des variables vers la zone Colonnes.
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Situation d'activité* et choisissez Modalités et totaux dans le menu contextuel.
- ▶ Sélectionnez 2 dans la liste Valeur(s).
- ▶ Cliquez sur Ajouter un sous-total pour afficher la boîte de dialogue Définir le sous-total.
- ▶ Saisissez Travail dans la zone de texte Etiquette.
- ▶ Sélectionnez Masquer les catégories des sous-totaux dans le tableau.
- ▶ Cliquez ensuite sur Poursuivre.
Cette action insère une ligne contenant le sous-total des deux premières catégories d'activité.
- ▶ Sélectionnez 8 dans la liste Valeur(s).
- ▶ Cliquez sur Ajouter un sous-total pour afficher la boîte de dialogue Définir le sous-total.
- ▶ Saisissez Sans travail dans la zone de texte Etiquette.
- ▶ Sélectionnez Masquer les catégories des sous-totaux.

- Cliquez ensuite sur Poursuivre.

Cette action insère une ligne contenant le sous-total des autres catégories d'activité.

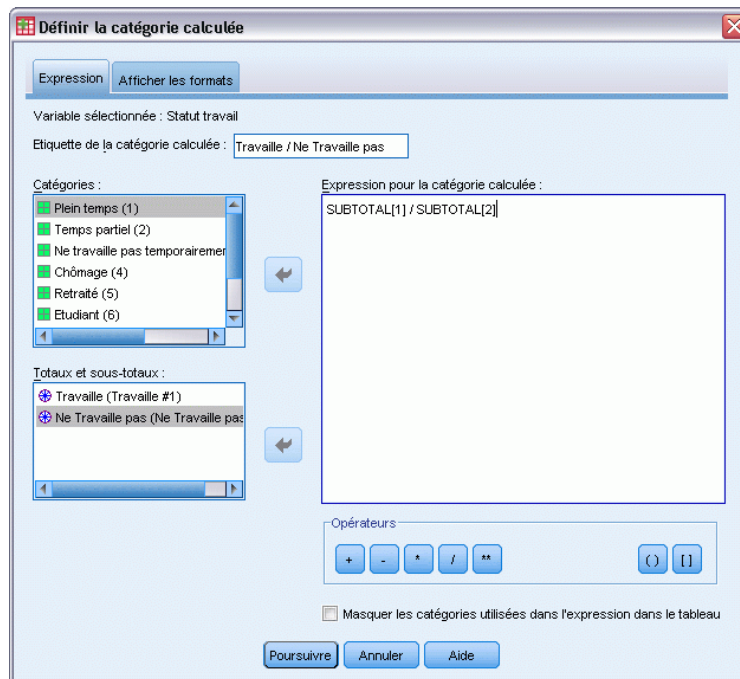
Figure 6-6

Boîte de dialogue Modalités et totaux



- Dans la liste des valeurs, sélectionnez le sous-total *Sans travail*.
- Cliquez sur Ajouter une catégorie pour afficher la boîte de dialogue Définir la catégorie calculée.
- Saisissez Travail / Sans travail dans la zone de texte Etiquette de la catégorie calculée.
- Sélectionnez Travail (Travail n°1) dans la liste des totaux et sous-totaux et cliquez sur le bouton fléché pour copier cette valeur dans la zone de texte Expression pour la catégorie calculée.
- Cliquez sur le bouton de division (/) du clavier numérique de la boîte de dialogue (ou appuyez sur la touche / sur votre clavier).
- Sélectionnez Sans travail (Sans travail n°2) dans la liste des totaux et sous-totaux et cliquez sur le bouton fléché pour copier cette valeur dans la zone de texte Expression pour la catégorie calculée.

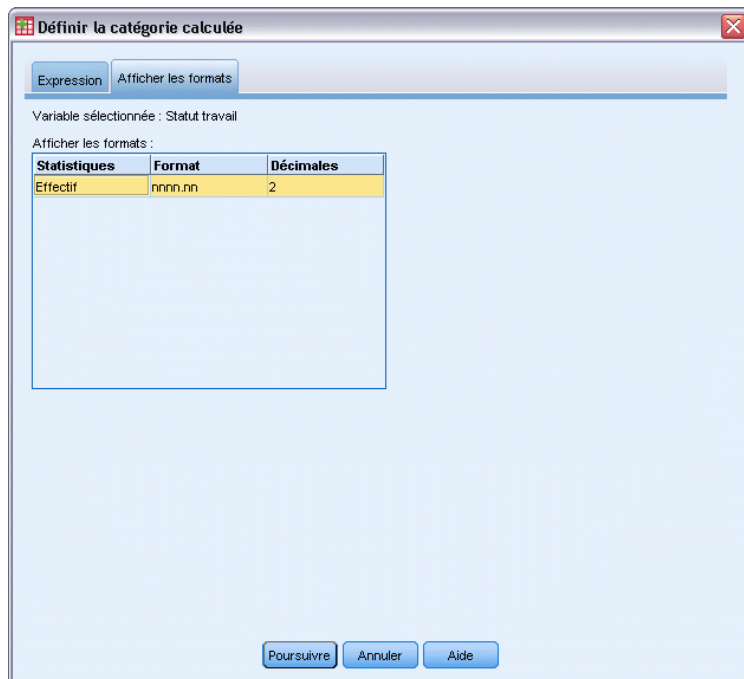
Figure 6-7
Boîte de dialogue Définir la catégorie calculée



Par défaut, la catégorie calculée utilise le même format que la statistique de la variable, Effectif dans ce cas. Parce que nous voulons afficher les décimales résultant de la division dans l'expression de la catégorie calculée, et que le format par défaut de la statistique Effectif n'affiche pas les décimales, nous devons modifier le format.

- ▶ Cliquez sur l'onglet Formats d'affichage.
- ▶ Changez le paramètre Décimales d'Effectif sur 2.

Figure 6-8
Onglet Formats d'affichage dans la boîte de dialogue Définir la catégorie calculée



- ▶ Cliquez ensuite sur Poursuivre.
- ▶ Dans le générateur de tableaux, cliquez sur Appliquer, puis sur OK pour créer le tableau.

Figure 6-9
Catégorie calculée affichant le rapport des sous-totaux

		Statut marital				
		Marrié	Veuf/veuve	Divorcé	Séparé	Célibataire
		Effectif	Effectif	Effectif	Effectif	Effectif
Statut travail	Travaille	916	64	330	67	494
	Ne Travaille pas	429	219	116	26	169
	Travaille / Ne Travaille pas	2.14	.29	2.84	2.58	2.92

Le tableau comprend deux sous-totaux et une catégorie calculée. La catégorie calculée affiche le rapport des sous-totaux, et vous permet de comparer facilement les groupes représentés par chaque sous-total. Le rapport entre les répondants veufs avec travail et les répondants veufs sans travail est plus faible que celui des autres groupes.. Il existe aussi un rapport légèrement plus faible de répondants mariés, résultant peut-être de conjoints qui ont arrêtés de travailler pour s'occuper des enfants.

Utilisation des catégories calculées pour afficher des sous-totaux non-exhaustifs

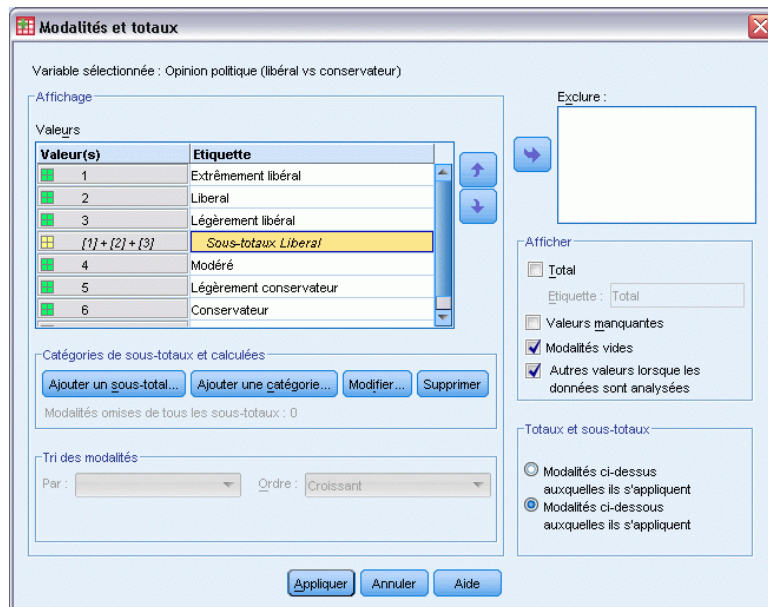
Les sous-totaux sont exhaustifs. C'est à dire que tous les sous-totaux d'un tableau comprennent toutes les valeurs au-dessus ou au-dessous de leur position dans le tableau. Les catégories calculées, elles, ne sont pas exhaustives et vous permettent d'ajouter et de mélanger les catégories d'un tableau.

- ▶ A partir du menu, sélectionnez :
Analyse > Tableaux > Tabuler...
- ▶ Cliquez sur Restaurer pour effacer les paramètres précédents dans le générateur de tableaux.
- ▶ Dans le générateur de tableaux, faites glisser la variable *Se considère libéral ou conservateur* de la liste des variables vers la zone Lignes du panneau de canevas.
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur la variable *Se considère libéral ou conservateur* et choisissez Modalités et totaux dans le menu contextuel.
- ▶ Sélectionnez 3 dans la liste Valeur(s).
- ▶ Cliquez sur Ajouter une catégorie pour afficher la boîte de dialogue Définir la catégorie calculée.
- ▶ Saisissez `Sous-total Libéraux` dans la zone de texte Etiquette de la catégorie calculée.
Remarque : il y a quatre espaces avant le texte. Ces espaces sont utilisés pour l'indentation du tableau des résultats.
- ▶ Sélectionnez Extrêmement libéral (1) dans la liste des catégories et cliquez sur le bouton fléché pour copier cette valeur dans la zone de texte Expression pour la catégorie calculée.
- ▶ Cliquez sur le bouton plus (+) du clavier numérique de la boîte de dialogue (ou appuyez sur la touche + sur votre clavier).
- ▶ Sélectionnez Libéral (2) dans la liste des catégories et cliquez sur le bouton fléché pour copier cette valeur dans la zone de texte Expression pour la catégorie calculée.
- ▶ Cliquez sur le bouton plus (+) du clavier numérique de la boîte de dialogue (ou appuyez sur la touche + sur votre clavier).
- ▶ Sélectionnez Légèrement libéral (3) dans la liste des catégories et cliquez sur le bouton fléché pour copier cette valeur dans la zone de texte Expression pour la catégorie calculée.

- Cliquez sur Poursuivre.

Cette action insère une ligne contenant le sous-total des catégories Libéraux.

Figure 6-10
Boîte de dialogue Modalités et totaux



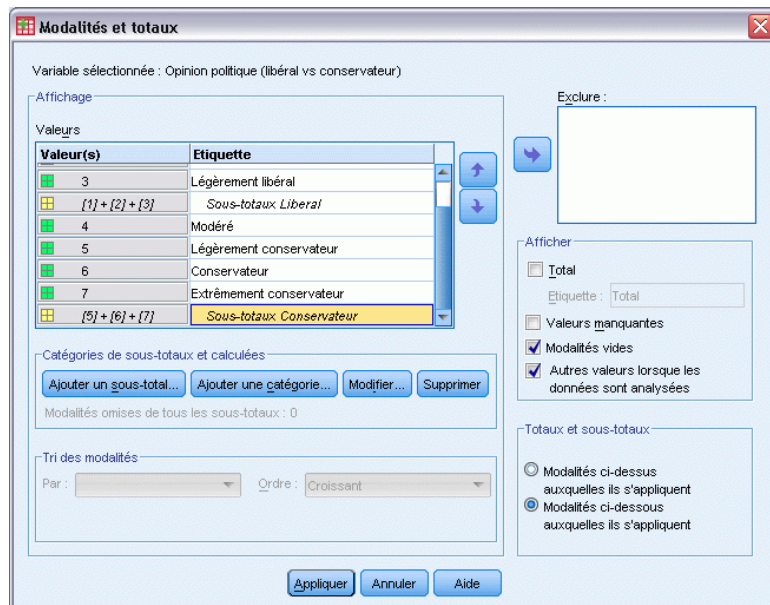
- Sélectionnez 7 dans la liste Valeur(s).
- Cliquez sur Ajouter une catégorie pour afficher la boîte de dialogue Définir la catégorie calculée.
- Saisissez `Sous-total Conservateurs` dans la zone de texte Etiquette de la catégorie calculée. Remarque : il y a quatre espaces avant le texte. Ces espaces sont utilisés pour l'indentation du tableau des résultats.
- Sélectionnez Légèrement conservateur (5) dans la liste des catégories et cliquez sur le bouton fléché pour copier cette valeur dans la zone de texte Expression pour la catégorie calculée.
- Cliquez sur le bouton plus (+) du clavier numérique de la boîte de dialogue (ou appuyez sur la touche + sur votre clavier).
- Sélectionnez Conservateur (6) dans la liste des catégories et cliquez sur le bouton fléché pour copier cette valeur dans la zone de texte Expression pour la catégorie calculée.
- Cliquez sur le bouton plus (+) du clavier numérique de la boîte de dialogue (ou appuyez sur la touche + sur votre clavier).
- Sélectionnez Extrêmement conservateur (7) dans la liste des catégories et cliquez sur le bouton fléché pour copier cette valeur dans la zone de texte Expression pour la catégorie calculée.

- Cliquez sur Poursuivre.

Cette action insère une ligne contenant le sous-total des catégories Conservateurs.

Figure 6-11

Boîte de dialogue Modalités et totaux



- Dans le générateur de tableaux, cliquez sur Appliquer, puis sur OK pour créer le tableau.

Figure 6-12

Catégories calculées affichant les sous-totaux non-exhaustifs

	Effectif
Opinion politique (libéral vs conservateur)	
Extrêmement libéral	64
Liberal	357
Légèrement libéral	351
Sous-totaux Liberal	772
Modéré	986
Légèrement conservateur	432
Conservateur	415
Extrêmement conservateur	86
Sous-totaux Conservateur	933

Le tableau comprend deux catégories calculées qui n'incluent pas toutes les catégories du tableau. La catégorie *Modérés* n'est comprise dans aucune catégorie calculée. Il est impossible de créer le même tableau avec des sous-totaux car ils doivent être exhaustifs.

Tableaux pour variables avec modalités partagées

Les enquêtes contiennent souvent de nombreuses questions caractérisées par un ensemble commun de réponses possibles. Notre échantillon d'enquête contient ainsi un certain nombre de variables relatives à la confiance accordée à divers services et institutions publics et privés, qui présentent toutes le même ensemble de modalités de réponses : 1 = *Beaucoup*, 2 = *Seulement certaines* et 3 = *Presque aucune*. Vous pouvez utiliser la technique de l'empilement pour afficher ces variables liées dans le même tableau et afficher les modalités de réponses partagées dans les colonnes du tableau. Ces fonctions sont également disponibles si vous utilisez des catégories calculées, à la condition que l'étiquette et l'expression de chaque catégorie calculée soient identiques dans toutes les variables.

Figure 7-1
Tableau de variables avec modalités partagées

	Enormém	Un peu	Pas du
Confiance dans les institutions bancaires et financières	490	1068	306
Confiance dans les multinationales	500	1078	243
Confiance dans l'éducation	511	1055	315
Confiance dans la presse	176	878	808
Confiance dans la médecine	844	864	167
Confiance dans la télévision	196	936	744

Remarque : Dans la version précédente de la fonction Tableaux personnalisés, cet élément était nommé "tableau de fréquences".

Fichier de données d'exemple

Les exemples de ce chapitre utilisent le fichier de données *survey_sample.sav*. [Pour plus d'informations, reportez-vous à la section Fichiers d'exemple dans l'annexe A sur p. 197.](#)

Tous les exemples fournis ici affichent des étiquettes de variable dans des boîtes de dialogue ; ces étiquettes sont triées par ordre alphabétique. Les propriétés d'affichage de la liste des variables sont paramétrées dans l'onglet Général de la boîte de dialogue Options (menu Modifier, Options).

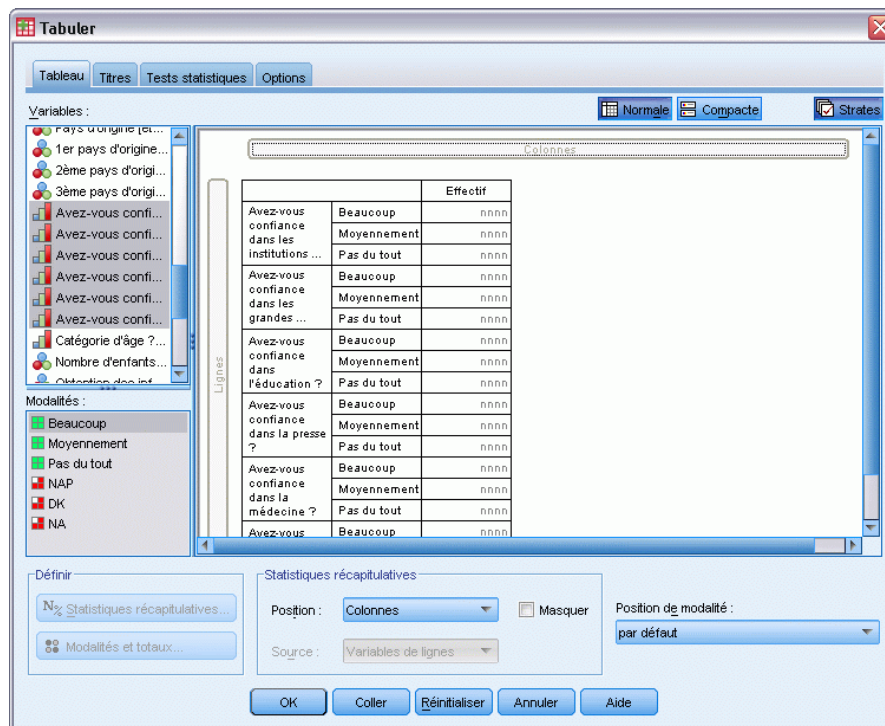
Tableau d'effectifs

- ▶ A partir du menu, sélectionnez :
Analyse > Tableaux > Tabuler...
- ▶ Dans la liste des variables du générateur de tableaux, cliquez sur *Confiance en les banques*, puis, tout en maintenant la touche Maj enfoncée, cliquez sur *Confiance en la télévision* pour

sélectionner toutes les variables commençant par le terme “ confiance “. (*Remarque* : Ceci implique que les étiquettes de variable soient affichées dans l’ordre alphabétique, et non dans l’ordre du fichier, dans la liste des variables.)

- Faites glisser les six variables de confiance vers la zone Lignes du panneau de canevas.

Figure 7-2
Variables de confiance empilées en lignes



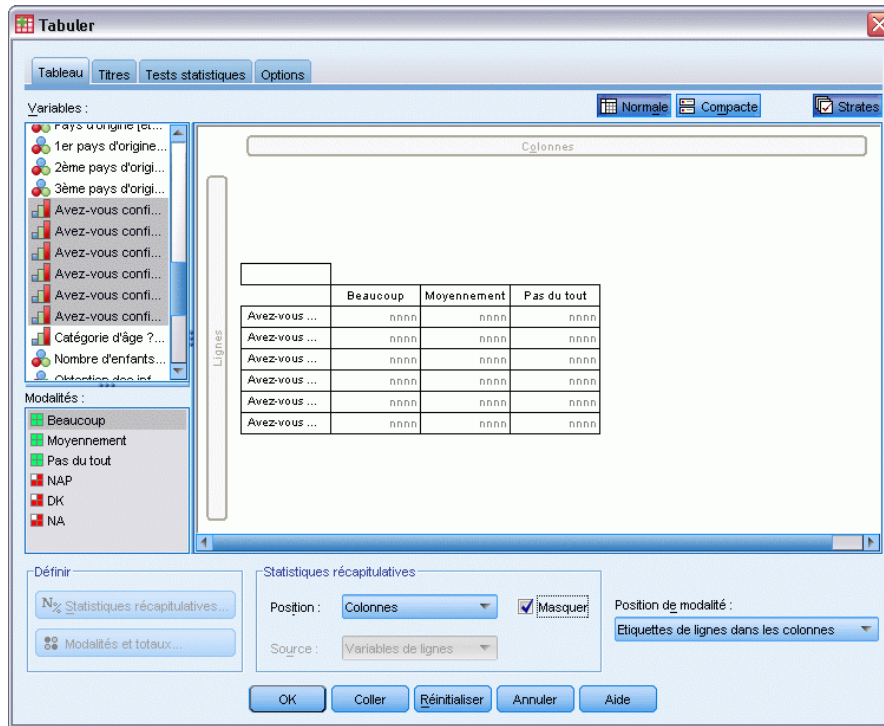
Les variables sont ainsi empilées dans la dimension de ligne. Par défaut, les étiquettes de modalité de chaque variable apparaissent également dans les lignes, ce qui donne un tableau très long et étroit (6 variables x 3 modalités = 18 lignes). Cependant, puisque les six variables partagent les mêmes étiquettes de modalité définies (étiquettes de valeurs), vous pouvez placer les étiquettes de modalité dans la dimension de colonne.

- Dans la liste déroulante Emplacement de la modalité, sélectionnez Etiquettes de ligne dans les colonnes.

Le tableau comporte désormais seulement six lignes, une pour chaque variable empilée, et les modalités définies deviennent les colonnes du tableau.

- Avant de créer le tableau, sélectionnez Masquer pour Position dans le groupe Statistiques récapitulatives, puisque l’étiquette de statistique récapitulative *Effectif* n’est pas vraiment nécessaire.

Figure 7-3
Étiquettes de modalité disposées en colonnes



- Cliquez sur OK pour créer le tableau.

Figure 7-4
Tableau de variables en ligne empilées avec étiquettes de modalité partagées disposées en colonnes

	Enormém	Un peu	Pas du
Confiance dans les institutions bancaires et financières	490	1068	306
Confiance dans les multinationales	500	1078	243
Confiance dans l'éducation	511	1055	315
Confiance dans la presse	176	878	808
Confiance dans la médecine	844	864	167
Confiance dans la télévision	196	936	744

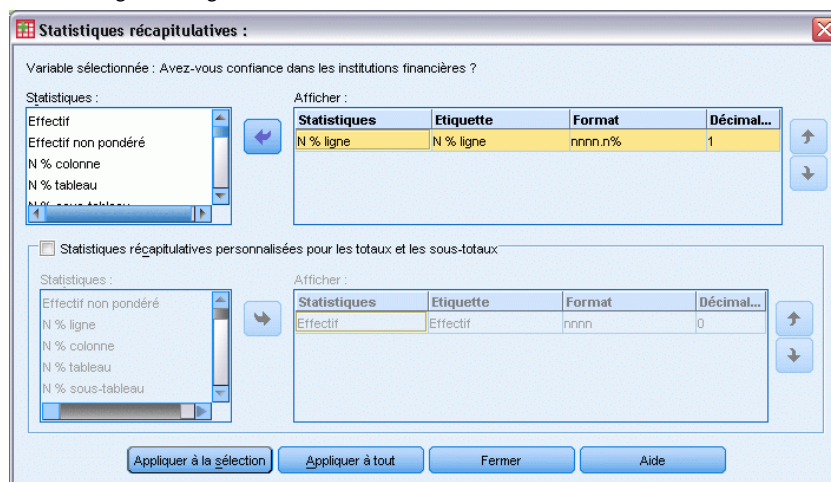
Au lieu d'afficher les variables dans les lignes et les modalités dans les colonnes, vous pouvez créer un tableau dans lequel les variables sont empilées dans les colonnes et les modalités affichées dans les lignes. Cette solution est préférable si le nombre de modalités est supérieur au nombre de variables, ce qui n'est pas le cas dans notre exemple, puisque le nombre de variables est supérieur au nombre de modalités.

Tableau de pourcentages

Pour un tableau dans lequel les variables sont empilées dans les lignes et les modalités affichées dans les colonnes, il semble pertinent (ou du moins compréhensible) de faire apparaître les pourcentages en ligne. (Pour un tableau dans lequel les variables sont empilées dans les colonnes et les modalités affichées dans les lignes, les pourcentages en colonne peuvent sembler plus appropriés.)

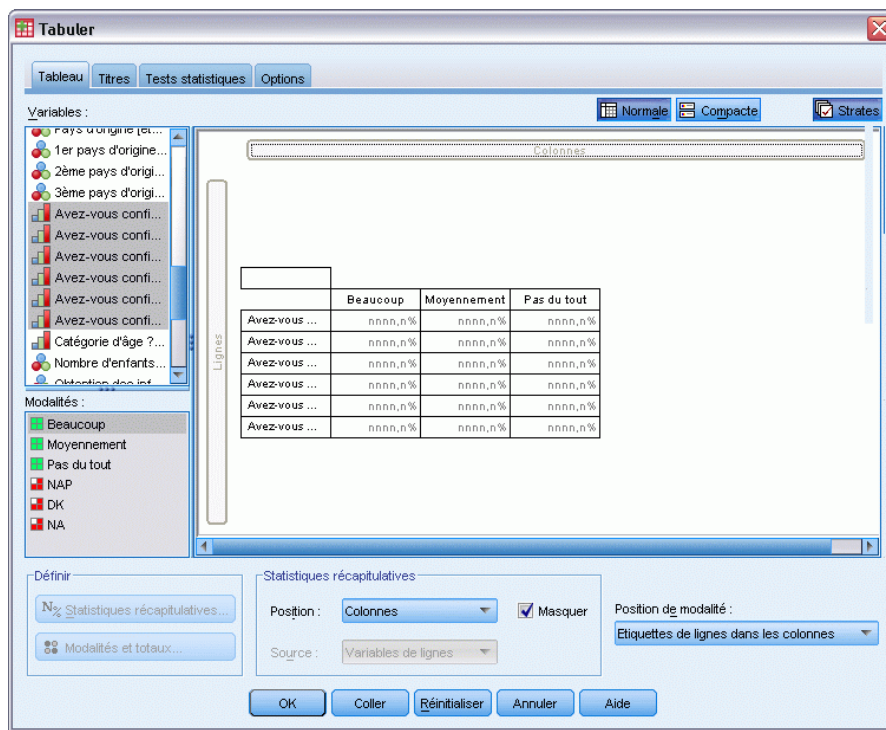
- ▶ Ouvrez à nouveau le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur une des variables de confiance dans l'aperçu du tableau et choisissez Statistiques récapitulatives dans le menu contextuel.
- ▶ Dans la liste Statistiques, sélectionnez N % ligne et cliquez sur le bouton fléché pour ajouter cet élément à la liste Afficher.
- ▶ Cliquez sur une cellule de la ligne *Effectif* dans la liste Afficher, puis sur le bouton fléché pour enlever cet élément de cette liste et le replacer dans la liste Statistiques.

Figure 7-5
Pourcentages en ligne sélectionnés.



- ▶ Cliquez sur Appliquer à tout pour appliquer la modification de la statistique récapitulative à toutes les variables empilées du tableau.

Figure 7-6
Pourcentages en ligne dans l'aperçu du tableau du panneau de canevas



Remarque : Si l'aperçu du tableau ne ressemble pas à la figure ci-dessus, vous avez probablement cliqué sur Appliquer à la sélection au lieu de cliquer sur Appliquer à tout, ce qui applique la nouvelle statistique récapitulative uniquement à la variable sélectionnée. Dans cet exemple, ceci donne deux colonnes pour chaque modalité : une dans laquelle des valeurs de substitution apparaissent pour l'effectif pour toutes les autres variables et une dans laquelle une valeur de substitution apparaît pour le pourcentage en ligne de la variable sélectionnée. Il s'agit exactement du tableau qui serait généré, mais *pas* de celui souhaité dans cet exemple.

- Cliquez sur OK pour créer le tableau.

Figure 7-7
Tableau des pourcentages en ligne pour les variables empilées dans les lignes, avec affichage des modalités dans les colonnes

	Enormément	Un peu	Pas du tout
Confiance dans les institutions bancaires et financières	26,3%	57,3%	16,4%
Confiance dans les	27,5%	59,2%	13,3%
Confiance dans	27,2%	56,1%	16,7%
Confiance dans la presse	9,5%	47,2%	43,4%
Confiance dans la	45,0%	46,1%	8,9%
Confiance dans la	10,4%	49,9%	39,7%

Remarque : Vous pouvez inclure un nombre quelconque de statistiques récapitulatives dans un tableau de variables avec modalités partagées. Par souci de simplicité, nos exemples illustrent une seule statistique à la fois.

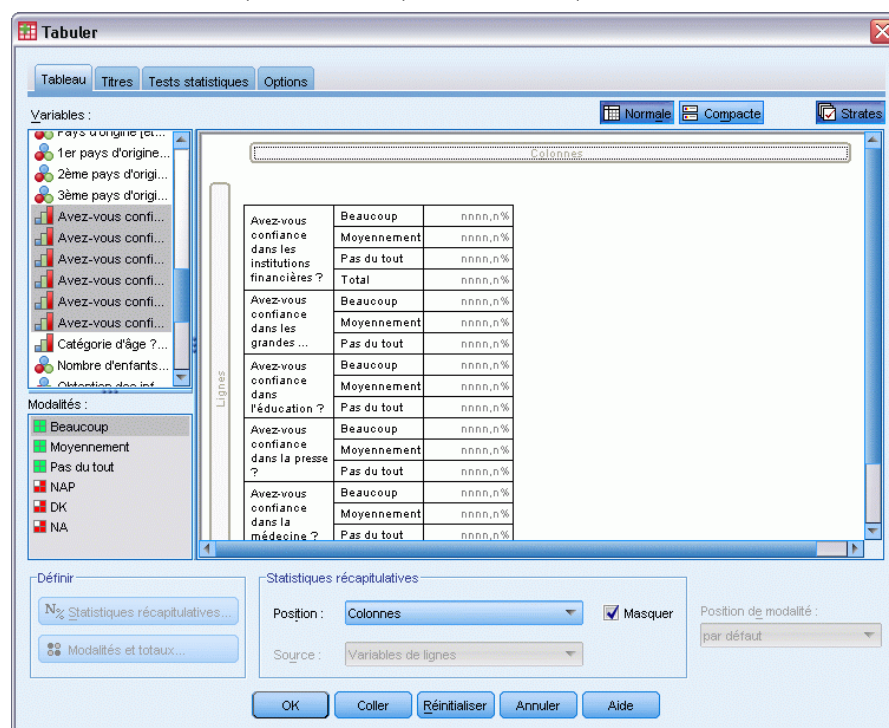
Contrôle des totaux et des modalités

Vous pouvez créer des tableaux avec des modalités situées dans la dimension opposée à celle des variables uniquement si toutes les variables du tableau comportent les mêmes modalités, affichées dans le même ordre. Ceci inclut les totaux, les sous-totaux et tout autre ajustement de modalité que vous effectuez. Par conséquent, toute modification apportée dans la boîte de dialogue Modalités et totaux doit être effectuée pour toutes les variables du tableau partageant les modalités.

- ▶ Ouvrez à nouveau le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur la première variable de confiance dans l'aperçu du tableau et choisissez Modalités et totaux dans le menu contextuel.
- ▶ Dans la boîte de dialogue Modalités et totaux, sélectionnez Total, puis cliquez sur Appliquer.

Figure 7-8

Ces résultats ne sont probablement pas ceux escomptés.

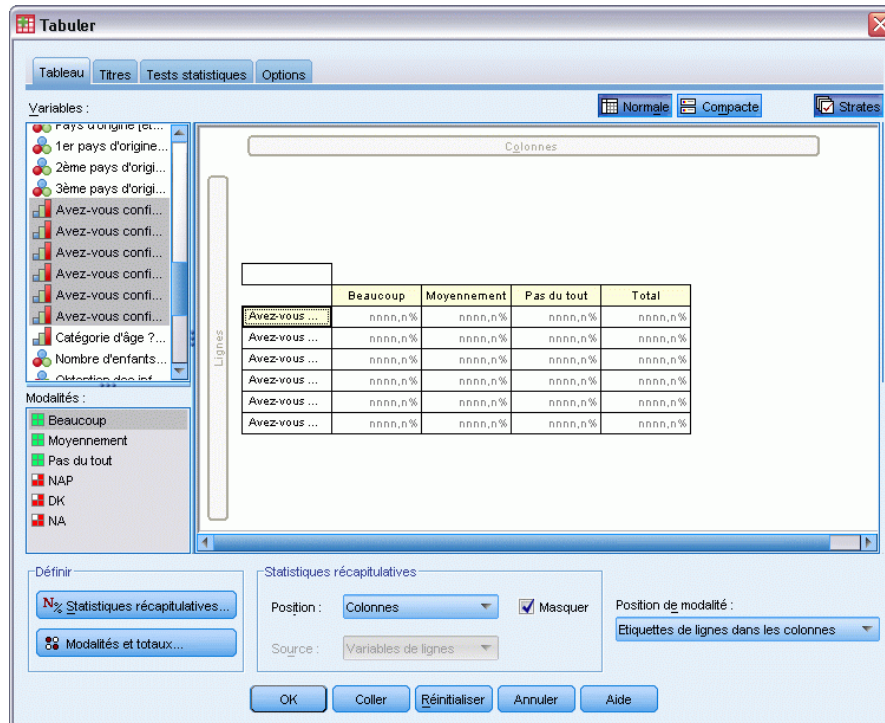


La première chose que vous pouvez remarquer est que les étiquettes de modalité ont été déplacées des colonnes vers les lignes. Vous pouvez également noter que le contrôle Emplacement de la modalité est désormais désactivé. En effet, les variables ne partagent plus exactement le même ensemble de "modalités". Une des variables comporte désormais une modalité Total.

- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur une des variables de confiance, puis sélectionnez Sélectionner toutes les variables en ligne dans le menu contextuel ou, tout en maintenant la touche Ctrl enfoncée, cliquez sur chaque variable empilée dans le panneau de canevas jusqu'à ce qu'elles soient toutes sélectionnées (vous devrez peut-être parcourir le panneau ou agrandir la fenêtre du générateur de tableaux).

- ▶ Cliquez sur Modalités et totaux dans le groupe Définir.
- ▶ Si l'élément Total n'est pas déjà sélectionné dans la boîte de dialogue Modalités et totaux, sélectionnez-le, puis cliquez sur Appliquer.
- ▶ La liste déroulante Emplacement de la modalité doit de nouveau être activée, puisque toutes les variables comportent désormais la modalité Total supplémentaire ; vous pouvez sélectionner Etiquettes de ligne dans les colonnes.

Figure 7-9
Modalités et totaux dans les colonnes



- ▶ Cliquez sur OK pour créer le tableau.

Figure 7-10
Tableau des pourcentages en ligne pour les variables empilées dans les lignes, avec affichage des modalités et des totaux dans les colonnes

	Enormément	Un peu	Pas du tout	Total
Confiance dans les institutions bancaires et financières	26,3%	57,3%	16,4%	100,0%
Confiance dans les	27,5%	59,2%	13,3%	100,0%
Confiance dans	27,2%	56,1%	16,7%	100,0%
Confiance dans la presse	9,5%	47,2%	43,4%	100,0%
Confiance dans la	45,0%	46,1%	8,9%	100,0%
Confiance dans la	10,4%	49,9%	39,7%	100,0%

Emboîtement dans des tableaux comportant des modalités partagées

Dans les tableaux emboîtés, les variables empilées comportant des modalités partagées doivent être situées au niveau d'emboîtement le plus interne dans leur dimension si vous souhaitez afficher les étiquettes de modalité dans la dimension opposée.

- Ouvrez à nouveau le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- Faites glisser la variable *Sexe* de la liste des variables vers la gauche de la zone Lignes.

Figure 7-11
Variables emboîtées avec modalités partagées

The screenshot shows the Tabuler software interface. On the left, a list of variables includes 'Sexe [sex]'. Below it, the 'Modalités' section shows 'Homme' and 'Femme' selected. The main area displays a table with 'Lignes' on the left and 'Colonnes' at the top. The table is nested, with 'Sexe' as the primary row category, split into 'Homme' and 'Femme'. Each of these is further divided into four sub-categories: 'Beaucoup', 'Moyennement', 'Pas du tout', and 'Total'. Each cell in the table contains a percentage value, represented as 'nnnn,n%'. At the bottom, the 'Statistiques récapitulatives' section has 'Position' set to 'Colonnes' and 'Position de modalité' set to 'Etiquettes de lignes dans les colonnes'. Buttons for 'OK', 'Coller', 'Réinitialiser', 'Annuler', and 'Aide' are visible at the bottom of the window.

Les variables empilées comportant des modalités partagées sont désormais emboîtées dans les modalités de la variable *Sexe* dans l'aperçu du tableau.

- Faites glisser la variable *Sexe* vers la droite de l'une des variables de confiance empilées dans l'aperçu du tableau.

Figure 7-12
Autre exemple de résultats non escomptés

The screenshot shows the Tabuler software interface. The main window displays a pivot table with the following structure:

Avez-vous confiance dans l'éducation ?		Sexe	Homme	nnnn
Beaucoup	Sexe		Femme	nnnn
	Sexe		Homme	nnnn
	Sexe		Femme	nnnn
Moyennement	Sexe		Homme	nnnn
	Sexe		Femme	nnnn
	Sexe		Homme	nnnn
Pas du tout	Sexe		Homme	nnnn
	Sexe		Femme	nnnn
	Sexe		Homme	nnnn
Total	Sexe		Homme	nnnn
	Sexe		Femme	nnnn
Avez-vous confiance dans la médecine ?		Beaucoup		nnnn,n%
Avez-vous confiance dans la médecine ?		Moyennement		nnnn,n%
Avez-vous confiance dans la médecine ?		Pas du tout		nnnn,n%
Avez-vous confiance dans la médecine ?		Total		nnnn,n%
Avez-vous confiance dans la presse ?		Beaucoup		nnnn,n%
Avez-vous confiance dans la presse ?		Moyennement		nnnn,n%
Avez-vous confiance dans la presse ?		Pas du tout		nnnn,n%
Avez-vous confiance dans la presse ?		Total		nnnn,n%
Avez-vous confiance dans les grandes		Beaucoup		nnnn,n%
Avez-vous confiance dans les grandes		Moyennement		nnnn,n%
Avez-vous confiance dans les grandes		Pas du tout		nnnn,n%

The interface includes a 'Variables' list on the left, a 'Modalités' section with 'Homme' and 'Femme' selected, and a 'Statistiques récapitulatives' section at the bottom with 'Position' set to 'Colonnes' and 'Masquer' checked. Buttons for 'OK', 'Coller', 'Réinitialiser', 'Annuler', and 'Aide' are visible at the bottom.

De nouveau, les étiquettes de modalité ont été replacées dans la dimension de ligne et le contrôle Emplacement de la modalité est désactivé. En outre, la variable *Sexe* est désormais emboîtée dans l'une des variables empilées, alors que les autres variables empilées ne contiennent aucune variable emboîtée. Vous pouvez ajouter la variable *Sexe* comme variable emboîtée à chaque variable empilée, mais le déplacement des étiquettes des lignes vers les colonnes aboutirait à l'affichage des étiquettes de modalité de la variable *Sexe* dans les colonnes, au lieu des étiquettes de modalité des variables empilées comportant des modalités partagées. En effet, la variable *Sexe* deviendrait la variable emboîtée la plus interne ; or, la modification de la position d'une modalité s'applique systématiquement à la variable emboîtée la plus interne.

Statistiques récapitulatives

Les **statistiques récapitulatives** comprennent une large variété d'éléments, des effectifs simples pour les variables qualitatives aux mesures de dispersion, comme l'erreur standard de la moyenne pour les variables d'échelle. Elles n'incluent *pas* les tests de signification disponibles dans l'onglet Statistiques de test de la boîte de dialogue Tableaux personnalisés. [Pour plus d'informations, reportez-vous à la section Statistiques de test dans le chapitre 10 sur p. 145.](#)

Les statistiques récapitulatives pour les variables qualitatives et les vecteurs multiréponses incluent des effectifs et une grande variété de calculs de pourcentage, notamment :

- Pourcentages en ligne
- Pourcentages en colonne
- Pourcentages en sous-tableau
- Pourcentages en tableau
- Pourcentages en N valide

Outre les statistiques récapitulatives disponibles pour les variables qualitatives, les statistiques récapitulatives pour les variables d'échelle et d'autres statistiques en total pour les variables qualitatives incluent :

- Moyenne
- Médiane
- Centiles
- Somme
- Ecart type
- Intervalle
- Valeurs minimale et maximale

Des statistiques récapitulatives supplémentaires sont disponibles pour les vecteurs multiréponses. [Pour plus d'informations, reportez-vous à la section Effectifs, réponses, pourcentages et totaux dans le chapitre 11 sur p. 165.](#) Une liste complète des statistiques récapitulatives est également disponible. [Pour plus d'informations, reportez-vous à la section Statistiques récapitulatives dans le chapitre 2 sur p. 34.](#)

Fichier de données d'exemple

Les exemples de cette du chapitre utilisent le fichier de données *survey_sample.sav*. [Pour plus d'informations, reportez-vous à la section Fichiers d'exemple dans l'annexe A sur p. 197.](#)

Tous les exemples fournis ici affichent des étiquettes de variable dans des boîtes de dialogue ; ces étiquettes sont triées par ordre alphabétique. Les propriétés d'affichage de la liste des variables sont paramétrées dans l'onglet Général de la boîte de dialogue Options (menu Modifier, Options).

Variable source des statistiques récapitulatives

Les statistiques récapitulatives disponibles dépendent du niveau de mesure de leur variable source. La source des statistiques récapitulatives (variable sur laquelle sont basées les statistiques) est déterminée par les éléments suivants :

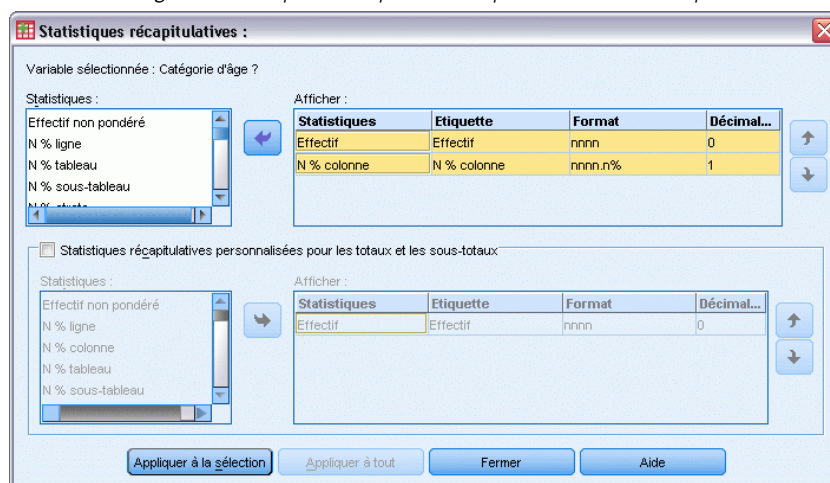
- **Niveau de mesure.** Si un tableau (ou une section dans un tableau de type empilé) contient une variable d'échelle, les statistiques récapitulatives sont basées sur cette dernière.
- **Ordre de sélection des variables.** La dimension source par défaut (ligne ou colonne) des statistiques pour les variables qualitatives est basée sur l'ordre dans lequel vous avez glissé les variables vers le panneau de canevas. Par exemple, si vous faites glisser d'abord une variable vers la zone Lignes, la dimension de ligne est la dimension source par défaut des statistiques.
- **Emboîtement :** Pour les variables qualitatives, les statistiques récapitulatives sont basées sur la variable la plus interne dans la dimension source.

Un tableau empilé peut comporter plusieurs variables sources de statistiques récapitulatives (variable d'échelle et variable qualitative), mais chaque section du tableau comporte une seule source de statistiques récapitulatives.

Source des statistiques récapitulatives pour les variables qualitatives

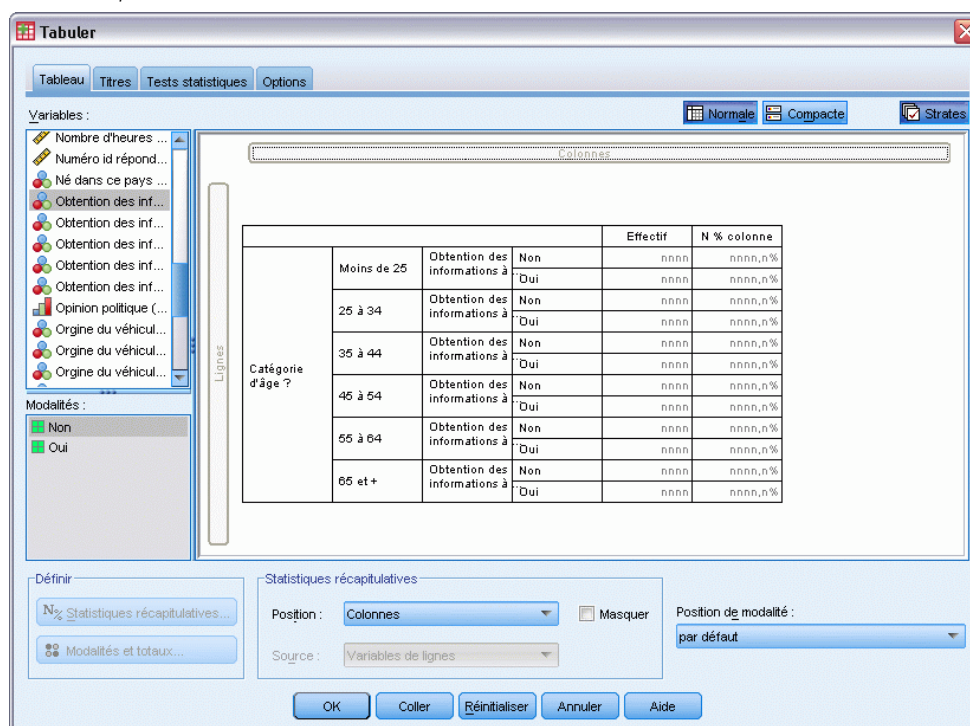
- ▶ A partir du menu, sélectionnez :
Analyse > Tableaux > Tableaux personnalisés...
- ▶ Dans le générateur de tableaux, faites glisser la variable *Tranche d'âge* de la liste des variables vers la zone Lignes du panneau de canevas.
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Tranche d'âge* et sélectionnez Statistiques récapitulatives dans le menu contextuel. (Puisqu'il s'agit de la seule variable du tableau, elle est la variable source des statistiques.)
- ▶ Dans la boîte de dialogue Statistiques récapitulatives, sélectionnez *N % colonne* dans la liste Statistiques et cliquez sur la flèche pour ajouter l'élément à la liste Afficher.

Figure 8-1
Boîte de dialogue Statistiques récapitulatives pour les variables qualitatives



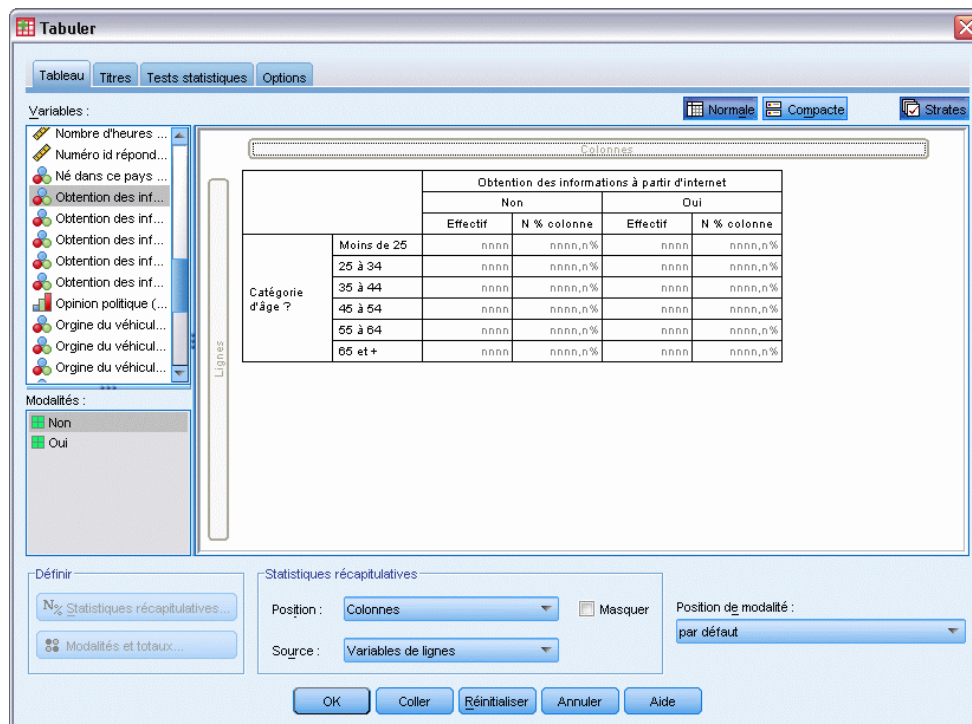
- Cliquez sur Appliquer à la sélection.
- Dans le générateur de tableaux, faites glisser la variable *Informations sur Internet* vers la droite de la variable *Tranche d'âge* dans le panneau de canevas.

Figure 8-2
Variables qualitatives emboîtées



- Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Tranche d'âge*. L'élément Statistiques récapitulatives du menu contextuel est désormais désactivé, car *Tranche d'âge* n'est pas la variable emboîtée la plus interne dans la dimension source des statistiques.
- Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Informations sur Internet*. L'item Statistiques récapitulatives est activé car, s'agissant de la variable emboîtée la plus interne dans la dimension source, elle représente désormais la variable source des statistiques récapitulatives. (Le tableau ne comportant qu'une seule dimension—Lignes—, il s'agit de la dimension source des statistiques.)
- Faites glisser *Informations sur Internet* de la zone Lignes du panneau de canevas vers la zone Colonnes.

Figure 8-3
Variables qualitatives de tableau croisé



- Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Informations sur Internet*. L'item Statistiques récapitulatives du menu contextuel est désormais désactivé, car la variable ne se trouve plus dans la dimension source des statistiques.

Tranche d'âge est à nouveau la variable source des statistiques car la dimension source par défaut pour les variables qualitatives est la première dimension dans laquelle vous placez des variables lors de la création du tableau. Dans cet exemple, la première opération a consisté à placer des variables dans la dimension de ligne. Par conséquent, la dimension de ligne est la dimension source par défaut des statistiques ; d'autre part, *Tranche d'âge* étant désormais la seule variable de cette dimension, il s'agit de la variable source des statistiques.

Source des statistiques récapitulatives pour les variables d'échelle

- Faites glisser la variable d'échelle *Heures par jour devant la télé* vers la gauche de la variable *Tranche d'âge* dans la zone Lignes du panneau de canevas.

Figure 8-4

Tableau croisé avec variable d'échelle et statistiques récapitulatives

The screenshot shows the 'Tabuler' software interface. On the left, a list of variables includes 'Nombre d'heures passées à regarder la télévision ?' and 'Catégorie d'âge ?'. The main canvas displays a pivot table with 'Lignes' (Rows) and 'Colonnes' (Columns). The table structure is as follows:

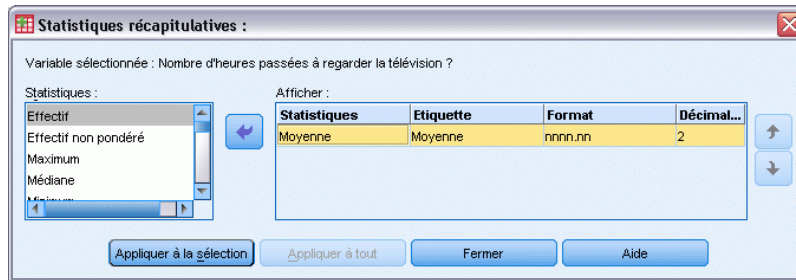
		Obtention des informations à ...	
		Non	Oui
		Moyenne	Moyenne
Nombre d'heures passées à regarder la télévision ?	Moins de 25	nnnn	nnnn
	25 à 34	nnnn	nnnn
	35 à 44	nnnn	nnnn
	45 à 54	nnnn	nnnn
	55 à 64	nnnn	nnnn
	65 et +	nnnn	nnnn

At the bottom, the 'Statistiques récapitulatives' section is configured with 'Position : Colonnes', 'Source : Variables de lignes', and 'Format : Moyenne'.

La première chose qui saute aux yeux est que les récapitulatifs *Effectif* et *N % colonne* ont été remplacés par *Moyenne* ; en outre, si vous cliquez avec le bouton droit de la souris sur *Heures par jour devant la télé* dans le panneau de canevas, vous constatez qu'il s'agit désormais de la variable source des statistiques récapitulatives. Pour un tableau comportant une variable d'échelle, cette dernière est toujours la variable source des statistiques, quel que soit son niveau d'emboîtement ou sa dimension, et la statistique récapitulative par défaut pour les variables d'échelle est la moyenne.

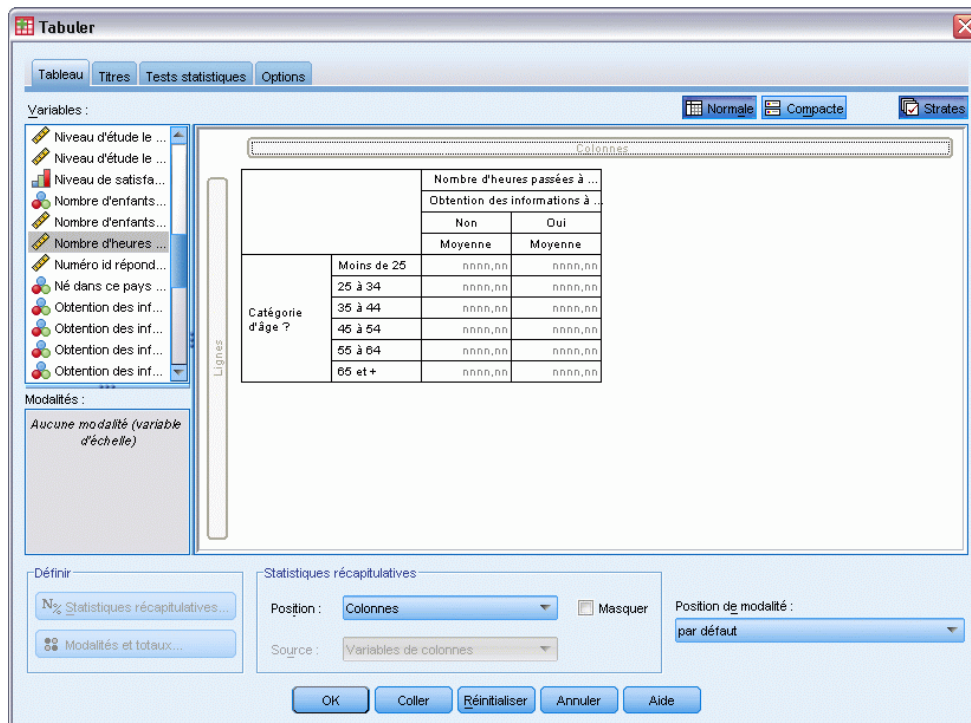
- Faites glisser *Heures par jour devant la télé* de la zone Lignes vers la zone Colonnes, au-dessus de *Informations sur Internet*.
- Cliquez avec le bouton droit de la souris sur *Heures par jour devant la télé* et sélectionnez *Statistiques récapitulatives* dans le menu contextuel. (Elle reste la variable source des statistiques, même lorsque vous la déplacez vers une autre dimension.)
- Dans la boîte de dialogue *Statistiques récapitulatives*, cliquez sur la cellule *Format* pour la moyenne dans la liste *Afficher*, puis sélectionnez *nnnn* dans la liste déroulante *Format*. (Vous pouvez être amené à parcourir la liste pour accéder à cet élément.)
- Dans la cellule *Décimales*, saisissez 2.

Figure 8-5
Boîte de dialogue Statistiques récapitulatives pour les variables d'échelle



- Cliquez sur Appliquer à la sélection.

Figure 8-6
Statistique récapitulative d'échelle avec deux décimales



L'aperçu du tableau dans le panneau de canevas indique désormais que les valeurs moyennes apparaîtront avec deux décimales.

- Cliquez sur OK pour créer le tableau.

Figure 8-7

Variable d'échelle récapitulée avec des variables qualitatives de tableau croisé

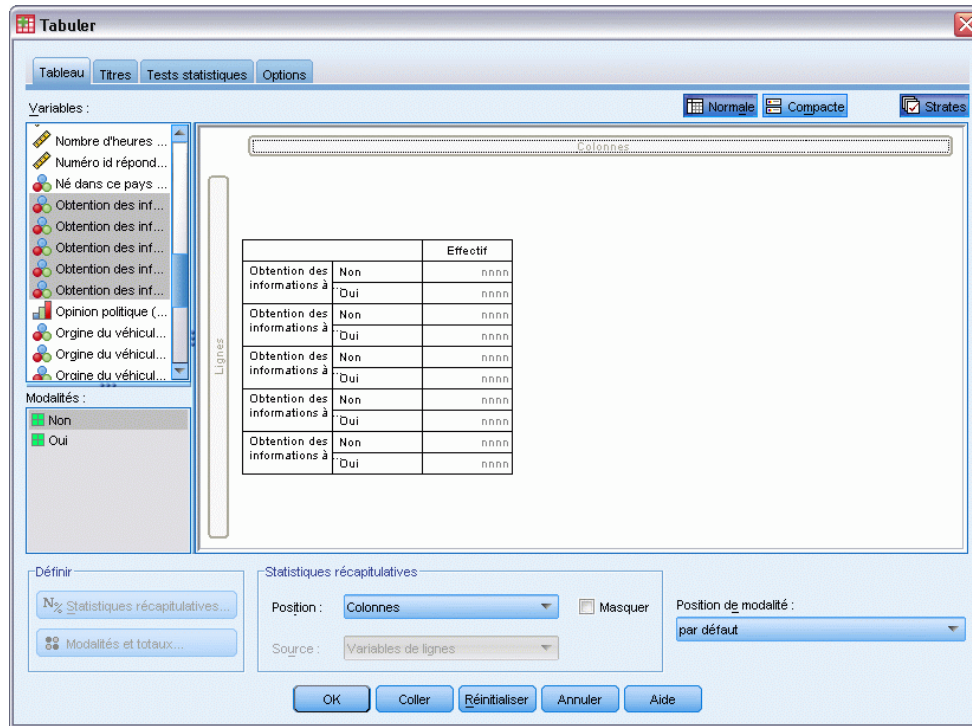
		# d'heures par jour passées devant la télévision ?	
		Obtention de l'information via internet	
		Non	Oui
		Moyenne	Moyenne
Catégories d'âges	Moins de 25	3,54	2,12
	25 à 34	3,42	2,14
	35 à 44	3,00	2,01
	45 à 54	2,83	2,06
	55 à 64	3,24	2,37
	65 ou plus	3,82	2,33

Variables empilées

Etant donné qu'un tableau empilé peut contenir plusieurs variables source de statistiques et que vous pouvez indiquer différentes statistiques récapitulatives pour chacune de ces variables, vous devez prendre en considération certains éléments avant d'indiquer les statistiques récapitulatives dans les tableaux empilés.

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Cliquez sur Restaurer pour effacer les paramètres précédents dans le générateur de tableaux.
- ▶ Cliquez sur *Informations sur Internet* dans la liste des variables, puis sur *Informations à la télévision* dans la liste des variables tout en maintenant la touche Maj enfoncée afin de sélectionner toutes les variables relatives aux informations. (*Remarque* : Ceci implique que les étiquettes de variable soient affichées dans l'ordre alphabétique, et non dans l'ordre du fichier, dans la liste des variables.)
- ▶ Faites glisser les cinq variables relatives aux informations vers la zone Lignes du panneau de canevas.

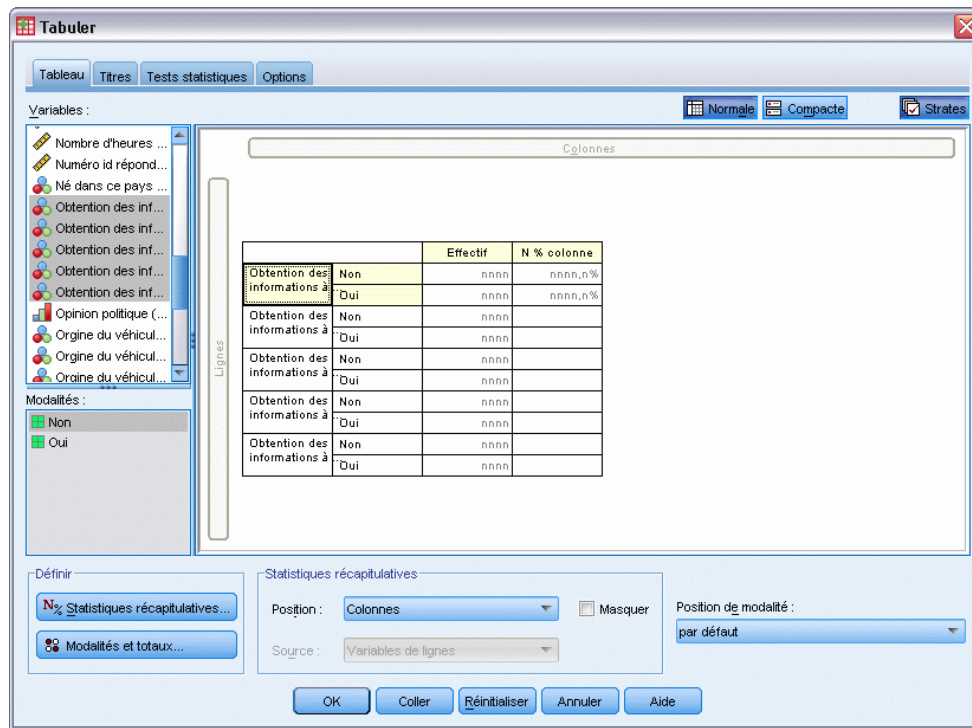
Figure 8-8
Variables relatives aux informations, empilées en lignes



Les cinq variables relatives aux informations sont empilées dans la dimension de ligne.

- ▶ Dans le panneau de canevas, cliquez sur *Informations sur Internet* de sorte que seule cette variable soit sélectionnée.
- ▶ Cliquez ensuite avec le bouton droit de la souris sur *Informations sur Internet* et sélectionnez *Statistiques récapitulatives* dans le menu contextuel.
- ▶ Dans la boîte de dialogue *Statistiques récapitulatives*, sélectionnez *N % colonne* dans la liste *Statistiques* et cliquez sur la flèche pour ajouter l'élément à la liste *Afficher*. (Pour déplacer les statistiques sélectionnées de la liste *Statistiques* vers la liste *Afficher*, vous pouvez utiliser la flèche ou faire glisser les statistiques.)
- ▶ Cliquez ensuite sur *Appliquer* à la sélection.

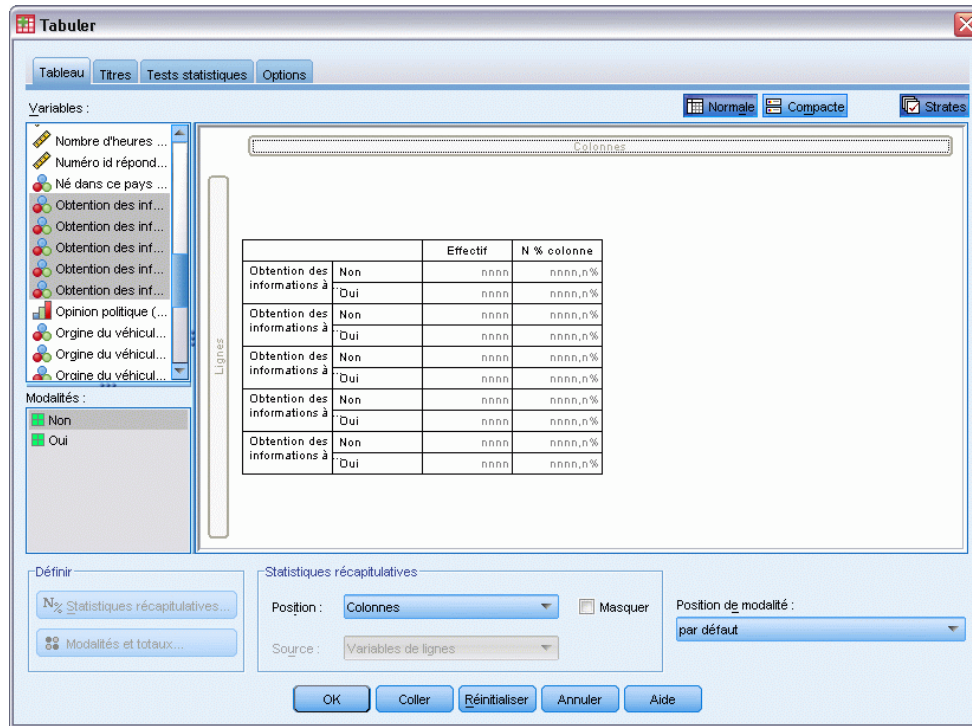
Figure 8-9
Statistique supplémentaire appliquée à une variable dans un tableau empilé



Une colonne est ajoutée pour les pourcentages en colonne, mais l'aperçu du tableau dans le panneau de canevas indique que les pourcentages en colonne apparaîtront pour une seule variable. Ceci est dû au fait qu'un tableau empilé comporte plusieurs variables source de statistiques et que chacune d'elles peut présenter des statistiques récapitulatives différentes. Dans cet exemple, toutefois, nous souhaitons afficher les mêmes statistiques récapitulatives pour toutes les variables.

- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Informations dans les journaux* et sélectionnez *Statistiques récapitulatives* dans le menu contextuel.
- ▶ Dans la boîte de dialogue *Statistiques récapitulatives*, sélectionnez *N % colonne* dans la liste *Statistiques* et cliquez sur la flèche pour ajouter l'élément à la liste *Afficher*.
- ▶ Cliquez ensuite sur *Appliquer à tout*.

Figure 8-10
Statistique supplémentaire appliquée à toutes les variables dans un tableau empilé



L'aperçu du tableau indique désormais que les pourcentages en colonne apparaîtront pour toutes les variables empilées.

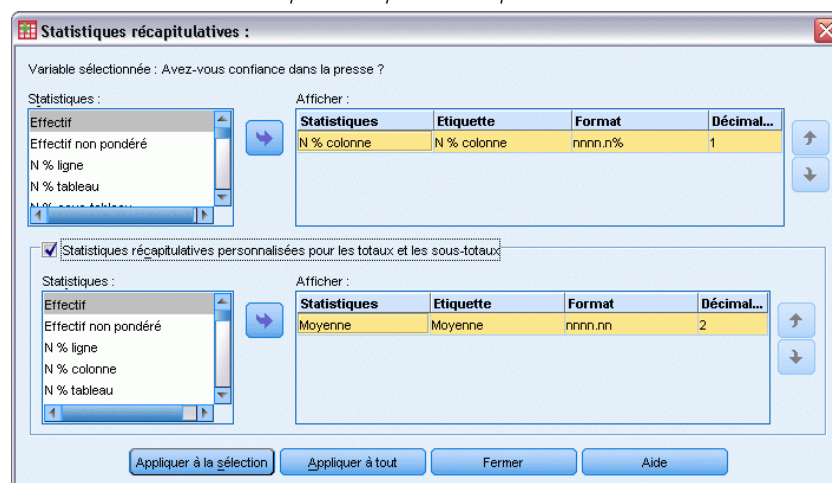
Autres statistiques récapitulatives en total pour les variables qualitatives

En ce qui concerne les variables source de statistiques qualitatives, vous pouvez inclure d'autres statistiques récapitulatives en total, différentes des statistiques affichées pour les modalités de la variable. Par exemple, pour une variable ordinaire, vous pouvez afficher des pourcentages pour chaque modalité, et la moyenne ou la médiane pour l'autre statistique récapitulative en total.

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Cliquez sur Restaurer pour effacer les paramètres précédents dans le générateur de tableaux.
- ▶ Cliquez sur *Confiance en la presse* dans la liste des variables, puis sur *Confiance en la télévision* tout en maintenant la touche Ctrl enfoncée afin de sélectionner les deux variables.
- ▶ Faites glisser les deux variables vers la zone Lignes du panneau de canevas. Les deux variables sont ainsi empilées dans la dimension de ligne.
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur une des deux variables choisissez Sélectionner toutes les variables en ligne dans le menu contextuel. (Elles sont peut-être déjà sélectionnées toutes les deux, mais cette action permet de le garantir.)

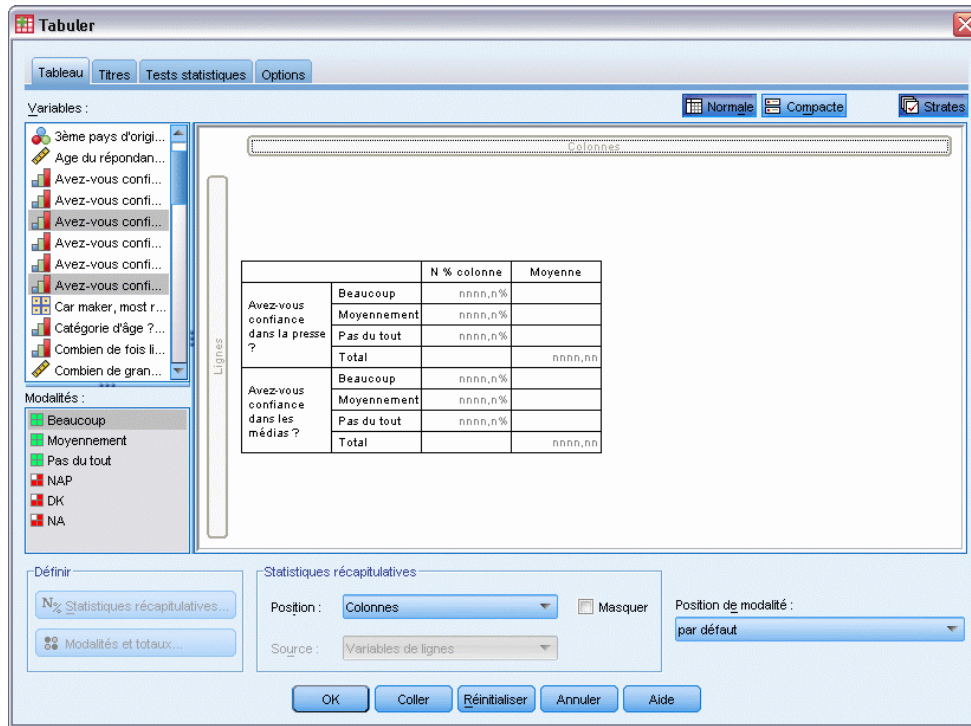
- ▶ Cliquez de nouveau avec le bouton droit de la souris sur la variable et sélectionnez Modalités et totaux dans le menu contextuel.
- ▶ Dans la boîte de dialogue Modalités et totaux, cochez Total, puis cliquez sur Appliquer.
L'aperçu du tableau dans le panneau de canevas affiche désormais une ligne de total pour les deux variables. Pour afficher d'autres statistiques récapitulatives en total, vous devez indiquer des totaux et/ou des sous-totaux pour le tableau.
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur une des deux variables et sélectionnez Statistiques récapitulatives dans le menu contextuel.
- ▶ Dans la boîte de dialogue Statistiques récapitulatives, cliquez sur *Effectif* dans la liste Afficher, puis sur la flèche pour déplacer l'élément vers la liste Statistiques (il disparaît de la liste Afficher).
- ▶ Cliquez sur *N % colonne* dans la liste Statistiques et cliquez sur le bouton fléché pour déplacer cet élément vers la liste Afficher.
- ▶ Cochez Personnaliser les statistiques récapitulatives pour les totaux et les sous-totaux.
- ▶ Cliquez sur *Effectif* dans la liste Afficher relative aux autres statistiques récapitulatives, puis sur la flèche pour déplacer l'élément vers la liste Statistiques (il disparaît de la liste Afficher).
- ▶ Cliquez sur *Moyenne* dans la liste Statistiques relative aux autres statistiques récapitulatives, puis sur la flèche pour déplacer l'élément vers la liste Afficher.
- ▶ Cliquez sur la cellule Format pour la moyenne dans la liste Afficher et sélectionnez nnnn.n dans la liste déroulante des formats. (Vous pouvez être amené à parcourir la liste pour accéder à cet élément.)
- ▶ Dans la cellule Décimales, saisissez 2.

Figure 8-11
Sélection d'autres statistiques récapitulatives pour les totaux



- ▶ Cliquez sur Appliquer à tout pour appliquer ces paramètres aux deux variables dans le tableau.

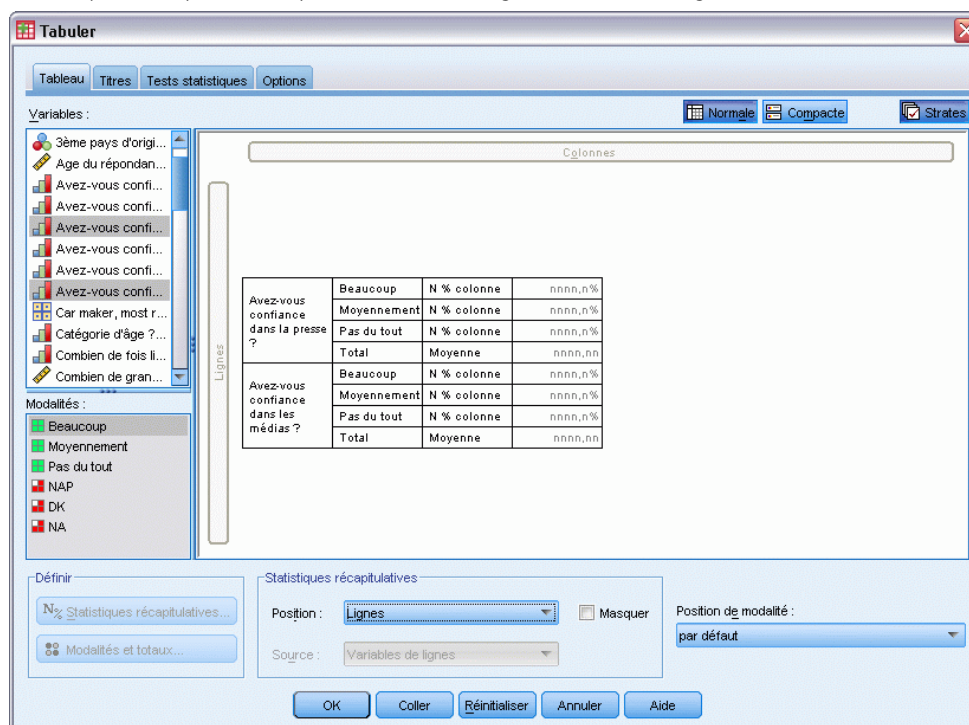
Figure 8-12
Autres statistiques récapitulatives en total pour les variables en ligne affichées en colonnes



Une nouvelle colonne a été ajoutée pour l'autre statistique récapitulative en total, mais peut-être n'est-ce pas ce que vous souhaitez, puisque l'aperçu du panneau de canevas indique clairement que ceci aboutira à un tableau contenant de nombreuses cellules vides.

- Dans le générateur de tableaux, dans le groupe Statistiques récapitulatives, sélectionnez Lignes dans la liste déroulante Position.

Figure 8-13
Statistiques récapitulatives pour variables en ligne affichées en lignes



Ceci déplace toutes les statistiques récapitulatives vers la dimension de ligne et les affiche dans une seule colonne du tableau.

- Cliquez sur OK pour créer le tableau.

Figure 8-14
Variables qualitatives avec d'autres statistiques récapitulatives en total

Confiance	Enormément	% colonne	9,5%
dans la	Un peu	% colonne	47,2%
presse	Pas du tout	% colonne	43,4%
	Total	Moyenne	2,34
Confiance	Enormément	% colonne	10,4%
dans la	Un peu	% colonne	49,9%
télévisinn	Pas du tout	% colonne	39,7%
	Total	Moyenne	2,29

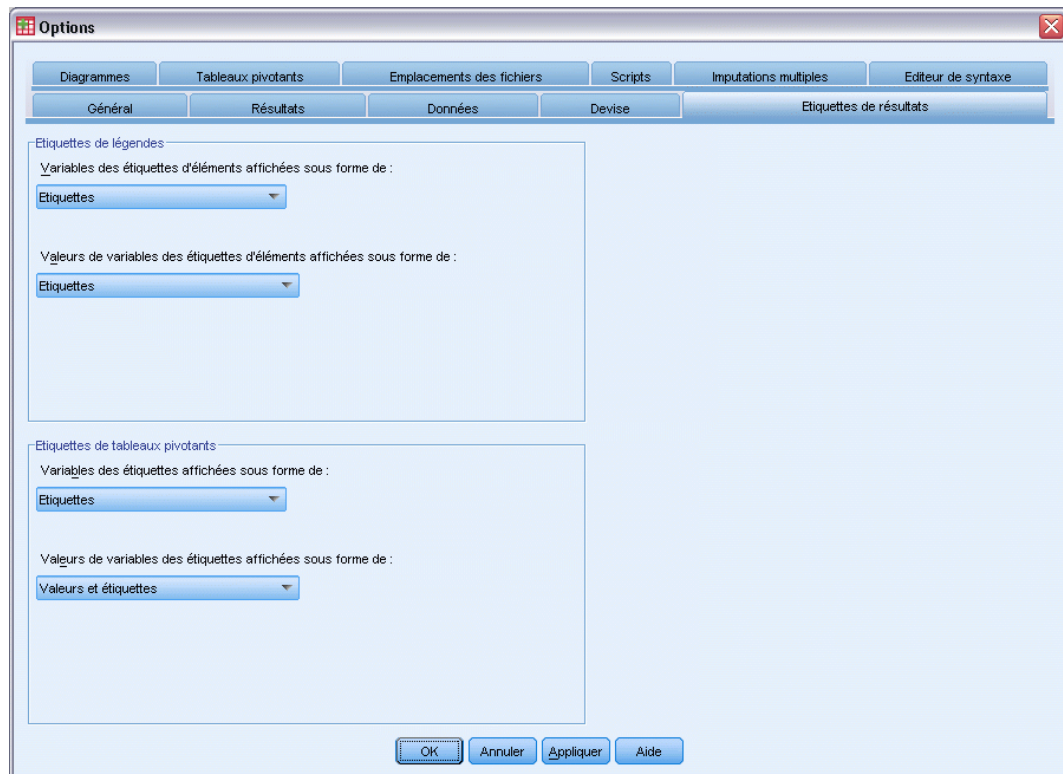
Affichage des valeurs des modalités

Le tableau précédent pose un petit problème : il peut s'avérer difficile d'interpréter la valeur moyenne sans connaître les valeurs de modalité sous-jacentes sur lesquelles elle est basée. Une moyenne de 2,34 se situe-t-elle quelque part entre *Beaucoup* et *Seulement certaines*, ou quelque part entre *Seulement certaines* et *Presque aucune* ?

Bien qu'il ne soit pas possible d'aborder ce problème directement dans Tableaux personnalisés, nous pouvons le traiter de façon plus générale.

- ▶ A partir du menu, sélectionnez :
Edition > Options...
- ▶ Dans la boîte de dialogue Options, cliquez sur l'onglet Etiquettes Résultats.
- ▶ Dans le groupe Etiquette tableau pivotant, sélectionnez Valeurs et étiquettes dans la liste déroulante Valeurs des variables en étiquettes montrées comme.

Figure 8-15
Options d'étiquetage des résultats



- ▶ Cliquez sur OK pour enregistrer ce paramètre.
- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tables (sélectionner tables et champs), Tableaux personnalisés), puis cliquez sur OK pour créer à nouveau le tableau.

Figure 8-16
Valeurs et étiquettes affichées pour les modalités de la variable

Confiance	1	Enormément	% colonne	9,5%
dans la presse	2	Un peu	% colonne	47,2%
	3	Pas du tout	% colonne	43,4%
		Total	Moyenne	2,34
Confiance	1	Enormément	% colonne	10,4%
dans la télévision	2	Un peu	% colonne	49,9%
	3	Pas du tout	% colonne	39,7%
		Total	Moyenne	2,29

Les valeurs de modalité indiquent clairement qu'une moyenne de 2,34 se situe quelque part entre *Seulement certaines* et *Presque aucune*. L'affichage des valeurs de modalité dans le tableau facilite grandement l'interprétation de la valeur des autres statistiques récapitulatives en total, comme la moyenne.

Ce paramètre d'affichage est un paramètre global qui affecte tous les résultats de tableau pivotant de toutes les procédures, et qui est conservé lors des sessions suivantes jusqu'à ce que vous le modifiiez. Pour modifier le paramètre et afficher de nouveau uniquement les étiquettes de valeur :

- ▶ A partir du menu, sélectionnez :
Edition > Options...
- ▶ Dans la boîte de dialogue Options, cliquez sur l'onglet *Etiquettes Résultats*.
- ▶ Dans le groupe *Etiquette tableau pivotant*, sélectionnez *Etiquettes* dans la liste déroulante *Valeurs des variables en étiquettes montrées comme*.
- ▶ Cliquez sur OK pour enregistrer ce paramètre.

Récapitulation des variables d'échelle

Une grande variété de statistiques récapitulatives sont disponibles pour les variables d'échelle. Outre les effectifs et les pourcentages disponibles pour les variables qualitatives, les statistiques récapitulatives pour les variables d'échelle incluent également :

- Moyenne
- Médiane
- Centiles
- Somme
- Ecart type
- Intervalle
- Valeurs minimale et maximale

Pour plus d'informations, reportez-vous à la section [Statistiques récapitulatives pour les variables d'échelle et autres statistiques récapitulatives en total pour les variables qualitatives dans le chapitre 2 sur p. 38.](#)

Fichier de données d'exemple

Les exemples de cette du chapitre utilisent le fichier de données *survey_sample.sav*. [Pour plus d'informations, reportez-vous à la section Fichiers d'exemple dans l'annexe A sur p. 197.](#)

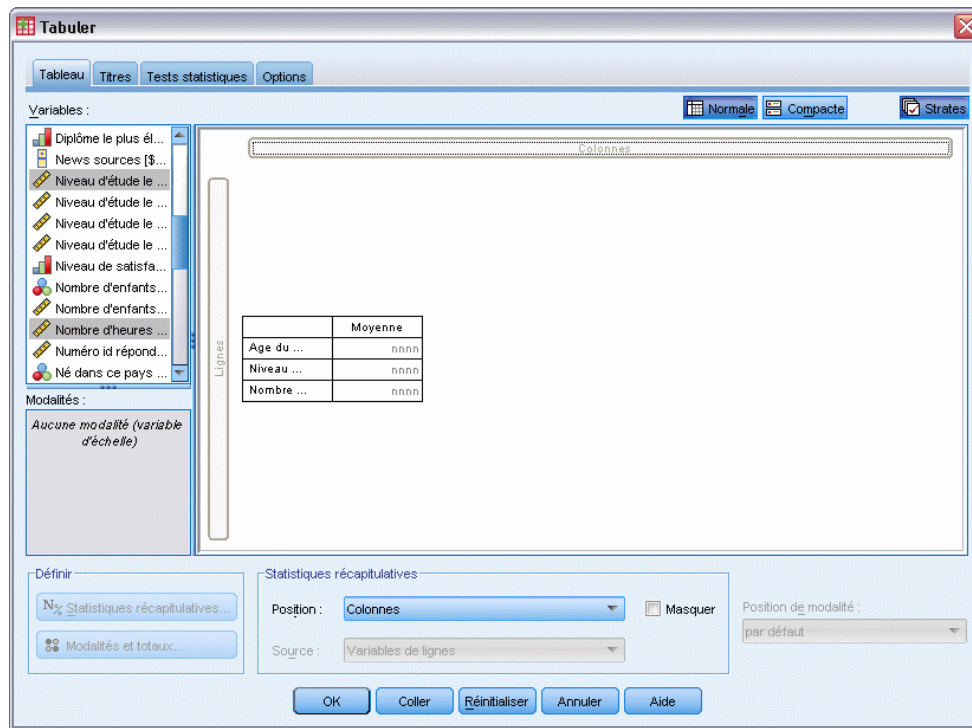
Tous les exemples fournis ici affichent des étiquettes de variable dans des boîtes de dialogue ; ces étiquettes sont triées par ordre alphabétique. Les propriétés d'affichage de la liste des variables sont spécifiées dans l'onglet Général de la boîte de dialogue Options (menu Modifier, Options).

Variables d'échelle empilées

Vous pouvez récapituler plusieurs variables d'échelle dans le même tableau en les empilant.

- ▶ A partir du menu, sélectionnez :
Analyse > Tableaux > Tableaux personnalisés...
- ▶ Dans le générateur de tableaux, cliquez sur *Age du répondant* dans la liste des variables, puis, tout en maintenant la touche Ctrl enfoncée, cliquez sur *Niveau d'étude le plus élevé* et sur *Heures par jour devant la télé* pour sélectionner les trois variables.
- ▶ Faites glisser les trois variables sélectionnées vers la zone Lignes du panneau de canevas.

Figure 9-1
Variables d'échelle empilées dans le générateur de tableaux



Les trois variables sont empilées dans la dimension de ligne. Les trois variables étant des variables d'échelle, aucune modalité n'est affichée et la statistique récapitulative par défaut est la moyenne.

- Cliquez sur OK pour créer le tableau.

Figure 9-2
Tableau des valeurs moyennes des variables d'échelle empilées

	Moyenne
Age du répondant	46
# d'années d'étude	13
# d'heures par jour passées devant la télévision ?	3

Statistiques récapitulatives multiples

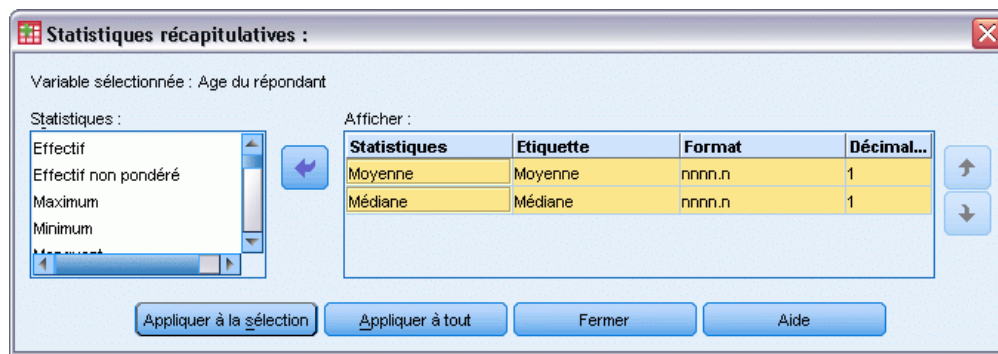
Par défaut, la moyenne apparaît pour les variables d'échelle ; toutefois, vous pouvez choisir d'autres statistiques récapitulatives pour les variables d'échelle et en afficher plusieurs.

- Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur l'une des trois variables d'échelle dans l'aperçu du tableau et sélectionnez Statistiques récapitulatives dans le menu contextuel.

- ▶ Dans la boîte de dialogue Statistiques récapitulatives, sélectionnez *Médiane* dans la liste Statistiques et cliquez sur la flèche pour ajouter cet élément à la liste Afficher. (Pour déplacer les statistiques sélectionnées de la liste Statistiques vers la liste Afficher, vous pouvez utiliser la flèche ou faire glisser les statistiques.)
- ▶ Cliquez sur la cellule Format pour la médiane dans la liste Afficher et sélectionnez nnnn.n dans la liste déroulante des formats.
- ▶ Dans la cellule Décimales, saisissez 1.
- ▶ Apportez les mêmes modifications pour la moyenne dans la liste Afficher.

Figure 9-3

Statistiques de moyenne et de médiane sélectionnées dans la boîte de dialogue Statistiques récapitulatives



- ▶ Cliquez sur Appliquer à tout pour appliquer ces modifications aux trois variables d'échelle.
- ▶ Dans le générateur de tableaux, cliquez sur OK pour créer le tableau.

Figure 9-4

Moyenne et médiane affichées dans le tableau des variables d'échelle empilées

	Moyenne	Médiane
Age du répondant	45,6	42,0
# d'années d'étude	13,3	13,0
# d'heures par jour passées devant la télévision ?	2,9	2,0

Effectif, N valide et valeurs manquantes

Il est souvent utile d'afficher le nombre d'observations utilisées pour calculer les statistiques récapitulatives, comme la moyenne, et vous pouvez raisonnablement penser que la statistique récapitulative *Effectif* est susceptible de fournir cette information. Toutefois, ceci n'offrira pas une base d'observations précise s'il existe des valeurs manquantes. Pour obtenir une base d'observations précise, utilisez la statistique *N valide*.

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur l'une des trois variables d'échelle dans l'aperçu du tableau et sélectionnez Statistiques récapitulatives dans le menu contextuel.

- ▶ Dans la boîte de dialogue Statistiques récapitulatives, sélectionnez **Effectif** dans la liste Statistiques et cliquez sur la flèche pour ajouter cet élément à la liste Afficher.
- ▶ Sélectionnez ensuite **N valide** dans la liste Statistiques et cliquez sur la flèche pour ajouter cet élément à la liste Afficher.
- ▶ Cliquez sur **Appliquer à tout** pour appliquer ces modifications aux trois variables d'échelle.
- ▶ Dans le générateur de tableaux, cliquez sur **OK** pour créer le tableau.

Figure 9-5
Effectif et N valide

	Moyenne	Médiane	Effectif	N valide
Age du répondant	45,6	42,0	2832	2828
Niveau d'étude le plus élevé	13,3	13,0	2832	2820
Nombre d'heures passées à regarder la télévision ?	2,9	2,0	2832	2337

Pour les trois variables, la statistique *Effectif* est identique : 2,832. Il ne s'agit pas d'un hasard, puisqu'elle représente le nombre total d'observations dans le fichier de données. Puisque les variables d'échelle ne sont pas emboîtées dans des variables qualitatives, la statistique *Effectif* représente simplement le nombre total d'observations dans le fichier de données.

D'autre part, la statistique *N valide* est différente pour chaque variable et se détache de façon significative de la statistique *Effectif* pour *Heures par jour devant la télé*. En effet, il existe un grand nombre de **valeurs manquantes** pour cette variable, c'est-à-dire des observations sans valeur enregistrée pour cette variable ou des valeurs définies en tant que données manquantes (le code 99 peut représenter *Sans objet* dans le cas de grossesse pour hommes).

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur l'une des trois variables d'échelle dans l'aperçu du tableau et sélectionnez Statistiques récapitulatives dans le menu contextuel.
- ▶ Dans la boîte de dialogue Statistiques récapitulatives, sélectionnez **N valide** dans la liste Afficher, puis cliquez sur le bouton fléché pour renvoyer cet élément vers la liste Statistiques (il n'apparaît plus dans la liste Afficher).
- ▶ Sélectionnez **Effectif** dans la liste Afficher, puis cliquez sur le bouton fléché pour renvoyer cet élément vers la liste Statistiques (il n'apparaît plus dans la liste Afficher).
- ▶ Sélectionnez **Manquant** dans la liste Statistiques et cliquez sur le bouton fléché pour ajouter cet élément à la liste Afficher.
- ▶ Cliquez sur **Appliquer à tout** pour appliquer ces modifications aux trois variables d'échelle.
- ▶ Dans le générateur de tableaux, cliquez sur **OK** pour créer le tableau.

Figure 9-6
Nombre de valeurs manquantes affichées dans le tableau des statistiques récapitulatives de variable d'échelle

	Moyenne	Médiane	Manquant
Age du répondant	45,6	42,0	4
# d'années d'étude	13,3	13,0	12
# d'heures par jour passées devant la télévision ?	2,9	2,0	495

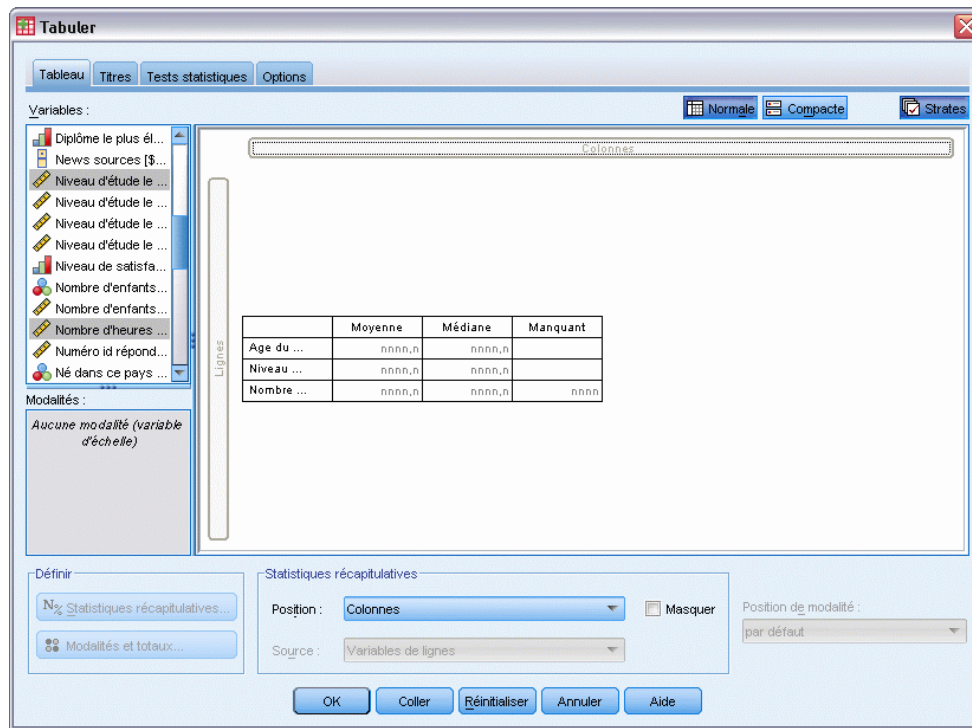
Le tableau indique désormais le nombre de valeurs manquantes pour chaque variable d'échelle. Il apparaît évident que la variable *Heures par jour devant la télé* présente un grand nombre de valeurs manquantes, tandis que les deux autres variables en présentent très peu. Il peut s'agir ici d'un facteur à prendre en compte avant d'accorder une grande confiance aux valeurs récapitulatives pour cette variable.

Différents récapitulatifs pour différentes variables

Outre l'affichage de plusieurs statistiques récapitulatives, vous pouvez afficher, dans un tableau empilé, différentes statistiques récapitulatives pour différentes variables d'échelle. Par exemple, le tableau précédent a montré qu'une seule des trois variables présentait un grand nombre de valeurs manquantes ; vous pouvez dès lors choisir d'afficher le nombre de valeurs manquantes uniquement pour cette variable.

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Dans le panneau de canevas, cliquez sur *Age du répondant* dans l'aperçu du tableau, puis cliquez sur *Niveau d'étude le plus élevé* tout en maintenant la touche Ctrl enfoncée pour sélectionner les deux variables.
- ▶ Cliquez avec le bouton droit de la souris sur l'une des deux variables sélectionnées, puis choisissez Statistiques récapitulatives dans le menu contextuel.
- ▶ Dans la boîte de dialogue Statistiques récapitulatives, sélectionnez Manquant dans la liste Afficher, puis cliquez sur le bouton fléché pour renvoyer cet élément vers la liste Statistiques (il n'apparaît plus dans la liste Afficher).
- ▶ Cliquez sur Appliquer à la sélection pour appliquer la modification uniquement aux deux variables sélectionnées.

Figure 9-7
Aperçu de tableau pour différentes statistiques récapitulatives pour différentes variables



Les valeurs de substitution dans les cellules de données du tableau indiquent que le nombre de valeurs manquantes apparaîtra uniquement pour la variable *Heures par jour devant la télé*.

- Cliquez sur OK pour créer le tableau.

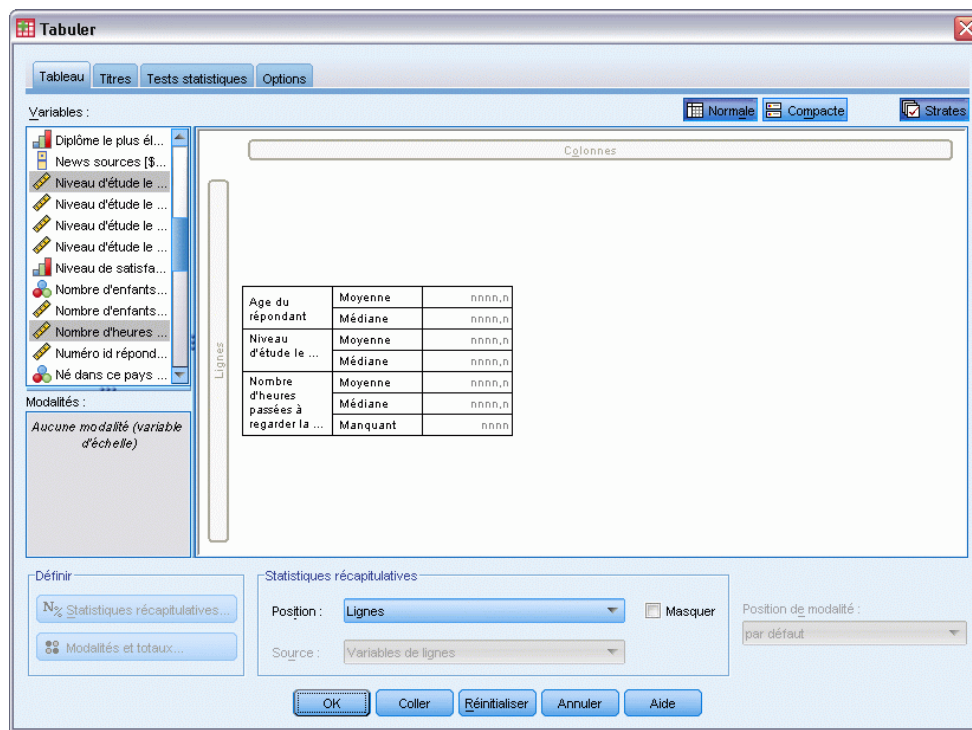
Figure 9-8
Tableau de différentes statistiques récapitulatives pour différentes variables

	Moyenne	Médiane	Manquant
Age du répondant	45,6	42,0	
# d'années d'étude	13,3	13,0	
# d'heures par jour passées devant la télévision ?	2,9	2,0	495

Bien que ce tableau fournisse les informations souhaitées, sa présentation peut compliquer son interprétation. Si vous visualisez le tableau, vous pouvez penser que les cellules vides de la colonne *Manquant* indique une absence de valeurs manquantes pour ces variables.

- Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- Dans le générateur de tableaux, dans le groupe Statistiques récapitulatives, sélectionnez Lignes dans la liste déroulante Position.

Figure 9-9
Déplacement des statistiques récapitulatives de la dimension de colonne vers la dimension de ligne



- Cliquez sur OK pour créer le tableau.

Figure 9-10
Statistiques récapitulatives et variables affichées dans la dimension de ligne

Age du répondant	Moyenne	45,6
	Médiane	42,0
# d'années d'étude	Moyenne	13,3
	Médiane	13,0
# d'heures par jour passées devant la télévision ?	Moyenne	2,9
	Médiane	2,0
	Manquant	495

Il est désormais évident que le tableau fait état du nombre de valeurs manquantes pour une seule variable.

Récapitulatifs de groupe dans des modalités

Vous pouvez utiliser des variables qualitatives comme variables de regroupement afin d'afficher des récapitulatifs de variables d'échelle au sein de groupes définis par les modalités de la variable qualitative.

- Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- Faites glisser la variable *Sexe* de la liste des variables vers la zone Colonnes du panneau de canevas.

Dans le panneau de canevas, si vous cliquez avec le bouton droit de la souris sur la variable *Sexe* dans l'aperçu du tableau, vous remarquez que l'option *Statistiques récapitulatives* est désactivée dans le menu contextuel. En effet, dans un tableau comportant des variables d'échelle, ces variables sont toujours les variables source des statistiques.

- Cliquez sur OK pour créer le tableau.

Figure 9-11

Récapitulatifs groupés de variable d'échelle avec une variable qualitative de colonne

		Sexe du répondant	
		Homme	Femme
Age du répondant	Moyenne	44,6	46,3
	Médiane	42,0	43,0
# d'années d'étude	Moyenne	13,4	13,2
	Médiane	13,0	13,0
# d'heures par jour passées devant la télévision ?	Moyenne	2,8	2,9
	Médiane	2,0	2,0
	Manquant	213	282

Ce tableau permet de comparer facilement les moyennes (moyenne et médiane) pour les hommes et les femmes, et indique clairement qu'il n'existe pas une grande différence entre elles (information peu pertinente au premier abord, mais susceptible d'être utile).

Variables de regroupement multiples

Vous pouvez diviser davantage les groupes en emboîtant, et/ou en utilisant des variables de regroupement qualitatives de ligne et de colonne.

- Ouvrez le générateur de tableaux (menu *Analyse*, *Tableaux*, *Tableaux personnalisés*).
- Faites glisser *Informations sur Internet* de la liste des variables vers l'extrême gauche de la zone Lignes du panneau de canevas. Veillez à positionner cet élément de sorte que les trois variables d'échelle (et non pas seulement une) soient emboîtées dedans.

Figure 9-12
Correct : Les trois variables d'échelle sont emboîtées dans la variable qualitative

Variables :

Tableau Titres Tests statistiques Options

Normale Compacte Strates

				Sexe	
				Homme	Femme
Obtention des informations à partir d'internet	Non	Age du répondant	Moyenne	nnnn,n	nnnn,n
			Médiane	nnnn,n	nnnn,n
		Niveau d'étude le ...	Moyenne	nnnn,n	nnnn,n
		Médiane	nnnn,n	nnnn,n	
		Nombre d'heures passées à regarder la ...	Moyenne	nnnn,n	nnnn,n
			Médiane	nnnn,n	nnnn,n
Oui	Age du répondant	Moyenne	nnnn,n	nnnn,n	
			Médiane	nnnn,n	nnnn,n
		Niveau d'étude le ...	Moyenne	nnnn,n	nnnn,n
		Médiane	nnnn,n	nnnn,n	
		Nombre d'heures passées à regarder la ...	Moyenne	nnnn,n	nnnn,n
			Médiane	nnnn,n	nnnn,n
		Manquant	nnnn	nnnn	

Définir

Statistiques récapitulatives

Position : Lignes Masquer Position de modalité : par défaut

Source : Variables de lignes

OK Coller Réinitialiser Annuler Aide

Figure 9-13
Incorrect : Une seule variable d'échelle est emboîtée dans la variable qualitative

Variables :

Tableau Titres Tests statistiques Options

Normale Compacte Strates

				Sexe	
				Homme	Femme
Obtention des informations à partir d'internet	Age du répondant	Moyenne		nnnn,n	nnnn,n
		Médiane		nnnn,n	nnnn,n
	Non	Niveau d'étude le ...	Moyenne	nnnn,n	nnnn,n
			Médiane	nnnn,n	nnnn,n
	Oui	Niveau d'étude le ...	Moyenne	nnnn,n	nnnn,n
			Médiane	nnnn,n	nnnn,n
Nombre d'heures passées à regarder la ...	Moyenne		nnnn,n	nnnn,n	
	Médiane		nnnn,n	nnnn,n	
	Manquant		nnnn	nnnn	

Définir

Statistiques récapitulatives

Position : Lignes Masquer Position de modalité : par défaut

Source : Variables de lignes

OK Coller Réinitialiser Annuler Aide

Bien que le second exemple présente un intérêt certain, ce n'est pas ce que nous allons choisir dans ce cas.

- Cliquez sur OK pour créer le tableau.

Figure 9-14

Récapitulatifs de variable d'échelle groupés par variables qualitatives de ligne et de colonne

				Sexe du répondant	
				Homme	Femme
Obtention de l'information via internet	Non	Age du répondant	Moyenne	47,0	48,8
			Médiane	45,0	46,0
		# d'années d'étude	Moyenne	13,4	13,1
			Médiane	13,0	12,0
			Manquant	213	282
	Oui	Age du répondant	Moyenne	38,7	41,1
			Médiane	35,0	38,0
		# d'années d'étude	Moyenne	13,2	13,3
			Médiane	13,0	13,0
			Manquant	0	0

Emboîtement de variables qualitatives dans des variables d'échelle

Bien que le tableau précédent puisse fournir les informations souhaitées, le format de ces informations n'est pas le plus aisé pour l'interprétation. Par exemple, vous pouvez comparer l'âge moyen des hommes qui utilisent Internet pour s'informer et celui des hommes qui ne l'utilisent pas, mais cette comparaison serait facilitée si les valeurs se trouvaient côte à côte, plutôt que séparées. Vous pouvez améliorer le tableau en permutant les positions des deux variables de ligne et en emboîtant la variable de regroupement qualitative dans les trois variables d'échelle. En ce qui concerne les variables d'échelle, le niveau d'emboîtement n'a aucun effet sur la variable source des statistiques. La variable d'échelle est toujours la variable source des statistiques, quel que soit le niveau d'emboîtement.

- Ouvrez le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- Dans le panneau de canevas, cliquez sur *Age du répondant* dans l'aperçu du tableau, puis, tout en maintenant la touche Ctrl enfoncée, cliquez sur *Niveau d'étude le plus élevé* et sur *Heures par jour devant la télé* pour sélectionner les trois variables d'échelle.
- Faites glisser les trois variables d'échelle vers l'extrême gauche de la zone Lignes, ce qui a pour effet d'emboîter la variable qualitative *Informations sur Internet* dans chacune des trois variables d'échelle.
- Cliquez sur OK pour créer le tableau.

Figure 9-15
Variable qualitative de ligne emboîtée dans les variables d'échelle empilées

				Sexe du répondant	
				Homme	Femme
Age du répondant	Obtention de l'information via internet	Non	Moyenne	47,0	48,8
			Médiane	45,0	46,0
	Oui	Moyenne	38,7	41,1	
		Médiane	35,0	38,0	
# d'années d'étude	Obtention de l'information via internet	Non	Moyenne	13,4	13,1
			Médiane	13,0	12,0
	Oui	Moyenne	13,2	13,3	
		Médiane	13,0	13,0	
# d'heures par jour passées devant la télévision ?	Obtention de l'information via internet	Non	Moyenne	3,2	3,4
			Médiane	2,0	3,0
		Manquant	213	282	
	Oui	Moyenne	2,1	2,1	
		Médiane	2,0	2,0	
	Manquant	0	0		

Le choix de l'ordre d'emboîtement dépend des relations ou des comparaisons à mettre en évidence dans le tableau. La modification de l'ordre d'emboîtement des variables d'échelle n'implique pas la modification des valeurs des statistiques récapitulatives ; seule leur position relative dans le tableau est modifiée.

Statistiques de test

Trois différents tests de signification sont disponibles pour l'étude de la relation entre les variables en ligne et en colonne. Ce chapitre décrit le résultat de chacun de ces tests, notamment les effets de l'emboîtement et de l'empilement. [Pour plus d'informations, reportez-vous à la section Empilement, emboîtement et strates avec des variables qualitatives dans le chapitre 4 sur p. 70.](#)

Fichier de données d'exemple

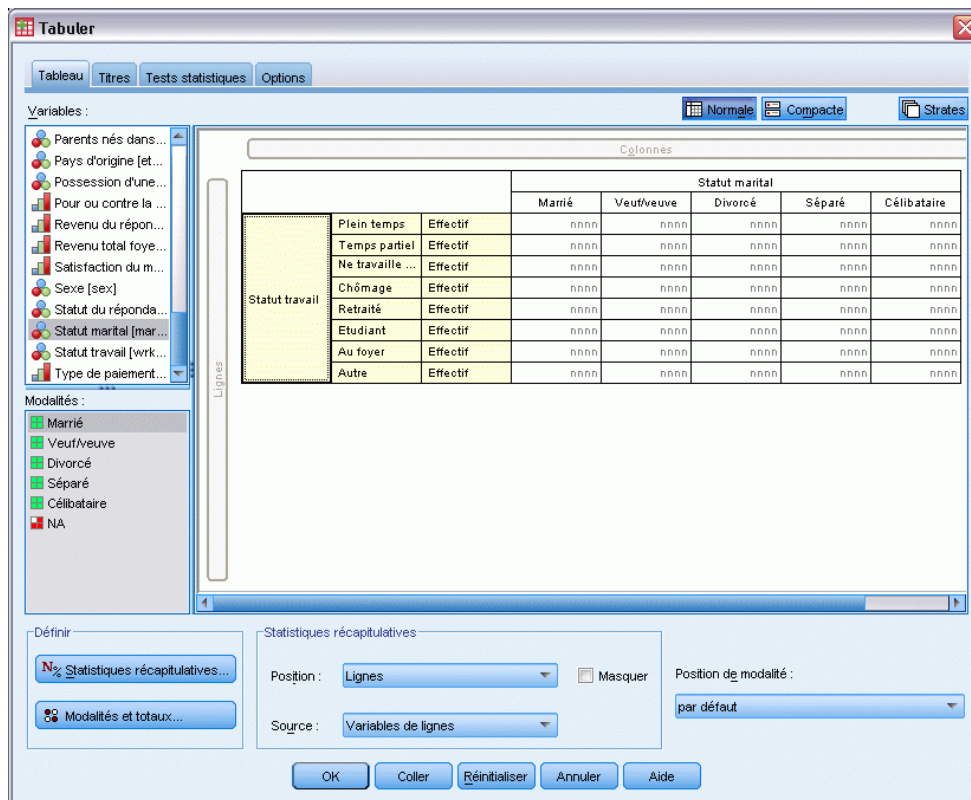
Les exemples de ce chapitre utilisent le fichier de données *survey_sample.sav*. [Pour plus d'informations, reportez-vous à la section Fichiers d'exemple dans l'annexe A sur p. 197.](#)

Tests d'indépendance (Khi-deux)

Le test d'indépendance Khi-deux sert à déterminer l'existence d'une relation entre deux variables qualitatives. Par exemple, vous pouvez déterminer si la variable *Situation d'activité* est liée à la variable *Situation familiale*.

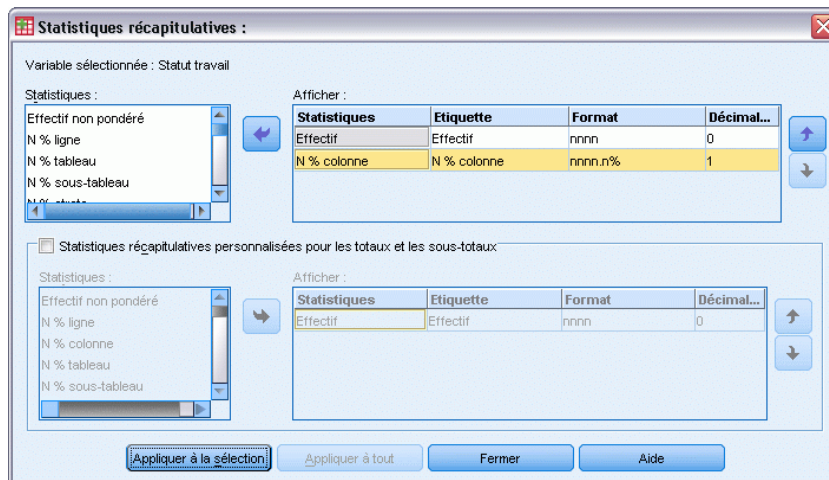
- ▶ A partir du menu, sélectionnez :
Analyse > Tableaux > Tabuler...
- ▶ Dans le générateur de tableaux, faites glisser la variable *Situation d'activité* de la liste des variables vers la zone Lignes du panneau de canevas.
- ▶ Faites glisser la variable *Situation familiale* de la liste des variables vers la zone Colonnes.

Figure 10-1
Variables affichées dans le panneau de canevas



- Sélectionnez Lignes comme position pour les statistiques récapitulatives.
- Sélectionnez *Situation d'activité*, puis cliquez sur Statistiques récapitulatives dans le groupe Définir.

Figure 10-2
Boîte de dialogue Statistiques récapitulatives



- Dans la liste Statistiques, sélectionnez la statistique N % colonne, puis ajoutez-la à la liste Afficher.

- ▶ Cliquez sur Appliquer à la sélection.
- ▶ Dans la boîte de dialogue Tableaux personnalisés, cliquez sur l'onglet Statistiques de test.

Figure 10-3

Onglet Statistiques de test dans lequel l'élément Tests d'indépendance (Khi-deux) est sélectionné

The screenshot shows the 'Tabuler' dialog box with the 'Tests statistiques' tab selected. The 'Tests d'indépendance (Khi-deux)' checkbox is checked. The 'Alpha de Cronbach' is set to 0.05. There are also checkboxes for 'Comparer les moyennes de colonne (tests-t)', 'Comparer les proportions de colonne (tests-z)', 'Ajuster les valeurs p pour des comparaisons multiples (méthode Bonferroni)', and 'Estimer la variance uniquement à partir des modalités comparées'. There are also buttons for 'OK', 'Coller', 'Réinitialiser', 'Annuler', and 'Aide'.

- ▶ Sélectionnez Tests d'indépendance (Khi-deux).
- ▶ Cliquez sur OK pour créer le tableau et obtenir le test du Khi-deux.

Figure 10-4

Variables Situation d'activité et Situation familiale

			Statut marital				
			Marié(e)	Veuf/veuve	Divorcé(e)	Séparé(e)	Jamais marié(e)
Statut professionnel	Travail à plein temps	Effectif	778	44	295	58	392
		% colonne	57,8%	15,5%	66,1%	62,4%	59,1%
	Travail à temps partiel	Effectif	138	20	35	9	102
		% colonne	10,3%	7,1%	7,8%	9,7%	15,4%
	Ne travaille pas temporairement	Effectif	23	2	9	1	11
		% colonne	1,7%	,7%	2,0%	1,1%	1,7%
	Chômage	Effectif	13	3	10	0	32
		% colonne	1,0%	1,1%	2,2%	,0%	4,8%
Retraité(e)	Effectif	168	150	53	6	17	
	% colonne	12,5%	53,0%	11,9%	6,5%	2,6%	
Ecolier(e)	Effectif	9	1	7	2	60	
	% colonne	,7%	,4%	1,6%	2,2%	9,0%	
Femme/Homme au foyer	Effectif	200	55	25	13	35	
	% colonne	14,9%	19,4%	5,6%	14,0%	5,3%	
Autre	Effectif	16	8	12	4	14	
	% colonne	1,2%	2,8%	2,7%	4,3%	2,1%	

Ce tableau est un tableau croisé des variables *Situation d'activité* et *Situation familiale*, dans lequel les effectifs et les proportions de colonne constituent les statistiques récapitulatives. Les proportions de colonne sont calculées de sorte que leur somme soit égale à 100 % pour chaque colonne. Si ces deux variables ne sont pas liées, dans chaque ligne les proportions doivent être identiques d'une colonne à l'autre. Des différences semblent exister entre les proportions ; vous pouvez effectuer une vérification à l'aide du test du Khi-deux.

Figure 10-5
Test du Khi-deux de Pearson

		Statut marital
Statut professionnel	Khi-deux	729,242
	ddl	28
	Sig.	,000*

*. Les statistiques de Khi-Deux sont significatives au niveau 0.05.

Le test d'indépendance part de l'hypothèse que les variables *Situation d'activité* et *Situation familiale* ne sont pas liées, c'est-à-dire que les proportions de colonne sont identiques d'une colonne à l'autre et que toute différence observée est due à une variation aléatoire. La statistique du Khi-deux mesure la différence globale entre les effectifs de cellules observés et les effectifs attendus si les proportions de colonne étaient identiques d'une colonne à l'autre. Plus la valeur de la statistique du Khi-deux est élevée, plus la différence entre les effectifs de cellules observés et théoriques est importante, et plus il apparaît que les proportions de colonne ne sont pas égales, que l'hypothèse d'indépendance est incorrecte et, par conséquent, que les variables *Situation d'activité* et *Situation familiale* sont liées.

La statistique du Khi-deux donne la valeur 729,242. Pour déterminer si ceci suffit pour rejeter l'hypothèse d'indépendance, la valeur de signification de la statistique est calculée. La valeur de signification est la probabilité qu'une variable aléatoire issue d'une distribution Khi-deux avec 28 degrés de liberté soit supérieure à 729,242. Dans la mesure où cette valeur est inférieure au niveau alpha spécifié dans l'onglet Statistiques de test, vous pouvez rejeter l'hypothèse d'indépendance au niveau 0,05. Par conséquent, les variables *Situation d'activité* et *Situation familiale* sont en fait liées.

Effets de l'emboîtement et de l'empilement sur les tests d'indépendance

La règle suivante s'applique aux tests d'indépendance : Un test distinct est exécuté pour chacun des sous-tableaux les plus internes. Pour voir l'impact de l'emboîtement sur les tests, examinons l'exemple précédent, mais avec la variable *Situation familiale* emboîtée dans les niveaux de la variable *Sexe*.

- ▶ Ouvrez à nouveau le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Faites glisser la variable *Sexe* de la liste des variables vers la zone Colonnes du panneau de canevas, au-dessus de la variable *Situation familiale*.
- ▶ Cliquez sur OK pour créer le tableau.

Figure 10-6
Test du Khi-deux de Pearson

Test du Khi-Deux de Pearson

		Sexe du répondant	
		Homme	Femme
		Statut marital	Statut marital
Statut professionnel	Khi-deux	246,637	542,589
	ddl	28	28
	Sig.	,000 ^{a,b}	,000 ^{a,b}

Les résultats sont basés sur les lignes et les colonnes non vides de chacun des sous-tableaux les plus internes.

*. Les statistiques de Khi-Deux sont significatives au niveau 0,05.

a. Plus de 20 % des cellules de ce sous-tableau ont un effectif de cellule attendu inférieur à 5. Les résultats du Khi-Deux risquent de ne pas être valides.

b. L'effectif de cellule minimal attendu pour ce sous-tableau est inférieur à un. Les résultats du Khi-Deux risquent de ne pas être valides.

Une fois la variable *Situation familiale* emboîtée dans les niveaux de la variable *Sexe*, deux tests sont exécutés (un pour chaque niveau de la variable *Sexe*). La valeur de signification de chaque test indique que vous pouvez rejeter l'hypothèse d'indépendance entre les variables *Situation familiale* et *Situation d'activité* pour les hommes et pour les femmes. Toutefois, le tableau indique que plus de 20 % des cellules de chaque tableau comportent des effectifs théoriques inférieurs à 5 et que l'effectif de cellule théorique minimum est inférieur à 1. Ces remarques montrent que les hypothèses du test du Khi-deux ne sont pas vérifiées dans ces tableaux et que, par conséquent, les résultats des tests sont suspects.

Remarque : Les bordures des cellules peuvent masquer les notes de bas de page. Pour les afficher, vous pouvez modifier l'alignement des cellules concernées à l'aide de la boîte de dialogue Propriétés de la cellule.

Pour voir l'impact de l'empilement sur les tests :

- ▶ Ouvrez à nouveau le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Faites glisser la variable *Diplôme le plus important* de la liste des variables vers la zone Lignes, sous la variable *Situation d'activité*.
- ▶ Cliquez sur OK pour créer le tableau.

Figure 10-7
Test du Khi-deux de Pearson

Test du Khi-Deux de Pearson

		Sexe du répondant	
		Homme	Femme
		Statut marital	Statut marital
Statut professionnel	Khi-deux	246,637	542,589
	ddl	28	28
	Sig.	,000 ^{a,b}	,000 ^{a,b}
# d'années d'étude	Khi-deux	43,844	105,506
	ddl	16	16
	Sig.	,000 ^{a,b}	,000 ^a

Les résultats sont basés sur les lignes et les colonnes non vides de chacun des sous-tableaux les plus internes.

*. Les statistiques de Khi-Deux sont significatives au niveau 0.05.

- a. Plus de 20 % des cellules de ce sous-tableau ont un effectif de cellule attendu inférieur à 5. Les résultats du Khi-Deux risquent de ne pas être valides.
- b. L'effectif de cellule minimal attendu pour ce sous-tableau est inférieur à un. Les résultats du Khi-Deux risquent de ne pas être valides.

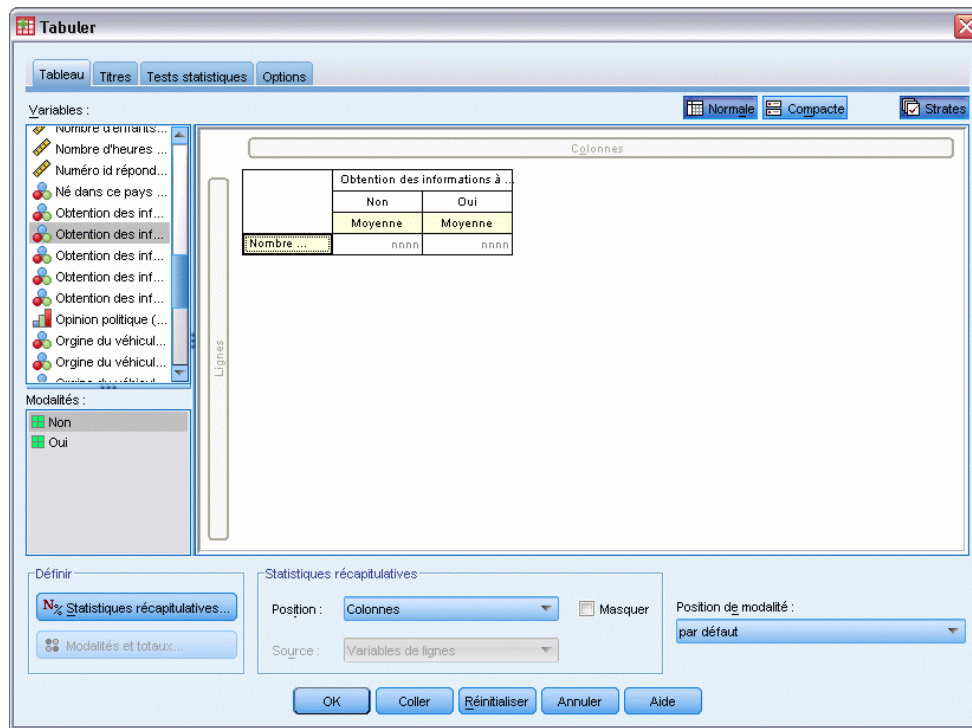
Une fois la variable *Diplôme le plus important* empilée avec la variable *Situation d'activité*, quatre tests sont exécutés : un test d'indépendance des variables *Situation familiale* et *Situation d'activité*, et un test des variables *Situation familiale* et *Diplôme le plus important* pour chaque niveau de la variable *Sexe*. Les résultats des tests pour les variables *Situation familiale* et *Situation d'activité* sont les mêmes que précédemment. Les résultats des tests pour les variables *Situation familiale* et *Diplôme le plus important* indiquent que ces variables ne sont pas indépendantes.

Comparaison des moyennes de colonne

Les tests sur les moyennes de colonne permettent de déterminer l'existence d'une relation entre une variable qualitative dans les colonnes et une variable continue dans les lignes. En outre, vous pouvez utiliser les résultats des tests pour déterminer le classement relatif des modalités de la variable qualitative sur le plan de la valeur moyenne de la variable continue. Par exemple, supposons que vous souhaitez déterminer si la variable *Heures par jour devant la télé* est liée à la variable *Informations dans les journaux*.

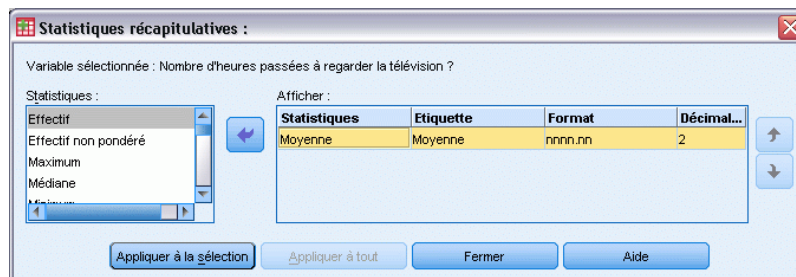
- ▶ A partir du menu, sélectionnez :
Analyse > Tableaux > Tabuler...
- ▶ Cliquez sur Restaurer pour rétablir les valeurs par défaut dans tous les onglets.
- ▶ Dans le générateur de tableaux, faites glisser la variable *Heures par jour devant la télé* de la liste des variables vers la zone Lignes du panneau de canevas.
- ▶ Faites glisser la variable *Informations dans les journaux* de la liste des variables vers la zone Colonnes.

Figure 10-8
Variables affichées dans le panneau de canevas



- Sélectionnez la variable *Heures par jour devant la télé*, puis cliquez sur *Statistiques récapitulatives* dans le groupe *Définir*.

Figure 10-9
Boîte de dialogue *Statistiques récapitulatives*



- Sélectionnez le format *nnnn*.
- Sélectionnez *2* comme nombre de décimales à afficher. Suite à cette opération, le format devient *nnnn.nn*.
- Cliquez sur *Appliquer à la sélection*.
- Dans la boîte de dialogue *Tableaux personnalisés*, cliquez sur l'onglet *Statistiques de test*.

Figure 10-10
Onglet Statistiques de test dans lequel l'élément Comparer les moyennes de colonne (tests-t) est sélectionné



- ▶ Sélectionnez Comparer les moyennes de colonne (tests-t).
- ▶ Cliquez sur OK pour créer le tableau et obtenir les tests des moyennes de colonne.

Figure 10-11
Variables Informations dans les journaux et Heures par jour devant la télé

	Obtention de l'information via la presse écrite	
	Non	Oui
	Moyenne	Moyenne
# d'heures par jour passées devant la télévision ?	2,92	2,74

Ce tableau indique la moyenne *Heures par jour devant la télé* pour les personnes qui s'informent ou non avec les journaux. La différence observée entre ces moyennes semble indiquer que les personnes qui ne s'informent pas avec les journaux passent environ 0,18 heure de plus devant leur poste de télévision que les autres. Pour voir si cette différence est due à une variation aléatoire, vérifiez les tests des moyennes de colonne.

Figure 10-12
Comparaisons des moyennes de colonne

	Obtention de l'information via la presse écrite	
	Non	Oui
	(A)	(B)
# d'heures par jour passées devant la télévision ?		

Le tableau des tests des moyennes de colonne attribue une clé (lettre) à chaque modalité de la variable en colonne. Pour la variable *Informations dans les journaux*, la modalité *Non* se voit attribuer la lettre A, tandis que la lettre B est attribuée à la modalité *Oui*. Pour chaque paire de colonnes, les moyennes de colonne sont comparées à l'aide d'un test *t*. Comme il n'y a que deux colonnes, un seul test est exécuté. Pour chaque paire significative, la clé de la modalité dont la moyenne est la plus basse est placée sous la modalité dont la moyenne est la plus élevée. Aucune clé n'étant reportée dans les cellules du tableau, les moyennes de colonne ne sont pas statistiquement différentes.

Résultats de signification dans la notation de style APA

Si vous ne souhaitez afficher les résultats de signification dans un tableau distinct, vous pouvez les afficher dans le tableau principal. Les résultats de signification sont identifiés à l'aide de la notation de style APA avec des indices. Réalisez les étapes précédentes pour comparer les moyennes des colonnes, mais effectuez la modification suivante sur l'onglet Statistiques de test.

- Dans la zone Identifier les différences significatives, sélectionnez Dans le tableau principal à l'aide d'indices de style APA.

Figure 10-13
 Onglet *Statistiques de test* dans lequel l'élément « Comparer les moyennes de colonne (tests-t) » et l'option « Dans le tableau principal à l'aide d'indices » sont sélectionnés



- Cliquez sur OK pour créer le tableau et obtenir les tests des moyennes de colonne avec la notation de style APA.

Figure 10-14
 Comparaison des moyennes de colonnes à l'aide de la notation de style APA

	Obtention des informations à partir de la presse quotidienne	
	Non	Oui
	Moyenne	Moyenne
Nombre d'heures passées à regarder la télévision ?	2.92 _a	2.74 _a

Le tableau des tests des moyennes de colonne attribue un indice aux catégories de la variable de colonne. Pour chaque paire de colonnes, la moyenne des colonnes est comparée à l'aide d'un test t . Si une paire de valeurs est différente de manière significative, les valeurs se voient attribuées des indices *différents*. Comme il n'y a que deux colonnes, un seul test est exécuté. Les moyennes des colonnes de cet exemple partageant le même indice, il apparaît que les moyennes ne sont pas différentes.

Effets de l'emboîtement et de l'empilement sur les tests des moyennes de colonne

La règle suivante s'applique aux tests des moyennes de colonne : Un groupe distinct de tests par paire est exécuté pour chacun des sous-tableaux les plus internes. Pour voir l'impact de l'emboîtement sur les tests, examinez l'exemple précédent, mais dans lequel la variable *Heures par jour devant la télé* est emboîtée dans les niveaux de la variable *Situation d'activité*.

- ▶ Ouvrez à nouveau le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Faites glisser la variable *Situation d'activité* de la liste des variables vers la zone Lignes du panneau de canevas.
- ▶ Cliquez sur OK pour créer le tableau.

Figure 10-15
Comparaisons des moyennes de colonne

			Obtention de l'information via la presse écrite	
			Non (A)	Oui (B)
Statut professionnel	Travail à plein temps	# d'heures par jour passées devant la télévision ?	B	
	Travail à temps partiel	# d'heures par jour passées devant la télévision ?		
	Ne travaille pas temporairement	# d'heures par jour passées devant la télévision ?		
	Chômage	# d'heures par jour passées devant la télévision ?		
	Retraité(e)	# d'heures par jour passées devant la télévision ?		
	Ecolier(e)	# d'heures par jour passées devant la télévision ?		
	Femme/Homme au foyer	# d'heures par jour passées devant la télévision ?		
	Autre	# d'heures par jour passées devant la télévision ?		

Une fois la variable *Heures par jour devant la télé* emboîtée dans les niveaux de la variable *Situation d'activité*, sept groupes de tests des moyennes de colonne sont exécutés : un pour chaque niveau de la variable *Situation d'activité*. Les mêmes clés (lettres) sont attribuées aux modalités de la variable *Informations dans les journaux*. Pour les répondants *qui travaillent à temps plein*, la clé B apparaît dans la colonne A. Par conséquent, pour les salariés à temps plein, la valeur moyenne de la variable *Heures par jour devant la télé* est inférieure pour les personnes qui s'informent avec les journaux. L'absence de toute autre clé dans les colonnes indique qu'il n'existe aucune autre différence statistiquement significative dans les moyennes de colonne.

Ajustements Bonferroni : Lorsque plusieurs tests sont exécutés, l'ajustement Bonferroni est appliqué aux tests des moyennes de colonne pour garantir que le niveau alpha (ou test du faux positif) spécifié dans l'onglet Statistiques de test s'applique à chaque *groupe* de tests. Par

conséquent, dans ce tableau, aucun ajustement Bonferroni n'a été appliqué car, bien que sept groupes de tests soient exécutés, une seule paire de colonnes est comparée au sein de chaque groupe.

Pour voir l'impact de l'empilement sur les tests :

- ▶ Ouvrez à nouveau le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Faites glisser la variable *Informations sur Internet* de la liste des variables vers la zone Colonnes, à gauche de la variable *Informations dans les journaux*.
- ▶ Cliquez sur OK pour créer le tableau.

Figure 10-16
Comparaisons des moyennes de colonne

			Obtention de l'information via internet		Obtention de l'information via la presse écrite	
			Non (A)	Oui (B)	Non (A)	Oui (B)
Statut professionnel	Travail à plein temps	# d'heures par jour passées devant la télévision ?	B		B	
	Travail à temps partiel	# d'heures par jour passées devant la télévision ?	B			
	Ne travaille pas temporairement	# d'heures par jour passées devant la télévision ?				
	Chômage	# d'heures par jour passées devant la télévision ?	B			
	Retraité(e)	# d'heures par jour passées devant la télévision ?	B			
	Ecolier(e)	# d'heures par jour passées devant la télévision ?	B			
	Femme/Homme au foyer	# d'heures par jour passées devant la télévision ?	B			
	Autre	# d'heures par jour passées devant la télévision ?	B			

Une fois la variable *Informations sur Internet* empilée avec la variable *Informations dans les journaux*, 14 groupes de tests des moyennes de colonne sont exécutés : un pour chaque niveau de la variable *Situation d'activité*, pour les variables *Informations sur Internet* et *Informations dans les journaux*. Là encore, aucun ajustement Bonferroni n'est appliqué car, au sein de chaque groupe, une seule paire de colonnes est comparée. Les tests pour la variable *Informations dans les journaux* sont les mêmes que précédemment. Pour la variable *Informations sur Internet*, la lettre A est attribuée à la modalité *Non*, tandis que la lettre B est attribuée à la modalité *Oui*. La clé B est reportée dans la colonne A pour chaque groupe de tests des moyennes de colonne, sauf pour les personnes ayant répondu Temporairement sans activité. Par conséquent, la valeur moyenne de la variable *Heures par jour devant la télé* est plus faible pour les personnes qui s'informent au moyen d'Internet que pour celles qui ne s'informent pas avec les journaux. Aucune clé n'étant

reportée pour le groupe *Temporairement sans activité*, les moyennes de colonne ne sont pas statistiquement différentes pour ces répondants.

Comparaison des proportions de colonne

Les tests des proportions de colonne permettent de déterminer le classement relatif des modalités de la variable qualitative Colonnes sur le plan des proportions de modalité de la variable qualitative Lignes. Par exemple, après avoir utilisé un test du Khi-deux révélant que les variables *Situation d'activité* et *Situation familiale* ne sont pas indépendantes, vous pouvez déterminer les lignes et les colonnes responsables de cette relation.

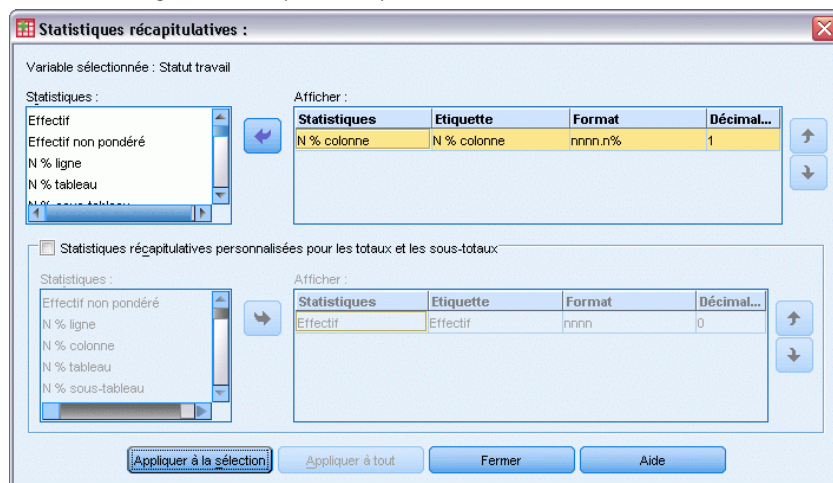
- ▶ A partir du menu, sélectionnez :
Analyse > Tableaux > Tabuler...
- ▶ Cliquez sur Restaurer pour rétablir les valeurs par défaut dans tous les onglets.
- ▶ Dans le générateur de tableaux, faites glisser la variable *Situation d'activité* de la liste des variables vers la zone Lignes du panneau de canevas.
- ▶ Faites glisser la variable *Situation familiale* de la liste des variables vers la zone Colonnes.

Figure 10-17
Variables affichées dans le panneau de canevas

		Statut marital					
		Marié	Veuf/veuve	Divorcé	Séparé	Célibataire	
Statut travail	Plein temps	Effectif	nnnn	nnnn	nnnn	nnnn	nnnn
	Temps partiel	Effectif	nnnn	nnnn	nnnn	nnnn	nnnn
	Ne travaille ...	Effectif	nnnn	nnnn	nnnn	nnnn	nnnn
	Chômage	Effectif	nnnn	nnnn	nnnn	nnnn	nnnn
	Retraité	Effectif	nnnn	nnnn	nnnn	nnnn	nnnn
	Etudiant	Effectif	nnnn	nnnn	nnnn	nnnn	nnnn
	Au foyer	Effectif	nnnn	nnnn	nnnn	nnnn	nnnn
	Autre	Effectif	nnnn	nnnn	nnnn	nnnn	nnnn

- ▶ Sélectionnez *Situation d'activité*, puis cliquez sur Statistiques récapitulatives dans le groupe Définir.

Figure 10-18
Boîte de dialogue Statistiques récapitulatives



- ▶ Dans la liste Statistiques, sélectionnez la statistique N % colonne, puis ajoutez-la à la liste Afficher.
- ▶ Désélectionnez Effectif dans la liste Afficher.
- ▶ Cliquez sur Appliquer à la sélection.
- ▶ Dans la boîte de dialogue Tableaux personnalisés, cliquez sur l'onglet Statistiques de test.

Figure 10-19
Onglet Statistiques de test, dans lequel l'élément Comparer les proportions de colonne (tests-z) est sélectionné

Tabuler

Tableau Titres Tests statistiques Options

Comparer les moyennes de colonne (tests-t)
Alpha de Cronbach : 0,05
 Ajuster les valeurs p pour des comparaisons multiples (méthode Bonferroni)
 Estimer la variance uniquement à partir des modalités comparées (ceci est toujours effectué pour les variables multiréponses)

Comparer les proportions de colonne (tests-z)
Alpha de Cronbach : 0,05
 Ajuster les valeurs p pour des comparaisons multiples (méthode Bonferroni)

Identifier les différences significatives
 Dans un tableau séparé Dans le tableau principal à l'aide des indices de style APA

Tests d'indépendance (Khi-deux) Alpha de Cronbach : 0,05

Utiliser les sous-totaux à la place des catégories de sous-totaux
 Inclure les variables multiréponses dans les tests

i Les tests Khi-deux et les tests de proportion des colonnes s'appliquent aux tableaux dans lesquels des variables qualitatives sont présentes à la fois dans les lignes et dans les colonnes.

Les tests sur les moyennes de colonne s'appliquent aux tableaux dans lesquels des variables d'échelles sont présentes dans les lignes et des variables qualitatives dans les colonnes.

Les tests ne sont pas effectués pour les tableaux dans lesquels les étiquettes de modalité sont déplacées hors de leur dimension de tableau par défaut.

Les totaux sont exclus de tous les tests. Les sous-totaux ne sont utilisés que si les modalités auxquelles ils s'appliquent sont masquées ou spécifiées au-dessus.

Les modalités calculées sont exclues des tests de signification.

OK Coller Réinitialiser Annuler Aide

- ▶ Sélectionnez Comparer les proportions de colonne (tests-z).
- ▶ Cliquez sur OK pour créer le tableau et obtenir les tests des proportions de colonne.

Figure 10-20
Variables Situation d'activité et Situation familiale

		Statut marital				
		Marié(e)	Veuf/Veuve	Divorcé(e)	Séparé(e)	Jamais marié(e)
		% colonne	% colonne	% colonne	% colonne	% colonne
Statut professionnel	Travail à plein temps	57,8%	15,5%	66,1%	62,4%	59,1%
	Travail à temps partiel	10,3%	7,1%	7,8%	9,7%	15,4%
	Ne travaille pas temporairement	1,7%	,7%	2,0%	1,1%	1,7%
	Chômage	1,0%	1,1%	2,2%	,0%	4,8%
	Retraité(e)	12,5%	53,0%	11,9%	6,5%	2,6%
	Ecolier(e)	,7%	,4%	1,6%	2,2%	9,0%
	Femme/Homme au foyer	14,9%	19,4%	5,6%	14,0%	5,3%
	Autre	1,2%	2,8%	2,7%	4,3%	2,1%

Ce tableau est un tableau croisé des variables *Situation d'activité* et *Situation familiale*, dans lequel les proportions de colonne constituent la statistique récapitulative.

Figure 10-21
Comparaisons des proportions de colonne

		Statut marital				
		Marié(e)	Veuf/Veuve	Divorcé(e)	Séparé(e)	Jamais marié(e)
		(A)	(B)	(C)	(D)	(E)
Statut professionnel	Travail à plein temps	B		A B	B	B
	Travail à temps partiel					A B C
	Ne travaille pas temporairement					
	Chômage				. ^a	A B
	Retraité(e)	E	A C D E	E		
	Ecolier(e)					A B C
	Femme/Homme au foyer	C E	C E		C E	
Autre						

Le tableau des tests sur les proportions de colonne attribue une clé (lettre) à chaque modalité des variables en colonne. Pour la variable *Situation familiale*, la lettre A est attribuée à la modalité *Marié*, la lettre B à la modalité *Veuf*, et ainsi de suite jusqu'à la modalité *N'a jamais été marié*, à laquelle est affectée la lettre E. Pour chaque paire de colonnes, les proportions de colonne sont comparées à l'aide d'un test z . Sept groupes de tests des proportions de colonne sont exécutés : un pour chaque niveau de la variable *Situation d'activité*. La variable *Situation familiale* comportant cinq niveaux, $(5*4/2) = 10$ paires de colonnes sont comparées dans chaque groupe de tests et des ajustements Bonferroni permettent d'ajuster les valeurs de signification. Pour chaque paire significative, la clé de la modalité dont la proportion est la plus basse est placée sous la modalité dont la proportion est la plus élevée.

Pour le groupe de tests associé à la réponse *Salarié à temps plein*, la clé B apparaît dans chacune des autres colonnes. En outre, la clé A apparaît dans la colonne C. Aucune autre clé n'est reportée dans les autres colonnes. Par conséquent, vous pouvez arriver à la conclusion que la proportion des personnes divorcées travaillant à temps plein est supérieure à celle des personnes mariées travaillant à temps plein, elle-même supérieure à celle des personnes veuves travaillant à temps plein. La proportion des personnes séparées ou n'ayant jamais été mariées, et travaillant à temps plein, ne peut pas être différenciée de celle des personnes divorcées ou mariées, et travaillant à temps plein ; toutefois, cette proportion est supérieure à celle des personnes veuves travaillant à temps plein.

Pour les tests associés aux réponses *Salarié à temps partiel* ou *Etudiant*, les clés A, B et C apparaissent dans la colonne E. Aucune autre clé n'est reportée dans les autres colonnes. Par conséquent, la proportion des personnes qui n'ont jamais été mariées, et qui fréquentent une école ou travaillent à temps partiel, est supérieure à la proportion des personnes mariées, veuves ou divorcées qui fréquentent une école ou travaillent à temps partiel.

Pour les tests associés aux réponses *Temporairement sans activité* ou *Autre*, aucune autre clé n'est reportée dans les colonnes. Par conséquent, il n'y a pas de différence perceptible entre les proportions des personnes mariées, veuves, divorcées, séparées ou qui n'ont jamais été mariées, et qui sont sans emploi ou dans une situation d'activité non répertoriée.

Les tests associés à la réponse *Retraité* indiquent que la proportion des personnes veuves retraitées est supérieure aux proportions des personnes retraitées relevant des autres situations familiales. En outre, les proportions des personnes retraitées mariées ou divorcées sont supérieures à celle des personnes retraitées n'ayant jamais été mariées.

La proportion des personnes ayant répondu Homme/Femme au foyer est plus importante parmi les personnes mariées, veuves ou séparées que parmi les personnes divorcées ou n'ayant jamais été mariées.

La proportion des personnes qui n'ont jamais été mariées et qui ont répondu *Sans emploi* est supérieure à celles des personnes mariées ou veuves sans emploi. En outre, la colonne *Séparé* comporte le symbole “. “, qui indique que la proportion observée des personnes séparées dans la ligne *Sans emploi* est égale à 0 ou 1 et, par conséquent, que cette colonne ne permet pas d'effectuer des comparaisons pour les répondants sans emploi.

Résultats de signification dans la notation de style APA

Si vous ne souhaitez afficher les résultats de signification dans un tableau distinct, vous pouvez les afficher dans le tableau principal. Les résultats de signification sont identifiés à l'aide de la notation de style APA avec des indices. Réalisez les étapes précédentes pour comparer les proportions des colonnes, mais effectuez la modification suivante sur l'onglet Statistiques de test :

- Dans la zone Identifier les différences significatives, sélectionnez Dans le tableau principal à l'aide d'indices de style APA.

Figure 10-22

Onglet Statistiques de test dans lequel l'élément « Comparer les proportions de colonne (tests-z) » et l'option « Dans le tableau principal à l'aide d'indices » sont sélectionnés



- Cliquez sur OK pour créer le tableau et obtenir les tests des moyennes de colonne avec la notation de style APA.

Figure 10-23

Comparaison des proportions de colonnes à l'aide de la notation de style APA

		Statut marital				
		Marrié	Veuf/veuve	Divorcé	Séparé	Célibataire
		N % colonne	N % colonne	N % colonne	N % colonne	N % colonne
Statut travail	Plein temps	57.8% _a	15.5% _b	66.1% _c	62.4% _{a,c}	59.1% _{a,c}
	Temps partiel	10.3% _a	7.1% _a	7.8% _a	9.7% _{a,b}	15.4% _b
	Ne travaille pas temporairement	1.7% _a	.7% _a	2.0% _a	1.1% _a	1.7% _a
	Chômage	1.0% _a	1.1% _a	2.2% _{a,b}	.0% ¹	4.8% _b
	Retraité	12.5% _a	53.0% _b	11.9% _a	6.5% _{a,c}	2.6% _c
	Etudiant	.7% _a	.4% _a	1.6% _a	2.2% _{a,b}	9.0% _b
	Au foyer	14.9% _a	19.4% _a	5.6% _b	14.0% _a	5.3% _b
	Autre	1.2% _a	2.8% _a	2.7% _a	4.3% _a	2.1% _a

Le tableau des tests des proportions de colonne attribue un indice aux catégories de la variable de colonne. Pour chaque paire de colonnes, les proportions des colonnes sont comparées à l'aide d'un test *z*. Si une paire de valeurs est différente de manière significative, les valeurs se voient attribuées des indices *différents*.

Pour le groupe de tests associé à la réponse *Salarié à temps plein*, la catégorie Veuf apparaît avec un indice non utilisé dans les autres colonnes. De plus, les catégories Marié et Divorcé partagent le même indice, ainsi que les catégories Séparé et N'a jamais été marié. Grâce à ces indices et aux proportions affichées dans le tableau, vous pouvez effectuer la même analyse que pour le premier exemple à base de tableaux distincts. Par conséquent, vous pouvez arriver à la conclusion que la proportion des personnes divorcées travaillant à temps plein est supérieure à celle des personnes mariées travaillant à temps plein, elle-même supérieure à celle des personnes veuves travaillant à temps plein. La proportion des personnes séparées ou n'ayant jamais été mariées, et travaillant à temps plein, ne peut pas être différenciée de celle des personnes divorcées ou mariées, et travaillant à temps plein ; toutefois, cette proportion est supérieure à celle des personnes veuves travaillant à temps plein. Le reste de l'analyse provenant de l'exemple précédent s'applique.

Effets de l'emboîtement et de l'empilement sur les tests des proportions de colonne

La règle suivante s'applique aux tests des proportions de colonne : Un groupe distinct de tests par paire est exécuté pour chacun des sous-tableaux les plus internes. Pour voir l'impact de l'emboîtement sur les tests, examinez l'exemple précédent, mais avec la variable *Situation d'activité* emboîtée dans les niveaux de la variable *Sexe*.

- Ouvrez à nouveau le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- Faites glisser la variable *Sexe* de la liste des variables vers la zone Lignes du panneau de canevas.
- Cliquez sur OK pour créer le tableau.

Figure 10-24
Comparaisons des proportions de colonne

				Statut marital				
				Marié(e) (A)	Veuf/veuve (B)	Divorcé(e) (C)	Séparé(e) (D)	Jamais marié(e) (E)
Sexe du répondant	Homme	Statut professionnel	Travail à plein temps	B		B	B	B
			Travail à temps partiel					A
			Ne travaille pas temporairement		a			
			Chômage		a		a	A
			Retraité(e)	E	A C D E	E		
	Femme	Statut professionnel	Travail à plein temps	B		A B	B	B
			Travail à temps partiel	B				B
			Ne travaille pas temporairement				a	
			Chômage				a	A
			Retraité(e)	E	A C D E	E		
		Femme/Homme au foyer					A C	
		Autre				A		
		Autre				B		
		Autre				a		
		Autre				a	A	
		Autre	E	A C D E	E			
		Autre					A B C	
		Autre	C E	C E		C		

Une fois la variable *Situation d'activité* emboîtée dans les niveaux de la variable *Sexe*, 14 groupes de tests des proportions de colonne sont exécutés : un pour chaque niveau de la variable *Situation d'activité*, pour chaque niveau de la variable *Sexe*. Les mêmes clés (lettres) sont attribuées aux modalités de la variable *Situation familiale*.

Les résultats du tableau font apparaître les deux points suivants :

- L'exécution d'un nombre supérieur de tests génère davantage de colonnes avec une proportion de colonne égale à zéro. Ceci concerne essentiellement les répondants séparés et les hommes veufs.
- Les différences de colonne précédemment observées parmi les personnes ayant répondu *Homme/Femme au foyer* semblent être entièrement imputables aux personnes de sexe féminin.

Pour voir l'impact de l'empilement sur les tests :

- ▶ Ouvrez à nouveau le générateur de tableaux (menu Analyse, Tableaux, Tableaux personnalisés).
- ▶ Faites glisser la variable *Diplôme le plus important* de la liste des variables vers la zone Lignes, sous la variable *Sexe*.
- ▶ Cliquez sur OK pour créer le tableau.

Figure 10-25
Comparaisons des proportions de colonne

				Statut marital				
				Marié(e) (A)	Veuf/Veuve (B)	Divorcé(e) (C)	Séparé(e) (D)	Jamais marié(e) (E)
Sexe du répondant	Homme	Statut professionnel	Travail à plein temps	B		B	B	B
			Travail à temps partiel					A
			Ne travaille pas temporairement		a			
			Chômage		a		a	A
			Retraité(e)	E	A C D E	E		
	Femme	Statut professionnel	Travail à plein temps	B		AB	B	B
			Travail à temps partiel	B				B
			Ne travaille pas temporairement				a	
			Chômage				a	A
			Retraité(e)	E	A C D E	E		
Diplôme le plus élevé	< Bac	Autre	Ecuyer(e)		a			AC
			Femme/Homme au foyer					
			Autre				A	
			Travail à plein temps	B		AB	B	B
			Travail à temps partiel	B				B
Diplôme le plus élevé	Bac	Autre	Ne travaille pas temporairement				a	
			Chômage				a	A
			Retraité(e)	E	A C D E	E		
			Ecuyer(e)					
			Femme/Homme au foyer	C E	C E		C	ABC
Diplôme le plus élevé	Licence	Autre	Autre					
			Autre					
			Autre					
			Autre					
Diplôme le plus élevé	Maîtrise	Autre	Autre					
			Autre					
			Autre					
			Autre					
Diplôme le plus élevé	Bac + 5	Autre	Autre					
			Autre					
			Autre					
			Autre					

Une fois la variable *Diplôme le plus important* empilée avec la variable *Sexe*, 19 groupes de tests des moyennes de colonne sont exécutés : les 14 tests indiqués précédemment, ainsi qu'un test pour chaque niveau de la variable *Diplôme le plus important*. Les mêmes clés (lettres) sont attribuées aux modalités de la variable *Situation familiale*.

Les résultats du tableau font apparaître les points suivants :

- Les résultats des 14 groupes de tests précédemment exécutés sont les mêmes.
- La proportion des personnes dont le niveau d'études est inférieur à un diplôme d'école secondaire est plus élevée chez les personnes veuves que parmi les répondants mariés, divorcés ou n'ayant jamais été mariés.
- La proportion des personnes ayant suivi un enseignement post-secondaire semble plus importante parmi les personnes mariées, divorcées et n'ayant jamais été mariées que parmi les personnes veuves.

Une remarque sur les pondérations et les vecteurs multiréponses

Les pondérations sont toujours basées sur l'effectif et non sur les réponses, et ce même si l'une des variables est une variable multiréponses.

Vecteurs de réponses multiples

Les options Tableaux personnalisés et le Générateur de diagrammes prennent également en charge un type spécifique de variable nommé **vecteur de réponses multiples**. Les vecteurs multiréponses ne sont pas réellement des “ variables “ au sens habituel du terme. En effet, vous ne pouvez pas les voir dans l’éditeur de données et les autres procédures ne les reconnaissent pas. Les vecteurs multiréponses utilisent plusieurs variables pour enregistrer les réponses à des questions auxquelles le répondant peut donner plusieurs réponses. Les vecteurs multiréponses sont traités comme des variables qualitatives, et la plupart des actions appliquées à ces dernières peuvent également l’être aux vecteurs multiréponses.

Les vecteurs multiréponses sont construits à partir de plusieurs variables dans le fichier de données. Un vecteur multiréponses est une structure spéciale dans un fichier de données. Vous pouvez définir et enregistrer les vecteurs multiréponses dans un fichier de données IBM® SPSS® Statistics, mais vous ne pouvez ni importer ni exporter des vecteurs multiréponses à partir/vers d’autres formats de fichiers. (Vous pouvez copier les vecteurs multiréponses d’autres fichiers de données SPSS Statistics via l’option Copier des propriétés de données accessibles à partir du menu Données dans la fenêtre de l’Editeur de données.

Fichier de données d’exemple

Les exemples de ce chapitre utilisent le fichier de données *survey_sample.sav*. [Pour plus d’informations, reportez-vous à la section Fichiers d’exemple dans l’annexe A sur p. 197.](#)

Tous les exemples fournis ici affichent des étiquettes de variable dans des boîtes de dialogue ; ces étiquettes sont triées par ordre alphabétique. Les propriétés d’affichage de la liste des variables sont spécifiées dans l’onglet Général de la boîte de dialogue Options (menu Modifier, Options).

Effectifs, réponses, pourcentages et totaux

Toutes les statistiques récapitulatives disponibles pour les variables qualitatives le sont également pour les vecteurs multiréponses. Des statistiques supplémentaires sont également disponibles pour les vecteurs multiréponses.

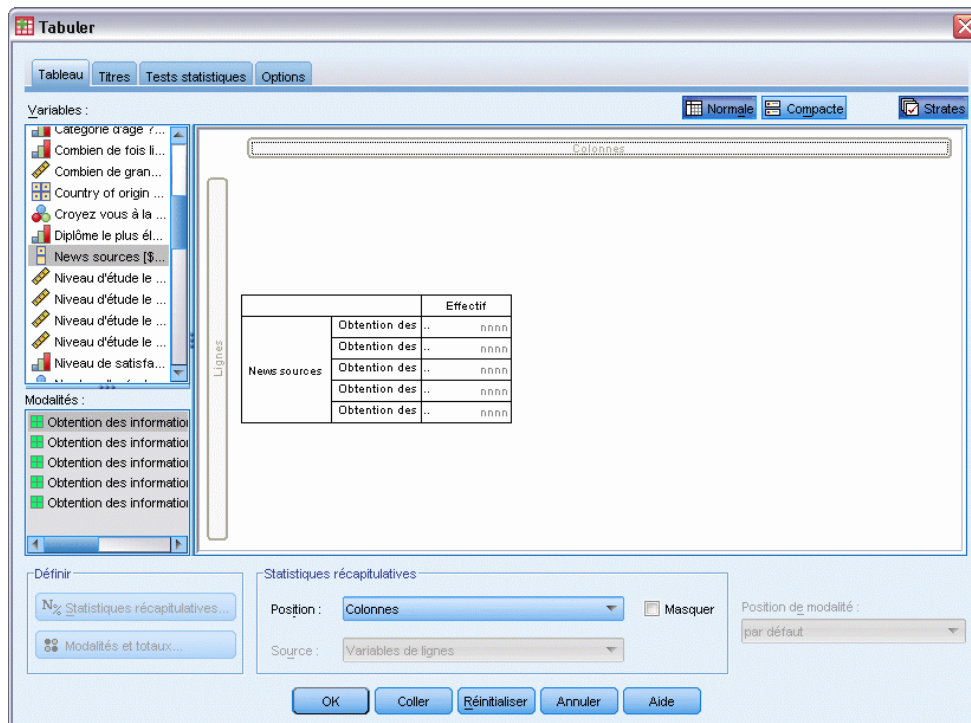
- ▶ A partir du menu, sélectionnez :
Analyse > Tableaux > Tabuler...
- ▶ Faites glisser *Sources d’informations* (étiquette descriptive du vecteur multiréponses *\$mltnews*) de la liste des variables vers la zone Lignes du panneau de canevas.

Dans la liste de variables, l’icône figurant en regard de la “ variable “ identifie celle-ci comme étant un vecteur de dichotomies multiples.

Figure 11-1
Icône du vecteur de dichotomies multiples



Figure 11-2
Vecteur de dichotomies multiples affiché dans l'aperçu du tableau



Pour un vecteur de dichotomies multiples, chaque "modalité" est en fait une variable distincte et les étiquettes de modalité sont les étiquettes de variable (ou noms de variable ne comportant aucune étiquette définie). Dans cet exemple, les effectifs qui seront affichés représentent le nombre d'observations recevant la réponse *Oui* pour chaque variable du vecteur.

- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Sources d'informations* dans l'aperçu du tableau et sélectionnez Modalités et totaux dans le menu contextuel.
- ▶ Dans la boîte de dialogue Modalités et totaux, sélectionnez Total en cliquant dessus, puis cliquez sur Appliquer.
- ▶ Cliquez à nouveau avec le bouton droit de la souris sur *Sources d'informations* et sélectionnez Statistiques récapitulatives dans le menu contextuel.
- ▶ Dans la boîte de dialogue Statistiques récapitulatives, sélectionnez N % colonne dans la liste Statistiques et cliquez sur la flèche pour ajouter l'élément à la liste Afficher.
- ▶ Cliquez sur Appliquer à la sélection, puis sur OK pour créer le tableau.

Figure 11-3
Dichotomies multiples - Effectifs et pourcentages en colonne

		Effectif	% colonne
Sources d'informations	Obtention de l'information via internet	867	41,7%
	Obtention de l'information via la radio	551	26,5%
	Obtention de l'information via la télévision	1077	51,8%
	Obtention de l'information via la presse people	294	14,1%
	Obtention de l'information via la presse écrite	805	38,7%
	Total	2081	100,0%

Totaux ne pouvant pas s'additionner

Si vous examinez les chiffres dans le tableau, vous pouvez remarquer qu'il existe une différence relativement importante entre les " totaux " et les valeurs soi-disant totalisées ; plus précisément, les totaux sont beaucoup plus bas qu'ils ne devraient l'être. Ceci est dû au fait que l'effectif de chaque " modalité " du tableau correspond au nombre d'observations dont la valeur est 1 (c'est-à-dire dont la réponse est *Oui*) pour cette variable, et que le nombre total de réponses *Oui* pour les cinq variables du vecteur de dichotomies multiples pourrait facilement excéder le nombre total d'observations dans le fichier de données.

Toutefois, l'" effectif " total correspond au nombre total d'observations dont la réponse est *Oui* pour au moins une variable du vecteur, ce nombre ne pouvant jamais dépasser le nombre total d'observations dans le fichier de données. Dans cet exemple, le total de 2 081 est inférieur de presque 800 unités au nombre total d'observations dans le fichier de données. Si aucune de ces variables ne comporte de valeurs manquantes, cela signifie que près de 800 répondants ont indiqué qu'ils ne tirent leurs informations d'aucune de ces sources. Le total est la base des pourcentages en colonne ; par conséquent, dans cet exemple, la somme des pourcentages en colonne est supérieure aux 100 % affichés pour le pourcentage en colonne total.

Totaux pouvant s'additionner

Bien que le terme " effectif " n'est pas ambigu, l'exemple ci-dessus montre qu'il peut être source de confusion s'agissant des totaux dans les vecteurs multiréponses, pour lesquels la *réponse* est souvent la statistique récapitulative que vous recherchez vraiment.

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tabuler, Tableaux personnalisés).
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Sources d'informations* dans l'aperçu du tableau et sélectionnez Statistiques récapitulatives dans le menu contextuel.
- ▶ Dans la boîte de dialogue Statistiques récapitulatives, sélectionnez Réponses dans la liste Statistiques et cliquez sur la flèche pour ajouter cet élément à la liste Afficher.
- ▶ Dans la liste Statistiques, sélectionnez % réponses colonne et cliquez sur la flèche pour ajouter cet élément à la liste Afficher.
- ▶ Cliquez sur Appliquer à la sélection, puis sur OK pour créer le tableau.

Figure 11-4
Dichotomies multiples - Réponses et réponses de la colonne en pourcentage

		Effectif	% colonne	Réponses	% réponses colonne
Sources d'informations	Obtention de l'information via internet	867	41,7%	867	24,1%
	Obtention de l'information via la radio	551	26,5%	551	15,3%
	Obtention de l'information via la télévision	1077	51,8%	1077	30,0%
	Obtention de l'information via la presse people	294	14,1%	294	8,2%
	Obtention de l'information via la presse écrite	805	38,7%	805	22,4%
Total		2081	100,0%	3594	100,0%

Pour chaque “ modalité “ du vecteur de dichotomies multiples, *Réponses* est identique à *Effectif* ; il en sera toujours ainsi dans les vecteurs de dichotomie multiples. Cependant, les totaux sont très différents. Le nombre total de réponses est de 3 594 ; il est donc supérieur de plus de 1 500 unités au total et de plus de 700 unités au nombre total d’observations dans le fichier de données.

En ce qui concerne les pourcentages, les totaux de *N % colonne* et de *% réponses colonne* sont tous deux de 100 % ; cependant, les pourcentages pour chaque catégorie du vecteur de dichotomies multiples sont bien inférieurs pour les réponses de la colonne en pourcentage. Ceci est dû au fait que la base de pourcentage pour les réponses de la colonne en pourcentage est le nombre total de réponses (dans ce cas 3 594), ce qui donne des pourcentages bien inférieurs à la base des pourcentages en colonne, qui est de 2 081.

Totaux de pourcentages supérieurs à 100 %

Les pourcentages en colonne et les réponses de la colonne en pourcentage génèrent des totaux de 100 % même si, dans notre exemple, la somme des valeurs individuelles de la colonne *N % colonne* est largement supérieure à 100 %. Par conséquent, que se passe-t-il si vous souhaitez afficher les pourcentages sur la base du total plutôt que sur le nombre total de réponses, mais que vous souhaitez également que le pourcentage “ total “ reflète précisément la somme des pourcentages de chaque modalité ?

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tabuler, Tableaux personnalisés).
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Sources d'informations* dans l’aperçu du tableau et sélectionnez Statistiques récapitulatives dans le menu contextuel.
- ▶ Dans la boîte de dialogue Statistiques récapitulatives, sélectionnez % réponses colonne (Base : Effectif) dans la liste Statistiques et cliquez sur la flèche pour ajouter cet élément à la liste Afficher.
- ▶ Cliquez sur Appliquer à la sélection, puis sur OK pour créer le tableau.

Figure 11-5
Réponses de la colonne en pourcentage, dont la base est l'effectif

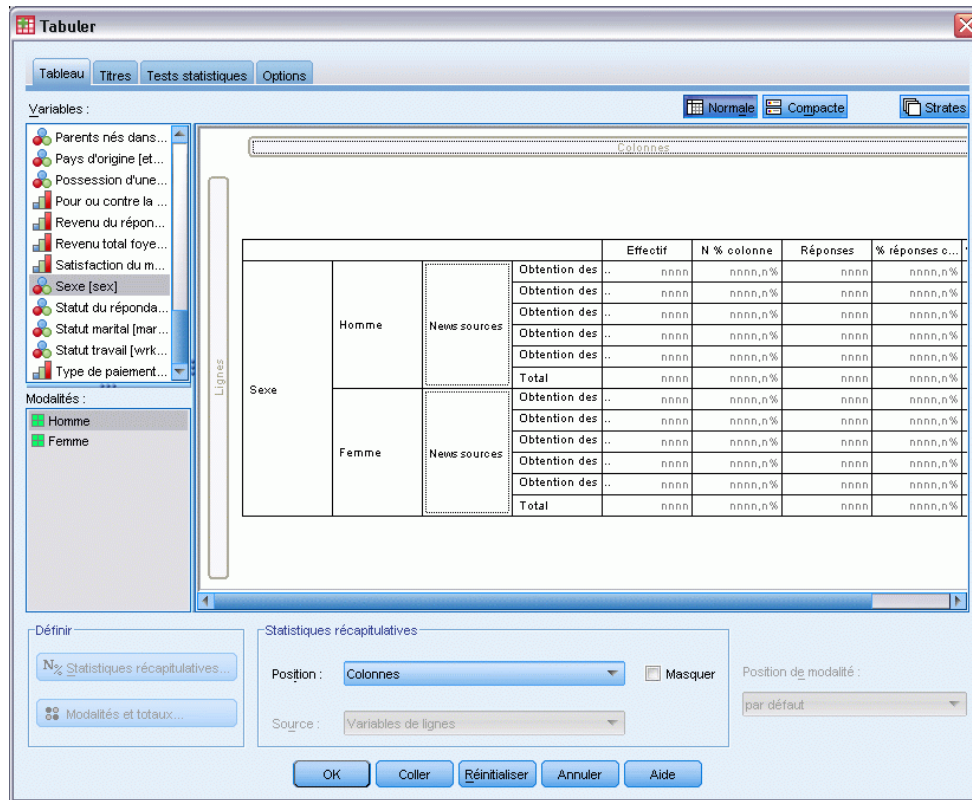
		Effectif	% colonne	Réponses	% réponses colonne	% réponses colonne (Base : Effectif)
Sources d'informations	Obtention de l'information via internet	867	41,7%	867	24,1%	41,7%
	Obtention de l'information via la radio	551	26,5%	551	15,3%	26,5%
	Obtention de l'information via la télévision	1077	51,8%	1077	30,0%	51,8%
	Obtention de l'information via la presse people	294	14,1%	294	8,2%	14,1%
	Obtention de l'information via la presse écrite	805	38,7%	805	22,4%	38,7%
	Total	2081	100,0%	3594	100,0%	172,7%

Utilisation de vecteurs multiréponses avec d'autres variables

En général, vous pouvez utiliser plusieurs vecteurs multiréponses comme des variables qualitatives. Par exemple, vous pouvez croiser un vecteur multiréponses avec une variable qualitative, ou emboîter un vecteur multiréponses dans une variable qualitative.

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tabuler, Tableaux personnalisés).
- ▶ Faites glisser la variable *Sexe* de la liste des variables vers la gauche de la zone Lignes dans le panneau d'aperçu, emboîtant ainsi le vecteur multiréponses *Sources d'informations* dans les modalités de la variable *Sexe*.

Figure 11-6
Aperçu dans le tableau du vecteur multiréponses emboîté



- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Sexe* dans l'aperçu du tableau et désélectionnez *Afficher l'étiquette des variables* dans le menu contextuel.
- ▶ Répétez l'opération pour *Sources d'informations*.
Cette opération entraîne la suppression des colonnes comportant les étiquettes de variable du tableau (puisqu'elles ne sont pas vraiment nécessaires dans ce cas).
- ▶ Cliquez sur OK pour créer le tableau.

Figure 11-7
Vecteurs multiréponses emboîtés dans une variable qualitative

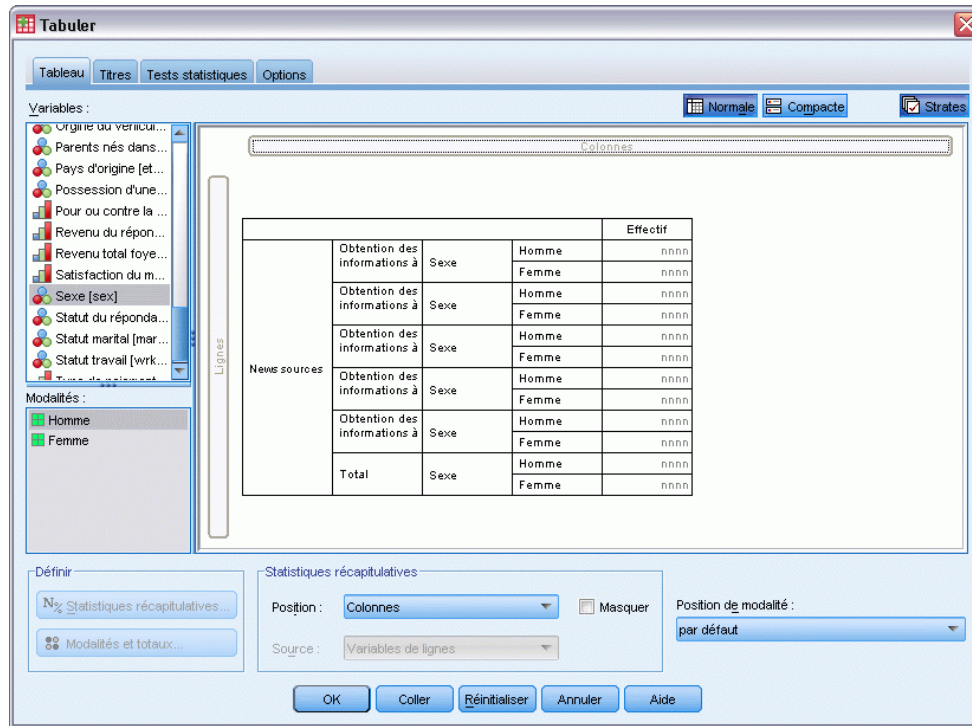
		Effectif	N % colonne	Réponses	% réponses colonne	% réponses colonne (Base : Effectif)
Homme	Obtention des informations à partir d'Internet	359	40,1%	359	23,3%	40,1%
	Obtention des informations à partir de la radio	233	26,0%	233	15,1%	26,0%
	Obtention des informations à partir de la télévision	451	50,3%	451	29,3%	50,3%
	Obtention des informations à partir des magazines	121	13,5%	121	7,9%	13,5%
	Obtention des informations à partir de la presse quotidienne	375	41,9%	375	24,4%	41,9%
	Total	896	100,0%	1539	100,0%	171,8%
Femme	Obtention des informations à partir d'Internet	508	42,9%	508	24,7%	42,9%
	Obtention des informations à partir de la radio	318	26,8%	318	15,5%	26,8%
	Obtention des informations à partir de la télévision	626	52,8%	626	30,5%	52,8%
	Obtention des informations à partir des magazines	173	14,6%	173	8,4%	14,6%
	Obtention des informations à partir de la presse quotidienne	430	36,3%	430	20,9%	36,3%
	Total	1185	100,0%	2055	100,0%	173,4%

Variable source de statistiques et statistiques récapitulatives disponibles

En l'absence d'une variable d'échelle dans un tableau, les variables qualitatives et les vecteurs multiréponses sont traités de la même manière en ce qui concerne la variable source de statistiques : La variable emboîtée la plus interne dans la dimension source des statistiques est la variable source de statistiques. Dans la mesure où certaines statistiques récapitulatives peuvent être affectées uniquement à des vecteurs multiréponses, le vecteur multiréponses doit être la variable emboîtée la plus interne dans la dimension source des statistiques si vous souhaitez obtenir l'une des statistiques récapitulatives multiréponses spéciales.

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tabuler, Tableaux personnalisés).
- ▶ Dans l'aperçu du tableau du panneau de canevas, faites glisser *Sources d'informations* vers la gauche de *Sexe* pour modifier l'ordre d'emboîtement.

Figure 11-8
Variable qualitative emboîtée dans un vecteur multiréponses



Toutes les statistiques récapitulatives multiréponses spéciales (réponses, réponses de la colonne en pourcentage) sont supprimées de l’aperçu du tableau car la variable qualitative *Sexe* est désormais la variable emboîtée la plus interne et, par conséquent, la variable source des statistiques.

Heureusement, le générateur de tableaux “ se souvient “ de ces paramètres. Si vous remplacez la variable *Sources d’informations* à sa position précédente, c’est-à-dire emboîtée dans la variable *Sexe*, toutes les statistiques récapitulatives liées aux réponses sont restaurées dans l’aperçu du tableau.

Vecteurs de modalités multiples et réponses en double

Les vecteurs de modalités multiples comportent une fonction non disponible pour les vecteurs de dichotomies multiples : cette fonction permet de compter les réponses en double. Dans de nombreux cas, les réponses en double dans les vecteurs de modalités multiples représentent probablement des erreurs de codage. Par exemple, dans le cadre de la question “ Selon vous, quels sont les trois pays à l’origine des meilleures voitures ? “, une réponse du type *Suède, Allemagne et Suède* n’est probablement pas valide.

Cependant, dans d’autres cas, des réponses en double peuvent être parfaitement valides. Par exemple, pour la question “ Dans quels pays vos trois dernières voitures ont-elles été fabriquées ? “, une réponse du type *Suède, Allemagne et Suède* est parfaitement logique.

Dans les tableaux personnalisés, vous pouvez choisir les réponses en double dans les vecteurs de modalités multiples. Par défaut, les réponses en double ne sont pas comptées, mais vous pouvez demander à les inclure.

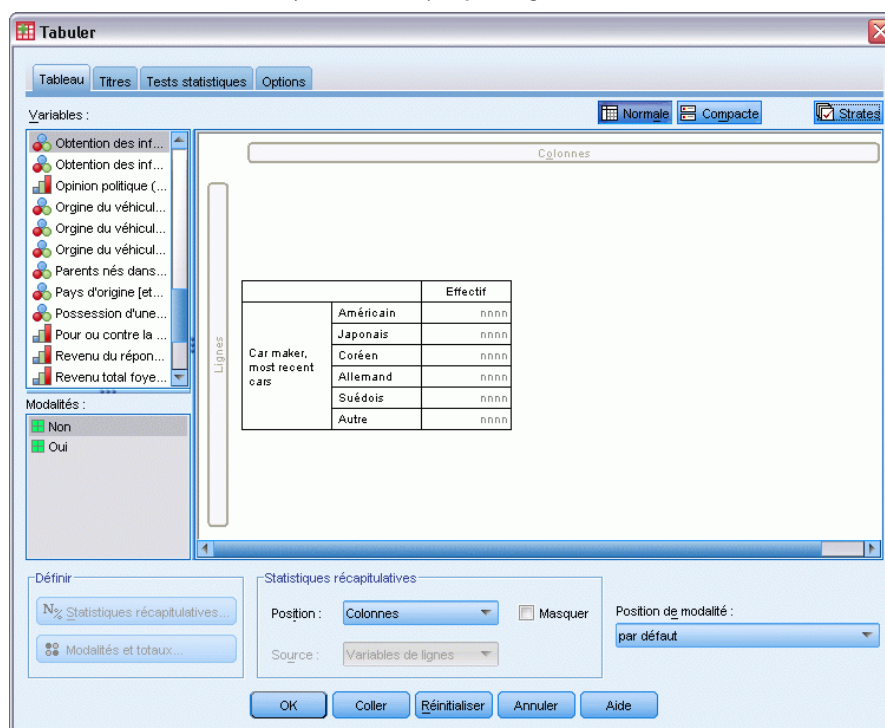
- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tabuler, Tableaux personnalisés).
- ▶ Cliquez sur Restaurer pour effacer les paramètres précédents.
- ▶ Faites glisser *Constructeur automobile, voitures les plus récentes* de la liste des variables vers la zone Lignes du panneau de canevas.

Dans la liste des variables, l'icône figurant en regard de la "variable" identifie celle-ci comme étant un vecteur de modalités multiples.

Figure 11-9
Icône du vecteur de modalités multiples



Figure 11-10
Vecteur de modalités multiples dans l'aperçu du générateur de tableaux



Pour les vecteurs de modalités multiples, les modalités affichées représentent l'ensemble commun d'étiquettes de valeurs définies pour toutes les variables du vecteur (pour les vecteurs de dichotomies multiples, les "modalités" sont en réalité les étiquettes de variable de chaque variable du vecteur).

- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Constructeur automobile, voitures les plus récentes* dans l'aperçu du tableau et sélectionnez Modalités et totaux dans le menu contextuel.
- ▶ Dans la boîte de dialogue Modalités et totaux, sélectionnez Total en cliquant dessus, puis cliquez sur Appliquer.

- ▶ Cliquez de nouveau avec le bouton droit de la souris sur *Constructeur automobile, voitures les plus récentes* et sélectionnez *Statistiques récapitulatives* dans le menu contextuel.
- ▶ Dans la boîte de dialogue *Statistiques récapitulatives*, sélectionnez *Réponses* dans la liste *Statistiques* et cliquez sur la flèche pour ajouter cet élément à la liste *Afficher*.
- ▶ Cliquez sur *Appliquer à la sélection*, puis sur *OK* pour créer le tableau.

Figure 11-11

Vecteur de modalités multiples : Effectifs et réponses sans doublons

		Effectif	Réponses
Constructeur automobile, voitures les plus récentes	Américain	1938	1938
	Japonnais	1327	1327
	Coréen	695	695
	Allemand	693	693
	Suèdois	360	360
	Autre	343	343
	Total	2832	5356

Par défaut, les réponses en double ne sont pas comptées ; par conséquent, dans ce tableau, les valeurs de chaque modalité dans les colonnes *Effectif* et *Réponses* sont identiques. Seuls les totaux diffèrent.

- ▶ Ouvrez le générateur de tableaux (menu *Analyse, Tabuler, Tableaux personnalisés*).
- ▶ Cliquez sur l'onglet *Options*.
- ▶ Cliquez sur *Dénombrer les réponses en double* pour les vecteurs de modalités multiples.
- ▶ Cliquez sur *OK* pour créer le tableau.

Figure 11-12

Vecteur de modalités multiples dans lesquels les réponses en double sont incluses

		Effectif	Réponses
Constructeur automobile, voitures les plus récentes	Américain	1938	2797
	Japonnais	1327	1717
	Coréen	695	760
	Allemand	693	754
	Suèdois	360	383
	Autre	343	359
	Total	2832	6770

Dans ce tableau, on remarque une différence notable entre les valeurs des colonnes *Effectif* et *Réponses*, en particulier pour les voitures américaines, ce qui indique que de nombreux répondants ont possédé plusieurs voitures américaines.

Test de signification pour les vecteurs de réponses multiples

Vous pouvez utiliser des vecteurs de réponses multiples dans des tests de signification de la même façon que vous utiliseriez des variables qualitatives.

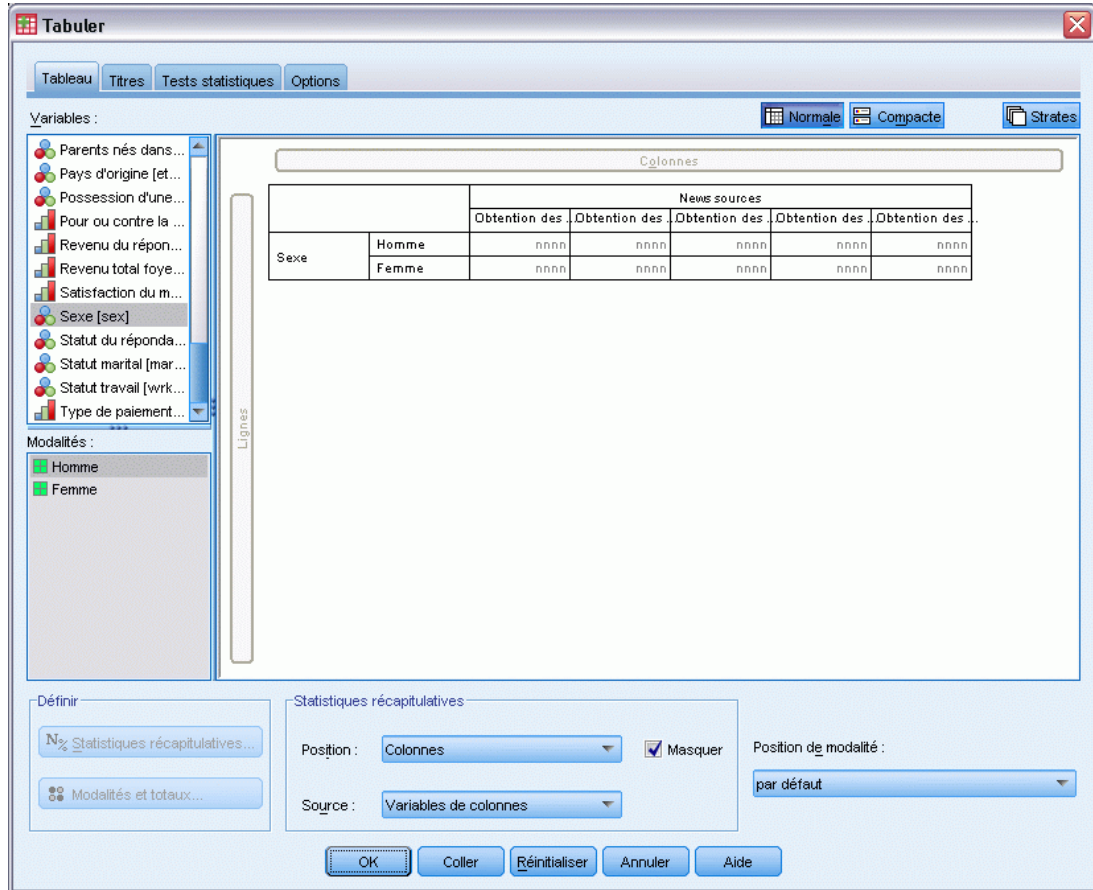
- Pour les tests d'indépendance (Khi-deux) ou pour la comparaison des proportions de colonne (tests-z), les tests sont réalisés sur la base des effectifs et l'Effectif doit être une des statistiques récapitulatives affichées dans le tableau.
- Pour les vecteurs de catégories multiples, les tests de comparaison des proportions de colonne ou des moyennes de colonnes (t-tests) ne sont pas réalisés si l'option Dénombrer les réponses en double pour les vecteurs de catégories multiples est sélectionnée dans l'onglet Options. [Pour plus d'informations, reportez-vous à la section Tableaux personnalisés : Onglet Options dans le chapitre 2 sur p. 49.](#)

Tests d'indépendance avec des vecteurs de réponses multiples

Cet exemple crée un tableau croisé de variables qualitatives et un vecteur de réponses multiples exécute et effectue un test d'indépendance Khi-deux sur le tableau croisé.

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tabuler, Tableaux personnalisés).
- ▶ Cliquez sur Restaurer pour effacer les paramètres précédents.
- ▶ Faites glisser *Sources d'informations* (étiquette descriptive du vecteur de dichotomies multiples *\$mltnews*) de la liste des variables vers la zone Colonnes du panneau de canevas.
- ▶ Faites glisser la variable *Sexe* de la liste des variables vers la zone Lignes du panneau de canevas.

Figure 11-13
Nouvelles sources (vecteur de réponses multiples) et Sexe sélectionnés en tant que variables de ligne et de colonne



- ▶ Cliquez sur l'onglet Statistiques de test.
- ▶ Sélectionnez (cochez) Tests d'indépendance (Khi-deux).
- ▶ Si cela n'est pas déjà fait, sélectionnez Inclure des variables de réponses multiples dans le test.

Figure 11-14
Onglet Statistiques de test

Tabuler

Tableau Titres **Tests statistiques** Options

Comparer les moyennes de colonne (tests-t)
Alpha de Cronbach : 0.05
 Ajuster les valeurs p pour des comparaisons multiples (méthode Bonferroni)
 Estimer la variance uniquement à partir des modalités comparées (ceci est toujours effectué pour les variables multiréponses)

Comparer les proportions de colonne (tests-z)
Alpha de Cronbach : 0.05
 Ajuster les valeurs p pour des comparaisons multiples (méthode Bonferroni)

Identifier les différences significatives
 Dans un tableau séparé Dans le tableau principal à l'aide des indices de style APA

Tests d'indépendance (Khi-deux) Alpha de Cronbach : 0.05

Utiliser les sous-totaux à la place des catégories de sous-totaux
 Inclure les variables multiréponses dans les tests

i Les tests Khi-deux et les tests de proportion des colonnes s'appliquent aux tableaux dans lesquels des variables qualitatives sont présentes à la fois dans les lignes et dans les colonnes.

Les tests sur les moyennes de colonne s'appliquent aux tableaux dans lesquels des variables d'échelles sont présentes dans les lignes et des variables qualitatives dans les colonnes.

Les tests ne sont pas effectués pour les tableaux dans lesquels les étiquettes de modalité sont déplacées hors de leur dimension de tableau par défaut.

Les totaux sont exclus de tous les tests. Les sous-totaux ne sont utilisés que si les modalités auxquelles ils s'appliquent sont masquées ou spécifiées au-dessus.

Les modalités calculées sont exclues des tests de signification.

OK Coller Réinitialiser Annuler Aide

- Cliquez sur OK pour exécuter la procédure.

Figure 11-15
Résultats Khi-deux

		News sources
Sexe	Chi-Deux	10,266
	ddl	5
	Sig.	,068

Le niveau de signification de 0,068 pour le test Khi-deux indique que les hommes et les femmes ne diffèrent probablement pas de manière significative dans leurs choix de nouvelles sources (en considérant que vous utilisez une valeur significative de 0,05 ou inférieur comme critère de détermination de la signification statistique)

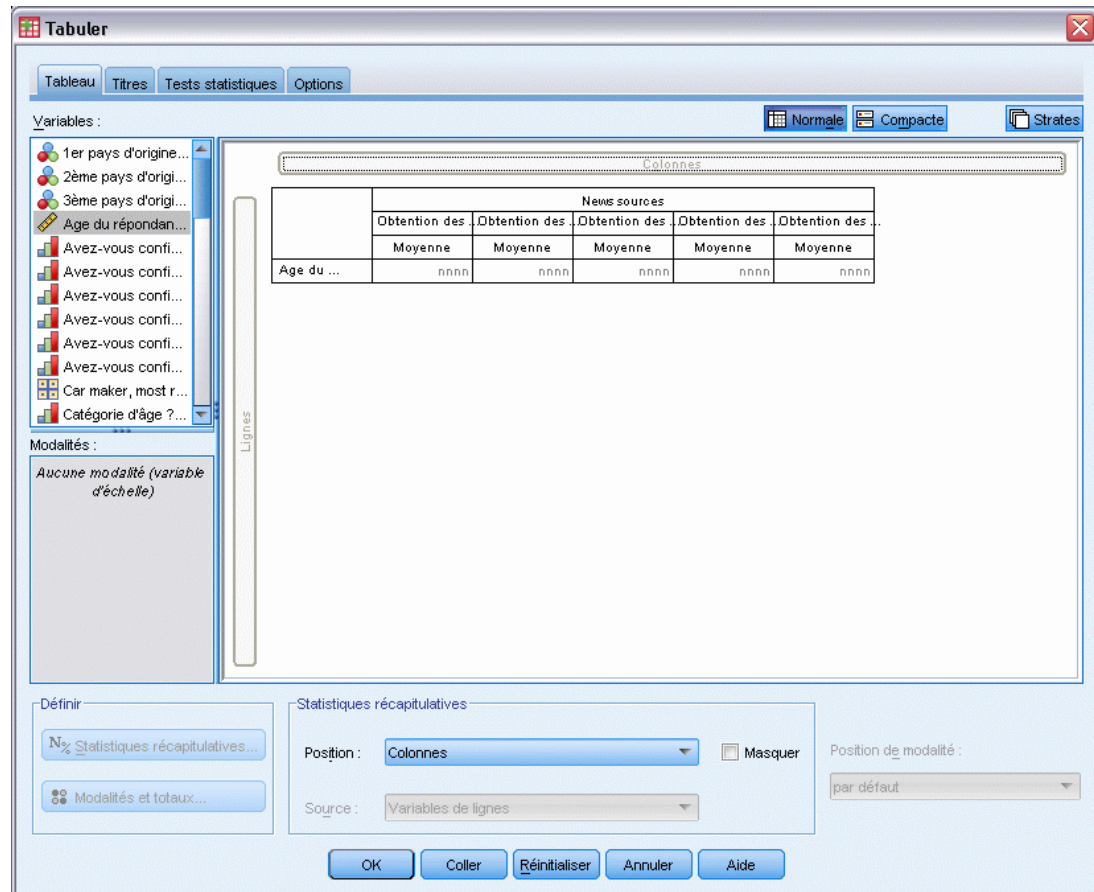
Comparaison des moyennes de colonne avec des vecteurs de réponses multiples

Cet exemple calcule les moyennes d'une variable d'échelle au sein des catégories définies par un vecteur de réponses multiples et compare la moyenne de chaque catégorie avec celle des autres catégories afin de déterminer des différences significatives.

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tabuler, Tableaux personnalisés).
- ▶ Cliquez sur Restaurer pour effacer les paramètres précédents.
- ▶ Faites glisser *Sources d'informations* (étiquette descriptive du vecteur de dichotomies multiples *Smltnews*) de la liste des variables vers la zone Colonnes du panneau de canevas.
- ▶ Faites glisser *Age du répondant* vers la zone Lignes du panneau de canevas.

Figure 11-16

Comparaison de l'âge moyen à travers des catégories de réponses multiples



- ▶ Cliquez sur l'onglet Statistiques de test.
- ▶ Sélectionnez (cochez) Comparer les moyennes de colonne (tests-t).
- ▶ Si cela n'est pas déjà fait, sélectionnez Inclure des variables de réponses multiples dans le test.
- ▶ Cliquez sur OK pour exécuter la procédure.

Figure 11-17
Résultats du test de signification

News sources					
Obtention des informations à partir d'internet	Obtention des informations à partir de la radio	Obtention des informations à partir de la télévision	Obtention des informations à partir des magazines	Obtention des informations à partir de la presse quotidienne	
Moyenne	Moyenne	Moyenne	Moyenne	Moyenne	
Age du répondant	40	40	48	40	52

Comparaisons des moyennes de colonne				
News sources				
Obtention des informations à partir d'internet	Obtention des informations à partir de la radio	Obtention des informations à partir de la télévision	Obtention des informations à partir des magazines	Obtention des informations à partir de la presse quotidienne
(A)	(B)	(C)	(D)	(E)
Age du répondant		A B D		A B C D

Les résultats sont basés sur des tests à deux phases, en supposant des variances égales, avec le seuil de signification 0.05. Pour chaque paire significative, la clé de la modalité la plus petite apparaît sous la modalité dont la moyenne est la plus élevée.

- Chaque catégorie de vecteur de réponses multiples est identifiée par une lettre (A, B, C, D, E) ; pour chaque catégorie pour laquelle la moyenne d'une autre catégorie est à la fois inférieure et diffère de manière significative de la moyenne de la catégorie en question, la lettre représentant la catégorie comportant la moyenne la plus basse est affichée.
- *Informations dans les journaux* (A) comprend l'âge moyen le plus élevé et toutes les autres catégories en diffèrent de manière significative.
- *Informations à la télévision* (C) comprend l'âge moyen suivant le plus élevé et toutes les autres moyennes de catégories (B, D et E) en diffèrent de manière significative. (C diffère également de manière significative de A, comme précédemment indiqué.)
- Les âges moyens pour *Informations dans les magazines* (B), *Informations à la radio* (D) et *Informations sur Internet* (E) ne diffèrent pas de manière significative les uns des autres.

Valeurs manquantes

De nombreux fichiers de données contiennent une certaine quantité de données manquantes. De nombreux facteurs peuvent être à l'origine des données manquantes. Par exemple, il arrive que des personnes interrogées ne répondent pas à chaque question, que certaines variables ne soient pas applicables à certaines observations et que des erreurs de codage aboutissent à la suppression de certaines valeurs.

Il existe deux sortes de valeur manquante dans IBM® SPSS® Statistics :

- **Valeurs manquantes spécifiées par l'utilisateur.** Valeurs définies comme contenant des données manquantes. Des étiquettes de valeurs peuvent être affectées à ces valeurs pour identifier la raison de l'absence des données (par exemple, le code 99 et l'étiquette de valeur *Sans objet* pour la grossesse chez les hommes).
- **Manquante par défaut.** Si aucune valeur n'est présente pour une variable numérique, la valeur manquante par défaut lui est affectée. Ceci est indiqué par un point dans l'affichage des données de l'éditeur de données.

Un ensemble de fonctionnalités permet de compenser les effets des données manquantes et d'analyser le type de ces dernières. Toutefois, l'objet de cette de ce chapitre est beaucoup plus simple : décrire comment la fonction Tableaux personnalisés gère les données manquantes et comment celles-ci affectent le calcul des statistiques récapitulatives.

Fichier de données d'exemple

Les exemples de du chapitre utilisent le fichier de données *missing_values.sav*. [Pour plus d'informations, reportez-vous à la section Fichiers d'exemple dans l'annexe A sur p. 197.](#) Il s'agit d'un fichier extrêmement simple et complètement artificiel, comportant une seule variable et dix observations, conçu pour illustrer les concepts de base des valeurs manquantes.

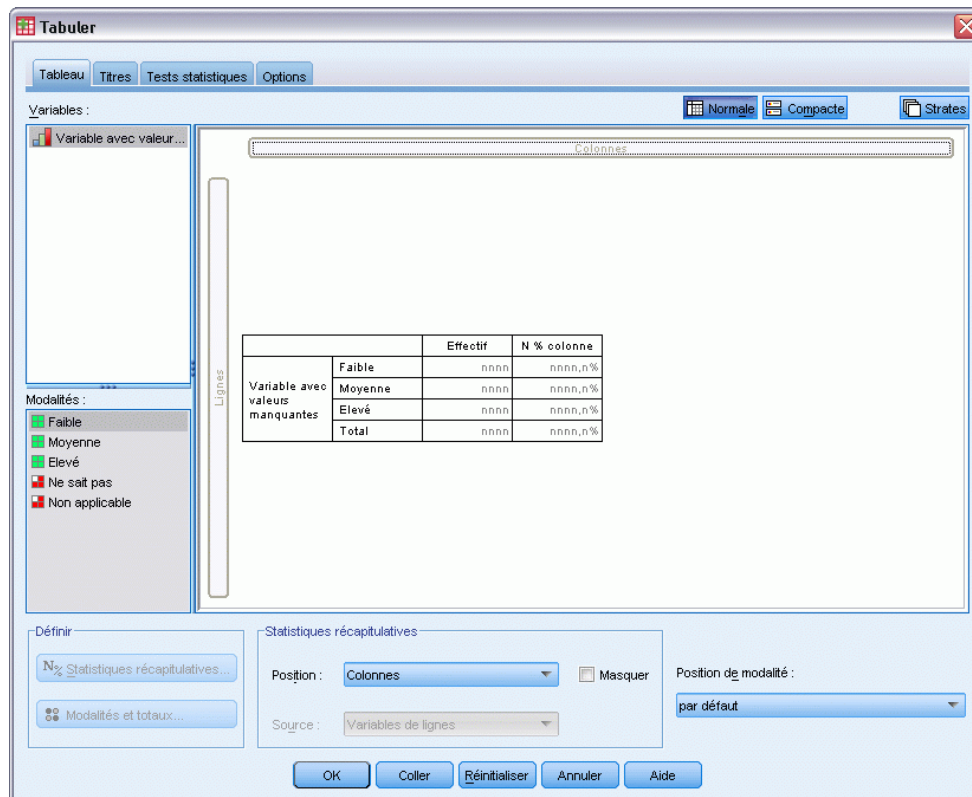
Tableaux dépourvus de valeurs manquantes

Par défaut, les modalités manquantes spécifiées par l'utilisateur n'apparaissent pas dans les tableaux personnalisés (et les valeurs manquantes par défaut n'apparaissent jamais).

- ▶ A partir du menu, sélectionnez :
Analyse > Tableaux > Tableaux personnalisés...
- ▶ Dans le générateur de tableaux, faites glisser la variable *Variable avec valeurs manquantes* (la seule variable du fichier) de la liste des variables vers la zone Lignes du panneau de canevas.
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur la variable et sélectionnez Modalités et totaux dans le menu contextuel.
- ▶ Dans la boîte de dialogue Modalités et totaux, cliquez sur Total, puis sur Appliquer.

- ▶ Dans le panneau de canevas, cliquez à nouveau avec le bouton droit de la souris sur *Variable avec valeurs manquantes* dans l'aperçu du tableau et sélectionnez Statistiques récapitulatives dans le menu contextuel.
- ▶ Dans la boîte de dialogue Statistiques récapitulatives, sélectionnez N % colonne dans la liste Statistiques et cliquez sur la flèche pour ajouter l'élément à la liste Afficher.
- ▶ Cliquez sur Appliquer à la sélection.

Figure 12-1
Aperçu d'un tableau dépourvu de valeurs manquantes



Vous pouvez noter une légère différence entre les modalités affichées dans l'aperçu du tableau dans le panneau de canevas et celles affichées dans la liste Modalités (sous la liste des variables, dans la partie gauche du générateur de tableaux). La liste Modalités contient une modalité portant l'étiquette *Valeurs manquantes* qui n'est pas incluse dans l'aperçu du tableau, car les modalités de valeurs manquantes sont exclues par défaut. Puisque le terme " valeur " apparaît au pluriel dans l'étiquette, la variable comporte au moins deux modalités manquantes spécifiées par l'utilisateur.

- ▶ Cliquez sur OK pour créer le tableau.

Figure 12-2
Tableau dépourvu de valeurs manquantes

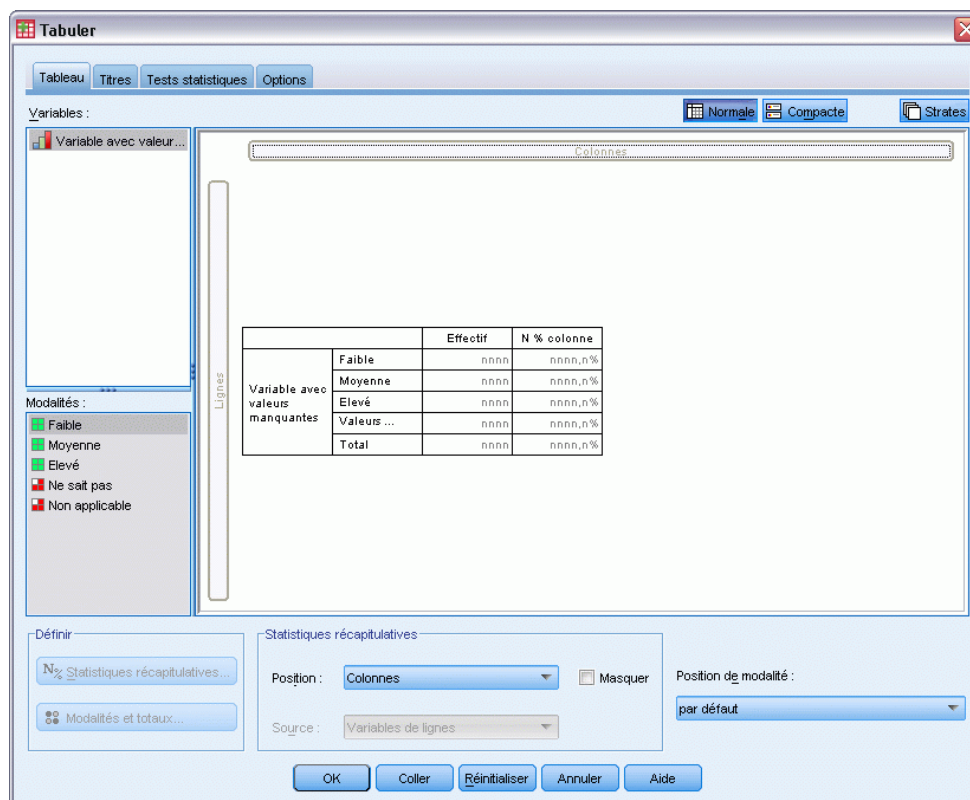
		Effectif	% colonne
Variable contenant des valeurs manquantes	Faible	2	28,6%
	Moyen	3	42,9%
	Elevé	2	28,6%
	Total	7	100,0%

Le contenu de ce tableau est parfait. Les valeurs des modalités s'additionnent pour donner des totaux et les pourcentages reflètent précisément les valeurs que vous obtiendriez si vous utilisiez le total comme base de pourcentage (par exemple, $3/7 = 0,429$ ou 42,9 %). Toutefois, le total n'est pas le nombre total d'observations dans le fichier de données, mais celui des observations comportant des valeurs **non manquantes**, ou des observations dépourvues de valeurs manquantes spécifiées par l'utilisateur ou de valeurs manquantes par défaut pour cette variable.

Inclusion de valeurs manquantes dans les tableaux

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tabuler, Tableaux personnalisés).
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Variable avec valeurs manquantes* dans l'aperçu du tableau et sélectionnez Modalités et totaux dans le menu contextuel.
- ▶ Dans la boîte de dialogue Modalités et totaux, cliquez sur Valeurs manquantes, puis sur Appliquer.

Figure 12-3
Aperçu du tableau, dans lequel une modalité de valeurs manquantes est affichée



L'aperçu du tableau comporte désormais une modalité *Valeurs manquantes*. Bien que l'aperçu du tableau ne fasse apparaître qu'une modalité pour les valeurs manquantes, toutes les modalités manquantes spécifiées par l'utilisateur seront affichées dans le tableau.

- ▶ Dans le panneau de canevas, cliquez à nouveau avec le bouton droit de la souris sur *Variable avec valeurs manquantes* dans l'aperçu du tableau et sélectionnez *Statistiques récapitulatives* dans le menu contextuel.
- ▶ Dans la boîte de dialogue *Statistiques récapitulatives*, cliquez sur *Personnaliser les statistiques récapitulatives pour les totaux et les sous-totaux*.
- ▶ Dans la liste *Statistiques relative aux autres statistiques récapitulatives*, sélectionnez *N valide*, puis cliquez sur la flèche pour ajouter cet élément à la liste *Afficher*.
- ▶ Répétez l'opération pour la statistique *N total*.
- ▶ Dans le générateur de tableaux, cliquez sur *Appliquer à la sélection*, puis sur *OK* pour créer le tableau.

Figure 12-4
Tableau avec valeurs manquantes

		Effectif	% colonne	N valide	N total
Variable contenant des valeurs manquantes	Faible	2	22,2%		
	Moyen	3	33,3%		
	Elevé	2	22,2%		
	Ne sais pas	1	11,1%		
	Sans objet	1	11,1%		
Total		9	100,0%	7	10

Les deux modalités manquantes spécifiées par l'utilisateur définies, à savoir *Ne sais pas* et *Sans objet*, apparaissent désormais dans le tableau, et le total est de 9 au lieu de 7, ce qui reflète l'ajout des deux observations comportant des valeurs manquantes spécifiées par l'utilisateur (une par modalité manquante spécifiée par l'utilisateur). Les pourcentages en colonne sont également différents car ils sont désormais basés sur le nombre de valeurs non manquantes et de valeurs manquantes spécifiées par l'utilisateur. Seules les valeurs manquantes par défaut ne sont pas incluses dans le calcul des pourcentages.

La statistique *N valide* indique le nombre total d'observations non manquantes (7), tandis que la statistique *N total* indique le nombre total d'observations, y compris les observations manquantes spécifiées par l'utilisateur et les observations manquantes par défaut. Le nombre total d'observations est 10, soit une de plus que le total des valeurs non manquantes et des valeurs manquantes spécifiées par l'utilisateur affiché dans la colonne *Effectif*. En effet, une observation comporte une valeur manquante par défaut.

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tabuler, Tableaux personnalisés).
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Variable avec valeurs manquantes* dans l'aperçu du tableau et sélectionnez Statistiques récapitulatives dans le menu contextuel.
- ▶ Sélectionnez N valide % colonne dans la liste Statistiques supérieure (pas les récapitulatifs personnalisés des totaux et sous-totaux) et cliquez sur la flèche pour ajouter cet élément à la liste Afficher.
- ▶ Répétez l'opération pour la statistique N total % colonne.
- ▶ Vous pouvez également les ajouter à la liste des statistiques récapitulatives personnalisées des totaux et sous-totaux.
- ▶ Cliquez sur Appliquer à la sélection, puis sur OK pour créer le tableau.

Figure 12-5
Tableau comportant valeurs manquantes et pourcentages totaux valides

		Effectif	% colonne	N valide % colonne	N total % colonne	N valide	N total
Variable contenant des valeurs manquantes	Faible	2	22,2%	28,6%	20,0%		
	Moyen	3	33,3%	42,9%	30,0%		
	Elevé	2	22,2%	28,6%	20,0%		
	Ne sais pas	1	11,1%	,0%	10,0%		
	Sans objet	1	11,1%	,0%	10,0%		
Total		9	100,0%	100,0%	100,0%	7	10

- La statistique *N % colonne* est le pourcentage dans chaque modalité sur la base du nombre de valeurs non manquantes et de valeurs manquantes spécifiées par l'utilisateur (dans la mesure où les valeurs manquantes spécifiées par l'utilisateur ont été explicitement incluses dans le tableau).
- La statistique *N valide % colonne* est le pourcentage dans chaque modalité uniquement sur la base des observations valides non manquantes. Ces valeurs sont les mêmes que les pourcentages en colonne du tableau d'origine qui n'incluait pas les valeurs manquantes spécifiées par l'utilisateur.
- La statistique *N total % colonne* est le pourcentage dans chaque modalité sur la base de toutes les observations, y compris les observations manquantes spécifiées par l'utilisateur et les observations manquantes par défaut. Si vous ajoutez les différents pourcentages dans cette modalité, vous obtenez seulement 90 % ; en effet, sur les 10 observations, une seule (soit 10 %) comporte la valeur manquante par défaut. Bien que cette observation soit incluse dans la base du calcul des pourcentages, aucune modalité n'est fournie dans le tableau pour les observations comportant des valeurs manquantes par défaut.

Formatage et personnalisation de tableaux

La fonction Tableaux personnalisés permet de contrôler un certain nombre de propriétés de formatage de tableau dans le cadre de la génération de tableaux, notamment :

- Format d’affichage et étiquettes des statistiques récapitulatives
- Largeur minimum et maximum des colonnes de données
- Texte ou valeur apparaissant dans les cellules vides

Ces paramètres persistent dans l’interface de génération de tableaux (jusqu’à leur modification, la réinitialisation des paramètres du générateur de tableaux ou l’ouverture d’un autre fichier de données), ce qui vous permet de créer plusieurs tableaux dotés des mêmes propriétés de formatage sans devoir les modifier manuellement après leur création. Vous pouvez aussi enregistrer ces paramètres de formatage, ainsi que tous les autres paramètres du tableau, à l’aide du bouton Coller de l’interface de génération de tableaux pour coller la syntaxe de commande dans une fenêtre de syntaxe, que vous pouvez ensuite enregistrer sous la forme d’un fichier.

De plus, vous pouvez modifier de nombreuses propriétés de formatage d’un tableau nouvellement créé à l’aide de toutes les fonctionnalités de formatage disponibles dans le Viewer pour les tableaux pivotants. Toutefois, ce chapitre est consacré au contrôle des propriétés de formatage d’un tableau avant sa création. Pour plus d’informations sur les tableaux pivotants, dans l’onglet Index du système d’aide, saisissez tableaux pivotants comme mot-clé.

Fichier de données d’exemple

Les exemples de ce chapitre utilisent le fichier de données *survey_sample.sav*. [Pour plus d’informations, reportez-vous à la section Fichiers d’exemple dans l’annexe A sur p. 197.](#)

Tous les exemples fournis ici affichent des étiquettes de variable dans des boîtes de dialogue ; ces étiquettes sont triées par ordre alphabétique. Les propriétés d’affichage de la liste des variables sont paramétrées dans l’onglet Général de la boîte de dialogue Options (menu Modifier, Options).

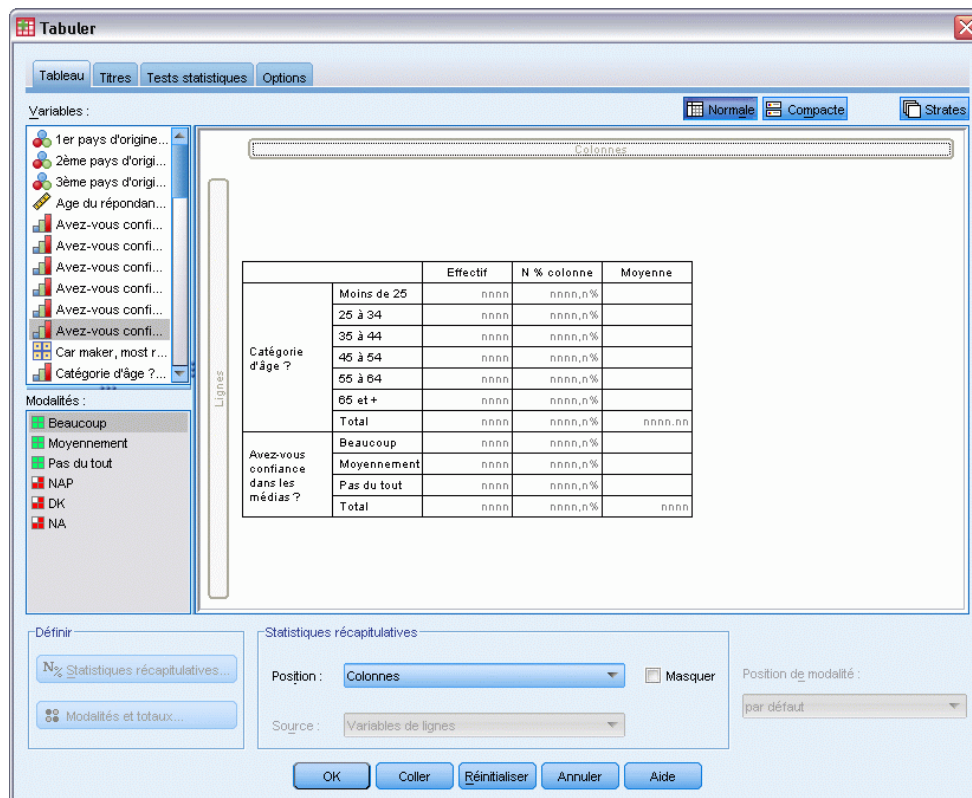
Formats d’affichage des statistiques récapitulatives

La fonction Tableaux personnalisés tente d’appliquer des formats par défaut relativement intelligents aux statistiques récapitulatives, mais vous pouvez ignorer ces valeurs par défaut.

- ▶ A partir du menu, sélectionnez :
Analyse > Tableaux > Tableaux personnalisés...
- ▶ Dans le générateur de tableaux, faites glisser la variable *Tranche d’âge* de la liste des variables vers la zone Lignes du panneau de canevas.
- ▶ Faites glisser la variable *Confiance en la télévision* sous la variable *Tranche d’âge* dans la zone Lignes, ce qui a pour effet d’empiler les deux variables dans la dimension de ligne.

- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur la variable *Tranche d'âge* dans l'aperçu du tableau, puis sélectionnez Sélectionner toutes les variables en ligne dans le menu contextuel.
- ▶ Cliquez de nouveau avec le bouton droit de la souris sur la variable *Tranche d'âge*, puis sélectionnez Modalités et totaux dans le menu contextuel.
- ▶ Dans la boîte de dialogue Modalités et totaux, sélectionnez (cochez) Total, puis cliquez sur Appliquer.
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur une variable dans l'aperçu du tableau, puis sélectionnez Statistiques récapitulatives dans le menu contextuel.
- ▶ Sélectionnez % colonne dans la liste Statistiques et cliquez sur le bouton fléché pour ajouter cet élément à la liste Afficher.
- ▶ Sélectionnez (cochez) Personnaliser les statistiques récapitulatives pour les totaux et les sous-totaux.
- ▶ Dans la liste Statistiques des autres statistiques récapitulatives, sélectionnez N % colonne et cliquez sur la flèche pour ajouter cet élément à la liste Afficher.
- ▶ Répétez l'opération pour la moyenne.
- ▶ Cliquez ensuite sur Appliquer à tout.

Figure 13-1
Formats d'affichage par défaut dans l'aperçu du tableau



Dans l'aperçu du tableau, les valeurs de substitution reflètent le format par défaut de chaque statistique récapitulative.

- Pour les effectifs, le format d'affichage par défaut est nnnn, c'est-à-dire des valeurs entières sans décimales.
- Pour les pourcentages, le format d'affichage par défaut est nnn.n%, c'est-à-dire des nombres avec une décimale, suivis du signe %.
- Pour la moyenne, le format d'affichage par défaut est *différent* pour les deux variables.

Pour les statistiques récapitulatives qui ne sont ni un effectif (notamment N valide et N total), ni un pourcentage, le format d'affichage par défaut est le format défini pour la variable dans l'éditeur de données. Si vous observez les variables dans l'affichage des variables de l'éditeur de données, vous pouvez constater que la variable *Tranche d'âge* (*agecat*) est définie comme ayant deux décimales, tandis que la variable *Confiance en la télévision* (variable *contv*) est définie comme n'ayant aucune décimale.

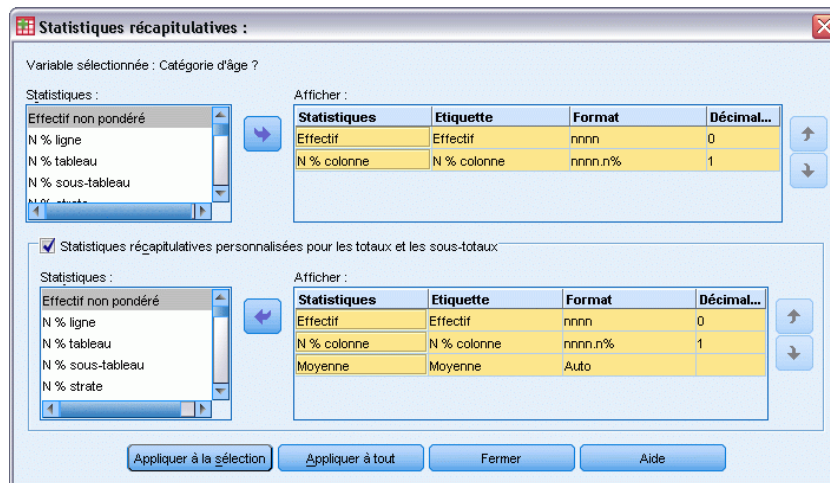
Figure 13-2
Affichage des variables dans l'éditeur de données

	Nom	Entrez	Largeur	Décimales	Etiquette
36	contv	Numérique	1	0	Avez-vous confiance dans les ...
37	agecat	Numérique	8	2	Catégorie d'âge ?
38	childcat	Numérique	8	2	Nombre d'enfants (catégorie)
39	news1	Numérique	6	0	Obtention des informations à p...
40	news2	Numérique	6	0	Obtention des informations à p...
41	news3	Numérique	6	0	Obtention des informations à p...
42	news4	Numérique	6	0	Obtention des informations à p...
43	news5	Numérique	6	0	Obtention des informations à p...
44	car1	Numérique	8	0	Orgine du véhicule 1
45	car2	Numérique	8	0	Orgine du véhicule 2

Il s'agit de l'un des cas dans lesquels le format par défaut ne vous convient probablement pas ; il serait sûrement préférable que les deux valeurs de moyenne comportent le même nombre de décimales.

- Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur une variable dans l'aperçu du tableau, puis sélectionnez Statistiques récapitulatives dans le menu contextuel.

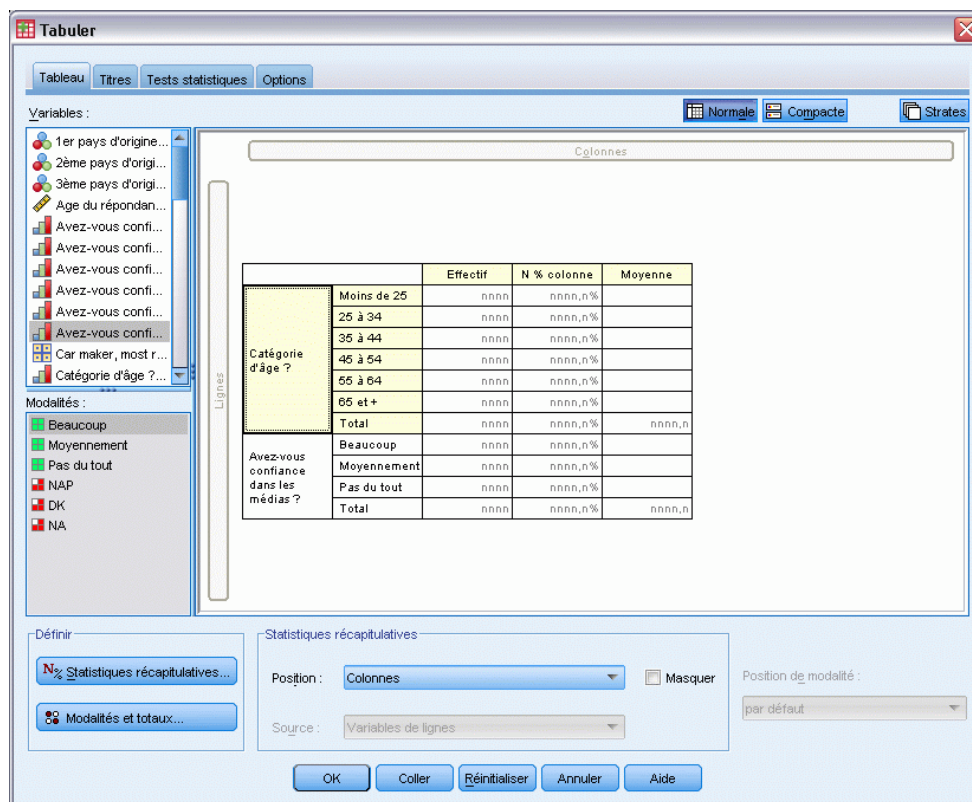
Figure 13-3
Boîte de dialogue Statistiques récapitulatives



Pour la moyenne, la cellule Format de la liste Afficher indique le format *Auto*, ce qui signifie que le format d'affichage défini pour la variable est utilisé, tandis que la cellule Décimales est désactivée. Pour spécifier le nombre de décimales, vous devez d'abord sélectionner un format différent.

- ▶ Dans la liste Afficher des statistiques récapitulatives personnalisées, cliquez sur la cellule Format de la moyenne et sélectionnez nnnn dans la liste déroulante des formats.
- ▶ Dans la cellule Décimales, saisissez la valeur 1.
- ▶ Cliquez ensuite sur Appliquer à tout pour appliquer ce paramètre aux deux variables.

Figure 13-4
Aperçu du tableau avec formats d'affichage de statistiques récapitulatives définis par l'utilisateur



L'aperçu du tableau indique désormais que les deux valeurs de moyenne apparaîtront avec une décimale. (Vous pourriez continuer et créer ce tableau maintenant, mais la valeur "moyenne" de la variable *Tranche d'âge* peut vous sembler un peu difficile à interpréter car les codes numériques réels de cette variable sont uniquement compris entre 1 et 6.)

Affichage des étiquettes pour les statistiques récapitulatives

Outre les formats d'affichage des statistiques récapitulatives, vous pouvez contrôler les étiquettes descriptives de chacune d'elles.

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tabuler, Tableaux personnalisés).
- ▶ Cliquez sur Restaurer pour effacer les paramètres précédents dans le générateur de tableaux.
- ▶ Dans le générateur de tableaux, faites glisser la variable *Tranche d'âge* de la liste des variables vers la zone Lignes du panneau de canevas.
- ▶ Faites glisser la variable *Mode de paiement la semaine dernière* de la liste des variables vers la zone Colonnes du panneau de canevas.
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur *Tranche d'âge* dans l'aperçu du tableau et sélectionnez Statistiques récapitulatives dans le menu contextuel.

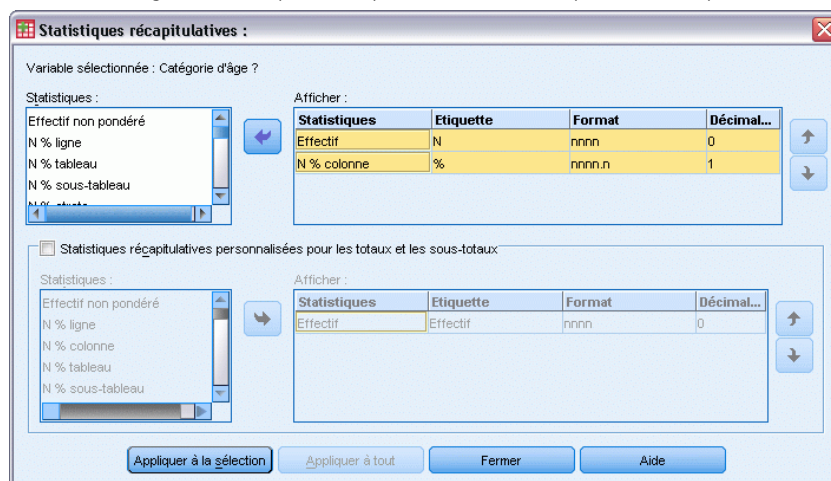
- ▶ Sélectionnez % colonne dans la liste Statistiques et cliquez sur le bouton fléché pour ajouter cet élément à la liste Afficher.
- ▶ Double-cliquez n'importe où sur le mot *Colonne* dans la cellule Etiquette de la liste Afficher pour modifier le contenu de la cellule. Supprimez le mot *Colonne* de l'étiquette afin que cette dernière se limite à %.
- ▶ Modifiez la cellule Etiquette de la statistique d'*effectif* de la même façon, de sorte que son étiquette se limite à *N*.

Nous allons modifier le format de la statistique N % colonne afin de supprimer le signe % superflu (puisque l'étiquette de la colonne indique que celle-ci contient des pourcentages).

- ▶ Cliquez sur la cellule Format de la statistique *N % colonne* et sélectionnez nnnn.n dans la liste déroulante des formats.

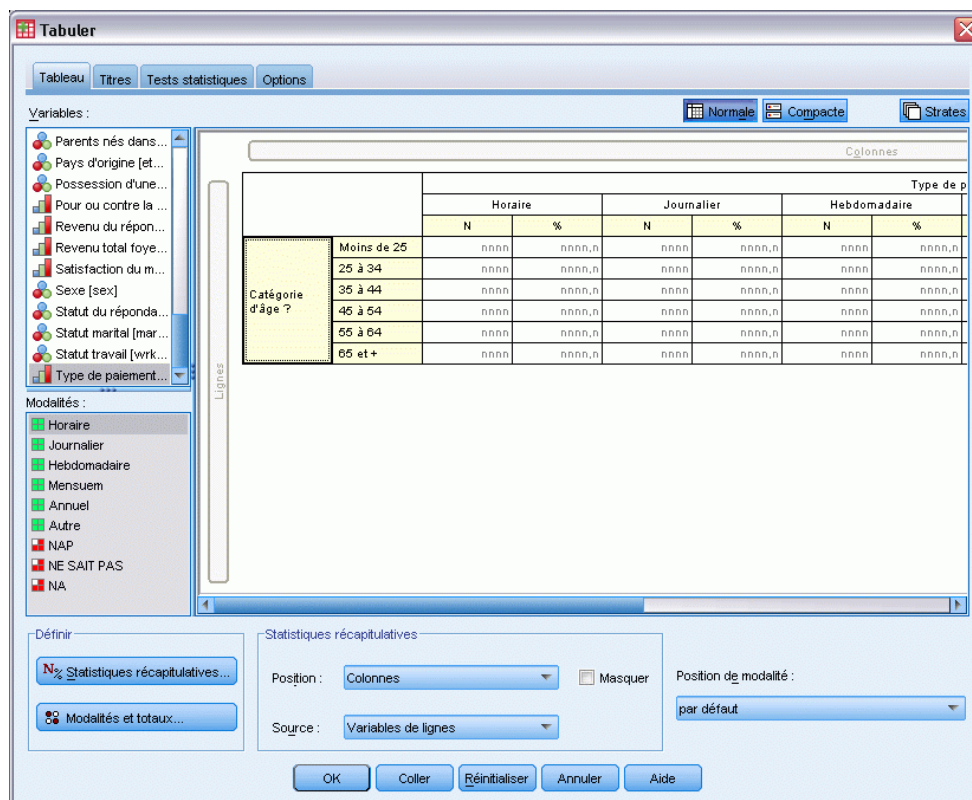
Figure 13-5

Boîte de dialogue Statistiques récapitulatives, dans laquelle les étiquettes et les formats ont été modifiés



- ▶ Cliquez ensuite sur Appliquer à la sélection.

Figure 13-6
Aperçu du tableau dans lequel les étiquettes des statistiques récapitulatives ont été modifiées



L'aperçu du tableau affiche les étiquettes et le format d'affichage modifiés.

- Cliquez sur OK pour créer le tableau.

Figure 13-7
Tableau dans lequel les étiquettes des statistiques récapitulatives ont été modifiées

		Quel a été votre mode de rémunération la semaine dernière ?											
		Paiement horaire		Paiement journalier		Paiement hebdomadaire		Paiement mensuel		Paiement annuel		Autre type de paiement	
		N	%	N	%	N	%	N	%	N	%	N	%
Catégories d'âges	Moins de 25	91	14,0	0	,0	12	9,7	3	2,0	7	3,1	14	7,7
	25 à 34	175	26,9	5	29,4	33	26,6	37	24,8	63	28,0	31	17,1
	35 à 44	185	28,5	5	29,4	42	33,9	45	30,2	66	29,3	61	33,7
	45 à 54	124	19,1	5	29,4	25	20,2	38	25,5	58	25,8	41	22,7
	55 à 64	52	8,0	0	,0	10	8,1	23	15,4	29	12,9	19	10,5
65 ou plus	23	3,5	2	11,8	2	1,6	3	2,0	2	,9	15	8,3	

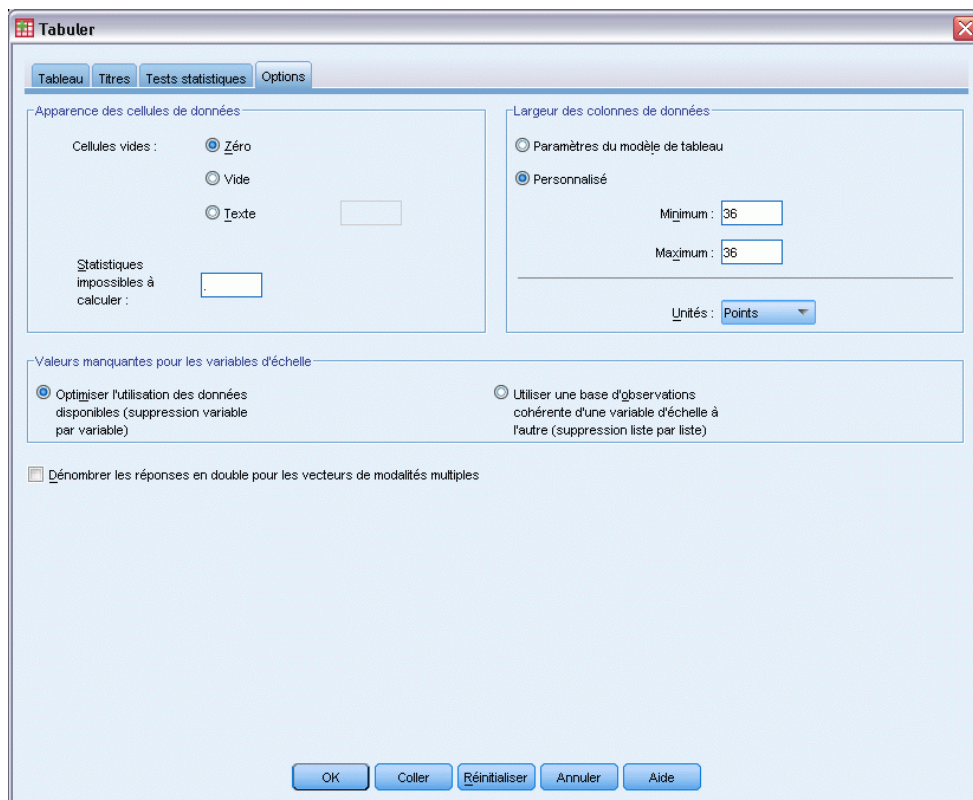
Largeur des colonnes

Vous avez peut-être remarqué que le tableau de l'exemple ci-dessus est assez volumineux. Une solution à ce problème consiste tout simplement à permuter les variables en ligne et les variables en colonne. Une autre solution consiste à réduire la largeur des colonnes, car celles-ci semblent beaucoup trop larges. (De fait, nous avons raccourci les étiquettes des statistiques récapitulatives afin de pouvoir réduire la largeur des colonnes.)

- Ouvrez le générateur de tableaux (menu Analyse, Tabuler, Tableaux personnalisés).

- ▶ Cliquez sur l'onglet Options.
- ▶ Dans le groupe Largeur des colonnes de données, sélectionnez Personnalisé.
- ▶ Dans la zone Maximum, saisissez 36. (Vérifiez que le paramètre Unités a pour valeur Points.)

Figure 13-8
Boîte de dialogue Tabuler, onglet Options



- ▶ Cliquez sur OK pour créer le tableau.

Figure 13-9
Tableau dans lequel la largeur des colonnes a été réduite

		Quel a été votre mode de rémunération la semaine dernière ?											
		Paiement horaire		Paiement journalier		Paiement hebdomadaire		Paiement mensuel		Paiement annuel		Autre type de paiement	
		N	%	N	%	N	%	N	%	N	%	N	%
Catégories d'âges	Moins de 25	91	14,0	0	,0	12	9,7	3	2,0	7	3,1	14	7,7
	25 à 34	175	26,9	5	29,4	33	26,6	37	24,8	63	28,0	31	17,1
	35 à 44	185	28,5	5	29,4	42	33,9	45	30,2	66	29,3	61	33,7
	45 à 54	124	19,1	5	29,4	25	20,2	38	25,5	58	25,8	41	22,7
	55 à 64	52	8,0	0	,0	10	8,1	23	15,4	29	12,9	19	10,5
	65 ou plus	23	3,5	2	11,8	2	1,6	3	2,0	2	,9	15	8,3

Le tableau est désormais beaucoup plus compact.

Affichage d'une valeur dans les cellules vides

Par défaut, un 0 apparaît dans les cellules vides (cellules dépourvues d'observations). Vous pouvez choisir de ne rien afficher dans ces cellules (les laisser vierges) ou indiquer la chaîne de caractères à afficher.

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tabuler, Tableaux personnalisés).
- ▶ Cliquez sur l'onglet Options.
- ▶ Dans le groupe Apparence des cellules de données, pour Cellules vides, sélectionnez Texte et tapez Aucune.
- ▶ Cliquez sur OK pour créer le tableau.

Figure 13-10

Tableau dont les cellules vides contiennent la mention "Aucune"

		Quel a été votre mode de rémunération la semaine dernière ?											
		Paiement horaire		Paiement journalier		Paiement hebdomadaire		Paiement mensuel		Paiement annuel		Autre type de paiement	
		N	%	N	%	N	%	N	%	N	%	N	%
Catégories d'âges	Moins de 25	91	14,0	aucun	aucun	12	9,7	3	2,0	7	3,1	14	7,7
	25 à 34	175	26,9	5	29,4	33	26,6	37	24,8	63	28,0	31	17,1
	35 à 44	185	28,5	5	29,4	42	33,9	45	30,2	66	29,3	61	33,7
	45 à 54	124	19,1	5	29,4	25	20,2	38	25,5	58	25,8	41	22,7
	55 à 64	52	8,0	aucun	aucun	10	8,1	23	15,4	29	12,9	19	10,5
	65 ou plus	23	3,5	2	11,8	2	1,6	3	2,0	2	,9	15	8,3

Les quatre cellules vides du tableau contiennent désormais le texte *Aucune* à la place de la valeur 0.

Affichage d'une valeur pour les statistiques manquantes

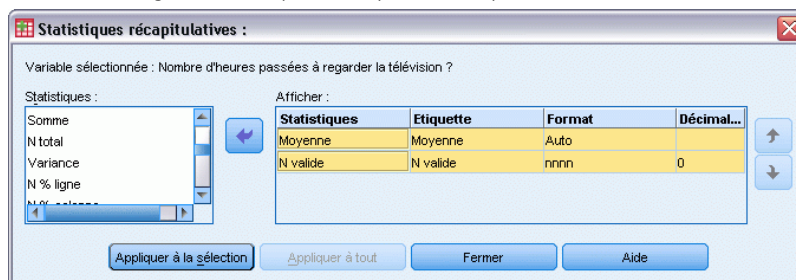
Si une statistique ne peut pas être calculée, la valeur apparaissant par défaut est un point (.), symbole utilisé pour indiquer la valeur manquante par défaut. Ceci est différent d'une cellule "vide" ; par conséquent, la valeur apparaissant pour les statistiques manquantes est contrôlée indépendamment de celle des cellules dépourvues d'observations.

- ▶ Ouvrez le générateur de tableaux (menu Analyse, Tabuler, Tableaux personnalisés).
- ▶ Faites glisser la variable *Heures par jour devant la télé* de la liste de variables vers le sommet de la zone Colonnes dans le canevas, au-dessus de la variable *Mode de paiement la semaine dernière*.

La variable *Heures par jour devant la télé* étant une variable d'échelle, elle devient automatiquement la variable source des statistiques et la moyenne devient la statistique récapitulative.

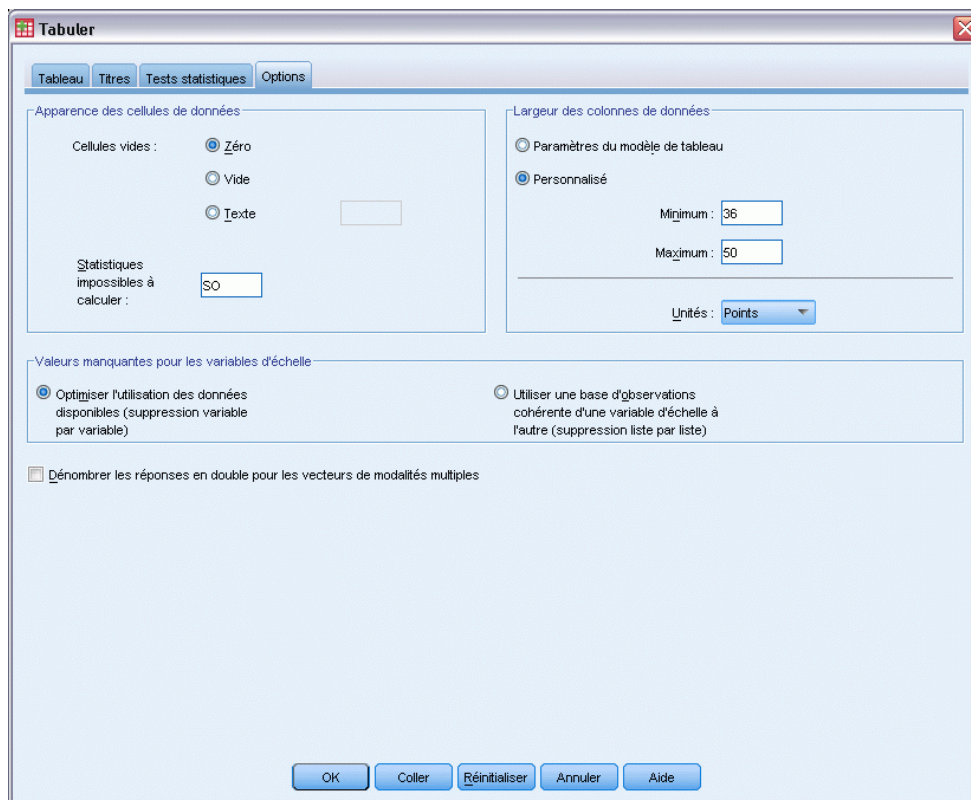
- ▶ Dans le panneau de canevas, cliquez avec le bouton droit de la souris sur la variable *Heures par jour devant la télé* dans l'aperçu du tableau et sélectionnez Statistiques récapitulatives dans le menu contextuel.
- ▶ Dans la liste Statistiques, sélectionnez N valide, puis cliquez sur le bouton fléché pour ajouter cet élément à la liste Afficher.

Figure 13-11
Boîte de dialogue Statistiques récapitulatives pour les variables d'échelle



- ▶ Cliquez sur Appliquer à la sélection.
- ▶ Cliquez sur l'onglet Options.
- ▶ Dans le champ texte correspondant à l'élément Statistiques impossibles à calculer, saisissez SO.

Figure 13-12
Modification de la valeur apparaissant pour les statistiques impossibles à calculer



- ▶ Cliquez sur OK pour créer le tableau.

Figure 13-13

Tableau dans lequel la mention " SO " apparaît pour les statistiques manquantes

		Nombre d'heures passées à regarder la télévision											
		Type de paiement											
		Horaire		Journalier		Hebdomadaire		Mensuem		Annuel		Autre	
		Moyen ne	N valide	Moyen ne	N valide	Moyen ne	N valide	Moyen ne	N valide	Moyen ne	N valide	Moyen ne	N valide
Catégorie	Moins de 25	3	71	NA	Aucun	3	10	2	3	2	6	2	8
d'âge	25 à 34	3	134	5	2	2	30	2	29	2	52	2	22
	35 à 44	3	136	2	5	3	30	2	34	2	47	3	46
	45 à 54	2	90	2	4	2	22	2	36	2	45	2	34
	55 à 64	3	40	NA	Aucun	3	7	2	15	2	23	3	15
	65 et +	3	18	2	2	1	1	NA	0	1	2	3	11

Le texte *SO* apparaît pour la moyenne dans trois cellules du tableau. Dans chaque cas, la valeur *N valide* correspondante explique pourquoi : Aucune observation ne permet de calculer la moyenne.

Toutefois, vous pouvez remarquer une légère différence ; l'une des trois valeurs *N valides* est exprimée par 0, et non par l'étiquette *Aucune* supposée apparaître dans les cellules dépourvues d'observations. En effet, malgré l'absence d'observations permettant de calculer la moyenne, la modalité n'est pas vraiment vide. Si vous revenez au tableau d'origine comportant uniquement les deux variables qualitatives, vous pouvez constater que cette modalité croisée comporte en réalité trois observations. Toutefois, aucune des observations n'est valide, car toutes trois comportent des valeurs manquantes pour la variable d'échelle *Heures par jour devant la télé.*

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'alfatoxine en parties par milliard (PPB).
- **anorectic.sav.** En cherchant à développer une symptomatologie standardisée du comportement anorexique/boulimique, des chercheurs 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.

- **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, 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, 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 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, 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'applicateur, 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. 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'entreprend 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'entreprend 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*.
- **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. 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 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 a présenté un tableau pour illustrer les groupes sociaux possibles. Guttman 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 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*, *génér(ation)* 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 , , 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.
- **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.csplan*. 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. À 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..
- **shampoo_ph.sav.** Ce fichier de données d'hypothèse concerne le processus de contrôle qualité observé dans une usine de produits capillaires. À 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 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.
- **stocks.sav** Ce fichier de données hypothétiques contient le cours et le volume des actions pour un an.
- **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 .
- **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 .
- **verd1985.sav.** Ce fichier de données concerne une enquête . 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 . 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.

- **worldsales.sav** Ce fichier de données hypothétiques contient les revenus des ventes par continent et par produit.

Remarques

Ces informations ont été développées pour les produits et services offerts dans le monde.

Il est possible qu'IBM n'offre pas dans les autres pays les produits, services et fonctionnalités décrits dans ce document. Contactez votre représentant local IBM pour obtenir des informations sur les produits et services actuellement disponibles dans votre région. Toute référence à un produit, programme ou service IBM n'implique pas que les seuls les produits, programmes ou services IBM peuvent être utilisés. Tout produit, programme ou service de fonctionnalité équivalente qui ne viole pas la propriété intellectuelle IBM peut être utilisé à la place. Cependant l'utilisateur doit évaluer et vérifier l'utilisation d'un produit, programme ou service non IBM.

IBM peut posséder des brevets ou des applications de brevet en attente qui couvrent les sujets décrits dans ce document. L'octroi de ce document n'équivaut aucunement à celui d'une licence pour ces brevets. Vous pouvez envoyer par écrit des questions concernant la licence à :

IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785, États-Unis

Pour obtenir des informations de licence concernant la configuration de caractères codés sur deux octets (DBCS), veuillez contacter dans votre pays le département chargé de la propriété intellectuelle chez IBM ou envoyez vos commentaires par écrit à :

Intellectual Property Licensing, Legal and Intellectual Property Law, IBM Japan Ltd., 1623-14, Shimotsuruma, Yamato-shi, Kanagawa 242-8502 Japon.

Le paragraphe suivant ne s'applique pas au Royaume-Uni ni à aucun pays dans lequel ces dispositions sont contraires au droit local : INTERNATIONAL BUSINESS MACHINES FOURNIT CETTE PUBLICATION « EN L'ÉTAT » SANS GARANTIE D'AUCUNE SORTE, IMPLICITE OU EXPLICITE, Y COMPRIS, MAIS SANS ETRE LIMITE AUX GARANTIES IMPLICITES DE NON VIOLATION, DE QUALITE MARCHANDE OU D'ADAPTATION POUR UN USAGE PARTICULIER. Certains états n'autorisent pas l'exclusion de garanties explicites ou implicites lors de certaines transactions, par conséquent, il est possible que cet énoncé ne vous concerne pas.

Ces informations peuvent contenir des erreurs techniques ou des erreurs typographiques. Ces informations sont modifiées de temps en temps ; ces modifications seront intégrées aux nouvelles versions de la publication. IBM peut apporter des améliorations et/ou modifications des produits et/ou des programmes décrits dans cette publications à tout moment sans avertissement préalable.

Toute référence dans ces informations à des sites Web autres qu'IBM est fournie dans un but pratique uniquement et ne sert en aucun cas de recommandation pour ces sites Web. Le matériel contenu sur ces sites Web ne fait pas partie du matériel de ce produit IBM et l'utilisation de ces sites Web se fait à vos propres risques.

IBM peut utiliser ou distribuer les informations que vous lui fournissez, de la façon dont il le souhaite, sans encourir aucune obligation envers vous.

Les personnes disposant d'une licence pour ce programme et qui souhaitent obtenir des informations sur celui-ci pour activer : (i) l'échange d'informations entre des programmes créés de manière indépendante et d'autres programmes (notamment celui-ci) et (ii) l'utilisation mutuelle des informations qui ont été échangées, doivent contacter :

IBM Software Group, Attention: Licensing, 233 S. Wacker Dr., Chicago, IL 60606, États-Unis.

Ces informations peuvent être disponibles, soumises à des conditions générales, et dans certains cas payantes.

Le programme sous licence décrit dans ce document et toute la documentation sous licence disponible pour ce programme sont fournis par IBM en conformité avec les conditions de l'accord du client IBM, avec l'accord de licence du programme international IBM et avec tout accord équivalent entre nous.

Les informations concernant les produits autres qu'IBM ont été obtenues auprès des fabricants de ces produits, leurs annonces publiques ou d'autres sources publiques disponibles. IBM n'a pas testé ces produits et ne peut confirmer l'exactitude de leurs performances, leur compatibilité ou toute autre fonctionnalité associée à des produits autres qu'IBM. Les questions sur les capacités de produits autres qu'IBM doivent être adressées aux fabricants de ces produits.

Ces informations contiennent des exemples de données et de rapports utilisés au cours d'opérations quotidiennes standard. Pour les illustrer le mieux possible, ces exemples contiennent des noms d'individus, d'entreprises, de marques et de produits. Tous ces noms sont fictifs et toute ressemblance avec des noms et des adresses utilisés par une entreprise réelle ne serait que pure coïncidence.

Si vous consultez la version papier de ces informations, il est possible que certaines photographies et illustrations en couleurs n'apparaissent pas.

Marques commerciales

IBM, le logo IBM, ibm.com et SPSS sont des marques commerciales d'IBM Corporation, déposées dans de nombreuses juridictions du monde entier. Une liste à jour des marques IBM est disponible sur Internet à l'adresse <http://www.ibm.com/legal/copytrade.shtml>.

Adobe, le logo Adobe, PostScript et le logo PostScript sont des marques déposées ou des marques commerciales de Adobe Systems Incorporated aux États-Unis et/ou dans d'autres pays.

Intel, le logo Intel, Intel Inside, le logo Intel Inside, Intel Centrino, le logo Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium, et Pentium sont des marques commerciales ou des marques déposées de Intel Corporation ou de ses filiales aux États-Unis et dans d'autres pays.

Java et toutes les marques et logos Java sont des marques commerciales de Sun Microsystems, Inc. aux États-Unis et/ou dans d'autres pays.

Linux est une marque déposée de Linus Torvalds aux États-Unis et/ou dans d'autres pays.

Microsoft, Windows, Windows NT et le logo Windows sont des marques commerciales de Microsoft Corporation aux États-Unis et/ou dans d'autres pays.

UNIX est une marque déposée de The Open Group aux États-Unis et dans d'autres pays.

Ce produit utilise WinWrap Basic, Copyright 1993-2007, Polar Engineering and Consulting, <http://www.winwrap.com/>.

Les autres noms de produits et de services peuvent être des marques d'IBM ou d'autres sociétés.

Les captures d'écran des produits Adobe sont reproduites avec l'autorisation de Adobe Systems Incorporated.

Les captures d'écran des produits Microsoft sont reproduites avec l'autorisation de Microsoft Corporation.



Index

- Affichage des valeurs de modalité, 131
- Autres statistiques récapitulatives en total, 128

- catégories calculées
 - à partir des sous-totaux, 103
 - Formats d'affichage, 47
 - masquer des catégories dans une expression, 101
 - Tableaux personnalisés, 46, 99
- catégories post-calculées
 - Tableaux personnalisés, 46, 99
- cellules vides
 - Valeur affichée dans les tableaux personnalisés, 49, 194
- Comperimeter tables, 48, 110
- Contrôle du nombre de décimales affichées, 63

- date
 - Insertion de la date dans les tableaux personnalisés, 51
- Décimales
 - Contrôle du nombre de décimales affichées dans les tableaux personnalisés, 34, 63, 186
- Différentes statistiques récapitulatives pour différentes variables
 - Tableaux empilés, 138

- écart-type
 - Tableaux personnalisés, 38
- Effectif .
 - N valide, 136
- Emboîtement de variables
 - Tableaux personnalisés, 73, 78
 - Variables d'échelle, 143
- Empilement des variables
 - Différentes statistiques récapitulatives pour différentes variables, 138
 - Empilement de variables de strate, 83
 - Plusieurs variables source de statistiques récapitulatives, 125
 - Tableaux personnalisés, 70–71
 - Variables d'échelle, 134
- Étiquettes
 - Modification du texte des étiquettes des statistiques récapitulatives, 190
- étiquettes de coins
 - Tableaux personnalisés, 51
- étiquettes de variable
 - Suppression de l'affichage dans les tableaux personnalisés, 33
- exclusion de modalités
 - Tableaux personnalisés, 41
- Exclusion de modalités
 - Tableaux personnalisés, 65

- fichiers d'exemple
 - emplacement, 197

- Formats d'affichage, 63
 - Statistiques récapitulatives dans des tableaux personnalisés, 40, 186
- fusion de modalités
 - Tableaux personnalisés, 95

- heure
 - Insertion de l'heure dans les tableaux personnalisés, 51

- impression de tableaux avec strates, 85

- Khi-deux
 - Tableaux personnalisés, 145

- largeur de colonne
 - Contrôle dans les tableaux personnalisés, 49, 192
- légendes
 - Tableaux personnalisés, 51

- marques commerciales, 209
- masquage des étiquettes de statistiques dans des tableaux personnalisés, 57
- maximum
 - Tableaux personnalisés, 38
- médiane, 135
 - Tableaux personnalisés, 38
- mentions légales, 208
- minimum
 - Tableaux personnalisés, 38
- mode
 - Tableaux personnalisés, 38
- moyenne, 134
- moyenne .
 - Tableaux personnalisés, 38
- Moyenne - tableaux de fréquences, 128
- Moyenne - tableaux d'effectif, 39

- N total, 182
- N valide, 136, 182
 - Tableaux personnalisés, 38
- niveau de mesure
 - Modification dans des tableaux personnalisés, 26

- Omission de modalités
 - Tableaux personnalisés, 65

- plage
 - Tableaux personnalisés, 38
- pourcentages
 - dans les tableaux personnalisés, 36–37, 58, 62
 - Valeurs manquantes, 182

- vecteurs multiréponses, 37
- Récapitulatifs groupés
 - Variables d'échelle, 140
- réorganisation des modalités
 - Tableaux personnalisés, 41
- somme
 - Tableaux personnalisés, 38
- sous-totaux
 - Tableaux personnalisés, 41
- Sous-totaux, 93
 - masquage des modalités de sous-total, 95
 - Tableaux personnalisés, 86
- statistiques
 - Autres statistiques récapitulatives en total, 128
 - Statistiques récapitulatives, 119
 - Tableaux empilés, 125
- statistiques de test
 - Tableaux personnalisés, 145
- Statistiques des moyennes de colonne
 - Tableaux personnalisés, 150
- Statistiques des proportions de colonne
 - Tableaux personnalisés, 157
- Statistiques récapitulatives, 119
 - Autres statistiques récapitulatives en total, 128
 - Différents récapitulatifs pour différentes variables dans des tableaux empilés, 138
 - Dimension source, 120
 - Format d'affichage, 186
 - Modification du texte d'une étiquette, 190
 - Tableaux empilés, 125
 - Variable source, 120
- suppression de modalités
 - Tableaux personnalisés, 41
- Suppression de modalités
 - Tableaux personnalisés, 65
- Tableau croisé
 - Tableaux personnalisés, 61
- Tableau d'effectifs
 - Tableaux personnalisés, 48
- Tabuler, 110
- tableaux
 - Tableaux personnalisés, 26
- Tableaux personnalisés
 - Affichage et masquage des noms et étiquettes de variable, 33
 - Autres totaux, 39
 - catégories calculées, 41, 46, 99
 - catégories post-calculées, 46, 99
 - Cellules vides, 49
 - Comperimeter tables, 48, 110
 - Contrôle du nombre de décimales affichées, 34
 - Dimension source des statistiques, 62
 - Emboîtement de variables, 73, 78
 - Emboîtement de variables de strate, 84
 - Empilement des variables, 70–71
 - Étiquettes de coins, 51
 - Étiquettes de valeurs pour variables qualitatives, 26
 - exclusion de modalités, 41
 - Exclusion de modalités, 65
 - Exclusion des valeurs manquantes pour les récapitulatifs d'échelle, 49
 - Formats d'affichage, 34
 - Formats d'affichage des statistiques récapitulatives, 40
 - fusion de modalités, 95
 - Impression de tableaux en strates, 85
 - Largeur de colonne, 49
 - Légendes, 51
 - Masquage des étiquettes de statistiques, 57
 - masquage des modalités de sous-total, 95
 - Mode Compact, 78
 - Modification de la dimension des statistiques récapitulatives, 39
 - Modification des étiquettes pour les statistiques récapitulatives, 58
 - Modification du niveau de mesure, 26
 - Moyenne - tableaux d'effectif, 39
 - Permutation des variables en ligne et en colonne, 79
 - Pour construire un tableau, 29
 - pourcentages, 36–37, 58, 62
 - Pourcentages en ligne et pourcentages en colonne, 58
 - Pourcentages pour vecteurs multiréponses, 37
 - Réorganisation des modalités, 41
 - Sous-totaux, 41, 86
 - statistiques de test, 52, 145
 - Statistiques récapitulatives, 36–38
 - Tableau croisé, 61
 - Tableau d'effectifs, 48, 110
 - Tableaux de variables avec modalités partagées, 48, 110
 - Tableaux simples pour variables qualitatives, 57
 - test de signification et multiréponses, 164
 - titres, 51
 - Totaux, 41, 60, 86
 - Totaux dans des tableaux dans lesquels des modalités ont été exclues, 65
 - Totaux marginaux, 64
 - traitement d'un fichier scindé, 31
 - Tri de modalités, 65
 - Variables de strate, 81, 83–84
 - Variables d'échelle, 26
 - Variables qualitatives, 26
 - Vecteurs de modalités multiples, 49
 - vecteurs multiréponses, 26, 165
- tests de signification
 - Tableaux personnalisés, 52
- Tests de signification
 - vecteurs multiréponses, 175, 178
- tests statistiques
 - Tableaux personnalisés, 52
- titres
 - Tableaux personnalisés, 51

- totaux
 - Tableaux personnalisés, 41
- Totaux
 - Modalités exclues, 87
 - Position d'affichage, 88
 - Strates, 91
 - Tableaux emboîtés, 89
 - Tableaux personnalisés, 60, 86
 - Totaux de groupe, 89
 - Totaux marginaux pour tableaux personnalisés, 64
- Totaux de groupe, 89
- Totaux de sous-groupe, 89
- traitement d'un fichier scindé
 - Tableaux personnalisés, 31
- Tri de modalités
 - Tableaux personnalisés, 65

- Valeurs
 - Affichage des étiquettes et des valeurs de modalité, 131
- Valeurs et étiquettes de valeur, 131
- Valeurs manquantes, 136, 180
 - Effet sur le calcul des pourcentages, 182
 - Inclusion dans les tableaux personnalisés, 182
- Valeurs manquantes par défaut, 180
- Valeurs manquantes spécifiées par l'utilisateur, 180
- Variable source des statistiques récapitulatives
 - Variables d'échelle, 143
- Variables de strate
 - Emboîtement de variables de strate, 84
 - Empilement de variables de strate, 83
 - Impression de tableaux en strates, 85
 - Tableaux personnalisés, 81, 83–84
- Variables d'échelle
 - Emboîtement, 143
 - Empilement, 134
 - Récapitulatifs groupés, 140
 - Récapitulatifs groupés par variables qualitatives de ligne et de colonne, 141
 - Statistiques récapitulatives, 134
 - Statistiques récapitulatives multiples, 135
- Variance
 - Tableaux personnalisés, 38
- vecteurs multiréponses, 165
 - pourcentages, 37
 - Réponses en double dans les vecteurs de modalités multiples, 49
 - test de signification, 164, 175, 178