

CoTab v2.15

Logiciel d'édition de tableaux statistiques

Sommaire

1. Présentation générale.....	6
2. Les études.....	7
2.1 Démarrage.....	7
2.2 Nouvelle étude.....	9
2.2.1 Définitions.....	9
2.2.2 Mise en œuvre : déclaration de la structure de l'étude.....	10
2.3 Importation d'études.....	14
2.3.1 Importation Conversoft.....	14
2.3.2 Importation Cosi.....	15
2.3.3 Importation Quancept.....	16
2.3.4 Importation SPSS.....	18
2.3.5 Importation TripleS.....	19
2.3.6 Importation Excel (xls ou xlsx).....	20
2.4 Ouverture d'une étude.....	21
2.5 Ecran « Propriétés ».....	21
2.6 Suppression d'une étude.....	22
3. Les variables.....	23
3.1 Généralités.....	23
3.2 Attributs des variables.....	24
3.3 Fonctionnement général.....	25
3.4 Formats.....	30
3.4.1 Généralités.....	30
3.4.2 Création d'un format.....	31
3.4.3 Génération automatique de formats.....	33
3.4.4 Exemples de formats.....	34
3.5 Fonctions.....	35
3.5.1 Fonction Saisie.....	36
3.5.2 Fonction SaisieCsv.....	38
3.5.3 Fonction Import.....	39
3.5.4 Fonction ImportCsv.....	40
3.5.5 Arithm : Opération arithmétique.....	41
3.5.6 Logic : Opération logique.....	43
3.5.7 Recod.....	45
3.5.8 Nbrep : nombre de valeurs d'une variable multiple.....	48
3.5.9 Nbval : Nombre de valeurs d'une variable dimensionnée.....	49
3.5.10 Somval : Somme des valeurs d'une variable dimensionnée.....	50
3.5.11 Dicho1 : Passage d'une variable de type multiple à une variable dimensionnée ..	51

3.5.12	Dicho2 : Passage d'une variable dimensionnée à une variable de type multiple. .	52
3.5.13	NbNiv : Nombre de valeurs d'une variable de niveau inférieur.	53
3.5.14	SomNiv : Somme des valeurs d'une variable d'un niveau inférieur.	54
3.5.15	Const : Table de constantes.	55
3.5.16	Table : Sélection d'une valeur dans une table au moyen d'une variable-indice. ...	56
3.5.17	Concat : concaténation de variables.	57
3.5.18	Min et Max : Minimum et maximum d'une liste de variables.	59
3.5.19	Rang et RangNF : rang d'enregistrement.	61
3.5.20	GetNiv : obtention du niveau.	61
3.5.21	Conv : conversions.	62
3.5.22	Fonctions de traitement des variables temporelles.	65
3.5.23	Chaine. Concaténation de chaînes de caractères.	69
3.5.24	ExtrChn : Extraire une sous-chaîne à partir d'une variable texte.	70
3.5.25	PosChn :	71
3.5.26	TmutS : transformation d'une variable Texte en une variable Simple.	72
3.6	Tables externes.	73
3.6.1	Principe.	73
3.6.2	Lecture des variables.	74
3.7	Bases.	75
3.8	Généralités sur les exécutions.	76
3.8.1	Cas d'une étude sans niveau, avec des individus mono-enregistrements :	77
3.8.2	Cas d'une étude sans niveau, avec des individus multi-enregistrements :	77
3.8.3	Cas d'une étude à niveaux :	77
3.8.4	Exécution multi-fichiers	78
3.8.5	Tables externes.	78
3.8.6	Fichier « cache ».	78
3.9	Fusion d'études.	80
3.9.1	Principe.	80
3.9.2	Mise en œuvre.	80
3.10	Prévisualisation	81
3.11	Tris à plat.	82
4.	Les modèles.	83
4.1	Principes.	83
4.2	Gestion des modèles.	84
4.3	Contenu d'un modèle.	85
5.	Les tableaux.	86
5.1	Gestion des demandes.	86
5.2	Contenu d'une demande.	87
5.3	Contenu d'une ligne (pages, lignes ou colonnes).	89
5.3.1.	Les variables.	89
5.3.2.	Sélection/filtre de modalités.	90
5.3.3	Case « étalement ».	91
5.3.4	Modèle :	91
5.3.5	Pourcentages (%) :	93
5.3.6	Changement de page (break) :	93
5.4	Exécution d'une demande.	94
5.4.1	Contrôle de la demande.	94
5.4.2	Exécution.	94
5.4.3	Ventilation dans un tableau élémentaire :	95

5.5 Exemples de calculs de pourcentages.	98
5.5.1 Pourcentages horizontaux.	99
5.5.2 Pourcentages verticaux.	99
5.6 Présentation des résultats.	100
5.6.1 Mise en forme du tableau.	101
5.6.2 Mise en page du tableau.	108
5.6.3 Index.	110
5.6.4 Tests : Gamma (Goodman et Kruskal) et Chi ²	111
5.6.5 Impression du tableau.	112
5.6.6 Visualisation avant impression.	112
5.6.7 Exportation.	113
5.6.8 Demandes faites à partir d'un dossier.	114
5.7 Opérations sur chaînes.	115
5.7.1 Demandes.	115
5.7.2 Exécution.	117
6. Exportations de fichiers.	118
6.1 Généralités.	118
6.2 Gestion des demandes.	118
6.2.1 Dans le cas CoTab :	119
6.2.2 Dans le cas SPSS :	121
6.2.3 Dans le cas SPAD:	122
6.2.4 Dans le cas Cosi :	122
7. Redressement d'échantillon.	123
7.1 Principe.	123
7.2 Demande.	123
7.2.1 Onglet « options » :	124
7.2.2 Onglet Critères.	125
7.3 Exécution.	127
8. Modèle prédictif.	129
8.1 Principe.	129
8.2 Réalisation.	129
8.3 Construction du modèle prédictif.	130
8.4 Utilisation du modèle prédictif.	132
8.5 Simulation.	133
9. Export CoTab Reader.	135
9.1 Principe.	135
9.2 Réalisation.	135
9.3 Exécution.	136
10. Rapports.	137
10.1 Principe.	137
10.2 Réalisation.	137

Annexe : Procédures

1. Généralités
 2. Création d'une procédure
 3. Le langage CPL.
 - 3.1 Symboles et notations
 - 3.2 Types de variables et conversion implicite.
 - 3.3 Structure d'une procédure CPL
 - 3.4 Instructions de base CPL
 - 3.5 Expressions de conditions et structures de contrôle (CPL)
 - 3.6 Instructions spéciales (CPL)
 - 3.7 Exemple de procédure
- Annexe : Liste des mots-clé

1. Présentation générale.

CoTab est un logiciel d'édition de tableaux statistiques, opérationnel sur ordinateur fonctionnant sous système Windows (XP, VISTA, Windows7), équipé d'une version avec **Framework 3.5**.

Les **données à traiter** peuvent avoir différents formats :

- fichiers ASCII, (éventuellement multi-formats)
- fichiers issus de logiciels spécialisés dans les enquêtes : Conversoft, Interviewer, Quancept, Triple S,....
- fichiers issus du logiciel Cosi
- bases de données : Access,...(dans une version ultérieure)

Il n'y a pas de limites sur le volume des données d'entrée, qui peuvent présenter un ou **plusieurs niveaux** d'unités statistiques pour la version complète du logiciel (**la version *light* ne permet pas de traiter les enquêtes à niveaux**, donc tous les articles de ce manuel relatifs aux niveaux ne la concernent pas).

CoTab présente 2 grandes fonctionnalités :

- **la construction des variables**, et leur construction à partir du fichier d'entrée et de toute opération de recodification nécessaire : il s'agit donc de décrire toute l'information utile au niveau des individus,
- **la construction des tableaux**, dans lesquels on ventilera les informations précédentes. Ils peuvent être produits en grand nombre, et avec de multiples options.

L'utilisateur dispose également de **modules annexes** lui permettant d'accomplir :

- **des redressements d'échantillons**
- **des exportations** de fichiers en différents formats.
- **un module de calcul prédictif**

Le passage des unes aux autres est immédiat, ce qui fait que l'utilisateur peut définir des variables, puis des tableaux, les exécuter, les visualiser, modifier des variables, revenir aux tableaux, etc.

Exécution multi-fichiers : toutes les exécutions de CoTab(tableaux, exportations, etc..) peuvent s'effectuer sur un ou plusieurs fichiers, (par exemple plusieurs vagues d'enquêtes), qui seront donc traités séquentiellement, comme un seul fichier concaténé : il suffit de nommer tous ces fichiers dans la fenêtre demandant le nom du fichier à traiter.

CoTab est conçu de manière à offrir au chargé d'études un instrument sûr, puissant, et facile à utiliser. De plus, sa conception récente (2004) garantit sa pérennité dans le cadre des évolutions techniques futures (Windows, Internet,...).

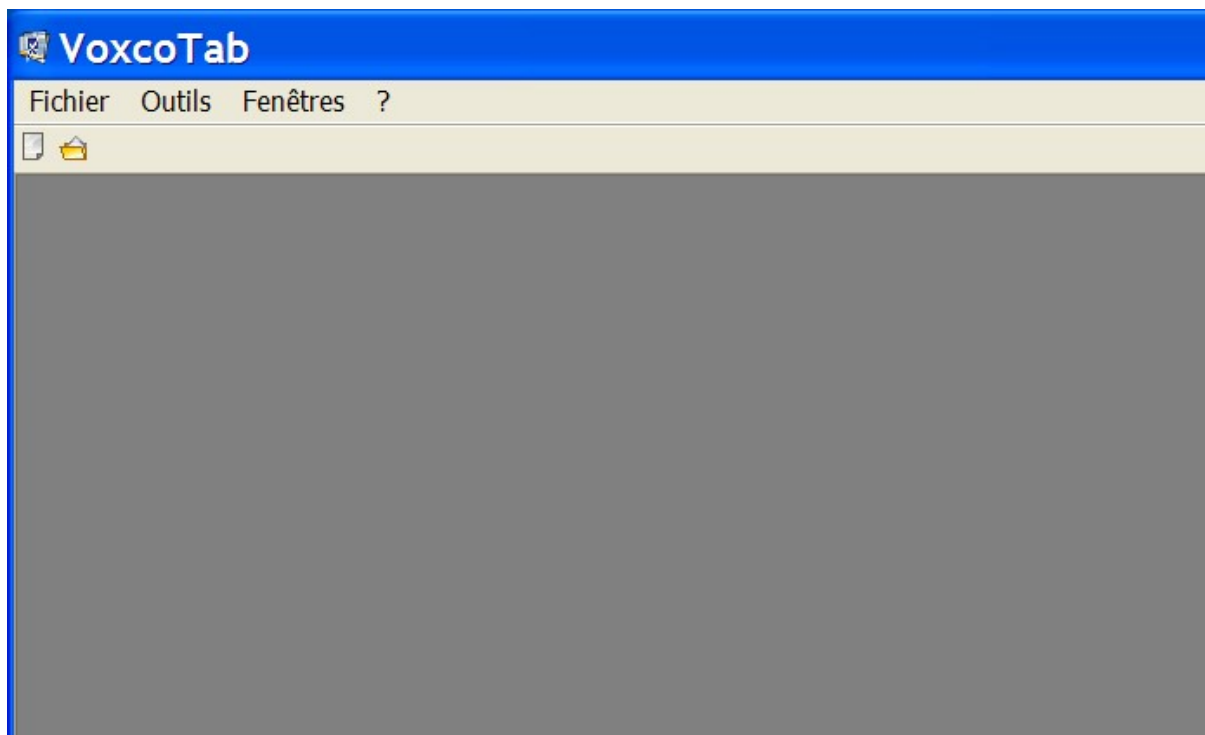
Au point de vue **ergonomique**, l'utilisateur familier avec Windows retrouvera son environnement habituel : menus, icônes, clic droit, raccourcis clavier. La souris est obligatoire.

2. Les études.

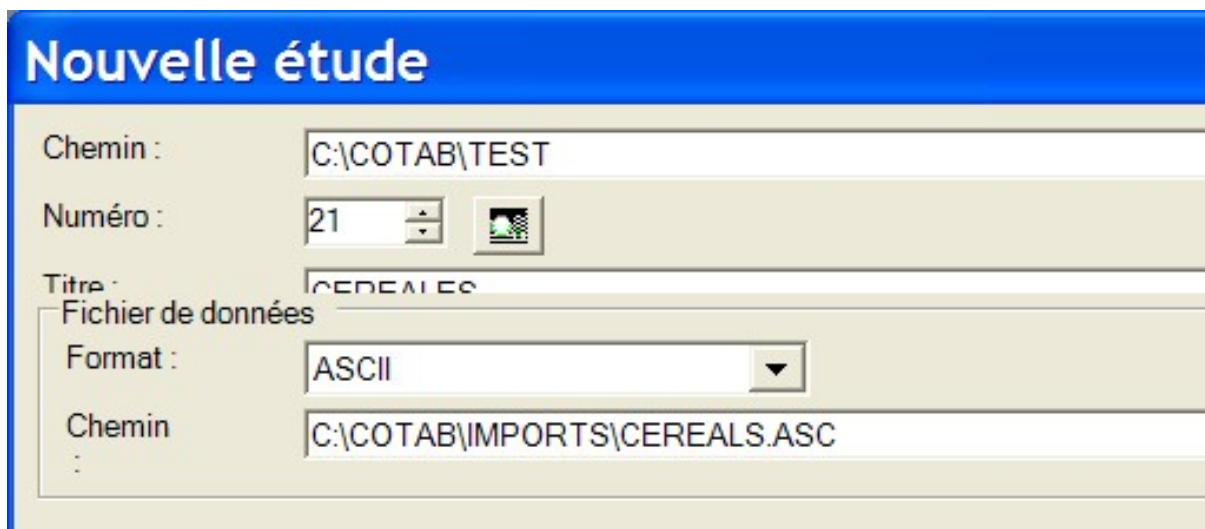
2.1 Démarrage.

Une étude correspond à un traitement donné (variables, tableaux, modules annexes) qui sera appliqué à un fichier de données.

Dès l'ouverture de CoTab, le menu initial propose de **créer**, **importer**, **ouvrir** ou **supprimer** une étude.

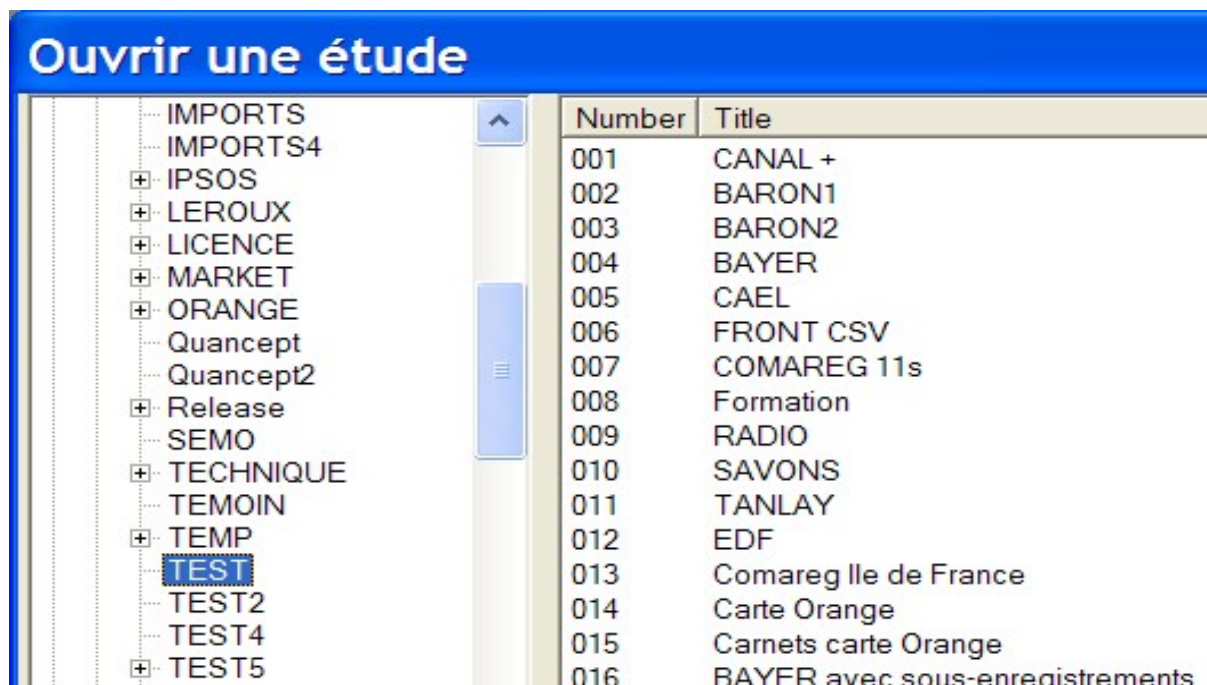


En cas de création, on doit indiquer le chemin, ainsi que le titre de l'étude et un numéro (1 à 999) qui sera utilisé dans le nom des fichiers de l'étude.

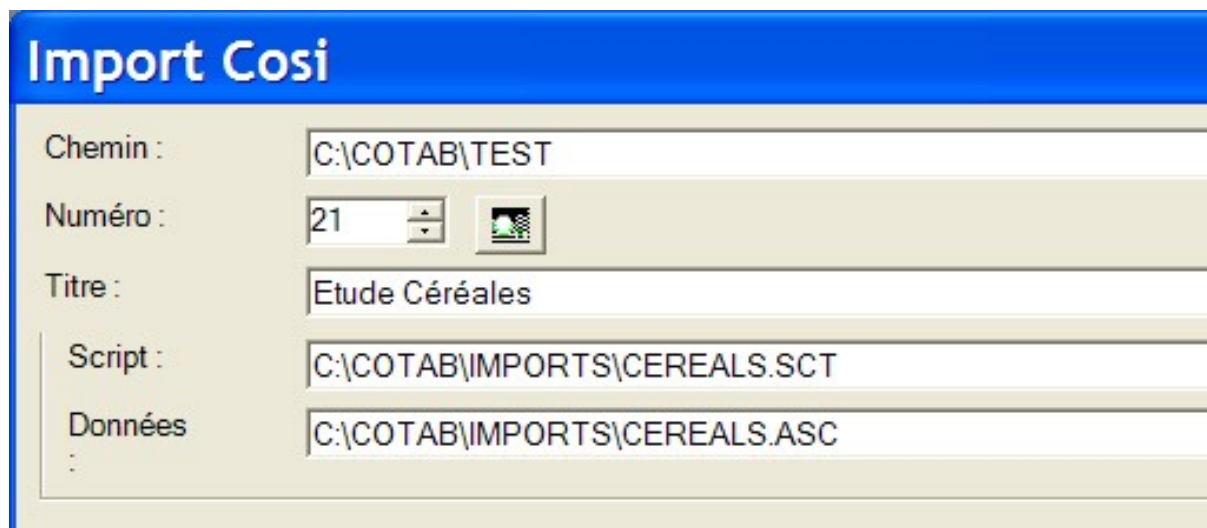


On doit préciser également s'il s'agit d'une étude à niveaux.
On peut également donner le nom du fichier de données à traiter, si on le connaît déjà (sinon, on le donnera lors d'une exécution ultérieure).

En cas d'ouverture d'une étude existante :



Dans le cas d'une importation :



On peut à tout moment sauvegarder l'étude en cours, et également la sauvegarder sous un autre numéro. De plus, quand on quitte une étude, une fenêtre demande à l'utilisateur s'il veut sauvegarder son travail.

2.2 Nouvelle étude.

2.2.1 Définitions.

Une nouvelle étude peut être avec ou sans niveaux : une étude avec niveaux contient plusieurs natures d'unités statistiques, un fichier sans niveau contient une seule nature d'unités statistiques.

Etude sans niveaux.

Dans le fichier à traiter, chaque individu statistique comporte zéro, un ou plusieurs types d'enregistrements.

Dans le cas, où il y a **zéro ou un type d'enregistrement par unité statistique**, il n'est pas obligatoire de renseigner **l'identifiant** (un enregistrement correspond à une unité statistique).

En revanche, on peut renseigner le type d'enregistrement pour sélectionner seulement les enregistrements correspondant au type indiqué, ou pour lire des informations variables selon le type d'enregistrement.

S'il y a **plusieurs types d'enregistrements** à prendre en compte, concernant **la même unité statistique**, il est obligatoire de définir un identifiant et les types d'enregistrements.

L'identifiant est un champ alphanumérique ou numérique, situé dans n'importe quelle position des enregistrements.

Les types d'enregistrements seront définis par une ou plusieurs conditions.

Par exemple : Identifiant en position 1 à 5 et types d'enregistrements définis par une lettre A, B, C, ou D en position 10.

ATTENTION : il est impératif que le fichier en entrée soit trié sur l'identifiant, mais il n'est pas nécessaire qu'il soit trié sur les types d'enregistrements

Tous les enregistrements n'appartenant pas aux types déclarés sont ignorés.

Pour un même identifiant, si on rencontre plusieurs fois le même type, c'est le dernier rencontré qui est pris en compte.

Remarque : si on déclare plusieurs types d'enregistrement, sans déclarer d'identifiant pour les unités statistiques, chaque enregistrement sera considéré comme une unité statistique, quel que soit son type.

Etude avec niveaux.

Il peut y avoir plusieurs niveaux, hiérarchisés ou non, en nombre non limité, correspondant chacun à une unité statistique et un ensemble de variables bien spécifié. Par exemple :

NIVEAUX	VARIABLES
1 - MENAGE	taille du foyer, région, équipement
2 - PERSONNE	sexe, âge, profession
3 - DEPLACEMENT	mode de transport, origine, destination

Quand il y a des niveaux, il faut en indiquer la hiérarchie, le premier NIVEAU est le NIVEAU 1 ; chaque niveau est rattaché à un seul niveau, appelé niveau Père.

Par rapport au fichier d'entrée, chaque niveau correspond :

- à un ou plusieurs types d'enregistrements, un type d'enregistrement étant défini par une ou plusieurs conditions,
- ou bien, il peut être défini uniquement par l'identifiant, la rupture sur cet identifiant provoquant la création d'un nouvel individu de ce niveau.

Cet identifiant est nécessaire s'il y a plusieurs types d'enregistrements à ce niveau, ou si le niveau a des descendants.

Seuls les niveaux sans descendants peuvent n'avoir ni identifiant, ni type. Dans ce cas, tout enregistrement déclaré, ou tous les enregistrements si aucun type n'est déclaré pour ce niveau, seront réputés appartenir à ce dernier niveau.

Si, pour un niveau, on ne précise pas d'identifiant, c'est l'apparition du type (ou tous les types) d'enregistrement correspondant qui déterminera la rupture de niveau pour ce niveau. En revanche, s'il y a un identifiant pour un niveau, il doit exister pour tous les types d'enregistrements de ce niveau et de ses descendants.

Le fichier doit être trié sur l'identifiant du niveau 1.

En revanche, tous les autres enregistrements d'une même famille (c'est-à-dire avec le même identifiant de niveau 1) peuvent apparaître dans n'importe quel ordre.

2.2.2 Mise en œuvre : déclaration de la structure de l'étude.

- **Si l'étude est sans niveau, avec un seul type d'enregistrement** (fichier ASCII à positions fixes, ou fichier CSV, à champs délimités) : dans ce cas, il n'y a rien à déclarer, on peut tout de suite commencer à créer les variables.
- **Si l'étude est multi-enregistrements, et/ou multi-niveaux** (fichier ASCII à positions fixes). Dans ce cas, l'icône « définie structure », représentant deux flèches obliques inversées donne accès à un écran permettant de définir la structure du fichier d'entrée (c'est-à-dire les types d'enregistrement), ainsi que, si nécessaire, la structure des niveaux.

Structure de l'étude					
Types d'enregistrements					
N°	Règle				
1	(1-1 = "1")				
2	(1-2 = "2")				
Structure des individus					
N°	Nom	Types d'enregistrements	De	A	Univers

Définition des types d'enregistrements : un click droit permet de créer un nouveau type d'enregistrements :

Type d'enregistrement

Numéro d'enregistrement : 2

Règles

N°	De	A	Opérateur	Code
1	1	2	=	2

Sous enregistrements

Chaque type d'enregistrement est défini par une ou plusieurs règles, chaque règle étant définie par un champ et des valeurs possibles (au moyen des opérateurs =, != ; >, <, <=, >=

Un clic droit permet de créer ou supprimer une condition.

On peut définir également des sous-enregistrements, lorsque l'enregistrement contient plusieurs unités statistiques, pour lesquelles toutes les variables sont décalées d'un pas constant ; on donne alors :

- le pas, supposé constant séparant les valeurs de deux individus consécutifs,
- le nombre de sous enregistrements (donc d'unités statistiques à créer).

Exemple : l'enregistrement correspond à un individu, les sous-enregistrements correspondent à ses n enfants possibles, pour lesquels on aura les variables Sexe, Age, etc..

Remarques :

- Un enregistrement du fichier peut appartenir à plusieurs types.
- **Si l'étude n'est pas déclarée à niveaux**, et si on définit des sous-enregistrements, les seules unités statistiques prises en compte seront celles correspondant aux sous-enregistrements (dans l'exemple ci-dessus, on ne traitera donc que les enfants, pas les individus).
- En utilisant l'option « sous-enregistrements », si les différentes variables correspondent à des pas non identiques (par exemple, le pas étant de 1 pour le Sexe, et de 2 pour l'Age), il faut alors déclarer 2 types

d'enregistrement, l'un comportant des sous-enregistrements avec un pas de 1, l'autre comportant des sous-enregistrements avec un pas de 2, et on lira les variables Sexe et Age dans ces 2 type respectivement.

Définition des niveaux.

Structure de l'étude						
Types d'enregistrements						
NÂ°	Règle					
1	(1-3 = "F20")					
2	(1-3 = "F21")					
3	(1-3 = "F22")					
4	(1-3 = "F23")					

Structure des individus						
N...	Nom	Parent	Types d'enregistrements	De	A	Univers
1	Individus	0	1	10	15	
2	Jours	1	2	16	16	
3	Déplacements	2	3	17	18	
4	Trajets	3	4	19	20	

Pour chaque niveau, on doit déclarer :

- le nom de l'unité statistique
- le niveau de rattachement
- le, ou les types d'enregistrements concernés par ce niveau
- le cas échéant, la position de l'identifiant de ce niveau
- le cas échéant, l'**univers** conditionnant l'existence de ce niveau

Remarques :

- si un niveau appelle un type d'enregistrement contenant des sous-enregistrements, ce niveau doit être terminal (sans descendants), et **il ne faut pas déclarer d'identifiant pour ce niveau.**
- si un niveau appelle un type d'enregistrement contenant des sous-enregistrements, il peut appeler d'autres types d'enregistrement générant le même nombre de sous-enregistrements, mais il ne doit pas appeler des types d'enregistrements « normaux » (ne générant pas de sous-enregistrements).
- Pour le dernier niveau, (niveau sans descendant), on accepte qu'il n'y ait pas d'identifiant, à condition qu'il appelle soit des types avec sous-enregistrements, soit au maximum un seul type d'enregistrement de type normal
- Si, pour un niveau, on n'a déclaré aucun type d'enregistrement, par convention, tous les enregistrements sont censés lui appartenir.
- Tout ce qui précède est valable pour les études sans niveau, en considérant implicitement qu'il s'agit d'un niveau 1 sans descendants.

Univers.

On peut mettre dans ce champ **une variable logique** permettant de filtrer les individus du niveau concerné par l'écriture d'un univers, en ne conservant que ceux pour lesquels la variable a la valeur *vrai*.

Les individus ainsi filtrés seront les seuls transmis au programme en cours d'exécution (que ce soit une demande de tableau, une demande de redressement, une exportation de fichiers, etc.)

Remarque : s'il n'y a pas de niveaux, on peut néanmoins utiliser la première ligne pour définir un univers permettant de filtrer les individus traités.

2.3 Importation d'études.

Les études importées sont sans niveaux.

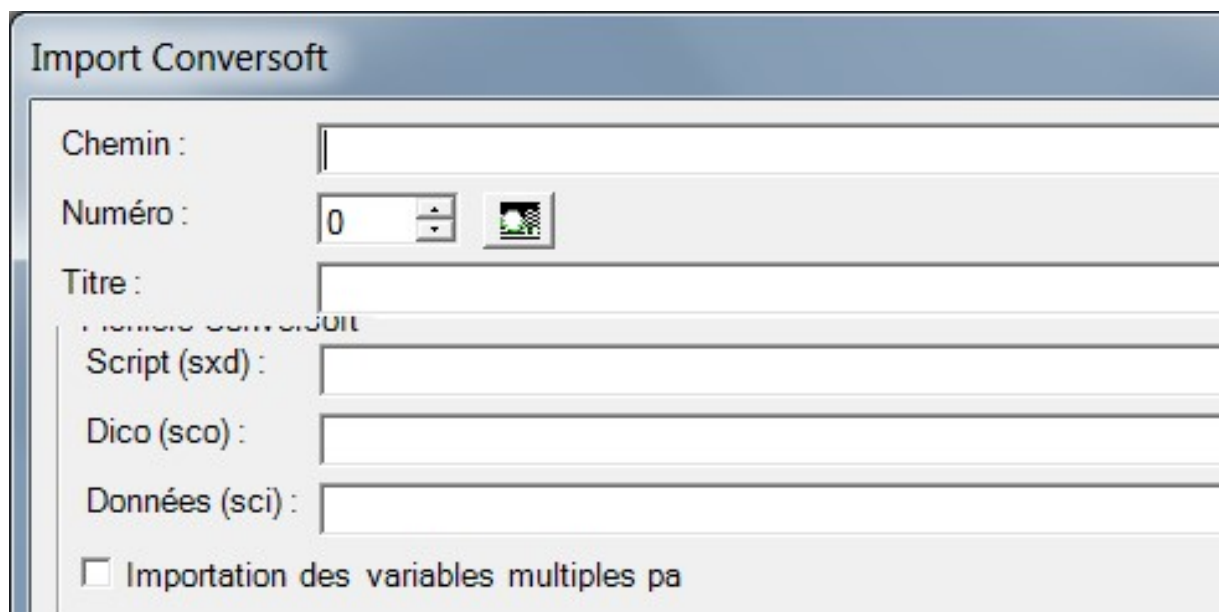
On peut importer des études provenant des logiciels (ou ayant le format) suivants :

- Conversoft
- Cosi
- Quancept
- SPSS
- Triple S-xml ou Txt
- Excel (xls ou xlsx)

2.3.1 Importation Conversoft.

Il faut disposer des fichiers issus de Conversoft d'extension *sxd*, *sco* et *sci*.

Dans le menu *fichier/imports*, on clique sur Conversoft :



On indique :

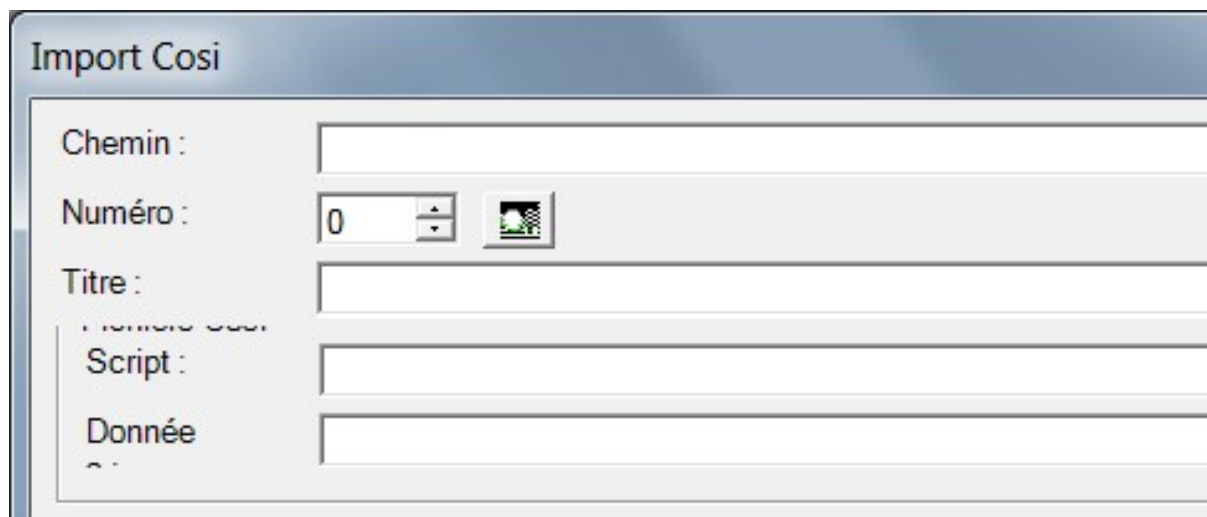
- le numéro de l'étude que l'on veut créer ainsi que son emplacement et son titre.
- Les fichiers issus de Conversoft d'extension *sxd*, *sco* et *sci*.
- Eventuellement le codage « par rangs » des variables multiples.
- Si on constate que l'importation standard ne respecte pas les caractères accentués, il suffit de la recommencer en cochant la case « encodage » et de choisir dans la liste proposée la ligne pertinente. En général, il s'agit de la ligne « ASMO » ou de la ligne « UTF-8 ».

En faisant OK, l'étude est alors créée. On peut la modifier en ajoutant toutes les nouvelles variables désirées, en supprimant des variables origines, etc..

2.3.2 Importation Cosi.

On peut importer les fichiers issus du menu *export* de Cosi : il s'agit de fichiers d'extension *sct* (script) et *asc* (données).

En cliquant sur Cosi dans le menu *fichier/imports*, on obtient :



The screenshot shows a dialog box titled "Import Cosi". It contains the following fields and controls:

- Chemin :** A text input field.
- Numéro :** A numeric input field containing the value "0", with a small icon to its right.
- Titre :** A text input field.
- Script :** A text input field with a small icon to its right.
- Donnée :** A text input field with a small icon to its right.

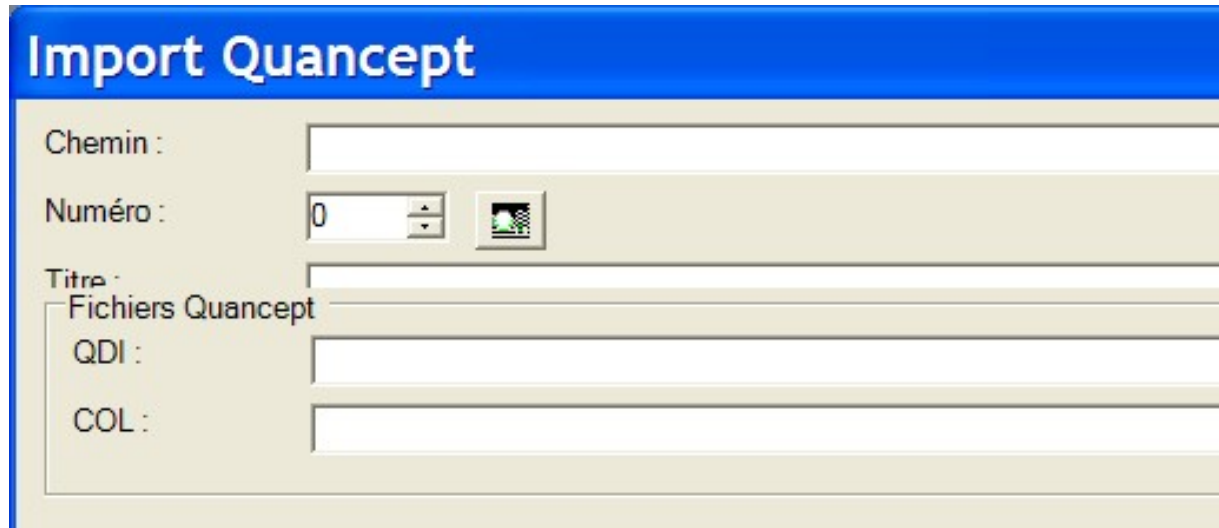
On indique :

- le numéro de l'étude que l'on veut créer ainsi que son emplacement et son titre.
- Les fichiers issus de Cosi d'extension *sct*, *asc*.
- Si on constate que l'importation standard ne respecte pas les caractères accentués, il suffit de la recommencer en cochant la case « encodage » et de choisir dans la liste proposée la ligne pertinente. En général, il s'agit de la ligne « ASMO » ou de la ligne « UTF-8 ».

En faisant OK, l'étude est alors créée. On peut la modifier en ajoutant toutes les nouvelles variables désirées, en supprimant des variables originales, etc..

2.3.3 Importation Quancept.

Le fichier de données doit être en mode ASCII (un répondeur = un enregistrement), et non en mode binaire (un répondeur = n enregistrements).



On indique :

- le numéro de l'étude que l'on veut créer ainsi que son emplacement et son titre.
- Les fichiers issus de Quancept d'extension *qdi et col*. (ou tout au moins des fichiers correspondant à ces extensions, rien n'empêche que l'extension soit *qda* par exemple).

Traitements spécifiques au format Quancept. Ce qui suit suppose la connaissance de l'attribut *format* des variables, décrit dans un chapitre ultérieur).

Les fichiers issus de Quancept présentent des valeurs particulières pour les variables nominales (simples ou multiples), à savoir :

- « 0 » : signifie autres valeurs (OS)
- « - » : signifie No answer (NA)
- « & » : signifie don't know (DK)

La valeur « 0 » ne pose pas de problème, étant une valeur admise dans CoTab aussi bien pour les simples que pour les multiples.

En revanche, les 2 autres valeurs sont transformées ainsi :

- **Si la variable est multiple** : les codes spéciaux - et & seront transformés, aussi bien dans le format associé à la variable que dans les valeurs lues dans le fichier de données, dans les nouvelles valeurs -1 et -2.

Formats

Nom : Type : Multiple

NÂ°	Code	Libellé	Pas d...	NE	Valeur
1	1	Ce numéro est trop cher	<input type="checkbox"/>	<input type="checkbox"/>	
2	2	Vous n'aimez pas les publicités de ce numéro	<input type="checkbox"/>	<input type="checkbox"/>	
3	3	Vous n'avez pas confiance en ce numéro	<input type="checkbox"/>	<input type="checkbox"/>	
4	4	Vous avez déjà appelé ce numéro et vous a...	<input type="checkbox"/>	<input type="checkbox"/>	
5	5	Quand vous appelez les renseignements, v...	<input type="checkbox"/>	<input type="checkbox"/>	
6	6	Ce numéro ne propose pas les services don...	<input type="checkbox"/>	<input type="checkbox"/>	
7	7	Vous avez l'habitude de composer un autre...	<input type="checkbox"/>	<input type="checkbox"/>	
8	8	Vous ne pouvez pas accéder à ce service d...	<input type="checkbox"/>	<input type="checkbox"/>	
9	9	Les télé-opérateurs ne sont pas compétents	<input type="checkbox"/>	<input type="checkbox"/>	
10	10	Les télé-opérateurs ne sont pas agréables a...	<input type="checkbox"/>	<input type="checkbox"/>	
11	0	Other (specify)	<input type="checkbox"/>	<input type="checkbox"/>	
12	-2	Don't know	<input type="checkbox"/>	<input type="checkbox"/>	

- **Si la variable est simple** : afin de prendre en compte ces 2 nouvelles valeurs qui ne sont pas des nombres, la variable simple de Quancept sera transformée dans CoTab en une variable **de type multiple**. Les codes spéciaux - et & seront transformés, aussi bien dans le format associé à la variable que dans les valeurs lues dans le fichier de données, dans les nouvelles valeurs -1 et -2. Toutefois, dans la mesure où cette variable pourrait donner lieu à des calculs de moyennes, la colonne « valeur » du format créé sera initialisée avec les valeurs numériques 1 à n (pas d'initialisation pour les 3 codes spéciaux ci-dessus).

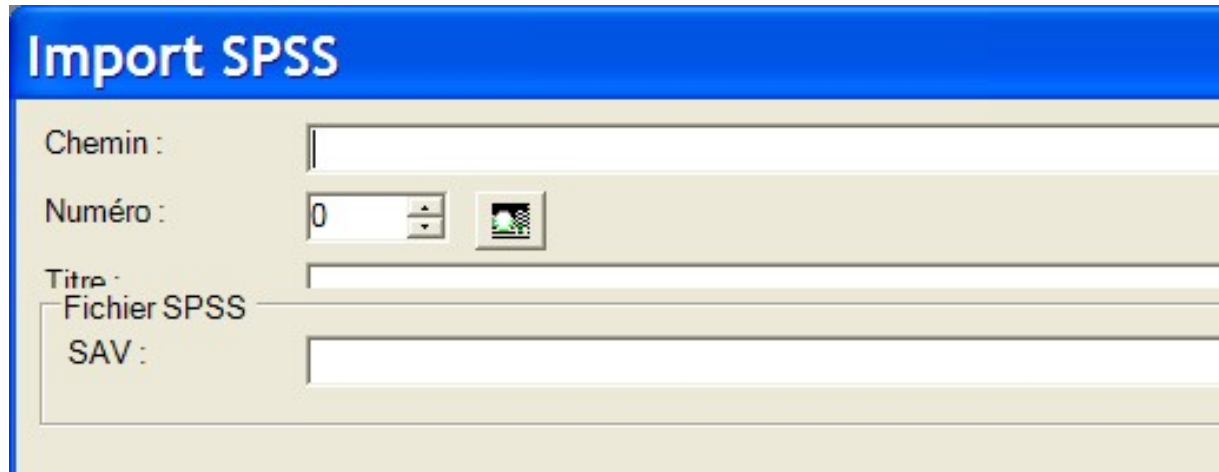
Formats

Nom : Type : Multiple

NÂ°	Code	Libellé	Pas d...	NE	Valeur
1	1	Tout à fait d'accord	<input type="checkbox"/>	<input type="checkbox"/>	1.000
2	2	Plutôt d'accord	<input type="checkbox"/>	<input type="checkbox"/>	2.000
3	3	Plutôt pas d'accord	<input type="checkbox"/>	<input type="checkbox"/>	3.000
4	4	Pas d'accord du tout	<input type="checkbox"/>	<input type="checkbox"/>	4.000
5	-2	Don't know	<input type="checkbox"/>	<input type="checkbox"/>	

2.3.4 Importation SPSS.

On peut importer les fichiers d'extension *sav* issus de SPSS .
En cliquant sur SPSS dans le menu *fichier/imports*, on obtient



On indique :

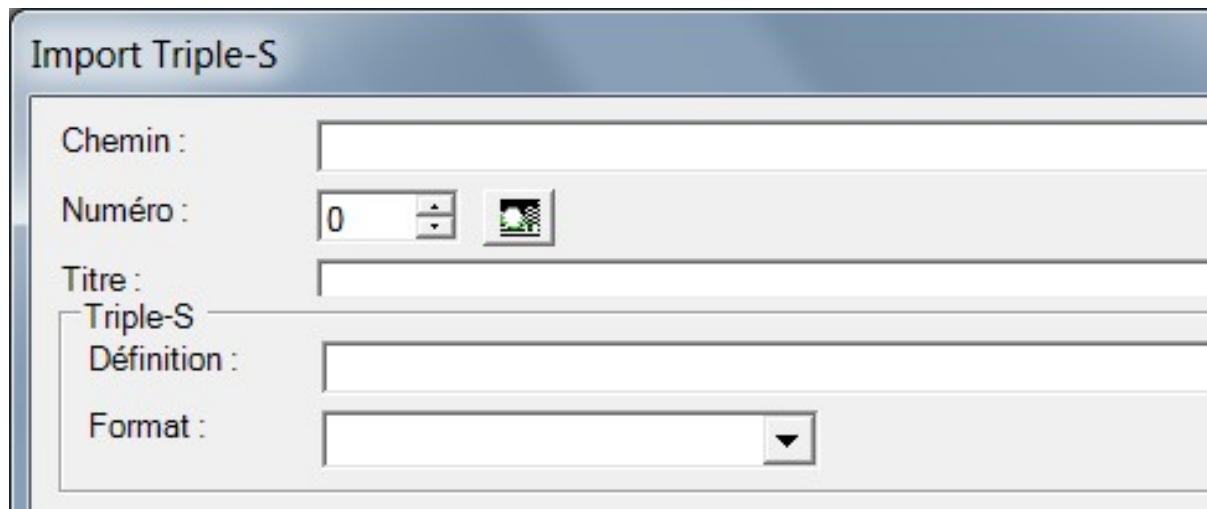
- le numéro de l'étude que l'on veut créer ainsi que son emplacement et son titre.
- Le fichier issu de SPSS d'extension *sav*.

En faisant OK, l'étude est alors créée. On peut la modifier en ajoutant toutes les nouvelles variables désirées, en supprimant des variables originales, etc..

Le fichier de données correspondant sera également créé, sous le nom E_nnn.asc, nnn étant le numéro de l'étude demandé.

2.3.5 Importation TripleS.

Les fichiers TripleS doivent être créés au **format XML**. Toutefois, les fichiers en mode TXT sont acceptés par CoTab à partir de la version 1.8.



On indique :

- le numéro de l'étude que l'on veut créer ainsi que son emplacement et son titre.
- Le fichier avec le format tripleS.
- Dans le champ « Format », on doit préciser si le fichier de données est ASCII ou CSV.
- Si on constate que l'importation standard ne respecte pas les caractères accentués, il suffit de la recommencer en cochant la case « encodage » et de choisir dans la liste proposée la ligne pertinente. En général, il s'agit de la ligne « ASMO » ou de la ligne « UTF-8 ».

Attention : ce fichier ne doit pas contenir l'instruction :

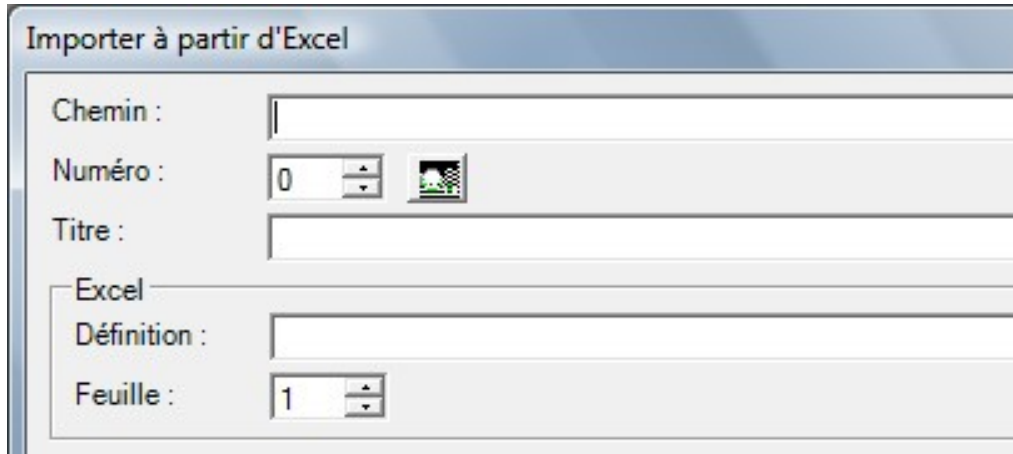
```
<!DOCTYPE sss PUBLIC "-//triple-s//DTD Survey Interchange v1.2//EN"  
    "http://www.triple-s.org/dtd/sss_v12.dtd">
```

La présence de ce fichier suppose que l'on a une connexion Internet active.

Si c'est le cas, il suffit de supprimer cette instruction avec un éditeur quelconque (par exemple le bloc notes).

2.3.6 Importation Excel (xls ouxlsx).

L'importation d'un fichier Excel se fait dans le menu « Fichier », « import », et elle se traduit par la création du dictionnaire de variables, et du fichier de données en format csv:



- les variables sont créées à partir de la 1ère ligne du fichier Excel :

Le nom des variables est V_{nn} , nn étant le numéro de colonne

Le type est « texte »

Le titre est le nom de la colonne

La fonction est Input csv, avec le bon numéro de colonne

- un fichier Ennn_csv est créé comme fichier de données, en supprimant la 1ère ligne du fichier Excel, et en le mettant sous forme csv, avec « ; » comme séparateur.

- **Attention :**

1. si certaines colonnes sont constituées de texte contenant le signe « ; » il faut au préalable supprimer ce signe, ou le remplacer par exemple par « , », pour éviter la création de **colonnes parasites** dans le fichier csv.

2. si certaines colonnes sont constituées de texte contenant un caractère « saut de ligne », il faut au préalable supprimer ce signe, ou le remplacer par « , » par exemple, pour éviter la création de **lignes parasites** dans le fichier csv.

Dans les 2 cas, le remplacement peut se faire directement dans le fichier Excel en recherchant suivant le cas [;] ou [Ctrl + J] .

Conseils pour améliorer le traitement.

L'utilisateur est libre de modifier les types des variables, à ses risques et périls (attention aux quantités, CoTab demande un point décimal et non une virgule).

L'utilisateur peut demander la création automatique de format pour les variables qu'il souhaite tabuler.

2.4 Ouverture d'une étude.

Ayant ouvert une étude (menu *fichier/ouvrir*), le dictionnaire de variables est affiché. L'utilisateur peut alors :

- Modifier le titre de l'étude (menu *fichier/propriétés*),
- Modifier le dictionnaire de variables, en ajoutant, modifiant ou supprimant des variables. Attention : la suppression de variables utilisées dans les différents modules exécutables de CoTab (tableaux, exportations, redressements), conduira à l'échec dans l'exécution des modules correspondants
- Construire des tableaux (onglet *tableaux*).
- Demander une exportation de fichiers (menu *outils/exports*)
- Demander un redressement d'échantillon (menu *outils/redressement*)

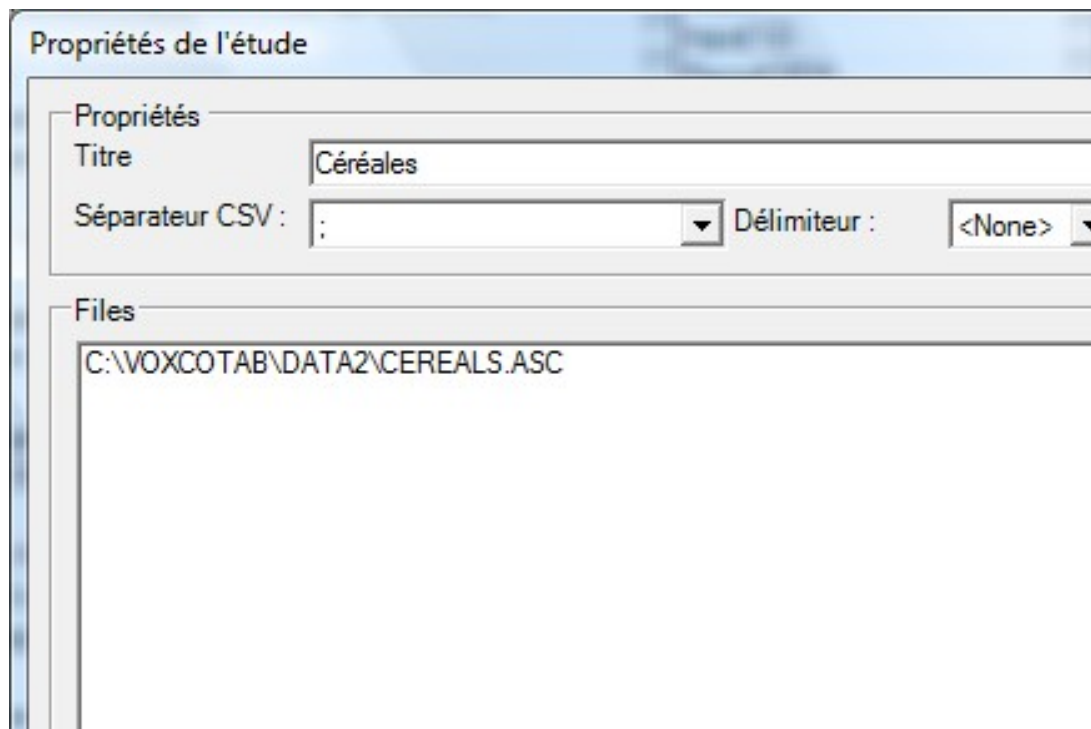
Sauvegardes : toutes les modifications du dictionnaire de variables seront sauvegardées soit en cliquant sur l'icône de sauvegarde, soit en cliquant sur *sauvegarde* dans le menu *fichier/enregistrer* ou *fichier/enregistrer sous*.

On peut donc sauvegarder l'étude sous un autre répertoire.

Si on sort du programme sans avoir effectué l'une de ces opérations, les modifications faites seront perdues. Toutefois, toute sortie du programme génère une demande de sauvegarde afin de minimiser les risques d'oubli.

2.5 Ecran « Propriétés ».

L'appel à cet écran se fait dans le menu *outils* :



Cet écran permet de :

- créer ou modifier le titre de l'étude,
- donner les séparateurs et délimiteurs des fichiers csv éventuellement traités
- donner le nom du fichier de données à traiter. On peut mettre plusieurs noms de fichiers dans cette fenêtre, auquel cas le traitement portera sur la concaténation de tous ces fichiers.
- d'activer l'option « cache » : voir paragraphe 3.8.6.

2.6 Suppression d'une étude.

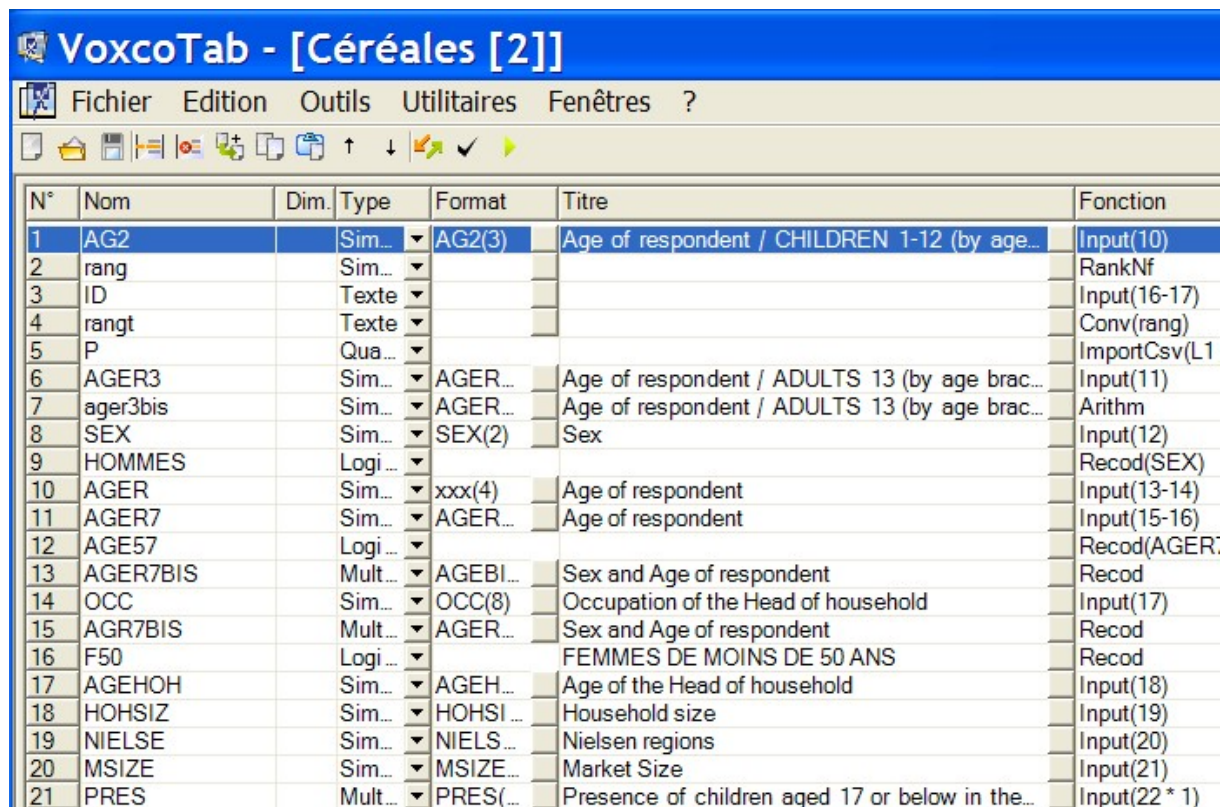
Dans le menu *fichier/supprimer*, on peut sélectionner une étude dans le répertoire de son choix, et la supprimer. Une confirmation sera demandée à l'utilisateur.

3. Les variables.

3.1 Généralités.

Les variables contiennent l'information entrant dans les tableaux statistiques. Chaque individu statistique du fichier à traiter a une valeur pour chaque variable.

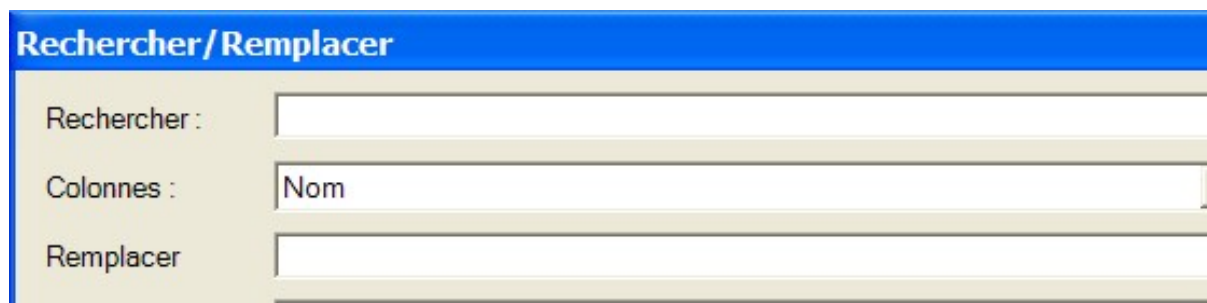
Les variables sont créées automatiquement dans le cas d'ouverture d'une étude existante ou d'une importation (on peut les modifier). Dans le cas d'une étude nouvelle, il faut les créer.



N°	Nom	Dim.	Type	Format	Titre	Fonction
1	AG2		Sim...	AG2(3)	Age of respondent / CHILDREN 1-12 (by age...	Input(10)
2	rang		Sim...			RankNf
3	ID		Texte			Input(16-17)
4	rangt		Texte			Conv(rang)
5	P		Qua...			ImportCsv(L1
6	AGER3		Sim...	AGER...	Age of respondent / ADULTS 13 (by age brac...	Input(11)
7	ager3bis		Sim...	AGER...	Age of respondent / ADULTS 13 (by age brac...	Arithm
8	SEX		Sim...	SEX(2)	Sex	Input(12)
9	HOMMES		Logi...			Recod(SEX)
10	AGER		Sim...	xxx(4)	Age of respondent	Input(13-14)
11	AGER7		Sim...	AGER...	Age of respondent	Input(15-16)
12	AGE57		Logi...			Recod(AGER...
13	AGER7BIS		Mult...	AGEBI...	Sex and Age of respondent	Recod
14	OCC		Sim...	OCC(8)	Occupation of the Head of household	Input(17)
15	AGR7BIS		Mult...	AGER...	Sex and Age of respondent	Recod
16	F50		Logi...		FEMMES DE MOINS DE 50 ANS	Recod
17	AGEHOH		Sim...	AGEH...	Age of the Head of household	Input(18)
18	HOHSIZ		Sim...	HOHSI...	Household size	Input(19)
19	NIELSE		Sim...	NIELS...	Nielsen regions	Input(20)
20	MSIZE		Sim...	MSIZE...	Market Size	Input(21)
21	PRES		Mult...	PRES(...	Presence of children aged 17 or below in the...	Input(22 * 1)

Une icône « check » permet de tester qu'il n'y a pas d'erreur dans la liste des variables.

La **recherche d'une variable**, (ou d'un format, ou d'une fonction) se fait au moyen des touches ctrl F : un menu permet alors d'effectuer cette recherche.



Rechercher/Remplacer

Rechercher :

Colonnes :

Remplacer :

3.2 Attributs des variables.

Chaque variable est décrite avec les attributs suivants :

- le **nom** (16 caractères alphanumériques, plus le caractère « _ », le premier étant une lettre). **Les caractères accentués ne sont pas recommandés**, en particulier **ils ne sont pas reconnus dans les procédures**.
- éventuellement la **dimension** dans le cas où la variable est constituée de n occurrences (n étant la dimension de la variable) : la même question concerne les n occurrences (ce qui évite de décrire n questions identiques, mis à part le titre).
- éventuellement le **niveau** dans le cas où l'étude a été déclarée à niveau lors de sa création.
- le **type** :
 - **simple** : la valeur de la variable pour un individu est un nombre entier, positif, négatif ou nul, ou à défaut la valeur conventionnelle REBUT.
 - **multiple** : un individu peut avoir plusieurs valeurs simultanément parmi l'ensemble (-2, -1, 0, 1, 2, 3, ...), ou à défaut (si aucune de ces valeurs n'est présente), la valeur conventionnelle REBUT.
 - **quantité** : la valeur est un nombre réel, ou à défaut la valeur conventionnelle REBUT.
 - **texte** : la valeur est une suite de n caractères (n étant défini par la fonction de construction de la variable), ou à défaut (zone blanche) la valeur conventionnelle REBUT.
 - **logique** : la valeur peut être *vraie* ou *fausse*.
 - **remarque** : le REBUT pour une variable signifie donc que l'information est manquante, ou non-conforme (par exemple, cas de valeurs alphabétiques dans un champ supposé contenir une valeur numérique).
- le **format** associé : les variables simples, multiples ou textes, peuvent appeler un format, donnant :
 - la liste des codes à faire figurer dans les tableaux lorsque l'on veut la distribution de la variable,
 - les libellés associés,
 - éventuellement les valeurs (parfois appelées « poids ») remplaçant les codes dans les calculs de moyennes
 - et différentes consignes d'édition.
- le **titre** à faire figurer dans les tableaux
- la **fonction** de construction de la variable
- éventuellement la **base** : variable logique, qui servira de filtre automatique pour la variable dans tous les tableaux où elle sera appelée.

3.3 Fonctionnement général.

Si on sélectionne toute la ligne d'une variable, un clic droit permet de :

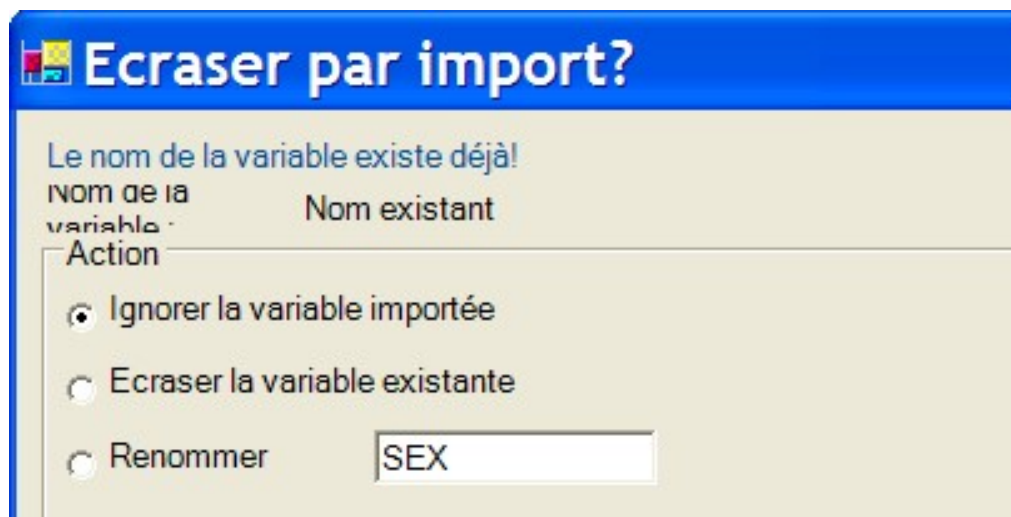
- *insérer*,
- *supprimer*,
- *dupliquer*,
- *copier*,
- *coller*
- ou *importer* une (ou des) variable(s)

Ces 3 dernières fonctions se faisant d'une étude à l'autre, on peut ainsi aisément reprendre dans une étude des variables déjà définies dans une autre étude.

Remarques :

- dans le cas « *coller* », il faut s'assurer que les variables à copier, en provenance donc d'une autre étude, ont toutes des noms différents de ceux existant dans l'étude, sinon le collage sera refusé. Par ailleurs, il vaut mieux coller les variables à la fin du dictionnaire, en ouvrant une nouvelle ligne, pour ne pas risquer d'écraser des variables existantes. Enfin, si une variable à coller possède un format déjà existant, c'est ce dernier qui sera conservé.

- dans le cas d'une *importation*, l'utilisateur doit indiquer l'étude à rechercher, et cocher les variables de cette étude qu'il désire importer. Au cas où une variable importée a le même nom qu'une variable dans l'étude réceptrice, le programme propose plusieurs options : ignorer la variable importée, la renommer, ou écraser la variable déjà existante.



Les variables importées le sont avec tous leurs attributs, y compris leurs formats : si un format a le même nom qu'un format existant, le programme propose également les mêmes options que pour le nom de la variable.

On peut également **déplacer les variables** avec ctrl + flèches : ce déplacement est automatiquement limité pour rester compatible avec les fonctions de la variable en cours et des variables qui en dépendent.

Ces fonctions sont également accessibles par des icônes, par les raccourcis claviers habituels, et par le menu « édition ».

Option de recherche/remplacement. L'action des touches Ctrl + F donne accès à un écran de recherche permettant de rechercher des variables selon leur nom, le contenu de leur titre, etc.

Quand on est positionné sur un champ :

nom, dimension, type, format,... on peut le créer ou le modifier : en tapant directement la valeur pour les champs *nom, dimension* et *niveau*, ou en cliquant sur le bouton à droite du champ pour *type, format, titre, fonction* et *base*.

Exception : quand une variable est associée à un format, on ne peut pas modifier son type s'il s'agit d'une variable texte ; on ne peut le modifier que de simple à multiple et inversement.

Un clic droit sur les champs *format, titre, fonction, base* donne accès aux fonctions *copier, coller* et *reset* (pour annuler le contenu du champ). Les copier/coller peuvent concerner toute une séquence (plusieurs lignes) de chacun de ces champs.

Le **type** des variables peut aussi être modifié, lorsqu'on est positionné sur le champ correspondant, en tapant la première lettre du nouveau type (S, M, Q, T ou L).

Le **niveau** des variables, pour une étude à niveaux, doit être un de ceux définis dans la structure de l'étude, ou bien être égal à zéro : une variable de niveau 0 appartiendra par convention à tous les individus, alors qu'une variable de niveau $j > 0$ appartiendra uniquement aux individus de niveau j .

Le **titre** des variables peut être laissé en blanc.

Options supplémentaires du clic droit :

- **Fusion** : pour fusionner des études (voir § 3.9)
- **Tris à plat** : pour obtenir un tri à plat immédiat de la variable

- **Création du format à partir des données**
- **Thèmes** : pour définir des thèmes et classer la variable dans un des thèmes
- **Références** : donne la liste des variables utilisées pour construire la variable en cours
- **Référencé par** : donne la liste des variables dépendant de la variable en cours

Variables dimensionnées :

Si on clique sur le signe + précédant la dimension, on ouvre la variable, faisant apparaître toutes les occurrences : on peut alors donner un titre à chacune des occurrences (éventuellement par copier/coller).

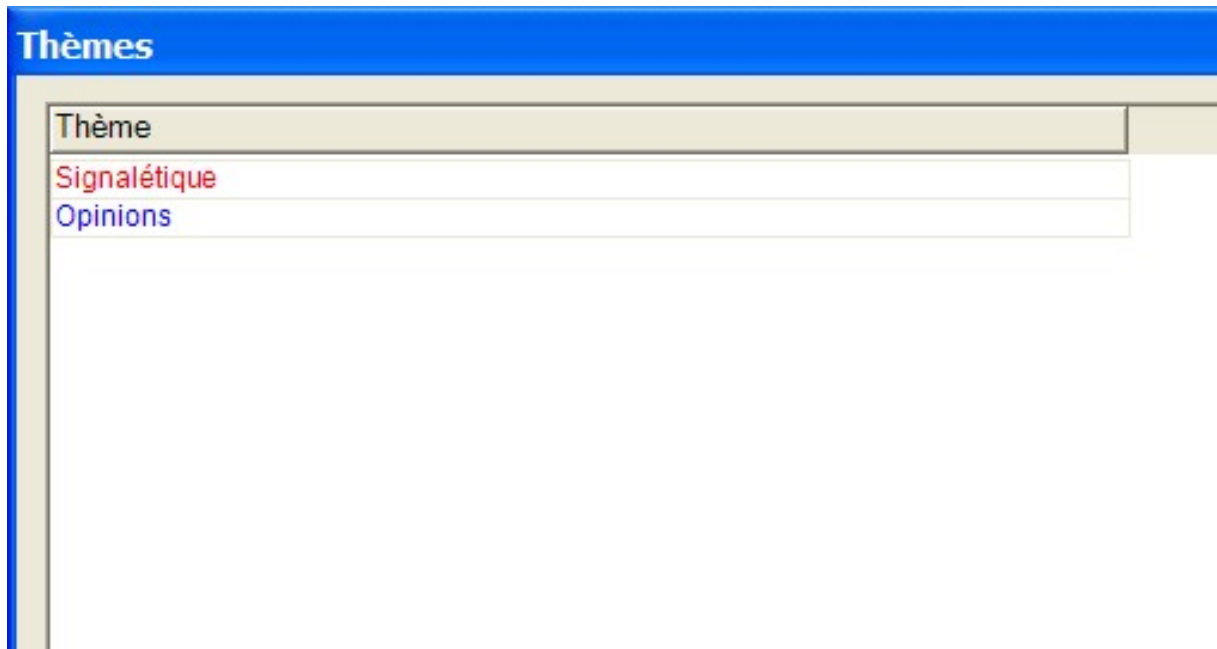
N°	Nom	Dim.	Type	Format	Titre	Fonction
24	QCI		Mult...	QCI(7)	QC - Time when started breakfast	Input(1843 *
25	QCUCI		Sim...	QCUCI...	QC - Time for breakfast	Input(1845)
26	QDI		Mult...	QDI(1)	QD - Had breakfast or not	Input(1846 *
27	V1CL2		Mult...	V1CL2...	Liste 2 (products consumed at breakfast) on...	Input(1925-
28	WQAD		Qua...		Weighting factor - Questionnaire Adults 13	Input(1938-1
29	WQCH		Qua...		Weighting factor - Questionnaire Children 1-12	Input(1948-1
30	WDIARY		Qua...		Weighting factor - Diary /children adults	Input(1968-1
31	NBQ6		Sim...		Q6- Average number of brands known	Input(1978-1
32	V8		Mult...	V8(17)	V8	Input(2083-
33	Z1R2X	12	Sim...	Z1R2X...	Opinions	Input(2085 /
...	<1>				I like trying new food products	
...	<2>				I often eat ready-made cooked meals (fresh, fro...	
...	<3>				I try not to snack between meals	
...	<4>				Breakfast is my favourite meal of the day	
...	<5>				A full breakfast is essential to get my day off to...	
...	<6>				A balanced breakfast helps prevent you feeling t...	
...	<7>				Breakfast is a meal where each person can mak...	
...	<8>				I frequently vary what I eat at breakfast	
...	<9>				During the weekend or on holiday, I eat bigger b...	
...	<10>				I really have to force myself to eat in the morning	
...	<11>				Between summer and winter, I change the way I...	

Thèmes des variables (topic).

Cette nouvelle option permet de classer les variables en plusieurs thèmes, visualisables aisément.

Un clic droit sur une variable donne accès à une fenêtre « thèmes » (topics) : un nouveau clic droit permet de créer autant de thèmes que l'on veut : pour chacun d'eux on donne un titre, et on peut également choisir la police (taille, style, couleur,...) de ce thème, qui affectera le nom de la variable.

Le bouton *Apply* permet alors d'affecter ce type à la variable en cours.



VoxcoTab - [Céréales [2]]

Fichier Edition Outils Utilitaires Fenêtres ?

N°	Nom	Dim.	Type	Format	Titre	Fonction
15	AGR7BIS		Mult...	AGER...	Sex and Age of respondent	Recod
16	F50		Logi...		FEMMES DE MOINS DE 50 ANS	Recod
17	AGEHOH		Sim...	AGEH...	Age of the Head of household	Input(18)
18	HOHSIZ		Sim...	HOHSI...	Household size	Input(19)
19	NIELSE		Sim...	NIELS...	Nielsen regions	Input(20)
20	MSIZE		Sim...	MSIZE...	Market Size	Input(21)
21	PRES		Mult...	PRES(...)	Presence of children aged 17 or below in the...	Input(22 * 1)
22	VENT1		Mult...	VENT1...	Age & Gender (by age brackets) / ADULTS 13	Input(36-37 *
23	VA1		Mult...	VA1(5)	Age & Gender (by age brackets) / ADULTS 13	Input(43 * 1)
24	QCI		Mult...	QCI(7)	QC - Time when started breakfast	Input(1843 *
25	QCUCI		Sim...	QCDU...	QC - Time for breakfast	Input(1845)
26	QDI		Mult...	QDI(1)	QD - Had breakfast or not	Input(1846 *
27	V1CL2		Mult...	V1CL2...	Liste 2 (products consumed at breakfast) on...	Input(1925-...
28	WQAD		Qua...		Weighting factor - Questionnaire Adults 13	Input(1938-1
29	WQCH		Qua...		Weighting factor - Questionnaire Children 1-12	Input(1948-1
30	WDIARY		Qua...		Weighting factor - Diary /children adults	Input(1968-1
31	NBQ6		Sim...		Q6- Average number of brands known	Input(1978-1
32	V8		Mult...	V8(17)	V8	Input(2083-...
33	Z1R2X	12	Sim...	Z1R2X...	Opinions	Input(2085 /
34	VT5		Mult...	VT5(15)	Breakfast features	Input(2131-...
35	Q8T		Mult...	Q8T(96)	Q8/Q8ter. Total products eaten	Input(2237-...

Références des variables.

Ces nouvelles options permettent à l'utilisateur de connaître les implications entre variables.

Un clic droit sur une variable donne accès aux écrans *References* et *References by*. Le premier donne toutes les variables qui ont servi à construire la variable en cours, le deuxième donne toutes les variables construites à partir de la variable en cours.

Variables référencées par AGR7BIS	
N...	Annuler
11	AGER7
8	SEX
Introuvable	

La variable en cours AGR7BIS utilise pour sa construction les variables AGER7 et SEX.

Variable AGER7 référencée par ...	
N...	Variable
12	AGE57
16	F50
15	AGR7BIS
13	AGER7BIS

La variable en cours AGER7 est utilisée pour construire les variables AGE57, F50, etc...

3.4 Formats.

3.4.1 Généralités.

Ils concernent les variables simples, multiples ou textes, et permettent de déclarer les codes qu'il faut éditer dans le cas où l'on souhaite la **distribution de la variable** (codes isolés et/ou sous-totaux), et quels sont les **libellés** à éditer à la place de ces codes. Ils permettent également de donner des valeurs aux différents codes, en vue de **calcul de moyennes** non pas sur les codes origines, mais sur les valeurs correspondantes fournies.

Les formats sont **automatiquement générés** en cas d'importation d'étude, avec comme nom le nom de la variable correspondante.

Pour affecter un format à une variable, on clique sur le bouton à droite du format, et on peut alors :



- **utiliser** un format existant, qui sera donc associé également à la variable en cours : une variable de type texte ne peut appeler qu'un format associé à une variable de type texte, une variable de type simple ou multiple ne peut appeler qu'un format associé à une variable de type simple ou multiple.
- **dupliquer** un format existant, sous un autre nom. Cette option est utilisée si on veut apporter des modifications à ce format, mais seulement pour la variable en cours. Le format appelé est alors affiché dans une fenêtre pour qu'on puisse le modifier.
- **importer** un format présent dans une autre étude (même restriction que ci-dessus).
- **créer** un nouveau format : il faut alors lui donner un nom (1 à 8 caractères alphanumériques), on obtient alors un nouvel écran.

Police par défaut des libellés dans les formats.

Elle peut être définie grâce à une option du menu « Outils »

3.4.2 Création d'un format.

Formats

Nom : Type : Simple

NÂ°	Code	Libellé	Pas d...	NE	Valeur
1	1	Less than 30	<input type="checkbox"/>	<input type="checkbox"/>	
2	2	30 - 39 year old	<input type="checkbox"/>	<input type="checkbox"/>	
3	1-2	Less than 40	<input type="checkbox"/>	<input type="checkbox"/>	
4	3	40 - 49 year old	<input type="checkbox"/>	<input type="checkbox"/>	
5	4	50 - 59 year old	<input type="checkbox"/>	<input type="checkbox"/>	
6	5	60 year old and over	<input type="checkbox"/>	<input type="checkbox"/>	
7	3-5	40 year old and over	<input type="checkbox"/>	<input type="checkbox"/>	

Chaque ligne renseignée du format donnera naissance à une position (ligne, colonne ou page selon le cas) dans le tableau appelant la variable considérée.

Cette position peut correspondre à **un code** de la variable (nombre entier pour les variables simples ou multiples, chaîne de caractères pour les variables textes), ou bien à **une (ou des) plage(s) de codes** (auquel cas la position correspondra au sous-total logique de tous ces codes), ou même éventuellement à aucun code, ce qui fera apparaître une position « commentaire » dans le tableau.

A ce code (ou plage) on associe un **libellé** (maximum 256 caractères), destiné à remplacer le code dans les tableaux concernés par la variable affectée de ce format.

Il faut remplir au minimum les 2 colonnes « codes » et « libellés », donnant la liste des codes de la variable à considérer dans les tableaux dans le cas où l'on veut la distribution de la variable, et les libellés remplaçant ces codes. Pour ce faire, plusieurs méthodes sont possibles :

- **en sélectionnant une ligne**, le clic droit de la souris donne accès aux fonctions d'insertion, suppression, duplication d'une ligne, ainsi qu'au copier/coller.
- **en se positionnant sur un champ**, on peut le créer directement, et se déplacer ensuite avec les flèches.
- on peut également faire des copier/coller dans un champ : par exemple prendre toute une liste de codes, ou de libellés dans un tableau Excel, et les coller dans la colonne correspondante.

- la fonction « **somme** » dans le clic droit de la souris, après avoir sélectionné plusieurs lignes, permet de créer une nouvelle ligne qui sera un **sous-total** de l'ensemble des codes sélectionnés.

- Ce **sous-total** pourra également être obtenu directement, en tapant la liste des codes avec les séparateurs « - » et « ; » : dans ce cas, on peut appeler n'importe quels codes licites, même s'ils ne sont pas déjà présents dans le format.

Exemple : 1-3 ; 5-22 ; 35 ; 43-299

- On peut également **cliquer sur le bouton à droite du champ code**, et sélectionner les codes à prendre en compte dans la liste affichée.

- l'option « **police** », à partir du clic droit de la souris, permet de donner des attributs de police, style, taille au libellé du (ou des) ligne(s) sélectionnée(s).

- toujours avec le clic droit de la souris, on peut commander l'**indentation** du libellé à droite ou à gauche.

- **cas particulier** du code absent : si on fournit un libellé, sans remplir la colonne « code », cela permettra d'insérer une ligne ou colonne vide dans le tableau concernant cette variable.

- la colonne « **sup.tri** », quand elle est cochée, permet d'exclure la ligne correspondante d'un tri éventuel du tableau ; on peut sélectionner plusieurs lignes, et un clic droit donne accès à une option « cocher tout ».

- la colonne « **sup.édition** », si elle est cochée, permet d'exclure la ligne correspondante du tableau : on peut sélectionner plusieurs lignes, et un clic droit donne accès à une option « cocher tout ».

- la colonne « **valeurs** » permet de donner des valeurs avec décimales aux codes correspondants. Si des valeurs sont ainsi données, et si un tableau demande la **moyenne ou le cumul** de la variable, alors ces moyennes ou cumuls seront calculées à partir des valeurs ainsi données dans cette colonne.

Les codes pour lesquels aucune valeur ne sera donnée seront donc exclus du calcul des moyennes ou cumuls.

Les seules valeurs qui seront considérées sont celles figurant dans des lignes correspondant à **un seul code**. Les valeurs figurant dans des lignes correspondant à une liste de plusieurs codes, ou bien à aucun code, seront ignorées.

Si la **variable est multiple**, la valeur retenue sera la somme des valeurs correspondant aux différents codes présents ; si un de ces codes n'a pas de valeur associée, la somme de ces valeurs sera prise égale à *Rebut*.

Si une valeur au moins figure dans la colonne « valeur », cela entraîne que les moyennes seront calculées à partir des seuls nombres figurant dans cette colonne, et jamais à partir des codes eux-mêmes, même si aucune valeur ne figure pour un code : dans ce cas, ce code sera exclu du calcul de la moyenne.

Formats					
Nom : S7DC		Type : Simple			
NÂ°	Code	Libellé	Pas d...	NE	Valeur
1	1	1 - 5 km	<input type="checkbox"/>	<input type="checkbox"/>	2.500
2	2	6 - 10 km	<input type="checkbox"/>	<input type="checkbox"/>	8.000
3	3	11 - 15 km	<input type="checkbox"/>	<input type="checkbox"/>	13.000
4	4	16 - 20 km	<input type="checkbox"/>	<input type="checkbox"/>	18.000
5	5	21 - 25 km	<input type="checkbox"/>	<input type="checkbox"/>	23.000
6	6	26 - 30 km	<input type="checkbox"/>	<input type="checkbox"/>	28.000
7	7	31 - 40 km	<input type="checkbox"/>	<input type="checkbox"/>	35.000
8	8	41 - 50 km	<input type="checkbox"/>	<input type="checkbox"/>	45.000
9	9	A plus de 50 km	<input type="checkbox"/>	<input type="checkbox"/>	70.000
10	10	Sans réponse	<input type="checkbox"/>	<input type="checkbox"/>	

- on peut **déplacer les lignes du format**, au moyen de ctrl + flèche, ou bien des icônes « flèches » : la présentation ainsi faite des lignes du format se retrouvera dans tous les tableaux.
- l'utilisation des formats permet donc d'éditer dans les tableaux la présentation souhaitée de la variable, sans avoir besoin de créer de nouvelles variables définissant les codes ou plages de codes à considérer.
- le nombre de lignes ainsi décrites dans le format sera rappelé dans l'écran principal des variables, entre parenthèses, à la suite du nom du format.
- si un format est appelé par plusieurs variables, on peut le modifier à partir de n'importe laquelle de ces variables.

3.4.3 Génération automatique de formats.

Cette option, présente dans le menu Outils, génère automatiquement des formats pour toutes les variables de type Simple, Multiple, ou Texte n'ayant pas encore de format : elle procède par lecture du fichier de données, recherche de toutes les valeurs possibles pour une telle variable sans format, et génération d'un format en mettant dans les colonnes Codes et Libellés toutes ces valeurs rencontrées. L'utilisateur peut alors modifier ces libellés, faire des regroupements de codes au moyen de l'option Somme, etc...

Cette option est à utiliser pour obtenir la distribution de variables pour lesquelles la liste des modalités est inconnue.

Cette option peut également être demandée **pour une seule variable** : il faut se positionner sur la ligne de la variable, et un clic droit permet de sélectionner l'option « création du automatique du format pour cette variable » (option qui ne sera pas prise en compte si la variable a déjà un format).

3.4.4 Exemples de formats.

1. Edition de plages de valeurs seulement.

Formats

Nom : Type : Simple

NÂ°	Code	Libellé	Pas d...	NE	Valeur
1	13-25	13-25	<input type="checkbox"/>	<input type="checkbox"/>	
2	26-34	26-34	<input type="checkbox"/>	<input type="checkbox"/>	
3	35-49	35-49	<input type="checkbox"/>	<input type="checkbox"/>	
4	50-99	50-99	<input type="checkbox"/>	<input type="checkbox"/>	

Seules les tranches d'âge seront éditées, sans recodification de la variable, qui conservera donc sa vraie valeur pour les calculs de moyennes.

2. Edition de codes, de sous-totaux, et moyennes calculées sur de nouvelles valeurs.

Formats

Nom : Type : Simple

NÂ°	Code	Libellé	Pas d...	NE	Valeur
1	1-2	total satisfait	<input type="checkbox"/>	<input type="checkbox"/>	
2	1	tout à fait satisfait(e)	<input type="checkbox"/>	<input type="checkbox"/>	4.000
3	2	assez satisfait(e)	<input type="checkbox"/>	<input type="checkbox"/>	3.000
4	3-4	total insatisfait	<input type="checkbox"/>	<input type="checkbox"/>	
5	3	peu satisfait(e)	<input type="checkbox"/>	<input type="checkbox"/>	2.000
6	4	pas du tout satisfait(e)	<input type="checkbox"/>	<input type="checkbox"/>	1.000

3.5 Fonctions.

Elles ont pour rôle de donner des valeurs aux variables, par lecture dans le fichier à traiter, ou par lecture dans des tables externes, ou par toute opération de recodification de variables déjà définies.

Quand on appelle une fonction, en cliquant sur le bouton à droite de ce champ, la liste proposée comprend les seules fonctions valides compte tenu des attributs de la variable : type, dimension, niveau, présence de format.

Sélectionner la fonction à exécuter	
Liste des fonctions disponibles	
Nom	Titre
Saisie	lit une valeur dans une zone ASCII, tout type de variable, sous contrôle éventuel d'un type d'enregistre...
SaisieCsv	lit une valeur dans un champ d'un fichier CSV, tout type de variable, sous contrôle éventuel d'un type ...
Import	lit une valeur (tout type de variable) dans une zone d'une table externe ASCII.
ImportCSV	lit une valeur (tout type de variable) dans un champ d'une table externe CSV.
Arithm	opérations arithmétiques, sur variables et constantes.
NbRep	donne le nombre de réponses d'une variable multiple.
NbVal	Nombre de valeurs différentes de rebut d'une variable dimensionnée.
SomVal	Somme(arithmétique ou logique, selon le type de la variable résultat) des valeurs d'une variable dim...
Const	Table de constantes.
Table	Sélection d'une valeur dans une table au moyen d'une variable-indice.
Min	minimum d'une liste de variables (variable dimensionnée ou d'un niveau inférieur).
Max	maximum d'une liste de variables (variable dimensionnée ou d'un niveau inférieur).
PosChn	Position de la première apparition d'une sous-chaîne à l'intérieur d'une chaîne de caractères.
Rang	rang d'enregistrement des individus, incrémenté pour les individus appartenant à l'univers.
RangNF	rang d'enregistrement des individus, incrémenté pour tous les individus.
Conv	conversion de types.
Année	retourne l'année d'une variable temporelle.
Mois	retourne le mois d'une variable temporelle.
Trimestre	retourne le trimestre d'une variable temporelle.
JourSem	retourne le jour de semaine d'une variable temporelle.
Quantième	retourne le quantième d'une variable temporelle.
Semaine	retourne le numéro de semaine d'une variable temporelle.
DiffTemps	retourne la différence entre 2 variables temporelles.
CPL	Fonction CPL.

Pour une variable donnée, la fonction correspondante ne peut appeler comme arguments que des **variables la précédant dans la liste des variables**.

Conversions. Plusieurs fonctions nécessitent des conversions de valeurs avant exécution. Par exemple, dans le cas d'une opération arithmétique dont le résultat est une variable simple (donc entière), et dont un opérande est une variable de type quantité, la valeur de la quantité sera convertie en entier avant de faire l'opération, **par troncature**.

Procédures. La dernière fonction proposée (CPL function) permet d'écrire une procédure dans le langage CPL, pour donner des valeurs à la variable en cours, ou n'importe quelle variable la précédant. Le langage CPL est décrit dans l'annexe « CoTab- Procédures ».

Remarque : La variable en cours ne doit pas être dimensionnée, afin d'éviter les conflits de dimensions avec les instructions éventuellement dimensionnées à l'intérieur de la procédure.

3.5.1 Fonction Saisie.

Une variable affectée de la fonction « saisie » prendra sa valeur dans l'enregistrement en cours de lecture. Si on est dans le cas simple et habituel où il n'y a qu'un enregistrement par individu, il suffit de donner le champ de lecture dans l'enregistrement. Dans les autres cas, il faut donner d'abord le type d'enregistrement concerné, puis le champ dans cet enregistrement.

Si on ne donne pas de type, la lecture se fera dans tous les types d'enregistrement.

Le champ de lecture est donné par ses **positions de départ et d'arrivée** (la première position est appelée 1, et non zéro).

Si la variable est multiple, les positions de départ et d'arrivée sont celles de la 1^{ère} valeur. Le pas entre les différentes valeurs, ainsi que le nombre de valeurs sont fournis comme des options de la fonction.

Si la variable est dimensionnée, les positions de départ et d'arrivée sont celles de la 1^{ère} occurrence, **le pas** indique l'écart entre les champs correspondant à 2 occurrences successives de la variable.

Les valeurs attendues dans le champ de lecture dépendent du type de la variable :

- **variable simple** : nombre entier, éventuellement signé, avec éventuellement des blancs à droite ou à gauche : **[bb][S]nnn[bbb]** donne la valeur **nnn**, avec son signe éventuel, toute autre chose donne REBUT.
- **variable quantité** : nombre avec un point décimal, éventuellement signé, avec éventuellement des blancs à droite ou à gauche : **[bb][S]nnn.nn[bbb]** donne la valeur **nnn.nn**, avec son signe éventuel (mais on peut aussi avoir un nombre entier donc sans point décimal), toute autre chose donne REBUT.
- **exception pour les variables S ou Q** : si on a l'argument « espace →0 » dans la fonction saisie, un champ totalement blanc donne la valeur 0.

- **variable multiple** : il y a 2 modes de saisie :

- **option par valeurs** : elle prend sa valeur dans les k champs demandés : dans chaque champ élémentaire, on doit trouver :
 - . un nombre entier positif ou nul, non signé, avec éventuellement des blancs à droite ou à gauche : **[b]nn[b]** donne la valeur n.
 - . ou bien les valeurs « - » ou « & », qui seront transformées à la lecture en donnant les valeurs -1 et -2 (voir importation des fichiers Quancept).
 - . Toute autre chose est ignorée.

Attention, si la variable est à lire dans 4 champs par exemple, et si on trouve dans ces champs les valeurs 6, 35, XY et blanc, le résultat est l'ensemble des 2 valeurs 6 et 35. La valeur REBUT n'est affectée à la variable que si aucun des champs de saisie ne contient une valeur licite.

- **option par rang** : dans ce cas, les n positions du champ de saisie correspondent aux valeurs 1 à n de la variable respectivement : la variable aura la valeur j si et seulement si il y a un caractère différent de blanc ou zéro dans la jème position du champ de saisie.

- **variable texte** : on prend tout le texte sans aucune conversion. Les blancs à droite et à gauche sont supprimés. La longueur du champ de saisie sera attachée à la variable comme étant son nombre de caractères.

- **variable logique** : un nombre entier positif, éventuellement signé, avec éventuellement des blancs à droite ou à gauche : **[bb][S]nnn[bbb]** donne la valeur **true**, toute autre chose donne **false**.

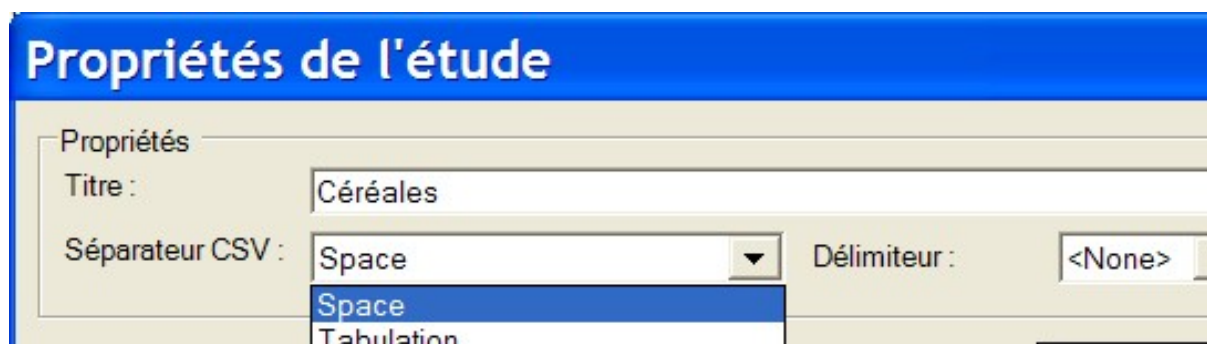
3.5.2 Fonction SaisieCsv.

Cette fonction permet de lire des variables dans les fichiers *csv*. L'utilisation de fichiers *csv* suppose que l'étude est **sans niveaux, et avec un seul type d'enregistrements**.

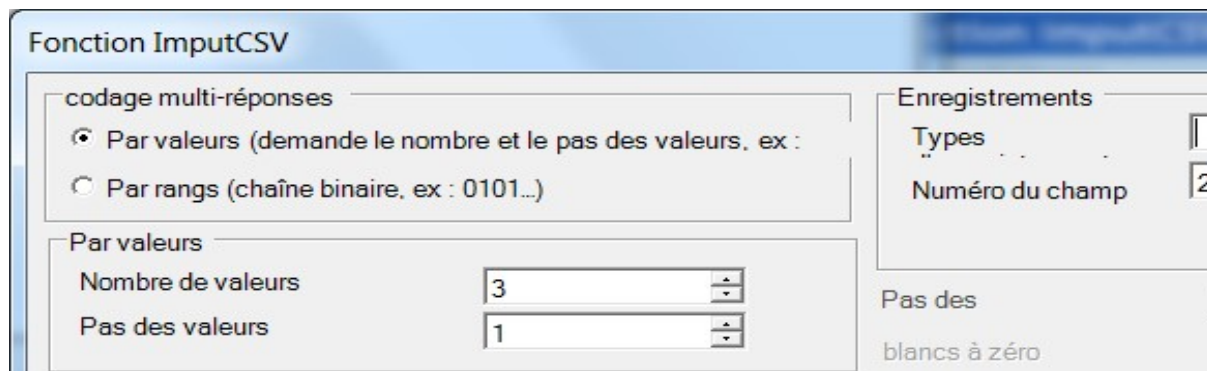
Le caractère séparateur peut être au choix :

- espace
- tabulation
- ;
- ,

L'utilisateur doit définir les caractères « séparateur » et « délimiteur » dans le menu *fichier/propriétés* :



Pour lire les variables, la fonction *saisie csv* permet à l'utilisateur de choisir le champ où figure cette variable.



Comme dans l'autre mode de saisie, on peut renseigner :

- pour les variables multiples, les options par valeur ou par rang, le nombre de valeurs, le pas entre les valeurs (exprimé en nombre de champs)
- pour les variables dimensionnées, le pas entre les occurrences.

Attention pour les variables multiples lues par valeurs :

- s'il n'y a pas de délimiteur, les différentes valeurs seront lues dans des champs différents, espacés du pas donné.

- s'il y a un délimiteur, les différentes valeurs seront lues dans le même champ.

Les valeurs attendues dans chaque champ sont les mêmes que pour la fonction *saisie*.

Pour les variables multiples lues par rangs : on lit une suite de 0 et de 1 dans le champ de saisie.

3.5.3 Fonction Import.

Une variable lue au moyen de la fonction *import* trouvera sa valeur non pas dans le fichier de données, mais dans une *table externe* (voir plus loin le paragraphe tables externes).

L'écran de demande est analogue à celui de la fonction saisie, le type d'enregistrement étant ici remplacé par le lien d'importation (numéro du lien externe).

Une fois trouvé l'enregistrement correspondant à l'individu en cours de traitement, (égalité des clés positions-variables clés), le processus de lecture est exactement le même que celui de la fonction *saisie*.

Remarque : la recherche de l'enregistrement dans la table externe se fait dynamiquement, au moment où le programme doit calculer une variable avec la fonction *import*, donc avec la valeur courante des variables clés.

Si l'utilisateur fait varier la valeur de ces variables (au moyen de procédures par exemple) entre 2 variables toutes 2 lues par la fonction *import* dans la même table externe, alors ces 2 variables pourront être lues dans 2 enregistrements différents de la table externe.

Cas particulier des variables dimensionnées.

Si les clés de la table externe ne sont pas dimensionnées, il y aura lecture de la variable dimensionnée dans l'enregistrement sélectionné, le pas servant à calculer la position des différentes occurrences (comme dans la fonction *saisie*).

Si une au moins des clés de la table externe est dimensionnée, alors le fonctionnement est différent pour les variables dimensionnées : chaque occurrence sera lue dans un enregistrement différent (défini par l'occurrence de la clé), et donc le pas éventuellement indiqué dans la fonction *import* ne sera pas utilisé, seule sera utilisée la position initiale. Une variable lue par *import* dans une table externe avec une clé dimensionnée doit être dimensionnée, et de même dimension que celle de la clé de la table externe.

3.5.4 Fonction ImportCsv.

Cette fonction permet de lire des variables dans une table externe au format csv.
Le caractère séparateur peut être au choix :

- espace
- tabulation
- ;
- ,

L'utilisateur doit définir ce caractère dans le menu *fichier/propriétés* :

L'écran de demande est analogue à celui de la fonction *import*, la position de lecture étant remplacée par un numéro de champ.

The screenshot shows the 'Fonction ImportCSV' dialog box. The title bar is blue with the text 'Fonction ImportCSV'. The main area is divided into several sections. On the left, there is a section for 'codage multi-réponses' with two radio buttons: 'Par valeurs (demande des nombres, ex : 0204...)' which is selected, and 'Par rangs (chaîne binaire, ex : 0101...)'. Below this is a section for 'Par valeurs' with two spinners: 'Nombre de valeurs' set to 1 and 'Pas des valeurs' set to 1. On the right side, there is a section for 'Enregistrements' with a 'Lien d'importation' button and a 'Numéro du champ' spinner set to 2. At the bottom right, there is a section for 'Pas des blancs à zéro'.

Tout le reste du fonctionnement est le même que celui de la fonction *import*.
Pour les variables multiples, voir la fonction *SaisieCsv*.

3.5.5 Arithm : Opération arithmétique.

Les opérandes peuvent être :

- des variables non dimensionnées, ou des occurrences de variables dimensionnées,
- des variables dimensionnées,
- des constantes.

On peut avoir au maximum 20 opérandes.

Opérations arithmétiques	
Formule	
Variable ou constante	Opéra...
WQAD	*
20054	+
1	

Les variables utilisées comme opérandes et comme variable résultat doivent être numériques (de type S ou Q).

Si la variable résultat est de type S, les valeurs des opérandes sont tronquées en entier, avant le calcul et le calcul est fait "en entier" (résultat tronqué).

Si la variable résultat est de type Q, les calculs sont effectués en réels. Si la variable résultat est dimensionnée, alors, toutes les variables opérandes doivent être de même dimension ou non dimensionnées (ou occurrences). Les calculs sont alors effectués pour chacune des occurrences.

On dispose des opérateurs * / + - et ^ (pour l'exponentiation).

Les priorités sont les suivantes : d'abord ^, puis * et /, enfin + et -.

En cas d'égalité de priorité, les opérations se font de la gauche vers la droite.

Le résultat du calcul sera REBUT dans les 3 cas suivants :

- division par zéro,
- au moins un opérande à REBUT,
- l'univers conditionnant la fonction n'est pas vérifié.

Tableau de contrôle de la fonction arithmétique :

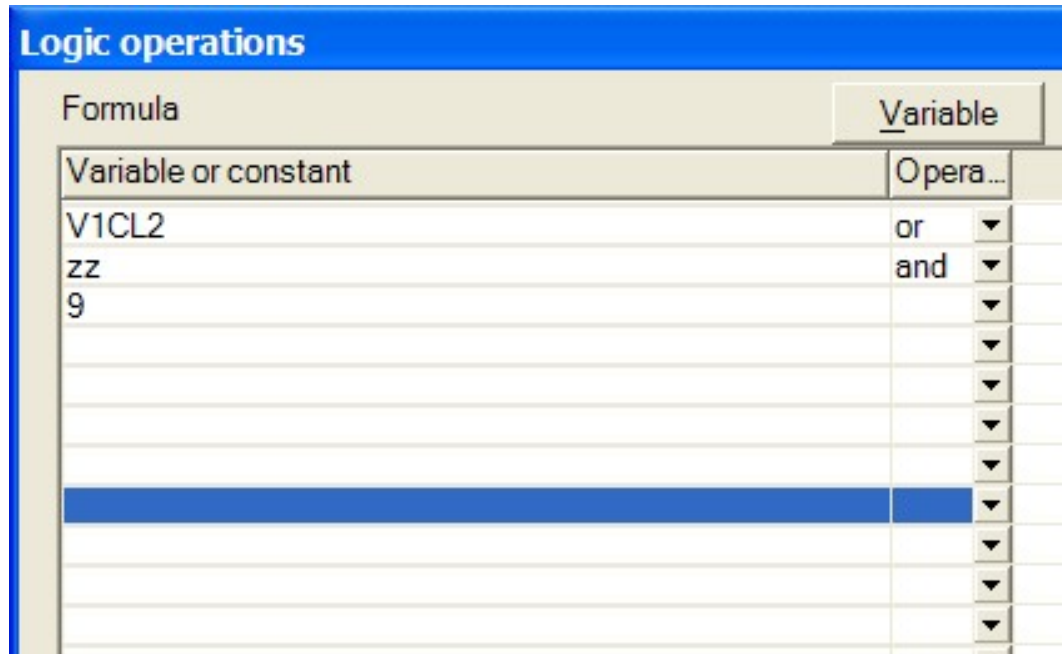
Contrôles	Var. résultat	Var. argument	Base
Type	S ou Q	S ou Q	L
Niveau	i	i ou <i	i ou <i
Dimension	non dimens.	non dimens.	non dimens.
ou	n	non dimens. ou dim n	non dimens. ou dim n

Si la base est dimensionnée, on ne considère que les occurrences de la variable **résultat** pour lesquelles l'occurrence correspondante de la base est *true*. (dans le cas contraire, le résultat de l'occurrence sera Rebut).

3.5.6 Logic : Opération logique.

Une variable multiple peut être construite par une opération logique ET, OU, EX (ou exclusif), SAUF à partir d'autres variables de type M ou S. On peut également faire intervenir des constantes entières.

On peut avoir au maximum 20 opérandes.



Par exemple :

Supposons que la 1ère variable argument : B1 vaut : 1, 2, 3, 4
et la 2ème variable B2 : 3, 4, 5, 6.

L'opération

B1 ou B2	donne	1, 2, 3, 4, 5, 6
B1 ex B2		1, 2, 5, 6
B1 et B2		3, 4
B1 sauf B2		1, 2

Remarques :

X et Rebut → Rebut
X ou Rebut → X
X ex Rebut → X
X sauf Rebut → X

Il n'y a pas de priorité entre les opérateurs, les opérations se font de gauche à droite :
B1 ou B2 et B3 est interprété : (B1 ou B2) et B3.

La constante Rebut est représentée par @NA.

Les variables opérandes doivent avoir la même dimension que la variable résultat, ou être des variables non dimensionnées, ou des occurrences.

Les valeurs des opérandes de type S sont converties en type multiple avant de faire l'opération ($0 \text{ à } n \rightarrow 0 \text{ à } n$, les autres valeurs donnent Rebut).

Tableau de contrôle de la fonction logique.

Contrôles	Var. résultat	Var. argument	Base
Type	M	M ou S	L
Niveau	i	i ou < i	i ou < i
Dimension	Non dimens	non dimens.	non dimens.
ou	n	non dimens. ou dim=n	non dimens. ou dim=n

Si la base est dimensionnée, on ne considère que les occurrences de la variable **résultat** pour lesquelles l'occurrence correspondante de la base est *true*. (dans le cas contraire, le résultat de l'occurrence sera Rebut).

3.5.7 Recod.

Cette fonction permet d'effectuer des **recodifications en groupes de valeurs**.

La variable résultat peut être de type Simple, Multiple ou Logique. Le cas d'une variable logique est un peu particulier et fait l'objet d'un paragraphe spécial.

L'écran de définition de la recodification comporte en haut un bouton pour sélectionner la (ou les) variable(s) argument(s), suivi d'une case à cocher "unique".

Fonction recod			
Sélectionner			
Code	Libellé	Variable source	Groupement
1	MALE	SEX "Sex "	1
2	FEMALE	SEX "Sex "	2
3	13 - 17 years old	AGER "Age of respo...	13 - 17
4	18 - 24 years old	AGER "Age of respo...	18 - 24
5	25 - 34 years old	AGER "Age of respo...	25 - 34
6	35 - 44 years old	AGER "Age of respo...	35 - 44
7	45 - 54 years old	AGER "Age of respo...	45 - 54
8	55 - 64 years old	AGER "Age of respo...	54 - 64
9	65 and over	AGER "Age of respo...	65 - 99

Si l'on veut utiliser des variables d'origine différentes selon les modalités de la variable résultat :

- sélectionnez d'abord la variable argument à utiliser pour la première modalité,
- décochez la case "unique" : les variables arguments s'afficheront dans une zone grisée "variable" pour chacune des modalités de la variable résultat, et l'on pourra sélectionner les variables en double cliquant dans cette zone grisée.

La colonne « label » donne la liste des libellés contenus dans le **format de la variable résultat, correspondant à un seul code** (donc les lignes du format correspondant à des sous-totaux ne seront pas reprises dans l'instruction Recod).

La colonne « recodification » sert à entrer les tranches de valeurs à prendre en compte ; ces valeurs sont des nombres réels dans le cas d'une variable argument de type quantité, de nombres entiers positifs ou nuls dans le cas d'une variable argument de type simple ou multiple, des chaînes de caractères dans le cas d'une variable argument de type texte, ou la valeur REBUT (représentée par @NA) dans tous les cas. La limite d'une tranche peut être INDEFINIE.

Sur une ligne, correspondant donc à un code de la variable résultat, on peut mettre une ou plusieurs valeurs ou tranches de valeurs.

Il y a trois façons de saisir les tranches de valeurs :

- **en tapant directement les bornes** de tranches en face du libellé correspondant, il faut alors respecter la syntaxe suivante :
 - 0-44 signifie de 0 à 44,
 - 18-24 ; 35-44 signifie de 18 à 24 ou de 35 à 44,
 - 1-3 ; @NA signifie de 1 à 3 ou non renseigné,
 - 45- signifie 45 ou plus,
 - -44 signifie 44 ou moins
 - A-D signifie depuis le caractère ASCII A majuscule jusqu'au caractère D majuscule, selon l'ordre du code ASCII.

- en actionnant le bouton **F5 par valeurs**, ou la touche F5 ; dans ce cas, pour chaque tranche correspondant à une modalité de la variable résultat :
 - entrez dans la colonne de gauche la borne inférieure de la tranche, et dans la colonne de droite sa borne supérieure ; on dispose d'un bouton en haut à gauche pour saisir la valeur REBUT (on peut aussi taper F8), et d'un autre à droite pour valeur INDEFINIE, c'est à dire "ou plus", ou "ou moins", on peut aussi taper F7,
 - puis appuyez sur le bouton OK et se positionner sur la modalité suivante, appuyer de nouveau sur le bouton F5 - par valeurs, ou la touche F5, et ainsi de suite...

- uniquement si la variable argument a un format, en actionnant le bouton **F6 par liste** : on peut alors cliquer sur les modalités à prendre en compte pour une modalité de la variable recodifiée, appuyer sur OK, se positionner sur la modalité suivante, et ainsi de suite...**Attention** : avec cette manière, seules les modalités correspondant à un seul code sont accessibles (celles correspondant à des sous-totaux ne sont pas montrées).

On peut par un clic droit appeler un menu qui permet de copier / coller des valeurs.

Si la variable **argument est de type texte** :

- ne pas encadrer les valeurs que l'on saisit par des guillemets ; ceux-ci seront générés pour l'affichage des valeurs de ce type de variables, pour rappeler qu'il s'agit de variables texte.
- l'ordre des caractères est l'ordre ASCII : de 0 à 9, puis de A à Z, puis de a à z.

Si la variable **argument est de type logique** : les valeurs à faire figurer sont 0 et 1 (pour *false* et *true*).

Dans le cas où la **variable résultat est dimensionnée** :

- si la variable argument l'est aussi (de même dimension), la recodification se fait occurrence par occurrence,
- si la variable argument n'est pas dimensionnée, ou si c'est une occurrence, la recodification se fait d'après les valeurs de cette variable non dimensionnée, pour toutes les occurrences de la variable résultat ; cette règle s'applique pour chacune des modalités de la variable résultat.

Tableau de contrôle de la fonction Recod :

Contrôles	Var. résultat	Var. argument	Base
Type	M, S ou L	M, S, Q, T, L	L
Niveau	i	i ou < i	i ou < i
Dimension	non dimens.	non dimens.	non dimens.
ou	n	non dimens ou dim n	non dimens. ou dim n

Exécution.

Si la variable résultat est de type simple, elle prend comme valeur le premier code pour laquelle la condition définie par le couple (variable, tranches de valeurs) est vérifiée.

Si la variable résultat est multiple, elle prend comme valeurs tous les codes pour lesquels la condition est vérifiée.

ATTENTION : pour une variable **argument de type texte**, si la comparaison porte sur des chaînes de caractères de longueurs différentes, elle sera faite par convention en se limitant au nombre de caractères de la plus petite des 2 chaînes à comparer.

Cas particulier des variables résultats logiques.

Quand la variable résultat est logique, le fonctionnement est un peu différent, compte-tenu de ce qu'une variable logique n'a pas de format.

Le fonctionnement est alors le suivant : une seule ligne sera proposée, et la valeur *true* sera écrite automatiquement sous le champ *label*.

On peut alors insérer autant de lignes que l'on veut, elles auront toutes le même libellé *true* sous le champ *label*.

Les autres parties de l'écran (variables, recodification) seront sans changement.

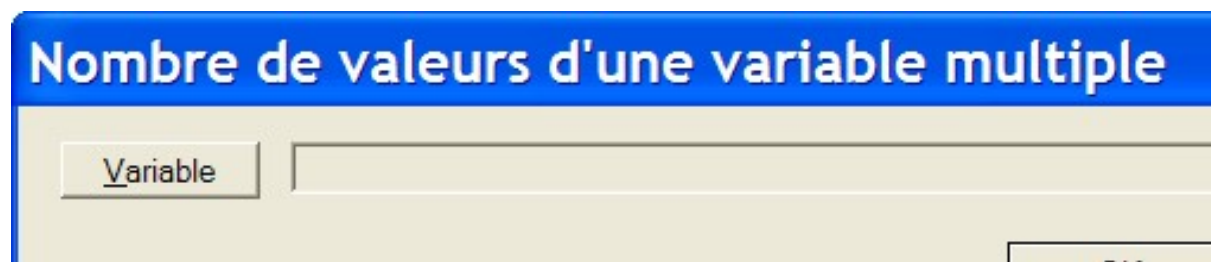
L'exécution se fera ainsi : la variable aura comme résultat la valeur *true* **si toutes les lignes** ainsi décrites sont vérifiées, sinon elle aura la valeur *false*. Exemple :

Fonction recod

Code	Label	Source Variable	Grouping
1	true	SEX "Sex "	1
1	true	AGER "Age of respo...	15 - 40

La variable résultat sera *true* pour les femmes de moins de 50 ans.

3.5.8 Nbrep : nombre de valeurs d'une variable multiple



- La variable résultat doit être de type simple, la variable argument multiple.
- On compte le nombre de codes de la variable multiple, sans se préoccuper des codes éventuellement définis dans le format associé.
- si la base est dimensionnée, on ne renseigne que les occurrences de la variable argument pour lesquelles l'occurrence correspondante de la base est *true*.
- le résultat final est donc un nombre de 0 à x si la base est *true*, et Rebut si la base est *false*.

Tableau de contrôle de la fonction Nbrep :

Contrôles	Var. résultat	Var. argument	Base
Type	S	M	L
Niveau	i	i	i ou <i
Dimension	non dimens.	non dimens.	non dimens.
ou	n	non dimens ou dim=n	non dimens ou dim=n

3.5.9 Nbval : Nombre de valeurs d'une variable dimensionnée.

Il s'agit du nombre d'occurrences différentes de Rebut (ou de *false* dans le cas d'une variable logique).

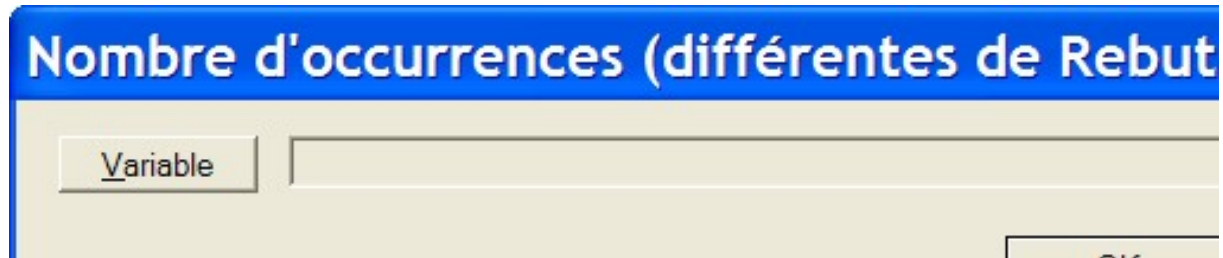


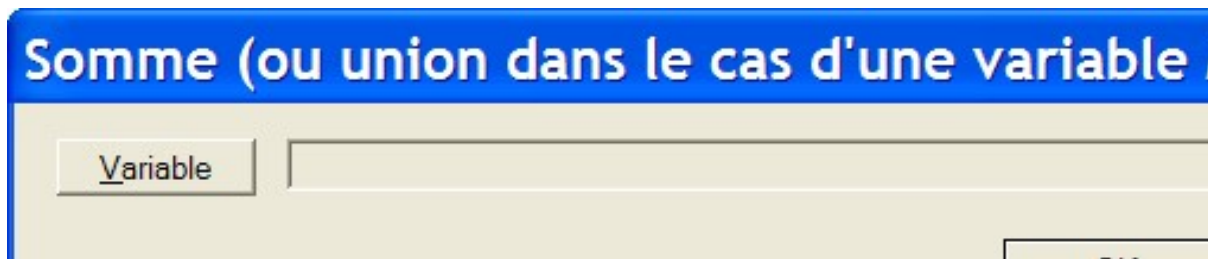
Tableau de contrôle de la fonction Nbval :

Contrôles	Var. résultat	Var. argument	Base
Type	S	M, S, Q, T, L	L
Niveau	i	i	i ou < i
Dimension	non dimens.	Dimension = n	non dimens. ou dim=n

Exécution :

- le résultat final est donc un nombre de 0 à x si la base est *true*, et Rebut si la base est *false*.
- si la base est dimensionnée, on ne considère que les occurrences de la variable Argument pour lesquelles l'occurrence correspondante de la base est *true*. Le résultat est Rebut si toutes les occurrences de la base sont *false*.

3.5.10 Somval : Somme des valeurs d'une variable dimensionnée.



Il y a 2 fonctionnements différents selon que la variable résultat est de type simple ou quantité d'une part, multiple d'autre part :

- dans le cas d'une variable résultat simple ou quantité, on fait la somme des différentes occurrences de la variable argument.
- Dans le cas d'une variable résultat de type multiple, on fait l'union des différentes occurrences de la variable argument.

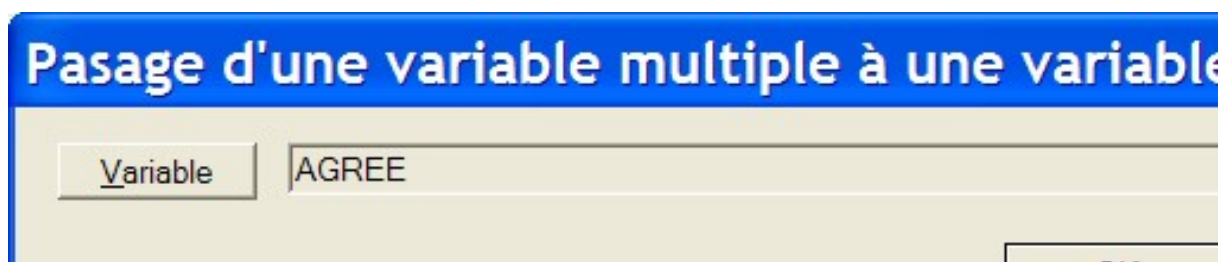
Tableau de contrôle de la fonction Somval :

Contrôles	Var. résultat	Var. argument	Base
Type	M	M ou S	L
type	S ou Q	S ou Q	L
Niveau	i	i	i ou < i
Dimension	non dimens.	Dimension =n	non dimens. ou dim=n

Exécution :

- Les conversions nécessaires seront faites le cas échéant avant de faire la sommation (simple en multiple, quantité en simple, etc..) : c'est la variable résultat qui impose son type.
- Si la variable résultat est de type S ou Q, et si une des occurrences de la variable argument vaut REBUT, la variable résultat vaudra aussi REBUT.
- le résultat final est donc un nombre de 0 à x si la base est *true*, et Rebut si la base est *false*.
- si la base est dimensionnée, on ne considère que les occurrences de la variable Argument pour lesquelles l'occurrence correspondante de la base est *true*. Le résultat est Rebut si toutes les occurrences de la base sont *false*.

3.5.11 Dichol : Passage d'une variable de type multiple à une variable dimensionnée



On accepte aussi les variables argument de type simple.

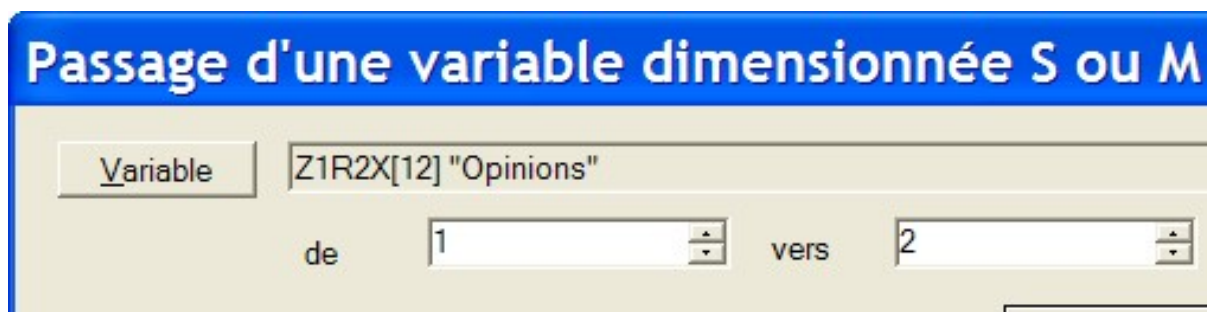
Tableau de contrôle de la fonction Dichol :

Contrôles	Var. résultat	Var. argument	Base
Type	S	M	L
Niveau	i	i	i ou < i
Dimension	Dimension = n	non dimens.	non dimens. ou dim=n

Exécution :

- les codes de la variable argument à considérer sont les codes 1 à n, n étant la dimension de la variable résultat (les codes 0, -1, et -2 sont ignorés).
- si la variable argument contient le code j, l'occurrence j de la variable résultat vaudra 1, Rebut sinon
- si la base est dimensionnée, on ne renseigne que les occurrences de la variable résultat pour lesquelles l'occurrence correspondante de la base est *true*.

3.5.12 Dicho2 : Passage d'une variable dimensionnée à une variable de type multiple.



Outre la variable argument, il faut renseigner les bornes de cette variable à considérer, seules les valeurs des occurrences comprises entre ces bornes génèrent une valeur de code de la variable multiple, correspondant au rang de l'occurrence dont la valeur est comprise entre les bornes.

Tableau de contrôle de la fonction Dicho2 :

Contrôles	Var. résultat	Var. argument	Base
Type	M	S, M	L
Niveau	i	i	i ou < i
Dimension	non dimens.	Dimension = n	non dimens. ou dim=n

Exécution :

- si l'occurrence j de la variable argument a une valeur comprise entre les bornes indiquées, la variable résultat aura le code j.
- Si la base est dimensionnée, on ne considère que les occurrences de la variable argument pour lesquelles l'occurrence correspondante de la base est *true*.

3.5.13 NbNiv : Nombre de valeurs d'une variable de niveau inférieur.

Cette fonction est réservée aux études à niveaux.

On obtient le nombre d'enfants de l'individu statistique en cours de traitement pour lesquels la variable argument est différente de Rebut.

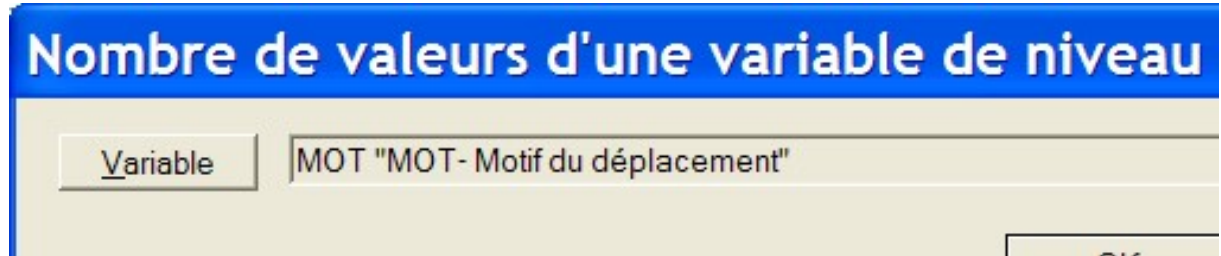


Tableau de contrôle de la fonction Nbniv :

Contrôles	Var. résultat	Var. argument	Base
Type	S	M, S, Q, T, L	L
Niveau	i	j > i et j descendant de i	j ou < j
Dimension	non dimens.	non dimens.	non dimens.
	Dimension = n	Dimension = n	Dimension =n ou non dimens.

Exécution :

- La notion de descendant résulte de la structure de l'étude (écran de définition des niveaux)
- le résultat final est un nombre de 0 à x si la base est *true*, et Rebut si la base est *false*.
- Si la base est dimensionnée, on ne considère que les occurrences de la variable argument pour lesquelles l'occurrence correspondante de la base est *true*.

3.5.14 SomNiv : Somme des valeurs d'une variable d'un niveau inférieur.

Cette fonction est réservée aux études à niveaux.

On cumule les valeurs de la variable argument pour tous les enfants de l'individu statistique en cours de traitement.

Selon que la variable résultat est arithmétique (de type S ou Q) ou multiple (de type M), on fera un cumul arithmétique ou logique (union) de la variable origine de niveau fils, vers la variable résultat de niveau père.

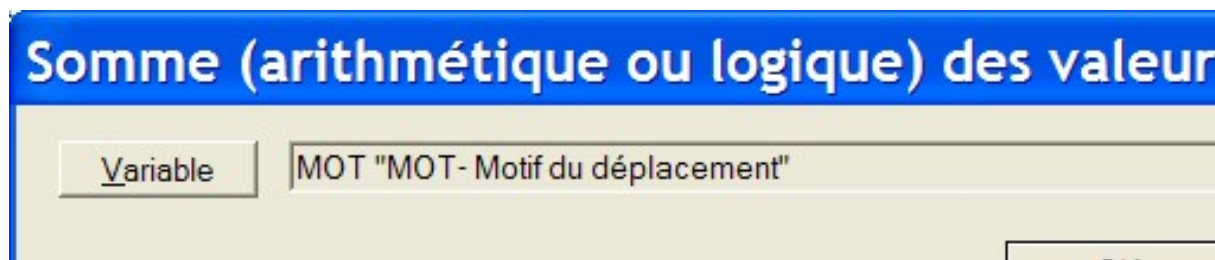


Tableau de contrôle de la fonction Somniv :

Contrôles	Var. résultat	Var. argument	Base
Type	M	M, S	L
ou	S	S, Q	L
ou	Q	S, Q	L
Niveau	i	j > i et j descendant de i	j ou < j
Dimension	non dimens.	non dimens.	non dimens.
ou	Dimension = n	Dimension = n	Dimension = n ou non dimens.

Exécution :

- Si la variable résultat est arithmétique et si la variable argument prend au moins une fois la valeur REBUT , le résultat sera REBUT
- Les conversions nécessaires seront faites automatiquement avant cumul.
- le résultat final est donc un nombre de 0 à x si la base est *true*, et Rebut si la base est *false*
- Si la base est dimensionnée, on ne considère que les occurrences de la variable argument pour lesquelles l'occurrence correspondante de la base est *true*.

3.5.15 Const : Table de constantes.

Cette fonction permet de constituer une table de constantes sous la forme d'une variable **dimensionnée ou non**. Il ne doit pas y avoir de Base (elle sera ignorée).

Le type de la variable résultat peut être M, S, Q ou T.

Pour les multiples, les différentes valeurs doivent être séparées par des virgules.

Définition des valeurs des constantes	
Constantes	
#	Valeur de la constante
1	22,55,8

Définition des valeurs des constantes	
Constantes	
#	Valeur de la constante
1	1
2	2
3	3
4	4
5	5
6	6

3.5.16 Table : Sélection d'une valeur dans une table au moyen d'une variable-indice.

La variable résultat prend comme valeur celle donnée par une variable dimensionnée (premier argument) indexée par la variable indice (2^{ème} argument).

RESULTAT = TABLE [Index]

The screenshot shows a dialog box titled "Sélection d'une valeur dans une table au moyen d'". It contains two input fields. The first field, labeled "Table variable", contains the text "Z1R2X[12] \"Opinions\"". The second field, labeled "Index variable", contains the text "XX".

Tableau de contrôle de la fonction Table :

Contrôles	Var. résultat	Var. table	Var. indice	Base
Type	M	M, S	S	L
ou	S	S, Q	S	L
ou	Q	S, Q	S	L
ou	T	T	S	L
Niveau	i	i ou < i	i ou < i	i ou < i
Dimension	non dimens.	dimensionnée	non dimens.	non dimens.
ou	Dimension = n	dimensionnée	Dimension = n	Dimension = n ou non dimensionnée

Exécution :

- les conversions nécessaires seront faites automatiquement
- dans le cas où la variable résultat est dimensionnée, sa dimension n n'a pas à être forcément égale à la dimension de la variable table : on a

RES [j] = TABLE [INDICE (j)]

- Si la base est dimensionnée, on ne considère que les occurrences de la variable résultat pour lesquelles l'occurrence correspondante de la base est *true*.
- le résultat final est Rebut si la base est *false*

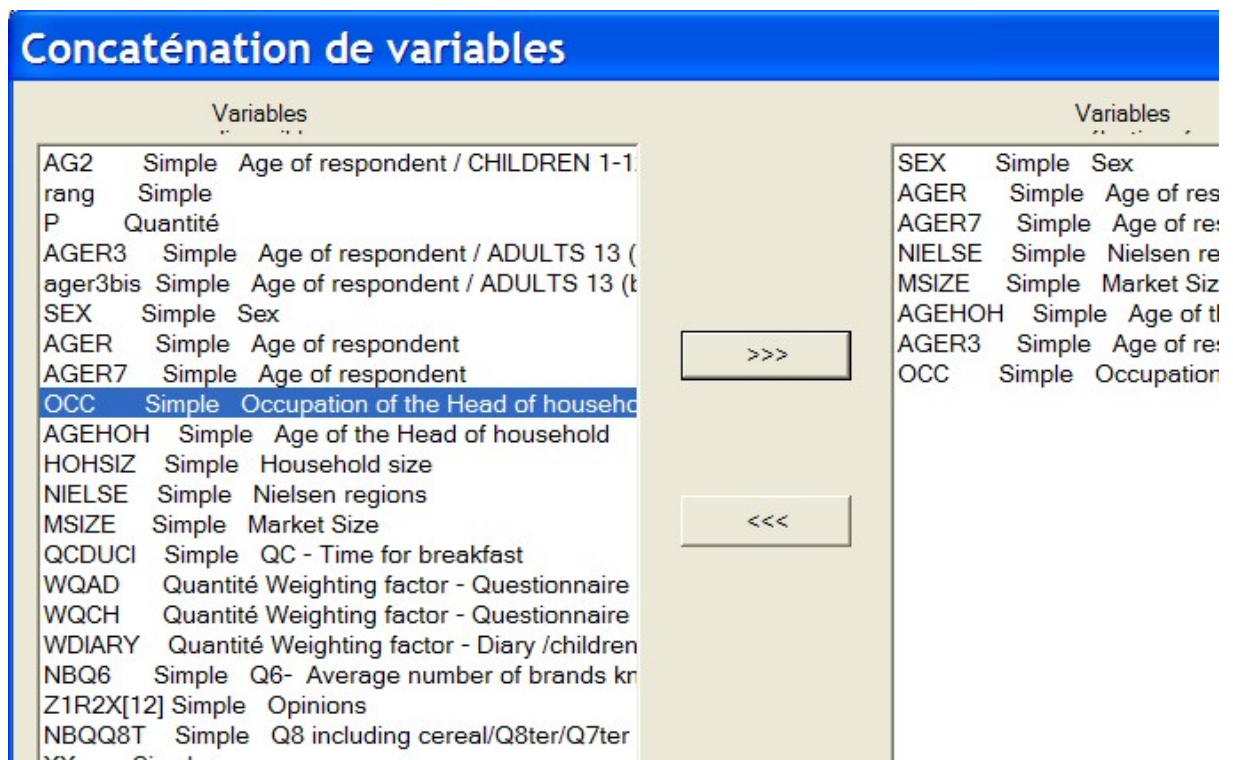
3.5.17 Concat : concaténation de variables.

Une variable dimensionnée peut être constituée par concaténation de plusieurs variables dimensionnées ou non. La variable résultat doit avoir pour dimension la somme des dimensions des variables d'origine.

Une variable de type multiple peut être obtenue par concaténation de variables de type multiple et simple.

Une variable de type simple ou quantité peut être obtenue par concaténation de variables de type simple ou quantité.

Une variable de type texte peut être obtenue par concaténation de variables de type texte.



Attention : pour faire apparaître et sélectionner des occurrences de variables dimensionnées, il faut se positionner sur la variable dimensionnée choisie et appuyer sur la barre espace, cela fera apparaître les occurrences que l'on pourra alors sélectionner.

Tableau de contrôle de la fonction Concat :

Contrôles	Var. résultat	Var. argument	Base
Type	M	M, S	L
ou	S	S, Q	L
ou	Q	S, Q	L
ou	T	T	L
Niveau	i	i ou < i	i ou < i
Dimension	Dimension = n	Dimensionnées ou non, somme des dimensions=n	Dimension = n ou non dimens.

Si on le désire, les titres des variables concaténées deviendront les titres d'occurrences de la variable résultat (option : mettre à jour les titres d'occurrences).

Remarque : dans la partie droite de l'écran, on peut déplacer les variables sélectionnées avec les touches ctrl + flèches

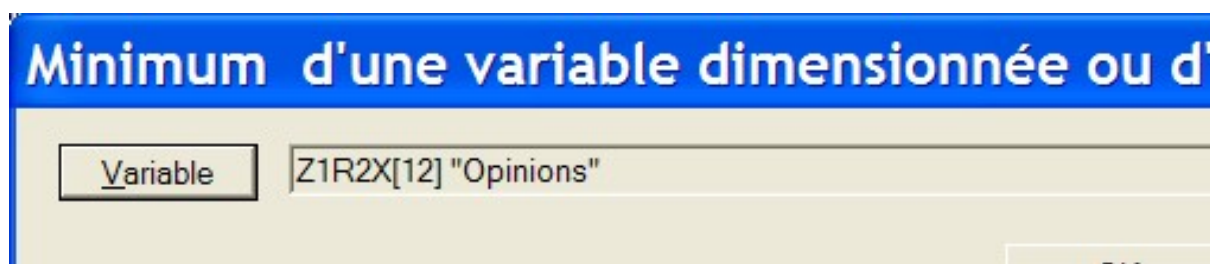
Exécution :

- les conversions sont faites le cas échéant
- Si la base est dimensionnée, on ne considère que les occurrences de la variable résultat pour lesquelles l'occurrence correspondante de la base est *true* (les autres sont en rebut).
- le résultat final est Rebut si la base est *false*

3.5.18 Min et Max : Minimum et maximum d'une liste de variables.

Les fonctions Min et MAX s'appliquent :

- soit à des variables dimensionnées pour rechercher la valeur MINIMUM ou MAXIMUM entre les différentes occurrences,
- soit dans le cas avec niveaux, pour rechercher le MINIMUM ou MAXIMUM d'une variable d'un niveau fils.



En cliquant sur le bouton Variable, on obtient un écran pour choisir la variable dont on veut calculer le minimum ou le maximum (variable à occurrences, ou d'un niveau inférieur) : le programme ne propose que les seules variables candidates en fonction du tableau de contrôle.

Tableaux de contrôle de la fonction Min ou Max :

1er cas : minimum ou maximum entre les occurrences (s'applique si les niveaux des variables résultat et argument sont égaux, ou si l'étude est sans niveau)

Contrôles	Var. résultat	Var. argument	Base
Type	S	S ou Q	L
Niveau	i	i	i ou < i
Dimension	non dimens.	Dimension =n	dim = n ou non dimens.

2ème cas : minimum ou maximum entre niveaux (s'applique s'il y a des niveaux et si le niveau j de la variable argument est > niveau de la variable résultat).

Contrôles	Var. résultat	Var. argument	Base
Type	S	S, Q	L
Niveau	i	j > i et j descendant de i	i ou < i
Dimension	non dimens.	non dimens.	non dimens.
ou	Dimension =n	Dimension =n	non dimens. Ou dim=n

Exécution (dans les 2 cas) :

- les conversions sont faites le cas échéant
- par convention, la valeur `Rebut` est considérée comme la plus petite valeur
- Si la base est dimensionnée, on ne considère que les occurrences de la variable argument pour lesquelles l'occurrence correspondante de la base est *true*.
- le résultat final est `Rebut` si la base est *false*

3.5.19 Rang et RangNF : rang d'enregistrement.

Il s'agit du rang (numéro séquentiel) de l'unité statistique traité. Dans le cas Rang, ce rang sera incrémenté pour les seuls individus appartenant à l'univers souhaité, c'est à dire ayant la valeur *true* pour la variable logique éventuellement spécifiée dans l'univers de la structure d'entrée/sortie, alors que dans le cas RangNF, le rang est incrémenté pour tous les individus, appartenant ou non à l'univers souhaité.

ATTENTION : s'il y a plusieurs fichiers de données, le rang est défini fichier par fichier, et non globalement, et présente donc des doublons.

Dans le cas d'une étude à niveau, le rang d'un individu de niveau $j > 1$ est son rang dans sa fratrie.

Il n'y a pas de nouvel écran pour cette fonction (aucun argument).

Tableau de contrôle de la fonction Rang ou RangNF.

Contrôles	Var. résultat	Base
Type	S	L
Niveau	i	i ou < i
Dimension	non dimens.	non dimens.

Exécution :

Cas sans niveaux :

La variable renseignée par cette fonction aura pour valeur le rang de l'individu dans le traitement (numéro séquentiel éventuellement limité aux individus filtrés).

Cas avec niveaux :

Si la variable à renseigner est de niveau 1, elle aura pour valeur le rang de l'individu de niveau 1 dans l'ensemble du fichier.

Si la variable à renseigner est d'un autre niveau, les individus ayant le même père ont un rang repartant de la valeur 1.

ATTENTION : une variable construite avec la fonction Rang ne doit pas être utilisée comme argument dans la fonction d'une autre variable.

3.5.20 GetNiv : obtention du niveau.

Cette fonction rend comme valeur le niveau en cours. On l'utilise en général avec une variable de niveau 0.

Il n'y a pas de nouvel écran pour cette fonction (aucun argument).

Tableau de contrôle de la fonction GetNiv :

Contrôles	Var. résultat	Base
Type	S	L
Niveau	i	i ou < i
Dimension	non dimens.	non dimens.

3.5.21 Conv : conversions.

Ces fonctions permettent d'opérer des conversions entre différents types de valeurs.

Tableau de contrôle de la fonction Conv :

Contrôles	Var. résultat	Var. argument	Base
Type	S, T, L	M, S, Q, T	L
Niveau	i	i	i ou < i
Dimension	non dimens.	non dimens.	non dimens
ou	Dimension = n	non dimens. ou dim=n	non dimens. ou dim=n

Exécution :

3.5.21.1. Variable résultat de type texte :

D'une manière générale, variable argument = rebut entraîne variable résultat = rebut

Simple vers texte.

Conversion d'un type à un autre. Simple vers

Variable: AGER3 "Age of respondent / ADULTS 13 (by age brackets)"

Format: "0013" " 13"

A droite A gauche

Longueur d'une nouvelle variable: 1

Les valeurs pourront être cadrées à droite avec des blancs ou des zéros à gauche, ou cadrées à gauche.

Attention : si la longueur prévue est insuffisante pour la valeur considérée, le résultat sera Rebut.

Quantité vers texte :

Conversion d'un type à un autre. Quantité vers

Variable: P

Format: "0013.5" " 13.5"

A droite A gauche

Longueur d'une nouvelle variable: 1

Même chose que dans le cas précédent, mais il faut donner également le nombre de décimales.

Multiple vers texte :

The screenshot shows a dialog box titled "Conversion d'un type à un autre. Multiple vers". It has a "Variable" field containing "AGER7BIS \"Sex and Age of respondent \"". Below this, there are two radio buttons for "Format": "Par valeurs" (which is selected) and "Par rangs". At the bottom, there is a field for "Longueur des valeurs de la nouvelle" with the value "1".

Les deux formats proposés sont :

- **Par valeurs** (en format fixe) : on spécifie la longueur à utiliser pour les valeurs.
- **Par rang** (suite de 0 et de 1) : on spécifie la longueur totale de la variable résultat, c'est-à-dire le nombre de valeurs à prendre en compte.
Attention : la 1^{ère} position (à gauche) correspond à **la valeur 0 de la variable** argument, la 2^{ème} position à la valeur 1, etc..C'est à dire que si on spécifie une longueur totale « n », les valeurs éditées correspondront aux valeurs 0, 1, ... n-1 (la valeur 1 dans la position j signifiant que la variable argument a la valeur j-1).

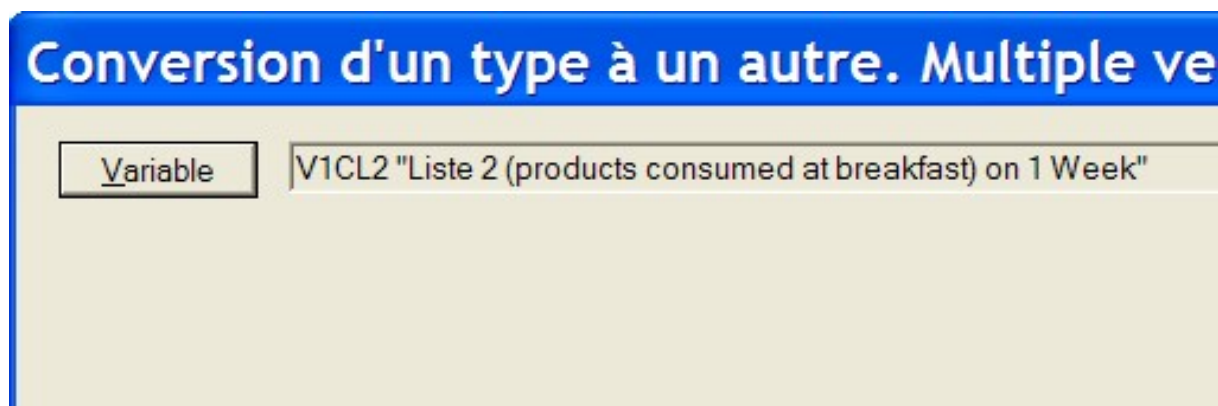
3.5.21.2. Variable résultat de type logique.

Pas d'option ; quel que soit le type de la variable argument, le résultat est *false* si la variable argument est rebut, *true* dans le cas contraire.

This is an identical screenshot to the one above, showing the same dialog box with the variable "AGER7BIS \"Sex and Age of respondent \"", the "Par valeurs" format selected, and the length set to 1.

3.5.21.3 Variable résultat de type simple.

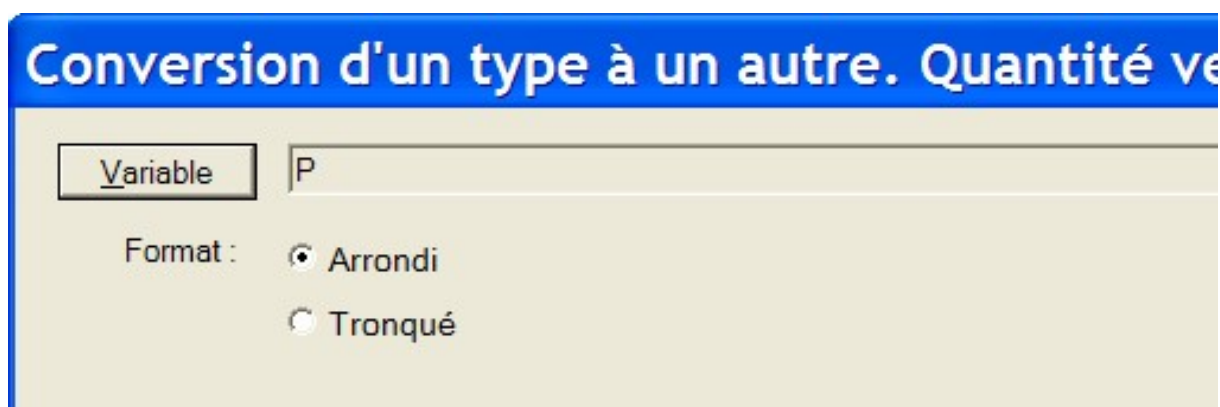
D'une manière générale, si variable argument = rebut → variable résultat = rebut

Multiple vers simple.

Il n'y a pas d'option : le résultat est un nombre obtenu en sommant les puissances de 2 des valeurs de la variable argument

0 → 1
 1 → 2
 2 → 4
 3 → 8
 4 → 16
 5 → 32, etc..

Exemple : si la variable argument a les valeurs (1,2,5), le résultat sera 38.
 L'intérêt de cette conversion est qu'elle fait apparaître **les combinaisons possibles** des valeurs prises par la variable argument.

Quantité vers simple.

On a le choix entre arrondi ou troncature.

Texte vers simple.

Il n'y a pas d'option. Un nombre entier, éventuellement signé, avec éventuellement des blancs à droite ou à gauche : **[bb][S]nnn[bbb]** donne la valeur **nnn**, avec son signe éventuel, toute autre chose donne REBUT

3.5.22 Fonctions de traitement des variables temporelles.

Pour toutes ces fonctions, le temps est exprimé par une variable de type texte de la forme:
AAAAMMJJ : année, mois, jour (soit 8 caractères), ou
AAAAMMJJHHMMSS : année, mois, jour, heure, minute, seconde (soit 14 caractères)
(Selon les fonctions cette variable de type texte intervient en argument ou en résultat).

3.5.22.1 Fonctions DateCour et HeureCour

Pour ces fonctions, il n'y a pas d'écran (pas d'argument).
La variable résultat doit être de **type texte, non dimensionnée**.
Le résultat sera de longueur 8, de la forme AAAAMMJJ dans le cas DateCour, et de longueur 6, de la forme HHMMSS dans le cas HeureCour.

3.5.22.2 Année, Mois, Trimestre, JourSem, Quantième, NoSem

La variable résultat doit être de type simple, non dimensionnée. La variable argument doit être de **type texte, non dimensionnée**, de la forme décrite ci-dessus.
Ces fonctions restituent respectivement l'année, le mois, le trimestre, le jour de la semaine, le quantième et le numéro de semaine.
Dans le cas Mois et JourSem, un format sera créé automatiquement, contenant les mois de Janvier à Décembre, et les jours de Lundi à Dimanche

Retourne l'année d'une variable temporelle

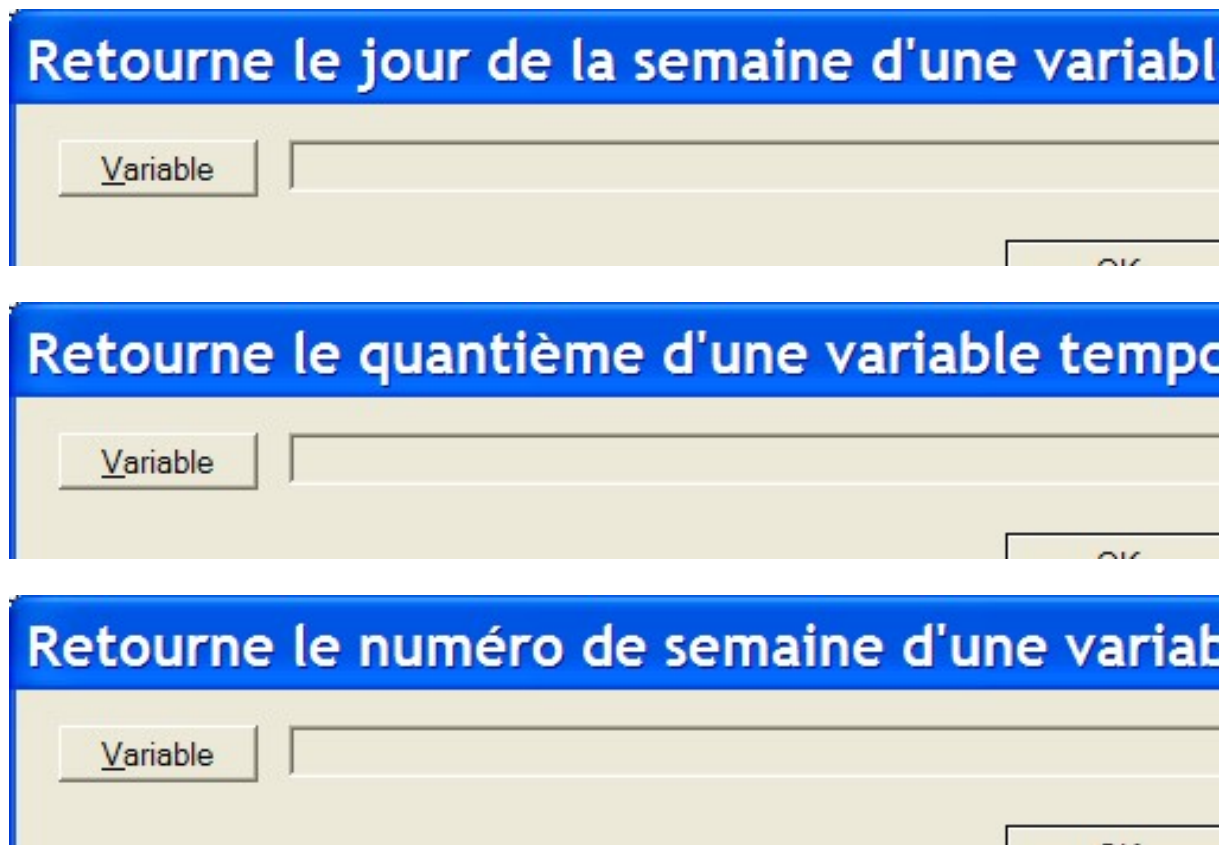
Variable

Retourne le mois d'une variable temporelle

Variable

Retourne le trimestre d'une variable temporelle

Variable



La fonction NOSem retourne le nombre (année *100 + numéro de semaine), sous la forme 200614.

Tableau de contrôle :

Contrôles	Var. résultat	Var. argument	Base
Type	S	T	L
Niveau	i	i ou < i	i ou < i
Dimension	non dimens.	non dimens.	non dimens

3.5.22.3. DiffTemps : différence de dates.

On spécifie deux dates (sous la forme de variables littérales ou de constantes), et l'unité dans laquelle on veut obtenir le résultat (en années, en mois, jour, heures, minutes, secondes).

La variable résultat doit être de type simple, non dimensionnée.

Les variables arguments doivent être de type texte, non dimensionnées, de la forme décrite ci-dessus

Tableau de contrôle :

Contrôles	Var. résultat	Var. argument	Base
Type	S	T	L
Niveau	i	i ou < i	i ou < i
Dimension	non dimens.	non dimens.	non dimens

3.5.22.4. AjoutTemps : ajouter une certaine durée à une date.

On spécifie la date d'origine (constante ou variable texte non dimensionnée), le nombre à ajouter (qui peut être une constante ou une variable Simple, de valeur positive ou négative) et l'unité dans laquelle ce nombre est exprimé.

The screenshot shows a software interface for calculating a new value for a variable. It is divided into two main sections: 'Temps' and 'Ajouter une durée'. Each section contains a 'Variable' input field and a 'Constante' input field. In the 'Temps' section, the 'Constante' field is populated with a date and time: 'samedi 26 mai 2007' and '12:56:32'. In the 'Ajouter une durée' section, the 'Constante' field is set to '0' and has a unit selection dropdown menu.

Le résultat est une date (variable texte de longueur 8, ou 14 selon celle de la date d'origine).

Tableau de contrôle :

Contrôles	Var. résultat	Var. argument	Base
Type	T	T et S	L
Niveau	i	i ou < i	i ou < i
Dimension	non dimens.	non dimens.	non dimens

3.5.23 Chaîne. Concaténation de chaînes de caractères.

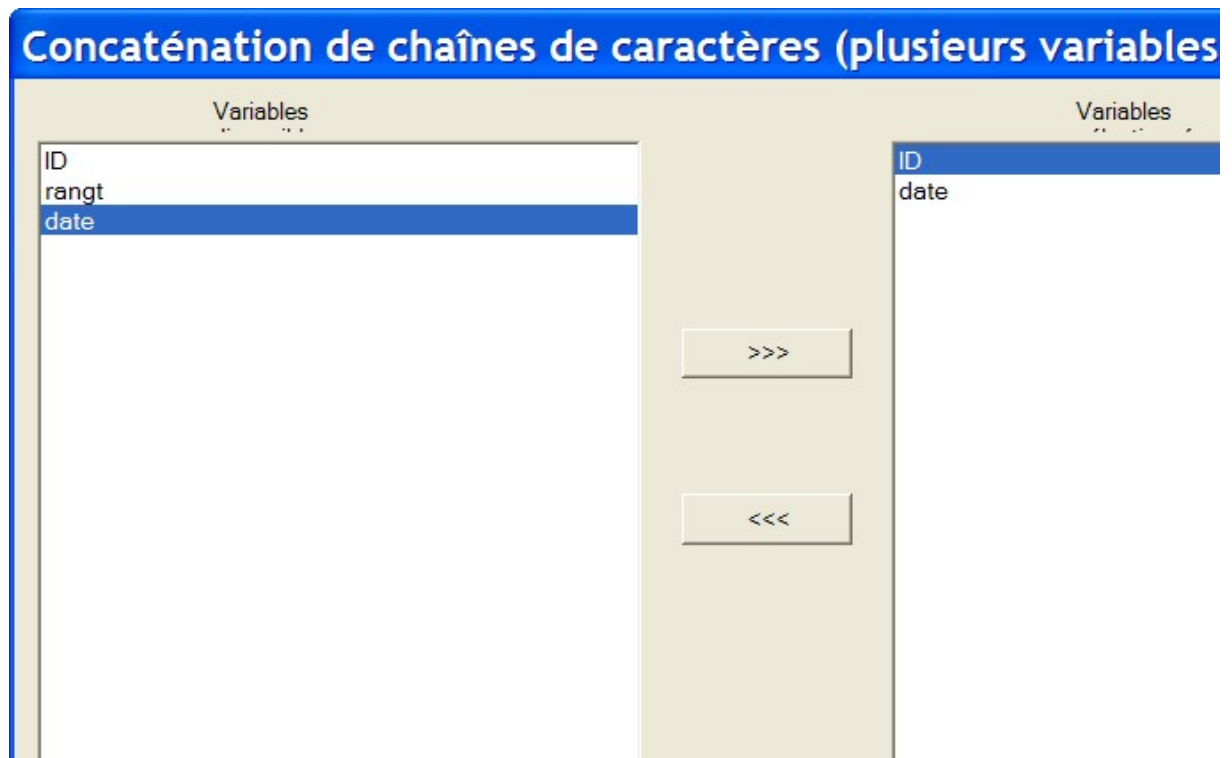


Tableau de contrôle :

Contrôles	Var. résultat	Var. argument	Base
Type	T	T	L
Niveau	i	i ou < i	i ou < i
Dimension	non dimens.	non dimens ou dimension =n	non dimens

Exécution :

- La longueur de la variable résultat est prise égale à **la somme des longueurs** de chaque élément à concaténer.
- Les occurrences éventuelles des variables à concaténer seront mises bout à bout,
- le résultat final est Rebut si la base est *false*

3.5.24 ExtrChn : Extraire une sous-chaîne à partir d'une variable texte.

On indique la position de départ, ainsi que la longueur de la sous-chaîne à extraire dans la variable origine :

- soit en entrant sa valeur,
- soit en spécifiant une variable simple qui porte cette valeur.

Tableau de contrôle :

Contrôles	var. résultat	Var.origine	Var position et longueur	Base
Type	T	T	S	L
Niveau	i	i ou < i	i ou < i	i ou < i
Dimension	non dimens.	Non dimens.	non dimens	non dimens.
ou	dim=n	dim=n	non dimens	non dimens.ou dim=n

Exécution :

- les caractères commençant à la position indiquée de la variable origine sont transférés dans la variable résultat, pour la longueur totale donnée.
- la longueur de la sous-chaîne à extraire sera limitée éventuellement par la longueur de la chaîne origine.

3.5.25 PosChn :

Position de la première apparition d'une sous-chaîne à l'intérieur d'une chaîne de caractères.

La chaîne et la sous-chaîne sont données comme des variables ou des constantes.

Tableau de contrôle :

Contrôles	var. résultat	Var.argument1	Var.argument2	Base
Type	S	T	T	L
Niveau	i	i ou < i	i ou < i	i ou < i
Dimension	non dimens.	Non dimens.	Non dimens.	non dimens.
ou	dim=n	dim=n	Non dimens.	non dimens.ou dim=n

3.5.26 TmutS : transformation d'une variable Texte en une variable Simple.

Cette nouvelle fonction a pour but de donner ses valeurs à une variable de type S à partir d'une variable de type Texte, ces valeurs 1 à n correspondant aux n lignes du format de la variable texte :

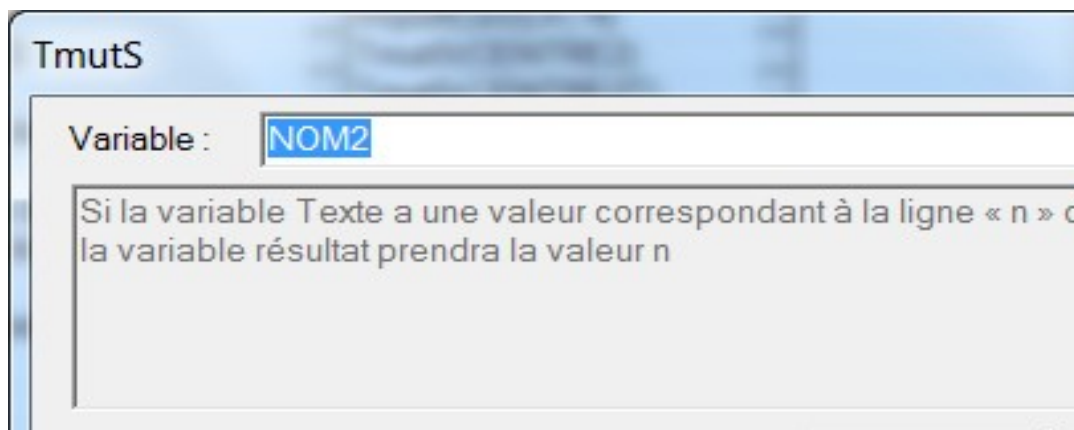


Tableau de contrôle de la fonction TmutS :

Contrôles	Var. résultat	Var. argument	Base
Type	S	T avec format	L
Niveau	i	i	i ou <i
Dimension	non dimens.	non dimens.	non dimens.
ou	n	n	non dimens. ou dim n

La variable argument doit être de type texte, avec un format. La variable résultat de type Simple, prendra comme la valeur n si pour l'individu en cours de traitement, la variable Texte a la valeur correspondant à la ligne n de son format.

Remarques :

- 1) Il faut que le format de la variable-argument Texte ne contienne que des valeurs uniques (pas de sous-totaux), sous peine de résultats inadéquats.
- 2) Il est bien entendu utile de donner un format à la variable-résultat Simple, avec les codes 1 à n dans les lignes 1 à n, et en recopiant les n libellés de la variable-argument Texte.
- 3) L'usage de TmutS, permet de travailler sur une variable Simple plutôt que sur une variable Texte, ce qui présente des avantages appréciables : sous-totaux dans les formats plus aisés, recodifications plus simples et plus riches, gain de temps, etc.
Il est donc recommandé de l'utiliser systématiquement, sauf si l'usage de la variable Texte se limite à une simple tabulation, sans recodification aucune.

3.6 Tables externes.

3.6.1 Principe.

Outre la saisie directe sur le fichier des données, on peut renseigner des variables au moyen de tables externes.

Ces tables externes sont des fichiers ASCII, avec un seul type d'enregistrements, en format standard ou csv.

Dans le cas standard, les enregistrements doivent être de longueur fixe.

Dans le cas csv, le caractère séparateur peut être au choix :

- espace
- tabulation
- ;
- ,

L'utilisateur doit définir ce caractère dans le menu *fichier/propriétés* :

Chaque individu du fichier de données sera associé à un enregistrement de la table externe au moyen de **clés de fusion**.

Dans chaque table externe, on saisit les variables que l'on veut sans limitation de nombre, au moyen de la fonction *import* .

Il faut d'abord créer **un lien**, définissant le fichier externe et les clés de fusion, en cliquant sur *table externe* dans le menu *outils*.

From	To	Variable
0	0	▼
0	0	▼
0	0	▼
0	0	▼
0	0	▼

On peut alors modifier un lien existant, créer un nouveau lien, supprimer un lien.

Quand on crée un nouveau lien, on clique sur le bouton *New Std* ou *New CSV* selon le cas. Le numéro de lien est alors créé automatiquement.

On renseigne alors le fichier correspondant à la table externe.

Puis, dans l'onglet correspondant au cas traité (format standard ou csv), on donne, pour chaque clé et suivant le cas :

- les positions de début et de fin dans le fichier externe,
- ou le numéro du champ correspondant,
- et la variable du dictionnaire associée. Cette variable est obligatoirement **de type texte**. **Elle peut être dimensionnée**. Si plusieurs clés sont dimensionnées pour la même table externe, elles doivent toutes avoir même dimension (mais il peut y avoir en plus des clés non dimensionnées).

3.6.2 Lecture des variables.

Elle se fait au moyen de la fonction *import*, ou *importCSV* (voir dans la partie fonctions ci-dessus), analogue à la fonction *saisie*, mais s'exerçant sur la table externe et non sur le fichier de données principal.

Si une des clés de fusion est dimensionnée, chaque individu du fichier de données sera associé non pas à un, mais à plusieurs enregistrements de la table externe (en nombre égal à la dimension de la clé).

Dans ce cas, les variables lues dans cette table doivent être dimensionnées, et la j^{ème} occurrence de chaque variable sera lue dans l'enregistrement correspondant à la j^{ème} occurrence de la clé de fusion.

Dans ce cas **le pas de lecture doit être mis à zéro**, pour bien spécifier que les positions de lecture sont les mêmes dans tous les enregistrements.

3.7 Bases.

Ce champ peut recevoir un nom de **variable logique**, qui doit exister et précéder la variable appelante dans la liste des variables.

Quand ce champ « base » est renseigné pour une variable, cela entraînera :

- que la fonction d'exécution ne lui sera pas appliquée si la variable logique dans la base a la valeur *fausse*
- qu'elle sera filtrée automatiquement par cette base dans tous les tableaux où elle apparaîtra.

La variable logique dans le champ « base » peut être **dimensionnée**, si et seulement si la variable appelante (ou une variable argument dans certains cas, voir tableaux de contrôle) est dimensionnée, avec la même dimension.

Cas particulier : pour les fonctions *saisie* et *saisiecsv*, la base ne sera pas prise en compte au moment de l'exécution de ces fonctions, qui seront donc exécutées même si la variable logique a la valeur *fausse*.

La base ne sera prise en compte (et la valeur de la variable mise à la valeur REBUT si la variable logique a la valeur *fausse*), qu'après que toutes les variables auront été calculées, donc au moment de passer la main au programme appelant (tableaux, exportation, etc..).

Une variable avec une fonction *saisie* ou *saisiecsv* peut donc avoir une valeur différente de REBUT au moment du calcul de toutes les autres variables, et une valeur REBUT finalement.

3.8 Généralités sur les exécutions.

Les variables sont calculées séquentiellement, de la première décrite à la dernière, quels que soient leurs niveaux ou leurs types, avec les cas particuliers suivants :

- les fonctions « *saisie* » sont exécutées les premières, dès la lecture sur le fichier de données de l'enregistrement correspondant.
- Les autres fonctions sont exécutées dans un 2^{ème} passage,
- La fonction « *rang* » est exécutée à la fin, après toutes les autres (puisque toutes les variables doivent avoir été traitées avant de décider si l'individu en cours sera accepté. Donc il ne faut pas utiliser une variable construite avec *rang* comme argument dans une fonction d'une autre variable.
- Dans le cas d'une étude à niveaux, les fonctions seront exécutées niveau par niveau : cela signifie que pour un individu statistique de niveau *i*, seules seront exécutées les fonctions relatives à des variables résultats de niveau *i*.

Traitement des Rebut (NA)

Pendant la phase de construction des variables, une variable est en Rebut si elle n'a pas de valeur, ou bien si sa valeur est incompatible avec son type (par exemple une lettre pour une variable Single). Cette définition est indépendante des valeurs figurant dans le **format** affecté à la variable, et elle reste la même dans les phases de construction de tableaux, d'exportation de fichiers, etc..

Cette définition **change pour les variables de type S, M ou T ayant un format** : en effet, dans les tableaux, il est naturel d'envisager que tous les individus n'entrant dans aucune des lignes du format entre dans la ligne N (non réponse). Cela est fait ainsi :

Après que toutes les variables aient été calculées pour un individu statistique, et avant de passer la main au programme en cours (tableaux, redressement d'échantillon, exportation, etc...), **pour toutes les variables ayant un format**, le logiciel vérifie que leurs valeurs font partie des valeurs figurant dans le format, (que ces valeurs figurent isolément ou dans des pages) : si ce n'est pas le cas, la variable correspondante est mise en Rebut.

Par exemple, si la variable X a un format dans lequel figurent les valeurs 1 à 10, et si elle prend la valeur 99, cette valeur 99 sera utilisable pour une variable Y construite à partir de X, et cependant elle sera remplacée par Rebut lors de la transmission de l'ensemble des variables au programme appelant.

Cette valeur 99 ne sera donc pas visible en preview, et pourtant elle aura pu être utilisée pour construire la variable Y.

Grâce à cette formule, dans les tableaux, l'option N des modèles correspond à tout ce qui n'est pas compris dans le format.

En résumé, la mise en Rebut des valeurs hors format intervient donc après le calcul de toutes les variables pour un individu, et elle est effective pour tous les programmes : prévu, tris à plat, export, redressement, tableaux.

Le processus est donc le suivant :

3.8.1 Cas d'une étude sans niveau, avec des individus mono-enregistrements :

- lecture d'un enregistrement,
- exécution de toutes les fonctions *saisie*,
- exécution de toutes les autres fonctions, sous contrôle de leurs bases,
- exécution de la fonction rang,
- application de la base aux variables avec une fonction *saisie*
- vérification des valeurs pour toutes les variables ayant un format, et mise en Rebut des variables dont la valeur est hors format.
- envoi de l'individu vers le programme appelant (tableaux, redressement, etc.), si l'univers éventuellement défini dans l'écran de structure est vérifié.
- boucle sur ce qui précède jusqu'à la fin du fichier

3.8.2 Cas d'une étude sans niveau, avec des individus multi-enregistrements :

- lecture d'un enregistrement,
- exécution de toutes les fonctions saisie concernées par cet enregistrement,
- boucle sur ce qui précède jusqu'à épuisement des enregistrements de l'individu en cours
- exécution de toutes les autres fonctions, sous contrôle de leurs bases,
- exécution de la fonction rang,
- application de la base aux variables avec une fonction *saisie*
- vérification des valeurs pour toutes les variables ayant un format, et mise en Rebut des variables dont la valeur est hors format.
- envoi de l'individu vers le programme appelant (tableaux, redressement, etc.) si l'univers éventuellement défini est vérifié.
- boucle sur tout ce qui précède jusqu'à la fin du fichier.

3.8.3 Cas d'une étude à niveaux :

- lecture d'un enregistrement,
- exécution de toutes les fonctions saisie concernées par cet enregistrement,
- boucle sur ce qui précède jusqu'à épuisement des enregistrements de la famille en cours (une famille étant définie pas un individu de niveau 1 et de tous ses descendants).
- exécution de toutes les autres fonctions, sous contrôle de leurs bases, séquentiellement dans l'ordre où elles sont placées. Pour une variable de niveau 1, comme il n'y a qu'un représentant de niveau 1 par famille, il n'y a donc qu'une exécution de la fonction correspondante. Pour une variable de niveau plus élevé (2, 3, etc.), comme il peut y avoir plusieurs représentants de ce niveau, **la fonction est exécutée pour tous les individus de ce niveau, avant de passer à la variable suivante.**
- exécution de la fonction rang pour chaque individu,
- application de la base aux variables avec une fonction *saisie*
- vérification des valeurs pour toutes les variables ayant un format, et mise en Rebut des variables dont la valeur est hors format.
- envoi de la famille vers le programme appelant (tableaux, redressement, etc.) : le programme envoie les individus un par un, en commençant par l'individu de niveau 1, suivi du premier de ses enfants, et ainsi de suite en explorant toute l'arborescence de la

famille selon un ordre lexicographique, de sorte que chaque père précèdera toujours ses enfants, et qu'il n'y aura pas de mélange des différentes branches (un individu ne se trouvera pas au milieu de ses cousins).

- boucle sur tout ce qui précède jusqu'à la fin du fichier

Cas particulier des variables de niveau 0.

Dans une étude à niveau, les variables appartiennent généralement à un des niveaux définis dans la structure. Elles peuvent aussi appartenir au niveau 0 : par convention, une variable de niveau 0 sera considérée comme appartenant à tous les niveaux, ce qui signifie qu'elle sera calculée pour chaque individu de chaque niveau.

En général, on utilise le niveau 0 pour définir les identifiants des niveaux, éventuellement des numéros séquentiels obtenus à partir de la fonction Rang, ou encore pour enregistrer explicitement le niveau de l'individu (par exemple en vue d'une exportation vers un fichier extérieur).

3.8.4 Exécution multi-fichiers

Toutes les exécutions de CoTab(tableaux, exportations, etc..) peuvent s'effectuer sur un ou plusieurs fichiers, (par exemple plusieurs vagues d'enquêtes), qui seront donc traités séquentiellement, comme un seul fichier concaténé : il suffit de nommer tous ces fichiers dans la fenêtre demandant le nom du fichier à traiter.

3.8.5 Tables externes.

Chaque fois qu'une fonction *import* doit être exécutée, le programme va chercher dans la table externe appelée l'enregistrement correspondant aux clés spécifiées, compte tenu de leur valeur courante (en cas de doublon, on prend le premier rencontré).

La variable à importer est lue dans cet enregistrement. Il se peut donc que, pour le même individu, 2 variables avec la fonction *import* dans la même table externe soient lues dans 2 enregistrements différents, si la valeur de la variable clé a été modifiée entre les 2 (par l'action d'une procédure).

Rappel : si la clé est dimensionnée, les variables lues par *import* dans la table externe correspondante doivent être dimensionnées, chaque occurrence étant lue dans l'enregistrement externe correspondant à l'occurrence de même indice de la clé.

3.8.6 Fichier « cache ».

Cette nouvelle option concerne les traitements lourds (plusieurs dizaines de milliers d'individus, par exemple) et permet d'améliorer considérablement les temps d'exécution.

En effet, le traitement d'un programme quelconque (exécution d'une demande de tableaux, ou d'une demande d'exportation, etc..) consiste à calculer dans une première phase l'ensemble des variables pour un individu, puis dans une deuxième phase à exécuter le programme de tableaux (ou d'export, etc..) pour cet individu, et ainsi de suite pour chaque individu du fichier de données.

Or la 1^{ère} phase peut être beaucoup plus longue que la 2^{ème} (s'il y a beaucoup de variables, ou si les procédures de calcul des variables sont très complexes), et donc si l'utilisateur doit exécuter successivement plusieurs demandes de tableaux, il devra exécuter autant de fois toute la 1^{ère} phase, ce qui sera très coûteux.

Si on utilise l'option *fichier cache*, (en cliquant sur la case **use cache** dans les propriétés de l'étude, menu *tools*), la 1^{ère} exécution faite pour un dictionnaire et un fichier data donnés écrira dans un fichier *cache* le résultat de toutes les variables calculées pour tous les individus, et lors de la demande suivante d'exécution de tableaux (ou export, ou autre), le programme relira automatiquement ce fichier cache au lieu de lire le fichier de données et de calculer à nouveau toutes les variables.

Ce fichier *cache* restera valable tant que le dictionnaire ne sera pas modifié : si le dictionnaire est modifié (nouvelle variable, modification d'une fonction, etc.), **le fichier cache sera automatiquement détruit et donc reconstitué.**

Exception : ni les formats, ni les bases des variables ne sont pris en compte pour évaluer si le dictionnaire a été modifié, ni également l'utilisation des univers dans la structure de l'étude. Dans ces 3 cas, le fichier *cache* n'est pas détruit.

L'utilisateur peut à tout moment écraser ce fichier *cache* (*reset cache* dans le menu *tools*).

Exécution sur plusieurs fichiers de données : dans ce cas, que ces fichiers soient appelés dans des exécutions successives ou dans la même exécution, le fichier cache contient un sous-fichier *cache* par fichier de données. Donc, si on demande l'exécution sur un fichier de données, le programme partira ou non du fichier *cache* selon qu'il contiendra ou non le sous-fichier *cache* correspondant au fichier demandé.

Application : une étude portant sur plusieurs vagues : si on demande une exploitation portant sur plusieurs vagues, le programme utilisera les sous-fichiers *cache* présents, et créera ceux correspondant à de nouvelles vagues.

3.9 Fusion d'études.

3.9.1 Principe.

Il s'agit d'enrichir l'étude principale traitée avec des données produites par une étude secondaire. L'étude secondaire porte sur des individus communs (au moins en partie) avec l'étude principale, et contient des variables autres que celles figurant dans l'étude principale. Une application est par exemple un panel d'individus, pour lesquels on dispose d'une part d'une étude comportant les données de recrutement, d'autre part d'une étude ponctuelle sur un sujet donné.

Lorsque l'on dépouille cette étude ponctuelle, on souhaite bien sûr pouvoir disposer pour chaque individu des variables de l'étude de recrutement.

On peut avoir comme cela plusieurs études secondaires si nécessaire.

3.9.2 Mise en œuvre.

En fait, toute étude secondaire correspondra à une table externe.

La marche à suivre est la suivante :

Il faut définir une table externe correspondant au fichier data de l'étude secondaire.

Puis, dans l'écran des variables de l'étude principale, on choisit l'option « link » dans le menu obtenu par le clic droit : on est alors invité à choisir une enquête, et à sélectionner les variables souhaitées de cette enquête.

Puis on est invité à choisir le lien de table externe défini préalablement.

Toutes les variables ainsi sélectionnées de l'étude secondaire sont alors incluses dans l'étude principale, toutes les fonctions « Input » ayant été automatiquement remplacées par des fonctions « Import » correspondant au lien externe choisi.

En cas de sélection d'une variable déjà présente dans l'étude principale, l'utilisateur peut au choix écraser celle de l'étude principale, ignorer ou renommer celle de l'étude secondaire.

Des choix identiques seront proposés pour les formats, si on rencontre des noms identiques.

Remarques :

La fonction « Import » générée sera CSV ou non selon la nature du lien demandé. Il faut prendre soin de la cohérence de ce choix avec la nature des fonctions « Input » de l'étude secondaire.

L'étude secondaire peut avoir elle-même ses propres tables externes : les fonctions « Import » correspondantes seront conservées en l'état, mais l'utilisateur devra lui-même déclarer ces tables externes dans l'étude principale, avec le même numéro que ceux figurant dans l'étude secondaire.

3.10 Prévisualisation

Dans le menu « outils », si on sélectionne « prévisualisation », (ou avec l'icône d'exécution), cela déclenche la lecture du fichier d'entrée, et le calcul des variables pour le premier individu ; les valeurs de ces variables sont visibles dans la colonne « prévisualisation ». En cliquant sur le haut de la colonne, on passe à l'individu suivant.

N°	Nom	Dim.	Type	Format	Titre	Fonction	Bas
1	AG2		Sim...	AG2(3)	Age of respondent / CHILDREN 1-1...	Input(10)	
2	rang		Sim...			RankNf	
3	ID		Texte			Input(16-17)	
4	rangt		Texte			Conv(rang)	
5	P		Qua...			ImportCsv(L1 / 2)	
6	AGER3		Sim...	AGER...	Age of respondent / ADULTS 13 (by...	Input(11)	
7	ager3bis		Sim...	AGER...	Age of respondent / ADULTS 13 (by...	Arithm	
8	SEX		Sim...	SEX(2)	Sex	Input(12)	
9	HOMMES		Logi...			Recod(SEX)	
10	AGER		Sim...	xxx(4)	Age of respondent	Input(13-14)	
11	AGER7		Sim...	AGER...	Age of respondent	Input(15-16)	
12	AGE57		Logi...			Recod(AGER7)	
13	AGER7BIS		Mult...	AGEBI...	Sex and Age of respondent	Recod	
14	OCC		Sim...	OCC(8)	Occupation of the Head of household	Input(17)	
15	AGR7BIS		Mult...	AGER...	Sex and Age of respondent	Recod	HOM
16	F50		Logi...		FEMMES DE MOINS DE 50 ANS	Recod	
17	AGEHOH		Sim...	AGEH...	Age of the Head of household	Input(18)	HOM
18	HOHSIZ		Sim...	HOHSI...	Household size	Input(19)	F50
19	NIELSE		Sim...	NIELS...	Nielsen regions	Input(20)	
20	MSIZE		Sim...	MSIZE...	Market Size	Input(21)	
21	PRES		Mult...	PRES(...)	Presence of children aged 17 or bel...	Input(22 * 1)	
22	VENT1		Mult...	VENT1...	Age & Gender (by age brackets) / A...	Input(36-37 * 1)	
23	VA1		Mult...	VA1(5)	Age & Gender (by age brackets) / A...	Input(43 * 1)	
24	QCI		Mult...	QCI(7)	QC - Time when started breakfast	Input(1843 * 1)	
25	QCI		Sim...	QCI	QC - Time for breakfast	Input(1845)	

Remarque : si on ouvre une variable dimensionnée (en cliquant sur le signe + à gauche de la dimension), on fait apparaître les valeurs de toutes les occurrences.

3.11 Tris à plat.

Cette nouvelle option permet d'avoir immédiatement la distribution d'une ou plusieurs variables, Ils sont accessibles soit dans le menu *Utilities* (auquel cas on peut choisir tout ou partie des variables), soit avec un clic droit sur une variable (auquel cas on ne s'intéresse qu'à cette variable), et concernent toutes les variables, quel que soit leur type.

N°	Nom	Type	Titre
1	AG2	Simple	Age of respondent / CHILDREN 1-12 (by age brackets)
2	rang	Simple	
3	ID	Texte	
4	rangt	Texte	
5	P	Quantité	
6	AGER3	Simple	Age of respondent / ADULTS 13 (by age brackets)
7	ager3bis	Simple	Age of respondent / ADULTS 13 (by age brackets)
8	SEX	Simple	Sex
9	HOMMES	Logique	
10	AGER	Simple	Age of respondent
11	AGER7	Simple	Age of respondent
12	AGE57	Logique	
13	AGER7BIS	Multiple	Sex and Age of respondent
14	OCC	Simple	Occupation of the Head of household
15	AGR7BIS	Multiple	Sex and Age of respondent
16	F50	Logique	FEMMES DE MOINS DE 50 ANS
17	AGEHOH	Simple	Age of the Head of household
18	HOHSIZ	Simple	Household size
19	NIELSE	Simple	Nielsen regions

Une fois choisi le fichier de données à traiter, on obtient la distribution de la (ou des) variable(s) choisie(s), c'est-à-dire la liste des valeurs rencontrées (colonne de gauche) et le nombre d'apparitions (colonne de droite).

Tris à plat	
AGER3	
Rebut	0
1	65
2	36
3	67
4	81
5	75
6	122
7	54

Pour les variables avec un format, il s'agit de la distribution des valeurs trouvées dans le fichier, et non pas des catégories présentes dans le format. La seule influence du format est que les valeurs hors format sont mises en Rebut (NA) et donc n'apparaissent pas dans la liste affichée.

Pour les variables de type « quantité », on obtient la moyenne, l'écart type, la médiane, et les 1er et 3ème déciles, en plus de la distribution des valeurs.

4. Les modèles.

4.1 Principes.

Dans les tableaux, chaque variable appelée sera associée à un modèle, dont le rôle est de définir les **options** de la variable à faire figurer.

Ces options seront choisies dans la liste suivante (les libellés associés peuvent être modifiés dans le menu « *préférences* ») :

T : Total : le total de la variable tiendra compte de la base éventuellement associée à la variable.

N : Non réponse : correspond au Rebut de la variable, tel qu'il a été défini dans le paragraphe sur les types de variables : il s'agit du vrai Rebut pour les variables sans format, et de toutes les valeurs ne figurant pas dans la liste des codes du format pour les variables avec format.

D: Distribution : il s'agit de la liste fournie dans le format associé à la variable, cette option impose donc que la variable ait un format.

A: Au moins une réponse : il s'agit du total des répondants, donc de l'option contraire à l'option N.

C : Cumul : il s'agit donc de sommer la valeur de la variable (valeur réelle pour les variables de type Simple ou Quantité, ou valeur éventuellement indiquée dans le format pour une variable avec format).

M: Moyenne : il s'agit de la moyenne de la valeur de la variable, valeur définie comme pour le cumul.

Mi : Minimum de la valeur de la variable définie comme ci-dessus

Ma : Maximum de la valeur de la variable définie comme ci-dessus

Md : Médiane : il s'agit de la médiane de la variable définie comme ci-dessus

NB : Nombre moyen de réponses, pour une variable multiple

NT : Nombre de réponses, pour une variable multiple

LO : Liste des occurrences : pour une variable dimensionnée, cette liste permet donc d'individualiser les occurrences, en les éditant dans des positions (lignes, colonnes ou pages) différentes.

W : Weighting (pondération)

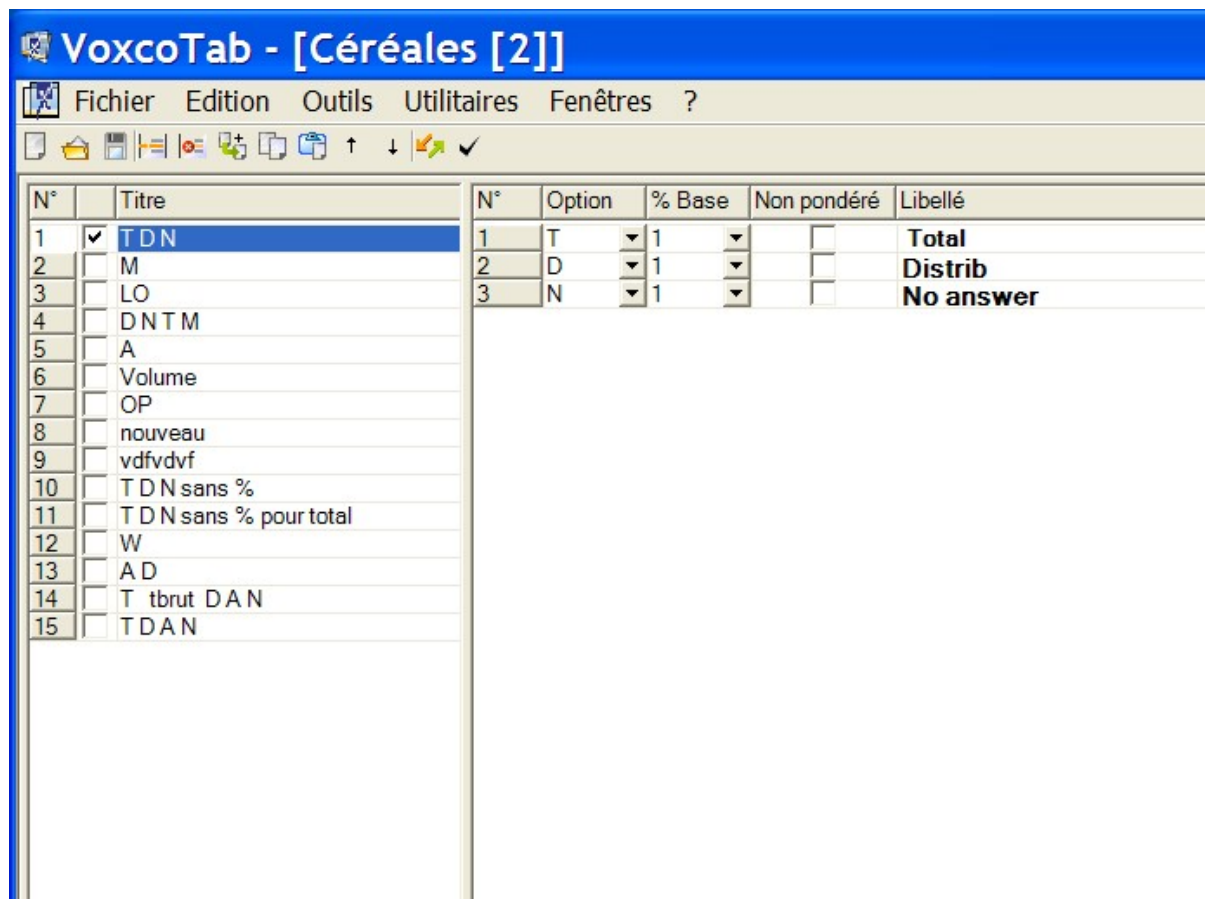
Un modèle comprendra donc une ou plusieurs de ces options, assorties de plusieurs attributs.

L'icône « check » permet de tester qu'il n'y a pas d'erreur dans les variables ni dans les modèles.

4.2 Gestion des modèles.

L'écran obtenu avec l'onglet « modèles » comprend 2 parties : dans la partie gauche, on va créer les modèles, en nombre quelconque.

La partie droite donne le contenu du modèle sélectionné à gauche.



Dans la partie gauche, la gestion des modèles se fait au moyen :

- du clic droit de la souris, on peut insérer un nouveau modèle, **supprimer, dupliquer, renommer, importer, copier ou coller** des modèles d'une autre étude. En création, on doit donner un numéro de modèle (de 1 à 999) et un titre.
- les mêmes opérations peuvent se faire également avec les icônes, et avec le menu « édition ».
- **il faut également spécifier un modèle comme le modèle par défaut**, en cliquant dans la case située à gauche. Ce modèle par défaut sera celui automatiquement associé aux variables pour lesquelles on ne spécifiera pas de modèle.
- en cas d'importation de modèles dont le numéro existe déjà, on est invité à modifier le numéro, ou à écraser l'ancien modèle de même nom.

- quand on sélectionne le champ « *titre* », le clic droit donne accès aux fonctions copier/coller dans ce champ, ou reset du titre.

4.3 Contenu d'un modèle.

Dans la partie droite de l'écran, après avoir sélectionné un numéro de modèle, on peut créer un nombre quelconque de lignes, chacune correspondant à une option.

En sélectionnant une ligne, le clic droit donne accès aux fonctions habituelles : ***insertion, suppression, duplication, copier/coller*** de lignes, ainsi qu'au menu permettant de choisir la police associée à l'option en cours.

Les mêmes fonctions sont accessibles par les icônes, ainsi que par le menu édition.

Les icônes « **flèches** » (ou au clavier ctrl + flèche) permettent de déplacer les lignes.

Les différents champs à remplir sont les suivants :

- **Option** : on peut taper directement la lettre correspondant à l'option désirée, ou cliquer sur le bouton à droite du champ pour choisir cette option.
- **Base de %** : on peut faire figurer un numéro de ligne qui désignera la base du pourcentage qui sera calculé. **La base** doit se rapporter à une ligne existante de type T, C, **NT** ou A exclusivement. **Les demandes de %** (lignes où une base est nommée) ne peuvent pas exister pour les types M, NB, Mi et Ma.
- **Non pondéré** : si on coche cette case, les éléments correspondant à cette option ne seront pas affectés par la pondération éventuelle.
- **Titre** : il s'agit du titre qui figurera dans les tableaux pour cette option, **sauf pour les options D et LO**. Ce titre est généré à partir de ceux indiqués dans le menu « préférences ». Il peut bien entendu être modifié par l'utilisateur.
- **Décimales** : nombre de décimales désiré pour les pourcentages, les moyennes, ainsi que pour les Cumuls, Médianes, Minimum et Maximum.
- **Place** : réservé pour une extension future
- **Non édition** : si on coche cette case, l'élément correspondant ne sera pas édité (utilisé simplement pour des calculs d'autres éléments)

En sélectionnant un champ, le clic droit donne accès aux fonctions *copier/coller* et *reset*. Ces fonctions peuvent s'appliquer à plusieurs lignes, si on a sélectionné un champ sur plusieurs lignes.

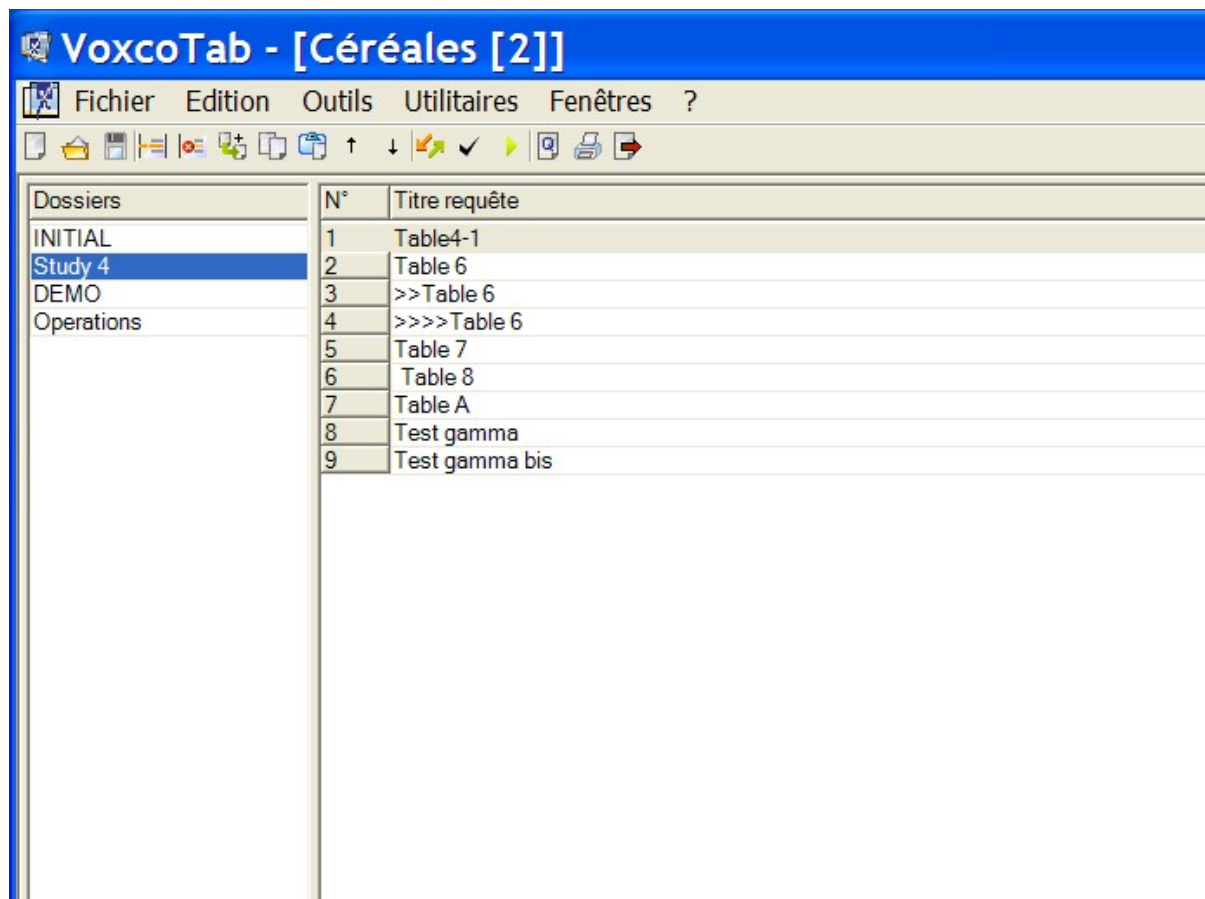
Sur le champ « *titre* », le clic droit donne également accès au menu « **police** », permettant de choisir police et style.

5. Les tableaux.

5.1 Gestion des demandes.

L'onglet « tableaux » donne accès à un écran en 2 parties :

- à gauche, la partie « *dossiers* » permet de définir un nombre quelconque de **dossiers**
- à droite, pour un dossier donné, on peut définir un nombre quelconque de **requêtes** pour un dossier.



Dans ces 2 parties, le clic droit (ou les icônes, ou le menu édition) permet de **créer**, **supprimer**, **dupliquer**, **importer**, **copier et coller** des dossiers (ou des demandes dans un dossier).

Les 3 dernières fonctions se rapportent à une autre étude (ou à la même pour copier/coller).

Pour les dossiers comme pour les demandes, on ne donne qu'un titre.

L'icône « **vérifier** » teste la demande de tableaux sélectionnée, ou, si on est positionné sur un dossier, toutes les demandes du dossier.

5.2 Contenu d'une demande.

En cliquant sur une demande, on ouvre l'écran correspondant, dans lequel l'utilisateur peut renseigner différentes parties :

The screenshot shows the VoxcoTab interface for a study named 'Céréales [2]'. The main configuration area includes:

- Dossier:** Study 4
- Nom:** Table4-1
- Pondération:** WQAD
- Univers:** (empty field)

Below the configuration fields are two tables:

Pages			
Etage	Variable	Etaler	Modèle
1			
2			

Lignes				Colonnes(*)			
N°	Variables	Etaler	Modèle	N°	Variables	Etaler	Modèle
1	V1CL2			1	SEX		
				2	AGER7		

- **Pondération** : nom d'une variable quantité, non dimensionnée (utilisable si le tableau ne nécessite qu'une variable de pondération).

- **Univers** : expression faisant intervenir des **variables logiques exclusivement**, reliées par des opérateurs ET, OU et EX (ou exclusif) : si le résultat de cette expression est vrai pour un individu donné, cet individu sera ventilé dans le tableau, dans le cas contraire il en sera exclu.

Des parenthèses peuvent être utilisées pour définir des priorités entre les opérateurs. En l'absence de parenthèses (ou à l'intérieur de parenthèses), les opérations se font de la gauche vers la droite.

Il peut y avoir des variables dimensionnées dans cette expression, mais toutes de même dimension.

Signification des opérateurs :

A ET B est vraie si les 2 variables logiques A et B sont vraies

A OU B est vraie si l'une au moins des variables A et B est vraie

A EX B est vraie si l'une des 2 variables A et B, et une seulement, est vraie.

Exemple d'univers :

(Q18 ET HOMME) OU (Q25 ET FEMME)

- **Pages** : on peut y inscrire une ou 2 variables servant à composer les pages du tableau
- **Lignes** : on peut y inscrire un nombre quelconque de variables servant à composer les lignes du tableau
- **Colonnes** : on peut y inscrire un nombre quelconque de variables servant à composer les colonnes du tableau

On peut se positionner dans chaque sous-écran correspondant à ces 3 parties (pages, ligne ou colonnes), en cliquant n'importe où dans ces sous-écrans : un signe * viendra alors suivre le nom du sous-écran, permettant de savoir où l'on se trouve.

Dans les écrans lignes ou colonnes, en sélectionnant une ligne quelconque, le **clic droit** donne accès aux fonctions de création, suppression, duplication de lignes, ainsi qu'au copier/coller (pour les pages, ce n'est pas nécessaire, le nombre de lignes possibles étant limité à 2).

Les mêmes fonctions peuvent être effectuées par les icônes, ainsi que par le menu édition.

Les lignes peuvent être déplacées avec ctrl + flèches verticales (ou les icônes flèches).

Notion d'étages :

Dans la partie « pages », la 1^{ère} variable est à l'étage 1, la 2^{ème} variable, si elle existe, est automatiquement à l'étage 2 : cela signifie qu'elle sera croisée avec la 1^{ère} variable : ses différentes valeurs seront étudiées à l'intérieur de chaque valeur de la 1^{ère} variable.

Dans les parties lignes ou colonnes, les variables peuvent être à volonté à l'étage 1 ou à l'étage 2. Toutefois, la 1^{ère} variable décrite doit être à l'étage 1. Une variable à l'étage 2 sera croisée avec la variable à l'étage 1 qui la précède.

Dans l'exemple ci-dessus, en lignes, la variable MSIZE est croisée avec les variables à l'étage 2 : OCC et SEX.

Pour déplacer une variable de l'étage 1 à l'étage 2 (ou vice-versa), on utilise ctrl + flèches horizontales.

Le tableau final demandé comprendra le croisement de toutes les variables pages par toutes les variables colonnes et par toutes les variables lignes.

Pour sortir d'une demande, et revenir à la liste des demandes du folder, on clique sur l'icône représentant une flèche verte gauche.

5.3 Contenu d'une ligne (pages, lignes ou colonnes).

5.3.1. Les variables

Les variables intervenant en pages, lignes ou colonnes peuvent être de n'importe quel type (simple, multiple, texte, quantité ou logique), de n'importe quel niveau (si l'étude est à niveau), et être dimensionnées ou non.

S'il n'y a des **variables dimensionnées non étalées** qu'en lignes (et ni en univers, ni en pages, ni en colonnes), elles peuvent avoir des dimensions différentes, hormis le cas du couple (variable_étage_1 , variable_étage_2) : ces 2 variables devant avoir même dimension si elles sont dimensionnées.

Même chose s'il n'y a des **variables dimensionnées non étalées** qu'en colonnes (et ni en lignes, ni en pages, ni en univers).

Dans tous les autres cas, toutes les **variables dimensionnées non étalées** apparaissant dans la demande doivent avoir même dimension.

Pour inscrire une variable dans une ligne, on **clique sur le bouton à droite du champ**, pour faire apparaître l'écran de sélection de variable :

Variables / Expressions			
Variables		Expressions	
N...	Nom	T...	Titre
1	AG2	Simp...	Age of respondent / CHILDREN 1-1...
2	rang	Simp...	
3	ID	Texte	
4	rangt	Texte	
5	P	Quan...	
6	AGER3	Simp...	Age of respondent / ADULTS 13 (by ...
7	ager3bis	Simp...	Age of respondent / ADULTS 13 (by ...
8	SEX	Simp...	Sex
9	HOMMES	Logiq...	
10	AGER	Simp...	Age of respondent
11	AGER7	Simp...	Age of respondent
12	AGE57	Logiq...	
13	AGER7BIS	Multi...	Sex and Age of respondent
14	OCC	Simp...	Occupation of the Head of household
15	AGR7BIS	Multi...	Sex and Age of respondent

Le bouton « **ligne en cours** » envoie la variable choisie sur la ligne courante, au 1^{er} étage

Les boutons « **ajout 1^{er} étage** » et « **ajout 2^{ème} étage** » permettent d'envoyer la (ou les) variable(s) sélectionnée(s) à la suite de la ligne courante (on peut donc sélectionner plusieurs variables dans cet écran de sélection).

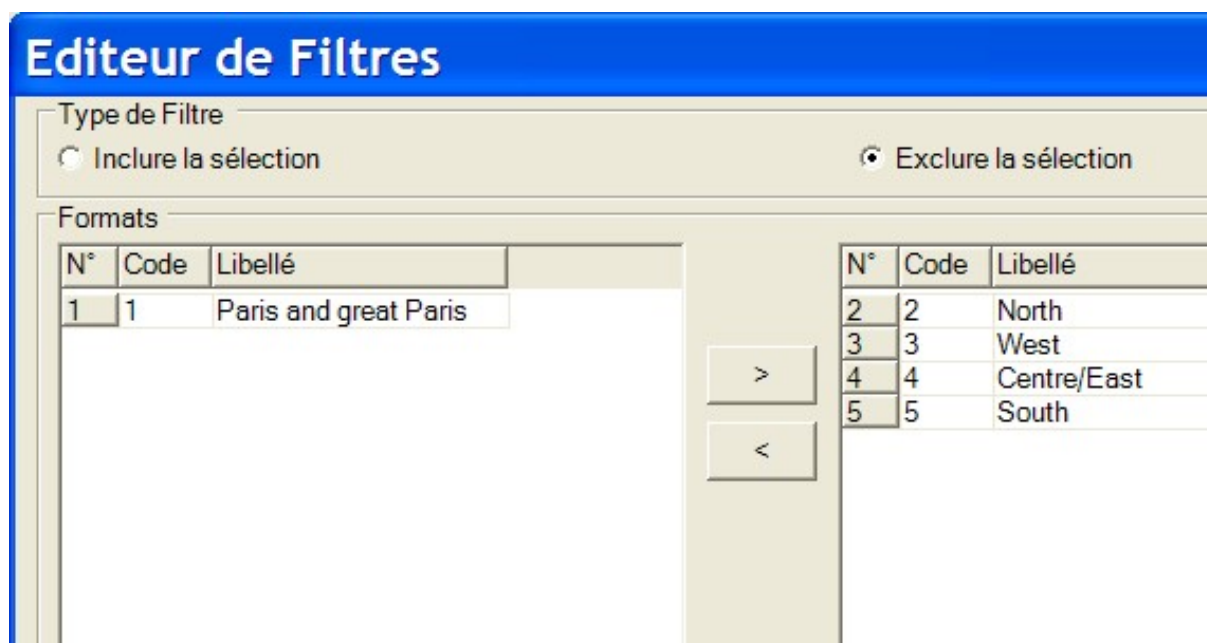
Pour faciliter la sélection, on peut cliquer sur les champs « nom » ou « type », ce qui permet de réduire l'offre de variables en jouant sur le début du nom ou sur le type.

Les variables dimensionnées peuvent être ouvertes en cliquant sur le signe +, ce qui permet de ne choisir qu'une (ou certaines) occurrence(s).

5.3.2. Sélection/filtre de modalités.

Quand on est positionné sur une variable (en lignes, colonnes ou pages), un clic droit donne accès, en plus des options *copier*, *coller*, *restaurer*, aux options *filtrer* et *restauration du filtre*. Ces options ne concernent que les variables associées à un modèle contenant l'option D.

Filtrer donne accès à un écran permettant d'exclure (ou de conserver) certaines modalités du format.



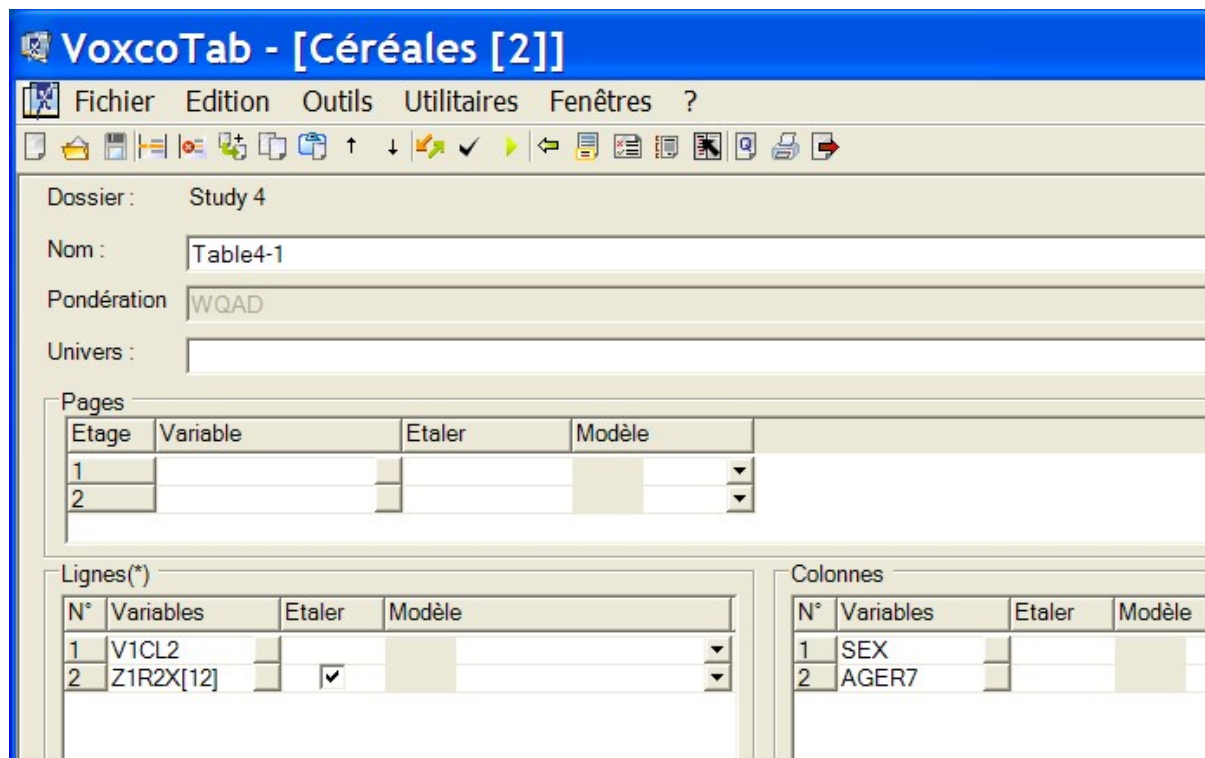
Les modalités initiales du format sont à gauche, les valeurs sélectionnées (à exclure ou à conserver, selon le choix fait), à droite.

Cette option sera effective à l'affichage seulement, c'est-à-dire qu'on peut la modifier une fois le tableau exécuté. Elle concerne exclusivement l'endroit où elle est demandée, c'est-à-dire que si dans le même tableau, la même variable figure ailleurs, elle n'y sera pas affectée par ce filtre.

Les variables filtrées dans une demande de tableau seront mises automatiquement en *Italique*.

5.3.3 Case « étalement ».

Les variables dimensionnées sélectionnées globalement peuvent être **étalées** en cochant la case correspondante : cela a exactement le même effet que si l'utilisateur avait sélectionné toutes les occurrences individuellement (et donc dans ce cas les règles décrites ci-dessus pour les variables dimensionnées ne s'appliquent pas).



5.3.4 Modèle :

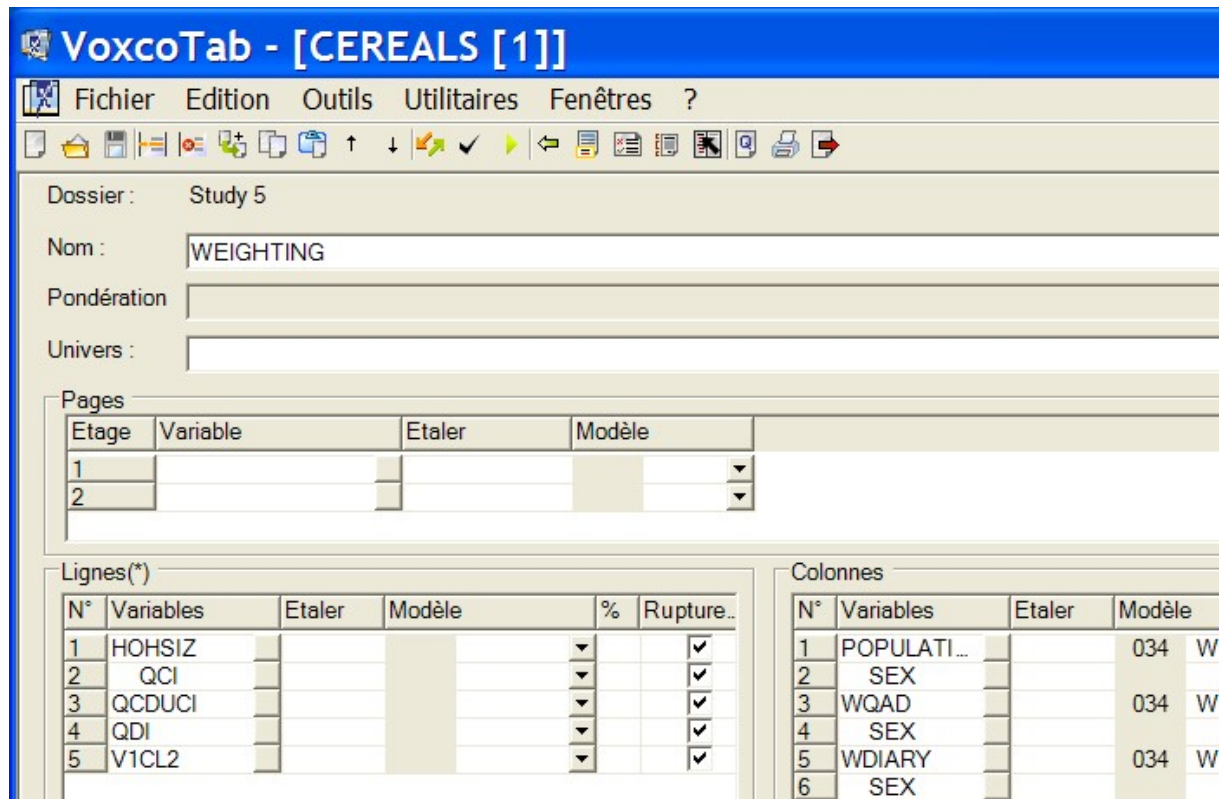
On peut y mettre le numéro de modèle que l'on veut associer à la variable, soit en le choisissant en cliquant sur le bouton à droite, soit en le tapant directement dans le champ : dans ce cas, en tapant le premier chiffre du modèle (1 par exemple), et en répétant cette opération, on obtient successivement tous les modèles commençant par 1 (1, 10, 11, ...)

Si aucun modèle n'est spécifié, le modèle décrit par défaut sera utilisé.

Remarque : Le modèle associé à une **variable logique** ne peut contenir que les options T, N, A ou LO : l'option T correspond au total de la variable, l'option A correspond à la valeur « vraie », l'option N à la valeur « fausse », enfin LO correspond à la liste des occurrences (pour une variable dimensionnée).

Utilisation de l'option W :

Cette option permet de construire des tableaux intégrant plusieurs variables de pondération. Elle peut figurer dans les pages, les colonnes ou les lignes d'un tableau.



Dans l'exemple ci-dessus, le modèle 34 contient uniquement l'option W. L'utilisation d'une option W a le même effet que l'utilisation d'une variable dans le champ Weighting de la demande de tableau, tout en ne concernant que les tableaux élémentaires en découlant.

Remarques :

- 1) que W apparaisse en lignes, colonnes ou pages, il est nécessaire qu'elle soit accompagnée d'une option autre que W, au 1^{er} étage si W est au 2^{ème} étage, et au 2^{ème} étage si W est au 1^{er} étage. Donc en particulier si on n'a qu'un seul étage, W ne peut figurer dans les options présentes.
- 2) S'il y a plusieurs options W dans le tableau, une seule est choisie dans chaque tableau élémentaire, avec l'ordre de priorité suivant : d'abord en colonne (étage 2 sinon étage 1), sinon en lignes, sinon en pages, sinon dans le champ weighting du tableau.

5.3.5 Pourcentages (%) :

Ce champ est en général **non renseigné**. Son utilisation permet de **modifier la définition de la base de pourcentages** qui est demandée dans le modèle associé à la variable : si on y inscrit un numéro de ligne, la base de pourcentage sera prise dans le couple variable/modèle inscrit dans la ligne de la demande de tableaux ayant ce numéro.

Cela permet d'avoir une base de pourcentages complètement indépendante de la variable qui demande ce pourcentage. Néanmoins les contraintes suivantes doivent être respectées :

Quand une variable A appelle comme pourcentage une ligne de la demande de tableaux correspondant à la variable B :

- si A est au 1^{er} étage, B doit être au 1^{er} étage
- si A est au 2^{ème} étage :
 - soit B est au 2^{ème} étage **dans le même groupe**, c'est-à-dire que la variable du 1^{er} étage dont elles dépendent doit être la même.
 - soit B est au 1^{er} étage, dans un autre groupe que celui de A.

De plus, si dans le modèle associé à A, la base de % est la ligne j de ce modèle, la ligne j du modèle associé à B doit être de type T, C, ou A exclusivement.

En général, ce champ « pourcentages » est utilisé en se référant à **une ligne contenant simplement une variable logique, associée à un modèle avec une seule option** : son utilisation est alors très simple.

Remarque concernant les pourcentages.

Ils sont toujours calculés au **dernier étage présent** : cela veut dire que lorsqu'il y a 2 étages, on ne tient pas compte des demandes de pourcentages faites avec la variable du 1^{er} étage.

De même, on ne tient pas compte de la nouvelle base demandée (cf ci-dessus) si la ligne appelée est au premier étage et si elle est suivie d'un deuxième étage.

5.3.6 Changement de page (break) :

En cliquant dans cette case, on provoquera à l'édition un changement de page avant l'édition des tableaux concernant cette variable (valable pour les éditions *pdf*, Excel-XML).

On peut cocher ou décocher en une seule fois la case *break* de plusieurs variables d'une demande de tableau.

Il suffit de sélectionner plusieurs variables, et de cliquer sur **Check all (clic droit)** pour les cocher simultanément.

Pour les champs « variables » et « modèles », le copier/coller et le reset peut concerner plusieurs lignes.

5.4 Exécution d'une demande.

5.4.1 Contrôle de la demande.

Quand on clique sur l'icône de check, ou l'icône d'exécution d'une demande de tableaux (petit triangle vert), le logiciel vérifie les variables, les modèles, et le contenu de la demande choisie. En cas d'erreurs, elles sont signalées et l'exécution n'a pas lieu. Les vérifications faites sont les suivantes :

- les variables appelées existent
- la variable pondération est de type quantité, non dimensionnée
- les variables Univers sont des variables de type Logique
- les variables Page, Ligne ou Colonne dont le modèle inclut une option D (distribution) doivent avoir un format
- les options C, M, Mi, Ma, Md dans un modèle impliquent que la variable appelante soit de type simple ou quantité, ou encore de type multiple ou texte à condition qu'il y ait des coefficients (valeurs) dans le format correspondant
- l'option W dans un modèle implique que la variable appelante soit de type quantité, et qu'elle soit associée au 1^{er} étage (ou 2^{ème} étage suivant le cas) à une option autre que W.
- les options NB et NT impliquent que la variable appelante est multiple
- l'option LO implique que la variable appelante est dimensionnée
- En cas de présence de variables dimensionnées non « étalées » (c'est-à-dire dont la case « expand » est non cochée) : elles peuvent figurer en page, ligne, colonne ou univers, et elles doivent toutes avoir même dimension, hormis le cas où elles ne sont présentes qu'en ligne, et ni en pages, ni en colonnes ni en univers (ou qu'en colonnes, et ni en pages, ni en lignes ni en univers) auquel cas il peut y avoir coexistence de variables de dimensions différentes, à condition que ce soit dans des blocs différents (un bloc est formé d'un couple de variables inscrites dans les étages 1 et 2, d'une demande « lignes » ou « colonnes »).

Les erreurs sont signalées dans une nouvelle fenêtre en bas des demandes : en cliquant sur un message d'erreur, cela introduit un fond rouge sur la ligne concernée de la demande.

5.4.2 Exécution.

Sinon, le fichier de données est lu individu par individu. Si les conditions d'entrée dans les tableaux sont satisfaites (c'est-à-dire si l'expression figurant dans *l'univers* est vérifiée), les données sont ventilées dans tous **les tableaux élémentaires** définis par une variable page (ou 2 s'il y a une variable au 2^{ème} étage), une variable colonne (ou 2 s'il y a une variable au 2^{ème} étage), et une variable ligne (ou 2 s'il y a une variable au 2^{ème} étage), en tenant compte des différentes modalités de leurs modèles associés.

La ventilation dans un tableau élémentaire dépend cependant **des bases** éventuelles associées dans le dictionnaire aux variables ligne, colonne ou page, selon les règles suivantes :

- **une base** associée à une variable de croisement (ligne, colonne ou page) sera prise en compte sauf dans les 2 cas suivants :
 - . sa dimension est supérieure à la dimension de la variable (ce qui veut dire que la base est dimensionnée, la variable ne l'étant pas : ce cas peut se produire avec la fonction Nbval par exemple).
 - . son niveau est inférieur au niveau de la variable (donc le niveau de la base est de numéro plus grand que le niveau de la variable)
- si une base est prise en compte, et si sa valeur est *false*, alors il n'y aura pas de ventilation dans ce tableau élémentaire pour l'individu traité.

Si le champ *pondération* est renseigné, la valeur *Rebut* de la variable correspondante entraîne qu'on n'entrera dans le tableau que dans les cases « *non redressé* », c'est-à-dire pour lesquelles un de leur modèle (lignes, colonnes ou pages a l'option « *non redressé* »).

Exemple :

- variable Page : SEXE avec modalités T, D, N de son modèle (total, distribution, non-réponse)
- variable Colonne : REG avec modalités T, D, N de son modèle
- variable Ligne : Q5 avec modalités T, D, N, M de son modèle

Les tableaux élémentaires sont par exemple :

SEXE (T) X REG (D) X Q5 (D), ou
SEXE (N) X REG (D) X Q5 (M)

L'association d'une variable avec une des modalités de son modèle associé est appelée **élément**. Une **chaîne (page, colonne ou ligne)** est définie par un élément du 1^{er} étage associé avec un élément du 2^{ème} étage s'il existe.

Un **tableau élémentaire** est donc le croisement de 3 chaînes au maximum (chaînes pages, colonnes et lignes), contenant 6 éléments au maximum.

5.4.3 Ventilation dans un tableau élémentaire :

- **tableau élémentaire non quantitatif** : c'est le cas où dans aucune des chaînes qui composent ce tableau, il n'apparaît les modalités C, M, Mi, Ma, Md, NT ou NB. On ventile alors une unité (éventuellement pondérée en cas d'existence d'une pondération) dans toutes les cases du tableau élémentaire concernées par l'individu traité, c'est-à-dire définies par le croisement de toutes les pages, toutes les colonnes et toutes les lignes vérifiées par l'individu dans ce tableau élémentaire, les bases prises en compte ayant la valeur *vraie*.

- **tableau élémentaire quantitatif** : c'est le cas où au moins une des chaînes qui composent ce tableau, il apparaît un des modalités C, M, Mi, Ma, Md, NT ou NB : auquel cas c'est la valeur de la variable correspondante qui va être ventilée dans toutes les cases du tableau élémentaire définies comme précédemment :

- avec C, la valeur de la variable correspondante est simplement cumulée (éventuellement multipliée par la pondération) dans les cases.

- avec M, on cumule la valeur de la variable, ainsi que son carré et l'effectif correspondant (éventuellement pondérés), afin de calculer à la fin du traitement la moyenne et l'écart-type.
- avec Mi ou Ma, on enregistre le minimum (hors Rebut) ou le maximum de la variable dans chacune des cases concernées.
- avec Md, on enregistre toutes les valeurs (hors Rebut) de la variable dans chacune des cases concernées, afin de calculer et d'éditer à la fin du traitement la médiane.
- avec NT, le nombre de valeurs de la variable correspondante est cumulé (éventuellement multiplié par la pondération) dans les cases
- avec NB, on cumule le nombre de valeurs (éventuellement pondéré) de la variable multiple associée afin de calculer sa moyenne en fin de traitement.

Cas où le tableau élémentaire contient plusieurs quantités : c'est-à-dire que les modalités C, M, Mi, Ma, Md, NT ou NB apparaissent plus d'une fois dans les six éléments possibles constituant le tableau élémentaire : comme dans une case du tableau élémentaire, on ne peut faire intervenir qu'une quantité, elle est choisie dans l'ordre de priorité suivant : d'abord en colonnes, à défaut en lignes, à défaut en pages, et dans chacune de ces dimensions, d'abord au 2^{ème} étage, à défaut au 1^{er} étage.

Tableau élémentaire dimensionné : c'est le cas où une au moins des variables servant à composer ce tableau élémentaire est dimensionnée et non étalée (case *étalement* non cochée), c'est-à-dire **une des variables univers, ou une des variables des chaînes pages, colonnes ou lignes.**

Dans ce cas, soit n le nombre d'occurrences de la variable dimensionnée (on rappelle que s'il y a plusieurs variables dimensionnées dans le tableau élémentaire, elles ont toutes la même dimension). L'individu traité sera ventilé n fois dans le tableau élémentaire, avec successivement les occurrences 1, puis 2, etc.

L'option LO des modèles permet alors de ventiler et d'afficher séparément les différentes occurrences.

Cas des fichiers à niveau.

La ventilation se fait alors en fonction des niveaux du questionnaire en cours de traitement d'une part, des différentes variables composant le tableau d'autre part.

Le **niveau d'un tableau élémentaire**, défini par la pondération, l'univers, les variables en pages, lignes et colonnes, est égal **au plus bas niveau** rencontré dans cette énumération (c'est-à-dire dont le numéro de niveau est le plus élevé).

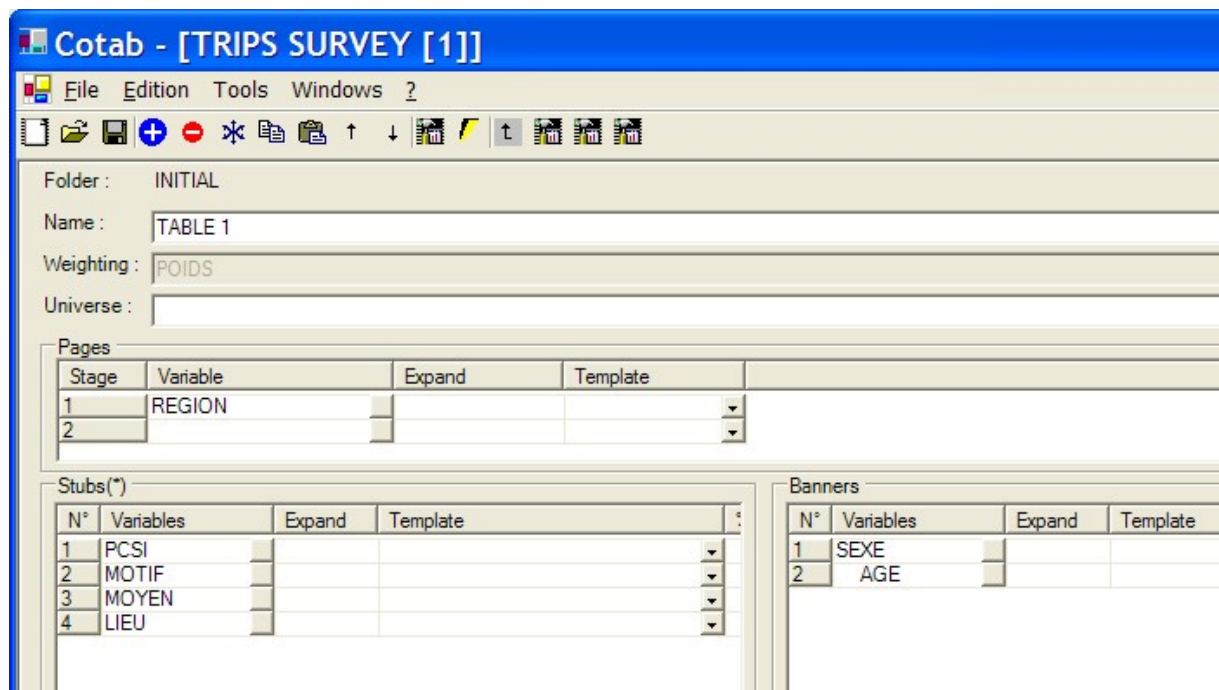
Le principe général est **qu'un tableau élémentaire de niveau j est rempli par les questionnaires de même niveau j** (ce qui n'empêche pas que certains composants du tableau peuvent être de niveau supérieur, donc de numéro plus petit).

Exemple : supposons que l'on ait les variables suivantes :

- au niveau 1 (foyers) : REGION, HABITAT
- au niveau 2 (individus) : SEXE, AGE, PCSI, POIDS

- au niveau 3 (trajets) : MOTIF, MOYEN, LIEU

et la demande de tableau suivante :



- le tableau élémentaire (REGION, SEXE-AGE, PCSI) contiendra des individus

- les tableaux élémentaires (REGION, SEXE-AGE, MODE), (REGION, SEXE-AGE, MOYEN), (REGION, SEXE-AGE, LIEU) contiendront des trajets.

5.5 Exemples de calculs de pourcentages.

Ci-dessous, figure le modèle et la demande de tableaux :

N°	Title	N°	Option	% Basis	No Weighting	Headings
1	<input checked="" type="checkbox"/> T D A N et T non pondéré	1	T	1	<input type="checkbox"/>	TOTAL
2	<input type="checkbox"/> T D N M	2	D	3	<input type="checkbox"/>	D
3	<input type="checkbox"/> MEAN and volume	3	A	1	<input type="checkbox"/>	Respondents
4	<input type="checkbox"/> T N B N	4	N	1	<input type="checkbox"/>	No answer
5	<input type="checkbox"/> T D N et A	5	T		<input checked="" type="checkbox"/>	TOTAL NON PONDERE
6	<input type="checkbox"/> MEAN					
7	<input type="checkbox"/> T D N M C					
8	<input type="checkbox"/> T N					
9	<input type="checkbox"/> T D N M					
10	<input type="checkbox"/> T D N N B M					
11	<input type="checkbox"/> T					
12	<input type="checkbox"/> volume					
13	<input type="checkbox"/> M T A					
14	<input type="checkbox"/> T D M C					
15	<input type="checkbox"/> A					

Folder : INITIAL

Name : First Request

Weighting :

Universe :

Stage	Variable	Expand	Template
1			
2			

N°	Variables	E...	Template	%	Pag
1	REG		001 T D A N et T non pondéré		
2	SEXE		001 T D A N et T non pondéré		
3	GROUP		005 T D N et A	4	
4	B11		015 A	4	

N°	Variables	E..	Template
1	Q5		001 T D A N et T
2	Q14		005 T D N et A
3	B11		015 A
4	Q10		001 T D A N et T

5.5.1 Pourcentages horizontaux.

- Q5 : les colonnes Total, Non Réponse et Au moins une réponse ont la colonne Total de Q5 comme base de pourcentages. Les colonnes correspondant à la distribution de Q5 ont comme base de pourcentages la colonne Au moins une réponse de Q5.
- Q14 : toutes les colonnes ont la colonne B11 comme base de pourcentages.
- Q10 : les colonnes Total, Non Réponse et Au moins une réponse ont la colonne Total de Q5 comme base de pourcentages. Les colonnes correspondant à la distribution de Q10 ont comme base de pourcentages la colonne Au moins une réponse de Q5.

5.5.2 Pourcentages verticaux.

- REG : les lignes Total, Non Réponse et Au moins une réponse ont la ligne Total de REG comme base de pourcentages. Les lignes correspondant à la distribution de REG ont comme base de pourcentages la ligne Au moins une réponse de REG.
- GROUP : toutes les lignes pour lesquelles un pourcentage est demandé, c'est-à-dire les lignes Total, Non Réponse, Au moins une réponse et Distribution, ont la ligne B11 (au moins une réponse, c'est-à-dire valeur *vraie*) comme base de pourcentages, et ce, pour toutes les modalités associées à SEXE, c'est-à-dire TOTAL, Hommes, Femmes, Non réponse, Au moins une réponse, Total non pondéré.
- B11 : la base de pourcentages est la ligne B11 elle-même.

5.6 Présentation des résultats.

L'icône de présentation des résultats fait apparaître le tableau demandé (la même icône sert à revenir sur la demande).

Si des **variables pages** étaient présentes, on peut sélectionner les différentes pages en cliquant sur la flèche de la fenêtre « page » en haut de l'écran.

On peut **redimensionner** les largeurs des colonnes et la hauteur des lignes en se positionnant sur les séparateurs colonnes et lignes, en cliquant dessus, et en se déplaçant avec la souris.

VoxcoTab - [CEREALS [7]]

Fichier Edition Outils Utilitaires Fenêtres ?

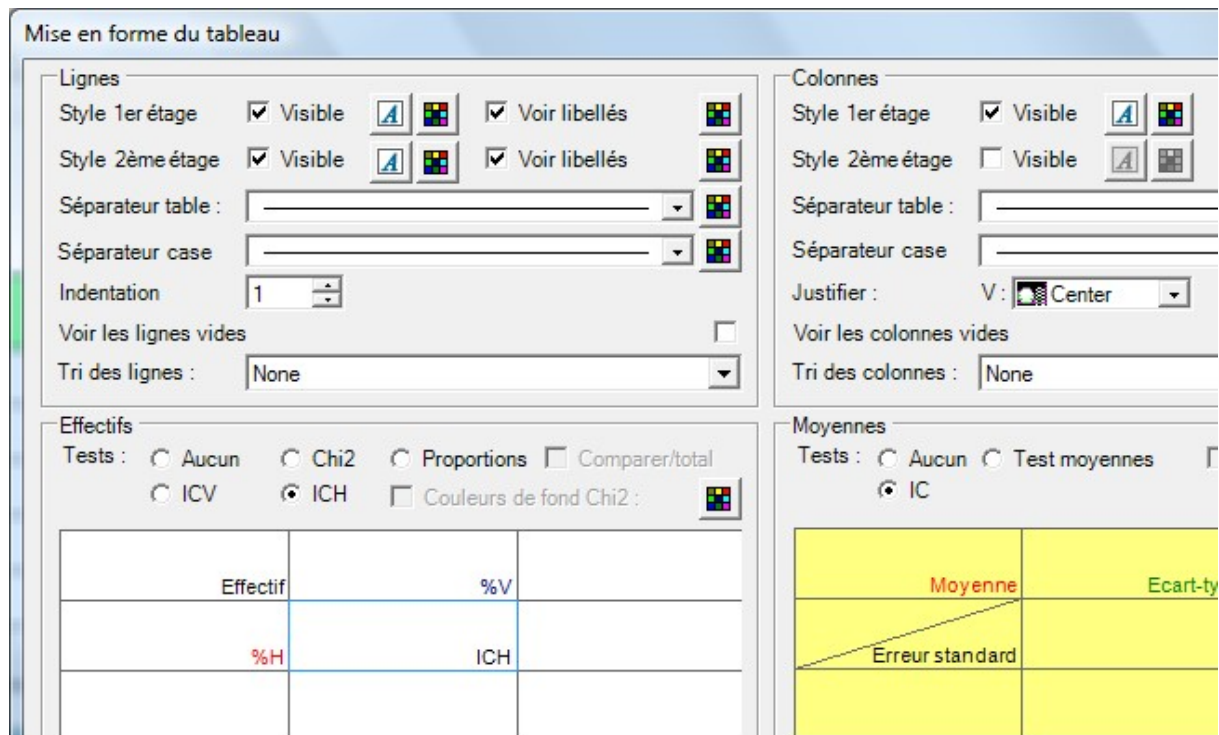
Dossier : Study 4
 Demande : Table4-1
 Page :

	Sex							
	Total		TOTAL unweighted		Male		Female	
QC - Time when started breakfast								
Total	517	100.0%	500	100.0%	243	100.0%	275	100.0%
	100.0%				46.9%		53.1%	
TOTAL unweighted	500		500		229		271	
	100.0%				45.8%		54.2%	
Before 6 am	41	7.9%	52	10.4%	30	12.2%	11	4.1%
	100.0%				72.3%		27.7%	
6:00 - 6:59	71	13.8%	89	17.8%	52	21.3%	20	7.2%
	100.0%				72.3%		27.7%	
7:00 - 7:59	236	45.5%	255	51.0%	161	66.5%	74	27.0%
	100.0%				68.5%		31.5%	
8:00 - 8:59	169	32.7%	104	20.8%			169	61.6%
	100.0%						100.0%	
Liste 2 (products consumed at breakfast) on 1 Week								
Total	517	100.0%	500	100.0%	243	100.0%	275	100.0%
	100.0%				46.9%		53.1%	

Un clic droit sur cet écran permet d'exécuter le tableau, d'appeler les écrans « **mise en forme** », « **mise en page** », « **index** », **test gamma**, **test chi2**, rafraîchissement de l'écran.

5.6.1 Mise en forme du tableau.

La mise en forme du tableau se fait en cliquant sur l'icône de même nom.



Les différentes options de cet écran permettent de modifier la présentation du tableau, que ce soit à l'écran, en édition, ou en exportation Excel.

Dans la partie gauche de l'écran, et de haut en bas, on trouve :

- **la présentation concernant les variables lignes** du 1^{er} ou 2^{ème} étage, pour lesquelles on peut :
 - cacher les titres et/ou les libellés : cette option ne sera prise en compte que si la ligne correspondante n'est pas une ligne de résultats,
 - choisir la police des titres (police, taille, style, couleur,...)
 - (pour les libellés, la police a été donnée dans le format des variables)
 - la couleur de fonds pour les titres et/ou les libellés
- la définition de la **ligne bordure extérieure**, et de **la ligne séparateur des cellules** : on a le choix entre un trait fin, un trait fort, ou pas de trait du tout. De plus, on peut choisir la couleur du trait.
- **l'indentation** des titres et libellés : pour les décaler de n caractères
- **afficher les lignes vides** : case à cocher. Si elle n'est pas cochée, les lignes vides seront supprimées, et si toutes les lignes sont vides, le tableau sera supprimé.
- **tri des lignes** : on peut choisir le tri ascendant ou descendant. Le tri ne se fait qu'à l'intérieur d'une même chaîne correspondant au dernier étage à une option D (distribution) ou LO (liste des occurrences).

Le tri se fait toujours **selon les valeurs de la 1^{ère} colonne**, c'est-à-dire sur **les effectifs**, sauf en présence de moyennes, auquel cas le tri se fait **sur les moyennes**.

Dans le cas d'une distribution, si les libellés du format correspondant font apparaître des indentations, le tri se fera en respectant les hiérarchies induites par ces indentations, comme dans l'exemple ci-dessous.

	Sex				
	Total	TOTAL unweighted	Male	Female	
TOTAL unweighted	500 100.0%	500	229 45.8%	271 54.2%	100
Bread	423 81.8% 100.0%	415 83.0%	167 39.5% 68.9%	256 60.5% 93.1%	100
Fresh	238 46.1% 100.0%	208 41.6%	95 40.0% 39.4%	143 60.0% 52.0%	100
Fresh white bread	219 42.4% 100.0%	150 30.0%	90 41.1% 37.1%	129 58.9% 47.0%	100
Fresh speciality bread	19 3.7% 100.0%	58 11.6%	5 28.0% 2.2%	14 72.0% 5.1%	100
Pre-wrapped	21 4.0% 100.0%	16 3.2%		21 100.0% 7.6%	100
Pre-wrapped speciality bread	21 4.0% 100.0%	16 3.2%		21 100.0% 7.6%	100
Bread something	164 31.6% 100.0%	191 38.2%	72 43.7% 29.5%	92 56.3% 33.5%	100
Viennoiseries	64 12.4% 100.0%	47 9.4%	56 88.1% 23.2%	8 11.9% 2.8%	100
Fresh viennoiseries	20 3.9% 100.0%	28 5.6%	13 46.4% 5.2%	8 29.6% 2.8%	100

- **tests pour les effectifs** : on peut choisir le test de proportions, le test du Chi2, ou les intervalles de confiance pour les pourcentages verticaux (ICV) ou horizontaux (ICH). Ces tests se font sur chaque tableau défini par les variables lignes et colonnes du dernier étage. Ils supposent que les variables définies au dernier étage en lignes et colonnes sont associées à un modèle comprenant l'option T (Total), ou à défaut l'option A (Répondants), de manière à présenter une ligne et une colonne Total.

Si le tableau contient des parties pondérées et des parties non pondérées, les tests seront toujours faits séparément pour chacune de ces parties. Donc, si le tableau présente la distribution pondérée d'une variable, il est nécessaire qu'il y ait un total ligne (demandé par l'option T ou l'option A) pondéré.

Enfin, ces tests supposent que ce tableau n'est affecté d'aucune option M (moyenne), Mi (minimum), Ma (maximum), NB (nombre moyen de réponses).

Le test de Khi2 permet de savoir si l'effectif d'une case est sur-représenté ou sous-représenté : le premier cas est signalé par la présence de signes « + », le deuxième cas est signalé par la présence de signes « - » ; il y a un, deux ou trois signes selon que le test est vérifié à 90%, 95% ou 99%.

Le test de Khi2 suppose la présence d'une colonne et une ligne Total (correspondant à l'option T du modèle, à défaut à l'option A) homogènes (c'est-à-dire pondérées ou non) avec la case testée.

Le test de proportions consiste à comparer les effectifs de tous les couples de 2 colonnes (incluant la colonne total si cela est demandé). Les colonnes sont numérotées A, B, C, etc.. La colonne total, si elle est demandée dans le test, est référencée Z. La colonne correspondant à l'option A (répondants) est référencée Y. Le résultat du test pour une case consistera à l'édition de toutes les lettres correspondant aux colonnes pour lesquelles le test de comparaison avec la colonne étudiée est positif. Toutefois, le résultat du test n'est édité que dans la colonne ayant le plus fort effectif (sauf quand on compare avec la colonne Total, auquel cas le test est édité dans l'autre colonne)

Le test de proportions suppose la présence d'une colonne Total si la comparaison avec le Total est demandée (option T, ou à défaut option A) et d'une ligne Total (correspondant à l'option T du modèle, à défaut à l'option A) homogènes (c'est-à-dire pondérées ou non) avec la case testée.

Remarques :

- s'il y a en colonne à la fois un total pondéré et un total non pondéré, ils seront tous les 2 référencés Z (ou Y), sachant que la colonne Total de comparaison sera celle homogène avec la partie de tableau testée.
- si le tableau présente une distribution pondérée, **pour que le test soit fait en tenant compte des effectifs bruts**, il faut qu'il y ait en ligne dans le tableau à la fois l'option T pondérée, et l'option T non pondérée (ou à défaut, l'option A pondérée, et l'option A non pondérée).

- **Couleur de fonds pour les effectifs** : permet de choisir une couleur de fonds pour les cases « effectifs »

- **Couleur de fonds pour le Khi2** : si on coche cette case (après avoir sélectionné le test de Khi2), les cases significatives (correspondant à +++, ++, ...) reçoivent la couleur de fonds définie grâce au bouton situé à droite.



Après avoir défini les couleurs, leur application est commandée par le bouton « utiliser les couleurs de fonds pour le Khi2 ».

- **présentation des résultats dans les cellules** : ne concerne que les cellules dans lesquelles **il n'y a pas de moyennes**. On dispose de 9 sous-cellules dans lesquelles on peut déplacer, par « drag and drop », les effectifs, les pourcentages horizontaux, les pourcentages verticaux, et les tests ou intervalles de confiance. En cliquant sur l'un de ces éléments, on peut :

- le rendre invisible (en décochant la case « visible ») : pour le test de Khi2, cela rend invisible les signes +++,..., mais n'a pas d'action sur les couleurs de fonds éventuellement demandées pour le Khi2.

- choisir une police, en cliquant sur le bouton « police » (police, taille, style, couleur, etc.)

Remarques concernant les effectifs :

- si la cellule correspond à une option C du modèle, les résultats seront présentés avec décimales, si on a indiqué un nombre de décimales non nul dans l'option C du modèle utilisé. **Attention** : dans ce cas particulier, le nombre de décimales indique sert à la fois pour le Cumul et pour le pourcentage s'il a été demandé.

- lorsque l'on cache les effectifs, les cases du tableau ne contiennent donc plus que des pourcentages. S'il y a des cases pour lesquelles aucun pourcentage n'est demandé, la case devrait donc être vide : pour éviter ce cas, par convention, **les effectifs seront imprimés dans les cases vides à la place du pourcentage manquant**. Cela permet d'avoir des tableaux présentant dans la même colonne les effectifs de la base pour les totaux, et les pourcentages pour les cases du tableau comme dans l'exemple ci-dessous :

The screenshot shows the VoxcoTab software interface with a menu bar (Fichier, Edition, Outils, Utilitaires, Fenêtres) and a toolbar. The main window displays a table with the following data:

Liste 2 (products consumed at breakfast) on 1 Week					
	Total	Bread	Fresh	Fresh white bread	Fresh speciality bread
Gender					
<i>Base non pondérée</i>	500	415	208	150	58
Male	45.8	42.4	40.9	38.7	46.6
Female	54.2	57.6	59.1	61.3	53.4
Répondants	100	100	100	100	100
Age of respondent					
<i>Base non pondérée</i>	500	415	208	150	58
18 - 24 years old	13.0	8.2	9.6	13.3	
25 - 34 years old	7.2	5.1	10.1	14.0	

- **seuil pour les tests** de proportions et de moyennes : quand ces tests sont demandés, on peut choisir les tests 90/95, 95/99, 90, 95 ou 99%. Les résultats seront imprimés en majuscules si le test le plus élevé du couple choisi est satisfait, en minuscules si seul le test le plus bas est satisfait. Dans le cas de seuils uniques (90 par exemple), la lettre sera en majuscules.
- **seuil pour les intervalles de confiance** : il faut choisir les tests 90, 95 ou 99% (sinon c'est le test à 95% qui s'applique).

- **afficher le signe %** : si cette case est cochée, le signe % sera édité après les pourcentages.

Ajustement des pourcentages.

Cette option permet d'ajuster les pourcentages de manière à ce que leur somme soit égale à 100.

La demande se fait en cochant la case à côté de la case « police » des %, active pour le % (H ou V) sélectionné.

Cette option ne concerne **que les options D des modèles**.

- pour un % H, la variable colonne au dernier étage doit être de type S ou T.
- pour un % V, la variable ligne au dernier étage doit être de type S ou T.

Par ailleurs, le % en question doit être **sans décimale** (sinon on ignore la demande).

Pour un % H, dans une ligne, on ne considère que les valeurs correspondant à des **codes simples du format** de la variable colonne au dernier étage, donc on ignore les colonnes correspondant à des codes multiples du format (l'ajustement ne les concernera pas).

Attention : cet ajustement n'est fait que sur les codes apparents. Donc ne pas utiliser cette option **s'il y a des codes cachés** (dont la case NE est cochée dans le format), sous peine de résultats inappropriés.

Par ailleurs, on a les 2 cas suivants :

- si le % est demandé sur une base T, on prendra toutes les valeurs correspondant à des codes simples du format + **la valeur dans la colonne N si elle existe**
- si le % est demandé sur une base A, on prendra toutes les valeurs correspondant à des codes simples du format.

Le programme sommerá les pourcentages correspondant à toutes les valeurs ainsi choisies, et si la somme est différente de 100, il modifierá un ou plusieurs pourcentages afin d'arriver à 100.

Dans la partie droite de l'écran, et de haut en bas, on trouve :

- **la présentation concernant les variables colonnes** du 1^{er} ou 2^{ème} étage, pour lesquelles on peut :
 - cacher les titres et/ou les libellés,
 - choisir la police des titres (police, taille, style, couleur,...)
 - (pour les libellés, la police a été donnée dans le format des variables)
 - la couleur de fonds pour les titres et/ou les libellés
- la définition de la **colonne bordure extérieure**, et de la **colonne séparateur** des cellules : on a le choix entre un trait fin, un trait fort, ou pas de trait du tout. De plus, on peut choisir la couleur du trait.
- la **justification** des titres et libellés en colonnes.
- **afficher les colonnes vides** : case à cocher. Si elle n'est pas cochée, les colonnes vides seront supprimées, et si toutes les colonnes sont vides, le tableau sera supprimé.
- **tri des colonnes** : on peut choisir le tri ascendant ou descendant. Le tri ne se fait qu'à l'intérieur d'une même chaîne correspondant au dernier étage à une option D (distribution) ou LO (liste des occurrences).

Le tri se fait toujours **sur les effectifs de la 1^{ère} ligne**. Si cette 1^{ère} colonne contient la moyenne d'une variable, **le tri se fait sur les moyennes**.

- **tests pour les moyennes** : si l'un au moins des effectifs est inférieur à 30, le test sera fait au moyen du *t* de Student, sinon le test sera fait au moyen de l'écart-réduit.

Le test est fait en comparant les résultats des cases situées dans 2 colonnes. Les colonnes sont référencées A, B, C,... comme dans le test de proportion, le Total étant référencé Z, et la colonne « répondants » (option A du modèle) étant référencée Y. Toutefois, le résultat du test n'est édité que dans la colonne ayant la plus forte moyenne (sauf quand on compare avec la colonne Total, auquel cas le test est édité dans l'autre colonne).

- **intervalles de confiance IC** : en cochant cette case, on obtient l'intervalle de confiance pour les moyennes (choisir le seuil 90,95 ou 99%).
- **couleur de fonds pour les moyennes** : en choisissant ainsi une couleur de fonds pour les cases contenant des moyennes, cela permet de les différencier visuellement des cases contenant des effectifs.
- **présentation des résultats dans les cellules** : ne concerne que les cellules dans lesquelles il y a des **moyennes**. On dispose de 9 sous-cellules dans lesquelles on peut déplacer, par « drag and drop », les effectifs, les moyennes, écarts-type, erreur-standard, tests et intervalles de confiance. En cliquant sur l'un de ces éléments, on peut :
 - le rendre invisible (en décochant la case « visible »)
 - choisir une police, en cliquant sur le bouton « police » (police, taille, style, couleur, etc.)
- **caractère pour case vide** : on peut y inscrire le caractère de son choix, pour remplacer le blanc en cas de case vide (ne concerne que la sous-cellule « effectifs »).
- **boutons « load » et « save »** : la mise en forme du tableau peut être sauvegardée (bouton « save ») ou rappelée « bouton « load »). L'utilisateur donne un nom quelconque quand il sauvegarde.
- **Exemple** de tableau édité en utilisant les couleurs de fonds :

VoxcoTab - [CEREALS [7]]

Fichier Edition Outils Utilitaires Fenêtres ?

Dossier : Study 4
Demande : Table4-1
Page :

	Age of respondent in clear				
	15-30	31-40	41-50	51-99	No answe
QC - Time when started breakfast					
Total	138 27%	55 11%	102 20%	221 43%	0 0%
TOTAL unweighted	144 29%	42 8%	115 23%	194 39%	5 1%
Before 6 am		11 28%	30 72%		
6:00 - 6:59	51 71%		3 2.8%	18 8.1%	
7:00 - 7:59	85 36%	43 18%	48 20%	59 26.5%	0 0%
8:00 - 8:59	2 1.7%		22 21.6%	145 86%	

QC - Time for breakfast

5.6.2 Mise en page du tableau.

Icône « mise en page ».

- bouton « **d'abord à droite** » « **d'abord en bas** » : permet de choisir l'ordre de présentation lors d'un découpage du tableau sur plusieurs pages (tableau trop grand ou découpage causé par une « rupture »).
- **ajustement automatique** : si le tableau est trop grand pour tenir sur une page, et si cette option est cochée, le programme essaiera de réduire la largeur des colonnes afin de faire tenir le tableau dans la page, tout en gardant la lisibilité des nombres dans les cellules). En cas d'échec, les largeurs initiales choisies par l'utilisateur seront conservées.
- **centrer verticalement ou horizontalement** : permettent de centrer le tableau pour l'impression.
- **numéro de la 1^{ère} page** : utilisé si se numéro de page est demandé (voir après)
- **marges** : pour donner les marges (en mm) en haut, à gauche, à droite, en bas de la page, ainsi qu'avant l'en-tête et le pied de page.
- **en-tête** : pour donner l'en-tête du tableau : on peut insérer plusieurs lignes (clic droit ou touche insert).

Le clic droit permet également de définir la hauteur de la ligne : il est vivement conseillé, (sauf cas particulier comme images, etc..) **de choisir l'option « auto »**. Le choix de l'option *fixe* avec la valeur zéro ferait disparaître la ligne de l'impression, même si elle contient des informations.

Chaque ligne est divisée en 3 parties : gauche, centre et droite.

Dans chaque partie, **un clic gauche** donne accès à un écran : on peut écrire un texte quelconque dans la ligne du bas, ou sélectionner un texte dans la liste fournie :

Expression	
Code	Description
Etude	Titre de l'étude
Dossier	Nom du dossier
Demande	Titre de la demande
Pondération	Pondération
Univers	Univers
Page	Titre Page
Date	Date
Heure	Heure
PageNr	Numéro de la page courante
NbPages	Nombre de pages
VLPAGE	Titre de la page logique verticale
HLPAGE	Titre de la page logique horizontale

Par exemple « *titre de l'étude* » permettra d'imprimer le titre de l'étude, « *titre page* » permettra d'imprimer le titre complet de la (ou des) variable(s) page du tableau, avec la valeur de la modalité courante, par exemple : Sexe : Hommes.

Les options **VLPAGE** et **HLPAGE** se rapportent aux titres définis dans le paragraphe suivant (Index).

L'option « **Seuil** » se rapporte au seuil choisi pour les tests.

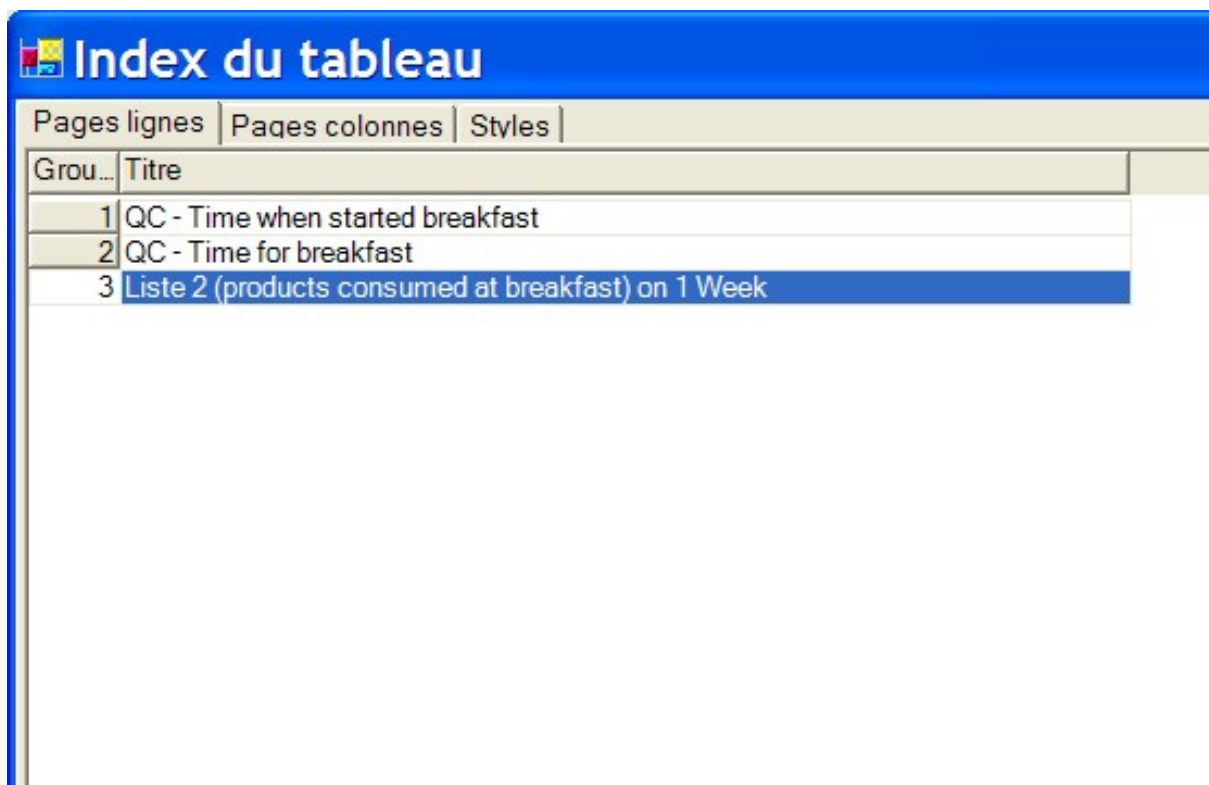
Les options « **Légende effectifs** » et « **Légende moyennes** » permettent d'éditer les légendes correspondantes.

Un **clic droit** sur le même champ permet de choisir la police, et également **d'insérer un logo**.

- **separateur** : permet de mettre un trait après l'en-tête, on peut également choisir sa couleur.
- **fonds de page de l'en-tête** : permet de choisir une couleur de fond pour l'en-tête
- **espace après** : donne l'espace en mm après l'en-tête
- **pied de page** : fonctionne de même manière que l'en-tête.
- **charger et enregistrer** : pour charger un modèle défini précédemment, ou sauver le modèle actuel, sous un nom quelconque choisi par l'utilisateur.

5.6.3 Index.

Icône « *index* ». Lorsque le tableau est découpé en plusieurs pages logiques obtenues en cochant les cases « *rupture* » (*break*) des variables situées en lignes ou colonnes, chacune de ces pages logiques peut être éditée (pdf, excel ou HTML), avec des titres spécifiques figurant dans la table index.



The screenshot shows a window titled "Index du tableau" with a blue header. Below the header is a navigation bar with three tabs: "Pages lignes", "Pages colonnes", and "Styles". The main area contains a table with two columns: "Grou..." and "Titre".

Grou...	Titre
1	QC - Time when started breakfast
2	QC - Time for breakfast
3	Liste 2 (products consumed at breakfast) on 1 Week

Ces titres sont initialisées avec le nom des variables ayant provoqué la *rupture*, l'utilisateur pouvant les remplacer par le texte de son choix.

Ces titres peuvent être ensuite placés dans l'*en-tête* ou *pied de page* du tableau (voir paragraphe précédent), ce qui permet d'identifier chaque page logique individuellement.

L'onglet *Styles* permet de donner des styles (police, taille, couleur, etc..) aux différents titres (Etude, dossier, requête, pages, colonnes, lignes).

5.6.4 Tests : Gamma (Goodman et Kruskal) et Chi2.

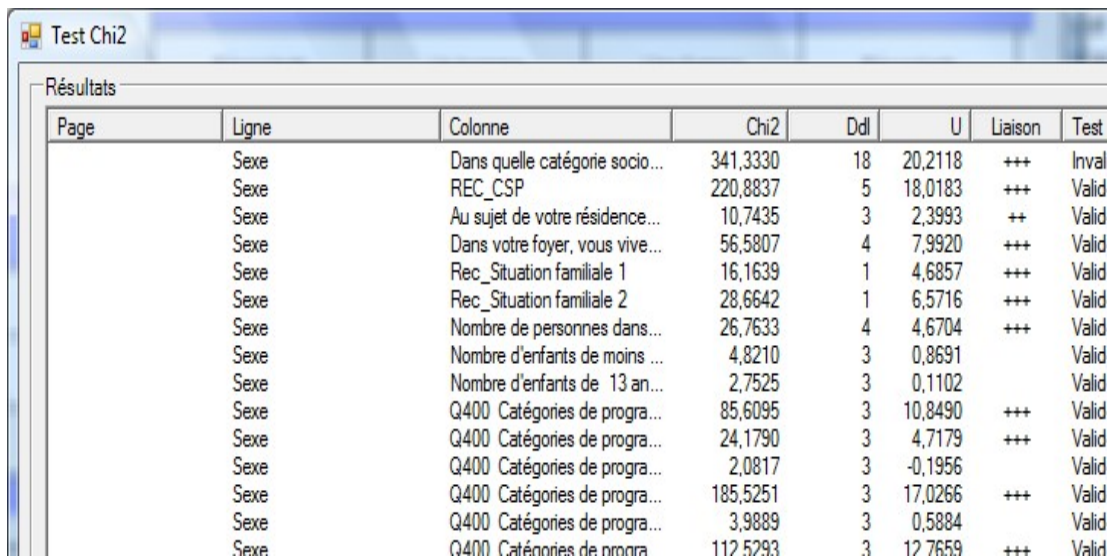
Ils s'obtiennent avec un clic droit sur la visu du tableau. On obtient dans les 2 cas une fenêtre avec les résultats (une ligne de résultat par tableau élémentaire de la demande). Le test gamma s'utilise lorsque les variables en lignes et colonnes sont ordinales.

Remarques sur le test de Chi2 : pour que le test puisse être calculé, il faut que le tableau comprenne une ligne et une colonne Total (obtenues avec l'option T du modèle, ou NT, ou A), pondérées ou non ainsi que les cases du tableau.

Il ne doit pas y avoir plus d'une option D dans les modèles en lignes et colonnes.

Enfin le format des variables en lignes et colonnes doit comprendre des valeurs individuelles, et non seulement des sous-totaux : le chi2 est calculé sur le sous-tableau composé uniquement de ces lignes et colonnes, et toutes les lignes et colonnes correspondant aux sous-totaux sont exclues du calcul du chi2.

S'il y a plusieurs totaux (pondérés ou non de même que le tableau), ils seront pris en compte dans l'ordre suivant : d'abord NT (totaux arithmétiques), à défaut T, à défaut A. Donc, en cas de variables multiples, l'utilisateur a le choix entre un calcul du Chi2 sur les totaux arithmétiques du tableau (en utilisant l'option NT), ou sur les totaux logiques (nombre de répondants, en l'absence de l'option NT).



Page	Ligne	Colonne	Chi2	Ddl	U	Liaison	Test
	Sexe	Dans quelle catégorie socio...	341,3330	18	20,2118	+++	Inval
	Sexe	REC_CSP	220,8837	5	18,0183	+++	Valid
	Sexe	Au sujet de votre résidence...	10,7435	3	2,3993	++	Valid
	Sexe	Dans votre foyer, vous vive...	56,5807	4	7,9920	+++	Valid
	Sexe	Rec_Situation familiale 1	16,1639	1	4,6857	+++	Valid
	Sexe	Rec_Situation familiale 2	28,6642	1	6,5716	+++	Valid
	Sexe	Nombre de personnes dans...	26,7633	4	4,6704	+++	Valid
	Sexe	Nombre d'enfants de moins ...	4,8210	3	0,8691		Valid
	Sexe	Nombre d'enfants de 13 an...	2,7525	3	0,1102		Valid
	Sexe	Q400 Catégories de progra...	85,6095	3	10,8490	+++	Valid
	Sexe	Q400 Catégories de progra...	24,1790	3	4,7179	+++	Valid
	Sexe	Q400 Catégories de progra...	2,0817	3	-0,1956		Valid
	Sexe	Q400 Catégories de progra...	185,5251	3	17,0266	+++	Valid
	Sexe	Q400 Catégories de progra...	3,9889	3	0,5884		Valid
	Sexe	Q400 Catégories de progra...	112,5293	3	12,7659	+++	Valid

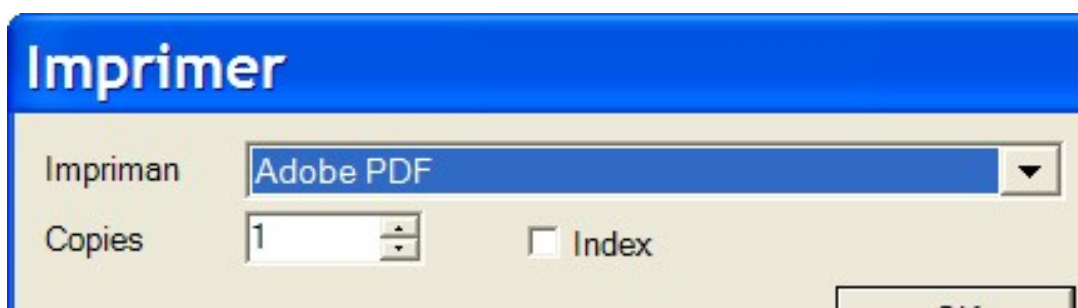
Les résultats pour chaque tableau sont : la valeur du chi2, le nombre de degrés de liberté, la valeur de U, la qualité de la liaison (+, ++ ou +++ pour une liaison à 90, 95 ou 99%), enfin l'indication de validité du test.

5.6.5 Impression du tableau.

En cliquant sur l'icône, on peut imprimer le tableau (imprimante ou fichier pdf), après avoir choisi les options d'impression.

Les options « *rupture* » éventuellement associées à des variables colonnes ou lignes généreront des sauts de page **avant** impression des résultats correspondant.

Si on coche la case « index », un sommaire sera imprimé en tête du tableau, donnant pour chaque *page logique*, (obtenue avec les options *rupture*), les titres donnés dans l'écran « *index* » (voir ci-dessus), ainsi que le numéro de page correspondant.



Remarque : l'option « portrait » ou « paysage » choisie une première fois sera conservée automatiquement pour le tableau. Il ne sera donc plus nécessaire de la redéfinir à chaque impression.

Impression de plusieurs demandes. Si on clique sur l'icône d'impression après avoir sélectionné plusieurs demandes d'un même dossier, on obtiendra l'impression des tableaux résultant de toutes ces demandes.

De même, si on demande l'impression en sélectionnant un (ou plusieurs) dossier, on obtiendra les tableaux de toutes les demandes du dossier.

Remarque : si on coche la case « index », la numérotation de l'ensemble des tableaux sera séquentielle, sans tenir compte du numéro de la première page éventuellement indiquée dans l'écran de mise en page des demandes.

Utilisation de PDF Creator. L'impression pdf standard utilise Acrobat : les temps d'exécution sont très longs, et même prohibitifs dès qu'il s'agit de centaines de pages. C'est pourquoi nous recommandons d'installer le logiciel gratuit PDF Creator, livré en même temps que CoTab, et dont les temps d'exécution sont très satisfaisants. Une fois installé, PDF Creator apparaît dans la liste des imprimantes disponibles.

5.6.6 Visualisation avant impression.

En cliquant sur l'icône, on peut visualiser à l'écran l'impression du tableau, avec les mêmes caractéristiques que celles figurant pour l'impression. L'écran « choix de l'imprimante » permet de sélectionner l'option « portrait » ou « paysage », ainsi que la taille de la page.

Cette visualisation peut concerner une ou plusieurs demandes, un ou plusieurs dossiers.

5.6.7 Exportation.

On peut choisir l'option « CSV », les options « Excel XLS » et « Excel XLSX » ou l'option HTML.

Ce choix peut être fait pour une requête ou pour un dossier.

Dans tous ces cas, on obtient l'ensemble du contenu du tableau, avec toutes les valeurs affichées (effectifs, pourcentages, etc.), et les titres et libellés.

Les options du tableau (cacher lignes et colonnes vides, cacher les pourcentages, etc..) sont prises en compte pour ces exportations.

Avec l'option XLS, (ou XLSX) on obtient de plus un fichier lisible par Excel dans lequel la mise en page est conservée (gras, couleurs, taille, etc.)

Le fichier CSV ne contient dans tous les cas qu'un seul feuillet. Il contient, en plus des titres et libellés (2^{ème} colonne), 2 colonnes de codes supplémentaires :

- la 1^{ère} colonne contient un code 0 à 3 indiquant si la ligne correspondante correspond à un titre (codes 0 ou 2) ou à des résultats (codes 1 ou 3), et s'il s'agit d'un premier étage (codes 0 ou 1) ou d'un second étage (codes 2 ou 3).
- la 3^{ème} colonne contient un code 0 à 2 indiquant à quelle ligne de sous-cellules correspondent les résultats présentés.

Le fichier XLS ou **XLSX** peut contenir plusieurs feuillets :

Si le tableau contient une variable « *page* », il y aura autant de feuillets Excel que de pages. De même, en présence de « *rupture* » dans la définition des lignes et colonnes du tableau, un nouveau feuillet sera généré à chaque apparition d'une « *rupture* ». Si le tableau comprend plus de 256 colonnes, il faut choisir l'option XLSX.

L'index est toujours fourni, sous forme d'un onglet supplémentaire « Index » donnant la liste des autres onglets.

L'option HTML peut être demandée soit sous forme mini-site, soit sous forme d'un seul fichier.

Dans le cas d'un seul fichier, le résultat est chargeable sous Excel, la mise en page étant conservée (polices, styles, etc.). Mais, contrairement au cas XML, le chargement sous Excel d'un fichier HTML réussit même avec de très gros tableaux.

Dans le cas mini-site, on obtient dans le répertoire sélectionné, un ensemble de fichiers consultables par Internet Explorer : le document étant découpé en pages selon les options break du tableau.

La légende apparaît dans la partie gauche, à partir du contenu de l'écran Table index, obtenu avec l'icône Index.

PONDERATION					
Household size					
2					
QC - Time when started breakfast					
Total	122	100	149	100	42
	100				34
Unweighted Total	149		149		51
	100				34
Before 6 am	25	20	30	20	9
	100				37
6:00 - 6:59	7	6	9	6	7
	100				100
7:00 - 7:59	68	56	83	56	25
	100				37
8:00 - 8:59	22	18	27	18	

5.6.8 Demandes faites à partir d'un dossier.

On dispose au niveau d'un *dossier*, et non seulement au niveau d'une *demande de tableau*, des icônes suivantes :

- exécution : toutes les demandes du dossier seront exécutées automatiquement
- mise en forme des tableaux : cela permet de définir la mise en forme des tableaux pour toutes les demandes du dossier
- mise en forme des pages : cela permet de définir la mise en forme des pages pour toutes les demandes du dossier
- aperçu avant impression
- impression : tous les tableaux du dossier seront imprimés comme un seul document
- exportation : tous les tableaux du dossier seront exportés dans un seul fichier

5.7 Opérations sur chaînes.

Il est possible de calculer de nouvelles colonnes ou lignes au moyen d'opérations arithmétiques portant sur d'autres lignes ou colonnes.

5.7.1 Demandes.

Dans la demande de tableau, avec l'écran permettant de sélectionner une variable pour la placer au 1^{er} ou 2^{ème} étage, on peut cliquer sur l'onglet « expression » afin de mettre à la place d'une variable un calcul portant sur différentes lignes ou colonnes (selon que l'on est dans la partie définissant les lignes ou les colonnes).

On peut ainsi définir une nouvelle colonne comme étant la somme de plusieurs autres colonnes, ou le rapport de 2 colonnes, etc...

Ces nouvelles lignes et/ou colonnes résultats de calculs doivent être créées avant d'effectuer la ventilation (exécution de la demande de tableau), même si le détail des calculs peut être modifié après la ventilation.

L'écran obtenu en choisissant l'onglet *expression* contient une ligne titre, et une zone dans laquelle on écrit l'expression en format libre.

The screenshot shows a software interface window titled "Variables / Expressions". It has two tabs: "Variables" and "Expressions", with "Expressions" being the active tab. Below the tabs, there is a label "Titre:" followed by a text input field containing the text "SQRT(L3/L2*100) AAA". Below this, there is a larger text area containing the text "SQRT(L3/L2*100)".

Le **titre** sera présent dans les tableaux au même niveau que le titre des variables.

L'expression sera **une formule arithmétique**, (avec des parenthèses éventuellement) faisant intervenir :

- des **numéros de lignes** existantes, écrits sous la forme **Lnn** (nn étant le numéro figurant à gauche dans une ligne de la demande, que l'on soit dans la partie définissant les *lignes* ou les *colonnes* du tableau).

Remarques :

- les lignes nnn référencées doivent contenir des **variables**, et non pas des opérations.
- si l'opération est au 1^{er} étage, elle ne doit pas être suivie d'un 2^{ème} étage, et les lignes appelées doivent être au premier étage également sans 2^{ème} étage.
- si l'opération est au 2^{ème} étage, les lignes appelées doivent être au 2^{ème} étage dans le même groupe.
- les lignes appelées doivent correspondre à des variables avec un modèle contenant **une seule option**, cette **option n'étant ni D, ni LO**.
- attention en cas de suppression, insertion, ou déplacement de lignes, il faut modifier les numéros de lignes dans les formules (contrairement à ce qui est fait pour les bases de %).
- on peut référencer, non pas une ligne (ou colonne) tout entière, mais un simple élément dont on connaît la position dans le tableau tel qu'il apparaîtra à l'écran : **L3/I1** référencera non pas tous les éléments de L3 successivement, mais celui situé dans la 1^{ère} colonne du tableau résultat (en cas d'opérations sur lignes) ou la 1^{ère} ligne (en cas d'opération sur colonnes). Il faut prendre garde que cet indice ne tient pas compte des non éditions éventuelles des lignes ou colonnes blanches.
- des **constantes** entières ou réelles.
- des **opérateurs** : + - * /
- des **fonctions** :

ABS : valeur absolue

SQRT : racine carrée

ROUND : arrondi

TRUNC : troncature

RANDOM : nombre au hasard, compris entre 0 et 1

EXP : exponentielle

LOG : logarithme népérien

ISP : retourne 1 si l'argument est positif ou nul, 0 sinon

ISZ : retourne 1 si l'argument est égal à zéro, 0 sinon

ISSP : retourne 1 si l'argument est strictement positif, 0 sinon

Remarque : avant de fermer l'écran (bouton *close*), il faut valider l'expression avec les boutons *placer* ou *ajouter*.

Modèle associé.

Il faut associer à l'expression un modèle qui servira à récupérer un titre pour l'édition (sur la ligne correspondant aux libellés des variables), le nombre de décimales, et l'indication « non édition ». Les autres champs du modèle (Option, %, pondération) seront ignorés. Ce modèle doit avoir une seule ligne.

5.7.2 Exécution.

L'exécution se fera après la ventilation, et après le calcul des moyennes et pourcentages. **On pourra donc modifier la formule de calcul après avoir visualisé le tableau résultat.**

Les opérations se font en à partir des effectifs (cas général), des cumuls (option C du modèle), ou des moyennes (option M du modèle) de la case opérande concernée.

On applique la formule demandée. Le résultat sera vide dans l'un des cas suivants :

- calcul impossible (division par zéro, etc)
- un des opérandes est déjà vide

L'ordre des opérations est : d'abord lignes puis colonnes.

6. Exportations de fichiers.

6.1 Généralités.

On peut exporter tout ou partie des variables de l'étude dans des fichiers de différents formats :

- format CoTab,
- format SPSS (*sav*)
- format Cosi
- format Spad (*sda*)

Le **format CoTab** générera 3 fichiers :

- le fichier de données d'extension *.asc*, ASCII, en format fixe ou délimité
- un fichier d'extension *.sdb*, c'est-à-dire une nouvelle étude CoTab
- un fichier d'extension *.map*, donnant la liste des variables du fichier et leur position.

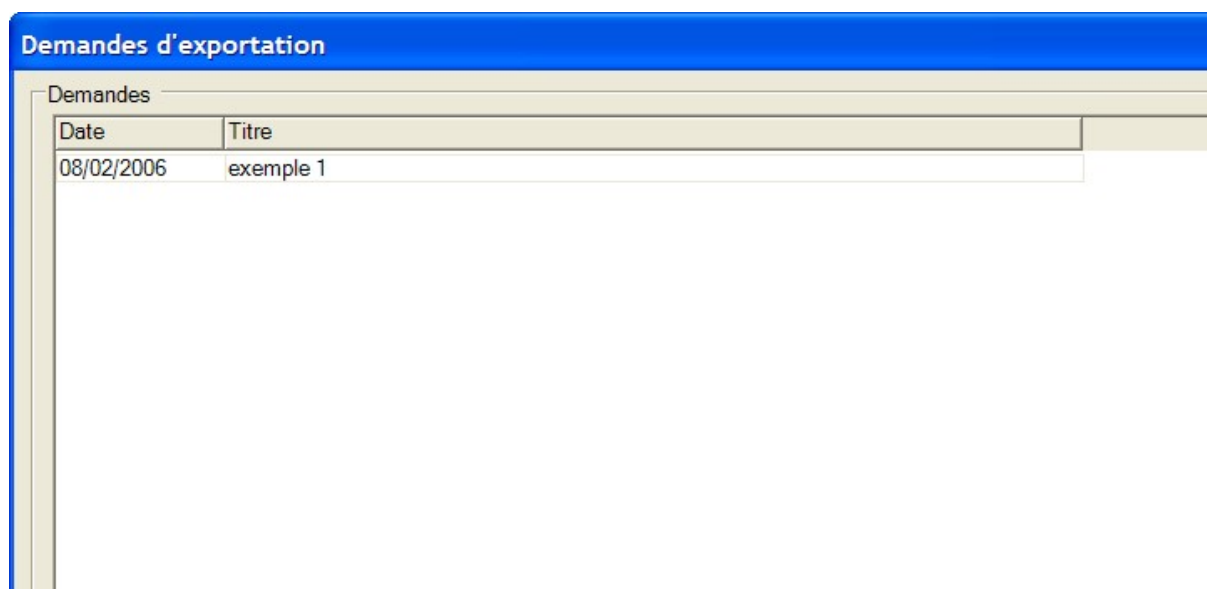
L'utilisateur dispose ainsi d'une étude CoTab lui permettant de traiter immédiatement avec CoTab le fichier exporté.

Le **format Cosi** générera 2 fichiers :

- le fichier de données d'extension *.asc*, ASCII, en format fixe
- un fichier d'extension *.sct*, c'est-à-dire le script de Cosi (dictionnaire des variables).

6.2 Gestion des demandes.

En cliquant sur *export données* dans le menu *outil*, on obtient l'écran suivant :



On peut créer des demandes, les modifier, les dupliquer, les supprimer, enfin les exécuter.

Une demande correspond à l'écran ci-après :

- titre : titre quelconque
- univers : peut contenir le nom d'une variable logique : seuls les individus ayant la valeur *vraie* pour cette variable seront exportés
- choix de l'onglet, selon que l'on veuille exporter un fichier CoTab (ASCII), Cosi, ou SPSS (*sav*), ou SPAD (*sda*).

6.2.1 Dans le cas CoTab :

il faut donner :

- L'emplacement du fichier créé (il ne faut pas fournir de nom de fichier)
- Le numéro de l'étude créée
- Le format ASCII ou *csv* : dans le cas *csv*, il faut donner le séparateur et le délimiteur
- Si on coche le bouton « *code en libellés* », ce sont les libellés dans les formats qui seront exportés, et non les valeurs des variables (quand le format existe).

Demande d'export

Titre : exemple 1

Univers :

VoxcoTab | SPSS | Cosi

Chemin : C:\Documents and Settings\Jean SOUSSELIER\Bureau

Numéro : 8 ?

Format : ASCII Séparateur : <None> Délimiteur : <None> Code en libellés

N°	Variable	Lg	Déc	Par val...	Nombre...	Rebut
1	AG2	1				
2	DSDS	1		<input type="checkbox"/>	2	
3	AGR7BIS	1		<input type="checkbox"/>	9	
4	AGER	2				
5	AGER7	2				
6	OCC	1				
7	AGEHOH	1				
8	HOHSIZ	1				
9	VA1	1		<input checked="" type="checkbox"/>	1	

Les fichiers créés s'appelleront (nnn étant le numéro indiqué dans la demande) : :

- E_nnn.asc
- E-nnn.sdb
- E-nnn.map

On peut alors choisir les variables que l'on veut exporter en cliquant sur le bouton dans le champ « variable » : on obtient l'écran suivant, permettant de sélectionner une ou plusieurs variables.

Choisir une variable			
N°	Nom	Type	Titre
8	SEX	Simple	Sex
9	HOMMES	Logique	
10	AGER	Simple	Age of respondent in clear
11	AGER7	Simple	Age of respondent
12	cccc	Simple	
13	AGE57	Logique	
14	AGER7B...	Multiple	Sex and Age of respondent
15	OCC	Simple	Occupation of the Head of household
16	AGR7BIS	Multiple	Sex and Age of respondent
17	F50	Logique	FEMMES DE MOINS DE 50 ANS
18	AGEHOH	Simple	Age of the Head of household
19	HOHSIZ	Simple	Household size
20	NIELSE	Simple	Nielsen regions
21	MSIZE	Simple	Market Size
22	PRFS	Multiple	Presence of children aged 17 or below in the househ

On renseigne pour chaque variable :

- la longueur du champ (inutile dans le cas d'un fichier csv)
- le nombre de décimales
- pour les variables multiples : l'option **par valeur** (succession des valeurs dans des champs successifs) ou **par rang** (succession de valeurs 0 ou 1, la position des 1 correspondant aux valeurs de la variable **1 à n**, n étant le nombre de valeurs demandées ; les valeurs extérieures à cet intervalle, en particulier la valeur 0, sont ignorées).
- pour les variables multiples : le nombre de valeurs à considérer
- la valeur à éditer pour le Rebut (blanc par défaut)

Attention pour les variables multiples avec l'option par valeurs dans le cas csv :

- s'il n' a pas de délimiteur, les valeurs seront placés dans des champs successifs
- s'il y a un délimiteur, les valeurs seront dans le même champ, séparées par le séparateur.

Pour les variables multiples avec l'option par rangs dans le cas csv : la suite de 0 et de 1 est placée dans le même champ.

Pour les variables dimensionnées, les occurrences sont placées automatiquement les unes à la suite des autres.

Option « taille automatique » : en appuyant sur le bouton « taille auto », toutes les valeurs indiquées ci-dessus (longueur des champs, nombre de valeurs,..) seront automatiquement fournies par le programme, qui lira pour cela le fichier de données afin de trouver les valeurs adéquates. La seule indication à fournir concerne le choix « par valeurs » ou « par rang » pour les variables multiples). L'utilisateur peut modifier ces valeurs s'il le souhaite.

Cas des études à niveau.

Les variables sélectionnées peuvent appartenir à différents niveaux.

Une nouvelle case peut être cochée : « *multi-niveaux* ».

- **Si cette case n'est pas cochée**, cela signifie qu'on exportera les individus d'un seul niveau, celui correspondant au numéro le plus élevé des variables sélectionnées. Les variables ayant un numéro de niveau inférieur seront également exportées, mais l'étude nouvelle créée sera sans niveau.
- **Si cette case est cochée**, cela entraîne que la nouvelle étude créée sera multi-niveaux, et le fichier ASCII créé sera multi-enregistrements (un par niveau). L'exportation ajoutera automatiquement dans chaque enregistrement, avant les variables explicitement demandées, le code niveau et les identifiants des niveaux.

6.2.2 Dans le cas SPSS :

il faut donner le nom et l'emplacement du fichier créé.

Il ne faut pas fournir l'extension, elle sera fournie par le programme.

Demande d'export

Titre: EXPORT SPSS

Univers:

VoxcoTab: SPSS

Chemi: C:\COTAB\IMPORTS4\XXX.sav

N...	Variable	Lg	Déc	Par val...	Nombre...	Rebut
1	P	2				
2	AGER3	2				
3	ager3bis	2				
4	SEX	2				
5	MEN	2				
6	AGER	2				
7	AGER7	2				
8	AGE57	2				
9	AGER7BIS	2		<input type="checkbox"/>		

On sélectionne les variables comme dans le cas précédent CoTab, et on renseigne les mêmes champs (longueur des champs, nombre de décimales, etc...).

Le fichier obtenu peut s'ouvrir avec SPSS.

Remarques :

- pour les variables multiples : l'option **par valeur** (succession des valeurs dans des champs successifs) ou **par rang** (succession de valeurs 0 ou 1, la position des 1 correspondant aux valeurs de la variable **0 à (n-1)**, n étant le nombre de valeurs demandées ; les valeurs extérieures à cet intervalle sont ignorées).

- les variables multiples sont décomposées en autant de variables que de valeurs demandées

- les variables dimensionnées donnent naissance à autant de variables qu'il y a d'occurrences.
- l'option « **taille automatique** » fonctionne comme dans le cas précédent.

6.2.3 Dans le cas SPAD:

Le fonctionnement est le même qu'avec SPSS; il faut sélectionner l'extension sda pour le fichier à créer.

6.2.4 Dans le cas Cosi :

Il faut donner le nom et l'emplacement des fichiers script (.sct) et données (.asc).

Export Request

Title: Export Cosi

Universe:

VoxcoTab | SPSS | Cosi

SCT File: C:\COTAB\IMPORTS4\new.sct

Data File: C:\COTAB\IMPORTS4\new.asc

N°	Variable	Lg	Dec	By Valu...	Nb Valu...	NA
4	AGER3	1				
5	AGER	2				
6	AGER2	1				
7	SEX	1				
8	MEN	1				
9	YOUNG_MEN	1				
10	VENT1	1		<input checked="" type="checkbox"/>	1	
11	VA1	1		<input checked="" type="checkbox"/>	1	
12	QCI	1		<input type="checkbox"/>	4	
13	QCDUCI	1				
14	QDI	1		<input type="checkbox"/>	1	
15	V1C12	1		<input type="checkbox"/>	11	

On sélectionne les variables comme dans le cas précédent CoTab, et on renseigne les mêmes champs (longueur des champs, nombre de décimales, etc...).

L'option « **taille automatique** » fonctionne comme dans le cas CoTab.

Remarque : dans le cas d'exportation Cosi, il ne faut pas exporter d'occurrences de variables dimensionnées, car Cosi ne saurait pas les importer.

7. Redressement d'échantillon.

7.1 Principe.

Ce module calcule, pour chaque individu, un coefficient de redressement ou poids, tel que la structure, selon un ou plusieurs critères, de l'échantillon pondéré par ces coefficients soit conforme à la structure théorique supposée connue.

La procédure de calcul est itérative.

Le redressement peut porter soit sur des nombres d'individus, soit sur une quantité qui leur est associée.

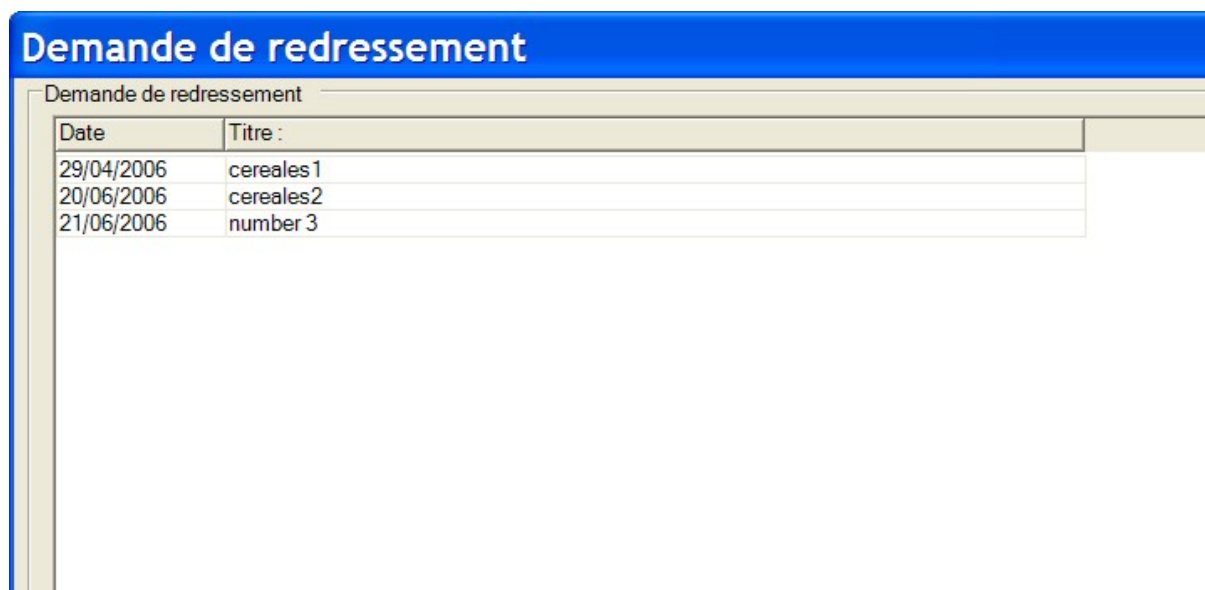
Le résultat est fourni sous forme d'un fichier contenant le poids pour chaque individu traité. Ce fichier devra alors être déclaré en table externe dans l'étude, afin de pouvoir importer le poids.

Les éléments indispensables sont les variables :

- **Pondération**, de type Quantité, (recevant le coefficient de redressement),
- **Identification**, de type texte, (pour identifier les individus dans le fichier résultat),
- **Critères** de redressement, de type simple, et qui seront donnés individuellement ou en couples (critères croisés).

7.2 Demande.

En cliquant sur *Redressement* dans le menu *utitaires*, on obtient l'écran suivant :



Date	Titre :
29/04/2006	cereales1
20/06/2006	cereales2
21/06/2006	number 3

On peut créer des demandes, les modifier, les dupliquer, les supprimer, enfin les exécuter. Une demande correspond à l'écran ci-après :

- **titre** : titre quelconque, qui sera utilisé dans le nom des fichiers créés
- **onglets** : il y a un onglet dans lequel on donne les options générales de la demande, et un onglet dans lequel on donne les critères de redressement, et les objectifs pour chaque critère.

Demande de redressement

Titre :

Options | Critères

Global

ID :

Base :

Poids :

Poids initial :

Quantité à ajuster :

Paramètres

Nombre max d'itéra	<input type="text" value="300"/>	Seuil :	<input type="text" value="0,100"/>
Poids nul si inférieur à :	<input type="text" value="0,00000"/>	Un poids nul est considéré :	<input type="text" value="NA"/>
Limites des poids Min :	<input type="text" value="0,00000"/>	Max :	<input type="text" value="99999"/>

7.2.1 Onglet « options » :

- **ID** : obligatoire, de type texte, pour identifier les individus dans le fichier des poids
- **Base** : facultative, recevant un nom de variable logique, déterminant l'univers sur lequel faire porter le redressement
- **Poids** : obligatoire, de type quantité, recevra le poids créé pour l'individu
- **Poids initial** : facultatif, de type quantité. Les valeurs de cette variable serviront de poids initiaux pour démarrer les itérations. Cette option est introduite à la demande de certains utilisateurs. Pour notre part, nous ne conseillons pas de l'utiliser.
- **Quantité à ajuster** : facultatif, de type Quantité. Si ce champ est renseigné, le redressement portera non pas sur des nombres (ou pourcentages) d'individus, mais sur des cumuls (ou pourcentages de ces cumuls) de cette quantité. Par exemple, si on traite une enquête portant sur des exploitations agricoles, on pourra redresser non pas sur des nombres d'exploitation de telle région et telle classe de superficie, mais sur des quantités de blé produites, selon la région et la classe de superficie.
- **Nombre maximum d'itérations** (300 par défaut) : il est conseillé de laisser cette valeur par défaut, sauf cas très particuliers.
- **Seuil** : écart entre valeurs recherchées et obtenues au-dessous duquel le processus itératif est arrêté. Il est conseillé de conserver la valeur par défaut (0.1).
- **Poids nul si inférieur** : les poids inférieurs à cette valeur seront remplacés par zéro.

- **Un poids nul est considéré**: les poids nuls peuvent être mis à zéro ou à Rebut.
- **Minimum-maximum** : pour limiter la valeur des poids à ces valeurs mini ou maxi.
- **Cibles en** : valeurs ou pourcentages. Les objectifs fournis pour les différents critères seront donnés en valeurs ou pourcentages.
- **Extrapolation** : si cette case n'est pas cochée, le redressement conduira à un effectif total inchangé par rapport à l'échantillon traité. Si cette case est cochée, on doit alors donner le nombre d'individus (ou le total de la quantité si elle existe) que l'on veut obtenir après redressement. Il faut noter que si la case est cochée, on doit donner un nombre d'individus, même si on a demandé l'option pourcentages.
 - Il est obligatoire de cocher cette case s'il y a une quantité à ajuster.
 - Mais il est inutile de cocher cette case quand les objectifs sont donnés en valeurs.

Cas des études à niveau :

Le redressement se fait à un seul niveau, qui est par convention le niveau de la variable ID. Il est alors nécessaire que la variable Poids soit du même niveau, ainsi que les variables Poids initial et Quantité si elles existent. Les critères doivent être de niveau supérieur ou égal (c'est-à-dire de numéro plus petit ou égal).

7.2.2 Onglet Critères.

La boîte « Références » permet de sélectionner un ou plusieurs critères, soit individuels, soit croisés.

Ces critères doivent être des variables de **type simples, non dimensionnées (ou occurrences d'une variable dimensionnée), et elles doivent avoir un format.**

Les objectifs seront fournis pour toutes les modalités simples du format, c'est-à-dire celles correspondant à un seul code (les lignes du format correspondant à des regroupements de codes seront ignorées).

Demande de redressement

Titre : cereales1

Options Critres

Variables : SEX

...	-
Male	45
Female	55
NA	0
Total	100

Les objectifs seront donnés pour toutes les modalités ainsi affichées, ainsi que pour la valeur Rebut.

Des copier/coller sont possibles dans la zone des objectifs.

Case à cocher : Conserver valeurs du R . Si cette case est cochée, on ne donne plus d'objectifs pour le rebut (du reste il n'est plus affiché), mais le programme calculera un coefficient de redressement conservant dans l'échantillon redressé le pourcentage du rebut trouvé dans l'échantillon non redressé.

Cas des critères croisés.

Dans ce cas, il faut donner les objectifs dans toutes les cases du tableau issu du croisement des 2 variables sélectionnées.

Demande de redressement

Titre : cereales2

Options Critres

Variables : AGER3 - SEX

...	Male	Female	NA	Total
13 - 17...	0	0	0	0
18 - 24...	0	0	0	0
25 - 34...	0	0	0	0
35 - 44...	0	0	0	0
45 - 54...	0	0	0	0
55 - 64...	0	0	0	0
65 and...	0	0	0	0
NA	0	0	0	0
Total	0	0	0	

Remarque : dans le cas d'un critère croisé, on ne peut plus cocher la case *Conserver valeurs du R*.

7.3 Exécution.

L'exécution faite, on obtient 2 fichiers (nnn : numéro de l'étude, ddddddd : titre de la demande) :

- le fichier des poids : Wnnn-dddddd.txt
- le fichier compte-rendu : Wnnn-dddddd.log

Le fichier des poids contient une ligne par individu appartenant à l'univers choisi, avec l'ID et le poids trouvé pour cet individu. Il faudra alors introduire ce fichier comme **table externe csv**, avec la variable ID comme clé associée au premier champ de la table, et lire le poids au moyen de la fonction « importsCSV » dans le deuxième champ.

Les individus dont l'ID, ou le poids initial sont en Rebut ne figurent pas dans ce fichier.

Le fichier log donne le compte-rendu de l'exécution, c'est-à-dire le rappel des options, et, pour chaque critère, les colonnes objectifs, valeurs initiales dans l'échantillon, valeurs redressées. Il contient aussi le **coefficient d'efficacité**, calculé ainsi :

$$\text{Efficacité} = 100 * (\sum \text{poids})^2 / n \sum \text{poids}^2$$

```

W7-cereales1.log - Bloc-notes
Fichier Edition Format Affichage ?

Coefficient de redressement      22/10/2014 13:18:14
Poids: P Quantité "WEIGHT"
NBQ: 500
Poids minimum rencontré: 0.55689
Poids maximum rencontré: 2.24860
Poids minimum: 0.00000
Poids maximum: 99999999.00000
Nul si Rebut: 65
Nombre maximum d'itérations: 300
Itérations faites: 8
Taille de l'ID: 4
Longueur du poids: 14
Efficacité: 86.95
SEX simple Sex
-----
Male                               Demandée      Initial      Pondérat
Female                             45.00000     45.80000     45.0767
NA                                  55.00000     54.20000     54.9232
AGER7 simple Age of respondent
-----
1 - 4 years old                     Demandée      Initial      Pondérat
5 - 7 years old                     0.00000     0.00000     0.0000
8 - 12 years old                    0.00000     0.00000     0.0000
13 - 17 years old                   0.00000     13.00000     0.0000
18 - 24 years old                   15.00000     7.20000     14.9904
25 - 34 years old                   20.00000     13.40000     19.9630
35 - 44 years old                   20.00000     16.20000     19.9910
45 - 54 years old                   20.00000     15.00000     19.9788
55 - 64 years old                   15.00000     24.40000     15.0247
65 and over                         10.00000     10.80000     10.0518
NO ANSWER                           0.00000     0.00000     0.0000
NA                                  0.00000     0.00000     0.0000

```


8. Modèle prédictif.

8.1 Principe.

Le problème de la **prédiction de données** se pose dans de nombreux cas :

- **données manquantes** dans une enquête,
- évaluation d'un risque de désabonnement, de défaillance, etc., pour une population de clients (**scoring**)
- **fusion de données**, afin d'enrichir un fichier client par exemple avec les données recueillies par enquête sur un échantillon. Il ne s'agit pas de créer de nouveaux individus, mais de calculer des variables non renseignées pour des individus.

Pour traiter ce problème, on part d'une **étude d'apprentissage** (ou de construction du modèle), dans lequel sont renseignées, d'une part, les variables qu'on cherche à prédire (**variables à expliquer**), d'autre part un ensemble de **variables explicatives**.

Un **modèle de données** est alors construit, ce modèle permettant par la suite de calculer les variables à expliquer pour tout individu pour lequel les variables explicatives sont renseignées.

Ce modèle utilise une **méthode exclusive**, appelée RESCUT, fonctionnant dans le cas où les **variables explicatives sont toutes nominales**, (c'est-à-dire simples, multiples ou textes, mais toutes avec Format), les **variables à expliquer pouvant être de type quelconque** (sauf Texte).

Ce modèle peut être utilisé pour calculer les variables à expliquer, soit dans la même étude (sur le même fichier de données ou sur un autre fichier de données), soit dans une autre étude dans laquelle les mêmes variables explicatives existent.

8.2 Réalisation.

Cette prédiction de données se fait à l'aide de 2 modules :

8.2.1 **Module de construction.** On déclare dans ce modèle :

- **les variables explicatives**, qui doivent être nominales, c'est-à-dire de type Simple ou Multiple, ou Texte, dans ces 3 cas avec un format. Les non réponses sont admises.
- **la (ou les) variable(s) à expliquer**, qui peuvent être de type Quantité, ou Simple sans format, ou Simple avec format, ou Multiple avec format, ou Logique.
L'exécution fournit alors le modèle prédictif, stocké automatiquement dans un fichier.

Attention: si la variable à expliquer est Simple, et s'il s'agit d'une échelle plutôt que d'une variable strictement nominale, il vaut mieux la définir **sans format** pour la construction du modèle.

Si la variable à expliquer est Multiple, les codes existant dans le format doivent être 1,2,3 etc. **Le code 0 est interdit**, de même que des codes ne commençant pas par 1.

Remarque : quand la variable à expliquer est Simple ou Multiple avec format, la variable prédite aura exclusivement comme valeurs les valeurs figurant dans ce format (a, b, c, ...), ou bien la valeur *Rebut*.

8.2.2 **Module d'utilisation.** On peut appeler ce module soit dans l'étude de construction, soit dans une autre étude. Dans ce dernier cas, le dictionnaire de variables doit contenir les variables explicatives et à expliquer avec les mêmes attributs (niveau, type, dimension et format) que dans l'étude d'apprentissage.

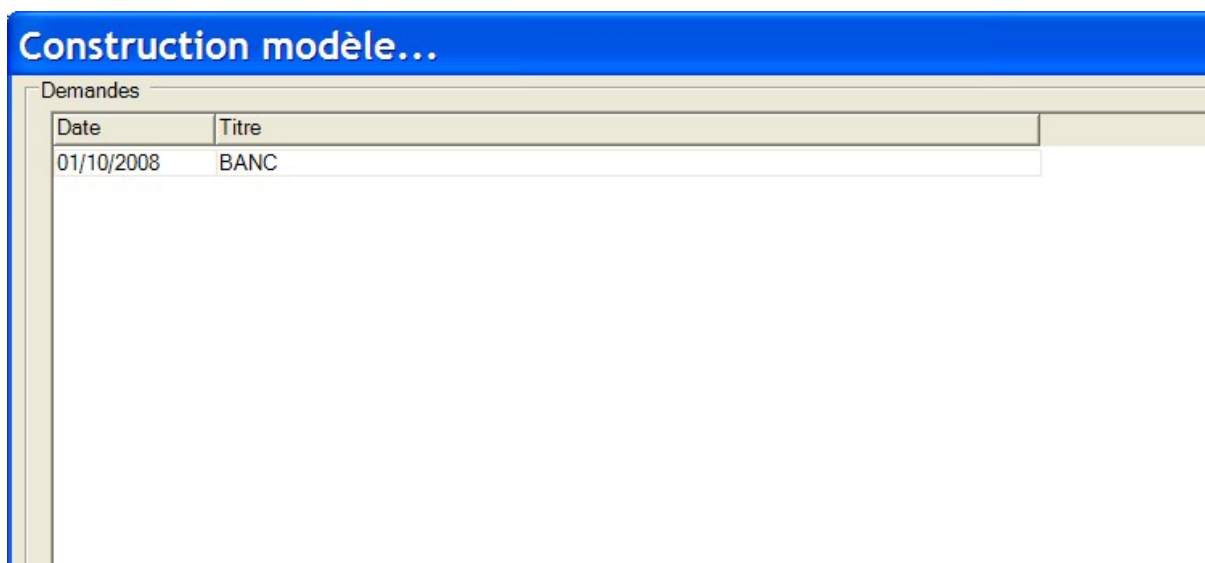
L'exécution de ce module fournit un fichier avec, pour chaque individu, les valeurs des variables à prédire. Si on veut introduire ces variables dans l'étude pour construire des tableaux par exemple, il suffit de déclarer ce fichier comme une table externe, et de définir les fonctions ImportCsv pour toutes les variables à prédire.

8.2.3 Remarque : Le modèle fonctionne même quand il y a des non-réponses pour les variables à expliquer. Cependant, ce taux de non-réponses sera reproduit dans la prédiction, ce qui est légitime si la non-réponse à une variable à expliquer a une signification précise (par exemple « ne possède pas »), mais cela ne l'est pas si la non-réponse correspond à « ne sait pas ».

Dans ce dernier cas, on aura intérêt à construire le modèle **en filtrant la demande sur les répondants à la variable à expliquer.**

8.3 Construction du modèle prédictif.

En cliquant sur *construction du modèle* dans le menu *utilitaires*, on obtient l'écran suivant :



On peut créer des demandes, les modifier, les dupliquer, les supprimer, enfin les exécuter. Une demande correspond à l'écran ci-après :

- **titre** : titre quelconque, qui sera utilisé dans le nom des fichiers créés
- **Base** : facultative, recevant un nom de variable logique, déterminant l'univers sur lequel construire le modèle
- **Poids** : facultatif, de type quantité, si on veut que le modèle soit construit en tenant compte du poids de chaque individu.

Model request

Titre : BANC

Général

Base:

Poids : Redressement

Variables

Variables explicatives		Variables à expliquer	
N°	Nom	N°	Nom
1	Sexe	1	Bancarisation
2	Age		
3	PCS		
4	Revenu		
5	segment		
6	Secteur		
7	FREQ_guichet		
8	FREQ_automates		
9	Anciennete		

Variables explicatives : en cliquant sur le bouton à droite du nom, on peut choisir des **variables nécessairement avec format** (de type Simple, Multiple ou Texte), éventuellement dimensionnées.

Variables à expliquer : en cliquant sur le bouton à droite du nom, on peut choisir une ou plusieurs **variables de tout type (sauf Texte)**, avec ou sans format, éventuellement dimensionnées (voir restrictions en 8.2.1).

S'il y a plusieurs variables à expliquer, il y aura un modèle construit pour chaque variable.

L'exécution d'un modèle sur un fichier de données fournira, dans le répertoire de l'étude, un fichier nommé **Mnnn_titre.txt**, contenant le modèle de toutes les variables à expliquer. (nnn étant le numéro de l'étude, titre étant le titre de la demande)

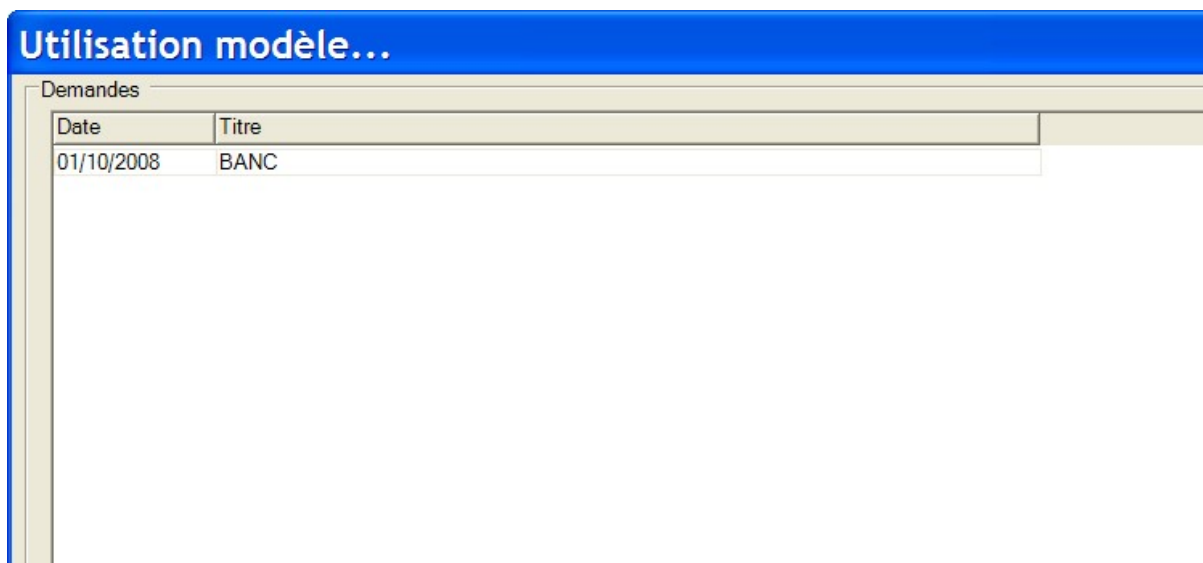
Cas des études à niveaux :

Dans le cas d'une étude à niveaux, le niveau du modèle est par convention celui de la variable de plus bas niveau figurant dans la demande (c'est-à-dire de numéro le plus élevé). Le modèle sera construit sur les individus de ce niveau.

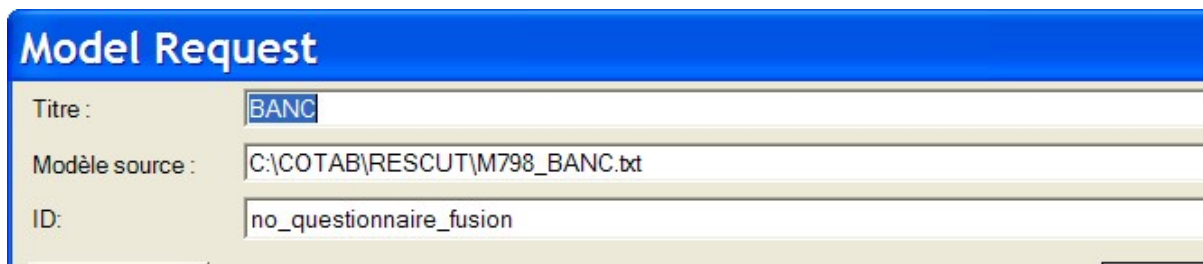
Il faut que toutes les variables à expliquer appartiennent à ce niveau, sous peine de résultats imprévisibles.

8.4 Utilisation du modèle prédictif.

En cliquant sur *utilisation du modèle* dans le menu *utilitaires*, on obtient l'écran suivant :



On peut créer des demandes, les modifier, les dupliquer, les supprimer, enfin les exécuter. Une demande correspond à l'écran ci-après :



On doit donner le nom du fichier contenant le(s) modèle(s) à utiliser , ainsi qu'une variable permettant d'identifier les individus.

L'étude d'utilisation doit contenir les variables à expliquer (renseignées bien sûr) et les variables explicatives, avec les mêmes attributs que dans le module de construction.

L'exécution fournira, dans le répertoire de l'étude, un fichier csv, de nom **Rnnn_titre.csv**, (nnn étant le numéro de l'étude d'utilisation, titre étant le titre de la demande), avec, pour chaque individu :

- premier champ : valeur de la variable ID
- champs suivants : valeurs des variables à expliquer, dans l'ordre dans lequel elles ont été déclarées dans le module de construction. (Ecriture **par rangs** pour les variables multiples).

Il faudra ensuite déclarer ce fichier comme une table externe csv, et déclarer une fonction Import csv pour toutes les variables à expliquer.

8.5 Simulation.

8.5.1 Principe.

Ce nouveau module, figurant dans le menu « Utilitaires », répond à la problématique suivante : on se pose souvent la question, dans les études de satisfaction par exemple, de savoir « quelle est la contribution des items à la satisfaction globale ? ».

Intuitivement, on comprend bien l'interrogation sous-jacente, il n'empêche que cette formulation nous paraît floue et impropre.

Aussi, nous préférons la remplacer par cet énoncé beaucoup plus précis et non contestable : que se passe-t-il (what-if ?) si, pour chaque item successivement et indépendamment, les réponses étaient toutes les mêmes. Par exemple, si l'étude de satisfaction comprend comme questions d'une part la satisfaction globale, d'autre part des items comme « rapidité », « amabilité », « qualité de la réponse », etc..., on cherche à savoir ce que devient la satisfaction globale si par exemple tous les interviewés sont tout à fait satisfaits de la rapidité du traitement.

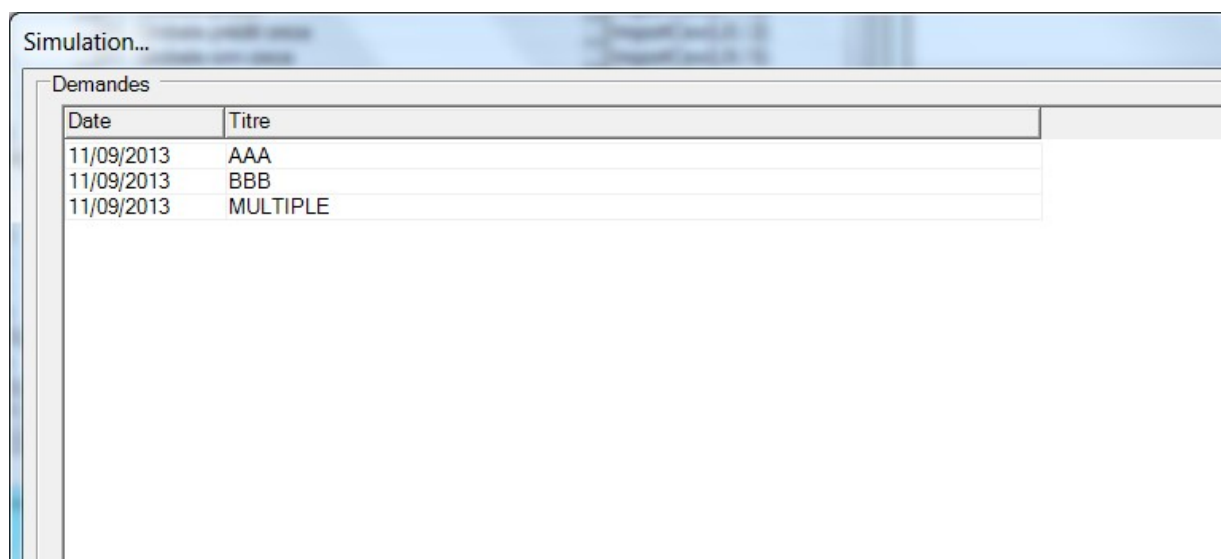
Cette manière de faire donne donc l'influence de chaque item, toutes choses égales par ailleurs.

Pour faire une simulation, il faut donc que l'on ait **une variable à simuler** (de type Quantité, ou Simple sans format, ou Simple avec format, ou Multiple avec format, ou Logique), et **des variables explicatives nominales**, c'est-à-dire de type Simple ou Multiple, ou Texte, dans ces 3 cas avec un format.

Bien entendu, la simulation est basée sur **un modèle prédictif construit auparavant**.

8.5.2 Mise en œuvre.

Pour faire une simulation, il faut d'abord construire le modèle prédictif (menu « utilitaires ») correspondant aux variables à expliquer et explicatives concernées. Ensuite, en cliquant sur la ligne « simulation » dans le menu « utilitaires », on obtient un écran permettant de créer des demandes de simulation.



The screenshot shows a window titled 'Simulation...' with a table of simulation requests. The table has two columns: 'Date' and 'Titre'. There are three rows of data.

Date	Titre
11/09/2013	AAA
11/09/2013	BBB
11/09/2013	MULTIPLE

On peut créer des demandes, les modifier, les dupliquer, les supprimer, enfin les exécuter.

Une demande correspond à l'écran ci-dessous. On doit donner :

- le nom du fichier, issu de la construction, contenant le(s) modèle(s) à utiliser,
- une variable ID permettant d'identifier les individus
- pour chaque variable explicative listée, la « cible », c'est-à-dire la valeur à donner à cette variable pour tous les individus du fichier.

Remarques :

- 1) S'il y a **plusieurs** variables à expliquer dans le modèle prédictif, la simulation ne concerne que la 1^{ère}.
- 2) Les variables explicatives **multiples** ne sont pas étudiées dans la simulation.
- 3) Les variables explicatives de type **texte** sont bien prises en compte, mais il faut écrire avec soin les textes des modalités-cibles (par exemple « Très satisfait »)

Simulation	
Titre :	AAA
Modèle source :	C:\COTAB\COURS\M024_GLOBALE.txt
ID:	RANGT
Name	Target
OnceJ	1
Delai	1
Ecoute	1
Prise	1
Qualite	1
Clarte	1
Rapidite	1
Rebond	1
Pertinence	1
Adhésion	1
OnceTr	1
Image	1

L'exécution fournira, dans le répertoire de l'étude, un fichier csv, de nom **Snnn_titre.csv**, (nnn étant le numéro de l'étude d'utilisation, titre étant le titre de la demande), avec, pour chaque individu :

- premier champ : valeur de la variable ID
- champs suivants : valeurs de la variable simulée, pour chaque simulation faite, une simulation correspondant à son calcul prédit quand une variable explicative prend, pour tous les individus, la valeur indiquée dans le champ « target » (cible).

On pourra ensuite déclarer ce fichier comme une table externe csv, et déclarer une fonction Import-csv pour obtenir la valeur de la variable pour toutes les simulations. La comparaison de toutes ces valeurs permettra de mesurer l'action respective de tous les items explicatifs.

9. Export CoTab Reader.

9.1 Principe.

Ce module du menu Utilitaires a pour but de préparer pour un client final un outil appelé Cotab Reader lui permettant de visualiser les tableaux de l'étude, et de composer tous les autres tableaux qu'il souhaite.

L'outil ainsi livré au client final comprend :

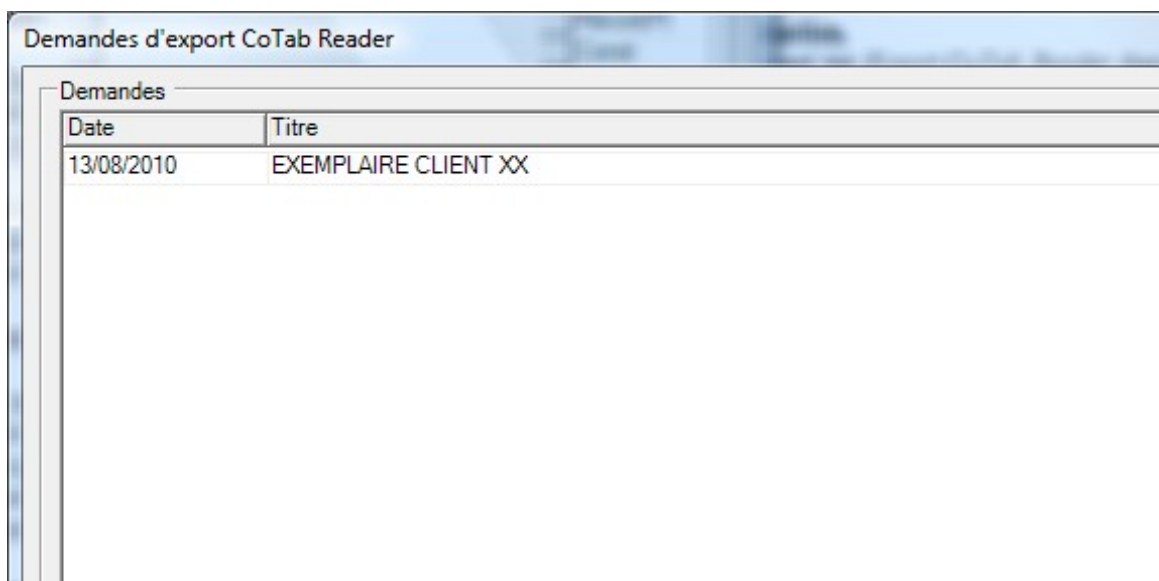
- une base de données constituée d'une partie des individus de l'étude, une partie des variables, une partie des tableaux
- le logiciel, sous-ensemble de CoTab, permettant de construire de nouveaux tableaux.

Ce logiciel CoTab Reader est libre d'emploi (aucune licence n'est nécessaire), mais il ne permet de travailler que sur l'étude associée, et il ne donne pas accès aux modules Exports de données, Redressement, et Modèle prédictif.

Les variables de la base ne peuvent pas être modifiées par l'utilisateur de CoTab Reader, mais en revanche il pourra construire de nouvelles variables à partir des variables existantes et de la plupart des fonctions de CoTab.

9.2 Réalisation.

En cliquant sur *Export CoTab Reader* dans le menu *utilitaires*, on obtient l'écran suivant :



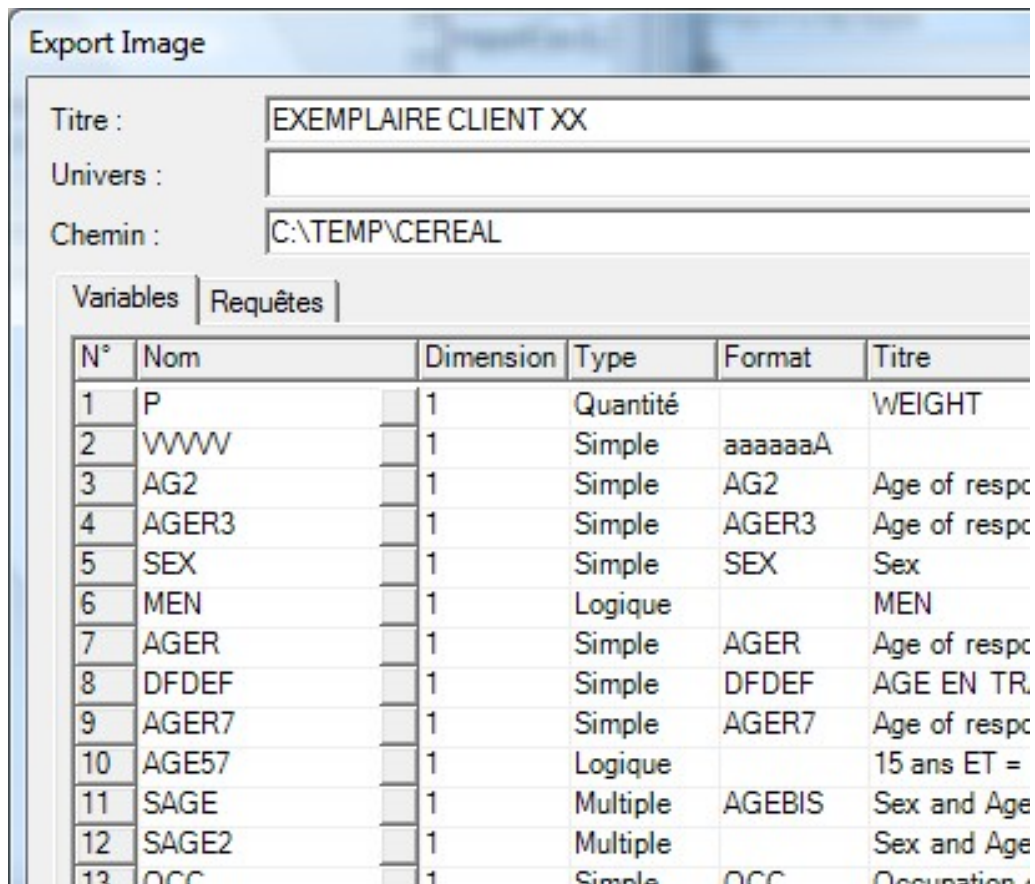
The screenshot shows a window titled "Demandes d'export CoTab Reader". Inside the window, there is a table with the following data:

Date	Titre
13/08/2010	EXEMPLAIRE CLIENT XX

On peut créer des demandes, les modifier, les dupliquer, les supprimer, enfin les exécuter. Une demande correspond à l'écran ci-après :

- **titre** : titre quelconque, pour identifier la demande
- **univers** : peut contenir le nom d'une variable logique, pour filtrer la base exportée sur la valeur *vrai* de cette variable.
- **chemin** : on donne ici l'emplacement du fichier créé (il ne faut pas fournir de nom de fichier)

- **onglets** : l'onglet *Variables* permet de sélectionner les variables que l'on veut exporter dans la base, l'onglet *Requêtes* permet de sélectionner les dossiers de la partie Tableaux que l'on veut exporter.



On peut donc exporter tout ou partie des individus, tout ou partie des variables, tout ou partie des tableaux.

La demande doit être cohérente, sinon un test d'erreur apparaît au moment de l'exécution, et l'exécution n'a pas lieu.

Il faut en particulier que :

- si une variable a une base dans le dictionnaire, la variable qui sert de base doit faire partie de la sélection
- les variables apparaissant dans les demandes de tableaux sélectionnées doivent toutes être présentes dans la sélection.

9.3 Exécution.

L'exécution produit sur le répertoire demandé le logiciel Cotab Reader associé à la base exportée. L'ensemble de ce répertoire peut être recopié sur n'importe quel ordinateur pour être exécuté librement.

10. Rapports.

10.1 Principe.

L'utilisateur ayant exporté ses tableaux sous Excel construit très souvent un rapport à partir de toutes les ressources d'Excel. Tout se passe bien si ces tableaux sont satisfaisants dès le premier essai, et si l'utilisateur n'a pas besoin de refaire l'exploitation sur d'autres données. En effet, une nouvelle exportation des mêmes tableaux portant sur un fichier de données différent écraserait le rapport construit la première fois, et tout serait à recommencer.

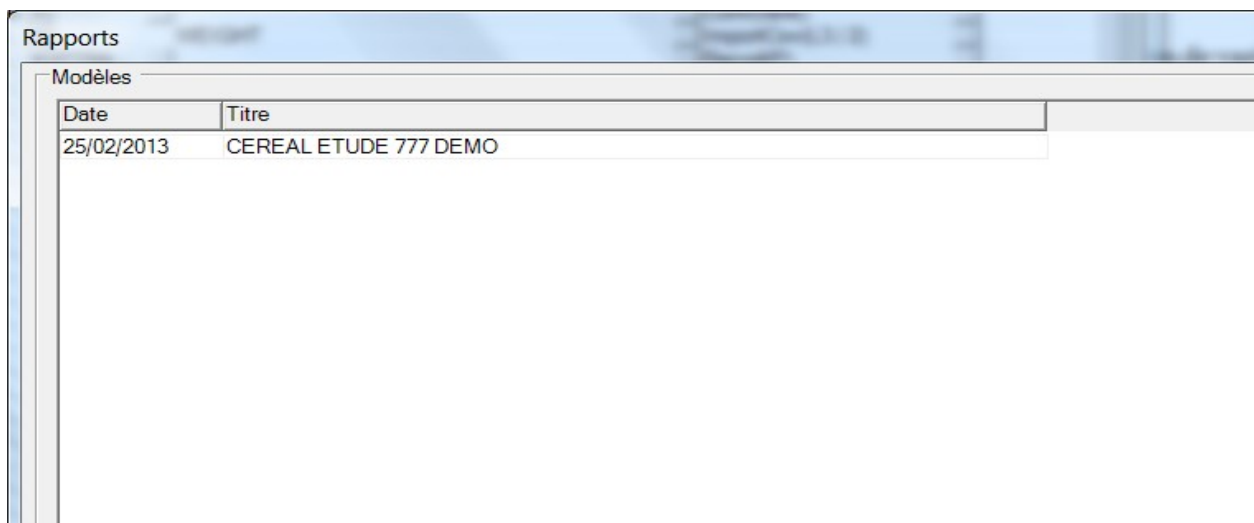
Le module « Rapports » permet de corriger cet inconvénient.

Il fonctionne de la manière suivante :

- L'utilisateur sélectionne les requêtes figurant dans un ou plusieurs dossiers contenant les tableaux qui serviront de base à son rapport.
- Il les exporte dans un fichier Excel « *modèle* », contenant le résultat de tous les tableaux de ces requêtes **dans un seul onglet « Data »**(contrairement aux exportations Excel classiques).
- Il construit son rapport sous Excel, dans d'autres onglets que celui contenant les tableaux fournis.
- S'il a besoin d'éditer à nouveau les mêmes tableaux, par exemple avec un fichier de données différent, une nouvelle exportation dans le même fichier Excel « *modèle* » ne détruira pas son rapport.
- Il pourra donc ainsi alterner les phases « édition de tableaux » et « mise au point du rapport ».
- Quand tout sera correct, il peut générer le rapport final.
- Une option permet de plus de générer ce rapport **pour toutes les valeurs d'une variable de contrôle** (variable dite « population »), il obtiendra donc autant de rapports (de fichiers Excel) que de valeurs pour cette variable.
- Une **macro Excel** est fournie pour copier les tableaux depuis l'onglet « Data » vers les onglets contenant le rapport, en respectant les mises en forme, et en intégrant les liens nécessaires (afin qu'un changement des valeurs des tableaux soit immédiatement transféré dans le rapport).

10.2 Réalisation.

Le module « *Rapports* » est inclus dans le menu « *Utilitaires* ».



The screenshot shows a window titled 'Rapports' with a sub-section 'Modèles'. It contains a table with two columns: 'Date' and 'Titre'. The table has one row with the date '25/02/2013' and the title 'CEREAL ETUDE 777 DEMO'.

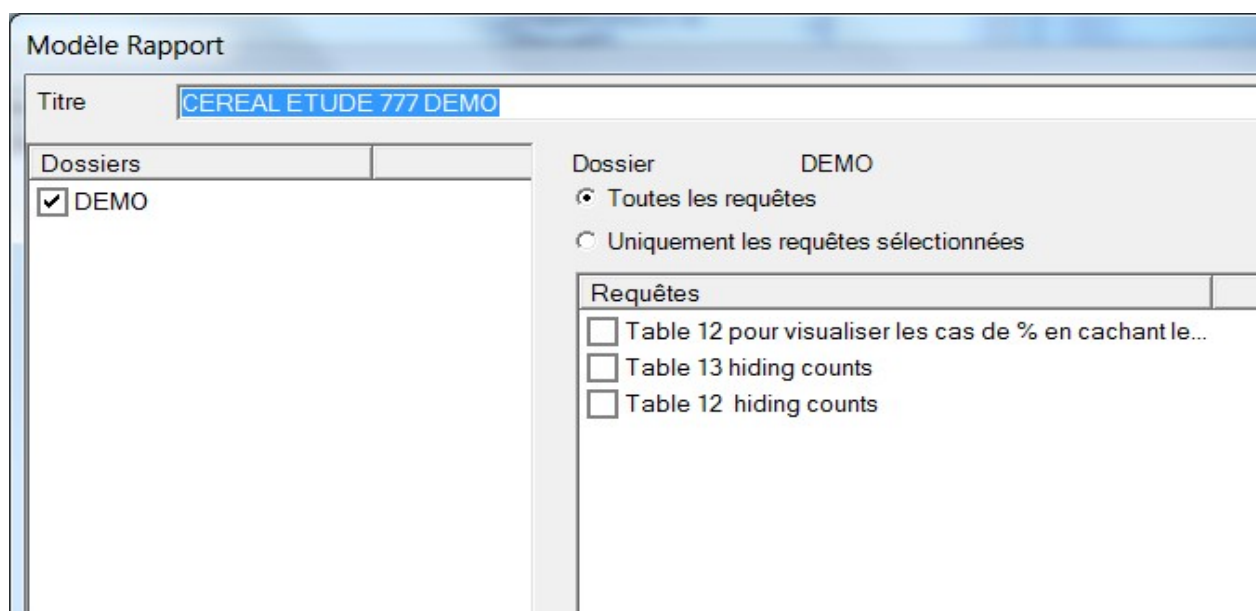
Date	Titre
25/02/2013	CEREAL ETUDE 777 DEMO

Des boutons permettent comme d'habitude de créer un nouveau rapport, de le modifier, de le supprimer, de le dupliquer.

Ce module comprend 3 phases :

- Création (ou modification)
- Construction
- Génération

10.2.1 Création. On sélectionne d'abord les dossiers et demandes de tableaux concernés.



The screenshot shows a window titled 'Modèle Rapport'. It has a 'Titre' field containing 'CEREAL ETUDE 777 DEMO'. Below this is a 'Dossiers' table with a checked checkbox for 'DEMO'. To the right, there are radio buttons for 'Toutes les requêtes' (selected) and 'Uniquement les requêtes sélectionnées'. Below these are three checkboxes for 'Requêtes': 'Table 12 pour visualiser les cas de % en cachant le...', 'Table 13 hiding counts', and 'Table 12 hiding counts'.

Dossiers	Dossier
<input checked="" type="checkbox"/> DEMO	DEMO

Toutes les requêtes
 Uniquement les requêtes sélectionnées

Requêtes
<input type="checkbox"/> Table 12 pour visualiser les cas de % en cachant le...
<input type="checkbox"/> Table 13 hiding counts
<input type="checkbox"/> Table 12 hiding counts

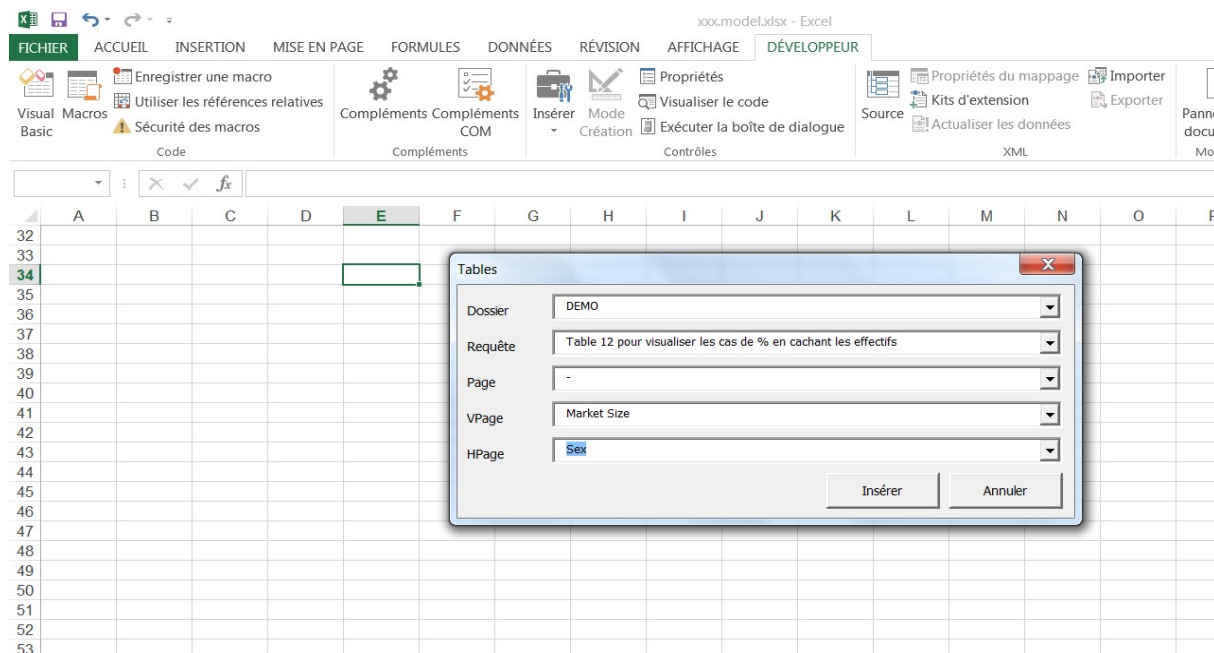
10.2.2 Construction.

Le bouton *Construction* fournit alors un fichier Excel avec un onglet Data contenant les résultats de tous les tableaux demandés, (les uns sous les autres dans un seul

onglet, quelles que soient les requêtes sélectionnées) et un onglet Map. Ce fichier est de la forme *XXX.model.xlsx* (XXX étant le nom donné par l'utilisateur).

L'utilisateur peut alors construire son rapport dans les onglets 3, 4 etc.

La **macro Excel *CotabReports.xla*** qui est fournie avec le logiciel permet de copier aisément des tableaux de l'onglet Data et de les coller dans un autre onglet en respectant les mises en forme et en conservant les liens (*collage spécial, coller avec liaison*).



Remarques :

- **il n'est pas nécessaire d'exécuter** les tableaux avant de lancer cette construction : la construction commande elle-même l'exécution des tableaux sélectionnés pour éditer le fichier Excel.
- Pour que la *Construction* fonctionne, aucune erreur ne doit exister dans l'étude, y compris dans les tableaux non sélectionnés par la demande de *Rapport*. Il faudra donc s'en assurer au préalable au moyen de l'icône *Check*.
- Bien entendu, les noms de dossiers doivent être uniques, de même que les noms des requêtes à l'intérieur d'un dossier.
- L'utilisateur ne doit pas modifier les contenus des onglets *Data* et *Map*.
- Après avoir construit son rapport, si l'utilisateur relance la *Construction* (par exemple s'il a un nouveau fichier de données), le fichier *XXX.model.xlsx* sera réécrit, en plaçant les nouveaux onglets *Data* et *Map* à la suite des onglets créés par l'utilisateur pour son rapport. L'utilisateur doit s'assurer de leur cohérence avec les précédents, afin que les liens présents dans son rapport correspondent aux zones adéquates des tableaux, sous peine de résultats aberrants. D'une construction à l'autre, les tableaux doivent donc être les mêmes (sauf les adjonctions), avec les mêmes structures (pas de changement dans les formats des variables, modifiant le nombre de valeurs), etc.

10.2.2 Génération.

Le bouton *Générer* permettra de créer le rapport définitif dans un autre fichier *XXX.xlsx* (les onglets *Data* et *Map* n'y figureront pas).



L'option *Tout* donnera un seul fichier, correspondant à la totalité du fichier de données. L'option *Répartir* permet, **en sélectionnant une variable** (de type Simple, Multiple, ou Texte) avec un format, d'éditer autant de rapports qu'il y a de valeurs dans le format de la variable, y compris les sous-totaux. Les noms des fichiers correspondants seront complétés par le libellé de la valeur concernée.

Remarque : dans le cas de l'utilisation d'une telle variable *Population*, il est important de pouvoir mettre dans chaque rapport le nom de la population concernée. Dans ce but, l'onglet *Data* comprendra le nom de la variable *Population* dans la case A1, et le libellé de la valeur concernée dans la case B1.

10.3 Remarques.

Un certain nombre de précautions doivent être prises en utilisant le module *Rapports*, dont certaines ont déjà été mentionnées plus haut.

- Aucune erreur ne doit exister dans l'étude, même dans les tableaux non sélectionnés par la demande.
- En construction comme en génération, les fichiers Excel concernés déjà existants doivent être fermés.
- La version d'Excel utilisée ne doit pas être antérieure à 2007.
- Ne pas modifier les onglets *Data* et *Map*.
- D'une construction à l'autre pour la même demande de rapports, ainsi que de la construction à la génération, **les tableaux ne doivent pas être modifiés** : pas de nouveau tableau (sauf à la fin), pas de changement dans le nombre de colonnes ou de lignes d'un tableau. En résumé il faut que la structure des onglets *Data* et *Map* restent inchangée, sinon les liens permettant de copier des cases entre les onglets *Data* et rapports ne seraient plus adéquats. C'est donc à l'utilisateur de vérifier cette cohérence.
- Pour la même raison, il **est fortement déconseillé d'utiliser dans les tableaux les options « supprimer les colonnes vides » et « supprimer les lignes vides »**, car d'un fichier de données à l'autre, on pourrait avoir un nombre de lignes ou de colonnes fluctuant.

Annexe : Procédures

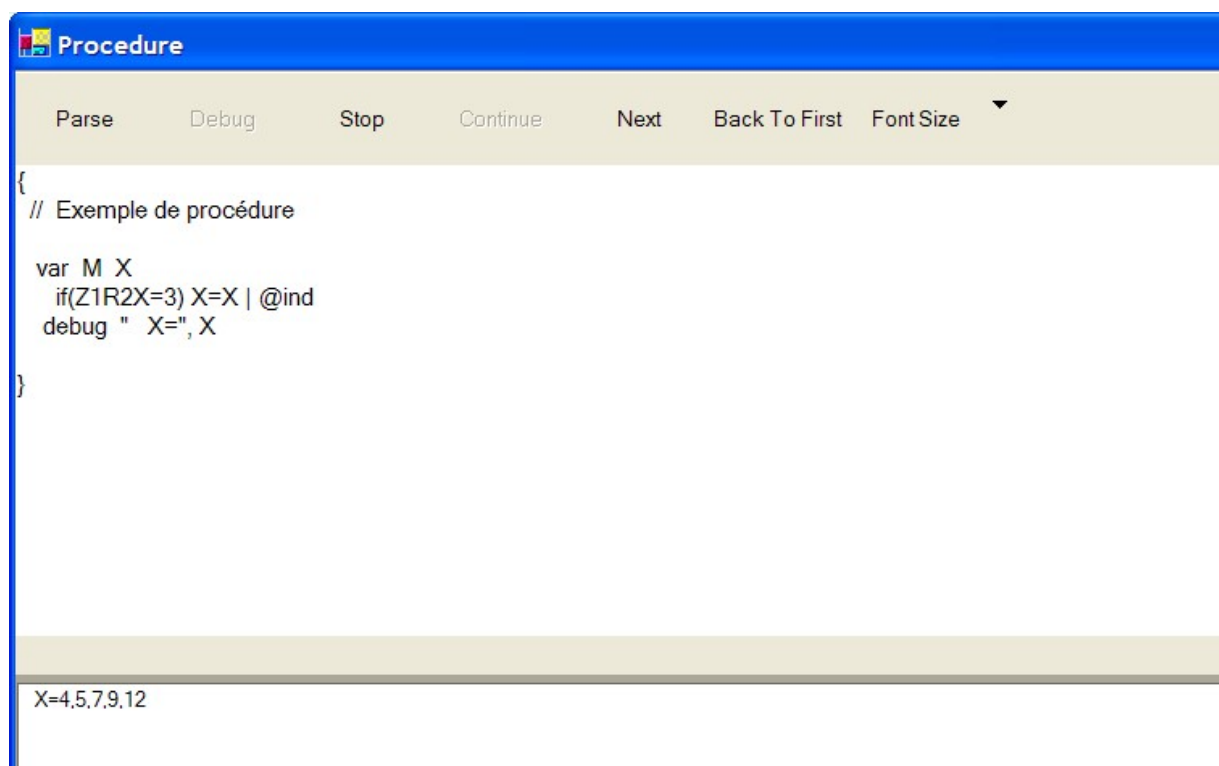
1) Généralités.

Une procédure est une fonction écrite dans le langage CPL (Cotab Procedural Language) : ce langage permet d'exécuter toute sorte d'opérations arithmétiques, logiques, des tests, des boucles, etc., afin de donner une valeur à la variable qui la porte, ou à toute autre variable située avant l'apparition de cette procédure.

On peut définir plusieurs procédures dans une étude. **La variable portant la procédure ne doit pas être dimensionnée.**

Pour créer une procédure, on choisit, dans la zone « *fonction* », la fonction CPL fonction.

On peut importer une procédure d'une autre étude, en important la variable qui la porte.



```
{
// Exemple de procédure

var M X
  if(Z1R2X=3) X=X | @ind
  debug " X=", X
}
```

X=4.5.7.9.12

2) Création d'une procédure.

L'écran de création d'une procédure est divisé en 2 parties : dans la partie supérieure, on écrit la procédure proprement dite.

On dispose pour cela d'un **éditeur** permettant de rédiger une suite d'instructions en format libre, et utilisant les touches classiques : Ctrl C, Ctrl V, Ctrl Z, Ctrl A, Inser, Suppr, Entrée.

En appuyant sur le bouton « *Parse* », on peut tester la syntaxe. En appuyant sur le bouton « *Debug* », on passe en mode test : en appuyant alors sur le bouton « *Suivant* », on voit apparaître dans la partie inférieure de l'écran le résultat sur le premier individu traité des instructions « *debug* » figurant dans la procédure. En appuyant à nouveau sur *Suivant*, on passe à l'individu suivant.

Retour au début permet de revenir au début du fichier.

Stop permet de revenir au mode Editeur, pour modifier la procédure.

```
Procédure
Parse Debug Stop Continue Next Back To First Font Size
{
var M x,y
x=(REC4 | REC2B) & depars[3]
y=(REC4 & REC2B) | depars[3]
debug " REC4=", REC4, " REC2B=", REC2B, " depars[3]=", depars[3], " x=", x, " y=", y
YY=y
}
REC4=3,6,7 REC2B=1 depars[3]=4 x=@na y=4
```

3) Le langage CPL.

3.1 Symboles et notations.

3.1.1 Définitions

CPL est un langage doté d'une syntaxe assez simple.

Les symboles que l'on peut utiliser en CPL peuvent être classés en :

- symboles systèmes
- mots clés
- variables de l'étude
- variables locales
- constantes

Remarque :

Les noms de variables sont sensibles à la casse (c'est-à-dire qu'il faut respecter les majuscules et minuscules), en revanche les symboles systèmes et les mots clés ne le sont pas.

3.1.2 Les symboles systèmes

Ces symboles sont identifiés par leur premier caractère qui doit être '@', suivi d'un ou plusieurs caractères, par exemple : @r, @ind, @ecrit, etc.,

3.1.3. Les mots clés

Les mots clés réservés sont sur un ou plusieurs caractères : *si, pour, continue,*.

3.1.4 Les variables de l'étude

Il s'agit de variables définies dans le dictionnaire **avant la procédure.**

Attention :

- a) ne pas utiliser de **variables accentuées** dans les procédures.
- b) il n'y a **pas de vérification automatique** que les variables appelées sont bien situées avant la procédure. Une variable située après la procédure vaudra toujours **Rebut.**

3.1.5 Les variables locales

Elles doivent être déclarées avant leur première utilisation (par une instruction du type : **var S a,b**), une variable locale peut être déclarée n'importe où dans la procédure. Le premier caractère du nom doit être une lettre, les autres caractères peuvent être alphanumériques ou '_', ce nom ne doit pas être un mot clé réservé de CPL. Il est à noter qu'il y a plusieurs différences entre les variables locales et les variables de l'étude :

- les variables locales sont initialisées, **au début de l'exécution**, avec la valeur 0 pour les types S ou Q, REBUT pour le type M, "" pour le type T, *faux* pour les types L, (alors que les variables de l'étude sont toutes initialisées à REBUT pour chaque nouvel individu statistique),
- la valeur d'une variable locale n'est pas affectée par le moteur principal de l'exécution de l'étude.

3.1.6 Les constantes

Les constantes numériques peuvent être de type S ou Q, CPL reconnaît automatiquement le type de ces constantes, par exemple : 123 est de type S, alors que 123.0 ou 123. sont de type Q. Les constantes littérales doivent être entourées de "", si un " fait partie d'une constante littérale, on le fait précéder du caractère ^, par exemple, "alors, ^"ok^"

Les constantes logiques sont 0 et 1, pour *faux* et *vrai*.

3.1.7 Liste des Symboles systèmes

. **@r** : constante REBUT

@ecrit : indique le fichier (nom et répertoire) dans lequel écrire les messages obtenus par la fonction *write*.

Par exemple :

```
@ecrit nom_fichier, « c:\xxx\test.log »
```

nom_fichier est le nom, interne à la procédure, auquel les instructions d'écriture feront référence (ce nom peut être un simple numéro).

Si ce n'est pas un numéro, il doit être défini comme une variable de type simple.

On peut définir un ou plusieurs fichiers dans le(s)quel(s) la procédure pourra écrire.

@ind : variable système de type S, retourne l'indice de l'occurrence dans la boucle implicite d'une instruction dimensionnée, par exemple, pour initialiser une variable dimensionnée STAB, les deux instructions suivantes sont équivalentes :

```
STAB = @ind
```

```
pour (i = 0 ; i <= dim ; i++) {STAB[i] = i}
```

3.1.8 Commentaires : on peut utiliser deux marques de commentaires :

// permet d'ignorer toute la partie droite d'une ligne,

/* */ permet d'ignorer plusieurs lignes d'une procédure.

3.2 Types de variables et conversion implicite.

3.2.1 Les types de variables

Les variables de l'étude ou locales peuvent être de type :

M : Multiréponses

S : Simple réponse (numérique sans décimales)

Q : Quantité (numérique avec décimales)

T : Texte

L : Logiques

La déclaration de variables locales se fait comme indiqué ci-après; il faut toujours préciser le type des variables à déclarer :

```
var M m1, m2, m3[200], m4[10]
```

Lorsqu'une variable ou constante est utilisée dans une instruction, CPL effectue la conversion implicite automatiquement si c'est nécessaire. La conversion ne détruit pas le type, ni la valeur d'origine de la variable. Il est à noter qu'une instruction nécessitant une conversion automatique est plus lente en exécution qu'une instruction sans conversion.

3.2.2 Règles de conversion entre les types :

M → T : 2,3,5 → "2,3,5"

Q → S : la valeur est arrondie, par exemple, 2.34 ⇒ 2 et 3.5 ⇒ 4.

S → M : si la valeur est ≥ 0 on affecte la valeur dans M, sinon M est mis à REBUT,

Q → M se traduit par Q → S et S → M

M → S ou Q : si la variable M est REBUT, le résultat est REBUT, sinon c'est **la valeur minimum** qui est prise.

S → T : 2005 → "2005 "

Q → T : 2.34 → "2.34"

et REBUT → "" ou " "

Grâce à la conversion automatique de types, on peut écrire les instructions ci-dessous en CPL :

```
var M m
var T text
var S i
var Q q
m = "2,4,10" // T => M
q = "2.34" // T => Q
text = m // M => T
i = q // Q => S
i = m // M => S
```

3.3 Structure d'une procédure CPL.

Une procédure est une routine utilisateur que le moteur principal appelle lors de l'exécution de l'étude.

Elle peut comprendre 2 sections : le début et le corps de la procédure.

Voici un exemple de procédure très simple qui comporte les deux sections, avec, sous forme de commentaire, les explications sur le fonctionnement de chacune d'elles :

```
var Q cumul_POIDS // init à 0.0
@debut
{
  var S Total // init à 0
  // Les instructions dans cette section sont exécutées une seule fois : au début de
  // l'exécution de l'étude, avant toute lecture du fichier de données.
  // Appelée par le moteur principal dans l'ordre du dictionnaire.
  Total = 0
}

{
  // Les instructions dans cette section sont exécutées pour chaque individu statistique
  // du niveau de la procédure.
  // Appelée par le moteur principal dans l'ordre du dictionnaire.
  Total++
  cumul_POIDS += POIDS
}
```

Les 2 sections sont toutes facultatives, chacune doit comporter, si elle est présente, au moins une instruction autre que les déclarations de variables.

L'instruction exit n'est pas permise dans la section @debut

Les variables de l'étude ne sont pas accessibles dans la section @debut

Les variables locales (propres à la procédure) peuvent être déclarées n'importe où, à l'intérieur ou en dehors d'une section.

3.4 Instructions de base du CPL.

CPL (Cotab procedural language) est un langage structuré à format libre.

Les instructions sont généralement terminées par une ou plusieurs fins de ligne, mais certaines expressions peuvent s'écrire plusieurs lignes; en particulier, certains caractères comme ')', '=', etc. peuvent être suivis d'une fin de ligne au milieu d'une instruction, car CPL sait que l'instruction n'est pas encore terminée dans ce cas.

3.4.1. Dimension des instructions (CPL)

La notion de dimension des instructions est un aspect très important du langage CPL, qui permet d'effectuer une instruction ou un bloc d'instructions (avec une structure de contrôles, cf. ci-après), occurrence par occurrence sur des variables dimensionnées (forcément de même dimension ou non dimensionnées).

Il s'agit d'une généralisation naturelle de l'exécution des fonctions Cotab lorsqu'elles s'appliquent à des variables dimensionnées.

Règles générales :

- la dimension d'une instruction de base est déterminée par celle des variables intervenant dans l'instruction
- elles doivent être toutes de même dimension ou non dimensionnées.

Le moteur de CPL exécute une instruction dimensionnée avec une boucle implicite de 1 jusqu'à la dimension de l'instruction.

La variable système *@ind* peut être utilisée dans une instruction dimensionnée, elle prend comme valeur l'indice de l'occurrence dans la boucle implicite (*@ind* est toujours 0 dans une instruction non dimensionnée).

Par exemple, supposons que l'on a une variable *CNBHAB* dimensionnée, *cumul* étant une variable locale non dimensionnée, alors les 2 instructions suivantes sont équivalentes,

cumul += CNBHAB // l'instruction est de dimension = CNBHAB.dim

```
pour (i=1; i<=CNBHAB.dim; i++)
{
  cumul += CNBHAB[i] // instruction non dimensionnée
}
```

Remarque : si la variable portant la procédure est dimensionnée, toute la procédure sera automatiquement exécutée n fois (n étant la dimension de la variable). Cela risque d'être difficile à concevoir pour l'utilisateur, c'est pourquoi il est conseillé de ne pas mettre de procédures sur les variables dimensionnées.

3.4.2. Accès à une occurrence d'une variable dimensionnée

Pour une variable dimensionnée nommée *var_tab*, l'expression *var_tab[expr_arithm]* permet d'obtenir l'occurrence de *var_tab* où l'évaluation de *expr_num* retourne l'indice de l'occurrence voulue.

Exemple : *var_tab[3]*, *var_tab[i*3+j]*, *var_tab[(i+j)*3+k]*, etc.

3.4.3. Comment incrémenter et décrémenter

On a souvent besoin d'incrémenter ou de décrémenter la valeur d'une variable numérique (de type S ou de type Q), CPL fournit pour cela deux instructions :

```
variable++ // variable est incrémentée de 1
variable-- // variable est décrémentée de 1
variable doit être de type S ou Q.
```

3.4.4. Les instructions d'affectation

```
variable = expr
expr peut être ( expr )
    ou      expr_num
    ou      expr_log
    ou      expr_litterale

expr_num peut être      expr_num + expr_num // addition
    ou      expr_num - expr_num // soustraction
    ou      expr_num * expr_num // multiplication
    ou      expr_num / expr_num // division
    ou      expr_num % expr_num // modulo
    ou      expr_num ** expr_num // puissance

    ou      foncmath(expr_num)
    ou      Min(arg1[, arg2, ...]) // arg doivent être non dimens.
    ou      Max(arg1[, arg2, ...]) // arg doivent être non dim.
    ou      NbRep(variable_M)
    ou      Long(variable_T)
    ou      arg_num (variable ou constante de type S ou Q)
```

où foncmath peut être une des fonctions mathématiques suivantes :

ABS : valeur absolue
 SQRT : racine carrée
 ROUND : arrondi
 TRUNC : troncature
 RANDOM : nombre au hasard, compris entre 0 et 1 (pas d'argument pour cette fonction)
 EXP : exponentielle
 LOG : logarithme népérien
 ISP : retourne 1 si l'argument est positif ou nul, 0 sinon
 ISZ : retourne 1 si l'argument est égal à zéro, 0 sinon
 ISSP : retourne 1 si l'argument est strictement positif, 0 sinon

```
expr_log peut être      expr_log & expr_log // et
    ou      expr_log | expr_log // ou
    ou      expr_log ~ expr_log // sauf
    ou      expr_log ^ expr_log // ex
    ou      arg_log (variable ou constante de type M)
```

`expr_litterale` peut être `arg_lit` (variable ou constante de type T)

Les priorités entre opérateurs arithmétiques sont :

** puis *, /, % puis +, -.

Il n'y pas de priorités entre les opérateurs logiques.

Remarque : les calculs sont toujours faits en nombres réels, même si les variables résultat ou arguments sont simples, et le résultat est arrondi si la variable résultat est simple.

3.4.6. Exemple

The screenshot shows a window titled "Procédure" with a menu bar containing "Parse", "Debug", "Stop", "Continue", "Next", "Back To First", and "Font Size". The code in the editor is as follows:

```
{ // exemple d'expressions arithmétiques ou logiques
var Q x,y
var S a
  a=rang % 10
x= (WQAD-a)/ (WQAD+a)
debug x

var M u
u=(QDI ^ 1) | (2 & VENT1)
}|
```

At the bottom of the window, the result of the calculation is displayed as 0.285714285714286.

3.4.6. Autres formes d'affectations (avec raccourci d'écriture)

<code>v_SQ += expr_num</code>	<code>// v_SQ = v_SQ + expr_num</code>
<code>v_SQ -= expr_num</code>	<code>// v_SQ = v_SQ - expr_num</code>
<code>v_SQ *= expr_num</code>	<code>// v_SQ = v_SQ * expr_num</code>
<code>v_SQ /= expr_num</code>	<code>// v_SQ = v_SQ / expr_num</code>
<code>v_SQ %= expr_num</code>	<code>// v_SQ = v_SQ % expr_num</code>
<code>v_SQ **= expr_num</code>	<code>// v_SQ = v_SQ ** expr_num</code>
<code>v_M &= expr_log</code>	<code>// v_M = v_M & expr_log</code>
<code>v_M = expr_log</code>	<code>// v_M = v_M expr_log</code>
<code>v_M ~= expr_log</code>	<code>// v_M = v_M ~ expr_log</code>
<code>v_M ^= expr_log</code>	<code>// v_M = v_M ^ expr_log</code>

3.4.7. Autres fonctions:

Concaténation de chaînes de caractères :

variable = **Chain** (*CHN1*, *CHN2*,...)

CHN1 est soit une variable texte, soit une constante texte (« »)

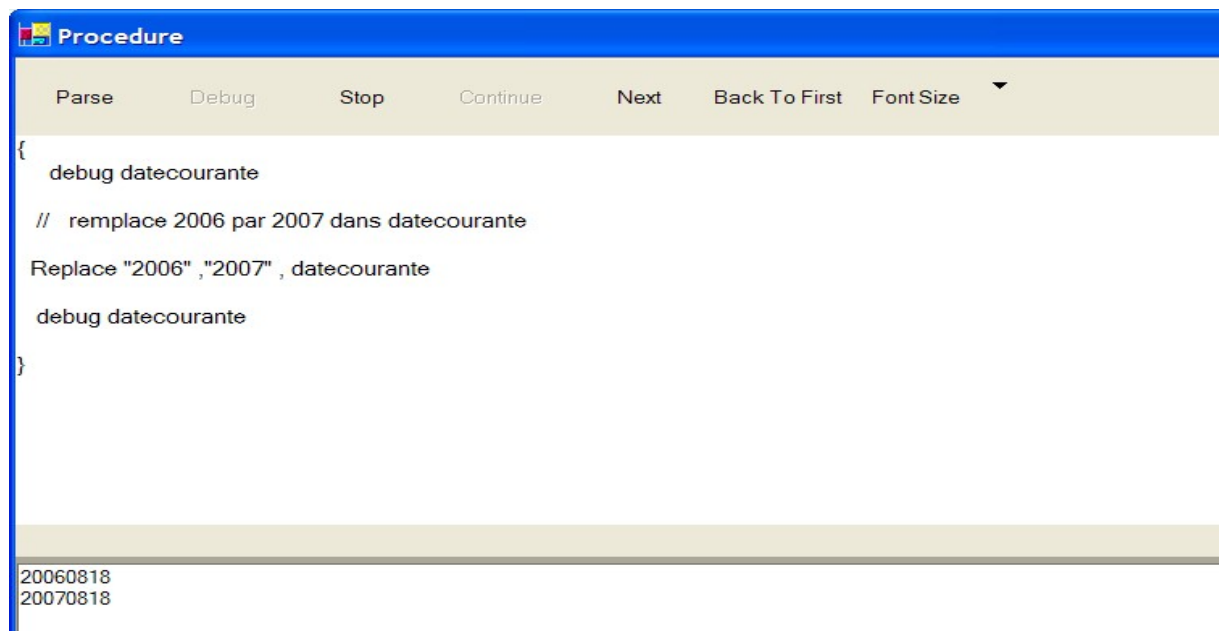
Remplacement de la 1^{ère} apparition d'une chaîne (1^{er} argument) par une autre (2^{ème} argument) dans une variable texte (3^{ème} argument).

Replace *CHN_RECH* , *CHN_REMPLACE* , *CHN_RES*

Remplacement de toutes les apparitions d'une chaîne :

ReplaceAll *CHN_RECH* , *CHN_REMPLACE* , *CHN_RES*

Exemple :



The screenshot shows a debugger window titled "Procédure" with a menu bar containing "Parse", "Debug", "Stop", "Continue", "Next", "Back To First", and "Font Size". The code in the editor is as follows:

```
{
  debug datecourante
  // remplace 2006 par 2007 dans datecourante
  Replace "2006" ,"2007" , datecourante
  debug datecourante
}
```

At the bottom of the window, the execution output is displayed:

```
20060818
20070818
```

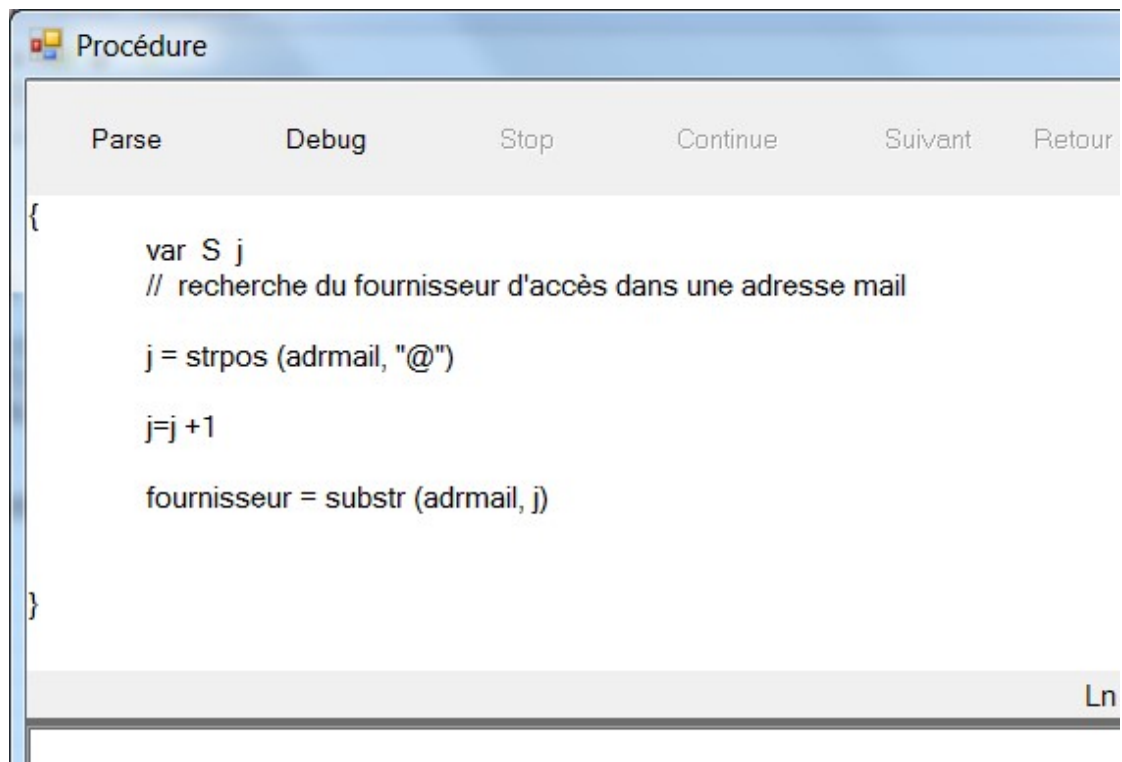
Strlen (varT) : fournit la **longueur** de la chaîne de caractères contenue dans la variable texte varT

Strpos (varT, « xyz ») : fournit la **position** de la première apparition de la sous-chaîne xyz dans la variable Texte varT (valeur -1 si la sous-chaîne xyz n'existe pas).

Substr (varT, j, k) : fournit la **sous-chaîne** de k caractères de la chaîne contenue dans varT, à partir de la position j

ou **Substr** (varT, j) : fournit la **sous-chaîne** de la chaîne contenue dans varT, à partir de la position j

Exemples :



```
Procédure  
Parse Debug Stop Continue Suivant Retour  
{  
    var S j  
    // recherche du fournisseur d'accès dans une adresse mail  
    j = strpos (adrmal, "@")  
    j=j +1  
    fournisseur = substr (adrmal, j)  
}
```

Ln

3.5 Expressions de conditions et structures de contrôle (CPL).

3.5.1. Expressions de condition

Les structures de contrôles utilisent des expressions de conditions pour contrôler le flux d'exécution des instructions, l'évaluation d'une condition retourne la valeur VRAI ou FAUX.

Syntaxe : **toute condition doit être mise entre parenthèses.**

condition peut être	(condition condition)	(ou)
ou	(condition & condition)	(et)
ou	(condition ~ condition)	(sauf)
ou	(condition ^ condition)	(ex)
ou	(! Condition)	(non)
ou	(arg1 > arg2)	
ou	(arg1 >= arg2)	
ou	(arg1 < arg2)	
ou	(arg1 <= arg2)	
ou	(arg1 = arg2)	
ou	(arg1 <> arg2)	
ou	(arg1)	// signifie : arg1 <> @rebut

où : arg1, arg2, etc. sont des variables ou constantes

Dans le cas « = » ou « <> », on peut mettre plusieurs valeurs séparées par des virgules :
 (arg1 = 1, 2, 3) signifie ((arg1 = 1) | (arg1 = 2) | (arg1 = 3))
 (arg1 <> 1, 2, 3) signifie ((arg1 <> 1) & (arg1 <> 2) & (arg1 <> 3))

Cas particulier pour les variables multiples :

(VM1 = VM2) VRAI si l'intersection de VM1 et de VM2 n'est pas vide,
 OU si VM1 et VM2 sont tous les deux en rebut (@r),

(VM1 <> VM2) : VRAI si l'intersection de VM1 et de VM2 est vide
 ET VM1 et VM2 ne sont pas tous les deux en rebut (@r),

(VM1 > VM2) VRAI si VM1 <> @r ET sa plus grande valeur > celle de VM2,

(VM1 < VM2) VRAI si VM2 <> @r ET sa plus grande valeur < celle de VM1.

Il est à noter que pour les variables de type Q, les tests = et <> peuvent être faussés pour des raisons de précision dans la représentation machine des nombres flottants, car certains nombres flottants ne peuvent pas être codés exactement par la machine (Cotab utilise IEEE double précision pour représenter en interne les valeurs des variables et constantes de type Q). Il est donc conseillé d'utiliser un test du genre ((VQ > -0.00001) & (VQ < 0.00001)) au lieu de (VQ = 0.0).

La **dimension** d'une condition simple est déterminée par celle des variables dimensionnées utilisées par la condition (qui doivent être de même dimension ou non dimensionnées). La dimension d'une combinaison de conditions est déterminée par celle des conditions simples impliquées dans la combinaison (qui doivent être de même dimension ou non dimensionnées).

3.5.2. Structure de contrôle

Ce sont des instructions conditionnelles simples (*si ... sinon...*) ou des boucles (*pour...*). Etant des instructions, elles peuvent être utilisées à l'intérieur d'autres instructions conditionnelles ou boucles. Il est à noter qu'il faut utiliser les accolades '{' et '}' pour regrouper une liste de 2 ou plusieurs instructions pour former *instruction* dans les syntaxes qui suivent.

La dimension d'une instruction de contrôle est déterminée par la dimension de l'expression de la condition utilisée.

3.5.2.1 Instructions conditionnelles simples

si (condition) instruction

CPL évalue d'abord condition : si la condition est VRAIE, CPL exécute *instruction*.

Exemple :

si ((v = 1) & (VQ <= poids1) | (VQ >= poids2)) myvar = VS

si (condition) instruction1

sinon instruction2

CPL évalue d'abord condition : si elle est VRAIE, CPL exécute instruction1, si elle est FAUSSE, CPL exécute instruction2.

The screenshot shows a debugger window titled "Procédure" with a menu bar containing "Parse", "Debug", "Stop", "Continue", "Next", "Back To First", and "Font Size". The code being debugged is as follows:

```
{
var S j,k, z[4]
  z=0

if(depars>0){z=depars+1
              debug " depars=",depars," z=",z
              }
else z=999
debug depars, " z=",z
}
```

The output window at the bottom shows the following execution results:

```
depars=55 z=56
depars=9 z=10
55 z=56
...
```

3.5.2.2 Boucles

pour (*[instruction_start]*; *[condition]*; *[instruction_boucle]*)
instruction

CPL exécute d'abord *instruction_start* qui est facultative.

Boucle : CPL évalue d'abord *condition*, si elle est VRAIE, CPL exécute *instruction* puis *instruction_boucle* puis recommence la boucle, sinon CPL sort de la boucle,

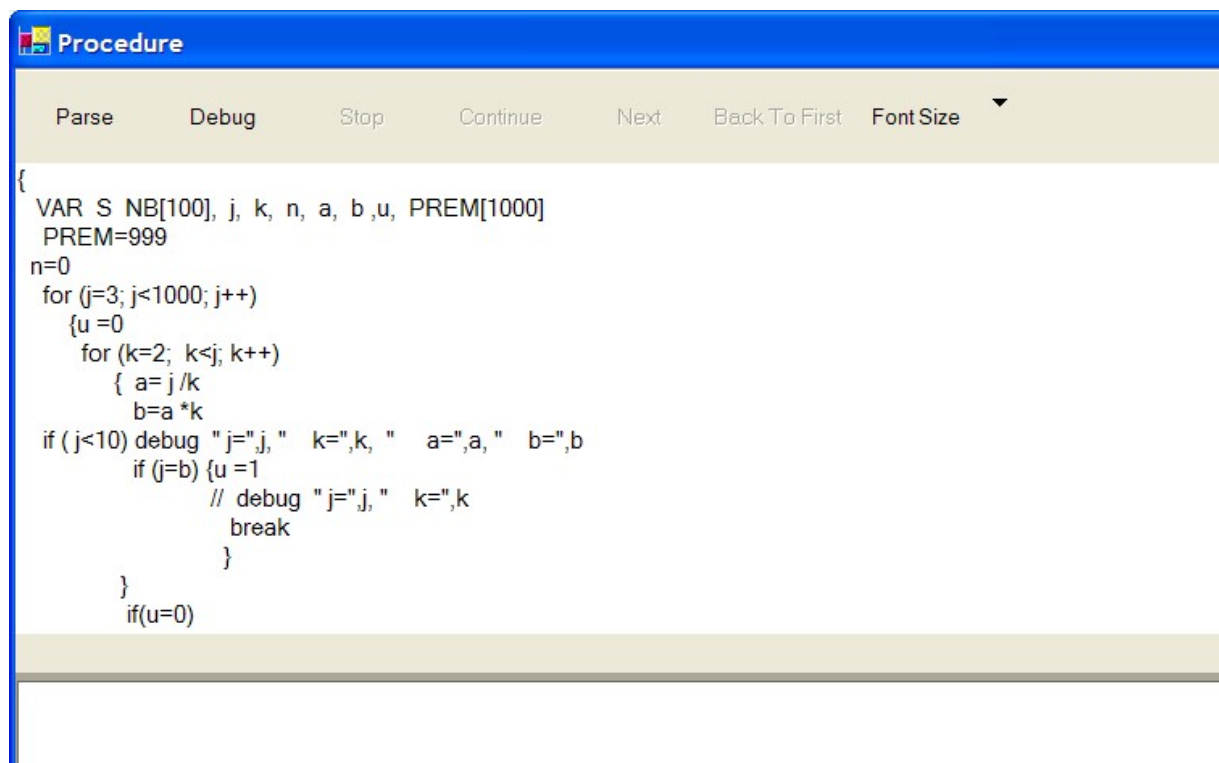
Exemple :

```
pour (i = 1; i <= 10; i++)
{
  vs_tab[i] = i
}
```

ATTENTION !!!

CPL ne contrôle pas la présence d'une éventuelle boucle sans fin qui fera planter l'exécution de l'étude et donc Cotab

Exemple :



```
{
VAR S NB[100], j, k, n, a, b,u, PREM[1000]
PREM=999
n=0
for (j=3; j<1000; j++)
  {u=0
  for (k=2; k<j; k++)
    { a=j/k
    b=a*k
if (j<10) debug "j=",j, " k=",k, " a=",a, " b=",b
  if (j=b) {u=1
    // debug "j=",j, " k=",k
    break
  }
}
if(u=0)
```

3.5.2.3 Les instructions "break" et "continue"

On a parfois besoin de quitter une boucle immédiatement lorsqu'une certaine condition est vérifiée (on appelle cela une exception), l'instruction break permet de quitter immédiatement une boucle.

Exemple :

```
pour (i = 1; i <= 10; i++)  
{  
  // Garde-fou : lorsque i est plus grand que la  
  // dimension de la variable vs_tab, quitter la boucle  
  si (i > vs_tab.dim) break  
  vs_tab[i] = i  
}
```

Un autre besoin concernant les boucles : retourner immédiatement au début de la boucle tout en ignorant les instructions restantes dans la boucle, l'instruction continue permet de réaliser cette action.

Exemple :

```
pour (i = 1; i <= 10; i++)  
{  
  mytab[i] = i * 2  
  si (i < 3) continue // lorsque i < 3 ignorer la suite  
  // de la boucle  
  vs_tab[i] = i  
}
```

3.6 Instructions spéciales (CPL) .

3.6.1 exit et stop

Pour faciliter le débogage d'une procédure et permettre à l'utilisateur de générer un compte rendu d'exécution, CPL fournit les instructions spéciales ci-dessous :

exit : permet de quitter immédiatement une des trois sections d'une procédure.

stop : permet d'arrêter l'exécution de l'étude

Etant donné qu'une procédure est une routine utilisateur que le moteur principal appelle, l'instruction *stop* ne peut pas fonctionner comme un arrêt brutal de l'exécution du moteur principal, au lieu de cela, le moteur de CPL met la valeur VRAI au flag *bUserArret* du moteur principal (cela simule en fait un clic sur le bouton "Arrêter" ou plutôt la fin de fichier), ce flag étant testé par le moteur principal au début de chaque lecture dans le fichier en entrée, le moteur s'arrête tout seul avant la prochaine lecture dans le fichier en entrée.

Par exemple, dans la mise au point d'une procédure, on souhaite arrêter l'exécution lorsque 3 enregistrements sont traités (pour ne pas avoir trop de messages de débogage), pour ceci, il suffit d'écrire :

```
{ var S eff
  eff++
  si( eff = 3) stop
}
```

3.6.2 reset

Cette instruction permet de supprimer un fichier ; elle s'emploie généralement dans la section *begin* d'une procédure.

```
@reset «c:\xxxxx »
```

(xxxx étant le nom du fichier)

3.6.3 debug et écrit

```
debug arg1, arg2, ...
```

```
écrit nomfichier , arg1, arg2, ...
```

Ces deux instructions, ayant la même syntaxe, permettent d'afficher à l'écran ou d'écrire dans un fichier des lignes de messages. La dimension de ces 2 instructions sont déterminées par la dimension de *arg1*, *arg2*, ... (forcément de même dimension ou non dimensionnés).

Debug affiche le message dans le bas de l'écran de procédure.

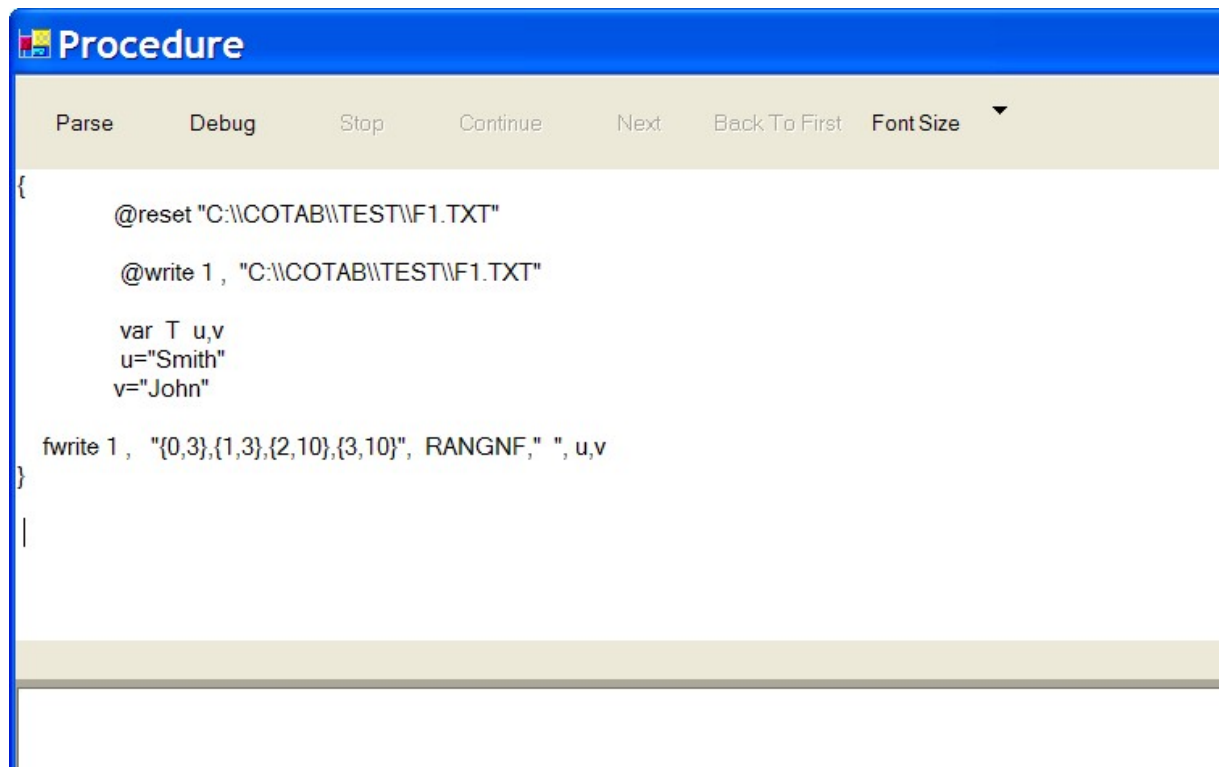
écrit va écrire les arguments spécifiés dans le fichier *nomfichier* (défini dans une instruction @écrit)

3.6.4 fecrit

Cette instruction permet d'écrire dans un fichier avec un format

fecrit « *nomfichier* », « {0,n1},{1,n2},{2,n3},... », *arg1*, *arg2*, ...

Par rapport à l'instruction *ecrit*, on ajoute entre accolades le nombre de positions ni de chaque argument *i* dans le fichier de sortie.



```
Parse   Debug   Stop   Continue   Next   Back To First   Font Size ▼
{
    @reset "C:\\COTAB\\TEST\\F1.TXT"

    @write 1 , "C:\\COTAB\\TEST\\F1.TXT"

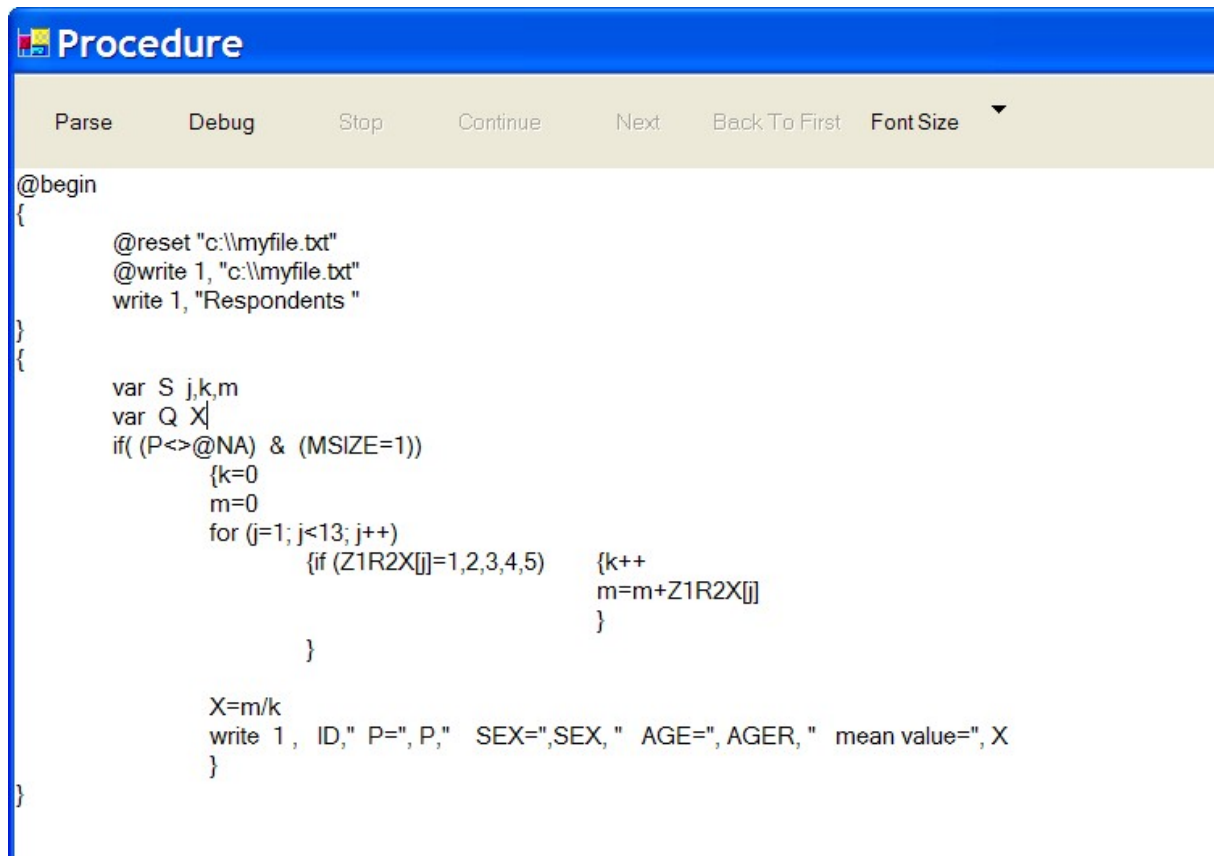
    var T u,v
    u="Smith"
    v="John"

    fwrite 1 , "{0,3},{1,3},{2,10},{3,10}", RANGNF," ", u,v
}
|
```

3.6.5 Remarque importante pour *ecrit* et *fecrit*

nomfichier défini dans l'instruction *@ecrit* n'est pas un nom interne à la procédure où il est défini : **il est utilisable dans toutes les autres procédures**. On peut donc écrire dans le même fichier à partir de plusieurs procédures.

3.7 Exemple de procédure



```
@begin
{
  @reset "c:\\myfile.txt"
  @write 1, "c:\\myfile.txt"
  write 1, "Respondents "
}
{
  var S j,k,m
  var Q X
  if( (P<>@NA) & (MSIZE=1))
    {k=0
     m=0
     for (j=1; j<13; j++)
       {if (Z1R2X[j]=1,2,3,4,5) {k++
                               m=m+Z1R2X[j]
                               }
       }

     X=m/k
     write 1, ID," P=", P," SEX=",SEX," AGE=", AGER," mean value=", X
  }
}
```

Annexe : Liste des mots-clé.

Français	Anglais
@debut	@begin
@r	@na
@ecrit	@write
ecrit	write
vrai	true
faux	false
@ind	@ind
@nomproc	@procname
debug	debug
var	var
chain	chain
nbval	nbval
somval	sumval
min	min
max	max
nbrep	nbrep
long	length
sqrt	sqrt
abs	abs
round	round
trunc	trunc
random	random
exp	exp
log	log
isp	isp
isz	isz
issp	issp
replace	replace
replaceall	replaceall
si	if
sinon	else
pour	for
break	break
continue	continue
exit	exit
stop	stop
debug	debug
@reset	@reset
fecrit	fwrite