

ADE-4



CategVar

CategVar : Categ->Bloc-----	1
CategVar : Categ -> Burt-----	3
CategVar : Categ -> Disj-----	7
CategVar : CrossingCateg -----	8
CategVar : Quant -> Categ -----	10
CategVar : Quant -> Sign-----	12
CategVar : Read Categ File-----	14
CategVar : Read Label File -----	17
CategVar : Read Taxa File -----	19
CategVar : ReadTestBurt -----	23
CategVar : Recoding -----	24
CategVar : VerifCateg -----	28

CategVar : Categ->Bloc



Utilitaire de manipulation de fichiers.



L'option assure l'inverse de la fonction de TextToBin : BlocTo1Categ. Un fichier de variables qualitatives contient une seule variable qui prend successivement la valeur 1 (n_1 fois), la valeur 2 (n_2 fois), ..., la valeur m (n_m fois). Il est transformé en un fichier à une seule colonne contenant l'indicatrice des blocs (voir fichiers indicateurs de blocs dans KtabUtil).



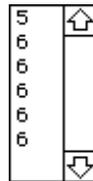
L'option utilise une seule fenêtre de dialogue :



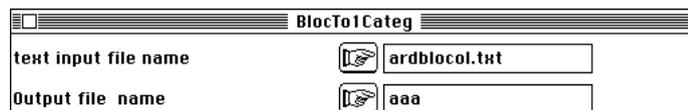
 Nom du fichier d'entrée de type .cat.



Sur la carte Ardèche de la pile ADE-4•Data, obtenir le fichier ardbloco1.txt :



Le transformer en une variable qualitative :



```
Qualitative variables file: aaa_c
Number of rows: 35, variables: 1, categories: 6
```

Description of categories:

Variable number 1 has 6 categories

```
[ 1]Category:      1 Num:      5 Freq.:      0.1429
[ 2]Category:      2 Num:      6 Freq.:      0.1714
[ 3]Category:      3 Num:      6 Freq.:      0.1714
[ 4]Category:      4 Num:      6 Freq.:      0.1714
[ 5]Category:      5 Num:      6 Freq.:      0.1714
[ 6]Category:      6 Num:      6 Freq.:      0.1714
```

Faire l'opération inverse :



```
File aaa_c.blo contains the number of sample units for each modality
It has 6 rows (modalities) and 1 column
```

aaa_c.blo	
	1
1	5.0000
2	6.0000
3	6.0000
4	6.0000
5	6.0000
6	6.0000



Un message d'erreur apparaît si le fichier de variables qualitatives a plus de une colonne ou si les valeurs qu'on y trouve ne sont pas dans l'ordre 1, ..., 1, 2, ..., 2, ..., n , ..., n .

CategVar : Categ -> Burt



Manipulations de tableaux et tests élémentaires.



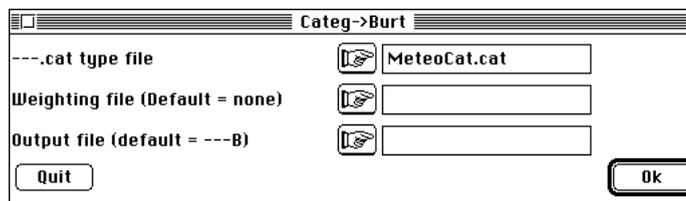
L'option permet de constituer un tableau de Burt à partir d'un tableau de variables qualitatives. Le tableau de Burt est celui qui contient pour chaque couple de modalités le nombre d'individus porteurs de ces deux modalités. Le tableau de Burt est aux variables qualitatives ce que la matrice des corrélations est aux variables quantitatives. On dispose de deux options :

1—Pondération uniforme : chaque ligne du tableau est un individu ;

2—Pondération non uniforme : chaque ligne du tableau est un groupe d'individus identiques dont on lit l'effectif dans un fichier unicolonne annexe.



L'option utilise une seule fenêtre de dialogue :



 Nom de fichier de type ---.cat décrivant un fichier de variables qualitatives.

 Option de pondération : par défaut, ce qui est le cas le plus fréquent, chaque ligne du fichier de variables qualitatives représente un individu.

 Nom du fichier de sortie : par défaut c'est le nom du fichier de variables qualitatives augmenté de "B".



Le fichier .cat doit être obtenu avec l'**option sortie complète** de Read Categ File.



Utiliser le fichier décrit sur le fiche de [CategVar : Read Categ File](#) :

```
File MeteoCatB contains Burt matrix from file MeteoCat
It has 21 rows (categories) and 21 columns (categories)
```

Éditer ce fichier avec Excel© et souligner sa structure en encadrant les blocs de cellules associés à chaque couple de variables :

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	8	0	0	6	2	0	6	2	0	4	4	0	3	5	0	2	1	5	3	3	2
2	0	10	0	4	5	1	2	5	3	2	5	3	5	4	1	4	4	2	5	3	2
3	0	0	5	0	1	4	0	0	5	1	0	4	4	0	1	3	1	1	3	1	1
4	6	4	0	10	0	0	5	4	1	5	5	0	3	6	1	1	2	7	2	4	4
5	2	5	1	0	8	0	3	3	2	2	4	2	4	3	1	4	3	1	5	2	1
6	0	1	4	0	0	5	0	0	5	0	0	5	5	0	0	4	1	0	4	1	0
7	6	2	0	5	3	0	8	0	0	5	3	0	4	4	0	2	2	4	4	2	2
8	2	5	0	4	3	0	0	7	0	1	5	1	2	5	0	2	3	2	2	4	1
9	0	3	5	1	2	5	0	0	8	1	1	6	6	0	2	5	1	2	5	1	2
10	4	2	1	5	2	0	5	1	1	7	0	0	2	4	1	0	2	5	2	2	3
11	4	5	0	5	4	0	3	5	1	0	9	0	4	4	1	3	3	3	4	3	2
12	0	3	4	0	2	5	0	1	6	0	0	7	6	1	0	6	1	0	5	2	0
13	3	5	4	3	4	5	4	2	6	2	4	6	12	0	0	8	4	0	10	2	0
14	5	4	0	6	3	0	4	5	0	4	4	1	0	9	0	1	2	6	1	5	3
15	0	1	1	1	1	0	0	0	2	1	1	0	0	0	2	0	0	2	0	0	2
16	2	4	3	1	4	4	2	2	5	0	3	6	8	1	0	9	0	0	8	1	0
17	1	4	1	2	3	1	2	3	1	2	3	1	4	2	0	0	6	0	3	3	0
18	5	2	1	7	1	0	4	2	2	5	3	0	0	6	2	0	0	8	0	3	5
19	3	5	3	2	5	4	4	2	5	2	4	5	10	1	0	8	3	0	11	0	0
20	3	3	1	4	2	1	2	4	1	2	3	2	2	5	0	1	3	3	0	7	0
21	2	2	1	4	1	0	2	1	2	3	2	0	0	3	2	0	0	5	0	0	5

Pour deux modalités distinctes de la même variable, l'effectif est évidemment nul. Le tableau de Burt juxtapose les tables de contingence élémentaire associées à un couple de variables.

File MeteoCatB.blo contains the number of modalities for each variable
It has 7 rows (categories) and 1 column

Le fichier MeteoCatB.blo est un indicateur de blocs.

```

-----
                Khi2 tests on Burt's crossed table MeteoCatB
-----
Var    2 versus var    1:Khi2 =  1.5410e+01 ddl =    4 Proba =  0.004097
-----
Var    3 versus var    1:Khi2 =  1.7907e+01 ddl =    4 Proba =  0.001416
Var    3 versus var    2:Khi2 =  1.2475e+01 ddl =    4 Proba =  0.014198
-----
Var    4 versus var    1:Khi2 =  1.0514e+01 ddl =    4 Proba =  0.032281
Var    4 versus var    2:Khi2 =  1.6429e+01 ddl =    4 Proba =  0.002656
Var    4 versus var    3:Khi2 =  1.5724e+01 ddl =    4 Proba =  0.003584
-----
...
-----
Var    7 versus var    1:Khi2 =  7.1390e-01 ddl =    4 Proba =  0.947195
Var    7 versus var    2:Khi2 =  6.5535e+00 ddl =    4 Proba =  0.159955
Var    7 versus var    3:Khi2 =  3.7316e+00 ddl =    4 Proba =  0.444877
Var    7 versus var    4:Khi2 =  4.3791e+00 ddl =    4 Proba =  0.357296
Var    7 versus var    5:Khi2 =  1.8679e+01 ddl =    4 Proba =  0.001020
Var    7 versus var    6:Khi2 =  1.8370e+01 ddl =    4 Proba =  0.001163
-----

```

L'option donne le niveau de signification du test Khi2 associé à chaque table de contingence élémentaire. Cette possibilité donne une idée générale du niveau d'association entre variables et de la pertinence d'une analyse multivariée (ACM) du tableau



En cas de données manquantes, les différentes tables constituant le tableau de Burt n'ont pas les mêmes effectifs totaux mais les tests sont encore valides.

Le fichier de pondération des individus permet de travailler sur les patrons de réponse dans une enquête plutôt que sur les réponses individuelles et d'améliorer sensiblement la qualité des représentations numériques¹.

Utiliser la carte Enquête de la pile ADE•Data pour obtenir les fichiers Ouv (319-4) et OuvEff (319-1). Ouv contient 319 configurations de réponses utilisées à 4 variables et OuvEff contient pour chaque configuration le nombre de fois qu'elle a été utilisé. Consulter l'ouvrage cité ¹ (pages

283 e suivantes) pour un descriptif des données, leur analyse et plus généralement pour le support théorique des analyses à un tableau.

Lire Ouv avec CategVar : Read Categ file :

```
Categorical variables: file Ouv
Rows: 319, Variables: 4, Categories: 32, Missing data: 0
```

Description of categories:

Variable number 1 has 8 categories

```
-----
[ 1]Category: 1 Num: 70 Freq.: 0.219
[ 2]Category: 2 Num: 42 Freq.: 0.132
[ 3]Category: 3 Num: 42 Freq.: 0.132
[ 4]Category: 4 Num: 16 Freq.: 0.0502
[ 5]Category: 5 Num: 48 Freq.: 0.15
[ 6]Category: 6 Num: 32 Freq.: 0.1
[ 7]Category: 7 Num: 32 Freq.: 0.1
[ 8]Category: 8 Num: 37 Freq.: 0.116
...

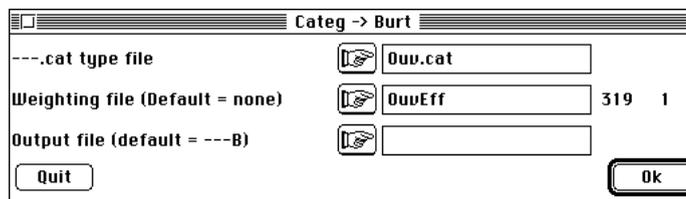
```

Variable number 4 has 8 categories

```
-----
[ 25]Category: 1 Num: 43 Freq.: 0.135
[ 26]Category: 2 Num: 78 Freq.: 0.245
[ 27]Category: 3 Num: 21 Freq.: 0.0658
[ 28]Category: 4 Num: 52 Freq.: 0.163
[ 29]Category: 5 Num: 8 Freq.: 0.0251
[ 30]Category: 6 Num: 25 Freq.: 0.0784
[ 31]Category: 7 Num: 31 Freq.: 0.0972
[ 32]Category: 8 Num: 61 Freq.: 0.191

```

Noter que cette lecture ne concerne que les patrons de réponse mais n'utilise pas les effectifs de réponse. Utiliser la présente option avec :



```
File OuvB contains Burt matrix from file Ouv
It has 32 rows (categories) and 32 columns (categories)
```

```
File OuvB.blo contains the number of modalities for each variable)
It has 4 rows (categories) and 1 column
```

185	3	2	2	1	3	147	3	155	15	2	10	31	2	37	94	148	63	18	15	1	5	17	79
4	47	0	0	0	0	41	0	8	12	4	6	24	1	22	15	8	25	1	24	1	5	11	17
2	1	19	0	1	2	57	0	10	8	1	2	19	1	23	18	5	19	2	19	3	3	12	19
0	3	0	3	0	0	20	0	1	1	1	2	3	1	16	1	2	4	2	4	0	2	9	3
8	0	2	0	15	3	84	1	12	4	2	4	28	1	39	23	5	21	4	25	2	6	26	24
7	0	0	0	1	1	151	0	16	4	0	2	18	0	40	80	11	19	7	20	0	12	23	68
6	0	0	0	3	0	101	0	17	3	0	2	15	0	39	34	14	15	3	11	1	5	27	34
9	1	1	0	1	3	98	7	14	2	0	2	11	1	29	61	10	12	8	7	1	2	15	65

On peut visualiser, dans le tableau de Burt, la bande de Burt qui croise la première variable avec les trois autres.

Tester les Khi2 des tables de continence avec CategVar: ReadTestBurt.



¹ Rouanet, H. & Le Roux, B. (1993) *Analyse des données multidimensionnelles*. Dunod, Paris. 1-310.

CategVar : Categ -> Disj



Utilitaire de manipulation de fichier.



L'option permet de constituer un tableau disjonctif complet à partir d'un tableau de variables qualitatives. Le tableau disjonctif complet est celui qui juxtapose les indicatrices (1 si l'individu porte la modalité, 0 sinon) des modalités de chaque variable.



L'option utilise une seule fenêtre de dialogue :



Nom de fichier de type ---.cat décrivant un fichier de variables qualitatives.



Utiliser le fichier décrit sur le fiche de CategVar : Read Categ File :

File MeteoCat01 contains dummy coded variable file
It has 23 rows and 21 columns (categories)

File MeteoCat01.num contains the number of modalities for each variable
It has 7 rows (variables) and 1 column

Éditer ce fichier MeteoCat01 avec Excel© et souligner sa structure en encadrant les blocs de modalités associées à chaque variables :

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	0	1	0	1	0	0	0	0	1	0	1	0	0	0	1	0	0	1	0	0	1
2	0	1	0	1	0	0	1	0	0	1	0	0	0	1	0	0	0	1	0	0	1
3	1	0	0	1	0	0	0	1	0	0	1	0	0	1	0	0	0	1	0	0	1
4	1	0	0	1	0	0	1	0	0	1	0	0	0	1	0	0	0	1	0	0	1
5	1	0	0	1	0	0	1	0	0	0	1	0	0	1	0	0	0	1	0	1	0
6	1	0	0	1	0	0	0	1	0	1	0	0	0	1	0	0	0	1	0	1	0
7	0	0	1	0	1	0	0	0	1	1	0	0	0	0	1	0	0	1	0	0	1
8	1	0	0	1	0	0	1	0	0	1	0	0	0	1	0	0	0	1	0	1	0
9	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	1	0	0
10	0	1	0	1	0	0	0	1	0	0	1	0	1	0	0	1	0	0	1	0	0
11	0	1	0	0	1	0	1	0	0	1	0	0	1	0	0	0	1	0	1	0	0
12	0	0	1	0	0	1	0	0	1	0	0	1	1	0	0	0	1	0	0	1	0
13	0	1	0	1	0	0	0	1	0	0	1	0	1	0	0	0	1	0	0	1	0
14	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	0	1	0	1	0	0
15	1	0	0	0	1	0	1	0	0	0	1	0	1	0	0	1	0	0	1	0	0
16	0	0	1	0	0	1	0	0	1	0	0	1	1	0	0	1	0	0	1	0	0
17	0	0	1	0	0	1	0	0	1	0	0	1	1	0	0	1	0	0	1	0	0
18	0	1	0	0	1	0	0	1	0	0	0	1	0	1	0	1	0	0	0	1	0
19	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0
20	0	1	0	0	1	0	0	0	1	0	0	1	1	0	0	1	0	0	1	0	0
21	1	0	0	0	1	0	1	0	0	0	1	0	1	0	0	1	0	0	1	0	0
22	0	0	1	0	0	1	0	0	1	0	0	1	1	0	0	1	0	0	1	0	0
23	0	1	0	0	0	1	0	0	1	0	0	1	1	0	0	1	0	0	1	0	0



L'AFC du tableau de disjonctif complet donne les mêmes résultats que l'ACM du tableau de départ. L'usage du tableau disjonctif complet se retrouve dans la préparation des K-tableaux (module KTabUtil).

CategVar : CrossingCateg



Utilitaire de manipulation de fichiers et tests élémentaires.



L'option donne le tableau de Burt croisé à partir de deux fichiers de variables qualitatives et les niveaux de signification des Khi2 des tables de contingence définies par une variable du premier fichier et une variable du second.



L'option utilise une seule fenêtre de dialogue :

First ---.cat type file (Table X)

Second ---.cat type file (table Y)

Weighting file (Default = none)

Output file (XtDY)

 Nom de fichier de type ---.cat décrivant un fichier de variables qualitatives X.

 Nom de fichier de type ---.cat décrivant un fichier de variables qualitatives Y.

 Option de pondération : par défaut, chaque ligne des fichiers de variables qualitatives représentent un individu.

 Nom du fichier de sortie. La notation XtDY permet de se souvenir que le tableau croisé aura en lignes les modalités de X et en colonnes les modalités de Y.



En cas de données manquantes, les différentes tables constituant le tableau de Burt croisé n'ont pas les mêmes effectifs totaux mais les tests sont encore valides.

Le fichier de pondération des individus permet de travailler sur les patrons de réponse dans une enquête plutôt que sur les réponses individuelles : voir [CategVar : Categ->Burt](#).



Utiliser la carte Mafragh+2 de la pile ADE•Data pour obtenir par [TextToBin : Char->Binary](#) les fichier Mil (97-11) et Cla (97-1). Lire ces deux fichiers avec [CategVar : Read Categ File](#). Utiliser la présente option :

```
File ClaMilBB contains crossed Burt's matrix X'DY from
coding matrix X and coding matrix Y.
It has 7 rows (categories) and 35 columns (categories)
Access for X: file Cla.cat
Access for Y: file Mil.cat
```

Le tableau ClaMilBB est le tableau de Burt croisé. Éditer ce fichier avec Excel© et souligner sa structure en encadrant les blocs de cellules associés à chaque couple de variables :

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI
1	0	0	1	5	5	0	1	0	6	0	3	3	1	2	3	0	3	1	2	0	6	0	4	1	1	2	4	1	4	1	0	3	3	0	0
2	4	2	2	17	7	10	5	3	22	3	20	5	2	18	5	1	8	12	4	11	14	2	6	11	6	16	9	3	13	9	1	17	6	1	0
3	2	1	5	0	0	2	4	2	7	1	6	2	0	5	3	0	4	3	1	5	3	1	3	1	3	3	5	1	4	3	3	5	0	0	0
4	4	4	5	8	4	6	5	6	18	3	16	5	5	13	3	3	9	5	4	18	3	3	11	5	2	9	12	6	12	3	1	6	6	4	4
5	6	8	2	4	4	5	6	5	15	5	18	2	7	12	1	5	8	5	2	19	1	8	7	4	1	11	9	8	9	3	0	4	3	3	10
6	1	1	3	2	2	3	1	1	7	0	5	2	1	5	1	4	3	0	0	7	0	6	0	1	0	7	0	6	1	0	0	0	4	3	0
7	4	2	4	0	1	2	7	0	4	6	10	0	4	6	0	4	5	1	0	10	0	7	2	1	0	10	0	7	3	0	0	1	2	4	3

Le fichier ClaMilBB.bli est l'indicateur des blocs de lignes du tableau croisé. Le fichier ClaMilBB.bco est l'indicateur des blocs de colonnes du tableau croisé.

```
File ClaMilBB.bli contains row indicator (number of modalities for each
variable)
It has 1 rows (categories) and 1 column
File ClaMilBB.bco contains column indicator (number of modalities for each
variable)
It has 11 rows (categories) and 1 column
```

```
-----
                Khi2 tests on Burt's crossed table ClaMilBB
-----
Var   1 versus var   1:Khi2 =  4.2315e+01 ddl =   18 Proba =  0.001042
Var   1 versus var   2:Khi2 =  3.0054e+01 ddl =   18 Proba =  0.036875
Var   1 versus var   3:Khi2 =  1.6035e+01 ddl =    6 Proba =  0.013652
Var   1 versus var   4:Khi2 =  7.8742e+00 ddl =    6 Proba =  0.246588
Var   1 versus var   5:Khi2 =  1.8725e+01 ddl =   12 Proba =  0.094926
Var   1 versus var   6:Khi2 =  2.8287e+01 ddl =   18 Proba =  0.057660
Var   1 versus var   7:Khi2 =  3.9469e+01 ddl =    6 Proba =  0.000001
Var   1 versus var   8:Khi2 =  4.5775e+01 ddl =   18 Proba =  0.000342
Var   1 versus var   9:Khi2 =  1.7713e+01 ddl =    6 Proba =  0.007134
Var   1 versus var  10:Khi2 =  2.7229e+01 ddl =   12 Proba =  0.007263
Var   1 versus var  11:Khi2 =  7.3432e+01 ddl =   24 Proba =  0.000001
-----
```



L'AFC du tableau de Burt croisé et la projection en individus supplémentaires des lignes des deux tableaux de départ est l'analyse canonique sur variables qualitatives de P. Cazes¹. C'est exactement l'analyse de co-inertie des deux ACM des tableaux de départ².



¹ Cazes, P. (1980) L'analyse de certains tableaux rectangulaires décomposé en blocs : généralisation des propriétés rencontrées dans l'étude des correspondances multiples. I. Définitions et applications à l'analyse canonique des variables qualitatives. II Questionnaires : variantes des codages et nouveaux calculs de contributions. *Les Cahiers de l'Analyse des Données* : 5, 145-161 & 387-406.

Cazes, P. (1981) L'analyse de certains tableaux rectangulaires décomposé en blocs : généralisation des propriétés rencontrées dans l'étude des correspondances multiples. III Codage simultané de variables qualitatives et quantitatives. IV Cas modèles. *Les Cahiers de l'Analyse des Données* : 6, 9-18 & 135-143.

² Chessel, D. & Mercier, P. (1993) Couplage de triplets statistiques et liaisons espèces-environnement. In : *Biométrie et Environnement*. Lebreton, J.D. & Asselain, B. (Eds.) Masson, Paris. 15-44.

CategVar : Quant -> Categ



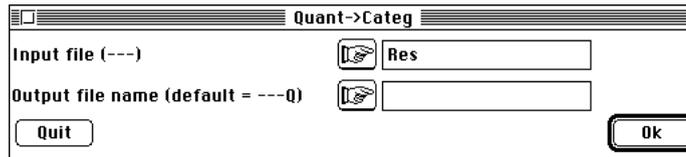
Utilitaire de changements de variables



L'option transforme des variables quantitatives en variables qualitatives par mise en classe.



L'option utilise une seule fenêtre de dialogue :

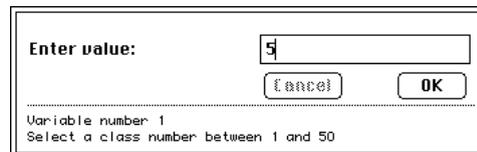


 Nom du fichier d'entrée (variables quantitatives).

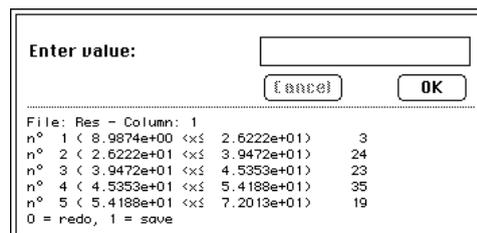
 Nom générique des fichiers de sortie (variables qualitatives). Par défaut, c'est le nom du fichier d'entrée augmenté de "Q".



L'option exécute une boucle dont chaque tour correspond à une variable du fichier de départ. Elle demande le nombre de classes :



Elle propose alors un histogramme (classes et effectifs dans chaque classe) :



Si cet histogramme convient, taper 1 : on passe à la variable suivante. Sinon, taper 0 : on recommence le tour de la boucle en entrant un nouveau nombre de classes.



La procédure est efficace pour des variables ayant une distribution suffisamment ordinaire. Les valeurs aberrantes ou une forte dissymétrie demande une transformation préalable.

On attribue pour chaque variable, le numéro de la classe à laquelle appartient l'individu. Le fichier de sortie est un fichier de variables qualitatives qui reçoit automatiquement le traitement assuré par CategVar : Read Categ File dont l'exécution est inutile.



Utiliser la carte Deug de la pile ADE•Data pour obtenir le fichier Res (94-9) :

File: Res - Column: 1

```

n° 1 [ 8.9843e+00, 3.4667e+01[ 9
n° 2 [ 3.4667e+01, 4.3839e+01[ 37
n° 3 [ 4.3839e+01, 5.0634e+01[ 30
n° 4 [ 5.0634e+01, 7.2016e+01[ 28
-----

```

```

File: Res - Column: 2
n° 1 [ 1.5990e+01, 2.9257e+01[ 27
n° 2 [ 2.9257e+01, 3.5867e+01[ 42
n° 3 [ 3.5867e+01, 4.4625e+01[ 18
n° 4 [ 4.4625e+01, 5.8011e+01[ 17
-----

```

```

...
File: Res - Column: 9
n° 1 [-3.7500e-03, 4.4512e+00[ 21
n° 2 [ 4.4512e+00, 1.0828e+01[ 22
n° 3 [ 1.0828e+01, 1.2341e+01[ 42
n° 4 [ 1.2341e+01, 1.5004e+01[ 19
-----

```

```

*****
* Description of a coding matrix *
*****
Categorical variables: file ResQ
Rows: 104, Variables: 9, Categories: 33, Missing data: 0

```

```

Description of categories:
-----
Variable number 1 has 4 categories
-----
[ 1]Category: 1 Num: 9 Freq.: 0.0865
[ 2]Category: 2 Num: 37 Freq.: 0.356
[ 3]Category: 3 Num: 30 Freq.: 0.288
[ 4]Category: 4 Num: 28 Freq.: 0.269

```

```

...
Variable number 9 has 4 categories
-----
[ 30]Category: 1 Num: 21 Freq.: 0.202
[ 31]Category: 2 Num: 22 Freq.: 0.212
[ 32]Category: 3 Num: 42 Freq.: 0.404
[ 33]Category: 4 Num: 19 Freq.: 0.183

```

```

-----
Auxiliary binary output file ResQModa: Indicator vector of modalities
It contains variable number for each modality
It has 33 rows (modalities) and one column

```

```

Auxiliary ASCII output file ResQ.123: labels (two characters) for 33
modalities
It contains one label for each modality
It has 33 rows (modalities) and labels 1a,1b, ..., 2a, 2b, ...
Variable number 1,2, ..., A, ..., Z,+, Modality number a,b, ..., z,+
-----

```



La procédure est rustique, mais on en trouve de plus raffinées dans les logiciels classiques.

CategVar : Quant -> Sign



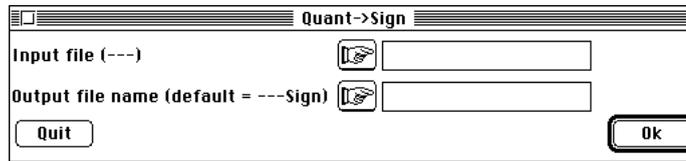
Utilitaire de changement de variables.



Réduit chaque variable quantitative du fichier d'entrée à une variable qualitative à deux modalités en faisant deux classes les plus équilibrées possibles. Les tests Khi2 issus de CategVar : Categ->Burt sont alors les tests classiques de la médiane. Les données en présence-absence donnent deux modalités (1 = absence, 2 = présence).



L'option utilise une seule fenêtre de dialogue :



 Nom du fichier d'entrée (variables quantitatives).

 Nom générique des fichiers de sortie (variables qualitatives). Par défaut, c'est le nom du fichier d'entrée augmenté de "Sign".



Utiliser la carte Deug de la pile ADE•Data pour obtenir le fichier Res (94-9) :

```
*****
* Description of a coding matrix *
*****
Categorical variables: file ResSign
Rows: 104, Variables: 9, Categories: 18, Missing data: 0

Description of categories:
-----
Variable number 1 has 2 categories
-----
[ 1]Category:  1 Num:   50 Freq.:   0.481
[ 2]Category:  2 Num:   54 Freq.:   0.519

...
Variable number 9 has 2 categories
-----
[ 17]Category:  1 Num:   47 Freq.:   0.452
[ 18]Category:  2 Num:   57 Freq.:   0.548

-----
Auxiliary binary output file ResSignModa: Indicator vector of modalities
It contains variable number for each modality
It has 18 rows (modalities) and one column

Auxiliary ASCII output file ResSign.123: labels (two characters) for 18
modalities
It contains one label for each modality
It has 18 rows (modalities) and labels 1a,1b, ..., 2a, 2b, ...
Variable number 1,2, ..., A, ..., Z,+, Modality number a,b, ..., z,+
-----
```



Le fichier de sortie est un fichier de variables qualitatives qui reçoit automatiquement le traitement assuré par CategVar : Read Categ File dont l'exécution est inutile.

CategVar : Read Categ File



Utilitaire de manipulation de fichier.

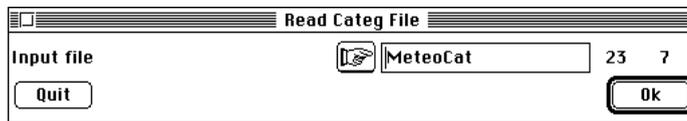


L'option constitue la première étape obligatoire pour toute utilisation d'un fichier de variables qualitatives. On désigne par variable qualitative l'enregistrement de l'information sous la forme d'un numéro de modalité.

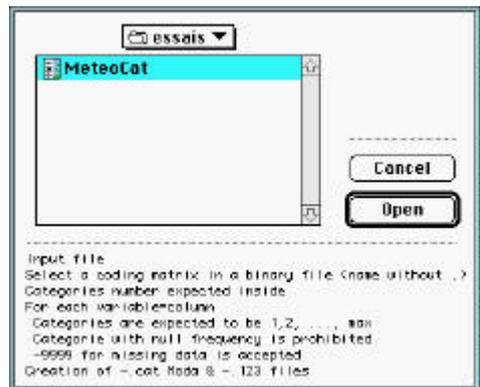
Dans un fichier de variables qualitatives, il convient de connaître le nombre de modalités par variables et le nombre d'individus par modalités dès qu'on l'ouvre. Ceci est assuré par l'exécution préliminaire (avant tout autre usage) de cette option qui crée un fichier auxiliaire ---.cat qui servira de référence pour tout appel au fichier lui-même.



L'option utilise une seule fenêtre de dialogue :



 Nom du fichier d'entrée (binaire) :



L'option n'est pas exécutable si on rencontre dans le fichier un nombre fractionnaire (seuls des entiers sont acceptables), une valeur nulle (les modalités doivent être codée 1, 2, ..., max), une valeur négative autre que -9999 (acceptée pour donnée manquante) ou une valeur supérieure à 200 (nombre maximum de modalités accepté pour une variable). L'option renvoie un message d'erreur si elle rencontre une modalité non représentée (les modalités doivent être codée 1, 2, ..., max et chaque modalité doit être portée au moins une fois).



Utiliser la carte Meteq de la pile ADE•Data pour obtenir par TextToBin : Char->Binary le fichier MeteoCat (23-7).

```
*****
* Description of a coding matrix *
*****
Categorical variables: file MeteoCat
Rows: 23, Variables: 7, Categories: 21, Missing data: 0

Description of categories:
```

```

-----
Variable number 1 has 3 categories
-----
[ 1]Category:  1 Num:    8 Freq.:  0.348
[ 2]Category:  2 Num:   10 Freq.:  0.435
[ 3]Category:  3 Num:    5 Freq.:  0.217

Variable number 2 has 3 categories
-----
[ 4]Category:  1 Num:   10 Freq.:  0.435
[ 5]Category:  2 Num:    8 Freq.:  0.348
[ 6]Category:  3 Num:    5 Freq.:  0.217
Variable number 3 has 3 categories
...
Variable number 7 has 3 categories
-----
[ 19]Category: 1 Num:   11 Freq.:  0.478
[ 20]Category: 2 Num:    7 Freq.:  0.304
[ 21]Category: 3 Num:    5 Freq.:  0.217

-----
Auxiliary binary output file MeteoCatModa: Indicator vector of modalities
It contains variable number for each modality
It has 21 rows (modalities) and one column

Auxiliary ASCII output file MeteoCat.123: labels (two characters) for 21
modalities
It contains one label for each modality
It has 21 rows (modalities) and labels 1a,1b, ..., 2a, 2b, ...
Variable number 1,2, ..., A, ..., Z,+, Modality number a,b, ..., z,+
-----

```

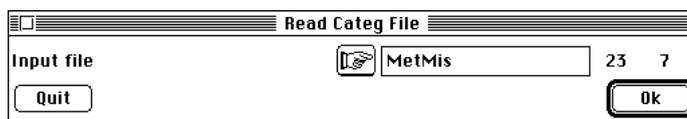
Le fichier MeteoCat.cat contient de l'information sous forme texte, respectivement le nombre d'individus (23), le nombre de variables (7), le nombre de modalités (21), le nombre de modalités de la variable 1 (3), ..., le nombre de modalités de la variable 7 (3), le nombre de porteurs des modalités de la variable 1 (8, 10, 5), ..., le nombre de porteurs des modalités de la variable 9 (11, 7, 5).

Le fichier MeteoCat.123 contient des étiquettes pour les modalités dans le format : 1a, 1b, 1c, 2a, 2b, 2c, 3a, 3b, 3c ...

Les trois fichiers MeteoCat (21-1), MeteoCat.123 et MeteoCat.cat (TEXT) contiennent des informations relatives aux modalités sous la forme de la variable qualitative qui donne à chaque modalité le numéro de la variable à laquelle elle appartient. Ces fichiers peuvent être utilisés dans certains modules graphiques quand on veut répartir les modalités par fenêtre.



Le module accepte les données manquantes dans le fichier de départ : le numéro de modalité est alors codé -9999. Utiliser la carte MetMissing de la pile ADE•Data pour obtenir par TextToBin : StatView©->Binary le fichier MetMis (23-7). Dans le fichier texte les données manquantes sont saisies par • et le symbole dans le fichier binaire donne la valeur -9999.



```

*****
* Description of a coding matrix *
*****
Categorical variables: file MetMis

```

Rows: 23, Variables: 7, Categories: 21, Missing data: 8

Description of categories:

Variable number 1 has 3 categories

[1]Category: 1 Num: 9 Freq.: 0.391
[2]Category: 2 Num: 8 Freq.: 0.348
[3]Category: 3 Num: 5 Freq.: 0.217
[]Missing data Num: 1 Freq.: 0.0435

Variable number 2 has 3 categories

[4]Category: 1 Num: 7 Freq.: 0.304
[5]Category: 2 Num: 10 Freq.: 0.435
[6]Category: 3 Num: 5 Freq.: 0.217
[]Missing data Num: 1 Freq.: 0.0435
...

Variable number 7 has 3 categories

[19]Category: 1 Num: 9 Freq.: 0.391
[20]Category: 2 Num: 5 Freq.: 0.217
[21]Category: 3 Num: 8 Freq.: 0.348
[]Missing data Num: 1 Freq.: 0.0435

Auxiliary binary output file MetMisModa: Indicator vector of modalities
It contains variable number for each modality
It has 21 rows (modalities) and one column

Auxiliary ASCII output file MetMis.123: labels (two characters) for 21 modalities

It contains one label for each modality
It has 21 rows (modalities) and labels 1a,1b, ..., 2a, 2b, ...
Variable number 1,2, ..., A, ..., Z,+, Modality number a,b, ..., z,+

Le fichier MetMis.cat contient de l'information sous forme texte, respectivement le nombre d'individus (23), le nombre de variables (7), le nombre de modalités (21), le nombre de modalités de la variable 1 (3), ..., le nombre de modalités de la variable 7 (3), le nombre de porteurs des modalités de la variable 1 (9, 8, 5), ..., le nombre de porteurs des modalités de la variable 9 (9, 5, 8).



Ne pas oublier que :

```
*****  
* Les programmes qui manipulent des fichiers de *  
* variables qualitatives supposent l'exécution *  
* préalable de ce programme qui permet en outre *  
* d'éditer les codes modalité. *  
*****
```

Voir aussi CategVar : VerifCateg en cas de difficultés avec les numéros de modalités.

CategVar : Read Label File



Utilitaire de lecture de fichier.



L'objectif est de lire un tableau de variables qualitatives saisi uniquement par des chaînes de caractères. L'option équivaut à CategVar:Read Categ File sur un tableau de variables qualitatives saisi par numéro de modalités et passé en binaire.



L'option utilise une seule fenêtre de dialogue :



Nom du fichier texte d'entrée. **Il doit porter un nom du type ---.lab.**



Nom du fichier binaire de sortie (création). Par défaut c'est ---.



Implanter un fichier dans un tableur :

	1	2	3	4	5
1	non	JT	aucun	1	a
2	oui	JT	aucun	1	b
3	non	DC	aucun	5	c
4	non	DC	aucun	5	d
5	non	SD	peu	5	e
6	oui	JT	beaucoup	1	f
7	oui	DC	beaucoup	1	g
8	oui	SD	peu	5	h
9	oui	DC	peu	5	i

Sauvegarder en texte seulement avec **séparateur tabulation (impératif)**.

```
-----  
|      Description of a coded matrix      |  
-----  
Qualitative variables file: essalabels  
Number of rows: 9, variables: 5, categories: 19  
  
Description of categories:  
-----  
Variable number 1 has 2 categories  
-----  
[  1]Category:   1 Num:    4 Freq.:   0.4444  
[  2]Category:   2 Num:    5 Freq.:   0.5556  
  
Variable number 2 has 3 categories  
-----  
[  3]Category:   1 Num:    3 Freq.:   0.3333  
...  
  
Variable number 5 has 9 categories  
-----  
[ 11]Category:   1 Num:    1 Freq.:   0.1111  
[ 12]Category:   2 Num:    1 Freq.:   0.1111  
[ 13]Category:   3 Num:    1 Freq.:   0.1111  
[ 14]Category:   4 Num:    1 Freq.:   0.1111  
[ 15]Category:   5 Num:    1 Freq.:   0.1111
```

```
[ 16]Category: 6 Num: 1 Freq.: 0.1111
[ 17]Category: 7 Num: 1 Freq.: 0.1111
[ 18]Category: 8 Num: 1 Freq.: 0.1111
[ 19]Category: 9 Num: 1 Freq.: 0.1111
```

```
-----
Auxiliary ASCII output file essalabels.123: labels for modalities
-----
```



Le fichier .123 ainsi créé attribue aux modalités les étiquettes du fichier de départ :



La présente option étend TextToBin: LabelToCateg à plusieurs colonnes comme souhaité sur Adelist par Nicolas DEGALLIER (degallie@solar.com.br) :

Subject: A propos de TextToBin/LabelToCateg

J'aimerais savoir s'il est normal de ne pouvoir transformer qu'une seule colonne à la fois avec l'option LabelToCateg de TextToBin. J'ai un tableau Excel/Text de 130 lignes (individus) sur 11 colonnes (variables), ces dernières étant codées par modalités non numériques (comme l'exemple de la fiche - mise à jour 19/9/95 - : Pierre, Paul, Jean etc.).

Je n'ai pas réussi à traiter le tout en une fois. J'ai dû transformer mes données alphabétiques en numériques, avec Excel, puis les traiter avec TextToBin/Text-Binary puis CategVar:ReadCategFile... c'est un peu fastidieux et on peut faire des erreurs.

Où alors y-aurait-il une autre option ou un autre module, ou une astuce qui permettrait de le faire?



Modification du 11/11/99 : l'option accepte la chaîne de caractères «**unknown**» pour coder une donnée manquante qui sera enregistrée -9999.

CategVar : Read Taxa File



Utilitaire de lecture et représentation d'un fichier taxonomique.



Un fichier taxonomique est un fichier texte (séparateur tabulation) qui contient le descriptif taxonomique d'une liste d'espèce compilée par chaîne de caractères.

	1	2	3	4	5	6
1	Bivalves	Yeneroide	Sphaeriidae	Pisidium	Pisidium	Pisidium casertanum
2	Bivalves	Yeneroide	Sphaeriidae	Pisidium	Pisidium	Pisidium heroldianum
3	Bivalves	Yeneroide	Sphaeriidae	Pisidium	Pisidium	Pisidium hibernicum
4	Bivalves	Yeneroide	Sphaeriidae	Pisidium	Pisidium	Pisidium milium
5	Bivalves	Yeneroide	Sphaeriidae	Pisidium	Pisidium	Pisidium nitidum
6	Bivalves	Yeneroide	Sphaeriidae	Pisidium	Pisidium	Pisidium obtusale
7	Bivalves	Yeneroide	Sphaeriidae	Pisidium	Pisidium	Pisidium personatum
8	Bivalves	Yeneroide	Sphaeriidae	Pisidium	Pisidium	Pisidium pseudosphaerium
9	Bivalves	Yeneroide	Sphaeriidae	Pisidium	Pisidium	Pisidium subtruncatum
10	Bivalves	Yeneroide	Sphaeriidae	Sphaerium	Sphaerium	Sphaerium cornutum
11	Bivalves	Yeneroide	Sphaeriidae	Sphaerium	Sphaerium	Sphaerium lacustre
12	Gastéropodes	Caenogastropode	Bithyniidae	Bithynia	Bithynia	Bithynia tentaculata
13	Gastéropodes	Caenogastropode	Valvatidae	Valvata	Valvata	Valvata cristata
14	Gastéropodes	Hydrophile	Acroloxidae	Acroloxus	Acroloxus	Acroloxus lacustris
15	Gastéropodes	Hydrophile	Lymnaeidae	Galba	Galba	Galba corvus
16	Gastéropodes	Hydrophile	Lymnaeidae	Lymnaea	Lymnaea	Lymnaea stagnalis

Les ordres de détermination sont rangés par colonnes du plus grossier au plus fin. Dans l'exemple ci-dessus dans la première colonne on trouve la classe, dans la seconde l'ordre, puis la famille, le genre et l'espèce. On peut utiliser les niveaux que l'on veut à condition de respecter les règles suivantes. Dans la dernière colonne, chacune des chaînes de caractères doivent être uniques et il vaut donc mieux utiliser le binom linéen (ou un numéro de catalogue unique). On suppose que deux genres d'étiquettes identiques ne sont pas présents dans deux ordres différents. Chacune des colonnes définit une variable qualitative exactement comme dans CategVar: Read Label File. Mais l'option testera que chaque partition est moins fine que la suivante (une classe de la colonne j est entièrement contenue dans une classe de la colonne $j - 1$, à tous les niveaux). Sauvegarder en texte (séparateur tabulation) et vérifier le fichier comme d'habitude (attention aux tabulations indésirables, au retour-charriot en fin de fichier, à l'unicité des orthographes, ...). **Donner au fichier texte un nom avec l'extension .tax** qui précise clairement l'intention.



L'option utilise une seule fenêtre de dialogue :



Nom du fichier binaire d'entrée (.tax).



Implanter dans le fichier aa l'information suivante :

X	A	a	a1
X	A	a	a2
X	A	b	b1
X	B	c	c1
X	B	d	d1
Y	C	e	e1
Y	C	e	e2
Y	C	e	e3

qui simule une liste faunistique à 8 espèces, 5 genres, 3 familles et 2 ordres.
L'option assure les fonctions de lecture :

```
-----
|      Description of a coded matrix      |
-----
Qualitative variables file: aa
Number of rows: 8, variables: 5, categories: 19

Description of categories:
-----
Variable number 1 has 1 categories
-----
[  1]Category:      1 Num:      8 Freq.:      1
```

La variable qui prend la valeur 1 pour toutes les lignes est ajoutée en tête du tableau. Elle définit le niveau de référence. On appellera sa modalité unique ROOT.

```
-----
Variable number 2 has 2 categories
-----
[  2]Category:      1 Num:      5 Freq.:      0.625
[  3]Category:      2 Num:      3 Freq.:      0.375

Variable number 3 has 3 categories
-----
[  4]Category:      1 Num:      3 Freq.:      0.375
[  5]Category:      2 Num:      2 Freq.:      0.25
[  6]Category:      3 Num:      3 Freq.:      0.375

Variable number 4 has 5 categories
-----
[  7]Category:      1 Num:      2 Freq.:      0.25
[  8]Category:      2 Num:      1 Freq.:      0.125
[  9]Category:      3 Num:      1 Freq.:      0.125
[ 10]Category:      4 Num:      1 Freq.:      0.125
[ 11]Category:      5 Num:      3 Freq.:      0.375

Variable number 5 has 8 categories
-----
[ 12]Category:      1 Num:      1 Freq.:      0.125
[ 13]Category:      2 Num:      1 Freq.:      0.125
[ 14]Category:      3 Num:      1 Freq.:      0.125
[ 15]Category:      4 Num:      1 Freq.:      0.125
[ 16]Category:      5 Num:      1 Freq.:      0.125
[ 17]Category:      6 Num:      1 Freq.:      0.125
[ 18]Category:      7 Num:      1 Freq.:      0.125
[ 19]Category:      8 Num:      1 Freq.:      0.125
-----
```

Le fichier est lu comme un fichier de variables qualitatives saisies par étiquettes. **Un fichier .cat est créé :**

```
8  Nombre de lignes
5  Nombre de variables
19 Nombre de modalités total
1  Nombre de modalité de la variable 1 (root)
2  Nombre de modalité de la variable 2 (2 ordres)
3  3 familles
5  5 genres
8  8 espèces
8  Nombre de lignes portant la modalité 1 de la variable 1 (unique)
5  Nombre de lignes portant la modalité 1 de la variable 2
3
...
```

1
1 *Nombre de lignes portant la modalité 8 de la variable 4*

Auxiliary ASCII output file aa.123: labels for modalities

Un fichier .123 est créé :

```
ROOT   Label de la racine
X      Label de l'ordre 1
Y      Label de l'ordre 2
A
B
C      Label de la famille 3
a      Label du genre 1
...
a1     Label de l'espèce 1
a2
...
e2
e3     Label de l'espèce 8
```

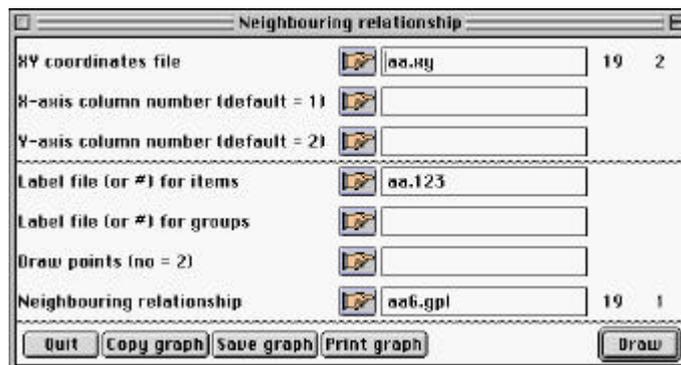
La chaîne de caractères indique que les contraintes d'utilisation sont respectées :

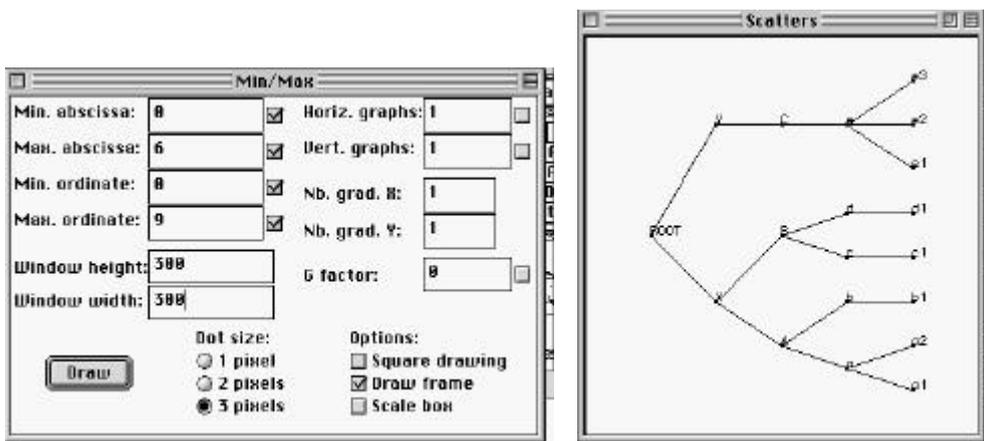
Convenient file for taxonomic hierarchy

Sinon des indications sont données sur l'origine du problème.

```
Category coordinates for graphical display if taxonomic tree in aa.xy
Neighbouring relationship in text file: aaG
It contains graph matrix (LEBART's M) with 19 rows and columns
Neighbouring weights in binary file: aaG.gpl
It contains 19 rows and 1 column
```

Les fichiers créés permettent de représenter l'arbre phylétique avec Scatters:Neighbouring relationship :





Sauvegarder en PICT et améliorer avec un grapheur. Voir un exemple en vraie grandeur dans la fiche Thema83.

File aa.@ob contains an orthonormal basis
It has 8 rows and 8 columns

La partition définit une base orthonormée dans l'espace des espèces qui sera utile dans certaines circonstances comme dans DMAUse.



Fonctionne comme CategVar: Read Label File, mais s'adresse à un type spécial d'information. Les fichiers .tax seront appelés dans d'autres options : leur utilisation supposera l'emploi préalable de celle-ci. Permet de représenter les niveaux taxonomiques sur les cartes factorielles d'espèces par le fichier .cat.



Il n'est pas indispensable de ranger les espèces d'un même genre, les genres d'une même famille ... par blocs consécutifs mais, pour une expression graphique optimale, c'est préférable.

CategVar : ReadTestBurt



Tests élémentaires sur les tables de contingence.



L'option lit directement un tableau de Burt et exécute les tests Khi2 sur chaque sous-table de contingence associée à chaque couple de deux variables qualitatives. Les tableaux de Burt sont constitués à partir des fichiers de variables qualitatives par CategVar : Categ->Burt. On suppose ici qu'on a obtenu et saisi directement de tels tableaux, fréquents dans la littérature.



L'option utilise une seule fenêtre de dialogue :



Nom du fichier d'entrée (binaire).



Nom du fichier indicateur des blocs de lignes (contient sur une colonne les effectifs des lignes dans chaque bloc).



Nom du fichier indicateur des blocs de colonnes (contient sur une colonne les effectifs des colonnes dans chaque bloc).



L'exemple ci-dessous est proposé par A. Leclerc¹ pour illustrer les propriétés de l'AFC des tableaux de Burt :

		Age du chef de ménage				Profession du chef de ménage								Ancienneté logement			
		1	2	3	4	1	2	3	4	5	6	7	8	1	2	3	4
Médecin de famille	oui	93	379	291	233	76	200	172	122	150	160	59	57	122	169	188	517
	non	145	226	114	77	31	125	105	59	55	99	46	42	160	131	93	178
Dentiste habituel	oui	128	417	296	209	79	262	199	130	129	133	52	66	156	192	201	501
	non	110	188	109	101	28	63	78	51	76	126	53	33	126	108	80	194

Utiliser la carte Ménages de la pile ADE•Data pour obtenir le fichier MenBurt (4-16) et les fichiers MenBloLig (2-1) et MenBloCol (3-1) qui définissent les blocs :

```

-----
                Khi2 tests on Burt's crossed table MenBurt
-----
Var   1 versus var   1:Khi2 =  9.2169e+01 ddl =    3 Proba =  0.000000
Var   1 versus var   2:Khi2 =  1.7069e+01 ddl =    7 Proba =  0.017007
Var   1 versus var   3:Khi2 =  9.3787e+01 ddl =    3 Proba =  0.000000
-----
Var   2 versus var   1:Khi2 =  2.6687e+01 ddl =    3 Proba =  0.000013
Var   2 versus var   2:Khi2 =  7.9459e+01 ddl =    7 Proba =  0.000000
Var   2 versus var   3:Khi2 =  2.9435e+01 ddl =    3 Proba =  0.000004
-----

```



Leclerc, A. (1975) L'analyse des correspondances sur juxtaposition de tableaux de contingence. *Revue de Statistique Appliquée* : 23, 5-16.

CategVar : Recoding



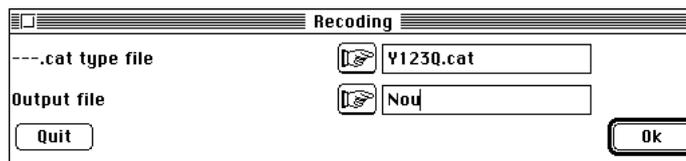
Utilitaire de changement de variables.



L'option permet de permuter et/ou regrouper des modalités. L'opération s'impose pour éliminer les modalités rares par regroupement.



L'option utilise une seule fenêtre de dialogue :

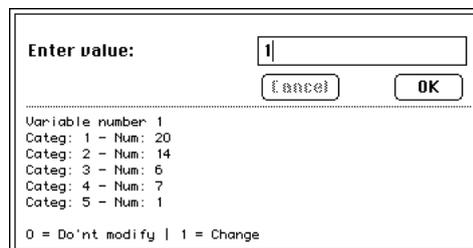


 Nom de fichier de type ---.cat décrivant un fichier de variables qualitatives.

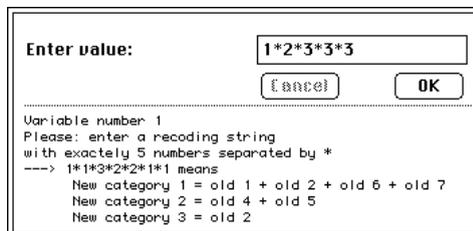
 Nom générique des fichiers de sortie.



L'option exécute une boucle dont chaque tour correspond à une variable du fichier de départ. Elle demande si la variable doit être recodée. Taper 0 pour passer à la variable suivante, 1 pour recoder cette variable :



La fenêtre de saisie affiche au plus les effectifs des huit premières modalités. En cas de dépassement, se reporter au listing de lecture qui correspond au fichier ---.cat sélectionné. Si l'option 1 est choisie, le programme demande l'entrée d'un message de recodage :



Un tel message est une chaîne de caractères formée d'entiers qui attribuent à chaque modalité de la variable de départ un numéro de modalité de la variable d'arrivée. Il ne peut y avoir plus de modalités après que de modalités avant le recodage. Les modalités finales doivent être numérotées de 1 à max et être toutes utilisées ou moins une fois. En cas d'erreur de frappe ou d'incohérence du message le choix est à nouveau proposé à moins qu'on quitte le programme.

Si le message est interprétable, le recodage est effectué en mémoire et on passe à la variable suivante. Les données manquantes (-9999) sont conservées.



Utiliser la carte Monde de la pile ADE•Data pour obtenir le fichier Y123 (48-3). Transformer en variables qualitatives par CategVar : Quant->Categ :

Imposer sans plus de considération cinq classes pour chaque variable :

On obtient :

```
File: Y123 - Column: 1
n° 1 [ 6.9598e+00, 3.1531e+01[      20
n° 2 [ 3.1531e+01, 8.9722e+01[      14
n° 3 [ 8.9722e+01, 1.1767e+02[       6
n° 4 [ 1.1767e+02, 1.5517e+02[       7
n° 5 [ 1.5517e+02, 2.0804e+02[       1
-----
File: Y123 - Column: 2
n° 1 [-1.8000e-02, 8.8438e+00[      20
n° 2 [ 8.8438e+00, 3.4400e+01[      12
...
-----
File: Y123 - Column: 3
n° 1 [ 7.9820e+00, 2.6083e+01[       6
n° 2 [ 2.6083e+01, 4.7235e+01[       9
n° 3 [ 4.7235e+01, 6.1632e+01[       7
n° 4 [ 6.1632e+01, 8.3360e+01[      13
n° 5 [ 8.3360e+01, 9.8018e+01[      13
-----
```

On décide alors de regrouper les modalités 3, 4 et 5 de la variable 1, les modalités 2 et 3, puis 4 et 5 de la variable 2, enfin les modalités 1, 2 et 3 de la variable 3. Lancer l'option :

Demander le recodage de la variable 1 :

Entrer le message de recodage :

Enter value:

Variable number 1
Please: enter a recoding string
with exactly 5 numbers separated by *
----> 1*1*3*2*2*1*1 means
New category 1 = old 1 + old 2 + old 6 + old 7
New category 2 = old 4 + old 5
New category 3 = old 2

Refaire la même opération pour la variable 2 :

Enter value:

Variable number 2
Please: enter a recoding string
with exactly 5 numbers separated by *
----> 1*1*3*2*2*1*1 means
New category 1 = old 1 + old 2 + old 6 + old 7
New category 2 = old 4 + old 5
New category 3 = old 2

Achever avec la troisième variable :

Enter value:

Variable number 3
Please: enter a recoding string
with exactly 5 numbers separated by *
----> 1*1*3*2*2*1*1 means
New category 1 = old 1 + old 2 + old 6 + old 7
New category 2 = old 4 + old 5
New category 3 = old 2

Input file: Y123Q

Variable number 1 modified with the message 1*2*3*3*3
Variable number 2 modified with the message 1*2*2*3*3
Variable number 3 modified with the message 1*1*1*2*3

Output file: Nou

```
*****
* Description of a coding matrix *
*****
Categorical variables: file Nou
Rows: 48, Variables: 3, Categories: 9, Missing data: 0
```

Description of categories:

Variable number 1 has 3 categories

[1]Category: 1 Num: 20 Freq.: 0.417
[2]Category: 2 Num: 14 Freq.: 0.292
[3]Category: 3 Num: 14 Freq.: 0.292

Variable number 2 has 3 categories

[4]Category: 1 Num: 20 Freq.: 0.417
[5]Category: 2 Num: 17 Freq.: 0.354
[6]Category: 3 Num: 11 Freq.: 0.229

Variable number 3 has 3 categories

[7]Category: 1 Num: 22 Freq.: 0.458
[8]Category: 2 Num: 13 Freq.: 0.271
[9]Category: 3 Num: 13 Freq.: 0.271

Auxiliary binary output file NouModa: Indicator vector of modalities
It contains variable number for each modality
It has 9 rows (modalities) and one column

Auxiliary ASCII output file Nou.123: labels (two characters) for 9 modalities
It contains one label for each modality
It has 9 rows (modalities) and labels 1a,1b, ..., 2a, 2b, ...
Variable number 1,2, ..., A, ..., Z,+, Modality number a,b, ..., z,+



Le fichier de sortie est une fichier de variables qualitatives qui reçoit automatiquement le traitement assuré par CategVar : Read Categ File dont l'exécution est inutile.



L'analyse des correspondances multiples (ACM) est très sensible aux modalités faiblement représentées : les éliminer est, dans la plupart des cas, impératif. De gros écarts entre nombres de modalités par variables sont, d'autre part, fortement déconseillés. Le codage et le recodage minutieux des variables est la base du traitement des variables qualitatives. Ils doivent être fait *a priori* pour garder une signification statistique.

CategVar : VerifCateg



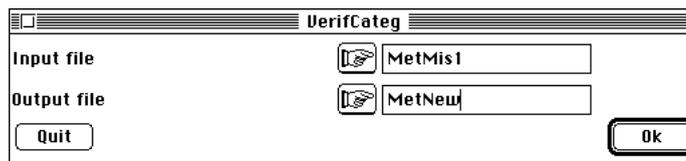
Utilitaire de changement de variables.



L'option élimine la modalité 0, parfois utilisée dans certains logiciels, et les modalités non représentées en renumérotant l'ensemble des modalités dans l'ordre naturel, en commençant à 1.



L'option utilise une seule fenêtre de dialogue :

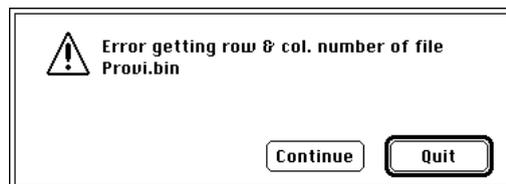
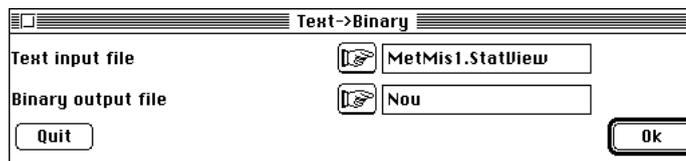


 Nom du fichier binaire d'entrée.

 Nom générique des fichiers de sortie.

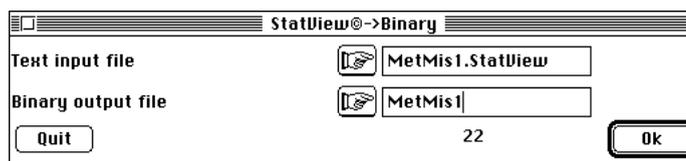


Utiliser la carte MetMissing de la pile ADE•Data. Obtenir le fichier texte MetMis1.StatView. La lecture par TextToBin : Text->Binary envoie le message :

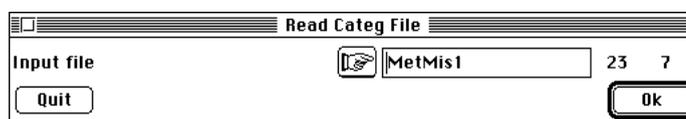


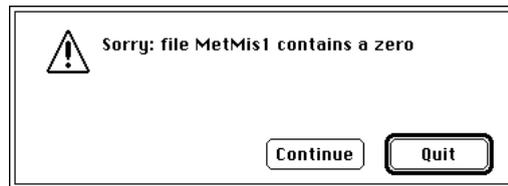
```
Rec1: file MetMis1.StatView does not contain a
constant number of numerical figures by row
-----
No file creation
```

Utiliser TextToBin : StatView@->Binary :

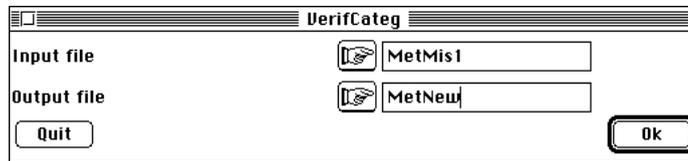


MetMis1 est un fichier binaire (23-7). La lecture de ce fichier par CategVar : Read categ File envoie le message :





Utiliser alors la présente option :



Chaque variable est analysée. La transformation effectuée est indiquée :

```
File: MetMis1 - Column: 1
Zero is not a valid number for category
Some categories are missing
old =    0 -> new =    1 || old =    1 -> new =    2 || old =    4 -> new =    3 ||
-----
File: MetMis1 - Column: 2
Zero is not a valid number for category
Some categories are missing
old =    0 -> new =    1 || old =    1 -> new =    2 || old =    4 -> new =    3 ||
-----
...
-----
Input data MetMis1
It has 23 rows and 7 columns
Output data MetNew
It has 23 rows and 7 columns
*****
* Description of a coding matrix *
*****
Categorical variables: file MetNew
Rows: 23, Variables: 7, Categories: 21, Missing data: 8

Description of categories:
-----
Variable number 1 has 3 categories
-----
[ 1]Category:  1 Num:    9 Freq.:  0.391
[ 2]Category:  2 Num:    8 Freq.:  0.348
[ 3]Category:  3 Num:    5 Freq.:  0.217
[   ]Missing data Num:    1 Freq.:  0.0435
-----
...
-----
Variable number 7 has 3 categories
-----
[ 19]Category:  1 Num:    9 Freq.:  0.391
[ 20]Category:  2 Num:    5 Freq.:  0.217
[ 21]Category:  3 Num:    8 Freq.:  0.348
[   ]Missing data Num:    1 Freq.:  0.0435
-----
...
-----
```



Le fichier de sortie est un fichier de variables qualitatives qui reçoit automatiquement le traitement assuré par CategVar : Read Categ File dont l'exécution est inutile. Les données manquantes (-9999) sont conservées.