

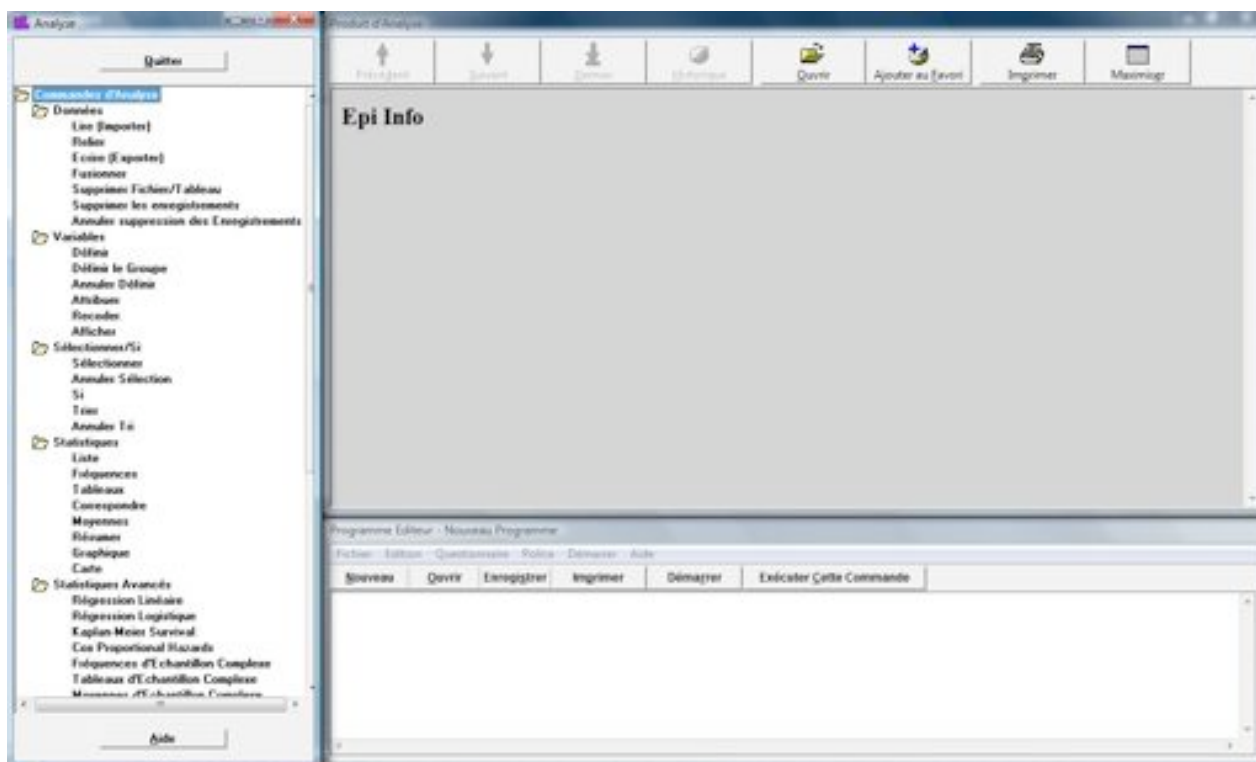
IV. ANALYSE des DONNEES AVEC le MODULE ANALYSE

Le module **ANALYSE Des Données** permet de visualiser les données, de calculer des fréquences, des tableaux croisés, des statistiques, de tester des modèles et de produire différents types de graphiques. Il s'utilise en entrant des commandes. Ces commandes peuvent être tapées une par une à partir du clavier, ou sélectionnées avec la souris, à partir d'une liste.

Pour lancer le module ANALYSE Des DONNEES

☞ Cliquez sur le bouton **Analyse Des Données** dans la fenêtre d'accueil d'Epi Info.

L'écran du module ANALYSE Des DONNEES



L'écran du module ANALYSE Des DONNEES comporte trois fenêtres:
la fenêtre "Analyse" sur la gauche, montrant toutes les commandes disponibles;
la fenêtre "Produit d'Analyse" où les résultats des commandes utilisées vont apparaître;
la fenêtre "Programme Editeur" où les commandes que vous sélectionnez vont s'écrire.

Pour sélectionner des commandes, cliquer dessus dans la fenêtre de gauche, une boîte de dialogue apparaîtra pour compléter la commande qui s'écrira dans l'écran en dessous.

☞ Pour quitter le module ANALYSE à tout moment, cliquez sur le bouton **Quitter** en haut à gauche.

Les commandes principales du module ANALYSE Des DONNEES

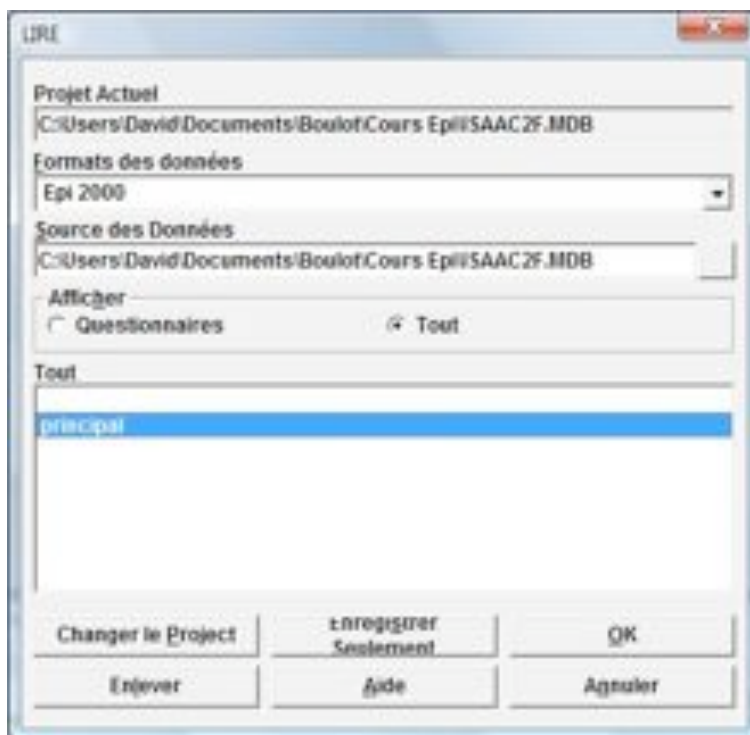
Pour apprendre à utiliser les commandes principales du module d'analyse, nous allons prendre comme exemple un extrait (incomplet, donc les résultats sont fictifs) d'une enquête réalisée chez des enfants scolarisés en France, pour étudier la relation entre des symptômes respiratoires et allergiques et la pollution atmosphérique des salles de classe de leur école. Ces données sont dans le fichier "Isaac2f.mdb".

LIRE

Tout d'abord, vous avez besoin de sélectionner la base de données sur laquelle vous allez travailler, avec la commande LIRE.

Pour ouvrir le projet appelé "Isaac2f.mdb":

- ☞ Cliquez sur la commande **Lire (Importer)** depuis la fenêtre de gauche, dans le groupe des commandes **■ Données**;
- ☞ dans la boîte de dialogue LIRE, cliquez sur le bouton **Changer le Project** et sélectionnez le projet (ici "Isaac2f.mdb"), dont le nom doit alors apparaître dans le champ appelé **Source des Données**;
- ☞ puis choisissez le format **Epi 2000** depuis le menu déroulant appelé "**Formats des données**";
- ☞ sélectionnez "**Tout**" dans le champ appelé "**Afficher**";
- ☞ choisissez une vue dans la liste en dessous (ici "principal"), et cliquez sur **OK**.



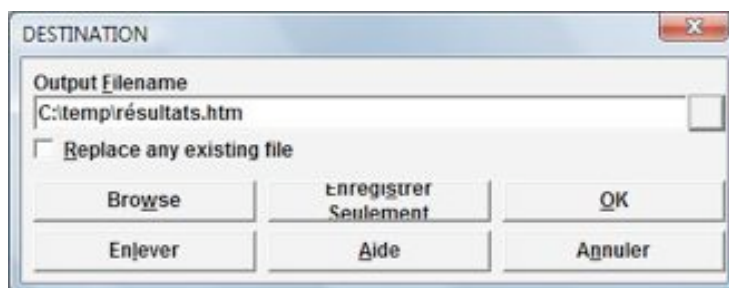
DESTINATION des résultats, FERMER fichier résultats

La commande **Destination** suivie du nom d'un fichier de votre choix, envoie tous les résultats vers ce fichier (de type HTML) et vers l'écran simultanément. Ainsi vous pourrez relire tous les résultats de vos analyses avec n'importe quel éditeur HTML (comme Internet Explorer ou FireFox).

La commande **Fermer** permet de refermer le fichier résultats après l'avoir rempli.

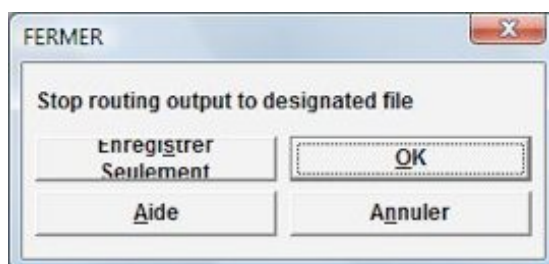
Pour envoyer tous les résultats vers un fichier appelé "resultats.htm", sur le répertoire "c:\temp\" :

- ☞ cliquez sur la commande **Destination** qui se trouve dans la fenêtre de gauche, dans le groupe des commandes ■ Sortie;
- ☞ dans la boîte de dialogue DESTINATION, dans le champ appelé "**Output Filename**", cliquez sur le bouton à droite ;
- ☞ dans la fenêtre de navigation qui apparaît, sélectionner le répertoire dans lequel se trouvera le fichier résultat "c:\temp\", puis tapez le nom du fichier "**résultats.htm**", et cliquez sur le bouton ouvrir.
- ☞ De retour sur la boîte de dialogue DESTINATION, cliquez sur le bouton **OK**.



Pour refermer ensuite le fichier et envoyer les sorties vers l'écran seulement :

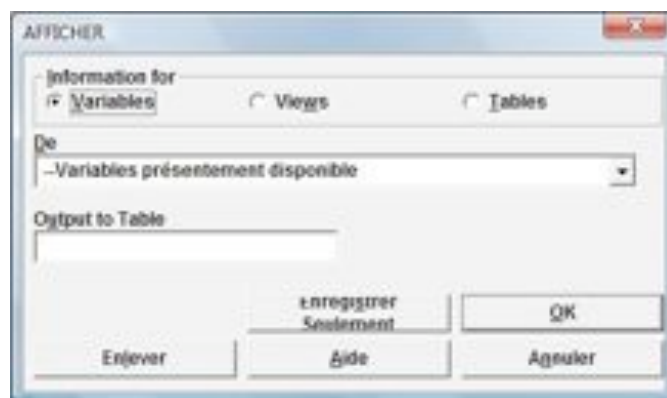
- ☞ cliquez sur la commande **Fermer** qui se trouve dans la fenêtre de gauche, dans le groupe des commandes ■ Sortie;
- ☞ dans la fenêtre de dialogue FERMER qui apparaît, cliquez juste sur le bouton **OK**.



AFFICHER VARIABLES

Une fois que vous avez sélectionné une table de données, la commande **Afficher variables**, avec l'option "Variables actuellement disponibles", donne le nom de la table, et la liste des variables avec leur nom et leur type.

- ☞ Cliquez sur la commande **Afficher** depuis la fenêtre de gauche "Analyse", dans le groupe des commandes ■ Variables;
- ☞ dans la boîte de dialogue AFFICHER, sélectionnez "Variables présentement disponibles" dans le menu déroulant appelé "De";
- ☞ cliquez sur **OK**.



⇒ Résultat de la commande **Afficher variables**:

Variable	Tableau	Type du champ	Format/Valeur	Info Spéciale	Invite
age	principal	Nombre			age
crisessiff	principal	Texte			crisessiff
dateentrevue	principal	Date			dateentrevue
eczema	principal	OuiNon			eczema
identification	principal	Nombre			identification
PM2_5	principal	Nombre			PM2_5
poids	principal	Nombre			poids
sexe	principal	Texte			sexe
sifflements	principal	OuiNon			sifflements
taille	principal	Nombre			taille
Language	Defined	Texte	FRANCAIS	Predefined	

Le résultat se présente sous forme d'un tableau.

Dans la première colonne apparaissent les noms des variables du questionnaire, dans la deuxième le nom du masque de saisie (ici principal) ; puis le type de champ (le type de la variable) ; puis deux colonnes avec des informations complémentaires qui sont vides ici ; enfin apparaît l'invite, soit le texte de la question correspondant à la variable.

Le questionnaire principal du projet "Isaac2f.mdb", ressemble à celui créé dans la partie II "Création d'un questionnaire" de ce document (cf p 10).

On retrouve le nom des variables de la partie identification :

Identification : numéro d'identification ;

Dateentrevue : date de l'entrevue ;

Age : âge des enfants en année ;

Sexe : sexe des enfants.

Le nom des variables de la partie symptômes respiratoires :

Sifflements : avoir des sifflements dans la poitrine ;

crisessiff : la fréquence des sifflements ;

eczema : avoir de l'eczéma.

Trois variables apparaissent en plus par rapport à l'exemple de questionnaire page 10 :

Poids : poids des enfants en kg ;

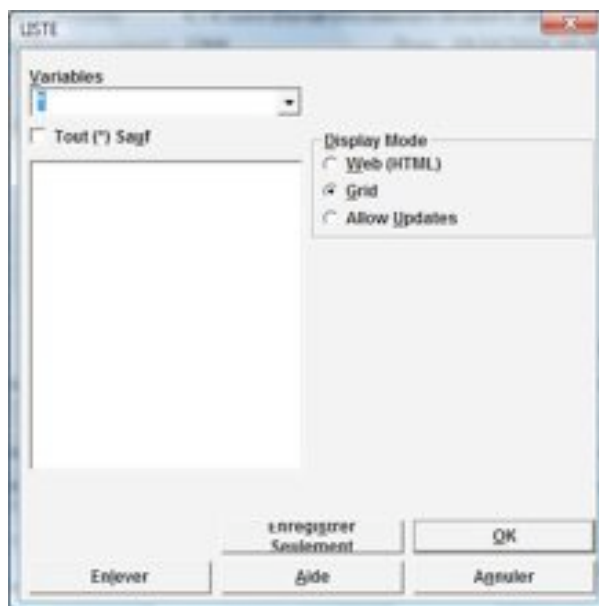
Taille : taille des enfants en cm ;

PM2_5 : concentration de microparticules de poids moléculaire 2,5 en $\mu\text{g}/\text{m}^3$ dans l'air des salles de classe (c'est une mesure de pollution de l'air).

LISTE

La première étape dans une étude, c'est d'examiner les données visuellement, pour avoir une vue d'ensemble et décider qu'elles sont les analyses appropriées à effectuer. Un tableau avec les observations en ligne et les variables en colonne est utile pour cela. La commande **Liste** permet d'obtenir un tel tableau à partir des enregistrements du fichier de données. La commande **Liste** suivie d'un ou plusieurs noms de variable n'affiche le tableau que de ces variables.

- ☞ Cliquez sur la commande **Liste** depuis la fenêtre "Analyse", dans le groupe des commandes ■ Statistiques;
- ☞ dans la boîte de dialogue LISTE, l'option **Grid** est sélectionnée dans le champ "Display Mode";
- ☞ cliquez sur **OK**.



⇒ Le résultat apparaît dans une nouvelle fenêtre:

identificati	dateentrev	age	sexe	affections	crisesulf	ecroma	poids	taille	PM2.5
100001	26/10/1999	10	M	Missing	Missing	No	41	1,54	23
100002	21/06/1999	9	M	Missing	Missing	No	37	1,43	31
100003	26/10/1999	10	M	Missing	Missing	No	40	1,51	23
100004	21/06/1999	10	M	Missing	Missing	No	50	1,43	31
100005	26/10/1999	10	F	Missing	Missing	No	48	1,5	23
100006	26/10/1999	11	M	Missing	Missing	No	34	1,51	23
100007	26/10/1999	11	F	Missing	Missing	No	34	1,55	23
100042	11/06/1999	10	M	Missing	Missing	No	45	1,44	13,9
100044	11/06/1999	12	M	Missing	Missing	No	34	1,43	13,9
100046	11/06/1999	10	F	Missing	Missing	No	30	1,45	13,9
100066	11/06/1999	10	M	Missing	Missing	Missing	30	1,39	11,2
100068	11/06/1999	10	F	Missing	Missing	No	30	1,38	11,2
100073	22/06/1999	9	F	Missing	Missing	No	37	1,43	23,5
100075	22/06/1999	9	M	Missing	Missing	No	40	1,46	23,5
100077	11/06/1999	10	F	Missing	Missing	No	50	1,49	11,2
100079	11/06/1999	10	M	Missing	Missing	No	37	1,4	11,2
100080	11/06/1999	10	F	Missing	Missing	No	46	1,49	11,2
100105	25/06/1999	10	M	Missing	Missing	Yes	26	1,25	13,9
100107	25/06/1999	9	M	Missing	Missing	Yes	44	1,34	13,9
100108	25/06/1999	10	F	Missing	Missing	No	32	1,49	13,9
100109	25/06/1999	10	F	Missing	Missing	No	47	1,53	13,9
100111	25/06/1999	9	F	Missing	Missing	No	33	1,42	13,9
100167	04/06/1999	11	F	Missing	Missing	Missing	Missing	Missing	14,6
100333	21/06/1999	11	F	Missing	Missing	No	45	1,4	20
100513	01/06/1999	10	F	No	Missing	No	33,5	1,37	12
100518	04/06/1999	11	M	No	Missing	No	37	1,48	14,6
100519	04/06/1999	10	F	No	Missing	No	39	1,45	14,6
100520	04/06/1999	11	F	No	Missing	Yes	48	1,54	14,6
100521	04/06/1999	11	M	Missing	Missing	No	33	1,43	14,6
100537	04/06/1999	10	F	Missing	Missing	Missing	34,6	1,51	10,9
100541	04/06/1999	11	M	No	Missing	No	42,6	1,54	10,9
100542	04/06/1999	9	M	No	Missing	No	34,6	1,42	10,9
100543	04/06/1999	10	M	No	Missing	No	26,4	1,4	10,9
100544	04/06/1999	10	F	No	Missing	No	40	1,38	10,9
100546	12/10/1999	8	M	Missing	Missing	No	27	1,29	15,6
100554	01/06/1999	9	F	No	Missing	No	33	1,39	12
100555	01/06/1999	9	M	No	Missing	No	39	1,46	12
100556	01/06/1999	9	F	No	Missing	No	27	1,4	12

FREQUENCES

Pour une variable en classes, la commande **Fréquences** va compter le nombre de réponse pour chaque catégorie de la variable, et les fréquences absolues et relatives (en pourcentage), ainsi que l'intervalle de confiance à 95 % (voir en annexe page 71).

- ☞ Cliquez sur la commande **Fréquences** depuis la fenêtre "Analyse", dans le groupe des commandes ■ Statistiques;
- ☞ dans la boîte de dialogue **FREQ**, sélectionnez la variable "Sifflements" à partir du menu déroulant appelé "Fréquence de";
- ☞ cliquez sur **OK**.



⇒ Les fréquences de la variable "Sifflements" apparaissent alors dans l'écran de sortie des résultats:

Ensemble de données Précédent Ensemble de données suivant Bibliothèque de Résultats

FREQ sifflements

Procédure Suivante

State

sifflements	Fréquence	Pourcent	Pourcent. Cum.	
Yes	112	9,8%	9,8%	
No	1032	90,2%	100,0%	
Total	1144	100,0%	100,0%	

Intervalle de confiance à 95 %

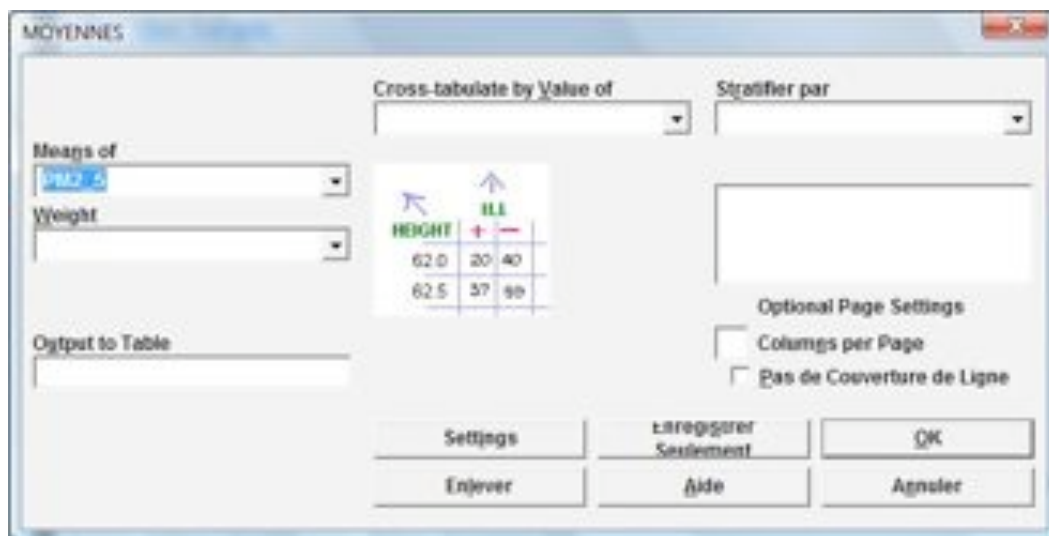
Yes 8,2% 11,7%

No 88,3% 91,8%

MOYENNES avec une variable

La commande **Moyennes** donne une description des variables numériques continues, avec les valeurs minimale et maximale, la médiane, les quartiles et le mode, la moyenne et la variance (voir les définitions en annexe page 67).

- ☞ Cliquez sur la commande **Moyennes** depuis la fenêtre "Analyse", dans le groupe des commandes ■ Statistiques;
- ☞ dans la boîte de dialogue MOYENNES, sélectionnez la variable PM2.5 (la concentration en micro particules) de notre exemple de données, depuis le menu déroulant appelé "Means of" ("Moyenne de");
- ☞ cliquez sur **OK**.



⇒ Les statistiques apparaissent dans la fenêtre de sortie des résultats:

Value	Count	Percentage	Cumulative Percentage
29,3	5	0,4%	96,0%
29,4	5	0,4%	96,3%
29,6	5	0,4%	96,7%
29,9	10	0,7%	97,4%
31	5	0,4%	97,8%
31,2	5	0,4%	98,2%
31,4	5	0,4%	98,5%
33	10	0,7%	99,3%
35	5	0,4%	99,6%
70,3	5	0,4%	100,0%
Total	1366	100,0%	100,0%

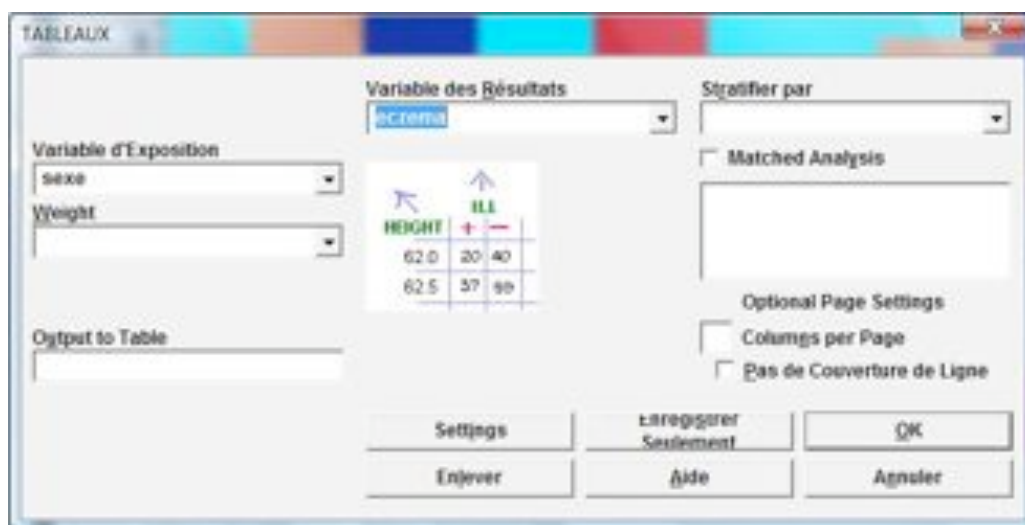
Obs	Total	Moyenne	Variance	Ecart type	
1366	22699,2000	16,6173	49,2394	7,0171	
Minimum	25%	Mé-diane	75%	Maximum	Mode
5,4000	11,5000	14,9000	19,8000	70,3000	10,2000

TABLEAUX

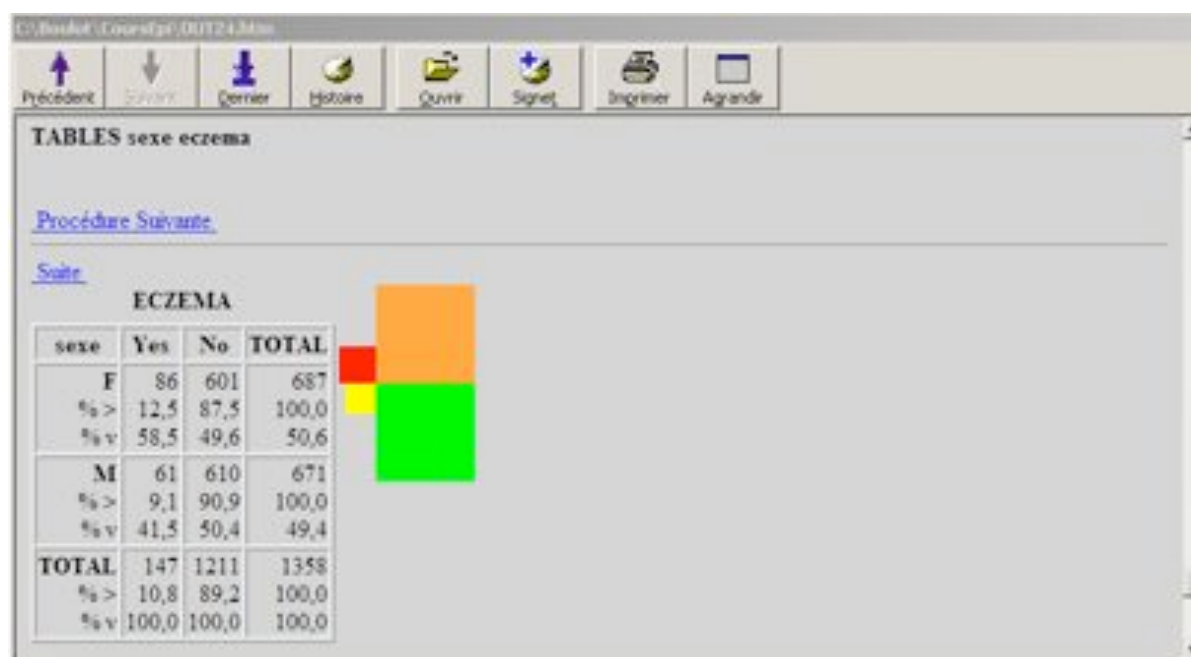
La commande **Tableaux** donne des tableaux croisés avec les statistiques pour tester l'association entre deux variables qualitatives (voir en annexe pages 68 à 70). Elle compte le nombre d'observations qui ont la même valeur pour ces deux variables.

Dans notre exemple de données, pour étudier l'association entre l'eczéma et le sexe:

- ☞ Cliquez sur la commande **Tableaux**, dans le groupe des commandes **■ Statistiques**;
- ☞ dans la boîte de dialogue TABLEAUX, sélectionnez la variable "sexe" depuis le menu déroulant appelé "**Variable d'Exposition**", et la variable "eczema" depuis le menu déroulant appelé "**Variable des Résultats**";
- ☞ Cliquez sur **OK**.



⇒ Les résultats apparaissent dans la fenêtre de sortie:



TABLES sexe eczema

[Procédure Suivante](#)

[Suite](#)

ECZEMA			
sexe	Yes	No	TOTAL
F	86	601	687
% >	12,5	87,5	100,0
% v	58,5	49,6	50,6
M	61	610	671
% >	9,1	90,9	100,0
% v	41,5	50,4	49,4
TOTAL	147	1211	1358
% >	10,8	89,2	100,0
% v	100,0	100,0	100,0

Table d'Analyse Unique			
	Point	95% Intervalle de Confiance	
	Estimer	Inferieur	Supérieur
PARAMETRES: Basé sur Odds			
Odds Ratio (cross product)	1,4309	1,0115	2,0243 (T)
Odds Ratio (MLE)	1,4306	1,0121	2,0305 (O)
		0,9980	2,0606 (F)
PARAMETRES: Basé sur Risk			
Risk Ratio (RR)	1,3770	1,0096	1,8781 (T)
Risk Difference (RD%)	3,4273	0,1326	6,7220 (T)
(T=Taylor series; C=Cornfield; M=Mid-P; F=Fisher Exact)			
TESTS STATISTIQUES			
	Chi-quarré	1-tailed p	2-tailed p
Chi-quarré - non corrigé	4,1306		0,0421143314
Chi-quarré - Mantel-Haenszel	4,1276		0,0421900966
Chi-quarré - corrigé (Yates)	3,7832		0,0517697792
Mid-p exact		0,0212617287	
Fisher exact		0,0256913870	

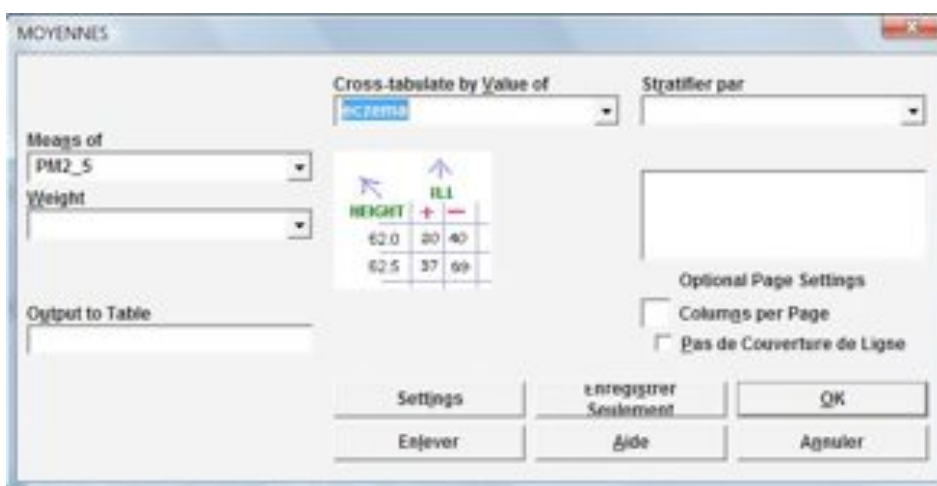
Remarquez que dans la fenêtre d'édition des commandes, TABLES est suivi d'abord par la variable sexe (le facteur "d'exposition"), puis par la variable eczéma (la "maladie"). Ainsi le risque relatif calculé représente le risque d'avoir de l'eczéma chez les filles par rapport aux garçons. Comme la valeur n'est pas incluse dans l'intervalle de confiance à 95 % du risque relatif (RR=1,38; IC 1,01 – 1,88) et de l'odds ratio (OR=1,43; IC 1,01 – 2,02), l'association entre l'eczéma et le sexe est statistiquement significative. De même, les valeurs p du Chi 2 sont inférieures à 0,05. Donc on peut dire que dans cette étude, les filles ont une probabilité plus élevée que les garçons d'avoir de l'eczéma.

MOYENNES avec deux variables (ANOVA)

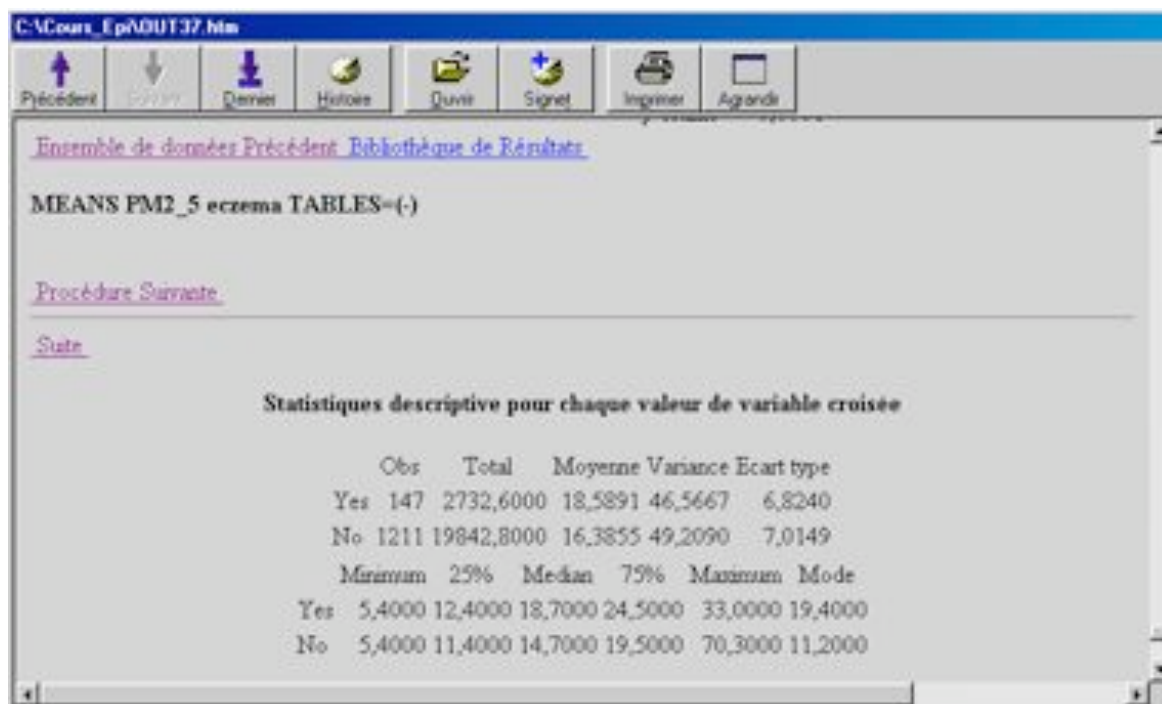
La commande TABLEAUX s'utilise avec les variables qualitatives, avec par exemple des valeurs binaires comme oui/ non (comme la variable ECZEMA) ou Masculin et Féminin (comme la variable SEXE). Les variables quantitatives, comme AGE et PM2_5, nécessitent un autre type d'analyse. La commande **Moyennes** permet de comparer les valeurs d'une variable quantitative entre les groupes définis par les catégories d'une variable qualitative en effectuant une analyse de la variance (ANOVA).

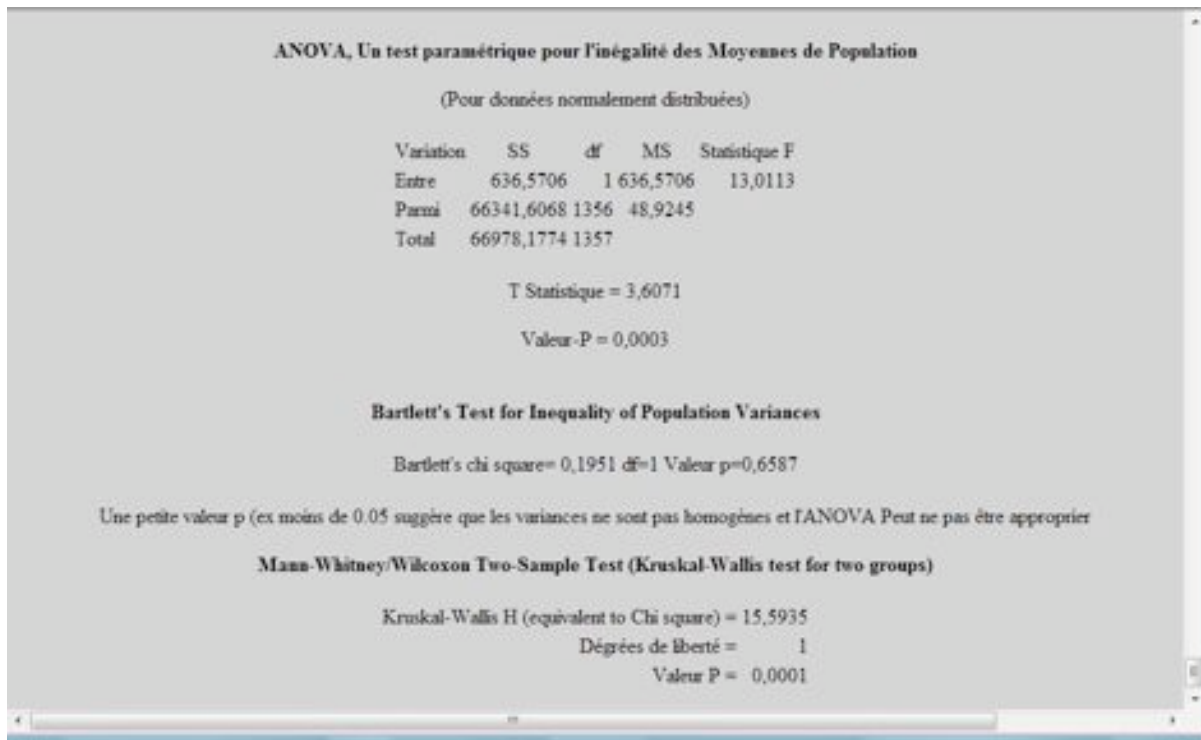
Pour comparer les valeurs des PM2.5 (les concentrations en micro particules dans l'air) entre les enfants qui ont de l'eczéma et ceux qui n'en ont pas:

- ☞ Cliquez sur la commande **Moyennes**, dans le groupe des commandes **Statistiques**;
- ☞ dans la boîte de dialogue MOYENNE, sélectionnez la variable "PM2_5" depuis le menu déroulant appelé "**Moyenne de**", et sélectionnez la variable "eczema" depuis le menu déroulant "**Tableaux croisés par Valeur de**";
- ☞ cliquez sur **OK**.



⇒ Les moyennes apparaissent dans la fenêtre de sortie, avec d'autres statistiques:





La moyenne des concentrations dans l'air des salles de classe des microparticules PM2.5 est de 18,6 $\mu\text{g}/\text{m}^3$ pour les enfants avec de l'eczéma, et de 16,4 $\mu\text{g}/\text{m}^3$ pour les enfants sans eczéma. Les valeurs p pour les tests ANOVA (test paramétrique) et de Kruskal-Wallis (test non paramétrique) sont très inférieures à 0,05, donc on peut conclure que les différences d'exposition aux PM2.5 sont significatives entre les enfants avec de l'eczéma et sans. Les enfants qui sont plus exposés aux PM2.5 ont une probabilité plus élevée d'avoir de l'eczéma.

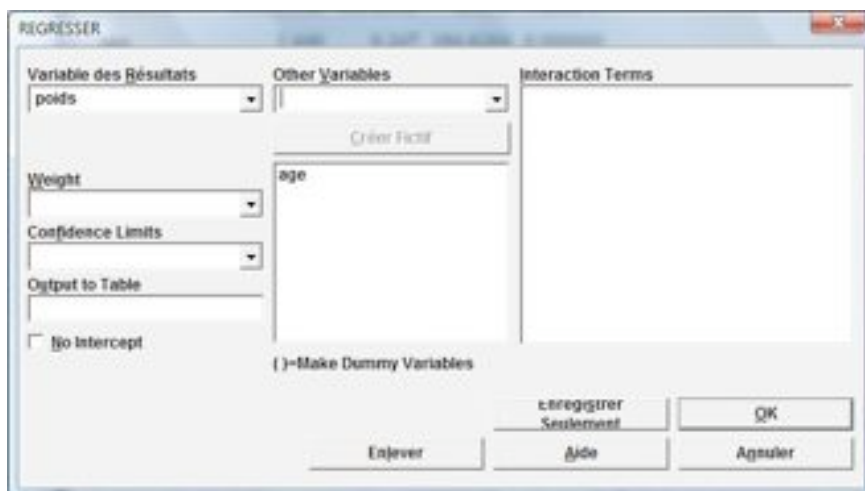
Le test de Bartlett nous renseigne sur lequel des deux tests de comparaison utiliser. Si les variances entre les groupes ne sont pas les mêmes (ne sont pas homogènes), il faut utiliser le test non paramétrique (le test de Kruskal-Wallis) plutôt que le test paramétrique (ANOVA). Ici les variances sont statistiquement égales.

REGRESSION LINEAIRE

La commande REGRESSION LINEAIRE permet de tester l'association entre deux ou plus de deux variables quantitatives. Elle calcule les coefficients de la régression linéaire, et le coefficient de corrélation linéaire entre les variables.

Dans notre exemple de données, pour tester l'association entre le poids et l'âge, et calculer la droite de régression du poids par l'âge (relation linéaire de la forme : $\text{poids} = a * \text{age} + b$):

- ☞ Cliquez sur la commande **Régression Linéaire**, dans le groupe des commandes ■ Statistiques Avancées;
- ☞ dans la boîte de dialogue REGRESSER, sélectionnez la variable "Poids" depuis le menu déroulant appelé "**Variable des Résultats**", et sélectionnez la variable "Age" depuis le menu déroulant appelé "**Other Variables**";
- ☞ cliquez sur OK.



⇒ Les coefficients de la régression apparaissent dans la fenêtre de sortie, avec d'autres statistiques:

REGRESS poids = age

[Procédez Suivante](#)

Régression Linéaire

Variable	Coefficient	Std Error	Test-F	Valeur-P
age	3,440	0,247	194,6204	0,000000
CONSTANT	2,105	2,472	0,7251	0,394464

Correlation Coefficient: $r^2 = 0,13$

Source	df	Somme des Carrés	Moyenne Carré	Statistique-F
Régression	1	11701,244	11701,244	194,620
Résiduels	1361	81827,996	60,123	
Total	1362	93529,240		

Le coefficient de corrélation linéaire entre le poids et l'âge est de 0,13, les coefficients a et b de la droite de régression valent respectivement 3,44 et 2,11 ($\text{poids} = 3,44 * \text{age} + 2,11$).

SELECTIONNER, Annuler SELECTION

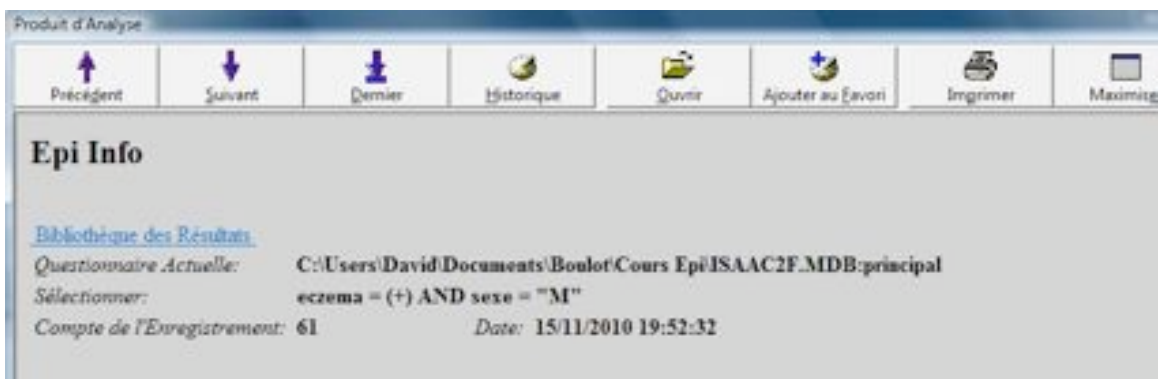
La commande **Sélectionner** vous permet de sélectionner des observations particulières selon vos critères, pour les analyser. La commande **Annuler Sélection** supprime l'effet de la sélection.

Avec nos données, pour sélectionner seulement les garçons avec de l'eczéma:

- ☞ cliquez sur la commande **Select**, dans le groupe des commandes ■ Sélectionner/Si;
- ☞ dans la boîte de dialogue SELECT, dans le champ de saisie appelé "**Critères de sélection**", écrivez à l'aide des boutons ou au clavier: **eczema=(+) AND sexe="M"** ;
- ☞ Cliquez sur **OK**.

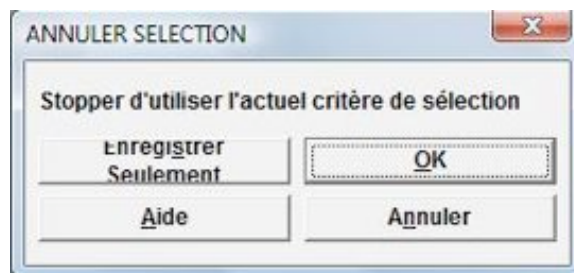


⇒ Dans la table de données, 61 enfants sont sélectionnés.



Pour annuler la sélection:

- ☞ Cliquez sur la commande **Annuler Sélection**;
- ☞ Cliquez sur **OK** dans la boîte de dialogue qui apparaît.



DEFINIR, Annuler DEFINIR

Le commande **Définir** permet de créer de nouvelles variables pour recoder ou combiner des variables existantes. La commande **Annuler Définir** supprime les nouvelles variables.

Pour créer, dans notre table de données, une nouvelle variable appelée "cIPM2_5" :

☞ cliquez sur la commande **Définir** dans le groupe des commandes ■ Variables;



☞ dans la boîte de dialogue DEFINIR, qui apparaît, dans le champ de saisie appelé "Variable Name", écrivez le nom de la nouvelle variable: **cIPM2_5**;

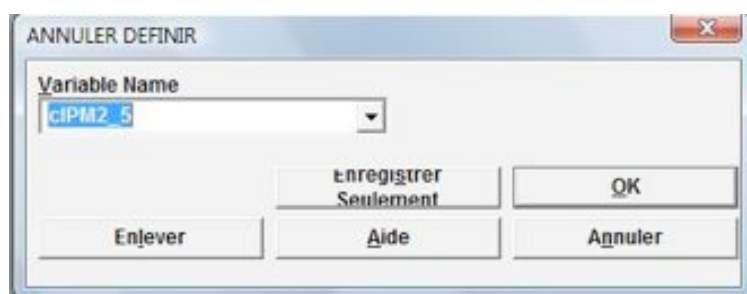
☞ dans le menu déroulant appelé "Variable Type" sélectionnez le type de la variable créée, ici **Numérique**.

☞ Cliquez sur **OK**.

Pour supprimer la nouvelle variable "cIPM2_5" :

☞ cliquez sur la commande **Annuler définir** dans le groupe des commandes ■ Variables;

☞ dans la boîte de dialogue ANNULER DEFINIR, qui apparaît, dans le champ de saisie appelé "Nom de la variable", écrivez le nom de la variable à supprimer: **cIPM2_5**;



☞ Cliquez sur **OK**.

RECODER

Le commande **Recoder** permet de recoder une variable existante et de mettre le résultat dans une nouvelle variable.

Pour recoder la variable "PM2_5" dans la variable "cIPM2_5" :

☞ cliquez sur la commande **Recoder** dans le groupe des commandes ■ Variables;

RECODER

From: [] To: []

Dates must be in US format

Valeur (vide = autre)	Valeur (si c'est le c	Valeur Recodée

Supprimer les intervalles Enregistrer Seulement OK

Enlever Aide Annuler

☞ dans la boîte de dialogue RECODER, en haut dans le champ de saisie "From" (De), écrivez le nom de la variable à recoder **PM2_5** et dans le champ "To" (à), mettez le nom de la nouvelle variable **cIPM2_5** ;

RECODER

From: PM2_5 To: cIPM2_5

Dates must be in US format

Valeur (vide = autre)	Valeur (si c'est le c	Valeur Recodée
LOVALUE	15	1
15 HIVALUE		2

Supprimer les intervalles Enregistrer Seulement OK

Enlever Aide Annuler

- ☞ dans les champ de saisie en dessous mettez l'intervalle des valeurs qui vont prendre un nouveau code, la valeur inférieure dans la colonne "**Valeur**", la valeur supérieure dans la colonne "**A la valeur**", puis le nouveau code dans la colonne "**Valeur Recodée**", en appuyant sur la touche Entrée à chaque fois.
- ☞ Pour finir, cliquez sur le bouton **OK**.

Vous avez défini ainsi une nouvelle variable en deux classes, qui vaudra 1 pour les valeurs basses de PM2_5 (entre la plus petite valeur, mot clé LOVALUE et 15), et 2 pour les valeurs élevées (entre 15 et la plus grande valeur, mot clé HIVALUE).

FREQ cIPM2_5

[Procédure Suivante](#)

[En avant](#)

cIPM2_5	Fréquence	Pourcentage	Pourcentage Cum
1	698	51,1%	51,1%
2	668	48,9%	100,0%
Total	1366	100,0%	100,0%

95% Conf Limits

1 48,4% 53,8%

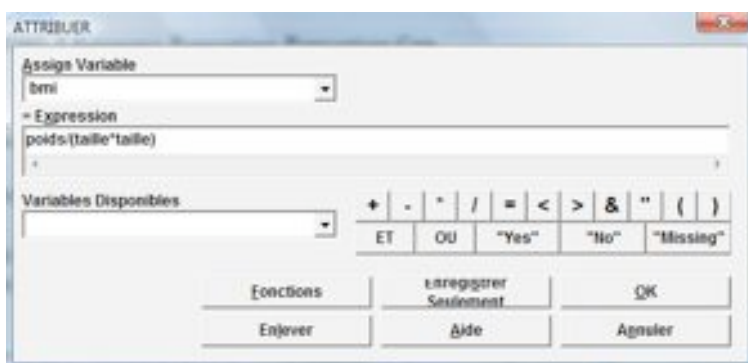
2 46,2% 51,6%

ATTRIBUER

Le commande **Attribuer** permet d'assigner à une nouvelle variable le résultat d'un calcul arithmétique à partir d'une ou de plusieurs variables existantes.

Pour créer une variable correspondant à l'indice de masse corporelle (bmi en anglais) des enfants, qui correspond au poids en kg divisé par le carré de la taille en cm, après avoir défini, avec la commande DEFINIR (voir page 43), une variable appelée "bmi" :

☞ cliquez sur la commande **Attribuer** dans le groupe des commandes ■ Variables;



☞ dans le champ "Assign Variable", sélectionnez depuis la liste déroulante la variable **bmi**, et dans le champ "Expression" tapez **poids/(taille*taille)**, puis cliquez sur le bouton **OK**;

Vous avez défini ainsi une nouvelle variable qui contiendra le résultat du calcul de l'indice de masse corporelle à partir du poids et de la taille des enfants, dont on peut déterminer la moyenne.

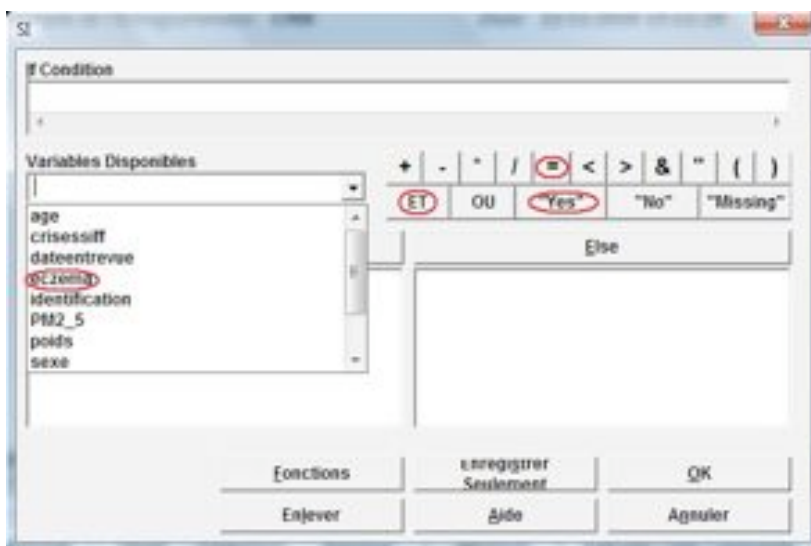
Obs	Total	Moyenne	Variance	Std Dev	Minimum	25%	Médiane	75%	Maximum	Mode
1363	24490,4772	17,9681	8,4465	2,9063	12,2449	15,8932	17,3469	19,4637	30,2618	15,3061

SI, ATTRIBUER

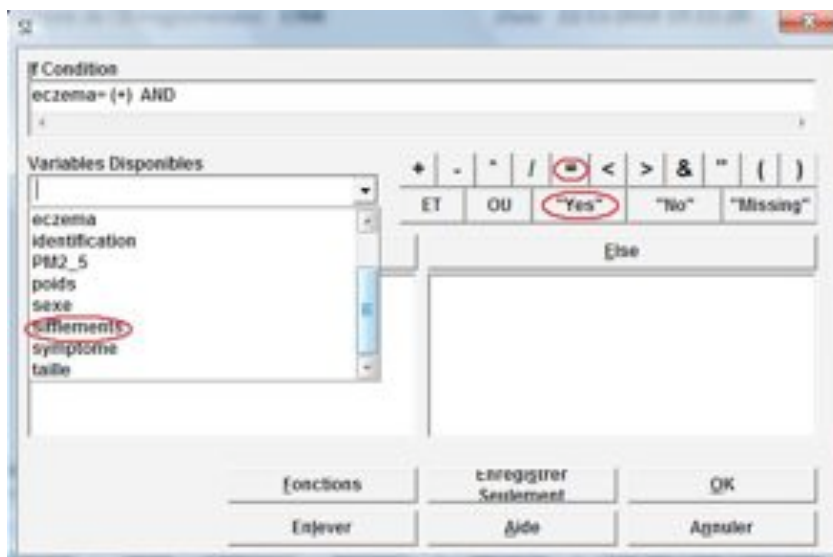
Le commande **Si**, combinée à la commande **Attribuer**, permet de coder une nouvelle variable suivant le contenu d'une ou de plusieurs variables existantes.

Avec notre table de données, pour créer une variable indiquant si un sujet cumule les symptômes de sifflements et d'eczéma, après avoir défini avec la commande DEFINIR (voir page 43) une variable appelée "symptome" :

☞ cliquez sur la commande **Si** dans le groupe des commandes ■ Sélectionner/Si;

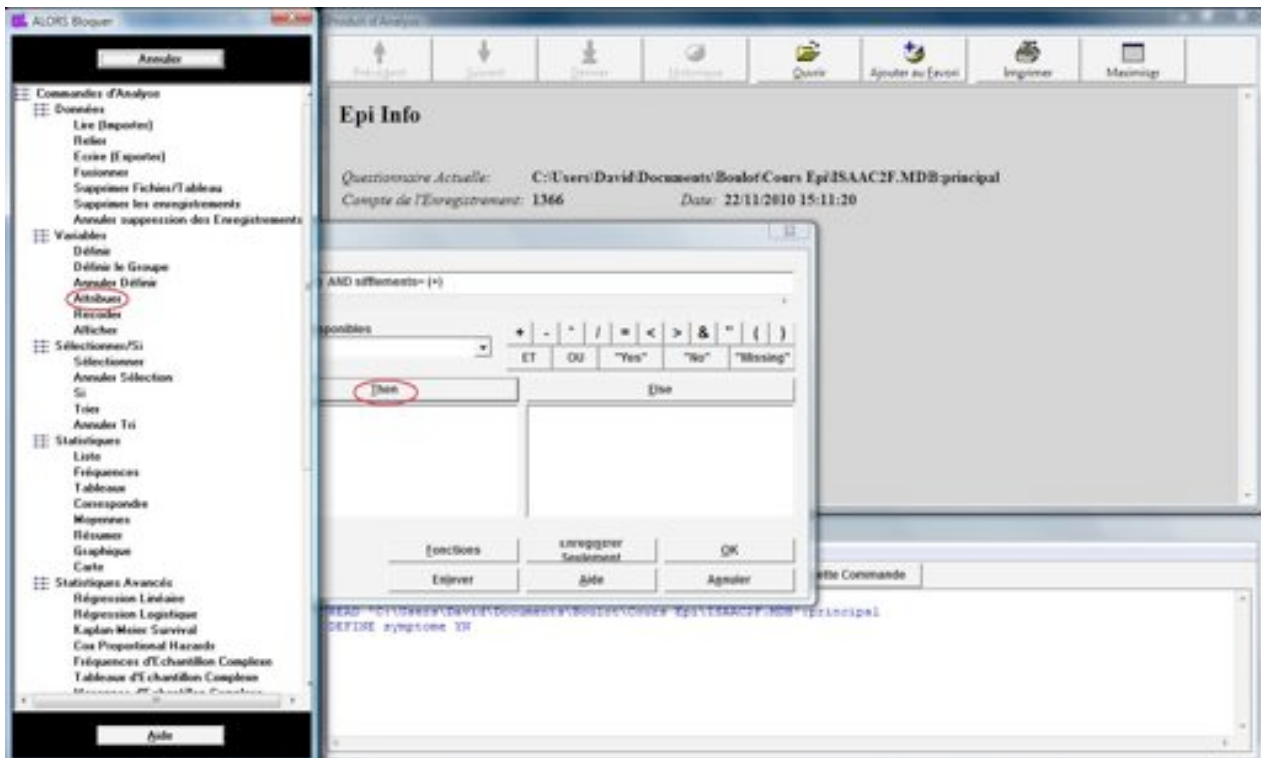


☞ Dans la boîte de dialogue SI, dans le champ "If Condition", sélectionnez depuis la liste déroulante en dessous "Variables disponibles", la variable **EczeMa** , et à partir du clavier virtuel de droite, cliquez sur les boutons =, "Yes", **ET**.



☞ Dans le champ "If Condition", sélectionnez depuis la liste déroulante "Variables disponibles", la variable **Sifflements**, et à partir du clavier virtuel de droite, cliquez sur les boutons =, "Yes".

☞ Cliquez sur le bouton **Then**, apparaît alors une liste de commandes sur la gauche;

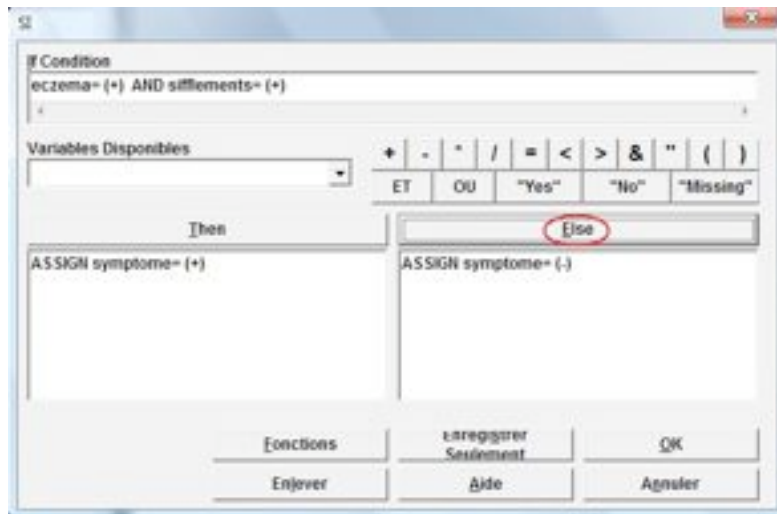


☞ choisissez la commande **Attribuer** ;



☞ dans le champ "Assign Variable", sélectionnez depuis la liste déroulante la variable **Symptome**, et à partir du clavier virtuel de droite, cliquez sur le bouton "Yes", puis cliquez sur le bouton **Add** (ajouter en anglais);

☞ dans la boîte de dialogue SI qui réapparaît, cliquez sur le bouton **Else** (autrement en anglais), puis dans la liste des commandes qui ressurgit, choisissez la commande **Attribuer**, et comme précédemment attribuer la valeur "No" à la variable **symptome** ;



☞ Pour finir, cliquez sur le bouton **OK** dans la boîte de dialogue SI.

Vous avez défini ainsi une nouvelle variable "symptome" qui vaudra "Yes" (oui) pour les sujets qui auront déclaré de l'eczéma et des sifflements, ou "No" (non) autrement.

symptome	Fréquence	Pourcentage	Pourcentage Cum
Yes	27	2,0%	2,0%
No	1339	98,0%	100,0%
Total	1366	100,0%	100,0%

95% Conf Limits
 Yes 1,3% 2,9%
 No 97,1% 98,7%