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 <u>A</u>nalyse Des Données dans la fenêtre d'accueil d'Epi Info.

Andys Description Data Description Descreseription D

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.

A 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 <u>P</u>roject et sélectionnez le projet (ici "Isaac2f.mdb"), dont le nom doit alors apparaître dans le champ appelé <u>Source des Données;</u>
- 🕼 puis choisissez le format **Epi 2000** depuis le menu déroulant appelé "<u>F</u>ormats des données";
- Sélectionnez "**Tout**" dans le champ appelé "Affic<u>h</u>er";
- 🕼 choisissez une vue dans la liste en dessous (ici "principal"), et cliquez sur **OK**.

yet Actuel	Baulati Canas Califord	C 25 MIDD
Users David Document	S Boulon Cours Epilisaa	C24-M08
rmats des données		
pi 2000		
ource des Données	and and a starting	and the second second
Risers/David/Document	s/Boulof/Cours Epil/SAA	C2F.MDB
Afficher		
Questionnaires	(# Tout	
rincipal		
Changer le Project	Enropistrer	ОК
Changer le <u>Project</u>	Enrogistrer Septement	QК

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.
- Se retour **sur la** boîte de dialogue DESTINATION, cliquez sur le bouton **OK**.

ESTINATION		
Output <u>F</u> ilename		
C:\temp\résultats.htm		
Replace any existing	g file	_
Browse	Enregistrer Seulement	Ōĸ
Enlever	Aide	Annuler

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

- Is 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.

ERMER	
Stop routing output to	designated file
Enregistrer	<u>О</u> К
Seulement	Summer Sum

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.

PRICHER		-
Information for	⊂ Vie <u>w</u> s	⊂ <u>I</u> ables
De		
All sale bling and south	an and discontinue.	1.00
-Variables present	ement ersponible	•
-Variables present	ement erspondie	-
-Variables present Oytput to Table	arment erspenitive	-
-variables present	Enrequisirer	
-variables present Oytput to Table	Enrogistrer Soulement	ОК

 \Rightarrow Résultat de la commande Afficher variables:

Précégent	Su	vant De	mier His	3 torique	Querie	3 Ajouter au Eavori	Imprimer	Maximisgr
Epi Info								
Questionnai Compte de l	re Actuelle Enregistri	e: C:/User	s David Docum	ents/Boulot/Co Date: 19/10/20	ours Epi/ISA. 10 10:59:53	AC2F.MDB:princi	pal	
ISPLAY D	BVARIA	BLES						
Variable	Tableau	Type du champ	FormatValen	Info Spéciale	Invite			
	and a strend							
age	prescipas	Nombre			age			
age crisessiff	principal	Nombre Texte			age crisessiff			
age crisessiff dateestrevue	principal principal	Nombre Texte Date			age crisessiff dateentrevue			
age crisessiff dateentrevue oczema	principal principal principal	Nombre Texte Date OuiNon			age crisessiff dateentrevue eczema			
age crisessiff dateentrevue eczema identification	principal principal principal principal	Nombre Texte Date OuiNon Nombre			age crisessiff dateentrevue eczema identification			
age crisessiff dateentrevue eczema identification PM2_5	principal principal principal principal principal	Nombre Texte Date OuiNon Nombre Nombre			age crisessiff dateentrevue eczema identification PM2_5			
age crisessiff dateentrevue eczema identification PM2_5 poids	principal principal principal principal principal principal	Nombre Teste Date OsiNon Nombre Nombre Nombre			age crisessiff dateentrevue eczema identification PM2_5 poids			
age crisessiff dateentrevue eczema identification PM2_5 poids sexe	principal principal principal principal principal principal principal	Nombre Teste Date OuiNon Nombre Nombre Nombre Teste			age crisessiff dateentrevue eczema identification PM2_5 poids sexe			
age crisessiff dateentrevue eczema identification PM2_5 poids sexe siffements	principal principal principal principal principal principal principal	Nombre Teste Date OuiNon Nombre Nombre Nombre Teste OuiNon			age crisessiff dateentrevue eczema identification PM2_5 poids sexe siffements			
age crisessiff dateentrevue eczema identification PM2_5 poids sexe siffements taile	principal principal principal principal principal principal principal principal principal	Nombre Texte Date OuiNon Nombre Nombre Texte OuiNon Nombre			age crisessiff dateentrevue eczema identification PM2_5 poids sexe siffements taile			

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 $\mu g/m^3$ 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.

Display Mo ⊂ Web (H G Grid ⊂ Allow Us	de ML) sdates
	Display Mo C Web (H) Grid C Allow (b)

 \Rightarrow Le résultat apparaît dans une nouvelle fenêtre:

identificati	dateentiev	age	10.00	sittlements	crisessiff	eczena	poids	taille	PM2_5	
00001	26/10/1999	10	M	Missing	Missing	No	41	1,54	23	
200002	21/06/1999	9	M	Masing	Missing	No	37	1,43	31	
000003	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	54	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	54	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	1,45	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	45	1,49	11.2	
100105	25/06/1999	10	M	Missing	Missing	Yes	28	1,29	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	
101333	21/06/1999	11	F	Missing	Missing	No	45	1.4	20	
101513	01/06/1999	10	F	No	Missing	No	33.5	1,37	12	
101518	04/06/1999	11	M	No	Missing:	No	37	1,48	14.6	
101519	04/06/1999	10	F	No	Missing	No	39	1.45	14.6	
101520	04/06/1999	11	F	No	Missing	Yes	48	1.54	14,6	
101521	04/06/1999	11	M	Missing	Missing	No	33	1,43	14.6	
101537	04/06/1999	10	F	Missing	Missing	Masing	34.6	1.51	10.9	
101541	04/06/1999	11	M	No	Missing	No	42.6	1,54	10.9	
101542	04/06/1999	9	M	No	Missing	No	34.6	1.42	10,9	
101543	04/06/1999	10	M	No	Missing	No	35,4	1.4	10.9	
101544	04/06/1999	10	£	No	Missing	No	40	1.38	10.9	
101546	12/10/1999	8	M	Missing	Missing	No	27	1,29	15.6	
101554	01/06/1999	9	F	No	Missing	No	33	1.29	12	
101555	01/06/1999	9	M.	No	Missing	No	29	1,45	12	
101556	01/06/1999	9	F	No	Missing	No	27	1,4	12	
	sheet which are shorten.	4	1.1	10	44		-140 M	4.4	1.5	

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 **S**tatistiques;
- dans la boîte de dialogue FREQ, sélectionnez la variable "Sifflements" à partir du menu déroulant appelé "Fréquence de";
- 🔊 cliquez sur OK.

				Eréquence de		Stratifier pa	r
111	freq	5					
+	29	35%		Tout (") Sagt			
-	37	65%		siffements			
fotal	57	1995					
hitpu	et to Te	abie	•				
				Settings	Enreg Seule	ment	QK

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

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 **S**tatistiques;
- 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.

	Cross-tat	bulate by ⊻a	lue of	Stratifier pa	r
Means of Weight Ogtput to Table	• K HEDGHT 620 625	* HL + - 20 40 57 60		Options Column Pas de	al Page Settings Is per Page Couverture de Ligne
	s	ettings	Enrog	ment	QK
	0	nlever	Ai	de	Annuler

 \Rightarrow Les statistiques apparaissent dans la fenêtre de sortie des résultats:

ACours EpADUT37.htm								
4	*	timier History	e	Signet	Ingeimer	Agranda		
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%					
		1 Mini 5,4	Obs Total 1366 22699,20 mum 25% 1 1000 11,5000 1	Moye 00 16,61 dédane (4,9000 1	ne Varian 173 49,239 75% Mi 19,8000 7	e Ecart ty 4 7,01 ccimum 3 0,3000 10	npe 71 Mode 0,2000	

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.

			14420150	-		
		Variable des B	ésultats		Stratifier pa	r
Variable d'Exposition sexe	•			-	T Matched	Analysis
Weight	•	HEIGHT + 62.0 20 62.5 57	40		Ontion	al Danas Settings
Ogtput to Table	-				Column Pas de	is per Page Couverture de Ligne
		Setting	25	Enreg	igtrer ment	QК
		Enjeve	e l	Ai	de	Agnuler

 \Rightarrow Les résultats apparaissent dans la fenêtre de sortie:

Doubut To	ursi(pi)	001243	Maria										
+ Yécéderk	+	Que	tier Ho	3 Soire	Querr	Spret	Ingrise	Agri	ndr				
TABLES	sexe	eczem	a -						- 11	11			3
200000													
Procedur	e Surva	me.											
Suite	FC7	MA											
Sexe	Yes	No	TOTAL										
F	86	601	687										
962	12.5	87.5	100.0										
	58,5	49,6	50,6										
М	61	610	671										
56>	9,1	90,9	100,0										
56 V	41,5	50,4	49,4										
TOTAL	147	1211	1358										
16>	10,8	89,2	100,0										- 194 194
	100,0	100,0	100,0										

Table	d'Analyse U	nique			<u>^</u>
	Point	95% Intervalle	de Confia	nce	
	Estimer	Inferieur	Supe	rieur	
PARAMETRES: Basé sur Odd	5				
Odds Ratio (cross product)	1,4309	1,0115	2,0243	(T)	
Odds Ratio (MLE)	1,4306	1,0121	2,0305	(H)	
		0,9980	2,0606	(2)	
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=Cor	nfield; M=	Mid-P: F=Fi	ther Exac	:=)	
TESTS STATISTIQUES	Chi-quaré	1-tailed p	2-tailed p		
Chi-quarré - non corrigé	4,1306		0,042114	3314	
Chi-quarré - Mantel-Haenszel	4,1276		0,042190	0966	
Chi-quarré - corrigé (Yates)	3,7832		0,051769	7792	
Mid-p exact		0,0212617287			
Fisher exact	3	0,0256913870			
					2
					2

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 un 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.

		Cross-tat	oulate	by Val	lue of	Stratifier pa	r
Meags of		1 martenue			-	a	-
PM2_5	*		1	1		-	
Keight	_	HEKAT	11	<u> </u>			
	•	62.0	20	40			
		62.5	37	69		Ontiona	Pane Settines
Dytput to Table	-					Colume Pas de	is per Page Couverture de Ligne
		5	etting	15	Enreg Seale	pigurer	ŌK
		E	njeve	r.	A	ide	Agnuler

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

C:\Court_Ep/\001	37.htm				
	Dernier Histoire	Quvit Signet	ingrimer Agro] rdir	
Ensemble de do MEANS PM2	nnées Précédent Biblio 5 eczema TABLES=(thèque de Rénátati (•)	-		-
Procédure Suiv	ante				
	Statistiques d	lescriptive pour ch	aque valeur de v	uriahle croisée	
	0 Yes 1 No 12	bs Total Me 47 2732,6000 18 11 19842,8000 16	yenne Variance Ec 1,5891 46,5667 1,3855 49,2090	art type 6,8240 7,0149	
	Minim Yes 5,40 No 5,40	um 25% Media 00 12,4000 18,700 00 11,4000 14,700	n 75% Maxim 0.24,5000 33,00 0.19,5000 70,30	um Mode 00 19,4000 00 11,2000	-
4					

	ANOVA, Un test paramétrique pour l'inégalité des Moyennes de Population	ŕ
	(Pour données normalement distribuées)	
	Variation SS df MS Statistique F Entre 636,5706 1 636,5706 13,0113 Parmi 66341,6068 1356 48,9245 Total 66978,1774 1357	
	T Statistique = 3,6071	
	Valeur-P = 0,0003	
	Bartlett's Test for Inequality of Population Variances	
	Bartlett's chi square= 0,1951 df=1 Valeur p=0,6587	
	Une petite valeur p (ex moins de 0.05 suggère que les variances ne sont pas homogènes et l'ANOVA Peut ne pas être approprier	
	Mann-Whitney/Wilcoxon Two-Sample Test (Kruskal-Wallis test for two groups)	
	Kruskal-Wallis H (equivalent to Chi square) = 15,5935	
	Dégrées de liberté = 1	
	Valeur P = 0,0001	5
•		

La moyenne des concentrations dans l'air des salles de classe des microparticules PM2.5 est de 18,6 μ g/m³ pour les enfants avec de l'eczéma, et de 16,4 μ g/m³ 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 : poids= $a^*age +b$):

- Cliquez sur la commande **Régression Linéaire**, dans le groupe des commandes **E** 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.

a de la d	1 848 8 247 194	A Real Property lies	
Variable des <u>B</u> ésultats	Other Variables	Interaction Terms	
poids	•	-	
	Créor Fictif		
Weight	age		
Confidence Limits	-		
Oytput to Table	-		
No Intercept			
	()-Make Dummy Variable	5	
		E.tiregistrer Seulement	QК
	Enjever	Aide	Agnuler

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

REGRESS po	ids = :	age					
Procédare Sai	ante						
Régression Li	inéaire						
Variable	Coe	fficient St	td Error	Test-F	Valeur-	P	
age		(3,440)	0,247	194,6204	0,00000	0	
CONSTANT	ŝ	2,105	2,472	0,7251	0,39446	4	
Correlation (Coeffs	cient: r^2=	0,13				
Source	df	Somme d	es Carré	s Moyem	ne Carré	Statistique-l	F
Régression	1	1	1701,24	4 11	1701,244	194,62	0
Résiduels	1361	5	1827,99	6	60,123		
Total	1362	5	3529,24	0			

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 (poids=3,44*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.

Selectionner Les Criteres oczema- (+) AND sexe-"M"					
4					
Variables Disponibles	+	-1-1	1 = <	> 8	" ()
	• ET	OU	"Yes"	"No"	"Missing"
Enection		Enregi	strer	1	Diec
	1				

 \Rightarrow Dans la table de données, 61 enfants sont sélectionnés.

Précégent	Suivant	2 Demier	3 Historique	Quvir	🤣 Ajouter au Eavori	er and the second secon	Maximita
Epi Info							
Bibliothèque de	es Résultats						
Bibliothèque de Questionnaire	n Réndrata Actuelle: (C:\Users\David	Documents/Boulot	Cours EpëlS	AAC2F.MDB:princi	pal	
Bibliothèque de Questionnaire Sélectionner:	ss Résultats Actuelle:	C:\Users\David eczema = (+) A2	Documents\Boulot ND sexe = "M"	Cours Epi IS	AAC2F.MDB:princi	pal	

Pour annuler la sélection:

Cliquez sur la commande Annuler Sélection;

🖙 Cliquez sur **OK** dans la boîte de dialogue qui apparaît.

Stopper d'utiliser l'actu	el critère de sélection
nopper a admoer racta	er entere de selection
Enregistrer	OK
Jenieniem	
Aide	Annuler

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 "clPM2_5" :

🕼 cliquez sur la commande **Définir** dans le groupe des commandes 🖿 Variables;

CHINEK		-
Variable Name		
clPM2_5		
Etendue		
Stangard		○ Permanent
Optional Settings		
variable Type	-	
Annacodic card		
Prompt		
4		
	Facediğinet	ок
	seumbert	
Enjever	Aide	Ageuler

- dans la boîte de dialogue DEFINIR, qui apparaît, dans le champ de saisie appelé "Variable Name", écrivez le nom de la nouvelle variable: clPM2_5;
- I dans le menu déroulant appelé "Variable Type" sélectionnez le type de la variable créée, ici Numérique.
- Cliquez sur OK.

Pour supprimer la nouvelle variable "clPM2_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: clPM2_5;

ariable Name		
cIPM2_5	•	
	Enregistrer Seulement	Ōĸ
Enlever	Aide	Annuler

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 "clPM2_5" :

🔊 cliquez sur la commande **Recoder** dans le groupe des commandes 🖿 Variables;

rom	To	
Da	tes must be in US form	et.
Valeur (vide = autre)	a Valeur (si c'est le c	Valeur Recodée
	_	
Bemper Les Intervalles	Larepigtrer Souliement	<u>Ok</u>

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 clPM2_5;

rom	To	
PM2_5	• CIPM2_5	
Da	tes must be in US form	st
Valeur (vide = autre)	a Valeur (si c'est le c	Valeur Recodée
LOVALUE	15	
15	HIVALUE	
1		
Esemper Les Intercetos	Enropsorer Seulement	Ök

- 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).

Précégen	2	↓ ivant	Remier H	istorique	Quvrie	Jouter au Eavori	(B) Imprimer	Meximiser	
FREQ clP	'M2_5								
Procédure	Suivante								
En avant.									
clPM2_5	Fréquence	Pourcentage	Pourcentage Cu						
1	698	51,1%	51,1	16					
2	668	48,9%	100,0	16					
Total	1366	100,0%	100,0	16	-				
95% C Limit 1 48,4% 5 2 46,2% 5	onf 15 13,8% 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;

Assidia Aquigand						
bmi						
- Expression	- T.V.					
poids (taille*taille)						
4						
Variables Disponibles		+ -	1 1	= <	> 8	"()
	-	ET	OU	"Yes"	"No"	"Missing"
		1	Enregis	trer		ж
	Eonctions		Sections	ent		

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.



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;

# Condition			
Variables Disponibles J age crisessiff dateentrevue @Czeeni0 identification PM2_5 poids sexe	•	- * / (1) < E) OU (1995) E	, > & " () "No" "Missing" se
	Enctions	Enregistrer Seulement	QК
	Enjever	Aide	Agnuler

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.

a a			- X
Condition eczema= (+) AND			
Variables Disponibles eczenta identification Pft2_5 polds sexe Efferments symptome tallie		ET OU Yes	< > & " () "No" "Missing" Else
	Enctions	Enregigtrer Seutement	<u>Ok</u>
	Enjever	Aide	Agauler

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;

ALLOIS BOQUE !	A second s		1 1 1			1		
Aseula	1	*	*		Davrie	Aposter au Escori	linging	Maximize
Commandes d'Analyse	Epi Info Questionnary Conney de ll	r Actuelle:	C:Users/David/D	ocuments/Boulof Date: 22/11	Cours Epid5/ 2010 15:11:20	AAC2F.MDB:princ	ipal	
Suppresenties enongentrements Annules suppression des Enongistrements	Compile in the	on a Dimension	1.140	1000 1011				
Définit Définit le Genape Annaber Définit Antaber	AND siffements-	1+2						
Allicher	ponibles		- + / = +	(> & " (10			
Selectioner Annules Selection Sil			1 00 115	Dise 1985				
Annuler Tri Statistiques								
Error Fréquences Tableaux Conespondie								
Mopennes Résumer Graphique		(onclives	Linepigner					
Carle E Statistiques Avancels		Enjever	Bite	Agnuler	ette C	ommande		
Régression Linduire Réglession Logistique Kaplan Mérics Survival Con Proportional Hazards Prégnances d'Echardillon Complexe Tablecou d'Echardillon Complexe Manuel d'Echardillon Complexe	DEFISE Sympto	NUTRIAL PURCH		** ********	- Non-Course	ipal		
Fréquences d'Echantilion Complexe Tableaux d'Echantilion Complexe Montes d'Antilion Complexe Asile								

🖙 choisissez la commande Attribuer ;

Assign Variable				
Symptome				
- Expression				
(+)				
4				
No. of the second se				
Variables Disposibles	•		< > 8	-1(1)
Variables Disposibles	<u> </u>	- * / = • T OU _ Yes*	< > &	" () "Missing"
Variables Disponibles	enctions	- * / = • T OU • Yes*	< > 8 "No"	" () "Missing"

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 ;

Condition						
<pre>eczema= (+) AND simements= (+</pre>)					
Variables Disponibles	•	+ . ET	• /	"Yes"	> &	" () "Missing"
Ites				E		
ASSIGN symptome= (+)		ASS	iiGN sym	ptome+ (-)		
E	nctions	1_	Enrege	purer		<u>OK</u>
	alover	1	Aid	.	An	nuter

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.

Précédent	Suiv	ant	ternier Histo	3 srique	Quvrir	3 Ajouter au Eavori	Imgrimer	Maximisgr	
FREQ sym	ptome								*
Procédare	Suivante.								
En avant									
symptome	Fréquence	Pourcentage	Pourcentage Cum	6					
Yes	27	2,0%	2,0%						
No	1339	98,0%	100,0%						
Total	1366	100,0%	100,0%	_					
95% Coal Yes 1,3% No 97,1%	Limits 2,9% 98,7%								