

TD : Régression linéaire

Exercice 1 : Régression linéaire simple

On souhaite étudier le niveau de stress en fonction du bruit. Pour cette étude, télécharger le fichier "bruit-stress.sta" sur le cours **Méthodes statistiques pour l'analyse des données en psychologie** (téléchargeable à l'adresse <http://coursenligne.u-paris10.fr>).

On considère 30 individus soumis à différents niveaux de bruit (variable BRUIT) (mesurés en décibels) et on mesure pour chacun d'entre eux le niveau de stress (variable STRESS).

1. Décrire les variables.
2. Représenter graphiquement le nuage de points. Un modèle linéaire semble-t-il pertinent ?
3. Rajouter sur le graphique la droite de régression, l'équation de la droite, les coefficients de corrélation et de détermination.
4. Interpréter les résultats de la régression. Relever les valeurs estimées des deux coefficients. Tester séparément la nullité de chacun des coefficients.
5. Calculer les valeurs prévues et les résidus pour toutes les observations.
6. Prévoir le niveau de stress que provoquerait un niveau de bruit de 90db.
7. Vérifier les conditions :
 - (a) Vérifier la normalité des résidus.
 - (b) Vérifier l'hypothèse d'homoscédasticité.
8. Analyse des valeurs atypiques.

Exercice 2 : Régression linéaire multiple

Cet exemple est tiré de "Weight, Shape, and Body Image" de David C. Howell basé sur un article de Geller, Johnston, and Madsen, 1997. Pour plus d'information, se référer à la page http://www.uvm.edu/~dhowell/StatPages/More_Stuff/Geller.html

Dans cet article, le but des auteurs était de montrer chez les femmes que la variable SAWBS (Shape And Weight Based Self-esteem inventory) a un rôle dans la prévision d'un dérèglement du comportement de l'alimentation, rôle indépendant des variables traditionnellement mises en cause dans ce comportement telles que la dépression, l'estime de soi...

Dans cet exercice, on va utiliser ces données (télécharger le fichier "image-corporelle.sta") pour illustrer un problème de régression linéaire multiple. On mesure 10 variables sur un échantillon de 84 femmes :

- SAWBS : indice indiquant dans quelle mesure les sentiments de valeur personnelle sont fondés sur l'image de son corps (mesure d'influence et non de satisfaction).
- WtPercep : score à l'échelle d'auto-évaluation de la perception de son poids, de 1 : en surcharge pondérale extrême à 7 : très maigre.
- ShPercep : score à l'échelle d'auto-évaluation de la perception de sa ligne, de 1 : pas du tout attirante à 7 : très attirante.

- HIQ (Health Inventory Questionnaire) : mesure la présence et la gravité de certaines pratiques alimentaires perturbées, valeurs de 0 à 69.
- EDIcomp (Eating Disorders Composite Index) : indice composé de la somme des scores à 3 échelles d'évaluation de désordre alimentaire.
- RSES : score à l'échelle d'estime de soi de Rosenberg, valeurs de 10 (faible estime de soi) à 50 (grande estime de soi).
- BDI (Beck Depression Inventory) : mesure de la dépression, valeurs de 0 à 63, plus le score est élevé, plus la dépression est importante.
- BMI (Body Mass Index) : mesure de la masse corporelle basée sur le poids et la taille.
- SES : statut socio-économique.
- SocDesir : échelle de "désirabilité sociale" prenant des valeurs de 0 à 10 exprimant la tendance d'un individu à avoir une bonne réponse aux différentes sollicitations sociales.

Remarque : Certaines données confidentielles ont été simulées et de ce fait, certaines observations tombent en dehors des échelles décrites précédemment.

1. Etude descriptive des variables.

2. Etude des corrélations entre variables.

- (a) Déterminer la matrice de corrélations.
- (b) Tracer les nuages de points pour tous les couples de variables.

3. Etude du modèle de régression multiple

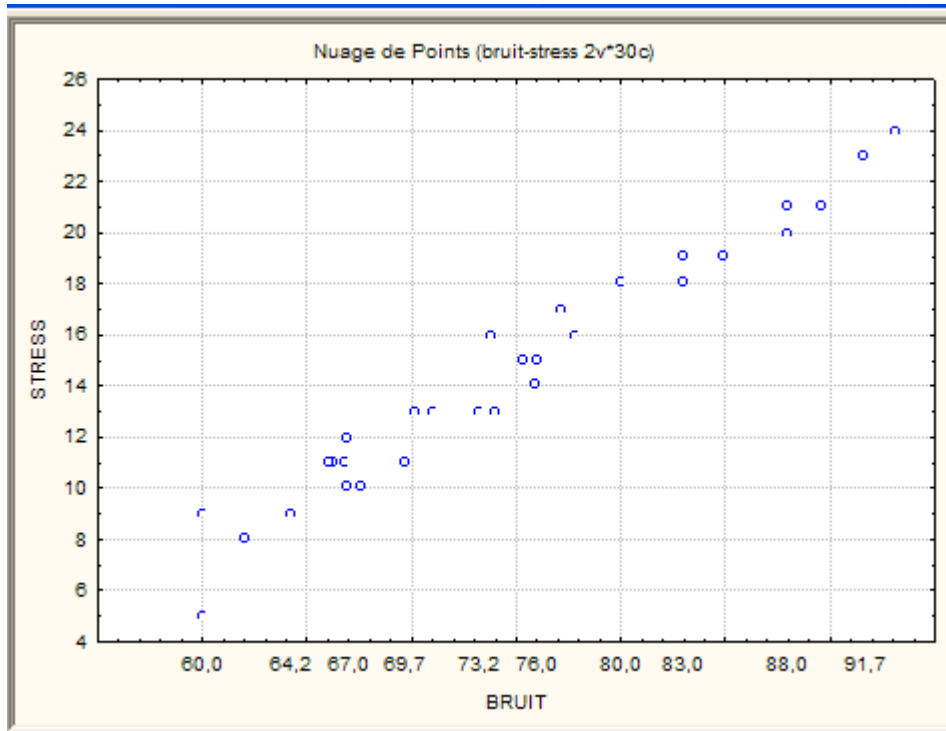
- (a) Pour répondre au problème initial, que peut-on choisir comme VI et VD ?
- (b) Dans la suite, on considère la variable dépendante : EDIcomp et 8 variables indépendantes : SAWBS, WtPercep, ShPercep, RSES, BDI, BMI, SES, SocDesir.
 - i. Donner les coefficients de l'équation de la régression linéaire multiple de EDIcomp en fonction de SAWBS, WtPercep, ShPercep, RSES, BDI, BMI, SES, SocDesir.
 - ii. Tester l'hypothèse " tous les coefficients $b_j, j = 1, \dots, 8$ sont nuls".
 - iii. Tester individuellement la nullité de chaque paramètre $b_j, j = 0, \dots, 8$. Que peut-on conclure ? En choisissant le modèle restreint (obtenu en retirant les variables X_j dont le coefficient correspondant b_j dans l'équation est non significatif), que vaut le coefficient R^2 de ce nouveau modèle ?
 - iv. Etudier la redondance des variables.
 - v. Analyse des résidus du modèle complet.
 - vi. Suppression de variables : méthode pas à pas.
Donner le modèle retenu par la méthode descendante et vérifier sa qualité.
Donner le modèle retenu par la méthode ascendante et vérifier sa qualité.
- (c) Conclure.

Exercice 1 : Régression simple

1. Spécifier la variable indépendante et la variable dépendante. Calculer les résumés standard des deux variables : minimum, maximum, moyenne,... (voir TD1).
2. Représenter graphiquement le nuage de points. Un modèle linéaire semble-t-il pertinent ?

Utiliser le menu

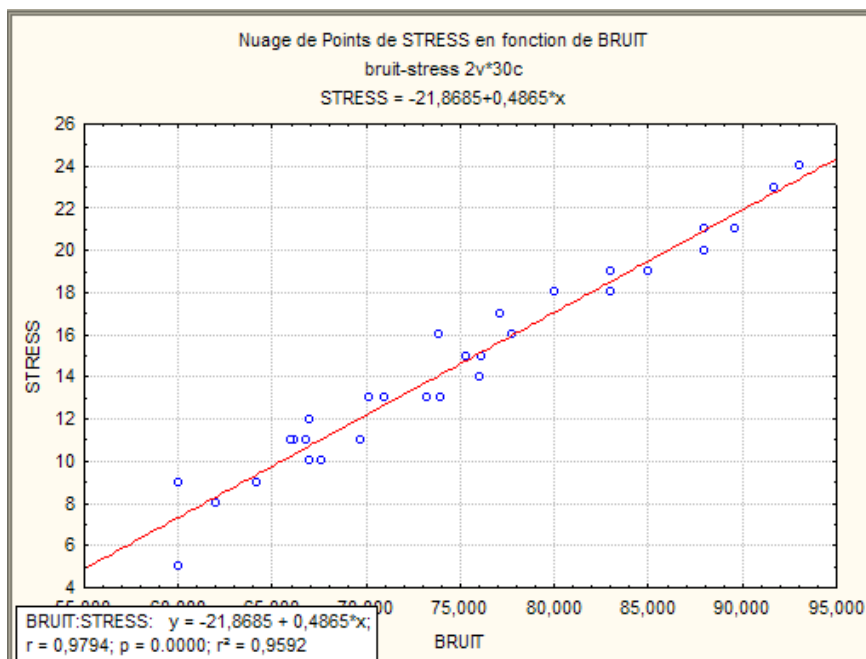
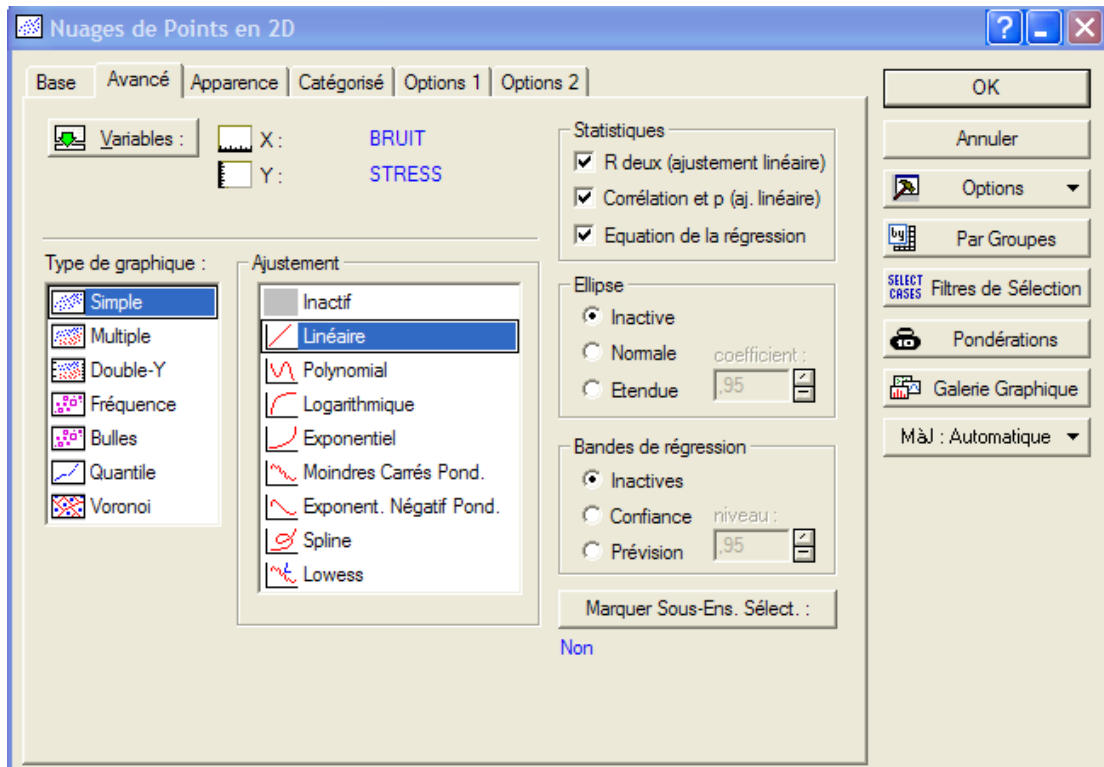
Graphiques / Graphiques en 2D / Nuage de points en 2D / cliquer sur variables / mettre pour X : BRUIT et pour Y : STRESS / Désactiver l'option type d'ajustement linéaire / cliquer sur OK



3. Rajouter sur le graphique la droite de régression, l'équation de cette droite et les coefficients de corrélation et détermination.

Utiliser le menu

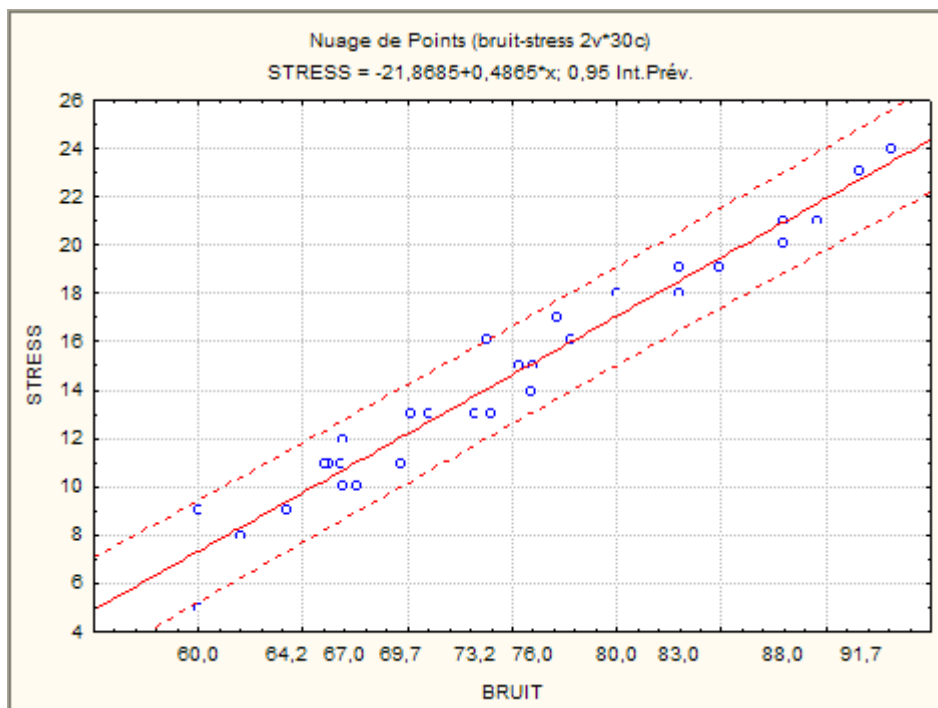
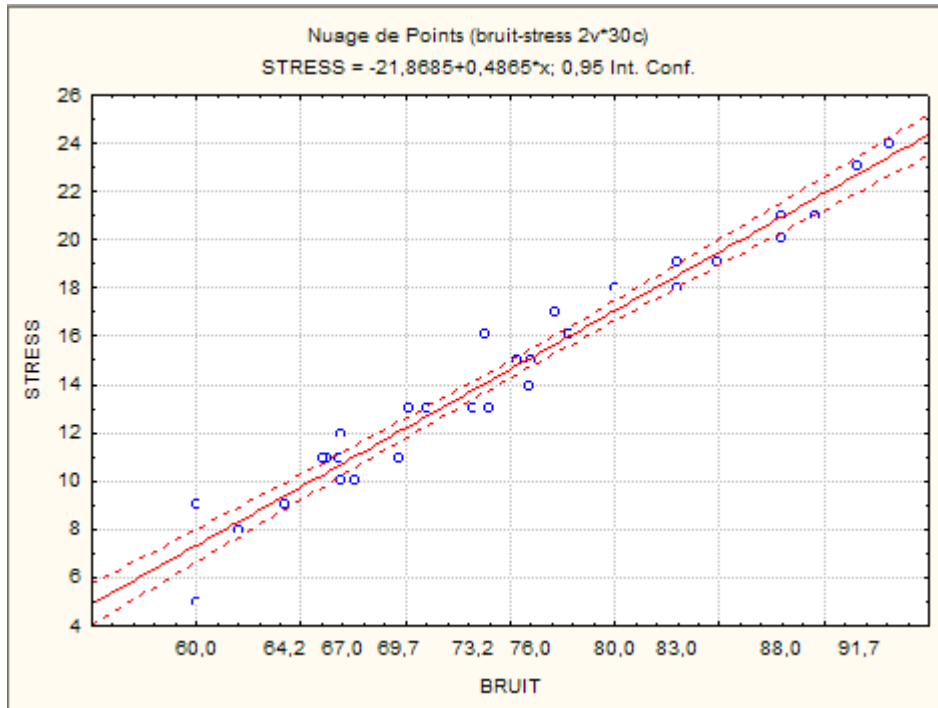
Graphiques / Graphiques en 2D / Nuage de points en 2D / cliquer sur variables / mettre pour X : BRUIT et pour Y : STRESS / cocher type d'ajustement linéaire / dans Avancé, cocher dans l'option Statistiques : R deux, corrélation et équation de la droite



$$stress = -21,8685 + 0,4865 * bruit$$

Utiliser le menu

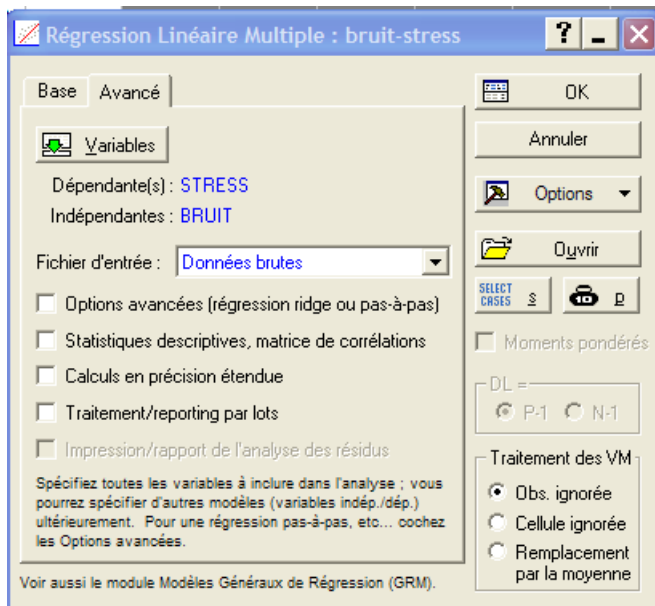
Graphiques / Graphiques en 2D / Nuage de points en 2D / cocher type d'ajustement linéaire / bandes de régression soit confiance soit prévision



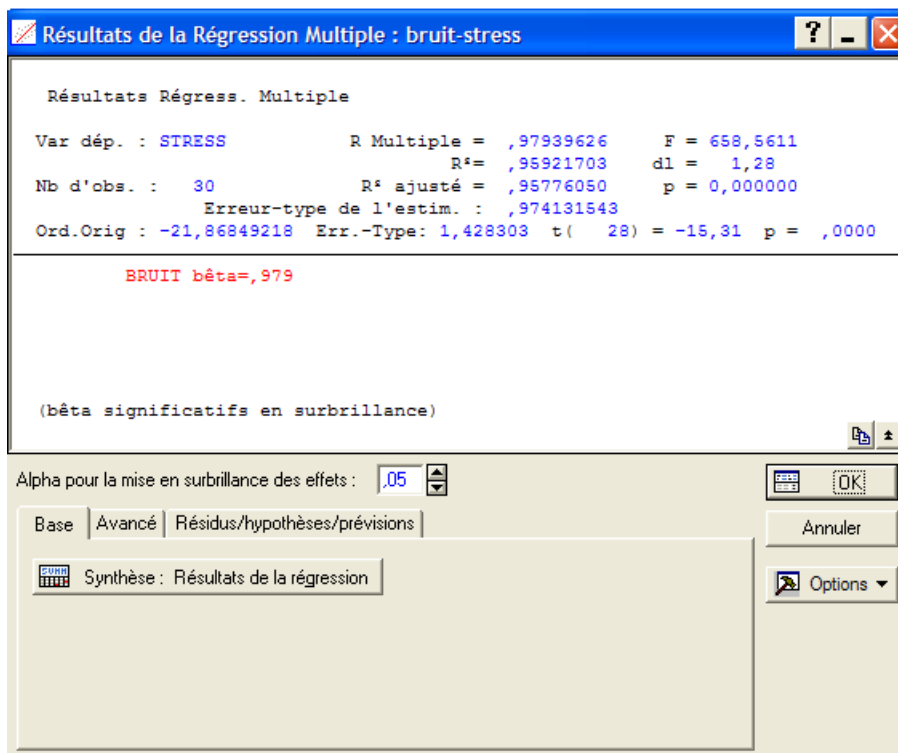
Remarquer que l'intervalle de prévision est plus large que l'intervalle de confiance (cf formules cours)

- Interpréter les résultats de la régression. Relever les valeurs estimées des deux coefficients. Tester séparément la nullité de chacun des coefficients.

Utiliser le menu *Statistiques / Régression multiple*



Taper OK



Cliquer sur synthèse des résultats dans base

Synthèse de la Régression; Variable Dép. : STRESS (bruit-stress)						
R= ,97939626 R²= ,95921703 R² Ajusté = ,95776050						
F(1,28)=658,56 p<0,0000 Err-Type de l'Estim.: ,97413						
N=30	Bêta	Err-Type de Bêta	B	Err-Type de B	t(28)	niveau p
OrdOrig.			-21,8685	1,428304	-15,3108	0,000000
BRUIT	0,979396	0,038165	0,4865	0,018959	25,6624	0,000000

On lit les estimations de b_0 et b_1 dans la colonne B.

La colonne beta correspond aux coefficients estimés pour les variables BRUIT et STRESS centrées réduites, dans le cas simple pas d'intérêt, par contre cela aura de l'intérêt dans le cas multiple (voir ex 2).

$R^2 = 0,959$ très bon ajustement

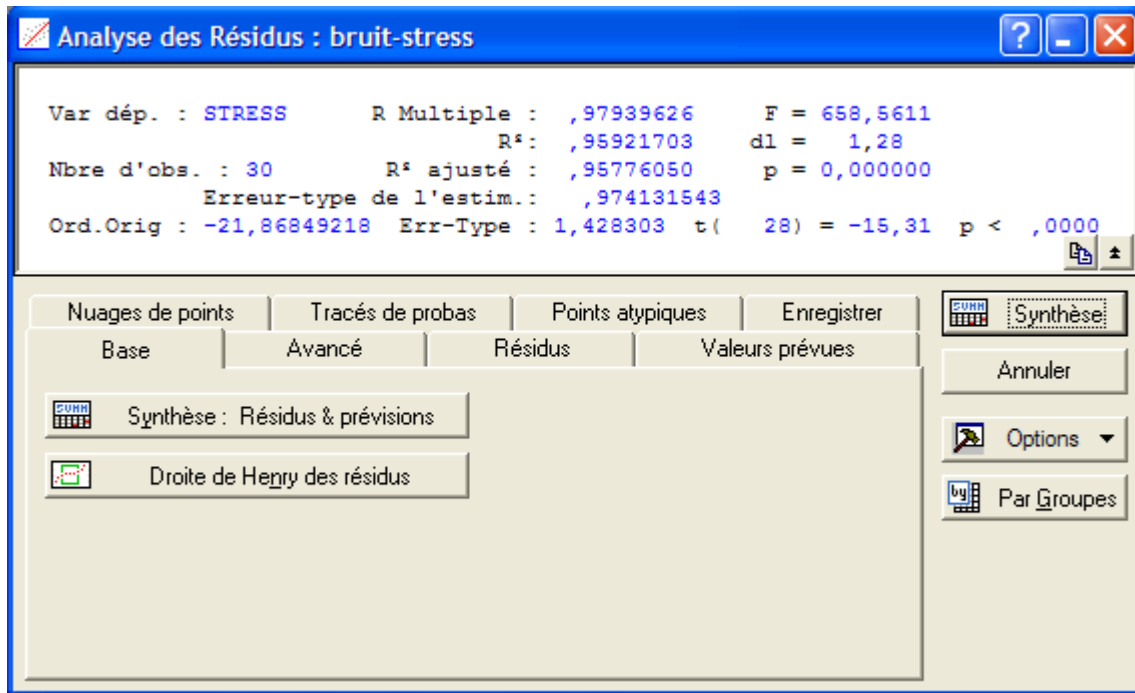
$t(28)$, expliquer ddl = 28 = 30-2 et de plus pour b_1 , on a $(F(1, 28))^{1/2} = t(28)$, $25,6624^2 = 658,56$

On rejette individuellement la nullité de b_0 et b_1 .

- Calculer les valeurs ajustées et les résidus pour toutes les observations.

Rouvrir la boîte de dialogue, cliquer sur Résidus/ hypothèses/Prévisions

Cliquer sur analyse des résidus



Cliquer sur synthèse : résidus et prévisions

N° d'Obs.	Valeurs Prévues & Résidus (bruit-stress)								
	Var. dépendante : STRESS								
	Valeur Observée	Valeur Prévue	Résidus	Val.Pré Standard	Standard Résidus	Err.Type Val.Prév	Mahalanobis (dist.)	Résidus Supprim.	Cook (dist.)
1	11,00000	10,34013	0,65987	-0,89611	0,67740	0,240640	0,803021	0,70276	0,015880
2	11,00000	12,04300	-1,04300	-0,52928	-1,07070	0,201985	0,280142	-1,08986	0,026908
3	23,00000	22,74677	0,25323	1,77651	0,25996	0,367288	3,155980	0,29520	0,006527
4	21,00000	21,72504	-0,72504	1,55641	-0,74430	0,333012	2,422411	-0,82099	0,041504
5	9,00000	9,36706	-0,36706	-1,10573	-0,37680	0,267653	1,222643	-0,39703	0,006270
6	18,00000	18,51391	-0,51391	0,86467	-0,52756	0,236845	0,747658	-0,54620	0,009293
7	9,00000	7,32361	1,67639	-1,54593	1,72091	0,331410	2,389896	1,89582	0,219193
8	14,00000	15,10817	-1,10817	0,13101	-1,13760	0,179423	0,017164	-1,14708	0,023521
9	21,00000	20,94659	0,05341	1,38872	0,05483	0,307792	1,928531	0,05934	0,000185
10	16,00000	15,98393	0,01607	0,31967	0,01650	0,187016	0,102187	0,01668	0,000005
11	8,00000	8,29668	-0,29668	-1,33631	-0,30456	0,300106	1,785728	-0,32779	0,005373
12	11,00000	10,24282	0,75718	-0,91708	0,77729	0,243210	0,841029	0,80752	0,021417
13	10,00000	11,02127	-1,02127	-0,74938	-1,04840	0,223622	0,561574	-1,07809	0,032273

6. Prévoir le niveau de stress que provoquerait un niveau de bruit de 90db.

Utiliser le menu *Statistiques / Régression multiple / Résidus , hypothèses , prévision / Prévoir les valeurs de la variable dépendante*

Résultats de la Régression Multiple : bruit-stress

Résultats Régress. Multiple

Var dép. : STRESS R Multiple = ,97939626 F = 658,5611
R² = ,95921703 dl = 1,28
Nb d'obs. : 30 R² ajusté = ,95776050 p = 0,000000
Erreur-type de l'estim. : ,974131543
Ord.Orig : -21,86849218 Err.-Type: 1,428303 t(28) = -15,31 p = ,0000

BRUIT bêta=,979

(bêta significatifs en surbrillance)

Alpha pour la mise en surbrillance des effets : .05

Base | Avancé | Résidus/hypothèses/prévisions

Analyse des résidus | Statistiques descriptives | Générateur de code

Valeurs prévues
?▶ Prévoir les valeurs de la variable dépendante
 Calculer les limites de confiance Alpha : .05
 Calculer les limites de prévision

OK | Annuler | Options

Résultats de la Régression Multiple : bruit-stress

Résultats Régress. Multiple

Var dép. : STRESS R Multiple = ,97939626 F = 658,5611
R² = ,95921703 dl = 1,28
Nb d'obs. : 30 R² ajusté = ,95776050 p = 0,000000
Erreur-type de l'estim. : ,974131543
Ord.Orig : -21,86849218 Err.-Type: 1,428303 t(28) = -15,31 p = ,0000

BRUIT bêta=,979

(bêta significatifs en surbrillance)

Alpha pour la mise en surbrillance des effets : .05

Base | Avancé | Résidus/hypothèses/prévisions

Analyse des résidus | Statistiques descriptives | Générateur de code

Valeurs prévues
?▶ Prévoir les valeurs de la variable dépendante
 Calculer les limites de confiance Alpha : .05
 Calculer les limites de prévision

Spécifiez les valeurs des var. indép.
BRUIT 90
OK | Annuler
Valeur commune 0
Appliquer

OK | Annuler | Options | Par Groupes

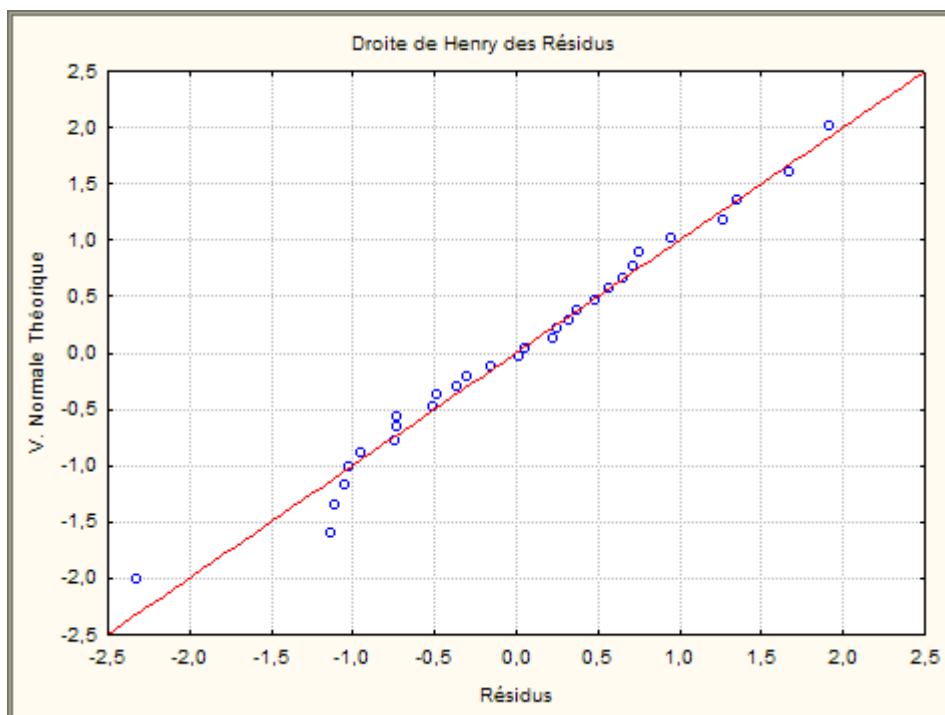
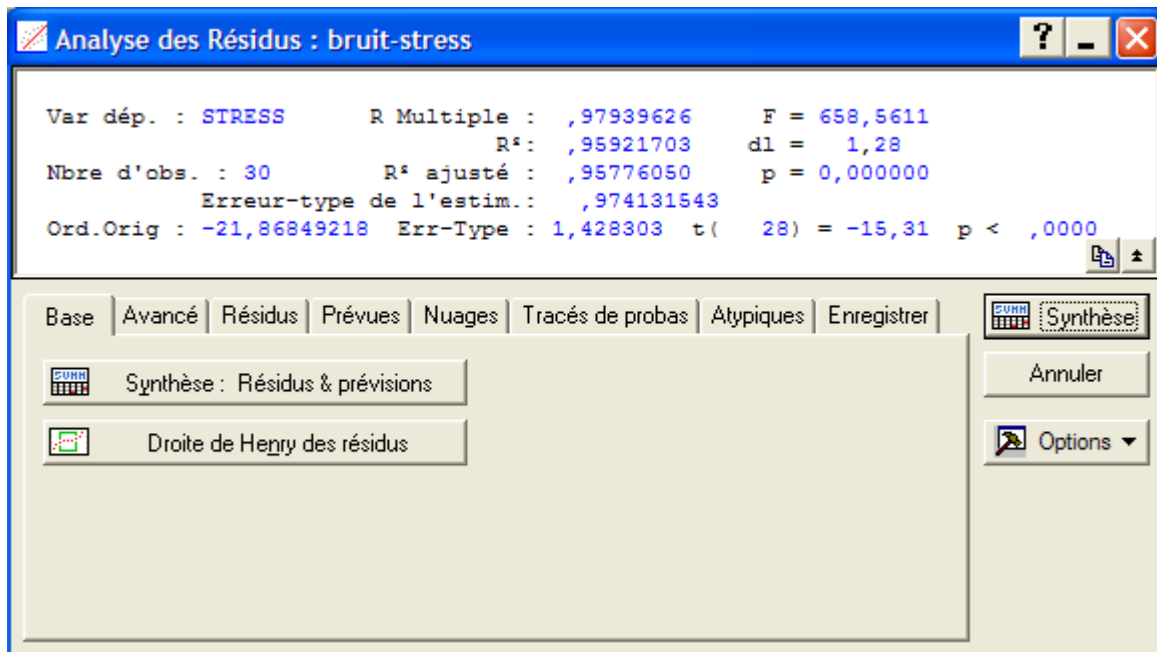
Valeurs Prévues (bruit-stress)			
Variable: STRESS			
Variable	Pond-B	Valeur	Pond-B * Valeur
BRUIT	0,486535	90,00000	43,7882
Ord.Orig			-21,8685
Prévision			21,9197
-95,0%LC			21,2243
+95,0%LC			22,6150

7. Vérifier les conditions.

(a) Vérifier la normalité des résidus.

Utiliser le menu *Statistiques / Régression multiple / Résidus , hypothèses , prévision / Analyse des résidus*

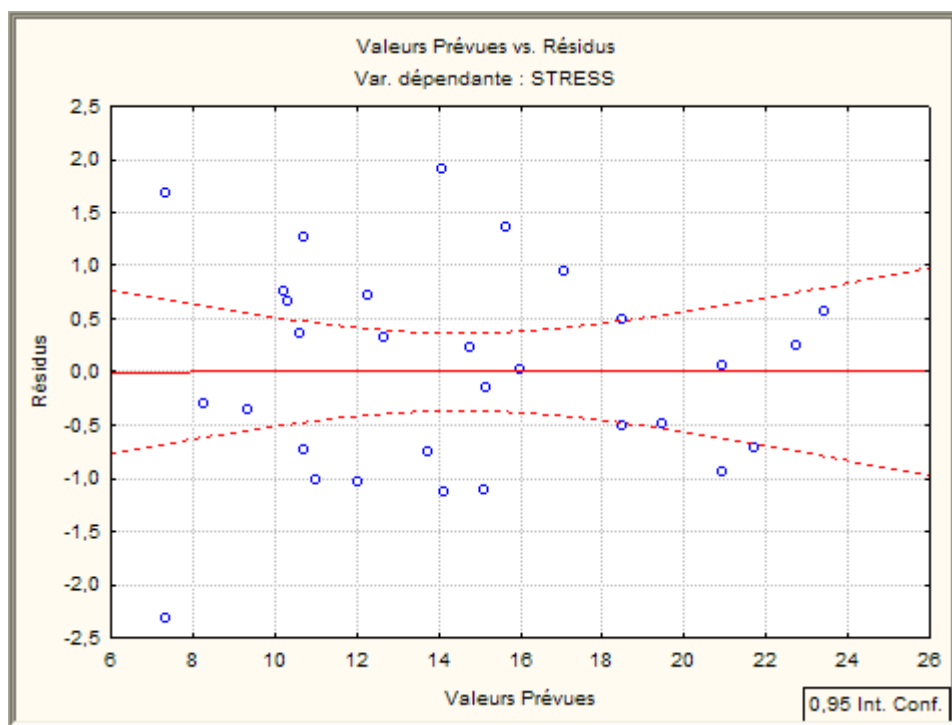
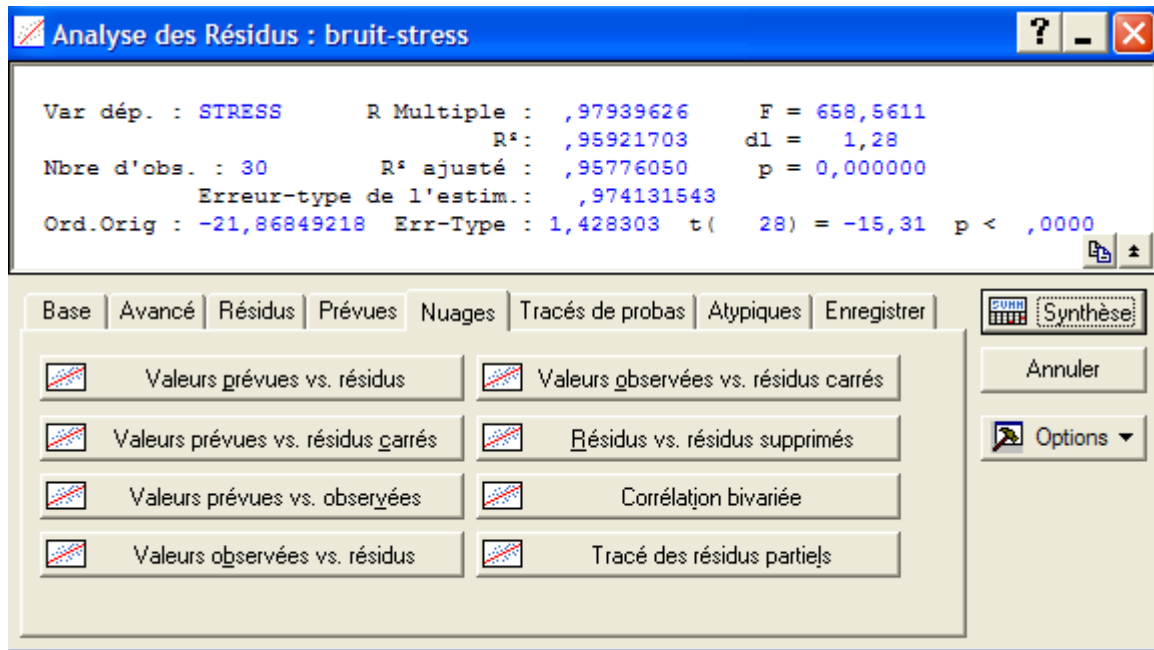
Cliquer sur base puis droite de Henry



Les points sont proches de la droite, on peut accepter la normalité

(b) Vérifier l'hypothèse d'homoscédasticité.

Utiliser le menu *Statistiques / Régression multiple / Résidus, hypothèses, prévision / Analyse des résidus / Nuages / valeurs prévues vs résidus*



pas de forme particulière du nuage, donc homoscedasticité, cadre général permet aussi de vérifier l'adéquation du modèle car nuage de points impossible

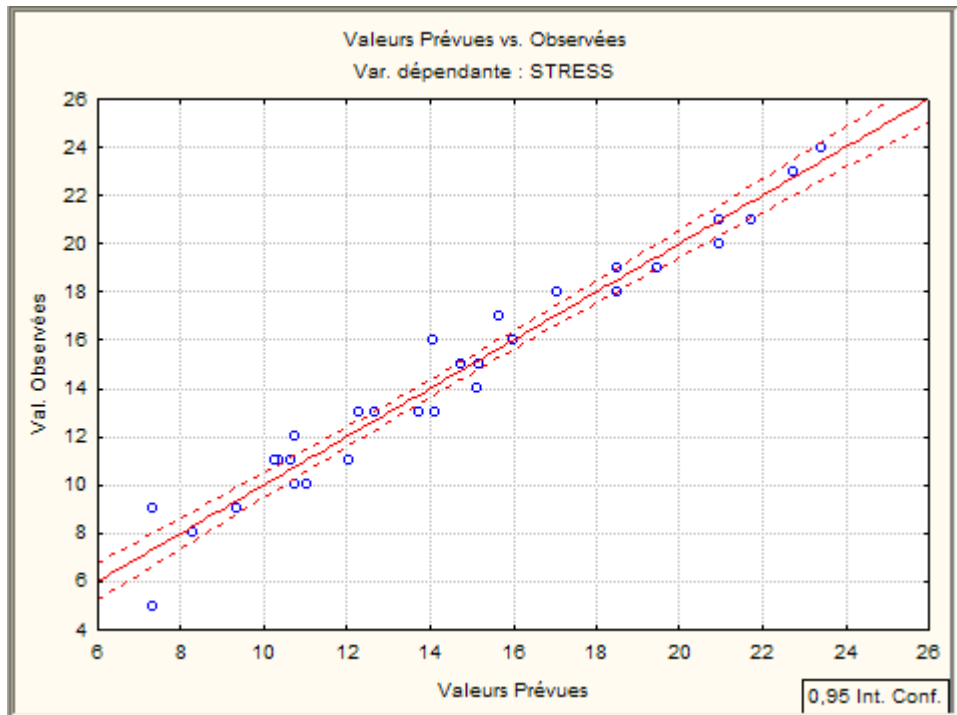
8. Analyse des valeurs atypiques.

Il y a deux graphiques pour repérer les observations "atypiques".

Utiliser le menu *Statistiques / Régression multiple / Résidus, hypothèses, prévision / Analyse des résidus / Nuages/ valeurs prévues vs observées*

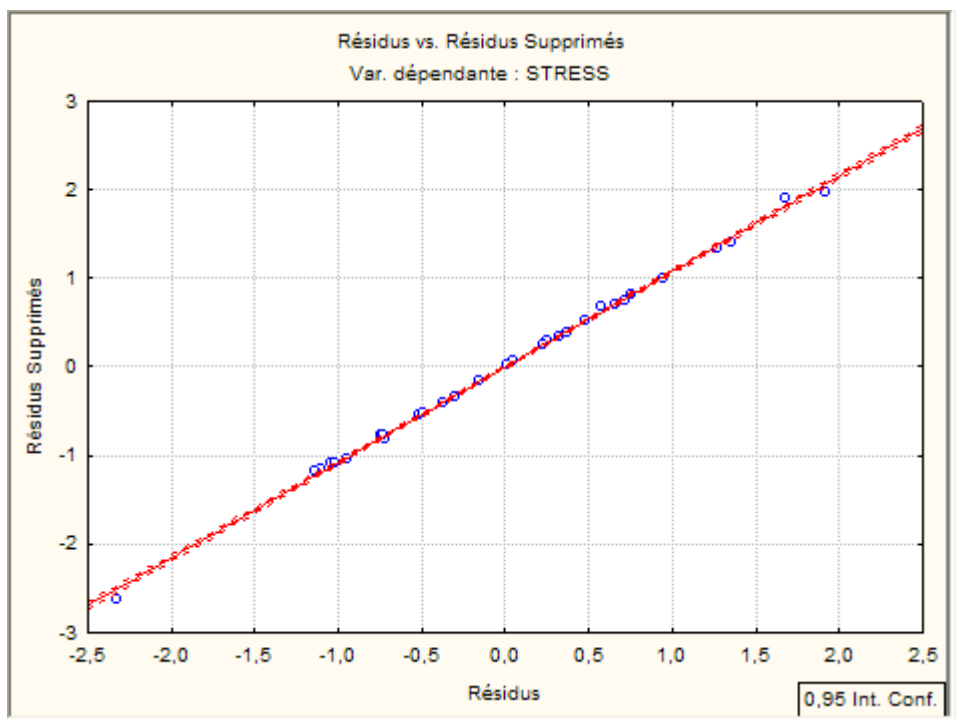
Ce graphique permet de détecter les observations mal prévues par le modèle (résidus anormalement

élevés).



Ce graphique permet de détecter les observations qui ont une influence exagérée sur l'estimation des coefficients.

Utiliser le menu *Statistiques / Régression multiple / Résidus, hypothèses , prévision / Analyse des résidus / Nuages/résidus vs résidus supprimés*

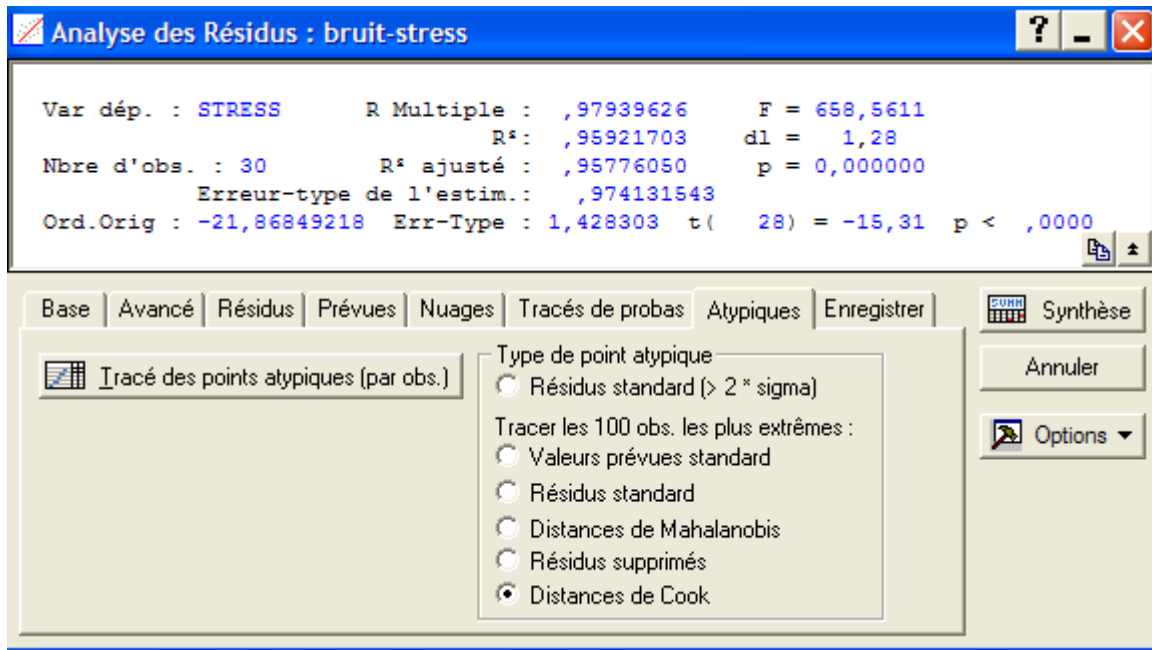


On ne repère pas de points particuliers

Utiliser le menu *Statistiques / Régression multiple / Résidus, hypothèses, prévision / Analyse des résidus / Atypique*

Cliquer sur *Tracé des points atypiques*

							Résidus Standard : STRESS (bruit-stress)					
							Pts Atypiques					
Résidus Standard							Valeur	Valeur	Résidus	Standard	Standard	Err.Type
Obs.	-5.	-4.	-3.	±2.	3.	4.	Observée	Prévue		Val.Prév	Résidus	Val.Prév
24	.	.	.	*	.	.	5,000000	7,323608	-2,32361	-1,54593	-2,38531	0,331410
Minimum	.	.	.	*	.	.	5,000000	7,323608	-2,32361	-1,54593	-2,38531	0,331410
Maximum	.	.	.	*	.	.	5,000000	7,323608	-2,32361	-1,54593	-2,38531	0,331410
Moyenne	.	.	.	*	.	.	5,000000	7,323608	-2,32361	-1,54593	-2,38531	0,331410
Médiane	.	.	.	*	.	.	5,000000	7,323608	-2,32361	-1,54593	-2,38531	0,331410



Cliquer sur *Tracé des points atypiques*

		Dist. de Cook : STRESS (bruit-stress) (trié)									
Obs.	,000	Dist. de Cook		Valeur Observée	Valeur Prévue	Résidus	Standard Val.Prév	Standard Résidus	Err.Type Val.Prév	Mahalanobis (dist.)	Résidus Supprim.
24	.	.	.	5,00000	7,32361	-2,32361	-1,54593	-2,38531	0,331410	2,389896	-2,62775
7	.	.	*	9,00000	7,32361	1,67639	-1,54593	1,72091	0,331410	2,389896	1,89582
16	.	*	.	16,00000	14,08644	1,91355	-0,08909	1,96437	0,178580	0,007937	1,98010
27	.	*	.	20,00000	20,94659	-0,94659	1,38872	-0,97173	0,307792	1,928531	-1,05157
26	.	*	.	12,00000	10,72935	1,27065	-0,81227	1,30439	0,230695	0,659779	1,34614
4	.	*	.	21,00000	21,72504	-0,72504	1,55641	-0,74430	0,333012	2,422411	-0,82099
22	.	*	.	24,00000	23,42792	0,57208	1,92324	0,58727	0,390723	3,698853	0,68176
17	.	*	.	17,00000	15,64336	1,35664	0,24630	1,39267	0,183347	0,060664	1,40647
13	.	*	.	10,00000	11,02127	-1,02127	-0,74938	-1,04840	0,223622	0,561574	-1,07809
2	.	*	.	11,00000	12,04300	-1,04300	-0,52928	-1,07070	0,201985	0,280142	-1,08986
23	.	*	.	13,00000	14,13510	-1,13510	-0,07861	-1,16524	0,178419	0,006179	-1,17450
8	.	*	.	14,00000	15,10817	-1,10817	0,13101	-1,13760	0,179423	0,017164	-1,14708
18	.	*	.	18,00000	17,05431	0,94569	0,55025	0,97081	0,203809	0,302771	0,98898
12	.	*	.	11,00000	10,24282	0,75718	-0,91708	0,77729	0,243210	0,841029	0,80752
29	.	*	.	10,00000	10,72935	-0,72935	-0,81227	-0,74872	0,230695	0,659779	-0,77269
1	.	*	.	11,00000	10,34013	0,65987	-0,89611	0,67740	0,240640	0,803021	0,70276
25	.	*	.	13,00000	12,28627	0,71373	-0,47688	0,73269	0,197668	0,227414	0,74438
28	.	*	.	13,00000	13,74587	-0,74587	-0,16245	-0,76568	0,180263	0,026391	-0,77232
14	.	*	.	19,00000	19,48698	-0,48698	1,07429	-0,49992	0,263430	1,154098	-0,52541
6	.	*	.	18,00000	18,51391	-0,51391	0,86467	-0,52756	0,236845	0,747658	-0,54620
21	.	*	.	19,00000	18,51391	0,48609	0,86467	0,49899	0,236845	0,747658	0,51663
3	.	*	.	23,00000	22,74677	0,25323	1,77651	0,25996	0,367288	3,155980	0,29520
5	.	*	.	9,00000	9,36706	-0,36706	-1,10573	-0,37680	0,267653	1,222643	-0,39703
11	.	*	.	8,00000	8,29668	-0,29668	-1,33631	-0,30456	0,300106	1,785728	-0,32779
19	.	*	.	11,00000	10,63205	0,36795	-0,83323	0,37772	0,233128	0,694271	0,39031
30	.	*	.	13,00000	12,67549	0,32451	-0,39303	0,33312	0,191535	0,154475	0,33756
15	.	*	.	15,00000	14,76759	0,23241	0,05764	0,23858	0,178157	0,003323	0,24045
20	.	*	.	15,00000	15,15682	-0,15682	0,14149	-0,16099	0,179684	0,020020	-0,16235

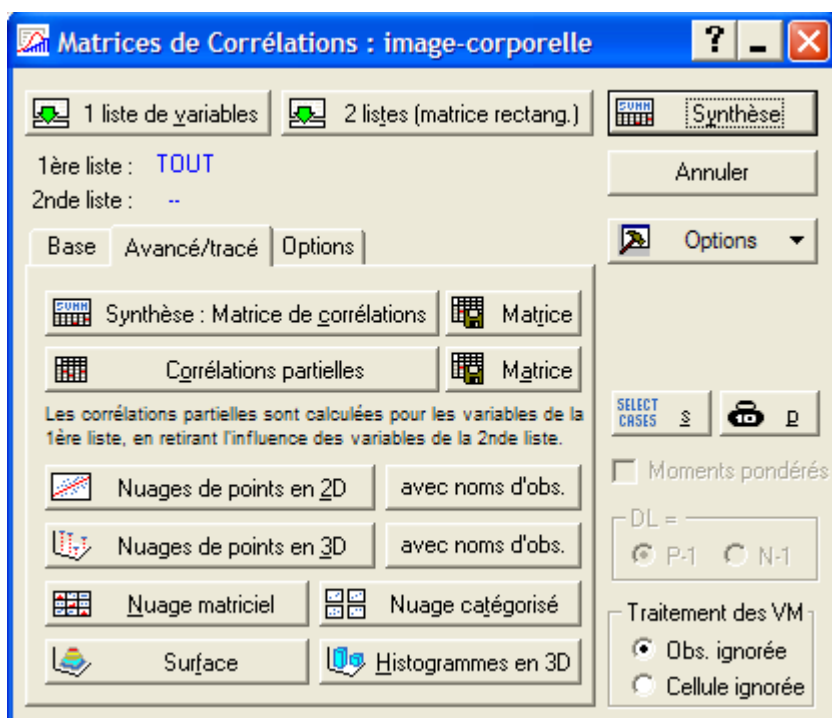
Les résidus standard permettent de détecter les valeurs atypiques au niveau de la variable expliquée (mal prévue), la distance de Mahalanobis permet de détecter les points leviers, les résidus supprimés permettent de détecter les observations influentes, la distance de Cook permet de détecter tous ces points. Outil important en régression multiple car on ne peut pas représenter les variables. L'individu 24 est légèrement atypique, pour bruit=60, stress = 5, voir sur le nuage de points les observations 24, 7, 16,...

Exercice 2 : Régression multiple

Le but de cette étude est de montrer que la variable SAWBS (Shape And Weight Based Self-esteem inventory) a un rôle dans la prévision d'un dérèglement du comportement de l'alimentation, rôle indépendant des variables traditionnellement mis en cause dans ce comportement tels que la dépression, l'estime de soi...

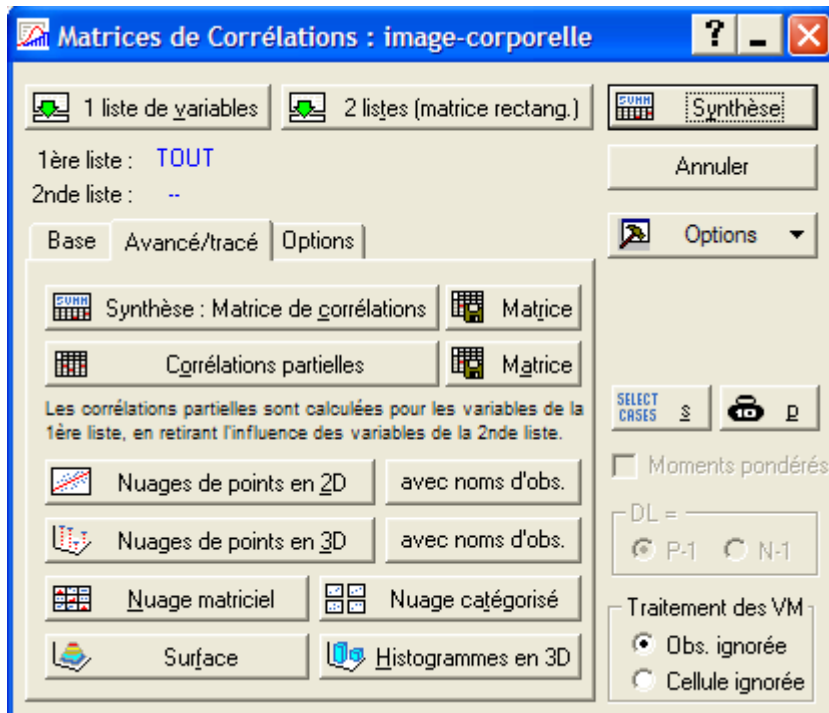
1. Etude des relations entre variables.

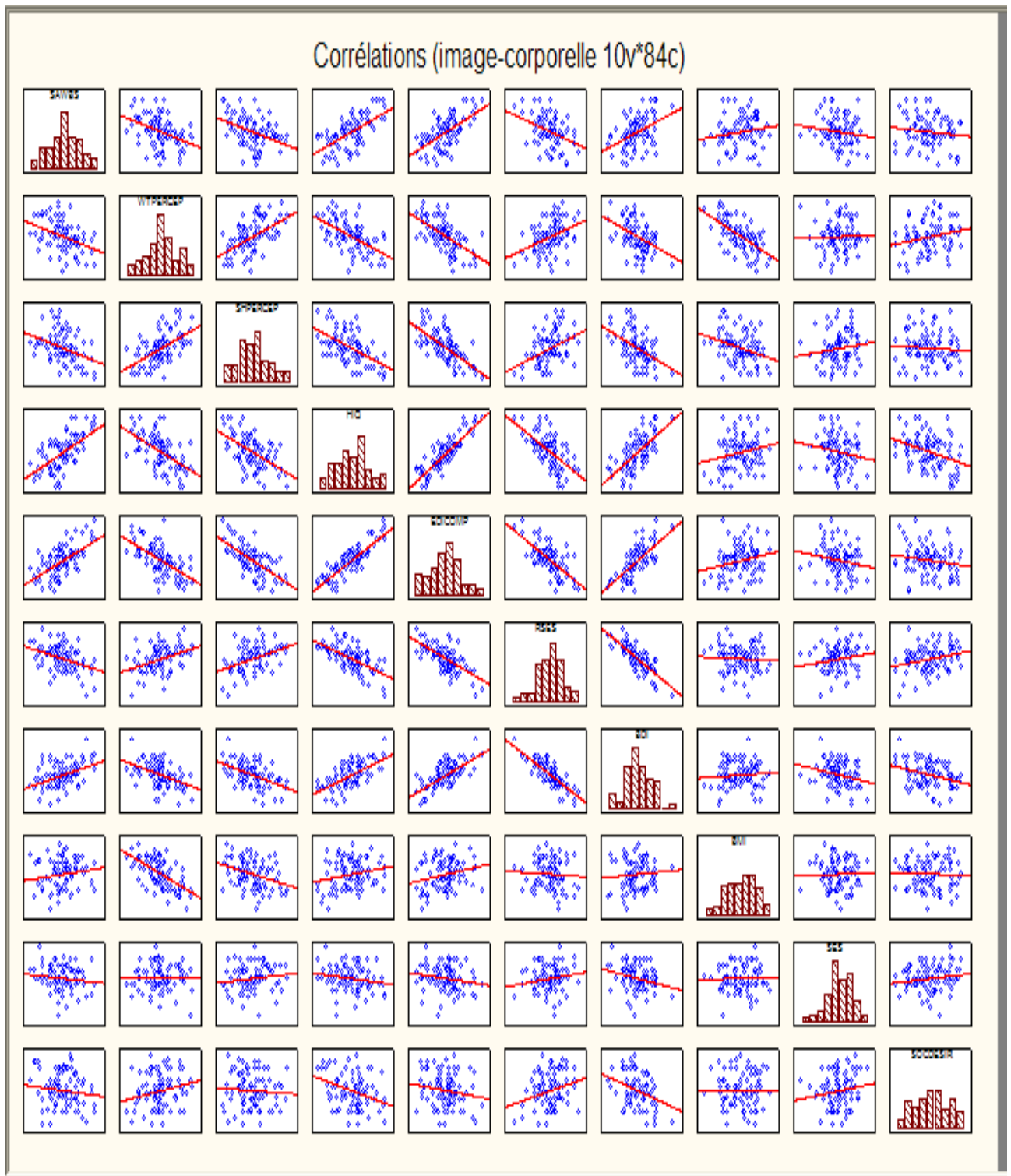
Utiliser le menu *Statistiques / Statistiques élémentaires / Matrice de corrélations* Sélectionner toutes les variables et cliquer sur Synthèse : Matrice de corrélations



Corrélations (image-corporelle)										
Corrélations significatives marquées à $p < ,05000$										
N=84 (Observations à VM ignorées)										
Variable	SAWBS	WTPERCEP	SHPERCEP	HIQ	EDICOMP	RSES	BDI	BMI	SES	SOCDESIR
SAWBS	1,00	-0,39	-0,39	0,61	0,61	-0,38	0,42	0,17	-0,13	-0,13
WTPERCEP	-0,39	1,00	0,56	-0,56	-0,61	0,38	-0,45	-0,61	0,00	0,23
SHPERCEP	-0,39	0,56	1,00	-0,55	-0,67	0,42	-0,47	-0,32	0,14	-0,07
HIQ	0,61	-0,56	-0,55	1,00	0,86	-0,61	0,66	0,22	-0,18	-0,36
EDICOMP	0,61	-0,61	-0,67	0,86	1,00	-0,68	0,71	0,24	-0,18	-0,17
RSES	-0,38	0,38	0,42	-0,61	-0,68	1,00	-0,79	-0,06	0,18	0,26
BDI	0,42	-0,45	-0,47	0,66	0,71	-0,79	1,00	0,08	-0,25	-0,33
BMI	0,17	-0,61	-0,32	0,22	0,24	-0,06	0,08	1,00	0,02	-0,01
SES	-0,13	0,00	0,14	-0,18	-0,18	0,18	-0,25	0,02	1,00	0,16
SOCDESIR	-0,13	0,23	-0,07	-0,36	-0,17	0,26	-0,33	-0,01	0,16	1,00

Utiliser le menu *Statistiques / Statistiques élémentaires / Matrice de corrélations* Sélectionner toutes les variables et cliquer sur Nuage matriciel





fortes corrélations d'edicomp avec hq, bdi,..., corrélation faible entre sawbx et bmi=0,17

2. Etude du modèle de régression multiple

(a) Pour répondre au problème initial, que peut-on choisir comme VI et VD ?

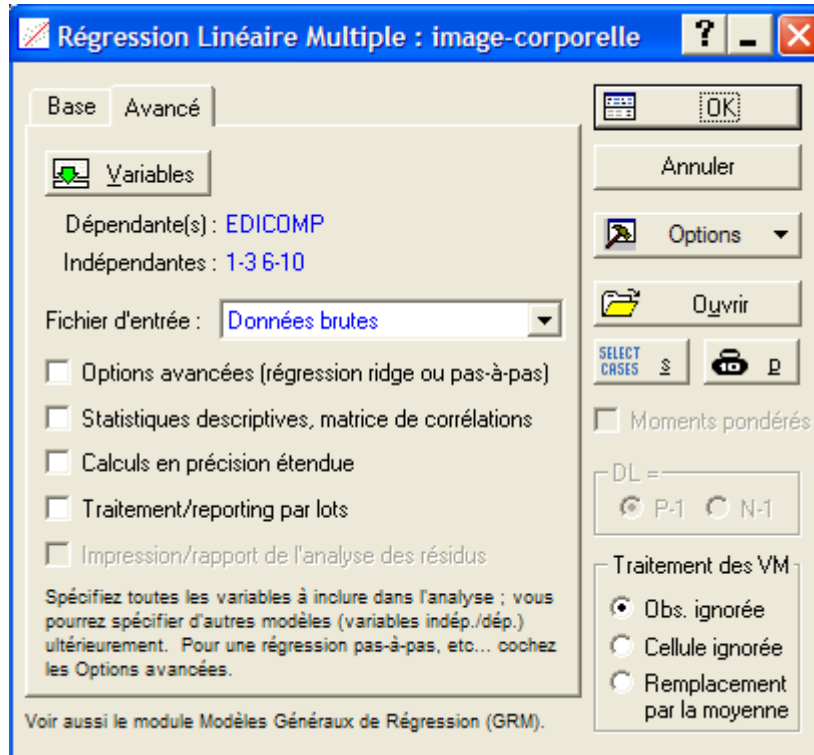
on peut prendre comme VD : edicomp ou hq et VI les autres

(b) Dans la suite, on considère la variable dépendante : EDIcomp et 8 variables indépendantes :

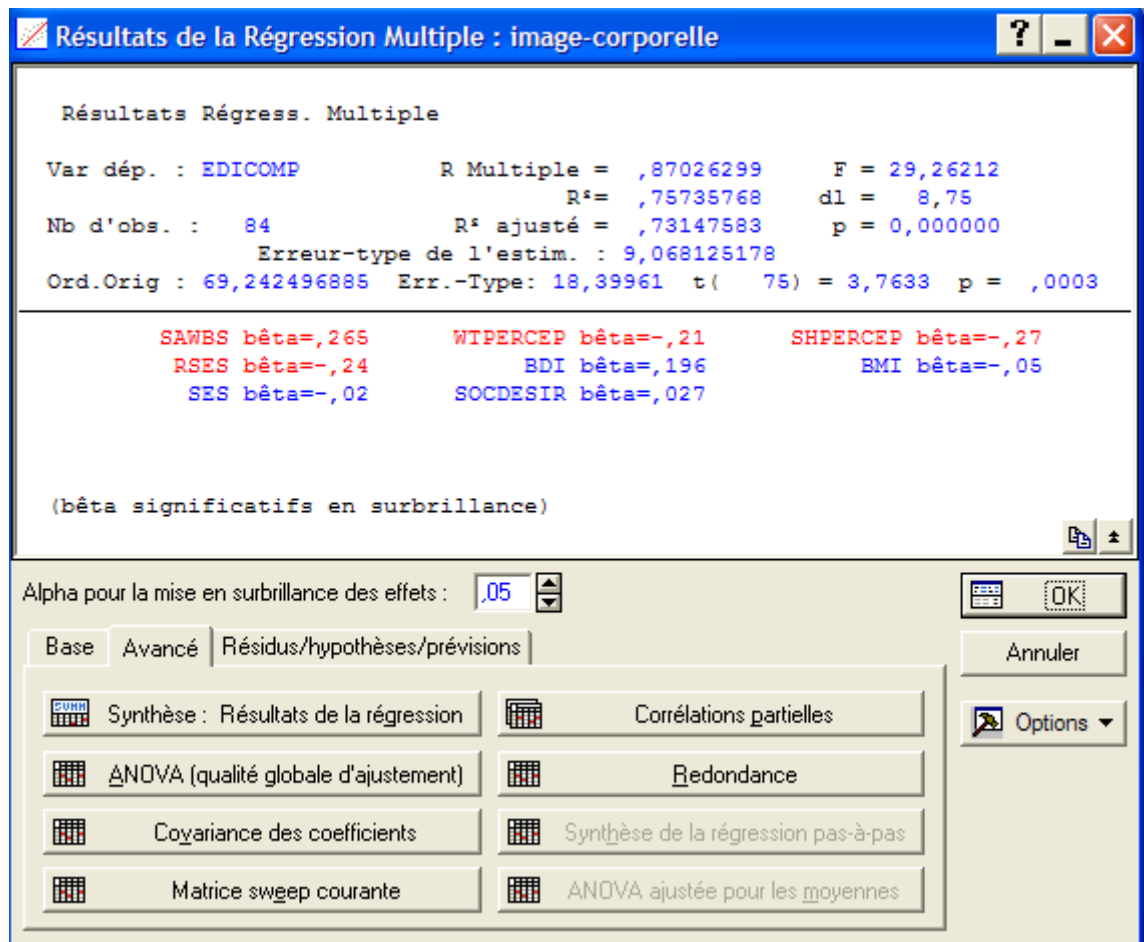
SAWBS, WtPercep, ShPercep, RSES, BDI, BMI, SES, SocDesir.

- i. Donner l'équation de régression multiple de EDIcomp en fonction de SAWBS, WtPercep, ShPercep, RSES, BDI, BMI, SES, SocDesir.

Utiliser le menu *Statistiques / Régression multiple*



Taper OK



Cliquer sur synthèse des résultats dans base

Synthèse de la Régression; Variable Dép. : EDICOMP (image-corporelle)						
R= .87026299 R²= .75735768 R² Ajusté = .73147583						
F(8,75)=29,262 p<0,0000 Err-Type de l'Estim.: 9,0681						
N=84	Bêta	Err-Type de Bêta	B	Err-Type de B	t(75)	niveau p
OrdOrig.			69,24250	18,39961	3,76326	0,000331
SAWBS	0,265329	0,065739	0,09985	0,02474	4,03607	0,000130
WTPERCEP	-0,210000	0,094154	-5,24591	2,35200	-2,23040	0,028712
SHPERCEP	-0,266485	0,078214	-3,88609	1,14057	-3,40715	0,001058
RSES	-0,240468	0,094372	-0,51320	0,20141	-2,54807	0,012878
BDI	0,196144	0,104126	0,47670	0,25306	1,88372	0,063479
BMI	-0,052406	0,075707	-0,29583	0,42736	-0,69222	0,490937
SES	-0,016656	0,060339	-0,29142	1,05569	-0,27605	0,783272
SOCDESIR	0,026794	0,065594	0,23441	0,57384	0,40849	0,684080

ii. Tester l'hypothèse "Tous les coefficients $b_j, j = 1, \dots, 8$ " sont nuls.

$$F = \frac{R^2}{1-R^2} \frac{n-p-1}{p}, \text{ on rejette la nullité de tous les coefficients}$$

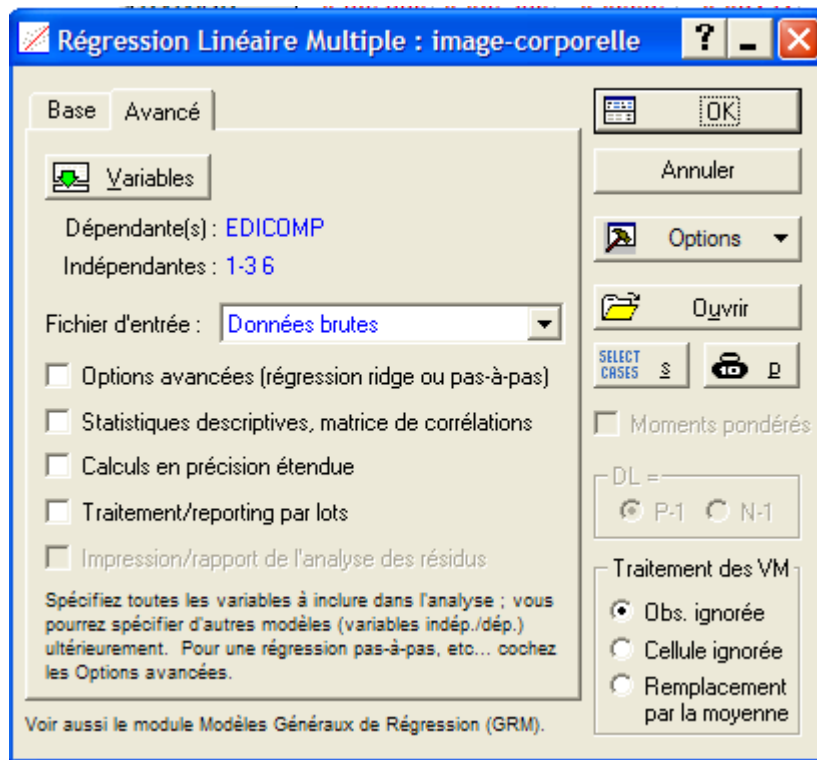
iii. Tester individuellement la nullité des paramètres $b_j, j = 0, \dots, 8$.

on rejette la nullité de $b_0, b_1, b_2, b_3, b_4, b_4$, on accepte la nullité $b_5, b_6, b_7, b_8, t(75), 75 = 84 - 8 - 1$

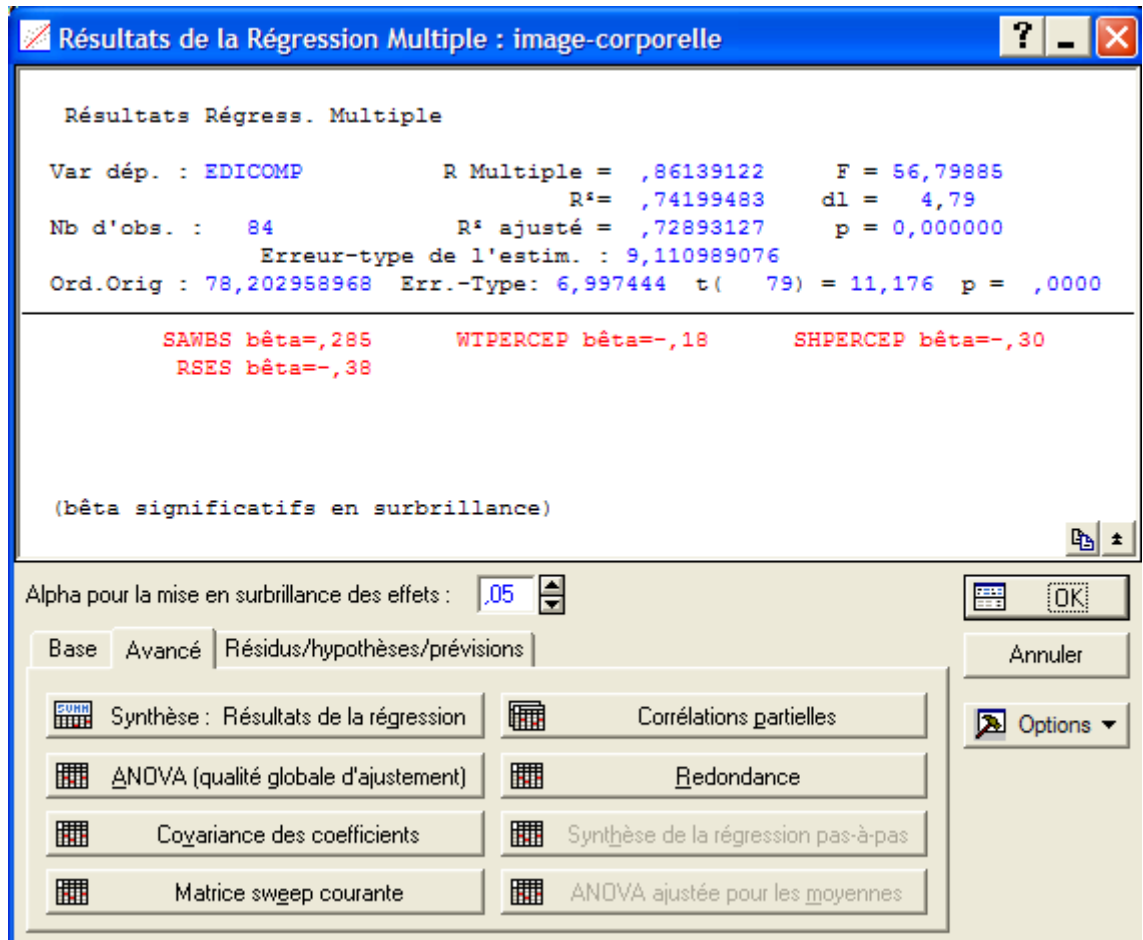
Que peut-on conclure ? En choisissant le modèle restreint, que vaut le coefficient R^2 de ce nouveau modèle ?

$R^2 = 0,7574$ avec modèle complet, nouveau modèle avec les 4 VI (variables 1,2,3,6), $R^2 = 0,742$, on ne perd pas grand chose.

Utiliser le menu *Statistiques / Régression multiple*

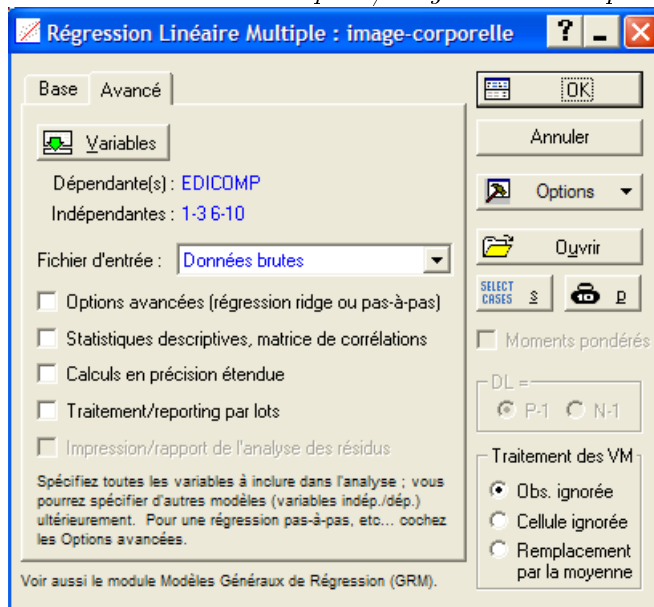


Taper OK

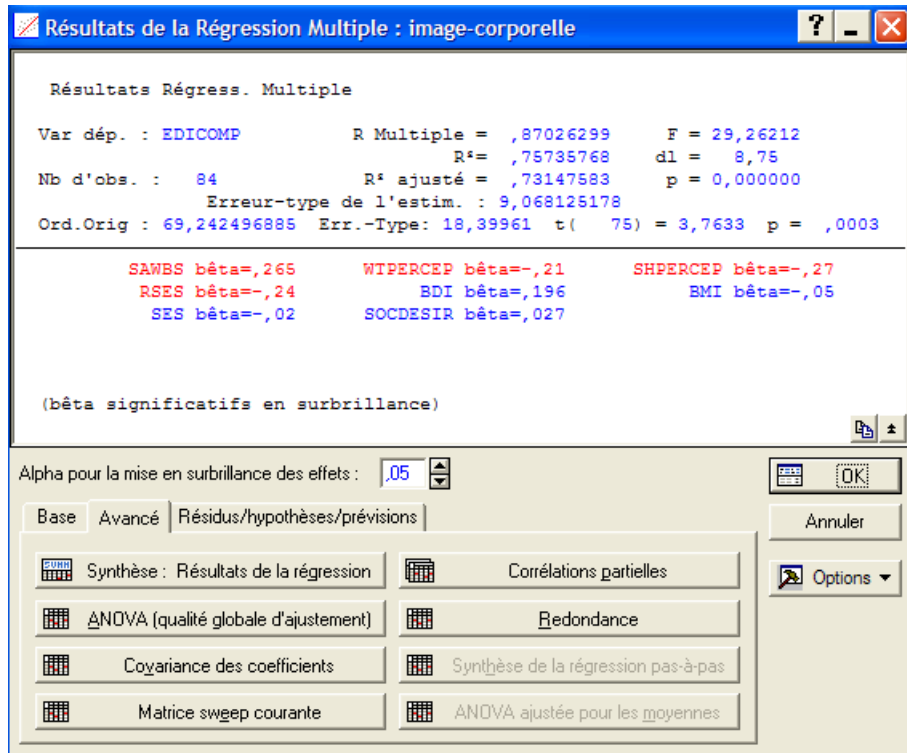


iv. Etudier la redondance des variables.

Utiliser le menu *Statistiques / Régression multiple*



Taper OK



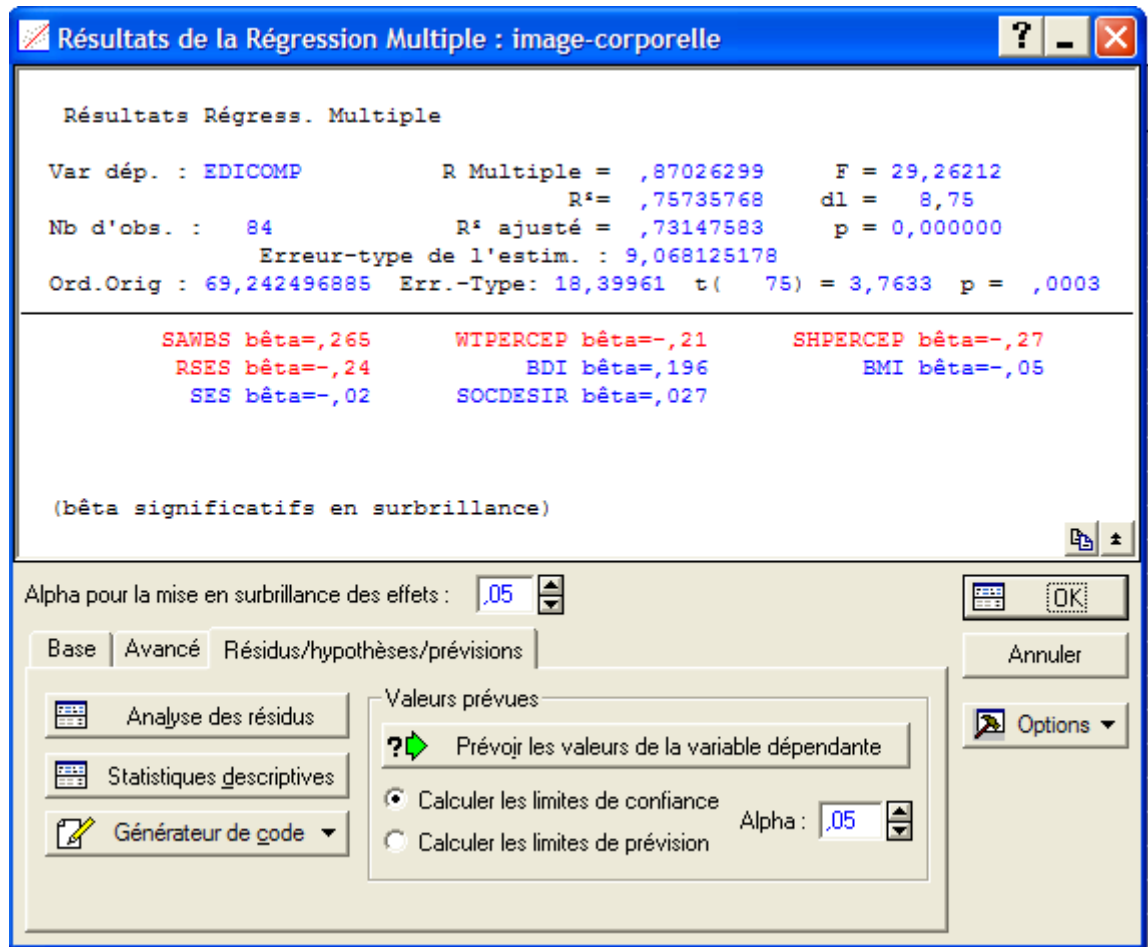
Cliquer sur redondance

Redondance des var. indépendantes ; VD: EDICOMP (image-corporelle) La colonne R ² contient les R ² des variables respectives avec toutes les autres variables indépendantes					
Variable	Toléranc.	R ²	Corrél. Partiel.	Corrél. Semipart	
SAWBS	0,748607	0,251393	0,422423	0,229568	
WTPERCEP	0,364948	0,635052	-0,249406	-0,126863	
SHPERCEP	0,528861	0,471139	-0,366109	-0,193796	
RSES	0,363259	0,636741	-0,282262	-0,144932	
BDI	0,298393	0,701607	0,212543	0,107144	
BMI	0,564454	0,435546	-0,079677	-0,039373	
SES	0,888614	0,111386	-0,031859	-0,015701	
SOCDESIR	0,751941	0,248059	0,047116	0,023234	

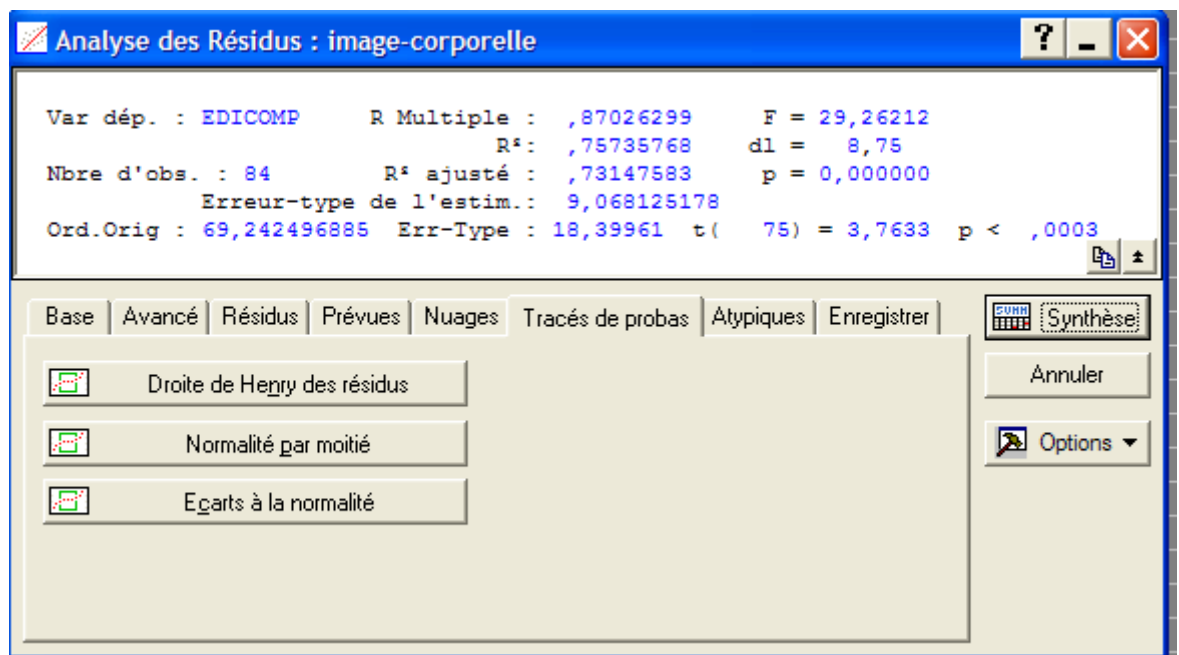
Tolérance : $1 - R^2$, si en dessous de 0,1, variable redondante, c'est à dire que par rapport aux autres, elle n'apporte rien, elle est fortement corrélée avec l'une ou combinaison des autres. Pour la corrélation partielle, on retire l'influence des autres. La plus petite est 0,29, pour la corrélation partielle , c'est bon pour les 4 premières variables, cela confirme ce qu'on a vu précédemment

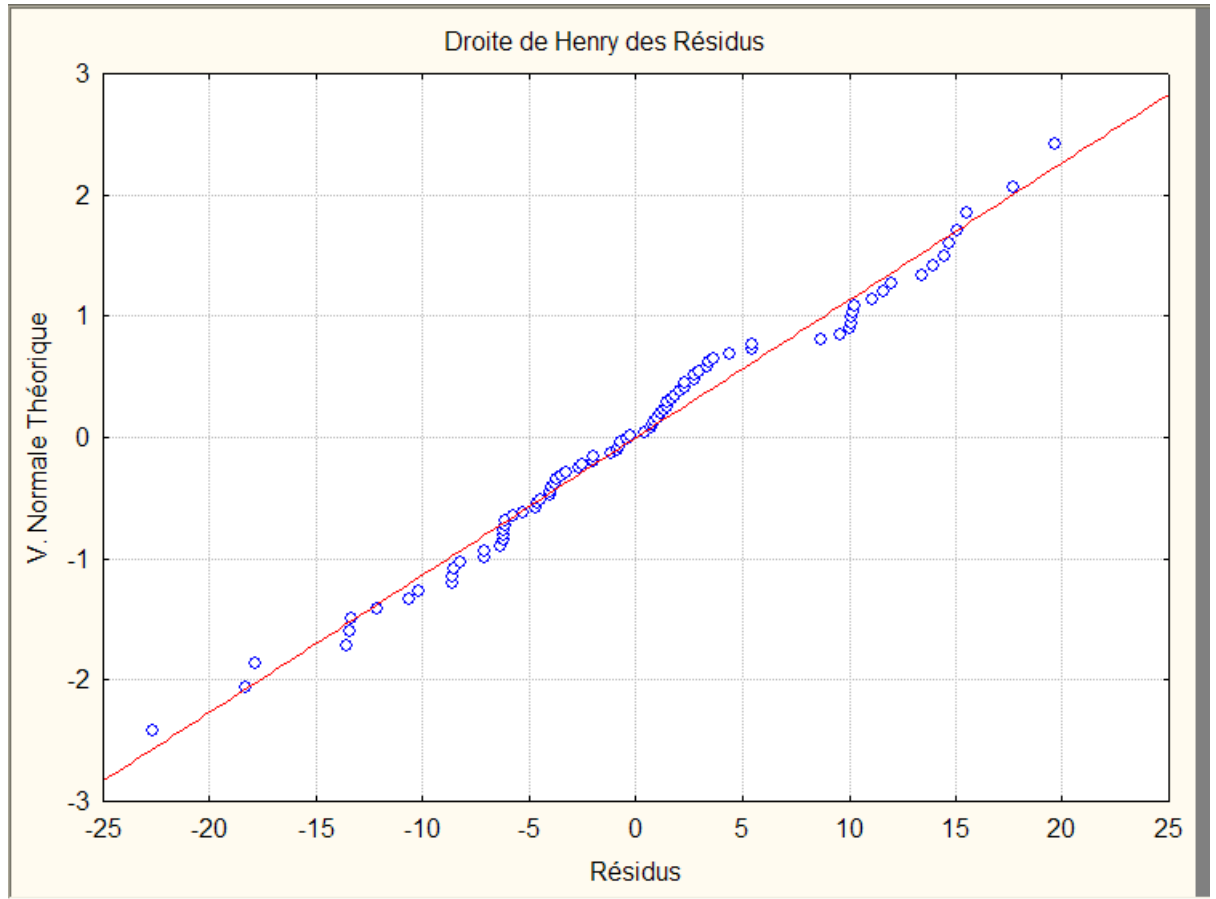
v. Analyse des résidus du modèle complet.

Utiliser le menu *Statistiques / Régression multiple / Résidus , hypothèses , prévision / Analyse des résidus*



Cliquer sur base puis droite de Henry





Utiliser le menu *Statistiques / Régression multiple / Résidus, hypothèses , prévision / Analyse des résidus / Nuages / valeurs prévues vs résidus*

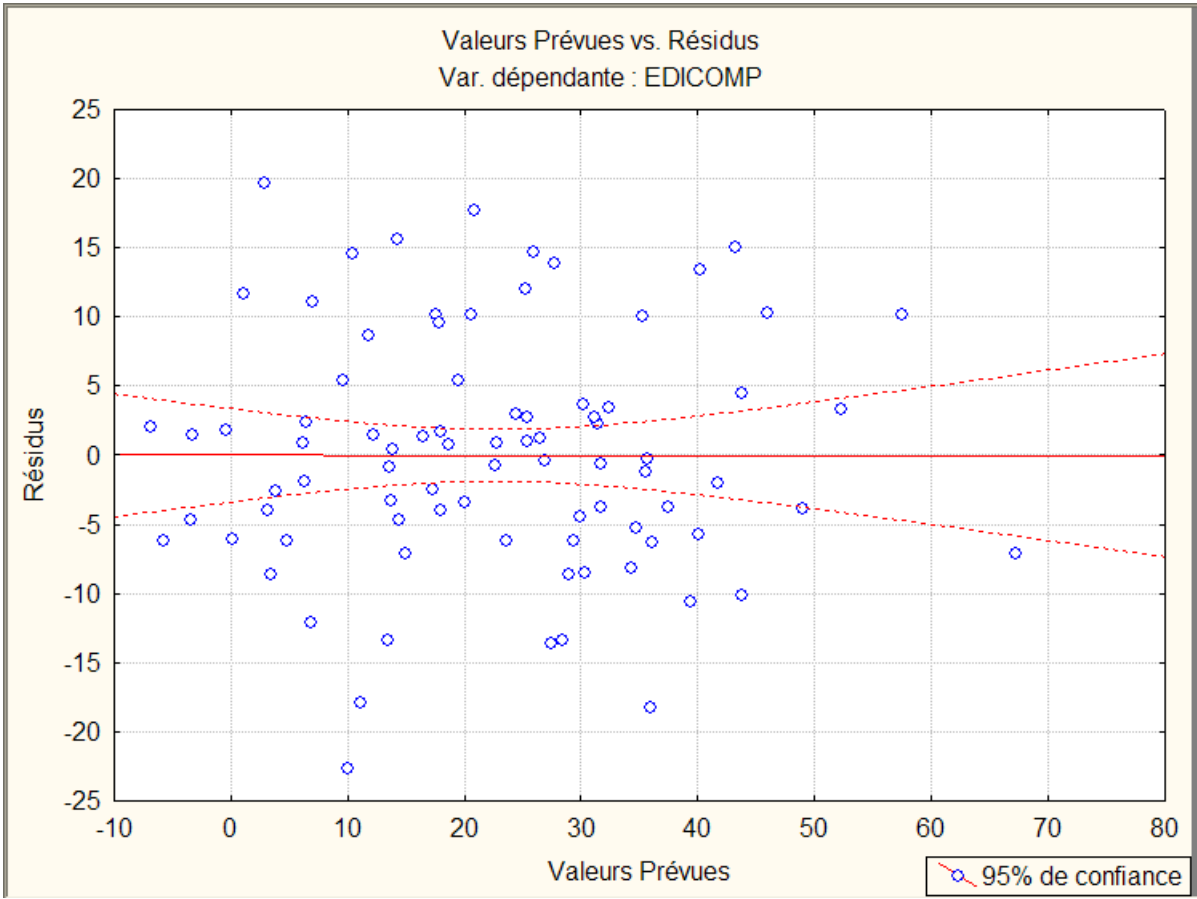
Analyse des Résidus : image-corporelle

Var dép. : EDICOMP R Multiple : ,87026299 F = 29,26212
 R² : ,75735768 dl = 8,75
 Nombre d'obs. : 84 R² ajusté : ,73147583 p = 0,000000
 Erreur-type de l'estim. : 9,068125178
 Ord. Orig : 69,242496885 Err-Type : 18,39961 t(75) = 3,7633 p < ,0003

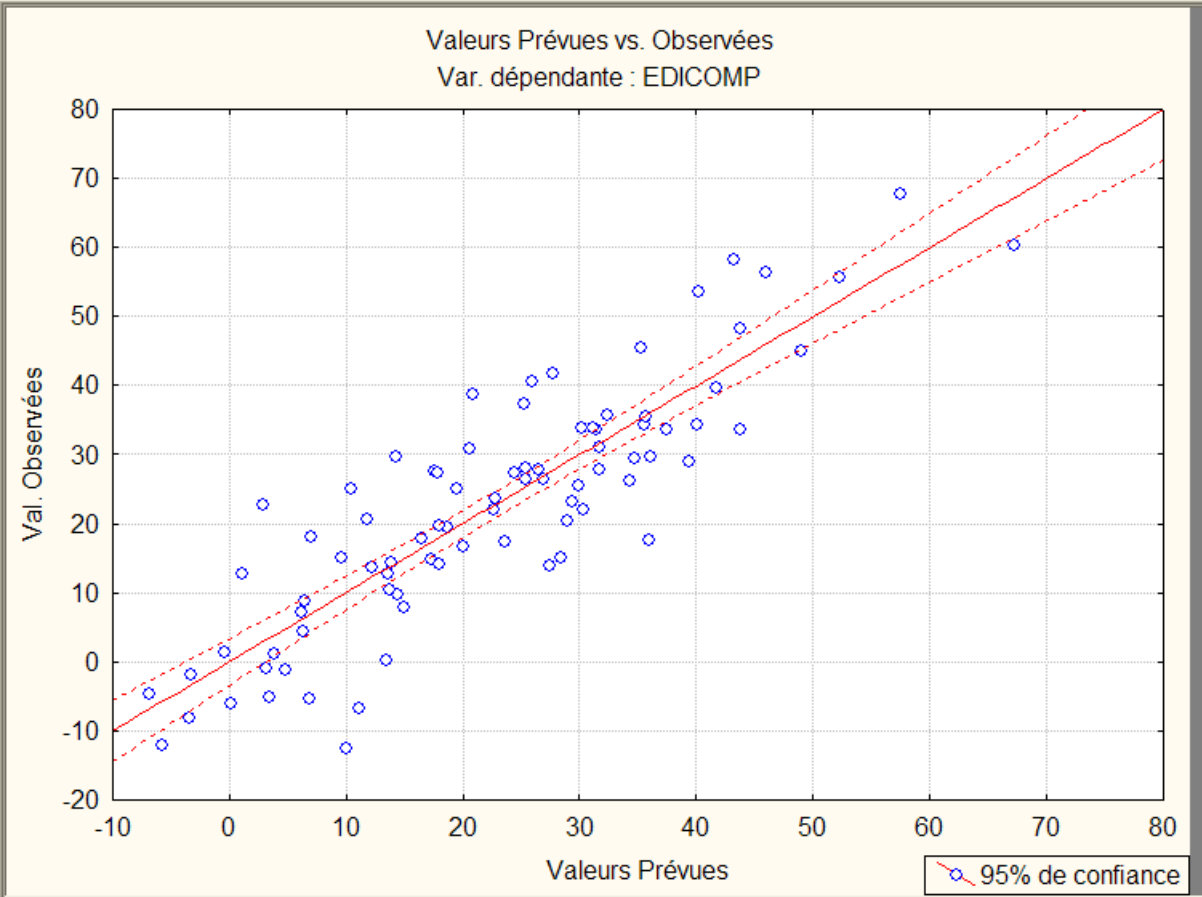
Base | Avancé | Résidus | Prévués | Nuages | Tracés de probas | Atypiques | Enregistrer

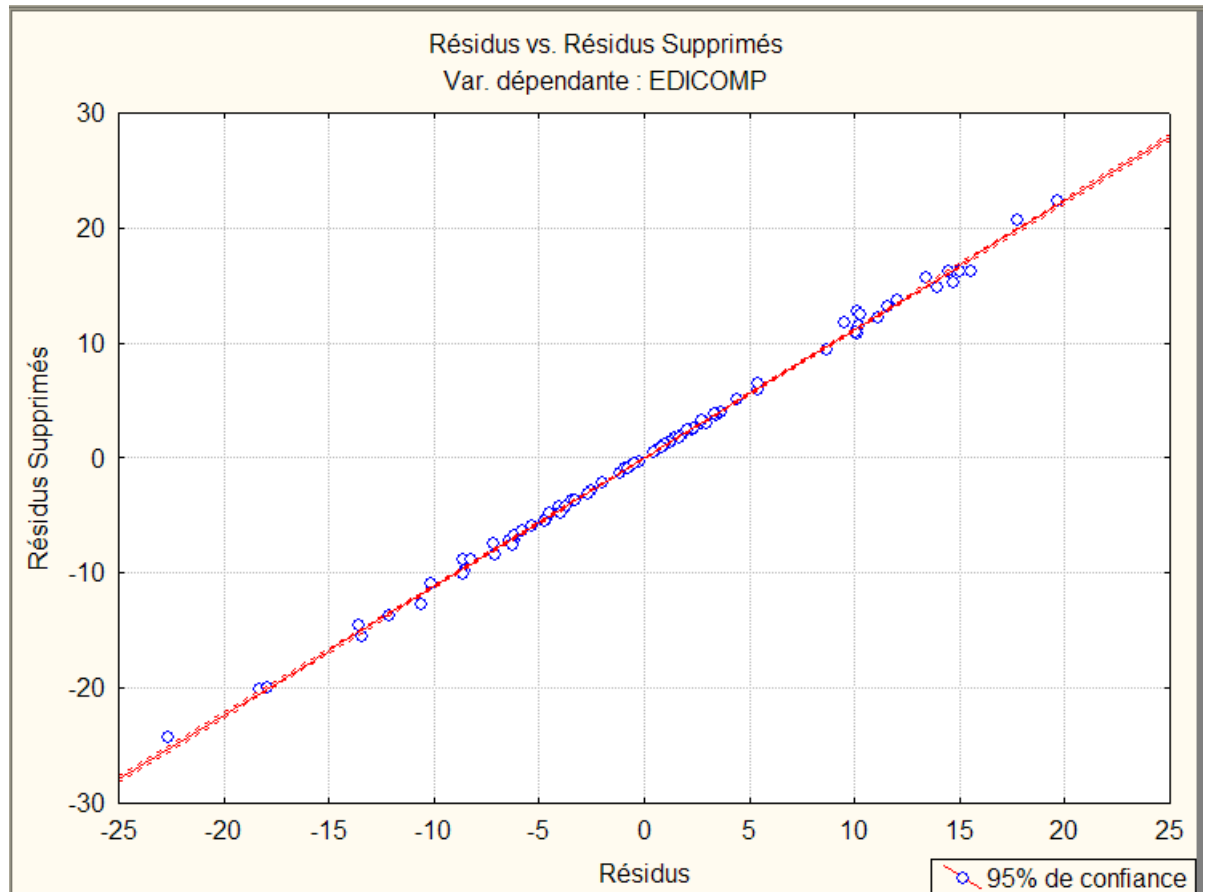
Valeurs prévues vs. résidus Valeurs observées vs. résidus carrés
 Valeurs prévues vs. résidus carrés Résidus vs. résidus supprimés
 Valeurs prévues vs. observées Corrélation bivariée
 Valeurs observées vs. résidus Tracé des résidus partiels

Synthèse
 Annuler
 Options



Utiliser le menu *Statistiques / Régression multiple / Résidus, hypothèses, prévision / Analyse des résidus / Nuages*





Utiliser le menu *Statistiques / Régression multiple / Résidus, hypothèses, prévision / Analyse des résidus / Atypique*

Analyse des Résidus : image-corporelle

Var dép. : EDICOMP	R Multiple : ,87026299	F = 29,26212
	R ² : ,75735768	dl = 8,75
Nbre d'obs. : 84	R ² ajusté : ,73147583	p = 0,000000
	Erreur-type de l'estim. : 9,068125178	
Ord. Orig : 69,242496885	Err-Type : 18,39961	t(75) = 3,7633 p < ,0003

Base | Avancé | Résidus | Prévues | Nuages | Tracés de probas | Atypiques | Enregistrer

Tracé des points atypiques (par obs.)

Type de point atypique

- Résidus standard (> 2 * sigma)
- Valeurs prévues standard
- Résidus standard
- Distances de Mahalanobis
- Résidus supprimés
- Distances de Cook

Tracer les 100 obs. les plus extrêmes :

Annuler

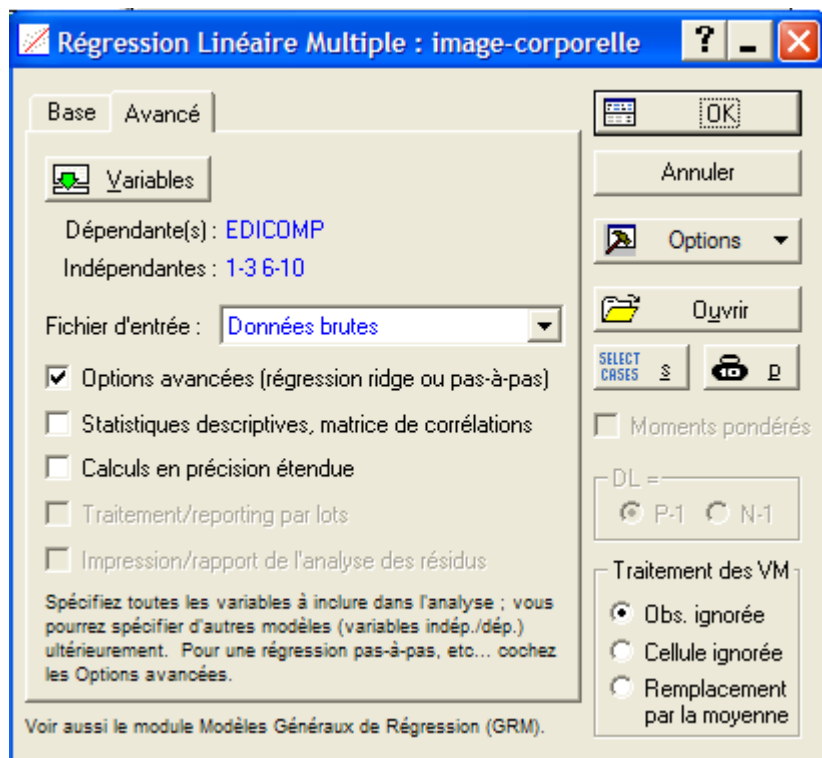
Options ▾

Résidus Standard : EDICOMP (image-corporelle)													
Pts Atypiques													
Obs.	Résidus Standard					Valeur Observée	Valeur Prévüe	Résidus	Standard Val.Prév	Standard Résidus	Err.Type Val.Prév		
	-5.	-4.	-3.	±2.	3.	4.	5.						
33	.	.	.	*	.	.	.	-12,6700	9,97345	-22,6435	-0,82254	-2,49704	2,373859
41	.	.	.	*	.	.	.	17,6800	35,95903	-18,2790	0,88376	-2,01574	2,760483
59	.	.	.	*	.	.	.	22,6100	2,94928	19,6607	-1,28378	2,16811	3,176330
Minimum	.	.	.	*	.	.	.	-12,6700	2,94928	-22,6435	-1,28378	-2,49704	2,373859
Maximum	.	.	.	*	.	.	.	22,6100	35,95903	19,6607	0,88376	2,16811	3,176330
Moyenne	.	.	.	*	.	.	.	9,2067	16,29392	-7,0873	-0,40752	-0,78156	2,770224
Médiane	.	.	.	*	.	.	.	17,6800	9,97345	-18,2790	-0,82254	-2,01574	2,760483

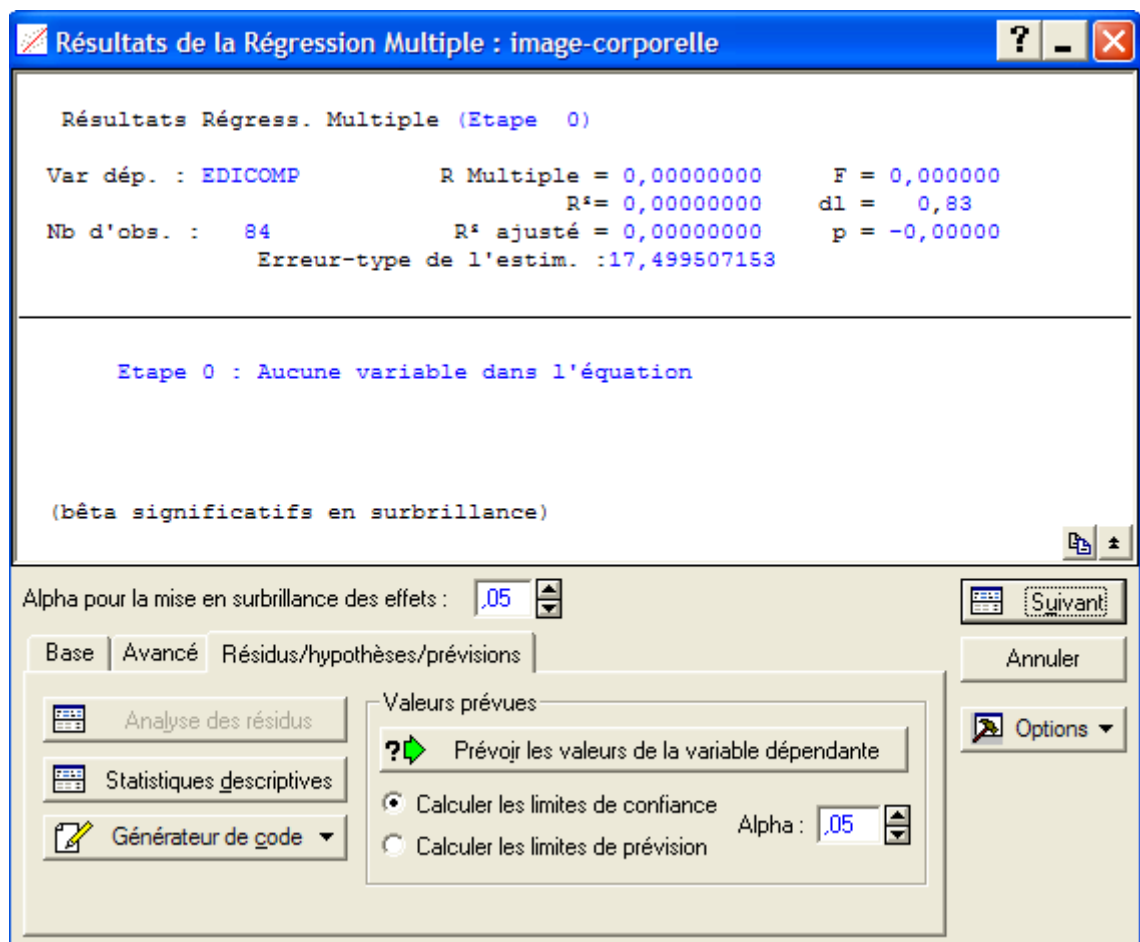
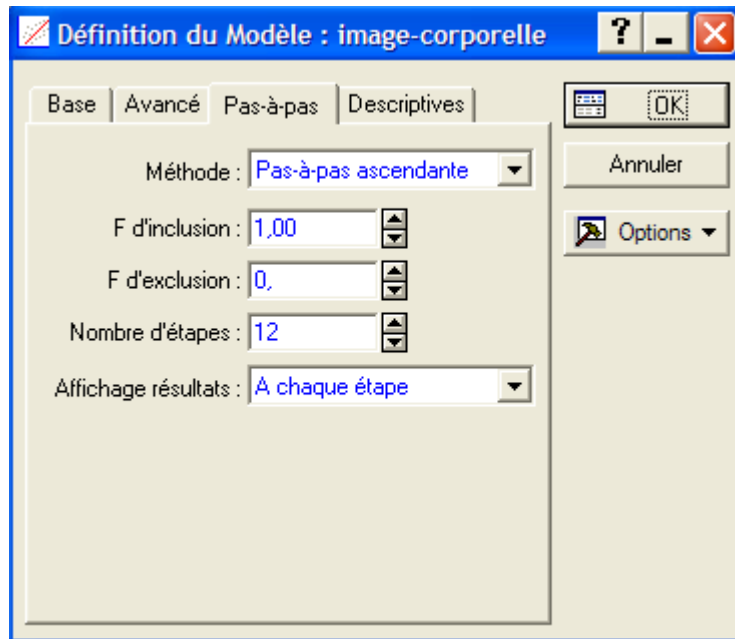
vi. Suppression de variables : méthode pas à pas.

Donner le modèle retenu par la méthode ascendante et vérifier sa qualité.

Utiliser le menu *Statistiques / Régression multiple*) cocher options avancées : (régression ridge ou pas à pas)



Cliquer sur OK. Dans le menu pas à pas, sélectionner



Résultats de la Régression Multiple : image-corporelle

Résultats Régress. Multiple (Etape 1)

Var dép. : EDICOMP	R Multiple = ,71122922	F = 83,94051
	R² = ,50584700	dl = 1,82
Nb d'obs. : 84	R² ajusté = ,49982074	p = ,000000
	Erreur-type de l'estim. : 12,376238121	
Ord.Orig : 4,350881283	Err.-Type: 2,397418	t(82) = 1,8148 p = ,0732

BDI bêta=,711

(bêta significatifs en surbrillance)

Alpha pour la mise en surbrillance des effets : ,05

Base | Avancé | Résidus/hypothèses/prévisions

Analyse des résidus

Statistiques descriptives

Générateur de code

Valeurs prévues

Prévoir les valeurs de la variable dépendante

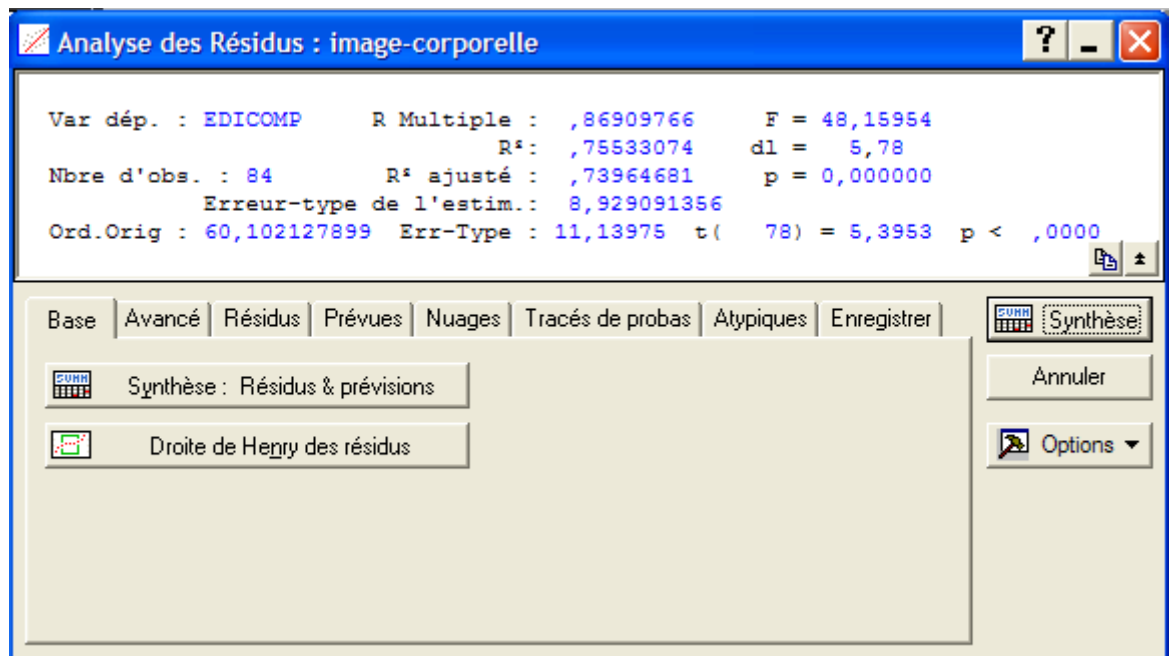
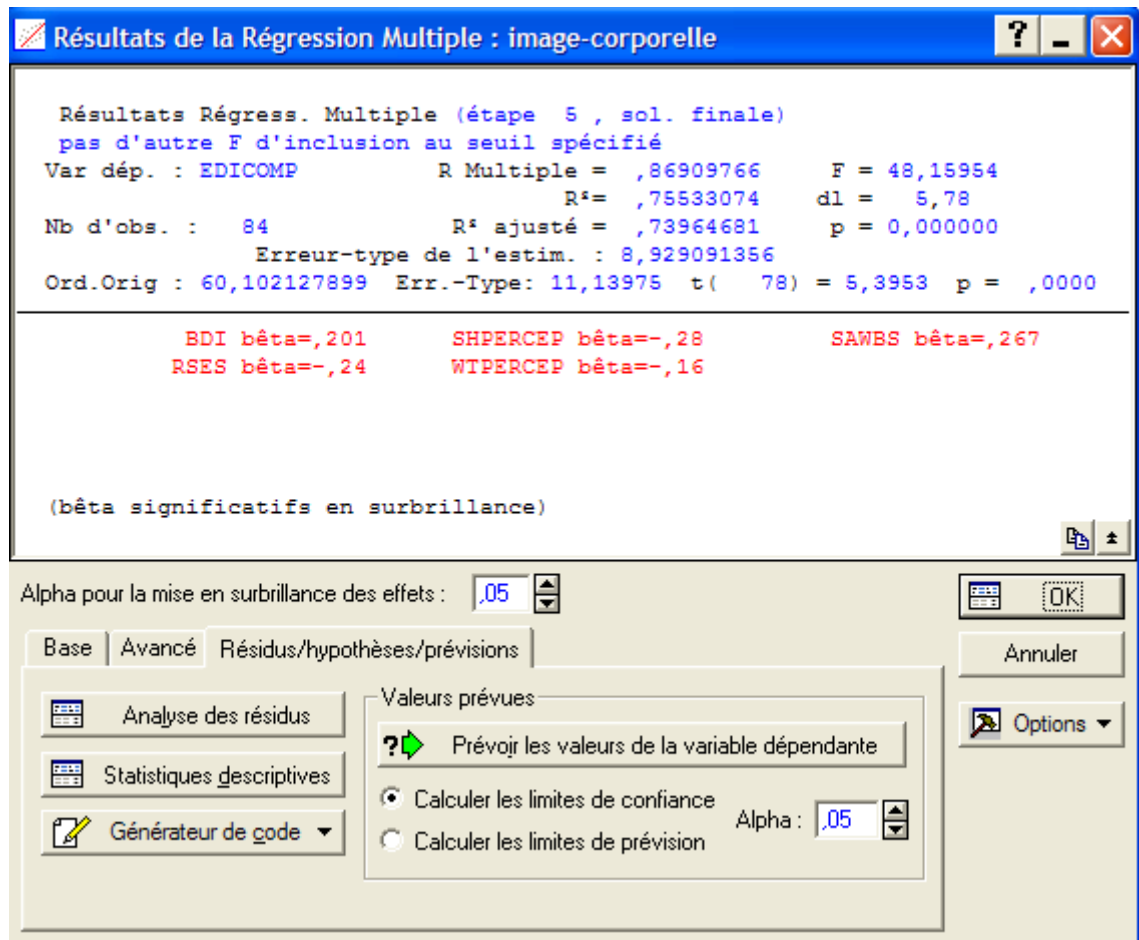
Calculer les limites de confiance Alpha : ,05

Calculer les limites de prévision

Suivant

Annuler

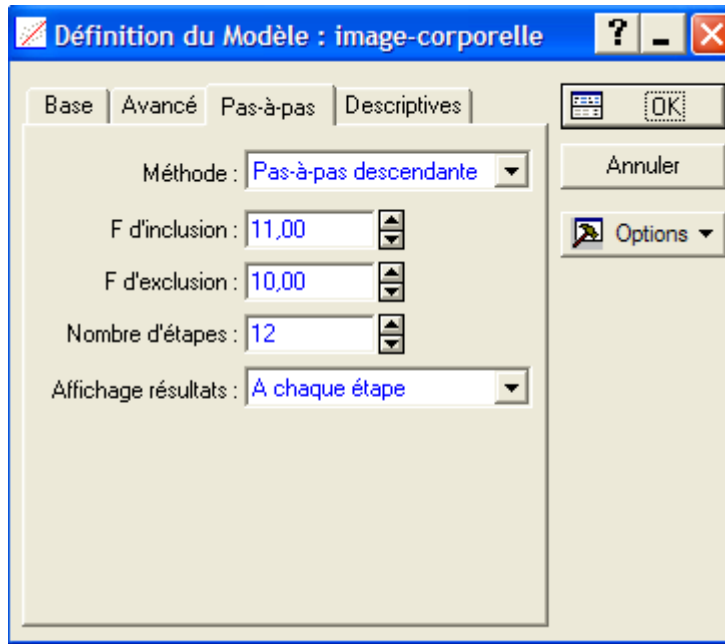
Options



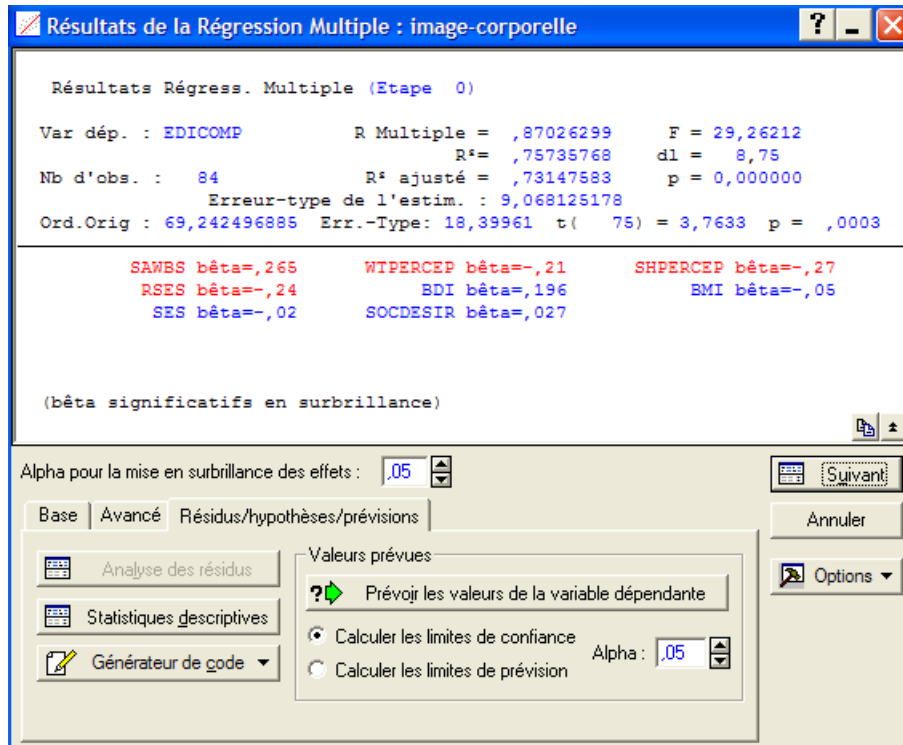
On garde le modèle avec 5 VI : 1-5

Donner le modèle retenu par la méthode descendante et vérifier sa qualité.

Utiliser le menu Statistiques / Régression multiple) cocher options avancées : (régression ridge ou pas à pas)



Cliquer sur OK. Dans le menu pas à pas, sélectionner



Résultats de la Régression Multiple : image-corporelle

Résultats Régress. Multiple (Etape 1)

Var dép. : EDICOMP R Multiple = ,87012134 F = 33,84290
 R² = ,75711115 dl = 7,76
 Nb d'obs. : 84 R² ajusté = ,73473981 p = 0,000000
 Erreur-type de l'estim. : 9,012844051
 Ord.Orig : 68,074160014 Err.-Type: 17,79706 t(76) = 3,8250 p = ,0003

SAWBS bêta=,266 WIPERCEP bêta=-,21 SHPERCEP bêta=-,27
 RSES bêta=-,24 BDI bêta=,201 BMI bêta=-,05
 SOCDESIR bêta=,024

(bêta significatifs en surbrillance)

Alpha pour la mise en surbrillance des effets : .05

Base Avancé Résidus/hypothèses/prévisions

Analyse des résidus
 Statistiques descriptives
 Générateur de code

Valeurs prévues
 Prévoir les valeurs de la variable dépendante
 Calculer les limites de confiance
 Calculer les limites de prévision Alpha : .05

Suivant
 Annuler
 Options

Résultats de la Régression Multiple : image-corporelle

Résultats Régress. Multiple (Etape 2)

Var dép. : EDICOMP R Multiple = ,86986329 F = 39,90541
 R² = ,75666214 dl = 6,77
 Nb d'obs. : 84 R² ajusté = ,73770075 p = 0,000000
 Erreur-type de l'estim. : 8,962400472
 Ord.Orig : 68,893506508 Err.-Type: 17,56345 t(77) = 3,9225 p = ,0002

SAWBS bêta=,266 WIPERCEP bêta=-,20 SHPERCEP bêta=-,28
 RSES bêta=-,24 BDI bêta=,192 BMI bêta=-,05

(bêta significatifs en surbrillance)

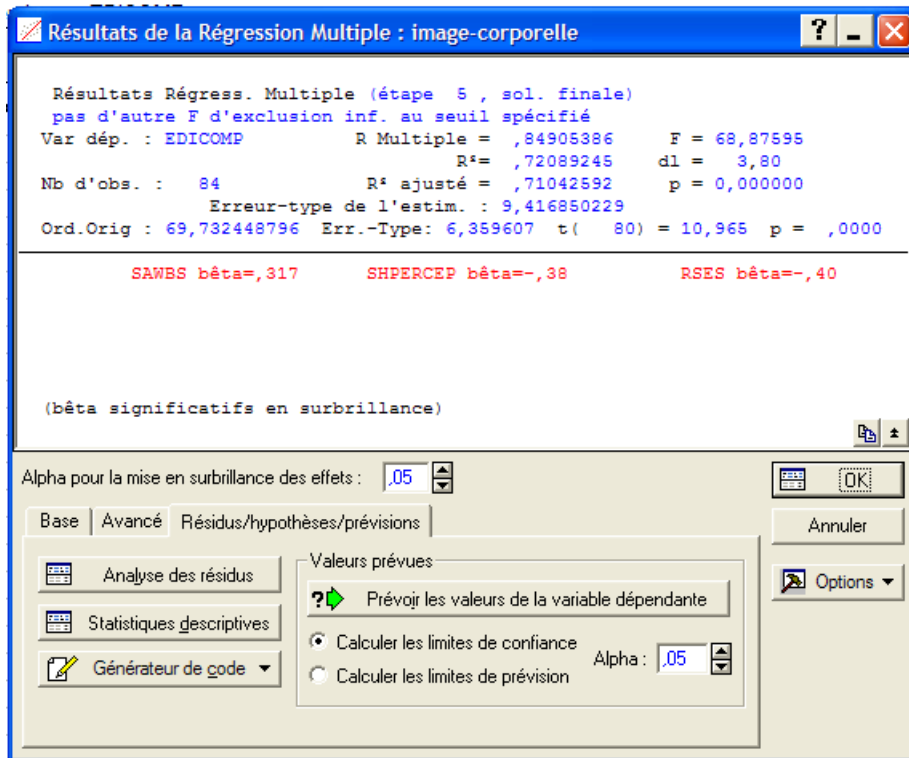
Alpha pour la mise en surbrillance des effets : .05

Base Avancé Résidus/hypothèses/prévisions

Analyse des résidus
 Statistiques descriptives
 Générateur de code

Valeurs prévues
 Prévoir les valeurs de la variable dépendante
 Calculer les limites de confiance
 Calculer les limites de prévision Alpha : .05

Suivant
 Annuler
 Options



On garde le modèle avec 3 VI : 1,2,4