الارتباط & الانحدار
البسيط ، الجزئي و المتعدد

د. سبأ محمد علوان

أستاذ مساعد قسم الاحصاء وبحوث العمليات

جامعة الملك سعود

2014

مقدمة

على الرغم من أهمية استخدام الإحصاء في تحليل البيانات للبحوث والرسائل العلمية فإن هناك بعض المشكلات التي قد تنشأ عن هذا الاستخدام والتي تكون غالباً بدون قصد وبسبب عدم التخصص في الإحصاء لدى البعض وبالتالي عدم الإلمام ببعض الجوانب العلمية الإحصائية الدقيقة .

وفي هذه الورشة نحاول معرفة التحليل الاحصائي المناسب من خلال طريقة سهلة وهي الأمثلة المباشرة والتي تحاكي بعض الحالات التي قد يكون فيها الباحث في مساره البحثي، مع ملاحظة اننا نركز على البحوث العلمية وبما هو متاح من الوقت لهذه الورشة .

SPSS = Statistical Package for Social Sciences

الارتباط والانحدار

العلاقات بين المتغيرات
"معاملات الارتباط"

- فالارتباط يقيس درجة أو قوة العلاقة بين المتغيرا ت بقيمة رقمية محصورة بين -١و+١.
- وهناك طرق متعددة لحساب درجة الارتباط بين متغيرين:
  - على سبيل المثال معامل ارتباط بيرسون إذا كان المتغيران كميين ،
  - معامل ارتباط سبيرمان إذا كان المتغيران كميين أو ترتيبين ، أو أحدهما كمي والآخر ترتيبي).

التنبؤ
"الانحدار"

- الانحدار يعبّر عن هذه العلاقة بمعادلة رياضية خطية (أو غير خطية) تفيد في التنبؤ بقيم المتغير التابع بافتراض قيم معينة للمتغير (أو المتغيرات) المستقلة.
- الانحدار يفيد أيضا في تحديد مدى تأثير المتغير المستقل على المتغير التابع .

معاملات الارتباط

معلمي

بيرسون

١ـ كلا المتغيرين كمي

٢ـ و كلا المتغيرين ينحدران من مجتمعات طبيعية

٣ـ و العلاقة المراد تقديرها هي علاقة خطية

سبيرمان الرتبي

١ـمتغيرات رتبيه

٢ـ أو متغير كمي+ متغير اسمي في مستويين

٣ـ أو المتغيرات من مجتمعات لا يعرف توزيعها

يفضل استخدامه اذا كانت حجم العينة أكبر من ١٠

لا معلمية

كيندال

١ـ متغيرين كميين أو متغيرين وصفيين ترتيبين

٢ـ أو المتغيرات من مجتمعات لا يعرف توزيعها

يفضل استخدامه اذا كانت حجم العينة أقل من ١٠

بوينت بايسيريال

يستخدم لقياس الارتباط بين متغير كمي X و متغير اسمي Yمستويين (نعم – لا) أو (ذكر – أنثى) و غيرها.

اشارة معامل الارتباط ليس لها معنى في حالة المتغيرات الوصفية فتقاس قوة العلاقة و ليس اتجاهها.

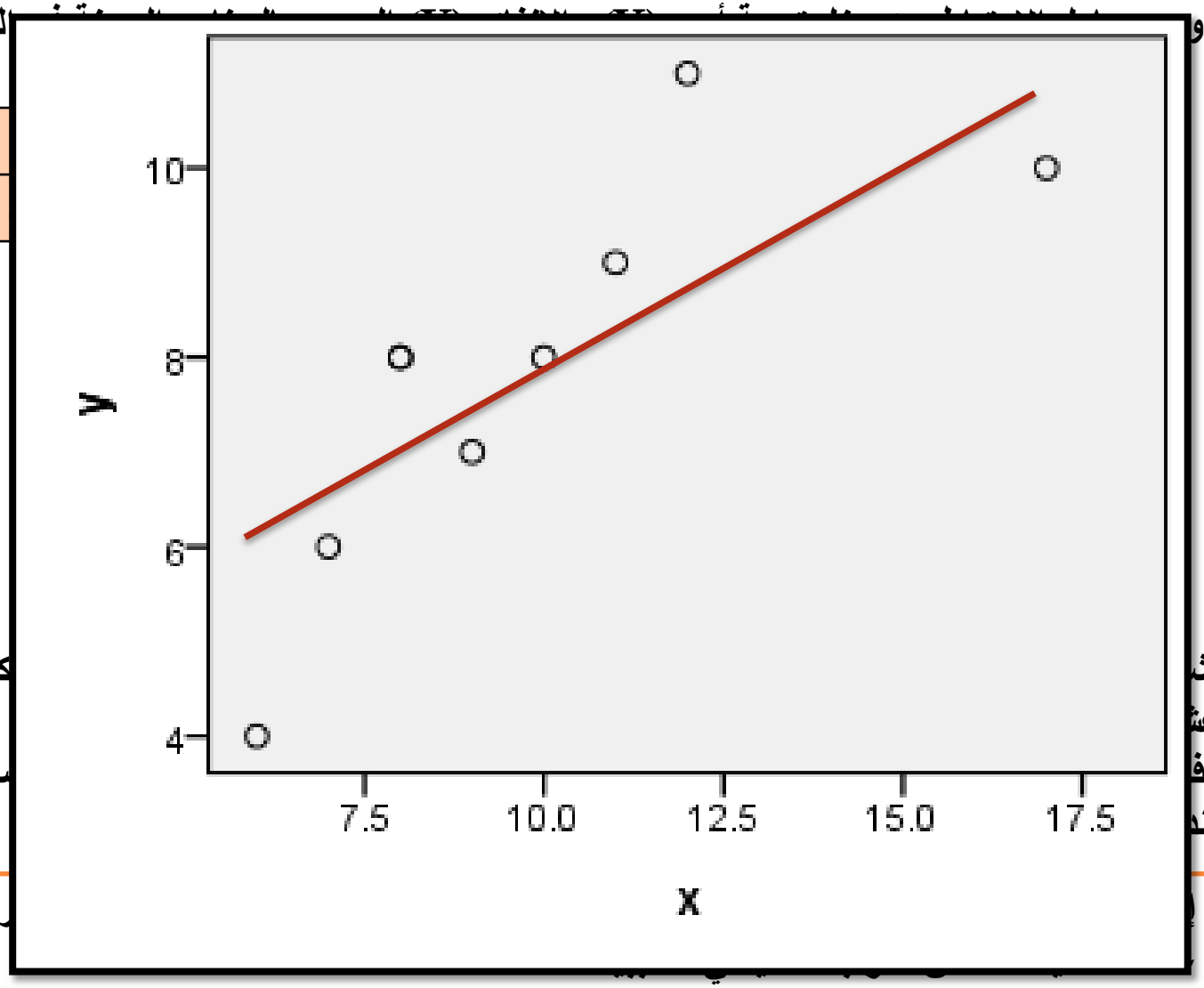
- وقد صنف بعض الإحصائيين درجات لقوة العلاقة يمكن تمثيلها على الشكل التالي:

معامل بيرسون للارتباط الخطي

مثال 1 Pearson Linear Correlation Coefficient

: أو الجدول الآتي:

X	
Y	



كون العينة

سبيرمان أو

رين بل لا توجد

إختبار الطبيعية

1- Analyze→ Descriptive Statistics → Explore
2- select the variable to the "Dependent List
3- Click "Plots" , then select Normality plot with test
4- Click Continue then "Ok"

فرضية العدم :
البيانات تتبع التوزيع الطبيعي
الفرضية البديلة :
البيانات لا تتبع التوزيع الطبيعي

Tests of Normality

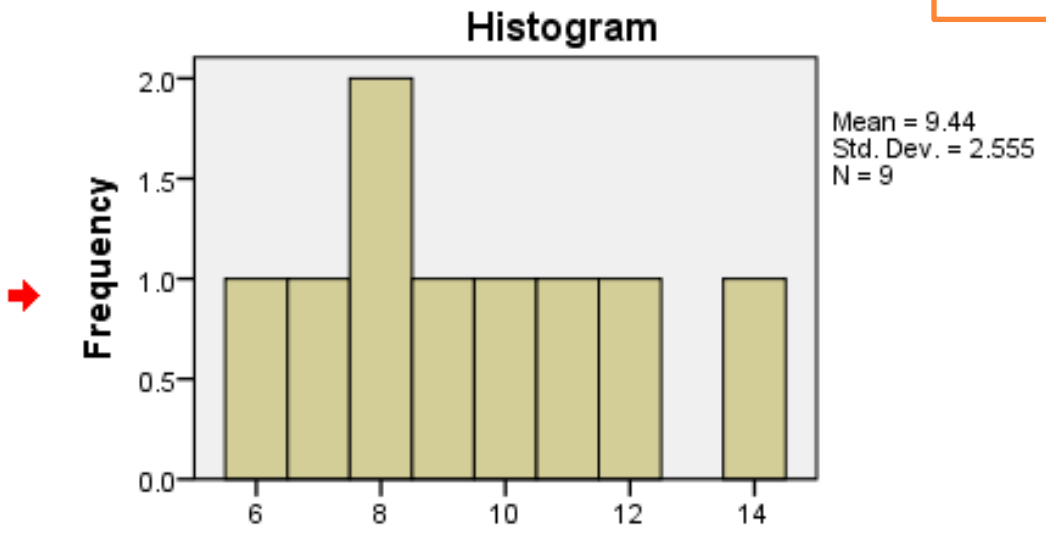
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
X	.159	9	.200*	.970	9	.890
Y	.188	9	.200*	.969	9	.885

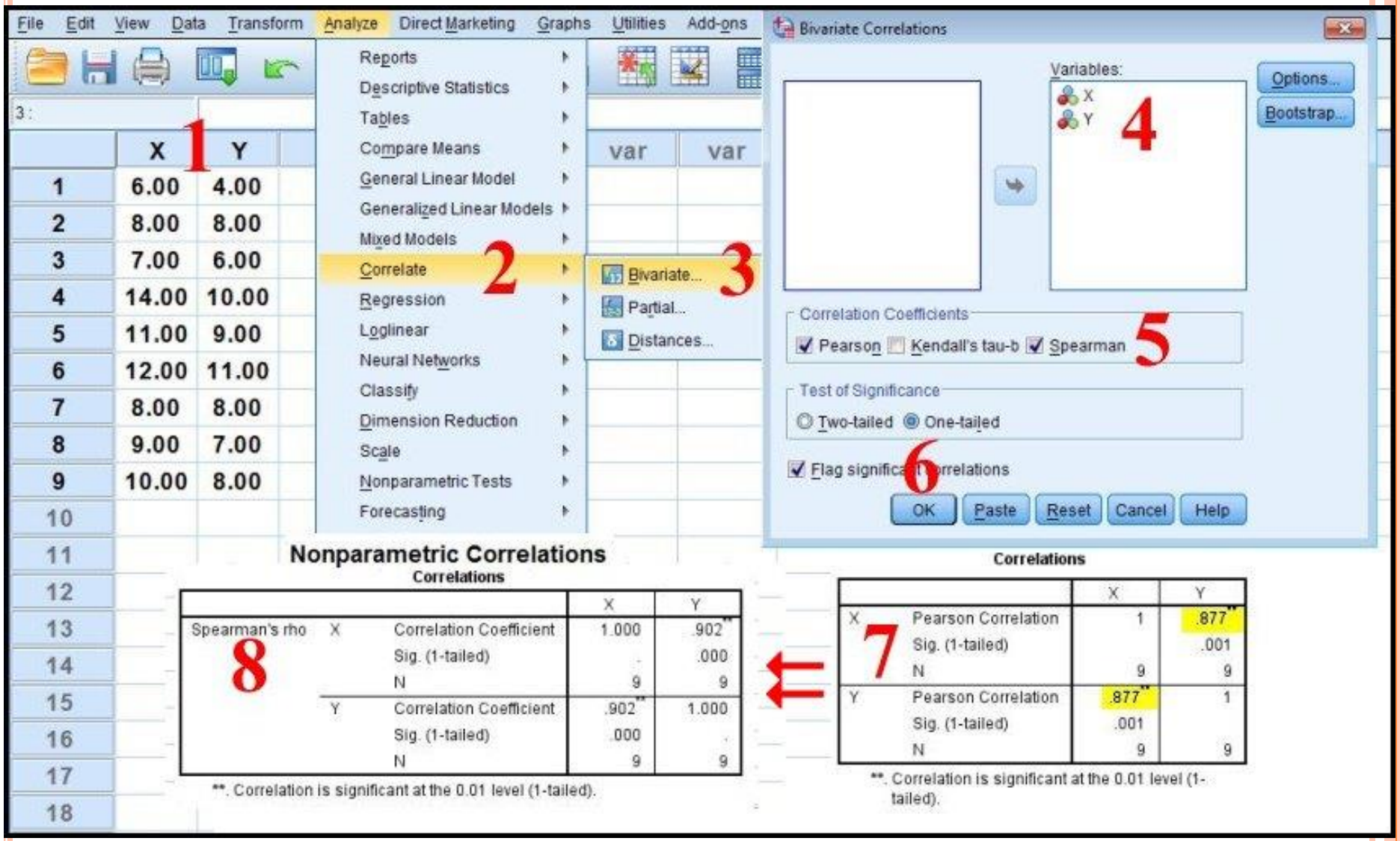
*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

وهي أكبر من 0.05
إذن نقبل الفرضية أن البيانات تتوزع طبيعيا

X





الاختبار هنا: لا توجد علاقة أي $H_o : \rho = 0$

توجد علاقة أي $H_1 : \rho > 0$ One tails test)

من الناتج المبينة نجد أن $r = 0.877$ بمستوى معنوية 0.01 فنرفض الفرضية الصفرية بعدم وجود علاقة خطية ونقبل الفرضية البديلة بوجود علاقة خطية ذات مستوى دلالة 0.01

مثال: 3

يبين الجدول التالي نسبة التغير في أسعار الوقود و في أسعار السلع و الخدمات الأخرى لمدة ١١ سنة
أوجد معامل الارتباط بين المتغيرين

السنة	٢٠٠٠	٢٠٠١	٢٠٠٢	٢٠٠٣	٢٠٠٤	٢٠٠٥	٢٠٠٦	٢٠٠٧	٢٠٠٨	٢٠٠٩	٢٠١٠
الوقود	٥١,٧	٥٣,٧	٥٣,٨	٥٥,٧	٧٩,٢	٨٩,٦	٩٧,٦	١٠٠	١٠٦,٦	١٥٣,٣	٢٠٦,٢
السلع	٦٤	٦٦,٨	٦٩	٧٣,٣	٨١,٤	٨٨,٨	٩٣,٩	١٠٠	١٠٧,٧	١١٩,٨	١٣٦

حيث أن العينة صغيرة ولانعلم ان كان المجتمع طبيعيا ام لا...نختبر إذا ما كانت البيانات لها توزيع طبيعي باساخدام اختبار Shapiro

١.فرضية العدم : البيانات لها توزيع طبيعي

٢-الفرضية البديلة البيانات لا تتبع التوزيع الطبيعي

نختبر إذا ما كانت البيانات لها توزيع طبيعي باساخدام اختبار Shapiro

فرضية العدم : البيانات لها توزيع طبيعي
الفرضية البديلة: البيانات لا تتبع التوزيع الطبيعي

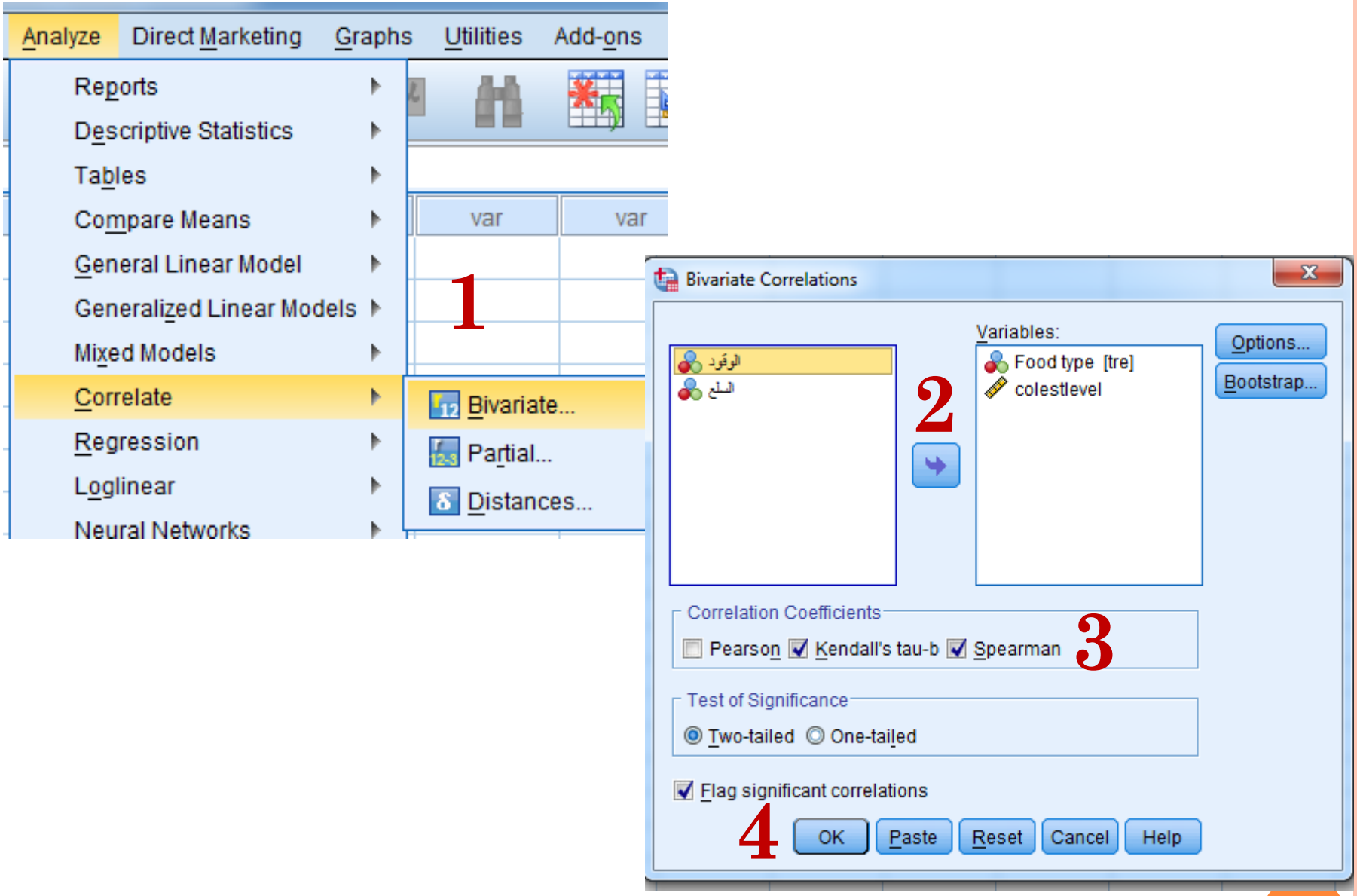
Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
نسبة التغير في الوقود	.224	11	.127	.840	11	.032
نسبة التغير في السلع و الخدمات	.223	11	.134	.745	11	.002

a. Lilliefors Significance Correction

نجد أن قيمة مستوى الدلالى في المتغيرين أقل من ٠,٠٥ لذلك نرفض فرضية العدم و نقبل الفرضية البديلة

(أي أن البيانات **لا** تتبع التوزيع الطبيعي ، عند مستوى دلالة 0.05
لذلك
نستخدم اختبار **لا معلمي**


Analyze
Direct Marketing
Graphs
Utilities
Add-ons
Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Bivariate...
Partial...
Distances...
var
var
1
Bivariate Correlations
الوقود
السلع
Variables:
Food type [tre]
colestlevel
2
Options...
Bootstrap...
Correlation Coefficients
Pearson
Kendall's tau-b
Spearman
3
Test of Significance
Two-tailed
One-tailed
Flag significant correlations
4
OK
Paste
Reset
Cancel
Help

Correlations

			نسبة التغير في الوقود	نسبة التغير في السلع و الخدمات
Kendall's tau_b	نسبة التغير في الوقود	Correlation Coefficient	1.000	.891**
		Sig. (2-tailed)	.	.000
		N	11	11
	نسبة التغير في السلع و الخدمات	Correlation Coefficient	.891**	1.000
		Sig. (2-tailed)	.000	.
		N	11	11
Spearman's rho	نسبة التغير في الوقود	Correlation Coefficient	1.000	.964**
		Sig. (2-tailed)	.	.000
		N	11	11
	نسبة التغير في السلع و الخدمات	Correlation Coefficient	.964**	1.000
		Sig. (2-tailed)	.000	.
		N	11	11

**. Correlation is significant at the 0.01 level (2-tailed).

Spearman
لأن حجم العينة أكبر من ١٠

بتطبيق معامل كندال و معامل سبيرمان للرتب نجد أن :
قيمة معامل كندال = ٠,٨٩١ أي وجود علاقة طردية قوية بين المتغيرين و هو دال احصائياً في مستوى ٠,٠١
قيمة معامل سبيرمان للرتب = ٠,٩٦٤ أي وجود علاقة طردية قوية جداً بين المتغيرين عند مستوى دلالة ٠,٠١

مثال 4

إدرس وجود علاقة بين تقدير الطالبة في الإحصاء و تقديرها في الرياضيات ، اخترنا خمس طالبات و كانت تقديراتهم كالتالي:

تقدير الإحصاء X	A	C	D	F	A
تقدير الرياضيات y	B	C	B	D	A

هل يمكن حساب معامل بيرسون

لا يمكن لأن المتغيرات ليست كمية

هل يمكن حساب معامل بوينت

لا يمكن لأن المتغير الوصفي ليس على مستويين و المتغير الأخر ليس كمياً.

معامل سيبرمان لارتباط الرتب

Spearman Rank Correlation Coefficient

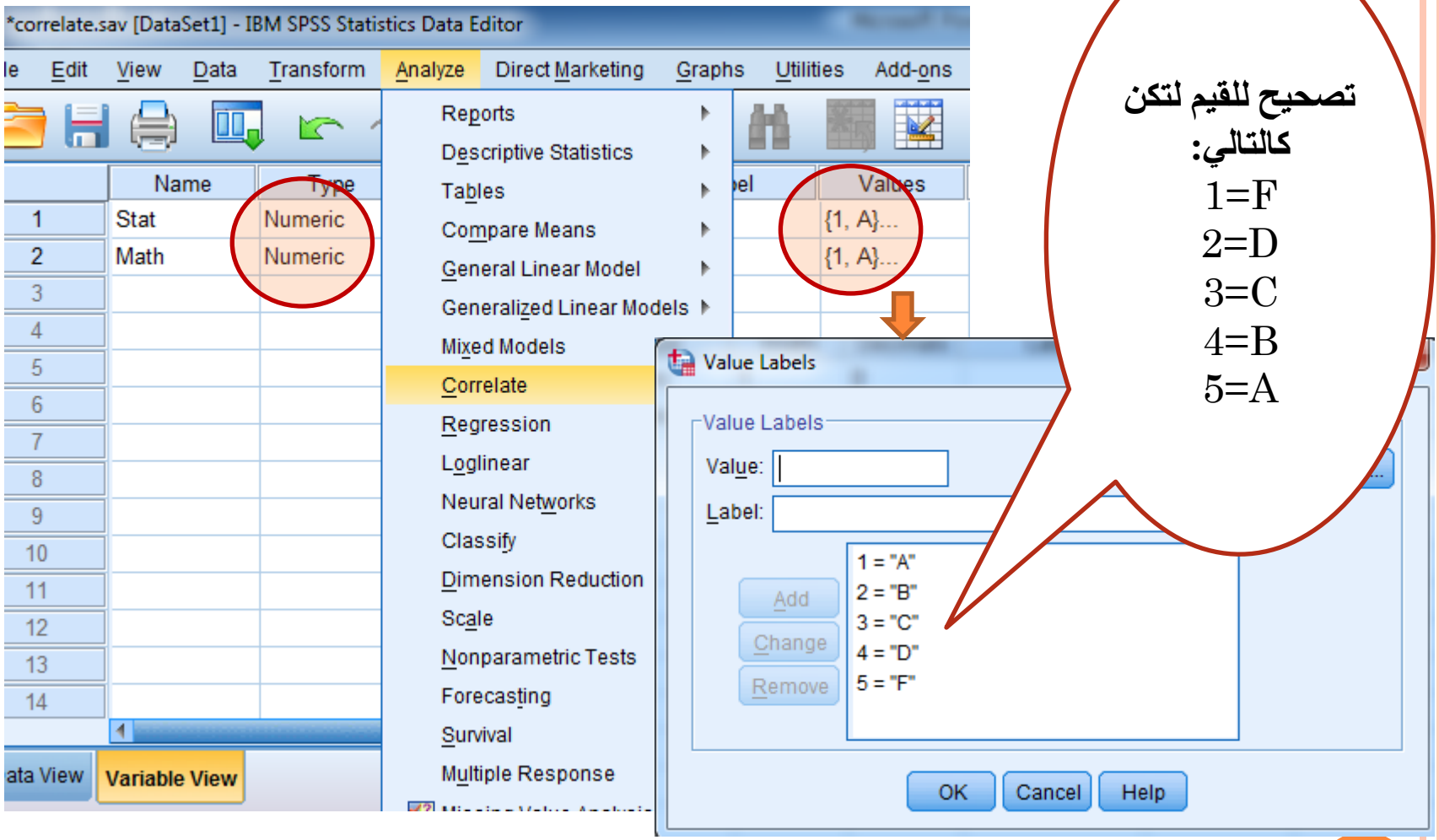
هل توجد علاقة ارتباط؟ ما نوعها ومدى قوتها؟

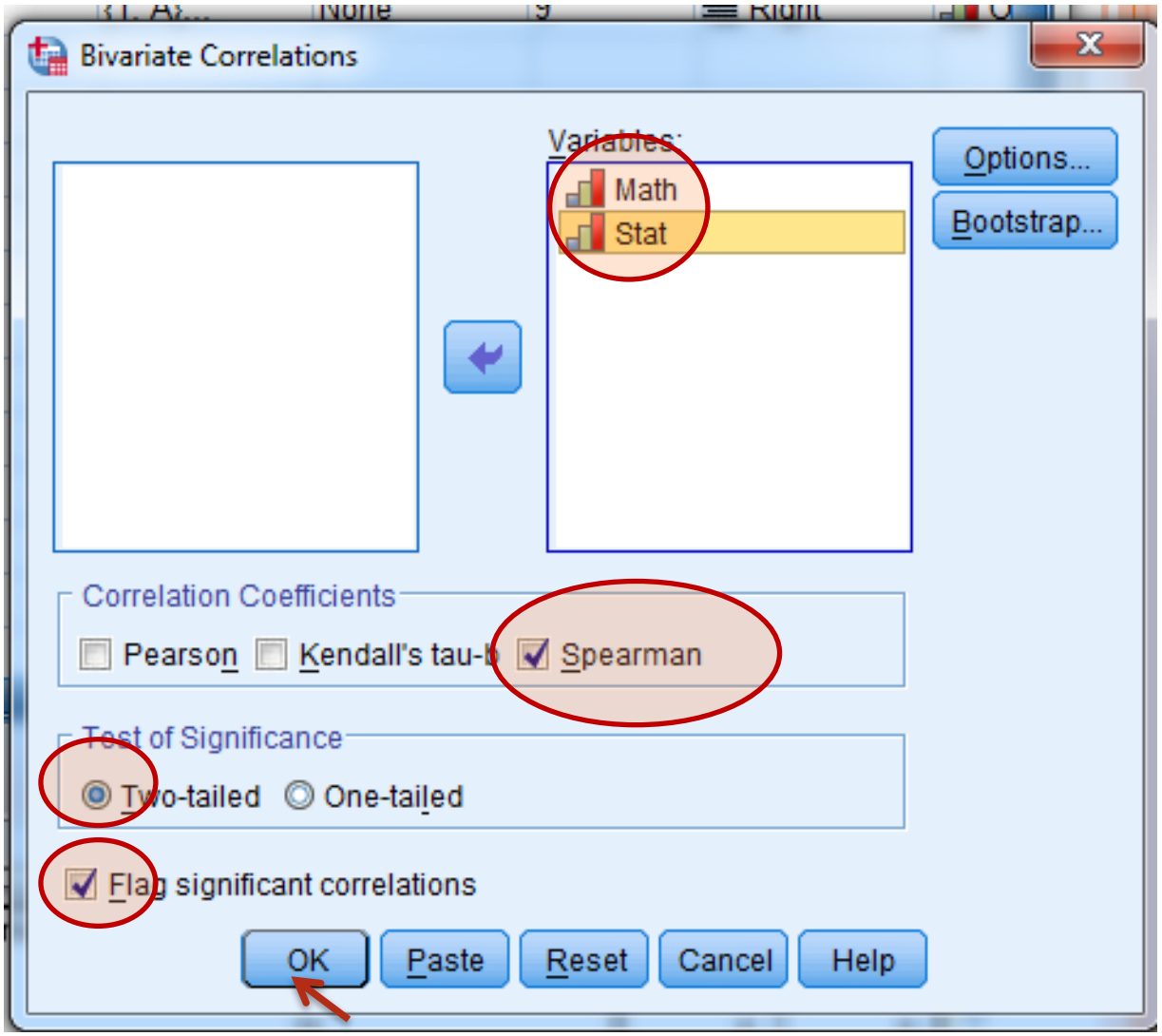
نستخدم معامل سيبرمان لارتباط الرتب
إذا كان قياس المتغيرين كليهما مقياس ترتيبي .

نعطي لكل تقدير رقم مقابل

F	D	C	B	A
1	2	3	4	5

Analyze----correlate------ Bivariate




Bivariate Correlations
Variables:
Math
Stat
Options...
Bootstrap...
Correlation Coefficients
Pearson
Kendall's tau-b
Spearman
Test of Significance
Two-tailed
One-tailed
Flag significant correlations
OK
Paste
Reset
Cancel
Help

Nonparametric Correlations

[DataSet1] C:\Users\hp\Documents\SPSS data\correlate.sav

Correlations

			Stat	Math
Spearman's rho	Stat	Correlation Coefficient	1.000	.763
		Sig. (2-tailed)	.	.133
		N	5	5
	Math	Correlation Coefficient	.763	1.000
		Sig. (2-tailed)	.133	.
		N	5	5

هل هذا التعبير صحيح؟

تصحيح

وعليه .. يوجد علاقة طردية قوية ذات دلالة احصائية عند مستوى دلالة 0.05 بين تقديرات الطلاب في مادة الإحصاء وتقديراتهم في مادة الرياضيات

معامل الا قتران (فاي)=

$$\sqrt{\frac{\text{كاي مربع}}{\text{العينة حجم}}}$$

مثال: 5

في دراسة على ٢٥ شخص لمعرفة العلاقة بين الجنس والتدخين فجمعت البيانات في الجدول التالي والمطلوب هل هناك ارتباط معنوي بين الجنس والتدخين.

	ذكر	أنثى
يدخن	10	2
لا يدخن	8	5

هل يمكن حساب معامل بيرسون

لا يمكن لأن المتغيرات ليست كمية

هل يمكن حساب معامل بوينت

لا يمكن لأنه لا يوجد متغير كمي

هل يمكن سبيرمان

لا يمكن لأن المتغيرات ليست ترتيبية.

معامل الاقتران (معامل فاي)= $\sqrt{\frac{\text{كاي معامل}}{\text{العينة حجم}}}$

- ويستخدم للبيانات المعبر عنها بالتكرار في مستويين للعلاقة بين متغيرين اسميين كل منهما ثنائي التقسيم.
- اشارة معامل فاي ليس لها معنى فهو يقيس قوة العلاقة دون اتجاهها.

	X_1	X_2	المجموع
Y_1	a	b	$a+b$
Y_2	c	d	$c+d$
المجموع	$a+c$	$b+d$	

$$r_{\emptyset} = \frac{a \times d - b \times C}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}$$

	أنثى	ذكر	المجموع
يدخن	2	10	12
لا يدخن	5	8	13
المجموع	7	18	25

نكمل الجدول بعملية الجمع بالصورة الآتية:

وبتطبيق القانون نجد أن:

R= 0.243

القيمة تدل على ضعف العلاقة بين الجنس والتدخين .

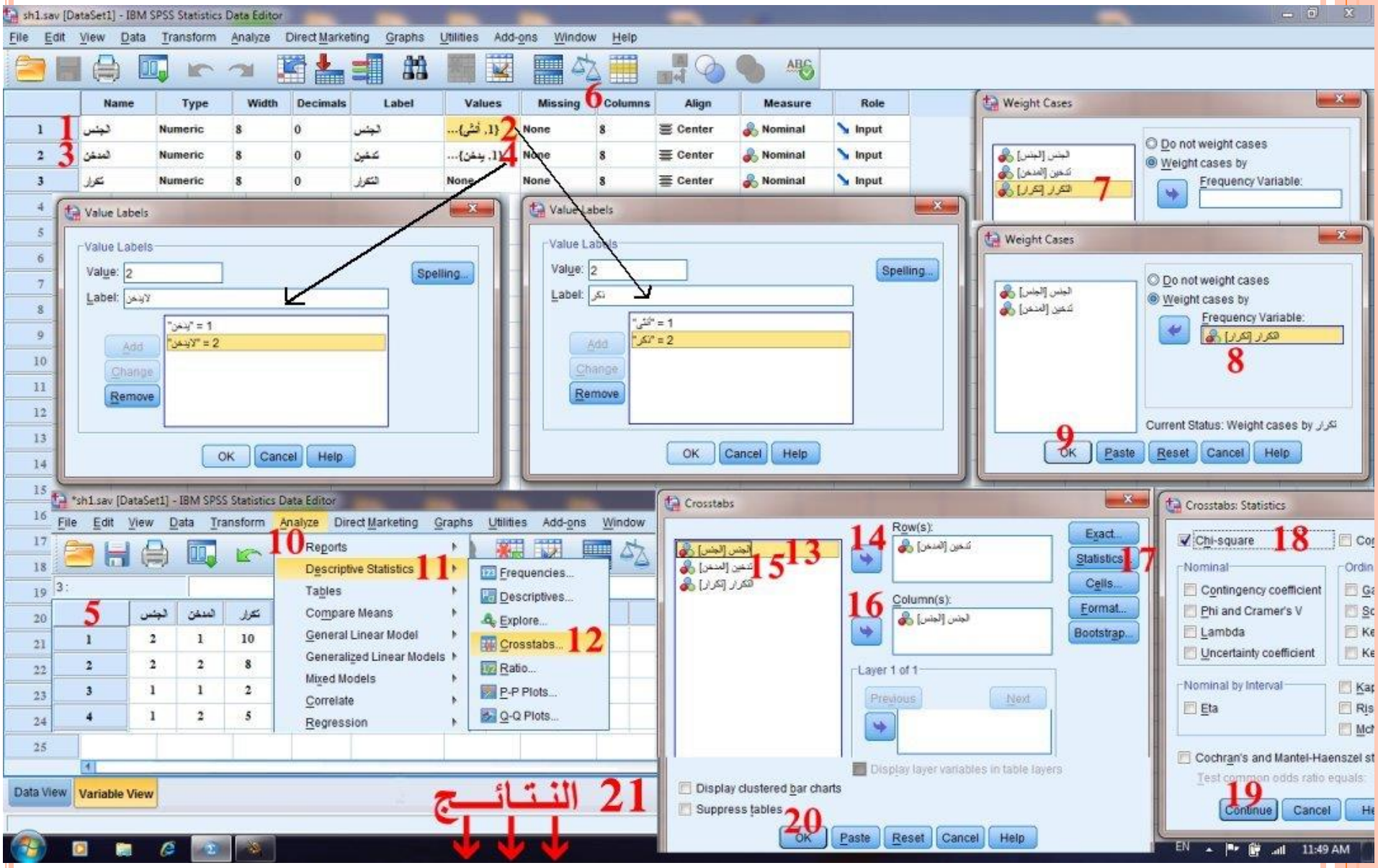
ملاحظة: عند تفسير نتيجة معامل الاقتران ننظر الى قوة العلاقة فقط (ضعيفة/متوسطة/قوية) وليس اتجاه العلاقة (طردية/عكسية) وهذا منطقي فالمتغيرات ليست بالكمية ولابالرتبية فكيف تزيد أو تنقص!

	x_i	x_i	Total
y_i	a	b	a + b
y_i	c	d	c + d
Total	a + c	b + d	a + b + c + d

$$r_\phi = \frac{bc - ad}{\sqrt{(a+c)(b+d)(c+d)(a+b)}}$$

كما ان العلاقة التالية صحيحة فلا دلالة للاشارة

خطوات ايجاد عامل ارتباط فاي باستخدام SPSS



Crosstabs

[DataSet1] C:\Users\shukri.AGU\Documents\sh1.sav

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
تدخين * الجنس	25	100.0%	0	.0%	25	100.0%

Crosstabulation تدخين * الجنس

Count

		الجنس		Total
		أنثى	ذكر	
تدخين	يدخن	2	10	12
	لايدخن	5	8	13
Total		7	18	25

معامل ارتباط فاي =

$$= \sqrt{\frac{1.470}{25}} = 0.243$$

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1.470[a]	1	.225		
Continuity Correction[b]	.588	1	.443		
Likelihood Ratio	1.511	1	.219		
Fisher's Exact Test				.378	.223
Linear-by-Linear Association	1.411	1	.235		
N of Valid Cases	25				

a. 2 cells (50.0%) have expected count less than 5. The minimum expected count is 3.36.
b. Computed only for a 2x2 table

مثال 6

أوجد قيمة معامل الارتباط بين النوع (ذكر/انثى) وبين الاصابة بمرض الاكتئاب وفقا للبيانات التالية:

الاكتئاب / النوع	مصاب	غير مصاب
ذكر	12	8
أنثى	4	6

الاكتئاب / النوع	مصاب	غير مصاب	المجموع
ذكر	12	8	20
أنثى	4	6	10
المجموع	16	14	30

$$r_\phi = \frac{ad-bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}} = \frac{12\times6-8\times4}{\sqrt{20\times10\times16\times14}} = \frac{40}{211.66} \approx 0.19$$

أن العلاقة بين النوع والاصابة بالاكتئاب ضعيفة.

لكن لابد من اختبار معنوية معامل الارتبط

معاملات الارتباط حسب نوع المتغير

	كمي	اسمي	رتبي
كمي	(معامل بيرسون)	(معامل الارتباط التسليلي، معامل ايتا، معامل ايبسلون).	(معامل الارتباط المتسليل المتعدد).
اسمي		معامل كرايمر ، معامل فاي، معامل التوافق، معامل لامدا)	**معامل فيتا، معامل الثنائي للرتب)**
رتبي			معامل سبيرمان، معامل جاما، معامل كندال).

تصحيح

معاملات الارتباط المتعددة:

7

مثال : إذا كان طلاب فصل دراسي في أحد المدارس ويضم ٣٠ طالباً يملكون مستوى تحصيلي إيجابي في أحد فروع الرياضيات فإننا نفترض أنه لا بد من أنهم يملكون مستوى تحصيلي إيجابي في الفروع الأخرى في مادة الرياضيات المقررة والتي تضم الفروع الأربعة وهي الجبر (Algebra) و الهندسة التحليلية (Analysis Geometry) و الهندسة المستوية (Plane Geometry) والهندسة الفراغية (Space Geometry) والمطلوب فحص هذه الافتراضية

Analysis_Geometry	Algebre	Plane_Geometry	Space_Geometry
٤٥	٤٦	٤٤	٤٧
٤٤	٤٠	٤٠	٤٥
٣٨	٣٥	٤١	٤٦
٥٠	٥٠	٤٦	٥٠
٤٠	٣٨	٣٨	٤١
٤٤	٤٧	٤٢	٤٩
٤٦			٥٠
٤٩			٤٩
٥٠			٤٨
٥٠	٥٠	٤٩	٤٨
٤٦	٤٥	٣٨	٣٧
٤٨	٤٥	٤٠	٤٤
٣٨	٤٠	٣٦	٤٠
٣٨	٣٥	٣٠	٣٢
٣٣	٣٠	٢٨	٣٢
٣٠	٢٣	٣٠	٣٥
٣٠	٣٣	٢٧	٣٠
٢٠	١٢	٢٠	٢٠

الفرضية:
هل توجد علاقة بين فروع مادة الرياضيات؟

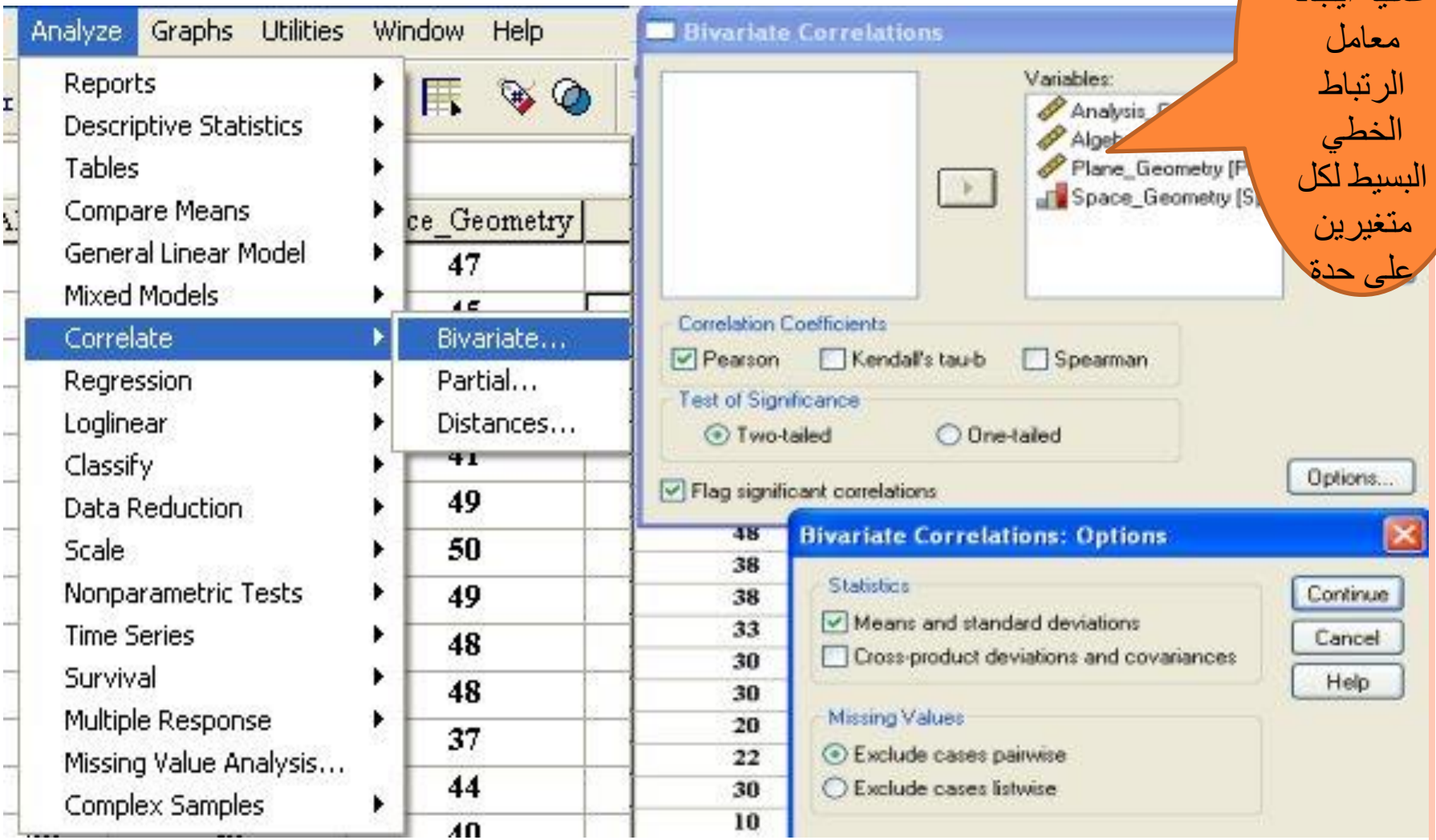
كل متغيرين مع بعض

نستخدم برنامج SPSS

ادخال المتغيرات كلها تختصر عملية ايجاد معامل الرتباط الخطي البسيط لكل متغيرين على حدة

: كما موضح في الشكل التالي:

Analyze → Correlate → Bivariate → Options → Continue



Correlations

Descriptive Statistics

	Mean	Std. Deviation	N
Analysis_Geometry	37.33	10.987	30
Algebre	36.03	11.755	30
Plane_Geometry	35.87	10.527	30
Space_Geometry	38.60	10.966	30

مصفوفة معاملات الارتباط بين أبعاد الفروع الأربعة

Correlations

		Analysis_ Geometry	Algebre	Plane_ Geometry	Space_ Geometry
Analysis_Geometry	Pearson Correlation	1	.960**	.921**	.926**
	Sig. (2-tailed)		.000	.000	.000
	N	30	30	30	30
Algebre	Pearson Correlation	.960**	1	.907**	.917**
	Sig. (2-tailed)	.000		.000	.000
	N	30	30	30	30
Plane_Geometry	Pearson Correlation	.921**	.907**	1	.970**
	Sig. (2-tailed)	.000	.000		.000
	N	30	30	30	30
Space_Geometry	Pearson Correlation	.926**	.917**	.970**	1
	Sig. (2-tailed)	.000	.000	.000	
	N	30	30	30	30

**. Correlation is significant at the 0.01 level (2-tailed).

* تعني دال عند مستوى 0.05

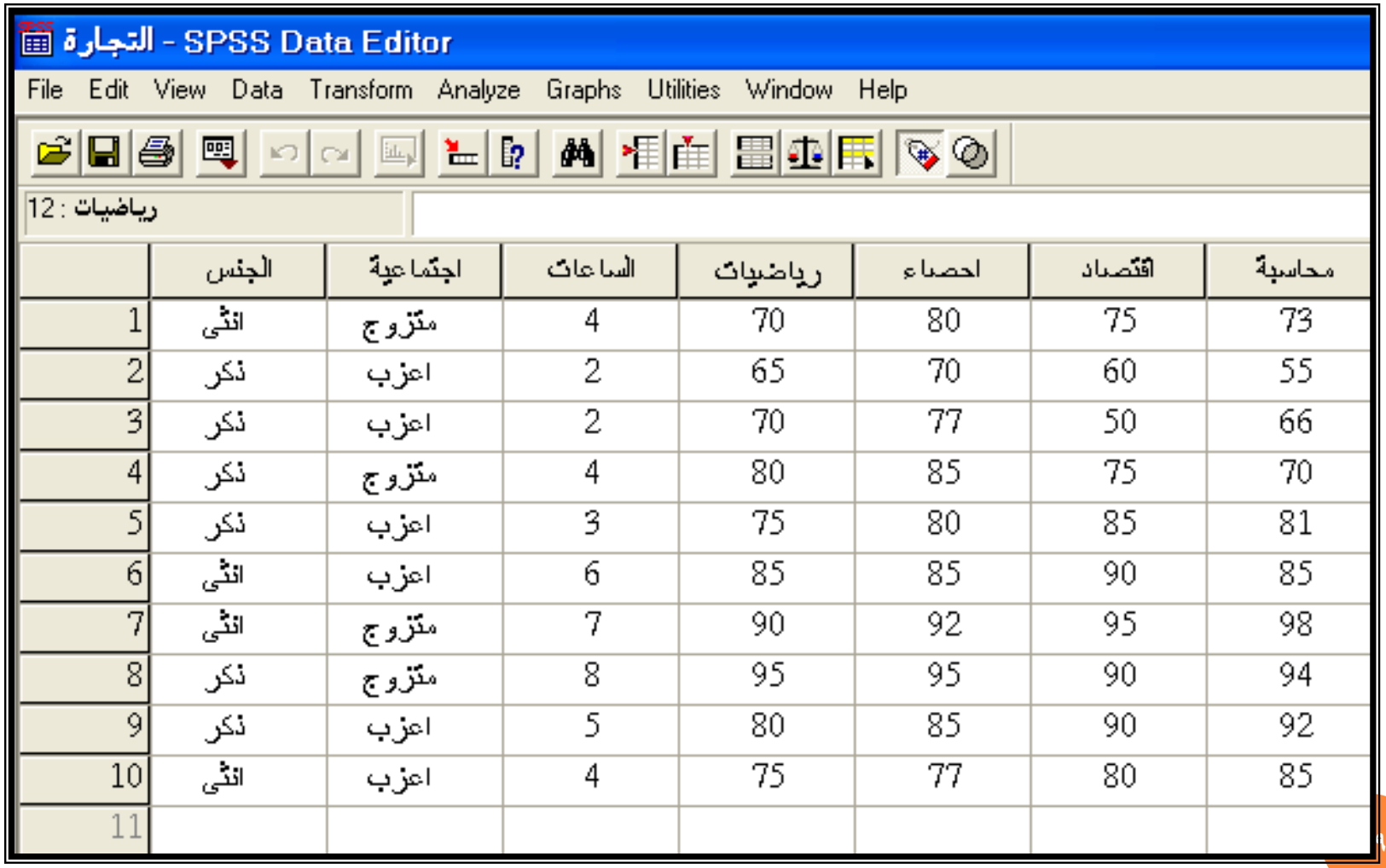
** تعني دال عند مستوى 0.01

مثال 8

في المثال السابق يشك الباحث في أن هنالك تأثير لفهم أحد الفروع على الآخر لوجود تشابه في المفردات أو القدرات المطلوبة للفهم .. فماذا يعمل ؟

ملاحظة:

إن معاملات الارتباط السابقة ليست دقيقة وصافية للعلاقة بين كل متغيرين فيعتبر كل واحد منهم دخيل على الآخرين فقد يضخم أو يقلل معامل الارتباط بينهما.



	الجنس	اجتماعية	الساعات	رياضيات	احصاء	اقتصاد	محاسبة
1	انثى	متزوج	4	70	80	75	73
2	ذكر	اعزب	2	65	70	60	55
3	ذكر	اعزب	2	70	77	50	66
4	ذكر	متزوج	4	80	85	75	70
5	ذكر	اعزب	3	75	80	85	81
6	انثى	اعزب	6	85	85	90	85
7	انثى	متزوج	7	90	92	95	98
8	ذكر	متزوج	8	95	95	90	94
9	ذكر	اعزب	5	80	85	90	92
10	انثى	اعزب	4	75	77	80	85
11							

بغرض المقارنة نقوم بإيجاد معاملات الارتباط المتعددة عن طريق المسار (كماسبق)

Analyze → Correlate → Bivariate → Options → Continue

فنحصل على المصفوفة التالية:

Correlations

		رياضيات	احصاء	اقتصاد	محاسبة
رياضيات	Pearson Correlation	1	.959**	.780**	.833**
	Sig. (2-tailed)	.	.000	.008	.003
	N	10	10	10	10
احصاء	Pearson Correlation	.959**	1	.746*	.811**
	Sig. (2-tailed)	.000	.	.013	.004
	N	10	10	10	10
اقتصاد	Pearson Correlation	.780**	.746*	1	.890**
	Sig. (2-tailed)	.008	.013	.	.001
	N	10	10	10	10
محاسبة	Pearson Correlation	.833**	.811**	.890**	1
	Sig. (2-tailed)	.003	.004	.001	.
	N	10	10	10	10

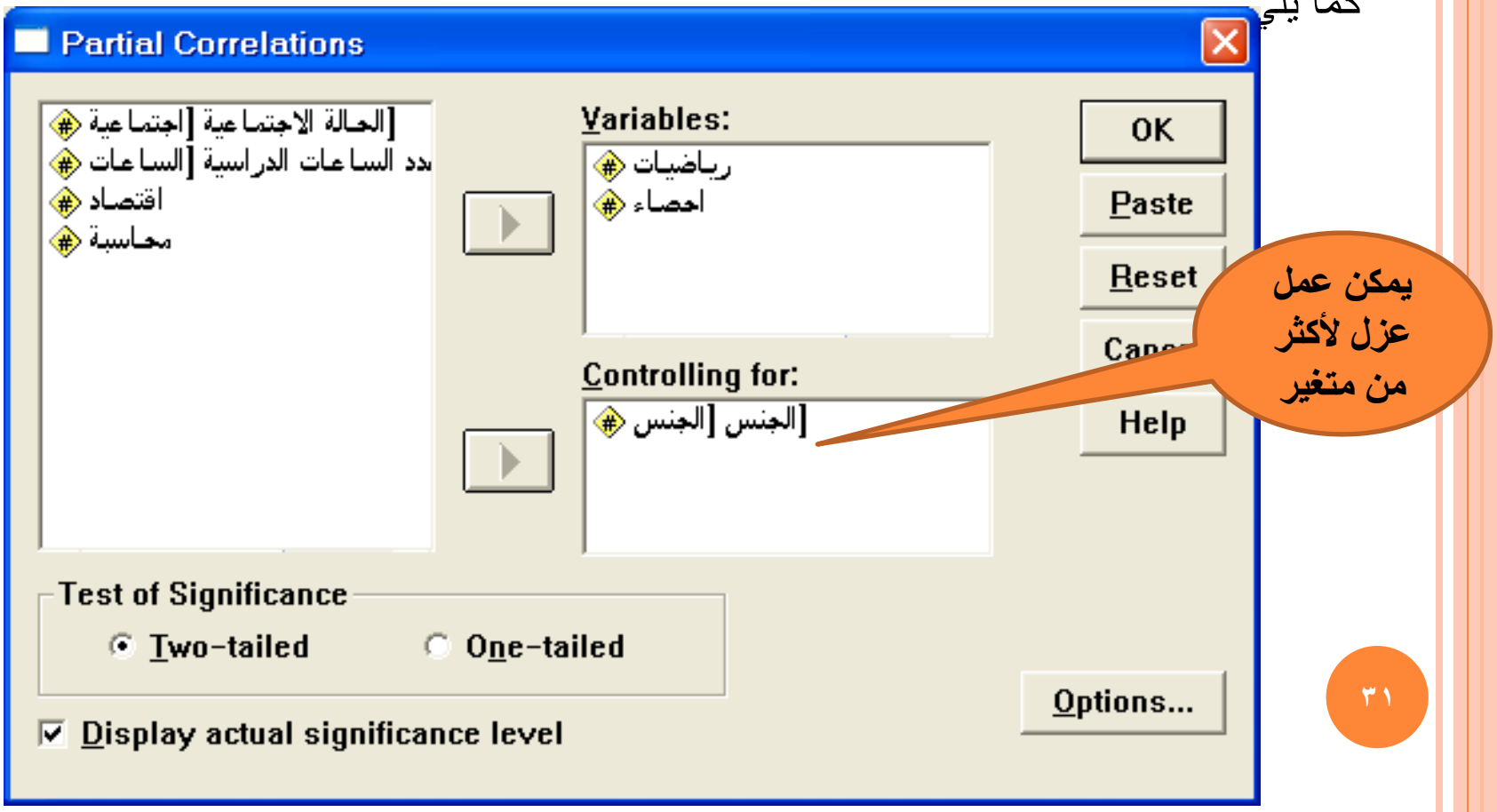
**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

الآن نقوم بإيجاد معامل الارتباط الجزئي عن طريق المسار

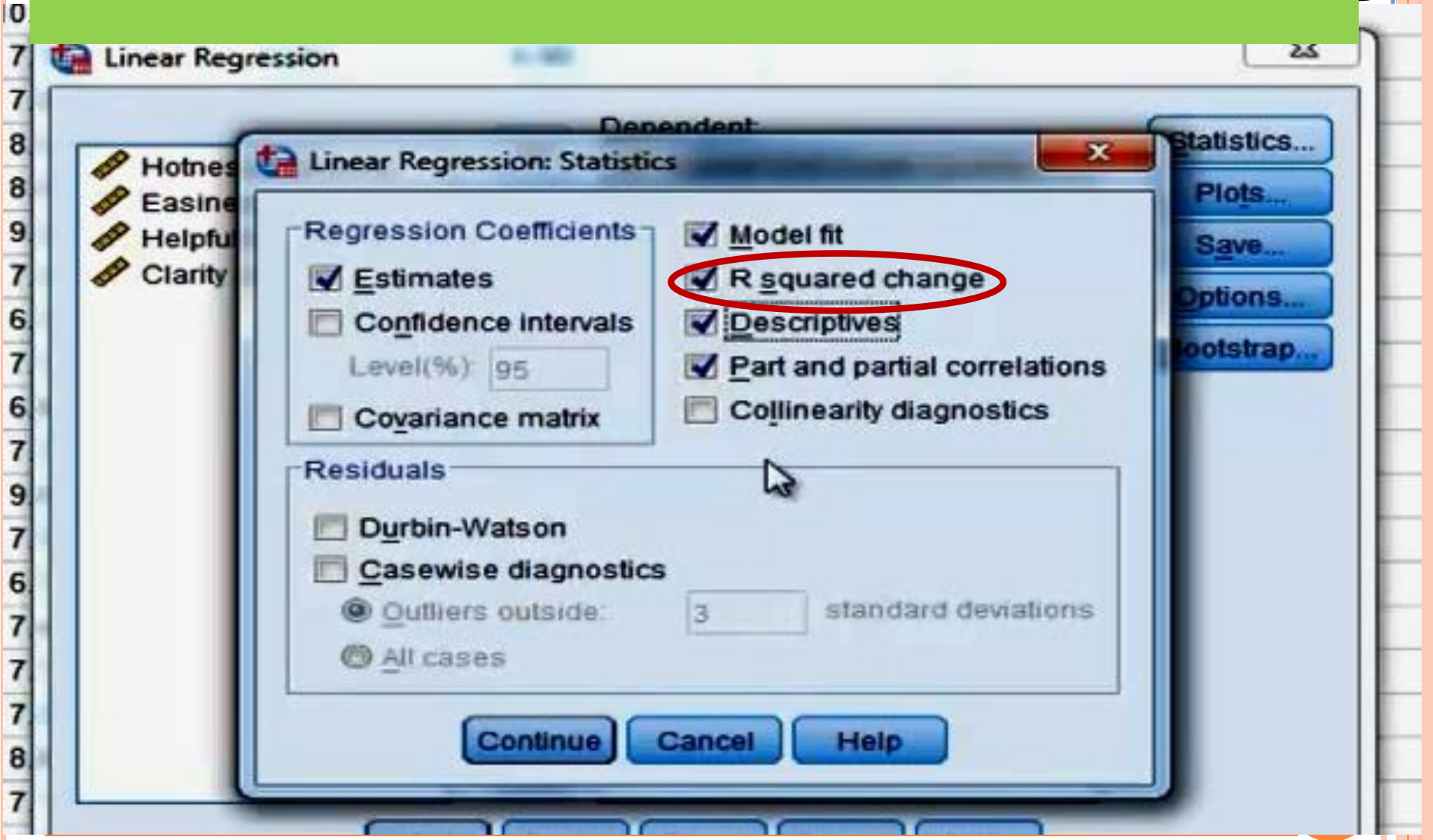
Analyze → Correlate → Partial→

ومن القائمة الفرعية اختر Partial يظهر مربع الحوار التالي:
ادخل المتغيرين " رياضيات " و " إحصاء " داخل المستطيل Variables ومتغير " الجنس " في المستطيل اسفل Controlling for: .ثم اضغط على زر Ok
كما يلي



لايجاد معامل الارتباط المتعدد (بين كل المتغيرات في آن واحد) : يكفي عمل انحدار متعدد (كما سيؤخذ لاحقا) واختيار

Linear Regression :statistics من R squared change



الانحدار الخطي البسيط

أما إذا كان المطلوب معرفة مدى تأثير متغير مستقل على متغير تابع أو التنبؤ بقيم المتغير التابع عند تحديد قيم معينة للمتغير المستقل فإن تحليل الانحدار يكون هو المناسب .

الانحدار الخطي البسيط (تنبؤ)

مثال 9

مزرعة للعجول ، يتم إعطاء العجول كميات من البروتين اليومي بغرض زيادة في الوزن، فإذا كان المطلوب زيادة في وزن العجل الرضيع 18كجم، فما هي كمية البروتين اليومية اللازم إعطاءها له؟

الحل قام الباحث بالتجريب على عينة من العجول وحصل على البيانات التالية التي تعبر ن كمية البروتين اليومي بالجرام التي يحتاجها العجل الرضيع، ومقدار الزيادة في وزن العجل بالكجم، وذلك لعينة من العجول الرضيعة حجمها 10

كمية البروتين	10	11	14	15	20	25	46	50	59	70
الزيادة في الوزن	10	10	12	12	13	13	19	15	16	20

بفرض أن هي x كمية البروتين، yهي مقدار الزيادة في الوزن، فإن معادلة الانحدار المقدرة، هي:

$$\hat{y} = 9.44 + 0.143x$$

لابد من تفسير لهذه المعادلة:

- الثابت: $\hat{\beta}_0 = 9.44$ يدل على أنه في حالة عدم استخدام البروتين قي التغذية، فإن الوزن يزيد 9.44 كجم.

معامل الانحدار: $\hat{\beta}_1 = 0.143$: يدل على أنه كلما زادت كمية البروتين جرام واحد، حدث زيادة في وزن العجل بمقدار 0.143 كجم، أى زيادة مقدارها 143 جرام.

$$\hat{y} = 9.44 + 0.143x$$

زيادة في الوزن ← $\hat{y}$ ، x → البروتين كمية

إذا أردنا زيادة في الوزن مقدارها 18 فما هي كمية البروتين اللازمة للحصول على هذه الزيادة

$$18 = 9.44 + 0.143\ x \quad ----> x = ?$$

$$x = 59.87$$

- إذا كانت كمية البروتين $x = 55$ فإن الزيادة في الوزن :

$$\hat{y} = 9.44 + 0.143(55) = 17.305$$

- إذا كانت كمية البروتين $x = 85$ فكم ستكون الزيادة في الوزن :

$$\hat{y} = 9.4 \quad .143(85)$$

لأنها قيمة خارج المجال والصحيح أن نتنبأ بقيم ضمن المجال حيث تم تقدير هذه المعادلة في هذا المجال أما في قيم خارج المجال فإننا نحتاج للسلاسل الزمنية وتحديد معادلة الاتجاه العام

للاطلاع

٦- الاختبارات اللا معلمية (أو اللا برامترية) :

* والمقصود بالإحصاء اللا معلمي وبالتالي الاختبارات اللا معلمية هي الحالات التي لا يكون التوزيع الاحتمالي (أو الإحصائي) للمجتمع محل الدراسة معروف . ويضاف إلى ذلك أيضاً العينات الصغيرة ، في هذه الحالات والتي قد تكون أكثر واقعية (لأنه في الغالب لا يعرف الباحث أي شيء عن توزيع المجتمع) فإن الاختبارات المناسبة هي الاختبارات اللا معلمية (توزيع المجتمع غير معروف والعينات صغيرة) .

أولاً : اختبار الوسط الحسابي :

أ- اختبار " ذو الحدين" . ب- اختبار مربع كاي لحسن المطابقة .

ج- اختبار كولموجروف – سمير نوف لعينة واحدة Kolmogrov – Smirnov .

ثانياً : اختبار الفرق بين وسطين في حالة الاستقلال :

أ- اختبار مان – وتني Mann- Whitney .

ب- اختبار كولموجروف – سمير نوف .

للاطلاع

ثالثاً : اختبار الفرق بين وسطين في حالة عدم الاستقلال :

أ- اختبار ويلكوكسون Wilcoxon

ب- اختبار ماك ينمار Mc Nemar

ج- اختبار الإشارة Sign-test .

رابعاً : اختبار الفروق بين أكثر من متوسطين في حالة الاستقلال :

أ- اختبار كروسكال – واليس Kruskal- Wallis

ب- اختبار الوسيط Median- test

خامساً : اختبار الفروق بين أكثر من متوسطين في حالة عدم الاستقلال :

أ- اختبار فريدمان Frideman test .

ب- اختبار كوكران Cochran test .

وهنا يجب على الطالب تحرى الدقة عند اختيار الاختبار المناسب والتأكد من أنه ينطبق على الحالة التي يقوم بدراستها .

الانحـــــدار المتـــعدد
Multiple Linear regression

مثال : 10

سنفترض البيانا ت التالية وهي ل ١٥ محل تجاري كعينة عشوائية في يوم معين ولسلعة ما، اليوم ودرجات الحرارة المسجلة في منتصف اليوم وسنوات الخبرة للبائع الذي يقوم بالخدمة .

لتكن:

- المبيعات Sales هي المتغير التابع dependent variable
- سنوات الخبره , Years درجة الحرارة Temperature هما المتغيراتن المستقلان،
- وسنستخدم الرموز التالية:
  - Y=Sales
  - *X*1=Temperature
  - X2= Years

المطلوب بناء نموذج خطي يصف ويتنبأ بحجم المبيعات بمعلومية درجة الحرارة وسنوات الخبرة للبائع

Sales	Temperature	Years
15	21	1
15	18	1
21	22	1
28	24	2
30	25	2
35	25	2
40	26	2
35	34	3
30	25	3
45	38	3
50	40	4
60	41	4
45	39	5
60	37	5
50	40	6

إن النموذج الخطي المقترح له الشكل التالي :

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$$

معالم مجهولة Parameters

الخطأ في النموذج

وعلى خطأ النموذج بعض الشروط وهي

1. مجموع الاخطاء العشوائية مساويا للصفر وبذلك يكون متوسط الأخطاء مسويا الصفر
2. تباين الخطأ مقدار ثابت لكل المشاهدات ومساويا σ^2 هناك تجانس
3. التغاير بين اى خطأين مساويا الصفر وبذك لا يوجد ارتباط بين الاخطاء وبعضها البعض
4. يفترض أن الخطأ يتوزع حسب التوزيع الطبيعى بمتوسط صفر وتباين مشترك σ^2

كيف يمكن الحصول على هذه المعادلة؟

يوجد ثلاثة أنواع من نماذج الانحدار :

الانحدار العياري : Standard or Simultaneous Regression

وفيها يتم ادخال **كل** المتغيرات المستقلة دفعة واحدة .
ولا نتعرض لمناقشة هل المتغيرات المستقلة مرتبطة ببعضها البعض ام مستقلة

الانحدار االهرمي : Hierachical Regression

وفي هذا النوع ندخل المتغيرات تباعا (**واحدا واحدا**)

الانحدار ال: Stepwise Regression

وفي هذا النوع ندخل **عدد** من المتغيرات (**وليس كل** المتغيرات).
وقد يكون الادخال

للأمام (forward)
للخلف (backward)
مزج بين الأسلوبين

- ويمكن الان التعرض لبعض انواع الانحدار وكيفيه استخدام الحزمة SPSS فى ذلك وسوف نهتم هنا بالانواع التاليه:

1. الانحدار الخطى المتعدد Multiple Linear Regression
2. الانحدار التدريجى Stepwise Regression
3. الانحدار الهرمى Hierachical Regression
4. الانحدار الغير خطى Curve Linear Regression

Enter لإدخال المتغيرات فى النموذج دفعة واحدة

وتستخدم اذا اردنا ان يحتوي نموذج الانحدار على كل المتغيرات المستقلة المدرجة حسب تسلسل الادخال.

Remove لحذف المتغيرات فى النموذج دفعة واحدة.

Forward لإدخال المتغيرات فى النموذج واحد تلو الآخر

وتستخدم اذا اردنا ان يحتوي نموذج الانحدار فقط على المتغيرات المستقلة المدرجة ذات الارتباط الدال (المعنوي) مع المتغير التابع مرتبة حسب قوة الارتباط (وليست على حسب تسلسل الادخال).

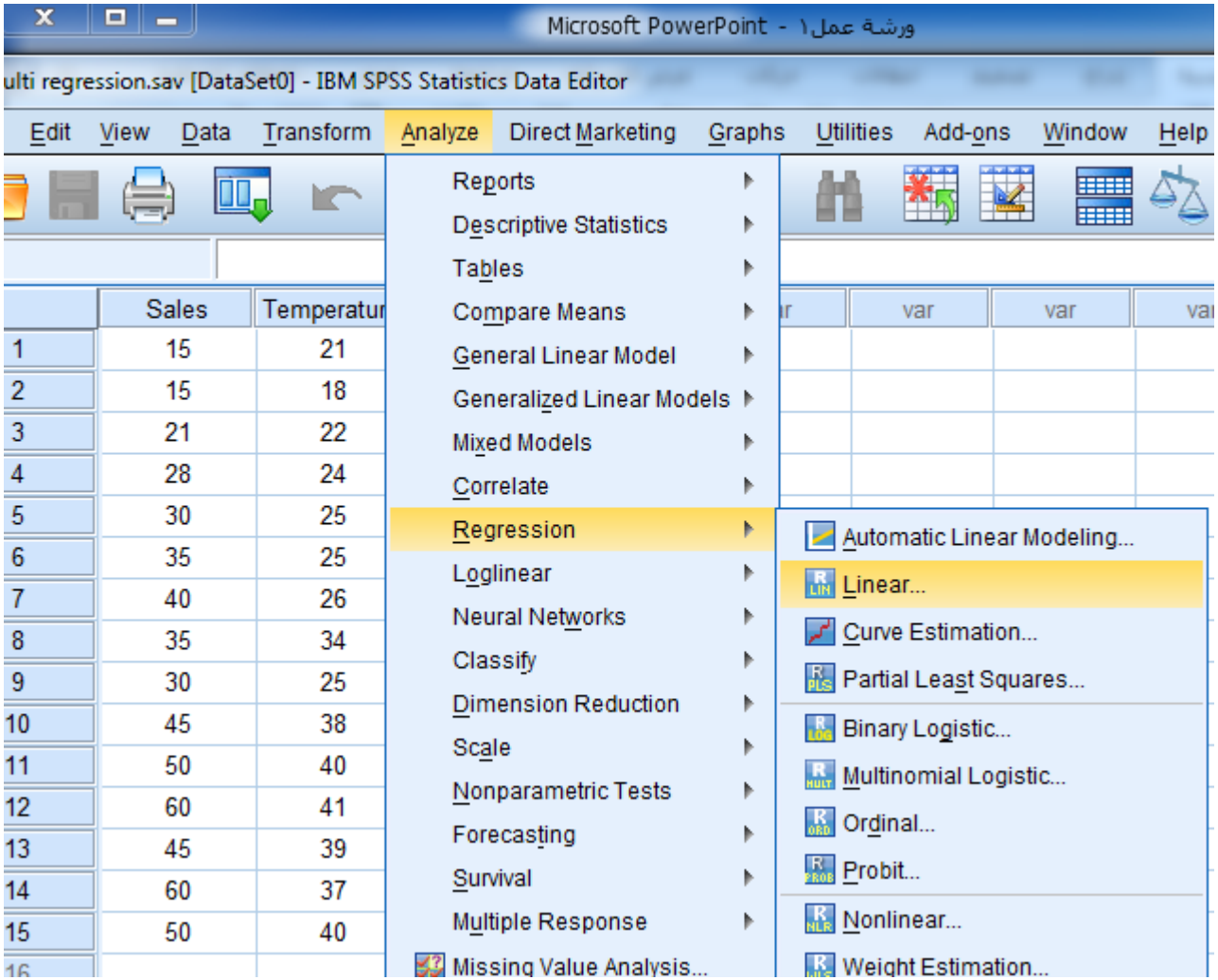
Backward لإدخال جميع المتغيرات فى النموذج دفعة واحدة ثم حذفها واحد تلو الآخر حسب درجة معنوية ارتباطها مع المتغير التابع.

Stepwise لإدخال المتغيرات فى النموذج واخراجها واختبارها واحد تلو الآخر

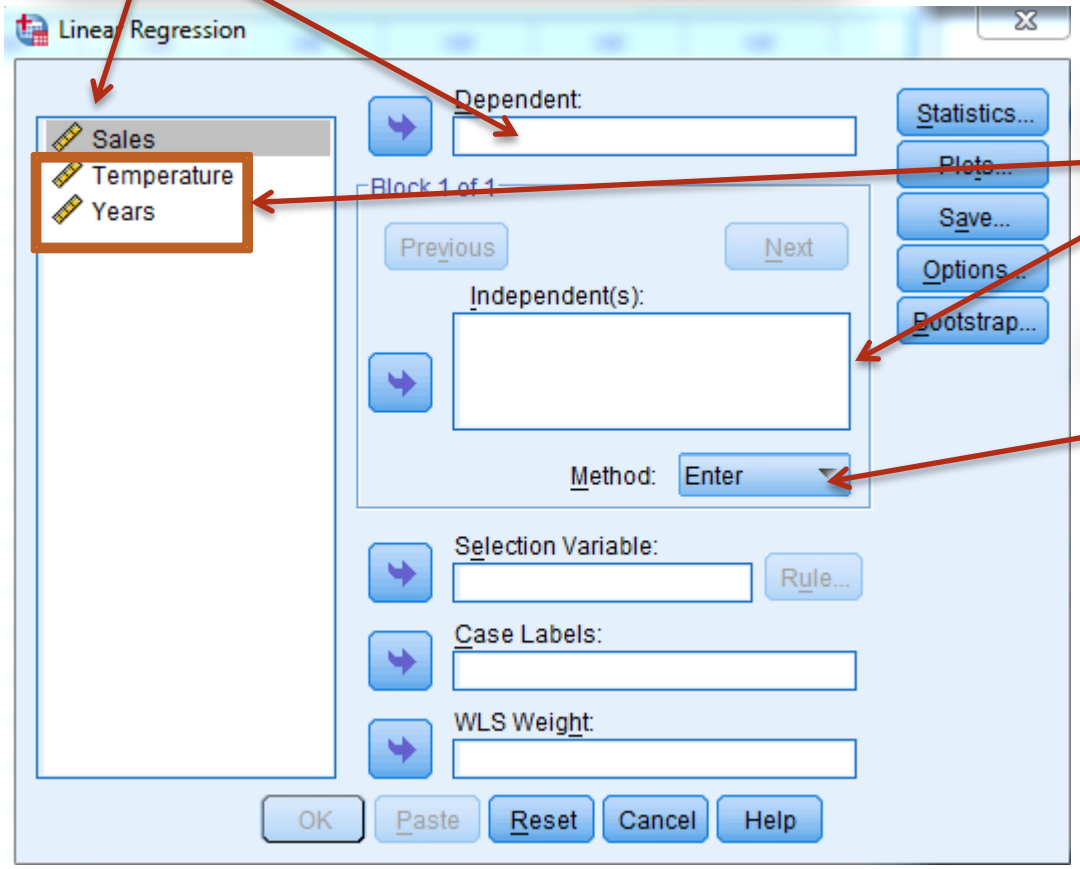
وتستخدم اذا اردنا ان يحتوي نموذج الانحدار فقط على المتغيرات المستقلة المدرجة ذات الارتباط الجزئي الدال (المعنوي) مع المتغير التابع فى وجود المتغيرات المستقلة مرتبة حسب قوة الارتباط (وليست على حسب تسلسل الادخال).

الانحدار العياري أو الانحدار الخطي المتعدد
Standard or Simultaneous Regression
Or
Multiple Linear Regression

Analyze-------Regression-----Linear



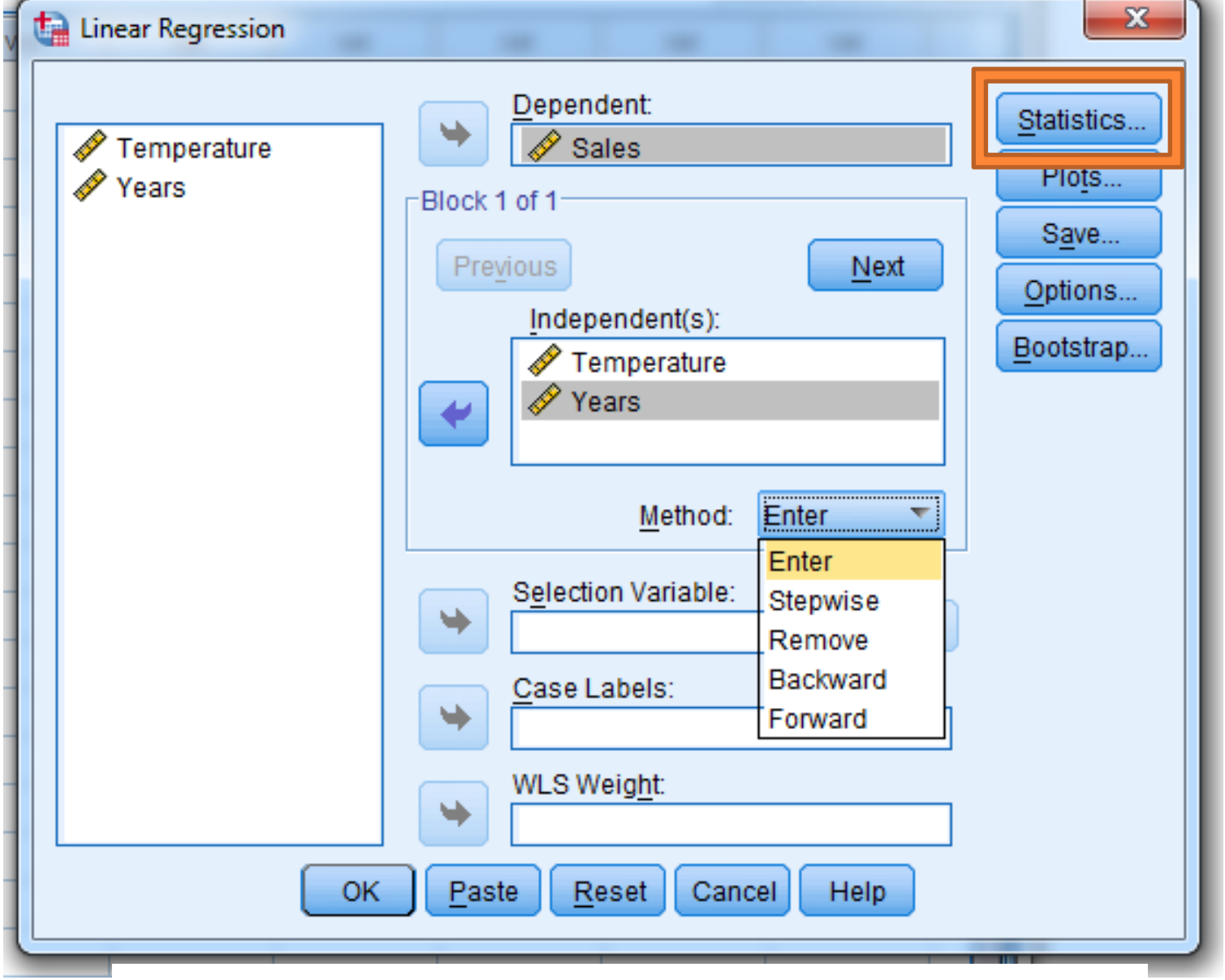
ننقل المتغير التابع Sales لخانة Dependent



ننقل المتغيرات المستقله temp, years
لخانة Independent

نختار نوع الانحدار من خانة Method

الطريقة العيارية وهى Enter

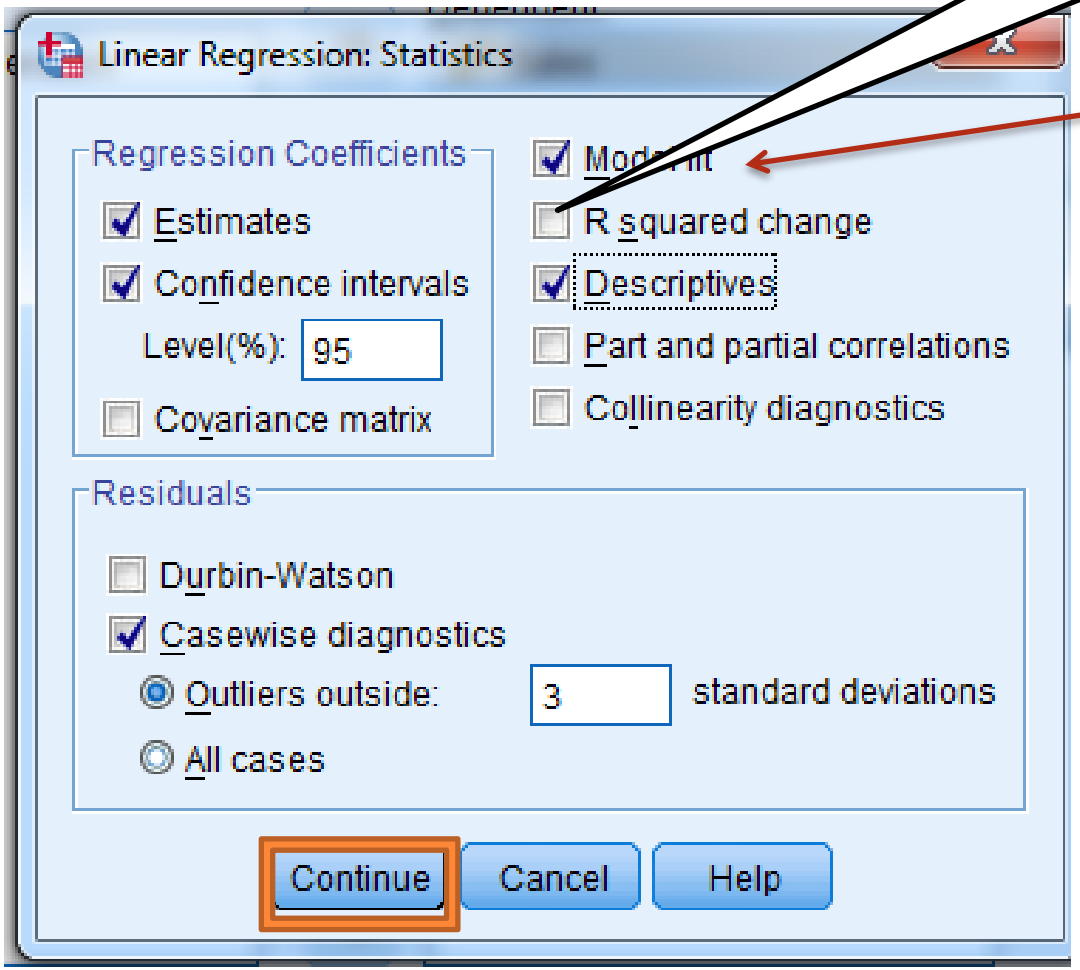


نضغط على الامر Statistics تظهر شاشه جديده بعنوان Linear Regression: Statistics

هذه الاختيار يعطيك معامل الارتباط بين الثلاثة المتغيرات معا

نتأكد أن Model fit, Estimates, Casewise diagnostics

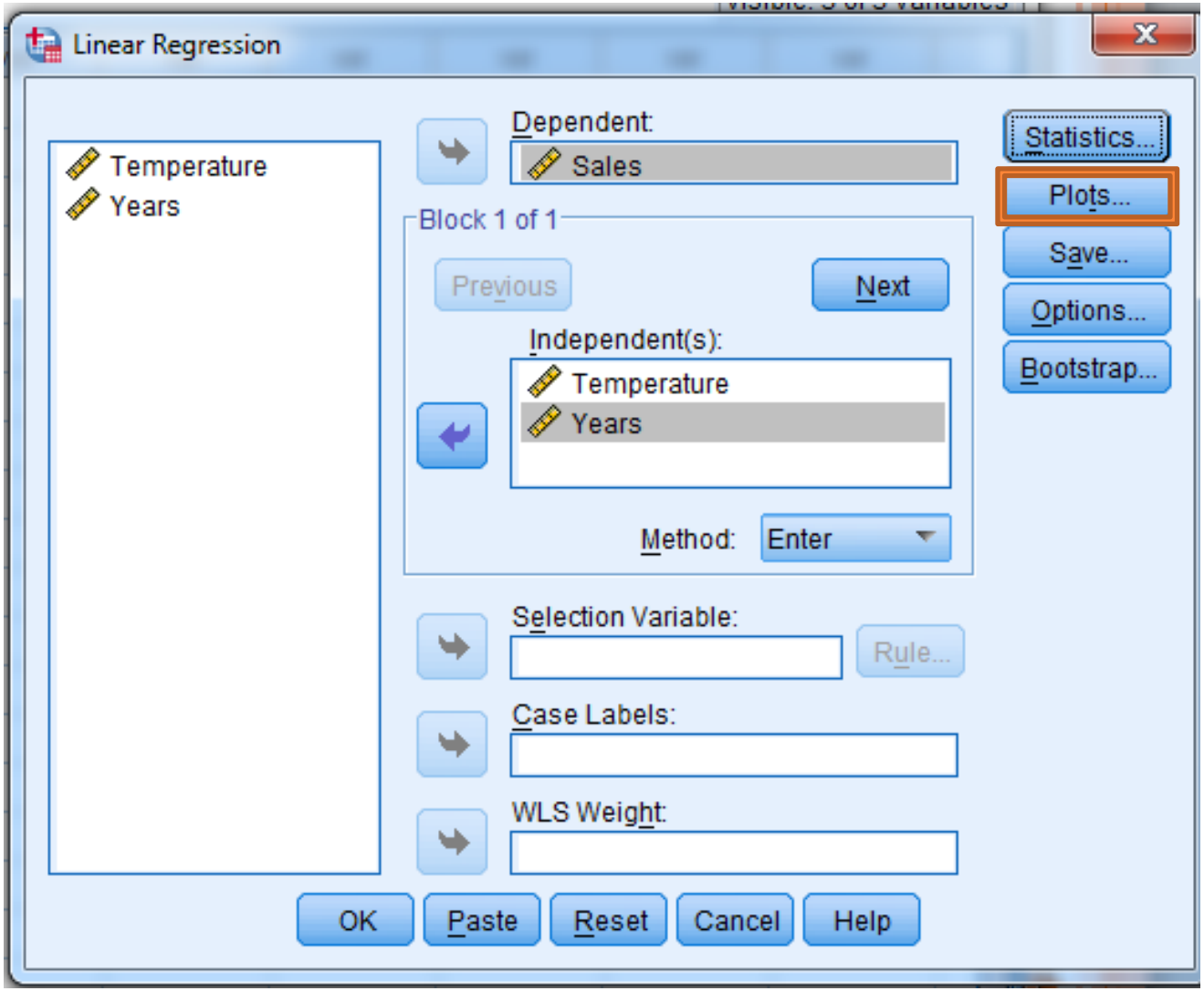
مختاره ويمكن ايضا اضافه اختيارات اخرى



موائمة النموذج (Model fit)

حيث يتم عرض معاملات او مؤشرات حسن المطابقة او الموائمة **Goodness of Fit** للمتغيرات التي يتم ادخالها او اخراجها من النموذج ، وتعرض المؤشرات : معامل الارتباط المتعدد **R** ، ومعامل التحديد R^2 ، وجدول تحليل التباين .

نختار Continue لنعود للشاشة السابقة



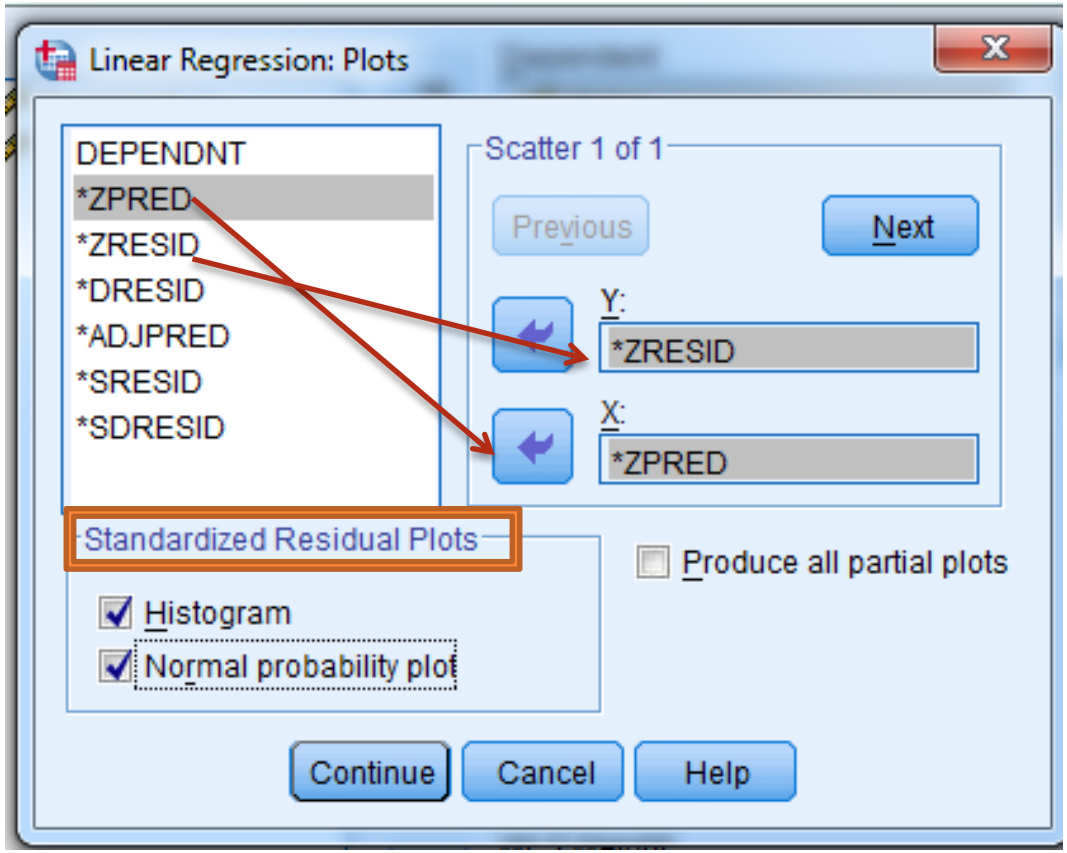
نضغط على الامر Plot فتظهر شاشه جديده بعنوان Linear Regression: Plots

الرسوم البيانية تساعد في اختبار صدق الافتراضات الأساسية مثل الإعتدالية والخطية وتجانس التباين ، وكذلك تفيد في الكشف عن الحالات أو الدرجات المتطرفة في التوزيع .

وبالنافذة السابقة يوجد ما يلي :

رسم بياني لمخطط شكل الانتشار – **Scatter plots**

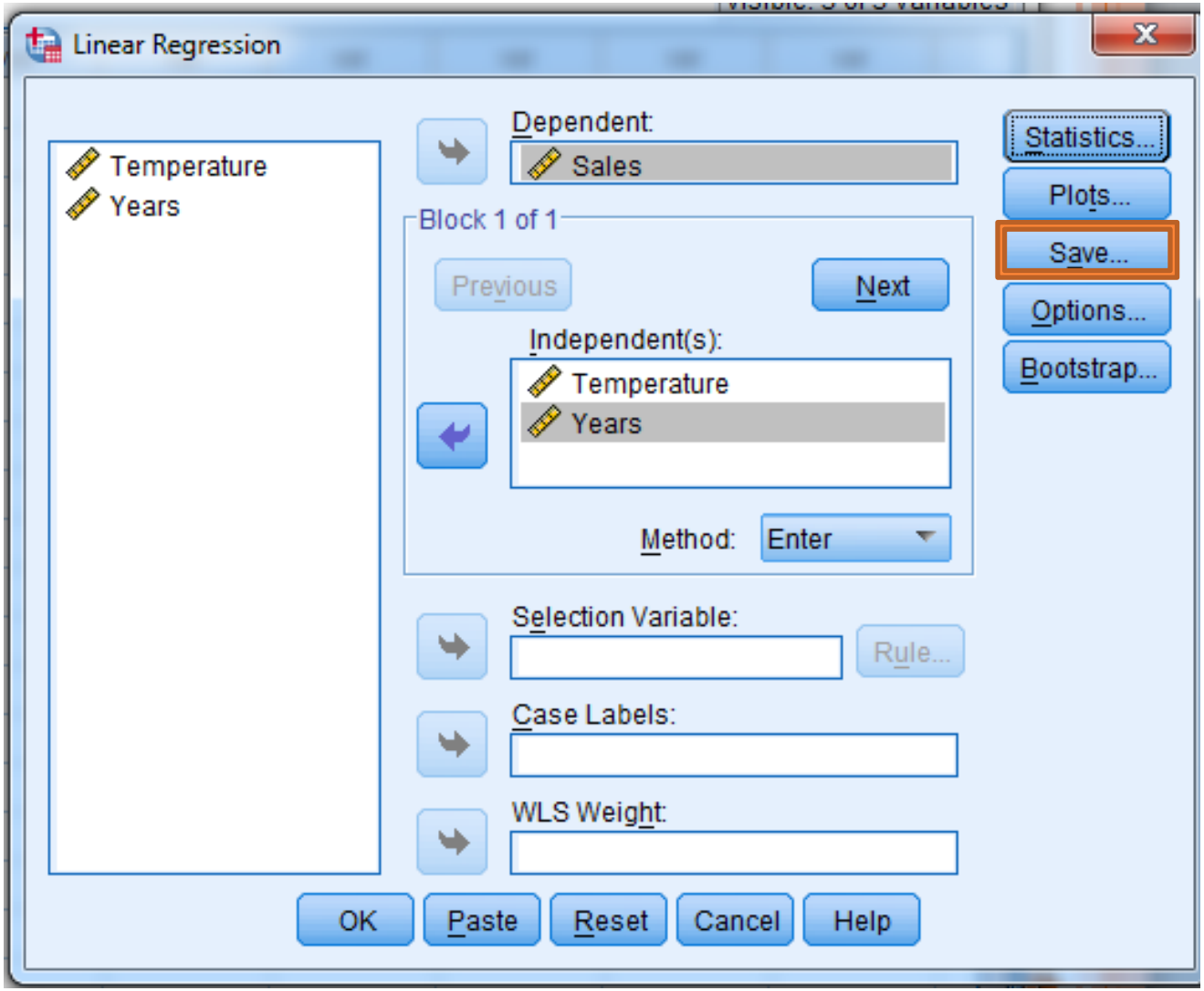
حيث يسمح بتحديد أكثر من رسم لمخططات الانتشار في الخانات **X , Y** ويمكن التحول بينهما بالزرين **Previous , next** ويمكننا من رسم شكل الانتشار بين أي اثنين من الآتي : المتغير التابع ، القيم المنبئة **Predicted Values**، البواقي ، البواقي المحذوفة ، القيم المنبئة المعدلة **Adjusted** ، البواقي المعيارية **Standardized** ويمكنك عمل رسم لشكل الانتشار بين البواقي مع القيم المنبئة لاختبار الخطية وتجانس التباين



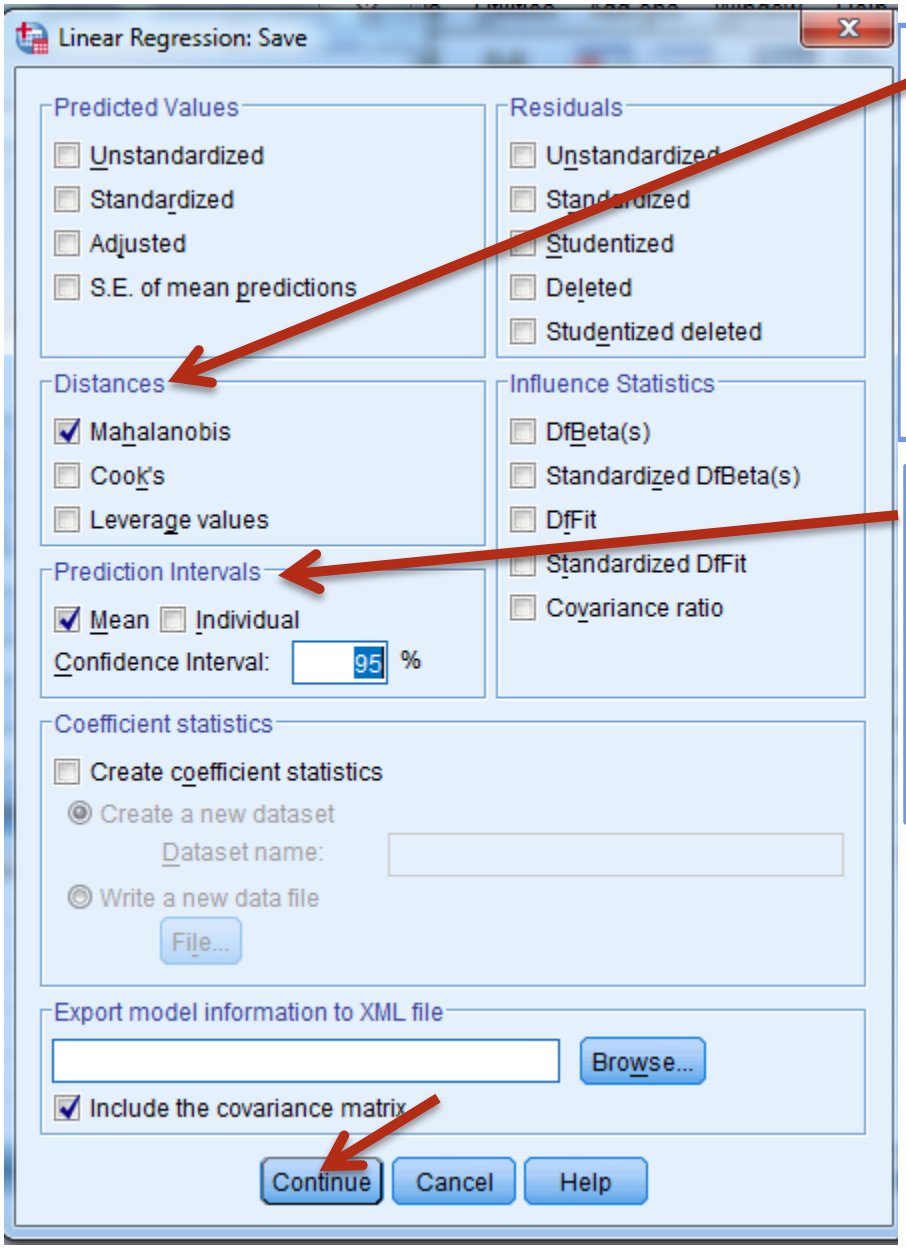
ننقل *ZRESID للمستطيل المقابل لـ Y: وايضا *ZPRED للمستطيل المقابل X:

من قائمة Standardized Residual Plots نختار كلا من Histogram, Normal probability plot

نضغط على Continue فنعود للشاشة السابقة



نضغط على الامر Save تظهر شاشة جديده بعنوان Linear Regression: Save



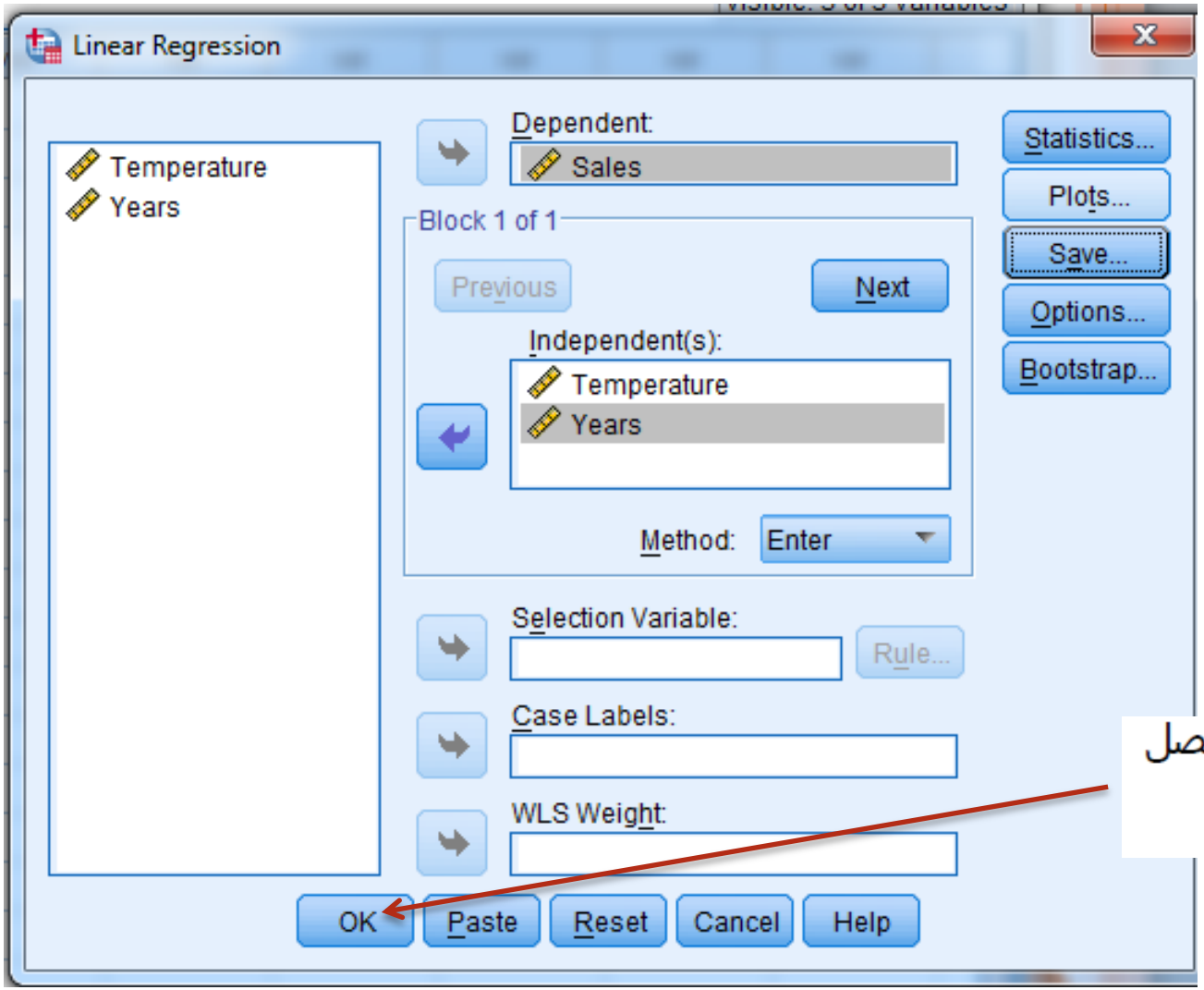
* Distances

وهي مقاييس تهدف إلى تحديد القيم أو الدرجات التي تشمل على تجمعات غير عادية للمتغير المستقل وكذلك الدرجات التي لها تأثير كبير على نموذج الانحدار ، وتقدم الاختبارات لتحديد ذلك وتشمل **Mahalanobis** **Cooks , Leverage** الفئات التنبؤية – **Prediction Intervals**

Prediction intervals

الفئات التنبؤية

وتقدم الحدود العليا والدنيا لكل من فئات المتوسط والدرجات الفردية المنبئة وتمكن المستخدم من تحديد مستوى الدلالة أو الثقة Confidence Interval .



نضغط على Ok فنحصل على النتائج التاليه.

الجدول الأول:

بعنوان Descriptive Statistics ويعطى لنا المتوسط والانحراف المعيارى وعدد الحالات لكل متغير على حده.

Descriptive Statistics

	Mean	Std. Deviation	N
Sales	37.27	14.499	15
Temperature	30.33	8.217	15
Years	2.93	1.580	15

الجدول الثانى:

بعنوان Correlations وهى مصفوفه الارتباط بين جميع المتغيرات وايضا معنويه الارتباط ونلاحظ انه لا يوجد ارتباط تام بين المتغيرات المستقله وبعضها الاخر.

Correlations

		Sales	Temperature	Years
Pearson Correlation	Sales	1.000	.907	.849
	Temperature	.907	1.000	.888
	Years	.849	.888	1.000
Sig. (1-tailed)	Sales	.	.000	.000
	Temperature	.000	.	.000
	Years	.000	.000	.
N	Sales	15	15	15
	Temperature	15	15	15
	Years	15	15	15

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Years, Temperature[b]	.	Enter

a. Dependent Variable: Sales

b. All requested variables entered.

الجدول الثالث

بعنوان Variables Entered/Removed ويحتوى على أسماء المتغيرات التى دخلت فى معادله الانحدار وهما متغيرين year, temp والمتغيرات التى استبعدت من الدخول فى المعادلة وهنا فى الطريقه العياريه لا تستبعد متغيرات.

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.912[a]	.832	.804	6.426

a. Predictors: (Constant), Years, Temperature

b. Dependent Variable: Sales

الجدول الرابع:

بعنوان Model Summary يحتوى على بعض المقاييس التى تم حسابها للنموذج المقدر وهى

1. قيمة معامل الارتباط R=0.912 وهو عالى جدا
2. مربع معامل الارتباط ويستخدم تعين مدى البيانات المستخدمة من المتغيرات المستقله فى تقدير المتغير التابع ونلاحظ أن النموذج المقدر يعبر عن (المتغيرين المستقلين معا) 80% من البيانات وزيادة قيمه هذا المقياس يفسر أن النموذج المقترح ملائم.
3. تعيين مربع معامل الارتباط المعدل Adjusted R Square ويستخدم لنفس الغرض السابق ولكنه أدق.
4. تعيين خطأ التقدير Std. Error of the Estimate وهو هنا 6.42616 كلما قل دل على خطأ أقل للنموذج.

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2447.388	2	1223.694	29.633	.000[b]
	Residual	495.546	12	41.295		
	Total	2942.933	14			

a. Dependent Variable: Sales

b. Predictors: (Constant), Years, Temperature

الجدول الخامس:

يحتوى على نتائج تحليل التباين ANOVA لاختبار معنويه الانحدار

الفرض الصفرى: الانحدار غير معنوى

الفرض البديل: الانحدار معنوى

ومن جدول ANOVA نجد ان Sig. = 000 وهى أقل من مستوى المعنويه 0.05 لذا سوف نرفض فرض العدم ونقبل الفرض البديل وهى أن الانحدار معنوى وبالتالى توجد علاقة ما بين المتغيرات المستقله والمتغير التابع.

النموذج ككل معنوي بدون تحديد أي من المعاملات هو السبب في المعنوية

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-6.993-	8.428		-.830-	.423	-25.356-	11.371
	Temperature	1.275	.454	.722	2.807	.016	.285	2.264
	Years	1.907	2.362	.208	.807	.435	-3.240-	7.054

a. Dependent Variable: Sales

الجدول السادس:

بعنوان Coefficients ويساعد هذا الجدول فى الحصول على كلا من

1. معادلة خط الانحدار المقدرة والخطأ فى التقدير لكل معامل وذلك من العمود Unstandarized Coefficients حيث

$$Sales = -6.993 + 1.275\,temp + 1.907\,years$$

قيمة sig اكبر من 0.05 لذا فان الثابت ومعامل السنوات غير معنوى لكن فى حاله درجات الحرارة sig. =0.016 أقل من 0.05 لذا فان معامل درجات الحرارة معنوى وهو سبب معنوية تحليل التباين للانحدار.

العمود الاخير يقدم التقدير بفترة لمعاملات خط الانحدار والثابت

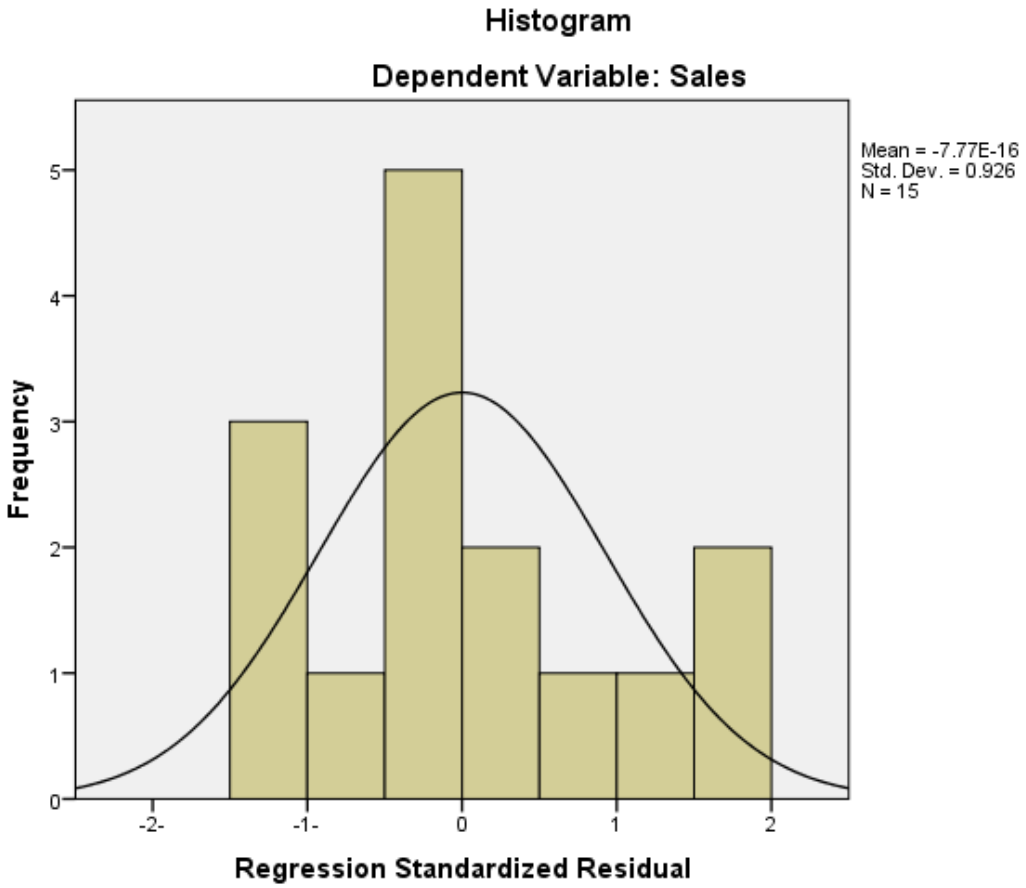
Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	17.86	55.44	37.27	13.222	15
Std. Predicted Value	-1.468-	1.374	.000	1.000	15
Standard Error of Predicted Value	1.945	4.247	2.801	.667	15
Adjusted Predicted Value	18.72	59.65	37.52	13.420	15
Residual	-7.255-	10.294	.000	5.949	15
Std. Residual	-1.129-	1.602	.000	.926	15
Stud. Residual	-1.257-	1.826	-.017-	1.034	15
Deleted Residual	-9.654-	13.374	-.254-	7.483	15
Stud. Deleted Residual	-1.291-	2.057	.008	1.090	15
Mahal. Distance	.349	5.182	1.867	1.347	15
Cook's Distance	.000	.333	.090	.113	15
Centered Leverage Value	.025	.370	.133	.096	15

a. Dependent Variable: Sales

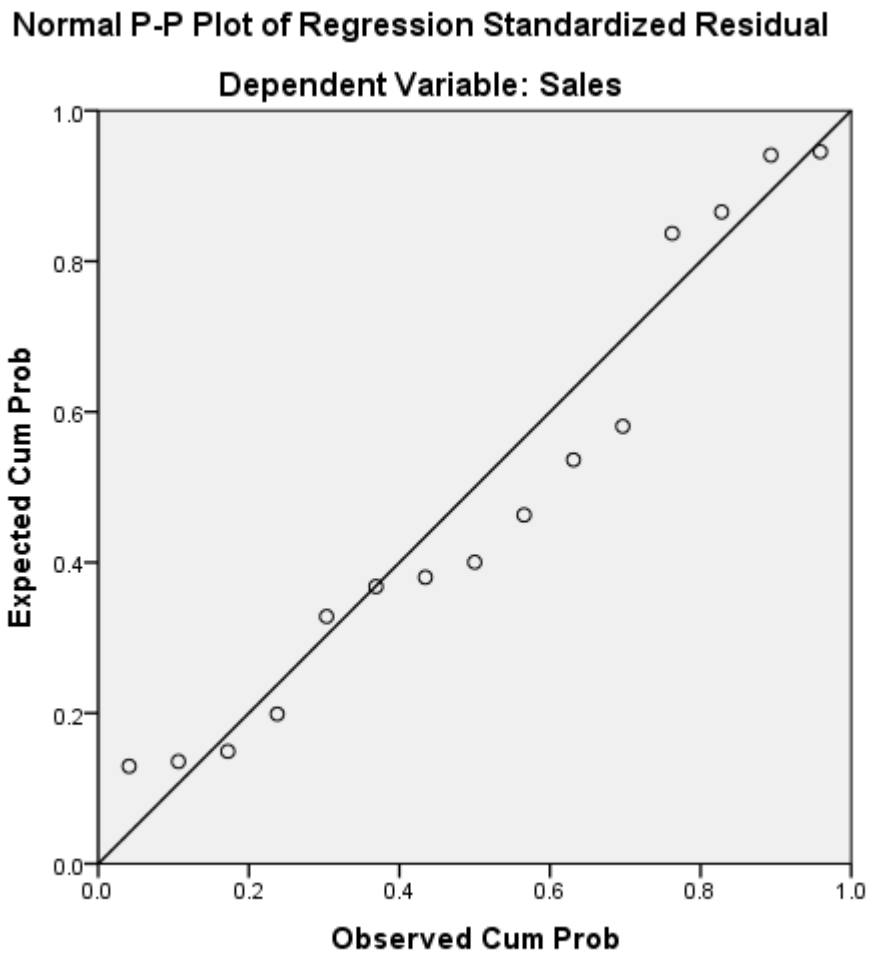
الجدول الثامن:

بعنوان Residuals Statistics يستخدم لمعرفه بعض المقاييس الخاصه بالبواقى



الشكل البيانى:

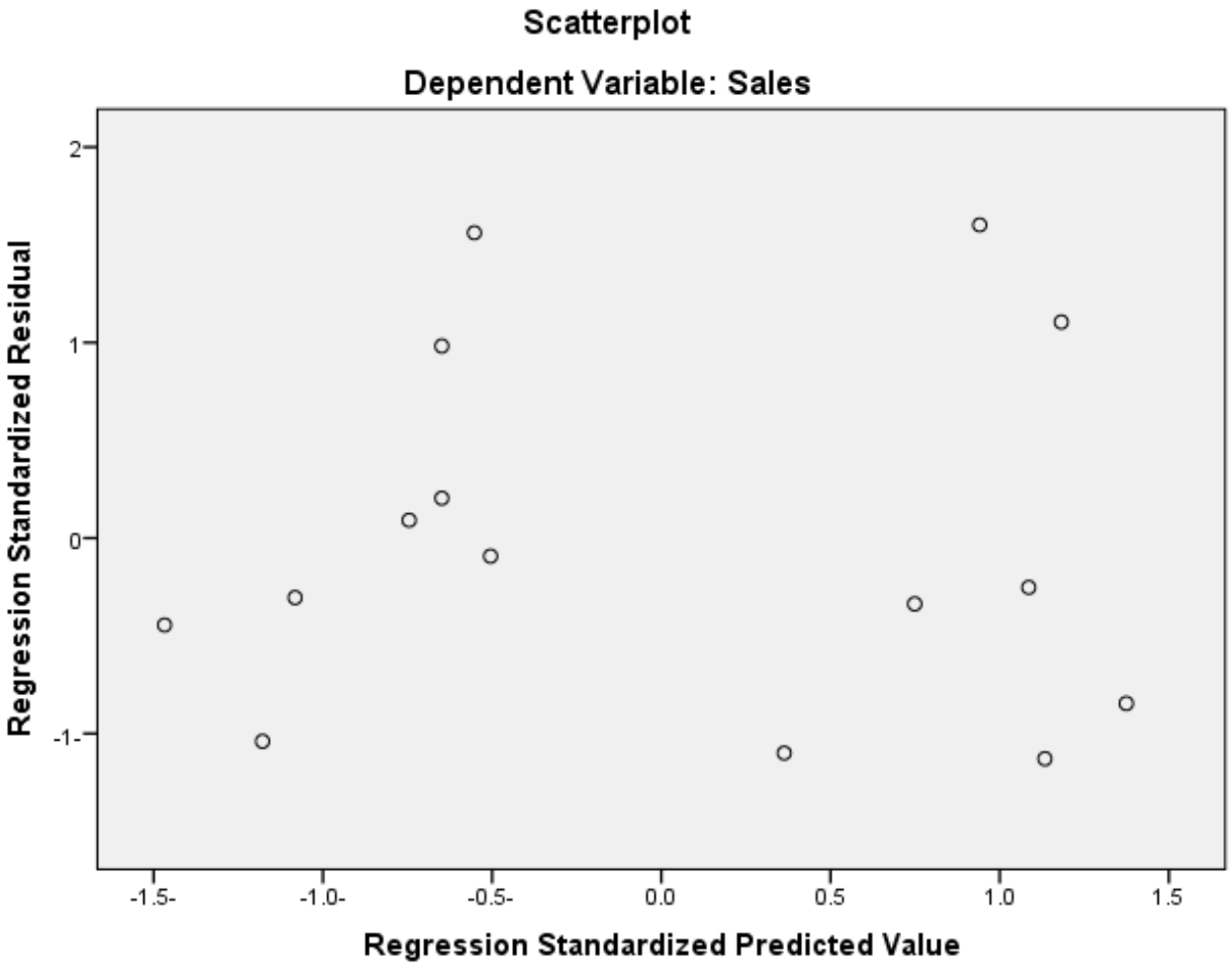
هو المدرج التكرارى ويستخدم للتعرف هل البيانات تتوزع حسب التوزيع الطبيعى أم لا؟



الشكل البيانى:

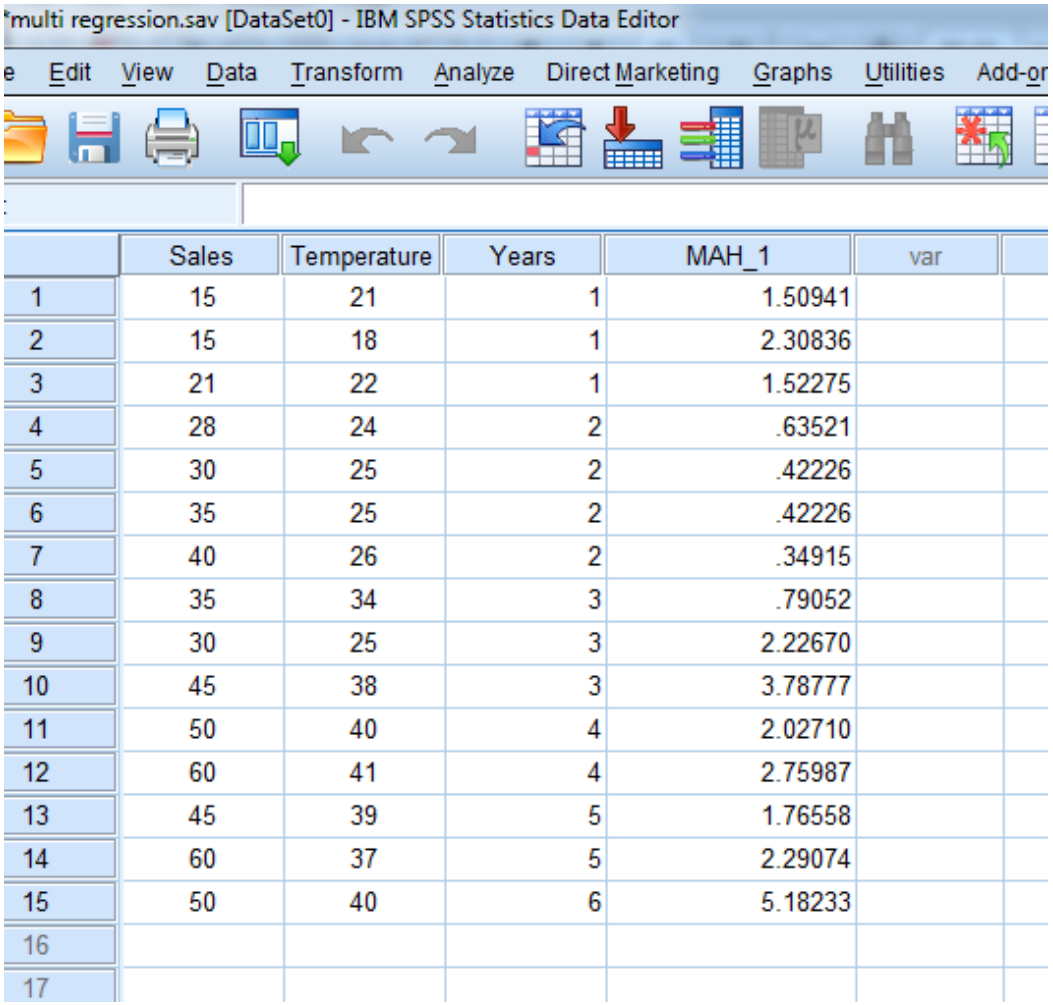
يختبر هل البواقى تتبع التوزيع الطبيعى ام لا؟

ومن الشكل نجد أن النقاط تتجمع حول الخط وبالتالى فان البيانات (البواقى) تتوزع حسب التوزيع الطبيعى.



الشكل البياني:

يمثل شكل الانتشار للبواقى مع القيم المتوقعه ومنه يتضح عدم وجود نمط معين للنقاط فى الشكل وهذا يتسق مع شرط الخطيه.



	Sales	Temperature	Years	MAH_1	var
1	15	21	1	1.50941	
2	15	18	1	2.30836	
3	21	22	1	1.52275	
4	28	24	2	.63521	
5	30	25	2	.42226	
6	35	25	2	.42226	
7	40	26	2	.34915	
8	35	34	3	.79052	
9	30	25	3	2.22670	
10	45	38	3	3.78777	
11	50	40	4	2.02710	
12	60	41	4	2.75987	
13	45	39	5	1.76558	
14	60	37	5	2.29074	
15	50	40	6	5.18233	
16					
17					

بالعودة لملف البيانات نجد انه قد أضيف متغير جديد mah_1 وذلك لأننا طلبنا اختبار Mahalanobis فنقوم بمقارنة قيم هذا المتغير بقمية Chi-Square عند درجة حريه n-1 = 2 ومستوى معنويه مثلا 0.001 فنجد أن chi-Square = 13.8 وجميع قيم المتغير أقل من هذه القيمه لذا فانه لا يوجد قيم متطرفه متعدده.

١ـ إن ادخال عدد كبير من المتغيرات في النموذج يجعل من الصعب تحقيق شروط تطبيق الانحدار (الارتباط الذاتي والخطية والتجانس،.....

٢ـ هناك شرط لابد من تحققه في عدد القراءات وعدد المتغيرات

(وهو أن عدد القراءات عشرون ضعف لعدد المتغيرات) فإذا كان عدد المتغيرات لم يتحقق ؟

٣ـ كيف نعالج عدم معنوية بعض المتغيرات.

سنستخدم الانحدار الهرمي (خطوة خطوة) والذي يهدف أساسا إلى إيجاد علاقة بين المتغير التابع والمتغيرات المستقلة الأكثر ارتباطا به ويتم ذلك بشكل تدريجي

مثال : 11

للمثال رقم 10 جد معاملة الانحدار بطريقة Stepwise Regression

الانحدار التدريجي
Stepwise Regression

للاطلاع

الخطوات النظرية التى تقوم عليها طريقة الانحدار التدريجى

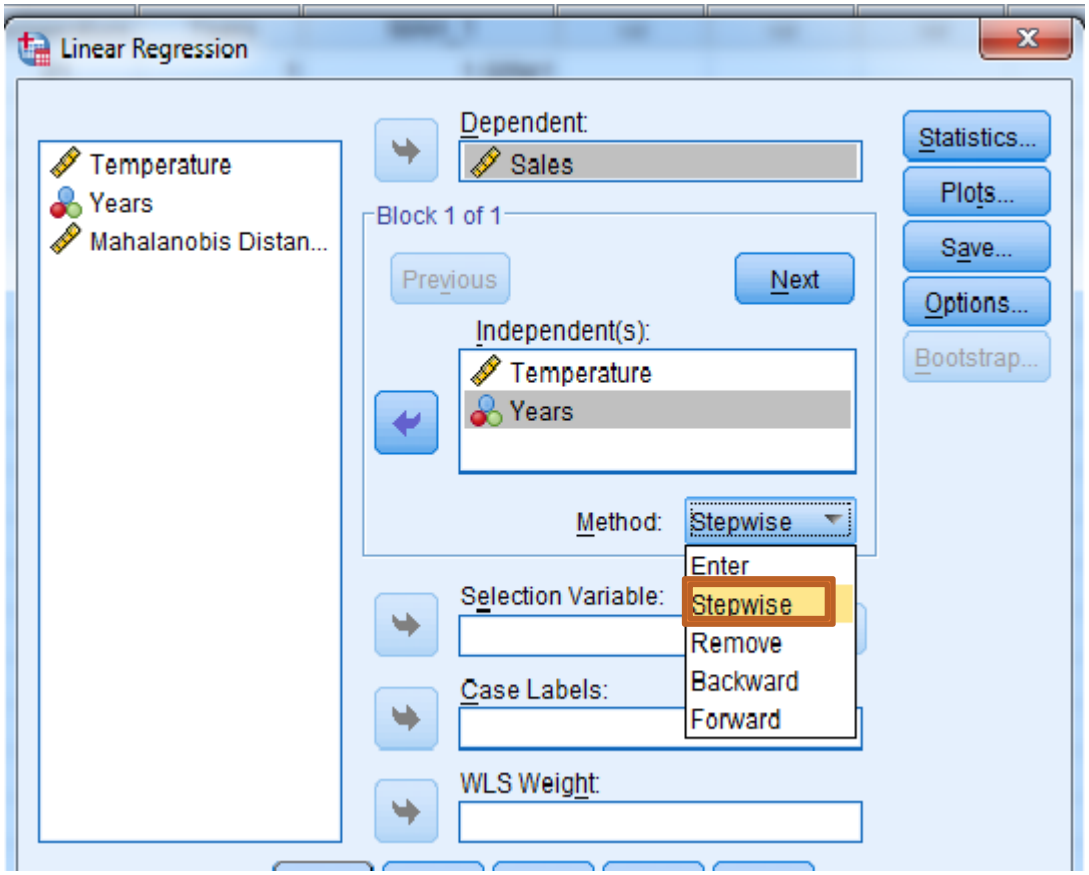
- يبدأ الانحدار التدريجى بالخطوات التالية:

1. نحسب مصفوفة الانحدار لجميع المتغيرات
2. نختار المتغير المستقل الذى له اكبر ارتباط بالمتغير التابع وندخله فى معادله الانحدار
3. نختار المتغير المستقل الثانى الذى له اكبر ارتباط بعد المتغير الذى دخل المعادله فنبحث اولا هل هناك ارتباط كبير بينه وبين المتغير الذى اختير اذا كانت الاجابه بنعم يستبعد ذلك المتغير واذا كانت بلا ندخله فى الاختيار
4. نكرر هذه العمليه مع بقيه المتغيرات المستقله مع استبعاد المتغيرات التى لها ارتباط كبير مع المتغيرات المختاره
5. تكون عملية الاضافه مجديه اذا كان هناك تأثير على معامل التحديد Coefficient of Determination (مربع معامل الارتباط) وكذلك قيمه F من جدول تحليل التباين
6. نتوقف عن الاضافه اذا لم يكن للاضافه تأثير على معامل التحديد وقيمة F المحسوبه (أو له تأثير ضعيف).

خطوات الانحدار التدريجي باستخدام SPSS

خطوات الانحدار التدريجي باستخدام SPSS

هي نفس خطوات الانحدار العياري باستثناء الخطوة التي فيها يتم اختيار الطريقة حيث سنختار Stepwise



النتائج

Variables Entered/Removed[a]			
Model	Variables Entered	Variables Removed	Method
1	Temperature	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).

الجدول الاول:

بعنوان Variables Entered/removed ويوضح المتغيرات الداخله فى المعادله وطريقة المعالجه ويتضح أن المتغير temp هو المتغير الوحيد الذى تم ادخاله فى معادله الانحدار.

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.907[a]	.822	.809	6.339

a. Predictors: (Constant), Temperature

b. Dependent Variable: Sales

الجدول الثاني:

بعنوان Model Summary ويعطى بعض المقاييس الهامه والمحسوبه من البيانات وأهمها معامل التوافق (مربع معامل الارتباط) ويستخدم للحكم على عمليه التوفيق ومنه نجد أن معادله الانحدار تمثل 80% من البيانات.

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2420.480	1	2420.480	60.228	.000[b]
	Residual	522.453	13	40.189		
	Total	2942.933	14			

a. Dependent Variable: Sales

b. Predictors: (Constant), Temperature

الجدول الثالث:

بعنوان ANOVA وهو تحليل التباين للانحدار ويتضح أن الانحدار معنوى حيث Sig. = 0.00 وهى أقل من 0.05

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-11.271-	6.465		-1.743-	.105
	Temperature	1.600	.206	.907	7.761	.000

a. Dependent Variable: Sales

الجدول الرابع:

بعنوان Coefficients ومنه يمكن ايجاد معادله الانحدار بين المبيعات ودرجة الحرارة فقط ومقدار الخطأ فى التقدير واختبار معنويه المعاملات والتقدير بفتره للمعاملات.

$$Sales = -11.271 + 1.600\,temp$$

لاحظ أن Temperature هو المتغير المعنوي في طريقة الانحدار العياري

Excluded Variables[a]

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics: Tolerance
1	Years	.208[b]	.807	.435	.227	.212

a. Dependent Variable: Sales

b. Predictors in the Model: (Constant), Temperature

الجدول الخامس:

بعنوان Excluded Variables ويعرض بيانات تخص المتغير الذى استبعد وهو سنوات الخبره

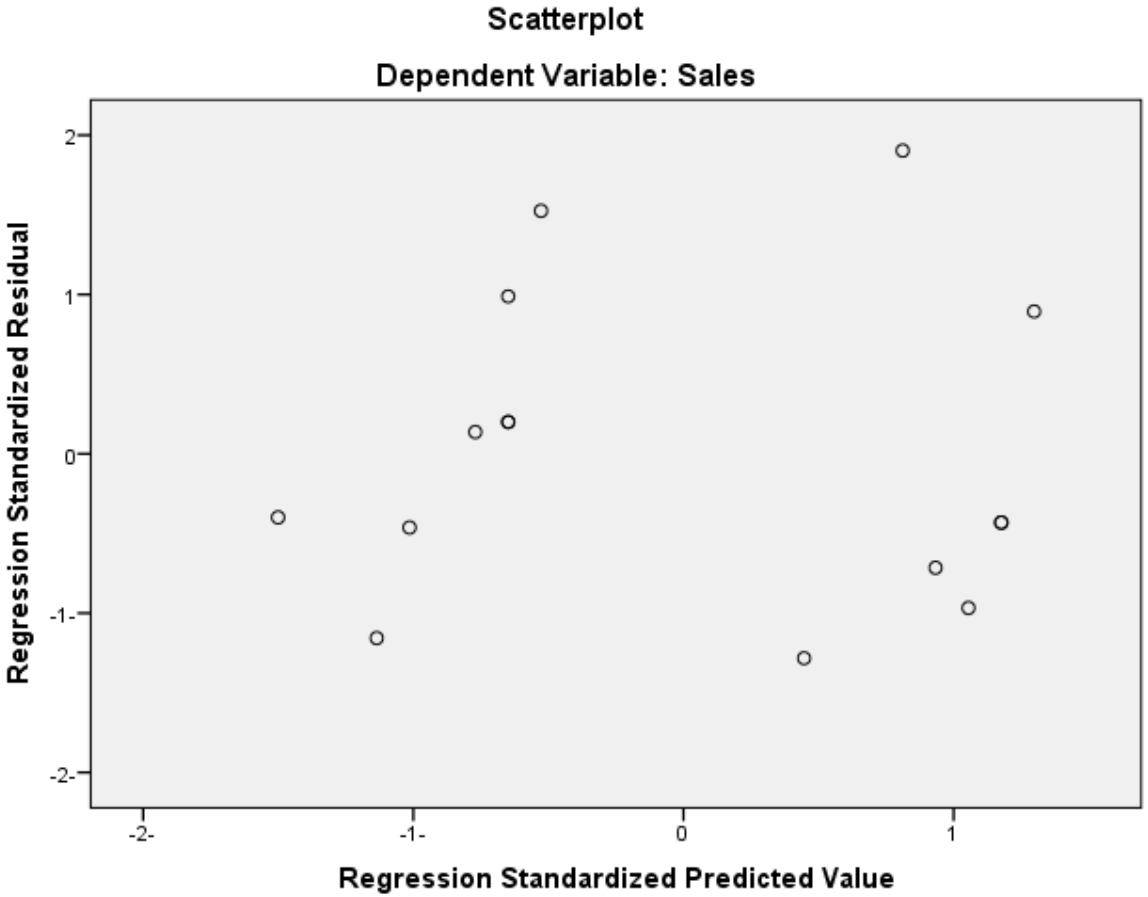
Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	17.53	54.33	37.27	13.149	15
Std. Predicted Value	-1.501-	1.298	.000	1.000	15
Standard Error of Predicted Value	1.803	3.024	2.289	.356	15
Adjusted Predicted Value	18.28	53.28	37.35	13.062	15
Residual	-8.134-	12.066	.000	6.109	15
Std. Residual	-1.283-	1.903	.000	.964	15
Stud. Residual	-1.338-	2.022	-.006-	1.027	15
Deleted Residual	-8.850-	13.613	-.088-	6.947	15
Stud. Deleted Residual	-1.385-	2.346	.021	1.092	15
Mahal. Distance	.199	2.253	.933	.587	15
Cook's Distance	.001	.262	.068	.072	15
Centered Leverage Value	.014	.161	.067	.042	15

a. Dependent Variable: Sales

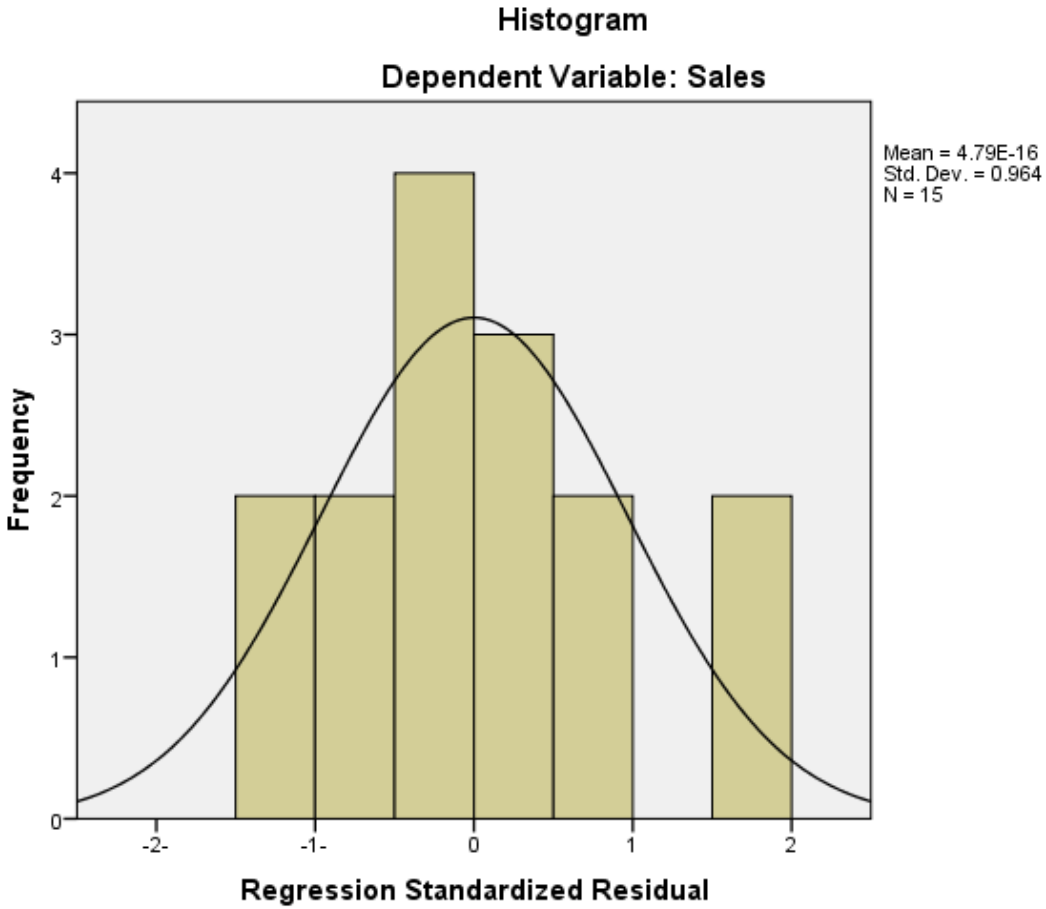
الجدول السادس:

بعنوان Residuals Statistics ويعرض بيانات خاصه بتحليل البواقى.



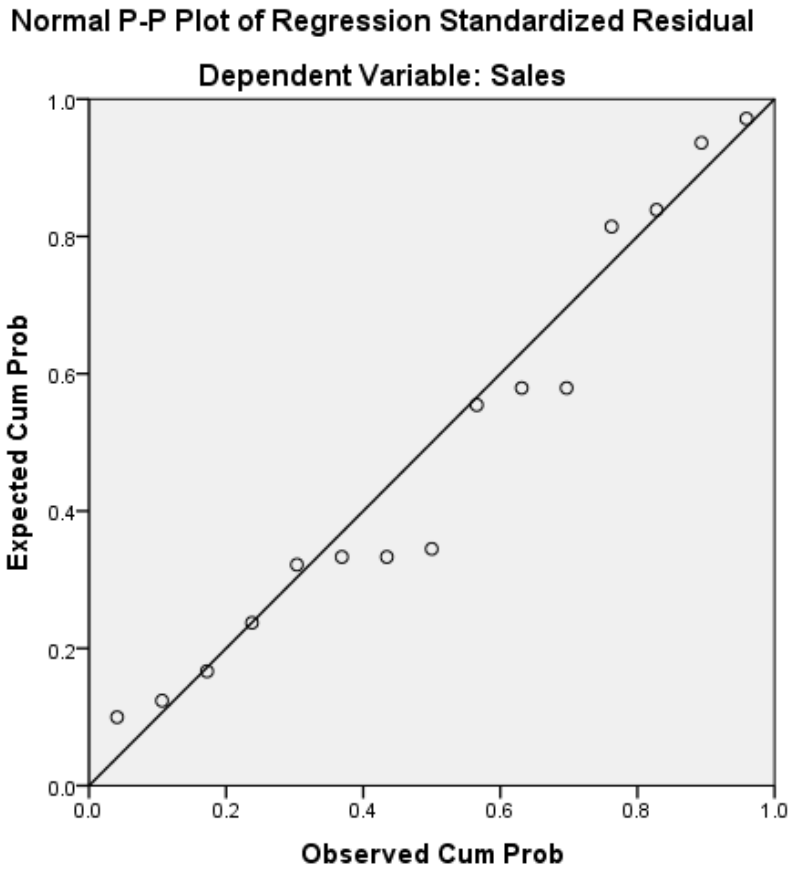
الشكل البيانى الاول:

يمثل شكل الانتشار للبواقى مع القيم المتوقعه ومنه يتضح عدم وجود نمط معين للنقاط فى الشكل وهذا يتسق مع شرط الخطيه.



الشكل البيانى الثالث:

هو المدرج التكرارى ويستخدم للتعرف هل البيانات تتوزع حسب التوزيع الطبيعى أم لا؟



الشكل البيانى الثانى:

يختبر هل البواقى تتبع التوزيع الطبيعى ام لا؟

ومن الشكل نجد أن النقاط تتجمع حول الخط وبالتالى فان البيانات (البواقى) تتوزع حسب التوزيع الطبيعى.

مثال : 12

للمثال رقم 10 جد معاملة الانحدار بطريقة Hierarchical Regression

الانحدار الهرمي
Hierarchical Regression

الانحدار الهرمي
Hierarchical Regression

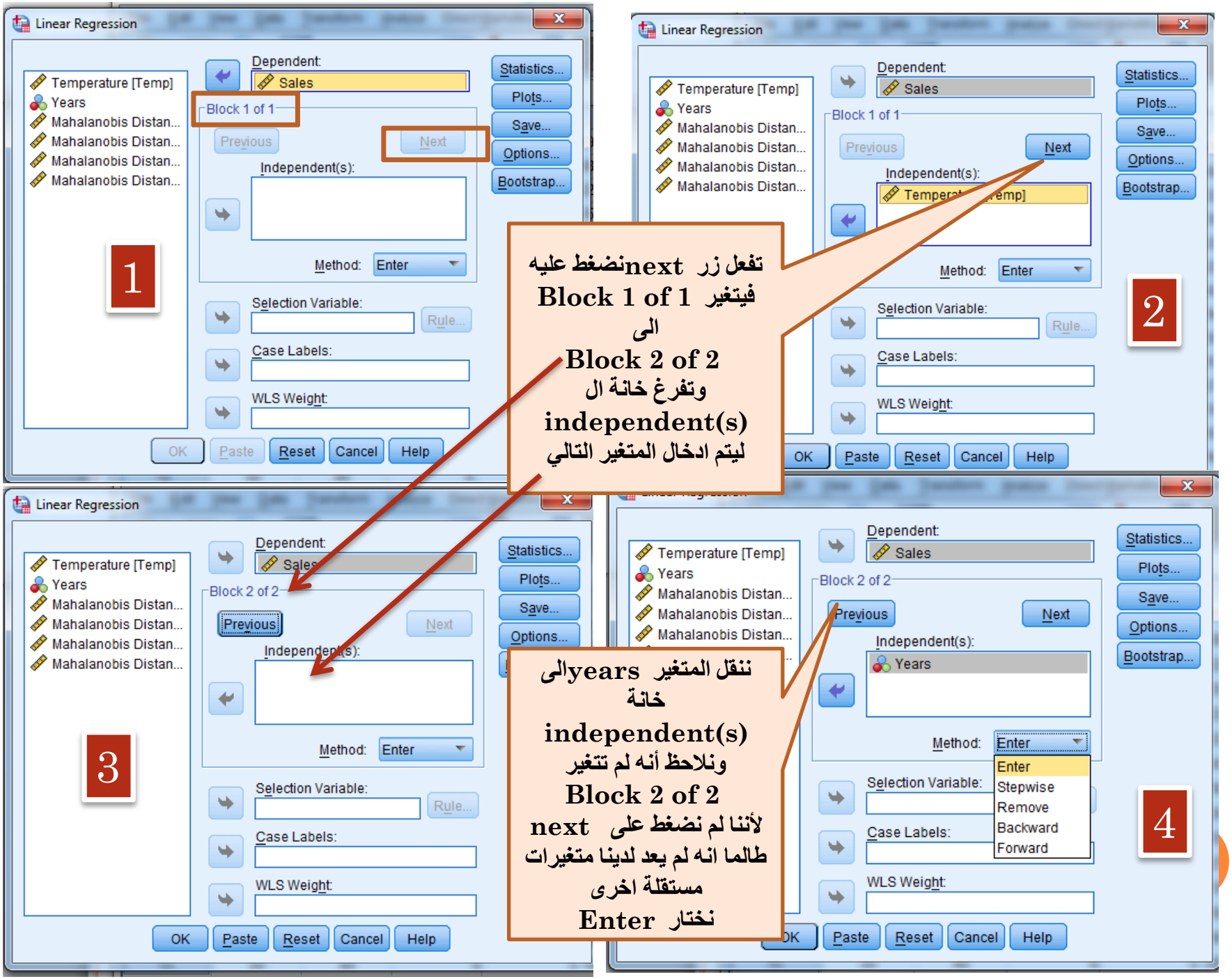
يتم فى هذه الحاله استخدام الطريقة العياريه فى تحديد معادلة الانحدار ولكن ليس لكل المتغيرات بل ندخل المتغيرات تباعا فندخل أول متغير ثم يليه المتغير التالى وهكذا.

يتم التخلص من كل الاختيارات السابقة وابسط الطرق هي حفظ البيانات ومن ثم قفل البرنامج وفتحه من جديد

خطوات الانحدار الهرمي:

Analyze----- Regression------ Linear

لتظهر النافذة التالية


Linear Regression
Temperature [Temp]
Years
Mahalanobis Distan...
Dependent:
Sales
Block 1 of 1
Previous
Next
Independent(s):
Method: Enter
Selection Variable:
Rule...
Case Labels:
WLS Weight:
Statistics...
Plots...
Save...
Options...
Bootstrap...
OK
Paste
Reset
Cancel
Help
1
2
3
4
تفعل زر next نضغط عليه
فيتغير Block 1 of 1
الى
Block 2 of 2
وتفرغ خانة ال
independent(s)
ليتم ادخال المتغير التالي
ننقل المتغير years الى
خانة
independent(s)
ونلاحظ أنه لم تتغير
Block 2 of 2
لأننا لم نضغط على next
طالما انه لم يعد لدينا متغيرات
مستقلة اخرى
نختار Enter
Enter
Stepwise
Remove
Backward
Forward

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Temperature[b]	.	Enter
2	Years[b]	.	Enter

a. Dependent Variable: Sales

b. All requested variables entered.

الجدول الاول:

بعنوان Variables Entered/Removed يوضح المتغيرات التى ادخلت للنموذج وطريقة الاختيار

Model Summary[c]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.907[a]	.822	.809	6.339	.822	60.228	1	13	.000
2	.912[b]	.832	.804	6.426	.009	.652	1	12	.435

a. Predictors: (Constant), Temperature

b. Predictors: (Constant), Temperature, Years

c. Dependent Variable: Sales

الجدول الثانى:

بعنوان Model Summary ويعطى ملخص عن النموذج لاحظ كلا من R square Change, F, Sig. نجد أن اضافة المتغير الأول temp معنوى وله جدوى بخلاف المتغير الثانى years

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2420.480	1	2420.480	60.228	.000[b]
	Residual	522.453	13	40.189		
	Total	2942.933	14			
2	Regression	2447.388	2	1223.694	29.633	.000[c]
	Residual	495.546	12	41.295		
	Total	2942.933	14			

a. Dependent Variable: Sales

b. Predictors: (Constant), Temperature

c. Predictors: (Constant), Temperature, Years

الجدول الثالث:

بعنوان ANOVA ويعطى تحليل التباين لكل خطوه ادخال ومنه يتضح معنويه الانحدار فى كل خطوه.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-11.271-	6.465		-1.743-	.105		
	Temperature	1.600	.206	.907	7.761	.000	1.000	1.000
2	(Constant)	-6.993-	8.428		-.830-	.423		
	Temperature	1.275	.454	.722	2.807	.016	.212	4.721
	Years	1.907	2.362	.208	.807	.435	.212	4.721

a. Dependent Variable: Sales

الجدول الرابع:

بعنوان Coefficients ومنه يحسب خط الانحدار المقترح (المقدر) واختبار معنويه المعاملات وخطأ التقدير. ومن الجدول نجد أن:

1. الخطوة الاولى تم ادخال درجات الحرارة temp وكان الانحدار معنوى.
2. الخطوة الثانية تم ادخال سنوات الخبره مع درجات الحرارة فكانت غير معنويه

وعلى ذلك يكون أفضل خط مقدر هو

$$Sales = -11.271 + 1.600\,temp$$

Excluded Variables[a]

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics		
						Tolerance	VIF	Minimum Tolerance
1	Years	.208[b]	.807	.435	.227	.212	4.721	.212

a. Dependent Variable: Sales

b. Predictors in the Model: (Constant), Temperature

الجدول الخامس:

بعنوان Excluded Variables وهو خاص بالمتغيرات المستبعده

أما بقية النتائج فهي مشابهة لما سبق

مثال 13 :

حيث أن متغير sales لم يكن ارتباطه الخطي ب year معنوي فماهي العلاقة بينهما إذن؟

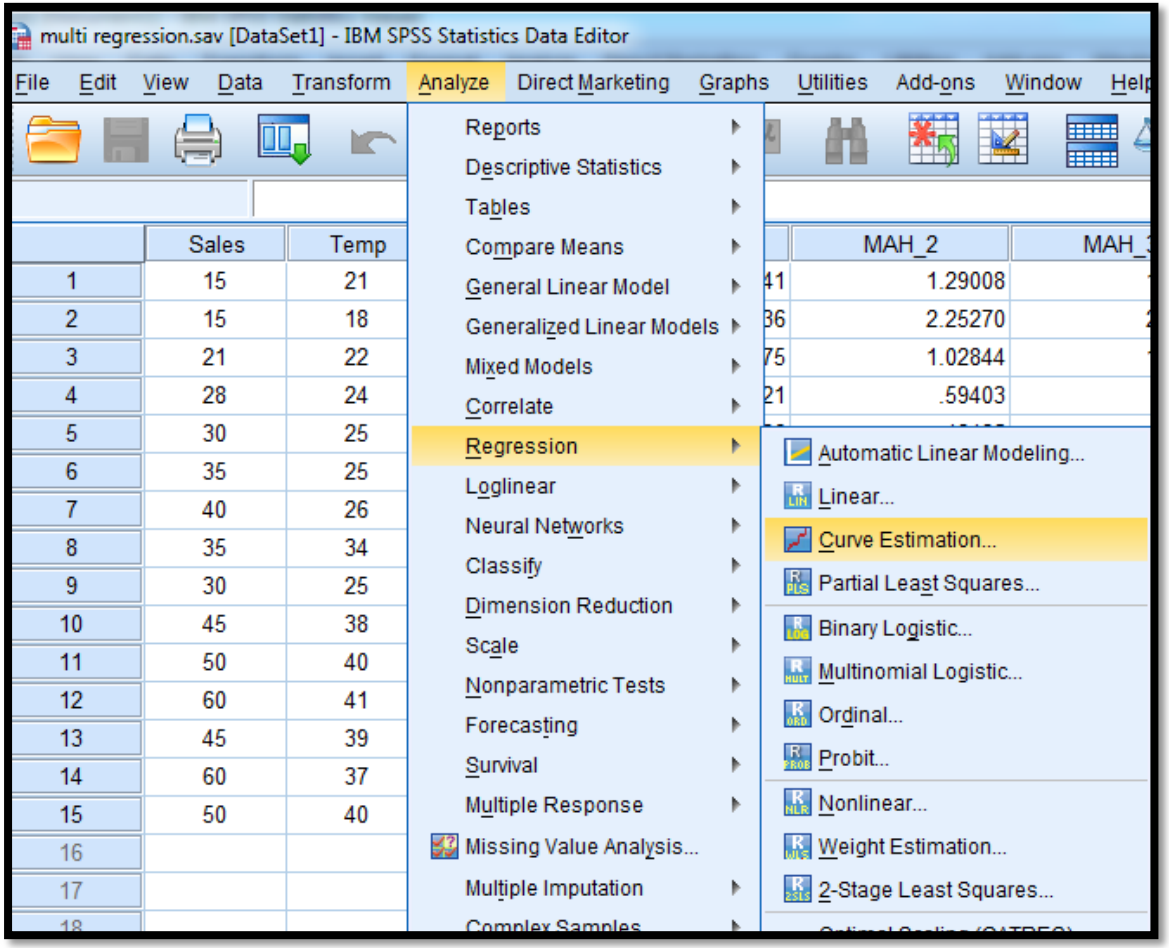
الانحدار غير الخطي
Curve Regression

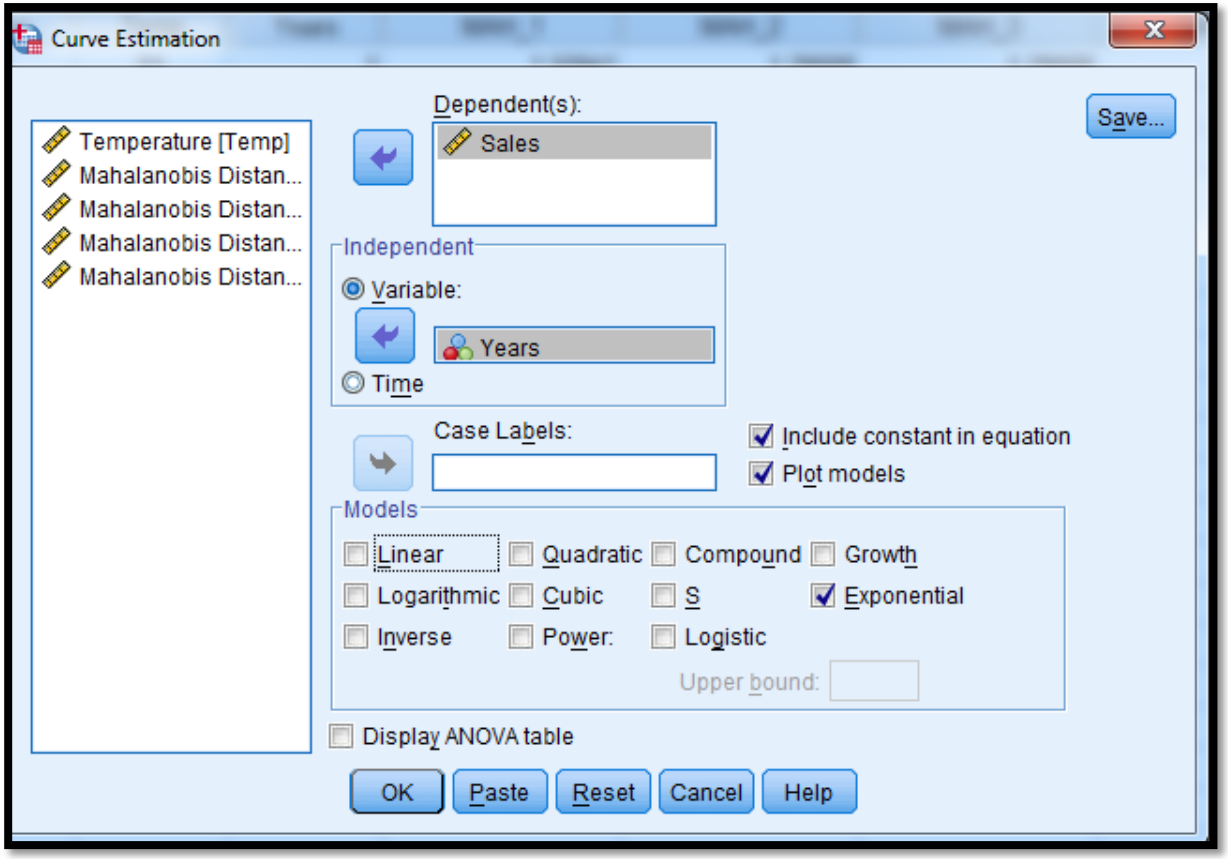
الانحدار غير الخطي
Curve Regression

- بفرض أن لدينا متغيرين وأن العلاقة بينهما غير خطيه. ونقصد بكلمة غير خطيه أن العلاقة تأخذ أى صورة من صورة العلاقات الغير خطيه مثل الدرجة الثانيه أو الثالثة ... الخ.
- وقد تكون العلاقة بينهما أسيه أو لوغاتيميه أو ... الخ.
- وتساعد الحزمة SPSS على ايجاد معادلة الانحدار فى حالة الانحدار الغير خطى وكذلك اجراء الاختبارات المعنويه اللازمة للتقدير وذلك لعدة نماذج غير خطيه منها:

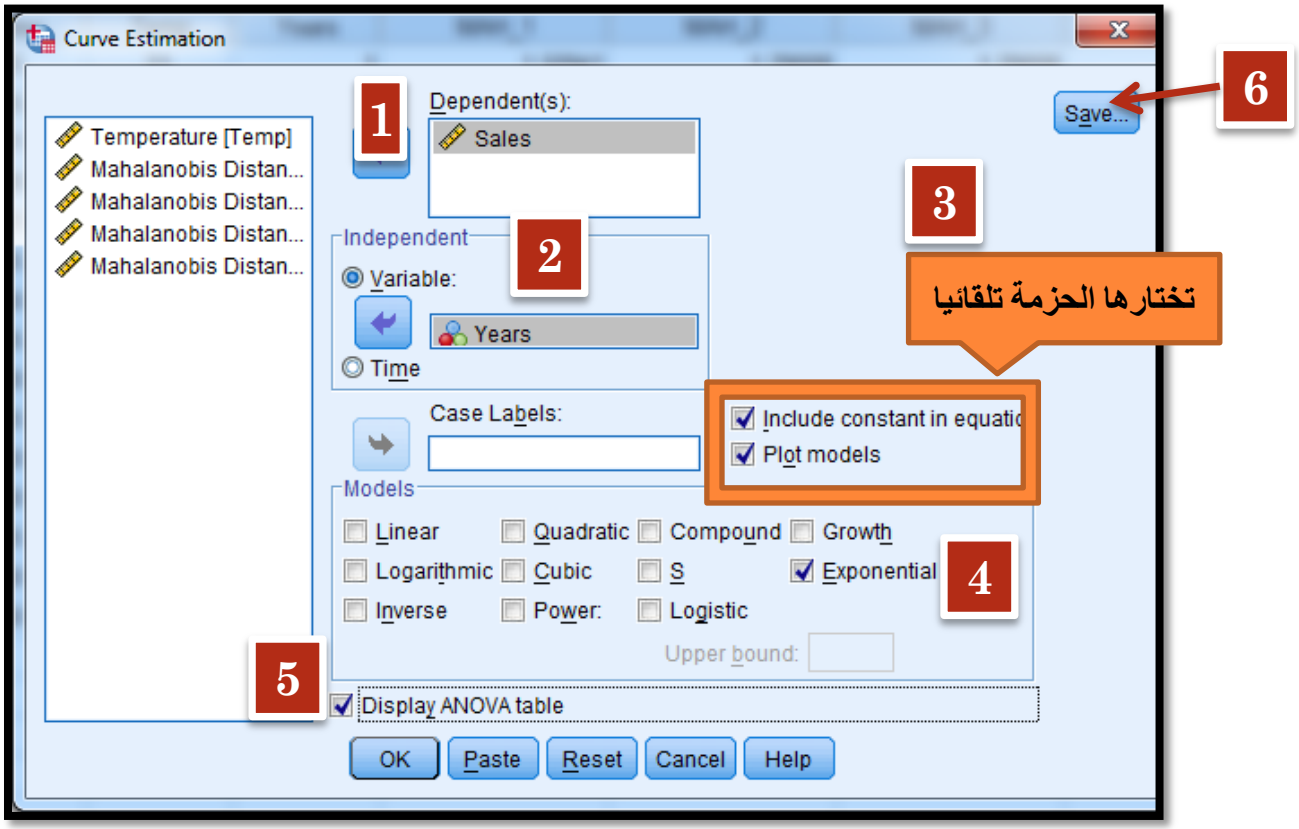
1. العلاقات الخطيه Linear ومن الدرجة الثانيه Quadratic والثالثة Cubic
2. العلاقات الآسيه Exponential
3. العلاقات اللوغاتيميه Logarithmic
4. العلاقات المعكوسه Inverse والعلاقات المركبه Compound
5. العلاقات اللوجستيه Logistic وعلاقة القوة Power

الخطوات:


multi regression.sav [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help
Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Complex Samples
Automatic Linear Modeling...
Linear...
Curve Estimation...
Partial Least Squares...
Binary Logistic...
Multinomial Logistic...
Ordinal...
Probit...
Nonlinear...
Weight Estimation...
2-Stage Least Squares...
Sales Temp MAH_2
1 15 21 1.29008
2 15 18 2.25270
3 21 22 1.02844
4 28 24 .59403
5 30 25
6 35 25
7 40 26
8 35 34
9 30 25
10 45 38
11 50 40
12 60 41
13 45 39
14 60 37
15 50 40
16
17


Curve Estimation
Temperature [Temp]
Mahalanobis Distan...
Mahalanobis Distan...
Mahalanobis Distan...
Mahalanobis Distan...
Dependent(s):
Sales
Save...
Independent
Variable:
Years
Time
Case Labels:
Include constant in equation
Plot models
Models
Linear
Quadratic
Compound
Growth
Logarithmic
Cubic
S
Exponential
Inverse
Power:
Logistic
Upper bound:
Display ANOVA table
OK
Paste
Reset
Cancel
Help

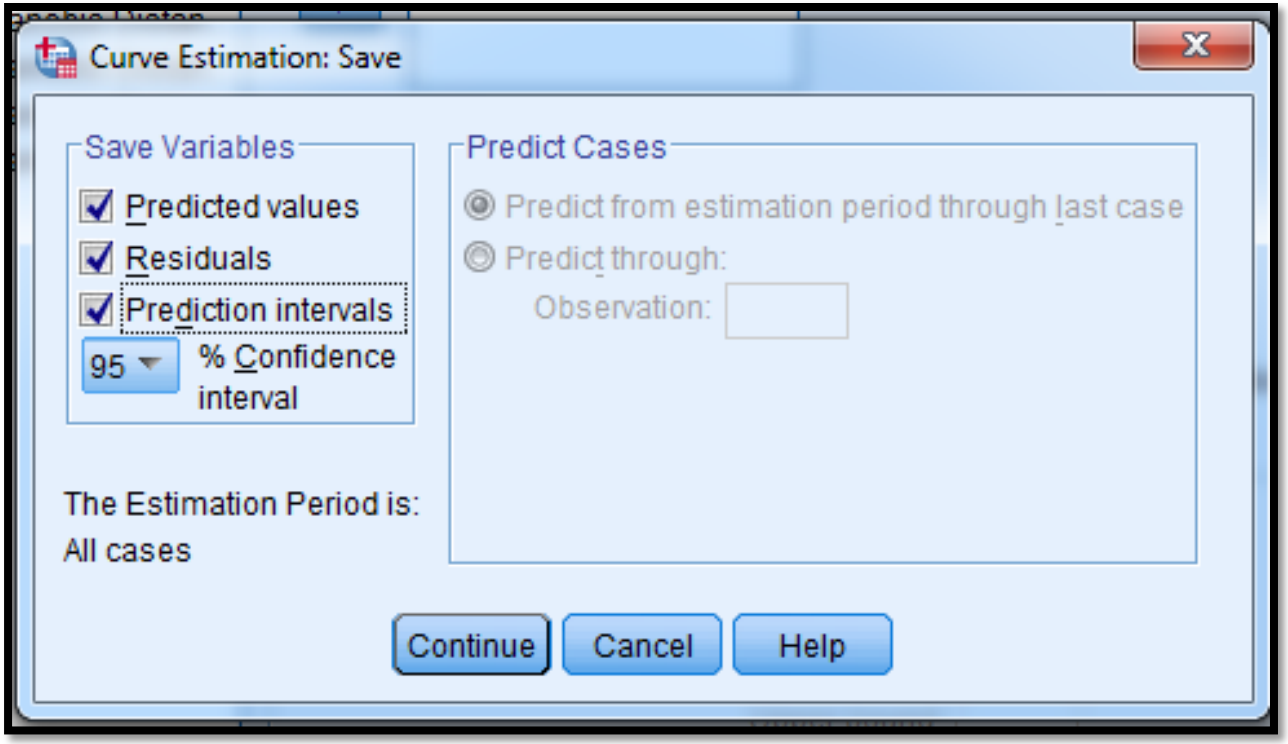
تظهر شاشة جديده بعنوان Curve Estimation ننقل المتغير التابع Sales لخانة Dependent والمتغير المستقل years لخانة Independent مع ملاحظه الاختيار Variable



نختار شكل المنحنى المطلوب وليكن Exponential

نلاحظ ان الامرين Include constant in Equation, Plot Models تم اختيارهما من قبل الحزمه.

تظهر شاشة جديده بعنوان Curve Estimation: Save نختار منها Predicted Values, Residuals, Predicted Interval ونلاحظ حدود الثقة المختارة هى 95% ويمكن تعديلها



نضغط على Continue فنعود للشاشه السابقه

نضغط على Ok فتظهر النتائج التاليه:

Model Description

Model Name		MOD_1
Dependent Variable	1	Sales
Equation	1	Exponential[a]
Independent Variable		Years
Constant		Included
Variable Whose Values Label Observations in Plots		Unspecified

a. The model requires all non-missing values to be positive.

الجدول الأول:

بعنوان Model Description ويحتوى على معلومات خاصه بالنموذج المحدد

اسم المتغير التابع والمستقل ونوع النموذج.

Sales

Exponential

Model Summary

R	R Square	Adjusted R Square	Std. Error of the Estimate
.832	.692	.669	.256

The independent variable is Years.

الجدول الرابع:

بعنوان Model Summary ويعطى معامل الارتباط ومربع معامل الارتباط

وايضا مربع معامل الارتباط المصحح وخطأ التقدير.

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Regression	1.916	1	1.916	29.237	.000
Residual	.852	13	.066		
Total	2.767	14			

The independent variable is Years.

الجدول الخامس:

بعنوان ANOVA ويعطى تحليل التباين لنموذج الانحدار المحدد (الانحدار الآسى) ونجد أن الانحدار معنوى.

Coefficients

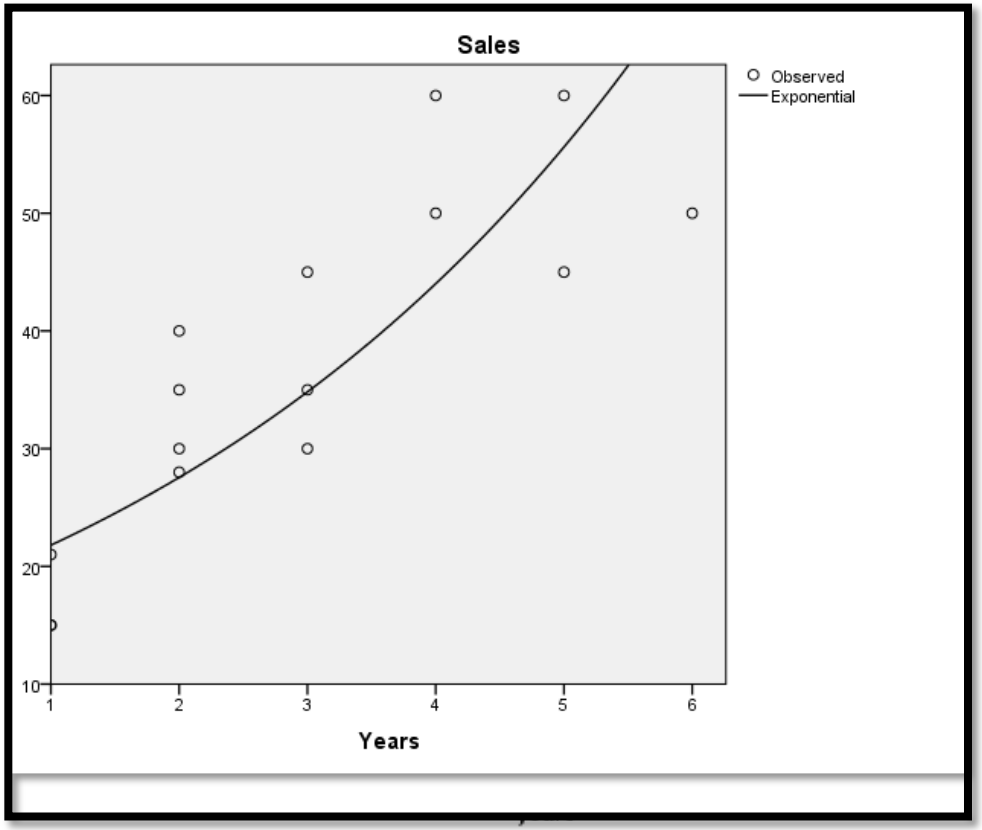
	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
Years	.234	.043	.832	5.407	.000
(Constant)	17.250	2.470		6.983	.000

The dependent variable is ln(Sales).

الجدول السادس:

بعنوان Coefficients ويعطى قيم معاملات نموذج الانحدار ومعنوية كل معامل على حده.

$$sales = 17.25\, e^{0.23419\, years}$$



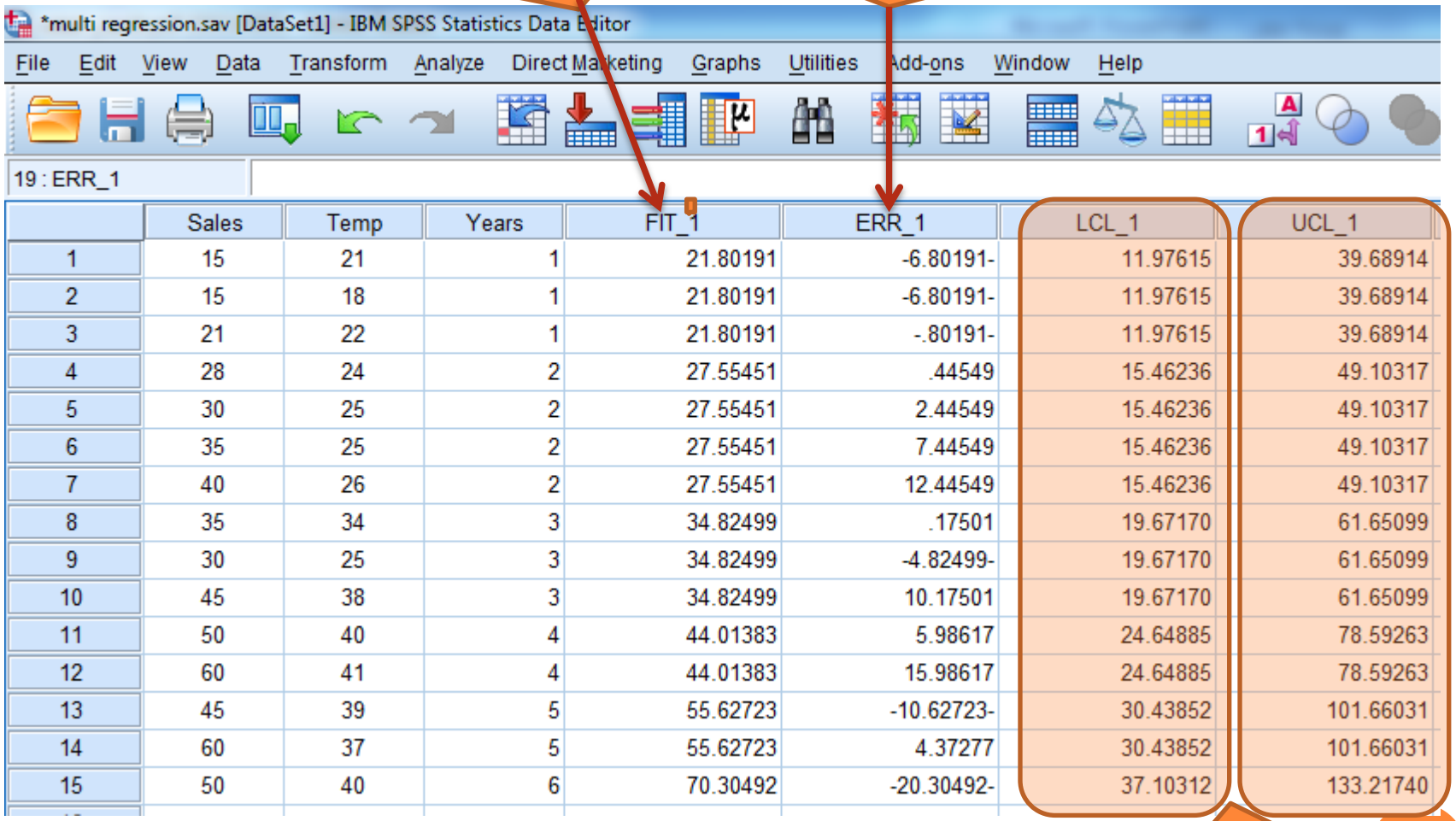
الشكل البياني: شكل الانتشار للبيانات والمنحنى المقدر للبيانات من نموذج الانحدار.

- بالعودة للملف regression.sav نجد أن الحزمة قد أضافت بعض المتغيرات الجديده

بالعودة الى ملف البيانات سنجد اضافة اربعة اعمدة

القيم المقدرة للمتغير Sales من النموذج
مثلا اول قيمة = (1) 17.25e^0.23419

الاخطاء
اول خطأ =15-21.80191



	Sales	Temp	Years	FIT_1	ERR_1	LCL_1	UCL_1
1	15	21	1	21.80191	-6.80191-	11.97615	39.68914
2	15	18	1	21.80191	-6.80191-	11.97615	39.68914
3	21	22	1	21.80191	-.80191-	11.97615	39.68914
4	28	24	2	27.55451	.44549	15.46236	49.10317
5	30	25	2	27.55451	2.44549	15.46236	49.10317
6	35	25	2	27.55451	7.44549	15.46236	49.10317
7	40	26	2	27.55451	12.44549	15.46236	49.10317
8	35	34	3	34.82499	.17501	19.67170	61.65099
9	30	25	3	34.82499	-4.82499-	19.67170	61.65099
10	45	38	3	34.82499	10.17501	19.67170	61.65099
11	50	40	4	44.01383	5.98617	24.64885	78.59263
12	60	41	4	44.01383	15.98617	24.64885	78.59263
13	45	39	5	55.62723	-10.62723-	30.43852	101.66031
14	60	37	5	55.62723	4.37277	30.43852	101.66031
15	50	40	6	70.30492	-20.30492-	37.10312	133.21740

الحدود العليا والدنيا لفترات الثقة للمتغير التابع عند كل قيمة للمستقل

المراجـــــــــــــــع

١)ابو سريع، رضا.(٢٠٠٤).تحليل البيانات باستخدام برنامج spss ،دار الفكر،عمان.

٢)البشير، سعد.(٢٠٠٣)،دليلك الى البرنامج الاحصائي spss ،المعهد العربي للتدريب والبحوث الاحصائية ،العراق

٣)الاختبارات الاحصائية البارامترية واللابارامترية باستخدام SPSSالاصدار السابع عشر ، د.سوسن ابراهيم أبو العلا شلبي، جامعة الملك سعود

٤)الارتباط والانحدار د. كامل أبو ضاهر ، الجامعة الاسلامية –غزة

٥)-بعض المعادلات الإحصائية المستخدمة في تحديد عينة البحث، الهزاع، هزاع محمد. فسيولوجيا الجهد البدني: الأسس النظرية والإجراءات المعملية للقياسات الفسيولوجية. الفصل الرابع. كتاب مقدم للنشر.

٦)Multiple Regression, Abdelfatah Mustafa, Mansuora University

http://www.nca.umich.edu/sample_size_chart
http://www.surveysystem.com/sscalc.htm