الاختبارات غير المعلمية Nonparametric Tests

في بعض الحالات قد لا تتوافر في المجتمع موضع الدراسة أن يكون توزيع هذا المجتمع له توزيع طبيعي أو يقترب منه، لذلك فإن استخدام الاختبارات المعلمية في مثل هذه الحالات قد يؤدي إلى نتائج غير دقيقة، كذلك يفترض أن تكون بيانات الظاهرة موضع الدراسة دقيقة، ولكن في بعض الأحيان يتعذر أخذ قياسات عددية دقيقة على بعض الظواهر، لذلك فإننا نستخدم طرق غير معلمية لا تعتمد على شروط معينة تتعلق بتوزيع المجتمع ولا تحتاج إلى قياسات دقيقة.

مزايا استخدام الاختبارات غيرالمعلمية:

١. سهولة العمليات الحسابية المستخدمة.
٢. لا تحتاج إلى شروط كثيرة لذلك فإن إمكانية إساءة استعمالها قليلة جداً.
٣. تستخدم عندما لا تتحقق الشروط اللازمة لتطبيق الاختبارات المعلمية مثل أن يكون توزيع المجتمع طبيعياً.
٤. تستخدم في حالة صعوبة الحصول على بيانات دقيقة.
٥. لا يتطلب استخدامها معرفة دقيقة في مجال الرياضيات أو الإحصاء.
٦. لا تشترط استخدامها أن يكون حجم العينات كبيراً، لذلك فإن عملية جمع البيانات في هذه الحالة توفر الوقت والمجهود والتكلفة.

عيوب استخدام الاختبارات غيرالمعلمية:

١. تستخدم أحياناً في الحالات التي يجب استخدام الاختبارات المعلمية وذلك لسهولة استخدامها.
٢. صعوبة الحصول على توزيع دوال الاختبار المستخدمة في هذه الاختبارات.

يمكن استخدام الاختبارات غيرالمعلمية في الحالات التالية:

١. للحصول على قرار سريع.
٢. إذا كانت البيانات المتوفرة عن ظاهرة ما لا تتفق مع الاختبارات المعلمية.
٣. إذا كانت الشروط المطلوب توافرها في الاختبار المعلمي غير متحققة.

سنعرض فيما يلي استخدام برنامج SPSS في الاختبارات غيرالمعلمية التالية:

١. استخدام اختبار كولمجروف – سمرنوف "One-Sample Kolmogorov-Smirnov Test" لمعرفة ما إذا كانت البيانات تتبع التوزيع الطبيعي.
٢. اختبار الإشارة "Sign Test"لاختبار فرضيات حول متوسط مجتمع واحد.
٣. اختبار ويلكوكسن "Wilcoxon Test" لاختبار فرضيات حول مقارنة متوسطي مجتمعين في حالة العينات المرتبطة.

٤. اختبار مان – وتني "Mann Whitney Test" لاختبار الفرضيات حول الفرق بين متوسطي مجتمعين في حالة العينات المستقلة.
٥. اختبار كروسكال – والاس "Kruskal-Wallis Test"لاختبار فرضيات لمقارنة متوسطات عدة مجتمعات مستقلة (تحليل التباين في حالة العينات المستقلة).
٦. اختبار فريدمان "Friedman Test"الذي يعالج موضوع تحليل التباين في حالة المشاهدات المتكررة (Repeated Measures)أو العينات المرتبطة .

اختبار التوزيع الطبيعي: كولمجروف– سمرنوف

مثال(١):

تمثل البيانات التالية كمية الإنتاج بالطن لسلعة ما في الأسبوع في أحد المصانع:

74 83 94 68 76 60 90 70 80 90 80 68 82 79 65 50 70
60 92 82 68 93 71 86 92 90 80 82 65 76

المطلوب: استخدم اختبار كولمجروف – سمرنوف لمعرفة أن البيانات السابقة لها توزيع طبيعي أم لا مستخدماً مستوى دلالة $\alpha = .05$.

الحل العملي:

One-Sample Kolmogorov-Smirnov Test

		QUANTITY
N		30
Normal Parameters [a,b]	Mean	77.2000
	Std. Deviation	11.3058
Most Extreme Differences	Absolute	.105
	Positive	.075
	Negative	-.105
Kolmogorov-Smirnov Z		.573
Asymp. Sig. (2-tailed)		.898

a. Test distribution is Normal.

b. Calculated from data.

من الجدول السابق: Sig.=.898 لذلك لا يمكن رفض فرضية العدم القائلة بأن كمية الإنتاج لها توزيع طبيعي وذلك على مستوى دلالة $\alpha = .05$.

اختبار ويلكوكسن "Wilcoxon Test"

مثال(٢):

لمعرفة تأثير إشارة ضوئية جديدة، تم الحصول على البيانات التالية التي تمثل عدد الحوادث في ١٢ مفترق خطر خلال أربعة أسابيع قبل وبعد تركيب الإشارة الضوئية:

(2,1), (3,2), (2,0), (1,3), (2,1), (6,3), (5,3), (4,1), (5,2), (3,2), (2,3), (4,2).

اختبر الفرضية المبدئية القائلة أنه لا يوجد تأثير للإشارة الضوئية الجديدة مستخدماً مستوى دلالة $\alpha = .05$.

الحل العملي:

من الجدول السابق Sig.=.039 لذلك نرفض فرضية العدم القائلة بأنه لا يوجد تأثير للإشارة الضوئية الجديدة مستخدماً مستوى دلالة $\alpha = .05$، ونستنتج أنه يوجد تأثير للإشارة الضوئية الجديدة، بمعنى أن معدل عدد الحوادث قد تناقص بعد تركيب الإشارة الضوئية الجديدة.

اختبار مان – وتني "Mann Whitney Test"

مثال(٣):

قامت إحدى الشركات بتدريب بعض عمالها على العمل على آلات جديدة وردت إلى مصانع الشركة، واستخدمت برنامجين للتدريب، البرنامج الأول محاضرات نظرية لمدة أسبوعين ومن ثم القيام بالتدريب العملي، والبرنامج الثاني محاضرات نظرية تتبعها تطبيقات عملية في نفس اليوم ولمدة أسبوعين. وكان الزمن اللازم للمتدربين لاكتساب المهارات المطلوبة مقدرة بالأيام كما يلي:

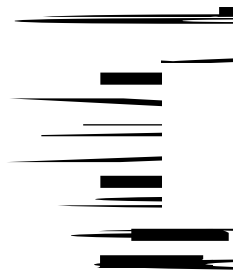
البرنامج الأول	البرنامج الثاني
٤٠	٢٩
٤٤	٢٧
٣٣	٣٢
٢٦	٢٥
٣١	٢٧
٢٩	٢٨
٣٤	٣١
٣١	٢٣
٣٨	٣٧
٣٣	٢٨
٤٢	٢٢
٣٥	٣١
	٢٤

هل تستطيع أن تستنتج أن البرنامج الثاني أكثر فاعلية من البرنامج الأول على مستوى دلالة $\alpha = .05$؟

الحل العملي:

Ranks

	FACTOR	N	Mean Rank	Sum of Ranks
TIME	first	12	17.54	210.50
	second	13	8.81	114.50
	Total	25		



من النتائج السابقة: Sig.=.002 لذلك نرفض فرضية العدم القائلة بأنه لا يوجد فرق بين فاعلية البرنامجين، ونستنتج بأنه يوجد فرق بين فاعلية البرنامجين، أو يمكن القول بأن البرنامج الثاني أكثر فاعلية من البرنامج الأول على مستوى دلالة $\alpha = .05$.

اختبار كروسكال – والاس "Kruskal-Wallis Test"

مثال(٤):

البيانات التالية تمثل درجات طلاب مساق الإحصاء التحليلي بكلية التجارة في الجامعة الإسلامية باستخدام ثلاثة أساليب مختلفة:

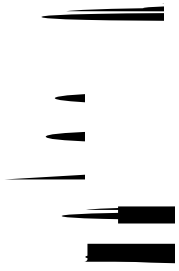
الأسلوب الأول	الأسلوب الثاني	الأسلوب الثالث
٨٦	٨٢	٧٥
٨١	٦٦	٧٨
٨٤	٦٩	٦١
٧١	٧٢	٦٩
٨١	٦٧	٧٥
٨٨	٦٨	
٧٩	٧٧	
٧٧		

هل تعطي هذه البيانات دليلاً كافياً على وجود فروق معنوية بين متوسطات درجات الطلاب باستخدام الأساليب الثلاثة السابقة على مستوى دلالة $\alpha = .05$؟

الحل العملي:

Ranks

	FACTOR	N	Mean Rank
MARK	Method_1	8	15.06
	Method_2	7	7.29
	Method_3	5	7.70
	Total	20	



من النتائج السابقة $\chi^2 = 7.968$ ، **Sig.=.019 لذلك نرفض فرضية العدم القائلة بأنه لا يوجد فرق بين درجات الطلاب باستخدام الأساليب الثلاثة ، ونستنتج بأنه يوجد فرق بين درجات الطلاب في مساق الإحصاء التحليلي باستخدام الأساليب الثلاثة على مستوى دلالة** $\alpha = .05$.

اختبار فريدمان "Friedman Test"

مثال(٥):

البيانات التالية تمثل زمن الشفاء (مقدراً بالأيام) من مرض معين عند تناول المرضى أربعة أنواع مختلفة من الأدوية.

النوع الأول	النوع الثاني	النوع الثالث	النوع الرابع
١٠	٧	١١	١٣
٨	١٣	٦	١٠
٧	١٥	١١	٩
١١	١١	٩	١٤
٩	١٢	٨	١١
٧	٨	٧	١٢
٨	١٤	٥	١٠
١١	١٠	١٠	١٣

هل تعطي هذه البيانات دليلاً كافياً على وجود فروق معنوية بين متوسطات الزمن باستخدام أنواع الأدوية الأربعة على مستوى دلالة $\alpha = .05$؟

الحل العملي:

Ranks

	Mean Rank
TYPE_1	2.00
TYPE_2	3.00
TYPE_3	1.63
TYPE_4	3.38

Test Statistics[a]

N	8
Chi-Square	10.130
df	3
Asymp. Sig.	.017

a. Friedman Test

من النتائج السابقة: $\chi^2 = 10.13$، **Sig.=.017 لذلك نرفض فرضية العدم القائلة بأنه لا يوجد فرقاً بين أنواع الأدوية الأربعة، ونستنتج بأنه يوجد فرق بين تلك الأنواع السابقة على مستوى دلالة** $\alpha = .05$.

الارتباط الخطي البسيط Simple Linear Regression

في معظم التطبيقات العملية نجد أن هناك علاقة بين متغيرين (أو أكثر)، فمثلاً نجد أن هناك علاقة وارتباط بين درجة الطالب وعدد ساعات الدراسة. يوجد نوعان من المتغيرات هما:

المتغير التابع Dependent (Response) Variable: هو المتغير الذي يقيس نتيجة دراسة ما، وعادة يرمز له بالرمز Y.

المتغير المستقل Independent (Explanatory) Variable:

هو المتغير الذي يُفسِّر أو يسبب التغيرات في المتغير التابع، أي هو الذي يؤثر في تقدير قيمة المتغير التابع، وعادة يرمز له بالرمز X. فمثلاً عدد أيام الغياب X و درجة الطالب في الإحصاء Y، العُمر X والإصابة بضغط الدم Y.

في بعض التطبيقات العملية يكون لدينا أكثر من متغيرين تحت الدراسة، فمثلاً قد توجد علاقة خطية بين ضغط الدم وكل من العُمر والوزن، ويسمى الارتباط في هذه الحالة الارتباط الخطي المتعدد.

عند دراسة العلاقة بين متغيرين X, Y فإن شكل الانتشار Scatter plot يمكن أن يوضح طبيعة هذه العلاقة، وتكون العلاقة بين X, Y قوية جداً إذا وقعت معظم نقاط شكل الانتشار على منحنى أو خط مستقيم، وتكون ضعيفة كلما تناثرت نقاط شكل الانتشار حول منحنى أو خط مستقيم يمر بتلك النقاط.

معامل الارتباط Correlation Coefficient:

هو مقياس لدرجة العلاقة بين المتغيرين X, Y ويرمز له بالرمز r، ويحقق معامل الارتباط الخطي المتباينة:

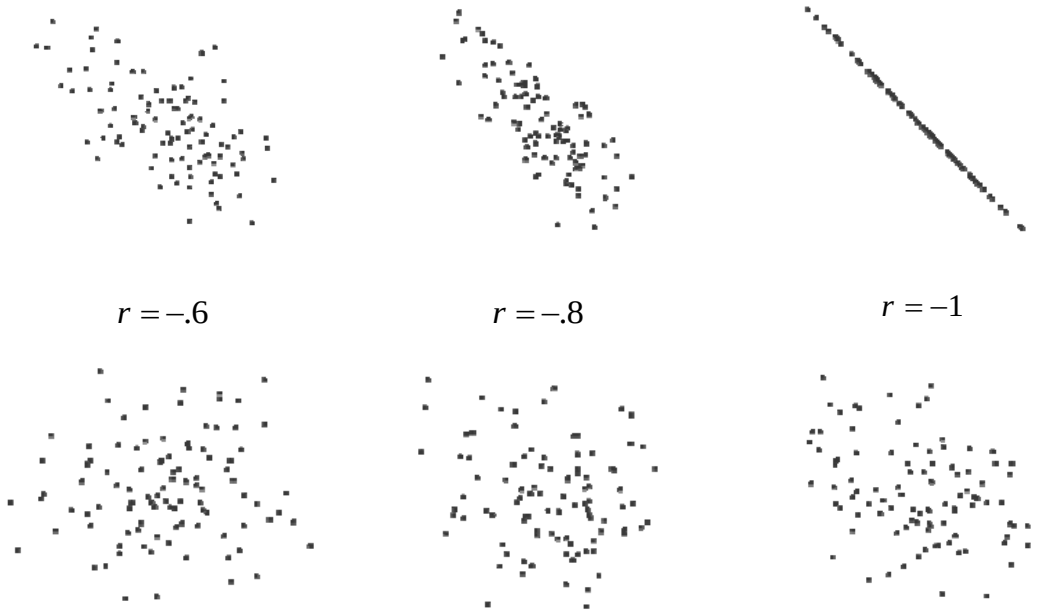
$$-1 \le r \le 1$$

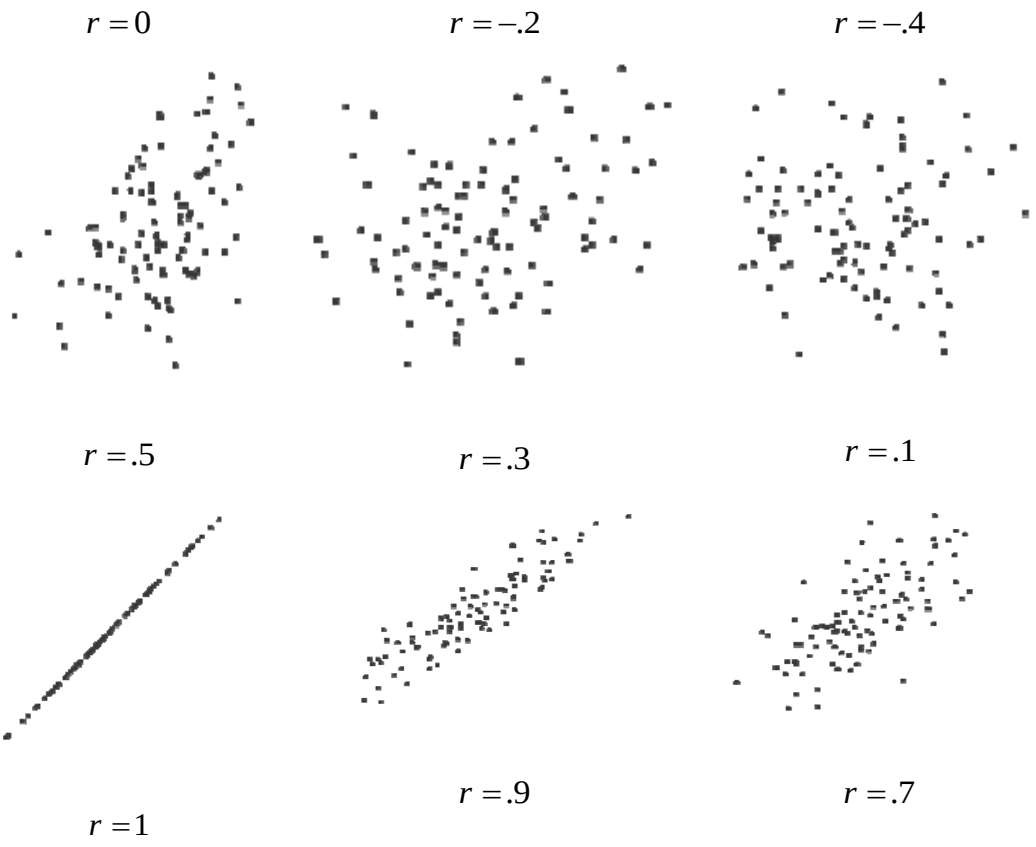
أي أن قيمة معامل الارتباط محصورة بين 1−، 1+ وتدل قيمته على درجة العلاقة بين المتغيرين أو المتغيرات موضع الدراسة من حيث أنها قوية، متوسطة، أو ضعيفة، وأما الإشارة فإنها تصف نوعية العلاقة هل هي

عكسية أم طردية، فالإشارة السالبة تدل على وجود علاقة عكسية أما الموجبة فتدل على وجود علاقة طردية بين المتغيرين موضع الدراسة.

- إذا كانت قيمة معامل الارتباط مساوية للواحد الصحيح فهذا يدل على أن الارتباط بين المتغيرين ارتباطاً طردياً تاماً، أما إذا كانت قيمته مساوية لـ 1– فهذا يدل على أن الارتباط بين المتغيرين ارتباطاً عكسياً تاماً.
- إذا كانت قيمة معامل الارتباط مساوية للصفر (r = 0) فهذا يدل على عدم وجود ارتباط خطي بين المتغيرين موضع الدراسة، بمعنى أنه إذا عرفنا اتجاه تغير أحد المتغيرين استحال علينا تحديد أو معرفة اتجاه المتغير الآخر.
- أما إذا ابتعدت بعض نقاط شكل الانتشار عن الخط المستقيم فإن الارتباط يكون غير تاماً، وتزداد قوة الارتباط كلما اقتربت قيمة r من القيمة 1+ أو القيمة 1–. فمثلاً الطول والوزن لمجموعة من الأشخاص قد يوجد بينها ارتباطاً طردياً ولكن ليس ارتباطاً تاماً. العلاقة بين X, Y تكون:
  - طردية ضعيفة عندما $0 < r < \frac{1}{2}$.
  - طردية متوسطة عندما $\frac{1}{2} \le r < \frac{3}{4}$.
  - طردية قوية عندما $\frac{3}{4} \le r < 1$
  - عكسية ضعيفة عندما $-\frac{1}{2} < r < 0$
  - عكسية متوسطة عندما $-\frac{3}{4} < r \le -\frac{1}{2}$
  - عكسية قوية عندما $-1 < r \le -\frac{3}{4}$

برسم لوحة الانتشار لقيم مختارة من معاملات الارتباط الخطي يمكن الحصول على أحد الأشكال التالية:





حساب قيمة معامل الارتباط:

يمكن حساب قيمة معامل الارتباط بعدة طرق مختلفة تبعاً لنوع البيانات.

الارتباط بين المتغيرات الرقمية: معامل بيرسون للارتباط.

الارتباط بين المتغيرات الترتيبية: معامل سبيرمان للرتب

الارتباط بين المتغيرات الوصفية: مربع كاي Chi-Square.

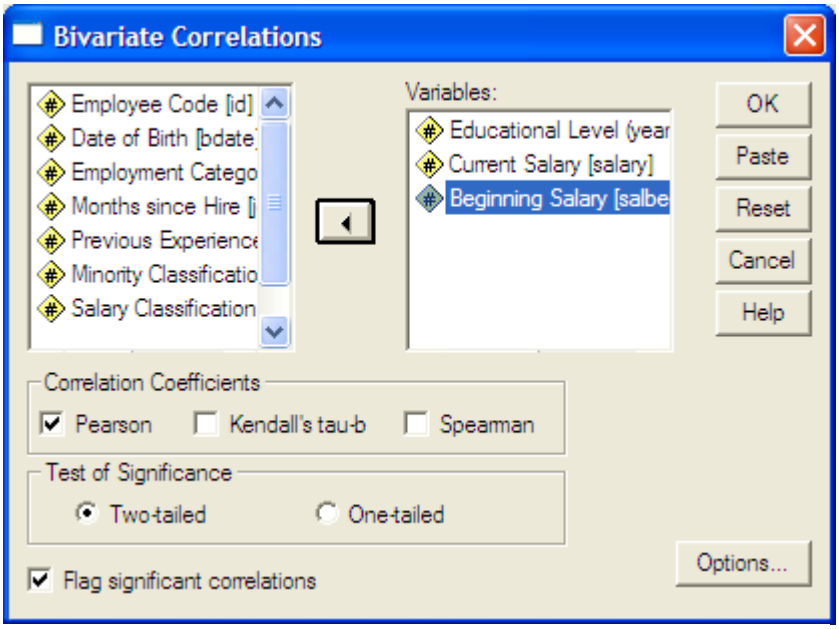
مثال (٥)

افتح الملف Employee Data. المطلوب إيجاد قيمة معامل الارتباط الخطي بين كلاً من المتغيرات salary, salbegin, educ

SPSS STEP BY STEP

Analyze ⇒ Correlate ⇒ Bivariate

أكمل المربع الحواري كما يلي:



من النتائج السابقة يمكن استنتاج ما يلي::

قيمة معامل الارتباط بين educ (Educational Level), salary (Current Salary) تساوي 661. وهذا يدل على أن الارتباط بينهما طردي، Sig.=.000 مما يدل على وجود ارتباط معنوي بين المتغيرين.

قيمة معامل الارتباط بين educ (Educational Level), salbegin (Beginning Salary) تساوي 633. وهذا يدل على أن الارتباط بينهما طردي، Sig.=.000 مما يدل على وجود ارتباط معنوي بين المتغيرين.

قيمة معامل الارتباط بين salbegin (Beginning Salary), salary (Current Salary) تساوي 880. وهذا يدل على أن الارتباط بينهما طردي، Sig.=.000 مما يدل على وجود ارتباط معنوي بين المتغيرين.

مثال (٦)

فيما يلي تقديرات عشرة من طلاب في امتحان مادتي الرياضيات والإحصاء:

الرياضيات	راسب	جيد	مقبول	جيـــد جداً	مقبول	مقبول	جيد	جيد جداً	جيد
الإحصاء	مقبول	جيد جداً	جيد	ممتاز	راسب	جيد	جيد جداً	جيد	ممتاز	راسب

المطلوب: احسب معامل الارتباط بين تقديرات المادتين.

بعد ادخال البيانات واتباع خطوات المثال السابق اختر Spearman فنحصل على النتائج التالية.

Correlations

			الرياضيات	الإحصاء
Spearman's rho	الرياضيات	Correlation Coefficient	1.000	.718*
		Sig. (2-tailed)	.	.019
		N	10	10
	الإحصاء	Correlation Coefficient	.718*	1.000
		Sig. (2-tailed)	.019	.
		N	10	10

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

من النتائج السابقة يمكن استنتاج ما يلي:

معامل سبيرمان للرتب = ٠,٧١٨ فبالتالي يوجد ارتباط طردي بين تقديرات الطلاب في المادتين، وذلك على أساس معامل سبيرمان للرتب. Sig.=.019 مما يدل على وجود ارتباط معنوي بين تقديرات الطلبة في مادتي الرياضيات والإحصاء.

الانحدار الخطي البسيط Simple Linear Regression

الانحدار هو دراسة للتوزيع المشترك لمتغيرين أحدهما متغير يقاس دون خطأ ويسمى متغير مستقل Independent variable ويرمز له بالرمز x والآخر يأخذ قيماً تعتمد على قيمة المتغير المستقل ويسمى التابع Dependent variable ويرمز له بالرمز y.

الهدف من دراسة الانحدار هو إيجاد دالة العلاقة بين المتغيرين المستقل والتابع والتي تساعد في تفسير التغير الذي قد يطرأ على المتغير التابع (y) تبعاً لتغير في قيم المتغير المستقل (x).

مثال (٧)

لدراسة العلاقة بين الدخل والاستهلاك بالدنانير في مدينة غزة، أخذت عينة مكونة من عشرة أسر فأعطت النتائج التالية:

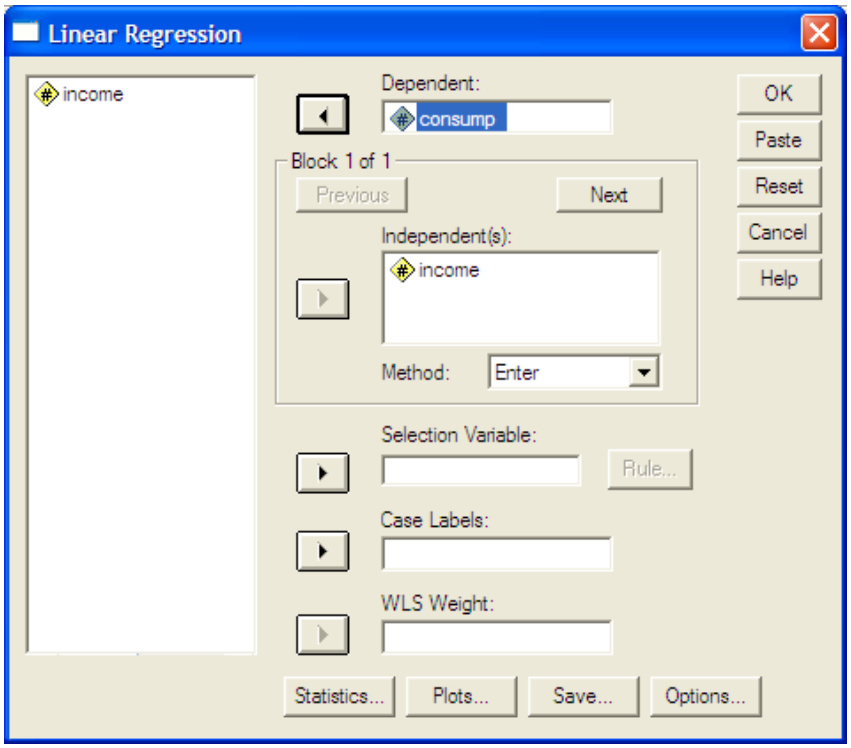
الدخل	٣٠٠	٣٥٠	٥٠٠	٦٠٠	٩٠٠	١٠٠٠	٩٠٠	١٢٠٠	١٠٥٠	٢٥٠
الاستهلاك	٢٨٠	٣٤٠	٥٠٠	٥٥٠	٨٠٠	٧٥٠	٨٥٠	١٠٥٠	١٠٠٠	٢٥٠

المطلوب: إيجاد نموذج انحدار الاستهلاك على الدخل.

SPSS STEP BY STEP

Analyze ⇒ Regression ⇒ Linear

أكمل المربع الحواري كما يلي:



Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.982[a]	.965	.960	58.60899

a. Predictors: (Constant), income

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	751329.9	1	751329.889	218.727	.000[a]
	Residual	27480.111	8	3435.014		
	Total	778810.0	9			

a. Predictors: (Constant), income

b. Dependent Variable: consump

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	48.229	43.913		1.098	.304
	income	.835	.056	.982	14.789	.000

a. Dependent Variable: consump

من النتائج السابقة يمكن استنتاج ما يلي:

١- نموذج انحدار الاستهلاك على الدخل هو:

Consump. = 48.229 + 0.835 * Income

٢- معامل الارتباط بين الدخل والاستهلاك = ٠,٩٨٢ وهو يدل على وجود ارتباط طردي قوي بينهما،($Sig. = 0.000 < \alpha = 0.05$)

٣- معامل التحديد $R^2 = 0.965$ ، ومعامل التحديد المُعدَّل= ٠,٩٦، الخطأ المعياري للتقدير = ٥٨,٦١٩٠.

تفسير قيمة معامل التحديد:

٩٦,٥% من تغير قيمة الاستهلاك (المتغير التابع) يمكن أن يفسر باستخدام العلاقة الخطية بين الدخل والاستهلاك والنسبة المتبقية ٣,٥% ترجع إلى عوامل أخرى تؤثر على قيمة الاستهلاك.

٤- F= 218.727، $Sig. = 0.000$ وهذا يدل على وجود علاقة معنوية بين الدخل والاستهلاك وأن نموذج الانحدار السابق جيد.

٥- $t_{b_0} = 1.098$ ، Sig.=0.304 وهذا يدل على أن نموذج انحدار الاستهلاك على الدخل يمر بنقطة الأصل.

٦- $t_{b_1} = 14.789$ ، Sig.=0.000 وهذا يدل على أن الدخل متغير مؤثر في تقدير قيمة الاستهلاك ويجب أن يكون ضمن نموذج خط الانحدار.