

ورشة العمل على برنامج SPSS تحليل البيانات وإيجاد العلاقات بينها (2)

إعداد

أ.سنة عبدالله أبو نصره

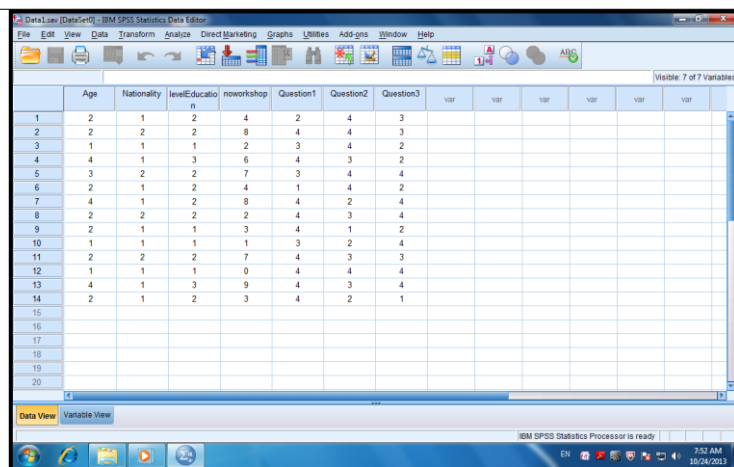
Sabunasrah@ksu.edu.sa

قسم الإحصاء وبحوث العمليات

جامعة الملك سعود

2

سنقوم بهذه الورشة بوصف وتحليل البيانات
البيانات التي قمنا بإعطائها في ورشة العمل
الأولى



	Age	Nationality	levelEducation	noworkshop	Question1	Question2	Question3				
1	2	1	2	4	2	4	3				
2	2	2	2	8	4	4	3				
3	1	1	1	2	3	4	2				
4	4	1	3	6	4	3	2				
5	3	2	2	7	3	4	4				
6	2	1	2	4	1	4	2				
7	4	1	2	8	4	2	4				
8	2	2	2	2	4	3	4				
9	2	1	1	3	4	1	2				
10	1	1	1	1	3	2	4				
11	2	2	2	7	4	3	3				
12	1	1	1	0	4	4	4				
13	4	1	3	9	4	3	4				
14	2	1	2	3	4	2	1				
15											
16											
17											
18											
19											
20											

بعد انتهائنا من عمل الرسومات سنقوم بتحليل البيانات

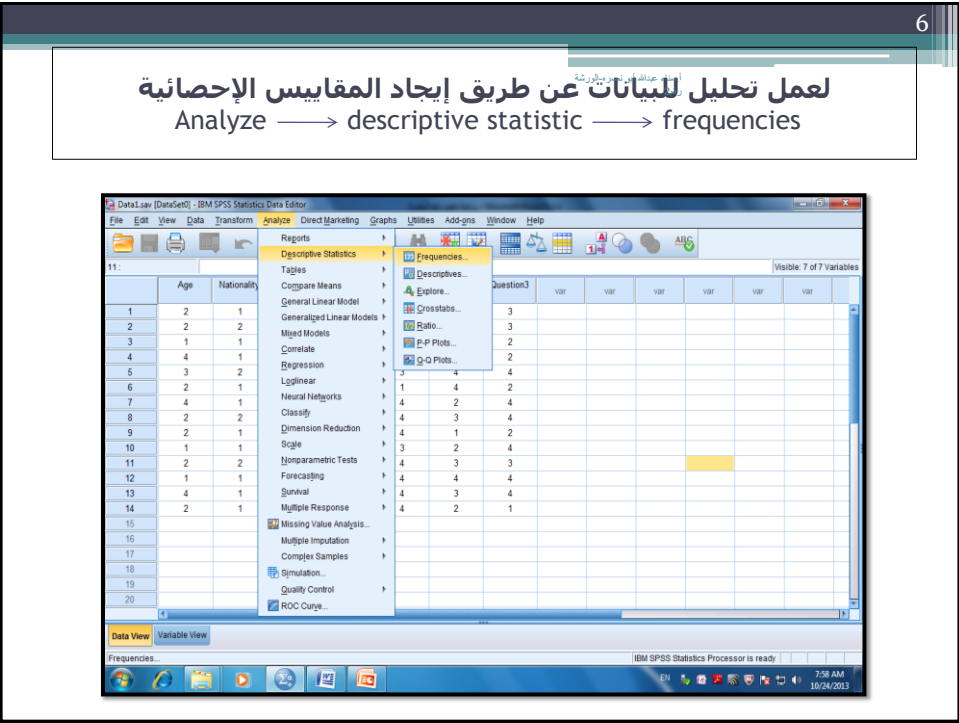
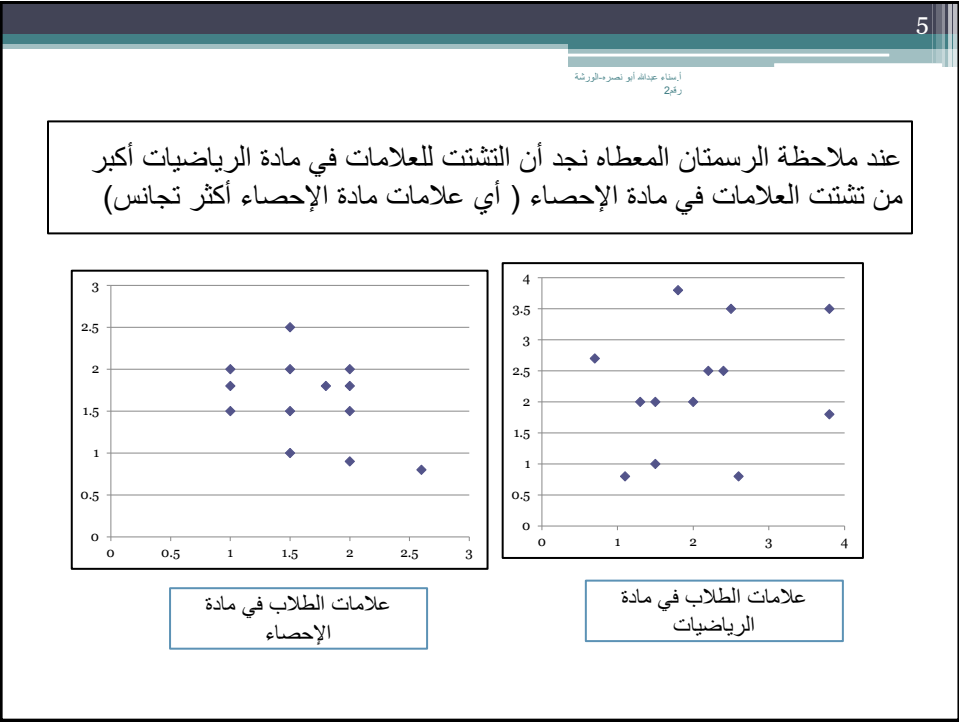
- معنى تحليل البيانات أي إيجاد المقاييس الإحصائية مثل الوسط الحسابي , الوسيط , المنوال , التباين , الانحراف المعياري, المدى والربيعيات .
- يمكن إيجاد المقاييس الإحصائية للبيانات اللفظية و الكمية منها وبالتالي يمكن إيجاد المقاييس الإحصائية للمتغير "عدد ورش العمل".
- لا يمكن إيجاد المقاييس الإحصائية للمتغير "العمر" مع أنه متغير كمي لأنه مقسم إلى فترات إلا إن وجدت البيانات الأصلية التي تم تقسيمها.
- عندما نتبع الخطوات لإيجاد المقاييس سنجد بنفس الخطوة الجدول التكراري للمتغير.
- **المقاييس تنقسم إلى قسمين :**
- **(أ)مقاييس النزعة المركزية : وهي قيم عديدة تلخص البيانات والتي تبحث عن مركز البيانات .**
 - (1) **الوسط الحسابي : Mean**
وهو من أهم المقاييس الإحصائية والأكثر استخداما ويتم حسابه للبيانات الرقمية فقط ويتم استخدامه الفروقات في الوسط الحسابي في اختبارات ANOVA, T- test
 - (2) **الوسيط : Median**
وهو القيمة التي تقع في وسط البيانات بعد ترتيبها تصاعديا أو تنازليا ويتم حسابه للبيانات الرقمية واللفظية (الوصفية) التي يمكن ترتيبها .
 - (3) **المنوال : Mode**
وهو القيمة الأكثر تكرارا ويتم حسابه للبيانات الرقمية والوصفية .

4) الربيعيات (Percentiles):

- وهي ثلاثة قيم تقسم البيانات إلى أربعة أقسام .
- الربيع الأول : يكون 25% من البيانات تسبقه .
- الربيع الثاني : يكون 50% من البيانات تسبقه .
- الربيع الثالث : يكون 75% من البيانات تسبقه .

ب) مقاييس التشتت : وهي المقاييس التي تبحث عن مقدار التباعد أو التقارب بين قيم المشاهدات التابعة لمتغير ما وكلما كان مقياس التشتت كبيرا دل على عدم التجانس بين القيم .

- (1) **المدى : Range**
وهو الفرق بين أكبر قيمة وأصغر قيمة في البيانات الإحصائية .
- (2) **التباين : Variance**
وهو متوسط مربع انحرافات القيم عن وسطها الحسابي .
- (3) **الانحراف المعياري : Standard deviation**
وهو الجذر التربيعي للتباين .



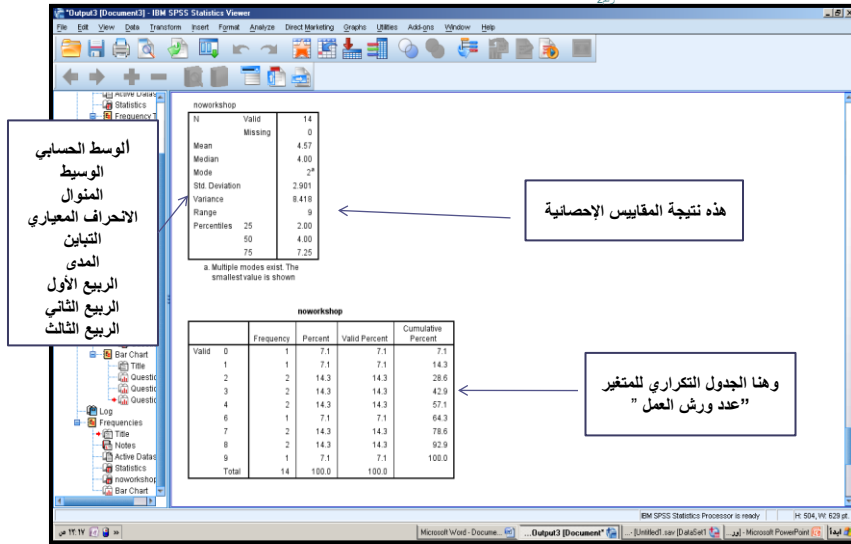
7

إسماء عبدالله أبو نصره-الورشة رقم2

نقل المتغير الذي نرغب بعمل تحليل له :

8

إسماء عبدالله أبو نصره-الورشة رقم2



ملاحظة :

ولو أردنا إيجاد المقاييس الإحصائية لأكثر من متغير في جدول واحد للمقارنة بينهم مباشرة. فإننا نستخدم نفس الطريقة السابقة ولكننا ننقل المتغيرات التي نرغب بحساب المقاييس لها

العلاقة بين المتغيرات

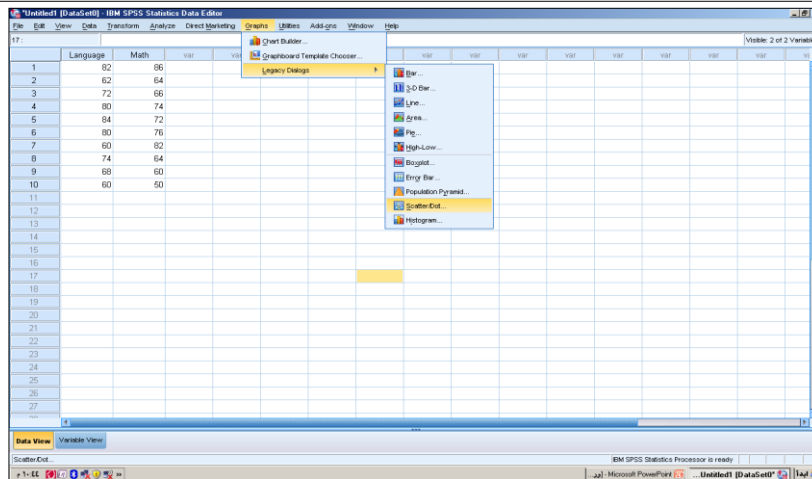
- سنقيس العلاقة بين متغيرين عن طريق حساب ما يسمى معامل الارتباط وهو يقيس العلاقة فيحدد درجة قوتها (قوية , ضعيفة , تامة) واتجاهها (عكسية , طردية) كالآتي :
 - قيمة معامل الارتباط تنحصر بين الصفر والواحد .
 - كانت قيمة معامل الارتباط = صفر ----- لا يوجد علاقة (Uncorrelated)
 - إذا كانت قيمة معامل الارتباط = 1 ----- علاقة تامة وهو أقوى صورة من صور الارتباط.
 - إذا اقترب المعامل من الصفر يكون الارتباط ضعيف , وإذا اقترب من الواحد يكون الارتباط قويا . يسمى الارتباط قويا جدا إذا كان بين 0.8 وواحد ويسمى قويا إذا كان بين [0.5,0.8] ومتوسطة إذا كانت بين [0.2,0.5] وضعيفة إذا كانت أقل من 0.2 .
 - إذا كانت إشارة معامل الارتباط موجبة فهذا يعني أن الارتباط طردي وإذا كانت سالبة فالارتباط عكسي .
 - إذا كانت المتغيرات مستقلة فإن معامل الارتباط بينهما يساوي صفر ولكن العكس غير صحيح .
- هناك نوعان من معامل الارتباط سنتطرق إليهما وهما :
 - معامل ارتباط بيرسون Pearson ويستعمل عندما تكون قيم المتغيرات عددية .
 - معامل ارتباط سبيرمان Spearman ويستعمل عندما تكون قيم المتغيرات لفظية ولكن يمكن ترتيبها أو عددية.

مثال : سنقوم بإدخال علامات عشرة طلاب في مادتي اللغة والرياضيات :

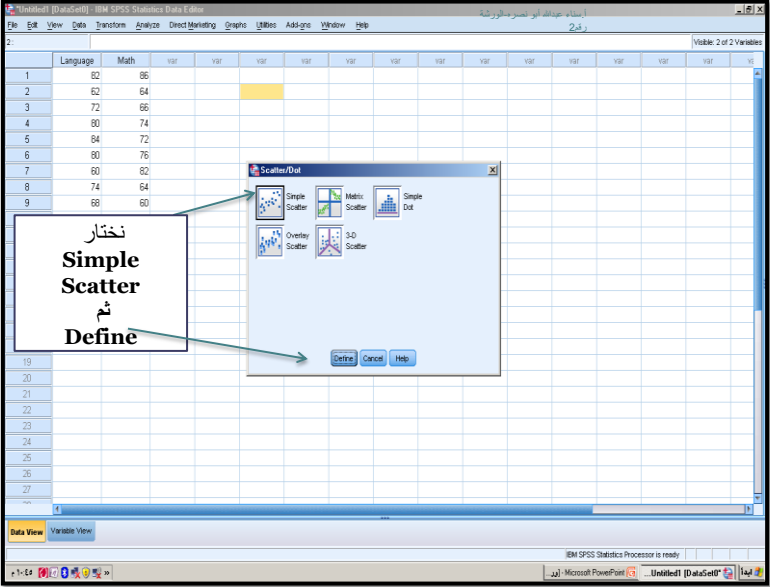
علامات مادة اللغة : 60,68,60,74,80,84,80,72,62,82

علامات مادة الرياضيات : 56,60,64,82,76,72,74,66,64,86

نريد أن نعرف العلاقة بين درجة الطلاب بين المادتين فنرسم رسم مبدئي لمعرفة العلاقة مبدئيا قبل إيجاد معامل الارتباط

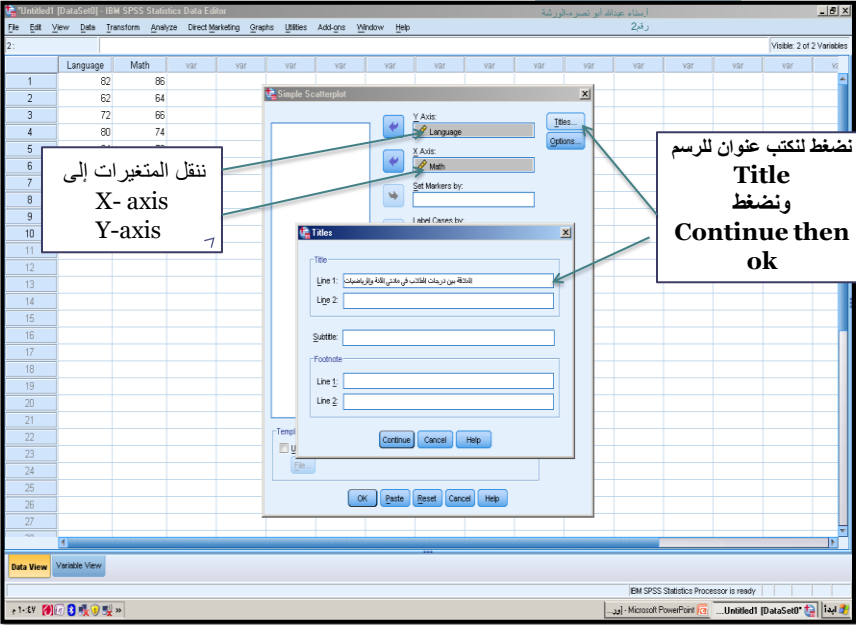


13



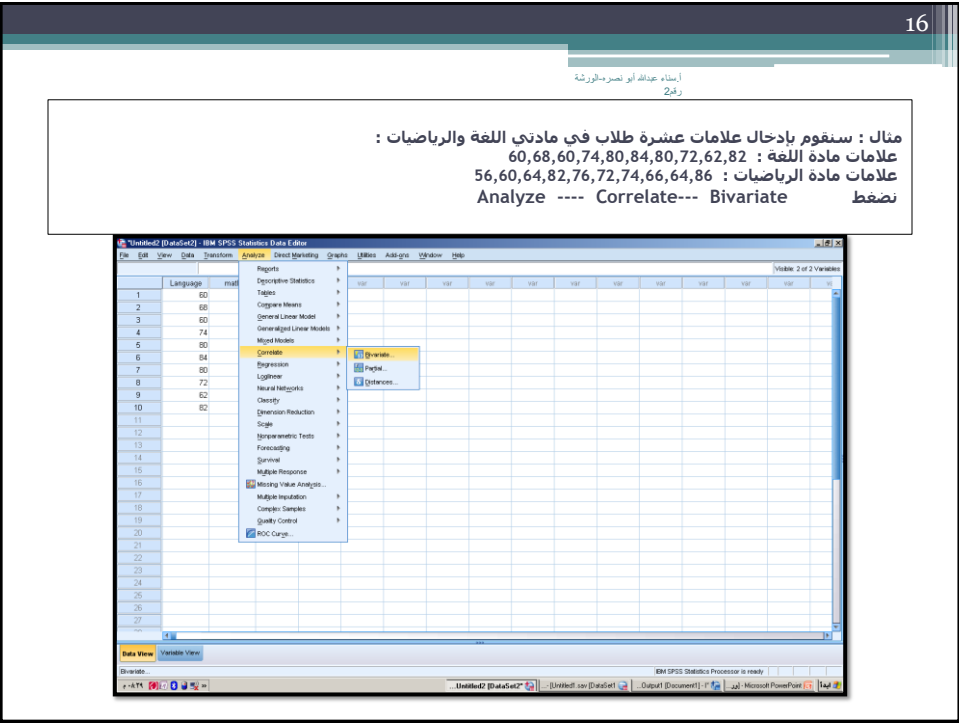
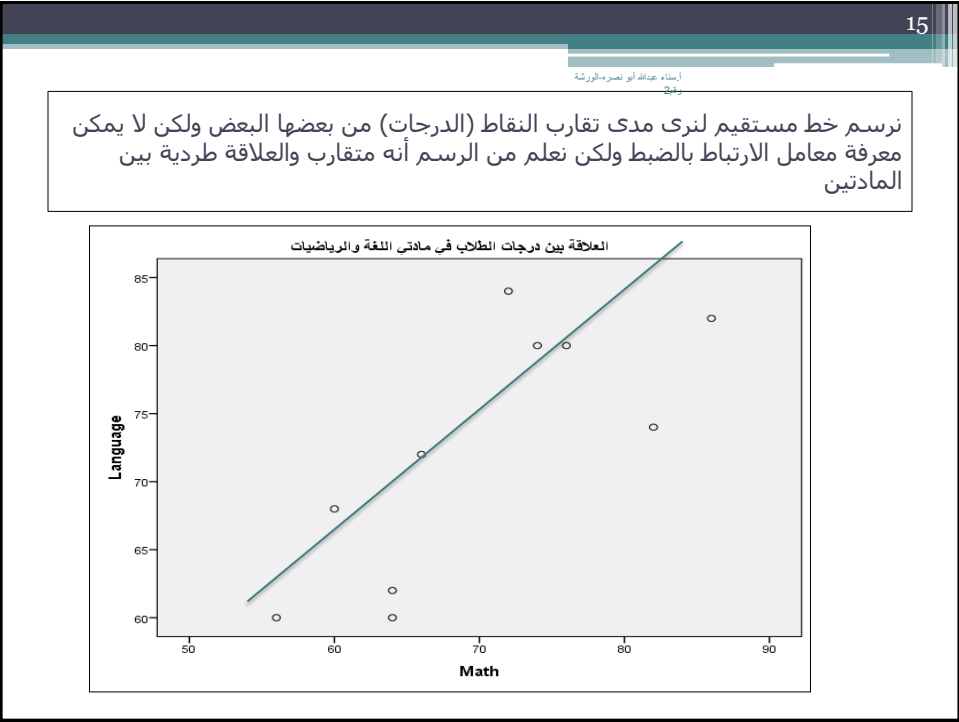
نختار
Simple Scatter
ثم
Define

14



ننقل المتغيرات إلى
X-axis
Y-axis

نضغط لنكتب عنوان للرسم
Title
ونضغط
Continue then ok



17

نقل المتغيرات التي نريد إيجاد العلاقة بينها ثم نختار أي من معاملات الارتباط

18

معامل الارتباط = 0.776 (علاقة قوية)

		Language	math
Language	Pearson Correlation	1	.776
	Sig. (2-tailed)		.000
	N	10	10
math	Pearson Correlation	.776	1
	Sig. (2-tailed)	.000	
	N	10	10

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

اختبار الاستقلال بين متغيرين اختبار مربع - كاي (chi-square test)

- كثيرا ما نحتاج عند إجراء دراسات تطبيقية أو نظرية إلى دراسة متغيرين إلى البحث في كونهما مستقلين أم لا (يوجد علاقة بينهما أم لا).
- ويمكن استخدام اختبار مربع - كاي للاستقلال في اختبار فرضيتين :
- الأساسية : المتغيرين مستقلين (لا يوجد علاقة).
- البديلة : المتغيرين غير مستقلين (يوجد علاقة بينهما).
- وتستخدم سواء كانت المتغيرات كمية أو وصفية .

طريقة أخرى لمعرفة إذا كان هناك علاقة بين متغيرين وهي اختبار مربع - كاي
Chi-Square Test

- مثال:
- سنقوم بعمل تجربة على 22 مشاهدة للمرضى الذين خضعوا للعلاج (A) والذين لم يخضعوا للعلاج (B)
- وهو المتغير (treatment)
- والمتغير الثاني هو (recover) ويمثل الشفاء من المرض.
- ويرمز R : الشفاء و N: عدم الشفاء.

treat	A	B	B	B	A	A	B	A	B	A	A	B	A	B	A	B	B	B	A	B	A	B
reco	R	R	N	N	R	R	N	N	N	R	R	R	R	N	R	N	N	N	R	R	N	N
gender	M	M	M	M	M	M	M	M	M	M	M	M	M	M	F	F	F	F	F	F	F	F

21

أستاذة د. أميرة أبو نصر - الورشة رقم 2

عمل اختبار الاستقلالية (independence) وهو اختبار **اختبار مربع - كاي** (chi-square)
فتختبر فرضية العدم (الأساسية) : لا يوجد علاقة بين أخذ العلاج والشفاء (من المثال السابق)
(treatment & recover)
نضغط Analyze ----- Descriptive statistic ---crosstab

22

أستاذة د. أميرة أبو نصر - الورشة رقم 2

نضغط على **Cell**
ليقوم بعمل جدول يضم العلاقات التي نطلبها بالجدول الذي سيظهر

ننقل المتغيرات المراد اختبار الاستقلال للصف والعمود

نضغط **Display cluster by chart** لتمثيل بيانات الجدول بالأعمدة البيانية المزدوجة

23

تضغط
Observed
ويظهر التكرار المشاهد
Expected
لويظهر التكرار المتوقع
ويستخدم لحساب اختبار
الاستقلال فلا نحتاج له هنا.
ثم نضغط
Continue

Percentages:
لحساب النسب المئوية
Row: للسطر
column: للعمود
Total: للمجموع

تضغط
Continue
ثم نضغط
Statistics

24

تضغط
Chi-square
لتجري اختبار الاستقلال
ونختار نوع الاختبار
Phi and
cramer's
بما أن المتغيرين
Nominal

وإذا أردنا حساب
Correlation
نضغط عليها
ثم
Continue

25

IBM SPSS Statistics Viewer

Case Processing Summary

	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
treat * recover	22	100.0%	0	0.0%	22	100.0%

تلخيص للبيانات المعطاة الموجودة ولا يوجد بيانات مفقودة

تreat * recover Crosstabulation

treat	A	B	recover		Total
			N	R	
treat A	Count		2	8	10
	% within treat		20.0%	80.0%	100.0%
	% within recover		18.2%	72.7%	45.5%
treat B	Count		9	3	12
	% within treat		75.0%	25.0%	100.0%
	% within recover		81.8%	27.3%	54.5%
Total	Count		11	11	22
	% within treat		50.0%	50.0%	100.0%
	% within recover		100.0%	100.0%	100.0%

الجدول السابق هو الجدول المزدوج للمتغيرين Count: هو التكرار المشاهد within treat: هو النسبة للنصف within recover: هو النسبة للعمود within total: هو النسبة للكللي لم يأخذ العلاج B: يأخذ العلاج A: لا يشفى N: يشفى R:

عدد المرضى الذين أخذوا العلاج وتم شفائهم = 8 نسبة المرضى الذين أخذوا العلاج وتم شفائهم بالنسبة للعلاج = 80% نسبة المرضى الذين أخذوا العلاج وتم شفائهم بالنسبة للشفاء = 72.7%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	6.600 ^a	1	.010		
Continuity Correction ^b	4.583	1	.032		
Likelihood Ratio	6.994	1	.008		
Fisher's Exact Test				.030	.015

26

سنقوم باختبار الفرضية الأساسية بأن العلاج والشفاء مستقلين (أي لا يوجد علاقة بينهما) ضد الاختبار البديل أنه يوجد علاقة

بما أن قيمة $P\text{-value} = 0.01 < \alpha$ حيث $\alpha = 0.05$ وبالتالي نرفض الفرضية الأساسية ونقبل بالبديلة أي أنه يوجد علاقة بين تناول العلاج والشفاء

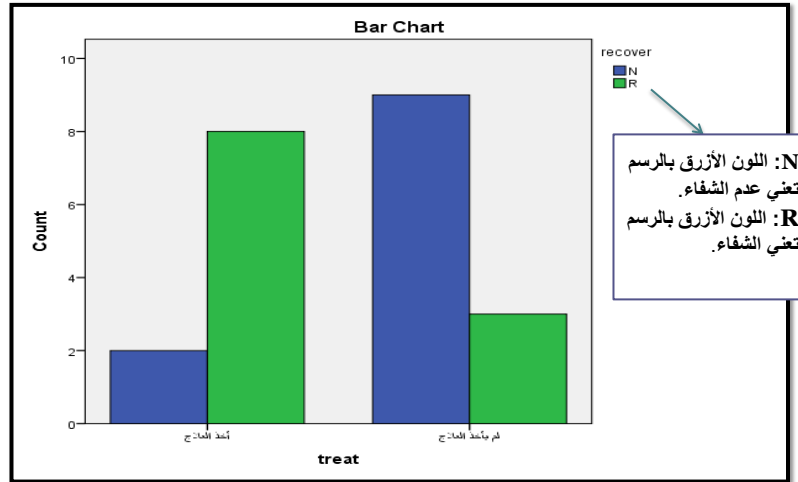
قيمة مربع كاي = 6.6 وهي قيمة الاختبار

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	6.600 ^a	1	.010		
Continuity Correction ^b	4.583	1	.032		
Likelihood Ratio	6.994	1	.008		
Fisher's Exact Test				.030	.015
N of Valid Cases	22				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 5.00.
b. Computed only for a 2x2 table

الرسم البياني يبين عدد الأشخاص الذين لم يتم شفائهم (اللون الأزرق) عند أخذ العلاج (A) ومن لم يأخذوا العلاج (B) واللون الأخضر هو عدد الأشخاص الذين يتم شفائهم .



السلام عليكم ورحمة الله

