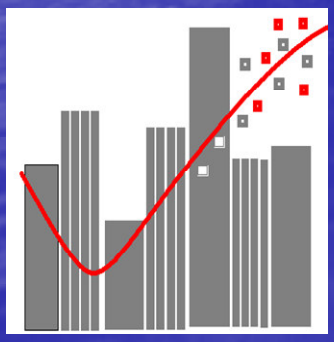
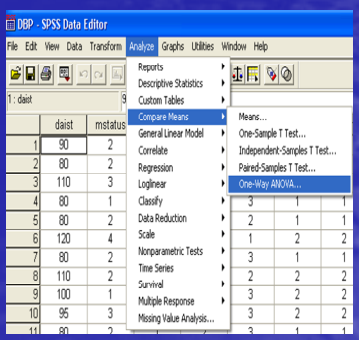
التحليل الإحصائي باستخدام SPSS

د. عز حسن عبد الفتاح
ezzhassan@hotmail.com

2005





1. إدخال البيانات

1.1 رموز الصفحة (شريط الأدوات)

فتح ملف تم حفظه مسبقاً

حفظ ملف جديد

طباعة ملف

عرض آخر عمليات تم حسابها

التراجع عن آخر عملية تمت

اعادت آخر عملية تم التراجع عنها

الذهاب الى حالة (case) معينة

المتغيرات

البحث عن رقم محدد

ادراج حالات

ادراج متغير

فصل الملف

وزن الحالات

اختيار حالات

عرض تعريفات المتغيرات

تحديد فئات محددة للإستعمال

1.2 أنواع المقاييس Measurement scales.

هناك 4 أنواع مختلفة لقياس المتغيرات:

الاسمية Nominal (مثل النوع والحالة الاجتماعية)
الترتيبية Ordinal (مثل المستوي التعليمي و الكادر الوظيفي)
الفترة Interval (مثل درجات الحرارة)
النسبة Ratio (مثل الأوزان والأطوال أو أي مقياس له صفر قياس)

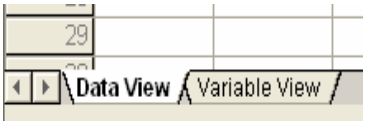
المقاييس الإسمية هي أدنى أنواع المقاييس يليها المقاييس الترتيبية فمقاييس الفترة. أما اعلي المقاييس فهي مقاييس النسبة والتي تصلح معها جميع التحليلات الإحصائية.

(ملاحظة: يعتبر SPSS كل من مقياس النسبة Ratio ومقياس الفترة Interval على أنهما مقاييس كمية Scale)

1.3 إدخال البيانات Data Entry

حقول تعريف المتغير

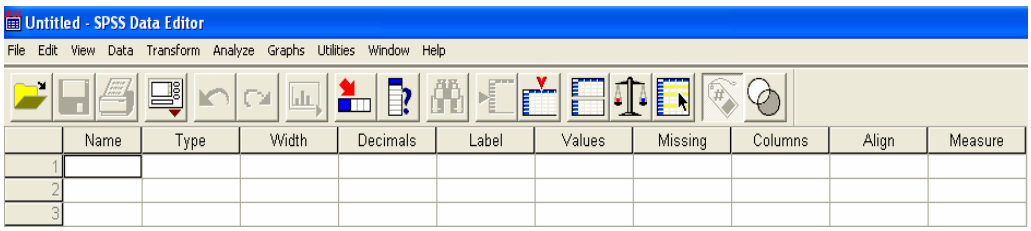
تحتوى صفحة إدخال البيانات على اختيارين أساسيين:



الأول: Data View وهى خاصة بإدخال البيانات الرقمية

الثاني: Variable View وهى خاصة بتعريف المتغيرات.

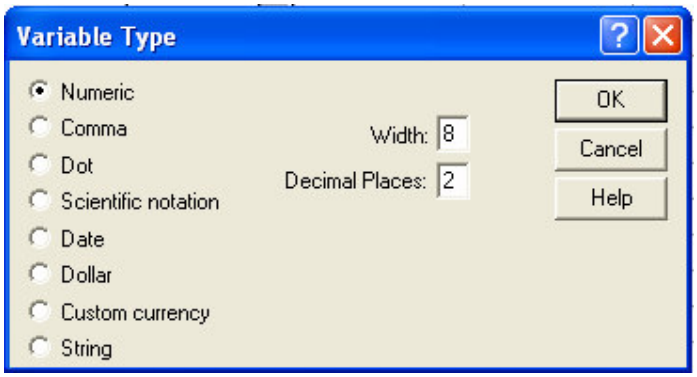
بالضغط على Variable View نتحول إلى صفحة تعريف المتغيرات، والتي تحتوى على عشرة حقول، تظهر كما في الشكل التالي:



لكل حقل استخدام خاص:

1- حقل Name لتعريف اسم المتغير(يجب أن يبدأ بحرف ولا يحتوي على اى مسافات)

2- حقل Type لتعريف نوع المتغير وهى 8 أنواع كما يظهر من الشكل التالي:



التعريف المحدد من قبل البرنامج (العرف) (default)هو المتغير العددي (Numeric) وهو معرف بعدد يحتوي 8 أرقام منهما رقمين بعد العلامة العشرية. يمكن تغيير هذا العرف عند الضرورة.

إذا كان المتغير عبارة عن تاريخ يمكن اختيار Date.

إذا كان المتغير عبارة عن عملة محلية يمكن اختيار Custom currency.

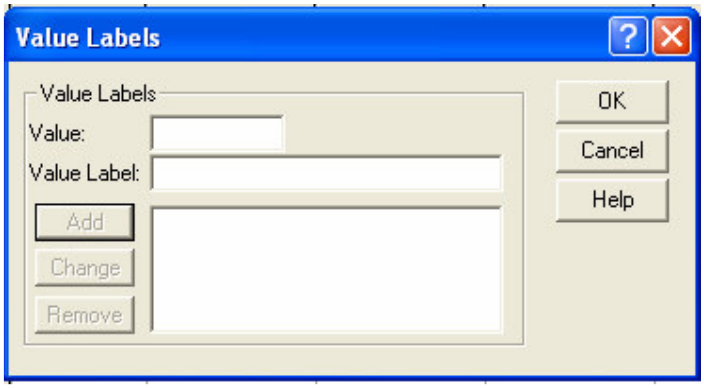
إذا كان المتغير عبارة عن أسماء (مثل أسماء الأشخاص) يمكن اختيار String.

3- حقل Width يضع البرنامج العدد 8 كعرف للدلالة على أن العدد مكون من 8 أرقام (يمكن تزويدها أو تقليلها)

4- حقل Decimals يضع البرنامج العدد 2 كعرف للدلالة على أن العدد مكون من رقمين عشريين (يمكن تزويدها أو إلغائها في حالة الأرقام الصحيحة)

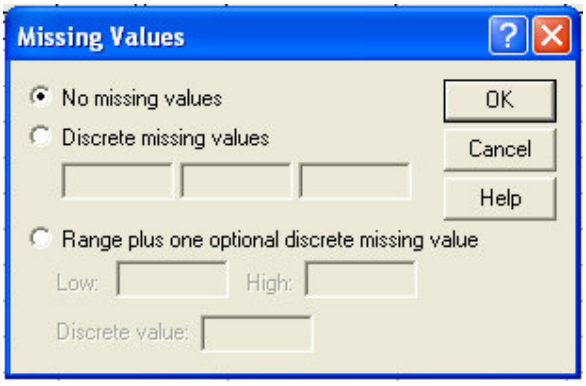
5- حقل Label يستخدم لوصف المتغير بكتابة مايعبر عنه المتغير.

6- حقل Values يستخدم لتعريف عناصر المتغير الترتيبي (Ordinal) أو الإسمي (Nominal). فبالضغط على الجزء الأيمن من الخلية ... يظهر الشكل التالي:



فى الخلية المناظرة ل Value: يتم وضع القيمة
فى الخلية المناظرة ل Value Label: يتم وضع ما تعبر عنه القيمة

7- حقل Missing يستخدم لتعريف القيم المفقودة. فبالضغط على الجزء الأيمن من الخلية ... يظهر الشكل التالي:



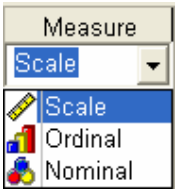
يمكن قبول وجود قيم مفقودة فى التحليل او الإبقاء على اختيار البرنامج والمعبر عن عدم وجود قيم مفقودة.

8- حقل Columns يستخدم لتحديد عرض العمود، والعرف هو 8 خانات يمكن زيادتها أو تقليلها.

9- حقل Align يستخدم لتحديد موقع الرقم داخل الخلية، والعرف هو اتجاه اليسار (Left). فبالضغط على الجزء الأيمن من الخلية ... يظهر الشكل التالي:

يمكن اختيار اتجاه اليمين (Right) او اتجاه الوسط (Center).

10- حقل Measure يستخدم لتحديد نوع المتغير، والعرف هو المتغير الكمي (Scale). فبالضغط على الجزء الايمن من الخلية ... يظهر الشكل التالي:



يمكن اختيار نوع المتغير الترتيبي (Ordinal) او المتغير الإسمي (Nominal).

1.4 مثال لإدخال البيانات:

مثال (1.1) البيانات الآتية تعبر عن أعمار عشرة من الأشخاص ومعرفة حسب النوع والمؤهل التعليمي:

الشخص	1	2	3	4	5	6	7	8	9	10
العمر	26	21	18	35	41	22	22	31	19	26
النوع	أنثي	أنثي	ذكر	ذكر	أنثي	ذكر	ذكر	أنثي	ذكر	ذكر
المؤهل التعليمي	عالي	متوسط	متوسط	ابتدائي	عالي	عالي	متوسط	ابتدائي	عالي	عالي

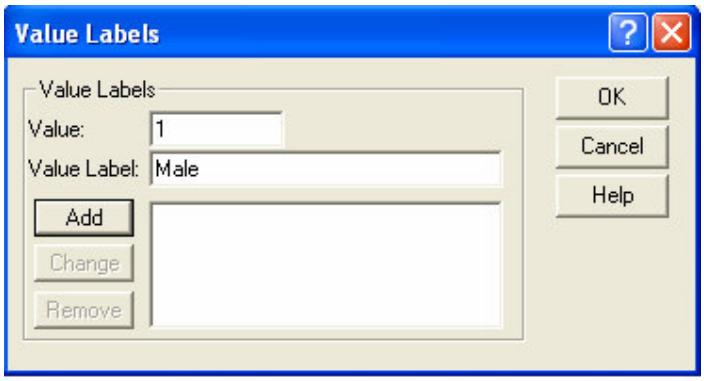
من الواضح أن :
العمر له مقياس كمي (Scale)
النوع له مقياس اسمي (Nominal)
المؤهل التعليمي له مقياس ترتيبي (Ordinal)

المقياس الكمي لايحتاج إلي تعريف القيم (values) أما كلا من المقياس الاسمي والترتيبي فلابد من تعريف القيم لكل منهما.

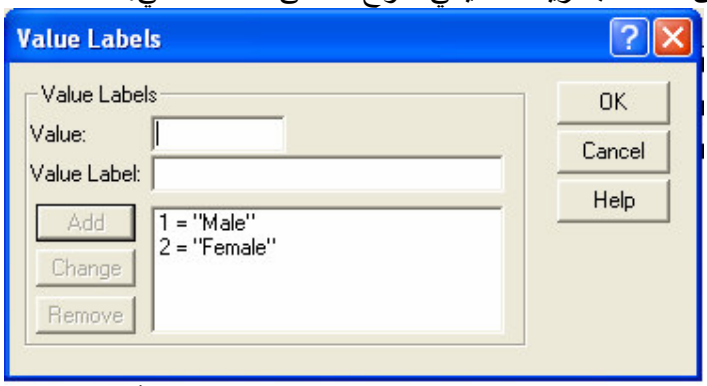
بالنسبة للمتغير الإسمي (النوع) يمكن اختيار "1" للدلالة علي الذكر(Male) و "2" للدلالة على الأنثى (Female) أو العكس، وذلك لأنه متغير اسمي لايهم الترتيب فيه.
يمكن تعريف القيم باللغة العربية إذا كان إصدار البرنامج يدعم اللغة العربية. أو باللغة الانجليزية بغض النظر عن الإصدار.

كما يمكن اختيار حروف مثل "M" للدلالة على الذكر و "F" للدلالة علي الأنثى. وفى هذه الحالة يعرف نوع المتغير علي انه "String" وان كنا لانتصح بذلك حيث انه لا يمكن التعامل مع ال"String" فى كثير من الاختبارات.

لتعريف تقسيمات النوع، نختار أولا Variable View فى خانة Value: نكتب "1" وفى خانة Value Label: نكتب " Male "، كما يتضح من الشكل التالي:

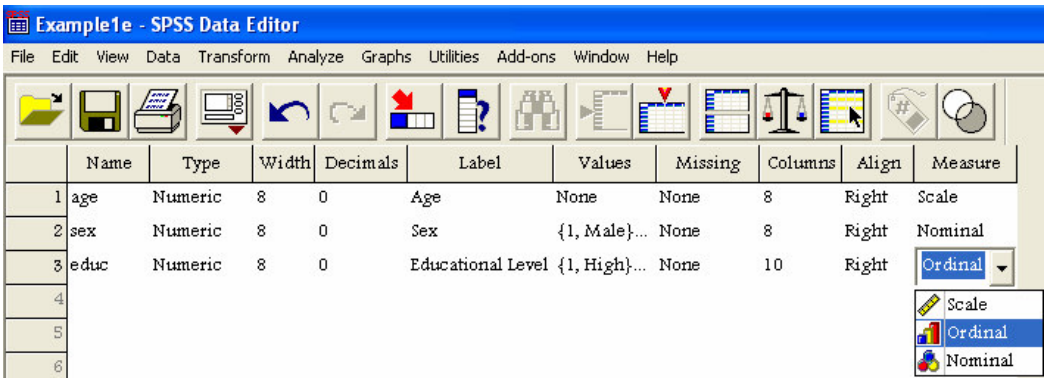


بالضغط على Add وعمل نفس الشيء بالنسبة للأنثى (Female) ثم الضغط على OK نكون قد قمنا بتعريف تقسيمي النوع كما فى الشكل التالي:

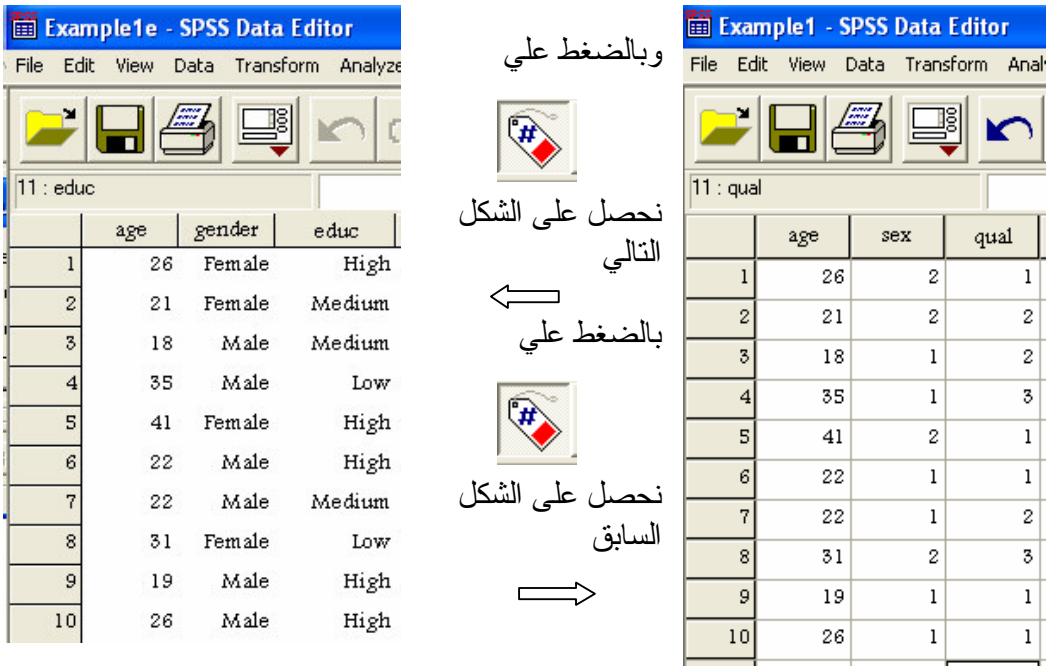


بالنسبة للمتغير الترتيبي (المؤهل التعليمي) يمكن الترتيب تصاعديا أو تنازلياً. فباختيار الترتيب التنازلي نعطي الرقم "1" للدلالة علي المؤهل التعليمي العالي(University) و الرقم "2" للدلالة على المؤهل التعليمي المتوسط(High school) والرقم "3" للدلالة علي المؤهل الابتدائي(Elementary).

فبتعريف المتغيرات الثلاث كما سبق نحصل على شكل مثل الشكل التالي:



وبالتحول إلى Data View وإدخال البيانات نحصل على الشكل التالي (على اليمين):



أهم الاختبارات الإحصائية

الهدف	نوع البيانات		
	كمية (من توزيع طبيعي)	ترتيبية (او كمية ليست من توزيع طبيعي)	اسمية (ثنائية)
وصف مجموعة واحدة	الوسط الحسابي والانحراف المعياري	الوسيط والانحراف الربيعي	المنوال والنسب
مقارنة مجموعة واحدة مع قيمة افتراضية	اختبار "ت" لعينة واحدة	اختبار "ويلكوكسون"	اختبار "مربع كاي" أو اختبار "ذي الحدين"
مقارنة مجموعتين مستقلتين	اختبار "ت" لعينتين مستقلتين	اختبار "مان ويتني"	اختبار "فيشر"
مقارنة مجموعتين مرتبطتين	اختبار "ت" لعينتين مرتبطتين	اختبار "ويلكوكسون"	اختبار "ماكنيمر"
مقارنة ثلاثة مجموعات او اكثر مستقلة	تحليل التباين فى اتجاه واحد	اختبار "كروسكال واليس"	اختبار "مربع كاي"
مقارنة ثلاثة مجموعات او اكثر مرتبطة		اختبار "فريدمان"	اختبار "كوكران كيو"
العلاقة بين مجموعتين	معامل "بيرسون" للارتباط	معامل "سبيرمان" للارتباط	اختبار "مربع كاي" أو "معامل التوافق"
التقدير بالاعتماد علي متغير مستقل واحد	الانحدار الخطي البسيط		الانحدار اللوجستي البسيط
التقدير بالاعتماد علي عدة متغيرات مستقلة	الانحدار الخطي المتعدد		الانحدار اللوجستي المتعدد

Goal	Type of Data		
	Measurement (from Normal Population)	Rank, Score, or Measurement (from Non-Normal Population)	Binomial (Two Possible Outcomes)
Describe one Group	Mean, SD	Median, Interquartile range	Proportion
Compare one group to a hypothetical value	One-sample t test	Wilcoxon test	Chi-square or Binomial test
Compare two unpaired groups	Unpaired t test	Mann-Whitney test	Fisher's test (Chi-square for large sample
Compare two paired groups	Paired t test	Wilcoxon test	McNemar's test
Compare three or more unmatched groups	One-way ANOVA	Kruskal-Wallis test	Chi-square test
Compare three or more matched groups	Repeated-measured ANOVA	Friedman test	Cochran Q
Quantify association between two groups	Pearson Correlation	Spearman Correlation	Contingency coefficient
Predict value from another measured variable	Simple linear regression or Nonlinear regression	Nonparametric regression	Simple logistic regression
Predict value from several measured or binomial variables	Multiple linear regression or Multiple nonlinear regression		Multiple logistic regression

2. الإحصاء الوصفي

2.1 بعض المفاهيم الإحصائية Some Statistical Concepts

إن كلمة "إحصاء" (Statistics) غالباً ما تحمل معان مختلفة ، وذلك لأن كلاً من الإحصائي وغير المتخصص يستخدم التعبير للدلالة على شيء مختلف . فعلى سبيل المثال فإنه لغير المتخصص ، فهذه الكلمة قد تعني التنبؤ الجوى أو التعداد السكاني أو نتائج المباريات وغيرها .

وعلى الجانب الآخر فإن الإحصائيين يستخدمون التعبير " إحصاء " للإشارة إلى الأساليب الإحصائية (Statistical Techniques) التي تطورت لمساعدة الباحثين في كل مجال في تحليل وفهم البيانات الناتجة من دراسة محل الاهتمام .

ومن الشائع تقسيم الإحصاء إلى درجتين بصفة عامة وهما :
الإحصاء الوصفي والإحصاء الاستدلالي (Descriptive and Inferential Statistics) .
ويقصد بالإحصاء الوصفي مجموعة الطرق الإحصائية التي تستخدم في جمع البيانات وعرضها بالجداول أو الرسوم الإحصائية وكذلك وصف الجوانب المختلفة للبيانات المختلفة للبيانات مثل حساب المتوسطات ومقاييس التشتت .
أما الإحصاء الاستدلالي (أو الاستدلال الإحصائي) فإنه يشمل مجموعة النظريات والطرق الإحصائية التي تستخدم في تقدير معالم المجتمعات Population Parameters واختبارات الفروض بشأنها ، وذلك باستخدام عينات Samples مسحوبة من تلك المجتمعات .
والعينة ليست هي موضوع الاهتمام ، ولكن المجتمع . ومع ذلك ، فنظراً لأنه يصعب قياس المجتمع ككل ، فإن بيانات العينة تستخدم لرسم أو إيضاح التصورات أو الاستنتاجات Conclusions عن المجتمع .

والفارق الجوهري بين الإحصاء الوصفي والاستدلالي هو الغرض Purpose فوصف البيانات هو غرض الإحصاء الوصفي بينما التعبير أو الامتداد من العينة إلى المجتمع هو غرض الإحصاء الاستدلالي وكلا النوعين له طرق مختلفة تستخدم في " فهم " البيانات واتخاذ القرارات Making Decisions .

ينقسم الإحصاء الاستدلالي أو الاستدلال الإحصائي إلى موضوعين رئيسيين :-
1- تقدير معالم المجتمعات Estimation .
2- اختبارات فروض بشأن معالم المجتمعات Testing Hypotheses .
ويتم ذلك عن طريق سحب عينة أو عينات من المجتمع أو المجتمعات المراد تقدير معالمها أو اختبارات فروض بشأنها .
يستخدم التقدير لمعالم المجتمع إذا كان الهدف هو تحديد قيمة معلمة مجهولة (Unknown Parameter) .
أما اختبارات الفروض فتستخدم بهدف الوصول إلى قرار بشأن رفض أو عدم رفض فرض إحصائي عن معلمة مجهولة (مزعومة) .

التقدير Estimation

إذا استخدمت قيمة إحصاءه (Statistic) كتقدير لمعلمة مجتمع (Population parameter) فإن تلك القيمة تسمى تقدير بنقطة (Point Estimation) .
وذلك مثل استخدام الوسط الحسابي للعينة لتقدير متوسط المجتمع μ.
اما التقدير بفترة Interval Estimation فهو تمكيننا من صياغة نتيجة ما "بهوامش للخطأ"

اختبارات الفروض Hypotheses Tests

نقوم بإجراء اختبار فرض فقط عند البحث عن قرار بشأن معلمة للمجتمع (Population parameter) بناءً على قيمة الإحصاءة في العينة (Sample statistic) .

المعنوية الإحصائية Statistical significance (p-level).
المعنوية الإحصائية لنتيجة ما هى قيمة مقدره لدرجة " تمثيل المجتمع".
يعبر احتمال المعنوية عن احتمال الخطأ فى تمثيل العينة للمجتمع عند اتخاذ القرار
فى كثير من العلوم تؤخذ القيمة 0.05 كحد للمعنوية
فإذا كانت قيمة احتمال المعنوية اقل من 0.05 (او اقل من القيمة المحددة من قبل الباحث) فإن الاختبار الإحصائي يعتبر معنوياً او يقال انه توجد دلالة إحصائية.

2.2 الإحصاء الوصفي Descriptive Statistics

الوسط الحسابي: The Mean

يعد الوسط الحسابي أحد مقاييس النزعة المركزية وهو يستخدم لحساب القيمة المركزية لمجموعة من البيانات الكمية.

الوسيط : The Median

يعد الوسيط أحد مقاييس النزعة المركزية وهو يستخدم لحساب القيمة الوسطى لمجموعة من البيانات الترتيبية او الكمية.

المنوال : The Mode

يعد المنوال أحد مقاييس النزعة المركزية وهو يستخدم لحساب القيمة الأكثر شيوعاً لمجموعة من البيانات الإسمية أو الترتيبية او الكمية.

التباين : The Variance

يعتبر التباين أحد أهم المقاييس الإحصائية لقياس تشتت البيانات ويعتمد فى حسابه على مجموع مربعات انحرافات البيانات عن وسطها الحسابي

الانحراف المعياري : The Standard deviation

يستخدم الانحراف المعياري لقياس تشتت البيانات ويتميز عن التباين فى أن له وحدة تمييز مثل الوحدة الأصلية لتمييز البيانات ويتم حسابه عن طريق الجذر التربيعي للتباين.

المدى : The Range
يستخدم المدى لقياس تشتت البيانات وهو يقيس الفرق بين اكبر قيمة للبيانات واقلها.

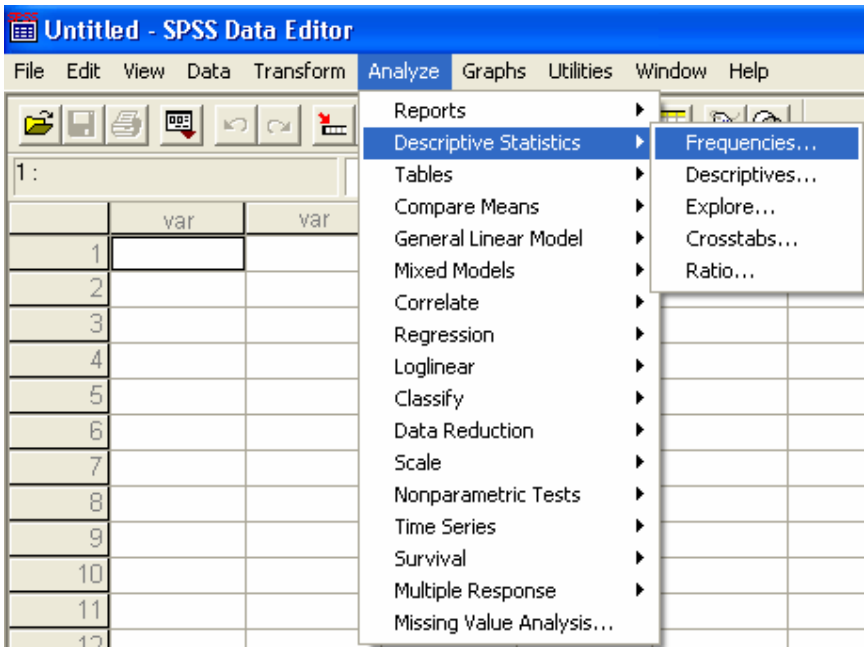
الخطأ المعياري للمتوسط : Standard error of the mean
يحسب الخطأ المعياري للمتوسط بقسمة الانحراف المعياري على جذر حجم العينة.

الالتواء : Skewness
الالتواء هو مقياس لتماثل التوزيع وهو يساوى الصفر للتوزيعات المتماثلة.

التفرطح : Kurtosis
التفرطح هو مقياس لتحدب التوزيع وهو يساوى الصفر للتوزيعات الطبيعية.

2.3 وصف المتغيرات الإسمية أو الترتيبية:

Analyze
Descriptive Statistics
Frequencies…

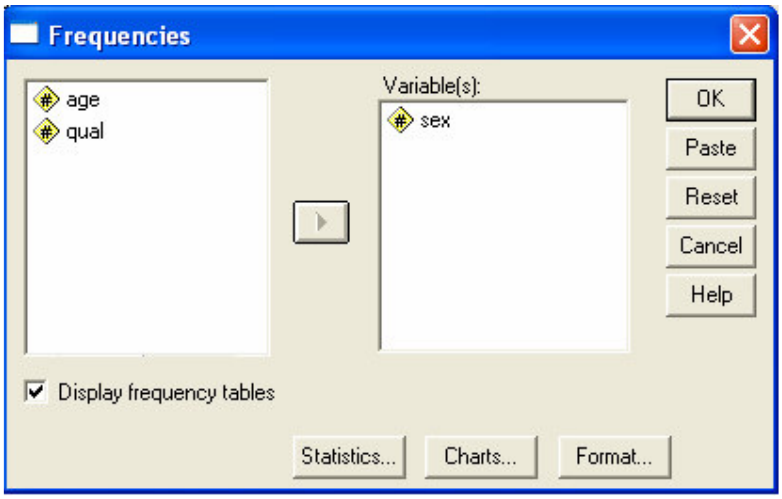


كيفية التنفيذ :
- أد راج المتغير المراد وصفه فى Variable(s)
- الضغط علي OK.

مثال (2.1) باستخدام بيانات مثال (1.1) يمكن وصف متغير النوع Sex (وهو متغير اسمي) أو وصف المؤهل التعليمي Qualification (وهو متغير ترتيبي). فبإتباع الخطوات السابقة لوصف النوع نتبع الخطوات التالية:

Analyze Descriptive Statistics Frequencies…

باختيار النوع (sex) فى Variable(s) والضغط على ▶ نحصل على الشكل التالي:



بالضغط على OK نحصل على الجدول التالي:

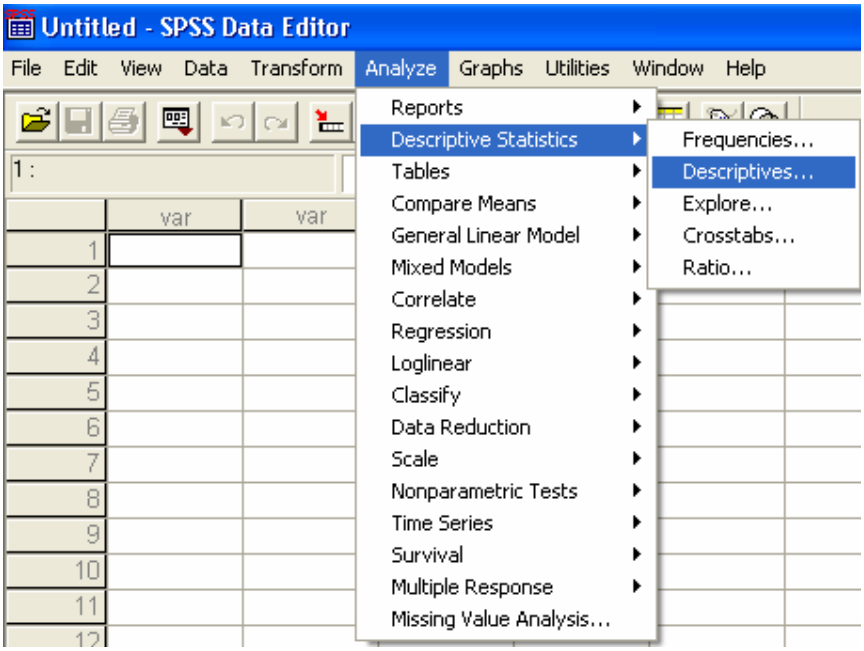
Sex

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	6	60.0	60.0	60.0
	Female	4	40.0	40.0	100.0
	Total	10	100.0	100.0	

من الجدول السابق يمكن ملاحظة التكرار (Frequency) والنسبة (Percent). ويلاحظ أن العمود (Valid Percent) له نفس قيم عمود التكرارات، وهما يختلفان فقط اذا كانت هناك قيم مفقودة (Missing values) حيث تحسب النسب حين ذاك تبعا للقيم الموجودة وليس القيم الإجمالية (اى بدون القيم المفقودة). اما عمود النسب التراكمية(Cumulative Percent) فهو يحسب النسب التراكمية (التجميعية).

2.4 وصف المتغيرات الكمية:

Analyze
Descriptive Statistics
Descriptives…



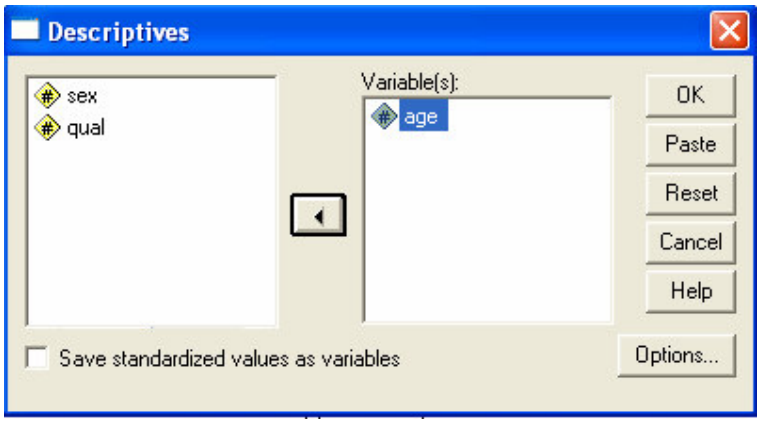
كيفية التنفيذ :

- أد راج المتغير المراد وصفه فى Variable(s)
- الضغط علي OK.

مثال (2.2) باستخدام بيانات مثال (1.1) يمكن وصف متغير العمر Age (وهو متغير كمي).
فبإتباع الخطوات السابقة لوصف العمر نتبع الخطوات التالية:

Analyze Descriptive Statistics Descriptives…

باختيار العمر (Age) فى Variable(s) والضغط على ▶ نحصل على الشكل التالي:

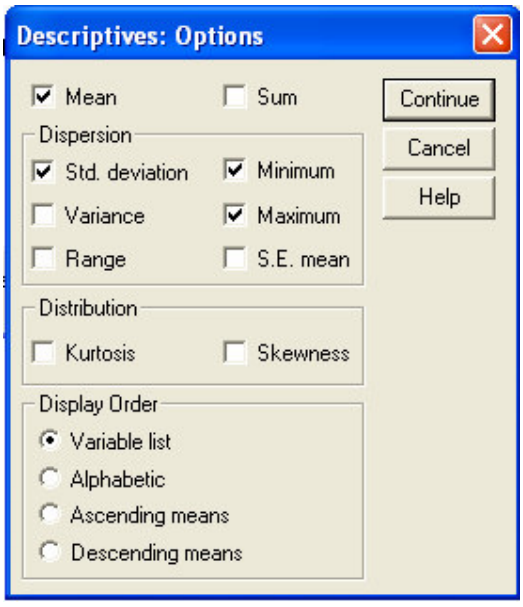


بالضغط على OK نحصل على الجدول التالي:

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Age	10	18	41	26.10	7.460
Valid N (listwise)	10				

من الجدول السابق يمكن ملاحظة عدد التكرارات (المشاهدات) والقيمة الصغرى والقيمة العظمي والمتوسط الحسابي والانحراف المعياري.
جدير بالذكر ان النواتج (الإحصاءات) السابقة هى العرف المقترح من البرنامج. أما إذا أردنا وصف المتغير بمزيد من الإحصاءات، فيمكن الضغط علي Options... من الشكل السابق فنحصل علي الشكل التالي:



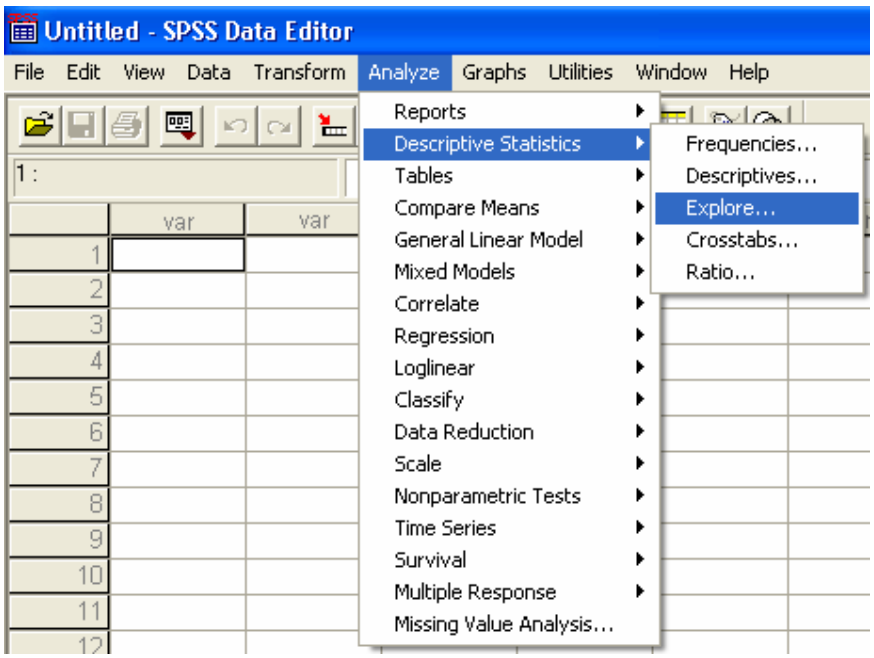
من خلال الشكل السابق يمكن التأشير على الإحصاءات المطلوب الحصول عليها ثم الضغط علي Continue ثم على OK.

2.5 أمر استكشاف (Explore):

يستخدم هذا الأمر:

1- **لاكتشاف وجود قيم شاذة (outliers) فى البيانات.**
2- **لاكتشاف وجود أخطاء فى البيانات قبل بدء التحليل.**
3- **لوصف المتغير الكمي شاملاً فترة الثقة وبعض الأشكال البيانية.**
4- **لإجراء اختبار الإعتدالية (Normality)**

Analyze
Descriptive Statistics
Explore…



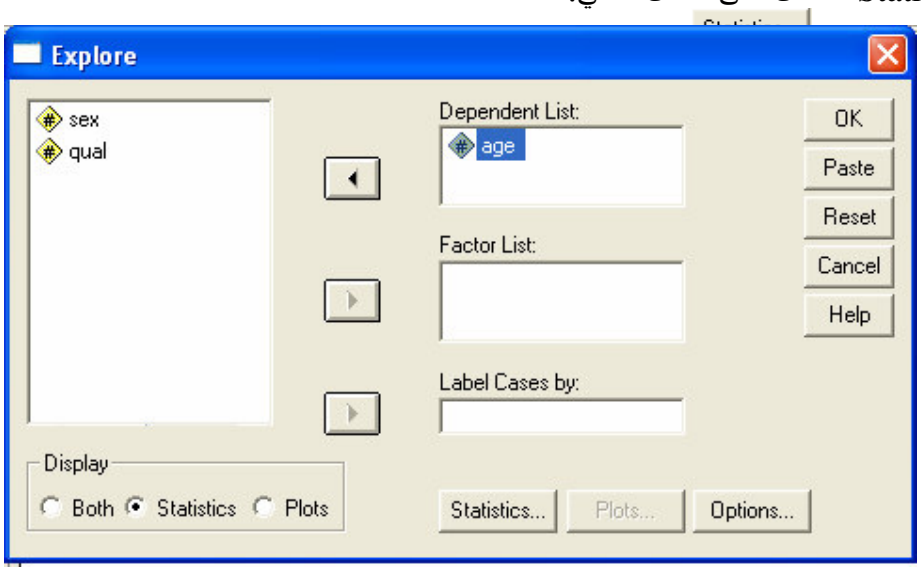
2.5.1 استخدام امر Explore لاكتشاف وجود أخطاء فى البيانات

مثال (2.3) نبدأ باستخدام بيانات مثال (1.1) وتغيير القراءة التاسعة من متغير العمر Age إلى 91 بدلا من 19 (قد يكون خطأ فى الإدخال).

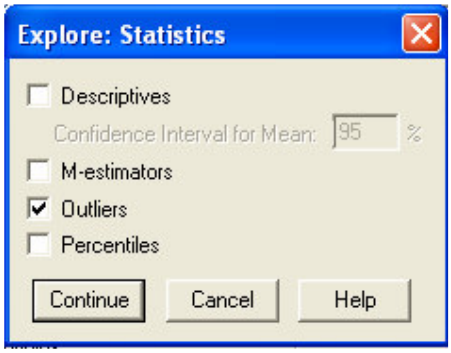
سواء لاكتشاف وجود قيم شاذة فى البيانات أو لاكتشاف وجود أخطاء فى البيانات قبل بدء التحليل لمتغير العمر نتبع الخطوات التالية:

Analyze Descriptive Statistics Explore…

باختيار العمر (Age) فى :Dependent List والضغط على ▶ مع التأشير على Statistics نحصل على الشكل التالي:



والضغط على التأشير مع التأشير على Outliers نحصل على الشكل التالي:



والضغط على Continue ثم OK نحصل على الجدول التالي:

Extreme Values

			Case Number	Value
Age	Highest	1	9	91
		2	5	41
		3	4	35
		4	8	31
		5	1	26[a]
	Lowest	1	3	18
		2	2	21
		3	7	22
		4	6	22
		5	10	26[b]

a. Only a partial list of cases with the value 26 are shown in the table of upper extremes.

b. Only a partial list of cases with the value 26 are shown in the table of lower extremes.

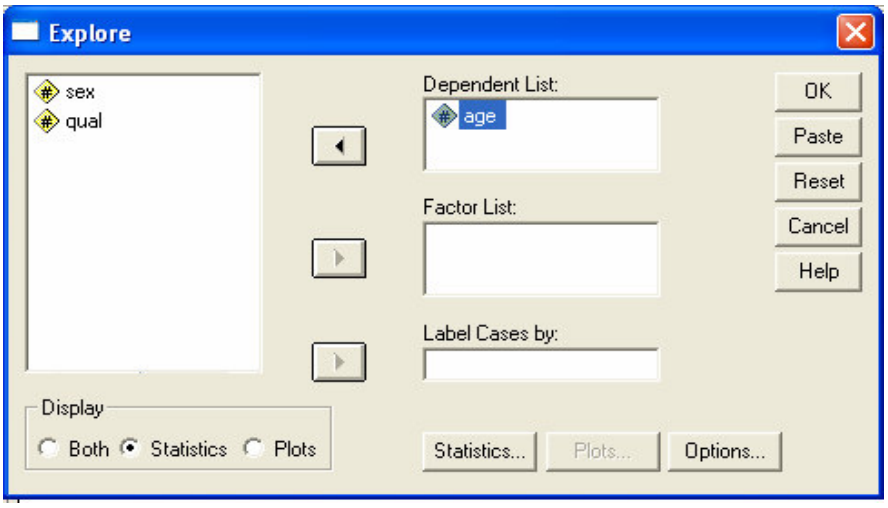
يلاحظ من الجدول السابق ان القيمة 91 للعمر إما أن تكون قيمة شاذة اوتكون خطأ كتابياً. فى مثالنا هذا تعتبر القيمة 91 خطأ كتابياً (بدلا من 19)، من ثم يتم تصحيحه قبل بداية التحليل. ويلاحظ أن الجدول أيضا يعطي رقم الحالة (Case Number) وهى رقم 9 (أي أن الخطأ موجود فى الحالة رقم 9) وذلك لسرعة الوصول إلى القيمة وتصحيحها.

2.5.2 استخدام امر Explore لوصف المتغير الكمى.

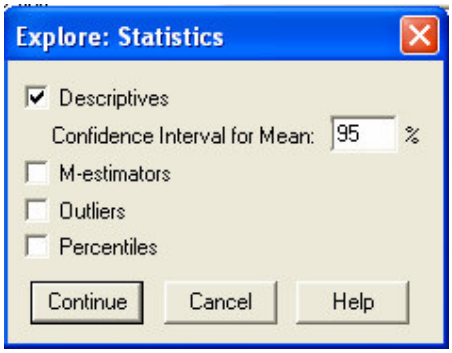
مثال (2.4) باستخدام بيانات المثال السابق بعد تعديل الخطأ او التأكد من عدم وجود خطأ لمتغير العمر نتبع الخطوات التالية:

Analyze Descriptive Statistics Explore…

باختيار العمر (Age) فى Dependent List: والضغط على ▶ مع التأشير على Statistics نحصل على الشكل التالي:



بالضغط على Statistics... مع التأشير على Descriptives نحصل على الشكل التالي:



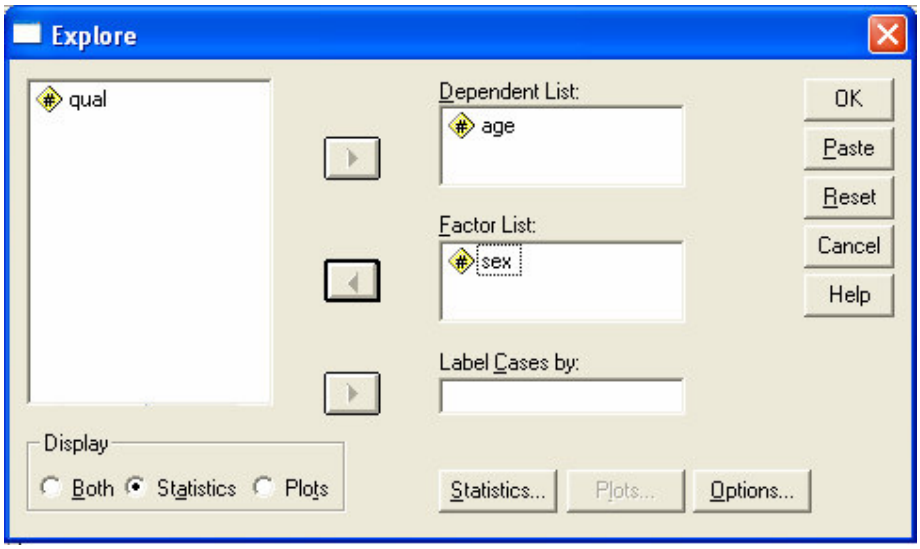
والضغط على Continue ثم OK نحصل على الجدول التالي:

Descriptives

			Statistic	Std. Error
Age	Mean		26.10	2.359
	95% Confidence Interval for Mean	Lower Bound	20.76	
		Upper Bound	31.44	
	5% Trimmed Mean		25.72	
	Median		24.00	
	Variance		55.656	
	Std. Deviation		7.460	
	Minimum		18	
	Maximum		41	
	Range		23	
	Interquartile Range		12	
	Skewness		.994	.687
	Kurtosis		.186	1.334

يوضح الجدول السابق وصفا كاملا للمتغير الكمي (العمر) شاملاً ايضا فترة الثقة عند مستوي معنوية 95%.

من الممكن تكرار الخطوة السابقة مع تقسيم المتغير الكمي حسب المتغير الإسمي او الترتيبي، فباختيار العمر (Age) فى Dependent List: وإدخال متغير النوع (sex) الى Factor List وتكرار الخطوات السابقة نحصل على الشكل التالي:



بالضغط علي OK نحصل على الجدول التالي:

Descriptives

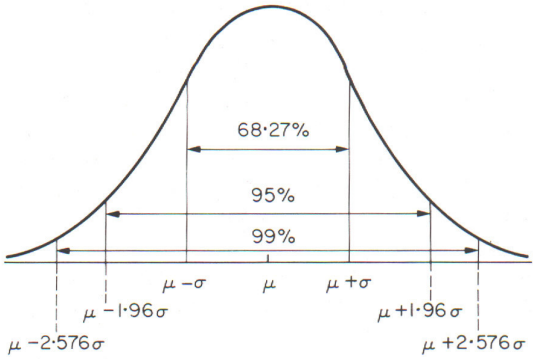
	Sex			Statistic	Std. Error
Age	Male	Mean		23.67	2.539
		95% Confidence Interval for Mean	Lower Bound	17.14	
			Upper Bound	30.19	
		5% Trimmed Mean		23.35	
		Median		22.00	
		Variance		38.667	
		Std. Deviation		6.218	
		Minimum		18	
		Maximum		35	
		Range		17	
		Interquartile Range		10	
		Skewness		1.467	.845
		Kurtosis		2.200	1.741
	Female	Mean		29.75	4.270
		95% Confidence Interval for Mean	Lower Bound	16.16	
			Upper Bound	43.34	
		5% Trimmed Mean		29.61	
		Median		28.50	
		Variance		72.917	
		Std. Deviation		8.539	
		Minimum		21	
		Maximum		41	
		Range		20	
		Interquartile Range		16	
		Skewness		.753	1.014
		Kurtosis		.343	2.619

يوضح الجدول السابق وصفا كاملا للمتغير الكمي (العمر) مقسماً حسب النوع. ويمكن بالطبع تكرار ما سبق مع أي متغير آخر (اسمي او ترتيبي)

ملاحظة: كان من الممكن أيضا تكرار ما سبق مع اختيار Plots وهى تعطي فقط رسوماً بيانيه أو اختيار Both وهى تعطي الإحصاءات الوصفية السابق الإشارة إليها بالإضافة للرسومات البيانية.

2.5.3 استخدام امر Explore لإجراء اختبار الإعتدالية (Normality)

التوزيع الطبيعى The Normal distribution يعتبر واحد من اهم التوزيعات الإحصائية وأكثرها استخداماً فى الاستدلال الإحصائي



2.5.3.a اختبار كولمجورف- سمرنوف- ليليفورز للإعتدالية Kolmogorov-Smirnov- Lilliefors test

*** الهدف**

يستخدم اختبار Kolmogorov-Smirnov-Lilliefors test لاختبار ماذا كانت العينة قد أخذت من توزيع طبيعي(معتدل) عندما يكون كلا من المتوسط والتباين مجهول ويتم تقديرهما من البيانات

طالما كان التوزيع الطبيعي مهم جدا فى الاستدلال الاحصائى، فإننا غالبا ما نختبر فرضية ان بياناتنا آتية من توزيع طبيعي (Normal Distribution). ان أحد طرق عمل ذلك تكون باستخدام Normal probability plot. فى ال Normal probability plot، يتم ازدواج كل قيمة مشاهدة (observed) مع قيمتها المتوقعة (expected) من التوزيع الطبيعي مكونة مجموعة من النقاط.

فاذا كانت العينة مأخوذة من توزيع طبيعي فإننا نتوقع ان تقع النقاط "تقريبا" على خط مستقيم.

* كيفية الوصول إلى الاختبار:

Analyze
Descriptive Statistics
Explore…

* كيفية تنفيذ الاختبار:

- أدراج المتغيرات المراد اختبارها Dependent list

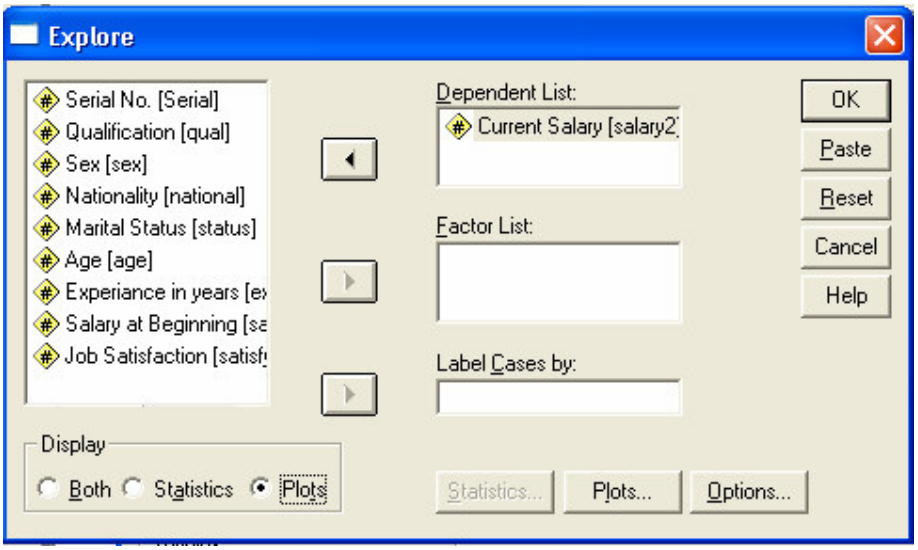
- يمكن اختيار Factor list اذا كان المتغير يمكن تقسيمه الي مجموعات
- اختيار Plots.
- الضغط علي Plots.
- التأشير علي Normality plots with tests.
- الضغط علي Continue
- الضغط علي OK

*** المعنوية تعنى عدم وجود اعتدالية Normality في البيانات.**

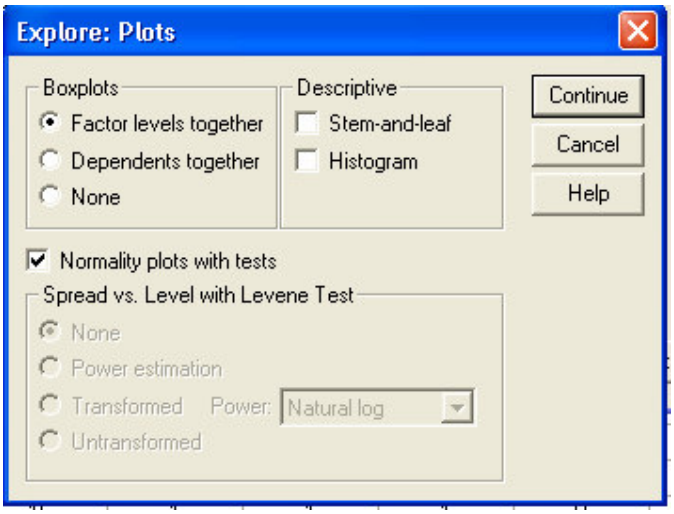
مثال (2.5) باستخدام بيانات المثال الخاص بالموظفين لدراسة ماذا كان الراتب الحالي Current Salary موزع توزيعاً طبيعياً نتبع الخطوات التالية:

Analyze Descriptive Statistics Explore…

باختيار المتغير Current Salary ضمن Dependent List مع التأشير على Plots نحصل على الشكل التالي:



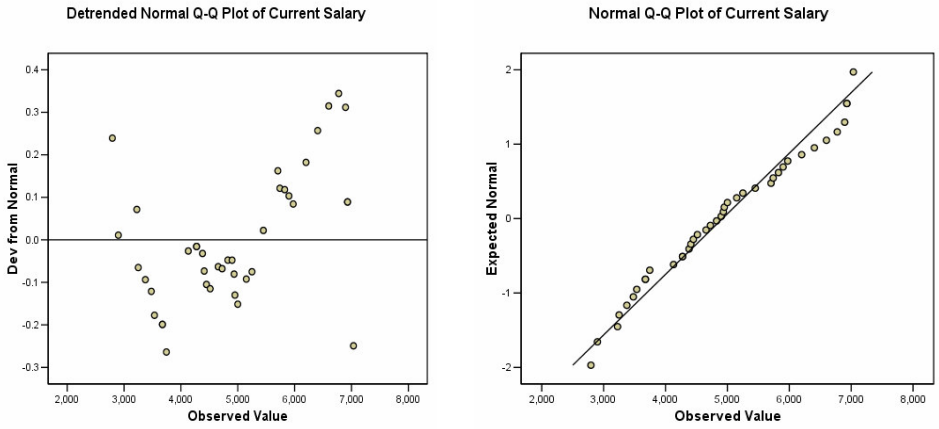
بالضغط على Plots... نحصل على الشكل التالي:



بالتأشير على Normality plots with tests ثم الضغط على Continue نعود للشكل السابق ثم بالضغط على OK نحصل على الناتج التالي:

الرسم على اليمين من الشكل التالى هو ال Normal probability plot للعينة الخاصة بالرواتب الحالية. ونلاحظ قرب النقاط من الخط المستقيم. ذلك يؤكد وجود الإعتدالية.

من الممكن ايضا رسم الانحرافات الفعلية للنقاط عن الخط المستقيم وهو مايعرف باسم Detrended Normal plot كما يتضح من الرسم علي اليسار من الشكل التالى.



فاذا كانت العينة مأخوذة من توزيع طبيعي فإننا نتوقع ان تتقسم النقاط حول الخط الافقى بدون شكل محدد. وايضاً يتضح ان الانحرافات عن الخط الافقي موزعة بشكل عشوائي . ذلك يؤكد وجود الإعتدالية .(وجود شكل محدد يقترح "البعد" عن الإعتدالية).

على الرغم من ال Normal probability plot يمدنا بقاعدة مرئية لاختبار الإعتدالية ، فاننا غالباً ما نقوم بحساب اختبار إحصائي ذو الفرض ان البيانات آتية من توزيع طبيعي (معتدل) . عدم وجود معنوية لهذا الاختبار تعني ان فرض الإعتدالية لايجب رفضة ، وهو ما يؤكده ايضا اختبار Lilliefors والقائم على تعديل اختبار كولمجورف- سمرنوف (Kolmogorov- Smirnov).

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Current Salary	.081	40	.200*	.960	40	.170

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

a. Lilliefors Significance Correction

ويلاحظ ان قيمة احتمال المعنوية فى اختبار كولمجورف- سمرنوف- ليليفورز (Sig.) تساوي 0.200 (اى اكبر من 0.05) مما **يعنى عدم وجود معنوية و يؤكد مطابقة البيانات للتوزيع الطبيعي.**

2.6 اختبار مربع كاي للإستقلالية Chi-square test for independence

*** الهدف**

يستخدم اختبار مربع كاي لدراسة الاستقلال بين متغيرين اسميين

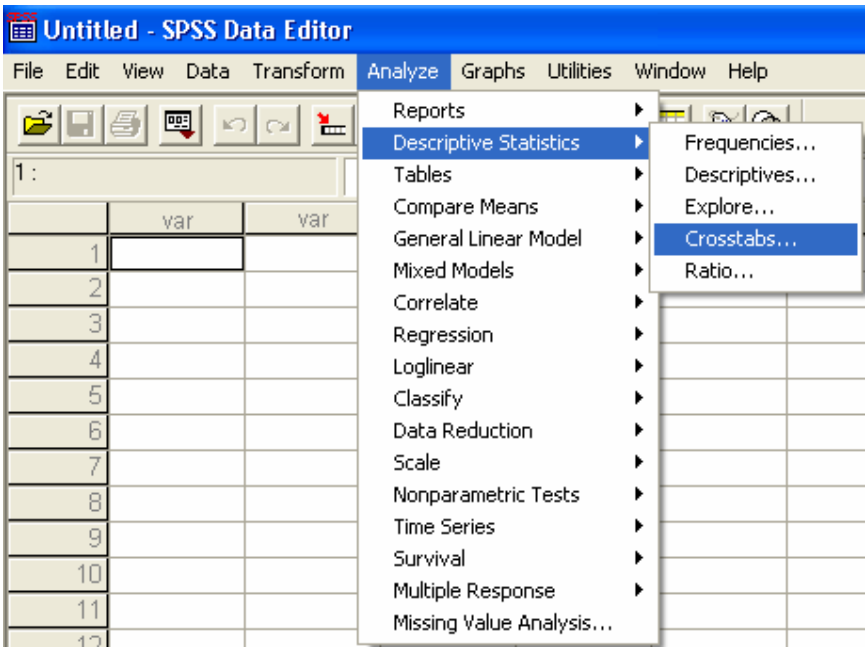
يعتبر اختبار مربع كاي اختباراً لامعلميا هاما، يستخدم لدراسة وجود علاقة بين متغيرين اسميين او متغير اسمي وآخر ترتيبي. ويعتمد على مقارنة القيم المشاهدة (observed) مع القيم المتوقعة (expected)

ومن ضمن شروط صحة تطبيقه ان تكون جميع القيم المتوقعة أكبر من 5.

ويعطي برنامج SPSS تحت الجدول عدد الخلايا المتوقعة ذات القيم أقل من 5 ونسبتها المئوية، حيث انه فى بعض الأحيان يمكن قبول نسبة حتى 20% من القيم المتوقعة اقل من 5.

* كيفية الوصول إلى الاختبار:

Analyze
Descriptive Statistics
Crosstabs…



* كيفية تنفيذ الاختبار:

- أدراج المتغير الأول فى الصف :Row(s)
- أدراج المتغير الثاني فى العمود :Column(s)
- الضغط علي Statistics
- التأشير علي Chi-square
- الضغط علي Continue
- الضغط علي OK.

*** المعنوية تعنى وجود علاقة بين المتغيرين (عدم وجود استقلالية).**

2.7 اعادة تعريف المتغير (ضم الخلايا):

يتم اعادة تعريف المتغير اوضم الخلايا من خلال اختيار :

Transform
Recode
Into Different Variables..

3. الارتباط والانحدار

3.1 معامل بيرسون للارتباط الخطي Pearson's Linear Correlation Coefficient

*** الهدف**

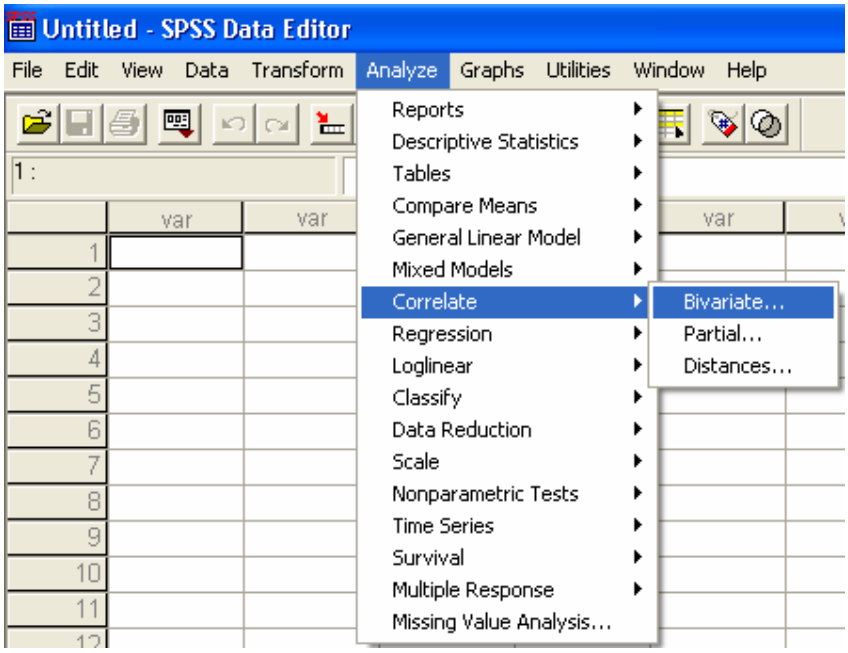
يستخدم هذا الإختبار للتحقق من وجود علاقة إرتباط خطية بين أزواج المشاهدات البيانات ، إضافة إلى مدى قوة واتجاه العلاقة بين المتغيرين

ان اهم اهداف اى بحث هى ايجاد علاقات بين المتغيرات وذلك هو الهدف الأساسي لعلم الاحصاء.
وتتميز دراسة العلاقة بين اى متغيرين بالخاصيتين التاليتين:
1- المقدار (الحجم) magnitude
2- الصلاحية (الصدق) reliability
ان صلاحية العلاقة هى تعبير عن حجم "تمثيل" النتيجة التى نحصل عليها من العينة للمجتمع.

تنحصر قيمة معامل الارتباط بين 1- و 1+. إذا كانت قيمة معامل الإرتباط مساوية 1+ عندها يكون الإرتباط طردي تام، وكذلك عندما تكون قيمة معامل الإرتباط مساوية 1 - عندها يكون الإرتباط عكسي تام.
ويهتم معامل بيرسون بدراسة وجود علاقة خطية. وإذا كانت قيمة معامل الإرتباط مساوية للصفر، فهذا يشير إلى عدم وجود علاقة إرتباط خطي بين المتغيرين (قد تكون هناك علاقة بين المتغيرين ولكنها غير خطية). ويشترط للمتغيرات ان تتبع التوزيع الطبيعي الثنائي .
وتكمن أهمية هذا الإختبار، عندما يأخذ معامل الإرتباط، أي من القيم المحصورة مابين 1- و 1+، عندها نبحث عن معنوية معامل الإرتباط.

* كيفية الوصول إلى الاختبار:

Analyze
Correlate
Bivariate …



* كيفية تنفيذ الاختبار:

- أد راج المتغيرين او المتغيرات المراد اختبارها في Variables
- التأشير علي Pearson
- اختيار نوع الاختبار : ذو طرفان Two-tailed او ذو طرف واحد One-tailed
- الضغط علي OK

*** المعنوية تعنى وجود علاقة ارتباط خطية .**

يلاحظ استخدام معامل بيرسون فى حالة وجود علاقة خطية لبيانات كمية تتبع توزيع طبيعى، واذا لم يتحقق ذلك او كانت البيانات ترتيبة يستخدم معامل سبيرمان (لدراسة اضطراد العلاقة) او معامل كندال تاو (لدراسة وجود توافق)

الإنحـــــــــدار

Regression

1

الإنحدار الخطي Linear Regression

*** الهدف**

يستخدم الانحدار الخطي لتقدير معامل المتغير المستقل للمعادلة الخطية بغرض تقدير المتغير التابع

يعرف خط الانحدار بأنه الخط الذى تكون مجموع مربعات انحرافات النقاط عنه اقل ما يمكن.

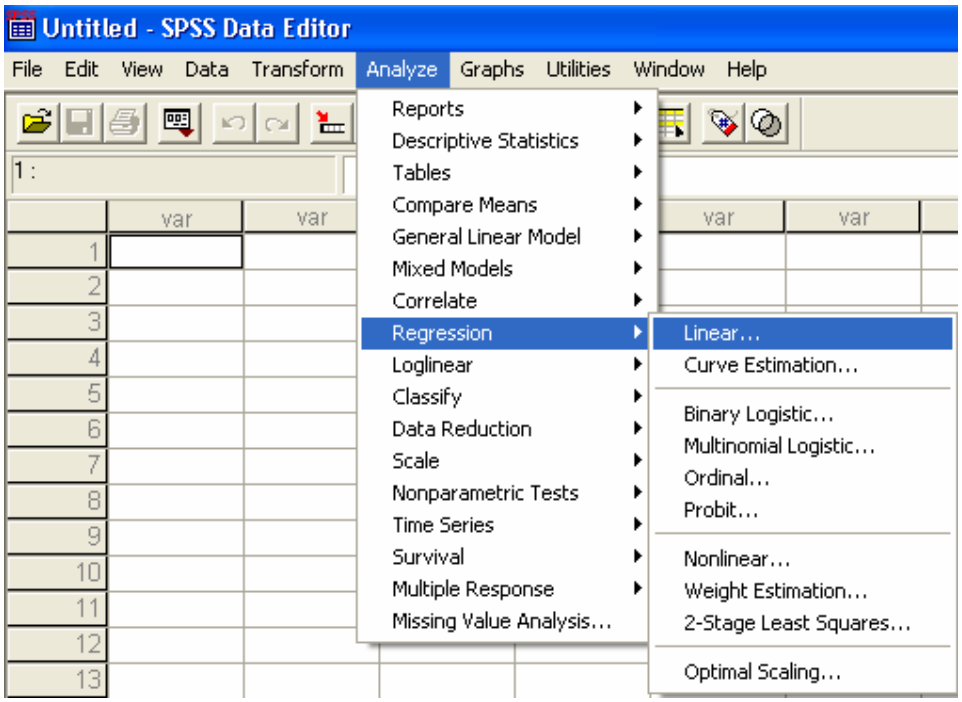
فى حالة وجود متغير مستقل واحد فإن معادلة الخط تأخذ الصورة:

$$Y = a + b*X$$

حيث تعبر X عن المتغير المستقل وتعبر Y عن المتغير التابع.

* كيفية الوصول إلى الاختبار:

Analyze
Regression
Linear…



* كيفية تنفيذ الاختبار:
- أد راج المتغير التابع فى Dependent
- أد راج المتغير المستقل فى Independent(s)
- الضغط علي OK.

*** المعنوية تعنى مناسبة النموذج للبيانات.**

2
الإنحدار الخطى المتعدد Multiple Linear Regression

*** الهدف**

يستخدم الانحدار الخطي المتعدد لتقدير معاملات المتغيرات المستقلة للمعادلة الخطية بغرض تقدير المتغير التابع

فى حالة وجود عدة متغيرات مستقلة فإن معادلة الخط تأخذ الصورة:

$$Y = a + b_1*X_1+b_2*X_2+\ldots+b_k*X_k$$

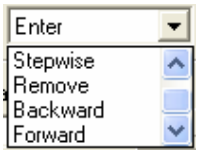
حيث تعبر $X_1, X_2,\ldots, X_k$ عن المتغيرات المستقلة وتعبر Y عن المتغير التابع.

* كيفية الوصول إلى الاختبار:

Analyze
Regression
Linear…

* كيفية تنفيذ الاختبار:
- أد راج المتغير التابع فى Dependent
- أد راج المتغيرات المستقلة فى Independent(s)
- الضغط علي OK.

*** المعنوية تعنى مناسبة النموذج للبيانات.**



فى حالة المتغيرات المتعددة توجد خمسة اختيارات:

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).

4. مقارنة بالمتوسطات

4.1 إختبار " ت" لمتوسط المجتمع T-test for a Population Mean

*** الهدف**

يستخدم إختبار " ت" للتحقق من ماإذا كان هناك فرق معنوي بين متوسط المجتمع μ ومتوسط فرضي μ_0 ، حيث μ_0 يمثل رقم ثابت.

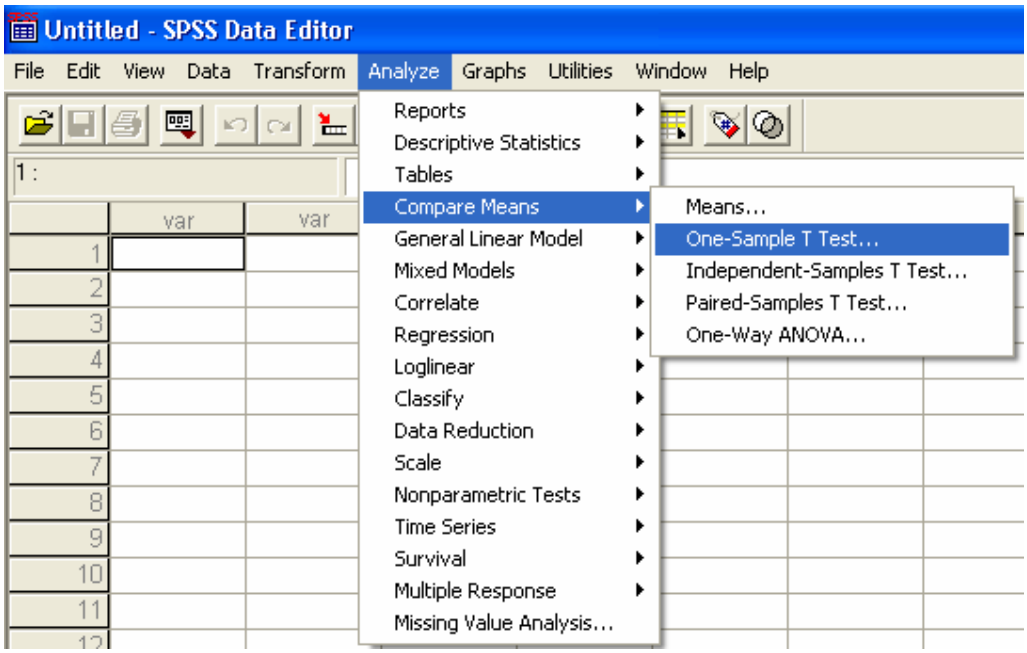
يعتبر اختبار "ت" اختباراً معلمياً هاما، يستخدم لدراسة متوسط المجتمع في حالة العينات الصغيرة (اقل من 30 مشاهدة) ويشترط لتطبيقه ان تكون:

1. البيانات مسحوبة من التوزيع الطبيعي.
2. الإنحراف المعياري σ للمجتمع غير معلوم.

(ملاحظة: فى حالة العينات الكبيرة (ذات الحجم 30 فأكثر) يمكن الاستغناء عن القيود المذكورة، وعلي الرغم من ان البرنامج يعتبر الاختبار أيضا اختبار "ت" t-test الا انه في الواقع يعرف إحصائيا باسم اختبار "ص" او Z test)

- **كيفية الوصول الى الإختبار**

Analyze
Compare Means
One-Sample T-Test …



- **كيفية تنفيذ الإختبار**

1. إدراج المتغير المراد إختياره الي :Test variable(s)
2. تحديد قيمة μ_0 (Test value) ، البرنامج عادة يضع قيمة $\mu_0 = 0$ ولكن بالإمكان تغيير هذه القيمة بما يتلاءم مع القيمة المراد اختبارها.
3. لحساب فترة الثقة Confidence Interval ، يتم الضغط على كلمة إختيارات (Options).
4. البرنامج عادة يحسب فترة الثقة عند 95% ، ولكن بالإمكان تغيير هذه القيمة بما يتلائم مع طبيعة البحث وليكن 99% او غيرها، ثم يتم الضغط على كلمة مواصلة (Continue).

المعنوية تعني أن متوسط المجتمع لا يساوي μ_0

4.2 إختبار " ت" لعينتين مستقلتين Independent-Samples T-Test

*** الهدف**

يستخدم اختبار " ت" لعينتين مستقلتين للتحقق من ما إذا كانت هناك فروق معنوية بين متوسطي مجتمعين مستقلين

يعتبر اختبار "ت" لعينتين مستقلتين اختباراً معلمياً هاما، يستخدم لدراسة تساوي متوسطي مجتمعين مستقلتين في حالة العينات الصغيرة (اقل من 30 مشاهدة) ويشترط لتطبيقه ان تكون:

1. البيانات لكل من المجتمع الأول والمجتمع الثاني تتبع التوزيع الطبيعي.
2. الإنحراف المعياري لكل من المجتمع الأول والمجتمع الثاني غير معلوم.
3. الإنحراف المعياري للمجتمع الأول σ_1، يساوي الإنحراف المعياري للمجتمع الثاني σ_2 (متجانسان).
4. بيانات المجتمع الأول مستقلة عن بيانات المجتمع الثاني.

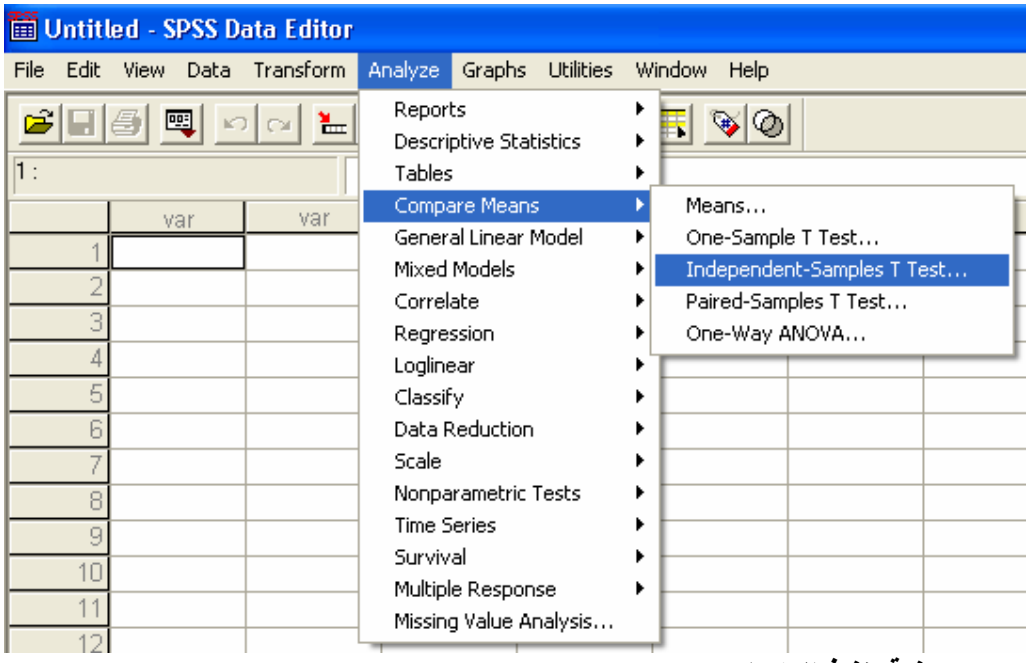
(كما ذكرنا في الاختبار السابق انه في حالة العينات الكبيرة (ذات الحجم 30 فأكثر) يمكن الاستغناء عن القيود المذكورة، وان الاختبار يعرف إحصائيا باسم اختبار "ص" او Z test)

- **كيفية الوصول الى الاختبار**

Analyze

Compare Means

Independent-Sample T-Test …



- **كيفية تنفيذ الإختبار**

1. إدراج المتغير المراد إختياره الي :Test variable(s)
2. تحديد قيمة μ_0 (Test value) ، البرنامج عادة يضع قيمة $\mu_0 = 0$ ولكن بالإمكان تغيير هذه القيمة بما يتلاءم مع القيمة المراد اختبارها.

3. لحساب فترة الثقة Confidence Interval ، يتم الضغط على كلمة إختيارات (Options) .
4. البرنامج عادة يحسب فترة الثقة عند 95% ، ولكن بالإمكان تغيير هذه القيمة بما يتلائم مع طبيعة البحث وليكن 99% او غيرها، ثم يتم الضغط على كلمة مواصلة (Continue) .

المعنوية تعني أن متوسط المجتمع الأول لا يساوي متوسط المجتمع الثاني

4.3 إختبار " ت" لعينتين مرتبطتين Paired-Samples T-Test

*** الهدف**

يستخدم اختبار " ت" لعينتين مرتبطتين للتحقق من ما إذا كانت هناك فروق معنوية بين متوسطي مجتمعين مرتبطين (اختبار قبل – بعد)

يعتبر اختبار "ت" لعينتين مستقلتين اختباراً معلمياً هاما، وهو يفترض بأن طبيعة البيانات تعتمد على بعضها البعض أي انها غير مستقلة (مرتبطة) Dependent Data .

يستخدم الاختبار لدراسة تساوي متوسطي مجتمعين مرتبطين ويعرف باسم الاختبار قبل وبعد . في حالة العينات الصغيرة (اقل من 30 مشاهدة) يشترط لتطبيقه ان تكون البيانات مسحوبة من التوزيع الطبيعي.

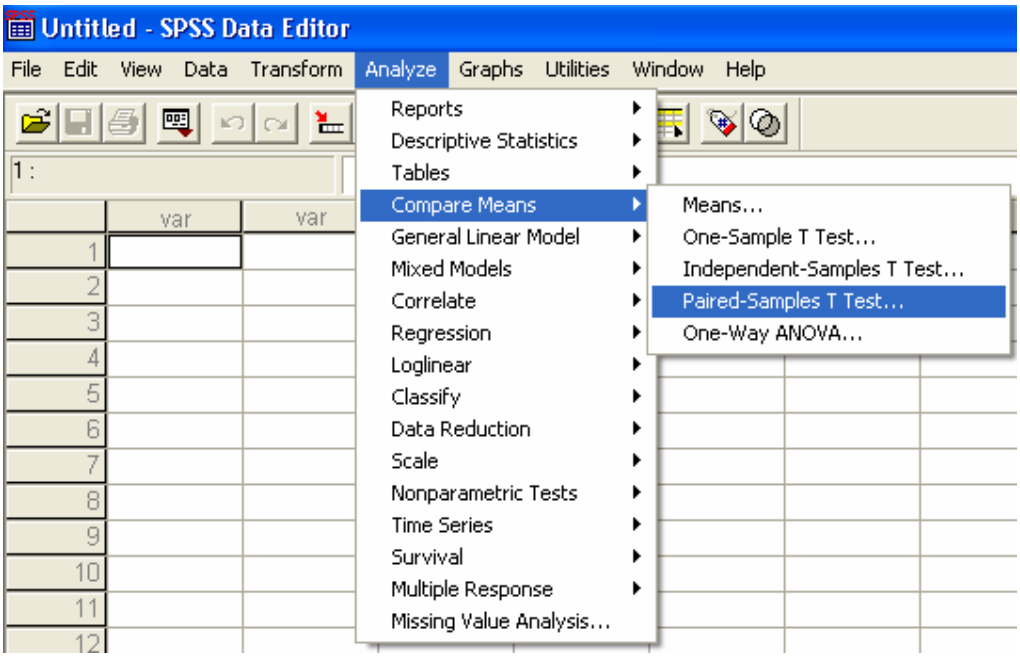
(فى حالة العينات الكبيرة (ذات الحجم 30 فأكثر) يمكن الاستغناء عن القيد المذكور، وان الاختبار يعرف إحصائيا باسم اختبار "ص" او Z test).

- **كيفية الوصول الى الاختبار**

Analyze

Compare Means

Paired-Sample T-Test …



- **كيفية تنفيذ الإختبار**

1. إدراج المتغير الأول (قبل) والمتغير الثاني (بعد) ، إلى Paired Variables .
2. لحساب فترة الثقة Confidence Interval ، يتم الضغط على كلمة اختيارات (Options) .
3. البرنامج عادة يحسب فترة الثقة عند 95% ، ولكن بالإمكان تغيير هذه القيمة بما يناسب رغبة الباحث وليكن 99% أو غيرها، ثم يتم الضغط على كلمة مواصلة (Continue) .
4. الضغط على OK .

المعنوية تعني أن متوسط المجتمع قبل التجربة لا يساوي متوسط المجتمع بعد التجربة.

4.4 إختبار تحليل التباين فى اتجاه واحد One-Way-ANOVA

*** الهدف**

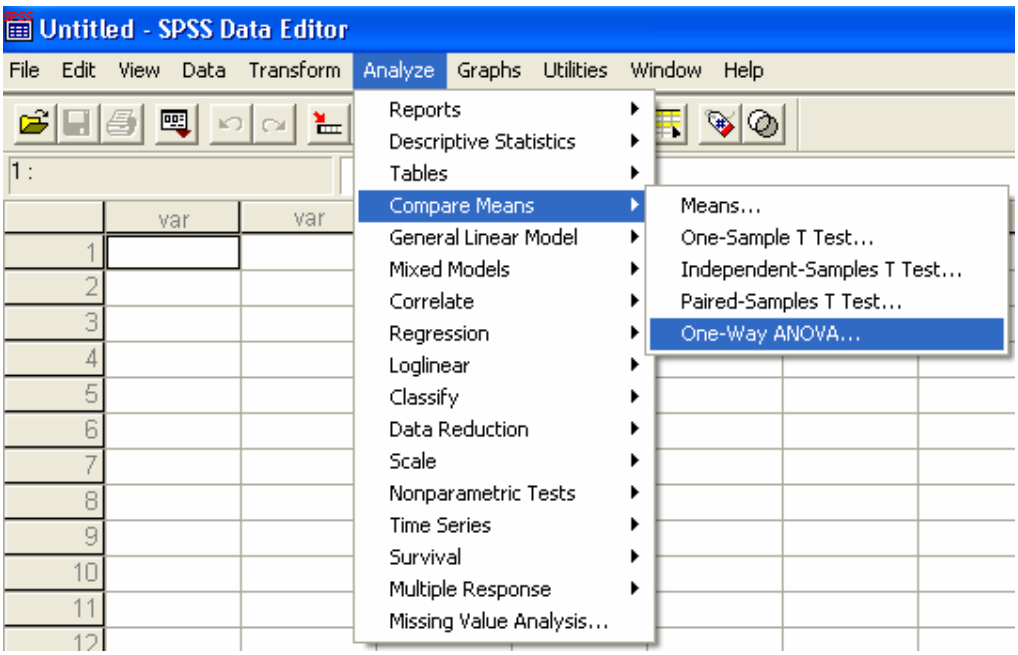
يستخدم إختبار تحليل التباين فى اتجاه واحد او مايعرف باختبار " ف" لعينات مستقلة للتحقق من ما إذا كانت هناك فروق معنوية بين متوسطات اكثر من مجتمعين مستقلين

يعتبر اختبار تحليل التباين (او اختبار "ف") اختبارا معلمياً هاما، يستخدم لدراسة تساوي متوسطات اكثر من مجتمعين مستقلتين. في حالة العينات الصغيرة (اقل من 30 مشاهدة) يشترط لتطبيقه ان تكون:

1. البيانات تتبع التوزيع الطبيعي.
2. بيانات المجتمعات مستقلة عن بعضها البعض.

- **كيفية الوصول الى الاختبار**

Analyze
Compare Means
One-Way ANOVA…



- **كيفية تنفيذ الإختبار**

1. إدراج المتغير المراد اختباره الى Dependent List
2. ادراج المتغير الذى يعبر عن العينات المستقلة الى Factor
3. الضغط على OK .

المعنوية تعني أن متوسطات المجتمعات المستقلة غير متساوية.

5. اختبارات لا معلمية

5.1 اختبار جودة التوفيق Goodness of fit test

*** الهدف**

يستخدم اختبار جودة التوفيق للتأكد من ان نموذج ما يناسب البيانات بشكل جيد

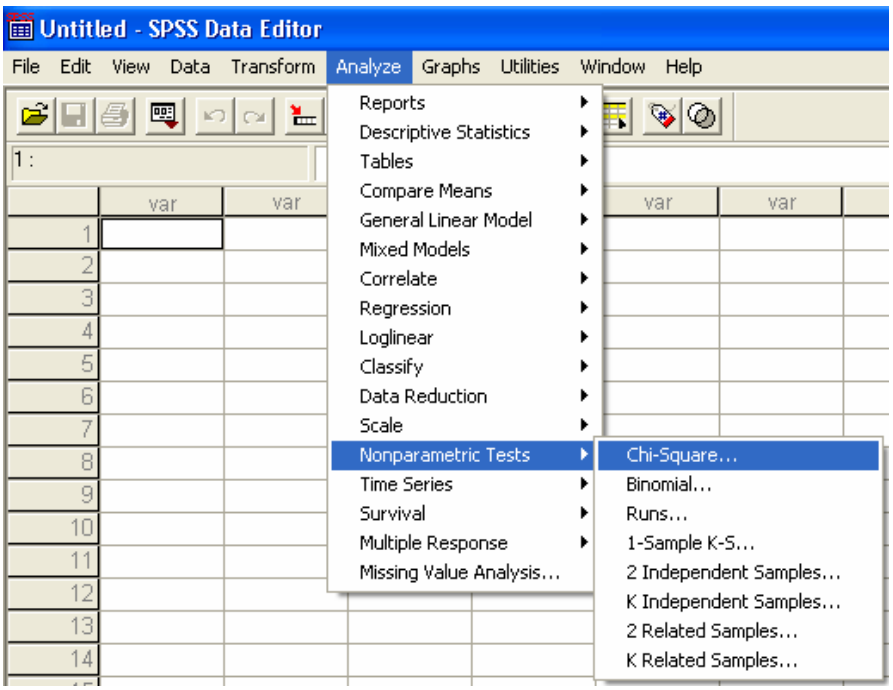
يعتبر اختبار مربع كاي اختباراً لجودة التوفيق اختباراً لامعلمياً هاما، يستخدم لدراسة التوزيع الاحتمالي للمتغير. ويعتمد على مقارنة القيم المشاهدة (observed) مع القيم المتوقعة (expected) من التوزيع الاحتمالي.

ومن ضمن شروط صحة تطبيقة ان تكون جميع القيم المتوقعة أكبر من 5.

ويعطي برنامج SPSS تحت الجدول عدد الخلايا المتوقعة ذات القيم أقل من 5 ونسبتها المئوية، حيث انه في بعض الأحيان يمكن قبول نسبة حتى 20% من القيم المتوقعة اقل من 5.

* كيفية الوصول إلى الاختبار:

Analyze
Nonparametric tests
Chi-square…



* كيفية تنفيذ الاختبار:

- أد راج المتغير المراد اختباره Test variable
- اختيار All categories equal في حالة تساوي الاحتمالات (توزيع منتظم) **أو:**
- اختيار Values في حالة عدم تساوي الاحتمالات مع تحديد قيم الاحتمالات المختلفة.
- الضغط علي OK.

*** المعنوية تعنى عدم مناسبة النموذج للبيانات.**

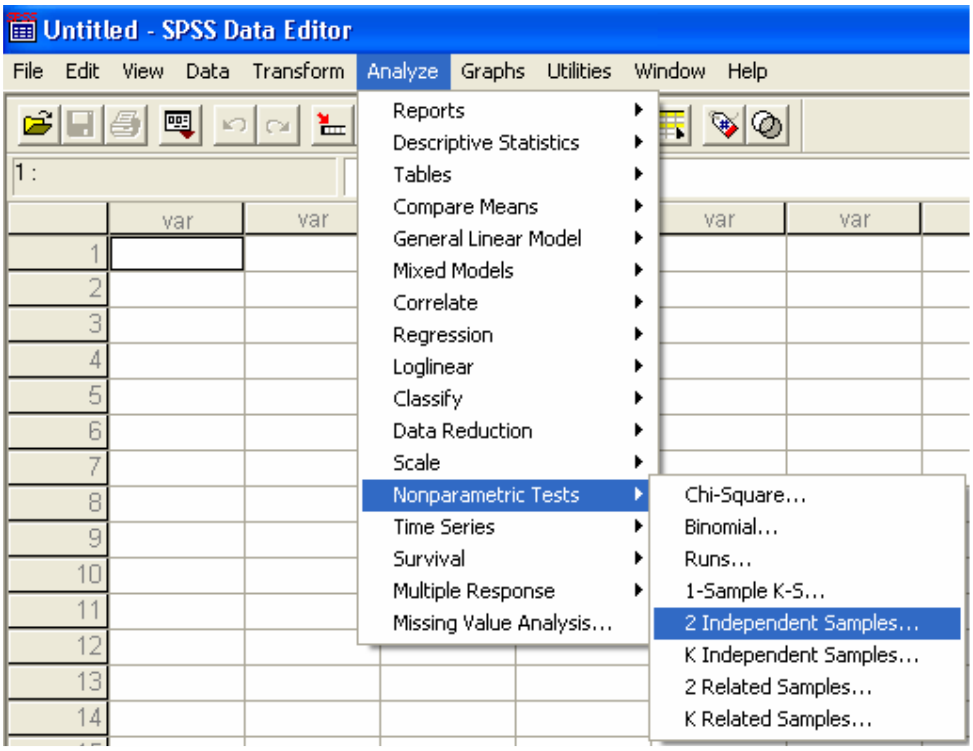
5.2 اختبار مان - ويتني Mann - Whitney U test

*** الهدف**

يستخدم اختبار Mann Whitney U test لاختبارتساوي وسيطي مجتمعين مستقلتين

* كيفية الوصول إلى الاختبار:

Analyze
Nonparametric tests
2- Independent Samples …



* كيفية تنفيذ الاختبار:

- أد راج المتغيرات المراد اختبارها Test variable list
- اختيار المتغير المقسم الي عدة مجموعات
- تحديد اية مجموعتين يراد اجراء الاختبار لهما (من خلال Grouping variable باستخدام Define groups... وذلك بتحديد المجموعة الأولي والمجموعة الثانية)
- التأشير علي Mann-Whitney U
- الضغط علي OK.

*** المعنوية تعنى عدم مطابقة وسيطي المجتمعين.**

5.3 اختبار كروسكال - واليس Kruskal - Wallis H test

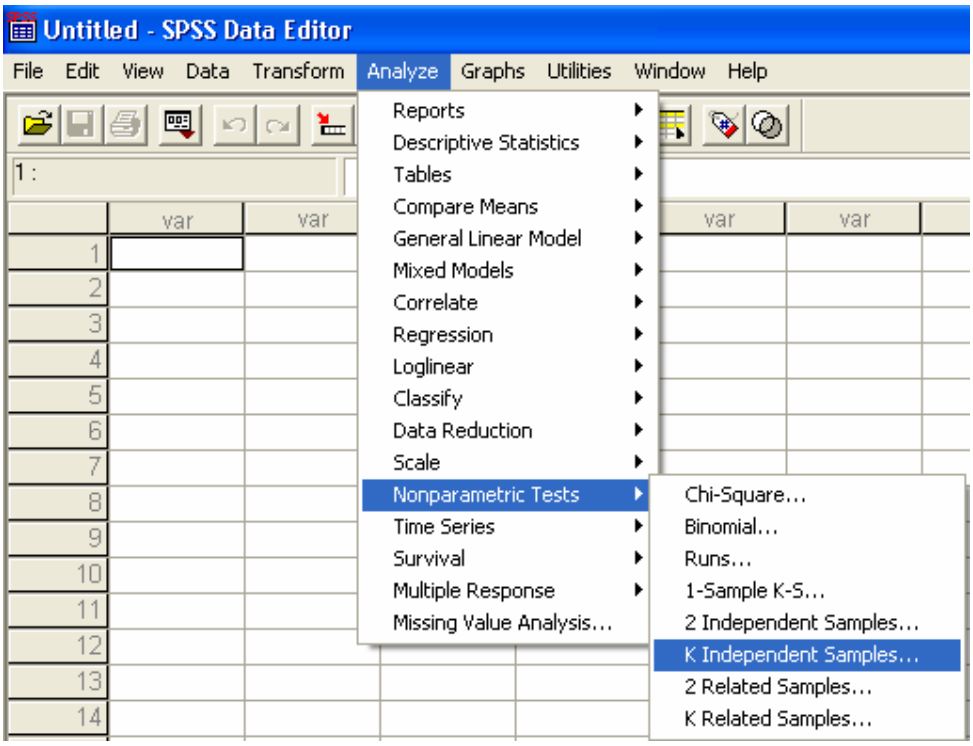
*** الهدف**

يستخدم اختبار Kruskal-Wallis H لاختبار تساوي أوسطة عدة مجتمعات مستقلة

يعتبر اختبار كروسكال واليس اختباراً لامعلمياً هاماً مناظراً لاختبار تحليل التباين ANOVA المستخدم فى الحالة المعلمية لمقارنة متوسطات عدة مجتمعات مستقلة. وهو يستخدم لاختبار تساوي أوسطة عدة مجتمعات مستقلة او لاختبار ماذا كانت k عينة مستقلة معرفة بعدة مجموعات جميعها تتبع نفس التوزيع.

* كيفية الوصول إلى الاختبار:

Analyze
Nonparametric tests
K- Independent Samples …



* كيفية تنفيذ الاختبار:

- أدراج المتغيرات المراد اختبارها Test variable list
- اختيار المتغير المقسم الي عدة مجموعات
- تحديد اية مجتمعات يراد إجراء الاختبار لها (من خلال Grouping variable باستخدام Define range... وذلك بتحديد المجموعة الأدنى والمجموعة الأعلى)
- التأشير علي Kruskal-Wallis H
- الضغط علي OK.

*** المعنوية تعنى عدم مطابقة أوسطة المجتمعات.**

5.4 اختبار ويلكوكسون لعينتان مرتبطتان Wilcoxon two related samples test

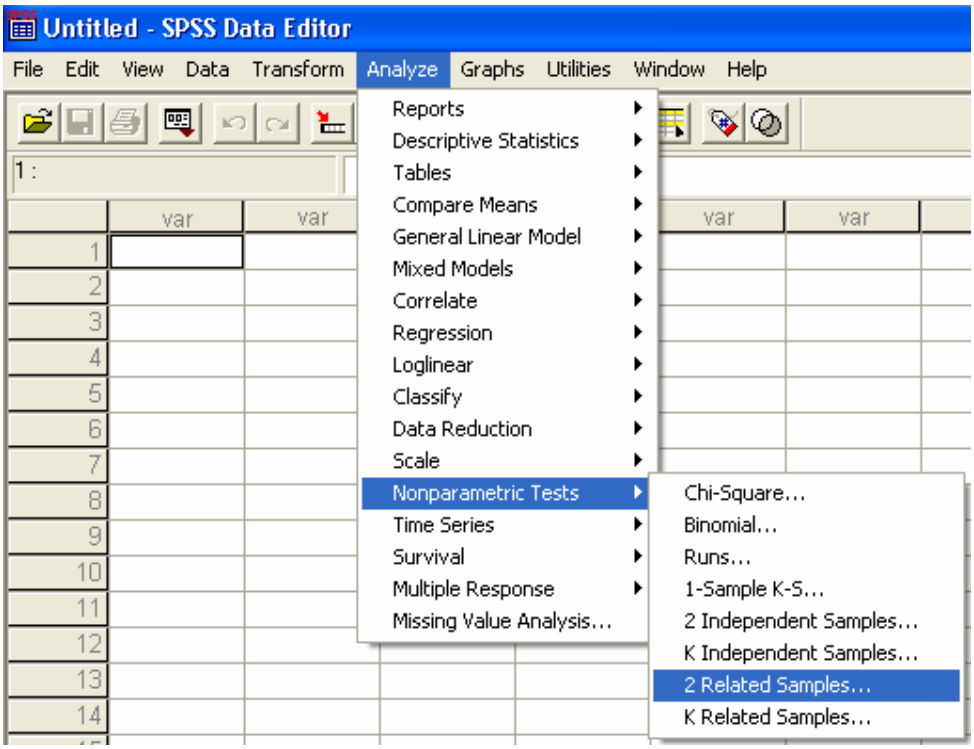
*** الهدف**

يستخدم اختبار Wilcoxon لاختبار تساوي وسيطي مجتمعين مرتبطين (اختبار قبل – بعد)

يعد اختبار ويلكوكسون اختبار مناظر ولكن اقوي من اختبار الإشارة *ويستخدم ايضا لمقارنة وسيطي مجتمعين مرتبطين.*

* كيفية الوصول إلى الاختبار:

Analyze
Nonparametric tests
2- Related Samples …



* كيفية تنفيذ الاختبار:

- أد راج المتغيرات المراد اختبارها في Test pairs list وذلك عن طريق اختيار المتغيرين آنيا (تباعاً).
- التأشير علي Wilcoxon.
- الضغط علي OK.

*** المعنوية تعنى عدم مطابقة الوسيطين.**

5.5 اختبار فريدمان لعدة عينات مرتبطة Friedman k- related samples test

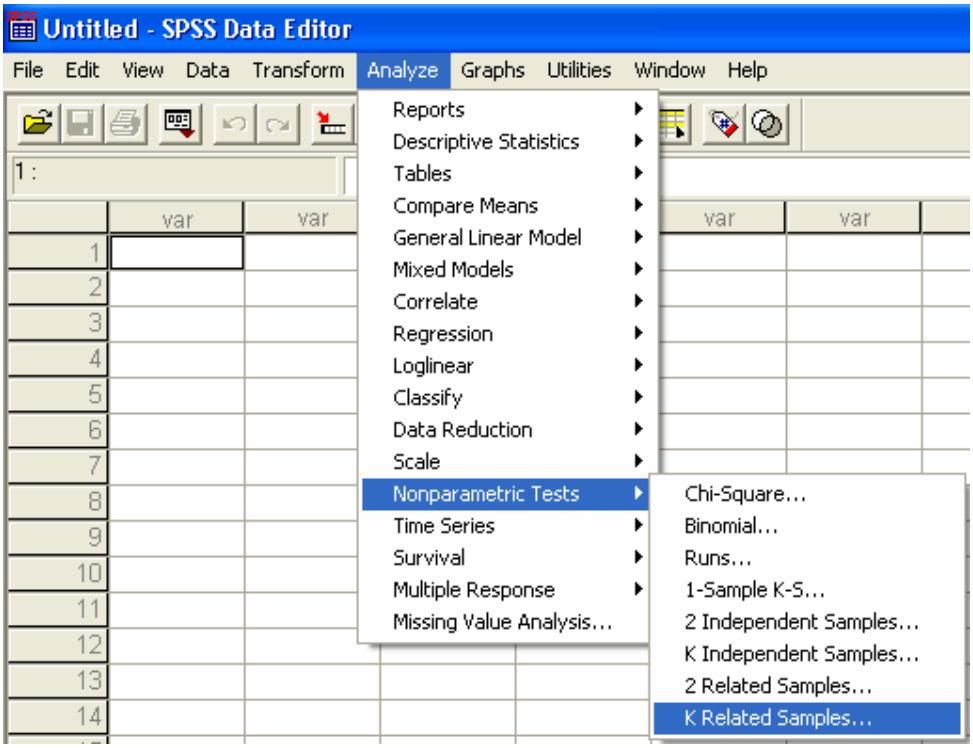
*** الهدف**

يستخدم اختبار Friedman لاختبار تساوي أوسطة عدة مجتمعات مرتبطة

يعتبر اختبار فريدمان اختباراً لامعلميا مكافئاً لتصميم قياسات متكررة لعينة واحدة او تحليل التباين فى اتجاهين مع وجود مشاهدة واحدة داخل كل خلية. يعتمد الاختبار علي وجود بيانات ترتيبية متكررة.

* كيفية الوصول إلى الاختبار:

Analyze
Nonparametric tests
k- Related Samples …



* كيفية تنفيذ الاختبار:

- أدراج المتغيرات المراد اختبارها في Test variables (وذلك عن طريق اختيار المتغيرات آنيا.
- التأشير علي Friedman
- الضغط علي OK.

*** المعنوية تعنى عدم مطابقة الأوسطة.**

5.6 معامل كندال دبليو للتوافق Kendall's W Coefficient of Concordance

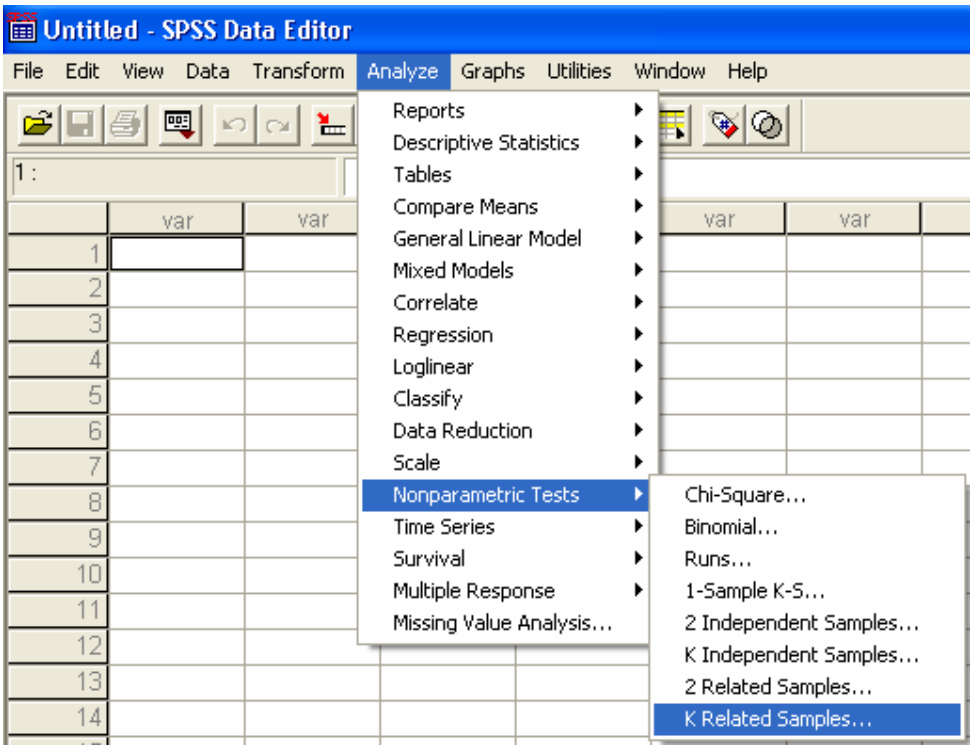
*** الهدف**

يستخدم معامل Kendall W لدراسة مدي التوافق بين عدة متغيرات

يعتبر اختبار كندال دبليو اختبارا ًلامعلميا لقياس مدى التوافق بين المحكمين. كل حالة تعتبر محكم وكل متغير يعتبر "وحدة" او "شخص" يقيم. تتراوح قيمة معامل كندال W بين الصفر (لا توافق) والواحد (توافق تام).

* كيفية الوصول إلى الاختبار:

Analyze
Nonparametric tests
k- Related Samples …



* كيفية تنفيذ الاختبار:

- أد راج المتغيرات المراد اختبارها في Test variables وذلك عن طريق اختيار المتغيرات آنيا.
- التأشير علي Kendall' s W
- الضغط علي OK.

*** المعنوية تعنى عدم التوافق.**

5.7 معامل كوكران كيو للتوافق Cochran's Q Coefficient of Concordance

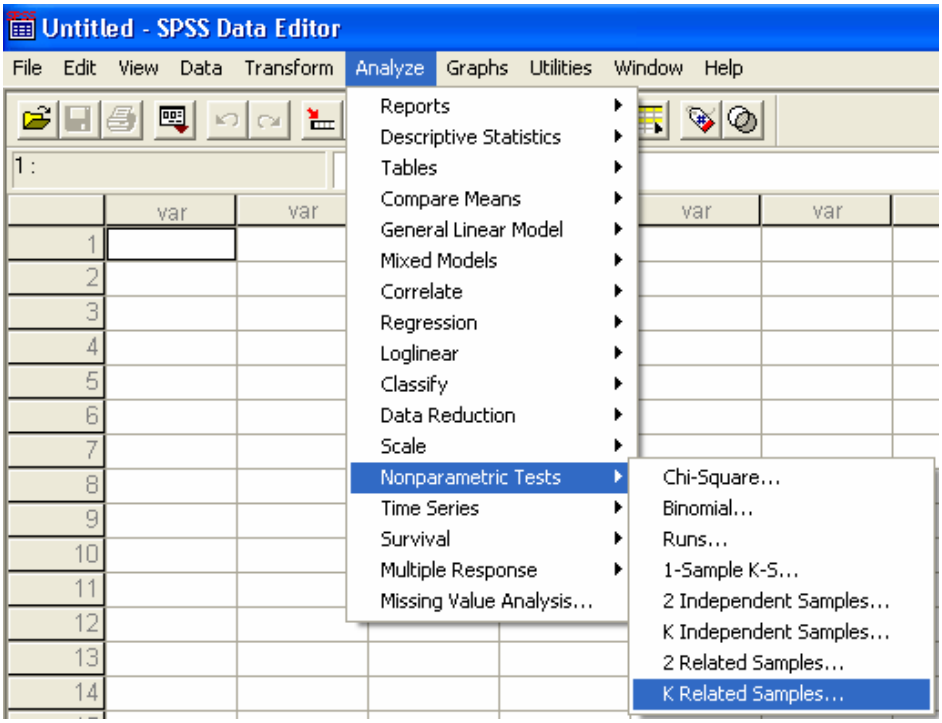
*** الهدف**

يستخدم معامل Cochran Q لاختبار تساوي تأثيرات عدة معالجات ثنائية dichotomous و مرتبطة

يعتبر اختبار كوكران كيو اختباراً لامعلمياً مطابقاً لاختبار فريدمان Friedman test ولكنه يطبق عندما تكون الاستجابات ثنائية (Binary). وهو يعد امتداداً لاختبار ماكنيمار McNemar test لحالة K من العينات.

* كيفية الوصول إلى الاختبار:

Analyze
Nonparametric tests
k- Related Samples …



* كيفية تنفيذ الاختبار:

- أدراج المتغيرات المراد اختبارها في Test variables وذلك عن طريق اختيار المتغيرات آنيا.
- التأشير علي Cochran's Q
- الضغط علي OK.

*** المعنوية تعنى عدم مطابقة الأوسطة.**