

استخدامات الحاسوب في الاقتصاد



استخدامات الحاسوب في الاقتصاد

الدكتور
محمد كبيه
أستاذ بقسم العلوم المالية والمصرفية
UNIVERSITY
ALEPPO

مديرية الكتب والمطبوعات الجامعية

١٤٢٧ هـ - ٢٠٠٧ م

لطلاب السنة الرابعة

قسم العلوم المالية والمصرفية





استخدامات الحاسوب في الاقتصاد



١٦٨	٢-٨ تحليل الانحدار Regression Analysis
١٦٩	١-٢ فرضيات الانحدار الخطي البسيط Assumptions Underlying Simple Linear Regression
١٦٩	٢-٢ طريقة المربعات الصغرى Least Square Method
١٧٠	٣-٢ معامل التحديد The Coefficient of Determination
١٧١	٤-٢ اختبار تحليل التباين
١٧١	١-٤-٢ اختبار الفرضية الأولية باستخدام الإحصائية t
١٧٦	٢-٤-٢ جدول تحليل التباين
١٨١	٣-٨ الانحدار الخطي المتعدد Multiple Linear Regression
١٨٦	١-٣ اختبار وجود مشكلة الازدواج الخطي
	الفصل الخامس
	استكشاف البيانات Explore
١٩٧	١-٨ استكشاف البيانات بواسطة الامر Explore
١٩٩	١-١ مخطط Boxplots
١٩٩	أولاً: الصندوق Box
٢٠١	ثانياً: الاستطالات خارج الصندوق Whiskers
٢٠١	ثالثاً: القيم الشاذة Outliers والقيم المتطرفة Extremes
٢٠١	٢-١ المخططات الوصفية Descriptives
٢٠١	أولاً: مخطط الجذع والورقة Stem – and – Leaf
٢٠٤	ثانياً: المدرج التكراري Histogram
٢٠٥	٣-١ مخططات لاختبار التوزيع الطبيعي Normality Plots with Tests
٢٠٥	٤-١ اختبار تجانس التباين (اختبار ليفن) Spread vs. Level with Levene Test
٢١٠	٥-١ اختبار فرضية كون البيانات تتبع التوزيع الطبيعي من نسبة معامل الالتواء
٢١٤	٦-١ اختبارات التوزيع الطبيعي Normality Plots with Tests
٢١٤	١-٦-١ اختبار كولموغروف سميرونوف
٢١٥	٢-٦-١ مخطط Normal Q-Q Plot
٢١٦	٣-٦-١ مخطط Detrended Normal Q-Q Plot
	الفصل السادس
	تحليل الاستجابات المتعددة
	Multiple Response Analysis
٢١٧	مقدمة
٢١٨	١-٨ أسلوب التكرارات متعددة الاستجابة Multiple Response

٨٠	٢-٨ الحصول على الدرجات المعيارية Producing Z Scores
٨٢	٣-٨ بناء مجال ثقة حول المتوسط
	Producing Confidence Intervals Around a Mean
٨٤	٤-٨ العلاقات بين متغيرين : الجدول التقاطعي
٨٤	١-٤ إنشاء جدول بمتغيرين Producing Bivariate Tables
١٠٠	٢-٤ معامل ارتباط فاي : Phi Coefficient
١٠٦	٣-٤ معامل التوافق (التصاحب) Contingency Coefficient
١١١	٤-٤ معامل الاقتران الرباعي Coefficient of Association
١١٥	٥-٨ اختبار كاي تربيع Chi-Square Test
١١٩	١-٥ اختبار مدى تطابق التوزيع المتوقع مع التوزيع الحقيقي
	الفصل الثالث
	مقارنة المتوسطات Compare Means
١٢٣	١-٨ مقارنة المتوسطات Compare Means
١٢٣	١-١ تحليل المتوسطات Means Analysis
١٣١	٢-١ اختبار التأثير الخطي Test of Linearity
١٣٦	٣-١ اختبار T لعينة واحدة One Sample T-Test
١٣٩	٤-١ اختبار الفرق بين متوسطين
	Producing a Test of Mean Differences
١٣٩	أولاً: اختبار T للفرق بين عينتين مستقلتين
	Independent-Samples T Test
١٤٢	ثانياً: اختبار T للمشاهدات المزدوجة Paired-Samples T Test
١٤٥	٢-٨ تحليل التباين في اتجاه واحد
	One-Way analysis of Variance
١٤٨	١-٢ الفرضيات Hypothesis
١٥٣	٢-٢ تجانس التباين Homogeneity – of- variance
	الفصل الرابع
	تحليل الارتباط والانحدار
	Correlation and Regression Analysis
١٥٧	١-٨ الارتباط Correlation
١٥٧	١-١ معامل الارتباط The Correlation Coefficient
١٥٨	٢-١ معامل سبيرمان لارتباط الرتب Spearman's Rank Correlation
	Coefficient
١٦٠	٣-١ استخدام برنامج SPSS في دراسة الارتباط
١٦٢	٤-١ الارتباط الجزئي Partial Correlation

الفصل الأول

البرنامج الاحصائي SPSS

حزمة البرامج الإحصائية للعلوم الاجتماعية

١٥	مقدمة
١٦	١-١ تشغيل برنامج SPSS
١٧	٢-١ أنواع الملفات في برنامج SPSS
١٨	٣-١ تهيئة الملفات في برنامج SPSS
١٨	٤-١ الاستبيان The Questionnaire
٢٠	١-٤ عملية الانتقال من الاستبيان إلى برنامج SPSS
٢٢	٥-١ محرر البيانات في برنامج SPSS
٢٣	١-٥ عرض المتغيرات Variables View
٢٩	٢-٥ جدول البيانات Data View
٣٤	٦-١ تحويل البيانات Data Transformation
٣٥	١-٦ استخدام القيم لمتغير جديد باستخدام الأمر Compute
٣٧	١-٦-١ الآلة الحاسبة Calculator Pad
٤٥	٢-٦ التعبيرات الشرطية Conditional Expressions
٤٦	٣-٦ استخدام الرموز العلائقية في التعبيرات الشرطية
٤٨	٤-٦ استخدام الأمر Compute مع تعبير شرطي
٥١	١-٤-٦ تحديد مجموعة من الحالات Selecting Cases
٥٨	أسئلة وتمارين

الفصل الثاني

الإحصاءات الوصفية

Descriptive Statistics

٦٥	١-١ إنشاء التوزيع التكراري Producing Frequency Distributions
٦٩	١-١ الأمر Frequency
٧٧	٢-١ الأمر Descriptive

	Frequencies
٢١٨	١-١ أسلوب المتغيرات المتعددة ذات الفئتين Multiple Dichotomy Method
٢٢١	٢-١ أسلوب الفئات المتعددة Multiple Category Method
٢٢٥	٢-٢ أسلوب جداول التقاطع متعددة الاستجابة Multiple Response Crosstabs
٢٣١	ملخص الفصل السادس
	الفصل السابع الاختبارات اللامعلمية Non-Parametric Tests
٢٣٣	مقدمة
٢٣٤	١-٢ الاستدلال حول عدد k من العينات غير المستقلة لبيانات رتبية
٢٣٧	١-١ اختبار فريدمان لتحليل التباين للترتيب في اتجاهين The Friedman Two – Way Analysis of Variance by Ranks Test
٢٤٠	٢-١ الدلالة الإحصائية لاختبار فريدمان في حالة $k \geq 5$ أو $n \geq 10$
٢٤١	٣-١ الدلالة الإحصائية لاختبار فريدمان $2 \leq k \leq 4$ أو $2 \leq n < 10$
٢٤٩	٢-٢ الاستدلال حول k من العينات غير المستقلة (بيانات إسمية) K- Sample Inference: Related Samples
٢٤٩	١-٢ اختبار كوكران Q Cochran's Q Test
٢٥١	٢-٢ الدلالة الإحصائية لاختبار كوكران Q
٢٥٤	٣-٢ اختبار كروسكال واليز لتحليل التباين في اتجاه واحد Kruskal – Wallis One way analysis of variance
٢٥٤	١-٣ اختبار كروسكال – واليز $n_j > 5$
٢٥٥	٢-٣ الدلالة الإحصائية لأسلوب كروسكال – واليز $k > 3$ و $n_j > 5$
٢٥٥	٣-٣ الدلالة الإحصائية لأسلوب كروسكال – واليز $k \leq 5$ و $n_j \leq 5$
٢٦٥	٤-٢ توزيع كاي تربيع وتحليل التكرارات The Chi-Square Distribution and the Analysis of Frequencies
٢٧١	٥-٢ اختبارات جودة التوفيق Tests of goodness- of -fit
٢٨٥	٦-٢ الاستدلال اللامعلمي لبيانات عينتين غير مستقلتين لبيانات رتبية Two Related Samples Inference
٢٨٥	١-٦ اختبار الإشارة Sign Test
٢٨٥	١-١-٦ خطوات الاختبار
٢٨٧	٢-١-٦ الدلالة الإحصائية لاختبار الإشارة – عدد الثنائيات $N \leq 30$

المقدمة

لقد تطورت تقنيات الحاسوب في السنوات الأخيرة تطوراً سريعاً وتعددت مجالات استخدامه لتشمل كافة المجالات الاقتصادية والاجتماعية، ويعد استخدام الحاسوب في مجال البحوث الاقتصادية من أهم الاستخدامات التي أدت إلى تطورها بسبب السرعة والدقة التي يتصف بها الحاسوب وبشكل خاص عند إنجاز العمليات الرياضية المعقدة وتبويب البيانات الإحصائية ومن ثم استخراج أهم المقاييس وتطبيق أساليب الاستدلال الإحصائي في تحليل البيانات.

لقد وقع اختيارنا على أحد أهم البرامج الإحصائية الجاهزة المستخدمة للأغراض الإحصائية وتطبيقاته في مجال العلوم الاجتماعية والمالية والإدارية والمصرفية وهو برنامج SPSS، نظراً لشهرته الواسعة واستخدامه من غالبية المتخصصين في الاقتصاد والإدارة والمحاسبة والعلوم المالية والمصرفية^٢ على أنه نظام لإدارة قواعد البيانات، من حيث سهولة إدخال البيانات وحفظها واستعادتها وتحليلها، بالإضافة إلى قراءة البيانات من ملفات موجودة ببرامج أخرى مثل برنامج الجداول الإلكترونية المعروف بـ Microsoft Excel .

بمساعدة برنامج SPSS يكون باستطاعة الباحث الاستفادة من البيانات في الحصول على التقارير Reports ، الإحصاءات الوصفية Descriptive Statistics الجداول التقاطعية Cross-Tabulation ، الأشكال البيانية Charts ، مقارنة المتوسطات Compare Means ، النماذج الخطية العامة General Linear Model ، الارتباط Correlation ، الانحدار Regression ، النماذج الخطية اللوغاريتمية Loglinear ، التصنيف Classify ، إنقاص المتغيرات Reduction Variables ، المقاييس Scales ، الاختبارات اللامعلمية Non-Parametric Tests ،

جداول الحياة Survivals ، والاستجابات المتعددة Multiple Responses .

^١Chava Frankfort-Nachmais, Anna Leon Guerrero “ Social Statistics for a Diverse Society ” Pine Forge Press a Sage Publication Company , Second Edition 2000 .
R. Mark Sirkin “ Statistics for the Social Sciences “ Sage Publications. Inc 1999.

٣٣١	٤-٢ مقياس التشتت في المنتج و التوزيع الطبيعي
٣٣٢	٥-٢ التعامل مع الاختلافات التصنيعية
٣٣٢	٦-٢ الضبط الإحصائي للجودة Statistical Quality Control
٣٣٣	١-٦-٢ تقنيات الضبط الإحصائي للجودة Statistical Quality Control Tools
٣٣٣	٢-٦-٢ تقنية التوزيع التكراري
٣٣٤	٣-§ طرق مراقبة الجودة Methods of Quality Control
٣٣٤	١-٣ معاينة القبول <i>Acceptance Sampling</i>
٣٣٤	٢-٣ ضبط العملية إحصائياً (SPC) Statistical Process Control
٣٣٥	٤-§ خرائط ضبط الجودة
٣٣٦	١-٤ خريطة الوسط الحسابي $\bar{X}$ charts
٣٤٣	٢-٤ خريطة المدى R charts
٣٥١	٣-٤ خريطة ضبط الانحراف المعياري للعينات S charts
٣٥٥	٤-٤ خرائط نسبة الوحدات المعيبة $P - Chart$
٣٦٠	٥-٤ خريطة عدد العيوب في الوحدة $C - Chart$
٣٦٣	٥-§ خرائط الضبط في SPSS
٣٧٠	١-٥ مؤشرات المقدرة Capability indices
٣٧١	٢-٥ مؤشرات الأداء Performance indices
٣٧٧	أسئلة وتمارين
٣٧٩	دليل المصطلحات العلمية انكليزي - عربي
٣٩٥	المراجع العربية والمراجع الأجنبية

٢٨٧	٣-١-٦ الدلالة الإحصائية لاختبار الإشارة – عدد الثنائيات $N > 30$
٢٩١	٢-٦ اختبار ولكوكسن لإشارة الرتب Wilcoxon Signed- Rank
٣٠١	٧-§ الاستدلال اللامعلمي لبيانات عينتين غير مستقلتين لبيانات اسمية
٣٠١	١-٧ اختبار ماكنمار McNemar's Test
٣٠١	١-١-٧ الدلالة الإحصائية لاختبار ماكنمار $10 \leq y_{12} + y_{21} \leq 20$
٣٠٢	٢-١-٧ الدلالة الإحصائية لاختبار ماكنمار $y_{12} + y_{21} > 20$
٣٠٥	٢-٧ اختبار الوسيط

الفصل الثامن

كتابة الأوامر

Syntax Commands

٣١٥	١-§ ملف الأوامر
٣١٥	١-١ الطرق المساعدة في بناء الملفات
٣١٦	٢-§ كتابة الأوامر من صناديق الحوار Dialog boxes
٣٢٢	٣-§ نسخ الأوامر من Log في مخرجات برنامج SPSS
٣٢٤	٤-§ نسخ الأوامر من Journal File
٣٢٤	٥-§ تنفيذ الأوامر البرمجية Run Command Syntax

الفصل التاسع

الضبط الإحصائي للجودة

Statistical Quality Control

٣٢٨	١-§ مفهوم الضبط الإحصائي للجودة
٣٣٠	٢-§ تقنيات الضبط الإحصائي للجودة ' Statistical Quality Control Tools
٣٣٠	١-٢ ضبط الجودة للمنتج الصناعي
٣٣٠	٢-٢ الاختلافات التصنيعية للمنتجات
٣٣١	٣-٢ العلاقة بين الاختلافات التصنيعية و جودة المنتج

¹ د. محمد عيشوني أستاذ مساعد – قسم التقنية الميكانيكية – ٢٠٠٤ m_aichouni@yahoo.co.uk

لقد وضع هذا الكتاب ليكون مادة علمية لمقرر استخدامات الحاسوب في الاقتصاد لطلبة السنة الرابعة قسم العلوم المالية والمصرفية.

يتضمن الكتاب مقدمة وتسعة فصول دراسية خصص الفصل الأول للتعريف ببرنامج SPSS وكيفية تشغيله والتعامل معه، أنواع الملفات التي يشتمل عليها وكيفية تهيئتها، وآلية الانتقال من الاستبيان إلى البرنامج . خصص الفصل الثاني للإحصاءات الوصفية وكيفية إنشاء التوزيع التكراري، بناء مجال ثقة حول المتوسط، تشكيل الجدول التقاطعي. خصص الفصل الثالث لمقارنة المتوسطات وساطة اختبار T لعينة واحدة ، اختبار الفرق بين متوسطين، اختبار T للملاحظات المزدوجة وتحليل التباين. أما الفصل الرابع فكان بعنوان تحليل الارتباط والانحدار الخطي البسيط والمتعدد بوساطة برنامج SPSS. في الفصل الخامس تم عرض موضوع استكشاف البيانات Explore بوساطة مخطط Boxplots، ومخطط الجذع والورقة. وفي الفصل السادس تم عرض تحليل الاستجابات المتعددة -مميزين بين أسلوب التكرارات وأسلوب جداول التقاطع متعددة الاستجابة. خصص الفصل السابع لبعض الاختبارات اللامعلمية مثل اختبار فريدمان، اختبار كوكران Q، اختبار كروسكال واليز، توزيع كاي تربيع وتحليل التكرارات، اختبار الإشارة ، اختبار ولكوكسن، اختبار ماكنمار، وخصص الفصل الثامن لكتابة الأوامر والتعليمات في برنامج SPSS . وأخيراً كان الفصل التاسع بعنوان الضبط الإحصائي للجودة Statistical Quality Control.

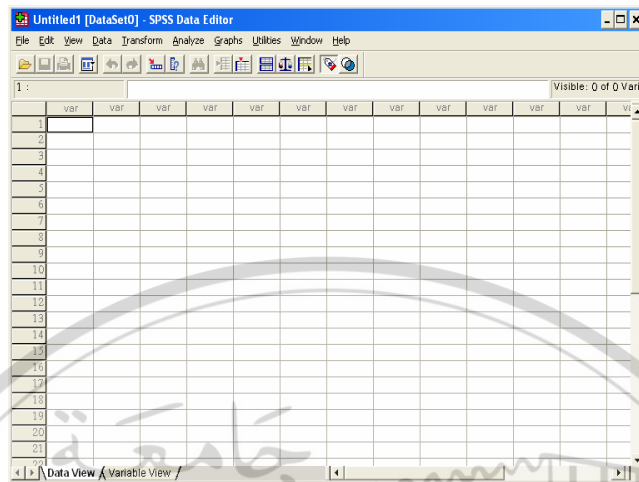
ونحن إذ نضع هذا الكتاب المتواضع بين أيدي أبنائنا الطلبة وزملائنا المهتمين بهذا المجال، نأمل أن نكون قد وفقنا في عرض موضوعاته عرضاً مبسطاً، وسنكون من الشاكرين لكل من يتقدم بأية ملاحظة تفيد في إغنائه للأخذ بها في الطبعة القادمة. وأخيراً لا يسعنا إلا أن نتقدم بجزيل الشكر لكل من أسهم في تقييم هذا الكتاب ولكل من ساعد على إخراجه بالشكل المطلوب.

والله ولي التوفيق

حلب في ١ / ١٢ / ٢٠٠٧ م

المؤلف

في الحالتين تظهر تلقائياً نافذة الحوار محرر البيانات Data Editor كما في الشكل
المعنونة Untitled1 [Dataset0] – SPSS Data Editor :



§ ٢- أنواع الملفات في برنامج SPSS:

هناك ثلاثة أنواع من الملفات في برنامج SPSS :

١- ملف البيانات DATA Files : تتشكل هذه الملفات باستخدام محرر البيانات Data Editor وهي تضم البيانات المستخدمة في التحليل الإحصائي، وعادة ما يكون لهذا النوع من الملفات الامتداد SAV .

٢- ملف المخرجات الإحصائية Output Files : وهي الملفات التي تضم مخرجات التحليل الإحصائي والمخططات البيانية، ويكون لهذا النوع من الملفات الامتداد SPO .

٣- ملف التعليمات Syntax : وهي الملفات التي تضم الإجراءات الإحصائية المخزنة على شكل أوامر برمجية، ويكون لهذا النوع من الملفات الامتداد SPS .

§ ٣- تهيئة الملفات في برنامج SPSS:

إن تهيئة الملفات في برنامج SPSS يتم بواسطة محرر البيانات Data Editor الذي يعتبر بمنزلة ورقة عمل Spread Sheet، هذه النافذة تفتح تلقائياً بمجرد تشغيل برنامج SPSS ، وتشتمل هذه النافذة على نوعين من العرض للبيانات:

١- عرض البيانات Data View : وهي عبارة عن ورقة عمل تضم صفوفاً وأعمدة حيث تمثل الأعمدة المتغيرات Variables وتمثل الصفوف الحالات Cases، إذ تمثل كل حالة منفصلة مستجوباً، والخلية تنتج من تقاطع المتغير مع الحالة.

الإلكترونية المعروف بـ Microsoft Excel .

بمساعدة برنامج SPSS يكون باستطاعة الباحث الاقتصادي الاستفادة من البيانات في الحصول على التقارير Reports ، الإحصاءات الوصفية Descriptive Statistics الجداول التقاطعية Cross-Tabulation ، الأشكال البيانية Charts ، مقارنة المتوسطات Compare Means ، النماذج الخطية العامة General Linear Model ، الارتباط Correlation ، الانحدار Regression ، النماذج الخطية اللوغاريتمية Loglinear ، التصنيف Classify ، إنقاص المتغيرات Reduction Variables ، المقاييس Scales ، الاختبارات اللامعلمية Non-Parametric Tests ، جداول الحياة Survivals ، والاستجابات المتعددة Multiple Responses .

ظهرت أول نسخة من هذا البرنامج عام ١٩٨٣ م ، ونتيجة لسرعة انتشاره وتداوله شرائح وقطاعات مختلفة من الباحثين ظهرت العديد من الإصدارات التي تعمل تحت بيئة النوافذ Windows ، والنسخة التي بين أيدينا هي SPSS 14.0 وليست أحدث الإصدارات وهي ثمّكن المستخدمين من إجراء عمليات إحصائية جداً متطورة ، كما أن ملفات البيانات تستوعب مئات المتغيرات Variables وآلاف المشاهدات أو الأشخاص (حالات) Cases .

تجدر الإشارة هنا بأن كافة المعالجات الإحصائية التي سترد لاحقاً باستخدام برنامج SPSS يمكن تنفيذها بسهولة بوساطة برنامج Microsoft EXCEL . سنحاول قبل الدخول في تفاصيل التطبيقات العملية أن نعرض بشكل موجز لأهمية الاستبيان وكيفية التعامل معه من حيث تحويل أسئلة الاستبيان إلى متغيرات ومن ثم تعريف هذه المتغيرات، وإدخال المتغيرات التي تم تعريفها والبيانات في ملف خاص ببرنامج SPSS ، وأخيراً كيفية التعامل مع هذه الملفات من حيث حفظها وفتحها وتعديلها وتجزئتها ودمجها .

١-٨ تشغيل برنامج SPSS

يمكن تشغيل برنامج SPSS بطريقتين :

١- بالنقر المزدوج Double Click بوساطة الطرف الأيسر من الماوس على أيقونة

برنامج SPSS في حال وجود اختصار لها على سطح المكتب.

٢- بالنقر على زر إبدأ Start الموجود أسفل الشاشة الرئيسية وحسب التسلسل التالي:

Start → All Programs → SPSS for Windows → SPSS 14.0 for Windows

. Evaluation Version

الفصل الأول

البرنامج الإحصائي SPSS

حزمة البرامج الإحصائية للعلوم الاجتماعية

Statistical Package for Sciences Socials (SPSS)

مقدمة:

لقد تطورت تقنيات الحاسوب في السنوات الأخيرة تطوراً سريعاً وتعددت مجالات استخدامه لتشمل كافة المجالات الاقتصادية والاجتماعية، ويعد استخدام الحاسوب في مجال البحوث الاقتصادية من أهم الاستخدامات التي أدت إلى تطورها بسبب السرعة والدقة التي يتصف بها الحاسوب وبشكل خاص عند إنجاز العمليات الرياضية المعقدة وتبويب البيانات الإحصائية ومن ثم استخراج أهم المقاييس وتطبيق أساليب الاستدلال الإحصائي في تحليل البيانات.

سنحاول استخدام أحد أهم البرامج الإحصائية الجاهزة المستخدمة للأغراض الإحصائية وتطبيقاته في مجال العلوم الاجتماعية والاقتصادية والإدارية وهو برنامج SPSS (حزمة البرامج الإحصائية للعلوم الاجتماعية Statistical Package for Sciences Socials).

لقد وقع اختيارنا على هذا البرنامج الإحصائي الجاهز لشهرته الواسعة واستخدامه من قبل غالبية المتخصصين في الإحصاء والاقتصاد والإدارة والمحاسبة، واعتماد أغلب المراجع الإحصائية العالمية المتخصصة في مجال العلوم الاقتصادية والاجتماعية¹ عليه، ومن مزايا هذا البرنامج إمكانية استخدامه كنظام لإدارة البيانات، أي التعامل معها على أنها قواعد بيانات من حيث سهولة إدخالها وحفظها واستعادتها وتحليلها، كما أن لدى برنامج SPSS إمكانية لاستخدام وقراءة البيانات من ملفات موجودة ببرامج أخرى أو موجودة بأنظمة قواعد البيانات الأخرى وفي مقدمتها برنامج الجداول

¹Chava Frankfort-Nachmais, Anna Leon Guerrero “ Social Statistics for a Diverse Society ” Pine Forge Press a Sage Publication Company , Second Edition 2000 .
R. Mark Sirkin “ Statistics for the Social Sciences “ Sage Publications. Inc 1999.

٢- عرض المتغيرات Variable View : وهي عبارة عن ورقة عمل تحتوي على وصف للمتغيرات، إذ تكون الصفوف هي المتغيرات ، والأعمدة تمثل وصفاً لكل متغير ، يمكن إضافة أو تغيير صفات أي متغير.

§-٤ الاستبيان^٢ The Questionnaire

إذا قرر الباحث جمع بيانات ميدانية سواء عن طريق الحصر الشامل أو العينات ، فإنه يحتاج إلى إعداد استمارة يضمنها جميع الأسئلة التي تمكنه من الحصول على البيانات التي يريدها بهدف التحقق من فرضيات البحث أو الإجابة عن أسئلة البحث ، وتسمى استمارة البحث أو استمارة جمع البيانات ، ورغم أن تصميم الاستمارة يختلف حسب طبيعة البحث إلا أن هناك قواعد وشروطاً عامة ينبغي مراعاتها عند تصميم الاستمارة، من أهمها :

- ١ . البساطة وعدم الغموض .
- ٢ . تحديد التعاريف ووحدات القياس .
- ٣ . الترتيب المنطقي للأسئلة وتقسيمها إلى مجموعات متجانسة .
- ٤ . صياغة الأسئلة بحيث يسهل تبويبها وتحليلها .
- ٥ . الابتعاد عن أنواع معينة من الأسئلة .

لنفترض أنك تقوم بدراسة أو إجراء بحث علمي حول موضوع يتعلق بالمعوقات السلوكية والتنظيمية والخدمات لكفاءة أعضاء هيئة التدريس في الجامعات السورية^٣، ولهذا السبب سوف تحتاج إلى إعداد استبيان يحتوي على مجموعة من الأسئلة المتعلقة بموضوع البحث ، ومن ثم توزيع هذا الاستبيان، الذي يشتمل على مجموعة من الأسئلة، على عينة من الأشخاص لهم علاقة مباشرة بمضمون الأسئلة الواردة في هذا الاستبيان ، وفي العادة يفضل أن يحتوي الاستبيان على متغيرات تصنيفية تتطلب الإجابة عن التساؤل المطلوب ، والاختيار من عدة إجابات ، بالإضافة إلى أسئلة تكون إجاباتها رقمية محددة ، وسوف نتعامل مع أجزاء محددة من هذا الاستبيان، ولمن يود معرفة التفاصيل عن الاستبيان يمكنه العودة إلى كتاب مبادئ الإحصاء لطلاب السنة الثانية بكلية الآداب قسم علم الاجتماع .

ونورد فيما يلي بعض الأسئلة المختارة من الاستبيان بهدف توضيح كيفية تعريف المتغيرات وترميزها وإدخال البيانات في برنامج SPSS .

^٢ د . محمد كبيه " مبادئ الإحصاء " السنة الثانية ، قسم علم الاجتماع ، منشورات جامعة حلب ٢٠٠٢ م .
^٣ هذا البحث تم إجراؤه فعلاً من قبل فريق بحث بكلية الاقتصاد قوامه الأستاذ الدكتور زكي حنوش ، الدكتور محمد خشروم ، الدكتور محمد كبيه ووافق مجلس جامعة حلب على إجرائه وتمويله .

- الوكيل الإداري : ()
- جهات أخرى خارج الجامعة ()
- رئيس القسم : ()
- وكيل الجامعة : ()
- لا ألبأ إلى أحد ()

وفي مثل هذه الحالة لا بد من تعريف ثمانية متغيرات بشكل يأخذ كل متغير قيمتين، الأولى صفر في حالة عدم اختياره، والثانية ١ في حالة وقوع الاختيار عليه.

ج - أسئلة مفتوحة جزئياً : وهذا النوع من الأسئلة يسمح باختيار إجابة محددة موجودة أو كتابة إجابة أخرى غير موجودة ، ومن الأمثلة على هذا السؤال من الاستبيان السابق نذكر :
- وسيلة النقل المستخدمة إلى العمل :

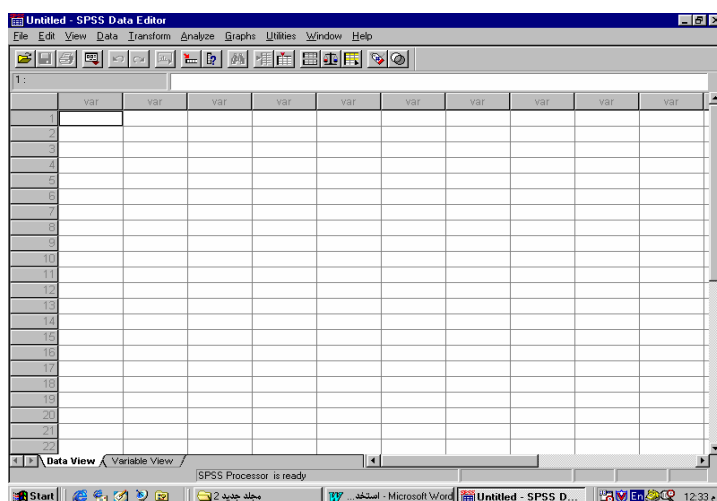
□ خاصة □ عامة □ غير ذلك

نكتفي بهذه الحالة بتعريف متغير واحد ليمثل السؤال لأن الإجابة هو اختيار واحد فقط ولهذا يمكن إعطاء المتغير القيمة ١ في حال اختيار الإجابة الأولى ، والرقم ٢ في حال اختيار الإجابة الثانية ، والرقم ٣ في حال اختيار الإجابة الثالثة مهما كانت وسيلة النقل المستخدمة.

§-٥ محرر البيانات في برنامج SPSS :

بمجرد الدخول إلى برنامج SPSS نجد أنفسنا أمام محرر البيانات غير المعنون Untitled SPSS Data Editor . الذي يعد ملفاً جديداً جاهزاً لاستقبال المتغيرات والقيم التي تأخذها هذه المتغيرات، ومن أسفل الصفحة يمكن الاختيار من بديلين : الأول وهو جدول تعريف المتغيرات ، والثاني جدول البيانات وذلك بالنقر على الخيار Variables View .

٥-١ عرض المتغيرات Variables View : بمجرد النقر على هذا الخيار الموجود أسفل صفحة محرر البيانات نحصل على جدول فارغ يشتمل على 10 أعمدة معنونة .



- الوكيل العلمي : ()
- الوكيل الإداري : ()
- جهات أخرى خارج الجامعة ()
- رئيس القسم : ()
- وكيل الجامعة : ()
- لا ألبأ إلى أحد ()

٤-١ عملية الانتقال من الاستبيان إلى برنامج SPSS :

للانتقال من الاستبيان إلى البرنامج لا بد من التعرف إلى كيفية تعريف المتغيرات وإدخالها ثم تفريغ البيانات، آخذين في الاعتبار أن الأشخاص أو المشاهدات التي تقوم بإملاء الاستبيان تعتبر حسب مفهوم البرنامج بمنزلة حالات Cases وكل سؤال يمكن اعتباره بمنزلة متغير في حين تعتبر إجابات الأشخاص بمنزلة قيم للمتغيرات ، ونوعية المتغيرات وعددها والقيم التي تأخذها تعتمد على نوع الأسئلة التي يحتويها الاستبيان وهي على أنواع :

أ - أسئلة بسيطة تسمح بإجابة واحدة فقط : وهذا النوع من الأسئلة يتم التعامل معه بسهولة ، ومن الأمثلة على هذا السؤال من الاستبيان السابق نذكر :

- هل تمارس عملاً خارج الجامعة ؟

□ نعم □ لا

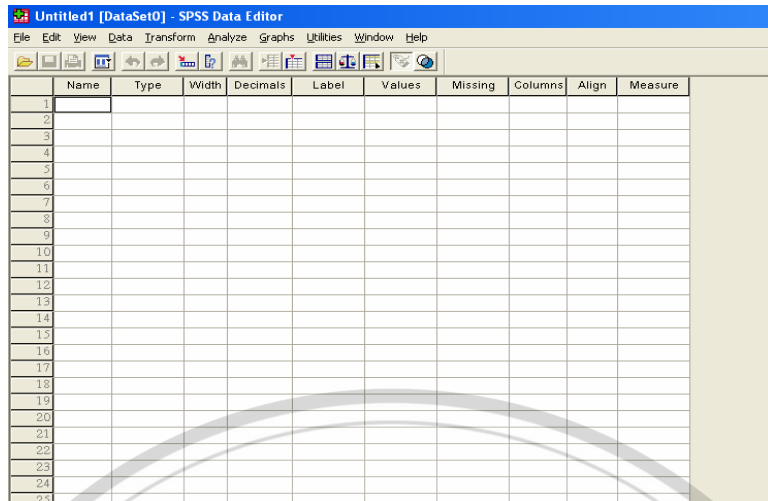
وفي هذه الحالة نعرف متغيراً واحداً وليكن Exjob وبما أن هناك احتمالاً لإجابتين يمكننا إعطاء الرمز ١ ليمثل الإجابة نعم Yes ، والرمز ٠ ليمثل الإجابة لا No ، ويمكن اختيار رموز حرفية مثل الرمز Y ليمثل الإجابة نعم Yes ، والرمز N ليمثل الإجابة لا No ، إلا أننا نفضل التعامل مع الرموز الرقمية كما يفضلها برنامج SPSS.

ب - أسئلة تسمح بإجابات متعددة : وهذا النوع من الأسئلة يعطي المستجوب الفرصة للاختيار وقد يكون بإمكان الشخص اختيار أكثر من إجابة، ومن الأمثلة على هذا السؤال من الاستبيان السابق نذكر :

- في حال اعترضتك مشكلة ما في الجامعة ، هل تلجأ إلى :

- العميد : ()
- رئيس الجامعة : ()
- الوكيل العلمي : ()

- الجنس : ☐ ذكر ☐ أنثى
- العمر : (سنة)
- الجامعة :
- سنوات الخدمة (تاريخ التعيين) :
- المرتبة العلمية :
- ☐ معيد عائد ☐ مدرس ☐ أستاذ مساعد ☐ أستاذ
- الراتب الشهري :
- أقل من 5000 ل.س. 5000 إلى أقل من 10000 ل.س. 10000 فأكثر
- الدخل الشهري غير راتب الجامعة إن وجد () ل.س.
- وسيلة النقل المستخدمة إلى العمل :
خاصة ☐ عامة ☐ غير ذلك
- أوزع ساعات عملي اليومي على الأنشطة التالية :
- المواصلات () ساعة
- التحضير () ساعة
- التدريس () ساعة
- غير ذلك ، اذكرها .
- هل تمارس عملاً خارج الجامعة .
☐ نعم ☐ لا
- أشعر بنقص في بعض المهارات التقنية (كاستخدام الحاسوب مثلاً) .
معارض ☐ محايد ☐ موافق ☐ موافق جداً ☐
- إن المركز الاجتماعي لأستاذ الجامعة جيد مقارنة بالشرائح الاجتماعية الأخرى .
معارض ☐ محايد ☐ موافق ☐ موافق جداً ☐
- يتيح لي عملي فرصاً دائمة للتجديد وللابتكار وللتميز :
معارض ☐ محايد ☐ موافق ☐ موافق جداً ☐
- في حال اعترضتك مشكلة ما في الجامعة، هل تلجأ إلى:
- العميد : ()
- رئيس الجامعة : ()



العمود الأول بمسمى Name (أسم المتغير) :

وفيه يتم تحويل كل سؤال في الاستبيان إلى متغير يعطى اسماً معيناً مختصراً يعبر عن السؤال مشتقاً من السؤال نفسه وبما لا يزيد عن ثمانية حروف ، ويفضل أن يخصص خانة أو متغير يمثل رقم السؤال في الاستبيان أي يعطى لكل شخص شارك في الإجابة رقماً يميزه عن غيره ويفيد في اكتشاف مواقع الأخطاء في حال حدوثها . لقد أعطينا السؤال الأول وهو جنس المستجيب اسماً يميزه كمتغير وهو Sex ، والنسبة للسؤال الثاني تم تعريف متغير باسم Age ليمثل عمر المستجوب ، والسؤال الثالث عرفنا متغير بمسمى unversty ليمثل الجامعة التي ينتمي إليها عضو هيئة التدريس وهكذا ... الخ.

لكتابة أسماء المتغيرات في برنامج SPSS يجب إتباع الخطوات التالية:

- يجب ألا يزيد طول الاسم عن ثمانية رموز Characters .
- يجب أن يبدأ اسم المتغير بحرف أما بقية الرموز فقد تكون أحرفاً أو أرقاماً أو نقطة.
- لا يجوز أن ينتهي اسم المتغير بنقطة.
- يجب ألا يتضمن اسم المتغير فراغات وبعض الرموز الخاصة مثل * , ' , ? , .
- لا يميز برنامج SPSS بين الحروف الصغيرة والحروف الكبيرة فالأسماء SEX و sex تعتبر متماثلة إذ أن البرنامج لا يقبل سوى الحروف الصغيرة لأسماء المتغيرات.

العمود الثاني بمسمى Type (نوع المتغير) :

وفيه يتم وصف نوع المتغير Variable Type حيث يفترض برنامج SPSS أن جميع المتغيرات رقمية ، ولتغيير نوع المتغير نضغط مرتين متتاليتين على الكلمة الافتراضية

وفيه يتم تحديد عدد الخانات العشرية إذ يفترض برنامج SPSS أن جميع المتغيرات رقمية مفترضاً خانتين عشريتين لكل متغير ، وبإمكاننا تغيير عدد الخانات العشرية حسب حاجتنا .

العمود الخامس بمسمى Label (وصف المتغير) :

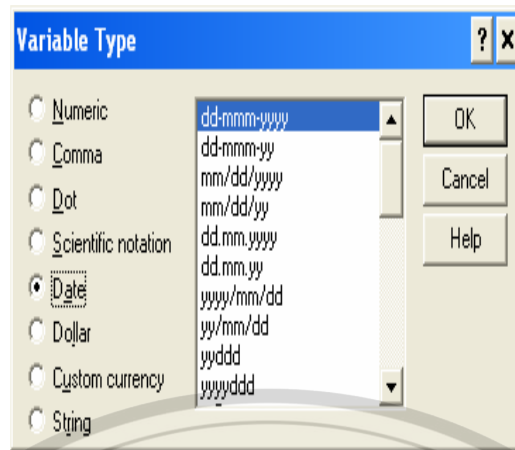
وفيه يتم وصف كل متغير بجملة موجزة تعبر عن السؤال وتوضحه حيث تظهر أوصاف المتغير وقيمته في النتائج الإحصائية وعلى الرسوم البيانية مما يساعد على تفسير النتائج بسهولة فمثلاً يمكن وصف رقم السؤال على الشكل number of respondent ، وعمر المستجوب age of respondent وكذلك جنس المستجوب sex of respondent وهكذا....

العمود السادس بمسمى Values (قيم ورموز المتغير)

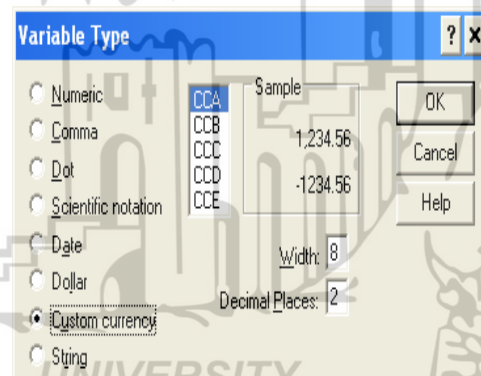
وفيه يتم تعيين رموز وأرقام ، فمثلاً بالنسبة للسؤال الثاني نضع الرمز F لتدل على الأنثى Female والرمز M ليدل على الذكر Male . وبالنسبة للسؤال : أشعر بنقص في بعض المهارات التقنية (كاستخدام الحاسوب مثلاً) نضع Disagree لتعبر عن معارض ، و Neuter لتعبر عن محايد ، ونضع agree لتعبر عن موافق ، و Strongly agree لتعبر عن موافق جداً .

Value	Value Label
1	Disagree
2	agree
3	agree
4	Strongly agree

العمود السابع بمسمى Missing (القيم المفقودة)



- **Dollar** : متغير رقمي يراد عرض قيمه بحيث يشتمل على إشارة الدولار \$ مع فاصلة كل ثلاثة مواقع للأرقام ، وبمجرد اختياره يقوم البرنامج بعرض الخيارات المتعددة ليتم اختيار ما يناسبك مثال \$80000.
- **Custom Currency** : متغير رقمي يراد عرض قيمه إذ يشتمل على عملة دولة معينة ثم تعريف مواصفاتها حسب الطلب وذلك باختيار الأمر Preference من القائمة Edit.



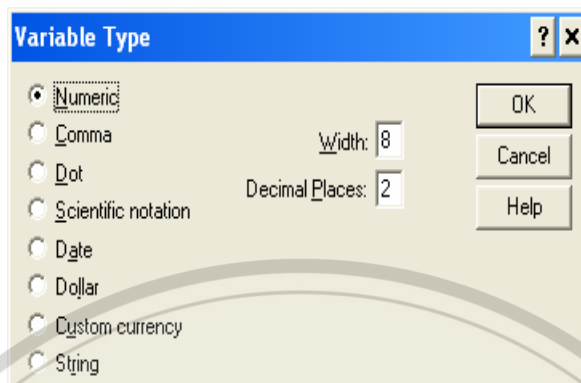
- **String** : متغير حرفي يراد عرض قيمه بحيث يشتمل على أحرف أو رموز أخرى ، وهنا من الضروري التمييز بين الأحرف الكبيرة والأحرف الصغيرة حيث لكل منهما معنى مختلف بالنسبة للبرنامج .

العمود الثالث بمسمى Width (عرض المتغير) :

وفيه يتم تحديد عرض المتغير Width حيث يفترض برنامج SPSS أن جميع المتغيرات رقمية مفترضاً 8 خانات لكل متغير وبإمكاننا تغيير العرض حسب حاجتنا أخذين بالاعتبار تخصيص كل حرف أو رقم أو فاصلة أو نقطة أو إشارة موجبة أو سالبة.

العمود الرابع بمسمى Decimal (عرض المتغير) :

Numeric فيظهر صندوق حوار تحت مسمى Variable Type كما هو واضح في صندوق الحوار التالي:

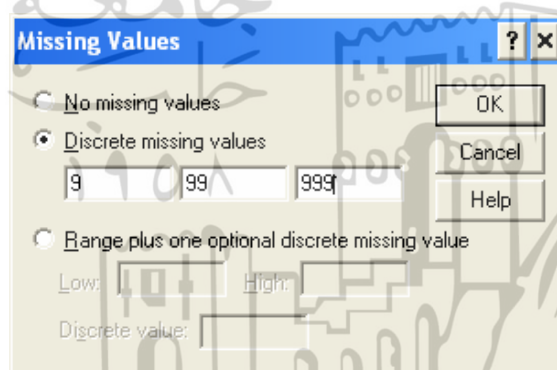


إن صندوق الحوار يحتوي على عدة أنواع من المتغيرات الممكنة كما يلي :

- **Numeric** : متغير رقمي يمكنك تحديد عرضه (عدد خاناته، وعدد الأرقام العشرية التي يتضمنها) ، وهو النوع الافتراضي للمتغيرات في ورقة عرض البيانات Data View.
- **Comma** : متغير رقمي يراد عرض قيمه بحيث يشتمل على فاصلة لكل ثلاثة مواقع مع نقطة تفصل الخانات العشرية مثال الرقم 8,000.00 .
- **Dot** : متغير رقمي يراد عرض قيمه بحيث يشتمل على نقطة لكل ثلاثة مواقع مع فاصلة تفصل الخانات العشرية مثال الرقم 8.000,00 .
- **Scientific notation** : متغير رقمي يراد عرض قيمه على شكل تعبير أسّي إذ يستخدم الحرفين E أو D ليقوم مقام الرقم 10 فمثلاً يمكن أن نكتب الرقم 80000 على الشكل 8D+4 أو 8D4 أو 8E+4 أو 8E4 .
- **Date** : متغير رقمي يراد عرض قيمه على شكل تاريخ أو وقت وبمجرد اختياره يقوم البرنامج بعرض الخيارات المتعددة لكتابة التاريخ أو الوقت ليتم اختيار ما يناسبك فمثلاً إذا تم اختيار الشكل (mm/dd/yy) حيث تدل mm على الشهر وتدل dd على اليوم كما تدل yy على السنة ، كما هو موضح في صندوق الحوار التالي:

وفيه يتم إعلام البرنامج بهذه القيم المفقودة والتي تحل محل الإجابات غير المتوفرة بسبب عدم قيام بعضهم على الإجابة على أنواع معينة من الأسئلة ، مع العلم أن البرنامج يفترض عدم وجود قيم مفقودة Non وفي حالة عدم إعلامه يقوم بوضع نقطة بدلاً من الإجابة غير المتاحة وتسمى بهذه الحالة بـ القيم المفقودة ، أما إذا قمنا بتعيين رمز لهذه القيم فتسمى بـ القيم المستخدمة المفقودة .

فمثلاً عند عدم إجابة أحد الأشخاص على العمر فإنه بإمكانك وضع الصفر أو أي رقم آخر وليكن الرقم 9 مكان القيم المفقودة وبهذه الحالة يتوجب عليك إعلام البرنامج بذلك. لتعريف القيم المفقودة لمتغير ما، نقر زر الخلية في عمود missing لذلك المتغير يظهر صندوق حوار بمسمى القيم المفقودة Missing Values كما يلي :



يمكن أن نختار قيمة واحدة أو عدة قيم مفقودة كما يظهر في صندوق الحوار، بالإضافة إلى وجود إمكانية لاختيار مجال من القيم مضافاً له قيمة مفقودة معينة يحددها الباحث.

العمود الثامن والتاسع بمسمى Columns & Align :

وفيه يتم إعلام البرنامج بعرض العمود وموقع البيانات داخل العمود يميناً أو يساراً أو في منتصفه ويتعلق هذان العمودان بالنواحي الجمالية للجدول .

العمود العاشر بمسمى Measure (وحدة القياس) :

وفيه يتم إعلام البرنامج بوحدة القياس التي يتم بموجبها قياس المتغير وهناك ثلاثة أنواع من المتغيرات:

- نسبية أو فئوية Scale : يستعمل للبيانات العددية القابلة للقياس الكمي وهو المقياس المستخدم غالباً لقياس متغيرات الطول والوزن ودرجات الحرارة.
- اسمية Nominal : يستعمل لقياس المتغيرات الاسمية وهي متغيرات لها عدد من الفئات دون أفضلية لإحداها على الأخرى ، لا يمكن ترتيبها تصاعدياً أو تنازلياً ولكن يمكن تحديد الفروق بينها بدقة كبيرة.

Column Width: 6 Alignment: Center

Print Format: F8.2

Write Format: F8.2

SEX sex of respondent جنس المستجوب

2

Measurement Level: Nominal

Column Width: 5 Alignment: Center

Print Format: A8

Write Format: A8

Value Label

F Female

M Male

AGE age of respondent

عمر المستجوب

3

Measurement Level: Scale

Column Width: 7 Alignment: Center

Print Format: F8.2

Write Format: F8.2

Missing Values: 9.00

UNVERSTY University of respondent الجامعة التي ينتمي لها المستجوب

4

Measurement Level: Nominal

Column Width: 5 Alignment: Center

ونورد فيما يلي جدول البيانات الخاص بمثالنا بعد أن تم تفريغ كافة الاستبيانات .
يمكننا عند اختيار الأمر File Info من القائمة Utilities الموجود في شريط القوائم أن
نعرض في نافذة المخرجات SPSS Viewer - Output معلومات تتعلق بالمتغيرات

	qustnob	sex	age	unvest	workdate	scntde	salary	exsalary	transpr	extjob	skills	socialst	jobinnov
1	1.00	F	35.00	A	Weekness of skills	00	12000.00	2000.0	1.00	1.00	4.00	1.00	2.00
2	2.00	F	40.00	A	02.05.90	2.00	11000.00	.	2.00	.00	3.00	1.00	3.00
3	3.00	M	44.00	B	12.07.80	2.00	12500.00	.	3.00	.00	2.00	2.00	3.00
4	4.00	M	52.00	T	07.09.92	1.00	15000.00	4000.0	1.00	1.00	3.00	2.00	1.00
5	5.00	M	38.00	D	12.10.86	2.00	10500.00	.	3.00	.00	3.00	4.00	1.00
6	6.00	M	39.00	D	12.03.87	2.00	9000.00	6000.0	2.00	1.00	1.00	4.00	2.00
7	7.00	M	29.00	D	03.07.87	4.00	4900.00	1500.0	1.00	1.00	1.00	4.00	2.00
8	8.00	M	28.00	A	04.12.85	4.00	5600.00	.	1.00	.00	2.00	4.00	4.00
9	9.00	F	46.00	A	02.05.90	2.00	11000.00	.	3.00	.00	2.00	3.00	4.00
10	10.00	F	37.00	A	12.07.80	2.00	8700.00	3400.0	3.00	1.00	4.00	2.00	4.00
11	11.00	M	29.00	A	07.09.92	4.00	7000.00	.	2.00	.00	4.00	3.00	2.00
12	12.00	M	32.00	B	12.10.86	3.00	9800.00	5000.0	2.00	1.00	4.00	3.00	3.00
13	13.00	F	34.00	T	12.03.87	3.00	10000.00	.	3.00	.00	3.00	1.00	3.00
14	14.00	M	36.00	D	03.07.87	2.00	9900.00	5700.0	3.00	1.00	2.00	1.00	1.00
15	15.00	M	29.00	D	12.04.98	4.00	6000.00	.	1.00	.00	3.00	2.00	4.00
16	16.00	M	50.00	D	18.04.80	1.00	14800.00	1200.0	1.00	1.00	3.00	2.00	4.00
17	17.00	M	47.00	A	06.09.91	1.00	15200.00	15000	3.00	1.00	1.00	4.00	3.00
18	18.00	M	39.00	A	02.05.75	2.00	13300.00	2000.0	1.00	1.00	1.00	4.00	2.00
19	19.00	M	44.00	B	12.03.87	1.00	12000.00	.	1.00	.00	2.00	4.00	4.00
20	20.00	F	41.00	T	03.07.87	2.00	9760.00	.	2.00	.00	2.00	2.00	1.00
21	21.00	M	43.00	D	12.04.98	2.00	11000.00	12000	3.00	1.00	4.00	3.00	3.00
22	22.00	F	50.00	A	18.04.80	1.00	16000.00	.	1.00	.00	4.00	3.00	4.00

الموجودة في ملف البيانات الذي نعمل عليه حالياً ، وتوضح هذه المعلومات كل ما يتعلق
بالمتغيرات أو الأسئلة التي تم تفريغها وذلك على النحو التالي:
قائمة بالمتغيرات الموجودة في ملف العمل

List of variables on the working file

اسم المتغير وموقعه المتسلسل

Name

Position

QUSTNOB number of respondent

رقم المستجوب

1

Measurement Level: Scale

- ترتيبية Ordinal : يستعمل لقياس المتغيرات الترتيبية إذ أن هذا المتغير له عدد محدد من الفئات يمكن ترتيبها تصاعدياً أو تنازلياً ولكن لا يمكن تحديد الفروق بينها بدقة كبيرة مثال تقدير طالب في امتحان (ممتاز ، جيد جداً ، جيد ، متوسط ، مقبول ، ضعيف) ومن الممكن أن يكون المتغير رمزياً أو عددياً . غالباً لا نحتاج لتحديد وحدة القياس إذ يقوم برنامج SPSS بتعيين وحدة القياس تلقائياً بمجرد قيامنا بتعيين رموز تستخدم على أنها أوصاف لقيم المتغير . ونورد فيما يلي جدول تعريف المتغيرات Variables View لمثالنا السابق

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	responda	Numeric	12	2	رقم المستجوب	None	None	8	Center	Scale
2	sex	String	8	0		{F, Female}	None	8	Center	Nominal
3	age	Numeric	8	2		None	None	8	Center	Scale
4	university	String	8	0		{A, Aleppo}	None	8	Center	Nominal
5	workdate	Date	8	0		None	None	8	Center	Scale
6	Scient_d	Numeric	8	2		{1.00, Profe}	None	8	Center	Scale
7	Salary	Numeric	8	2		{1.00, Less}	None	8	Center	Scale
8	ex_Salar	Numeric	8	2		None	None	9	Center	Scale
9	transport	Numeric	8	2		{1.00, Privat}	None	8	Center	Scale
10	exter_job	Numeric	8	2		{0.00, No}	None	8	Center	Scale
11	skills	Numeric	8	2		{1.00, Disag}	None	8	Center	Scale
12	social_ce	Numeric	8	2		{1.00, Disag}	None	8	Center	Scale
13	job_inno	Numeric	8	2		{1.00, Disag}	None	8	Center	Scale
14	Averg	Numeric	8	2		None	None	10	Right	Scale
15	New_Sal	Numeric	8	2		None	None	12	Right	Scale
16	aveg	Numeric	8	2		None	None	10	Right	Scale
17										
18										
19										
20										
21										
22										

٢-٥ جدول البيانات Data View :

بمجرد النقر على هذا الخيار والموجود أسفل صفحة محرر البيانات نحصل على جدول فارغ من البيانات يشتمل على أعمدة معنونة بأسماء المتغيرات التي سبق لنا تعريفها في جدول تعريف المتغيرات ، يكون هذا الجدول جاهزاً لاستقبال البيانات من الاستبيانات التي قمت بجمعها بعد أن تمت تعبئتها من قبل المستجوبين .

لإدخال البيانات في برنامج SPSS هناك طريقتان:

الأولى تتلخص بتفريغ كل استمارة دفعة واحدة ، والاستمارة تعبر عن شخص محدد وبالتالي ستشغل حالة محددة Case .

والثانية تتلخص بتفريغ البيانات الخاصة بسؤال واحد لكافة المستجوبين وعند الانتهاء من هذا السؤال يتم الانتقال إلى السؤال الذي يليه .

Print Format: A8

Write Format: A8

Value	Label
A	Aleppo
B	Baath
D	Damascus
T	Teshrin

WORKDATE Starting date of working

تاريخ بدء العمل

5

Measurement Level: Scale

Column Width: 8 Alignment: Center

Print Format: EDATE10

Write Format: EDATE10

SCNTDEGR Scientific degree of respondent

المرتبة العلمية

6

Measurement Level: Scale

Column Width: 5 Alignment: Center

Print Format: F8.2

Write Format: F8.2

Value Label

1.00	professor
2.00	associate professor
3.00	teacher
4.00	assistant

SALARY Salary of respondent

الدخل للمستجوب

7

Print Format: F8.2

Write Format: F8.2

Value Label

1.00 Disagree

2.00 Neuter

3.00 Agree

4.00 Strongly agree

JOBINNOV My work destines to me opportunity

13

عملي يتيح لي فرص الإنجاز والابتكار

Measurement Level: Scale

Column Width: 6 Alignment: Center

Print Format: F8.2

Write Format: F8.2

Value Label

1.00 Disagree

2.00 Neuter

3.00 Agree

4.00 Strongly agree

♦ بالنسبة لكل متغير تم عرض اسم المتغير ووصف للمتغير .

♦ مستوى القياس : نسبي ، اسمي ، ترتيبي .

♦ عرض العمود : موقع البيانات في العمود: يسار ، يمين ، في الوسط .

♦ نوع المتغير وعدد الخانات المخصصة له بالإضافة لعدد الخانات العشرية، فمثلاً **F8.2**

تعني أن المتغير رقمي ، وأن سعته ثمانية رموز وخانتان عشريتان .

A8 تعني أن المتغير حرفي وأن سعة هذا المتغير هي ثمانية رموز أو 8 خانات.

EXTJOB Agree with working out of university

10

هل توافق على العمل خارج الجامعة

Measurement Level: Scale

Column Width: 5 Alignment: Center

Print Format: F8.2

Write Format: F8.2

Value Label

.00 NO

1.00 YES

SKILLS Weakness of skills

11

هل تعاني من نقص في المهارات

Measurement Level: Scale

Column Width: 5 Alignment : Center

Print Format: F8.2

Write Format: F8.2

Value Label

1.00 Disagree

2.00 Neuter

3.00 Agree

4.00 Strongly agree

SOCIALST socialist Center

المركز الاجتماعي لأستاذ الجامعة

12

Measurement Level: Scale

Column Width: 6 Alignment: Center

Measurement Level: Scale

Column Width: 8 Alignment: Center

Print Format: F8.2

Write Format: F8.2

Value Label

1.00 less than 5000

2.00 5000 to 10000

3.00 10000 or More

EXSALARY Salary of respondent

الدخل الخارجي للمستجوب

8

Measurement Level: Scale

Column Width: 6 Alignment: Center

Print Format: F8.2

Write Format: F8.2

TRANSPRT Utilization of transportatin

وسيلة النقل المستخدمة

9

Measurement Level: Scale

Column Width: 5 Alignment: Center

Print Format: F8.2

Write Format: F8.2

Value Label

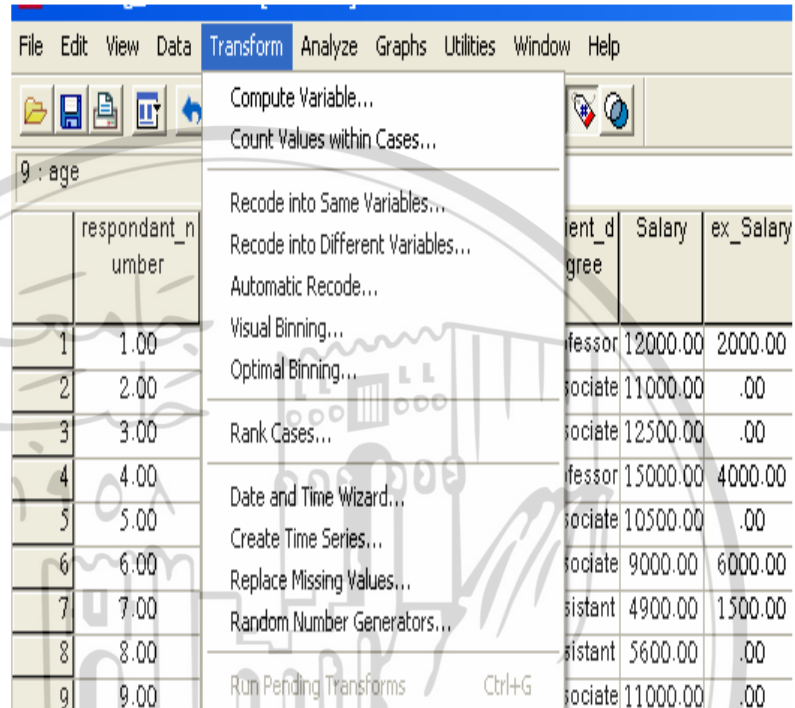
1.00 private

2.00 public

3.00 others

٦-٨ تحويل البيانات Data Transformation

يمكن إنجاز تحويل على البيانات الأصلية Raw Data تتراوح بين تجميع البيانات في فئات معينة إلى تكوين متغيرات جديدة باستخدام معادلات و تعبيرات شرطية، تضم القائمة Transformation الأوامر التالية:



٦-١ استخدام القيم لمتغير جديد باستخدام الأمر Compute

باستخدام برنامج SPSS يمكننا تكوين متغيرات جديدة يتم حسابها بالاستناد لقيم المتغيرات الأصلية الموجودة في ملف البيانات المستخدم. لنفترض أننا نرغب إضافة الدخل الذي يحصل عليه عضو هيئة التدريس من خارج الجامعة EX_Salary إلى الدخل الذي يحصل عليه من الجامعة Salary بعد أن تم زيادة الدخل الأصلي بنسبة 10% مضافاً إليها علاوة شهرية بمبلغ 1000 ل.س، لهذا السبب نرغب في إنشاء متغير جديد في ملف البيانات المسمى Banking_Student بإسم New_Salary يشتمل على القيم الجديدة للدخل بعد إضافة الراتب الذي يحصل عليه من خارج الجامعة ، وزيادة العلاوة ونسبة الزيادة.

لحساب الدخل الجديد يمكننا تكوين المعادلة التالية التي تحسب قيم المتغير الجديد New_Salary بالاعتماد على قيم المتغيرات الأصلية وذلك على النحو التالي:

$$\text{New\_Salary} = \text{Salary} + \text{Ex\_Salary} + 0.10 * \text{Salary} + 1000$$

- الرموز المنطقية:

تستخدم الرموز المنطقية في ربط علاقيتين أو عكس ناتج لتعبير شرطي جوابه إما نعم أو لا .

- التتابع Functions:

تحتوي قائمة التتابع في برنامج SPSS على أكثر من ٧٠ تابع، من أهم هذه التتابع:

- ١- التتابع الحسابية Arithmetic Functions .
- ٢- التتابع الإحصائية Statistical Functions .
- ٣- التتابع المنطقية Logical Functions .
- ٤- توابع الوقت والتاريخ Date and Time Functions .
- ٥- توابع القيم المفقودة Missing Value Functions .
- ٦- التتابع النصية String Functions .
- ٧- توابع التوزيع Distribution Functions .
- ٨- توابع المتغيرات العشوائية Random Variable Functions .
- ٩- توابع الدرجات Scoring Functions .

ولمعرفة المزيد عن التتابع في برنامج SPSS يمكن الضغط على الزر مساعدة Help الموجود في صندوق الحوار Compute Variable .

مثال:

لنفترض أن الباحث يرغب في حساب الوسط الحسابي لعدد من المتغيرات موجودة أصلاً في ملف البيانات، فمثلاً لحساب المتوسط الحسابي للمتغيرات التالية⁴ :

- هل تمارس عملاً خارج الجامعة EXte_Job .
- هل تمتلك بعض المهارات Skills .
- المركز الاجتماعي للمستجوب Social_Cent .
- عملي يتيح لي فرص الابتكار Job_Inno .

⁴ إن المتغيرات التي يشتمل عليها الملف لا تساعد كثيراً في حساب بعض الدوال الإحصائية كالمتوسط الحسابي ، إلا أننا سنقوم على إيراد عملية الحساب لبيان كيفية حسابها فقط مدركين أن لا معنى لها معنوياً.

٦-١-١ الآلة الحاسبة Calculator Pad

يحتوي صندوق الحوار Compute Variable على آلة حاسبة تضم أرقاماً ، ورموزاً حسابية، ورموزاً علائقية، ورموزاً منطقية، وذلك كما هو موضح في الجدول التالي:

الرموز الحسابية		الرموز العلائقية		الرموز المنطقية	
الرمز	العملية	الرمز	العملية	الرمز	العملية
+	الجمع	<	أصغر من	&	و (AND) العلاقات يجب أن تكون جميعها صحيحة
-	الطرح	>	أكبر من		
*	الضرب	<=	أصغر أو يساوي		أو (OR) واحدة من العلاقات على الأقل يجب أن تكون صحيحة
/	القسمة	>=	أكبر أو يساوي		
**	رفع الى قوة	=	يساوي	~	تفيد النفي (NOT) تعكس القيمة الناتجة التعبير جوابه نعم/لا
()	ترتيب العمليات	~=	لا يساوي		

- الرموز الحسابية:

إن أي تعبير حسابي يجب أن يحتوي على أحد الرموز الحسابية، وتتبعي مراعاة تسلسل تنفيذ العمليات الحسابية في أي تعبير حسابي، وذلك وفقاً لما يسمى بقاعدة الأولوية، ويكون التسلسل على النحو التالي:

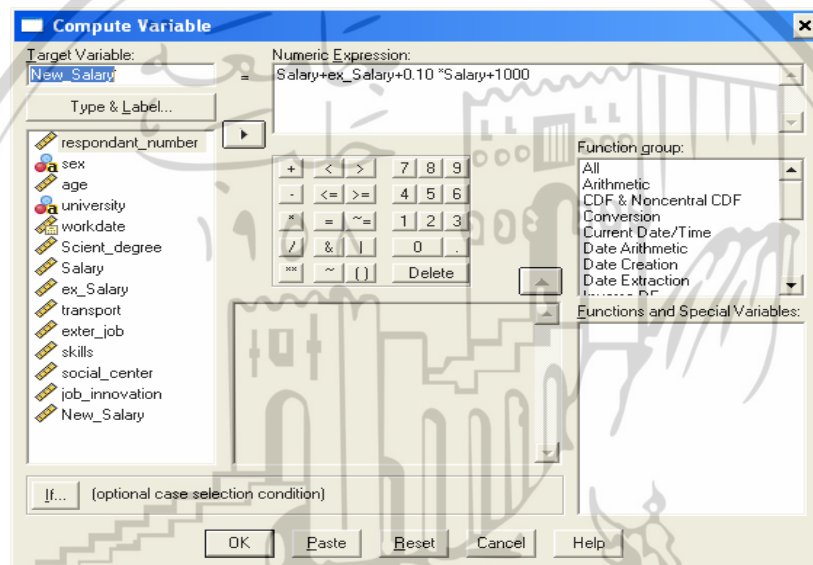
- ١- الاقتران.
- ٢- الرفع إلى قوة (الأس).
- ٣- الضرب أو القسمة.
- ٤- الجمع والطرح.

إن استعمال الأقواس في التعبيرات الحسابية يبطل قاعدة الأولوية، إذ أن الكمية الداخلة بين الأقواس تحسب قبل أي حدود خارج الأقواس، أي أن جميع العمليات داخل الأقواس تجري معاً حتى يتم إغلاق الأقواس.

- الرموز العلائقية:

العلاقة هي تعبير منطقي يقارن بين قيمتين باستخدام الرموز العلائقية، وتستخدم هذه الرموز عادة في التعبيرات الشرطية.

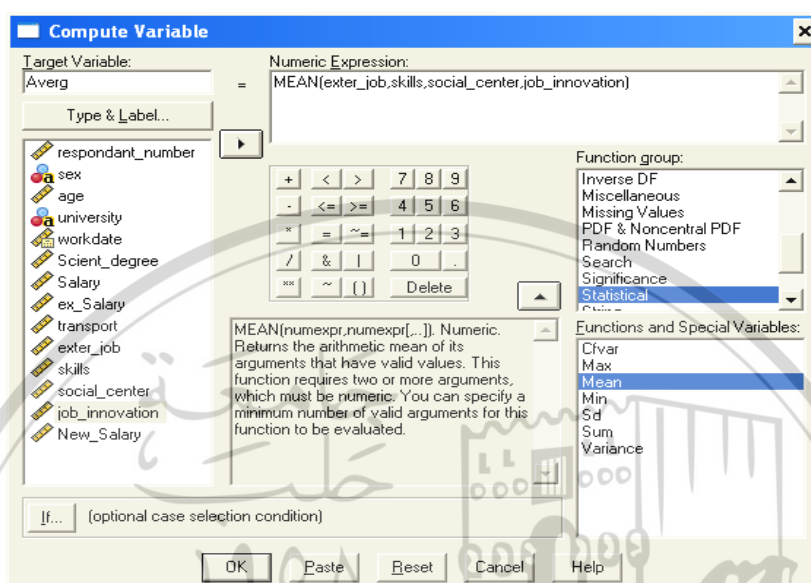
- باستخدام برنامج SPSS يمكننا حساب قيم المتغير الجديد بإتباع الخطوات التالية:
- من القائمة Transform نختار Compute نحصل على صندوق الحوار التالي تحت مسمى Compute Variable.
 - في المستطيل المعنون Target Variable (المتغير الهدف) نكتب بوساطة لوحة المفاتيح اسم المتغير الجديد المطلوب حساب قيمه، وهو في مثالنا New_Salary .
 - في المستطيل المعنون Numeric Expression (الصيغة الرقمية الحسابية) نحاول كتابة الصيغة بوساطة الآلة الحاسبة الظاهرة في نافذة الحوار مع استخدام لوحة المفاتيح إذا تطلب الأمر ذلك.



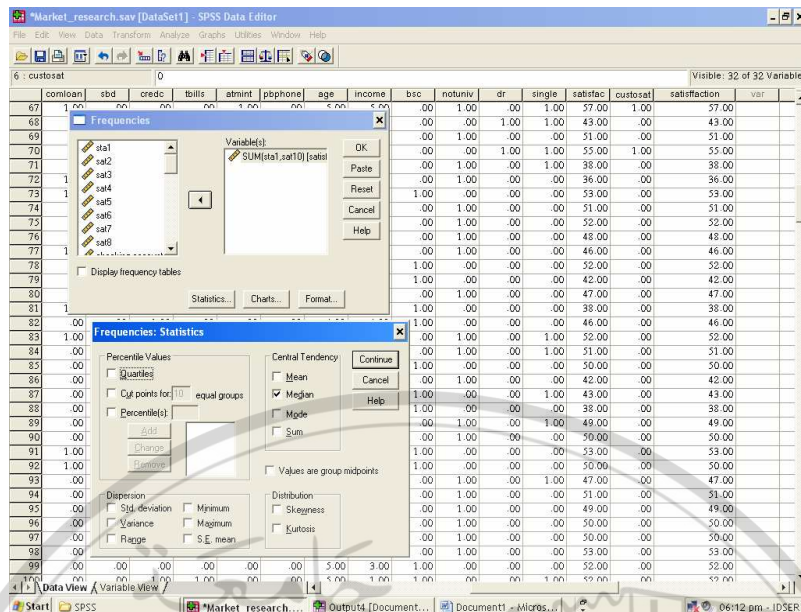
- بالضغط على الزر OK نحصل على المتغير الجديد New_Salary في نهاية ملف البيانات ويحتوي القيم الجديدة المحسوبة بناء على الصيغة الرياضية السابقة.

respondent_number	sex	age	university	workdate	Scient_degree	Salary	ex_Salary	transport	exter_job	skills	social_center	job_innovation	New_Salary
1	1.00	F	35.00	A	05/12/00	1.00	12000.00	2000.00	1.00	1.00	4.00	1.00	16200.00
2	2.00	F	40.00	A	02/07/90	2.00	11000.00	.00	2.00	.00	3.00	1.00	13100.00
3	3.00	M	44.00	B	12/07/80	2.00	12500.00	.00	3.00	.00	2.00	3.00	14750.00
4	4.00	M	52.00	T	07/09/92	1.00	15000.00	4000.00	1.00	1.00	3.00	2.00	21500.00
5	5.00	M	38.00	D	12/10/86	2.00	10500.00	.00	3.00	.00	3.00	4.00	12550.00
6	6.00	M	39.00	D	12/03/87	2.00	9000.00	6000.00	2.00	1.00	1.00	4.00	16900.00
7	7.00	M	29.00	D	03/07/87	4.00	4900.00	1500.00	1.00	1.00	1.00	4.00	7890.00
8	8.00	M	28.00	A	04/12/85	4.00	5600.00	.00	1.00	.00	2.00	4.00	7160.00
9	9.00	F	46.00	A	02/05/90	2.00	11000.00	.00	3.00	.00	2.00	3.00	13100.00
10	10.00	F	37.00	A	12/07/80	2.00	8700.00	3400.00	3.00	1.00	4.00	2.00	13970.00
11	11.00	M	29.00	A	07/09/92	4.00	7000.00	.00	2.00	.00	4.00	3.00	8700.00
12	12.00	M	32.00	B	12/10/86	3.00	9800.00	5000.00	2.00	1.00	4.00	3.00	16780.00
13	13.00	F	34.00	T	12/03/87	3.00	10000.00	.00	3.00	.00	3.00	1.00	12000.00
14	14.00	M	36.00	D	03/07/87	2.00	9900.00	5700.00	3.00	1.00	2.00	1.00	17590.00
15	15.00	M	29.00	D	12/04/98	4.00	6000.00	.00	1.00	.00	3.00	2.00	7600.00
16	16.00	M	50.00	D	04/13/80	1.00	14800.00	1200.00	1.00	1.00	3.00	2.00	18480.00
17	17.00	M	47.00	A	09/06/91	1.00	15200.00	1500.00	3.00	1.00	1.00	4.00	19220.00
18	18.00	M	39.00	A	05/02/75	2.00	13900.00	2000.00	1.00	1.00	1.00	4.00	17630.00
19	19.00	M	44.00	B	03/12/87	1.00	12000.00	.00	1.00	.00	2.00	4.00	14200.00

إن برنامج SPSS يساعدنا في حساب الوسط الحسابي المطلوب عن طريق اختيار الدالة المطلوبة وهي MEAN من نافذة الحوار Function group الموجودة في صندوق الحوار .Compute Variable



- في المستطيل المعنون Target Variable (المتغير الهدف) نكتب بواسطة لوحة المفاتيح اسم المتغير الجديد المطلوب حساب قيمه، وهو في مثالنا Averg.
- في المستطيل المعنون Function group نختار Statistical.
- في المستطيل المعنون Functions and Special Variables نختار MEAN ونضغط عليها مرتين فتظهر بالشكل التالي MEAN(?,?).
- في المستطيل المعنون Numeric Expression (الصيغة الرقمية الحسابية) نحاول كتابة الصيغة بواسطة الآلة الحاسبة الظاهرة في نافذة الحوار مع استخدام لوحة المفاتيح إذا تطلب الأمر ذلك، كما يوضحه صندوق الحوار أعلاه.
- نورد فيما يلي بعض التتابع الاحصائية في برنامج SPSS .
- CFVAR(numexpr,numexpr[,...]). تحسب معامل الاختلاف لقيم رقمية.
- LAG. LAG(variable). تحسب قيم المتغير في الحالة السابقة بملف البيانات سواء كانت البيانات رقمية أو نصية.
- MAX(value,value[,...]). تعطي أكبر قيمة سواء كانت البيانات رقمية أو نصية.
- MEAN(numexpr,numexpr[,...]) تحسب الوسط الحسابي للقيم الرقمية لوسطاء الدالة.



لترميز القيم التي هي أصغر من قيمة الوسيط بالرمز ٠ والقيم التي هي أكبر من قيمة الوسيط بالرمز 1 ، نتبع الخطوات التالية:

Transform Recorde into Different Variables

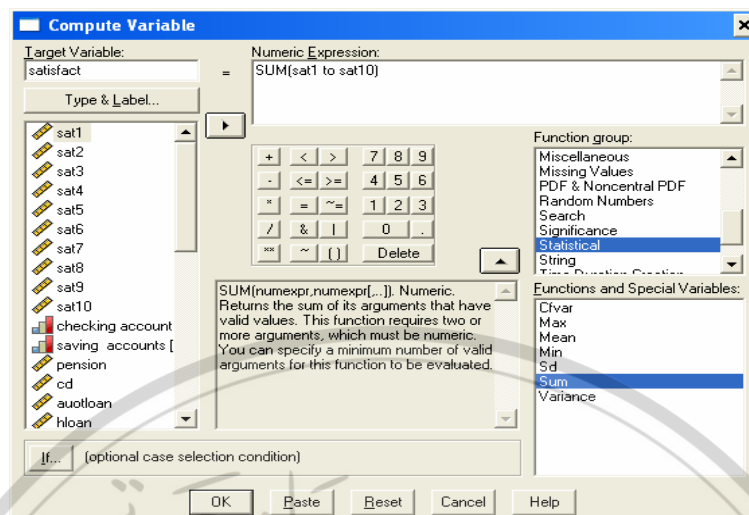
ننقل المتغير Satisfaction إلى نافذة الحوار:

numeric Variable → Output Variable

نكتب اسم المتغير الناتج Customer نضغط على الزر Old and New Values نحصل على صندوق حوار جديد :

Recorde into Different Variables: Old and New Values

Sum(sat1 to sat10)



بالضغط على الزر OK نحصل على المتغير الجديد Satisfact في نهاية ملف البيانات ويحتوي القيم الجديدة المحسوبة بناء على الصيغة الرياضية السابقة. للحصول على رضى المستهلكين عن الخدمات المقدمة من البنك، سيتم تقسيمهم إلى راضين وغير راضين وذلك بعد الحصول على قيمة الوسيط Median على أنه مؤشر على رضى المستهلك، وذلك على النحو التالي:

Analysis Descriptive Statistics Frequencies

ننقل المتغير Satisfaction إلى النافذة Variables وبالنقر على الزر Statistics نحصل على صندوق حوار بعنوان:

Frequencies: statistics

نفعل الزر الوسيط Median ثم continue نعود إلى صندوق الحوار الأصلي ، وبالضغط على الزر Ok نحصل على قيمة الوسيط التي تساوي 53 .

Statistics

SUM(sta1,sat10)

N	Valid	349
	Missing	1
Median		53.0000

- $\text{MIN}(\text{value}, \text{value}[\dots])$. تعطي أصغر قيمة سواء كانت البيانات رقمية أو نصية
- $\text{SD}(\text{numexpr}, \text{numexpr}[\dots])$. تحسب الانحراف المعياري للقيم الرقمية لوسطاء الدالة.

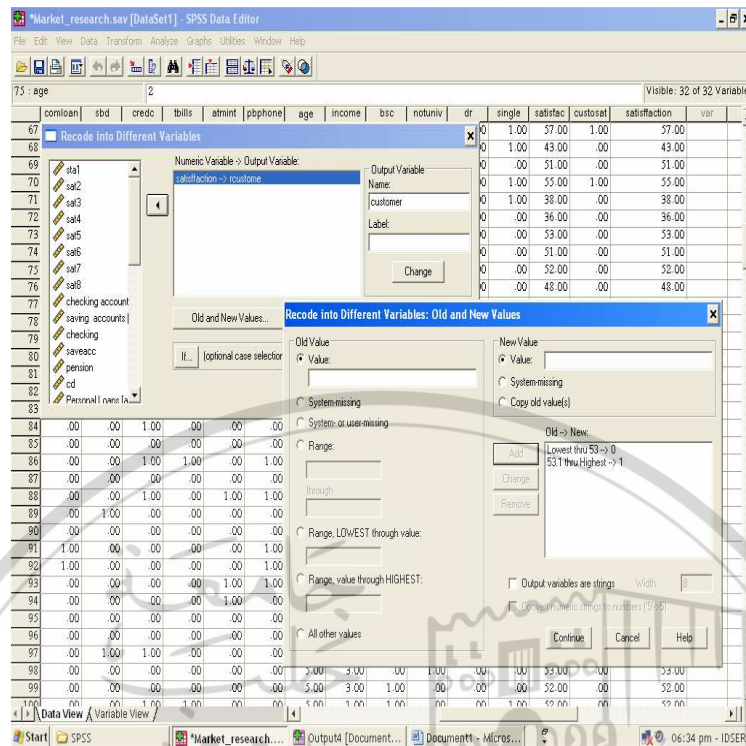
- $\text{VARIANCE}(\text{numexpr}, \text{numexpr}[\dots])$. تحسب التباين للقيم الرقمية لوسطاء الدالة.
- $\text{SUM}(\text{numexpr}, \text{numexpr}[\dots])$. تحسب المجموع للقيم الرقمية لوسطاء الدالة.

في كثير من الحالات يجد الباحث أنه بحاجة لإجراء عمليات مشابهة للعملية السابقة لتكوين ما يسمى بالمجالات Domains، فالمجال ببساطة هو عبارة عن متغير يتم الحصول عليه بإيجاد المجموع الكلي SUM لقيم عدة متغيرات موجودة في ملف البيانات وتشارك في صفة معينة، فمثلاً لو كان لدينا المتغيرات:

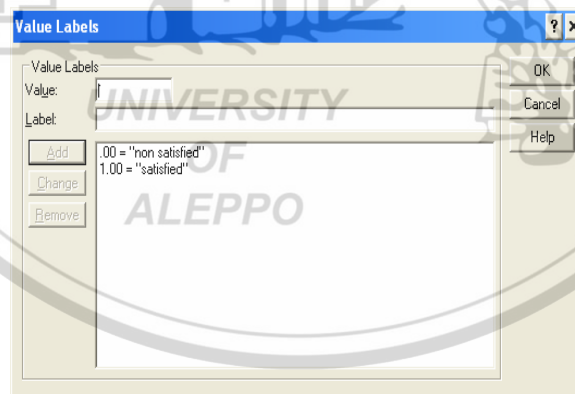
Sat1 , Sat2 , Sat3 , ... , Sat10

وأردنا تكوين مجالاً باسم Satisfact يشتمل على جميع تلك المتغيرات، لعمل ذلك نتبع الخطوات التالية:

- من القائمة Transform نختار Compute نحصل على صندوق الحوار التالي تحت مسمى Compute Variable.
- في المستطيل المعنون Target Variable نكتب بوساطة لوحة المفاتيح اسم المتغير الجديد المطلوب حساب قيمه، وهو في مثالنا Satisfact .
- في المستطيل المعنون Numeric Expression نحاول كتابة الصيغة بوساطة الآلة الحاسبة الظاهرة في نافذة الحوار مع استخدام لوحة المفاتيح إذا تطلب الأمر ذلك، بشكل تظهر على الصورة التالية.



في الطرف الأيسر وفي النافذة Range, Lowest through values نكتب قيمة الوسيط وهي 53. وفي الطرف الأيمن في النافذة Value نكتب الرقم 0. في الطرف الأيمن وفي النافذة Range, values through Highest نكتب قيمة الوسيط وهي 53.1. وفي الطرف الأيمن في النافذة Value نكتب الرقم 1.



المثال	الرمز العلائقي المستخدم	الحالات التي سيتم اختيارها
Salary<=12000	<=	الحالات التي دخلها الشهري أصغر أو تساوي 12000
Age>36	>	الحالات التي عمرها أكبر من 36
Sex='M'	=	الحالات التي جنسها ذكر
University='B'	=	الحالات التي تنتمي لجامعة البعث

يمكن ربط التعبيرات الشرطية باستخدام الرموز المنطقية ، حيث يمكن ربط تعبيرين شرطين أو أكثر باستخدام الرموز المنطقية (&) و (|) فمثلاً في الصيغ التالية نجد :

age >= 35 & sex='F'

يتم اختيار الحالات التي تحقق شرط العمر و شرط الجنس معاً.

age >= 35 | sex="F"

يتم اختيار الحالات التي تحقق إما شرط العمر أو شرط الجنس .

age >= 35 | Salary<15000

يتم اختيار الحالات التي تحقق إما شرط العمر أو شرط الدخل .

Salary >= 15000 & sex='M'

يتم اختيار الحالات التي تحقق شرط الدخل و شرط الجنس معاً .

- إن قيم المتغيرات الحرفية String يجب أن توضع بين:
- علامتي اختصار apostrophes كما في الحالة: sex='M'
- أو بين علامتي تنصيص quotation كما في الحالة: sex="F".
- إن نوعية الحروف لقيمة المتغير في الشرط يجب أن تكون نفس نوعية الحروف لقيمة المتغير الموجود في ملف البيانات المستخدم، لأن SPSS يفرق بين الأحرف الكبيرة والأحرف الصغيرة.
- إذا كانت قيمة المتغير الحرفي تحتوي على علامة اختصار (') يجب وضع تلك القيمة بين علامتي تنصيص (" ").
- المتغيرات الموجودة داخل صيغ التوابع يجب أن تفصل باستخدام فواصل (,) ويمكن وضع فراغ بين المتغيرات إذا أراد المستخدم ذلك.

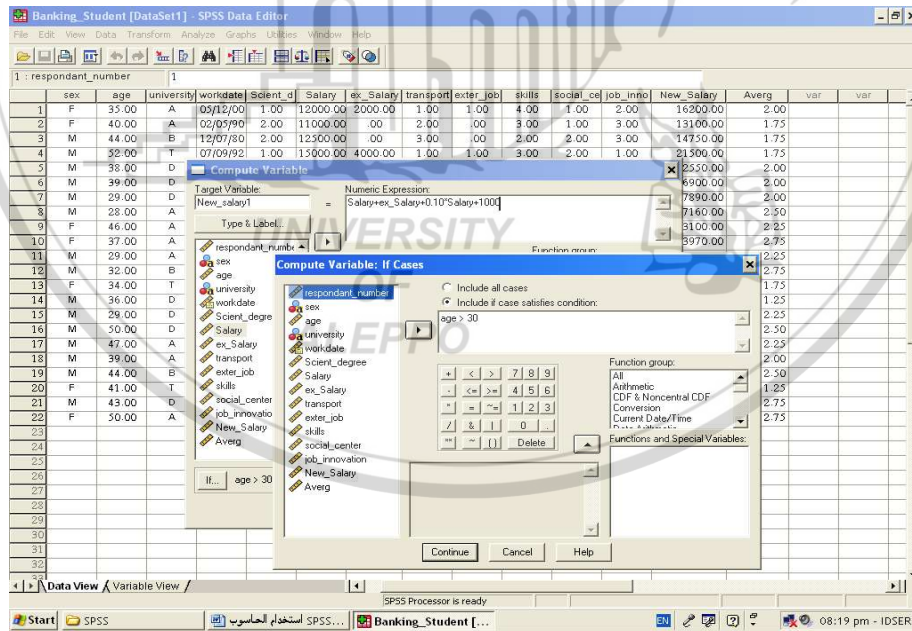
customer

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	non satisfied	182	52.0	52.1	52.1
	satisfied	167	47.7	47.9	100.0
	Total	349	99.7	100.0	
Missing	System	1	.3		
Total		350	100.0		

٢-٦ التعبيرات الشرطية Conditional Expressions

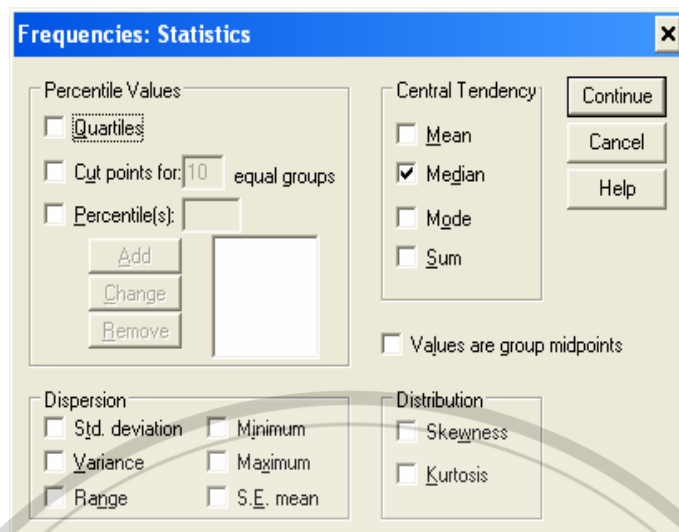
تستخدم التعبيرات الشرطية لإجراء تعديلات على مجموعة محددة من الحالات التي تشتمل عليها ملف البيانات، فمثلاً إذا كانت رغبة الباحث أو المسؤول زيادة الدخل السنوية أو الشهرية للموظفين الذين تزيد أعمارهم عن ٣٠ عاماً فقط، بهذه الحالة سيتم زيادة دخل الموظفين الذين يحققون الشرط أي زيادة دخل الحالات التي تحقق الشرط التالي: $Age > 30$.

يقوم التعبير الشرطي باختبار قيم كافة الحالات في ملف البيانات سواء كانت القيمة صحيحة أو خاطئة أو مفقودة، فإذا كانت نتيجة التعبير الشرطي صحيحة فإن الزيادة تطبق على الحالة، وإذا كانت النتيجة خاطئة أو أن القيمة مفقودة فلا تطبق الزيادة على تلك الحالة.



٣-٦ استخدام الرموز العلائقية في التعبيرات الشرطية:

تحتوي معظم التعبيرات الشرطية على رمز علائقي واحد على الأقل كما هو موضح بالأمثلة التالية:



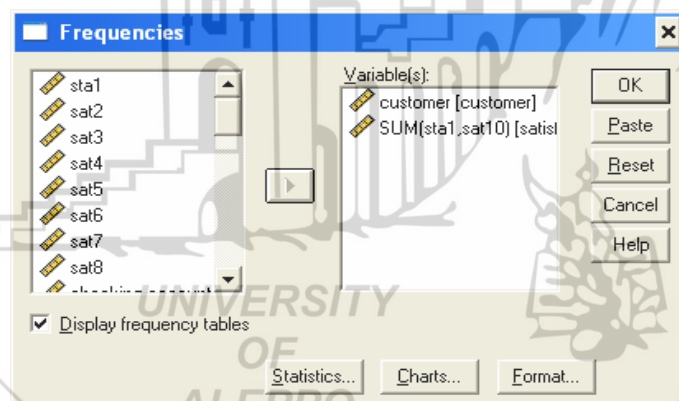
نضغط Continue نعود إلى صندوق الحوار الأصلي.

وللحصول على جول التوزيع التكراري للمتغيرين

Customer , Satisfaction(Sum(sat1 to sat2)

نضغط Frequencies → Descriptive Statistics → Analysis يظهر

صندوق الحوار التالي بعنوان Frequencies .



نضغط OK نحصل على النتائج التالية:

Statistics

		customer	SUM(sta1, sat10)
N	Valid	349	349
	Missing	1	1

- المتغيرات الخاصة بالتتابع يجب أن توضع بين قوسين () ، ويمكن وضع فراغ بين اسم المتغير وبين الأقواس إذا أراد المستخدم ذلك.
- كل علاقة في تعبير شرطي يجب أن تكون مكتملة ، كما في الأمثلة التالية:

$(age \geq 35 \ \& \ sex='F')$

هذه الصيغة صحيحة.

$(age \geq 35 \ \& \ age < 50)$

هذه الصيغة صحيحة.

$(age \geq 40 \ \& \ < 50)$

هذه الصيغة غير صحيحة.

$(age \geq 35 \ | \ Salary < 15000)$

هذه الصيغة صحيحة.

$(age \geq 35 \ | \ Salary \geq 15000)$

هذه الصيغة صحيحة.

$(age \geq 45 \ | \ = 15000)$

هذه الصيغة غير صحيحة.

- النقطة (.) هي فقط الفاصل العشري الصحيح في التعبيرات.

٤-٦ استخدام الأمر Compute مع تعبير شرطي

لنفترض أن الباحث يرغب في زيادة دخل أعضاء الهيئة التدريسية الذكور (بنفس الطريقة التي زیدت بها سابقاً) وممن تزيد أعمارهم عن ٥٠ سنة وذلك في ملف البيانات Banking_Student وبهذه الحالة سيحتاج إلى تكوين متغير جديد بمسمى New_Salary2 والذي يتم حساب قيمه بالاعتماد على قيم المتغيرات الأصلية حسب المعادلة التالية:

$$New\_Salary2 = Salary + Ex\_Salary + 0.10 * Salary + 1000$$

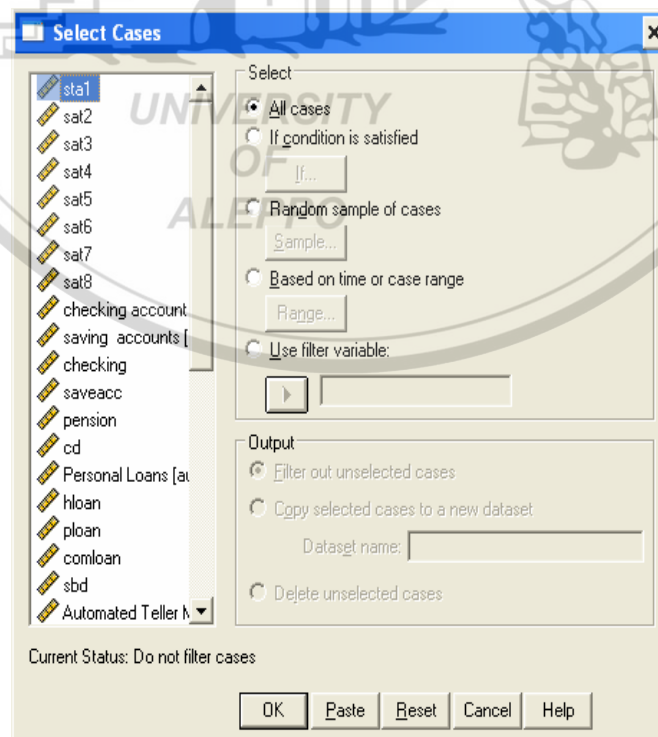
بالإضافة إلى استخدام التعبير الشرطي المركب التالي:

$(age > 50 \ \& \ sex='M')$

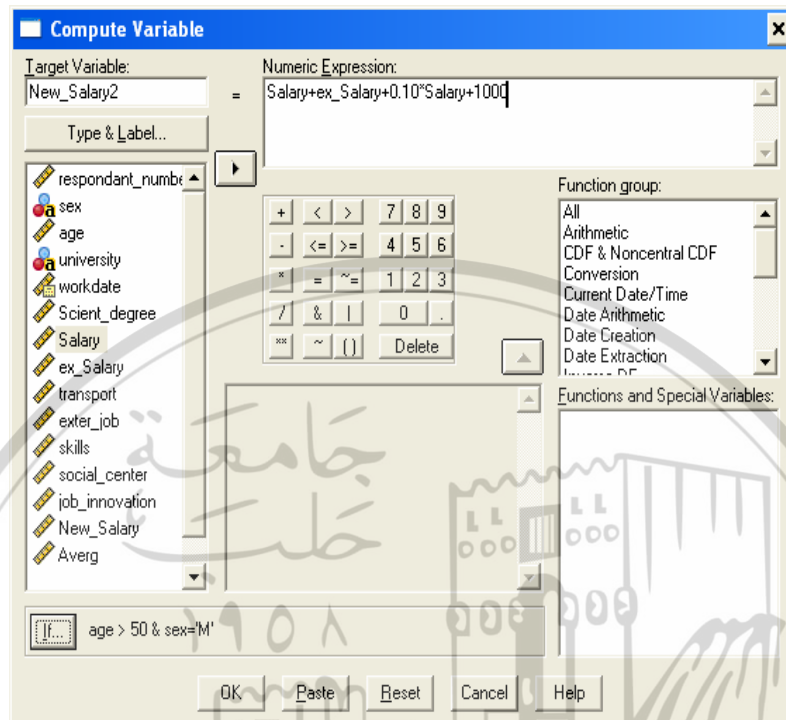
لتنفيذ ما تقدم باستخدام برنامج SPSS نفتح ملف البيانات Banking_Student ثم نتبع الخطوات التالية:

يتمتع برنامج SPSS بميزة حصر التحليل على مجموعة جزئية من الحالات وذلك اعتماداً على معايير تشتمل على متغيرات وتعبيرات شرطية مركبة، كما يتيح البرنامج إمكانية اختيار عينة عشوائية من الحالات.

- لتحديد مجموعة جزئية من الحالات، نختار من القائمة Data الأمر select cases نحصل على صندوق الحوار المسمى select cases الذي يشتمل على :
 - اختيار كل الحالات All cases، وبهذه الحالة يتم استخدام التحليل بشكل اعتيادي.
 - اختيار الحالات التي تحقق الشرط .
 - اختيار عينة عشوائية من الحالات .
 - اختيار حالات ضمن مجال محدد أو بالاعتماد على تاريخ معين.
 - استخدام تصفية للمتغيرات.
- وفي جميع الحالات يمكن اختيار واحدة من الخيارات المتوضعة تحت عنوان النتائج:
- عدم إدخال الحالات غير المحددة في التحليل ولكنها تبقى موجودة في ملف البيانات، ويمكن استخدامها لاحقاً في التحليل Filter out unselected cases .
- نسخ البيانات التي لم يتم إدخالها في التحليل وحفظها في ملف جديد copy selected cases to a new dataset
- حذف الحالات غير المحددة في التحليل من ملف البيانات.



بالنقر على الزر continue نحصل على صندوق حوار بعنوان : Compute Variable، مع ظهور عبارة الشرط بجانب الزر if أسفل صندوق الحوار.



- نضغط موافق نحصل على متغير جديد في نهاية الملف يحمل الاسم New_Salary2 ويشتمل على قيم جديدة حسبت بالاستناد إلى الصيغة الرياضية والتعبير الشرطي المركب، مع ملاحظة ظهور نقطة مقابل كل حالة لا تحقق التعبير الشرطي المركب، وذلك كما هو واضح في الشاشة التالية:

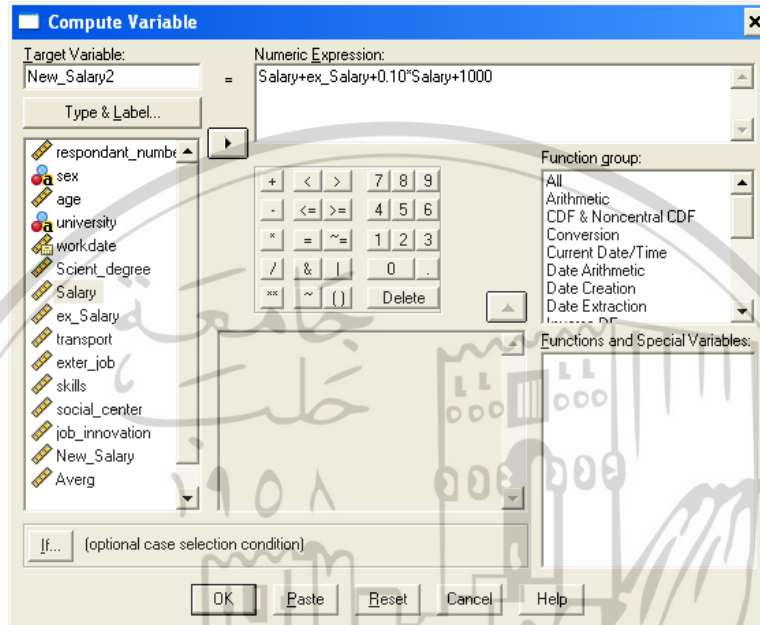
	workdate	Scient_d	Salary	ex_Salary	transport	exter_job	skills	social_ce	job_inno	New_Salary	Averg	New_Salary2	var
1	05/12/00	1.00	12000.00	2000.00	1.00	1.00	4.00	1.00	2.00	16200.00	2.00	-	-
2	02/05/90	2.00	11000.00	.00	2.00	.00	3.00	1.00	3.00	13100.00	1.75	-	-
3	12/07/80	2.00	12500.00	.00	3.00	.00	2.00	2.00	3.00	14750.00	1.75	-	-
4	07/09/92	1.00	15000.00	4000.00	1.00	1.00	3.00	2.00	1.00	21500.00	1.75	21500.00	-
5	12/10/86	2.00	10500.00	.00	3.00	.00	3.00	4.00	1.00	12550.00	2.00	-	-
6	12/03/87	2.00	9000.00	6000.00	2.00	1.00	1.00	4.00	2.00	16900.00	2.00	-	-
7	05/07/87	4.00	4900.00	1500.00	1.00	1.00	1.00	4.00	2.00	7890.00	2.00	-	-
8	04/12/83	4.00	5600.00	.00	1.00	.00	2.00	4.00	4.00	7160.00	2.50	-	-
9	02/05/90	2.00	11000.00	.00	3.00	.00	2.00	3.00	4.00	13100.00	2.25	-	-
10	12/07/80	2.00	8700.00	3400.00	3.00	1.00	4.00	2.00	4.00	13970.00	2.75	-	-
11	07/09/92	4.00	7000.00	.00	2.00	.00	4.00	3.00	2.00	8700.00	2.25	-	-
12	12/10/86	3.00	9800.00	5000.00	2.00	1.00	4.00	3.00	3.00	16780.00	2.75	-	-
13	12/03/87	3.00	10000.00	.00	3.00	.00	3.00	1.00	3.00	12000.00	1.75	-	-
14	03/07/87	2.00	9900.00	5700.00	3.00	1.00	2.00	1.00	1.00	17590.00	1.25	-	-
15	12/04/98	4.00	6000.00	.00	1.00	.00	3.00	2.00	4.00	7600.00	2.25	-	-
16	04/18/80	1.00	14800.00	1200.00	1.00	1.00	3.00	2.00	4.00	18480.00	2.50	-	-
17	09/06/91	1.00	15200.00	1500.00	3.00	1.00	1.00	4.00	3.00	19220.00	2.25	-	-
18	05/02/75	2.00	13300.00	2000.00	1.00	1.00	1.00	4.00	2.00	17630.00	2.00	-	-
19	03/12/87	1.00	12000.00	.00	1.00	.00	2.00	4.00	4.00	14200.00	2.50	-	-
20	07/03/87	2.00	9760.00	.00	2.00	.00	2.00	2.00	1.00	11736.00	1.25	-	-
21	04/12/83	2.00	11000.00	1200.00	3.00	1.00	4.00	3.00	3.00	14200.00	2.75	-	-

١-٤-٦ تحديد مجموعة من الحالات selecting Cases

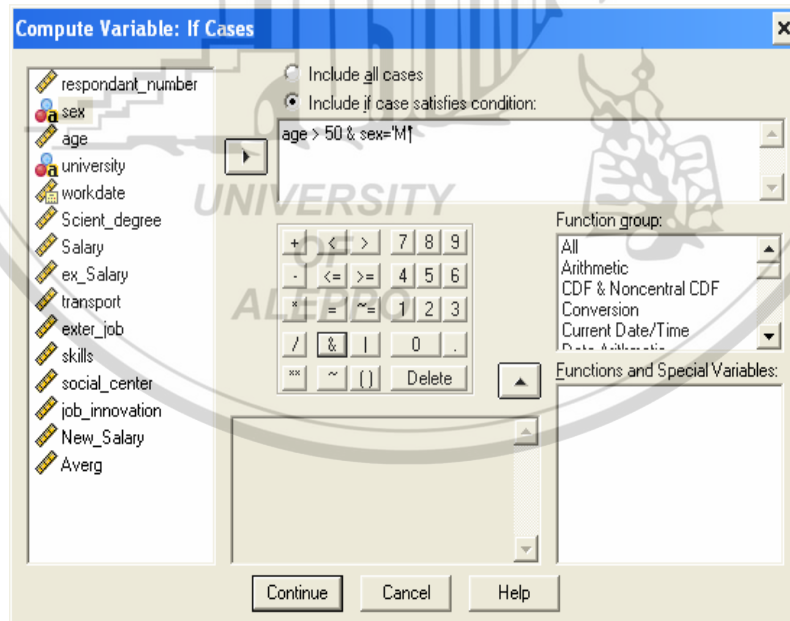
- من القائمة Transform نختار Compute نحصل على صندوق الحوار Compute Variable.

- في المستطيل المعنون Target Variable نكتب New_Salary2 .

- في المستطيل المعنون Numeric Expression نحاول كتابة الصيغة الرياضية.



- نضغط على الزر if نحصل على صندوق الحوار التالي:



- نختار الخيار Include if case satisfies condition .

- نقوم بإدخال التعبير الشرطي المطلوب $age > 50 \& sex='M'$ وذلك بالنقر المزدوج على

المتغيرات المطلوبة ونقر الأزرار الخاصة بالرموز من الآلة الحاسبة.

ملخص

الفصل الأول

تشغيل برنامج SPSS

- ١- بالنقر المزدوج Double Click على أيقونة SPSS في حال وجود اختصار لها.
 - ٢- بالنقر على زر ابدأ Start الموجود أسفل الشاشة الرئيسية:
- Start→All Programs→SPSS for Windows→SPSS 15 for Windows
Evaluation Version في الحالتين تظهر النافذة Data Editor بعنوان Untitled1
[Dataset0] – SPSS Data Editor

أنواع الملفات في برنامج SPSS:

- ١- ملف البيانات DATA Files : تضم البيانات المستخدمة في التحليل الإحصائي، وعادة ما يكون لهذا النوع من الملفات الامتداد SAV .
- ٢- ملف المخرجات الإحصائية Output Files : وهي الملفات التي تضم مخرجات التحليل الإحصائي والمخططات البيانية، ويكون لهذا النوع من الملفات الامتداد SPO .
- ٣- ملف التعليمات Syntax : وهي الملفات التي تضم الإجراءات الإحصائية المخزنة على شكل أوامر برمجية، ويكون لهذا النوع من الملفات الامتداد SPS .

تهيئة الملفات في برنامج SPSS:

- يتم بواسطة محرر البيانات Data Editor الذي يعتبر بمنزلة ورقة عمل Spread Sheet، النافذة تفتح تلقائياً بمجرد تشغيل البرنامج، تشتمل على نوعين من العرض:
- ١- عرض البيانات Data View: عبارة عن ورقة عمل تضم صفوفاً وأعمدة، تمثل الأعمدة المتغيرات Variables والصفوف الحالات Cases، حيث تمثل كل حالة منفصلة مستجوباً، والخلية تنتج من تقاطع المتغير مع الحالة.
 - ٢- عرض المتغيرات Variable View: ورقة عمل تحتوي على وصف للمتغيرات، الصفوف هي المتغيرات، والأعمدة تمثل وصفاً لكل متغير، يمكن إضافة أو تغيير الصفات.

تحويل البيانات Data Transformation : يمكن إنجاز تحويل على البيانات الأصلية Raw Data يتراوح بين تجميع البيانات في فئات معينة إلى تكوين متغيرات جديدة باستخدام معادلات و تعبيرات شرطية.

استخدام القيم لمتغير جديد باستخدام الأمر Compute: من القائمة Transform نختار Compute نحصل على صندوق حوار يسمى Compute Variable.

في المستطيل Target Variable نكتب اسم المتغير الجديد المطلوب حساب قيمه .
- في المستطيل المعنون Numeric Expression نحاول كتابة الصيغة بوساطة الألة الحاسبة الظاهرة في نافذة الحوار مع استخدام لوحة المفاتيح إذا تطلب الأمر ذلك.
بالضغط على الزر OK نحصل على المتغير الجديد في نهاية ملف البيانات ويحتوي القيم الجديدة المحسوبة بناء على الصيغة الرياضية السابقة.
بعض التوابيع الاحصائية في برنامج SPSS .

- $CFVAR(numexpr, numexpr[,...])$ تحسب معامل الاختلاف لقيم رقمية.
 - $LAG. LAG(variable)$ تحسب قيم المتغير في الحالة السابقة بملف البيانات سواء كانت البيانات رقمية أو نصية.
 - $MAX(value, value[,...])$ تعطي أكبر قيمة سواء كانت البيانات رقمية أو نصية.
 - $MEAN(numexpr, numexpr[,...])$ تحسب الوسط الحسابي للقيم الرقمية لوسطاء الدالة.
 - $MIN(value, value[,...])$ تعطي أصغر قيمة سواء كانت البيانات رقمية أو نصية.
 - $SD(numexpr, numexpr[,...])$ تحسب الانحراف المعياري للقيم الرقمية لوسطاء الدالة.
 - $VARIANCE(numexpr, numexpr[,...])$ تحسب التباين للقيم الرقمية لوسطاء الدالة.
 - $SUM(numexpr, numexpr[,...])$ تحسب المجموع للقيم الرقمية لوسطاء الدالة.
- التعبيرات الشرطية Conditional Expressions**: تستخدم لإجراء تعديلات على مجموعة محددة من الحالات التي تشتمل عليها ملف البيانات.

وصف المتغير Label: وفيه يتم وصف كل متغير بجملة موجزة تعبر عن السؤال وتوضحه حيث تظهر أوصاف المتغير وقيمته في النتائج الإحصائية وعلى الرسوم البيانية مما يساعد على تفسير النتائج .

قيم ورموز المتغير Values: وفيه يتم تعيين رموز وأرقام ، فمثلاً بالنسبة إلى السؤال الثاني نضع الرمز F لتدل على الأنثى Female والرمز M ليدل على الذكر Male .

القيم المفقودة Missing: تحل محل الإجابات غير المتوفرة بسبب عدم قيام بعضهم بالإجابة على أنواع معينة من الأسئلة ، البرنامج يفترض عدم وجود قيم مفقودة Non وفي حالة عدم إعلامه يقوم بوضع نقطة بدلاً من الإجابة غير المتاحة وتسمى بهذه الحالة بقيم النظام المفقودة ، أما إذا قمنا بتعيين رمز لهذه القيم فتسمى بقيم المستخدم المفقودة .

عرض العمود وموقع البيانات Columns & Align: وفيه يتم إعلام البرنامج بعرض العمود وموقع البيانات داخل العمود يميناً أو يساراً أو في منتصفه ويتعلق هذان العمودان بالنواحي الجمالية للجدول .

وحدة القياس Measure: وفيه يتم إعلام البرنامج بوحدة القياس التي يتم بموجبها قياس المتغير وهناك ثلاثة أنواع من المتغيرات:

- نسبية أو فئوية Scale : يستعمل للبيانات العددية القابلة للقياس الكمي وهو المقياس المستخدم غالباً لقياس متغيرات الطول والوزن ودرجات الحرارة.

- اسمية Nominal: يستعمل لقياس المتغيرات الاسمية وهي متغيرات لها عدد من الفئات دون أفضلية لإحداها على الأخرى، لا يمكن ترتيبها تصاعدياً أو تنازلياً ولكن يمكن تحديد الفروق بينها بدقة كبيرة.

- ترتيبية Ordinal : يستعمل لقياس المتغيرات الترتيبية إذ إن هذا المتغير له عدد محدد من الفئات يمكن ترتيبها تصاعدياً أو تنازلياً ولكن لا يمكن تحديد الفروق بينها بدقة كبيرة.

جدول البيانات Data View : جدول فارغ من البيانات يشتمل على أعمدة معنونة بأسماء المتغيرات، يكون هذا الجدول جاهزاً لاستقبال البيانات من الاستبيانات التي قمت بجمعها.

◆ **F8.2** تعني أن المتغير رقمي ، وأن سعته ثمانية رموز وخانتان عشريتان.

◆ **A8** تعني أن المتغير حرفي وأن سعة هذا المتغير هي ثمانية أو 8 خانات.

محرك البيانات في برنامج SPSS :

يعتبر بمثابة ملف جديد جاهز لاستقبال المتغيرات وقيمها، ومن أسفل الصفحة يمكن الاختيار من بين جدول تعريف المتغيرات، وجدول البيانات بالنقر على الخيار Variables View .
جدول تعريف المتغيرات Variables View: بمجرد النقر على هذا الخيار الموجود أسفل صفحة محرك البيانات نحصل على جدول فارغ يشتمل على 10 أعمدة معنونة .

اسم المتغير Name: وفيه يتم تحويل كل سؤال في الاستبيان إلى متغير يعطى اسماً معيناً مختصراً يعبر عن السؤال مشتقاً من السؤال نفسه وبما لا يزيد عن ثمانية حروف، يفضل أن يخصص خانة أو متغير يمثل رقم السؤال في الاستبيان أي يعطى لكل شخص رقماً يميزه عن غيره ويفيد في اكتشاف مواقع الأخطاء في حال حدوثها .

كتابة أسماء المتغيرات في برنامج SPSS يجب إتباع الخطوات التالية:

- يجب ألا يزيد طول الاسم عن ثمانية رموز Characters .
- يجب أن يبدأ اسم المتغير بحرف أما بقية الرموز فقد تكون أحرفاً أو أرقاماً أو نقطة .
- لا يجوز أن ينتهي اسم المتغير بنقطة .
- يجب ألا يتضمن اسم المتغير فراغات وبعض الرموز الخاصة مثل * , ' , ? , ! .
- لا يميز برنامج SPSS بين الحروف الصغيرة والحروف الكبيرة فالأسماء SEX و sex تعتبر متماثلة إذ أن البرنامج لا يقبل سوى الحروف الصغيرة لأسماء المتغيرات .

نوع المتغير Type: وفيه يتم وصف نوع المتغير Variable Type إذ يفترض برنامج SPSS أن جميع المتغيرات رقمية ، ولتغيير نوع المتغير نضغط مرتين متتاليتين على الكلمة الافتراضية Numeric فيظهر صندوق حوار تحت مسمى Variable Type .

عرض المتغير Width: وفيه يتم تحديد عرض المتغير Width إذ يفترض البرنامج أن جميع المتغيرات رقمية مفترضاً 8 خانات لكل متغير وبإمكاننا تغيير العرض حسب الحاجة آخذين بالاعتبار تخصيص كل حرف أو رقم أو فاصلة أو نقطة أو إشارة .

عدد الخانات العشرية Decimal: وفيه يتم تحديد عدد الخانات العشرية إذ يفترض برنامج SPSS أن جميع المتغيرات رقمية مفترضاً خانتين عشريتين لكل متغير ، وبإمكاننا تغيير عدد الخانات العشرية حسب حاجتنا .

-استخدام الرموز العلائقية في التعبيرات الشرطية: تحتوي معظم التعبيرات الشرطية على رمز علائقي واحد على الأقل . يمكن ربط التعبيرات الشرطية باستخدام الرموز المنطقية ، حيث يمكن ربط تعبيرين شرطين أو أكثر باستخدام الرموز المنطقية (&) و (|) فمثلاً :

`age >= 35 & sex='F'`

يتم اختيار الحالات التي تحقق شرط العمر و شرط الجنس معاً.

`age >= 35 | sex="F"`

يتم اختيار الحالات التي تحقق إما شرط العمر أو شرط الجنس .



A: الاستبيان المفقودة ، B: المستخدم المفقودة ، C: النظام المفقودة ، D: البيانات المفقودة ،
E: (A+B).

- إذا قمنا بتعيين رمز للقيم المفقودة بعمود Missing تسمى القيم بهذه الحالة بقيم :

A: الاستبيان المفقودة ، B: المستخدم المفقودة ، C: النظام المفقودة ، D: البيانات المفقودة ، E:
(A+B+C).

- هناك عدة أنواع للملفات في البرنامج الإحصائي SPSS هي ملفات:

A: الاستكشاف ، B: التقاطع، الانحدار ، C: العودة، والاستيطان، D: البيانات، المخرجات
الإحصائية، التعليمات ، E: (A+B).

- جدول عرض البيانات Data View :

A: يكون جاهز لاستقبال البيانات ، B: يتم فيه وصف المتغيرات ، C: يحتوي على ٨ أعمدة،
D: يتم فيه تحديد نوع المتغير ، E: (A+C).

- ملف البيانات DATA Files:

A : يتشكل باستخدام محرر البيانات ، B : يضم بيانات التحليل الإحصائي ، C : له
الامتداد SAV ، D : (A+B+C) ، E : (A+C).

- ملف التعليمات Syntax :

A: له الامتداد SPO ، B: له الامتداد SPS ، C: يضم الإجراءات المخزنة على شكل
أوامر برمجية ، D : (B+C) ، E : (A+C).

- تهيئة الملفات في برنامج SPSS:

A: يتم بواسطة محرر البيانات ، B: يعتبر ورقة عمل، C: النافذة تفتح تلقائياً، D: نافذة تضم
نوعين لعرض للبيانات، E: كل ماسبق صحيح.

- عرض المتغيرات Variable View:

A: ورقة عمل تصف المتغيرات ، B : الصفوف هي المتغيرات ، C: الأعمدة تمثل وصفاً لكل
متغير ، D : يمكن إضافة أو تغيير صفات أي متغير ، E: كل ما سبق صحيح.

- محرر البيانات في برنامج SPSS:

A: ورقة عمل تصف المتغيرات ، B يعتبر ملف جديد جاهز لاستقبال المتغيرات وقيمها، C:
يمكننا من الاختيار بين جدول المتغيرات و جدول البيانات، D: (B+C) ، E: (A+B+ C).

- جدول تعريف المتغيرات Variables View عبارة عن جدول فارغ يشتمل على:

- الرمز F8.2 تعني أن المتغير حرفي ، وأن سعته ثمانية رموز وخانتين عشريتين .
- ملف البيانات DATA Files يتشكل باستخدام محرر البيانات Data Editor و له الامتداد .SPSS
- لكتابة أسماء المتغيرات في برنامج SPSS، يجب ألا يزيد طول الاسم عن ثمانية رموز Characters.
- Date : متغير رقمي يراد عرض قيمه بحيث يشتمل على إشارة الدولار \$ مع فاصلة كل ثلاثة مواقع.
- Custom Currency: متغير حرفي يراد عرض قيمه بحيث يشتمل على أحرف أو رموز أخرى .
- هناك ثلاثة أنواع من الملفات في برنامج SPSS تأخذ الامتدادات التالية : SPSS , SAV , SPO.
- في الصندوق Box ضمن المخطط Boxplots تبعد القيم المتطرفة عن حافتي الصندوق أقل من 3 من طول الصندوق.
- في العمود Type يتم وصف نوع المتغير Variable Type إذ يفترض SPSS أن جميع المتغيرات حرفية.
- محرر البيانات في برنامج SPSS عبارة ورقة عمل تصف المتغيرات وملف جديد جاهز لاستقبال المتغيرات وقيمتها .
- لكتابة أسماء المتغيرات في برنامج SPSS يجب أن يبدأ الاسم بحرف ولا يجوز أن ينتهي هذا الاسم بنقطة.
- أختر الإجابة الصحيحة فقط معناها ظلل المربع (A أو B أو C أو D أو E) .**

- هناك ثلاث وحدات لقياس المتغير هي:

A: نسبية \حرفية \رتبيه ، B: كمية وصفية ، C: عشرية \عملة \رتبيه ، D: رتبيه \اسمية \عملة ، E: نسبية \اسمية \رتبيه.

- في العمود Type (نوع المتغير) ، إذا كان المتغير المراد إدخاله اسماً فإننا نختار:

A: String ، B: Numeric ، C: Data ، D: Dollar ، E: Missing Values.

- في العمود Missing وفي حال عدم إعلام البرنامج توضع نقطة بدلاً من الإجابة المتاحة وتسمى قيم :

أسئلة وتمارين

أجب بصح أو خطأ (A تعنى أن الإجابة صحيحة، B تعنى أن الإجابة خاطئة)

- Scientific notation : متغير رقمي يراد عرض قيمه على شكل تعبير أسّي يستخدم الحرفين E أو D ليقوم مقام الرقم 10.
- لكتابة أسماء المتغيرات في برنامج SPSS يجب ألا يبدأ الاسم بحرف أما بقية الرموز فقد تكون أحرفاً أو أرقاماً أو نقطة.
- الرمز A8 : يعني أن المتغير رقمي وأن سعة هذا المتغير هي ثمانية رموز أو 8 خانات
- في الصندوق Box ضمن المخطط Boxplots تبعد القيم الشاذة عن حافتي الصندوق أقل من 1.5 من طول الصندوق.
- لا يميز برنامج SPSS بين الحروف الصغيرة والحروف الكبيرة ولا يقبل سوى الحروف الصغيرة لأسماء المتغيرات.
- جدول تعريف المتغيرات Variables View عبارة عن جدول فارغ يشتمل على 10 أعمدة معنونة.
- عرض البيانات Data View في برنامج SPSS عبارة عن ورقة عمل تضم صفوفاً وأعمدة .
- لكتابة أسماء المتغيرات في برنامج SPSS، يجب ألا يتضمن اسم المتغير فراغات وبعض الرموز الخاصة مثل * , ' , ? , !.
- Comma : متغير رقمي يمكنك تحديد عرضه، عدد خاناته، وعدد أرقامه العشرية، وهو النوع الافتراضي.
- ملف المخرجات الإحصائية Output Files له الامتداد SPSS و يضم مخرجات التحليل الإحصائي.
- يكون التوزيع طبيعياً إذا كانت نسبة الالتواء إلى الخطأ المعياري هي أكبر من 2.
- في الصندوق Box ضمن المخطط Boxplots تبعد القيم المتطرفة والشاذة عن حافتي الصندوق بمقدار طول الصندوق .
- Numeric : متغير رقمي يراد عرض قيمه بحيث يشتمل على فاصلة لكل ثلاثة مواقع مع نقطة تفصل الخانات العشرية.
- Dollar : متغير رقمي يراد عرض قيمه بحيث يشتمل على عملة دولة معينة ثم تعريف مواصفاتها .

A: 8 أعمدة معنونة ، B : 6 أعمدة معنونة ، C: 10 أعمدة معنونة ، D : 10 أعمدة و 10 حالات ، E: 8 أعمدة و 6 حالات.

- في العمود المسمى Name يتم تحويل كل سؤال في الاستبيان لمتغير يعطى اسماً مختصراً يعبر عنه وبما لا يزيد عن:

A: ثمانية حروف ، B : عشرة حروف ، C: ستة عشر حرفاً ، D : إثني عشر حرفاً ، E: ستة عشر رمزاً.

- لكتابة أسماء المتغيرات في برنامج SPSS يجب التقيد بما يلي:

A: ألا يزيد طول الاسم عن ثمانية رموز ، B : أن يبدأ اسم المتغير بحرف صغير ، C: ألا ينتهي اسم المتغير بنقطة ، D: ألا يتضمن اسم المتغير فراغات وبعض الرموز الخاصة مثل ' , ? , ! , * ، E: كل ما سبق صحيح.

- إذا كانت لديك المعلومات: $N=58$, $Std\ Deviation = 17.56$, $Mean=68.16$ فإن الخطأ المعياري هو:

A: 0.2519 ، B : 2.305741 ، C: -50.988 ، D : 1170.50 ، E: 17.56.

- من السؤال السابق إذا كانت $t=2$ فإن مجال ثقة ٩٥ % للمتوسط الحسابي هو:

A: [53.59 , 74.75] ، B : [63.55 , 72.77] ، C : [65.85 , 75.78] ، D : [51.5 , 72.7] ، E : [-2.58 , +2.58].

- يكون التوزيع طبيعياً إذا كانت نسبة الالتواء إلى الخطأ المعياري هي:

A : أصغر من +2 ، B : أكبر من +2 ، C: أصغر من -2 ، D : أكبر من -2 ، E : [-2 , +2].

- في العمود المسمى Type يتم وصف نوع المتغير Variable Type إذ يفترض برنامج SPSS أن جميع المتغيرات:

A: حرفية ، B : رقمية ، C: من نوع Date ، D : من نوع Dollar ، E: من نوع Scientific notation.

- لتعريف القيم المفقودة لمتغير، ننقر الخلية بعمود missing يظهر صندوق حوار القيم المفقودة نختار:

A: قيمة واحدة أو عدة قيم مفقودة ، B : كلمة السر ، C: مجال من القيم مضافاً له قيمة معينة ، D : (A+B) ، E : (A+C).

- لعرض الإحصاءات الوصفية لمتغير أو عدة متغيرات إضافة لعرض القيم المعيارية Zscores لها نتبع :

Analysis → Descriptive Statistics → : B ، Frequency → Statistic : A
Analyze → Descriptive : D ، Frequency → Record : C ، Descriptive
Frequency → Standardize : E ، Statistics

- في برنامج SPSS يتم إنشاء جدول تقاطعي سواء بالنسبة لمتغير أو عدة متغيرات عن طريق الخيار:

Analyze → : C ، Crosstabs : B ، Frequency → Statistics → Crosstabs: A
Descriptive Statistics → Crosstabs : D ، Descriptive Statistics → Crosstabs
Analyze → Descriptive Statistics : E ،

- لحساب معاملات التوافق، فاي وكرامرز، لمبدأ، ومعامل عدم التأكد نختار من صندوق الحوار **Crosstabs : Statistics**

: D ، Nominal by Interval : C ، Correlations : B ، Chi-square: A
Ordinal : E ، Nominal

- لحساب معامل ارتباط بيرسون ومعامل ارتباط الرتب لسبيرمان نختار من الصندوق **Crosstabs: Statistics المربع:**

، Nominal : D ، Nominal by Interval : C ، Correlations : B ، Chi-square : A
Ordinal : E

- لتطبيق أمر مقارنة المتوسطات لعدد من المتغيرات نختار ما يلي:

:C ، Analysis → Compute → Means :B ، Analysis → Means :A
Analysis → Compare Means → Means :D ، Statistics → Compare Means
Analysis → Statistics → Means :E ،

- يستخدم الأمر **Compare Means → Means** لحساب ومقارنة الأوساط الحسابية للمتغيرات بالإضافة لإجراء:

A : تحليل التباين ، B : اختبار التأثير الخطي للمعاملات ، C : اختبار الاقتران Eta ، D :
(A+B+C) : E ، (A+C)

A: أصغر قيمة ، B: التباين للقيم الرقمية ، C: معامل الاختلاف ، D : الوسط الحسابي ، E: الانحراف المعياري قيم رقمية.

- تستخدم التعبيرات الشرطية **Conditional Expressions** لإجراء تعديلات على مجموعة محددة من:

A: التوابع بملف البيانات، B: المتغيرات بملف البيانات، C: الحالات بملف البيانات، D : الرموز المنطقية، E: الوسائط.

- باستخدام التعبير الشرطي التالي في برنامج SPSS يتم اختيار الحالات التي تحقق شرط العمر وشرط الجنس معاً:

A: `age >= 35 | sex="F"` ، B: `age >= 35 OR sex="F"` ، C: `age >= 35 & sex="F"` ، D: `age >= 35 and sex="F"` ، E: كل ما سبق غير صحيح.

- باستخدام التعبير الشرطي التالي في برنامج SPSS يتم اختيار الحالات التي تحقق إما شرط العمر أو شرط الجنس:

A: `age >= 35 | sex="F"` ، B: `age >= 35 OR sex="F"` ، C: `age >= 35 & sex="F"` ، D: `age >= 35 and sex="F"` ، E: كل ما سبق غير صحيح.

- في برنامج SPSS يتم تشكيل جدول التوزيع التكراري سواء بالنسبة لمتغير أو عدة متغيرات عن طريق:

A: Analyze → Statistics → Frequency ، B: Compute → Frequency ، C: Analyze → Record ، D: Analyze → Descriptive ، E: Statistics → Frequency Variable View.

- نحصل على الأعمدة البيانية، الدوائر البيانية، المدرج التكراري مع أو بدون المنحنى الطبيعي من صندوق الحوار .

A: Graphs : Frequency ، B: Charts : Graphs ، C: Graphs : ، D: Analyze : ، E: Charts : Frequencies .

- نحصل على تكرار قيم المتغير بالإضافة للنسب والربيعات والمئينات ومقاييس النزعة والتشتت من صندوق الحوار:

A: Graphs : Frequency ، B: Statistics : Graphs ، C: Graphs : ، D: Statistics : Frequencies .

- يتكون المخطط Boxplots في الأمر Explore من:

A: الصندوق Box فقط ، B : القيم الشاذة Outliners ، C: الاستطالات خارج الصندوق
Whiskers ، D : القيم المتطرفة Extremes ، E: كل ماسبق صحيح.

- الصندوق Box يشمل المكونات التالية:

A: الربع الثاني Q_2 ، B : الربع الأول Q ، C: الربع الثالث Q_3 ، D : $(A + B + C)$ ،
E: $(A+B)$.

- من المخططات الوصفية الموحدة ضمن الأمر plots في الأمر Explore :

A: Stem and Leaf ، B: Normality plots ، C: Histogram ، D: Boxplots ، E
: كل ما سبق صحيح.

- هناك ثلاثة أنواع من الملفات في برنامج SPSS تأخذ الامتدادات التالية:

A: (ExL,SPO,Msf) ، B : (SAV,SPO,SPSS) ، C: (SAV,SPO,SMS) ،
D : (SAV,SPO,SPS) ، E: (DBM).

- يمكن حساب متغيرات جديدة استناداً لقيم المتغيرات الأصلية وذلك باستخدام الأمر.

A: Analyze ، B : Compute ، C : Record ، D : Target ، E : Variable View.

- يحتوي صندوق الحوار Compute Variable على آلة حاسبة تضم أرقاماً، ورموزاً
حسابية، ورموزاً:

A: علائقية وشرطية ، B: منطقية وشرطية ، C: علائقية ومنطقية ، D : شرطية ، E: رياضية
منطقية.

- ينبغي مراعاة تسلسل تنفيذ العمليات الحسابية في أي تعبير حسابي، وفقاً لقاعدة الأولوية
على النحو:

A: الرفع لقوة، الضرب ، الجمع ، B: الضرب، القسمة، الجمع والطرح ، C: اقتران، قوة،
ضرب ، قسمة ، D : الاقتران، رفع لقوة، الضرب أو القسمة، الجمع والطرح ، E: كل ما سبق
غير صحيح.

- في برنامج SPSS تستخدم الدالة CFVAR(numexpr,numexpr[,...]) لحساب:

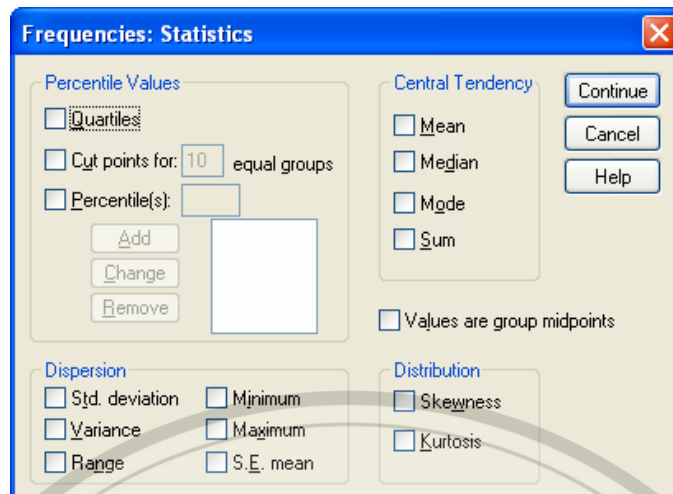
A: أكبر قيمة ، B: التباين لقيم رقمية ، C: مجموع القيم ، D : معامل الاختلاف لقيم رقمية ، E:
معامل الاختلاف لقيم نوعية.

- في برنامج SPSS تستخدم الدالة SD(numexpr,numexpr[,...]) لحساب:

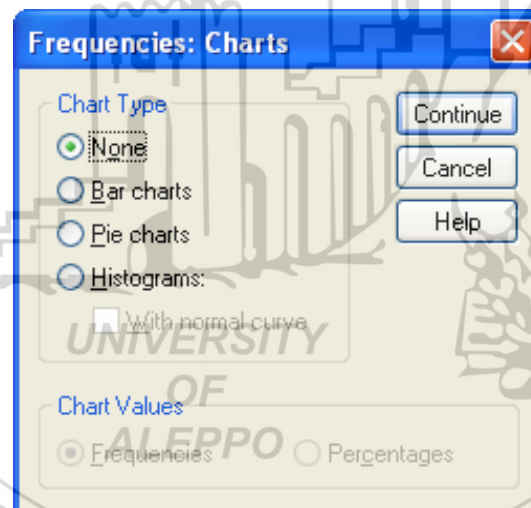
- ملف المخرجات الإحصائية Output Files:

A: يضم مخرجات التحليل الإحصائي ، B: له الامتداد SPO ، C: له الامتداد SPSS ،
D : (A+B) ، E : (A+C).





بالضغط على المربع Continue ننتقل ثانية إلى صندوق الحوار Frequency ، فإذا تم الضغط على المربع Ok نحصل على الإحصاءات الوصفية التي تم تفعيلها . كما أن هناك خيارات أخرى تتعلق بالأشكال البيانية وذلك عن طريق الضغط على الخيار Charts الموجود ضمن المربع الموجود أسفل صندوق الحوار Frequency فنحصل على صندوق حوار جديد يسمى Frequencies : Charts .



من هذا الخيار يمكن للباحث تفعيل ما يناسبه من الأشكال البيانية :

- الأعمدة البيانية Bar Charts .
- الدوائر البيانية Pie Charts .
- المدرج التكراري مع المنحنى الطبيعي أو بدونه Histogram with normal curve.

Better for Man to Work , Woman tend Home

من الأفضل للرجل أن يعمل وللمرأة أن تهتم بالمنزل ؟

وللإجابة عن هذا السؤال هناك خيارات متعددة يمكن أن يختار من بينها الشخص الذي يتم توجيه السؤال إليه وهي:

موافق جداً Strongly Agree ، موافق Agree ، غير موافق Disagree ، غير موافق جداً Strongly Disagree .

وبمجرد النقر على الخيار OK في صندوق الحوار Frequency نحصل جدول التوزيع التكراري المدرج أدناه ، إذا اكتفينا بتفعيل الخيار عرض جدول التوزيع التكراري Display frequency tables .

Better for Man to Work , Woman tend Home

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly agree	85	32.2	32.8	32.8
	Agree	98	37.1	37.8	70.7
	Disagree	71	26.9	27.4	98.1
	Strongly Disagree	5	1.9	1.9	100.0
	Total	259	98.1	100.0	
Missing	System	5	1.9		
Total		264	100.0		

مع العلم أن هناك خيارات أخرى وهي الحصول على كافة الإحصاءات الوصفية وذلك عن طريق الضغط على الخيار Statistics الموجود ضمن مربع أسفل صندوق الحوار Frequency فنحصل على صندوق حوار جديد يسمى Statistics : Frequencies ومن هذا الخيار يمكن للباحث تفعيل ما يناسبه من الإحصاءات الوصفية بدءاً من الربيعات والمئينات إلى مقاييس النزعة المركزية : الوسط الحسابي ، الوسيط ، المنوال ، المجموع ، ومقاييس التشتت : الانحراف المعياري ، التباين ، المدى ، الخطأ المعياري ، أصغر قيمة ، أكبر قيمة وانتهاء بمقاييس الالتواء والتفرطح .

Analysis → Descriptive Statistics → Frequencies

Frequencies → Statistics → Frequencies: Statistics

الفصل الثاني

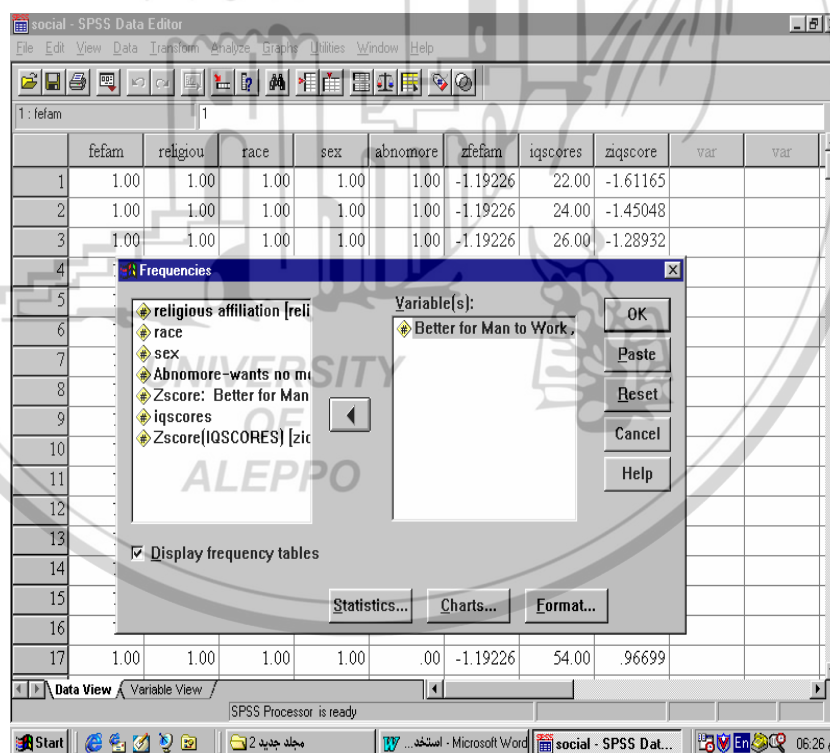
الإحصاءات الوصفية

Descriptive Statistics

§-1 إنشاء التوزيع التكراري Producing Frequency Distributions

في برنامج SPSS يمكنك تشكيل جدول التوزيع التكراري سواء بالنسبة لمتغير واحد أو لعدة متغيرات بآن واحد ، ويمكن الحصول على التوزيع التكراري لأحد المتغيرات عن طريق الخيار المسمى التكرار Frequency من الإحصاءات الوصفية Descriptive Statistics الذي نحصل عليه بعد النقر على الخيار تحليل "Analyze" من شريط القوائم فيظهر صندوق حوار بعنوان Frequencies كما هو موضح بالشكل التالي:

من الطرف الأيسر في صندوق الحوار Frequency نضغط مرتين متتاليتين على المتغير الذي نرغب إجراء التحليل عليه فينتقل تلقائياً للطرف الأيمن تحت عنوان المتغيرات وفي حالتنا



نضغط على المتغير fefam والذي يمثل السؤال التالي :

Statistics

Better for Man to Work , Woman tend Home

N	Valid	259
	Missing	5
Mean		1.9846
Std. Error of Mean		5.131E-02
Median		1.9508 ^a
Mode		2.00
Std. Deviation		.8258
Variance		.6819
Skewness		.237
Std. Error of Skewness		.151
Kurtosis		-1.006
Std. Error of Kurtosis		.302
Range		3.00
Minimum		1.00
Maximum		4.00
Sum		514.00
Percentiles	10	b, c
	20	1.1016
	25	1.2432
	30	1.3847
	40	1.6678
	50	1.9508
	60	2.2533
	70	2.5598
	75	2.7130
	80	2.8663
	90	3.3842

a. Calculated from grouped data.

b. The lower bound of the first interval or the upper bound of the last interval is not known. Some percentiles are undefined.

c. Percentiles are calculated from grouped data.

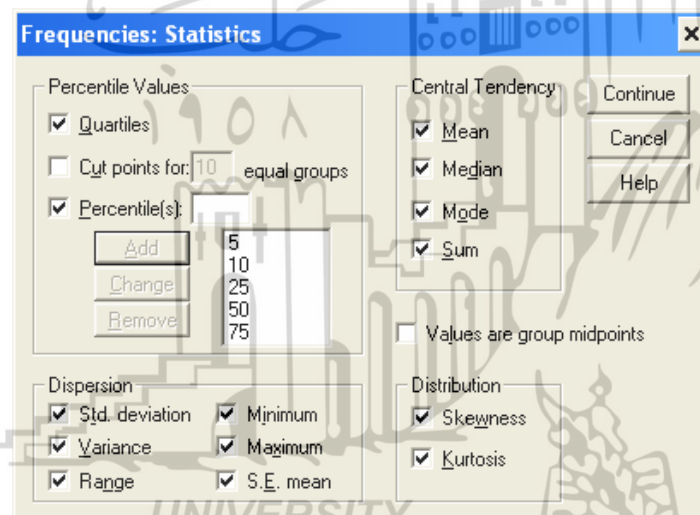
وبالضغط على المربع Continue ننتقل ثانية إلى صندوق الحوار Frequency ، فإذا تم انتقاء الخيار " المدرج التكراري مع المنحنى الطبيعي Histogram with normal curve ، وبالضغط على المربع Ok نحصل على الأشكال البيانية التي تم تفعيلها ، وهي في مثالنا المدرج التكراري مع المنحنى الطبيعي ، وتظهر إلى جانبه الوسط الحسابي والانحراف المعياري بالإضافة لعدد المشاهدات .

17	15200	47
18	13300	39
19	12000	44
20	9760	41
21	11000	43
22	16000	50

Analysis → Descriptive Statistics → Frequencies

يظهر صندوق حوار Frequencies نختار المتغير Salary ونفعل الخيار عرض جدول التوزيع التكراري Display frequency tables .

- **Statistics** : يعرض بعض الإحصائيات حسب حاجة المستخدم، كما يوضحه صندوق الحوار:



- **Percentile Values** : تشمل الربعات Quartile والمئينات Percentiles وقد حددنا المئين الخامس والعاشر والخامس والعشرون...الخ.
- **Cut points for....equal groups** : يحدد القيم التي تقسم البيانات إلى عدد من الفئات المتساوية في الطول يحددها الباحث بنفسه.
- **Dispersion** : لحساب مقاييس التشتت.
- **Central Tendency** : لحساب مقاييس النزعة المركزية.
- **Distribution** : لعرض مقاييس التوزيع وتشمل الالتواء والتفرطح.
- **Values are group midpoints** : في حالة تفعيل هذا الخيار فإن برنامج SPSS سيقوم بحساب الوسيط وقيم المئينات Median , Percentiles Values باعتبار أن بيانات المتغير المدروس تمثل مراكز الفئات للبيانات المبوبة.

Frequencies: Statistics

Percentile Values

☐ **Quartiles**

☐ Cut points for: equal groups

☐ Percentile(s):

Central Tendency

☐ Mean

☐ Median

☐ Mode

☐ Sum

☐ Values are group midpoints

Dispersion

☐ Std. deviation

☐ Variance

☐ Range

☐ Minimum

☐ Maximum

☐ S.E. mean

Distribution

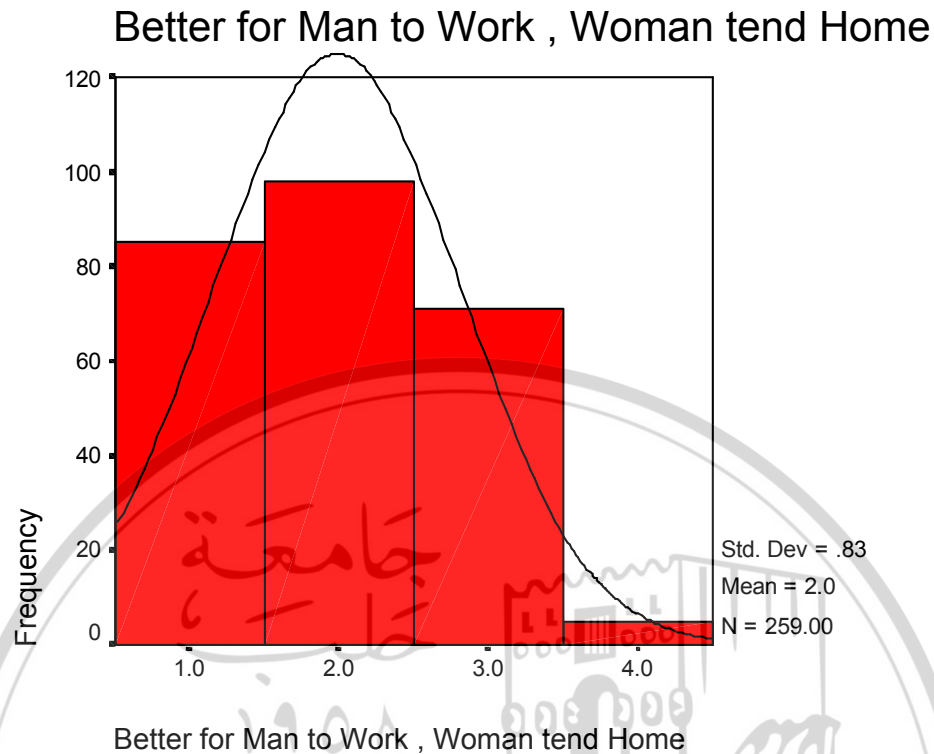
☐ Skewness

☐ Kurtosis

مثال (١):

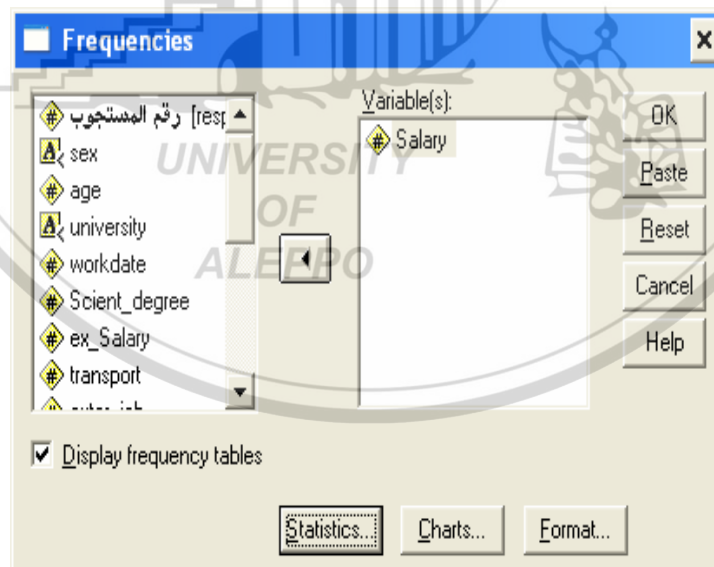
ليكن لدينا المتغير **Salary** الموضح في الجدول التالي والذي يشتمل على دخل 22 عضو هيئة تدريسية، لحساب تكرار قيم المتغير المدروس مع حساب بعض المقاييس الإحصائية نتبع ما يلي:

Respd_num	Salary	Age
1	12000	35
2	11000	40
3	12500	44
4	15000	52
5	10500	38
6	9000	39
7	4900	29
8	5600	28
9	11000	46
10	8700	37
11	7000	29
12	9800	32
13	10000	34
14	9900	36
15	6000	29
16	14800	50



١-١ الأمر Frequency

يستخدم هذا الأمر Frequencies لعرض تكرار كل قيمة من قيم المتغير قيد الدراسة بعد تفعيل المربع Display frequency tables كما هو موضح بالشكل:



بالإضافة إلى حساب بعض الإحصائيات Statistics كالنسب والربيعات والمئينات وبعض مقاييس النزعة المركزية والتشتت والتوزيع وعرض بعض المخططات البيانية Charts مثل الأعمدة البيانية والدوائر والمدرج التكراري.

Chart - يعرض هذا المربع بعض المخططات البيانية، وفي حالة اختياره يظهر صندوق حوار بعنوان Frequencies Charts يشتمل على ما يلي:

Chat Type : نوع التخطيط:

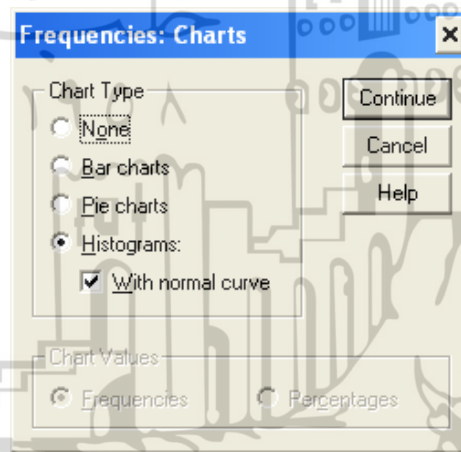
Non : عدم عرض أية مخططات .

Bar charts : يستخدم لعرض الأعمدة البيانية.

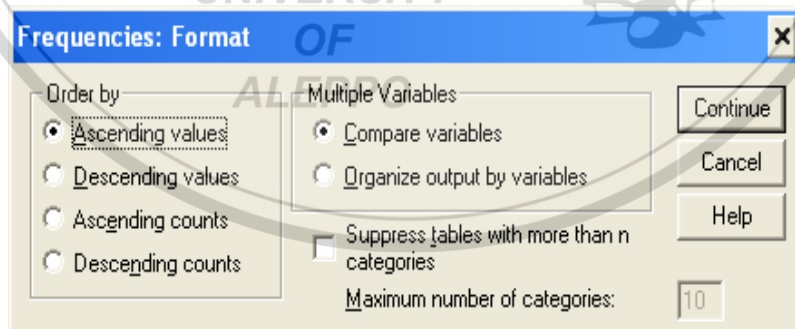
Pie charts : يستخدم لعرض الدوائر.

في حالة تفعيل أحد الخيارين السابقين على المستخدم أن يختار نوع القيم المستخدمة في التخطيط هل هي تكرارات أم نسب مئوية كما يوضحه الجزء السفلي من صندوق الحوار.

Histograms : يستخدم لعرض المدرج التكراري فقط أو المدرج التكراري مع منحنى التوزيع الطبيعي.



Format : يعرض هذا المربع بعض الخيارات الترتيبية والمقارنة والحذف كما يلي:



Order by : تستخدم لترتيب المشاهدات في الجدول التكراري تصاعدياً أو تنازلياً حسب القيم Values أو التكرارات أي عدد القيم Counts.

Multiple Variables : يستخدم في حالة وجود أكثر من متغير نرغب بدراسته، ويشتمل على خيارين الأول: compare variables يعرض المؤشرات الإحصائية لكافة المتغيرات في

يشتمل الجدول على عدة أعمدة هي:

. القيم الفعلية Valid

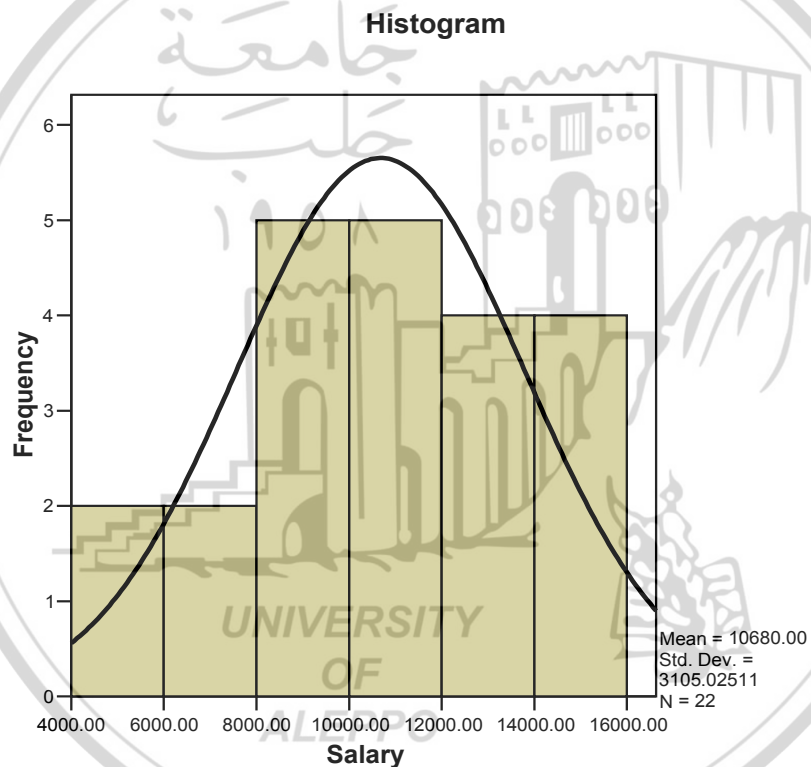
. التكرارات Frequency

. النسبة المئوية Percent

. النسبة المئوية الفعلية Valid Percent

. النسبة المئوية التجميعية Cumulative Percent

مع ملاحظة أن المشاهدات في الجدول تم ترتيبها تصاعدياً حسب قيم المتغير المدرس Salary



٢-١ الأمر Descriptive

يستخدم هذا الأمر **Descriptive** لعرض مقاييس الإحصاء الوصفي لمتغير واحد أو لمجموعة من المتغيرات بالإضافة إلى عرض القيم المعيارية **Zscores** لهذه المتغيرات وذلك بإتباع ما يلي :

- Std Error of Mean : تمثل الخطأ المعياري للمتوسط الحسابي.
- Median : تمثل قيمة الوسيط للراتب .
- mode : تمثل قيمة المنوال للراتب.
- Std. Deviation : تمثل قيمة الانحراف المعياري للراتب.
- Skewness : تمثل قيمة معامل الالتواء للتوزيع.
- Std Error of Skewness : تمثل قيمة الخطأ المعياري لمعامل الالتواء .
- Kurtosis : تمثل قيمة معامل التفرطح للتوزيع.
- Std Error of Kurtosis : تمثل قيمة الخطأ المعياري لمعامل التفرطح .
- Minimum : تمثل أقل قيمة من القيم المدروسة وهنا تمثل أقل راتب.
- Maximum : تمثل أكبر قيمة من القيم المدروسة وهنا تمثل أكبر راتب.
- Sum : تمثل المجموع ، وهنا تمثل مجموع الرواتب .
- Percentiles : وتمثل قيم الربيعات والمئينات التي يرغبها المستخدم.

Salary					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	4900.00	1	4.5	4.5	4.5
	5600.00	1	4.5	4.5	9.1
	6000.00	1	4.5	4.5	13.6
	7000.00	1	4.5	4.5	18.2
	8700.00	1	4.5	4.5	22.7
	9000.00	1	4.5	4.5	27.3
	9760.00	1	4.5	4.5	31.8
	9800.00	1	4.5	4.5	36.4
	9900.00	1	4.5	4.5	40.9
	10000.00	1	4.5	4.5	45.5
	10500.00	1	4.5	4.5	50.0
	11000.00	3	13.6	13.6	63.6
	12000.00	2	9.1	9.1	72.7
	12500.00	1	4.5	4.5	77.3
	13300.00	1	4.5	4.5	81.8
	14800.00	1	4.5	4.5	86.4
	15000.00	1	4.5	4.5	90.9
	15200.00	1	4.5	4.5	95.5
	16000.00	1	4.5	4.5	100.0
Total		22	100.0	100.0	

جدول واحد الثاني: Organize output by variables يعرض الإحصائيات لكل متغير في جدول مستقل.

- **suppress tables with more than.....categories** : يستخدم لإخفاء جدول التوزيع التكراري للمتغيرات التي يزيد عدد فئاتها عن العدد المحدد من المستفيد.

Frequencies

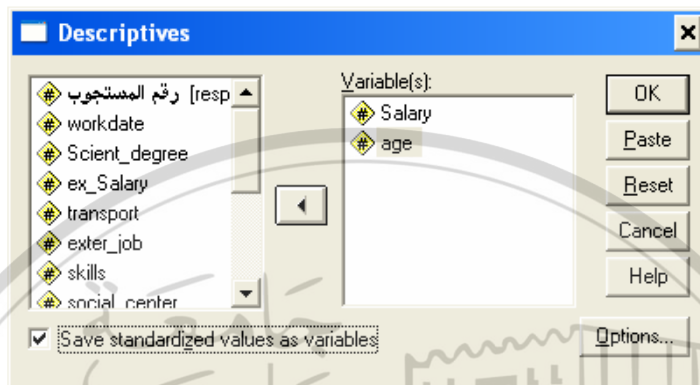
Statistics		
Salary		
N	Valid	22
	Missing	0
Mean		10680.00
Std. Error of Mean		661.99358
Median		10750.00
Mode		11000.00
Std. Deviation		3105.025
Variance		9641181
Skewness		-.121
Std. Error of Skewness		.491
Kurtosis		-.487
Std. Error of Kurtosis		.953
Range		11100.00
Minimum		4900.00
Maximum		16000.00
Sum		234960.00
Percentiles	5	5005.0000
	10	5720.0000
	25	8925.0000
	50	10750.00
	75	12700.00

إذ أن :

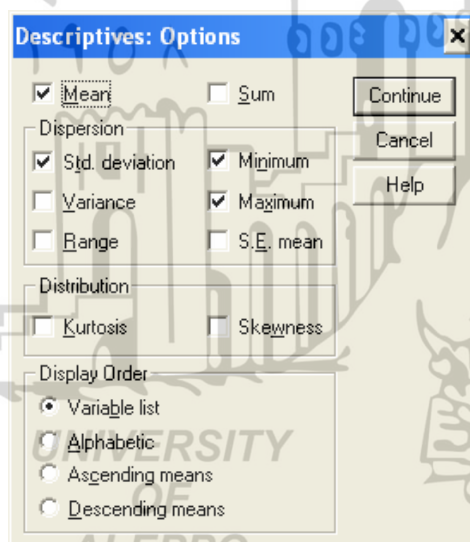
- Quartiles (25,50,75) الربعيات
- Percentiles (5,10,50) المئينيات
- Valid : تمثل القيم الفعلية التي تم دراستها (الغير مفقودة).
- Missing : تمثل القيم المفقودة.
- Mean : الوسط الحسابي للراتب.

Analysis → Descriptive Statistics → Descriptive

يظهر صندوق حوار **Descriptives** نختار المتغيرين Salary و Age ونفعل الخيار حفظ القيم المعيارية على أنها متغيرات Save Standardized Values as Variables الذي يقوم بإضافة المتغيرات المعيارية إلى محرر البيانات في برنامج SPSS.



نختار خيارات Options يظهر صندوق حوار بمسمى Descriptive: Options



نقوم باختيار بعض الإحصاءات الوصفية حسب حاجتنا ، في الوقت الذي يقوم فيه الخيار المسمى Display Order بترتيب عرض المقاييس الوصفية حسب ورودها في قائمة المتغيرات Variable List ، هجائياً Alphabetic ، عرض المقاييس الإحصائية حسب الترتيب التصاعدي للأوساط الحسابية للمتغيرات Ascending means ، عرض المقاييس الإحصائية حسب الترتيب التنازلي للأوساط الحسابية للمتغيرات Descending means .
بالنقر على الزر OK نحصل على النتائج التالية:

Descriptives

14	48	0.48349
15	50	0.64466
16	52	0.80582
17	54	0.96699
18	56	1.12815
19	58	1.28932
20	60	1.45048
21	62	1.61165

§-٣ بناء مجال ثقة حول المتوسط

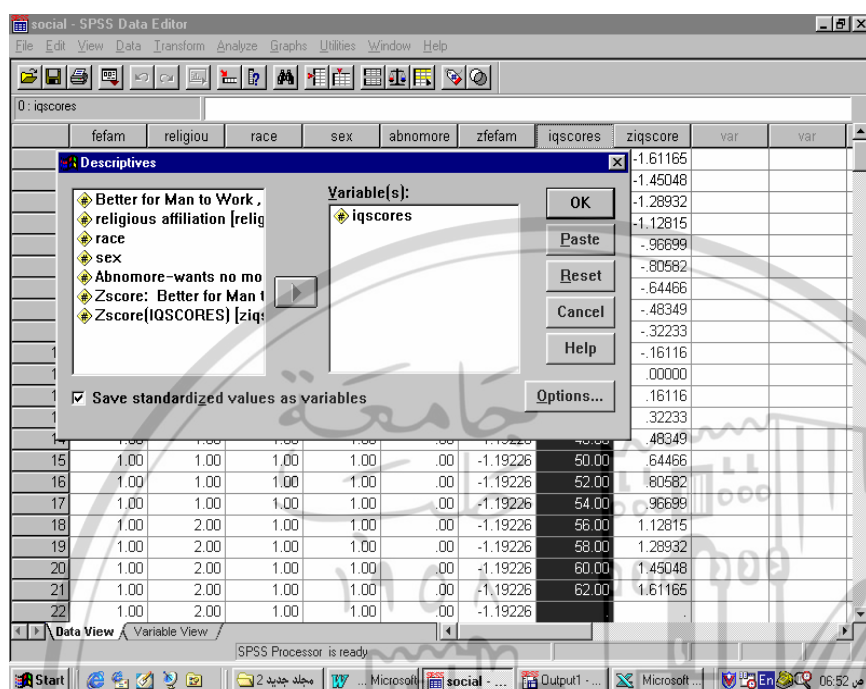
Producing Confidence Intervals Around a Mean

باستخدام برنامج SPSS يمكننا بناء مجال ثقة حول المتوسط وذلك بوساطة التعليمات Explore من الإحصاءات الوصفية Descriptive Statistics الذي تحصل عليه بعد النقر على الخيار تحليل " Analyze " من شريط القوائم وبعد تحديد المتغير الذي نرغب بناء مجال الثقة حول متوسطه ، وليكن درجات الذكاء التي حصل عليها عينة من الطلاب عددها 21 طالب وتوزعهم ما بين الذكور والإناث ، وقبل الضغط على الزر OK نتأكد من تفعيل مجال الثقة عن طريق المربع Statistics أسفل صندوق الحوار Explore ، وبالضغط على OK نحصل على النتائج المطلوبة.

Case Processing Summary

	Cases					
	Valid		Missing		Total	
SEX	N	Percent	N	Percent	N	Percent
IQSCORE female	12	100.0%	0	.0%	12	100.0%
Male	9	100.0%	0	.0%	9	100.0%

وبالضغط على الزر OK نحصل على النتائج المطلوبة وهي الدرجات المعيارية ZIQScores على أنه متغير جديد وفي عمود مستقل .



Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
IQSCORES	21	22.00	62.00	42.0000	12.4097
Valid N (listwis	21				

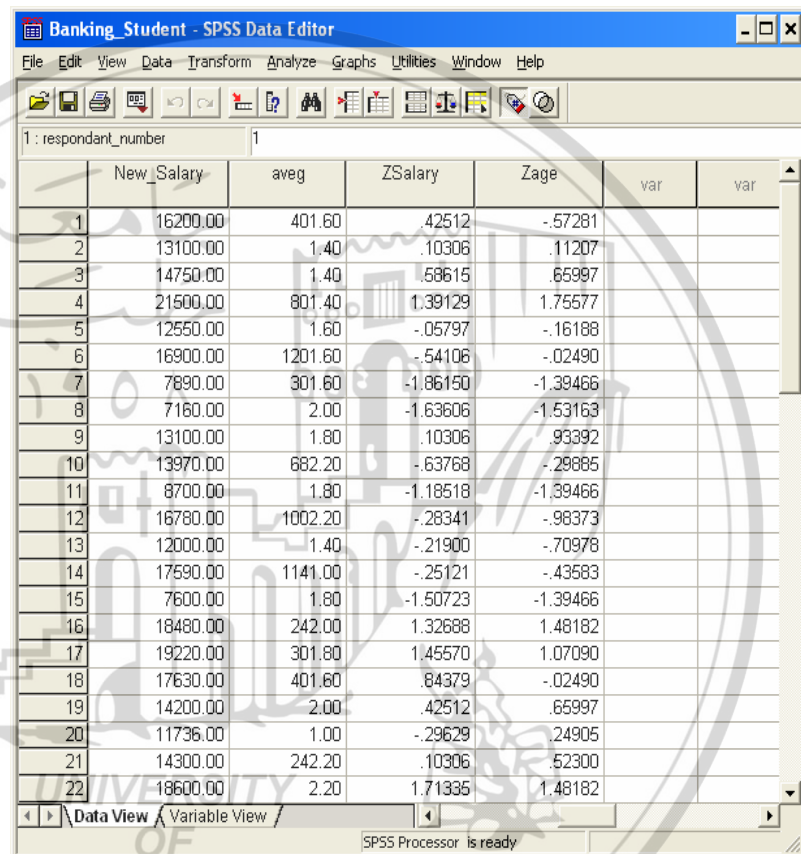
الدرجات المعيارية ZIQScores

الطالبة	درجات الذكاء IQScores	الدرجات المعيارية ZIQScores
1	22	-1.61165
2	24	-1.45048
3	26	-1.28932
4	28	-1.12815
5	30	-0.96699
6	32	-0.80582
7	34	-0.64466
8	36	-0.48349
9	38	-0.32233
10	40	-0.16116
11	42	0
12	44	0.16116
13	46	0.32233

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Salary	22	4900.00	16000.00	10680.00	3105.02511
age	22	28.00	52.00	39.1818	7.30060
Valid N (listwise)	22				

لقد تم عرض المقاييس الإحصائية حسب ترتيب المتغيرات في قائمة المتغيرات، مع إضافة الدرجات المعيارية للمتغيرات الى محرر البيانات وذلك على النحو التالي:



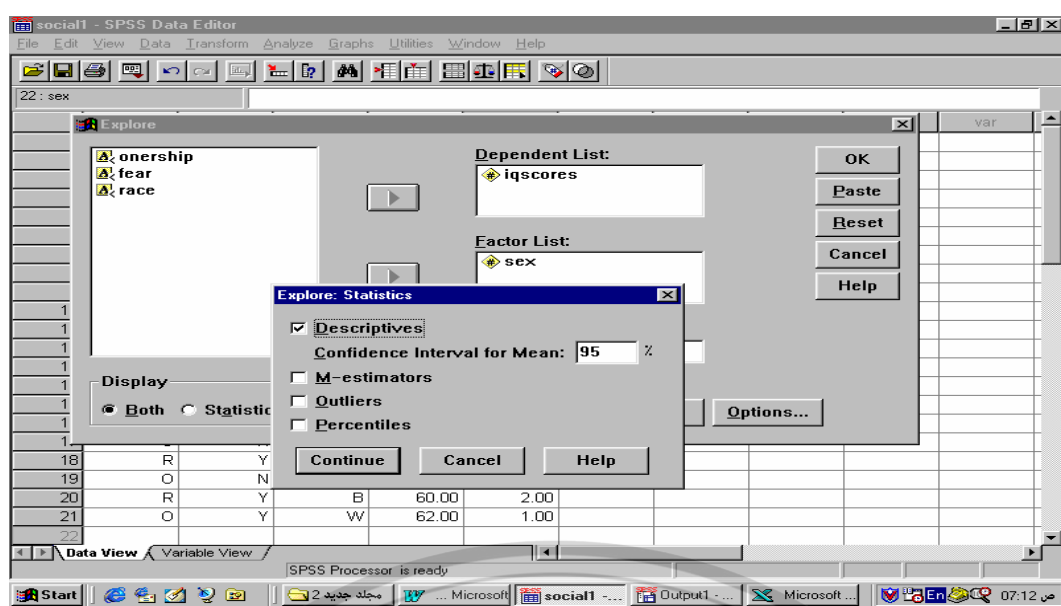
The screenshot shows the SPSS Data Editor window with the following data:

1: respondent_number	New_Salary	aveg	ZSalary	Zage	var	var
1	16200.00	401.60	.42512	-.57281		
2	13100.00	1.40	.10306	.11207		
3	14750.00	1.40	.58615	.65997		
4	21500.00	801.40	1.39129	1.75577		
5	12550.00	1.60	-.05797	-.16188		
6	16900.00	1201.60	-.54106	-.02490		
7	7890.00	301.60	-1.86150	-1.39466		
8	7160.00	2.00	-1.63606	-1.53163		
9	13100.00	1.80	.10306	.93392		
10	13970.00	682.20	-.63768	-.29885		
11	8700.00	1.80	-1.18518	-1.39466		
12	16780.00	1002.20	-.28341	-.98373		
13	12000.00	1.40	-.21900	-.70978		
14	17590.00	1141.00	-.25121	-.43583		
15	7600.00	1.80	-1.50723	-1.39466		
16	18480.00	242.00	1.32688	1.48182		
17	19220.00	301.80	1.45570	1.07090		
18	17630.00	401.60	.84379	-.02490		
19	14200.00	2.00	.42512	.65997		
20	11736.00	1.00	-.29629	.24905		
21	14300.00	242.20	.10306	.52300		
22	18600.00	2.20	1.71335	1.48182		

٢-٨ الحصول على الدرجات المعيارية Producing Z Scores

باستخدام برنامج SPSS يمكننا حساب الدرجات المعيارية Z من أجل أي توزيع طبيعي وذلك عن طريق استخدام التعليمة Descriptive الموجودة ضمن الإحصاءات الوصفية Descriptive Statistics الذي تحصل عليه بعد النقر على الخيار تحليل "Analyze" من شريط القوائم وبعد تحديد المتغير الذي نرغب حساب الدرجات المعيارية الخاصة به وليكن درجات الذكاء IQScores التي حصلت عليها عينة من الطلاب عددها 21 طالباً ، وقبل الضغط على الزر OK نفعل الخيار الموجود أسفل مربع الخيار وهو حفظ القيم المعيارية بمثابة متغير جديد:

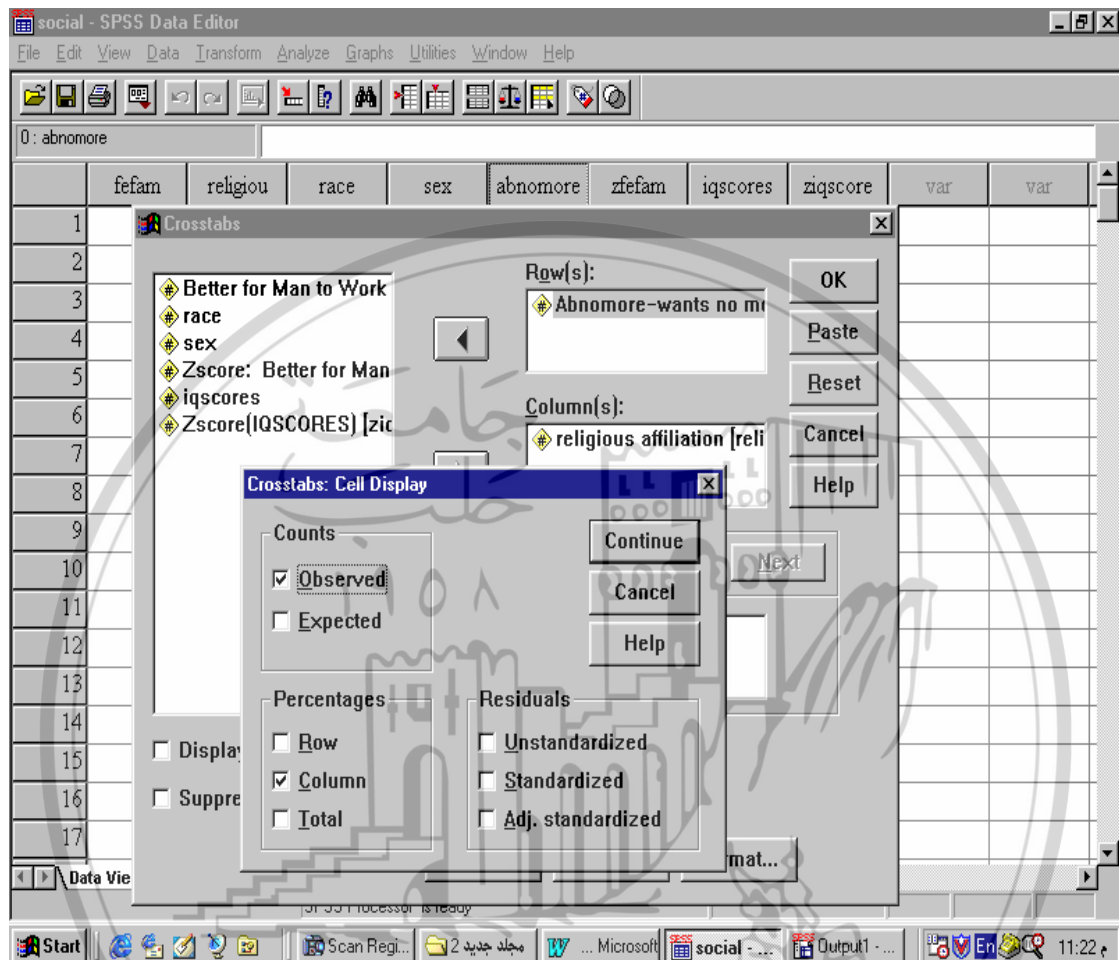
Save a Standardized values as Variables



من الجدول التالي نلاحظ ما يلي: إن الحد الأدنى لمجال الثقة لمتوسط درجات الذكاء لدى الإناث يساوي 34.3340 درجة ، بينما بلغ الحد الأعلى لمجال الثقة لمتوسط درجات الذكاء لدى الإناث 50.666 درجة وذلك عند درجة ثقة 95 % وبلغ الخطأ المعياري للمتوسط 3.7101 درجة ، ومتوسط درجات الذكاء لدى الإناث 42.5.

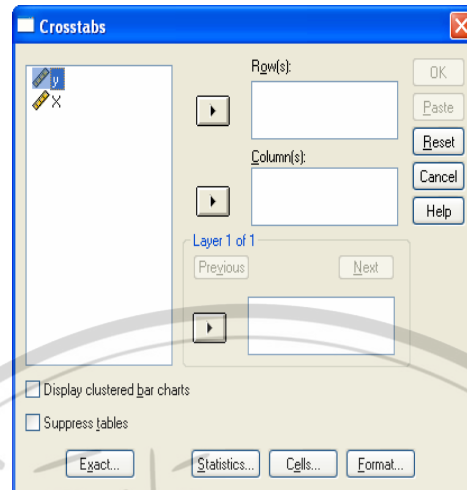


وبعد تحديد النسب Percentages في صندوق الحوار Cell Display Crosstabs : نضغط على المربع Continue فننتقل ثانية إلى صندوق الحوار Crosstabs .



وبالضغط على المربع Continue ننتقل ثانية إلى صندوق الحوار المعنون Crosstabs ، فإذا تم الضغط على المربع Ok نحصل على الجدول التقاطعي المبين أعلاه ، والذي يوضح بدقة وتفصيل العلاقة بين الانتماء الديني من جهة ووجهة النظر حول موضوع الإجهاض الشرعي بالنسبة للمرأة المتزوجة وغير الراغبة بمزيد من الأطفال .

Analysis → Descriptive Statistics → Crosstabs



مثال (٢) :

سنحاول استقصاء العلاقة بين المؤيدين والمعارضين لعملية الإجهاض الشرعي Legal abortion للمرأة المتزوجة وغير الراغبة بإنجاب المزيد من الأطفال والانتماء الديني Religious affiliation .

بشكل طبيعي يقوم برنامج SPSS بعرض عدد التكرارات بكل خلية وبإمكانك أن تطلب من البرنامج أن يعرض بالإضافة إلى عدد التكرارات النسب المئوية .

إن المتغير المستقل سيكون الانتماء الديني والذي يجب وضعه في الأعمدة ، بينما يتم عرض المتغير التابع الإجهاض الشرعي في الأسطر ، ولعرض النسب المئوية نفعل المربع خلايا Cells الموجود أسفل صندوق الحوار Crosstabs .

إن الحد الأدنى لمجال الثقة لمتوسط درجات الذكاء لدى الذكور يساوي 31.702 درجة ، بينما بلغ الحد الأعلى لمجال الثقة لمتوسط درجات الذكاء لدى الذكور 50.9647 درجة وذلك عند درجة ثقة 95 % وبلغ الخطأ المعياري للمتوسط 4.1767 درجة ، ومتوسط درجات الذكاء لدى الذكور 41.3333 .

Descriptives				Statistic	Std. Error
SEX					
IQSCORES	female	Mean		42.5000	3.7101
		95% Confidence Interval for Mean	Lower Bound	34.3340	
			Upper Bound	50.6660	
		5% Trimmed Mean		42.5556	
		Median		41.0000	
		Variance		165.182	
		Std. Deviation		12.8523	
		Minimum		22.00	
		Maximum		62.00	
		Range		40.00	
		Interquartile Range		22.0000	
		Skewness		-.077	.637
		Kurtosis		-1.097	1.232
	Male	Mean		41.3333	4.1767
		95% Confidence Interval for Mean	Lower Bound	31.7020	
			Upper Bound	50.9647	
		5% Trimmed Mean		41.2593	
		Median		44.0000	
		Variance		157.000	
		Std. Deviation		12.5300	
		Minimum		24.00	
		Maximum		60.00	
		Range		36.00	
		Interquartile Range		22.0000	
		Skewness		.105	.717
		Kurtosis		-1.284	1.400

§-٤ العلاقات بين متغيرين : الجدول التقاطعي

Relations between Two Variables: Cross-Tabulation

٤-١ إنشاء جدول بمتغيرين Producing Bivariate Tables

يتمتع برنامج SPSS بميزة إنشاء جدول تقاطعي والذي نحصل عليه بالطريقة نفسها التي حصلنا بها على جدول التوزيع التكراري ، وذلك عن طريق الخيار الجدول التقاطعي Crosstabs من الإحصاءات الوصفية Descriptive Statistics الذي تحصل عليه بعد النقر على الخيار تحليل " Analyze " من شريط القوائم ، فيظهر صندوق حوار بعنوان Crosstabs .

يتطلب إنشاء الجدول التقاطعي منا تحديد كلا المتغيرين اللذين سيتم تمثيل أحدهما بواسطة الأسطر والثاني بواسطة الأعمدة وذلك على النحو التالي:

Abnmore-wants no more childern * religious affiliation Crosstabulation						
		religious affiliation				Total
		protestant	catholic	muslem	others	
Abnmore-wai No no more childern	Count	5	23	96	30	154
	% within religio affiliation	29.4%	65.7%	66.7%	45.5%	58.8%
Yes	Count	12	12	48	36	108
	% within religio affiliation	70.6%	34.3%	33.3%	54.5%	41.2%
Total	Count	17	35	144	66	262
	% within religio affiliation	100.0%	100.0%	100.0%	100.0%	100.0%

مثال (٣) :

لتكن لدينا البيانات التالية والتي تعبر عن استطلاع للرأي يهدف لمعرفة العلاقة بين ثلاثة متغيرات وهي : الخوف أثناء المشي وحيداً في الليل ، العرق ، ملكية البيت والمطلوب : بناء جدول توزيع تكراري بمتغيرين لكل من العرق والخوف أثناء المشي وحيداً في الليل والذي يعتبر المتغير المستقل بهذه الحالة ؟

والمطلوب :

- حساب النسب المئوية للجدول مع بيان العلاقة ما بين العرق والخوف ليلاً ؟
- استخدم هذه البيانات لبناء جدول بمتغيرين يمكننا من إجراء المقارنة للخائفين من المشي ليلاً بين الأشخاص المالكين للبيت والأشخاص المستأجرين^١ ؟

RESPONDENT	RACE	FEAR OF WALKING ALONE	RENT / OWN
1	B	N	O
2	B	Y	O
3	W	N	R
4	W	N	O
5	w	Y	R
6	B	N	R
7	W	N	O
8	B	N	R

^١ Race: B = black, W = white; Fear: Y = yes, N = no; R = rent, O = own

العرق : أسود ، أبيض ، الخوف : نعم ، لا ، الملكية : مستأجر ، مالك

FEAR * RACE * ONERSHIP Crosstabulation

ONERSHIP			RACE		Total
			black	white	
own	FEAR no	Count	1	5	6
		% within RAC	33.3%	62.5%	54.5%
	yes	Count	2	3	5
		% within RAC	66.7%	37.5%	45.5%
	Total	Count	3	8	11
		% within RAC	100.0%	100.0%	100.0%
R	FEAR no	Count	3	2	5
		% within RAC	42.9%	66.7%	50.0%
	yes	Count	4	1	5
		% within RAC	57.1%	33.3%	50.0%
	Total	Count	7	3	10
		% within RAC	100.0%	100.0%	100.0%

يستعمل الأمر Crosstabs في عمل جداول الاقتران المؤلفة من سطرين وعمودين 2x2 والجداول المتعددة المؤلفة من أكثر من سطرين وأكثر من عمودين بهدف حساب مؤشرات الارتباط بين المتغيرات في هذه الجداول والإحصائيات المرتبطة بها، إذ إن لهذا الأمر فائدة عملية في البحوث التطبيقية التي تسهل عملية إعداد جداول الاقتران والجداول المتعددة لأسئلة الاستبيانات الإحصائية.

مثال (٤) :

يمثل الجدول التالي عينة حجمها 29 شخصاً مريضاً خضع قسم منهم للعلاج تم إعطاءهم الرمز A والذين لم يخضعوا للعلاج تم إعطاءهم الرمز B وتم التعبير عن ذلك بوساطة المتغير treatment ، أما المتغير recover فيمثل الشفاء من المرض ولهذا تم إعطاء الرقم ١ للشفاء من المرض والرقم ٠ لعدم الشفاء من المرض ، والمتغير الثالث gender فيمثل جنس المريض حيث تم إعطاء الرمز F للأنثى والرمز M للذكر .

gender	recover	treatment
F	1	A
F	1	B
M	0	B
M	0	B

FEAR * RACE Crosstabulation

Count

		RACE		Total
		black	white	
FEAR	no	4	7	11
	yes	6	4	10
Total		10	11	21

- حساب النسب المئوية للجدول مع بيان العلاقة ما بين العرق والخوف ليلاً ؟

FEAR * RACE Crosstabulation

			RACE		Total
			black	white	
FEAR	no	Count	4	7	11
		% within RACE	40.0%	63.6%	52.4%
	yes	Count	6	4	10
		% within RACE	60.0%	36.4%	47.6%
Total		Count	10	11	21
		% within RACE	100.0%	100.0%	100.0%

- جدول بمتغيرين يوضح العلاقة بين الخائفين من المشي ليلاً وبين الأشخاص المالكين والأشخاص المستأجرين ؟

9	W	Y	O
10	B	Y	R
11	W	N	O
12	B	Y	O
13	B	Y	R
14	W	Y	O
15	W	N	R
16	B	N	R
17	W	N	O
18	B	Y	R
19	W	N	O
20	B	Y	R
21	W	Y	O

الحل :

لإنشاء جدول بمتغيرين نتبع الخطوات التالية : ننتقي الخيار الجدول التقاطعي Crosstabs من الإحصاءات الوصفية Descriptive Statistics الذي نحصل عليه بعد النقر على الخيار تحليل " Analyze " من شريط القوائم ، فيظهر صندوق حوار بعنوان Crosstabs . إن المتغير المستقل سيكون الخوف من المشي ليلاً والذي يجب وضعه في الأعمدة ، بينما يتم عرض المتغير التابع وهو العرق في الأسطر ، ولعرض النسب المئوية نفعل المربع خلايا Cells الموجود أسفل صندوق الحوار Crosstabs . وبعد تحديد النسب Percentages في صندوق الحوار Cell Display : Crosstabs نضغط على المربع Continue فننتقل ثانية إلى صندوق الحوار Crosstabs . وبالضغط على المربع Ok نحصل على الجدول التقاطعي المبين أدناه والذي يوضح بدقة العلاقة بين الخوف من جهة والعرق من جهة ثانية .

M	1	A
F	1	A
M	0	B
F	1	A
F	0	B
M	0	A
M	1	A
M	1	B
F	0	A
M	0	B
F	1	B
F	1	B
M	0	A
M	1	A
M	0	B
F	1	A
M	1	B
F	0	A
F	0	A
M	1	B
M	1	A
M	0	B
F	1	A
M	0	B
M	0	B

والمطلوب :

تكوين جدول اقتران للعلاقة بين المتغيرين خضوع المريض للمعالجة treatment والشفاء من المرض recover مع إجراء اختبار الاستقلالية بين المتغيرين أي هل هناك علاقة بين تعاطي العلاج والشفاء من المرض وذلك في حالتين : الأولى اختبار معنوية العلاقة بالنسبة لجميع المرضى ، الثانية إجراء اختبار الاستقلالية بين المتغيرين حسب متغير الجنس.

الحل:

لتكوين جدول الاقتران بين المتغيرين نتبع الخطوات التالية:

Crosstabulation خضوع المريض للمعالجة * الشفاء من المرض

Count

		الشفاء من المرض		Total
		عدم الشفاء من المرض	الشفاء من المرض	
خضوع المريض للمعالجة	خضع للمعالجة	5	9	14
لم يخضع للمعالجة	لم يخضع للمعالجة	9	6	15
Total		14	15	29

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1.710 ^b	1	.191		
Continuity Correction	.876	1	.349		
Likelihood Ratio	1.729	1	.189		
Fisher's Exact Test				.272	.175
N of Valid Cases	29				

a. Computed only for a 2x2 table

b. 0 cells (.0%) have expected count less than 5. The minimum expected count was .76.

يتم اختبار فرضية عدم القائلية باستقلالية الأسطر عن الأعمدة أي استقلالية خضوع المريض للمعالجة عن الشفاء من المرض مقابل الفرضية البديلة القائلة بوجود علاقة بين العلاج والشفاء من المرض باستخدام إحصائية الاختبار Chi-square مع:
 $(C-1)(r-1)$ درجة حرية.

إن قيمة إحصائية الاختبار لكاي تربيع تساوي 1.710 ودرجات الحرية تساوي 1 ومستوى المعنوية الدقيق يساوي 0.191 .

بما أن $P\text{-value}=0.191 > 0.05$ هذا يعني لا يمكننا رفض فرضية عدم عند مستوى معنوية 0.05 وبالتالي نقبل هذه الفرضية القائلة باستقلالية الأسطر عن الأعمدة أي استقلالية العلاج عن الشفاء من المرض.

يشتمل صندوق الحوار Crosstabs : Statistics على ما يلي:

Chi-square : تستخدم لحساب إحصائية كاي تربيع بهدف اختبار استقلالية الأسطر عن الأعمدة أو استقلالية المتغيرين خضوع المريض للمعالجة، الشفاء من المرض.

Correlations : تستخدم لحساب معامل الارتباط الخطي لبيرسون Pearson ومعامل ارتباط الرتب لسبيرمان Spearman في آن واحد ، مع العلم أنه يجب ان تكون قيم المتغيرات نوعية أو كمية ليتم حساب معامل ارتباط سبيرمان باعتبار أن كلاً من الأسطر والأعمدة عبارة عن قيم ترتيبية Ordinal values، كما يتم حساب معامل الارتباط الخطي لبيرسون باعتبار أن المتغيرات في كل من الأسطر والأعمدة هي متغيرات كمية .

Nominal : توفر هذه القائمة حساب أربع معاملات للاقتتران أو الارتباط (معامل التوافق، معامل فاي وكرامرز، معامل لميدا، ومعامل عدم التأكد) في حالة كون كل من الأسطر والأعمدة عوامل غير كمية بالإضافة إلى عدم إمكانية ترتيب البيانات مثل متغير الجنس ذكر أو أنثى أو متغير الحالة الاجتماعية (متزوج، أعزب، أرمل، مطلق) أو متغير المحافظة (دمشق، حلب ن حمص، دير الزور) .

Ordinal : توفر هذه القائمة حساب أربعة معاملات للاقتتران أو الارتباط (معامل غاما، معامل سومر، معامل كندال تو بي ، معامل كندال تو سي) في حالة كون كل من الأسطر والأعمدة متغيرات رتبية.

Nominal by Interval : تستخدم لحساب إحصائية إيتا للاقتتران بين متغيرين أحدهما المتغير التابع ويقاس ضمن مجال مثل متغير الدخل والمتغير الآخر هو المتغير المستقل له عدد محدد من الخصائص أو الفئات مثل متغير النوع gender.

نظراً لكون المتغيرات المدروسة في مثالنا هي متغيرات اسمية سنقوم بحساب إحصائية كاي تربيع مع مقياس فاي وكرامرز للاقتتران بين المتغيرات الاسمية.

بالنقر على الزر Continue في صندوق حوار Crosstabs : Statistics والنقر على الزر

OK في صندوق حوار Crosstabs نحصل على النتائج التالية:

Analysis → Descriptive Statistics → Crosstabs

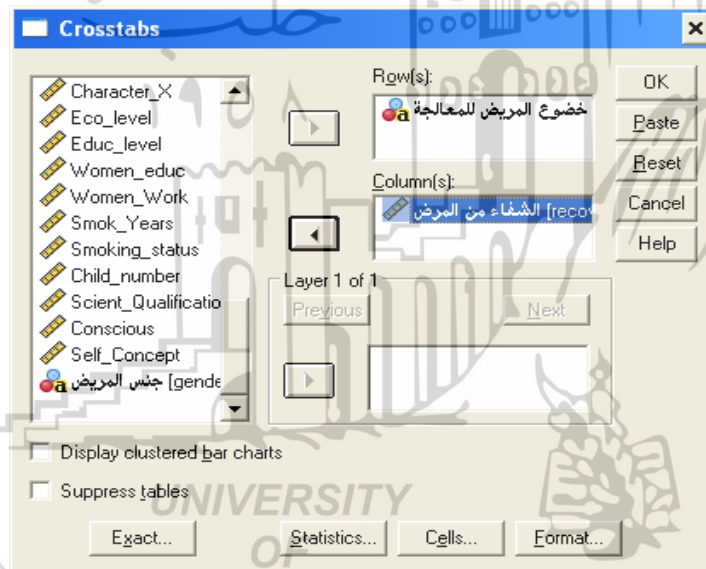
يظهر صندوق الحوار التالي المسمى Crosstabs الذي يشتمل على متغيرات ملف البيانات كافة في الطرف الأيسر من صندوق الحوار ، وفي الطرف الأيسر ننقل المتغير treatment إلى داخل النافذة Rows والمتغير recover إلى داخل النافذة Columns .
إذ أن :

Rows : يتضمن المتغير الذي نرغب أن تكون مكوناته أسطراً في جدول الاقتران.

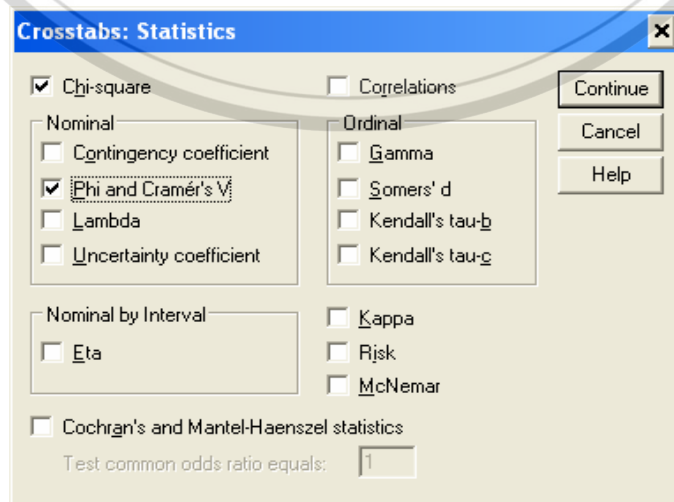
Columns : يتضمن المتغير الذي نرغب أن تكون مكوناته أعمدة في جدول الاقتران.

Display Clusterd bar Charts : تستخدم لعرض الأشرطة البيانية عندما يرغب الباحث في ذلك.

Supress tables : لإخفاء جدول الاقتران والجداول المتعددة عندما يرغب الباحث في ذلك.



بالنقر على الزر Statistics يظهر صندوق حوار بمسمى Crosstabs : Statistics



Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	-.243	.191
	Cramer's V	.243	.191
N of Valid Cases		29	

- Not assuming the null hypothesis.
- Using the asymptotic standard error assuming the null hypothesis.

أما مقياس الاقتران فاي $phi = \sqrt{\frac{\chi^2}{n}} = \sqrt{\frac{1.710}{29}} = 0.243$ يشير إلى عدم معنوية الارتباط

أو الاقتران ما بين العلاج والشفاء من المرض عند مستوى معنوية 0.05 ، حيث أن P-value=0.191 > 0.05

إن مقياس الاقتران فاي يستعمل غالباً مع جداول الاقتران التي تضم سطرين وعمودين والحالة العامة لهذا المقياس هي معامل كرامر Cramer Coefficient C الذي يتم حسابه لجداول تضم أكثر من سطرين وأكثر من عمودين بموجب العلاقة التالية:

$$Cramer\ coefficient = \sqrt{\frac{\chi^2}{n(L-1)}} = \sqrt{\frac{1.710}{29(1)}} = 0.243$$

حيث تمثل n حجم العينة ، و تمثل L العدد الأصغر بين الصفوف والأعمدة في جدول الاقتران حيث تستعمل إحصائية كاي تربيع لاختبار الاستقلالية مع (r-1)(C-1) درجة حرية مع ملاحظة أن قيمة هذا المقياس مساوية تماماً لمقياس الاقتران فاي.

ملاحظة : هناك مقياس آخر يعرف بمعامل التوافق Contingency Coefficient يتم حسابه لأي جدول توافقي يضم (rxc) سطر وعمود ويحسب بموجب العلاقة التالية:

$$Contingency\ coefficient = \sqrt{\frac{\chi^2}{\chi^2 + N}} = \sqrt{\frac{1.710}{1.71 + 29}} = 0.236$$

Chi-Square Tests

جنس المريض		Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
أنثى	Pearson Chi-Square	.188 ^b	1	.665		
	Continuity Correction	.000	1	1.000		
	Likelihood Ratio	.193	1	.661		
	Fisher's Exact Test				1.000	.594
	N of Valid Cases	12				
ذكر	Pearson Chi-Square	2.487 ^c	1	.115		
	Continuity Correction	1.127	1	.288		
	Likelihood Ratio	2.506	1	.113		
	Fisher's Exact Test				.162	.145
	N of Valid Cases	17				

a. Computed only for a 2x2 table

b. 3 cells (75.0%) have expected count less than 5. The minimum expected count is 1.3

c. 3 cells (75.0%) have expected count less than 5. The minimum expected count is 2.4

Symmetric Measures

جنس المريض			Value	Approx. Sig.
أنثى	Nominal by Nominal	Phi	.125	.665
		Cramer's V	.125	.665
	N of Valid Cases		12	
ذكر	Nominal by Nominal	Phi	-.383	.115
		Cramer's V	.383	.115
	N of Valid Cases		17	

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Crosstabs

Row(s):
 Row(s): خضوع المريض للمعالجة

Column(s):
 Column(s): الشفاء من المرض [reco]

Layer 1 of 1
 Previous Next

Display clustered bar charts ☒
 Suppress tables ☐

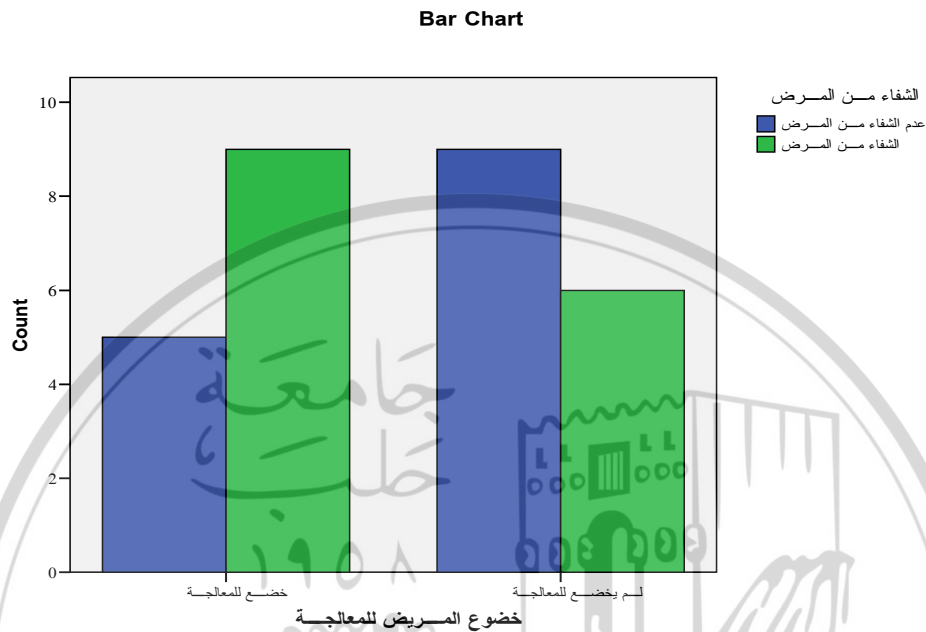
Exact... Statistics... Cells... Format...

نضغط OK نحصل على النتائج التالية:

Crosstabulation خضوع المريض للمعالجة * الشفاء من المرض * جنس المريض

Count		الشفاء من المرض		Total
جنس المريض		عدم الشفاء من المرض	الشفاء من المرض	
أنثى	خضع للمعالجة	3	5	8
	خضوع المريض للمعالجة	1	3	4
	Total	4	8	12
ذكر	خضع للمعالجة	2	4	6
	خضوع المريض للمعالجة	8	3	11
	Total	10	7	17

إن تفعيل الخيار Display Clusterd bar Charts في صندوق حوار Crosstabs يعرض الأعمدة البيانية كما في الشكل:

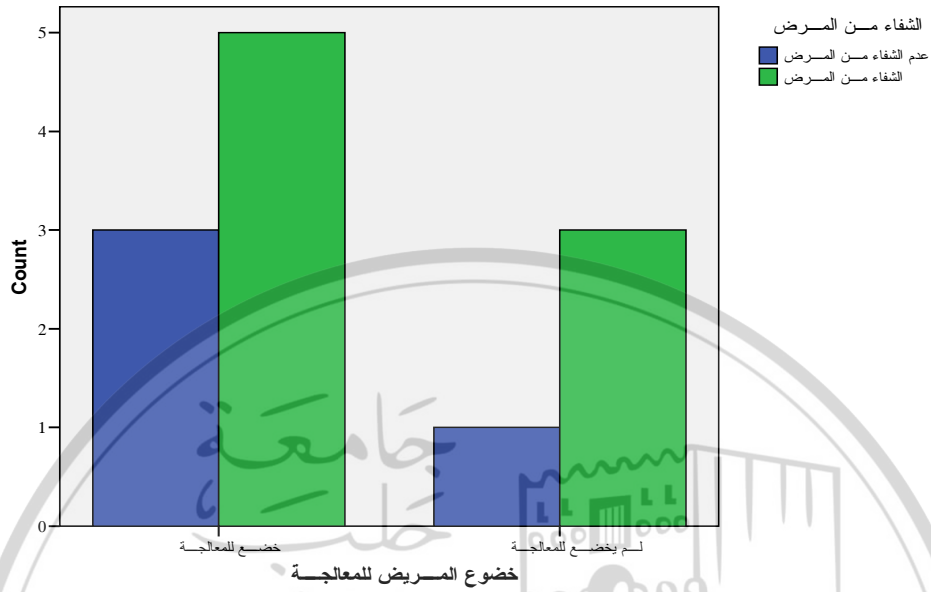


ثانياً : لإجراء اختبار الاستقلالية بين المتغيرين حسب متغير الجنس. نتبع الخطوات التالية:

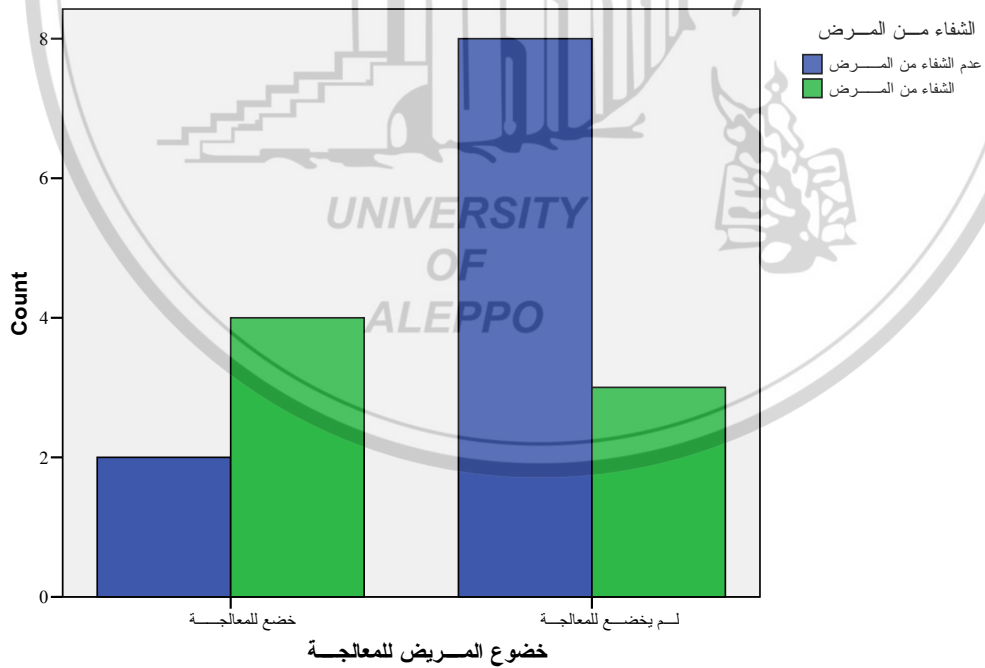
Analysis → Descriptive Statistics → Crosstabs

يظهر صندوق الحوار التالي المسمى Crosstabs الذي يشتمل على كافة متغيرات ملف البيانات في الطرف الأيسر من صندوق الحوار ، وفي الطرف الأيسر ننقل المتغير treatment إلى داخل النافذة Rows والمتغير recover إلى داخل النافذة Columns والمتغير gender إلى داخل النافذة: Layer 1 of 1.

جنس المريض=أنثى



جنس المريض=ذكر



٢-٤ معامل ارتباط فاي Phi Coefficient :

يستخدم هذا المعامل في حالة المتغيرات النوعية التي ينقسم كل منها انقساماً ثنائياً في صورة اسمية مثل (نعم ، لا) ، (صح ، خطأ) ، (٠ ، ١) ، (ناجح ، راسب) ، (ذكر ، أنثى

والمطلوب: هل يمكننا أن نستنتج وجود علاقة ارتباطية بين إجابات أفراد العينة على هذين السؤالين وذلك عند مستوى معنوية ٠,٠٥ و ٠,٠١ ؟

الحل: نحسب أولاً معامل الارتباط Φ بين المتغيرين المدروسين (الأول: أشعر بالتوتر أغلب الأوقات والثاني: أنام بسهولة في الليل) بموجب العلاقة التالية :

$$\Phi = \frac{34(14) - 16(6)}{\sqrt{(34 + 16)(6 + 14)(34 + 6)(16 + 14)}} = 0.347$$

ثانياً : نحسب قيمة إحصائية الاختبار Z بموجب العلاقة التالية :

$$Z = \Phi \cdot \sqrt{n} = 0.347 \sqrt{70} = 2.90$$

إن قيمة إحصائية الاختبار هي أكبر في الحالتين من قيمة Z الجدولية أو القيمة الحرجة والتي تساوي ٢,٥٨ عند مستوى معنوية ٠,٠١ و تساوي ١,٩٦ عند مستوى معنوية ٠,٠٥ ، وبالنتيجة يمكن القول بأن هناك علاقة ارتباطية طردية ذات دلالة إحصائية بين استجابات أفراد العينة على هذين السؤالين بمعنى أن من أجابوا بنعم على السؤال الأول يميلون إلى الإجابة بنعم على السؤال الثاني ، وقد تخالف النتائج الواقع أو المنطق فكيف لمن يشعر بالتوتر أغلب الوقت يستطيع أن ينام بسهولة .

وعلى الرغم من مخالفة النتائج لأبسط قواعد المنطق إلا أن الوقائع الأخيرة دلت بوضوح على ذلك ، إذ كانت الغالبية من أبناء الشعب العربي تنام بهدوء ولا نعلم إن كانت تعاني من وخز الضمير أو لا

ملاحظة :

على الرغم من إمكانية استخدام معامل فاي كتقدير جيد للارتباط بين متغيرات ثنائية التقسيم إلا أن هناك مأخذ على هذا المعامل تتمثل في أن قيمته لا تصل في غالب الأحيان إلى الواحد الصحيح سلباً أو إيجاباً إلا في حالة تساوي مجموع العمودين مع مجموع السطرين .

أنام بسهولة في الليل	أشعر بالتوتر أغلب الوقت	أنام بسهولة في الليل	أشعر بالتوتر أغلب الوقت
1	1	1	1
1	1	1	1
1	1	1	1
1	1	1	1
1	1	1	1
1	1	1	1
1	1	1	1

(، أو في حالة تحويل المتغيرات المستمرة الى متغيرات ثنائية ، وبما أن الهدف من دراسة هذا المعامل هو معرفة العلاقة بين هذين المتغيرين لا بد من تنظيم البيانات في جدول اقتران رباعي بطريقة يسهل التعامل معها ، فمثلاً إذا كانت الإجابة عن سؤالين مختلفين في الاختبارات السيكولوجية تتلخص بنعم أو لا ، وكان المطلوب معرفة الارتباط بين درجات الإجابة على هذين السؤالين . نضع البيانات في جدول يضم فقط سطرين وعمودين يشتملان على المشاهدات الخاصة بالمتغيرين في صورة تكرارات موزعة على الأقسام المختلفة للمتغيرين حيث a, b, c, d هي خلايا التكرارات المشاهدة في الجدول كما يلي :

السؤال الأول X	السؤال الثاني Y	
	نعم	لا
نعم	a	b
لا	c	d

إن معامل الارتباط Φ بين المتغيرين يحسب بموجب العلاقة التالية :

$$\Phi = \frac{a.d - b.c}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}$$

لمعرفة الدلالة الإحصائية لمعامل الارتباط Φ نحسب القيمة المعيارية له والتي تتبع التوزيع الطبيعي المعياري ، وذلك بموجب الصيغة التالية :

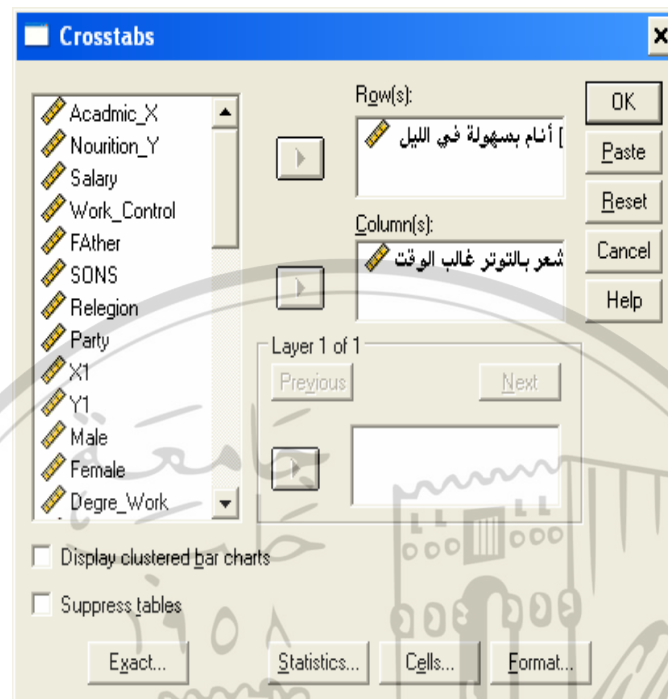
$$Z = \Phi . \sqrt{n}$$

تقارن القيمة المحسوبة لإحصائية الاختبار Z مع القيم النظرية أو القيم الحرجة المستخرجة من جدول التوزيع الطبيعي المعياري عند مستويات معنوية مختلفة وذلك طبقاً لاختبار من طرفين .

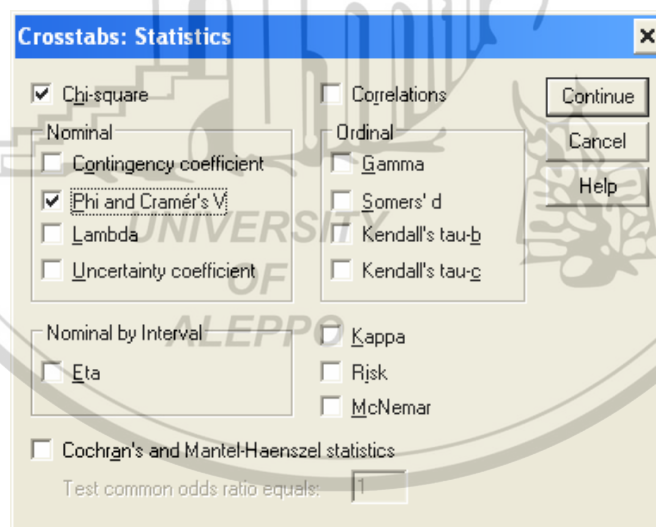
مثال (٥) :

في دراسة حديثة أجراها فريق بحث اجتماعي على طلبة كلية الاقتصاد نتيجة الأحداث الدامية في فلسطين والمجازر التي يرتكبها شارون بحق الشعب الفلسطيني البطل ، هدفها التعرف على وجود علاقة إرتباطية بين الإجابات على سؤالين (متغيرين) ، الأول : أشعر بالتوتر أغلب الأوقات والثاني : أنام بسهولة في الليل ، فجاءت البيانات نتيجة اختبارات تجريبية ، طبقت على 70 طالباً تم اختيارهم بطريقة عشوائية.

في الليل الى داخل النافذة Rows والمتغير أشعر بالتوتر غالب الوقت إلى داخل النافذة Columns .



بالنقر على الزر Statistics يظهر صندوق حوار بمسمى Crosstabs : Statistics .



بالنقر على الزر Continue في صندوق حوار Crosstabs : Statistics والنقر على الزر OK في صندوق حوار Crosstabs نحصل على النتائج التالية:

إن معامل التوافق أو التصاحب r_{Cont} بين المتغيرين يحسب بموجب العلاقة التالية :

$$r_{Cont} = \sqrt{1 - \frac{1}{C}}$$

$$c = \sum \frac{E_{ij}^2}{\sum C_i \cdot \sum R_j}$$

حيث :

E_{ij}^2 تمثل مربع تكرار كل خلية في الجدول التكراري .

$\sum C_i$ مجموع تكرارات السطر i .

$\sum R_j$ مجموع تكرارات العمود j .

ولمعرفة الدلالة الإحصائية لمعامل التوافق r_{Cont} نحسب قيمة إحصائية الاختبار كاي تربيع:

$$\chi^2 = \frac{n \cdot r_{Cont}^2}{1 - r_{Cont}^2}$$

تقارن القيمة المحسوبة لكاي تربيع مع القيم النظرية أو القيم الحرجة لكاي تربيع المستخرجة من جدول كاي تربيع عند مستويات معنوية مختلفة ودرجات حرية تساوي دوماً : (عدد الأعمدة - ١) (عدد الأسطر - ١) .

مثال (٦) :

في دراسة للتعرف على وجود علاقة بين لون العيون لدى الأبناء ولونها لدى الآباء ، حيث جاءت البيانات التي حصل عليها أحد الباحثين الاجتماعيين بنتيجة دراسة تجريبية شملت عدداً من الأسر كما هي موضحة في الجدول التالي :

الأبناء الآباء	زرقاء	خضراء	بنية	المجموع
زرقاء	6	12	12	30
خضراء	9	3	18	30
بنية	15	6	9	30
المجموع	30	21	39	90

Symmetric Measures

		Value	Approx. Sig.
Nominal by	Phi	.347	.004
Nominal	Cramer's V	.347	.004
N of Valid Cases		70	

- Not assuming the null hypothesis.
- Using the asymptotic standard error assuming the null hypothesis.

أما مقياس الاقتران فاي $phi = \sqrt{\frac{\chi^2}{n}} = \sqrt{\frac{8.423}{70}} = 0.34688$ يشير الى معنوية الارتباط أو الاقتران ما بين الشعور بالتوتر والنوم بسهولة في الليل عند مستوى معنوية 0.05 وعند مستوى معنوية 0.01، إذ أن $0.01 < 0.05 < P\text{-value} = 0.004$. إن مقياس الاقتران فاي يستعمل غالباً مع جداول الاقتران التي تضم سطرين وعمودين والحالة العامة لهذا المقياس هي معامل كرامر Cramer Coefficient الذي يتم حسابه لجداول تضم أكثر من سطرين وأكثر من عمودين بموجب العلاقة التالية:

$$Cramer\ coefficient = \sqrt{\frac{\chi^2}{n(L-1)}} = \sqrt{\frac{8.423}{70(1)}} = 0.34688$$

حيث تمثل n حجم العينة، و تمثل L العدد الاصغر بين الصفوف والأعمدة في جدول الاقتران حيث تستعمل إحصائية كاي تربيع لاختبار الاستقلالية مع $(r-1)(C-1)$ درجة حرية مع ملاحظة أن قيمة هذا المقياس مساوية تماماً لمقياس الاقتران فاي.

٣-٤ معامل التوافق (التصاحب) Contingency Coefficient

يستخدم هذا المعامل في حالة المتغيرات النوعية التي لا يمكن قياسها أو حتى مجرد تقسيمها لرتب مثل متغير الحالة الاجتماعية (متزوج ، أعزب ، مطلق ، أرمل) ، متغير الجنسية (سوري ، أمريكي ، مصري ، سعودي) ، أو متغير يمثل نوع الجريمة المرتكبة (جنس ، سرقة ، تعاطي مخدرات ، قتل) .

تتطلب هذه المتغيرات في جداول تكرارية مزدوجة تحوي صفوفاً وأعمدة ليس من الضروري أن تتساوى أعدادها .

Crosstabulation أنام بسهولة في الليل * أشعر بالتوتر غالب الوقت

Count

		أشعر بالتوتر غالب الوقت		Total
		Yes	No	
أنام بسهولة	Yes	34	16	50
في الليل	No	6	14	20
Total		40	30	70

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	8.423 ^b	1	.004		
Continuity Correction ^a	6.943	1	.008		
Likelihood Ratio	8.486	1	.004		
Fisher's Exact Test				.007	.004
Linear-by-Linear Association	8.303	1	.004		
N of Valid Cases	70				

a. Computed only for a 2x2 table

b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 8.57.

يتم اختبار فرضية عدم القائلة باستقلالية الأسطر عن الأعمدة أي استقلالية الشعور بالتوتر والنوم بسهولة مقابل الفرضية البديلة القائلة بوجود علاقة بين الشعور بالتوتر والنوم بسهولة باستخدام إحصائية الاختبار Chi-square مع $(r-1)(C-1)$ درجة حرية. إن قيمة إحصائية الاختبار لكاي تربيع تساوي 1.710 ودرجات الحرية تساوي 1 ومستوى المعنوية الدقيق يساوي 0.191 .

بما أن $P\text{-value}=0.004 < 0.05$ هذا يعني رفض فرضية عدم عند مستوى معنوية 0.05 وعند مستوى معنوية 0.01 وبالتالي نقبل الفرضية البديلة القائلة بعدم استقلالية الأسطر عن الأعمدة أي عدم استقلالية الشعور بالتوتر والنوم بسهولة في الليل.

الحل :

لحساب معامل التوافق أو التصاحب r_{Cont} بين المتغيرين ، لون العيون لدى الأبناء ولونها لدى الآباء ، لابد من حساب C أولاً بموجب العلاقة التالية :

$$c = \sum \frac{E_{ij}^2}{\sum C_i \cdot \sum R_j} = 1.149231$$

حصلنا على حاصل قسمة مربع كل خلية على جداء مجموع سطرها في مجموع عمودها كما هو وارد في الجدول التالي :

الأبناء	الأباء	زرقاء	خضراء	بنية	المجموع
زرقاء	0.04	0.228571	0.12307	0.39165	
خضراء	0.09	0.014286	0.27692	0.38121	
بنية	0.25	0.057143	0.06923	0.37637	
المجموع	0.38	0.3	0.46923	1.14923	

فمثلاً ناتج الخلية الأولى الموجودة في السطر الأول والعمود الأول حصلنا عليه كما يلي :

$$(6)^2 / 30(30) = 36 / 900 = 0.04$$

نحسب معامل التوافق بموجب العلاقة :

$$r_{Cont} = \sqrt{1 - \frac{1}{C}} = \sqrt{1 - 1/1.14923} \\ = \sqrt{0.129853} = 0.36035$$

ولمعرفة الدلالة الإحصائية لمعامل التوافق أو التصاحب r_{Cont} نحسب قيمة إحصائية الاختبار كاي تربيع من العلاقة :

$$\chi^2 = \frac{n \cdot r_{Cont}^2}{1 - r_{Cont}^2} = \frac{90(0.1296)}{1 - 0.1296} = 13.4$$

وتقارن القيمة المحسوبة لكاي تربيع مع القيمة النظرية 9.49 المستخرجة من جدول كاي تربيع عند مستوى معنوية 5 % و 4 درجات حرية ، أو مع القيمة النظرية 13.28 المستخرجة من جدول كاي تربيع عند مستوى معنوية 1 % و 4 درجات حرية ، نجد في الحالتين أن قيمة كاي

إن قيمة احصائية الاختبار لكاي تربيع تساوي 14.431 ودرجات الحرية تساوي 4 ومستوى المعنوية الدقيق يساوي 0.009 .

بما أن $P\text{-value}=0.009 < 0.05$ هذا يعني رفض فرضية العدم عند مستوى معنوية 0.05 وعند مستوى معنوية 0.01 وبالتالي نقبل الفرضية البديلة القائلة بعدم استقلالية لون العيون للأباء وللأبناء.

أما معامل التوافق فيساوي 0.36 ومستوى المعنوية الدقيق 0.009 يشير الى معنوية الارتباط أو الاقتران ما بين لون العيون لدى الأبناء ولونها لدى الآباء ونستنتج بأن هناك علاقة بين الخصائص الوراثية للون العيون لدى الأبناء ولونها لدى الآباء

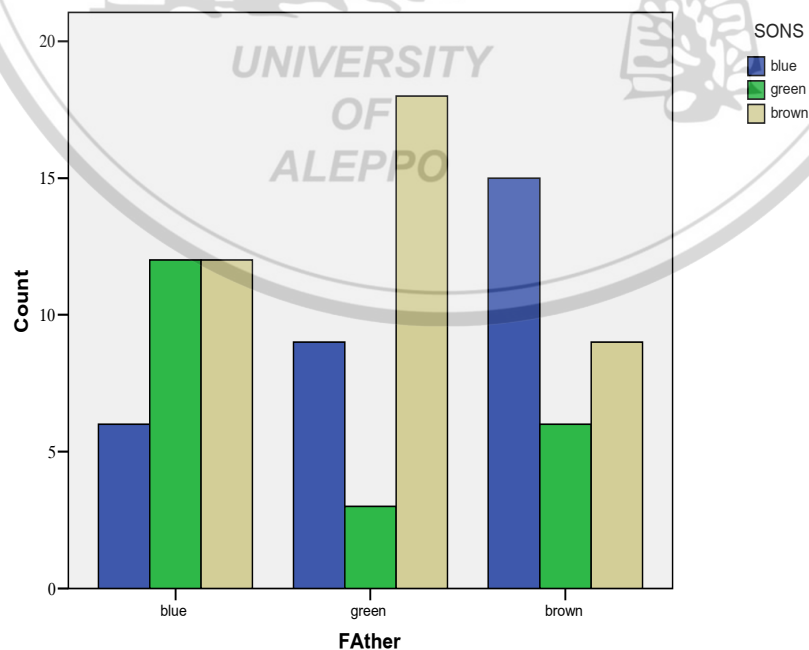
Symmetric Measures

	Value	Approx. Sig.
Nominal by Nominal - Contingency Coefficient	.360	.009
N of Valid Cases	90	

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Bar Chart



Crosstabs: Statistics

☒ **Chi-square**

☐ Correlations

☐ Continuity

☐ Gamma

☐ Somers' d

☐ Kendall's tau-b

☐ Kendall's tau-c

☐ Kappa

☐ RISK

☐ McNemar

☐ Cramer's V

☐ Lambda

☐ Uncertainty coefficient

☐ Eta

☐ Cochran's and Mantel-Haenszel statistics

Test common odds ratio equals: 1

Continue

Cancel

Help

بالنقر على الزر Continue في صندوق حوار Crosstabs : Statistics والنقر على الزر OK في صندوق حوار Crosstabs نحصل على النتائج التالية:

FATHER * SONS Crosstabulation

Count		SONS			Total
		blue	green	brown	
FATHER	blue	6	12	12	30
	green	9	3	18	30
	brown	15	6	9	30
Total		30	21	39	90

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	13.431 ^a	4	.009
Likelihood Ratio	13.315	4	.010
Linear-by-Linear Association	3.137	1	.077
N of Valid Cases	90		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 7.00.

يتم اختبار فرضية عدم القائلة باستقلالية لون العيون للأباء عن لون العيون للأبناء مقابل الفرضية البديلة القائلة بوجود علاقة بين لون العيون للأباء وللأبناء باستخدام إحصائية الاختبار Chi-square مع $(r-1)(C-1)$ درجة حرية.

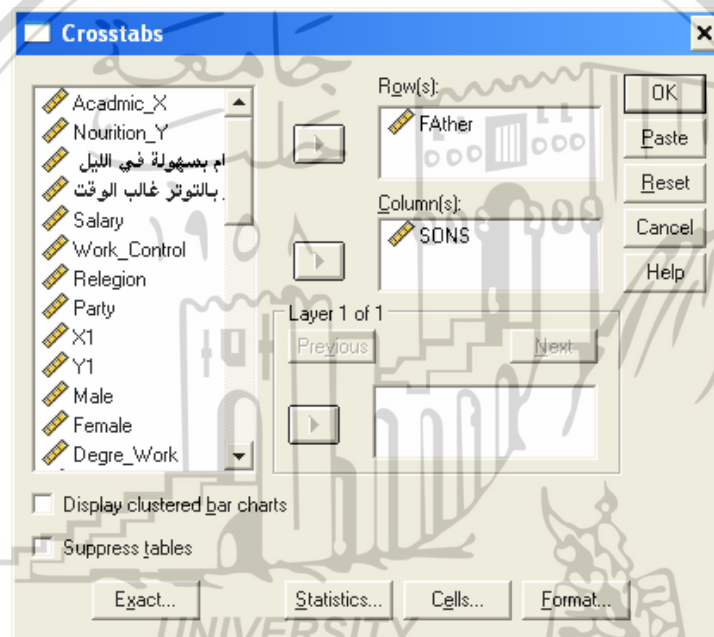
تربيع المحسوبة أكبر من القيم الحرجة وبالتالي نرفض فرض العدم ونستنتج بأن هناك علاقة بين الخصائص الوراثية للون العيون لدى الأبناء ولونها لدى الآباء .

الحل باستخدام برنامج SPSS

لتكوين جدول الاقتران او التوافق بين المتغيرين نتبع الخطوات التالية:

Analysis → Descriptive Statistics → Crosstabs

يظهر صندوق الحوار التالي المسمى Crosstabs ننقل المتغير الآباء إلى داخل النافذة المعنونة الأسطر Rows وننقل المتغير الأبناء إلى داخل النافذة المعنونة Columns كما هو موضح بصندوق الحوار التالي:



بالنقر على الزر Statistics يظهر صندوق حوار بمسمى Statistics : Crosstabs.

٤-٤ معامل الاقتران الرباعي Coefficient of Association

يستخدم هذا المعامل في حالة المتغيرات النوعية حيث يقتصر معيارا التصنيف على خاصيتين أو مستويين فقط مثل (ملتزم ، غير ملتزم) ، (حضر ، لم يحضر) بحيث توضع البيانات في جدول يضم فقط سطرين وعمودين ، يطلق على هذا الجدول جدول التوافق 2×2 .

فإذا كان لدينا متغيران نوعيان X و Y يضم كل منهما صفتين فقط ، فإن جدول التوافق 2×2 يشتمل على المشاهدات الخاصة بالمتغيرين في صورة تكرارات موزعة على الأقسام المختلفة للمتغيرين حيث a, b, c, d هي خلايا التكرارات المشاهدة في الجدول كما يلي :

المتغير الأول X	المتغير الثاني Y	
	1	2
1	a	b
2	c	d

إن معامل الاقتران الرباعي بين المتغيرين X و Y يحسب بموجب العلاقة التالية :

$$CA = \frac{a \cdot d - b \cdot c}{a \cdot d + b \cdot c}$$

لمعرفة الدلالة الإحصائية لمعامل الاقتران الرباعي نحسب القيمة المعيارية له والتي تتبع التوزيع الطبيعي المعياري ، وذلك بموجب الصيغة التالية :

$$Z = \frac{2CA}{1 - (CA)^2} \cdot \frac{1}{\sqrt{1/a + 1/b + 1/c + 1/d}}$$

وتقارن القيمة المحسوبة لإحصائية الاختبار Z مع القيم النظرية أو القيم الحرجة المستخرجة من جدول التوزيع الطبيعي المعياري عند مستويات معنوية مختلفة وذلك طبقاً لاختبار من طرفين .

مثال (٧)

في دراسة حديثة أجراها فريق بحث اجتماعي هدفها التعرف على أهمية الحالة الغذائية للطفل ومدى تأثيرها على الأداء الأكاديمي له ، تم اختيار 74 طفلاً من عدد من المدارس الابتدائية وبشكل عشوائي، فجاءت البيانات نتيجة اختبارات تجريبية قام بها الفريق الباحث

في دراسة أجراها فريق بحث اجتماعي اقتصادي هدفها التعرف على أهمية زيادة الرواتب والأجور على الانضباط في العمل في مؤسسات وشركات القطاع العام في سورية ، جاءت البيانات نتيجة اختبار تجريبي (طبق على عينة عشوائية تضم ٧٠ موظفاً في عدد من القطاعات) كما هي موضحة في الجدول التالي :

الأجور والرواتب الانضباط في العمل	زيادة	عدم زيادة
أفضل	34	16
كما هو	6	14

والمطلوب :

هل يمكننا أن نستنتج وجود علاقة بين زيادة الرواتب والأجور والانضباط في العمل، أي حساب معامل الاقتران الرباعي ودلالته الإحصائية عند مستوى دلالة 0.05 و 0.01 ؟

الحل :

أولاً : نحسب معامل الاقتران الرباعي بين المتغيرين المدروسين (زيادة الرواتب والأجور والانضباط في العمل) بموجب العلاقة التالية :

$$CA = \frac{a.d - b.c}{a.d + b.c} = \frac{34(14) - 16(6)}{34(14) + 16(6)} = 0.664$$

ثانياً : نحسب قيمة إحصائية الاختبار Z بموجب العلاقة التالية :

$$Z = \frac{2(0.664)}{1 - (0.664)^2} \cdot \frac{1}{\sqrt{1/34 + 1/16 + 1/6 + 1/14}}$$

$$= \frac{1.328}{0.559} \cdot \frac{1}{\sqrt{0.33}} = 2.376(1.74) = 4.14$$

إن قيمة إحصائية الاختبار المحسوبة هي أكبر في الحالتين من قيمة Z الجدولية أو الحرجة والتي تساوي 2.58 عند مستوى معنوية 0.01 وتساوي 1.96 عند مستوى معنوية 0.05 ، أي نرفض فرضية العدم ، وبالنتيجة يمكن القول إن هناك علاقة ارتباطية ذات دلالة إحصائية بين تحسن الرواتب والأجور والانضباط في العمل .

Good	Good	Bad	Good	Bad	Bad
Good	Good	Bad	Good	Bad	Bad
Good	Good	Bad	Good	Bad	Bad
Good	Good	Bad	Good	Bad	Bad
Bad	Bad	Bad	Bad		

الأداء الأكاديمي للطفل X	الحالة الغذائية للطفل Y	
	جيدة	ضعيفة
جيد	32	18
ضعيف	4	20

والمطلوب : هل يمكننا أن نستنتج وجود علاقة بين الحالة الغذائية للطفل والأداء الأكاديمي له ، أي حساب معامل الاقتران الرباعي ودلالته الإحصائية عند مستوى معنوية 0.05 و 0.01 ؟

الحل :

نحسب أولاً معامل الاقتران الرباعي بين المتغيرين المدروسين (الحالة الغذائية للطفل والأداء الأكاديمي له) بموجب العلاقة التالية :

$$CA = \frac{a.d - b.c}{a.d + b.c} = \frac{32(20) - 18(4)}{32(20) + 18(4)} = 0.798$$

ثانياً : نحسب قيمة إحصائية الاختبار Z بموجب العلاقة التالية :

$$Z = \frac{2(0.798)}{1 - (0.798)^2} \cdot \frac{1}{\sqrt{1/32 + 1/18 + 1/4 + 1/20}}$$

$$= \frac{1.596}{0.363196} \cdot \frac{1}{\sqrt{0.386806}} = 4.4(1.60788) = 7.1$$

إن قيمة إحصائية الاختبار المحسوبة هي أكبر في الحالتين من قيمة Z الجدولية 2.58 عند مستوى معنوية 0.01 ، و 1.96 عند مستوى معنوية 0.05 ، أي نرفض فرضية العدم ، وبالنتيجة يمكن القول إن هناك علاقة ارتباطية ذات دلالة إحصائية بين تحسن الحالة الغذائية للطفل والأداء الأكاديمي له .

مثال (٨) :

(بالتعاون مع إدارة المدرسة وأهالي الأطفال الذين وقع عليهم الاختيار) كما هي موضحة في الجدول التالي :

الأداء الأكاديمي للطفل X	الحالة الغذائية للطفل Y	الأداء الأكاديمي للطفل X	الحالة الغذائية للطفل Y	الأداء الأكاديمي للطفل X	الحالة الغذائية للطفل Y
Good	Good	Good	Good	Bad	Good
Good	Good	Good	Good	Bad	Good
Good	Good	Good	Good	Good	Bad
Good	Good	Good	Good	Good	Bad
Good	Good	Good	Good	Good	Bad
Good	Good	Good	Good	Good	Bad
Good	Good	Good	Good	Bad	Bad
Good	Good	Good	Good	Bad	Bad
Good	Good	Bad	Good	Bad	Bad
Good	Good	Bad	Good	Bad	Bad
Good	Good	Bad	Good	Bad	Bad
Good	Good	Bad	Good	Bad	Bad
Good	Good	Bad	Good	Bad	Bad
Good	Good	Bad	Good	Bad	Bad
Good	Good	Bad	Good	Bad	Bad
Good	Good	Bad	Good	Bad	Bad
Good	Good	Bad	Good	Bad	Bad
Good	Good	Bad	Good	Bad	Bad
Good	Good	Bad	Good	Bad	Bad
Good	Good	Bad	Good	Bad	Bad

أي بازدياد الرواتب والأجور سيتنامى الإحساس بالمسؤولية أكثر أو سيزداد الإحساس بالخوف من عدم الالتزام وما سيترتب عليه من مساءلة وعقوبات قد تكون مادية ولها تأثير سلبي على التعويضات وبالتالي على مستوى معيشة للعامل .

§-٥ اختبار كاي تربيع Chi-Square Test

باستخدام برنامج SPSS يمكننا إجراء اختبار كاي تربيع لتحليل البيانات الاسمية واختبار مدى تطابق التوزيع المتوقع مع التوزيع الحقيقي . وللقيام بذلك واستخدام البيانات الاسمية المرتبة في الجدول التقاطعي في عمليات التحليل الإحصائي، لا بد من إعادة تشكيل البيانات الموجودة في الجدول التقاطعي بشكل مناسب ومن ثم إعلام البرنامج بأن المدخلات في المتغير frequency هي تكرارات وليس درجات .

وللتوضيح نورد المثال التالي الذي يلقي الضوء على مشكلة وجود أو عدم وجود فروق في وجهات النظر بين الجنسين الذكور والإناث بمسألة متابعة الدراسات العليا للإناث ، في بحث اجتماعي تم سؤال 30 طالباً و 30 طالبة السؤال التالي : هل ينبغي أن تواصل الخريجة دراساتها العليا للحصول على درجة الماجستير والدكتوراه ؟ والإجابة عن هذا السؤال تتضمن 3 اختيارات وهي نعم ، لا ، محايد . يهدف الباحث الاجتماعي إلى التأكد من وجود فرق بالرأي يرجع إلى الجنس . والجدول التالي يوضح وجهات النظر بين الجنسين حول مسألة متابعة الدراسات العليا

المجموع	النساء	الرجال	الاستجابات
56	36	20	نعم
30	5	25	لا
64	34	30	محايد
150	75	75	المجموع

الحل :

سنقوم بإعادة تشكيل البيانات الموجودة في الجدول التقاطعي على الشكل الموضح بالجدول التالي:

الجنس	الاستجابات	التكرار
1	1	20

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	18.155 ^a	2	.000
Likelihood Ratio	19.441	2	.000
Linear-by-Linear Association	1.196	1	.274
N of Valid Cases	150		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 15.00.

المهم في هذه النتائج هو الصف المسمى Pearson Chi-Square ، إذ تشير هذه النتائج إلى وجود علاقة ذات دلالة إحصائية بين متغير الجنس ومتغير الاستجابات إذ أن قيمة كاي تربيع المحسوبة تساوي 18.155 وهي معنوية عند مستوى دلالة أقل من 0.0001 لأن قيمة Asymp.Sig تساوي الصفر، وبناء عليه فإن فرضية العدم مرفوضة ، أي لا توجد علاقة ذات دلالة إحصائية بين متغير الجنس ومتغير الاستجابات .

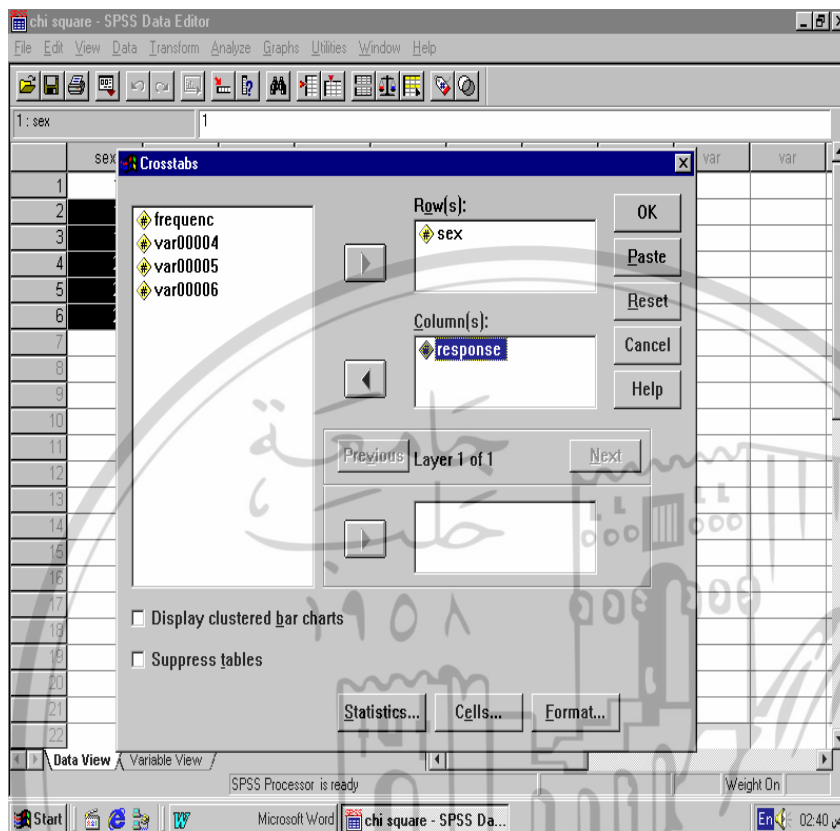
١-٥ اختبار مدى تطابق التوزيع المتوقع مع التوزيع الحقيقي

مثال : لتحقيق رغبات الطلاب بمدرسة إعدادية أثناء حصص التربية الرياضية ، اختيرت عينة عشوائية من الطلاب وطلب إلى كل منهم تحديد نوع اللعبة المفضلة أكثر من غيرها ، وظهرت النتائج كما يلي :

نوع اللعبة المفضلة	عدد الطلاب (التكرار)
السلة	35
الجمباز	20
تنس	40
الطائرة	25

والمطلوب:

هل يستطيع المشرف الرياضي في المدرسة أن يعتمد هذه النتائج فيقدم التنس والسلة أكثر مما يقدم لعبتي الجمباز والطائرة وذلك عند مستوى معنوية 0.05 ؟



SEX * RESPONSE Crosstabulation

			RESPONSE			Total
			1.00	2.00	3.00	
SEX	1.00	Count	20	25	30	75
		Expected Count	28.0	15.0	32.0	75.0
	2.00	Count	36	5	34	75
		Expected Count	28.0	15.0	32.0	75.0
Total	Count		56	30	64	150
	Expected Count		56.0	30.0	64.0	150.0

25	2	1
30	3	1
36	1	2
5	2	2
34	3	2

إذ تم استخدام المتغيرات التالية :

- الجنس وله مستويان : ذكر = 1 ، أنثى = 2 .
- الاستجابات ولها ثلاثة مستويات : نعم = 1 ، لا = 2 ، محايد = 3 .
- التكرار أو العدد .

من القائمة Data نختار Weight cases فيظهر صندوق بالعنوان نفسه نختار منها Weight cases by ومن قائمة المتغيرات في الجهة اليسرى يتم النقر على المتغير frequency ومن ثم ننقر على الخيار OK ليتم تنفيذ الإجراء ، وبهذا يمكننا استخدام البيانات السابقة في عمليات التحليل الإحصائي ، أي تطبيق اختبار كاي تربيع للتحقق من صحة الفرضية الأولية القائلة بعدم وجود علاقة ذات دلالة إحصائية بين متغير الجنس ومتغير الاستجابات .

- من القائمة Analysis نختار Descriptive Statistics فتظهر قائمة أوامر فرعية ، نختار منها Crosstabs فيظهر صندوق حوار بالعنوان نفسه.

- من قائمة المتغيرات في الجهة اليسرى ننقر المتغير Sex لينتقل إلى داخل المستطيل Rows ، ومن قائمة المتغيرات أيضاً في الجهة اليسرى ننقر المتغير Responses لينتقل إلى داخل المستطيل Columns .

- وفي صندوق الحوار ننقر الزر Statistics فيظهر صندوق حوار جديد بنفس الاسم نختار منها Chi-Square ومن ثم ننقر الزر Continue للعودة إلى صندوق الحوار Crosstabs .

- ننقر على الزر Cells فيظهر صندوق حوار باسم Crosstabs : Cell Display .

- من المستطيل Counts ننقر مربع الخيار Expected ، ومن ثم نضغط Continue وبعدها مباشرة الزر OK فنحصل على النتائج كما هي موضحة أدناه .

الحل:

إن التوزيع المتوقع لعدد الألعاب المفضلة هو أن يكون منتظماً ، أي 30 طالباً لكل نوع من أنواع اللعب ، إلا أن التوزيع المشاهد أو التجريبي (التوزيع الحقيقي) جاء غير منتظم .
الفرضية الأولية (العدم) : لا توجد فروق ذات دلالة إحصائية بين تفضيلات الطلبة لنوع الألعاب التي يفضلونها ، مقابل الفرضية البديلة القائلة بوجود فروق ذات دلالة إحصائية بين تفضيلات الطلبة لنوع الألعاب التي يفضلونها ، وذلك عند مستوى معنوية 0.05 ؟ ولاختبار صحة الفرضية الأولية نطبق اختبار كاي تربيع ، بعد إدخال البيانات في محرر بيانات برنامج SPSS واستخدام المتغيرات التالية :

السلة = 1 ، الجمباز = 2 ، تنس = 3 ، الطائرة = 4

فحصل على الجدول التالي الذي يضم متغيرين :

نوع اللعبة المفضلة Preference Game	التكرار (frequency)
1	35
2	20
3	40
4	25

- سنقوم بإعلام البرنامج SPSS بأن المدخلات في المتغير frequency هي تكرارات وذلك لترجيح كل حالة بالتكرار المقابل لها ، وذلك من خلال القائمة Data نختار Weight cases فيظهر صندوق بنفس العنوان نختار منها Weight cases by ومن قائمة المتغيرات في الجهة اليسرى يتم النقر على المتغير frequency ومن ثم ننقر على الخيار OK ليتم تنفيذ الإجراء ، وبهذا يمكننا استخدام البيانات السابقة في عمليات التحليل الإحصائي ، أي تطبيق اختبار كاي تربيع للتحقق من صحة الفرضية الأولية (العدم) وذلك باتباع الخطوات التالية :

- من القائمة Analysis نختار Nonparametric Tests فتظهر قائمة أوامر فرعية ، نختار منها Chi-Square فيظهر صندوق حوار بالعنوان نفسه على

الفصل الثالث

مقارنة المتوسطات

Compare Means

§-1 مقارنة المتوسطات Compare Means

1-1 تحليل المتوسطات Means Analysis

يفيد هذا التحليل في حساب ومقارنة الأوساط الحسابية للمجاميع الجزئية Subgroups لمتغير معين بالإضافة إلى بعض المؤشرات الإحصائية الأخرى مثل المجموع عدد الحالات للمجموعة الجزئية ، الانحراف المعياري للمجموعة الجزئية ، الوسيط ، المنوال ، الخطأ المعياري للمتوسطات..... الخ بالإضافة إلى إجراء تحليل التباين أحادي التصنيف – One Way Analysis of Variance واختبار التأثير الخطي للمعاملات Test of linearity واختبار الاقتران Eta .

مثال (١)

لنفترض لدينا المتغير الذي يمثل درجات مقرر الرياضيات المأخوذ من ملف البيانات hsb2 الذي يشتمل على 200 مشاهدة تخص عينة عشوائية لطلاب المدرسة العليا تتضمن المعلومات الشخصية عن الطلبة مثل نوعهم (gender) ، الحالة الاقتصادية الاجتماعية (ses) socio-economic status الخلفية العرقية (race) ethnic background ، كما تتضمن عدداً من درجات الاختبارات النمطية مثل اختبار القراءة (read) ، الكتابة (write) ، الرياضيات (math) ، والدراسات الاجتماعية (socst) ، نوع البرنامج المتبع (Type of Program (prog) ، نوع المدرسة خاصة أو عامة (Type of school (schtyp) . ونرغب في حساب متوسط درجات مقرر الرياضيات حسب المجاميع الفرعية ، general , vocation , academic وعدد الحالات لكل مجموعة والانحراف المعياري والخطأ المعياري.

id	gender	schtyp	math
1	male	general	41
2	female	vocation	53

Test Statistics

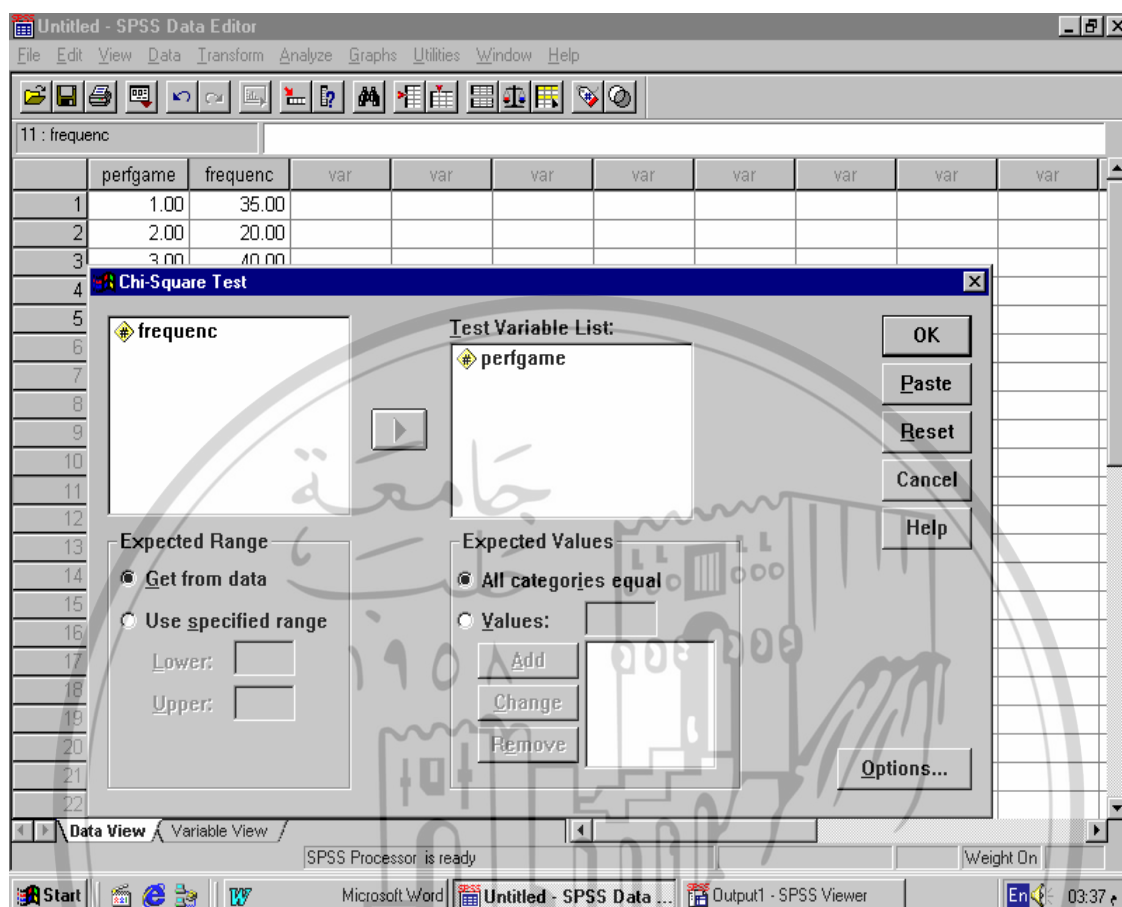
	PERFGAME
Chi-Square ^a	8.333
df	3
Asymp. Sig.	.040

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 30.0.

المهم في هذه النتائج هو أن قيمة إحصائية الاختبار كاي تربيع 8.33 هي معنوية عند مستوى دلالة أصغر من $\alpha < 0.05$ ، لأن قيمة Asymp.Sig تساوي 0.04 ، وبناء عليه فإن فرضية العدم مرفوضة ، أي إن هناك فروقاً ذات دلالة إحصائية بين تفضيلات الطلبة لنوع الألعاب ، وذلك عند مستوى معنوية 0.05 .



الشكل التالي :



- من قائمة المتغيرات في الجهة اليسرى ننقر المتغير PrefGame لينتقل إلى داخل المستطيل Test Variable List ، ننقر الزر OK فنحصل على النتائج المطلوبة كما هو موضح في الجدول التالي :

PERFGAME

	Observed N	Expected N	Residual
1.00	35	30.0	5.0
2.00	20	30.0	-10.0
3.00	40	30.0	10.0
4.00	25	30.0	-5.0
Total	120		

3	male	general	54
4	male	vocation	47
5	male	academic	57
6	male	academic	51
7	male	general	42
8	male	academic	45
9	male	general	54
10	male	academic	52
11	male	vocation	51
12	male	academic	51
13	male	academic	71
14	male	academic	57
15	male	academic	50
16	male	general	43
17	male	academic	51
18	male	general	60
19	male	academic	62
20	male	general	57
21	male	general	35
22	male	vocation	75
23	male	academic	45
24	male	academic	41
.....			
.....			
200			

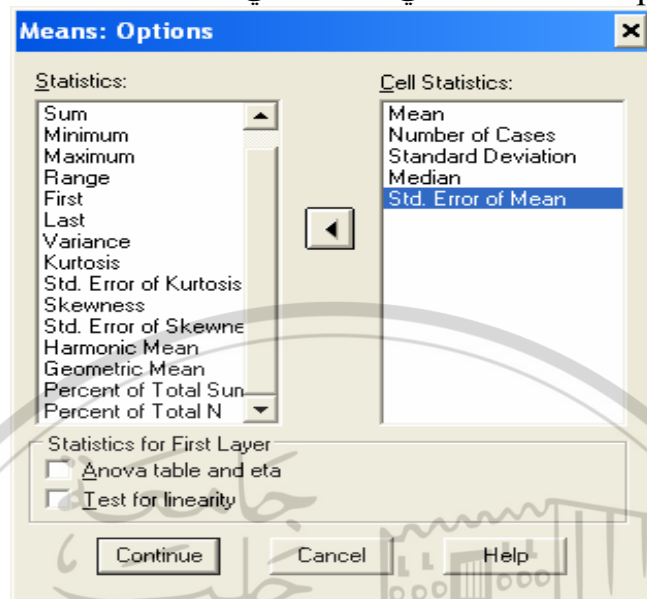
الحل:

لتطبيق أمر مقارنة المتوسطات على هذا المتغير نتبع الخطوات التالية:
من شريط القوائم نختار :

Analysis → Compare Means → Means

يظهر صندوق حوار يسمى Means على النحو التالي:

بالضغط على الزر خيارات Options في صندوق حوار Means نحصل على صندوق حوار جديد يسمى Means: Options كما في الشكل التالي:



بالنقر على الزر Continue ثم OK تظهر النتائج التالية:

math score * type of program

math score

type of program	Mean	N	Std. Deviation	Median	Std. Error of Mean
general	50.0222	45	7.44217	49.0000	1.10941
academic	56.7333	105	8.73022	57.0000	.85198
vocation	46.4200	50	7.95418	45.0000	1.12489
Total	52.6450	200	9.36845	52.0000	.66245

math score * gender

math score

gender	Mean	N	Std. Deviation	Median	Std. Error of Mean
male	52.9451	91	9.66478	52.0000	1.01314
female	52.3945	109	9.15102	53.0000	.87651
Total	52.6450	200	9.36845	52.0000	.66245

من الجدول نلاحظ تقارب الوسط الحسابي لدرجات الذكور مع الوسط الحسابي لدرجات الإناث مع أفضلية بسيطة للذكور، مع العلم بأن وسيط الدرجات في مقرر الرياضيات للإناث هو أعلى منه بالنسبة للذكور.

Case Processing Summary

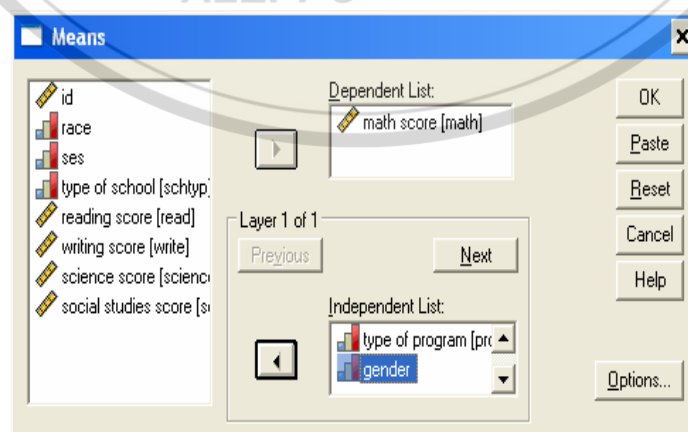
	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
math score * type of program	200	100.0%	0	.0%	200	100.0%

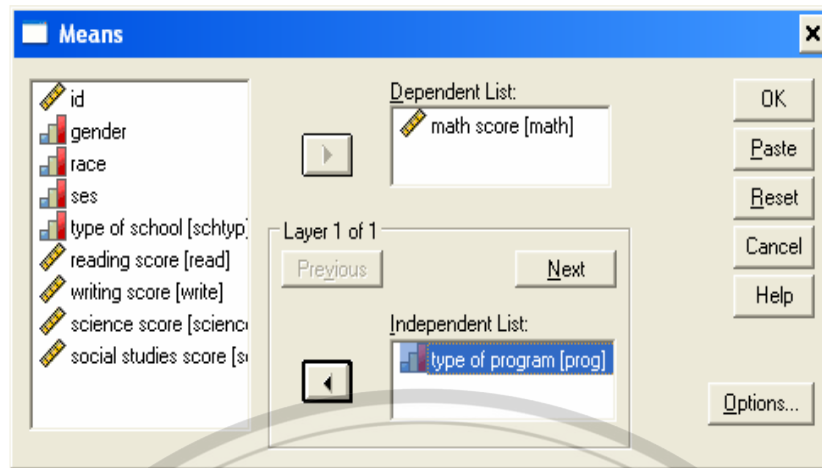
Report

math score					
type of program	Mean	N	Std. Deviation	Median	Std. Error of Mean
general	50.0222	45	7.44217	49.0000	1.10941
academic	56.7333	105	8.73022	57.0000	.85198
vocation	46.4200	50	7.95418	45.0000	1.12489
Total	52.6450	200	9.36845	52.0000	.66245

نلاحظ من الجدول أن الوسط الحسابي لدرجات الطلبة الذين نوع برنامجهم أكاديمي في مقرر الرياضيات هو أعلى من متوسط الدرجات للطلبة الذين يتبعون برامج تعليمية من النوع العام أو من النوع المهني، بشكل عام لقد تم حساب المؤشرات الإحصائية المطلوبة للمتغير math حسب كل مجموعة جزئية (نوع البرنامج التعليمي المتبع) .

- لحساب المؤشرات الإحصائية التالية : الأوساط الحسابية للمجاميع الفرعية ، عدد الحالات ، الانحراف المعياري ، الوسيط ، الخطأ المعياري للأوساط الحسابية حسب المجاميع الجزئية نوع البرنامج التعليمي المتبع type of program وحسب النوع gender بصورة مستقلة ، نتبع الخطوات السابقة نفسها لنصل إلى صندوق حوار المتوسطات ونقوم بإدخال متغير النوع كمتغير مستقل إلى جانب نوع البرنامج:



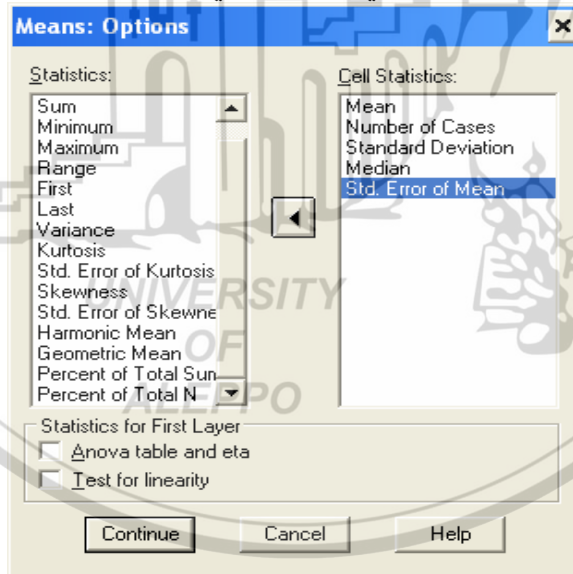


إذ أن :

Dependent List : المتغير التابع أو ما يسمى بمتغير الاستجابة Response variable لتأثره بالمتغير المستقل، إن المتغير التابع في هذا التطبيق هو درجات الرياضيات math score (math).

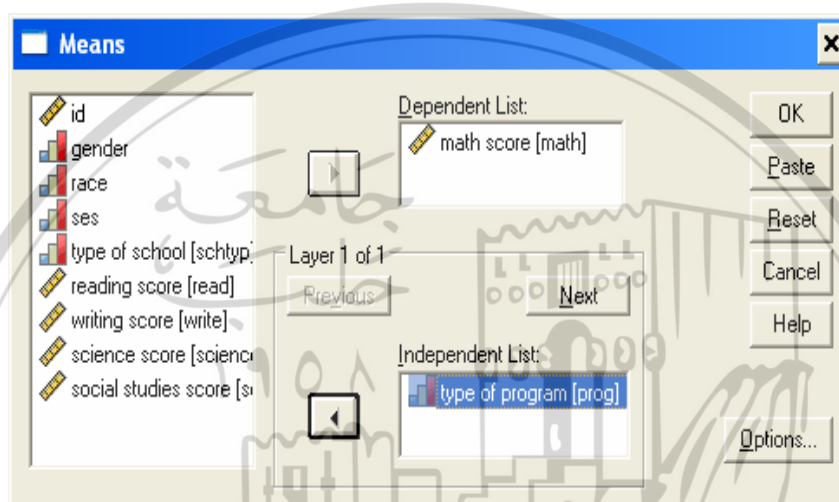
Independent List : المتغير المستقل أو المؤثر ، إن المتغير المستقل في هذا التطبيق هو المجاميع الجزئية نوع البرنامج المتبع Type of Program (prog) الذي يعمل كمتغير تجزئة لدرجات الطلبة.

بالضغط على الزر خيارات Options في صندوق حوار Means نحصل على صندوق حوار جديد يسمى Means: Options كما في الشكل التالي:

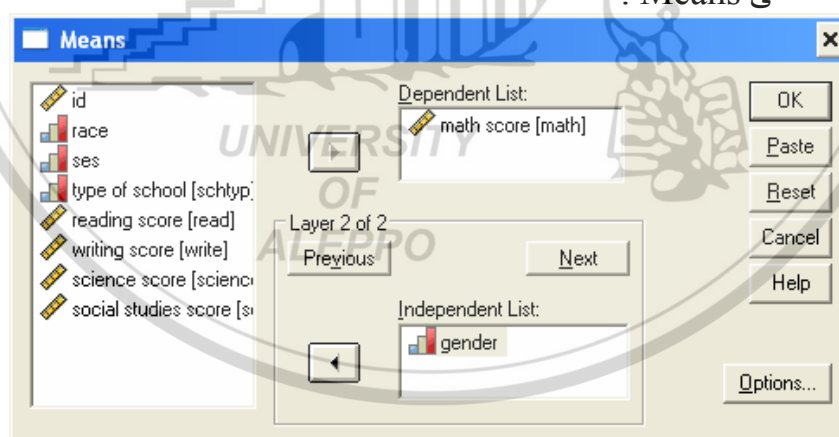


نختار المؤشرات الإحصائية التالية: الأوساط الحسابية للمجاميع الفرعية ، عدد الحالات ، الانحراف المعياري ، الوسيط ، الخطأ المعياري للأوساط الحسابية. بالنقر على الزر Continue ثم OK تظهر النتائج التالية:

- لحساب المؤشرات الإحصائية التالية : الأوساط الحسابية ، عدد الحالات ، الانحراف المعياري ، الوسيط ، الخطأ المعياري للأوساط الحسابية للمتغير math حسب النوع gender في داخل كل من المجاميع الجزئية vocation , academic , general التي تمثل نوع البرنامج التعليمي المتبع type of program من الطلبة ، نتبع الخطوات السابقة نفسها لنصل إلى صندوق حوار المتوسطات ونقوم بإدخال المتغير prog على أنه متغير في الطبقة الأولى Layer 1 وذلك على النحو التالي:



بالنقر على الزر Next أعلى خانة Independent List للانتقال إلى الطبقة الثانية Layer 2 نقوم بنقل المتغير gender إلى خانة Independent List في الطبقة الثانية Layer 2 كما في الشكل التالي المسمى Means .



يمكن الرجوع إلى الطبقة الأولى بالضغط على الزر previous .
بالضغط على الزر خيارات Options في صندوق حوار Means نحصل على صندوق الحوار Means: Options نفعل الخيارات المطلوبة للحصول على المؤشرات الإحصائية.
بالنقر على الزر Continue ثم OK تظهر النتائج التالية:

لقد طبق أحد الباحثين اختبار جالتون Galton – Bar لقياس قدرة الأفراد في معرفة الأطوال والخطوط matching lines for length حيث كان يعرض على كل فرد من أفراد العينة خطأ أفقياً طوله ١١٥ مم ويطلب إليه أن يجد خطأ آخر يبدو له أنه يساوي الخط الأول . تم تطبيق هذه العملية، على عدد من الأشخاص (وحدات التجربة) تم اختيارهم عشوائياً، تحت خمسة ظروف مختلفة يطلق عليها في تحليل التباين اسم المعالجات treatments ، ونورد فيما يلي نتائج هذه التجربة ، إذ تدل الأرقام على الدرجة التي حصل عليها الفرد في كل معالجة من المعالجات الخمس .

الدرجات التي حصل عليها الأشخاص (وحدات التجربة)

المجموع	5	4	3	2	1	الطريقة
24	5	3	7	5	4	1
25	6	2	8	6	3	2
28	6	4	6	7	5	3
31	4	5	8	8	6	4
36	7	7	7	7	8	5
27	5	3	9	6	4	6
34	8	4	9	8	5	7
35	7	4	10	9	5	8
240	48	32	64	56	40	المجموع

والمطلوب:

معرفة إذا كان هناك فرق بين المعالجات الخمس من حيث الدرجات التي حصلت عليها وحدات التجربة.

الحل :

سنقوم بتعريف متغيرين الأول treatment ويشتمل على كافة المشاهدات التي تخص المتغيرات التابعة الخمسة مرتبة الواحد تلو الآخر وهي في مثالنا خمس معالجات مختلفة ، والمتغير الثاني factor ويشتمل على عامل يحدد تبعية كل مشاهدة (وحدات التجربة) للمعالجة المعنية .

treatment	factor	treatment	factor
4	1	3	4
3	1	2	4
5	1	4	4

Report

الدخل الاجمالي

الجنس	وسيلة النقل	Mean	N	Std. Deviation
انثى	خاصة	16523.81	21	6454.6034
	للجامعة	15000.00	10	7888.1064
	Total	16032.26	31	6853.1446
ذكر	عامة	9390.1262	206	10749.6433
	خاصة	9266.6667	15	1944.4671
	للجامعة	8000.0000	5	.0000
	Total	9351.1770	226	10274.2876
Total	عامة	9390.1262	206	10749.6433
	خاصة	13500.00	36	6203.6855
	للجامعة	12666.67	15	7187.9529
	Total	10157.07	257	10150.6275

نلاحظ من الجدول أن متوسط الدخل لجميع أعضاء الهيئة التدريسية الذين شملهم الاستقصاء هو 10157 ل.س .

- متوسط الدخل للنساء اللاتي يستخدمن وسيلة نقل خاصة يساوي 16523 ل.س .
- متوسط الدخل للنساء اللاتي يستخدمن وسيلة نقل للجامعة يساوي 15000 ل.س . وهذا يمثل الطبقة الأولى Layer 1 .
- متوسط الدخل للنساء اللاتي يستخدمن وسيلة نقل خاصة يساوي 16523 ل.س .
- متوسط الدخل للرجال الذين يستخدمون وسيلة نقل عامة يساوي 9390 ل.س .
- متوسط الدخل للرجال الذين يستخدمون وسيلة نقل خاصة يساوي 13500 ل.س .
- متوسط الدخل للرجال الذين يستخدمون وسيلة نقل للجامعة يساوي 12667 ل.س . وهذا يمثل الطبقة الثانية Layer 2 .

بالإضافة إلى ما تقدم نلاحظ بأن الجدول يحتوي على الانحراف المعياري وعدد الأشخاص في كل تصنيف معروض .

٢-١ اختبار التأثير الخطي Test of Linearity

يفيد هذا التحليل في اكتشاف وجود اتجاه عام خطي Linear trend معنوي لأوساط المعالجات في تحليل التباين أحادي التصنيف ، فمثلاً إذا رغبتنا في اختبار التأثير الخطي للمعاملات للبيانات الواردة في المثال (٣) التالي :

مثال (٣) :

Report

math score

type of program	gender	Mean	N	Std. Deviation	Median	Std. Error of Mean
general	male	50.1905	21	8.29228	49.0000	1.80952
	female	49.8750	24	6.79074	49.5000	1.38615
	Total	50.0222	45	7.44217	49.0000	1.10941
academic	male	57.1277	47	8.96519	57.0000	1.30771
	female	56.4138	58	8.60036	55.5000	1.12928
	Total	56.7333	105	8.73022	57.0000	.85198
vocation	male	46.9130	23	8.30067	46.0000	1.73081
	female	46.0000	27	7.78065	45.0000	1.49739
	Total	46.4200	50	7.95418	45.0000	1.12489
Total	male	52.9451	91	9.66478	52.0000	1.01314
	female	52.3945	109	9.15102	53.0000	.87651
	Total	52.6450	200	9.36845	52.0000	.66245

مثال (٢) :

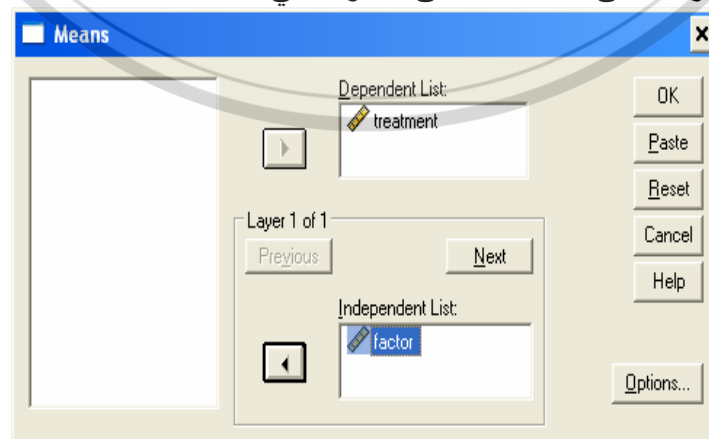
لحساب المتوسطات الحسابية للدخل الإجمالي لأعضاء هيئة التدريس من الذكور والإناث الذين يستخدمون وسائل النقل (عامة ، خاصة ، للجامعة) ، سنحتاج لاستخدام المتغيرات الثلاثة التالية : الدخل الإجمالي Totincome ، وسائل النقل Transprt ، الجنس Sex .

من القائمة Analysis نختار Compare Means فنظهر قائمة أوامر فرعية نختار منها Means فيظهر صندوق حوار ، نختار من قائمة المتغيرات في الجهة اليسرى المتغير Totincome لينتقل إلى المستطيل المعنون Dependent List ، ونختار المتغير Sex لينتقل إلى المستطيل المعنون Independent List .

بالنقر على الزر Next سيؤدي ذلك إلى إنشاء طبقة أخرى Layer 2 of 2 ، نختار المتغير Transprt لينتقل إلى المستطيل المعنون Independent List .
وبالنقر على الزر OK نحصل على النتائج التالية :

6	1	5	4
8	1	7	4
4	1	3	4
5	1	4	4
5	1	4	4
5	2	5	5
6	2	6	5
7	2	6	5
8	2	4	5
7	2	7	5
6	2	5	5
8	2	8	5
9	2	7	5
7	3		
8	3		
6	3		
8	3		
7	3		
9	3		
9	3		
10	3		

من شريط القوائم نختار :
 Analysis → Compare Means → Means
 يظهر صندوق حوار بمسمى Means على النحو التالي:



٣-١ اختبار T لعينة واحدة One Sample T-Test

يستخدم اختبار T لعينة واحدة One Sample T-Test للتعرف فيما إذا كان هناك فرق معنوي Significant Differences بين متوسط المجتمع الذي سحبت منه العينة عن قيمة ثابتة مفترضة، بالإضافة إلى إمكانية تقدير مجال ثقة لمتوسط المجتمع Confidence Interval، ويستخدم عادة في حالة العينات صغيرة الحجم أي $n < 30$.

مثال (٤) :

لنفترض لدينا المتغير Salary الذي يمثل الدخل الشهري الذي يحصل عليه عضو هيئة التدريس في الجامعة، سحبت عينة عشوائية تضم 22 شخصاً .

المطلوب:

- اختبار فرضية العدم القائلة إن الوسط الحسابي للمجتمع الذي سحبت منه العينة هو 10000 ل.س ، مقابل الفرضية البديلة القائلة إن الوسط الحسابي للمجتمع الذي سحبت منه العينة هو مختلف عن 10000 ل.س .
- تقدير مجال ثقة 95 % للوسط الحسابي للمجتمع .
الحل : يمكن كتابة الفرضيات الإحصائية كما يلي:

$$H_0 : \mu = 10000$$

$$H_1 : \mu \neq 10000$$

إحصائية الاختبار :

إن إحصائية الاختبار $t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$ تتبع توزيع ستودنت مع $n-1$ درجة حرية .

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{10680 - 10000}{3105.02511/\sqrt{22}} = 1.0272$$

إذ تم حساب الوسط الحسابي والانحراف المعياري لمتغير الدخل بشكل مسبق عن طريق البرنامج، وبهذه النتيجة لا يمكننا رفض فرضية العدم لأن قيمة إحصائية الاختبار المحسوبة هي أصغر من القيمة الحرجة التي يمكن استخراجها من جدول توزيع ستودنت أو تم حسابها بوساطة الدالة : $tinv(0.05;21) = 2.0796138$ في برنامج Microsoft – Excel. لتطبيق أمر اختبار T لعينة واحدة One Sample T-Test على هذا المتغير نتبع الخطوات التالية:

Measures of Association

	R	R Squared	Eta	Eta Squared
treatment * factor	-.074	.005	.735	.541

إن جدول تحليل التباين يشير الى وجود فروق معنوية بين متوسطات المعالجات إذ إن قيمة

إحصائية الاختبار لفيشر $F = 10.294$ ومستوى المعنوية $Sig = 0.000$.

يتم الحصول على مجموع مربعات الانحرافات عن الخطية على النحو التالي:

$$SS \text{ deviation from linearity} = SS \text{ (combined)} - SS \text{.linearity}$$

$$SS d = 80 - 0.80 = 79.20$$

لحساب قيمة إحصائية الاختبار F لاختبار التأثير الخطي نطبق الصيغة التالية:

$$F = MS \text{. Deviation from linearity} / MS \text{ . within groups}$$

$$F = 26.40 / 1.943 = 13.588$$

إن قيمة مستوى المعنوية $Sig = 0.000$ المقابلة تدعونا إلى رفض فرضية العدم القائلة بأن

أوساط المعالجات تقع على خط مستقيم في حالة إجراء الاختبار بمستوى معنوية ٥ % أي أن الأوساط تنحرف عن الخط المستقيم .

إن قيمة المعامل إيتا تمثل مقياساً للاقتران بين المتغير التابع والمتغير المستقل

وتنحصر قيمته ما بين الصفر أي إن الارتباط بين المتغيرين معدوم ، والواحد الصحيح أي إن

الارتباط بين المتغيرين تام ، أما مربع المعامل إيتا $Eta\text{-}Square$ تمثل نسبة التباينات الكلية في

المتغير التابع treatment التي يفسرها المتغير المستقل factor إذ أن :

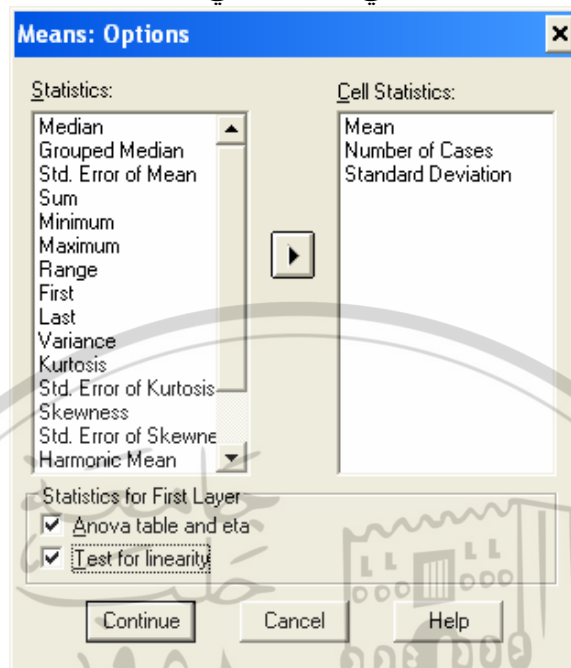
$$Eta\text{-}Square = SS \text{.between groups} / SS \text{.total}$$

$$Eta\text{-}Square = 80.00 / 148.00 = 0.541$$

إن قيمة R تمثل معامل الارتباط البسيط بين المتغير التابع والمتغير المستقل ، أما $R\text{-}square$

تمثل معامل التحديد وهي مقياس لجودة توفيق نموذج الانحدار الخطي.

بالضغط على الزر خيارات Options في صندوق حوار Means نحصل على صندوق حوار جديد يسمى Means: Options كما في الشكل التالي:



نفعل الخيارين : جدول تحليل التباين أحادي التصنيف ومعامل إيتا Anova table and eta واختبار التأثير الخطي Test for Linearity . بالنقر على الزر Continue ثم OK تظهر النتائج التالية:

Report

treatment			
factor	Mean	N	Std. Deviation
1.00	5.0000	8	1.51186
2.00	7.0000	8	1.30931
3.00	8.0000	8	1.30931
4.00	4.0000	8	1.51186
5.00	6.0000	8	1.30931
Total	6.0000	40	1.94804

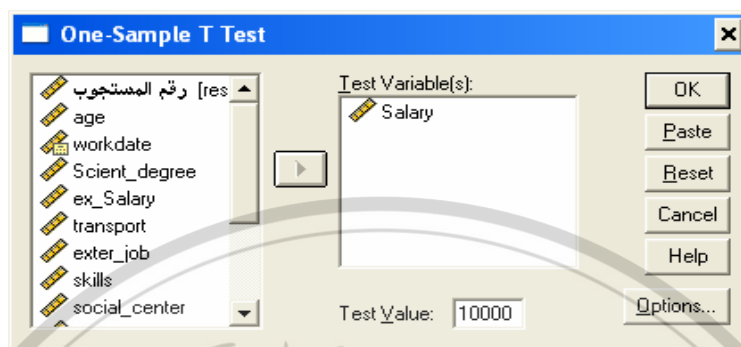
ANOVA Table

		Sum of Squares	df	Mean Square	F	Sig.
treatment * factor	Between Groups (Combined)	80.000	4	20.000	10.294	.000
	Linearity	.800	1	.800	.412	.525
	Deviation from Linearity	79.200	3	26.400	13.588	.000
	Within Groups	68.000	35	1.943		
	Total	148.000	39			

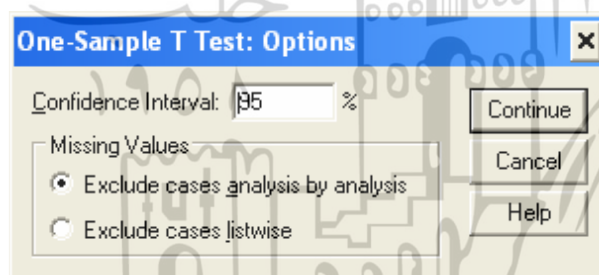
من شريط القوائم نختار :

Analysis → Compare Means → One Sample T- test

يظهر صندوق حوار يسمى One Sample T- test على النحو التالي:



ننقل المتغير المطلوب اختبارَه Salary إلى النافذة Test Variable(s) ثم نضع القيمة المختبرة أو الافتراضية مساوية إلى 10000. نضغط خيارات Options نحصل على صندوق الحوار التالي :



نحدد فترة الثقة المطلوبة 95 % أو 99 % أو 90 % حسب الرغبة. بالنقر على الزر Continue ثم OK تظهر النتائج التالية:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Salary	22	10680.00	3105.02511	661.9936

One-Sample Test

Test Value = 10000						
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Salary	1.027	21	.316	680.00000	-696.691	2056.691

من الجدول نلاحظ بأن مستوى المعنوية:

$$\text{Sig. (2-tailed)} = 0.316 > 0.05$$

9760	F	0	41
11000	M	1	43
16000	F	0	50

المطلوب:

اختبار وجود فرق معنوي بين متوسطي الدخل الشهري لـ أعضاء الهيئة التدريسية الذكور والإناث عند مستوى معنوية 0.05 ، أي اختبار الفرضية التالية:

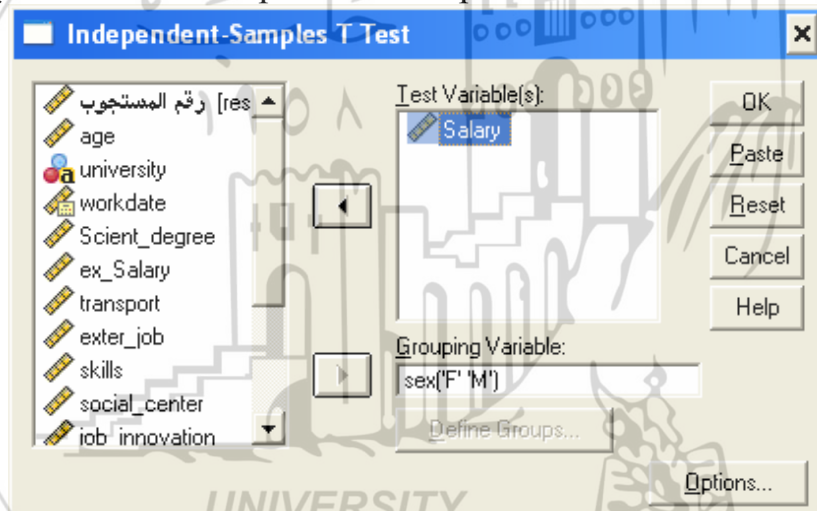
$$H_0 : \mu_F = \mu_M$$

$$H_1 : \mu_F \neq \mu_M$$

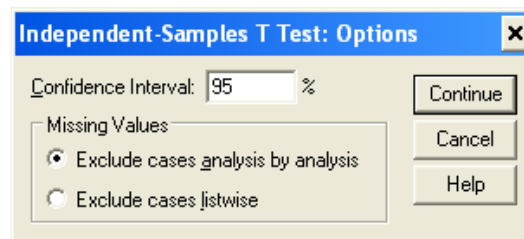
لتنفيذ ذلك نتبع الخطوات التالية:
من شريط القوائم نختار :

Analysis → Compare Means → Independent-Samples T Test

يظهر صندوق حوار يسمى Independent-Samples T Test على النحو التالي:



نقوم بنقل المتغير Salary إلى النافذة المعنونة Test Variable(s) والمتغير Sex (متغير التجزئة) إلى النافذة Grouping Variable ثم تعريف المجاميع الجزئية M , F عن طريق ضغط الزر Define Group إذ أن هذا المتغير له قيمتان متميزتان هما : M , F . بالضغط على خيارات نحصل على صندوق الحوار المسمى اختبار T لعينتين مستقلتين : خيارات يمكننا من خلالها تحديد مجال الثقة المطلوب بالإضافة إلى استبعاد القيم المفقودة من التحليل بالنسبة للمتغيرات التي تحتويها أو استبعاد القيم المفقودة لمتغير وما يقابلها في المتغيرات الأخرى.



يمكن إجراء اختبار T للفرق بين عينتين مستقلتين عن طريق التعليمات Independent-Samples T Test الموجودة ضمن مقارنة المتوسطات Compare Means الذي تحصل عليه بعد النقر على الخيار تحليل " Analyze " من شريط القوائم وبعد الحصول على صندوق الحوار المعنون Independent-Samples T Test، الذي يتطلب توصيف عدة متغيرات لإجراء الاختبار وهي : المتغير التابع والمتغير المستقل أو متغير التجميع (متغير التجزئة) Grouping Variable .

سنحاول اختبار فرضية عدم القائل بأن متوسط الدخل الذي يحصل عليه أعضاء هيئة التدريس للذكور والإناث هو نفسه أي لا يوجد فرق معنوي بين المتوسطين . مع العلم أن البيانات هي الدخل الذي حصل عليه عينة من أعضاء هيئة التدريس وعددها 22 شخصاً وتوزعهم ما بين ذكور وإناث وذلك كما هو موضح بالجدول التالي مع العلم أن الجدول يوضح متغيراً ثالثاً يمثل كون عضو هيئة التدريس يمارس عملاً خارج الجامعة أو لا يمارس .

Salary	gender	Ex_Job	Age
12000	F	1	35
11000	F	0	40
12500	M	0	44
15000	M	1	52
10500	M	0	38
9000	M	1	39
4900	M	1	29
5600	M	0	28
11000	F	0	46
8700	F	1	37
7000	M	0	29
9800	M	1	32
10000	F	0	34
9900	M	1	36
6000	M	0	29
14800	M	1	50
15200	M	1	47
13300	M	1	39
12000	M	0	44

هو أكبر من مستوى المعنوية المطلوب 0.05 لهذا نقبل فرضية العدم القائلة إن متوسط الدخل يساوي 10000 ل.س.

- نلاحظ من الجدول أن مجال ثقة 95 % للفرق بين الوسط الحسابي للعينة والقيمة المختبرة تساوي: [2056.691 : -696.691].

أما لتقدير مجال ثقة 95 % للوسط الحسابي للمجتمع يمكن حسابه عن طريق المعادلة التالية:

$$\bar{x} \pm t_{(0.05;21)} \cdot s / \sqrt{n} = 10680 \pm 2.0796138(661.9936) = [9303.309 : 12056.69]$$

حيث أن :

$$tinv(0.05;21) = 2.0796138$$

ويمكن الحصول على مجال الثقة المطلوب باستخدام الأمر Explore كما هو موضح بالجدول التالي:

Descriptives			Statistic	Std. Error
Salary	Mean		10680.00	661.9936
	95% Confidence Interval for Mean	Lower Bound	9303.309	
		Upper Bound	12056.69	
	5% Trimmed Mean		10706.06	
	Median		10750.00	
	Variance		9641181	
	Std. Deviation		3105.025	
	Minimum		4900.00	
	Maximum		16000.00	
	Range		11100.00	
	Interquartile Range		3775.00	
	Skewness		-.121	.491
	Kurtosis		-.487	.953

١-٤ اختبار الفرق بين متوسطين

Producing a Test of Mean Differences

باستخدام برنامج SPSS يمكننا اختبار الفرق بين متوسطين وذلك في حالتين :

أولاً: اختبار T للفرق بين عينتين مستقلتين Independent-Samples T Test

وبالضغط على Continue ثم OK نحصل على النتائج المطلوبة .

Group Statistics

sex	N	Mean	Std. Deviation	Std. Error Mean
Salary Female	7	11208.57	2363.72184	893.4029
Male	15	10433.33	3444.18074	889.2836

Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	5% Confidence Interval of the Difference	
								Lower	Upper
Salary Equal variances assumed	2.022	.170	.536	20	.598	75.23810	1446.033	-2241.13	791.609
Equal variances not assumed			.615	16.738	.547	75.23810	1260.553	-1887.47	437.950

يتضح من الجدول عدم وجود فرق معنوي بين متوسط الدخل للذكور ومتوسط الدخل للإناث في الحالتين:

الأولى : حالة تساوي تباين المجتمعين .

والثانية : حالة عدم تساوي تباين المجتمعين .

إذ إن قيمة p_value في الحالتين أكبر من 0.05 ، ما يعني أن فرضية العدم مقبولة ، كما أن قيمة إحصائية الاختبار المحسوبة في الحالتين التي تساوي تقريباً 0.536 مفترضين أن تباين المجتمع معلوم ، وتساوي تقريباً 0.615 مفترضين أن تباين المجتمع غير معلوم ، ستكون أصغر بكثير من قيمة t الجدولية التي نجدها من اجل 21 درجة حرية ومستوى معنوية 0.05 مساوية إلى 2.079614

أما اختبار Leven لتجانس التباين يدعونا لقبول فرضية العدم القائلة بتجانس أو تساوي تباين المجتمعين المأخوذة منهما العينتين لأن قيمة $p\_value = 0.17$ أكبر من 0.05 .

ثانياً: اختبار T للملاحظات المزدوجة Paired-Samples T Test

يمكن إجراء اختبار T للملاحظات المزدوجة لاكتشاف معنوية الفروق بين متوسطي عينتين تنتميان إلى المجتمع نفسه حيث تكون مشاهدات العينة على شكل ثنائيات أو أزواج ، مثلاً نختبر

معنوية الفرق بين متوسط ضغط الدم قبل تعاطي دواء معين

وبعده في عينة تضم 25 شخصاً.

مثال (٥)

٨-٢ تحليل التباين في اتجاه واحد One-Way analysis of Variance

سنتناول تحليل التباين في اتجاه واحد في إطار التجارب والاختبارات المصممة Controlled experiments إذ جرى العرف ، في هذه التجارب ، على تسمية الأشياء المراد دراسة تأثيرها بالمعالجات ، وتسمية الوحدات التي تطبق عليها المعالجات بوحدات التجربة experimental units بحيث تطبق كل معالجة على وحدتين على الأقل وبطريقة عشوائية ، ولهذا السبب نجدها في غالبية المراجع المتخصصة ترد تحت مسمى " تصميم تام العشوائية " .

لنفترض أن شركة ما ترغب في مقارنة k طريقة لعرض منتجاتها بالمناجر التابعة لها لمعرفة مدى تأثيرها على المشتري وذلك من خلال تجربة صممت بصورة تطبق كلاً من الطرق (المعالجات) وعددها j ، حيث $(j = 1 , 2 , 3 , \dots , k)$ تمثل عدد الطرق في التجارب المصممة ، على عدد n من المناجر مختارة عشوائياً بحيث يكون لكل متجر (وحدات التجربة) الفرصة نفسها في أن تطبق عليها كل معالجة .

فإذا رمزنا لحجم المبيعات من المنتجات بعد فترة محددة والناجم عن تطبيق المعالجة رقم j على المتجر i بـ x_{ij} يمكن تمثيل نواتج التجربة في الجدول رقم (1) .

x_{12} تمثل حجم المبيعات في المتجر رقم ١ نتيجة قيامه بتطبيق طريقة العرض 2 .

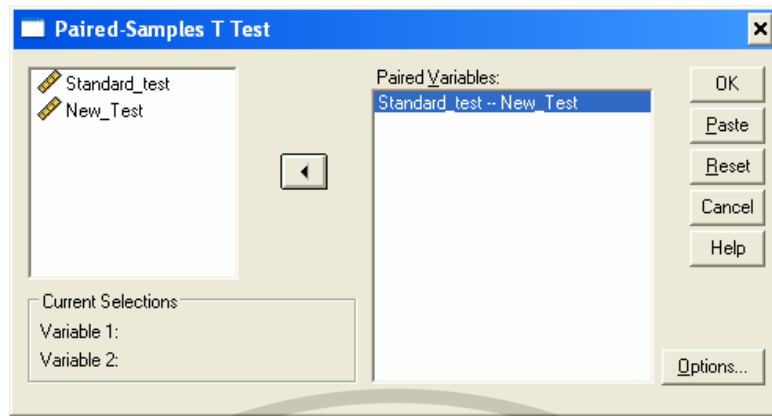
x_{21} تمثل حجم المبيعات في المتجر رقم ٢ نتيجة قيامه بتطبيق طريقة العرض 1 .

x_{nk} تمثل حجم المبيعات في المتجر رقم n نتيجة قيامه بتطبيق طريقة العرض k .

قد تكون طريقة العرض الأولى هي المنتجات من المزرعة أو من المصنع مباشرة إلى المستهلك دون إجراء أية تعديلات عليها ، وقد تكون طريقة العرض الثانية هي التركيز على نوعية الفرز والتدريج والتعبئة والتغليف التي تعطي طابعاً جمالياً مميزاً يركز عليه مجموعة غير قليلة من الزبائن وهكذا ...

جدول رقم (1)
قيم العينة في تصميم تام العشوائية

المتاجر وحدات التجربة	طرق العرض (المعالجات)				
	1	2	3		k



المتغيران Standard_Test و New_Test يمثلان المشاهدات الخاصة بدرجات الاختبار النمطي ودرجات الاختبار الجديد، حيث يتوجب إدخال المتغيرين في قائمة Paired Variables باعتبار أن المشاهدات مزدوجة. وبالضغط على OK نحصل على النتائج المطلوبة.

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Standard_test	83.2727	11	19.48892	5.87613
	New_Test	75.0000	11	16.58312	5.00000

يوضح الجدول السابق الإحصاءات المتعلقة بالعينتين: الوسط الحسابي للعينتين، عدد المشاهدات، الانحراف المعياري للعينتين، الخطأ المعياري للعينتين.

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Standard_test & New_Test	11	.956	.000

وهذا المعامل معنوي بمستوى 0.956. لقد تم حساب معامل الارتباط بين المتغيرين ويساوي 0.05 و 0.01 حيث أن مستوى المعنوية الدقيق أصغر. 0.01 ومستوى دلالة 0.05 دلالة

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error	5% Confidence Interval of the Difference				
				Lower	Upper			
Pair 1 Standard_test – New	8.27273	6.06780	1.82951	4.19632	12.34913	4.522	10	.001

بالنسبة إلى اختبار T للمشاهدات المزدوجة Paired-Samples T Test إن قيمة مستوى المعنوية الدقيق لاختبار من طرفين هي معنوية عند مستوى معنوية 0.05 وعند مستوى معنوية 0.01، إذ أن قيمة:

$$\text{Sig. (2-tailed)} = 0.001 < 0.01 < 0.05$$

ما يدعونا إلى رفض فرضية العدم القائلة بتساوي متوسطي درجات العينتين.

فريق متخصص في أمراض الصحة العقلية يعمل لفترة طويلة في مشفى للأمراض النفسية ، يرغب في قياس مستوى استجابة المرضى لبرنامج فيزيائي تحفيزي، ولتحقيق الغرض يوجد اختبار نمطي ، إلا انه مكلف ويستغرق وقتاً طويلاً ، ولحل المشكلة طور الفريق اختباراً جديداً يسهل تطبيقه وللتأكد من فاعلية الاختبار الجديد لقياس استجابة المرضى ، تم فحص العلاقة بين المعدلات المتحققة نتيجة تطبيق الاختبار الجديد ومعدلات الاختبار النمطي . خضع 11 مريضاً للاختبار الجديد فحققوا الدرجات: 50 ، 55 ، ، 95 ، 100 على الترتيب ، ثم جرى إخضاعهم من جديد للاختبار النمطي فكانت النتائج الموضحة في الجدول .

المطلوب : اختبار الفرضية القائلة بتساوي متوسطي درجات الاختبار النمطي ودرجات الاختبار الجديد عند مستوى معنوية 0.05 ؟ مقابل الفرضية البديلة القائلة بعدم تساوي متوسطي درجات الاختبار النمطي ودرجات الاختبار الجديد

$$H_0 : \mu_S = \mu_N$$

$$H_1 : \mu_S \neq \mu_N$$

درجات الاختبار النمطي	درجات الاختبار الجديد	رقم المريض
61	50	1
61	55	2
59	60	3
71	65	4
80	70	5
76	75	6
90	80	7
106	85	8
98	90	9
100	95	10
114	100	11

لتطبيق أمر اختبار T للملاحظات المزدوجة Paired-Samples T Test نتبع الخطوات التالية:

من شريط القوائم نختار :

Analysis → Compare Means → Paired-Samples T Test

يظهر صندوق حوار يسمى Paired-Samples T Test على النحو التالي:

1	x_{11}	x_{12}			x_{1k}	
2	x_{21}	x_{22}			x_{2k}	
...	...	...			...	
...	...	...			...	
n	x_{n1}	x_{n2}		...	x_{nk}	
المجموع	$T_{.1}$	$T_{.2}$			$T_{.k}$	$T_{..}$
المتوسط	$\bar{x}_{.1}$	$\bar{x}_{.2}$			$\bar{x}_{.k}$	$\bar{x}_{..}$

وبشكل عام نقول بأن :

x_{ij} تمثل المشاهدة رقم i الناتجة عن المعالجة رقم j ، $j = 1, 2, \dots, k$ ، $i = 1, 2, \dots, n_j$ إن عدد المعالجات الإجمالية يساوي k .

$$T_{.j} = \sum_{i=1}^{n_j} x_{ij} \quad \text{مجموع المشاهدات الخاضعة للمعالجة } j.$$

$$\bar{x}_{.j} = T_{.j} / n_j \quad \text{متوسط المعالجة رقم } j.$$

إجمالي قيم كافة المشاهدات .

$$T_{..} = \sum_{j=1}^k T_{.j} = \sum_{j=1}^k \sum_{i=1}^{n_j} x_{ij}$$

إجمالي عدد المشاهدات والمتوسط العام .

$$\bar{x}_{..} = \frac{T_{..}}{N} , N = \sum_{j=1}^k n_j$$

مثال (٦)

أجريت تجربة لمقارنة أربع طرق مختلفة للتعليم (الطريقة التقليدية A وتتمثل بطريقة التلقين ، الطريقة الحديثة B باستخدام وسائل إيضاح سمعية بصرية ، الطريقة الثالثة وهي الطريقة التفاعلية C ، الطريقة الرابعة وهي الطريقة الالكترونية D ، والمعيار المستخدم في المقارنة هو متوسط الدرجات التي يحصل عليها الطلبة بعد مرور فترة زمنية محددة .

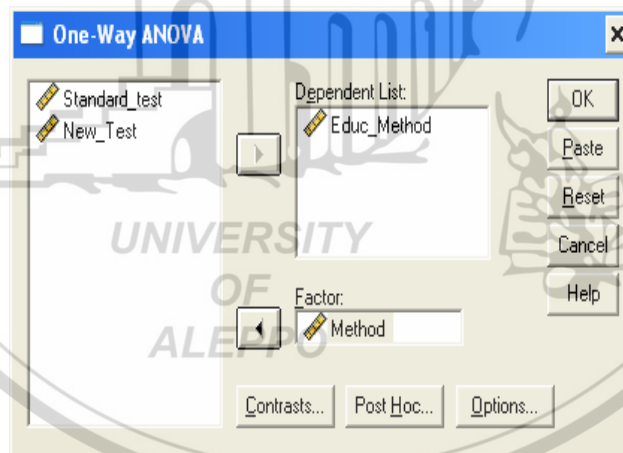
تم إخضاع أربع مجموعات من طلبة الهندسة المعلوماتية حيث تضم كل مجموعة ٧ طلاب لدورات تدريبية استخدمت فيها الطرق التعليمية المختلفة (المعالجات المختلفة)، وفي نهاية الفترة المحددة للدورة ، خضع الطلبة لاختبار محدد بهدف قياس درجة استيعابهم

يشتمل الجدول على: مصدر التباين Source of Variation بين المجموعات Between Groups وداخل المجموعات Within Groups ، درجات الحرية df : Degrees of freedom ، مجموع المربعات Sum of squares ، متوسط المربعات Mean squares ، قيمة إحصائية الاختبار F ratio ، المعنوية الإحصائية Significance . إن قيمة مستوى المعنوية الدقيق لتحليل التباين من طرف واحد هي معنوية عند مستوى معنوية 0.05 وعند مستوى معنوية 0.01، حيث أن قيمة:

$$\text{Sig} = 0.000 < 0.01 < 0.05$$

مما يدعونا إلى رفض فرضية العدم القائلة بتساوي متوسطات الطرق التعليمية المختلفة، أي توجد فروق معنوية بين متوسطات الطرق التعليمية المختلفة.

٢- نظراً لوجود فروق معنوية بين متوسطات الطرق التعليمية المختلفة فهذا يعني عدم تساوي متوسطي طريقتين على الأقل واختبار معنوية الفرق بين كل زوج من الطرق التعليمية نلجأ إلى المقارنات المتعددة باستخدام طريقة الفرق المعنوي الأصغر L.S.D والطريقة القياسية (السيطرة) Dunnett وذلك بالنقر على زر Post Hoc في صندوق الحوار المعنون تحليل التباين أحادي التصنيف One-Way ANOVA كما في الشكل التالي:



فيظهر صندوق حوار يسمى One-Way ANOVA: Post Hoc Multiple Comparison كما في الشكل التالي الذي يوضح وجود العديد من الطرق للمقارنات المتعددة.

مقابل $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$: ليست كل المتوسطات μ_j متساوية أي :

$$H_1 : \mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4 .$$

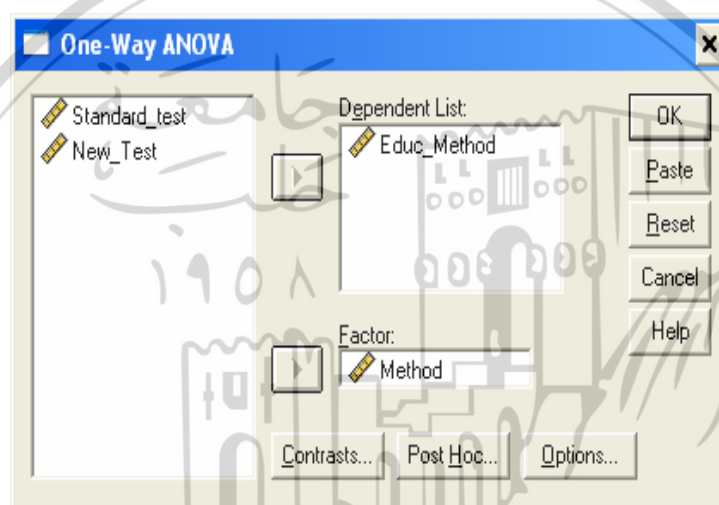
لإجراء تحليل التباين نحاول إدخال البيانات في محرر البيانات عن طريق المتغيرين

Educ_Method و Method.

من شريط القوائم نختار :

Analysis → Compare Means → One-Way ANOVA

يظهر صندوق حوار يسمى One-Way ANOVA على النحو التالي:



إذ أن:

Dependent list : يمثل المتغير التابع أو المعتمد ويمثل نتيجة المعالجات أو التجارب وفي مثالنا يجب أن يكون متغيراً عددياً.

Factor : يمثل المتغير المستقل الذي يستعمل في تعريف المعالجات أو التجارب ويمثل نوع الطريقة التعليمية المتبعة .

وبالضغط على OK نحصل على النتائج المطلوبة:

ANOVA

Educ_Method

	Sum of Squares	df	Mean Square	F	Sig.
Between Group	233.857	3	77.952	10.231	.000
Within Groups	182.857	24	7.619		
Total	416.714	27			

واستجاباتهم للطرق التعليمية المختلفة فكانت نتائج الاختبار بمنزلة المعدلات الموضحة بالجدول التالي :

المعالجات (طرق التعليم)			
A	B	C	D
13	12	14	16
12	13	15	14
6	10	16	18
10	15	18	20
11	18	20	15
8	14	14	13
3	16	15	17

والمطلوب:

١- هل هذه البيانات كافية لتشير إلى وجود فرق بين طرق التعليم المختلفة عند مستوى المعنوية 0.05 وذلك باستخدام تحليل التباين؟

٢- في حالة ظهور معنوية الفروق بين الطرق التعليمية المختلفة باستخدام اختبار F يطلب ما يلي:

٣- اختبار معنوية الفروق بين متوسطي كل طريقتين باستخدام طريقة الفرق المعنوي الأصغر L.S.D للمستوى معنوية ٠,٠٥ ؟

٤- اختبار معنوية الفروق بين متوسط كل من الطرق ٢ و ٣ و ٤ وبين متوسط الطريقة الأولى بافتراضها الطريقة القياسية (السيطرة) باستخدام طريقة Dunnett .

- الفرضيات :

اختبار فرضية العدم القائلة بأن متوسطات كل المجتمعات (المعالجات) متساوية مقابل الفرضية البديلة القائلة إن هناك متوسطين على الأقل غير متساويين ، ونعبر عن ذلك كالتالي :

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_k \quad \text{مقابل} \quad H_A : \text{ليست كل المتوسطات } \mu_j \text{ متساوية} .$$

إذا كان متوسط المجتمعات متساوية فإن تأثير كل معالجة سيكون مساوياً للصفر وبذلك تأخذ الفرضيتان الشكل التالي :

$$H_0 : \tau_j = 0, j = 1, 2, \dots, k \quad \text{مقابل} \quad H_A : \text{ليست كل } \tau_j = 0 .$$

وفي مثالنا ستكون الفرضيتان الأولى (العدم) والبديلة كما يلي :

One-Way ANOVA: Post Hoc Multiple Comparisons

Equal Variances Assumed

☒ LSD ☐ S-N-K ☐ Waller-Duncan
☐ Bonferroni ☐ Tukey Type I/Type II Error Ratio: 100
☐ Sidak ☐ Tukey's-b ☒ Dunnett
☐ Scheffe ☐ Duncan Control Category: First
☐ R-E-G-W F ☐ Hochberg's GT2 Test
☐ R-E-G-W Q ☐ Gabriel ☒ 2-sided ☐ < Control ☐ > Control

Equal Variances Not Assumed

☐ Tamhane's T2 ☐ Dunnett's T3 ☐ Games-Howell ☐ Dunnett's C

Significance level: .05

Continue Cancel Help

نقوم بتفعيل الطريقتين LSD و Dunnett وفيما يتعلق بالطريقة الأخيرة يجب تحديد مجموعة السيطرة (First , Last) وقد تم اختيار First على اعتبار أن الطريقة الأولى هي المسيطرة والتي سيتم المقارنة بها مع تحديد نوع الاختبار من طرف واحد أو طرفين مع تحديد مستوى المعنوية المطلوب 0.05 .
 بالضغط على الزر Continue ثم OK نحصل على النتائج المطلوبة كما يلي:

Multiple Comparisons

Dependent Variable: Educ_Method

	(I) Method	(J) Method	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
LSD	1.00	2.00	-5.00000*	1.47542	.002	-8.0451	-1.9549
		3.00	-7.00000*	1.47542	.000	-10.0451	-3.9549
		4.00	-7.14286*	1.47542	.000	-10.1880	-4.0977
	2.00	1.00	5.00000*	1.47542	.002	1.9549	8.0451
		3.00	-2.00000	1.47542	.188	-5.0451	1.0451
		4.00	-2.14286	1.47542	.159	-5.1880	.9023
	3.00	1.00	7.00000*	1.47542	.000	3.9549	10.0451
		2.00	2.00000	1.47542	.188	-1.0451	5.0451
		4.00	-.14286	1.47542	.924	-3.1880	2.9023
	4.00	1.00	7.14286*	1.47542	.000	4.0977	10.1880
		2.00	2.14286	1.47542	.159	-.9023	5.1880
		3.00	.14286	1.47542	.924	-2.9023	3.1880
Dunnett t (2-sided)	2.00	1.00	5.00000*	1.47542	.007	1.3016	8.6984
	3.00	1.00	7.00000*	1.47542	.000	3.3016	10.6984
	4.00	1.00	7.14286*	1.47542	.000	3.4445	10.8413

*. The mean difference is significant at the .05 level.

a. Dunnett t-tests treat one group as a control, and compare all other groups against it.

من الجدول نلاحظ أنه باستعمال الطريقة LSD ظهرت فروق معنوية بمستوى معنوية 5 % بين متوسط المعالجات (1,2) , (1,3) , (1,4) إذ إن قيمة مستوى المعنوية الدقيق للطرق الثلاث هي :

Sig.=(0.02, 0.00, 0.00)

إذ إن:

Descriptive : تستخدم لعرض بعض المقاييس الوصفية مثل الوسط الحسابي ، الانحراف المعياري ، مجال الثقة ، أصغر وأكبر قيمة.

Homogeneity – of- variance : تستخدم لاختبار تجانس التباين للطرق باستخدام إحصائية Levene إذ يعتبر تجانس التباين أحد الفروض المهمة في إجراء تحليل التباين . عرض بعض المقاييس الوصفية مثل الوسط الحسابي ، الانحراف المعياري ، مجال الثقة ، أصغر وأكبر قيمة.

Means Plot : لعرض التمثيل البياني لمتوسط المعالجات.

Exclude Cases analysis by anlysis : تستخدم لاستبعاد الحالات التي تحتوي قيماً مفقودة للمتغيرات الداخلة في التحليل فقط.

Exclude Cases Listwise : تستخدم لاستبعاد الحالات التي تحتوي قيماً مفقودة لأي واحد من المتغيرات.

بالضغط على الزر Continue ثم OK نحصل على النتائج المطلوبة كما يلي:

Descriptives									
Educ_Method									
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum	
					Lower Bound	Upper Bound			
1.00	7	9.0000	3.55903	1.34519	5.7084	12.2916	3.00	13.00	
2.00	7	14.0000	2.64575	1.00000	11.5531	16.4469	10.00	18.00	
3.00	7	16.0000	2.23607	.84515	13.9320	18.0680	14.00	20.00	
4.00	7	16.1429	2.41030	.91101	13.9137	18.3720	13.00	20.00	
Total	28	13.7857	3.92860	.74243	12.2624	15.3091	3.00	20.00	

Test of Homogeneity of Variances

Educ_Method			
Levene Statistic	df1	df2	Sig.
.834	3	24	.489

يبين الجدول المعنون نتيجة اختبار تجانس التباين (الفرضية الأولية القائلة بتجانس التباين مقابل الفرضية البديلة القائلة بعدم تجانس التباين) باستخدام إحصائية اختبار Levene التي توضح أن قيمة مستوى المعنوية الدقيق هو: $\text{Sig.} = 0.489 > 0.05$ التي تشير إلى قبول الفرضية الأولية القائلة بتجانس التباينات، والشكل البياني التالي يمثل الوسط الحسابي لكل طريقة من طرق التعليم.

Multiple Comparisons

Dependent Variable: Educ_Method

	(I) Method	(J) Method	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
LSD	1.00	2.00	-5.00000*	1.47542	.002	-8.0451	-1.9549
		3.00	-7.00000*	1.47542	.000	-10.0451	-3.9549
		4.00	-7.14286*	1.47542	.000	-10.1880	-4.0977
	2.00	1.00	5.00000*	1.47542	.002	1.9549	8.0451
		3.00	-2.00000	1.47542	.188	-5.0451	1.0451
		4.00	-2.14286	1.47542	.159	-5.1880	.9023
	3.00	1.00	7.00000*	1.47542	.000	3.9549	10.0451
		2.00	2.00000	1.47542	.188	-1.0451	5.0451
		4.00	-.14286	1.47542	.924	-3.1880	2.9023
	4.00	1.00	7.14286*	1.47542	.000	4.0977	10.1880
		2.00	2.14286	1.47542	.159	-.9023	5.1880
		3.00	.14286	1.47542	.924	-2.9023	3.1880
Dunnett t (2-sided)*	1.00	4.00	-7.14286*	1.47542	.000	-10.8413	-3.4445
	2.00	4.00	-2.14286	1.47542	.351	-5.8413	1.5555
	3.00	4.00	-.14286	1.47542	.999	-3.8413	3.5555

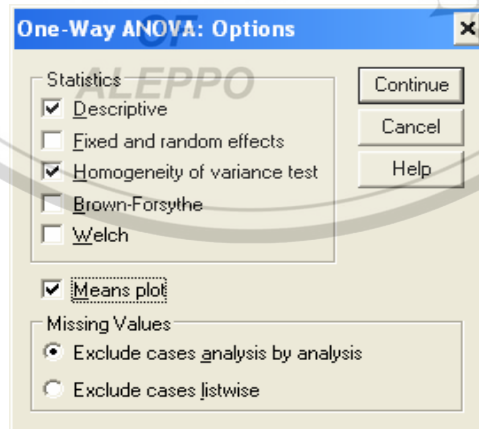
*. The mean difference is significant at the .05 level.

a. Dunnett t-tests treat one group as a control, and compare all other groups against it.

من الجدول نلاحظ أنه باستعمال اختبار Dunnett وجود فروق معنوية بمستوى معنوية ٥ % بين متوسطي الطريقة الاولى والطريقة الأخيرة فقط الأخرى (4,1) إذ إن قيمة مستوى المعنوية الدقيق هي: Sig.= (0.00) وهي أقل من ٠,٠٥ مع وجود علامة * على الفرق بين المتوسطين.

١-٢ تجانس التباين Homogeneity – of- variance :

من صندوق حوار One-Way ANOVA نختار Options نحصل على صندوق حوار بعنوان One-Way ANOVA: Options على النحو التالي:

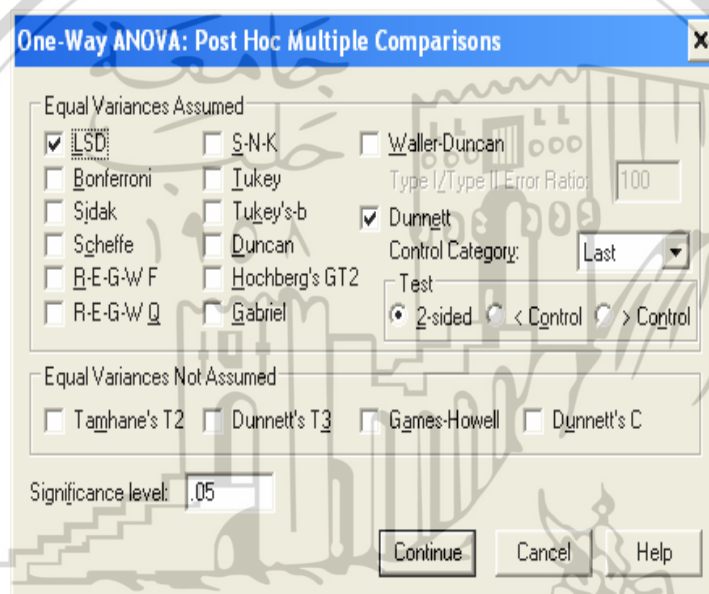


وهي أقل من 0.05 مع وجود علامة * على فروق المتوسطات. كما أن اختبار Dunnett يشير أيضاً إلى وجود فروق معنوية بمستوى معنوية 5 % بين متوسطي الطريقة الأولى والطرق الثلاث الأخرى (1,4), (1,3), (1,2) إذ إن قيمة مستوى المعنوية الدقيق للطرق الثلاث هي:

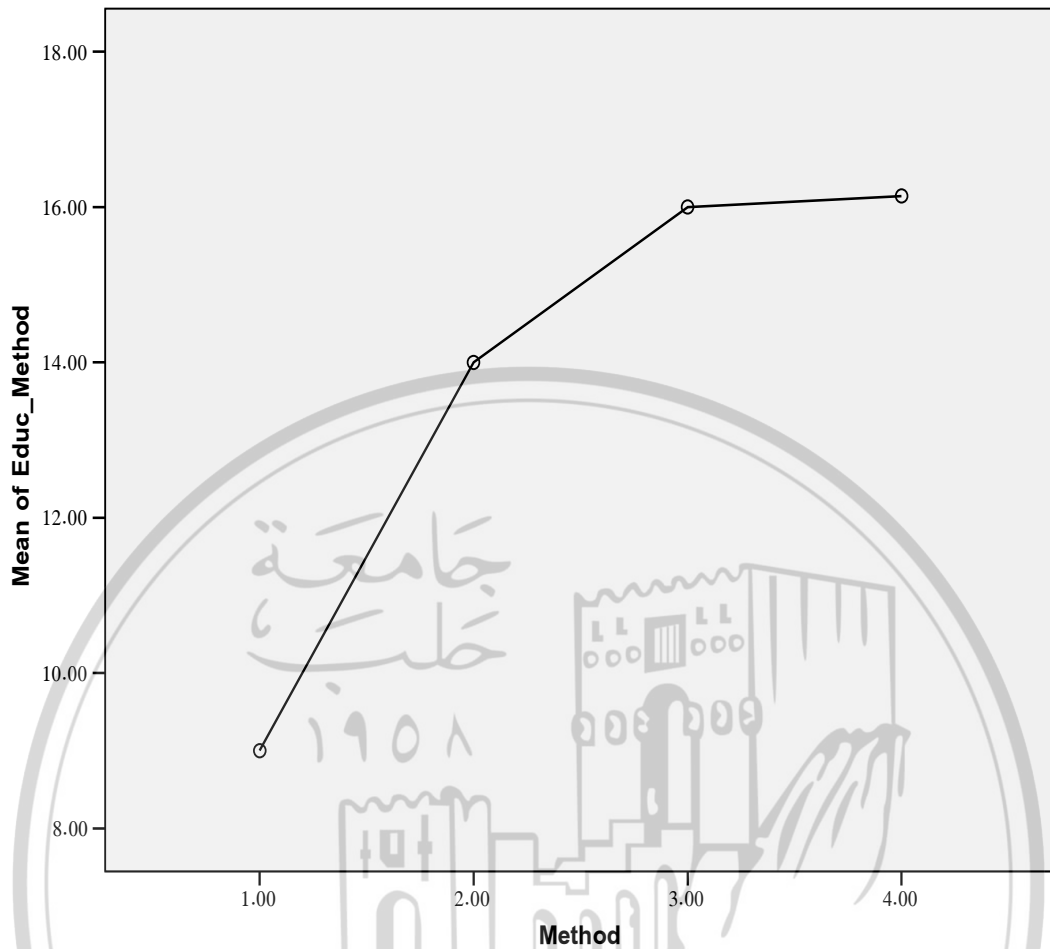
$$\text{Sig.} = (0.07, 0.00, 0.00)$$

هي أقل من 0.05 مع وجود علامة * على الفروق بين المتوسطات.

- إذا قمنا بتفعيل الطريقتين LSD و Dunnett وفيما يتعلق بالطريقة الأخيرة Control (First , Last) نختار Last على اعتبار أن الطريقة الأخيرة هي المسيطرة والتي سيتم المقارنة بها .



بالضغط على الزر Continue ثم OK نحصل على النتائج المطلوبة كما يلي:



مثال (٧) :

لقد طبق أحد الباحثين اختبار جالتون Galton – Bar لقياس قدرة الأفراد في معرفة الأطوال والخطوط matching lines for length إذ كان يعرض على كل فرد من أفراد العينة خطأ أفقياً طوله ١١٥ مم ويطلب إليه أن يجد خطأ آخر يبدو له أنه يساوي الخط الأول . تم تطبيق هذه العملية، على عدد من الأشخاص (وحدات التجربة) تم اختيارهم عشوائياً، تحت خمسة ظروف مختلفة يطلق عليها في تحليل التباين اسم المعالجات treatments ، ونورد فيما يلي نتائج هذه التجربة ، إذ تدل الأرقام على الدرجة التي حصل عليها الفرد في كل معالجة من المعالجات الخمس .

الدرجات التي حصل عليها الأشخاص (وحدات التجربة)

الطريقة	1	2	3	4	5	المجموع
1	4	5	7	3	5	24
2	3	6	8	2	6	25
3	5	7	6	4	6	28
4	6	8	8	5	4	31

تتراوح قيمة معامل الارتباط بين سالب واحد والواحد: $(-1 \leq r \leq +1)$ ، وكلما اقتربت قيمته من الواحد كان الارتباط بين المتغيرين أقوى بغض النظر عن الإشارة، وكلما ابتعدت قيمته عن الواحد، ضعف الارتباط حتى يتلاشى عندما تكون قيمة معامل الارتباط مساوية للصفر ويدل على عدم وجود علاقة خطية بين المتغيرين، أي إن التغير في y لا يمكن إرجاعه إلى x . ومن جهة أخرى إذا كانت قيمة معامل الارتباط موجبة فهذا يعني أنه ارتباط طردي وإذا كانت قيمته سالبة فذلك يعني أن الارتباط بين المتغيرين عكسي. إن الهدف هو معرفة اتجاه العلاقة بين المتغيرين وقوتها فقط وليس حساب الأثر أو الارتباط كما في الانحدار، وليس مهما معرفة أي

المتغيرين هو المستقل أو التابع ولهذا نجد أن: $r_{xy} = r_{yx}$.

نرمز عادة بـ ρ لمعامل الارتباط الذي يصف العلاقة بين المتغيرين في المجتمع بينما تشير r لمعامل الارتباط الذي يصف العلاقة بين المتغيرين في العينة. وغالباً نهتم بمعرفة فيما إذا كانت $\rho \neq 0$ أي أن المتغيرين X و Y مرتبطان، وبما أن ρ غالباً مجهولة نسحب عينة عشوائية من المجتمع المدروس ونحسب r كتقدير لـ ρ ثم نختبر الفرضية: $H_0: \rho=0$ مقابل الفرضية البديلة $H_A: \rho \neq 0$. يمكن حساب معامل الارتباط بطريقتين الأولى مباشرة:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}}$$

والثانية معامل الارتباط بالطريقة غير المباشرة:

$$r = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}}$$

٢-١ معامل سبيرمان لارتباط الرتب Spearman's Rank Correlation Coefficient

في كثير من الأحيان يصعب قياس متغير ما رقمياً ولكن يسهل تعيين رتب للصفة أو الخاصية المراد دراستها عن هذا المتغير، فمثلاً إذا كان هدف إدارة التسويق بإحدى المنشآت الوطنية المنتجة لنوع من العصائر التعرف على مدى تقبل المستهلك لهذا المنتج وذلك من الإجابة على سؤال يحدد درجة تفضيله لهذا المنتج، نجد أنه من السهل عليه أن يرتب درجات تفضيله من الدرجة ١ إلى الدرجة ٥ حيث الدرجة ٥ تعني أن المنتج يستحوذ على اهتمام كبير جداً لديه في حين أن الدرجة ١ تعني أن المنتج غير مقبول، وربما كان من الصعب على المستجوب إعطاء

الفصل الرابع

تحليل الارتباط والانحدار

Correlation and Regression Analysis

§-١ الارتباط Correlation

تتنوع العلاقات وتتعدد بين المتغيرات في الحياة العملية بشكل عام وفي مجال العلوم الاجتماعية بشكل خاص، حيث غالباً ما نجد أنفسنا أننا بحاجة إلى أن نتعلم شيئاً ما حول العلاقة بين متغيرين، فمثلاً قد نهتم بدراسة العلاقة بين سنوات الخبرة التدريسية والأداء الأكاديمي للمشرفين الاجتماعيين، العلاقة بين ذكاء الأبناء والآباء، العلاقة بين الذكاء على أنه قدرة عامة والتحصيل العلمي، العلاقة بين الطول والوزن، مستوى استهلاك بعض المواد الغذائية ومعدل زيادة الوزن، إجمالي دخل العائلة والإنفاق على الرعاية الصحي.....الخ. إن طبيعة العلاقة بين مثل هذه المتغيرات وقوتها يمكن اكتشافها بواسطة تحليل الانحدار والارتباط. فمثلاً قد يهتم مدير الأفراد بإحدى الشركات بدراسة العلاقة بين درجات اختبار القدرات للعاملين بالشركة وبين أدائهم للعمل. فإذا وجدت علاقة بين المتغيرين، عندها قد نرغب في تحديد درجة هذه العلاقة الخطية بين المتغيرين وقوتها باستخدام أساليب تحليل الارتباط Correlation Analysis. إن معاملات الارتباط توضح عادة اتجاه العلاقة^١ فقط بين المتغيرين هل هي طردية أو عكسية ومدى قوة هذه العلاقة.

١-١ معامل الارتباط The Correlation Coefficient :

يتم عادة قياس الارتباط الخطي بين متغيرين بمعامل يسمى معامل الارتباط وهو الذي يبين لنا إحصائياً ما إذا كانت هذه العلاقة طردية أو عكسية، وكذلك يبين لنا مدى قوة هذا الارتباط.

يستخدم معامل الارتباط الخطي البسيط في حالة كون المتغيرات مقاسه بشكل رقمي وكمي سواء كانت البيانات المتاحة عن هذه المتغيرات مبوبة أو غير مبوبة، بافتراض أن العلاقة بين المتغيرين خطية.

^١ المقصود باتجاه العلاقة بين المتغيرين : هل يسيران باتجاه واحد أي كلما تزايد الأول إزداد الثاني ، أم يسيران باتجاهين مختلفين ، أي إذا إزداد الأول تناقص الثاني .

5	8	7	7	7	7	36
6	4	6	9	3	5	27
7	5	8	9	4	8	34
8	5	9	10	4	7	35
المجموع	40	56	64	32	48	240

والمطلوب:

معرفة إذا كان هناك فرق بين المعالجات الخمس من حيث الدرجات التي حصلت عليها وحدات التجربة.

الحل :

- نقوم بتعريف متغيرين الأول يشتمل على كافة المشاهدات التي تخص المتغيرات التابعة الخمسة مرتبة الواحد تلو الآخر وهي في مثالنا خمس معالجات مختلفة ، والثاني يشتمل على عامل يحدد تبعية كل مشاهدة (وحدات التجربة) للمعالجة المعنية .
- من القائمة Analysis نختار Compare Means فتظهر قائمة أوامر فرعية بعنوان One – Way ANOVA .
- من قائمة المتغيرات في الجهة اليسرى ننقر المتغير المعالجات treatments لينتقل إلى المستطيل Dependent list .
- من قائمة المتغيرات في الجهة اليسرى ننقر المتغير أنواع المعالجات type of treatments لينتقل إلى المستطيل Factor، ويترك إكمال الحل للطالب.

قيم عددية لكل مستوى أو درجة من درجات تفضيله للسلعة والأمر نفسه ينطبق على كثير من الأمور التسويقية والظواهر التربوية والاجتماعية .

يستخدم معامل سبيرمان لارتباط الرتب بديلا عن معامل بيرسون في الحالة التي لا تتوافر فيها لدينا قيم المتغيرين X , Y وإنما فقط رتبها، أو التي تكون فيه قيم المتغيرين X , Y متوفرة إلا أننا وللتبسيط نريد إيجاد الارتباط بين رتب X ورتب Y بدلا من قيمهما. في هذه الحالة نرتب قيم X من ١ إلى n بدءا من اصغر رقم ومن ثم نرتب قيم Y أيضا من ١ إلى n . فمثلا إذا كانت لدينا مجموعة من تلاميذ المرحلة الابتدائية و أعطينا رتبا لهؤلاء التلاميذ بخصوص صفتين لكل منهما : الذكاء مثلا والمستوى الاجتماعي ، المستوى الاقتصادي مثلا والمستوى الثقافي لأوليائهم فإنه يتعذر علينا معرفة العلاقة بين المتغيرين باستعمال معامل ارتباط بيرسون لعدم توافر البيانات الرقمية عن هذه المتغيرات ، وبهذه الحالة يمكن استخدام معامل ارتباط الرتب لسبيرمان الذي يأخذ الصيغة التالية :

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}, \dots$$

حيث : n عدد أزواج البيانات للمتغيرين المدروسين . d الفرق بين رتب المتغيرين X , Y .

يستند معامل سبيرمان على الاختلافات d بين الرتب المتقابلة لكل من X , Y . فإذا كان الارتباط بين الرتب تاما وموجبا تكون جميع الفروق مساوية إلى الصفر، أما إذا كان الارتباط تاما وسالبا فإن مجموع مربعات الفروق بين الرتب لكل من المتغيرين سيأخذ أكبر قيمة ممكنة لها وتكون : $r_s = -1$ وفيما عدا هاتين الحالتين تتراوح قيمة معامل ارتباط الرتب لسبيرمان بين -١ ، +١ .

مثال (١)

يوضح الجدول التالي قيمة المنفق على الإعلان أو على مندوبي مبيعات الأدوية على أنها مزايا عينية (أدبية ، ترفيهية ، ثقافية ، صحية ومعنوية) أو تدريبية (x) بآلاف الليرات سنويا ورقم المبيعات السنوي من الأدوية (y) بآلاف الليرات . الأرقام تمثل مبيعات فروع الشركة من الأدوية في المدن الرئيسية مقارنة بما يخصصه كل فرع من مصاريف للدعاية والإعلان ولمندوبي المبيعات .

رقم المبيعات السنوي y

المنفق على الإعلان x

المعنوية الدقيق هي أصغر من ٠,٠٥ وأصغر من ٠,٠١ : = (Sig. (2-tailed))
 $0.000 < 0.01 < 0.05$

Correlations

			Advertising	Sales
Spearman's rho	Advertising	Correlation Coefficient	1.000	.966**
		Sig. (2-tailed)	.	.000
		N	10	10
	Sales	Correlation Coefficient	.966**	1.000
		Sig. (2-tailed)	.000	.
		N	10	10

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

١-٤ الارتباط الجزئي Partial Correlation :

يقيس معامل الارتباط الجزئي قوة العلاقة بين متغيرين مع ثبات متغير ثالث، فمثلاً قد نحصل على قيمة عالية لمعامل الارتباط البسيط للعلاقة بين أسعار اللحوم البيضاء واللحوم الحمراء فقد لا توجد علاقة فعلية بين المتغيرين ولكن كلاهما يتأثر بعامل ثالث هو المستوى العام للأسعار، فإذا استبعدنا هذا المتغير أو ثبتناه فإن قيمة معامل الارتباط بين أسعار اللحوم البيضاء واللحوم الحمراء التي نحسبها ستكون أقل في هذه الحالة، وهذا ما يعرف بالارتباط الجزئي، مع العلم يمكننا استبعاد أي عدد من المتغيرات عند قياس العلاقة بين ظاهرتين.

مثال (٢)

البيانات التالية تمثل العلاقة بين القيمة المقدرة لبنايات سكنية ، عمر البناء ، عدد المداخل، عدد المكاتب ، المساحة الطابقية.

Floor space (x1)	Offices (x2)	Entrances (x3)	Age (x4)	Assessed value (y)
2310	2	2	20	142000
2333	2	2	12	144000
2356	3	1.5	33	151000
2379	3	2	43	150000
2402	2	3	53	139000
2425	4	2	23	169000
2448	2	1.5	99	126000
2471	2	2	34	142900
2494	3	3	23	163000

Bivariate Correlations

Variables:

- Advertising
- Sales

Correlation Coefficients

☒ Pearson ☐ Kendall's tau-b ☒ Spearman

Test of Significance

☒ Two-tailed ☐ One-tailed

☒ Flag significant correlations

Options...

OK Paste Reset Cancel Help

بالضغط على الزر OK نحصل على النتائج التالية:

Correlations

		Advertising	Sales
Advertising	Pearson Correlation	1	.959**
	Sig. (2-tailed)		.000
	N	10	10
Sales	Pearson Correlation	.959**	1
	Sig. (2-tailed)	.000	
	N	10	10

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

- إن قيمة معامل الارتباط الخطي لبيرسون بين الإنفاق الإعلاني والمبيعات تساوي ٠,٩٥٩ أي أن العلاقة بين المتغيرين هي طردية وقوية جداً.
- إن العلاقة الارتباطية بين الإنفاق الإعلاني ورقم المبيعات يختلف معنوياً عن الصفر بمستوى معنوية ٥ % كما أن العلاقة الارتباطية معنوية أيضاً بمستوى معنوية ١ %، لأن قيمة مستوى المعنوية الدقيق هي أصغر من ٠,٠٥ وأصغر من ٠,٠١ : $\text{Sig. (2-tailed)} = 0.000 < 0.01 < 0.05$.
- إن قيمة معامل ارتباط الرتب لسبيرمان بين الإنفاق الإعلاني والمبيعات تساوي ٠,٩٦٦ أي أن العلاقة بين المتغيرين هي طردية وقوية جداً.
- إن العلاقة الارتباطية بين الإنفاق الإعلاني ورقم المبيعات يختلف معنوياً عن الصفر بمستوى معنوية ٥ % كما أن العلاقة الارتباطية معنوية أيضاً بمستوى معنوية ١ %، لأن قيمة مستوى

120	12
150	15
110	13
110	10
130	14
150	16
160	18
170	20
170	21
180	21
1450	160

والمطلوب :

- حساب قيمة معامل الارتباط الارتباط الخطي البسيط لبيرسون Pearson بين المنفق على الإعلان (x) ورقم المبيعات السنوي من الأدوية عند مستوى معنوية ٠,٠٥ .

- حساب معامل ارتباط الرتب لـ سبيرمان Spearman.

٣-١ استخدام برنامج SPSS في دراسة الارتباط

لحساب معامل الارتباط نتبع الخطوات التالية:

من شريط القوائم نختار :

Analysis → Correlate → Bivariate

يظهر صندوق حوار بعنوان Bivariate Correlations يشتمل على ما يلي:

Pearson : يستخدم لحساب معامل الارتباط الخطي لبيرسون.

Kendall's Tau : يستخدم لحساب معامل الارتباط لكندال بالطرق اللامعلمية وباستعمال الرتب.

Spearman : يستخدم لحساب معامل ارتباط الرتب .

Test Significance : يستخدم لاختبار الفرضية من طرفين Two Tailed أو لاختبار الفرضية من طرف واحد One Tailed .

Flag Significance Correlation : يستخدم لتأشير الارتباطات المعنوية عن طريق وضع إشارة النجمة *.

2517	4	4	55	169000
2540	2	3	22	149000

المطلوب : حساب معامل الارتباط الجزئي بين المتغيرين Y و X1 مع ثبات X4 .

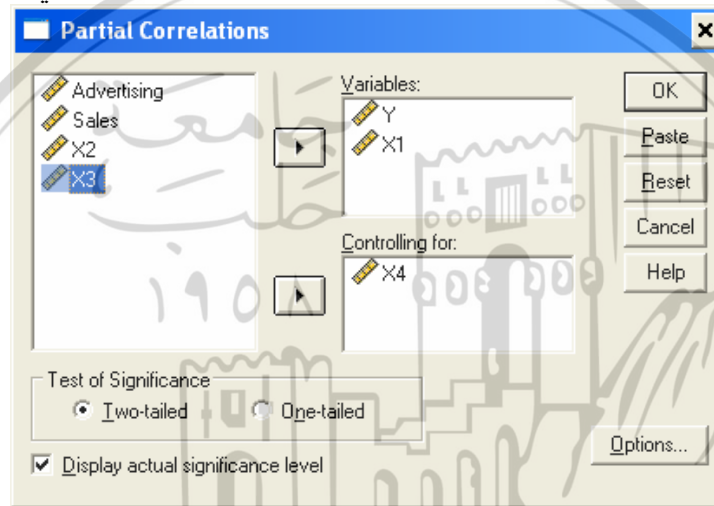
حساب معامل الارتباط الجزئي بين المتغيرين Y و X1 مع ثبات X4 و X3 .

لحساب معامل الارتباط الجزئي بين Y و X1 مع ثبات X4 نتبع الخطوات التالية:

من شريط القوائم نختار :

Analysis → Correlate → Partial

يظهر صندوق حوار بعنوان Bivariate Correlations يشتمل على ما يلي:



في النافذة المعنونة Variables يتم إدخال المتغيرات التي نرغب حساب معامل الارتباط الجزئي لها، وفي النافذة Controlling for يتم إدخال المتغيرات التي نرغب استبعاد أثرها. بالضغط على الزر OK نحصل على النتائج التالية:

Correlations

Control Variables			Y	X1
X4	Y	Correlation	1.000	.491
		Significance (2-tailed)	.	.150
		df	0	8
	X1	Correlation	.491	1.000
		Significance (2-tailed)	.150	.
		df	8	0

إن معامل الارتباط الجزئي بين المتغيرين Y و X1 مع ثبات X4 تساوي ٠,٤٩١ وهذا المعامل غير معنوي لأن قيمة مستوى المعنوية الدقيق أكبر من ٠,٠٥ : $(\text{Sig. (2-tailed)}) = 0.15 > 0.05$.

- لحساب معامل الارتباط الجزئي بين المتغيرين Y و X1 مع ثبات X4 و X3 نتبع الخطوات نفسها السابقة.

- لإيجاد مصفوفة معاملات الارتباط بين عدة متغيرات وليكن الدخل الإجمالي الذي يحصل عليه عضو هيئة التدريس، متوسط الانفاق الشهري، العمر، عدد الأولاد، سنوات الخدمة بالجامعة. نتبع نفس الخطوات السابقة فنحصل على النتائج كما هي موضحة بالجدول، إذ أن النجمة (*) تشير إلى أن معاملات الارتباط هي معنوية إحصائياً عند مستوى معنوية 0.05.

Correlations

		الدخل الإجمالي	متوسط الانفاق الشهري	العمر	عدد الأولاد	سنوات الخدمة بالجامعة
الدخل الإجمالي	Pearson Correlation	1.000	-.131*	-.068	-.153*	-.040
	Sig. (2-tailed)	.	.035	.284	.014	.537
	N	257	257	253	257	240
متوسط الانفاق الشهري	Pearson Correlation	-.131*	1.000	.005	.094	-.082
	Sig. (2-tailed)	.035	.	.932	.135	.208
	N	257	257	253	257	240
العمر	Pearson Correlation	-.068	.005	1.000	.026	-.040
	Sig. (2-tailed)	.284	.932	.	.681	.542
	N	253	253	253	253	240
عدد الأولاد	Pearson Correlation	-.153*	.094	.026	1.000	.061
	Sig. (2-tailed)	.014	.135	.681	.	.345
	N	257	257	253	258	240
سنوات الخدمة بالجامعة	Pearson Correlation	-.040	-.082	-.040	.061	1.000
	Sig. (2-tailed)	.537	.208	.542	.345	.
	N	240	240	240	240	240

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

- لإيجاد معامل الارتباط الجزئي بين الدخل الإجمالي الذي يحصل عليه عضو هيئة التدريس ومتوسط الإنفاق الشهري وذلك بعد عزل تأثير عدد الأولاد من جهة وسنوات الخدمة في الجامعة من جهة أخرى (أي بعد تثبيت المتغيرين). نختار من القائمة Analysis في شريط القوائم الخيار Correlate تظهر قائمة أوامر فرعية، نختار منها Partial فيظهر صندوق حوار Partial Correlation. من قائمة المتغيرات نختار بشكل متتالي المتغيرين Totincom و Consumption فينتقلان بالتالي إلى المستطيل المعنون Variables، ومن قائمة المتغيرات نختار بشكل متتالي المتغيرين Yservin, Childnob فينتقلان بالتالي إلى المستطيل المعنون Controlling for، حيث يسمى كل من المتغيرين الأخيرين بمتغير التحكم وعددها يحدد درجة معامل الارتباط الجزئي.

--- PARTIAL CORRELATION COEFFICIENTS			
Controlling for.. YSERVIN CHILDNOB			
TOTINCOM CONSUMPT			
TOTINCOM	1.0000	-.1275	
	(0)	(236)	

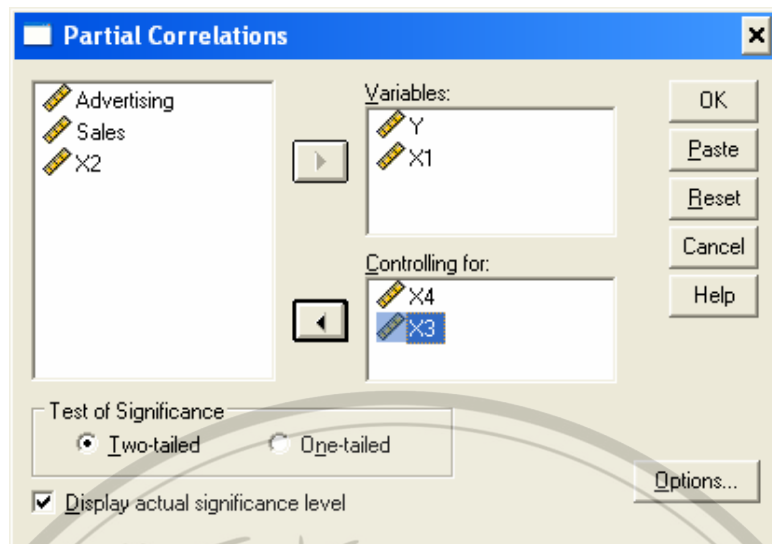
- **Administrative Work** العمل الإداري (لا يوجد = 1 ، رئيس القسم = 2 ، وكيل الجامعة = 3 ، العميد = 4 ، مدير مستشفى = 5) .
- **University** الجامعة (دمشق = 1 ، حلب = 2 ، تشرين = 3 ، البعث = 4) .
- **Socialist Status** الوضع الاجتماعي (أعزب = 0 ، متزوج = 1) .
- **Child Number** عدد الأولاد ، **Salary** الراتب الشهري .
- **Total Income** الدخل الإجمالي .
- **Consumption** متوسط الإنفاق الشهري
- لإيجاد معامل الارتباط بين متغيرين وليكن الدخل الإجمالي الذي يحصل عليه عضو هيئة التدريس ، ومتوسط الإنفاق الشهري ، نختار من القائمة Analysis في شريط القوائم الخيار Correlate فتظهر قائمة أوامر فرعية ، نختار منها Bivariate فيظهر صندوق حوار بالعنوان نفسه.
- من قائمة المتغيرات بصندوق الحوار الواقع في الجهة اليسرى ننقر بشكل متتالي على المتغيرين Totincom و Consumption فينتقلان بالتالي إلى المستطيل المعنون Variables .
- بشكل طبيعي يقوم البرنامج باستخدام معامل الارتباط بيرسون Pearson من بين المعاملات الأخرى الواردة تحت المستطيل Correlation Coefficients .
- بالنقر على الزر OK نحصل على النتائج كما في الجدول التالي :

Correlations

		الدخل الإجمالي	متوسط الإنفاق الشهري
الدخل الإجمالي	Pearson Correlation	1.000	-.131*
	Sig. (2-tailed)	.	.035
	N	257	257
متوسط الإنفاق الشهري	Pearson Correlation	-.131*	1.000
	Sig. (2-tailed)	.035	.
	N	257	257

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

من الجدول نلاحظ أن المعنوية الإحصائية (significance . 2- tailed = 0.035) هي أقل من 0.05 ، ومعنى ذلك أن هناك ارتباطاً ذات دلالة إحصائية بين إجمالي الدخل ومتوسط الإنفاق الشهري ، على الرغم من كون العلاقة الارتباطية عكسية وضعيفة جداً بين المتغيرين التي بلغت 0.131 - والتي لا تتطابق مع ما تقره وتؤكدته النظرية الاقتصادية بخصوص العلاقة ما بين الدخل والاستهلاك .



بالضغط على الزر OK نحصل على النتائج التالية:

Correlations

Control Variables			Y	X1
X4 & X3	Y	Correlation	1.000	.214
		Significance (2-tailed)	.	.580
		df	0	7
	X1	Correlation	.214	1.000
		Significance (2-tailed)	.580	.
		df	7	0

إن معامل الارتباط الجزئي بين المتغيرين Y و X1 مع ثبات X4 و X3 تساوي ٠,٢١٤ وهذا المعامل غير معنوي لأن قيمة مستوى المعنوية الدقيق أكبر من ٠,٠٥ : (Sig. (2-tailed) = 0.214 > 0.05).

مثال (٣)

لدراسة الارتباط باستخدام برنامج SPSS تم إدخال البيانات الشخصية الخاصة بـ 258 عضو هيئة تدريسية في الجامعات السورية شاركوا في الإجابة على استبيان خصص لإلقاء الضوء على المعوقات السلوكية والتنظيمية والخدماتية والمادية لكفاءة أعضاء هيئة التدريس ، وما لذلك من انعكاسات سلبية على مردودية عضو هيئة التدريس وروحه المعنوية.

تضمنت البيانات أوصافاً للمتغيرات وقيمها في ملف البيانات على الشكل التالي :

- Sex الجنس (Male = 1 ; Female =0) ، Age العمر .
- Transportation وسيلة النقل (عامة = 1 ، خاصة = 2 ، للجامعة = 3) .
- Residence نوعية السكن (إيجار = 1 ، بلا = 2 ، ملك = 3) .

	P= .	P= .049
CONSUMPT	-.1275	1.0000
	(236)	(0)
	P= .049	P= .
(Coefficient / (D.F.) / 2-tailed Significance)		
" . " is printed if a coefficient cannot be computed		

من الجدول نلاحظ بأن المعنوية الإحصائية (significance.2- tailed = 0.049) أقل من ٠,٠٥ ومعنى ذلك أن هناك ارتباطاً معنوياً بين إجمالي الدخل ومتوسط الإنفاق الشهري ، على الرغم من انخفاض قيمة معامل الارتباط قليلاً جداً، وبقيت العلاقة الارتباطية عكسية وضعيفة جداً بين المتغيرين 0.128 - حتى بعد عزل أثر المتغيرين ، عدد الأولاد وسنوات الخدمة خارج الجامعة .

٢-٨ تحليل الانحدار Regression Analysis

عند الحديث عن الانحدار الخطي البسيط فذلك يعني تناول بعض المشاكل التي تتضمن متغيرين، فمثلاً قد يهتم مدير التسويق في إحدى الشركات الدوائية بالعلاقة بين حجم مبيعات الأدوية وبين مصروفات الدعاية والإعلان في شركته، كذلك قد يهتم مدير الأفراد بإحدى الشركات بدراسة العلاقة بين درجات اختبار القدرات للعاملين في الشركة وبين أدائهم للعمل. فإذا وجدت علاقة بين المتغيرين، عندها قد نرغب في تحديد درجة متانة هذه العلاقة الخطية بين المتغيرين باستخدام تحليل الارتباط، كذلك قد نرغب في التنبؤ بقيمة أحد المتغيرين باستخدام قيمة المتغير الآخر، وتسمى الأساليب المستخدمة في تناول هذا النوع من المشاكل بأساليب الانحدار. " regression techniques ."

١-٢ فرضيات الانحدار الخطي البسيط Assumptions Underlying Simple Linear Regression

تشير أساليب الانحدار إلى الطرق المستخدمة للتوصل إلى معادلة لتوفيق البيانات المتاحة ، إذ يتم استخدام هذه المعادلة في التقدير والتنبؤ ، يسمى المتغير الذي نرغب في التنبؤ به أو تقديره بالمتغير التابع Y dependent variable ، بينما يسمى المتغير الآخر بالمتغير المستقل X independent variable ، ففي نموذج الانحدار البسيط يوجد متغيران أحدهما تابع والآخر مستقل ، أما في نموذج الانحدار المتعدد فيوجد متغير تابع واحد فقط ومتغيران مستقلان على الأقل . سنهتم بنموذج الانحدار البسيط الذي يمكن تمثيل معادلته ببيانيا بخط مستقيم

لاختبار الفرضية حول β فان إحصائية الاختبار : $t = \frac{b - \beta_0}{s_b}$ تتبع توزيع t مع $n - 2$

درجة حرية . وللحصول على قيمة s_b يجب أن نقدر $\sigma_{y/x}^2$ بوساطة التباين غير المفسر

$$s_{y/x}^2 = \frac{\sum (y_i - \hat{y})^2}{n - 2}$$

المحسوب من بيانات العينة

$$b \pm t_{(1-\alpha/2)} \sqrt{\frac{s_{y/x}^2}{\sum (x_i - \bar{x})^2}}$$

كما أن ٩٥ % مجال ثقة لـ β يعطى بالعلاقة :

مثال (٤)

لنأخذ بيانات التطبيق (٧) المتعلق بقيمة المنفق على الإعلان (x) بآلاف الليرات ورقم المبيعات السنوي من الأدوية (y) بآلاف الليرات . والمطلوب :

- حساب معادلة انحدار Y على X مفترضاً العلاقة الخطية واختبار مؤشرات النموذج.
- حساب مجال ثقة ٩٥ % لكل من الثابت وميل خط الانحدار للنموذج الخطي.
- اختبار جودة توفيق النموذج الخطي باستعمال جدول تحليل التباين ANOVA وباستعمال معامل التحديد مع تحليل الأخطاء العشوائية بالرسم البياني.
- اختبار التوزيع الطبيعي للأخطاء العشوائية بيانياً .

الحل: من شريط القوائم نختار :

Analysis → Regression → Linear

يظهر صندوق حوار بعنوان Linear Regression يشتمل على ما يلي:

٣-٢ معامل التحديد : The Coefficient of Determination

يعبر المقياس SST عن مجموع مربعات الانحرافات

$$SST = \sum (y_i - \bar{y})^2 = \sum (\hat{y} - \bar{y})^2 + \sum (y_i - \hat{y})^2$$

مجموع مربع الانحرافات المفسرة ونرمز له بـ SSR ، أما المقدار

$$\sum (y_i - \hat{y})^2$$
 فيقيس مجموع مربع الانحرافات غير المفسرة ونرمز له بـ SSE

، ويطلق عليه مجموع مربعات الأخطاء (البواقي the residual) وهذا المقدار هو الذي يتم

تصغيره minimized عندما نحصل على خط المربعات الصغرى Least-squares line .

ويمكن التعبير عن العلاقة بين المقادير الثلاثة السابقة كما يلي : $SST = SSR + SSE$ ، يمكننا

القول إن معادلة الانحدار أدت عملاً جيداً من خلال وصفها للعلاقة بين المتغيرين إذا كان التباين

المفسر يشكل نسبة كبيرة من التباين الإجمالي ، وهذه النسبة ، أي نسبة التباين المفسر من التباين

الإجمالي تدعى معامل تحديد العينة r^2 .

$$0 \leq r^2 \leq 1$$

إذ أن:

وتحسب بموجب العلاقة التالية :

$$r^2 = \frac{\sum (\hat{y} - \bar{y})^2}{\sum (y_i - \bar{y})^2} = \frac{b^2 [\sum x_i^2 - (\sum x_i)^2 / n]}{\sum y_i^2 - (\sum y_i)^2 / n} = \frac{SSR}{SST}$$

٤-٢ اختبار تحليل التباين :

يستخدم لقياس جودة توفيق ومعنوية النموذج ويتناول اختبار الفرضية $H_0 : \beta = 0$ باستخدام

الإحصائية F ، هذا الاختبار هو إحدى الطرق للوصول لنتيجة بخصوص العلاقة بين X , Y

- اختبار الفرضية $H_0 : \beta = 0$ باستخدام الإحصائية t :

إذا تحققت فرضيات النموذج الخطي فإن توزيع المعاينة لكل من a و b يتبع التوزيع الطبيعي

بمتوسط وتباين :

$$\mu_a = \alpha, \sigma_a^2 = \frac{\sigma_{y/x}^2 \sum x_i^2}{n \sum (x_i - \bar{x})^2}, \quad \mu_b = \beta, \sigma_b^2 = \frac{\sigma_{y/x}^2}{\sum (x_i - \bar{x})^2}$$

مع العلم أن العديد من نماذج الانحدار تمثل معادلتها بمنحنى وليس بخط مستقيم. وفيما يلي الفرضيات الواجب تحققها في النموذج :

- إن قيم المتغير المستقل X ثابتة، غير عشوائية و المتغير X مقاس بدون أخطاء.
- الأخطاء العشوائية تتوزع طبيعياً بمتوسط يساوي الصفر.
- الأخطاء العشوائية لها تباين ثابت (فرضية تجانس تباين الخطأ العشوائي (Homoscedasticity).

- عدم وجود ارتباط ذاتي Autocorrelation بين الأخطاء العشوائية.
- إن متوسط كافة المشاهدات التي يأخذها Y تقع على نفس الخط المستقيم وتعرف بفرض

الخطية assumption of Linearity ويعبر عنها رياضياً : $\mu_{y/x} = \alpha + \beta x$

٢-٢ طريقة المربعات الصغرى : تكتب معادلة الانحدار الخطي البسيط $Y = a + bX$ حيث تمثل a الجزء القاطع على المحور الراسي وتمثل متوسط قيمة Y عندما X تساوي الصفر، بينما تمثل b ميل خط الانحدار، أي التغير في Y المناظر للتغير بوحدة واحدة في X ، وللتوصل لمعادلة الانحدار المقدرة يجب الحصول على قيمة كل من a , b عن طريق الحل المشترك للمعادلتين الطبيعييتين:

$$\sum y_i = na + b \sum x_i$$

$$\sum x_i y_i = a \sum x_i + b \sum x_i^2$$

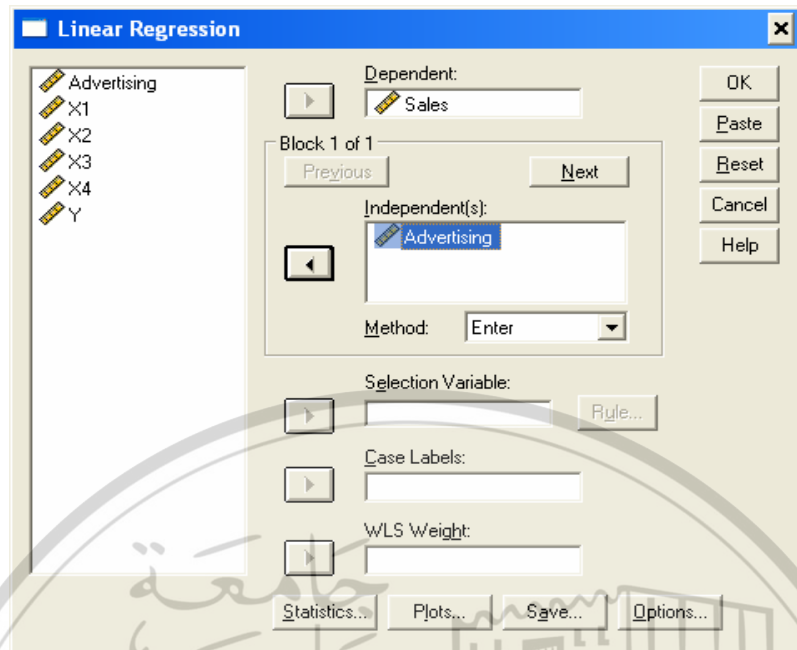
أو بموجب العلاقات:

$$\hat{b} = \frac{n \sum xy - (\sum x)(\sum y)}{n \sum x^2 - (\sum x)^2} = \frac{\sum xy - (\sum x)(\sum y)/n}{\sum x^2 - (\sum x)^2/n}$$

$$\hat{a} = \frac{\sum y}{n} - \hat{b} \frac{\sum x}{n} = \bar{y} - b\bar{x}$$

$$\hat{b} = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sum (x - \bar{x})^2}$$

$$\hat{a} = \bar{y} - \hat{b}\bar{x}$$



إذ أن:

- Dependent : يمثل المتغير التابع.
- Independent : يمثل المتغير أو المتغيرات المستقلة . يمكن إدخال أكثر من مجموعة من المتغيرات المستقلة ، حيث تدخل كل مجموعة ضمن مربع Block له رقم تسلسلي ويمكن الانتقال من Block إلى آخر بواسطة الزرين Previous و Next . إذا كان لدينا نموذجين لاحدهما متغير مستقل X والنموذج الآخر له متغير مستقل Z مع متغير تابع واحد Y لكلا النموذجين ، وبهذه الحالة يتم إدخال المتغير المستقل X في Block1 وإدخال المتغير المستقل Z في Block2 .
- Method : تحدد نوع الطريقة المستخدمة في الانحدار الاعتيادية Enter أو غيرها من الطرق Stepwise , Remove , Backward , Forward.
- Selection Variable : يستخدم في تحديد التحليل لمجموعة معينة من الحالات التي لها قيمة معينة لمتغير الاختيار كأن يقتصر نموذج الانحدار على الحالات التي تكون فيها قيمة المتغير أكبر من ٥) يتم التحديد بواسطة الزر Rule.
- Case Labels : متغير تستخدم قيمه على أنها عناوين لنقاط شكل الانتشار.
- WLS Weight : تستخدم لحساب المربعات الصغرى الموزونة.
- بالضغط على الزر Statistics نحصل على صندوق حوار بعنوان Linear Regression Statistics كما يلي:

إن ميل معادلة خط الانحدار يشير إلى أن زيادة الإنفاق الإعلاني بمقدار ١٠٠٠ ل.س يؤدي إلى زيادة المبيعات بمقدار ٦٣٩٧ ل.س، الأرقام الموجودة ضمن الأقواس وأسفل المقدرات تمثل الخطأ المعياري لثوابت المعادلة المقدرة.

- تستعمل إحصائية الاختبار t لاختبار الفرضيات التالية بالنسبة إلى الثابت والميل:

$$H_0 = \beta_0 = 0 \quad , \quad H_1 = \beta_0 \neq 0$$

$$H_0 = \beta_1 = 0 \quad , \quad H_1 = \beta_1 \neq 0$$

- تستخدم قيمة مستوى المعنوية الدقيق Sig لرفض أو قبول فرضية العدم.

إذا كانت $\text{Sig} < 0.05$ نرفض فرضية العدم عند مستوى معنوية ٠,٠٥ .

إذا كانت $\text{Sig} < 0.01$ نرفض فرضية العدم عند مستوى معنوية ٠,٠١ . وإذا كانت في

الحالتين قيمتها أكبر من مستوى المعنوية المرغوب نقبل فرضية العدم.

إن مستوى المعنوية الدقيق Sig لثابت النموذج يساوي 0.005 وهي أقل من 0.01 ، كما أن

مستوى المعنوية الدقيق Sig لميل خط الانحدار يساوي 0.00 وهي أقل من 0.01، ولهذا نرفض

فرضية العدم في الحالتين أي أن كلا من الثابت والميل يختلفان معنوياً عن الصفر، إن معنوية

ميل خط الانحدار يعكس أهمية متغير الإنفاق الإعلاني في النموذج.

Beta : تمثل المعامل المعياري Standardized coefficient أي ميل خط الانحدار للنموذج

المقدر باستعمال القيم المعيارية $(X - \bar{X})/S$ لكل من المتغير المستقل، والمتغير التابع

$(Y - \bar{Y})/S$ إذ لا يتضمن هذا النموذج ثابتاً.

يمكن كتابة مجال ثقة ٩٥% للحد الثابت بمعادلة الانحدار الخطي كما يلي:

$$\text{Pr}(17.379 \leq \beta_0 \leq 67.916) = 0.95$$

يمكن كتابة مجال ثقة ٩٥% لميل خط الانحدار كما يلي:

$$\text{Pr}(4.858 \leq \beta_1 \leq 7.936) = 0.95$$

Linear Regression: Save

Predicted Values

☒ Unstandardized

☐ Standardized

☐ Adjusted

☐ S.E. of mean predictions

Residuals

☐ Unstandardized

☒ Standardized

☐ Studentized

☐ Deleted

☐ Studentized deleted

Distances

☐ Mahalanobis

☐ Cook's

☐ Leverage values

Influence Statistics

☐ DfBeta(s)

☐ Standardized DfBeta(s)

☐ DfFit

☐ Standardized DfFit

☐ Covariance ratio

Prediction Intervals

☐ Mean ☐ Individual

Confidence Interval: 95 %

Coefficient statistics

☐ Create coefficient statistics

☒ Create a new dataset

Dataset name:

☐ Write a new data file

File...

Export model information to XML file

Browse

☒ Include the covariance matrix

Continue

Cancel

Help

لقد قمنا بتفعيل الخيارين Unstandardized Predicted Values أي قيم $\hat{y}$ المقدرة وكذلك Standardized Residuals أي e_s ، إذ سيتم استخدام المتغيرين لرسم شكل الانتشار ، مع العلم أنه سيتم إضافة هذين المتغيرين إلى محرر البيانات إلى جانب المتغيرات الأخرى الموجودة في ملف البيانات حيث يعطى المتغير Unstandardized Predicted Values الاسم Pre_1 ويعطى المتغير Standardized Residuals الاسم Zre_1.

إن الزر خيارات Options يمكننا من تضمين الحد الثابت في النموذج الخطي أو حذفه. بالضغط على الزر OK نحصل على النتائج التالية:

١- جدول المعاملات:

Coefficients^a

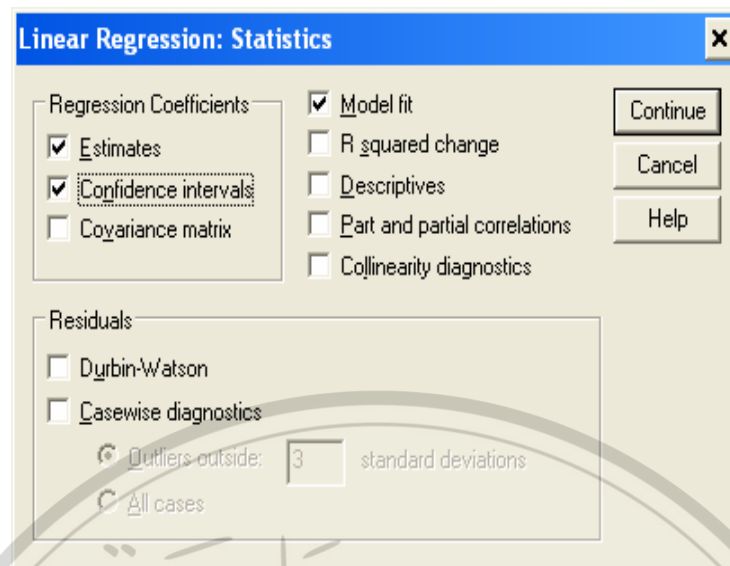
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	42.647	10.958		3.892	.005	17.379	67.916
	Advertising	6.397	.667	.959	9.586	.000	4.858	7.936

a. Dependent Variable: Sales

من الجدول يمكننا كتابة نموذج الانحدار الخطي المقدر كما يلي:

$$\hat{y} = 42.647 + 6.397x$$

(10.958) (0.667)



لقد تم تفعيل الخيارات التالية:

Estimate: لتقدير معالم نموذج الانحدار واختبارات t المرافقة لمعالم النموذج.

Confidence Interval: تستخدم لتقدير مجال ثقة ٩٥% لكل من الثابت وميل خط الانحدار.

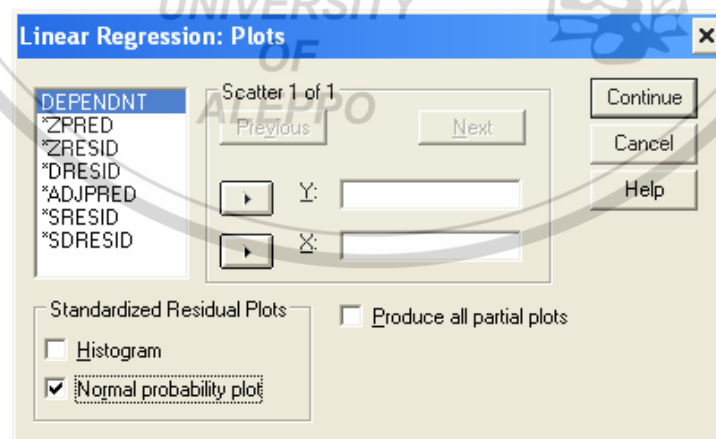
Model Fit: تستخدم لعرض قيم معامل التحديد وجدول تحليل التباين أحادي التصنيف

ANOVA. لتقدير معالم نموذج الانحدار واختبارات t المرافقة لمعالم النموذج.

بالضغط على الزر Continue نعود الى صندوق الحوار السابق Statistics، نضغط على

الزر Plots نحصل على صندوق حوار بعنوان Linear Regression: Plots.

نفعّل الخيار Normal Probability Plot لاختبار التوزيع الطبيعي للأخطاء العشوائية المعيارية.



بالنقر على الزر Save يظهر صندوق حوار بعنوان Linear Regression: save.

٢- جدول تحليل التباين:

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5565.441	1	5565.441	91.885	.000 ^a
	Residual	484.559	8	60.570		
	Total	6050.000	9			

a. Predictors: (Constant), Advertising

b. Dependent Variable: Sales

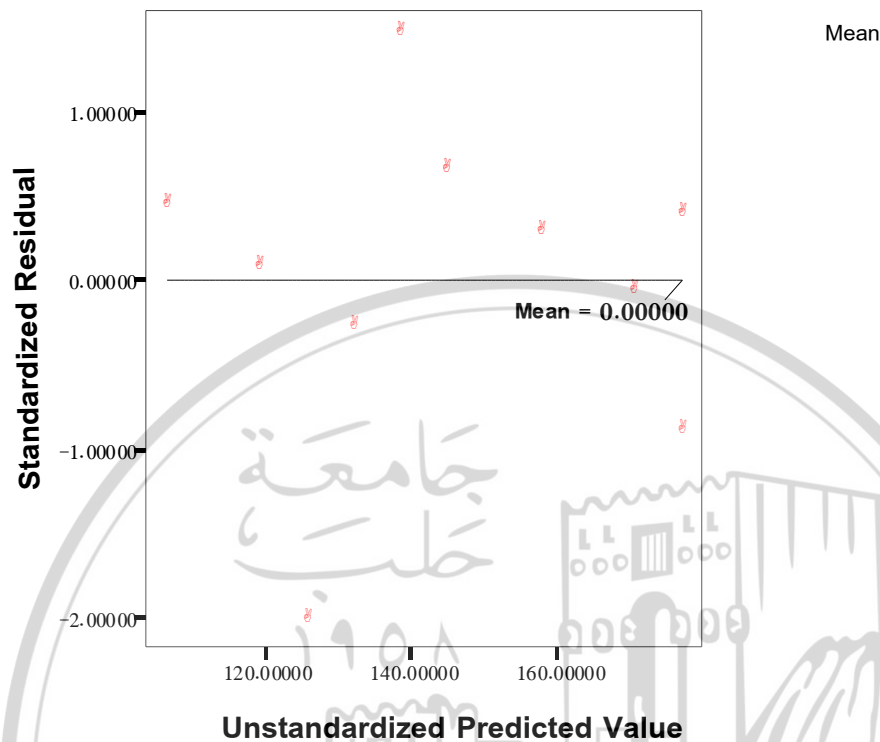
يتضح من الجدول أن إحصائية الاختبار F لفيشر تستخدم لاختبار الفرضية الخاصة بميل خط الانحدار وهذا الاختبار مكافئ تماماً لاختبار t لميل خط الانحدار (حيث أن مستوى المعنوية الدقيق متساوية في الاختبارين) مع العلم أن $F = T^2$. إن قيمة إحصائية الاختبار F لفيشر وقيمة مستوى المعنوية الدقيق يستخدمان للتحقق من مدى ملاءمة النموذج الخطي ومدى جودة توفيق النموذج.

- من جدول تحليل التباين يمكن حساب معامل التحديد Determination Coefficient الذي يعتبر مقياساً لجودة توفيق النموذج، ويمكن حسابه من جدول تحليل التباين:

$$R^2 = \frac{\text{Explained Variations}}{\text{Total Variations}} = \frac{SSR}{SST} = \frac{5565.441}{6050} = 0.919908 \approx 0.92$$

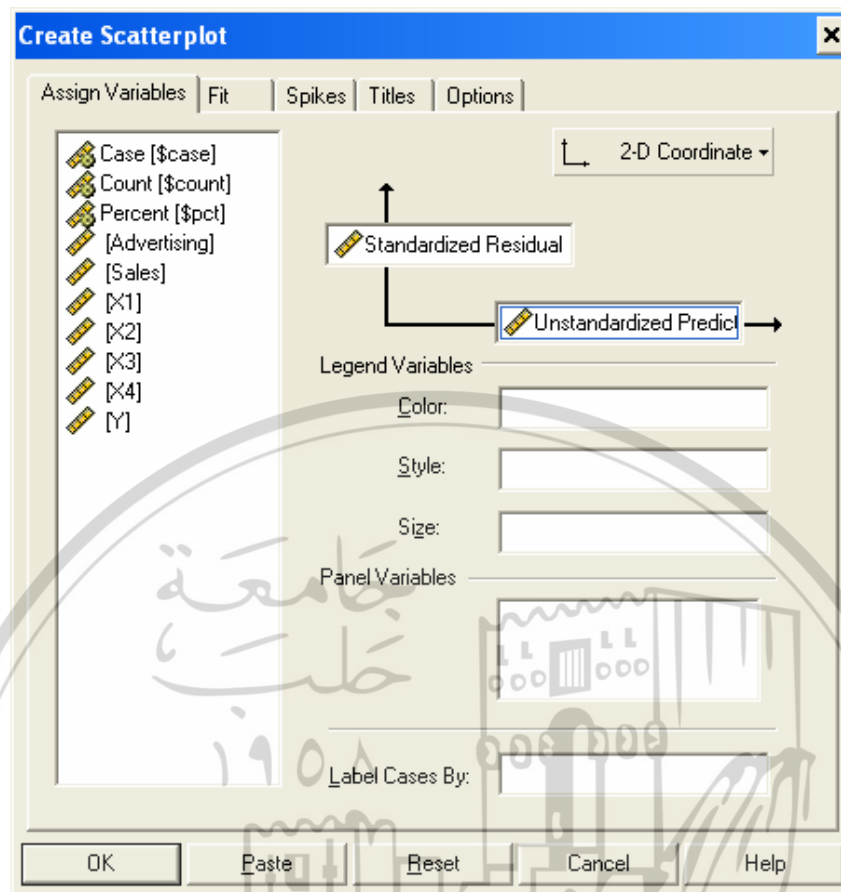
ومعنى ذلك أن ٠,٩٢ من التباينات أو التغيرات (الانحرافات الكلية في المتغير التابع Y) تفسرها العلاقة الخطية أي نموذج الانحدار الخطي، أي ٩٢% من التغير الحاصل في رقم المبيعات مرده التغير في الإنفاق الإعلاني، وأن ٠,٠٨ من التباينات يعود إلى عوامل عشوائية كأن تكون هناك متغيرات مهمة لم تُضمن في النموذج، وبشكل عام كلما اقتربت قيمة معامل التحديد من ١٠٠% دل ذلك على جودة توفيق النموذج.

إن معامل الارتباط الخطي البسيط لبيرسون نحصل عليه عن طريق الجذر التربيعي لمعامل التحديد $r = \sqrt{R^2} = \sqrt{0.9199} = 0.959$ ، إن إشارة معامل الارتباط هي من إشارة ميل خط الانحدار نفسها.



من الشكل نلاحظ توزع النقاط بشكل شريط أفقي حول الصفر مما يشير إلى أن النموذج لا يعاني من مشكلة عدم تجانس تباين الخطأ العشوائي ولا توجد حاجة إلى استخدام علاقة من درجة أعلى.

- يمكن تحليل الأخطاء العشوائية بيانياً أيضاً بأن نقوم بتمثيل القيم الحقيقية للمتغير التابع y على المحور الأفقي والأخطاء المعيارية على المحور العمودي وذلك بالنقر على الزر Plot في صندوق حوار Linear Regression فيظهر صندوق حوار كما يلي:



بالضغط على الزر OK يظهر المخطط في شاشة SPSS Viewer، ننقر مرتين متتاليتين للانتقال إلى شاشة SPSS Chart Editor، حيث نختار من شريط القوائم Mean Fit لإضافة خط إلى المحور العمودي حول الصفر فيظهر المخطط على النحو التالي:

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.959 ^a	.920	.910	7.78266

a. Predictors: (Constant), Advertising

b. Dependent Variable: Sales

- إذا أضيف متغير مستقل للنموذج الخطي فإن قيمة معامل التحديد سترتفع حتى لو لم تكن هناك أهمية للمتغير المستقل في النموذج ، إن إضافة متغير مستقل للنموذج يؤدي إلى زيادة قيمة معامل التحديد والسبب في ذلك هو الزيادة في مجموع مربعات الانحدار SSR وثبات مجموع المربعات الكلي SST ولهذا السبب يتم حساب معامل التحديد المصحح Adjusted R Square الذي يأخذ بالاعتبار النقصان الحاصل في درجات الحرية وقيمه دوماً أقل من قيمة معامل التحديد وهو في مثالنا ٠,٩١ .

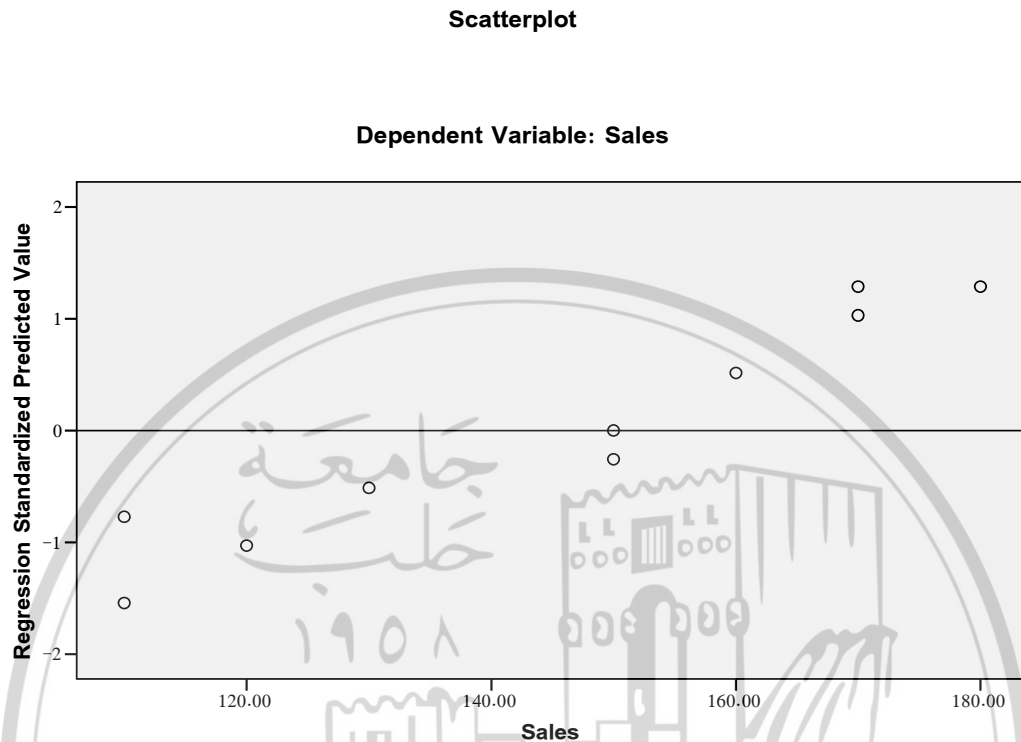
$$\tilde{r}^2 = 1 - \frac{\sum (y_i - \hat{y})^2 / (n-2)}{\sum (y_i - \bar{y})^2 / (n-1)} = 1 - \frac{SSE / (n-2)}{SST / (n-1)} = 0.91$$

- الخطأ المعياري للتقدير Std. error of the Estimate يقيس تشتت القيم المشاهدة عن خط الانحدار وكلما كانت قيمتها صغيرة دل على صغر الأخطاء العشوائية وبالتالي جودة تمثيل خط الانحدار لنقاط شكل الانتشار.

لتحليل الأخطاء العشوائية بيانياً نقوم بتمثيل القيم المقدرة $\hat{y}$ للمتغير التابع Unstandardized Predicted Values (Pre_1) على المحور الأفقي والأخطاء المعيارية على المحور العمودي Standardized Residuals (Zre_1) وذلك بأن نختار من شريط القوائم:

Graph → Interactive → Scatterplot

بالضغط على الزر Continue ثم OK يتم الحصول على المخطط التالي:



- إن اختبار فرضية التوزيع الطبيعي للأخطاء العشوائية تتم بطريقتين الأولى: من خلال رصد الأخطاء المعيارية فإذا وقعت ٩٥% من الأخطاء ضمن المجال (+2 ; -2) فإن الأخطاء تتوزع طبيعياً، من الشكل البياني السابق نلاحظ أن الأخطاء المعيارية لا تتجاوز المجال (-1.6 ; +1.4) تقريباً مما يشير إلى توزع الأخطاء طبيعياً.

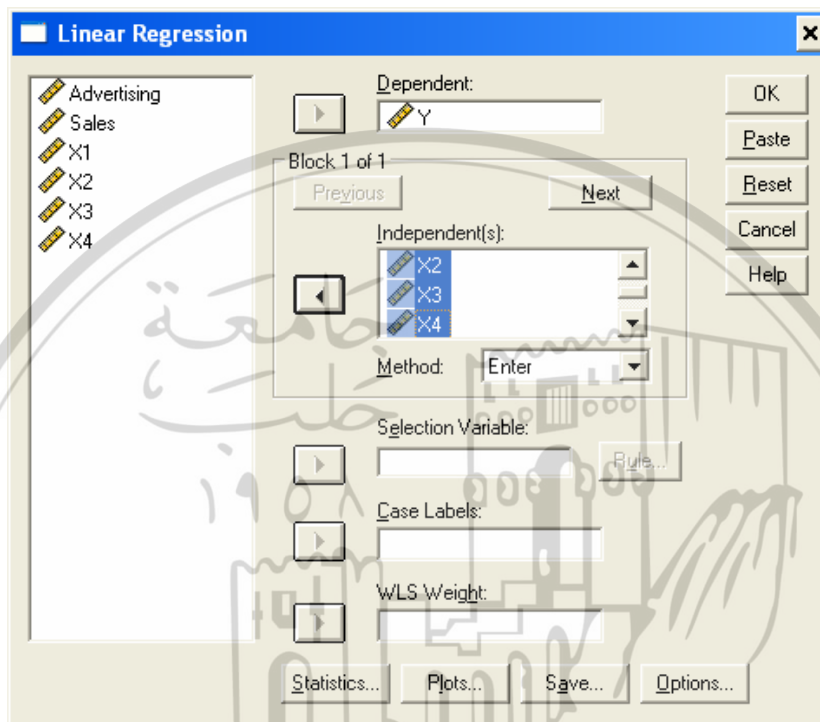
الثانية: عن طريق عرض مخطط Normal Probability Plot الذي قمنا بتفعيله في صندوق حوار Plot وبنتيجة ذلك نحصل على المخطط التالي:

الحل:

من شريط القوائم نختار :

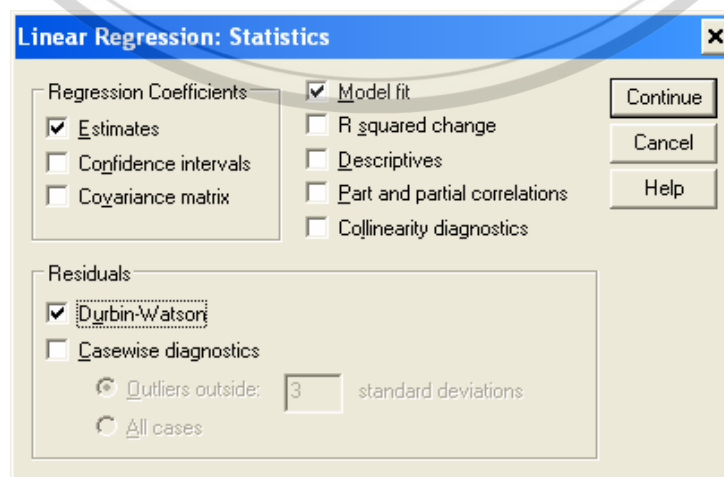
Analysis → Regression → Linear

يظهر صندوق حوار بعنوان Linear Regression يشتمل على ما يلي:



ننقل المتغير التابع y إلى داخل النافذة المعنونة Dependent ، كما ننقل المتغيرات المستقلة الأربعة إلى داخل النافذة المعنونة Independent(s) ، وفي النافذة المعنونة Method نختار الطريقة الاعتيادية enter على أنها طريقة للانحدار.

بالضغط على الزر Statistics نحصل على صندوق حوار بعنوان Linear Regression Statistics كما يلي:



y	القيمة المقدرة للبناء الذي يضم عدة مكاتب
x1	المساحة الأرضية بالأمتار المربعة
x2	عدد المكاتب
x3	عدد المداخل
x4	عمر المبنى بالسنوات

يفترض المثال وجود علاقة خطية بين كل متغير مستقل والمتغير التابع، قامت إدارة المكتب العقاري بسحب عينة عشوائية تضم ١٥ بناءً طابقياً من أصل ١٥٠٠ بناية، فحصل على المعلومات المدونة بالجدول، مع العلم أن نصف مدخل يعني وجود مدخل يستخدم لعمليات الاستلام والمناولة.

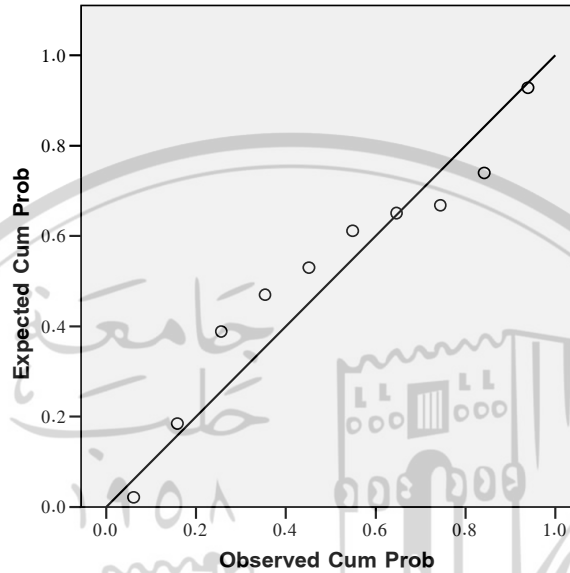
Floor space (x1)	Offices (x2)	Entrances (x3)	Age (x4)	Assessed value (y)
2310	2	2	20	142000
2333	2	2	12	144000
2356	3	1.5	33	151000
2379	3	2	43	150000
2402	2	3	53	139000
2425	4	2	23	169000
2448	2	1.5	99	126000
2471	2	2	34	142900
2494	3	3	23	163000
2517	4	4	55	169000
2540	2	3	22	149000
2460	3	2	16	150000
3000	5	4	17	156000
2780	3	3	14	150000
2760	4	2	21	145000

المطلوب :

- حساب معادلة انحدار المتغير التابع y على المتغيرات المستقلة وتفسير النتائج.
- تكوين جدول تحليل التباين واختبار معنوية المعاملات.
- اختبار مشكلة الارتباط الذاتي التسلسلي باستخدام إحصائية دوربين واتسون WD.

Normal P-P Plot of Regression Standardized Residual

Dependent Variable: Sales



من الشكل البياني نلاحظ أن معظم النقاط تتجمع بالقرب من الخط المستقيم مما يشير إلى التوزيع الطبيعي للأخطاء العشوائية.

§-٣ الانحدار الخطي المتعدد Multiple Linear Regression

إن فرضيات النموذج الخطي المتعدد هي نفسها فرضيات النموذج البسيط يضاف إليها فرضية عدم وجود ارتباط خطي متعدد بين المتغيرات المستقلة Multicollinearity. يأخذ النموذج الخطي المتعدد الصيغة التالية:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + e$$

إذ أن : e ، $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ تمثل الحد الثابت ، معاملات الانحدار الجزئية، الخطأ العشوائي.

مثال (٥)

لنفترض أن أحد المكاتب العقارية تشتري مجموعة من المباني التي تضم مكاتب صغيرة في منطقة تجارية، باستطاعة المكتب العقاري استخدام تحليل الانحدار الخطي المتعدد لتقدير قيمة المبنى في منطقة معينة بالاستناد إلى المتغيرات التالية

المتغير

البيان

لقد تم تفعيل الخيارات التالية:

Estimate: لتقدير معالم نموذج الانحدار واختبارات t المرافقة لمعالم النموذج.

Model Fit: لعرض قيم معامل التحديد وجدول تحليل التباين ANOVA.

Durbin - Watson: لحساب إحصائية دوربين واتسون DW.

بالضغط على الزر Continue ثم OK نحصل على النتائج التالية:

Variables Entered/Removed^b

Model	Variables Entered	Variables Removed	Method
1	X4, X3, ^a X2, X1	.	Enter

a. All requested variables entered.

b. Dependent Variable: Y

تم إدخال كافة المتغيرات المستقلة في التحليل، لم يتم استبعاد أي متغير مستقل، طريقة الانحدار اعتيادية.

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.913 ^a	.834	.767	5472.06133	1.825

a. Predictors: (Constant), X4, X3, X2, X1

b. Dependent Variable: Y

عند اختبار وجود الارتباط الذاتي للأخطاء العشوائية عند مستوى معنوية ٠,٠٥ فإن إحصائية دوربين واتسون $DW = 1.825$ المحسوبة تشير إلى وجود مشكلة يمكن البحث عنها، يترك الأمر للطالب للإجابة على ذلك؟؟

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	222623.2	23985.45		9.282	.000
	X1	-45.872	11.284	-.776	-4.065	.002
	X2	10262.89	2055.257	.870	4.993	.001
	X3	7101.410	2226.658	.509	3.189	.010
	X4	-164.036	66.507	-.331	-2.466	.033

a. Dependent Variable: Y

إن معادلة الانحدار الخطي المتعدد تأخذ الشكل التالي:

$$\hat{y} = 222623.2 - 45.872x_1 + 10262.89x_2 + 7101.41x_3 - 164.036x_4$$

$$(23985.45) \quad (11.284) \quad (2055.257) \quad (2226.658) \quad (66.507)$$

$$R^2 = 0.834 \quad \bar{R}^2 = 0.767$$

Collinearity Diagnostics: لتشخيص مشكلة الازدواج الخطي .

Linear Regression: Statistics

Regression Coefficients

☒ Estimates

☐ Confidence intervals

☐ Covariance matrix

☒ Model fit

☐ R squared change

☐ Descriptives

☒ Part and partial correlations

☒ Collinearity diagnostics

Residuals

☐ Durbin-Watson

☐ Casewise diagnostics

☒ Outliers outside: 3 standard deviations

☐ All cases

Continue

Cancel

Help

بالضغط على الزر Continue ثم OK نحصل على النتائج التالية:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.800 ^a	.641	.497	8044.13303

a. Predictors: (Constant), X4, X3, X2, X1

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1E+009	4	288302143	4.455	.025 ^a
	Residual	6E+008	10	64708076.2		
	Total	2E+009	14			

a. Predictors: (Constant), X4, X3, X2, X1

b. Dependent Variable: Y

إذ أن : $P = k+1 = 5$ تمثل عدد المؤشرات أو ثوابت النموذج ، إن مستوى المعنوية الدقيق $\text{Sig.} = 0.025 < 0.05$ يدعونا إلى رفض فرضية العدم عند مستوى معنوية 0,05 أي أن الانحدار معنوي أو أن المتغيرات المستقلة الأربعة مجتمعة لها تأثير معنوي على الانحدار أو أن واحدة منها على الأقل أو اثنتين أو أكثر من المعالم تختلف معنوياً عن الصفر.

لمعرفة تأثير كل متغير مستقل على المتغير التابع بصورة انفرادية نلجأ لإحصائية الاختبار t في جدول المعاملات Coefficients .

- يتضح من جدول المعاملات Coefficients أن معاملات المتغيرين X_4, X_2 يختلفان معنوياً عن الصفر عند مستوى معنوية 0,05 وبالتالي فإن هذين المتغيرين يؤثران في المتغير التابع

2333	2	2	12	144000
2356	3	1.5	33	151000
2379	3	2	43	150000
2402	4	3	53	139000
2425	4	2	23	169000
2448	2	1.5	60	126000
2471	2	2	34	142900
2494	3	3	23	163000
2517	5	4	55	169000
2540	2	3	22	149000
2460	3	2	16	150000
3000	5	4	17	156000
2800	3	3	14	150000
2760	4	2	21	145000

المطلوب :

- إيجاد معادلة انحدار المتغير التابع y على المتغيرات المستقلة (الانحدار الخطي المتعدد).
- اختبار مشكلة الازدواج الخطي Multicollinearity بين المتغيرات المستقلة.
- إيجاد أفضل معادلة انحدار باستخدام أسلوب الانحدار على مراحل Stepwise

Regressin

الحل: من شريط القوائم نختار :

Analysis → Regression → Linear

يظهر صندوق حوار بعنوان Linear Regression ، ننقل المتغير التابع y إلى داخل النافذة المعنونة Dependent ، كما ننقل المتغيرات المستقلة الأربعة إلى داخل النافذة المعنونة Independent(s) ، وفي النافذة المعنونة Method نختار الطريقة الاعتيادية enter على أنها طريقة للانحدار.

بالضغط على الزر Statistics نحصل على صندوق حوار بعنوان Linear Regression Statistics . إذ نقوم بتفعيل الخيارات التالية:

Estimate: لتقدير معالم نموذج الانحدار واختبارات t المرافقة لمعالم النموذج.

Model Fit: لعرض قيم معامل التحديد وجدول تحليل التباين ANOVA .

Part & Partial Correlations: لتقدير معاملات الارتباط .

هذه الأرقام تشير إلى أن زيادة مساحة البناء بمقدار ١ متر مربع يؤدي إلى نقصان القيمة التقديرية للبناء بمقدار ٤٥,٨٧٢ دولار بافتراض ثبات قيم المتغيرات المستقلة الأخرى، كما أن زيادة عدد المكاتب في المبنى بمقدار مكتب واحد يؤدي إلى زيادة القيمة التقديرية للبناء بمقدار ١٠٢٦٢,٨٩ دولار بافتراض ثبات قيم المتغيرات المستقلة الأخرى، كما أن زيادة عدد المداخل في المبنى بمقدار مدخل واحد يؤدي إلى زيادة القيمة التقديرية للبناء بمقدار ٧١٠١,٤١ دولار بافتراض ثبات قيم المتغيرات المستقلة الأخرى، إن زيادة عمر المبنى بمقدار سنة واحدة يؤدي إلى نقصان القيمة التقديرية للبناء بمقدار ١٦٤,٠٣٦ دولار بافتراض ثبات قيم المتغيرات المستقلة الأخرى. إن قيمة معامل التحديد المعدل $\bar{R}^2 = 0.767$ تشير إلى جودة جيدة نوعاً ما للنموذج الخطي.

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2E+009	4	433098330	459.754	.000 ^a
	Residual	5652135	6	942022.553		
	Total	2E+009	10			

a. Predictors: (Constant), X4, X3, X2, X1

b. Dependent Variable: Y

إذ أن :

$P = k+1 = 5$ تمثل عدد المؤشرات أو ثوابت النموذج ، إن مستوى المعنوية الدقيق $\text{Sig.} = 0.000 < 0.05$ يدعونا إلى رفض فرضية العدم عند مستوى معنوية ٠,٠٥ أي أن الانحدار معنوي أو أن المتغيرات المستقلة الأربعة مجتمعة لها تأثير معنوي على الانحدار أو أن واحدة منها على الأقل أو اثنتين أو أكثر من المعالم تختلف معنوياً عن الصفر. لمعرفة تأثير كل متغير مستقل على المتغير التابع بصورة انفرادية نلجأ إلى إحصائية الاختبار t في جدول المعاملات Coefficients . يتضح من هذا الجدول بأن جميع المعاملات معنوية عند مستوى معنوية ٠,٠٥ وعليه فإن المتغيرات الأربعة تؤثر في المتغير التابع ويوصى بإبقائها في النموذج.

٣-١ اختبار وجود مشكلة الازدواج الخطي:

لنأخذ معطيات التطبيق السابق مع إجراء بعض التعديلات الطفيفة في بعض البيانات:

Floor space (x1)	Offices (x2)	Entrances (x3)	Age (x4)	Assessed value (y)
2310	2	2	20	142000

ويوصى بابقائهما في النموذج ، في حين أن معاملات المتغيرين X_3, X_1 لا تختلفان معنوياً عن الصفر عند مستوى معنوية ٠,٠٥ ، وبالتالي فإن هذين المتغيرين لايؤثران في المتغير التابع ويوصى باستبعادهما من النموذج.

Coefficients ^a										
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
	B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	209050.2	35992.77							
	X1	-34.204	15.695	-.585	-2.179	.054	.177	-.567	-.413	.499
	X2	7779.555	2819.258	.727	2.759	.020	.592	.657	.523	.517
	X3	5386.263	3690.329	.386	1.460	.175	.500	.419	.277	.515
	X4	-371.024	153.883	-.520	-2.411	.037	-.232	-.606	-.457	.773

a. Dependent Variable: Y

إن معادلة الانحدار الخطي المتعدد تأخذ الشكل التالي:

$$\hat{y} = 209050.2 - 34.204x_1 + 7779.55x_2 + 5386.26x_3 - 371.024x_4$$

$$(35992.8) \quad (15.69) \quad (2819.26) \quad (3690.33) \quad (153.88)$$

$$R^2 = 0.641 \quad \bar{R}^2 = 0.497$$

هذه الأرقام تشير إلى أن زيادة مساحة البناء بمقدار ١ متر مربع يؤدي إلى نقصان القيمة التقديرية للبناء بمقدار ٣٤,٢ دولار بافتراض ثبات قيم المتغيرات المستقلة الأخرى، كما أن زيادة عدد المكاتب في المبنى بمقدار مكتب واحد يؤدي إلى زيادة القيمة التقديرية للبناء بمقدار ٧٧٧٩,٥ دولار بافتراض ثبات قيم المتغيرات المستقلة الأخرى، كما أن زيادة عدد المداخل في المبنى بمقدار مدخل واحد يؤدي إلى زيادة القيمة التقديرية للبناء بمقدار ٥٣٨٦,٣ دولار بافتراض ثبات قيم المتغيرات المستقلة الأخرى، إن زيادة عمر المبنى بمقدار سنة واحدة يؤدي إلى نقصان القيمة التقديرية للبناء بمقدار ٣٧١,٠٢٤ دولار بافتراض ثبات قيم المتغيرات المستقلة الأخرى.

إن قيمة معامل التحديد المعدل $\bar{R}^2 = 0.497$ منخفضة نسبياً وتشير إلى أن النموذج لا يعبر عن العلاقة الخطية بشكل جيد.

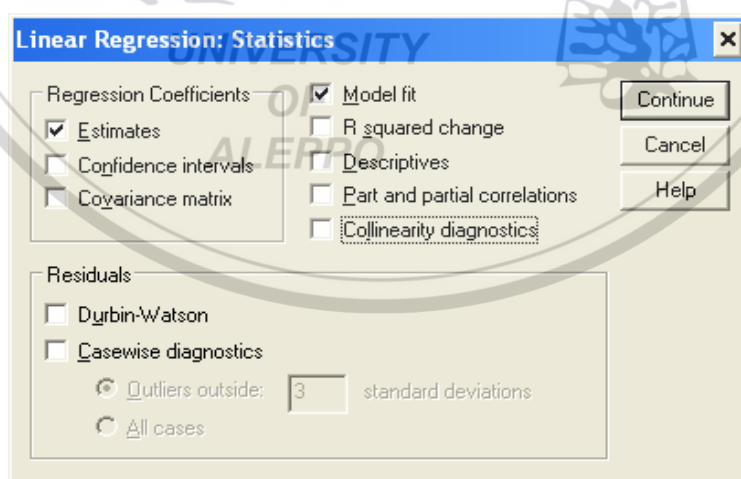
- يتضح من جدول المعاملات Coefficients أن هناك ثلاثة أنواع من الارتباطات وهي: Zero-Order-Correlation: وتمثل معامل الارتباط البسيط لبيرسون بين المتغير المستقل والمتغير التابع.

Partial Correlation: وتمثل معامل الارتباط الجزئي بين المتغير التابع والمتغير المستقل مع ثبات المتغيرات المستقلة الأخرى.

Analysis → Regression → Linear

يظهر صندوق حوار بعنوان Linear Regression ، ننقل المتغير التابع y إلى داخل النافذة المعنونة Dependent ، كما ننقل المتغيرات المستقلة الأربعة إلى داخل النافذة المعنونة Independent(s) ، وفي النافذة المعنونة Method نختار الطريقة Stepwise كطريقة للانحدار. مع العلم هناك خيارات:

- Enter: لإدخال كافة المتغيرات المستقلة في نموذج الانحدار.
 - Stepwise: لإدخال المتغيرات المستقلة واحداً بعد الآخر بخطوات متسلسلة في نموذج الانحدار مع استبعاد المتغيرات التي تصبح غير مؤثرة بوجود بقية المتغيرات.
 - Remove: لاستبعاد المتغيرات غير المهمة بخطوة واحدة من نموذج الانحدار.
 - Backward: استبعاد المتغيرات غير المؤثرة واحداً بعد الآخر بخطوات متسلسلة في نموذج الانحدار.
 - Forward: لإدخال المتغيرات المستقلة واحداً بعد الآخر إلى نموذج الانحدار ولا يتم استبعاد المتغيرات التي تصبح غير مؤثرة فيما بعد من النموذج.
- بالضغط على الزر Statistics نحصل على صندوق حوار بعنوان Linear Regression Statistics . حيث نقوم بتفعيل الخيارات التالية:
- Estimate: لتقدير معالم نموذج الانحدار واختبارات t المرافقة لمعالم النموذج.
 - Model Fit: لعرض قيم معامل التحديد وجدول تحليل التباين ANOVA .



بالضغط على الزر Continue نحصل على صندوق حوار بعنوان Linear Regression Statistics نضغط الزر Option نحصل على صندوق حوار بعنوان خيارات الانحدار الخطي .Linear Regression: Options

- تستخدم الجذور المميزة Eigenvalue للمصفوفة $(X'X)$ في تشخيص مشكلة الازدواج الخطي، حيث أن X هي مصفوفة المتغيرات المستقلة، ففي حالة وجود عدة جذور مميزة قريبة من الصفر فهذا دليل على مشكلة الازدواج الخطي بين المتغيرات المستقلة.

- لاختبار وجود ارتباط خطي بين المتغيرات المستقلة يستعمل ما يعرف بدليل الشرط Condition Index وهو عبارة عن الجذر التربيعي لحاصل قسمة أكبر جذر مميز على كل من الجذور المميزة الأخرى، فمثلاً يتم حساب هذا الدليل للحد الثابت وبالنسبة إلى كافة المتغيرات المستقلة كما يلي:

$$\begin{aligned} \text{ConditionIndex}_c &= \sqrt{\frac{4.709}{4.709}} = 1, \quad \text{ConditionIndex}_{X_1} = \sqrt{\frac{4.709}{0.189}} = 4.989 \\ \text{ConditionIndex}_{X_2} &= \sqrt{\frac{4.709}{0.066}} = 8.433, \quad \text{ConditionIndex}_{X_3} = \sqrt{\frac{4.709}{0.034}} = 11.804 \\ \text{ConditionIndex}_{X_4} &= \sqrt{\frac{4.709}{0.002}} = 55.753 \end{aligned}$$

- إذا زادت قيمة دليل الشرط عن ١٥ فهذا مؤشر على إمكانية وجود مشكلة الازدواج الخطي، أما إذا زادت قيمة المؤشر عن ٣٠ فهذا دليل على خطورة المشكلة.

- يتضح من الجدول أن أكبر قيمة لدليل الشرط هي ٥٥ أي إن مشكلة الازدواجية الخطية خطيرة جداً.

المقياس الآخر لمشكلة الازدواجية الخطية هو ما يعرف بنسبة التباين Variance Proportion وهو يمثل نسبة تباين التقدير المفسر بواسطة المركبة الأساسية principal component المرافقة لكل جذر مميز، إذ تعتبر مشكلة الازدواج الخطي مؤثرة إذا كانت المركبة الأساسية المرافقة لدليل الشرط مرتفعة نسبياً وتساهم بصورة أساسية في تباين اثنين أو أكثر من المتغيرات المستقلة.

- يتضح من جدول تشخيص التعدد الخطي السابق بأن المركبة الأساسية الخامسة تساهم بصورة كبيرة في تباين المتغير الأول وتباين الحد الثابت، المركبة الأساسية الرابعة تساهم بصورة كبيرة في تباين المتغير الثالث، كما أن المركبة الأساسية الثانية تساهم بصورة كبيرة في تباين المتغير الرابع. ولهذا تعتبر مشكلة الازدواجية الخطية مؤثرة بشكل كبير في البيانات.

لإيجاد أفضل معادلة انحدار باستخدام أسلوب Stepwise Regressin من شريط القوائم نختار

:

Part Correlation: وتمثل معامل الارتباط الجزئي بين المتغير التابع والمتغير المستقل المحدد باستبعاد أثر المتغيرات المستقلة الأخرى عن المتغير المستقل المحدد فقط.

- إن نموذج الانحدار المتعدد يعتمد الارتباطات الجزئية بين كل من المتغيرات المستقلة والمتغير التابع ونلاحظ أن المتغير X_2 له أعلى ارتباط جزئي مع المتغير التابع وعليه يكون له أعلى قيمة لإحصائية الاختبار t يليه المتغير المستقل X_4 إذ له ثاني أعلى ارتباط جزئي مع المتغير التابع وعليه يكون له ثاني أعلى قيمة لإحصائية الاختبار t .

- بهدف تشخيص مشكلة الارتباط الخطي المتعدد (الازدواج الخطي) يتم في البداية حساب المعامل $Tolerance$ لكل من المتغيرات المستقلة إذ إن: $Tolerance = 1 - R_{X_i, Others}^2$ و

$$R_{X_i, Others}^2$$

تمثل مربع معامل الارتباط المتعدد بين المتغير المستقل X_i وبقية المتغيرات المستقلة.

نستخرج معامل تضخم التباين VIF Variance Inflation Factor إذ إن:

$$VIF = \frac{1}{Tolerance}$$

يعتبر هذا المعامل مقياساً لتأثير الارتباط بين المتغيرات المستقلة على زيادة تباين معلمة المتغير المستقل (تتميز مشكلة الازدواج الخطي بارتفاع تباين معالم النموذج وبالتالي عدم ظهور المعلمة معنوية نتيجة انخفاض قيمة إحصائية الاختبار t).

- إذا كانت قيمة معامل تضخم التباين VIF لأحد المتغيرات المستقلة تزيد عن ٥ أو ١٠ فهذا دليل على أن تقدير المعلمة المرافقة لهذا المتغير المستقل تتأثر ومشكلة الازدواج الخطي.
- إن قيم معامل تضخم التباين VIF لمتغيرات النموذج تشير إلى عدم تأثر أي من المتغيرات ومشكلة الازدواج الخطي.

Collinearity Diagnostics^a

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions				
				(Constant)	X1	X2	X3	X4
1	1	4.709	1.000	.00	.00	.00	.00	.01
	2	.189	4.989	.00	.00	.01	.03	.68
	3	.066	8.433	.02	.01	.23	.10	.04
	4	.034	11.804	.00	.00	.66	.75	.03
	5	.002	55.753	.98	.99	.10	.11	.24

a. Dependent Variable: Y

حيث نقوم بتفعيل الخيارات التالية:

- تحديد مستوى المعنوية أو قيمة F التي سيتم بموجبها إدخال واستبعاد المتغيرات من النموذج وذلك في حقل معيار طريقة الانحدار على مراحل Stepping Method Criteria وذلك بموجب الخيارين:

Use probability of F : لتحديد مستوى المعنوية المستخدم في كل خطوة لإدخال واستبعاد المتغيرات من النموذج، إذا أردنا متغيرات أكثر في النموذج نزيد من قيمة Entry ولاستبعاد عدد أكبر من المتغيرات من النموذج نقلل من قيمة Removal .

Use F Value : لتحديد قيمة F التي سيتم بموجبها إدخال واستبعاد المتغيرات من النموذج ، إذا أردنا متغيرات أكثر في النموذج نقلل من قيمة Entry ولاستبعاد عدد أكبر من المتغيرات من النموذج نزيد من قيمة Removal .

- سنعتمد القيم الافتراضية للخيار الأول Use probability of F أي ٥ % لإدخال المتغير و ١٠ %

لاستبعاد المتغير المستقل من التحليل.

بالضغط على الزر Continue نحصل على صندوق حوار بعنوان Linear Regression Statistics نضغط الزر OK نحصل على النتائج التالية:

الفصل الخامس

استكشاف البيانات Explore

§-١ استكشاف البيانات بواسطة الأمر Explore :

يعد أسلوب الاستكشاف Explore الخطوة الأولى في التعامل مع البيانات حيث يقدم ملخصاً إحصائياً للبيانات، إعداد مخططات لكل الحالات أو إعداد مخططات لمجاميع لهذه الحالات. يستند هذا الأمر في فحص البيانات Screening لإيجاد القيم الشاذة، القيم المتطرفة، المقاييس الوصفية للبيانات ، اختبار الفرضيات للاستدلال وذلك إذا كان الأسلوب الإحصائي المستخدم ملائماً أم لا ، معرفة توزيع البيانات هل هو طبيعي أو لا ؟ ، وهل هناك حاجة لتحويل البيانات Transformation.

مثال (١):

لنفترض لدينا المتغير درجات الدراسات الاقتصادية Socst المأخوذ من ملف البيانات hsb2 الذي يشتمل على 200 مشاهدة تخص عينة عشوائية لطلاب المدرسة العليا تتضمن المعلومات الشخصية عن الطالبة مثل نوعهم (female) their gender ، الحالة الاقتصادية الاجتماعية (ses) socio-economic status الخلفية العرقية (race) ethnic background ، كما تتضمن عدداً من درجات الاختبارات النمطية مثل اختبار القراءة (read)، الكتابة (write)، الرياضيات (math)، والدراسات الاجتماعية (socst).

لتطبيق أمر الاستكشاف على هذا المتغير نتبع الخطوات التالية:
من شريط القوائم نختار :

Analysis → Descriptive statistics → Explore

يظهر صندوق حوار بمسمى Explore على النحو التالي:

بموجب طريقة الانحدار على مراحل يتم إدخال المتغيرات المستقلة الأربعة واحداً بعد الآخر إلى النموذج مع العلم أن المتغير الداخل سيكون عرضة للاستبعاد في الخطوات التالية إذا ثبت عدم معنويته بوجود المتغيرات الأخرى.

- في مثالنا كان المتغير X2 أول المتغيرات الداخلة في النموذج لأن له أكبر معامل ارتباط بسيط مع المتغير التابع وأكبر معامل ارتباط جزئي أيضاً مع المتغير التابع، وبالتالي أكبر قيمة لإحصائية الاختبار t. من الجدول نجد أن قيمة مستوى المعنوية الدقيق تساوي 0.02 وهي أصغر من مستوى المعنوية المعتمد لإدخال المتغير وهو 0.05 ولهذا تم إدخال المتغير X2 في التحليل وعليه يصبح النموذج كما يلي:

$$\hat{y} = 129881.8 + 6333.475x_2$$

وهذا النموذج هو أفضل نموذج انحدار تم اعتماده ، حيث لا يتضمن المتغيرات X3 , X4 , X1.

Excluded Variables^b

					Partial	Collinearity
Model		Beta In	t	Sig.	Correlation	Statistics
						Tolerance
1	X1	-.178 ^a	-.667	.517	-.189	.731
	X3	.211 ^a	.721	.484	.204	.606
	X4	-.323 ^a	-1.496	.160	-.396	.980

a. Predictors in the Model: (Constant), X2

b. Dependent Variable: Y

يبين الجدول المتغيرات التي تم استبعادها من التحليل اعتباراً من الخطوة الأولى، إذ إن:

Beta In : تمثل المعلمة المعيارية للمتغير لو أُدخل إلى النموذج في الخطوة التالية.

t : تمثل قيمة إحصائية الاختبار t لستودنت.

Sig. : يمثل مستوى المعنوية الدقيق ويجب أن يكون أقل من ٠,٠٥ ليصار لإدخال المتغير بالتحليل.

Partial Correlation : تمثل معاملات الارتباط الجزئية.

Collinearity Statistics Tolerance : تمثل إحصائيات الازدواج الخطي والتحمل.

Variables Entered/Removed^a

Model	Variables Entered	Variables Removed	Method
1	X2	.	Stepwise (Criteria: Probability -of- F-to-enter ≤ .050, Probability -of- F-to-remove ≥ .100).

a. Dependent Variable: Y

المرحلة الأولى تم إدخال المتغير X2 في التحليل طبقاً لقيم F المعتمدة في الإدخال والاستبعاد.

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	129881.8	7881.461		16.479	.000
	X2	6333.475	2390.884	.592	2.649	.020

a. Dependent Variable: Y

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.592 ^a	.351	.301	9483.50675

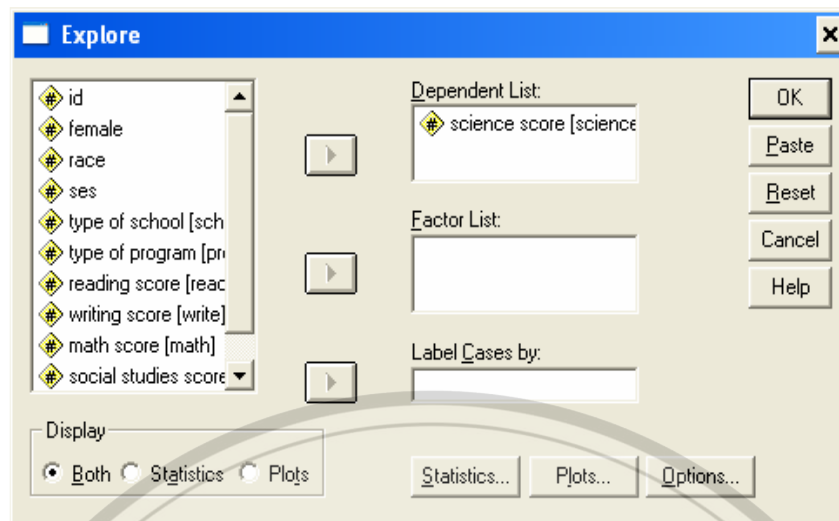
a. Predictors: (Constant), X2

ANOVA^b

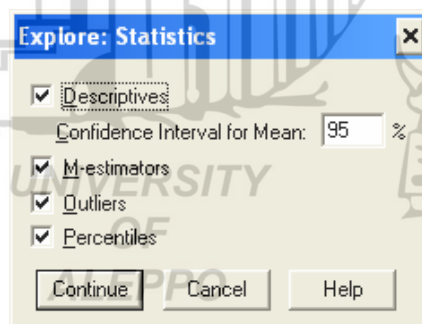
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6E+008	1	631109630	7.017	.020 ^a
	Residual	1E+009	13	89936900.3		
	Total	2E+009	14			

a. Predictors: (Constant), X2

b. Dependent Variable: Y



ننقل المتغير المطلوب إجراء التحليل عليه إلى نافذة الحوار المسماة Dependent List. والنافذة المسماة Factor List تستخدم لإجراء التحليل الإحصائي لعدة مجاميع من حالات المتغير المدروس وبصورة مستقلة (ويعتبر متغير تجزئة Break down variable). يمكن إعطاء تعريف للحالات بواسطة متغير معين وذلك باستخدام نافذة الحوار Label Cases by ويكون الغرض من التعريف إعطاء عنوان للقيم المتطرفة والشاذة في مخطط Box plot. بالنقر على الزر Statistics يظهر صندوق حوار بمسمى استكشاف الإحصائيات : Explore Statistics الذي يشتمل على الخيارات التالية:



- Descriptives لإظهار الإحصاءات الوصفية مثل الوسط الحسابي، الانحراف المعياري، الالتواء، التفرطح .
 - Confidence Interval for Means لتشكيل مجال ثقة حسب رغبة الباحث ٩٥% أو ٩٩%.

- M-estimators تحسب المقدرات الدقيقة والموزونة للنزعة المركزية مثل: Robust maximum Likelihood Estimators إذ يتم إعطاء القيم البعيدة عن المركز أوزان أقل من

٢-١ المخططات الوصفية Descriptives: يسمح الخيار Descriptive بعرض المخططات التالية:

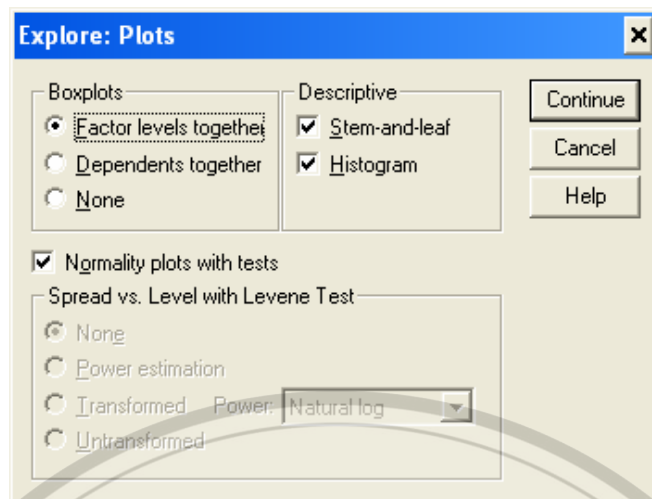
أولاً : مخطط الجذع والورقة Stem – and – Leaf: يتم في هذا المخطط قسمة أي مشاهدة (أي رقم) إلى جزأين : الأول هو الجذع Stem ويمثل الجزء الأيسر، والثاني الورقة Leaf ويمثل الجزء الأيمن. فإذا كانت لدينا القيم التالية :

60
46
42
40
32
30
24
14
10
57
8
4
54
19

نقوم بتقسيم كل رقم إلى جزأين الأول الجذع ويمثل خانة العشرات والثاني الورقة الذي يمثل خانة الأحاد، وكأن المتغير تم تقسيمه إلى فئات طول كل فئة منها ١٠ درجات.

VAR00014 Stem-and-Leaf Plot

Leaf &	Stem	Frequency
٤٨	. ٠	٢,٠٠
٠٤٩	. ١	٣,٠٠
٤	. ٢	١,٠٠
٠٢	. ٣	٢,٠٠
٠٢٦	. ٤	٣,٠٠
٤٧	. ٥	٢,٠٠
٠	. ٦	١,٠٠



إن المخطط **Boxplots** يعطي فكرة عن توزيع المشاهدات وفيما إذا كان التوزيع ملتوياً **Skewness** أو لا ؟ إذا كان الوسيط في منتصف الصندوق فإن توزيع المشاهدات يكون طبيعياً ، وإن لم يكن الوسيط في منتصف الصندوق فإن توزيع المشاهدات يكون ملتوياً .

- إذا كان الوسيط أقرب إلى الربع الأول فإن التوزيع يكون ملتوياً نحو اليمين ويسمى (التواء موجب) .

- إذا كان الوسيط أقرب إلى الربع الثالث فإن التوزيع يكون ملتوياً نحو اليسار ويسمى (التواء سالب) .

كما يشتمل المربع **Boxplots** الخيارات التالية:

- **Factor Levels together** : يسمح هذا الخيار بعرض المخطط للمجاميع الجزئية للمتغير التابع.

- **Dependent together** : يسمح هذا الخيار بعرض المخطط للمتغيرات التابعة داخل كل مجموعة.

- **None** : يسمح هذا الخيار بعدم عرض المخطط **Boxplot**.

ثانياً- **الاستطلاات خارج الصندوق Whiskers** : وتمتد من حافتي الصندوق إلى أعلى وأقل قيمة غير شاذة، إذ يعطي الخطان الخارجيان من الصندوق من الأعلى والأسفل (الاستطلاات Whiskers) فكرة عن طول طرفي التوزيع .

ثالثاً- **القيم الشاذة Outliers والقيم المتطرفة Extremes**: تبعد القيم الشاذة عن حافتي الصندوق أكثر من ١,٥ من طول الصندوق ، وتبعد القيم المتطرفة عن حافتي الصندوق أكثر من ٣ من طول الصندوق

القيم القريبة منه، وهناك أربعة أنواع من المقدرات هي: Tukey , Hample , Andrew .Huber.

- Outliers لإظهار أكبر ستة قيم وأصغر ست قيم للمتغير ويطلق عليها اسم القيم المتطرفة Extreme Values في مخرجات البرنامج.

- Percentiles لإظهار القيم المئينية الخامسة والعاشره فمثلاً 5th percentile هو قيم المتغير المدروس التي يسبقها 5% من القيم ويليهها 95% من القيم بعد ترتيب قيم المتغير تصاعدياً.

للحصول على هذه القيم نقوم بتفعيل المربعات المخصصة لكل منها في صندوق الحوار المسمى Explore : Plots

الذي نحصل عليه بالضغط على الزر Plots ، الذي يشتمل على الخيارات التالية:

١-١ **مخطط Boxplots** : ويطلق عليه تسمية Box-and- whisker plot الذي يتكون من ثلاثة أجزاء على النحو التالي:

أولاً- الصندوق Box : ويحتوي المكونات التالية :

الربيع الأول Q1 : وهي القيمة التي تسبقها 25% من قيم المتغير المدروس عند ترتيبها تصاعدياً.

الربيع الثاني Q2 : وهي القيمة التي تقسم البيانات إلى قسمين متساويين أي يسبقها 50% من قيم المتغير المدروس عند ترتيبها تصاعدياً، وهذه القيمة هي الوسيط تماماً.

الربيع الثالث Q3 : وهي القيمة التي تسبقها 75% من قيم المتغير المدروس عند ترتيبها تصاعدياً.

إن Q1 و Q3 تمثلان حافتي الصندوق ويطلق عليها hinges.

أما طول الصندوق Q3-Q1 ويطلق عليه hspread ويسمى المدى الربيعي Inter Quartile Range. أما الخط الأوسط داخل الصندوق فهو الوسيط Median.

Stem width: 10.00

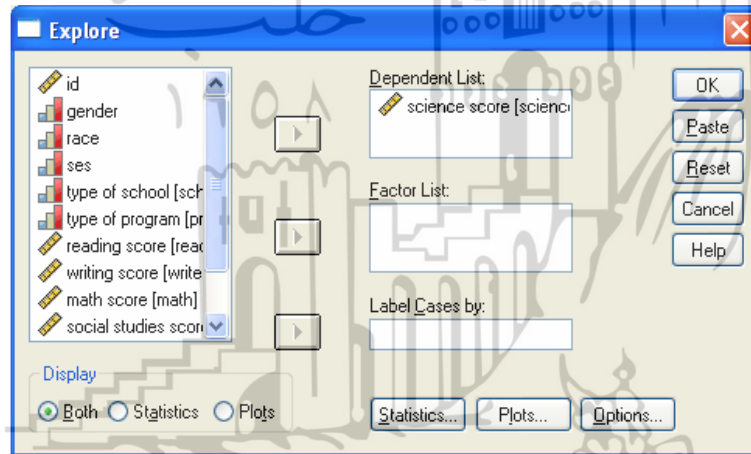
(Each leaf: 1 case(s)

يلاحظ أن هذا المخطط يشبه المدرج التكراري Histogram والفرق بينهما أن التكرارات في المدرج التكراري تمثل بوساطة أعمدة في حين تمثل بوسيقية في حالة الجذع والورقة وبذلك يعكس معلومات عن طبيعة القيم الموجودة .

لنأخذ المتغير درجات المواد العلمية science score الذي يشتمل على ٢٠٠ مشاهدة ، وكان المطلوب إنشاء مخطط الجذع والشجرة لهذا المتغير ، نتبع الخطوات التالية :
من شريط القوائم نختار :

Analysis → Descriptive statistics → Explore

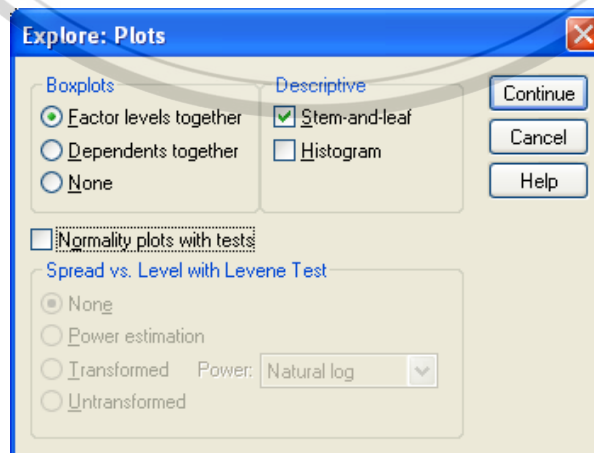
يظهر صندوق حوار يسمى Explore على النحو التالي:



- ننقل المتغير science score المطلوب إجراء التحليل عليه إلى النافذة Dependent List.

- ننقر الزر Plots نحصل على صندوق الحوار Explore: Plots.

- نفعل المربع Stem-and-Leaf ثم نضغط OK.



Descriptives

			Statistic	Std. Error
science score	Mean		51.8500	.70010
	95% Confidence Interval for Mean	Lower Bound	50.4694	
		Upper Bound	53.2306	
	5% Trimmed Mean		51.9556	
	Median		53.0000	
	Variance		98.028	
	Std. Deviation		9.90089	
	Minimum		26.00	
	Maximum		74.00	
	Range		48.00	
	Interquartile Range		14.00	
	Skewness		-.189	.172
	Kurtosis		-.556	.342

- الوسط الحسابي Mean = 51.85

- الانحراف المعياري std.Deviation = 9.90089

- الخطأ المعياري = std.Error

$$std.Error = SD / \sqrt{n} = 9.90089 / \sqrt{200} = 0.70010$$

يتم حساب الخطأ المعياري عن طريق قسمة الانحراف المعياري للعينة على جذر حجم العينة.

- تشكيل فترة ثقة 95 % لمتوسط المجتمع μ : يتم الحصول على فترة الثقة باستخدام العلاقة :

$$\bar{X} \pm t_{(0.025,199)} \cdot std.Error$$

إذ إن t تستخرج من جدول توزيع t لتودنت عند مستوى معنوية 0.025 ودرجات حرية: n-

1=199 أو يمكن استخراجها مباشرة من برنامج SPSS باستخدام الأمر :

Transform → Compute

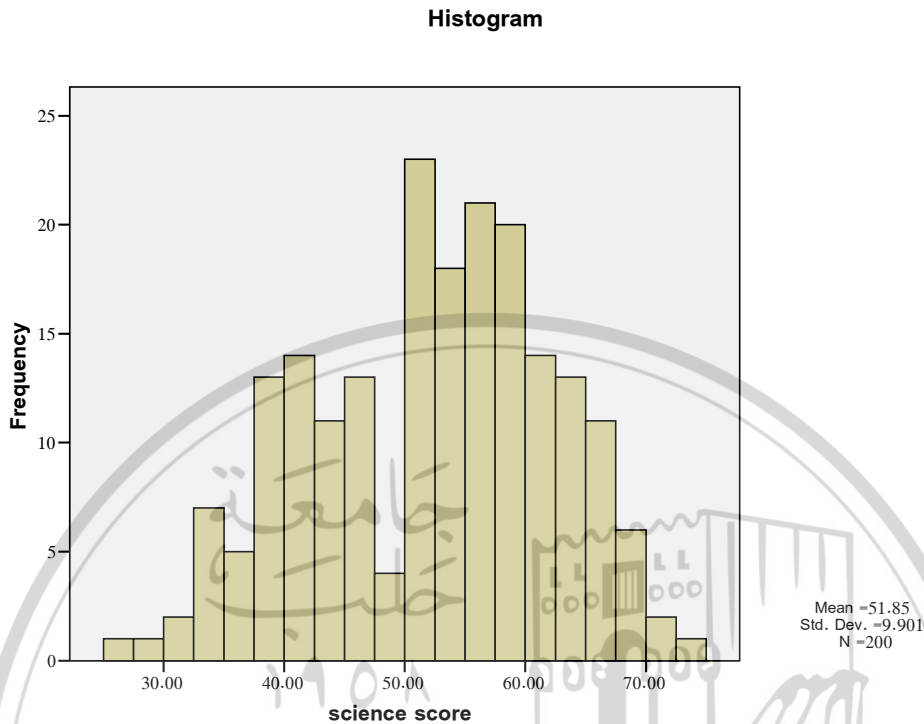
- من النافذة Function group نختار الدالة Inverse DF .

- من النافذة Functions and special variables نختار الدالة idf.t وبالنقر مرتين يتم نقلها

إلى النافذة المعنونة Numeric Expression ثم نكتب الدالة أو نكتب الوسطاء الرقمية للدالة

لتأخذ الصيغة التالية : $= idf.t(0.975,199)$

- نعطى اسماً للمتغير الهدف Target variable وليكن y .



٣-١ مخططات لاختبار التوزيع الطبيعي **Normality Plots with Tests** : يستخدم هذا الخيار لعمل مخططات لاختبار التوزيع الطبيعي للبيانات .

٤-١ اختبار تجانس التباين (اختبار ليفن) **Spread vs. Level with Levene Test** : يستخدم هذا الخيار لاختبار تجانس التباين في تصميم التجارب مع العلم أن هذا الاختبار لا يكون فعالاً إلا عندما يكون هناك متغير تجزئة Factor Variable .

- عند نقر الزر OK في صندوق الحوار Explore نحصل على النتائج التالية :
- ملخص الحالات التي تم معالجتها :

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
science score	200	100.0%	0	.0%	200	100.0%

- الإحصاءات الوصفية للمتغير المدروس:

science score Stem-and-Leaf Plot

Frequency	Stem & Leaf
1	1 0
1	1 1
1	1 2
1	1 3
1	1 4
1	1 5
1	1 6
1	1 7
1	1 8
1	1 9
1	2 0
1	2 1
1	2 2
1	2 3
1	2 4
1	2 5
1	2 6
1	2 7
1	2 8
1	2 9
1	3 0
1	3 1
1	3 2
1	3 3
1	3 4
1	3 5
1	3 6
1	3 7
1	3 8
1	3 9
1	4 0
1	4 1
1	4 2
1	4 3
1	4 4
1	4 5
1	4 6
1	4 7
1	4 8
1	4 9
1	5 0
1	5 1
1	5 2
1	5 3
1	5 4
1	5 5
1	5 6
1	5 7
1	5 8
1	5 9
1	6 0
1	6 1
1	6 2
1	6 3
1	6 4
1	6 5
1	6 6
1	6 7
1	6 8
1	6 9
1	7 0
1	7 1
1	7 2
1	7 3
1	7 4
1	7 5
1	7 6
1	7 7
1	7 8
1	7 9
1	8 0
1	8 1
1	8 2
1	8 3
1	8 4
1	8 5
1	8 6
1	8 7
1	8 8
1	8 9
1	9 0
1	9 1
1	9 2
1	9 3
1	9 4
1	9 5
1	9 6
1	9 7
1	9 8
1	9 9

[illegible]

Stem width: 10.00

(Each leaf: 1 case(s

ثانياً : المدرج التكراري Histogram: يسمح هذا الخيار بعرض المدرج التكراري للمتغير المدروس .

- نضغط على الزر OK .

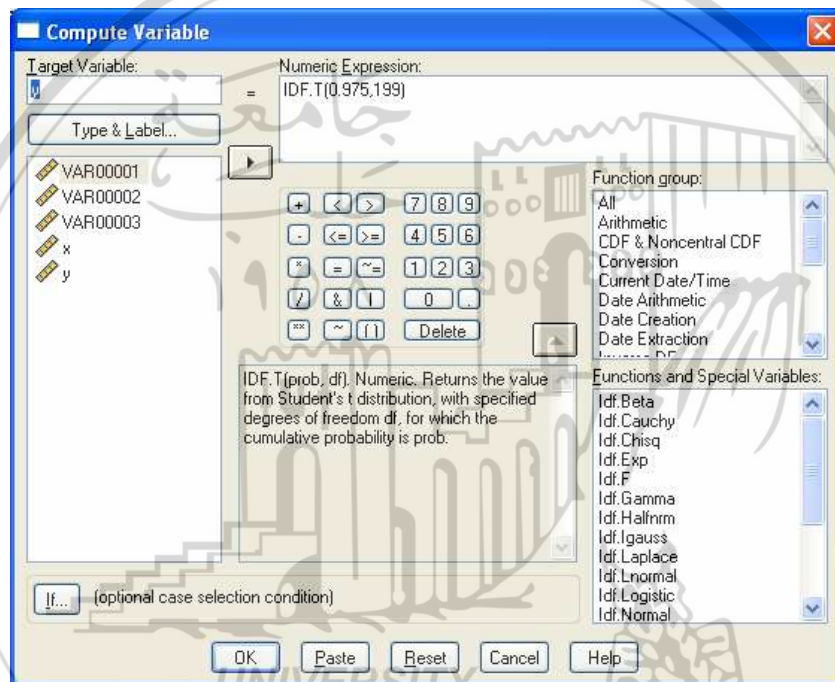
نحصل على النتيجة في عمود يسمى y وتكون النتيجة مساوية إلى ١,٩٧ .
حيث أن :

$$p = 1 - 0.025 = 0.975$$

$$df = n - 1 = 200 - 1 = 199$$

كما يمكن الحصول على قيمة t لستودنت باستخدام برنامج Excel وذلك بواسطة الدالة :

$$tinv(0.05;199) = 1.971956$$



وبناء على ما تقدم تكون حدود فترة الثقة كما يلي :

$$\text{الحد الأعلى لفترة الثقة : } 51.85 + 1.971956(0.70010) = 53.23057$$

$$\text{الحد الأدنى لفترة الثقة : } 51.85 - 1.971956(0.70010) = 50.46943$$

وتكتب فترة الثقة على النحو التالي :

$$Pr(50.46943 < \mu < 53.23057) = 95\%$$

أي إن احتمال وقوع متوسط المجتمع بين القيمتين ٥٠,٤٦٩٤٣ و ٥٣,٢٣٠٥٧ يساوي ٩٥ % .

- الوسط الحسابي المشذب **Trimed Mean = 51.9556** :

يتم حسابه بعد ترتيب القيم تصاعدياً ثم حذف ٥ % من القيم من الأعلى و ٥ % من القيم من الأسفل ثم حساب المتوسط للقيم المتبقية الذي لن يتأثر بالقيم الشاذة .

٦٥، وعليه يمكن الحصول على قيمة المئين التسعين بإجراء عملية الاستكمال الخطي Interpolation بين القيمتين المستخرجتين على النحو التالي :

$$90th\ Percentile = 64(0.10) + 65(0.90) = 64.90$$

٥-١ اختبار فرضية كون البيانات تتبع التوزيع الطبيعي من نسبة معامل الالتواء:

يمكن اختبار فرضية كون البيانات تتبع التوزيع الطبيعي عن طريق ملاحظة حاصل قسمة معامل الالتواء على الخطأ المعياري لمعامل الالتواء ، وفي مثالنا كان معامل الالتواء:

$$Skewness = -0.189 \text{ أما الخطأ المعياري لمعامل الالتواء : } std. Error = 0.172$$

$$Ratio = \frac{Skewness}{Std. Error} = \frac{-0.189}{0.172} = -1.09884$$

بما أن هذه النسبة لا تقع ضمن المجال $[-2, +2]$ إذن لا يمكننا قبول فرضية العدم القائلة بأن البيانات أو المتغير المدروس Science Score يتبع التوزيع الطبيعي. أما إذا كانت النسبة أكبر من $+2$ فهذا يعني أن التوزيع ملتو نحو اليمين أي ملتو إلتواءً موجباً ، وإذا كانت النسبة أصغر من -2 فهذا يعني أن التوزيع ملتو نحو اليسار أي ملتو إلتواءً سالباً. وعليه فإن توزيع البيانات ملتو نحو اليسار.

والجدول التالي يمثل مخرجات برنامج SPSS بالنسبة للمتغير Science Score فيما يتعلق بتقديرات الوسط الحسابي بالطرق الأربعة .

M-Estimators

	Huber's M-Estimator ^a	Tukey's Biweight ^b	Hampel's M-Estimator ^c	Andrews' Wave ^d
science score	52.2102	52.2260	52.0555	52.2245

- The weighting constant is 1.339.
- The weighting constant is 4.685.
- The weighting constants are 1.700, 3.400, and 8.500
- The weighting constant is $1.340 \cdot \pi$.

يتضح من الجدول وجود فروق ذات أهمية بين الوسط الحسابي المحسوب بموجب الطرق الأربع والذي يساوي تقريباً ٥٢,٢ درجة والوسط الحسابي للمتغير Science Score المحسوب بالطريقة الاعتيادية والذي يساوي ٥١,٨٥ .

- **الربيعات والمئينات Quartiles & Percentiles**: تتكون الربيعات Quartiles من ثلاثة قيم هي الربع الأول والربع الثاني والربع الثالث Q_1 ، Q_2 ، Q_3 وتقسم المجموعة إلى أربعة أقسام متساوية . أما المئينات Percentiles فتقسم المجموعة إلى مئة قسم متساوي فالمئين الخامس تسبقه ٥% من البيانات عند ترتيبها تصاعدياً وتليها ٩٥% من البيانات، المئين الخامس عشر تسبقه ١٥% من البيانات عند ترتيبها تصاعدياً وتليها ٨٥% من البيانات الخ، مع العلم بأن الربع الأول Q_1 يقابل المئين الخامس والعشرين (٢٥th Percentile) ، والربع الثاني Q_2 (الوسيط) يقابل المئين الخمسين (٥٠th Percentile) ، والربع الثالث Q_3 يقابل المئين الخامس والسبعين (٧٥th Percentile) . والجدول التالي يمثل مخرجات برنامج SPSS بالنسبة للمتغير Science Score

		Percentiles						
		5	10	25	50	75	90	95
Weighted science score	Average(Definition)	34.0000	39.0000	44.0000	53.0000	58.0000	64.9000	66.9500
Tukey's Hinges science score				44.0000	53.0000	58.0000		

وعليه فإن الربع الأول Q_1 يساوي ٤٤ ، والربع الثاني أي الوسيط Q_2 يساوي ٧٠ ، والربع الثالث ، Q_3 يساوي ٥٨ . أما المدى الربيعي فيساوي ١٤ أي :

$$Inter\ Quartile\ Range = Q_3 - Q_1 = 58 - 44 = 14$$

تستخدم طريقة المتوسط الموزون Weighted Average Method لحساب المئينات ، فمثلاً تستخدم الصيغة التالية لحساب ترتيب المئين الخامس : $(n + 1).P = 201(0.05) = 10.05$ إن ترتيب المئين الخامس هو ١٠,٠٥ عند ترتيب القيم تصاعدياً وعليه فإن قيمة المئين الخامس تقع بين القيمة التي ترتيبها ١٠ وهذه القيمة ٣٤ والقيمة التي ترتيبها ١١ وهذه القيمة تساوي ٣٤ أيضاً ، وعليه يمكن الحصول على قيمة المئين الخامس بإجراء عملية الاستكمال الخطي Interpolation بين القيمتين المستخرجتين على النحو التالي:

$$5th\ Percentile = 34(0.95) + 34(0.05) = 34$$

$$\text{لحساب ترتيب المئين التسعين : } (n + 1).P = 201(0.90) = 180.9$$

إن ترتيب المئين رقم ٩٠ هو ١٨٠,٩ عند ترتيب القيم تصاعدياً وعليه فإن قيمة المئين التسعين تقع بين القيمة التي ترتيبها ١٨٠ وهذه القيمة ٦٤ والقيمة التي ترتيبها ١٨١ وهذه القيمة تساوي

في المثال عدد القيم للمتغير المدروس هو ٢٠٠ وأن ٥% من القيم يساوي ١٠ أي أنه يجب حذف ١٠ قيم من الأعلى و ١٠ قيم من الأسفل، أما القيم المتبقية بعد الحذف فهي

$$200 * 0.90 = 180$$

حيث تستخرج قيمة الوسط الحسابي المشذب على النحو التالي:

$$\text{Trimmed Mean} = \frac{51.85 * 200 - 699 - 319}{180} = 51.95556$$

إذ أن القيمة ٥١,٨٥ تمثل الوسط الحسابي لـ ٢٠٠ قيمة وأن القيمة ٦٩٩ تمثل مجموع القيم العشرة الأعلى والقيمة ٣١٩ تمثل مجموع القيم العشرة الأدنى .

Extreme Values

			Case Number	Value
science score	Highest	1	74	74.00
		2	22	72.00
		3	34	72.00
		4	46	69.00
		5	55	69.00 ^a
	Lowest	1	56	26.00
		2	130	29.00
		3	173	31.00
		4	15	31.00
		5	103	33.00 ^b

- a. Only a partial list of cases with the value 69.00 are shown in the table of upper extremes.
b. Only a partial list of cases with the value 33.00 are shown in the table of lower extremes.

القيم العشرة من الأعلى	القيم العشرة من الأسفل
67	26
69	29
69	31
69	31
69	33
69	33
69	34
72	34
72	34
74	34

[illegible]

Analysis → Descriptive statistics → Frequencies

The screenshot shows the 'Frequencies' dialog box in SPSS. The variable 'science score [science]' is selected in the 'Variable(s):' list. The 'Display frequency tables' checkbox is checked. The dialog box has a blue title bar with the text 'Frequencies' and a large 'ALEPPO' watermark in the background. The list of variables on the left includes 'id', 'gender', 'race', 'ses', 'type of school [sch', 'type of program [pr', 'reading score [read', 'writing score [write', and 'math score [math1'. The 'OK', 'Paste', 'Reset', 'Cancel', and 'Help' buttons are on the right. At the bottom, there are buttons for 'Statistics...', 'Charts...', and 'Format...'.

Frequencies: Charts

- نضغط Continue نعود إلى صندوق الحوار السابق Explore .
- نضغط الزر OK نحصل على الجدول التالي:

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
science score	.076	200	.007	.985	200	.035

a. Lilliefors Significance Correction

حيث تستخدم الإحصائية D في الاختبار والتي تحسب بموجب الصيغة التالية :

$$D = \sup_x |F_S(x) - F_T(x)|$$

إذ إن :

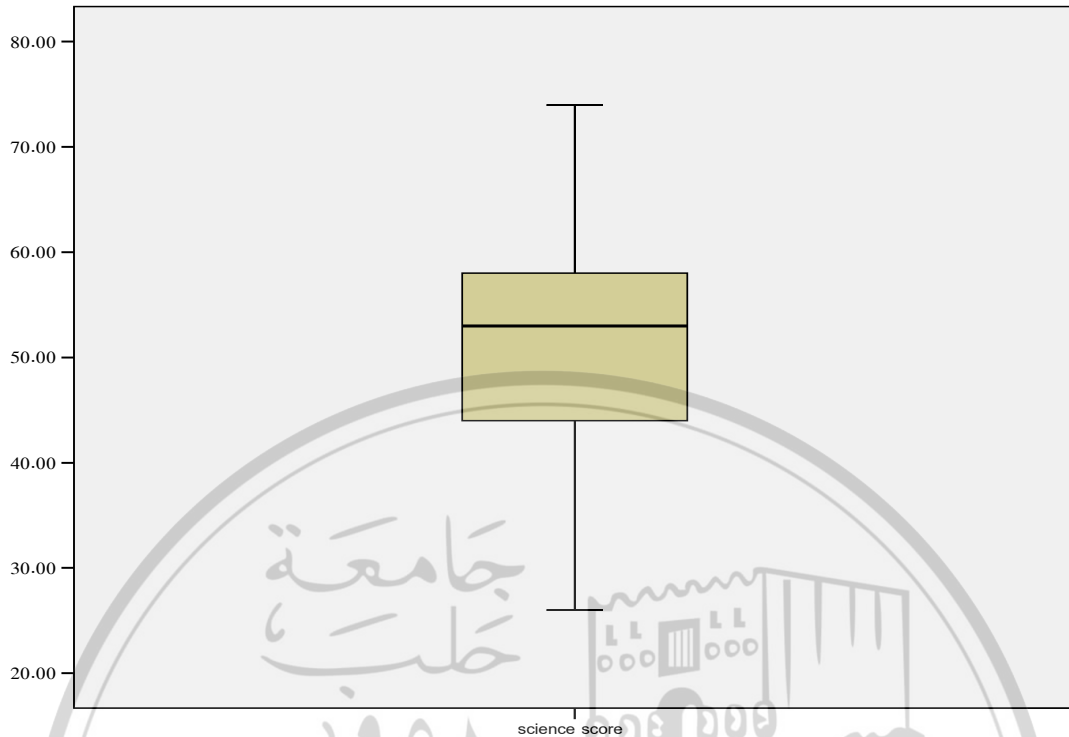
$F_S(x)$: تمثل تابع التوزيع التجمعي للعينة .

$F_T(x)$: تمثل تابع التوزيع النظري أو تابع التوزيع الطبيعي .

تقارن القيمة المحسوبة لإحصائية الاختبار مع القيمة النظرية أو القيمة الحرجة المستخرجة من جدول كولموغروف عند مستوى معنوية محدد ودرجات حرية تساوي حجم العينة . .
في مثالنا نجد أن قيمة إحصائية الاختبار المحسوبة $D = 0.076$ هي أكبر من قيمة مستوى المعنوية الفعلي التي تساوي $P\text{-Value} = \text{Sig} = 0.007$ مما يدعونا لرفض فرضية العدم القائلة إن المتغير المدروس يتبع التوزيع الطبيعي عند مستوى معنوية ٠,٠٥ .

٢- مخطط Normal Q-Q Plot :

للحصول على هذا المخطط يتم رسم كل قيمة من قيم البيانات الأصلية على المحور الأفقي مقابل قيم التوزيع الطبيعي القياسي المتوقعة لها Expected Z Score كما هو موضح في الشكل البياني أدناه الخاص بالمتغير science score.



٦-١ اختبارات التوزيع الطبيعي : Normality Plots with Tests

يستخدم هذا الخيار لاختبار التوزيع الطبيعي للبيانات باستخدام:

- اختبار كولموغوروف سميرنوف Kolmogrov-Samirnov .

بالإضافة إلى اختبار التوزيع الطبيعي للبيانات باستخدام نوعين من المخططات وهما:

- Normal Q-Q Plot .

- Detrended Normal Q-Q Plot .

١- اختبار كولموغوروف سميرنوف Kolmogrov-Samirnov :

يعتبر من الاختبارات اللامعلمية لقياس جودة توفيق البيانات بواسطة التوزيع الطبيعي

Non-ParaMETRIC Goodness of Fit Test حيث يتم اختبار فرضية عدم القائلة إن

البيانات تتبع التوزيع الطبيعي مقابل الفرضية البديلة القائلة إن المتغير المدروس لا يتبع التوزيع

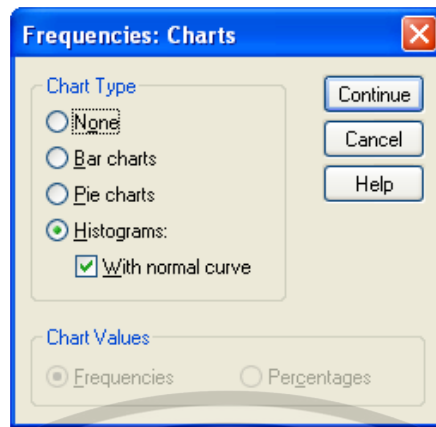
الطبيعي (البيانات لا تتبع التوزيع الطبيعي) ، ونحصل على نتائج هذا الاختبار كما يلي:

من شريط القوائم نختار :

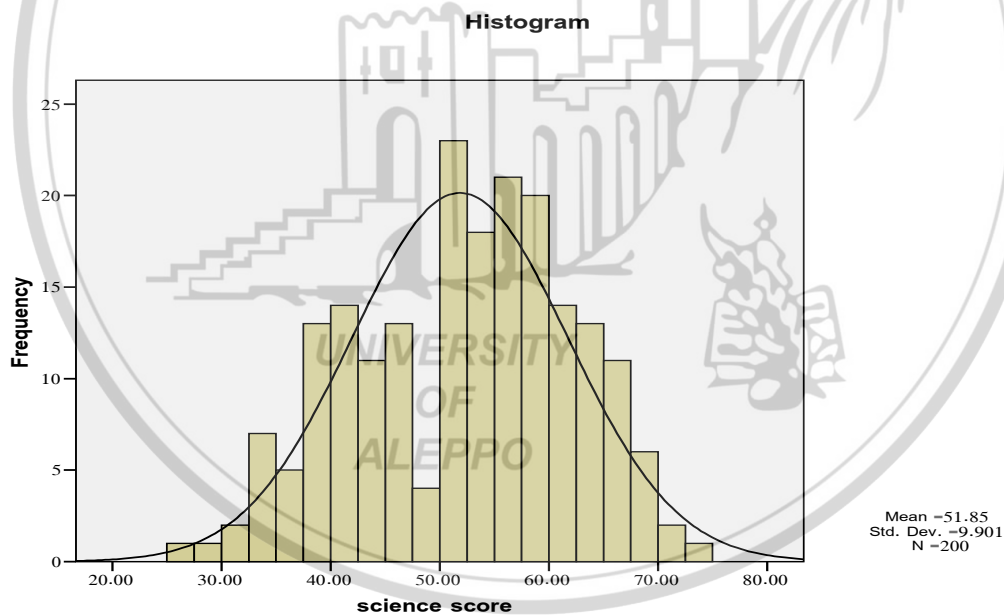
Analysis → Descriptive statistics → Explore

يظهر صندوق حوار بمسمى Explore.

- نُفعل الخيار Normality Plots with Tests .



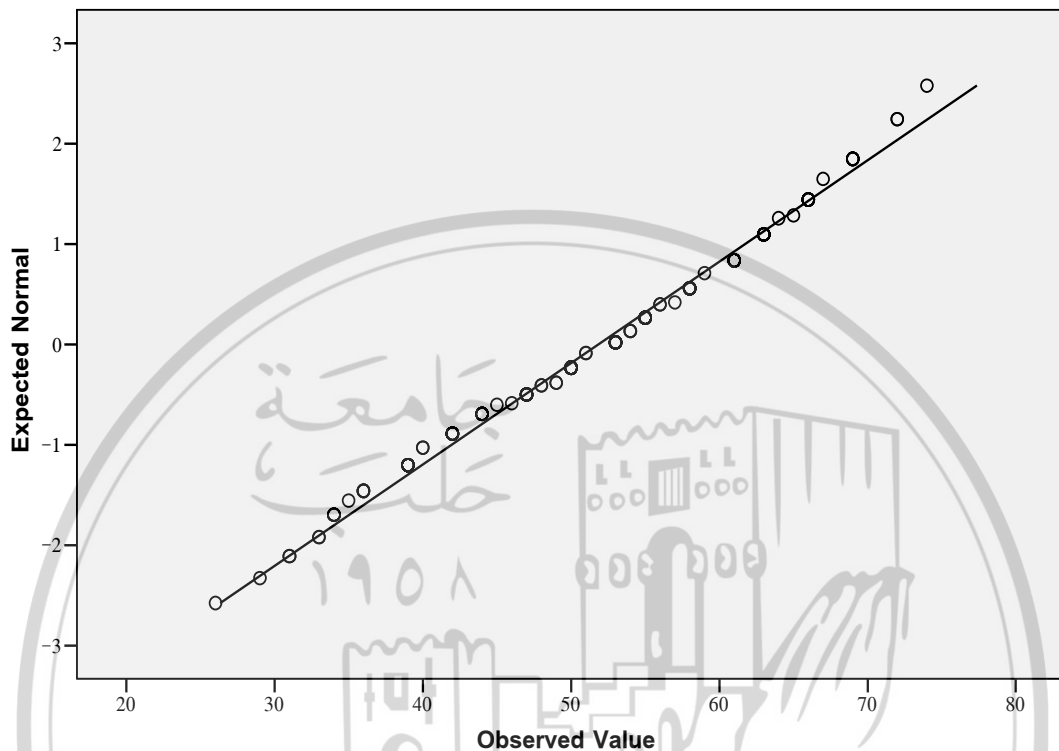
- نُفعل الخيار Histogram مع إضافة منحنى التوزيع الطبيعي With normal curve .
- نضغط Continue نعود إلى صندوق الحوار السابق Frequencies .
- نضغط الزر OK نحصل على المدرج التكراري مع إضافة منحنى التوزيع الطبيعي الذي لا يبدو متماثلاً إلى حد ما حول الوسط الحسابي للبيانات.



- مخطط Boxplots :

يمثل المخطط التالي **مخطط Boxplots** لمتغير science score حيث نلاحظ عدم وجود قيم متطرفة أو قيم شاذة ، كما أن موقع الوسيط لا يقع في منتصف الصندوق مما يشير إلى حد ما إلى أن هذا المتغير لا يتبع التوزيع الطبيعي (بسبب وجود التواء سالب كون الوسيط يقترب من الربع الثالث للبيانات) .

Normal Q-Q Plot of science score



إذا كانت البيانات تتبع التوزيع الطبيعي أي إذا كانت العينة مسحوبة من مجتمع يتبع التوزيع الطبيعي فإن نقاط شكل الانتشار ستقع تقوياً بمحاذاة الخط المستقيم، أما إذا كانت النقاط تقع بعيدة عن الخط المستقيم فهذا دليل على أن العينة غير مسحوبة من مجتمع طبيعي أي أن البيانات لا تتبع التوزيع الطبيعي.

٣- مخطط Detrended Normal Q-Q Plot :

للحصول على هذا المخطط يتم تمثيل قيم البيانات الأصلية الطبيعية على المحور الأفقي ، أما انحرافات القيم المعيارية للبيانات عن قيم التوزيع الطبيعي المتوقعة لها فتتمثل على المحور العمودي، فإذا كانت معظم نقاط الانتشار (٩٠ % أو ٩٥ % من قيم البيانات) تقع ضمن المجال $[-0.2, +0.2]$ فيمكن أن نقول بأن البيانات تتبع التوزيع الطبيعي.

في الشكل البياني التالي نجد أن قسماً كبيراً من النقاط لا تقع ضمن المجال $[-0.2, +0.2]$ وعليه يمكن القول إن المتغير المدروس لا يخضع للتوزيع الطبيعي.

ب - أسلوب جدول التقاطع متعددة الاستجابة Multiple Response Crosstabs ومنه نحصل على جداول تكرارية ذات بعدين أو ذات ثلاثة أبعاد.

٨-١ أسلوب التكرارات متعددة الاستجابة Multiple Response Frequencies

مثال: في استبيان حول المناهج والاختصاصات الجامعية موجه لطلبة السنوات الأخيرة ولأعضاء الهيئة التدريسية في الجامعات الحكومية السورية كان أحد الأسئلة على النحو التالي: المصدر المفضل للمعلومات المتعلقة باختصاصك حسب رأي عضو هيئة التدريس : (المرجع الأجنبي ، الكتاب الجامعي ، المرجع العربي ، الانترنت ، المحاضرات المكتوبة ، الأملية) . لمعرفة التكرارات الخاصة بكل مصدر يمكننا الاختيار بين أسلوبين:

١-١ أسلوب المتغيرات المتعددة ذات الفئتين Multiple Dichotomy Method

بموجب هذه الطريقة نخصص متغير لكل مصدر من المصادر الستة يأخذ قيمتين فقط وليكن الصفر والواحد ، فإذا كان المستجوب يفضل المصدر المعين فإننا نعطي القيمة ١ لمتغير المصدر المفضل وصفر فيما عدا ذلك ، ونبين فيما يلي طريقة ترتيب البيانات في شاشة محرر البيانات Data Editor .

Quest7a المرجع الأجنبي	Quest7b الكتاب الجامعي	Quest7c المرجع العربي	Quest7d الانترنت	Quest7e المحاضرات المكتوبة	Quest7f الأملية
1	0	1	0	1	0
0	0	1	0	0	0
0	1	1	0	0	0
0	0	0	1	1	0
1	0	0	0	1	0
1	0	1	0	0	0
1	1	0	0	0	0

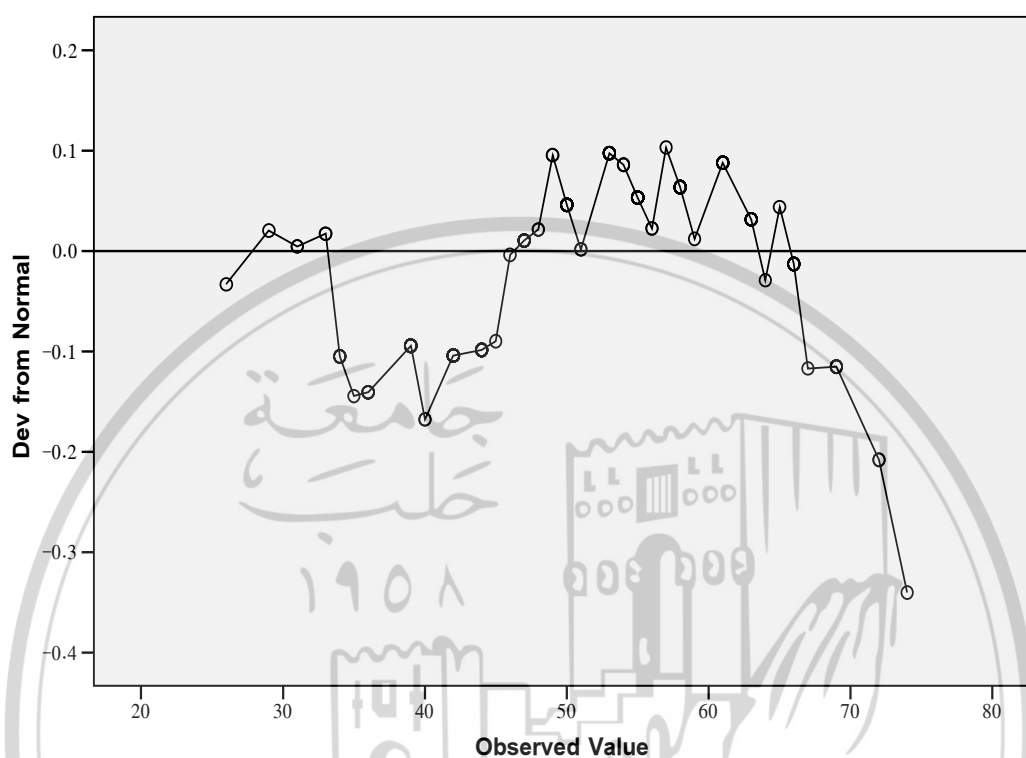
لإيجاد التكرارات المقابلة لكل مصدر نتبع الخطوات التالية :

- من شريط القوائم في محرر البيانات نختار:

Analysis → Multiple Response → Define variable sets

يظهر صندوق حوار بعنوان Define Multiple Response sets

Detrended Normal Q-Q Plot of science score



الفصل السادس

تحليل الاستجابات المتعددة

Multiple Response Analysis

مقدمة:

في بعض الحالات قد يكون المطلوب من المستجوب (مع إفتراض أن الأشخاص الذين شملهم الاستقصاء هم طلبة السنوات الأخيرة في الجامعات السورية) الإجابة عن سؤال يمثل مدى استعداده للتحويل إلى أساليب تعلم جديدة (أي التخلي عن أسلوب التلقين) واختيار واحدة من أساليب التعلم المقترحة :

✓ أسلوب التعلم التشاركي.

✓ أسلوب التعلم التفاعلي.

✓ أسلوب التعلم الإلكتروني .

✓ أسلوب التعلم الذاتي

✓ غير مستعد للتحويل .

فإن بإمكانه اختيار أحد أساليب التعلم المقترحة (اختيار واحدة فقط من الخيارات المطروحة) أو الإبقاء على أسلوب التلقين في حالة عدم استعداده للتحويل.

أما إذا كان المطلوب من المستجوب الإجابة على سؤال وهناك احتمال أن يختار أكثر من إجابة في الوقت نفسه ، فمثلاً إذا كان السؤال المطروح على طلبة السنوات الأخيرة على النحو التالي:

النظام الامتحاني الحالي يختبر مقدرة الطالب :

✓ على حل المشاكل .

✓ على التحليل .

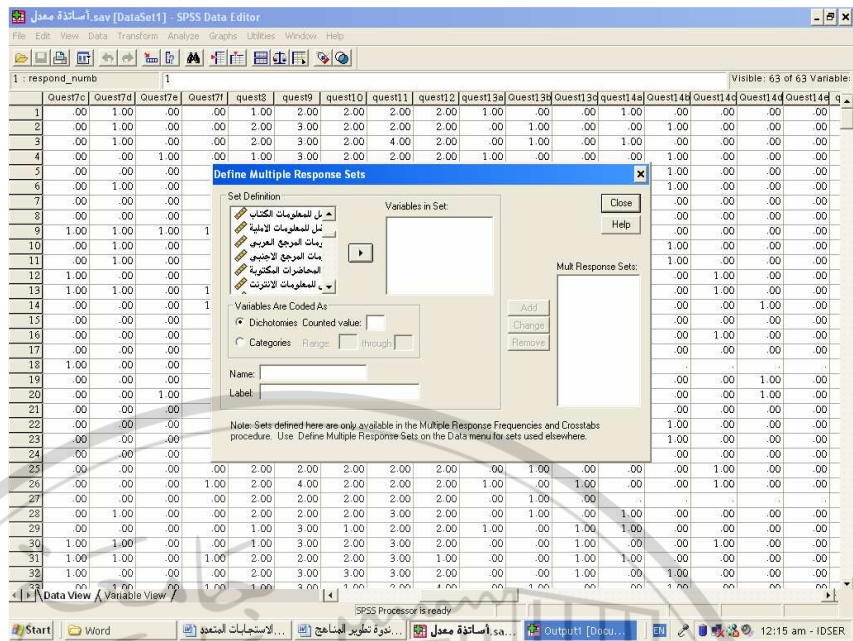
✓ على تقدير وجهات النظر .

✓ على تقبل الأفكار

✓ يشجع الطالب على الإبداع.

في هذه الحالة فإن إيجاد التكرارات Frequencies يتم عن طريق تحليل الاستجابات المتعددة وهناك أسلوبين في التحليل:

أ- أسلوب التكرارات متعددة الاستجابة Multiple Response Frequencies ومنه نحصل على جدول تكراري بسيط.



الذي نرتبه بإتباع الخطوات التالية:

- اختيار المتغيرات الستة من النافذة Set Definition ونقلها إلى النافذة Variable in set .
- كتابة الرقم ١ بعد اختيار المربع Dichotomies Counted Value .
- إعطاء اسم لمجموعة الاستجابة المتعددة يتكون من سبعة رموز كحد أعلى في المربع Name .
- ولكن هذا الاسم Information_preference_Source .
- ننقر الزر إضافة Add لإضافة المجموعة إلى خانة Mult Response sets .
- عند نقر الزر إغلاق Close في صندوق الحوار يتم تخزين مجموعة الاستجابة المتعددة باسم Infor_pref_Source للاستفادة منها فيما بعد بعملية التحليل.

X ₁	X ₂	X ₃
1	4	6
5	6	.
3	.	.
1	4	5
1	3	.
3	5	.
2	.	.
1	6	.
5	.	.

لإيجاد التكرارات المقابلة لكل مصدر (لكل رمز أو كود) نتبع الخطوات التالية :
- من شريط القوائم في محرر البيانات نختار:

Analysis → Multiple Response → Define variable sets

يظهر صندوق حوار بعنوان Define Multiple Response sets

الذي نرتبه بإتباع الخطوات التالية:

- اختيار المتغيرات الثلاثة من النافذة Set Definition ونقلها إلى النافذة Variable in set .
- اختيار Categories من النافذة Variables are coded as وتعيين فئات مصادر المعلومات من ١ إلى ٦ (Range 1 through 6) .

\$Infor_Pref_Source Frequencies

	Responses		Percent of Cases
	N	Percent	
\$Infor_Pref_Source\$ المصدر المفضل للمعلومات الكتاب	226	20.8%	34.5%
المصدر المفضل للمعلومات الأملية	58	5.3%	8.8%
المصدر المفضل للمعلومات المرجع العربي	184	16.9%	28.0%
المصدر المفضل للمعلومات المرجع الاجنبي	364	33.5%	55.5%
المصدر المفضل للمعلومات المحاضرات المكتوبة	82	7.5%	12.5%
المصدر المفضل للمعلومات الانترنت	174	16.0%	26.5%
Total	1088	100.0%	165.9%

a. Dichotomy group tabulated at value 1.

إن المتغيرات التي ظهرت في شاشة محرر البيانات , quest7a , quest7b , quest7c , quest7d , quest7e , quest7f , Infor_pref_Source فتعرف باسم مجموعة الاستجابات المتعددة Multiple Response . sets

٢-١ أسلوب الفئات المتعددة Multiple Category Method

بموجب هذه الطريقة يتوجب على المحلل معرفة أكبر عدد من الإجابات على السؤال (أي أكبر عدد من مصادر المعلومات المفضلة لدى المستجوب) ، نفترض أن هذا العدد هو ثلاثة مصادر كحد أعلى وفي المقابل نكوّن العدد نفسه من المتغيرات ولتكن X_1 , X_2 , X_3 حيث أن قيم كل منها معرفة على النحو التالي :

($X_1=1$ المرجع الأجنبي ، $X_2=2$ الكتاب الجامعي ، $X_3=3$ المرجع العربي ، $X_4=4$ الانترنت ، $X_5=5$ المحاضرات المكتوبة ، $X_6=6$ الأملية)

فمثلاً إذا كان المستجوب يفضل ثلاثة مصادر للمعلومات هي المرجع الأجنبي، الإنترنت، الأملية، فإن المتغير X_1 سيأخذ القيمة ١ ، والمتغير X_2 سيأخذ القيمة ٤ ، والمتغير X_3 سيأخذ القيمة 6 ، أما إذا كان المستجوب يفضل المحاضرات المكتوبة والأملية فقط فإننا سنخصص المتغير X_1 بالقيمة ٥ ، والمتغير X_2 بالقيمة ٦ ، أما المتغير X_3 سة مفقودة Missing Value . ونبين فيما يلي طريقة ترتيب البيانات

في شاشة محرر البيانات Data Editor .

Define Multiple Response Sets

Set Definition

Variables in Set:

Variables Are Coded As:

☒ Dichotomies Counted value:

☐ Categories Range: through

Name:

Label:

Mult Response Sets:

\$Infor_Pref_Source

Add Change Remove

Note: Sets defined here are only available in the Multiple Response Frequencies and Crosstabs procedure. Use Define Multiple Response Sets on the Data menu for sets used elsewhere.

- من شريط القوائم نختار Analysis → Multiple Response → Frequencies يظهر صندوق الحوار Multiple Response Frequencies الذي نقوم من خلاله بنقل مجموعة الاستجابة Infor_pref_Source إلى خانة Table(s) for .

Multiple Response Frequencies

Mult Response Sets:

Table(s) for:

\$Infor_Pref_Source

OK Paste Reset Cancel Help

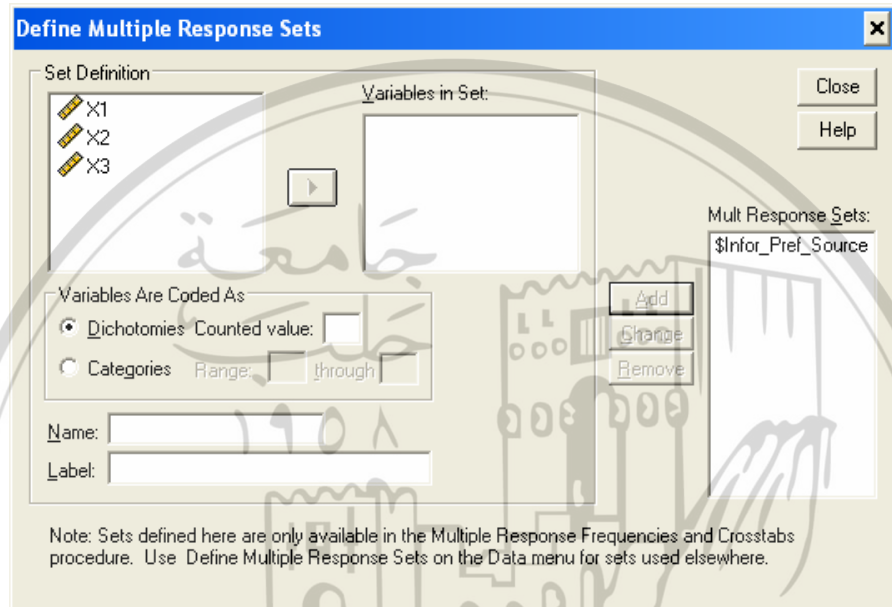
Missing Values

☐ Exclude cases listwise within dichotomies

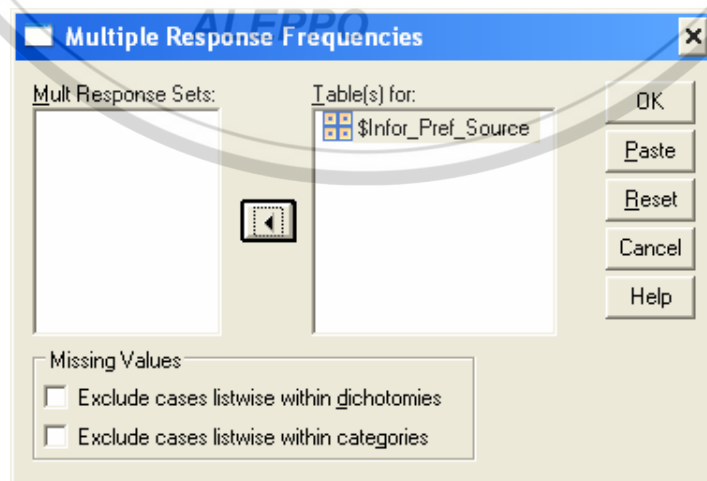
☐ Exclude cases listwise within categories

عند نقر الزر OK نحصل على جدول التوزيع التكراري البسيط الذي يبين عدد التكرارات الخاصة بكل مصدر من مصادر المعلومات والتكرارات النسبية للاستجابات بالإضافة إلى التكرارات النسبية للحالات .

- إعطاء اسم لمجموعة الاستجابة المتعددة في المربع Name وليكن هذا الاسم Infor_pref_Source (ويمكن إعطاء عنوان Label للمجموعة اختياريًا).
- ننقر الزر إضافة Add لإضافة المجموعة إلى خانة Mult Response sets.
- عند نقر الزر إغلاق Close في صندوق الحوار يتم تخزين مجموعة الاستجابة المتعددة باسم Infor_pref_Source للاستفادة منها فيما بعد بعملية التحليل.



- من شريط القوائم نختار Analysis → Multiple Response → Frequencies يظهر صندوق الحوار Multiple Response Frequencies الذي نقوم من خلاله بنقل مجموعة الإستجابة Infor_pref_Source إلى خانة Table(s) for.

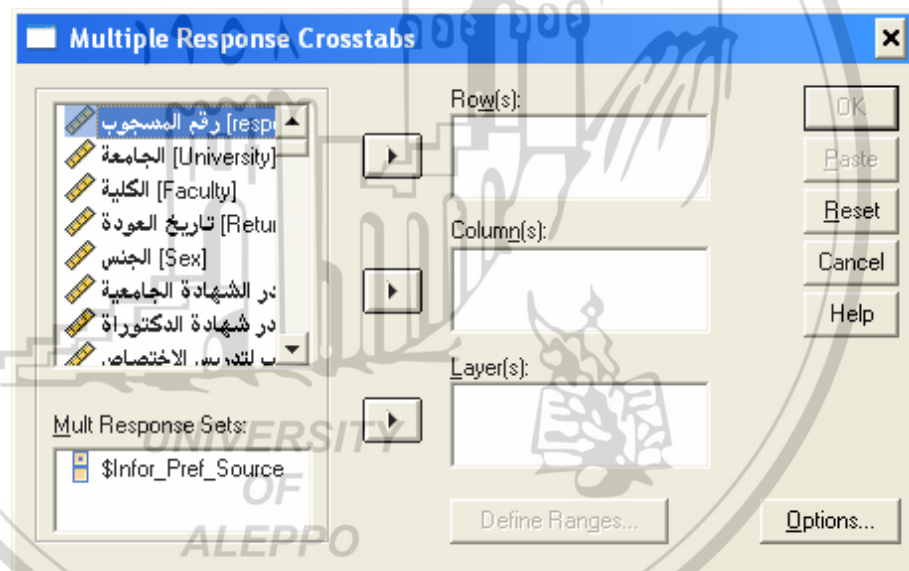


- نختار المتغيرات الستة من النافذة Set Definition ونقلها إلى النافذة Variable in set .
- نكتب الرقم ١ بعد اختيار المربع Dichotomies Counted Value .
- إعطاء اسم لمجموعة الاستجابة المتعددة وليكن هذا الاسم Infor_pref_Source .
- ننقر الزر إضافة Add لإضافة المجموعة إلى خانة Mult Response sets .
- عند نقر الزر إغلاق Close في صندوق الحوار يتم تخزين مجموعة الاستجابة المتعددة باسم Infor_pref_Source للاستفادة منها فيما بعد بعملية التحليل.
- من شريط القوائم في محرر البيانات نختار:

Analysis → Multiple Response → Crosstabs

يظهر صندوق حوار بعنوان Multiple Response Crosstabs

- نختار المتغير Infor_pref_Source من النافذة Mult Response sets ونقله إلى النافذة المعنونة الصفوف Rows .



- نختار المتغير Sex من النافذة Multiple Response Crosstabs ونقله إلى النافذة المعنونة الأعمدة Column(s) .
- تحديد الحدين الأعلى والأدنى لفئات المتغير الأولي الجنس وهي (١ و ٢) ويتم ذلك بالنقر على الزر Define Ranges في النافذة المعنونة الأعمدة Column(s) .

\$Infor_Pref_Source Frequencies

	Responses		Percent of Cases
	N	Percent	
\$Infor_Pref_Source ^a المرجع الأجنبي	4	23.5%	44.4%
الكتاب الجامعي	1	5.9%	11.1%
المرجع العربي	3	17.6%	33.3%
الانترنت	2	11.8%	22.2%
المحاضرات المكتوبة	4	23.5%	44.4%
الأملية	3	17.6%	33.3%
Total	17	100.0%	188.9%

a. Group

§ ٢- أسلوب جداول التقاطع متعددة الإستجابة Multiple Response Crosstabs

تستخدم هذه الطريقة لتكوين جداول تقاطع للمتغيرات الأولية Elementary Variables أو للمجاميع المعرفة Multiple Response sets أو لكليهما معاً .
 لنفترض في مثالنا السابق، أننا نرغب في تكوين جدول توزيع تكراري مزدوج (جدول تقاطع) يبين تقاضيات المستجوبين لمصادر المعلومات تبعاً لمتغير جديد وهو جنس المستجوب الذي يأخذ قيمتين : القيمة ١ للذكور والقيمة ٢ للإناث .

Quest7a المرجع الأجنبي	Quest7b الكتاب الجامعي	Quest7c المرجع العربي	Quest7d الانترنت	Quest7e المحاضرات المكتوبة	Quest7f الأملية	Sex	University
1	0	1	0	1	0	2	5
0	0	1	0	0	0	1	4
0	1	1	0	0	0	2	3
0	0	0	1	1	0	1	1
1	0	0	0	1	0	1	2
1	0	1	0	0	0	1	2
1	1	0	0	0	0	2	3

لتكوين جدول التقاطع بحيث تكون فئات التفضيل للمصادر في الصفوف ومتغير الجنس في الأعمدة نتبع الخطوات التالية :

- من شريط القوائم في محرر البيانات نختار:

Analysis → Multiple Response → Define variable sets

يظهر صندوق حوار بعنوان Define Multiple Response sets

عند نقر الزر OK نحصل على جدول التوزيع التكراري البسيط الذي يبين عدد التكرارات الخاصة بكل مصدر من مصادر المعلومات والتكرارات النسبية للإستجابات بالإضافة إلى التكرارات النسبية للحالات .

\$Infor_Pref_Source Frequencies

		Responses		Percent of Cases
		N	Percent	
\$Infor_Pref_Source ^a	1.00	4	23.5%	44.4%
	2.00	1	5.9%	11.1%
	3.00	3	17.6%	33.3%
	4.00	2	11.8%	22.2%
	5.00	4	23.5%	44.4%
	6.00	3	17.6%	33.3%
Total		17	100.0%	188.9%

a. Group

في العمود السادس المسمى Values (قيم ورموز المتغير) نقوم بتعيين رموز وأرقام لمصادر المعلومات وذلك على النحو التالي:

Value Labels

Value Labels:

Value:

Label:

Add

Change

Remove

1.00 = "المرجع الأجنبي"

2.00 = "الكتاب الجامعي"

3.00 = "المرجع العربي"

4.00 = "الإنترنت"

5.00 = "المحاضرات المكتوبة"

6.00 = "الأولية"

OK

Cancel

Help

بتكرار الخطوات السابقة نفسها نتوصل إلى نتيجة مشابهة تماماً لما توصلنا إليه بطريقة المتغيرات المتعددة ذات الفئتين .

Multiple Response Crosstabs

Row(s): \$Infor_Pref_Source

Column(s): Sex(1 2)

Layer(s):

Define Ranges... Options...

OK Paste Reset Cancel Help

Multi Response Sets:

رغم المسجوب [resp]
 الجامعة [University]
 الكلية [Faculty]
 تاريخ العودة [Retu]
 شهادة الجامعة
 شهادة الدكتوراة
 ب لتدريس الاختصاص
 تنسأة ضم المنعاج

عند نقر الزر OK نحصل على جدول التوزيع التكراري المزدوج الذي يبين عدد التكرارات الخاصة بكل مصدر من مصادر المعلومات والتكرارات النسبية للإستجابات بالإضافة إلى التكرارات النسبية للحالات .

\$Infor_Pref_Source*Sex Crosstabulation

		الجنس		Total
		Male	Female	
\$Infor_Pref_Source	المصدر المفضل للمعلومات الكتاب	193	20	213
	المصدر المفضل للمعلومات الاملية	47	7	54
	المصدر المفضل للمعلومات	143	29	172
	المصدر المفضل للمعلومات	299	51	350
	المصدر المفضل للمعلومات	65	12	77
	المصدر المفضل للمعلومات	140	27	167
Total	Count	537	90	627

Percentages and totals are based on respondents.

a. Dichotomy group tabulated at value 1.

- نختار المتغير Sex من النافذة Mult Response sets ونقله إلى الصفوف Rows

- نختار المتغير Infor_pref_Source ونقله إلى الأعمدة Column(s).

Sex*\$Infor_Pref_Source Crosstabulation

			\$Infor_Pref_Source ^a						Total
			المصدر المفضل للمعلومات الكتاب	المصدر المفضل للمعلومات الاملية	المصدر المفضل للمعلومات المرجع العربي	المصدر المفضل للمعلومات المرجع الاجنبي	المصدر المفضل للمعلومات المحاضرات المكتوبة	المصدر المفضل للمعلومات الانترنت	
الجنس	Male	Count	193	47	143	299	65	140	537
	Female	Count	20	7	29	51	12	27	90
Total		Count	213	54	172	350	77	167	627

Percentages and totals are based on respondents.

a. Dichotomy group tabulated at value 1.

Sex*\$Infor_Pref_Source*University Crosstabulation

				\$Infor_Pref_Source					Total
				المصدر المفضل للمعلومات					
				المصدر المفضل للمعلومات المكتوبة	المصدر المفضل للمعلومات الإلكترونية	المصدر المفضل للمعلومات المرجعية الأجنبية	المصدر المفضل للمعلومات المرجعية العربية	المصدر المفضل للمعلومات الإلكترونية	
الجامعة	الجنس	المرجع	المصدر المفضل للمعلومات	المصدر المفضل للمعلومات المكتوبة	المصدر المفضل للمعلومات الإلكترونية	المصدر المفضل للمعلومات المرجعية الأجنبية	المصدر المفضل للمعلومات المرجعية العربية	المصدر المفضل للمعلومات الإلكترونية	Total
Damascu	Male	Count	54	14	47	86	18	35	152
		% of Tot	27.4%	7.1%	23.9%	43.7%	9.1%	17.8%	77.2%
	Female	Count	10	6	15	24	5	9	45
		% of Tot	5.1%	3.0%	7.6%	12.2%	2.5%	4.6%	22.8%
	Total	Count	64	20	62	110	23	44	197
		% of Tot	32.5%	10.2%	31.5%	55.8%	11.7%	22.3%	100.0%
Aleppo	Male	Count	82	23	57	130	33	80	229
		% of Tot	30.8%	8.6%	21.4%	48.9%	12.4%	30.1%	86.1%
	Female	Count	9	1	11	21	7	14	37
		% of Tot	3.4%	.4%	4.1%	7.9%	2.6%	5.3%	13.9%
	Total	Count	91	24	68	151	40	94	266
		% of Tot	34.2%	9.0%	25.6%	56.8%	15.0%	35.3%	100.0%
Techreen	Male	Count	23	7	18	51	5	15	77
		% of Tot	28.0%	8.5%	22.0%	62.2%	6.1%	18.3%	93.9%
	Female	Count	0	0	2	4	0	4	5
		% of Tot	.0%	.0%	2.4%	4.9%	.0%	4.9%	6.1%
	Total	Count	23	7	20	55	5	19	82
		% of Tot	28.0%	8.5%	24.4%	67.1%	6.1%	23.2%	100.0%
Baath	Male	Count	16	1	15	27	7	6	49
		% of Tot	32.0%	2.0%	30.0%	54.0%	14.0%	12.0%	98.0%
	Female	Count	1	0	0	0	0	0	1
		% of Tot	2.0%	.0%	.0%	.0%	.0%	.0%	2.0%
	Total	Count	17	1	15	27	7	6	50
		% of Tot	34.0%	2.0%	30.0%	54.0%	14.0%	12.0%	100.0%
Forat	Male	Count	18	2	6	5	2	4	30
		% of Tot	56.3%	6.3%	18.8%	15.6%	6.3%	12.5%	93.8%
	Female	Count	0	0	1	2	0	0	2
		% of Tot	.0%	.0%	3.1%	6.3%	.0%	.0%	6.3%
	Total	Count	18	2	7	7	2	4	32
		% of Tot	56.3%	6.3%	21.9%	21.9%	6.3%	12.5%	100.0%

Percentages and totals are based on respondents.

a. Dichotomy group tabulated at value 1.

Sex*\$Infor_Pref_Source*University Crosstabulation

الجامعة	\$Infor_Pref_Source						Total		
			مصدر المفضل للمعلومات	مصدر المفضل للمعلومات	مصدر المفضل للمعلومات	مصدر المفضل للمعلومات			
	الجنس	معلومات الانترنت	المعلومات المكتوبة	المرجع الاجنبي	المرجع العربي	مصدر المفضل للمعلومات	مصدر المفضل للمعلومات		
Damasci	Male	Coun	54	14	47	86	18	35	152
	Femal	Coun	10	6	15	24	5	9	45
	Total	Coun	64	20	62	110	23	44	197
Aleppo	Male	Coun	82	23	57	130	33	80	229
	Femal	Coun	9	1	11	21	7	14	37
	Total	Coun	91	24	68	151	40	94	266
Techreer	Male	Coun	23	7	18	51	5	15	77
	Femal	Coun	0	0	2	4	0	4	5
	Total	Coun	23	7	20	55	5	19	82
Baath	Male	Coun	16	1	15	27	7	6	49
	Femal	Coun	1	0	0	0	0	0	1
	Total	Coun	17	1	15	27	7	6	50
Forat	Male	Coun	18	2	6	5	2	4	30
	Femal	Coun	0	0	1	2	0	0	2
	Total	Coun	18	2	7	7	2	4	32

Percentages and totals are based on respondents.

a.Dichotomy group tabulated at value 1.

University*\$Infor_Pref_Source Crosstabulation

		\$Infor_Pref_Source						Total
		المصدر المفضل للمعلومات الكتاب	المصدر المفضل للمعلومات الاملية	المصدر المفضل للمعلومات المرجع العربي	المصدر المفضل للمعلومات المرجع الاجنبي	المصدر المفضل للمعلومات المحاضرات المكتوبة	المصدر المفضل للمعلومات الانترنت	
الجامعة	Damascus Count	64	20	62	110	24	44	198
	Aleppo Count	100	26	73	161	42	100	284
	Techreen Count	27	7	23	59	5	20	88
	Baath Count	17	3	19	27	9	6	54
	Forat Count	18	2	7	7	2	4	32
Total	Count	226	58	184	364	82	174	656

Percentages and totals are based on respondents.

a. Dichotomy group tabulated at value 1.

يمكن الحصول على تكرارات كل صف أو عمود من جدول التقاطع إما كنسبة مئوية من الاستجابات Responses أو كنسبة مئوية من المستجيبين Respondents أو كنسبة مئوية من الاستجابات و المستجيبين وذلك عن طريق الخيار Options.

Sex*\$Infor_Pref_Source Crosstabulation

		\$Infor_Pref_Source						Total
		المصدر المفضل للمعلومات الكتاب	المصدر المفضل للمعلومات الاملية	المصدر المفضل للمعلومات المرجع العربي	المصدر المفضل للمعلومات المرجع الاجنبي	المصدر المفضل للمعلومات المحاضرات المكتوبة	المصدر المفضل للمعلومات الانترنت	
الجنس	Male Count	193	47	143	299	65	140	537
	% of Total	30.8%	7.5%	22.8%	47.7%	10.4%	22.3%	85.6%
	Female Count	20	7	29	51	12	27	90
	% of Total	3.2%	1.1%	4.6%	8.1%	1.9%	4.3%	14.4%
Total	Count	213	54	172	350	77	167	627
	% of Total	34.0%	8.6%	27.4%	55.8%	12.3%	26.6%	100.0%

Percentages and totals are based on respondents.

a. Dichotomy group tabulated at value 1.

يمكننا الحصول على تفضيلات المستجيبين لمصادر المعلومات المفضلة حسب جنس المستجوب من جهة وحسب انتماء المستجوب للجامعة وذلك كما هو موضح في جدول التوزيع التكراري المزدوج التالي.

ملخص الفصل السادس

تحليل الاستجابات المتعددة : Multiple Response Analysis

في بعض الحالات قد يكون المطلوب من المستجوب الإجابة عن سؤال محدد يشتمل على خيارات متعددة وبالتالي اختيار واحدة من الإجابات المقترحة .
في هذه الحالة فإن إيجاد التكرارات Frequencies الخاصة بكل خيار يتم عن طريق تحليل الاستجابات المتعددة وهناك أسلوبان في التحليل:

أ- أسلوب التكرارات متعددة الاستجابة Multiple Response Frequencies ومنه نحصل على جدول تكراري بسيط.

ب - أسلوب جدول التقاطع متعددة الاستجابة Multiple Response Crosstabs ومنه نحصل على جداول تكرارية ذات بعدين أو ذات ثلاثة أبعاد.

أسلوب التكرارات متعددة الاستجابة Multiple Response Frequencies

لمعرفة التكرارات الخاصة بكل مصدر يمكننا الاختيار بين أسلوبين:

١ - أسلوب المتغيرات المتعددة ذات الفئتين Multiple Dichotomy Method

بموجب هذه الطريقة نخصص متغير لكل خيار من الخيارات المتاحة ويأخذ قيمتين فقط وليكن الصفر والواحد.

لإيجاد التكرارات المقابلة لكل خيار نتبع الخطوات التالية :

- من شريط القوائم في محرر البيانات نختار:

Analysis → Multiple Response → Define variable sets

يظهر صندوق حوار بعنوان Define Multiple Response sets

٢ - أسلوب الفئات المتعددة Multiple Category Method

بموجب هذه الطريقة يتوجب على المحلل معرفة أكبر عدد من الإجابات على السؤال (أي أكبر عدد من الخيارات المفضلة لدى المستجوب) ، لنفترض أن هذا العدد هو ثلاثة خيارات كحد أعلى وفي المقابل نكون العدد نفسه من المتغيرات ولتكن X_1 , X_2 , X_3 .

لإيجاد التكرارات المقابلة لكل خيار نتبع الخطوات التالية :

- من شريط القوائم في محرر البيانات نختار:

Analysis → Multiple Response → Define variable sets

التصنيف للبيانات الخام المشاهدة u_{ij} الذي يضم n قطاعاً مستقلاً يضم كل منها k مشاهدة مرتبطاً ببعضها ببعض داخل كل قطاع، وعادة تكون هذه المشاهدات مستمرة ومن الممكن أن تكون متكررة .

جدول رقم (١)
جدول ثنائي التصنيف للبيانات الخام المشاهدة

تصنيف القطاعات Block ID	المعالجات Treatments			
	1	2	...	k
1	u_{11}	u_{12}	...	u_{1k}
2	u_{21}	u_{22}	...	u_{2k}
...	...	...	...	...
...	...	...	...	...
n	u_{n1}	u_{n2}	...	u_{nk}

إن أول خطوة نبدأها بهدف إجراء اختبار الفروق بين العينات غير المستقلة هو استبدال البيانات الخام u_{ij} بقياسات متوافقة معها، وهذه القياسات هي رتب البيانات في كل عينة على حدة. فمثلاً يمكننا أن نستبدل k مشاهدة $(u_{i1}, u_{i2}, \dots, u_{ik})$ في القطاع i برتب متوافقة معها $(r_{i1}, r_{i2}, \dots, r_{ik})$ أو بأنصاف الرتب في حال وجود مشاهدات متكررة فنحصل بذلك على مخطط جديد ثنائي التصنيف للرتب الخاصة بعدد k من العينات غير المستقلة. إن لم تكن هناك مشاهدات متكررة ضمن الجدول u_{ij} يمكننا تخصيص k عدد صحيح $(1, 2, \dots, k)$ على أنها رتب لهذه المشاهدات عددها k ، وليس ضرورياً أن تكون الرتب تصاعدية أو تنازلية. في حالة وجود مشاهدات متكررة نحسب متوسط الرتب أو أنصاف الرتب للمشاهدات ذات الصلة كما يلي:

لنفترض أن هناك k مشاهدة في القطاع الأول تأخذ e_1 قيمة مختلفة متميزة مع وجود

d_{11} مشاهدة متكررة تساوي أصغر قيمة، d_{12} مشاهدة متكررة تساوي أصغر ثاني قيمة، d_{13} مشاهدة متكررة تساوي أصغر ثالث قيمة وهكذا ... ، وكذلك هناك k مشاهدة في القطاع الثاني تأخذ e_2 قيمة مختلفة متميزة مع وجود d_{21} مشاهدة متكررة تساوي أصغر قيمة ، d_{22} مشاهدة متكررة تساوي أصغر ثاني قيمة ، d_{23} مشاهدة متكررة تساوي أصغر ثالث قيمة وهكذا ... ،

الفصل السابع

الاختبارات اللامعلمية

Non-Parametric Tests

مقدمة:

تصنف الأساليب الإحصائية الاستدلالية إلى أساليب معلمية Parametric Methods لأنها تهتم بمعالم المجتمع الإحصائي مثل المتوسط والنسبة والتباين، وتتطلب تحقق افتراضات معينة حول توزيع المجتمع الإحصائي الذي سحبت منه عينة الدراسة كأن يكون طبيعياً أو أن يكون تباينه متجانساً، وأساليب لامعلمية Nonparametric Methods تستخدم في الحالات التي لا يكون فيها نوع التوزيع الاحتمالي للمجتمع الإحصائي معروفاً أو في حالة عدم تحقق شرط كون التوزيع النظري للمجتمع الإحصائي طبيعياً. وإذا كانت الأساليب المعلمية تناسب البيانات الإحصائية الكمية، فإن الأساليب اللامعلمية تناسب البيانات الإحصائية الكمية والنوعية المعطاة في صورة اسمية أو رتبية حيث تفشل الأساليب المعلمية في معالجتها.

سنتناول في هذا الفصل عرض الأساس النظري لبعض الاختبارات اللامعلمية للاستدلال حول، عينتين مرتبطتين، عينتين مرتبطتين، k عينة مرتبطة، k عينة مستقلة مميزين بين نوعين من البيانات رتبية و اسمية ومن ثم سنحاول إجراء تطبيقات عملية باستخدام برنامج SPSS على كل نوع. آخذين في الحسبان أنه عندما يريد الباحث تحليل البيانات الخاصة بعينة الدراسة، تصادفه مشكلة أساسية وهي تحديد الطريقة اللامعلمية المناسبة، ويتوقف ذلك على عدة اعتبارات منها:

- هدف الدراسة هل هو دراسة علاقات ارتباطية بين المتغيرات أو دراسة الاختلافات؟
- عدد العينات المدروسة هل هي عينة واحدة، أو عينتان أو أكثر من ذلك؟
- هل هذه العينات مستقلة أو عينات مرتبطة هل هذه العينات متماثلة أو مختلفة؟
- نوع البيانات المستخدمة، هل هي بيانات كمية أو نوعية، هل البيانات اسمية أو رتبية؟ وتكمن أهمية البحث في تحديد الطريقة اللامعلمية المناسبة لتحليل الظاهرة.

§1- الاستدلال حول عدد k من العينات غير المستقلة لبيانات رتبية

إن البيانات اللازمة لاختبار الفروق بين عدد k من العينات غير المستقلة تمثل عادة بجدول ثنائي التصنيف Two- way layout (في اتجاهين). نشير إليه عادةً بمخطط ثنائي

يظهر صندوق حوار بعنوان Define Multiple Response sets

الذي نرتبه باتباع الخطوات التالية:

- اختيار المتغيرات المحددة من النافذة Set Definition ونقلها الى النافذة Variable in set

- اختيار Categories من النافذة Variables are coded as وتعيين فئات الخيارات (

..... through Range) .

- إعطاء إسم لمجموعة الاستجابة المتعددة في المربع Name .

- ننقر الزر إضافة Add لإضافة المجموعة إلى خانة Mult Response sets .

- عند نقر الزر إغلاق Close يتم خزن مجموعة الاستجابة المتعددة بإسم محدد للاستفادة منها

فيما بعد بعملية التحليل.

أسلوب جداول التقاطع متعددة الإستجابة Multiple Response Crosstabs

تستخدم هذه الطريقة لتكوين جداول تقاطع للمتغيرات الأولية Elementary Variables أو

للمجاميع المعرفة Multiple Response sets أو لكليهما معاً .

وأخيراً هناك k مشاهدة في القطاع الأخير n تأخذ e_n قيمة مختلفة متميزة مع وجود d_{n1} مشاهدة متكررة تساوي أصغر قيمة ، d_{n2} مشاهدة متكررة تساوي أصغر ثاني قيمة ، d_{n3} مشاهدة متكررة تساوي أصغر ثالث قيمة وهكذا ... ، يمكننا الآن أن نعرف بدقة أنصاف الرتب . فمن أجل $i = 1, 2, \dots, n$ فإن أنصاف الرتب المتميزة e_i the distinct mid-ranks في القطاع i يمكن أن تعطى بموجب العلاقة:

$$\begin{aligned} r_{i1}^* &= (d_{i1} + 1) / 2 , \\ r_{i2}^* &= d_{i1} + (d_{i2} + 1) / 2 , \\ &\dots \\ r_{i,e_i}^* &= d_{i1} + d_{i2} + \dots + (d_{i,e_i} + 1) / 2 , \end{aligned}$$

يمكننا الآن أن نستبدل البيانات الأصلية $(u_{i1}, u_{i2}, \dots, u_{ik})$ في القطاع i برتب متوافقة معها $(r_{i1}, r_{i2}, \dots, r_{ik})$ أو بأنصاف الرتب في حال وجود مشاهدات متكررة حيث r_{ij} تمثل أنصاف الرتب المتميزة $(r_{i1}^* < r_{i2}^* < \dots < r_{i,e_i}^*)$ ، وبذلك نحصل على مخطط جديد ثنائي التصنيف للرتب الخاصة بـ k عينة مرتبطة (غير مستقلة).

جدول رقم (٢)

جدول ثنائي التصنيف للرتب الخاصة بـ k عينة مرتبطة

تصنيف القطاعات	المعالجات Treatments			
Block ID	1	2	...	k
1	r_{11}	r_{12}	...	r_{1k}
2	r_{21}	r_{22}	...	r_{2k}
...	...	...	...	...
...	...	...	...	...
n	r_{n1}	r_{n2}	...	r_{nk}

لنفترض أن لدينا 5 عينات غير مستقلة $k = 5$ ، بياناتها الخام تتوزع على قطاعين فقط كما في الجدول التالي:

جدول ثنائي التصنيف للبيانات الخام لـ 5 عينات مع قطاعين

تصنيف القطاعات	العينات Samples				
Block ID	1	2	3	4	5
1	1.3	1.1	1.1	1.6	1.1
2	1.9	1.7	1.9	1.9	1.7

جدول رقم (٣)
ثاني التصنيف للبيانات الخام المشاهدة

تصنيف القطاعات Block ID	المعالجات Treatments			
	1	2	...	k
1	u_{11}	u_{12}	...	u_{1k}
2	u_{21}	u_{22}	...	u_{2k}
...	...	...	...	...
...	...	...	...	...
n	u_{n_1}	u_{n_2}	...	u_{n_k}

- **الشروط:** بالإضافة إلى تقسيم الوحدات إلى أعمدة وصفوف نفترض أن المتغير مستمراً وأن المشاهدات في كل قطاع قابلة للتحويل إلى رتب تصاعدية وأن لا يكون هناك تفاعل بين الصفوف والأعمدة.

- **الفروض:** سنختبر فرضية عدم القائلة إن تأثير المعالجات متطابق، مقابل الفرض البديل : تأثير المعالجات غير متطابق.

- **إحصائية الاختبار:** للحصول على إحصائية الاختبار نرتب بيانات كل قطاع ونعطي رتباً مناظرة للقيم والترتيب يكون تصاعدياً، نجمع رتب كل قطاع فنحصل على R_i التي تمثل مجموع رتب القطاع i ، وإحصائية الاختبار تحسب بالعلاقة :

$$\chi_c^2 = \left(\frac{12 \sum R_i^2}{nk(k+1)} \right) - 3n(k+1) \quad (1)$$

- نستخرج من جداول خاصة باختبار فريدمان وبمساعدة α ، n ، k قيمة جدولية تقارن بالقيمة المحسوبة، نرفض فرضية عدم إذا زادت القيمة المحسوبة عن القيمة الجدولية ويمكن استخدام جدول توزيع كاي تربيع عند $k-1$ درجة حرية.

- يمكن إجراء تصحيح عند وجود تداخلات مما يؤدي إلى تحسين النتائج ودقة الاختبار وذلك بقسمة إحصائية الاختبار على المقدار:

$$1 - \left(\sum_{i=1}^n T_i \right) / nk(k^2 - 1) \quad , \quad T_i = \sum t_i^3 - \sum t_i \quad (2)$$

ومتوسط مجموع الرتب لكل عينة أو معالجة، ومتوسط مجموع الرتب لكل العينات (أو المعالجات) $r_{..}$ ، $r_{.j}$ و r_{ij} وكل منها يختلف عن الآخر.

إذا كانت T هي إحصائية الاختبار، فإن فرضية العدم والفرضية البديلة ستأخذان الشكل التالي:

H_0 : لا توجد فروق ذات دلالة إحصائية بين k معالجة أو k عينة غير مستقلة¹.

H_1 : هناك على الأقل معالجتان (أو عينتان غير مستقلتين) مختلفتان.

١-١ اختبار فريدمان لتحليل التباين للرتب في اتجاهين:

The Friedman Two – Way Analysis of Variance by Ranks Test

يعتبر اختبار فريدمان أحد أهم الاختبارات اللامعلمية الشائعة الاستخدام لمقارنة k عينة غير مستقلة، فمثلاً إذا كان لدينا عدداً من المواقع التجريبية التي تتعرض لها المشاهدة أو وحدة التجربة بشكل تعطي لها درجة بكل موقف، أو أن يكون هناك بدائل (تخصصات أو مهن) تعرض على وحدة التجربة وعليه أن يربتها حسب تفضيلاته أو يربتها تصاعدياً من الأكثر أهمية إلى الأقل. في مثل هذه الحالات يستخدم اختبار فريدمان لتحليل التباين ثنائي التصنيف للعينات المترابطة.

نفرض أن لدينا تصميم قطاعات كاملة العشوائية يحتوي على n قطاع (صف) وعلى k عمود بحيث أن كل قطاع يكون مستقلاً عن الآخر بالتبادل Mutually Independent معنى ذلك أن تكون u_{ij} هي نتيجة المشاهدة رقم j في الصف رقم i كما هو موضح بالجدول ، ويكون هدفنا في النهاية معرفة هل هناك اختلاف بين المعالجات أو لا، وإن وجد يمكننا أن نستخدم أسلوب المقارنات المتعددة لمعرفة أي المعالجات هي السبب في وجود هذا الفرق.

¹ عندما يكون لدينا أكثر من عینتين غير مستقلتين يصعب على الباحث من الناحية العملية أن يحقق شرط الاستقلال بين وداخل المعالجات، ويكون أسهل له أن تكون لديه وحدات متشابهة تماماً فيما بينها ويتم تقسيمها إلى مجموعات ويستخدم في كل مجموعة معالجة مختلفة وبعد انتهاء التجربة نبحث عن تأثير المعالجات هل هو مختلف أو متساوي. في هذه الحالة شرط الاستقلال يكون غير متحقق والنتائج في حالتنا هذه تتشابه مع تصميم القطاعات الكاملة العشوائية حيث تمثل فيها الأعمدة المعالجات وتمثل الصفوف مجموعات من العينات الغير مستقلة وغالباً نضع في الصفوف بعض المعالم غير المهمة مثل عامل الزمن (يوم، شهر، سنة). ستهتم فروض الاختبار بالفرق بين الأعمدة.

نلاحظ من الجدول أنه لدينا في القطاع الأول: $e_1 = 3$ مع $d_{11} = 3, d_{12} = 1, d_{13} = 1$ وبتطبيق العلاقة السابقة نحصل على أنصاف الرتب على النحو التالي:

$$r_{11}^* = (3+1)/2 = 2, r_{12}^* = 4, r_{13}^* = 5$$

وفي القطاع الثاني لدينا: $e_2 = 2$ مع $d_{21} = 2, d_{22} = 3$ وبتطبيق العلاقة نفسها نحصل على:

$$r_{21}^* = (2+1)/2 = 1.5, r_{22}^* = 2 + (3+1)/2 = 4$$

يمكننا أن نستخدم أنصاف الرتب هذه لتحل محل البيانات الأصلية فنحصل على جدول ثنائي التصنيف للرتب الخاصة بالعينات الخمس المرتبطة كما يلي:

جدول ثنائي التصنيف للرتب لـ 5 عينات مع قطاعين

تصنيف القطاعات Block ID	العينات Samples				
	1	2	3	4	5
1	4	2	2	5	2
2	4	1.5	4	4	1.5

إن كافة الاختبارات تعتمد على إحصائيات الاختبار التي سنستخدمها في حساباتها على الجدول ثنائي التصنيف للرتب أو لأنصاف الرتب رقم (2) في حالة وجود مشاهدات متكررة، ولهذا السبب وقبل توصيف هذه الإحصائيات سنعرف:

- مجموع الرتب rank-sum لكل عينة أو معالجة j حيث $j = 1, 2, \dots, k$ بالعلاقة:

$$r_j = \sum_{i=1}^n r_{ij}$$

- متوسط مجموع الرتب average rank-sum لكل عينة أو معالجة j بالعلاقة:

$$r_{.j} = \frac{r_j}{N}$$

- متوسط مجموع الرتب average rank-sum لكل العينات أو المعالجات بالعلاقة:

$$r_{..} = \frac{\sum_{j=1}^k r_j}{k} = \frac{k+1}{2}$$

إن إحصائيات الاختبار لكل من فريدمان، كندال W، كوكران Q، كواد و باج Friedman, Kendal W, Cochran Q, Quade and Page هي عبارة عن توابع لمجموع الرتب

لنفترض أن النموذج الملائم من أجل تصميم قطاع تام العشوائية، الذي يولد البيانات الخام المشاهدة في الجدول رقم (٣) ثنائي التصنيف يأخذ الصيغة التالية:

$$U_{ij} = \mu + \beta_i + \tau_j + \varepsilon_{ij} \quad (3)$$

حيث: U_{ij} تمثل قيمة نموذجية لكافة المشاهدات المدروسة وذلك من أجل كل $i = 1, 2, \dots, n_j$ وكذلك $j = 1, 2, \dots, k$.

μ تمثل في هذا النموذج المتوسط العام.

τ_j تمثل تأثير المعالجة j تعكس حقيقة تلقي وحدة التجربة المعالجة j .

β_i تمثل تأثير القطاع i وتعكس حقيقة وقوع وحدة التجربة المشاهدة في القطاع i .

ε_{ij} تمثل حدود الخطأ العشوائي غير المشاهدة وتمثل كل مصادر التباين باستثناء المعالجات والقطاعات وهي مستقلة وتتبع توزيعاً احتمالياً غير معلوم بمتوسط صفر وكل وسطائه غير معلومة.

إن تأثيرات القطاع والمعالجة هي تجميعية هذا يعني أنه لا وجود لتأثيرات متبادلة

Interaction بين القطاع والمعالجة ورياضياً نكتب:

$$\sum_{i=1}^{n_j} \beta_i = \sum_{j=1}^k \tau_j = 0$$

إن فرضية عدم القائلة بعدم وجود تأثير للمعالجة تصاغ على النحو التالي:

$$H_0: \tau_1 = \tau_2 = \dots = \tau_k$$

$$H_1: \tau_{j1} \neq \tau_{j2}$$

مقابل الفرضية البديلة:

وذلك من أجل زوج واحد على الأقل (j_1, j_2) .

هذه الفرضية لا تحدد أي ترتيب للمعالجات بدلالة مستوى الاستجابات المتزايد، فالفرضية البديلة هنا تحدد فقط بأن المعالجات مختلفة (أي لا تحدد أن هذه المعالجة أو تلك هي أكثر فعالية أو أكثر استجابة more responsive مقارنة بغيرها).

إن اختبار فريدمان يستخدم إحصائية الاختبار T_F المعرفة على الجدول ثنائي التصنيف للترتيب، ولحسابها نتبع الخطوات التالية:

أولاً: نبدأ بإعطاء رتب لجميع المشاهدات في كل عينة على حدة، فنعطي الرتبة ١ لأصغر مشاهدة ثم الرتبة ٢ للمشاهدة التي تليها في الكبير وهكذا....

E	3	4	5	6	2	1
F	2	5	6	3	1	4
G	6	5	3	4	2	1
H	5	1	4	3	6	2
مجموع الرتب	37	26	28	24	25	28
متوسط مجموع الرتب	4.625	3.25	3.5	3	3.125	3.5

- إن البيانات الخام هي رتب أو تفضيلات الطلاب للتخصصات.

- من أجل كل معالجة أو تخصص نحسب مجموع الرتب لجميع أفراد العينة، نحسب متوسط مجموع الرتب لكل تخصص، كما نحسب متوسط مجموع الرتب لكل المعالجات بالعلاقة:

$$r_{..} = \frac{\sum_{j=1}^k r_j}{k} = \frac{k+1}{2} = 7/2 = 3.5$$

- نحسب القيمة R بموجب العلاقة التالية:

$$R = \left(\sum_{i=1}^{n_1} r_{i1} - nr_{..} \right)^2 + \left(\sum_{i=1}^{n_2} r_{i2} - nr_{..} \right)^2 + \dots + \left(\sum_{i=1}^{n_k} r_{ik} - nr_{..} \right)^2 = 110$$

- نحسب إحصائية الاختبار T_F بموجب العلاقة التالية:

$$T_F = \frac{12R}{nk(k+1)} = \frac{12(110)}{8(6)(6+1)} = 3.93$$

بمقارنة قيمة إحصائية الاختبار المحسوبة T_F مباشرة مع قيمة كاي تربيع النظرية عند مستوى معنوية ٠,٠٥ وكذلك ٠,٠١ ودرجات حرية تساوي $k-1=5$. نقبل فرضية عدم القائلة بعدم وجود فروق ذات دلالة إحصائية بين المعالجات المختلفة لأن قيمة إحصائية الاختبار المحسوبة T_F أصغر من قيمة كاي تربيع النظرية في الحالتين التي تساوي ١١,٠٧ عند مستوى معنوية ٥ %، وتساوي ١٥,٠٩ عند مستوى معنوية ١ %. أي أنه لا توجد فروق ذات دلالة إحصائية بين ترتيب التخصصات أو التفضيلات كما تصورهما أفراد العينة.

اختبار فريدمان مخرجات برنامج MINITAB (Friedman Test (MINITAB MINITAB Outputs)

Friedman test for Subjects by Treatmen blocked by Blocks

S = 3.93 DF = 5 P = 0.560

ملاحظة هامة:

عندما تكون بعض قيم المشاهدات متكررة نضطر لاستخدام أنصاف الرتب، وبهذه الحالة فإن إحصائية الاختبار T_F^2 تحسب بموجب العلاقة التالية:

$$T_F = \frac{12 \sum_{j=1}^k (r_j - nr_{..})^2}{nk(k+1) - (k-1)^{-1} \sum_{i=1}^n [\sum_{j=1}^{e_i} d_{ij}^2 - k]}, \quad (6)$$

إذ إن: n تمثل عدد القطاعات (وتمثل عدد أفراد العينة) ، k تمثل عدد المعالجات أو عدد البدائل ، في القطاع الأول i هناك k مشاهدة تأخذ e_i قيمة مختلفة متميزة مع وجود d_{i1} مشاهدة متكررة تساوي أصغر قيمة ، d_{i2} مشاهدة متكررة تساوي أصغر ثان قيمة ، d_{i3} مشاهدة متكررة تساوي أصغر ثالث قيمة وهكذا

مثال (١):

لوقوف على رأي مجموعة من الطلاب الحاصلين على الشهادة الثانوية (وقد تم تمييز كل طالب بالحرف الأول من أسمه) في التخصصات ذات الأفضلية (علم اجتماع، فلسفة، تاريخ، جغرافيا، أدب عربي، لغة إنكليزية) إذا أجبروا على القبول في كلية الآداب، وقد هدف الباحث للكشف عما إذا كانت هناك رغبة متميزة لتخصص دون الآخر، وللتحقق من صحة فرض العدم القائل بعدم وجود فروق ذات دلالة إحصائية بين رغبات الطلاب في التخصصات بكلية الآداب " قام الباحث بجمع البيانات التالية (إذ تعطى الرغبة الأولى 6 درجات والأخيرة 1 درجة وذلك حسب درجة التفضيل للفرع لديه .

البيان	علم اجتماع	فلسفة	تاريخ	جغرافيا	أدب عربي	لغة إنكليزية
A	6	3	2	1	4	5
B	4	3	1	2	5	6
C	5	1	2	3	4	6
D	6	4	5	2	1	3

² معامل الاتفاق أو اختبار كندال Kendall's W or Coefficient of Concordance تم تطويره على أنه مقياس للاقتتران مع n قطاع يمثلون n حكماً مستقلاً، كل واحد منهم يعطي رتب لنفس المجموعة k من المرشحين. يقيس المعامل مدى اتفاق الحكام في الترتيب الذي يخصصونها للمرشحين. ونظراً للعلاقة الوثيقة ما بين اختبار فريدمان ومعامل الاتفاق لكندال فإن إحصائية الاختبار تحسب كما يلي:

$$W = \frac{T_F}{N(k-1)}$$

إذ إن : $W = 1$ في حالة الاتفاق التام بين الحكام في ترتيب المرشحين ، $W = 0$ في حالة عدم الاتفاق التام بين الحكام في ترتيبهم للمرشحين.

ثانياً: من أجل كل معالجة أو طريقة $j = 1, 2, \dots, k$ ، نحسب مجموع الرتب لكل

$$r_j = \sum_{i=1}^n r_{ij} \quad \text{معالجة بموجب العلاقة التالية:}$$

ثالثاً: نحسب متوسط مجموع الرتب لكل معالجة بالعلاقة: $r_{.j} = \frac{r_j}{N}$ وكذلك متوسط مجموع

$$r_{..} = \frac{\sum_{j=1}^k r_{.j}}{k} \approx \frac{k+1}{2} \quad \text{الرتب لكل المعالجات بالعلاقة:}$$

رابعاً: نحسب القيمة R بموجب العلاقة التالية:

$$R = \left(\sum_{i=1}^{n_1} r_{i1} - nr_{.1} \right)^2 + \left(\sum_{i=1}^{n_2} r_{i2} - nr_{.2} \right)^2 + \dots + \left(\sum_{i=1}^{n_k} r_{ik} - nr_{.k} \right)^2, \quad (4)$$

خامساً: في حال عدم وجود مشاهدات متكررة نحسب إحصائية الاختبار T_F بالعلاقة:

$$T_F = \frac{12 R}{nk (k+1)}, \quad (5)$$

حيث: n تمثل عدد القطاعات (وتمثل عدد أفراد العينة).

k تمثل عدد المعالجات أو عدد البدائل.

٢-١ الدلالة الإحصائية لاختبار فريدمان في حالة $n \geq 10$ أو $k \geq 5$:

نقارن قيمة إحصائية الاختبار المحسوبة T_F مباشرة مع قيمة كاي تربيع النظرية أو

الدرجة عند مستوى معنوية محدد ودرجات حرية تساوي $k-1$. نرفض فرضية العدم القائلة

بعدم وجود فروق ذات دلالة إحصائية بين المعالجات المختلفة إذا كانت

قيمة إحصائية الاختبار المحسوبة T_F أكبر من قيمة كاي تربيع النظرية.

٣-١ الدلالة الإحصائية لاختبار فريدمان في حالة $2 \leq n < 10$ أو $2 \leq k \leq 4$:

نقارن قيمة إحصائية الاختبار المحسوبة T_F بهذه الحالة مع القيمة النظرية أو الدرجة

المستخرجة من جدول فريدمان والتي تتحدد طبقاً لعدد أفراد العينة n وعدد المعالجات k .

نرفض فرضية العدم القائلة بعدم وجود فروق ذات دلالة إحصائية بين المعالجات المختلفة إذا

كانت قيمة إحصائية الاختبار المحسوبة T_F أكبر من القيمة النظرية أو الدرجة لفريدمان.

Treatmen	N	Est Median	Sum of Ranks
1	8	3.437	28.0
2	8	3.354	25.0
3	8	3.021	24.0
4	8	3.771	28.0
5	8	3.521	26.0
6	8	5.021	37.0

Grand median = 3.688

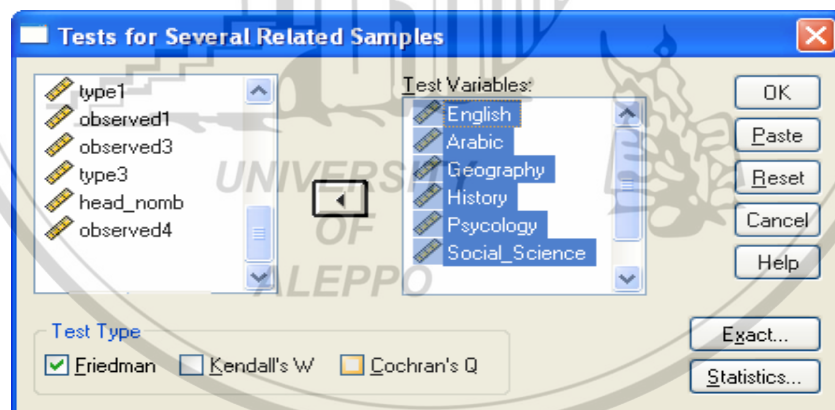
الحل :

من شريط القوائم نختار :

Analysis-→ Non Parametric-→ K Related Samples

يظهر صندوق حوار بعنوان :

Tests for Several Related Samples



- ننقل المتغيرات الواجب مقارنة الفروق بينها وهي :

English ,Arabic , Geography , History , Psychology , Social_Science إلى

النافذة المعنونة Test Variable List

- نضغط على المربع Statistics نحصل على صندوق حوار بعنوان:

. Several Related Samples: Stat...

$$1 - (\sum_{i=1}^n T_i) / n(k^3 - k) = 1 - (6) / 3(5^3 - 5) = 0.983333$$

$$\sum_{i=1}^n T_i = \sum_{i=1}^n (t_i^3 - t_i) = 2^3 - 2 = 6$$

Friedman Test

Friedman test for C10 by C12 blocked by C13

S = 10.07 DF = 4 P = 0.039

S = 10.24 DF = 4 P = 0.037 (adjusted for ties)

C12	N	Est	Sum of
		Median	Ranks
1	3	30.00	10.5
2	3	20.80	5.0
3	3	10.20	4.5
4	3	54.00	15.0
5	3	25.00	10.0

Grand median = 28.00

الحل :

من شريط القوائم نختار :

Analysis → Non Parametric → K Related Samples

يظهر صندوق حوار بعنوان : Tests for Several Related Samples

مثال (٢):

في دراسة لتأثيرات التنويم المغناطيسي على قوة الجلد effects of hypnosis on skin potential، فإن مشاعر وأحاسيس: الخوف القلق والفرع، السعادة البهجة والغبطة، الاكتئاب والحزن والانحطاط، الهدوء الطمأنينة والسكينة، والتهيج أو الانفعال والاضطراب كانت مطلوبة (بترتيب عشوائي) من أجل كل واحدة من ثلاثة مواضيع، والجدول التالي يعطي نتائج القياسات بالميلي فولت milivolts على تحمل الجلد (تم معايرتها عند مستوى أولي).

جدول ثنائي التصنيف للبيانات الخام لـ ٥ عينات مع ثلاث قطاعات

تصنيف القطاعات Subjects Id	المعالجات				
	الخوف	السعادة	الاكتئاب	الهدوء	التهيج
1	33	58	11	24	34
2	23	53	10	20	40
3	23	54	22	21	22

الرتب الخاصة بـ 5 عينات مع ثلاث قطاعات:

تصنيف القطاعات Subjects Id	المعالجات				
	الخوف	السعادة	الاكتئاب	الهدوء	التهيج
1	3	5	1	2	4
2	3	5	1	2	4
3	4	5	2.5	1	2.5
r_j مجموع الرتب	10	15	4.5	5	10.5
متوسط الرتب	3.33	5.00	1.50	1.67	3.50
r_j^2	100.00	225.00	20.25	25.00	110.25

$$R = \sum_{j=1}^k r_j^2 = 110.25 + 25 + 20.25 + 225 + 100 = 480.5$$

$$T_F = \frac{12R}{nk(k+1)} - 3n(k+1) = \frac{12(480.5)}{3(5)(5+1)} - 3(3)(6) = 10.0667 \approx 10.07$$

- يمكن إجراء تصحيح عند وجود تداخلات ووجود مشاهدات متكررة مما يؤدي إلى تحسين النتائج ودقة الاختبار وذلك بقسمة إحصائية الاختبار على المقدار:

$$1 - \left(\sum_{i=1}^n T_i \right) / n(k^3 - k)$$

$$T_F = \left[\frac{12R}{nk(k+1)} - 3n(k+1) \right] / 1 - \left(\sum_{i=1}^n T_i \right) / n(k^3 - k) = 10.07 / 0.9833 = 10.24$$



- نُفعل الخيارين Descriptive , Quartiles .

- نضغط Continue فنعود إلى صندوق الحوار السابق .

- نُفعل الخيار Friedman .

- نضغط OK فنحصل على النتائج التالية :

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles		
						25th	50th (Median)	75th
English	8	3.5000	2.07020	1.00	6.00	1.2500	3.5000	5.7500
Arabic	8	3.1250	1.88509	1.00	6.00	1.2500	3.0000	4.7500
Geography	8	3.0000	1.51186	1.00	6.00	2.0000	3.0000	3.7500
History	8	3.5000	1.77281	1.00	6.00	2.0000	3.5000	5.0000
Psychology	8	3.2500	1.58114	1.00	5.00	1.5000	3.5000	4.7500
Social_Science	8	4.6250	1.50594	2.00	6.00	3.2500	5.0000	6.0000

Friedman Test (SPSS Outputs)

اختبار فريدمان مخرجات برنامج SPSS

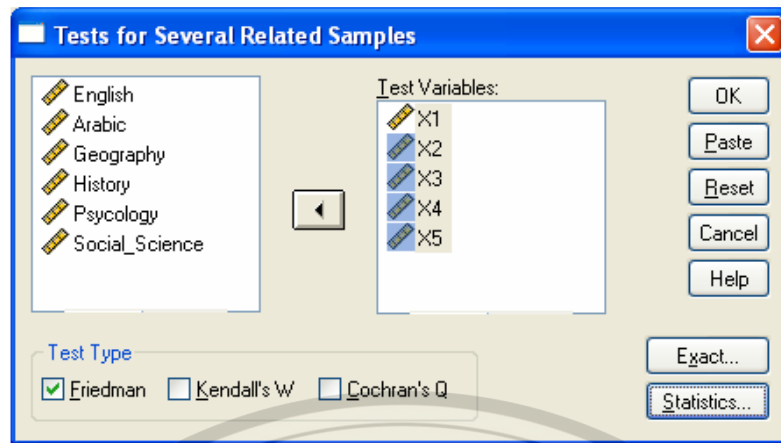
Ranks

	Mean Rank
English	3.50
Arabic	3.13
Geography	3.00
History	3.50
Psychology	3.25
Social_Science	4.63

Test Statistics ^a

N	8
Chi-Square	3.929
df	5
Asymp. Sig.	.560

a. Friedman Test



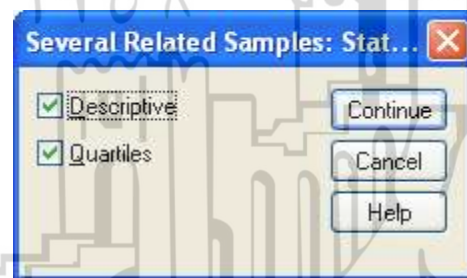
- ننقل المتغيرات الواجب مقارنة الفروق بينها وهي :

التهيج X_1 ، الهدوء X_2 ، الإكتئاب X_3 ، السعادة X_4 ، الخوف X_5 ، إلى النافذة

المعنونة Test Variable List

- نضغط على المربع Statistics نحصل على صندوق حوار بعنوان:

. Several Related Samples: Stat...



- نُفعل الخيارين Descriptive , Quartiles .

- نضغط Continue فنعود إلى صندوق الحوار السابق .

- نُفعل الخيار Friedman ، ثم نضغط OK فنحصل على النتائج التالية :

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles		
						25th	50th (Median)	75th
X1	3	32.0000	9.16515	22.00	40.00	22.0000	34.0000	40.0000
X2	3	21.6667	2.08167	20.00	24.00	20.0000	21.0000	24.0000
X3	3	14.3333	6.65833	10.00	22.00	10.0000	11.0000	22.0000
X4	3	55.0000	2.64575	53.00	58.00	53.0000	54.0000	58.0000
X5	3	26.3333	5.77350	23.00	33.00	23.0000	23.0000	33.0000

Friedman Test (SPSS Outputs)

اختبار فريدمان مخرجات برنامج SPSS

$$Q = \frac{k(k-1) \sum_{j=1}^k (B_j - \bar{B})^2}{k \sum_{i=1}^n L_i - \sum_{i=1}^n L_i^2} \quad (8)$$

التي تتبع توزيع كاي تربيع مع $(k-1)$ درجة حرية.

حيث B_j هي العدد الإجمالي للنجاحات في المعالجة j ، L_i العدد الإجمالي للنجاحات في القطاع i ، و $\bar{B} = (B_1 + B_2 + \dots + B_k) / k$ هي متوسط عدد حالات النجاح في كافة المعالجات.

٢-٢ الدلالة الإحصائية لاختبار كوكران Q :

نقارن قيمة إحصائية الاختبار المحسوبة Q مباشرة مع قيمة كاي تربيع النظرية أو الدرجة عند مستوى معنوية محدد ودرجات حرية تساوي $k-1$ عدد المعالجات أو الاختبارات المتتالية - ١ . نرفض فرضية العدم القائلة بعدم وجود فروق ذات دلالة إحصائية بين المعالجات المختلفة إذا كانت قيمة إحصائية الاختبار المحسوبة Q أكبر من قيمة كاي تربيع النظرية.

مثال (٣):

استخدمت أربعة برامج للنمو الاجتماعي للأطفال، حيث تم تطبيق كل برنامج على عينة مختلفة، تم مماثلة العينات الأربع قبل تطبيق البرامج وبعده، تم إجراء اختبار للنمو الاجتماعي على أطفال العينات الأربع فكانت النتائج كما هي موضحة في الجدول التالي، حيث يشير الرقم ١ إلى اجتياز الطفل والرقم ٠ إلى عدم اجتياز الطفل. والمطلوب: التأكد من صحة فرضية العدم القائلة بعدم وجود فروق ذات دلالة إحصائية في النمو الاجتماعي للأطفال باختلاف نوع البرنامج المطبق.

المجموعة رقم الطفل	X	Y	Z	W	عدد النجاحات L_i	L_i^2
1	1	0	1	0	2	4
2	1	0	1	1	3	9
3	1	1	1	1	4	16
4	0	1	1	1	3	9
5	0	0	1	0	1	1
6	0	0	0	1	1	1
7	1	1	1	1	4	16
8	1	1	1	1	4	16
9	1	1	1	0	3	9
10	1	0	1	0	2	4
11	0	1	1	1	3	9

١-٢ اختبار كوكران Q Cochran's Q Test

إن أحد أهم النماذج الرياضية الشائعة الاستخدام والتي تولد بيانات ثنائية القيمة (0 ، 1)
بجدول ثنائي التصنيف هو نموذج الانحدار اللوجستي (المنطقي) the logistic regression model .

$$\log \frac{\pi_{ij}}{1-\pi_{ij}} = \mu + \beta_i + \tau_j \quad (7)$$

حيث من أجل كل $i = 1, 2, \dots, n$ $j = 1, 2, \dots, k$ فإن : $\pi_{ij} = \Pr(U_{ij} = 1)$

μ تمثل في هذا النموذج المتوسط العام background log-odds of response .

τ_j تمثل تأثير المعالجة j .

β_i تمثل تأثير القطاع i وتعكس حقيقة وقوع وحدة التجربة المشاهدة في القطاع i .

يفترض النموذج عدم وجود تأثيرات متبادلة بين القطاع والمعالجة أي:

$$\sum_{i=1}^n \beta_i = \sum_{j=1}^k \tau_j = 0$$

إن تطبيق اختبار فريدمان على مثل هذه البيانات الاسمية يعرف باختبار كوكران Q .

إن فرضية العدم القائلة بعدم وجود تأثير للمعالجة تصاغ على النحو التالي:

$$H_0 : \tau_1 = \tau_2 = \dots = \tau_k$$

مقابل الفرضية البديلة : $H_1 : \tau_{j1} \neq \tau_{j2}$ وذلك من أجل زوج واحد على الأقل (j_1, j_2) .

إن الفرضية البديلة هنا تحدد فقط بأن المعالجات مختلفة (أي لا تحدد أن هذه المعالجة أو تلك هي أكثر فعالية أو أكثر استجابة more responsive مقارنة بغيرها) .

خطوات الاختبار: نعطي الرقم ١ لحالة الموافقة أو النجاح والرقم ٠ لحالة الرفض أو الفشل،
نحسب عدد حالات النجاح في كل معالجة أو عينة ومن ثم نحسب مجموع عدد حالات النجاح
لجميع أفراد العينة، نحسب عدد حالات النجاح في كل قطاع أو لكل فرد في المجموعة موضع
الاختبار.

إحصائية الاختبار: إن اختبار كوكران Q يستخدم إحصائية الاختبار T_F لفريدمان، ونظراً

لكون المشاهدات اسمية تأخذ القيمتين ٠ و ١ فإن إحصائية الاختبار T_F تختصر إلى:

Ranks

	Mean Rank
X1	3.50
X2	1.67
X3	1.50
X4	5.00
X5	3.33

Test Statistics^a

N	3
Chi-Square	10.237
df	4
Asymp. Sig.	.037

a. Friedman Test

§٢- الاستدلال حول عدد k من العينات غير المستقلة (لبيانات إسمية)

K- Sample Inference: Related Samples

إن البيانات اللازمة لاختبار الفروق بين عدد k من العينات غير المستقلة تمثل عادة بجدول ثنائي التصنيف Two- way layout (في اتجاهين) الجدول رقم (١)، نشير إلى هذا الجدول عادة بمخطط ثنائي التصنيف للبيانات الخام المشاهدة u_{ij} الذي يضم n قطاعاً مستقلاً يضم كل منها k مشاهدة مرتبطة بعضها ببعض داخل كل قطاع، وعادة تكون مستمرة ومن الممكن أن تكون متكررة .

نفترض أن القيم u_{ij} في الجدول رقم (1) ثنائي التصنيف Two- way layout جميعها ثنائية القيمة Binary من الشكل (0 ، 1) حيث يشير الرقم ١ إلى النجاح والرقم ٠ يشير إلى الفشل، إن هذا الجدول يتضمن عادة عدداً من المجموعات المتماثلة تماماً أو أن هناك تناظراً في اختيار أفرادها، وقد يكون لدينا مجموعة واحدة يطبق عليها اختبار في الاتجاهات أو الميول أو النزعات عدداً من المرات، وبهذا الشكل يكون لدينا عدد من العينات غير المستقلة، وباعتبار أن الاستجابة على الاختبارات ستكون على النحو: (موافق ، غير موافق) ، (نعم ، لا) ، (0 ، 1) ، (ناجح ، راسب) ، (أجتاز ، تعثر) ، والهدف هو معرفة فيما إذا كانت الفروق بين نتائج الاختبارات المختلفة هي معنوية أو لا .

12	0	1	0	1	2	4
عدد النجاحات B_j	7	7	10	8	32	98

- نحسب متوسط عدد حالات النجاح في كافة المعالجات:

$$\bar{B} = (B_1 + B_2 + \dots + B_k) / k = (7 + 7 + 10 + 8) / 4 = 8$$

- نحسب إحصائية الاختبار:

$$Q = \frac{k(k-1) \sum_{j=1}^k (B_j - \bar{B})^2}{k \sum_{i=1}^n L_i - \sum_{i=1}^n L_i^2} = \frac{4(4-1)(6)}{4(32) - 98} = \frac{72}{30} = 2.4$$

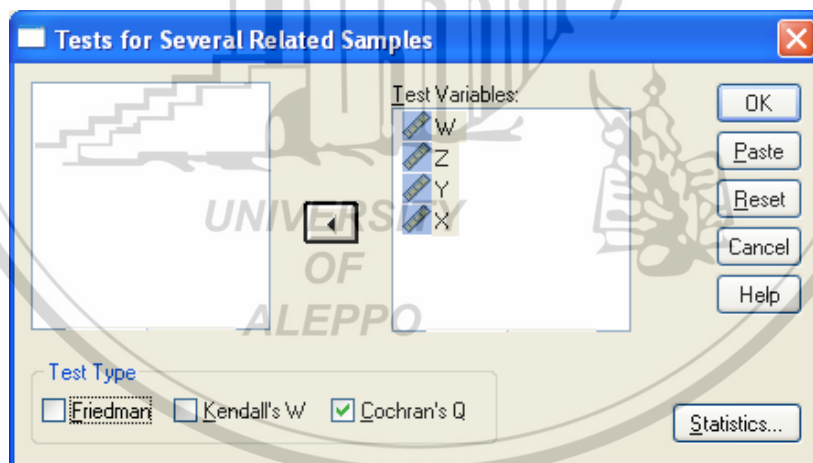
وعند مستوى معنوية 0.05 و 4 درجات حرية فإن قيمة كاي تربيع الجدولية تساوي 7.82 وهي أكبر من قيمة إحصائية الاختبار المحسوبة وبالتالي نقبل فرضية العدم القائلة بعدم وجود فروق ذات دلالة إحصائية في النمو الاجتماعي للأطفال باختلاف نوع البرنامج المطبق.

الحل :

من شريط القوائم نختار :

Analysis → Non Parametric → K Related Samples

يظهر صندوق حوار بعنوان : Tests for Several Related Samples :



ننقل المتغيرات الواجب مقارنة الفروق بينها وهي : X , Y , Z , W إلى النافذة المعنونة Test

Variables

- نضغط على المربع Statistics (واختيار هذا المربع اختياري) للحصول على الإحصاءات

الوصفية والربيعات فنحصل على صندوق حوار بعنوان : Several Related Samples:

. Stat...

ثالثاً: نربع مجموع الرتب لكل مجموعة أو لكل عينة.

نحسب إحصائية الاختبار بموجب العلاقة:

$$H = \frac{12}{N(N+1)} \sum_{j=1}^K \frac{R_j^2}{n_j} - 3(N+1) \quad (4)$$

تتبع إحصائية الاختبار H توزيع كاي تربيع χ^2 بدرجات حرية $(k-1)$.

٢-٣ الدلالة الإحصائية لأسلوب كروسكال – واليز عندما تكون $k > 3$ و $n_j > 5$

يُختبر فرض العدم القائل بتطابق توزيعات هذه المجتمعات بمقارنة قيمة إحصائية الاختبار المحسوبة H مع القيمة الحرجة المستخرجة من جداول χ^2 عند مستوى المعنوية المحدد ودرجات حرية $(k-1)$ ، يرفض فرض العدم إذا كانت قيمة H أكبر من أو تساوي قيمة χ^2 الحرجة.

٣-٣ الدلالة الإحصائية لأسلوب كروسكال – واليز عندما تكون $k \leq 5$ و $n_j \leq 5$

عندما يكون عدد العينات h فأقل وتضم كل عينة h مشاهدات فأقل نستخدم جدول كروسكال – واليز المرفق كملحق والذي يمكن الدخول إليه مباشرة بتحديد عدد العينات موضع المقارنة وعدد مشاهدات كل عينة وتكون قيمة إحصائية الاختبار المحسوبة H ذات دلالة إحصائية إذا تجاوزت القيمة الجدولية.

مثال (٤):

قامت الوحدة المهنية للاستشارات الاقتصادية بتنظيم دورات تدريبية لتعليم اللغة الإنكليزية باستخدام ثلاث طرق تعليمية (A, B, C) مختلفة وذلك لثلاث مجموعات من طلبة كلية الاقتصاد اختيروا بطريقة عشوائية. بعد انتهاء الدورة تم إخضاعهم لاختبار موحد، فكانت النتائج المبينة في الجدول التالي:

درجات اختبار المجموعات الثلاث

التي تعلمت اللغة الإنكليزية باستخدام الطرق التعليمية A, B, C

أفراد المجموعة	المجموعة الأولى	المجموعة الثانية	المجموعة الثالثة
	A	B	C
1	95	86	79
2	79	87	81
3	80	82	76
4	91	74	82
5	97	81	92
6	94	92	89

عند إجراء اختبارات الفروض المتعلقة بالفروق بين متوسطات ثلاثة مجتمعات أو أكثر (تحليل التباين) افترضنا أن المجتمعات الأصلية طبيعية وأن تبايناتها مجهولة ومتساوية. فإذا تحققت هذه الافتراضات فإنه يمكننا استخدام اختبار f لمعرفة ما إذا كانت الفروق المشاهدة بين متوسطات العينات راجعة للصدفة أم أنها تشير إلى فروق حقيقية بين متوسطات المجتمعات المناظرة. وفي حالة عدم تحقق هذه الافتراضات وأصبح من غير الممكن استخدام توزيع f لدراسة هذا النوع من المشاكل، فيمكن استخدام طريقة لا معلمية ملائمة تسمى اختبار مجموع الرتب rank sum test الذي يستخدم لاختبار تطابق توزيعات عدة مجتمعات ويُعرف باسم اختبار كروسكال واليز Kruskal – Wallis test ويطلق عليه أحياناً اسم اختبار H ، لأن إحصائية الاختبار يرمز لها بالرمز H ، تتطلب هذه الطريقة أن توزيعات المجتمعات مستمرة و لا تتطلب أي افتراض آخر، يعتبر أسلوب كروسكال واليز بديلاً لتحليل التباين أحادي الاتجاه في الأساليب اللامعلمية، ويصلح للمقارنة بين عدة عينات مستقلة حجم كل منها صغير جداً قد يصل إلى واحد أو اثنين و لا يتطلب تساوي أحجام العينات.

- الشروط: بالإضافة إلى تقسيم الوحدات إلى أعمدة وصفوف نفترض أن المتغير مستمراً وأن المشاهدات في كل قطاع قابلة للتحويل إلى رتب تصاعدية وأن لا يكون هناك تفاعل بين الصفوف والأعمدة.

- الفروض: فرض العدم ينص على عدم اختلاف توزيعات المجتمعات بينما ينص الفرض البديل على اختلافها.

نفترض أن مشكلة معينة تتضمن k من المجتمعات التي سحبت منها k عينة عشوائية (عينة عشوائية من كل مجتمع، حيث n_1 حجم العينة الأولى، n_2 حجم العينة الثانية، n_k حجم العينة رقم k).

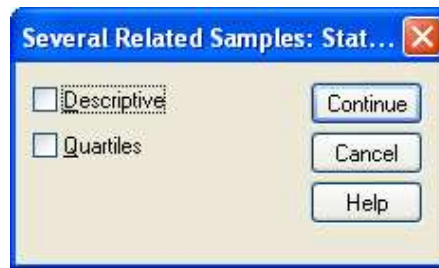
٣-١ اختبار كروسكال-واليز عندما يكون بكل عينة أكثر من خمسة مشاهدات

لإجراء اختبار كروسكال واليز نتبع الخطوات التالية:

أولاً: ندمج جميع مشاهدات العينات مع بعضها ثم نرتبها تصاعدياً من الأصغر إلى الأكبر، فنعطي الرتبة ١ لأصغر مشاهدة ثم الرتبة ٢ للمشاهدة التي تليها في الكبر وهكذا.. وفي النهاية تعطى أكبر قيمة الرتبة N ، حيث ترمز N إلى عدد المشاهدات في جميع العينات.

ثانياً: نحسب مجموع الرتب لكل عينة ونرمز لهذا المجموع بالرمز R_j ، حيث:

$$j = 1, 2, \dots, k$$



- نُفعل الخيارين Descriptive , Quartiles ثم نضغط Continue نعود إلى صندوق الحوار السابق حيث نقوم بتفعيل الخيار Cochran's Q ثم نضغط OK فنحصل على النتائج التالية :

Descriptive Statistics								
	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles		
						25th	50th (Median)	75th
W	12	.6667	.49237	.00	1.00	.0000	1.0000	1.0000
Z	12	.8333	.38925	.00	1.00	1.0000	1.0000	1.0000
Y	12	.5833	.51493	.00	1.00	.0000	1.0000	1.0000
X	12	.5833	.51493	.00	1.00	.0000	1.0000	1.0000

Cochran Test

Frequencies

	Value	
	0	1
W	4	8
Z	2	10
Y	5	7
X	5	7

Test Statistics

N	12
Cochran's Q	2.400 ^a
df	3
Asymp. Sig.	.494

a. 0 is treated as a success.

§-٣ اختبار كروسكال واليز لتحليل التباين في اتجاه واحد

Kruskal – Wallis One way analysis of variance

87	79	72	7
95	69	89	8
70	70	76	9
73	72	78	10
98	65	-	11
94	-	-	12

إذا علمت أن:

المجموعة الأولى تضم 10 طلاب، والمجموعة الثانية تضم 11 طالباً، المجموعة الثالثة تضم 12 طالباً.

والمطلوب:

اختبار فرض العدم القائل بتساوي مجتمعات درجات اختبار طرق التعليم الثلاث مقابل الفرض البديل القائل باختلافها عند مستوى معنوية 0.05.
إن مجموع المشاهدات في جميع العينات يساوي:

$$N = n_1 + n_2 + n_3 = 10 + 11 + 12 = 33$$

لإجراء الاختبار:

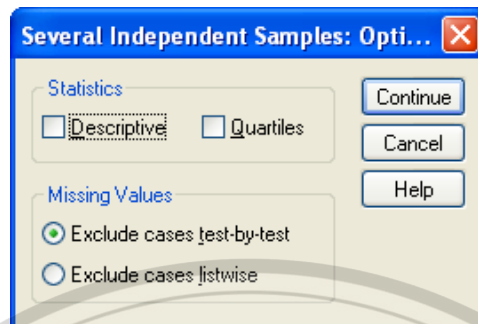
- ندمج جميع هذه المشاهدات.
- نرتبها ترتيباً تصاعدياً.
- نعطي كل مشاهدة الرتبة الملائمة.
- نقوم بحساب مجموع رتب كل مجموعة كما في الجدول التالي:

درجات اختبار ورتب المجموعات الثلاث

التي تعلمت اللغة الإنكليزية باستخدام الطرق التعليمية A , B , C

المجموعة الثالثة		المجموعة الثانية		المجموعة الأولى		أفراد المجموعة
الرتب	C	الرتب	B	الرتب	A	
13	79	20	86	30.5	95	1
16.5	81	21.5	87	13	79	2
9.5	76	18.5	82	15	80	3
18.5	82	8	74	25	91	4
26.5	92	16.5	81	32	97	5
23.5	89	26.5	92	28.5	94	6
21.5	87	13	79	5.5	72	7
30.5	95	2	69	23.5	89	8
3.5	70	3.5	70	9.5	76	9
7	73	5.5	72	11	78	10
33	98	1	65	-	-	11
28.5	94	-	-	-	-	12
$R_1 = 231.5$		$R_2 = 136$		$R_3 = 193.5$		مجموع الرتب

- نحدد القيمة الأدنى والقيمة الأعلى للمجال الذي يتحرك ضمنه المتغير Grouping Variable
- نضغط Continue نعود إلى صندوق الحوار السابق حيث نضغط على المربع Options .



- نُفعل الخيارين Descriptive , Quartiles ، ثم نضغط Continue فنعود إلى صندوق الحوار السابق .
- نُفعل الخيار Kruskal-Wallis H .
- نضغط OK فنحصل على النتائج التالية :

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles		
						25th	50th (Median)	75th
Education_Meth	33	82.5455	9.27729	65.00	98.00	75.0000	81.0000	91.5000
Groups2	33	2.0303	.84723	1.00	3.00	1.0000	2.0000	3.0000

Kruskal-Wallis Test

Ranks

	Groups2	N	Mean Rank
Education_Methods	1.00	11	19.41
	2.00	10	11.60
	3.00	12	19.29
	Total	33	

1	10	84.50	19.4	0.92
2	11	79.00	12.4	-1.95
3	12	84.50	19.3	1.03
Overall	33		17.0	

$H = 3.79$ $DF = 2$ $P = 0.150$

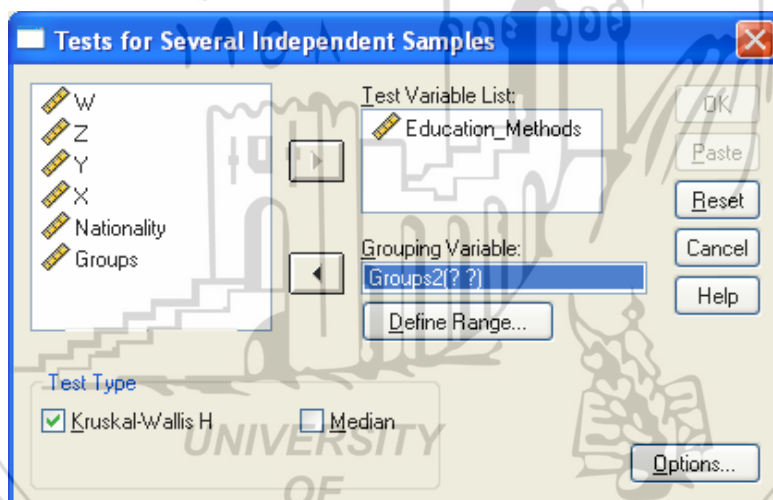
$H = 3.80$ $DF = 2$ $P = 0.149$ (adjusted for ties)

الحل :

من شريط القوائم نختار :

Analysis → Non Parametric → K Independent Samples

يظهر صندوق حوار بعنوان : Tests for Several Independent Samples :



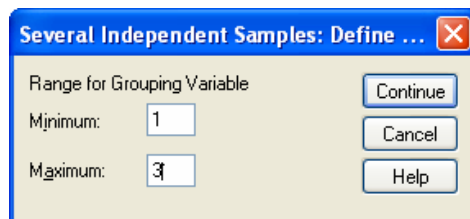
- ننقل المتغيرات الواجب مقارنة الفروق بينها وهي :

Education_Methods إلى النافذة المعنونة Test Variable List

Groups2 إلى النافذة المعنونة Grouping Variable List

- نضغط على المربع Define Range نحصل على صندوق حوار بعنوان :

Several Independent Samples: Define...



19.29167	12.36364	19.35	متوسط الرتب

نحسب قيمة إحصائية الاختبار:

$$\begin{aligned}
 H &= \frac{12}{N(N+1)} \sum_{j=1}^K \frac{R_j^2}{n_j} - 3(N+1) \\
 &= \frac{12}{33(33+1)} \left[\frac{(231.5)^2}{10} + \frac{(136)^2}{11} + \frac{(193.5)^2}{12} \right] - 3(33+1) \\
 &= 0.010695[9891.7] - 102 = 3.793587
 \end{aligned}$$

عند مستوى معنوية 0.05 ودرجات حرية $k-1 = 2$ فإن قيمة كاي تربيع الدرجة تساوي 5.99 وهذه القيمة أكبر من قيمة إحصائية الاختبار H المحسوبة وبالتالي لا نستطيع أن نرفض فرضية العدم أي أن طرق تعليم اللغة الإنكليزية الثلاث ليس لها تأثير معنوي مختلف على درجات الطلاب وذلك عند مستوى معنوية 0.05 ، جدير بالذكر أن اختبار مجموع الرتب هو دائماً اختبار ذو طرف أيمن.

- يمكن إجراء تصحيح عند وجود تداخلات ووجود مشاهدات متكررة مما يؤدي إلى تحسين النتائج ودقة الاختبار وذلك بقسمة إحصائية الاختبار على المقدار:

$$1 - \left(\sum_{i=1}^n T_i \right) / (N^3 - N) \quad , \quad \sum_{i=1}^n T_i = \sum_{i=1}^n (t_i^3 - t_i) \quad (3)$$

حيث: t_1 عدد التكرارات المتشابهة في حالة كونها الأولى.

t_2 عدد التكرارات المتشابهة في حالة كونها الثانية.

t_3 عدد التكرارات المتشابهة في حالة كونها الثالثة.

$$1 - \left(\sum_{i=1}^n T_i \right) / N^3 - N = 1 - (120) / 33^3 - 33 = 0.996658$$

$$H = 3.793587 / 0.996658 = 3.803 \approx 3.80$$

Kruskal-Wallis Test (MINITAB Outputs)

Kruskal-Wallis Test on C1

C2 N Median Ave Rank Z

Test Statistics^{a,b}

	Education_ Methods
Chi-Square	4.486
df	2
Asymp. Sig.	.106

a. Kruskal Wallis Test

b. Grouping Variable: Groups2

مثال (٥) :

قامت الوحدة المهنية للدراسات والبحوث الاجتماعية بكلية الآداب بتطبيق اختبار للاتزان الانفعالي على أربع مجموعات من الطلبة ينتمون لجنسيات مختلفة فكانت الدرجات على النحو التالي:

درجات اختبار المجموعات الأربع التي خضعت لاختبار للاتزان الانفعالي

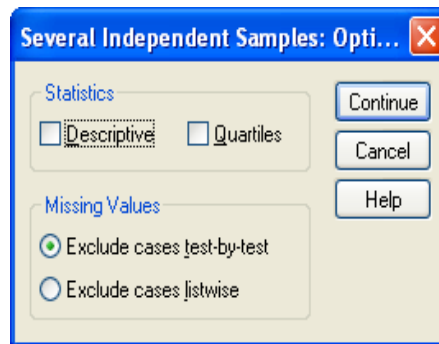
فرنسي الجنسية	عربي الجنسية	إنكليزي الجنسية	ياباني الجنسية
17	31	25	35
33	31	30	50
33	11	40	48
18	21	45	50
			52

إذا علمت أن المجموعة الأولى تضم ٤ طلاب، والمجموعة الثانية تضم 4 طلاب، المجموعة الثالثة تضم 4 طلاب، والمجموعة الرابعة تضم 5 طلاب.. والمطلوب اختبار فرض العدم القائل بتساوي مجتمعات درجات اختبار الاتزان الانفعالي مقابل الفرض البديل القائل باختلافها عند مستوى معنوية 0.05.

درجات ورتب المجموعات الأربع التي خضعت لاختبار الاتزان الانفعالي

فرنسي	الرتب	عربي	الرتب	إنكليزي	الرتب	ياباني	الرتب
17	2	31	7.5	25	5	35	11
33	9.5	31	7.5	30	6	50	15.5
33	9.5	11	1	40	12	48	14
18	3	21	4	45	13	50	15.5
						52	17
	24		20		36		73
	6		5		9		14.6

نحسب قيمة إحصائية الاختبار:



- نُفعل الخيارين Descriptive , Quartiles ، ثم نضغط Continue فنعود إلى صندوق الحوار السابق .

- نُفعل الخيار Kruskal-Wallis H .

- نضغط OK فنحصل على النتائج التالية :

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles		
						25th	50th (Median)	75th
Nationality	17	33.5294	12.61506	11.00	52.00	23.0000	33.0000	46.5000
Groups	17	2.5882	1.17574	1.00	4.00	1.5000	3.0000	4.0000

Kruskal-Wallis Test

Ranks

	Groups	N	Mean Rank
Nationality	1.00	4	6.00
	2.00	4	5.00
	3.00	4	9.00
	4.00	5	14.60
Total		17	

Test Statistics^{a,b}

	Nationality
Chi-Square	10.108
df	3
Asymp. Sig.	.018

a. Kruskal Wallis Test

b. Grouping Variable: Groups

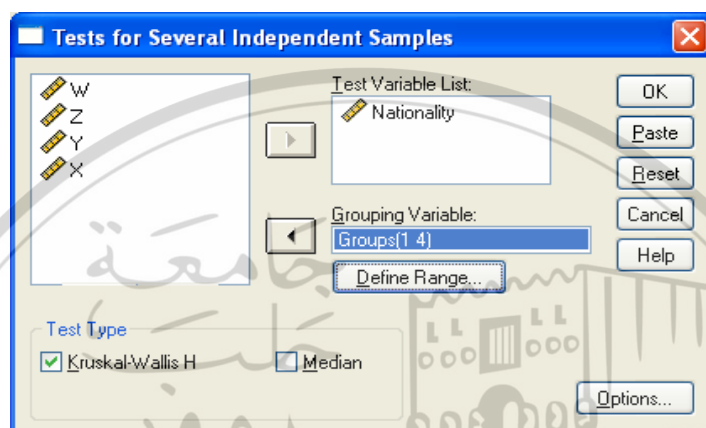
الحل :

من شريط القوائم نختار :

Analysis-→ Non Parametric-→ K Independent Samples

يظهر صندوق حوار بعنوان :

Tests for Several Independent Samples



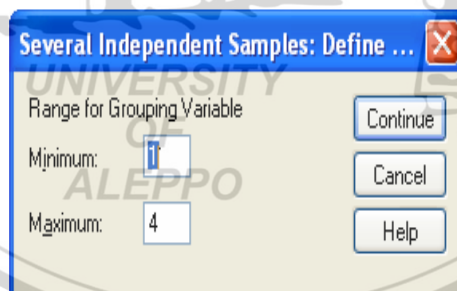
- ننقل المتغيرات الواجب مقارنة الفروق بينها وهي :

Nationality إلى النافذة المعنونة Test Variable List

Groups إلى النافذة المعنونة Grouping Variable List

- نضغط على المربع Define Range نحصل على صندوق حوار بعنوان:

Several Independent Samples: Define...



- نحدد القيمة الأدنى والقيمة الأعلى للمجال الذي يتحرك ضمنه متغير التجميع Grouping Variable

- نضغط Continue نعود إلى صندوق الحوار السابق حيث نقوم بتفعيل الخيار Options .

$$H = \frac{12}{N(N+1)} \sum_{j=1}^K \frac{R_j^2}{n_j} - 3(N+1)$$

$$= \frac{12}{17(17+1)} \left[\frac{(24)^2}{4} + \frac{(20)^2}{4} + \frac{(36)^2}{4} + \frac{(73)^2}{5} \right] - 3(17+1)$$

$$= 0.039216[1633.8] - 54 = 10.07$$

عند مستوى معنوية 0.01 ودرجات حرية $k-1 = 3$ فإن قيمة كاي تربيع الدرجة تساوي 11.34 وهذه القيمة أكبر من قيمة إحصائية الاختبار H المحسوبة وبالتالي لا نستطيع أن نرفض فرضية العدم أي أنه لا توجد فروق جوهرية في الاتزان الانفعالي باختلاف الجنسية وذلك عند مستوى معنوية 0.01 ، وعند مستوى معنوية 0.05 ودرجات حرية $k-1 = 3$ فإن قيمة كاي تربيع الدرجة تساوي 7.82 وبالتالي فإن فرضية العدم مرفوضة أي أن هناك فروق جوهرية في الاتزان الانفعالي باختلاف الجنسية وذلك عند مستوى معنوية 0.05 . ونظراً لوجود بعض المشاهدات المتكررة نقوم بإجراء تعديل على قيمة إحصائية الاختبار على الشكل التالي:

$$H = 10.07 / 0.996324 = 10.107 \approx 10.11$$

$$\sum_{i=1}^n T_i = \sum_{i=1}^n (t_i^3 - t_i) = (2^3 - 2) + (2^3 - 2) + (2^3 - 2) = 18$$

$$1 - (\sum_{i=1}^n T_i) / N^3 - N = 1 - (18) / 17^3 - 17 = 0.996324$$

مع العلم أن الإنكليز واليابانيين أفضل في الاتزان الانفعالي من الفرنسيين والعرب إلا أن هذا الأمر يتطلب التحقق من ذلك بين كل مجموعتين باستخدام اختبار مان ويتني.

Kruskal-Wallis Test

Kruskal-Wallis Test on C3

C4	N	Median	Ave Rank	Z
1	4	25.50	6.0	-1.36
2	4	26.00	5.0	-1.81
3	4	35.00	9.0	0.00
4	5	50.00	14.6	2.95
Overall	17		9.0	

$$H = 10.07 \quad DF = 3 \quad P = 0.018$$

$$H = 10.11 \quad DF = 3 \quad P = 0.018 \text{ (adjusted for ties)}$$

* NOTE * One or more small samples

The Chi-Square Distribution and the Analysis of Frequencies

مقدمة

كما نعلم بأن اختبار الفرضيات "اختبار Z" و " اختبار T " يستخدمان لقياس الدلالة الإحصائية للفروق وللتحقق من صحة الفروض المعلمية التي نضعها وذلك بالنسبة للبيانات الكمية، إلا أننا سنحاول هنا دراسة اختبار كاي تربيع الشائع الاستخدام لعدد من الحالات المختلفة والتميزة والتي لا يصلح معها استخدام الأسلوبين السابقين، حيث تكون البيانات المتاحة للتحليل معطاة إما على شكل تكرارات أو كمعطيات نوعية.

يمتاز مقياس كاي تربيع كتحليل التباين بأنه يضع أيدينا على الفروق دفعة واحدة وليس واحداً واحداً، كما هو الحال في مقياس الفرق بين متوسطين أو الفرق بين نسبتيين أو كل متوسطين أو كل نسبتيين على حدة، ومن هذه الاختبارات نذكر اختبار جودة التوفيق Tests of goodness of fit ، اختبار الاستقلالية Tests of independence ، واختبار التجانس Tests of homogeneity ، وسوف نكتشف بالتطبيق العملي أن كافة اختبارات كاي تربيع هي اختبارات جودة توفيق التكرارات المشاهدة مع التكرارات المتوقعة حيث نستخدم عبارة جودة التوفيق للإشارة فقط إلى مقارنة توزيع العينة مع توزيع نظري نعتقد بأنه يصف المجتمع الذي سحبت منه العينة.

يعتبر اختبار كاي تربيع ملائماً جداً للظواهر أو المتغيرات التي لها تصنيفات مختلفة مثل درجة التعليم وتصنيفاتها، السلعة المنتجة وأشكالها، المنفق على الدعاية والإعلان ومستوياته، العمر وتصنيفاته ومن ثم يكون المطلوب معرفة ما إذا كانت هناك علاقة معنوية بين هذه التصنيفات وتكراراتها، أي مدى تجانسها ، وذلك إذا كنا بصدد دراسة متغير واحد وهو ما يعرف باختبار التجانس . أو قد يكون المطلوب معرفة ما إذا كانت هناك علاقة أو لا بين تصنيف ظاهرتين أو متغيرين وهو ما يعرف باختبار الاستقلالية.

هل ينبغي أن تواصل الخريجة دراساتها العليا للحصول على درجة الماجستير والدكتوراه ؟
والإجابة على هذا السؤال تتضمن 3 اختيارات وهي: نعم ، لا ، محايد .
يهدف الباحث الاجتماعي إلى التأكد من وجود فرق بالرأي يرجع إلى الجنس.
إن فرضية العدم بهذه الحالة تتلخص بعدم وجود فرق في الاستجابات يرجع إلى الجنس . إن
القاعدة العامة لإيجاد قيمة كاي تربيع هي المعادلة التالية:

$$X^2 = \sum \left[\frac{(O_i - E_i)^2}{E_i} \right]$$

هناك إمكانية لإيجاد قيمة كاي تربيع لكل جنس على حدة ثم نجمع القيمتين، ولكن الأفضل إيجاد
معنوية الفروق دفعة واحدة.

والجدول رقم (٤) يوضح عدد الرجال وعدد النساء الذين أجابوا بنعم أو أجابوا لا أو كانت
الإجابة محايد (ويسمى بجدول التكرارات المشاهدة أو التجريبية).

جدول رقم (٤)
وجهات النظر بين الجنسين حول مسألة متابعة الدراسات العليا

الاستجابات	الرجال	النساء	المجموع
نعم	20	36	56
لا	25	5	30
محايد	30	34	64
المجموع	75	75	150

إذا لم يكن هناك فرق بين الجنسين، فإننا نتوقع أن نجد عدداً متساوياً في استجابات الجنسين
بالقول " نعم " وخاصة أن عدد الذكور يساوي عدد الإناث، وكذلك من المتوقع أن نجد عدداً
متساوياً من الاستجابات القائلة " لا " من قبل كلا الجنسين، والأمر نفسه ينطبق على الاستجابة
القائلة " محايد " بالنسبة لكلا الجنسين، وبهذا المنطق يمكننا أن نتوقع الحصول على العدد
المشترك لاستجابة الجنسين وهو 56 بخصوص الاستجابة " نعم " والعدد المشترك لاستجابة
الجنسين وهو 30 بخصوص الاستجابة " لا " أيضاً العدد المشترك لاستجابة الجنسين وهو 64
بخصوص الاستجابة " محايد " .

إذا لم يكن هناك من فرق بين الاستجابات يرجع إلى الجنس، فإننا نتوقع أن تكون هذه الاستجابات
مقسمة بالتساوي بين الجنسين أي 50 % للذكور و 50 % للإناث، فيكون لدينا 28 استجابة " نعم
للذكور و 28 استجابة " نعم " للإناث، وكذلك سيكون لدينا 15 استجابة " لا " للذكور و
15 استجابة " لا " للإناث، وأخيراً سيكون لدينا 32 استجابة " محايد " للذكور و 32 استجابة "

يساوي 35 ، فهل هذا التكرار 35 يختلف اختلافاً جوهرياً عن التكرار النظري أو المتوقع وهو 25 أي التكرار الواجب الحصول عليه استناداً إلى فرضية العدم القائلة بعدم وجود فروق ذات دلالة إحصائية ؟ وتعتبر هذه القيمة هي متوسط التكرارات النظرية.

يمكن الحصول على قيمة كاي تربيع بموجب المعادلة التالية:

$$X^2 = \frac{(O_i - E_i)^2}{E_i} = \frac{2(35-25)^2}{25} = \frac{200}{25} = 8$$

ولتفسير هذه القيمة نقارنها مع القيمة الحرجة لكاي تربيع عند مستوى معنوية 1 % ودرجة حرية تساوي الواحد والتي نجدها في جداول معدة خصيصاً لذلك مساوية إلى 6.635 وهي أصغر من القيمة المحسوبة لكاي تربيع مما يمكننا من رفض فرضية العدم ، كما أن القيمة الحرجة لكاي تربيع عند مستوى معنوية 5 % ودرجة حرية واحدة، والتي نجدها في جدول كاي تربيع مساوية إلى 3.84 والتي يمكننا من رفض فرضية العدم القائلة بعدم وجود فرق جوهري في وجهات نظر المجتمع الإحصائي المدروس بخصوص فكرة الزواج للخريج.

لقد تعاملنا في مثالنا السابق مع تكرار تجريبي واحد وهو 35، ولكن في الواقع لدينا تكراران هما 15 ويمثل عدد الأشخاص الذين رفضوا فكرة الزواج للخريج و 35 ويمثل عدد الأشخاص الذين أيدوا فكرة الزواج للخريج، وبالإمكان إيجاد قيمة كاي تربيع باستخدام هذين التكرارين دون إيجاد التكرار المتوقع النظري عن طريق المعادلة:

$$X^2 = \frac{(O_1 - O_2)^2}{O_1 + O_2} = \frac{(35-15)^2}{35+15} = \frac{400}{50} = 8$$

وهي القيمة نفسها التي حصلنا عليها سابقاً بالطريقة الأولى، وبالتدقيق في الطريقة الأولى نلاحظ بأننا استخدمنا الفرق بين التكرارين المشاهد والمتوقع، وضرربنا الناتج في 2 ، بينما في الطريقة الثانية لم نضرب بهذا الرقم أو غيره .

لقد كانت المشكلة المطروحة متعلقة بالمقارنة بين استجابتين فقط هما موافق أو معارض، ولكن يصادف الباحث الاقتصادي في الحياة العملية مشاكل تتطلب المقارنة بين عدة استجابات مثل موافق جداً ، موافق ، محايد ، معارض ، معارض جداً .

للتوضيح نورد المثال التالي الذي يلقي الضوء على مشكلة وجود أو عدم وجود فروق في وجهات النظر بين الجنسين الذكور والإناث بمسألة متابعة الدراسات العليا للإناث، في بحث اجتماعي تم سؤال 30 طالباً و 30 طالبة من طلبة دبلوم الدراسات العليا كافة التخصصات بكلية الاقتصاد السؤال التالي:

يتميز هذا النوع من الاختبارات باستخدامه للبيانات النوعية إلى جانب البيانات الكمية، شريطة أن تكون العينة احتمالية وأن لا يقل عدد المفردات في كل خلية من خلايا التكرار المتوقع عن خمس مفردات. بالإضافة لذلك فإن صياغة الفروض المتعلقة باختبار كاي تربيع، لاسيما الفرض البديل، ذات طبيعة خاصة حيث يعتبر دوماً اختباراً من طرف أيمن. إن إحصائية الاختبار X^2 تأخذ الصيغة التالية:

$$X^2 = \sum \left[\frac{(O_i - E_i)^2}{E_i} \right]$$

حيث نرمز عادة للتكرارات المشاهدة Observed Frequency بالحرف O وللتكرارات المتوقعة Expected Frequency بالحرف E، (عندما تكون فرضية العدم H_0 صحيحة). إن المقدار X^2 يقيس المدى الذي تقترب أو تبتعد فيه التكرارات المشاهدة من التكرارات المتوقعة، وعندما يكون هذا المقدار كبيراً سيؤدي ذلك إلى رفض فرضية العدم H_0 ، وعندما يكون التطابق تاماً بين التكرارات المشاهدة والتكرارات المتوقعة فإن قيمة X^2 ستكون معدومة وبالتالي لا نستطيع رفض فرضية العدم H_0 .
لتوضيح استخدامات كاي تربيع نورد المثال التالي:

لنفترض أننا أجرينا بحثاً اجتماعياً يهدف لاستطلاع رأي مجموعة من خريجي كلية الاقتصاد قسم العلوم المالية والمصرفية من الذكور المتزوجين وعددهم 50 شخصاً، وبنتيجة الاستطلاع تبين أن 35 شخصاً منهم قد أجابوا بأن الزواج فكرة طيبة للخريج وأن 15 شخصاً منهم رفضوا فكرة الزواج، والمطلوب معرفة ما إذا كان لهذا الفرق دلالة حقيقية بمعنى هل يوجد هناك فروقاً في وجهات النظر بين أفراد المجتمع الأصلي، إذ توضح الغالبية العظمى تقريباً من خريجي كلية الاقتصاد قسم العلوم المالية والمصرفية موافقتها على فكرة الزواج للخريج.

إن فرضية العدم بهذه الحالة تقول بعدم وجود فروق حقيقية في الرأي بين أفراد المجتمع الأصلي، وبالتالي تقسم المجتمع إلى نصفين متساويين أي 50% يوافقون على فكرة الزواج للخريج، و 50% يرفضون هذه الفكرة وبالتالي يتساوى المؤيدون مع المعارضين لفكرة الزواج. واستناداً إلى هذا التقسيم: 50% إلى 50% فإن التكرارات التي نتوقعها تصبح 25، ويعرف هذا بالتكرار المتوقع أو التكرار النظري ونرمز عادة للتكرار النظري أو التكرار المتوقع بالرمز E، أما التكرار الحقيقي الذي حصلنا عليه من واقع التجربة والاستقصاء الفعلي فنطلق عليه التكرار المشاهد أو التكرار التجريبي ونرمز له بالرمز O، وهو في مثالنا التجريبي

محايد " للإناث. وهذه الاستجابات المتماثلة للجنسين تمثل التكرارات المتوقعة في ضوء فرضية العدم.

والجدول رقم (٥) التالي يوضح التكرارات المتوقعة أو النظرية المتعلقة بوجهات النظر بين الجنسين حول مسألة متابعة الدراسات العليا.

جدول رقم (٥)
التكرارات المتوقعة أو النظرية

الاستجابات	الرجال	النساء	المجموع
نعم	28	28	56
لا	15	15	30
محايد	32	32	64
المجموع	75	75	150

لحساب قيمة كاي تربيع نوجد الفرق بين التكرارات المشاهدة والتكرارات المتوقعة $(O - E)$ لكل استجابة ولكل جنس .

- نربع الفرق بين التكرارات المشاهدة والتكرارات المتوقعة $(O - E)^2$ لكل استجابة ولكل جنس .

- نقسم مربع الفرق بين التكرارات المشاهدة والتكرارات المتوقعة على قيمة التكرار المتوقع المقابل لكل قيمة $(O - E)^2 / E$ بالنسبة لكل استجابة ولكل جنس ثم نجمع حاصل القسمة بالنسبة للجنسين فنحصل على المطلوب، بعد الاستعانة بجدول الحسابات المساعد .

جدول رقم (٦)
العمليات الحسابية المطلوبة لإيجاد قيمة كاي تربيع

الاستجابات	$(O - E)$		$(O - E)^2$		$(O - E)^2 / E$	
	الرجال	النساء	الرجال	النساء	الرجال	النساء
نعم	-8	8	64	64	2.286	2.286
لا	10	-10	100	100	6.667	6.667
محايد	-2	2	4	4	0.125	0.125
المجموع	0	0	164	164	9.077	9.077

إن قيمة كاي تربيع هي حاصل جمع مربع الفرق بين التكرارات المشاهدة والتكرارات المتوقعة على قيمة التكرار المتوقع المقابل لكل قيمة $(O - E)^2 / E$ بالنسبة لكل من الرجال والنساء وتساوي بمثلنا إلى 18.15 أي أننا جمعنا بالنسبة للخلايا الست الموجودة في التجربة.

إن القيمة الكبيرة لكاي تربيع المحسوبة وهي 18.15 توضح مدى تباعد التكرارات المشاهدة عن التكرارات المتوقعة التي كنا سنحصل عليها لو لم يكن هناك فرق يرجع إلى الجنس . لتفسير هذه القيمة نقارنها مع القيمة الحرجة لكاي تربيع عند مستوى معنوية 1 % ودرجات حرية تساوي

وعند مستوى معنوية 0.05 فإن القيمة الجدولية أو القيمة الحرجة لكاي تربيع و 5 درجات حرية $\nu = J - 1 = 6 - 1 = 5$ تساوي 11.070.

- لحساب قيمة إحصائية الاختبار نتبع الخطوات التالية وفق تسلسلها في جدول الحسابات المساعد .

أولاً : نخصص العمود الأول للنتائج الممكنة والتي يمكن الحصول عليها نتيجة رمي قطعة النرد، وهي معروفة وستكون واحدة من النتائج التالية: 1، 2، 3، 4، 5، 6.

ثانياً : نخصص العمود الثاني لل تكرارات المشاهدات التجريبية O التي حصلنا عليها فعلياً عند رمي قطعة النرد 480 مرة .

ثالثاً : نخصص العمود الثالث لاحتمال تحقق كل نتيجة من النتائج الممكنة التي يمكن الحصول عليها بالاستناد إلى صحة فرضية العدم القائلة بأن الزهر كامل الاتزان وبالتالي فإن احتمال ظهور كل وجه من الأوجه الستة يساوي $1/6$.

رابعاً : نحدد التكرارات النظرية أو التكرارات المتوقعة E استناداً لصحة فرضية العدم وأن احتمال ظهور كل وجه من الأوجه الستة يساوي $1/6$ ، وحيث أن عدد المحاولات $n = 480$ فإن التكرار المتوقع لكل خلية أو المقابل لكل نتيجة ممكنة سيكون مساوياً إلى : $(1/6) \cdot 480 = 80$.

خامساً : نحسب الفرق بين التكرارات المشاهدات والتكرارات المتوقعة $(O - E)$ لكل نتيجة من النتائج الست .

سادساً : نربع الفرق بين التكرارات المشاهدات والتكرارات المتوقعة $(O - E)^2$ لكل نتيجة من النتائج الست .

سابعاً : نقسم مربع الفرق بين التكرارات المشاهدات والتكرارات المتوقعة على قيمة التكرار المتوقع المقابل لكل قيمة $(O - E)^2 / E$ لكل نتيجة من النتائج الست ثم نجمع حاصل القسمة بالنسبة للنتائج الست المختلفة، فنحصل على المطلوب بعد الاستعانة بجدول الحسابات المساعد .

جدول رقم (٧)

العمليات الحسابية المطلوبة لحساب قيمة إحصائية الاختبار

النتائج الممكنة	التكرارات المشاهدة	الاحتمال	التكرار المتوقع	$(O - E)$	$(O - E)^2$	$(O - E)^2 / E$
1	65	0.1667	80	15	225	2.8125

إحصائية الاختبار X^2 تتبع تقريباً توزيع كاي تربيع بدرجات حرية تساوي $(J - 1)$ ، حيث J تمثل عدد الفئات في جدول التوزيع التكراري.

فإذا كانت قيمة الإحصائية صغيرة جداً فإن هذا يدل على جودة التوفيق وبالتالي لا نستطيع رفض فرض العدم، ويتم رفض فرض العدم إذا كانت قيمة إحصائية الاختبار أكبر من أو تساوي القيمة الحرجة المستخرجة من جداول كاي تربيع بدرجات حرية: $(J - 1)$ والتي تساوي المساحة الواقعة على يمينها مستوى المعنوية .

مثال (٥) :

يرغب أحد المشرفين الاجتماعيين في المدينة الجامعية التأكد ما إذا كان جميع الطلبة الذين يشرف عليهم متمائلين في هواياتهم الرياضية والثقافية، وهل باستطاعته أن يوزعهم على اللجان الفرعية (رياضية ، ثقافية ، فنية ، علاقات عامة واجتماعية...) بالتساوي من منطلق تحقيق العدالة ومبدأ تكافؤ الفرص. لهذا السبب رغب في إجراء تجربة بسيطة يستدل من خلالها، إذا كان زهر نرد معين كامل الاتزان أو لا (وهل من الضروري في حالة كون الزهر متزناً أن تكون النتائج للأوجه الستة المختلفة للنرد متساوية) ، ولتحقيق هذا الغرض ألقى الزهر 480 مرة وسجلت نتائجها . وبالطبع فإن النتائج الست الممكنة لهذه التجربة هي الأوجه: 1 ، 2 ، 3 ، 4 ، 5 ، 6. فإذا علمت أن التكرارات المشاهدة التجريبية المناظرة لكل من هذه النتائج الممكنة هي على التوالي:

65 ، 85 ، 90 ، 78 ، 80 ، 82

هل تؤدي هذه البيانات إلى عدم رفض الفرض الصفري (العدم) القائل بعدم وجود فروق معنوية بين النتائج الممكنة لزهر النرد (القائل بأن الزهر كامل الاتزان) وذلك عند مستويي معنوية 0.05 و 0.01 ؟

الحل:

الفرضيات: سنختبر فرضية العدم القائلة بأن الزهر كامل الاتزان (عدم وجود فروق معنوية بين الأوجه المختلفة لقطعة النرد ، مقابل الفرضية البديلة القائلة بوجود فروق معنوية بين الأوجه المختلفة لقطعة النرد (النرد غير متزن) .

- عند مستوى معنوية 0.01 ، إن القيمة الجدولية أو القيمة الحرجة لكاي تربيع المستخرجة من جداول معدة خصيصاً لذلك و 5 درجات حرية $5 = 6 - 1 = J - 1 = \nu$ تساوي 15.0863 ،

² التي نجدها في جداول معدة خصيصاً لذلك مساوية إلى 9.21 وهي أصغر من القيمة المحسوبة لكاي تربيع مما يمكننا من رفض فرضية العدم .

كما أن القيمة الحرجة لكاي تربيع عند مستوى معنوية 5 % ودرجات حرية تساوي 2 والتي نجدها في جدول كاي تربيع مساوية إلى 5.99 التي يمكننا من رفض فرضية العدم القائلة بعدم وجود فرق جوهري في استجابات الجنسين يرجع إلى الجنس في المجتمع الإحصائي المدروس ، ونستنتج بوجود فرق في الاستجابات بين النساء والرجال في الرغبة في مواصلة الدراسات العليا.

لقد اتضح لنا بجلاء، أن هناك بشكل عام فرقاً بين الجنسين في هذه الاستجابات ولكن لمعرفة موقع هذا الفرق بالضبط، أي يعود لأي نوع من الاستجابات، ننظر في الجدول السابق الذي يوضح العمليات الحسابية اللازمة لحساب قيمة كاي تربيع فنجد أن معظم قيم كاي تربيع الفرعية في الجدول هي صغيرة في جميع الحالات باستثناء ما يخص استجابة واحدة وهي الاستجابة " لا " حيث كان هناك فرق واضح في هذه الاستجابة بين الرجال والنساء (25 رجلاً مقابل 5 نساء فقط) وقيمة كاي تربيع المقابلة تعادل 13.35، وبالنظر إلى جدول كاي تربيع عند درجة حرية واحدة سنجد بأن هناك فرقاً ذا دلالة إحصائية تفوق مستوى معنوية 1 % ومعنى ذلك سيكولوجياً أن الرجال يقررون بحزم عدم موافقتهم على استمرار النساء بمتابعة مرحلة الدراسات العليا، أما إذا أخذنا بعين الاعتبار الذين لم يقرروا بعد والذين قالوا نعم فإن هذا الفرق يصبح أقل نوعاً ما.

§-5 اختبارات جودة التوفيق Tests of goodness-of-fit

يستخدم اختبار كاي تربيع لجودة التوفيق بهدف الاستدلال إحصائياً على توزيع مجتمع وذلك على أساس التوزيع الذي نحصل عليه من بيانات العينة، فمثلاً قد نرغب في تحديد ما إذا كانت عينة من القيم المشاهدة لمتغير عشوائي تتوافق مع فرضية كونها مسحوبة من مجتمع طبيعي أم لا ؟ ولإجراء الاختبار يجب وضع بيانات العينة في صورة توزيع تكراري يسمى التكرار المشاهد ومن ثم تقارن بالتكرارات المتوقعة التي نحصل عليها باستخدام التوزيع النظري المحدد في فرضية العدم وهو التوزيع الطبيعي.

تستخدم إحصائية الاختبار X^2 كمقياس للفرق بين التوزيعين المشاهد والمتوقع ، فإذا كان حجم العينة كبيراً بدرجة تكفي لأن يكون التكرار المتوقع في أية خلية لا يقل عن 5 وحدات فإن

$$^2 \text{ درجات الحرية} = (\text{عدد الأعمدة} - 1) (\text{عدد الأسطر} - 1) = (3 - 1) (2 - 1) = 2$$

2	85	0.1667	80	-5	25	0.3125
3	90	0.1667	80	-10	100	1.25
4	78	0.1667	80	2	4	0.05
5	80	0.1667	80	0	0	0
6	82	0.1667	80	-2	4	0.05
	480	1	480			4.475

إن قيمة كاي تربيع هي حاصل جمع مربع الفرق بين التكرارات المشاهدة والتكرارات المتوقعة على قيمة التكرار المتوقع المقابل لكل قيمة $(O - E)^2 / E$ بالنسبة لكل نتيجة وتساوي بمثلنا 4.475 أي أننا جمعنا بالنسبة للخلايا الست الموجودة في التجربة وذلك كما يلي:

$$X^2 = \sum_{i=1}^6 \left[\frac{(O_i - E_i)^2}{E_i} \right] = 4.475$$

إن القيمة الصغيرة لكاي تربيع المحسوبة وهي 4.475 توضح مدى قرب بعض التكرارات المشاهدة عن التكرارات المتوقعة التي كنا سنحصل عليها لو لم يكن هناك فرق بين الأوجه الستة لقطعة النرد.

لتفسير هذه القيمة نقارنها مع القيمة الحرجة لكاي تربيع عند مستوى معنوية 1 % ودرجات حرية تساوي 5 التي تساوي 15.086 وهي أكبر من القيمة المحسوبة لكاي تربيع مما لا يمكننا من رفض فرضية العدم.

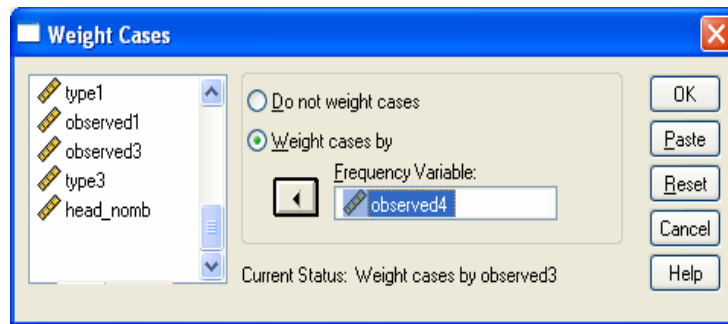
كما أن القيمة الحرجة لكاي تربيع عند مستوى معنوية 5 % و 5 درجات حرية تساوي 11.070 والتي لا يمكننا أيضاً من رفض فرضية العدم القائلة بعدم وجود فرق جوهري بين الأوجه الستة لقطعة النرد ، ونستنتج عدم وجود فرق بين الأوجه المختلفة لقطعة النرد وبالتالي فإن النرد كامل الاتزان ، والفروق التي حصلنا عليها نتيجة التجربة هي ليست ذات دلالة إحصائية وترجع لعامل المصادفة.

الحل :

لاختبار الفرضية السابقة باستخدام اختبار كاي تربيع Chi-Square Test نتبع ما يلي:

- نقوم بترتيب البيانات في محرر البيانات Data Editor على النحو التالي:

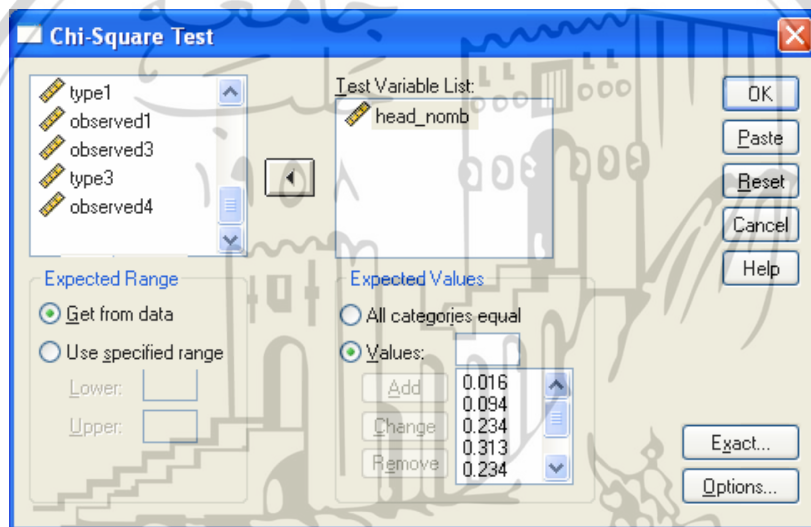
Type	Observed3
1	65
2	85
3	90
4	78



نُفعل الخيار by Weight Cases ثم نقوم بإدخال المتغير Observed4 إلى خانة الخيار الأخير Frequency Variable ثم نضغط OK .
من شريط القوائم نختار :

Analysis → Non Parametric → Chi-Square

يظهر صندوق حوار بعنوان : Chi-Square-Test الذي نقوم بترتيبه على النحو المبين في الشكل أدناه.



- ننقل المتغير head_nomb إلى النافذة المعنونة Test Variable List

في خانة Expected Values :

يتم تفعيل الخيار Values عندما تكون التكرارات المتوقعة مختلفة وبالتالي تكون النسب المئوية لعدد الصور مختلفة ، ولهذا السبب يتم إدخال النسب المختلفة في المربع Values الواحدة تلو الأخرى وننقر الزر Add بالنسبة لكل نسبة ثم نضغط OK فنحصل على النتائج التالية:
إن قيمة إحصائية الاختبار لكاي تربيع تساوي 156.023 ، عدد درجات الحرية يساوي 6 ، إن قيمة مستوى المعنوية الفعلي :

$$P\text{-Value} = \text{Asymp.Sig} = 0.000$$

أصغر من مستوى المعنوية المفترض ٠,٠٥ مما يدعونا إلى رفض فرضية العدم أي أن عدد الصور المشاهدة لا تتبع توزيع ثنائي الحدين.

type3

	Observed N	Expected N	Residual
1.00	65	80.0	-15.0
2.00	85	80.0	5.0
3.00	90	80.0	10.0
4.00	78	80.0	-2.0
5.00	80	80.0	.0
6.00	82	80.0	2.0
Total	480		

Test Statistics

	type3
Chi-Square ^a	4.475
df	5
Asymp. Sig.	.483

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 80.0.

مثال (٢) :

تم رمي ست قطع نقود 1000 مرة وفي كل مرة حسبت عدد الصور Heads في كل مرة وكانت النتائج على الشكل التالي: ٣٥ ، ١٢٥ ، ٣٥٠ ، ٢٨٠ ، ١٧٠ ، ٣٥ ، ٥ .

باستخدام اختبار كاي تربيع Chi-Square اختبر الفرضية القائلة بأن عدد مرات ظهور الصورة يتبع قانون توزيع ثنائي الحدين Binomial عند مستوى معنوية % .

الحل: لاختبار الفرضية باستخدام اختبار كاي تربيع Chi-Square Test :

- نقوم بترتيب البيانات في محرر البيانات Data Editor على النحو التالي:

Head_nomb	Observed
0	35
1	125
2	350
3	280
4	170
5	35
6	5

من شريط القوائم نختار :

Data → Weight Cases

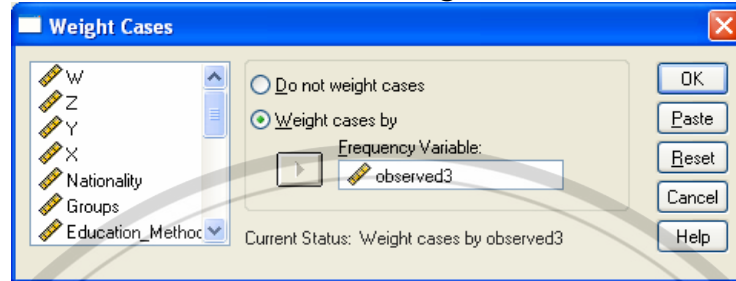
يظهر صندوق حوار بعنوان : Weight Cases .

5	80
6	82

من شريط القوائم نختار :

Data → Weight Cases

يظهر صندوق حوار بعنوان : Weight Cases .

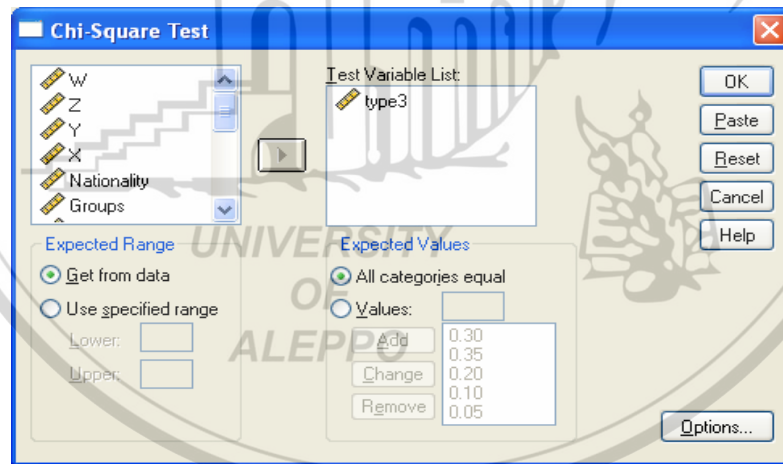


نُفعل الخيار Weight Cases by ثم نقوم بإدخال المتغير Observed3 إلى خانة الخيار الأخير Frequency Variable ثم نضغط OK .

من شريط القوائم نختار :

Analysis → Non Parametric → Chi-Square

يظهر صندوق حوار بعنوان : Chi-Square-Test الذي نقوم بترتيبه على النحو المبين في الشكل أدناه.



- ننقل المتغير Type3 إلى النافذة المعنونة Test Variable List

في خانة Expected Values نختار المربع :

١- All Categories equal : جميع الفئات لها نفس التكرار المتوقع أو النسبة المئوية لكل صفة أو فئة متساوية، ثم نضغط OK فنحصل على النتائج التالية:

NPar Tests

Chi-Square Test - Frequencies

NPar Tests

Chi-Square Test - Frequencies

head_nomb

	Observed N	Expected N	Residual
.0000000	35	16.0	19.0
1.0000000	125	93.9	31.1
2.0000000	350	233.8	116.2
3.0000000	280	312.7	-32.7
4.0000000	170	233.8	-63.8
5.0000000	35	93.9	-58.9
6.0000000	5	16.0	-11.0
Total	1000		

Test Statistics

	head_nomb
Chi-Square ^a	156.023
df	6
Asymp. Sig.	.000

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 16.0.

باستخدام الدالة : $\text{Binmomdist}(0;6;0.50;0)$ تم حساب احتمال ظهور الصورة من صفر إلى ستة وذلك كما هو موضح في الجدول التالي :

عدد الصور	الدالة	الاحتمال
0	$\text{Binmomdist}(0;6;0.50;0)$	0.015625
1	$\text{Binmomdist}(1;6;0.50;0)$	0.09375
2	$\text{Binmomdist}(2;6;0.50;0)$	0.234375
3	$\text{Binmomdist}(3;6;0.50;0)$	0.3125
4	$\text{Binmomdist}(4;6;0.50;0)$	0.234375
5	$\text{Binmomdist}(5;6;0.50;0)$	0.09375
6	$\text{Binmomdist}(6;6;0.50;0)$	0.015625

مثال (٧) :

يهتم الباحثون في الجمهورية الفرنسية بمعرفة الحالة التعليمية لجميع من لهم حق الانتخاب، وهل طرأ تغير إيجابي أو سلبي على هذه الحالة بعد مرور الزمن، وهل قامت

كما أن القيمة الحرجة لكاي تربيع عند مستوى معنوية 5 % و 4 درجات حرية تساوي 9.49 والتي تمكننا أيضاً من رفض فرضية العدم القائلة بعدم وجود فرق جوهري بين الحالة التعليمية لمن لهم حق الانتخاب في فرنسا الآن والحالة التعليمية منذ 10 سنوات ماضية. ونستنتج وجود فرق بين الحالتين مما يدل بوضوح أن الحالة

التعليمية لمن لهم حق الانتخاب قد طرأ عليها تغيرات جذرية لصالح المتعلمين وحملة الشهادات العليا .

لقد اتضح لنا بجلاء من أن هناك بشكل عام فرقاً بين الحالة التعليمية لمن لهم حق الانتخاب حالياً والحالة التعليمية منذ 10 سنوات، ولكن لمعرفة موقع هذا الفرق أو التغير بالضبط (أي يعود لأي حالة من حالات التعليم) ننظر في الجدول السابق الذي يوضح العمليات الحسابية اللازمة لحساب قيمة كاي تربيع فنجد أن معظم قيم كاي تربيع الفرعية في الجدول هي كبيرة في جميع الحالات وهي أكثر حدة في مرحلة التعليم الابتدائي ومرحلة التعليم الثانوي ومرحلة الدراسات العليا حيث كان التغير في هذه المراحل متماثلاً وكبيراً مقارنة بالتغير الذي طرأ على الحالتين التعليميتين المتوسطة والجامعية.

الحل : لاختبار الفرضية باستخدام اختبار كاي تربيع Chi-Square Test نتبع الخطوات التالية:

- نقوم بترتيب البيانات في محرر البيانات Data Editor على النحو التالي:

الحالة التعليمية	Type	Observed
ابتدائي	1	200
متوسط	2	310
ثانوي	3	280
جامعي	4	120
دراسات عليا	5	90

- لقد استعضنا عن الصفات (الحالة التعليمية) بالمتغير العددي Type لعدم إمكانية التعامل مع المتغيرات الرمزية أي عدم إمكانية ذكر أسماء الصفات مع العلم أنه بلامكان إضافة عنوان القيمة Value Label إلى المتغير Type لكي تعرض عناوين القيم في جدول المخرجات ، أما المتغير Observed فيمثل التكرار المشاهد في كل فئة وتستعمل قيم هذا المتغير كأوزان للحالات.

- في ورقة العمل المسماة عرض المتغيرات Variable View نقوم بإضافة عنوان القيمة Value Label إلى المتغير Type لكي تعرض عناوين القيم في جداول

ثالثاً : نخصص العمود الثالث لاحتمال تحقق كل حالة (النسبة المخصصة لكل حالة)

من الحالات الممكنة والتي يمكن الحصول عليها بالاستناد إلى الاستطلاع السابق .

رابعاً : نحدد التكرارات النظرية أو التكرارات المتوقعة E استناداً لصحة فرضية العدم وأن احتمال ظهور كل حالة من الحالات الخمس يساوي النسب نفسها التي تحققت في الاستطلاع الأول ، وبما أن عدد أفراد العينة يساوي $n = 1000$ فإن التكرار المتوقع لكل خلية أو المقابل لكل حالة تعليمية ممكنة سيكون مساوياً إلى حاصل ضرب العدد 1000 في الاحتمال أو النسبة المقابلة لها .

خامساً : نحسب الفرق بين التكرارات المشاهدة والتكرارات المتوقعة $(O - E)$.

سادساً : نربع الفرق بين التكرارات المشاهدة والتكرارات المتوقعة $(O - E)^2$.

سابعاً : نقسم مربع الفرق بين التكرارات المشاهدة والتكرارات المتوقعة على قيمة التكرار المتوقع المقابل لكل قيمة $(O - E)^2 / E$ ثم نجمع حاصل القسمة ، فنحصل على المطلوب ، بعد الاستعانة بجدول العمليات الحسابية .

جدول رقم (٨)
العمليات الحسابية المطلوبة لحساب قيمة إحصائية الاختبار

الحالة التعليمية	التكرارات المشاهدة	الاحتمال	التكرار المتوقع	$(O - E)$	$(O - E)^2$	$(O - E)^2 / E$
ابتدائي	200	0.3	300	-100	10000	33.33
متوسط	310	0.35	350	-40	1600	4.57
ثانوي	280	0.2	200	80	6400	32.00
جامعي	120	0.1	100	20	400	4.00
دراسات عليا	90	0.05	50	40	1600	32.00
المجموع	1000	1	1000	0		105.90

إن قيمة كاي تربيع في مثالنا تساوي 105.9 .

$$X^2 = \sum_{i=1}^5 \left[\frac{(O_i - E_i)^2}{E_i} \right] = 105.9$$

وهذه القيمة كبيرة جداً توضح مدى تباعد بعض التكرارات المشاهدة عن التكرارات المتوقعة التي كنا سنحصل عليها لو لم يكن هناك فرق بين الحالة التعليمية حالياً وبين الحالة التعليمية لمن لهم حق الانتخاب في فرنسا منذ 10 سنوات .

لتفسير هذه القيمة نقارنها مع القيمة الحرجة لكاي تربيع عند مستوى معنوية 1 % ودرجات حرية تساوي 4 والتي تساوي 13.28 وهي أصغر من القيمة المحسوبة لكاي تربيع مما يمكننا من رفض فرضية العدم .

الحكومة بتحقيق قفزة في مجال التعليم أو لا، مفترضين أنه قد تم منذ 10 سنوات ماضية تصنيف جميع من لهم حق الانتخاب حسب حالتهم التعليمية إلى خمسة أقسام متنافية وشاملة وهي:
الصف السادس الابتدائي أو أقل ، التعليم المتوسط ، التعليم الثانوي ، التعليم الجامعي ، مرحلة الدراسات العليا، وقد كانت النسب المناظرة لهذه الأقسام هي على الترتيب :

$$0.05 ، 0.10 ، 0.20 ، 0.35 ، 0.30$$

يرغب الباحثون في معرفة ما إذا كان التوزيع الحالي للحالة التعليمية لمن لهم حق الانتخاب في فرنسا ما زال هو نفس التوزيع السابق الذي تم الحصول عليه منذ 10 سنوات ماضية، ولتحقيق هذا الغرض سحبت عينة عشوائية تضم 1000 شخص ممن لهم حق الانتخاب حيث حصلنا على التكرارات التجريبية المشاهدة المناظرة للأقسام التعليمية الخمسة السابقة على الترتيب وهي :

$$90 ، 120 ، 280 ، 310 ، 200$$

والمطلوب:

معرفة ما إذا كانت هذه البيانات تدل على تغير توزيع الحالة التعليمية في الجمهورية الفرنسية لمن لهم حق الانتخاب عما كان عليه في الماضي وذلك عند مستوى معنوية 0.05 وكذلك 0.01 .

الحل :

- **الفرضيات:** سنختبر فرضية العدم القائلة بأن الحالة التعليمية لمن لهم حق الانتخاب في فرنسا لم يطرأ عليها تغيير (عدم وجود فروق معنوية بين الحالة التعليمية الآن والحالة التعليمية منذ 10 سنوات ماضية) مقابل الفرضية البديلة القائلة بوجود فرق معنوي بين الحالة التعليمية الآن لمن لهم حق الانتخاب والحالة التعليمية لهؤلاء الأشخاص منذ 10 سنوات ماضية.

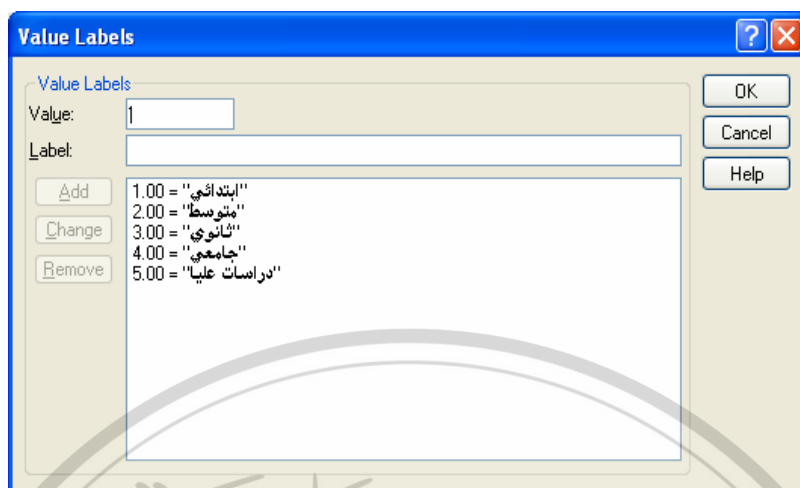
- عند مستوى معنوية 0.01 فإن القيمة الجدولية أو القيمة الحرجة لكاي تربيع المستخرجة من جداول معدة خصيصاً لذلك وعند ٤ درجات حرية $J - 1 = 5 - 1 = 4$ تساوي 13.28 ، أما عند مستوى معنوية 0.05 فإن القيمة الجدولية أو القيمة الحرجة لكاي تربيع عند 5 درجات حرية $J - 1 = 5 - 1 = 4$ تساوي 9.49 .

- لحساب قيمة إحصائية الاختبار نتبع نفس الخطوات السابقة وفق تسلسلها في جدول الحسابات المساعد:

أولاً : نخصص العمود الأول للحالات التعليمية لمن لهم حق الانتخاب.

ثانياً : نخصص العمود الثاني للتكرارات المشاهدة أو التجريبية O التي حصلنا عليها فعلياً عند سحب عينة عشوائية تضم 1000 شخص .

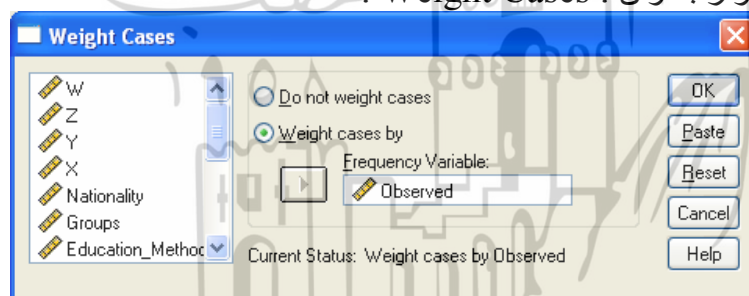
المخرجات.



من شريط القوائم نختار :

Data → Weight Cases

يظهر صندوق حوار بعنوان : Weight Cases .



نُفعل الخيار Weight Cases by ثم نقوم بإدخال المتغير Observed إلى خانة الخيار الأخير

Frequency Variable ثم نضغط OK .

من شريط القوائم نختار :

Analysis → Non Parametric → Chi-Square

يظهر صندوق حوار بعنوان : Chi-Square-Test

الذي نقوم بترتيبه على النحو المبين في الشكل أدناه.

المجتمعين لهما نفس التوزيع . ولاختبار فرضية أن المجتمعين لهما نفس التوزيع نستخدم اختبار الإشارة الذي يشترط أن تكون البيانات المتاحة عن العينتين رتبيه على الأقل، سواء تم اختيار العينة بطريقة عشوائية أو بطريقة غير عشوائية.

خطوات الاختبار:

- إن البيانات اللازمة لاختبار الإشارة بين عينتين غير مستقلتين هي N ثنائية من المشاهدات الخاصة بمتغيرين X و Y مرتبطين عادة من خلال الأزواج المتطابقة matched-pairs ، حيث نضع بيانات كل متغير في عمود مستقل (عادة يخصص الأول لبيانات الاختبار الأولي والثاني لبيانات الاختبار اللاحق).

- نكتب إشارات الفروق بين المتغيرين السابقين موجبة كانت أو سالبة بدون النظر إلى قيمة الفرق.

- نحسب عدد الثنائيات ذات الإشارات الموجبة D^+ و عدد الثنائيات ذات الإشارات السالبة D^- ونختار منهما القيمة الأقل لتكون قيمة إحصائية الاختبار المحسوبة أو القيمة التي سيتم تحويلها إلى قيمة معيارية⁴.

إذا كان فرض عدم القائل بتطابق المجتمعين المدروسين صحيحاً فإننا نتوقع أن يكون عدد الفروق الموجبة مساوياً لعدد الفروق السالبة، ويمكن صياغة فرض عدم على الشكل: $H_0 : P = 0.5$ أي أن نسبة الإشارات الموجبة (أو الإشارات السالبة) في المجتمع هي 0.50. أو صياغة فرض عدم على النحو: $H_0 : M_e = M_0$ مقابل الفرضية البديلة $H_1 : M_e \neq M_0$ ، حيث M_0 هي قيمة الوسيط الفعلية للظاهرة المدروسة مقترضين أن ٥٠ % من القياسات أقل منها و ٥٠ % أكبر منها .

- الدلالة الإحصائية^٥ لاختبار الإشارة عندما يكون عدد الثنائيات $N \leq 30$:

⁴ إن الفروق الصفرية بالنسبة للثنائيات المتماثلة يتم إهمالها وبذلك فإن حجم العينة سيكون مساوياً لجميع الأزواج مطروحاً منها عدد الأزواج التي لها فروق صفرية.

^٥ إن الاستدلال يعتمد على توزيع إحصائية الاختبار : $T_s = \sum_{i=1}^N I(D_i > 0)$ وقيمتها المشاهدة

مع العلم أن t_s هي عدد الفروقات الموجبة من أصل الفروقات الإجمالية N . يتم حساب

قيمة T_s عن طريق تخصيص الفروقات الإجمالية بإشارات موجبة وإشارات سالبة بكل الطرق الممكنة والتي

عددها يساوي إلى 2^N ، حيث التخصيصات لها نفس الاحتمال وهو 2^{-N} :

$\Gamma = \{(\text{sgn}(D_1), \text{sgn}(D_2), \dots, \text{sgn}(D_N)) : \text{sgn}(D_i) = 1 \text{ or } -1 \text{ } i = 1, 2, \dots, N\}$

Test Statistics

	Type
Chi-Square ^a	105.905
df	4
Asymp. Sig.	.000

- a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 50.0.

٦-٨ الاستدلال الالامعلمي لبيانات عينتين غير مستقلتين لبيانات رتبية:

Two Related Samples Inference

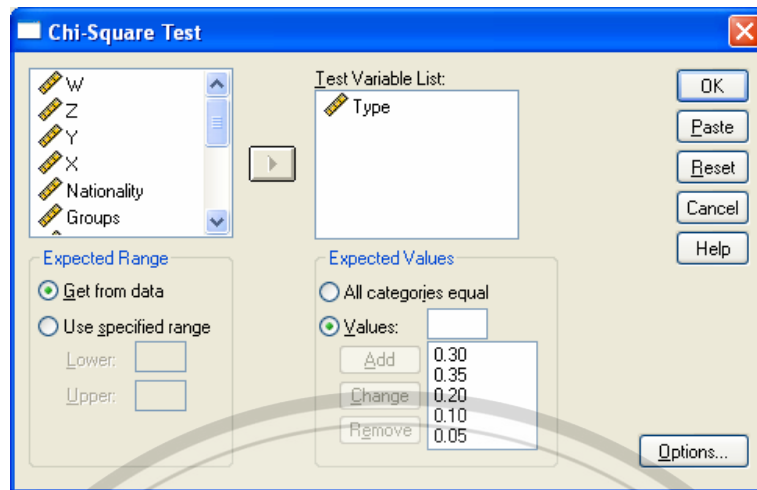
إن الاختبارات التي سنعالجها تتمتع بخصائص مشتركة وهي أنها قابلة للتطبيق على مجموعات من البيانات التي تضم ثنائيات مرتبطة ببعضها (غير مستقلة)، وخاصة عندما تؤخذ المشاهدات على شكل ثنائيات متناظرة من البيانات مثل نتائج اختبار أولي (سابق) pre test ونتائج اختبار تالي (لاحق) post test على العينة نفسها، أو كما هو الحال في الثنائيات المتطابقة matched - pairs .

إن الهدف من اختبار بيانات عينتين مرتبطتين هو معرفة فيما إذا كان العنصر الأول من الثنائية يخضع لتوزيع احتمالي مختلف عن العنصر الآخر. كما أن اختيار نوع الاختبار يعتمد على نوعية البيانات موضع الدراسة هل هي مستمرة ؟ ثنائية binary ؟ أو تصنيفية categorical ؟.

١-٦ اختبار الإشارة Sign Test :

يعتبر اختبار الإشارة من الاختبارات الالامعلمية الهامة ويعتبر بديلاً لاختبار t للعينات غير المستقلة المرتبطة، ويتميز بالكشف عن اتجاه الفروق بين أزواج المشاهدات، إذ إننا في بعض الأحيان لا نمتلك معلومات فعلية عن الفروق بين الثنائيات وكل ما لدينا هو إشارات هذه الفروق، وبهذه الحالة لا يمكننا إعطاء رتب للفروق لنتمكن من حساب إحصائية ولكوكسن لإشارة لرتب، بل نستطيع أن نستخدم المعلومات المتاحة عن إشارات الفروق لتطبيق اختبار الإشارة الذي يستفاد منه للكشف عن اتجاهات الفروق بين عينتين مرتبطتين.

عندما يكون لدينا عينتين صغيرتين نستخدم عادة اختبار t لمعرفة ما إذا كان متوسطا المجتمعين متساويين، مفترضين أن توزيع المجتمعين طبيعي وتباينهما متساو وغير معلوم، وغالباً لا يتحقق أحد الشرطين أو كلاهما، ما يضطرنا لإعادة صياغة فرض العدم القائل بتساوي متوسطي المجتمعين أي : $H_0 : \mu_1 = \mu_2$ على الشكل : $H_0 : F_1(x) = F_2(x)$ أي اختبار أن



- ننقل المتغير Type إلى النافذة المعنونة Test Variable List

في خانة Expected Values هناك خياران :

١- All Categories equal : جميع الفئات لها نفس التكرار المتوقع أو النسبة المئوية لكل صفة أو فئة متساوية.

٢- Values : يتم تفعيل هذا الخيار عندما تكون التكرارات المتوقعة مختلفة وبالتالي تكون النسب المئوية للصفات مختلفة كما هو الحال في نسب الحالة التعليمية في مثالنا. ولهذا السبب يتم:

- إدخال النسب المختلفة في المربع Values .

- نقر الزر Add بالنسبة لكل نسبة .

- نضغط OK .

نحصل على النتائج التالية:

NPar Tests

Chi-Square Test-Frequencies

Type

	Observed N	Expected N	Residual
ابتدائي	200	300.0	-100.0
متوسط	310	350.0	-40.0
ثانوي	280	200.0	80.0
جامعي	120	100.0	20.0
دراسات عليا	90	50.0	40.0
Total	1000		

نقارن القيمة الأقل $\min(D^+, D^-)$ من عدد الإشارات الموجبة أو السالبة مع القيم الحرجة أو النظرية المستخرجة من جدول معد خصيصاً لاختبار الإشارة لعينتين غير مستقلتين وذلك من أجل مستويات معنوية مختلفة و أحجام مختلفة للعينتين. فإذا كانت قيمة إحصائية الاختبار أكبر من القيمة الجدولية فإن الفروق بين العينتين ليست ذات دلالة إحصائية، أما إذا كانت قيمة إحصائية الاختبار أصغر أو تساوي القيمة الجدولية فإن الفروق بين العينتين ستكون ذات دلالة إحصائية. كما يمكننا الحصول على القيمة الحرجة (من جداول توزيع ذو الحدين) أو عن طريق تطبيق قانون توزيع ذو الحدين للمتغير العشوائي D الذي يمثل العدد الأقل للإشارات السالبة أو الموجبة.

- الدلالة الإحصائية لاختبار الإشارة عندما يكون عدد الثنائيات $N > 30$:

في هذه الحالة يتم تحويل القيمة الأقل $\min(D^+, D^-)$ من عدد الإشارات لقيمة معيارية^٦ كما يلي :

$$Z = \frac{(D^+ - 1/2) - N/2}{\sqrt{N/4}}, \text{ if } (D^+ > N/2) \quad (1)$$

$$Z = \frac{(D^+ + 1/2) - N/2}{\sqrt{N/4}}, \text{ if } (D^+ < N/2) \quad (2)$$

حيث Z تتبع التوزيع الطبيعي المعياري بمتوسط صفر وانحراف معياري ١ .

مثال (٨):

يرغب الباحثون في معرفة فيما إذا كانت التعليمات الخاصة بالعناية الشخصية والتبرج التزيين personal care and grooming في مدرسة للمعوقين عقلياً ستؤدي إلى تحسين ظهور الفتيات المتخلفات عقلياً appearance of mentally retarded girls أو لا ؟ تم اختيار ١٥ فتاة عشوائياً وتم تدريبهن على تعليمات خاصة بالعناية الشخصية والتزيين بعد مرور أسبوعين على انتهاء الدورة، تم إجراء مقابلة لهن من قبل ممرضة متخصصة ومشرف اجتماعي وحصلت كل واحدة منهن على درجة على مظهرها العام كما في الجدول . يعتقد الباحثون بأن الدرجة المعطاة تعكس بأفضل الحالات مقياس ترتيبي an ordinal scale ، فهم

إن القيمة المتوقعة وتباين إحصائية الاختبار في حالة اختبار الإشارة هو: $E(T_s) = N/2$ ، $\sigma^2(T_s) = N/4$ ويمكن أن تستبدل إحصائية الاختبار بالقيمة المعيارية :

$$Z = \frac{T - E(T)}{\sigma(T)}$$
التي تتبع التوزيع الطبيعي المعياري

^٦Wayne W . Daniel “ Biostatistics: A Foundation for Analysis in the Health Sciences” John Wiley & Sons. Inc 1991,pp585.

- نُفعل الخيار Sign ، ثم نضغط OK فنحصل على النتائج التالية :

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles		
						25th	50th (Median)	75th
Median	10	5.0000	.00000	5.00	5.00	5.0000	5.0000	5.0000
X_Degrees	10	6.9000	1.85293	4.00	10.00	5.7500	6.5000	8.2500

Sign Test

Frequencies

	N
X_Degrees - Median Negative Differences ^a	1
Positive Differences ^b	8
Ties ^c	1
Total	10

- a. X_Degrees < Median
- b. X_Degrees > Median
- c. X_Degrees = Median

Test Statistics^b

	X_Degrees - Median
Exact Sig. (2-tailed)	.039 ^a

- a. Binomial distribution used.
- b. Sign Test

- سنحاول أن نختبر فرضية العدم بتطبيق قانون توزيع ثنائي الحدين الذي يمكننا من تحديد احتمال الحصول على x إشارة سالبة أو أقل في عينة بحجم n ومعلمة p بموجب العلاقة التالية:

$$P(k \leq x/n, p) = \sum_{k=0}^n C_n^k p^k q^{n-k} \quad (9)$$

وفي حالتنا نحسب : $C_9^0 (0.5)^0 (0.5)^{9-0} + C_9^1 (0.5)^1 (0.5)^{9-1} = 0.0195$

وكان من الممكن استخراج القيمة مباشرة من الجدول: $P(k \leq 1/9, 0.5) = 0.0195$

نستنتج مما تقدم بأن قيمة الوسيط لا تساوي ٥ ، وأن مستوى المعنوية (أو الاحتمال الدقيق-p

value) لهذا الاختبار هو $2(0.0195) = 0.0390$

الاختبار أصغر أو تساوي القيمة الجدولية عند مستوى معنوية ٠,٠٥ فان الفروق بين العينتين^٧ ستكون ذات دلالة إحصائية أي أن الدورة التدريبية أعطت نتائج إيجابية ، أما عند مستوى معنوية 0.01 فان الفروق ليست ذات دلالة إحصائية .

وباستخدام برنامج SPSS نجد أن الفروق بين العينتين ذات دلالة إحصائية عند مستوى معنوية 0.039.

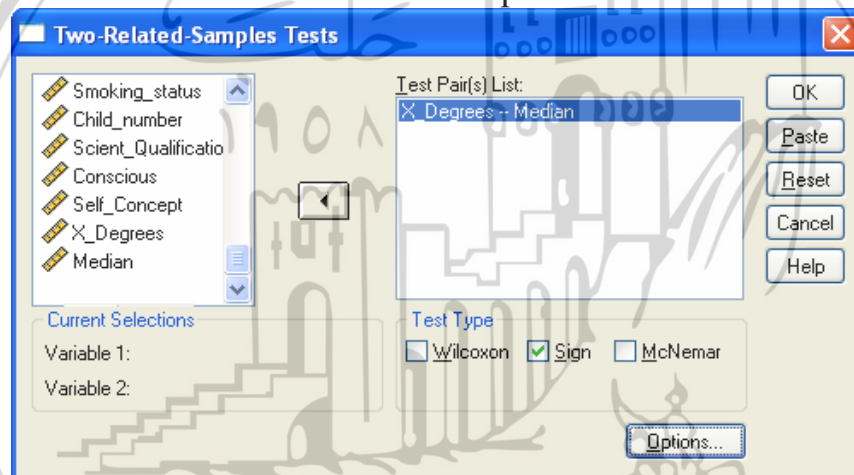
الحل :

من شريط القوائم نختار :

Analysis→ Non Parametric→ 2 Related Samples

يظهر صندوق حوار بعنوان :

Two-Related-Samples-Tests

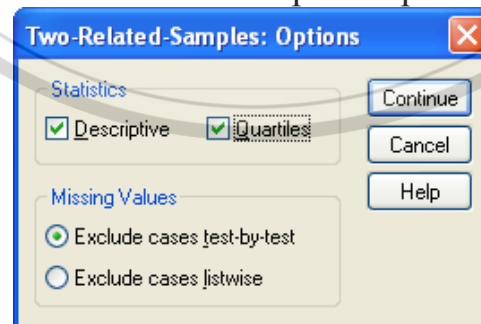


- ننقل المتغيرات الواجب مقارنة الفروق بينها وهي :

X_Degrees , Median إلى النافذة المعنونة Test Variable List

- نضغط على المربع Options نحصل على صندوق حوار بعنوان:

Two-Related-Samples: Options



- نُفعل الخيارين Descriptive , Quartiles ، ثم نضغط Continue فنعود إلى صندوق الحوار السابق .

^٧ مفترضين أن أفراد العينة الأولى قد حصلوا على قيمة تساوي ٥ وهي القيمة المفترضة للوسيط .

يشعرون بأن الدرجة ٨ تمثل مظهراً أفضل من الدرجة ٦ إلا أنهم غير ميالين للقول إن الفرق بين الدرجتين ٨ و ٦ يساوي تماماً الفرق بين الدرجتين 8 و 10 أو أن الفرق بينهما يمثل ضعف الفرق بين الدرجتين 5 و 6 .

والمطلوب :

هل نستطيع أن نستنتج بأن وسيط درجات المجتمع الذي سحبت منه هذه العينة هو مختلف عن ٥ القيمة المفترضة للوسيط hypothesized median ؟

الدرجة مقارنة بالقيمة المفترضة للوسيط	الدرجة X	رقم الفتاة
-	4	1
0	5	2
+	8	3
+	8	4
+	9	5
+	6	6
+	10	7
+	7	8
+	6	9
+	6	10

الحل :

سنختبر فرضية العدم القائلة بأن وسيط المجتمع $H_0 : M_e = 5$ مقابل الفرضية البديلة $H_1 : M_e \neq 5$.

إن إحصائية الاختبار في اختبار الإشارة هي القيمة الأقل $\min(D^+, D^-)$ من عدد الإشارات الموجبة المشاهدة أو عدد الإشارات السالبة المشاهدة. لتحديد طبيعة إحصائية الاختبار نفحص بيانات الجدول لتحديد الدرجات الواقعة أسفل القيمة المفترضة للوسيط فنعطيهما إشارة سالب ، وكذلك الدرجات الواقعة أعلى القيمة المفترضة للوسيط فنعطيهما إشارة موجب ونضع النتائج في العمود الثالث من الجدول ، (إن الفروق الصفرية بالنسبة للثنائيات المتماثلة تم إهمالها وبذلك فإن حجم العينة سيكون مساوياً ٩ في مثالنا).

نقارن القيمة الأقل $\min(D^+ = 8, D^- = 1)$ وهي في مثالنا $D^- = 1$ مع القيم الحرجة لاختبار الإشارة والتي تساوي ١ عند مستوى معنوية ٠,٠٥ وحجم العينة يساوي ٩ ، وتساوي ٠ عند مستوى معنوية ٠,٠١ وحجم العينة يساوي ٩ ، وبما أن قيمة إحصائية

٢-٦ اختبار ولوكوكسن لإشارة الرتب^٨ Wilcoxon Signed- Rank

يعتبر اختبار ولوكوكسن من الاختبارات اللامعلمية الهامة، كما يعتبر بديلاً لاختبار t للعينات غير المستقلة المرتبطة، ويتميز بالكشف عن اتجاه الفروق بين أزواج المشاهدات وكذلك حجم هذه الفروق .

- إن البيانات اللازمة لاختبار الفروق بين عيتين غير مستقلتين هي : N ثنائية من المشاهدات : $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$ ، حيث X و Y متغيرين عشوائيين مرتبطين عادة من خلال الأزواج المتطابقة matched - pairs ، حيث نضع بيانات كل متغير في عمود مستقل (عادة يخصص الأول لبيانات الاختبار الأولي والثاني لبيانات الاختبار اللاحق).

- نعرف متغيراً عشوائياً جديداً يساوي الفرق بين المتغيرين السابقين : $D = X - Y$ ، لتكن : $d_i = x_i - y_i$ ، $i = 1, 2, \dots, N$ ممثلة للفروق المشاهدة بين المتغيرين والتي عددها N متغير .

- نحذف إشارات الفروق ونفترض أن $|d_i| > 0$ ، $\forall i$ ، إن توزيع المتغير العشوائي D_i يحقق الافتراضات التالية :

- إن توزيع كل الفروق D_i هو متناظر symmetric .
- إن الفروق D_i هي مستقلة تبادلياً mutually independent .
- إن الفروق D_i لها نفس الوسيط .
- نرتب تصاعدياً القيمة المطلقة لفروق الثنائيات المشاهدة كما يلي :

$$|d_{[1]}| \leq |d_{[2]}| \leq \dots \leq |d_{[N]}|$$

- نحسب الرتب المتوافقة مع الفروق (أو أنصاف الرتب في حالة البيانات المتكررة)

وبشكل محدد أكثر إذا كانت هناك m_j مشاهدة متكررة عند أصغر قيمة مطلقة ذات الترتيب j فإننا نخصصهم جميعاً بالترتبة التالية :

$$r_{[j]} = m_1 + m_2 + \dots + m_{j-1} + 1/2(m_j + 1)$$

^٨ إن القيمة المتوقعة وتباين إحصائية الاختبار في حالة اختبار ولوكوكسن لإشارة الرتب هو :

$$E(T_{SR}) = \sum_{i=1}^N r_i / 2 = N(N+1)/4, \sigma^2(T_{SR}) = \sum_{i=1}^N r_i^2 / 4 = \frac{N(N+1)(2N+1)}{24}$$

وبما أن قيمة إحصائية الاختبار المحسوبة $T_p = 5.5$ هي أكبر من القيمة الجدولية في

الحالتين فإننا نرفض الفرض الصفري ونستنتج بوجود فروق ذات دلالة إحصائية لصالح الدورة المكثفة للتقانات الحديثة .

ملاحظة :

للتعرف على الدلالة الإحصائية لاختبار ولكوكسن في حالة وجود أزواج متطابقة وبالتالي وجود فروق صفرية يمكننا استخدام إحصائية الاختبار المعيارية Z ، وفي مثالنا هناك زوج متماثل من المشاهدات لهذا نستخدم إحصائية الاختبار :

$$Z = \frac{4T - N(N+1)}{\sqrt{\frac{2N(N+1)(2N+1)}{3}}} = \frac{4(5.5) - 9(10)}{\sqrt{\frac{18(10)(19)}{3}}} = -2.014$$

إن قيمة إحصائية الاختبار Z المحسوبة هي أكبر من قيمة Z الحرجة عند مستوى معنوية 0.05 والتي تساوي -1.96 وبالتالي نرفض فرض العدم ونستنتج وجود فروق ذات دلالة إحصائية بين نتائج الاختبارين لصالح الدورة المكثفة للتقانات الحديثة عند مستوى معنوية 0.05 ، في حين نقبل فرض العدم عند مستوى معنوية 0.01 لأن قيمة Z الحرجة تساوي -2.58 .

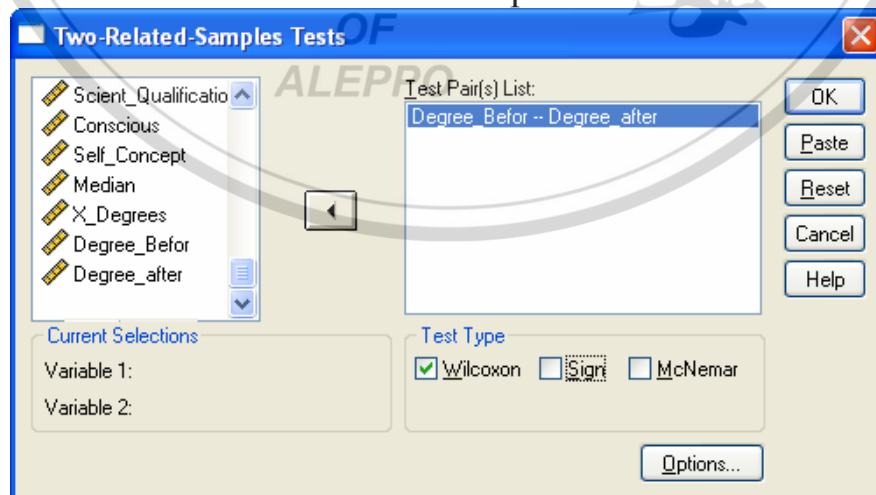
الحل :

من شريط القوائم نختار :

Analysis → Non Parametric → 2 Related Samples

يظهر صندوق حوار بعنوان :

Two-Related-Samples-Tests



- ننقل المتغيرات الواجب مقارنة الفروق بينها وهي :

Test Variable List Degree_befor , Degree_after إلى النافذة المعنونة

والمطلوب: التحقق من صحة الفرض الصفري القائل بعدم وجود فروق ذات دلالة إحصائية بين نتائج الاختبارين ؟

الحل :

بعد وضع نتائج كل اختبار في عمود مستقل الأول لبيانات الاختبار الأولي X والثاني لبيانات الاختبار اللاحق Y .

- تم إنشاء جدول الحسابات المساعد ومن خلاله حسبنا الفرق بين المتغيرين السابقين

$D = X - Y$ في العمود الأول من الجدول . تم حذف إشارات الفروق

$|d_i| > 0$, $\forall i$ في العمود الثاني من الجدول .

- تم ترتيب القيمة المطلقة لفروق الثنائيات المشاهدة تصاعدياً في العمود الثالث .

- تم حساب الرتب المتوافقة مع الفروق (أو أنصاف الرتب في حالة البيانات المتكررة) في العمود الرابع .

- أعدنا كتابة الرتب في عمود جديد خامس ووضعنا أمامها الإشارات التي تم حذفها.

إن مجموع الفروق الموجبة يساوي $T_p = 5.5$ ، في حين بلغ مجموع الفروق السالبة $T_M = 39.5$

، نأخذ منها الأقل فتكون هي قيمة إحصائية الاختبار المحسوبة وتساوي 5.5 .

الفروق بالاتجاه المطلق $D = X - Y$	الفرق بالقيمة المطلقة	ترتيب الفروق تصاعدياً	رتب الفروق	رتب الفروق مع الإشارة
-4	4	0	-	-
0	0	1	1.5	1.5
1	1	1	1.5	-1.5
-1	1	3	4	-4
-5	5	3	4	4
-7	7	3	4	-4
-3	3	4	6.5	-6.5
-4	4	4	6.5	-6.5
3	3	5	8	-8
-3	3	7	9	-9

نلاحظ أن حجم العينة يساوي 9 لأننا أهملنا الفروق الصفرية بالنسبة للثنائيات المتماثلة إذ كان

هناك زوج واحد متماثل ، إن القيمة الجدولية في اختبار ولوكوكسن من طرفين وعند مستوى

معنوية 0.05 تساوي 5 وعند مستوى معنوية 0.01 فإن القيمة الجدولية تساوي 2 .

- نعيد كتابة الرتب في عمود جديد ونضع أمامها الإشارات التي تم حذفها سابقاً ، ثم نجمع الفروق الموجبة ونرمز لها بـ T_p ، ثم نجمع الفروق السالبة ونرمز لها بـ T_M ونأخذ منها الأقل فتكون هي قيمة إحصائية الاختبار المحسوبة (مع العلم أن الفروق الصفرية بالنسبة للثنائيات المتماثلة يتم إهمالها وبذلك فإن حجم العينة سيكون مساوياً لجميع الأزواج مطروحاً منها عدد الأزواج التي لها فروق صفرية) .

- **الدلالة الإحصائية لاختبار ولكوكسن عندما يكون عدد الثنائيات $6 \leq N \leq 25$:**

نقارن قيمة إحصائية الاختبار المحسوبة سابقاً مع القيم الحرجة أو النظرية المستخرجة من جداول معدة خصيصاً لاختبار ولكوكسن لعينتين غير مستقلتين وذلك من أجل مستويات معنوية مختلفة و أحجام مختلفة للعينتين. فإذا كانت قيمة إحصائية الاختبار أكبر من القيمة الجدولية فإن الفروق بين العينتين ليست ذات دلالة إحصائية ، أما إذا كانت قيمة إحصائية الاختبار أصغر من القيمة الجدولية فإن الفروق بين العينتين ستكون ذات دلالة إحصائية .

- **الدلالة الإحصائية لاختبار ولكوكسن عندما يكون عدد الثنائيات $N > 25$:**

في هذه الحالة سنرمز لإحصائية الاختبار بـ T وهي القيمة الأقل لمجموع رتب الفروق ونستبدلها بإحصائية الاختبار المعيارية على النحو التالي :

$$Z = \frac{T - E(T)}{\sigma_T} = \frac{T - N(N+1)/4}{\sqrt{\frac{N(N+1)(2N+1)}{24}}} = \frac{4T - N(N+1)}{\sqrt{\frac{2N(N+1)(2N+1)}{3}}} \quad (10)$$

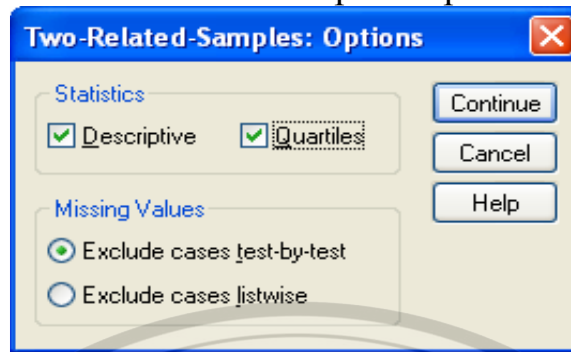
حيث Z تتبع التوزيع الطبيعي المعياري.

مثال (9):

تم تطبيق اختبار لدرجة الاستجابة لدى مجموعة من الأطفال قبل وبعد إتباعهم لدورة تقنية مكثفة في هذا المجال، تهدف الدورة إلى تعزيز درجة الاستجابة لدى الأطفال فكانت النتائج كما يلي:

الدرجة قبل الدورة X	الدرجة بعد الدورة Y
14	18
13	13
17	16
19	20
10	15
9	16
11	14
7	11
15	12
13	16

- نضغط على المربع Options نحصل على صندوق حوار بعنوان:
Two-Related-Samples: Options



- نُفعل الخيارين Descriptive , Quartiles ، ثم نضغط Continue فنعود إلى صندوق الحوار السابق .
- نُفعل الخيار Wilcoxon ، ثم نضغط OK فنحصل على النتائج التالية :

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles		
						25th	50th (Median)	75th
Degree_Bef	10	15.1000	2.72641	11.00	20.00	12.7500	15.5000	16.5000
Degree_afte	10	12.8000	3.67575	7.00	19.00	9.7500	13.0000	15.5000

Wilcoxon Signed Ranks Test

Ranks

		N	Mean Rank	Sum of Ranks
Degree_after -	Negative Ranks	7 ^a	5.64	39.50
Degree_Befor	Positive Ranks	2 ^b	2.75	5.50
	Ties	1 ^c		
	Total	10		

a. Degree_after < Degree_Befor

b. Degree_after > Degree_Befor

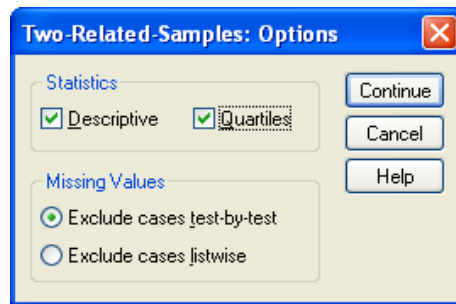
c. Degree_after = Degree_Befor

Test Statistics^b

	Degree_after - Degree_Befor
Z	-2.025 ^a
Asymp. Sig. (2-tailed)	.043

a. Based on positive ranks.

b. Wilcoxon Signed Ranks Test



- نُفعل الخيارين Descriptive , Quartiles .
- نضغط Continue فنعود إلى صندوق الحوار السابق .
- نُفعل الخيار Wilcoxon .
- نضغط OK .
- نحصل على النتائج التالية :

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles		
						25th	50th (Median)	75th
normal_rate	15	5.0500	.00000	5.05	5.05	5.0500	5.0500	5.0500
blood_rate	15	5.6413	1.40929	3.14	7.50	4.6400	5.8000	6.7400

Wilcoxon Signed Ranks Test

Ranks

	N	Mean Rank	Sum of Ranks
blood_rate - normal_rate Negative Ranks	5 ^a	6.80	34.00
Positive Ranks	10 ^b	8.60	86.00
Ties	0 ^c		
Total	15		

a. blood_rate < normal_rate

b. blood_rate > normal_rate

c. blood_rate = normal_rate

Test Statistics^b

	blood_rate - normal_rate
Z	-1.477 ^a
Asymp. Sig. (2-tailed)	.140

a. Based on negative ranks.

b. Wilcoxon Signed Ranks Test

§-٧ الاستدلال الالمعلمي لبيانات عينتين غير مستقلتين لبيانات اسمية:

في جدول الحسابات المساعد حسبنا الفرق بين المتغيرين السابقين $d_i = x_i - 5.05$ في العمود الأول من الجدول ، حذفنا إشارات الفروق $\forall i, |d_i| > 0$ في العمود الثاني من الجدول ، تم ترتيب القيمة المطلقة لفروق الثنائيات المشاهدة تصاعدياً في العمود الثالث ، وحسبنا الرتب المتوافقة مع الفروق في العمود الرابع من الجدول ، ثم أعدنا كتابة الرتب في عمود جديد خامس ووضعنا أمامها الإشارات التي تم حذفها سابقاً .

- إن مجموع الفروق الموجبة يساوي $T_p = 86$ ، في حين بلغ مجموع الفروق السالبة $T_M = 34$ ، نأخذ منها الأقل فتكون هي قيمة إحصائية الاختبار المحسوبة و تساوي 34 .
نحسب إحصائية الاختبار :

$$Z = \frac{4T - N(N+1)}{\sqrt{\frac{2N(N+1)(2N+1)}{3}}} = \frac{4(34) - 15(16)}{\sqrt{\frac{30(16)(31)}{3}}} = \frac{-104}{70.42727} = -1.4767$$

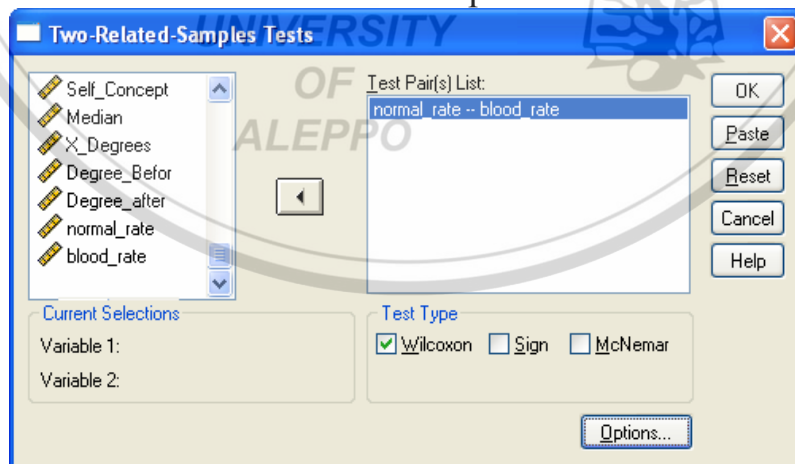
إن قيمة إحصائية الاختبار Z المحسوبة هي أصغر من قيمة Z الحرجة عند مستوى معنوية ٠,٠٥ والتي تساوي -١,٩٦ وبالتالي نقبل فرض العدم ونستنتج أن متوسط المجتمع الإحصائي يمكن أن يكون ٥,٠٥ لتر دم / دقيقة .
الحل :

من شريط القوائم نختار :

Analysis → Non Parametric → 2 Related Samples

يظهر صندوق حوار بعنوان :

Two-Related-Samples-Tests



- ننقل المتغيرات الواجب مقارنة الفروق بينها وهي :

Test Variable List Normal_rate , Blood_rate إلى النافذة المعنونة

- نضغط على المربع Options نحصل على صندوق حوار بعنوان:

Two-Related-Samples: Options

مثال (١٠) :

تم قياس معدل ما يضخه القلب من الدم في الوضع الوحشي الأيسر (لتر دم / دقيقة) بوساطة التمديد الحراري وذلك لعينة تضم ١٥ مريضاً خضعوا لعمليات جراحية للقلب فكانت النتائج كما يلي:

المعدل الطبيعي X	معدل ما يضخه القلب من الدم Y
5.05	4.91
5.05	4.1
5.05	6.74
5.05	7.27
5.05	7.42
5.05	7.5
5.05	6.56
5.05	4.64
5.05	5.98
5.05	3.14
5.05	3.23
5.05	5.8
5.05	6.17
5.05	5.39
5.05	5.77

والمطلوب :

هل يمكننا أن نستنتج بأن متوسط المجتمع الإحصائي هو مختلف عن 5.05 لتر/دقيقة؟

رتب الفروق مع الإشارة	رتب الفروق	ترتيب الفروق تصاعدياً	الفروق بالقيمة المطلقة	الفروق $d_i = x_i - 5.05$
-1	1	0.14	0.14	-0.14
2	2	0.34	0.95	-0.95
-3	3	0.41	1.69	1.69
4	4	0.72	2.22	2.22
5	5	0.75	2.37	2.37
6	6	0.93	2.45	2.45
-7	7	0.95	1.51	1.51
8	8	1.12	0.41	-0.41
9	9	1.51	0.93	0.93
10	10	1.69	1.91	-1.91
-11	11	1.82	1.82	-1.82
-12	12	1.91	0.75	0.75
13	13	2.22	1.12	1.12
14	14	2.37	0.34	0.34
15	15	2.45	0.72	0.72

الحل :

١-٧ اختبار ماكنمار McNemar's Test

هو اختبار لا معلمي لاختبار متغيرين أو عینتین اسمیتین مرتبطين (ومن الممكن أن تكون بيانات نفس العينة قبل إجراء التجربة وبعدها) ، يستخدم اختبار ماكنمار McNemar's Test لاختبار المساواة بين معدلات الاستجابة الثنائية من مجتمعين حيث البيانات عبارة عن ثنائيات أو أزواج متطابقة من عینتین مرتبطين أو أنها تخص عينة واحدة قبل إخضاعها لمعالجة معينة (إجراء التجربة) وبعدها، وبالتالي فإن الاختبار سيحدد فيما إذا كان معدل الاستجابة الأولية يتساوى مع معدل الاستجابة اللاحقة أي التحقق من عدم وجود تغير ذي دلالة إحصائية بين العینتین موضع المقارنة، لإجراء الاختبار يتطلب الأمر تنظيم البيانات في جدول يحوي سطرین وعمودین فقط:

الاختبار الأولي السابق PRE-TEST	الاختبار اللاحق POST-TEST	
	- (0)	+ (1)
- (0)	y_{11}	y_{12}
+ (1)	y_{21}	y_{22}

مفترضین أن :

y_{11} تمثل عدد الأشخاص الذين استجابتهم الأولية واللاحقة سالبة ، y_{22} تمثل عدد الأشخاص الذين استجابتهم الأولية واللاحقة موجبة ، y_{12} تمثل عدد الأشخاص الذين استجابتهم الأولية سالبة واستجابتهم اللاحقة موجبة ، y_{21} تمثل عدد الأشخاص الذين استجابتهم الأولية موجبة واستجابتهم اللاحقة سالبة .

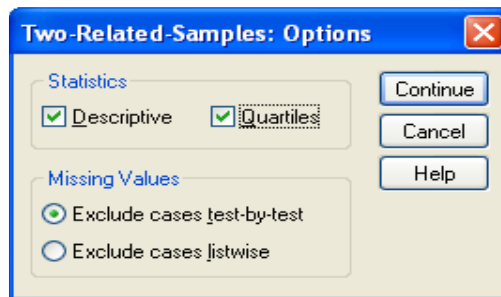
- الدلالة الإحصائية لاختبار ماكنمار في حالة $10 \leq y_{12} + y_{21} \leq 20$:

يتم حساب قيمة إحصائية الاختبار من جدول الاقتران السابق بموجب العلاقة التالية :

$$\chi^2 = \frac{[|y_{21} - y_{12}| - 1]^2}{y_{21} + y_{12}} \quad (11)$$

التي تتبع توزيع كاي تربيع مع درجات حرية تساوي الواحد دوماً .

تقارن قيمة إحصائية الاختبار المحسوبة سابقاً مع القيم الحرجة أو النظرية المستخرجة من جدول كاي تربيع عند درجة حرية واحدة ، وذلك من أجل مستويات معنوية مختلفة . إذا كانت قيمة إحصائية الاختبار أكبر من القيمة الجدولية فإن الفروق بين العینتین أو الاستجابتين الأولية واللاحقة ليست ذات دلالة إحصائية ، أما إذا كانت قيمة إحصائية الاختبار أصغر من القيمة الجدولية فإن الفروق بين العینتین أو الاستجابتين الأولية واللاحقة ستكون ذات دلالة إحصائية .



- نُفعل الخيارين Descriptive , Quartiles ، ثم نضغط Continue فنعود إلى صندوق الحوار السابق .

- نُفعل الخيار McNemar ، ثم نضغط OK فنحصل على النتائج التالية :

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles		
						25th	50th (Median)	75th
response_bef	19	.5263	.51299	.00	1.00	.0000	1.0000	1.0000
response_afte	19	.3684	.49559	.00	1.00	.0000	.0000	1.0000

McNemar Test - Crosstabs

response_befor & response_after

response_befor	response_after	
	0	1
0	7	2
1	5	5

Test Statistics^b

	response_befor & response_ after
N	19
Exact Sig. (2-tailed)	.453 ^a

a. Binomial distribution used.

b. McNemar Test

٢-٧ الاستدلال حول عدد من المجموعات المستقلة لبيانات اسمية (اختبار الوسيط)

للمقارنة بين عدد من العينات المستقلة نستخدم اختبار الوسيط للتحقق من دلالة الفروق بين وسيط عدد من العينات المستقلة، الذي يعتمد على إيجاد قيمة الوسيط للعينات موضع

المقارنة أولاً، وهي عملية لا معلمية لاختبار فرضية العدم $H_0: F_1 = F_2 = \dots = F_k$

- (0)	7	2
+ (1)	5	5

ونظراً لكون $10 \leq y_{12} + y_{21} \leq 20$ نقوم بحساب قيمة إحصائية الاختبار من جدول الاقتران كما يلي :

$$\chi^2 = \frac{[|y_{21} - y_{12}| - 1]^2}{y_{21} + y_{12}} = \frac{[|5 - 2| - 1]^2}{5 + 2} = \frac{4}{7} = 0.571429$$

وبمقارنة قيمة إحصائية الاختبار المحسوبة مع القيمة الحرجة المستخرجة من جدول كاي تربيع عند مستوى معنوية ٠,٠٥ ودرجة حرية واحدة والتي تساوي ٦,٦٤ فنجدها أقل بكثير وبناء عليه نستنتج عدم وجود فروق ذات دلالة إحصائية بين اتجاهات الشباب قبل مشاهدة الفيلم واتجاهاتهم بعد عرض الفيلم أي نقبل الفرض الصفري. أي أن الفيلم ليس له تأثير على تغير استجابات الشباب نحو مهنة رجل الإطفاء .

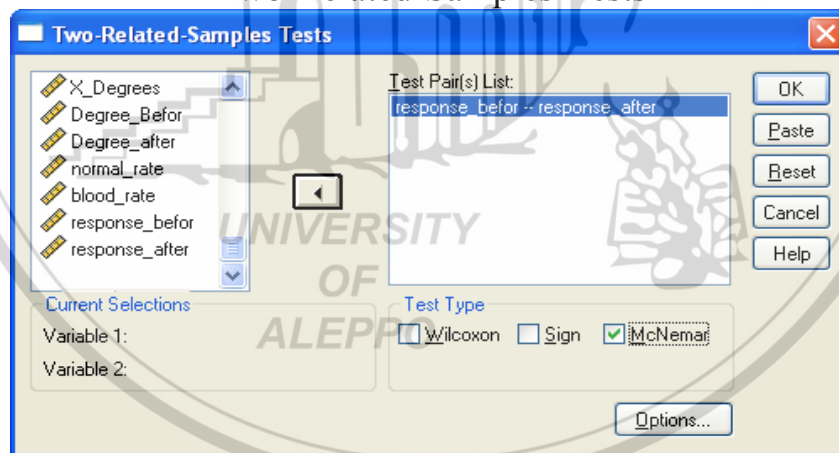
الحل :

من شريط القوائم نختار :

Analysis → Non Parametric → 2 Related Samples

يظهر صندوق حوار بعنوان :

Two-Related-Samples-Tests



- ننقل المتغيرات الواجب مقارنة الفروق بينها وهي :

Test Variable List Response_befor , response_after إلى النافذة المعنونة

- نضغط على المربع Options نحصل على صندوق حوار بعنوان:

Two-Related-Samples: Options

جدير بالتنويه أن اختبار ماكنمار لا يصلح في حالة كون $y_{12} + y_{21} < 20$ ويستخدم بدلاً عنه اختبار ذي الحدين .

- الدلالة الإحصائية لاختبار ماكنمار في حالة $y_{12} + y_{21} > 20$:

في هذه الحالة يتم حساب قيمة إحصائية الاختبار من العلاقة التالية :

$$\chi^2 = \frac{(y_{21} - y_{12})^2}{y_{21} + y_{12}} \quad (12)$$

التي تتبع توزيع كاي تربيع مع درجات حرية تساوي الواحد دوماً .

مثال (١١)

في دراسة للاتجاهات نحو مهنة رجل الإطفاء، تم عرض فيلم سينمائي (يوضح مخاطر ومزايا المهنة) أمام مجموعة من الشبان الحاصلين على الثانوية العامة نحو مهنة رجل الإطفاء، فإذا تم طرح سؤال للاتجاهات قبل الفيلم وبعده . والمطلوب : التحقق من صحة الفرض القائل " لا تأثير للفيلم على اتجاهات الحاصلين على الثانوية نحو هذه المهنة " وذلك عند مستوى معنوية 0.05، فكانت النتائج كما يوضحه الجدول التالي :

الاستجابة الأولية قبل الفيلم		الاستجابة اللاحقة بعد مشاهدة الفيلم	
نعم	1	نعم	1
نعم	1	لا	0
نعم	1	لا	0
لا	0	لا	0
لا	0	لا	0
لا	0	لا	0
نعم	1	نعم	1
نعم	1	نعم	1
لا	0	لا	0
لا	0	لا	0
نعم	1	لا	0
نعم	1	لا	0
نعم	1	نعم	1
لا	0	نعم	1
لا	0	لا	0
لا	0	نعم	1
نعم	1	نعم	1
نعم	1	لا	0
لا	0	لا	0

لإجراء الاختبار نقوم بتنظيم البيانات في جدول يحوي سطرين وعمودين:

الاختبار الأولي السابق PRE-TEST	الاختبار اللاحق POST-TEST	
	(0) -	(1) +

مقابل الفرضية البديلة H_1 القائلة بأن هناك على الأقل زوج واحد (j_1, j_2) فقط غير متساويان $F_{j_1} \neq F_{j_2}$. يستخدم اختبار الوسيط عندما يطلب منا عدم وضع افتراضات حول التوزيعات المختبرة $F_j, j=1,2,\dots,k$ ومع ذلك إذا كنا نمتلك معلومات إضافية حول هذه التوزيعات، كأن نعتقد بأن لها نفس الشكل ولكنها تختلف عن بعضها بعضاً بنقل موقع المؤشرات في حال تحقق الفرضية البديلة.

سنحاول الآن تعريف إحصائية الاختبار بالنسبة لاختبار الوسيط، مع العلم أننا نعلم بأن أول خطوة في الاختبار هي تحويل البيانات الخام الأصلية إلى قياسات متوافقة معها. ولكي نحدد هذه القياسات نحسب أولاً الوسيط M_e لبيانات العينة المجمعة ذات الحجم N (حيث يتم دمج العينات موضع المقارنة وكأنها درجات عينة واحدة). للحصول على الوسيط نرتب البيانات u_{ij}

تصاعدياً، ومن ثم نحسب الوسيط بوحدة من الطرق التالية:

- الوسيط هو العدد الذي ترتيبه يساوي إلى $\frac{N+1}{2}$ في حال كون N عدداً فردياً .

- الوسيط هو عبارة عن الوسط الحسابي للعددين الذين ترتيبهما $\frac{N}{2}$ و $\frac{N}{2}+1$ ، وذلك في

حال كون N عدداً زوجياً .

- الوسيط من بيانات مبوبة يتطلب تكوين جدول التوزيع التكراري المتجمع الصاعد، حساب

ترتيب الوسيط من العلاقة $\frac{\sum f}{2} = \frac{n}{2}$ ، تحديد الفئة الوسيطة من جدول التوزيع التكراري

المتجمع الصاعد، وأخيراً نحسب قيمة الوسيط بموجب العلاقة:

$$M_e = L_i + \frac{\frac{n}{2} - f_1}{f_2 - f_1} \cdot w$$

حيث أن :

L_i الحد الأدنى للفئة الوسيطة

f_1 التكرار المتجمع المناظر للحد الأدنى للفئة الوسيطة .

f_2 التكرار المتجمع المناظر للحد الأعلى للفئة الوسيطة .

15	11	14
6	5	11
5	13	15
4	10	14
8	9	13
7	13	-

- إن أول خطوة في اختبار الوسيط هي تحويل البيانات الخام الأصلية إلى قياسات متوافقة معها، ولكي نحدد هذه القياسات نحسب أولاً الوسيط M_e لبيانات العينة المجمعة ذات الحجم N أي يتم دمج المجموعات الثلاث موضع المقارنة وكأنها درجات مجموعة (عينة) واحدة^١.

- للحصول على الوسيط نرتب البيانات u_{ij} تصاعدياً أولاً، ومن ثم سيكون الوسيط هو العدد

الذي ترتيبه $\frac{N+1}{2} = \frac{23+1}{2} = 12$ بسبب كون حجم العينة المجمعة عدداً فردياً أي:

$$N = 8 + 8 + 7 = 23 ، وبالتالي فإن وسيط الدرجات يساوي $M_e = 10$.$$

- تصنف القياسات w_{ij} المتوافقة مع الملاحظة الخام u_{ij} لكل عينة كما كانت قبل الدمج في ضوء قيمة الوسيط بحيث نضع عوضاً عن قيمة الملاحظة الرقم ١ إذا كانت قيمتها أقل أو تساوي قيمة الوسيط ، ونضع عوضاً عن قيمة الملاحظة الرقم ٠ إذا كانت قيمتها أكبر من قيمة الوسيط ، وذلك كما في الجدول التالي :

(جامعيين)		(حملة الثانوية)		(حملة الإعدادية)	
u_{ij}	w_{ij}	u_{ij}	w_{ij}	u_{ij}	w_{ij}
9	1	10	1	10	1
11	0	7	1	11	0
15	0	11	0	15	0
6	1	5	1	6	1
5	1	13	0	5	1
4	1	10	1	4	1
8	1	9	1	8	1
7	1	13	0	7	1

- نحدد عدد (التكرارات) المشاهدات التي تكون قيمها أقل من أو تساوي قيمة الوسيط.
 - وكذلك عدد (التكرارات) المشاهدات التي تكون قيمها أكبر من قيمة الوسيط في كل عينة من العينات موضع المقارنة قبل الدمج .

^{١٠} إن اختبار الوسيط لا يشترط تساوي حجم العينات موضع المقارنة حيث أن حجم العينة الثالثة هو ٧ في حين حجم العينتين الأولى والثانية هو ٨ .

$$E_{11} = \frac{n_1 \cdot m}{N}, \quad E_{22} = \frac{n_2 \cdot (N - m)}{N}, \quad E_{2k} = \frac{n_k \cdot (N - m)}{N}$$

إحصائية الاختبار T لاختبار الوسيط تحسب بموجب العلاقة التالية:

$$T = \sum_{j=1}^k \frac{(\hat{w} - n_j m / N)^2}{n_j m / N} + \sum_{j=1}^k \frac{(n_j - \hat{w} - n_j (N - m) / N)^2}{n_j (N - m) / N}$$

وتتبع توزيع كاي تربيع بدرجات حرية تساوي $(k - 1)(r - 1) = (k - 1)(2 - 1) = k - 1$ حيث k عدد العينات و r عدد الأسطر .
لذلك إذا كانت t هي القيمة الفعلية المشاهدة لـ T ، فإن القيمة الاحتمالية الدقيقة لـ p من أجل اختبار من طرفين للوسيط تعطى بموجب العلاقة التالية:

$$p_2 = \sum_{T \geq t} h(\hat{w})$$

حيث المجموع يشمل كل قيم $\hat{w} \in W$ والتي من أجلها : $T(\hat{w}) \leq t$.
- إن إحصائية الاختبار السابقة هي تماماً إحصائية كاي تربيع ليبرسون التي تحسب بموجب الصيغة التالية :

$$X^2 = \sum_{i=1}^2 \sum_{j=1}^K \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

O_{ij} : تمثل التكرار المشاهد في الخلية الموجودة في الصف i والعمود k .

مثال (١٢)⁹:

تم تطبيق اختبار على ثلاث مجموعات الأولى من الحاصلين على الإعدادية والمجموعة الثانية من الحاصلين على الشهادة الثانوية، والمجموعة الثالثة من الحاصلين على الشهادة الجامعية، بهدف التحقق من مدى استيعابهم واستفادتهم من مضمون الدورة التدريبية التي تقيمها هيئة مكافحة البطالة في إدارة المشروعات الصغيرة للعاطلين عن العمل بهدف منحهم القروض اللازمة لإقامة المشروعات. والمطلوب التحقق من صحة الفرض القائل " لا توجد فروق ذات دلالة إحصائية في مستوى الاستيعاب والاستفادة بين المجموعات الثلاث حملة الإعدادية، حملة الثانوية ، الجامعيين " .

المجموعة الأولى (حملة الإعدادية)	المجموعة الثانية (حملة الثانوية)	المجموعة الثالثة (جامعيين)
9	10	8
11	7	11

⁹ د . زكريا الشربيني " الإحصاء اللابارامتري في العلوم النفسية والتربوية والاجتماعية " مكتبة الانجلو المصرية ، القاهرة ١٩٩٠ م .

w طول الفئة الوسيطة .

- إن القياسات w_{ij} المتوافقة مع الملاحظة الخام u_{ij} يمكن تعريفها كما يلي :

$$w_{ij} = \begin{cases} 1 & \text{if } u_{ij} \leq M_e \\ 0 & \text{if } u_{ij} > M_e \end{cases}$$

أي تصنف مشاهدات كل عينة كما كانت قبل الدمج في ضوء قيمة الوسيط إذ نضع عوضاً عن قيمة الملاحظة:

الرقم 1 إذا كانت قيمتها أقل أو تساوي قيمة الوسيط، ونضع عوضاً عن قيمة الملاحظة الرقم 0 إذا كانت قيمتها أكبر من قيمة الوسيط.

- نحدد عدد (تكرارات) المشاهدات التي تكون قيمها أقل أو تساوي قيمة الوسيط وكذلك عدد (تكرارات) المشاهدات التي تكون قيمها أكبر من قيمة الوسيط في كل عينة من العينات موضع المقارنة قبل الدمج وتصنف التكرارات في جدول مفترضين أن:

$$w_j = \sum_{i=1}^{n_j} w_{ij}$$

تمثل العدد الإجمالي لمشاهدات العينة z التي تساوي أو تقل عن الوسيط. $m = \sum_{j=1}^k w_j$ تمثل العدد الإجمالي لمشاهدات العينة المجمعة التي هي مساوية أو أقل من الوسيط.

إن إحصائية الاختبار للوسيط معرفة على جدول التوافق $2 \times k$ ، حيث يمثل:

- السطر الأول من هذا الجدول عدد الاستجابات (المشاهدات) في كل عينة التي تكون قيمتها مساوية أو أقل من قيمة الوسيط

- السطر الثاني عدد الاستجابات في كل عينة التي تكون قيمتها أكبر من قيمة الوسيط.

جدول رقم (٣)

بيانات جدول التوافق $2 \times k$ اللازمة لاختبار الوسيط

تصنيف المجموعات	العينات Samples				مجموع الأسطر
	١	٢		k	
أصغر أو تساوي الوسيط $\leq M_e$	w_1	w_2		w_k	m
أكبر من الوسيط $> M_e$	$n_1 - w_1$	$n_2 - w_2$		$n_k - w_k$	$N - m$
مجموع الأعمدة	n_1	n_2		n_k	N

- نحسب التكرار المتوقع لكل خلية باستخدام المجاميع الهامشية لكل من اسطر وأعمدة جدول التوافق والمجموع الكلي للمشاهدات وذلك على النحو التالي:

- تصنف التكرارات في جدول التوافق 2×3 كما يلي:
بيانات جدول التوافق 2×3 اللازمة لاختبار الوسيط

تصنيف المجموعات	العينات Samples			مجموع الأسطر
	الأولى	الثانية	الثالثة	
أصغر أو تساوي الوسيط $M_e \leq$	٦	٥	١	١٢
أكبر من الوسيط $M_e >$	٢	٣	٦	١١
مجموع الأعمدة	٨	٨	٧	٢٣

- نحسب التكرار المتوقع لكل خلية باستخدام المجاميع الهامشية للأسطر وأعمدة جدول التوافق والمجموع الكلي للملاحظات وذلك على النحو التالي :

$$E_{11} = \frac{n_{1.} \cdot m_{.1}}{N} = \frac{8(12)}{23} = 4.173913 \quad , \quad E_{22} = \frac{n_{2.} \cdot (N - m_{.1})}{N} = \frac{8(11)}{23} = 3.826087$$

باتباع نفس الخطوات نحصل على جدول التكرارات المتوقعة كما يلي :

تصنيف المجموعات	العينات Samples			مجموع الأسطر
	الأولى	الثانية	الثالثة	
أصغر أو تساوي الوسيط $M_e \leq$	4.17	4.17	3.65	12
أكبر من الوسيط $M_e >$	3.83	3.83	3.35	11
مجموع الأعمدة	8	8	7	23

- إن إحصائية الاختبار T لاختبار الوسيط تحسب كما يلي:

$$T = \frac{(6 - 4.17)^2}{4.17} + \frac{(5 - 4.17)^2}{4.17} + \frac{(1 - 3.65)^2}{3.65} + \frac{(2 - 3.83)^2}{3.83} + \frac{(3 - 3.83)^2}{3.83} + \frac{(6 - 3.35)^2}{3.35} = 6.04$$

وتتبع توزيع كاي تربيع بدرجات حرية تساوي $k - 1 = 3 - 1 = 2$.

عند مستوى معنوية ٠,٠٥ نقارن قيمة إحصائية الاختبار المحسوبة مع قيمة كاي تربيع الدرجة عند ٢ درجة حرية والتي نجدها مساوية إلى ٥,٩٩ ، فنجد أن قيمة إحصائية الاختبار المحسوبة أكبر من قيمة كاي تربيع الدرجة لذلك نرفض فرض العدم القائل بعدم وجود فروق جوهرية ذات دلالة إحصائية في مستوى الاستيعاب والاستفادة بين المجموعات الثلاث حملة الإعدادية، حملة الثانوية ، والجامعيين ."

وعند مستوى معنوية ٠,٠١ نجد أن قيمة إحصائية الاختبار المحسوبة أصغر من قيمة كاي تربيع الدرجة ٩,٢١ لذلك نقبل فرض العدم القائل بعدم وجود فروق جوهرية ذات دلالة إحصائية في مستوى الاستيعاب والاستفادة بين المجموعات الثلاث حملة الإعدادية، حملة الثانوية ، والجامعيين ."

الفصل الثامن

كتابة الأوامر

Syntax Commands

يوفر برنامج SPSS للمستخدم معظم الأوامر التي تمكنه من تنفيذ التحليل الإحصائي المرغوب ، يمكن الحصول على هذه الاوامر من القوائم أو من صناديق الحوار ، باستثناء بعض الأوامر التي لا تتوفر إلا باستخدام لغة الاوامر Command Language التي تساعد المستخدم على تنفيذ أعمال وأنشطة غير متوفرة ضمن أوامر SPSS، تسمح هذه اللغة بتخزين الأعمال والأنشطة في ملف الأوامر Syntax File حيث يمكن طلبها بأي وقت وتنفيذها لأي مجموعة من المتغيرات.

§-١ ملف الأوامر Syntax File

- عبارة عن ملف نصوص Text File يحتوي على أوامر مكتوبة Commands بلغة برمجية بسيطة، حيث يجب مراعاة بعض القواعد في كتابة هذه الأوامر.
- يجب أن يبدأ كل أمر بسطر جديد وينتهي بنقطة.
 - تفصل الأوامر الفرعية Sub Command بواسطة العلامة سلاش (/)، مع العلم أن وضع العلامة (/) قبل أو أمر فرعي يكون اختياريًا.
 - يجب كتابة الاسماء الكاملة للمتغيرات.
 - يجب كتابة النص المحصور بين علامتي اقتباس Quotation Marks في سطر واحد.
 - يجب ألا يزيد طول السطر في ملف الأوامر عن ٨٠ رمزا.
 - تستخدم النقطة (.) للتعبير عن الفاصلة العشرية.

١-١ الطرق المساعدة في بناء ملفات الأوامر

إن كتابة الاوامر البرمجية تحتاج الى معرفة وممارسة في كتابة البرامج مما قد لا يتوفر لدى الكثيرين من مستخدمي برنامج SPSS، ولهذا السبب يمكن للمستخدم أن يستعين بأحد الوسائل الثلاث التالية التي تساعد في كتابة الأوامر البرمجية الخاصة ببرنامج SPSS، وهي:

- كتابة الأوامر من صناديق الحوار Dialog boxes.
- نسخ الأوامر من Log في مخرجات برنامج SPSS.
- نسخ الأوامر من Journal File .

Median Test

Frequencies

		Grouping		
		1.00	2.00	3.00
Academic	> Median	2	3	6
	<= Median	6	5	1

Test Statistics^b

	Academic
N	23
Median	10.0000
Chi-Square	6.039 ^a
df	2
Asymp. Sig.	.049

a. 6 cells (100.0%) have expected frequencies less than 5. The minimum expected cell frequency is 3.3.

b. Grouping Variable: Grouping

Mood Median Test

Chi-Square = 6.04 DF = 2 P = 0.049

Individual 95.0% CIs					
C2	N<=	N>	Median	Q3-Q1	
1	6	2	7.50	5.25	(-----+-----)
2	5	3	10.00	5.00	(-----+-----)
3	1	6	13.00	3.00	(-----+-----)
				6.0 9.0 12.0 15.0	

Overall median = 10.00

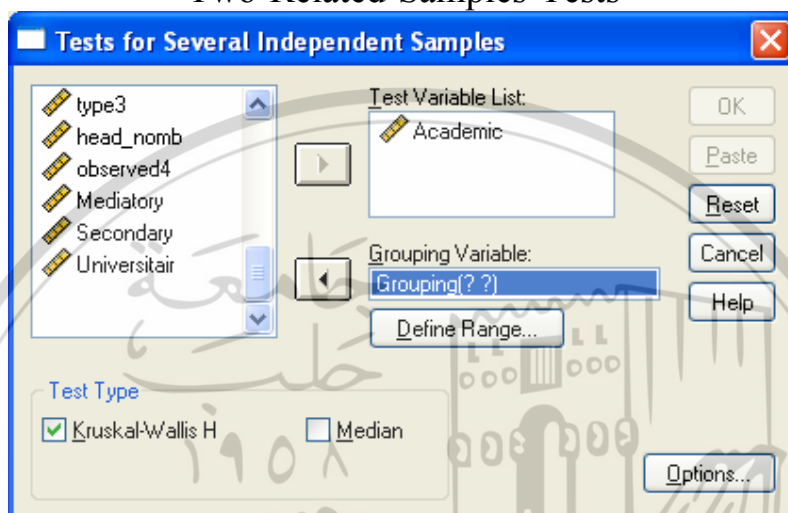
الحل :

من شريط القوائم نختار :

Analysis-→ Non Parametric-→ 2 Related Samples

يظهر صندوق حوار بعنوان :

Two-Related-Samples-Tests

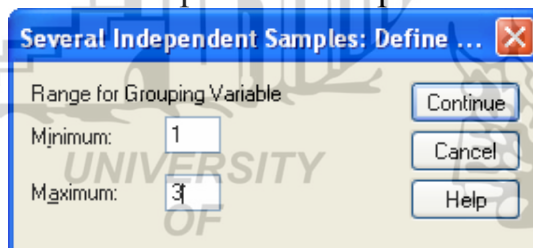


- ننقل المتغيرات الواجب مقارنة الفروق بينها وهي :

Academic إلى النافذة المعنونة Test Variable List

- نضغط على المربع Define Range نحصل على صندوق حوار بعنوان:

Several Independent Samples: Define



- نحدد القيمة الدنيا والقيمة العليا، ثم نضغط Continue فنعود إلى صندوق الحوار السابق .

- نُفعل الخيار Median ، ثم نضغط OK فنحصل على النتائج التالية :

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles		
						25th	50th (Median)	75th
Academic	23	9.9565	3.32317	4.00	15.00	7.0000	10.0000	13.0000
Grouping	24	2.0000	.83406	1.00	3.00	1.0000	2.0000	3.0000

٢-٨ كتابة الأوامر من صناديق الحوار Dialog boxes

تعتبر هذه الطريقة من أسهل الطرق لكتابة الأوامر، حيث يتم فتح صندوق الحوار لتنفيذ تحليل إحصائي معين، ثم استعمال الزر Paste في لصق الأوامر في ملف الأوامر Syntax File ، ونورد فيما يلي بعض الأمثلة:

مثال (١):

ليكن لدينا المتغيران Salary و Age الموضحان في الجدول التالي الذي يشتمل على عمر ودخل ٢٢ عضو هيئة تدريسية، لحساب تكرار قيم المتغير المدروس مع حساب بعض المقاييس الإحصائية نتبع ما يلي:

Respd_num	Salary	Age
1	12000	35
2	11000	40
3	12500	44
...		
...		
22	16000	50

إدخال المتغيرين في Salary و Age في شاشة محرر البيانات Data Editor ونرغب في تشكيل ملف أوامر بطريقة الكتابة من صندوق الحوار يشتمل على التحاليل الإحصائية التالية:

- إنشاء جدول التوزيع التكراري للدخل والعمر.
 - إنشاء جدول التوزيع التكراري للدخل فقط مع حساب المقاييس الإحصائية: الوسط الحسابي، الوسيط، الانحراف المعياري، التباين ، المجموع.
 - إنشاء جدول التوزيع التكراري للدخل مع رسم المدرج التكراري والمنحنى الطبيعي.
 - إنشاء جدول التوزيع التكراري للدخل مع حساب الربيعات والمئين الخامس
- العاشر الخامس عشر والمئين الأربعين.

ملاحظة:

في حال كان المطلوب حساب الربيعات والمئينات لمتغير الدخل بدون إظهار جدول التوزيع التكراري، فإن الأوامر البرمجية ستأخذ الشكل التالي:

نفس الصيغة الأخيرة الموجودة في الجدول باستثناء إضافة العبارة:

/FORMAT=NOTABLE

لصيغة البرمجية.

EXAMINE

VARIABLES=Salary BY sex

/ PLOT BOXPLOT STEMLEAF

/ COMPARE GROUP

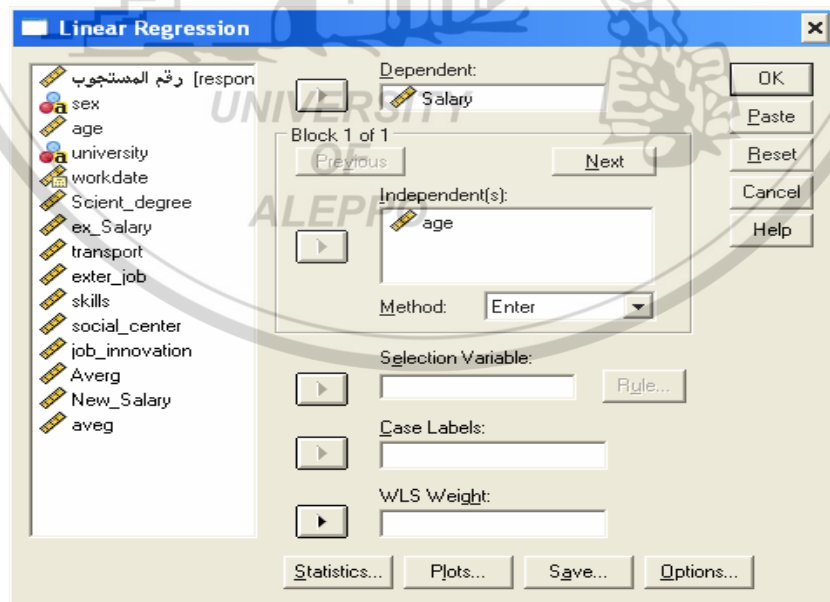
/ STATISTICS DESCRIPTIVES

/ CINTERVAL 95

/ MISSING LISTWISE

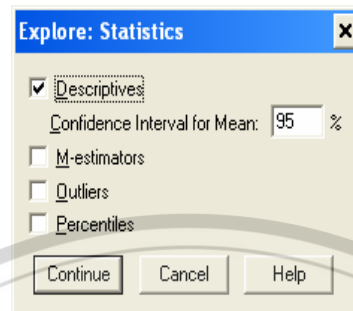
/ NOTOTAL.

- مثال (٢): من بيانات التطبيق السابق نرغب تشكيل ملف أوامر بطريقة الكتابة من صندوق الحوار يشتمل على:
- استخراج معاملات Coefficients انحدار المتغير التابع الدخل Salary على المتغير المستقل الجنس Sex بالطريقة الاعتيادية Enter.
 - استخراج جدول تحليل التباين ANOVA.
 - استخراج إحصائية الاختبار لدوربين واتسون Durbin Watson .
- من شريط القوائم نختار Linear → Regression → Analysis يظهر صندوق الحوار التالي:

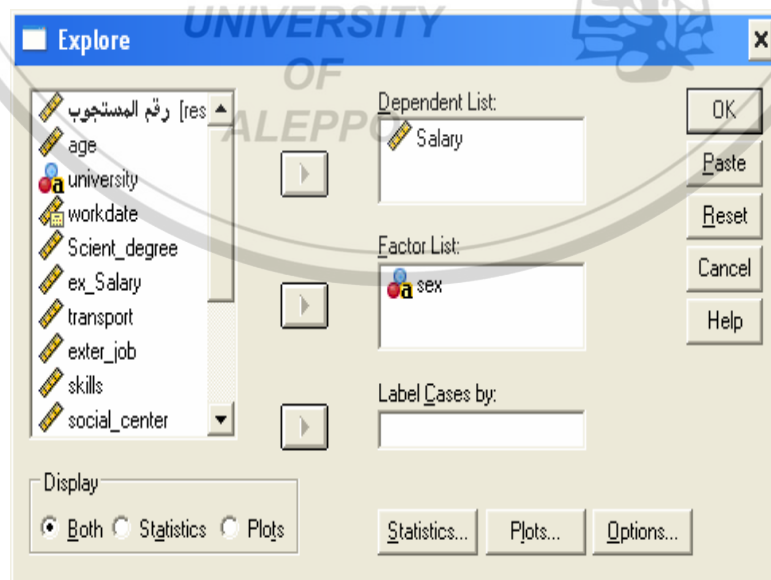
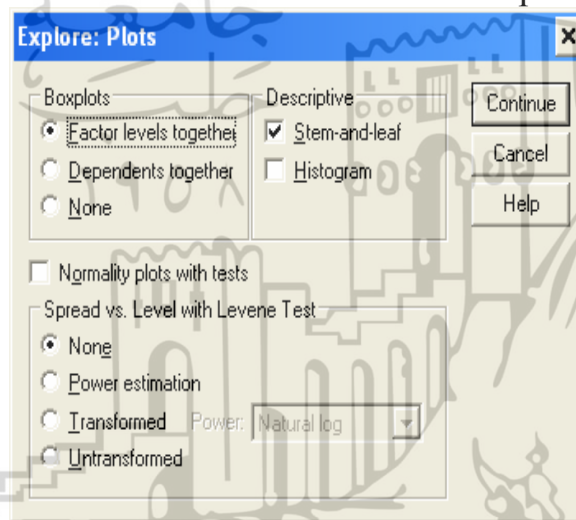


نفعل الخيارات التالية بعد الضغط على الزر Statistics: Estimate : لتقدير معالم النموذج ، Model Fit لبناء جدول تحليل التباين وحساب قيمة معامل التحديد.

تشكيل ملف أوامر بطريقة الكتابة من صندوق الحوار يشتمل على فحص واستكشاف متغير الدخل حسب الجنس مع تفعيل الخيار عرض كل من الإحصائيات والأشكال البيانية Display both: Statistics & Plots ، إذ إن المقصود بالإحصائيات هي المقاييس الوصفية وبناء مجال ثقة ٩٥% حول المتوسط .



المقصود بالأشكال البيانية: الشكل الصندوقي والمدرج التكراري على شكل الجذع والورقة . Boxplot and stem-and-leaf



FREQUENCIES

VARIABLES=Salary /FORMAT=NOTABLE

/NTILES= 4

/PERCENTILES= 5 10 15 40

/HISTOGRAM NORMAL

/ORDER= ANALYSIS .

تشكيل ملف أوامر بطريقة الكتابة من صندوق الحوار يشتمل على:

- حساب الإحصاءات الوصفية لمتغير الدخل مع حفظ القيم المعيارية للدخل كمتغير.

DESCRIPTIVES

VARIABLES=Salary /SAVE

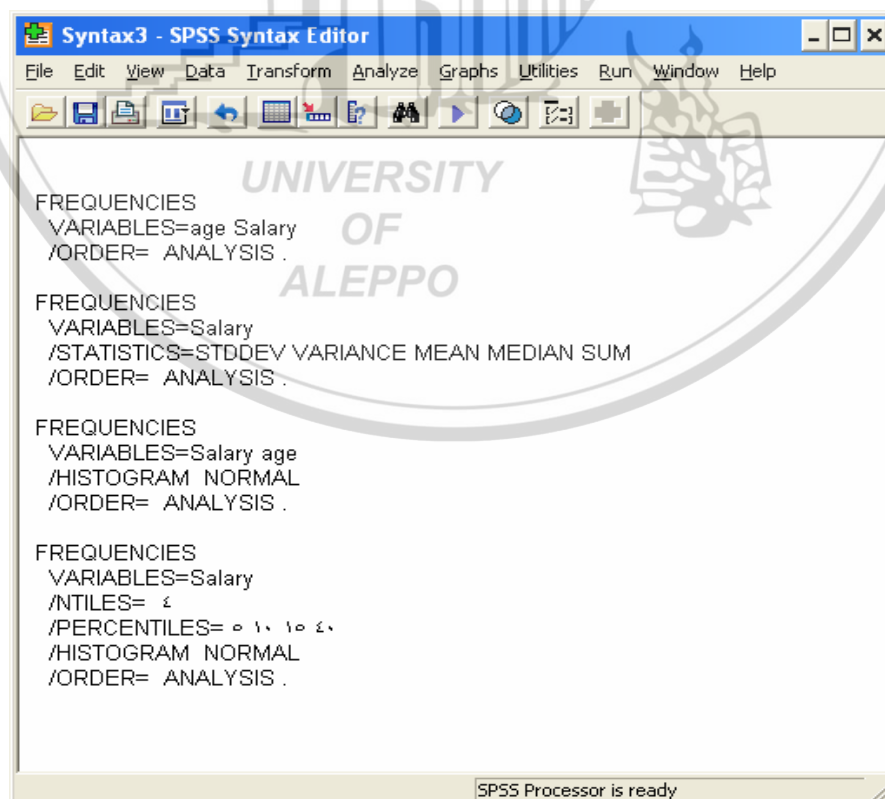
/STATISTICS=MEAN STDDEV MIN MAX .

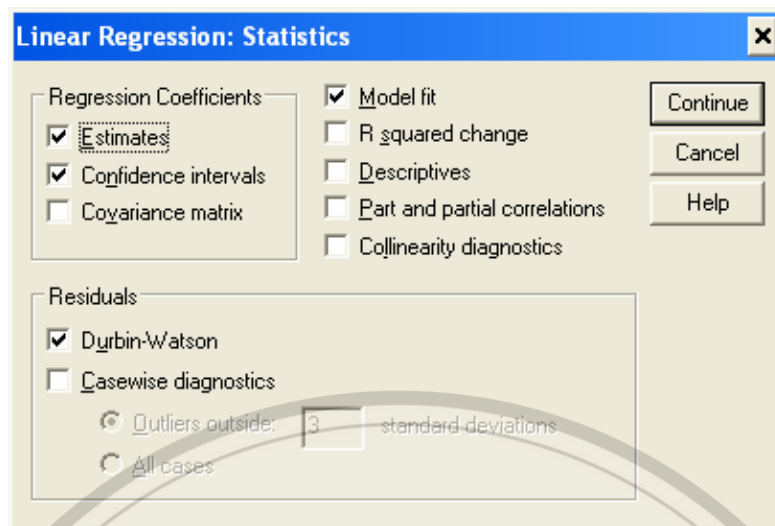
- حساب الإحصاءات الوصفية لمتغير الدخل.

DESCRIPTIVES

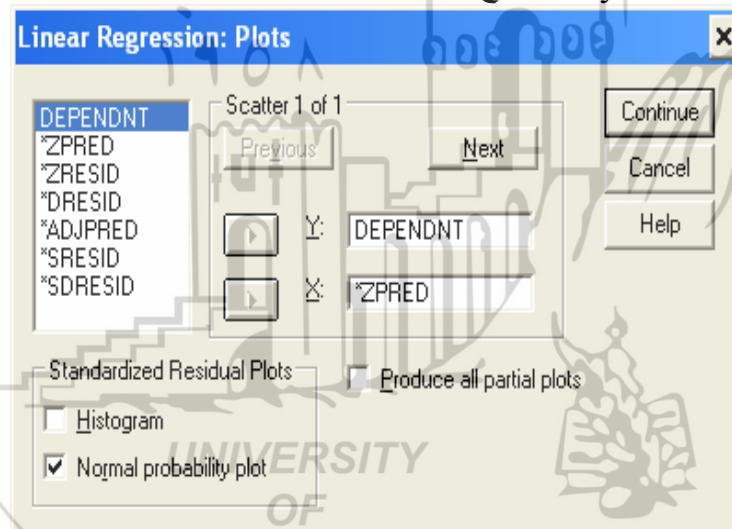
VARIABLES=Salary

/STATISTICS=MEAN STDDEV MIN MAX .

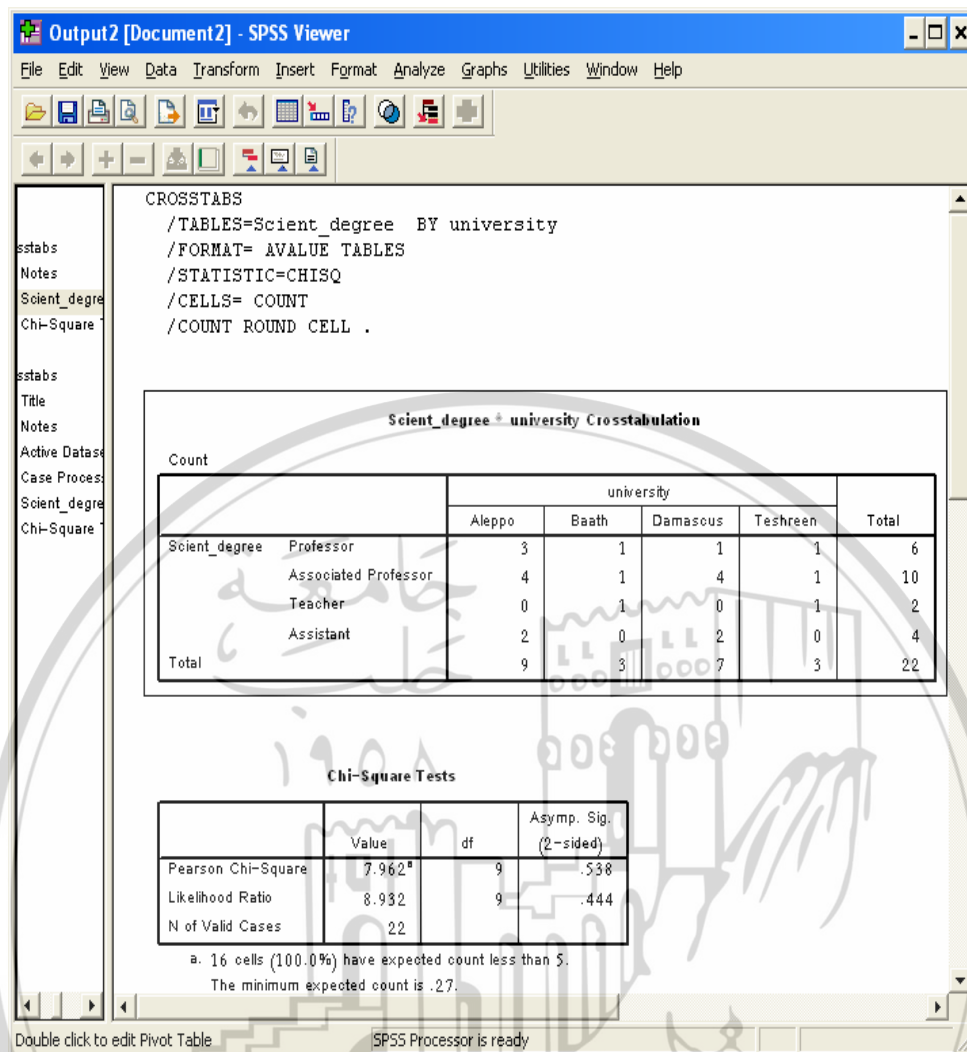




نضغط Continue للعودة الى صندوق الحوار المعنون Linear Regression .
نفعل الخيارات التالية بعد الضغط على الزر Plots:
Normal Probability Plot: لاختبار كون المتغير المدروس يتبع التوزيع الطبيعي.
رسم شكل الانتشار للمتغير Salary مع المتغير " القيمة المقدرة المعيارية " .



بالضغط على الزر Paste يتم فتح ملف الأوامر Syntax File بشكل تلقائي باسم افتراضي
Syntax2:SPSS Syntax Editor ولصق أوامر الانحدار المختارة فيه كما هو موضح
بالشكل التالي:



§-٤ نسخ الأوامر من Journal File

إن كافة الأوامر التي تنفذ خلال جلسات إجراء التحاليل الإحصائية المختلفة يتم تنفيذها في Journal File تحت مسمى spss.jnl وعليه يمكن استدعاء هذا الملف الموجود ضمن الموقع الافتراضي C:\Windows\Temp.

- تنفيذ الأوامر البرمجية Run Command Syntax

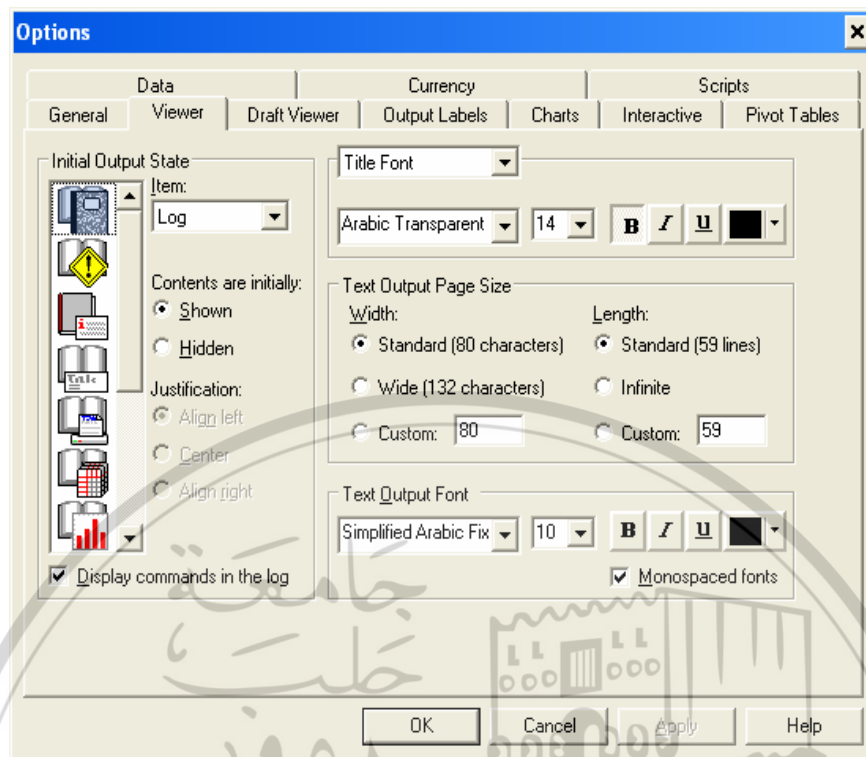
لتنفيذ ملف الأوامر أو أي جزء منه نتبع الخطوات التالية:

فتح ملف الأوامر Syntax File المخزن سابقاً.

File → Open → Syntax

من شريط القوائم في شاشة SPSS Syntax Editor يتم اختيار Run لتنفيذ الأوامر البرمجية إذ تتوفر الخيارات التالية:

All : لتنفيذ كافة الأوامر البرمجية في الملف.



مثال (٣) :

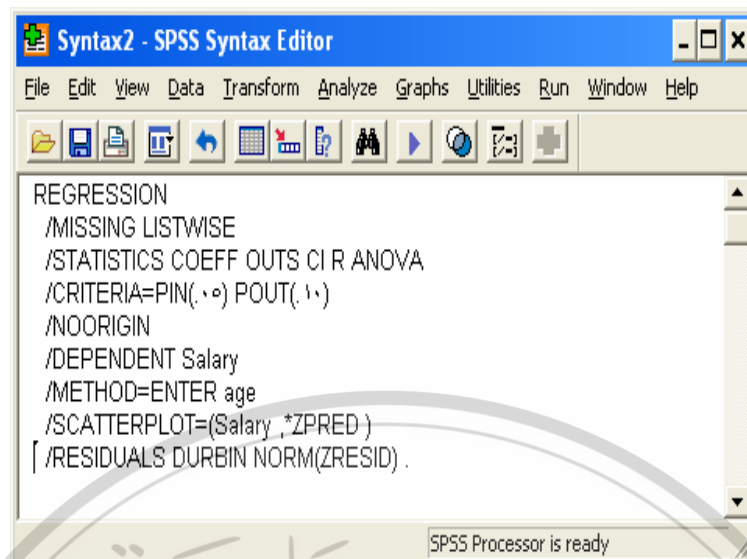
لنفترض أننا نرغب تشكيل جدول تقاطعي يضم أعضاء الهيئة التدريسية في الجامعات السورية حسب مرتبتهم العلمية مع حساب إحصائية كاي تربيع نتبع ما يلي:
من شريط القوائم نختار:

Analysis → Descriptive Statistics → Crosstabs

يظهر صندوق حوار بعنوان Crosstabs ننقل المتغير المرتبة العلمية Scient_Degree للنافذة المعنونة Row(s) وننقل المتغير الجامعة التي ينتمي إليها University للنافذة المعنونة Column(s).

بالضغط على الزر Statistics يظهر صندوق حوار بعنوان Crosstabs: Statistics إذ نختار Chi-Square.

نضغط Continue ثم OK فنحصل على مخرجات الجدول التقاطعي في شاشة .



§-٣ نسخ الأوامر من Log في مخرجات برنامج SPSS

يوفر برنامج SPSS للمستخدم إمكانية نسخ الأوامر من Log في مخرجات البرنامج وذلك عند تنفيذ أمر أي تحليل إحصائي مثل إنشاء جدول تقاطعي، تشكيل جدول توزيع تكراري، حساب المقاييس الإحصائية الوصفية، ولكن قبل ذلك يجب إظهار الـ Log في مخرجات البرنامج كما يلي:

من شريط القوائم لشاشة محرر البيانات Data Editor نختار Options → Edit يظهر صندوق حوار بعنوان Options .

بالنقر على الزر Viewer في صندوق حوار Options للانتقال إلى نافذة Viewer ، حيث نقوم بتفعيل الخيار عرض الأوامر في المربع المعنون بـ:

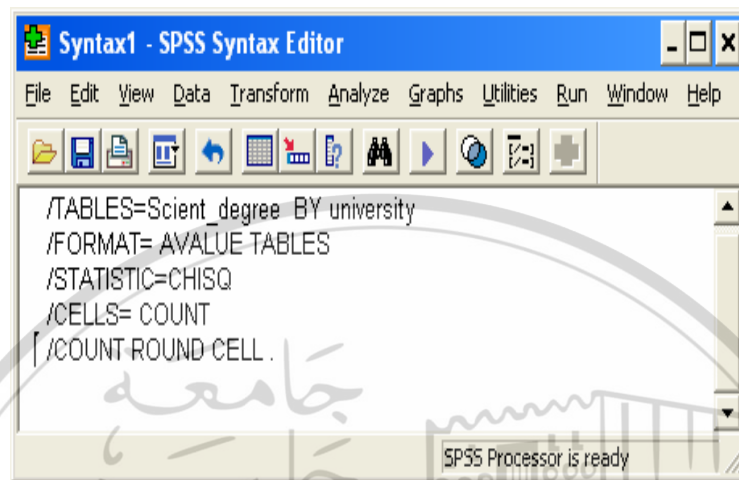
Display Commands in the log.

بالنقر على الزر OK ، نكون بذلك قد فعلنا خيار إظهار الأوامر البرمجية مع مخرجات أي تحليل إحصائي ينفذ في شاشة مخرجات البرنامج.

Selection : لتنفيذ الأوامر المختارة المظللة في الملف.

Current : لتنفيذ الأوامر في موقع المؤشر.

To End : لتنفيذ الأوامر من موقع المؤشر الى نهاية الملف.



الأخطاء وليس البحث عنها ومعالجتها فقط بل إتباع مفهوم التحسين المستمر باعتباره المنطق الأساسي لإدارة الجودة الشاملة (Total Quality Management (TQM). أما مقدرة العملية فهي تقنية رقابية تهدف لتحسين جودة المنتج، وتعني قدرة العملية على أن تطابق أو تفوق المواصفات المحددة مسبقاً من قبل مهندسي التصميم لمقابلة وظيفة محددة للمنتج.

يشمل ضبط الجودة الجهود اللازمة لتحقيق الجودة، والحفاظ على استمرارية جودة مرتفعة للمنتجات والخدمات المزمع استخدامها. بمعنى أنها الجهود المبذولة لضمان صلاحية المنتج أو الخدمة للاستخدام وهذه الصلاحية يتم قياسها وفقاً لمواصفات ومحددات معينة، إلا أن طرق قياس هذه المواصفات تخضع للأخطاء، وعندما تختلف المقاييس فإن سبب هذه الاختلافات يمكن دراسته باستخدام الطرق الإحصائية. وبالنتيجة فإن الضبط الإحصائي للجودة يعني تطبيق المبادئ والأساليب الإحصائية في جميع مراحل المنتج أو الخدمة بهدف اكتشاف الانحرافات عن المواصفات المحددة للسلعة وإلى إظهار ضرورة اتخاذ إجراءات علاجية في عملية الإنتاج، ولهذه العملية طابع الرقابة المستمرة على جودة ومواصفات المنتجات.

يؤثر في العملية الإنتاجية نوعين من التغيرات:

الأول: يمكن التعرف عليه وبالتالي يمكن التحكم فيه (مثل عدم خبرة العامل، تقادم الآلات والمعدات والتجهيزات، نوعية المواد الخام المستخدمة في العملية الإنتاجية).

الثاني: ذو طبيعة عشوائية لا يمكن التحكم فيه.

ولهذا السبب تقسم الاختلافات عن المواصفات المطلوبة للمنتج إلى اختلافات يمكن تحديدها Assignable (أي التعرف على هذه الاختلافات وتصحيحها) واختلافات عشوائية Random لا يمكن التحكم فيها.

يمكن النظر إلى طرق ضبط جودة الإنتاج³ على أنه تطبيق عملي لمقرر الإحصاء. إذ تستخدم خرائط ضبط الجودة لمعرفة مدى تحكمنا في العملية الإنتاجية، وتستخدم مقاييس النزعة المركزية ومقاييس التشتت مثل الوسط الحسابي والمدى والانحراف المعياري في إنشاء خرائط ضبط الجودة. وغالباً ما يعتمد بناء هذه الخرائط على عينات صغيرة جداً، وباستخدام نتائج الاختبارات التي نحصل عليها من هذه العينات، تعمم النتائج على جميع الوحدات المنتجة.

³ لنكولن تشاو " الإحصاء في الإدارة " ص ٩٨١-٩٨٣ .

٤- العامل Operator : يشمل هذا المصدر للتغير الطريقة التي يؤدي بها العامل العملية. وتسهم حالة العامل الصحية والنفسية أيضاً في التغير.

المصادر الأربعة السابقة تسهم في التغير الحقيقي ، كما توجد تغيرات تسجل بسبب عملية الفحص Inspection ، فمعدات الفحص الخاطئة أو التطبيق غير الصحيح لنمطيات الجودة يمكن أن يتسبب في تسجيل تغيرات غير صحيحة. بشكل عام يجب أن يكون التغير في الفحص مساوياً الى عُشر الأربعة مصادر الأخرى للتغير، مع ملاحظة ان ثلاثة من هذه المصادر تكون موجودة في نشاط الفحص وهي : الفاحص ، ومعدة الفحص، والبيئة.

وبما أن هذه المصادر للتغير تتذبذب بطريقة طبيعية أو متوقعة، فيتطور نمط مستقر من المسببات العشوائية للتغير تعود للحظ وتسمى مسببات الفرصة Chance Causes.

إن مسببات الفرصة للتغير تكون حتمية، ونظراً لكونها عديدة وأهميتها بسيطة فهي صعبة الاكتشاف أو التحديد . أما مسببات التغير كبيرة الحجم فتصنف على أنها مسببات محددة Assignable Causes . وعندما لا توجد إلا مسببات الفرصة فقط في العملية، تعتبر العملية في حالة مراقبة إحصائية، وتكون مستقرة ويمكن التنبؤ بها. إلا أنه عندما توجد مسببات محددة للتغير أيضاً، فيكون التغير زائداً وتصنف العملية على أنها خارج المراقبة أو خارج التغير الطبيعي المتوقع.

§-١ مفهوم الضبط الإحصائي للجودة:

يلعب الضبط الإحصائي للجودة دوراً أساسياً في العملية الإنتاجية ، حيث تتركز وظيفته على سلامة التصميم الهندسي وكفاءة المنتج من حيث مطابقته للمواصفات القياسية . ولما كان الفحص الشامل لكل المنتجات عملية مكلفة وغير اقتصادية، جاء الضبط الإحصائي للجودة Statistical Quality Control أداة اقتصادية فعالة لضبط جودة المنتجات من خلال أخذ عينات عشوائية باستمرار ، واتخاذ القرار بقبول المنتج أو رفضه .

يعرّف الضبط الإحصائي للجودة على أنه مجموعة من الإجراءات المنتظمة التي تتبعها الشركة لقياس الجودة الحقيقية لأداء المنتج، ومقارنته بالمواصفات المحددة والإجراءات التصحيحية التي تتخذ في حال وجود أي اختلاف. وهذا معناه استبعاد الوحدات المعيبة ودراسة أسبابها لوضع برنامج لمعالجتها، أي أن الهدف هو العمل قدر الإمكان على منع حدوث

الفصل التاسع

الضبط الإحصائي للجودة^١

Statistical Quality Control

مقدمة

أحد الافتراضات أو الحقائق البديهية للتصنيع^٢ هي أنه لا يمكن أن ينتج جزأين متشابهين بدقة تماماً. وفي الحقيقة، مفهوم التغير هو قانون طبيعي في أنه لا يوجد عنصران طبيعيان يقعان في نفس الفئة ويكونان الشيء نفسه تماماً، والتغير يمكن أن يكون كبيراً جداً وسهل الملاحظة مثل أطوال الأشخاص، أو يمكن أن يكون صغيراً جداً كرقائق الثلج المتساقطة في الأجواء الجليدية، عندما تكون التغيرات صغيرة جداً لدرجة تبدو متطابقة، إلا أن أجهزة القياس الدقيقة يمكن أن تبين الفروقات، وكلما أصبحت أجهزة القياس دقيقة، استمرت التغيرات في الوجود.

يوجد ثلاث فئات للتغيرات في إنتاج أجزاء القطع:

- تغير داخل القطعة: التغير بخشونة السطح للقطعة.
 - تغير من قطعة لأخرى: التغير الحاصل عبر القطع المنتجة بنفس الوقت.
 - التغير من وقت لآخر: التغير في المنتج الذي ينتج في أوقات مختلفة من اليوم.
- تحدث التغيرات في كل عملية بسبب خلط المعدات، والمواد، والبيئة، والعامل.
- مما تقدم يمكن القول بأن هناك أربعة مصادر للتغير هي:

- ١- المعدات Equipment : يشمل هذا المصدر تآكل العدد وذبذبات الآلات والتغيرات الهيدروليكية والكهربائية.
- ٢- المواد Material: يشمل هذا المصدر التغير في خواص الجودة مثل قوة الشد أو اللدونة، أو السماكة أو محتوى الرطوبة والتي يمكن أن تسهم في التغير الكلي للمنتج النهائي.

- ٣- البيئة Environment : إن درجة الحرارة، والضوء، والضغط والرطوبة يمكن أن تسهم في التغير في المنتج.

^١ د. محمد كبيه ، د. أميرة عبيدو " مقدرة العملية والضبط الإحصائي للجودة " بحث مقبول للنشر في مجلة بحوث جامعة حلب ، العدد ٣٠ لعام ٢٠٠٣ م .

^٢ أ.د. دال بستر فيلد " الرقابة على الجودة " ترجمة ومراجعة الدكتور سرور ابراهيم سرور . ص ١٢١-١٢٣ ، الناشر المكتبة الأكاديمية ١٩٩٥ م.

٢-٤ تقنيات ضبط الإحصائي للجودة Tools Statistical Quality Control

١-٢ ضبط الجودة للمنتج الصناعي

لقد عرفنا سابقاً الجودة بأنها تعبر عن مدى ملاءمة مواصفات و خصائص المنتج مع رغبات ومتطلبات المستهلك. ولتحقيق ذلك يتوجب اللجوء إلى التقنيات العملية لضبط الجودة، ومن أهم التقنيات التي تحقق ضبط الجودة على المستوى الصناعي نذكر:

أولاً : ضمان جودة المنتج عن طريق ضبط العمليات الإنتاجية Process Control ويتم ذلك بواسطة ضبط الجودة الإحصائي Statistical Quality Control.

- التوزيع التكراري Histograms .
- خرائط التحكم للمتغيرات Control Charts for Variables .
- خرائط التحكم للخواص Control Charts for Attributes .

ثانياً : ضمان الجودة عن طريق فحص واختبار المنتج Inspection and Testing ويتم ذلك بواسطة خطط الفحص و المعاينة Acceptance Sampling Plans

ثالثاً: ضمان الجودة عن طريق تطبيق مفهوم إدارة الجودة الشاملة TQM .

٢-٢ الاختلافات التصنيعية للمنتجات

- يتسم الإنتاج الصناعي الحديث بأنه إنتاج كمي (Mass Production) ينتج أعداداً كبيرة من الوحدات المتماثلة.
- على الرغم بأن جميع الوحدات مصنعة بالتصاميم نفسها وبالمواصفات نفسها، إلا أن خواصها الميكانيكية، الفيزيائية أو الكيماوية قد تكون فيها بعض الاختلافات والتباينات.
- يطلق على الاختلافات في خصائص المنتج بالاختلافات التصنيعية Manufacturing Variations.

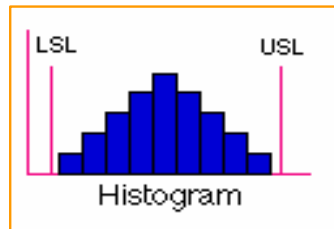
- أسباب الاختلافات التصنيعية

- أولاً : المادة الخام Raw Material .
- ثانياً: النظام الصناعي Manufacturing System ويشتمل على: الآلات Machine ، العمالة الفنية Man، طرق التصنيع Methods، بيئة العمل Environment
- ثالثاً : المنتج الصناعي Product

Raw Material → Manufacturing System → Product → Customer

⁴ د. محمد عيشوني أستاذ مساعد – قسم التقنية الميكانيكية – ٢٠٠٤ m_aichouni@yahoo.co.uk

فئات معينة و من ثم تحديد قيم الأعداد لكل فئة، وهذا ما نطلق عليه اسم تكرار الفئة (Frequency).



Histograms / Frequency Distribution

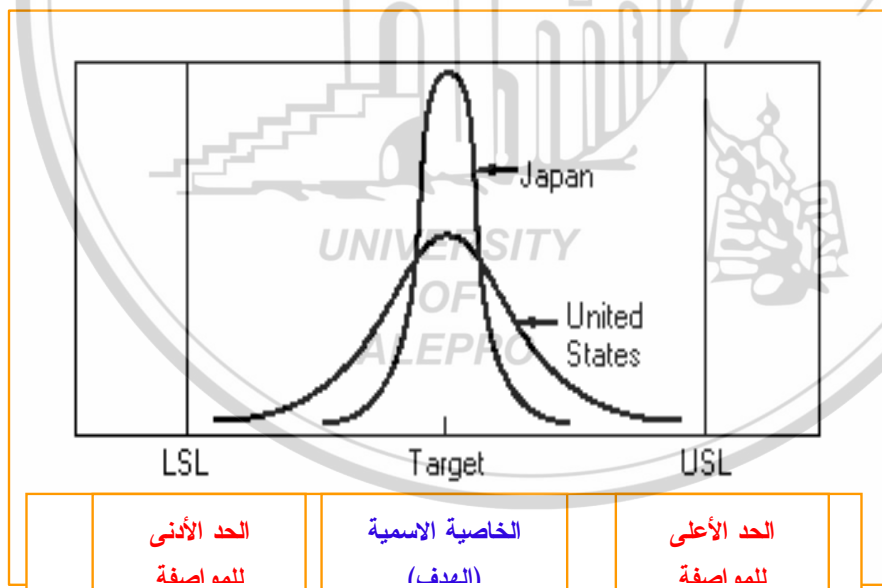
بهدف دراسة نتائج البيانات الخام للجودة نقوم بحساب المعاملات الإحصائية التالية :

• مدى القيم Range

• القيمة المتوسطة Average / Mean Value

• الانحراف المعياري Standard Deviation

عن طريق التوزيع التكراري والمعاملات الإحصائية يمكن دراسة مدى اتساع مجال الاختلافات التصنيعية وتشتت خصائص الجودة عن المواصفات المطلوبة.



§-٣ طرق مراقبة الجودة Methods of Quality Control

١-٣ معاينة القبول Acceptance Sampling

تتلخص الاجراءات بفحص المنتجات النهائية، حيث يتم اختيار عينة من المنتجات أو العناصر من كل دفعة، يتم عد العناصر غير المطابقة defective items وتسجل. باستخدام

٥-٢ التعامل مع الاختلافات التصنيعية

- نأخذ عينة من الإنتاج .
- نقوم بعمليات الفحص والقياس.
- نسجل نتائج قياسات خصائص المنتج. فمثلاً نقوم بتسجيل أبعاد القطع المنتجة: القطعة الأولى X_1 الثانية X_2 الثالثة X_3 القطعة X_n ، إذ: $X_1 \neq X_2 \neq X_3 \neq \dots \neq X_n$
- تمثل نتائج القياسات التي جمعت والمتعلقة بخواص المنتج بيانات خام للجودة.
- عادة ما تكون هذه البيانات على شكل قيم عددية غير منتظمة و يشوبها تباين كثير.
- عن طريق الأساليب الإحصائية يمكن استخلاص معلومات مهمة عن جودة المنتج.

٦-٢ الضبط الإحصائي للجودة Statistical Quality Control

تعتمد طريقة الضبط الإحصائي للجودة على تحليل نتائج الفحوصات والاختبارات لخواص الجودة باستعمال الطرق الإحصائية (Statistics). وتعتبر هذه الطريقة من بين أكثر الأساليب أهمية في مجال ضبط الجودة للمنتجات الصناعية إذ تم استعمالها في هذا المجال منذ نهاية الحرب العالمية الثانية من طرف كبرى الشركات العالمية. مع التطور المذهل في برامج الحاسب الآلي فقد زادت أهمية هذه الطريقة و عرفت تطبيقات مهمة في المجال الصناعي.

١-٦-٢ تقنيات الضبط الإحصائي للجودة Statistical Quality Control Tools

١. التوزيع التكراري Histograms / Frequency Distribution
٢. خرائط ضبط الجودة للمتغيرات Control Charts for Variables
٣. خرائط ضبط الجودة للخواص Control Charts for Attributes
٤. طريقة أبشيكوا Cause and Effect Diagram
٥. مخطط باريتو Pareto Chart
٦. مخطط التلاحق Run Chart
٧. شكل الانتشار Scatter Diagram
٨. مخططات التدفق Flow Charts

٢-٦-٢ تقنية التوزيع التكراري :

يعتبر التوزيع التكراري أهم تقنية لدراسة حالة جودة المنتج الصناعي في مرحلة الإنتاج. يمكن من خلال هذه التقنية تلخيص البيانات الخام لجودة المنتج و تصنيفها (تبويبها) إلى

المواد الأولية <-- نظام التصنيع <-- المنتج <-- المستهلك

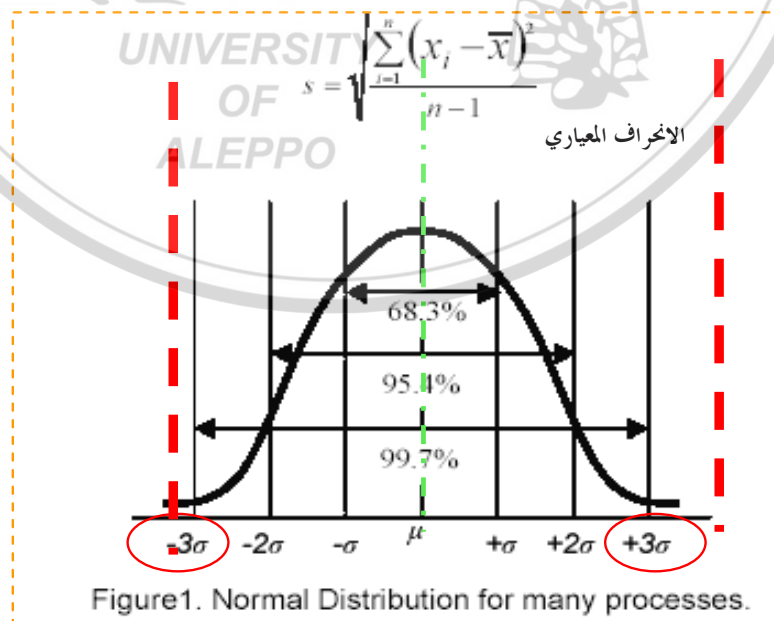
٣-٢ العلاقة بين الاختلافات التصنيعية و جودة المنتج

إن للاختلافات التصنيعية أثر سيئ على جودة المنتج. فكلما زادت نسبة الاختلافات وكبر مجالها ساء مستوى الجودة، لأن خصائص المنتج تكون متباعدة ومشتتة عن المواصفات القياسية للمنتج المصممة خصيصاً قصد إرضاء المستهلك. إن تحسين مستوى الجودة Quality Improvement يعني التقليل من الاختلافات التصنيعية، كما يعني كذلك تقليص و تقليل عدد الوحدات المعيبة.

٤-٢ مقياس التشتت في المنتج و التوزيع الطبيعي:

أكدت الدراسات أن معظم العمليات الإنتاجية تتبع التوزيع الطبيعي (Normal Distribution)، كما أن مقدرة العملية الإنتاجية تحدد كالتالي:

$\mu \pm 1\sigma$	٦٨,٢٧% من المنتج تكون في حدود
$\mu \pm 2\sigma$	٩٥,٤٥% من المنتج تكون في حدود
$\mu \pm 3\sigma$	٩٩,٧٣% من المنتج تكون في حدود
$\mu \pm 1.96$	٩٥% من المنتج تكون في حدود
$\mu \pm 2.58\sigma$	٩٩% من المنتج تكون في حدود



القواعد الإحصائية يتم الحكم على الدفعة بالقبول acceptable أو عدم قبول الدفعة unacceptable. إن نظام معاينة القبول MIL-STD-105E هو الأكثر شيوعاً.

٢-٣ ضبط العملية إحصائياً: *Statistical Process Control (SPC)*

- هناك مجموعة من الإجراءات المتبعة لمراقبة وفحص عمليات التصنيع.
- زيادة فرصة كون المنتجات النهائية مطابقة للمواصفات .
- مخططات الضبط Control Charts: تقنيات بيانية لمراقبة ورصد الانتاج مرحلة بمرحلة.
- الطرق البيانية Graphical Methods: مجموعة من التقنيات لتحديد الاسباب المحتملة للمنتجات غير المحققة لشروط الضبط .

- تحليل باريتو Pareto Analysis : أدوات لوضع الاولويات

- تصميم التجارب *Design of Experiments*.

مجموعة من الإجراءات المستخدمة لتحديد افضل تصميم للمنتج

٤-٤ خرائط ضبط الجودة:

تعتبر خرائط الضبط من أهم التقنيات الإحصائية المستخدمة في ضبط وتحسين الجودة وتسمى أحياناً خرائط شيوارت Shewart نسبة لمكتشفها، وهي طريقة إحصائية بيانية تظهر الحدود الدنيا والعليا للعملية التي نريد مراقبتها، وتتكون من ثلاثة خطوط أفقية يمثل الخط الأوسط منها مستوى الجودة المرغوب فيه، ويمثل الخطان الآخران الحد الأدنى والحد الأعلى للجودة المتوقعة (كما تساعد على معرفة طبيعة الانحرافات^٥ في مستوى جودة المنتج).

لبناء خريطة الضبط يتم حساب الوسط الحسابي (وسط أوساط العينة، وسط قياسات المدى، وسط نسبة الوحدات المعيبة) للعملية الإنتاجية، ولتحديد حدي الضبط يضاف إلى الوسط الحسابي، ومن ثم يطرح من الوسط الحسابي غالباً ثلاثة انحرافات معيارية (أي إذا كانت ظروف العملية الإنتاجية طبيعية فإن احتمال أن تقع الانحرافات في جودة المنتج ضمن حدي الضبط الأدنى والأعلى سيكون مساوياً إلى ٩٩,٧٣ %).

يوجد نوعان من خرائط ضبط الجودة هما:

- خرائط ضبط المتغيرات (خريطة الوسط الحسابي والمدى ، خريطة الوسط الحسابي والانحراف المعياري للعينات) .
- خرائط ضبط المواصفات (خريطة نسبة الوحدات المعيبة، وعدد العيوب في الوحدة).

^٥ مرد الانحرافات أو الاختلافات عن مواصفات الإنتاج إما عوامل طبيعية متوقعة وإما عوامل غير طبيعية من العمال أو المواد أو الآلات أو مجتمعة، إن الانحرافات في العملية الإنتاجية تخضع للتوزيع الطبيعي.

$$Lower \ \& \ Upper \ CL_{\bar{X}} = \bar{\bar{X}} \mp A_2 \bar{R} ,$$

A_2 : عامل factor يتتغير بتغير حجم المجموعات الجزئية Subgroups وتستخرج من

جداول معدة خصيصاً لحسابها، والجدول التالي يمثل قيم بعض العوامل.

$\bar{R}$: متوسط قياسات المدى للعينات.

Control Chart Factors Table.			
Subgroup size	D3	D4	A2
2	0	3.267	1.880
3	0	2.575	1.023
4	0	2.282	0.729
5	0	2.114	0.577
6	0	2.004	0.483
7	0.076	1.924	0.419
8	0.136	1.864	0.373
9	0.184	1.816	0.337
10	0.223	1.777	0.308
11	0.256	1.744	0.285
12	0.283	1.717	0.266
13	0.307	1.693	0.249
14	0.328	1.672	0.235
15	0.347	1.653	0.223
16	0.363	1.637	0.212
17	0.378	1.622	0.203
18	0.391	1.609	0.194
19	0.404	1.596	0.187
20	0.415	1.585	0.180
21	0.425	1.575	0.173
22	0.435	1.565	0.167
23	0.443	1.557	0.162
24	0.452	1.548	0.157
25	0.459	1.541	0.153

يتم رسم خريطة الضبط وتؤشر عليها القياسات المعبرة عن الجودة ومن ثم تدرس وتحلل النقاط المنتشرة في المستوي الإحداثي الممثل لخريطة ضبط الجودة، فإذا كانت كل النقاط الممثلة

ومنها يتم حساب الوسط الحسابي العام والذي يمثل الحد الأوسط على خريطة الضبط ، وتحسب بموجب العلاقة التالية :

$$\bar{\bar{X}} = \frac{\sum_{i=1}^m \bar{x}_i}{m} = \frac{\bar{x}_1 + \bar{x}_2 + \dots + \bar{x}_m}{m} = CL$$

حيث:

$\bar{\bar{X}} = CL$ تمثل الحد الأوسط لخريطة الوسط الحسابي.

$\bar{x}_i$, m تمثل على الترتيب عدد العينات، متوسط العينة رقم i .

لقد حددت حدود الضبط التجريبية لخرائط الضبط عند ± 3 انحرافات معيارية من الحد الأوسط لخريطة الضبط (في عملية الضبط الإحصائي للجودة ينتشر استخدام مجالات ثقة ذات 6 أخطاء معيارية على طرفي الوسط الحسابي العام ، مفترضين أن احتمال وقوع الوسط الحسابي للعينة خارج هذا المجال يعتبر صغيراً جداً لا يتجاوز 0.0027 ، وذلك استناداً للخاصية التي يتمتع بها التوزيع الطبيعي حيث أن 99.73 % من المشاهدات المدروسة تقع في المساحة المحصورة تحت منحنى التوزيع الطبيعي والمحور الأفقي الممتد مابين $\bar{x} - 3\sigma$ و $\bar{x} + 3\sigma$) ، وذلك بموجب العلاقة التالية :

$$UCL_{\bar{X}} = \bar{\bar{X}} + 3\sigma_{\bar{x}} , \quad LCL_{\bar{X}} = \bar{\bar{X}} - 3\sigma_{\bar{x}}$$

حيث أن :

UCL , LCL : تمثلان على الترتيب حدي الضبط الأدنى والأعلى .

$\sigma_{\bar{X}}$: الانحراف المعياري لقياسات الأوساط الحسابية للعينات.

وإذا كان الانحراف المعياري للمجتمع مجهولاً، نستخدم الانحراف المعياري للعينة على أنه تقدير للانحراف المعياري للمجتمع، وتصبح حدود الضبط التجريبية لخرائط الوسط الحسابي للضبط:

$$UCL_{\bar{X}} = \bar{\bar{X}} + 3s_{\bar{x}} = \bar{\bar{X}} + 3\left(\frac{s}{\sqrt{n}}\right) ,$$

$$LCL_{\bar{X}} = \bar{\bar{X}} - 3s_{\bar{x}} = \bar{\bar{X}} - 3\left(\frac{s}{\sqrt{n}}\right)$$

ولتبسيط العمليات الحسابية تستخدم الصيغة التالية لحدود الضبط التجريبية:

أولاً: خرائط ضبط المتغيرات Control Charts for Variables

تهدف الأساليب الإحصائية المستخدمة في ضبط جودة الإنتاج التحكم في العملية الإنتاجية^٦ نفسها، فمثلاً تعبئة عبوات زجاجية متتالية سعة ٢ كغ من رب البندورة (عصير الطماطم) باستخدام أجهزة ومعدات مصممة لتعبئة كل عبوة بالكمية نفسها يعتبر مثلاً لعملية إنتاجية، كما أن إنتاج أنابيب حديدية بطول ١٠٠ سم وقطر ٥ مم بوساطة ماكينات خراطة وتسوية مصممة لإنتاج كل أنبوب بالطول نفسه والقطر نفسه يعتبر مثلاً لعملية إنتاجية . إن الهدف الأول من ضبط العملية الإنتاجية هو توحيد مقاييس المنتج الصناعي لتتفق مع معايير ومواصفات قياسية معروفة وسبق تحديدها من قسم التصميم في الشركة. فعند تعبئة العبوات الزجاجية بعصير الطماطم وإنتاج الأنابيب الحديدية، يعتبر الوزن في الحالة الأولى والطول والقطر في الحالة الثانية بمنزلة متغيرات مستمرة يمكن قياسها.

تعتبر خرائط ضبط الجودة وسيلة لاكتشاف أو لمعرفة الانحرافات التي تحدث في مقاييس النزعة المركزية ومقاييس التشتت لمجموعة من المفردات وبيان ما إذا كانت العملية الإنتاجية مستقرة أم لا .

ومن المعلوم أن هناك مسببات طبيعية (ذات طبيعة عشوائية لا يمكن التحكم فيها) للانحراف وهي حتمية لا نستطيع عمل شيء بشأنها وهي صعبة التحديد والاكتشاف وفي حال وجودها فقط تكون العملية في حالة ضبط إحصائي.

أما في حالة وجود مسببات غير طبيعية (عدم خبرة العامل، تقادم الآلات والتجهيزات، نوعية المواد الخام) يكون الانحراف كبيراً وتكون العملية خارج الضبط الإحصائي .

١-٤ خريطة الوسط الحسابي $\bar{X}$ charts

تستخدم خريطة الوسط الحسابي لإظهار اختلافات الأوساط الحسابية للعينات عن الوسط الحسابي العام لمشاهدات جميع العينات (إذا كانت أحجام العينات متساوية فإن الوسط الحسابي العام يساوي الوسط الحسابي للأوساط الحسابية لجميع العينات)، أخذين في الاعتبار ضرورة كون المتغير المراد رسم خريطة له صفة للجودة يمكن قياسها والتعبير عنها رقمياً. يتم اختيار المجموعات الجزئية (العينات العشوائية) الرشيدة^٧ حجماً وعدداً ومن ثم تؤخذ القياسات ،

^٦ مجموعة عمليات يفترض تكرارها تحت نفس الظروف.

^٧ المجموعات التي يعود الانحراف فيها لأسباب طبيعية فقط.

للعينات واقعة ضمن حدي الضبط فإن العملية الإنتاجية تكون مضبوطة إحصائياً وتكون مسببات الانحراف طبيعية. وفي حالة وقوع إحدى النقاط الممثلة للعينات خارج حدي الضبط، تكون العملية الإنتاجية خارج حدود الضبط الإحصائية، وتكون مسببات الانحراف غير طبيعية ومردّها أسباب محددة يمكن معرفتها (كفاءة العامل، المادة الخام، الآلة) ويمكن تحد يد حدود جديدة للضبط عن طريق استبعاد بعض العينات وبالتالي أوساطها الحسابية.

مثال (١):

تقوم إحدى المعدات الأوتوماتيكية الموجودة بمصنع لحفظ المواد الغذائية بتعبئة الفاكهة المجففة في صناديق. قام مهندس ضبط جودة الإنتاج بسحب عينة عشوائية مكونة من خمسة صناديق من إنتاج هذه الآلة كل ساعة وذلك خلال ١٠ ساعات، ثم قام بتسجيل وزن كل صندوق منها (لا توجد مواصفات قياسية معتمدة خاصة بوزن الصندوق) في الجدول التالي:

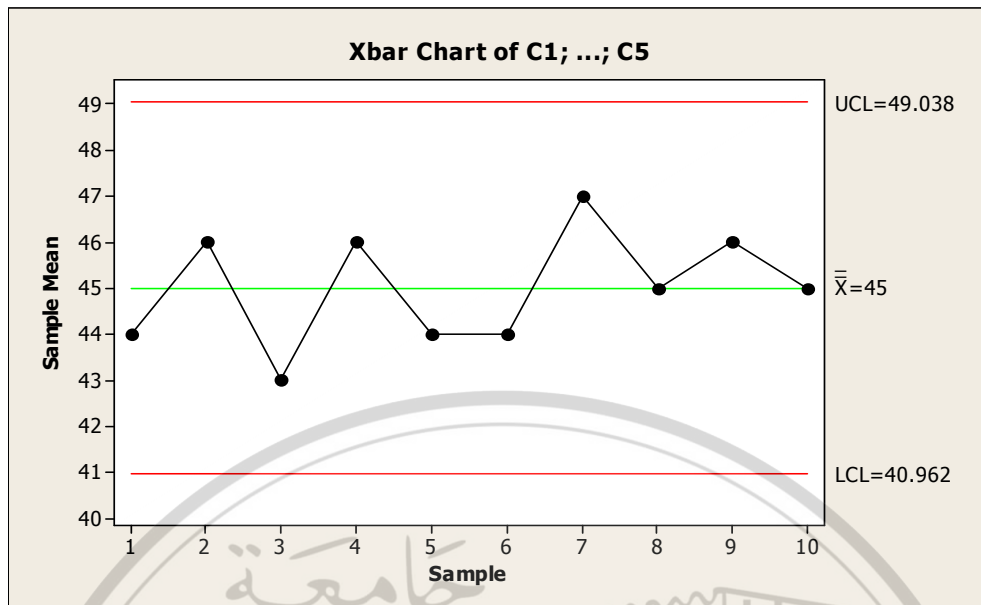
والمطلوب :

- وقّع بيانات العشر ساعات على خريطة الوسط الحسابي .
- هل العملية الإنتاجية محكمة (مضبوطة إحصائياً أو لا) ؟

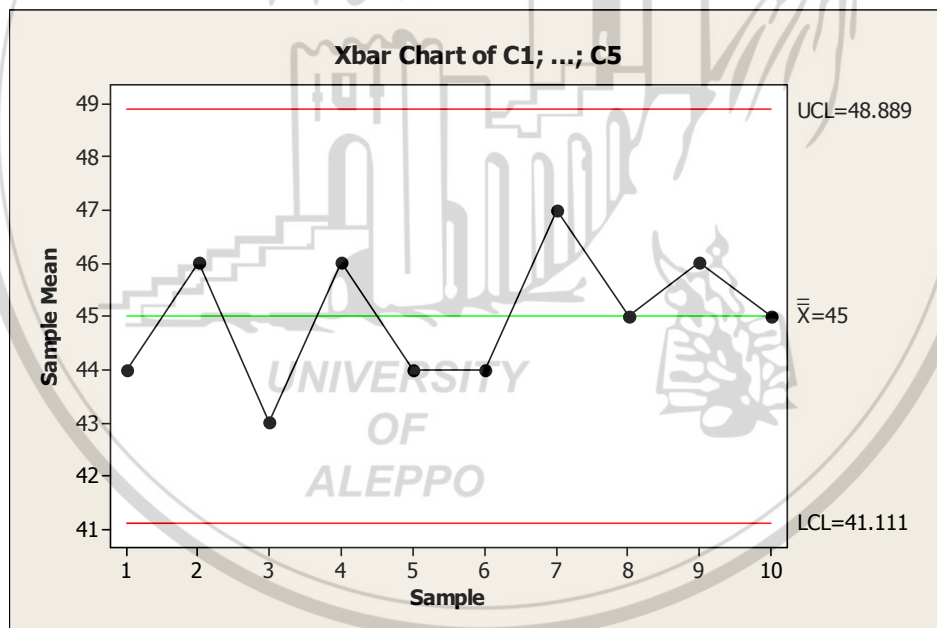
العينات	الصناديق				
	1	2	3	4	5
1	42	44	45	47	42
2	44	46	48	47	45
3	40	41	42	44	48
4	47	48	46	45	44
5	47	44	46	44	39
6	38	47	45	46	44
7	50	46	47	47	45
8	42	48	46	44	45
9	47	50	45	48	40
10	47	44	47	49	38

الحل:

نفترض أن X تمثل وزن صندوق الفاكهة بالكيلو غرامات، إن كل عينة تتضمن خمس مشاهدات وبالتالي نستخدم هذه المشاهدات لنحصل على الوسط الحسابي لكل عينة. لهذا السبب سنخصص الجدول التالي لحساب مجموع قيم مشاهدات كل عينة، الوسط الحسابي لكل عينة وبالتالي يحسب الوسط الحسابي العام باستخدام الصيغة التالية:



الشكل رقم (1) : خريطة الوسط الحسابي $\bar{X}$ charts
 ملاحظة : إذا وقع الوسط الحسابي لأية عينة فوق حد الضبط الأعلى أو أسفل حد الضبط الأدنى، فيجب عندئذ اتخاذ الإجراءات المناسبة للتعرف على أسباب الاختلافات لتصحيحها.



الشكل رقم (2) : خريطة الوسط الحسابي $\bar{X}$ charts

٢-٤ خريطة المدى R charts :

تستخدم خريطة المدى لبيان تغيرات مدى العينات حول الوسط الحسابي لمدى جميع العينات، يتم رسم خريطة المدى بطريقة مماثلة لطريقة رسم خريطة الوسط الحسابي. وهنا أيضاً يجب أن يكون المتغير المراد رسم خريطة له صفة للجودة يمكن قياسها والتعبير عنها رقمياً، ومن ثم يتم اختيار المجموعات الجزئية (العينات) الرشيدة حجماً وعدداً ومن ثم تؤخذ القياسات

وبالتالي فإن الخطأ المعياري يساوي: $s_{\bar{x}} = \sqrt{1.66} = 1.28841$

وعليه يمكننا حساب حدي الضبط الأعلى والضبط الأدنى لخريطة الوسط الحسابي:

$$UCL_{\bar{x}} = 45 + 3 (1.28841) = 45 + 3.86523 = 48.865$$

$$LCL_{\bar{x}} = 45 - 3 (1.28841) = 45 - 3.86523 = 41.135$$

وبين الشكل رقم (1) حدي الضبط الأعلى والأدنى بالإضافة إلى الأوساط الحسابية للعينات. وبما أن الأوساط الحسابية لجميع العينات تقع في نطاق الضبط، فإننا نستنتج أن العملية الإنتاجية محكمة (مضبوطة إحصائياً).

لتبسيط العمليات الحسابية تستخدم الصيغ التالية لحدود الضبط التجريبية:

$$Upper \quad UCL_{\bar{x}} = \bar{\bar{X}} + A_2 \bar{R} = 45 + 0.577 (7) = 49.0390$$

$$Lower \quad LCL_{\bar{x}} = \bar{\bar{X}} - A_2 \bar{R} = 45 - 0.577 (7) = 40.9610$$

n	A2	D4
2	1.88	3.27
3	1.02	2.57
4	0.73	2.28
5	0.577	2.11

باستخدام برنامج Minitab حصلنا على خريطة الوسط الحسابي، في حالة كون الطريقة المستخدمة بتقدير الانحراف هي متوسط المدى لقياسات مدى العينات وحجم العينة أكبر من ١.

$$\bar{\bar{X}} = \frac{\sum_{i=1}^m \bar{x}_i}{m} = \frac{\bar{x}_1 + \bar{x}_2 + \dots + \bar{x}_m}{m} = \frac{450}{10} = 45$$

كذلك يبين الجدول نتائج حساب تباين العينات (لأن تباين المجتمع مجهول فإننا سنقوم بتقديره باستخدام تباين العينة) ، وذلك على النحو التالي :

$$s_1^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{18}{4} = 4.5$$

$$s_2^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{10}{4} = 2.5$$

$$s_{10}^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{74}{4} = 18.5$$

من المتوقع أن نصل إلى تقدير أفضل لتباين المجتمع باستخدام تباينات جميع العينات بدلاً من استخدام تباين عينة واحدة فقط.

العينات	مجموع المشاهدات	متوسط العينات	مجموع مربعات الانحرافات	تباين العينات
1	220	44	18	4.5
2	230	46	10	2.5
3	215	43	40	10
4	230	46	10	2.5
5	220	44	38	9.5
6	220	44	50	12.5
7	235	47	14	3.5
8	225	45	20	5
9	230	46	58	14.5
10	225	45	74	18.5
المجموع	2250	450	332	83
الوسط الحسابي		45		8.3

يمكن استخدام متوسط تباينات العينات كتقدير لتباين المجتمع، ومن الجدول نجد أن مجموع تباينات العينات يساوي 83 وعليه يمكننا الحصول على متوسط تباينات العينات كما يلي:

$$s_x^2 = \frac{\sum_{i=1}^m s_i^2}{m} = \frac{83}{10} = 8.3$$

باستخدام تباين العينة يمكن حساب تباين توزيع المعاينة للوسط الحسابي بموجب العلاقة:

$$s_{\bar{x}}^2 = \frac{s_x^2}{n} = \frac{8.3}{5} = 1.66$$

ومنها يتم حساب الوسط الحسابي العام ووسط قياسات المدى واللذين يمثلان الحد الأوسط على خريطة الضبط وذلك بموجب الصيغة التالية :

$$\bar{R} = \frac{\sum_{i=1}^m R_i}{m} = \frac{R_1 + R_2 + \dots + R_m}{m} = CL \quad (2)$$

حيث :

$\bar{R} = CL$ تمثل الحد الأوسط لخريطة المدى.

R_i , m تمثل على الترتيب عدد العينات، مدى العينة.

لقد حددت حدود الضبط التجريبية لخرائط الضبط عند ± 3 انحرافات معيارية من الحد الأوسط لخريطة الضبط أي أن:

$$\begin{aligned} UCL_{\bar{X}} &= \bar{R} + 3\sigma_R, & LCL_R &= \bar{R} - 3\sigma_R \\ UCL_{\bar{X}} &= \bar{R} + 3s_R, & LCL_R &= \bar{R} - 3s_R \end{aligned} \quad (3)$$

LCL , UCL : حدا الضبط الأدنى والأعلى .

s_R : الانحراف المعياري لقياسات المدى المحسوب وفقاً للصيغة التالية:

$$s_R = \sqrt{\frac{\sum_{i=1}^m (R_i - \bar{R})^2}{m-1}} \quad (4)$$

ولتبسيط العمليات الحسابية تستخدم الصيغ التالية لحدود الضبط التجريبية:

$$\begin{aligned} UCL_R &= D_4 \bar{R} \\ LCL_R &= D_3 \bar{R} \end{aligned} \quad (5)$$

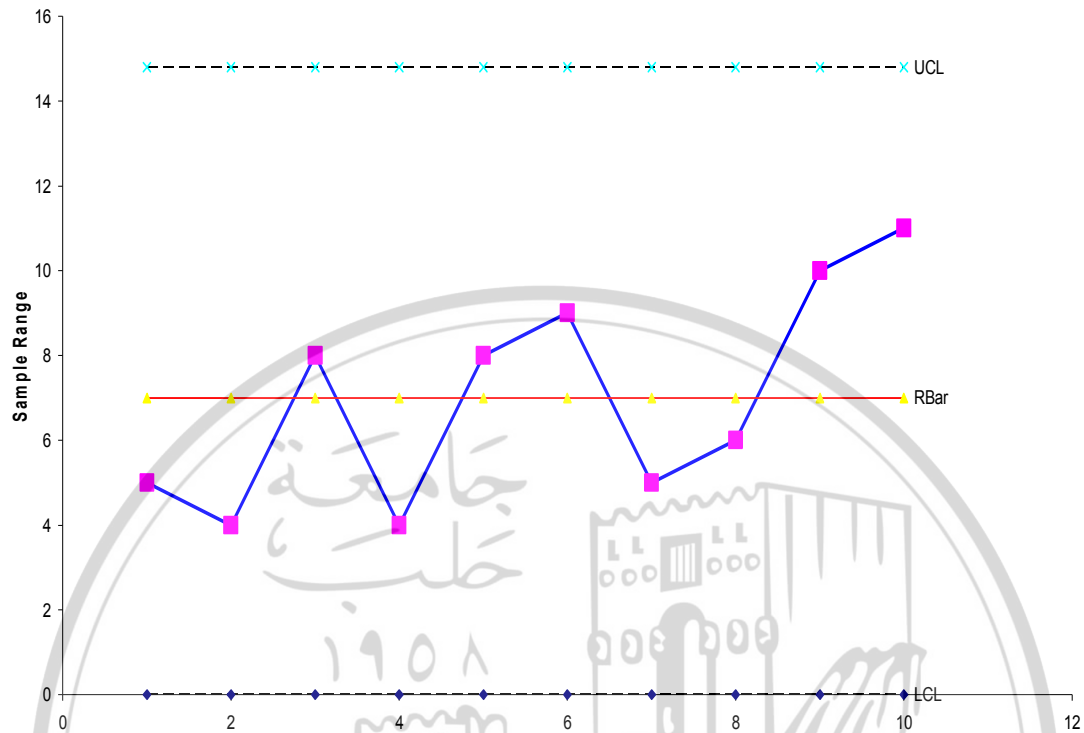
D_3 , D_4 : عوامل factors تتغير بتغير حجم المجموعات الجزئية Subgroups .

يتم رسم خريطة الضبط وتؤشر عليها القياسات المعبرة عن الجودة ومن ثم تدرس وتحلل النقاط المنتشرة في المستوي الإحداثي الممثل لخريطة ضبط الجودة ، فإذا كانت

كل النقاط الممثلة للعينات واقعة ضمن حدي الضبط فإن العملية الإنتاجية تكون مضبوطة إحصائياً وتكون مسببات الانحراف طبيعية. وفي حالة وقوع إحدى النقاط الممثلة للعينات خارج حدي الضبط، تكون العملية الإنتاجية خارج حدود الضبط الإحصائية، وتكون مسببات الانحراف غير طبيعية ومردّها أسباب محددة يمكن معرفتها (كفاءة العامل، المادة الخام، الآلة).

ملاحظة (2): في حالة كون حد الضبط الأدنى لخريطة المدى سالباً (ونظراً لعدم وجود إمكانية لأن يكون المدى سالباً)، يجب أن يتم وضع حد الضبط الأدنى مساوياً للصفر.

مثال (٢) :



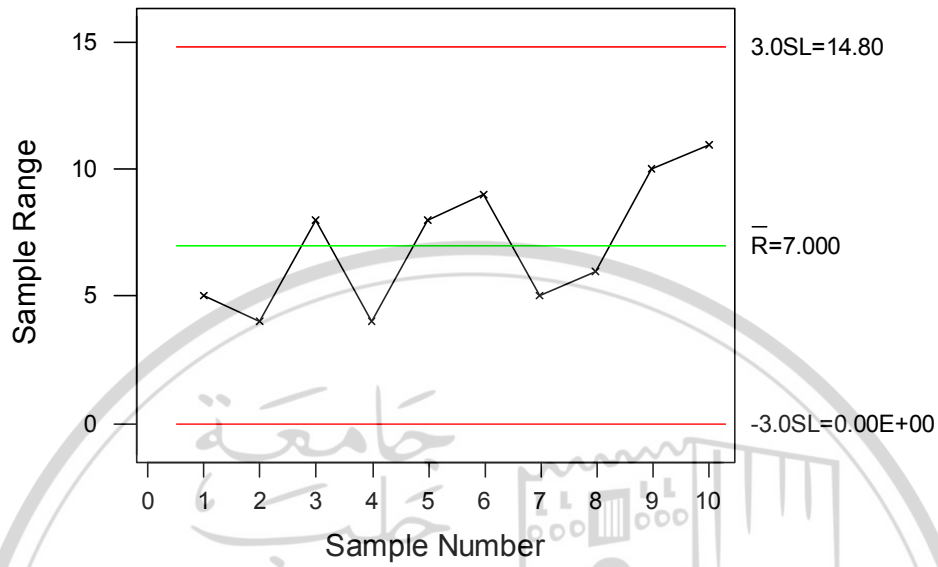
مثال (٣) :

للمحافظة على سمعة ممتازة على أنه مقدم لخدمات الطعام، يحتاج مراقب الجودة الى قياس جودة الانتاج وجودة تقديم الخدمة على أساس منتظم.

المطلوب :

استخدام برنامج Microsoft Excel لتطوير عملية قياس جودة الانتاج وجودة الخدمة. ملف البيانات Cattleman.xls يشتمل على ساعة تقديم الخدمة، زمن تقديم الخدمة.

R Chart for Fruit bo



الشكل رقم (٣) : خريطة المدى R charts :

ولتبسيط العمليات الحسابية تستخدم الصيغ التالية لحدود الضبط التجريبية:

$$UCL_R = D_4 \bar{R} = 2.114(7) = 14.798$$

$$LCL_R = D_3 \bar{R} = 0(7) = 0$$

UNIVERSITY
OF
ALEPPO

المطلوب: وقّع بيانات العشر ساعات المتعلقة بالمثال السابق على خريطة المدى.

- هل العملية الإنتاجية محكمة (مضبوطة إحصائياً أم لا) ؟

العينات	المدى R_i	$\sum_{i=1}^m (R_i - \bar{R})^2$
1	5	4
2	4	9
3	8	1
4	4	9
5	8	1
6	9	4
7	5	4
8	6	1
9	10	9
10	11	16
المجموع	70	58
الوسط الحسابي		45

يبين الجدول السابق مدى كل عينة من العينات المتاحة، ويتم حساب الوسط الحسابي للمدى :

$$\bar{R} = \frac{\sum_{i=1}^m R_i}{m} = \frac{70}{10} = 7 \quad (6)$$

وبما أن مجموع مربعات انحرافات مدى كل عينة عن الوسط الحسابي للمدى يساوي 58 ، فإن الانحراف المعياري للمدى يساوي:

$$s_R = \sqrt{\frac{\sum_{i=1}^m (R_i - \bar{R})^2}{m-1}} = \sqrt{\frac{58}{10-1}} = \sqrt{6.44} = 2.538$$

بالاستناد إلى ما تقدم يمكننا حساب حدي الضبط الأعلى والأدنى كما يلي:

$$UCL_R = 7 + 3 (2.539) = 7 + 7.617 = 14.617$$

$$LCL_R = 7 - 3 (2.539) = 7 - 7.617 = -0.617 = 0$$

يبين الشكل التالي رقم (2) حدي الضبط الأعلى والأدنى بالإضافة إلى قياسات المدى للعينات،

وبما أن قياسات المدى لجميع العينات تقع في نطاق الضبط، فإننا نستنتج أن العملية الإنتاجية

محكمة. وإذا وقع المدى لأية عينة فوق حد الضبط الأعلى أو أسفل حد الضبط الأدنى، فيجب

عندئذ اتخاذ الإجراءات المناسبة للتعرف على أسباب الاختلافات حتى يمكن تصحيحها.

	A	B	C	D	E	F	G	H
1	Hour	Table 1	Table 2	Table 3	Table 4		Mean	Range
2	1	16	18	21	23		19.5	7
3	2	26	20	19	19		21	7
4	3	20	22	18	18		19.5	4
5	4	24	16	22	20		20.5	8
6	5	17	19	24	17		19.25	7
7	6	17	17	15	18		16.75	3
8	7	22	12	20	22		19	10
9	8	24	19	19	17		19.75	7
10	9	18	18	20	14		17.5	6
11	10	17	23	19	15		18.5	8
12	11	20	20	17	21		19.5	4
13	12	21	17	21	23		20.5	6
14	13	22	17	22	17		19.5	5
15	14	16	19	18	19		18	3
16	15	17	18	15	23		18.25	8
17	16	19	17	21	17		18.5	4
18	17	19	19	13	16		16.75	6
19	18	21	14	17	16		17	7
20	19	18	17	25	18		19.5	8
21	20	20	18	20	19		19.25	2
22	21	23	21	23	21		22	2
23	22	20	20	20	14		18.5	6
24	23	18	18	26	15		19.25	11
25	24	20	22	23	21		21.5	3
26	25	23	22	21	24		22.5	3
27	26	22	14	21	19		19	8
28	27	18	20	18	22		19.5	4
29	28	19	20	16	14		17.25	6
30	29	21	19	16	20		19	5
31	30	22	22	18	21		20.75	4

الحل:

- نفتح الملف، نحسب متوسط زمن تقديم الخدمة ونضع الناتج في المجال (G2:G31) بوساطة

الدالة Average(G2:G31) .

- نحسب مدى زمن تقديم الخدمة ونضع الناتج في المجال (H2:H31) بوساطة الدالة

Max(B2:B31)-Min(B2:B31) .

- نضغط على PHStat من شريط القوائم ونختار Control Charts and R & Xbar

Charts . يظهر صندوق حوار بمسمى R & Xbar Charts كما في الشكل:

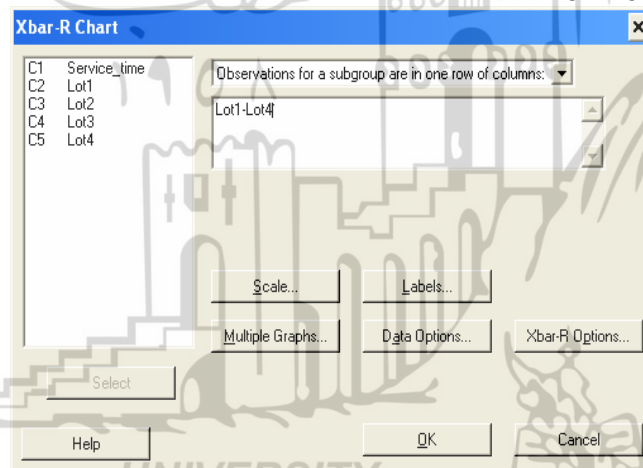
Center	5.733333333
Upper Control Limit	13.08346667
XBar Chart Intermediate Calculations	
Average of Subgroup Averages	19.24166667
A2 Factor	0.729
A2 Factor * RBar	4.1796
XBar Chart Control Limits	
Lower Control Limit	15.06206667
Center	19.24166667
Upper Control Limit	23.42126667

باستخدام برنامج MINITAB يمكن الحصول على خريطتي الوسط الحسابي وخريطة المدى وذلك باتباع الخطوات التالية:

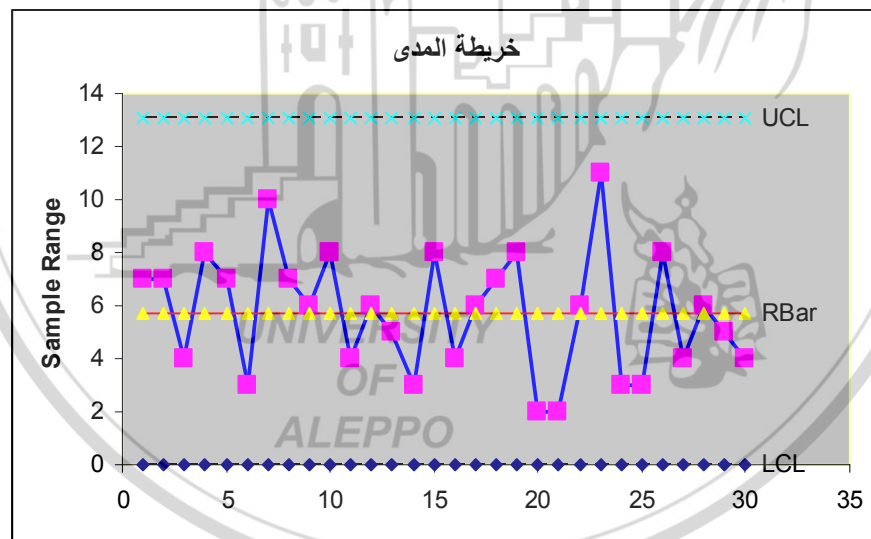
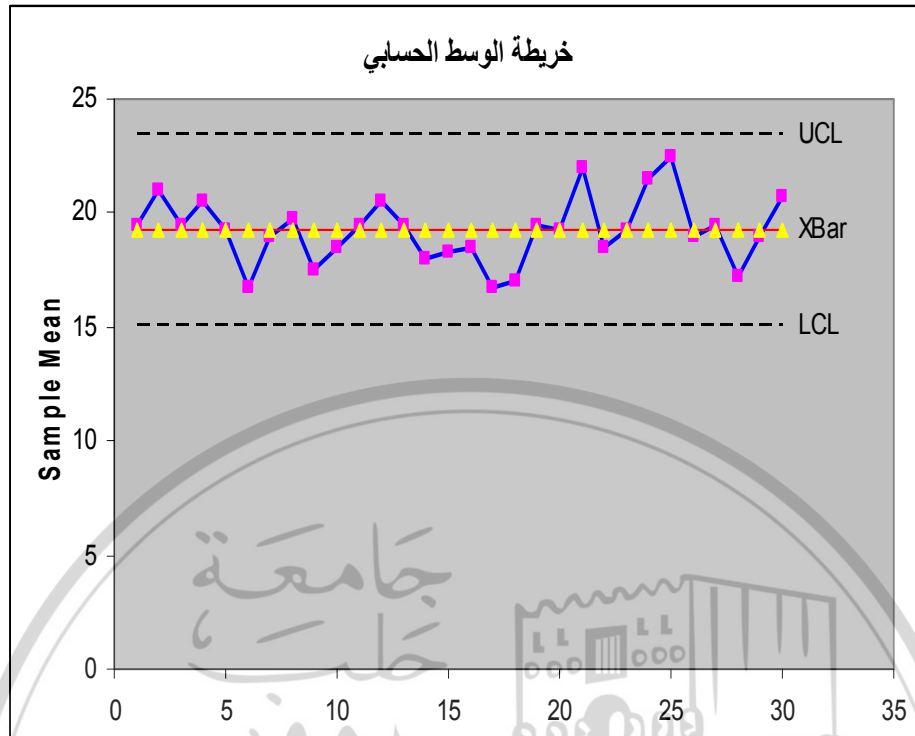
من شريط القوائم نختار :

Stat → Control Charts → Variables Charts for Subgroups → Xbar-R

نحصل على نافذة حوار بعنوان Xbar-R Chart:



نضغط OK نحصل خريطتي المراقبة.



أما الجدول التالي فيوضح النتائج الخاصة بحساب الخط الأوسط وحدود خريطتي المراقبة.

Control Chart Calculations

Data	
Sample/Subgroup Size	4
R Chart Intermediate Calculations	
RBar	5.733333333
D3 Factor	0
D4 Factor	2.282
R Chart Control Limits	
Lower Control Limit	0

R and XBar Charts

Data

Subgroup/Sample Size:

Subgroup Ranges Cell Range:

☒ First cell contains label

Chart Options

☒ R Chart Only

☐ R and XBar Charts

Subgroup Means Cell Range:

☒ First cell contains label

Output Options

Title:

Help OK Cancel

- نضع الرقم 4 في النافذة المعنونة حجم المجموعات الجزئية Subgroup/Sample Size أو حجم العينة.
- نكتب المجال الذي تتواجد فيه قيم المدى والمجال الذي تتواجد فيه قيم الوسط الحسابي بما فيها عناوين الأعمدة وذلك في كل من النافذتين المعنونتين: Subgroup Range Cell Range لقيم المدى و Subgroup Means Cell Range لقيم الوسط الحسابي، ونكتب العنوان المطلوب للخريطة.

R and XBar Charts

Data

Subgroup/Sample Size:

Subgroup Ranges Cell Range:

☒ First cell contains label

Chart Options

☐ R Chart Only

☒ R and XBar Charts

Subgroup Means Cell Range:

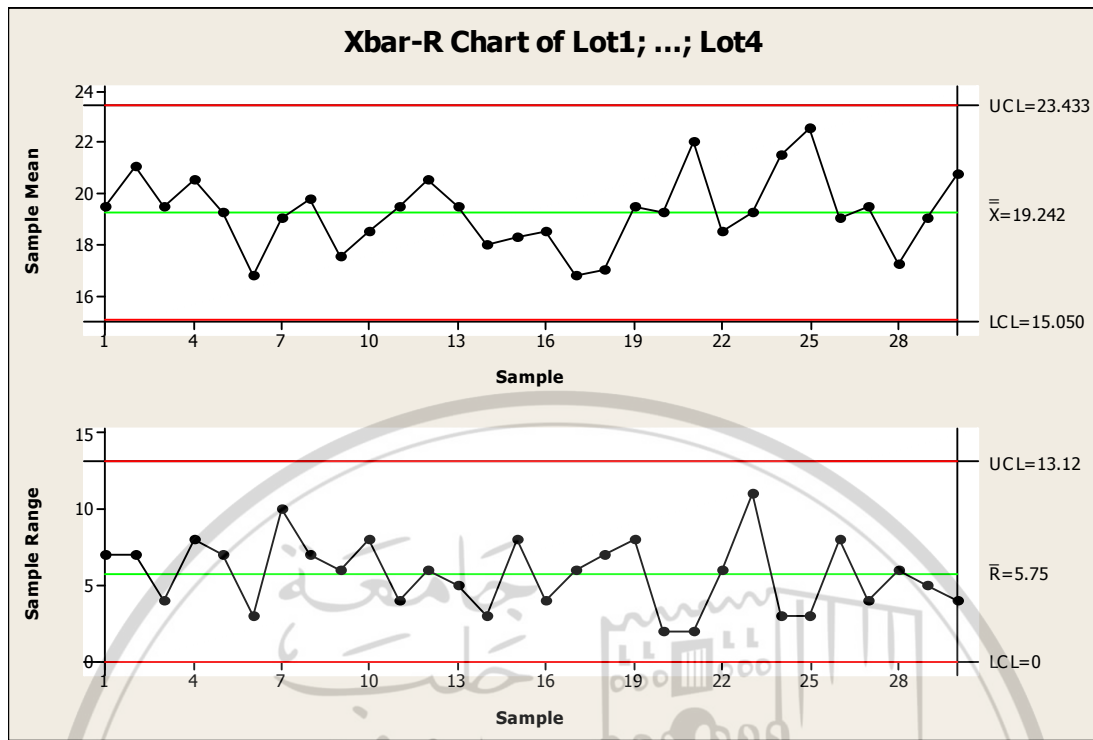
☒ First cell contains label

Output Options

Title:

Help OK Cancel

- نضغط OK نحصل على النتائج كما يلي:



٣-٤ خريطة ضبط الانحراف المعياري للعينات S charts :

The Sample Standard Deviation Control Chart

تستخدم خرائط الانحراف المعياري للعينات كمقياس لتشنت العينات وتعتبر أكثر دقة من خريطة المدى لأنها تحسب باستخدام البيانات كافة ، يتم رسم خريطة المدى بطريقة مماثلة لطريقة رسم خريطة الوسط الحسابي وخريطة المدى. وهنا أيضاً يجب أن يكون المتغير المراد رسم خريطة له صفة للجودة يمكن قياسها والتعبير عنها رقمياً، ومن ثم يتم اختيار المجموعات الجزئية (العينات) الرشيدة حجماً وعدداً ومن ثم تؤخذ القياسات ومنها يتم حساب الوسط الحسابي للانحرافات المعيارية وبشكل يمثل الحد الأوسط على خريطة الضبط.

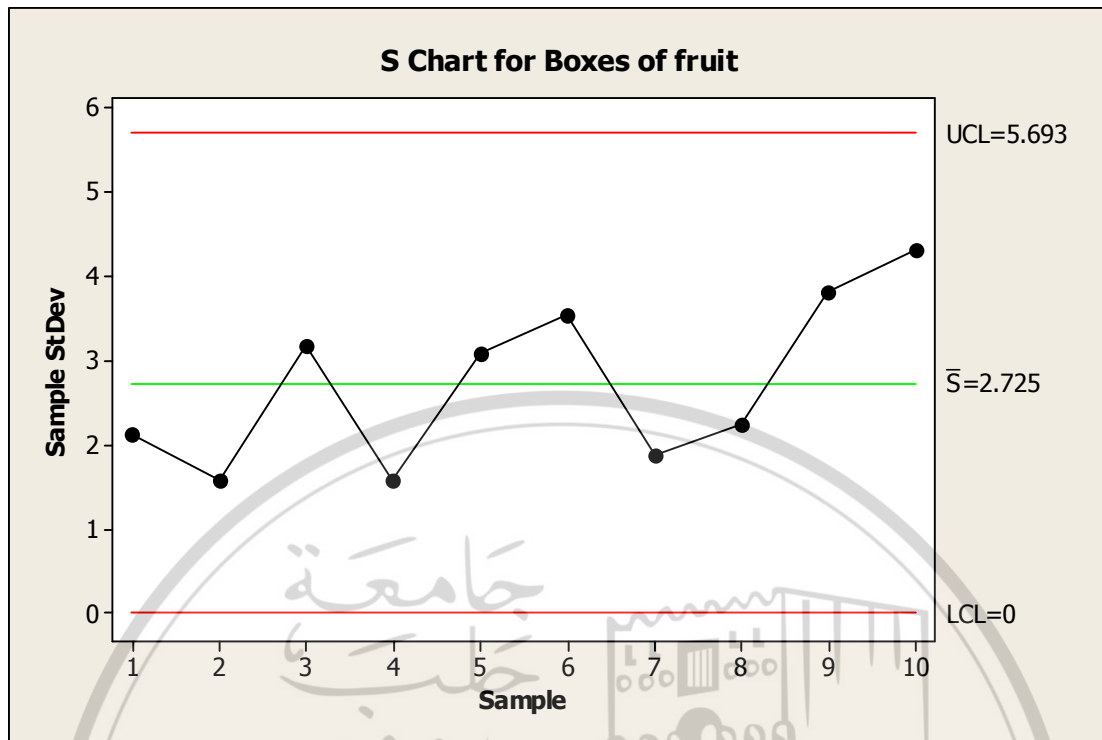
تستخدم الصيغ التالية في حساب الحد الأوسط لخريطة الانحراف المعياري وكذلك حدود الضبط أيضاً لخريطة الانحراف المعياري:

$$\bar{S} = \frac{\sum_{i=1}^m S_i}{m}, \quad S_i = \sqrt{\frac{\sum_{j=1}^n (S_{ij} - \bar{S}_i)^2}{n-1}} \quad (7)$$

حيث:

m تمثل عدد العينات، و n تمثل حجم العينة.

S_i : الانحراف المعياري للقياسات الأصلية للظاهرة المدروسة.



الشكل رقم (4) : خريطة الانحراف المعياري S charts ولتبسيط العمليات الحسابية تستخدم الصيغ التالية لحدود الضبط التجريبية:

$$UCL_S = B_4 \bar{S} = 2.114(2.728) = 5.766992$$

$$LCL_S = B_3 \bar{S} = 0(2.728) = 0$$

ثانياً : خرائط ضبط المواصفات : Control Charts of Attributes

في الحياة العملية كثيراً ما نصادف منتجات لا يمكن التعبير عن جودتها أو مواصفاتها بواسطة متغيرات يمكن قياسها بشكل كمي، بل تكون خصائص جودة المنتج وصفية وبالتالي لا يمكن التعبير عنها على أنها متغير مستمر مثل الطول والوزن، فإما أن يحقق المنتج مواصفات معينة أو لا يحققها.

يشير مفهوم المواصفات في ضبط الجودة إلى أن المنتج إما مطابق أو غير مطابق للمواصفات (الوحدات المنتجة معيبة وغير معيبة، عدد العيوب في المنتج أو في وحدة المساحة والحجم) ، ومن الأمثلة على ذلك المصابيح الكهربائية تضيء أو لا تضيء ، الأجهزة الكهربائية أو أجزاء من هذه الأجهزة تعمل أو لا تعمل.

وهناك نوعان لخرائط ضبط المواصفات:

٤-٤ : خرائط نسبة الوحدات المعيبة $P - Chart$:

$$S_1 = \sqrt{\frac{\sum_{j=1}^n (X_i - \bar{X}_i)^2}{n-1}} = \sqrt{\frac{18}{5-1}} = 2.12132$$

$$S_2 = \sqrt{\frac{10}{5-1}} = 1.581139$$

$$S_{10} = \sqrt{\frac{47}{5-1}} = 4.301163$$

$$\bar{S} = \frac{\sum_{i=1}^m S_i}{m} = \frac{27.280}{10} = 2.728$$

$$s_s = \sqrt{\frac{\sum_{i=1}^m (S_i - \bar{S})^2}{m-1}} = \sqrt{\frac{8.5825}{10-1}} = 0.977$$

وبالاستناد إلى ما تقدم يمكننا حساب حدي الضبط الأعلى والأدنى كما يلي:

$$UCL_S = 2.728 + 3 (0.977) = 2.728 + 2.930 = 5.658$$

$$LCL_S = 2.728 - 3 (0.977) = 2.728 - 2.930 = -0.202 = 0$$

يبين الشكل التالي رقم (٤) حدي الضبط الأعلى والأدنى بالإضافة إلى الانحرافات المعيارية للعينات.

وبما أن الانحرافات المعيارية لجميع العينات تقع في نطاق الضبط، فإننا نستنتج أن العملية الإنتاجية محكمة (مضبوطة إحصائياً).

إذا وقع الانحراف المعياري لأية عينة فوق حد الضبط الأعلى أو أسفل حد الضبط الأدنى، فيجب عندئذ اتخاذ الإجراءات المناسبة للتعرف على أسباب الاختلافات حتى يمكن تصحيحها.

$$LCL_{\bar{S}} = \bar{S} + 3s_s, \quad LCL_R = \bar{S} - 3s_s \quad (8)$$

ولتبسيط العمليات الحسابية تستخدم الصيغ التالية لحدود الضبط التجريبية:

$$UCL_S = B_4 \bar{S}, \quad LCL_S = B_3 \bar{S}$$

حيث: B_3, B_4 عوامل تستخرج من جداول وتؤخذ للحصول على حدود ضبط ضمن ثلاثة انحرافات معيارية لخرائط الانحراف المعياري.

ولرسم خريطة الانحراف المعياري S للعينات المدروسة يتم استخدام الانحراف المعياري المجمع.

$$s_{\bar{S}} = \sqrt{\frac{\sum_{i=1}^m (S_i - \bar{S})^2}{m-1}}, \quad (9)$$

يتم رسم خريطة الضبط وتؤشر عليها القياسات المعبرة عن الجودة ومن ثم تدرس وتحلل النقاط المنتشرة في المستوي الإحداثي الممثل لخريطة ضبط الجودة.

مثال (٤):

المطلوب: وقّع بيانات العشر ساعات المتعلقة بالمثل السابق على خريطة الانحراف المعياري S charts.

- هل العملية الإنتاجية محكمة (مضبوطة إحصائياً أو لا) ؟

العينات	الانحراف المعياري S_i	$\sum_{i=1}^m (S_i - \bar{S})^2$
1	2.121	0.3680
2	1.581	1.3152
3	3.162	0.1886
4	1.581	1.3152
5	3.082	0.1255
6	3.536	0.6522
7	1.871	0.7347
8	2.236	0.2420
9	3.808	1.1662
10	4.301	2.4750
المجموع	27.280	8.5825
الوسط الحسابي	2.728	الانحراف المعياري 0.977

يبين الجدول السابق الانحراف المعياري لكل عينة من العينات المتاحة، وكذلك الوسط الحسابي والانحراف المعياري للانحرافات المعيارية والتي حسبت باستخدام العلاقات التالية:

تعتبر من أهم خرائط المواصفات وبموجبها تصنف الوحدات الخاضعة للفحص إلى وحدات معيبة أو غير معيبة وبالإمكان تطبيق هذه الخريطة على مواصفات اللون، الرائحة وقدرة المنتج على الاستعمال وتعتمد على مبدأ حجم العينة الثابت (يتم اختيار صفة للجودة، يحدد عدد العينات، وعادة 25 عينة على الأقل بحجم 50 وحدة كحد أدنى [أنظر المرجع (1) Besterfield, p239).

يتم حساب وسط العملية الإنتاجية (وسطي نسبة الوحدات المعيبة) وكذلك الانحراف المعياري لمتوسط نسبة الوحدات المعيبة بموجب العلاقة التالية:

$$CL = \bar{P} = \frac{\sum_{i=1}^m P_i}{m}$$

$$S_{\bar{P}} = \sqrt{\frac{\bar{P}(1-\bar{P})}{n}} \quad (10)$$

وتحدد حدود الضبط التجريبية بموجب العلاقة:

$$UCL_p = \bar{P} + 3\sqrt{\frac{\bar{P}(1-\bar{P})}{n}}, \quad LCL_p = \bar{P} - 3\sqrt{\frac{\bar{P}(1-\bar{P})}{n}} \quad (11)$$

وفي حال وجود نقطة واحدة خارج حدي الضبط تعتبر العملية خارج الضبط الإحصائي وهناك أسباب غير طبيعية محددة تؤثر بالعملية يجب إزالتها.

مثال (٥):

لنفترض أن خط التجميع بأحد المصانع الوطنية يقوم بإنتاج محركات أجهزة البرادات المنزلية بأحجام مختلفة قد تكون جيدة أو معيبة. سحبت 20 عينة عشوائياً (من إنتاج الشركة الذي يخضع للفحص قبل عرضه للزبائن) في فترات زمنية متساوية بحيث تحتوي كل منها على 50 محركاً لفحصها، وتم تسجيل عدد المحركات المعيبة في كل عينة منها كما في الجدول.

المطلوب:

إنشاء خريطة نسبة الوحدات المعيبة $P - Chart$ باستخدام بيانات هذه العينات ، هل العملية الإنتاجية محكمة (مضبوطة إحصائياً أو لا) ؟

- عدد المحركات المعيبة في عشرين عينة تضم كل منها 50 محركاً .

العينة	حجم العينة	عدد الوحدات المعيبة	نسبة المعيب
1	50	4	0.08
2	50	2	0.04
3	50	3	0.06
4	50	5	0.1

وحيث أن:

$$n = 50, \bar{P} = 0.056842 \Rightarrow (1 - \bar{P}) = 0.943158$$

فإن الانحراف المعياري للمتغير $\bar{P}$ يساوي:

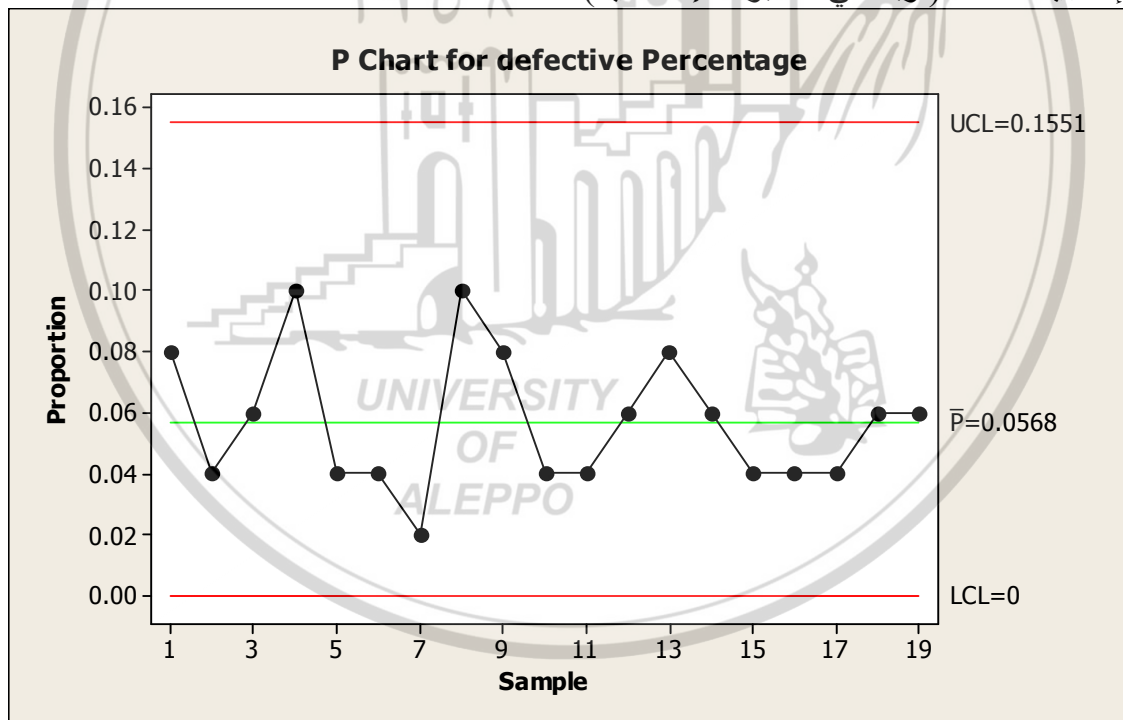
$$S_{\bar{P}} = \sqrt{\frac{\bar{P}(1 - \bar{P})}{n}} = \sqrt{\frac{0.056842(1 - 0.056842)}{50}} = 0.032745$$

استناداً إلى ما تقدم فإن حد الضبط الأعلى وحد الضبط الأدنى لنسبة المحركات المعيبة لأجهزة البرادات المنزلية في العينة يحسب كما يلي:

$$UCL_p = \bar{P} + 3\sqrt{\frac{\bar{P}(1 - \bar{P})}{n}} \\ = 0.056842 + 3(0.032745) = 0.152234$$

$$LCL_p = \bar{P} - 3\sqrt{\frac{\bar{P}(1 - \bar{P})}{n}} \\ = 0.056842 - 3(0.032745) = -0.02823 = 0$$

يبين الشكل التالي رقم (6) حدي الضبط الأعلى والأدنى بالإضافة إلى نسب القطع المعيبة للعينات. وحيث أن الانحرافات المعيارية لجميع العينات تقع في نطاق الضبط، نستنتج أن العملية الإنتاجية محكمة (وبالتالي مضبوطة إحصائياً).



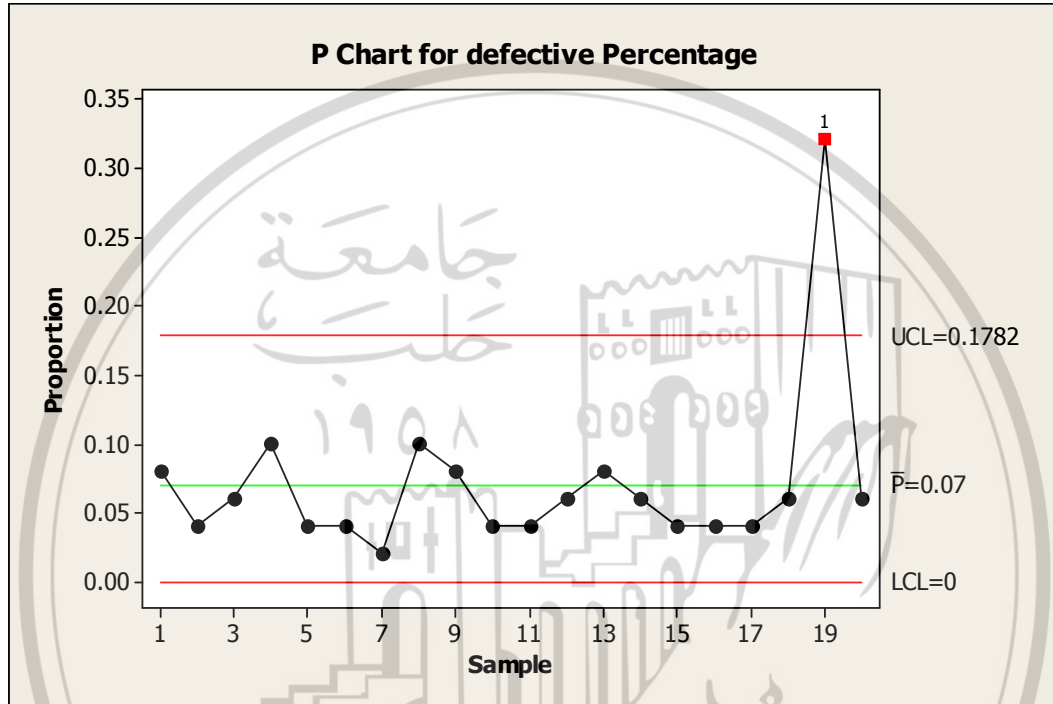
الشكل رقم (5): خريطة نسبة الوحدات المعيبة $P - Chart$

باستخدام برنامج Microsoft Excel نحصل على النتائج السابقة نفسها:

Number		Defective numbers	p	LCL	Center	UCL
1	50	4	0.08		0.07	0.17825
2	50	2	0.04		0.07	0.17825

يبين الشكل التالي رقم (4) حدي الضبط الأعلى والأدنى بالإضافة إلى نسب القطع المعيبة للعينات.

وحيث أن الانحرافات المعيارية لجميع العينات تقع في نطاق الضبط، باستثناء العينة رقم 19 فإنها تقع خارج حدود الضبط الإحصائي، أي أن العملية الإنتاجية غير محكمة (وبالتالي غير مضبوطة إحصائياً).



الشكل رقم (5) : خريطة نسبة الوحدات المعيبة *P - Chart*

يتضح من الشكل رقم (5) أن نسبة المعيب في العينة التاسعة عشرة تقع فوق حد الضبط الأعلى، لذا يفترض وجود أسباب وقت سحب هذه العينة، أدت إلى زيادة نسبة المعيب فيها وبالتالي يجب التعرف عليها بدرجة أفضل لمعرفة الأسباب وتصحيحها ، وللحصول على متوسط للنسب يمكن الاعتماد عليها بدرجة أفضل، نستبعد بيانات هذه العينة من الحسابات فنحصل على النتائج التالية:
بعد حذف العينة رقم 19 يصبح عدد المحركات المعيبة 54 محركاً من بين 950 محركاً، أي أن متوسط نسبة المعيب في العينة يساوي إلى:

$$\bar{P} = \frac{54}{950} = 0.056842$$

ويمكن الحصول على متوسط نسبة الوحدات المعيبة باستخدام القانون التالي:

$$CL = \bar{P} = \frac{\sum_{i=1}^m P_i}{m} = \frac{1.08}{19} = 0.056842$$

0.04	2	50	5
0.04	2	50	6
0.02	1	50	7
0.1	5	50	8
0.08	4	50	9
0.04	2	50	10
0.04	2	50	11
0.06	3	50	12
0.08	4	50	13
0.06	3	50	14
0.04	2	50	15
0.04	2	50	16
0.04	2	50	17
0.06	3	50	18
0.32	16	50	19
0.06	3	50	20
1.4	70	1000	

الحل :

يتضح من الجدول وجود ٧٠ محركاً معيماً من بين 1000 محرك، أي أن متوسط نسبة المعيب في العينة المدروسة يساوي إلى:

$$\bar{P} = \frac{70}{1000} = 0.07$$

ويمكننا الحصول على متوسط نسبة الوحدات المعيبة باستخدام القانون التالي:

$$CL = \bar{P} = \frac{\sum_{i=1}^m P_i}{m} = \frac{1.4}{20} = 0.07$$

وحيث أن:

$$n = 50, \bar{P} = 0.07 \Rightarrow (1 - \bar{P}) = 0.93$$

فإن الانحراف المعياري للمتغير $\bar{P}$ يساوي:

$$S_{\bar{P}} = \sqrt{\frac{\bar{P}(1 - \bar{P})}{n}} = \sqrt{\frac{0.07(1 - 0.07)}{50}} = 0.036083$$

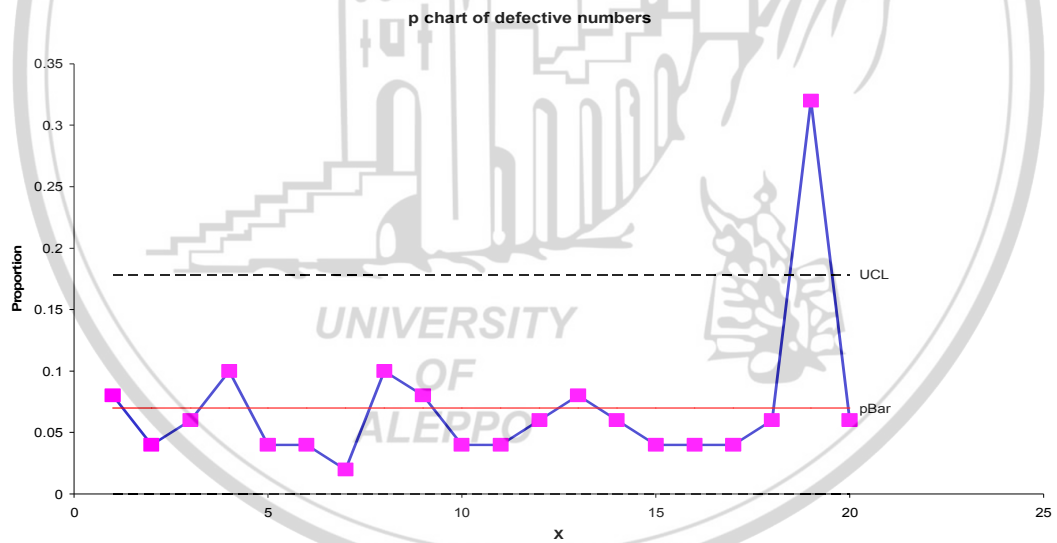
استناداً إلى ما تقدم فإن كل من حد الضبط الأعلى وحد الضبط الأدنى لنسبة المحركات المعيبة لأجهزة البرادات المنزلية في العينة يحسب كما يلي:

$$UCL_p = \bar{P} + 3\sqrt{\frac{\bar{P}(1 - \bar{P})}{n}} \\ = 0.07 + 3(0.036083) = 0.17825$$

$$LCL_p = \bar{P} - 3\sqrt{\frac{\bar{P}(1 - \bar{P})}{n}} \\ = 0.07 - 3(0.036083) = -0.03825$$

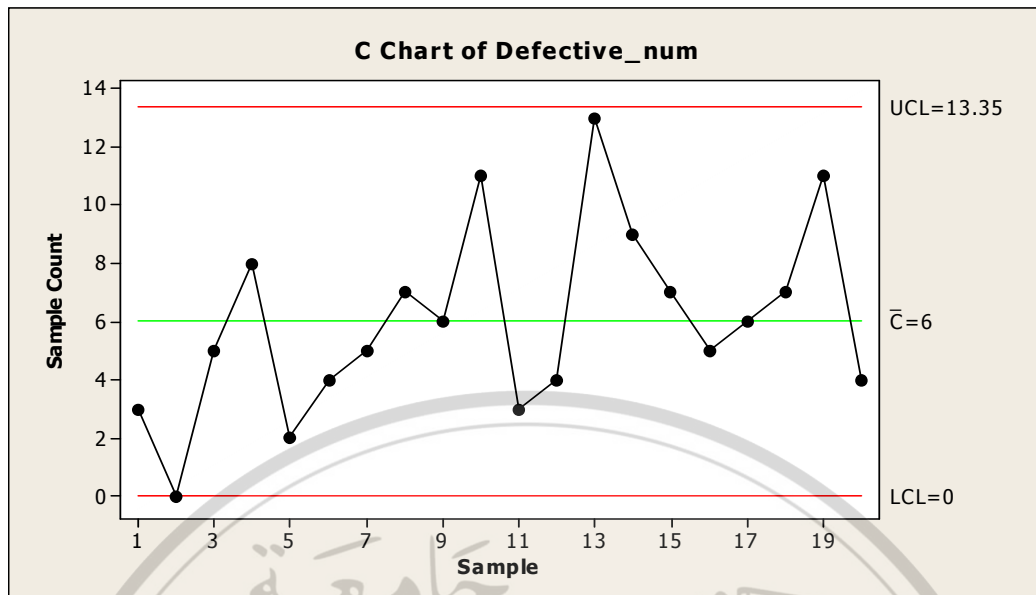
وبما أن حد الضبط الأدنى قيمة سالبة فإننا نضع هذه القيمة مساوية للصفر أي أن $LCL_{\bar{P}} = 0$.

3	50	3	0.06		0.07	0.17825
4	50	5	0.1		0.07	0.17825
5	50	2	0.04		0.07	0.17825
6	50	2	0.04		0.07	0.17825
7	50	1	0.02		0.07	0.17825
8	50	5	0.1		0.07	0.17825
9	50	4	0.08		0.07	0.17825
10	50	2	0.04		0.07	0.17825
11	50	2	0.04		0.07	0.17825
12	50	3	0.06		0.07	0.17825
13	50	4	0.08		0.07	0.17825
14	50	3	0.06		0.07	0.17825
15	50	2	0.04		0.07	0.17825
16	50	2	0.04		0.07	0.17825
17	50	2	0.04		0.07	0.17825
18	50	3	0.06		0.07	0.17825
19	50	16	0.32		0.07	0.17825
20	50	3	0.06		0.07	0.17825



٥-٤ خريطة عدد العيوب في الوحدة C – Chart

في الحياة العملية تواجهنا بعض الحالات لا نستطيع تحديد نسبة المعيب في العملية الإنتاجية، أو لا ينطبق على مخرجات العملية الإنتاجية مفهوم نسبة المعيب لعدم وجود عدد مناسب نعتمده كحجم للعينة المدروسة. فمثلاً قد يكون هدف قسم الرقابة التحكم في عدد الثقوب الصغيرة في كل متر من سلك كهربائي معزول، أو عدد العيوب الموجودة في المتر المربع الواحد من السجاد الوطني.



الشكل رقم (6) خريطة عدد العيوب في الوحدة C – Chart

§-٥ خرائط الضبط في SPSS

تسمح مخططات الضبط في SPSS للمستخدم الاختيار وتحديد نوع المخطط الذي يلائمه، وذلك بالنقر على الأيقونة الخاصة بنوع المخطط.

مثال (٧) :

مدير مصنع متخصص بصنع منتجات العناية بالشعر يحتاج لمراقبة مستوى الـ PH في الشامبو الذي ينتجه، إن المستوى الطبيعي للـ PH في الشعر هو حمضي نوعاً ما slightly acidic حيث تتراوح نسبة الحموضة ما بين 4.5 و 5.5 للمحافظة على قوة الشعر وحيويته. إن مستوى الـ PH في الشامبو والمنتجات الأخرى يجب أن يكون متوازناً، بهدف تأكيد وضمان جودة المنتجات تم قياس 6 دفعات منفصلة من الإنتاج على فترات زمنية منتظمة وتم تسجيل مستوى الـ PH في هذه المشاهدات.

جُمعت البيانات وتم إدخالها في محرر البيانات ببرنامج SPSS وتم تسمية الملف باسم:

[shampoo_ph.sav](#)

المطلوب:

إستخدام مخططي الوسط الحسابي والمدى X-bar and R charts لمراقبة العملية.

	time	batch	shift	ph
1	1	1	1	4.98
2	1	2	1	4.65
3	1	3	1	4.91

- إنشاء خريطة عدد العيوب في الوحدة $C - Chart$.

- هل العملية الإنتاجية محكمة (مضبوطة إحصائياً أم لا) ؟

عدد الثقوب الموجودة في كل متر من الأسلاك
الكهربائية المعزولة في عشرين عينة تضم كل منها متراً

العينة	عدد الثقوب	العينة	عدد الثقوب
1	3	11	3
2	0	12	4
3	5	13	13
4	8	14	9
5	2	15	7
6	4	16	5
7	5	17	6
8	7	18	7
9	6	19	11
10	11	20	4
المجموع	51	المجموع	69

الحل:

يتضح من الجدول أعلاه وجود 120 ثقباً صغيراً في العينة المكونة من 20 سلكاً كهربائياً معزولاً بطول متر لكل منها، أي أن متوسط عدد العيوب في كافة العينات يساوي إلى:

$$CL = \bar{C} = \frac{\sum_{i=1}^m C_i}{m} = \frac{C_1 + C_2 + \dots + C_m}{m} = \frac{120}{20} = 6$$

ولحساب حدود الضبط التجريبية نستخدم العلاقة:

$$UCL_C = \bar{C} + 3\sqrt{\bar{C}} = 6 + 3\sqrt{6} = 13.34847$$

$$LCL_C = \bar{C} - 3\sqrt{\bar{C}} = 6 - 3\sqrt{6} = -1.34847$$

وحيث أن حد الضبط الأدنى قيمة سالبة فإننا نضع هذه القيمة مساوية للصفر أي أن:

$$LCL_C = 0$$

يبين الشكل التالي رقم (6) حدي الضبط الأعلى والأدنى بالإضافة إلى عدد الوحدات المعيبة لكافة العينات، وحيث أن الانحرافات المعيارية لجميع العينات تقع في نطاق الضبط⁹، نستنتج أن العملية الإنتاجية محكمة (وبالتالي مضبوطة إحصائياً) .

⁹ باستثناء العينة رقم (٢) التي تقع تماماً على خط الضبط الأدنى، ووقوع هذه العينة على حد الضبط الأدنى يعني تماماً أن عدد العيوب في الوحدة المفحوصة هو معدوم وبالتالي فإن العملية الإنتاجية مضبوطة حكماً.

فإذا كان هدف الباحث أو المحلل التحكم في عدد العيوب الموجودة في كل قطعة من الوحدات المنتجة، وليس في نسبة الوحدات المعيبة، نجد من الضروري استخدام خريطة أخرى هي خريطة C ، وهي الخريطة التي نوقع عليها عدد العيوب الموجودة في كل وحدة من وحدات العينة المأخوذة من عملية إنتاجية معينة .

لرسم خريطة الضبط لعدد العيوب في الوحدة المفحوصة يؤخذ 25 عينة كحد أدنى من العملية الإنتاجية وبحجم وحدة واحدة فقط لكل عينة.

سوف نرمز لعدد العيوب الموجودة في كل وحدة من وحدات العينة بالرمز C^8 ، فإذا كانت العينة تحتوي على m من الوحدات ، وكانت : $C_1, C_2, \dots, C_m$ تمثل عدد العيوب في كل وحدة من هذه الوحدات .

يتم حساب وسط العملية الإنتاجية (متوسط عدد العيوب في كافة العينات) بالعلاقة:

$$CL = \bar{C} = \frac{\sum_{i=1}^m C_i}{m} = \frac{C_1 + C_2 + \dots + C_m}{m} \quad (12)$$

ومثلما رأينا في خرائط الوسط الحسابي والانحراف المعياري والمدى والنسبة، سيتم تمثيل خط المنتصف بوساطة الوسط الحسابي لعدد العيوب الموجودة في وحدات العينة، $CL = \bar{C}$ ، ويبعد كل حد من حدود الضبط بمقدار ثلاثة انحرافات معيارية عن خط المنتصف . وبهذه الحالة تحدد حدود الضبط التجريبية بموجب العلاقة:

$$UCL_C = \bar{C} + 3\sqrt{\bar{C}} \quad , \quad LCL_C = \bar{C} - 3\sqrt{\bar{C}} \quad (13)$$

حيث UCL_C ، LCL_C هما الحد الأعلى والخريطة عدد العيوب في الوحدة.

$\sqrt{\bar{C}}$ الانحراف المعياري لعدد العيوب في كافة العينات.

مثال (٦):

تستخدم عملية إنتاجية معينة في إنتاج الأسلاك الكهربائية المعزولة، ولضبط جودة الأسلاك الكهربائية المعزولة المنتجة، قام مهندس ضبط جودة الإنتاج باختيار 20 متراً بطريقة عشوائية وسجل عدد الثقوب في كل متر منها، وذلك كما هو موضح بالجدول

والمطلوب:

⁸ إن عدد العيوب الموجودة في كل وحدة من منتج معين يتبع تقريباً توزيع بواسون، حيث يتساوى الوسط الحسابي لهذا التوزيع مع تباينه ، وغالباً يكون تباين المجتمع ووسطه الحسابي مجهولين .

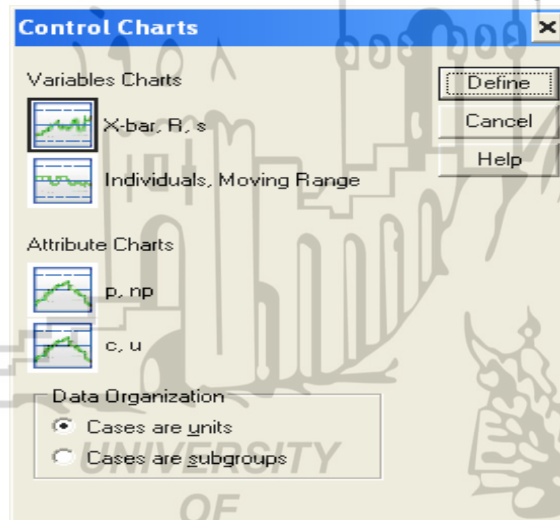
4	1	4	1	4.62
5	1	5	1	4.84
6	1	6	1	4.69
7	2	1	1	4.52
8	2	2	1	4.73
9	2	3	1	5.05
.....				
.....				
239	40	5	2	5.16
240	40	6	2	4.76

الحل:

لاستخدام تقنية مخططات الضبط ، نختار من شريط القوائم :

Analyze → Quality Control → Control Charts

يظهر صندوق حوار بمسمى مخططات الضبط Control Charts .



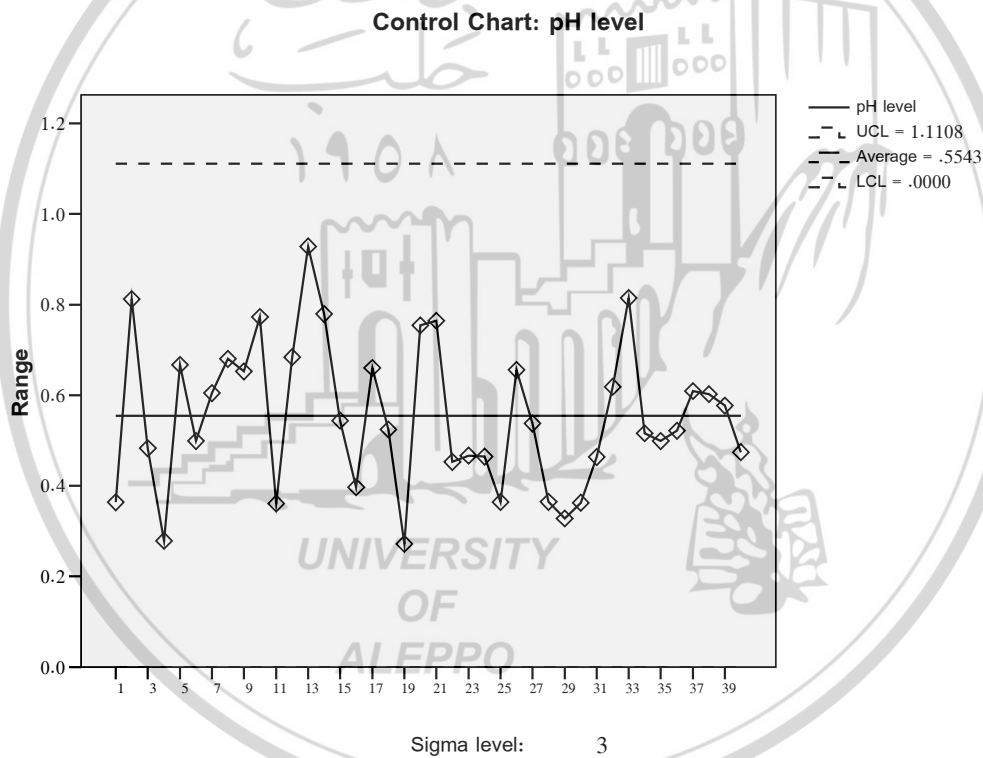
نختار المربع المعنون X-bar, R , s ، ومن تنظيم البيانات Data Organization نختار

Cases are Units ثم نضغط على Define، يظهر صندوق حوار جديد بعنوان: X-bar, R ,

s: Cases are Units

إن خريطة المدى تبدو محكمة ومضبوطة إذ إن جميع النقاط هي ضمن حدي الضبط، في الطرف الأيمن من خريطة المدى يمكن أن تكون هناك بدايات لمشكلة مع تغير طفيف جداً ولكن لا يمكننا أن نؤكد ذلك بدون الحصول على قياسات إضافية.

إذا كانت خريطة المدى تشير إلى أن تغيرات العملية هي خارج المراقبة أو الضبط، عندها يمكنك غض النظر أو تجاهل خريطة الوسط الحسابي، إذ إن متوسط مدى العينات يستخدم في حساب حدي الضبط لخريطة الوسط الحسابي، وعليه إذا كانت خريطة المدى تشير بأن العملية غير مضبوطة، فإن حدي الضبط لخريطة الوسط سوف تثير التساؤلات، لكن في حال أشارت خريطة المدى إلى أن العملية تحت المراقبة أي مضبوطة، يمكنك بسهولة فحص خريطة الوسط الحسابي.



تستخدم خريطة الوسط الحسابي الأوساط الحسابية للعينات لمراقبة ورصد التغيرات في العملية، يتم تمثيل الأوساط الحسابية للعينات على أنها نقاط انتشار متصلة بخطوط مستقيمة. الخط المستقيم الغامق في الوسط يمثل الوسط الحسابي لأوساط العينات، الخطان المنقطان يمثلان حدي المراقبة الأعلى (UCL) والأدنى (LCL).

إن شكل ولون النقاط يشير إلى أن نقطة ما تخترق أو تنتهك violates حدود المراقبة أو حدود الضبط المحددة. الوسط الحسابي العام يساوي 4.9915، إلا أن الشكل البياني لخريطة

- نختار CP, CpU, CpL, K, CpM, and Z-out من مجموعة مؤشرات مقدرة العملية .
- نختار PP, PpU, PpL, PpM, and Z-out من مجموعة مؤشرات أداء العملية .
- نختار النسبة المئوية الفعلية خارج حدود المواصفة Actual % outside specification limits.
- نختار كل قواعد الضبط All control rules.



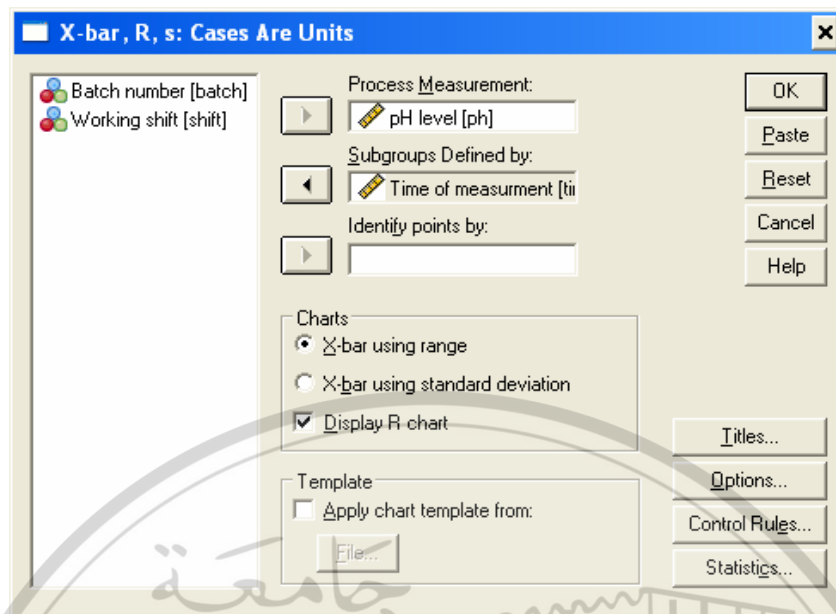
ننقر على الزر Continue.

نعود إلى صندوق الحوار السابق X-bar, R, s: Cases are Units
نضغط OK نحصل على :

- مخطط المدى range chart.
- مخطط الوسط الحسابي X-bar chart.
- جدول اختراق القواعد a rule violations table.
- جدول إحصاءات العملية a process statistics table.

نفحص أولاً خريطة المدى R chart حيث تستخدم هذه الخريطة مدى العينات لمراقبة ورصد التغيرات على امتداد العملية، يتم رسم مدى العينات على أنها نقاط إنتشار متصلة بخطوط مستقيمة. الخط المستقيم الغامق في الوسط يمثل الوسط الحسابي لمدى العينات، الخطان المنقطان يمثلان حدي المراقبة الأعلى (UCL) والأدنى (LCL)، مع ملاحظة أن قيمة حد المراقبة الأدنى يساوي الصفر، لذلك لا يبدو واضحاً أو مرئياً بسبب التصاقه بالمحور الأفقي.

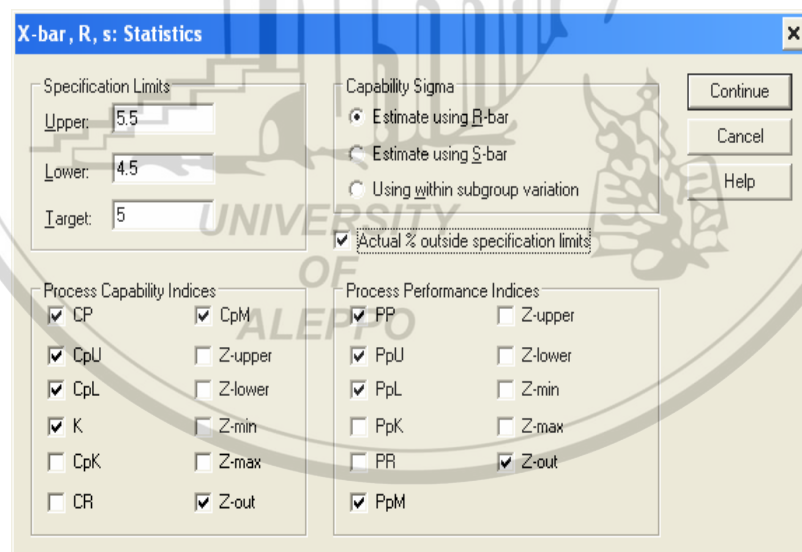
حدود المراقبة أو الضبط حُسبت بطريقة إذا كانت العملية مضبوطة أو تحت المراقبة، إذا كانت كل نقاط العينة تقريباً واقعة بين الحدين.



نختار pH level كمتغير قياسات العملية ونضعه في النافذة المعنونة قياسات العملية Process Measurement.

نختار زمن القياسات Time of measurement ونضعه في النافذة المعنونة المجموعات الجزئية معرفة بواسطة subgroups Defined by.

ننقر على الزر Statistics ، يظهر صندوق حوار بعنوان X-bar, R, s: Statistics

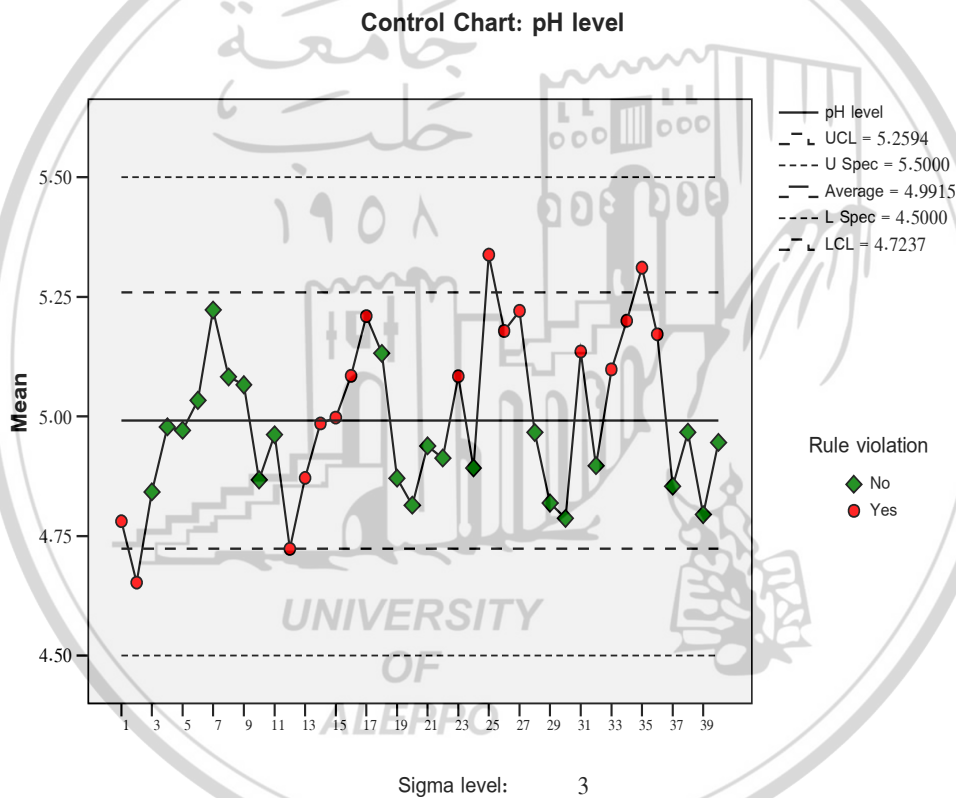


- نضع الرقم 5.5 بمنزلة الحد الأعلى للمواصفة The upper specification limit (USL)

- نضع الرقم 4.5 بمنزلة الحد الأدنى للمواصفة The lower specification limit (LSL)

- نضع الرقم 5 بمنزلة القيمة المستهدفة The target value.

الضبط توضح أن العملية خارج التحكم. حيث توجد نقاط خارج حدي الضبط وهذه النقاط هي التي تم تسجيلها خلال الفترات الزمنية 2 , 25 , 35 . هناك اتجاه للتزايد تمثله النقاط التي تم تسجيلها خلال الفترات الزمنية من 12 إلى 17 ، بالإضافة إلى ذلك هناك عدة اتجاهات على وشك أن تأخذ شكل نموذج دوري، من المؤكد أن في العملية نقاط ذروة أو نتوءات خلال الفترات الزمنية 35, 25 , 17 , 7 . ونقاط تنجبه نحو الأسفل خلال الفترات الزمنية 30 , 20 , 12 , 2 بالعودة الى الفترات الزمنية التي تم فيها تسجيل النقاط نجد أن نقاط الانخفاض حدثت خلال الوردية الأولى بينما نقاط الذروة حدثت خلال الوردية الثالثة.



يوضح الجدول التالي النقاط التي تنتهك قواعد الضبط.

زمن القياس Time of measurement: هو متغير يحدد ويوصف النقاط، وبشكل تلقائي يعرف المجموعات الجزئية ضمن مخطط الضبط.

النقاط التي تنتهك قواعد الضبط Violation for Points: هي النقاط التي تنتهك قواعد الضبط المحددة.

بالإجمال هناك ١٧ نقطة إنتهكت قواعد الضبط.

المستهدفة وهي 5، إلا أن متوسط العملية يظهر تذبذباً كبيراً. بالاستناد الى ما أظهرته خريطة الوسط الحسابي فإن الخطوة التالية يجب أن تتركز حول فحص الفروقات في مستويات الـ PH في الشامبو المنتج خلال الورديات المختلفة.

مثال (٨):

مدير مصنع متخصص بصنع الأقمشة كلف مراقب الجودة بمراقبة عملية التصنيع من خلال كل دفعة يتم إنتاجها في المصنع، قام المراقب بأخذ عينات من الأقمشة المصنعة وسجل عدد القطع غير المقبولة، جُمعت البيانات وتم إدخالها في محرر البيانات ببرنامج SPSS وتم تسمية الملف باسم:

[.clothing_defects.sav](#)

المطلوب:

استخدام مخططات الضبط لمراقبة نسبة القطع المعيبة، أي رسم مخطط نسبة القطع المعيبة p chart.

lot	sampled	defects
1	193	6
2	198	1
3	211	1
4	210	0
5	204	2
6	214	1
7	208	1
8	210	3
9	198	2
10	174	2
11	195	1
12	192	6
13	162	1
14	197	2
15	213	3
16	220	3
17	198	0
18	220	1
19	212	1
20	215	1
21	201	0
22	210	5
23	186	0
24	190	4
25	193	0
26	204	5
27	206	2
28	195	2
29	182	3

$CP = 0.672$: هذا دليل على أن التذبذب أو التغيرات في مستوى الـ pH في العينات التي تم تسجيلها تتجاوز المواصفات.

CpL and CpU : تساعد الباحث أو الفاحص بما إذا كانت تغيرات العملية متناظرة، إذا كانت قيمهما قريبة من قيمة CP ، عندها نقول بأن تغيرات العملية متناظرة نسبياً.
K : هي مقياس لمقدرة متوسط العملية بأن يقترب من خط الوسط الفاصل بين حدي المواصفة، وكلما كانت قيمتها صغيرة مع قيمة كبيرة لـ CP ، ستكون إشارة إلى كون العملية ضمن حدود المواصفات.

CpM : تقيس مقدرة العملية process capability لتبقى ضمن الهدف المحدد، إذا كانت قيمتها قريبة من قيمة CP كما في مثالنا فالعملية تبقى ضمن الهدف المحدد.
يوضح الجدول التالي مؤشرات المقدرة والأداء capability and performance indices.

Process Statistics		
Capability Indices	CP ^a	.762
	CpL ^a	.762
	CpU ^a	.749
	K	.017
Performance Indices	CpM ^{a,b}	.761
	PP	.652
	PpL	.641
	PpU	.663
	PpM ^b	.652
	Z-upper	1.990

The normal distribution is assumed. LSL = 4.5 and USL = 5.5.

a. The estimated capability sigma is based on the mean of the sample group ranges.

b. The target value is 5.

٢-٥ مؤشرات الأداء :Performance indices

هذه المقاييس مماثلة لمؤشرات مقدرة العملية، إلا أنها تتجه لتكون أكثر تحفظاً لأنها مبنية على تغيرات العملية بأكملها وليس على تغيرات العينة.
إذا كانت قيم مؤشرات الأداء أصغر بكثير من قيم مؤشرات مقدرة العملية، مما يشير على الأغلب إلى وجود تغيرات غير عشوائية في العملية، في مثالنا لا تبدو قيم مؤشرات الأداء أصغر بكثير من قيم مؤشرات مقدرة العملية.

باستخدام خريطتي المدى والوسط الحسابي، توصلنا لنتيجة مفادها بأن عملية إنتاج الشامبو هي خارج التحكم أي غير مضبوطة إحصائياً بشكل وسطي، إن متوسط العملية هو قريب من القيمة

Rule Violations

Time of measurment	Violations for Points
1	2 points out of the last 3 below -2 sigma
2	Less than -3 sigma
2	2 points out of the last 3 below -2 sigma
12	Less than -3 sigma
12	6 points in a row trending up
13	6 points in a row trending up
14	6 points in a row trending up
15	6 points in a row trending up
16	6 points in a row trending up
17	6 points in a row trending up
23	4 points out of the last 5 above +1 sigma
25	Greater than +3 sigma
25	2 points out of the last 3 above +2 sigma
25	4 points out of the last 5 above +1 sigma
26	2 points out of the last 3 above +2 sigma
26	4 points out of the last 5 above +1 sigma
27	2 points out of the last 3 above +2 sigma
27	4 points out of the last 5 above +1 sigma
31	4 points out of the last 5 above +1 sigma
33	4 points out of the last 5 above +1 sigma
34	2 points out of the last 3 above +2 sigma
34	4 points out of the last 5 above +1 sigma
35	Greater than +3 sigma
35	2 points out of the last 3 above +2 sigma
35	4 points out of the last 5 above +1 sigma
36	2 points out of the last 3 above +2 sigma
36	4 points out of the last 5 above +1 sigma

17 points violate control rules.

١-٥ مؤشرات المقدرة Capability indices:

هذه المقاييس تساعد أو تقيم قدرة العملية على أن تحقق المواصفات، إذ إن كل مؤشر يقيس أوجه مختلفة للعملية.

CP : عبارة عن نسبة الفرق بين حدي المواصفة وتغيرات العملية المشاهدة.

$CP \geq 1$: هذا دليل على أن مقدرة العملية تحقق المواصفات المطلوبة .

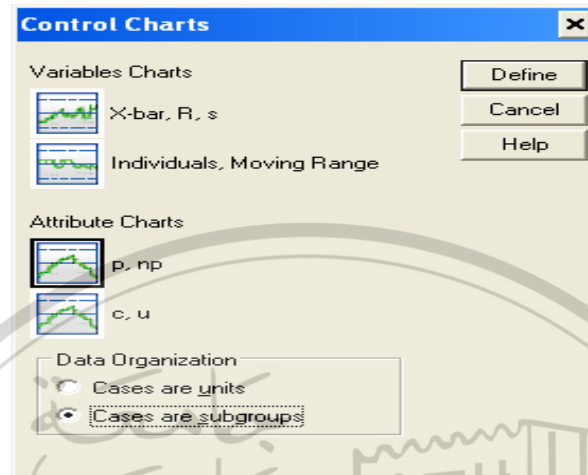
$CP < 1$: هذا دليل على تذبذب العملية.

30	189	2

الحل: لاستخدام تقنية مخططات الضبط ، نختار من شريط القوائم :

Analyze → Quality Control → Control Charts..

يظهر صندوق حوار بمسمى مخططات الضبط Control Charts .

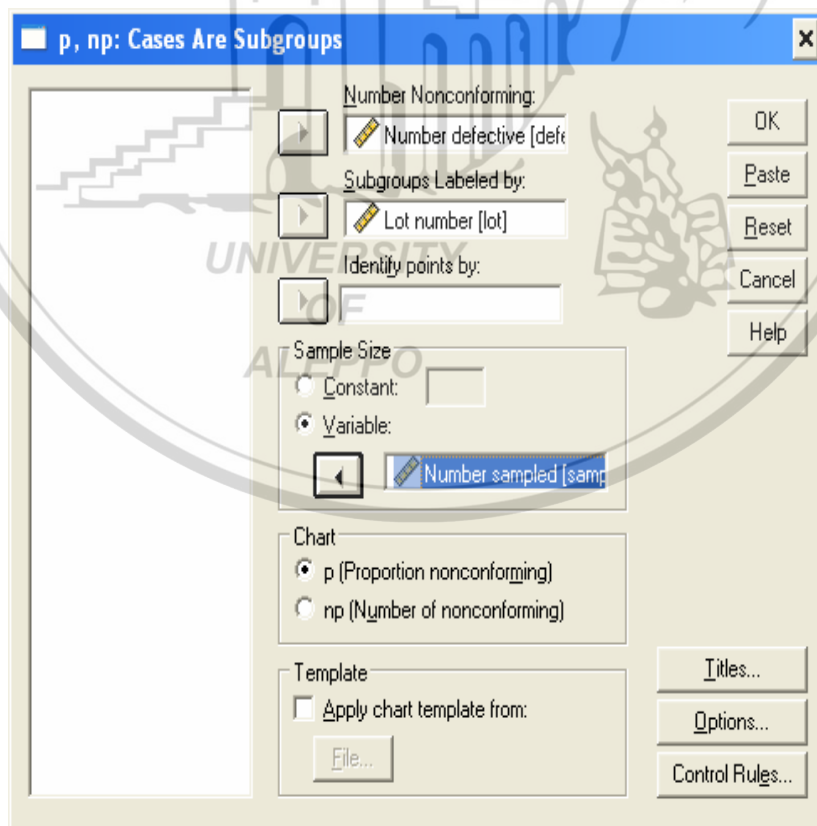


- نختار المربع المعنون p, np .

- من تنظيم البيانات Data Organization نختار Cases are subgroups - نضغط على Define .

يظهر صندوق حوار جديد بعنوان:

p, np: Cases are subgroups



أسئلة وتمارين

١ - تستخدم إحدى العمليات الإنتاجية في إنتاج جزء معين يقاس طوله بالسنتيمتر. سحبت عينة عشوائية مكونة من خمسة أجزاء في كل عشر دقائق وقيست أطوالها لأقرب سنتيمتر. والجدول التالي يمثل البيانات التي تم الحصول عليها لعشرة من هذه العينات.

الفترة الزمنية	القياسات				
	1	2	3	4	5
1	13	16	17	15	14
2	17	16	9	14	9
3	12	18	16	15	13
4	11	13	15	11	16
5	10	9	16	15	15
6	12	15	13	14	11
7	17	14	11	15	18
8	16	16	18	12	13
9	14	12	11	15	13
10	16	15	11	15	18

والمطلوب:

- رسم خريطة الوسط الحسابي وخريطة المدى ووقع البيانات على كل منهما.
 - هل تدل خريطة الوسط الحسابي على أن العملية الإنتاجية محكمة؟
 - ما هو قرارك بالنسبة لخريطة المدى؟
- ٢ - لنفترض أن إحدى الشركات الوطنية تقوم بتعبئة معجون الأسنان في عبوات بلاستيكية، يفترض أن تحتوي كل منها على 200 غرام من هذا المعجون. سحبت ثماني عينات عشوائية تحتوي كل منها على خمس عبوات وسجل الوزن في الجدول:

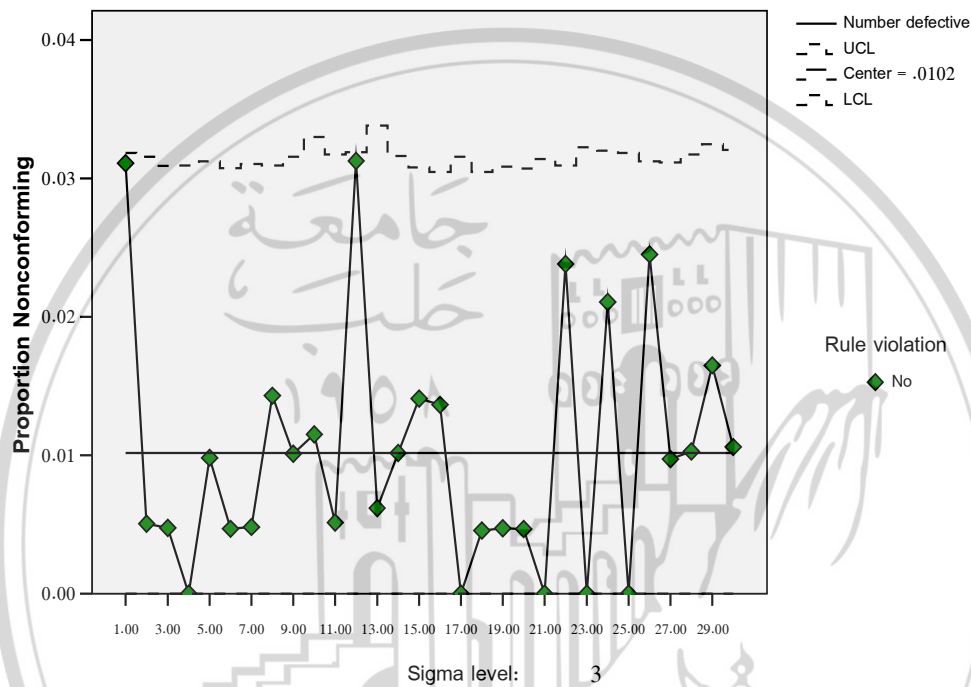
العينة	الوزن بالغرام				
	1	2	3	4	5
1	197	200	192	201	198
2	193	205	200	197	202
3	205	192	204	203	204
4	209	192	198	201	205
5	200	195	208	203	212
6	210	205	196	197	202
7	203	198	198	199	194
8	191	196	204	203	196

والمطلوب:

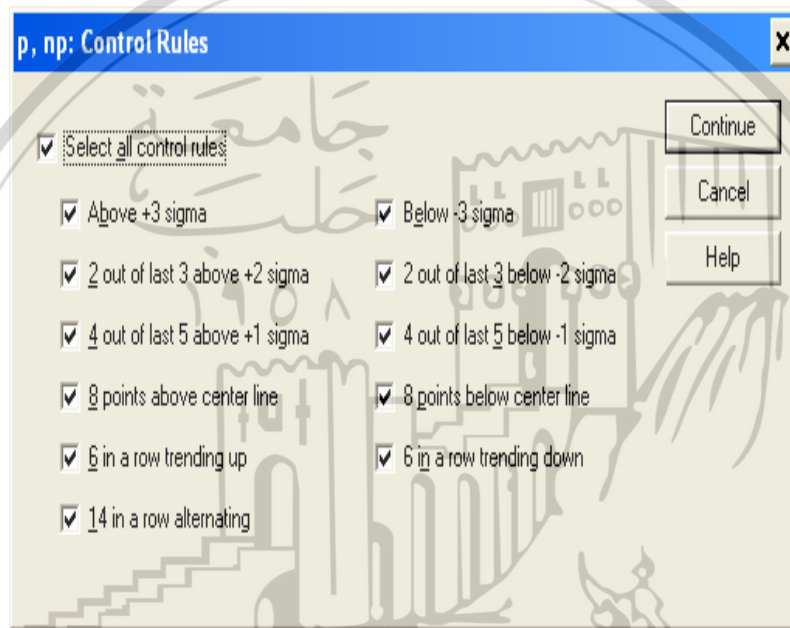
- رسم خريطتي الوسط الحسابي والمدى ووقع البيانات على كل منهما
- هل تدل خريطة الوسط الحسابي على أن العملية الإنتاجية محكمة؟

إن حدي الضبط في الخريطة يعكسان كمية التغيرات المتوقعة في عدد القطع المعيبة فقط في حالة كون أن اسباب التغير هي شائعة common causes .
 إن خريطة النسبة تبدو جيدة ومحكمة حيث أن جميع النقاط هي ضمن حدي الضبط.

Control Chart: Number defective



- نختار Number defective كمتغير يمثل عدد القطع غير المطابقة للمواصفات ونضعه في النافذة المعنونة عدد غير المطابقات number Nonconforming.
- نختار عدد الدفعات Lot number ونضعه في النافذة المعنونة المجموعات الجزئية معنونة بواسطة subgroups Labeled by .
- نختار عدد العينات التي تم فحصها Number sampled ونضعه في النافذة المعنونة حجم العينات sample sizes .
- نضغط الزر قواعد الضبط Control Rules ، ثم نختار كل القواعد .



ننقر على الزر Continue ، نعود الى صندوق الحوار السابق المعنون:

. p, np: Cases are Subgroups

نضغط OK نحصل على :

- مخطط نسبة القطع المعيبة p chart .

تستخدم هذه الخريطة لمراقبة ورصد التغيرات في نسبة القطع المعيبة أو عدد الوحدات غير المطابقة للمواصفات، يتم تمثيل نسبة القطع المعيبة كنقاط انتشار متصلة بخطوط مستقيمة. الخط المستقيم الغامق في الوسط يمثل الوسط الحسابي للنسب في العينات، الخطان المنقطان يمثلان حدي المراقبة الأعلى (UCL) والأدنى (LCL)، مع ملاحظة أن قيمة حد المراقبة الأعلى يتغير لأن حجم العينة يتغير خلال فترات جمع المعلومات. إن قيمة حد الضبط الأدنى قريبة من الصفر، لذلك لا يبدو واضحاً أو مرئياً بسبب التصاقه بالمحور الأفقي.

- ما هو قرارك بالنسبة لخريطة المدى ؟ اشرح إجابتك .

٣- يعتقد مراقب ضبط الجودة في مصنع سكر مسكنة بأن العملية الإنتاجية تسير وفقاً للمواصفات المحددة . ولهذا السبب قام المهندس المسؤول عن ضبط جودة الإنتاج باختيار عينة عشوائية مكونة من تسعة أكياس من السكر في كل ساعة خلال أربعين ساعة وسجل الوزن لكل منها فوجد أن الوسط الحسابي العام لوزن الكيس يساوي ٤,٥ كيلو غرام بخطأ معياري قدره 0.2 كغ ، كما وجد أن الوسط الحسابي لمدى الأربعين عينة يساوي 0.25 كيلو غرام بانحراف معياري قدره 0.05 كيلو غراماً .

والمطلوب:

- إنشاء خريطة الوسط الحسابي والمدى باستخدام هذه البيانات.
٤- لنفترض أننا سحبنا عشر عينات جديدة تتكون كل منها من تسعة أكياس. والجدول التالي يعطي الوسط الحسابي والمدى لكل عينة من هذه العينات الجديدة:

والمطلوب:

هل تدل خريطة الوسط الحسابي على أن العملية الإنتاجية محكمة ؟ ما هو قرارك بالنسبة لخريطة المدى ؟

- اشرح إجابتك .

المدى	الوسط الحسابي	الساعة
0.22	4.8	1
0.3	3.5	2
0.28	4.2	3
0.22	4.8	4
0.44	5	5
0.24	4.6	6
0.4	5.4	7
0.32	4.8	8
0.22	4.6	9
0.2	4.4	10

Block diagrams
Box-Cox power transformation
Brainstorming Process

رسومات المربعات
تحويلة بوكس – كوكس الأسية
عملية عصف الأفكار

C

Calculator Pad
Cases
Central Tendency
Charts
Chat Type
Chi-square Statistic
Chi-Square Test
Classify
Cochran's Q Test
Coefficient Contingency
Coefficient of Association
Coefficient Phi
Collinearity Statistics Tolerance
Columns & Align
Compare Means
Compare Means
Condition Index
Confidence Interval
Controlled experiments
Correlation
Correlation and Regression Analysis
Counts
Cramer Coefficient C
Cross-Tabulation
Cumulative Percent
Custom Currency
Capability analysis
Capability Analysis for Nonnormal
Data
capability and performance indices
Capability Index Cp
Capability Studies
Cause and Effect Diagram
Check sheets
Collection of Data

الألة الحاسبة
حالات
مقاييس النزعة المركزية
الأشكال البيانية
نوع التخطيط
إحصائية الاختبار كاي تربيع
اختبار كاي تربيع
التصنيف
اختبار كوكران Q
معامل التوافق (التصاحب)
معامل الاقتران الرباعي
معامل ارتباط فاي
إحصائيات الارتداد الخطي والتحمل
لتحديد عرض العمود وموقع البيانات
مقارنة المتوسطات
مقارنة المتوسطات
دليل الشرط
مجال ثقة لمتوسط المجتمع
التجارب والاختبارات المصممة
الارتباط
تحليل الارتباط والانحدار
عدد القيم
معامل كرامر
الجدول التقاطعية
النسبة المئوية التجميعية
عملة دولة معينة
تحليل المقدرة
تحليل المقدرة لبيانات لا تتبع التوزيع
الطبيعي
مؤشرات المقدرة والآداء
مؤشر (دليل) المقدرة
دراسات المقدرة
مخطط السبب والأثر
قوائم الفحص
جمع البيانات

دليل المصطلحات العلمية

إنكليزي - عربي

A

Analysis of Frequencies	تحليل التكرارات
Analyze	تحليل
table and eta Anova	تحليل التباين أحادي التصنيف ومعامل إيتا
Functions Arithmetic	التوابع الحسابية
means Ascending	الترتيب التصاعدي للأوساط الحسابية
of Linearity Assumption	فرض الخطية
Underlying Simple Assumptions	فرضيات الانحدار الخطي البسيط
Linear Regression	ارتباط ذاتي
Autocorrelation	كل قواعد الضبط
All control rules	
American Society for Quality (ASQ)	الجمعية الأمريكية للجودة
American Society for Quality Control	الجمعية الأمريكية للتحكم بالجودة
Analyze	تحليل
Anderson-Darling (AD)	إحصائية الاختبار لأندرسون - دارلينغ
Appraisal Costs	تكاليف التقييم
Arithmetic Mean	الوسط الحسابي
artificial subgroups	المجموعات الجزئية الاصطناعية
Assumptions of Process Capability	فرضيات مقدرة العملية
ATM	الصراف الآلي
AUTOMATE A	مكنة الإجراءات
Average of subgroup ranges	متوسط المدى للمجموعات الجزئية
Average of subgroup standard deviations	متوسط الانحرافات المعيارية للمجموعات الجزئية

B

Charts Bar	الأعمدة البيانية
Between Groups	بين المجموعات
Boxplots	مخطط صندوقي
down variable Break	متغير تجزئة
Blank cause-and-effect diagram	إنشاء مخطط فارغ للسبب والآخر
Blank cause-and-effect diagram with empty branches	إنشاء مخطط السبب والآخر مع أغصان فارغة

- يعتقد مدير الشركة العامة للصناعات الجلدية (قسم تصنيع الأحذية) بوجود نسبة مرتفعة من النعال المعيبة، فقام مسؤول ضبط الجودة في المصنع باختيار عينة عشوائية في كل ساعة خلال عشرين ساعة وسجل عدد النعال المعيبة في كل عينة منها في الجدول التالي:

الساعة	حجم العينة	عدد الوحدات المعيبة	الساعة	حجم العينة	عدد الوحدات المعيبة
1	50	8	11	50	6
2	50	4	12	50	5
3	50	8	13	50	4
4	50	7	14	50	5
5	50	6	15	50	4
6	50	7	16	50	3
7	50	3	17	50	5
8	50	5	18	50	4
9	50	4	19	50	6
10	50	6	20	50	4

والمطلوب:

- إنشاء خريطة نسبة المعيب باستخدام هذه البيانات.

- هل العملية الإنتاجية محكمة ؟

- اشرح إجابتك.

Control Charts
Control Charts for Attributes
Control Charts for Variables
Control graph
Control matrices
Control Process
Control sheet

مخططات الضبط
مخططات الضبط للمواصفات
مخططات الضبط للمتغيرات
مخطط الضبط
مصفوفات المراقبة
ضبط العملية
ورقة المراقبة

D

Data Editor
DATA Files
Data Transformation
Data View
Date and Time Functions
Decimal
Degrees of freedom
Descending means
Descriptive Statistics
Dispersion
Display frequency tables
Distribution Functions
Domains
Double Click
Data Organization
Descriptive statistics
Design of Experiments
Dispersion diagram
Distribution identification

محرر البيانات
ملف البيانات
تحويل البيانات
عرض البيانات
توابع الوقت والتاريخ
لتحديد عدد الخانات العشرية
درجات الحرية
الترتيب التنازلي للأوساط الحسابية
الإحصاءات الوصفية
التشتت
عرض جدول التوزيع التكراري
توابع التوزيع
المجالات
النقر المزدوج
تنظيم البيانات
الإحصاءات الوصفية
تصميم التجارب
شكل الانتشار
توصيف (مطابقة) التوزيع

E

Eigenvalues
Variables Elementary
Enter
Estimate
Exclude Cases
Frequency Expected
experimental units
Explore
Extreme Values

الجزور المميزة
المتغيرات الأولية
ادخال كافة المتغيرات المستقلة
لتقدير معالم نموذج الانحدار
استبعاد الحالات
التكرارات المتوقعة
وحدات التجربة
استكشاف البيانات
القيم المتطرفة

Means Plot	التمثيل البياني لمتوسط المعالجات
Measure	وحدة القياس
Median	الوسيط
M-estimators	المقدرات الموزونة
Microsoft Excel	برنامج الجداول الإلكترونية
Minimum	أقل قيمة
Missing Value Functions	توابع القيم المفقودة
Missing Values	القيم المفقودة
mode	المنوال
Multicollinearity	الازدواج الخطي
Multiple Category Method	أسلوب الفئات المتعددة
Multiple Dichotomy Method	أسلوب المتغيرات المتعددة ذات الفئتين
Multiple Response Analysis	تحليل الاستجابات المتعددة
Multiple Response Crosstabs	جداول التقاطع متعددة الإستجابة
Multiple Response Frequencies	التكرارات متعددة الاستجابة
Multiple Responses	الاستجابات المتعددة
Manufacturing System	النظام الصناعي (نظام التصنيع)
Manufacturing Variations	التغيرات التصنيعية
Mass Production	إنتاج كمي
<i>Measures of Central Tendency</i>	مقاييس النزعة المركزية والتشتت
Measures of Dispersion	مقاييس التشتت
Median of moving range	وسيط المدى المتحرك
Moving Range Chart	مخطط المدى المتحرك
Position & Name	اسم المتغير وموقعه المتسلسل
Nominal	اسمية
Nonparametric Methods	أساليب لامعلمية
Tests Non-Parametric	الاختبارات اللامعلمية
Non-Parametric Tests	الاختبارات اللامعلمية
Normality Plots with Tests	اختبارات التوزيع الطبيعي
Plots with Tests Normality of respondent number	مخططات لاختبار التوزيع الطبيعي
Numeric	رقم المستجوب
Expression Numeric	متغير رقمي
number Nonconforming	الصيغة الرقمية الحسابية
Number sampled	عدد الوحدات غير المطابقة
	عدد العينات التي تم فحصها

Interaction	تأثيرات متبادلة
Interpolation	الاستكمال الخطي
Improvement	تحسين
Indicator	مؤشر
Individual Distribution Identification	اختيار التوزيع الأمثل للبيانات
Individuals Charts	مخططات القيم الفردية
Inspection and Testing	الفحص والاختبار
Inspection By Attributes	الفحص بالموصفات
Inspection By Variables	الفحص بالمتغيرات
International Quality Standards	المواصفات القياسية الدولية للجودة
International Standardization	
Organization	منظمة المواصفات العالمية
Ishikawa diagram	مخطط إيشيكوا

J

K- Sample Inference: Related Samples	الاستدلال حول عدد من العينات غير المستقلة
Kolmogorov-Smirnov	اختبار كولموغوروف سميرنوف
Kruskal – Wallis test	اختبار كروسكال واليز
Kurtosis	معامل التفرطح للتوزيع
Just-in-time	فلسفة الوقت المحدد

L

Label	وصف المتغير
Layer 2	الطبقة الثانية
Least-squares line	خط المربعات الصغرى
List of variables on the working file	قائمة بالمتغيرات الموجودة في ملف العمل
Functions Logical	التوابع المنطقية
Loglinear	النماذج الخطية اللوغاريتمية
Lots number	عدد الدفعات
Lower Control Limit	الحد الأدنى للضبط

M

McNemar's Test	اختبار ماكنمار
Mean	الوسط الحسابي
Mean squares	متوسط المربعات
Means Analysis	تحليل المتوسطات

Expected " overall" performance
External Failure Costs

مؤشرات الأداء المتوقعة الكلية
تكاليف الفشل الخارجي

F

ratio F
Flag Significance Correlation
Friedman Test (MINITAB Outputs
Friedman Two – Way Analysis of
Variance by Ranks Test
Functions
Facilities
Flow Charts
Fundamental of Statistics

قيمة إحصائية الاختبار لفيشر
تأشير الارتباطات المعنوية بنجمة
اختبار فريدمان مخرجات برنامج
اختبار فريدمان لتحليل التباين للرتب
في اتجاهين
التتابع
التسهيلات
خرائط إنسيابية
أساسيات الإحصاء

G

Linear Model General
Grouping Variable
Geometric Mean
goodness-of-fit test results
goodness-of-fit tests
Graphical Methods

النماذج الخطية العامة
متغير التجميع (متغير التجزئة)
الوسط الهندسي
نتائج اختبار جودة التوفيق
اختبارات جودة التوفيق
طرق بيانية

H

with normal curve Histogram
Histograms
Homogeneity – of- variance
Homogeneity – of- variance
Homoscedasticity
Hspread
Heterogeneity
Highly Capable
Hypergeometric Distribution
Hypothesis Testing

المدرج التكراري مع المنحنى
الطبيعي
المدرج التكراري
اختبار تجانس التباين
تجانس التباين
تجانس تباين الخطأ العشوائي
طول الصندوق
عدم التجانس
العملية ذات مقدرة عالية
التوزيع فوق الهندسي
اختبار الفرضيات

I

T Test Independent-Samples
Quartile Range Inter

اختبار T للفرق بين عينتين مستقلتين
المدى الربيعي

O

Observed Frequency	التكرارات المشاهدة
One – Way Analysis of Variance	تحليل التباين أحادي التصنيف
One Sample T-Test	اختبار T لعينة واحدة
One Tailed test	اختبار الفرضية من طرف واحد
Options	خيارات
Ordinal	ترتيبية
Outliers	القيم الشاذة
SPSS Viewer – Output	نافذة المخرجات
File Output	ملف المخرجات الاحصائية
Operator	عامل تشغيل
Optimal lambda	قيمة لمبدأ المثلى
Output Quality	جودة المخرجات
Outside specification limits	خارج حدود المواصفة
Overall capability	المقدرة الكلية
P	
Paired-Samples T Test	اختبار T للملاحظات المزدوجة
Parametric Methods	أساليب معلمية
Partial Correlation	معامل الارتباط الجزئي
Partial Correlation	الارتباط الجزئي
Pearson	معامل الارتباط الخطي لبيرسون
Percent	النسبة المئوية
Percentiles	المئينات
Percentiles Values	قيم المئينات
Pie Charts	الدوائر البيانية
principal component	المركبة الأساسية
Producing a Test of Mean Differences	اختبار الفرق بين متوسطين
Producing Bivariate Tables	إنشاء جدول بمتغيرين
Producing Confidence Intervals Around a Mean	مجال ثقة حول المتوسط
Pareto Chart	مخطط باريتو
Pareto chart with a by variable	مخطط باريتو للمتغيرات
Pareto chart with raw data	مخطط باريتو للبيانات الخام
Pareto graph	رسم باريتو البياني
Parts per million (ppm)	عدد القطع المعيبة من المليون
Pay-By-Phone	الدفع بوساطة الهاتف
percentage graphs	رسومات النسب المئوية

Sampling Distributions	توزيعات المعاينة
Sampling Error	أخطاء المعاينة
Sampling or measurement problems	مشاكل المعاينة أو القياس
Saving Account	حساب الإدخار
Scatter Diagram	خريطة التبعثر (شكل الانتشار)
six-sigma program	برنامج الستة سيغما
specification spread	مدى المواصفة
Standard Deviation	الانحراف المعياري
Statistical Process Control Charts	خرائط التحكم للعمليات
Statistical Quality Control Tools	أدوات ضبط الإحصائي للجودة
Successes and Failures	نجاح وفشل

Target Variable	المتغير الهدف
Target variable	المتغير الهدف
Test of linearity	اختبار التأثير الخطي للمعاملات
Tests of goodness- of -fit	اختبارات جودة التوفيق
The Chi-Square Distribution	توزيع كاي تربيع
The Coefficient of Determination	معامل التحديد
The Correlation Coefficient	معامل الارتباط
the distinct mid-ranks	أنصاف الرتب المتميزة
the logistic regression model	نموذج الانحدار اللوجستي
The Questionnaire	الاستبيان
the residual	مجموع مربعات الأخطاء البواقي
Transformation	تحويل البيانات
Trimed Mean	الوسط الحسابي المشذب
Two Related Samples Inference	الاستدلال عينتين غير مستقلتين
Two Tailed test	اختبار الفرضية من طرفين
Two- way layout	جدول ثنائي التصنيف
Type of Program	نوع البرنامج المتبع
Tables of Random Numbers	جدول الأرقام العشوائية
Tabular Presentation	العرض الجدولي
The Binomial Distribution	توزيع ثنائي الحدين
The Coefficient of Variation	معامل الاختلاف
The lower specification limit LSL	الحد الأدنى للمواصفة
The Normal Distribution	التوزيع الطبيعي
The Poisson Distribution	توزيع بواسون
The Sample Standard Deviation	خريطة ضبط الانحراف المعياري
Control Chart	للعينات

S

Save Standardized Values as Variables	حفظ القيم المعيارية كمتغيرات
Scale	نسبية أو فنوية
Scales	المقاييس
Screening	فحص البيانات
selecting Cases	تحديد مجموعة من الحالات
Sign Test	اختبار الإشارة
Significance	المعنوية الإحصائية
Significant Differences	فروق معنوية
Skewness	معامل الالتواء للتوزيع
Source of Variation	مصدر التباين
Spearman	معامل ارتباط الرتب
Spearman's Rank Correlation Coefficient	معامل سبيرمان لارتباط الرتب
Spread Sheet	ورقة عمل
Spread vs. Level with Levene Test	اختبار تجانس التباين (اختبار ليفن)
SS deviation from linearity	مجموع مربعات الانحرافات عن الخطية
Standardized coefficient	المعامل المعياري
Statistical Functions	التوابع الإحصائية
Statistical Package for Sciences Socials	حزمة البرامج الإحصائية للعلوم الاجتماعية
Std Error of Kurtosis	الخطأ المعياري لمعامل التفرطح
Std Error of Mean	الخطأ المعياري للمتوسط الحسابي
Std Error of Skewness	الخطأ المعياري لمعامل الالتواء
Std .Deviation	الانحراف المعياري
Std .error of the Estimate	الخطأ المعياري للتقدير
Stem – and – Leaf	مخطط الجذع والورقة
Stepwise	إدخال المتغيرات المستقلة واحداً بعد الآخر
Stepwise Regressin	أسلوب الانحدار على مراحل
String	متغير حرفي
String Functions	التوابع النصية
Subgroups	المجاميع الجزئية
Sum of squares	مجموع المربعات
Survivals	جداول الحياة
symmetric	متناظر
Syntax	ملف التعليمات
sample sizes	حجم العينات

Perform all tests for special causes
Pooled standard deviation
Potential capability
probability plots
Process Capability for Variables
Process Capability Formulas
Process Capability Studies
Process charts (Flow diagrams)
Process Design
Process Mapping
Process Measurement
process statistics table
p-value for a distribution

إنجاز إختبارات الاسباب الخاصة
الانحراف المعياري المجمع
المقدرة المحتملة (الكامنة)
التمثيل البياني للتوزيع الاحتمالي
مقدرة العملية للمتغيرات
الصيغ الرياضية لمقدرة العملية
دراسات مقدرة العملية
خرائط التدفق
تصميم العملية
رسم وتحليل العمليات
قياسات العملية
جدول إحصاءات المقدرة
مستوى المعنوية الفعلي للتوزيع

Quartile
Percentiles & Quartiles
Qualitative Data
Quality Improvement Techniques
Quantitative Data
Quantitative (Metric) Data

الربيعات
الربيعات والمئينات
بيانات نوعية
تقنيات تحسين الجودة
بيانات كمية
بيانات كمية (قياسية)

Random Variable Functions
rank sum test
rank-sum
Reduction Variables
Regression
Relations between Two Variables
Remove
Reports
Response variable
R chart
Random Error
Random Experiment
Range chart
Reliability
Rule violations table
Run Chart

توابع المتغيرات العشوائية
اختبار مجموع الرتب
مجموع الرتب
إنقاص المتغيرات
الانحدار
العلاقات بين متغيرين
لاستبعاد المتغيرات غير المهمة
التقارير
متغير الاستجابة
مخطط المدى
خطأ عشوائي
التجربة العشوائية
مخطط المدى
الموثوقية (الصلاحية)
جدول انتهاك القواعد
مخطط التلاحق (مخطط الدورة)

The Standard Normal Distribution
The target value
The upper specification limit USL
Tolerance limits

التوزيع الطبيعي المعياري
القيمة الهدف (المستهدفة)
الحد الأعلى للمواصفة
حدود التحمل

U

SPSS Data Editor Untitled
Unacceptable
Upper Control Limit

محرر البيانات غير المعنون
عدم قبول الدفعة
حد الضبط الأعلى

V

Percent Valid
Type Variable
View Variable
List Variables
Variance Proportion
Violation for Points

النسبة المئوية الفعلية
نوع المتغير
عرض المتغيرات
قائمة المتغيرات
نسبة التباين
قواعد الضبط (المراقبة)

W

Weighted Average Method
Whiskers
Width
Wilcoxon Signed- Rank
Windows
Within Groups
WLS Weight
World -Class Company

طريقة المتوسط الموزون
الاستطلاات خارج الصندوق
عرض المتغير
اختبار ولكوكسن لإشارة الرتب
بيئة النوافذ
داخل المجموعات
المربعات الصغرى الموزونة
الشركات العالمية للتصنيف

X

X-bar chart

خريطة الوسط الحسابي

Z

Zero-Order-Correlation
Zone Chart

معامل ارتباط بيرسون
مخطط نطاق

" حقوق الطبع والترجمة والنشر
محفوظة لمديرية الكتب والمطبوعات "

تم تدقيق الكتاب من قبل اللجنة العلمية المؤلفة من السادة:

الدكتور

محمد سمير دركزلي

أستاذ بقسم الإحصاء والبرمجة

الدكتور

محمد كبيه

أستاذ بقسم العلوم المالية والمصرفية

UNIVERSITY
OF
ALEPPO

تم تدقيق الكتاب لغوياً من قبل

الدكتورة سويس البطمان

مدرسة بقسم اللغة العربية

المراجع الأجنبية

1. Chava Frankfort-Nachmais, Anna Leon Guerrero “ Social Statistics for a Diverse Society ” Pine Forge Press a Sage Publication Company , Second Edition 2000.
2. R. Mark Sirkin “ Statistics for the Social Sciences “ Sage Publications. Inc 1999.
3. Cao, G., Clarke, S., and Lehaney, B. (2000) “A systemic View of Organizational Change and TQM,” The TQM Magazine, Vol. 12, No. 3, pp.186-193.
4. Love, P. E., Li, H., Irani, Z., and Holt, G. (2000) “Re-thinking TQM: Toward a Framework for Facilitating Learning and Change in Construction Organizations,” The TQM Magazine. Vol. 12, Issue. 2, pp.107-117.
5. Bayeh Mirna. " Testing TQM Practices on customer Satisfaction in Byblos Bank in North Lebanon " DEA's Memoire , Universite Saint-Esprit, Kaslik - Faculte de Gestion et de Sciences Commerciales . 1997 .

المراجع العربية

- (١) شرجي ، عبد الرزاق ، ومكداشي ، طارق " بحوث التسويق واستخدام SPSS ، الجامعة اللبنانية الأمريكية LAU ، ١٩٩٩ م.
- (٢) كبيه، محمد ، وعبيدو أميرة ، مقدرة العملية والضبط الإحصائي للجودة، مجلة بحوث جامعة حلب، العدد رقم (٣٢) لعام ٢٠٠٣ م.
- (٣) كبيه، محمد ، مراقبة الجودة الإحصائية " تطبيقات عملية باستخدام برنامج MINITAB الندوة العلمية السادسة (تطبيقات المعلوماتية في العلوم الهندسية)، الجمعية العلمية السورية للمعلوماتية بجامعة حلب اكتوبر ٢٠٠١ م
- (٤) كبيه، محمد ، جودة الخدمات التعليمية "الندوة الوطنية الثامنة للجودة" الجودة سبيلنا إلى التحديث والتطوير، دمشق ٢٠٠٣ م .
- (٥) كبيه ، محمد " الأساليب الكمية في التسويق " مذكرة لطلاب دبلوم التسويق التطبيقي في الغرفة التجارية الصناعية بالرياض المملكة العربية السعودية ١٩٩٨ م.
- (٦) كبيه ، محمد " جودة الخدمة المصرفية (مفهومها - أبعادها - طرق قياسها)، حالة عملية باستخدام SPSS لقياس رضا المتعاملين مع المصارف السورية" الاسبوع الاقتصادي بكلية الاقتصاد ٢٠٠٠ م.



