

المعهد العربي للتدريب والبحوث الإحصائية



دليلك الى .. البرنامج الإحصائي *SPSS*

Version 10

الإصدار العاشر

إعداد

سعد ز غول بشير

رئيس باحثين أقدم

الجهاز المركزي للإحصاء / جمهورية العراق

2003

تقديم

أن تطوير العمل الإحصائي باستخدام البرامج الجاهزة يأتي في مقدمة الأولويات التي يسعى المعهد العربي للتدريب والبحوث الإحصائية الى تحقيقها .

ومن هذا المنطلق فقد جاء هذا المؤلف ليلقي الضوء على الاستخدامات المختلفة للبرنامج الجاهز SPSS في إعداد الإحصاءات التطبيقية وليكون دليل عمل تفصيلي للعاملين في المجال الإحصائي من طلبة وباحثين .

ولقد ارتأينا ان نكلف السيد سعد زغلول بشير بأعداد هذا المؤلف لما عهدناه فيه من دراية وممارسة عملية في موضوع البرامجيات الجاهزة .

نرجو أن يكون لهذا المؤلف فائدة ملموسة في مجالات الإحصاء في الوطن العربي .

ومن الله التوفيق .

الدكتور خالد خواجه

مدير عام المعهد العربي للتدريب

والبحوث الإحصائية

بغداد / 2003

مقدمة

يعتبر البرنامج الجاهز SPSS من أكثر البرامج الإحصائية استخداماً من قبل شريحة واسعة من الطلبة والباحثين في مختلف الاختصاصات الإحصائية والطبية والهندسية والزراعية فقد أصبح علم الإحصاء في السنوات الأخيرة أداة أساسية لاغنى عنها لتوصيف البيانات وتحليلها وأعداد التقديرات والتنبؤات المستقبلية ونظراً لكبر حجم البيانات التي يتعامل معها علم الإحصاء من جهة واعتماده على أساليب كمية مطولة فقد برزت الحاجة الى ضرورة استخدام الحاسب الشخصي لإتجاز العمليات الإحصائية اختصاراً للجهد والوقت .

ونظراً لقلة من يجيد استخدام البرنامج SPSS بصورة وافية إضافة الى افتقار المكتبة العربية الى كتب تعليمية حول هذا البرنامج فقد كان هدفنا ومن خلال هذه المحاولة المتواضعة أن نضع مصدراً تفصيلياً بين أيدي طلبتنا في أقسام الإحصاء والاختصاصات المختلفة الأخرى سواء في الدراسات الأولية أو العليا والباحثين في مختلف المجالات .

يتضمن الكتاب تعريفاً بالجوانب الأساسية لبرنامج SPSS ويهدف الى إكساب المستفيد المهارات اللازمة للاستفادة القصوى من إمكانيات البرنامج المتاحة أخذين بنظر الاعتبار أن هناك عدداً كبيراً من المستفيدين لا يمتلكون خلفية إحصائية وافية تمكنهم من التعامل مع البرنامج بصورة صحيحة حيث تضمن الكتاب عرضاً موجزاً للجانب النظري للأسلوب الإحصائي المستخدم بالإضافة الى التفسير الإحصائي لمخرجات البرنامج لبعض التطبيقات الإحصائية المهمة مثل اختبار الفرضيات ، التحليل العاملي ، الاختبارات اللامعلمية ، تحليل التباين ، الانحدار كما أنه يتضمن الرسوم البيانية والجداول التكرارية والمقاييس الوصفية . وقد تم التعامل مع تطبيقات البرنامج من خلال أمثلة مبسطة تتيح للقارئ الانتقال الى خطوات متقدمة بسهولة وإن معظم هذه الأمثلة مأخوذ من مصادر عربية وأجنبية معتمدة عدا ما تم وضعه من قبلنا .

كما يتضمن الكتاب نواحي استخدام البرنامج كقاعدة بيانات فيما يتعلق بدمج الملفات وترتيبها واختيار الحالات وتبادل البيانات مع البرامج الأخرى .

أخيراً يطيب لي وأنا أنتهي من أعداد هذا الكتاب أن أتقدم بالشكر الجزيل الى كل من مد يد العون لإظهاره بصورته الحالية وخص منهم بالذكر الدكتور مهدي محسن العلاق خبير الجهاز المركزي للإحصاء لتكده عناء مراجعة المادة العلمية ولما أبداه من ملاحظات أغنت الكتاب كثيراً . كما يطال شكري الزملاء والأخوة في الجهاز المركزي للإحصاء الذين ساهموا بملاحظاتهم وأرائهم فلمني كل التقدير . أخيراً نأمل أن يحقق هذا المطبوع الفائدة المتوخاة منه .

والله ولي التوفيق .

سعد زغلول بشير

بغداد / 2003

المحتويات		الموضوع
الصفحة		
I-III		مقدمة عامة
1	تهيئة ملفات الإدخال لبرنامج SPSS	الفصل الأول
1	تهيئة الملفات	1-1
3	إدخال البيانات	2-1
4	أولاً : اسم المتغير	
5	ثانياً : نوع المتغير	
6	ثالثاً: عرض المتغير	
6	رابعاً : عدد المراتب العشرية	
6	خامساً : عنوان المتغير	
7	سادساً : عناوين القيمة	
7	سابعاً : تعريف القيم المفقودة	
8	ثامناً : عرض العمود	
9	تاسعاً : محاذاة النص	
9	عاشراً : القياس	
11	العمليات على المتغيرات و صفاتها في ورقة Data Editor	3-1
13	استعمال مجاميع جزئية من المتغيرات	4-1
16	أوامر القائمتين View و Data	الفصل الثاني
16	أوامر القائمة View	1-2
20	أوامر القائمة Data	2-2
20	1. تعريف التاريخ للسلسلة الزمنية Define Date	
21	2. الأمر Insert Variable	
21	3. الأمر Insert Case	
21	4. الأمر Go to Case	
21	5. الأمر Sort Cases	
23	6. الأمر Transpose	
24	7. دمج الملفات Merge Files	
24	أ . إضافة حالات Add Cases	
26	ب . إضافة متغيرات Add Variables	
31	8. فصل (تجزئة) الملفات Split Files	
35	9. تجميع البيانات Aggregate Data	
الصفحة		الموضوع
37	10 . اختيار الحالات Select Cases	
40	11. ترجيح الحالات Weight Cases	
42	تحويل البيانات Data Transformation	الفصل الثالث
42	1. الأمر Compute	
45	2. الأمر Random Number Seed	
45	3. الأمر Count	
47	4. الأمر Recode	
47	أ . الأمر Recode in to Same variables	

49	 Recode in to different Variable	ب. الأمر
50	 Categorize Variables	5. الأمر
51	 Automatic recode	6. الترميز التلقائي
52	 Rank Cases	7. الأمر
56	 Time Series	8. السلاسل الزمنية
60	 Replace missing Values	9. تقدير القيم المفقودة
62	 الإحصاءات الوصفية والجداول التكرارية	الفصل الرابع
62	 Frequencies	1-4
66	 Descriptives	2-4
68	 Pivot Tables	الفصل الخامس
68	 Pivot Table	1-5
68	 Edit Pivot Table	2-5
72	 Book Marks	3-5
74	 Explore	الفصل السادس
74	 Explore	1-6
79	 Standard error	الخطأ المعياري
79	 μ	تكوين فترة ثقة لمتوسط المجتمع
79	 Trimmed Mean	الوسط الحسابي المشذب
80		الربيعات والمئينات
81		اختبار التوزيع الطبيعي للبيانات من نسبة معامل الالتواء
81	 Stem-and-Leaf	مخطط
82	 Histogram	المرج التكراري
82	 Boxplot	مخطط

الصفحة	الموضوع
82	 Normality Plots with Tests
82	 1. اختبار Kolmogrov-Smirnov
83	 2. مخطط Normal Q-Q Plot
84	 3. مخطط Detrended Normal Q-Q Plot
86	 Test of Homogeneity of Variances
86	 Spread vs. Level with Levene Test
89	 1. الاختبار الإحصائي لتجانس التباين بواسطة إحصائية Levene
89	 2. اختبار تجانس التباين من خلال مخطط Spread vs. Level Plot
92	 التعامل مع القيم المفقودة
96	 Crosstabs
103	 Compare Means
103	 Means
108	 One Sample T-Test
111	 Independent Samples T-Test
113	 Paired Samples T-Test
115	 Analysis of Variance
115	 One way ANOVA

119	 Orthogonal Comparisons المقارنات المستقلة	1-1-9
122	 Trend Analysis تحليل الاتجاهات	2-1-9
123	 Tow Way ANOVA تحليل التباين لمعيارين	2-9
129	 Covariance Analysis تحليل التباين المشترك	3-9
132	Correlation and Regression Analysis تحليل الارتباط و الانحدار	الفصل العاشر
132	 Correlation الارتباط	1-10
132	 Simple Linear Correlation الارتباط الخطي البسيط	2-10
135	 Partial Correlation الارتباط الجزئي	3-10
138	 Regression Analysis تحليل الانحدار	4-10
138	 نموذج الانحدار الخطي البسيط	1-4-10
146	...Weighted Least Squares Method طريقة المربعات الصغرى الموزونة	2-4-10
149	 نموذج الانحدار الخطي المتعدد	3-4-10
159	 Factor Analysis التحليل العاملي	الفصل الحادي عشر
159	 التحليل العاملي	1-11
	الموضوع	الصفحة
159	... Principal Components Method طريقة المكونات الأساسية	2-11
168	 Factor Analysis Methods طرق التحليل العاملي	3-11
171	 Non Parametric Tests الاختبارات اللامعلمية	الفصل الثاني عشر
171	 Chi-Square اختبار	1-12
175	Tow Independent Samples Tests اختبارات عينتين مستقلتين	2-12
177	K-Related Samples Tests اختبارات K من العينات المرتبطة	3-12
180	 CHARTS المخططات البيانية	الفصل الثالث عشر
180	 Bar Charts الأعمدة البيانية	1-13
192	 Chart Template عمل قالب لمخطط بياني	2-13
193	 Line مخطط Bar الى مخطط	3-13
195	 Pie مخطط Bar الى مخطط	4-13
205	 Histogram المدرج التكراري	5-13
208	 Box Plot مخطط	6-13
213	 Scatterplot مخطط شكل الانتشار	7-13
214	 1. شكل الانتشار البسيط Simple	
218	 2. شكل انتشار Overlay	
220	 3. شكل الانتشار Matrix	
221	 4. شكل الانتشار ثلاثي الأبعاد 3-D	
223	 Data Exchange تبادل البيانات	الفصل الرابع عشر
223	 Importing Data Files استيراد الملفات	1-14
230	 Exporting Data Files تصدير الملفات	2-14
233	 Syntax Commands كتابة الأوامر	الفصل الخامس عشر
233	 Syntax File ملف الأوامر	1-15
233	 Command Syntax الطرق المساعدة في بناء ملفات الأوامر	2-15
233	 Dialog Boxes كتابة الأوامر من صناديق الحوار	1-2-15
235	 نسخ الأوامر من Log في مخرجات البرنامج	2-2-15

236	 Journal File	نسخ الأوامر من	3-2-15
237	 To Run Command Syntax	تمشية ملف الأوامر	3-15
241	 Multiple Response Analysis	تحليل الاستجابات المتعددة	الفصل السادس عشر
241	Multiple Response Frequencies	أسلوب التكرارات متعددة الاستجابة	1-16

الصفحة	الموضوع
241	 Multiple dichotomy method
243	Multiple Category Method
245	 Multiple Response Crosstabs
	2-16

مقدمة عامة

Introduction

البرنامج الإحصائي spss

يعد البرنامج الإحصائي spss (مختصر statistical package for social sciences) من أكثر البرامج الإحصائية استخداما من قبل الباحثين في المجالات التربوية و الاجتماعية والفنية و الهندسية والزراعية في إجراء التحليلات الإحصائية اللازمة.

وقد بدأت شركة (spss) بإعداد هذا النظام الذي كان يعمل تحت نظام تشغيل MS-DOS وقد تم تطويره ليعمل في بيئة نظام التشغيل WINDOWS في عام 1993 متلافيا بذلك الصعوبات التي كانت تواجه العاملين على هذا النظام في بيئة MS-DOS . وقد توالى الإصدارات لهذا النظام التي كان آخرها الإصدار العاشر spss 10.0 الذي صدر في 27 / 11 / 1999 حيث يوفر هذا النظام مجالا واسعا للتحليلات الإحصائية واعداد المخططات البيانية لتلبية حاجة المختصين و المهتمين في مجال الإحصاء كما يوفر إمكانية تناقل البيانات مع قواعد البيانات و برامج EXCEL و LOTUS وغيرها من البرامجيات.

النوافذ المتوفرة في برنامج SPSS

تتوفر في برنامج SPSS الأنواع التالية من النوافذ.

1. نافذة محرر البيانات Data Editor :- وهذه النافذة تعرض محتويات ملف معين من البيانات حيث يمكن تكوين ملف جديد أو تحويل ملف موجود وان هذه النافذة تفتح تلقائيا عند بدء تشغيل البرنامج .
2. نافذة المشاهد Viewer :- هذه النافذة تعرض جميع النتائج الإحصائية و الجداول و المخططات charts حيث يمكن تنقيح النتائج و خزنها.
3. نافذة مسودة المشاهد Draft viewer :- هذه النافذة تتيح عرض المخرجات كنص اعتيادي (بدلا من جداول محورية تفاعلية) ولهذا لايمكن تحويل الجداول والمخططات في هذه النافذة.
4. نافذة محرر الجدول المحوري Pivot Table Editor :- هذه النافذة تتيح إمكانية تحويل الجداول المحورية بعدة طرق.
5. نافذة محرر المخططات Chart Editor :- تتيح هذه النافذة إمكانية تحويل المخططات.
6. نافذة محرر النصوص Text output Editor :- تتيح هذه النافذة إمكانية تحويل المخرجات التي لا تعرض كجداول محورية.
7. نافذة محرر القواعد Syntax Editor :- تتيح هذه النوافذ إمكانية تخزين خيارات صناديق الحوار حيث يمكن تحويلها لإضافة أوامر و مميزات لا تتوفر في الاوامر القياسية لبرنامج SPSS .
8. نافذة محرر الخطوط Script Editor :- تتيح هذه النافذة إمكانية خلق و تحويل الخطوط الأساسية.

أنواع الملفات في برنامج SPSS

تتوفر في برنامج SPSS الأنواع الرئيسية التالية من الملفات :

1. ملفات البيانات Data Files :- وهي الملفات التي تتكون باستخدام محرر البيانات Data Editor وهي تحتوي على البيانات التي تستخدم في التحليل الإحصائي و يكون لهذا النوع من الملفات الاستطالة SAV .
2. ملفات المخرجات الإحصائية Output Files :- وهي الملفات التي تحتوي على مخرجات التحليل الإحصائي أو المخططات و تكون ذات الاستطالة SPO .
3. ملفات التعليمات (syntax) :- وهي الملفات التي تحوي الإجراءات الإحصائية التي تخزن على شكل أوامر وتكون ذات استطالة SPS .

تشغيل برنامج SPSS

يمكن تشغيل برنامج SPSS بأحد طريقتين :-

1. عن طريق النقر مرتين Double-click بزر الماوس الأيسر على أيقونة برنامج SPSS (في حال وجودها على سطح المكتب).
2. أو من خلال الزر Start وحسب التسلسل التالي

Start → Programs → SPSS V.10.0

حيث تظهر نافذة Data Editor تلقائياً الشكل (1-1) .

استخدامات جهاز الفأرة Mouse

- النقر click :- النقر بزر الماوس الأيسر وقد تشير إلى ذلك بالنقر (للاختصار).
- النقر المزدوج Double-click :- نقر زر الماوس الأيسر مرتين متتاليتين وقد تشير إلى ذلك بالنقر مرتين (للاختصار) .
- النقر click بزر الماوس الأيمن :- ويفيد في إظهار قائمة الأوامر المختصرة short command list وتعرف أيضاً بالقائمة الموضوعية Context List كما هو الحال لبرامج office كما تفيد أيضاً في إظهار تعليق لأي نص يرد في صندوق الحوار أو أي مؤشر محتسب في الجداول المحورية.

سنة طرق مباشرة للحصول على مساعدة

1. قائمة Help : وهذه القائمة موجودة ضمن شريط القوائم Menu bar لكل نافذة من نوافذ SPSS . وتتضمن Topics ومن خلالها يتم توفير ثلاثة أنواع من المساعدة بواسطة Contents و Index و Find . كما تتضمن Tutorial التي توفر مدخلاً تعليمياً إلى برنامج SPSS .
2. زر المساعدة في صندوق الحوار Dialog box help button هذا الزر موجود في أغلب صناديق الحوار لبرنامج SPSS ومن خلاله يمكن الحصول على معلومات عامة عن الموضوع المتعلق بصندوق الحوار .
3. المساعدة الموضوعية في صندوق الحوار Dialog box context menu help : يمكن الحصول على مساعدة عن أي نص يرد في صندوق الحوار بنقر ذلك النص بزر الماوس الأيمن Right-Click لعرض وصف عن ذلك النص .
4. المساعدة الموضوعية في الجدول المحوري Pivot table context menu help : يمكن الحصول على هذه المساعدة بنقر الجدول المحوري مرتين بزر الماوس الأيسر لتفعيله ثم نقر عنوان Label الصف أو العمود (في حالة أنه يمثل مؤشراً تم احتسابه من خلال البرنامج) بزر الماوس الأيمن في

الجدول المحوري في شاشة SPSS Viewer (علماً أن معظم مخرجات SPSS هي جداول محورية) ثم اختيار What's this من القائمة الموضوعية لعرض تعريف عن محتويات الصف او العمود .

5. مرشد النتائج Result Coach: يمكن الحصول على هذا النوع من المساعدة بنقر الجدول المحوري مرتين بزر الماوس الأيسر لتفعيله ثم نقر الجدول بزر الماوس الأيمن واختيار Result Coach من القائمة الموضوعية لعرض تفسير إحصائي مبسط للنتائج الإحصائية المضمنة في الجدول من خلال عدة نوافذ متسلسلة .

6. المرشد Tutorial: يمكن الحصول على هذه المساعدة باختيار Tutorial من قائمة help في أي نافذة للوصول الى مدخل تعليمي مباشر باستعمال عدد من النوافذ التعليمية المتسلسلة .

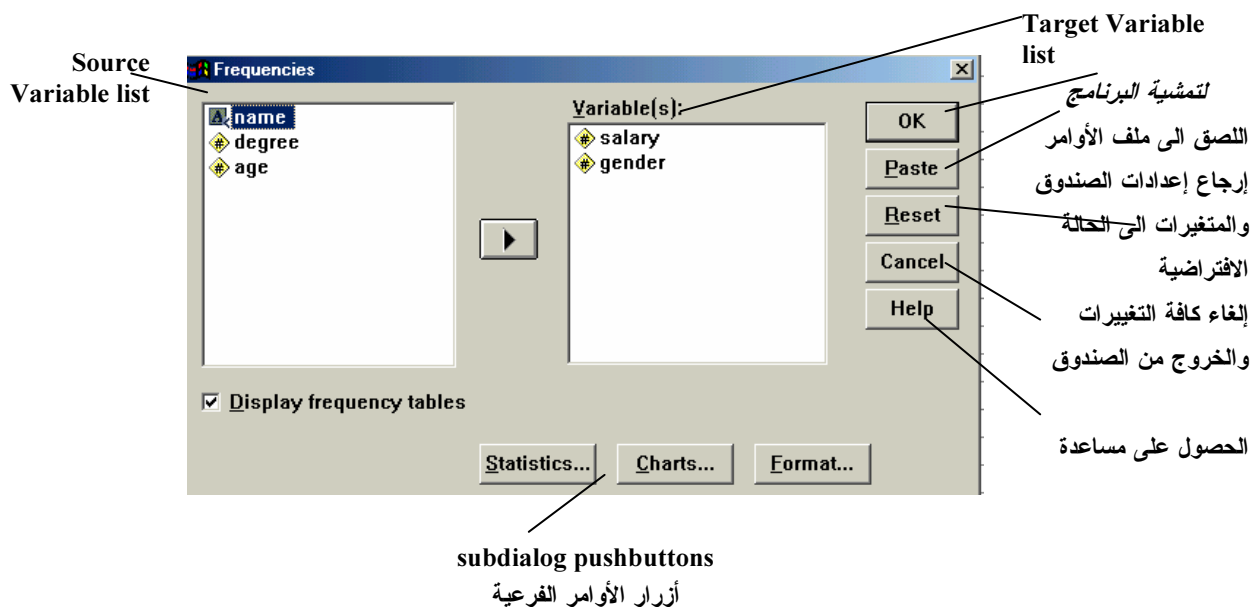
صناديق الحوار لبرنامج SPSS

أن صندوق الحوار Dialog Box في برنامج SPSS - وكما هو الحال للبرامج التي تعمل مع النظام Windows - يتيح اختيار المتغيرات التي نرغب في إجراء التحليلات الإحصائية عليها كما انه يعتبر بديلاً عن كتابة الأوامر البرمجية المعقدة التي يكون الهدف منها إنجاز أسلوب إحصائي معين (كما هو الحال لبرنامج SPSS الذي يعمل مع نظام التشغيل MS- DOS) ويتكون الصندوق من العناصر التالية :

قائمة المتغيرات المصدر Source Variables List : وتشمل كافة المتغيرات الموجودة في الملف الحالي ذات الأنواع المسموحة الاستخدام للأسلوب الإحصائي المختار .

قائمة (أو قوائم) المتغيرات الهدف Target Variable List(s) : واحدة أو أكثر من القوائم التي تتضمن أسماء المتغيرات المختارة للتحليل الإحصائي .

أزرار الأوامر Command pushbuttons : وهذه الأزرار تقوم بإعلام البرنامج لتنفيذ عمل معين مثلاً تمشية البرنامج أو الحصول على مساعدة ، كما هو واضح في صندوق الحوار التالي للأمر Frequencies



الفصل الأول

تهيئة ملفات الإدخال لبرنامج spss

Data Entry

(1-1) تهيئة الملفات

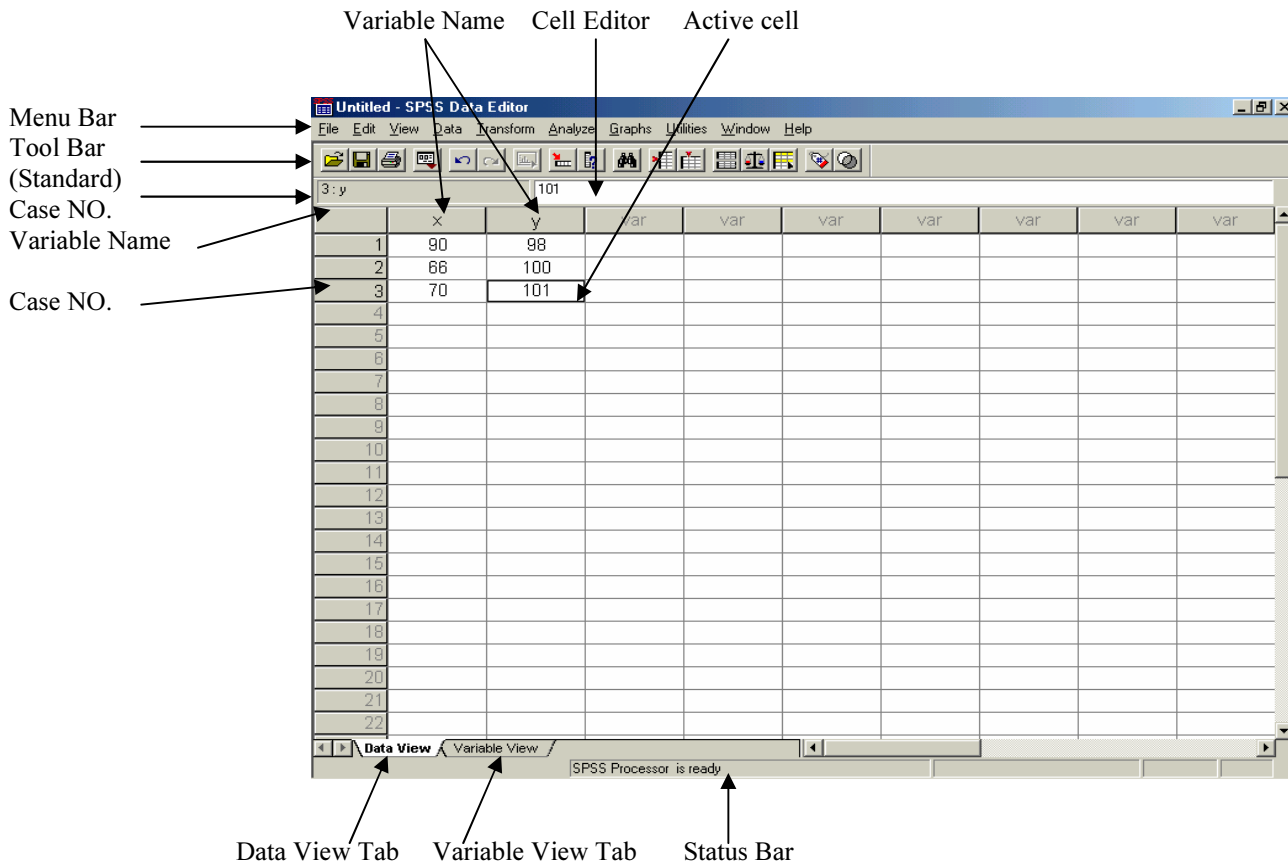
أن تهيئة الملفات في برنامج spss ينفذ بواسطة ما يعرف بمحرر البيانات Data Editor ، وهو عبارة عن ورقة نثر Spread sheet تشبه ورقة العمل في برنامج Excel و أن نافذة محرر البيانات تفتح تلقائياً عند بداية تشغيل البرنامج و يوفر محرر البيانات نوعين من العرض للبيانات :-

1. Data view :- هو عبارة عن ورقة نثر مقسمة إلى أعمدة و صفوف (الأعمدة تمثل المتغيرات Variables و الصفوف تمثل الحالات Cases) فعلى سبيل المثال فأن كل مستجيب لأسئلة استبيان ما يمثل حالة منفصلة و أن كل خلية cell هي عبارة عن تقاطع المتغير مع الحالة.

2. Variable view :- يتضمن وصفا لصفات كل متغير في ملف البيانات و في هذه الحالة تكون الصفوف هي المتغيرات و الأعمدة تمثل صفات المتغير و تشمل الصفات (اسم المتغير ، نوعه ، عدد الأرقام أو الرموز الخ) حيث يمكن إضافة أو حذف أو تحويل صفات المتغيرات. انظر الشكل (1-1) .

شكل (1-1)

محرر البيانات Data Editor لبرنامج spss



شكل (2-1) : شريط الأدوات القياسي Standard Tool Bar لورقة نثر Data Editor

الأيقونة	العنوان	الوظيفة
	Open file	فتح ملف
	Save File	خزن ملف
	Print	طبع ملف
	Dialog Recall	إظهار آخر 12 مجموعة من الإجراءات
	Undo	التراجع عن آخر تغيير
	Redo	اعادة إجراء التغيير
	Go To Chart	الانتقال إلى التخطيط
	Go to Case	الانتقال إلى الحالة
	Variables	عرض معلومات عن المتغيرات
	Find	البحث عن حالة ضمن متغير
	Insert Case	إضافة حالة
	Insert Variable	إضافة متغير
	Split File	تجزئة ملف
	Weight Cases	تحديد أوزان للحالات
	Select Cases	اختيار حالات
	Value Labels	إظهار أو إخفاء عناوين القيم
	Use Sets	استخدام مجموعات جزئية من المتغيرات المتوفرة في Data Editor

ملاحظات

1. يمكن إظهار ورقة Variable view عن طريق نقر عروة (tab) variable view أسفل ورقة Data view أو أنقر مرتين أسم المتغير (أعلى العمود) في ورقة Data view.
2. للانتقال من ورقة Variable view إلى ورقة Data view أنقر عروة Data view أسفل ورقة Variable view أو أنقر مرتين رقم (السطر) في Variable view.

(2 - 1) إدخال البيانات

لنفترض أننا نريد إدخال البيانات التالية التي تمثل قيود مجموعة معينة من الأشخاص في اختبار معين

id	gender	bdate	Grade
Ahmad	1	15.7.69	76
Khadim	1	12.4.70	80
Sabah	2	1.6.68	83
Mahdi	1	9.5.72	90
Zainab	2	20.9.74	80
Nabil	1	5.1.67	78

أن هذه البيانات يتم إدخالها في ورقة Data View (يمكن أن تكون الورقة خالية من البيانات أو يتم إضافة البيانات إلى الورقة) علماً أن البيانات في هذه الورقة هي عبارة عن متغيرات (كل عمود في الورقة يمثل متغيراً من المتغيرات) فبالنسبة للمتغيرات أعلاه يمكن تقسيمها كما يلي :-

المتغير الأول - متغير الاسم (id) وهو متغير رمزي string variable .

المتغير الثاني - متغير الجنس (gender) الرقم 1 يمثل الذكور و الرقم 2 يمثل الإناث وهو متغير رقمي Numeric variable .

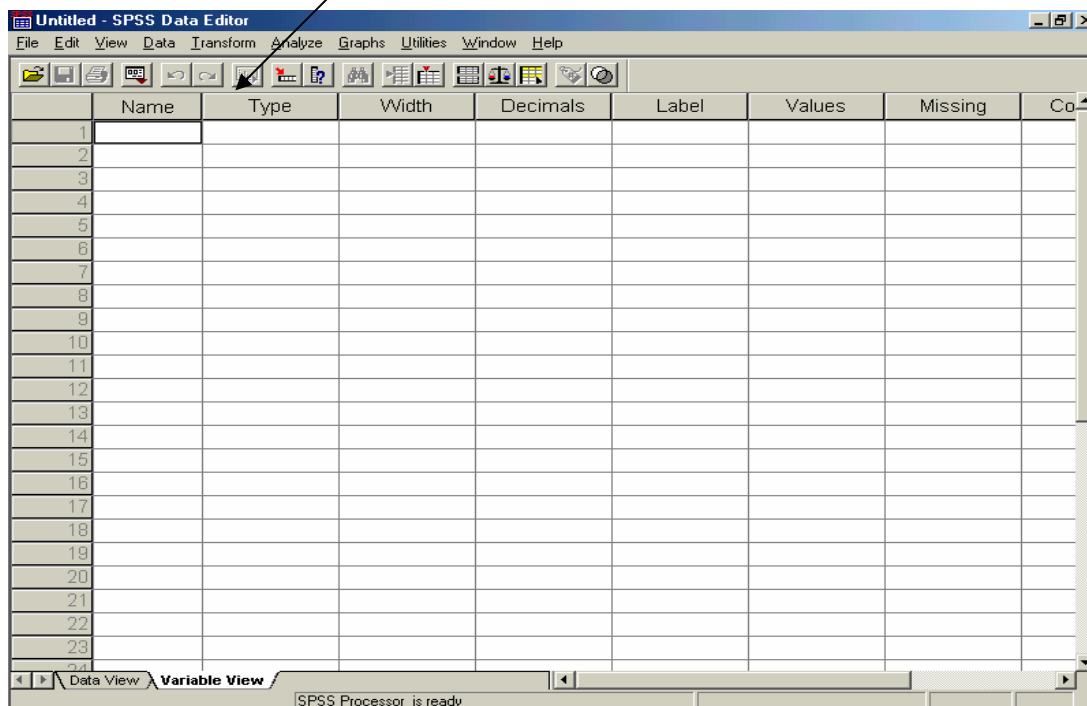
المتغير الثالث - متغير تاريخ الميلاد (bdate) ونوع هذا المتغير هو متغير تاريخ date .

المتغير الرابع - متغير درجة الاختبار (Grade) وهو متغير رقمي .

قبل أن نقوم بإدخال قيم البيانات في ورقة Data View نقوم بتعريف أسم وصفات المتغيرات الأربعة Name and Attributes ولذلك ننتقل إلى ورقة Variable View عن طريق نقر عروة Variable View أسفل ورقة Data View حيث تظهر الورقة كما في الشكل (3-1) والتي يمثل كل سطر فيها متغيراً من المتغيرات .

شكل (3-1)
ورقة نشر Variable View

Variables Names & Attributes



أن أسم وصفات المتغير Variable Name & Attributes تشمل ما يلي :

1. أسم المتغير، 2. نوع المتغير ، 3. عرض المتغير ، 4. عدد المراتب العشرية للمتغيرات العددية ،
 5. عنوان المتغير ، 6. عناوين القيم ، 7. القيم المفقودة ، 8. عرض العمود ، 9. محاذاة النص ، 10. القياس .
- حيث يتم تعريفها للمتغيرات الأربعة في ورقة Variable View كما يلي :

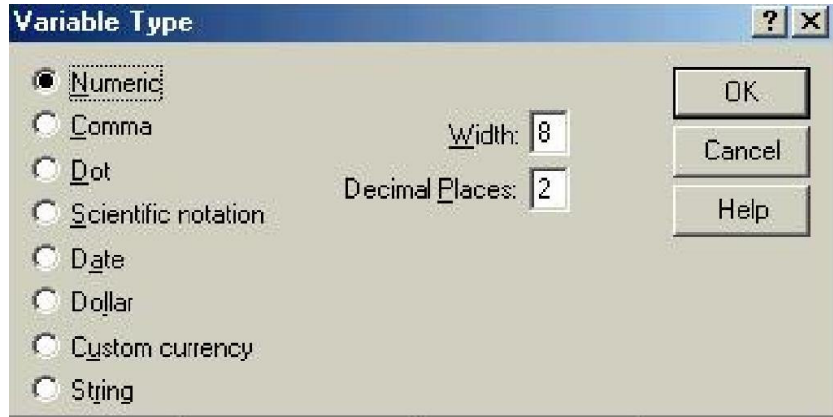
أولاً : اسم المتغير Variable Name

أنقر أي خلية في العمود الذي يحمل العنوان Name ثم أكتب أسم المتغير مثلاً الاسم Id في السطر الأول و الاسم Gender في السطر الثاني وهكذا لبقية المتغيرات علماً انه يتوجب اتباع القواعد التالية في كتابة أسماء المتغيرات في برنامج SPSS :-

- 1- لا يزيد طول الاسم عن ثمانية رموز characters.
- 2- يجب أن يبدأ أسم المتغير بحرف أما بقية الرموز فقد تكون أحرفاً أو أرقاماً أو فترة period و يرمز لها (.) أو بقية الرموز @ ، # ، - ، \$.
- 3- لا يمكن أن ينتهي أسم المتغير بفترة (.) .
- 4- لا يتضمن أسم المتغير فراغات و بعض الرموز الخاصة مثل ! ، ? ، ' ، * .
- 5- لا يميز برنامج SPSS بين الحروف الكبيرة و الحروف الصغيرة فالأسماء NEWVAR و newvar تعتبر متماثلة حيث أن البرنامج لا يتقبل سوى الحروف الصغيرة لأسماء المتغيرات .

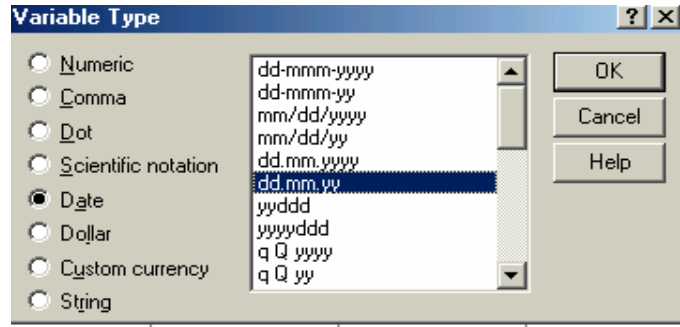
ثانيا : نوع المتغير Variable Type

بعد إدخال اسم المتغير انتقل إلى الخلية المجاورة التي تقع ضمن العمود الذي يحمل العنوان Type و عند الوقوف عليها يظهر زر button عند النقر عليه يظهر صندوق الحوار Dialog box كما يلي :
صندوق حوار Variable Type



حيث يمكن تحديد الأنواع التالية من المتغيرات :-

- Numeric : متغير عددي وهو النوع الافتراضي للمتغيرات في ورقة Data View .
- Comma : هو متغير عددي مع إضافة فاصلة (,) للفصل بين كل ثلاثة مراتب صحيحة مثلا العدد 722667.123 يكتب 722,667.123 بموجب هذا النوع .
- Dot : هو متغير عددي مع استخدام (.) لفصل كل ثلاثة مراتب صحيحة و تستخدم الفاصلة (,) للفصل بين الجزء الصحيح و الجزء العشري فالعدد أعلاه يكتب 722.667,123 بموجب هذا النوع.
- Scientific Notation : هو رمز مكتوب بصيغة التدوين اليائي E-notation مثلا العدد 10^7 يكتب 1.0E+07 والعدد 1234 يكتب 1.2E+03 .
- Date : متغير يمثل التاريخ أو الوقت بالساعات مثلا .
- Dollar : يستعمل كرمز للدولار الأمريكي.
- Custom Currency : متغير من تعريف المستفيد للدلالة على العملة يمكن ضبطه من الخيار Edit → Options → Currency .
- String : هو متغير رمزي (أسماء مثلا) .
- أما المربع Width فيبين عدد مراتب المتغير .
- أما مربع Decimal Places : فيمثل عدد المراتب العشرية للمتغيرات العددية فقط.
- بالنسبة للمتغير id فسنتار النوع string بما انه متغير رمزي في هذه الحالة سيختفي مربع Decimal Places من صندوق حوار Variable Rank . أما المتغيرين gender و Grade فسنتار النوع Numeric لكل منهما. أما متغير تاريخ الميلاد bdate فسنتار النوع Date حيث يظهر صندوق الحوار التالي :



حيث يمكن ان نختار أحد أنواع كتابة التاريخ وقد اخترنا dd .mm .yy ليعبر عن النوع المطلوب للتاريخ.

ملاحظات

1. أن النوع الافتراضي لنوع المتغير هو Numeric ولهذا ففي حالة المتغيرات العددية (نوع Numeric فقط) يمكن إدخال الأرقام مباشرة في شاشة Data View بدون تحديد نوع المتغير (بأنه متغير عددي) أما في حالة المتغير غير العددي (عدا numeric) يتوجب تحديد نوع المتغير قبل الإدخال.
2. بالنسبة للمتغيرات العددية Numeric و Comma و Dot فيمكن إدخال مراتب عشرية (لغاية 16 مرتبة) و يقوم Data View بعرض العدد المحدد للمراتب العشرية و تدوير القيم ذات المراتب العشرية الأعلى علما ان القيمة الكاملة تستخدم في الحسابات.
3. بالنسبة لسلاسل الحروف string (المتغيرات الرمزية) فان القيمة تمتد لغاية أعلى عرض فلو كان للمتغير الرمزي width = 6 فان كلمة No تخزن " No " وليست مساوية إلى "NO".

ثالثاً : عرض المتغير Variable width

أن عرض المتغير Width (هو نفسه الوارد في صندوق حوار variable type وهو العمود الثالث في ورقة Variable View) فيمثل عدد الرموز المخصصة للجزء العشري زائداً رمز للفاصلة العشرية وما يتبقى من الرموز تخصص للعدد الصحيح في حالة المتغيرات العددية أما بالنسبة للمتغيرات غير العددية فيمثل عرض المتغير عدد المراتب المخصصة للمتغير غير العددي . مثلاً عرض العدد التالي 333,333.02 نوع Comma هو 10. بما أن قيم المتغير Gender تتكون من رمز واحد (1 أو 2) فقد جعلنا عرض المتغير في عمود Width يساوي 1 (من الممكن اختيار أي رقم آخر أكبر من واحد) أما بالنسبة لمتغير الدرجات grade فقد اخترنا Width يساوي 3 لأن الدرجات الامتحانية تتراوح من صفر الى 100 وبالنسبة لمتغير التاريخ فأن البرنامج يحدد Width الافتراضي له وهو 8 رموز لنوع التاريخ الذي اخترناه في هذا المثال dd.mm.yy. أما متغير الاسم id فقد كانت القيمة الافتراضية لـ Width هي 8 . في حالة تحديد عدد قليل من الرموز لمتغير الاسم id مثلاً Width = 4 نلاحظ اختفاء جزء من الاسم في ورقة Data View إذا كان يحتوي على عدد من الرموز (الحروف) تزيد عن 4 .

رابعاً : عدد المراتب العشرية Decimals

يمثل عدد المراتب العشرية المخصصة للكسر العشري في المتغيرات العددية (Numeric ، Comma ، Dot) ويمكن زيادة او إنقاص المراتب العشرية بواسطة الأسهم الى الأعلى والى الأسفل علماً أن المراتب العشرية يمكن تحديدها من صندوق حوار Variable Type أيضا .

خامساً : عنوان المتغير Variable Label

يمكن أن يعطى المتغير عنواناً يصل عدد رموزه إلى 256 رمز يستعمل لوصف المتغير فمثلا يعطى

العنوان Date of birth للمتغير bdate حيث يستعمل العنوان بدلاً من اسم المتغير في مخرجات (جداول) برنامج spss.

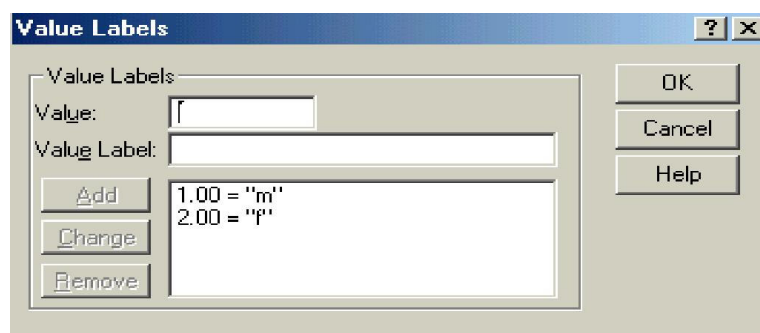
سادساً : عناوين القيمة Value Labels

أحياناً تبرز الحاجة إلى تعيين عنوان للقيمة كون المتغير يستعمل قيماً عددية للتعبير عن قيم غير عددية مثلاً متغير Gender يستعمل الرقم 1 للتعبير عن الذكور Males والرقم 2 للتعبير عن الإناث Females وكما هو وارد في الجدول التالي :

Value	Value Label
1	m
2	f

لإعطاء عنوان لقيم المتغير gender وحسب الجدول المذكور نتبع الخطوات التالية :

1. انقر الخلية التي تقع تحت العمود Value وفي سطر المتغير Gender في Variable view.
 2. يظهر صندوق حوار define Labels.
 3. انقر المستطيل المجاور لكلمة value واكتب فيه الرقم 1 ثم انقر المستطيل المجاور لكلمة value label و اكتب m ثم انقر الزر add لإضافة العنوان.
 4. انقر المستطيل المجاور لكلمة value واكتب الرقم 2 فيه ثم انقر المستطيل المجاور لكلمة value label و اكتب f فيه ثم انقر الزر Add لإضافة العنوان (الشكل التالي يوضح نتائج الإضافة).
- علماً انه يمكن إزالة عنوان بعد الوقوف على العنوان في المستطيل الأسفل ونقر زر Remove أو تغيير عنوان بنقر الزر change .
5. عند الانتهاء انقر زر OK .



ملاحظات

1. يمكن ان يصل طول عنوان القيمة الى 60 رمز.
2. يمكن ان يكون العنوان قيماً عددية ليمثل قيماً غير عددية مثلاً الرقم 1 عنوان للذكور m والرقم 2 عنوان للإناث f.
3. يمكن إظهار عنوان القيمة value label في ورقة Data view أما بالنقر على الأيقونة  في شريط الأدوات أو بتأشير value label من القائمة view في شريط القوائم.
4. يظهر عنوان القيمة بدلاً من القيمة نفسها في مخرجات (جداول) برنامج SPSS .

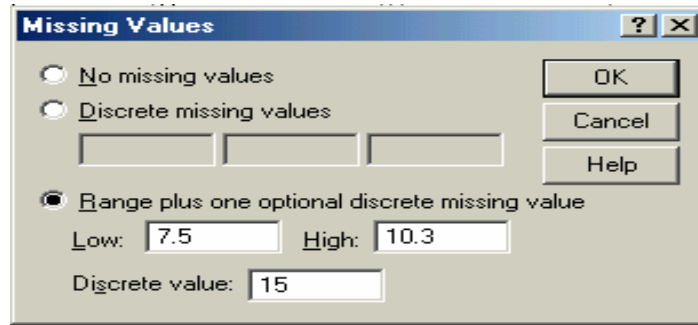
سابعاً : تعريف القيم المفقودة Missing Values

أحياناً نرغب في تعريف بعض قيم المتغير بأنها قيم مفقودة (أي إن هذه القيم موجودة أصلاً ولكننا لا نرغب إدخالها في التحليل الإحصائي كونها قيمة شاذة مثلاً أو بسبب كون السؤال لا ينطبق على المستجيب).
لتعريف القيم المفقودة لمتغير معين :

انقر زر الخلية التي تقع في عمود missing لهذا المتغير في ورقة variable view يظهر صندوق الحوار missing values الذي يحتوي الخيارات التالية :

- عدم وجود قيم مفقودة no missing values
- Discrete missing values يمكن تحديد ثلاثة قيم مفقودة كحد أعلى مثلاً القيم 100 ، 10 ، 12 سوف تعتبر مفقودة في حالة تحديدها.
- Range plus one discrete missing values مثلاً يتم تحديد القيم المفقودة الواقعة ضمن المدى بين 7.5 إلى 10.3 بالإضافة إلى إمكانية تحديد قيمة مفقودة واحدة اختياريًا كالقيمة 15 كما في صندوق الحوار التالي :

بعد إدخال القيم المفقودة انقر زر OK .



ملاحظات

1. يوجد نوعين من القيم المفقودة في برنامج spss
النوع الأول : هي القيم المفقودة التي تحدد من قبل المستخدم user- defined missing values (هي بالأصل ليست مفقودة) ويتم تعريفها بواسطة صندوق حوار missing values .
النوع الثاني : هي قيم المتغير المفقودة أصلاً (أي إنها خلايا فارغة) نتيجة عدم الاستجابة من قبل بعض الأشخاص لسؤال معين في استبيان ما و في هذه الحالة فإن الخلايا الفارغة تحول تلقائياً إلى قيم مفقودة للنظام system – missing values و هذا ينطبق على المتغيرات العددية إما بالنسبة للمتغيرات الرمزية string variables فإن الخلايا الفارغة تعتبر صحيحة valid أي أنها لا تعتبر قيماً مفقودة.
2. المديات Ranges في صندوق حوار missing values تستعمل لتحديد القيم المفقودة للمتغيرات العددية فقط أما المتغيرات الرمزية فيستعمل الخيار Discrete missing values معها.

ثامناً : عرض العمود column width

يمكن تحديد عرض العمود لمتغير معين بالوقوف على الخلية الواقعة ضمن العمود المعنون column في ورقة variable view حيث يمكن زيادة أو تقليل عرض العمود بواسطة الأسهم إلى الأعلى أو الأسفل (أو كتابة عرض العمود مباشرة).

Untitled - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

	Name	Type	Width	Decimals	Label	Values	Missing	
1	id	String	8	0		None	None	8
2	gender	Numeric	1	0		{1, m}...	None	8
3	bdate	Date	8	0	Date of Birth	None	None	8
4	grade	Numeric	3	0		None	None	8
5								

ملاحظات:

1. ان عرض العمود Column يمثل عدد الرموز المخصصة للعمود ويجب أن يكون عرض العمود أكبر أو يساوي عرض المتغير المضمن فيه .
2. يمكن تغيير عرض العمود لمتغير معين في ورقة Data view مباشرة عن طريق نقر و سحب حدود العمود clicking and dragging.

تاسعاً : محاذاة النص Alignment

لضبط محاذاة النص داخل خلايا المتغير انقر الخلية التابعة لمتغير معين في ورقة variable view الواقعة ضمن العمود المعنون Align ثم انقر السهم المتجه للأسفل لاختيار أمر مما يلي :-

Left : لمحاذاة النص إلى يسار الخلية.

Center : لمحاذاة النص إلى وسط الخلية.

Right : لمحاذاة النص إلى يمين الخلية.

علما ان المحاذاة الافتراضية هي (Right).

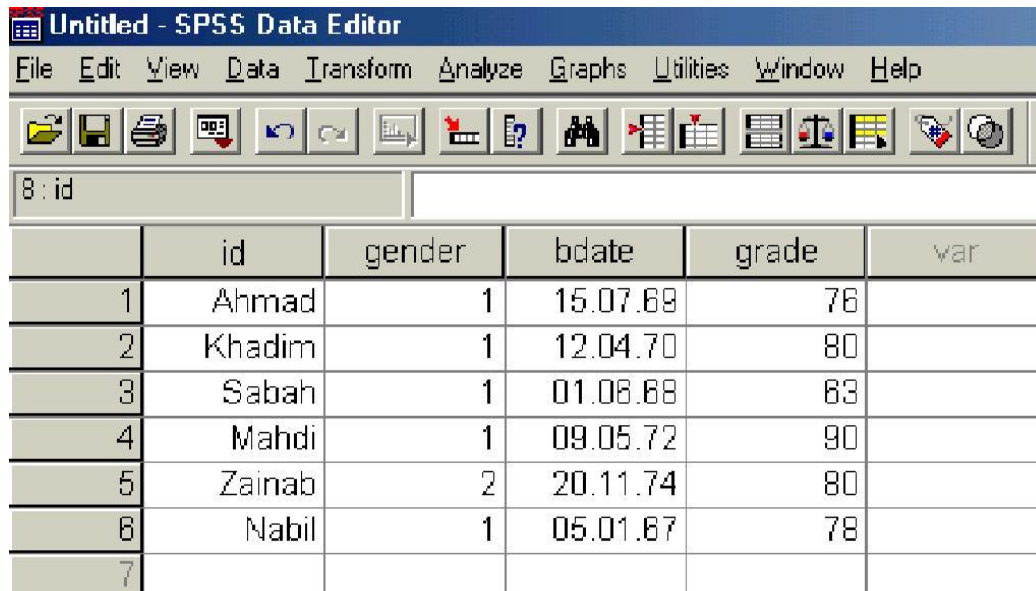
عاشراً : القياس Measurement

لغرض تعريف مقياس لمتغير معين انقر خلية المتغير التي تقع ضمن عمود measure في ورقة variable view حيث يظهر ثلاث خيارات

1. scale : يستعمل للبيانات العددية (القابلة للقياس الكمي) في قياس فترة أو نسبة وهذا المقياس المستعمل غالبا لقياس متغيرات الطول ، الوزن
2. ordinal : ويستعمل لقياس المتغيرات الترتيبية حيث أن هذا المتغير ذو عدد محدد من الفئات يمكن ترتيبها تصاعدياً أو تنازلياً ولكن لا يمكن تحديد الفروق بينها بدقة مثلاً تقدير طالب في امتحان (ممتاز ، جيد جداً ، جيد ، متوسط ، مقبول ، ضعيف) ويمكن أن يكون المتغير رمزياً أو عددياً على انه يفضل الأخير (عددي).
3. nominal : ويستعمل لقياس المتغيرات الاسمية وهي متغيرات لها عدد من الفئات دون أفضلية لاحداها على الأخرى (لا يمكن ترتيبها تصاعدياً أو تنازلياً) مثل تقسيم المجتمع إلى ذكور و إناث فمثلاً إذا رمزنا بالرقم 1 للذكور و الرقم 2 للإناث فإن هذين الرقمين لا يعطيان المعنى الحقيقي لهذا المتغير ولا يمكن إجراء العمليات الحسابية على هذا النوع من المتغيرات و يمكن ان يكون هذا النوع من المتغيرات عددياً أو رمزياً .مثلاً أيضاً متغير المحافظة (بغداد ، موصل ، بصرة) نستعمل معه nominal لعدم إمكانية ترتيب المحافظات بأسبوعية معينة .

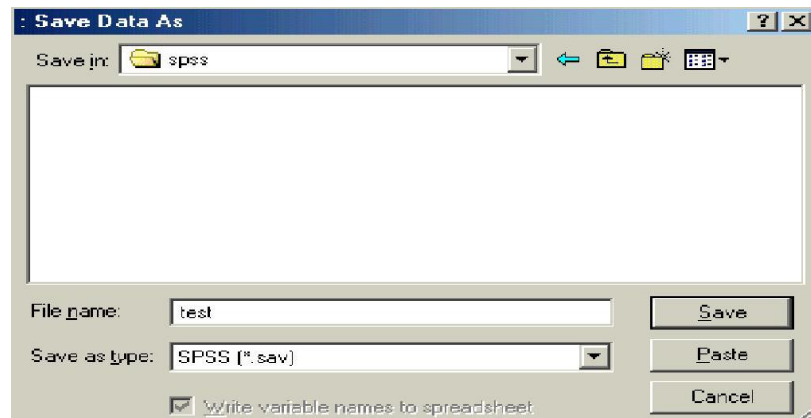
بعد تعريف المعلومات للمتغيرات الأربعة المذكورة في المثال تظهر شاشة variable view كما يلي :

عندها يمكن نقر عروة Data view للانتقال إلى إدخال البيانات في هذه الورقة حيث يمكن إدخال البيانات أما حسب المتغيرات أو حسب الحالات وذلك بالوقوف بالماوس على الخلية الأولى في المتغير لتصبح الخلية الفعالة Active cell وإدخال القيم و ضغط مفتاح enter أو مفاتيح الأسهم و الانتقال إلى الخلية الثانية و هكذا . وبعد إتمام عملية الإدخال تظهر ورقة Data view كما يلي :



	id	gender	bdate	grade	var
1	Ahmad	1	15.07.68	76	
2	Khadim	1	12.04.70	80	
3	Sabah	1	01.08.68	63	
4	Mahdi	1	09.05.72	90	
5	Zainab	2	20.11.74	80	
6	Nabil	1	05.01.67	78	
7					

لخزن الملف الذي تم تكوينه باسم test من القوائم نختار Save As → File فيظهر صندوق حوار Save Data As كما يلي :



نقوم بكتابة اسم الملف Test في المستطيل file name حيث يتم خزن الملف باسم test.sav حيث تكون الاستطالة لملفات المدخلات sav أما ملفات المخرجات فتكون ذات الاستطالة spo .
 لفتح ملف مخزن سابقا من القوائم نختار من شريط القوائم Open → file ثم نكتب اسم الملف المخزون و نوعه في صندوق حوار open . لخزن الملف (المفتوح) بنفس الاسم القديم بعد إجراء التغييرات عليه يتم ذلك بالأمر Save → File أو بواسطة المفاتيح **ctrl** + **s**

لخزن الملف بأسم جديد من القوائم نختار Save as → File .

(1-3) العمليات على المتغيرات وصفاتها في ورقة Data Editor

1. لاختيار select (تحديد أو تظليل) متغير variable ما انقر الخلية الحاوية على اسم المتغير في أعلى العمود في ورقة Data view بزر الماوس الأيسر.

2. لاختيار حالة case بأكملها انقر الخلية الحاوية على رقم الحالة في ورقة Data view بزر الماوس الأيسر.

3. لاختيار مجموعة من المتغيرات المتجاورة

◀ انقر الخلية الحاوية على اسم المتغير الأول.

◀ انقر الخلية الحاوية على اسم المتغير الأخير بعد ضغط مفتاح shift .

4. لاختيار مجموعة من المتغيرات المتباعدة

◀ انقر بزر الماوس الأيسر الخلية الحاوية على اسم المتغير الأول لاختياره.

◀ انقر الخلية الحاوية على اسم المتغير الثاني بعد ضغط مفتاح ctrl لاختياره وهكذا يتم اختيار بقية المتغيرات.

5. بنفس الطريقة المستخدمة لاختيار مجموعة من المتغيرات المتجاورة أو غير المتجاورة يمكن استعمالها لاختيار مجموعة من الحالات cases المتجاورة أو غير المتجاورة.

6. لإضافة متغير جديد بين متغيرين موجودين في Data view

◀ انقر بزر الماوس الأيسر اسم المتغير (أو أي خلية من خلايا المتغير) الذي يقع إلى يمين الموقع المراد إضافة المتغير الجديد إليه بعدها يمكن إضافة متغير بإحدى الطرق التالية :

الطريقة الأولى : من شريط القوائم اختر القائمة Data (الشكل التالي يبين محتويات القائمة Data) .



من القائمة أعلاه اختر insert variable فيضاف متغير جديد الى يسار المتغير الحالي حيث يمكن

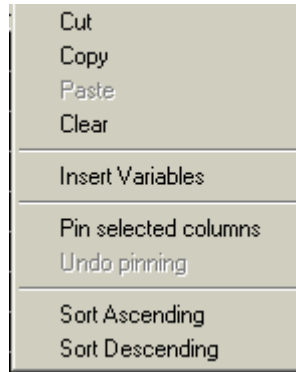
تغيير اسمه.

الطريقة الثانية : انقر الأيقونة  في شريط الأدوات Tool Bar فيضاف متغير جديد الى

يسار المتغير الحالي .

كما يمكن إضافة متغير جديد بنقر اسم المتغير الذي يقع الى يمين الموقع المراد إضافة المتغير الجديد إليه بزر الماوس الأيمن فتظهر القائمة المختصرة short list التالية :

و منها نختار الأمر insert variable فيضاف متغير جديد إلى يسار المتغير الحالي.

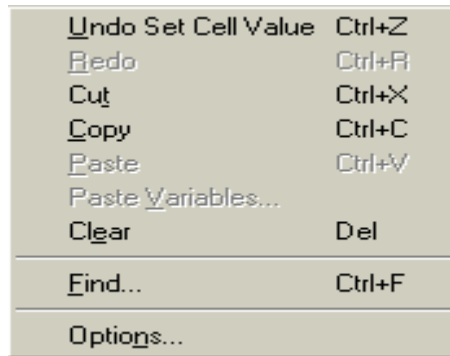


7. بنفس الطريقة التي استخدمناها لإضافة متغير يمكن إضافة حالة case فوق (أعلى) حالة موجودة (بعد تحديدها).

8. لحذف متغير

➤ انقر اسم المتغير في ورقة Data view بزر الماوس الأيسر لاختياره .

➤ من شريط القوائم اختر القائمة Edit التي تضم الخيارات التالية



➤ اختر الأمر clear فيتم حذف المتغير المختار من ورقة Data view .

. يمكن حذف المتغير بعد اختياره مباشرة باستعمال مفتاح Del .

و كطريقة أخرى لحذف متغير :

➤ انقر اسم المتغير بزر الماوس الأيمن فتظهر القائمة المختصرة short list ومنها نختار الأمر

clear فيتم حذف المتغير .

9. نفس طرق حذف متغير الواردة في الفقرة 8 يمكن استعمالها في حذف الحالة case .

10. لعمل نسخة من متغير معين copy اتبع الخطوات التالية:

➤ انقر اسم المتغير (المصدر) لتحديده .

➤ من شريط القوائم اختر Edit ← copy ثم انقر أسم المتغير الذي يراد نسخ المتغير المصدر إليه

➤ من شريط القوائم اختر paste → Edit .

كطريقة ثانية لنسخ متغير

- انقر اسم المتغير (المصدر) بزر الماوس الأيمن (لإظهار القائمة المختصرة short list) ثم اختر copy من القائمة المختصرة.
- انقر اسم المتغير الذي يراد نسخ المتغير (المصدر) إليه بزر الماوس الأيمن ثم اختر paste من القائمة المختصرة .

11. لتغيير موضع متغير معين

- انقر اسم المتغير (المصدر) لتحديده .
 - من شريط القوائم اختر cut Edit
 - انقر اسم المتغير الذي يراد تحريك المتغير (المصدر) إليه.
 - من شريط القوائم اختر paste Edit →
- كطريقة ثانية لتحريك متغير يمكن استعمال القائمة المختصرة كما هو في حالة النسخ copy .

12. للانتقال الى حالة معينة

- من شريط القوائم نختار Go to case → Data.
 - يظهر صندوق حوار Go to case .
 - نقوم بإدخال رقم الحالة و نقر زر ok .
- ويمكن إنجاز الفعالية السابقة بنقر الأيقونة  .
13. يمكن نسخ صفات متغير Attributes الى متغير (أو متغيرات أخرى) والتي تشمل (النوع type ، العرض width ، عدد المراتب العشرية الخ) باتباع الخطوات التالية:

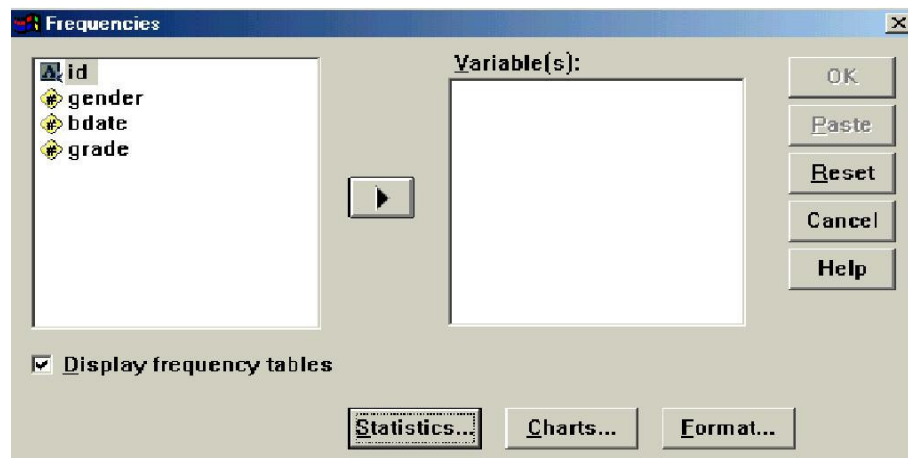
- في ورقة variable view انقر اسم المتغير الذي تريد نسخ صفاته إلى متغير آخر .
 - من شريط القوائم اختر copy Edit → .
 - انقر اسم المتغير الذي تريد نسخ الصفات إليه.
 - من شريط القوائم اختر paste Edit → .
- كما يمكن نسخ صفة واحدة للمتغير كما يلي:
- انقر خلية الصفة المراد نسخها مثلاً النوع type للمتغير المصدر.
 - اختر copy Edit → من شريط القوائم .
 - انقر خلية الصفة للمتغير المراد نسخ الصفة إليه.
 - اختر paste Edit → .

(1 - 4) استعمال مجاميع جزئية من المتغيرات

في بعض الأحيان يكون عدد المتغيرات في ورقة نثر Data Editor كبيراً جداً ويكون من الصعوبة التعامل مع هذا العدد الكبير من المتغيرات ولهذا يكون من المفضل تكوين مجاميع جزئية تحتوي على عدد صغير من المتغيرات مما يسهل من عملية إيجادها ويقلل من الوقت اللازم لذلك .

بالنسبة للمثال السابق فقد ضمن ورقة النثر أربعة متغيرات هي id ، gender ، bdate ، grade فعند استخدام أي أسلوب إحصائي فأن كافة هذه المتغيرات ستظهر في القائمة المصدر Source List فمثلاً لو

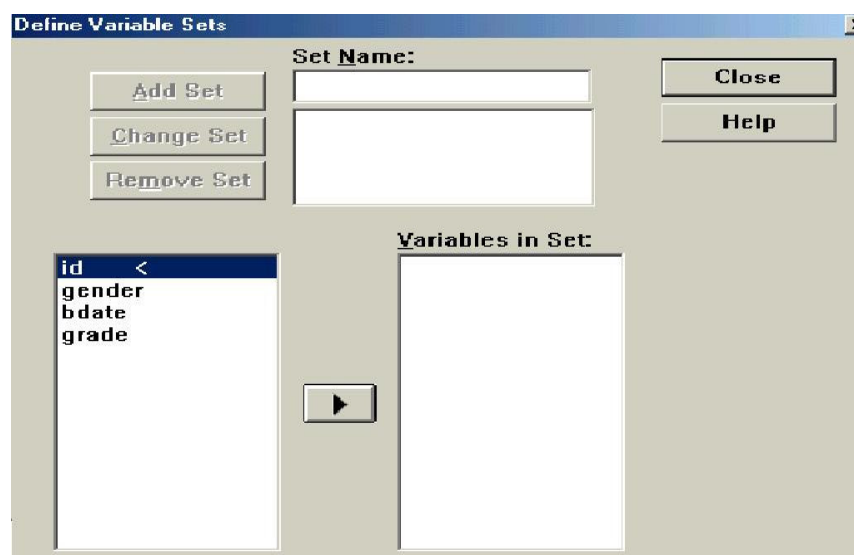
اخترنا من شريط القوائم الأسلوب Frequencies → Descriptive Statistics → Analyze فأن صندوق حوار الأمر Frequencies يظهر كما يلي :



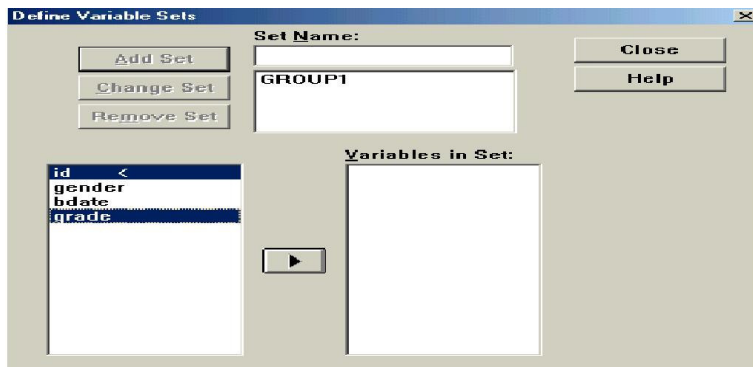
حيث أن جميع المتغيرات قد ظهرت في القائمة المصدر Source List في الجهة اليسرى من الصندوق حيث من الممكن أن نقوم بنقل أي من المتغيرات التي نرغب في إجراء العمليات الإحصائية عليها إلى قائمة Variables بنقر الزر  . إذا أردنا أن يظهر المتغيرين id و grade فقط في Source List أعلاه نتبع الخطوات التالية :

الخطوة الأولى

من شريط القوائم في شاشة Data Editor (لايهم Variable View أو Data View) أختار الأمر Define sets → Utilities فيظهر صندوق الحوار التالي :



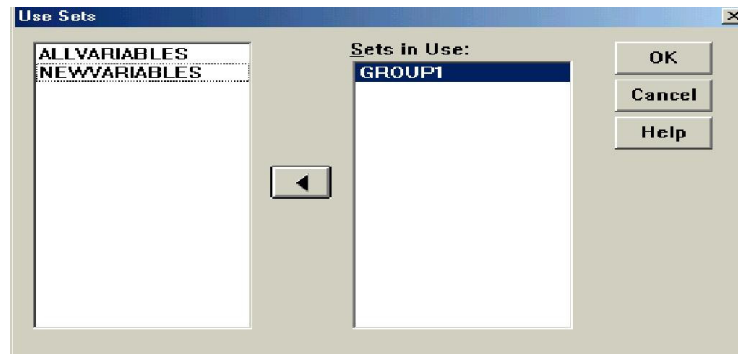
- لتعريف مجموعة باسم Group1 التي تحتوي المتغيرين id و grade نتبع الخطوات التالية :
- نكتب أسم المجموعة Group1 في قائمة Set Name في صندوق الحوار أعلاه .
 - إدخال المتغيرين id و grade الى قائمة Variables in set بنقر كل متغير بزر الماوس الأيسر ثم نقر الزر  .
 - أنقر الزر Add Set . (الزر Remove set لحذف مجموعة والزر Change لتغيير عناصر مجموعة) .
- فيظهر صندوق الحوار أعلاه بعد الترتيب كما يلي :
- عند نقر زر Close نحصل على مجموعة جزئية من المتغيرات باسم Group1 تحتوي المتغيرين id و gender .



الخطوة الثانية

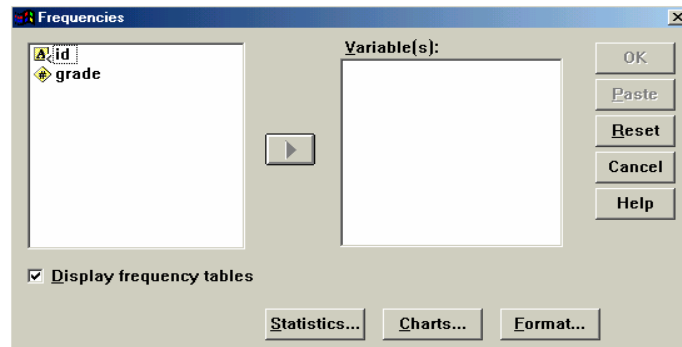
أن ناتج الخطوة الأولى لا يغير من عدد المتغيرات المضمنة في التحليل الإحصائي (وهي أربعة متغيرات) والتي تظهر كما هي في صندوق حوار Frequencies المذكور أنفاً لتحديد عدد المتغيرات بمتغيرين اثنين هما id و grade نتابع الخطوات التالية :

- من شريط القوائم اختر Use Sets → Utilities فيظهر صندوق حوار Use Sets الذي نرتبه كما يلي :



فقد قمنا بنقل Group1 من القائمة في يسار الصندوق الى قائمة Sets in Use كما نقلنا أيضاً مجموعتي All variables (مجموعة تحتوي كافة المتغيرات) و New Variables (مجموعة تحتوي المتغيرات الجديدة التي تم تكوينها في الجلسة الحالية) من قائمة Sets in Use الى القائمة في جهة اليسار. < عند نقر زر OK فإن المتغيرات المضمنة في مجموعة Group1 فقط (id و gender) ستستعمل في التحليل الإحصائي .

عند اختيار الأمر Frequencies → Descriptive Statistics → Analyze مثلاً فإن صندوق حوار frequencies سيحتوي المتغيرين id و gender فقط بدلاً من احتوائه على أربعة متغيرات وكما يلي :



الفصل الثاني

أوامر القائمتين Data و View

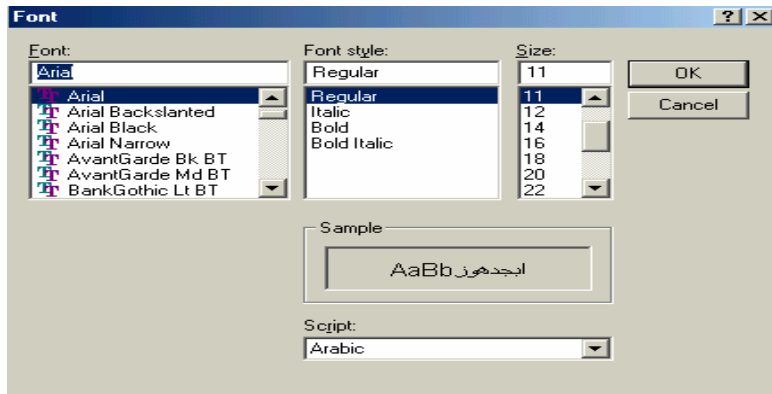
(2 - 1) أوامر القائمة View

يمكن إنجاز فعاليات مختلفة باستخدام قائمة view و التي تضم الخيارات التالية :



حيث أن :

1. status bar : لعرض شريط الحالة (أسفل الشاشة) عند نقره بزر الماوس الأيسر .
2. Font : لتغيير حجم ونوع الخط المستعمل عند الكتابة في شاشة Data Editor فعند نقر الأمر Font بزر الماوس الأيسر يظهر صندوق حوار Font التالي :



حيث أن :

القائمة Font : لتحديد نوع الخط المستعمل : تتوفر أنواع عديدة من الخطوط التي يمكن الاستفادة منها كما تتوفر خطوط للكتابة باللغة العربية مثل Arial ، Arabic Transparent ، Andalus ، Akhbar MT .

القائمة Font Style : لتحديد نمط الخط :

عادي Regular

مائل *Italic*

غامق **Bold**

غامق مائل ***Bold Italic***

القائمة size : لتحديد حجم الخط .

ملاحظة : يمكن تغيير نوع الأرقام المستعملة (عربية ، هندية ، موضعية) من قائمة start حسب التسلسل التالي :

Start → Settings → Control panel → Regional setting → Number style

بعدها نختار نوع الأرقام (context , Hindi ,Arabic).

3. Gridlines : لعرض خطوط الشبكة في Data Editor .

4. value labels : لعرض عناوين القيم لمتغير ما (في حال تعريفها) . علماً أنه يمكن استعمال

الأيقونة  لنفس الغرض .

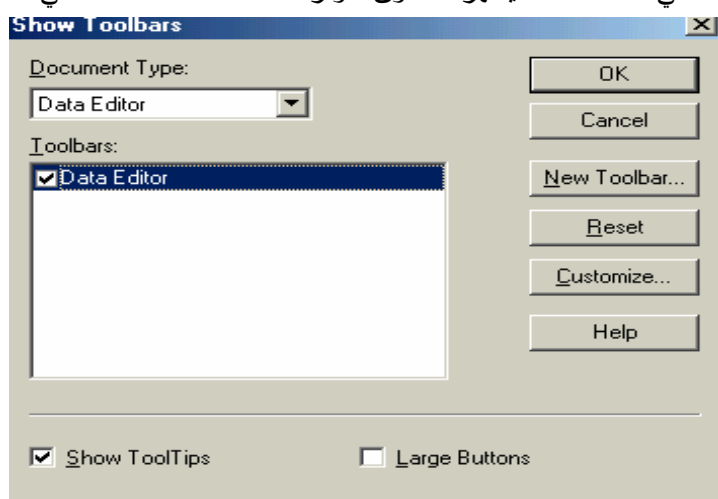
5. variables : للانتقال إلى ورقة variable view و بالعكس أي الرجوع إلى ورقة Data view .

6. toolbars : يستعمل هذا الأمر لتحقيق أحد هدفين

أ. إضافة أيقونات إلى شريط الأدوات القياسي standard toolbar .

ب. إنشاء شريط أدوات جديد (غير القياسي) .

عند نقر toolbars في قائمة view يظهر صندوق حوار Show Toolbars التالي :



حيث أن :

Document Type : يبين نوع الوثيقة (الوثيقة في صندوق الحوار الحالي هي ورقة Data Editor) . بنقر

السهم المتجه للأسفل يمكن أن نختار أحد الأنواع التالية من الوثائق :

1. all كافة الوثائق.

2. Data Editor وثيقة محرر البيانات.

3. Viewer شاشة عرض النتائج الإحصائية.

4. Draft Viewer مسودة شاشة العرض.

5. chart شاشة عرض التخطيطات.

6. syntax وثيقة كتابة البرامج.

7. script Editor محرر الوثائق.

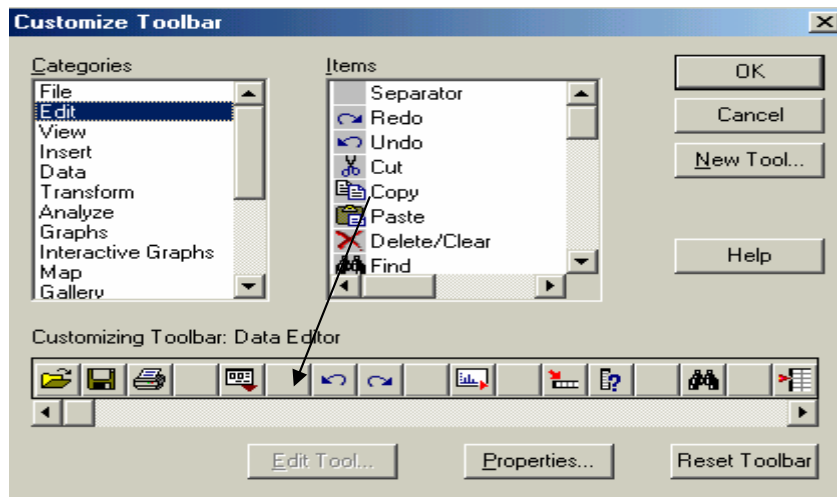
حيث أن لكل نوع من الوثائق (الأوراق أو الشاشات) أعلاه عدد معين من أشرطة الأدوات الخاصة

به .

Toolbars : يبين أشرطة الأدوات القياسية المتوفرة لكل من الأنواع السابقة فمثلا بالنسبة لورقة Data Editor يتوفر شريط أدوات قياسي واحد فقط هو شريط أدوات Data Editor كما هو واضح في صندوق الحوار السابق حيث يمكن إزالة هذا الشريط عن طريق نقر Check box المجاور لـ Data Editor أي إزالة العلامة .

Show ToolTips : لإظهار تعليق توضيحي عند تمرير الماوس على أيقونات شريط الأدوات Toolbar يوضح وظيفة الأيقونة (الأداة). إزالة العلامة يؤدي إلى عدم إظهار التعليق .
Large Buttons : عند نقر checkbox المجاور بزر الماوس الأيسر يقوم البرنامج بإظهار أيقونات شريط الأدوات ذات حجم كبير .

مثال 1: لإضافة أيقونة Copy إلى شريط الأدوات القياسي لورقة Data Editor نتبع الخطوات التالية:
✎ اختر toolbar → view من شريط القوائم فيظهر صندوق حوار show toolbars السابق .
✎ انقر زر customize في صندوق حوار show toolbars فيظهر صندوق حوار customize toolbar .
✎ انقر أيقونة Copy ضمن مفردات Edit حيث يتحول مؤشر الماوس إلى ما يشبه "قبضة اليد" استعمل هذه القبضة في سحب الأيقونة إلى شريط الأدوات في الأسفل كما في الشكل التالي:

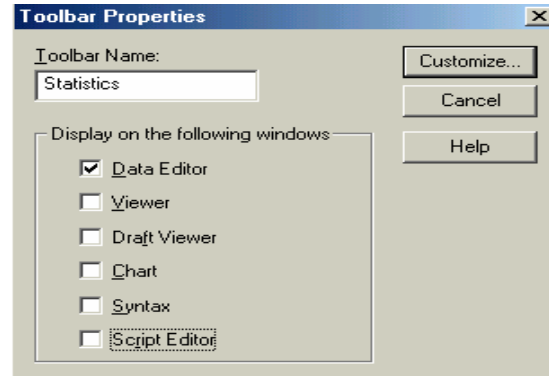


✎ عند نقر زر OK يظهر شريط الأدوات الجديد وقد أضيفت إليه أيقونة Copy.
ملاحظة :- يمكن إرجاع شريط الأدوات القياسي (بدون Copy) بعد تأشير الشريط toolbar (في هذا المثال شريط Data Editor) في قائمة Toolbars في صندوق حوار Show Toolbars ثم نقر زر Reset و زر OK .

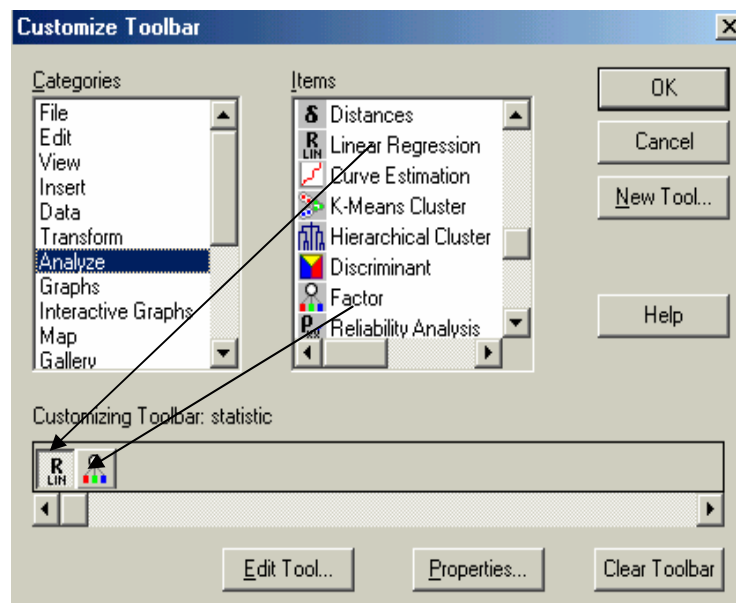
مثال 2 : (أعداد شريط أدوات جديد New Toolbar)

لغرض إضافة شريط أدوات جديد يحتوي عمليتين إحصائيتين Linear Regression و Factor لوثيقة Data Editor التي تحتوي شريط قياسي واحد هو Data Editor Toolbar نتبع الخطوات التالية:
✎ من القوائم اختر toolbar → view يظهر صندوق حوار Show Toolbars السابق .

انقر زر New Tool يظهر صندوق حوار properties : Toolbar .حيث نقوم بكتابة اسم شريط الأدوات الجديد وهو statistics في حقل Toolbar Name ثم نحدد الوثيقة التي يظهر فيها هذا الشريط نفترض إننا نريد عرضه في وثيقة Data Editor فقط كما في الشكل التالي:

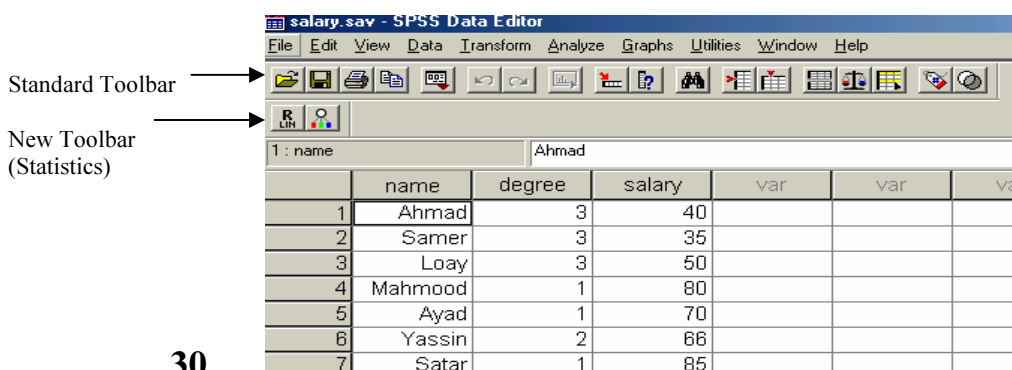


عند نقر زر customize يظهر صندوق حوار Toolbar customize : حيث نقوم بسحب أيقونة Linear Regression وأيقونة Factor بزر الماوس الأيسر إلى شريط الأدوات الجديد (statistics) اسفل الصندوق كما في الشكل التالي:



عند نقر زر OK يضاف شريط أدوات جديد باسم statistic إلى وثيقة Data Editor أي يصبح لدينا شريط أدوات قياسي هو شريط Data Editor و شريط أدوات غير قياسي هو شريط أدوات statistics

كما في الشكل التالي:



(2 - 2) أوامر القائمة Data : تضم القائمة Data الأوامر التالية :

1. تعريف التاريخ للسلسلة الزمنية Define Date : يقوم هذا الأمر بتوليد متغيرات التاريخ date Variables التي تستعمل كتاريخ لقيم السلسلة الزمنية فقط أي أن متغير التاريخ لا يستعمل في العمليات على السلسلة الزمنية التي تتجزأ باستخدام الأمر Create Time series التابع لقائمة Transform .

مثال

المتغير revenue يمثل عائدات منشأة معينة للفترة من (الشهر السادس /2000 ولغاية شهر كانون الثاني / 2002) ويظهر في شاشة Data editor كما في الشكل المجاور ¹ :
يطلب تعريف التاريخ للمتغير revenue حسب السنة والفصل والشهر خلال الفترة المذكورة .

لتنفيذ ذلك نتبع الخطوات التالية :

من شريط القوائم نختار Define Dates → Data فيظهر صندوق الحوار التالي الذي نقوم بترتيبه كما يلي :

revenue
117
120
130
145
150
190
220
250
243
257
260
340
360
362
380
340
350
420
389
400

لاحظ إننا اخترنا Years,quarters,months من قائمة cases ومن بين الخيارات الأخرى المتوفرة مايلي :

- years : لعرض قيم السلسلة الزمنية حسب السنوات فقط .
- years,quarters : لعرض قيم السلسلة الزمنية حسب السنوات والفصول .
- years,months : لعرض قيم السلسلة الزمنية حسب السنوات و الأشهر .
- Not dated : لإزالة أي متغير تاريخ سبق تعريفه .
- Custom : لعرض قيم السلسلة الزمنية حسب تاريخ معين يحدد من قبل المستخدم (غير متوفر ضمن خيارات قائمة Cases are) مثلاً أسبوع يتكون من أربعة أيام عمل وهذا يتم عادة باستخدام أوامر Syntax .

¹ في هذا المثال والأمثلة القادمة سنعتبر أن البيانات الواردة في المثال قد أدخلت في شاشة Data Editor بنفس الكيفية التي نعرضها تماماً فيالنسبة للملف أعلاه يتضمن عموداً واحداً باسم Revenue وفي شاشة Data Editor يمثل هذا العمود بمتغير اسمه Revenue أما الأرقام فتمثل الحالات وعددها 20 حالة .

في قائمة First Case is نقوم بتعريف التاريخ للحالة الأولى في السلسلة الزمنية: في حقل السنة ندخل سنة البداية 2000 .

في حقل الفصل ندخل فصل البداية 2 (شهر حزيران يقع في الفصل الثاني من السنة) .
في حقل الشهر ندخل شهر البداية 6 .

لاحظ أنه لا يمكن تنفيذ الأمر في حالة وجود تعارض بين الفصول والأشهر مثلاً استعمال الفصل 3 بدلاً من الفصل 2 لشهر حزيران.

أما Periodicity at higher level فيبين دورية التاريخ أو انه يعطي أكبر قيمة يمكن تزويدها للبرنامج مثلاً أعلى دورية للفصول هي 4 وأعلى دورية للأشهر هي 2 .

عند نقر OK في صندوق حوار Define Dates تضاف متغيرات التاريخ الى Data Editor كما يلي :

revenue	year_	quarter_	month_	date_
117	2000	2	6	JUN 2000
120	2000	3	7	JUL 2000
130	2000	3	8	AUG 2000
145	2000	3	9	SEP 2000
150	2000	4	10	OCT 2000
190	2000	4	11	NOV 2000
220	2000	4	12	DEC 2000
250	2001	1	1	JAN 2001
243	2001	1	2	FEB 2001
257	2001	1	3	MAR 2001
260	2001	2	4	APR 2001
340	2001	2	5	MAY 2001
360	2001	2	6	JUN 2001
362	2001	3	7	JUL 2001
380	2001	3	8	AUG 2001
340	2001	3	9	SEP 2001
350	2001	4	10	OCT 2001
420	2001	4	11	NOV 2001
389	2001	4	12	DEC 2001
400	2002	1	1	JAN 2002

2. الأمر Insert Variable : يستعمل لإضافة متغير الى يسار المؤشر في Data Editor وهي نفس الوظيفة التي تؤديها الأيقونة  في شريط الأدوات القياسي (راجع الفصل الأول) .

3. الأمر Insert Case : يستعمل لأضافة حالة Case أعلى المؤشر في Data Editor وهي نفس الوظيفة التي تؤديها الأيقونة  (راجع الفصل الأول) .

4. الأمر Go to Case : يستعمل للذهاب الى حالة معينة بتحديد رقمها Case Number وهي نفس الوظيفة التي تؤديها الأيقونة  .

5. الأمر Sort Cases : يستعمل لترتيب حالات ملف ما ترتيباً تصاعدياً أو تنازلياً حسب متغير(متغيرات) ترتيب معين Sorting Variable .

مثال (على الأمر Sort) :

الملف salary يحتوي أسماء مجموعة من الموظفين ودرجاتهم الوظيفية degree والأجر الشهري salary وكما يظهر في شاشة Data Editor في الشكل المجاور :

الملف Salary قبل الترتيب		
name	degree	salary
Ahmad	3	40
Samer	3	35
Loay	3	50
Mahmood	1	80
Ayad	1	70
Yassin	2	66
Satar	1	85
Razak	1	77
Kamal	2	59
Abas	3	45
Mahdi	1	90
Salim	2	62
Sabah	2	57
Falah	2	55
Imad	1	82

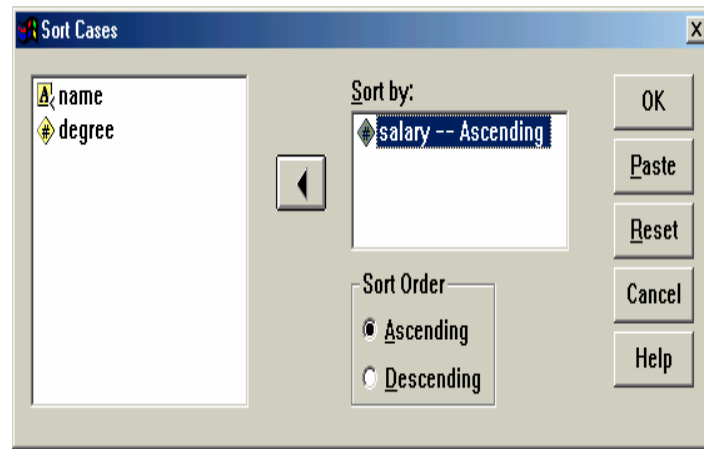
سنأخذ حالتين لترتيب الملف وكالتالي :

أ. الحالة الأولى : لترتيب الملف تصاعدياً حسب المتغير

Salary نتبع الخطوات التالية :

من شريط القوائم اختر Sort Cases → Data فيظهر

صندوق حوار Sort Cases الذي نرتبه كالآتي :



عند نقر زر ok يتم ترتيب الملف تصاعدياً حسب المتغير Salary وكما يلي :

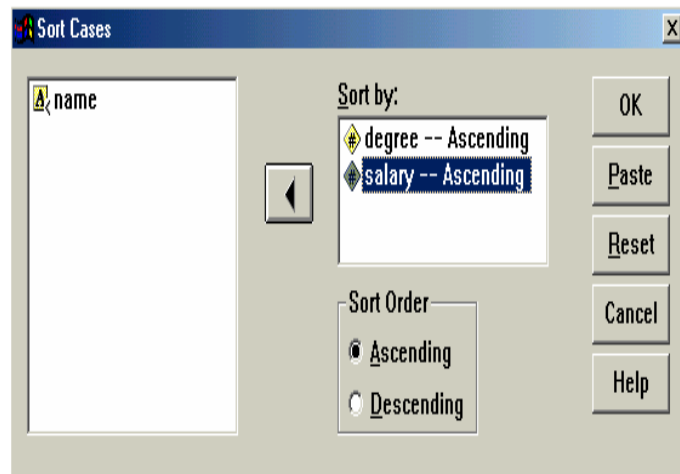
ب. الحالة الثانية : لترتيب الملف salary تصاعدياً حسب المتغير Salary ضمن فئات الدرجة

الوظيفية degree نتبع الخطوات التالية :

من شريط القوائم اختر Sort Cases → Data فيظهر

صندوق حوار Sort Cases الذي نرتبه كالآتي :

الملف المرتب حسب المتغير salary		
الحالة الأولى		
name	degree	salary
Samer	3	35
Ahmad	3	40
Abas	3	45
Loay	3	50
Falah	2	55
Sabah	2	57
Kamal	2	59
Salim	2	62
Yassin	2	66
Ayad	1	70
Razak	1	77
Mahmood	1	80



عند نقر زر OK نحصل على الترتيب التالي :

6. الأمر Transpose : يستعمل لقلب الصفوف إلى أعمدة وبالعكس أي تحويل المتغيرات Variables إلى حالات Cases وبالعكس .

مثال :

الملف التالي يحتوي المصفوفة المعرفة بأعمدتها بالمتغيرات x1,x2,x3 إضافة إلى متغير تسمية y وقد أدخلت البيانات في شاشة Data Editor وكما يلي :

x1	x2	x3	y
3	6	9	y1
4	7	10	y2
5	8	11	y3

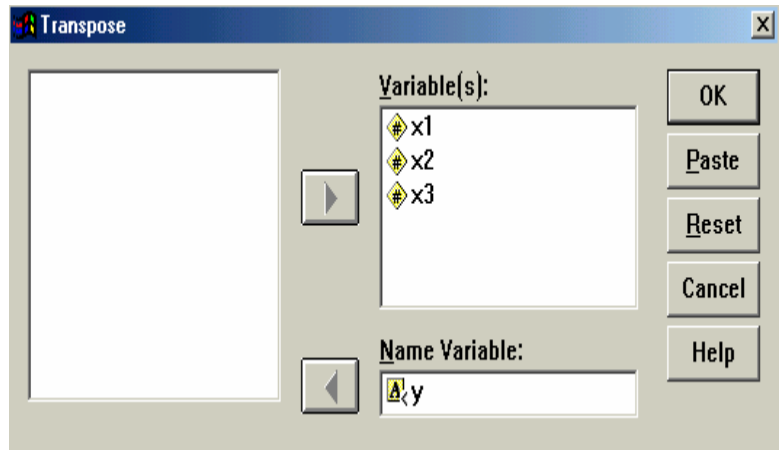
الملف المرتب تصاعدياً حسب المتغير		
salary	ضمن الدرجة الوظيفية	
الحالة الثانية		
Ayad	1	70
Razak	1	77
Mahmood	1	80
Imad	1	82
Satar	1	85
Mahdi	1	90
Falah	2	55
Sabah	2	57
Kamal	2	59
Salim	2	62
Yassin	2	66
Samer	3	35
Ahmad	3	40
Abas	3	45
Loay	3	50

يطلب إيجاد المبدلة للمصفوفة x وتسمية أعمدتها بالمتغير y

لتنفيذ ذلك نتبع الخطوات التالية :

من شريط القوائم اختر Transpose → Data

فيظهر صندوق حوار Transpose الذي نرتبه كما يلي :



عند نقر زر OK يتم الحصول على مبدلة المصفوفة وتظهر في شاشة Data Editor كما يلي :

case_lbl	y1	y2	y3
X1	3	4	5
X2	6	7	8
X3	9	10	11

حيث ان y1,y2,y3 تمثل أسماء أعمدة مبدلة المصفوفة وأن case_lbl هو متغير رمزي يكون تلقائياً ويمثل أسماء المتغيرات القديمة (أعمدة المصفوفة قبل استخراج المبدلة) .

7. دمج الملفات Merge files

يستعمل هذا الأمر لدمج ملفين وهذه العملية مهمة جداً في حالة استخدام برنامج SPSS كقاعدة بيانات

ويمكن أن يتم الدمج بإحدى الطريقتين التاليتين :

أ. إضافة حالات Add Cases ب. إضافة متغيرات Add Variables

أ. إضافة حالات Add Cases : أن هذا الأمر يتيح دمج ملفين يحتويان نفس المتغيرات وحالات مختلفة مثلاً دمج ملف درجات طلبة شعبة أ مع ملف درجات طلبة شعبة ب حيث يتشابه الملفان من حيث الدروس ويختلفان من حيث الطلبة في كل شعبة (الحالات) .

مثال : الملف Group1 يحتوي درجات طالبين (حالتين) في أربعة مواد (variables) و الملف Group2 يحتوي درجات مجموعة ثانية من الطلاب (ثلاثة حالات) في أربعة مواد ويظهر كل منهما كما يلي في شاشة : Data Editor

	Group1			
	name	math	chem	physc
	Samir	100	90	95
	Lubna	95	87	90

	Group2			
	name	math	chem	physc
	Yousif	85	90	77
	Ammar	95	83	82
	Sinan	90	92	86

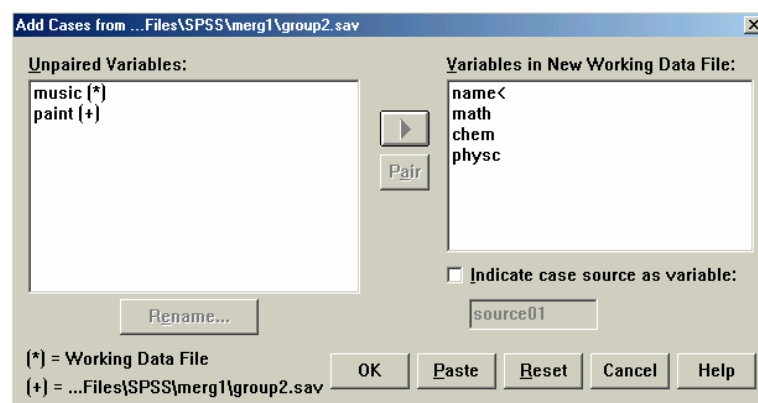
لدمج الملفين (إضافة حالات الملف الثاني إلى حالات الملف الأول) نتبع الخطوات التالية :

أفتح أحد الملفين (بالأمر open) وليكن الملف Group1 يعرف هذا الملف بأسم الملف العامل

working data file .

من شريط القوائم اختر Add Cases → Merge Files → Data فيظهر صندوق حوار Add Cases : Read File ومنه نختار الملف Group2 المخزون مسبقاً ويعرف بالملف الخارجي External Data File ، ونلاحظ أن الملفين يختلفان في مادة واحدة هي music في الملف الأول و paint في الملف الثاني .

وعند نقر زر open في صندوق حوار Add Cases : Read File يظهر صندوق الحوار التالي :



يشمل الصندوق المكونات التالية :

Variables in New Working Data File : وهي المتغيرات التي سوف تضمن في الملف الناتج من عملية الدمج حيث يتم تضمين كافة المتغيرات من كلا الملفين التي تتطابق من ناحية الاسم ونوع المتغير (عددي أم رمزي) ، حيث أنه بالإمكان حذف أي متغير من هذه القائمة وعدم تضمينه في الملف المدمج .

Unpaired Variables: تحتوي هذه القائمة أسماء المتغيرات التي سوف لن تضمن في الملف الناتج عن عملية الدمج merged File .

ويرمز للمتغيرات من الملف العامل Working Data File بالرمز *

ويرمز للمتغيرات من الملف الخارجي External Data File بالرمز +

وتشمل القائمة Unpaired Variables المتغيرات التالية :

- المتغيرات من الملفين المدمجين التي لا تتطابق من ناحية الاسم .
 - المتغيرات التي تم تعريفها كمتغيرات عددية في أحد الملفين وكمتغيرات رمزية في الملف الآخر حيث لا يمكن دمج المتغيرات العددية بالمتغيرات الرمزية .
 - المتغيرات الرمزية التي لها أطوال غير متساوية Unequal Width في كلا الملفين .
- في هذا المثال نلاحظ أن المتغيرين music في ملف Group1 و paint في ملف Group2 قد ضمنا في هذه القائمة بسبب عدم تطابق أسمائهما في الملفين .
- ملاحظة 1: يمكن تغيير اسم أي متغير في هذه القائمة بتأشير المتغير (نقره بزر الماوس الأيسر) ثم نقر زر Rename وإعطاء اسم جديد لهذا المتغير .

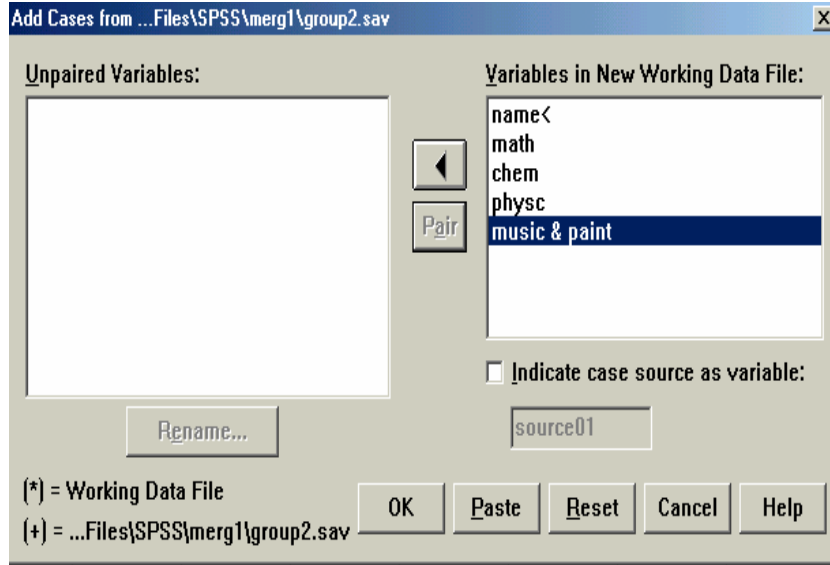
Indicate Case source variable: عند تأشير هذا الخيار يتم إضافة متغير جديد بأسم Source01 في الملف المدمج ويأخذ القيمة 0 لحالات الملف العامل والقيمة 1 لحالات الملف الخارجي .

عند نقر زر ok في صندوق حوار Add Cases From يتم دمج الملفين ويظهر الملف المدمج في شاشة Data editor كما يلي حيث يمكن إعطائه الاسم المرغوب (بالأمر Save As) :

Merged File			
name	math	chem	physc
Samir	100	90	95
Lubna	95	87	90
Yousif	85	90	77
Ammar	95	83	82
Sinan	90	92	86

ملاحظة 2: يمكن نقل أي زوج من المتغيرات في قائمة Unpaired إلى قائمة Variables in New Working Data File على شرط أن يكون أحد المتغيرين من الملف العامل *والآخر من الملف الخارجي + فننقل الملفين music و paint للذان يظهران في قائمة Unpaired في صندوق حوار Add Cases from نتبع الخطوات التالية :

1. نختار كلا المتغيرين (أنقر المتغير الأول ثم أضغط مفتاح CTRL مع النقر على اسم المتغير الثاني) .
2. أنقر زر Pair لنقل المتغيرين بلسم جديد هو music & paint حيث يظهر صندوق حوار Add Cases from كما يلي :



3. عند نقر زر OK في صندوق الحوار أعلاه يتم دمج الملفين كما يلي :

Merged File

name	math	chem	physc	music
Samir	100	90	95	87
Lubna	95	87	90	85
Yousif	85	90	77	88
Ammar	95	83	82	90
Sinan	90	92	86	95

يلاحظ أنه تم إعطاء أسم music للمتغير الناتج عن عملية الدمج .

ملاحظة 3 : يمكن تضمين أي من المتغيرين music و paint أو كلاهما في قائمة Variables in New Working Data File باختيار المتغير المطلوب (بنقره بزر الماوس الأيسر) ثم نقر زر  في صندوق حوار Add Cases From وبعد ذلك نقر زر OK (إذا رغبتنا بتضمين كلا المتغيرين) سنحصل على الملف المدمج الذي سيحتوي قيما مفقودة للحالات المضمنة في الملف الذي لا يحتوي المتغير المعني وكما يلي :

Merged File

name	math	chem	physc	music	paint
Samir	100	90	95	87	.
Lubna	95	87	90	85	.
Yousif	85	90	77	.	88
Ammar	95	83	82	.	90
Sinan	90	92	86	.	95

ب. أضافة متغيرات Add Variables

يتيح هذا الأمر إمكانية دمج الملف العامل (الحالي) مع الملف الخارجي والذان يحتويان نفس الحالات ولكن متغيرات مختلفة.

مثال 1

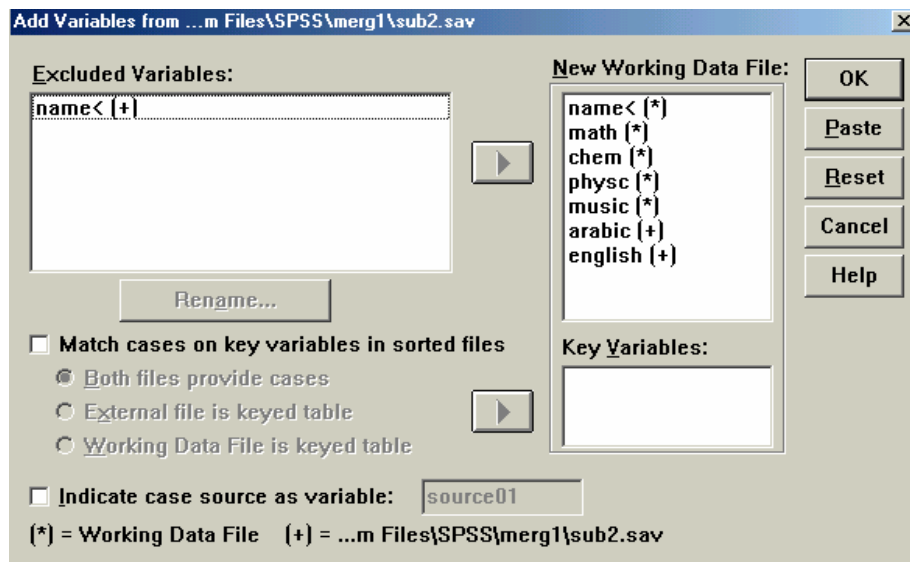
الملف Group1 يتضمن درجات طالبين في أربعة مواضيع والملف Sub2 يحتوي درجات نفس الطالبين في موضوعين آخرين وكما يلي :

Group1				
name	math	chem	physc	music
Samir	100	90	95	87
Lubna	95	87	90	85

sub2		
name	arabic	english
Samir	80	98
Lubna	85	95

لدمج الملفين (إضافة متغيرات الملف الثاني الى متغيرات الملف الأول) نتبع الخطوات التالية:

- أفتح ملف Group1 المخزون سابقاً بالأمر open (الملف العامل) .
- من شريط القوائم اختر Merge Files → Add Variables → Data فيظهر صندوق حوار Add Variables : Read File ومنه نختار الملف Sub2 المخزون مسبقاً (الملف الخارجي) .
- وعند نقر زر open في صندوق حوار Add Variables : Read File يظهر صندوق حوار Add Variables From وكما يلي :



يتضمن الصندوق مايلي :

- New Working data File** : وهي قائمة بأسماء المتغيرات التي سوف تضمن في الملف الجديد المدمج ويتم تضمين كافة المتغيرات التي لا تتشابه من ناحية الاسم في كلا الملفين المدمجين .
- Excluded Variables** : وهي قائمة بأسماء المتغيرات التي تستبعد من الملف المدمج الجديد وعادة تتضمن أسماء المتغيرات من الملف الخارجي sub2 التي تشابه أسماء متغيرات الملف العامل Group1 ونلاحظ ان القائمة تتضمن المتغير name (+) نظراً لتكراره في كلا الملفين .
- Key Variables** المتغيرات المفاتيح : تستعمل هذه المتغيرات في حالة عدم تطابق بعض الحالات في كلا الملفين المراد دمجهما مثلاً احتواء الملفين على بعض الحالات غير المتشابهة أو توجد حالات مفقودة . علماً أنه يتوجب توفر الشروط التالية في Key Variables .

1. يجب أن يضمن Key Variable في كلا الملفين المراد دمجهما وبنفس الاسم .

2. يجب ترتيب الملفين تصاعدياً Sorting Ascending بموجب Key Variable .
 بالنسبة لهذا المثال لا نحتاج إلى متغير مفتاح Key Variable .
 عند نقر زر OK في صندوق حوار Add Variables From يظهر الملف الناتج عن عملية الدمج كما يلي :

Merged File

name	math	chem	physc	music	arabic	english
Samir	100	90	95	87	80	98
Lubna	95	87	90	85	85	95

مثال 2 : (توضيح استعمال Key variable)

بافتراض أننا نريد دمج الملف Group1 الذي يحتوي درجات طالبين (المثال 1 أعلاه) مع الملف Sub3 الذي يحتوي درجات خمسة طلاب كما يلي :

	<u>Sub3</u>	
name	arabic	english
Yousif	90	85
Ammar	87	92
Sinan	85	91
Samir	80	98
Lubna	85	95

في حالة عدم استعمال Key Variable يظهر ناتج الدمج كما يلي :

name	math	chem	physc	music	arabic	english
Samir	100	90	95	87	90	85
Lubna	95	87	90	85	87	92
.	.	.	.	.	85	91
.	.	.	.	.	80	98
.	.	.	.	.	85	95

نلاحظ أن عملية الدمج غير صحيحة حيث لم تضاف درجات الطلبة في ملف Sub3 إلى درجات نفس الطلبة في ملف Group1 ، وعليه يتوجب استعمال Key Variable حسب الخطوات التالية :

ترتيب الملفين Group1 و sub3 تصاعدياً Sort Ascending حسب المتغير name بواسطة الأمر
 Sort Cases → Data مع خزن الملفين المرتبين باسم SGroup1 و SSub3 وكما يلي :

SGroup1

name	math	chem	physc	music
Lubna	95	87	90	85
Samir	100	90	95	87

SSub3

name	arabic	english
Ammar	87	92
Lubna	85	95
Samir	80	98
Sinan	85	91
Yousif	90	85

لاحظ أن الملفين قد رتبا حسب الحروف الألفبائية لأن المتغير name (متغير الترتيب) هو متغير رمزي .

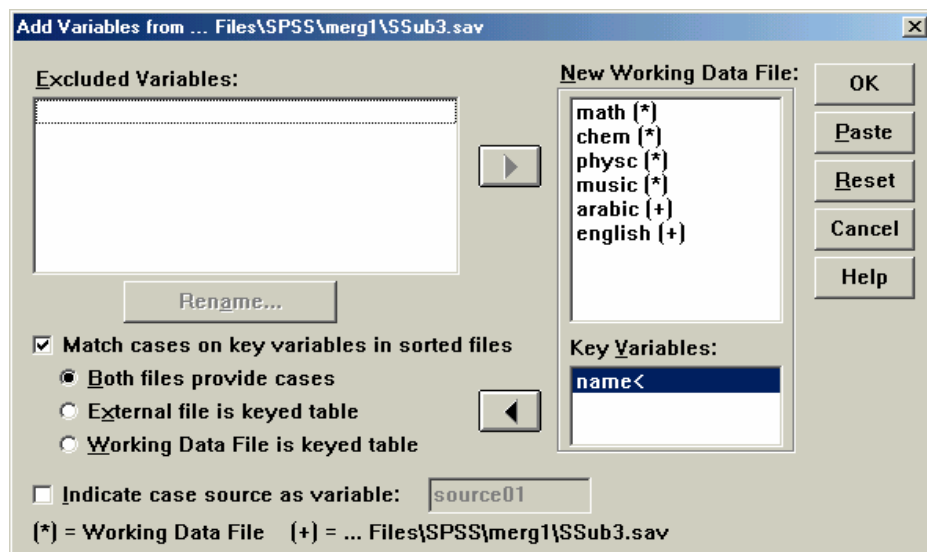
أفتح الملف المرتب SGroup1 .

من القوائم اختر Add Variables → Merge Files → Data فيظهر صندوق

حوار Add Variables From الذي نقوم بترتيبه على الشكل التالي :

- تأشير خانة Match Cases on Key Variable in Sorted files مع اختيار Both Files . Provide cases

- في قائمة Exclude Variables انقر المتغير name لتأشير ثم أنقله بواسطة  إلى خانة Key Variables فيظهر صندوق حوار Add Variables from على الشكل التالي :



عند نقر زر ok في الصندوق أعلاه يتم دمج الملفين وكما يلي :

Merged File

name	math	chem	physc	music	arabic	english
Ammar	.	.	.	.	87	92
Lubna	95	87	90	85	85	95
Samir	100	90	95	87	80	98
Sinan	.	.	.	.	85	91
Yousif	.	.	.	.	90	85

الملف الباحث Table Look up file أو Keyed Table : وهو ملف يحتوي عدة حالات Cases بحيث أن كل حالة يمكن أن تقابل (تطابق) عدة حالات في ملف آخر حيث يمكن الاستفادة من ملف كهذا في حالة التعدادات أو المسوحات الإحصائية لربط معلومات الوحدة السكنية مثلاً والتي هي وحدة العد بمعلومات أفراد الأسرة .

مثال 3 (الملف الباحث)

الملف household يحتوي معلومات عن أفراد الأسرة (الاسم name ، العمر Age ، التحصيل الدراسي Edu ، رقم المسكن housno) كما يلي :

household				
name	Age	Edu	housno	الأسرة الساكنة في المسكن
Ahmad	20	Sec	10	10
Zeki	35	Bsc	10	
Sabah	30	sec	10	
Zainab	15	Prim	10	
Ibrahim	17	Sec	12	12
Samir	40	Ma	12	
Selma	36	Bsc	12	

أما الملف house فيحتوي معلومات عن الأسرة كلها تتضمن حجم الأسرة Size ، موقع المسكن Location ، رقم المسكن housno كالتالي :

house		
Size	Location	housno
4	Baghdad	10
3	Baghdad	12

علماً أننا قد رتبنا الملفين أعلاه تصاعدياً بحسب المتغير housno الذي هو key Variable

→ بالأمر sort Cases Data وهذه نقطة البداية في دمج الملفين .

أن دمج الملفين أعلاه يعني عملية أقران معلومات المسكن (الملف الثاني house) بكل فرد من أفراد الأسرة في الملف الأول household ويمكن تنفيذ ذلك حسب الخطوات التالية :

⏪ أفتح ملف household الذي سيصبح الملف العامل Working File

→ من القوائم اختر Add Variables Merge Files Data ثم اختر الملف

house من صندوق حوار Add Variables : Read File الذي سيكون External File

⏪ عند نقر زر open في صندوق حوار Add Variables : Read File يظهر صندوق حوار

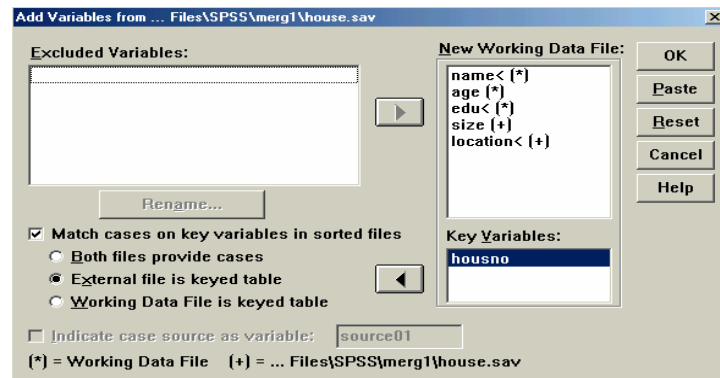
Add Variables From الذي يحتوي على المتغير housno ضمن قائمة Excluded Variables

حيث نقوم بأجراء مايلي :

• تأشير الخيار Match Cases on Key Variables in Sorted Files مع اختيار External File

is Keyed Table باعتبار أن الملف الخارجي هو الملف الباحث Table Look Up file .

- في قائمة Excluded Variables انقر المتغير housno من الملف الخارجي لتأشيرته ثم انقله بواسطة  الى خانة Key Variables باعتباره المتغير المفتاح حيث يظهر صندوق حوار Add Variables From بعد إجراء التحويلات المذكورة كما يلي :



عند نقر زر OK في صندوق الحوار السابق يتم دمج الملفين كما يلي :

Merged File					
name	Age	Edu	housno	Size	Location
Ahmad	20	Sec	10	4	Baghdad
Zeki	35	Bsc	10	4	Baghdad
Sabah	30	sec	10	4	Baghdad
Zainab	15	Prim	10	4	Baghdad
Ibrahim	17	Sec	12	3	Baghdad
Samir	40	Ma	12	3	Baghdad
Selma	36	Bsc	12	3	Baghdad

8. فصل (تجزئة الملفات) Split Files

يستعمل هذا الأمر لغرض تجزئة (فصل) ملف البيانات لأغراض التحليل الإحصائي .

مثال 1

الجدول التالي يمثل رواتب مجموعة من الأشخاص حسب الجنس :

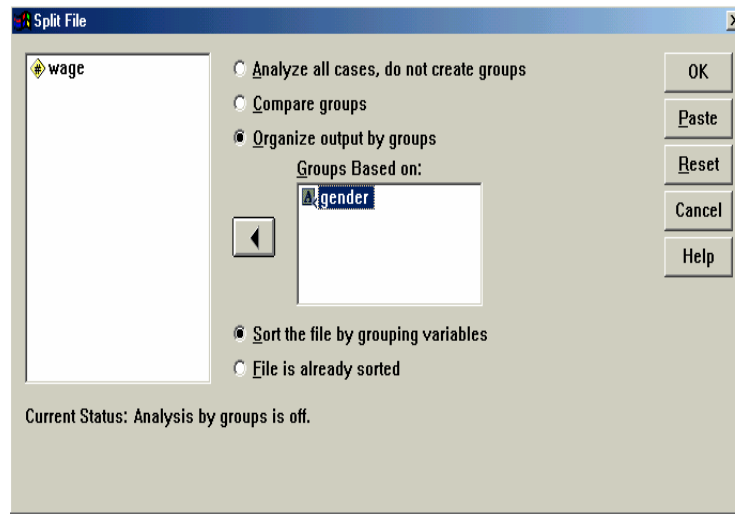
wage	gender
60	m
30	f
70	m
35	f
65	m
40	f

لتجزئة الملف إلى جزأين الأول يمثل رواتب الذكور m والثاني يمثل رواتب الإناث f نتبع الخطوات

التالية:

من شريط القوائم اختر  Split File → Data

يظهر صندوق حوار Split File الذي قمنا بترتيبه على الشكل التالي :



حيث أن :

Analyse All Cases,do not Creat Groups : عدم تجزئة الملف .

Compare Groups : يتم تجزئة الملف حسب فئات متغير معين (يحدد هذا المتغير أو مجموعة من المتغيرات في خانة Groups Based on) ويتم عرض نتائج عملية إحصائية معينة على شكل مقارنة النتائج بين المجاميع المختلفة لمتغير التجزئة .

Organize Output by Groups : هذا الخيار هو نفس الحالة السابقة Compare Groups ولكن يتم عرض النتائج بصورة مستقلة لكل مجموعة من مجاميع متغير التجزئة (في هذا المثال تم اعتماد هذا الخيار وقد استعملنا متغير التجزئة gender) . أن طريقة فصل الملف بموجب هذا الخيار تشابه طريقة الفصل لخيار Compare Groups ولكن الاختلاف يكمن في طريقة عرض نتائج العمليات الإحصائية مثلاً الأمر . Frequencies

Sort The File by Grouping Variable : يتم ترتيب الملف حسب مجاميع متغير (متغيرات) التجزئة .

File is Already Sorted: الملف لا يحتاج الى ترتيب ولكن يجب الانتباه الى أن البيانات مرتبة بصورة صحيحة حسب متغير التجزئة .فائدة هذا الخيار هي لاختصار الوقت اللازم لعملية الترتيب Sort وعلى الأغلب لا يتم اعتماد هذا الخيار وإنما نعتمد الخيار السابق (Sort The File by Grouping Variable) .

عند نقر زر OK في صندوق حوار Split File يتم تجزئة الملف إلى جزأين أحدهما للذكور m

والآخر للإناث f وقد تم الترتيب Sort حسب الحروف الأبجدية للمتغير gender وكما يلي :

wage	gender
30	f
35	f
40	f
60	m
70	m
65	m

علماً أننا نتوصل الى نفس الترتيب في حالة تأشير الخيار Compare groups ولكن الفرق بين الخيارين يظهر في مخرجات العمليات الإحصائية فلحساب المتوسط الحسابي للمتغير wage بالأمر Frequencies → Descriptive Statistics → Analyze (أنظر الفصل الرابع حول احتساب المتوسط الحسابي بالأمر Frequencies) تكون مخرجات الخيارين كالتالي :

1. مخرجات الخيار Organize Output by Groups

Frequencies GENDER = f

Statistics^a

WAGE

N	Valid	3
	Missing	0
Mean		35.00

a. GENDER = f

GENDER = m

Statistics^a

WAGE

N	Valid	3
	Missing	0
Mean		65.00

a. GENDER = m

2. مخرجات الخيار compare groups :

Frequencies

Statistics

WAGE

f	N	Valid	3
		Missing	0
	Mean		35.00
m	N	Valid	3
		Missing	0
	Mean		65.00

مثال 2 : الجدول التالي يبين الإنتاج لمحصول معين في سنتي 2000 و 2001 وحسب المناطق (شمالية،جنوبية)

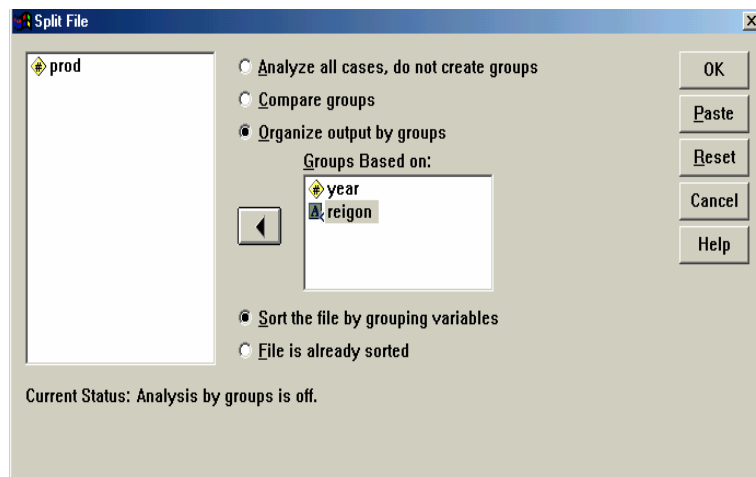
prod	year	region
800	2000	North
600	2000	South
1400	2001	North
900	2000	North

1090	2001	South
950	2000	North
1350	2001	North
1180	2001	South
700	2000	South
975	2000	North
1290	2001	North
1000	2001	South
750	2000	South
1310	2001	North
1150	2001	South

المطلوب تجزئة الملف (متغير الإنتاج prod) حسب السنة والمنطقة .

لغرض تجزئة الملف بحيث أن حالات المتغير prod تفصل حسب المناطق ضمن كل سنة نتبع الخطوات التالية :

من القوائم اختر Split File → Data فيظهر صندوق حوار Split File الذي نقوم بترتيبه على الشكل التالي :



عند نقر زر OK في الصندوق أعلاه يتم تجزئة الملف الى أربعة مجاميع كما في الشكل التالي :

prod	year	region	
800	2000	North	1st Group
900	2000	North	
950	2000	North	
975	2000	North	2nd Group
600	2000	South	
700	2000	South	
750	2000	South	
1400	2001	North	3rd Group
1350	2001	North	
1290	2001	North	
1310	2001	North	
1090	2001	South	4th Group

1180	2001	South
1000	2001	South
1150	2001	South

عند إجراء التحليل الإحصائي بالأمر frequencies لاستخراج متوسط المتغير prod فأن هذا التحليل
ينجز على المجاميع الأربعة Organize Output by Groups كما يلي :

Frequencies

YEAR = 2000, REIGON = North

Statistics^a

PROD

N	Valid	4
	Missing	0
Mean		906.25

a. YEAR = 2000, REIGON = North

YEAR = 2000, REIGON = South

Statistics^a

PROD

N	Valid	3
	Missing	0
Mean		683.33

a. YEAR = 2000, REIGON = South

YEAR = 2001, REIGON = North

Statistics^a

PROD

N	Valid	4
	Missing	0
Mean		1337.50

a. YEAR = 2001, REIGON = North

YEAR = 2001, REIGON = South

Statistics^a

PROD

N	Valid	4
	Missing	0
Mean		1105.00

a. YEAR = 2001, REIGON = South

9. تجميع البيانات Aggregate Data

يستعمل هذا الأمر لتلخيص المعلومات المتعلقة بمجموعة من الحالات cases في حالة تجميعية واحدة وتكوين ملف تجميعي جديد. فعلى سبيل المثال إذا توفرت قائمة تحتوي معدلات الطلاب لمجموعة من المدارس فقد نرغب في التعامل مع المدرسة (باعتبارها وحدة العد الإحصائية) بدلاً من الطالب فيمكن أن نعرض الوسط الحسابي لمعدلات الطلاب في كل مدرسة أو الانحراف المعياري مثلاً في ملف تجميعي جديد بدلاً من عرض معدلات جميع الطلاب لكافة المدارس .

مثال

الملف salary يحتوي أسماء مجموعة من الموظفين ودرجتهم الوظيفية degree والأجر الشهري salary وكما يظهر في شاشة Data Editor يلي :

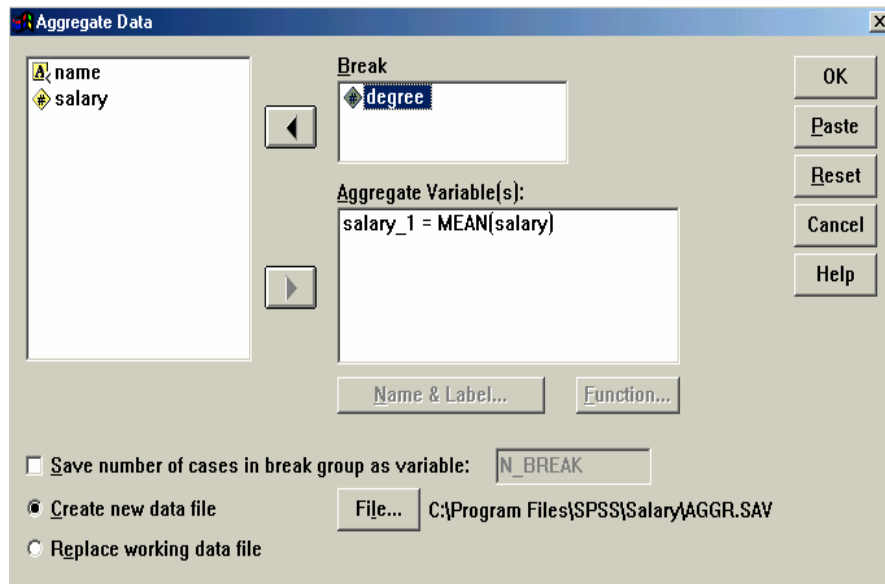
الملف salary

name	degree	salary
Ahmad	3	40
Samer	3	35
Loay	3	50
Mahmood	1	80
Ayad	1	70
Yassin	2	66
Satar	1	85
Razak	1	77
Kamal	2	59
Abas	3	45
Mahdi	1	90
Salim	2	62
Sabah	2	57
Falah	2	55
Imad	1	82

المطلوب تجميع متغير الراتب (استخراج المتوسط الحسابي) حسب الدرجة الوظيفية degree.

لتنفيذ ذلك نتبع الخطوات التالية :

من شريط القوائم اختر Aggregate → Data فيظهر صندوق حوار Aggregate Data الذي نرتبه كما يلي :



حيث أن :

Break: وهو متغير (متغيرات) تجزئة breakdown Variable(s) يستعمل لتعريف مجاميع (من الحالات) وفي هذا المثال استعملنا متغير الدرجة الوظيفية degree لتكوين المجاميع .

Aggregate Variable(s) : وهو المتغير (المتغيرات) الذي نرغب في تجميع الحالات التي يتكون منها حسب متغير التجزئة .

في هذا المثال فإن المتغير التجميعي هو salary وعند نقله الى قائمة Aggregate Variable(s) فإنه يظهر بأسم افتراضي هو salary_1 ويمثل الوسط الحسابي للأجر الشهري لكل درجة وظيفية علماً انه :
• يمكن تغيير الاسم الافتراضي للمتغير وعنوانه باستعمال الزر Name & Label بعد نقر اسم المتغير في قائمة Aggregate Variables .

• يمكن تغيير الدالة الافتراضية (الوسط الحسابي Mean) بنقر الزر Function بعد نقر اسم المتغير في قائمة Aggregate Variables حيث يمكن اختيار دوال أخرى مثل Standard Deviation ، Sum ، No. of Cases ، of Cases... الخ .

Save Number of cases in break Group as variable : عند تأشير المربع المجاور له يتم تكوين متغير جديد بأسم افتراضي هو N_Break يبين عدد الحالات لكل مجموعة فمثلاً للدرجة الوظيفية الأولى يأخذ المتغير القيمة 6 وللدرجة الوظيفية الثانية 5 وللدرجة الثالثة 4 .

Creat new Data File : هذا الخيار يتيح تكوين ملف جديد يحتوي المعلومات التجميعية بأسم افتراضي هو Aggr في نفس الدليل الذي يقع فيه الملف الأصلي salary وقد اخترنا هذا الخيار. علماً انه يمكن تغيير أسم الملف التجميعي وموقعه بنقر الزر File .

Replace Working Data File : لاحتلال الملف التجميعي محل الملف الحالي salary علماً أن ذلك لايلغي خزن الملف الأصلي وأن الملف التجميعي الناتج لا يخزن مالم يتم خزنه بأسم معين .

◀ عند نقر زر Ok يتم تكوين وخزن الملف التجميعي AGGR ، لقراءة محتوياته يتوجب فتحه بالأمر

Data → Open → File ويظهر الملف كما يلي :

الملف AGGR

degree salary

1 80.67

2 59.80

3 42.50

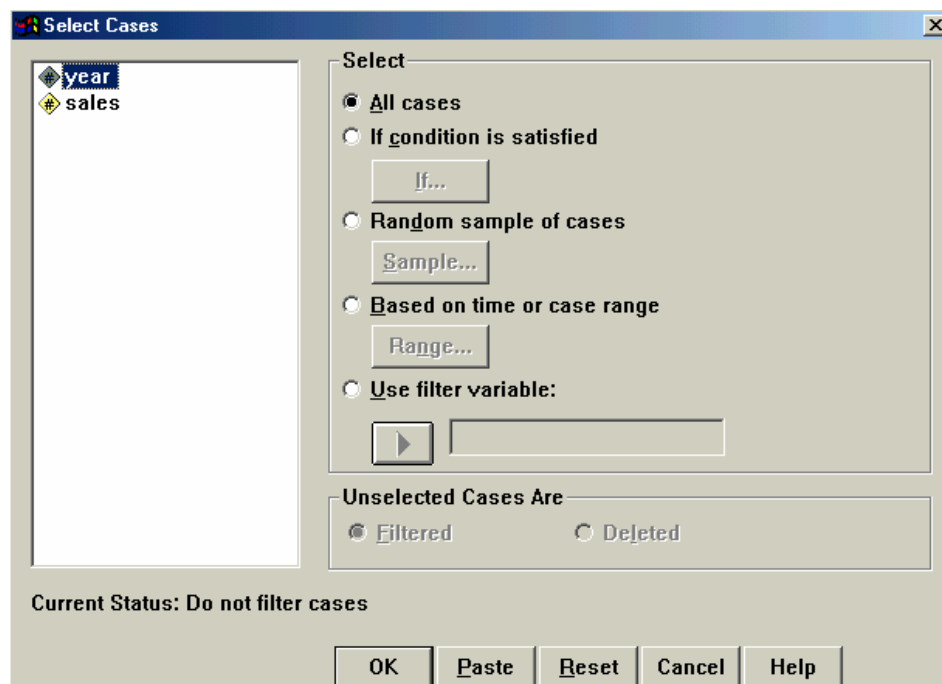
يمكن إدخال المتغير salary إلى قائمة Aggregate Variable(s) عدة مرات في نفس الوقت حيث

يأخذ تسمية برقم تسلسلي salary_1 ، salary_2 ... مثلاً المتغير salary_1 يمثل تجميع الحالات حسب المتوسط والمتغير salary_2 يمثل التجميع حسب الانحراف المعياري ... وهكذا .

10. أختيار الحالات Select Cases: يفيد هذا الأمر في اختيار جزء من الحالات لتضمينها في التحليل الإحصائي. وهي نفس الوظيفة التي تؤديها الأيقونة  في شريط الأدوات . لنفترض لدينا السلسلة الزمنية التالية التي تمثل مبيعات شركة ما خلال الفترة 1990 – 2002 .

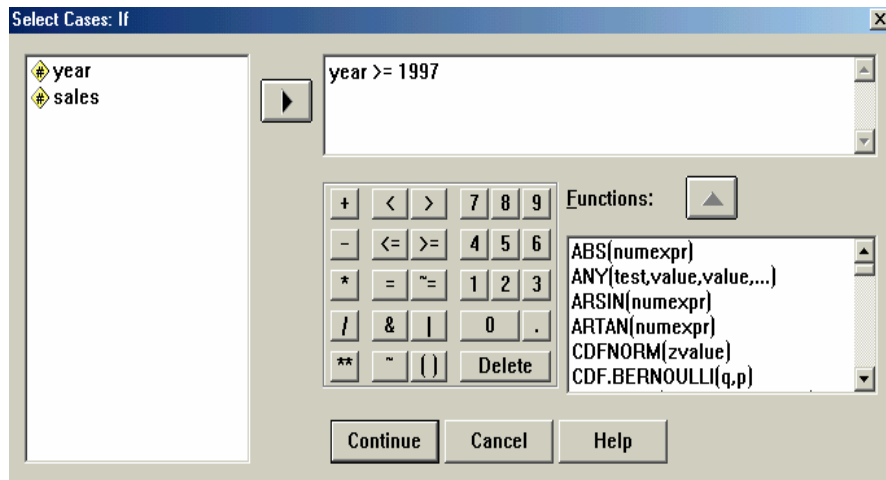
	year	sales
1	1990	110
2	1991	115
3	1992	118
4	1993	125
5	1994	130
6	1995	127
7	1996	150
8	1997	170
9	1998	177
10	1999	166
11	2000	184
12	2001	210
13	2002	220

نرغب في أختيار الحالات التي تقابل السنوات 1997-2002 ، لتنفيذ ذلك نتبع الخطوات التالية :-
 < من شريط القوائم أختار Select Cases → Data فيظهر صندوق حوار Select Cases كما يلي :-



حيث أن الخيار All Cases يعني أختيار كافة حالات الملف

➤ لأختيار الحالات التي تقابل السنوات 1997-2002 أنقر الخيار If Condition is Satisfied ثم أنقر زر IF فيظهر صندوق الحوار التالي الذي نقوم بترتيبه كما يلي :



يمكنك النقر على الأزرار في الصندوق أعلاه بزر الماوس الأيسر كما تستعمل الأزرار في الحاسبة اليدوية .

➤ عند نقر زر OK يتم اختيار الحالات المطلوبة (1997-2002) أما بقية الحالات فتكون مشطوبة Dashed ويتم إضافة متغير جديد الى الملف بأسم filter_\$ الذي يأخذ القيمة 1 للحالات المختارة والقيمة 0 للحالات المستبعدة كما يلي :-

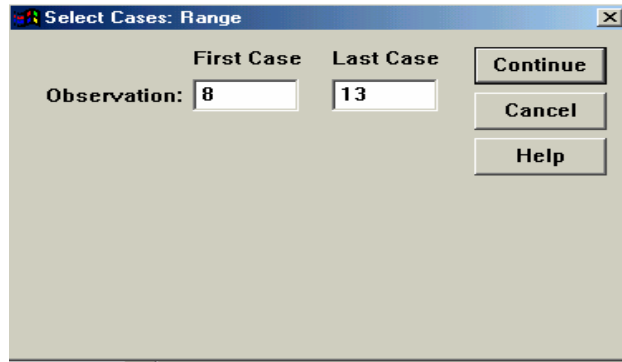
	year	sales	filter_\$
1	1990	110	0
2	1991	115	0
3	1992	118	0
4	1993	125	0
5	1994	130	0
6	1995	127	0
7	1996	150	0
8	1997	170	1
9	1998	177	1
10	1999	166	1
11	2000	184	1
12	2001	210	1
13	2002	220	1

لاحظ أن الحالات المستبعدة لم تحذف من الملف ولكنها لاتدخل التحليل الإحصائي حيث أننا قمنا بتأشير الخيار Filtered في خانة Unselected Cases Are في صندوق حوار Select Cases إذا أردنا

	year	sales	filter_\$
1	1997	170	1
2	1998	177	1
3	1999	166	1
4	2000	184	1
5	2001	210	1
6	2002	220	1

حذف الحالات المستبعدة نهائياً نقوم بتأشير الخيار Deleted حيث نحصل على النتيجة التالية:-

أن اختيار الحالات التي تقابل السنوات (1997-2002) يمكن ان يتم أيضاً بتأشير الخيار Based on Time or Case Range ثم نقر زر Range فيظهر صندوق الحوار التالي حيث نحدد الحالات المختارة (8-13) .

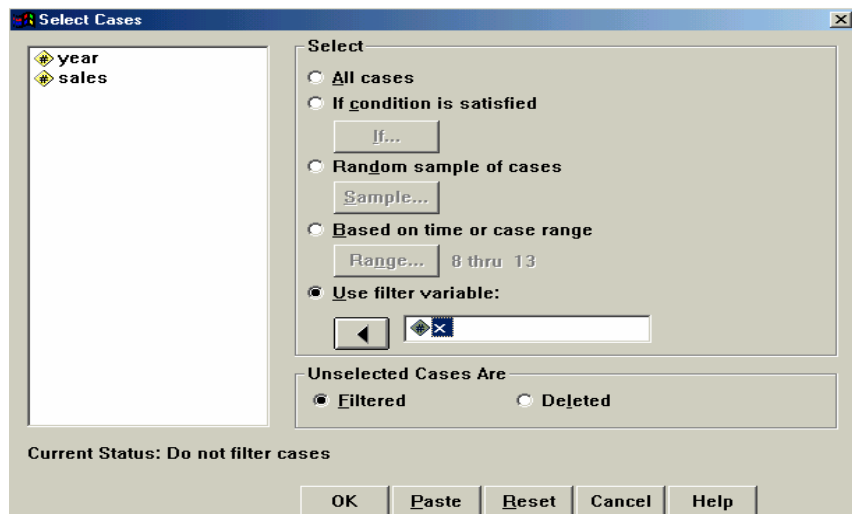


ملاحظات :

1. يمكن اختيار الحالات من خلال إضافة filter variable الى الملف حيث نعطي القيمة 1 للحالات التي نرغب في اختيارها والقيمة 0 للحالات التي لا نرغب في اختيارها فبالنسبة للمثال السابق تكون قيم هذا المتغير (X مثلاً) كما يلي :-

	year	sales	x
1	1990	110	0
2	1991	115	0
3	1992	118	0
4	1993	125	0
5	1994	130	0
6	1995	127	0
7	1996	150	0
8	1997	170	1
9	1998	177	1
10	1999	166	1
11	2000	184	1
12	2001	210	1
13	2002	220	1

ويتم ترتيب صندوق حوار Select Cases كما يلي :-



حيث نحصل على النتيجة ذاتها التي توصلنا اليها بالخيار if condition is satisfied

2. الخيار Random Number of Cases يتيح اختيار عينة عشوائية بسيطة من الحالات ويتم ذلك

أما بتحديد النسبة المئوية للعينة (مثلاً 5% من الحالات) أو تحديد العدد الصريح للحالات .

3. لإلغاء اختيار الحالات نقوم بتأشير All Cases في صندوق حوار Select Cases .

11. ترجيح الحالات Weight Cases : يتيح هذا الأمر إمكانية إعطاء أوزان لحالات Cases ملف معين

نظراً لاختلافها من ناحية الأهمية النسبية وهي نفس الوظيفة التي تؤديها الأيقونة  في شريط

الأدوات القياسي.

مثال : القيم التالية تمثل نتائج امتحان أحد الطلبة في مادة الإحصاء علماً أن لكل امتحان وزناً أو أهمية

نسبية معينة وان البيانات قد أدخلت في شاشة Data Editor لبرنامج SPSS.

الامتحان	الدرجة degree	الوزن weight
الأول	70	10
الثاني	60	30
الثالث	75	10
الرابع	55	50

يطلب ما يلي :

1. حساب الوسط الحسابي للامتحانات الأربعة .
2. حساب الوسط الحسابي المرجح للامتحانات الأربعة.
1. لحساب الوسط الحسابي من شريط القوائم نختار

Analyze → Descriptive Statistics → Frequencies

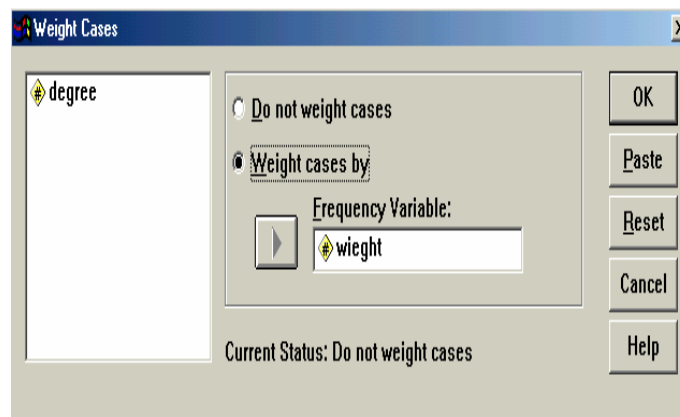
(راجع الفصل الرابع حول الأمر Frequencies) مع تأشير الخيار Mean حيث نحصل على النتيجة التالية :

Frequencies

Statistics

DEGREE		
N	Valid	4
	Missing	0
Mean		65.00

2. لحساب الوسط الحسابي المرجح بالمتغير weight للامتحانات الأربع نتبع الخطوات التالية :
- من شريط القوائم نختار Weight Cases → Data فيظهر صندوق حوار Weight Cases الذي نقوم بترتيبه بالشكل التالي :



عند نقر زر OK يتم وزن حالات الملف بالمتغير weight (علماً أنك لن تلاحظ أي تغيير في شاشة Data Editor) .

عند احتساب الوسط الحسابي بالأمر Frequencies سوف نحصل على النتيجة التالية :

Frequencies

Statistics

DEGREE		
N	Valid	100
	Missing	0
Mean		60.00

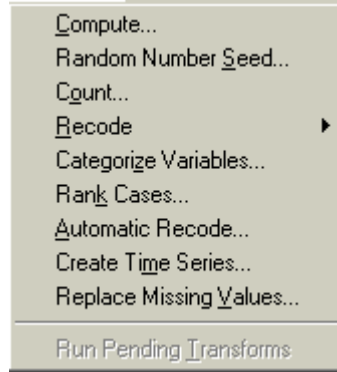
ملاحظة : لإلغاء ترجيح حالات الملف نقوم بتأشير الخيار Do not weight cases في صندوق حوار Weight Cases.

الفصل الثالث

تحويل البيانات

Data Transformation

يمكن انجاز تحويل على البيانات الأصلية Row Data تتراوح بين تجميع البيانات في فئات معينة إلى تكوين متغيرات جديدة بالاعتماد على معادلات و صيغ شرطية. تضم قائمة Transform تظهر الأوامر التالية:



1. الأمر Compute : يتيح هذا الأمر إمكانية حساب متغيرات جديدة باستخدام أكثر من 70 دالة تتضمن (دوال حسابية ، إحصائية ، توزيعات احتمالية) .

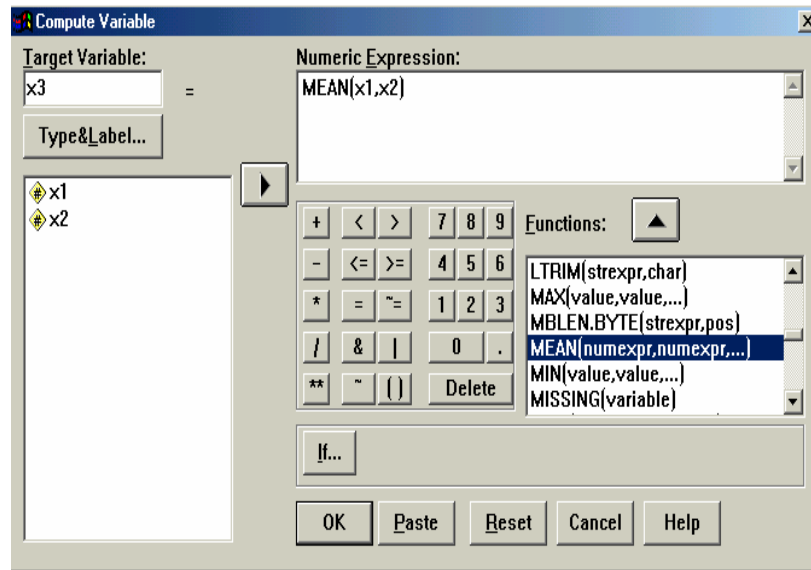
مثال :

الجدول التالي يمثل المتغيرين x_1 و x_2 اللذين تم إدخالهما إلى شاشة Data Editor وكما يلي :

x_1	x_2
60	90
87	88
70	43
90	80
57	55
73	47
95	90
66	50
40	55
55	80
85	75
88	86
35	70

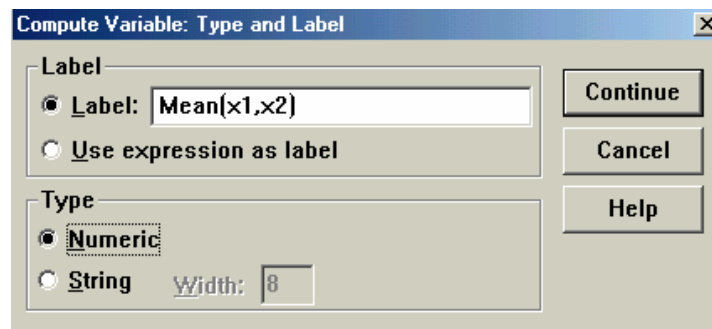
يطلب حساب الوسط الحسابي Mean (دالة إحصائية) لحالات المتغيرين x_1 و x_2 في حالة كون قيم $x_1, x_2 \geq 50$. لتنفيذ ذلك نتبع الخطوات التالية :

من شريط القوائم نختار $\rightarrow$ compute Transform فيظهر صندوق حوار $\leftarrow$ compute variable الذي نقوم بترتيبه على الشكل التالي :



وقد أجرينا العمليات التالية :

- أ. تحديد اسم المتغير الهدف Target Variable الذي هو عبارة عن الوسط الحسابي لـ X_1 و X_2 .
(أنقر المستطيل أسفل خانة Target Variable ثم أكتب اسم المتغير الهدف X_3).
- ب. اختيار الدالة و هي Mean من قائمة Functions ثم ضغط زر  .
- ت. اختيار المتغيرات الداخلة وهي X_1 و X_2 كل على حدة ثم نقر زر  لإدخالها في الدالة Mean وفي المواقع المحددة كما هو واضح في صندوق الحوار أعلاه .
- ث. لتحديد عنوان و نوع المتغير الهدف أنقر الزر Type & Label فيظهر صندوق حوار type & Label وقد رتبناه كما يلي:



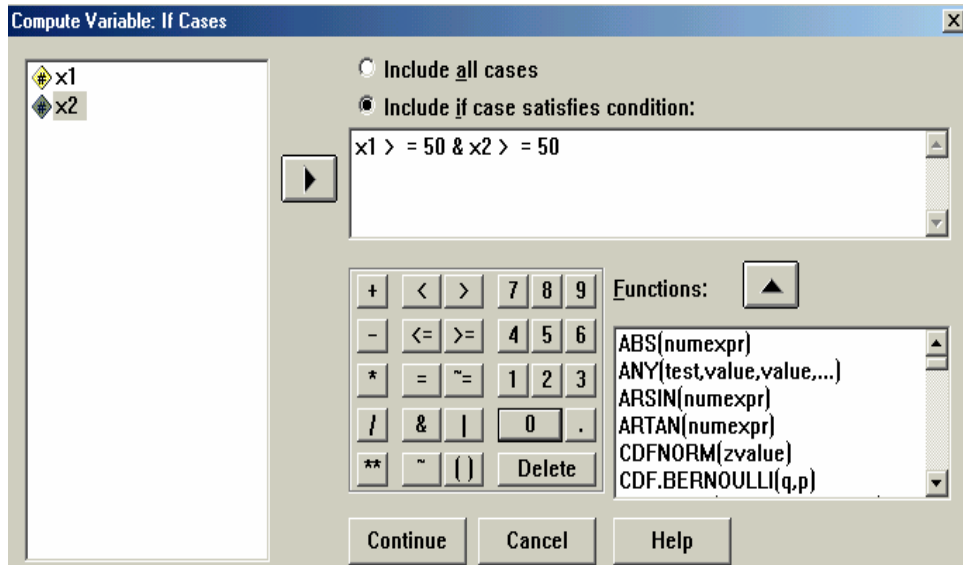
حيث أن :

label : من خلاله يمكن إعطاء عنوان للمتغير X_3 بنقر الدائرة المجاورة لكلمة label (يمكن أن يكون طول العنوان لغاية 120 رمز) وقد اخترنا العنوان $\text{Mean}(x_1, x_2)$ أو يكون التعبير الحسابي Numeric expression هو عنوان المتغير X_3 عند نقر الدائرة المجاورة لـ use expression . as label

type : أن النوع الافتراضي للمتغير الناتج هو عددي numeric أما في حالة كون المتغير الناتج رمزياً String فيتوجب تحديد نوع وطول المتغير width .

بعد الانتهاء انقر الزر continue. لاختيار جزء من الحالات cases التي تجرى عليها عمليات الاحتساب (نريد احتساب الحالات التي يكون فيها $x_1, x_2 \geq 50$) انقر if في صندوق حوار compute variable يظهر صندوق حوار if cases

- لتضمنين كافة الحالات انقر include all cases
 - لتضمنين جزء من الحالات انقر include if case satisfies condition
- بما أننا نريد احتساب جزء من الحالات سنقوم بنقر الخيار الأخير ثم إدخال الشرط ($x_1 \geq 50$, $x_2 \geq 50$) بالاستعانة بالأزرار الموجودة في أسفل المستطيل الخاص بالشرط و بالدوال المتوفرة حيث يظهر صندوق حوار if Cases بعد ترتيبه بالشكل التالي :



عند نقر زر continue في صندوق حوار If Cases ثم زر OK في صندوق حوار Compute Variable يتم الحصول على نتائج الاحتماب حيث يتم إضافة متغير آخر هو x3 (الوسط الحسابي للمتغيرين x1 و x2) إلى Data Editor كما يلي :

x1	x2	x3
60	90	75
87	88	88
70	43	.
90	80	85
57	55	56
73	47	.
95	90	93
66	50	58
40	55	.
55	80	68
85	75	80
88	86	87
35	70	.

ملاحظة: إن دوال التوزيعات الاحتمالية المتوفرة في برنامج SPSS تتيح إمكانية الحصول على الاحتمالات المقابلة لقيم متغير أي من التوزيعات الاحتمالية المعروفة و بذلك فهي تغني عن استخدام جداول التوزيعات (مثل التوزيع الطبيعي ، توزيع t ... الخ) فمثلا يمكن الحصول على الدالة التجميعية CDF المقابلة لأي قيمة من قيم المتغير العشوائي لتوزيع معين عن طريق تحديد المعالم parameters لذلك التوزيع كما في الشكل التالي :

CDFZ=CDFNORM[Z]

الدالة التجميعية للتوزيع
الطبيعي القياسي

CDFbinom=CDF.BINOM[x,10,0.5]

binomial r.v. $\begin{matrix} n & p & | & | \end{matrix}$

CDFt=CDF.T[t,15]

t-dist. r.v. $\begin{matrix} & df \end{matrix}$

	z	cdfz	var00003	x	cdfbinom	var00002	t	cdft
1	.00	.5000	.	0	.0010	.	.0	.5000
2	.05	.5199	.	1	.0107	.	.1	.5392
3	.10	.5398	.	2	.0547	.	.2	.5779
4	.15	.5598	.	3	.1719	.	.3	.6159
5	.20	.5793	.	4	.3770	.	.4	.6526
6	.25	.5987	.	5	.6230	.	.5	.6878
7	.30	.6179	.	6	.8281	.	1.0	.8334
8	.35	.6368	.	7	.9453	.	1.5	.9228
9	.40	.6554	.	8	.9893	.	2.0	.9680
10	.45	.6736	.	9	.9990	.	2.5	.9877
11	.50	.6915	.	10	1.0000	.	3.0	.9955
12	.55	.7088	.	.	.	.	3.5	.9984
13	.60	.7257	.	.	.	.	4.0	.9994
14	.65	.7422	.	.	.	.	4.5	.9998
15	.70	.7580	.	.	.	.	.	.
16	.75	.7734	.	.	.	.	.	.
17	.80	.7881	.	.	.	.	.	.
18	1.20	.8849	.	.	.	.	.	.
19	1.60	.9452	.	.	.	.	.	.
20	2.00	.9772	.	.	.	.	.	.

df : تمثل درجات الحرية Degrees of Freedom . :

2. الأمر Random Number Seed : يستعمل لضبط توليد الأعداد العشوائية .

3. الأمر Count :

إن لهذا الأمر أهمية خاصة في الاستبيانات الإحصائية لحساب عدد مرات تكرار نفس القيمة لمجموعة من المتغيرات و لكل حالة . فعلى سبيل المثال عند تنفيذ استطلاع حول مجموعة من المجالات لمعرفة أي من المجالات يقرأ المستفيد حيث يكون جواب السؤال نعم / لا حيث يمكن تكوين متغير جديد الذي يحسب عدد مرات تكرار الإجابة yes لمعرفة عدد المجالات المقروءة لكل مستفيد.

مثال: نفترض لدينا المتغيرين العدديين y1 و y2 (تم إدخال البيانات في Data Editor) كما في

الجدول التالي :

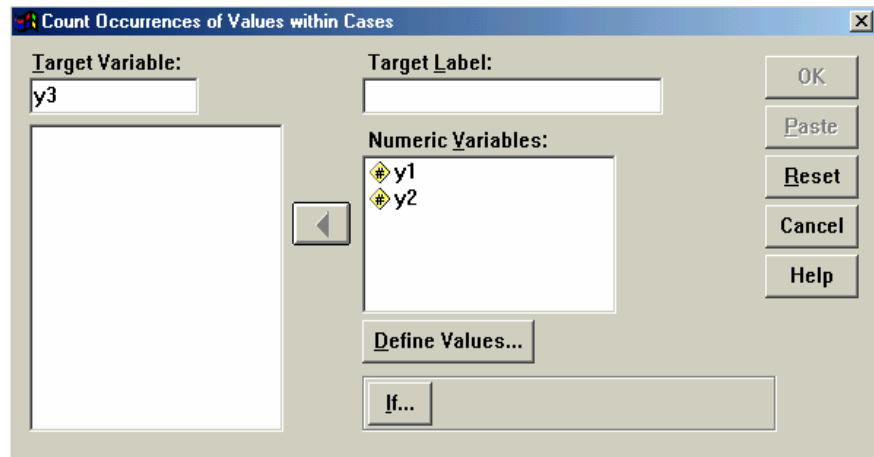
y1	y2
1	8
15	3
14	7
9	1
1	1
4	10
6	7
11	6
15	9
9	1
6	6
20	9
16	3

25	22
8	5
7	7

و نرغب في تكوين متغير ثالث y3 الذي يحسب عدد مرات تكرار الأرقام 1 و 6 و 20 فأكثر في المتغيرين المذكورين لتنفيذ ذلك نطبق الخطوات التالية:-

Transform → Count من شريط القوائم نختار

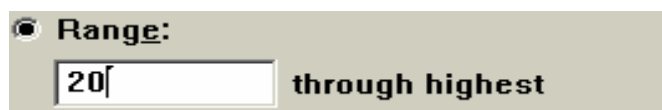
يظهر صندوق حوار Count Occurrence of values within cases الذي نقوم بترتيبه كما يلي :



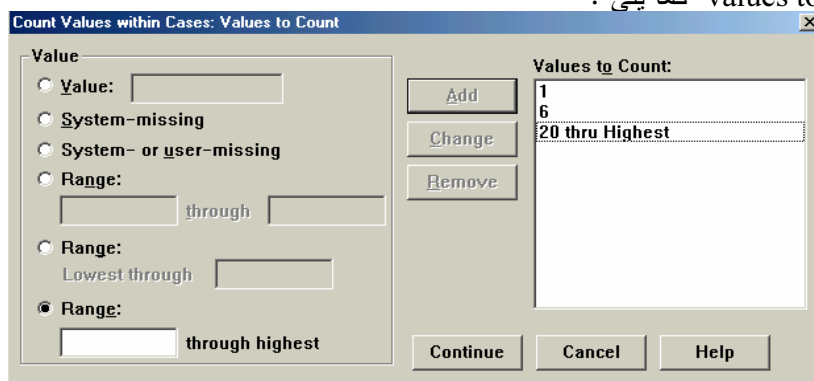
حيث يتطلب مايلي :

- إدخال اسم المتغير الهدف Target variable وهو y3 ويمكن إدخال عنوان المتغير الهدف (اختياري) .
- اختيار المتغيرين y1 و y2 من قائمة المتغيرات (يمكن اختيار اثنين أو أكثر من المتغيرات على ان تكون من نفس النوع (عددية أو رمزية) ثم نقلها إلى مستطيل Numeric variable .
- لغرض تعريف القيم التي نريد حساب عدد مرات تكرارها ننقر الزر Define values يظهر صندوق الحوار الخاص بالقيم Values to count حيث أننا نرغب بأدخال القيم 1،6،20 فأكثر وينم ذلك حسب الخطوات التالية:

- لإدخال القيمة 1 ننقر المستطيل المجاور لكلمة value وندخل القيمة 1 ثم ننقر زر Add لإضافة القيمة إلى مستطيل value to Count في اليمين.
- لإدخال القيمة 6 ننقر المربع المجاور لكلمة value وندخل القيمة 6 ننقر زر Add لإضافة القيمة إلى قائمة Values to count في اليمين.
- لإدخال القيمة 20 فأكثر ننقر الدائرة بجوار Range (أسفل الصندوق) ثم ننقر المستطيل أسفل كلمة Range ونكتب 20 أي أنها ستكون على الشكل التالي :



ثم نقر زر Add لإضافة المدى 20 فأكثر إلى قائمة Values to Count في اليمين .حيث يكون شكل صندوق حوار values to Count كما يلي :



علماً أن الزر Change يستعمل لتغيير قيمة والزر Remove لحذف قيمة بعد تأشيرها بزر الماوس الأيسر في قائمة Values to count .

عند نقر زر Continue للرجوع إلى صندوق حوار count occurrence of values within

y1	y2	y3
1	8	1
15	3	0
14	7	0
9	1	1
1	1	2
4	10	0
6	7	1
11	6	1
15	9	0
9	1	1
6	6	2
20	9	1
16	3	0
25	22	2
8	5	0
7	7	0

cases ثم نقر زر OK في هذا الأخير يضاف المتغير

y3 إلى Data Editor كما يلي :

4. الأمر Recode : يستفاد من هذا الأمر في إعطاء code (رمز) لكل قيمة من قيم متغير ما حيث

يستفاد منها في عمل الفئات ويتضمن الأمر نوعين من الترميز :

أ. الأمر Recode into same variables : يستفاد من هذا الأمر في تكوين متغير جديد قيمه عبارة

عن رموز لقيم متغير قديم ويأخذ هذا المتغير نفس أسم المتغير القديم .

مثال : لنفترض لدينا المتغير salary الذي يأخذ القيم التالية

salary : 20 16 95 88 65 53 35 46 90 22 30 28 51 60 85

وقد أدخلت قيم المتغير في شاشة Data Editor . المطلوب هو إعطاء رمز Code لكل قيمة من

قيم المتغير حسب الترتيب التالي:

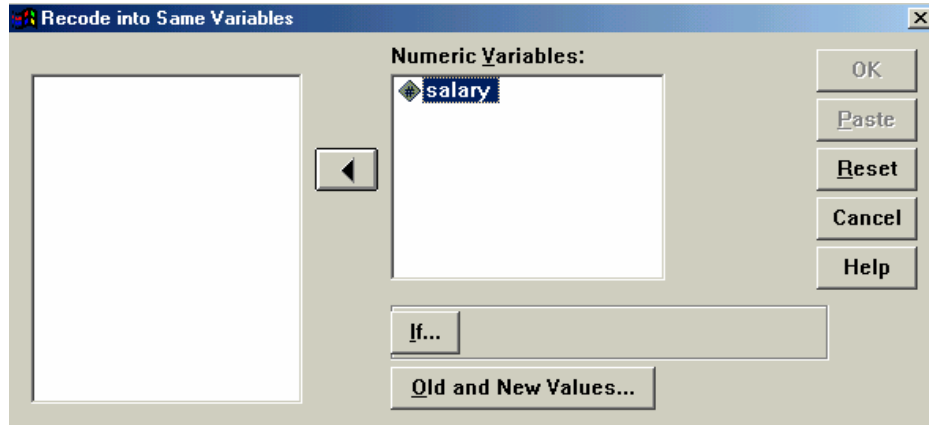
الكود	الفئة
1	24 فأقل
2	49-25

3 74 - 50

4 75 فأكثر

لتنفيذ ذلك نتبع الخطوات التالية:

من القوائم نختار Transform → Recode → into same variables
فيظهر صندوق حوار Recode into same variables الذي نرتبه كما يلي :



لقد اخترنا المتغير salary الذي نريد ترميز قيمة (علماً انه يمكن اختيار أكثر من متغير واحد على ان تكون كلها من نفس النوع (عددية او رمزية).

انقر زر Old and New Values لتحديد طريقة ترميز القيم فيظهر صندوق حوار

Old and New Values حيث يلاحظ وجود نوعين من القيم :

old value : وهي القيم التي نريد إعطاء رمز لها قد تكون (قيمة مفردة أو قيم مفقودة أو نطاق محدد من القيم).

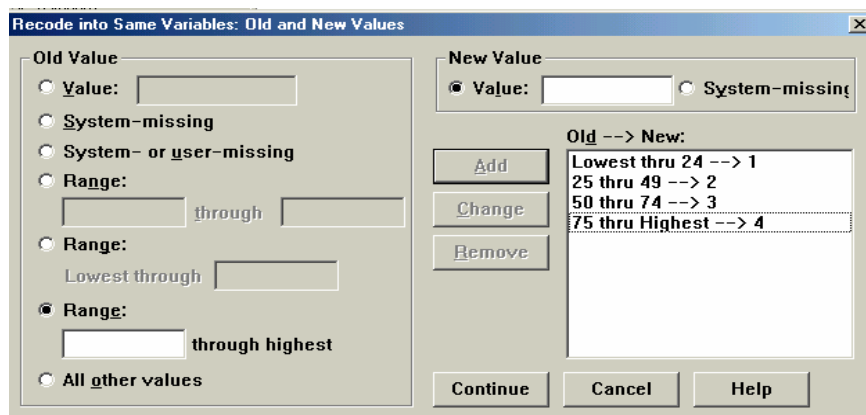
new value : وهي الرموز التي تعطى للقيم القديمة ومن الممكن أن تكون قيماً مفقودة.

- لإدخال الفئة الأولى 24 فأقل في قائمة old value انقر الدائرة المجاورة لـ range ويتم إدخال البيانات كما يلي :



- في قائمة New Value انقر المستطيل المجاور لـ value لإدخال الرمز المقابل للفئة (24 فأقل) وهو 1.

- انقر زر Add فيتم إضافة الفئة مع الرمز المقابل إلى قائمة new → old .
و بنفس الطريقة يتم إدخال بقية الفئات مع أكودها حيث يظهر صندوق حوار old and new values كما يلي :



علماً أنه يمكن بعد تأشير فئة معينة إنجاز الفعاليات التالية :

الزر Change لتغيير فئة بعد تأشيرها وأدخال القيمة القديمة في قائمة Old Value وأدخال القيمة الجديدة في قائمة New Value .

الزر Remove لحذف فئة بعد تأشيرها في قائمة Old → New.

بعد الانتهاء من إدخال الفئات انقر زر Continue لإنجاز الترميز حيث يصبح المتغير salary في Data Editor على الشكل التالي:-

salary : 1 1 4 4 3 3 2 2 4 1 2 2 3 3 4

أي أن قيم المتغير الأصلي Salary قبل الترميز سوف تفقد من Data Editor ويحل محلها الرموز Codes ، علماً أن الملف الذي يحتوي المتغير الأصلي لن يفقد ما لم يتم حفظ الملف الناتج بنفس اسم الملف القديم .

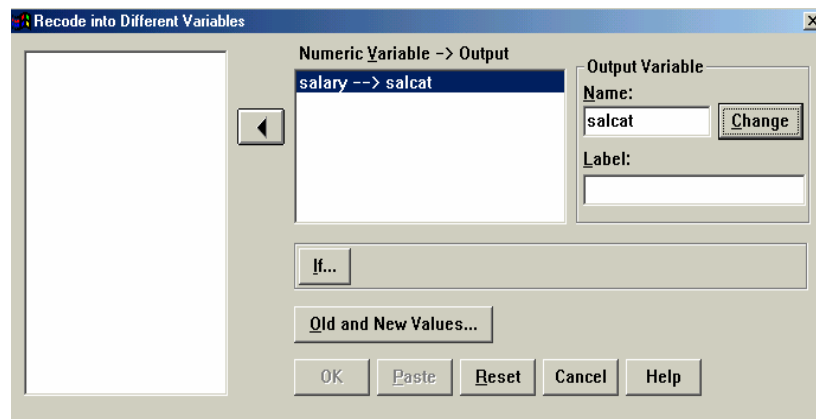
ب. الأمر Recede into different variable : يمكن هذا الأمر من إعطاء رموز لمتغير ما وتخزينها في متغير آخر مع الحفاظ على المتغير القديم وإن الأسلوب هو نفسه المتبع مع الترميز إلى نفس المتغير.

مثال : نرغب في ترميز المتغير Salary الوارد في المثال السابق حسب الفئات المذكورة وتخزين الرموز في متغير مختلف .

لترميز المتغير salary إلى متغير مختلف نتبع الخطوات التالية

من القوائم نختار

Transform → Recede → into different variables



فيظهر صندوق حوار Recode in to different variables الذي نرتبه كما يلي :

وقد قمنا بنقل المتغير Salary الى قائمة Output Numeric Variable بعدها أدخلنا أسم المتغير الناتج عن طريق نقر المستطيل في خانة Name ثم كتابة الاسم الجديد salcat مع نقر الزر Change وب نفس الطريقة يمكن تغيير الأسم عند الرغبة بذلك .

- عند نقر الزر Old and New Values يتم إدخال القيم القديمة و الجديدة بنفس طريقة الترميز إلى نفس المتغيرات ويظهر صندوق حوار Recode into different variables بعد الإدخال كما يلي:

ملاحظة : يمكن الاحتفاظ بالقيم القديمة نفسها للمتغير (التي لم تعطى كود معين) عن طريق اختيار All other values من خانة old value واختيار copy old values من خانة new value.

عند نقر Continue في صندوق حوار Recode into

different variables يضاف متغير جديد باسم salcat في Data Editor

كما في الشكل المجاور :

5. الأمر categorize variables (تبويب المتغيرات)

ان هذا الإجراء يقوم بتحويل متغير معين الى عدد منفصل من الفئات ويتطلب صندوق الحوار تحديد اسم المتغير (المتغيرات) و عدد الأصناف categories فمثلا إذا تم تحديد عدد الفئات مساويا إلى 4 للمتغير salary فسيتم تخصيص الرقم 1 لقيم المتغير في الربع الأول (ترتيب القيم اقل من 25%).

تخصيص الرقم 2 لقيم المتغير (التي ترتيبها من 25% الى 50%

).

تخصيص الرقم 3 لقيم المتغير (التي ترتيبها من 50% الى 75%).

تخصيص الرقم 4 القيم التي ترتيبها (75% فاكثر).

لتنفيذ هذه العملية لنفس المثال السابق (المتغير salary) نتبع الخطوات التالية :

من القوائم اختر Categorize variables → Transform

فيظهر صندوق حوار categorize Variables الذي نرتبه كما يلي :

ند نقر زر OK يتم إضافة المتغير nsalary الى Data Editor وكما يلي :

nsalary

ملاحظة : لا يمكن تبويب المتغيرات الرمزية.

6. الترميز التلقائي Automatic Recode

باستخدام فعالية الترميز التلقائي يمكن تكوين متغير جديد قيمه عبارة عن

أعداد متعاقبة (تصاعدية أو تنازلية) للمتغير القديم (سواء كانت متغيرات عددية أم رمزية)

مثال: الملف التالي يتضمن المتغيرين salary (متغير عددي) و name (متغير

رمزي) وكما يلي :

name	salary
Ahmad	40
Samer	35
Loay	50
Mahmood	80
Ayad	70
Yassin	66
Satar	85

يطلب ترميز المتغيرين بموجب فعالية الترميز التلقائي . لتنفيذ ذلك نتبع الخطوات التالية :

من القوائم نختار Automatic Recode → Transform فيظهر

صندوق حوار Automatic Recode حيث نقوم بإدخال المتغيرين salary و name في

قائمة New Name → Variable ثم نعطي أسمى المتغيرين الجديدين (بعد الترميز) مثلاً

rsalary و rname

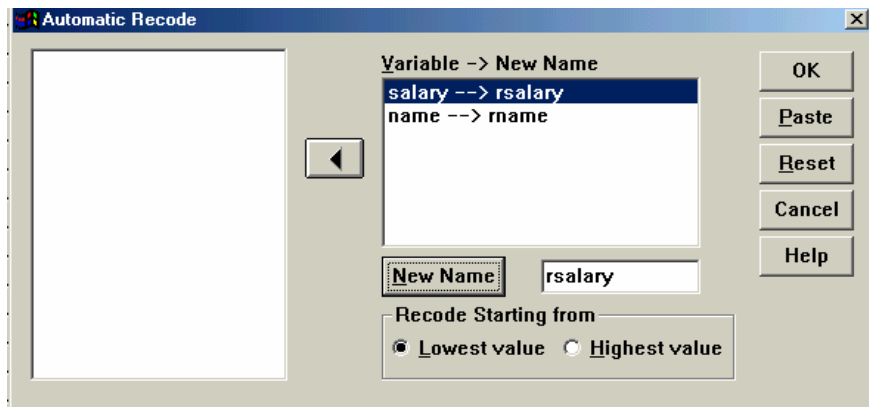
• يمكن إدخال (أو تغيير) اسم المتغير الجديد عن طريق تأشير المتغير بزر الماوس الأيسر ثم إدخال

اسم المتغير في المستطيل المجاور لزر New Name ثم نقر زر New Name .

ويمكن أن

تكون الكودات تصاعدية باختيار Recode starting from Lowest value

تكون الكودات تنازلية باختيار Recode starting from Highest value



يتم ترميز المتغيرات الرمزية حسب التسلسل الأبجائي للحروف وأن الحروف الكبيرة تسبق الحروف الصغيرة .

◀ عند نقر زر OK يضاف المتغيرين الجديدين إلى Data Editor ويظهران كما يلي :

name	salary	rname	rsalary
Ahmad	40	1	2
Samer	35	5	1
Loay	50	3	3
Mahmood	80	4	6
Ayad	70	2	5
Yassin	66	7	4
Satar	85	6	7

7. الأمر Rank Cases:

يمكن بواسطة هذه الفعالية تكوين متغيرات جديدة هي عبارة عن رتب لمتغيرات معينة و تكون هذه الرتب تصاعدية او تنازلية . كما و يمكن إعطاء رتب لمتغير معين بواسطة متغيرات أخرى .

مثال : الملف التالي يحتوي المتغيرات الأجر salary والجنس gender والمنطقة region وكما

يظهر في Data Editor :

region	gender	salary
1	2	30
1	1	70
1	1	100
1	1	50
1	2	45
1	2	36
1	1	70
1	2	25
1	2	22
1	1	42
2	2	15
2	1	100
2	1	110
2	1	88
2	1	92
2	2	55
2	2	32
2	1	47
2	2	20

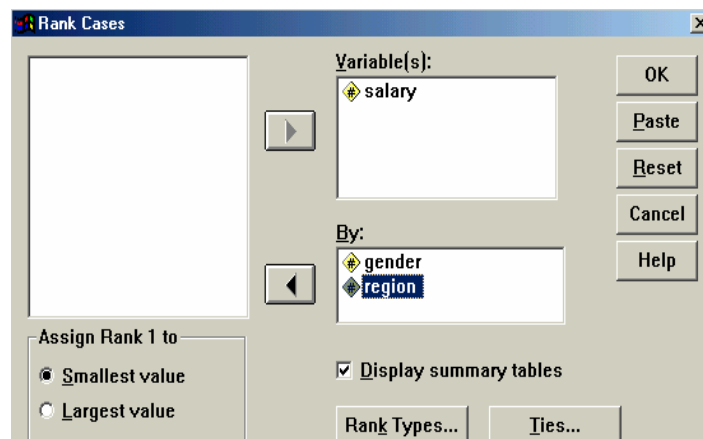
لنفترض أننا نريد إعطاء رتباً تصاعدية لمتغير الأجر salary ضمن فئات الجنس gender ضمن فئات المنطقة region .

لتنفيذ ذلك نتبع الخطوات التالية:

◀ من القوائم نختار Cases → Transform

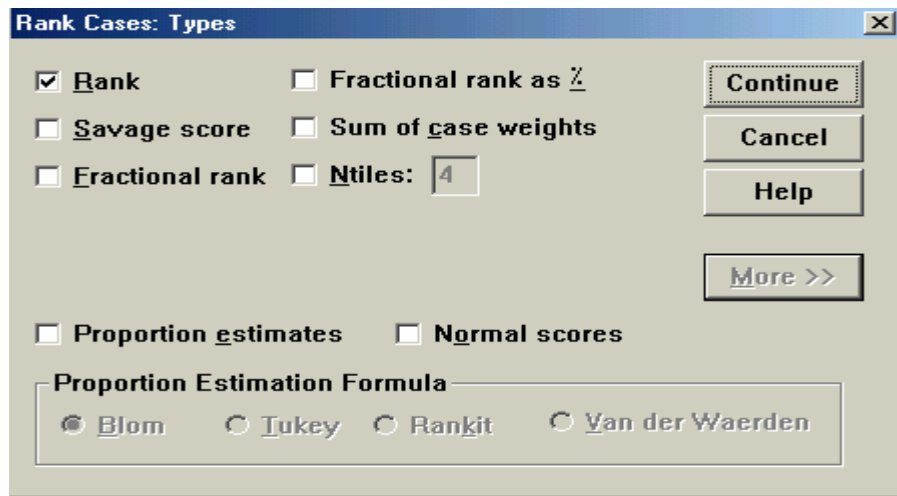
Rank فيظهر صندوق حوار Rank Cases الذي نرتبه

كما يلي :



لقد قمنا بإدخال المتغير الذي نريد إعطاؤه رتباً Salary في خانة Variables أما المتغيرين Gender و region اللذان يتم الترتيب بموجبهما ويعرفان بمتغيرات التجميع Grouping Variables فيتم إدخالهما في خانة By وفي خانة Assign Rank 1 to أشرنا الخيار Smallest Value لإعطاء الرتب تصاعدياً .

◀ اختيار نوع الرتبة انقر زر Rank Types في صندوق حوار Rank cases يظهر صندوق حوار Rank Cases:Types ومنه نختار النوع البسيط للترتيب Rank



◀ عند نقر زر continue ثم زر OK يضاف متغير جديد (متغير الرتب) باسم rsalary إلى Data Editor كما في الجدول التالي:

region	gender	salary	rsalary
1	2	30	3
1	1	70	4
1	1	100	5
1	1	50	2
1	2	45	5
1	2	36	4
1	1	70	4
1	2	25	2
1	2	22	1
1	1	42	1
2	2	15	1
2	1	100	4
2	1	110	5
2	1	88	2
2	1	92	3
2	2	55	4
2	2	32	3
2	1	47	1
2	2	20	2

لاحظ أنه تم إعطاء رتب للمتغير salary ضمن فئات Gender (ذكور ، أناث) ضمن فئات Region

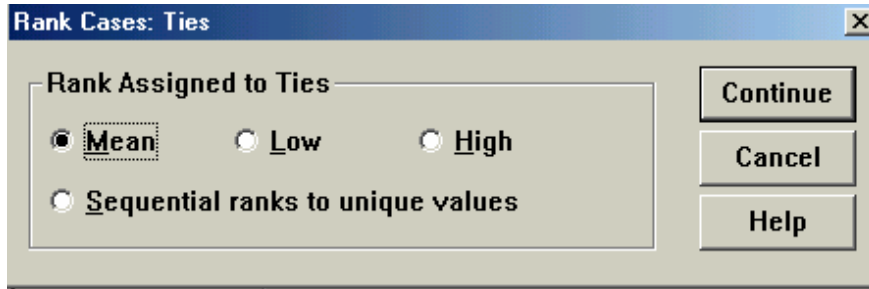
(2و1) .

ملاحظات:

1. يمكن إعطاء رتب للمتغيرات العددية فقط و لا يمكن إعطاء رتب للمتغيرات الرمزية كما يمكن إعطاء رتب لأكثر من متغير عددي في آن واحد.
2. يمكن إعطاء رتب لمتغير معين دون الحاجة إلى استعمال متغيرات التجميع Grouping Variables ففي المثال السابق يمكن إعطاء رتب للمتغير Salary دون استعمال متغيري التجميع

gender و region وفي هذه الحالة تكون خانة By في صندوق حوار Rank Cases خالية من المتغيرات .

3. لغرض تنظيم الرتب للحالات التي لها نفس القيم للمتغير الأصلي (أي تكرار قيمة معينة للمتغير عدة مرات) يتم نقر الزر Ties في صندوق Rank cases حيث يظهر صندوق الحوار التالي الذي يحتوي أربعة خيارات للقيم المكررة :



الجدول التالي يبين كيفية تخصيص الرتب بالطرق الأربعة للقيم المتشابهة :

Value	Mean	Low	High	sequential
10	1	1	1	1
15	3	2	4	2
15	3	2	4	2
15	3	2	4	2
16	5	5	5	3
20	6	6	6	4

→ قيم متشابهة

4. يمكن تكوين الأنواع التالية من الرتب و كما هو وارد في صندوق حوار Rank Cases : Types Rank : وهي الرتبة البسيطة (حيث يتم إعطاء رتبة لكل قيمة من قيم المتغير تعبر عن ترتيبه ضمن المجموعة) .

Savage Scores : تعطي رتب لقيم المتغير بموجب التوزيع الأسى .

Fractional Rank : وهي الرتبة الناتجة من قسمة الرتبة البسيطة لقيم المتغير على مجموع الأوزان لكافة الحالات (أو عدد الحالات في حالة عدم وجود أوزان أي يعتبر الوزن مساويا إلى الواحد) .

Fractional Rank as % : يتم الحصول على هذه الرتبة من حاصل ضرب الرتبة السابقة في 100 .
sum of cases weights : الرتبة تكون متساوية لكافة الحالات و تمثل مجموع الأوزان لكافة الحالات (أو عدد الحالات في حالة عدم وجود أوزان) .

Ntiles : يتم إعطاء رتب بعد تقسيم قيم المتغير إلى مجاميع تعطى كل منها رتبة معينة (بعد ترتيبها تصاعديا أو تنازليا) فإذا اخترنا 4Ntiles (حالة الترتيب تصاعدي) فإنه سيتم إعطاء الرتبة 1 للقيم التي ترتيبها أقل من 25% وتعطى الرتبة 2 للقيم من 25% إلى 50% و الرتبة 3 للقيم 50% إلى 75% والرتبة 4 لـ 75% فما فوق .

proportion estimates : التقديرات النسبية و هي أربعة أنواع :

Blom : يتم إعطاء رتبة لقيم المتغير حسب الصيغة التالية

$$(r - 3/8)/(w + 1/4)$$

حيث إن w يمثل مجموع أوزان الحالات و r يمثل رتبة الحالة

Tukey : يتم إعطاء رتبة حسب الصيغة التالية

$$(r - 1/3)/(w + 1/3)$$

RanKit : تعطى رتبة حسب الصيغة التالية

$$(r - 1/2)/w$$

حيث أن w هو عدد المشاهدات و r هي رتبة الحالة

Vander waerden : تعطى رتبة حسب الصيغة التالية

$$r/(w + 1)$$

حيث w هو مجموع أوزان الحالات و r تمثل الرتبة

المثال التالي يوضح الرتبة المحسوبة بموجب الطرق المختلفة لقيم (حالات) المتغير x

	x	rx	sx	nx	rfr001	per001	n001	px	pro001	pro002	pro003
1	10.00	1.000	-.8571	1	.1429	14.29	7	.0862	.0909	.0714	.1250
2	80.00	6.000	.5929	4	.8571	85.71	7	.7759	.7727	.7857	.7500
3	90.00	7.000	1.5929	4	1.0000	100.00	7	.9138	.9091	.9286	.8750
4	70.00	5.000	.0929	3	.7143	71.43	7	.6379	.6364	.6429	.6250
5	50.00	3.000	-.4905	2	.4286	42.86	7	.3621	.3636	.3571	.3750
6	60.00	4.000	-.2405	3	.5714	57.14	7	.5000	.5000	.5000	.5000
7	40.00	2.000	-.6905	2	.2857	28.57	7	.2241	.2273	.2143	.2500
8	.	.	.	.	.	.	.	.	.	.	.
9	.	.	.	.	.	.	.	.	.	.	.

حيث أن

X : Variable متغير

rx : Rank(simple)

sx : Savage Score

nx : Ntiles

rfr001 : Fractional Rank

$$6/7 = 0.8571$$

تحتسب رتبة الحالة الثانية (مثلاً) كما يلي

per001 : Fractional Rank as %

$$0.8571 * 100 = 85.71$$

تحتسب رتبة الحالة الثانية كما يلي

n001 : Sum of case Weights

px : Proportion Estimate (Blom)

$$(6-3/8)/(7+1/4) = 0.7759$$

تحتسب رتبة الحالة الثانية كما يلي

Pro001 : Proportion Estimate (Tukey)

$$(6-1/3)/(7+1/3) = 0.7727$$

تحتسب رتبة الحالة الثانية كما يلي

Pro002 : Proportion Estimate (Rankit)

$$(6-1/2)/7 = 0.7857$$

تحتسب رتبة الحالة الثانية كما يلي

Pro003 : Proportion Estimate (Vander Waeden)

$$6/(7+1) = 0.7500$$

تحتسب رتبة الحالة الثانية كما يلي

Normal Scores : أن الرتبة المحسوبة بموجب الخيار Normal Scores هي عبارة عن قيم المتغير

الطبيعي القياسي Z Scores التي تقابل التقديرات النسبية التجميعية Estimated Cumulative

Proportions التي هي عبارة عن احتمالات تجميعية والتي سبق وأن استخرجت بالطرق الأربعة المذكورة

(... ، Tukey ، Bloom) الجدول التالي يمثل رتب Normal Scores المحسوبة بالطرق الأربعة :

Blom Tukey Rankit Vander Waeden

	x	nx	nor001	nor002	nor003
1	10.00	-1.3645	-1.3352	-1.4652	-1.1503
2	80.00	.7583	.7479	.7916	.6745
3	90.00	1.3645	1.3352	1.4652	1.1503
4	70.00	.3529	.3488	.3661	.3186
5	50.00	-.3529	-.3488	-.3661	-.3186
6	60.00	.0000	.0000	.0000	.0000
7	40.00	-.7583	-.7479	-.7916	-.6745

مثلاً يمكن إيجاد رتب Normal Scores بصيغة Blom بالاعتماد على التقديرات النسبية Proportion Estimates المحتسبة بصيغة Blom (المتغير px) حيث أن قيم هذا المتغير هي عبارة عن احتمالات Cumulative probabilities → تجميعية

ويمكن إيجاد قيم متغير التوزيع الطبيعي القياسي Z المقابلة لهذه القيم (الاحتمالات التجميعية) باختيار Compute Transform من شريط القوائم ثم اختيار الدالة IDF للتوزيع الطبيعي بمتوسط مساوي للصفر وانحراف معياري مساوي للواحد كما يلي $IDF.NORMAL(px,0,1)$ حيث نحصل على نفس قيم المتغير nx في الجدول أعلاه .

8. السلاسل الزمنية Create Time Series :

السلسلة الزمنية Time Series هي عبارة عن قيم متغير معين خلال فترات زمنية متساوية كالأيام أو الأشهر أو السنين . وقد تعلمنا كيفية إنشاء سلسلة زمنية من خلال الأمر

Data → Define Dates

ونرغب الآن في إجراء بعض العمليات الإحصائية على السلسلة الزمنية من خلال عدة دوال إحصائية تتضمن :- الفروق Differences ، الأوساط المتحركة Moving Averages ، الوسيطات المتحركة running Medians ، المتغيرات الراجعة زمنياً lag ، lead function .

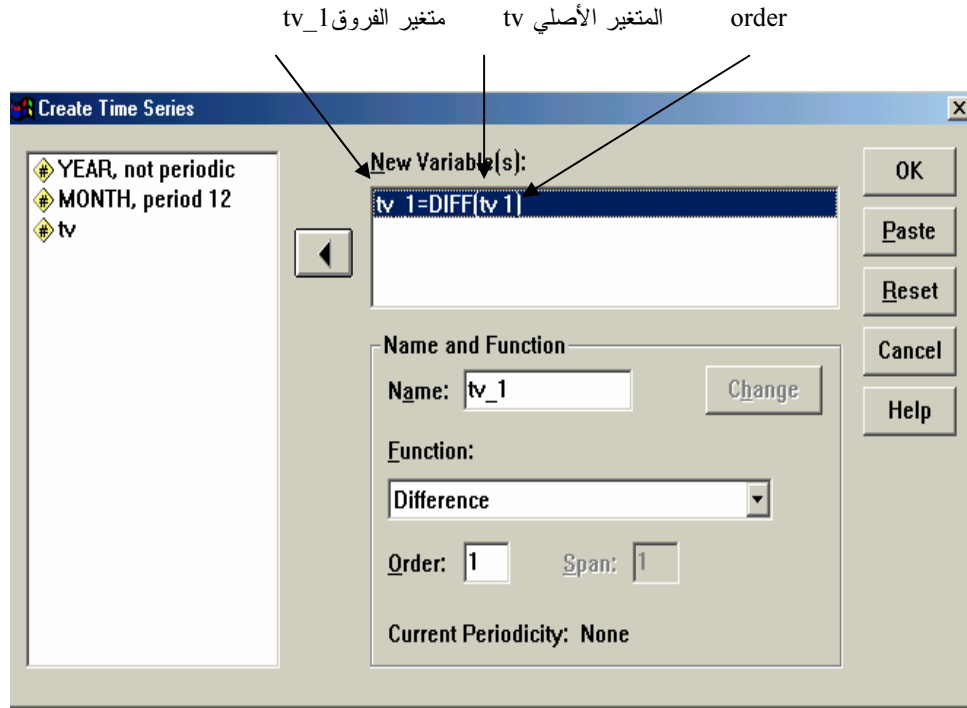
مثال 1 : نفرض إن المتغير tv يمثل المبيعات الشهرية من أجهزة التلفزيون خلال 17 شهراً في مؤسسة معينة (تم إنشاء التاريخ للسلسلة الزمنية بالأمر Define Date → Data) ونرغب في عمل فروق Differences من الدرجة الأولى لهذا المتغير .

year_	month_	date_	tv
2002	1	JAN 2002	274
2002	2	FEB 2002	207
2002	3	MAR 2002	255
2002	4	APR 2002	350
2002	5	MAY 2002	382
2002	6	JUN 2002	383
2002	7	JUL 2002	351
2002	8	AUG 2002	268
2002	9	SEP 2002	380
2002	10	OCT 2002	409
2002	11	NOV 2002	445
2002	12	DEC 2002	455
2003	1	JAN 2003	460
2003	2	FEB 2003	482
2003	3	MAR 2003	449

2003	4 APR 2003	389
2003	5 MAY 2003	398

لتنفيذ ذلك نتبع الخطوات التالية :

من القوائم نختار Create time series → Transform فيظهر صندوق حوار Create time series حيث نقوم بإدخال اسم المتغير tv في مستطيل New variables وذلك بعد اختيار الدالة Difference والرتبة Order مساوية للواحد (أي الفروقات من الدرجة الأولى) .



إن الاسم الافتراضي للمتغير الجديد هو عبارة عن المراتب الستة الأولى من اسم المتغير الأصلي tv متبوعاً بـ (underscore) يليه رقم تسلسلي (أي أنه يكون tv_1).

عند نقر OK يضاف متغير جديد باسم tv_1 إلى Data Editor وكما يلي :

year_	month_	date_	tv	tv_1
2002	1	JAN 2002	274	.
2002	2	FEB 2002	207	-67
2002	3	MAR 2002	255	48
2002	4	APR 2002	350	95
2002	5	MAY 2002	382	32
2002	6	JUN 2002	383	1
2002	7	JUL 2002	351	-32
2002	8	AUG 2002	268	-83
2002	9	SEP 2002	380	112
2002	10	OCT 2002	409	29
2002	11	NOV 2002	445	36
2002	12	DEC 2002	455	10
2003	1	JAN 2003	460	5
2003	2	FEB 2003	482	22
2003	3	MAR 2003	449	-33

2003	4 APR 2003	389	-60
2003	5 MAY 2003	398	9

ملاحظات

1. أن قيمة متغير الفروق في فترة زمنية معينة هي عبارة عن قيمة المتغير الأصلي في نفس الفترة مطروحاً منه قيمة المتغير الأصلي للفترة السابقة إذا كانت المرتبة تساوي 1 أي تكون معادلة الفروق بالشكل التالي $tv_t^* = tv_t - tv_{t-1}$ هو متغير الفروق من المرتبة الأولى و tv هو المتغير الأصلي و t يمثل الفترة الزمنية) ولهذا يكون لمتغير الفروق قيمة مفقودة في الحالة الأولى . أما إذا كانت مرتبة الفروق تساوي 2 فأن معادلة الفروق تنطبق على متغير الفروق من المرتبة الأولى بدلاً من القيم الأصلية ولهذا يكون لمتغير الفروق من المرتبة الثانية قيمتين مفقودتين في الحالتين الأولى والثانية .
2. يمكن تغيير أسم متغير الفروق tv_1 عن طريق كتابة الاسم الجديد في خانة Name ثم نقر زر Change . كما يمكن تغيير نوع الدالة Function بنقر السهم في خانة Function واختيار الدالة من القائمة المنسدلة ثم نقر الزر Change وبنفس الطريقة يتم تغيير المرتبة Order .

مثال 2 (الأوساط المتحركة) :

تستعمل الأوساط المتحركة Moving Averages في إزالة الآثار الموسمية والدورية وغير المنتظمة تاركة الاتجاه العام فقط وتتطلب تقدير طول الدورة Span.

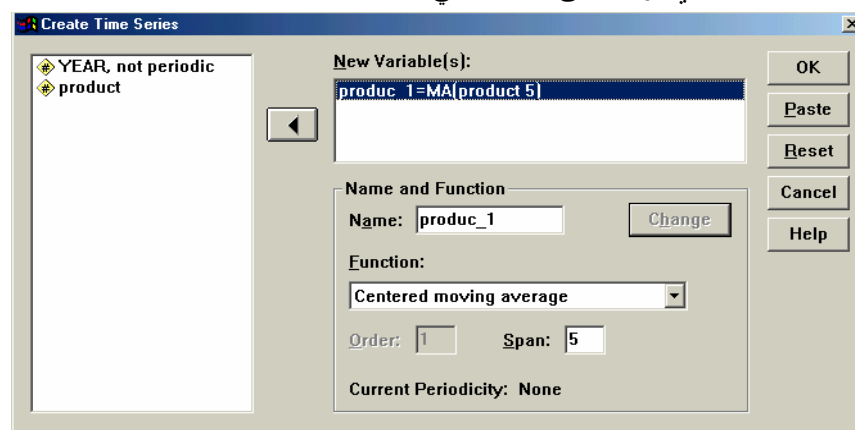
الملف التالي يحتوي على السلسلة الزمنية لقيمة الإنتاج Product للسنوات 1990-2000 وكما يلي

year_	date_	product
1990	1990	50.0
1991	1991	36.5
1992	1992	43.0
1993	1993	44.5
1994	1994	38.9
1995	1995	38.1
1996	1996	32.6
1997	1997	38.7
1998	1998	41.7
1999	1999	41.1
2000	2000	33.8

يطلب حساب الأوساط المتحركة المركزية Centered Moving Averages للسلسلة الزمنية باعتبار أن طول الدورة $Span=5$. لتنفيذ ذلك نتبع الخطوات التالية :

من القوائم نختار Create Time series → Transform فيظهر صندوق حوار Create

Time series الذي نرتبه على النحو التالي :



عند نقر زر OK يضاف متغير جديد بأسم Product_1 الى Data Editor يمثل المتوسطات

المتحركة للسلسلة product وكما يلي :

year_	date_	product	produc_1
1990	1990	50.0	.
1991	1991	36.5	.
1992	1992	43.0	42.6
1993	1993	44.5	40.2
1994	1994	38.9	39.4
1995	1995	38.1	38.6
1996	1996	32.6	38.0
1997	1997	38.7	38.4
1998	1998	41.7	37.6
1999	1999	41.1	.
2000	2000	33.8	.

حيث أن عدد القيم المفقودة في بداية السلسلة ونهايتها هو $n/2$ (n يمثل طول الدورة Span) ويلاحظ أنه تم ترك الحالتين الأولتين والأخيرتين خالية. يحتسب المتوسط المتحرك (في حالة أن طول الدورة = 5) كما يلي :

$$M_1 = \frac{50 + 36.5 + 43 + 44.5 + 38.9}{5} = \frac{212.9}{5} = 42.6$$

$$M_2 = \frac{36.5 + 43 + 44.5 + 38.9 + 38.1}{5} = \frac{201}{5} = \frac{212.9 - 50 + 38.1}{5} = 40.2$$

حيث يقابل المتوسط المتحرك القيمة الوسطية للمجموعة التي احتسب منها في حالة كون طول الدورة فردي أما في حالة كون طول الدورة زوجي span is even فإن المتوسط المتحرك يحتسب عن طريق حساب متوسط كل زوج من المتوسطات غير المركزية Uncentered Means كما في الجدول التالي (في حالة أن Span = 4).

product	متوسط 4 قيم غير مركزي	product_1 متوسط قيمتين Span=4 المتوسط المتحرك
50.0		
36.5		
43.0	43.500	42.113
44.5	40.725	40.925
38.9	41.125	39.825
38.1	38.525	37.800
32.6	37.075	37.425
38.7	37.775	38.150
41.7	38.525	38.675
41.1	38.825	
33.8		

9. تقدير القيم المفقودة Replace Missing Values

أن وجود قيم مفقودة لبعض المتغيرات تعتبر أحياناً عقبة كبيرة تواجه تطبيق أسلوب إحصائي معين ويتوجب في هذه الحالة تقدير القيمة المفقودة حيث يوفر برنامج SPSS هذه الإمكانية .

مثال :

يحتوي المتغير income قيمتين مفقودتين وكما يلي :

بطلب تقدير القيمتين المفقودتين لهذا المتغير .لتنفيذ ذلك نتبع الخطوات التالية :

من القوائم اختر Replace Missing Values → Transform

فيظهر صندوق حوار Replace Missing Values الذي نرتبه كما يلي :

income
95
100
11
120
100
140
.
145
147
150
166
170
190
210
199
.
215
217
230

عدد القيم المجاورة نوع الطريقة Method

حيث يقوم البرنامج بإعطاء أسم افتراضي للمتغير هو عبارة عن

المراتب الستة الأولى من اسم المتغير الأصلي متبوعاً بـ (underscore) يليه رقم تسلسلي ويمكن

تغيير هذا الاسم بكتابة الاسم الجديد في خانة Name ثم نقر زر Change .

في خانة Method يمكن اختيار أحد أنواع التقدير التالية عند نقر السهم المتجه للأسفل :

1. Series Mean : يتم تقدير القيم المفقودة بالمتوسط الكلي للسلسلة الزمنية .
2. Mean of Nearby Points : يتم تقدير القيمة المفقودة بالاعتماد على متوسط القيم المجاورة حيث يتوجب تحديد Span الذي يمثل عدد القيم المعتمدة في حساب المتوسط أعلى وأسفل القيمة المفقودة .
3. Median of nearby points : التقدير يعتمد على الوسيط للقيم المجاورة .
4. linear Interpolation : اعتماد أسلوب الاستكمال الخطي في تقدير القيم المفقودة .
5. linear trend at point : تقدير القيم المفقودة بالقيم التنبؤية Predicted Values المستحصلة من انحدار قيم السلسلة المتوفرة (المتغير المعتمد) على (المتغير المستقل) الذي يأخذ قيماً تسلسلية من 1 إلى n .

وقد اخترنا الأسلوب Mean of nearby points في التقدير .

Span of nearby points : ويستعمل لتحديد عدد القيم المجاورة المستعملة في التقدير ويتضمن خيارين :

1. Number : لتحديد عدد معين من القيم المجاورة .
 2. All : استعمال كافة قيم السلسلة في التقدير وهذا يعني أنه ستكون هناك قيمة واحدة تحل محل كافة القيم المفقودة في السلسلة الزمنية .
وفي هذا المثال تم تحديد Span =2 .
- ◀ عند نقر زر OK في صندوق حوار Replace Missing Values يضاف متغير بأسم income_1 الى ورقة Data Editor وكما يلي :

income	income_1
95	95.0
100	100.0
11	11.0
120	120.0
100	100.0
140	140.0
.	133.0
145	145.0
147	147.0
150	150.0
166	166.0
170	170.0
190	190.0
210	210.0
199	199.0
.	210.3
215	215.0
217	217.0
230	230.0

مثلاً تم حساب القيمة التقديرية للحالة رقم 7 كما يلي :

$$(100 + 140 + 145 + 147) / 4 = 133$$

الفصل الرابع

الإحصاءات الوصفية والجداول التكرارية

Descriptive Statistics

الأمـر (1- 4) Frequencies

مشاهدات المتغير Tall

Tall

80
84
71
72
35
93
91
74
60
63
79
80
70
68
90
92
80
70
63
76
48
90
92
85
83
76
61
99
83
88
74
70
65
51
73
71
72
95
82
70
33
37
32
41
44
49
47
50
59
55
53
56
52
64
60
66

يستعمل هذا الأمر لعرض تكرار كل قيمة لمتغير ما وحساب بعض مقاييس التمرکز والتشتت والربيعات والمئينات مع عرض بعض المخططات البيانية .

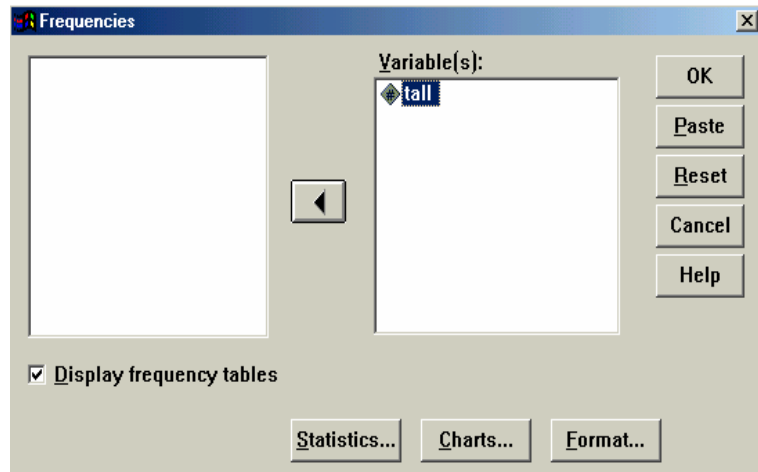
مثال 1 :

المتغير Tall يمثل أطوال 80 نباتاً من نباتات القطن مقدرة بالسنتمترات سنستخدم الأمر Frequencies في حساب تكرارات المشاهدات مع بعض المقاييس الإحصائية لهذا المتغير حسب الخطوات التالية :

← من شريط القوائم اختر

Analyze → Descriptive Statistics → Frequencies

فيظهر صندوق حوار Frequencies الذي يظهر على الشكل التالي بعد إدخال المتغير Tall في قائمة Variables عن طريق نقره بزر الماوس الأيسر (في القائمة الى جهة اليسار) ثم نقر الزر

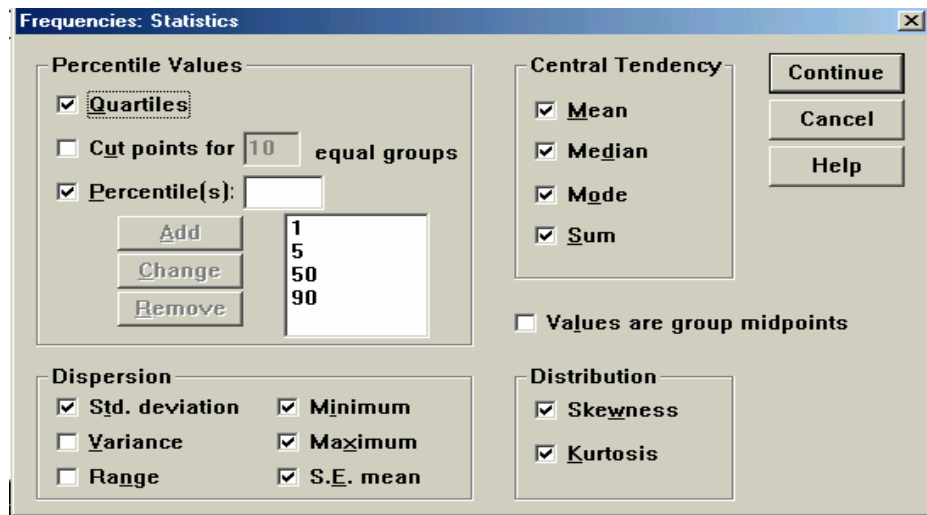


حيث أن :

Display frequency table : يعرض الجدول التكراري عند تأشير المربع المجاور له .

الزر **Statistics** : يعرض بعض المؤشرات الإحصائية حسب حاجة المستفيد ، عند نقره يظهر صندوق حوار Statistics وقد اخترنا المؤشرات التالية للعرض وكما يلي :

حيث أن :



Percentile Values تشمل الربعيات Quartiles و المئينات percentiles (راجع تخطيط Boxplots والربيعات والمئينات في البند (6 - 1) من الفصل السادس حول طريقة الاحتماب) . وقد حددنا المئين الأول ، الخامس ، الخمسين ، التسعين . لتحديد المئين الأول مثلاً نكتب الرقم واحد في المربع المجاور لـ Percentiles ثم ننقر زر Add فيتم إضافة الرقم 1 إلى المستطيل أسفل Percentiles ويستعمل الزر Change لتغيير قيمة المئين والزر Remove لإزالة المئين .

أما الخيار ☐ Groups Equal Cut points for فيحدد القيم التي تقسم البيانات إلى عدد من الفئات المتساوية في الطول والتي تحدد من قبل المستخدم .

Dispersion : لعرض مقاييس التشتت .

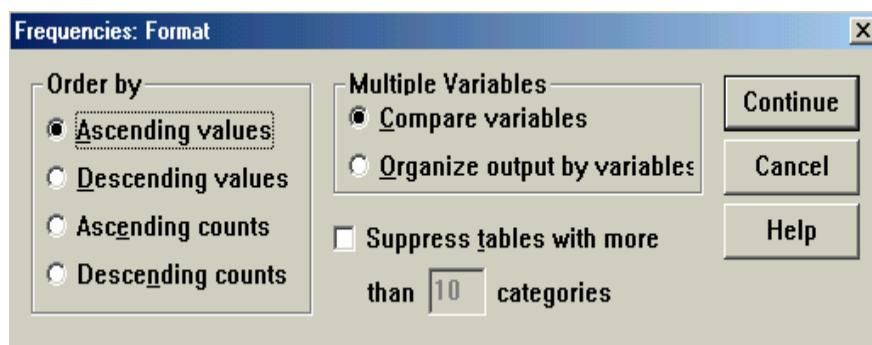
Central Tendency : لعرض مقاييس النزعة المركزية .

Distribution : لعرض مقاييس التوزيع وتشمل الالتواء والتفرطح .

ملاحظة: في حالة تأشير المربع المجاور لـ ☐ Values are group midpoints فإن برنامج SPSS سيقوم بحساب Median و Percentiles Values باعتبار أن بيانات المتغير المعني تمثل مراكز الفئات للبيانات المبوبة .

الزر **chart** : لعرض المخططات Bar ، Pie ...

الزر **format** : عند نقره يظهر صندوق الحوار التالي :



حيث أن :

Ordered by : لترتيب المشاهدات في الجدول التكراري تصاعدياً أو تنازلياً حسب القيم values أو التكرارات Counts وفي هذا المثال اخترنا الترتيب التصاعدي حسب القيم .

Multiple Variables : يستخدم هذا الأمر في حالة وجود أكثر من متغير في القائمة variables في صندوق حوار Frequencies ويشمل ما يلي :

Compare variables : لعرض المؤشرات الإحصائية للمتغيرات كافة في جدول واحد .

Organize output by variables : لعرض مؤشرات كل متغير في جدول مستقل .

وفي هذا المثال لايهم اختيار أي من الخيارين لوجود متغير واحد فقط .

الخيار **supress tables with more than** **categories** : لاختفاء الجدول التكراري للمتغيرات التي يزيد عدد فئاتها عن العدد المحدد من قبل المستفيد .

عند نقر زر OK في صندوق حوار Frequencies تظهر النتائج التالية:

Frequencies

Statistics

TALL

N	Valid	56
	Missing	0
Mean		68.16
Std. Error of Mean		2.29
Median		70.00
Mode		70
Std. Deviation		17.17
Skewness		-.314
Std. Error of Skewness		.319
Kurtosis		-.639
Std. Error of Kurtosis		.628
Minimum		32
Maximum		99
Sum		3817
Percentiles	1	32.00
	5	34.70
	25	55.25
	50	70.00
	75	81.50
	90	91.30

Quartiles(25,50,75)
Percentiles(1,5,50,90)

حيث أن :

Valid : تمثل القيم الصحيحة (الغير مفقودة)

Missing : تمثل القيم المفقودة .

TALL

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	32	1	1.8	1.8	1.8
	33	1	1.8	1.8	3.6
	35	1	1.8	1.8	5.4
	37	1	1.8	1.8	7.1
	41	1	1.8	1.8	8.9
	44	1	1.8	1.8	10.7
	47	1	1.8	1.8	12.5
	48	1	1.8	1.8	14.3
	49	1	1.8	1.8	16.1
	50	1	1.8	1.8	17.9
	51	1	1.8	1.8	19.6
	52	1	1.8	1.8	21.4
	53	1	1.8	1.8	23.2
	55	1	1.8	1.8	25.0
	56	1	1.8	1.8	26.8
	59	1	1.8	1.8	28.6
	60	2	3.6	3.6	32.1
	61	1	1.8	1.8	33.9
	63	2	3.6	3.6	37.5
	64	1	1.8	1.8	39.3
	65	1	1.8	1.8	41.1
	66	1	1.8	1.8	42.9
	68	1	1.8	1.8	44.6
	70	4	7.1	7.1	51.8
	71	2	3.6	3.6	55.4
	72	2	3.6	3.6	58.9
	73	1	1.8	1.8	60.7
	74	2	3.6	3.6	64.3
	76	2	3.6	3.6	67.9
	79	1	1.8	1.8	69.6
	80	3	5.4	5.4	75.0
	82	1	1.8	1.8	76.8
	83	2	3.6	3.6	80.4
	84	1	1.8	1.8	82.1
	85	1	1.8	1.8	83.9
	88	1	1.8	1.8	85.7
	90	2	3.6	3.6	89.3
	91	1	1.8	1.8	91.1
	92	2	3.6	3.6	94.6
	93	1	1.8	1.8	96.4
	95	1	1.8	1.8	98.2
	99	1	1.8	1.8	100.0
Total		56	100.0	100.0	

نلاحظ أن المشاهدات في الجدول قد رتبت تصاعدياً حسب قيم المتغير tall .

(4 - 2) Descriptives الأمر

يفيد هذا الأمر في عرض مقاييس الإحصاء الوصفي لمجموعة من المتغيرات في جدول واحد مع عرض المتغيرات المعيارية zscores .

مثال 2

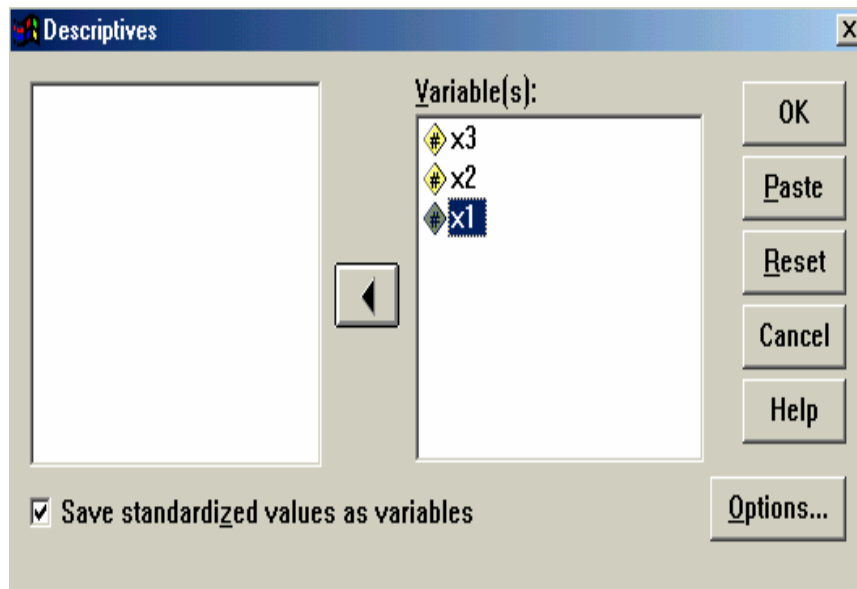
الجدول التالي يتضمن المتغيرات x1,x2,x3 والتي تم إدخالها في SPSS Data Editor لبرنامج

	x1	x2	x3
90	50	12	
70	52	15	
56	55	19	
65	60	22	
85	65	20	
60			
69			
57			
50			
75			
62			
51			
85			

لغرض استخراج المقاييس الوصفية بالأمر Descriptives نتبع الخطوات التالية :

من شريط القوائم اختر Descriptives → Descriptive statistics → Analyze

فيظهر صندوق حوار Descriptives الذي نقوم بترتيبه على الشكل التالي :

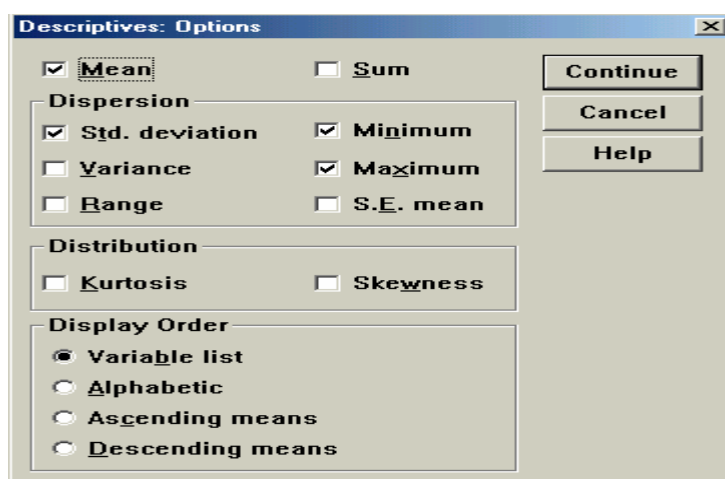


حيث أن :

Save standardized values as variables : لإضافة المتغيرات المعيارية $\frac{x - \bar{x}}{s}$ الى Data

Editor عند تأشير (نقر) المربع المجاور .

الزر **Options** : عند نقره يظهر صندوق الحوار التالي :



حيث يمكن اختيار المؤشرات الوصفية المرغوبة أما الخيار الأخير Display Order فيعمل على ترتيب عرض المقاييس الوصفية للمتغيرات حسب أحد الخيارات التالية :
Variable List : يعرض المقاييس الوصفية حسب تسلسل المتغيرات الوارد في خانة variables في صندوق حوار Descriptives .

Alphabetic : يعرض المقاييس الوصفية حسب الترتيب الأبجدي للمتغيرات .
Ascending means : يعرض المقاييس الوصفية حسب الترتيب التصاعدي للأوساط الحسابية للمتغيرات .
Descending means : يعرض المقاييس الوصفية حسب الترتيب التنازلي للأوساط الحسابية للمتغيرات .
عند نقر زر OK في صندوق حوار Descriptives يعرض المخرج التالي :

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
X3	5	12.00	22.00	17.6000	4.0373
X2	5	50	65	56.40	6.11
X1	13	50.00	90.00	67.3077	13.2186
Valid N (listwise)	5				

حيث تم عرض المؤشرات الوصفية المختارة حسب ترتيب المتغيرات في خانة variables في صندوق حوار Descriptives ، كما تم إضافة الدرجات المعيارية للمتغيرات إلى Data Editor وكما يلي :

x1	x2	x3	zx3	zx2	zx1
90	50	12	-1.387	-1.048	1.717
70	52	15	-.644	-.720	.204
56	55	19	.347	-.229	-.855
65	60	22	1.090	.589	-.175
85	65	20	.594	1.408	1.338
60	.	.	.	.	-.553
69	.	.	.	.	.128
57	.	.	.	.	-.780
50	.	.	.	.	-1.309
75	.	.	.	.	.582
62	.	.	.	.	-.402
51	.	.	.	.	-1.234
85	.	.	.	.	1.338

الفصل الخامس

الجدول المحورية

Pivot Tables

(5 - 1) الجدول المحوري Pivot Table

أن معظم جداول المخرجات في برنامج SPSS أن لم نقل كلها والتي تظهر في شاشة SPSS Viewer كنتاج لتنفيذ أمر معين هي جداول محورية ويتكون الجدول المحوري بصورة عامة من ثلاثة مكونات رئيسية هي :

1. الصفوف Rows .
2. الأعمدة Columns .
3. الطبقات Layers .

وليس من الضروري أن يحتوي الجدول المحوري المكونات أعلاه كلها فقد يحتوي صفوفاً ولا يحتوي أعمدة وبالعكس وقد يحتوي طبقات أو يكون خالياً منها . أن الجدول المحوري هو جدول تفاعلي حيث يمكن إعادة ترتيب الصفوف ، الأعمدة والطبقات فيه وكما ذكرنا فأن هذه الجداول هي مخرجات لأمر معين لبرنامج SPSS ويكون هناك ترتيب مسبق لصفوف وأعمدة وطبقات الجدول (افتراضي) ولكن يمكن محورة الجدول (إعادة ترتيبه) بالصيغة التي يريتها المستخدم ، أما الجداول المحورية في برنامج Excel فهي تشبه جداول SPSS ولكن يتطلب تصميمها أولاً من قبل المستخدم وبذلك يمكن محورتها مستقبلاً .

(5 - 2) تنقيح الجداول المحورية Edit Pivot Tables

يمكن تنقيح الجدول المحوري بنقره مرتين في شاشة SPSS Viewer لتفعيل منقح الجداول المحورية Pivot Tables Editor وكطريقة ثانية أنقر الجدول بزر الماوس الأيمن ثم اختر من القائمة المختصرة Edit → SPSS Pivot Table Object حيث يمكن إجراء فعاليات عديدة على الجدول المحوري .

مثال 1 :

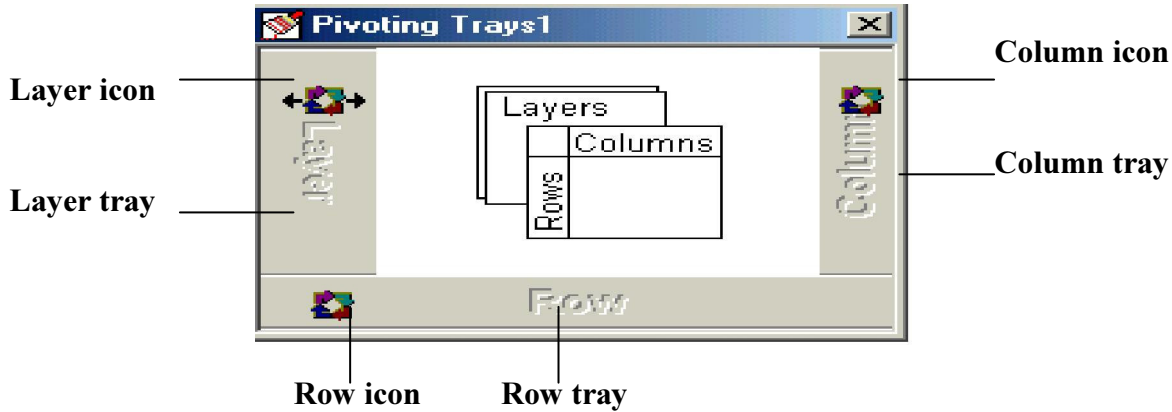
الجدول التالي هو ناتج الأمر Crosstabs → Descriptives → Analyze (للمطلوب الأول من المثال الأول في الفصل السابع) وكما يظهر في شاشة SPSS Viewer.

TREAT * RECOVER Crosstabulation

Count		RECOVER		Total
		a1	b1	
TREAT	a	8	2	10
	b	3	9	12
Total		11	11	22

لتنقيح الجدول أعلاه نتبع الخطوات التالية:

- ☐ أنقر الجدول مرتين لتفعيله .
- ☐ من شريط قوائم الجدول المحوري اختر Pivoting Trays Pivot لظهار صواني المحورة Pivoting Trays وكما يلي :



يتكون

الجدول المحوري من الأبعاد الثلاثة التالية :

1. صينية الصفوف وتحتوي على أيقونة (أيقونات) الصف وهي فئات المتغير Treat لهذا المثال.
2. صينية الأعمدة وتحتوي على أيقونة (أيقونات) العمود وهي فئات المتغير Recover لهذا المثال.
3. صينية الطبقات Layers وتحتوي على أيقونة (أيقونات) الطبقة في هذا المثال توجد طبقة واحدة هي القيم المشاهدة Observed ويمكن تكوين طبقة ثانية تمثل القيم المتوقعة لجدول الاقتران Expected (راجع الملاحظة رقم 5 بعد المطلوب الأول من المثال 1 الوارد في الفصل السابع). حيث يمكن تخيل الجدول بأنه عبارة عن طبقات ثنائية الأبعاد (صفوف وأعمدة) متراكمة واحدة فوق الأخرى وأن الطبقة العلوية وحدها تكون مرئية .

ملاحظة : بالإمكان إظهار عناوين كلاً من الصفوف والأعمدة والطبقات بنقر الأيقونة المقابلة في Pivoting trays مع السحب إلى الأسفل فمثلاً يكون ناتج العملية إظهار العنوان Recover عند نقر أيقونة العمود مع السحب للأسفل .

يمكن استبدال الأعمدة بالصفوف وبالعكس بطريقتين (بعد تفعيل الجدول) :

الطريقة الأولى: باستخدام صواني المحورة Pivoting Trays أسحب أيقونة الصف الى صينية العمود بالزر الأيسر للماوس وأسحب أيقونة العمود الى صينية الصف .

الطريقة الثانية: من شريط قوائم الجدول المحوري اختر Transpose Rows & Columns Pivot → فيكون ناتج كل من الطريقتين الجدول التالي :

TREAT * RECOVER Crosstabulation

Count		TREAT		Total
		a	b	
RECOVER	a1	8	3	11
	b1	2	9	11
Total		10	12	22

في الجدول أعلاه إذا نقلنا أيقونة العمود (المتغير Treat) الى صينية الصف في Pivoting Trays فسيظهر الجدول كالآتي :

TREAT * RECOVER Crosstabulation

Count				
RECOVER	a1	TREAT	a	8
			b	3
		Total		11
	b1	TREAT	a	2
			b	9
		Total		11
Total		TREAT	a	10
			b	12
		Total		22

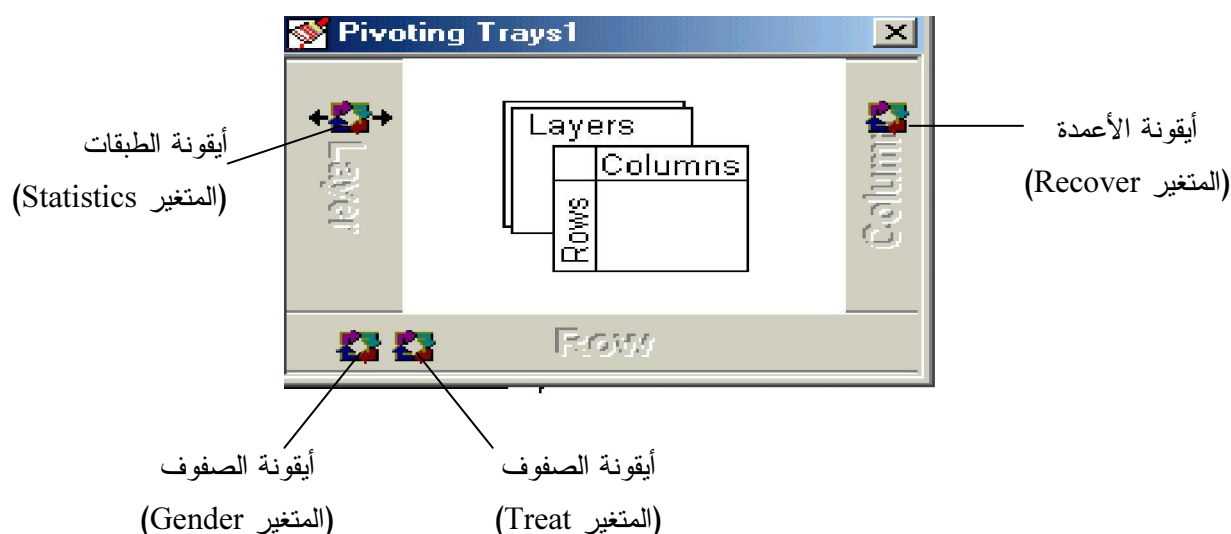
ملاحظة : يمكن إرجاع الإعدادات الأصلية للجدول (بعد تفعيله بنقره مرتين بزر الماوس الأيسر) ثم اختيار الأمر Pivot → Reset Pivots to defaults. SPSS Viewer.

مثال 2 : الجدول التالي هو ناتج الأمر Crosstabs → Descriptives → Analyze) للمطلوب الثاني من المثال الأول في الفصل السابع (وكما يظهر في شاشة SPSS Viewer.

TREAT * RECOVER * GENDER Crosstabulation

Count					
GENDER			RECOVER		Total
			a1	b1	
f	TREAT	a	2	1	3
		b	1	4	5
		Total	3	5	8
m	TREAT	a	6	1	7
		b	2	5	7
		Total	8	6	14

تظهر Pivoting Trays لهذا الجدول كما يلي :



نلاحظ وجود أيقونتين للصفوف إحداهما للمتغير Gender والأخرى للمتغير Treat. لتغيير ترتيب الجدول بحيث تكون هناك طبقتين إحداهما للذكور m والأخرى للإناث f فإننا نقوم بسحب أيقونة Gender من صينية الصفوف الى صينية الطبقات ليصبح لدينا متغيرين في صينية الطبقة أحدهما هو المتغير Statistics ويتكون من طبقة واحدة هي طبقة Count والاخر هو متغير Gender ويتكون من طبقتين أحدهما للذكور m والأخرى للإناث f ليظهر الجدول (بعد تفعيله) كما يلي :

TREAT * RECOVER * GENDER Crosstabulation				
Statistics	Count			
GENDER	f			
		RECOVER		Total
		a1	b1	
TREAT	a	2	1	3
	b	1	4	5
Total		3	5	8

أن الطبقة التي تظهر في الجدول للإناث f لإظهار طبقة معينة مثلاً طبقة الذكور m يمكن أتباع أحد الطرق التالية :

الطريقة الأولى : انقر السهم المتجه للأسفل المجاور للفئة f للمتغير Gender في الجدول أعلاه ثم اختر الطبقة m .

الطريقة الثانية : تلاحظ أن كل أيقونة في صينية الطبقات لها سهمين أيمن وأيسر تستعمل للانتقال من طبقة الى أخرى فعند نقر أحد السهمين في أيقونة Gender في صينية الطبقات يتم الانتقال الى طبقة (أو جدول) الذكور m (السهم الأيمن للانتقال الى الطبقة اللاحقة والسهم الأيسر للانتقال الى الطبقة السابقة) .

الطريقة الثالثة : من شريط قوائم لشاشة SPSS Viewer (و بعد تفعيل الجدول) اختر Pivot → Goto Layer فيظهر صندوق حوار Go to Layer Category وفيه نقوم بتحديد أسم المتغير (Gender) وبضمنه الفئة (الطبقة) التي تعرض حالياً وهي طبقة الإناث f ثم تحديد الفئة (الطبقة) التي نرغب في عرضها وهي طبقة الذكور m .

أن ناتج أي من الطرق أعلاه هو الجدول التالي :

TREAT * RECOVER * GENDER Crosstabulation				
Statistics	Count			
GENDER	m			
		RECOVER		Total
		a1	b1	
TREAT	a	6	1	7
	b	2	5	7
Total		8	6	14

ملاحظة : لاختفاء صف أو عمود في الجدول المحوري نتبع الخطوات التالية (بعد تفعيل الجدول) :

□ انقر عنوان الفئة للعمود أو الصف الي نرغب بإخفائه مثلاً الصف الذي يضم العنوان a للجدول أعلاه مثلاً.

□ طبق Ctrl+Alt+Click فيتم تظليل الصف أو العمود.

□ من قوائم الجدول المحوري اختر View → Hide . فيختفي الصف التابع للفئة a .

لحذف الصف أنقر مفتاح Del في لوحة المفاتيح .

لإظهار الصف الذي سبق إخفائه نتبع الخطوات التالية :

□ أنقر أي عنوان آخر لنفس المتغير (البعد) . للجدول السابق أنقر العنوان b للمتغير Treat .

□ → من

قوائم الجدول المحوري اختر View Show All categories in Treat .

وكطريقة أخرى بدون نقر عنوان آخر لنفس المتغير ، من قوائم الجدول المحوري اختر

View → Show All

(3 - 5) إشارات التعليم Book Marks

يستفاد من إشارات التعليم في خزن أوضاع مختلفة للجدول المحوري كخزن طبقة من طبقات الجدول

أو ترتيب معين للصفوف والأعمدة أو أسبقية العناصر ضمن الصف أو العمود .

مثال 3 :

الجدول التالي يحتوي خمسة مؤشرات إحصائية لثلاثة متغيرات x1,x2,x3 وهو ناتج الأمر

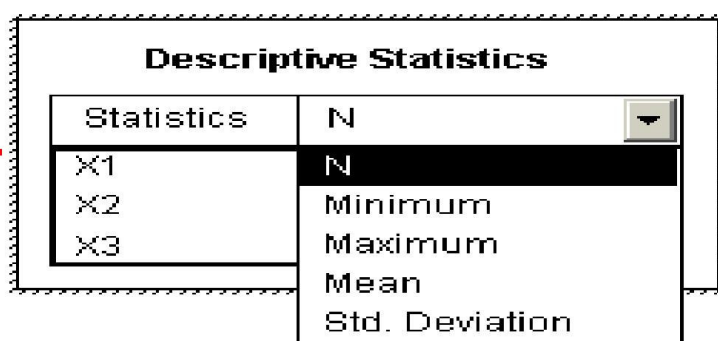
Analyze → descriptive statistics → Frequencies

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
X1	13	50.00	90.00	67.3077	13.2186
X2	5	50	65	56.40	6.11
X3	5	12.00	22.00	17.6000	4.0373

وقد قمنا بمحورة الجدول بحيث يمثل كل مؤشر من المؤشرات الخمس طبقة ضمن الجدول المحوري

وكما يلي (حاول أن تتفد ذلك بنفسك) :



لوضع إشارة تعليم Bookmarks لطبقة Mean نتبع الخطوات التالية :

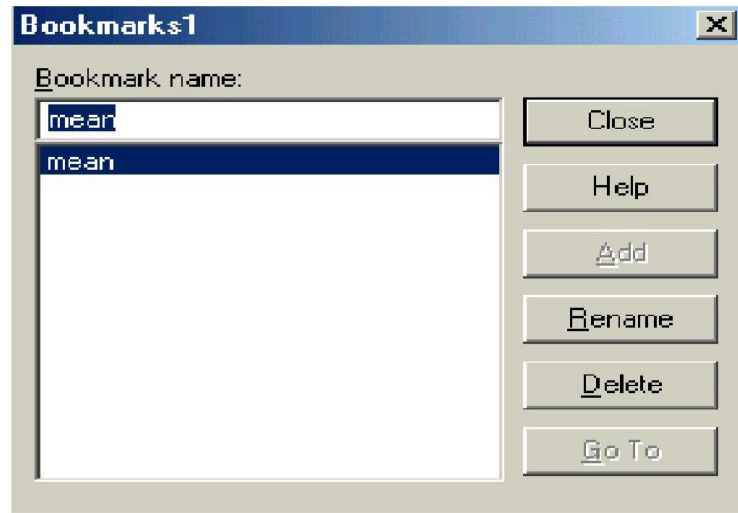
□ قم بتنفيذ الجدول المحوري بنقر الجدول في شاشة SPSS Viewer مرتين .

□ قم بإظهار الطبقة التي تمثل Mean كما يلي :

8

Descriptive Statistics	
Statistics	Mean
X1	67.3077
X2	56.40
X3	17.6000

- من قوائم شاشة SPSS Viewer اختر Bookmarks → Pivot فيظهر صندوق حوار Bookmarks ، أدخل أسم لإشارة التعليم (الاسم mean مثلاً) ثم أنقر زر Add فيضاف الاسم mean الى القائمة في الأسفل ويظهر صندوق bookmarks كما يلي :



وبذلك نكون قد خزنا طبقة Mean كإشارة تعليم باسم mean .ويمكن استرجاع هذه الطبقة حتى لو تم تغيير محورة الجدول الى أي شكل آخر .

- لعرض (استرجاع) إشارة التعليم (الطبقة mean) نتبع الخطوات التالية :
- قم بتنفيذ الجدول المحوري بنقر الجدول في شاشة SPSS Viewer مرتين .
- من القوائم اختر Bookmarks → Pivot فيظهر صندوق حوار Bookmarks .
- أنقر أسم إشارة التعليم mean في القائمة داخل صندوق الحوار المذكور .
- أنقر زر GO TO .
- يقوم البرنامج بعرض الطبقة mean .

الفصل السادس

استكشاف البيانات بالأمر Explore

(6-1) استكشاف البيانات بالأمر Explore

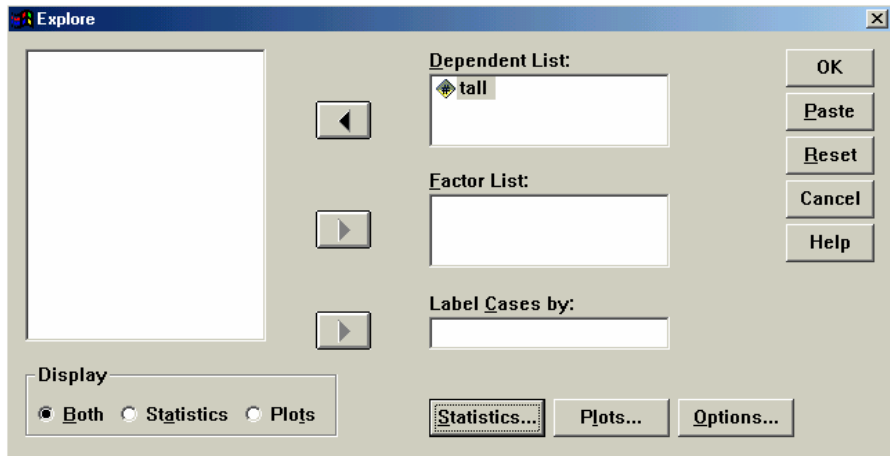
أن أسلوب Explore يعد الخطوة الأولى في التعامل مع البيانات فهو يقدم ملخصاً إحصائياً للبيانات وأعداد مخططات لكل الحالات أو لمجاميع معينة فيها .

حيث أنه يعتمد لفحص البيانات Screening ، أيجاد القيم الشاذة، المقاييس الوصفية ، اختبار الفرضيات فمن خلال هذه العملية يمكن الاستدلال فيما إذا كان التكنيك الإحصائي المستخدم ملائماً أم لا مثلاً يمكن أن يشير الاستكشاف الى الحاجة الى تحويل البيانات Transformation إذا كانت النظرية الإحصائية تفترض التوزيع الطبيعي للبيانات ، أو افتراض تجانس تباين المعاملات في بحوث تصميم التجارب .

مثال 1:

نفترض لدينا المتغير Tall (نفس المتغير الوارد في المثال التابع للبند (4-1) من الفصل الرابع) ونريد تطبيق أمر الاستكشاف Explore على هذا المتغير حيث نتبع الخطوات التالية :

من شريط القوائم اختر Explore → Descriptive Statistics → Analyze
فيظهر صندوق حوار Explore الذي نقوم بترتيبه كما يلي :

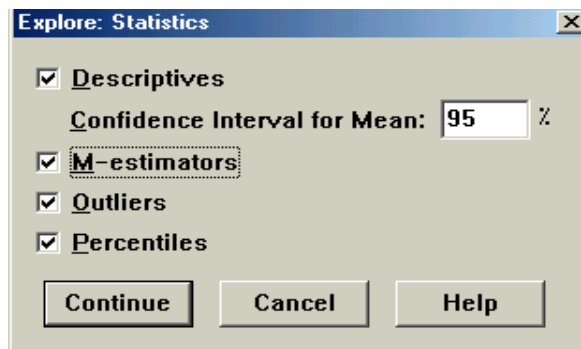


حيث أن :

dependent List: هو المتغير (المتغيرات) المعتمدة التي نرغب في إجراء التحليل الإحصائي عليها.
Factor List: هو متغير تجزئة Break down variable يمكن بواسطته إجراء التحليل الإحصائي لعدة مجاميع من حالات المتغير المعتمد وبصورة مستقلة (كما سيجري توضيحه لاحقاً) . في حالة وجود أكثر من متغير تجزئة فهذا يعني وجود أكثر من طريقة في تجزئة المتغير المعتمد وبالتالي يكون لكل طريقة تجزئة تحليل مستقل . لاحظ في هذا المثال أننا لم نستعمل متغير تجزئة وهذا يعني تحليل قيم المتغير Tall كافة بدون تجزئتها الى مجاميع أصغر . ويمكن أن يكون هذا المتغير عددياً أو رمزياً .

Label Cases by: يمكن إعطاء تعريف للحالات بواسطة متغير معين حيث تعرض الحالات حسب تسلسلها في ملف البيانات (رقم الحالة) ويكون الغرض من هذا التعريف هو إعطاء عنوان للقيم المتطرفة والشاذة في مخطط Box plot .

الزر **Statistics**: عند نقره يظهر صندوق الحوار التالي :



ويحتوي المفردات التالية :

Confidence Interval for Mean : لتكوين فترة ثقة حسب الرغبة 95% أو 99% ...

Descriptives: لإظهار مؤشرات الإحصاء الوصفي Mean ، Standard Deviation ، Skewness ، Kurtosis

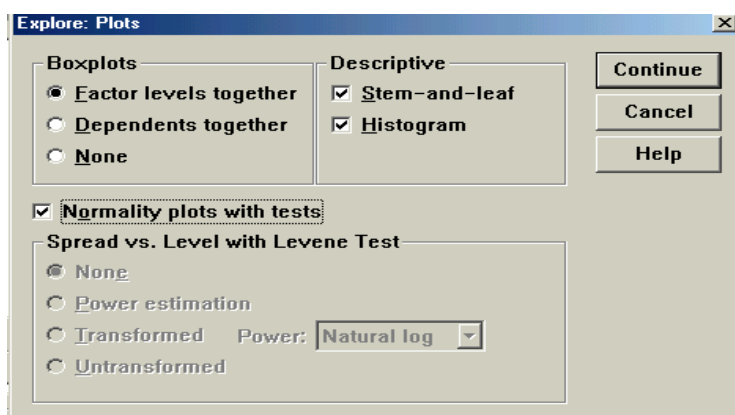
M-estimators : وهي المقدرات الحصينة للنزعة المركزية Robust maximum Likelihood Estimators حيث يتم إعطاء القيم البعيدة عن المركز أوزان أقل من القيم القريبة منه . وهناك أربعة أنواع من المقدرات وهي Tukey, Hample, Andrew, Huber .

Outliers: لإظهار الحالات التي هي أكبر ستة قيم وأقل ستة قيم لمتغير معين ويشار إليها باسم Extreme Values في مخرج برنامج SPSS .

Percentiles : يقوم بعرض القيم المئينية التالية 5،10،25،50،75،90،95، فالمئين الخامس مثلاً 5th Percentile هو قيم المتغير المعتمد التي يسبقها 5% من الحالات ويليها 95% من الحالات بعد ترتيب قيم المتغير تصاعدياً .

عند الرغبة في استخراج أحد هذه المؤشرات نقوم بتأشير المربع المجاور له Check Box .

الزر **Plots** : عند نقره يظهر صندوق الحوار التالي :



ويحتوي المفردات التالية :

1. مخطط Boxplots

ويطلق عليه أيضاً تسمية box-and-whisker plot يتكون هذا المخطط من ثلاثة اجزاء وكالتالي :

أ. الصندوق Box : ويشمل المكونات التالية :

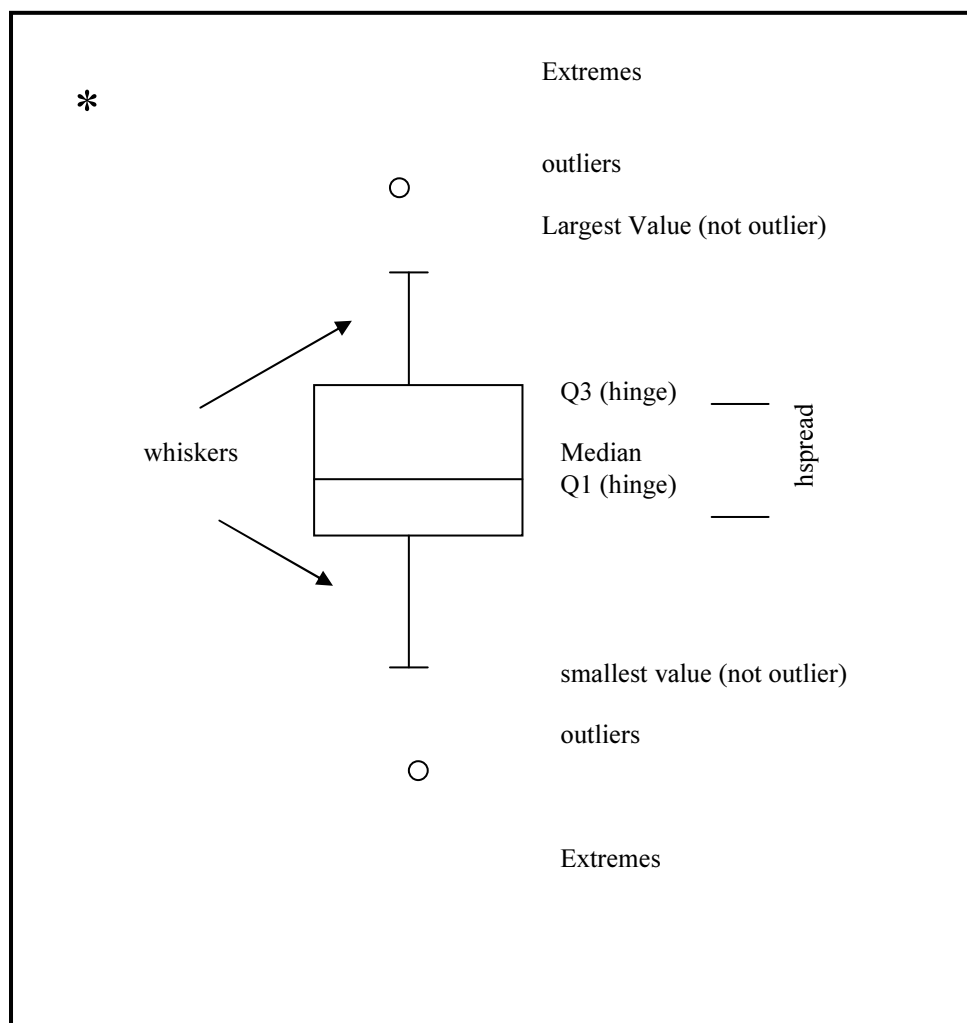
الربيع الأول Q1 : وهي القيمة التي تسبقها 25% من مشاهدات المتغير المعني عند ترتيبها تصاعدياً .

الربع الثاني Q2 : وهي القيمة التي تقسم البيانات الى قسمين متساويين أي يسبقها 50% من المشاهدات عند ترتيبها تصاعدياً وهذه القيمة هي الوسيط Median .

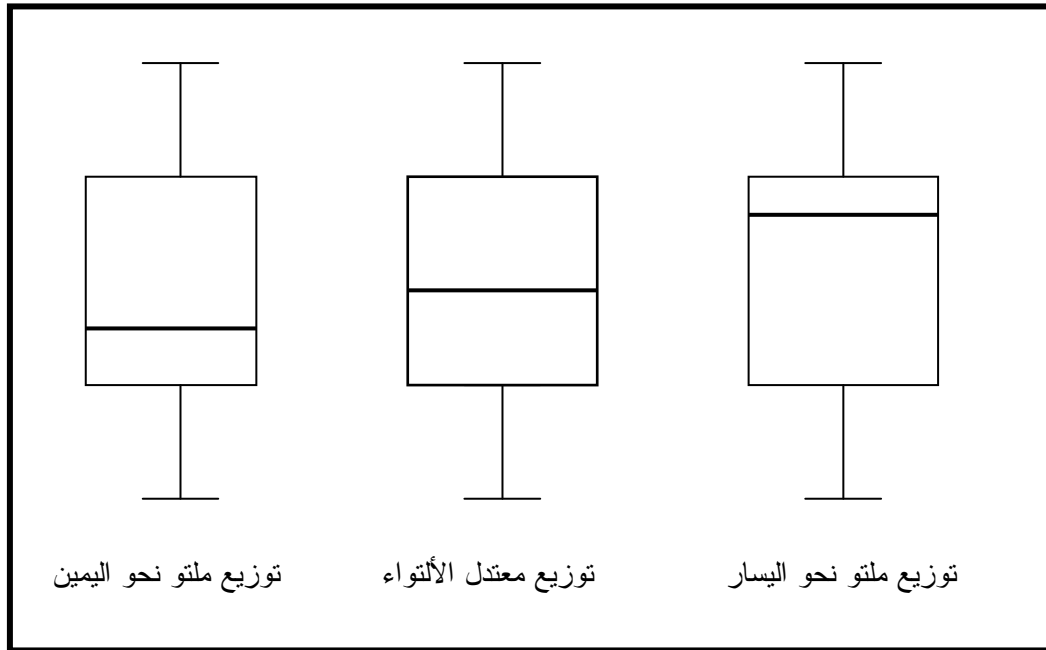
الربع الثالث Q3 : وهي القيمة التي تسبقها 75% من مشاهدات المتغير المعني عند ترتيبها تصاعدياً .
 أن Q1 و Q3 تمثلان حافتي الصندوق ويطلق عليها hinges . أما طول الصندوق فهو Q3-Q1
 ويطلق عليه hspread ويعرف بالمدى الربيعي Inter Quartile Range . أما الخط الوسطي
 داخل الصندوق فهو الوسيط Median

ب. الأسطوانات خارج الصندوق Whiskers: وتمتد من حافتي الصندوق الى أعلى و أقل قيمة غير شاذة .
 ج. القيم الشاذة outliers والقيم المتطرفة Extremes :

حيث أن القيم الشاذة تبعد عن حافتي الصندوق أكثر من 1.5 من طول الصندوق أما القيم المتطرفة فتبعد عن حافتي الصندوق أكثر من 3 من طول الصندوق وكما في الشكل التالي :



أن هذا المخطط يعطي فكرة عن توزيع المشاهدات (الالتواء Skewness) فإذا لم يكن الوسيط في منتصف الصندوق فأن التوزيع ملتو أما إذا كان الوسيط أقرب الى الربع الأول فأن التوزيع ملتو الى اليمين (موجب الالتواء) وإذا كان الوسيط أقرب الى الربع الثالث فأن التوزيع ملتو الى اليسار (سالب الالتواء) كما في الشكل التالي كما أن Whiskers تعطي فكرة عن طول ذيل التوزيع .



ويتضمن مخطط Boxplots الخيارات التالية :

Factor Levels together : يتم عرض المخطط للمجاميع الجزئية لكل متغير معتمد .

Dependent together: يتم عرض المخطط للمتغيرات المعتمدة داخل كل مجموعة جزئية .

None: عدم عرض مخطط Boxplot .

في هذا المثال يمكن تأشير أي من الخيارين الأول أو الثاني وذلك لعدم وجود متغير تجزئة Factor Variable .

2. المخططات الوصفية Descriptive

يمكن عرض المخططات التالية :

أ-مخطط Stem-and-leaf : في هذا المخطط تتم قسمة أي رقم الى جزأين الأول Stem (الجذع) والثاني Leaf (الورقة) ويمثل Stem الجزء الأيسر وleaf الجزء الأيمن . فإذا كانت لدينا القيم التالية 5،7،12،15،16،20،21،23،30 فأننا نقسمها الى جزأين الأول Stem الذي يمثل خانة العشرات والثاني Leaf الذي يمثل خانة الآحاد وكأن المتغير قسم الى فئات طول كل منها 10 درجات ويلاحظ أن هذا المخطط يشبه مخطط histogram والفرق بينهما أن التكرارات في Histogram تمثل بأعمدة في حين تمثل بالقيم الحقيقية في حالة Stem-and-Leaf ولذلك فإنه يعكس معلومات عن طبيعة القيم الموجودة .

VAR1 Stem-and-Leaf Plot		
Frequency	Stem	Leaf
2.00	0	57
3.00	1	256
3.00	2	013
1.00	3	0
Stem width: 10.00		
Each leaf: 1 case(s)		

ب- المدرج التكراري Histogram : لعرض المدرج التكراري .

Normality Plots with Tests.3 : عمل مخططات لاختبار التوزيع الطبيعي للبيانات .

4. spread vs. Level with Levene Test : لاختبار تجانس التباين في تصميم التجارب . علماً أن هذا الاختبار لا يكون فعالاً (نشطاً) ألا عندما يكون هناك متغير تجزئة Factor Variable . (سيتم توضيح هذا الخيار بالتفصيل في المثال 3) .

عند نقر زر OK في صندوق حوار Explore تظهر النتائج التالية :

Explore

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
TALL	56	100.0%	0	.0%	56	100.0%

Descriptives

			Statistic	Std. Error
TALL	Mean		68.1607	2.2948
	95% Confidence Interval for Mean	Lower Bound	63.5618	
		Upper Bound	72.7596	
	5% Trimmed Mean		68.5635	
	Median		70.0000	
	Variance		294.901	
	Std. Deviation		17.1727	
	Minimum		32.00	
	Maximum		99.00	
	Range		67.00	
	Interquartile Range		26.2500	
	Skewness		-.314	.319
	Kurtosis		-.639	.628

الخطأ المعياري للأوساط Standard Error

Mean = 68.1607, Std.Deviation = 17.1727, Std. Error = $SD/\sqrt{n} = 17.1727/\sqrt{56} = 2.2948$

حيث أن Std. Error هو الخطأ المعياري (الانحراف المعياري للأوساط الحسابية المحسوبة من العينة) .

تكوين فترة ثقة 95% لمتوسط المجتمع μ : تستخرج كما يلي :

$$\bar{X} \pm t_{0.025,55} * Std.Error$$

حيث أن t تستخرج من جدول توزيع t لاحتمال 0.025 (المساحة الى يمين قيمة المتغير t) ودرجة حرية

n-1=55 أو يمكن استخراجها مباشرة من برنامج SPSS ودون الرجوع الى الجداول من خلال الأمر

Compute Variable → Transform ثم اختيار الدالة IDF.T(p,df) في صندوق حوار Compute Variable

حيث أن p تمثل الاحتمال التجميعي 0.975 و df = 55 و p = 1-0.025 = 0.975 و عليه فإن IDF.T(0.975,55) =

2 t.0.025,55 = و عليه تكون حدود فترة ثقة 95% للمتوسط كما يلي :

$$\text{Upper Bound} = 68.1607 + 2 * 2.2948 = 72.75$$

$$\text{Lower Bound} = 68.1607 - 2 * 2.2948 = 63.57$$

وتكتب فترة الثقة على الصورة التالية :

$$\text{Pr}(63.57 < \mu < 72.57) = 95\%$$

حيث أن Pr تمثل الاحتمال أي أن احتمال وقوع متوسط المجتمع بين القيمتين 63.57 و 72.57 يساوي 95%

ملاحظة :

يمكن إنجاز العديد من الفعاليات على مخرجات برنامج SPSS مثلاً جدول Descriptives أعلاه حيث يمكن تغيير هيئة الجدول ، عدد المراتب العشرية Decimals ، نوع الخط Font وحجمه ولون الكتابة ، الموقع Alignment ، عرض الأعمدة ويتم ذلك بنقر الجدول المطلوب مرتين في SPSS Viewer (شاشة عرض المخرجات) ثم اختيار الأمر Table Properties → Format لأجراء فعاليات على الجدول ككل ولتغيير مواصفات خلية معينة في الجدول يتم بعد الوقوف على الخلية المعنية ثم اختيار الأمر Format → Cell Properties .

الوسط الحسابي المشذب Trimmed Mean

يستخرج بعد ترتيب القيم تصاعدياً ثم حذف 5% من القيم من الأعلى و 5% من القيم من الأسفل ثم حساب المتوسط للقيم المتبقية الذي لن يتأثر بالقيم الشاذة .

في هذا المثال كان عدد قيم المتغير tall هو 56 قيمة وأن 5% من القيم يساوي 2.8 أي أنه يجب حذف 2.8 قيمة من الأعلى و 2.8 قيمة من الأسفل أما القيم المتبقية بعد الحذف فهي $56 * 0.90 = 50.4$ حيث تستخرج قيمة الوسط المشذب كما يلي :

$$\text{TrimmedMean} = \frac{68.1607 * 56 - 99 - 95 - 32 - 33 - 0.8 * 93 - 0.8 * 35}{50.4} = \frac{3455.5992}{50.4} = 68.5635$$

حيث أن $68.1607 * 56$ تمثل مجموع 56 قيمة وأن 99 و 95 و 32 و 33 تمثلان أعلى قيمتين وأن 32 و 33 تمثلان أدنى قيمتين وأن 93 و 35 تمثلان القيمتين العليا والدنيا التالية للقيمتين العليا والدنيا على الترتيب . وقد حصلنا على هذه القيم من مخرجات برنامج SPSS للقيم المتطرفة Extremes كما في الجدول التالي :

Extreme Values

			Case Number	Value
TALL	Highest	1	28	99.00
		2	38	95.00
		3	6	93.00
		4	16	92.00
		5	23	92.00
	Lowest	1	43	32.00
		2	41	33.00
		3	5	35.00
		4	42	37.00
		5	44	41.00

الربيعات والمئينات

تتكون الربيعات Quartiles من ثلاثة قيم Q1, Q2, Q3 وتقسّم المجموعة الى أربعة أقسام متساوية (راجع مخطط Boxplot حول تعريف الربيعات) أما المئينات Percentiles فتقسم المجموعة الى مائة قسم متساوي فالمئين الخامس تسبقه 5% من البيانات عند ترتيبها تصاعدياً وتليه 95% من البيانات والمئين العاشر تسبقه 10% من البيانات عند ترتيبها تصاعدياً وتليه 90% من البيانات ... وهكذا .

علماً أن الربع الأول Q1 يقابل المئين 25 (25th Percentile)

وأن الربع الثاني (الوسيط) Q2 يقابل المئين 50 (50th Percentile)

وأن الربع الثالث Q3 يقابل المئين 75 (75th Percentile)

الجدول التالي يمثل مخرجات البرنامج للمئينات :

Percentiles

	Percentiles						
	5	10	25	50	75	90	95
Weighted Average(Definitior TALL	34.7000	43.1000	55.2500	70.0000	81.5000	91.3000	93.3000
Tukey's Hinges TALL			55.5000	70.0000	81.0000		

وعليه فإن $Q1 = 55.2$, $Q2 = \text{Median} = 70$, $Q3 = 81.5$

Inter quartile Range = $81.5 - 55.2 = 26.25$

أن الطريقة المتبعة في حساب المئينات هي طريقة المتوسط الموزون Weighted Average Method فلحساب ترتيب المئين الخامس مثلاً تستعمل الصيغة التالية $(n+1)*P = 57*0.05 = 2.85$ أي أن ترتيب المئين الخامس هو 2.85 عند ترتيب القيم تصاعدياً . وعليه فإن قيمة المئين الخامس تقع بين القيمة التي ترتيبها 2 وهي 33 والقيمة التي ترتيبها 3 وهي 35 (من جدول Extreme Values) وعليه نحصل على قيمة المئين بأجراء استكمال خطي Interpolation بين القيمتين المستخرجتين وكما يلي :

$$5\text{th Percentile} = 33 * 0.15 + 35 * 0.85 = 34.7$$

أختبار التوزيع الطبيعي للبيانات من نسبة معامل الالتواء

يمكن أختبار التوزيع الطبيعي من ملاحظة نسبة معامل الالتواء Skewness الى الخطأ المعياري له (جدول Descriptives) وفي هذا المثال كانت النسبة كما يلي $-0.98 = -0.314 / 0.319$ وبما أن هذه النسبة تقع ضمن المدى (-2,2) فأذن نقبل فرضية العدم القائلة بأن المتغير Tall يتبع التوزيع الطبيعي . أما إذا كانت النسبة اكبر من 2 فهذا يعني أن التوزيع ملتو التواءاً موجباً (الى اليمين) وإذا كانت النسبة أقل من -2 فهذا يعني أن التوزيع ملتو التواءاً سالباً (الى اليسار) .

المخرج التالي يبين التقديرات الحصينة للوسط الحسابي M-Estimators والمحسوبة بأربعة طرق ويلاحظ عدم وجود فرق محسوس بين المتوسط المحتسب بموجب هذه الطرق والمتوسط المحتسب بالطريقة الاعتيادية للمتغير tall.

M-Estimators

	Huber's M-Estimator ^a	Tukey's Biweight ^b	Hampel's M-Estimator ^c	Andrews' Wave ^d
TALL	69.1469	69.3859	68.9754	69.3749

- 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 $1.340 \cdot \pi$.

مخطط Stem-and-Leaf

المخرج التالي يمثل مخطط Stem-and-Leaf لمتغير الطول ومنه يستدل أن هذا المتغير يتبع التوزيع الطبيعي فهو يبدو متماثلاً تقريباً حول المتوسط .

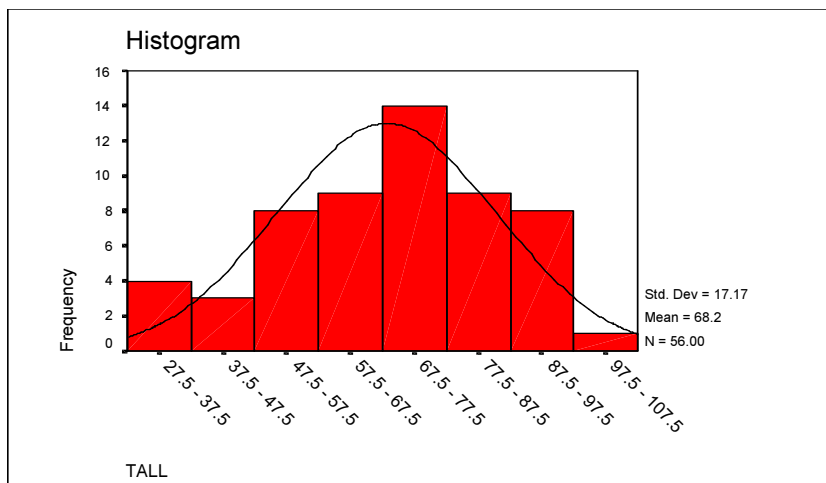
TALL Stem-and-Leaf Plot

Frequency	Stem & Leaf
4.00	3 . 2357
5.00	4 . 14789
7.00	5 . 0123569
9.00	6 . 001334568
14.00	7 . 00001122344669
9.00	8 . 000233458
8.00	9 . 00122359

Stem width: 10.00
Each leaf: 1 case(s)

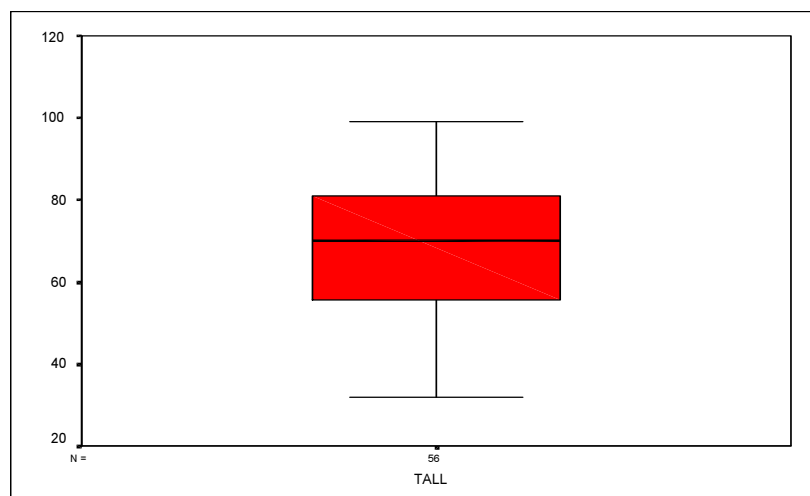
المدرج التكراري Histogram

المخطط التالي يمثل المدرج التكراري لمتغير الأطوال وقد اضيف إليه منحنى التوزيع الطبيعي الذي يبدو متماثلاً حول المتوسط (راجع فصل المخططات لمزيد من التفاصيل حول المدرج التكراري) .



Boxplots

المخرج التالي يمثل مخطط Boxplots للمتغير Tall ومنه يلاحظ عدم وجود قيم متطرفة أو شاذة وأن موقع الوسيط في وسط الصندوق تقريباً مما يشير إلى أن المتغير يتوزع طبيعياً (عدم وجود التواء) .



الخيار Normality Plots with Test : يتيح هذا الخيار إمكانية اختبار التوزيع الطبيعي للبيانات باستخدام

اختبار Kolmogorov-Smirnov بالإضافة إلى اختبار التوزيع الطبيعي للبيانات باستخدام نوعين من

المخططات هما :

- Normal Q-Q Plot

- Detrended Normal Q-Q Plot

1. اختبار Komogrov-Smirnov للتوزيع الطبيعي: وبعد من الاختبارات اللامعلمية للتوزيع الطبيعي

Non-Parametric Goodness of Fit Test حيث تختبر فرضية العدم القائلة بأن مشاهدات متغير معين

تتبع التوزيع الطبيعي ضد الفرضية البديلة القائلة بأن البيانات لا تتوزع طبيعياً. وقد تم الحصول على المخرج

التالي لاختبار التوزيع الطبيعي لملاحظات المتغير tall .

Tests of Normality

	Kolmogorov-Smirnov ^a		
	Statistic	df	Sig.
TALL	.096	56	.200*

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

a. Lilliefors Significance Correction

حيث تستخدم الإحصائية D في الاختبار $D = \sup_x |F_S(x) - F_T(x)|$ حيث أن $F_S(x)$ تمثل دالة التوزيع

التجميعي للعينة وأن $F_T(x)$ تمثل دالة التوزيع النظري (التوزيع الطبيعي في هذا المثال) والتي تقارن مع

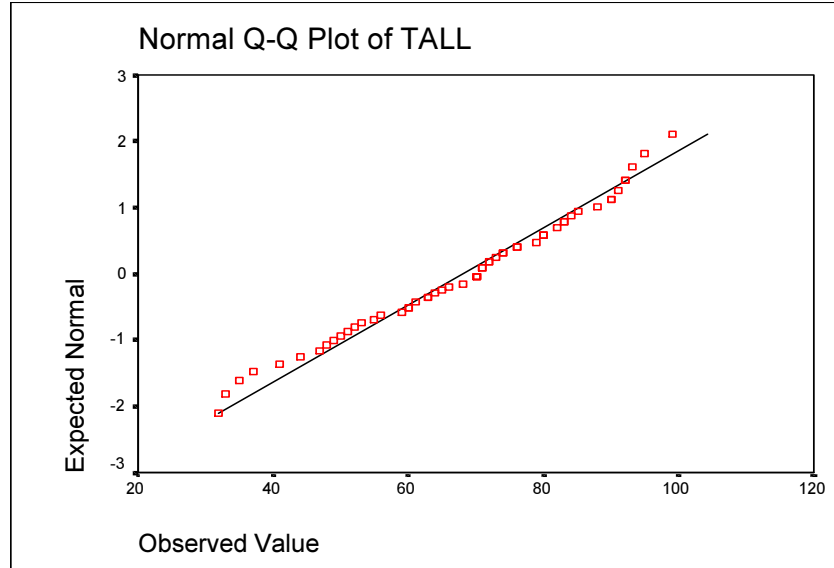
القيمة النظرية لـ D من جدول Kolmogrov بستوى دلالة معين ودرجت حرية n (حجم العينة) . نلاحظ

في هذا المثال أن D = .096 وأن P-Value = 0.20 > 0.05 مما يدعونا إلى قبول فرضية العدم بمستوى

دلالة 5% أي أن توزيع الأطوال يتبع التوزيع الطبيعي .

2. مخطط Normal Q-Q Plot

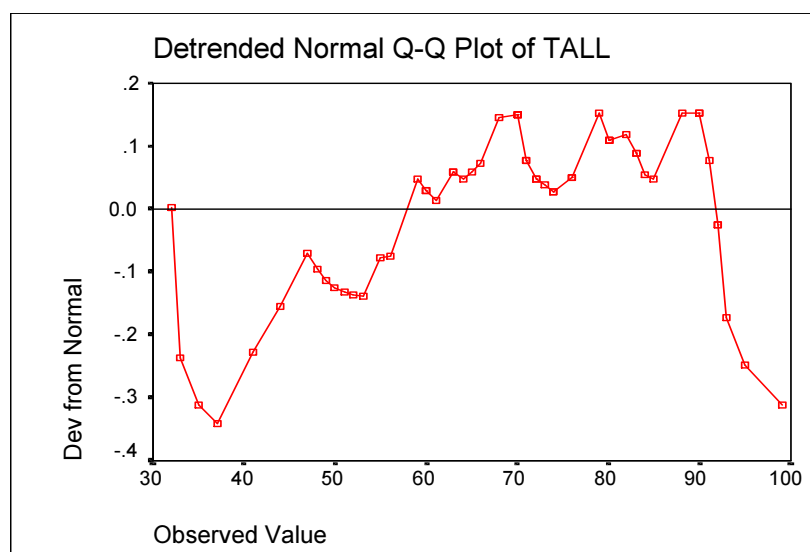
في هذا المخطط يتم رسم كل مشاهدة من البيانات الأصلية على المحور الأفقي مقابل قيم التوزيع الطبيعي القياسي المتوقعة لها Expected Z Score (راجع أمر Rank Cases حول طريقة احتساب القيم القياسية المتوقعة) كما في المخطط التالي للمتغير tall .



أن كل نقطة على الخط المستقيم في المخطط أعلاه تمثل بواسطة القيم المتوقعة لدرجات التوزيع الطبيعي على المحور العمودي مقابل الدرجات المعيارية للتوزيع الطبيعي للبيانات على المحور الأفقي وعلى الشكل التالي $(-1.5, -1.5)$, $(0, 0)$, $(1.5, 1.5)$... وهكذا . فإذا كانت العينة مسحوبة من مجتمع يتوزع طبيعياً فإن نقاط شكل الانتشار ستقع تقريباً بمحاذاة الخط المستقيم أما إذا كانت تقع بعيدة عن الخط المستقيم فهذا يعني أن البيانات لا تتبع التوزيع الطبيعي . أن انتشار النقاط في الشكل أعلاه بمحاذاة الخط المستقيم يشير إلى أن مشاهدات المتغير tall تتبع التوزيع الطبيعي .

3. مخطط Detrended Normal Q-Q Plot

كما ذكرنا بالنسبة للمخطط السابق أنه في حالة عدم انتشار المشاهدات بمحاذاة الخط المستقيم فأننا نستنتج أن البيانات لا تتبع التوزيع الطبيعي ولكن السؤال الذي يطرح إلى أي مدى تكون الانحرافات عن الخط المستقيم كافية لنقرر أن المشاهدات لا تتوزع طبيعياً . أن برنامج SPSS يعرض مخطط Detrended Normal Q-Q Plot لهذا الغرض حيث يتم تمثيل المشاهدات الأصلية على المحور الأفقي أما المحور العمودي فيمثل انحرافات القيم المعيارية للمشاهدات عن قيم التوزيع الطبيعي المتوقعة لها . فإذا كانت معظم نقاط شكل الانتشار (يمكن أن نقرر النسبة 95% أو 90%) تقع ضمن المدى $(-2, 2)$ فيمكن أن نقرر أن التوزيع الطبيعي للبيانات والعكس صحيح . في المخطط أدناه نلاحظ أن معظم الانحرافات للمتغير Tall تقع ضمن المدى $(-2, 2)$ عدا ستة نقاط تقريباً (أي ما يقارب 90% من الانحرافات) وعليه يمكن أن نقرر أن متغير الأطوال يتوزع توزيعاً طبيعياً .



تم ربط نقاط المخطط من خلال Linear Interpolation في Chart Editor لتبسيط عملية رصد الانحرافات والذي يمكن الدخول فيه بنقر المخطط بزر الماوس الأيسر مرتين .

ملاحظة : في حالة الاستنتاج بأن مشاهدات متغير ما لا تتوزع طبيعياً فإنه يمكن عمل تحويلات للبيانات Transformation لغرض جعل التوزيع طبيعي (نفس الطريقة المتبعة في المثال 3) .

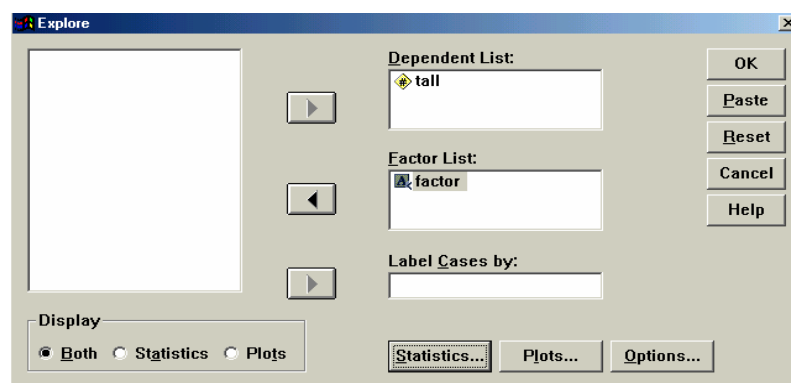
مثال 2

في المثال السابق للمتغير Tall لو اعتبرنا أن 22 حالة الأولى تمثل أطوال الفئة الأولى من البيانات وأن 34 حالة الثانية تمثل أطوال الفئة الثانية من البيانات ونرغب في إجراء تحليل إحصائي لكل مجموعة على حدة في هذه الحالة يتوجب إدخال متغير تجزئة (لنسمه Factor) في قائمة Factor List في صندوق حوار Explore الذي يتم ترتيبه بالشكل التالي :

أما ترتيب البيانات في Data Editor فيكون كما يلي :

البيانات في Data Editor

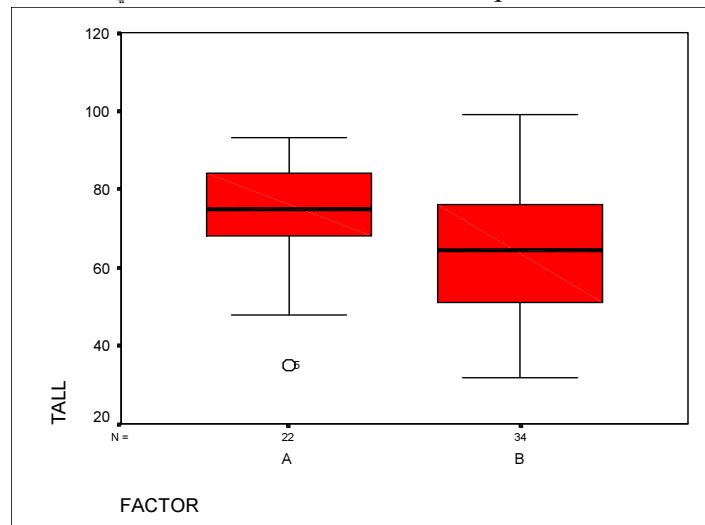
عند تنفيذ أمر Explore فإن المخرجات تعرض لكل من الفئتين A و B بصورة مستقلة مثلاً القيم المتطرفة تعرض كما يلي :



tall	factor
80	A
84	A
71	A
72	A
35	A
93	A
91	A
74	A
60	A
63	A
79	A
80	A
70	A
68	A
90	A
92	A
80	A
70	A
63	A
76	A
48	A
90	A
92	B
85	B
83	B
76	B
61	B
99	B
99	B
83	B
88	B
74	B
70	B
65	B
51	B
73	B
71	B
72	B
95	B
82	B
70	B
33	B
37	B
32	B
41	B
44	B
49	B
47	B
50	B
59	B
55	B
53	B
56	B
52	B
64	B
60	B
66	B

Extreme Values					
FACTOR				Case Number	Value
TALL	A	Highest	1	6	93
			2	16	92
			3	7	91
			4	15	90
			5	22	90
	Lowest		1	5	35
			2	21	48
			3	9	60
			4	19	63
			5	10	63
	B	Highest	1	28	99
			2	38	95
			3	23	92
			4	30	88
			5	24	85
	Lowest		1	43	32
			2	41	33
			3	42	37
			4	44	41
			5	45	44

كما ينفذ مخطط Boxplots لكل مجموعة بصورة منفصلة كما يلي :



نلاحظ وجود قيمة شاذة Outlier في الفئة A وهي الحالة رقم 5 وقيمتها 35 حيث أنها تبعد عن الربع الأدنى $Q1=66.75$ بمسافة تزيد عن 1.5 وتقل عن $Q3$ من طول الصندوق 18.75 ويرمز لها بالرمز في مخطط Boxplots وبجانب الرمز يعرض رقم الحالة 5 وهو الحالة الافتراضية في حالة عدم وجود Label للحالة (متغير يتم إدخاله في قائمة Label Cases by في صندوق حوار Explore) .

(6 - 2) اختبار تجانس التباين *Test of Homogeneity of Variances*

أن أحد الفرضيات التي ينبغي توفرها عند إجراء تحليل التباين ANOVA هو تجانس أو تساوي تباين المعالجات (المعاملات) ففي حالة عدم تحقق هذا الشرط فإن ذلك سوف يؤدي للتوصل الى قرارات خاطئة عند اختبار الفرضية الخاصة بتساوي تأثير المعاملات (أو تساوي أوساط المعالجات) .

يوفر برنامج SPSS اختباراً لتجانس التباين هو Levene Test حيث تختبر فرضية العدم القائلة بتجانس تباين المعالجات ضد الفرضية البديلة القائلة بعدم تجانس التباين . وغالباً عندما تكون التباينات غير متساوية فإن ذلك يؤدي الى عدم تحقق شرط التوزيع الطبيعي داخل المعالجات وهذا الشرط ضروري لاختبار الفرضيات حيث نفترض أن الخطأ العشوائي له توزيع طبيعي $\varepsilon \sim N(0, \sigma^2)$. وفي هذه الحالة يتوجب إجراء تحويلات على البيانات Transformation لضمان تجانس التباين وأن القيام بهذه التحويلات يجعل التوزيع يقترب من الطبيعي داخل كل معالجة أيضاً .

لقد أفرح Box & Cox أداة تشخيصية للوصول الى قرار فيما اذا كان سينجم عن التحويلات تجانس التباينات . تعتمد الأداة على تحديد وجود علاقة بين أوساط المعالجات وانحرافات المعيارية (يجب أن تكون الأوساط مستقلة عن الانحراف المعياري لضمان تجانس التباين) . حيث يتم حساب الميل من انحدار لوغاريتم الانحرافات عن المتوسط على لوغاريتم أوساط المعالجات لتكوين Power Transformation للبيانات حسب الصيغة التالية y^{1-b} حيث أن y تمثل البيانات الأصلية وأن $1-b$ تمثل الأس Power وهذه بعض التحويلات شائعة الاستعمال :

<u>Transformation</u>	<u>Power</u>	<u>Slop b</u>
Square	2	-1
None	1	0
Square Root	1/2	1/2
Logarithm	0	1
Reciprocal of Square Root	-1/2	3/2
Reciprocal	-1	2

يلاحظ أن الأس يتراوح من -1 الى 2 فاذا كانت قيمة الميل تساوي 1 فهذا يعني أن أوساط المعالجات مستقلة عن انحرافات المعيارية وبالتالي لا تحتاج البيانات الى تحويل نظراً لتجانس التباين أي أن الأس يساوي 0 واذا كانت معلمة الميل مساوية الى 0.671 فإن الأس $1-b$ سيكون 0.329 فأنتنا سنختار تحويله الجذر التربيعي للبيانات الأصلية لأن 0.329 قريبة من 1/2 أما في حالة أن الأس يساوي 1 فأنتنا نختار تحويله اللوغاريتم علماً أنه لا يوجد فرق بين استعمال اللوغاريتم الطبيعي أو اللوغاريتم للأساس 10 .

الخيار Spread vs. Level with Leven Test

يتيح هذا الخيار (الذي يظهر في صندوق حوار Plots كما ورد سابقاً) إمكانية اختبار تجانس تباين المعاملات (المعالجات) باستخدام Levene Statistics إضافة الى اختبار تجانس التباين بواسطة المخطط Spread –Versus Level Plot (يمثل نفس أسلوب Box & Cox مع تحوير بسيط) حيث يتم تمثيل لوغاريتم الوسيط للمعالجات (المستوى Level) على المحور الأفقي لهذا المخطط وتمثيل لوغاريتم المدى الربيعي Inter Quartile Range على المحور العمودي (التباعد Spread) ففي حالة عدم وجود علاقة بين المستوى والتباعد فإن نقاط هذا المخطط تتجمع في خط أفقي (Slop=0) الذي يشير الى تجانس تباين المعالجات وعكس ذلك فإن تجمع النقاط حول خط اتجاه عام Trend يعني وجود علاقة وبالتالي عدم تجانس تباين المعالجات مما يستلزم إجراء تحويل على البيانات حسب جدول التحويلات المذكور آنفاً . علماً أنه بالإمكان جعل البرنامج يقوم برسم المخطط بعد التحويل لغرض التحقق من أن عملية التحويل قد أثمرت عن تجانس التباين (أي أن الميل سيكون قريب من الصفر) .

مثال 3: الجدول التالي يمثل نتائج تجربة عاملية استخدمت فيها أربعة معالجات A,B,C,D مع 12 تكرار لكل معالجة :

Treatments Obs.	A	B	C	D
1	895	1520	43300	11000
2	540	1610	32800	8600
3	1020	1900	28800	8260
4	470	1350	34600	9830
5	428	980	27800	7600
6	620	1710	32800	9650
7	760	1930	28100	8900

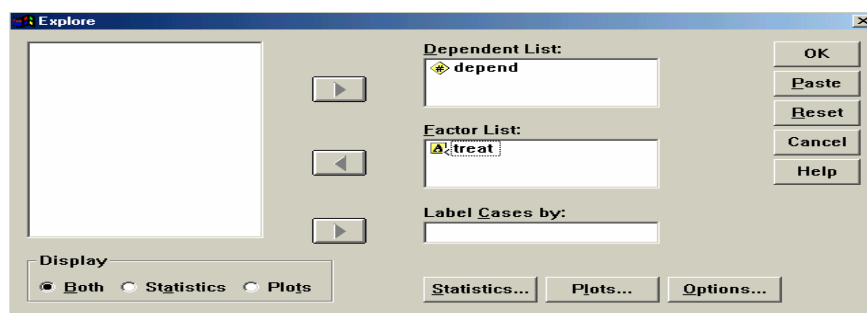
8	537	1960	18900	6060
9	845	1840	31400	10200
10	1050	2410	39500	15500
11	387	1520	29000	9250
12	497	1685	22300	7900
Mean	670.75	1701.25	30775	9395.83
Std. Deviation	233.92	356.54	6688.68	2326.04

المصدر: د.محمود المشهداني وكمال علوان المشهداني ، تصميم وتحليل التجارب ، جامعة بغداد ، ص 49 .
المطلوب اختبار تجانس تباين المعالجات مع تحديد التحويل المناسب للبيانات في حالة عدم التجانس .
يتم إدخال البيانات الى نافذة Data Editor لبرنامج SPSS كما هو في (الصفحة التالية) حيث أن المتغير depend يمثل المشاهدات (المتغير المعتمد) وأن treat هو متغير تجزئة للمتغير المعتمد ويمثل المعالجات A,B,C,D . لاختبار تجانس التباينات نتبع الخطوات التالية :

من شريط القوائم نختار

Explore → Descriptive Statistics → Analyze حيث يظهر صندوق حوار Explore

الذي نقوم بترتيبه على الشكل التالي :

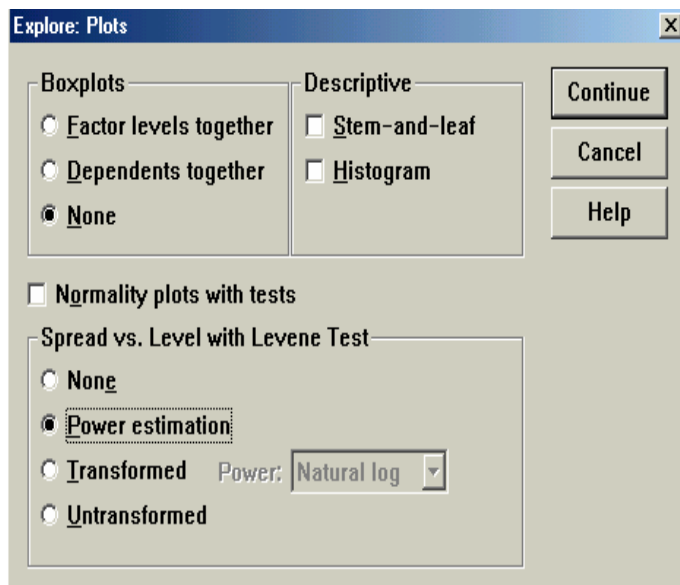


أقر الزر Plots في صندوق حوار Explore وفي خانة spread vs. Level with levene test

(الذي يكون فعالاً نظراً لوجود متغير تجزئة Treat) ثم نتبع الخطوتين الرئيسيتين التاليتين .

الخطوة الأولى: أختار Power Estimation حيث يظهر صندوق حوار Plots بعد ترتيبه كما يلي :

أن الهدف من هذه الخطوة هو التالي :



ترتيب البيانات في Data Editor للمثال 3

<u>Treat</u>	<u>Depend</u>
A	895
A	540
A	1020
A	470
A	428
A	620
A	760
A	537
A	845
A	1050
A	387
A	497
B	1520
B	1610
B	1900
B	1350
B	980
B	1710
B	1930
B	1960
B	1840
B	2410
B	1520
B	1685
C	43300
C	32800
C	28800
C	34600
C	27800
C	32800
C	28100
C	18900
C	31400
C	39500
C	29000
C	22300
D	4000
D	8600
D	8260
D	9830
D	7600
D	9650
D	8900
D	6060
D	10200
D	15500
D	9250
D	7900

1. الاختبار الإحصائي لتجانس التباين بواسطة إحصائية Levene .

2. اختبار العلاقة بين المستوى والانتشار من خلال الرسم البياني مع

تحديد نوع التحويل الملائم للبيانات Power Transformation في

حالة وجود علاقة (أي التباين غير متجانس) .

أقر زر Continue ثم زر OK فنحصل على النتائج التالية :

1. الاختبار الأحصائي لتجانس التباين بواسطة إحصائية Levene

المخرج أدناه يبين نتيجة اختبار فرضية العدم (تجانس التباين)

ضد الفرضية البديلة (عدم تجانس التباين) باستخدام إحصائية Levene

حيث أن قيمة هذه الإحصائية إما تكون مبنية على المتوسط Mean أو

الوسيط Median أو الوسط المشذب فبالنسبة للإحصائية المبنية على

المتوسط فقد بلغت 10.783 وبمعنوية عالية $(p\text{-value} \approx 0.000 < 0.05)$.

وعليه ترفض فرضية العدم بمستوى دلالة 5% وتؤخذ الفرضية البديلة

القائلة بعدم تجانس التباينات وهي نفس النتيجة التي نتوصل إليها عند

اعتماد قيمة الإحصائية المبنية على الوسيط والوسط المشذب .

Test of Homogeneity of Variance

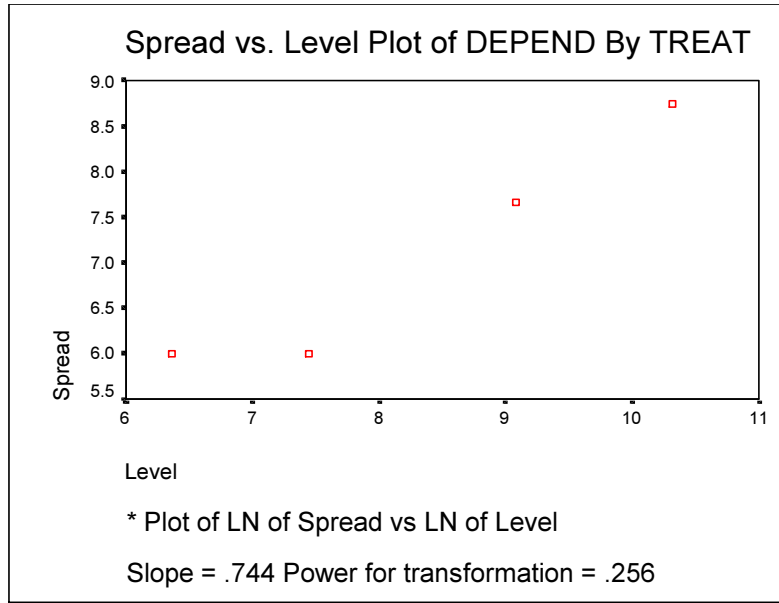
	Levene Statistic	df1	df2	Sig.
Mean	10.783	3	44	.000
Median	10.621	3	44	.000
Median and trimmed mean	10.621	3	15.853	.000
	10.779	3	44	.000

2. اختبار تجانس التباين من خلال مخطط Spread vs. Level Plot

يمكن من خلال هذا المخطط تشخيص مشكلة عدم تجانس التباين

بالإضافة الى تقدير الأس المناسب لتحويل البيانات في حال وجود المشكلة

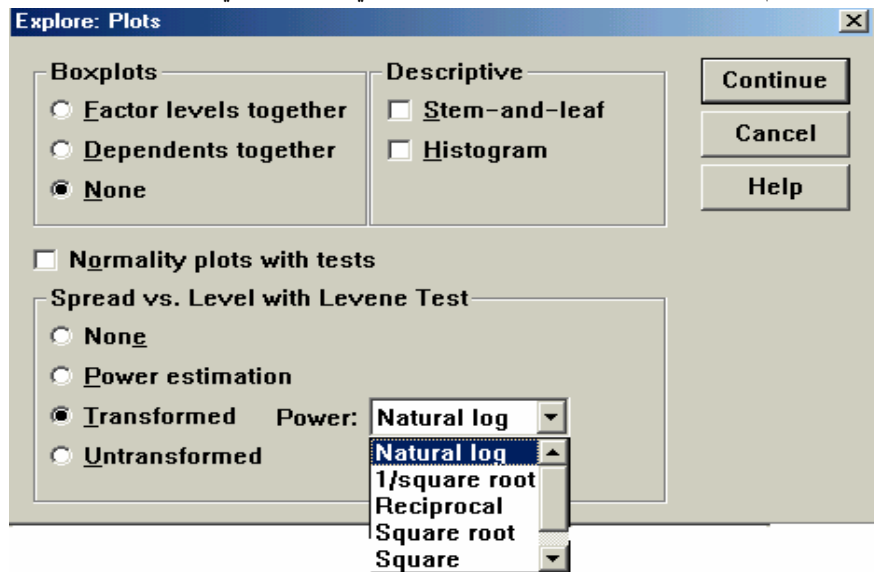
(الخيار Power estimation) كما في المخطط التالي :



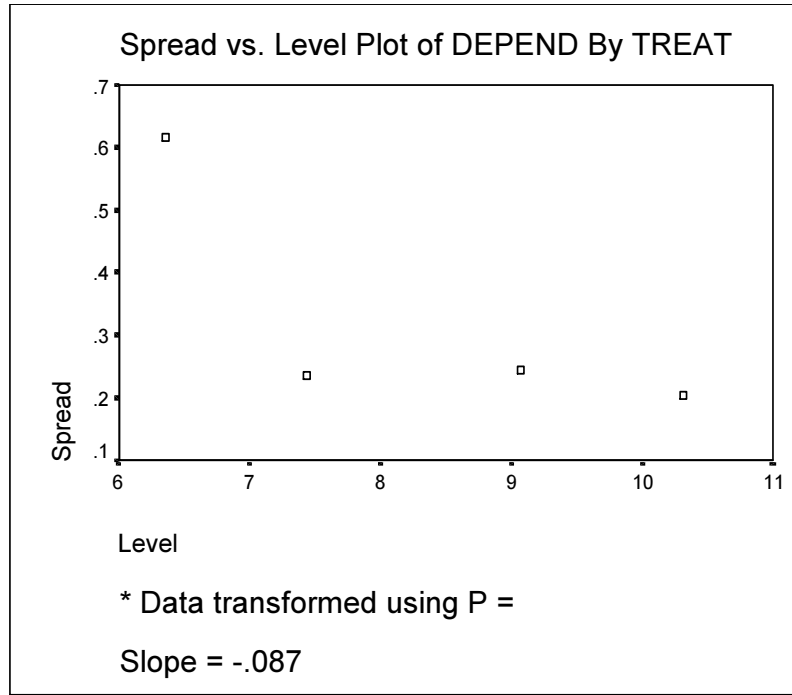
حيث يتضح من خلال الاتجاه العام لشكل الانتشار وجود علاقة بين لوغاريتم الوسيط للمعالجات A,B,C,D ولوغاريتم المدى الربيعي لها وأن الميل $b=0.744$ وأن الأس اللازم لتحويل البيانات هو $1-b = 0.256$ وبالرجوع الى جدول التحويلات الشائعة نجد أن قيمة الأس 0.256 تقع بين الأس 0 (تحويل لوغاريتمي) وبين الأس $1/2$ (تحويلة الجذر التربيعي) وفي هذه الحالة يمكن تجربة كلا النوعين في التحويل واعتماد التحويل الذي يعطي نتائج افضل (ميل قريب من الصفر أو الأس يساوي واحد) ان تنفيذ ذلك يتم مباشرة بواسطة الخيار Transformed في صندوق حوار Plots (الخطوة التالية) .

الخطوة الثانية: تحويل البيانات بواسطة الخيار Transformed

نستنتج من الخطوة الأولى أن تباين المعالجات غير متجانس وعليه نقوم بتحويل البيانات لجعل التباين متجانسا ويتم ذلك بالعودة الى صندوق حوار Explore:Plot ثم نقر Transformed فتصبح دالة التحويل Power فعالة ثم اختر التحويلة Natural Log وكما في الشكل التالي :

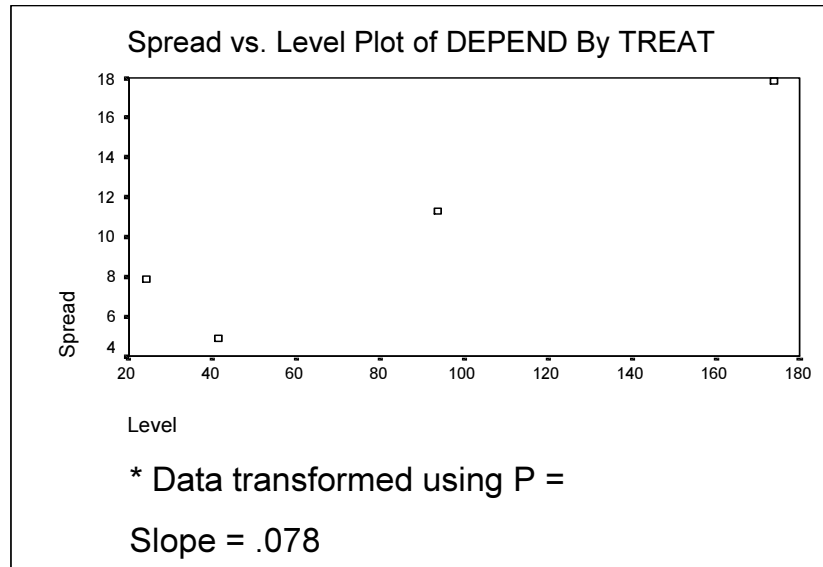


عند نقر زر Continue ثم زر OK يظهر مخطط البيانات المحولة (اللوغاريتمات كما يلي) :



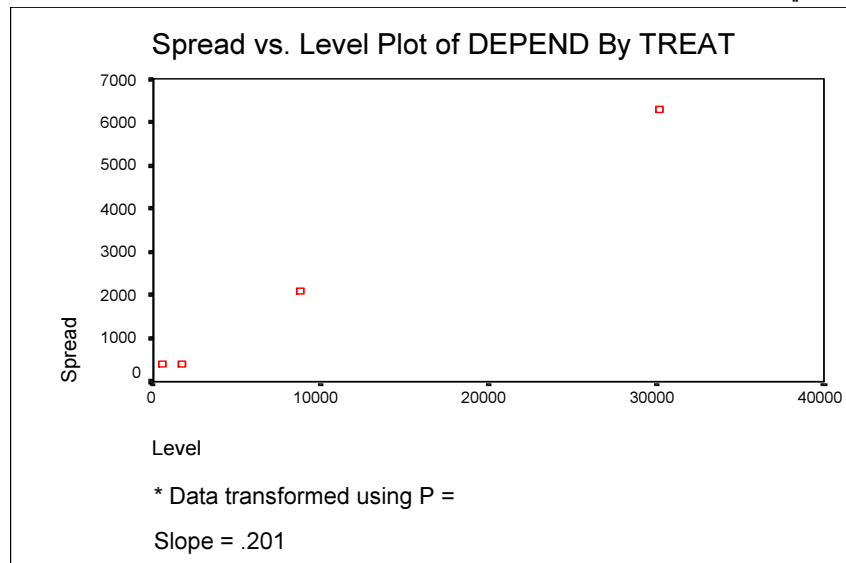
حيث يتضح عدم وجود اتجاه عام للعلاقة بين الوسيط والانتشار للبيانات المحولة وأما تتركز النقاط حول خط أفقي وهذا يشير الى انتفاء مشكلة عدم تجانس التباين للبيانات المحولة كما أن قيمة الميل $b = -0.087$ وعليه فإن الأس سيكون $1 - b = 1.087$ وهو قريب من الواحد أي نستطيع القول أن البيانات المحولة لاتعاني من مشكلة عدم تجانس التباين .

أما في حالة اختيار تحويلة الجذر التربيعي نحصل على المخطط التالي :



حيث أن الميل يساوي 0.078 وعليه فإن الأس سيكون 0.922 أي أنه قريب من الواحد وهذا يعني تجانس التباين للبيانات المحولة بالجذر التربيعي . وقد كانت النتائج متقاربة لكلا النوعين من التحويلات وعليه يمكن اعتماد أي منهما .

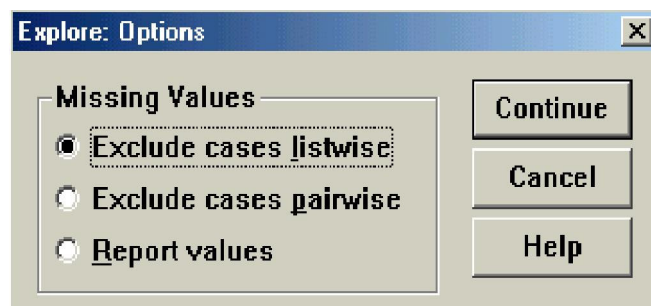
أما بالنسبة للخيار Untransformed في صندوق حوار Plots فإنه ينفذ المخطط بالاعتماد على القيم الأصلية (بدون تحويل البيانات) وذلك برسم العلاقة بين الوسيط والمدى الربيعي (IQR) بدون لوغاريتمات كما في الشكل التالي :



ويلاحظ وجود اتجاه عام للعلاقة بين الوسيط والمدى الربيعي علماً أن هذا المخطط لا يستعمل في تحديد نوع التحويل لأنه لا يمثل العلاقة بين لوغاريتم الوسيط ولوغاريتم المدى الربيعي للمعالجات.

(3-6) التعامل مع القيم المفقودة

يمكن تحديد كيفية استبعاد الحالات التي تحوي قيمةً مفقودة بنقر الزر Options في صندوق حوار Explore فيظهر صندوق حوار Options التالي :



حيث أن :

Exclude Cases Listwise: يستعمل لاستبعاد الحالة بأكملها (لكافة المتغيرات) إذا كانت تحتوي على قيمة مفقودة لأي من المتغير المعتمد **Dependent** أو متغير التجزئة **Factor** وهذا الخيار هو الحالة الافتراضية .

Exclude Cases Pairwise : يتم استبعاد الحالات التي لها قيم مفقودة لأي من أو كلا المتغيرات التي تستعمل في حساب أحصائية معينة .

Report Values : تعامل القيم المفقودة لمتغير التجزئة كفئة مستقلة ويتم عرض النتائج لهذه الفئة الإضافية .

لتوضيح ما سبق نفترض لدينا البيانات التالية التي أدخلت في شاشة Data Editor :

dep1	dep2	fac
1	10	1
2	11	1
.	.	1
4	13	2
5	14	.
6	.	2
7	8	2

□ في البداية ننفذ الأمر Analyze → Descriptive Statistics → Explore

□ في صندوق حوار Explore نتبع مايلي :

- انقر المتغيرين dep1 و dep2 وأدخلهما الى خانة Dependent List .
- انقر المتغير fac وأدخله الى خانة Factor List .
- انقر زر Options وأختر Exclude Cases list Wise .

□ انقر زر OK فتظهر النتيجة التالية :

Case Processing Summary

		Cases					
		Valid		Missing		Total	
		N	Percent	N	Percent	N	Percent
DEP1	1	2	66.7%	1	33.3%	3	100.0%
	2	2	66.7%	1	33.3%	3	100.0%
DEP2	1	2	66.7%	1	33.3%	3	100.0%
	2	2	66.7%	1	33.3%	3	100.0%

Descriptives

FAC			Statistic	Std. Error
DEP1	1	Mean	1.50	.50
	2	Mean	5.50	1.50
DEP2	1	Mean	10.50	.50
	2	Mean	10.50	2.50

لاحظ كيفية حساب المتوسطات (على سبيل المثال) في حالة Exclude Cases Listwise حيث تم استبعاد أي حالة تحتوي على قيمة مفقودة لأي من المتغيرات في قائمة Dependent List أو Factor List. فمثلاً أحتسب المتوسط للمتغير dep1 للقيمة 1 لمتغير التجزئة fac بالاعتماد على الحالتين 1 و 2 وعلى الحالات 4 و 7 للقيمة 2 لمتغير التجزئة أما بالنسبة لمتوسط dep2 للقيمة 1 من متغير التجزئة فقد أحتسب بالاعتماد على الحالتين 1 و 2 ... وهكذا .

أما في حالة الخيار Exclude Cases Pairwise فأن المتوسطات تحتسب كما يلي :

Case Processing Summary

		Cases					
		Valid		Missing		Total	
		N	Percent	N	Percent	N	Percent
DEP1	1	2	66.7%	1	33.3%	3	100.0%
	2	3	100.0%	0	.0%	3	100.0%
DEP2	1	2	66.7%	1	33.3%	3	100.0%
	2	2	66.7%	1	33.3%	3	100.0%

Descriptives

FAC			Statistic	Std. Error
DEP1	1	Mean	1.50	.50
	2	Mean	5.67	.88
DEP2	1	Mean	10.50	.50
	2	Mean	10.50	2.50

لاحظ أن متوسط المتغير dep1 لقيمة متغير التجزئة fac قد أحتسب من الحالات 1 و 2 أما بالنسبة لمتوسط dep1 للقيمة 2 لمتغير التجزئة فقد أحتسب من الحالات 4 و 6 و 7 ونفس الشيء للمتغير dep2 . الآن ماذا يحصل لو حذفنا المتغير fac من قائمة Factor List في صندوق حوار Explore مع الإبقاء على الخيار Exclude Cases Pairwise ؟ في هذه الحالة يتم استبعاد الحالات الحاوية على قيمة مفقودة لكل متغير وبصورة مستقلة حيث نحصل على القيم التالية للمتوسطات الحسابية .

Descriptives

		Statistic	Std. Error
DEP1	Mean	4.17	.95
DEP2	Mean	11.20	1.07

فالقيمة 4.17 هي متوسط الحالات 1 و 2 و 4 و 5 و 6 و 7 للمتغير dep1 والقيمة 11.02 هي متوسط الحالات 1 و 2 و 4 و 5 و 7 للمتغير dep2 .

سنحاول الآن تجربة الحالة الأخيرة للقيم المفقودة Report Values . أرجع المتغير fac الى خانة Factor List في صندوق حوار Explore مع تأشير الخيار Report Values في صندوق حوار Options . عند نقر زر OK في صندوق حوار Explore نحصل على النتائج التالية :

Case Processing Summary

FAC	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
DEP1 . (Missing)	1	100.0%	0	.0%	1	100.0%
1	2	66.7%	1	33.3%	3	100.0%
2	2	66.7%	1	33.3%	3	100.0%
DEP2 . (Missing)	1	100.0%	0	.0%	1	100.0%
1	2	66.7%	1	33.3%	3	100.0%
2	2	66.7%	1	33.3%	3	100.0%

لاحظ أن متغير التجزئة Fac له أصلاً فئتين هما 1 و 2 وفي حالة اختيار Values Report فقد أضيفت فئة ثالثة بأسم Missing للقيم المفقودة. يتم احتساب المتوسطات (على سبيل المثال) للخيار Report Values كما يلي :

Descriptives

FAC			Statistic	Std. Error
DEP 1	1	Mean	1.50	.50
	2	Mean	5.50	1.50
DEP 2	1	Mean	10.50	.50
	2	Mean	10.50	2.50

يلاحظ أن أستبعد القيم المفقودة و حساب المتوسط قد تم بنفس طريقة Exclude Listwise حيث حصلنا على نفس النتائج تماماً .

الفصل السابع

جداول التقاطع

Crosstabs

يستعمل الأمر Crosstabs في عمل جداول الاقتران 2×2 Tables (تتكون من صفين وعمودين) والجداول المتعددة Multiway Tables (تتكون من أكثر من صفين أو عمودين) مع احتساب مؤشرات الارتباط لهذه الجداول والاختبارات الإحصائية المرافقة. أن لهذا الأمر فائدة كبيرة في البحوث التطبيقية فمثلاً يسهل عملية إعداد جداول الاقتران والجداول المتعددة لأسئلة الاستبيانات الإحصائية .

مثال 1: الجدول التالي يمثل عينة حجمها 22 مشاهدة للمرضى الذين خضعوا لعلاج a والذين لم يخضعوا لعلاج b ويعبر عن ذلك المتغير treat أما المتغير recover فيمثل الشفاء من المرض يرمز للشفاء a1 وعدم الشفاء b2 أما المتغير gender فيمثل جنس المريض m للذكور و f للإناث وقد أدخلت البيانات في شاشة Data Editor كما يلي :

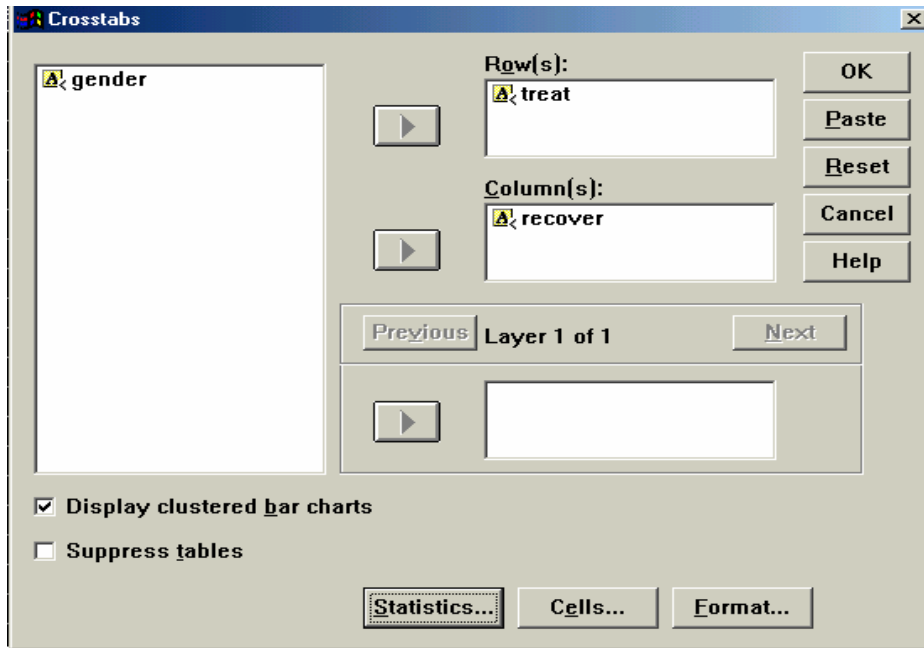
treat	recover	gender
a	a1	m
b	a1	m
b	b1	m
b	b1	m
a	a1	m
a	a1	m
b	b1	m
a	b1	m
b	b1	m
a	a1	m
a	a1	m
b	a1	m
a	a1	m
b	b1	m
a	a1	f
b	b1	f
b	b1	f
b	b1	f
a	a1	f
b	a1	f
a	b1	f
b	b1	f

يطلب مايلي :

1. تكوين جدول اقتران للعلاقة بين المتغيرين treat و recover مع اختبار الاستقلالية بين المتغيرين المذكورين أي اختبار وجود علاقة معنوية بين تعاطي العلاج والشفاء من المرض .
2. تكوين جدول اقتران للعلاقة بين المتغيرين treat و recover مع اختبار الاستقلالية بين المتغيرين المذكورين حسب متغير الجنس gender .

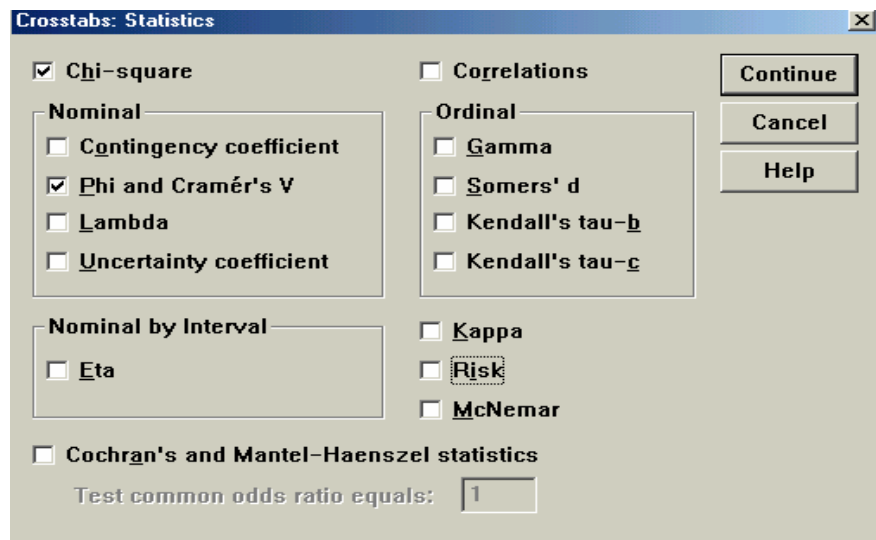
1. لتنفيذ المطلوب الأول نتبع الخطوات التالية :

من شريط القوائم نختار Crosstabs → Descriptive Statistics → Analyze
 فيظهر صندوق حوار Crosstabs الذي نرتبه كما يلي :



حيث أن :

- Row(s)** : يتضمن المتغير الذي نرغب في جعل فئاته صفوفاً في الجدول .
- Column(s)** : يتضمن المتغير الذي نرغب في جعل فئاته أعمدة في الجدول .
- Display Clustered bar Charts** : لعرض الأشرطة البيانية عند التأشير .
- Supress tables** : لاختفاء جداول الاقتران والجداول المتعددة عند التأشير .



عند نقر زر Statistics يظهر صندوق الحوار Statistics الذي نرتبه كما يلي :
 حيث أن :

- Chi-Square** : لحساب إحصائية chi-Square وملحقاتها لاختبار استقلالية الصفوف عن الأعمدة .
- Correlation** : لحساب معاملي الارتباط لـ Spearman و Pearson في أن معاً علماً أنه يجب أن تكون

قيم المتغيرات الداخلة عددياً Numeric فيتم حساب معامل ارتباط Spearman باعتبار أن كلا من الصفوف والأعمدة عبارة عن قيم ترتيبية Ordered Values (الخيار Ordinal في العمود Measure في شاشة Variable view) مثلاً الحالة العلمية (أمي، يقرأ ، يقرأ ويكتب ...) كما يحتسب معامل ارتباط Pearson باعتبار أن كلا من الصفوف والأعمدة متغيرات كمية Quantitative Variables ويعبر عنها بالخيار Scale في عمود Measure في شاشة Variable View ويشار لها أيضاً Interval.

Nominal: تتيح هذه القائمة حساب أربعة معاملات للاقتتران (ارتباط) في حالة كون كلا من الصفوف والأعمدة عوامل غير كمية بالإضافة إلى عدم إمكانية ترتيب البيانات مثلاً متغير الجنس (ذكور، إناث) أو متغير المحافظة (بغداد، موصل، بصرة) .

Ordinal: تتيح هذه القائمة حساب أربعة معاملات للاقتتران في حالة كون كلا الصفوف والأعمدة متغيرات ترتيبية .

Nominal by Interval: لحساب إحصائية Eta للاقتتران بين متغيرين أحدهما المعتمد يقاس ضمن فترة Interval Scale مثل متغير الدخل Income والمتغير الآخر مستقل له عدد محدود من الفئات Categories مثل الجنس gender .

يمكنك التعرف على بقية إحصائيات الصندوق بنقرها بزر الماوس الأيسر لظهور التعليق الخاص بها .

نظراً لكون الصفوف والأعمدة في هذا المثال متغيرات اسمية Nominal (لها الخيار Nominal في عمود Measure في شاشة Variable View) فقد اكتفينا بحساب إحصائية Chi-Square لاختبار الاستقلالية مع مقياس Phi and Gramer's v للاقتتران للمتغيرات الاسمية Nominal Variables.

◀ عند نقر زر Continue في صندوق حوار Statistics ثم ok في صندوق حوار Crosstabs نحصل على المخرجات التالية :

TREAT * RECOVER Crosstabulation

Count		RECOVER		Total
		a1	b1	
TREAT	a	8	2	10
	b	3	9	12
Total		11	11	22

Chi-Square Tests

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

a. Computed only for a 2x2 table

b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 5.00.

يتم اختبار فرضية العدم القائلة باستقلالية الصفوف عن الأعمدة أي استقلالية العلاج عن الشفاء ضد الفرضية البديلة القائلة بوجود علاقة بين العلاج و الشفاء باستخدام إحصائية Chi-Square التي لها درجة حرية $(r-1)(c-1)$ حيث ان r يمثل عدد الصفوف و c عدد الأعمدة و حسب الصيغة التالية :

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

حيث أن O_i يمثل التكرار المشاهد و E_i يمثل التكرار المتوقع وقد كانت قيمة الإحصائية 6.6 وأن قيمة $p\text{-value} = 0.010 < 0.05$ تدعونا الى رفض فرضية العدم بمستوى دلالة 5% أي توجد علاقة بين العلاج والشفاء من المرض .

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	.548	.010
	Cramer's V	.548	.010
N of Valid Cases		22	

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

أما مقياس الأقران $\Phi = 0.548$ $\sqrt{\frac{\chi^2}{n}} = \sqrt{\frac{6.6}{22}} = 0.548$ فإن قيمة $p\text{-value} = 0.010$ فتشير إلى

معنوية الأقران بمستوى دلالة 5 %، حيث تستعمل إحصائية χ^2 لاختبار الاستقلالية بدرجة حرية $(r-1)(c-1)$.

ملاحظة 1 : يحتسب تصحيح Yates للاستمرارية Continuity Correction في حالة أن درجة الحرية

تساوي واحد أي في جداول 2×2 حسب الصيغة التالية $\chi^2 = \sum \frac{(|O_i - E_i| - 1/2)^2}{E_i}$.

ملاحظة 2 : أن مقياس الأقران Φ غالباً ما يستعمل مع جداول الأقران 2×2 والحالة العامة لهذا المقياس هو ما يعرف بـ Cramer Coefficient C الذي يحتسب لجداول متعددة $r \times c$ من العلاقة التالية:

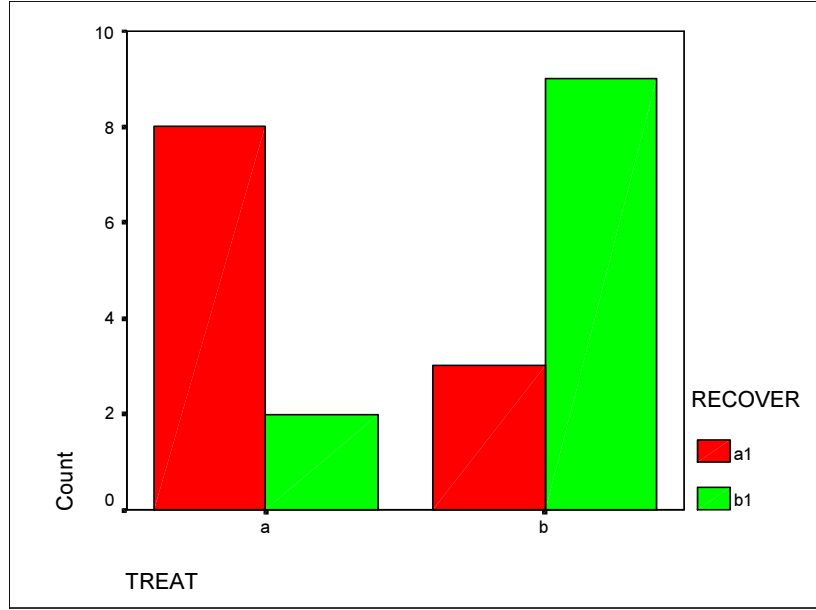
$$C = \sqrt{\frac{\chi^2}{N(L-1)}} = \sqrt{\frac{6.6}{22}} = 0.548$$

حيث أن N يمثل حجم العينة و L العدد الأصغر بين الصفوف والأعمدة في جدول التوافق حيث تستعمل إحصائية χ^2 لاختبار الاستقلالية بدرجة حرية $(r-1)(c-1)$ ونلاحظ أن قيمة هذا المقياس مساوية تماماً لمقياس الأقران Φ .

ملاحظة 3 : هناك مقياس اخر مقارب يعرف بمعامل التوافق Contingency Coefficient (موجود في خانة Nominal في صندوق حوار Crosstbs:Statistics المذكور أنفاً) ويحتسب لأي جدول توافقي $r \times c$ من العلاقة التالية:

$$C = \sqrt{\frac{\chi^2}{\chi^2 + N}} = \sqrt{\frac{6.6}{6.6 + 22}} = 0.480$$

حيث أن N يمثل حجم العينة و تستعمل إحصائية χ^2 لاختبار الاستقلالية بدرجة حرية (c-1)(r-1). حيث أن قيمة P-Value المستخدمة في اختبار الاستقلالية هي نفسها للمعاملات الثلاثة أعلاه وتساوي 0.010. أن ناتج تأشير Check Box المجاور لـ Display Clustered Bar Charts في صندوق حوار Crosstabs يعطينا المخطط التالي :



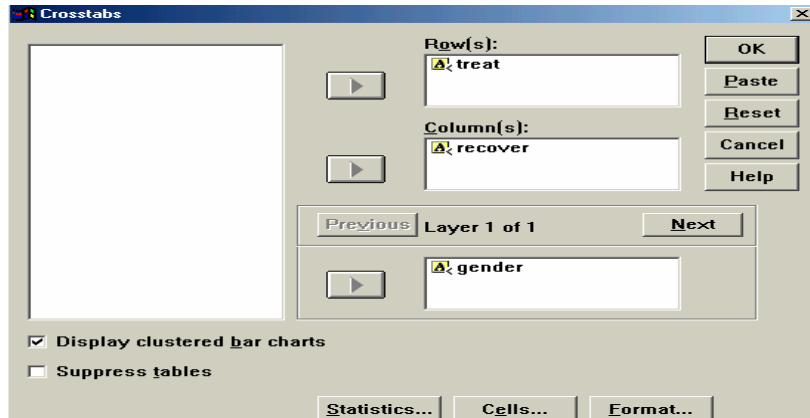
ملاحظة 4 :

- يمكن التحكم بعرض محتويات جدول Crosstab وذلك بنقر الزر Cells في صندوق حوار Crosstabs حيث يظهر صندوق حوار Cell Display الذي يحتوي الخيارات التالية :
 - Counts : الذي يتضمن خيارين :
 - Observed : وهو الخيار الافتراضي حيث تملأ خلايا الجدول بالتكرار المشاهد O_i .
 - Expected : بموجب هذا الخيار تملأ خلايا الجدول بالتكرار المتوقع E_i .
 - Percentages : الذي يتضمن ثلاثة خيارات :
 - Rows : تملأ خلايا الجدول بالنسب المئوية من مجموع الصف .
 - Columns : تملأ خلايا الجدول بالنسب المئوية من مجموع العمود .
 - Total : تملأ خلايا الجدول بالنسب المئوية من المجموع الكلي .
 - Residuals : ويتضمن ثلاثة خيارات :
 - Unstandardised : تملأ خلايا الجدول بالفرق بين التكرار المشاهد والتكرار المتوقع $O_i - E_i$.
 - Standardized : تملأ خلايا الجدول بالفرق بين التكرار المشاهد والتكرار المتوقع مقسوماً على الخطأ المعياري له .
 - Adj.Standardised : نفس الخيار السابق معبراً عنه بوحدة الانحراف المعياري عن المتوسط .

ملاحظة 5 : يمكن ترتيب صفوف الجدول تصاعدياً أو تنازلياً بنقر الزر Format في صندوق حوار Crosstabs

2. لتنفيذ المطلوب الثاني نتبع الخطوات التالية :

من شريط القوائم نختار Crosstabs → Descriptive Statistics → Analyze
فيظهر صندوق حوار Crosstabs الذي نرتبه كما يلي :



نلاحظ أنه تم إدخال المتغير gender ضمن قائمة Layer (طبقة) ويعرف بمتغير السيطرة Control Variable وهو متغير فئوي categorical Variable (له عدد محدد من القيم في هذا المثال للمتغير gender قيمتين m و f) حيث يتم تكوين الجدول Crosstabs للعلاقة بين المتغيرين treat و recover واحساب المقاييس لكل قيمة من قيم متغير السيطرة أي أنه يتم تكوين جدولين أحدهما للذكور m والآخر للإناث f .

عند نقر OK في صندوق حوار Crosstabs نحصل على النتائج التالية :

TREAT * RECOVER * GENDER Crosstabulation

Count			RECOVER		Total
GENDER			a1	b1	
f	TREAT	a	2	1	3
		b	1	4	5
	Total		3	5	8
m	TREAT	a	6	1	7
		b	2	5	7
	Total		8	6	14

Chi-Square Tests

GENDER		Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
f	Pearson Chi-Square	1.742 ^b	1	.187		
	Continuity Correction ^a	.320	1	.572		
	Likelihood Ratio	1.762	1	.184		
	Fisher's Exact Test				.464	.286
	N of Valid Cases	8				
m	Pearson Chi-Square	4.667 ^c	1	.031		
	Continuity Correction ^a	2.625	1	.105		
	Likelihood Ratio	5.004	1	.025		
	Fisher's Exact Test				.103	.051
	N of Valid Cases	14				

a. Computed only for a 2x2 table

b. 4 cells (100.0%) have expected count less than 5. The minimum expected count is 1.13.

c. 4 cells (100.0%) have expected count less than 5. The minimum expected count is 3.00.

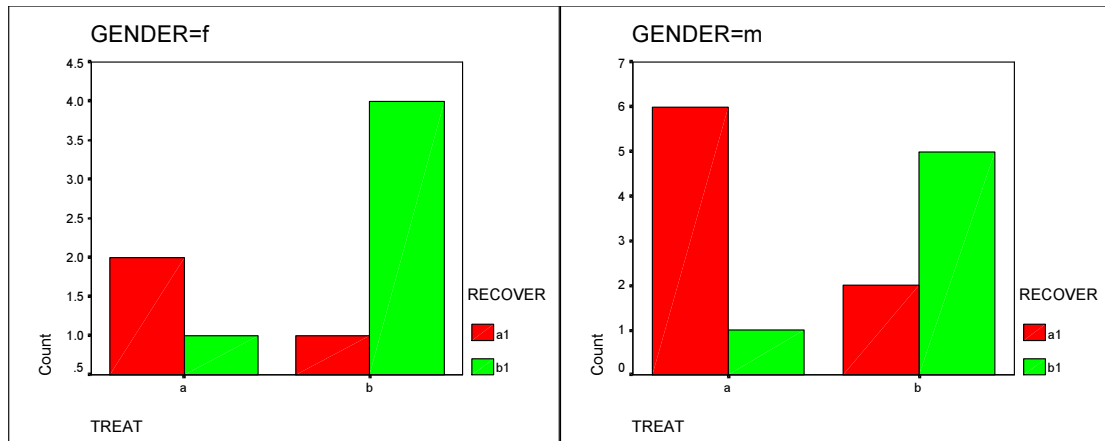
نلاحظ من خلال قيمة p-value لإحصائية Pearson Chi-Square عدم وجود علاقة بين تعاطي الدواء والشفاء من المرض بالنسبة للمرضى الإناث في حين تؤكد الإحصائية على وجود علاقة بالنسبة للمرضى الذكور ولمستوى دلالة 5% .

أن إحدى الفرضيات لاختبار Chi-Square في الجداول من نوع 2×2 تنص على أن التكرار المتوقع E_i يجب أن لا يقل عن 5 بأي حال لأي من خلايا الجدول وألا فأن الاختبار سوف يقود الى نتائج مظللة وفي هذا المثال نلاحظ أن التكرار المتوقع يقل عن 5 لكافة خلايا الجدول 100% ولكل من الذكور والإناث مما يستوجب توخي الحذر في اعتماد نتائج الاختبار .

Symmetric Measures

GENDER			Value	Approx. Sig.
f	Nominal by	Phi	.467	.187
	Nominal	Cramer's V	.467	.187
	N of Valid Cases		8	
m	Nominal by	Phi	.577	.031
	Nominal	Cramer's V	.577	.031
	N of Valid Cases		14	

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



الفصل الثامن

مقارنة المتوسطات

Compare Means

(8-1) تحليل المتوسطات Means

يفيد هذا التحليل في حساب المتوسطات للمجاميع الجزئية subgroups لمتغير معين بالإضافة الى مؤشرات أخرى مثل (المجموع، عدد الحالات للمجموعة الجزئية، الوسيط، المنوال، الانحراف المعياري... الخ) ويمكن أيضاً إجراء تحليل التباين لمعيار واحد واختبار التأثير الخطي للمعاملات Test of Linearity واختبار الاقتران eta .

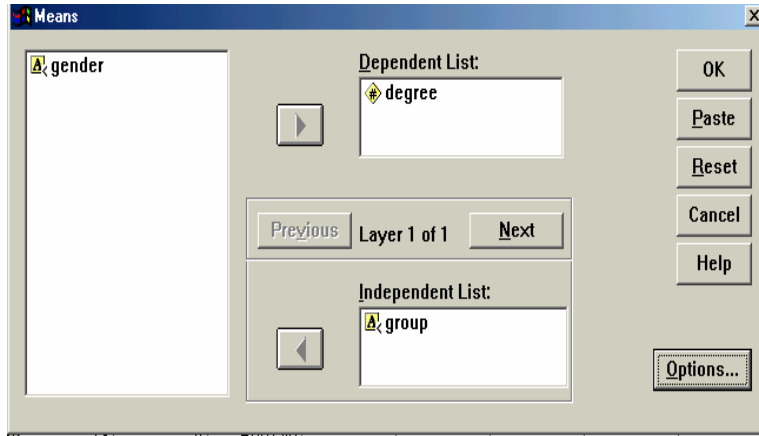
مثال 1

الجدول التالي يبين توزيع درجات 16 طالب ينتمون الى 3 مجاميع مختلفة A، B، C وحسب الجنس Gender والتي تم إدخالها في شاشة Data Editor على الشكل التالي :

<u>degree</u>	<u>group</u>	<u>gender</u>
70	A	Female
90	B	Male
88	A	Male
86	B	Male
68	C	Male
64	C	Male
76	B	Male
83	A	Female
79	B	Female
55	C	Female
97	B	Male
100	A	Male
64	C	Female
59	C	Female
90	A	Male
73	A	Female

لحساب متوسط الدرجات حسب المجاميع الفرعية A، B، C وعدد الحالات والانحراف المعياري لكل مجموعة نتبع الخطوات التالية :-

من شريط القوائم اختر Means → Compare means → Analyze فيظهر صندوق حوار Means الذي نقوم بترتيبه على الشكل التالي :-

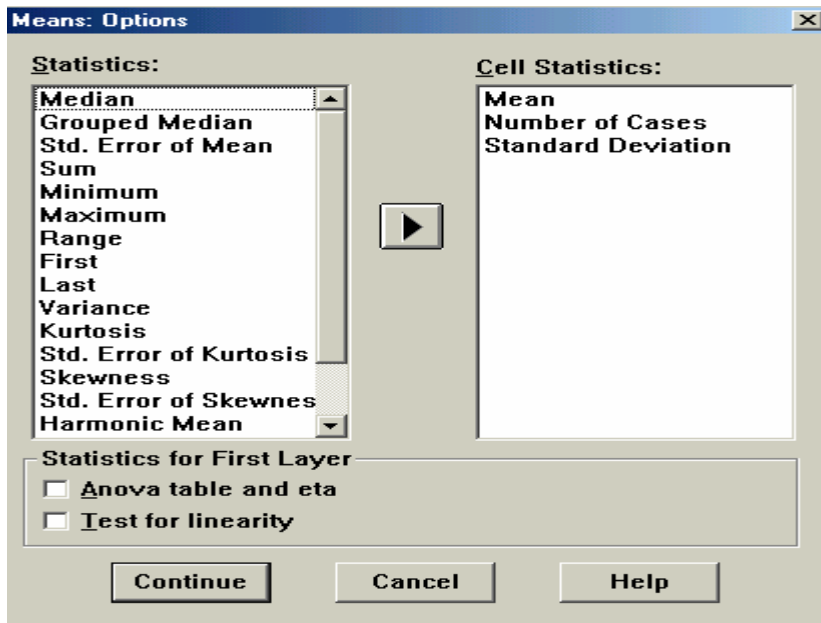


حيث أن :-

Dependent List : المتغير المعتمد أو ما يعرف بمتغير الاستجابة Response Variable نظراً لتأثره بالمتغير المستقل Independent Variable . أن المتغير المعتمد في هذا المثال هو درجات الطلبة Degree .

Independent List : وهو المتغير المستقل أو المؤثر ويمثل المعاملات Treatments في حالة تصميم التجارب أو تحليل التباين ANOVA . أن المتغير المستقل في هذا المثال هو المجاميع الجزئية group الذي يعمل كمتغير تجزئة لدرجات الطلبة .

أقر زر Options في صندوق حوار Means ثم أختار Standard ، No. of Cases ، Means Deviation في صندوق حوار Means:Options وهي المؤشرات المطلوب حسابها حسب المجاميع الجزئية حيث يظهر هذا الصندوق بعد الترتيب كما يلي:-



أقر زر Continue ثم OK فتظهر النتائج التالية :-

Report

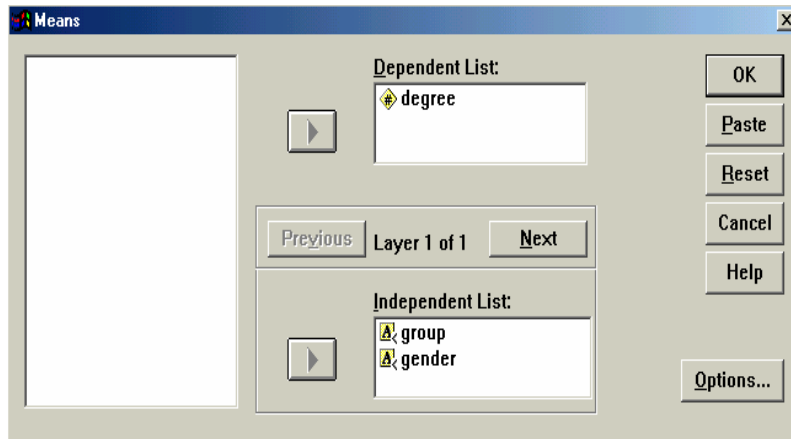
DEGREE

GROUP	Mean	N	Std. Deviation
A	84.00	6	11.19
B	85.60	5	8.44
C	62.00	5	5.05
Total	77.63	16	13.65

نلاحظ أنه تم احتساب المؤشرات المطلوبة للمتغير Degree حسب كل مجموعة جزئية.

مثال 2 :

لنفس بيانات المثال 1 يطلب حساب المتوسط وعدد الحالات والانحراف المعياري لمتغير الدرجات حسب المجاميع group وحسب الجنس gender بصورة مستقلة . لتنفيذ ذلك نتبع نفس الخطوات الواردة في المثال 1 ثم نقوم بترتيب صندوق حوار Means على الشكل التالي :



تظهر نتائج هذا الترتيب كما يلي

DEGREE * GROUP

DEGREE

GROUP	Mean	N	Std. Deviation
A	84.00	6	11.19
B	85.60	5	8.44
C	62.00	5	5.05
Total	77.63	16	13.65

DEGREE * GENDER

DEGREE

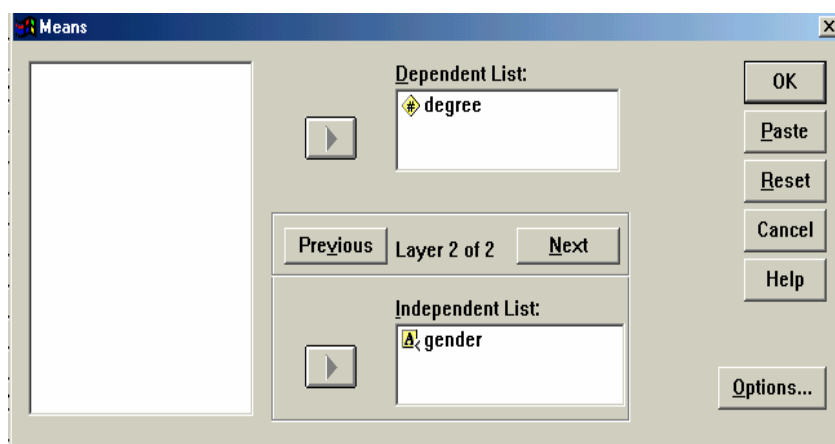
GENDER	Mean	N	Std. Deviation
Female	69.00	7	10.28
Male	84.33	9	12.43
Total	77.63	16	13.65

مثال 3

لنفس المثال 1 يطلب حساب المتوسط وعدد الحالات والانحراف المعياري للمتغير degree حسب الجنس (ذكور، إناث) في داخل كل من المجاميع الفرعية A ، B ، C . لتنفيذ ذلك نقوم بتكوين طبقتين Layers في صندوق حوار Means الطبقة الأولى تضم المتغير group والطبقة الثانية تضم المتغير gender ويتم عمل هذه الطبقات عند ظهور صندوق حوار Means حسب التسلسل التالي :

1. أنقل المتغير group الى خانة Independent List في الطبقة الأولى Layer1 .
2. أنقر زر Next أعلى خانة Independent List للانتقال الى الطبقة الثانية Layer2 .
3. أنقل المتغير gender الى خانة Independent List في الطبقة الثانية Layer2 .

الشكل التالي يبين صندوق حوار Means للطبقة الثانية :



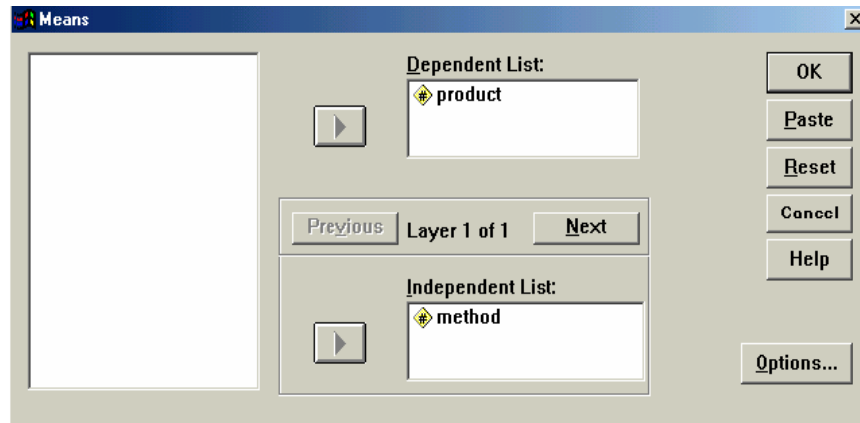
للرجوع الى الطبقة الأولى أنقر زر Previous أعلى خانة Independent List. تظهر نتائج هذا الترتيب عند نقر الزر OK كما يلي :

Report

DEGREE				
GROUP	GENDER	Mean	N	Std. Deviation
A	Female	75.33	3	6.81
	Male	92.67	3	6.43
	Total	84.00	6	11.19
B	Female	79.00	1	.
	Male	87.25	4	8.77
	Total	85.60	5	8.44
C	Female	59.33	3	4.51
	Male	66.00	2	2.83
	Total	62.00	5	5.05
Total	Female	69.00	7	10.28
	Male	84.33	9	12.43
	Total	77.63	16	13.65

مثال 4 : (اختبار التأثير الخطي Test of Linearity)

هذا الاختبار يفيد في اكتشاف وجود اتجاه عام خطي Linear trend معنوي لأوساط المعالجات في تحليل التباين لمعيار واحد One-Way ANOVA. إذا أردنا اختبار التأثير الخطي للمعاملات للبيانات الواردة في المثال 1 من البند (9-1) نقوم بترتيب صندوق حوار Means كما يلي :



ثم نقوم بتأشير الخيارين ANOVA and eta و Test for Linearity في صندوق حوار Means : Options الوارد أنفاً حيث تظهر النتائج كما يلي :

ANOVA Table

	Sum of Squares	df	Mean Square	F	Sig.
PRODUCT * METH Between Groups	402.000	3	134.000	7.053	.012
Linearity	86.400	1	86.400	4.547	.066
Deviation from Linearity	315.600	2	157.800	8.305	.011
Within Groups	152.000	8	19.000		
Total	554.000	11			

Measures of Association

	R	R Squared	Eta	Eta Squared
PRODUCT * METHOD	-.395	.156	.852	.726

أن جدول تحليل التباين يشير الى وجود فروق معنوية بين متوسطات المعالجات ($F=7.053$). يتم الحصول مجموع مربعات الانحرافات عن الخطية كما يلي :

$$SS \text{ Deviation From Linearity} = SS. (\text{Combined}) - SS. \text{Linearity} = 402 - 86.4 = 315.6$$

وتحتسب قيمة F لاختبار Test for Linearity كما يلي :

$$F = MS. \text{ Deviation From Linearity} / \text{within Groups MS.} = 157.8 / 19 = 8.31$$

وأن قيمة P-value المقابلة وتساوي 0.011 تدعونا إلى رفض فرضية العدم القائلة بأن أوساط المعالجات تقع على خط مستقيم في حالة إجراء الاختبار بمستوى دلالة 5% أي أن الأوساط تتحرف عن الخط المستقيم (أنظر البند 2-8 حول تعريف P-Value) .

أن قيمة Eta هي مقياس للاقتزان (الارتباط) Measure of association بين المتغير المعتمد والمتغير المستقل وتقع ضمن المدى 0 إلى 1 حيث أن القيمة 0 تشير إلى انعدام الارتباط والقيمة 1 تشير إلى ارتباط تام ، أما Eta-Square فتمثل نسبة التباينات الكلية في المتغير المعتمد (product) التي يفسرها المتغير المستقل (Method) حيث أن $Eta-Square = SS. \text{ between Groups} / SS. \text{ Total} = 402/554 = 0.726$.

أن قيمة R تمثل معامل الارتباط البسيط Simple Correlation بين المتغير (product) والمتغير (Method) باعتبار أن هذا الأخير يعبر عنه بكودات من 1 إلى 4 وفي حالة وجود أكثر من متغير مستقل فإن R يمثل معامل الارتباط المتعدد Multiple R أما R-Square فهو مقياس لجودة توفيق نموذج الانحدار الخطي Linear Regression .

(2 - 8) اختبار T لعينة واحدة One Sample T-Test

يفيد هذا الاختبار في اكتشاف وجود اختلاف معنوي Significant Difference لمتوسط المجتمع الذي سحبت منه العينة عن قيمة ثابتة Constant . إضافة إلى إمكانية تقدير فترة ثقة لمتوسط المجتمع Confidence Interval ويستعمل هذا الاختبار للعينات الصغيرة ($n < 30$) .

مثال

المتغير التالي Weight يمثل عينة حجمها 10 لأوزان لحم الدجاج بعمر 52 يوماً وقد تم إدخال المتغير في شاشة Data editor .

WEIGHT	1.10	.80	1.20	1.00	.90	1.20
1.10	.90	.80	1.20			

يطلب مايلي

1. اختبار فرضية العدم القائلة بأن المتوسط الحسابي للمجتمع الذي سحبت منه العينة هو 1.25 ضد الفرضية البديلة القائلة بأن متوسط المجتمع لا يساوي 1.25 بمستوى دلالة 5% و 1% .
2. تقدير فترة ثقة 99% للوسط الحسابي للمجتمع .

الحل

يمكن كتابة فرضيتي العدم H_0 و البديلة H_1 كما يلي :-

$$H_0 : \mu = 1.25 \text{ or } \mu - 1.25 = 0$$

$$H_1 : \mu \neq 1.25$$

تكون إحصائية t المستخدمة في الاختبار كما يلي :-

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

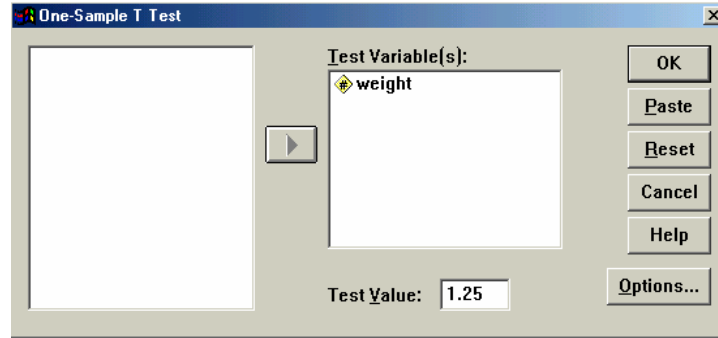
حيث أن $\bar{x} = 1.020$ يمثل الوسط الحسابي للعينة وأن $S = 0.162$ يمثل الانحراف المعياري للعينة و $n = 10$ حجم العينة وأن μ الوسط الحسابي للمجتمع بموجب فرضية العدم ويساوي 1.25 وان هذه الإحصائية تتبع

توزيع t بدرجة حرية $v = n - 1 = 9$ (تحتسب إحصائيتي $\bar{x}$ و S من قبل البرنامج) .
لتنفيذ هذا الاختبار نتبع الخطوات التالية :

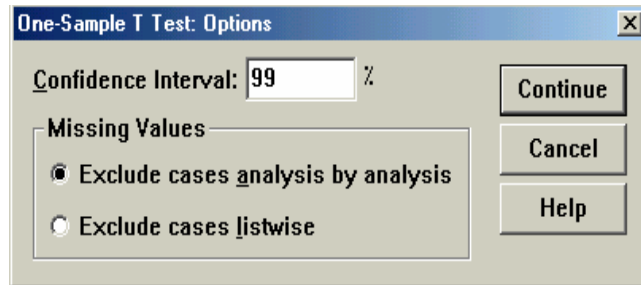
من شريط القوائم نختار

Analyze → Compare Means → One sample T-Test

فيظهر صندوق حوار One sample t-Test الذي نقوم بترتيبه كما يلي :



أنقر زر Options فيظهر صندوق حوار Options الذي نرتبه كما في الشكل التالي :



يمكن بواسطة الخيار Options تحديد فترة الثقة (المطلوب الثاني من المثال) حيث نقوم بكتابة الفترة 99% في حقل Confidence Interval.

أما حقل Missing Values فيستفاد منه في التعامل مع القيم المفقودة (في حالة إجراء T-Test لمجموعة متغيرات أي أن خانة Test Variables في صندوق حوار One sample T-Test تضم أكثر من متغير) وفي هذه الحالة يمكن تأشير أحد الخيارين التاليين :

Exclude Cases Analysis by Analysis : يتم استبعاد الحالات المفقودة Missing Cases للمتغير الذي يجرى عليه اختبار T في حالة احتوائه على قيم مفقودة وفي حالة وجود متغير آخر لا يحتوي قيم مفقودة فأن كافة قيم المتغير تستخدم لاحتساب إحصائية T أي أنه يمكن أن يكون حجم العينة مختلفاً للمتغيرات المضمنة في التحليل .

Exclude Cases List Wise : يتم تضمين الحالات الصحيحة Valid (غير المفقودة) فقط لكافة المتغيرات فإذا كانت الحالة 6 مثلاً مفقودة لأحد المتغيرات فإنه يتم استبعاد الحالة رقم 6 لبقية المتغيرات (وأن كانت غير مفقودة لهذه المتغيرات) وفي هذه الحالة يكون حجم العينة متساوياً للمتغيرات المضمنة في التحليل .
في هذا المثال لا يؤثر اختيار أي من الخيارين على النتائج النهائية للتحليل لعدم وجود قيم مفقودة أولاً ولاحتواء خانة Test Variables على متغير واحد ثانياً .

عند نقر زر OK في صندوق حوار One sample T-Test تظهر النتائج التالية للتحليل :

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
WEIGHT	10	1.020	.162	5.121E-02

One-Sample Test

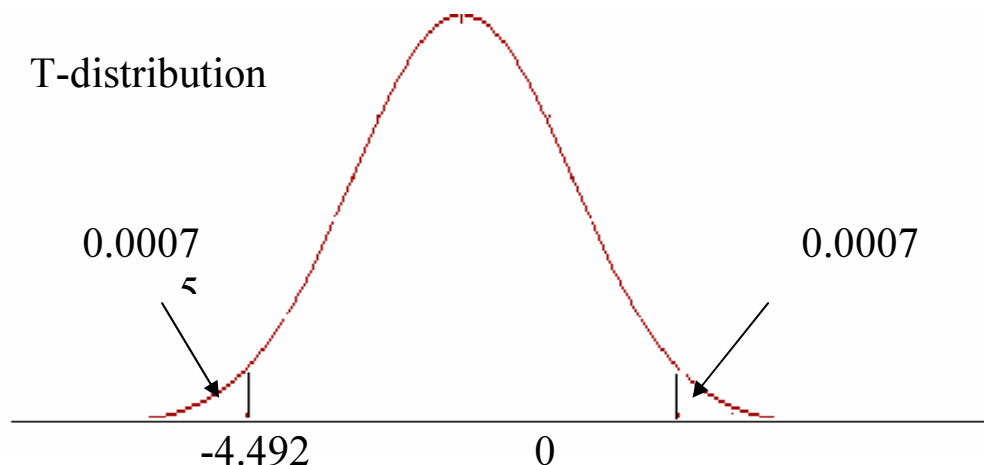
	Test Value = 1.25					
	t	df	Sig. (2-tailed)	Mean Difference	99% Confidence Interval of the Difference	
					Lower	Upper
WEIGHT	-4.492	9	.0015	-.230	-.396	-6.36E-02

حيث تحتسب إحصائية t بموجب الصيغة التالية $t = (1.020 - 1.25) / 0.0512 = -4.492$ وهذه تعرف بـ t المحسوبة ($t(cal.)$) ويجري مقارنتها بقيمة t الجدولية ($t(tab.)$) بدرجة حرية $v = 9$ ولمستوى دلالة α حيث ترفض H_0 في حالة أن $|t(cal.)| \geq t(tab.), v, \alpha/2$ حيث أن قيمة t الجدولية (تستخرج من جداول توزيع t) ولمستوى دلالة $\alpha/2$ (لأن الاختبار من طرفين Two Tailed test) هي كالآتي :

$t_{9,0.025} = 2.262 \quad (\alpha = 5\%)$
 $t_{9,0.005} = 3.250 \quad (\alpha = 1\%)$

وبما أن القيمة المطلقة لـ t المحسوبة (4.492) أكبر من الجدولية ، أذن نرفض فرضية العدم ونأخذ الفرضية البديلة أي أن الوسط الحسابي للمجتمع يختلف معنوياً عن القيمة 1.25 لمستويي دلالة 5% و 1% .

أن الطريقة الثانية المستخدمة في الاختبار تعتمد على القيمة P -value وتتميز على الطريقة الأولى كونها لا تحتاج لاستخدام جداول التوزيعات ويتم احتسابها مباشرة من قبل برنامج SPSS وهي القيمة Sig. في الجدول أعلاه . ويمكن تعريف P -value بأنها أقل قيمة لـ α التي ترفض عندها فرضية العدم . حيث نرفض فرضية العدم إذا كانت P -value أقل من α . المخطط التالي يوضح طريقة احتساب P -value .



$$P\text{-value} = \Pr(t \geq 4.492) + \Pr(t \leq -4.492) = 0.00075 + 0.00075 = 0.0015$$

Pr يمثل الاحتمال Probability

وبما أن $P\text{-value} < 0.05$ وأن $P\text{-value} < 0.01$ أذن نرفض فرضية العدم لمستويي دلالة 5% و 1%
أما بالنسبة لفترة ثقة 99% للفرق بين متوسط العينة ومتوسط المجتمع Mean Difference فيمكن كتابتها كما يلي :

$$\Pr(-0.396 < \bar{x} - \mu - 0.23 < -0.0636) = 99\%$$

أن تقدير فترة ثقة للفروق بين الأوساط قليل الاستعمال في التحليل الإحصائي وبدلاً من ذلك نقوم بتقدير فترة ثقة 99% لمتوسط المجتمع μ كما يلي :

$$\Pr(\bar{x} - t_{9,0.005} * (S/\sqrt{n}) < \mu < \bar{x} + t_{9,0.005} * (S/\sqrt{n})) = 99\%$$

$$\Pr(0.8536 < \mu < 1.1864) = 99\%$$

أي احتمال أن يقع متوسط المجتمع بين القيمتين 0.8536 و 1.1864 يساوي 99% وعليه نرفض فرضية العدم بمستوى دلالة 5% لأن القيمة 1.25 تقع خارج فترة الثقة وهذه تعتبر طريقة ثانية لاختبار الفرضيات .

(3-8) اختبار T للفرق بين متوسطي عينتين *Independent samples T-Test*

يستعمل هذا الاختبار للمقارنة بين متوسطي مجموعتين من الحالات وتستعمل إحصائية T لأجراء الاختبار .

مثال

في تجربة لمقارنة نسبة البروتين في صنفين من الحنطة A و B تم اختبار 12 نباتاً من كل صنف وقدرت نسبة البروتين فيها وكانت النتائج كالتالي :

صنف الحنطة A	صنف الحنطة B
12.5	9.4
9.4	8.4
11.7	11.6
11.3	7.2
9.9	9.7
8.7	7.0
9.6	10.4
11.5	8.2
10.3	6.9
10.6	12.7
9.6	7.3
9.7	9.2

المطلوب اختبار وجود فرق معنوي بين متوسطي نسبة البروتين في الصنفين لمستوي دلالة 5% و 1%. أي اختبار الفرضية التالية :

$$H_0 : \mu_A = \mu_B$$

$$H_1 : \mu_A \neq \mu_B$$

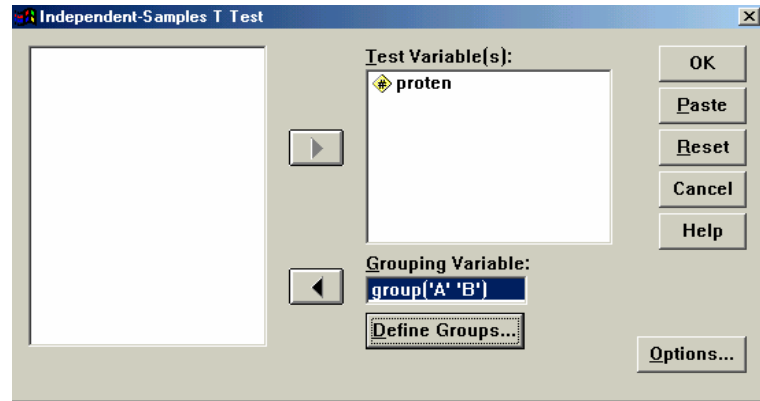
لتنفيذ ذلك نتبع الخطوات التالية :

➤ نقوم بإدخال البيانات في Data Editor كما يظهر في الشكل المجاور حيث أن المتغير Proten يمثل نسبة البروتين والمتغير Group هو متغير تجزئة .

➤ من شريط القوائم اختر

Analyze → Compare Means → Independent Samples T Test

فيظهر صندوق حوار Independent Samples T Test الذي نقوم ترتيبه كما يلي :



فقد قمنا بإدخال المتغير Proten في قائمة Test Variables والمتغير Group (متغير التجزئة) في قائمة Grouping Variable ثم تعريف المجاميع A و B عن طريق نقر الزر Define Groups حيث أن المتغير Group له قيمتين متميزتين A و B (هذا المتغير يسمى dichotomous Variable) أو يمكن أن يأخذ القيمتين 1 و 2 علماً أنه يتوجب التمييز بين الحروف الكبيرة والصغيرة لحالات متغير التجزئة .

كما أن هناك طريقة أخرى في تحديد المجاميع من خلال تحديد نقطة فصل Cut Point في صندوق حوار Define Groups (مثلاً القيمة 10) فكل الحالات التي لها قيمة أقل من 10 لمتغير التجزئة تكون مجموعة واحدة والحالات التي لها قيمة أكبر أو تساوي 10 تكون مجموعة ثانية (في هذه الحالة فأن متغير التجزئة يجب أن يكون عددياً Numeric) .

أما الزر Options في صندوق حوار Independent sample T Test فله نفس الوظيفة في صندوق حوار One Sample T Test .

ملاحظة : يمكن إدخال أكثر من متغير في قائمة Test Variables أي نقوم باختبار الفرق بين متوسطي مجموعتين لعدة متغيرات في أن معاً .

➤ عند نقر زر OK في صندوق حوار Independent sample T Test تظهر النتائج التالية :

Group Statistics

GROUP	N	Mean	Std. Deviation	Std. Error Mean
PROTEN A	12	10.4000	1.1314	.3266
B	12	9.0000	1.8742	.5410

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	95% Confidence Interval of the Difference	
									Lower	Upper
PROTEN	Equal variances assumed	2.776	.110	2.215	22	.037	1.400	.6320	8.936E-02	2.7106
	Equal variances not assumed			2.215	18.08	.040	1.400	.6320	7.267E-02	2.7273

نلاحظ أن الاختبار يجرى في حالتين الأولى تفترض تساوي تباين المجتمعين المأخوذة منهما العينتين أي $\sigma_A^2 = \sigma_B^2$ أما الحالة الثانية فتفترض عدم تساوي التباينات $\sigma_A^2 \neq \sigma_B^2$. ففي الحالة الأولى فأن قيمة $P\text{-Value} = 0.037 < 0.05$ تدفعنا الى رفض فرضية العدم بمستوى دلالة 5% أي توجد فروق معنوية بين متوسطي نسبة البروتين لصنفي الحنطة أما في حالة الاختبار بمستوى دلالة 1% فأن قيمة $P\text{-Value} > 0.01$ ولهذا نقبل فرضية العدم بمستوى دلالة 1% .

أما اختبار Leven لتجانس التباين فأن $P\text{-value} = 0.11 > 0.05$ تدعونا الى قبول فرضية العدم القائلة بتجانس (تساوي) تباين المجتمعين المأخوذة منهما العينتين .

(8-4) اختبار T للملاحظات المزدوجة Paired-Samples T-Test

يستعمل هذا الاختبار لاكتشاف معنوية الفروق بين متوسطي متغيرين لمجموعة (عينة) واحدة حيث تكون مشاهدات العينة على هيئة أزواج مثلاً اختبار معنوية الفرق بين متوسط نسبة الكوليسترول قبل تعاطي عقار معين وبعده في عينة مكونة من 12 شخصاً .

مثال:

زرع صنفين (A و B) من الذرة الصفراء في عشر مناطق واستخدمت قطعتان متساويتان في كل منطقة زرعت إحداهما بالصنف A وزرعت الأخرى بالصنف B والبيانات التالية تمثل كمية المحصول في كل قطعة :

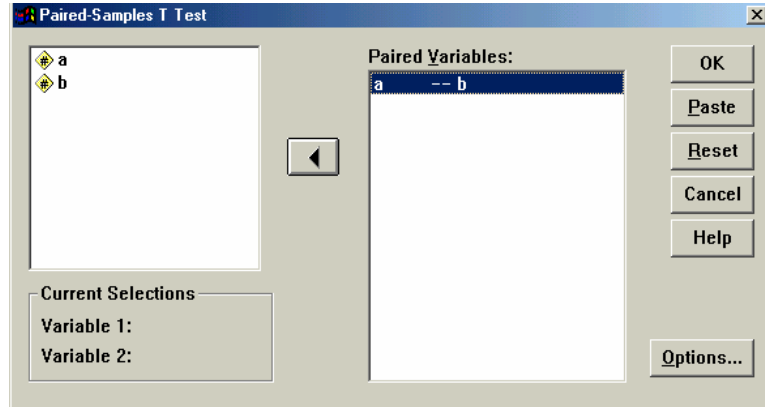
المنطقة	1	2	3	4	5	6	7	8	9	10
الصنف A	127	195	162	170	143	205	168	175	197	136
الصنف B	135	200	160	182	147	200	172	186	194	141

المطلوب اختبار الفرضية القائلة بتساوي متوسطي كمية الإنتاج للمحصولين بمستوى دلالة 5% .

لتنفيذ هذا الاختبار نتبع الخطوات التالية :

من شريط القوائم اختر Analyze → Compare Means → Paired Samples T Test

فيظهر صندوق حوار Paired Samples T Test الذي نقوم بترتيبه على الشكل التالي :



- المتغيران a و b يمثلان مشاهدات الصنفين a و b حيث يتوجب إدخال كلا المتغيرين في نفس الوقت في قائمة Paired variables باعتبار أن المشاهدات مزدوجة حسب الخطوات التالية :
1. نقر المتغير a بزر الماوس الأيسر لاختياره .
 2. أضغط مفتاح Shift مع نقر المتغير b بزر الماوس الأيسر لاختيار المتغيرين معاً .
 3. أنقر الزر  لنقل المتغيرين الى قائمة Paired variables
- أقر زر OK للحصول على نتائج الاختبار وكما يلي :

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	A	167.80	10	26.58	8.40
	B	171.70	10	24.60	7.78

Paired Samples Correlations

	N	Correlation	Sig.
Pair 1 A & B	10	.978	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	A - B	-3.90	5.74	1.82	-8.01	.21	-2.147	9	.060

حيث تم احتساب معامل الارتباط بين المتغيرين 0.978 وهو معنوي بمستوى دلالة 5% و 1%. أما بالنسبة لاختبار T فإن قيمة $P\text{-value}=0.060>0.05$ تدعونا الى قبول فرضية العدم $H_0: \mu A = \mu B$ أي عدم وجود اختلافات معنوية بين متوسطي الإنتاج لصنفي الحنطة .

الفصل التاسع

تحليل التباين

Analysis of Variance

يقصد بتحليل التباين العمليات الرياضية الخاصة بتقسيم مجموع المربعات الكلي لمجموعة من البيانات الى مصادره المختلفة وتلخص نتائج التحليل في جدول يعرف بجدول تحليل التباين ANOVA Table. أن الهدف من إجراء هذا التحليل هو اختبار فرضية تساوي متوسطات مجموعة من العينات وتعرف بالمعالجات (أو المعاملات Treatments) دفعة واحدة ولهذا فهو يعتبر توسيعاً لاختبار t الذي يستعمل لاختبار الفرضية الخاصة بتساوي متوسطي عينتين فقط .

(1 - 9) تحليل التباين لمعيار واحد One Way ANOVA

مثال 1 :

استخدمت أربعة طرق صناعية لإنتاج نوع معين من القماش وبثلاثة مكررات لكل طريقة وكانت نتيجة التجربة كما في الجدول التالي:

الطريقة	المكررا	1	2	3	المتوسط
الطريقة 1		55	47	48	50
الطريقة 2		55	64	64	61
الطريقة 3		55	49	52	52
الطريقة 4		50	44	41	45

المصدر : د. محمود المشهداني وكمال المشهداني ، تصميم وتحليل التجارب ، جامعة بغداد ، 1989 ، ص 63 .

المطلوب مايلي :

1. إجراء تحليل التباين واختبار معنوية الفرق بين متوسطات الطرق الصناعية لمستوى دلالة 5%.
2. في حالة ظهور معنوية الفروق بين الطرق الصناعية باستخدام اختبار F يطلب مايلي :
أ. اختبار معنوية الفروق بين متوسطي كل معالجتين (طريقتين) باستخدام طريقة الفرق المعنوي الأصغر L.S.D. لمستوى دلالة 5% .
ب. اختبار معنوية الفرق بين متوسط كل من الطرق 2 و 3 و 4 وبين متوسط الطريقة 1 بافتراضها الطريقة القياسية (السيطرة) باستخدام طريقة Dunnett .
1. أن الهدف من إجراء تحليل التباين هو اختبار فرضية عدم القائلة بتساوي متوسطات الطرق ضد الفرضية البديلة التي تنص على عدم تساوي متوسطي طريقتين على الأقل أي اختبار الفرضية التالية :

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

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

حيث ان μ يمثل متوسط المجتمع للطريقة أما المتوسط الوارد في الجدول أعلاه يمثل متوسط العينة لكل طريقة .

لأجراء تحليل التباين نتبع الخطوات التالية :

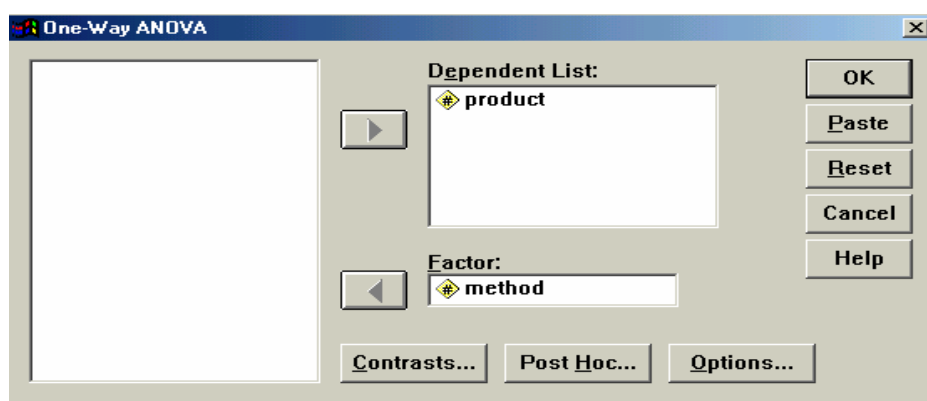
إدخال البيانات في ورقة Data Editor على الشكل التالي :

method	product
1	55
1	47
1	48
2	55
2	64
2	64
3	55
3	49
3	52
4	50
4	44
4	41

من شريط القوائم اختر

Analyze → Compare Means → One-Way NOVA

فيظهر صندوق حوار One-Way ANOVA الذي نقوم بترتيبه كما يلي :



حيث أن :

Dependent List : يمثل المتغير المعتمد ويمثل نتيجة التجربة في هذا المثال ويجب أن يكون متغيراً عددياً. علماً أنه بالامكان استخدام أكثر من متغير معتمد (المكررات) لنفس المعالجات Factor وفي هذه الحالة نحصل على عدد من جداول تحليل التباين بقدر عدد المتغيرات المضمنة في الخانة Dependent . Factor : يمثل المتغير المستقل الذي يستعمل في تعريف الفئات ويمثل نوع الطريقة الصناعية في هذا المثال وأن هذا المتغير يجب أن يكون متغيراً عددياً Numeric.

عند نقر زر OK يعرض البرنامج جدول تحليل التباين كالاتي :

ANOVA

PRODUCT					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	402.000	3	134.000	7.053	.012
Within Groups	152.000	8	19.000		
Total	554.000	11			

أن قيمة $P\text{-Value}=0.012$ المصاحبة لإحصائية F أقل من 0.05 ولهذا نستطيع رفض فرضية العدم لمستوى دلالة 5% أي توجد فروق معنوية بين متوسطات الطرق الصناعية الأربع .

2. نظراً لوجود فروق معنوية بين متوسطات الطرق فهذا يعني عدم تساوي متوسطي طريقتين على الأقل ولاختبار معنوية الفرق لكل زوج من المعالجات نلجأ الى المقارنات المتعددة Multiple Comparisons باستخدام طريقتي L.S.D. و Dunnett بأتباع الخطوات التالية :

أقر زر Post Hoc في صندوق حوار One-Way ANOVA فيظهر صندوق حوار Post Hoc Multiple comparisons الذي نرتبه كما يلي :

نلاحظ وجود طرق عديدة للمقارنات المتعددة حيث قمنا بتأشير LSD و Dunnett وبالنسبة للأخيرة يتوجب تحديد مجموعة السيطرة Control Category (First , Last) وقد اخترنا First لأن الطريقة الأولى هي طريقة السيطرة التي يطلب المقارنة بها مع تحديد نوع الاختبار (من طرف واحد أو طرفين في Test . واخيراً تحديد مستوى المعنوية المطلوب 5% في Significance Level .

عند نقر زر Continue ثم زر OK نحصل على المخرجات التالية :

Multiple Comparisons

Dependent Variable: PRODUCT

	(I) METHOD	(J) METHOD	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
LSD	1	2	-11.00*	3.56	.015	-19.21	-2.79
		3	-2.00	3.56	.590	-10.21	6.21
		4	5.00	3.56	.198	-3.21	13.21
	2	1	11.00*	3.56	.015	2.79	19.21
		3	9.00*	3.56	.035	.79	17.21
		4	16.00*	3.56	.002	7.79	24.21
	3	1	2.00	3.56	.590	-6.21	10.21
		2	-9.00*	3.56	.035	-17.21	-.79
		4	7.00	3.56	.085	-1.21	15.21
	4	1	-5.00	3.56	.198	-13.21	3.21
		2	-16.00*	3.56	.002	-24.21	-7.79
		3	-7.00	3.56	.085	-15.21	1.21
Dunnett t (2-sided) ^a	2	1	11.00*	3.56	.037	.75	21.25
	3	1	2.00	3.56	.896	-8.25	12.25
	4	1	-5.00	3.56	.409	-15.25	5.25

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

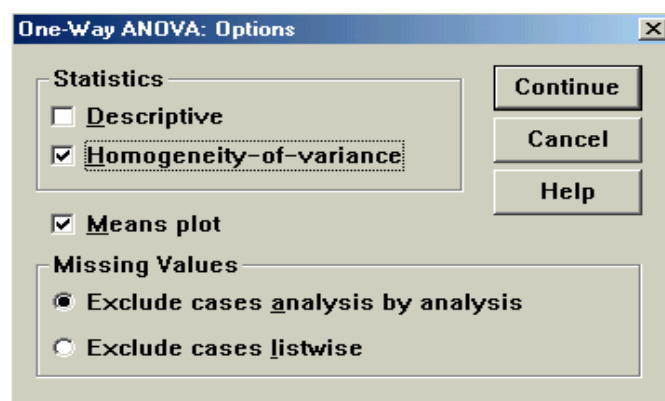
a. Dunnett t-tests treat one group as a control, and compare all other groups against it.

يلاحظ أنه باستعمال طريقة LSD فقد ظهرت فروق معنوية بمستوى دلالة 5% بين متوسطات المعالجات (2و1)، (3و2)، (4و2) حيث كانت قيمة P-Value أو Sig. أقل من 0.05 مع وجود علامة * على فروق المتوسطات .

أما اختبار Dunnett فيشير الى وجود فروق معنوية بمستوى دلالة 5% بين متوسطي الطريقة الأولى و الطريقة الثانية عند المقارنة بالطريقة الأولى باعتبارها مجموعة سيطرة .

ملاحظة :

عند نقر زر Options في صندوق حوار One Way ANOVA يظهر صندوق الحوار التالي :



حيث أن :

Descriptive : لعرض بعض المقاييس الوصفية مثل الوسط الحسابي ، الانحراف المعياري
Homogeneity-of-Variance : لاختبار تجانس تباين المعالجات (الطرق في هذا المثال) باستخدام إحصائية Levene .حيث يعتبر تجانس التباين أحد الفروض المهمة في إجراء تحليل التباين .
Means plot :لعرض تخطيط يمثل متوسط المعالجات .
Exclude Cases Analysis by Analysis :استبعاد الحالات التي تحتوي قيماً مفقودة للمتغيرات المشمولة بالتحليل فقط .

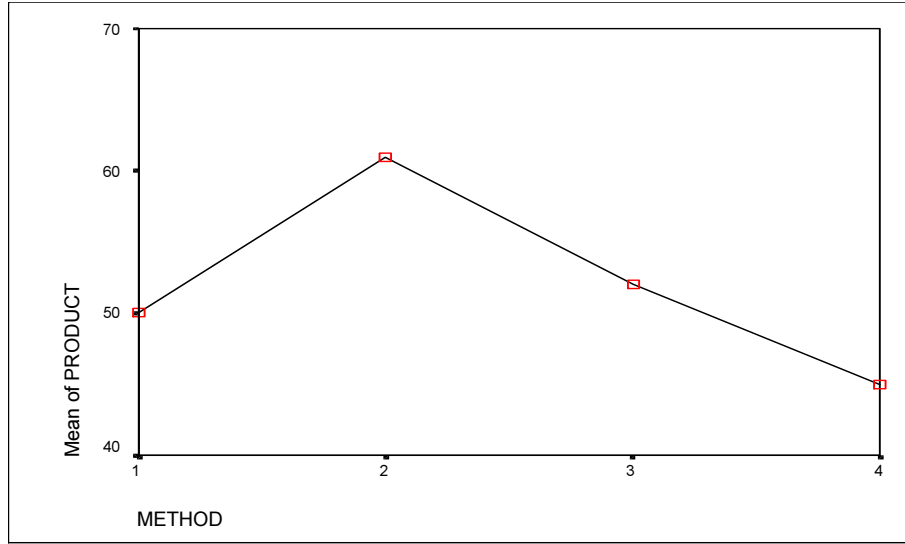
Exclude Cases Listwise :استبعاد الحالات التي تحتوي قيماً مفقودة لأي واحد من المتغيرات .
في هذا المثال لايم تأشير أي من الخيارين لعدم أحتواء البيانات على قيم مفقودة .
عند نقر زر Continue في صندوق الحوار أعلاه ثم زر OK في صندوق حوار One-Way ANOVA يتم عرض النتائج التالية :

Test of Homogeneity of Variances

PRODUCT			
Levene Statistic	df1	df2	Sig.
.667	3	8	.596

المخرج أعلاه يبين نتيجة اختبار فرضية العدم (تجانس التباين) ضد الفرضية البديلة (عدم تجانس التباين) باستخدام إحصائية Levene حيث أن قيمة $P\text{-Value} = 0.596 > 0.05$ تدعونا الى قبول فرضية العدم القائلة بتجانس التباينات.(راجع الأمر Explore حول إحصائية Levene) .
المخطط التالي يمثل الوسط الحسابي لكل معالجة :

Means Plots



(9-1-1) المقارنات المستقلة Orthogonal Comparisons

في معظم التجارب يهتم الباحث بأجراء مقارنة بين بعض المعالجات دون غيرها . فإذا كان لدينا معالجات (مجموعات) عددها t فإنه يمكن تكوين $t-1$ من المقارنات المستقلة وتعرف باسم Contrast (تقابلات) الذي تأخذ العلاقة الخطية التالية :

$$\bar{Q} = \sum C_i Y_{i.} \quad \text{with} \quad \sum C_i = 0$$

حيث ان $Y_{i.}$ يمثل مجموع المعالجة i وأن C_i هي معاملات Coefficients فإذا كان لدينا مقارنتين $Q_1 = \sum C_{1i} Y_{i.}$ وأن $Q_2 = \sum C_{2i} Y_{i.}$ فإنهما تكونان مستقلتين (متعامدتين) Orthogonal إذا كان $\sum C_{1i} C_{2i} = 0$ وعليه يكون من الممكن تقسيم مجموع مربعات المعالجات SS_t الذي له درجة حرية $t-1$ الى $t-1$ من المقارنات المستقلة ولكل منها درجة حرية واحدة ويكون بالإمكان اختبار معنوية كل من هذه المقارنات على حدة . لتوضيح ما سبق نأخذ المثال التالي :

مثال 2:

أجريت تجربة لدراسة تأثير خمسة أنواع من العلائق على نمو الفراخ وسجلت بيانات عن الوزن وبأربعة مكررات لكل نوع ونظمت في الجدول التالي

المجموع $Y_{i.}$	الوزن Weight				نوع العليقة
168	40	42	40	46	1
188	42	47	48	51	2
168	46	44	42	36	3
172	43	45	42	42	4
144	36	37	36	35	5

المصدر : د. خاشع الراوي وآخرون : تحليل وتصميم التجارب الزراعية ، جامعة

الموصل ، ص 56 .

يطلب مايلي :

1. أجراء تحليل التباين لاختبار معنوية الفرق بين متوسطات المعالجات .

2. اختبار معنوية المقارنات الأربعة المستقلة $t-1$ التالية :

$$Q_1 = 4Y_{1.} - Y_{2.} - Y_{3.} - Y_{4.} - Y_{5.} = 0$$

$$Q_2 = Y_{2.} + Y_{3.} - Y_{4.} - Y_{5.} = 0$$

$$Q_3 = Y_{2.} - Y_{3.} = 0$$

$$Q_4 = Y_{4.} - Y_{5.} = 0$$

في البداية يتم إدخال البيانات الى ورقة Data Editor كما يلي :

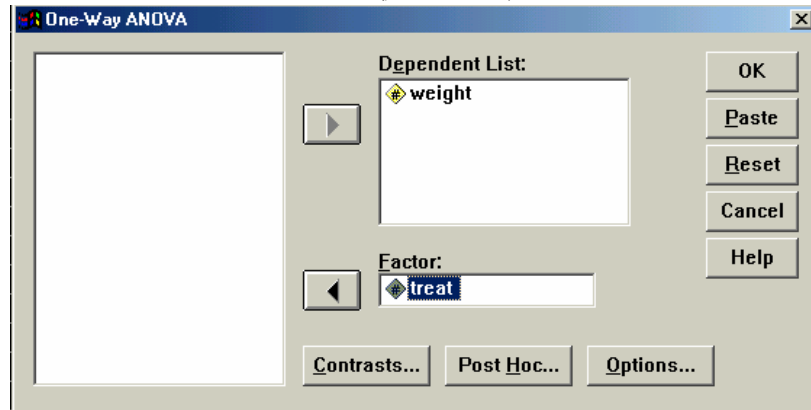
treat	weight
1	46
1	40
1	42
1	40
2	51
2	48
2	47
2	42
3	36
3	42
3	44
3	46
4	42
4	42
4	45
4	43
5	35
5	36
5	37
5	36

لتنفيذ المطلوبين الأول والثاني نتبع الخطوات التالية :

من القوائم اختر

Analyze → Compare Means → One-Way ANOVA

حوار One-Way ANOVA الذي نرتبه كما يلي :



أقر زر Contrasts في صندوق حوار One-Way ANOVA فيظهر صندوق حوار Contrasts حيث نقوم بإدخال المقارنات واحدة تلو الأخرى فبالنسبة للمقارنة الأولى يتم إدخال معاملاتها كما يلي :

- انقر المربع المجاور لكلمة Coefficients ثم أكتب الرقم 4 ثم انقر زر Add فيتم إضافة المعامل 4 الى المستطيل في الأسفل .
- أكتب الرقم 1- في المربع المجاور لكلمة Coefficients ثم انقر الزر Add فيتم إضافة المعامل - 1 الى المستطيل في الأسفل .
- أكتب الرقم 1- في المربع المجاور لكلمة Coefficients ثم انقر الزر Add فيتم إضافة المعامل - 1 الى المستطيل في الأسفل .
- أكتب الرقم 1- في المربع المجاور لكلمة Coefficients ثم انقر الزر Add فيتم إضافة المعامل - 1 الى المستطيل في الأسفل .
- أكتب الرقم 1- في المربع المجاور لكلمة Coefficients ثم انقر الزر Add فيتم إضافة المعامل - 1 الى المستطيل في الأسفل .

وبهذا نكون قد أدخلنا كافة معاملات المقارنة الأولى .لإدخال معاملات المقارنة الثانية انقر زر **Next** ثم أبدأ بإدخال المعاملات بنفس طريقة إدخال معاملات المقارنة الأولى . وهكذا لبقية المقارنات . تظهر معاملات المقارنة الأولى مثلاً في صندوق حوار **Contrasts** كما يلي :

لانتقال الى المقارنة التالية انقر زر **Next** وللمعودة الى مقارنة سابقة انقر الزر **Previous**.
 انقر زر **Continue** ثم زر **OK** في صندوق حوار **One-Way ANOVA** فيقوم البرنامج بعرض المخرج التالي :

ANOVA

WEIGHT

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	248.000	4	62.000	7.154	.002
Within Groups	130.000	15	8.667		
Total	378.000	19			

أن اختبار **F** يشير الى معنوية الفروق بين متوسطات المعالجات (العلائق) بمستوى دلالة 5% و 1% .

Contrast Coefficients

Contrast	TREAT				
	1	2	3	4	5
1	4	-1	-1	-1	-1
2	0	1	1	-1	-1
3	0	1	-1	0	0
4	0	0	0	1	-1

الجدول أعلاه يبين معاملات المقارنات الأربعة . أما الجدول التالي فيبين اختبار **t** لمعنوية المقارنات ومنه يتضح أن المقارنة الأولى غير معنوية ($Q1=0$) بمستوى دلالة 5% لأن $P\text{-Value} = 1 > 0.05$ أما بقية المقارنات فهي معنوية (تختلف جوهرياً عن الصفر) بمستوى دلالة 5% لأن $P\text{-Value} < 0.05$ في حالة افتراض تساوي التباينات . مع العلم أن قيمة المقارنة **Value of Contrast** تستخرج كما يلي :

$$\text{Value of Contrast} = \sum C_i Y_{i.} / r$$

r : يمثل عدد المشاهدات في كل مجموعة ويساوي 4 .

Contrast Tests

	Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
WEIGHT Assume equal variances	1	.00	6.58	.000	15	1.000
	2	10.00	2.94	3.397	15	.004
	3	5.00	2.08	2.402	15	.030
	4	7.00	2.08	3.363	15	.004
	Does not assume equal variances	1	.00	6.39	4.727	1.000
		2	10.00	2.97	6.823	.012
		3	5.00	2.86	5.880	.132
		4	7.00	.82	8.573	.000

Trend Analysis (2-1-9) تحليل الاتجاهات

يمكن تجزئة مجموع مربعات المعالجات الى مركبات خطية ، من الدرجة الثانية ، تكعيبية...حيث أن درجة متعدد الحدود Polynomial تساوي $t-1$ حيث أن t يمثل عدد المعالجات ويستفاد من هذا التحليل في التجارب التي تكون المعالجات فيها عبارة عن مستويات لعامل كمي مثلاً استخدام كميات مختلفة من سماد معين ويكون الهدف هو اختبار معنوية المركبات المختلفة لغرض تكوين فكرة عن استجابة الصفة المدروسة لمستوى المعالجة وبالتالي تقدير المستوى الأمثل من العامل المدروس .

لتحليل مجموع مربعات المعالجات للمثال السابق فإن درجة متعدد الحدود تساوي $5-1=4$ أي أنه بالإمكان الحصول على المركبة الخطية ، من الدرجة الثانية ، تكعيبية ، من الدرجة الرابعة .حيث نقوم بتحويل صندوق حوار Contrasts كما يلي :

علماً أن هذه المركبات تكون مستقلة عن بعضها ولها نفس صفات المقارنات Contrasts. لاحظ أننا لانحتاج الى كتابة المعاملات Coefficients بعد أن اخترنا Polynomial وحددنا درجة متعدد الحدود 4th. عند نقر زر Continue ثم زر OK في صندوق حوار One-way ANOVA نحصل على النتيجة التالية :

ANOVA

WEIGHT

		Sum of Squares	df	Mean Square	F	Sig.
Between Groups	(Combined)	248.000	4	62.000	7.154	.002
	Linear Term Contrast	102.400	1	102.400	11.815	.004
	Deviation	145.600	3	48.533	5.600	.009
	Quadratic Term Contrast	92.571	1	92.571	10.681	.005
	Deviation	53.029	2	26.514	3.059	.077
	Cubic Term Contrast	1.600	1	1.600	.185	.674
	Deviation	51.429	1	51.429	5.934	.028
	4th-order Term Contrast	51.429	1	51.429	5.934	.028
Within Groups		130.000	15	8.667		
Total		378.000	19			

لاحظ أنه تم توزيع مجموع المربعات بين المعالجات Between Groups على المركبات الأربعة مع إعطاء درجة حرية واحدة لكل مركبة كما أن جميع المركبات ظهرت معنوية عند اختبارها بمستوى دلالة 5% عدا المركبة التكعيبية Cubic term .

(2-9) تحليل التباين لمعيارين Two Way ANOVA

يستفاد من تحليل التباين لمعيارين في اختبار تساوي متوسطات معالجات العامل الأول إضافة إلى اختبار تساوي متوسطات العامل الثاني و هذا النموذج يقارب تصميم القطاعات الكاملة العشوائية RCB Design في تصميم التجارب .

مثال 3 :

في تجربة قطاعات كاملة عشوائية تم الحصول على الجدول التالي الذي يبين الفرق في سمك الأطار (المتغير المعتمد) بعد قطع مسافة معينة وقد استخدمت أربعة أنواع من الإطارات (العامل الأول) واستخدمت أربعة سيارات (القطاعات وسنعتبرها العامل الثاني) وقد كانت البيانات المختزلة (المبسطة) كما يلي :

نوع الإطار السيارة	A	B	C	D
1	4	1	-1	0
2	1	1	-1	-2
3	0	0	-3	-2
4	0	-5	-4	-4

المصدر: شارلز هيكس - ترجمة قيس سبع خماس ، المفاهيم الأساسية

في تصميم التجارب ، الجامعة المستنصرية ، 1984، ص 83 .

يرغب الباحث في تكوين جدول تحليل التباين لمعيارين لاختبار معنوية الفرق بين متوسطات أنواع الإطارات Tyre وكذلك معنوية الفرق بين أنواع السيارات car وبمستوى دلالة 5% .
في البداية نقوم بإدخال البيانات في Data Editor وكما يلي :

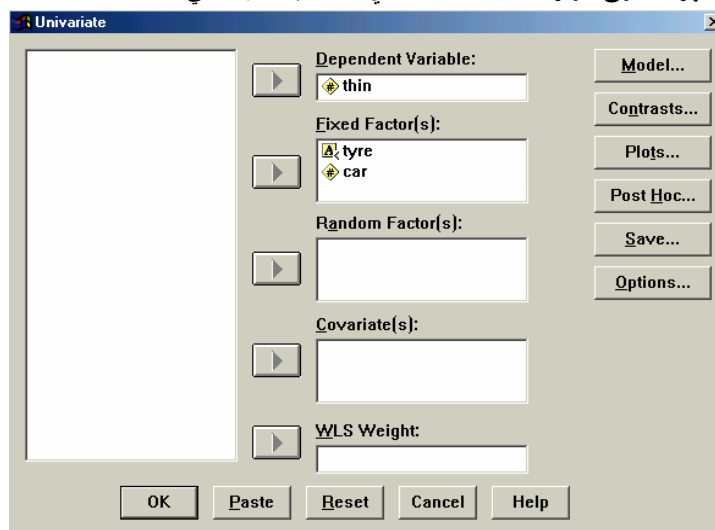
car	tyre	thin
A	1	4
A	2	1
A	3	0
A	4	0
B	1	1
B	2	1
B	3	0
B	4	-5
C	1	-1
C	2	-1
C	3	-3
C	4	-4
D	1	0
D	2	-2
D	3	-2
D	4	-4

لتكوين جدول تحليل التباين نتبع الخطوات التالية :

من شريط القوائم نختار :

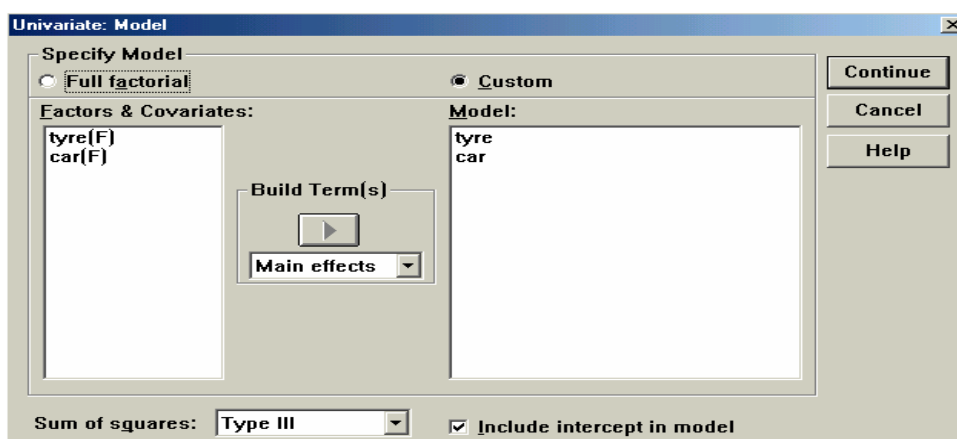
Analyze → General Linear Model → Univariate

فيظهر صندوق حوار Univariate الذي ننظمه بالشكل التالي :



أدخلنا العاملين في Fixed Factors على اعتبار أن كافة معالجات العامل قد ضمنت في التجربة أما في حالة أخذ عينة من معالجات العامل فأنا سنستعمل خانة Random Factors. لاحظ أن العامل الثابت Fixed Factor يمكن أن يكون متغيراً رمزياً على العكس منه في حالة One-Way ANOVA حيث يجب أن يكون Factor متغيراً عددياً .

أقر زر Model فيظهر صندوق حوار Model ، اختر Custom بدلاً من Full Factorial وذلك لأننا لا نرغب في ظهور التفاعل Interaction في جدول تحليل التباين (car*tyre) لعدم وجود درجات حرية كافية للخطأ التجريبي حيث يظهر صندوق حوار Model بعد ترتيبه كما يلي :



حيث أن تأثير Include Intercept in Model يعمل على تضمين الحد الثابت في النموذج الخطي العام باعتباره نموذج انحدار .

أما خانة Build Terms فتستعمل لتعيين نوع التأثيرات Effects التي يراد إظهارها في جدول تحليل التباين حيث قمنا بنقل المتغيرين tyre و car من خانة Factors & Covariates الى خانة Model بنقر

كل منهما بزر الأيسر ثم نقر زر  لتظهر كتأثيرات رئيسية Main Effects (بعد التأكد من أن خانة Build Terms تتضمن الخيار Main effects).

عند نقر زر Continue ثم OK في صندوق حوار Univariate نحصل على المخرج التالي :

Between-Subjects Factors

	N
TYRE A	4
B	4
C	4
D	4
CAR 1	4
2	4
3	4
4	4

Tests of Between-Subjects Effects

Dependent Variable: THIN

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	69.375 ^a	6	11.563	9.000	.002
Intercept	14.063	1	14.063	10.946	.009
TYRE	30.688	3	10.229	7.962	.007
CAR	38.688	3	12.896	10.038	.003
Error	11.563	9	1.285		
Total	95.000	16			
Corrected Total	80.938	15			

a. R Squared = .857 (Adjusted R Squared = .762)

لاحظ أن :

SS. Corrected Model = SS.TYRE + SS. CAR

SS. Corrected Total = SS. Total – SS.Intercept

أن هذا الجدول يخص النموذج الخطي العام GLM الذي هو نموذج انحدار ويمكن تبسيط هذا الجدول (راجع فصل الجداول المحورية) ليكون جدول تحليل تباين بمعيارين Two-Way ANOVA كما يلي :

Tests of Between-Subjects Effects

Dependent Variable: THIN

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
TYRE	30.688	3	10.229	7.962	.007
CAR	38.688	3	12.896	10.038	.003
Error	11.563	9	1.285		
Corrected Total	80.938	15			

ومنه يتضح معنوية الفروق بين الإطارات tyer وكذلك بين السيارات car بمستوى دلالة 5% .

ملاحظات :

1. لإظهار التفاعل بين tyre و car (علماً أنه لن نحصل على قيمة F لهذا المثال لعدم كفاية درجات الحرية) نتبع الخطوات التالية:

• أنقر السهم المتجه نحو الأسفل في خانة Build Terms وأختار Interaction في صندوق حوار Model.

• أختار المتغيرين tyre و car من الخانة Factors & Covariates في نفس الوقت (بنقر المتغير tyre ثم أضغط مفتاح Shift مع نقر المتغير car).

• أنقر زر  لنقل المتغيرين المذكورين الى خانة Model .
 علماً أن نفس الشيء يمكن أن ينجز باختيار Full Factorial في صندوق حوار Model .
 2. يمكن استعمال النموذج الخطي العام GLM في إجراء تحليل التباين لمعيار واحد .
 مثال 4 : (تحليل التباين لمعيارين مع تسجيل أكثر من مشاهدة لكل وحدة تجريبية)

الجدول التالي يمثل المبيعات الأسبوعية لسلسلة من المتاجر حسب موقع الرف Shelf Location (العامل الأول) و حجم المتجر Store Size (العامل الثاني) وقد أخذت مشاهدتين لكل من التوافيق الممكنة :

موقع الرف Shelf Location				حجم المتجر Size
D	C	B	A	
48	65	56	45	Small
53	71	63	50	
60	73	69	57	Medium
57	80	78	65	
71	82	75	70	Large
75	89	82	78	

يطلب تكوين جدول تحليل التباين لمعيارين مع اختبار معنوية الفروق بين متوسطات معالجات عامل الحجم Size وبين معالجات الموقع Location واختبار معنوية التفاعل Size*Location بمستوى دلالة 5% مع توضيح ذلك بيانياً .

في البداية نقوم بإدخال البيانات في Data Editor وكما يلي :

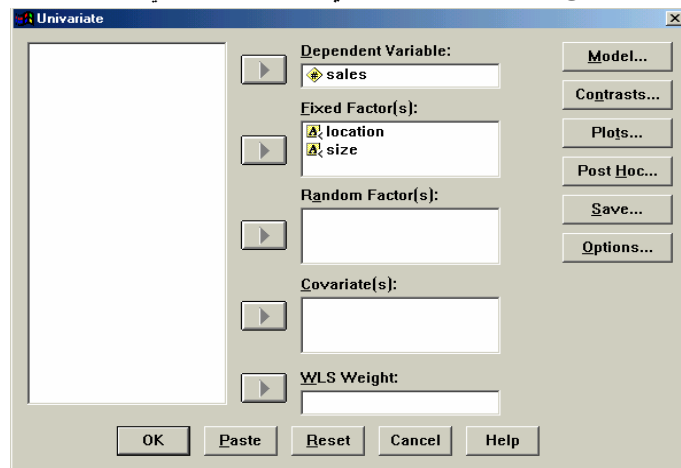
location	size	sales
A	Small	45
A	Small	50
A	Medium	57
A	Medium	65
A	Large	70
A	Large	78
B	Small	56
B	Small	63
B	Medium	69
B	Medium	78
B	Large	75
B	Large	82
C	Small	65
C	Small	71
C	Medium	73
C	Medium	80
C	Large	82
C	Large	89
D	Small	48
D	Small	53
D	Medium	60
D	Medium	57
D	Large	71
D	Large	75

يتم تنفيذ المطلوب حسب الخطوات التالية :

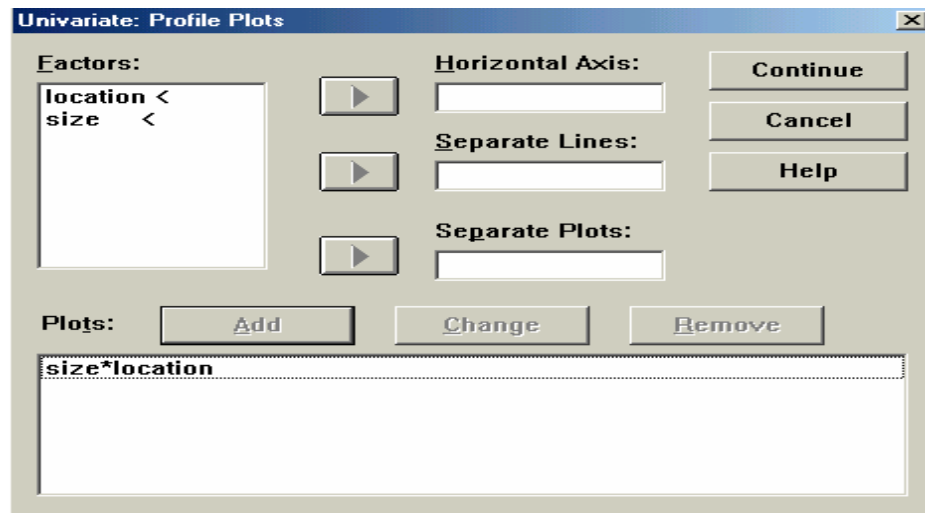
← من شريط القوائم نختار :

Analyze → General Linear Model → Univariate

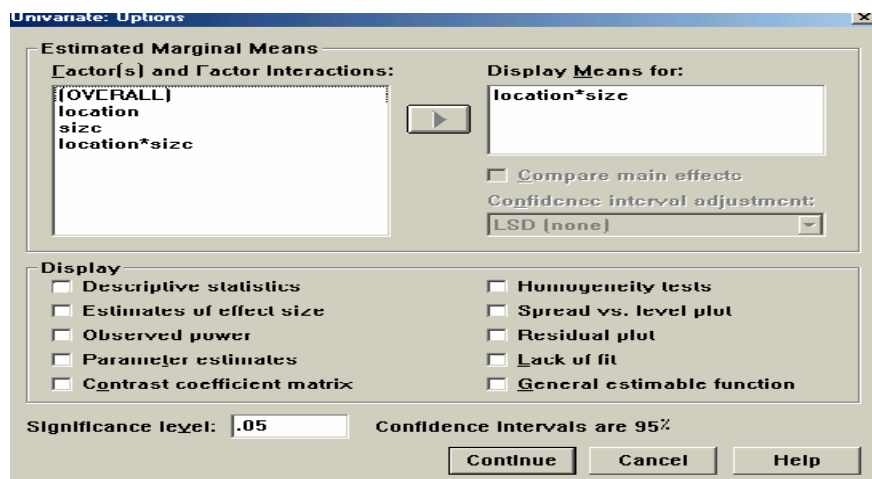
فيظهر صندوق حوار Univariate الذي ننظمه بالشكل التالي :



- أنقر زر Model وأختار Full Factorial لأننا نرغب في الحصول على التفاعل Interaction لوجود عدد كافي من درجات الحرية للخطأ العشوائي .
- (إذا كنت ترغب في اختيار جزء من التأثيرات أنقر Custom وأختار التأثيرات المطلوبة) علماً أن Full Factorial هو الخيار الافتراضي Default .
- أنقر زر Plots لأعداد تخطيط تمثل فيه فئات المتغير size على المحور الأفقي تقابله سلاسل المتوسطات الأربعة للمتغير Location فيظهر صندوق حوار Profile Plots أو Interaction Plot الذي يستفاد منه لمقارنة المتوسطات الحدية للمتغير المعتمد في النموذج و ينظم كما يلي :



- يتم إدخال المتغيرات في الصندوق أعلاه كما يلي :
- أنقل المتغير size من خانة Factors الى خانة Horizontal Axis.
 - أنقل المتغير Location الى خانة Separate Lines .
 - ستلاحظ أن الزر Add أصبح فعالاً أنقره فيضاف التفاعل size*location الى خانة Plot في الأسفل .
- أما قائمة Separate Plots فيستفاد منها في عمل تخطيط عند كل مستوى من مستويات عامل ثالث .
- أنقر زر Continue للرجوع الى صندوق حوار Univariate .
- أنقر زر Options لعرض المتوسطات الحدية Estimated Marginal Means للتفاعل location*size حيث ينظم صندوق حوار Options بالشكل التالي :



عند نقر زر Continue ثم OK في صندوق حوار Univariate نحصل على النتائج التالية .

Tests of Between-Subjects Effects

Dependent Variable: SALES

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	3019.333 ^a	11	274.485	12.767	.000
Intercept	108272.667	1	108272.667	5035.938	.000
LOCATION	1102.333	3	367.444	17.090	.000
SIZE	1828.083	2	914.042	42.514	.000
LOCATION * SIZE	88.917	6	14.819	.689	.663
Error	258.000	12	21.500		
Total	111550.000	24			
Corrected Total	3277.333	23			

a. R Squared = .921 (Adjusted R Squared = .849)

أن هذا الجدول يمثل نتائج الانحدار للنموذج الخطي العام وبالإمكان تبسيطه ليكون متلائماً مع نموذج تحليل التباين وكما يلي :

Tests of Between-Subjects Effects

Dependent Variable: SALES

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
LOCATION	1102.333	3	367.444	17.090	.000
SIZE	1828.083	2	914.042	42.514	.000
LOCATION * SIZE	88.917	6	14.819	.689	.663
Error	258.000	12	21.500		
Corrected Total	3277.333	23			

أن اختبار F يشير إلى وجود فروقات معنوية بين معالجات العامل Location وكذلك بين معالجات العامل size بينما لم تظهر معنوية حد التفاعل بمستوى دلالة 5% لأن $p\text{-Value} = 0.663 > 0.05$.

Estimated Marginal Means

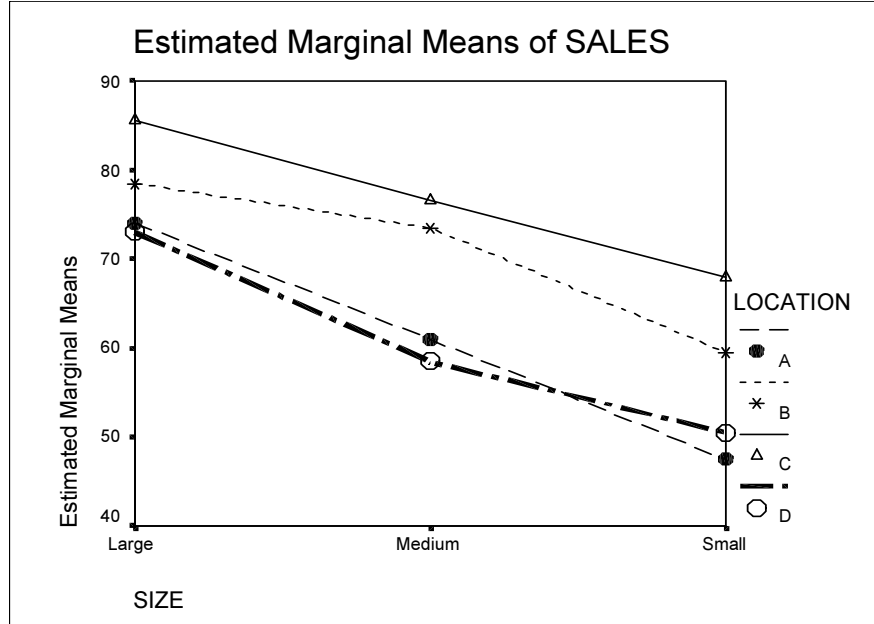
LOCATION * SIZE

Dependent Variable: SALES

LOCATION	SIZE	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
A	Large	74.000	3.279	66.856	81.144
	Medium	61.000	3.279	53.856	68.144
	Small	47.500	3.279	40.356	54.644
B	Large	78.500	3.279	71.356	85.644
	Medium	73.500	3.279	66.356	80.644
	Small	59.500	3.279	52.356	66.644
C	Large	85.500	3.279	78.356	92.644
	Medium	76.500	3.279	69.356	83.644
	Small	68.000	3.279	60.856	75.144
D	Large	73.000	3.279	65.856	80.144
	Medium	58.500	3.279	51.356	65.644
	Small	50.500	3.279	43.356	57.644

المخطط التالي يبين درجة التفاعل بين العاملين Location و Size حيث يلاحظ أن الخطوط متوازية تقريباً مما يشير إلى أن الفروق في المواقع لا تتغير بتغير حجم المتجر كما أن الخطوط المتوازية تؤثر عدم وجود تفاعل بين العاملين size و location وهذا يدعم ما توصلنا إليه بتطبيق اختبار F .

Profile Plots



(9-3) تحليل التباين المشترك Covariance Analysis

من المشاكل التي تواجه تحليل التباين أو تصميم التجارب أن هناك عوامل مستقلة (مصاحبة) Covariates يرمز لها بالرمز X التي تؤثر على الظاهرة المراد قياس تأثير المعالجات عليها ويرمز لها بالرمز Y وهو المتغير المعتمد Dependent Variable مثلاً عند قياس تأثير عدد من الأغذية (المعالجات) على زيادة وزن الحيوان Y فإن وزن الحيوان X في بداية التجربة يؤثر على مدى استجابته للغذاء وبالتالي فإن الباحث في نهاية التجربة لا يستطيع معرفة فيما إذا كانت الاختلافات بين الأوزان النهائية للحيوان (في نهاية التجربة) ناتجة عن أثر الغذاء أم عن أثر الوزن في بداية التجربة. لغرض التخلص من أثر المتغير المصاحب X نستخدم تحليل التباين المشترك وهو أسلوب يجمع بين تحليل التباين والانحدار .

مثال 5 :

أجريت تجربة على تغذية الأفراخ لمقارنة تأثير أربعة علائق (المعالجات) مختلفة على زيادة وزن الأفراخ وأستخدمت ستة مشاهدات لكل معالجة. الجدول التالي يبين الزيادة في وزن الأفراخ Y في نهاية التجربة كما يبين وزن الأفراخ X عند بداية التجربة .

المعالجات treat						المتغير	المشاهدات Observations
T1	X	Y	29	33	21	20	27
			151	167	156	130	170
T2	X	Y	25	20	26	20	31
			170	180	161	171	169
T3	X	Y	29	30	35	35	32
			172	160	190	138	189
T4	X	Y	36	28	35	30	32
			189	142	193	200	173

يطلب اختبار معنوية الفروق بين متوسطات المعالجات بعد إزالة أثر الوزن في بداية التجربة X (أسلوب تحليل التباين المشترك) .

في البداية نرتب بيانات الجدول في ورقة Data Editor بالشكل التالي :

لتنفيذ أسلوب تحليل التباين المشترك نتبع الخطوات التالية :

treat	X	Y
T1	30	165
T1	27	170
T1	20	130
T1	21	156
T1	33	167
T1	29	151
T2	24	180
T2	31	169
T2	20	171
T2	26	161
T2	20	180
T2	25	170
T3	34	156
T3	32	189
T3	35	138
T3	35	190
T3	30	160
T3	29	172
T4	41	201
T4	32	173
T4	30	200
T4	35	193
T4	28	142
T4	36	189

Analyze → General Linear Model → Univariate

فيظهر صندوق حوار Univariate الذي نرتبه كما يلي :

عند نقر زر OK نحصل على النتيجة التالية:

Tests of Between-Subjects Effects

Dependent Variable: Y

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	2845.956 ^a	4	711.489	2.572	.071
Intercept	6938.602	1	6938.602	25.087	.000
X	682.831	1	682.831	2.469	.133
TREAT	1609.595	3	536.532	1.940	.157
Error	5255.002	19	276.579		
Total	699323.000	24			
Corrected Total	8100.958	23			

a. R Squared = .351 (Adjusted R Squared = .215)

وقد بسطنا الجدول أعلاه ليكون في صورة تحليل التباين المشترك وكما يلي :

Tests of Between-Subjects Effects

Dependent Variable: Y

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
TREAT	1609.595	3	536.532	1.940	.157
Error	5255.002	19	276.579		
Total+Error	6864.597	22			

a. R Squared = .351 (Adjusted R Squared = .215)

يلاحظ أن قيمة $P\text{-Value} = 0.157 > 0.05$ مما يدعونا الى قبول فرضية العدم بمستوى دلالة 5% التي تنص على تساوي متوسطات الأنواع الأربعة من العلائق بعد إزالة أثر المتغير المصاحب X .

الفصل العاشر

تحليل الارتباط و الانحدار

Correlation & Regression Analysis

(10 - 1) الارتباط Correlation

تسمى العلاقة بين ظاهرتين بالارتباط Correlation مثلاً العلاقة بين الدخل والاستهلاك فمن البديهي أن زيادة دخل الفرد يؤدي الى زيادة استهلاكه من السلع والخدمات (علاقة طردية) كما أن ارتفاع سعر سلعة ما يؤدي الى تدني الطلب عليها (علاقة عكسية) علماً أن الارتباط قد يكون خطياً Linear أو غير خطي Non Linear. أن المقياس المستخدم الذي يقيس درجة الارتباط يعرف بمعامل الارتباط Correlation Coefficients ويرمز له r وتتراوح قيمته بين -1 الى 1 $(-1 \leq r \leq 1)$.

(10 - 2) الارتباط الخطي البسيط Simple Linear Correlation

يحتسب معامل الارتباط الخطي البسيط بافتراض وجود علاقة خطية بين اثنتين من المتغيرات فقط مع العلم أن الحصول على قيمة صغيرة (قريبة من الصفر) لهذا المعامل لا يعني عدم وجود علاقة بين المتغيرين فقد توجد علاقة من الدرجة الثانية (ارتباط غير خطي) .

مثال 1:

البيانات التالية تمثل الدرجات التي حققها 10 طلاب في اختبار اللغة Lang واختبار الرياضيات Math وتظهر في ورقة Data Editor لبرنامج SPSS كما يلي :

Lang	Math
60	56
68	60
60	64
74	82
80	76
84	72
80	74
72	66
62	64
82	86

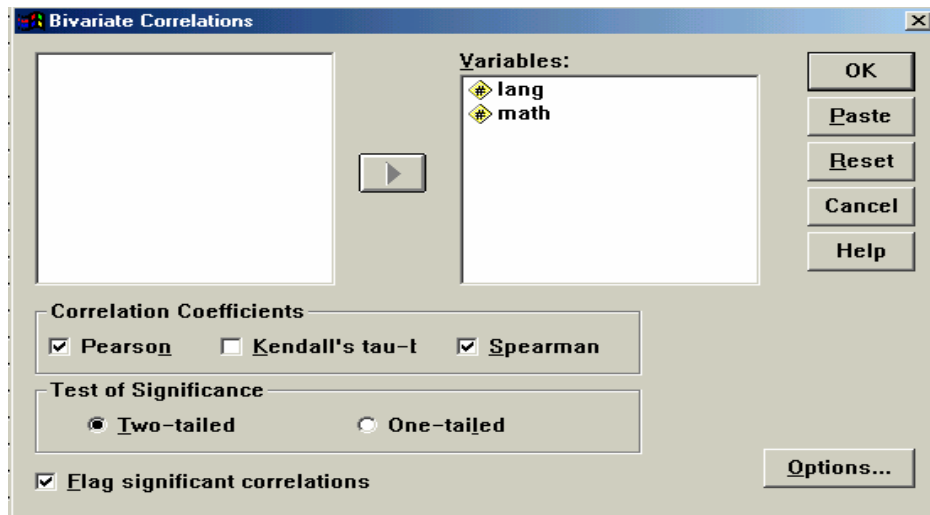
المطلوب مايلي :

1. أوجد معامل الارتباط الخطي البسيط لـ Pearson ، معامل ارتباط الرتب لـ Spearman .

2. أختبر معنوية معامل الارتباط بمستوى دلالة 5% .

لتنفيذ المطالبات أعلاه نتبع الخطوات التالية :

من شريط القوائم اختر Bivariate → Correlate → Analyze فيظهر صندوق حوار Bivariate Correlation الذي نرتبه كما يلي :



يتضمن حقل Correlation Coefficients الخيارات التالية :

Pearson : لاستخراج معامل الارتباط الخطي البسيط للمتغيرات الكمية .

Kendall's Tau : لاستخراج معامل الارتباط بالطرق اللامعلمية وباستعمال الرتب Ranks في حالة الرغبة في تقدير معامل الارتباط بصورة تقريبية . كما أنه من الضروري استخدام هذا المعامل في حالة كون إحدى الظاهرتين أو كلاهما غير مقاستين (ليست متغيرات كمية) ولكن يمكن إعطاؤها رتباً تصاعدياً أو تنازلياً وفي هذه الحالة يتم إدخال الرتب المتناظرة للظاهرتين بدلاً من القيم الأصلية .

Spearman : هو معامل ارتباط للرتب كما هو الحال بالنسبة لمعامل Kendall .

وقد أشرنا الخيارين Pearson و Spearman .

في حقل Test significance أشرنا الخيار Two Tailed لاختبار الفرضية من طرفين . لاختبار

الفرضية من طرف واحد نؤشر One Tailed .

Flag Significance Correlation : لتعليم الارتباطات المعنوية (إظهار علامة Star) .

عند نقر زر OK تظهر النتائج التالية :

Correlations

		LANG	MATH
LANG	Pearson Correlation	1.000	.776**
	Sig. (2-tailed)	.	.008
	N	10	10
MATH	Pearson Correlation	.776**	1.000
	Sig. (2-tailed)	.008	.
	N	10	10

** . Correlation is significant at the 0.01 level

أن قيمة معامل الارتباط الخطي لبيرسون بين LANG و MATH هي $r = 0.776$ لاختبار الفرضية

التالية (أختبار من طرفين) :

$$H_0: \rho = 0$$

$$H_1: \rho \neq 0$$

نستخدم إحصائية T التي تتبع توزيع T بدرجة حرية n-k حيث أن n يمثل حجم المجتمع و k عدد المتغيرات .

$$T = r \frac{\sqrt{n-k}}{\sqrt{1-r^2}} = 0.776 \frac{\sqrt{10-2}}{\sqrt{1-0.776^2}} = 3.48$$

→ يمكنك الحصول على P-

Value من خلال الأمر Compute Transform ثم استعمال الدالة التجميعية لتوزيع T وكما يلي
 $(1 - CDF.T(3.48,8)) * 2 = 0.008$. أن قيمة $P\text{-Value} = 0.008 < 0.05$ تشير الى أن الارتباط بين درجة امتحان اللغة و امتحان الرياضيات يختلف معنوياً عن الصفر بمستوى دلالة 5% (الاختلاف معنوي بمستوى دلالة 1% أيضاً) .

المخرج التالي يمثل معامل ارتباط الرتب لسبيرمان :

Nonparametric Correlations

Correlations

			LANG	MATH
Spearman's rho	LANG	Correlation Coefficient	1.000	.786**
		Sig. (2-tailed)	.	.007
		N	10	10
	MATH	Correlation Coefficient	.786**	1.000
		Sig. (2-tailed)	.007	.
		N	10	10

** . Correlation is significant at the .01 level (2-tailed).

لاختبار فرضية العدم $H_0 : \rho = 0$ ضد الفرضية البديلة $H_1 : \rho \neq 0$ لمعامل Spearman
تستخدم نفس إحصائية T لأختبار Pearson أعلاه . أن قيمة P-Value المرافقة لمعامل ارتباط Spearman تشير الى معنوية هذا المعامل بمستوى دلالة 5% وكذلك 1% .

ملاحظة :

لاختبار فرضية معامل الارتباط من طرف واحد نؤشر الخيار One Tailed في صندوق حوار Bivariate Correlations وهنا توجد حالتين للأختبار :

1. إذا كانت قيمة معامل الارتباط المقدرة موجبة يقوم البرنامج باختبار الفرضية التالية :

$$H_0 : \rho = 0$$

$$H_1 : \rho > 0$$

مثلاً كانت قيمة معامل ارتباط بيرسون موجبة 0.766 وتظهر نتيجة اختبار الفرضية أعلاه كما يلي :

Correlations

		LANG	MATH
LANG	Pearson Correlation	1.000	.776**
	Sig. (1-tailed)	.	.004
	N	10	10
MATH	Pearson Correlation	.776**	1.000
	Sig. (1-tailed)	.004	.
	N	10	10

** . Correlation is significant at the 0.01 level

حيث يتم احتساب إحصائية T نفسها في حالة الاختبار من طرفين وتساوي 3.48 ويمكن الحصول على P-Value المرافقة من خلال الأمر Compute → Transform ثم استعمال الدالة التجميعية لتوزيع T وكما يلي $T(3.48, 8) = 0.004$ لاحظ أن المقدار لا يضرب في 2 لأن الاختبار من طرف واحد. أن قيمة $P\text{-Value} = 0.004 < 0.05$ تشير الى أن الارتباط بين درجة امتحان اللغة و امتحان الرياضيات أكبر من الصفر بمستوى دلالة 5% (الاختلاف معنوي بمستوى دلالة 1% أيضاً) .

2. إذا كانت قيمة معامل الارتباط المقدرة سالبة يقوم البرنامج باختبار الفرضية التالية :

$$H_0: \rho = 0$$

$$H_1: \rho < 0$$

(10 - 3) الارتباط الجزئي Partial Correlation

يقيس معامل الارتباط الجزئي قوة العلاقة بين متغيرين بثبوت متغير ثالث . مثلاً قد نحصل على قيمة عالية لمعامل الارتباط البسيط للعلاقة بين أسعار اللحوم البيضاء واللحوم الحمراء فقد لا توجد علاقة فعلية بين المتغيرين ولكن كلا المتغيرين يتأثر بعامل ثالث هو المستوى العام للأسعار فإذا استبعدنا المستوى العام للأسعار (أو تثبيته) عند قياس العلاقة بين أسعار اللحوم البيضاء والحمراء فسيتم الحصول على قيمة أقل لمعامل الارتباط وهذا يعرف بالارتباط الجزئي . علماً أنه يمكن استبعاد أي عدد من المتغيرات عند قياس العلاقة بين ظاهرتين .

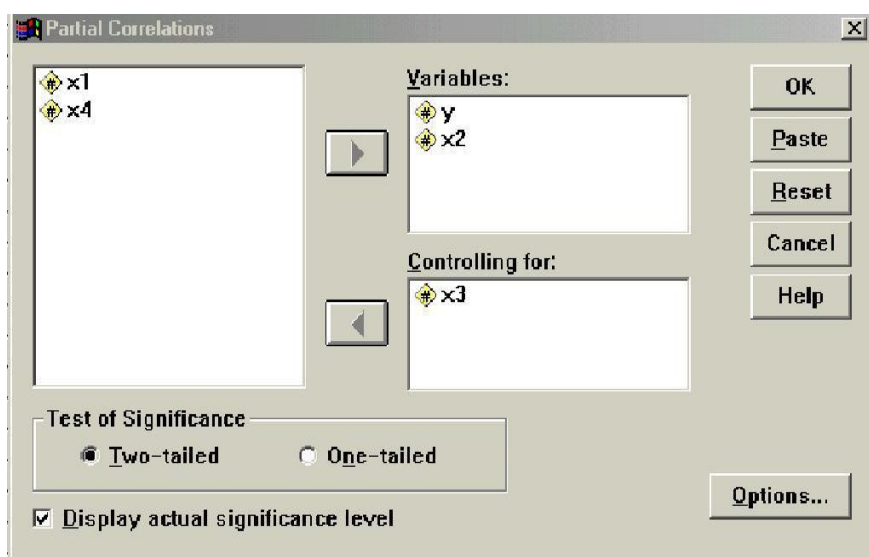
مثال 2:

للبينات الواردة في المثال 6 من البند (10 - 4 - 3) والتي يبلغ حجم العينة 13 يطلب ما يلي :

1. حساب معامل الارتباط الجزئي بين Y و X_2 بثبات X_3 .
2. حساب معامل الارتباط الجزئي بين Y و X_2 بثبات X_3 و X_4 .
3. أختبر معنوية معاملات الارتباط (من طرفين) بمستوى دلالة 5%.

الحل:

من شريط القوائم أختَر Partial → Correlate → Analyze فيظهر صندوق حوار Partial Correlation الذي نرتبه كما يلي للمطلوب الأول :



في خانة Variables يتم إدخال المتغيرات التي يراد حساب معامل الارتباط الجزئي لها ، وفي خانة Controlling for يتم إدخال المتغير (المتغيرات) الذي يراد استبعاد أثره .
 عند نقر زر Ok نحصل على النتيجة التالية .

--- PARTIAL CORRELATION COEFFICIENTS ---

Controlling for.. X3

	Y	X2
Y	1.0000 (0) P= .	.2851 (10) P= .369
X2	.2851 (10) P= .369	1.0000 (0) P= .

(Coefficient / (D.F.) / 2-tailed Significance)

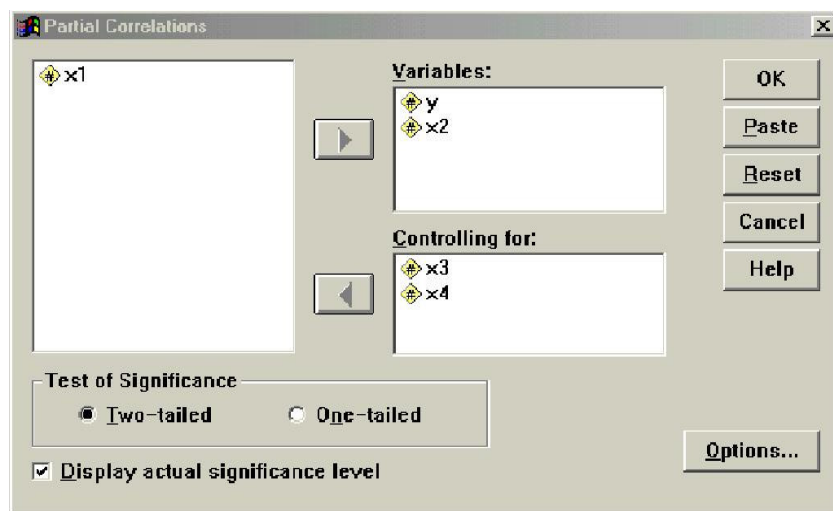
" . " is printed if a coefficient cannot be computed

حيث أن $r_{yx2.x3}=0.285$ لاختبار معنوية معامل الارتباط الجزئي (ذو طرفين) تستعمل نفس إحصائية T لاختبار معامل الارتباط البسيط. لاحظ أن الرقم 10 داخل قوسين والذي يظهر في الجدول أعلاه لا يمثل حجم العينة وإنما درجات الحرية اللازمة لاختبار T حيث $df = n - k = 13 - 3 = 10$ وأن T قد احتسبت كما يلي :

$$T = r \frac{\sqrt{n-k}}{\sqrt{1-r^2}} = (0.2851) \frac{\sqrt{13-3}}{\sqrt{1-0.2851^2}} = 0.941$$

وأن قيمة $P\text{-Value}=0.369 > 0.05$ ولهذا فإن معلمة المجتمع ρ لا تختلف معنوياً عن الصفر بمستوى دلالة 5% .

لحساب معامل الارتباط الجزئي بين Y و X_2 بثبات X_3 و X_4 (المطلوب الثاني) نكرر نفس الخطوات السابقة ونرتب صندوق حوار Partial Correlations على الشكل التالي :



وتظهر النتائج كما يلي :

--- PARTIAL CORRELATION COEFFICIENTS ---

	Controlling for..	
	X3	X4
	Y	X2
Y	1.0000 (0) P= .	.2338 (9) P= .489
X2	.2338 (9) P= .489	1.0000 (0) P= .

(Coefficient / (D.F.) / 2-tailed Significance)

" . " is printed if a coefficient cannot be computed

ومنه يظهر أن $r_{yx2.x3x4}$ لا يختلف معنوياً عن الصفر بمستوى دلالة 5% .

ملاحظة :

يمكن أستخراج معامل الارتباط البسيط لـ Pearson بنقر زر Options في صندوق حوار Partial Correlations ثم تأشير الخيار Zero Order Correlations .

(10 - 4) تحليل الانحدار Regression Analysis

أن نموذج الانحدار يعبر عن علاقة بين متغير معتمد Dependent Variable وبين واحد أو أكثر من المتغيرات المستقلة Independent Variables أو Regressors فإذا أحتوى النموذج على متغير مستقل واحد فيعرف بنموذج الانحدار البسيط Simple Regression model وإذا أحتوى أكثر من متغير مستقل فهو نموذج الانحدار المتعدد Multiple regression Model كما أن النموذج قد يكون خطياً Linear Model أو غير خطي Non Linear Model .

(10-4-1) نموذج الانحدار الخطي البسيط

بأخذ الصيغة العامة التالية : $Y = B_0 + B_1X + e$

حيث أن :

Y : المتغير المعتمد

X : المتغير المستقل

B_0 : الحد الثابت أو معلمة تقاطع خط الانحدار مع المحور الصادي Intersection Parameter .

$B_1 = \frac{\Delta Y}{\Delta X}$ Slope Parameter معلمة الميل

e : الخطأ العشوائي وهو الفرق بين القيمة الحقيقية Y والقيمة التقديرية $\hat{Y}$ ويعرف بالمتبقي residual حيث أن $e = Y - \hat{Y}$.

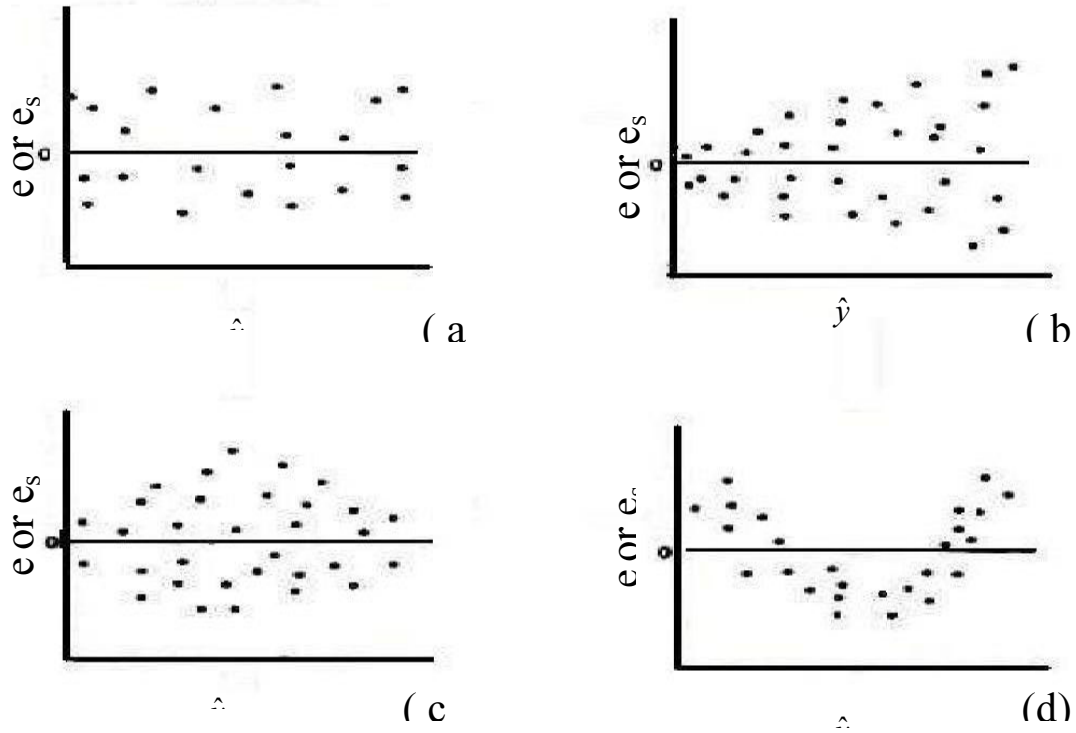
أن الطريقة المتبعة غالباً في تقدير معالم الانحدار B_0 و B_1 هي طريقة المربعات الصغرى

Least Squares Method (OLS) علماً أن نموذج الانحدار الخطي البسيط يجب أن يحقق الفرضيات التالية :

1. وجود علاقة خطية بين X و Y .
2. أن الأخطاء العشوائية تتوزع بمتوسط مساوي للصفر .
3. أن الأخطاء العشوائية لها تباين ثابت يساوي σ^2 (فرضية تجانس تباين الخطأ العشوائي Homoscedasticity) .
4. الأخطاء تتوزع طبيعياً وهذا الشرط ليس ضرورياً لتقدير المعالم بطريقة OLS ولكنه ضروري لاختبار الفرضيات المتعلقة بمعاملات الانحدار B_0 و B_1 .
5. عدم وجود ارتباط ذاتي Autocorrelation بين الأخطاء العشوائية .

يمكن التحقق من توفر فرضيات النموذج الخطي البسيط من خلال تخطيط Scatter plots بتمثيل $\hat{y}$ على المحور الأفقي (أو x) يقابله الخطأ العشوائي e أو الأخطاء المعيارية Standardized Residuals e_s على المحور الرأسي كما هو واضح في الشكل التالي .

أنماط الأخطاء العشوائية Residuals في نموذج الانحدار البسيط



(a) توفر فرضيات التحليل (عدم وجود مشكلة) .

(b) زيادة تباين الخطأ العشوائي بزيادة y .

(c) زيادة وتناقص في تباين الخطأ العشوائي (مشكلة عدم تجانس تباين الخطأ العشوائي) .

(d) عدم ملائمة العلاقة الخطية (يتوجب استعمال نماذج أخرى مثلاً نموذج من الدرجة الثانية) .

مثال 3 :

البيانات التالية تمثل العمر X وضغط الدم Y (ملم زئبق) لعينة مكونة من 10 أشخاص وقد تم

إدخال البيانات في ورقة Data Editor وكما يلي :

Obs.	X	Y
1	35	112
2	40	128
3	38	130
4	44	138
5	67	158
6	64	162
7	59	140
8	69	175
9	25	125
10	50	142

يطلب مايلي :

1. استخراج معادلة انحدار Y/X مفترضاً العلاقة الخطية واختبار معنوية معالم النموذج .

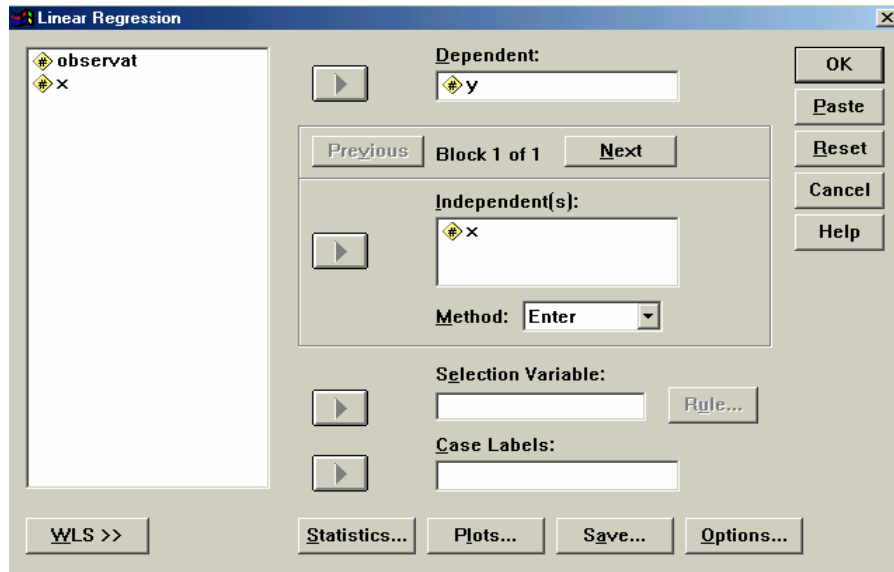
2. أستخرج فترة ثقة 95% لكل من معلمتي الانحدار B_0 و B_1 .

3. استخراج جدول تحليل التباين ANOVA .
4. أختبر جودة توفيق النموذج الخطي (باستعمال معامل التحديد R^2) مع تحليل الأخطاء العشوائية بالرسم البياني .
5. أختبر التوزيع الطبيعي للأخطاء العشوائية بيانياً .

الحل :

لتنفيذ المطالبات أعلاه نتبع الخطوات التالية :

□ من شريط القوائم اختر Linear Regression → Regression → Analyze فيظهر صندوق حوار Linear Regression الذي نرتبه كما يلي :



حيث أن :

Dependent: يمثل المتغير المعتمد .

Independent : المتغير (أو المتغيرات المستقلة). يمكن إدخال أكثر من مجموعة من المتغيرات المستقلة كل مجموعة تدخل ضمن Block له رقم تسلسلي ويمكن الانتقال من Block الى آخر بالزرين Previous و Next. فإذا كان لدينا نموذجين لأحدهما متغير مستقل X والنموذج الآخر له متغير مستقل Z مع متغير معتمد واحد هو Y لكلا النموذجين في هذه الحالة يتم إدخال المتغير X في Block1 والمتغير Z في Block2 .

Method : نوع الطريقة المستخدمة في الانحدار (الطريقة الاعتيادية هي Enter) .
Selection Variable : يستعمل في تحديد التحليل لمجموعة معينة من الحالات التي لها قيمة معينة لمتغير الاختيار (مثلاً اقتصار نموذج الانحدار على الحالات التي تكون فيها قيمة المتغير Observat أكبر من 5) يتم التحديد بواسطة الزر Rule .

Case Labels : متغير تستخدم قيمه كعناوين لنقاط شكل الانتشار Scatterplots .

□ انقر زر Statistics فيظهر صندوق حوار Statistics الذي نرتبه كما يلي :

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 standard deviations

☐ All cases

Continue

Cancel

Help

وقد تم تأشير الخيارات التالية :

Estimate: لتقدير معالم نموذج الانحدار واختبارات t المرافقة.

Confidence Interval: لتقدير فترة ثقة 95% لكل من معلمتي الانحدار.

Model Fit: لعرض R^2 و ANOVA .

☐ انقر زر Plots فيظهر صندوق حوار Plots أشر الخيار Normal Probability Plot لاختبار

التوزيع الطبيعي للأخطاء العشوائية المعيارية (المطلوب الخامس) .

☐ انقر زر Save فيظهر صندوق حوار Save الذي نرتبه كما يلي :

Linear Regression: Save

Predicted Values

☒ Unstandardized

☐ Standardized

☐ Adjusted

☐ S.E. of mean predictions

Distances

☐ Mahalanobis

☐ Cook's

☐ Leverage values

Prediction Intervals

☐ Mean ☐ Individual

Confidence Interval: %

Save to New File

☐ Coefficient statistics

Export model information to XML file

Residuals

☐ Unstandardized

☒ Standardized

☐ Studentized

☐ Deleted

☐ Studentized deleted

Influence Statistics

☐ DfBeta(s)

☐ Standardized DfBeta(s)

☐ DfFit

☐ Standardized DfFit

☐ Covariance ratio

Continue

Cancel

Help

لاحظ أننا اخترنا Unstandardized Predicted Values أي $\hat{y}$ وكذلك Standardized Residuals أي e_s سيتم استعمال هذين المتغيرين لرسم Scatterplots للشطر الثاني من المطلوب الرابع، علماً أنه سيتم إضافة هذين المتغيرين (عند تمشية البرنامج) الى ورقة Data Editor الى جانب متغيرات X و Y و Observat حيث يضاف المتغير Unstandardized Predicted Values بأسم Pre_1 ويضاف متغير Standardized Residuals بأسم Zre_1 .

يمكن الاختيار بين تضمين الحد الثابت في النموذج او حذفه من خلال خيارات الزر Options في صندوق حوار Linear Regression.

□ عند نقر زر OK في صندوق حوار Linear Regression نحصل على النتائج التالية :

Coefficients ^a								
		Unstandardize d Coefficients		Standa rized Coefficients	t	Sig.	95% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	85.044	9.970		8.530	.000027	62.052	108.036
	X	1.140	.195	.900	5.846	.00038	.690	1.589

a. Dependent Variable: Y

من خلال الجدول أعلاه يمكن كتابة النموذج كما يلي :

$$\hat{y} = 85.044 + 1.140x$$

$$(9.97) \quad (0.195)$$

أن معلمة الميل تشير الى أن زيادة العمر سنة واحدة يؤدي الى زيادة ضغط الدم بمقدار 1.140 ملم زئبق. الأرقام داخل الأقواس تمثل الخطأ المعياري للمعلمة المقابلة .

يستعمل اختبار T لاختبار الفرضية التالية لمعلمة الميل B_1 :

$$H_0 : B_1 = 0 \quad \text{فرضية العدم}$$

$$H_1 : B_1 \neq 0 \quad \text{الفرضية البديلة}$$

كما يستعمل اختبار T لاختبار الفرضية التالية لمعلمة التقاطع (الحد الثابت) B_0 :

$$H_0 : B_0 = 0 \quad \text{فرضية العدم}$$

$$H_1 : B_0 \neq 0 \quad \text{الفرضية البديلة}$$

نستخدم قيمة P-Value المرافقة لإحصائية T للمعلمة في الاختبار كما يلي :

إذا كانت $P\text{-Value} < 0.05$ نرفض فرضية العدم بمستوى دلالة 5% .

إذا كانت $P\text{-Value} < 0.01$ نرفض فرضية العدم بمستوى دلالة 1% .

عكس هذا نقبل فرضية العدم .

أن P-value لمعلمة الميل تساوي 0.00038 وهي أقل من 0.01 و أن P-value لمعلمة الحد الثابت تساوي 0.000027 وهي أقل من 0.01 ولهذا نرفض فرضيتنا العدم لكل من المعلمتين أي أن كلا المعلمتين تختلف جوهرياً عن الصفر أن ظهور معلمة الميل معنوية يعكس أهمية متغير العمر في النموذج .

أما معلمة Standardized Coefficient ويشار لها بـ Beta فهي معلمة الميل للنموذج المقدر باستعمال القيم المعيارية $(X - \bar{X})/S$ لكل من المتغير المستقل والمعتمد بدل القيم الأصلية ولا يحتوي هذا النموذج على معلمة تقاطع B_0 وكما يلي $\hat{y}^* = \beta x^*$ حيث أن $\hat{y}^*$ و x^* متغيرات معيارية . يمكن كتابة فترة ثقة 95% للحد الثابت كما يلي :

$$\Pr(62.052 \leq B_0 \leq 108.036) = 95\%$$

حيث أن Pr يمثل الاحتمال كما يمكن كتابة فترة ثقة 95% لمعلمة الميل كما يلي :

$$\Pr(.690 \leq B_0 \leq 1.589) = 95\%$$

الجدول التالي يعرف بجدول تحليل التباين ANOVA ويشتمل على إحصائية F لاختبار نفس الفرضية الخاصة بمعلمة الميل B_1 وهذا الاختبار مكافئ تماماً لاختبار T لمعلمة الميل (لاحظ أن قيمة P-Value متساوية لكلا الاختبارين) علماً أن $F=T^2$.

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2661.050	1	2661.050	34.174	.00038 ^a
	Residual	622.950	8	77.869		
	Total	3284.000	9			

a. Predictors: (Constant), X

b. Dependent Variable: Y

الجدول التالي يتضمن أهم مؤشر لنموذج الانحدار وهو معامل التحديد Coefficient Of Determination ويرمز له R^2 ويعتبر مقياساً لجودة توفيق النموذج . ويحتسب من جدول تحليل التباين كما يلي :

$$R^2 = \frac{\text{Explained Variations}}{\text{Total Variations}} = \frac{SSR}{SST} = \frac{2661.05}{3284} = 0.81 \quad 0 \leq R^2 \leq 1$$

وتفسير ذلك أن 81% من التباينات (الانحرافات الكلية في قيم المتغير Y) تفسرها العلاقة الخطية أي نموذج الانحدار وأن 19% من التباينات ترجع الى عوامل عشوائية كأن تكون هناك متغيرات مهمة لم تضمن في النموذج . وعلى العموم كلما اقتربت قيمة R^2 من 100% دل ذلك على جودة توفيق النموذج . هذا وأن $r = \sqrt{R^2}$ حيث أن r معامل الارتباط الخطي البسيط لبيرسون وأن إشارة r هي نفس إشارة معلمة الميل .

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.900 ^a	.810	.787	8.82

a. Predictors: (Constant), X

يتصف معامل التحديد بأنه لو أضيف متغير مستقل للنموذج فأن قيمته سترتفع حتى لو لم تكن هناك أهمية للمتغير المستقل في النموذج حيث أن إضافة متغير مستقل الى نموذج الانحدار يؤدي الى زيادة R^2

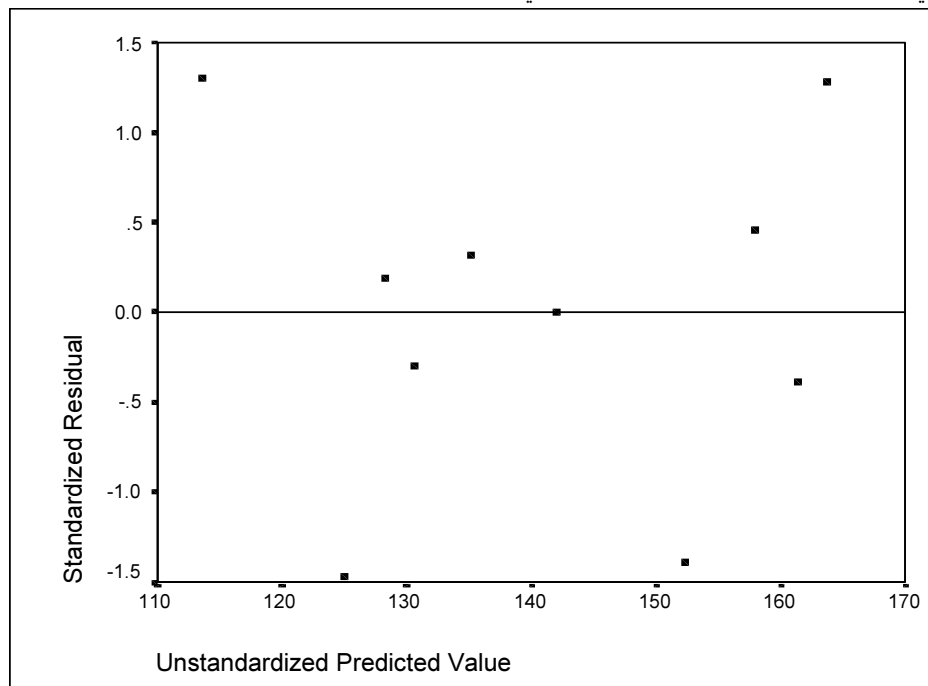
بسبب زيادة مجموع المربعات العائدة للانحدار SSR مع ثبات مجموع المربعات الكلية SST ولهذا يتم احتساب معامل التحديد المصحح Adjusted R Square الذي يأخذ بالاعتبار النقصان الحاصل في درجات الحرية وقيمتها دائماً أقل من قيمة معامل التحديد (غير المصحح) في هذا المثال 79% ولهذا يمكن القول أن النموذج جيد التوفيق.

أما الخطأ المعياري للتقدير Standard Error of Estimate فيقيس تشتت القيم المشاهدة عن خط الانحدار وأن الحصول على قيمة صغيرة لهذا المؤشر يعني صغر الأخطاء العشوائية وبالتالي جودة تمثيل خط الانحدار لنقاط شكل الانتشار .

لتحليل الأخطاء العشوائية بيانياً نكون شكل الانتشار Scatterplots بتمثيل القيم التقديرية $\hat{y}$ على المحور الأفقي والأخطاء المعيارية e_s على المحور العمودي وكما يلي (راجع فصل المخططات البيانية) :

□ من شريط القوائم اختر Simple → Scatter → Graphs فيظهر صندوق حوار Scatterplots .

- انقر المتغير Zre_1 لإدخاله الى خانة Y في صندوق حوار Scatter .
- انقر المتغير Pre_1 لإدخاله الى خانة X .
- انقر زر OK فيظهر المخطط في شاشة SPSS Viewer .
- انقر المخطط مرتين للانتقال الى شاشة SPSS Chart Editor .
- من شريط القوائم اختر Reference Line → Chart لإضافة Reference Line الى المحور الرأسي حول الصفر . فيظهر المخطط كما يلي :

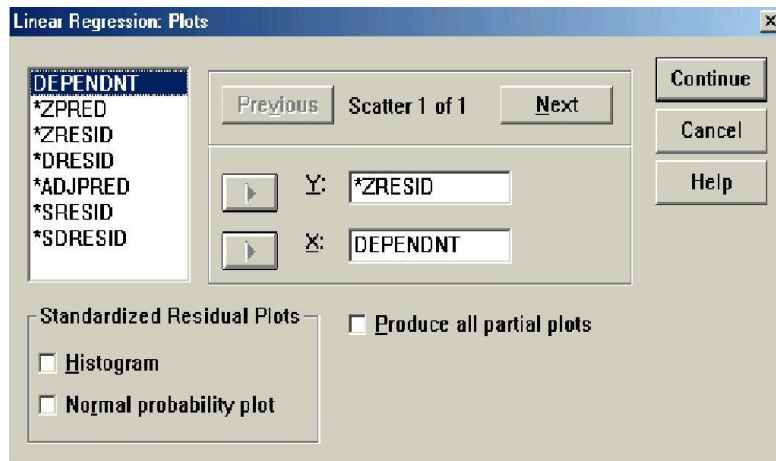


نلاحظ أن النقاط تتوزع بشكل شريط أفقي متساوي حول الصفر مما يدل على توفر فرضيات التحليل بصورة عامة حيث لا يعاني النموذج من مشكلة عدم تجانس تباين الخطأ العشوائي ولا توجد حاجة لاستخدام علاقة من درجات أعلى .

ملاحظة :

يمكن تحليل الأخطاء العشوائية بيانياً بطريقة مشابهة بتمثيل القيم الحقيقية للمتغير المعتمد y على المحور الأفقي والأخطاء المعيارية e_y على المحور العمودي حيث يمكن التوصل الى مخطط مقارب للمخطط السابق وكما يلي :

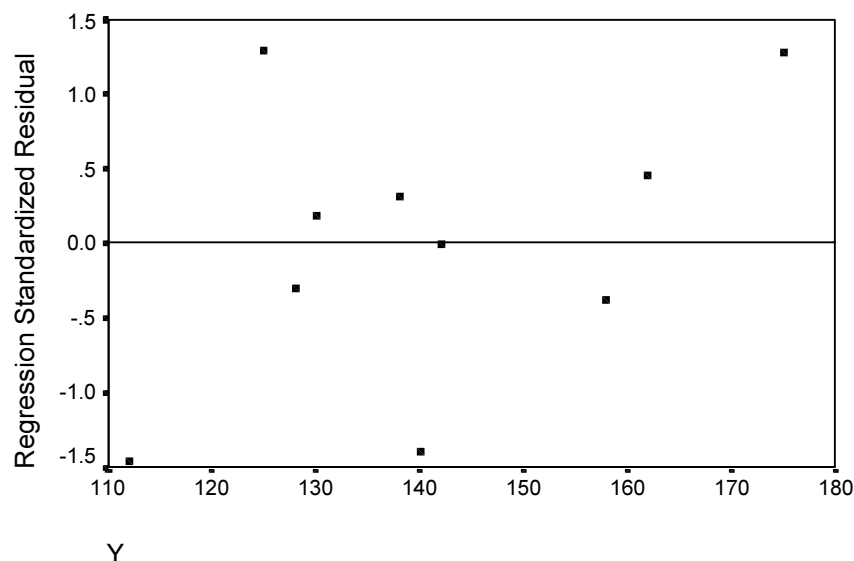
□ أنقر زر Plots في صندوق حوار Linear Regression فيظهر صندوق حوار Plots الذي نرتبه كما يلي :



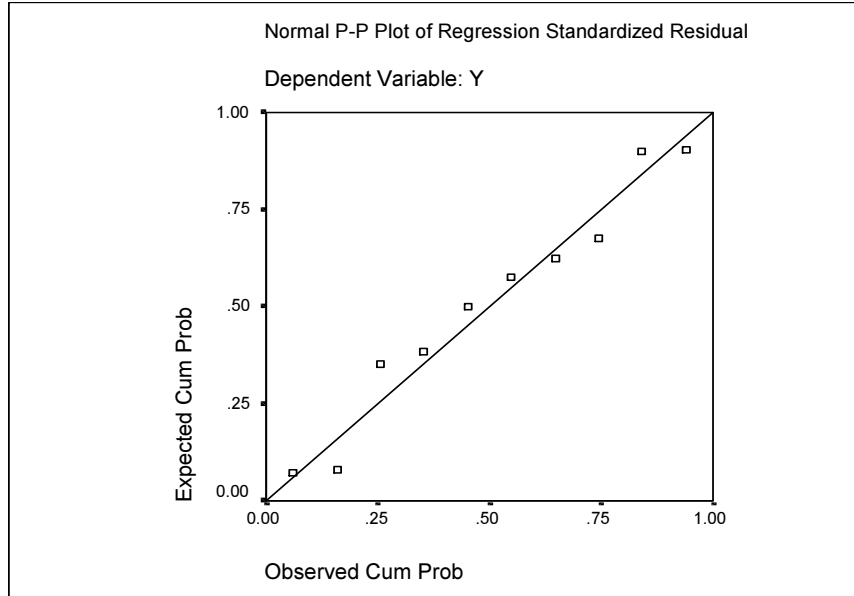
لقد قمنا بنقل المتغير DEPENDENT أي المتغير المعتمد من قائمة المتغيرات في جهة اليسار (هذه المتغيرات تحتسب تلقائياً) الى خانة الإحداثي الأفقي X كما نقلنا متغير الأخطاء (البواقي) المعيارية ZRESID الى خانة الإحداثي الرأسى Y وعند نقر زر Continue في هذا الصندوق وزر OK في صندوق Linear Regression يتم عرض مخطط لشكل الانتشار التالي والذي يقودنا الى نفس الاستنتاج الذي توصلنا اليه من خلال المخطط السابق:

Scatterplot

Dependent Variable: Y



أن اختبار التوزيع الطبيعي للأخطاء العشوائية (المطلوب الخامس) يمكن أن يتم بطريقتين : الأولى من خلال رصد الأخطاء المعيارية فإذا وقعت 95% من الأخطاء ضمن المدى (-2,2) فإن الأخطاء تتوزع طبيعياً ، من المخطط الأخير نلاحظ أن الأخطاء المعيارية لا تتعدى المدى (-1.5,1.5) ومنه نستدل على أن الأخطاء تتوزع طبيعياً . أما الطريقة الثانية فتتمثل في عرض مخطط Normal probability Plot الذي أشرناه في صندوق حوار Plot حيث يعرض المخطط التالي عند تنفيذ البرنامج :



نلاحظ أن معظم النقاط تقريباً تتجمع قرب الخط المستقيم وهذا يدل على التوزيع الطبيعي للأخطاء العشوائية .

(10 - 4 - 2) طريقة المربعات الصغرى الموزونة *Weighted Least Squares Method*

تستخدم هذه الطريقة في حالة عدم تحقق فرضية تجانس تباين الخطأ العشوائي Homoscedasticity وغالباً ما تظهر هذه المشكلة في بيانات المقطع العرضي Cross-Section Data ونادراً ما تظهر في بيانات السلاسل الزمنية .

فقد لوحظ مثلاً في بحوث ميزانية الأسرة أن تباين الاستهلاك C يتزايد مع تزايد الدخل المتاح Y وعليه فإن تباين الخطأ لا يكون ثابتاً وفي الغالب يكون متناسباً مع مربع الدخل Y حيث تكون صيغة تباين الخطأ العشوائي كما يلي $\text{var } e = \sigma^2 Y^2$ فإذا اعتبرنا المقدار $W = 1/Y^2$ يمثل الوزن Weight فإذا كان نموذج الانحدار يأخذ الصيغة التالية :

$$C = \beta_0 + \beta_1 Y + e$$

يكون بالإمكان تثبيت تباين الخطأ العشوائي بترجيح طرفي المعادلة بالمقدار $\sqrt{W} = 1/Y$ حيث نقوم بتقدير النموذج التالي :

$$\frac{C}{Y} = \frac{B_0}{Y} + B_1 + \frac{e}{Y}$$

أن تباين حد الخطأ العشوائي للنموذج الأخير يحسب كما يلي
أي أننا حصلنا على تباين ثابت للخطأ العشوائي . ويجب $\text{var } \frac{e}{Y} = \frac{1}{Y^2} \text{var } e = \frac{1}{Y^2} \sigma^2 Y^2 = \sigma^2$

ملاحظة أن B1 قد أصبحت حداً ثابتاً في النموذج الجديد وأن B0 قد أصبحت معلمة الميل .مع العلم أن

برنامج SPSS يقوم بحساب المعالم من النموذج الأخير ثم تستخدم هذه المعالم في النموذج الأصلي لاستخراج البواقي وجدول تحليل التباين وفي حساب R^2 , SEE , وغيرها من التقديرات . أن هذه الطريقة تعرف أيضاً بطريقة المربعات الصغرى المعممة GLS .

مثال 4 : (على طريقة المربعات الصغرى الموزونة)

الجدول التالي يمثل الاستهلاك c والدخل المتاح y لعدد من الأسر عددها 30 أسرة².وقد تم إدخال

المتغيرين المذكورين الى شاشة Data editor لبرنامج SPSS كما يلي (المتغير w أحتسب فيما بعد) :

c	y	w
10600	12000	6.9444E-09
10800	12000	6.9444E-09
11100	12000	6.9444E-09
11400	13000	5.9172E-09
11700	13000	5.9172E-09
12100	13000	5.9172E-09
12300	14000	5.1020E-09
12600	14000	5.1020E-09
13200	14000	5.1020E-09
13000	15000	4.4444E-09
13300	15000	4.4444E-09
13600	15000	4.4444E-09
13800	16000	3.9063E-09
14000	16000	3.9063E-09
14200	16000	3.9063E-09
14400	17000	3.4602E-09
14900	17000	3.4602E-09
15300	17000	3.4602E-09
15000	18000	3.0864E-09
15700	18000	3.0864E-09
16400	18000	3.0864E-09
15900	19000	2.7701E-09
16500	19000	2.7701E-09
16900	19000	2.7701E-09
16900	20000	2.5000E-09
17500	20000	2.5000E-09
18100	20000	2.5000E-09
17200	21000	2.2676E-09
17800	21000	2.2676E-09
18500	21000	2.2676E-09

يطلب ما يلي :

1. استخراج معادلة انحدار C على Y بطريقة المربعات الصغرى الاعتيادية OLS ثم اختبار تجانس تباين الخطأ العشوائي من الرسم البياني .

2. بافتراض أن تباين الخطأ العشوائي غير متجانس ويرتبط مع Y بالعلاقة التالية $\text{var} e = \sigma^2 Y^2$ أستخدم طريقة المربعات الصغرى الموزونة في تقدير نموذج الانحدار .

² دومنيك سالفاتور ، الإحصاء والاقتصاد القياسي ، ملخصات شوم، دار ماكروهيل ، 1982 ، ص . 217 .

الحل :

1. يمكن تقدير نموذج الانحدار الخطي البسيط بطريقة OLS بنفس الخطوات المتبعة في المثال السابق وقد تم التوصل الى النموذج التالي :

$$\hat{C} = 1408.0 + 0.788Y_d$$

$$R^2 = 0.97 \quad (0.27) \quad (449.6)$$

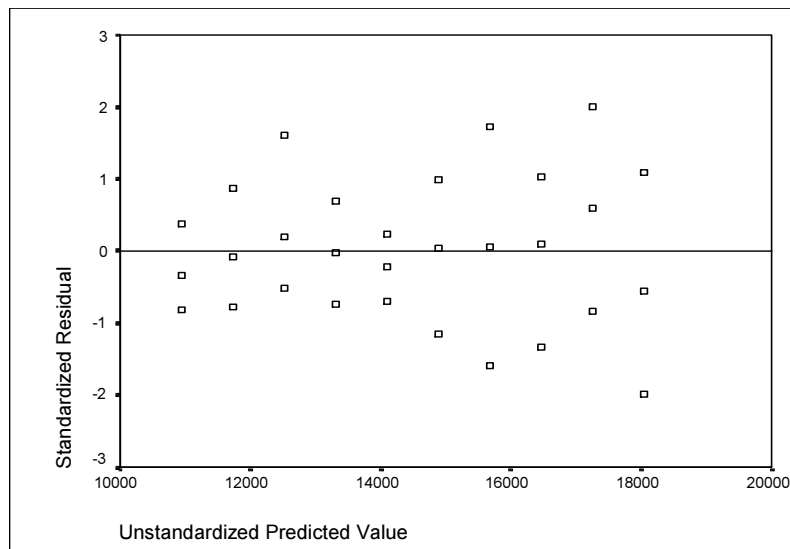
الأرقام داخل الأقواس تمثل الخطأ المعياري للمعالم . لاختبار وجود مشكلة عدم تجانس تباين الخطأ العشوائي بياناً نقوم بتمثيل القيم التنبؤية للمتغير المعتمد $\hat{C}$ على المحور الأفقي لشكل الانتشار والأخطاء المعيارية على المحور العمودي Standardized Residuals ، لتنفيذ ذلك نتبع الخطوات التالية:

□ اختر Linear → Regression → Analyze من شريط القوائم لفتح صندوق Linear Regression ثم انقر زر Save في هذا الصندوق لفتح صندوق Save في هذا الصندوق أشر الخيار Unstandardized في الحقل Predicted Values لكي يعرض البرنامج القيم التقديرية للمتغير المعتمد $\hat{C}$ في شاشة Data editor (تظهر في الشاشة بأسم Pre-1) ثم أشر الخيار Standardized Residuals في الحقل Residuals لكي يعرض البرنامج الأخطاء المعيارية (تظهر بأسم Zre-1) .

□ من شريط القوائم اختر Simple → Scatter → Graphs فيظهر صندوق حوار ScatterPlots يتم ترتيب هذا الصندوق كما يلي :

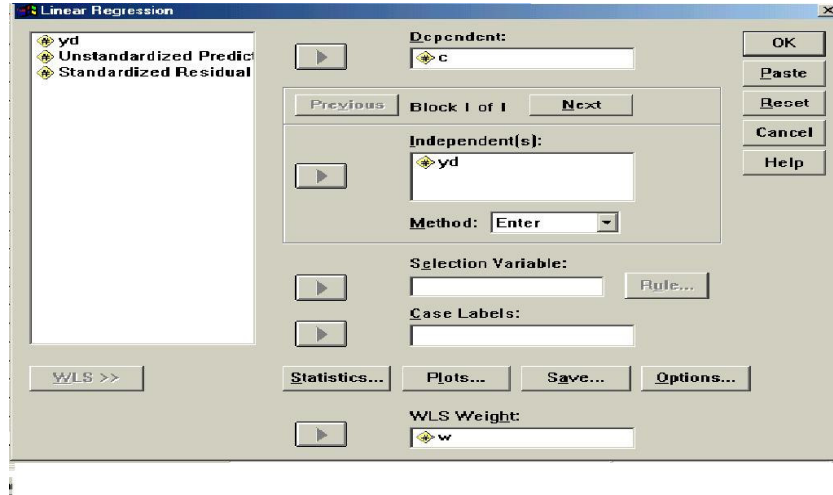
- أنقل المتغير Zre-1 الى حقل Y-axis .
- أنقل المتغير Pre-1 الى حقل X-axis .
- انقر زر OK .

فيظهر المخطط التالي بعد إضافة Reference Line إليه وكما يلي :



نلاحظ أن نقاط شكل الانتشار لا تتوزع بانتظام حول الصفر وأن هناك اتجاهها عاماً في زيادة تباين الخطأ العشوائي كلما زادت $\hat{Y}$ مما يشير الى وجود مشكلة عدم تجانس تباين الخطأ ويتوجب معالجتها .

2. بافتراض أن تباين الخطأ يرتبط مع المتغير المستقل Y بالعلاقة التالية $\text{var } e = \sigma^2 Y^2$ فلكي نطبق طريقة المربعات الصغرى الموزونة نحدد الوزن وهو مقلوب التباين $W = 1/y^2$ ونقوم بحساب هذا المتغير بالأمر Compute → Transform وإضافته الى شاشة Data Editor كما في الجدول أعلاه ثم نطبق نفس خطوات إجراء الانحدار الخطي حيث يظهر صندوق حوار Linear Regression بعد ترتيبه كما يلي :



لاحظ أننا قمنا بإدخال متغير الوزن W في خانة WLS Weight بعد نقر زر WLS .
 عند نقر زر OK يظهر المخرج التالي :

Coefficients^{a,b}

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1421.278	395.496		3.594	.001
	YD	.792	.025	.986	31.511	.000

a. Dependent Variable: C

b. Weighted Least Squares Regression - Weighted by W

ويمكن كتابة المعادلة المقدرة بطريقة المربعات الصغرى الموزونة كما يلي :

$$\hat{C} = 1421.278 + 0.792Y_d \quad R^2 = 0.97$$

(395.496) (0.25)

حاول أن ترسم شكل الانتشار لعلاقة القيم التقديرية بالأخطاء المعيارية للتأكد من أن البيانات تتوزع بشكل شريط أفقي حول الصفر وبالتالي ثبات تباين الخطأ العشوائي.

(10- 4- 3) نموذج الانحدار الخطي المتعدد

يأخذ النموذج الخطي العام الصيغة التالية :

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + e$$

حيث أن

β_0 : يمثل الحد الثابت .

$\beta_1, \beta_2, \dots, \beta_k$: معاملات الانحدار الجزئية Partial regression Coefficients أو الميول الجزئية .

e : الخطأ العشوائي .

حيث يكون عدد معالم النموذج الخطي العام هو $P = K + 1$ (K يمثل عدد المتغيرات المستقلة في النموذج) .

أن فرضيات النموذج الخطي المتعدد هي نفسها فرضيات النموذج البسيط يضاف الى ذلك فرضية عدم وجود ارتباط خطي متعدد بين المتغيرات المستقلة (Multicollinearity) .

مثال 5 :

الجدول التالي يتضمن دخل الفرد y (ألف دولار) مع نسبة القوة العاملة في الزراعة x_1 (نسبة مئوية) ومتوسط سنوات التعليم للسكان x_2 لـ 15 دولة في سنة 1981³. وقد أدخلت البيانات في شاشة Data Editor كما في الجدول التالي :

y	x_1	x_2
6	9	8
8	10	13
8	8	11
7	7	10
7	10	12
12	4	16
9	5	10
8	5	10
9	6	12
10	8	14
10	7	12
11	4	16
9	9	14
10	5	10
11	8	12

المطلوب :

1. حساب معادلة انحدار y على x_1 و x_2 وتفسير النتائج .
 2. تكوين جدول تحليل التباين ANOVA واختبار معنوية المعاملات .
 3. اختبار مشكلة الارتباط الذاتي المتسلسل باستخدام إحصائية DW .
- لتنفيذ المطالبات المذكورة نتبع الخطوات التالية :
- من شريط القوائم اختر Linear → Regression → Analyze فيظهر صندوق حوار Linear Regression الذي يرتب كالتالي :
- أنقر المتغير Y وادخله الى خانة Dependent .
 - أشر المتغيرين x_1, x_2 وانقلهما الى خانة Independent .
 - في خانة Method تأكد أن نوع طريقة الانحدار هي الطريقة الاعتيادية Enter .

³ دومينيك سالفاتور ، الإحصاء والاقتصاد القياسي ، دار ماكروهيل ، 1982 ، ص 172 .

أقر Statistics في صندوق حوار Linear Regression فيظهر صندوق حوار Statistics حيث

نقوم بتأشير كل مما يلي :

Estimate : لتقدير معالم النموذج والإحصاءات المرافقة .

Model Fit : لتقدير R^2 وANOVA.

Durbin - Watson : لحساب إحصائية DW .

عند نقر زر OK في صندوق حوار Linear regression يتم عرض النتائج التالية :

Model Summary^a

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.833 ^a	.693	.642	1.01	.946

a. Predictors: (Constant), X2, X1

b. Dependent Variable: Y

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	6.203	1.862		3.331	.006
	X1	-.376	.133	-.461	-2.834	.015
	X2	.453	.120	.615	3.786	.003

a. Dependent Variable: Y

يمكن معادلة كتابة نموذج الانحدار المتعدد كما يلي :

$$\hat{y} = 6.203 - 0.376X_1 + 0.453X_2 \quad \bar{R}^2 = 0.64$$

(1.862) (0.133) (0.120)

أن معلمة المتغير X1 تشير الى أن زيادة نسبة القوى العاملة في الزراعة بمقدار 1% من أجمالي القوى العاملة يؤدي الى نقصان في دخل الفرد بمقدار 376 دولار بافتراض ثبات المتغير X2 كما أن زيادة متوسط سنوات التعليم سنة واحدة يؤدي الى زيادة دخل الفرد بمقدار 453 دولار بافتراض ثبات المتغير X1 ، كما أن قيمة $\bar{R}^2$ تشير إلى جودة توفيق متوسطة للنموذج .

أن الغرض من حساب جدول تحليل التباين هو تحليل مجموع مربعات الانحرافات الكلية SST لقيم المتغير المعتمد $\sum (y - \bar{y})^2$ الى مجموع المربعات العائدة للانحدار SSR ومجموع مربعات الخطأ SSE . كما يتم احتساب إحصائية F التي يستفاد منها في اختبار الفرضية التالية :

$$H_0 : \beta_1 = \beta_2 = 0$$

$$H_1 : \beta_1 \neq \beta_2 \neq 0$$

لاحظ أن اختبار F لا يشمل معلمة التقاطع B₀ . لقد تم الحصول على جدول تحليل التباين التالي :

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	(SSR)27.728	(p-1) 2	(MSR) 13.864	13.557	.001 ^a
	Residual	(SSE)12.272	(n-p) 12	(MSE) 1.023	F = MSR/MSE	
	Total	(SST)40	(n-1) 14			

a. Predictors: (Constant), X2, X1

b. Dependent Variable: Y

حيث أن $P = 3$ ويمثل عدد المعالم و n يمثل حجم العينة. أن قيمة $P\text{-Value} = 0.001 < 0.05$ تدعونا الى رفض فرضية العدم بمستوى دلالة 5% أي أن الانحدار معنوي أو أن المتغيرين المستقلين مجتمعين لهما تأثير معنوي على الانحدار أو أن واحدة على الأقل من معلمتي الانحدار B_1 و B_2 تختلف معنوياً عن الصفر .

لمعرفة تأثير كل متغير مستقل على المتغير المعتمد بصورة انفرادية نلجأ الى اختبار t (جدول Coefficients في الأعلى) ومنه يتضح معنوية معلمتي الميل لكل من X_1 و X_2 بمستوى دلالة 5% وعليه فكل المتغيرين مؤثرين ويوصى بإبقائهما في نموذج الانحدار .

عند اختبار وجود الارتباط الذاتي للأخطاء العشوائية بمستوى دلالة 5% وبدرجة حرية $n=15$ و $p=3$ فإن إحصائية DW المحسوبة تشير الى وجود ارتباط ذاتي موجب في الأخطاء العشوائية .

مثال 6: (اختبار وجود مشكلة التعدد الخطي / مع اختيار أفضل نموذج انحدار)

الجدول التالي⁴ يتضمن المتغير المعتمد Y والمتغيرات المستقلة X_1, X_2, X_3 والتي أدخلت في ورقة

Data Editor كما يلي :

Y	X1	X2	X3	X4
43	5	3	18	12
63	9	5	27	9
71	10	7	34	11
61	8	4	24	10
81	11	6	33	6
44	12	5	22	8
58	9	4	28	9
71	7	7	32	7
72	8	5	23	8
67	13	8	20	5
64	4	5	21	4
69	10	9	36	10
68	11	10	30	11

1. أوجد معادلة الانحدار الخطي المتعدد .

2. أختبر وجود مشكلة التعدد الخطي Multicollinearity بين المتغيرات المستقلة .

3. أوجد أفضل معادلة انحدار بأسلوب Stepwise Regression .

1. لتنفيذ المطلوبين الأول والثاني من المثال نتبع الخطوات التالية :

بغداد، 1988، ص277.

⁴ د.أموري هادي كاظم ، ومحمد مناجد الدليمي ، مقدمة في تحليل الانحدار الخطي، جامعة

□ من شريط القوائم اختر Linear → Regression → Analyze فيظهر صندوق حوار Linear Regression الذي يرتب كالتالي :

- أشر المتغير Y وادخله الى خانة Dependent .
- أشر المتغيرات X1,X2,X3 وانقلها الى خانة Independent .
- في خانة Method تأكد أن نوع طريقة الانحدار هي الطريقة الاعتيادية Enter .

□ أنقر زر Statistics في صندوق حوار Linear Regression فيظهر صندوق حوار Statistics حيث نقوم بتأشير كل مما يلي

Estimate : لتقدير معالم النموذج والإحصاءات المرافقة .

Model Fit : لتقدير R^2 وANOVA.

Part & Partial Correlations : لتقدير معاملات الارتباط .

Collinearity Diagnostics : لتشخيص مشكلة التعدد الخطي .

◀ عند نقر زر OK في صندوق حوار Linear regression يتم عرض النتائج التالية :

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.810 ^a	.656	.484	7.72

a. Predictors: (Constant), X4, X1, X3, X2

Coefficients^a

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
	B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1 (Constant)	47.06	13.0		3.611	.007					
X1	-.430	1.010	-.104	-.425	.682	.210	-.149	-.088	.713	1.403
X2	1.199	1.486	.232	.807	.443	.540	.274	.167	.520	1.925
X3	1.153	.476	.631	2.423	.042	.634	.651	.502	.633	1.580
X4	-2.038	.950	-.462	-2.1	.064	-.331	-.604	-.445	.928	1.077

a. Dependent Variable: Y

يمكن كتابة معادلة نموذج الانحدار المتعدد كما يلي :

$$\hat{y} = 47.06 - 0.43X_1 + 1.199X_2 + 1.153X_3 - 2.038X_4 \quad \bar{R}^2 = 0.48$$

(13) (1.01) (1.486) (0.476) (0.950)

أن معلمة المتغير X1 تشير الى أن زيادة قدرها وحدة واحدة في قيم المتغير ينشأ عنها نقصان المتغير المعتمد بمقدار 0.430 وحدة بافتراض ثبات المتغيرين X2 و X3 .وبنفس الطريقة يتم تفسير بقية معالم النموذج .

يلاحظ ومن خلال قيمة P-value المرافقة لإحصائية t أن معلمة المتغير X_3 فقط قد ظهرت معنوية بمستوى دلالة 5% (بالإضافة الى معلمة الحد الثابت). كما أن قيمة معامل التحديد المصحح المنخفضة نسبياً تشير الى أن النموذج لا يعبر عن العلاقة الخطية بصورة جيدة .

لقد تم أستخراج ثلاثة أنواع من الارتباطات وكما يلي :

Zero-order Correlation : معامل الارتباط البسيط لبيرسون بين المتغير المستقل والمتغير المعتمد .
Partial Correlation : معامل الارتباط الجزئي بين المتغير المعتمد والمتغير المستقل (بثبات المتغيرات المستقلة الأخرى) .

Part Correlation : معامل الارتباط الجزء بين المتغير المعتمد والمتغير المستقل باستبعاد أثر المتغيرات المستقلة عن المتغير المستقل فقط .

أن نموذج الانحدار المتعدد يعتمد الارتباطات الجزئية بين كل من المتغيرات المستقلة والمتغير المعتمد ونلاحظ أن X_3 له أعلى معامل ارتباط جزئي مع المتغير المعتمد وعليه يكون له أعلى قيمة لإحصائية t يليه X_4 في حين يعتمد نموذج الانحدار البسيط معامل الارتباط البسيط بين المتغير المستقل والمتغير المعتمد .

لغرض تشخيص مشكلة الارتباط الخطي المتعدد يتم في البداية حساب المعامل Tolerance لكل من المتغيرات المستقلة حيث ان $Tolerance = 1 - R_{X_i, others}^2$ حيث أن $R_{X_i, others}^2$ يمثل مربع معامل

الارتباط المتعدد بين المتغير المستقل i وبقية المتغيرات المستقلة . ثم يستخرج معامل VIF لكل متغير مستقل (Variance Inflation Factor) حيث ان $VIF = \frac{1}{Tolerance}$. يعتبر هذا المعامل مقياساً لتأثير

الارتباط بين المتغيرات المستقلة على زيادة تباين معلمة المتغير المستقل (تتميز مشكلة التعدد الخطي بارتفاع تباين معالم النموذج وبالتالي عدم ظهور المعلمة معنوية نتيجة انخفاض قيمة إحصائية t بالرغم من أن المتغير قد يكون مهماً في النموذج) . أن الحصول على قيمة لمعامل VIF لأحد المتغيرات المستقلة تزيد عن 5 أو 10 تشير الى أن تقدير المعلمة المرافقة يتأثر بمشكلة التعدد الخطي ، أن قيمة VIF لمتغيرات النموذج تشير الى عدم تأثر أي منها بمشكلة التعدد الخطي ،

تستخدم الجذور المميزة لمصفوفة XX' في تشخيص مشكلة التعدد الخطي (X هي مصفوفة المتغيرات المستقلة) ففي حالة وجود عدة جذور مميزة قريبة من الصفر فهذا دليل على مشكلة التعدد الخطي . لاختبار وجود ارتباط خطي بين المتغيرات المستقلة يستعمل ما يعرف بدليل الحالة Condition Index وهو عبارة عن الجذر التربيعي لحاصل قسمة أكبر جذر مميز على كل من الجذور المميزة مثلاً يحتسب هذا الدليل للحد الثابت كما يلي :

$$Condition\ Index = \sqrt{\frac{4.813}{4.813}} = 1$$

وبالنسبة للمتغير X_1 يحتسب كما يلي :

$$condition\ Index = \sqrt{\frac{4.813}{0.09868}} = 7.020$$

فاذا زادت قيمة الدليل عن 15 فهذا مؤشر على إمكانية وجود مشكلة التعدد الخطي. أما إذا زادت عن 30 فهذا مؤشر على خطورة المشكلة. في الجدول التالي نلاحظ أن أكبر قيمة للدليل هي 16 أي إمكانية وجود المشكلة .

المقياس الآخر للمشكلة هو ما يعرف بـ Variance Proportion وهو يمثل نسبة تباين التقدير المفسر بواسطة المكون الأساسي Principle Component المرافق لكل جذر مميز حيث تعتبر مشكلة التعدد الخطي مؤثرة إذا كان المكون الأساسي المرافق لدليل حالة Condition Index مرتفع نسبياً يساهم بصورة أساسية في تباين أثنتين أو أكثر من المتغيرات المستقلة. في هذا المثال المكون الأساسي الخامس يساهم بصورة كبيرة في تباين المتغير X_3 فقط ولهذا لا تعتبر مشكلة التعدد الخطي مؤثرة بشكل كبير على البيانات وكما هو واضح في الجدول التالي :

Collinearity Diagnostics^a

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions				
				(Constant)	X1	X2	X3	X4
1	1	4.813	1.000	.00	.00	.00	.00	.00
	2	9.768E-02	7.020	.01	.06	.17	.00	.35
	3	4.345E-02	10.525	.01	.71	.29	.08	.01
	4	2.907E-02	12.868	.30	.05	.26	.22	.63
	5	1.669E-02	16.983	.68	.18	.28	.70	.01

a. Dependent Variable: Y

لتنفيذ لمطلوب الثالث من المثال نتبع الخطوات الاعتيادية للانحدار المتعدد وكما يلي :

□ من شريط القوائم اختر Linear Regression → Regression → Analyze فيظهر صندوق حوار Linear Regression الذي يرتب كالتالي :

- أشر المتغير Y وادخله الى خانة Dependent .
- أشر المتغيرات X_1, X_2, X_3 وانقلها الى خانة Independent .
- في خانة Method أنقر السهم المتجه للأسفل فتظهر الخيارات التالية لاختيار أفضل نموذج انحدار :

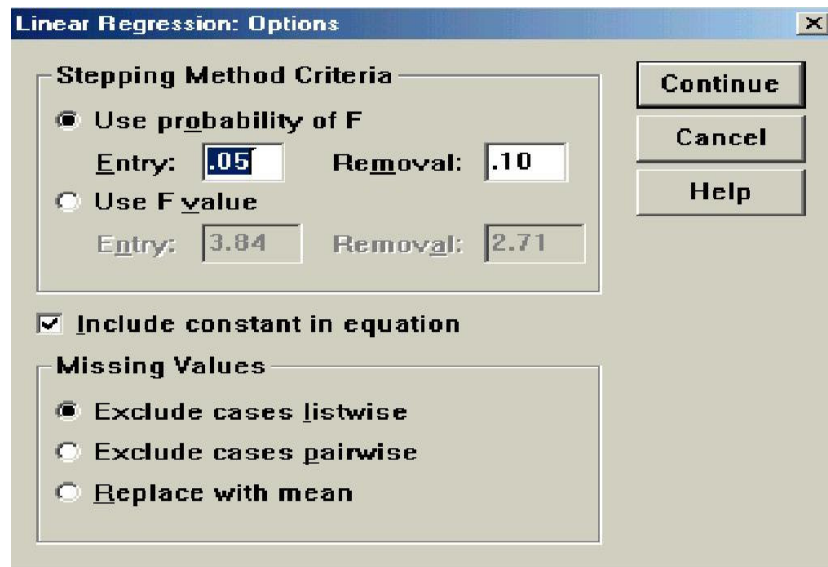
1. Enter : إدخال كافة المتغيرات المستقلة الى نموذج الانحدار (الأسلوب الاعتيادي لتنفيذ الانحدار) .
 2. Stepwise : إدخال المتغيرات واحداً بعد الآخر بخطوات متسلسلة إلى النموذج مع استبعاد المتغيرات التي تصبح غير مؤثرة بوجود بقية المتغيرات .
 3. Remove : استبعاد المتغيرات غير المهمة بخطوة واحدة من النموذج .
 4. Backward : استبعاد المتغيرات غير المؤثرة واحداً بعد الآخر بخطوات متسلسلة.
 5. Forward : إدخال المتغيرات واحداً بعد الآخر الى النموذج ولا يتم استبعاد المتغيرات التي تصبح غير مؤثرة فيما بعد من النموذج .
- أختار نوع طريقة الانحدار Stepwise (تعتبر أفضل الطرق) .

◀ أنقر زر Statistics في صندوق حوار Linear Regression فيظهر صندوق حوار Statistics حيث نقوم بتأشير كل مما يلي

Estimate : لتقدير معالم النموذج والإحصاءات المرافقة .

Model Fit : لتقدير R^2 وANOVA .

◀ أنقر زر Options في صندوق حوار Linear Regression فيظهر صندوق حوار Options وكما يلي :



تحتاج طريقة Stepwise الى تحديد مستوى المعنوية أو قيمة F التي يتم بموجبها إدخال واستبعاد المتغيرات من النموذج . في حقل Stepping Method criteria وذلك بموجب الخيارين التاليين :

- Use Probability of F : تحديد مستوى المعنوية الذي سيستخدم في كل خطوة لإدخال واستبعاد المتغيرات وان مستوى الدلالة لإدخال المتغيرات يجب أن يكون أقل من مستوى دلالة استبعاد المتغيرات أو مساويا له . لتضمنين متغيرات اكثر في النموذج زد من قيمة Entry . لاستبعاد عدد أكبر من المتغيرات من النموذج قلل من قيمة Removal .

- Use F Value : تحديد قيمة F الذي ستستخدم في كل خطوة لإدخال واستبعاد المتغيرات وان قيمة F لإدخال المتغيرات يجب أن تكون أعلى من قيمة F لاستبعاد المتغيرات . لتضمنين متغيرات اكثر في النموذج قلل من قيمة Entry . لاستبعاد عدد أكبر من المتغيرات من النموذج زد من قيمة Removal .

أن الاختبار المستخدم في إدخال واستبعاد المتغيرات هو اختبار F الجزئي Partial F Test الذي يستخدم في اختبار معنوية جزء من معالم النموذج ويستخدم هنا في اختبار معنوية معلمة واحدة فقط لمتغير واحد لاختبار معنوية مجموع المربعات التي يضيفها المتغير المستقل الى النموذج لكي نتوصل الى قرار بشأن استبعاده أو بقاءه في النموذج ، علما أن هذا الاختبار F يكون مكافئاً تماماً لاختبار t في حالة استعمال اختبار F لاختبار معنوية معلمة واحدة فقط .

لقد اعتمدنا القيم الافتراضية للخيار الأول Use Probability of F أي 0.05 للإدخال و0.10 للاستبعاد .

عند نقر زر OK في صندوق حوار Linear regression يتم عرض النتائج التالية :

Variables Entered/Removed^a

Model	Variables Entered	Variables Removed	Method
1	X3	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
2	X4	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).

a. Dependent Variable: Y

بموجب طريقة Stepwise يتم أخال المتغيرات الأربعة واحداً بعد الآخر الى النموذج علماً أن المتغير الداخل عرضة للاستبعاد في الخطوات اللاحقة إذا ثبتت عدم معنويته بوجود المتغيرات الأخرى . لقد كان X_3 أول المتغيرات الداخلة الى النموذج لأن له أكبر معامل ارتباط بسيط مع المتغير المعتمد وبالتالي أكبر قيمة لإحصائية t . من الجدول اللاحق Coefficients نلاحظ أن قيمة P-Value المرافقة لإحصائية t تساوي 0.02 وهي أقل من 0.05 (مستوى الدلالة للإدخال Entry) ولهذا يسمح بإدخال X_3 الى النموذج (لاحظ أن اختبار t الذي أستخدمناه هنا مكافئ تماماً لاختبار F الجزئي) وعليه يصبح النموذج في الخطوة الأولى كما يلي :

$$\hat{y} = 33.007 + 1.158X_3$$

في الخطوة الثانية يتم إدخال المتغير الذي له أعلى معامل ارتباط جزئي مع المتغير المعتمد بثبات المتغير X_3 وهو المتغير X_4 ولكن يجب أولاً التأكد من معنوية المتغير بحساب إحصائية T له (من الجدول اللاحق P-Value المرافقة لإحصائية t تساوي 0.033 وهي أقل من 0.05) (مستوى الدلالة للإدخال Entry) ولهذا يسمح بإدخال X_4 الى النموذج ليصبح على الشكل التالي :

$$\hat{y} = 46.152 + 1.345X_3 - 2.147X_4$$

في النموذج أعلاه نجد القيمة الأقل لإحصائية t للمتغيرين X_3 و X_4 أي القيمة الأكبر لـ P-Value المرافقة لإحصائية t وقد كانت للمتغير X_4 و تساوي 0.033 وهي أقل من 0.10 (مستوى الدلالة للاستبعاد Removal) ولهذا يبقى كلا المتغيرين في النموذج .

أن النموذج أعلاه هو النموذج النهائي ويمثل أفضل نموذج انحدار ولا يتضمن المتغيرين X_2 و X_1 حيث لم تظهر معنوية المتغير الذي له أعلى ارتباط بالمتغير المعتمد بمستوى دلالة 0.05 . الجدول التالي يبين النماذج التي تم اختبارها لحين الوصول الى النموذج النهائي .

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	33.007	11.648		2.834	.016
	X3	1.158	.426	.634	2.720	.020
2	(Constant)	46.152	11.011		4.191	.002
	X3	1.345	.360	.737	3.735	.004
	X4	-2.147	.871	-.486	-2.466	.033

a. Dependent Variable: Y

(1) النموذج الأول (2) النموذج الثاني (النهائي)

ملاحظة : في حالة اختيار Use F Value في Stepping Method Criteria فإن المقارنة تتم بين t^2 والقيمة المحددة لإحصائية F للإدخال والاستبعاد فعند إدخال متغير إذا كانت $t^2 > F$ (Enter) نسمح بدخول المتغير إلى النموذج . عند إخراج متغير إذا كانت $t^2 > F$ (Remove) نسمح ببقاء المتغير في النموذج .
الجدول التالي يبين المتغيرات التي تم استبعادها في كل من النموذجين الأول والثاني :

Excluded Variables^c

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics
						Tolerance
1	X1	.027 ^a	.108	.916	.034	.915
	X2	.267 ^a	.941	.369	.285	.682
	X4	-.486 ^a	-2.466	.033	-.615	.955
2	X1	-.012 ^b	-.058	.955	-.019	.909
	X2	.175 ^b	.721	.489	.234	.663

a. Predictors in the Model: (Constant), X3

b. Predictors in the Model: (Constant), X3, X4

c. Dependent Variable: Y

حيث أن Beta in تمثل المعلمة المعيارية للمتغير فيما لو أدخل إلى النموذج في الخطوة اللاحقة .

الفصل الحادي عشر

التحليل العاملي

Factor Analysis

(11-1) التحليل العاملي

تهدف طرق التحليل العاملي الى إيجاد مجموعة من العوامل Factors التي تكون مسؤولة عن توليد الاختلافات Variations في مجموعة مكونة من عدد كبير من متغيرات الاستجابة Response Variables حيث يمكن التعبير عن المتغيرات المشاهدة كدالة في عدد من العوامل المستترة Factors وغالباً ما يعبر عن متغيرات الاستجابة بتركيب خطي Linear Compounds من العوامل المستترة .حيث تكون العلاقة بين المتغيرات داخل العامل الواحد أقوى من العلاقة مع المتغيرات في عوامل أخرى .
أن التحليل العاملي يساعد على فهم تركيب مصفوفة الارتباط أو التباين المشترك من خلال عدد قليل من العوامل .

(11-2) طريقة المكونات الأساسية Principal Components Method

أن طريقة المكونات الأساسية هي واحدة من أهم طرق التحليل العاملي وتأتي في مقدمة الطرق لبساطتها .

أن المكون الأساسي (أو العامل) هو عبارة عن تركيب خطي من متغيرات الاستجابة Response Variables . باعتبار أن لدينا p من متغيرات الاستجابة فأن المكون الأساسي الأول يعبر عنه كما يلي :

$$Z_1 = a_{11}X_1 + a_{21}X_2 + + a_{p1}X_p$$

حيث أن a_{ij} تمثل تشبعات Loadings متغيرات الاستجابة بالعامل الأول . أما المكون الأساسي الثاني فيعبر عنه كما يلي :

$$Z_2 = a_{12}X_1 + a_{22}X_2 + + a_{p2}X_p$$

أن المكون الأول له أعظم تباين Variance (يفسر أكبر نسبة من هيكل التباينات لمتغيرات الاستجابة) يليه المكون الأساسي الثاني وهكذا . وأن هذه المكونات تكون متعامدة فيما بينها Orthogonal ويمكن حساب المكونات بطريقتين :

1. استعمال مصفوفة التباين المشترك Variance-Covariance Matrix لمتغيرات الاستجابة وفي هذه الحالة فأن المتغيرات تكون مقاسه بالانحرافات عن الوسط الحسابي $\bar{X} - X$.

2. استعمال مصفوفة الارتباطات Correlation Matrix لمتغيرات الاستجابة وفي هذه الحالة تستعمل المتغيرات المعيارية Standardized Variables ويكون ذلك ضرورياً في حالة اختلاف وحدات القياس لمتغيرات الاستجابة .

مثال 1 :

البيانات التالية تمثل بعض المتغيرات الاجتماعية لمستوى المعيشة حسب المناطق regions (المحافظات) في العراق لسنة 1977 والتي تظهر كما يلي بعد إدخالها في شاشة Data Editor لبرنامج SPSS :

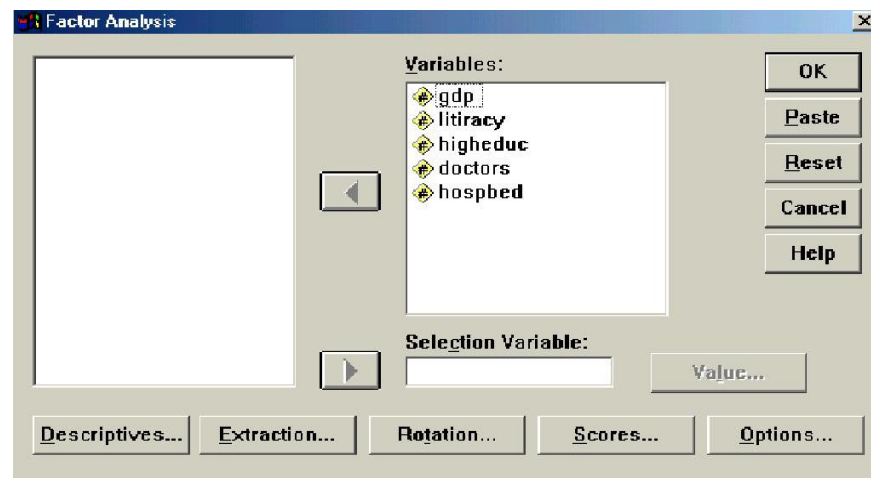
Region	gdp	literacy	higheduc	doctors	hospbed
Dohok	17.2	30.1	1.09	17	139
Nineveh	24.0	44.2	1.85	14	172
Arbil	22.2	35.2	1.18	13	163
Sulayman	16.2	33.5	1.01	10	115
Ta'meem	32.3	49.4	1.85	18	143
Salah AL-Deen	98.4	37.9	1.32	15	80
Diala	23.1	44.1	1.93	10	153
Anbar	22.7	44.3	1.58	16	144
Baghdad	75.0	61.6	4.04	36	280
Wasit	19.5	36.7	1.11	28	199
Babylon	22.8	44.1	1.82	18	145
Kerbala	21.5	47.7	1.53	24	173
Najaf	18.7	46.2	1.59	27	190
Qadisia	21.0	35.2	.95	9	195
Muthana	21.3	33.5	.84	18	178
Thi-Qar	18.1	33.8	.73	12	144
Maysan	20.4	34.4	.90	11	301
Basrah	19.0	53.6	2.24	25	219

حيث أن المتغير gdp يمثل معدل نصيب الفرد من خدمات التنمية الاجتماعية والمتغير literacy يمثل النسبة المئوية للمتعلمين من الكبار والمتغير higheduc يمثل النسبة المئوية للباحثين على شهادة عليا والمتغير doctors يمثل عدد الأطباء لكل 100000 من السكان والمتغير hospbed يمثل عدد أسرة المستشفيات لكل 100000 من السكان .

يطلب تحليل هيكل الارتباطات للمتغيرات المذكورة باستخدام طريقة المكونات الأساسية .

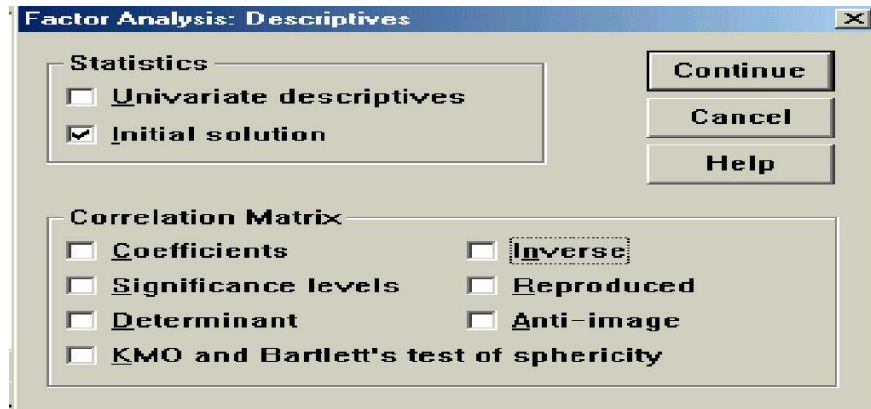
لتنفيذ ذلك نتبع الخطوات التالية :

من شريط القوائم اختر Factor → Data Reduction → Analyze فيظهر صندوق حوار Factor Analysis الذي نرتبه كما يلي :



لاحظ أن المتغير Region لا يدخل في التحليل كونه متغير رمزي .

□ عند نقر زر Descriptives يظهر صندوق الحوار التالي :



والذي يتضمن قسمين أساسيين :

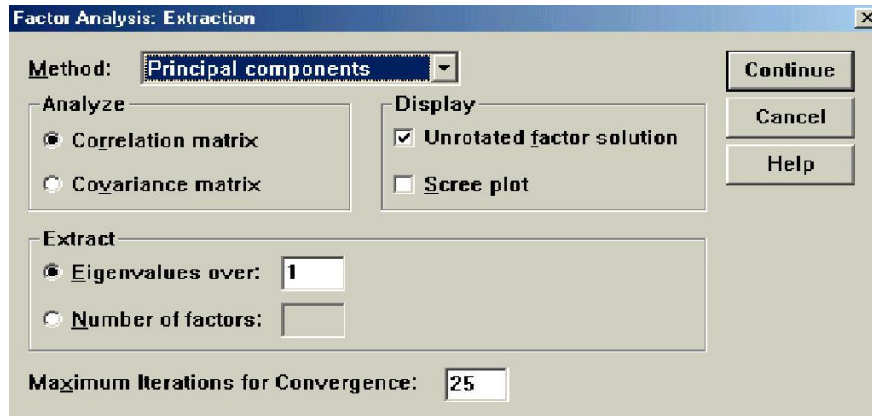
1. Statistics ويتضمن ما يلي :

univariate descriptives : ويعرض بعض الإحصائيات البسيطة للمتغيرات مثل Mean ، Standard ... Deviation

Initial solution : ويعرض القيم الأولية للاشتراكيات Communalities ، الجذور الكامنة (أو الكامنة) Eigen Values والنسبة المئوية للتباين المفسر .

2. Correlation Matrix : يعرض مصفوفة معاملات الارتباطات ، مستوى المعنوية ، المحددة ومعكوس مصفوفة الارتباطات Inverse.

□ عند نقر الزر Extraction في صندوق حوار Factor Analysis يظهر صندوق الحوار التالي الذي نرتبه كما يلي:



ويتضمن الصندوق التفاصيل التالية :

Method : لاختيار الطريقة المطلوبة في التحليل في هذا المثال اخترنا طريقة المكونات الأساسية (يتضمن البرنامج سبعة طرق من طرق التحليل العاملي) .

Analyse : ويتضمن ما يلي :

Correlation Matrix : يتم تحليل مصفوفة الارتباطات للمتغيرات المدروسة ويكون ذلك ضرورياً في حالة اختلاف وحدات القياس للمتغيرات المشمولة .

covariance Matrix : : يتم تحليل مصفوفة التباين والتباين المشترك للمتغيرات المدروسة ويمكن اعتماد ذلك في حالة أن المتغيرات المدروسة لها نفس وحدات القياس .

Extraction : ويتضمن أسلوبين لاستخلاص المكونات (أو العوامل) وكما يلي :

Eigenvalues Over : يتم استخلاص المكونات الأساسية التي لها جذور كامنة أو تعرف أيضاً (بالجذور المميزة) تزيد عن قيمة معينة يحددها المستفيد (على الأغلب تساوي واحد) .

Number of Factors : يتم استخلاص عدد معين من المكونات (العوامل) يحدد عددها من قبل المستفيد .

Maximum Iteration for Convergence : تحديد الحد الأعلى لعدد خطوات الخوارزمية اللازمة للوصول الى حل .

Display : يتضمن مايلي :

Unrotated factor solution يتم عرض تشبعات المتغيرات بالعوامل غير المدورة Factor Matrix وفي هذا المثال Component Matrix لأننا نستعمل طريقة المكونات الأساسية .

Scree Plot : يتم عرض مخطط يمثل المحور الأفقي رقم المكون أم المحور الصادي فيمثل الجذور الكامنة Eigen Values .

□ عند نقر زر Rotation في صندوق حوار Factor Analysis يظهر صندوق حوار Rotation

ويحتوي على خمسة طرق لتدوير المحاور حيث أن تدوير المحاور هي طريقة هندسية الغرض منها جعل التشبعات (Loadings) الكبيرة أكبر والتشبعات الصغيرة أصغر مما هي عليه قبل التدوير. كما يمكن أن تقلل من التشبعات السالبة وتزيد من التشبعات الصفرية في الحالات التي لا يكون هناك تفسير منطقي للإشارة السالبة للتشبع . ويوفر البرنامج خمسة طرق للتدوير وهي (Direct Oblimin ، Varimax ، Promax ، Equamax ، Quartimax) . في هذا المثال أشرنا الاختيار None أي عدم تدوير المحاور .

□ عند نقر زر Scores في صندوق حوار Factor Analysis يظهر صندوق حوار Factor Scores

وكما يلي :



يتضمن الصندوق الفقرات التالية :

Save as Variables : عند تأشير هذا الخيار يقوم البرنامج بحساب العوامل Factor Scores (في هذا المثال المكونات الأساسية) وتضاف هذه المكونات الى يمين المتغيرات الموجودة في شاشة Data Editor علماً أن العامل Factor Score أو المكون الأساسي في هذا المثال هو عبارة عن تركيب خطي من المتغيرات المعيارية Z Scores ويمكن حساب العامل (المكون الأساسي) رقم i (عدد العوامل الكلي يكون بقدر عدد المتغيرات ويساوي p) بموجب المعادلة التالية :

$$F_i = \sum_{j=1}^p w_{ij} z_j \quad (j = 1, 2, \dots, p)$$

حيث أن :

z : تمثل المتغيرات المعيارية ، w : تمثل الأوزان وهي معاملات العوامل Factor Scores Coefficients ونحصل عليها عند تأشير الخيار التالي :

Display Factor Score Coefficient Matrix علماً أن البرنامج يعرض العوامل المستخلصة فقط ومعاملاتها .

ملاحظة : يتم حساب العوامل بثلاثة طرق (Anderson-Rubin ، Bartlett، Regression) أما بالنسبة للمكونات فإن الطرق الثلاثة تعطي نفس النتائج عند تأشيرها حيث أن هناك متجه وحيد للمعاملات .

□ يمكن بواسطة الزر Option في صندوق حوار Factor Analysis تنظيم استبعاد الحالات الحاوية

على قيم مفقودة وكذلك ترتيب التشبعات في مصفوفة المكونات Components Matrix (التي سترد لاحقاً) حسب الحجم وعدم إظهار المعاملات التي تقل قيمتها المطلقة عن قيمة تحدد من قبل المستفيد .

□ عند نقر زر OK في صندوق حوار Factor Analysis نحصل على النتائج التالية :

Communalities

	Initial	Extraction
GDP	1.000	.797
LITIRACY	1.000	.843
HIGHEDUC	1.000	.896
DOCTORS	1.000	.704
HOSPBED	1.000	.772

Extraction Method: Principal Component Analysis.

الجدول أعلاه يمثل القيم الأولية والمستخلصة للاشتراكيات Communalities حيث أن القيم الأولية للاشتراكيات تؤخذ مساوية الى الواحد في طريقة المكونات الأساسية في حالة اعتماد مصفوفة الارتباطات وتؤخذ الاشتراكيات مساوية لتباين كل متغير في حالة اعتماد مصفوفة التباينات أما بقية الطرق فتستعمل R^2 الناتج من انحدار كافة المتغيرات على متغير معين كتقدير لاشترائية المتغير .

أن القيمة المستخلصة لاشترائية المتغير GDP مثلاً تشير الى أن 0.797 من التباينات في قيم المتغير GDP تفسرها العوامل المشتركة (تم استخلاص عاملين) أن قيمة الاشتراكية تتراوح من 0 الى 1 وهي تعبر عن مربع معامل الارتباط المتعدد Square Multiple Correlation للمتغير GDP مع المكونات (العوامل) وبصورة عامة نلاحظ أن العوامل المشتركة تفسر نسبة عالية من تباين المتغيرات حيث أن اقل نسبة هي 0.704 لمتغير Doctors . في حالة الحصول على قيمة صغيرة لاشترائية أحد المتغيرات فهذا يشير الى عدم أهمية المتغير ويوصى باستبعاده من التحليل .

الجدول التالي يبين الجذور الكامنة لمصفوفة الارتباطات (تباين المكونات) ومجموعها يساوي رتبة المصفوفة ويساوي 5 بقدر عدد المتغيرات حيث أن المكون الرئيسي الأول له أكبر جذر كامن (أو تباين المكون) ويساوي 2.900 ويفسر 58% من التباينات الكلية لمتغيرات التنمية الاجتماعية حيث أن :

$$\text{نسبة التباين المفسر للمكون الأول} = \frac{\text{الجذر الكامن}}{\text{مجموع الجذور الكامنة}} * 100 = \frac{2.9}{5} = 58\% .$$

وأن المكون الثاني يفسر 22.2% من التباينات ويفسر المكونان نسبة 80.2% من هيكل التباينات للمتغيرات الخمسة ، وقد أهمل البرنامج بقية المكونات نظراً لكون جذورها الكامنة تقل عن الواحد .

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.900	58.009	58.009	2.900	58.009	58.009
2	1.110	22.207	80.216	1.110	22.207	80.216
3	.523	10.457	90.673			
4	.393	7.864	98.537			
5	7.313E-02	1.463	100.000			

Extraction Method: Principal Component Analysis.

المخرج التالي يمثل مصفوفة المكونات Components Matrix التي تتضمن تشبعات Loadings المكونين الأول والثاني الذين تم استخلاصهما أن التشبع هو عبارة عن معامل الارتباط البسيط بين المكون (أو العامل) والمتغير . أن أقوى المتغيرات ارتباطاً بالعامل الأول هو متغير التعليم العالي HIGHEDEC حيث أن تشبع المتغير بالمكون الأساسي الأول هو 0.941 يليه متغير التعليم LITERACY ثم متغير عدد الأطباء DOCTORS وأن أضعف المتغيرات ارتباطاً بالعامل الأول هما متغيري GDP ومتغير عدد أسرة المستشفيات HOSPBED. أما أقوى المتغيرات ارتباطاً بالمكون الثاني فهو GDP ثم متغير HOSPBEDS ولكن باتجاه معاكس، ويرتبط المكون الثاني بعلاقة ضعيفة ببقية المتغيرات .

Component Matrix ^a

	Component	
	1	2
GDP	.477	.754
LITIRACY	.918	1.899E-02
HIGHEDEC	.941	.102
DOCTORS	.829	-.131
HOSPBED	.508	-.717

Extraction Method: Principal Component Analysis.

a. 2 components extracted.

ملاحظات :

1. أن اشتراكية المتغير هي مجموع مربعات تشبعات المتغير بالعوامل المستخلصة فاشتراكية المتغير GDP تساوي $0.477^2 + 0.754^2 = 0.797$ وهكذا لباقي المتغيرات .
2. أن مجموع مربعات تشبعات المتغيرات بالمكون (العامل) يساوي الجذر الكامن للمكون فمثلاً نحصل على الجذر الكامن للمكون (العامل) الأول كما يلي :

$$0.477^2 + 0.918^2 + 0.941^2 + 0.829^2 + 0.508^2 = 2.9$$
لاحظ أنه قد يحصل اختلاف بسيط في النتائج بسبب عمليات التقريب .

المخرج التالي يمثل معاملات المكونات (العوامل) Component Scores Coefficients وتحسب هذه المعاملات من مصفوفة المكونات السابقة Martix Components فمثلاً المعامل GDP للمكون الأول يحسب كما يلي :

$$\frac{0.477}{0.477^2 + 0.918^2 + 0.941^2 + 0.829^2 + 0.508^2} = 0.165$$

Component Score Coefficient Matrix

	Component	
	1	2
GDP	.165	.679
LITIRACY	.316	.017
HIGHEDUC	.324	.092
DOCTORS	.286	-.118
HOSPBED	.175	-.645

Extraction Method: Principal Component Analysis.

Component Scores.

أما قيم المكونات (العوامل) فتحسب بموجب الدالة الخطية التالية للمكون الأول مثلاً :

fac_1 =

0.165*GDP+0.316*LITIRACY+0.324*HIGHEDUC+0.286*DOCTORS+0.175*HOSPBED

علماً أنه يتم استعمال المتغيرات المعيارية Standardized variables في التركيب الخطي (أي

القيم المعيارية لمتغيرات GDP ، Literacy ، ...) وكما ذكرنا تضاف العوامل الى يمين المتغيرات

الموجودة في شاشة Data Editor وكما يلي :

Region	gdp	literacy	higheduc	doctors	hospbed	fac_1	fac_2
Dohok	17.2	30.1	1.09	17	139	-.84873	.00862
Nineveh	24.0	44.2	1.85	14	172	.05265	-.01091
Arbil	22.2	35.2	1.18	13	163	-.65458	-.04095
Sulayman	16.2	33.5	1.01	10	115	-1.10939	.37657
Ta'meem	32.3	49.4	1.85	18	143	.37097	.54501
Salah AL-Deen	98.4	37.9	1.32	15	80	-.11362	3.32500
Diala	23.1	44.1	1.93	10	153	-.14014	.26391
Anbar	22.7	44.3	1.58	16	144	-.08303	.22314
Baghdad	75.0	61.6	4.04	36	280	3.22748	.21850
Wasit	19.5	36.7	1.11	28	199	.04721	-.80457
Babylon	22.8	44.1	1.82	18	145	.09229	.21066
Kerbala	21.5	47.7	1.53	24	173	.41839	-.29133
Najaf	18.7	46.2	1.59	27	190	.53708	-.62771
Qadisia	21.0	35.2	.95	9	195	-.81013	-.42906
Muthana	21.3	33.5	.84	18	178	-.62869	-.37423
Thi-Qar	18.1	33.8	.73	12	144	-1.03052	.02027
Maysan	20.4	34.4	.90	11	301	-.44139	-1.76886
Basrah	19.0	53.6	2.24	25	219	1.11415	-.84406

علماً أن المكونات الأساسية هي متغيرات وهمية (نظرية) ليس لها أي تفسير محدد ولكن لها

استعمالات خاصة مثلاً معالجة مشكلة التعدد الخطي في نماذج الانحدار ، كما ويمكن الاستفادة من المكونات

(العوامل) في تحديد الحالات الشاذة . أن المكونات (العوامل) هي متغيرات معيارية بمتوسط مساو للصفر وانحراف معياري مساو للواحد فإذا كانت هذه العوامل تتبع التوزيع الطبيعي فعليه تعتبر قيم العامل التي تقع خارج المدى (2,2-) قيمة شاذة .

أن رسم مخطط الانتشار Scatter Plot للمكونين المستخلصين في هذا المثال يساعد في تحديد الحالات الشاذة . لرسم هذا المخطط نتبع الخطوات التالية :

من شريط القوائم اختر Scatter (Simple) → Graphs فيظهر صندوق حوار Scatter Plot .

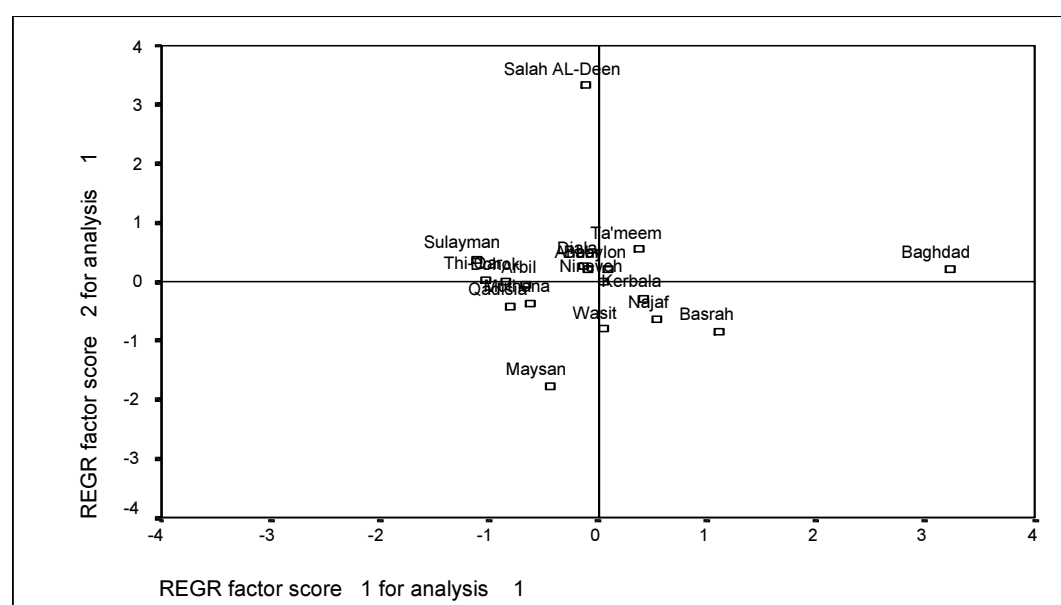
Y Axis : fac_2 عرف .

X Axis : fac_1 عرف .

Label Cases by : Region

أقر زر OK .

(عند ظهور الرسم أقره مرتين ثم اختر Reference Line → Chart من شريط قوائم شاشة Chart Editor لكل من X Axis و Y Axis عند الصفر) يظهر المخطط التالي :



في حالة أن التوزيع الاحتمالي للمكونات (العوامل) يقترب من التوزيع الطبيعي فإن نقاط شكل الانتشار تتوزع بشكل دائري حول النقطة (0,0). لرصد القيم الشاذة نلاحظ أن كافة نقاط المكونين تقع ضمن المدى (2,2-) عدا نقطتين الأولى لمحافظة بغداد التي لها قيمة شاذة للعامل الأول (تقرب من 3) والثانية لمحافظة صلاح الدين التي لها قيمة شاذة للعامل الثاني (تقرب من 3 أيضاً) ولغرض تحديد المتغيرات المسؤولة عن القيم الشاذة للمحافظتين (الحالتين) المذكورتين نقوم بمقارنة بيانات المحافظة مع متوسط كافة المحافظات للمتغيرات ذات التشعبات الكبيرة في العامل ، فبالنسبة لمحافظة بغداد فأن هناك ثلاثة متغيرات تنتشعب بدرجة عالية بالعامل الأول هي literacy و highedu و doctors وقد كانت قيمة المتغيرات لمحافظة بغداد أعلى بكثير من متوسط هذه المتغيرات لعموم محافظات القطر كما يبين ذلك الجدول التالي وهذا هو سبب الحصول على قيمة كبيرة لمحافظة بغداد للعامل الأول (3.227).

Case Summaries

Statistics: Mean

	LITIRACY	HIGHEDUC	DOCTORS
Baghdad	61.60	4.04	36.00
Total	41.42	1.53	17.83

اما سبب القيمة الشاذة (3.325) لمحافظة صلاح الدين فيرجع الى القيمة العالية لمتغير GDP (الذي يتشبع بدرجة عالية بالمكون الثاني) بالمقارنة مع متوسط المتغير لكافة المحافظات وعلى الرغم من انخفاض HOSBED في المحافظة بالمقارنة بمتوسط القطر الا أن هذا المتغير يتشبع بدرجة عالية بالمكون الثاني ولكن باشارة سالبة (أنظر مصفوفة المكونات Components Matrix) . المقارنة يوضحها الجدول التالي :

Case Summaries

Mean

	GDP	HOSPBED
Salah AL-Deen	98.40	80.00
Total	28.52	174.06

ملاحظات :

1. يمكن الحصول على المتجه الكامن (أو المميز) Eigen Vector المرافق للجذر الكامن بإيجاد طول المتجه NORM حسب المعادلة $\|V\| = \sqrt{X_1^2 + X_2^2 + \dots + X_n^2}$ ولكل مكون في المصفوفة Components Matrix وذلك بقسمة كل قيمة من قيم المكون (التشبع) على طول المكون Norm للحصول على المتجه المعياري Normalized Vector الذي يكون طوله واحد وهذا يكون متجه وحيد Unique Vector فمثلاً تحسب القيمة الأولى للمتجه الكامن الأول من تشبع المتغير GDP بالعامل الأول كما يلي :

$$\frac{0.477}{\sqrt{0.447^2 + 0.918^2 + 0.941^2 + 0.829^2 + 0.508^2}} = 0.280$$

الجدول التالي يمثل المتجهين الكامين الأول والثاني المرافقين للجذرين الكامين الأول (2.9) والثاني

(1.11) لمصفوفة الارتباطات .

	Eigen Vector 1	Eigen Vector 2
GDP	0.280278	0.715757
LITERACY	0.538845	0.018025
HIGHEDUC	0.552474	0.096474
DOCTORS	0.486656	-0.12474
HOSPBED	0.298378	-0.68007

2. أن المكونات الأساسية تكون متعامدة فيما بينها Orthogonal أي أن حاصل الضرب القياسي Dot Product لها يساوي صفر $\langle fac_1, fac_2 \rangle = \sum fac_1 * fac_2 = 0$ أي مجموع حاصل

ضرب القيم المتقابلة للمكونين . كذلك فأن حاصل الضرب القياسي لكل من متجهي معاملات المكونات Components Score Coefficients والمتجهين الكامنين يساوي صفر .

(3 - 11) طرق التحليل العاملي Factor Analysis Methods

أن طريقة المكونات الأساسية هي واحدة من طرق التحليل العاملي وتعتبر نقطة البداية في إجراء التحليل العاملي بأية طريقة أخرى وتتميز بالبساطة وبأنها تعتمد أسلوباً رياضياً في الاحتساب .

أما طرق التحليل العاملي بصورة عامة فيعبر عنها نموذج التحليل العاملي Factor Analysis Model الذي يعبر عن كل متغير من متغيرات الاستجابة $X_1, X_2, \dots, X_p$ كدالة في مجموعة من العوامل المشتركة (لمتغيرات الاستجابة) Common Factors وعامل وحيد خاص بذلك المتغير .

بافتراض التوزيع الطبيعي لمتغيرات الاستجابة يكون نموذج التحليل العاملي كما يلي :

$$X_1 = a_{11}Y_1 + \dots + a_{1m}Y_m + e_1$$

$$\dots$$

$$X_p = a_{p1}Y_1 + \dots + a_{pm}Y_m + e_p$$

حيث أن

Y_j : العامل المشترك j (عدد العوامل هو m أقل من عدد مكونات الاستجابة p) .

a_{ij} : التشبعات Loadings وهي معالم تعكس أهمية العامل j في تركيب متغير الاستجابة i .

e_i : العامل الوحيد الخاص بالمتغير X_i (Specific Factor) .

مثال 2 : (على طريقة الإمكان الأعظم Maximum Likelihood Method)

لبيانات المثال 1 سنحاول إجراء التحليل العاملي بطريقة الإمكان الأعظم التي تعتبر من الطرق المهمة في التحليل العاملي لمقارنة أوجه الاختلاف بين هذه الطريقة وطريقة المكونات الرئيسية .

بالنسبة لخطوات التنفيذ فهي مشابهة لما ورد في المثال الأول عدا أننا نختار نوع الطريقة

Maximum Likelihood في صندوق حوار Factor Analysis : Extraction حيث تظهر المخرجات التالية :

Factor Analysis

Communalities

	Initial	Extraction
GDP	.350	.999
LITIRACY	.852	.903
HIGHEDUC	.870	.935
DOCTORS	.503	.487
HOSPBED	.286	.227

Extraction Method: Maximum Likelihood.

أن التشبعات الأولية initial Communalities بالنسبة لطريقة المكونات الأساسية هي واحد دائماً أما باقي طرق التحليل العاملي فأن التشبع الأولي لأي متغير هو مربع معامل الارتباط المتعدد R^2 الذي نحصل عليه من أنحدار المتغير (Dependent Variable) على بقية المتغيرات (Independent Variables) . أن القيمة 0.999 لمتغير الاستجابة GDP تبين أن العاملين المستخلصين الأول والثاني يفسران 0.999 من الاختلافات الكلية للمتغير GDP .

Total Variance Explained

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.900	58.009	58.009	1.413	28.254	28.254
2	1.110	22.207	80.216	2.138	42.756	71.009
3	.523	10.457	90.673			
4	.393	7.864	98.537			
5	7.313E-02	1.463	100.000			

Extraction Method: Maximum Likelihood.

الجدول أعلاه يبين الجذور الكامنة الأولية Initial Eigenvalues وهي نفسها التي حصلنا عليها لطريقة المكونات الأساسية وفي طريقة المكونات الأساسية نلاحظ أن مجموع مربعات التشعبات المستخلص Extraction Sums Squared Loadings هو نفسه Initial eigenvalues وكذلك تتساوى نسبة التباين المفسر أما في طريقة الإمكان الأعظم فيوجد اختلاف بين القيم الأولية والمستخلصة وفي نسبة التباين المفسر كما هو واضح في الجدول أعلاه فمثلاً يتم احتساب مجموع مربعات تشعبات العامل المستخلص الأول من مصفوفة العوامل أدناه كما يلي :

$$0.999^2 + 0.333^2 + 0.469^2 + 0.275^2 - 8.71E-02^2 = 1.413$$

وأن نسبة ما يفسره العامل الأول هي $1.413 / 5 * 100 = 28.254\%$

Factor Matrix^a

	Factor	
	1	2
GDP	.999	-9.49E-03
LITIRACY	.333	.890
HIGHEDUC	.469	.846
DOCTORS	.275	.641
HOSPBED	-8.71E-02	.468

Extraction Method: Maximum Likelihood.

a. 2 factors extracted. 14 iterations required.

المصفوفة أعلاه هي مصفوفة العوامل المستخلصة ولها نفس خواص مصفوفة المكونات في المثال 1

Goodness-of-fit Test

Chi-Square	df	Sig.
1.337	1	.247

الجدول أعلاه يمثل اختباراً لفرضية العدم القائلة بأن العوامل المستخلصة كافية لتمثيل هيكل التباينات (أو الارتباطات) ضد الفرضية البديلة القائلة بعدم كفاية العوامل المستخلصة لتمثيل هيكل التباينات. وقد استخدمت لهذا الغرض أحصائية χ^2 حيث أن قيمة $P\text{-value} = 0.247$ تدعونا إلى قبول فرضية العدم

بمستوى دلالة 1% و 5% أي التسليم بكفاية العاملين المستخلصين لتمثيل هيكل التباينات لمتغيرات الاستجابة. ان هذا الاختبار غير موجود لطريقة المكونات الأساسية .

ملاحظة :

أن درجات العوامل Factor Scores وكما هو الحال بالنسبة للمكونات هي دالة خطية في متغيرات الاستجابة المعيارية وأن المعاملات Factor scores Coefficients تحتسب بطريقة واحدة لطريقة المكونات الأساسية ولهذا فهناك مكون (أو عامل) وحيد أما بالنسبة لطريقة الإمكان الأعظم فهناك 3 طرق لتقدير معاملات العوامل ولهذا نحصل على ثلاثة بدائل للعامل الواحد وهذه الطرق كما يلي (تظهر هذه الطرق عند نقر زر Scores في صندوق حوار Factor Analysis) :

Regression :العوامل الناتجة لها متوسط صفر وتباين مساوي لمربع معامل الارتباط المتعدد بين العوامل المقدرة والعامل الحقيقي .

Bartlett : العوامل الناتجة لها متوسط يساوي صفر وأن مجموع مربعات العوامل الخاصة اقل ما يمكن .
Anderson-Rubins :هي تحويل لطريقة Bartlett بحيث أن العوامل الناتجة لها متوسط يساوي صفر وانحراف معياري يساوي واحد وتكون العوامل مستقلة (غير مرتبطة) .

الفصل الثاني عشر

الاختبارات اللامعلمية

Non Parametric Tests

أن الاختبارات اللامعلمية هي اختبارات لا تعتمد إحصائية الاختبار فيها على معالم المجتمع كمعلمة المتوسط Mean أو التباين Variance كما أنها لا تفترض توزيع ما للبيانات ولهذا فهي تعرف أيضاً باختبارات التوزيع الحر Distribution – Free tests. أن سبب استعمال الاختبارات اللامعلمية يعود الى عدم توفر الفرضيات الخاصة بالاختبارات المعلمية فمثلاً يتوجب أن تتوزع بيانات المجتمع قريباً من التوزيع الطبيعي عند تطبيق اختبار T وهو أحد الاختبارات المعلمية شائعة الاستعمال وعند عدم توفر هذا الشرط نلجأ الى الاختبارات اللامعلمية علماً ان هذه الأخيرة لها شروط يجب توفرها ولكنها أسهل بكثير من شروط الاختبارات المعلمية كما أنه يجب استعمال الاختبارات المعلمية في حالة توفر الشروط الخاصة بها كونها أكثر دقة من الاختبارات اللامعلمية. يغطي برنامج SPSS عدداً كبيراً جداً من طرق الاختبار اللامعلمية وفيما يلي توضيح لكيفية تنفيذ بعض هذه الاختبارات .

(12 - 1) اختبار Chi-Square

يستعمل اختبار Chi-Square للمقارنة بين التكرار المشاهد للفئات Observed frequencies والتكرار المتوقع لها Expected Frequencies المحتسب على أساس فرضية العدم. فإذا كان لدينا فئات عددها K فإن أحصائية Chi-Square المستعملة في الاختبار تعطى كما يلي :

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

حيث أن O_i يمثل التكرار المشاهد وأن E_i يمثل التكرار المتوقع وأن درجة حرية الاختبار تساوي $k-1$.

مثال 1:

في تجربة لتهجين صنفين من الشعير تم الحصول على الصفات التالية :

التكرار المشاهد O_i	الصفات type	التسلسل
439	أسود بدون سفا	1
168	أسود ذو سفا	2
133	أبيض بدون سفا	3
60	أبيض ذو سفا	4
800	المجموع	

خاشع الراوي ، المدخل الى الأحصاء ، جامعة الموصل ، 1979، ص 375 .

المطلوب اختبار فرضية العدم التالية بمستوى دلالة 5% :

$$H_0 : P_1 = \frac{9}{16}, P_2 = \frac{3}{16}, P_3 = \frac{3}{16}, P_4 = \frac{1}{16}$$

اما الفرضية البديلة H_1 فتتص على أن النتائج تختلف عن هذه النسب (مثلاً P_1 تمثل نسبة الصفة الأولى أسود بدون سفا) وهكذا .

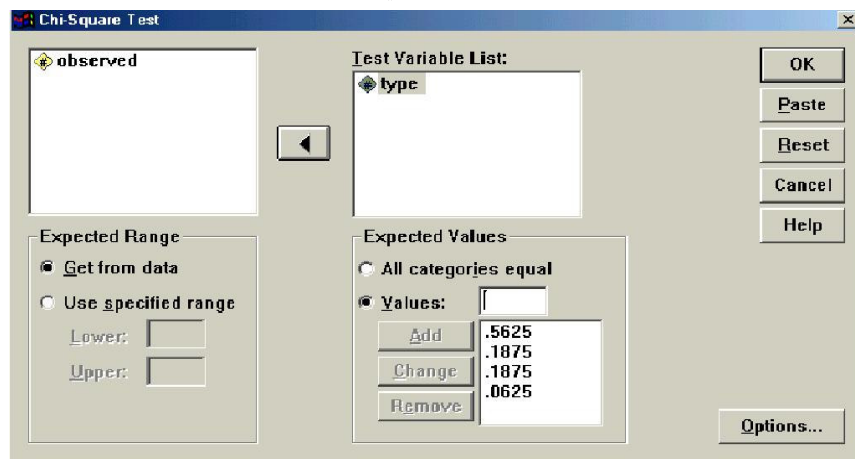
لاختبار الفرضية المذكورة نطبق اختبار Chi-Square وحسب الخطوات التالية :

◀ يتم ترتيب البيانات في ورقة Data Editor كما يلي :

type	Observed
1	439
2	168
3	133
4	60

لاحظ أننا استعصنا عن الصفات بالمتغير العددي Type لعدم إمكانية التعامل مع المتغيرات الرمزية أي عدم إمكانية ذكر أسماء الصفات مع العلم أنه بالإمكان إضافة عنوان القيمة Value Label الى المتغير Type لكي تعرض عناوين القيم في جداول المخرجات بدلاً من الأرقام التسلسلية. أما المتغير Observed فيمثل التكرار المشاهد في كل فئة وتستعمل قيم هذا المتغير كأوزان للحالات ويتم ذلك عن طريق اختيار Data → Weight Cases من شريط القوائم بعدها نؤشر الخيار Weight Cases by في صندوق حوار Weight Cases ثم إدخال المتغير Observed الى خانة الخيار الأخير (Weight Cases by) .

◀ من شريط القوائم اختر Chi-Square → Nonparametric Tests → Analyze فيظهر صندوق حوار Chi-Square Test الذي نرتبه كما يلي :



لاحظ أننا ادخلنا المتغير Type الى خانة Test Variable List .

في خانة Expected Values يوجد خيارين :

1. All categories equal : جميع الفئات لها نفس التكرار المتوقع أولها نفس النسبة .
2. Values : يتم تأشير هذا الخيار عندما يختلف التكرار المتوقع (وبالتالي تختلف النسب) من فئة الى أخرى كما هو الحال في هذا المثال فقد أدخلت النسب المحددة في فرضية العدم وكما يلي :

لأدخال القيمة الأولى $P_1 = \frac{9}{16} = 0.5625$ أنقر المستطيل المجاور لكلمة Value تم أكتب القيمة

0.5625 بعدها أنقر زر Add فتضاف أسفل الكلمة value (نفس الشيء لباقي القيم) .

◀ عند نقر زر OK نحصل على النتائج التالية :

TYPE

	Observed N	Expected N	Residual
1	439	np1 450	-11.0
2	168	np2 150	18.0
3	133	np3 150	-17.0
4	60	np4 50	10.0
Total	800		

الجدول أعلاه يمثل القيم المشاهدة والمتوقعة فمثلاً تحتسب القيمة المتوقعة للنوع الأول كما يلي
 $nP_1 = 800 * 0.5625 = 450$ لاحظ أن مجموع النسب يساوي 1 وعليه فإن حجم العينة للقيم المتوقعة هو نفسه للقيم المشاهدة .

Test Statistics

	TYPE
Chi-Square ^a	6.356
df	3
Asymp. Sig.	.096

a. 0 cells (.0%) have expected frequencies less than
 5. The minimum expected cell frequency is 50.0.

من الجدول أعلاه فإن قيمة إحصائية Chi-Square تساوي 6.356 وأن قيمة P-value المرافقة 0.096 تدعونا إلى قبول فرضية العدم بمستوى دلالة 5% أي قبول النسب الواردة في فرضية العدم
 → باعتبارها نسباً صحيحة
 يمكنك استخراج P-Value من الأمر Compute Transform ثم استعمال الدالة التالية لتوزيع مربع كاي :

$$1 - \text{CDF.CHISQ}(6.356, 3) = .096$$

مثال 2 :

رميت خمسة قطع نقود 1000 مرة وفي كل مرة حسبت عدد الصور Head وكانت النتائج كالتالي
 (بعد إدخالها في ورقة Data Editor) :

headno	observed
0	38
1	144
2	342
3	287
4	164
5	25

باستخدام اختبار Chi-Square أختبر الفرضية القائلة بأن عدد مرات ظهور الصورة يتبع توزيع Binomial وبمستوى دلالة 5% .

أن اختبار Chi-Square هنا يعرف باختبار حسن المطابقة Goodness of Fit. في البداية تقل (أوزن) الحالات بالمتغير Observed بنفس الطريقة المتبعة في المثال السابق ثم أتبع الخطوات التالية:
 ← من شريط القوائم اختر Chi-Square → Nonparametric Tests → Analyze فيظهر صندوق حوار Chi-Square Test (نفسه للمثال السابق) الذي نرتبه كما يلي :

- أنقل المتغير headno الى خانة Test Variable List .
- أستعمل دالة توزيع binomial لحساب احتمال ظهور الصورة من صفر الى خمسة وكانت النتائج كما يلي :

No. of heads	Probability
0	0.03125
1	0.15623
2	0.3125
3	0.3125
4	0.15623
5	0.03125

- أنقر الخيار values في خانة Expected values ثم أدخل النسب (الاحتمالات) أعلاه واحدة بعد الأخرى .
- عند نقر زر OK نحصل على النتائج التالية :

HEADNO

	Observed N	Expected N =1000*Pr	Residual
0	38	31.2	6.8
1	144	156.2	-12.2
2	342	312.6	29.4
3	287	312.6	-25.6
4	164	156.2	7.8
5	25	31.2	-6.2
Total	1000	1000.0	

يستخرج التكرار المتوقع للفئة الأولى مثلاً (عدد الصور يساوي صفر) كما يلي :

$$\text{Expected Frequency} = 1000 * 0.03125 = 31.25$$

Test Statistics

	HEADNO
Chi-Square ^a	8.920
df	5
Asymp. Sig.	.112

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 31.2.

أن قيمة $P\text{-value} = .112$ تدعونا الى قبول فرضية العدم بمستوى دلالة 5% أي أن عدد الصور المشاهد يتبع توزيع Binomial .

ملاحظة : يمكن أخذ مدى من الفئات بدلاً من كافة الفئات ففي المثال السابق سنهمل الفئة الأولى (عدد الصور = 0) وسنكتفي بالفئات الخمس المتبقية (عدد الصور = 1-5) وفي هذه الحالة سنكتفي بتزويد الاحتمالات لهذه الفئات فقط (1-5) في خانة Expected Values في صندوق حوار Chi-square Test فما دمنا قد أدخلنا الاحتمالات المقابلة لعدد الصور (0-6) في المثال السابق نقوم الآن بحذف الاحتمال

المقابل لعدد الصور =0 بتأشير هذا الاحتمال ويساوي 0.03125 في خانة Expected Values ثم نقر زر Remove .

بما أن الفئات المتبقية هي خمسة فئات (1-5) نقوم باختيار هذا المدى من الفئات في خانة Expected Range بتأشير الخيار Use Specified Range ثم إدخال قيمة Lower وتساوي 1 وقيمة Upper وتساوي 5 ،وعند نقر زر OK نحصل على النتيجة التالية :

Frequencies				
	HEADNO			
	Category	Observed N	Expected N	Residual
1	1	144	155.1	-11.1
2	2	342	310.4	31.6
3	3	287	310.4	-23.4
4	4	164	155.1	8.9
5	5	25	31.0	-6.0
Total		962	962	

(12 - 2) اختبارات عينتين مستقلتين *Two Independent Samples tests*

هذه الاختبارات مقارنة لاختبار T لمقارنة متوسطي عينتين مستقلتين حيث يستعمل اختبار T في حالة أن إحصائية الاختبار تتبع توزيع T عدا ذلك لا يمكن استعمال هذا الاختبار لعدم توفر شرط التوزيع الطبيعي لمجموعي العينتين ولذلك نلجأ إلى الاختبارات اللامعلمية حيث يوفر برنامج SPSS الاختبارات اللامعلمية التالية :

1. اختبار Mann-Whitney U
2. اختبار Kolmogorov-Smirnov Z
3. اختبار Moses Extreme Reactions
4. اختبار Wald-Wolfwitz Runs

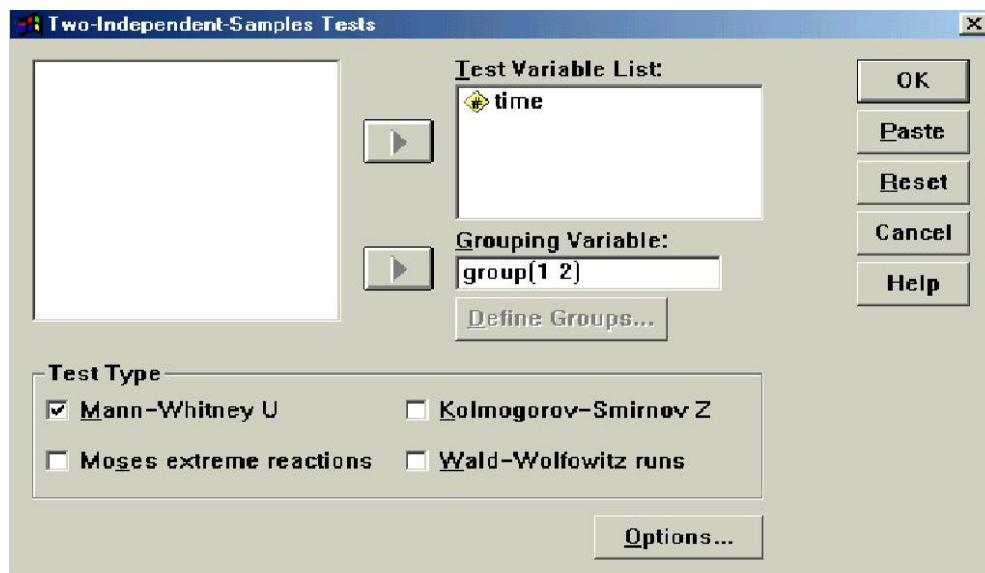
مثال 3: البيانات التالية تمثل عينتين مستقلتين الأولى بحجم 8 للأشخاص الأصحاء (يرمز لهم 1) والثانية بحجم 10 للأشخاص المرضى (الرمز 2) والمجموعتين من الأشخاص يضمها المتغير Group علماً أنه تم إضافة Value Label لهذا المتغير (القيمة 1 يقابلها العنوان Healthy والقيمة 2 يقابلها العنوان disease لكي تظهر العناوين بدلاً عن القيم في جداول الإخراج) وقد تم قياس الوقت المستغرق في فحص الجهد (المتغير Time) لكل شخص في العينتين وتظهر البيانات كما يلي في شاشة Data Editor لبرنامج SPSS :

time	group
1014	1
684	1
810	1
990	1
840	1
978	1
1002	1
1110	1
864	2
636	2
638	2
708	2
786	2
600	2
1320	2
750	2
594	2
750	2

يطلب اختبار فرضية عدم القائل بأن العينتين مسحوبتين من نفس المجتمع بمستوى دلالة 5% باستخدام اختبار Mann-Whitney (بمعنى آخر اختبار الفرضية عدم التالية $\mu_1 = \mu_2$ ضد البديلة $\mu_1 \neq \mu_2$).

لتنفيذ ذلك نتبع الخطوات التالية :

← من شريط القوائم اختر Analyze → Non Parametric Tests → 2 Independent samples
فيظهر صندوق حوار 2 Independent samples Test الذي نرتبه كما يلي :



لاحظ أننا أدخلنا المتغير Time (الذي يمثل مشاهدات العينتين) في خانة Test Variable List
أما المتغير group فقد أدخلناه في خانة Grouping Variable وهو متغير تجزئة لتعريف مجموعتي هذا المتغير انقر زر Define Group ثم عرف مجموعة الأصحاء 1 : Group1 ومجموعة المرضى 2 : Group2. أخيراً تم تأشير اختبار Mann-Whitney U .
← عند نقر زر OK تظهر النتائج التالية :

Ranks

	GROUP	N	Mean Rank	Sum of Ranks
TIME	healthy	8	12.63	101.00
	disease	10	7.00	70.00
	Total	18		

يعتمد هذا الاختبار على الرتب Ranks حيث يتم دمج العينتين (الأصحاء والمرضى) في عينة واحدة ثم ترتيب قيم العينة المدمجة تصاعدياً واعطاء رتبة لكل قيمة تم إيجاد مجموع رتب كل عينة (مجموع رتب الأصحاء 101 ومجموع رتب المرضى 70) كما في الجدول أعلاه ثم تحتسب إحصائية Mann-Whitney U بالصيغة التالية :

$$U = N_1 N_2 + \frac{N_1(N_1+1)}{2} - T_1$$

حيث أن N_1 يمثل حجم أي من العينتين (لنسمها الأولى $N_1=8$ و N_2 يمثل حجم العينة الأخرى (الثانية $N_2=10$) وأن T_1 يمثل مجموع رتب العينة الأولى (Sum of Ranks $T_1=101$). الجدول التالي يحتوي قيمة إحصائية U المحتسبة بالاعتماد على معطيات العينة الأولى بالإضافة الى إحصائية Wilcoxon W وهي مكافئة لإحصائية U . لاستخراج قيمة Z المرافقة يجب معرفة توزيع المعاينة لـ U حيث يستخرج المتوسط والانحراف المعياري للتوزيع كما يلي :

$$\mu = \frac{N_1 N_2}{2} = 40 \quad \sigma = \sqrt{\frac{N_1 N_2 (N_1 + N_2 + 1)}{12}} = 11.25$$

أن توزيع المعاينة لـ U يقارب التوزيع الطبيعي كلما كان حجم العينتين كبيراً وعليه يمكن استخراج القيمة المعيارية للتوزيع الطبيعي كما يلي $Z = (U - \mu) / \sigma = (15 - 40) / 11.25 = -2.222$ ويمكنك استخراج قيمة P -Value المرافقة لإحصائية Z بالأمر Compute Transform وباستعمال دالة التوزيع الطبيعي القياسي وكما يلي $2 * \text{CDFNORM}(-2.222) = 0.0263$ (الاختبار من طرفين) حيث أن قيمة P -value = 0.026 للتوزيع الطبيعي التقاربي تدعونا الى رفض فرضية العدم بمستوى دلالة 5% والأخذ بالفرض البديل أي أن العينتين لم تسحبا من نفس المجتمع أو أن متوسط الوقت المستغرق في فحص الجهد للأشخاص الأصحاء يختلف جوهرياً عن متوسط الوقت المستغرق للأشخاص المرضى .

Test Statistics^b

	TIME
Mann-Whitney U	15.000
Wilcoxon W	70.000
Z	-2.222
Asymp. Sig. (2-tailed)	.026
Exact Sig. [2*(1-tailed Sig.)]	.027 ^a

a. Not corrected for ties.

b. Grouping Variable: GROUP

ملاحظة : يمكنك الاعتماد على مجموع رتب العينة الثانية T_2 الكبيرة في حساب إحصائية U وفي هذه الحالة تحور إحصائية U كما يلي :

$$U = N_1 N_2 + \frac{N_2 (N_2 + 1)}{2} - T_2$$

N_2 يمثل حجم العينة الكبيرة وباستخراج القيمة المعيارية لإحصائية U وأجراء الاختبار يمكن

التوصل الى نفس النتيجة السابقة .

(12 - 3) اختبارات K من العينات المرتبطة K- Related Samples Tests

أن اختبار Mann-Whitney هو نسخة لامعلمية لاختبار T لعينتين مستقلتين وأن اختبار Kruskal-Wallis هو اختبار لا معلمي لتحليل التباين لمعيار واحد One- Way ANOVA. أما في حالة اختبار عينتين غير مستقلتين فيستعمل اختبار T لذلك وفي حالة عدم تحقق الشروط اللازمة لاختبار T أو ان

البيانات عبارة عن رتب Ranks في هذه الحالة يمكن إجراء أحد الاختبارات اللامعلمية التالية التي يوفرها برنامج SPSS :

1. اختبار Friedman .
2. اختبار Kendall's W .
3. اختبار Cochran's Q .

مثال 4 :

قام تسعة مختصين Therapists بتخصيص رتب لثلاثة نماذج للمحفزات الكهربائية a و b و c (الرتبة 1 تشير الى درجة التفضيل الأولى تليها الرتبتين 2 و 3) وكانت النتائج⁵ كما تظهر في شاشة Data Editor

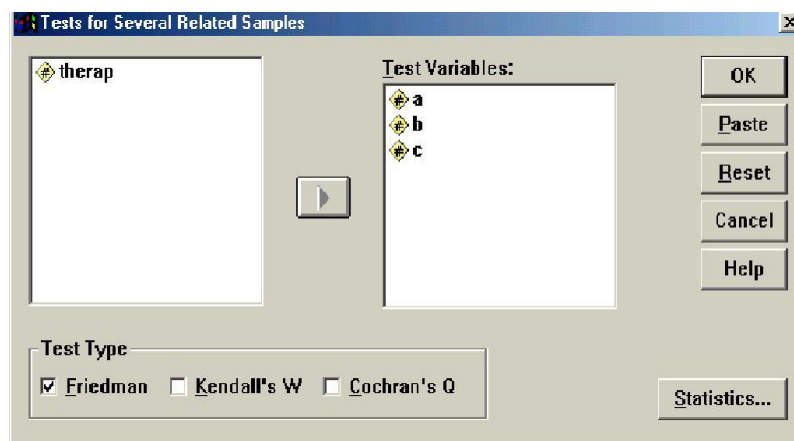
Therap	a	b	c
1	2	3	1
2	2	3	1
3	2	3	1
4	1	3	2
5	3	2	1
6	1	2	3
7	2	3	1
8	1	3	2
9	1	3	2

لبرنامج SPSS :

المطلوب اختبار فرضية العدم القائلة بعدم وجود فروقات معنوية في درجة التفضيل للنماذج الثلاثة باستخدام اختبار Friedman وبمستوى دلالة 5% .

لتنفيذ ذلك نتبع الخطوات التالية :

☐ من شريط القوائم اختر Analyze → Nonparametric Tests → K-Related samples فيظهر صندوق حوار Tests for several Related Samples الذي نرتبه كما يلي :



☐ عند نقر زر OK نحصل على النتائج التالية :

⁵Daniel W.W.(1978) , Biostatistics : A Foundation for analysis in The Health Sciences , 2nd Edition, pp.398 .

NPar Tests Friedman Test

Ranks

	Mean Rank
A	1.67
B	2.78
C	1.56

حيث أن Mean Rank تمثل متوسط الرتب في كل من النماذج الثلاثة . أن التجربة أعلاه تشبه تحليل التباين لمعيارين Two- way ANOVA أو ما يعرف بتصميم القطاعات العشوائية Randomized Blocks Experiment حيث تمثل الأعمدة (أو النماذج) المعالجات Treatments وان الصفوف تمثل القطاعات Blocks حيث يعتبر اختبار Friedman ملائماً لأجراء تحليل التباين لمعيارين حسب الرتب Ranks ويتطلب أن تكون البيانات ترتيبية Ordinal في الأقل وأن إحصائية Friedman المستخدمة تتبع توزيع χ^2 بدرجة حرية J-1 وتحتسب كما يلي:

$$\chi^2 = \frac{12}{KJ(J+1)} \left[\sum T_i^2 \right] - 3K(J+1) = 8.2$$

حيث أن K يمثل عدد الصفوف (القطاعات) وأن J يمثل عدد المعالجات (الأعمدة) وأن T_i يمثل مجموع الرتب لكل معالجة أي أن $K=9$ و $J=3$.

الجدول أدناه يبين قيمة إحصائية Friedman وقيمة P-Value المرافقة التي تدعونا الى رفض فرضية العدم بمستوى دلالة 5% أي أن هناك فروقاً معنوية في درجة التفضيل للنماذج الثلاثة .

Test Statistics^a

N	9
Chi-Square	8.222
df	2
Asymp. Sig.	.016

a. Friedman Test

الفصل الثالث عشر

المخططات البيانية

CHARTS

تعتبر المخططات البيانية أداة مهمة من أدوات الإحصاء الوصفي والتي يمكن بواسطتها عرض البيانات الإحصائية بطريقة مبسطة ومعبرة . من المخططات المهمة الأعمدة البيانية Bars والخطوط البيانية Lines والدوائر البيانية Pies... سنحاول في هذا الفصل أن نستعرض الإمكانيات والتسهيلات التي يوفرها برنامج SPSS في مجال معالجة الرسوم البيانية .

Bar Charts الأعمدة البيانية (1-13)

مثال 1 :

الجدول التالي يبين المبيعات Sales لإحدى المؤسسات حسب السنوات Year :

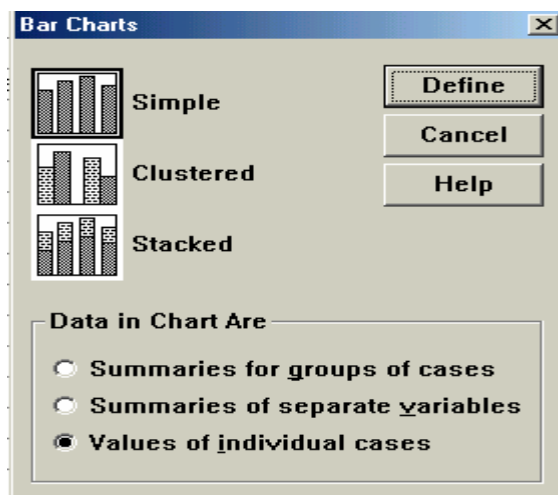
year	sales
1990	50
1991	52
1992	55
1993	60
1994	65

يطلب أعداد مخطط الأعمدة البيانية Bar Chart .

لتنفيذ ذلك نتبع الخطوات التالية :

من شريط القوائم اختر Bar → Graphs. فيظهر صندوق حوار Bar Chart وقد

رتبناه كالتالي:

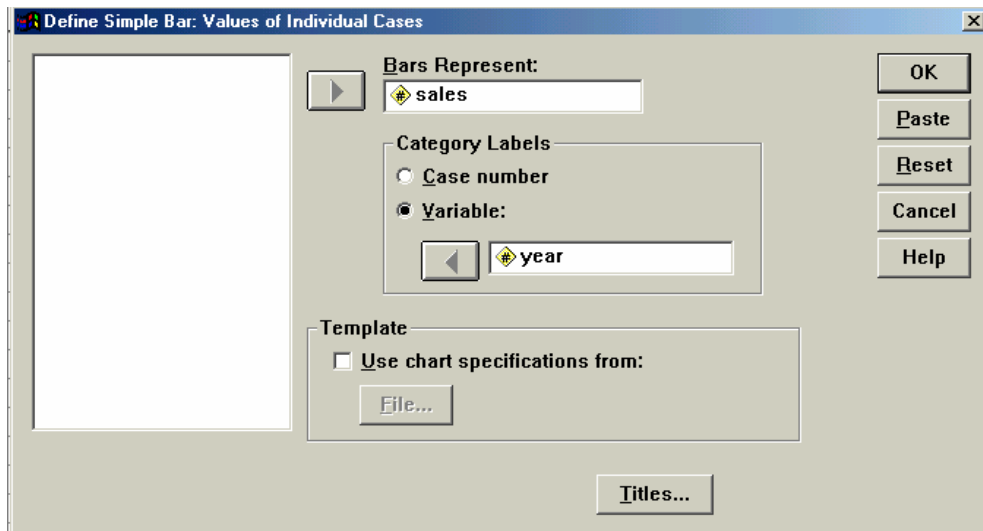


وقد اخترنا Values of individual cases في قائمة Data in Chart are ويقابل كل خيار في هذه

القائمة ثلاثة خيارات وهي Simple ، Clustered ، Stacked وقد اخترنا Simple وهو أبسط الحالات .

عند نقر زر Define في صندوق الحوار السابق يظهر صندوق حوار Define Simple Bar

الذي نرتبه كالتالي :



. حيث أن :

Bars Represent : متغير عددي كل قيمة من قيمه تمثل بشريط في المخطط .

Category Labels : وهو عناوين الفئات للمخطط البياني (المحور السيني) ويتضمن ما يلي :

Case number : يعرض رقم الحالة كعنوان للقيمة على المحور السيني .

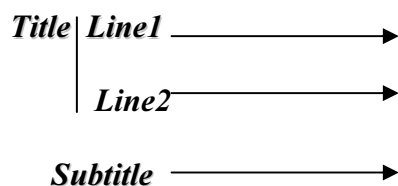
Variable : يعرض قيم المتغير الموجود في هذه القائمة كعناوين لقيم المتغير على المحور السيني (في هذا المثال المتغير Years) .

Title : لعرض عناوين المخطط Title , subtitle , Footnote .

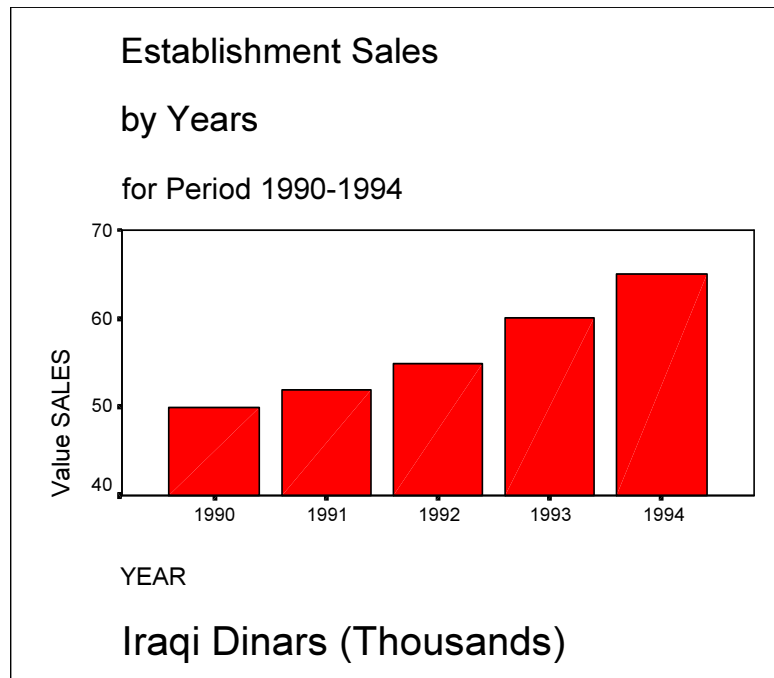
Template : لعمل قالب يحتوي مواصفات معينة يمكن تطبيقها مباشرة عند أعداد مخططات أخرى .

عند نقر زر OK يتم عرض المخطط في SPSS Viewer كما يلي :

مخطط رقم 1

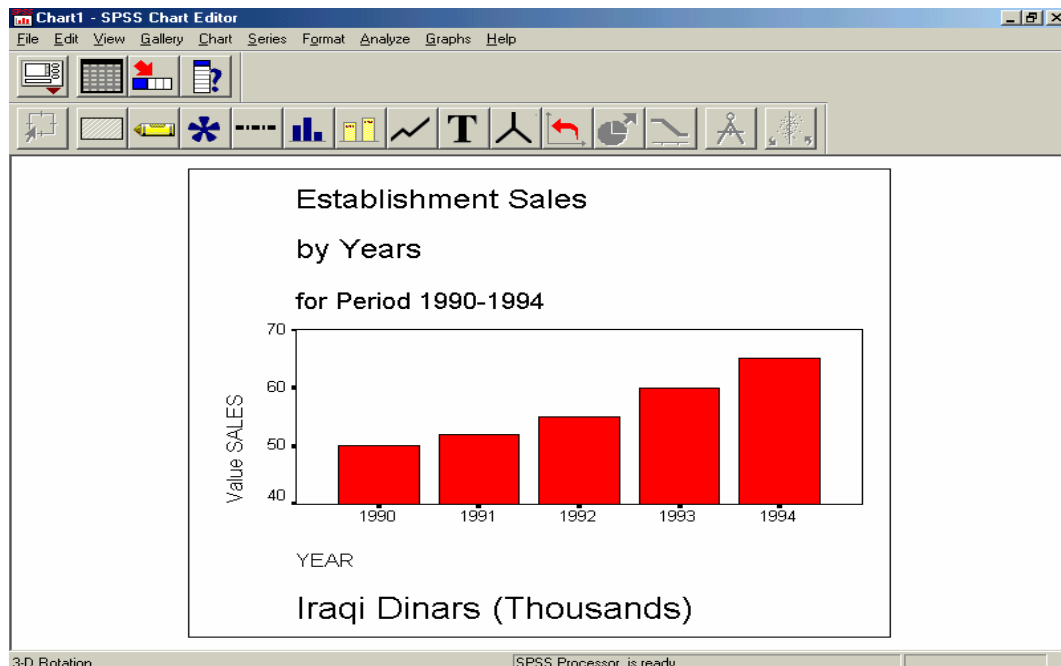


Footnote —————>



يمكن إجراء تعديلات على المخطط بنقره مرتين بزر الماوس الأيسر حيث يعرض المخطط في SPSS Chart Editor كما يلي :

شاشة تنقيح المخططات SPSS Chart Editor



حيث يمكن إجراء التعديلات التالية (كافة التعديلات التي ستجرى لاحقاً تتم في شاشة Chart Editor أي بعد نقر المخطط مرتين بزر الماوس الأيسر) :

1. تغيير لون الأعمدة : يتم تنفيذ ذلك كما يلي (كما ذكرنا أن التعديلات تجري في شاشة Chart Editor كما في الشكل أعلاه) :

➤ أنقر الأعمدة بزر الماوس الأيسر .

➤ من شريط القوائم في شاشة Chart Editor اختر Color → Format أو أنقر أيقونة

في شريط الأدوات فيظهر صندوق حوار Colors كما يلي :



حيث أن :

Fill : لتغيير لون إملاء الشريط .

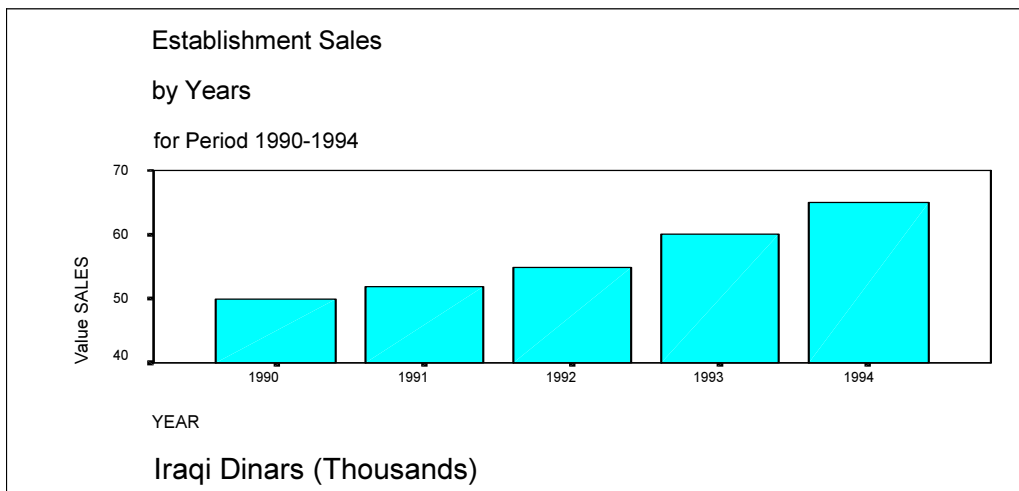
Border : لتغيير لون الحدود الخارجية للشريط .

◀ لتغيير لون إملاء الشريط الى اللون الأزرق مثلاً اختر اللون الأزرق بنقره بزر الماوس الأيسر بعد أن تتأكد من تأشير الخيار Fill ثم أنقر زر Apply فيتغير لون الأعمدة .

◀ أنقر زر Close لغلاق صندوق حوار Colors . ويظهر المخطط بعد تغيير لون الأعمدة كما

يلي :

مخطط رقم 2



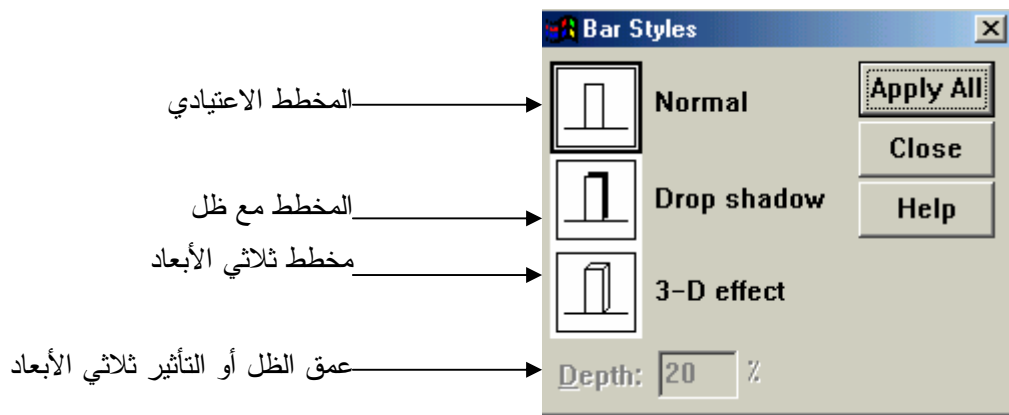
ملاحظة :

يمكن تغيير لون الخلفية مثلاً بنفس الخطوات السابقة (عدا الخطوة الأولى حيث نقوم بنقر خلفية المخطط بدلاً من نقر الأعمدة) .

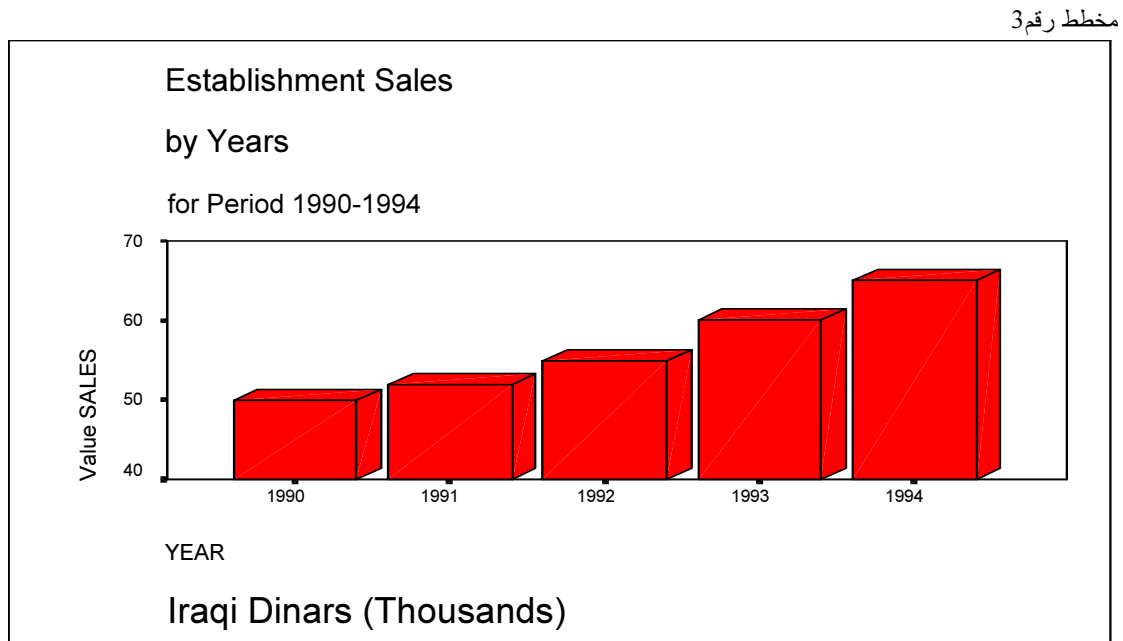
2. تغيير نمط الأعمدة Bar Style : لتنفيذ ذلك نتبع مايلي :

◀ من شريط القوائم في شاشة Chart Editor اختر Bar Style → Format أو أنقر

الأيقونة  في شريط الأدوات فيظهر صندوق الحوار التالي :



عند اختيار 3-D effect بنقره بزر الماوس الأيسر ثم نقر زر Apply All تظهر الأعمدة ثلاثية الأبعاد كما يلي :

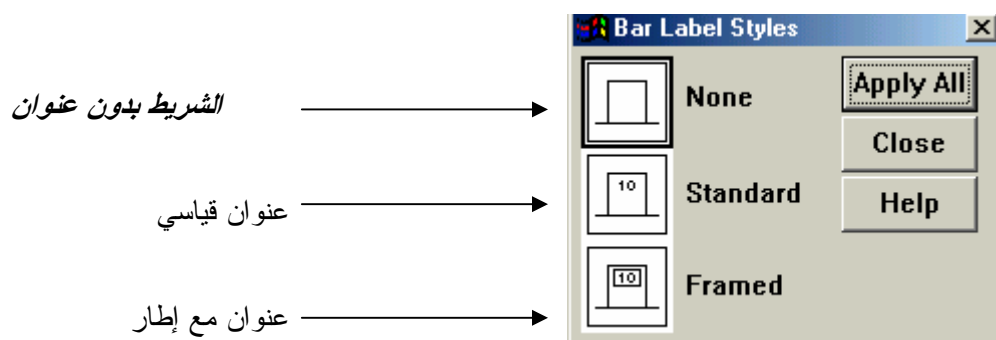


ملاحظة :

عند تحديد رقم موجب في Depth يتم إضافة التأثير الى يمين الشريط وعند تحديد رقم سالب يضاف التأثير الى يسار الشريط .

3. تغيير نمط عناوين الأعمدة Bar Label Style: لتنفيذ ذلك نتبع الخطوات التالية :

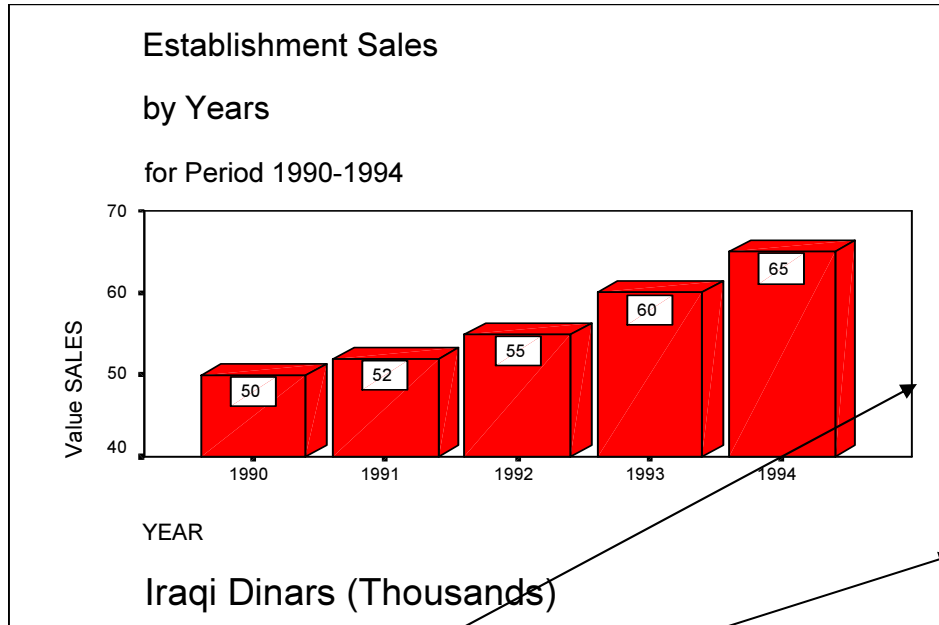
من شريط الأدوات اختر Bar Label Style → Format أو أنقر الأيقونة  فيظهر صندوق



الحوار التالي :

أختر Framed بنقره بزر الماوس الأيسر ثم أنقر زر Apply فنحصل على المخطط التالي :

مخطط رقم 4



الإطار الداخلي الإطار الخارجي

ملاحظات :

أ. لإزالة الإطار الخارجي للمخطط أختر Chart من شريط القوائم في شاشة Chart Editor ثم أنقر Outer Frame لإزالة العلامة منها .

ب. لإزالة الإطار الداخلي للمخطط أختر Chart من شريط القوائم في شاشة Chart Editor ثم أنقر Inner Frame لإزالة العلامة منها .

4. تغيير الإعدادات الخاصة بالمحور الرأسي Scale Axis

تكون خطوات تغيير الإعدادات كما يلي :

أنقر مرتين على عناوين المحور الرأسي (40،50،60...) للمخطط في شاشة Chart Editor فيظهر صندوق حوار Scale Axis كما يلي :

حيث أن :

- a. Display Axis Line: عرض أو إزالة خط المحور الرأسي (لإزالة المحور الرأسي يتوجب قبلها إزالة الإطار الداخلي بالأمر Chart ثم Inner Frame) .
- b. Title Justification: لتغيير موقع عنوان المحور الرأسي (Value Sales) .
- c. Scale: لتحديد نوع المقياس (خطي، لوغاريتمي) .
- d. Major divisions: التقسيمات الرئيسية .
- e. Minor divisions: التقسيمات الثانوية .
- f. Increment: المسافة بين تقسيمين علماً أن $\geq \text{Major Increment} \text{ Minor Increment}$
- g. Grid: عرض خطوط الشبكة .
- h. Ticks: لعرض Tick Marks .
- i. Display Labels: عرض أو إزالة عناوين التقسيمات (60, 50, 40, ...)

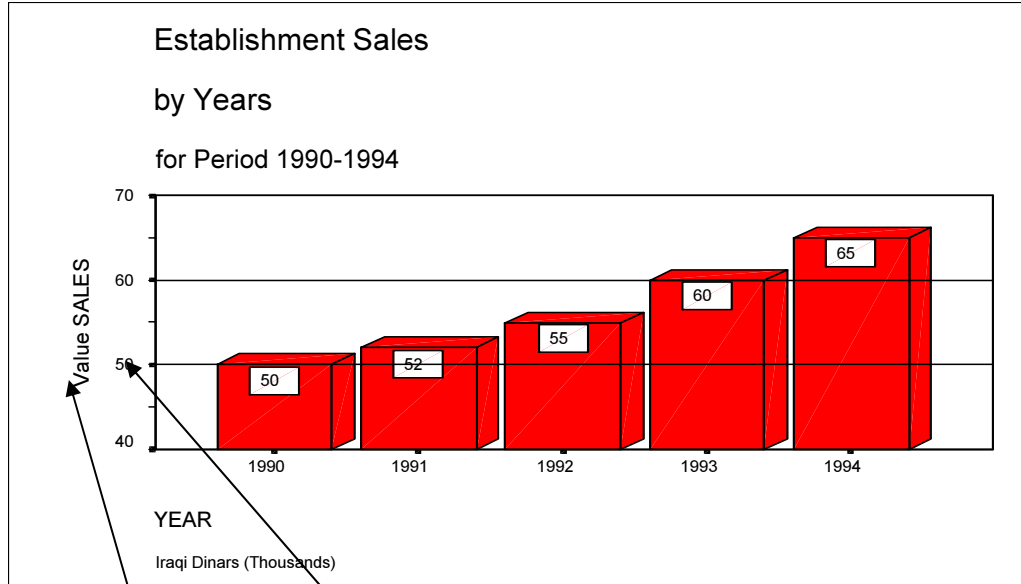
إذا أردنا إجراء الفعاليات التالية على المخطط رقم 4 :

- وضع عنوان المحور الرأسي (الذي هو Value Sales) في الوسط Center .
- جعل المسافة بين التقسيمات الثانوية Minor Increment هي 5 بدلاً من 10 مع إظهار علامات التقسيم الثانوية Tick Marks .
- إضافة خطوط الشبكة Grid Lines إلى التقسيمات الرئيسية .

فإن صندوق حوار Scale Axis يرتب كالتالي (نحصل عليه بنقر عناوين المحور الرأسي مرتين ثم نقوم بالتعديلات اللازمة) :

أما المخطط فيظهر كما يلي بعد نقر الزر OK :

مخطط رقم 5

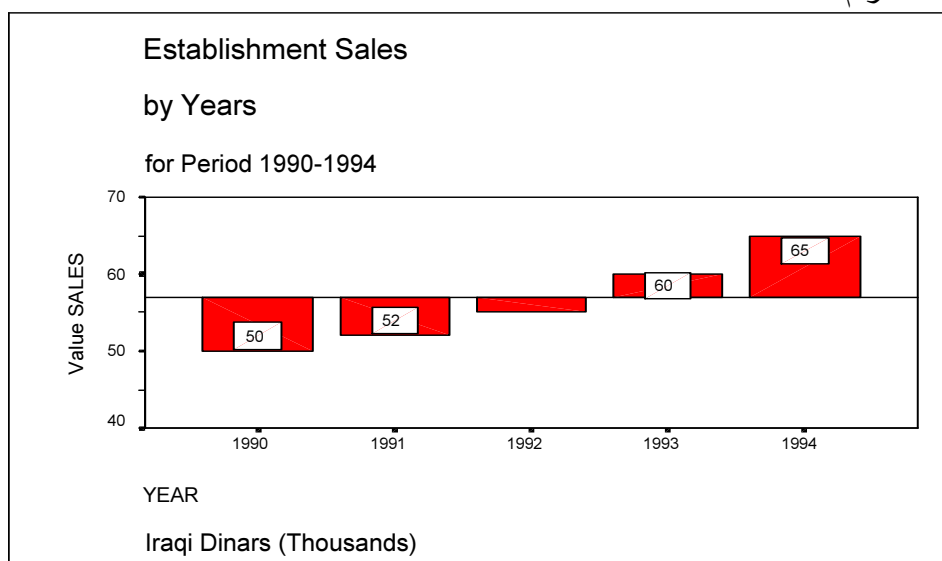


Axis Title

Axis Label

- J. Labels : عند نقر هذا الزر يظهر صندوق حوار Labels ومن خلاله يمكن القيام بالعمليات التالية:
- Decimal Places : لتحديد عدد المراتب العشرية لعناوين المحور الرأسي فإذا حددنا عدد المراتب العشرية يساوي واحد فستعرض العناوين كما يلي : 40.0، 50.0، 60.0، 70.0 .
 - Leading Character : وهو رمز يضاف الى بداية كل عنوان من عناوين المحور الرأسي فإذا أردنا إضافة رمز الدينار العراقي D فستظهر العناوين كما يلي : D40 ، D50 ، D60 ، D70 .
 - Trailing Character : وهو رمز يضاف الى نهاية كل عنوان من عناوين المحور الرأسي فإذا أردنا إضافة علامة % فستظهر العناوين كما يلي : 40%، 50%، 60%، 70% .
 - 1000 S Separator : لعرض القيم التي تزيد عن 1000 بفواصل (فارزة Comma أو فترة . (Period

K. Bar Origin Line : لرسم خط الأصل Origin Line وهو خط يظهر في المخطط البياني عند قيمة محددة بحيث أن الأعمدة التي تمثل قيماً أعلى من قيمة خط الأصل سوف تمتد فوق الخط أما الأعمدة التي تمثل قيماً أقل من قيمة الخط فسوف تمتد تحته. إذا حددنا قيمة Bar Origin Line تساوي 57 في صندوق حوار Scale Axis للمخطط رقم 4 (بعد إزالة التأثير الثلاثي الأبعاد) فسوف نحصل على المخطط التالي :



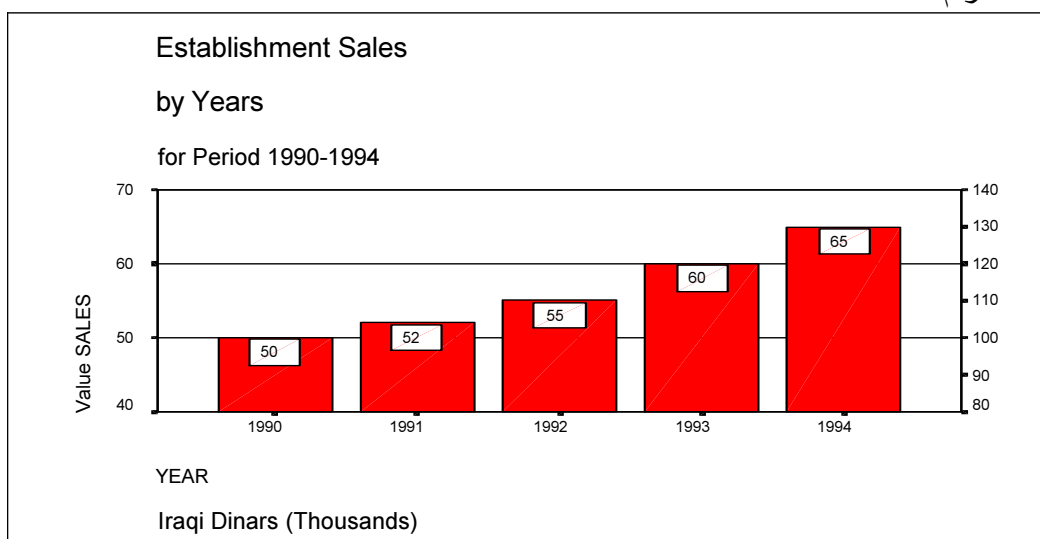
حيث أن الخط الوسطي Origin Line عند النقطة 57 .

L. Display Derived Axis : لعرض محور رأسي إضافي في الاتجاه المقابل للمحور الرأسي الأصلي ويكون له مقياس مختلف ويتم ذلك بتأشير Display Derived Axis في صندوق حوار Scale Axis ثم نقر زر Derived Axis فيظهر صندوق حوار Derived Axis كما يلي :

فاذا أردنا جعل كل وحدة من المحور الأصلي Scale Axis تقابل وحدتين من Derived Axis للمخطط رقم 1 فأن قائمة Definition في صندوق الحوار أعلاه ترتب كما يلي :

	Scale Axis		Derived Axis	
Ratio	1	Units Equal	2	Units
Match	0	Value Equal	0	Value

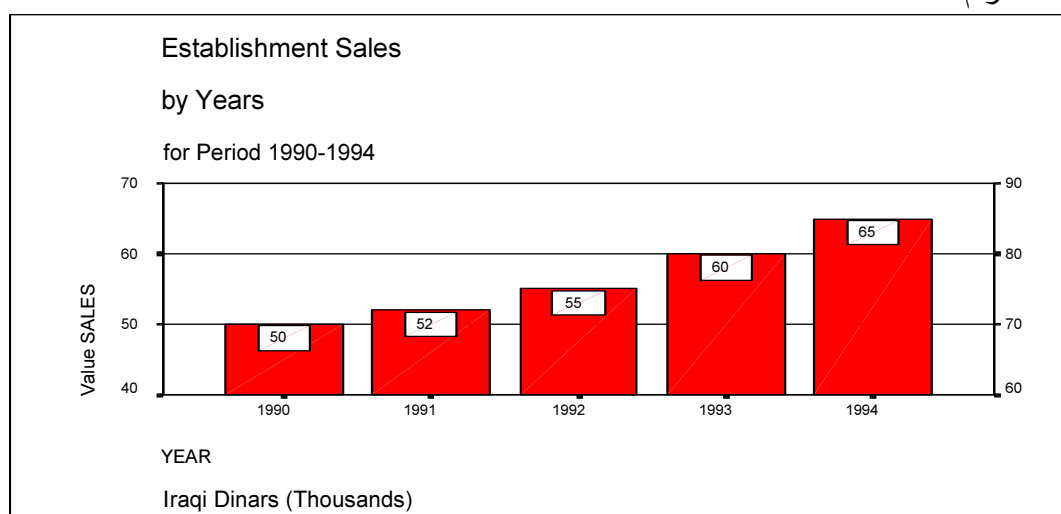
عند نقر زر Continue ثم زر OK نحصل على المخطط التالي :



وقد أضفنا العناوين Labels وكذا الشبكة لغرض التوضيح. نلاحظ أن كل قيمة على المحور الأصلي يقابلها الضعف على المحور المشتق Derived Axis .
 أما الخيار Match فيسمح بتحديد قيمتين مختلفتين (أو متساويتين) على كل من المحور الأصلي والمشتق بحيث تظهران في نفس الموقع. فإذا أردنا جعل القيمة 50 على المحور الأصلي تقابل القيمة 70 على المحور المشتق للمخطط 1 نرتب قائمة Definition في صندوق حوار Derived Axis كما يلي :

	Scale Axis		Derived Axis	
Ratio	1	Units Equal	1	Units
Match	50	Value Equal	70	Value

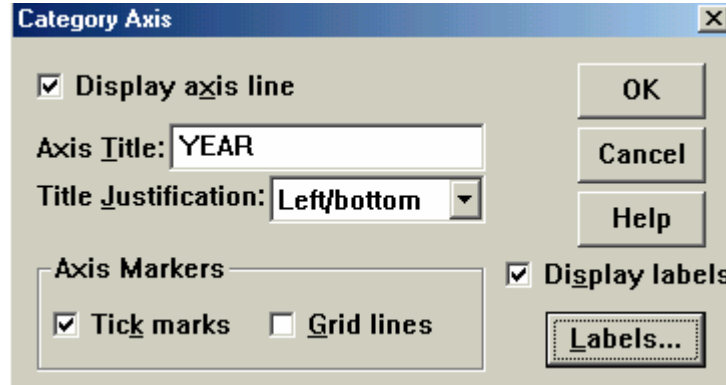
عند نقر زر Continue ثم زر OK نحصل على المخطط التالي :



نلاحظ أن القيمة 50 تقابل القيمة 70 والقيمة 60 تقابل القيمة 80 ... وهكذا .

5. تغيير الإعدادات الخاصة بالمحور الأفقي Category Axis

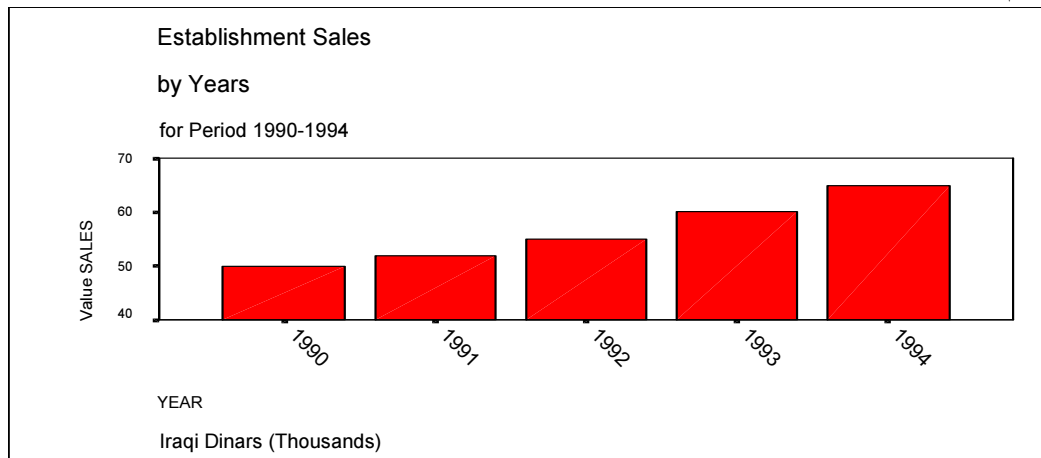
يمكن تغيير الإعدادات الخاصة بالمحور الأفقي بنقر التخطيط مرتين بزر الماوس الأيسر في SPSS Viewer ثم نقر عناوين المحور الأفقي للمخطط مرتين وهو في شاشة Chart Editor فيظهر صندوق الحوار التالي (للمخطط رقم 1 مثلاً) :



حيث يمكن القيام بالفعاليات التالية :

- عرض أو إزالة المحور الأفقي Display Axis Line (لإزالة المحور الأفقي يتوجب قبلها إزالة الإطار الداخلي بالأمر Chart ثم Inner Frame) .
- تغيير عنوان المحور بتظليل العنوان القديم في المستطيل المجاور لـ Axis Title ثم كتابة العنوان الجديد .
- موقع العنوان Title justification
- إظهار أو إزالة التقسيمات أو خطوط الشبكة .
- عند نقر زر Labels يظهر صندوق حوار Labels وفيه يمكن عرض عناوين المحور الأفقي جميعها أو جزء منها ، تغيير العناوين الموجودة ، التحكم بزاوية الميلان أو موقع العناوين Orientation حيث تتوفر الخيارات التالية: Horizontal ، Vertical ، Diagonal ، Staggered . فعند اختيار Diagonal لعرض عناوين المحور الأفقي للمخطط 1 فسيظهر المخطط كما يلي :

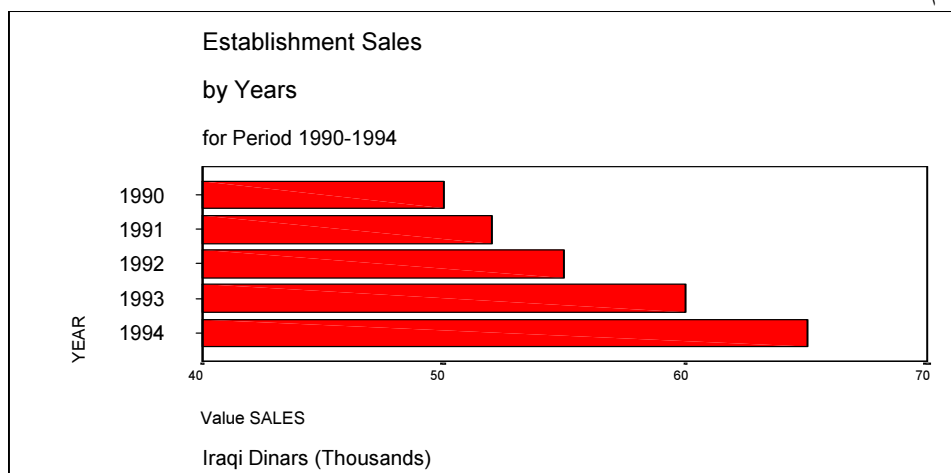
مخطط رقم 9



يكون من الضروري عرض العناوين الطويلة بهذه الطريقة تجنباً لتشابك العناوين فيما بينها كما يمكن الاستفادة من الخيارين Staggered و Vertical في هذه الحالة .

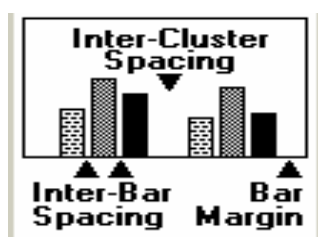
6. قلب المحاور Swap Axis : لقلب المحاور ورأي تحويل المخطط من أعمدة الى أشرطة وبالعكس من شريط القوائم نختار Format → Swap Axis أو نقر الأيقونة  في شريط الأدوات فمثلاً عند قلب محاور المخطط رقم 1 فإنه يظهر كما يلي :

مخطط رقم 10



ملاحظات :

- يمكن تغيير نوع الخط Font وحجمه Size لأي عنوان في المخطط (عندما يكون في شاشة SPSS Chart Editor) بنقر العنوان ثم اختيار Text → Format من شريط القوائم أو نقر الأيقونة .
- في حالة عدم ظهور الكتابة باللغة العربية أنقر المخطط مرتين للتحويل الى شاشة SPSS Chart Editor ثم أنقر الأيقونة  بعد نقر النص المطلوب بزر الماوس الأيسر ومن قائمة Font اختر نوع الخط (Arial(Arabic) أو Andalus أو Arabic Transparent).
- يمكن تغيير نمط الإملاء للأشربة بنقر أشربة المخطط (عندما يكون في شاشة SPSS Chart Editor) ثم اختيار Fill Pattern  من شريط القوائم أو نقر الأيقونة .
- يمكن إضافة وسيلة إيضاح Legend باختبار Legend → Chart من شريط القوائم في شاشة Chart Editor ثم تأشير Display Chart Legend في صندوق حوار Legend.
- يمكن التحكم بالمسافات بين الأعمدة وبين كل من العمود الأول والأخير وبين الإطار الداخلي للمخطط Inner Frame وكذلك بين العناقيد Cluster باختبار Bar Spacing → Chart من شريط الأدوات كما في الشكل التالي :



الجدير بالذكر أنه يمكن التحكم بعرض الأشربة بواسطة هذا الأمر.

(2-13) **عمل قالب لمخطط بياني Chart Template :** من الواضح أننا ننفذ عمليات مختلفة على المخطط الى أن يأخذ شكله النهائي وفي بعض الأحيان يتطلب الأمر تنفيذ نفس العمليات ولكن على سلسلة من المتغيرات وهذه تعتبر عملية مطولة خاصة إذا كان هناك عدد كبير من المتغيرات وفي هذه الحالة يكون بالإمكان عمل قالب Template يتضمن كافة العمليات التي أجريت على المخطط الأول حيث يمكن تطبيقه على متغيرات أخرى. فمثلاً المخطط رقم 4 يتضمن عمليات مختلفة (ثلاثي الأبعاد ، عناوين الأعمدة بإطار ، العناوين الأصلية والفرعية ...). فلغرض تطبيق نفس هذه العمليات على متغير آخر x (الذي يأخذ القيم التالية 80،88،110،90،77 للسنوات 90-94) نتبع الخطوات التالية :

➤ **خزن المخطط 4 بصيغة قالب بالأمر Save Chart Template** → File في شاشة Chart

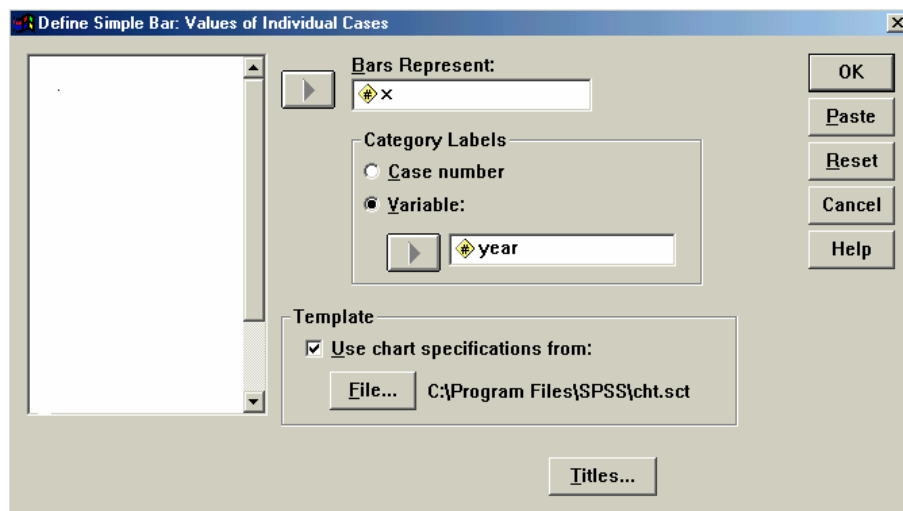
Editor حيث يأخذ القالب الاستطالة sct وليكن أسم القالب cht .

➤ **من شريط القوائم اختر Bar** → Graphs ثم اختر Values of Individual Cases

(simple) ثم أدخل المتغيرات x و year في صندوق حوار Values of Individual Cases أدناه .

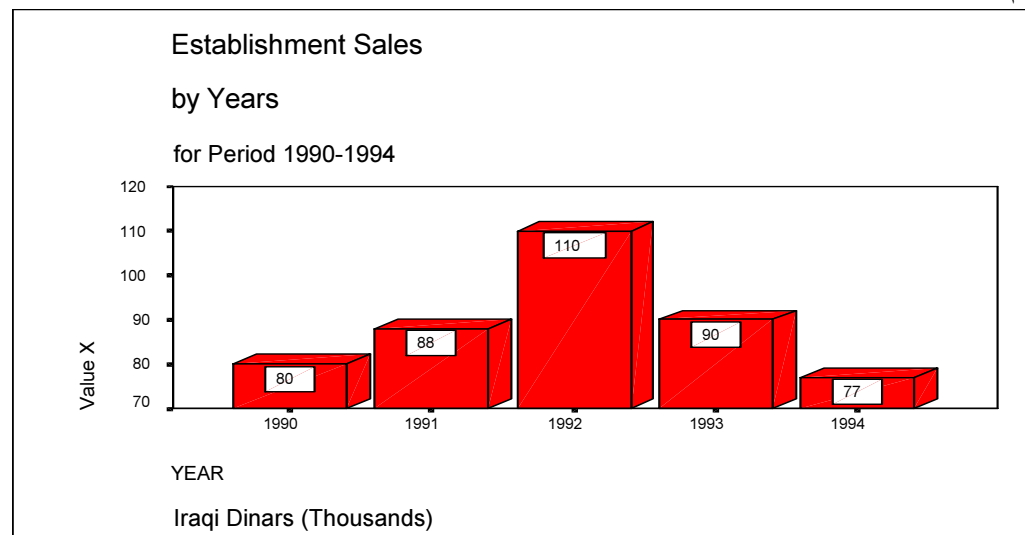
➤ **في خانة template قم بتأشير Chart Specification from** ثم انقر زر File لفتح القالب

cht حيث يتم ترتيب صندوق حوار Values of Individual Cases كما يلي :



➤ عند نقر زر Ok نحصل على المخطط التالي للمتغير x :

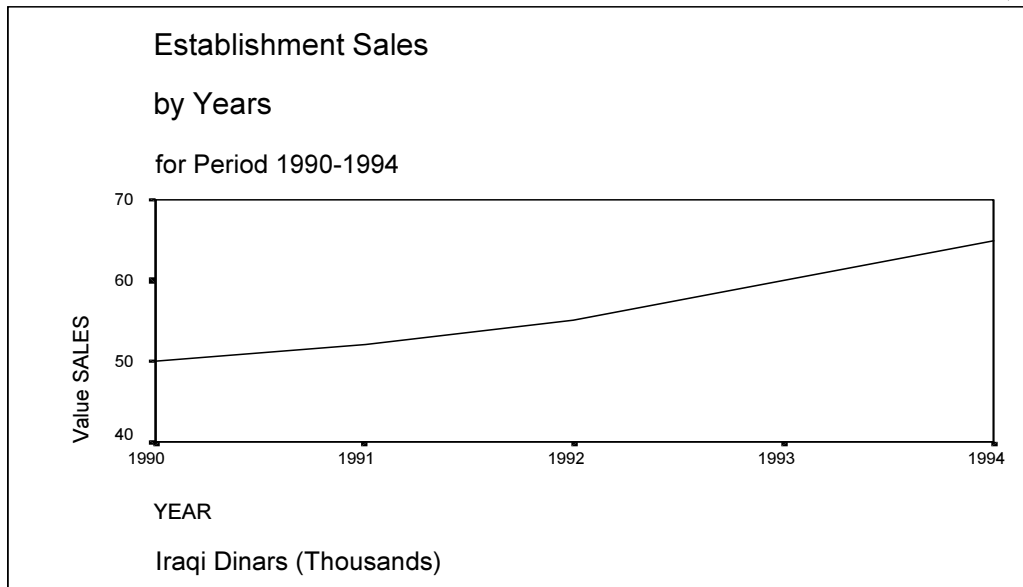
مخطط رقم 11



(13-3) تغيير مخطط Bar الى مخطط Line

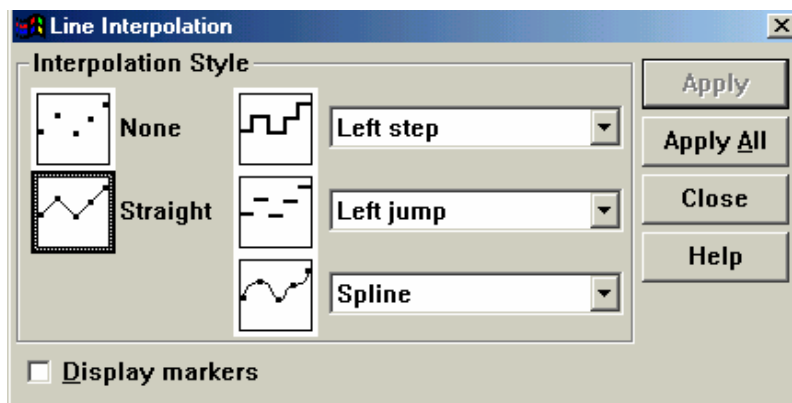
لتغيير المخطط رقم 1 من أعمدة Bar الى خطوط Line نتبع الخطوات التالية :

- انقر المخطط مرتين ليكون في شاشة Chart Editor .
 - من شريط القوائم اختر Simple → Line → Gallery فيظهر صندوق حوار Line Charts (اخترنا Simple لان المخطط يتضمن سلسلة واحدة فقط Sales في حالة وجود أكثر من سلسلة (أو متغير) ونريد تضمينها جميعها نختار Multiple) .
 - انقر Replace فيتحول المخطط الى Line كما يلي :
- مخطط رقم 12

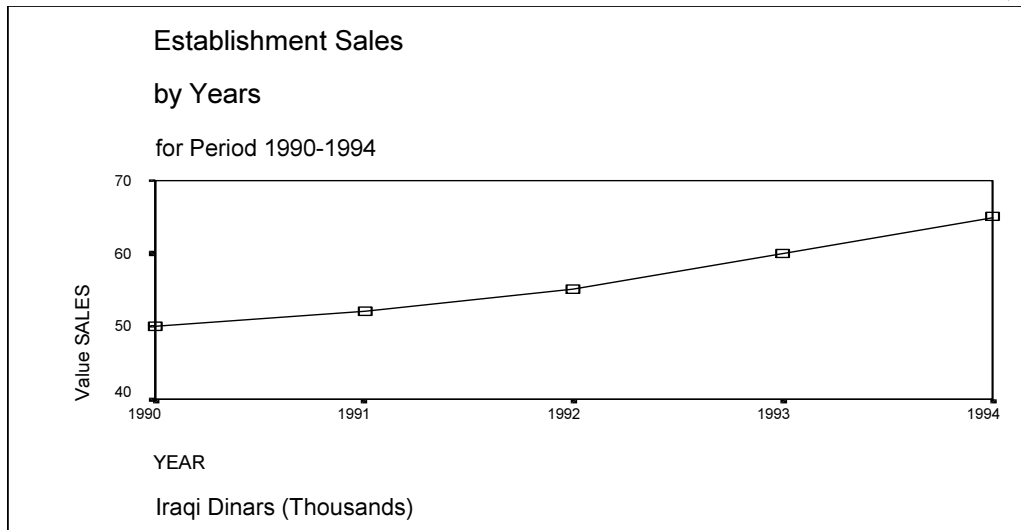


لعرض العلامات Markers على الخط البياني نتبع الخطوات التالية :

- من القوائم اختر Interpolation → Format أو انقر الأيقونة  في شريط الأدوات (عندما يكون المخطط في شاشة Chart Editor) فيظهر صندوق حوار Interpolation كما يلي :

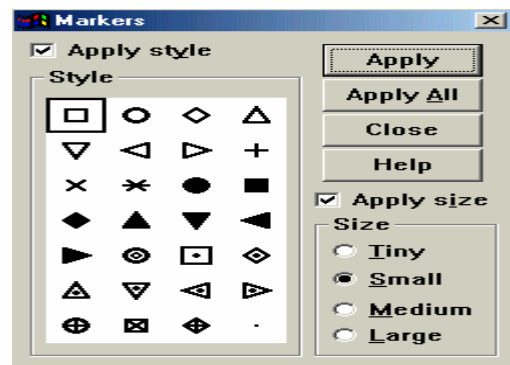


- انقر المربع الصغير Check Box المجاور لـ Display Markers لتأثيره ثم انقر زر Apply
- All فتضاف العلامات الى المخطط كما في الشكل التالي :



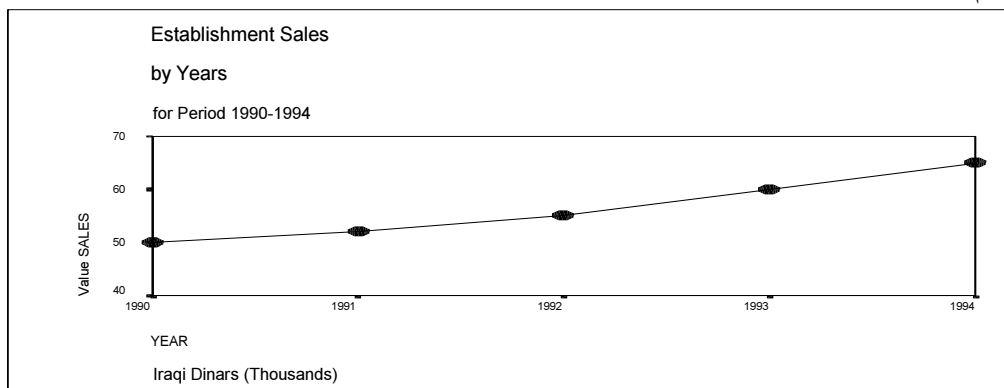
لتغيير شكل العلامة نتبع ما يلي :

- انقر المخطط مرتين ليكون في شاشة Chart Editor ثم انقر العلامات لتأشيرها .
- من القوائم اختر Markers → Format أو انقر الأيقونة  فيظهر صندوق حوار Markers التالي :



➤ اختر نوع العلامة وحجمها .

➤ انقر زر Apply فيتم تغيير العلامة كما يلي :



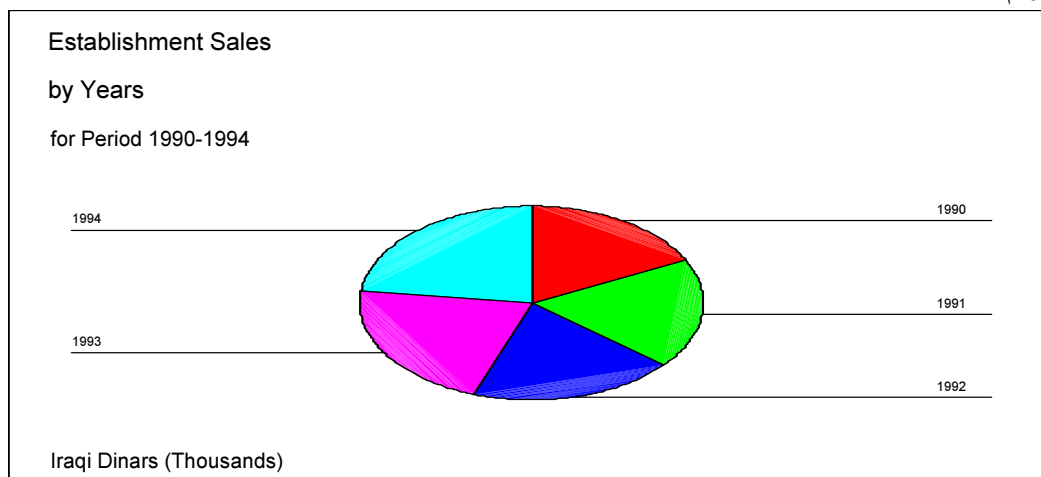
ملاحظة : في حالة وجود أكثر من سلسلة (متغير) ونريد تغيير شكل العلامة لها جميعاً انقر زر Apply All .

(13 - 4) تغيير مخطط Bar الى مخطط Pie

لتغيير المخطط رقم 1 من أعمدة Bar الى دائرة بيانية Pie نتبع الخطوات التالية :

- انقر المخطط مرتين ليكون في شاشة Chart Editor .
- من شريط القوائم اختر Pie → Gallery فيظهر صندوق حوار Pie Charts
- اختر Simple ثم انقر Replace فيتحول المخطط الى Pie كما يلي :

مخطط رقم 15

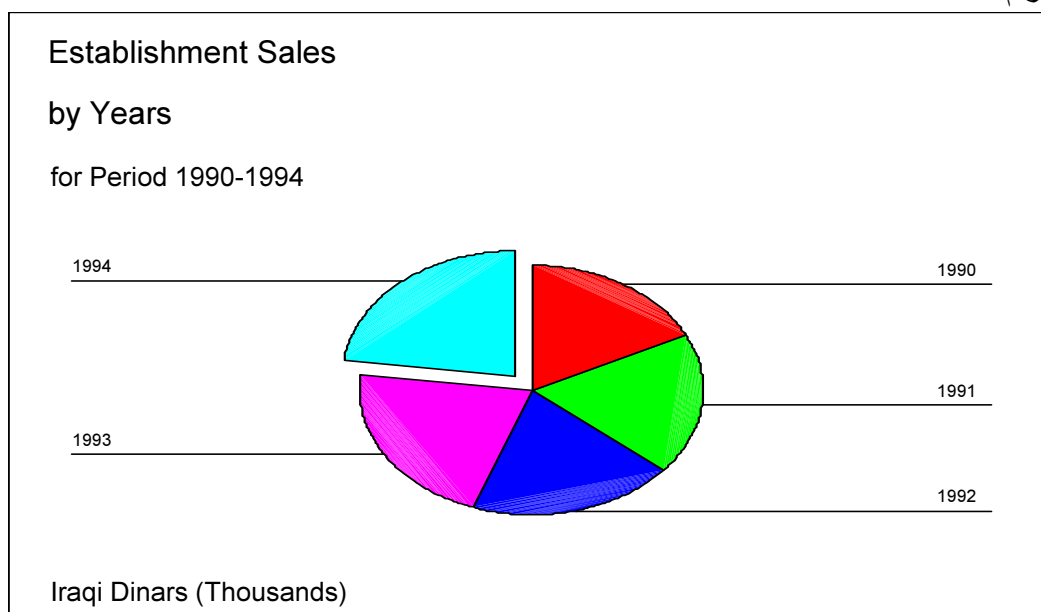


لفصل مقطع من الدائرة الخاص لسنة 1994 مثلاً نتبع الخطوات التالية :

- انقر المقطع الخاص بسنة 1994 لتحديده بعد أن تتأكد أن المخطط معروض في شاشة Chart Editor .

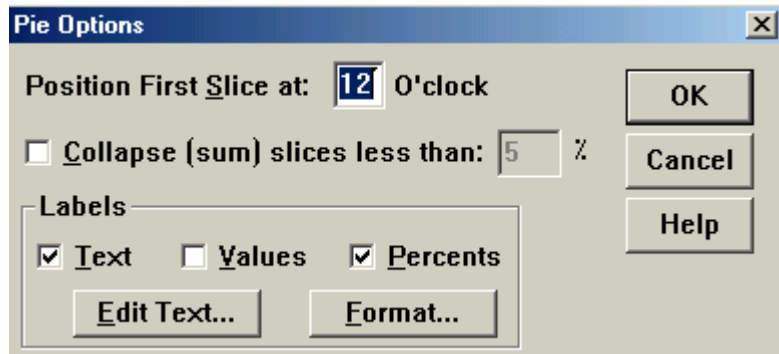
- من شريط القوائم اختر Explode Slice → Format أو انقر الأيقونة  فنحصل على المخطط التالي :

مخطط رقم 16

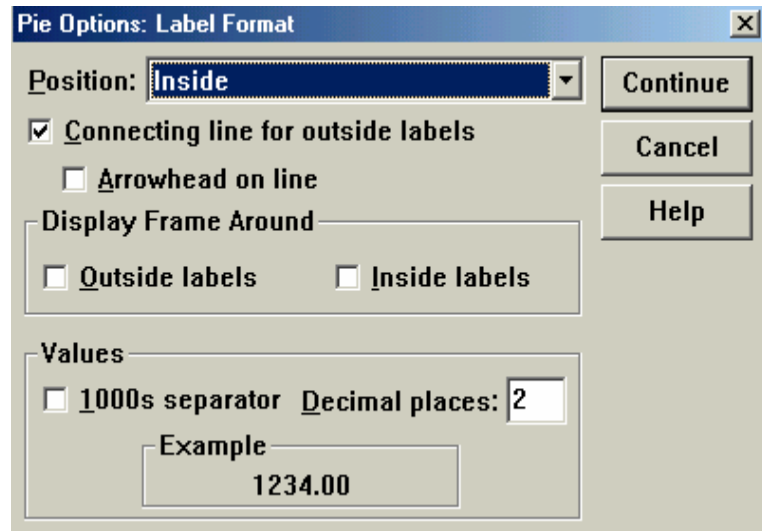


- لأعادة المقطع المفصول الى ما كان عليه أنقر المقطع بزر الماوس الأيسر ثم اختر Explode slice → Format أو أنقر الأيقونة  .

لعرض النسبة المئوية للمتغير في كل سنة في داخل كل مقطع Inside Slice نتبع الخطوات التالية :
☐ من القوائم أختَر Options → Chart (بعد التأكد أن المخطط في Chart Editor) فيظهر صندوق حوار Pie Options الذي نرتبه كما يلي :



لاحظ أننا أشرنا الخيار Percents في خانة Labels لعرض النسبة المئوية والسنة Text لكل مقطع .
 انقر زر Format فيظهر صندوق حوار Label Format الذي يرتب كما يلي :



حيث أن :

Position : يمثل موقع عرض النسبة المئوية وقيمة المتغير Value و Text (inside Outside Justified) ، Best fit، Outside، numbers inside Texts Outside) . وقد اخترنا Inside لأننا نرغب في عرض النسبة المئوية والسنة داخل المقطع .

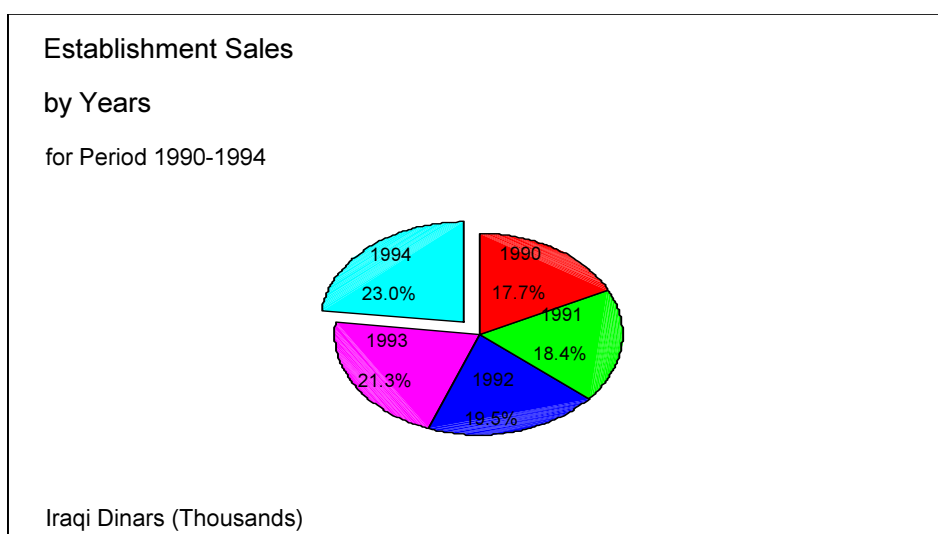
Connecting Line for Outside Labels : لعرض خطوط تربط بين المقاطع والعناوين الخارجية فقط .

Arrowhead on Line : استعمال خطوط سهمية الرأس .

Display Frame Around : لعرض إطار حول العناوين الداخلية أو الخارجية .

Values : لإضافة فواصل للأعداد التي تزيد عن 1000 وتحديد عدد المراتب العشرية للقيم فقط.

☐ عند نقر زر Continue ثم زر Ok نحصل على المخطط التالي :



مثال 2 :

الجدول التالي يمثل صادرات وأستيرادات بلد ما للسنوات 1998-2002 .

year	Exports	Impots
1998	190	180
1999	220	200
2000	219	215
2001	245	230
2002	250	244

يطلب ما يلي :

1. أعداد مخطط الأعمدة (الاشريطة) القطاعية Clustered Bars .

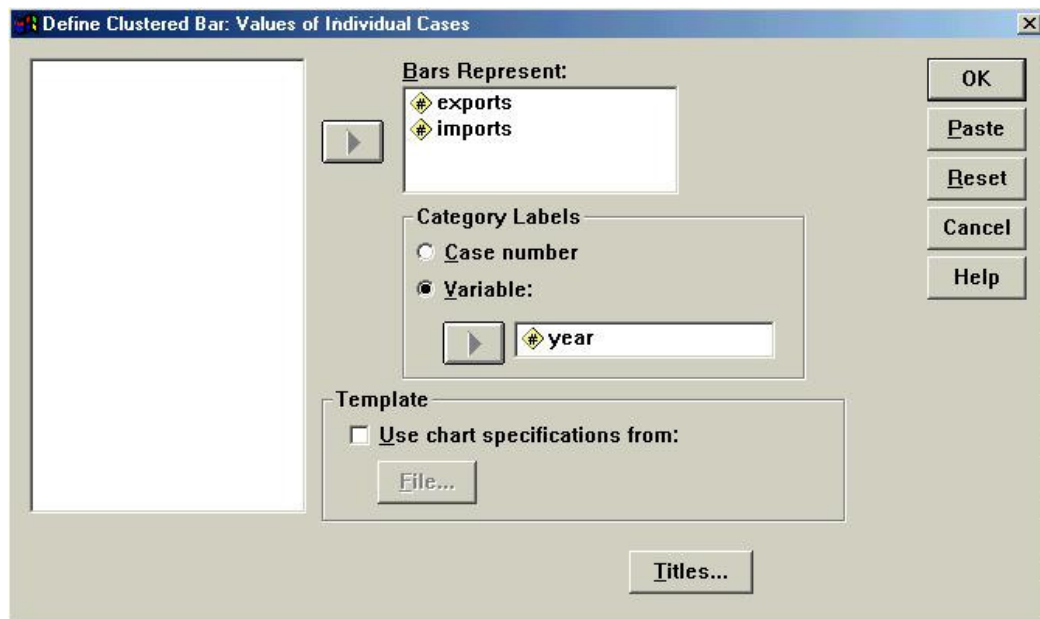
2. أعداد مخطط الأعمدة التراكمية Stacked Bars .

1. لأعداد الأعمدة القطاعية تتبع الخطوات التالية :

□ من شريط القوائم اختر Bar → Graph فيظهر صندوق Bar Chart ومنه اختر

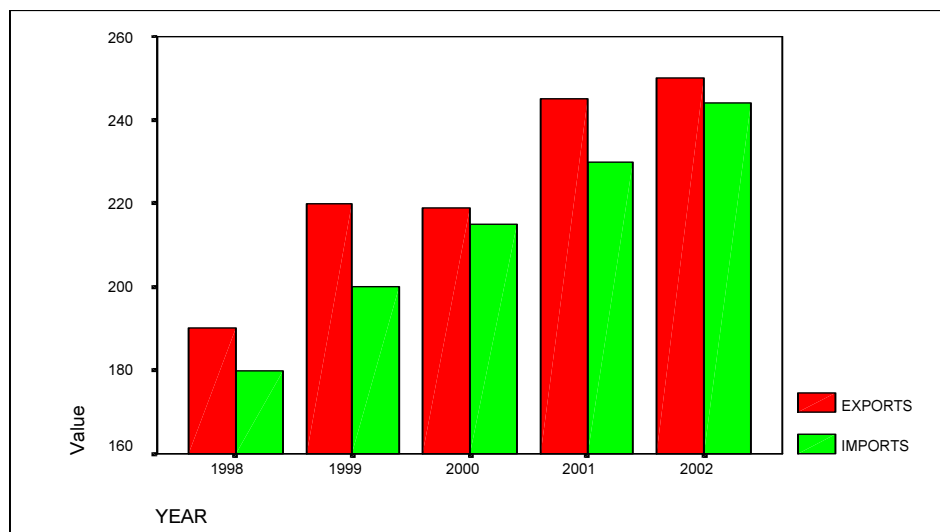
Values of Individual Cases / Clustered ثم انقر زر Define فيظهر صندوق حوار

Clustered Bar / Values of Individual Cases الذي نرتبه كما يلي :

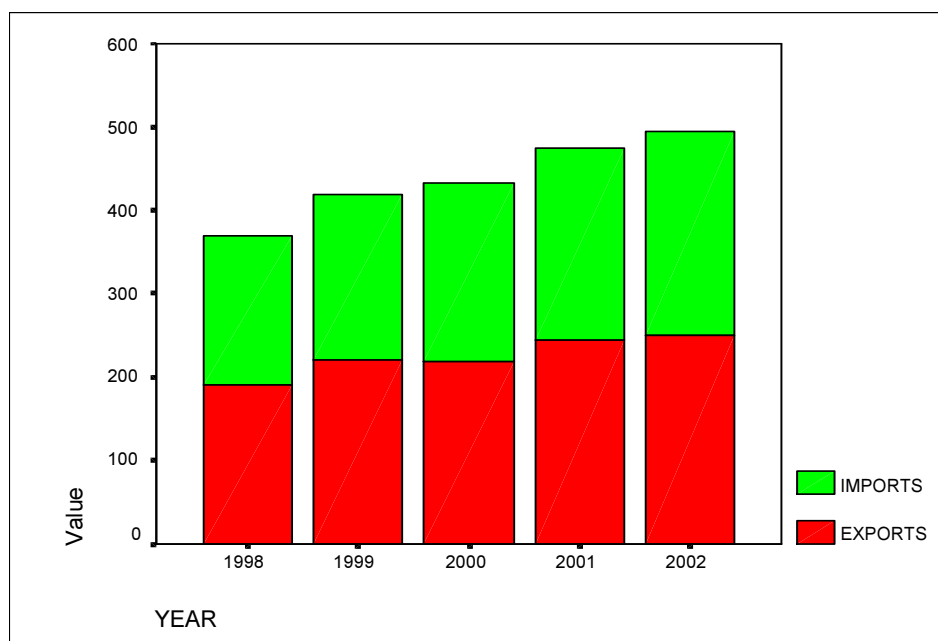


□ عند نقر زر Ok نحصل على المخطط التالي :

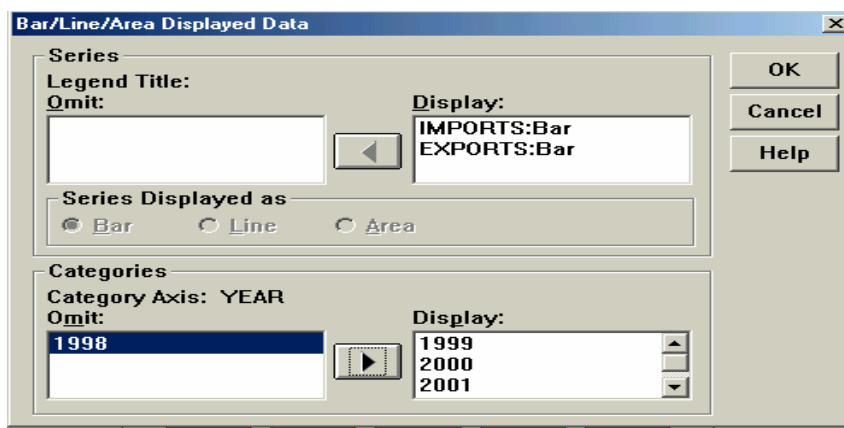
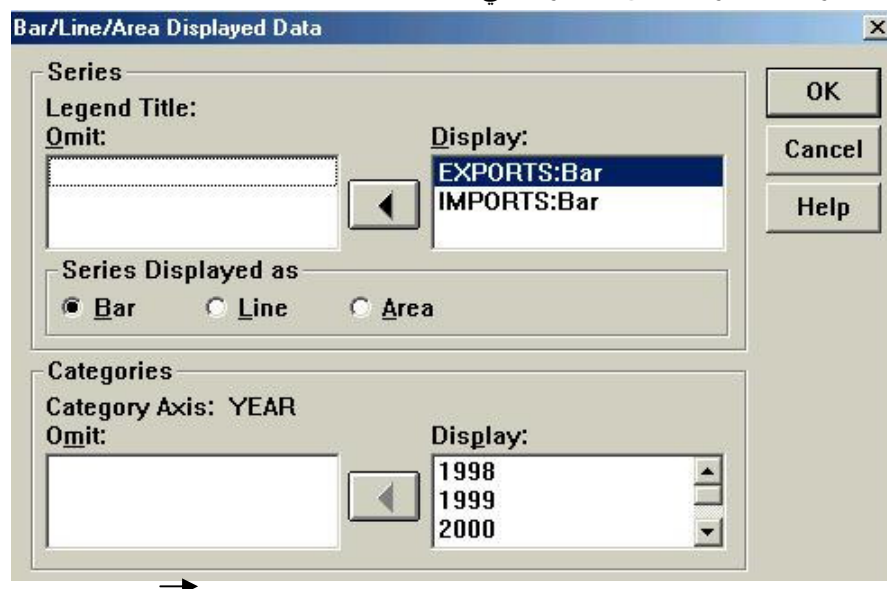
مخطط رقم 18



2. لتنفيذ مخطط Stacked Bar نتبع الخطوات السابقة نفسها عدا الخطوة رقم 2 حيث نختار Stacked بدل Clustered في صندوق حوار Bar Charts حيث نحصل على المخطط التالي :



بفرض أننا نريد جعل السلسلة Imports تعرض أولاً (في الأسفل) و السلسلة Exports تعرض ثانياً (في الأعلى) بالإضافة الى ذلك نرغب في إزالة فئة 1998 من المخطط لتنفيذ ذلك انقر المخطط مرتين ليكون في شاشة Chart Editor ثم اختر من شريط القوائم Series → Displayed أو انقر أي من السلسلتين مرتين فيظهر صندوق الحوار التالي :

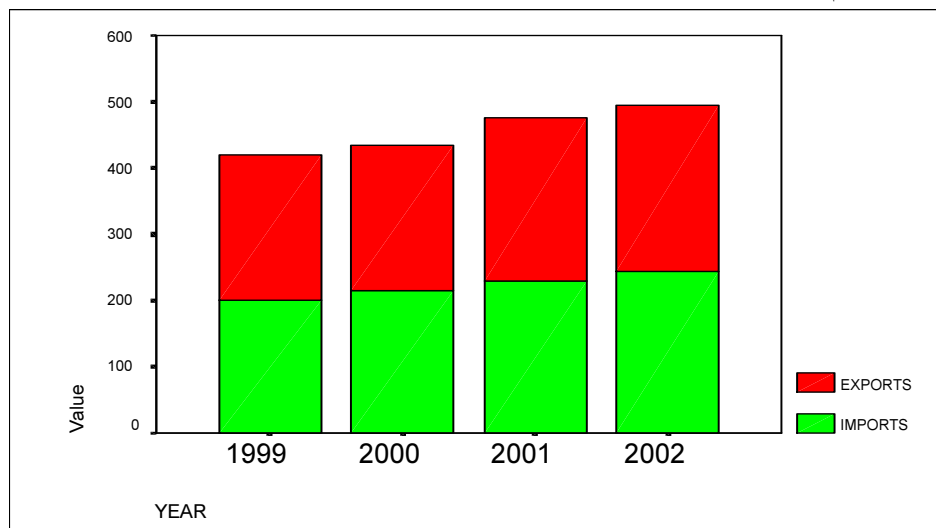


- أنقر السلسلة Exports في خانة Display (في الأعلى) لتحديد لها (إذا لم تحدد مسبقاً) ثم أنقر الزر  لنقلها الى خانة Omit (السلاسل المحذوفة من المخطط) في اليسار ثم أنقر الزر  لإرجاع السلسلة Exports من خانة Omit الى خانة Display (السلاسل المعروضة) وبهذا نكون قد غيرنا ترتيب السلاسل في خانة Display .
- أنقر العنوان 1998 في خانة Display (في الأسفل) لتحديده ثم أنقر الزر  لنقله الى خانة الحذف Omit .

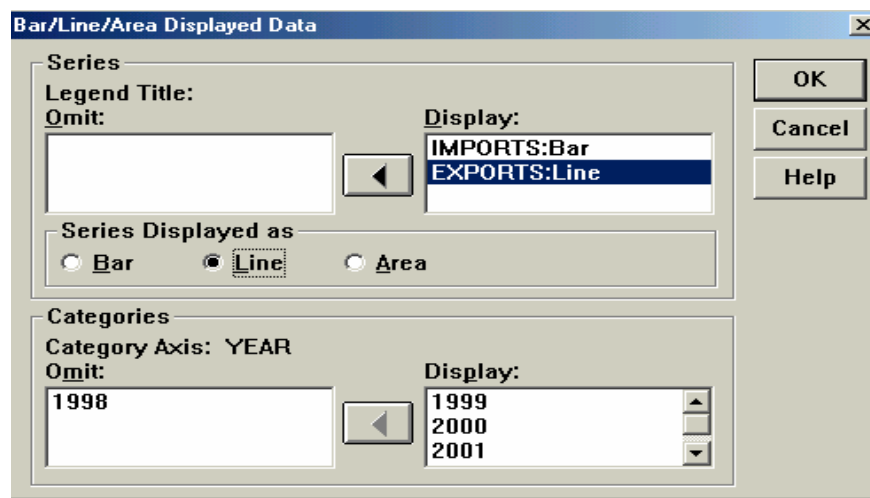
يظهر صندوق الحوار بعد التحوير كما يلي :

- عند نقر زر OK يظهر يعرض المخطط التالي :

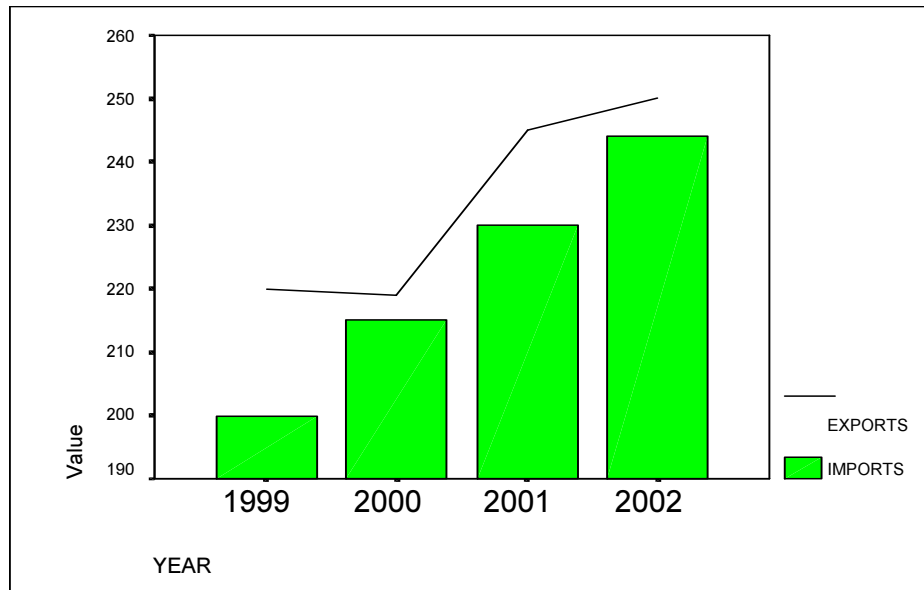
مخطط رقم 20



لعرض السلسلة Exports على شكل خط Line والسلسلة Imports على شكل أعمدة Bar نقوم بتحويل صندوق حوار Bar/Line/Area Displayed Data كما يلي :



لاحظ أننا أختارنا Bar للسلسلة Imports و Line للسلسلة Exports وعند نقر زر OK يعرض المخطط التالي :



ملاحظة :

لقد حصلنا على مخططي Line و Pie من مخطط Bar وبطبيعة الحال يمكن أن نحصل على هذه المخططات بصورة مستقلة كما حصلنا على مخطط Bar وكما يلي :

Line	Graph	Line	→
Pie	Graph	Pie	→
Area	Graph	Area	→

مثال 3 (على الخيار Summaries of Separate Variables)

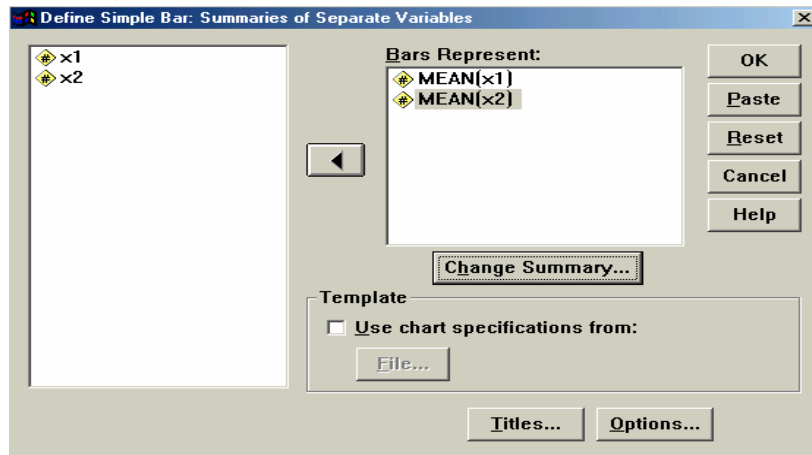
الجدول التالي يمثل المتغيرين x1 و x2 والمتغير g وقد أدخلت البيانات في شاشة Data Editor لبرنامج SPSS كما يلي :

x1	x2	g
100	10	a
200	20	b
300	30	a
400	40	b
500	50	b

يطلب مايلي :

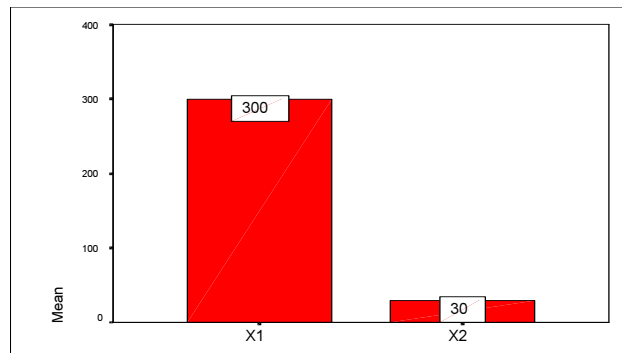
1. أعداد مخطط Bar لمتوسطي المتغيرين x1 و x2 .
 2. أعداد مخطط قطاعي Clustered Bar لمتوسطات فئتي (a و b) للمتغيرين x1 و x2 .
1. لتنفيذ مخطط Bar لمتوسطي المتغيرين x1 و x2 نتبع الخطوات التالية:
- ☐ من شريط القوائم نختار Bar → Graph فيظهر صندوق حوار Bar Charts ومنه نختار Summaries of Separate Variables / Simple .
 - ☐ عند نقر زر Define في هذا الصندوق الأخير يظهر صندوق حوار

Define Simple Bar : Summaries of Separate Variables الذي نرتبه كما يلي :



نلاحظ أن الأعمدة ستمثل متوسط كل من المتغيرين x_1 و x_2 حيث أن المتوسط هو الاختيار الافتراضي لنوع الدالة ويمكن تغيير المتوسط بنقر الزر Change Summary الذي يتيح اختيار مؤشرات أخرى مثل ... Median, Mode, No. of Cases, Min. Value, Max. Value .

□ عند نقر زر OK نحصل على المخرج التالي :



مخطط رقم 22

أن ارتفاع العمود x_1 هو الوسط الحسابي للمتغير ويساوي 300 وأن ارتفاع العمود x_2 هو الوسط الحسابي للمتغير ويساوي 30 .

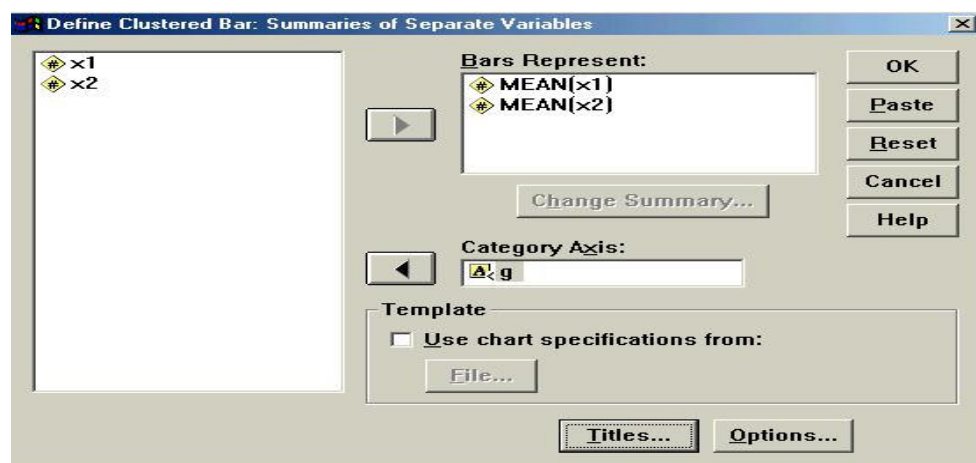
2. لتنفيذ المخطط القطاعي Clustered Bar لمتوسطات فئتي (a و b) للمتغيرين x_1 و x_2 نتبع

الخطوات التالية :

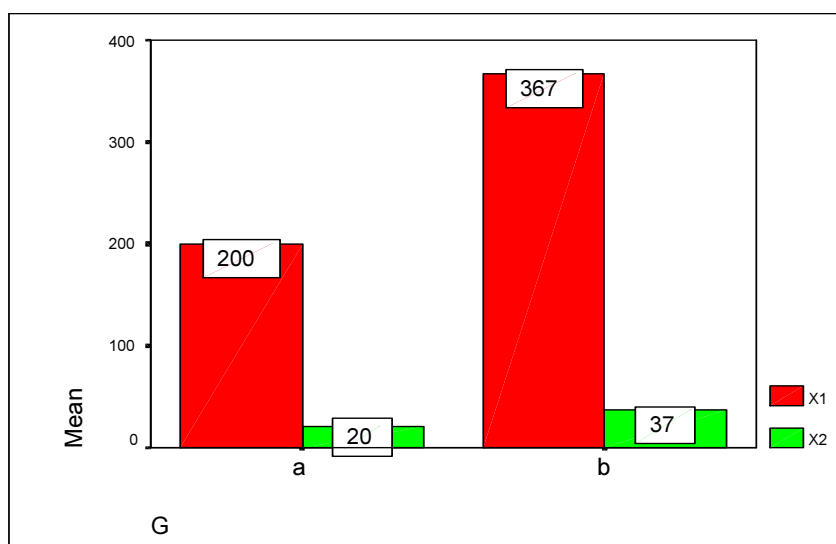
من شريط القوائم نختار Bar → Graph فيظهر صندوق حوار Bar Charts ومنه نختار Summaries of Separate Variables / Clustered .

□ عند نقر زر Define في هذا الصندوق الأخير يظهر صندوق حوار

Define Clustered Bar : Summaries of Separate Variables الذي نرتبه كما يلي :



□ عند نقر زر Ok نحصل على المخطط التالي :
مخطط رقم 23



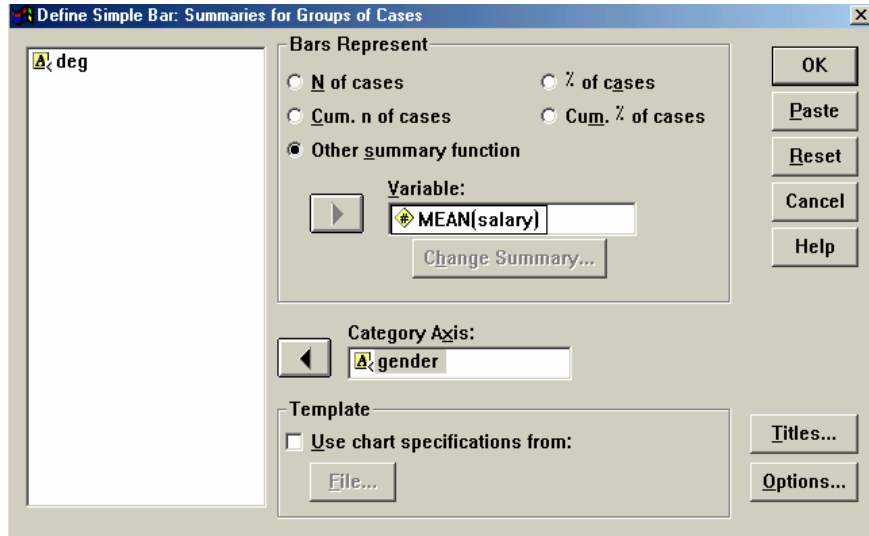
مثال 4 : (على الخيار Summaries of Groups of Cases)

الجدول التالي يمثل رواتب salary مجموعة من الموظفين حسب الدرجة الوظيفية deg والجنس gender .

deg	gender	salary
First	Male	90
First	Female	70
Third	Male	56
Second	Male	65
First	Male	85
Second	Female	60
Second	Male	69
Third	Male	57
Third	Female	50
First	Female	75
Second	Female	62
Third	Female	51
First	Male	85

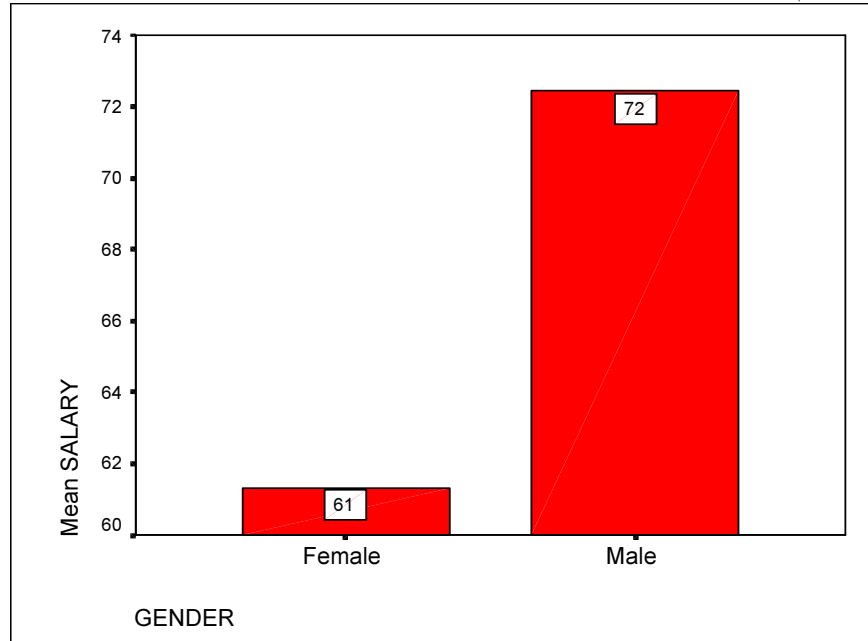
المطلوب مايلي :

1. أعداد مخطط بالأعمدة البيانية Bars يمثل متوسط الراتب لفئتي الذكور والإناث .
 2. أعداد مخطط بالأعمدة البيانية Bars يمثل متوسط الراتب لكل الذكور والإناث حسب كل درجة وظيفية .
1. لتنفيذ المطلوب الأول نتبع الخطوات التالية:
- ☐ من شريط القوائم نختار Bar → Graphs فيظهر صندوق حوار Bar Charts ومنه نختار Summaries for Group of Cases/ Simple
- ☐ عند نقر زر Define في صندوق الحوار أعلاه يظهر صندوق حوار Define Simple Bar : Summaries for Group of Cases الذي يرتب كما يلي :



☐ عند نقر زر OK نحصل على المخطط التالي :

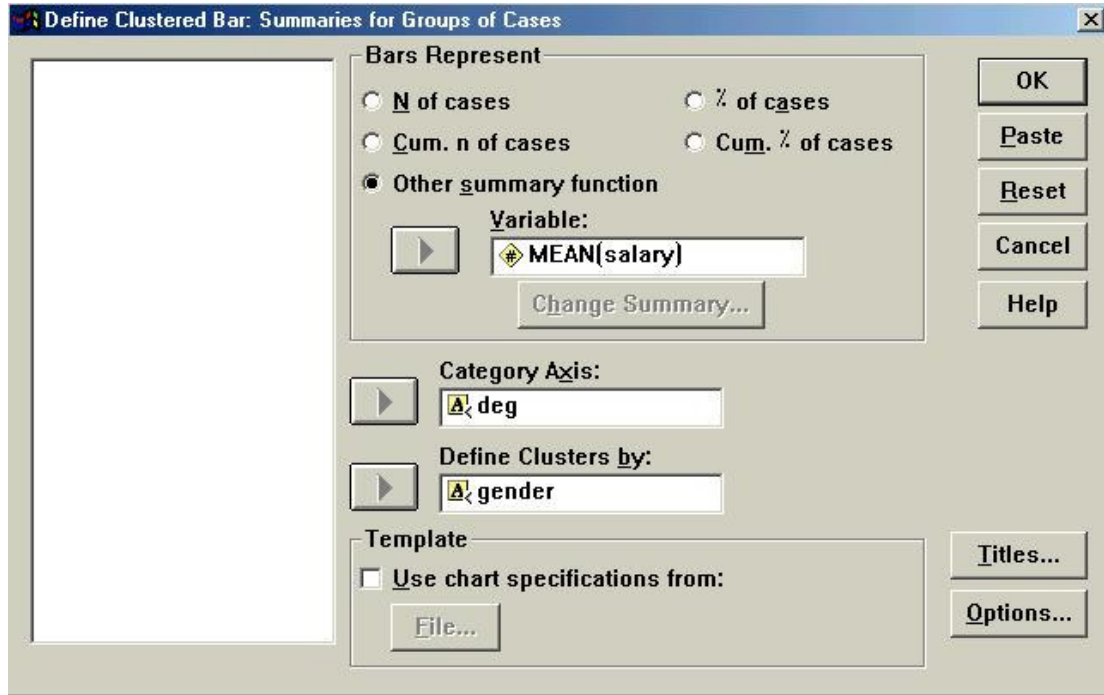
مخطط رقم 24



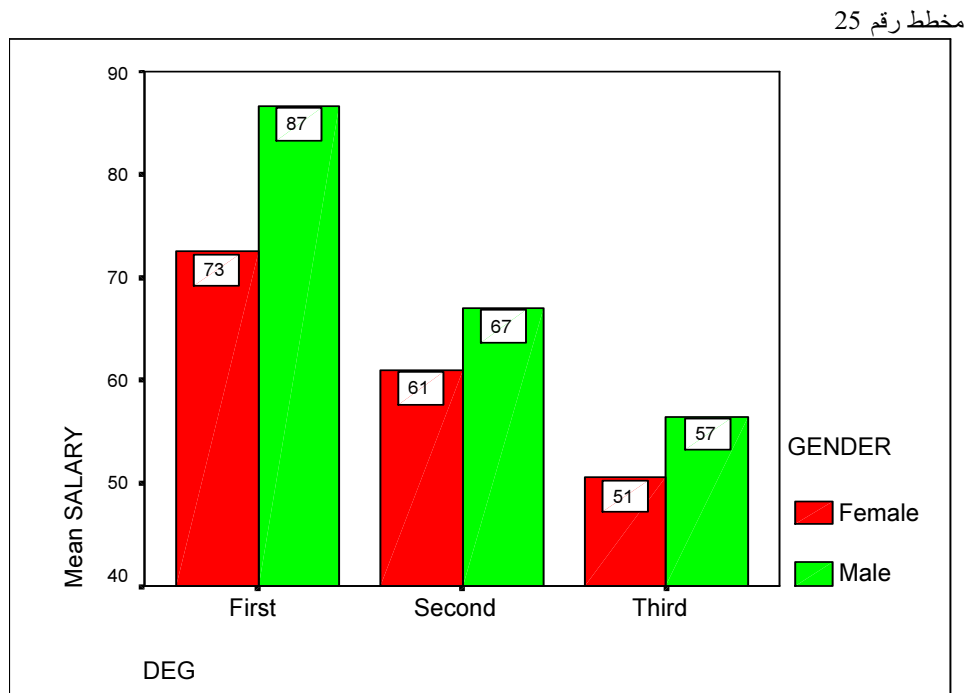
2. لتنفيذ المطلوب الثاني نتبع الخطوات التالية :

من شريط القوائم نختار **Bar Charts** → **Graphs** فيظهر صندوق حوار **Bar Charts** ومنه نختار **Summaries for Group of Cases/Clustered** حيث أننا سنمثل الدرجة الوظيفية على المحور الأفقي يقابل كل درجة وظيفية عمودين أحدهما لمتوسط رواتب الذكور والآخر لمتوسط رواتب الإناث **Clusters** .
عند نقر زر **Define** في صندوق الحوار أعلاه يظهر صندوق حوار

Define Clustered Bar: Summaries for Group of Cases الذي يرتب كما يلي :



عند نقر زر **OK** نحصل على المخطط التالي :



من خلال المثالين الأخيرين نلاحظ الميزة المهمة التي تتمتع بها مخططات برنامج SPSS والتي لا تتوفر في مخططات تطبيقات أخرى مثل EXCEL وهي المعالجة الإحصائية للبيانات حيث يمكن استخراج المؤشرات الإحصائية كالمتوسطات الحسابية مثلاً لفئات متغير معين وتمثيلها بيانياً وهذا يعد اختصاراً للجهود والوقت اللازمين لترتيب البيانات وخاصة إذا كان حجم البيانات كبيراً كما في استمارات الاستبيانات الإحصائية .

(13 - 5) المدرج التكراري Histogram

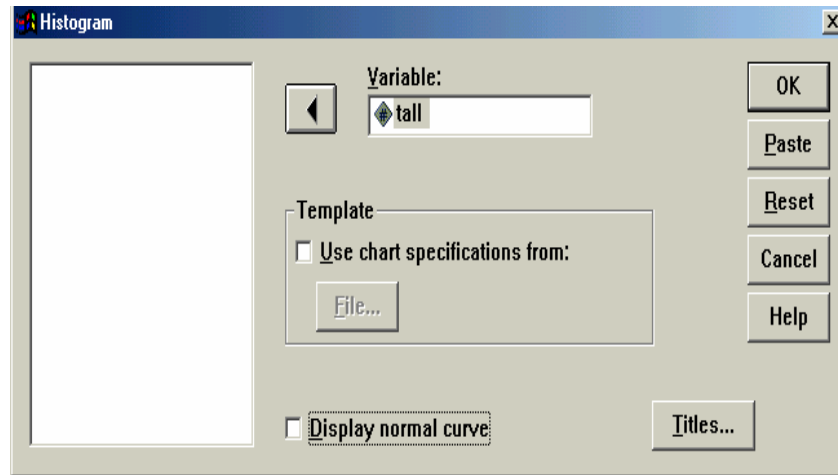
أن المخططات السابقة Bar و Line و pie تستعمل لعرض البيانات الخام Row Data أما المدرج التكراري فيستعمل لعرض البيانات المبوبة Grouped Data على الرغم من أننا نقوم بإدخال البيانات بصيغتها الخام Ungrouped الى برنامج SPSS حيث يقوم البرنامج بتصنيف البيانات الى جدول توزيع تكراري Frequency Table تلقائياً ومن ثم رسم المدرج التكراري لهذا الجدول حيث يمكن تغيير إعدادات المخطط (عدد الفئات ، طول الفئة ، ...) كما في المثال التالي .

مثال 5

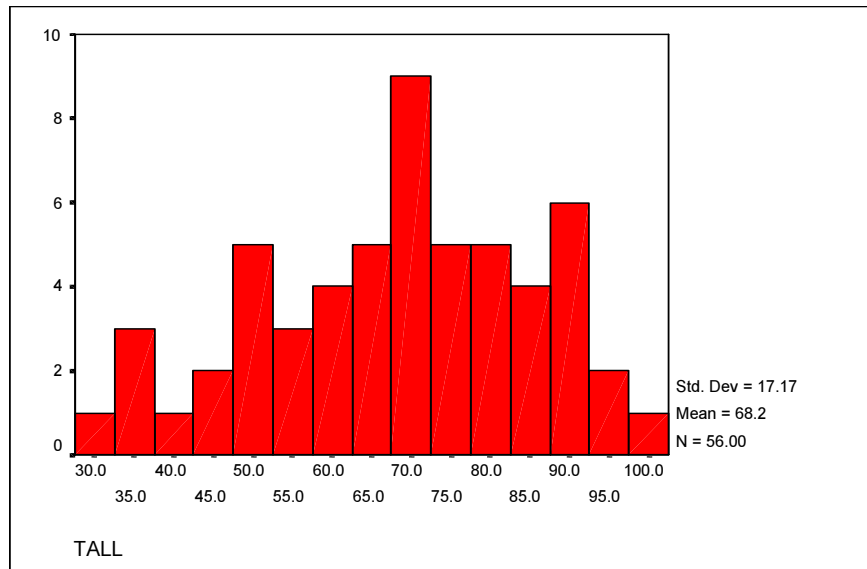
يطلب رسم المدرج التكراري للمتغير tall والذي أدخل مسبقاً في شاشة Data Editor (نفس المتغير الوارد في المثال التابع للبند (4-1) من الفصل الرابع) .

لرسم المدرج التكراري نتبع الخطوات التالية :

□ من شريط القوائم اختر Histogram → Graph فيظهر صندوق حوار Histogram الذي نرتبه كما يلي :

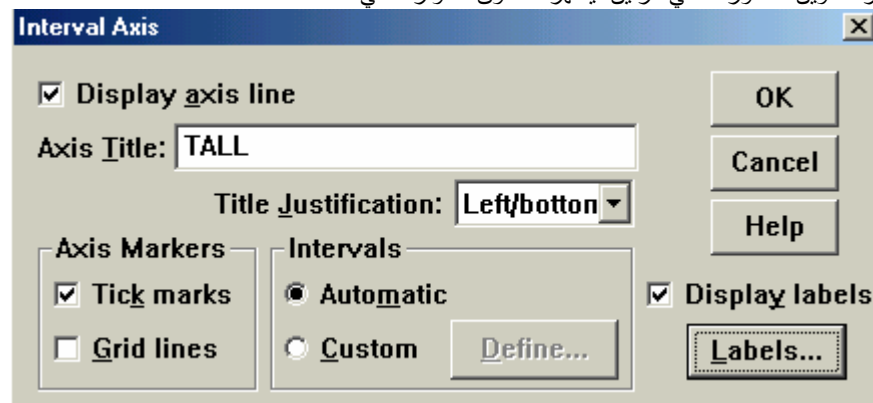


□ عند نقر زر OK يتم عرض المدرج التكراري التالي :
مخطط رقم 26



حيث نلاحظ أنه تم تمثيل مراكز الفئات Midpoints على المحور الأفقي نرغب بتمثيل الفئات Classes على المحور الأفقي عوضاً عن مراكز الفئات كما نود أيضاً تحديد طول الفئة Interval Width بـ 10. لتنفيذ ذلك نتبع الخطوات التالية:

- انقر مخطط Histogram مرتين لينتقل الى شاشة Chart Editor .
- انقر عناوين المحور الأفقي مرتين فيظهر صندوق الحوار التالي :



أقر زر Custom في خانة Intervals فيتحفز الزر Define وعند نقره يظهر صندوق الحوار Define Custom Intervals الذي نرتبه كما يلي :

لقد حددنا طول الفئة بـ 10 في Interval width وجعلنا اقل قيمة معروضة 30 وأكبر قيمة معروضة 100 وعليه فأن عدد الفئات يساوي 7 الذي يحتسب من العلاقة التالية :

$$No. of Classes = \frac{Range}{Class (Interval) Width} = \frac{70}{10} = 7$$

• أنقر زر continue للرجوع الى صندوق حوار Interval axis ثم أنقر زر Labels في هذا الصندوق فيظهر صندوق حوار Labels الذي نرتبه كما يلي :

حيث قمنا بالعمليات التالية :

1. اخترنا Range بدلاً من Midpoint في خانة Type لأننا نرغب في عرض الفئات على المحور الأفقي بدلاً من مراكز الفئات .
2. حددنا عدد المراتب العشرية بصفر للحدود العليا والدنيا للفئات في خانة Decimal Places .
3. في خانة Orientation اخترنا Diagonal كي لا تتشابك حدود الفئات فيما بينها .

أقر زر Continue للرجوع الى صندوق حوار Interval axis ثم أنقر زر OK فيعرض المخطط التالي :

مخطط رقم 27

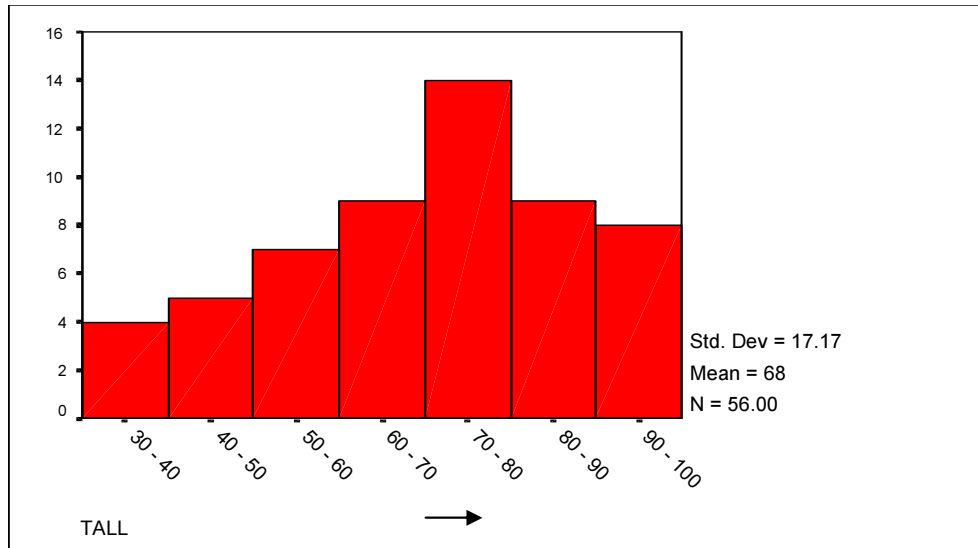


Chart Options عندما يكون المخطط في شاشة Chart Editor ثم انقر Normal curve في خانة Display أو يمكنك تأشير Display Normal Curve في صندوق حوار Histogram عند ترتيب هذا الصندوق في بداية العمل .

Box Plot (6 - 13) مخطط

يستعمل هذا المخطط لوصف توزيع مجموعة من المشاهدات حول الوسيط Median (راجع البند 6 - 1 لمزيد من التفاصيل حول مكونات المخطط) .

مثال 6 : الجدول التالي يتضمن مجموعة من المتغيرات التي أدخلت الى شاشة Data Editor في برنامج SPSS وكما يلي :

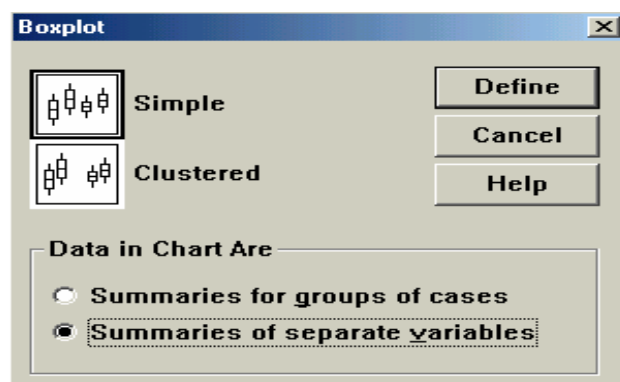
x	y	cat	z	g	gender
30	47	a	20	a	f
20	66	a	60	b	f
31	77	a	50	a	m
35	90	a	80	b	m
100	55	a	100	b	m
80	62	a	70	a	m
79	80	b	89	b	m
55	99	b	35	a	f
50	43	b	65	b	f
60	87	b	40	a	f
95	92	b	55	b	f
45	70	b	69	b	m
39	40	b	40	a	m

1. يطلب رسم مخطط Box Plots للمتغير x .
2. رسم مخطط Box plot للمتغيرين x و y حسب الفئتين a و b التي يحددها المتغير cat .
3. رسم مخطط Box Plot للمتغيرين x و y حسب الفئتين a و b التي يحددها المتغير cat .
4. رسم مخطط Box Plot لفئتي المتغير z (b و a) التي يحددها المتغير g .

5. رسم مخطط Box Plot لفئتي المتغير z (a و b) التي يحددها المتغير g علماً أن كل من الفئتين تتضمن بدورها فئتين إحداهما للذكور m والأخرى للإناث f .

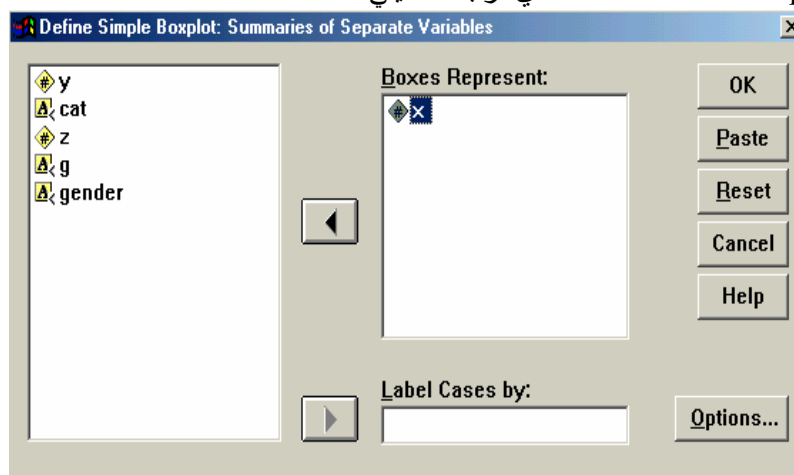
1. لتنفيذ المطلوب الأول نتبع مايلي :

من شريط القوائم اختر Box plot → Graphs فيظهر صندوق حوار Boxplot التالي :

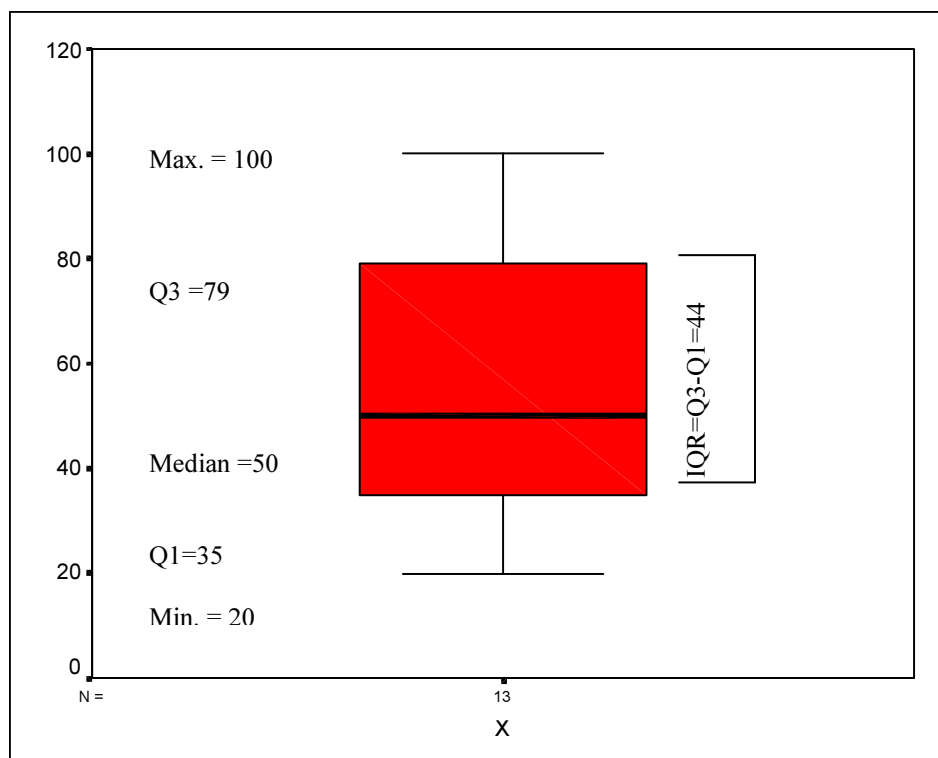


لاحظ أننا اخترنا summaries of Separate Variables مع الخيار Simple (حيث لايمكن اختيار Clustered لعدم وجود فئات للمتغير x) .

عند نقر زر Define في صندوق الحوار أعلاه يظهر صندوق حوار Define Simple boxplot: Summaries of Separate Variables الذي نرتبه كما يلي :

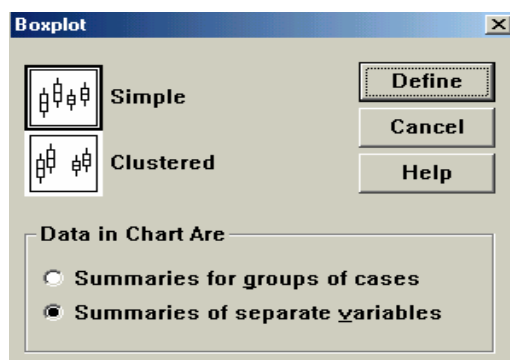


لقد تم إدخال أسم المتغير في Boxes Represent كما يتم إدخال متغير في خانة Label Cases by عند الرغبة في تحديد عنوان Label للقيم المتطرفة والشاذة (تظهر العناوين في المخطط) أما في حالة عدم إدخال متغير فيسمعمل رقم الحالة التي تقع ضمنها القيمة المتطرفة أو الشاذة كعنوان لها .
عند نقر زر OK نحصل على المخطط التالي :



2. لتنفيذ المطلوب الثاني نتبع مايلي :

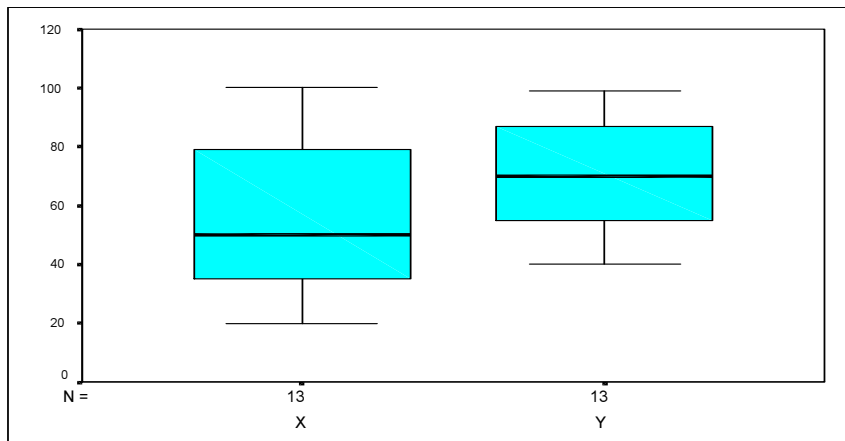
من شريط القوائم اختر Box plot → Graphs فيظهر صندوق حوار Boxplot الذي نرتبه كما يلي :



عند نقر زر Define يظهر صندوق حوار Define Simple Box plot : Summaries of

Separate Variables الذي نرتبه كالتالي :

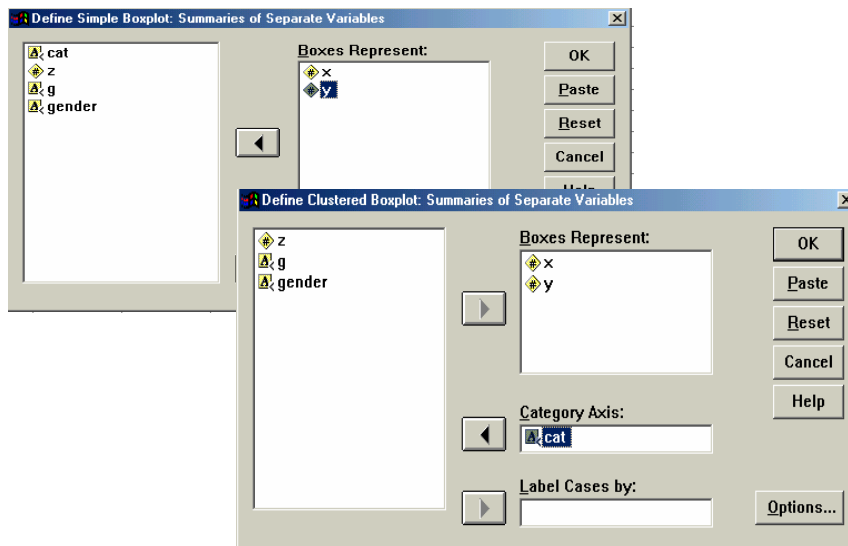
عند نقر زر OK يعرض المخطط التالي :



3. تنفيذ المطلوب الثالث نتبع الخطوات التالية :

Graphs

من القوائم نختار

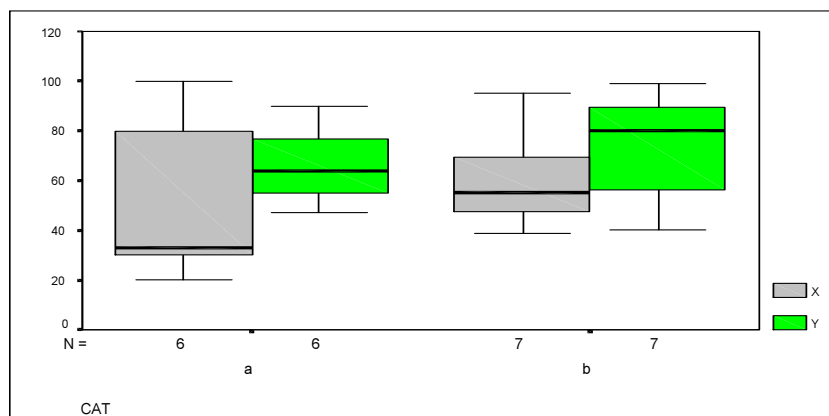


Boxplot يظهر صندوق حوار Box Plot ومنه نختار Summaries of Separate Variables مع اختيار Clustered وعند نقر زر Defined في صندوق الحوار الأخير يظهر صندوق حوار Define Clustered Box Plot : Summaries of Separate Variables الذي نرتبه كما يلي :

لاحظ أن خانة Box Represent يجب أن تتضمن متغيرين في الأقل وأن Category Axis هو المتغير الذي يتم تمثيل فئاته على المحور الأفقي .

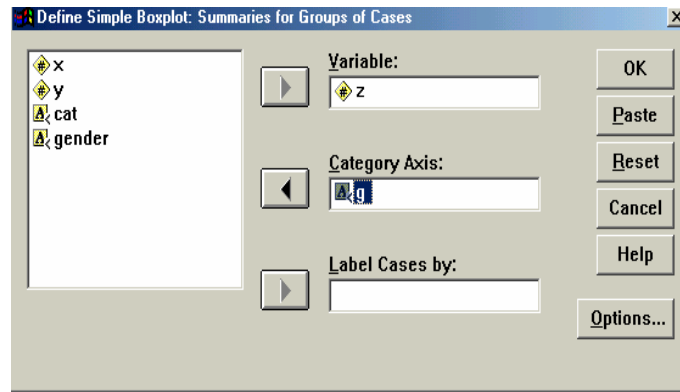
عند نقر زر OK يعرض المخطط التالي :

مخطط رقم 30



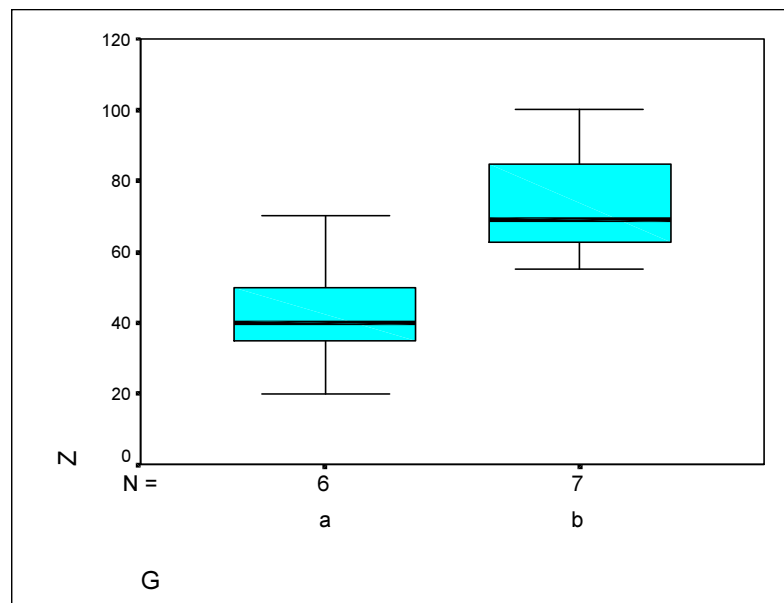
4. لتنفيذ المطلوب الرابع نتبع الخطوات التالية :

من القوائم نختار Boxplot → Graphs فيظهر صندوق حوار Box Plot ومنه نختار Summaries for Group of Cases مع اختيار Simple وعند نقر زر Defined في صندوق الحوار الأخير يظهر صندوق حوار Define Simple Box Plot : Summaries for Group of Cases الذي نرتبه كما يلي :



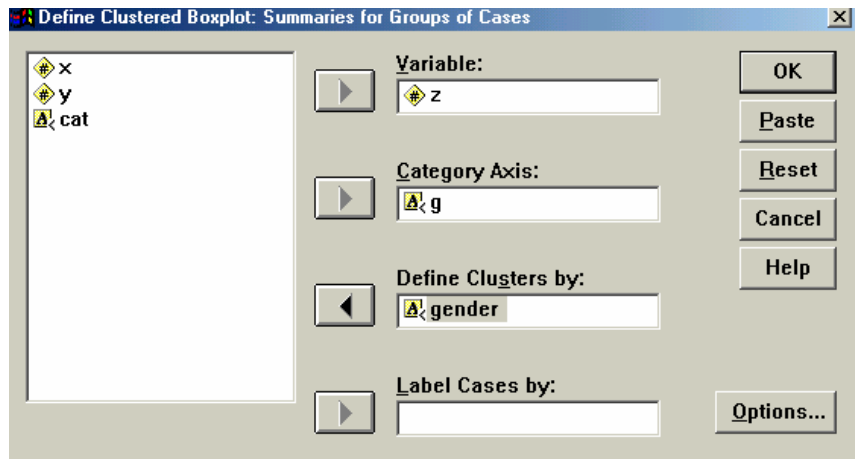
عند نقر زر OK يعرض المخطط التالي :

مخطط رقم 31



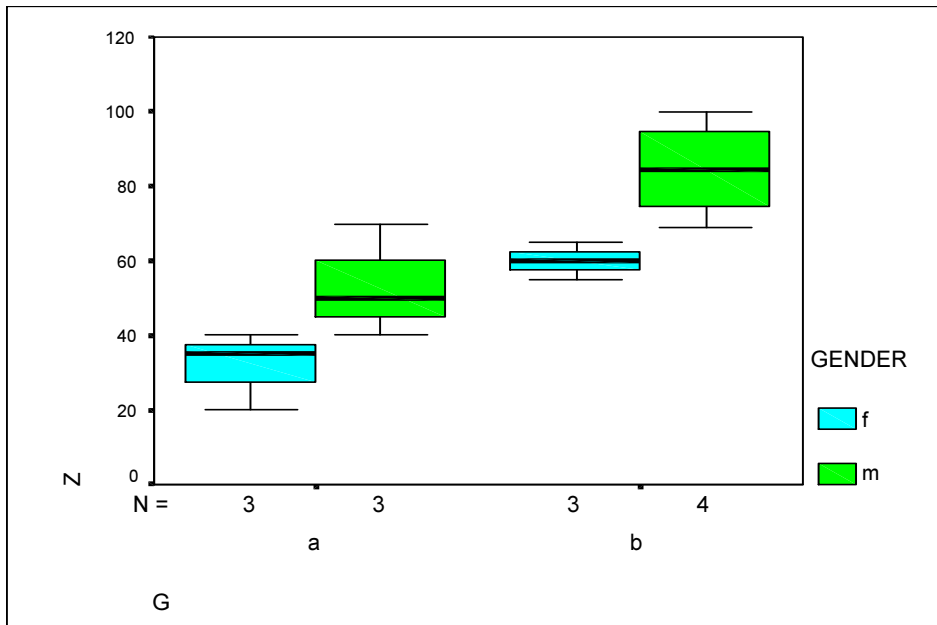
5. لتنفيذ المطلوب الخامس نتبع الخطوات التالية :

من القوائم نختار Boxplot → Graphs فيظهر صندوق حوار Box Plot ومنه نختار Summaries for Group of Cases مع اختيار Clustered وعند نقر زر Defined في صندوق الحوار الأخير يظهر صندوق حوار Define Clustered Box Plot : Summaries for Group of Cases الذي يتم ترتيبه كالتالي :



عند نقر زر OK يعرض المخطط التالي :

مخطط رقم 32



(7-13) مخطط شكل الانتشار Scatterplot

يستعمل هذا النوع من المخططات لتمثيل العلاقة بين متغيرين أو أكثر بمجموعة من النقاط المتناثرة ويتضمن أربعة أنواع لشكل الانتشار .

مثال 7 :

الجدول التالي يتضمن عدد من المتغيرات والتي أدخلت في ورقة Data Editor لبرنامج SPSS وكما يلي :

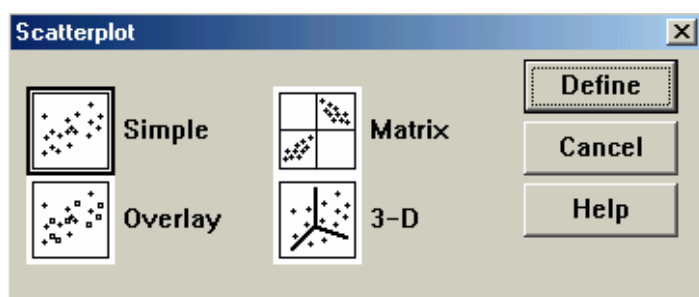
y	x1	x2	x3	marker	label
79	26	6	60	a	c1
74	29	15	52	a	c2
104	56	8	20	a	c3
87	31	8	47	a	c4
95	52	6	33	a	c5
109	55	9	22	a	c6
102	71	17	6	b	c7
72	31	22	44	b	c8
93	54	18	22	b	c9
115	47	4	26	b	c10
83	40	23	34	b	c11
113	66	9	12	b	c12
109	68	8	12	b	c13

سنحاول فيما يلي توضيح كيفية أعداد الأنواع الأربعة لشكل الانتشار :

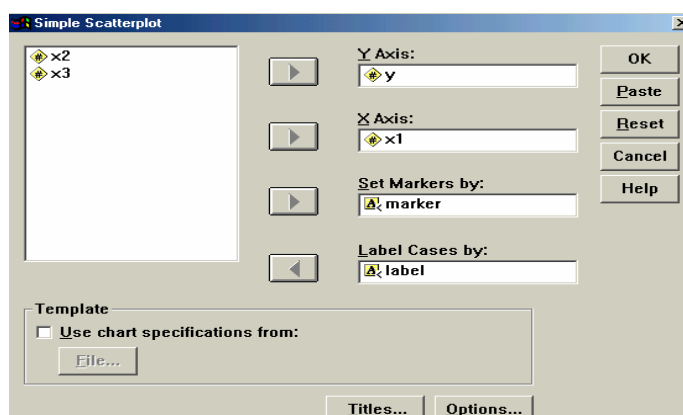
1. شكل الانتشار البسيط Simple

يستعمل لرسم العلاقة بين متغيرين فقط أحدهما يمثل المحور الأفقي للشكل والآخر المحور العمودي له وأن إحداثيات كل نقطة في الشكل تعتمد على هذين المتغيرين ، لأعداد شكل الانتشار البسيط للعلاقة بين y و x1 نتبع الخطوات التالية:

◀ من شريط القوائم اختر Scatter → Graphs فيظهر صندوق حوار Scatterplot التالي :



◀ في صندوق الحوار أعلاه أختَر Simple بنقره ثم أنقر Define فيظهر صندوق حوار Simple scatterplot الذي نرتبه كما يلي :



حيث أن :

Y Axis : تتضمن هذه الخانة المتغير الذي يمثل المحور الرأسي لشكل الانتشار .

X Axis : تتضمن هذه الخانة المتغير الذي يمثل المحور الأفقي لشكل الانتشار .

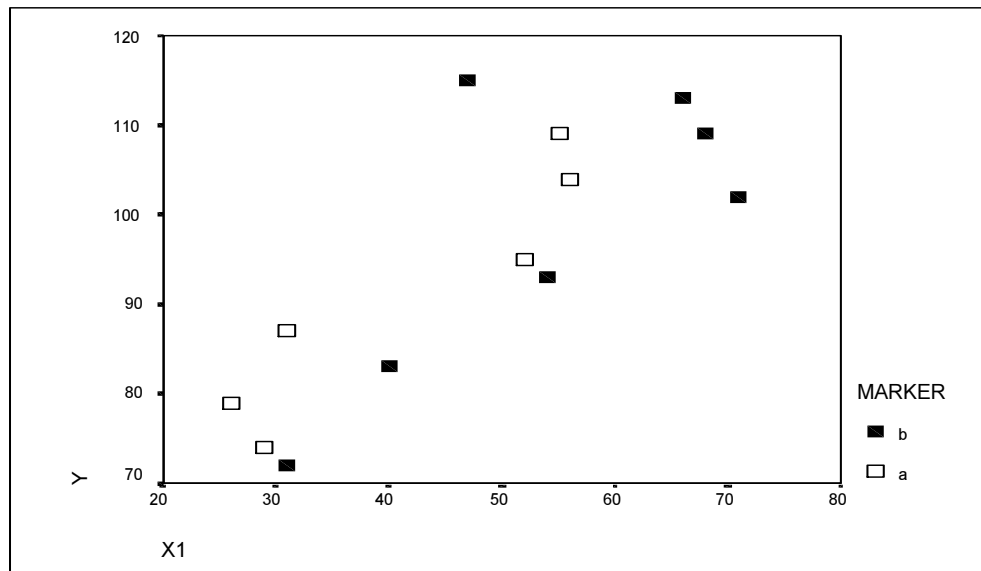
Set Marker by : تتضمن هذه الخانة متغير له فئتين أو أكثر وكل فئة تمثل بعلامة Marker خاصة بها في شكل الانتشار .

Label Cases by : تتضمن هذه الخانة متغير تستعمل قيمه كعناوين Labels لنقاط شكل الانتشار حيث أنه بالإمكان إظهار العناوين أو إخفائها .

علماً أن الخانتين Set Markers by و Label Cases by اختياريّتين Optional أي بالإمكان عدم إدخال متغير في أي منهما .

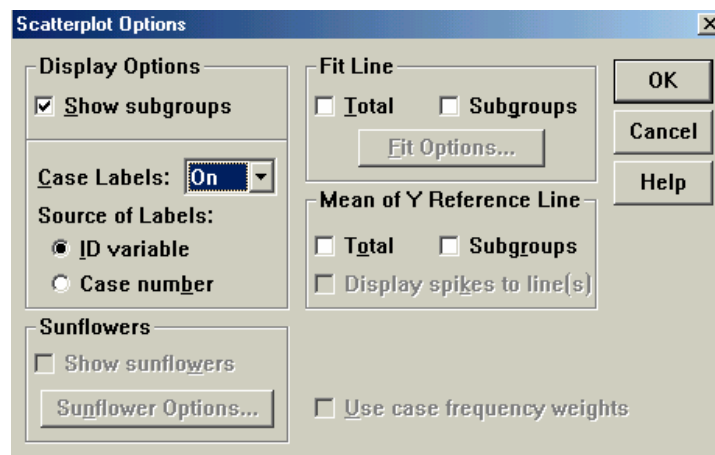
➤ عند نقر زر OK يعرض المخطط التالي :

مخطط رقم 33



لاحظ أن عناوين الحالات لم تظهر في المخطط على الرغم من أننا استعملنا قيم المتغير Label كعناوين للحالات ، ولإظهار العناوين نتبع مايلي :

- انقر المخطط مرتين بزر الماوس الأيسر ليتحول المخطط الى شاشة Chart Editor .
- في شاشة Chart Editor اختر من شريط القوائم Options → Chart فيظهر صندوق حوار Options الذي نرتبه كما يلي :



حيث أن خانة Case Labels تتظم عرض عناوين الحالات Case Labels وتحتوي ثلاثة خيارات :

off : لإزالة عناوين الحالات .

on : لعرض عناوين الحالات كافة .

As is : لعرض عناوين جزء من الحالات أنقر أيقونة تعريف النقطة  في شاشة Chart Editor فيصبح شكل الماوس مشابهاً للأيقونة وعند نقر أي نقطة في المخطط يعرض العنوان الخاص بها في جوار النقطة (لإنهاء عمل الأيقونة تنقر مرة ثانية) .

وبما أننا نرغب في عرض كافة عناوين الحالات فقد اخترنا on .

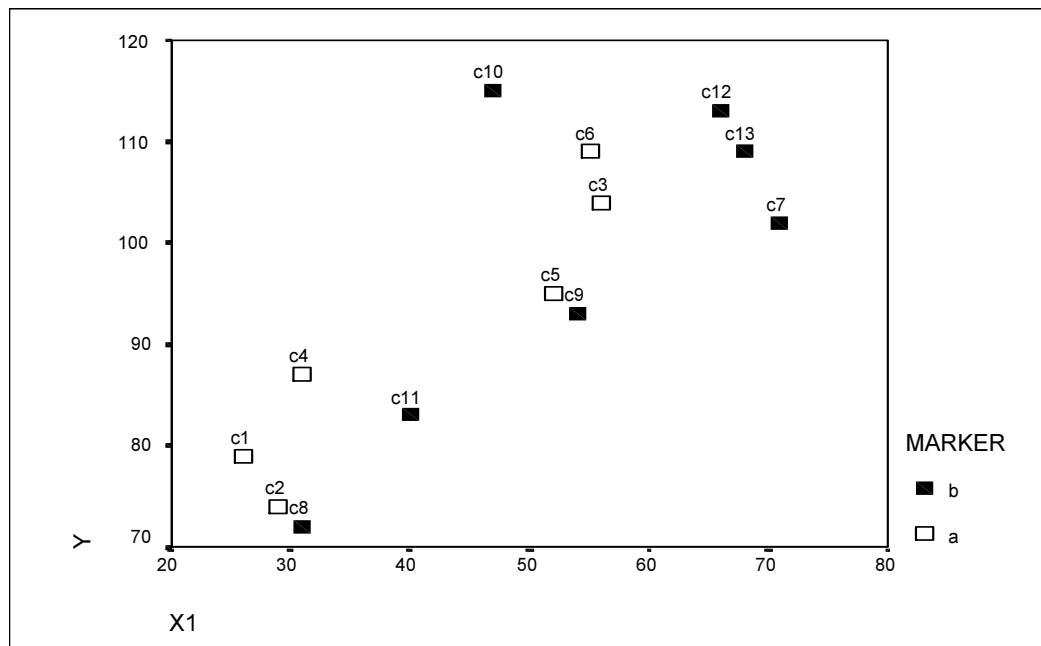
أما خانة Source of Labels فتبين مصدر عناوين الحالات حسب الخيارين التاليين :

ID Variable : يستعمل في حالة وجود متغير يمثل العناوين كما هو الحال في هذا المثال (متغير العناوين هو Label وقد أدخلناه في صندوق حوار Simple Scatterplot .

Case Number : يستعمل رقم الحالة كعنوان في حالة عدم وجود متغير لعناوين الحالات .

• عند نقر زر OK في صندوق الحوار أعلاه تعرض عناوين الحالات كما يلي :

مخطط رقم 34

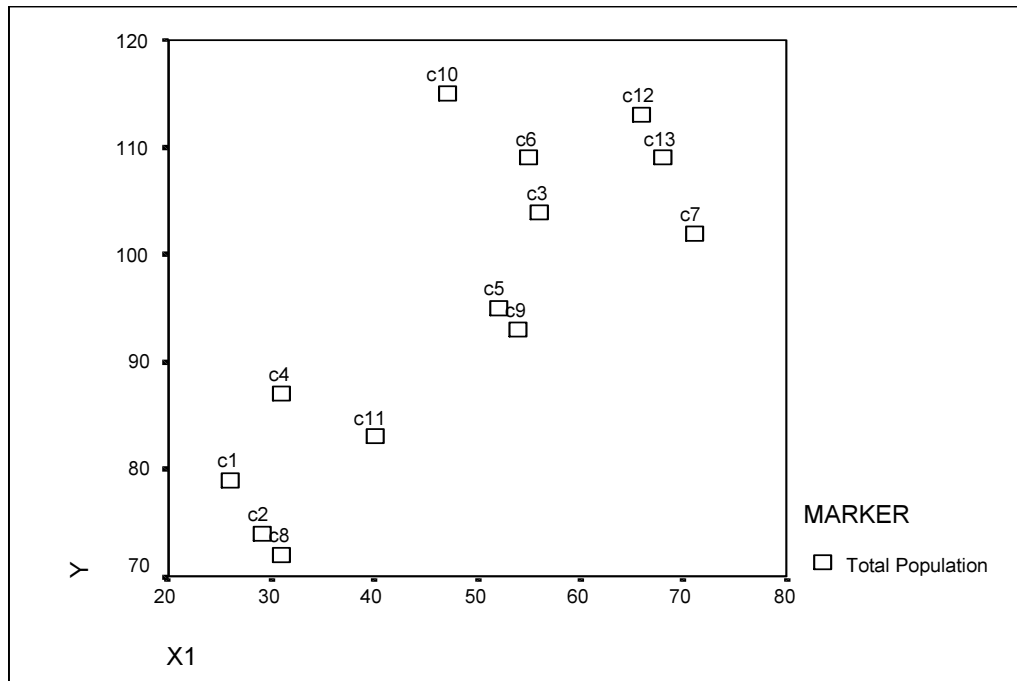


ملاحظة 1 : يمكن عرض عناوين الحالات مباشرة بنقر الزر Options في صندوق حوار Simple Scatterplot ثم تأشير الخيار Display Chart With Case Labels في صندوق حوار Options .

ملاحظة 2 : في صندوق حوار Scatterplot : Options إذا كان الخيار Show Subgroups (في خانة Display Options) مؤشراً يتم عرض المجاميع الفرعية (في هذا المثال المجموعتين a

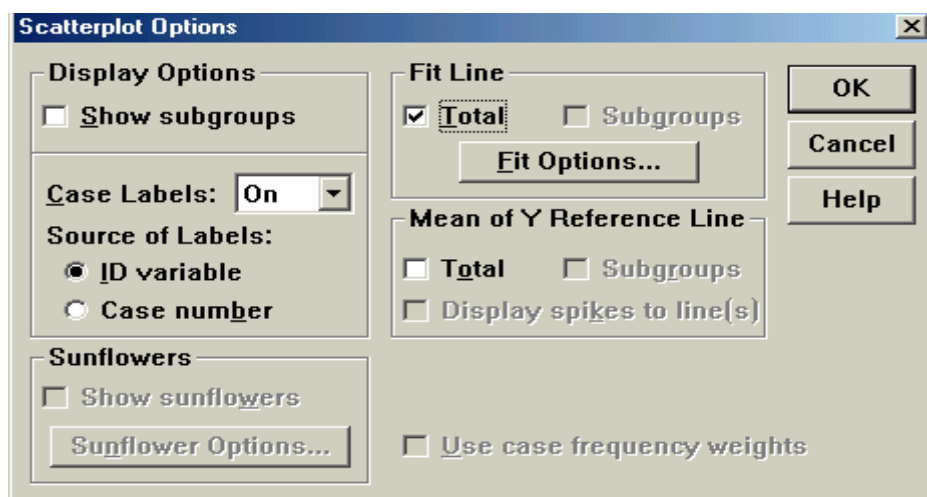
و b) بعلامات مختلفة لكل منهما كما في المخطط أعلاه (المربع الأسود لـ b والمربع الأبيض لـ a) وعند تأشير هذا الخيار يتم عرض المجموعتين بعلامة واحدة كما في المخطط التالي .

مخطط رقم 35



ملاحظة 3 : لإضافة خط الانحدار من الدرجة الأولى $\hat{Y} = B_0 + B_1X_1$ باعتبار أن Y هو المتغير المعتمد و X1 هو المتغير المستقل مع فترة ثقة 95% لمتوسط المتغير Y $(E(Y_0) = B_0 + B_1X_0)$ للمخطط السابق نتبع الخطوات التالية :

□ من شريط القوائم (شاشة Chart Editor) اختر Options → Chart فيظهر صندوق حوار Scatterplot Options الذي نرتبه كما يلي :



لاحظ أننا قمنا بتأشير Total في Fit Line الذي يعني إضافة خط انحدار واحد لمجمل نقاط شكل الانتشار أي أننا لم نأخذ بالاعتبار المجموعتين الجزئيتين a و b من النقاط التي يحددها المتغير Marker .

◀ عند نقر Fit Options في صندوق الحوار السابق يظهر صندوق حوار Fit Line الذي نرتبه كما يلي :

Scatterplot Options: Fit Line

Fit Method

☒ Linear regression ☐ Lowess

☐ Quadratic regression

☐ Cubic regression

% of points to fit: 50

of iterations: 3

Regression Prediction Line(s)

☒ Mean ☐ Individual

Confidence Interval: 95 %

Regression Options

☒ Include constant in equation

☒ Display R-square in legend

Continue

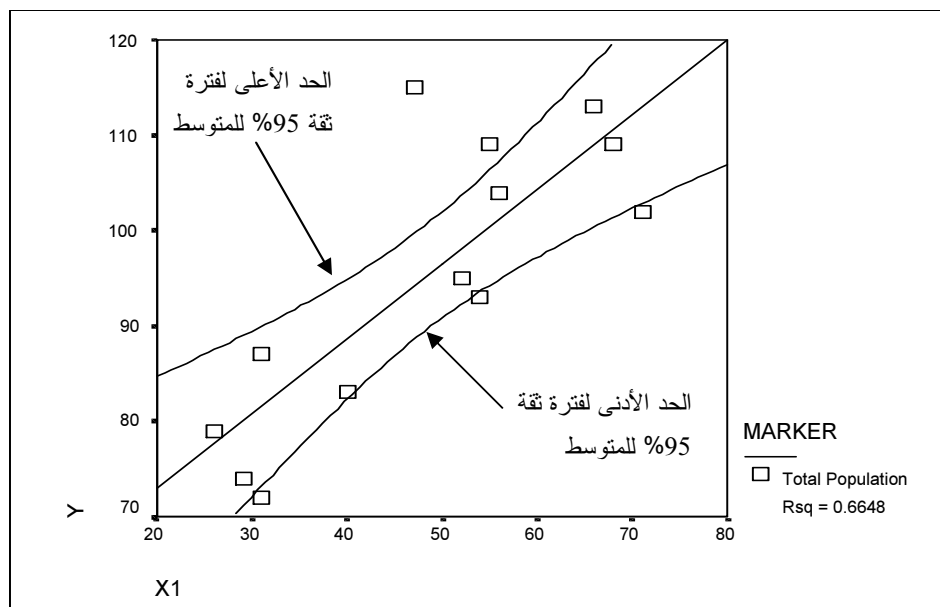
Cancel

Help

لاحظ أننا اخترنا Linear Regression انحدار خطي (يمكن اختيار انحدار تربيعي ، تكعيبي أو استخدام طريقة المربعات الصغرى الموزونة لتقدير منحنى الانحدار) وبما أننا نرغب في رسم فترة ثقة 95% للمتوسط فقد اخترنا Mean من خانة Regression Prediction Line (عند اختيار Individual يعرض البرنامج فترة ثقة للقيمة الكلية $Y_0 = B_0 + B_1X_0 + e_0$) .

□ عند نقر زر continue ثم زر OK يعرض المخطط التالي (بعد إزالة Case Labels) :

مخطط رقم 36

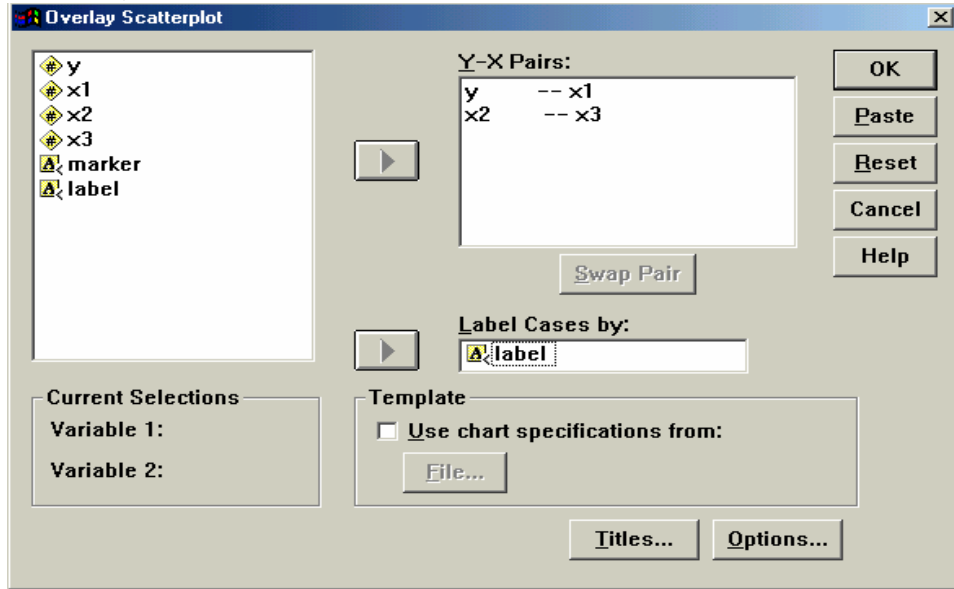


ملاحظة: إذا تم تأشير الخيار Subgroups في خانة Fit Line في صندوق حوار Scatterplot : Options فأن البرنامج يعرض خطي انحدار (خط انحدار لكل فئة من فئتي المتغير Marker) وفي هذه الحالة يجب أن يكون الخيار Show Subgroups في خانة Display Options مؤشراً .

2. شكل انتشار Overlay

يستعمل لتمثيل نقاط الانتشار لزوجين من المتغيرات على الأقل لتمثيل شكل انتشار للمتغيرين Y و X1 وكذلك للمتغيرين X2 و X3 في مخطط واحد تتبع الخطوات التالية :

□ من القوائم نختار Scatter → Graph فيظهر صندوق حوار Scatterplot ومنه نختار Overlay وعند نقر زر Define يظهر صندوق حوار Overlay Scatterplot الذي نرتبه كما يلي :



أن إدخال ازواج المتغيرات في خانة Y-X Pairs يتم كالتالي (مثلا Y و X1) :

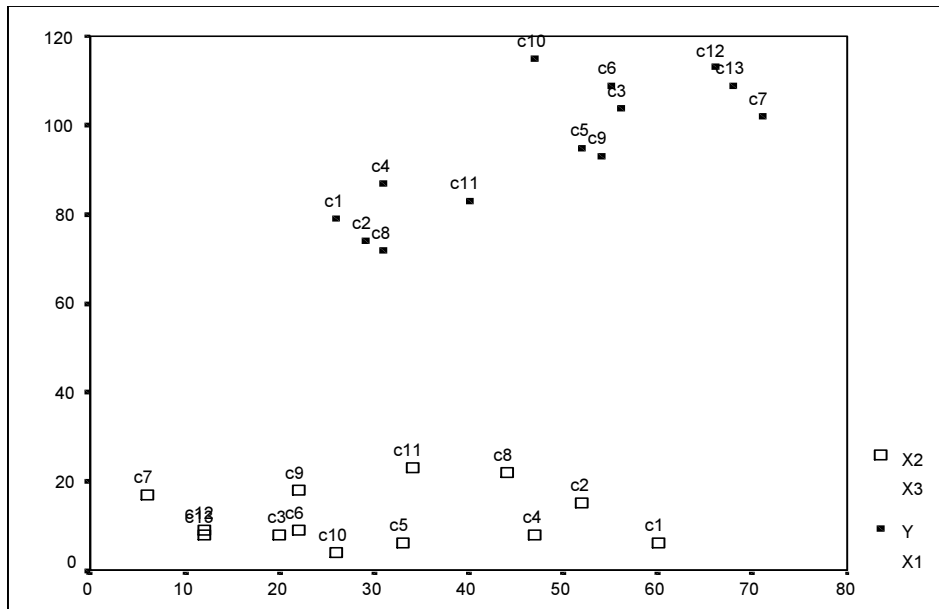
- أنقر المتغير Y لاختياره .
 - أنقر المتغير X1 لاختياره .
 - أنقر [] لنقل المتغيرين الى خانة Y-X Pairs .
- حيث أن المتغير Y يتم تمثيله على المحور العمودي أما المتغير X1 فيمثل على المحور الأفقي ، وب نفس الطريقة يتم إدخال المتغيرين X2 و X3 .

Swap Pair : تستعمل لعكس المحاور لأي زوج من المتغيرات بعد اختياره .

□ أنقر زر Options ثم أنقر Display Chart with Case Labels لعرض عناوين الحالات .

□ عند نقر زر OK في صندوق حوار Overlay Scatterplot يعرض المخطط التالي :

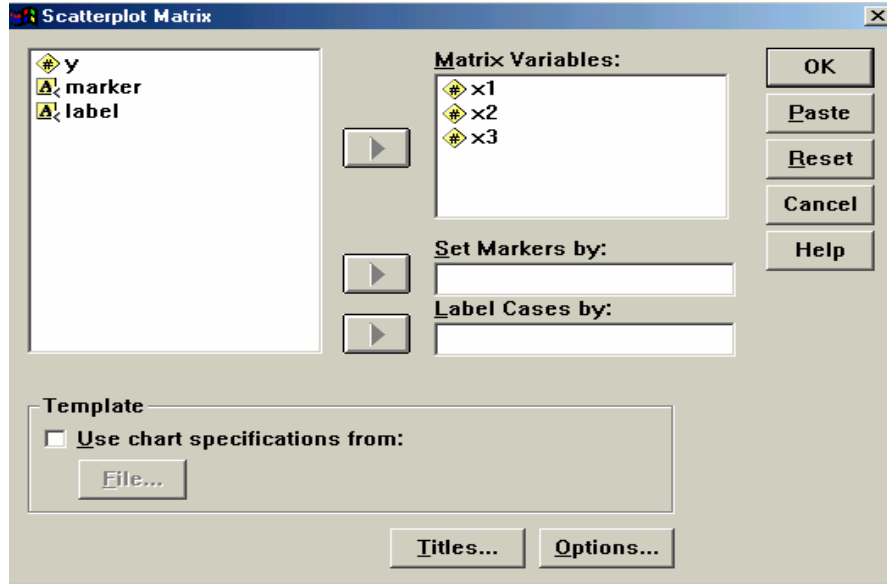
مخطط رقم 37



3. شكل الانتشار Matrix

يستعمل لرسم كل التوافق الممكنة من متغيرين أو أكثر . لو أردنا مثلاً رسم مخطط Matrix للمتغيرات X1 و X2 و X3 نتبع الخطوات التالية :

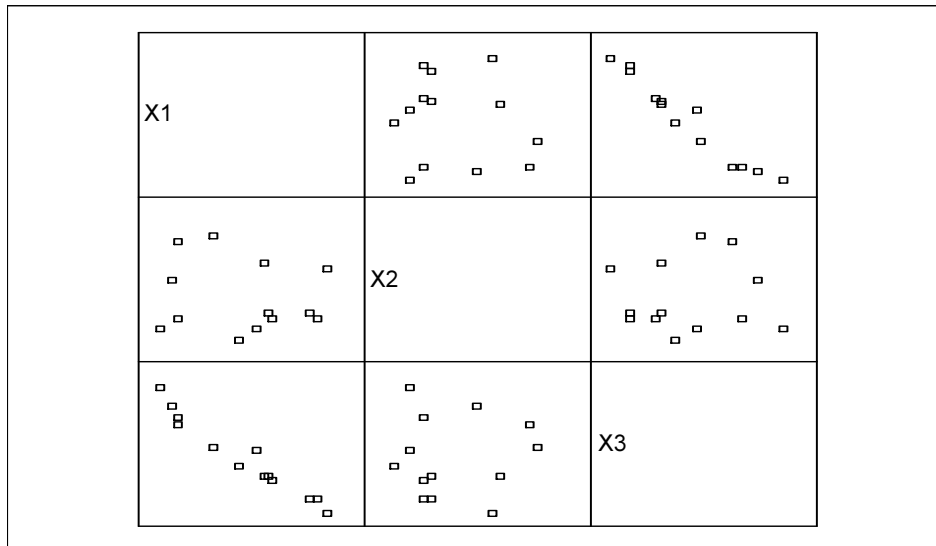
□ من القوائم نختار Scatter → Graph فيظهر صندوق حوار Scatterplot ومنه نختار Matrix فيظهر صندوق حوار Scatterplot Matrix الذي نرتبه كما يلي :



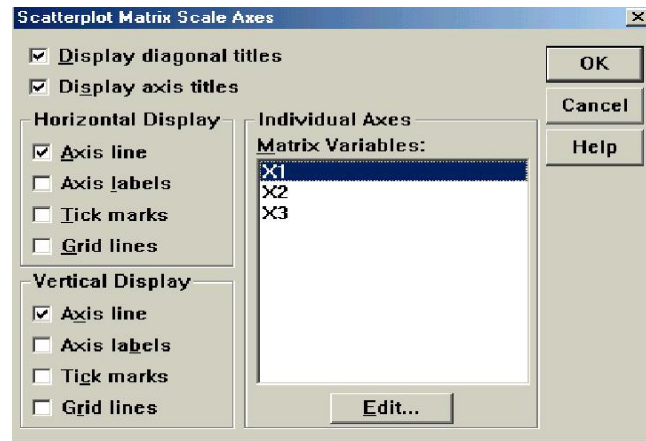
علماً أننا لم ندخل المتغيرين marker و Label اختياريّاً حيث أننا قد أوضحنا تأثير إدخالهما على المخطط.

□ عند نقر زر OK يظهر المخطط التالي :

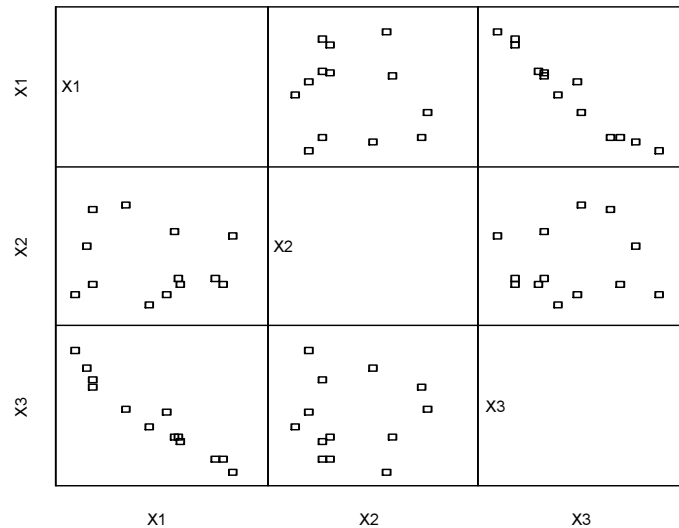
مخطط رقم 38



لإضافة عناوين المحاور Axis Titles أنقر المخطط مرتين بزر الماوس الأيسر ليتحول الى شاشة Chart Editor ثم اختر من شريط القوائم Axis → Chart فيظهر صندوق حوار Scatterplot Matrix Scale Axes حيث نقوم بتأشير الخيار Display Axis Titles ويظهر الصندوق كما يلي :



لتوسيط العناوين انقر زر Edit في صندوق الحوار السابق فيظهر صندوق حوار Edit Selected Axis ومنه اختر Center في خانة Justification. وبنقر زر Continue ثم OK نحصل على المخطط التالي :

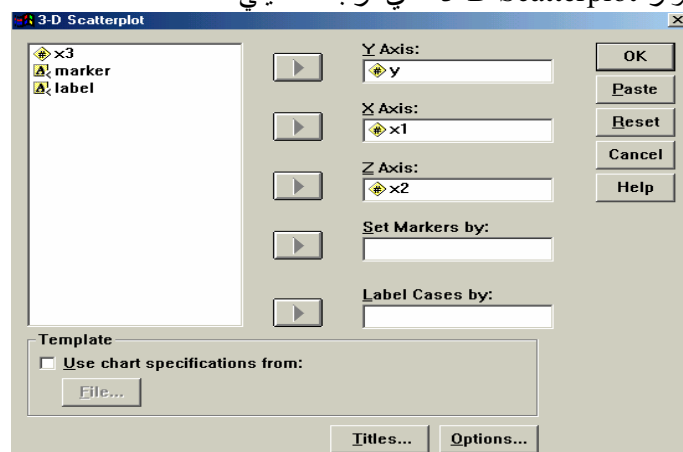


مخطط رقم 39

4. شكل الانتشار ثلاثي الأبعاد 3-D

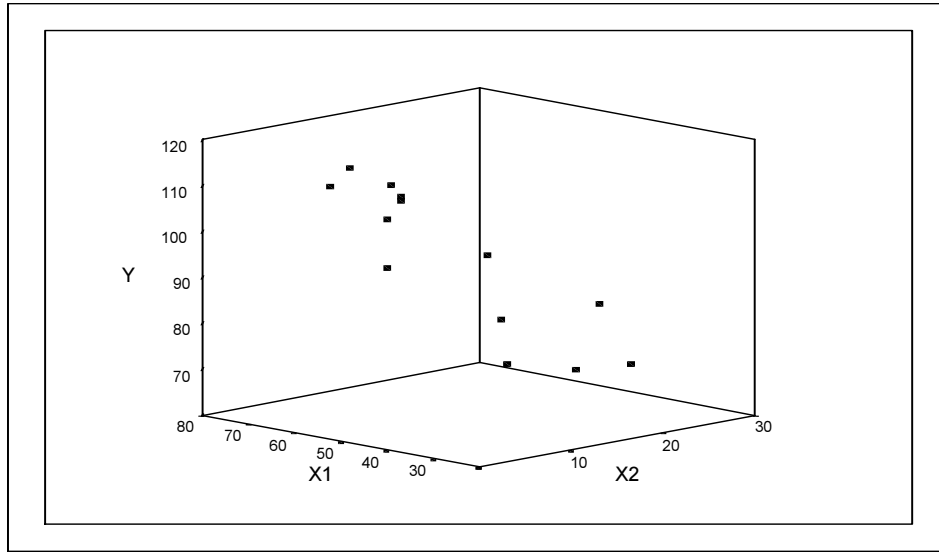
يستعمل لرسم ثلاثة متغيرات في مخطط ثلاثي الأبعاد . لو أردنا مثلاً رسم مخطط 3-D للمتغيرات Y و X1 و X2 نتبع الخطوات التالية :

□ من القوائم نختار Scatterplot → Graph فيظهر صندوق حوار Scatterplot ومنه نختار 3-D فيظهر صندوق حوار 3-D Scatterplot الذي نرتبه كما يلي :



□ عند نقر زر OK يعرض المخطط التالي :

مخطط رقم 40



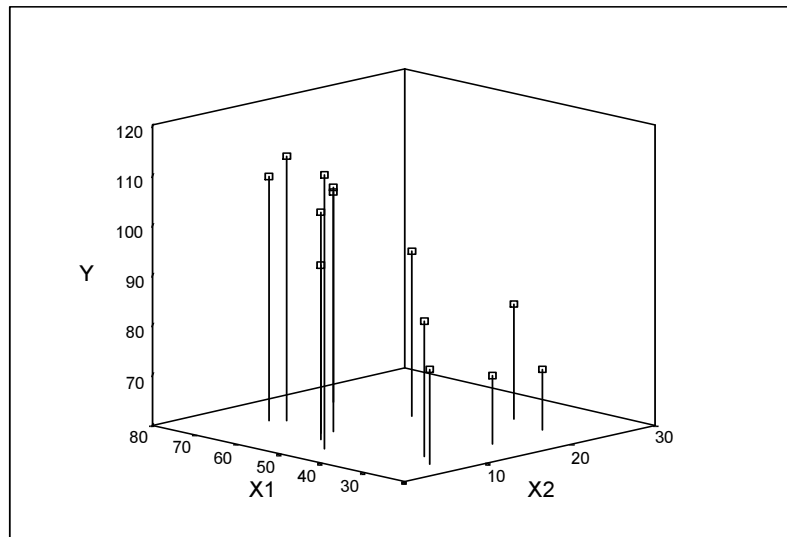
ملاحظة 1:

يمكن تدوير المخطط ثلاثي الأبعاد بنقر الأيقونة  في شريط الأدوات لشاشة Chart Editor أو اختيار Format → 3D-Rotation من شريط القوائم .

ملاحظة 2 :

يمكن إضافة خطوط تمتد من نقاط شكل الانتشار والى أرضية الشكل باختيار Options → Chart من شريط قوائم شاشة Chart Editor فيظهر صندوق حوار 3-D Scatterplot Options ومنه نختار Floor في خانة Spikes وبنقر زر OK نحصل على المخطط التالي :

مخطط رقم 41



علماً أنه يتوفر الخيارين التاليين لخانة Spikes في صندوق حوار 3-D Scatterplot Options :

Centroid : لرسم خط من كل نقطة من نقاط شكل الانتشار الى مركز كافة النقاط في الشكل Centroid .

Origin : لرسم خط من كل نقطة من نقاط شكل الانتشار الى نقطة الأصل Origin .

الفصل الرابع عشر

تبادل البيانات

Data Exchange

أن ملفات البيانات تأتي بصيغ مختلفة ويمكن لبرنامج SPSS أن يتعامل مع الأنواع التالية من الملفات:

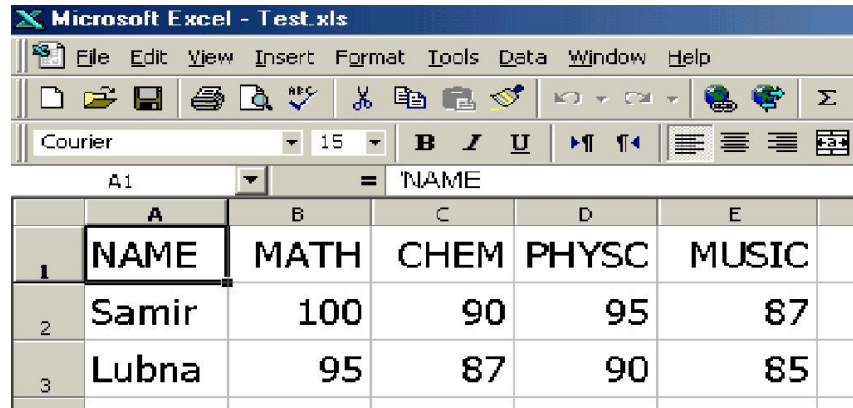
- ☐ أوراق النثر الخاصة ببرنامجي Lotus 1-2-3 و Excel .
 - ☐ ملفات dBase .
 - ☐ ملفات Tab-delimited, الأنواع الأخرى من ملفات نصوص ASCII .
 - ☐ ملفات SPSS المكونة بواسطة أنظمة تشغيل أخرى .
 - ☐ ملفات SYSTAT .
- حيث يمكن استيراد (فتح) ملفات من تطبيقات أخرى وتصدير (خزن) ملفات الى تطبيقات أخرى .

(14 - 1) استيراد الملفات *Importing Data files*

مثال 1 : (استيراد ملف من برنامج EXCEL 97)

الملف التالي Test المخزون في المجلد merge يحتوي قيود مجموعة من الأشخاص في برنامج

Excel97 كما يظهر في الشكل التالي :



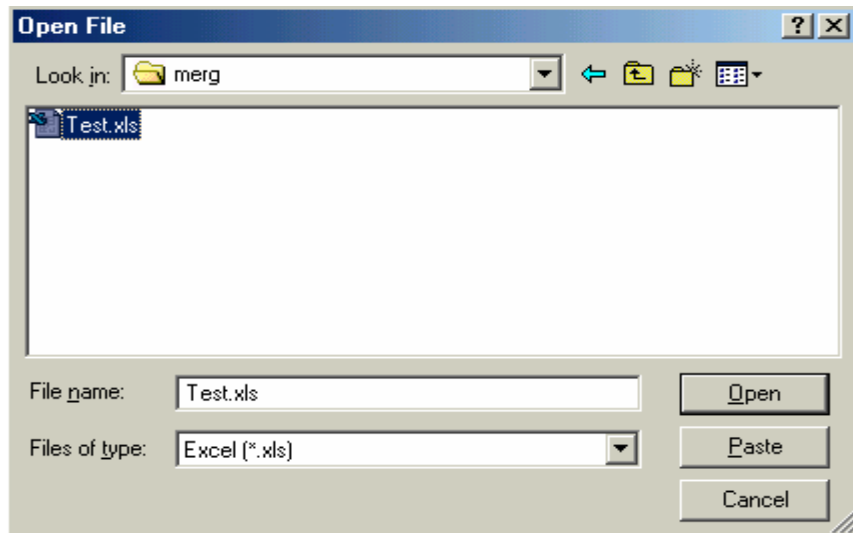
	A	B	C	D	E
1	NAME	MATH	CHEM	PHYSC	MUSIC
2	Samir	100	90	95	87
3	Lubna	95	87	90	85

يطلب استيراد الملف Test من صيغة Excel الى برنامج SPSS .

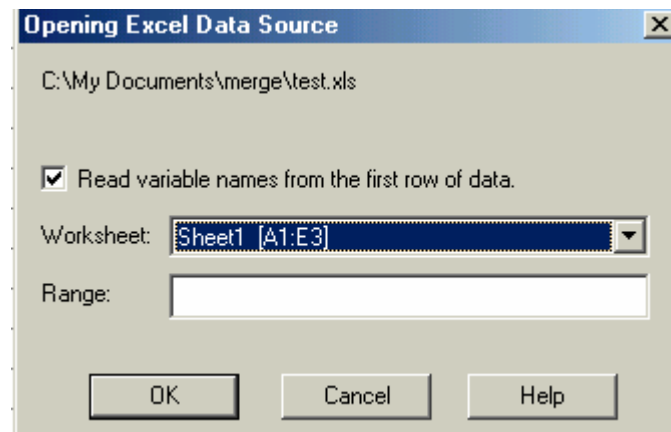
لتنفيذ ذلك نتبع الخطوات التالية :

- ☐ من شريط القوائم في برنامج SPSS اختر Data → Open → File فيظهر

صندوق حوار Open الذي يرتب كما يلي :



لاحظ أننا كتبنا اسم الملف Test في خانة File Name ، أما في خانة Files of Type فقد اخترنا ملفات Excel ذات الاستطالة xls حيث أنه بالإمكان فتح أنواع عديدة من الملفات مثل ملفات SPSS(*.SAV) (وهي ملفات بيانات برنامج SPSS للنظام Windows ذات الاستطالة SAV) وملفات SPSS للنظام DOS وهي SPSS/PC+(*.SYS) وملفات Lotus(*.W*) وملفات dBase(*.dbf) وملفات Text(*.txt) وغيرها من الملفات .
□ عند نقر زر open يظهر صندوق الحوار التالي :



حيث أن :

Read Variable Names From The First Row of data : لقراءة أسماء المتغيرات من الصف الأول في ملف Excel .

Worksheet : لتحديد الورقة التي تحتوي البيانات المراد نقلها الى برنامج SPSS .

Range : لتحديد مدى معين من البيانات التي يراد نقلها الى برنامج SPSS .

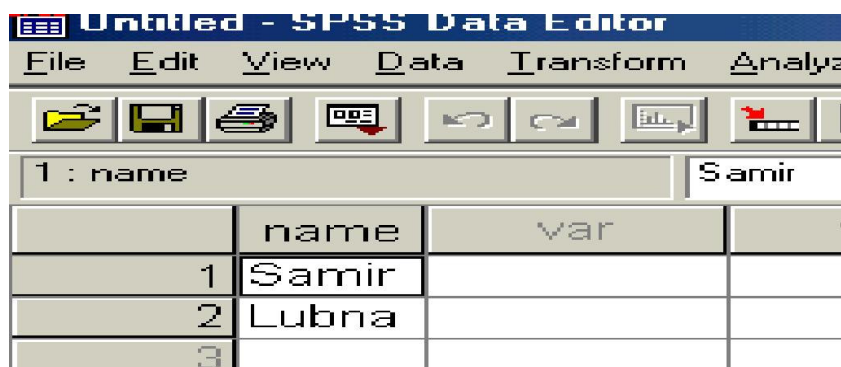
□ عند نقر زر OK يتم نقل ملف Excel الى SPSS وتظهر البيانات في شاشة Data Editor كما يلي :

Untitled - SPSS Data Editor						
File Edit View Data Transform Analyze Graphs Utilities Window Help						
1 : name Samir						
	name	math	chem	physc	music	
1	Samir	100	90	95	87	
2	Lubna	95	87	90	85	
3		.	.	.	.	

ويمكن تخزين الملف السابق كملف SPSS ذو الاستطالة SAV .

ملاحظات :

1. عند تحديد Range معين من البيانات في حقل Range في صندوق حوار Opening Excel Data Source (مثلاً A1:A3) يظهر الناتج التالي :



	name	var
1	Samir	
2	Lubna	
3		

2. يمكن تحديد أي ورقة عمل في الملف مثلاً Sheet2 بدلاً من Sheet1 في خانة Worksheet في صندوق حوار Opening Excel Data Source لاستيراد البيانات الموجودة في هذه الورقة أو أي Range من بيانات هذه الورقة .

3. عند عدم تأشير الخيار Read Variable Names From The First Row of data في المثال السابق فإن أسماء المتغيرات في ملف Excel سوف تضاف كأول قيمة لكل متغير في ملف SPSS ويتم إضافة أسماء متغيرات افتراضية بدلاً عنها .

4. عند وجود أكثر من متغير بنفس الاسم في ملف Excel فإن كل متغير سوف يكون له أسم وحيد Unique عند استيراد الملف الى برنامج SPSS كما أن أسم كل متغير في برنامج Excel له رموز Characters يزيد عددها عن 8 فسيتم قطع الاسم لغاية 8 رموز ويتم إضافة الاسم الكامل كعنوان للمتغير Label عند استيراد الملف الى برنامج SPSS .

5. يمكن استيراد بيانات ملف Excel بطريقة ثانية بتظليل البيانات المطلوب نقلها في برنامج Excel ثم اختيار Copy → Edit ثم الانتقال الى برنامج SPSS واختيار Paste → Edit من شريط القوائم في شاشة Data View فيتم نقل البيانات الى برنامج SPSS ولكن يجب ملاحظة فقدان أسماء المتغيرات الحقيقية وظهور أسماء افتراضية بدلاً عنها كما نلاحظ فقدان قيم المتغيرات الرمزية كون النوع الافتراضي لمتغيرات SPSS هو النوع العددي Numeric ويتوجب في هذه الحالة اختيار النوع String للمتغيرات الرمزية في شاشة Data Editor قبل إنجاز عملية النسخ .

مثال 2 : (استيراد ملفات النصوص Text Files)

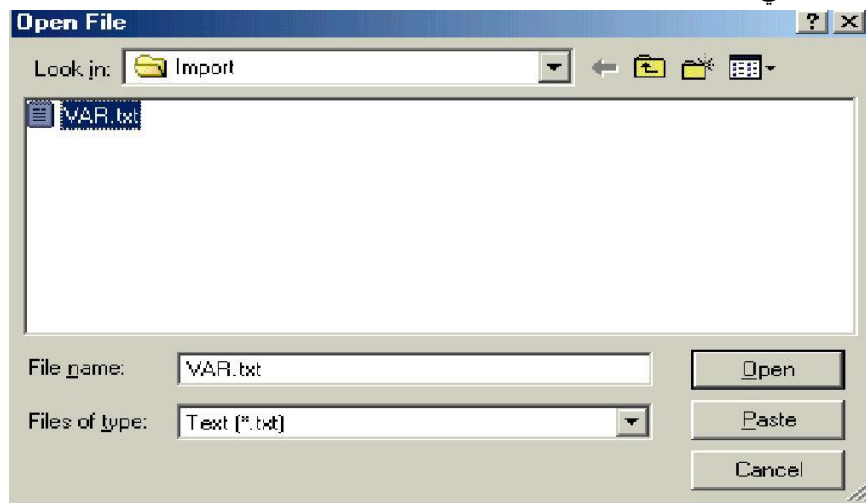
الملف التالي يحتوي بيانات تم إدخالها عن طريق معالج النصوص MS-DOS Editor لبرنامج MS-DOS والمخزون بأسم Var.txt وقد استخدمت الفراغات Spaces للفصل بين المتغيرات وكما يلي :

x1	x2	x3
12	30	33
16	45	60
29	64	88

لأستيراد الملف أعلاه الى برنامج SPSS نتبع الخطوات التالية:

□ من شريط القوائم اختر **File → Read Text Data** فيظهر صندوق حوار **Open File**

الذي نرتبه كما يلي :



كما يمكن انجاز هذه الخطوة باختيار **Data → Open → File** من شريط القوائم وفي هذه الحالة يتوجب تحديد نوع الملف **Text** في حقل **Files of type** في صندوق الحوار أعلاه .

□ عند نقر زر **Open** في صندوق الحوار أعلاه يظهر صندوق حوار **Text Import wizard** (step 1 of 6) وكما يلي :

Text Import Wizard - Step 1 of 6

Welcome to the text import wizard!

This wizard will help you read data from your text file and specify information about the variables.

Does your text file match a predefined format?

☐ Yes ☒ No

Text file:

	var1	var2	var3
1			
2			
3			
4			

1 2 3 4

0 10 20 30 40 50

1	x1	x2	x3
2	12	30	33
3	16	45	60
4	29	64	88

< Back Next > Finish Cancel Help

□ عند نقر زر Next يتم الانتقال الى الخطوة رقم 2 وكما يلي :

Text Import Wizard - Step 2 of 6

How are your variables arranged?

☒ Delimited - Variables are delimited by a specific character (i.e., comma, tab).
☐ Fixed width - Variables are aligned in fixed width columns.

Are variable names included at the top of your file?

☒ Yes
☐ No

Text file:

0 10 20 30 40 50

1	x1	x2	x3
2	12	30	33
3	16	45	60
4	29	64	88

< Back Next > Finish Cancel Help

في حقل How are your Variables arranged? فقد أشرنا الخيار Delimited الذي يستعمل للمتغيرات التي يفصل بينها فاصل من نوع معين (فارزة، فارزة منقوطة، فراغ، الخ) .
 في حقل Are Variables Names included at the top of your file? فقد أشرنا الخيار yes .
 عند نقر زر Next يتم الانتقال الى الخطوة 3 وكما يلي :

The first case of data begins on which line number? 2

How are your cases represented?

☒ Each line represents a case

☐ A specific number of variables represents a case: 3

How many cases do you want to import?

☒ All of the cases

☐ The first 1000 cases.

☐ A random percentage of the cases (approximate): 10 %

Data preview

1	12	30	33
2	16	45	60
3	29	64	88

< Back Next > Finish Cancel Help

في حقل the first Case of data Begins on Which Line Number ? فقد حددنا ابتداء الحالة الأولى بالسطر الثاني من الملف لأن السطر الأول مخصص لأسماء المتغيرات .
 في حقل How are Your Cases represented? فقد أشرنا الخيار Each Line Represents a Case أما في حالة وجود أكثر من حالة في السطر الواحد فنستخدم الخيار الثاني حيث يتوجب (في هذه الحالة) تحديد عدد المتغيرات التي تكون حالة واحدة .
 في حقل How Many Cases Do you want to Import ? فقد تم تأشير الخيار الأول - كافة الحالات .

□ عند نقر زر Next ننتقل الى الخطوة الرابعة من الـ Wizard وكما يلي :

Text Import Wizard - Delimited Step 4 of 6

Which delimiters appear between variables?

☐ Tab
 ☒ Space
 ☐ Comma
 ☐ Semicolon
 ☐ Other:

Data preview

x1	x2	x3
12	30	33
16	45	60
29	64	88

الغاية من هذه الخطوة هي تحديد نوع الفاصل Delimiter وكما ذكرنا فأنا قد أيسر عملنا الفراغات كفاصل بين المتغيرات عند الإدخال في معالج النصوص Edit وقد تم تأشير الخيار Space (يرجى ملاحظة أن معظم الخيارات تؤثر تلقائياً من قبل البرنامج) .

□ عند نقر زر Next يتم الانتقال الى الخطوة الخامسة من ال wizard

Text Import Wizard - Step 5 of 6

Specifications for variable(s) selected in the data preview

Variable name:

Data format:

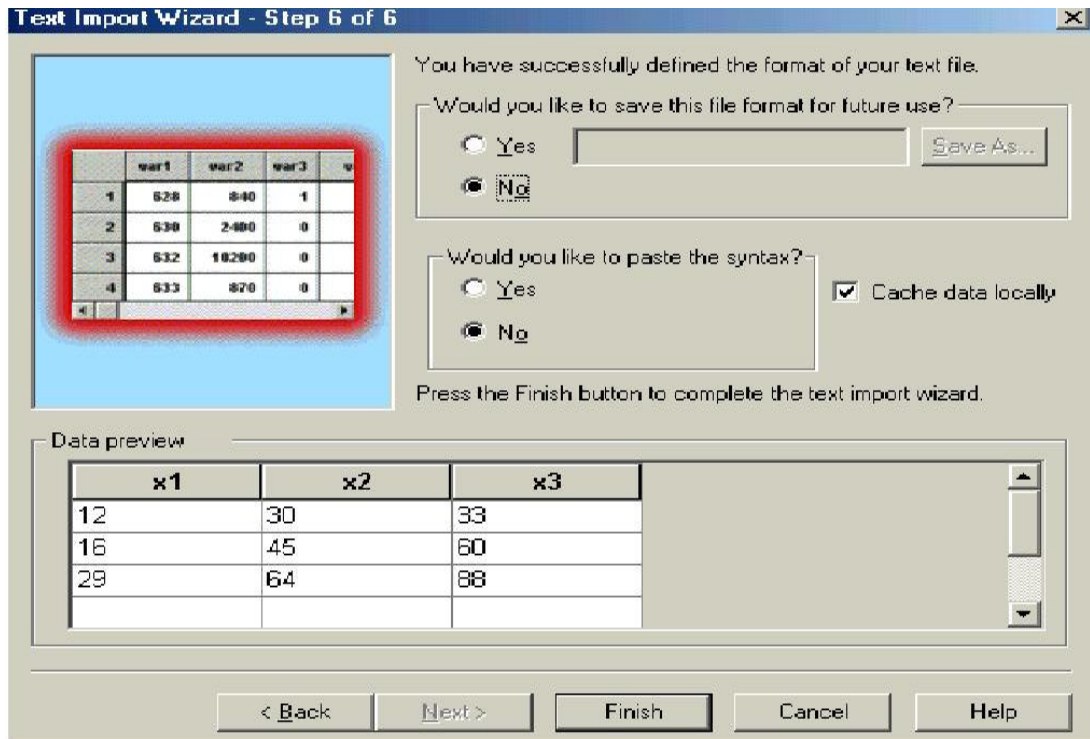
Do Not Import

Data preview

x1	x2	x3
12	30	33
16	45	60
29	64	88

يمكن في هذه الخطوة تغيير أسم ونوع Format & Name أي من المتغيرات بنقر خلية أسم المتغير في Data Preview أسفل الصندوق ثم كتابة الأسم الجديد ونوع المتغير في حقل Variable Name و Data Format .

□ عند نقر زر Next يتم الانتقال الى الخطوة السادسة والأخيرة من الـ Wizard



في حالة الرغبة في تخزين صيغة البيانات التي تم استيرادها لغرض تطبيقها في المستقبل على بيانات مستوردة مباشرة نؤشر الخيار Yes كإجابة عن السؤال Would You Like to save This Format for Future Use? وفي هذه الحالة انقر Save As لحزن الصيغة في ملف معين (نوع استضافة tpf) حيث يمكن استرجاع الصيغة المخزونة مسبقاً وتطبيقها على البيانات المستوردة عند بداية استيراد الملف في الخطوة الأولى من الـ Wizard حيث يتم الإجابة بـ Yes على السؤال Does Your Text File Match a Predefined Format? ثم نقر الزر Browse لاسترجاع ملف الصيغة. لعدم الرغبة في تخزين صيغة البيانات فقد أشرنا الخيار NO.

□ عند نقر زر Finish يتم إضافة البيانات الى شاشة Data Editor وكما يلي :

	x1	x2	x3
1	12.0	30.0	33.0
2	16.0	45.0	60.0
3	29.0	64.0	88.0

(14-2) تصدير الملفات Exporting Data Files

يمكن تصدير ملفات SPSS الى تطبيقات أخرى وذلك عن طريق تخزين الملف بصيغ عديدة منها مايلي :

1. SPSS(*.SAV) وهي ملفات البيانات للإصدار الحالي للبرنامج .
2. SPSS 7.0(*.SAV) وهي ملفات البيانات للإصدار السابع للبرنامج .
3. SPSS/PC+(*.SYS) وهي ملفات بيانات SPSS للنظام DOS .
4. Tab-delimited(*.dat) وهي ملفات النصوص Text .
5. Fixed ASCII(*.dat)
6. Excel(*.xls)
7. Lotus1-2-3(*.w*) للإصدارات 1,2,3 .
8. dBase(*.dbf) للإصدارات II و III و IV .

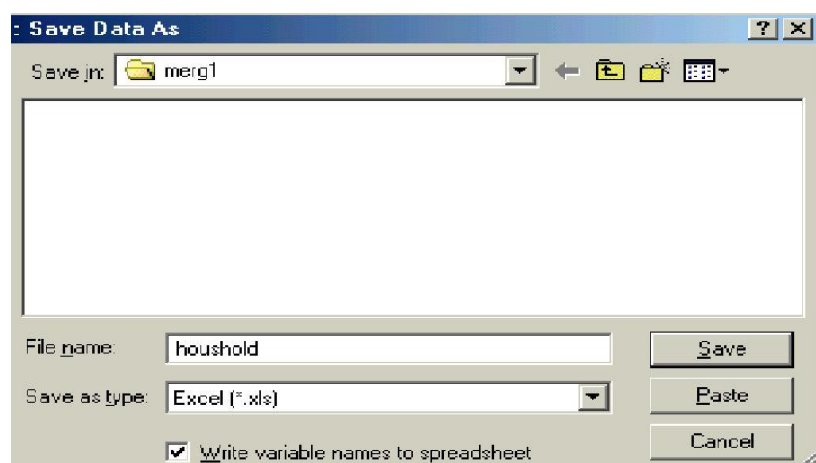
مثال 3: (على تصدير البيانات)

لدينا الملف household.sav الذي يظهر كما يلي في شاشة Data Editor لبرنامج SPSS :

household.sav - SPSS Data Editor					
File Edit View Data Transform Analyze Graphs Utilities Window Help					
1 : name Ahmad					
	name	age	edu	housno	
1	Ahmad	20	Sec	10	
2	Zeki	35	Bsc	10	
3	Sabah	30	sec	10	
4	Zainab	15	Prim	10	
5	Ibrahim	17	Sec	12	
6	Samir	40	Ma	12	
7	Selma	36	Bs	12	

المطلوب مايلي :

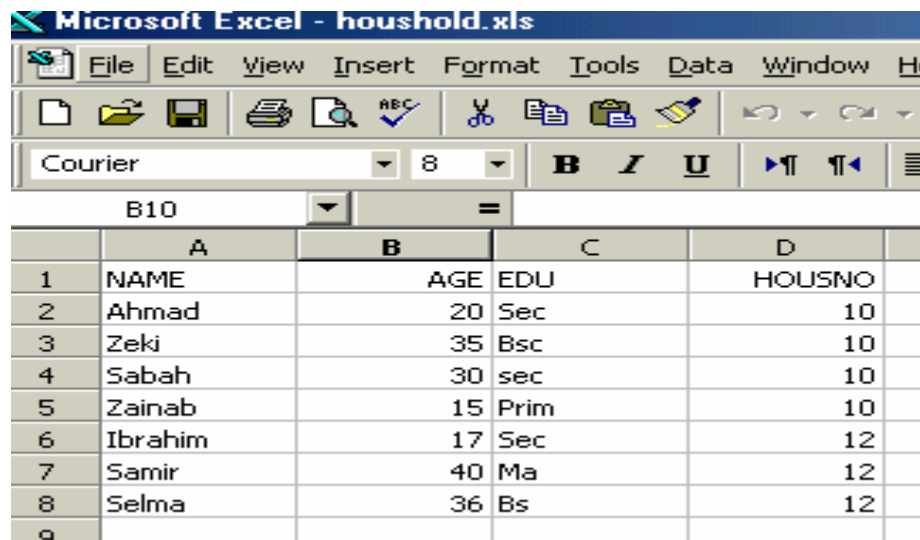
1. تصدير الملف الى برنامج Excel .
 2. تصدير الملف الى معالج النصوص MS-Editor .
1. لتصدير الملف الى برنامج Excel نتبع الخطوات التالية :
- من شريط القوائم اختر Save As → File فيظهر صندوق حوار Save AS الذي نرتبه كما



يلي :

لاحظ أنه تم إعطاء أسم للملف في حقل File Name وتم تحديد نوع الملف في حقل Save as Type كملف من ملفات Excel ذو الاستطالة xls .

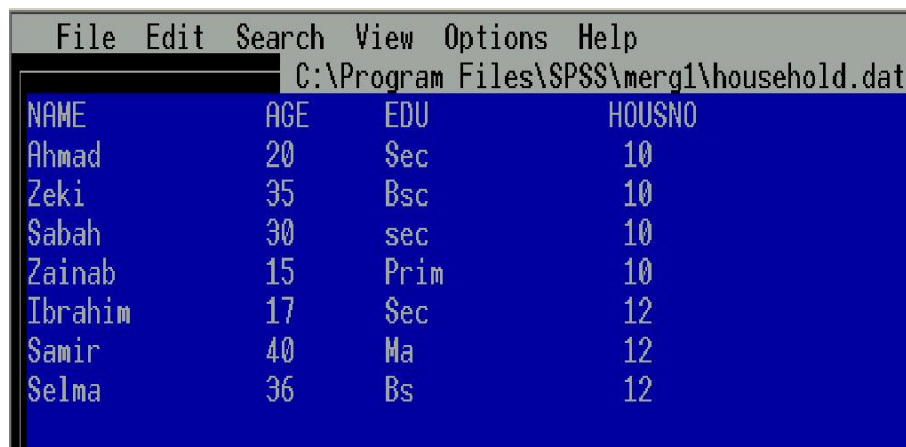
عند تأشير write variable names to spreadsheet يتم إضافة أسماء المتغيرات الى ورقة عمل
نشر Excel في السطر الأول منها أما عدم التأشير فيؤدي الى إزالة أسماء المتغيرات من السطر الأول .
□ عند نقر زر Save يتم تخزين الملف بصيغة ملفات Excel أي سيكون لدينا ملفين ضمن المجلد
merg1 أحدهما هو household ذو الاستطالة sav والآخر بنفس الاسم ولكن بالاستطالة xls .
يمكن فتح الملف household في برنامج Excel بأختيار Open → File ثم تحديد موقع الملف Path
حيث يظهر في ورقة Excel كما يلي :



The screenshot shows the Microsoft Excel interface with the file 'household.xls' open. The spreadsheet has columns labeled NAME, AGE, EDU, and HOUSNO. The data is as follows:

	A	B	C	D
	NAME	AGE	EDU	HOUSNO
2	Ahmad	20	Sec	10
3	Zeki	35	Bsc	10
4	Sabah	30	sec	10
5	Zainab	15	Prim	10
6	Ibrahim	17	Sec	12
7	Samir	40	Ma	12
8	Selma	36	Bs	12

2. لتصدير الملف الى MS-Editor نكرر نفس الخطوة الأولى (عند التصدير الى Excel) ولكن نختار
Tab-delimited في حقل Save as Type في صندوق حوار Save Data As ثم نعطي اسم
الملف household. وعند نقر زر Save يتم تخزين الملف بصيغة Tab-delimited أي يضاف ملف جديد
للملفين السابقين داخل المجلد merg1 باسم household أيضاً ولكن باستطالة dat ، ويظهر الملف
household بفواصل ثابتة عند فتحه في MS-Editor كما في الشكل التالي.



The screenshot shows the MS-Editor interface with the file 'C:\Program Files\SPSS\merg1\household.dat' open. The data is displayed in a text format with columns labeled NAME, AGE, EDU, and HOUSNO. The data is as follows:

NAME	AGE	EDU	HOUSNO
Ahmad	20	Sec	10
Zeki	35	Bsc	10
Sabah	30	sec	10
Zainab	15	Prim	10
Ibrahim	17	Sec	12
Samir	40	Ma	12
Selma	36	Bs	12

الفصل الخامس عشر

كتابة الأوامر

Syntax Commands

أن معظم الأوامر يمكن الحصول عليها من القوائم أو في صناديق الحوار ولكن هناك أوامر لا تتوفر إلا باستخدام لغة الأوامر Command Language حيث يمكن من خلال هذه اللغة تنفيذ فعاليات لا تتوفر ضمن أوامر SPSS وهذه اللغة تسمح بخزن الفعاليات Jobs في ملف الأوامر Syntax File حيث يمكن استرجاعها في أي وقت وتنفيذها لأي مجموعة من المتغيرات .

(1-15) ملف الأوامر Syntax File : وهو عبارة عن ملف نصوص Text File يحتوي على أوامر مكتوبة Commands (لغة برمجة) ويجب مراعاة القواعد التالية عند كتابة هذه الأوامر :

- يجب أن يبدأ كل أمر بسطر جديد وينتهي بالفترة (.) .
- الأوامر الفرعية Sub Commands تفصل بعلامة Slash (/) ، علماً أن وضع علامة Slash قبل أول أمر فرعي يكون اختياريًا .
- يجب كتابة الأسماء الكاملة للمتغيرات .
- النص المحصور بين علامتي اقتباس Quotation Marks يجب أن يكتب في سطر واحد .
- لا يزيد طول السطر في ملف الأوامر عن 80 رمز .
- تستخدم الفترة (.) للتعبير عن الفاصلة العشرية بغض النظر عن إعدادات Windows الموجودة في Regional Setting .

للقوف على تفاصيل كتابة البرامج نحيل القارئ إلى SPSS Syntax reference Guide .

(2- 15) الطرق المساعدة في بناء ملفات الأوامر Command Syntax

أن كتابة الأوامر تحتاج إلى معرفة تفصيلية وممارسة طويلة في كتابة البرامج مما لا يتوفر لدى الكثيرين من مستخدمي برنامج SPSS . عوضاً عن ذلك يمكن الاستعانة بأحد الوسائل الثلاثة المبسطة التالية في تكوين ملف الأوامر :

- كتابة الأوامر من صناديق الحوار Dialog boxes .
- نسخ الأوامر من Log في مخرجات البرنامج .
- نسخ الأوامر من Journal File .

(1 - 2 - 15) كتابة الأوامر من صناديق الحوار Dialog boxes : وهذه أسهل طريقة لبناء ملف

الأوامر وذلك يفتح صندوق الحوار لامر معين ثم استعمال الزر Paste في لصق الأوامر في ملف الأوامر syntax File كما هو واضح في المثال التالي :

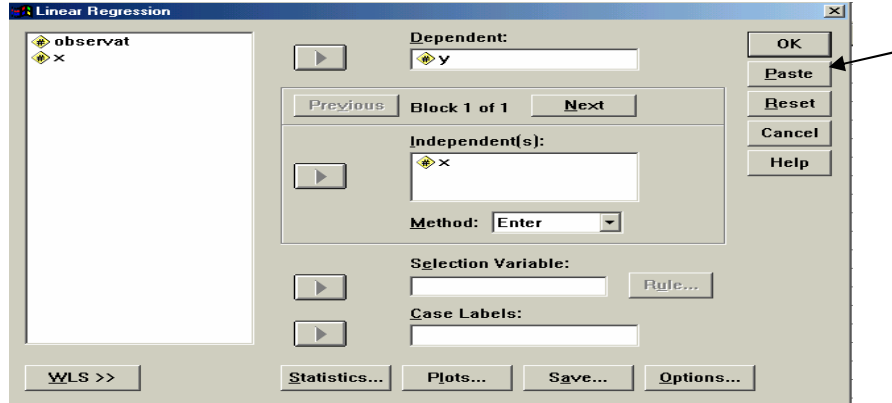
مثال 1 : لنفس بيانات المثال 3 الوارد في البند (10 - 4 - 1) حيث تم إدخال المتغيرين X و Y في شاشة Data Editor ونرغب في بناء ملف أوامر بطريقة الكتابة من صندوق الحوار يتضمن الفعاليات التالية :

1. استخراج معاملات Coefficients انحدار Y/X .
2. استخراج جدول تحليل التباين ANOVA .

3. استخراج معامل Durbin-Watson .

لتنفيذ ذلك نتبع الخطوات التالية :

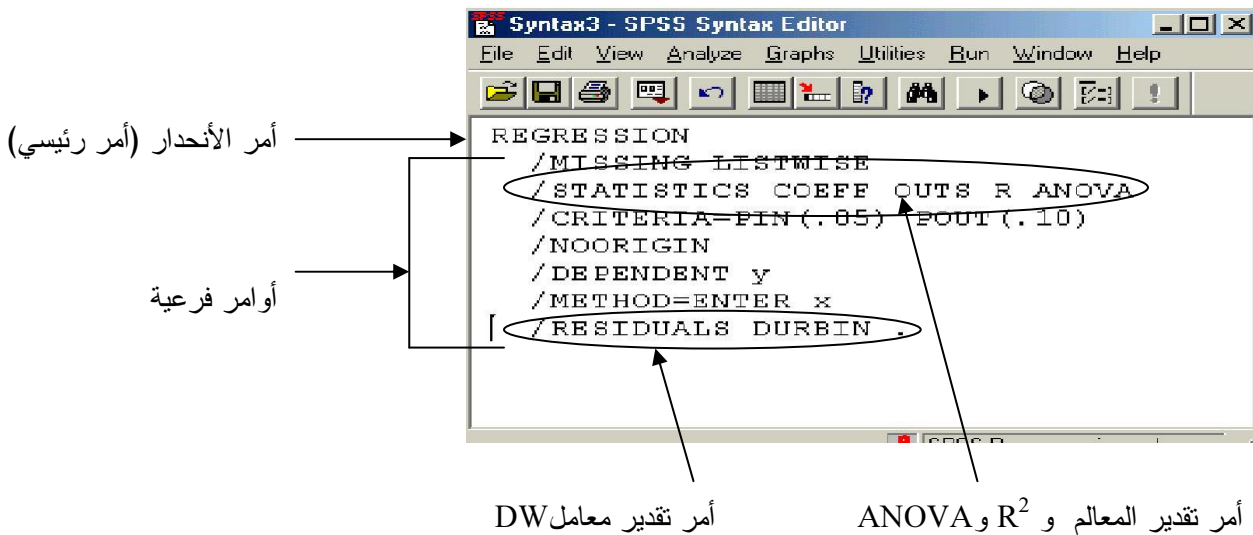
1. من شريط القوائم اختر Linear Regression → Analyze فيظهر صندوق حوار Linear Regression الذي نرتبه كما يلي :



2. في صندوق الحوار أعلاه انقر Statistics ثم أشر الخيارات التالية في صندوق حوار Statistics :

- Estimate لتقدير معالم النموذج .
- Model Fit لتكوين جدول تحليل التباين ANOVA واستخراج R^2 .
- Durbin - Watson لتقدير معامل DW .

3. انقر زر Paste في صندوق حوار Linear Regression فيتم فتح ملف الأوامر Syntax File تلقائياً (بأسم افتراضي) ولصق أوامر الانحدار المختارة فيه وكما يلي :



لحزن ملف الأوامر بأسم معين من شريط القوائم في نافذة Syntax Editor اختر File → Save As ثم أعطي أسم الملف ونوعه (SPSS Syntax File) الذي يكون ذو الاستطالة SPS .

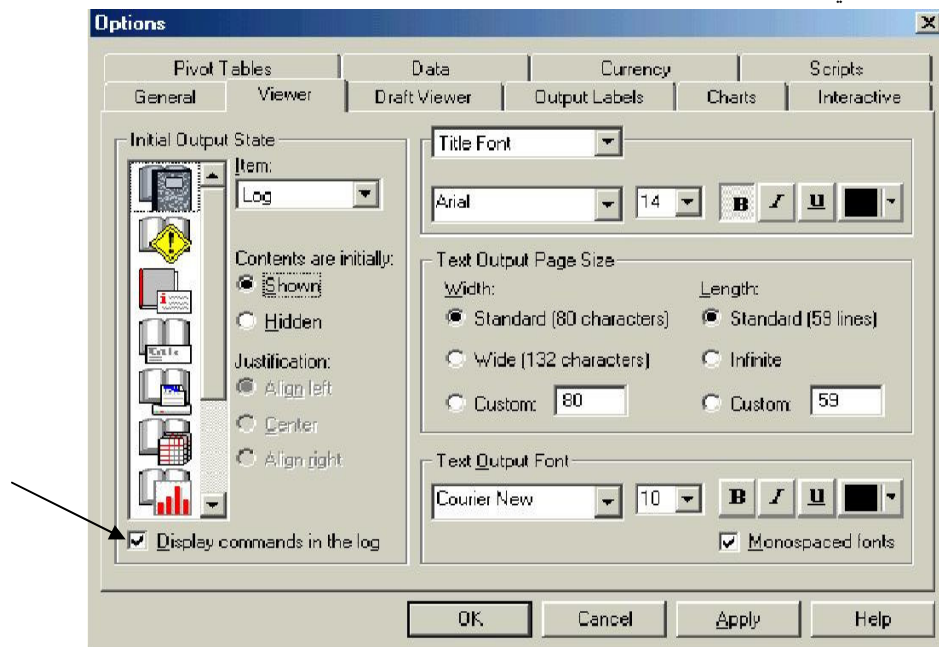
(2 - 2 - 15) نسخ الأوامر من Log في مخرجات البرنامج

يمكن نسخ أوامر من Log في مخرجات البرنامج عند تنفيذ أمر أي أسلوب إحصائي كالانحدار مثلاً وقبل ذلك يجب إظهار log في مخرجات البرنامج كما يلي :

□ من شريط القوائم لنافذة Data Editor اختر Options → Edit فيظهر صندوق حوار Options .

□ أنقر عروة Viewer في صندوق حوار Options للانتقال الى صندوق حوار Viewer.

□ أشر الخيار Display Commands in the Log إذا لم يكن مؤشراً .حيث يظهر صندوق حوار Viewer كما يلي :

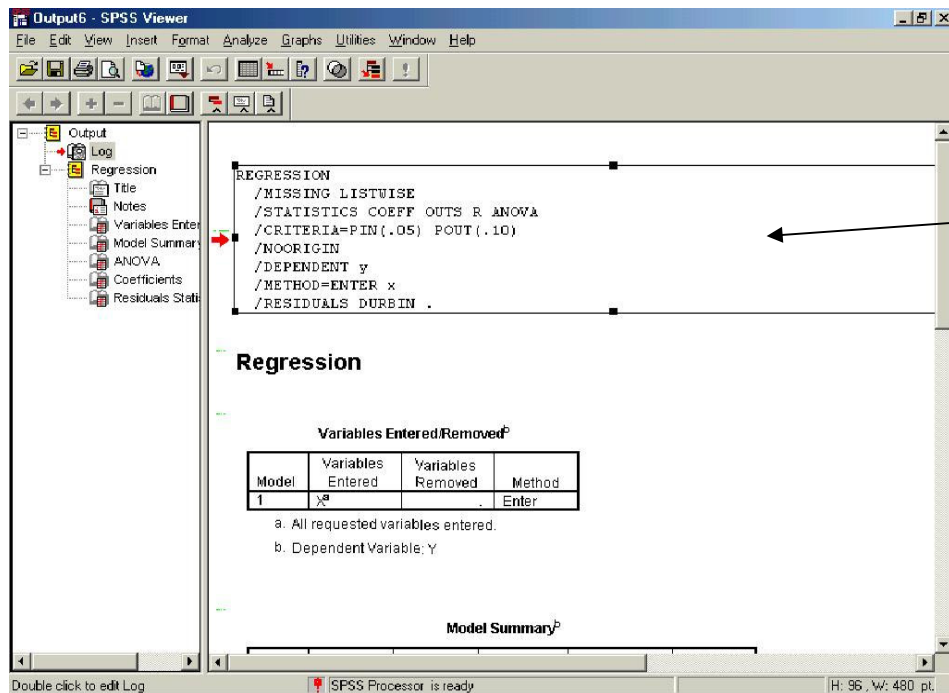


□ أنقر زر OK . أن هذا سيجعل الأوامر تظهر مع مخرجات أي أسلوب إحصائي في شاشة المخرجات Viewer.

مثال 2 :

لنفس المثال السابق يطلب تكوين ملف الأوامر بطريقة النسخ من Log .

□ لتنفيذ ذلك نتبع الخطوتين 1 و 2 للمثال السابق ثم أنقر زر OK في صندوق حوار Linear Regression فتظهر مخرجات الانحدار في شاشة viewer كما يلي :



Log

□ أفتح ملف أوامر جديد باختيار الأمر **File → New → Syntax** من شريط القوائم في شاشة **SPSS Viewer**.

□ انقر في شاشة **SPSS Viewer** مرتين بزر الماوس الأيسر لتفعيله.

□ ظلل الأوامر التي ترغب بنسخها (ترغب مثلاً بنسخ كافة الأوامر).

□ من قوائم **Viewer** اختر **Edit → Copy**.

□ من القوائم في نافذة **Syntax** اختر **Edit → Paste**.

□ أخرج ملف الأوامر **Syntax** الناتج للاستفادة منه فيما بعد.

(15 - 2 - 3) نسخ الأوامر من **Journal File**

أن كافة الأوامر التي تنفذ خلال الجلسة يتم تسجيلها في **Journal File** يسمى **spss.jnl** وعليه يمكن استدعاء هذا الملف (يتواجد ضمن الموقع الافتراضي **C:\WINDOWS\TEMP**) حيث يمكن تنقيح الملف بحذف الأوامر غير المطلوبة ورسائل الخطأ ثم خزن الملف الناتج بأسم مختلف كملف أوامر **Syntax** باستطالة **SPS** حيث يمكن استدعائه وتمشيته في أي وقت.

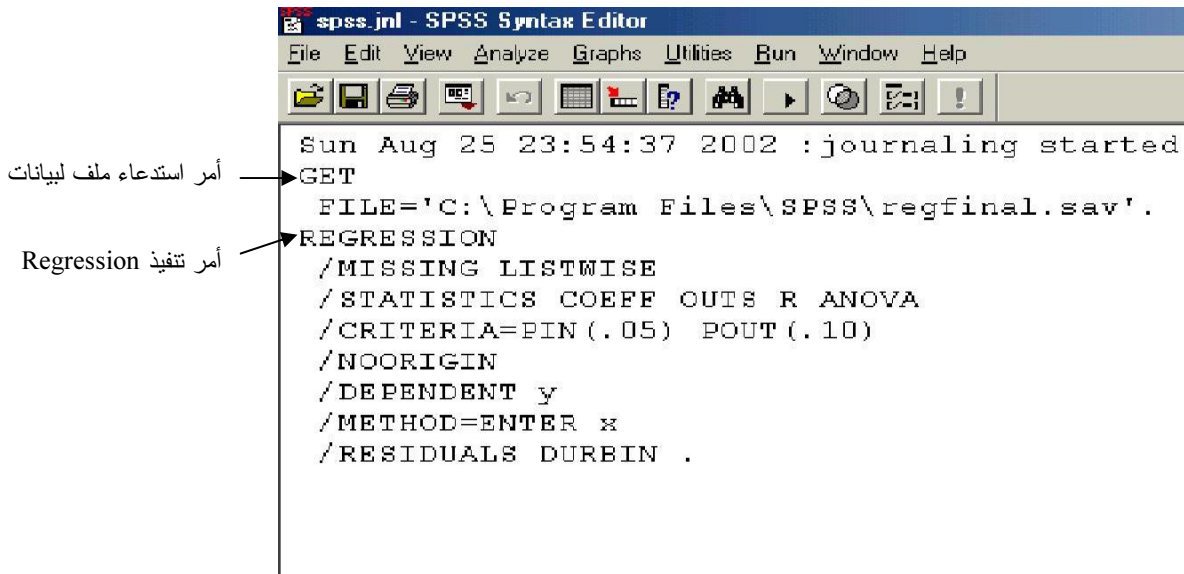
ملاحظة : يمكن التحكم بإعدادات **Journal File** باختبار الأمر **Edit → Options → General** في شريط القوائم لشاشة **Data Editor** حيث يمكن التحكم بموقع الملف. لغرض جعل البرنامج يقوم بتخزين الأوامر المنفذة في **Journal File** يجب تأشير الخيار **Record Syntax in Journal** (إذا لم يكن مؤشراً) حيث تتوفر طريقتين لخزن الأوامر :

Append : تخزن الأوامر المنفذة في الجلسات المتعاقبة ابتداء من تنصيب البرنامج.

Overwrite : تخزن الأوامر المنفذة خلال الجلسة الحالية فقط لبرنامج **SPSS**.

مثال 3 : للمثال الأول يطلب تكوين ملف الأوامر بطريقة النسخ من Journal File .

- لتتفيذ ذلك نتبع الخطوتين 1 و 2 للمثال الأول ثم أنقر زر OK في صندوق حوار Linear Regression فتظهر مخرجات الانحدار في شاشة SPSS Viewer .
- من شريط القوائم في شاشة SPSS Viewer أختار File → Open → Other ثم افتح الملف spss.jnl في الموقع C:\WINDOWS\TEMP فيظهر الملف كما يلي :



- يمكن تنقيح الملف spss.jnl في هذا المثال نرغب بالاحتفاظ به كما هو عدا أننا نقوم بحذف السطر الأول (تاريخ الخزن) لأنه ليس من أوامر البرمجة . الآن نقوم بخزن الملف بأسم مختلف كملف Syntax ذو الاستطالة SPS لأستدعائه وتمشيته عند الحاجة .
- ملاحظة : نوع الخزن المستخدم لـ journal file في هذا المثال هو Overwrite (أي أوامر الجلسة الحالية فقط) .

(15-3) تمشية ملف الأوامر To Run command syntax

- لتمشية ملف الأوامر (أو تمشية البرنامج) أو أي جزء منه نتبع الخطوات التالية :
- أفتح ملف الأوامر Syntax File المخزون مسبقاً بالأمر Syntax Open من File من شريط القوائم .
- من شريط القوائم في شاشة SPSS Syntax Editor أختار Run لتمشية البرنامج حيث تتوفر الخيارات التالية :
- All : لتمشية كافة أوامر الملف .

Selection : لتمشية الأوامر المختارة (المظللة) Highlighted .

Current : لتمشية الأوامر في موقع المؤشر Cursor .

To End : لتمشية الأوامر من موقع المؤشر والى نهاية الملف .

مثال 4 : لنفس بيانات المثال الأول (يتضمن المتغيرين X و Y) يطلب تكوين ملف أوامر يقوم بتنفيذ الفعاليات التالية :

1. أستخراج معاملات Coefficients انحدار Y/X و أستخراج جدول تحليل التباين ANOVA وأستخراج معامل Durbin-Watson .

2. تكوين شكل الانتشار Scatterplot الذي يوضح العلاقة بين X و Y .
 3. استخراج مجموع قيم المتغير X ومجموع قيم المتغير Y .
- لتكوين ملف الأوامر للمثال أعلاه (بافتراض أن المتغيرين X و Y قد تم إدخالهما في شاشة Data Editor لبرنامج SPSS) سنستعمل طريقة اللصق من صناديق الحوار بموجب الخطوات التالية :
- ☐ أفتح ملف أوامر جديد Syntax File بالأمر File → New → Syntax .
 - ☐ نفذ الخطوتين 1 و 2 نفسها للمثال الأول ثم انقر زر Paste في صندوق حوار Linear Regression فيتم لصق أوامر الانحدار في ملف الأوامر المفتوح حديثاً .
 - ☐ لتكوين أوامر مخطط Scatterplot للمتغيرين X و Y من شريط القوائم في شاشة Data Editor اختر Simple → Scatter → Graphs ثم انقر زر Define فيفتح صندوق حوار Simple Scatterplot .
 - ☐ في صندوق حوار Simple Scatterplot انقر المتغير Y وأدخله في قائمة : Y Axis ثم انقر المتغير X وأدخله في قائمة : X Axis .
 - ☐ انقر زر Paste فيتم إضافة أمر Scatterplot الى ملف الأوامر .
 - ☐ لجمع قيم كل من المتغيرين X و Y من شريط القوائم اختر Frequencies → Descriptive Statistics → Analyze فيفتح صندوق حوار Frequencies .
 - ☐ في صندوق حوار Frequencies انقل المتغيرين X و Y الى قائمة Variable(s) ثم أبلغ تأشير Display Frequency Tables. ثم انقر زر Statistics فيفتح صندوق حوار Statistics .
 - ☐ في صندوق حوار Statistics أشر الخيار Sum.
 - ☐ انقر زر Paste في صندوق حوار Frequencies فيتم إضافة أوامر Frequencies الى ملف الأوامر .
 - ☐ في نافذة Syntax Editor ستحصل على ملف أوامر بأسم افتراضي هو Syntax1 يتضمن الأوامر الثلاثة المطلوبة ، لخرن هذا الملف من شريط القوائم اختر File → Save As ثم أعط أسم الملف (مثلاً three jobs) بالاستطالة SPS ، حيث يظهر هذا الملف كما يلي :



```

three jobs.SPS - SPSS Syntax Editor
File Edit View Analyze Graphs Utilities Run Window Help
[Icons]

REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT y
  /METHOD=ENTER x
  /RESIDUALS DURBIN .
GRAPH
  /SCATTERPLOT(BIVAR)=x WITH y
  /MISSING=LISTWISE .
FREQUENCIES
  VARIABLES=x y /FORMAT=NOTABLE
  /STATISTICS=SUM
  /ORDER= ANALYSIS .
  
```

□ عند اختيار All → Run من شريط القوائم في النافذة أعلاه يتم تنفيذ الأوامر الثلاثة ونحصل على النتائج التالية :

Regression

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.900 ^a	.810	.787	8.82	1.885

a. Predictors: (Constant), X

b. Dependent Variable: Y

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2661.050	1	2661.050	34.174	.000 ^a
	Residual	622.950	8	77.869		
	Total	3284.000	9			

a. Predictors: (Constant), X

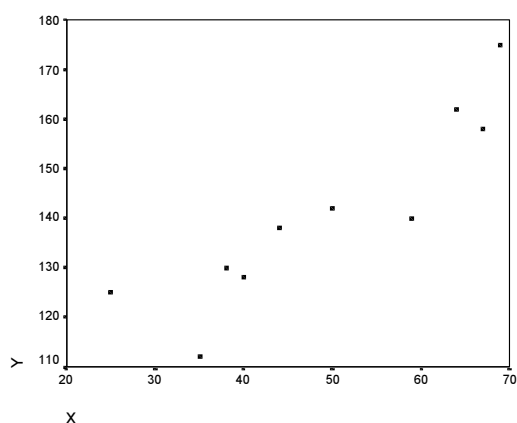
b. Dependent Variable: Y

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	85.044	9.970		8.530	.000
	X	1.140	.195	.900	5.846	.000

a. Dependent Variable: Y

Graph



Frequencies

Statistics

		X	Y
N	Valid	10	10
	Missing	0	0
Sum		491	1410

□ يمكن تنفيذ جزء من الأوامر فمثلاً لتنفيذ Regression نقوم بتظليل الأوامر الخاصة به ثم اختيار Run → Selection علماً أنه لا يمكن تنفيذ أمر فرعي دون غيره بل يتم تنفيذ كافة الأوامر الفرعية العائدة للأمر الرئيسي .

□ إذا كان موقع المؤشر على الأمر الرئيسي Graph فعند اختيار الأمر Run → Current من شريط القوائم أو نقر الأيقونة  في شريط الأدوات فيتم عرض المخطط البياني Scatterplot فقط. أما عند اختيار الأمر Run To End فيتم عرض المخطط البياني وناتج الأمر Frequency للمجموع Sum .

ملاحظة :

إذا رغبتنا بتطبيق الأوامر الثلاثة المذكورة آنفاً على المتغيرين X1 و Y1 بدلاً من المتغيرين X و Y فأن ذلك يتم حسب الخطوات التالية :

- إدخال المتغيرين X1 و Y1 في نافذة Data Editor .
- فتح ملف الأوامر الذي يحتوي الفعاليات الثلاث والمخزون مسبقاً بأسم three jobs .
- تنقيح الملف بإبدال كل X بـ X1 وكل Y بـ Y1 .
- تمشية البرنامج Run .

الفصل السادس عشر

تحليل الاستجابات المتعددة

Multiple Response Analysis

في حالة إجابة المستجيب عن سؤال يمثل رأيه مثلاً في برامج التلفزيون عند تحديد الفئات التالية (جيد ، متوسط ، رديء) فإن بإمكانه اختيار فئة واحدة فقط من الفئات الثلاث المذكورة . أما إذا سئل المستجيب عن الصحف الأثيرة لديه فإن هناك احتمال أن يختار أكثر من صحيفة واحدة وفي هذه الحالة فإن إيجاد التكرارات Frequencies يتم عن طريق تحليل الاستجابات المتعددة وهناك أسلوبين في التحليل :

1. أسلوب التكرارات متعددة الاستجابة Multiple response Frequencies : ومنه نحصل على جدول تكراري بسيط .

2. أسلوب جداول التقاطع متعددة الاستجابة Multiple Response Crosstabs : ومنه نحصل على جداول تكرارية ببعدين وثلاثة أبعاد .

(16 - 1) أسلوب التكرارات متعددة الاستجابة Multiple response Frequencies

مثال 1

في استبيان لمعرفة تفضيل المواطن لواحدة أو أكثر من الصحف المحلية التالية (بابل ، الجمهورية ، العراق ، القادسية ، الثورة) فلغرض معرفة التكرارات لكل صحيفة (عدد الأشخاص الذين يفضلون كل صحيفة من الصحف) فإنه يمكن الاختيار بين أسلوبين :

1. أسلوب المتغيرات المتعددة ذات الفئتين **Multiple dichotomy method** : بموجب هذه الطريقة فأنا نخصص متغير لكل صحيفة من الصحف الخمسة يأخذ قيمتين (فئتين) فقط (مثلاً صفر وواحد) فإذا كان المستجيب يفضل الصحيفة المعنية فأنا نعطي القيمة واحد لمتغير الصحيفة عدا ذلك يأخذ المتغير القيمة صفر . المثال المبسط التالي يبين طريقة ترتيب البيانات في شاشة Data Editor .

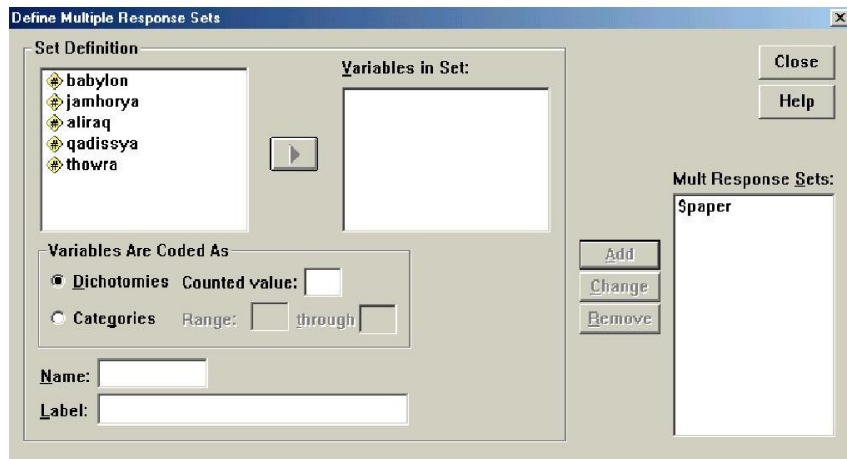
Babylon jamhorya aliraq qadissya thowra

1	0	1	1	0
1	1	0	0	0
0	0	1	0	0
1	0	0	1	1
0	1	1	0	0
1	0	0	1	1
1	0	1	0	0

لإيجاد التكرارات المقابلة لكل صحيفة نتبع الخطوات التالية :

من شريط القوائم لشاشة Data Editor اختر Define → Multiple Response → Analyze sets

فيظهر صندوق حوار Define Multiple Response Sets الذي نرتبه كما يلي :



وقد قمنا بالخطوات التالية لترتيب صندوق الحوار وكما يظهر أعلاه :

- اختيار المتغيرات الخمسة من خانة Set definition ونقلها الى خانة Variables in set في اليمين .

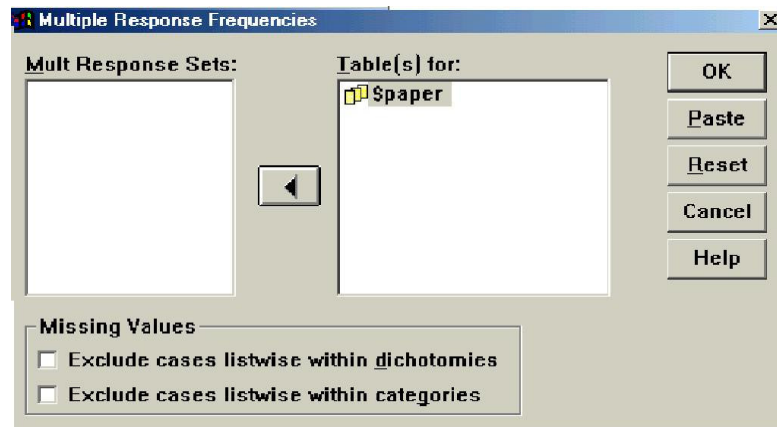
- كتابة فئة العد (واحد) بعد اختيار Dichotomies Counted value .

- إعطاء أسم لمجموعة الاستجابة المتعددة يتكون من سبعة رموز كحد أعلى في خانة Name وليكن paper (ويمكن إعطاء عنوان Label للمجموعة اختياريا) .

- نقر زر Add لإضافة المجموعة الى خانة Mult response sets

◀ عند نقر زر Close في صندوق الحوار يتم خزن مجموعة الاستجابة المتعددة بأسم paper للاستفادة منها فيما بعد .

◀ من القوائم اختر Frequencies → Multiple Response → Analyze فيظهر صندوق حوار Multiple Response Frequencies الذي نرتبه كالتالي :



حيث قمنا بنقل مجموعة الاستجابة \$paper الى خانة Tables for .

◀ عند نقر زر OK في صندوق الحوار أعلاه نحصل على الجدول التكراري البسيط التالي :

Multiple Response

Group \$PEPER

(Value tabulated = 1)

Dichotomy label Name	Pct of Count	Pct of Responses	Pct of Cases
BABYLON	5	31.3	71.4
JAMHORYA	2	12.5	28.6
ALIRAQ	4	25.0	57.1
QADISSYA	3	18.8	42.9
THOWRA	2	12.5	28.6
Total responses	16	100.0	228.6

0 missing cases; 7 valid cases

ملاحظات :

1. أن المتغيرات التي تظهر في شاشة Data Editor وهي Babylon و Jamhorya ... الخ تعرف بالمتغيرات الأولية Elementary Variables أما المجموعة التي تم تكوينها بأسم \$paper فتعرف بأسم مجموعة الاستجابات المتعددة Multiple response Set وهي تتضمن عدة متغيرات أولية.

2. في حالة وجود عناوين للمتغيرات المستخدمة في التحليل فأن عناوين هذه المتغيرات Labels تظهر في الجدول التكراري بدلاً من أسماء المتغيرات وفي هذا المثال فأن أسماء المتغيرات ظهرت في الجدول لأننا لم نعط عناوين للمتغيرات .

2. أسلوب الفئات المتعددة Multiple Category Method : لتطبيق هذا الأسلوب فانه يتوجب أولاً معرفة أكبر عدد من الإجابات على السؤال (أي أكبر عدد من الصحف المفضلة لدى المستجيب) لنفترض أن هذا العدد هو ثلاثة صحف كحد اعلى وفي المقابل نكون نفس العدد من المتغيرات لنسمها X1 و X2 و X3 حيث أن قيم كل منها الكودات التالية :

1 = Babylon , 2 = Jamhorya , 3 = Aliraq , 4 = Qadissya , 5 = Thowra
فمثلاً إذا كان المستجيب يفضل ثلاثة صحف هي بابل والعراق والقادسية فان المتغير X1 سيأخذ الكود 1 والمتغير X2 له الكود 3 والمتغير X3 له الكود 4 أما إذا كان المستجيب يفضل صحيفتي بابل والجمهورية فقط فأننا سنخصص الكود 1 للمتغير X1 والكود 2 للمتغير X2 أما X3 فستكون له قيمة مفقودة Missing Value . فبالنسبة للمثال السابق فأننا سترتب المتغيرات X1 و X2 و X3 في شاشة Data Editor

كما يلي :

X1	X2	X3
1	3	4
1	2.	
3.	.	
1	4	5
2	3.	
1	4	5
1	3.	

لإيجاد التكرارات المقابلة لكل صحيفة (أي لكل كود) نتبع الخطوات التالية وهي مشابهة تقريباً

لأسلوب المتغيرات المتعددة ذات الفئتين :

من شريط القوائم لشاشة Data Editor اختر Define sets → Multiple Response → Analyze

- فيظهر صندوق حوار Define Multiple Response Sets الوارد أنفاً الذي نرتبه كما يلي :
- اختيار المتغيرات الثلاثة X1 و X2 و X3 من خانة Set definition ونقلها الى خانة Variables in set في اليمين .
 - اختيار Categories في خانة Variables Are Coded As وتعيين فئات الصحف من 1 الى 5 (Range 1 through 5) .
 - إعطاء أسم لمجموعة الاستجابة المتعددة يتكون من سبعة رموز كحد أعلى في خانة Name وليكن paper (ويمكن إعطاء عنوان Label للمجموعة اختيارياً) .
 - نقر زر Add لإضافة المجموعة الى خانة Mult response sets
 - عند نقر زر Close في صندوق الحوار يتم خزن مجموعة الاستجابة المتعددة بأسم paper للاستفادة منها فيما بعد .
 - من القوائم اختر Frequencies → Multiple Response → Analyze فيظهر صندوق حوار Multiple Response Frequencies حيث قمنا بنقل مجموعة الاستجابة \$paper الى خانة Tables for .
 - عند نقر زر OK في صندوق الحوار المذكور نحصل على الجدول التكراري البسيط التالي :

Group \$PAPER

Category label	Count	Pct of Responses	Pct of Cases
Code			
1	5	31.3	71.4
2	2	12.5	28.6
3	4	25.0	57.1
4	3	18.8	42.9
5	2	12.5	28.6
	-----	-----	-----
Total responses	16	100.0	228.6

0 missing cases; 7 valid cases

أن هذه النتيجة مشابهة تماماً لما توصلنا إليه بأسلوب المتغيرات المتعددة ذات الفئتين .

ملاحظة : في حالة إعطاء قيم المتغيرات الثلاثة X1 و X2 و X3 عناوين للقيمة Value Label فأن عنوان القيمة يظهر بدلاً من القيمة نفسها في الجدول التكراري .

(16 - 2) أسلوب جداول التقاطع متعددة الاستجابة Multiple Response Crosstabs

أن هذا الأسلوب يتيح تكوين جداول تقاطع للمتغيرات الأولية Elementary Variables أو للمجاميع المعرفة Multiple response Sets أو لكليهما معاً .

مثال 2 :

لنفس بيانات المثال الأول لنفترض أن هناك سؤالاً يتعلق برغبة المستجيب في أن تكون الصحيفة بالألوان أو اسود و ابيض وفي هذه الحالة سوف نكون المتغير Colored الذي يتضمن قيمتين (القيمة 1 إذا كان المستجيب يفضل الألوان والقيمة 0 إذا كان لا يفضل الألوان) بالإضافة الى المتغيرات التي تمثل تفضيل المستجيب لكل صحيفة من الصحف .حيث تظهر المتغيرات في شاشة Data Editor كما يلي :

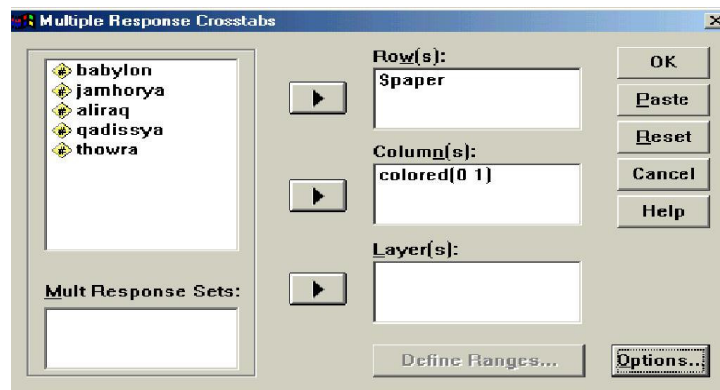
Babylon	jamhorya	aliraq	qadissya	thowra	colored
1	0	1	1	0	1
1	1	0	0	0	0
0	0	1	0	0	0
1	0	0	1	1	1
0	1	1	0	0	0
1	0	0	1	1	1
1	0	1	0	0	1

نرغب في تكوين جدول تكراري مزدوج (جدول تقاطع) بين تفضيل الصحف وبين السؤال الذي يتضمن تفضيل الألوان .

لتكوين هذا الجدول نحتاج للمتغير colored الذي يمثل تفضيل الألوان وهو متغير أولي Elementary Variable كما نحتاج الى مجموعة الاستجابات المتعددة \$paper التي تمثل تفضيل الصحف ، يجب أولاً تكوين هذه المجموعة بنفس الطريقة المذكورة في المثال الأول بأسلوب المتغيرات المتعددة ذات الفئتين Multiple dichotomy method (علماً أنه يمكن اعتماد أسلوب الفئات المتعددة Multiple Category Method).

لتكوين جدول التقاطع بحيث تكون فئات تفضيل الصحف في الصفوف وفئات تفضيل الألوان في الأعمدة نتبع الخطوات التالية :

من شريط القوائم اختر Crosstabs → Multiple Response → Analyze فيظهر صندوق حوار Multiple Response Crosstabs الذي نرتبه كما يلي :



علماً أنه يتوجب تحديد الحدين الأعلى والأدنى لفئات المتغير الأولي colored وهي (0,1) ويتم ذلك بنقر الزر Define Ranges في صندوق الحوار أعلاه .

عند نقر زر OK نحصل على الجدول التالي .

```

* * * C R O S S T A B U L A T I O N * * *
$PAPER (tabulating 1)
by COLORED

```

		COLORED		
\$PAPER	Count			Row Total
		0	1	
BABYLON	1	1	4	5 71.4
JAMHORYA	2	2	0	2 28.6
AL IRAQ	2	2	2	4 57.1
QADISSYA	0	0	3	3 42.9
THOWRA	0	0	2	2 28.6
Column Total		3 42.9	4 57.1	7 100.0

Percents and totals based on respondents

ملاحظات :

1. يمكن الحصول على تكرارات كل صف أو عمود في جدول التقاطع إما كنسبة مئوية من الاستجابات responses أو كنسبة مئوية من المستجيبين respondents (أو الحالات Cases) ويتم ذلك بنقر زر Options في صندوق حوار Multiple Response Crosstabs ثم تأشير أحد الحالتين التاليتين في خانة Percentages Based on :

- Cases : لغرض الحصول على تكرار الصف أو العمود كنسبة مئوية من الحالات Cases أو المستجيبين Respondents (عددهم في هذا المثال هو 7) .
- Responses : لغرض الحصول على تكرار الصف أو العمود كنسبة مئوية من الاستجابات Responses (عدد الاستجابات في هذا المثال هو 16) .

2. يمكن الحصول على النسبة المئوية لكل خلية من خلايا جدول التقاطع بنقر زر Options في صندوق حوار Multiple Response Crosstabs ثم تأشير أحد الخيارات التالية في خانة Cell Percentages :

- Row : لعرض تكرار الخلية كنسبة مئوية من مجموع تكرارات الصف .
- Column : لعرض تكرار الخلية كنسبة مئوية من مجموع تكرارات العمود .
- Total : لعرض تكرار الخلية كنسبة مئوية من مجموع التكرارات الكلية .

مثال 3

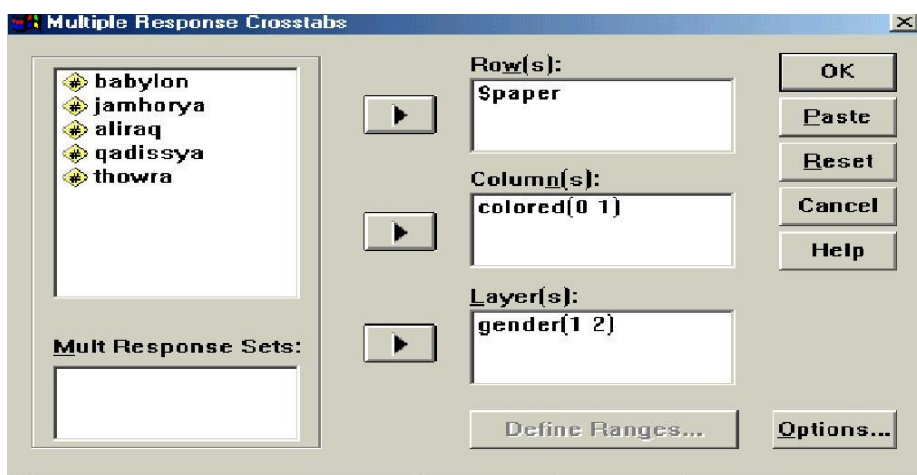
لنفس المثال السابق يطلب تكوين جدول تقاطع لكل من المستجيبين الذكور والإناث بصورة منفصلة. في هذه الحالة يتوجب إضافة متغير الجنس gender الى شاشة Data Editor للمثال السابق حيث أنه متغير ذو فئتين (1 يمثل الذكور و 2 يمثل الإناث وقد أعطينا عناوين القيم Value Label بحيث أن 1 = Male و 2 = Female) وتظهر الشاشة كما يلي :

Babylon	jamhorya	aliraq	qadissya	thowra	colored	gender
1	0	1	1	0	1	2
1	1	0	0	0	0	1
0	0	1	0	0	0	2
1	0	0	1	1	1	1
0	1	1	0	0	0	1
1	0	0	1	1	1	1
1	0	1	0	0	1	2

لتكوين هذا الجدول نحتاج الى مجموعة الاستجابات المتعددة \$paper التي تمثل تفضيل الصحف ، يجب أولاً تكوين هذه المجموعة بنفس الطريقة المذكورة في المثال الأول بأسلوب المتغيرات المتعددة ذات الفئتين Multiple dichotomy method (علماً أنه يمكن اعتماد أسلوب الفئات المتعددة Multiple Category Method).

لتكوين جدول التقاطع بحيث تكون فئات تفضيل الصحف في الصفوف وفئات تفضيل الألوان في الأعمدة وحسب فئتي الجنس نتبع الخطوات التالية :

من شريط القوائم اختر Crosstabs → Multiple Response → Analyze فيظهر صندوق حوار Multiple Response Crosstabs الذي نرتبه كما يلي :



لاحظ أننا حددنا الحدين الأعلى والأدنى لفئات المتغير Colored وكذلك فئات المتغير Gender بواسطة الزر Define Ranges بعد نقر أسم المتغير المعني . وعند نقر زر OK في صندوق الحوار أعلاه نحصل على المخرج التالي :

		COLORED		
\$PAPER	Count			Row Total
		0	1	
BABYLON		1	2	3 75.0
JAMHORYA		2	0	2 50.0
ALIRAQ		1	0	1 25.0
QADISSYA		0	2	2 50.0
THOWRA		0	2	2 50.0
	Column Total	2 50.0	2 50.0	4 100.0

\$PAPER (tabulating 1)

by COLORED

by GENDER

Category = 1 Male

Percents and totals based on respondents

لاحظ أن النسب المئوية مبنية على المستجيبين (الحالات) وأن عدد المستجيبين الذكور هو 4 (الجدول السابق) وأن عدد المستجيبين الإناث هو 3 (الجدول التالي) .

\$PAPER (tabulating 1)
 by COLORED
 by GENDER
 Category = 2 Female

		COLORED		
\$PAPER	Count.			Row Total
		0	1	
BABYLON		0	2	2 66.7
ALIRAQ		1	2	3 100.0
QADISSYA		0	1	1 33.3
Column Total		1 33.3	2 66.7	3 100.0

Percents and totals based on respondents