جامعة الجزائر 3

كلية العلوم الاقتصادية و العلوم التجارية وعلوم التسيير

قسم العلوم الاقتصادية

مطبوعة مقدمة لطلبة الماستر تخصص اقتصاد وتسيير المؤسسات بعنوان:

دروس وتطبيقات متقدمة في برنامج SPSS

إعداد:

الأستاذ: بوعراب رابح

أستاذ محاضر قسم " أ "

السنة الجامعية: 2019/2018

فهرس المحتويات

الفصل الأول: مفاهيم أساسية حول البيانات الاحصائية وطرق جمعها..... 3
1. بعض المفاهيم الاساسية حول البيانات الاحصائية:..... 3
1.1 المجتمع الاحصائي ومكوناته:..... 3
2.1 مفهوم المتغيرات وخصائصها:..... 3
أولا. تصنيف المتغيرات حسب طبيعتها:..... 4
ثانيا. تصنيف المتغيرات حسب علاقاتها مع بعضها البعض:..... 4
ثالثا. تصنيف المتغيرات حسب مستويات قياسها:..... 5
2. طرق جمع البيانات الإحصائية:..... 7
1.2 طرق اختيار العينة:..... 7
أولا. طرق اختيار العينة غير العشوائية: 7
ثانيا. طرق اختيار العينات العشوائية :..... 8
2.2 طرق جمع البيانات:..... 10
أولا. المقابلة الشخصية:..... 10
ثانيا. الملاحظة المباشرة:..... 10
ثالثا. الإستبيان :..... 10
2-3. الترميز (عملية الانتقال من الاستبيان إلى برنامج SPSS): 12
الفصل الثاني: بعض المبادئ الاساسية لاستعمال برنامج SPSS..... 15
1. التعرف على بيئة البرنامج الاحصائي SPSS: 15
1.1 تثبيت البرنامج الاحصائي SPSS:..... 15
2.1 تشغيل برنامج SPSS:..... 17
3.1 النوافذ الأساسية المتوفرة في برنامج SPSS: 17
4.1 أنواع الملفات في برنامج SPSS: 18

5.1 صناديق الحوار في برنامج SPSS:..... 18
6.1 القوائم الرئيسية للبرنامج SPSS: 19
7.1 شريط الأدوات (الأيقونات المختصرة): 20
2. تهيئة الملفات وتعبئة البيانات في برنامج SPSS: 22
1.2 تهيئة ملفات البيانات:..... 22
الفصل الثالث: بعض تقنيات ادارة البيانات باستخدام برنامج SPSS 32
1. العليات على البيانات:..... 32
أولا. ترتيب ملف البيانات بواسطة الأمر Sort cases:..... 32
ثانيا. تجزئة ملفات البيانات بالأمر Split file: 34
ثالثا. اختيار الحالات باستخدام الأمر Sort Cases: 37
2. العمليات على المتغيرات:..... 40
الفصل الرابع: التحليل الاحصائي للبيانات باستعمال برنامج SPSS..... 46
1. بعض المبادئ الأساسية للتحليل الوصفي للبيانات: 46
2. التحليل الاحصائي الوصفي باستخدام البرنامج SPSS:..... 48
2-1. الأمر Frequencies:..... 48
2-2. الأمر descriptives:..... 55
2-3. الأمرExplore:..... 57
أولا. استكشاف البيانات عن طريق المؤشرات الاحصائية والمخططات البيانية: 57
ثانيا. استكشاف البيانات عن طريق اختبار التوزيع الطبيعي وتجانس التباين: 61
2-4. الجداول التقاطعية Crosstabs:..... 69
أولا. الجدول التقاطعي لمتغيرين:..... 69
ثانيا. الجدول التقاطعي لأكثر من متغيرين: 76
3. مقارنة المتوسطات:..... 79
3-1. تحليل المتوسطات بواسطة الأمر Means:..... 79
أولا: تحليل المتوسط حسب متغير مستقل واحد: 79

ثانيا: تحليل المتوسط بأكثر من متغير مستقل: 82
3-2. اختبار t لعينة واحدة One sample t test:..... 85
3-3. اختبار t لعينتين مستقلتين Independent simple t test: 88
3-4. اختبار تساوي متوسطي عينة مزدوجة Paired-Simples T Test: 91
3-5. اختبار تساوي عدة متوسطات: 94
أولا: تحليل التباين بمعيار واحد One-way ANOVA: 94
ثانيا : تحليل التباين بمعيارين أو أكثر:. 108
4. تحليل الارتباط والانحدار الخطي:..... 112
4-1. الارتباط:..... 112
أولا. الارتباط الخطي البسيط:..... 112
ثانيا. الارتباط الجزئي:..... 117
4-2. تحليل الانحدار:..... 120
4-1. تحليل الانحدار الخطي البسيط:..... 120
أولا: فرضيات النموذج الخطي البسيط:..... 120
ثانيا. طريقة تقدير نموذج الانحدار:..... 121
4-2. تحليل الانحدار الخطي المتعدد: 129
خاتمة:..... 133
قائمة المراجع:..... 134

مقدمة:

يعتبر البرنامج الجاهز المعروف بمختصر SPSS وهي اختصار لـ Statistical Package for Social Sciences من بين البرامج الأكثر استعمالا من طرف الطلبة الباحثين في مختلف الميادين وعلى وجه الخصوص ميدان العلوم الاجتماعية والإنسانية. ففي الوقت الذي اصبح فيه التحليل الاحصائي للبيانات أمرا ضروريا وأداة اساسية لتحليل مختلف الظواهر، فإن الاستعانة بالبرنامج كذلك اصبح ضروريا نظرا لأن حجم البيانات التي يجب على الباحث أن يتعامل معها كبير جدًا، ومن الصعب اجراء التحليلات الاحصائية اللازمة بالطرق البسيطة (كإعداد الجداول الاحصائية ومختلف الحسابات بالاستعانة بالآلة الحاسبة)، بالإضافة إلى تطور الاساليب الكمية الممكن استعمالها والتي تتميز بنوع من التعقيد. ومن هذا المنطلق برزت الحاجة لاستخدام الحاسب الآلي لإعداد الدراسات الاحصائية لاختصار الوقت والجهد.

وقد عرفت الآونة الاخيرة تطوير عدة برمجيات جاهزة من شأنها التعامل مع البيانات ذات الحجم الكبير والمبرمجة لإجراء التحليلات الاحصائية المتقدمة، لكن تعتبر بمرمجية الـ SPSS من بين أهم البرمجيات المتطورة بشكل خاص للتعامل مع البيانات الإسمية، بطرق بسيطة وسهلة في متناول الطلبة والباحثين خاصة في مجال العلوم الانسانية والتي تتميز بهذا النوع من البيانات.

وارتأينا ان نضع بين ايدي طلبتنا في العلوم الاقتصادية هذه المطبوعة، التي تتناول شرح بعض طرق التعامل مع البرنامج في اعداد الدراسات الاحصائية والقياسية في ميدان العلوم الاقتصادية، من خلال استخدام بعض الأمثلة البسيطة التي تساعد الطالب في التدريب على البرنامج والتعود على استخدام الاوامر الأساسية للبرنامج.

تمهيد:

يبحث علم الاحصاء في طرق جمع البيانات وتبويبها وتحليلها من خلال مجموعة من الأساليب الرياضية والبيانية، وتهدف هذه العملية إلى استخلاص معلومات حول صفات متغير أو مجموعة من المتغيرات من البيانات المتوفرة كم عينة من هذه المتغيرات، من أجل الوصول إلى قرارات مناسبة يمكن تعميمها على المجتمع محل الدراسة والذي سحبت منه العينة.

ومن المعروف ان جمع المعلومات حول جميع أفراد المجتمع ليس بالأمر الهين ويستحيل تحقيقه في غالب الأحيان، نظرا للصعوبات التي يواجهها الباحث في هذه العملية بالاضافة إلى ما يتطلبه من جهد ووقت وتكلفة اضافية لتحقيق ذلك، وعليه يصبح سحب عينة من هذا المجتمع يسهل عملية جمع المعلومات من حيث الجهد والوقت والتكلفة في آن واحد.

وتعتمد البحوث التي تستعي استخدام الاساليب الاحصائية للحصول على نتائج تساعد على اتخاذ القرارات، على عدة خطوات، نذكرها فيما يلي:

الخطوة الأولى: تحديد مشكلة البحث المراد دراستها، في هذه المرحلة يكون الباحث بصدد تحديد المتغيرات المراد دراستها وشكل العلاقات التي تربط بين بعضها البعض؛

الخطوة الثانية: تحديد الاداة التي تستخدم لجمع البيانات، مثلا الاستعانة باستمارة استبيان يتم اعداده بطرق متخصصة لتسهيل عملية جمع معلومات صالحة للإجابة على الاشكالية المطروحة؛

الخطوة الثالثة: تحديد عينة الدراسة، وهي تحديد الأفراد أو الوحدات التي توجه اليها الاستبيان لتحصيل المعلومات اللازمة؛

الخطوة الرابعة: ترميز البيانات وتحويلها إلى أرقام أو حروف من أجل تسهيل عملية تحصيل المعلومات وإدخالها إلى الحاسوب الأمر الذي يسمح بربح الوقت وتسهيل التعامل معها، ثم ادخالها إلى الحاسوب وجعلها جاهزة للتحليل الإحصائي ومن ثم اجراء التحليل الاحصائي المناسب لتحقيق اهداف البحث المنشودة.

وسنتناول من خلال هذا الجزء بعض الركائز الاساسية لعلم الاحصاء حول المتغيرات وطرق جمع البيانات وطرق تحديد العينات، والتي من شأنها جعل البيانات صالحة للاستعمال والوصول إلى نتائج دقيقة وذات مصداقية احصائية. فاستعمال بيانات غير دقيقة حول أفراد لا تمثل المجتمع الاحصائي يؤدي إلى نتائج مضللة وغير صحيحة لا يمكن اعتماد نتائجها لتحليل المجتمع محل الدراسة.

الفصل الأول: مفاهيم أساسية حول البيانات الاحصائية وطرق جمعها

1. بعض المفاهيم الاساسية حول البيانات الاحصائية:

1.1 المجتمع الاحصائي ومكوناته:

يعتمد التحليل الاحصائي على توفر البيانات الاحصائية حول الظاهرة المدروسة، التي بدورها ترتبط بوحدات المجتمع الذي يتميز بخصائص الظاهرة، والتي تمثل في شكل متغير أو عدة متغيرات، وعليه سنقوم بتحديد مجموعة من المفاهيم الخاصة بالمجتمع الاحصائي.

أولا: الوحدة الاحصائية: هي الوحدة الاساسية لتكوين المجتمع الإحصائي، والتي غالبا ما تكون عبارة عن فرد من أفراد المجتمع الاحصائي (فرد، عائلة، مؤسسة....إلخ)

ثانيا: المجتمع الاحصائي: هو مجموعة المشاهدات و القياسات الخاصة بمجموعة من الوحدات الاحصائية والتي تخص ظاهرة من الظواهر القابلة للقياس، مثلا مجتمع من الأفراد، مجتمع من الطلبة، مجتمع من الأسر، مجتمع من المؤسسات، ...وغيرها.

ثالثا: الظاهرة الإحصائية: هي الخاصية المدروسة، أو المتغير المدروس في المجتمع الإحصائي، الاستهلاك، الدخل، البطالة، النمو الاقتصادي....وغيرها.

رابعا: العينة الإحصائية: هي جزء من المجتمع الإحصائي، وفيها يشترط أن تكون ممثلة للمجتمع أحسن تمثيل وليس أي جزء منه، بحيث يعتمد على العينة الاحصائية عوض المجتمع الاحصائي لتسهيل عملية تحصيل المعلومات الخاصة بمتغيرات الدراسة، وحجم العينة الاحصائية يتحدد حسب اهمية الدراسة والامكانيات المادية والبشرية التي يتوفر عليها الباحث للقيام بالدراسة. ويمكن التمييز بين عدة انواع من العينات الاحصائية التي سنفصل فيها لاحقا.

2.1 مفهوم المتغيرات وخصائصها:

تعرف المتغيرات بانها الخصائص أو الصفات التي يمكن قياسها، وتختلف درجاتها بين الأفراد أو المجموعات أو لفرد معين عبر الزمن. ويمكن التمييز بين نوعين من المتغيرات:

- المتغيرات الإحصائية: وهي عبارة عن مجموعة القيم التي يمكن أن تأخذها ظاهرة معينة، ويمكن قياسها بوحدات معينة؛
- المتغيرات العشوائية: وهي عبارة عن ظاهرة نوعية أو كمية لا يمكن التنبؤ بها بشكل مسبق وترتبط بقيم احتمالية.

ويمكن التمييز بين عدة اصناف من المتغيرات وفق عدة معايير نلخصها فيما يلي:

أولا. تصنيف المتغيرات حسب طبيعتها:

تصنف المتغيرات حسب طبيعتها إلى نوعين أساسيين:

1. المتغيرات المستمرة (المتصلة): وهي المتغيرات التي تقع درجات قياسها على تدريج كمي متصل، بمعنى انه يمكن الحصول على أي قيمة ضمن مدى هذا المتغير، ويكون لهذه المتغيرات توزيعا احتماليا مستمرا كالتوزيع الطبيعي مثلا.

2. المتغيرات المنقطعة (المنفصلة): وهي المتغيرات التي تأخذ قيما من مجموعة منتهية من القيم، وتنقسم بدورها إلى قسمين:

أ. المتغيرات النوعية: وهي عبارة عن الخصائص أو الصفات التي لا يمكن قياسها بشكل كمي أو رقمي، مثلا متغير الجنس الذي يحتمل بديلين اثنين: ذكر أو انثى، وهاتين الفئتين لا يمكن قياسها بمقدار كمي، وبالتالي لا يمكن اجراء اي عمليات حسابية على بيانات من هذا النوع، لكنها تساعد في اجراء المقارنات وتصنيف أفراد العينة أو المجتمع المدروس.

ب. المتغيرات الكمية المنفصلة: وهي المتغيرات التي تأخذ قيما عددية من مجموعة منتهية قابلة للقياس، مثلا عدد الطلبة في تخصص معين، عدد الأفراد في الأسرة... وغيرها، وفي هذا النوع من المتغيرات يمكن اجراء مختلف العمليات الحسابية على بياناتها.

ثانيا. تصنيف المتغيرات حسب علاقاتها مع بعضها البعض:

يمكن تصنيف المتغيرات وفق هذا المعيار إلى:

أ. المتغيرات المتسقلة: وهي المتغيرات التي يفترض الباحث انها تؤثر على متغيرات أخرى، مثلا اذا اردنا فحص أثر استخدام البرنامج SPSS في تحليل البيانات الخاصة بدراسة معينة على الوقت المستغرق في اعداد الدراسة، أو أثر استخدام اجهزة العرض (Data Show) في التدريس على تحصيل الطلبة في مقياس معين بالمقارنة مع الطريقة التقليدية (الصبورة)، تكون طريقة التدريس هي المتغير المستقل.

ب. المتغيرات التابعة: وهي المتغيرات التي يفترض الباحث انها تتأثر بمتغيرات أخرى، مثلا المعدل المحصل في مقياس معين يكون متغير تابع يتأثر بالطريقة المتبعة في التدريس.

ج. المتغيرات المُعَدِّلة: وهي متغيرات يفترض الباحث أنها تؤثر على قوة واتجاه العلاقة بين المتغيرات التابعة والمستقلة، فمثلا إذا اردنا فحص اثر استخدام اجهزة العرض (Data Show) على تحصيل الطلبة في مقياس معين، وإذا تبين أن هذه الطريقة أفضل للطلبة الذين لديهم مهارات قراءة عالية من الطلبة الذين لديهم مهارات قراءة متدنية، ففي هذه الحالة يكون متغير مهارات القراءة متغيرا معدلا للعلاقة بين طريقة التدريس والتحصيل في مقياس معين.

د. المتغيرات الوسيطة: وهي متغيرات تفسر العلاقة بين متغيرين، فمثلا الرغبة في التَعلُم (الدافعية للتعلُم)، ربما يكون متغيرا وسيطيا بين طريقة التدريس والتحصيل في مقياس معين لأنه يفسر العلاقة بينهما، ففي حالة عدم توفر الرغبة في التَعلُم فإنه لا توجد علاقة بين طريقة التدريس والتحصيل في المقياس المعني.

هـ. المتغيرات الضابطة: وهي المتغيرات التي تؤثر على العلاقة بين المتغيرين التابع والمستقل، ولكن الباحث يقوم بضبط أثر هذا المتغير لأن اهتمامه يكون فقط على أثر المتغير المستقل على المتغير التابع، فمثلا مستوى التحصيل العلمي في مقياس معين ربما يكون متغيرا ضابطا لأنه يؤثر على العلاقة بين طريقة التدريس والتحصيل في المقياس، ويتم ضبط أثر هذه المتغيرات اما من خلال الاساليب الاحصائية أو عند اختيار أفراد العينة.

ثالثا. تصنيف المتغيرات حسب مستويات قياسها:

يمكن تصنيف المتغيرات حسب طبيعة البيانات المتوفرة من عينة الدراسة فعليا، فمتغير الجنس مثلا لا يشبه من حيث طبيعته البيانات الخاصة بمتغير العمر (السن) والذي بدوره لا يشبه متغير درجة الاعتقاد بموضوع معين، وتصنيف المتغيرات حسب هذا المعيار يسمى مستوى القياس، ويمكن التمييز بين أربعة أنواع أساسية نذكرها فيما يلي:

أ. القياس الاسمي (Nominal Measurment):

وهو قياس خاص بالمتغيرات النوعية أو الإسمية، والتي تحتوي على عدد معين من الفئات دون أي معنى كمي لهذه الفئات، اذ يمكن تصنيف أفراد المجتمع إلى هذه الفئات دون أفضلية لإيحادها على الاخرى. ومن امثلة ذلك تصنيف الأفراد إلى ذكور واناث، تصنيفهم حسب المنطقة الجغرافية شمال، وسط وجنوب. ويتم ترميز هذه الفئات بإعطاء ارقام لكل منها مثلا الرمز (1) للذكور و (2) للإناث، لكن هذه الأرقام ليس لها أي معنى كمي، بحيث لا يمكن اجراء اي عمليات حسابية عليها بحيث أن المتوسط الحسابي مثلا لمتغير الجنس ليس له اي معنى من الناحية الاحصائية.

ب. القياس الترتيبي (Ordinal Measurment):

هو قياس لمتغير على شكل فئات، يمكن ترتيبها تصاعديا أو تنازليا، ولكن لا يمكن تحديد الفروق بين قيم الأفراد المختلفة بدقة، مثلا بدرجة كبيرة، متوسطة وضعيفة، هي ثلاث اجابات محتملة لوصف درجة تأثير الاشهار على قرار الشراء لدى الأفراد من منتوج معين، فنقول أن A أكبر من B، و B أكبر من C لكن لا يمكن تحديد كم يكبر A عن B، أو B عن C.

ويمكن لبعض المتغيرات الكمية أن تأخذ القياس الترتيبي إذا ما قسمت بياناته إلى فئات محددة بمجالات (أبواب)، مثلا متغير العمر (السن)، بحيث يمكن تقسيم هذا المتغير إلى فئات يحددها الباحث حسب طبيعة الدراسة، مثلا:

-العمر: 1. أقل من 30 سنة؛ 2. من 30 إلى 39 سنة؛ 3. من 40 إلى 49 سنة؛ 4. من 50 إلى 60 سنة؛ 5. أكثر من 60 سنة.

ملاحظة: عموما يتم التعامل مع المتغيرات الإسمية وفق معيار القياس بمستوى القياس الاسمي والترتيبي على حد سواء وهذا حسب طبيعة الدراسة والبيانات المتوفرة.

ج. القياس الفئوي (Interval Measurment):

هو مقياس للمتغيرات الكمية، وفيها يمكن معرفة ترتيب الأفراد حسب البيانات المتوفرة وكذلك يمكن تحديد درجة الاختلاف بين الأفراد، فمثلا متغير المعدل المحصل في مادة معينة، إذا عرفنا أن أحمد تحصل على 15 نقطة و علي تحصل على 12 نقطة و سعيد على 10 نقاط، ففي هذه الحالة يمكننا معرفة الترتيب الخاص بهؤلاء الأفراد ودرجة الاختلاف بينهم. وكذلك وفق هذا القياس يكون للقيمة 0 معنى مختلف عن المقاييس الأخرى، فحصول الفرد سليم على العلامة 0 في المادة فهذا لا يعني أنه لا يعرف شيء في المادة المعنية، بينما تسجيل مقياس درجة الحرارة الدرجة 0 لا تدل على أنه لا توجد درجة الحرارة. ونشير هنا إلى ان معظم الدراسات النفسية والتربوية والاجتماعية هي متغيرات تقاس كميا على مستوى القياس الفئوي.

د. القياس النسبي (Ratio Measurment):

وهو قياس خاص بالمتغيرات الكمية التي يكون فيها الصفر له معنى حقيقي أي انعدام الخاصية، فمثلا متغير المسافة المقطوعة أو الزمن المستغرق في قطع مسافة معينة. فأخذ هذه المتغيرات القيمة (0) يعني فعلا انعدامها أى عدم قطع المسافة أصلا.

ملاحظة: يتم التعامل مع النوعين (ج) و (د)، احصائيا بنفس الطريقة ويطلق عليها اسم المتغيرات الكمية. والبرنامج محل الدراسة يأخذ بعين الاعتبار 3 مستويات القياس: الإسمي، الترتيبي والكمي كما سنتناوله لاحقا.

2. طرق جمع البيانات الإحصائية:

يتم جمع البيانات الإحصائية عموما بإحدى الطرق التالية:

أ. **طريقة المسح الشامل:** فيها تجمع البيانات من جميع وحدات (أفراد) المجتمع دون استبعاد أي وحدة، فمثلا إذا أردنا التعرف على مستوى طلاب كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير في مادة الإحصاء نقوم برصد علامات جميع الطلبة في مادة الإحصاء وهكذا... ، وهذه الطريقة عادة تكون طويلة ومكلفة وتحتاج إلى الكثير من الوقت ناهيك عن عدم إمكانية تطبيقها في جميع الحالات.

ب. **طريقة العينة:** وفيها يتم اختيار عينة تمثل المجتمع وتجرى عليها الدراسة وتعمم النتائج على المجتمع، وكلما كانت العينة مختارة بطريقة صحيحة وممثلة تمثيلا صادقا للمجتمع كلما كانت النتائج دقيقة وذات مصداقية.

1.2 طرق اختيار العينة:

تصنف طرق المعاينة إلى الطرق غير العشوائية والطرق العشوائية أو الاحتمالية[1].

أولا. طرق اختيار العينة غير العشوائية:

تكون العينات في هذه الطريقة انتقائية ولا تمثل المجتمع تمثيلا صحيحا، وإنما تتم وفق اختيار الباحث، ولذلك لا تكون هناك فرصة متساوية لأفراد المجتمع في الظهور في العينة (الانتقاء في العية)، وهذه العينات تستخدم بهدف الحصول على نتائج استطلاعية نظرا لأن اختيار عينات عشوائية يتطلب وقتا أو تكلفة أو جهد كبير. وفي هذه العينات لا يمكن استخدام أساليب الإحصاء التحليلي والذي يقتصر استخدامه على العينات العشوائية، ومن العينات غير العشوائية نذكر ما يلي:

أ. **العينات العرضية:** وتحدث عندما يتم جمع بيانات من المواطنين أو العمال في مصنع كبير الذين يصادفونهم حول اتجاهاتهم نحو سلع معينة أو نحو إدارة مصنع أو نظم الرقابية فيه للحصول على بعض المعلومات والمؤشرات بأقل تكلفة أو جهد ممكن.

[1] لتفاصيل أكثر، أنظر في ذلك: محمد شامل بهاء الدين فهمي: **"الاحصاء بلا معاناة المفاهيم مع التطبيقات باستخدام البرنامج SPSS"** مركز البحوث بمعهد الادارة العامة، المملكة العربية السعودية، الجزء الأول 2005، ص 95-166.

ب. **المعاينة الطبقية غير العشوائية:** وتحدث على سبيل المثال عندما يقسم مجتمع الدراسة في مصنع إلى طبقة الإداريين وطبقة العمال، أو إلى إناث وذكور، وبذلك تراعى نسبة المجموعات الفرعية في الدراسة. ولكن العينة من كل طبقة لا تؤخذ بطريقة عشوائية وإنما يقوم الباحث باختيار الذين يصادفهم.

ج. **العينة الغرضية:** والتي تستخدم عند دراسة تكاليف صناعة ما على سبيل المثال، الأمر الذي يتطلب تعاونا من المستجوب لتوفير المعلومات.

ثانيا. طرق اختيار العينات العشوائية :

تسمح طرق اختيار العينات العشوائية بالحصول على عينات ممثلة للمجتمع، ويكون احتمال سحب أي مفردة معروفا ومتساويا ويمكن حسابه ولذلك تسمى عينة احتمالية. فمثلا إذا كان حجم العينة المختارة 25 مفردة من مجتمع حجمه 500 فان احتمال سحب كل مفردة هو $\frac{25}{500} = 5\%$

وتعرف العينة العشوائية بأنها العينة التي يكون فيها احتمال اختيار جميع المفردات متساوي ومعروف ويمكن حسابه. وهناك عدة طرق لاختيار العينة نذكر من أهمها:

1- العينة العشوائية البسيطة:

تتصف العينة العشوائية البسيطة بأنها مجموعة جزئية من المجتمع الأصلي وبحجم معين لها نفس الفرصة (الاحتمال) لتختار كعينة من ذلك المجتمع، وفيها يكون لكل فرد من أفراد المجتمع نفس الفرصة للظهور في هذه العينة، وهي تستخدم في حالة المجتمع الذي يتكون من وحدات متجانسة. ويمكن الحصول على عينات عشوائية بسيطة باستعمال جداول الأعداد العشوائية.

2- العينة المنتظمة:

يرى الكثيرون أن طريقة المعاينة المنتظمة هي في جوهرها شكل من أشكال المعاينة العشوائية البسيطة. وتعرف العينة المنتظمة بأنها العينة التي تؤخذ أفرادها بناءا عللى ترتيبهم في ضمن أفراد المجتمع، بحيث يتم إضافة رقم معين بشكل منتظم من قائمة كاملة مرتبة عشوائيا لأفراد المجتمع. وتعتبر العينة المنتظمة بديلا عن العينة العشوائية البسيطة للأسباب التالية:

(أ) العينة المنتظمة اكثر سهولة في التنفيذ من العينة العشوائية البسيطة.

(ب) العينة العشوائية يستطيع شخص غير مدرب لتعينها.

مثال (1-1):

يطلب اختيار عينة حجمها n=200 من مجموعة من بطاقات التسجيل في إحدى الجامعات التي يسجل فيها N = 3000 طالبا، وهذا بهدف دراسة البطاقات التي بها أخطاء.

الحل: إن طريقة العينة المنتظمة تقتضي أن يكون طول المجال الذي سيسحب منها أول مفردة بطريقة عشوائية وهي $\frac{3000}{200}=15$. ولذلك نختار رقما عشوائيا من 1 إلى 15 وليكن 8.

نختار الرقم 8 ومن ثم نضيف 15 للرقم 8 وبذلك نسحب الرقم 23، ثم نضيف الرقم 15 للرقم 23 لنسحب الرقم 38، وهكذا .... وتكون آخر بطاقة مسحوبة هي رقم 2993.

ونلاحظ هنا أنه إذا لم يكن طول المجال عددا صحيحا فإننا نقرب العدد إلي أقرب عدد صحيح.

3- <u>العينة الطبقية العشوائية</u>:

تعرف العينة الطبقية العشوائية بأنها العينة التي تؤخذ من خلال تقسيم وحدات المجتمع إلى طبقات متجانسة واختيار عينة عشوائية بسيطة أو منتظمة من كل منها.

وتستخدم هذه الطريقة عندما يكون المجتمع منقسما إلى طبقات طبيعية وتكون لدينا الرغبة في تمثيل جميع هذه الطبقات في العينة. وتتلخص الطريقة بتحديد حجم العينات الجزئية المتناسبة من كل طبقة على أساس المعادلة التالية: **حجم العينة الطبقية = (حجم الطبقة ÷ حجم المجتمع) × حجم العينة**

مثال: بافتراض أن مجتمع يتكون من طبقات تحتوي على عناصر كما يبينه الجدول التالي:

الطبقة الأولى	الطبقة الثانية	الطبقة الثالثة	الطبقة الرابعة	الطبقة الخامسة
500	400	280	200	220

وأراد باحث اختيار عينة حجمها 160 من هذا المجتمع، فما هو حجم العينة في كل طبقة.

الحل: حجم المجتمع الكلي = 500 + 400+ 280+ 200+ 220 = 1600

حجم العينة من الطبقة الأولى = $160\times\frac{500}{1600}=50$

حجم العينة من الطبقة الثانية = $160\times\frac{400}{1600}=40$

حجم العينة من الطبقة الثالثة = $160\times\frac{280}{1600}=28$

حجم العينة من الطبقة الرابعة = $160\times\frac{200}{1600}=20$

حجم العينة من الطبقة الخامسة = $160\times\frac{220}{1600}=22$

4- العينة العنقودية:

وتستخدم في حالة حجم المجتمع كبير جدًا، وعندما يكون تقسيم هذا المجتمع إلى مجموعات صغيرة تسمى عندئذ عناقيد، فإننا نختار عينة عشوائية من هذه العناقيد. مثلا إذا أردنا إجراء دراسة على أحد مناهج السنة الرابعة من التعليم الابتدائي، فإن مجتمع الدراسة يشمل كل مؤسسات التربية على مستوى التراب الوطني، ويصعب أخذ عينة عشوائية بسيطة منه، ولأن هذا المجتمع مقسم إلى مديريات التربية (عناقيد) وهذه العناقيد مقسمة إلى عناقيد صغيرة تتمثل في المدارس التابعة لهذه المديريات وهذه الأخيرة تحتوي على أقسام السنة الرابعة والتي تتمثل في عناقيد أصغر. ولاختيار عينة عشوائية يتم اختيار بشكل عشوائي قسم من كل مدرسة ويكون جميع التلاميذ من هذا القسم ضمن العينة الكلية.

2.2 طرق جمع البيانات:

يمكن التمييز بين عدة طرق لجمع البيانات نذكر منها:

أولا. المقابلة الشخصية:

وهي أن يقوم الباحث بمقابلة أفراد العينة والتحدث إليهم عن موضوع البحث المراد إنجازه، وبذلك فان كمية المعلومات التي سيقوم بجمعها ستكون دقيقة إلى حد ما، إلا أن تحليلها سيكون صعبا، وعليه يجب أن يَنتَبِه إلى تدوين البيانات أثناء المقابلة، لأن أي خطا في التدوين يؤدي إلى خطا في النتائج.

ثانيا. الملاحظة المباشرة:

عندما لا يكون هناك أفراد للعينة، فإننا نستخدم طريقة الملاحظة المباشرة، ومن أمثلة ذلك أن تقف عند نقطة تقاطع طرق (Intrsection)، وتحسب عدد السيارات التي تمر من هذا التقاطع خلال فترة زمنية معينة (من الساعة الثامنة إلى الساعة التاسعة صباحا مثلا) بهدف حصر كثافة السير في وقت ذهاب الموظفين إلى أعمالهم، أو أن تقوم بمراقبة تصرف مجموعة من الأطفال أثناء اللعب وتدوين الملاحظات بهدف التعرف على سلوكات الأطفال في بعض المواقف.

ثالثا. الإستبيان :

يعتبر الإستبيان من أكثر الوسائل استعمالا في جمع البيانات اللازمة للتحقق من فرضيات مشكلة معينة أو للإجابة على أسئلة بحث معين، وعند تصميم الإستبيان يجب مراعاة بعض الشروط حتى تضمن دقة النتائج وصحتها، ومن أهم هذه الشروط:

- **يجب أن تكون أسئلة الاستبيان بسيطة ومفهومة للجميع بنفس الطريقة ولا تكون غامضة.**

مثال: كم عدد الأطفال لديك؟

هنا يتحير المجيب ليسال هل الطفل من هو دون سن الخامسة أم السابعة أم العاشرة...

ولذلك على الباحث أن يعيد السؤال ليصبح مثلا:

كم عدد الأطفال الذين تقل أعمارهم عن 12 سنة لديك؟

- **يجب على الباحث أن يبتعد عن تلك الأسئلة التي توحي بالإجابة. وغالبا ما تكون الأسئلة المنفية موحية بالإجابة**

مثال: ألا تعتقد أن أسلوب هذا الكتاب مبسط للدارس؟ نعم ☐ لا ☐

فالمجيب سيقوم باختيار الإجابة الأولى، وكأن الباحث يريد أن يقوم المستجيب بالإجابة كما يريد الباحث.

- **يجب تحديد الكميات أو الوحدات عندما تكون الإجابات أرقاما.**

مثال: كم تحتاج من كمية الماء للشرب يوميا؟ ....

سيجيب أحد الأشخاص لتر ماء ويجيب آخر 5 كئوس، أو غير ذلك. لذلك يجب إعادة صياغة السؤال إلى كم لترا من الماء تشرب في اليوم؟ ...، أو كم كأسا من الماء تشرب في اليوم؟ أو حتى: ما هو متوسط استهلاكك اليومي للماء؟ مع تحديد مكان الاجابة بذكر الوحدة مثلا:.....لتر.

- **يجب أن تكون الأسئلة مباشرة وواضحة وان لا يفكر المُجِيب بعمق ليجيب على الأسئلة.**
- **يجب أن يكون الاستبيان قصيرا قدر الإمكان، حيث قد لا يكون عند المجيب وقتا طويلا للإجابة على أسئلة الاستبيان.**
- **يفضل أن يوزع الاستبيان على مجموعة صغيرة للتجريب وتعديل الأخطاء قبل التطبيق النهائي.**
- **يجب أن يكون الاستبيان صادقاً وثابتاً، فان لم يكن صادقا فلن تكون المعلومات دقيقة. أما إذا لم يكن ثابتا فلن نستطيع تعميم الاستبيان، ولن يكون قرارنا صالحا لفترة زمنية معينة.**

ونقصد بصدق الاستبيان تمثيله للمجتمع المدروس بشكل جيد، أي ان الاجابات المتحصل عليها تعطينا المعلومات المستهدفة، أما ثبات الاستبيان فيعني أنه اذا اردنا توزيع نفس الاستبيان على عينة

أخرى فإن النتائج ستكون متقاربة بتلك المحصلة من العينة الأولى، وتكون النتائج بين العينتين متساوية الاحتمال (معامل ثابت).

ويتم اختبار صدق وثبات الاستبيان بعدة أدوات من أشهرها معامل ألفا كرونباخ ومعامل التجزئة النصفية[2]، وقيم هذه المعاملات تكون محصورة بين الصفر والواحد الصحيح، وكلما كانت قريبة من الواحد نقول أن الاستبيان صادق وأنه ممثل للمجتمع المدروس، والعكس اذا كان يقترب من الصفر، وفي هذه الحالة يجب إعادة صياغة اسئلة الاستبيان وإعادة اختباره من جديد.

وتجدر الاشارة الى أن هذا الاختبار يتعلق بأسئلة الاستبان من دون الاسئلة المتعلقة بالمعلومات الشخصية للمستجوب.

2-3. الترميز (عملية الانتقال من الاستبيان إلى برنامج SPSS):

هذه الخطوة هي الخطوة التي تسبق عملية إدخالها إلى الحاسوب بهدف التحليل وتعرف بترميز البيانات. وهي عملية تحويل إجابات كل سؤال إلى أرقام أو حروف يسهل إدخالها إلى الحاسوب.

حسب مفهوم SPSS فان الأفراد (المشاهدات) الذين يقومون بالإجابة على أسئلة الاستبيان يطلق عليهم اسم الحالات (Cases) ، وكل سؤال (فقرة) في الاستبيان هو عبارة عن متغير، وتسمى إجابات الأشخاص على الأسئلة (الفقرات) بقيم المتغيرات.

يحتوي الاستبيان على عدة أنواع من الأسئلة، وهذه الأنواع هي:

أ. سؤال يسمح باختيار إجابة واحدة فقط:

مثال: هل أنت طالب أو موظف؟ نعم ☐ لا ☐

متغير واحد يكفي لتمثيل هذا السؤال، في هذه الحالة نرمز للإجابة " نعم " بالرمز 1 وللإجابة " لا " بالرمز 2 أو نرمز للإجابة" نعم " بالرمز N وللإجابة" لا " بالرمز Y ولكن يفضل استخدام الترميز بالأرقام لان عملية إدخال البيانات الرقمية في برنامج SPSS تتم بسهولة اكثر ولان الحاسوب يفرق بين الحروف الصغيرة والكبيرة وكذلك أن كثير من الأوامر في البرنامج SPSS تنفذ فقط مع المتغيرات الرقمية ولا تنفذ مع المتغيرات الحرفية.

[2] للتعمق أكثر في استخدام هذا المعامل، أنظر في ذلك: غيث البحر ، معن التنجي: " **التحليل الإحصائي للاستبيانات باستخدام برنامج** IBM SPSS Statistics"، مركز سبر للدرسات الإحصايية رالسياسات العامة، تركيا، 2014، ص 14-20.

مثال: هل توافق أن يكون تسجيل الطالب في الجامعة عبر الانترنت؟

موافق بشدة ☐ **موافق** ☐ **محايد** ☐ **غير موافق** ☐ **غير موافق بشدة** ☐

في هذا المثال ربما يستخدم الرقم 5 ليدل على الإجابة " موافق بشدة" والرقم 4 ليدل على الإجابة " موافق" والرقم 3 ليدل على الإجابة " محايد" والرقم 2 ليدل على الإجابة " غير موافق " والرقم 1 ليدل على الإجابة " غير موافق بشدة".

ب. سؤال يسمح بأكثر من إجابة:

مثال: ما هي أهم الهوايات التي تمارسها ؟

القراءة ☐ **الرياضة** ☐ **السباحة** ☐ **الموسيقى** ☐ **غير ذلك** ☐

ففي هذا السؤال نلاحظ أن الشخص يمكن أن يعطي اكثر من إجابة، لذلك فان متغيرا واحدا لا يكفي لتمثيل السؤال. في هذه الحالة يفضل إنشاء خمسة متغيرات، كل متغير له احتمال إجابتين نعم / لا ويستخدم لهما 1 للإجابة " نعم " و 0 للإجابة " لا"

ج. سؤال مفتوح جزئياً:

ويقصد بذلك، السؤال الذي يسمح للشخص باختيار إجابة موجودة ضمن الخيارات أو كتابة إجابة أخرى غير موجودة ضمن الخيارات.

مثال: عند سفرك للخارج أي خطوط الطيران تستخدم؟

الجزائرية ☐ التركية ☐ القطرية ☐ الأردنية ☐ غير ذلك اذكرها .....

في هذا النوع من الأسئلة فان متغيرا واحدا يكفي لتمثيل هذا السؤال لان المسموح به هو إجابة واحدة فقط (شريطة أن يستخدم المسافر شركة طيران واحدة) إلا أن عملية تعيين رموز تصف قيم المتغير (الإجابات) هي صعبة نوعا ما وتتم باستخدام عدة طرق يمكن تلخيصها كالتالي:

الطريقة الأولى: أن ترمز لكل شركة طيران وردت في الإجابة برقم من 1 إلى N حيث يمثل N عدد شركات الطيران الواردة بالإجابة وهذه طريقة سيئة لأنها تحتاج لوقت كبير، لأنه سيتعامل مع كل استبيان بشكل منفرد ليتم جمع البيانات كلها.

الطريقة الثانية: تعيين الرمز 5 ليصف الإجابة " غير ذلك " بحيث يتم معاملة هذه الإجابات كمجموعة واحدة عند تحليل الإجابات بغض النظر عما ذكر من أنواع شركات الطيران الممكنة. وهذه الطريقة سيئة

لأنها تمكننا من فقدان معلومات كثيرة، إلا أن هذا الفقدان من المعلومات قد لا يكون مشكلة إذا كان الاستبيان يركز على شركات الطيران الواردة في السؤال.

ولاختيار أي الطرق أفضل فإنه يجب الأخذ بعين الاعتبار العوامل التالية :

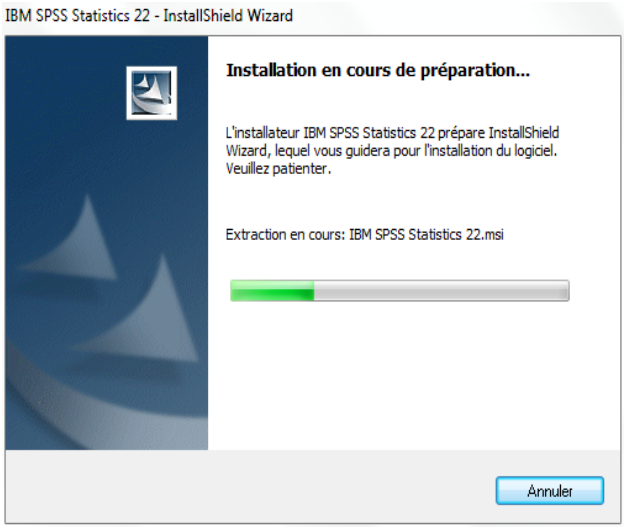
- الهدف من الاستبيان
- شكل الاستبيان الذي تم تقديمه للأشخاص وكيفية الإجابة علية.
- الوقت المتاح للباحث.
- الدعم المادي المتوفر للباحث.
- الدقة المطلوبة.

الفصل الثاني: بعض المبادئ الاساسية لاستعمال برنامج SPSS

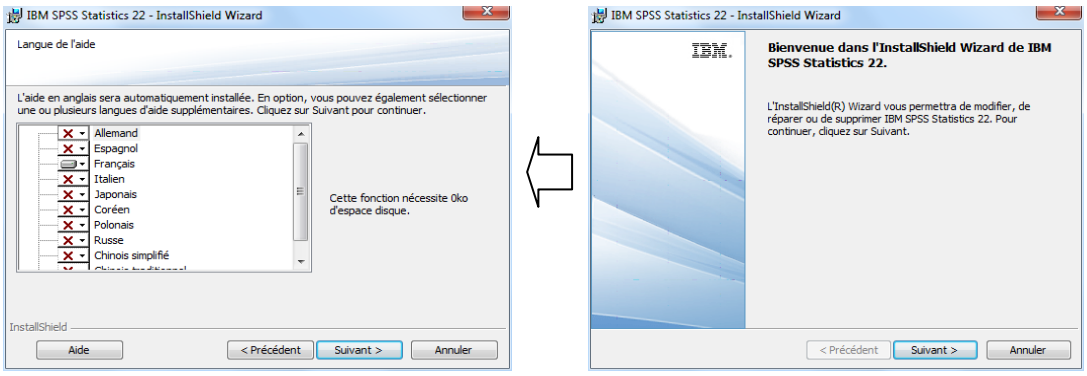
1. التعرف على بيئة البرنامج الاحصائي SPSS:

1.1 تثبيت البرنامج الاحصائي SPSS:

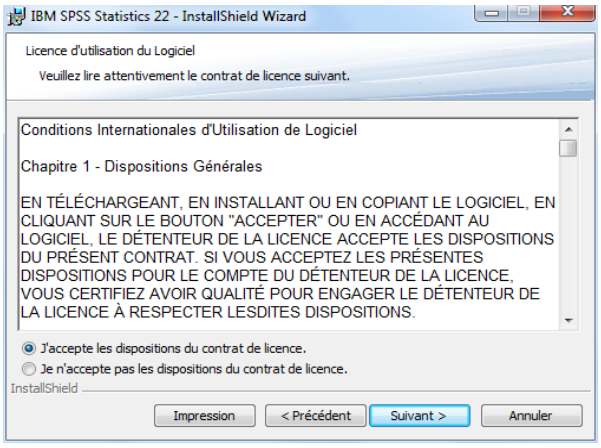
يتوفر البرنامج الاحصائي بإصدارات مختلفة طورت عبر الزمن لتتوافق مع متطلبات التحليل الإحصائي وتطور انظمة الاستغلال Microsoft، ولا تختلف طرق تثبيت البرناج على الحاسوب عن باقي البرامج الجاهزة الأخرى، ويتم تثبيت البرنامج بفتح الملف الخاص بالبرنامج من القرص المضغوط الذي يحتويه، بحيث يتم النقر مرتين بزر الماوس الايسر على الملف SPSS_Statistics_22_win32.exe مثلا للاصدار 22، فيظهر على الشاشة سيرورة تثبيت البرنامج كما توضحه الصورة المقابلة:



ومن ثم النقر على الزر Suivant في صناديق للحوار المعروضة على الشاشة لإتباع عملية التثبيت



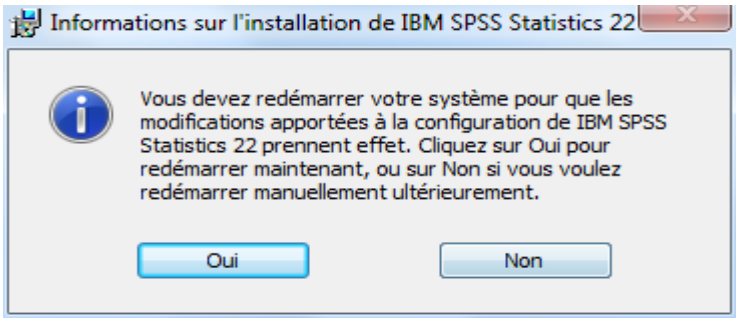
وصولا إلى الخطوة التالية، أين يظهر في صندوق الحوار الخيارين:



- J'accepte les dispositions du contrat de licence
- Je n'accepte pas les dispositions du contrat de licence

اين يتم التأشير على الخيار الاول أي الموافقة على احكام عقد الترخيص، كما تظهره صورة صندوق الحوار المقابلة، ومن ثم النقر على الزر Suivant لاستكمال عملية التثبيت.

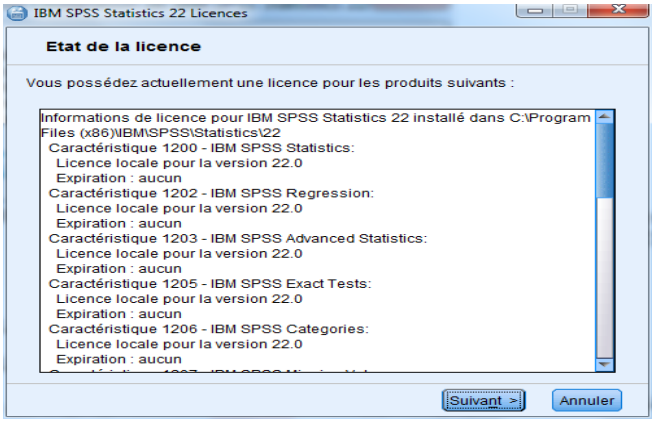
ويتم تثبيت البرنامج بعرض عدة صناديق الحوار، أين يتم دائما النقر على الزر Suivant للاستمرار في عملية التثبيت، وصولا إلى طلب البرنامج إعادة انطلاق البرنامج Redémarrer اين يتم النقر على الزر Ok كما يظهر في الصورة التالية.



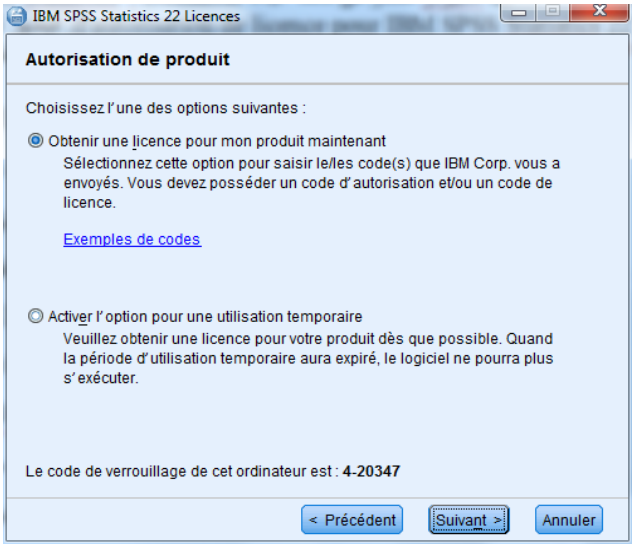
وبعد إعادة تشغيل الكبيوتر نختار من قائمة البدء Démarrer :

IBM SPSS Statistics؛ Assistant d'autorisation de licence pour IBM SPSS Statistics 22.

فيظهر لدينا صندوق الحوار على الشكل التالي:



أي يتم النقر على Suivant ليظهر صندوق حوار خاص بكود تفعيل البرنامج، على الشكل التالي:



بحيث يتيح البرنامج خيارين، إما تشغيل البرنامج باستعمال كود التفعيل الموجود على غلاف القرص المضغوط الذي يحتوي البرنامج في حالة شرائه مباشرة من شركة IBM، أو تشغيله لفترة زمنية محدودة باستعمال ملف الاستعمال الظرفي للبرنامج المتاح على الويب والذي يتم تحميله من موقع IBM.

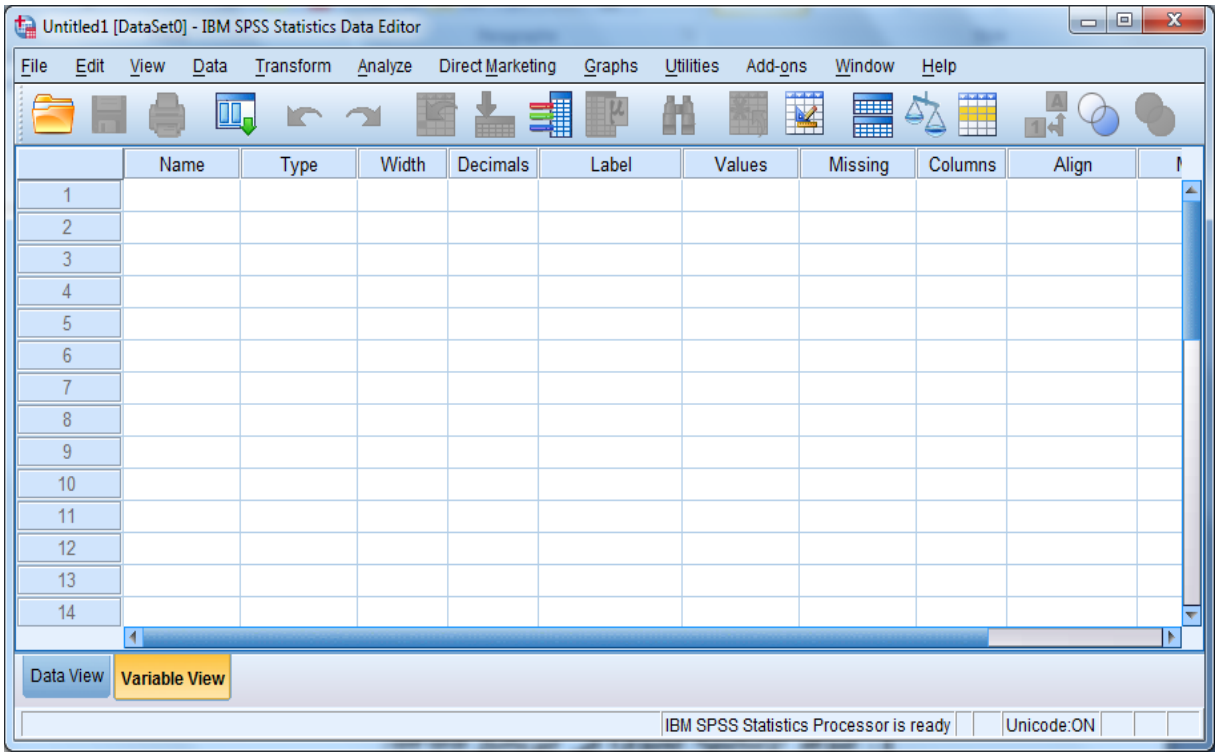
وبعد استكمال هذه الخطوات يصبح البرنامج صالح للاستعمال.

2.1 تشغيل برنامج SPSS:

لتشغيل البرنامج من خلال النوافذ Windows نتبع الخطوات التالية:

1. النقر على زر البدء من النافذة الرئيسية في نظام التشغيل، ثم اختيار قائمة البرامج ثم مجموعة IBM SPSS Statistics.

2. النقر على IBM SPSS Statistics 22 فتظهر النافذة الرئيسة للبرنامج الاحصائي SPSS



المبينة في الشكل التالي :

3.1 النوافذ الأساسية المتوفرة في برنامج SPSS:

يتوفر برنامج الـ SPSS على مجموعة من النوافذ نذكرها فيما يلي:

1. نافذة محرر البيانات Data Editor: تظهر تلقائيا مباشرة عند بداية تشغيل البرنامج، وهي نافذة مخصصة لعرض محتوى ملف البيانات، حيث يمكن أن يكون ملف جديد أو ملف موجود.

2. نافذة المشاهد Viewer: وهي نافذة خاصة بعرض نتائج التحليل الاحصائي من جداول وأشكال بيانية، وهي كذلك تفتح تلقائيا بعد ملف بيانات معين، لكن تكون في نافذة مستقلة تظهر في شريط أدوات النظام باسم « Output ».

3. نافذة محرر الجداول المحورية Pivot Table Editor: هذه النافذة تسمح بإعداد جداول محورية في شكل جداول تقاطع بعدة طرق حسب رغبة الباحث.

4. نافذة محرر المخططات Chart Editor: وهي تسمح بإعداد مختلف الاشكال البيانية.

5. نافذة محرر النصوص Text Output Editor: تتيح هذه النافذة تحوير المخرجات التي لا تعرض كجداول محورية.

6. نافذة Editor of syntax: وهي تسمح بتغيير syntaxe التحكم فى استخدام المهام الخاصة التى لا تتوفر فى صناديق الحوار، كما يمكن تسجيل الطلبات فى قائمة من اجل استخدامها فى دورات لاحقة.

4.1 أنواع الملفات في برنامج SPSS:

يتوفر في برنامج SPSS ثلاث نوافذ أساسية هي:

1. ملف البيانات Data File: وهي ملفات تتكون باستخدام محرر البيانات Data Editor، تحتوى على البيانات التي تستخدم في التحليل الإحصائي وتخزن باستطالة SAV.

2. ملف المخرجات Output Files: وهي ملفات تحتوي مخرجات عمليات التحليل الاحصائي المنجزة خطوة بخطوة، وهي تشمل نتائج التحليل من جداول واشكال بيانية وتخزن باستطالة SPV.

3. ملف التعليمات (Syntax): وهي ملفات تحتوي على الإجراءات الاحصائية المخزنة على شكل أوامر وتخزن باستطالة SPS.

5.1 صناديق الحوار في برنامج SPSS:

عند اجراء التحاليل الاحصائية على برنامج SPSS تظهر صناديق الحوار (Boite de dialogue) الخاصة بكل عملية، كغيره من البرامج الإحصائية، بحيث تسمح باختيار المتغيرات المعنية بالتحليل من بين المتغيرات الموجودة في ملف البيانات دون الحاجة إلى كتابتها، وكذلك اختيار مختلف النتائج المراد الحصول عليها من عملية التحليل حسب ما يوفره البرنامج، بحيث يمكن عرض صناديق حوار ثانوية خاصة بكل عملية احصائية معينة، (وهذا ما سنراه فيما بعد).

ويشمل صندوق الحوار عموما ثلاث اجزاء اساسية وهي:

1. قائمة المتغيرات المتوفرة في المصدر (Sourse Variable List): وهي تشمل كل المتغيرات الموجودة في ملف البيانات المستخدم، ويمكن في بعض النسخ ان تكون من النوع المسموح به لاستخدام التحليل الاحصائي المعتمد[3].

[3] في النسخ القديمة لبرنامج الـ Spss يتم عرض قائمة المتغيرات الممكن اجراء عليها نوع التحليل المختار بينما لا تظهر باقي المتغيرات. أما في النسخ المتطورة للبرنامج يتم عرض كل المتغيارت.

2. قائمة متغيرات الهدف Target Variables List(s): وهي قائمة المتغيرات المعنية بالتحليل والتي تكون واحد أو أكثر حسب الحاجة.

3. أزرار الأوامر Command pushbutons: تقوم باعلام البرنامج تنفيذ عمل معين حسب ما يختاره المستخدم من الأوامر المتاحة، وتنفيذ الأمر أو طلب المساعدة من خلال زر Help.

6.1 القوائم الرئيسية للبرنامج SPSS:

تتوفر النافذة الرئيسية للبرنامج SPSS على مجموعة من المفاتيح الاساسية للقيام بأي عملية في انظمة النوافذ، ويزودها البرنامج بعشر قوائم رئيسية تتخللها قوائم فرعية نستطيع من خلالها القيام بجميع العمليات التي يوفرها برنامج SPSS وتتمثل فيما يلي:

- قائمة ملف File:

تسمح هذه القائمة بالتعامل مع ملفات البيانات من خلال[4]: انشاء ملفات جديدة، فتح ملفات مخزنة (موجودة في ذاكرة الكمبيوتر أو أي مصدر آخر)، تخزين الملفات، طباعة الملفات أو الخروج من البرنامج.

- قائمة تحرير Edit:

تحتوي هذه القائمة على الكثير من الأدوات المهمة مثل نسخ ونقل البيانات من مكان إلى آخر، انشاء متغير جديد، البحث عن حالات معينة، والبحث عن حالات مهمة، وأخيرا الأدات Option والتي تتيح للمستخدم امكانية تغيير طرق العرض، اللغة المستخدمة، رموز الكتابة وغيرها من الخيارات.

- قائمة عرض View:

تتيح هذه القائمة امكانية اظهار شريط الادوات وتعديل الأيقونات المختصرة التي يتم عرضها في شريط الأيقونات المختصرة، والتي يتم استخدامها مباشرة بدون البحث في القوائم، وكذلك نستطيع من خلال هذه القائمة اظهار او اخفاء خطوط الشبكة، وتغيير نوع الخط المستخدم واظهار او اخفاء عناوين القيم.

- قائمة بيانات Data:

تسمح هذه القائمة باجراء مختلف العمليات على البيانات مثلا تعريف المتغيرات وتغيير اسمائها، فرز وتحويل و دمج بيانات مع بيانات أخرى، تجزئة وترتيب البيانات حسب متغير أو متغيرات معينة، وغيرها من العمليات التي تخص البيانات.

[4] نتناول فقط بعض الأوامر الاساسية الخاصة بالقائمة.

- قائمة التحويلات Transform:

نستطيع من خلال هذه القائمة اجراء تحويلات على البيانات من عمليات حسابية بمختلف الدوال الاحصائية والرياضية والمنطقية وغيرها، وإعادة ترميز البيانات الخاصة بالمتغيرات، تحديد الرتب الخاصة بالبيانات وغيرها من العمليات الخاصة بتحويل البيانات..

- قائمة التحليل Analyze:

وهي اهم قائمة في البرنامج، بحيث انها توفر جميع ادوات التحليل الاحصائي العادية والمتقدمة مثل الاحصاء الوصفي، تحليل الانحدار بمختلف انواعه، التحاليل الاحصائية الخاصة بالمتغيرات الكمية والنوعية على حد سواء، استعراض الجداول المحورية والتقاطعية بالاضافة بعض الاختبارات الاحصائية الخاصة بالفروض وغيرها.

- قائمة المخططات البيانية Graphs:

تتيح هذه القائمة امكانية عمل بعض المخططات البيانية بمختلف الأشكال باستعمال البيانات المتوفرة.

- قائمة الأدوات Utilities:

نستطيع من خلال هذه القائمة ايجاد معلومات مفصلة عن الملف المستخدم والمتغيرات التي يحتويها الملف، وتعريف واستخدام المجموعات Sets للمتغيرات المختلفة.

- قائمة الاضافات Add-ons:

تظهر في هذه القائمة مجموعة من البرامج (الخوارزميات) التي تعمل مع برنامج SPSS والتي اضافها المستخدم، بحيث يمكن للمستخدم تشغيل البرنامج من خلال هذه القائمة.

- قائمة نافذة Window:

يمكن من خلال هذه القائمة التنقل بين النوافذ المختلفة والتحكم في حجمها.

-قائمة المساعدة Help:

تزودنا هذه القائمة بنظام مساعدة من شأنه الاجابة على مختلف التساؤلات الخاصة بالمشاكل التي يمكن مواجهتها عند استعمال برنامج SPSS.

7.1 شريط الأدوات (الأيقونات المختصرة):

تتوفر النافذة الرئيسية للبرنامج SPSS على شريط ادوات يحتوي على ايقونات Icons رسومية تمثل وظائف أو عمليات معينة، يمكن تعديلها بالاستعانة بقائمة View، تسهل استخدام البرنامج دون

العودة إلى البحث عن بعض الاوامر من مختلف القوائم، ويظهر في الصورة التالية شريط الادوات الموجدة اسفل شريط القوائم الخاص ببرنامج SPSS كما يظهر في الصورة التالية:

والجدول التالي يلخص معنى كل أيقونة موجودة في الشريط:

الأيقونة	العنوان	الوظيفة
	open	فتح ملف مخزن
	Save	تخزين ملف
	Print	طباعة ملف
	Dialog Recall	إظهار آخر مجموعة من الإجراءات التي تم استخدامها
	Undo	تراجع عن آخر عملية قمت بها
	Redo	الرجوع عن آخر عملية تراجعت عنها
	Go to case	الانتقال إلى حالة (خانة أو صف)
	Go to Variable	الانتقال إلى متغير
	Variable	إعطاء معلومات عن المتغير
	Run descriptive statistics	عرض جداول الاحصاء الوصفي
	Find	بحث عن
	Insert Case	إدراج حالة جديدة إلى الملف
	Insert Variable	إدراج متغير جديد إلى الملف
	Split File	تجزئة الملف إلى جزأين
	Weight Cases	إعطاء أوزان للحالات
	Select Cases	اختيار مجموعة حالات
	Value Labels	إظهار (أو إخفاء) عناوين القيم
	Use Sets	استخدام مجموعات من المتغيرات
	Show all variable	اظهار كل المتغيرات
	Spell check	التدقيق الاملائي

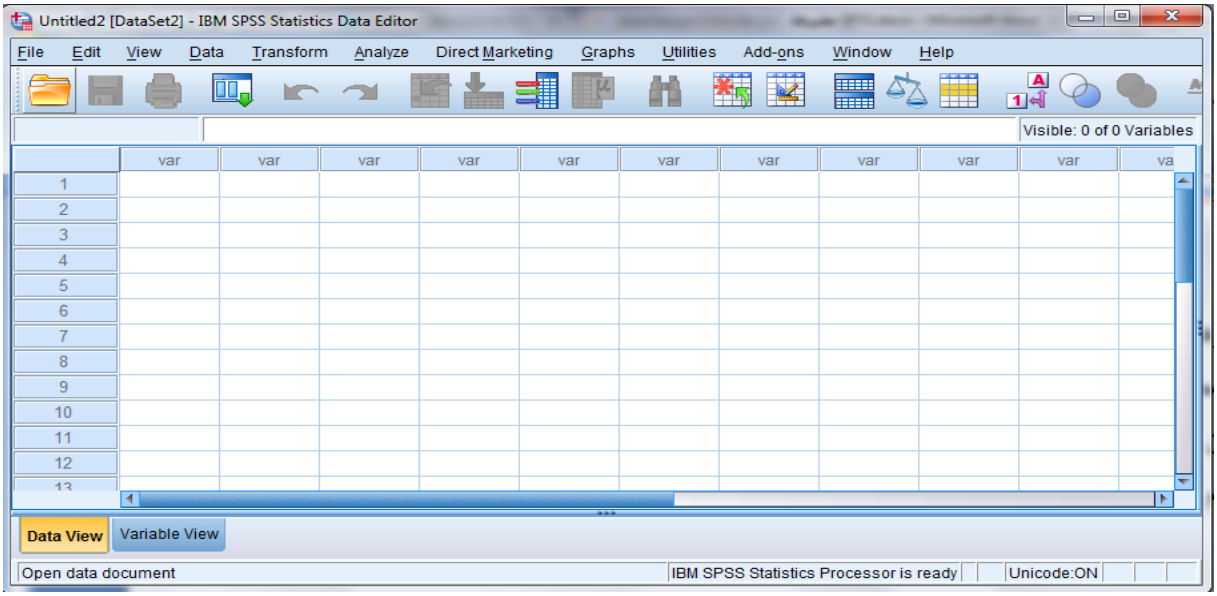
2. تهيئة الملفات وتعبئة البيانات في برنامج SPSS:

ان استخدام برنامج SPSS في التحليل الاحصائي للبيانات يتميز بنوع من التعقيد، ويتعلق الأمر في تقنيات التصريح بالمتغيرات المكونة لقاعدة البيانات، بحيث يختلف تعامل برنامج SPSS عن باقي البرمجيات كبرمجية Excel مثلا أي يكون التصريح بالمتغيرات بطريقة بسيطة جدا بحيث يكفي إعطاء اسماء المتغيرات في السطر الأول فقط ومن ثم يتم تعبئة البيانات في الخانات الخاصة بكل متغير. فتعبئة البيانات في برنامج SPSS يشترط تعريف المتغيرات وتحديد مختلف خصائصها وفق ما هو محدد في البرنامج كخطوة أولى ومن ثم تعبئة البيانات كخطوة ثانية. وعليه وقبل التطرق إلى طرق تعبئة البيانات يجب التطرق إلى بعض الخصائص المتعلقة بتهيئة الملفات في برنامج SPSS.

1.2 تهيئة ملفات البيانات:

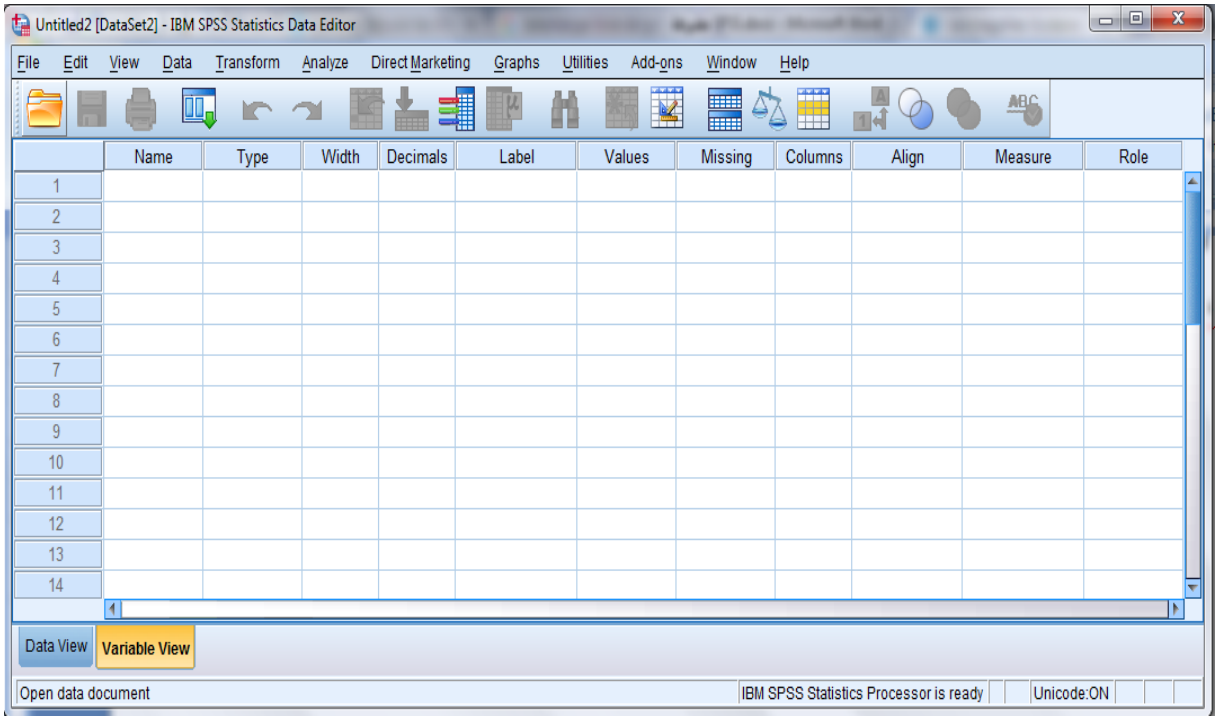
ان تهيئة الملفات في برنامج SPSS تكون من خلال محرر البيانات Data View، وهي تظهر مباشرة عند تشغيل البرنامج، وهو يوفر نوعين من العرض:

أ. **Data View**: وهو عبارة عن ورقة نثر (مصفوفة) تشبه ورقة العرض في برنامج Excel مقسمة إلى أسطر وأعمدة، بحيث يمثل كل عمود متغير معين بينما تكون الأسطر عبارة عن القيم التي تأخذها المتغيرات عند مستويات معينة، فكل خلية من هذه المصفوفة تمثل نقطة تقاطع المتغير مع الحالة.



ب. **Variable View**: وهي ورقة خاصة بصفات المتغيرات، وهي عبارة عن ورقة مقسمة إلى أسطر وأعمدة، بحيث تمثل الأسطر أسماء المتغيرات، والأعمدة تمثل صفات المتغيرات، وهي محددة

(مقيدة) في البرنامج وتتكون من: اسم المتغير، نوع المتغير، عدد الرموز، عدد المراتب العشرية، عنوان المتغير، عناوين القيم، القيم المفقودة، عرض العمود، القياس ودور المتغير.

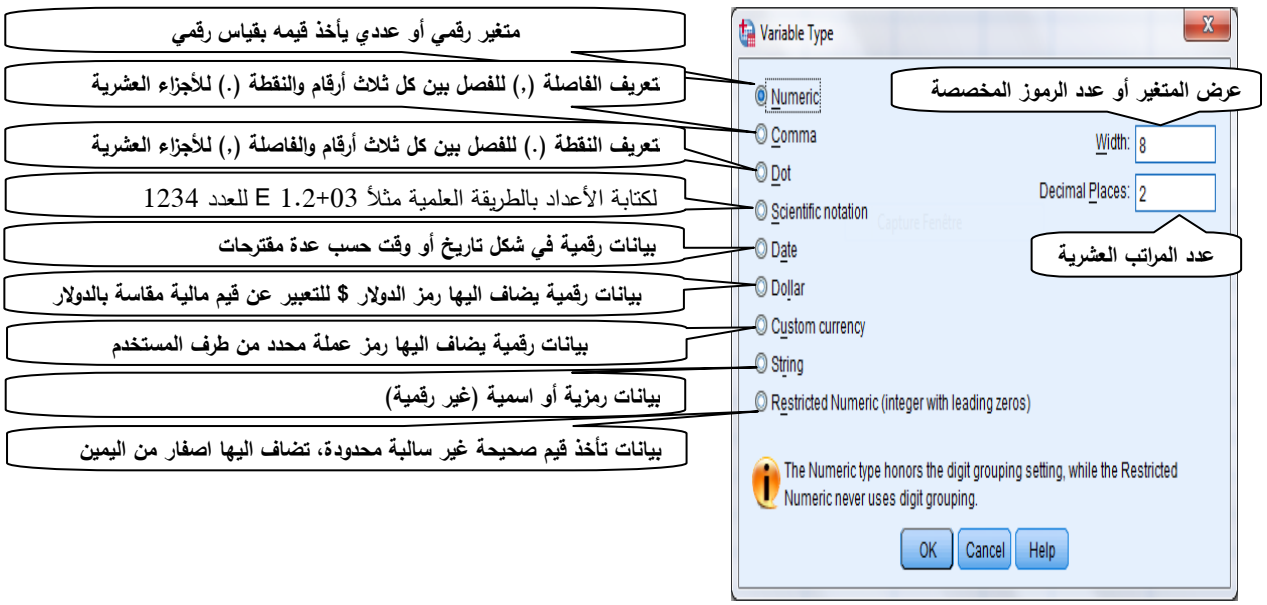


1. **اسم المتغير Name:** يتم استخدام عمود Name لإعطاء اسماء المتغيرات بحيث تمثل كل خلية من هذا العمود اسم متغير، لكن يشترط البرنامج بعض الشروط في تسمية المتغيرات منها:

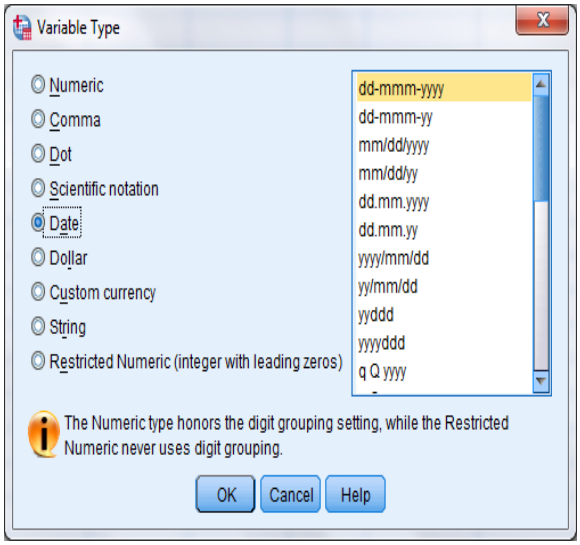
- أن لا يزيد طول الاسم عن 8 رموز Caractères،
- كما يجب ان يبدأ اسم المتغير بحرف والباقي يمكن ان يكون حروفا أو ارقام أو رموز.
- لا يمكن ان يكون في نهاية اسم المتغير فترة Espace.
- لا يمكن لاسم المتغير ان يتضمن فراغات أو بعض الرموز الخاصة مثل: ؟، !، '، *.

ونشير هنا إلى أنه يمكن للباحث إعطاء أي اسم للمتغير حسب ما يراه مناسبا، لكن من الأفضل أن يكون مختصرا لربح الوقت، ومن ثم إعطاء عنوان للمتغير، مثلا كتابة Q02 للتعبير عن سؤال خاص بالجنس مثلا.

2. **نوع المتغير Type:** وهو العمود الثاني في شاشة variable view ، وهو مخصص لتحديد نوع المتغير حسب ما هو محدد في البرنامج بالنقر على المربع الصغير في يمين الخانة فيظهر صندوق حوار variable type كما هو موضح في الشكل التالي:



لتعريف متغير معين يجب اختيار نوع المتغير، بالتأشير على الخيار المناسب من بسن الخيارات المتاحة في صندوق حوار Variable Type، ثم النقر على Ok. أما الزر Width و Places Decimal فهي نفسها الموجودة في عمودي Width و Decimal في شاشة Varible View.



وفي حالة اختيار نوع المتغير في شكل تاريخ (Date)، تظهر عدة خيارات في صندوق الحوار الخاص بنوع المتغير كما هو موضح في الصورة التالية:

بحيث يعرض البرنامج عدة خيارات لطريقة كتابة التاريخ، حسب رغبة المستخدم. فمثلا في حالة اختيار الرمز (-) للفصل بين اليوم والشهر والسنة، يجب على المستخدم التقيد بذلك فالبرنامج لا يعرض الكتابة التي لا تتوافق مع هذا الخيار.

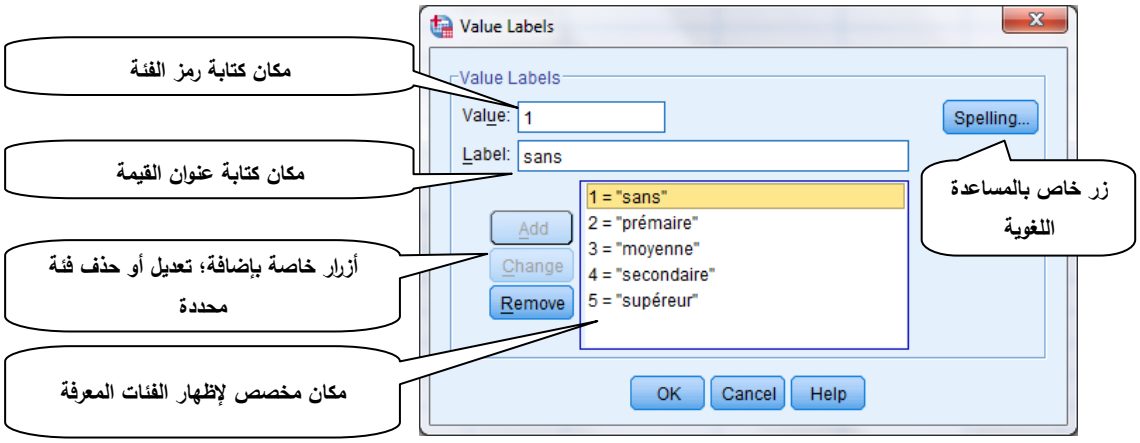
3. **عرض المتغير Width:**

يمثل عدد الرموز المخصصة لكتابة بيانات المتغير المعني، وهو يشمل عدد الارقام الصحيحة زائد عدد الارقام العشرية في حالة المتغير العددي (الرقمي)، وعدد الحروف بالنسبة للمتغير الاسمي.

4. **عدد المراتب العشرية Decimals:** يمثل عدد المراتب العشرية المخصصة للبيانات الكمية، بحيث يتم زيادة أو تخفيض عدد المراتب المخصصة للارقام بعد الفاصلة باستخدام الاسهم التي تظهر في اقصى يمين الخانة.

5. **عنوان المتغير Label:** هو مخصص لإعطاء عنوان المتغير، ككتابة الاسم الكامل للمتغير مثلا: إعطاء للمتغير Ned عنوان Niveau d'éducation، بحيث يسمح البرنامج بكتابة عنوان يتكون من 256 رمز كحد أقصى، كما يسمح باستخدام مختلف الرموز والحروف بدون قيد على خلاف اسم المتغير. والغرض من عنوان المتغير هو اظهاره في المخرجات بدل من الاسم المختصر للمتغير.

6. **عناوين القيم Values:** مخصص لإعطاء عناوين للقيم، والغرض منه هو تعريف الأرقام المعبرة عن فئات متغير معين باسم الفئة وهو خاص بالمتغيرات الإسمية المعبر عنها بفئات، مثلا ترميز متغير الجنس: 1 للذكور و 2 للإناث. ويتم تعريف عناوين القيم بالنقر على المربع الصغير يمين الخانة في عمود Values فيظهر صندوق الحوار التالي:



ان تعريف عناوين القيم يتم عبر عدة خطوات نختصرها فيما يلي:

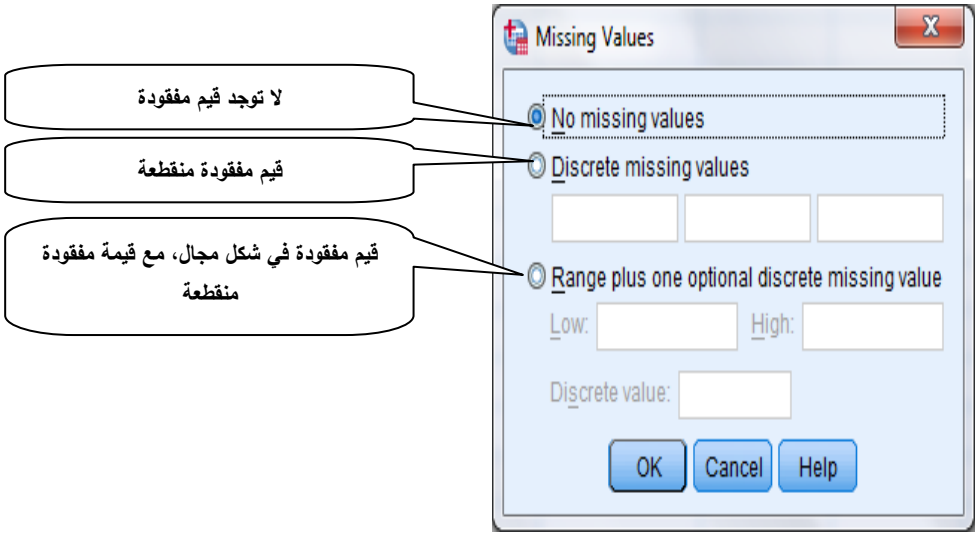
1. كتابة رمز الفئة في خانة Value؛
2. كتابة عنوان القيمة في خانة Label؛
3. النقر على الزر Add لإضافة الفئة.

فتظهر الفئة في المكان المخصص لإظهار عناوين القيم، وفي حالة الرغبة في تعديل إحدى الفئات يتم النقر عليها وتعديلها ثم النقر على الزر Change لتحويلها إلى الصيغة الجديدة، أما في حالة الرغبة في حذف احدى الفئات فيكفي النقر عليها ثم النقر على الزر Remove. وفي الاخير يكفي النقر على زر Ok لتخزين المعلومات، فيتم تحويل المتغير من شكل رموز إلى شكل اسمي حسب الفئات المحددة وتخزن في ذاكرة البرنامج، ولإظهارها يكفي النقر على الأيقونة في لوحة الادوات الخاصة بالبرنامج في شاشة Data View.

7. **القيم المفقودة Missing:** وهي خاصة بالتعريف بالقيم المفقودة ان وجدت، بحيث تسمح باعتبار بعض القيم الخاصة بالمتغير انها مفقودة، فمثلا في بعض الحالات لا يتم الاجابة على أحد أو بعض

اسئلة الاستبيان أو الاجابة عنها يكون خلاف لما هو مقترح، هذه الاجابات يجب استبعادها في التحليل الاحصائي للبيانات. ويعتبر البرنامج الخانات الفارغة والتي تظهر بالرمز الفاصلة قيم مفقودة تلقائيا في حالة البيانات العددية بينما القيم الاخرى المراد اعتبارها مفقودة يجب تعريفها، أما في حالة البيانات الإسمية فيعتبرها البرنامج عبارة عن سلسلة رموز فارغة وتؤخذ في الاعتبار في التحليل الاحصائي. ولتعريف القيم المفقودة نتبع الخطوات التالية:

1. انقر على المربع الصغير في يمين خانة missing الخاصة بالمتغير في شاشة Variable View، فيظهر صندوق الحوار Missing Values، بحيث يسمح بتحديد ثلاث قيم مفقودة، أو تحديد مجال معين مثلا في حالة متغير السن أو الطول وغيرها.



2. تحديد القيمة المفقودة (من قيمة واحدة إلى 3 قيم حسب الحاجة) في الخانات المخصصة مثلا كتابة القيمة (9) لذلك بالتأشير على Discrete Missing Values، بحيث عند تعبئة البيانات يجب كتابة الرقم 9 في حالة الاجابات غير المرغوبة أو عدم الاجابة في الاستبيان، أو اعتبار احد الاجابات مستبعدة (مفقودة)، أو تحديد القيمة الدنيا والقيمة العليا لمجال معين اذا تعلق الأمر بمتغير عددي مثلا من 55 سنة إلى 66 سنة، بالتأشير على Range plus one optional discret missing value، بالإضافة إلى امكانية تحديد قيمة وحيدة إضافة إلى المجال المحدد مثلا في حالة عدم الاجابة على السؤال الخاص بالسن.

أما في حالة البيانات الإسمية والتي تعرف بنوع String، وهي الحالة التي لا يعتبر البرنامج الخانات الفارغة مفقودة تلقائيا، فيجب تعريف القيم الفقودة بنفس الطريقة السابقة بتحديد العبارة « NR » كقيمة

مفقودة في خانة Discrete Missing Values، ونشير هنا إلى ان البرنامج لا يعتبر القيمة المعبر عنها بالأحرف الصغيرة مفقودة ويجب كتابتها بالأحرف الكبيرة.

3. بعد تعريف القيم المفقودة ننقر على الزر Ok لتطبيق الأمر.

4. إعطاء عنوان للقيم التي اعتبرناها مفقودة مثلا "عدم الاجابة" في حالة عدم الاجابة على سؤال من اسئلة الاستبيان لكي تظهر في مخرجات التحليل الاحصائي.

8. **عرض العمود Columns:** يسمح بتحديد عرض العمود بزيادة أو تخفيض عرض العمود الخاص بمتغير معين باستخدام الاسهم الموجودة يمين الخانة الخاصة بالمتغير في عمود Columns أو كتابة عرض العمود مباشرة مقاس بعدد الرموز المخصصة للمتغير.

9. **محاذاة النص Alignement:** يتعلق باختيار طريقة عرض البيانات الخاصة بالمتغير من حيث محاذاة النص، بحيث يتوفر البرنامج على ثلاث طرق هي: من اليسار إلى اليمين Left؛ في الوسط Center أو من اليمين إلى اليسار Right، بينما تكون محاذاة النص التلقائية من اليمين إلى اليسار. لكن هذه العملية ليس لها تأثير على عملية التحليل الاحصائي، إنما تتعلق بعملية عرض البيانات في شاشة Data View.

10. **القياس Measure:** وهو عمود خاص بتحديد مقياس المتغير، وهو أمر جد مهم بحيث ان مقياس المتغير يسهل معرفة صلاحية التحليل الاحصائي للمتغير. ويتوفر البرنامج على ثلاث قياسات للمتغيرات تتمثل في:

أ. Scale: بمعنى سلمي، وهو مقياس خاص بالمتغيرات الكمية ذات القياس العددي أو الرقمي، بحيث يتكون من بيانات عددية معبر عنها بأرقام، مثل قياس الطول، المسافة، التكلفة، المبلغ....

ب. Ordinal: بمعنى ترتيب، وهو مقياس خاص بمتغيرات ترتيبية تكون في شكل فئات مرتبة ترتيبا تصتعديا أو تنازليا حسب درجة الأهمية مثلا الاجابة على سؤال استبيان خاص بتحديد درجة أهمية بثلاث ابعاد: درجة كبيرة، درجة متوسطة، درجة ضعيفة. بحيث يتم التعبير عنها بقيم رقمية أو عددية (Numeric).

ج. Nominal: بمعنى اسمي، وهو خاص بالمتغيرات النوعية أو الإسمية سواء كان التعبير عنها برموز أو حروف أو أرقام عددية. والتي لا تتميز بالترتيب مثلا تقسيم أفراد المجتمع حسب المنطقة (1. الحضرية و 2. الريف)، بحيث لا يمكن ترتيبها تصاعديا أو تنازليا وليس لذلك أي معنى على خلاف المتغيرات الإسمية الترتيبية.

مثال (2-1):

لدينا المعلومات الخاصة بعشر طلبة اجتازوا امتحان شهادة البكالوريا:

Prénom	date de naissance	sexe	Série de Bac	moyenne
Sofiane	19.10.1995	Masculin	SNV	10,16
Mohamed	23.06.1993	Masculin	SNV	15,43
SIHEM	15.09.1995	Feminin	GE	10,60
Imed	22.02.1992	Masculin	SNV	12,35
WALID	12.01.1993	Masculin	SNV	10,53
SOUHA	10.09.1994	Feminin	Ge	11,68
NESRINE	01.04.1995	Feminin	GE	11,98
Miloud	20.10.1991	Masculin	GE	14,37
TAREK	25.11.1993	Masculin	TM	10,28
MERIEM	20.12.1994	Feminin	GM	13,55

والمطلوب: تعبئة البيانات في ملف Spss وتخزينه باسم BAC2015.

لتعبئة بيانات الجدول اعلاه في برنامج Spss نتبع الخطوات التالية:

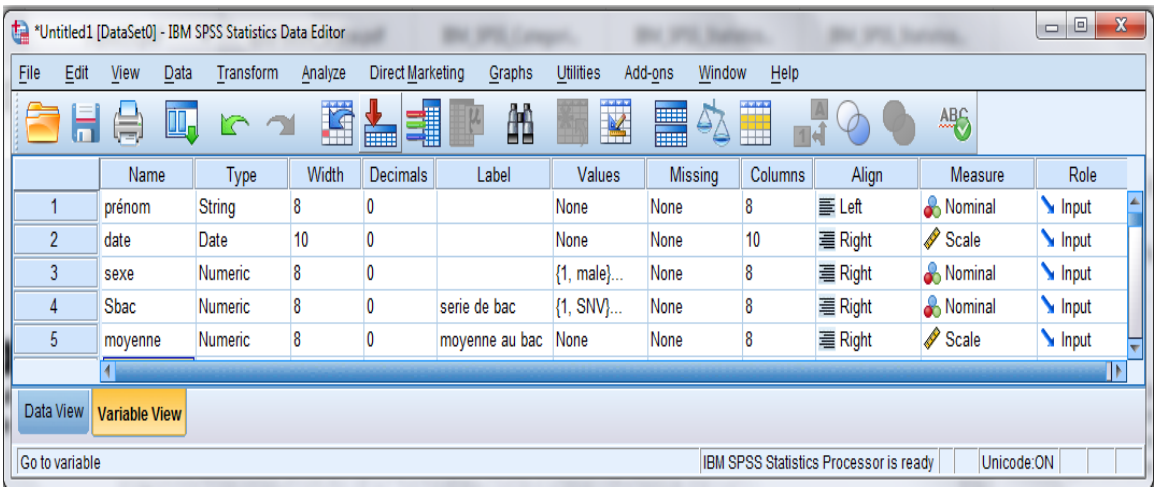
أولا. التصريح بالمتغيرات:

في هذه المرحلة يتم إعطاء اسماء المتغيرات ومختلف خصائصها في شاشة Varible View، بحيث يتم تشغيل البرنامج، والنقر على ايقونة Variable View، ومن ثم النقر على الخانة الأولى في عمود Name وبداية تسمية المتغيرات وتحديد صفاتها بالانتقال (التحرك) فوق سطر المتغير خانة بخانة، مثلا بالنسبة للمتغير:

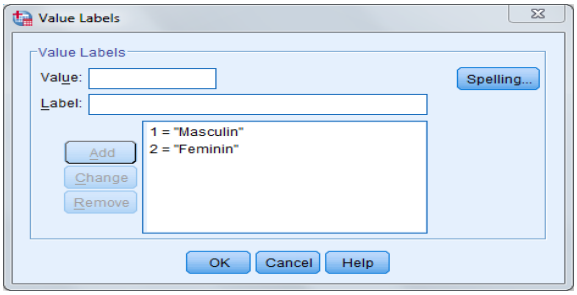
الاسم: Prénom، النوع: String، ونذهب مباشرة إلى عمود Measure ونختار Nominal.

تاريخ الميلاد: الاسم Date، النوع Date مع تحديد شكل التاريخ: dd.mm.yyyy، عرض المتغير 10، عنوان المتغير: Date de naissance، القياس: Scale

وهكذا مع بقية المتغيرات، فتظهر شاشة Variable View كما يلي:

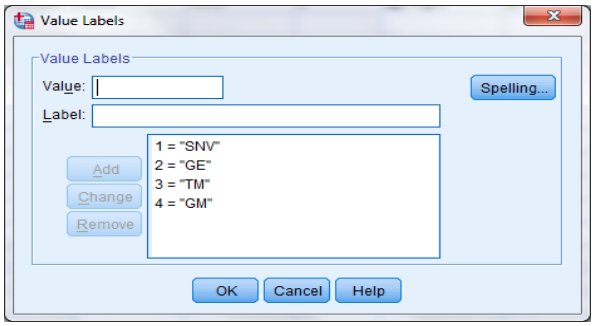


لاحظ أنه تم تحديد نوع مختلف المتغيرات حسب طبيعتها، بحيث تم تحديد النوع الرقمي لبيانات متغيرات الجنس وشعبة البكالوريا بالرغم من كونها اسمية بحيث يتم تعبئة البيانات في طبيعة رقمية عوض الإسمية لربح الوقت.



بالنسبة لمتغير الجنس، تم تعريف قيم المتغير بـ 1 لـ Masculin و 2 لـ Feminin كما تظهر في صندوق الحوار المقابل:

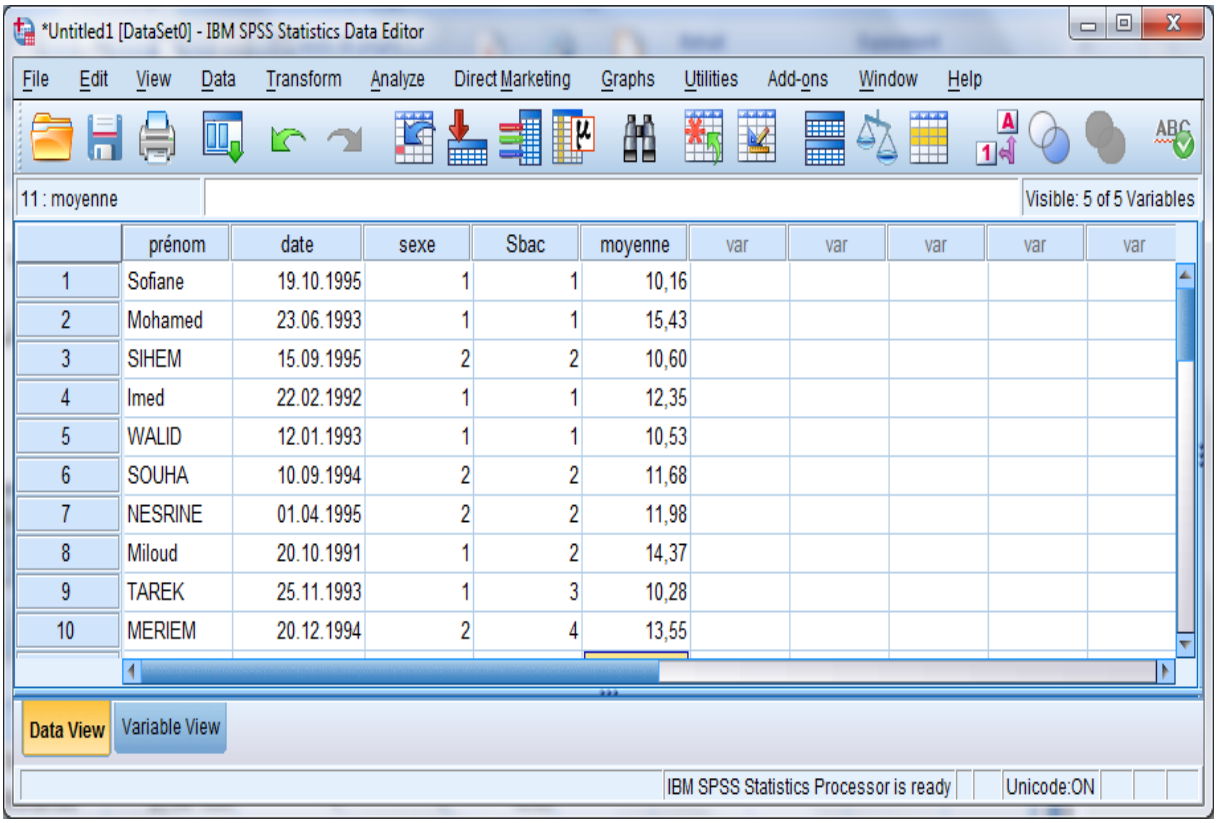
وكذلك الأمر بالنسبة لمتغير شعبة البكالوريا، بحيث تم إعطاء قيم 1، 2، 3 و 4 للشعب: SNV، GE، TM و GM على الترتيب كما تظهر في صندوق حوار Values المقابل:



بالاضافة إلى تحديد قياس كل متغير من المتغيرات حسب طبيعتها، بحيث حدد القياس Nominal لكل من متغير الاسم، الجنس وشعبة البكالوريا، بينما حدد القياس Scale لمتغير تاريخ الميلاد و معدل البكالوريا.

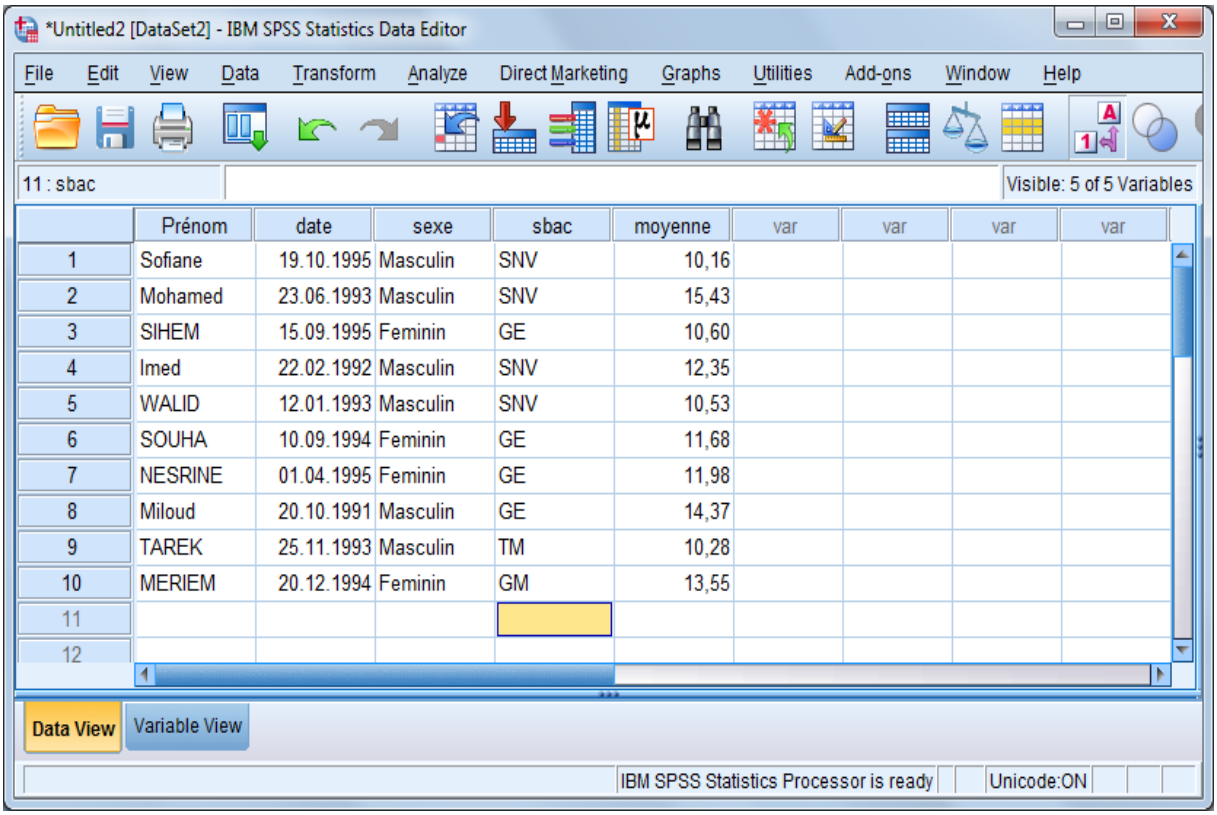
ثانيا. تعبئة البيانات:

بعد التصريح بالمتغيرات ننقر على ايقونة Data View في أسفل شاشة Data Editor، فنجد أن السطر الأول من الجدول يحمل اسماء مختلف المتغيرات المصرح بها، فنقوم عندئذ بملأ الخانات الخاصة بمختلف المتغيرات حسب ما تم التصريح به. فتظهر لدينا شاشة Data View كما يلي:



	prénom	date	sexe	Sbac	moyenne	var	var	var	var	var
1	Sofiane	19.10.1995	1	1	10,16					
2	Mohamed	23.06.1993	1	1	15,43					
3	SIHEM	15.09.1995	2	2	10,60					
4	Imed	22.02.1992	1	1	12,35					
5	WALID	12.01.1993	1	1	10,53					
6	SOUHA	10.09.1994	2	2	11,68					
7	NESRINE	01.04.1995	2	2	11,98					
8	Miloud	20.10.1991	1	2	14,37					
9	TAREK	25.11.1993	1	3	10,28					
10	MERIEM	20.12.1994	2	4	13,55					

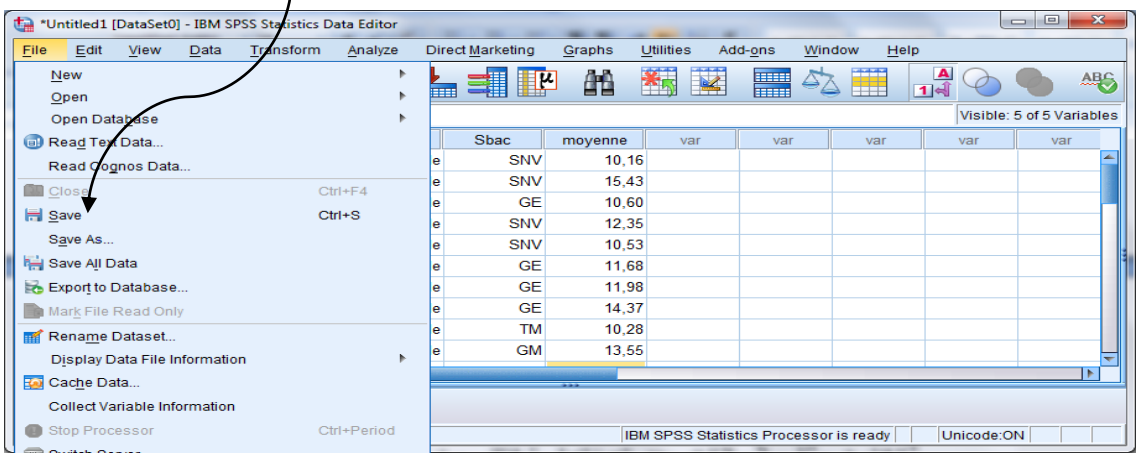
ويمكن اظهار البيانات كما هي في الجدول الخاص بالبيانات بالنقر على الايقونة في شريط الأدوات فتظهر لدينا الشاشة على الشكل التالي:

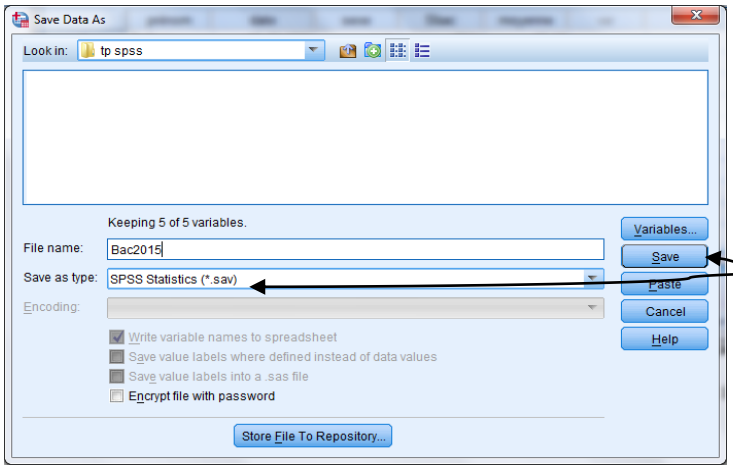


	Prénom	date	sexe	sbac	moyenne	var	var	var	var
1	Sofiane	19.10.1995	Masculin	SNV	10,16				
2	Mohamed	23.06.1993	Masculin	SNV	15,43				
3	SIHEM	15.09.1995	Feminin	GE	10,60				
4	Imed	22.02.1992	Masculin	SNV	12,35				
5	WALID	12.01.1993	Masculin	SNV	10,53				
6	SOUHA	10.09.1994	Feminin	GE	11,68				
7	NESRINE	01.04.1995	Feminin	GE	11,98				
8	Miloud	20.10.1991	Masculin	GE	14,37				
9	TAREK	25.11.1993	Masculin	TM	10,28				
10	MERIEM	20.12.1994	Feminin	GM	13,55				
11									
12									

ثالثا. تخزين الملف:

لتخزين ملف البيانات بعد تعبئتها، نختار من قائمة File الأمر Save As.





ومن ثم اختيار مكان التخزين وتحديد اسم للملف وفي حالتنا يكون الاسم Bac2015، والذي يكون باستطالة (sav) كونه ملف بيانات، ثم يتم النقر على Save لإتمام عملية التخزين كما هو موضح في الصورة المقابلة

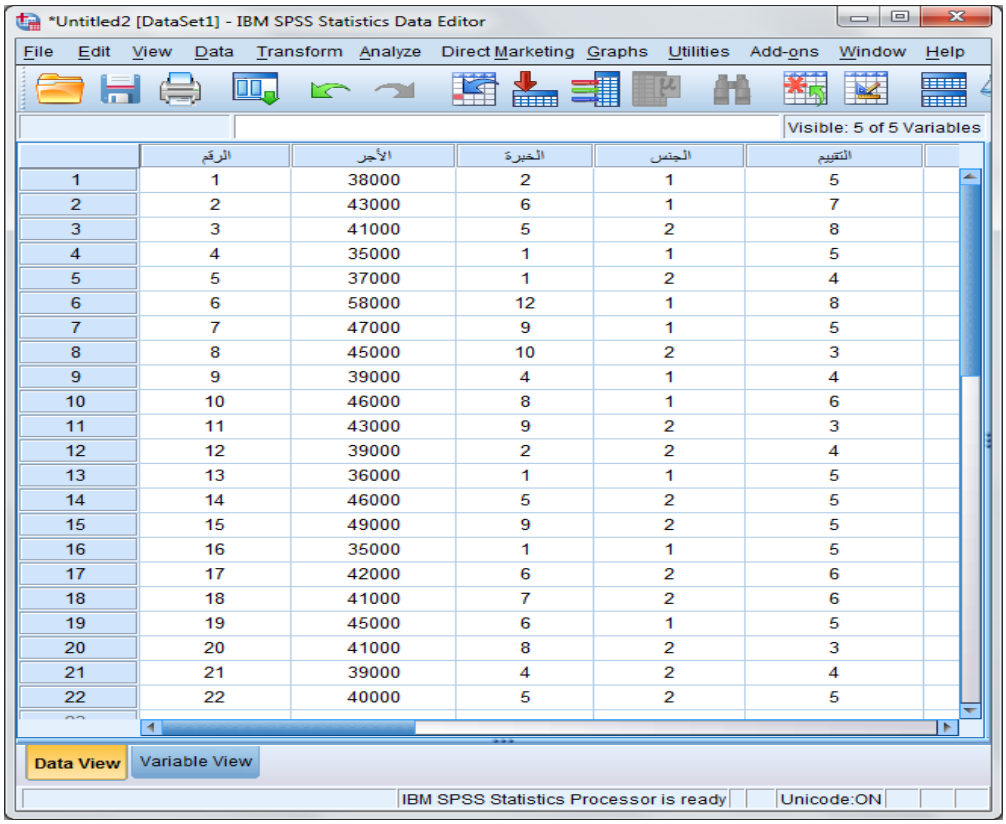
الفصل الثالث: بعض تقنيات ادارة البيانات باستخدام برنامج SPSS

1. العمليات على البيانات:

في غالب الاحيان يحتاج الباحث إلى ادارة ملف البيانات كخطوة اولية قبل اجراء التحليلات الإحصائية، وتتوفر القائمة Data على مجموعة من الأوامر التي تساعد على ادارة البيانات. وفي هذا الصدد نتناول أمرين مهمين من قائمة Data وهما الأمر Sort cases والأمر Split file.

أولا. ترتيب ملف البيانات بواسطة الأمر Sort cases:

يفيدنا الأمر Sort cases في ترتيب بيانات ملف معين حسب متغير أو عدة متغيرات معينة تصاعديا أو تنازليا، والغرض من ذلك تسهيل عملية قراءة البيانات والملاحظة المباشرة لها. ولتنفيذ الأمر نستعين بالبيانات الخاصة بكشف الراتب لـ 22 عامل في مؤسسة MATRIX[5]، والتي تتكون من 5 متغيرات تتمثل في: رقم الكشف، الأجر، الخبرة المهنية، الجنس والتقييم.

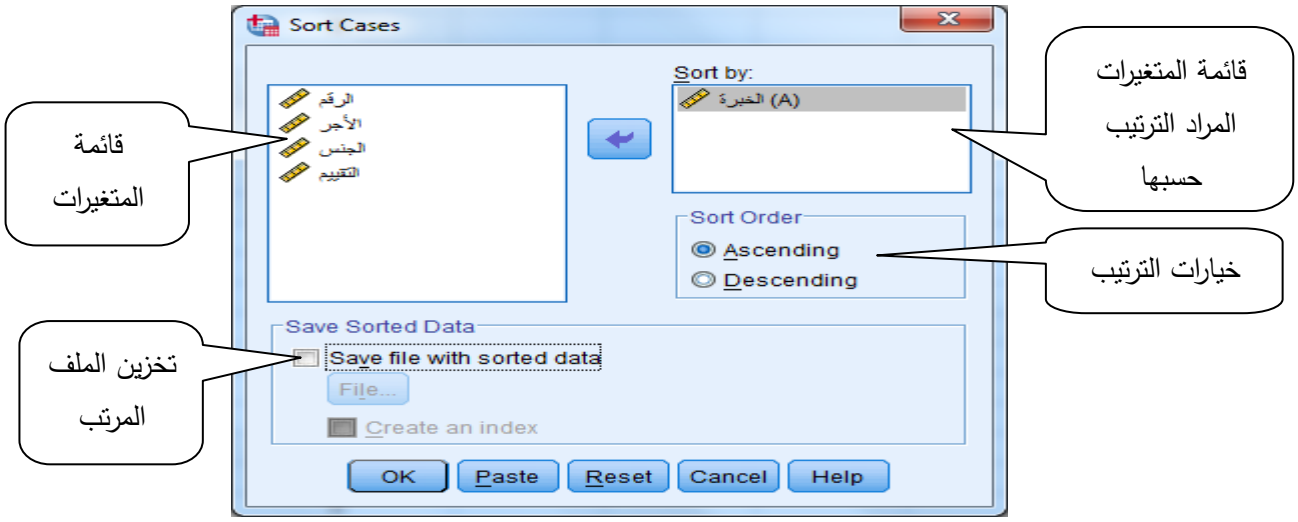


	الرقم	الأجر	الخبرة	الجنس	التقييم
1	1	38000	2	1	5
2	2	43000	6	1	7
3	3	41000	5	2	8
4	4	35000	1	1	5
5	5	37000	1	2	4
6	6	58000	12	1	8
7	7	47000	9	1	5
8	8	45000	10	2	3
9	9	39000	4	1	4
10	10	46000	8	1	6
11	11	43000	9	2	3
12	12	39000	2	2	4
13	13	36000	1	1	5
14	14	46000	5	2	5
15	15	49000	9	2	5
16	16	35000	1	1	5
17	17	42000	6	2	6
18	18	41000	7	2	6
19	19	45000	6	1	5
20	20	41000	8	2	3
21	21	39000	4	2	4
22	22	40000	5	2	5

لترتيب ملف البيانات حسب متغير عدد سنوات الخبرة المهنية للعمال، نتبع الخطوات التالية:

1. نختار من قائمة Data الأمر Sort cases، كما يوضحه صندوق حوار ترتيب الخانات التالي:

[5] Gérald Baillargeon, Fernando ouellet : «Analyse des données avec SPSS pour Windows» ; Les Edition SMG, Canada, 2008, p 20.

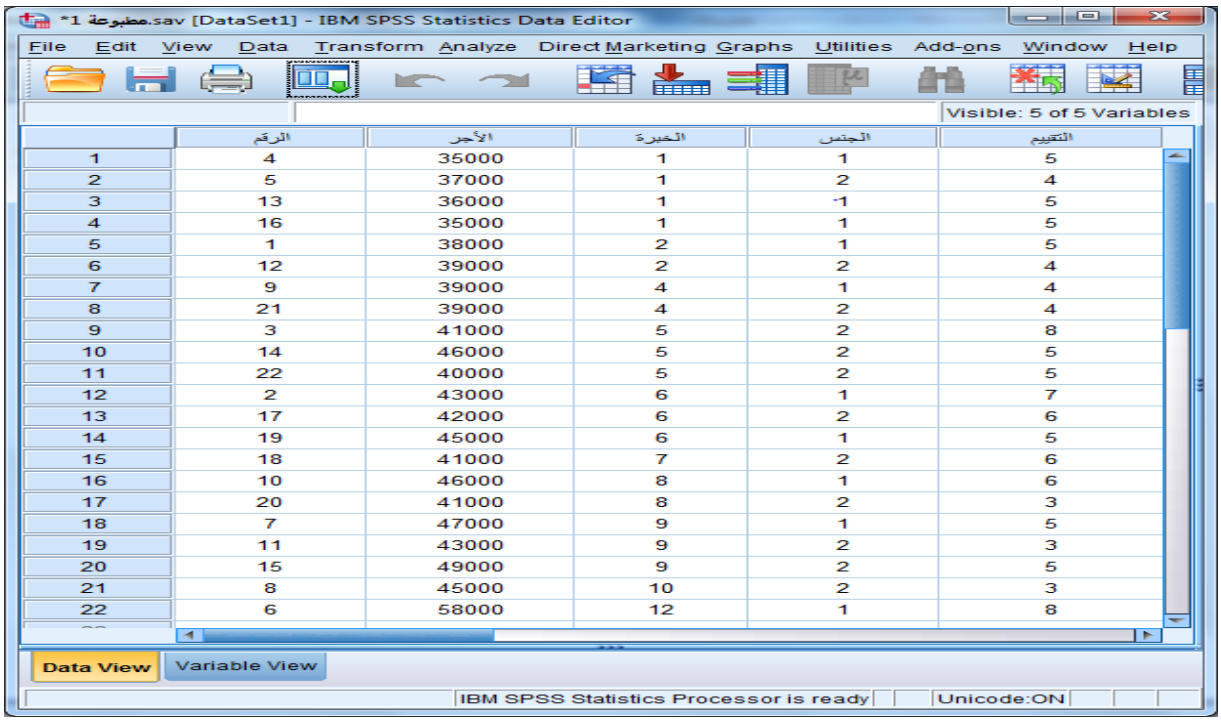


2. نختار المتغير الترتيبي (ترتيب الملف حسبه) وننقله إلى المجال الخاص بذلك: "Sort by".

3. نختار طريقة الترتيب تصاعديا أو تنازليا بالتأشير على الخيار المطلوب.

4. إذا اردنا الاحتفاظ بلمف بيانات مرتب يتم التأشير على المربع الصغير يسار العبارة save file with sorted data، بحيث يعرض البرنامج خيارات التخزين لملف جديد.

كما يتيح البرنامج من خلال صندوق الحوار Sort by أزرار خاصة بخيارات للمستخدم تتعلق بـ: اتمام العملية أو إعادة العملية من جديد دون الرجوع إلى المرحلة السابقة أو الغاء العملية بالإضافة إلى المساعدة عند الحاجة. وبعد اتمام هذه الإجراءات ننقر على الزر Ok فيظهر لدينا الملف مرتب حسب سنوات الخبرة كما يلي:



	الرقم	الأجر	الخبرة	الجنس	التقييم
1	4	35000	1	1	5
2	5	37000	1	2	4
3	13	36000	1	1	5
4	16	35000	1	1	5
5	1	38000	2	1	5
6	12	39000	2	2	4
7	9	39000	4	1	4
8	21	39000	4	2	4
9	3	41000	5	2	8
10	14	46000	5	2	5
11	22	40000	5	2	5
12	2	43000	6	1	7
13	17	42000	6	2	6
14	19	45000	6	1	5
15	18	41000	7	2	6
16	10	46000	8	1	6
17	20	41000	8	2	3
18	7	47000	9	1	5
19	11	43000	9	2	3
20	15	49000	9	2	5
21	8	45000	10	2	3
22	6	58000	12	1	8

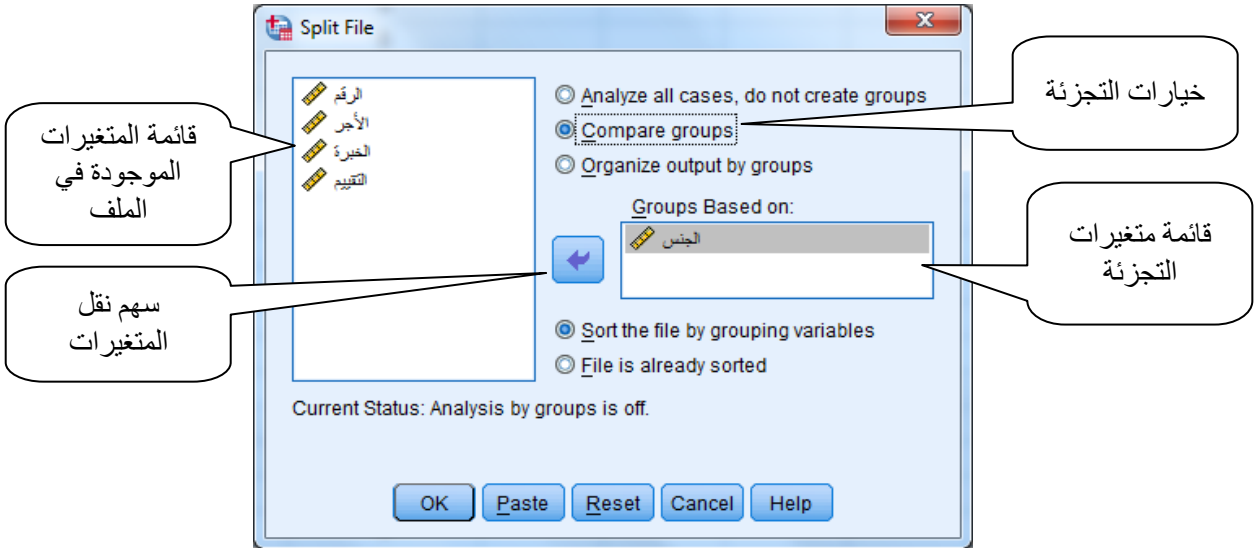
حيث يظهر البرنامج الملف مرتبا حسب عدد سنوات الخبرة تصاعديا بداية من العدد 1 إلى 12.

ثانيا. تجزئة ملفات البيانات بالأمر Split file:

يسمح تجزئة ملف البيانات إلى مجموعات حسب فئات متغير معين بتسهيل عملية تحليل البيانات، بحيث يسمح بتحليلها حسب فئات المتغير المعني عوض تحليل الملف الإجمالي، فمثلا تحليل بيانات الاجور حسب متغير الجنس، بمعنى تحليل الاجور لدى الاناث والذكور كل على حدى. ويسمح البرنامج بتجزئة الملف دون الحاجة إلى تكوين ملفين مستقلين لكل من الذكور والاناث، بل يتم التعامل مع ملف واحد، وذلك من خلال الأمر Split file من قائمة Data.

بالاستعانة بملف البيانات الخاص بكشوف أجور مؤسسة MATRIX السابق (مثال 2-1)، نقوم بتجزئة الملف حسب متغير الجنس بإتباع الخطوات التالية:

1. نختار من قائمة Data الأمر Split file، فيظهر لدينا صندوق الحوار الخاص به كما يلي:



2. اختيار متغير التجزئة ونقله إلى المجال Groups Based on عن طريق النقر مرتين بزر الفأرة الايسر على المتغير المعني أو نقله مباشرة باستعمال السهم وفي مثالنا يتمثل في متغير الجنس.

3. اختيار طريقة التعامل مع التجزئة بحيث يوفر البرنامج عدة خيارات نذكرها فيما يلي:

- Analyze all cases, do not create grops: يعني عدم تجزئة الملف وهو الخيار التلقائي للبرنامج، ويستعمل لالغاء عملية التجزئة والعودة إلى العمل على الملف دون التجزئة؛

- Compare groups: يسمح هذا الخيار بتجزئة الملف إلى مجموعات جزئية حسب متغير أو متغيرات التجزئة، مع عرض نتائج عملية احصائية معينة بشكل يسمح بمقارنة النتائج بين المجاميع الجزئية المختلفة لمتغير التجزئة. مثلا جدول واحد مقسم إلى جزئين لتسهيل عملية المقارنة؛

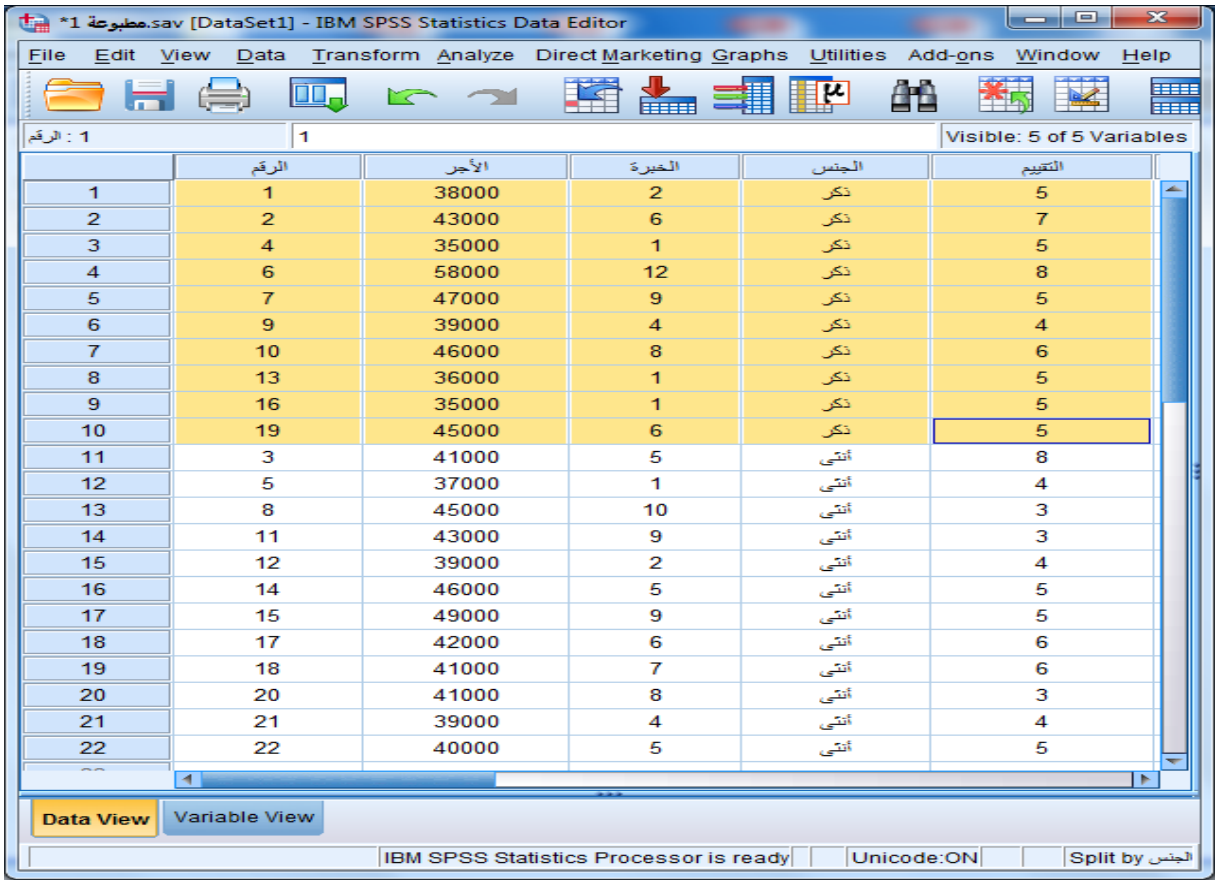
- Organize output by groups: يسمح هذا الخيار بتجزئة الملف، لكن طريقة عرض النتائج تختلف عن الخيار السابق، بحيث يعرض نتائج العملية الاحصائية بصفة مستقلة للمجاميع الجزئية لمتغير التجزئة. مثلا عرض النتائج في جداول مستقلة لتحليلها حسب مجاميع متغير التجزئة كل على حدى.

- Sort the file by grouping variables: وهو كذلك خيار تلقائي للبرنامج بحيث يسمح بترتيب ملف البيانات حسب متغير التجزئة مثل عمل الأمر Sort cases بحيث يظهر الملف مرتب حسب فئات متغير الجزئة.

- File is already sorted: يعني عدم ترتيب الملف، بحيث يتم باختياره التعامل مع ملف البيانات على حاله ويظهر ملف البيانات دون ترتيب، لكن يتعامل معه وهو مجزأ وفق أحد الخيارين الثاني أو الثالث من بين خيارات التجزئة. وفي الغالب يتم التأشير على الخيار الرابع لإظهار نتائج التجزئة والتمييز بين المجاميع الجزئية حسب متغير التجزئة.

4. اختيار أحد الأزرار التي تظهر في اسفل صندوق الحوار، ويتعلق الأمر بتنفيذ أمر التجزئة بالنقر على ok، أو إعادة العملية من جديد بالنقر على الزر Reset، أو الغاء الأمر بالنقر على الزر Cancel، أو طلب المساعدة بالزر Help.

وفي حالتنا نختار التجزئة بالخيار Compare groups أو Organize output by groups. وكلاهما إلى جانب خيار ترتيب الملف Sort the file by grouping variables، وبتنفيذ الأمر نتحصل على ملف بيانات مرتب تصاعديا حسب متغير الجنس ويظهر الجدول وكأنه مقسم إلى قسمين كما يلي:



	الرقم	الأجر	الخبرة	الجنس	التقييم
1	1	38000	2	ذكر	5
2	2	43000	6	ذكر	7
3	4	35000	1	ذكر	5
4	6	58000	12	ذكر	8
5	7	47000	9	ذكر	5
6	9	39000	4	ذكر	4
7	10	46000	8	ذكر	6
8	13	36000	1	ذكر	5
9	16	35000	1	ذكر	5
10	19	45000	6	ذكر	5
11	3	41000	5	أنثى	8
12	5	37000	1	أنثى	4
13	8	45000	10	أنثى	3
14	11	43000	9	أنثى	3
15	12	39000	2	أنثى	4
16	14	46000	5	أنثى	5
17	15	49000	9	أنثى	5
18	17	42000	6	أنثى	6
19	18	41000	7	أنثى	6
20	20	41000	8	أنثى	3
21	21	39000	4	أنثى	4
22	22	40000	5	أنثى	5

ويظهر الاختلاف في خياري التجزئة Compare groups و Organize output by groups عند اجراء عملية إحصائية. ففرضًا أننا نريد استخراج مؤشرات الاحصاء الوصفي[6] لمتغير الأجر، فإننا نتحصل على النتائج التالية:

- بدون اجراء عملية التجزئة: نتحصل على جدول المؤشرات الاحصائية الوصفية حول متغير الأجر لـ 22 عامل دون الأخذ في الاعتبار أي تجزئة كما تظهر في الجدول التالي:

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
الأجر	22	35000	58000	42045,45	5304,780
Valid N (listwise)	22				

- باجراء عملية التجزئة مع اختيار Compare groups: في هذه الحالة يعرض البرنامج النتائج في جدول واحد يشمل مؤشرات احصائية حول متغير الأجر عند 10 عمال من جنس ذكر و12 عامل من جنس أنثى من أجل المقارنة كما يبينه الجدول التالي:

[6] تم اختيار هذه العملية كونها سهلة لاننا لم نتطرق بعد لإجراءات التحليل الإحصائي، فالمؤشرات المقصودة هنا معروفة لدى الطلبة وهي حجم العينة N، الحد الأدنى والأعلى للأجور، المتوسط والانحراف المعياري لمتغير الأجر.

Descriptive Statistics

الجنس		N	Minimum	Maximum	Mean	Std. Deviation
ذكر	الأجر	10	35000	58000	42200,00	7192,589
	Valid N (listwise)	10				
أنثى	الأجر	12	37000	49000	41916,67	3369,875
	Valid N (listwise)	12				

بحيث يظهر جليا أن متوسط الأجر عند الرجال (42200) أكبر من متوسط الأجر عند الإناث (41916.76).

– اجراء عملية التجزئة مع اختيار Organize output by groups: وفي هذه الحالة يعرض البرنامج النتائج في جدولين مستقلين أحدهما يتعلق بالذكور والآخر بالاناث كما يظهر في ما يلي:

Descriptive Statistics[a]

	N	Minimum	Maximum	Mean	Std. Deviation
الأجر	10	35000	58000	42200,00	7192,589
Valid N (listwise)	10				

a. الجنس = ذكر

Descriptive Statistics[a]

	N	Minimum	Maximum	Mean	Std. Deviation
الأجر	12	37000	49000	41916,67	3369,875
Valid N (listwise)	12				

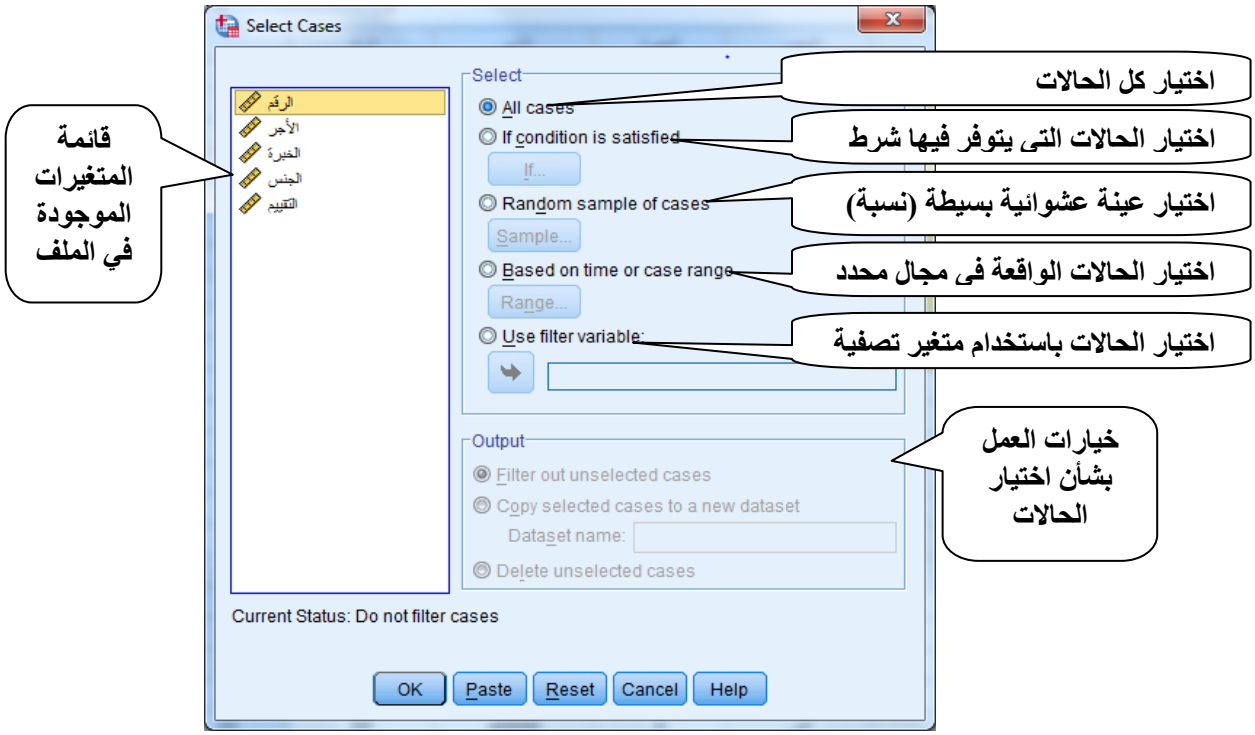
a. الجنس = أنثى

وفي هذه الحالة يستخدم الباحث هذا الاسلوب لتحليل نتائج العملية الاحصائية في صورة مستقلة حسب مجاميع متغير التجزئة وليس في شكل مقارنة.

ثالثا. اختيار الحالات باستخدام الأمر Sort Cases:

يتعلق الأمر هنا بتحديد فئات معينة من بين البيانات التي يحتويها ملف البيانات، من أجل اجراء عملية احصائية معينة عليها دون غيرها من الحالات، فإذا فرضنا اننا نريد تحليل مستوى الأجور لدى الأفراد ذوي الخبرة التي تتجاوز 5 سنوات (5 سنوات فأكثر) دون غيرهم وعوض تكوين ملف جديد يحتوي فقط على فئات العمال ذوي مستوى الخبرة المهنية التي تساوي أو تزيد عن 5 سنوات. وباستعمال برنامج SPSS يمكننا تحديد ذلك بإتباع الخطوات التالية:

1. اختيار الأمر Select cases من قائمة Data، فيظهر على الشاشة صندوق الحوار التالي:



يتوفر صندوق الحوار Select cases على عدة خيارات لاختيار الحالات من خلال الأمر Select، بحيث يعمل كل منها على تحديد فئات محددة من البيانات التي يجرى عليها التحليل نوضحها فيما يلي:

- All cases: يسمح بتحديد كل الحالات من البيانات وهو الخيار التلقائي للبرنامج، لكن يمكن للمستخدم التأشير عليه من اجل الغاء عملية اختيار الحالات التي تم اجراؤها من قبل والعودة للعمل على كل البيانات.

- If condition as satisfied: يسمح باختيار الحالات التي يتوفر فيها شرط معين، فبالتأشير على هذا الخيار يفتح البرنامج المجال للنقر على الزر If... الذي يلي هذا الخيار ومن ثم يظهر صندوق حوار لتحديد الشرط، وهو يحتوي على مختلف الدوال والعلاقات الرياضية والإحصائية والمنطقية وغيرها، والتي يستخدمها الباحث حسب الحاجة، ويعرض البرنامج صندوق الحوار على النحو التالي:

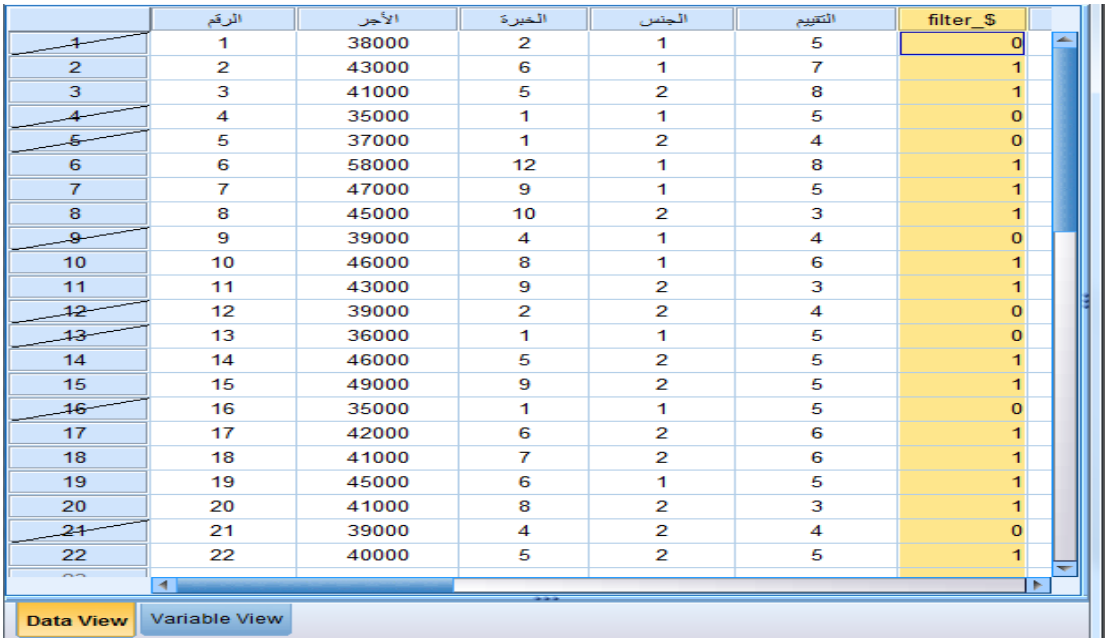
2. كتابة الشرط الخاص باختيار الحالات، بحيث يتم استعمال احد المتغيرات أو عدة متغيرات من بين المتغيرات المتوفرة، والاستعانة بمختلف الدوال والعلاقات التي يوفرها البرنامج، واستخدام الرموز المتوفرة في لوحة المفاتيح الخاصة بصندوق الحوار Select cases if. وفي حالتنا نختار متغير الخبرة ونظيف اليه العبارة (5=>) أي كتابة (5=<الأجر).

3. التأشير على الزر Continue لنعود إلى صندوق الحوار Select cases.

4. التأشير على الزر Ok لتنفيذ الأمر، بحيث يقوم البرنامج بتشطيب الحالات التي تستوفي شرط الخبرة المهنية دون حذفها من ملف البيانات غير أنه يعتمد الخانات غير المشطبة لإجراء العمليات الاحصائية المطلوبة.

لكن يتيح البرنامج للمستخدم خيارات تخزين الحالات المختارة في ملف جديد من خلال التأشير على الخيار Copy selected cases to a new dataset في المجال Output بكتابة اسم الملف الجديد، أو الاحتفاظ بالملف الموجود مع حذف الحالات غير المختارة، وذلك من خلال التأشير على الخيار Delet unselected cases. بينما يتيح الخيار الأول (filter out unselected cases) وهو الخيار التلقائي للبرنامج انشاء متغير جديد خاص بعملية اختيار الحالات بحيث يعطي القيمة 1 للحالة المختارة و القيمة 0 للحالة غير المختارة ويظهر في الملف تلقائيا بعد اتمام العملية، ويمكن استعمال هذا الاخير مرة أخرى عند الحاجة في الخيار الخامس (Use filter variable) دون الحاجة إلى إعادة تحديد شرط اختيار هذه الحالات.

وبإتمام عملية اختيار الحالات بشرط ان يكون للعامل خبرة أكبر أو تساوي 5 سنوات يقوم البرنامج بشطب الحالات التي لا تستوفي هذا الشرط وإنشاء متغير جديد باسم filter_$ كما يظهر في الصورة التالية:



	الرقم	الأجر	الخبرة	الجنس	التقييم	filter_$
1	1	38000	2	1	5	0
2	2	43000	6	1	7	1
3	3	41000	5	2	8	1
4	4	35000	1	1	5	0
5	5	37000	1	2	4	0
6	6	58000	12	1	8	1
7	7	47000	9	1	5	1
8	8	45000	10	2	3	1
9	9	39000	4	1	4	0
10	10	46000	8	1	6	1
11	11	43000	9	2	3	1
12	12	39000	2	2	4	0
13	13	36000	1	1	5	0
14	14	46000	5	2	5	1
15	15	49000	9	2	5	1
16	16	35000	1	1	5	0
17	17	42000	6	2	6	1
18	18	41000	7	2	6	1
19	19	45000	6	1	5	1
20	20	41000	8	2	3	1
21	21	39000	4	2	4	0
22	22	40000	5	2	5	1

Data View | Variable View

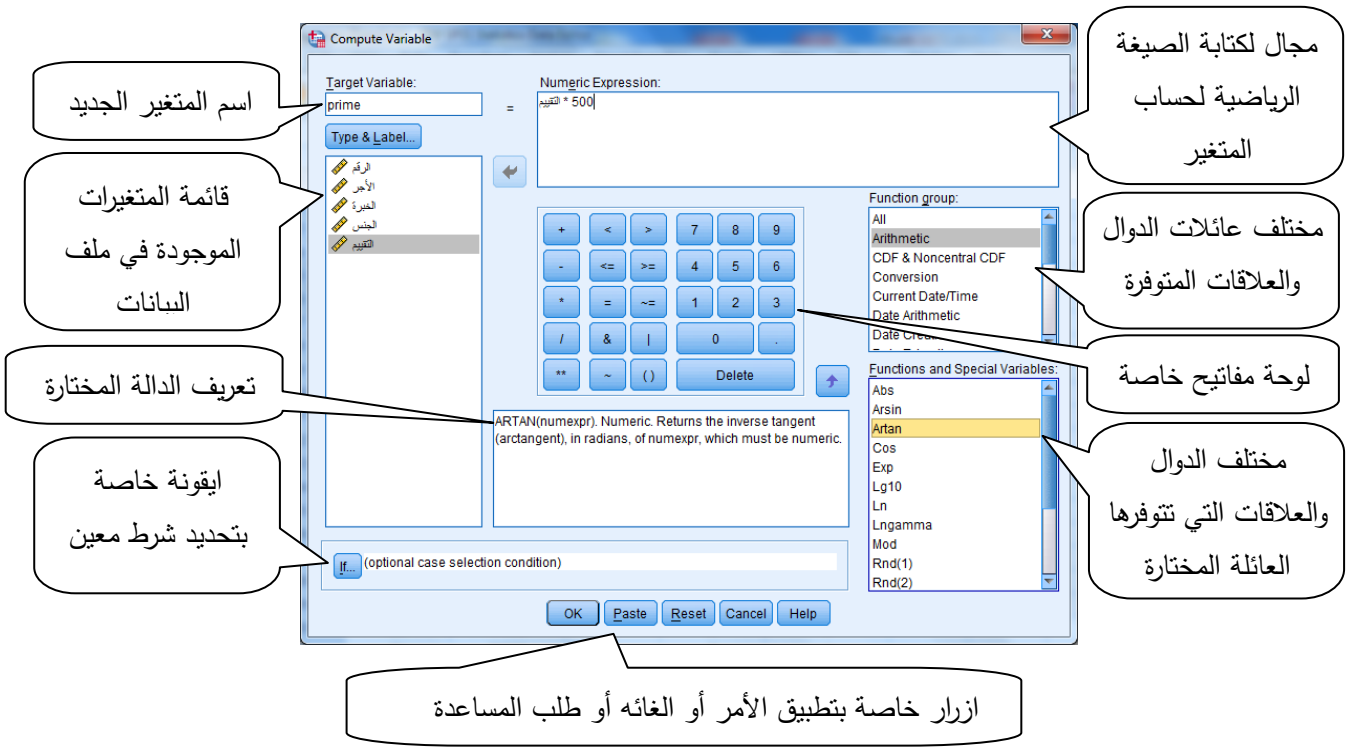
2. العمليات على المتغيرات:

المقصود بالعمليات على المتغيرات هو تحويل المتغيرات باستخدام متغيرات أخرى، كحساب متغير جديد مثلا بالاستعانة بمتغير أو عدة متغيرات، وباستعمال مختلف الدوال والعلاقات الشرطية. أو إعادة ترميز متغير وتحويله إلى الشكل الذي يساعد الباحث في عملية التحليل الإحصائي، كعملية تبويب متغير السن وجعله في شكل فئات مثلا.

ويتوفر البرنامج من خلال قائمة Transform على مجموعة من الخيارات التي من شأنها القيام بعمليات تحويل المتغيرات.

أولا. حساب متغير جديد:

يسمح البرنامج بحساب متغير جديد باستعمال علاقات ودوال تربطه بمتغير أو عدة متغيرات متوفرة في قاعدة البيانات. فباختيار عرض القائمة Transform نجد الخيار الأول Compute variable، وهو الأمر الذي يتم من خلاله حساب متغير جديد. فبافتراض اننا نريد حساب قيمة منحة تمنحها المؤسسة للعمال بقيمة 500 دج لكل نقطة من التقييم، نختار الأمر Compute variable من قائمة Transform فيظهر صندوق الحوار التالي:



في هذه الحالة يتم استخدام متغير التقييم بنقله إلى المجال numeric exepression لحساب قيمة المنحة التي يجب أن نعطيها اسما في Traget variable وعنوانا من خلال الزر Type&Label دون العودة الى شاشة Varible View، ونكتب الصيغة الرياضية للمتغيرة: التقييم*500، والنقر على ok، فيقوم البرنامج بحساب قيمة المنحة (Prime) في عمود جديد يضاف إلى المتغيرات الموجودة.

الرقم	الأجر	الخبرة	الجنس	التقييم	prime
1	38000	2	1	5	2500
2	43000	6	1	7	3500
3	41000	5	2	8	4000
4	35000	1	1	5	2500
5	37000	1	2	4	2000
6	58000	12	1	8	4000
7	47000	9	1	5	2500
8	45000	10	2	3	1500
9	39000	4	1	4	2000
10	46000	8	1	6	3000
11	43000	9	2	3	1500
12	39000	2	2	4	2000
13	36000	1	1	5	2500

كما يوفر البرنامج مختلف الدوال لاستعمالها حسب الحاجة، ويمكن كذلك وضع شرط معين حول البيانات بنفس الطريقة التي ذكرناها سابقا، بحيث يتم حساب المتغير الجديد فقط لفئات البيانات التي تتوفر على هذا الشرط.

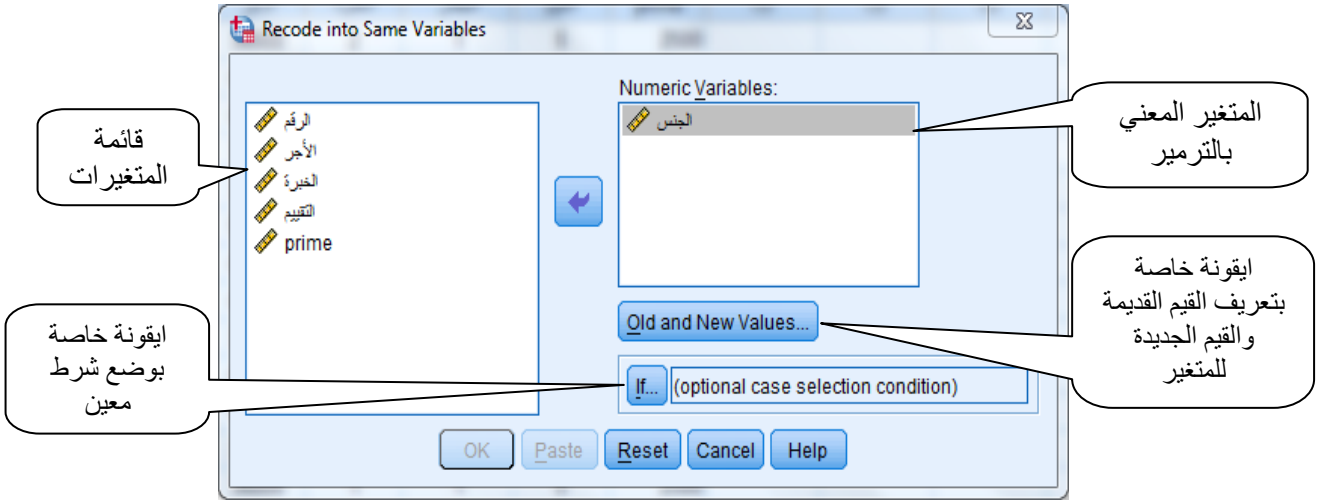
ثانيا. تبويب وإعادة ترميز البيانات:

في كثير من الحالات يحتاج الباحث لتحويل متغير أو عدة متغيرات إلى شكل متغير مستمر في شكل فئات لاغراض احصائية معينة، إلى جانب إعادة ترميز البيانات، وهي الحالة التي يقوم فيها الباحث بإعادة ترميز متغير معين إما برموز جديدة مثلا تحويل القيمة 0 إلى 1، أو تحويل متغير الجنس من شكله الاسمي (ذكر وانثى) إلى رموز (1 و 2). ولهذا الغرض يتوفر برنامج SPSS على أمرين أساسين للقيام بهذه العملية وهما:

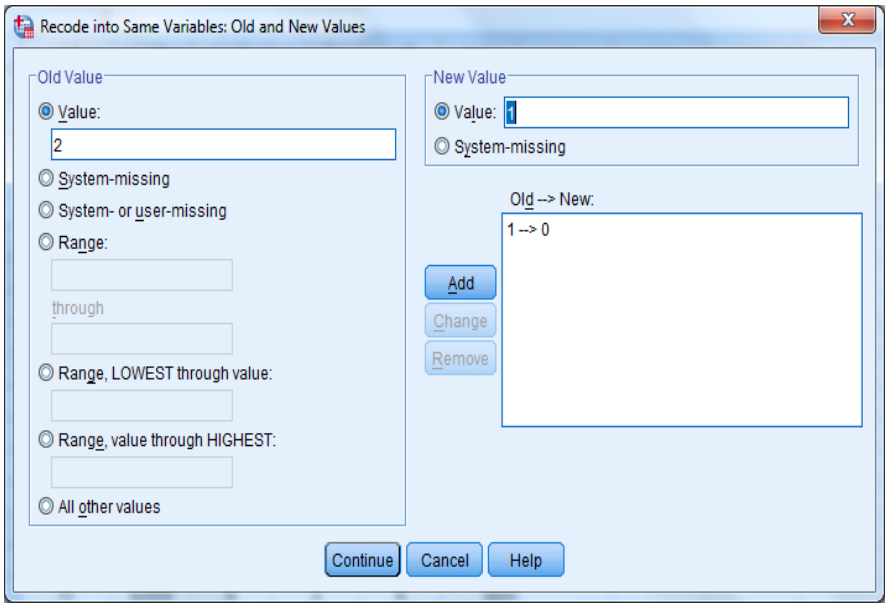
- Recode into same variable : يقوم بإعادة ترميز قيم المتغير محل المتغير الموجود دون الحاجة لانشاء متغير جديد، وفي هذه الحالة يقوم البرنامج بتحويل رموز المتغير المعني بالقيم الجديدة.

- Recode into different variables : يسمح بإعادة ترميز قيم المتغير في متغير جديد يجب تحديد اسمه وعنوانة من خلال صندوق الحوار الخاص بهذا الأمر، بحيث يقوم البرنامج بتحويل قيم المتغير المعني إلى القيم الجديدة التي يحددها المستخدم مع الاحتفاظ بالقيم القديمة للمتغير أى دون المساس به.

نفرض أنه نريد تحويل متغير الجنس إلى متغير من شكل ثنائي (1 و 0) بتحويل القيمة 1 إلى 0 و2 إلى 1. لتنفيذ ذلك نختار من قائمة Transform الأمر: Recode into same variable فيظهر صندوق الحوار على الشكل التالي:



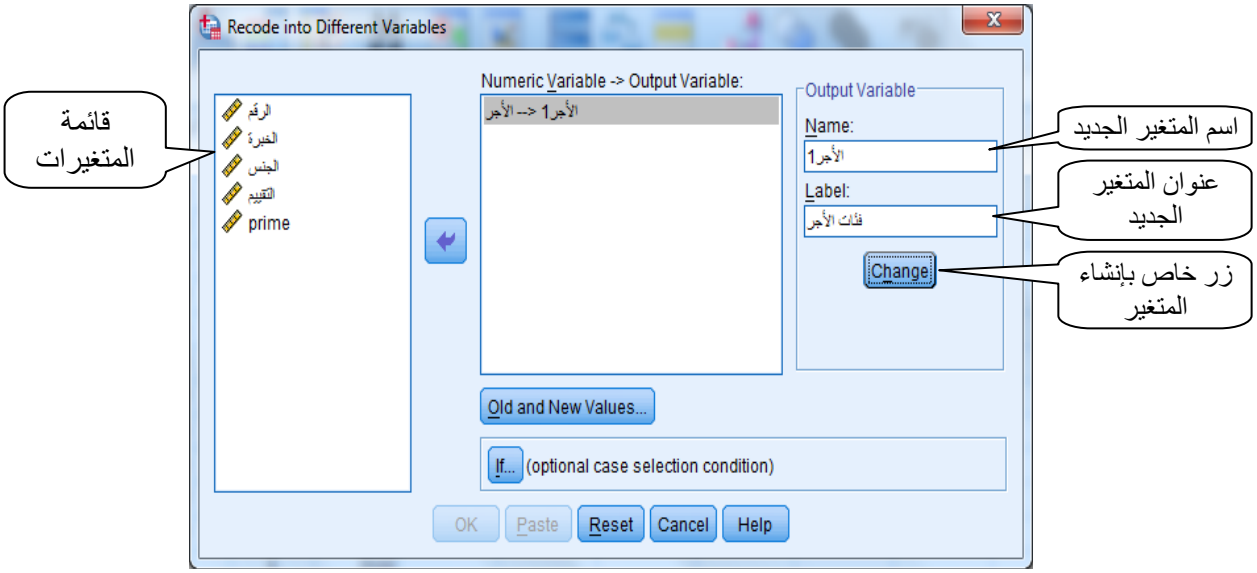
وبعد اختيار المتغير المعني بالتحويل وهو متغير الجنس بنقله من قائمة المتغيرات إلى المجال "Numeric Variable"، ننقر على الأيقونة Old and New Values لتعريف القيم القديمة والقيم الجديدة لمتغير الجنس لأن الأمر يتعلق بتحويل قيم المتغير إلى قيم جديدة وهي 0 و1، فيظهر على الشاشة صندوق حوار Old and New Values على الشكل التالي:



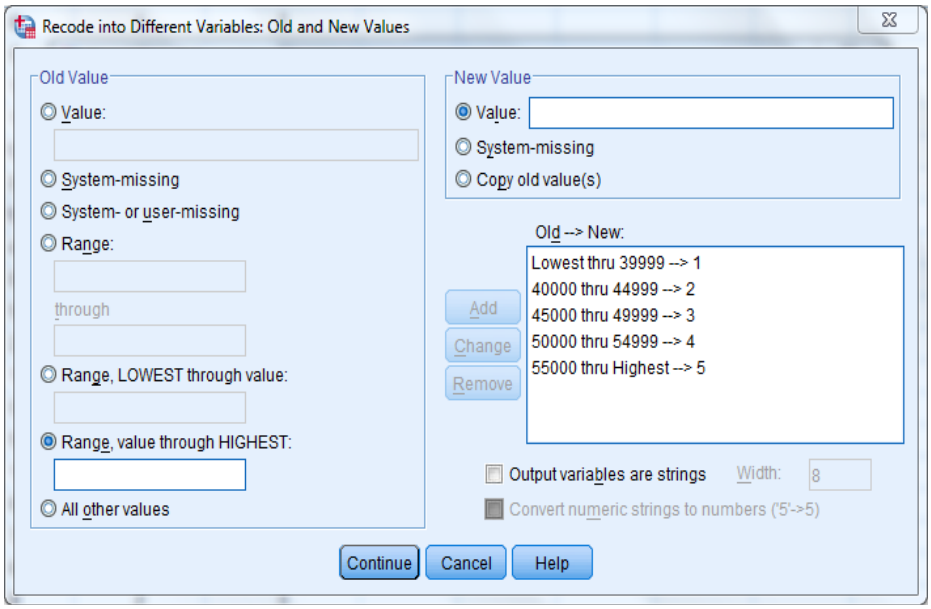
كذلك، نفرض انه نريد تبويب متغير الأجر إلى فئات كما يلي:

1: أقل من 40000؛ 2: من 40000 إلى 45000؛ 3: من 45000 إلى 50000؛ 4: من 50000 إلى 55000؛ 5: 55000 فأكثر.

لتنفيذ هذا الأمر نختار من قائمة Transform الأمر: Recode into different variable لأننا لا نريد تحويل متغير الأجر نفسه وفقدان القيم الحقيقية للأجر. فيظهر صندوق حوار على الشكل التالي:



في هذه الحالة يجب تحديد اسم المتغير الجديد وعنوانه في الخانات الخاصة بذلك في المجال Output Variable والنقر على الزر Change، ليفتح البرنامج المجال لتحديد القيم القديمة والقيم الجديدة للمتغير من خلال الزر Old and New Values بنفس الطريقة المذكورة سابقا، فيظهر لنا صندوق الحوار الخاص بهذه العملية، بحيث يتم تحديد فئات الأجر وإعطاء الرمز الخاص بكل فئة كما يظهر في الصورة التالية:



يتيح صندوق حوار Old and New Values عدة خيارات لتحديد الفئات نلخصها فيما يلي:

أولا: القيم القديمة Old Value: يتم من خلاله تحديد القيم القديمة التي تكون غالبا في شكل مجالات، بالاضافة إلى القيم المفقودة في حالة اعتبارها ضمن فئة معينة. والخيارات المتاحة هي:

- System-missing: وهي القيم المفقودة تلقائيا (عدم توفر المعلومة (.))

- System-or user-missing: وهي القيم المفقودة أو القيم التي تم تعريفها بأنها مفقودة من طرف المستخدم.

- Range: في هذه الحالة يتيح البرنامج تحديد القيمة الدنيا والقيمة العليا للفئة في الخانتين على الترتيب. كما هو الحال بالنسبة للفئات 2، 3 و4 في هذا المثال، بحيث نكتب القيمة الدنيا ثم القيمة العليا في كل حالة. ونشير هنا إلى أنه يجب تحديد االمجالات المحددة للفئات بعناية، بحيث أن القيمة الدنيا في الفئة الثانية يجب ان تكون أكبر من الفئة العليا في الفئة الأولى بمقدار 1 ان كانت الاعداد صحيحة وبمقدار 0.1 في حالة الأعداد العشرية برقم واحد بعد الفاصلة أو 0.01 في حالة رقمين بعد الفاصلة وهكذا....، وهذا تجنبا لاحتساب بعض القيم مرتين في فئتين مختلفتين وما يمكن أن ينجر عن ذلك من خلل.

- Rang, value through value: وهي الحالة التي يتم فيها تحديد الفئة التي تحتوي على قيمة معينة وأقل منها، دون تحديد القيمة الدنيا، وفي هذه الحالة يتم كتابة القيمة العليا للفئة فقط. في هذا المثال نكتب مثلا القيمة 39999 للفئة الأولى، لان هذه الفئة تشمل الأجر أقل تماما من 40000.

- Rang, value through HIGHEST: وهي الحالة التي يتم فيها تحديد الفئة التي تحتوي على قيمة معينة وأكثر منها دون تحديد القيمة العليا، وفي هذه الحالة يتم كتابة القيمة الدنيا للفئة فقط كما هو الحال بالنسبة للفئة الأخيرة في مثالنا هذا، وتكون بكتابة القيمة 55000.

- All other values: لاعتبار القيم الأخرى المتبقية والتي لا تنتمي لأي مجال من المجالات المحددة ضمن فئة أخرى (الفئة الأخيرة)، لكي لا يحتوى المتغير الجديد على قيم مفقودة، لأن البرنامج في حالة بقاء قيم لا تنتمي للفئات التي حددها الباحث يعتبرها قيم مفقودة وتظهر في شكل خانات فارغة.

ثانيا: القيم الجديدة New Value: يتم من خلاله إعطاء رموز للفئات المحددة والتي تكون في شكل رموز معبر عنها بأرقام لترتيب الفئات تنازليا أو تصاعديا حسب رغبة الباحث. ويتوفر هذا المجال على ثلاث خيارات هي:

- value: وفيه يتم كتابة القيمة العددية المعبرة عن الفئة، وفي مثالنا تكون بكتابة الأرقام 1، 2،...،5.
- System-missing: لاعتبار هذه الفئة قيم مفقودة وتظهر في شكل خانات فارغة.
- copy old value(s): لنسخ القيم القديمة أي ابقاء القيم القديمة لهذه الفئة.

ويتم تحويل متغير "الأجر" إلى متغير "الأجر1" في شكل فئات، بتحديد الفئات بطريقة متتالية بحيث يتم تحديد الفئات في Old value واعطائها الرمز المناسب في New Value، والنقر على الزر Add لاضافة الفئة بتحويل القيم القديمة إلى القيم الجديدة.

ملاحظة:

في حالة وجود خطأ في تحديد الفئات يمكن للمستخدم أن يصححه من خلال النقر على الفئة المعنية في المجال Old-->New وتصحيح الخطأ الوارد في تحديد الفئة أو ترميزها والنقر على الزر Change عوض Add فيتم تصحيح الخطأ، أو عن طريق حذف الفئة بالنقر عليها في المجال Old-->New و النقر على Remove، وإعادة تحديد الفئة من جديد.

كما يتيح البرنامج امكانية اعتبار المتغير الجديد متغيرا اسميا بالتأشير على المربع الصغير الذي يظهر اسفل صندوق الحوار بعنوان Output Variable as string وتحديد عرضه في With دون الحاجة للعودة إلى شاشة Varible View لتعريف المتغير الجديد.

وبالنقر على الزر Ok يقوم البرنامج بانشاء المتغير الجديد.

الفصل الرابع: التحليل الاحصائي للبيانات باستعمال برنامج SPSS

نتناول من خلال هذا الفصل، طرق استعمال برنامج Spss في التحليل الاحصائي للبيانات الإحصائية بحيث يشمل الجداول التكرارية والتمثيلات البيانية الملائمة حسب طبيعة البيانات المراد تحليلها، بالإضافة إلى طرق استخراج الجداول التقاطعية والاختبارات الاحصائية الموافقة حسب الحاجة.

يتوفر برنامج SPSS على مجموعة من الأوامر التي من خلالها يتم اجراء التحليل الاحصائي للبيانات بمختلف انواعها، بحيث يوفر كل أمر من هذه الأوامر امكانية الحصول على مختلف المقاييس الاحصائية والجداول والمخططات البيانية حسب طبيعة التحليل الذي يرغب المستخدم في اجرائه. وهذا على خلاف برنامج EXCEL الذي يوفر ذلك لكن يستوجب على الباحث في كل مرة كتابة دوال وعلاقات احصائية، الأمر الذي يتطلب الكثير من الوقت.

1. بعض المبادئ الأساسية للتحليل الوصفي للبيانات:

يتضمن التحليل الاحصائى الوصفى للبيانات تنظيم البيانات ووصفها بطريقة تجعلها اكثر وضوحا للمستخدم. وتتحدد الطريقة المثلى لوصف البيانات حسب نوعية البيانات ووحدة القياس الخاص بها (صفات البيانات)، ويمكن تلخيص طرق العرض الاحصائى فى ما يلي:

1-1. العرض الجدولي Tabulation: وهو وضع البيانات فى جدول احصائى ويختلف شكل الجدول الاحصائى حسب نوع البيانات وصفية أو كمية متقطعة أو مستمرة.

1-2. المخططات البيانية Graphical Representation: وهو التعبير عن البيانات بشكل بيانى يتلائم مع طبيعة البيانات. مثلا:

أ. البيانات التى لها وحدة قياس اسمية او ترتيبية ترسم باستخدام الأعمدة البيانيةBar Chart أو شكل الدائرة Pie Chart لكن بشرط ان تكون تقسيمات المتغير اقل من 10 تقسيمات.

ب. البيانات المستمرة (وحدة قياسها مجال او نسبة) الموضوعة فى جدول تكرارى بعد الترميز ترسم بالمدرج التكراري Histogram والمنحنى التكراري Frequency Curve، كما يمكن استخدام المدرج التكراري ايضا فى حالة البيانات الكمية المتقطعة.

ج.يمكن تمثيل البيانات المستمرة ايضا باستخدام شكل الصندوق Box Plot والتى تعتمد على الوسيط والربيعيات.

د. يمكن تمثيل البيانات الكمية (متقطعه أو مستمرة) باستخدام شكل الساق والأوراق Stem and leaf.

1-3. حساب بعض المقاييس الاحصائية Statistical Measures:

يمكن التعبير عن البيانات الاحصائية بحساب مقياس احصائى أو اكثر لها وتوجد عدة مقاييس احصائية هى:

1- مقاييس النزعة المركزية Measures of central Tendency: يعرف مقياس النزعة المركزيه للبيانات بأنه الرقم الذى تتمركز حوله البيانات. وتتوفر على عدة مقاييس منها:

أ. الوسط الحسابى Arithmetic Mean: ويصلح للبيانات الكميه فقط وهو وحيد ولكنه يتأثر بالقيم الشاذة.

ب. الوسيط Median: وهى القيمة التى تقع فى منتصف البيانات بعد الترتيب وهو وحيد ولا يتأثر بالقيم الشاذة.

جـ . المنوال Mode: وهى القيمة الأكثر شيوعا ولا يتأثر بالقيم الشاذة لكنه قد لا يكون وحيدا او قد لا يكون موجودا (لايمكن حسابه).

ملاحظة: مقاييس النزعة المركزيه كلها مقاييس مطلقة اي لها نفس مميزات وخصائص البيانات الأصلية وبذلك لا تصلح للمقارنة بين مجموعتين أو اكثر اذا اختلفت وحدة القياس.

2. مقاييس التشتت المطلق والنسبى: وتقيس مدى البعد بين مفردات المتغير أى أنها تقيس مدى التباعد والتنافر بين قيم المتغير أو قد تقيس مدى بعد البيانات عن مقياس النزعة المركزية الخاص بها وتوجد مقاييس مطلقة للتشتت ومقاييس نسبيه ومن مقاييس التشتت الأكثر استعمالا نذكر ما يلي:

أ. المدى Range: وهو البعد بين اكبر واقل قيمة وهو يتأثر بالقيم الشاذة

ب. الانحراف المعيارى Standard deviation: وهو الجذر التربيعى الموجب للتباين وهو وحيد ويقيس تشتت البيانات عن الوسط الحسابي للبيانات.

جـ. معامل الاختلاف Coefficient of Variation: وهو من المقاييس النسبيه للتشتت وهو من احسن مقاييس التشتت النسبية. ويستخدم معامل الاختلاف احيانا للتعرف على القيم الشاذه للبيانات.

د. مقياس الالتواء Measure of Skewness: وهو مقياس تشتت نسبى يحدد هل البيانات متماثلة أو ملتوية.

هـ. مقياس التفرطح Measure of Kurtosis: وهو مقياس نسبى يقيس ارتفاع قمة المنحنى.

ملاحظة: إذا كانت قيمة مقياس التشتت كبيرة فهذا يعنى وجود تشتت واختلاف كبير ويطلق عليه احيانا بمصطلح عدم التجانس في العينة. كما أن مقاييس التشتت المطلقة تتميز بقيمها الموجبه (لا يمكن أن تكون سالبة) لكن المقاييس النسبيه قد تكون موجبه أو سالبه. كما يعتبر في برنامج SPSS المقاييس (د) و(هـ) من مقاييس التوزيع كونها تتعلق بالتوزيع التكراري للبيانات .

2. التحليل الاحصائي الوصفي باستخدام البرنامج SPSS:

نتناول في هذا الصدد الأوامر الخاصة بالتحليل الوصفي للبيانات الاحصائية التي يتوفر عليها البرنامج SPSS من خلال القائمة Analyze وهي القائمة الخاصة بمختلف العمليات الاحصائية، بحيث نجد الأمر Descriptive Statistics الذي يحتوي بدوره على مجموعة من الأوامر التي تقوم باجراءات التحليل الوصفي للبيانات.

2-1. الأمر Frequencies:

يستخدم هذا الأمر عموما في التحليل الوصفي لمختلف البيانات الاحصائية، لكنه يستخدم على وجه الخصوص للمتغيرات الإسمية او الترتيبية التي تتكون من مجموعة محدودة من الفئات، لغرض وصف متغير أو عدة متغيرات من خلال الجداول التكرارية وبعض المخططات البيانية الخاصة بذلك، بالإضافة إلى حساب مجموعة من المؤشرات الاحصائية الوصفية وهي تتمثل في مقاييس النزعة المركزية (الوسط الحسابي، الوسيط، المنوال...) ومقاييس التشتت (الانحراف المعياري، التباين،....)، إلى جانب مقاييس لتحديد شكل التوزيع (معامل التفرطح والالتواء) وبعض المؤشرات المرتبطة بالرتب كالمئينات والربيعيات. بالرغم من أن هذه المؤشرات لا يمكن استعمالها في حالة المتغيرات النوعية ذات الطابع الاسمي، لكن تم تزويد هذا الأمر بها لاستعماله في حالات المتغيرات الترتيبية والكمية على حد سواء.

كما يمكن استعمال الأمر Frequencies كذلك لإعداد بعض الرسومات البيانية المرتبطة بالتحليل الوصفي كالأعمدة التكرارية والدوائر النسبية بالنسبة للمتغيرات النوعية ذات الطابع الاسمي أو الترتيبي، والمدرج التكراري الذي يستخدم فقط في حالة المتغيرات الكمية.

لتوضيح طريقة استخدام الأمر Frequencies في برنامج SPSS نستعين بالمثال التالي:

مثال (4-1):

لدينا جزء من البيانات الاحصائية الخاصة بنتائج استبيان حول استخدام الحافلة في احدى المدن، الاسئلة المطروحة في هذا الاستبيان كانت على الشكل التالي[7]:

1. الجنس: ذكر ☐ أنثى ☐

2. العمر أو السن: 1. 18 - 34 سنة ☐ 2. 35 - 54 سنة ☐ 3. 55 سنة فأكثر ☐

3. خلال الـ 12 شهرا الماضية كيف كان استعمالك للحافلة؟

1. دائما ☐ 2. أحيانا ☐ 3. لم استعمل الحافلة ☐

[7] Gérald Baillargeon, Fernando ouellet, Op cit. p 36.

والجدول التالي يبين نتائج الاجابات على هذه الاسئلة المحصلة لعينة من 75 شخص:

N	Q1	Q2	Q3	N	Q1	Q2	Q3	N	Q1	Q2	Q3
1	ذكر	34-18 سنة	دائما	26	أنثى	55 سنة فأكثر	لم استعمل الحافلة	51	أنثى	54-35 سنة	لم استعمل الحافلة
2	أنثى	55 سنة فأكثر	دائما	27	ذكر	55 سنة فأكثر	لم استعمل الحافلة	52	ذكر	54-35 سنة	لم استعمل الحافلة
3	أنثى	34-18 سنة	دائما	28	ذكر	55 سنة فأكثر	أحيانا	53	ذكر	54-35 سنة	أحيانا
4	أنثى	34-18 سنة	لم استعمل الحافلة	29	ذكر	55 سنة فأكثر	لم استعمل الحافلة	54	أنثى	54-35 سنة	لم استعمل الحافلة
5	ذكر	54-35 سنة	أحيانا	30	ذكر	34-18 سنة	لم استعمل الحافلة	55	ذكر	34-18 سنة	لم استعمل الحافلة
6	أنثى	34-18 سنة	لم استعمل الحافلة	31	أنثى	34-18 سنة	أحيانا	56	أنثى	55 سنة فأكثر	لم استعمل الحافلة
7	أنثى	34-18 سنة	لم استعمل الحافلة	32	ذكر	54-35 سنة	دائما	57	أنثى	54-35 سنة	لم استعمل الحافلة
8	أنثى	55 سنة فأكثر	أحيانا	33	أنثى	34-18 سنة	لم استعمل الحافلة	58	ذكر	34-18 سنة	أحيانا
9	ذكر	34-18 سنة	دائما	34	أنثى	34-18 سنة	أحيانا	59	أنثى	34-18 سنة	لم استعمل الحافلة
10	ذكر	54-35 سنة	لم استعمل الحافلة	35	أنثى	34-18 سنة	أحيانا	60	أنثى	54-35 سنة	لم استعمل الحافلة
11	ذكر	54-35 سنة	أحيانا	36	أنثى	54-35 سنة	لم استعمل الحافلة	61	ذكر	54-35 سنة	لم استعمل الحافلة
12	أنثى	55 سنة فأكثر	دائما	37	ذكر	55 سنة فأكثر	لم استعمل الحافلة	62	ذكر	54-35 سنة	أحيانا
13	أنثى	54-35 سنة	لم استعمل الحافلة	38	أنثى	34-18 سنة	لم استعمل الحافلة	63	ذكر	34-18 سنة	لم استعمل الحافلة
14	ذكر	54-35 سنة	لم استعمل الحافلة	39	أنثى	54-35 سنة	أحيانا	64	أنثى	54-35 سنة	لم استعمل الحافلة
15	ذكر	34-18 سنة	أحيانا	40	أنثى	54-35 سنة	لم استعمل الحافلة	65	أنثى	54-35 سنة	لم استعمل الحافلة
16	ذكر	34-18 سنة	لم استعمل الحافلة	41	ذكر	34-18 سنة	لم استعمل الحافلة	66	أنثى	34-18 سنة	لم استعمل الحافلة
17	أنثى	54-35 سنة	لم استعمل الحافلة	42	أنثى	34-18 سنة	لم استعمل الحافلة	67	ذكر	55 سنة فأكثر	أحيانا
18	أنثى	34-18 سنة	دائما	43	أنثى	34-18 سنة	لم استعمل الحافلة	68	أنثى	54-35 سنة	لم استعمل الحافلة
19	ذكر	54-35 سنة	لم استعمل الحافلة	44	أنثى	54-35 سنة	لم استعمل الحافلة	69	ذكر	34-18 سنة	لم استعمل الحافلة
20	ذكر	54-35 سنة	لم استعمل الحافلة	45	ذكر	55 سنة فأكثر	أحيانا	70	ذكر	34-18 سنة	لم استعمل الحافلة
21	ذكر	54-35 سنة	أحيانا	46	ذكر	55 سنة فأكثر	لم استعمل الحافلة	71	أنثى	54-35 سنة	أحيانا
22	أنثى	34-18 سنة	لم استعمل الحافلة	47	أنثى	34-18 سنة	أحيانا	72	أنثى	55 سنة فأكثر	لم استعمل الحافلة
23	أنثى	55 سنة فأكثر	لم استعمل الحافلة	48	أنثى	55 سنة فأكثر	لم استعمل الحافلة	73	أنثى	54-35 سنة	لم استعمل الحافلة
24	ذكر	34-18 سنة	أحيانا	49	أنثى	54-35 سنة	لم استعمل الحافلة	74	ذكر	34-18 سنة	لم استعمل الحافلة
25	أنثى	سنة 34-18	لم استعمل الحافلة	50	ذكر	سنة 54-35	لم استعمل الحافلة	75	ذكر	54-35 سنة	لم استعمل الحافلة

المصدر: Gérald Baillargeon, Fernando ouellet, Op cit. p 36.

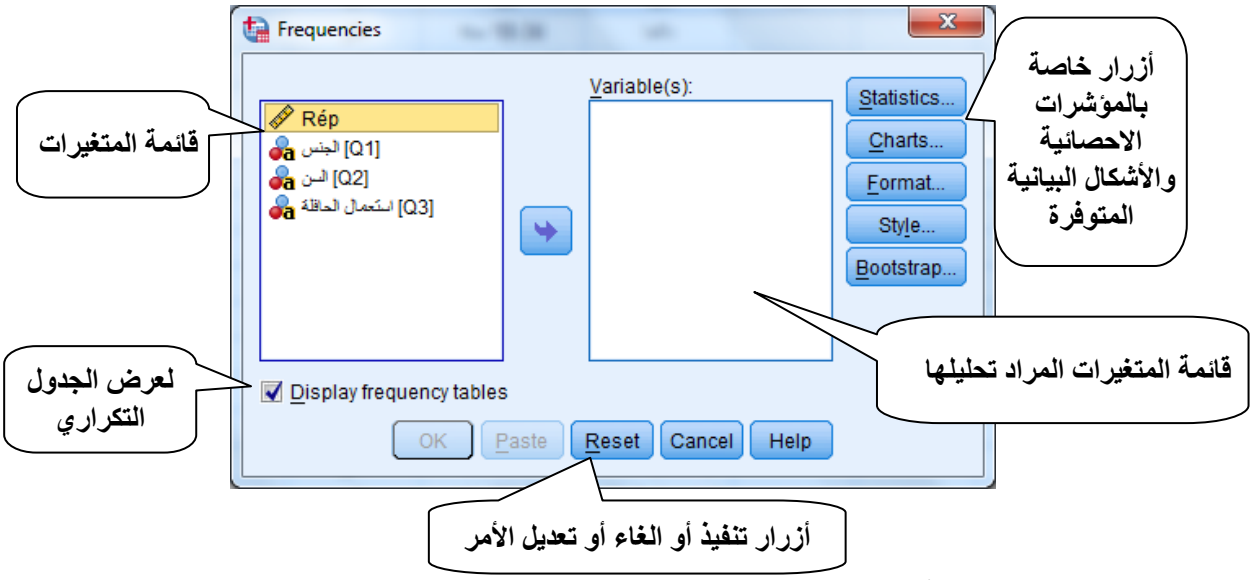
المطلوب:

تحليل بيانات الجدول باستخدام الأمر Frequencies.

ملاحظة:

يمكن تعبئة هذه البيانات في برنامج كما هي موضحة في الجدول، أو بإعطاء رموز للإجابات كما هي مبينة في الأسئلة وإعطاء عنوان لكل قيمة.

بعد تعبئة البيانات في برنامج SPSS، وإعطاء عناوين المتغيرات بحيث يكون لكل متغير عنوانه على الشكل التالي: Q1. الجنس، Q2. السن وQ3. استعمال الحافلة، من اجل اظهارها في المخرجات عوض اسمها. نختار من قائمة Analyze-> Descriptive Statistics-> Frequencies فيظهر صندوق حوار على الشكل التالي:

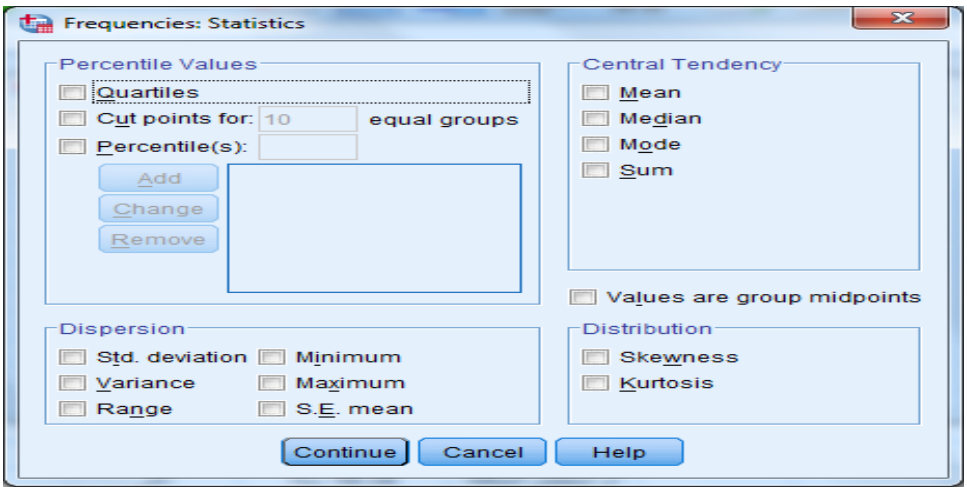


يتم اختيار المتغير أو المتغيرات المراد تحليلها بواسطة هذا الأمر ونقلها إلى المجال (Varible(s، ثم النقر على الأزرار الموجودة في اليمين من أجل تحديد المؤشرات الاحصائية ...Statistics والأشكال البيانية المراد إعدادها...Charts ، بحيث:

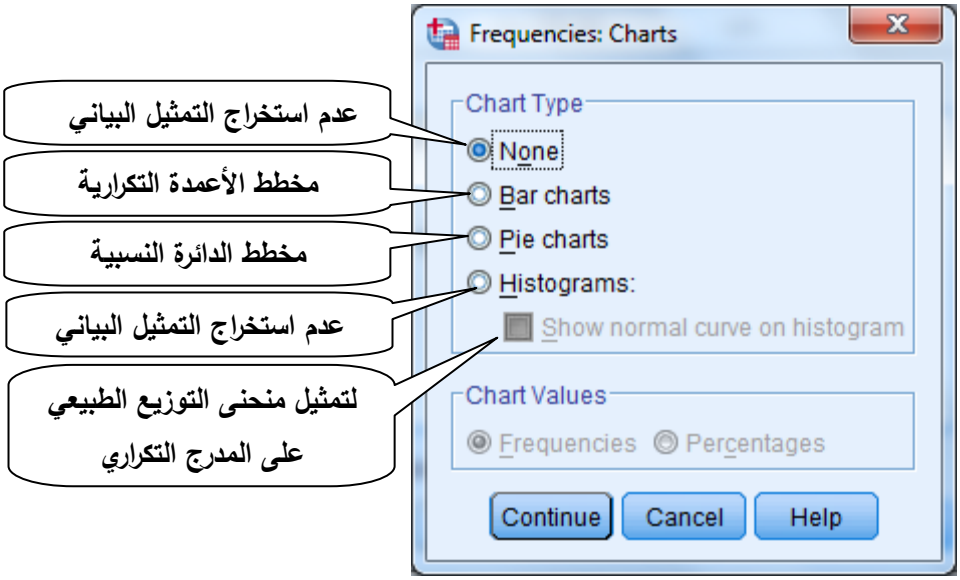
- **الزر** Statistics: يتوفر على مجموعة من المؤشرات الاحصائية التي تستعمل معظمها في البيانات الكمية وهي تتمثل في:

1. مؤشرات مرتبطة بالرتب Percentile Values: الربيعيات والمئينات يجب تحديد رتبها من طرف المستخدم بالتأشير على الخيارات المتاحة.
2. مؤشرات النزعة المركزية Central Tedency: الوسط الحسابي Mean، الوسيط Median، المنوال Mode والمجموع Sum.
3. مقاييس التشتت Dispersion: الانحراف المعياري Std. Deviation، التباين Variance، المدى Range، القيمة الدنيا Minimum، القيمة العظمى Maximum، والانحراف المتوسط SE. Mean.
4. مقاييس مرتبطة بالتوزيع Distribution: معامل الالتواء Skeweness ومعامل التفرطح Kurtosis.

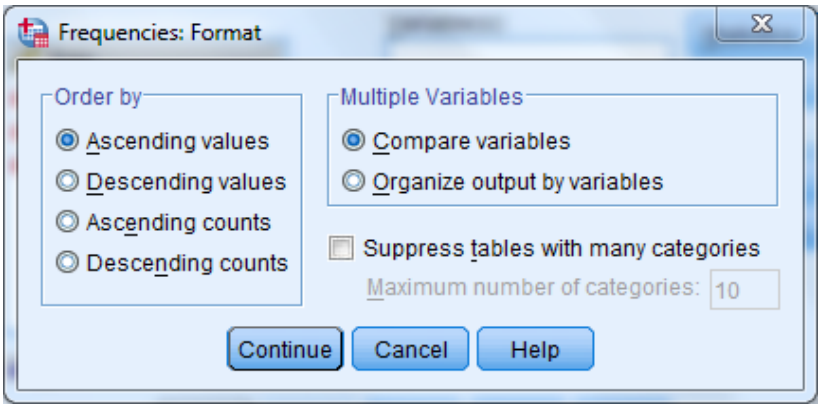
والشكل التالي يبين صندوق الحوار الخاص بالزر ...Statistics:



- **الزر Charts:** يتوفر هذا الزر على ثلاث خيارات للأشكال البيانية التي يمكن اعدادها من خلال الأمر Frequencies وهي موضحة في الشكل التالي:



- **الزر Format:** يفيدنا هذا الزر في اختيار طريقة عرض مخرجات العملية الاحصائية باستخدام الأمر Frequencies، بحيث يسمح البرنامج بترتيب البيانات في الجدول التكراري تصاعديا أو تنازليا حسب قيمة المتغير (ويتعلق الأمر بالقيم التي يأخذها المتغير أو عنوانها) من خلال المجال Order by، بحيث يتم ترتيبها ترتيبا ابجديا في حالة عناوين القيم وحسب القيم في حالة عدم وجود عناوين القيم، أو حسب القيم الاحصائية كالتكرارات مثلا، بحيث يتم ترتيب الجدول حسب قيم التكرارات تصاعديا أو تنازليا. كما يسمح هذا الزر بتحديد طريقة عرض نتائج العملية الاحصائية (في حالة تحليل عدة متغيرات في آن واحد) في جداول من أجل المقارنة أو في شكل مستقل حسب المتغيرات، من خلال Multiple Variables. أما الخيار Suppress tabeles with many categories فيفيد في استبعاد الجداول التكرارية للمتغيرات التي يفوق عدد الفئات فيها عدد معين يحدده المستخدم في المجال Maximum number of categories. والشكل التالي يمثل صندوق حوار Format.



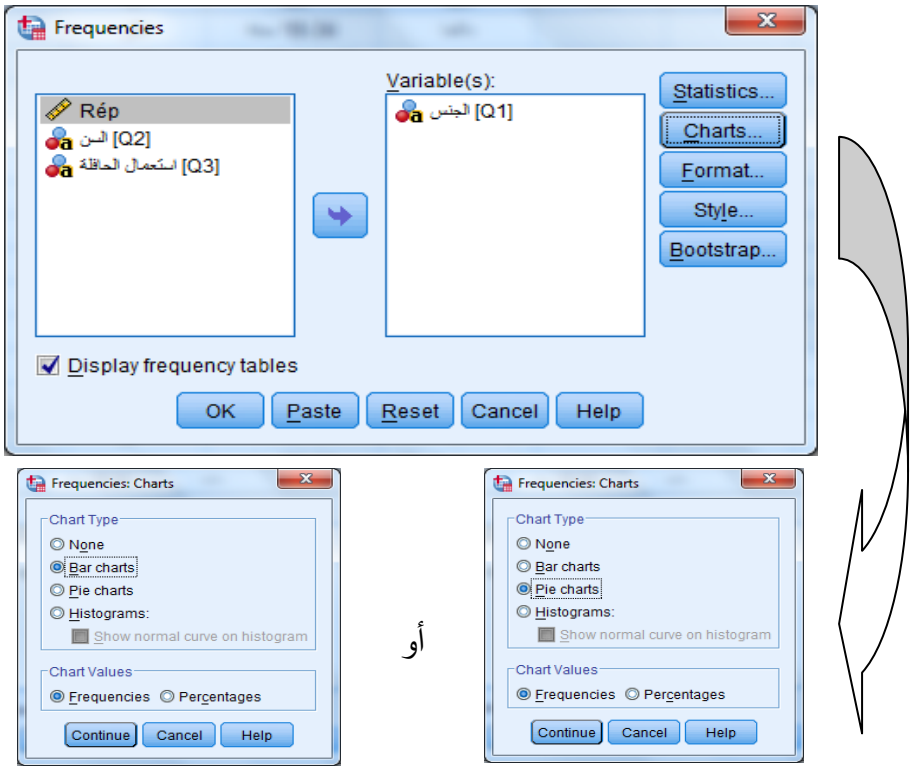
➢ تحليل بيانات الاستبيان باستخدام الأمر Frequencies:

بما أن البيانات المتعلقة بالاستبيان الخاص باستخدام الحافلة، عبارة عن متغيرات اسمية، فإن تحليلها باستخدام برنامج SPSS تكون بواسطة الأمر Frequencies، من خلال استخراج الجداول التكرارية الخاصة والأشكال البيانية على شكل أعمدة تكرارية أو دوائر نسبية الخاصة بكل متغيرة.

1- متغير الجنس (Q1):

يكون اجراء التحليل باختيار الأمر Frequencies من قائمة Analyze ونقل المتغير Q1 إلى المجال Variable(s)، واختيار شكل المخططات البيانية الملائمة باستخدام الزر Charts، وفي هذه الحالة يمكن اختيار المخطط Bar charts أو Pie charts لأن المتغير المعني اسمي. بينما لا نختار المؤشرات الاحصائية لأن المتغير اسمي لا يمكن حساب المؤشرات الاحصائية الخاصة به.

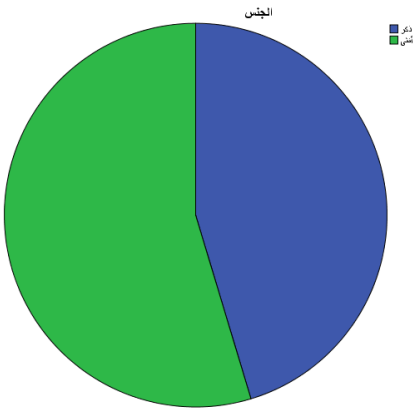
وتكون خطوات تحليل المتغير Q1 كما توضحه الصورة التالية:

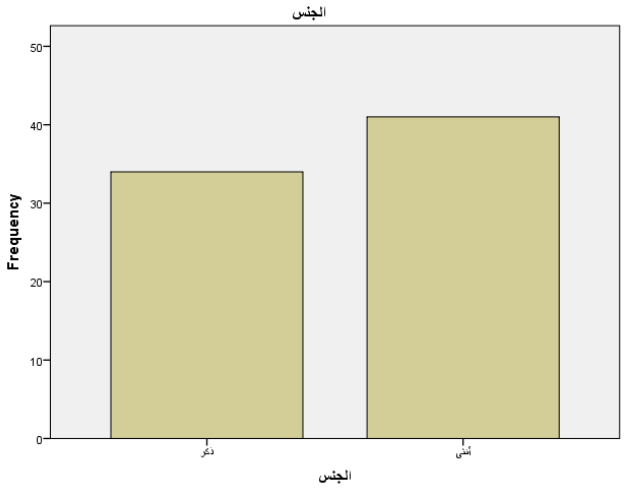


ثم ننقر على الزر Ok فيظهر لدينا الجدول التكرار والتمثيل البياني حسب الخيار على الشكل التالي:

الجنس

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ذكر	34	45,3	45,3	45,3
	أنثى	41	54,7	54,7	100,0
	Total	75	100,0	100,0	

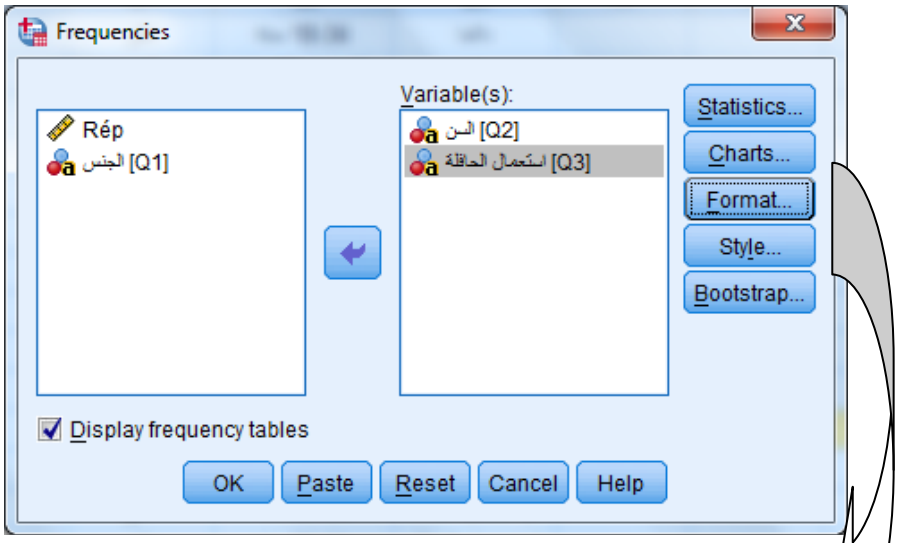


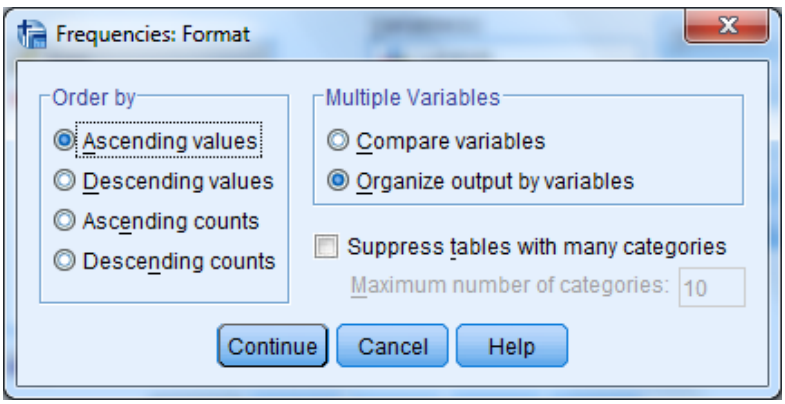


ومن خلال الجدول التكراري يظهر أن عدد الأفراد من الجنس " ذكر" المستجوبين يساوي 34 وهو يمثل 45.3% من العينة بينما كان عدد الاناث فيقدر بـ 41 وهو يمثل 54.7% من العينة.

المتغيرين Q2 و Q3 السن واستعمال الحافلة:

في هذه الحالة نقوم باجراء الأمر Frequencies على المتغيرين في آن واحد لكن العرض في جداول مستقلة لأن المتغرتين مستقلتين تماما عن بعضها. والصور التالية توضح تطبيق الأمر:



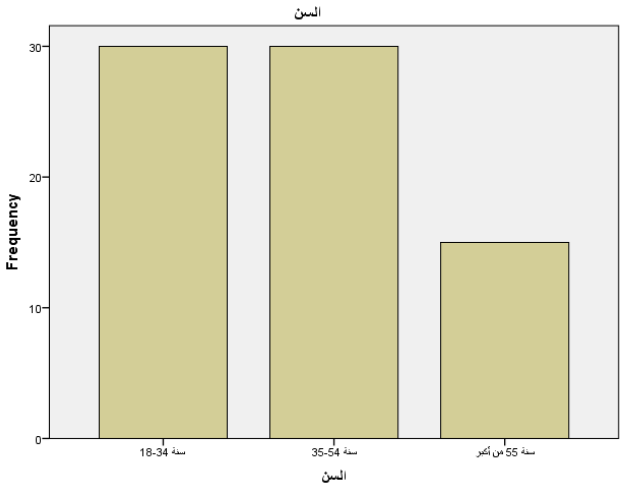


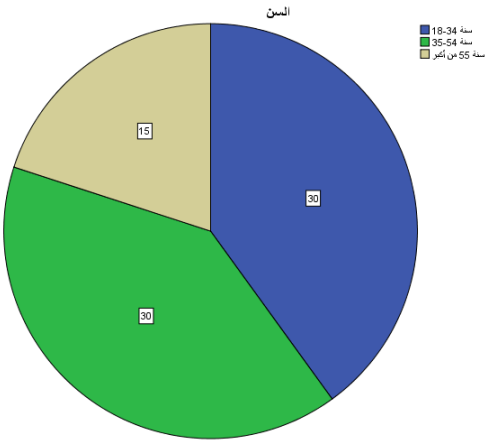
بحيث يتم اختيار Organize output by variables لعرض نتائج العملية في شكل مستقل حسب المتغيرات. ويتم اختيار الأشكال البيانية بالطريقة المذكورة أعلاه.

ـ متغير السن:

السن

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	18-34سنة	30	40,0	40,0	40,0
	35-54سنة	30	40,0	40,0	80,0
	أكبر من 55 سنة	15	20,0	20,0	100,0
	Total	75	100,0	100,0	

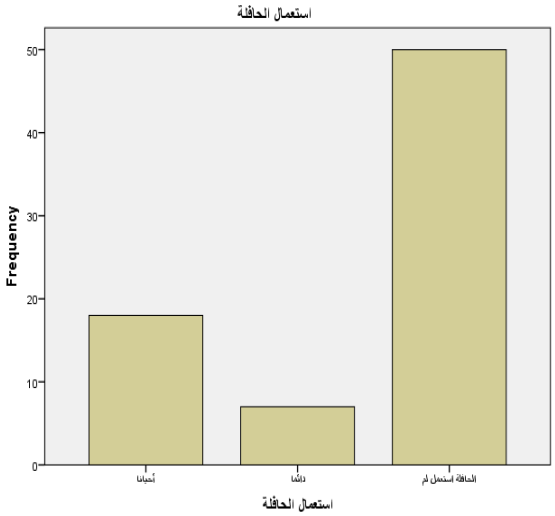


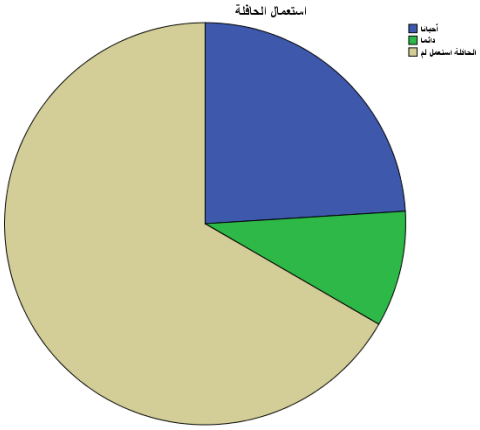


– متغير استعمال الحافلة:

استعمال الحافلة

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	أحيانا	18	24,0	24,0	24,0
	دائما	7	9,3	9,3	33,3
	لم استعمل الحافلة	50	66,7	66,7	100,0
	Total	75	100,0	100,0	





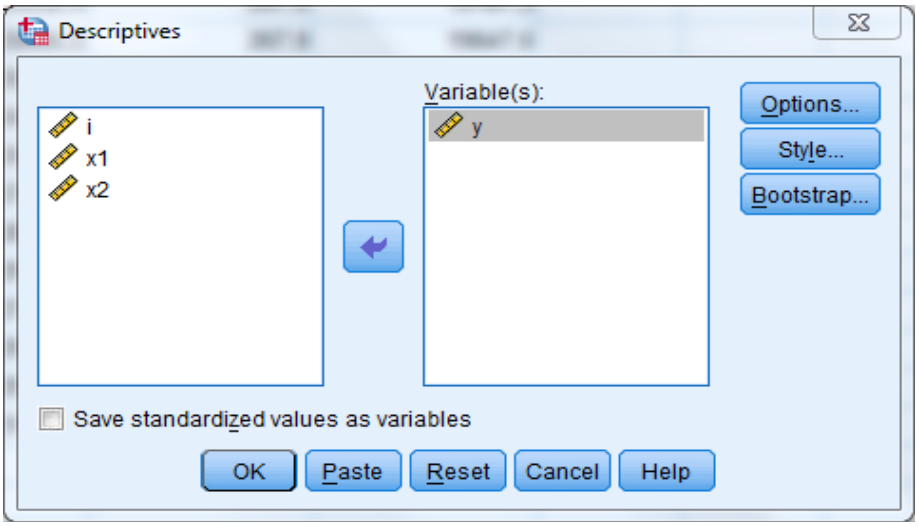
2-2. الأمر descriptives:

يفيدنا الأمر Descriptives في تحليل البيانات الكمية دون غيرها من المتغيرات بحساب مختلف المؤشرات الاحصائية الوصفية من مقاييس النزعة المركزية، التشتت والتوزيع حسب حاجة الباحث، بحيث يسمح باستخراج مختلف المؤشرات لمتغير واحد أو عدة متغيرات في جدول واحد مع امكانية المستخدم اختيار المؤشرات من خلال صندوق الحور Descriptives. ولتوضيح ذلك نستعين بالمثال التالي:

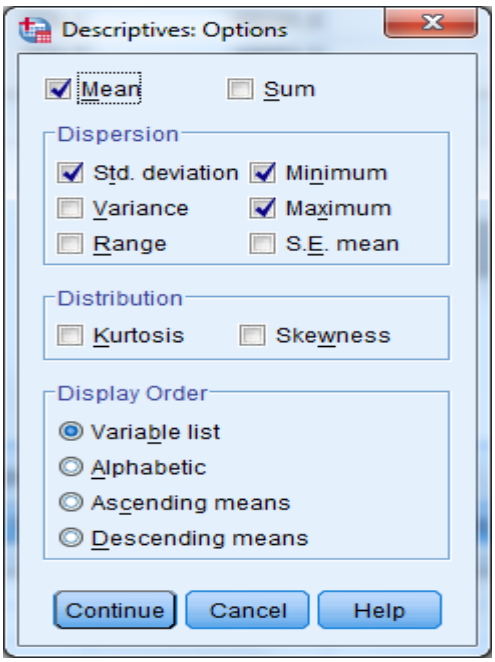
مثال (4-2): لدينا البيانات الخاصة بثلاث متغيرات كما يوضحها الجدول التالي:

i	y	x1	x2
1	16607,7	275,5	17803,7
2	17511,3	274,4	18096,8
3	20171,2	269,7	18271,8
4	20932,9	267	19167,3
5	20406	267,8	19647,6
6	20831,6	275	20803,5
7	24806,3	283	22076,6
8	26465,8	300,7	23445,2
9	27403	307,5	24939
10	28628,7	303,7	26713,7
11	29904,5	304,7	29957,8
12	27508,2	298,6	31585,9
13	29035,5	295,5	33474,5
14	29281,5	299	34821,8
15	31535,8	288,1	41794,3

لاستخراج المؤشرات الاحصائية الوصفية باستخدام البرنامج SPSS، نختار الأمر Descriptives من القائمة Descriptive Statistics من قائمة Analyze، فيظهر صندوق الحوار Descriptives على الشكل التالي:



بعد اختيار المتغيرات المعنية ونقلها إلى المجال Variable(s) وفي هذه الحالة نختار المتغيرات y، x1 و x2، نقوم بتحديد المؤشرات الاحصائية المراد استخراجها بالنقر على الزر Options فيظهر لدينا صندوق الحوار التالي:



بحيث نقوم بالتأشير على المؤشرات المراد استخراجها حسب الحاجة. وفي هذه الحالة يقوم البرنامج بعرض مختلف المؤشرات في جدول واحد، ويمنح البرنامج امكانية ترتيب المخرجات حسب اسم المتغيرات (أبجديا) أو حسب المتوسط الحسابي تصاعديا أو تنازليا، من خلال الخيارات المتوفرة في المجال Display list. أما الخيار Save Standarized values as variables فيسمح بحساب القيم المعيارية للمتغيرات وانشاء متغيرات خاصة بها باضافة الحرف Z لاسم المتغيرة لاستخدامها عند الحاجة، هذه القيم تحسب وفق الصياغة التالية:

$$ZX = \frac{X - \overline{X}}{S}$$

بحيث تمثل:ZX القيم المعيارية للمتغير X، X القيم الأصلية للمتغيرة، $\bar{X}$ المتوسط الحسابي للمتغير، S الانحراف المعياري للمتغير X .

ويكون جدول الاحصائيات الوصفية على الشكل التالي:

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
y	15	16607,7	31535,8	24735,333	4874,1735
x1	15	267,0	307,5	287,347	14,8066
x2	15	17803,7	41794,3	25506,633	7335,4815
Valid N (listwise)	15				

2-3. الأمر Explore:

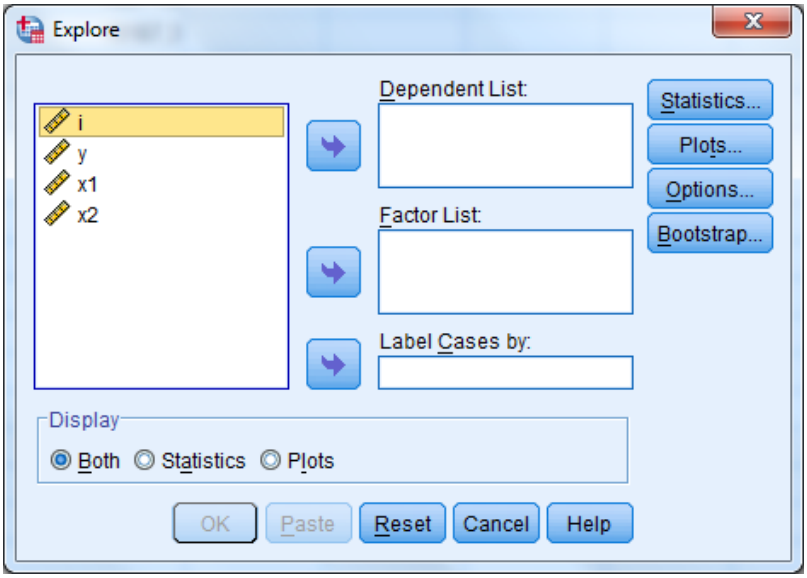
يفيد الأمر Explore في استكشاف البيانات، بحيث يسمح بتقديم ملخص احصائي للبيانات وإعداد مخططات لكل الحالات أو لمجاميع معينة، كما يسمح بإعطاء صورة للبيانات لمعرفة مدى صلاحيتها لإجراء تحليل احصائي معين، فمثلا يمكن من خلال الأمر Explore التحقق من توفر فرضية التوزيع الطبيعي للعينة أو تجانس تباين المعاملات من عدمها مما يستدعي اجراء تحويلات على البيانات مثلا.

أولا. استكشاف البيانات عن طريق المؤشرات الاحصائية والمخططات البيانية:

مثال:

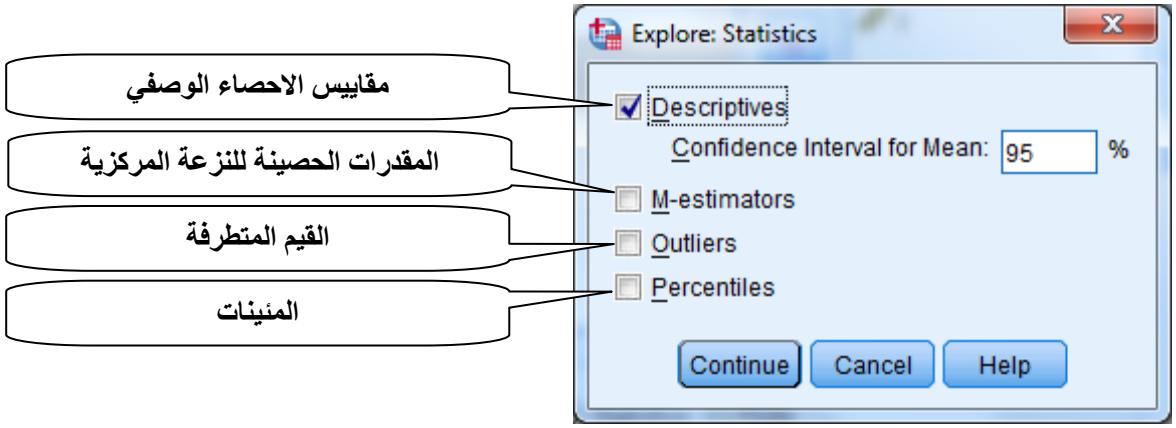
استخدام الأمر Explore لاستكشاف بيانات المثال السابق (2-4).

نختار من قائمة Aalyze؛ Descriptive Statistics؛ الأمر Explore، فيظهر صندوق الحوار الخاص باستكشاف البيانات كما يلي:



بحيث نختار المتغير المراد استكشافه ونقله إلى المجال Dependent List، واستخدام الأزرار لتحديد المؤشرات الاحصائية المراد استخراجها بحيث يتوفر الزر:

-Statistics: بالنقر على هذا الزر يظهر البرنامج صندوق حوار على الشكل التالي:



-Plots: يوفر هذا الزر امكانية اختيار المخططات البيانية للأمر Explore، وهي متمثلة في مخطط يتكون من ثلاث أجزاء هي:

أ. الصندوق Box: الجزء الأول Boxplot خاص بطريقة عرض المخططات البيانية وهو يشمل:

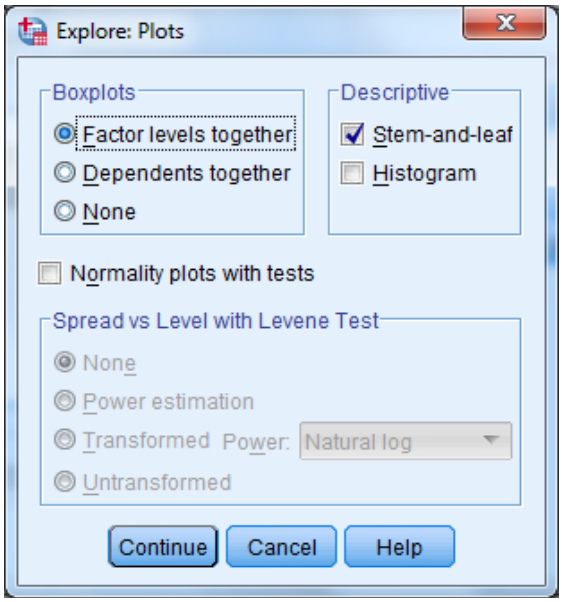
Factor Levels together: لعرض المخطط للمجاميع الجزئية لكل متغير معتمد؛

Dependent together: لعرض المخطط للمتغيرات المعتمدة داخل كل مجموعة جزئية؛

None: لعدم عرض مخطط Boxplot.

بحيث يتم اختيار الخيارين الأول أو الثاني في حالة توفر متغير التجزئة.

بحيث يظهر صندوق حوار Eplore : plots على الشكل التالي:



ويشمل هذا المخطط المكونات التالية:

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

– الربيع الثاني Q2: وهي القيمة التي تقسم البيانات إلى جزئين متساويين أي يسبقها 50% من المشاهدات عند ترتيبها تصاعديا، وهذه القيمة هي الوسيط Median.

– الربيع الثالث Q3: وهي القيمة التي تسبقها 75% من المشاهدات عند ترتيبها تصاعديا.

وتكون حافتي الصندوق عبارة عن قيمتي Q1 وQ3 أي أن طول الصندوق هو الفرق بين الربيع الثالث والربيع الأول، وهو ما يعرف بالمدى الربيعي، أما الخط الذي يظهر داخل الصندوق فهو قيمة الوسيط.

ب. الاستطالات خارج الصندوق Whiskers: وتمتد من حافتي الصندوق إلى أعلى وأدنى قيمة غير شاذة.

جـ. القيم الشاذة Outliers والقيم المتطرفة Extremes: القيم الشاذة التي تظهر في المخطط تتمثل في القيم التي تبعد عن حافتي الصندوق بطول 1.5 من طول الصندوق، والقيم المتطرفة هي القيم التي تبعد عن حافتي الصندوق بطول 3 من طول الصندوق.

ب. المخططات الوصفية Descriptive:

يسمح هذا الجزء بالاختيار بين مخططين هما:

– مخطط الساق والأوراق Stem and leaf: وفيه يتم تقسيم المتغير إلى فئات طول كل فئة 10 وحدات (درجات)، فيظهر العمود الأول التكرارات الفعلية لهذه الفئات، وفي العمود الثاني يتم تقسيم أي رقم إلى جزئين الأول هو الجذع أو الساق (الأيسر) والثاني هو الورقة (الجزء الأيمن). وهذا المخطط يشبه المدرج التكراري، والفرق بينهما هو أنه في المدرج التكراري يتم تمثيل التكرارات بأعمدة للقيم الحقيقية، بينما في الساق و الأوراق يتم تمثيلها بمعلومات عن طبيعة القيم الموجودة.

ونشير هنا إلى أن استخدام طريقة الساق والأوراق ليس سهلاً أو ليس كثير الاستخدام مقارنة بباقي المخططات التي يوفرها برنامج SPSS. إلا انه يعتمد على مبدأ بسيط جدا وهو وضع القيمة اليسرى في القيم الحقيقية للمتغيرة في الساق، والقيمة الأولى على اليمين في الورقة وتليها القيم الأولى وفق هذا الترتيب من القيم التي تتكرر في مجال تقسيم القيم الحقيقية (من نفس الدرجة)، فمثلا اذا كان لدينا الارقام التالية:

155	**127**	**122**	**117**	**77**	**73**	**37**	**17**	**15**	**13**	**12**

يمكن اعداد مخطط الساق والأوراق بتقسيم القيم الحقيقية على الشكل التالي:

الأوراق	الساق	التكرارات
2357	1	4
7	3	1
37	7	2
7	11	1
27	12	2
5	15	1

أو تقسيم هذه القيم على 100، ووضع طول العمود 100 كما يعمل برنامج SPSS، نحصل على المخطط على الشكل التالي:

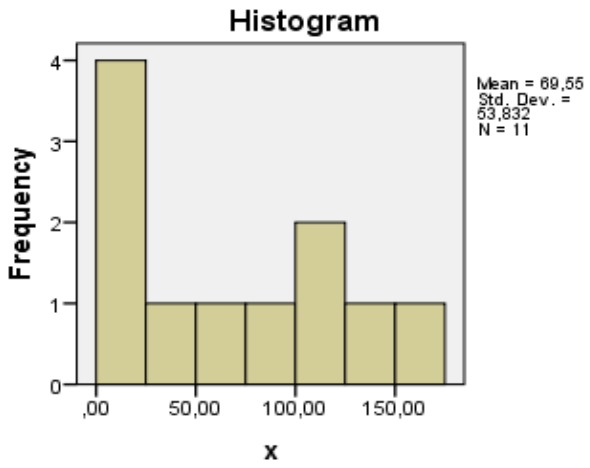
الأوراق	**الساق**	**التكرارات**
11113	**0**	**5**
77	**0**	**2**
122	**1**	**3**
5	**1**	**1**

وبتطبيق الأمر Explore على المتغيرة X التي تأخذ القيم المذكورة أعلاه ينتج لدينا مخطط المدرج التكراري ومخطط الساق والأوراق والصندوق بالإضافة إلى جدول ملخص احصائي وصفي يحتوي على مختلف المؤشرات الاحصائية.

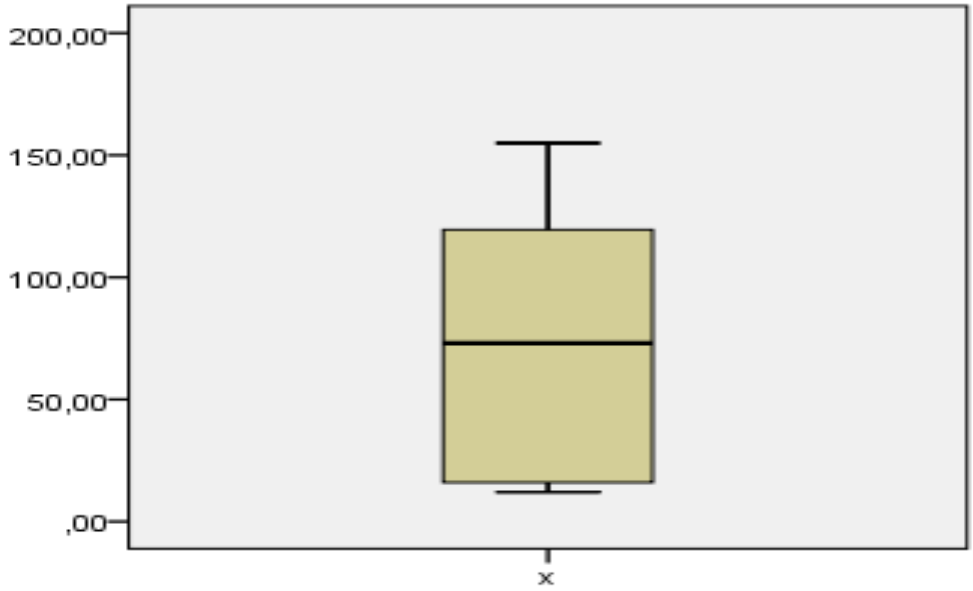
مخطط الساق والأوراق:

```
x Stem-and-Leaf Plot

 Frequency    Stem &  Leaf

     5,00        0 .  11113
     2,00        0 .  77
     3,00        1 .  122
     1,00        1 .  5

 Stem width:    100,00
 Each leaf:       1 case(s)
```

المدرج التكراري:



مخطط الصندوق



بحيث تمثل حافتي الصندوق قيمة الربيع الأول من الأسفل، وقيمة الربيع الثالث من الأعلى، بينما القيمة الدنيا فهي عبارة عن القيمة الأولى المشاهدة بعد ترتيب البيانات تصاعديا، وهي عبارة قيمة الحاصل: $Q_1 - 1.5 \times (Q_3 - Q_1)$ ، أما القيمة العظمى فهي القيمة الأخيرة المشاهدة (بعد استبعاد القيم المتطرفة والشاذة) وهي عبارة عن الحاصل: $Q_3 + 1.5 \times (Q_3 - Q_1)$. بينما تظهر القيم الشاذة والمتطرفة بعيدا عن الصندوق وتظهر في شكل نقاط انتشار بعيدة عن الصندوق (في حالتنا لا توجد قيم متطرفة). ويمكن الحكم على اعتدالية التوزيع والتوائه من خلال وضع الخط الذي يظهر في داخل الصندوق وهو قيمة الوسيط، فكلما كانت بعيدة عن وسط الصندوق كان توزيع البيانات ملتوي.

ثانيا. استكشاف البيانات عن طريق اختبار التوزيع الطبيعي وتجانس التباين:

يوفر الأمر Explore للمستخدم إلى جانب استخراج المؤشرات الاحصائية الوصفية والمخططات البيانية المذكورة سابقا، امكانية اختبار التوزيع الطبيعي للعينة وكذا اختبار تجانس التباين في حالة تعدد العينات. بحيث يتم اجراء هذه الاختبارات بطريقة بسيطة نشرحها فيما يلي:

أ. اختبار التوزيع الطبيعي للعينة:

إن عملية التحليل الاحصائي للبيانات واختبار الفروض يشترط توفر بعض الشروط في عينات الدراسة، خاصة في حالة العينات الصغيرة، ومن أهم هذه الشروط: ان تكون عينة الدراسة مسحوبة من مجتمع يتبع التوزيع الطبيعي، بينما يستغنى عن هذا الشرط في حالة تجاوز حجم العينة 30 مشاهدة ويفترض

انها تتبع التوزيع الطبيعي وفق نظريات النزعة المركزية. وشرط ثاني يتمثل في شرط تجانس التباين في حالة تعدد العينات، وكذا استقلاليتها.

فمثلا: اجراء تحليل التباين بمعيار واحد يشترط تجانس تباين العينات وان تكون مسحوبة من مجتمع يتبع التوزيع الطبيعي. وسنتناول من خلال هذا المثال طريقة اختبار التوزيع الطبيعي لعينة تتكون من أقل من 30 مشاهدة.

مثال (4-3):

نفرض انه لدينا معدلات 15 طالب من قسم الاقتصاد الكمي في كلية العلوم الاقتصادية المبينة في الجدول، والمطلوب اختبار ما اذا كانت عينة الدراسة مأخوذة من مجتمع يتبع التوزيع الطبيعي.

15	**11**	**8**	**10**	**12**
16	**14**	**3**	**11**	**15**
10	**16**	**7**	**10**	**15**

في هذه الحالة يتم اختبار الفرضية التالية:

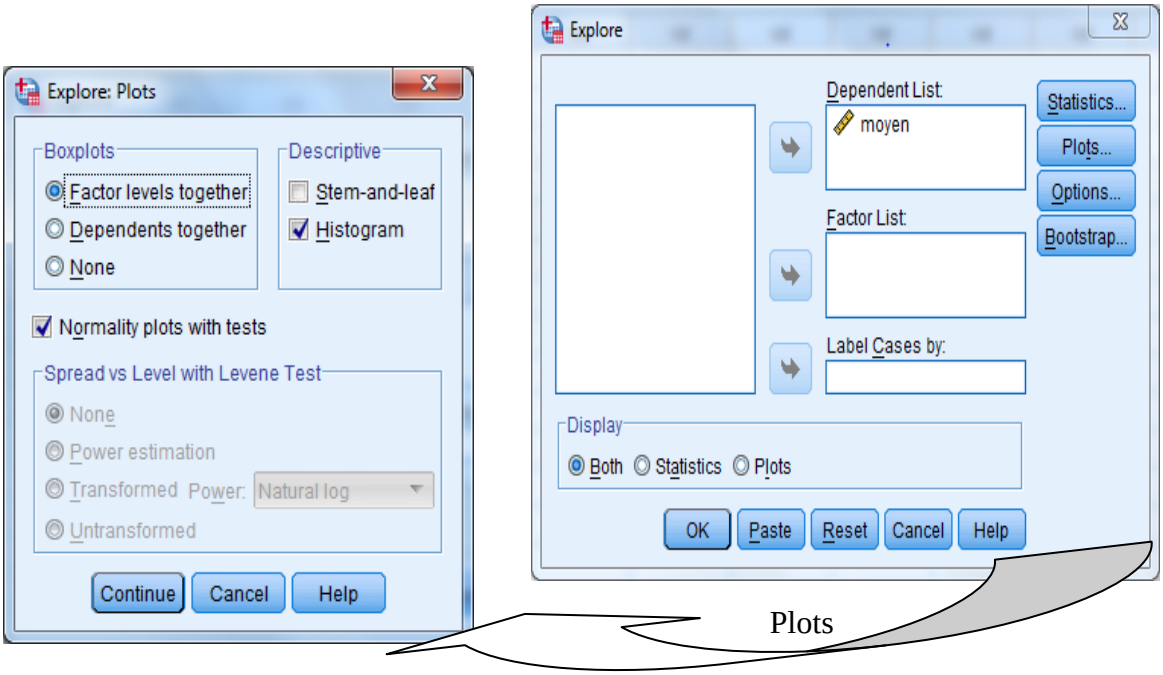
H_0: ان العينة مسحوبة من مجتمع تتبع بياناته التوزيع الطبيعي؛

H_1: ان العينة مسحوبة من مجتمع لاتتبع بياناته التوزيع الطبيعي.

ويوفر برنامج spss اختبارين أساسيين التوزيع الطبيعي وهما: اختبار كولموجوروف-سميرنوف واختبار شابيرو-ويليك، على خلاف برامج احصائية اخرى يمكن ان توفر اختبارات اخرى غير هذين الاختبارين.

ولاجراء هذا الاختبار باستعمال برنامج SPSS باتباع الخطوات التالية:

نختار من قائمة **Analyze: Descreptive statistics؛ Explore**. ثم نختار المتغير المراد اختباره في صندوق الحوار الخاص بـ **Explore** وننقله إلى المجال **Dependent List**، ومن ثم التأشير على مربع **histogram** ليظهر المدرج التكراري الموضح للتوزيع التكراري، والتأشير على مربع **Normality plots with test** لإظهار نتائج الاختبار، كما هو موضح في الصورة التالية:



أما في حالة تعدد العينات، فيجب اضافة متغير التجزئة الذي يفصل بين العينات إلى المجال Factor List. وعندئذ يمكن اختبار التوزيع الطبيعي للعينات في آن واحد دول اللجوء إلى اساليب أخرى كتحليل كل عينة على حدى باستخدام Select Cases.

بالنقر على الزر Continue يظهر لنا في المخرجات نتائج الاختبار في جدول ومخططات بيانية خاصة باختبار التوزيع الطبيعي. وتكون نتائج الاختبار الاحصائي في الجدول التالي:

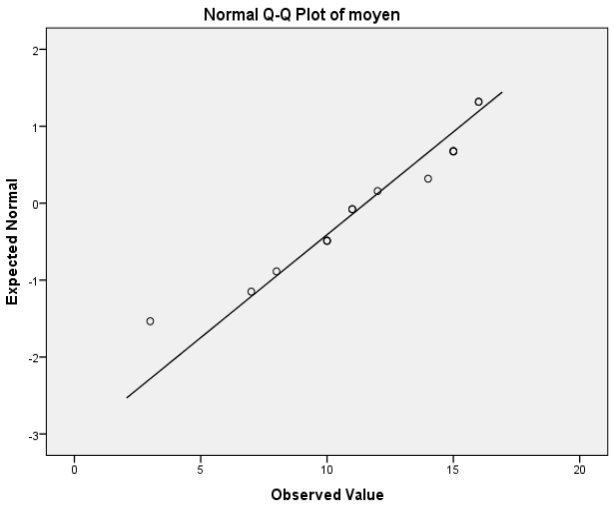
Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
moyen	0,156	15	0,200*	0,921	15	0,202

يظهر من خلال الجدول، أن القيمة الاحصائية المحسوبة لاختبار كولموغورف- سميرنوف تساوي إلى 0.156 بقيمة احتمالية تقدر بـ 0.2، ويتم اتخاذ القرار بمقارنة القيمة المحسوبة للاختبار بتلك المجدولة عند مستوى معنوية 5% بدرجة حرية (N=15) في جدول التوزيع الاحتمالي الخاص بكولموغوروف-سميرنوف، والتي تساوي 0.3376. لاحظ أن القيم المحسوبة 0.156 أقل من القيمة المجدولة k=0.3376 عند مستوى معنوية 5%، مما يقودنا الى قبول فرضية العدم H0 القائلة ان العينة مسحوبة من مجتمع يتبع التوزيع الطبيعي. كما يمكن مقارنة القيمة الاحمالية P-Value التي تظهر في الجدول "Sig=0.200" بمستوى المعنوية الذي يختاره الباحث والذي يفترض غالبا يساوي 0.05، فإذا كانت القيمة الاحتمالية المحسوبة أكبر من 0.05 يتم قبول فرضية العدم ورفض الفرضية البديلة.

كذلك الأمر بالنسبة لاختبار شابيرو–ويلك، بحيث قدرت القيمة المحسوبة للاختبار بـ 0.921 وهي أكبر من القيمة المجدولة عند مستوى 5% والمقدرة بـ W=0.881، وهذا يقودنا إلى قبول فرضية العدم القائلة ان العينة مسحوبة من مجتمع يتبع التوزيع الطبيعي، ويمكن كذلك استخدام القيمة الاحتمالية Sig=0.202 وهي أكبر من 0.05.

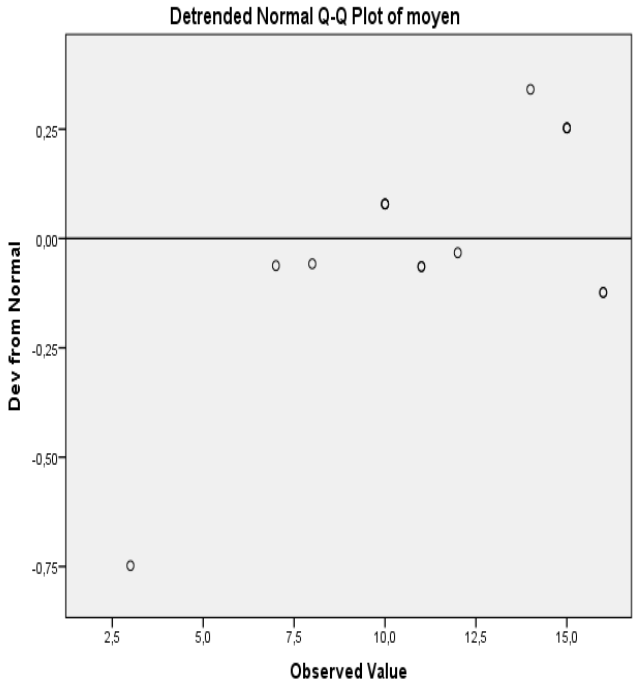
كما يظهر في ملف المخرجات مخططين بيانيين على الشكل التالي:



- مخطط Normal Q–Q plot:

في هذ المخطط يتم رسم كل مشاهدة من البيانات الأصلية على محور الفواصل (المحور الأفقي) مقابل قيم التوزيع القياسي المتوقعة لها والتي تعرف بـ Expected Zscore.

فإذا كانت نقاط التقاطع (الانتشار) قريبة من الخط المستقيم فهذا يدل على ان العينة مسحوبة من مجتمع يتبع التوزيع الطبيعي، والعكس اذا كانت بعيدة عنه.



- مخطط Detrended Normal Q–Q Plot:

يبين هذا المخطط درجة ابتعاد نقاط الانتشار عن الخط المستقيم وتبيين إلى أي مدى تكون الاتحرافات عن الخط المستقيم كافية للاقرار بأن المشاهدات الخاصة بالعينة لا تتوزع طبيعيا. بحيث يتم تمثيل انحرافات القيم المعيارية للمشاهدات عن قيم التوزيع الطبيعي المتوقعة لها. فإذا كانت معظم نقاط الانتشار تقع ضمن المجال (2،–2) يمكن الاقرار بأن العينة تتوزع طبيعيا، كما هو الحال بالنسبة لبيانات هذا المثال، حيث أن كل النقاط تقع ضمن المجال (2، –2).

ملاحظة:

يمكن استخدام متغير تجزئة لتحليل مجموعة من العينات باستخدام الأمر Explore، مثلا استخدام متغير القطاع، الشهادة المحصلة، المنطقة الجغرافية،...وغيرها، وادخاله إلى المجال Fctore List في صندوق حوار Explore. بحيث يقوم البرنامج بالتعامل مع المجاميع الجزئية عوض كامل البيانات، وهنا يمكن للباحث اختيار طريقة عرض النتائج اما من اجل المقارنة (عرض مخطط واحد يحتوى على الصندوق لكل مجموعة) أو لكل جزء مخطط خاص (عدة مخططات) حسب خيارات Boxplots.

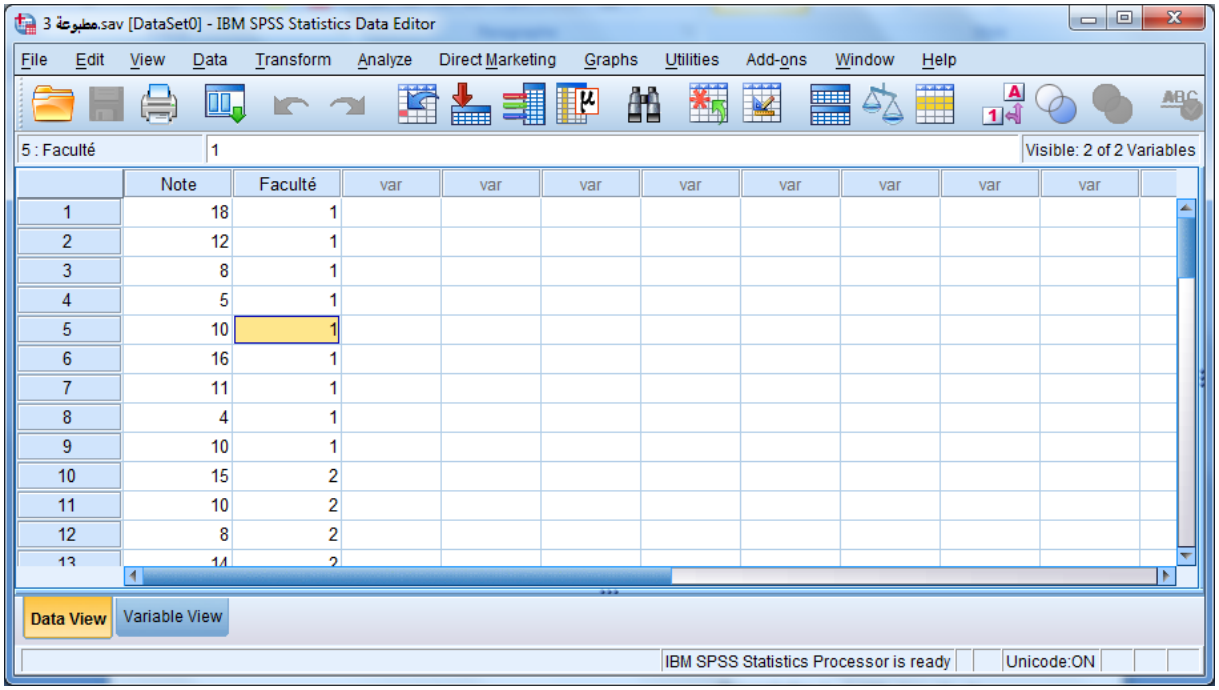
ب. <u>اختبار تجانس التباين</u>:

ان بعض العمليات الاحصائية كتحليل التباين مثلا يشترط توفر فرضية تجانس تباينات العينات، وعدم توفر هذه الفرضية يؤدي حتما إلى قرارات خاطئة عند اختبار تأثير المعاملات أو تساوي متوسطات العينات. وعليه يجب على الباحث اختبار فرضية تساوي التباين لعينات الدراسة من عدمها قبل اجراء العملية الاحصائية، ويوفر برنامج SPSS من خلال الأمر Explore امكانية اختبار هذه الفرضية باختبار Levene، بحيث يختبر فرضية العدم القائلة بتجانس تباينات العينات ضد الفرضية القائلة بعدم تجانس التباين. وغالبا ما يؤدي عدم تجانس التباين إلى عدم تحقق فرضية التوزيع الطبيعي داخل العينات وهو شرط ضروري لاختبار الفرضيات حيث نفترض ان الخطأ العشوائي يتوزع طبيعيا. وفي هذه الحالة يجب اجراء تحويلات على البيانات لضمان تجانس التباين أو جعل التوزيع داخل العينات يقترب من التوزيع الطبيعي. ولتوضيح خطوات اجراء اختبار تجانس التباين نستعين بالمثال التالي:

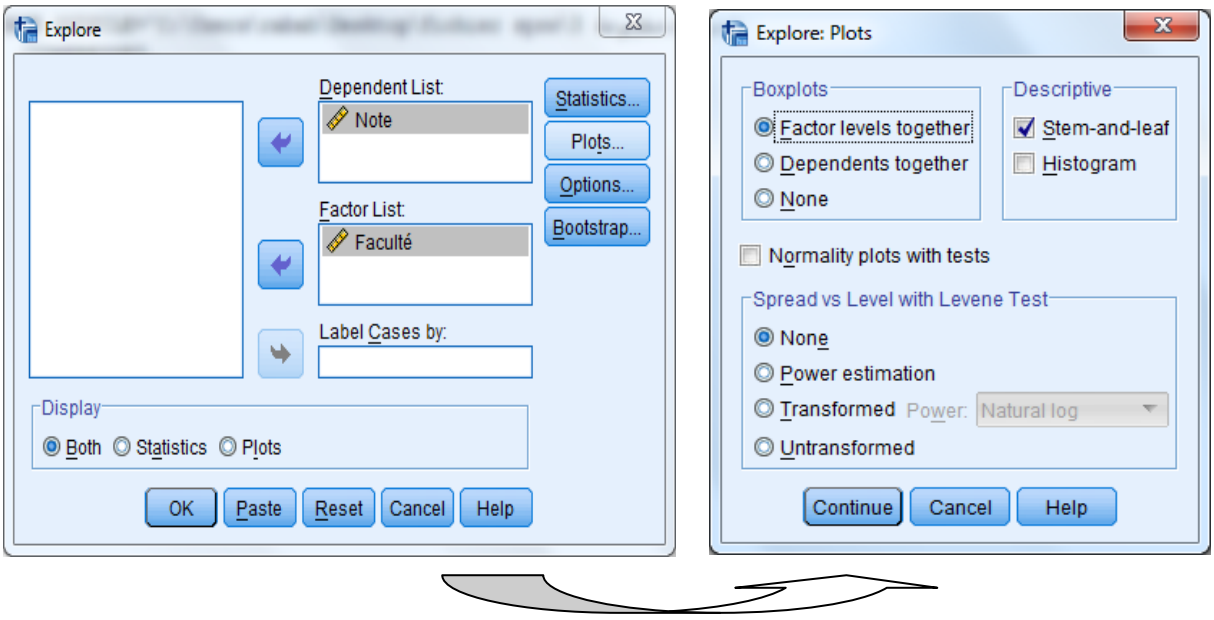
مثال (4-4): لدينا العلامات الخاصة بثلاث أفواج من طلبة السنة الاولى جذع مشترك في مادة الاحصاء في ثلاث كليات: العلوم الاقتصادية، الطب والتكنولوجيا.

كلية العلوم الاقتصادية	كلية العلوم الطبية	كلية العلوم التكنولوجية
18	15	10
12	10	12
8	8	15
5	14	20
10	3	18
16	18	17
11	10	19
4	14	14
10	7	10

ولاجراء الاختبار باستخدام برنامج SPSS، نشير إلى أنه يجب ادخال هذه البيانات بمتغيرين أساسيين هما: العلامات المحصلة لكل فوج (Notes) ومتغيرة خاصة بالكلية التي ينتمي اليها بإعطاء رموز 1،2،3 للتعبير على الكلية (Faculté) كما هي موضحة في الصورة التالية:



لإجراء اختبار تجانس التباين في نقاط الطلبة من مختلف الكليات نختار من قائمة Analyze؛ Descreptive Statistics؛ Explore. ثم ننقل المتغير (Notes) إلى المجال Dependent List، والمتغير العامل (Faculté) إلى المجال Factor List. ومن ثم ننقر على الزر Plots كما يظهر في الشكل التالي:



ونقوم بإجراء الاختبار بالتأشير على أحد الخيارات الموجود في المجال Sperad vs Level with Levene Test وهي على النحو التالي:

- None: عدم اجراء الاختبار الخاص بتجانس التباين، والاكتفاء بخيارات اختبار التوزيع الطبيعي.

- Power estimation: يسمح باجراء الاختبار الاحصائي لتجانس التبيان بحساب إحصائية Levene، واختبار العلاقة بين المستوى والانتشار من خلال الرسم البياني (Spread vs. level plot) مع تحديد نوع التحويل الملائم للبيانات في حالة عدم وجود علاقة أي عدم تجانس التباين.

- Transformed: وهو الخيار الذي يسمح باعداد الرسم البياني مع تحديد طريقة تحويل البيانات في حالة ظهور عدم تجانس التباين، بالتأشير على هذا الخيار يظهر لدينا مجموعة من طرق تحويل البيانات المتوفرة نذكرها في ما يلي:

* Natural Log: التحويل اللوغاريتمي.

*1/square root: التحويل بالقسمة على الجذر التربيعي.

*reciprocal: التحويل بالرفع إلى القوة -1.

* Square root: التحويل باستخدام الجذر التربيعي.

* Square: التحويل باستخدام التربيع (الرفع إلى القوة 2).

* Cube: التحويل عن طريق التكعيب (الرفع إلى القوة 3)

- Untransformed: يفيد هذا الخيار في اعداد الرسم البياني بالاعتماد على القيم الأصلية دون اجراء أي تحويل، وذلك برسم العلاقة بين الوسيط والمدى الربيعي بدون لوغاريتمات.

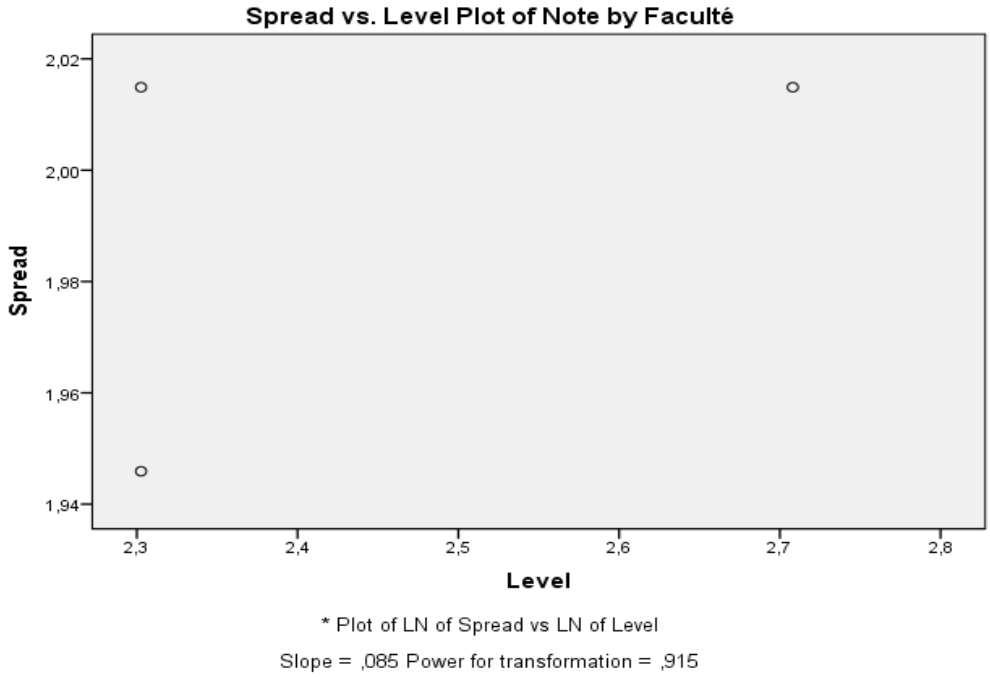
أولا. اجراء الاختبار باختيار Power estimation:

عند اختيارنا للخيار Power estimation، والنقر على الزر Continue يظهر لدينا الجدول الخاص باختبار livene للتجانس والرسم البياني للوغاريتم الوسيط ولوغاريتم المدى الربيعي على الشكل التالي:

Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
Note	Based on Mean	,176	2	24	,839
	Based on Median	,106	2	24	,900
	Based on Median and with adjusted df	,106	2	21,334	,900
	Based on trimmed mean	,180	2	24	,836

من خلال نتائج اختبار Livene الذي يظهر القيم المحسوبة المبنية على المتوسط، الوسيط والوسط الحسابي المشذب الموضحة في الجدول أعلاه، نلاحظ أن القيمة المحسوبة للاختبار المبنية على المتوسط (Based on Mean) تساوي 0.176 وهي أقل من القيمة المجدولة لفيشر بدرجة حرية (2، 24) وهي (عدد العينات−1 "K−1"، وحجم العينة−عدد العينات "N−K") المقدرة بـ 3.40، مما يقودنا إلى قبول فرضية العدم والإقرار بتجانس تباين العينات الثلاثة. كما يمكن اتخاذ القرار بالاستعانة بالقيمة الاحتمالية المحسوبة P−value والمشار اليها في الجدول بالرمز .Sig ومقارنتها بمستوى المعنوية 0.05، ويظهر هنا أن 0.839 أكبر من 0.05 وبالتالي قبول فرضية العدم التي تقر بتجانس التباين ورفض الفرضية البديلة القائلة بان تباين العينات غير متجانس.



ومن خلال الرسم البياني يظهر عدم وجود علاقة بين لوغاريتم الوسيط للعينات ولوغاريتم المدى الربيعي لها، وهذا يشير إلى تجانس تباين العينات، كما أن قيمة الميل Slope=0.085 تشير إلى أن قيمة الأس اللازم لتحويل البيانات يساوي 1−0.085=0.915 وهو قريب من الواحد الصحيح دليل على عدم الحاجة لتحويل البيانات وهي لا تعاني من مشكلة عدم تجانس التباين.

ملاحظة: لو أن نتائج الاختبار اظهرت مشكلة عدم تجانس التباين، نستعين بقمة الأس لمعرفة طبيعة التحويل الواجب اجراؤه على البيانات للتخلص من المشكلة.

4-2. الجداول التقاطعية Crosstabs:

في كثير من الأحيان يحتاج الباحث إلى اعداد جداول تقاطعية لتلخيص بيانات متغيرين أو أكثر في شكل مصفوفة متعددة الأبعاد، تحتوي على تكرارات فئات متغير معين حسب فئات متغير آخر، وهي بذلك عبارة عن جداول تكرارية لأكثر من متغير. ويفيدنا في تحليل العلاقة بين المتغيرات في حالات الاستبيان. ويعرف الجدول التقاطعي على أنه جدول توزيع أفراد عينة الدراسة حسب متغيرين أو أكثر وهو جدول يستخدم لعرض عدد الحالات (التكرارات) التي لها مجاميع مختلفة من قيم متغيرين مصنفين أو اكثر(Categorical Variables)، ويمكن أن يرافق الجدول التقاطعي حساب ملخصات إحصائية واختبارات. ويسمى الجدول التقاطعي لمتغيرين باسم (two-way crostabulation)، ويسمى الجدول التقاطعي لأكثر من متغيرين باسم (multi-way crostabulation).

وعموما يتم استخدام هذا النوع من الجداول في تحليل العلاقة بين متغيرات نوعية وكمية على حد سواء، إلا أنه اذا تعلق الأمر بالمتغيرات الكمية يجب تبويبها إلى فئات محددة، والتحليل على هذا الأساس يكون أكثر وضوحا. وفي هذا الصدد يوفر البرنامج SPSS من خلال الأمر Crosstabs امكانية استخراج الجدول التقاطعي إلى جانب اجراء اختبارات احصائية لفرضية استقلالية الأسطر عن الاعمدة بمختلف الاساليب حسب طبيعة المتغيرات، بالإضافة إلى اعداد مخططات بيانية تتماشى مع تحليل العلاقة بين المتغيرات.

أولا. الجدول التقاطعي لمتغيرين:

لتوضيح طريقة اجراء الجدول التقاطعي لمتغيرين (two-way crostabulation) نستعين بالمثال التالي:

مثال (4-5):

في دراسة ميدانية لإحدى الشركات تعمل في مجال الاتصالات، ومن أجل تحديد مدى معرفة زبائنها بالتكنولوجيات الحديثة، لدينا الإجابات الخاصة بـ 144 زبون بحيث تم اختيار سؤالين اساسين هما:

السؤال الأول: إلى أي مدى توافق على أهمية تعرفك على التكنولوجيات الجديدة مباشرة عند طرحها في السوق؟

1. موافق تماما 2. موافق 3.غير موافق 4 .غير موافق تماما

السؤال الثاني: حدد فيما يلي الفئة العمرية التي تنتمي اليها؟

1. أقل من 24 سنة؛ 2. من 25 إلى 34 سنة؛ من 35 إلى 44 سنة؛ من 45 إلى 54 سنة؛ 55 سنة فأكثر.

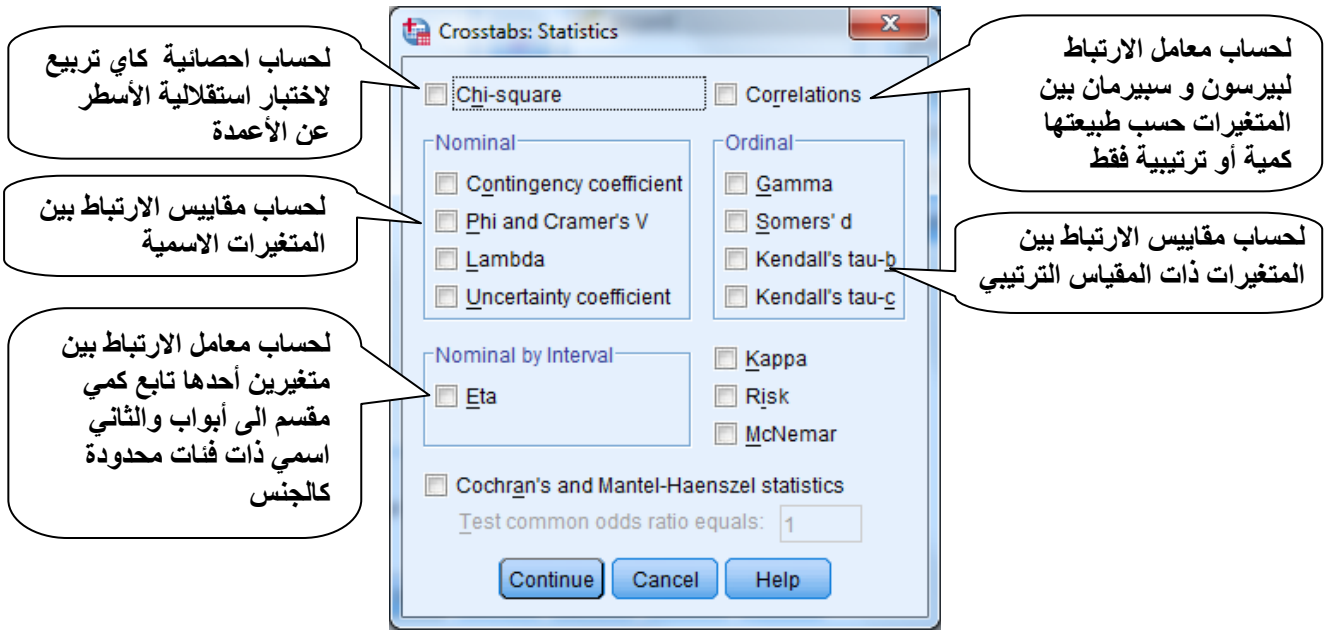
وتكمن الفرضية المراد التأكد منها في هذا الاستبيان هي وجود علاقة بين أهمية التعرف على التكنولوجيات الجديدة مباشرة عند طرحها في السوق وفئات العمر التي ينتمي اليها المجيب (الزبون).

للتحقق من هذه الفرضية وبناءا على البيانات الخاصة بالإجابة على السؤالين المذكورين اعلاه والتي كانت على الشكل التالي:

نختار من القائمة Analyze؛ Descreptive Statistics الأمر Crosstabs فيظهر لدينا صندوق الحوار الخاص بـ Crosstabs كما هو مبين في الصورة التالية:

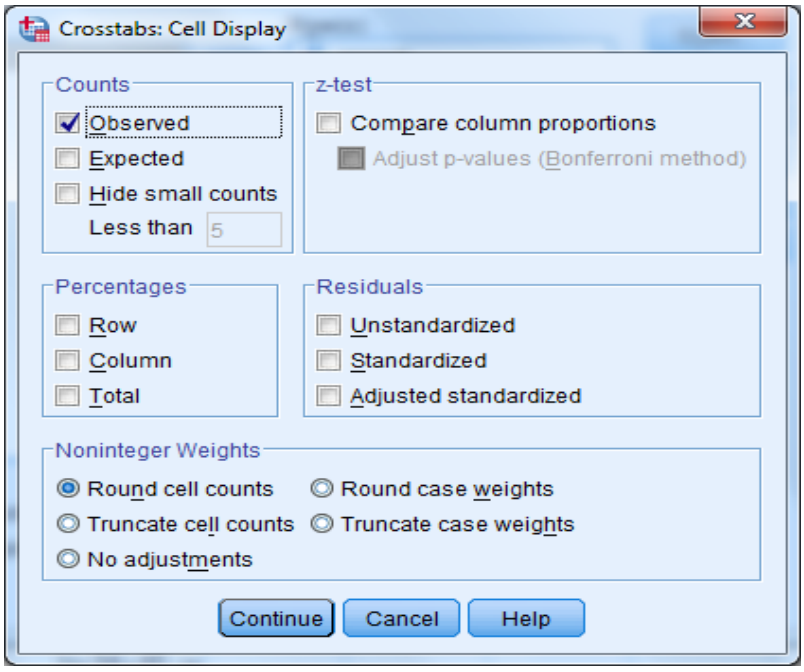
بحيث نختار المتغير accord في الأسطر والمتغير age في الأعمدة لكن هذا الاختيار ليس ضروري ان يكون بهذا الترتيب فيمكن قلبه حسب حاجة الباحث. ويظهر في هذا الصندوق أزرار مهمة تتمثل في:

- Statistics: يوفر هذا الزر امكانية اختيار مخرجات العملية الاحصائية بالتأشير على مجموعة من الخيارات يوفرها صندوق حوار Statistics على النحو التالي:



ونشير هنا إلى أن معاملات الارتباط المراد حسابها يتوقف على طبيعة المتغيرات، ففي حالة المتغيرات الإسمية يتم التأشير على أحد الخانات المتوفرة في المجال Nominal، بينما في حالة المتغيرات الترتيبية يتم التأشير على أحد الخيارات في المجال Ordinal. أما في حالة المتغير التابع ذات المقياس الكمي والمقسم إلى فئات ومتغير ثاني اسمي ذو فئات محدودة كمتغير الجنس مثلا فيمكن التأشير على Eta في المجال Nominal by Interval، ويتم التأشير على Correlation لحساب معاملي الارتباط لـ Spearman و Person في آن واحد علما أنه يجب أن تكون المتغيرات التابعة عددية أو رقمية فيتم حساب معامل ارتباط سبيرمان باعتبار أن كلا الأسطر والأعمدة عبارة عن قيم ترتيبية مثلا المستوى العلمي، كما يتم حساب معامل ارتباط بيرسون باعتبار ان كلا من الأسطر والأعمدة متغيرات كمية.

- Cells: يفيد هذا الزر في التحكم بمحتويات الجدول التقاطعي حيث يظهر صندوق الحوار الخاص به على الشكل التالي:



وهو يشمل:

- Counts: فيه ثلاث خيارات اساسية:

- Observed: وهو الخيار التلقائي المقصود به عرض قيم التكرار المشاهد Oi في كل خانة من خانات الجدول التقاطعي.
- Expected: والمقصود به التكرار المتوقع Ei.

- Hide small counts: الغرض منه اختيار إخفاء الحسابات الأقل من عدد صحيح محدد، سيتم عرض القيم المخفية على أنها <N، حيث N هي العدد الصحيح المحدد. ويشترط أن يكون العدد الصحيح المحدد أكبر من أو يساوي 2.

➢ Percentages: يخص اختيار النسب المؤوية وفق ثلاث خيارات هي:

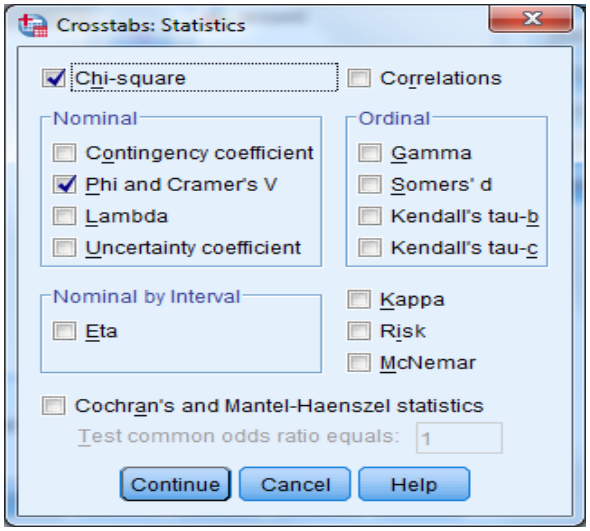
- Row: تملأ خلايا الجدول بنسب مئوية محسوبة بالنسبة لمجموع السطر؛
- Columns: تملأ خلايا الجدول بنسبة مئوية محسوبة بالنسبة لمجموع العمود؛
- Total: تملأ خلايا الجدول بنسبة مئوية محسوبة بالنسبة للمجموع الكلي للتكرارات.

➢ Residuals: يتكون من ثلاث خيارات:

- Unstandarised: تملأ خلايا الجدول بالفرق بين التكرار المشاهد والتكرار المتوقع Oi-Ei؛
- Standarized: تملأ خلايا الجدول بالفرق بين التكرار المشاهد والتكرار التوقع مقسوما على الخطأ المعياري له.
- Adj-Standarized: وهو نفس الخيار السابق معبرا عنه بوحدات الانحراف المعياري عن المتوسط.

كما يمكن اختيار ترتيب أسطر الجدول تصاعديا أو تنازليا من خلال الزر Format في صندوق حوار Crosstabs.

بالعودة إلى مثالنا، لدينا المتغيرتين accord و age كلاهما متغيرين اسميين (Nominal)، وبالتالي تكون الخيارات الخاصة بجدول التقاطع في هذه الحالة مقتصرا على عرض التكرار المشاهد والمتوقع، بالاضافة إلى احصائية كاي تربيع وأحد المعاملات الخاصة بالمجال Nominal وليكن مثلا Phi and Cramers V. كما هو موضح في الصورة المقابلة:



بحيث يتم حساب احصائية كاي تربيع وفق العلاقة التالية:

$$Chi-square=\sum_{i=1}^{l}\sum_{j=1}^{c}\frac{(O_{ij}-E_{ij})^2}{E_{ij}}$$

بدرجة حرية (l-1)×(c-1) أي عدد الاسطر ناقص واحد في عدد الاعمدة ناقص واحد.

أما احصائية phi وهي عبارة عن مقياس الارتباط بين المتغيرتين فهي تحسب باستخدام احصائية كاي تربيع المحسوبة وفق العلاقة:

$$Phi = \sqrt{\frac{Chi - square}{N}}$$

وهو يستخدم في حالة الجداول من الشكل 2×2، أما في الحالة العامة فيتم استعمال معامل كرامر (C) Cramer Coefficient، والذي يحسب للجداول المتعددة وفق العلاقة التالية:

$$C = \sqrt{\frac{Chi - square}{N(r-1)}}$$

بحيث تمثل N عدد المشاهدات و r هو العدد الاصغر بين الأسطر أو الأعمدة.

وبتطبيق الأمر والنقر على الزر Continue يظهر لنا جدول التقاطع على الشكل التالي:

accord * age Crosstabulation

			age					
			أقل من 24سنة	بين 25 و 34 سنة	بين 35 و 44 سنة	بين 45 و 54 سنة	55 سنة فأكثر	Total
accord	موافق تماما	Count	7	11	2	14	10	44
		% within accord	15,9%	25,0%	4,5%	31,8%	22,7%	100,0%
	موافق	Count	11	7	10	8	15	51
		% within accord	21,6%	13,7%	19,6%	15,7%	29,4%	100,0%
	غير موافق	Count	4	14	5	7	4	34
		% within accord	11,8%	41,2%	14,7%	20,6%	11,8%	100,0%
	غير موافق تماما	Count	0	4	3	3	5	15
		% within accord	0,0%	26,7%	20,0%	20,0%	33,3%	100,0%
Total		Count	22	36	20	32	34	144
		% within accord	15,3%	25,0%	13,9%	22,2%	23,6%	100,0%

ويتم قراءة الجدول حسب رغبة الباحث اما بالتكرار المطلق أو النسبي، بحيث يظهر أن 7 أفراد من بين 44 الموافقون تماما بمعرفة التكنولوجيات الجديدة مباشرة عند طرحها في السوق عمرهم يقل عن 24 سنة، 11 فردا من فئة العمر بين 25-34 سنة، 2 فقط أعمارهم بين 35 و44 سنة، 14 مفئة العمر 45-54 سنة و 10 من فئة العمر 55 فأكثر، وهكذا بالنسبة لكل أسطر الجدول. كما يمكن قراءة الجدول باستخدام النسب سواءً بالنسبة للأسطر أو الأعمدة.

أما بالنسبة لاختبار الفرضية القائلة بأنه توجد علاقة بين معرفة الزبائن بالتكنولوجيا الجديدة مباشرة بعد طرحها في السوق وفئات العمر المختلفة فتكون في الجدول الثاني المبين لنتائج اختبار كاي تربيع على النحو التالي:

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	20,694[a]	12	0,055
Likelihood Ratio	23,824	12	0,021
Linear-by-Linear Association	0,002	1	0,962
N of Valid Cases	144		

بحيث يظهر بأن القيمة المحسوبة لاحصائية كاي تربيع تقدر بـ 20.694 وهي أقل من القيمة المجدولة بدرجة حرية (4−1)×(5−1)=12 والمقدرة بـ 21.026، الأمر الذي يقودنا إلى قبول الفرضية البديلة التي تقر بعدم استقلالية الأسطر عن الأعمدة، وبالتالي توجد علاقة بين تعرف الزبائن على التكنولوجيات الجديدة مباشرة عند طرحها في السوق وعمر الزبون. كما يمكن اثبات ذلك من خلال مقارنة القيمة الاحتمالية المقدرة بـ 0.055 وهي أكبر من مستوى المعنوية 0.05.

أما بالنسبة لمعامل الارتباط لكرامر (Cramer Coefficient) والذي يظهر من خلال الجدول التالي:

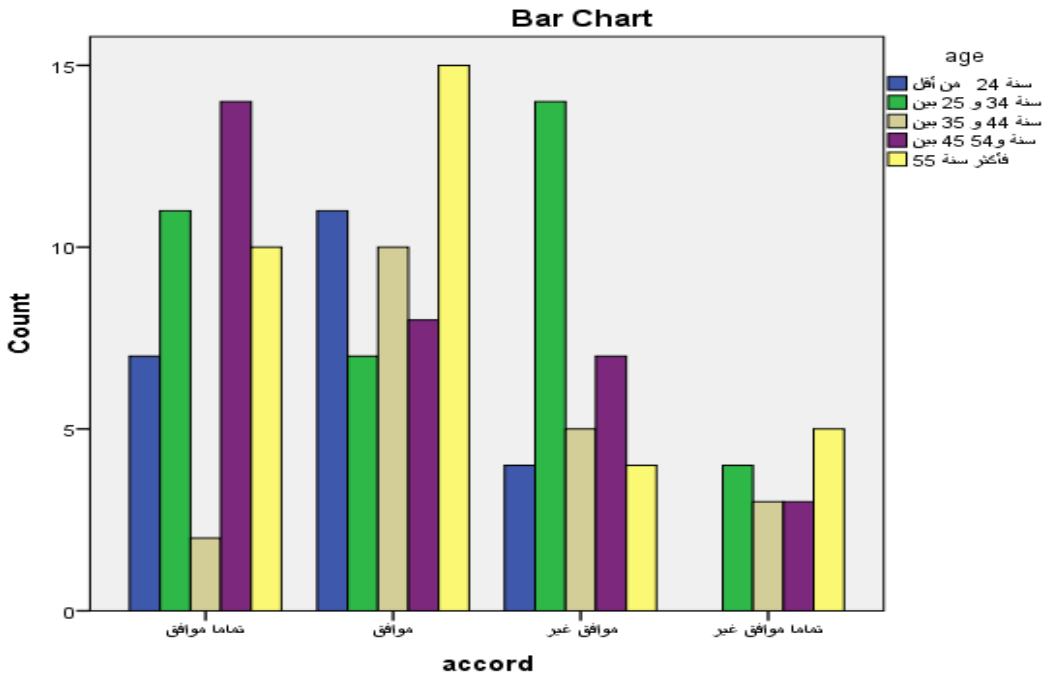
Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	0,379	0,055
	Cramer's V	0,219	0,055
N of Valid Cases		144	

فيشير إلى أن القيمة الاحتمالية لهذا المعامل المقدر بـ C=0.219، وهي كذلك أكبر من مستوى المعنوية 0.05 بحيث قدرت بـ 0.055.

وتجدر الاشارة إلى انه يمكن استخدام معاملات أخرى التي تتوفر في المجال Nominal كمعامل التوافق مثلا وغيرها، والتي ستبين نفس النتائج لاختبار فرضية استقلالية الأسطر عن الأعمدة.

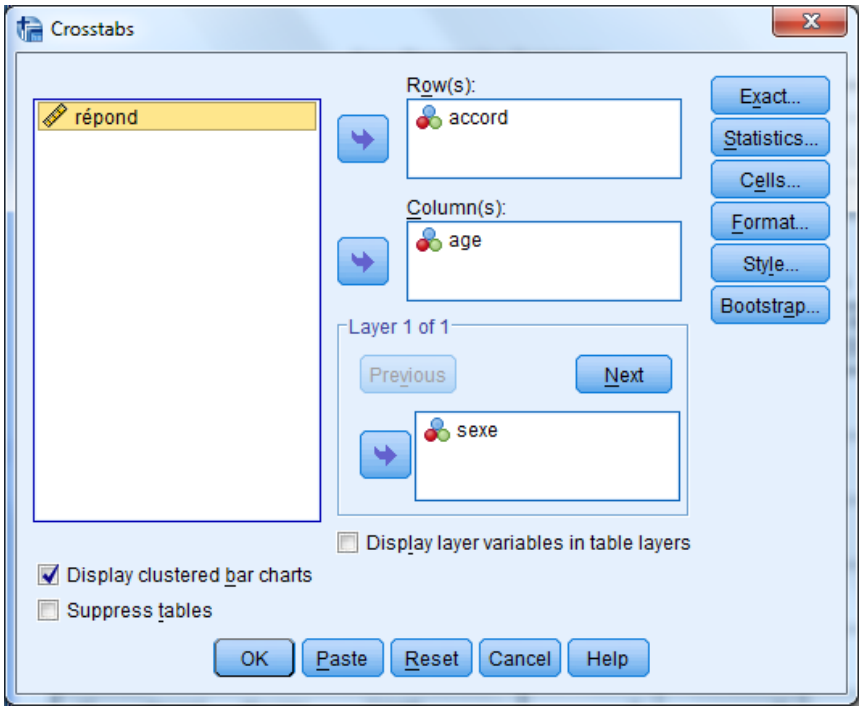
كما يمكن ترجمة الجدول التقاطعي في شكل مخطط الأعمدة التكرارية بالتأشير على المربع المجاور لـ display Clustered Bar Chart في صندوق حوار Crosstabs، بحيث يظهر في المخرجات المخطط البياني على الشكل التالي:



ثانيا. الجدول التقاطعي لأكثر من متغيرين:

في هذه الحالة نستعين بمتغير ثالث يتمثل في متغير الجنس، فتصبح الفرضية هي العلاقة بين متغير الموافقة (accord) والسن (age) حسب متغير الجنس والذي يصبح معروفا بمتغير السيطرة Control Variable، بحيث يتم اعداد جدولين تقاطعين حسب فئات هذا الأخير وهي ذكور و إناث.

ولتنفيذ هذا الأمر نتبع نفس الخطوات السابقة مع اضافة متغير الجنس إلى البيانات نقله إلى المجال Layer 1of1، كما هو موضح في الصورة التالية:



وبالنقر على الزر OK تظهر لدينا النتائج كما يلي:

-الجدول التقاطعي مقسم إلى قسمين حسب فئات متغير السيطرة وهو متغير الجنس:

accord * age * sexe Crosstabulation

sexe				age					Total
				أقل من 24سنة	بين 25 و 34سنة	بين 35 و 44سنة	بين 45 و 54سنة	55سنة فأكثر	
ذكر	accord	موافق تماما	Count	2	5	0	12	8	27
			% within accord	7,4%	18,5%	0,0%	44,4%	29,6%	100,0%
		موافق	Count	5	1	3	6	6	21
			% within accord	23,8%	4,8%	14,3%	28,6%	28,6%	100,0%
		غير موافق	Count	2	7	4	2	2	17
			% within accord	11,8%	41,2%	23,5%	11,8%	11,8%	100,0%
		غير موافق تماما	Count	0	2	1	2	2	7
			% within accord	0,0%	28,6%	14,3%	28,6%	28,6%	100,0%
		Total	Count	9	15	8	22	18	72
			% within accord	12,5%	20,8%	11,1%	30,6%	25,0%	100,0%
أنثى	accord	موافق تماما	Count	5	6	2	2	2	17
			% within accord	29,4%	35,3%	11,8%	11,8%	11,8%	100,0%
		موافق	Count	6	6	7	2	9	30
			% within accord	20,0%	20,0%	23,3%	6,7%	30,0%	100,0%
		غير موافق	Count	2	7	1	5	2	17
			% within accord	11,8%	41,2%	5,9%	29,4%	11,8%	100,0%
		غير موافق تماما	Count	0	2	2	1	3	8
			% within accord	0,0%	25,0%	25,0%	12,5%	37,5%	100,0%
		Total	Count	13	21	12	10	16	72
			% within accord	18,1%	29,2%	16,7%	13,9%	22,2%	100,0%
Total	accord	موافق تماما	Count	7	11	2	14	10	44
			% within accord	15,9%	25,0%	4,5%	31,8%	22,7%	100,0%
		موافق	Count	11	7	10	8	15	51
			% within accord	21,6%	13,7%	19,6%	15,7%	29,4%	100,0%
		غير موافق	Count	4	14	5	7	4	34
			% within accord	11,8%	41,2%	14,7%	20,6%	11,8%	100,0%
		غير موافق تماما	Count	0	4	3	3	5	15
			% within accord	0,0%	26,7%	20,0%	20,0%	33,3%	100,0%
		Total	Count	22	36	20	32	34	144
			% within accord	15,3%	25,0%	13,9%	22,2%	23,6%	100,0%

جدول اختبار الاستقلالية بين الأسطر والأعمدة حسب الجنس:

Chi-Square Tests

sexe		Value	df	Asymp. Sig. (2-sided)
ذكر	Pearson Chi-Square	20,727[b]	12	0,055
	Likelihood Ratio	24,734	12	0,016
	Linear-by-Linear Association	2,194	1	0,139
	N of Valid Cases	72		
أنثى	Pearson Chi-Square	15,103[c]	12	0,236
	Likelihood Ratio	16,368	12	0,175
	Linear-by-Linear Association	2,832	1	0,092
	N of Valid Cases	72		
Total	Pearson Chi-Square	20,694[a]	12	0,055
	Likelihood Ratio	23,824	12	0,021
	Linear-by-Linear Association	,002	1	0,962
	N of Valid Cases	144		

6 cells (30,0%) have expected count less than 5. The minimum expected count is 2,08.[a]

14 cells (70,0%) have expected count less than 5. The minimum expected count is ,78.[b]

16 cells (80,0%) have expected count less than 5. The minimum expected count is 1,11.[c]

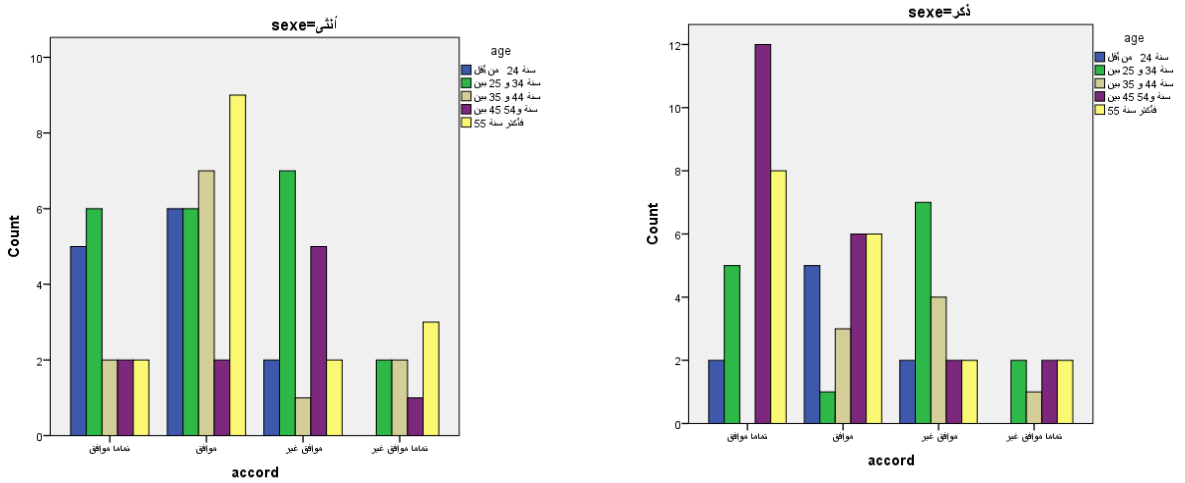
من خلال الجدول يظهر ان القيمة الاحتمالية المحسوبة لكاي تربيع أكبر من 0.05 بالنسبة للذكور والإناث على حد سواء مما يقودنا إلى قبول الفرضية البديلة، وبالتالي الاقرار بوجود علاقة بين المتغيرين accord و age.

ملاحظة:

تجدر الاشارة إلى ان الاعتماد على احصائية كاي تربيع في هذه الحالة يؤدي نتائج مظللة، نظرا لوجود نسب كبيرة من التكرارات المتوقعة تقل قيمتها عن 5 كما تشير إليه الملاحظات في آخر الجدول وهذا يتعارض مع فرضيات الاختبار النظرية، وعليه يجب استعمال اختبارات أخرى وعدم الأخذ بعين الاعتبار قيمة كاي تربيع.

- المخطط البياني للعلاقة:

في هذه الحالة يقوم البرنامج بإعداد مخططين أحدهما خاص بالذكور والآخر خاص بالإناث كما يظهر في الشكل الموالي:



تمرين تدريبي:

يطلب من الطالب إعادة تطبيق الخطوات المتبعة في هذا المثال على البيانات الخاصة باستعمال وسائل النقل المدروس في بداية هذا الفصل (الأمر Frequencies)، لدراسة العلاقة بين استعمال الحافلة ومتغير السن، ومن ثم دراسة هذه العلاقة حسب متغير الجنس.

3. مقارنة المتوسطات:

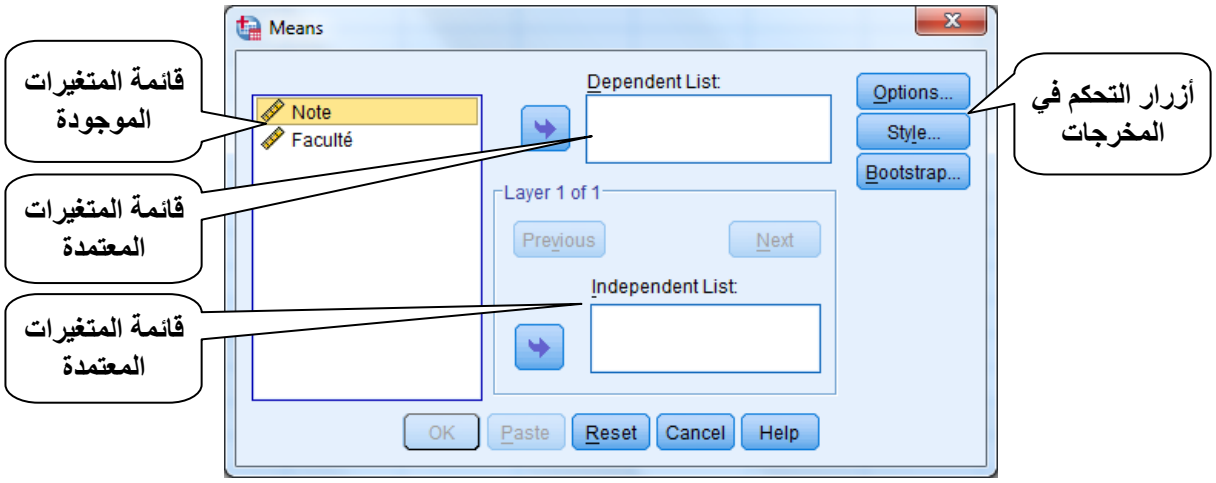
نتناول في هذا الجزء احد اجراءات الاحصاء الاستدلالي، ويتعلق الأمر باختبار فرضية تساوي متوسط عينة أو عدة عينات فيما بينها، مع الاشارة إلى خطوات هذا الإجراء باستخدام برنامج SPSS حسب الحالة، بحيث يتوفر البرنامج على عدة أوامر تتعلق بهذا الجانب من خلال الأمر Compare Means حسب الحالة.

3-1. تحليل المتوسطات بواسطة الأمر Means:

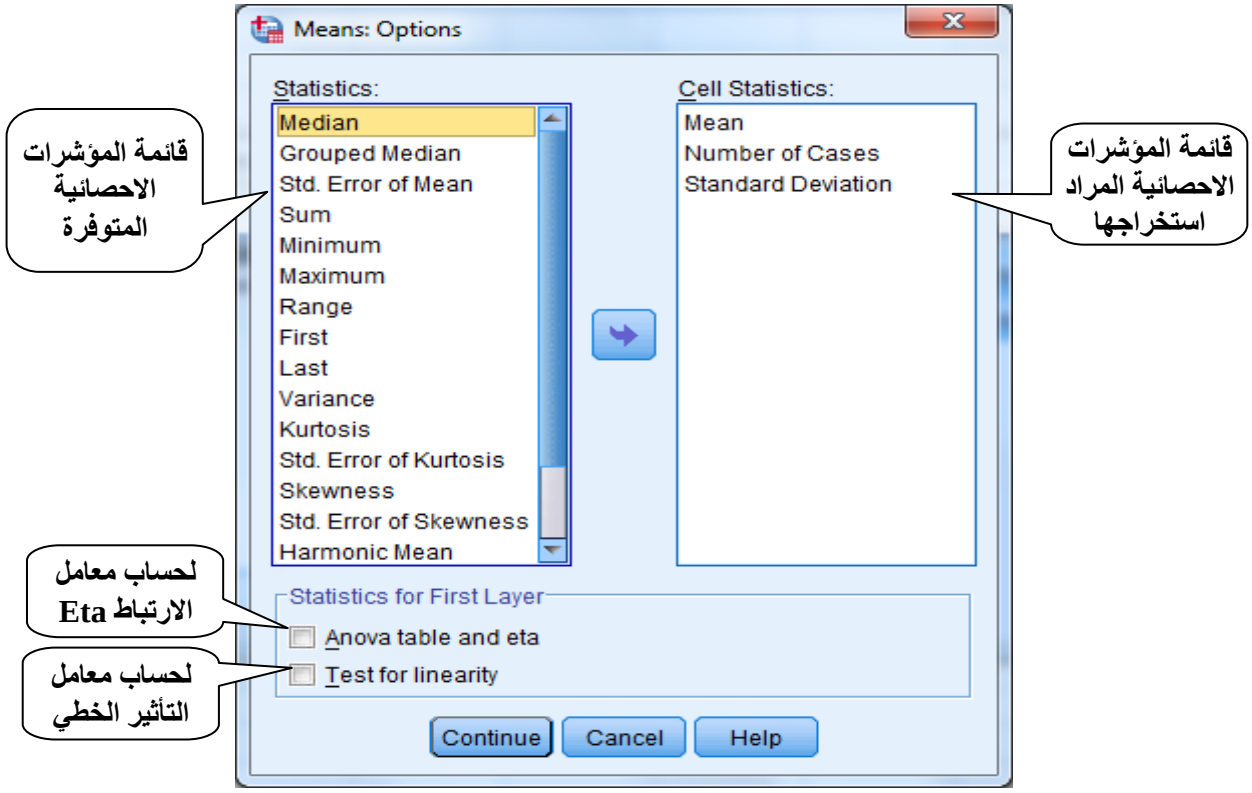
يفيد هذا التحليل في حساب المتوسطات للمجاميع الجزئية لمتغير معين، بالإضافة إلى مؤشرات أخرى مثل (المجموع، الوسيط، المنوال، الانحراف المعياري...وغيرها)، كما يمكن ايضا اجراء تحليل التباين بمعيار واحد واختبار التأثير الخطي للمعاملات، واختبار الارتباط (الازدواج) بحساب معامل الارتباط Eta.

أولا: تحليل المتوسط حسب متغير مستقل واحد:

مثال (4-6): بناء على البيانات المتعلقة بعلامات الطلبة في مادة الاحصاء لطلبة ثلاث أفواج من كليات مختلفة المدروس سابقا، نقوم بتطبيق الأمر Means من قائمة Analyze؛ Compare Means فيظهر صندوق حوار على الشكل التالي:



ولتنفيذ الأمر ننقل المتغير Note إلى المجال Dependent List باعتباره متغير معتمد، ونقل المتغير Faculté إلى المجال Independent List بصفته متغير عامل أو مستقل. ثم النقر على الزر Option لتحديد خيارات الإجراء الاحصائي فيظهر لنا صندوق حوار Option التالي:



ومن خلال هذا الصندوق يمكن للباحث اضافة المؤشرات الاحصائية التي يرغب في حسابها إلى جانب المؤشرات التلقائية الثلاثة الموجودة في المجال Cell Statistics بالطريقة المعتادة، ومن ثم التأشير على المربع الصغير المقابل لـ Anova table and eta التي من شأنها الكشف عن وجود اتجاه عام خطي معنوي بين متوسطات العينات في تحليل التباين بمعيار واحد، وكذلك اجراء اختبار التأثير الخطي من خلال التأشير على المربع الصغير المقابل لـ Test for linearity.

وبتطبيق الأمر ينتج لدينا النتائج التالية:

- جدول المؤشرات الإحصائية: وتحدد محتوياته من طرف المستخدم بإضافة المؤشرات الاحصائية حسب الحاجة، وفي الحالة التلقائية يعرض المتوسط الحسابي، عدد المشاهدات والانحراف المعياري لكل عينة.

Report

Note

Faculté	Mean	N	Std. Deviation
1	10,44	9	4,586
2	11,00	9	4,664
3	15,00	9	3,775
Total	12,15	27	4,672

- جدول تحليل التباين: يتم من خلاله اختبار وجود فروق معنوية بين متوسطات العينات باستخدام احصائية F، بحيث يختبر فرضية العدم القائلة بأنه يوجد فروقات معنية بين متوسطات العينات ضد الفرضية البديلة التي تقر بعدم وجود فروقات معنوية بين متوسطات العينات.

ANOVA Table

			Sum of Squares	df	Mean Square	F	Sig.
Note * Faculté	Between Groups	(Combined)	111,185	2	55,593	2,925	0,073
		Linearity	93,389	1	93,389	4,913	0,036
		Deviation from Linearity	17,796	1	17,796	0,936	0,343
	Within Groups		456,222	24	19,009		
	Total		567,407	26			

يظهر من خلال نتائج تحليل التباين للمتغيرين Note و faculté أن قيمة فيشر المحسوبة تساوي 2.925 وهي أقل من القيمة المجدولة ($F_{(2,26)}$=3.37) مما يدل على عدم وجود فروق معنوية بين متوسطات العينات الثلاثة (رفض فرضية العدم). بينما تشير احصائية F للعلاقة الخطية بين متوسطات العينات (Deviation from Linearity) المقدرة بـ 0.936 إلى قبول فرضية العدم القائلة بأن متوسطات العينات تقع على خط مستقيم ولا تنحرف عنه، لأن القيمة المحسوبة أقل من القيمة المجدولة ($F_{(1,24)}$=4.26). كما يمكن اتخاذ القرار الخاص باستخدام القيم الاحتمالية .Sig بحيث يتم قبول الفرضية الصفيرية اذا كانت القيمة المحسوبة أكبر من 0.05.

- جدول مقياس الاقتران (الارتباط): يتم من خلاله عرض نتائج حساب معامل الارتباط البسيط R بين المتغير المعتمد note والمتغير المستقل faculté، ومقياس جودة التوفيق الخطي R-squared بالإضافة إلى مقياس الاقتران Eta ونسبة التباينات الكلية للمتغير المعتمد المفسرة بالمتغير المستقل.

Measures of Association

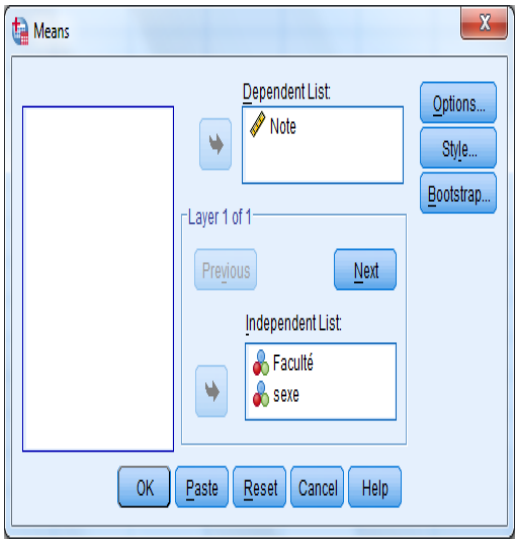
	R	R Squared	Eta	Eta Squared
Note * Faculté	0,406	0,165	0,443	0,196

بحيث تشير قيمة R=0.406 إلى معامل الارتباط بين المتغيرين المعتمد والمستقل، R-Squared إلى مقياس جودة توفيق نموذج الانحدار الخطي والتي تشير إلى أن 16.5% من التغير الحاصل في المتغير التابع راجع إلى التغير في المتغير المعتمد. أما قيمة مقياس الارتباط Eta=0.443 فهي تشير إلى وجود ارتباط بدرجة متوسطة بين المتغيرين، فهذا المعامل يكون محصور بين القيمة 0 المعبرة عن عدم وجود ارتباط بين المتغيرين و القيمة 1 التي تعبر عن وجود ارتباط تام. بينما تشير قيمة Eta Squared إلى أن 19.6% من التباينات الكلية في المتغير المعتمد يفسرها المتغير المستقل.

ثانيا: تحليل المتوسط بأكثر من متغير مستقل:

يمكن استخدام أكثر من متغير مستقل في تحليل المتوسطات بطريقتين مختلفتين:

- **الطريقة الأولى**: نقل كل المتغيرات المستقلة إلى المجال Independent List، وفي هذه الحالة يقوم البرنامج بإجراء تحليل المتوسطات حسب كل متغير في شكل أزواج متغير معتمد*متغير مستقل، وفي هذه الحالة يعرض البرنامج جداول المؤشرات الاحصائية بشكل مستقل حسب المتغيرات المدرجة.



بإضافة متغير الجنس إلى بيانات المثال السابق نقوم باجراء Compare Means-Means بادخال المتغير المعتمد Note في المجال Dependent List والمتغيرات faculté و sexe في المجال Independent List كما توضحه الصورة المقابلة:

بإتباع نفس الإجراءات المذكورة سابقا نحصل على المخرجات التالية:

Note * Faculté

Report

Note

Faculté	Mean	N	Std. Deviation
اقتصاد	10,44	9	4,586
طب	11,00	9	4,664
تكنولوجيا	15,00	9	3,775
Total	12,15	27	4,672

ANOVA Table

			Sum of Squares	df	Mean Square	F	Sig.
Note * Faculté	Between Groups	(Combined)	111,185	2	55,593	2,925	0,073
		Linearity	93,389	1	93,389	4,913	0,036
		Deviation from Linearity	17,796	1	17,796	0,936	0,343
	Within Groups		456,222	24	19,009		
	Total		567,407	26			

Measures of Association

	R	R Squared	Eta	Eta Squared
Note * Faculté	0,406	0,165	0,443	0,196

Note * sexe

Report

Note

sexe	Mean	N	Std. Deviation
ذكر	11,29	14	5,327
أنثى	13,08	13	3,840
Total	12,15	27	4,672

ANOVA Table[a]

			Sum of Squares	df	Mean Square	F	Sig.
Note * sexe	Between Groups	(Combined)	21,627	1	21,627	0,991	0,329
	Within Groups		545,780	25	21,831		
	Total		567,407	26			

a. With fewer than three groups, linearity measures for Note * sexe cannot be computed.

Measures of Association

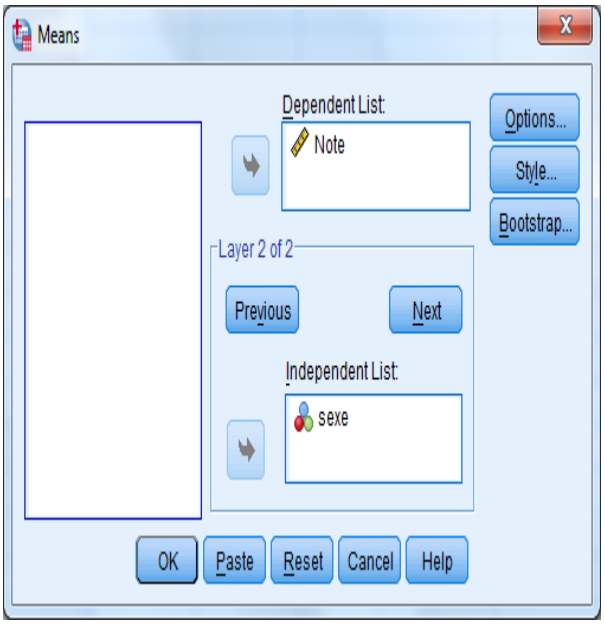
	Eta	Eta Squared
Note * sexe	0,195	0,038

- **الطريقة الثانية**: تحليل متوسط المتغير المعتمد حسب متغير مستقل ضمن فئات (مجاميع جزئية) لمتغير آخر، بحيث يعتبر أحد المتغيرات متغير مستقل والآخر متغير تجزئة، وهنا يتم عرض نتائج التحليل في جدول واحد حسب كل نوع من الجداول حسب المتغير المستقل.

وفي مثالنا متوسط المتغير note حسب متغير الجنس sexe ضمن داخل متغير الكلية التي ينتمي اليها الطالب Faculté. في هذه الحالة نستعين بالزر Next في المجال Layer 1 of 1 بحيث يتم نقل المتغير المستقل الأول faculté ثم النقر على الزر Next وإضافة متغير الجنس إلى مجال Independent List. وبإتمام العملية نحصل على النتائج التالية:

Report

Faculté	sexe	Mean	N	Std. Deviation
اقتصاد	ذكر	9,00	5	5,568
	أنثى	12,25	4	2,630
	Total	10,44	9	4,586
طب	ذكر	8,75	4	4,573
	أنثى	12,80	5	4,324
	Total	11,00	9	4,664
تكنولوجيا	ذكر	15,60	5	3,050
	أنثى	14,25	4	4,924
	Total	15,00	9	3,775
Total	ذكر	11,29	14	5,327
	أنثى	13,08	13	3,840
	Total	12,15	27	4,672



ANOVA Table

			Sum of Squares	df	Mean Square	F	Sig.
Note * Faculté	Between Groups	(Combined)	111,185	2	55,593	2,925	0,073
		Linearity	93,389	1	93,389	4,913	0,036
		Deviation from Linearity	17,796	1	17,796	0,936	0,343
	Within Groups		456,222	24	19,009		
	Total		567,407	26			

Measures of Association

	R	R Squared	Eta	Eta Squared
Note * Faculté	,406	,165	,443	,196

3-2. اختبار t لعينة واحدة One sample t test:

يفيد هذا الاختبار في الكشف عن وجود اختلاف معنوي لمتوسط المجتمع الذي حسبت منه العينة عن قيمة ثابتة. بالإضافة إلى امكانية تقدير مجال الثقة لمتوسط المجتمع ويستعمل هذا الاختبار للعينات الصغيرة (أقل من 30 مشاهدة).

وتكون الفرضية المختبرة في هذه الحالة هي:

$$H_0 : \overline{X} = \mu$$
$$H_1 : \overline{X} \neq \mu$$

بحيث u تمثل قيمة المتوسط النظري (الافتراضي)

وتكون صيغة الاختبار t بالصيغة التالية:

$$t = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$$

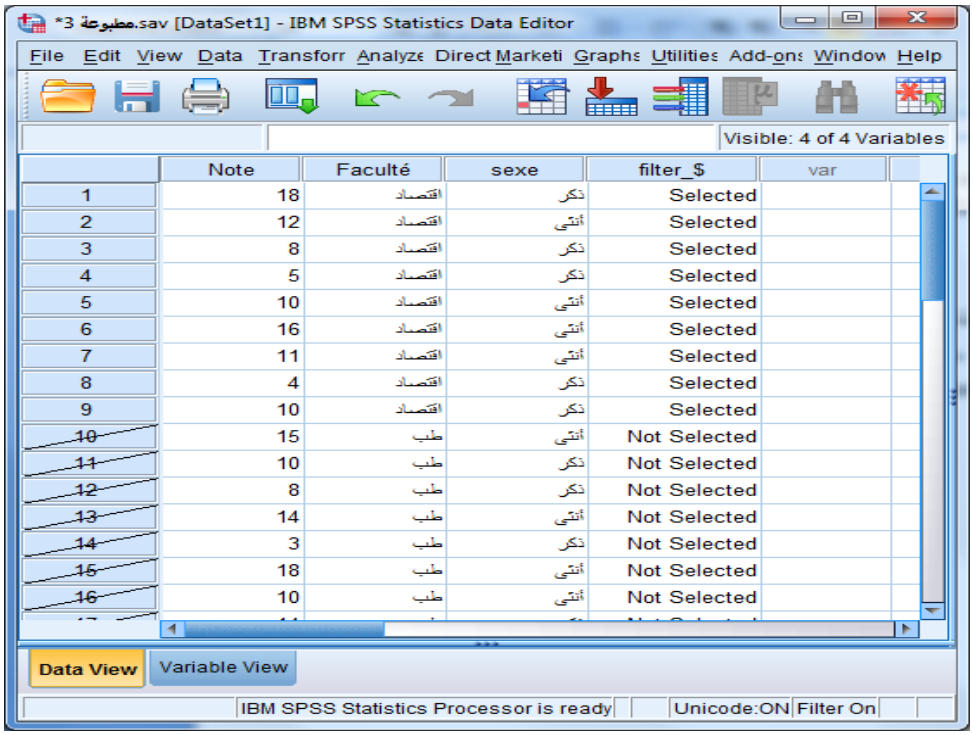
بحيث تمثل σ قيمة الانحراف المعياري و n عدد المشاهدات.

وللقيام بالاختبار باستخدام برنامج SPSS نختار من قائمة Analyze الأمر Compare means ثم One Simple T Test.

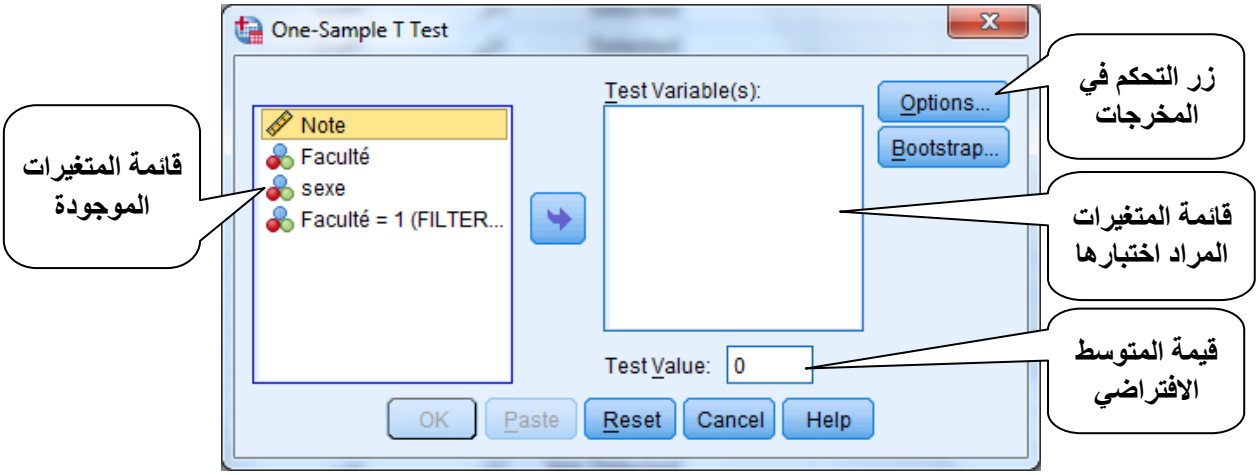
مثال (4-7):

لناخذ مثلا بيانات المثال السابق، بحيث نختار فقط طلبة كلية العلوم الاقتصادية ونختبر فرضية تساوي متوسط النتائج المحصلة بالمعدل 10 مثلا.

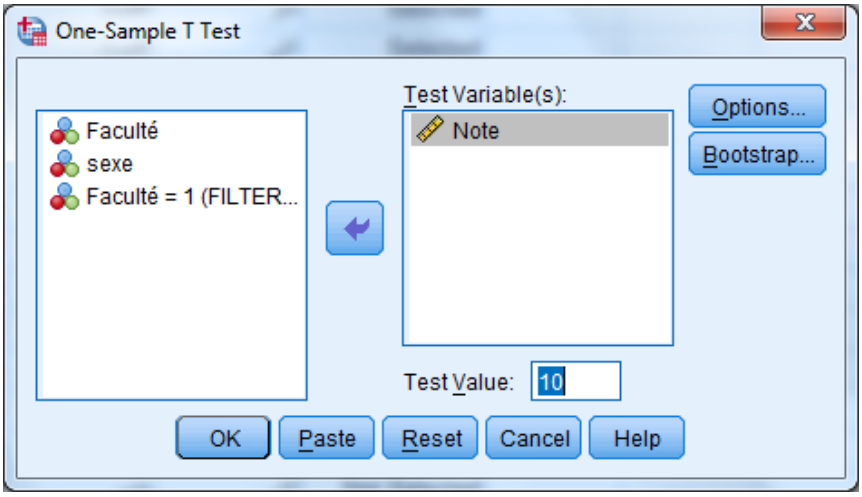
للقيام بهذا الإجراء نستعين بالأمر Select Cases لاختيار عينة الطلبة، ومن ثم نختبر الفرضية كما هو موضح ي الصورة التالية:



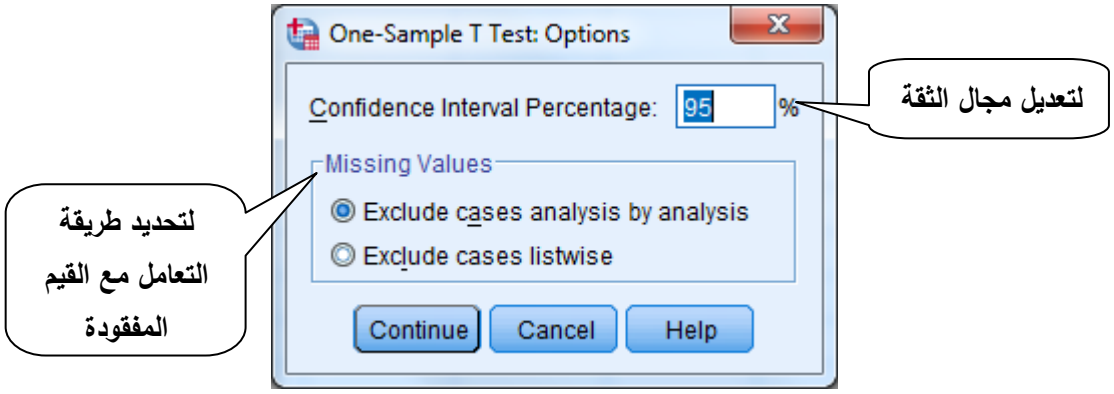
وباختيار الأمر One Simple T Test يظهر لدينا صندوق الحوار التالي:



بحيث ننقل المتغير note إلى المجال Test Variable(s) وكتابة القيمة الافتراضية للمتوسط (10) في test Value كما هو موضح في الصورة التالية:



ويمكننا تعديل مجال الثقة من خلال الزر Options الذي يظهر على الشكل التالي:



بحيث يمكن تعديل مجال الثقة من 95% إلى 99% مثلا حسب حاجة المستخدم من خلال كتابة 99 بدلا من 95 في الخانة الخاصة بذلك.

أما المجال Missing Values فهو يتيح امكانية اختيار طريقة التعامل مع القيم المفقودة في حالة اختبار T لعدة متغيرات في آن واحد (اختيار عدة متغيرات في المجال Test Variable(s))، بحيث يمكن التأشير على أحد الخيارين:

-Exclude cases analysis by analysis: من خلاله يتم استبعاد القيم المفقودة للمتغير المعني باختبار t في حالة احتوائه على قيم مفقودة، بينما يتم استخدام كل قيم المتغير الآخر الذي لا يحتوي على قيم مفقودة في حساب قيمة الاختبار، أي أنه يمكن اجراء الاختبار لمتغيرين أو أكثر يمكن أن تختلف فيها حجم العينة (احتواء احد المتغيرات مثلا على قيم مفقودة).

- Exclude cases Listwise: يتم من خلال هذا الخيار استبعاد عدد معين من الحالات يساوي عدد القيم المفقودة في احد المتغيرات بحيث لايأخذ في الاعتبار قيم من المتغيرات المقابلة للقيم المفقودة في أحد المتغيرات في حساب قيم t، وبالتالي يكون حجم العينة هو نفسه لكل المتغيرات المدرجة في التحليل بالرغم من احتواء بعضها على كل القيم.

أما في حالة المتغير الواحد فلا يستدعي التأشير على أحد الخيارين لأنه لا يحمل أي جدوى، فالاختبار يكون على القيم الموجودة دون المفقودة تلقائيا.

وبعد اتمام الخطوات المذكورة ننقر على الزر Ok فتظهر لنا نتائج الاختبار على الشكل التالي:

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Note	9	10,44	4,586	1,529

جدول المؤشرات الاحصائية الوصفية

الانحراف المعياري σ

$\sigma / \sqrt{n}$

One-Sample Test

	Test Value = 10					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Note	0,291	8	0,779	0,444	-3,08	3,97

قيمة t المحسوبة

درجة الحرية (n-1)

القيمة الاحتمالية

الفرق بين المتوسطين $\overline{X} - \mu$

مجال الثقة للفرق بين المتوسطين

ومن خلال جدول اختبار t، نلاحظ أن القيمة المحسوبة للاختبار تساوي 0.291 وهي محسوبة وفق العلاقة السالفة الذكر بحيث:

$$t = \frac{10.44 - 10}{4.586 / \sqrt{9}} = \frac{0.444}{1.528} = 0.291$$

ويتم مقارنتها بالقيمة المجدولة لـ t ستودنت عند مستوى معنوية 5% التي تقدر بـ 2.306. وبما أن القيمة المحسوبة للاختبار أقل من القيمة المجدولة عند مستوى 5%، فهذا يقودنا إلى قبول فرضية العدم القائلة بأن متوسط العينة يساوي إلى 10، ويمكن الاستعانة بالقيمة الاحتمالية (Sig.(2-tailed في الجدول ومقارنتها بالمستوى 0.05 بحيث كلما كانت أكبر نقبل فرضية العدم H0 والعكس صحيح.

3-3. اختبار t لعينتين مستقلتين Independent simple t test:

يقوم هذا الإجراء على المقارنة بين متوسطي عينتين مستقلتين تحت الفرضية:

$$H_0 : \overline{X}_1 = \overline{X}_2$$
$$H_1 : \overline{X}_1 \neq \overline{X}_2$$

باستعمال اختبار t ستودنت وفق العلاقة التالية:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sum(X_{i1} - \overline{X}_1)^2 + \sum(X_{i2} - \overline{X}_2)^2}{N_1 + N_2 - 1}} \sqrt{\frac{1}{N_1} + \frac{1}{N2}}}$$

بحيث N_1 و N_2 عدد المشاهدات في العينة الأولى والثانية على التوالي. وX_i تمثل القيم المشاهدة للعينة في الحالة i.

مثال (4-8):

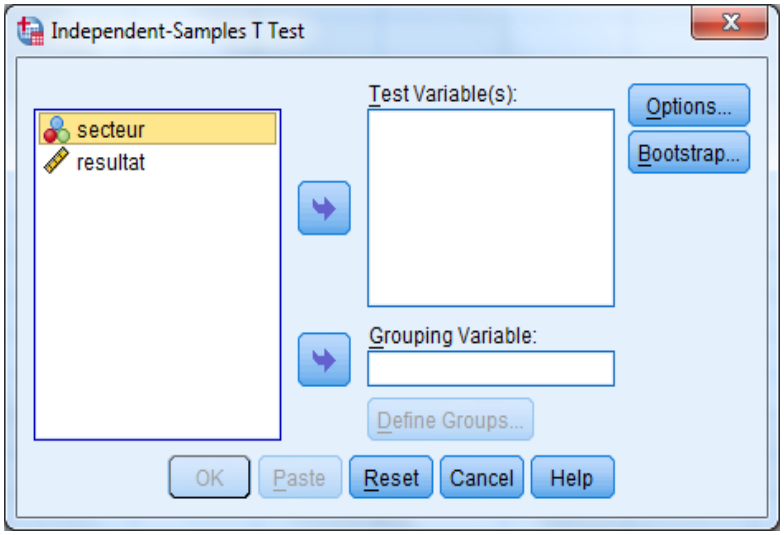
لدينا البيانات الخاصة بأداء الاطارات الوسيطية في شركتين تعمل الاولى في قطاع تجهيزات النقل (32 اطار) والثانية تعمل في انتاج المنتجات الالكترونية (34 اطار)، والناتجة عن اختبار تقييمي لمختلف الخصائص الادارية. وتمثلت هذه البيانات في مجموع النقاط المحصلة في الامتحان والموضحة في الجدول التالي:

cadre	Transport	Electronique
1	200	210
2	198	208
3	177	183
4	210	223
5	214	227
6	189	197
7	168	172
8	168	172
9	178	184
10	191	199
11	213	225
12	195	204
13	195	204
14	184	191
15	201	211
16	170	174
17	206	217

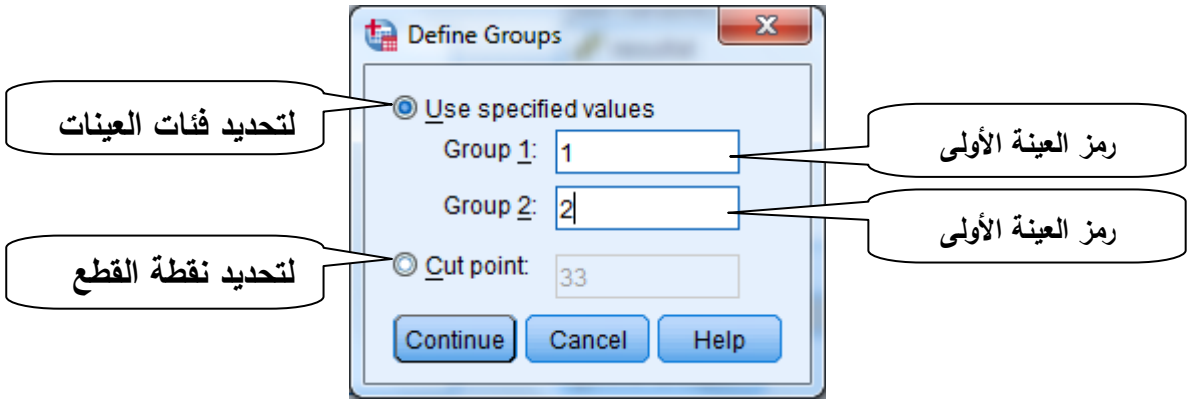
cadre	Transport	Electronique
18	159	162
19	188	196
20	204	215
21	183	190
22	129	155
23	221	235
24	178	184
25	206	218
26	147	147
27	212	225
28	178	183
29	182	189
30	150	150
31	192	201
32	174	179
33		199
34		186

قبل تعبئة هذه البيانات في برنامج SPSS، يجب الاشارة إلى أنه يجب تعريف متغيرين أساسين هما متغير النتائج (Resultat) ومتغير القطاع (Secteur)، ومن ثم تعبئة البيانات الخاصة بنتائج التقييم في المتغير الاول، بينما المتغير الثاني فيأخذ القيمة 1 لقطاع النقل و2 لقطاع الالكترونيك. وهي الطريقة الصحيحة لتعبئة البيانات التي تمكننا من اجراء اختبار وجود فروقات معنوية بين اداء الاطارات الوسيطية في القطاعين.

لإجراء الاختبار نختار من قائمة Analyze الأمر Compare Maens ثم نختار الأمر Independent-Simple T Test، فيظهر لدينا صندوق الحوار الخاص به على الشكل التالي:



يتم نقل متغير النتائج (Resultat) إلى المجال Test Variable(s) والمتغير Secteur إلى المجال Groupinge Variable، ليظهر لدينا الزر Define Groups لتعريف الرموز الدالة عن العينتين وفي حالتنا القطاعين 1 و 2 بحيث يظهر لدينا الشكل التالي:



في هذه الحالة يكفي كتابة الرقم 1 في المجال Group 1 و الرقم 2 في المجال Group 2 والنقر على الزر Contenue. بينما في حالة متغير التجزئة المعبر عن العينات يتم اللجوء إلى التأشير على الخيار Cut Point بتحديد القيمة الدنيا للعينة الثانية، فمثلا لو انشأنا متغير رقمي يعبر عن رقم العامل (Cadre) من 1 إلى 66 بنفس الترتيب بحيث يكون آخر عامل تابع للقطاع الأول نكتب الرقم 33 لتعريف العينة الثانية التي تكون فيها قيم المتغير الجديد أكبر أو يساوي 33، بينما الفئة الثانية فيهي التي تقابل قيم المتغير (Cadre) التي تقل تماما عن 33.

وبالنقر على الزر Contenue والعودة إلى صندوق حوار Independent-Simple T Test، والنقر على الزر OK نحصل على النتائج التالية:

Group Statistics

	secteur	N	Mean	Std. Deviation	Std. Error Mean
resultat	1	32	186,25	21,295	3,764
	2	34	194,56	22,670	3,888

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
resultat	Equal variances assumed	,241	,625	-1,532	64	,130	-8,309	5,422	-19,141	2,523
	Equal variances not assumed			-1,535	64,000	,130	-8,309	5,412	-19,120	2,502

من خلال الجدول نلاحظ أن اختبار الفرضية في هذه الحالة يجرى في حالتين، تفترض الأولى توفر شرط تجانس تباين العينتين ($\sigma_1^2 = \sigma_2^2$)، بينما يفترض في الحالة الثانية عدم تجانس التباين ($\sigma_1^2 \neq \sigma_2^2$). وقد قدرت قيمة احصائية t في الحالة الأولى بـ 1.532- بينما قدرت في الحالة الثانية بـ 1.535-، وهي في كلتا الحالتين أقل من القيمة المجدولة عند مستوى معنوية 5% ($t_{64}^{0.025} \approx 2.0003$)، وهذا يقودنا إلى قبول فرضية العدم القائلة بأنه لا توجد فروقات معنوية بين متوسطي العينتين أي أن $\overline{X}_1 = \overline{X}_2$ بالرغم من وجود فرق بين المتوسطين بمقدار 8.309.

بينما تشير نتائج اختبار Levene لتجانس التباين إلى أن القيمة الاحتمالية للإحصائية المحسوبة للاختبار تقدر بـ 0.625 وهي أكبر من 0.05، مما يقودنا إلى قبول فرضية العدم القائلة أن تباين العينتين متساوي (تحقق فرضية العدم) أي $\sigma_1^2 = \sigma_2^2$.

3-4. اختبار تساوي متوسطي عينة مزدوجة Paired-Simples T Test:

يفيد هذا الإجراء في مقارنة متوسط يتعلق بنفس العينة لكن بمقياس مزدوج وهو كثير الاستعمال مجال العلوم الطبية، ويتعلق الأمر ببيانات العينة الواحدة قبل حدث معين وبعده، فمثلا مراقبة مقياس ضغط الدم لدى عينة من الأشخاص المصابين باضطرابات في الضغط والأمراض القلبية قبل وبعد الخضوع لعلاج معين، أو مراقبة انتاجية مجموعة من العمال في مؤسسة انتاجية قبل وبعد تلقي دورة تكوينية مثلا. ويكون التعامل مع هذه الحالة مختلف إلى حد ما عن التعامل مع بيانات العينات المستقلة.

مثال (4-9):

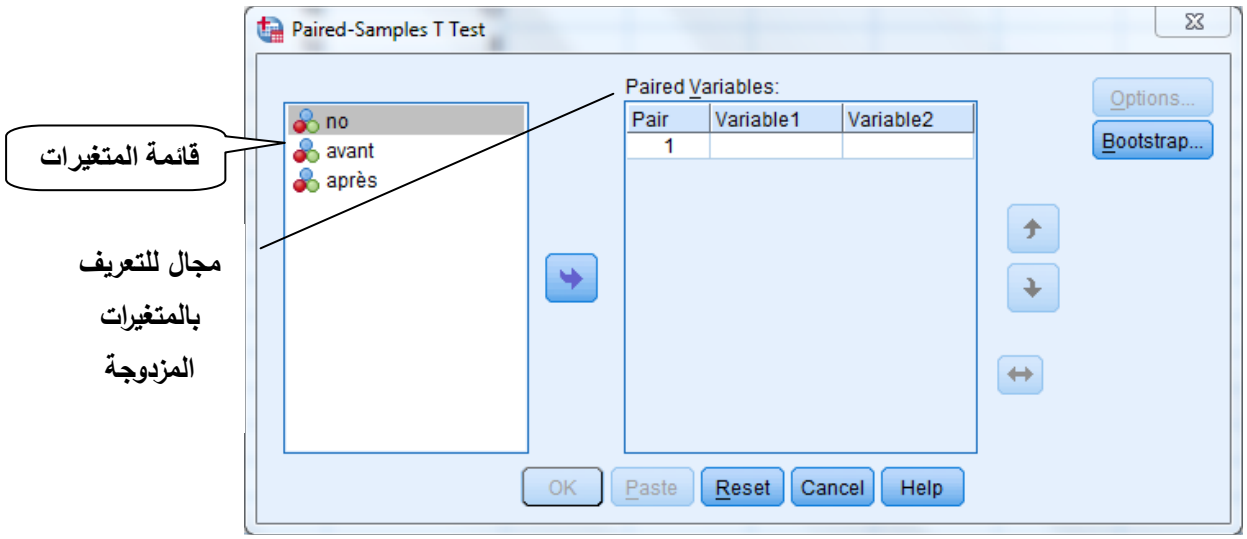
في دراسة لعينة عشوائية متكونة من 15 عامل من عمال شركة Gescom الذين تلقوا دورة تكوينية في مجال التغليف، بحيث تمت ملاحظة عدد القطع التي تم تغليفها من طرف كل عامل خلال فترة 3 أسابيع قبل وبعد اجراء الدورة التكوينية، والبيانات الخاصة بهذه الدراسة مبينة في الجدول التالي:

الرقم	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
قبل الدورة	15	13	8	9	7	12	11	12	11	9	10	12	11	7	12
بعد الدورة	17	16	10	9	9	13	14	15	14	11	14	11	13	10	13

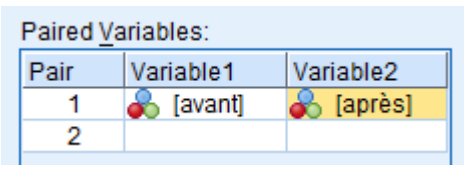
ويتم تعبئة البيانات في البرنامج كما هي في ثلاث أعمدة على خلاف المثال السابق، كما هي موضحة في الصورة التالية:

no	avant	après
1	15	17
2	13	16
3	8	10
4	9	9
5	7	9
6	12	13
7	11	14
8	12	15
9	11	14
10	9	11
11	10	14
12	12	11
13	11	13
14	7	10
15	12	13

ويتم اجراء تحليل المتوسطات في هذه الحالة باختيار الأمر Paired-Simples T Test من قائمة Compare Means من القائمة Analyze، فيظهر لدينا صندوق حوار على الشكل التالي:



بحيث يتم نقل كل زوج من المتغيرات إلى المجال Paired Variables وفق الترتيب المتغير الاول قبل وقوع الحدث والمتغير الثاني بعد وقوع الحدث، ويمكن تغيير المتغيرات باستخدام الاسهم الموجودة داخل الصندوق.



ويمكن اختيار مجال الثقة من خلال الزر Option، ثم النقر على الزر OK، فتظر لنا نتائج الاختبار على النحو التالي:

- **جدول المؤشرات الاحصائية الوصفية:**

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	avant	10,60	15	2,261	0,584
	après	12,60	15	2,501	0,646

- **معامل الارتباط بين المتغيرات:**

		N	Correlation	Sig.
Pair 1	avant & après	15	0,854	0,000

- **نتائج الاختبار الاحصائي للفرق بين متوسطي العينتين:**

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	avant - après	-2	1,309	0,338	-2,725	-1,275	-5,916	14	0,000

$\bar{d} = \bar{X}_1 - \bar{X}_2$ $\quad \sigma_d$ $\quad \frac{\sigma_d}{\sqrt{n}}$ $\quad t - student$

من خلال هذه النتائج يتبين أنه يوجد ارتباط بين المتغيرين يقدر بـ 0.854 وهو معنوي من الناحية الاحصائية عند مستويات متدنية حيث قدرت القيمة الاحتمالية (.Sig) بأقل من 0.0001، بينما قدر الفرق بين متوسطي العينتين بمقدار 2−، ونتيجة اختبار t المحسوبة وفق العلاقة التالية:

$$t = \frac{\bar{d}}{\sigma_d / \sqrt{n}} = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{X}_1 - \bar{X}_2} / \sqrt{n}} = -5.916 \neg \qquad t^{\alpha/2}_{(n-k-1)}$$

ان القيمة المحسوبة تساوي 5.916− وهي أكبر من القيمة المجدولة عند مستوى 5% ($t^{0.025}(14)=2.14$) مما يقودنا إلى رفض فرضية العدم القائلة بأنه لا يوجد اختلاف معنوي بين متوسطي العينتين ونقبل الفرضية البديلة $H_1 : \bar{X}_1 \neq \bar{X}_2$، مما يدل على أن انتاجية الاطارات تحسنت نتيجة تلقيهم التكوين نظرا لوجود فرق سالب معنوي بين المتوسطين.

ومن هنا يتضح أنه يمكن التفصيل في الفرضية المختبرة بحيث يتم الاقرار بوجود اختلاف معنوي بين متوسطي عينة مزدوجة أو عينتين مستقلتين بناء على نتائج اختبار t، مع الاشارة إلى ان المتوسط الأول أكبر من المتوسط الثاني أو العكس بناء على اشارة الفرق بينهما أي أن الفرضية H1 تصبح على الشكل التالي:

- $H_1 : \bar{X}_1 > \bar{X}_2$ إذا كانت الاحصائية المحسوبة أكبر من تلك المجدولة عند مستوى معنوية 5% الفرق بين المتوسطين موجب؛

- $H_1 : \overline{X}_1 < \overline{X}_2$ إذا كانت الاحصائية المحسوبة أكبر من تلك المجدولة عند مستوى معنوية 5% الفرق بين المتوسطين سالب.

3-5. اختبار تساوي عدة متوسطات:

أولا: تحليل التباين بمعيار واحد One-way ANOVA:

في هذه الحالة يصبح اختبار t غير قابل للاستعمال على خلاف الحالات المذكورة سابقا، بحيث يتعلق الأمر بالمقارنة بين أكثر من متوسطين (أكثر من عينتين)، من خلال اختبار فرضية العدم القائلة بان متوسطات العينات متساوي ضد فرضية عدم تساوي المتوسطات أي:

$$H_0 : \overline{X}_1 = \overline{X}_2 = ... = \overline{X}_k$$
$$H_1 : \overline{X}_1 \neq \overline{X}_2 \neq ... \neq \overline{X}_k$$

وهي عبارة عن اسلوب احصائي يعرف بتحليل التباين بمعيار واحد (One-Way ANOVA)، بإعتبار متغير كمي يعتبر تابع لمتغير مستقل كيفي يعبر عن عدد معين من العينات، ويعتمد هذا الاسلوب على البحث عن مصادر الاختلاف في المتوسطات من خلال جدول تحليل التباين، أين يتم تجزئة التباين الكلي للمتغير التابع إلى جزئين تباين بين المجموعات والذي يرجع مصدره إلى الفروقات في المتوسطات، وتباين ضمن المجموعات والذي لا يعرف مصدره (البواقي)، ويقارن بين تباين الجزئين باستعمال اختبار فيشر.

ويشترط اجراء هذا الأسلوب توفر ثلاث شروط أساسية وهي:

- شرط تجانس تباين المتغير التابع لكل مجتمع من مجتمعات المتغير المستقل ؛
- شرط التوزيع الطبيعي للمتغير التابع لكل مجتمع من مجتمعات المتغير العامل (المستقل)؛
- ان تكون العينات من كل مجتمع من مجتمعات المتغير العامل عشوائية.

وعدم توفر هذه الشروط وعلى وجه الخصوص الشرط الثاني والثالث، فإنه يتوجب على المستخدم استعمال الطرق غير المعلمية والتي لا تشترط الشروط المذكورة.

ملاحظة: ان الفرضية البديلة تنص على وجود أحد المتوسطات يختلف عن باقي المتوسطات وليس بالضرورة اختلاف كل المتوسطات فيما بينها.

إلى جانب الاختبار الاحصائي المذكور، يمكن استخدام المقارنات البعدية Post Hoc لمقارنة متوسطات المتغير التابع لكل زوجين من الفئات على حدة، فإذا كان لدينا ثلاث فئات سيتم المقارنة بين كل زوج منها كما يلي: الأولى-الثانية؛ الثانية-الثالثة والأولى-الثالثة. وفي هذه الحالة يتم استخدام اختبار

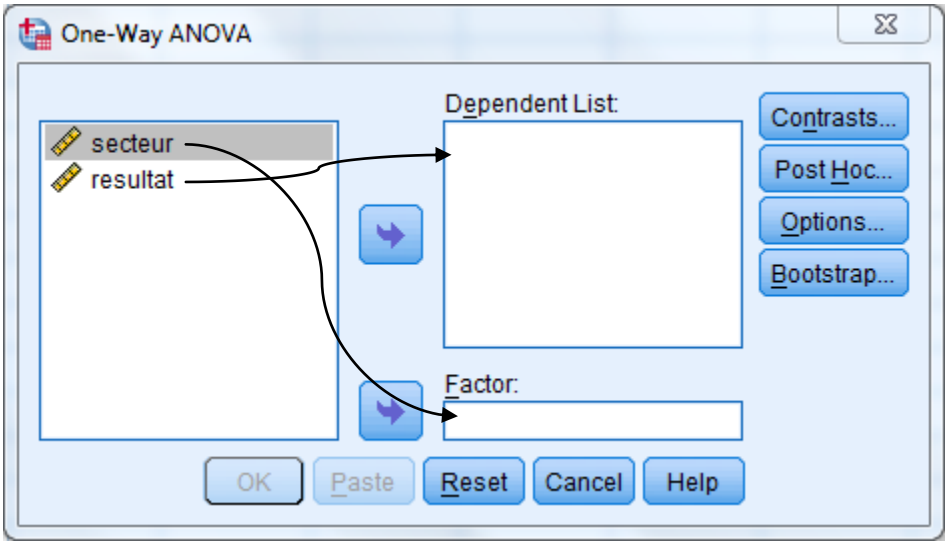
t للمقارنة بين كل زوج من المتوسطات وفق الطريقة السابقة. كما يمكن الاستعانة بمخططات بيانية لتحليل الفروقات بين المتوسطات كمخطط الصندوق والمنحنى البياني للمتوسطات.

ولتوضيح خطوات هذا الإجراء باستخدام برنامج SPSS نستعين بالمثال المدروس سابقا في تحليل متوسطي عينتين مستقلتين (3-2) مع اضافة عينة من قطاع انتاج التجهيزات الكهرومنزلية، بحيث يصبح لدينا ثلاث عينات مستقلة. كما هو موضح في الجدول التالي:

no	transport	electronique	electroménagé
1	200	210	202
2	198	208	256
3	177	183	225
4	210	223	202
5	214	227	246
6	189	197	168
7	168	172	204
8	168	172	181
9	178	184	164
10	191	199	168
11	213	225	170
12	195	204	205
13	195	204	234
14	184	191	189
15	201	211	202
16	170	174	187
17	206	217	208
18	159	162	198
19	188	196	233
20	204	215	222
21	183	190	239
22	129	155	165
23	221	235	172
24	178	184	172
25	206	218	163
26	147	147	210
27	212	225	204
28	178	183	200
29	182	189	196
30	150	150	203
31	192	201	-
32	174	179	-
33	-	199	-
34	-	186	-

تذكير: يتم تعبئة هذه البيانات في متغيرين: المتغير Secteur: متغير كيفي يأخذ القيم 1، 2 و3 للقطاعات الثلاثة على الترتيب، تقابلها قيم المتغير Resultat وهي نتيجة الاختبار للإطارات من كل قطاع.

بحيث نختار من قائمة Compare Means من قائمة Analyze الخيار One-Way Anova فيظهر لنا صندوق حوار على النحو التالي:



وبعد نقل المتغيرين إلى المجال Dependent List و Facteur كما هو مبين في الصورة، يمكن النقر على الزر OK لاجراء الاختبار فينتج لدينا النتائج التالية:

resultat

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2834,657	2	1417,329	2,612	0,079
Within Groups	50459,582	93	542,576		
Total	53294,240	95			

يظهر من خلال الجدول ثلاث مجاميع اساسية تستخدم لحساب القيمة الاحصائية للاختبار وهي:

- مجموع مربعات الفروق بين المجموعات Between Groups: $SSTR = \sum_{i=1}^{k} n_i \left(\bar{Y}_{i\bullet} - \bar{Y}_{\bullet\bullet}\right)^2$

- مجموع مربعات الفروق داخل المجموعات Within Groups: $SSE = \sum_{i=1}^{k} \sum_{j=1}^{n} \left(Y_{ij} - \bar{Y}_{i\bullet}\right)^2$

- المجموع الكلي لمربعات الفروق Total: $SSTO = \sum_{i=1}^{k} \sum_{i=1}^{n} \left(Y_{ij} - \bar{Y}_{\bullet\bullet}\right)^2$

ودرجات الحرية: $(k-1=3-1=2)$، $\left(\sum_{i=1}^{k} n_i - k = 96-3=93\right)$، $\left(\sum_{i=1}^{k} n_i - 1 = 96-1=95\right)$

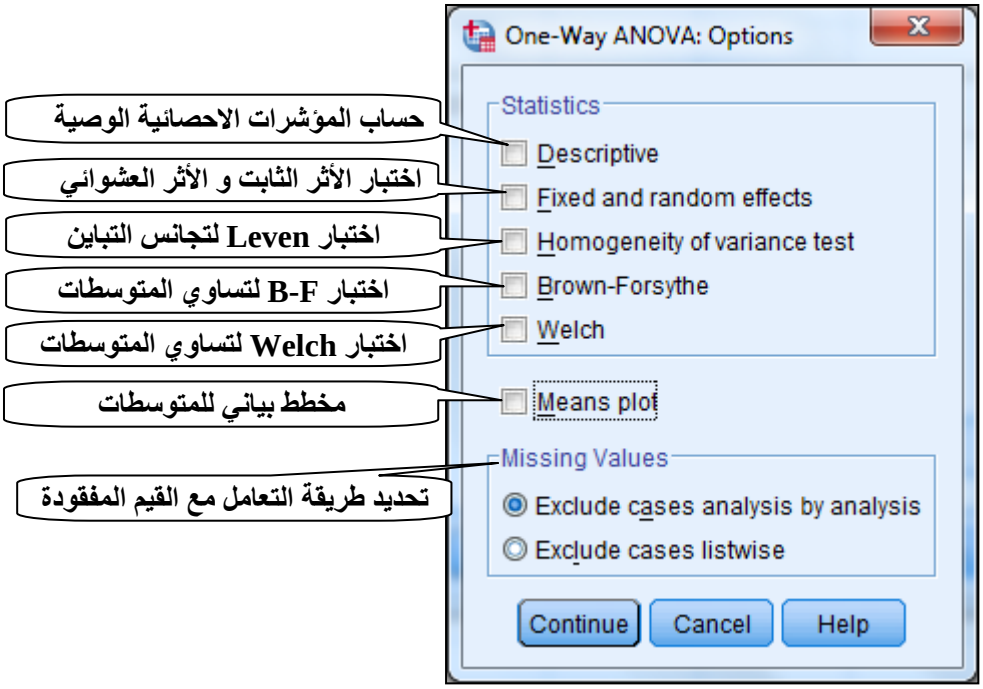
ويتم حساب قيمة فيشر وفق العلاقة التالية:

$$F = \frac{SSTR/k-1}{SSE/\sum_{i=1}^{n} n_i - 1} = \frac{2834.657/2}{50459.582/93} = \frac{1417.329}{542.576} = 2.612$$

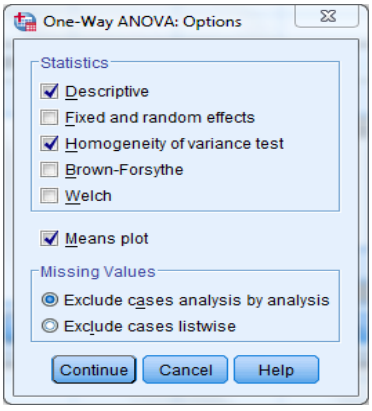
ومن خلال هذه النتائج يمكن الاقرار بتساوي المتوسطات وعدم وجود اختلافات معنوية بينها، لأن القيمة المحسوبة لإحصائية فيشر أقل من القيمة المجدولة عند مستوى معنوية 5% (3.40>2.92)، وكذلك القيمة الاحتمالية Sig.=0.07 وهي أكبر من 0.05.

➢ **المؤشرات الاحصائية الوصفية واختبار الفرضيات:**

كما يسمح هذا الأمر باختبار فرضية تجانس التباين وعرض مخطط بياني يساعد في التحليل والاجابة على التساؤل: هل توجد علاقة خطية بين متوسطات العينات المدروسة؟، بحيث يتيح الزر Option مجموعة من الخيارات كما يظهر في الصورة التالية:



ونشير هنا إلى أن اختبار Brown–Forsuthe و اختبار Welch هي اختبارات قوية (Robust) لتساوي المتوسطات تستخدم في حالة عدم تحقق فرضية تجانس التباين.



وفي هذه الحالة نكتفي بالتأشير على الخيارات التالية: Descriptive، Homogeneity of variance test و Means Plot كما توضحه الصورة المقابلة.

وباتمام العملية بالنقر على Continue و Ok في صندوق حوار One way anova ينتج لدينا النتائج التالية:

- جدول المؤشرات الاحصائية الوصفية:

Descriptives

resultat

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
transport	32	186,25	21,295	3,764	178,57	193,93	129	221
électronique	34	194,56	22,670	3,888	186,65	202,47	147	235
electroménagé	30	199,60	25,892	4,727	189,93	209,27	163	256
Total	96	193,36	23,685	2,417	188,57	198,16	129	256

- نتائج اختبار تجانس التباين:

Test of Homogeneity of Variances

resultat

Levene Statistic	df1	df2	Sig.
0,446	2	93	0,642

القيمة الاحتمالية لاختبار **Leven** أكبر من **0.05**، وبالتالي نقبل فرضية تجانس التباين

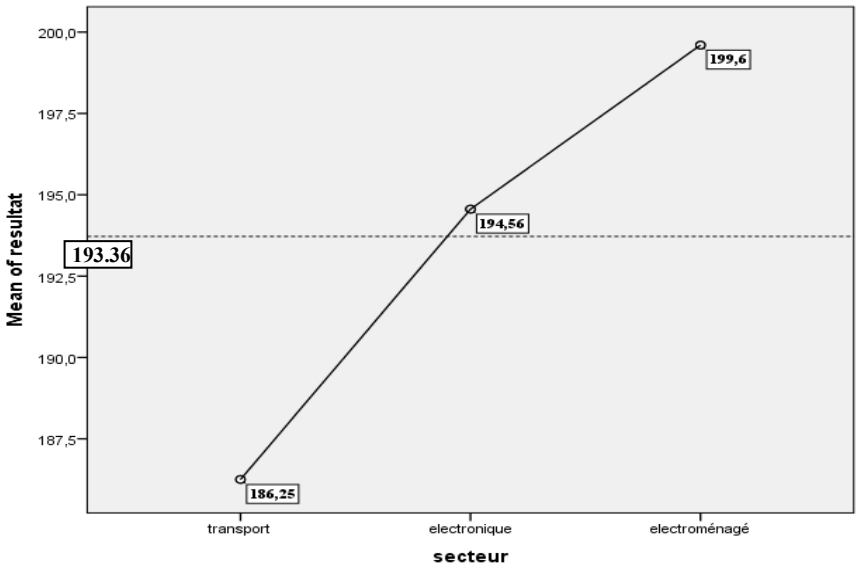
- جدول تحليل التباين:

ANOVA

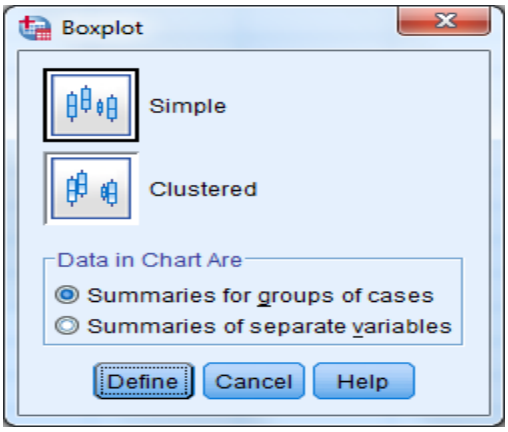
resultat

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2834,657	2	1417,329	2,612	0,079
Within Groups	50459,582	93	542,576		
Total	53294,240	95			

- التمثيل البياني للمتوسطات:

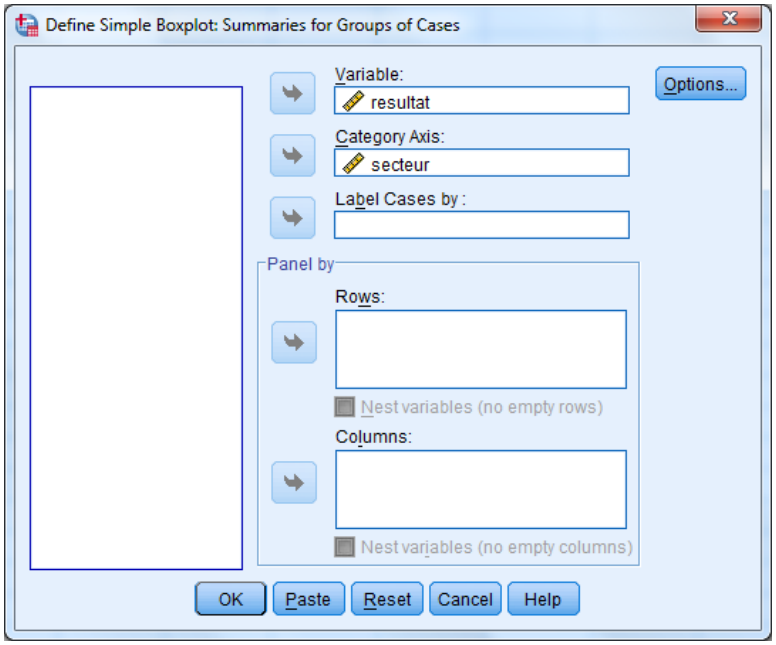


كما يمكن الاستعانة بمخطط الصندوق لتحليل توزيع كل عينة والتحقق من توفر فرضية ان كل عينة مسحوبة من مجتمع يتبع التوزيع الطبيعي، وذلك باختيار من قائمة الأدوات الأمر Graphs ثم اختيار الخيار Legacy Dialogs ثم Boxplot فيظهر لدينا صندوق حوار Boxplot المقابل:



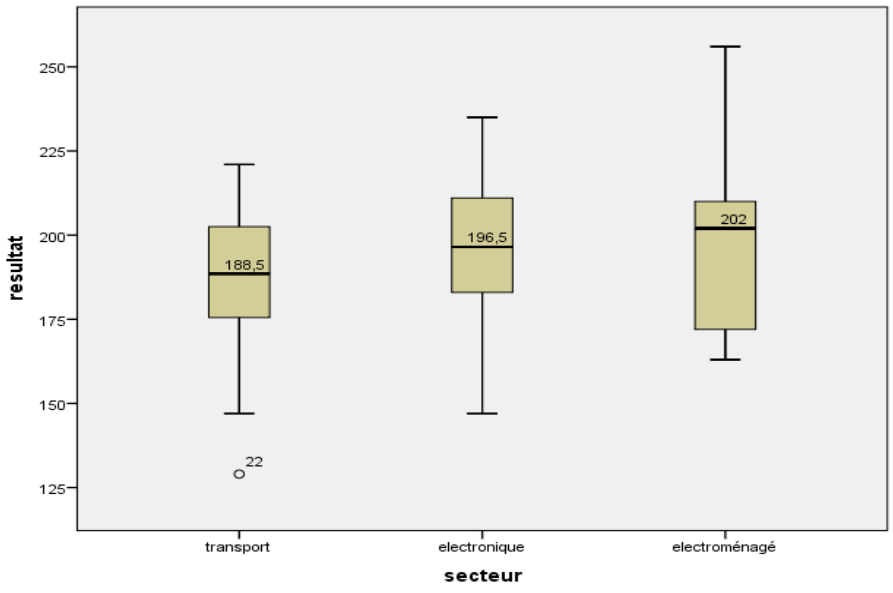
أو اجراء الاختبار باستعمال الأمر Explore وفق المنهجية المذكورة سابقا.

بحيث نختار الزر Simple ثم ننقر على Define ليظهر لنا صندوق حوار الخيارات كما يلي:



ونقوم بنقل المتغيرات إلى المجالات المتوفرة كما هو موضح في الصورة المقابلة، وهذا بمثابة أمر البرنامج باعداد مخطط الصندوق للمتغير التابع حسب المتغير العامل في معلم واحد من أجل المقارنة.

وباتمام العملية نتحصل على المخطط التالي:



➢ **المقارنات المتعددة Post Hoc:**

كما يوفر البرنامج من خلال صندوق حوار One-way Anova، عدة طرق للمقارنات المتعددة والتوسع أكثر في مقارنة المتوسطات في حالة توفر فرضية تجانس التباين أو عدم توفرها على حد سواء، وتجدر الاشارة إلى أن هذا الإجراء نلجأ اليه من أجل التحقق من فرضية وجود فروقات معنوية بين كل زوج من العينات في حالة قبول الفرضية البديلة في نتائج تحليل التباين (جدول anova).

ولتوضيح طريقة القيام بهذه العملية نستعين ببيانات المثال التالي:

مثال (4-10):

لدينا البيانات الخاصة بنتائج امتحان البكالوريا في مادة الرياضيات لعينة عشوائية من طلبة السنة ثالثة ثانوي من ثلاثة شعب: آداب ولغات أجنبية، علوم تجريبية، تسيير واقتصاد. ونريد اختبار تساوي متوسط العلامة المحصلة في مادة الرياضيات حسب الشعب. والبيانات موضحة في الجدول التالي:

no	note	filière	no	note	filière	no	note	filière	no	note	filière
1	13	آداب ولغات أجنبية	24	10	آداب ولغات أجنبية	47	12	علوم تجريبية	70	16	تسيير واقتصاد
2	10	آداب ولغات أجنبية	25	9	آداب ولغات أجنبية	48	13	علوم تجريبية	71	9	تسيير واقتصاد
3	12	آداب ولغات أجنبية	26	9	آداب ولغات أجنبية	49	11	علوم تجريبية	72	8	تسيير واقتصاد
4	9	آداب ولغات أجنبية	27	15	آداب ولغات أجنبية	50	17	علوم تجريبية	73	13	تسيير واقتصاد
5	8	آداب ولغات أجنبية	28	10	آداب ولغات أجنبية	51	8	علوم تجريبية	74	8	تسيير واقتصاد
6	10	آداب ولغات أجنبية	29	13	آداب ولغات أجنبية	52	10	علوم تجريبية	75	12	تسيير واقتصاد
7	4	آداب ولغات أجنبية	30	9	آداب ولغات أجنبية	53	13	علوم تجريبية	76	12	تسيير واقتصاد
8	12	آداب ولغات أجنبية	31	13	علوم تجريبية	54	13	علوم تجريبية	77	8	تسيير واقتصاد
9	10	آداب ولغات أجنبية	32	11	علوم تجريبية	55	10	علوم تجريبية	78	9	تسيير واقتصاد
10	9	آداب ولغات أجنبية	33	11	علوم تجريبية	56	15	علوم تجريبية	79	13	تسيير واقتصاد
11	7	آداب ولغات أجنبية	34	17	علوم تجريبية	57	13	علوم تجريبية	80	11	تسيير واقتصاد
12	12	آداب ولغات أجنبية	35	15	علوم تجريبية	58	16	علوم تجريبية	81	8	تسيير واقتصاد
13	7	آداب ولغات أجنبية	36	12	علوم تجريبية	59	13	علوم تجريبية	82	13	تسيير واقتصاد
14	9	آداب ولغات أجنبية	37	11	علوم تجريبية	60	15	علوم تجريبية	83	11	تسيير واقتصاد
15	9	آداب ولغات أجنبية	38	9	علوم تجريبية	61	10	تسيير واقتصاد	84	12	تسيير واقتصاد
16	11	آداب ولغات أجنبية	39	11	علوم تجريبية	62	14	تسيير واقتصاد	85	9	تسيير واقتصاد
17	11	آداب ولغات أجنبية	40	10	علوم تجريبية	63	14	تسيير واقتصاد	86	9	تسيير واقتصاد
18	11	آداب ولغات أجنبية	41	19	علوم تجريبية	64	8	تسيير واقتصاد	87	12	تسيير واقتصاد
19	8	آداب ولغات أجنبية	42	15	علوم تجريبية	65	9	تسيير واقتصاد	88	7	تسيير واقتصاد
20	15	آداب ولغات أجنبية	43	12	علوم تجريبية	66	7	تسيير واقتصاد	89	12	تسيير واقتصاد
21	9	آداب ولغات أجنبية	44	15	علوم تجريبية	67	12	تسيير واقتصاد	90	12	تسيير واقتصاد
22	9	آداب ولغات أجنبية	45	12	علوم تجريبية	68	11	تسيير واقتصاد			
23	10	آداب ولغات أجنبية	46	8	علوم تجريبية	69	11	تسيير واقتصاد			

باجراء اختبار فرضية تساوي متوسطات العينات، بينت نتائج تحليل التباين وفق الطريقة المذكورة سابقا تحصلنا على النتائج التالية:

Descriptives

note

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
تسيير واقتصاد	30	10,67	2,324	,424	9,80	11,53	7	16
آداب ولغات أجنبية	30	10,00	2,304	,421	9,14	10,86	4	15
علوم تجريبية	30	12,67	2,695	,492	11,66	13,67	8	19
Total	90	11,11	2,675	,282	10,55	11,67	4	19

Test of Homogeneity of Variances

note

Levene Statistic	df1	df2	Sig.
,800	2	87	0,452

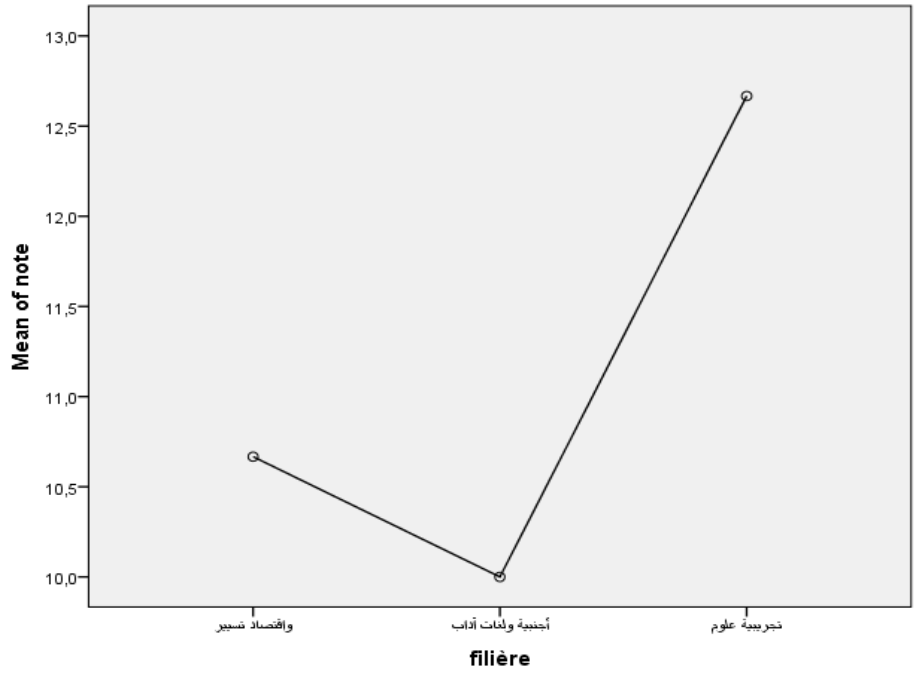
القيمة الاحتمالية 0.452 أكبر من 0.05 وبالتالي نقبل فرضية تجانس التباين

ANOVA

note

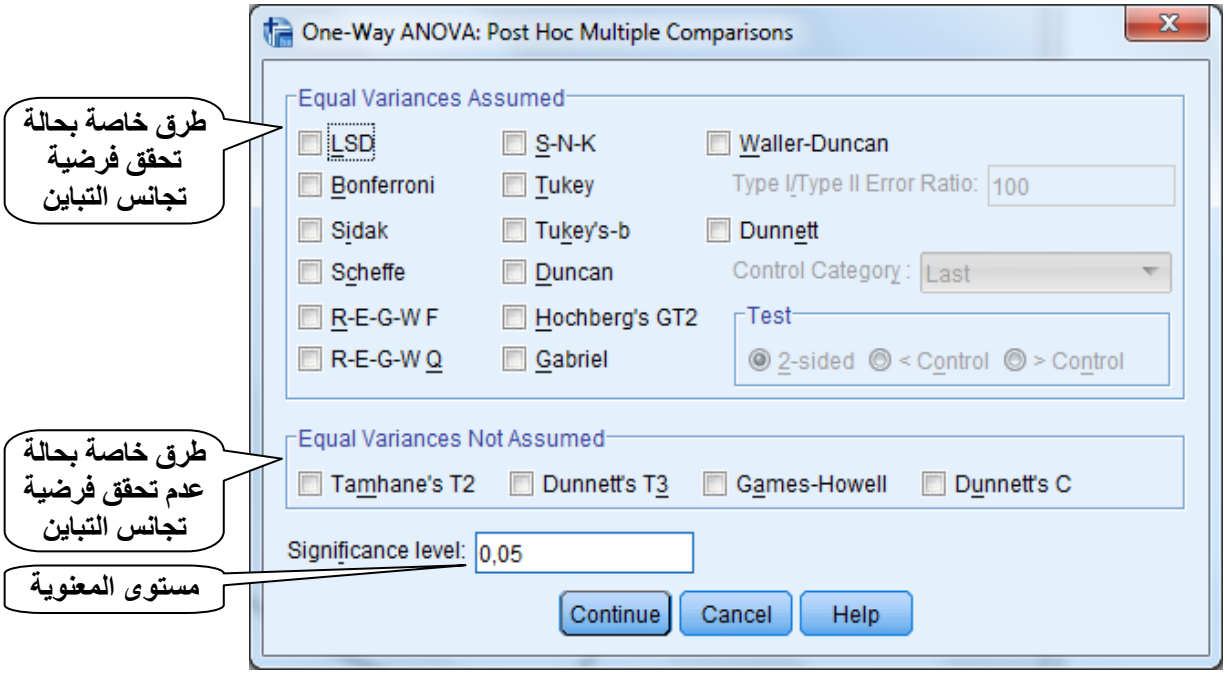
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	115,556	2	57,778	9,642	0,000
Within Groups	521,333	87	5,992		
Total	636,889	89			

القيمة الاحتمالية 0.000 أقل من 0.05 وبالتالي نقبل فرضية وجود فروقات معنوية في متوسطات علامات الطلبة حسب الشعب.



من خلال هذه النتائج يمكن الاقرار بوجود اختلاف (فرق) معنوي في تباين العلامات المحصلة في مادة الرياضيات، وان نسبة كبيرة من هذا التباين راجع إلى الاختلاف في المتوسطات حسب الشعب، وفي هذه الحالة نطرح السؤال: هل جميع المتوسطات مختلفة عن بعضها البعض؟ وللإجابة عليه سنعتمد على اجراء المقارنات المتعددة وفق طرق تتحقق فيها فرضية تجانس تباين العينات.

نختار من صندوق حوار One-way ANOVA الخيار Post Hoc فيظهر لنا صندوق الحوار التالي:



وبما أن فرضية تجانس التباين في هذه الحالة محققة، سنختار طريقة أو أكثر من بين الطرق الواقعة في المجال Equal Variances Assumed، بينما في حالة عدم تحقق الفرضية يتم اختيار طريقة واحدة فقط من بين الطرق المتوفرة، وسنركز على ثلاث طرق للمقارنات المتعددة نذكرها فيما يلي:

LSD (Least Significant Difference): وهي طريقة تعتمد على مقارنة متوسط كل زوج من العينات (الشعب) لعدد 2k زوج، واختبار فرضية تساوي المتوسطات باستعمال احصائية t-Student بدرجة حرية N-k.

$$LSD_{i-j} = \frac{\mu_i - \mu_j}{\sigma_{\mu_i - \mu_j}}$$

Bonferroni: وهي طريقة شبيهة بـ LSD، والاختلاف يكون في حساب قيمة الاختبار ومجال الثقة؛ بحيث يتم ضرب مستوى المعنوية المقدر في طريقة LSD بعدد الاختبارات التي تم اجراؤها أي مقدار (k*(k-1))/2 في مثالنا (3*(3-1))/2=3.

Dunnett: هو عبارة عن اختبار احصائي يستعمل في حالة وجود عينة (مجموعة) السيطرة Control Category يتم مقارنة باقي المجموعات بها، ولهذا نجد أن البرنامج يفتح المجال لتحديد هذه الأخيرة

باختيار الأولى First أو الأخيرة Last، بالاضافة إلى اختيار نوع الاختبار من طرف واحد أو من طرفين. ويتم حساب قيمة الاختبار بطريقة مشابهة لاختبار ستودنت، ويتم قبول الفرضية البديلة (وجود فروقات معنوية بين متوسطين) إذا كانت القيمة الاحتمالية المحسوبة أقل من 0.05.

ويمكن للباحث استخدام أي طريقة من الطرق المتوفرة حسب الحاجة وفي الغالب لا توجد اختلافات كبيرة في القرار الخاص باختبار فرضية تساوي المتوسطات لكل زوج من العينات، وفي هذه الحالة نختار الطرق الثلاثة بالتأشير عليها.

ملاحظة: بالرغم من أنه ليس لدينا مجموعة سيطرة بخصوص اختبار Dunnett الا اننا اخترنا المجموعة الاخيرة Last من أجل توضيح طريقة اجراء الاختبار فقط.

وباتمام العملية يظهر لدينا إلى جانب النتائج السابقة جدول المقارنات المتعدد على الشكل التالي:

Multiple Comparisons

Dependent Variable: note

	(I) filière	(J) filière	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
LSD	تسيير واقتصاد	آداب ولغات أجنبية	0,667	0,632	0,294	-0,59	1,92
		علوم تجريبية	-2,000*	0,632	0,002	-3,26	-0,74
	آداب ولغات أجنبية	تسيير واقتصاد	-0,667	0,632	0,294	-1,92	0,59
		علوم تجريبية	-2,667*	0,632	0,000	-3,92	-1,41
	علوم تجريبية	تسيير واقتصاد	2,000*	0,632	0,002	0,74	3,26
		آداب ولغات أجنبية	2,667*	0,632	0,000	1,41	3,92
Bonferroni	تسيير واقتصاد	آداب ولغات أجنبية	0,667	0,632	0,883	-0,88	2,21
		علوم تجريبية	-2,000*	0,632	0,006	-3,54	-0,46
	آداب ولغات أجنبية	تسيير واقتصاد	-0,667	0,632	0,883	-2,21	0,88
		علوم تجريبية	-2,667*	0,632	0,000	-4,21	-1,12
	علوم تجريبية	تسيير واقتصاد	2,000*	0,632	0,006	0,46	3,54
		آداب ولغات أجنبية	2,667*	0,632	0,000	1,12	4,21
Dunnett t (2-sided)[b]	تسيير واقتصاد	علوم تجريبية	-2,000*	0,632	0,004	-3,42	-0,58
	آداب ولغات أجنبية	علوم تجريبية	-2,667*	0,632	0,000	-4,09	-1,25

*. The mean difference is significant at the 0.05 level.

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

من خلال جدول المقارنات المتعددة يتضح أنه لا يوجد فرق بين القرار الخاص باختبار فرضية تساوي المتوسطات بين كل زوج من الشعب باستخدام اختبار LSD أو Bonferroni، بحيث أظهر الاختبار وجود فروق معنوية في العلامات المحصلة في الرياضيات بين طلبة شعبتي التسيير واقتصاد

والعلوم التجريبية، وكذلك بين شعبة الآداب واللغات الاجنبية وشعبة العلوم التجريبية حيث كانة القيمة الاحتمالية .Sig أقل من 0.05 اضافة إلى وجود الإشارة (*) على الفروقات المحسوبة بين متوسطات هذه الأزواج، بينما لا توجد فروقات في المتوسط بين شعبة التسيير والاقتصاد وشعبة الآداب واللغات الأجنبية لأن القيمة الاحتمالية المقدرة أكبر من 0.05. كما يمكن الاعتماد على مجالات الثقة المقدرة للاختبارات بحيث كلما كان الفرق بين العينتين ينتمي إلى مجال الثقة المقدر يتم الاقرار بوجود فرق معنوي بين علامات الطلبة لهذه الشعب.

لاحظ أن الفرق بين متوسطات العلامات بالنسبة لشعبتي تسيير واقتصاد وآداب ولغات أجنبية يساوي 0.667 وهو يقع خارج مجال الثقة المقدر لاختبار LSD [0.59,1.92−]، بينما الفروقات في المتوسط بالنسبة للزوجين المتبقيين يكون داخل مجال الثقة:

الزوج: تسيير واقتصاد−علوم تجريبية= **2−** وهو ينتمي إلى المجال **[3.26,−0.74−]**

الزوج: وآداب ولغات أجنبية−علوم تجريبية=**2.667−** وهو ينتمي إلى المجال **[3.92,−1.41−]**

أما بالنسبة لإختبار Dunnett ، فقد أثبت قبول الفرضية البديلة لنفس الازواج (1−3) و (2−3) لأن القيمة الاحتمالية تقل عن مستوى المعنوية 5% بالنسبة للأزواج (تسيير واقتصاد−علوم تجريبية) و(آداب ولغات أجنبية− علوم تجارية)، بينما لا تظهر مقارنة الزوج (تسيير واقتصاد−آداب ولغات أجنبية)، لأننا اخترنا الفئة الثالثة (علوم تجريبية) كمجموعة سيطرة وهذا افترضناه سابقا من أجل التوضيح فقط.

- **المقارنات المستقلة:**

يسمح الأمر One−way ANOVA بالاضافة الى ما سبق، اجراء مقارنات مستقلة من خلال الزر Contrasts ، ففي كثير من الحالات يهتم الباحث باجراء مقارنة بعض الحالات (الفئات) دون غيرها بناءً على معلومات مسبقة حول عينات الدراسة، بحيث إذا كانت لدينا فئات عددها K فإنه يمكن تكوين K−1 من المقارنات الممستقلة وتعرف باسم Contrast. ويتم من خلالها اختبار الفرضية من الشكل التالي:

$$\begin{cases} H_0 : C_1\overline{X}_1 + C_2\overline{X}_2 + ... + C_k\overline{X}_k = 0 \\ H_1 : C_1\overline{X}_1 + C_2\overline{X}_2 \neq .... + C_k\overline{X}_k \neq 0 \end{cases} \Rightarrow \begin{cases} H_0 : \sum_{i=1}^{k} C_i\overline{X}_i = 0 \\ H_1 : \sum_{i=1}^{k} C_i\overline{X}_i \neq 0 \end{cases}$$

بحيث C عبارة عن ثوابت حقيقية يتم تحديدها من طرف الباحث بناءا على معلومات مسبقة. مع اشتراط أن: $\sum_{i=1}^{K} C_i = 0$.

فإذا افترضنا انه لدينا مقارنتين مستقلتين بحيث تكون لدينا الثوابت $C_1, C_2, ..., C_k$ و $C'_1, C'_2, ..., C'_k$، فإنه وحتى تكون هذه المقارنات مستقلة متعامدة Orthogonal يشترط أن تكون $\sum_{i=1}^{k} C_i C'_i = 0$. وعندئذ يمكن التعبير عن مجموع المربعات الكلي بالصياغة التالية:

$$SS_T = SS_{C_1} + SS_{C_2}$$

والتي يتم استخدامها لاختبار الفرضيات الخاصة بمعنوية كل من هذه المقارنات على حدى، بحيث اذا تم وصف الاختلافات في المتوسط باحدى المقارنات، فإنه يتم حساب احصائية فيشر للمقارنة الأولى بالعلاقة التالية: $F = \frac{SS_{C_1}}{SSE}$ ومن المرجح أن تكون أكبر من $F = \frac{SS_T}{SSE}$.

ولتوضيح طريقة استخدام هذا الأمر باستعمال برنامج SPSS، وبالاستعانة ببيانات المثال السابق (4-10) نفرض أنه نريد اجراء المقارنة التالية:

Secteur	electro ménage	electronique	transport
Contrast 1	−0.5	1	−0.5

أي أن متوسط الأداء لاطارات قطاع انتاج التجهيزت الالكترونية يساوي نصف متوسط الأداء لاطارات قطاع انتاج وسائل النقل+ نصف متوسط الأداء لاطارات قطاع انتاج المنتجات الكهرومنزلية، أي اختبار فرضية أن مستوى الأداء في قطاع الالترونيك يساوي ضعف مستوى الاداء في قطاعي النقل والكهرومنزلي، والتي يمكن التعبير عنها كما يلي:

$$H_0 : \overline{X}_2 = \frac{\overline{X}_1 + \overline{X}_3}{2}$$

$$H_1 : \overline{X}_2 \neq \frac{\overline{X}_1 + \overline{X}_3}{2}$$

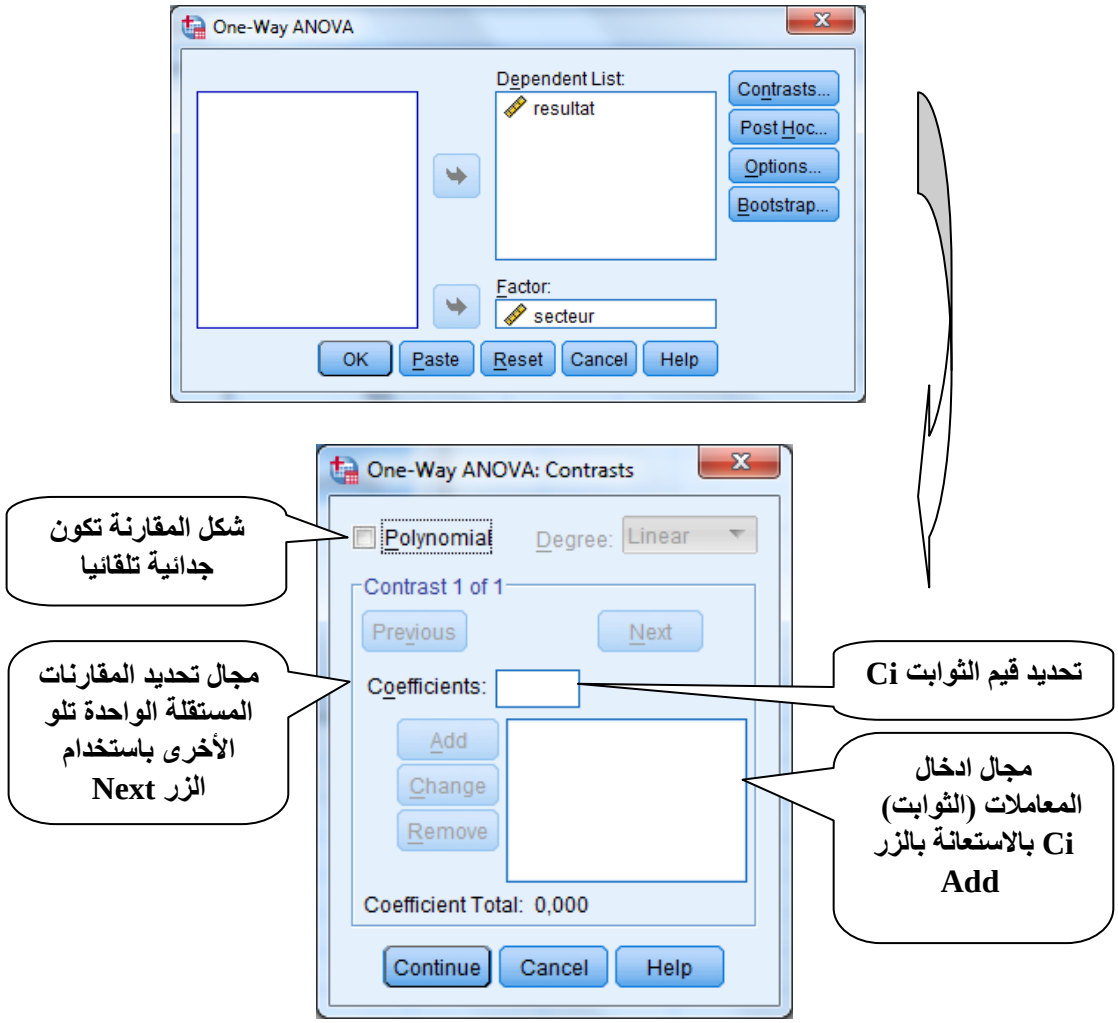
أو

$$H_0 : -0.5\overline{X}_1 + \overline{X}_2 - 0.5\overline{X}_3 = 0$$

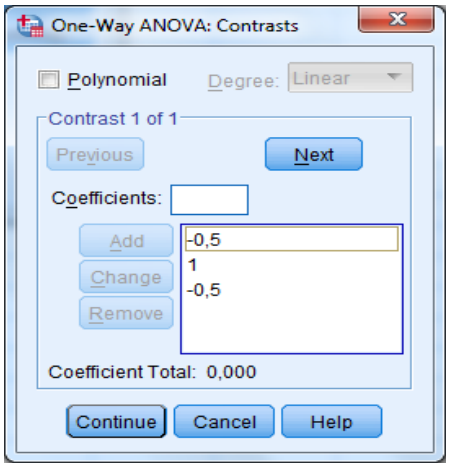
$$H_1 : -0.5\overline{X}_1 + \overline{X}_2 - 0.5\overline{X}_3 \neq 0$$

لاحظ أن الفرضية تحقق الشرط $\sum_{i=1}^{K} C_i = 0$

ولاجراء الاختبار ننقر على الزر Conrasts في صندوق حوار One-way ANOVA فيظهر لدينا صندوق حوار Conrasts التالي:



بحيث نقوم بكتابة قيمة كل معامل من المعاملات الخاصة بـ Contrast 1 الموضحة في الجدول أعلاه، الواحد تلو الآخر في الخانة Coefficients مع النقر على زر الاضافة Add لاضافته، وفي حالة وجود مقارنة أخرى الى جانب المقارنة الأولى نستعين بالزر Next لاضافتها فتصبح لدينا مقارنتين. وفي حالتنا بعد تحديد المقارنة المستقلة Contrast 1 يصبح صندوق الحوار أعلاه على الشكل المقابل:



ولإتمام العملية ننقر على الزر Continue فنحصل على النتائج التالية:

جدول (1) تحليل التباين:

ANOVA

resultat

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2834,657	2	1417,329	2,612	,079
Within Groups	50459,582	93	542,576		
Total	53294,240	95			

جدول (2) المعاملات:

Contrast Coefficients

	secteur		
Contrast	transport	electronique	electroménagé
1	-0,5	1	-0,5

جدول (3) نتائج اختبار الفرضية:

Contrast Tests

		Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
resultat	Assume equal variances	1	1,63	4,972	0,329	93	0,743
	Does not assume equal variances	1	1,63	4,924	0,332	69,938	0,741

يظهر من خلال الجدول (3) نتائج اختبار معنوية المقارنة المفترضة (Contrast 1)، في الحالتين حالة تحقق فرضية تجانس التباين للعينات الثلاث، وحالة عدم تحققها، بحيث تم تقدير قيمة المقارنة وفق الصيغة التالية $\frac{\sum_{i=1}^{3} C_i X_{i\bullet}}{k}$ بحيث C_i هي معاملات المقارنة المذكورة، k عدد العينات (k=3)، $X_{i.}$ هي قيم المتغير التابع بالنسبة لكل قطاع i. وبعبارة أخرى هي $\sum_{i=1}^{3} C_i \overline{X}_i$. قيمة الانحراف المعياري 4.972، وقيمة اختبار ستودنت t المقدرة بـ 0.329 وهي أقل من القيمة المجدولة $(t_{(0.025,\ 93)}=1.98)$، مما يقودنا للإقرار بعدم معنوية هذه المقارنة عند مستوى 5%، كما يمكن الاستعانة بالقيمة الاحتمالية للاختبار .Sig والتي تساوي 0.743 وهي أكبر من 0.05.

تجدر الاشارة الى أن هذه المقارنة المفترضة ليس لها أي غرض، وتم الاستعانة بها من أجل التوضيح فقط، بينما في الواقع يتم استعمال هذا النوع من أجل أغراض علمية وعملية بحتة تساعد في اتخاذ القرار وتحليل الظواهر خاصة في حالة تحليل التباين الثنائي والمتعدد. بالإضافة الى اننا تعرضنا فقط للمقارنات العمودية، وفي الواقع يمكن استعمال مقارنات غير عمودية.

ثانيا : تحليل التباين بمعيارين أو أكثر:

يهتم تحليل التباين بعدة عوامل بمعرفة مدى تأثير عدد معين من العوامل على المتغير التابع (المعتمد)، بحيث يكون لدينا متغير تابع من النوع الكمي يفسر بعدد من المتغيرات الإسمية أو النوعية تأخذ عدد معين من الفئات، ويتم اختبار تأثير كل عامل من هذه العوامل بصفة مستقلة (Main Effect) وكذا تفاعل هذه العوامل فيما بينها (Interaction) باستخدام اختبار فيشر F، وفي هذه الحالة يمكن التعبير عن النموذج العام لتحليل التباين بمعيارين بالعلاقة التالية:

$$y_{ijk} = \mu_{\bullet\bullet} + \alpha_i + \beta_j + \delta_{ij} + \varepsilon_{ijk}$$

بحيث y المتغير التابع، $\mu_{..}$ المتوسط العام، α_i تأثير المتغير المستقل الأول، β_j تأثير المتغير المستقل الثاني، δ_{ij} تأثير تفاعل المتغيرين المستقلين، ε الخطأ العشوائي.

وكمثال على ذلك نستعين ببيانات المثال السابق (4-10)، واضافة عامل جديد وهو عامل الجنس فيصبح لدينا امكانية تحليل التباين بعاملين هما شعبة البكالوريا والجنس. بحيث يتم تحليل تأثير عامل شعبة البكالوريا وعامل الجنس بالإضافة الى تأثير التفاعل بين الشعبة والجنس على المتغير التابع (العلامة المحصلة في مادة الرياضيات). فيصبح النموذج يعبر على أن المتغير التابع له متوسط عام لا يتأثر بالشعبة والجنس وهو $\mu_{..}$ إلا أن هذا المتوسط العام يتأثر بالشعبة التي يدرس فيها الطالب α_i وبنوع الجنس β_j وبتفاعل الشعبة والجنس δ_{ij}. لكن استعمال برنامج SPSS لإجراء هذه العملية لا يكون من خلال الأمر One-way ANOVA لأن هذا الأمر مخصص فقط لتحليل التباين بعامل واحد، لكن باستعمال الأمر General Linear Model، وهو الأمر الخاص بتحليل التباين بعدة عوامل.

مثال (4-11):

لتكن لدينا البيانات الخاصة بالمعدلات المحصلة (Note) لـ 90 طالب اجتازوا امتحان البكالوريا الى جانب المعلومات الخاصة بشعبة البكالوريا (Filière) والجنس (Sexe). المطلوب اجراء تحليل التباين بمعيارين باستخدام البيانات التالية:

no	note	filière	Sexe
24	10,00	آداب ولغات أجنبية	أنثى
25	9,00	آداب ولغات أجنبية	أنثى
26	9,00	آداب ولغات أجنبية	أنثى
27	15,00	آداب ولغات أجنبية	أنثى
28	10,00	آداب ولغات أجنبية	أنثى
29	13,00	آداب ولغات أجنبية	ذكر
30	9,00	آداب ولغات أجنبية	أنثى
31	13,00	علوم تجريبية	أنثى
32	11,00	علوم تجريبية	ذكر
33	11,00	علوم تجريبية	ذكر
34	17,00	علوم تجريبية	ذكر
35	15,00	علوم تجريبية	أنثى
36	12,00	علوم تجريبية	ذكر
37	11,00	علوم تجريبية	ذكر
38	9,00	علوم تجريبية	أنثى
39	11,00	علوم تجريبية	ذكر
40	10,00	علوم تجريبية	أنثى
41	19,00	علوم تجريبية	ذكر
42	15,00	علوم تجريبية	أنثى
43	13,79	علوم تجريبية	ذكر
44	10,63	علوم تجريبية	ذكر
45	11,33	علوم تجريبية	ذكر
46	11,79	علوم تجريبية	ذكر

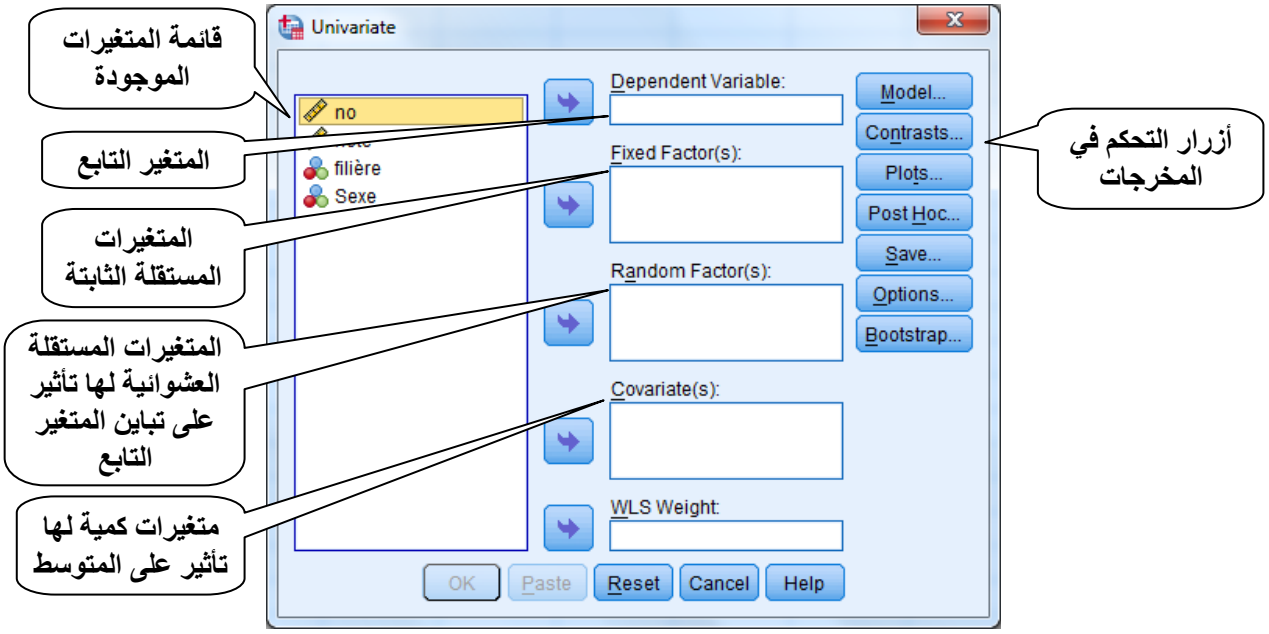
no	note	filière	Sexe
1	13,00	آداب ولغات أجنبية	أنثى
2	10,00	آداب ولغات أجنبية	ذكر
3	12,00	آداب ولغات أجنبية	أنثى
4	9,00	آداب ولغات أجنبية	أنثى
5	8,00	آداب ولغات أجنبية	أنثى
6	10,00	آداب ولغات أجنبية	أنثى
7	4,00	آداب ولغات أجنبية	أنثى
8	12,00	آداب ولغات أجنبية	أنثى
9	10,00	آداب ولغات أجنبية	أنثى
10	9,00	آداب ولغات أجنبية	ذكر
11	7,00	آداب ولغات أجنبية	أنثى
12	12,00	آداب ولغات أجنبية	ذكر
13	7,00	آداب ولغات أجنبية	ذكر
14	9,00	آداب ولغات أجنبية	ذكر
15	9,00	آداب ولغات أجنبية	أنثى
16	11,00	آداب ولغات أجنبية	ذكر
17	11,00	آداب ولغات أجنبية	أنثى
18	11,00	آداب ولغات أجنبية	أنثى
19	8,00	آداب ولغات أجنبية	أنثى
20	15,00	آداب ولغات أجنبية	أنثى
21	9,00	آداب ولغات أجنبية	أنثى
22	9,00	آداب ولغات أجنبية	أنثى
23	10,00	آداب ولغات أجنبية	أنثى

no	note	filière	Sexe
47	10,75	علوم تجريبية	أنثى
48	9,92	علوم تجريبية	ذكر
49	9,88	علوم تجريبية	أنثى
50	11,54	علوم تجريبية	ذكر
51	9,88	علوم تجريبية	أنثى
52	9,54	علوم تجريبية	ذكر
53	9,67	علوم تجريبية	ذكر
54	13,67	علوم تجريبية	ذكر
55	10,42	علوم تجريبية	ذكر
56	10,83	علوم تجريبية	ذكر
57	8,88	علوم تجريبية	أنثى
58	13,71	علوم تجريبية	أنثى
59	12,88	علوم تجريبية	ذكر
60	10,38	علوم تجريبية	ذكر
61	9,46	تسيير واقتصاد	ذكر
62	13,67	تسيير واقتصاد	ذكر
63	9,58	تسيير واقتصاد	أنثى
64	9,29	تسيير واقتصاد	أنثى
65	11,83	تسيير واقتصاد	أنثى
66	10,92	تسيير واقتصاد	ذكر
67	9,83	تسيير واقتصاد	أنثى
68	9,79	تسيير واقتصاد	أنثى
69	14,04	تسيير واقتصاد	ذكر

no	note	filière	Sexe
70	10,63	تسيير واقتصاد	ذكر
71	12,42	تسيير واقتصاد	أنثى
72	10,25	تسيير واقتصاد	ذكر
73	12,50	تسيير واقتصاد	ذكر
74	10,54	تسيير واقتصاد	ذكر
75	11,38	تسيير واقتصاد	ذكر
76	16,08	تسيير واقتصاد	أنثى
77	13,17	تسيير واقتصاد	أنثى
78	11,96	تسيير واقتصاد	ذكر
79	11,83	تسيير واقتصاد	أنثى
80	10,25	تسيير واقتصاد	أنثى
81	9,54	تسيير واقتصاد	أنثى
82	10,00	تسيير واقتصاد	ذكر
83	16,50	تسيير واقتصاد	أنثى
84	15,71	تسيير واقتصاد	أنثى
85	11,00	تسيير واقتصاد	ذكر
86	12,00	تسيير واقتصاد	أنثى
87	12,67	تسيير واقتصاد	ذكر
88	10,17	تسيير واقتصاد	أنثى
89	13,33	تسيير واقتصاد	أنثى
90	12,25	تسيير واقتصاد	ذكر

لإجراء تحليل التباين لمتغير معدل البكالوريا والجنس وشعبة البكالوريا وتحديد تأثير هذه العوامل على معدل البكالوريا باستخدام برنامج SPSS، نتبع الخطوات التالية:

- نختار من قائمة Analyze الأمر General Linear Model ثم Univariate، لأننا بصدد تحليل التباين لمتغير تابع واحد بعاملين (في اتجاهين). فيظهر لدينا صندوق الحوار على النحو التالي:



بحيث يتم تحديد المتغيرات بنقل المتغير Note الى المجال Dependent والذي يجب أن يكون بالضرورة متغير من القياس الكمي، والمتغيرين Filiere و Sexe الى المجال Fixed Facteur(s) وهي متغيرات اسمية (نوعية)، ومن ثم استخدام الأزرار الموجودة في اليمين حسب ما تمت الاشارة اليه في تحليل التباين بمعيار واحد.

وباستكمال اجراءات العملية بالنقر على الزر Ok دون استخدام الأزرار الموجودة في اليمين تحصلنا على النتائج التالية:

Tests of Between-Subjects Effects

Dependent Variable: note

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	65,268[a]	5	13,054	2,550	0,034
Intercept	9547,317	1	9547,317	1865,273	0,000
filière	43,793	2	21,896	4,278	0,017
Sexe	0,053	1	0,053	0,010	0,919
filière * Sexe	2,727	2	1,364	0,266	0,767
Error	429,950	84	5,118		
Total	11741,851	90			
Corrected Total	495,218	89			

a. R Squared = ,132 (Adjusted R Squared = ,080)

ويمكن استعاد تأثير المتوسط العام بالنقر على الزر Model والتأشير على المربع الصغير المقابل للعبارة Include Intercept in Model والتي تكون مؤشر عليها تلقائيا للاخذ في الاعتبار هذا التأثير. وبإعادة التقدير يظهر لنا جدول النتائج دون احتساب أثر المتوسط العام على الشكل التالي:

Dependent Variable: note

Source	Type IV Sum of Squares	df	Mean Square	F	Sig.
Model	11311,901[a]	6	1885,317	368,337	0,000
filière	43,793	2	21,896	4,278	0,017
Sexe	0,053	1	0,053	0,010	0,919
filière * Sexe	2,727	2	1,364	0,266	0,767
Error	429,950	84	5,118		
Total	11741,851	90			

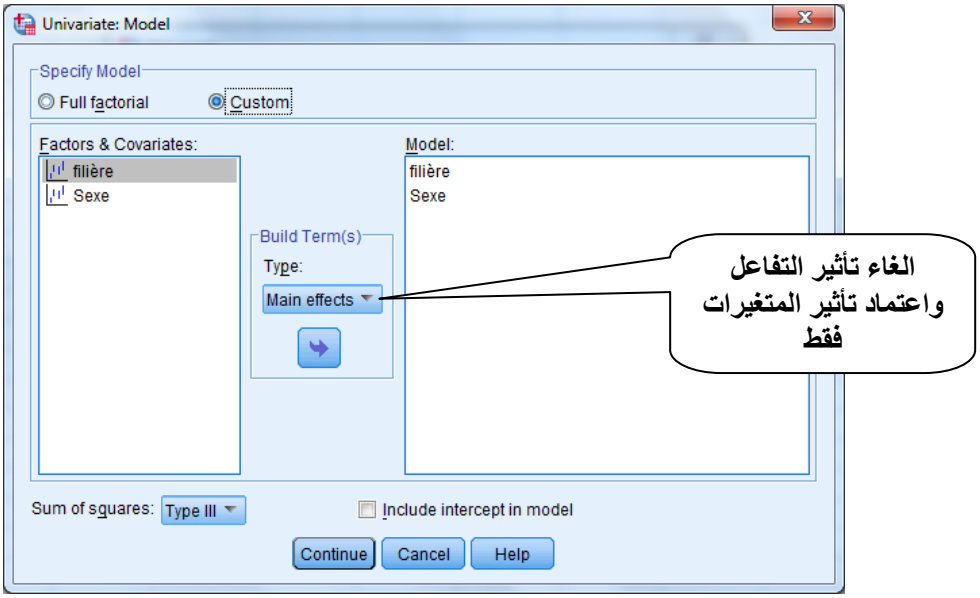
a. R Squared = ,963 (Adjusted R Squared = ,961)

ويتم من خلال هذه النتائج اختبار تأثير العوامل المدرجة في النموذج على المتغير التابع باستعمال احصائية فيشر لإختبار الفرضيات التالية:

$$H_0 : \alpha_i = 0 \quad H_0 : \beta_j = 0 \quad H_0 : \delta_{ij} = 0$$
$$H_1 : \alpha_i \neq 0 \quad \text{و} \quad H_1 : \beta_j \neq 0 \quad ، \quad H_1 : \delta_{ij} \neq 0$$

فإذا كانت القيمة المحسوبة للاختبار أكبر من القيمة المجدولة لفيشر يمكن رفض فرضية العدم والإقرار بتأثير المتغير المعني على نتائج الامتحان أو باستعمال القيمة الاحتمالية التي يجب أن تكون أقل من 0.05 للإقرار بمعنوية التأثير. وفي هذه الحالة وبناءا على نتائج التقدير في الجدول أعلاه يمكن الاقرار بتأثير الشعبة على نتائج امتحان البكالويا في مادة الرياضيات لأن القيمة الاحتمالية لفيشر أقل من 0.05، بينما لا يؤثر متغير الجنس على هذه النتائج وكذلك الأمر بالنسبة لتفاعل الشعبة والجنس نظرا لكون القيمة الاحتمالية لفيشر أكبر من 0.05.

ويمكن تقدير النموذج باستبعاد تأثير التفاعل بين الشعبة والجنس بالنقر على الزر Model في صندوق حوار Univariate وتحديد الخيارات الموضحة في الشكل التالي:



وباتمام الإجراء والنقر على الزر Ok في صندوق الحور الرئيسي نحصل على النتائج التالية:

Dependent Variable: note

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Model	11309,174[a]	4	2827,293	561,960	0,000
filière	55,725	2	27,862	5,538	0,005
Sexe	0,011	1	0,011	0,002	0,963
Error	432,678	86	5,031		
Total	11741,851	90			

a. R Squared = ,963 (Adjusted R Squared = ,961)

وفي الغالب يتم الاستعانة بالتمثيل البياني للمتوسطات الحدية لتوضيح تأثير العوامل على المتغير التابع في حالة وجود تأثير معنوي للعوامل المدرجة والتفاعل بينها من أجل توضيح قوة التأثير الخطي، وذلك بالاستعانة بالزر Plots

4. تحليل الارتباط والانحدار الخطي:

تناولنا سابقا تحليل العلاقة بين متغيرات معبر عنها بفئات ومتغير كمي في تحليل مقارنة المتوسطات وتحليل التباين، وكذا العلاقة بين المتغيرات الإسمية فيما بينها من خلال الجداول التكرارية والتقاطعية، وفي هذه المرحلة نتناول طرق التعامل مع البيانات الكمية من خلال موضوع الارتباط والانحدار.

4-1. الارتباط:

نواجه في الغالب مشكلة تحليل العلاقة بين متغيرين أو أكثر من النوع الكمي غالبا ما يتم اعتبار أحد المتغيرات تابع والأخرى مستقلة، وهذا النوع من التحليل يختلف الى حد ما عن ما تطرقنا اليه فيما سبق. بحيث يكون الهدف منه هو تحديد شكل العلاقة وكذا قوتها، والذي يتم تحديده من خلال معاملات الارتباط بين كل زوج من المتغيرات. ويمكن تعريف معامل الارتباط بأنه المعامل الذي يقيس درجة الارتباط بين المتغيرات ويرمز إليه بالرمز "r_{xy}" يأخذ قيمه في المجال [1-، 1]، بحيث تشير القيمة السالبة الى العلاقة العكسية بين المتغيرين، والإشارة الموجبة تشير الى العلاقة الطردية بينهما، بينما القيم التي تقترب من الواحد الصحيح (+أو-) فتدل على قوة الارتباط بين المتغيرين، أما اقترابها من الصفر فهذا يدل على ضعف الارتباط وليس انعدامها لأنه يمكن أن يكون هنا ارتباط غير خطي بين المتغيرين.

أولا. الارتباط الخطي البسيط:

يتم حساب معامل الارتباط البسيط الذي ينطلق من فرضية وجود علاقة خطية بين متغيرين، وبالتالي يتم تقدير نسبة الارتباط بينهما والتي أشرنا اليها فيما سبق، بأنها تعبر عن اتجاه وقوة العلاقة.
ويبين الجدول التالي مدى قوة معامل الارتباط مقارنة بالقيمة العددية له:

قوة معامل الارتباط	مدى معامل الارتباط
تام	1 أو -1
عال جدا	بين 0,80 و 0,99
عال جدا	بين 0,60 و 0,79
متوسط	بين 0,40 و 0,59
ضعيف	بين 0,20 و 0,39
ضعيف جدا	بين 0,01 و 0,19
لا توجد علاقة خطية	0

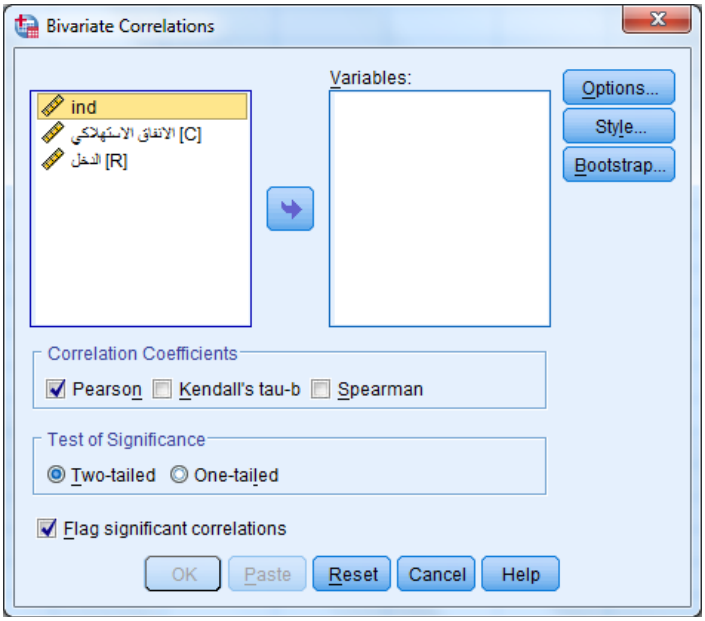
ويتوفر برنامج SPSS على الأوامر اللازمة لحساب هذا المعامل والتي نوضحها من خلال المثال التالي:

مثال (4-12):

لدينا البيانات الخاصة بمجموعة من الأفراد ومقدار الدخل والإنفاق الاستهلاكي الخاص بهم خلال فترة زمنية معينة. وانطلاقا من فرضيات النظرية الاقتصادية التي تقر بأنه توجد علاقة طردية بين الدخل والإنفاق الاستهلاكي نريد اختبار ما إذا كانت هذه الفرضية محققة وأنه توجد فعلا علاقة قوية بين هذين المتغيرين باستعمال بيانات من الواقع. والجدول التالي يبين هذه البيانات

الفرد	الانفاق الاستهلاكي (C)	الدخل (R)
1	17250	23000
2	16100	28750
3	21390	32200
4	25760	34500
5	25300	36800
6	29900	40250
7	34500	46000
8	39100	50600
9	44850	57500
10	52900	69000
11	48300	69000

بعد تعبئة هذه البيانات في نافذة Data Editor، نختار الأمر Correlate ثم Bivariate فيظهر لدينا صندوق حوار Bivariate Correlation التالي:



بحيث نقوم بنقل المتغيرين المعنيين بالارتباط وهما الانفاق الاستهلاكي والدخل الى المجال Variables.

كما يتوفر هذا الصندوق على الخيارات التالية:

1. المجال Correlation Coefficients:

- Pearson: لاستخراج معامل الارتباط البسيط لبيرسون، وهو مقياس للارتباط بين متغيرين كميين، وهو الخيار التلقائي للبرنامج.

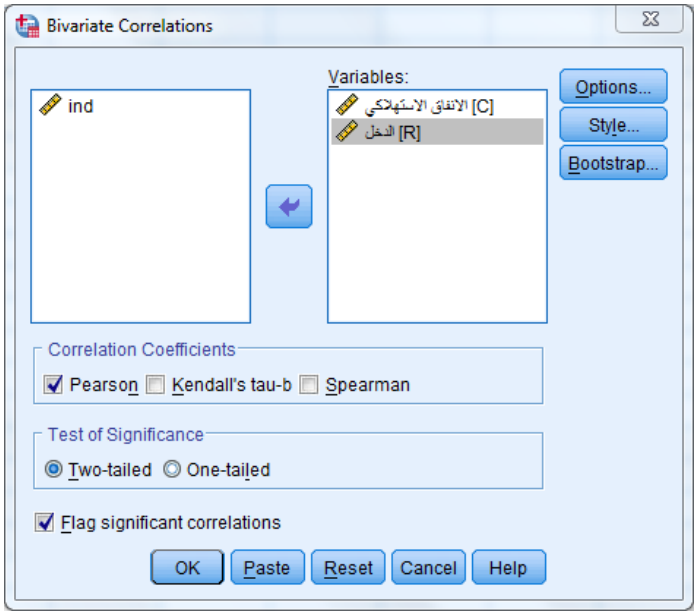
- Kendall's Tau: لاستخراج معامل الارتباط بالطرق غير المعلمية باستعمال الرتب في حالة الرغبة في تقدير معامل الارتباط بطريقة تقريبية، وهو المقياس الذي يستعمل في حالة المتغيرات الترتيبية (ليست كمية) وعندئذ يتم ادخال الرتب المتناظرة المتغيرين بدلا من القيم الأصلية.

Spearman: لاستخراج معامل الارتباط للمتغيرات الترتيبة كما هو الحال لمعامل Kendall.

2- المجال Test Significiance: لتحديد اختبار الفرضية من طرف واحد One Tailed أو من طرفين Two Tailed.

3. Flag Significiance Correlation: لتعليم الارتباطات المعنوية بوضع اشارة (*) أمامها دليل على معنويتها احصائيا.

وبالتأشير على الخيارات المختارة كما هو موضح في الصورة التالية:



وبالنقر على الزر OK ينتج لدينا في المخرجات Output النتائج التالية:

Correlations

		الاستهلاكي الانفاق	الدخل
الاستهلاكي الانفاق	Pearson Correlation	1	0,986**
	Sig. (2-tailed)		0,000
	N	11	11
الدخل	Pearson Correlation	0,986**	1
	Sig. (2-tailed)	0,000	
	N	11	11

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

هذه النتائج هي عبارة عن مصفوفة الارتباط بين المتغيرات المُدرجة، وفي حالة تعدد المتغيرات تكون رتبة المصفوفة تساوي عدد المتغيرات، بحيث يتم حساب معامل الارتباط البسيط وفق العلاقة التالية:

$$r_{XY} = r_{YX} = \frac{N\sum X_i Y_i - (\sum X_i)(\sum Y_i)}{\sqrt{\left[N\sum X_i^2 - (\sum X_i)^2\right]\left[N\sum Y_i^2 - (\sum Y_i)^2\right]}} = 0.986$$

وهي تشير الى وجود ارتباط قوي بين المتغيرين، بينما تشير القيم القطرية الى وجود ارتباط ذاتي تام للمتغيرين ($r_{yy}=r_{xx}=1$). بالإضافة الى القيمة الاحتمالية لاختبار ستودنت لمعنوية المعاملات المقدرة مع التأشير على المعاملات بالرمز (**) دليل على معنوية الاختبار عند مستوى 1% والتي تمت الاشارة اليها في آخر الجدول، بحيث يتم اختبار الفرضية التالية:

$$H_0 : r = 0$$
$$H_1 : r \neq 0$$

وقيمة الاختبار المحسوبة تكون عبارة عن t ستيودنت بدرجة حرية (N−k)، بحيث N هي عدد المشاهدات و k هي عدد المتغيرات و t المحسوبة وفق العلاقة التالية:

$$t^* = r.\frac{\sqrt{N-k}}{\sqrt{1-r^2}} = 0.986 \times \frac{\sqrt{11-2}}{\sqrt{1-(0.986)^2}} = 17.74$$

وهي أكبر من القيمة المجدولة (t=2.26)، مما يقودنا الى الاقرار بالفرضية البديلة أي أن معامل الارتباط بين الانفاق الاستهلاكي والدخل المقدر ب r=0.986 يختلف معنويا عن الصفر.

كما يمكن استخراج قيمة معامل ارتباط الرتب لسبيرمان بالتأشير على الخيار Spearman في المجال Correlation Coefficients، وبتطبيق الأمر ينتج لدينا النتائج المبينة في الجدول التالي:

Nonparametric Correlations

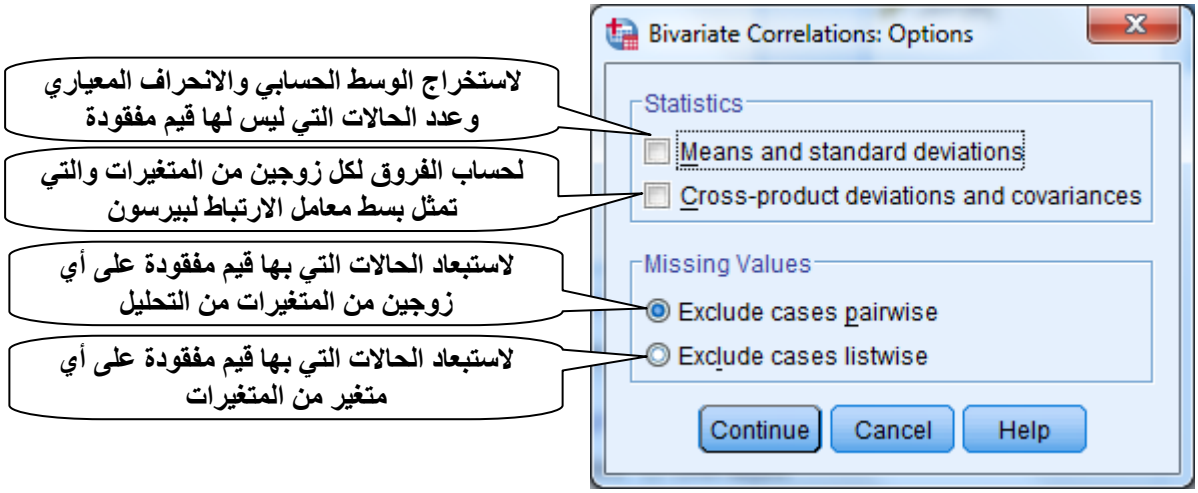
Correlations

			الاستهلاكي الانفاق	الدخل
Spearman's rho	الاستهلاكي الانفاق	Correlation Coefficient	1,000	0,980**
		Sig. (2-tailed)	.	0,000
		N	11	11
	الدخل	Correlation Coefficient	0,980**	1,000
		Sig. (2-tailed)	0,000	.
		N	11	11

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

بحيث تشير النتائج الى أن قيمة معامل الارتباط الرتبي بين الانفاق الاستهلاكي والدخل تساوي الى 0.98 وهي تختلف معنويا عن الصفر عند مستوى معنوية 1% بحيث أن القيمة الاحتمالية لاختبار ستودنت أقل من 0.01.

كما يمكن الاستعانة بالزر Option لعرض مؤشرات احصائية وصفية وطريقة التعامل مع القيم المفقودة إن وجدت من خلال التأشير على أحد الخيارات المتاحة، الذي يظهر من خلاله صندوق الحوار التالي:



وبالعودة الى الأمر Bivariate–Correlate والنقر على الزر Option، والتأشير على الخيارين في المجال Statistics ينتج لدينا النتائج التالية:

Descriptive Statistics

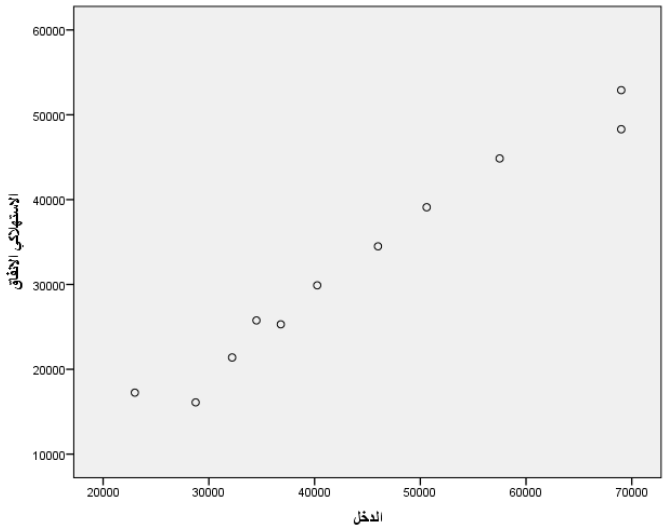
	Mean	Std. Deviation	N
الاستهلاكي الانفاق	32304,55	12627,131	11
الدخل	44327,27	15644,765	11

Correlations

		الاستهلاكي الانفاق	الدخل
الاستهلاكي الانفاق	Pearson Correlation	1	0,986**
	Sig. (2-tailed)		0,000
	Sum of Squares and Cross-products	1594444472,727	1947056636,364
	Covariance	159444447,273	194705663,636
	N	11	11
الدخل	Pearson Correlation	0,986**	1
	Sig. (2-tailed)	0,000	
	Sum of Squares and Cross-products	1947056636,364	2447586818,182
	Covariance	194705663,636	244758681,818
	N	11	11

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

ملاحظة: غالبا ما يتم الاستعانة بالتمثيل البياني لشكل الانتشار لتوضيح الارتباط بين المتغيرات والذي يتم استخراجه من خلال القائمة Graphs واختيار الخيار Legacy Dialogs ثم الخيار Scatter/Dot.



ثم اختيار المتغير التابع في محور التراتيب والمتغير المستقل في محور الفواصل.

وفي حالتنا ينتج لدينا الشكل المقابل:

بحيث يظهر جليا اقتراب شكل نقاط الانتشار من خط مستقيم مما يدل على وجود ارتباط خطي موجب بين المتغيرين الانفاق الاستهلاكي والدخل.

ثانيا. الارتباط الجزئي:

يستخدم هذا النوع من معاملات الارتباط في حالة قياس الارتباط بين متغيرين مع افتراض ثبات متغير ثالث (أو باقي المتغيرات المستقلة)، بحيث يقيس لنا درجة تأثير متغير مستقل على متغير تابع مع ثبات باقي المتغيرات المستقلة، على خلال معامل الارتباط البسيط الذي يقيس هذا التأثير مع اهمال باقي المتغيرات[8]. ولتوضيح طريقة حساب هذا المعامل باستخدام برنامج SPSS نستعين بالمثال التالي:

مثال (4-13):

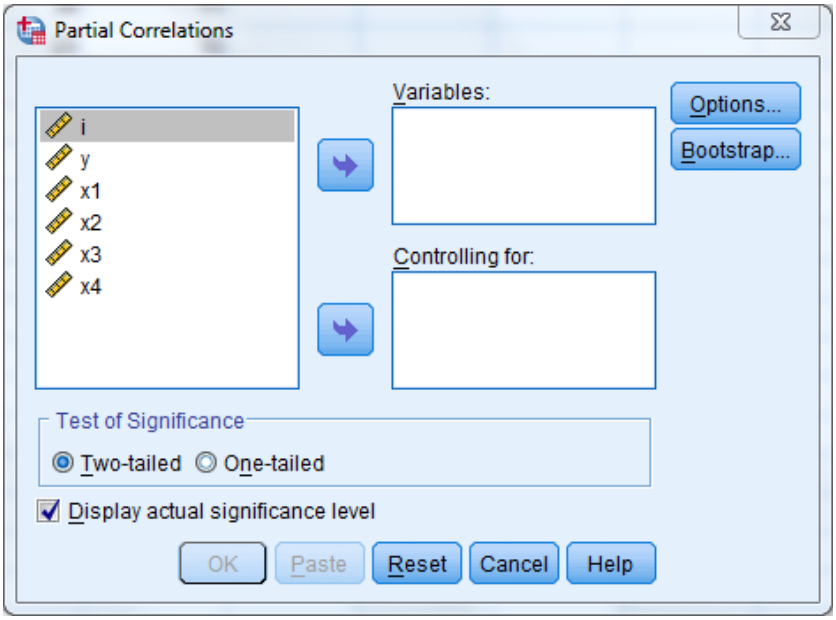
لدينا البيانات الخاصة بمجموعة من المتغيرات على النحو التالي:

i	Y	X1	X2	X3	X4
1	43	5	3	18	12
2	63	9	5	27	9
3	71	10	7	34	11
4	61	8	4	24	10
5	81	11	6	33	6
6	44	12	5	22	8
7	58	9	4	28	9
8	71	7	7	32	7
9	72	8	5	23	8
10	67	13	8	20	5
11	64	4	5	21	4
12	69	10	9	36	10
13	68	11	10	30	11

والمطلوب حساب معامل الارتباط الجزئي بين Y و X2 مع افتراض ثبات X3.

[8] للمزيد من التفاصيل حول هذا المعامل، أنظر في ذلك: محمد شامد بهاء الدين فهمي: "الاحصاء بلا معاناة المفاهيم مع التطبيقات باستخدام برنامج SPSS" الجزء الثاني، دار البحوث، الرياض، 1426هـ، ص 554.

للإجابة على هذا السؤال، نستعين بالأمر Correlation من قائمة Analyze ثم نختار الخيار Partial فيظهر صندوق حوار على الشكل التالي:



بحيث نقوم بنقل المتغيرات الذي نريد حساب معامل الارتباط بينها الى المجال Variables، والمتغيرات المراد استبعاد أثرها الى المجال Controlling for، بحيث ننقل المتغيرات Y، X2 الى المجال Variables، والمتغير X3 الى المجال Controlling for، وفي هذه الحالة المتغيرات الأخرى لا تؤخذ بعين الاعتبار. و بالنقر على الزر Ok يظهر لنا النتائج التالية:

جدول الارتباط الجزئي بين المتغيرين Y و X2.

Partial Corr

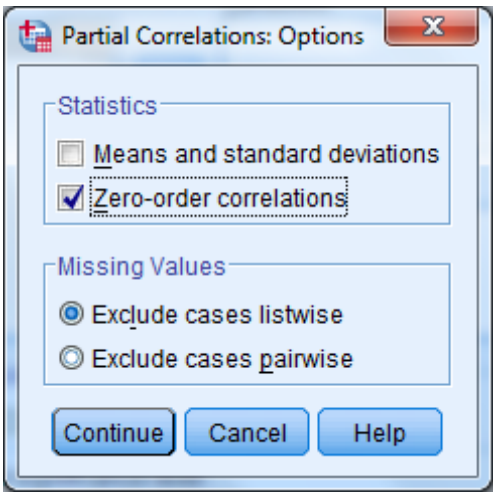
Correlations

Control Variables			y	x2
x3	y	Correlation	1,000	0,285
		Significance (2-tailed)	.	0,369
		df	0	10
	x2	Correlation	0,285	1,000
		Significance (2-tailed)	0,369	.
		df	10	0

معامل الارتباط الجزئي بين المتغيرين Y و X2

القيمة الاحتمالية لاختبار t لمعنوية المعامل

كما يمكننا النقر على الزر Option من اجل حساب معامل الارتباط البسيط بالتأشير على الخيار Zero Order Correlation كما تظهره الصورة التالية:



وبإتمام العملية بالنقر على الزر OK في صندوق حوار Partial Correlation يتنتج لدنا النتائج التالية:

Correlations

Control Variables			y	x2	x3
-none-[a]	y	Correlation	1,000	0,540	0,634
		Significance (2-tailed)	.	0,057	0,020
		df	0	11	11
	x2	Correlation	0,540	1,000	0,564
		Significance (2-tailed)	0,057	.	0,045
		df	11	0	11
	x3	Correlation	0,634	0,564	1,000
		Significance (2-tailed)	0,020	0,045	.
		df	11	11	0
x3	y	Correlation	1,000	0,285	
		Significance (2-tailed)	.	0,369	
		df	0	10	
	x2	Correlation	0,285	1,000	
		Significance (2-tailed)	0,369	.	
		df	10	0	

معاملات الارتباط البسيط بين المتغيرات Y و X2 و X3 ونتائج اختبار معنويتها

معاملات الارتباط الجوئي بين المتغيرات Y و X2 ونتائج اختبار معنويتها

a. Cells contain zero-order (Pearson) correlations.

4-2. تحليل الانحدار:

الانحدار الخطي عبارة عن اسلوب احصائي يستخدم لتحليل العلاقات بين المتغيرات الكمية، بحيث يكون أحدها عبارة عن متغير تابع للمتغيرات الأخرى والتي تعتبر مستقلة (مُفَسِّرة)، والغرض منه هو قياس مدى تأثير المتغيرات المستقلة على المتغير التابع بتقدير هذا الأخير واختباره احصائيا.

4-1. تحليل الانحدار الخطي البسيط:

يدرس نموذج الانحدار الخطي البسيط العلاقة بين متغيرين فقط أحدهما تابع الآخر مستقل وفق العلاقة التالية:

$$Y_i = \alpha + \beta X_i + \varepsilon_i \qquad i = 1,2,...,N$$

بحيث تمثل Y المتغير التابع وX المتغير المستقل، α و β معلمات للتقدير و ε الأخطاء العشوائية والتي تعرف ببواقي التقدير وهي تساوي الى القيم الحقيقية للمتغير التابع مطروحا منها القيم المقدرة ويعبر عنها رياضيا بـ: $\hat{\varepsilon}_i = Y_i - \hat{Y}_i$.

أولا: فرضيات النموذج الخطي البسيط:

يعتمد نموذج الانحدار الخطي البسيط على مجموعة من الفرضيات الأساسية حول المتغيرات المستعملة وكذلك الأخطاء العشوائية المقدرة يمكن اختصارها فيما يلي:

1- ان تكون العلاقة بين المتغير التابع والمتغير المستقل خطية.

2- ان يكون توقع الخطأ العشوائي في نموذج الانحدار معدوم $E(\varepsilon_i) = 0$.

3- أن يكون تباين الخطأ العشوائي ثابت (فرضية تجانس تباين الأخطاء العشوائية أي:

$$Var(\varepsilon_i) = E(\varepsilon_i^2) = \sigma^2, \forall i < \infty .$$

4- لا يوجد ارتباط بين الأخطاء العشوائية فيما بينها أي: $Cov(\varepsilon_i, \varepsilon_j) = E(\varepsilon_i \varepsilon_j) = 0, \forall i \neq j$.

5- لا توجد علاقة بين المتغير المستقل والخطأ العشوائي.

كما يشترط أن تكون العلاقة بين المتغير التابع والمتغير المستقل في اتجاه واحد، أي أن المتغير المستقل يؤثر في المتغير التابع والعكس غير صحيح.

ومن خلال الفرضيتين (2 و 3)، يتجلى شرط التوزيع الطبيعي بمتوسط معدوم وتباين σ^2 الذي يعتبر ضروريا لاختبار معنوية المعلمات المقدرة للنموذج (معاملات الانحدار).

ثانيا. طريقة تقدير نموذج الانحدار:

تتعدد طرق تقدير نموذج الانحدار الخطي(بتقدير المعلمات α و β)، لكن من أشهر هذه الطرق والأكثر استعمالا هي طريقة المربعات الصغرى العادية نظرا لخصائص مقدراتها التي تتميز بخاصية عدم التحيز وخاصية أقل تباين، بالإضافة إلى سهولة تطبيقها، بحيث تسمح بإيجاد أحسن تصحيح خطي بتدنئة مجموع مربعات الانحرافات بين القيم الحقيقية والمقدرة (بواقي التقدير). ويتم حساب هذه المقدرات بالعلاقات الرياضية التالية:

$$\hat{\beta} = \frac{\sum_{i=1}^{N}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{N}(Y_i - \overline{Y})^2}$$

$$\hat{\alpha} = \overline{Y} - \hat{\beta}.\overline{X} \quad \text{و}$$

ثالثا. اختبار الفرضيات:

بعد تقدير معلمات النموذج، يتم اختبارها احصائيا باستعمال الاختبار الاحصائية من أهمها:

1- اختبار المعنوية الاحصائية للمعلمات المقدرة:

يتم استعمال اختبار t ستودنت لاختبار الفرضية التالية:

$$H_0 : \beta = 0 \quad H_0 : \alpha = 0$$
$$H_1 : \beta \neq 0 \quad \text{و} \quad H_1 : \alpha \neq 0$$

بحيث يتم اختبار فرضية انعدام المعلمات المقدرة كل على حدى، تكون صياغة الاختبار على الشكل التالي:

$$t = \frac{\hat{\beta}}{\sigma_{\hat{\beta}}} \quad \text{و} \quad t = \frac{\hat{\alpha}}{\sigma_{\hat{\alpha}}}$$

وتقارن القيمة المحسوبة للاختبار بالقيمة المجدولة بدرجة حرية $N\text{-}k$، من أجل قبول أو رفض الفرضية الصفرية H_0 عند مستوى معنوية معين، عادة ما يستعمل مستوى 5%. بحيث نقبل الفرضية H_0 إذا كانت القيمة المحسوبة أكبر من القيمة المجدولة، أو بالاستعانة بالقيمة الاحتمالية للاختبار كما تمت الاشارة اليه سابقا.

2- تحليل التباين وإختبار القدرة التفسيرية للنموذج:

تعتمد هذه الطريقة على قيم بواقي التقدير التي تساعد في قياس مدى تمثيل معادلة الانحدار لمشاهدات العينة، بحيث ان القيمة الكبيرة للبواقي تدل على التمثيل غير الجيد للعلاقة بين المتغير المفسر

والمتغير التابع. وبالاستعانة بجدول تحليل التباين يمكن حساب مؤشر جودة التوفيق (معامل التحديد) R^2، وكذلك حساب قيمة اختبار فيشر.

و يتم استعمال اختبار فيشر F بدرجة حرية k-1 و N-k، بحيث k هي عدد المعلمات في النموذج (2)، لاختبار فرضية انعدام معلمات النموذج المقدرة من عدمها(*). وتكون صيغة الاختبار على الشكل التالي:

$$F = \frac{\sum_{i=1}^{N}(Y-\bar{Y}) \,/\, (k-1)}{\sum_{i=1}^{N}\varepsilon_i \,/\, (N-k)} \neg F_{(k-1,N-k)}$$

بحيث يتم قبول فرضية العدم التي تقر بانعدام معلمات النموذج إذا كانت القيمة المحسوبة أقل من القيمة المجدولة.

أما معامل التحديد والذي يشير الى نسبة الانحرافات المفسر بالنموذج المقدر، ويتم حسابه بالعلاقة:

$$R^2 = \frac{\sum_{i=1}^{N}(\hat{Y}-\bar{Y})}{\sum_{i=1}^{N}(Y-\bar{Y})} = 1 - \frac{\sum_{i=1}^{N}\varepsilon_i^2}{\sum_{i=1}^{N}(Y-\bar{Y})}$$

ويأخذ معامل التحديد قيمه في المجال [0،1]، فكلما كانت هذه القيمة قريبة من الواحد فهذا يعني ان نسبة كبيرة من المتغير التابع مفسرة بالمتغيرات المستقلة، والعكس صحيح.

3- اختبار الفرضيات الخاصة بالبواقي ε_i :

في هذه الحالة نركز على فرضية التوزيع الطبيعي لبواقي التقدير من خلال التمثيلات البيانية الخاصة بالبواقي، بالإضافة الى اختبار الارتباط الذاتي للأخطاء العشوائية من خلال احصائية داربين واتسون DW.

ولتوضح طريقة تقدير نموذج الانحدار الخطي البسيط باستعمال برنامج SPSS نستعين بالبيانات الخاصة بالدخل والإنفاق الاستهلاكي لعينة متكونة من 11 فرد، من أجل قياس أثر الدخل على الاستهلاك والتحقق من فرضيات النظرية الاقتصادية التي تقر بوجود علاقة طردية بين المتغيرين، مع تحديد درجة تأثير التغير في الدخل على التغير في الاستهلاك بتقدير معلمات النموذج التالي:

$$C_i = c_0 + c_1 . R_{1i} + \varepsilon_i$$

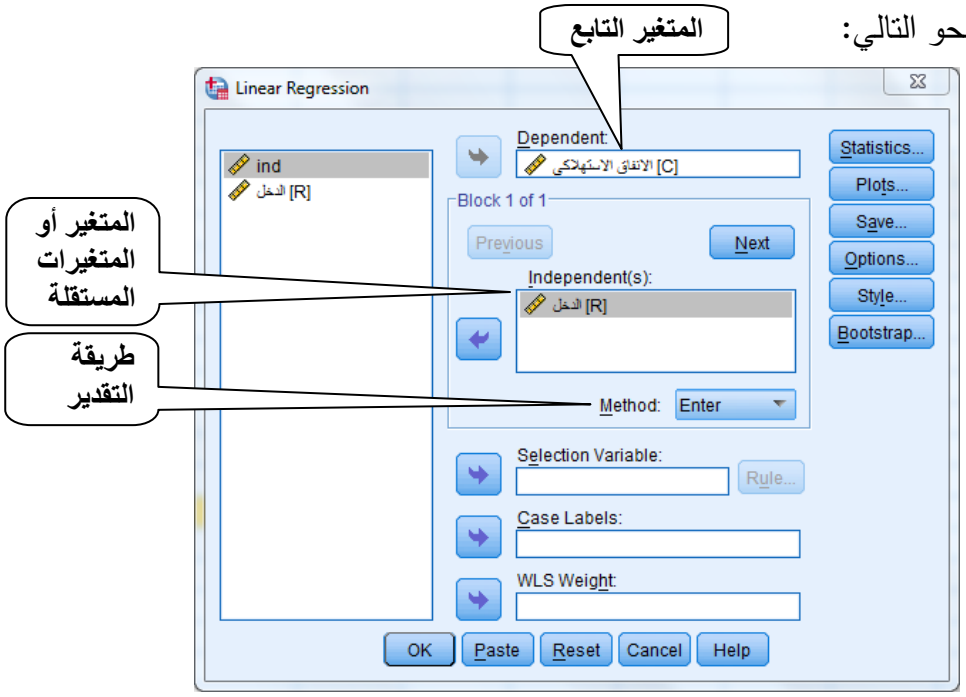
(*) نشير هنا الى أن اختبار فيشر يختبر فرضية انعدام معلمات النموذج من دون الحد الثابت، أي معلمات الانحدار فقط.

بحيث تمثل c_0 قيمة الاستهلاك المستقل عن الدخل (حد الكفاف)، و c_1 نسبة الدخل المخصص للاستهلاك.

ويبين الجدول التالي البيانات الخاصة بالاستهلاك والدخل: (الوحدة: ألف دينار)

الفرد	1	2	3	4	5	6	7	8	9	10	11
الاستهلاك	75	70	93	112	110	130	150	170	195	230	210
الدخل	100	125	140	150	160	175	200	220	250	300	300

وللقيام بهذا الإجراء نختار من قائمة Analyze الأمر Regression ثم Lineair فيظهر لنا صندوق حوار على النحو التالي:

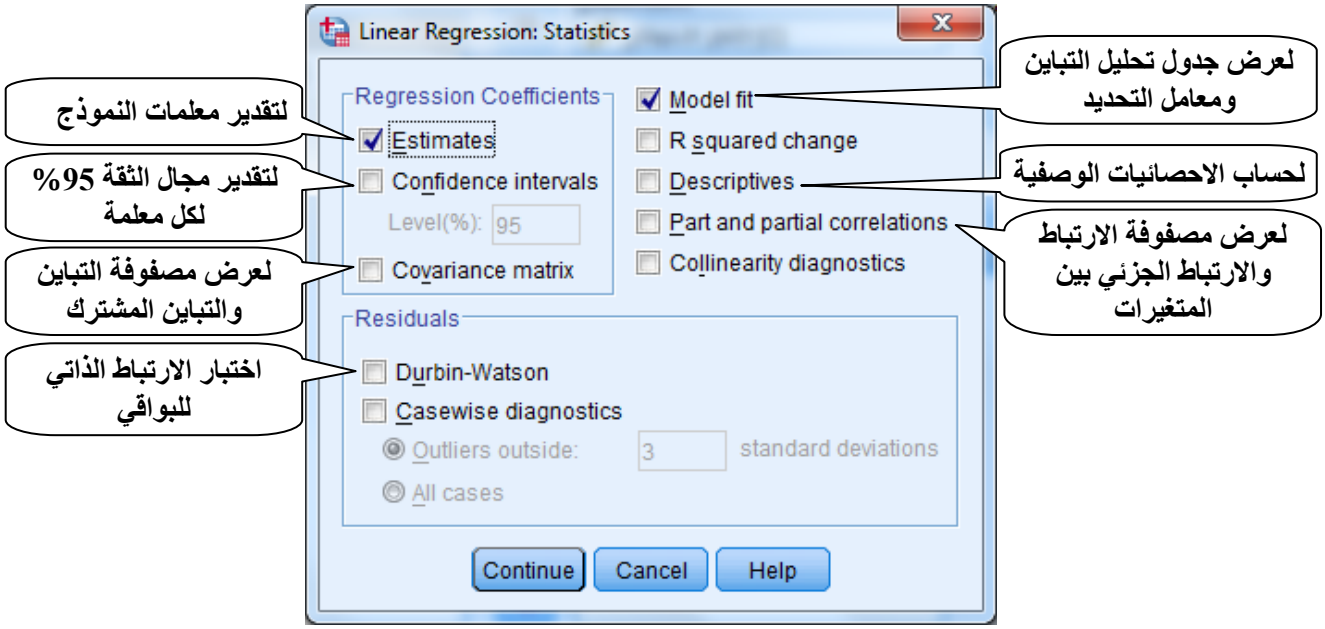


بحيث يتم نقل المتغير التابع الى المجال Dependent والمتغير المستقل الى المجال Independent، ويتوفر الزر Method على عدة خيارات خاصة بطريقة التقدير، والطريقة التلقائية للبرنامج هي طريقة المربعات الصغرى العادية المشار اليها بـ Enter. أما المجال الخاص بالمتغيرات المستقلة فيتوفر فيه امكانية تقدير مجموعة من النماذج باستخدام نفس المتغير التابع ومتغيرات مختلفة مستقلة مثلا استخدام المتغير X1 في النموذج (1) و X2 في النموذج (2) بنفس الإجراء، وذلك بالاستعانة بالزر Next في المجال Block 1 Of 1، بحيث يتم نقل المتغير X1 في Block1 والمتغير X2 في Block2، ونتائج التقدير تكون في شكل نموذجين مختلفين. بينما ادخال المتغيرين معا في المجال Independent فإنه يؤدي الى تقدير نموذج انحدار متعدد (بمتغيرين مستقلين).

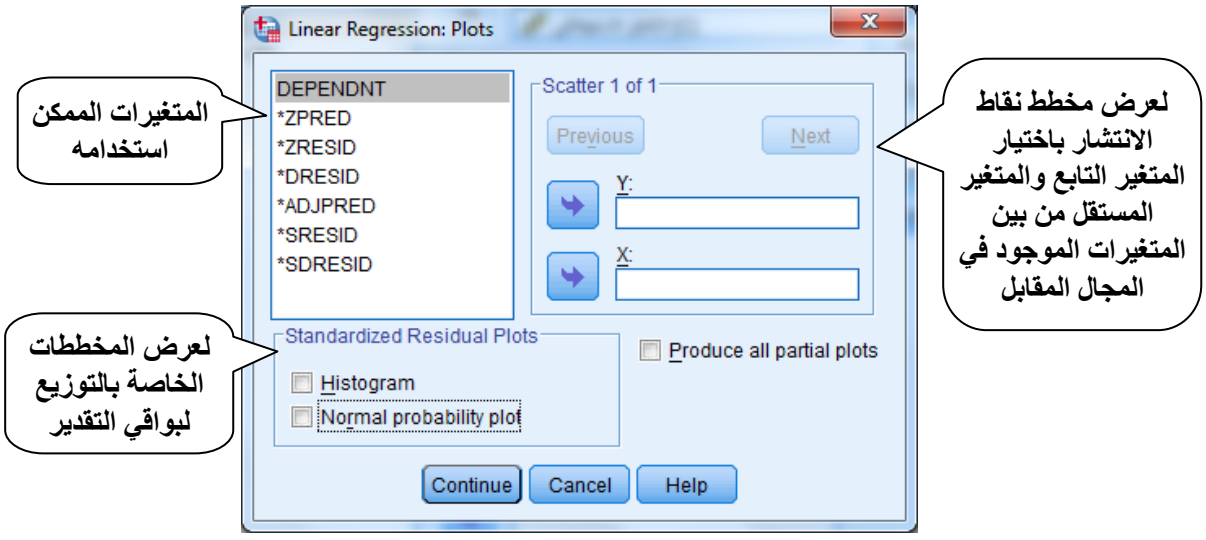
ويفيدنا المجال Selection Variable في تحديد متغير يحدد لنا مجموعة معينة من البيانات متعلقة بالتحليل دون غيرها، فمثلا استخدام متغير "الفرد" بتحديد رقم الأفراد المعنيين أكبر من 4 مثلا، وهذا بالاستعانة بالزر Rule.

كما يتوفر هذا الصندوق على مجموعة من الأزرار نختصرها فيما يلي:

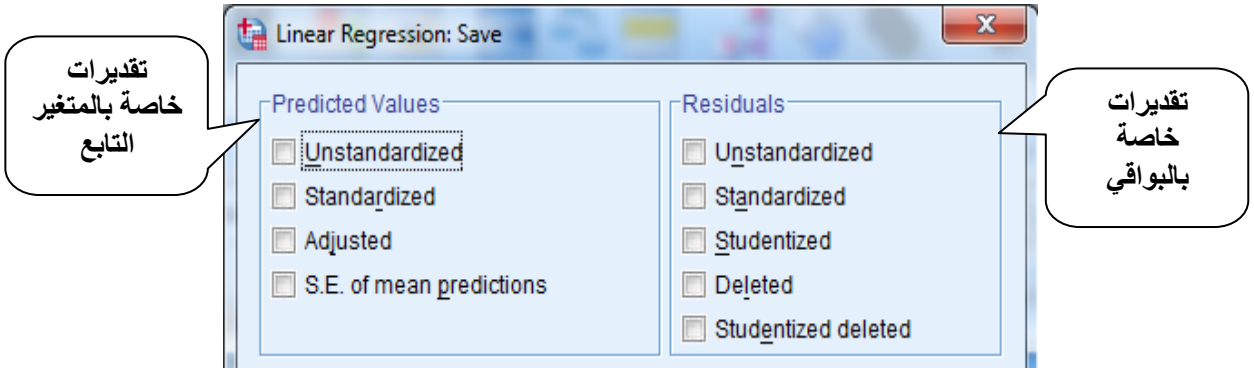
- الزر Statistics: بالنقر عملى هذا الزر يظهر لنا صندوق الحوار التالي:



وبالنقر على الزر Continue، والعودة الى صندوق الحوار السابق، ننقر على الزر Plots وهو خاص بعرض بعض المخططات البيانية باستخدام القيم المقدرة للمتغير التابع باستخدام معادلة الانحدار المقدرة وكذلك قيم البواقي وعدة قيم أخرى يتم تحديدها من خلال صندوق الحوار التالي:



- الزر Save: يسمح هذا الزر بحساب مجموعة من القيم والمؤشرات، ونذكر على وجه الخصوص القيم المقدرة للمتغير التابع والقيم المقدرة للبواقي. والتي تظهر في صندوق حوار Save التالي:



بحيث نقوم بالتأشير على unstandarized في المجال Predicted Values و Standarized في المجال Residuals، لحساب القيم المقدرة للمتغير التابع بالعلاقة المقدرة، والأخطاء العشوائية (بواقي التقدير) التي يتم اضافتها الى البيانات الموجودة في ورقة Data View. والهدف من ذلك هو اعداد مخطط بياني لجودة التوفيق.

الزر Option: يفيدنا في اضافة الحد الثابت (المعامل c_0) الى النموذج بالتأشير على الخيار Include Constant in equation، ونشير هنا إلى أننا قمنا بتحديد معامل الانحدار (c_1) من خلال اضافتنا المتغير R الى مجال المتغيرات المستقلة سابقا.

وباتمام كل الإجراءات المذكورة نحصل على نتائج تقدير النموذج ومختلف متطلباته المحددة على الشكل التالي:

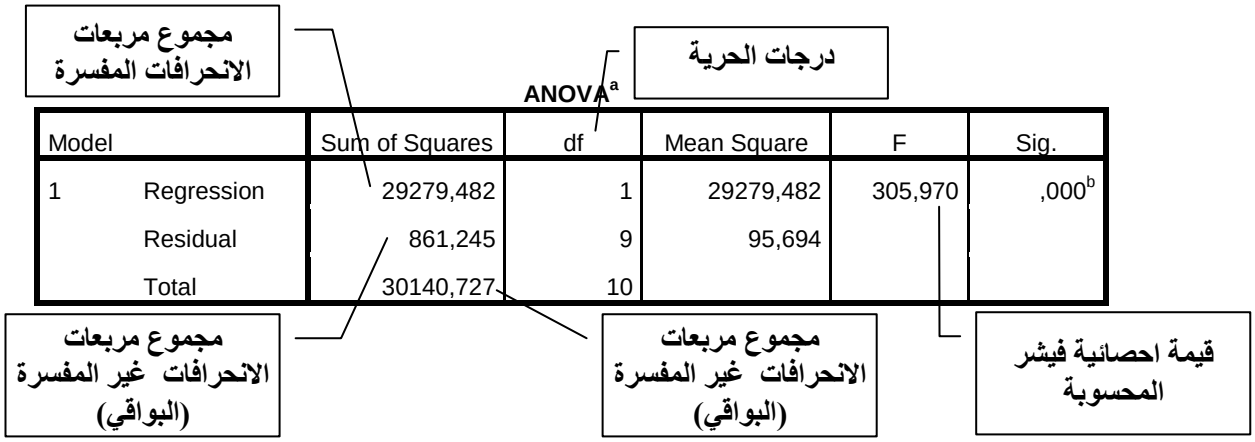
الجدول رقم (1): معامل التحديد وإحصائية داربين واتسون:

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	0,986[a]	0,971	0,968	9,782	1,706

يبين هذا الجدول أن القيمة المحسوبة لمعامل التحديد R Square=0.971، وهي تدل على أن 97.1% من الاستهلاك مفسرة بمتغير الدخل وتبقى نسبة 2.9% منه غير مفسرة ويعبر عنها ببواقي التقدير. بينما احصائية داربين واتسون فتساوي 1.706 والتي يتم مقارنتها بالقيمة المجدولة لداربين واتسون النظرية (المجدولة) حسب القيمة العليا d_U والقيمة الدنيا d_L، لمعرفة وجود ارتباط ذاتي للأخطاء العشوائية (البواقي)، في هذه الحالة القيمة المحسوبة أكبر من القيمة المجدولة d_U=1.32، وبالتالي يتم الاقرار بعدم وجود ارتباط ذاتي للأخطاء العشوائية. ونظرا للانتقادات الموجهة لمعامل التحديد يمكن الاستعانة بمعامل التحديد المصحح Adjusted R Square.

الجدول رقم (2): جدول تحليل التباين:



ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	29279,482	1	29279,482	305,970	,000[b]
	Residual	861,245	9	95,694		
	Total	30140,727	10			

ومن خلال هذه النتائج يتضح أن احصائية فيشر المحسوبة تقدر بـ 305.97 وهي أكبر من القيمة المجدولة ($F_{(1,9)}$=5.12)، مما يقودنا الى قبول الفرضية البديلة (H_0) التي تقر بالاختلاف المعنوي للمعلمات المقدرة عن الصفر. كما يتضح كذلك أن مجموع مربعات البواقي صغير نوعا ما مما يدل على أن نسبة كبيرة من الانحرافات في المتغير التابع (الاستهلاك) مفسرة بالتغير في المتغير المستقل (الدخل).

الجدول رقم (3): نتائج التقدير:

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	-12,860	9,248		-1,391	,198			
	الدخل	,796	,045	,986	17,492	,000	,986	,986	,986

a. Dependent Variable: الاستهلاكي الانفاق

من خلال هذا الجدول يمكن كتابة النموذج المقدر للعلاقة بين الاستهلاك والدخل على الشكل التالي:

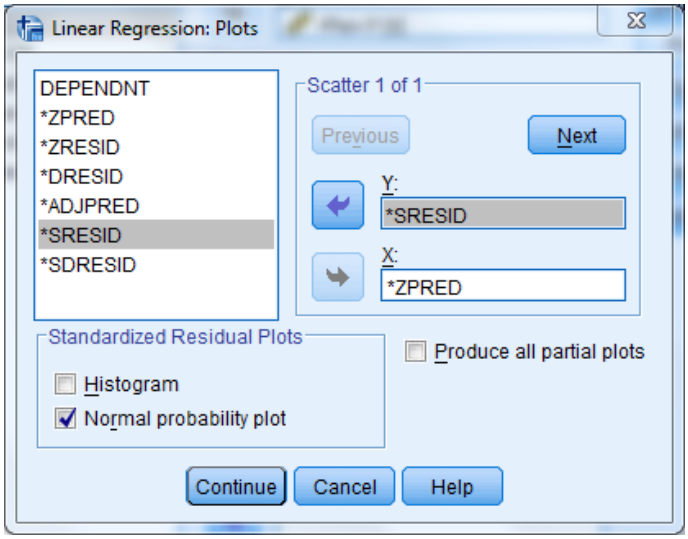
$$\hat{C} = -12.86 + 0.796.R$$

بحيث أن قيمة الحد الثابت معبر عنها بـ Constant وهي تساوي الى 12.86−، أما معلمة الميل (c_1) التي تظهر في السطر الأخير مقابل متغير الدخل وتساوي 0.796 وهي تشير الى الميل الحدي للاستهلاك بحيث أي زيادة في الدخل بنسبة 1% تؤدي الى زيادة في الاستهلاك بـ 0.796%. بينما تشير القيمة الموجودة في العمود Standarized Coefficients الى قيمة المعلمة c_1 المقدرة بـ

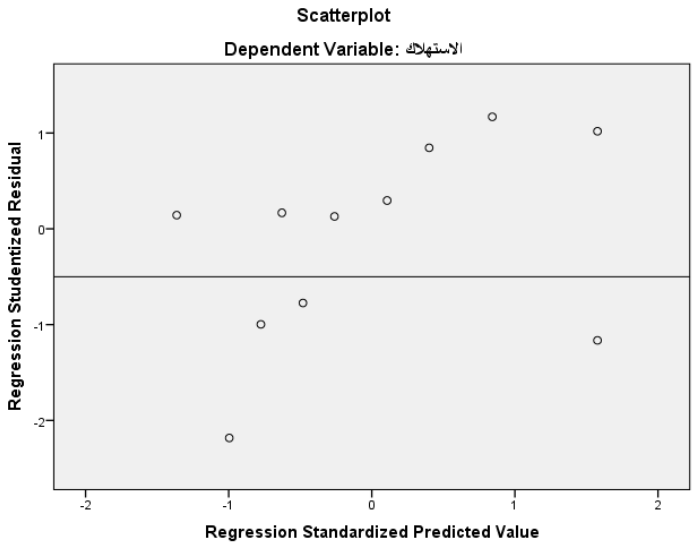
0.986 في النموذج باستعمال القيم المعيارية Standarized لكل من المتغير التابع والمتغير المستقل أي استعمال $(X-\bar{X})/S$ و $(Y-\bar{Y})/S$ عوض X و Y، والنموذج المقدر يكون بدون حد ثابت.

بالنسبة للعمود t فهو يشير الى قيمة احصائية ستودنت المحسوبة لكل معلمة، وهي 1.391- و17.492 لـ c_0 و c_1 على التوالي. وهي تشير الى عدم معنوية الحد الثابت لأن القيمة المحسوبة أقل من القيمة المجدولة عند مستوى معنوية 5%، بينما معلمة الميل فهي معنوية لأن القيمة المحسوبة أكبر من القيمة المجدول عند نفس المستوى والتي تقدر بـ 2.26، كما يمكن الاستعانة بالقيمة الاحتمالية Sig. ومقارنتها بمستوى المعنوية 0.05، بحيث تقبل H0 إذا كانت قيمة Sig. أكبر من 0.05.

وفي الاخير يعرض البرنامج مختلف المخططات البيانية لتحليل البواقي واختبار الفرضيات الخاصة بها، وذلك من خلال التأشير على الخيارات المتوفرة في صندوق حوار Plots المشار اليه سابقا، وغالبا ما يتم استخدام شكل الانتشار للعلاقة بين القيم المقدرة للمتغير التابع (PRED) والأخطاء العشوائية (RESID)، بحيث يتم نقل المتغيرات كما توضحها الصورة المقابلة:



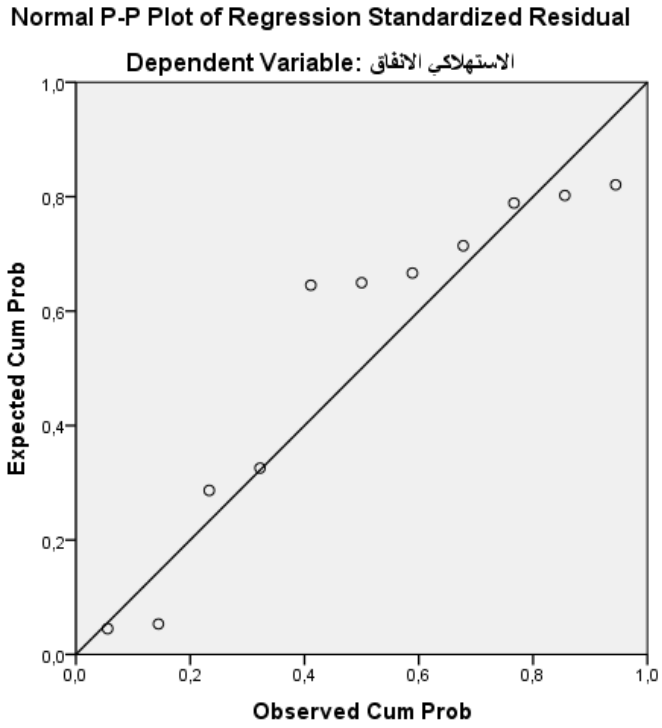
مع التأشير على الخيار Normal probability plot، وبتنفيذ الأمر يعرض البرنامج بالإضافة الى الجداول السابقة المخطط البياني التالي:



وهو يشير الى أن الأخطاء العشوائية تتوزع طبيعيا نظرا لكون 95% من نقاط الانتشار تقع داخل المجال [-2،2].

ويمكن تدعيم ذلك بمخطط بياني آخر

بتمثيل العلاقة بين الأخطاء المعيارية والقيم الحقيقية للمتغير التابع DEPENDNT محل ZPRED، في صندوق حوار Plots، فيظهر لنا الشكل التالي:



يظهر من خلال الشكل أن معظم نقاط الانتشار قريبة من الخط المستقيم، وهذا دليل على أن الأخطاء العشوائية تتوزع طبيعيا. كما يمكن للباحث استخدام طرق اختبار التوزيع الطبيعي للعينات التي أشرنا اليها سابقا على القيم المقدرة للبواقي.

ملاحظة:

يمكن للباحث استخدام الزر Save لحساب مختلف القيم المعيارية، والتي تضاف تلقائيا الى ورقة Data View والتي تفيد في اعداد مختلف المخططات البيانية التي تساعد في اختبار الفرضيات المتعلقة بالاخطاء العشوائية سواء باستخدام الأوامر المتاحة في البرنامج او المخططات البيانية المختلفة، كمخطط الصندوق والمدرج التكراري وغيرها. بحيث يظهر لنا بعض المتغيرات التي مكن اضافتها من خلال الزر Save كما يلي:

ind	C	R	PRE_1	ZRE_1
1	75	100	66,68995	,84950
2	70	125	86,57746	-1,69463
3	93	140	98,50997	-,56326
4	112	150	106,46498	,56582

4-2. تحليل الانحدار الخطي المتعدد:

يكمن الاختلاف بين نموذج الانحدار الخطي البسيط والمتعدد، في كون هذا الأخير يتكون من مجموعة من المتغيرات المستقلة بدلا من متغير واحد فقط. وعليه يأخذ الشكل العام لهذا النموذج الشكل التالي:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_k X_{ki} + \varepsilon_i$$

بحيث: k هو عدد المتغيرات المستقلة في النموذج.

ويقوم هذا النموذج على نفس فرضيات النموذج البسيط بالاضافة الى فرضيتين اساسيتين هما:

- لا توجد علاقة بين المتغيرات المستقلة فيما بينها في حالة النموذج متعدد المتغيرات (فرضية دم وجود تعدد خطي).
- ان تكون عدد المشاهدات أكبر من عدد المعلمات في النموذج، وهو شرط ضروري للتقدير بطريقة المربعات الصغرى العادية).

ولتوضيح خطوات استخدام البرنامج لتقدير هذا النوع من النماذج نستعين بالمثال التالي:

مثال (4-14):

لتكن لدينا البيانات بثلاث متغيرات Y، X1 و X2، الموضحة في الجدول التالي:

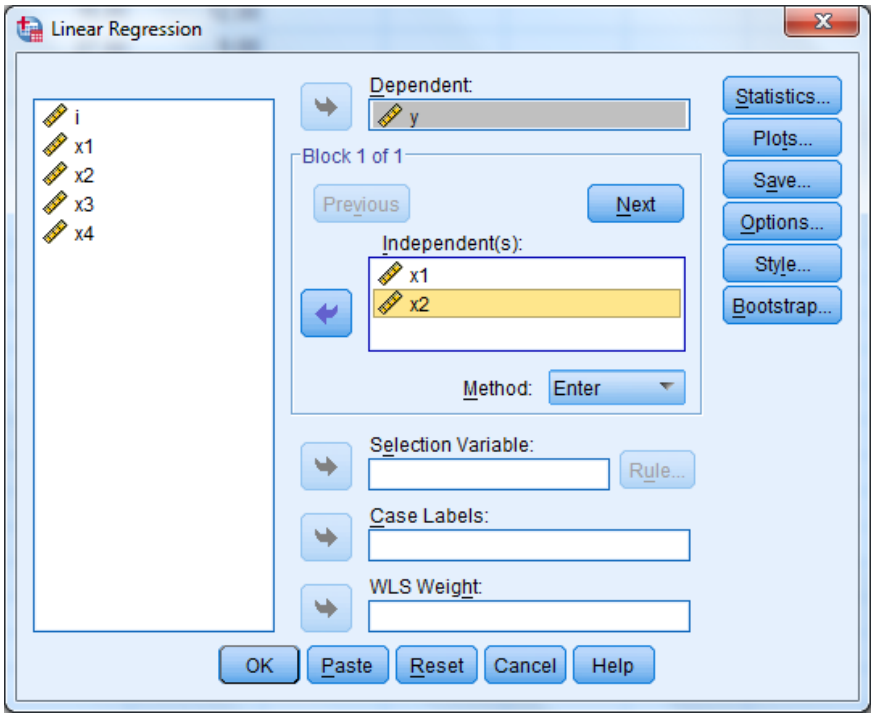
i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
y	11,48	9,34	8,42	10,07	9,24	8,86	6,21	8,25	8,03	7,47	5,91	7,95	6,13	5,86	3,16	5,87
x1	2,26	2,54	3,07	2,91	2,73	2,77	3,59	3,23	2,6	2,89	3,77	3,64	2,82	2,96	4,24	3,69
x2	3,49	2,85	4,06	3,64	3,21	3,66	3,76	3,49	3,13	3,2	3,65	3,6	2,94	3,12	3,58	3,53
x3	158,1	173	165	172,9	178	199	186	189	180	183	182	185	184	188	176	188

المطلوب:

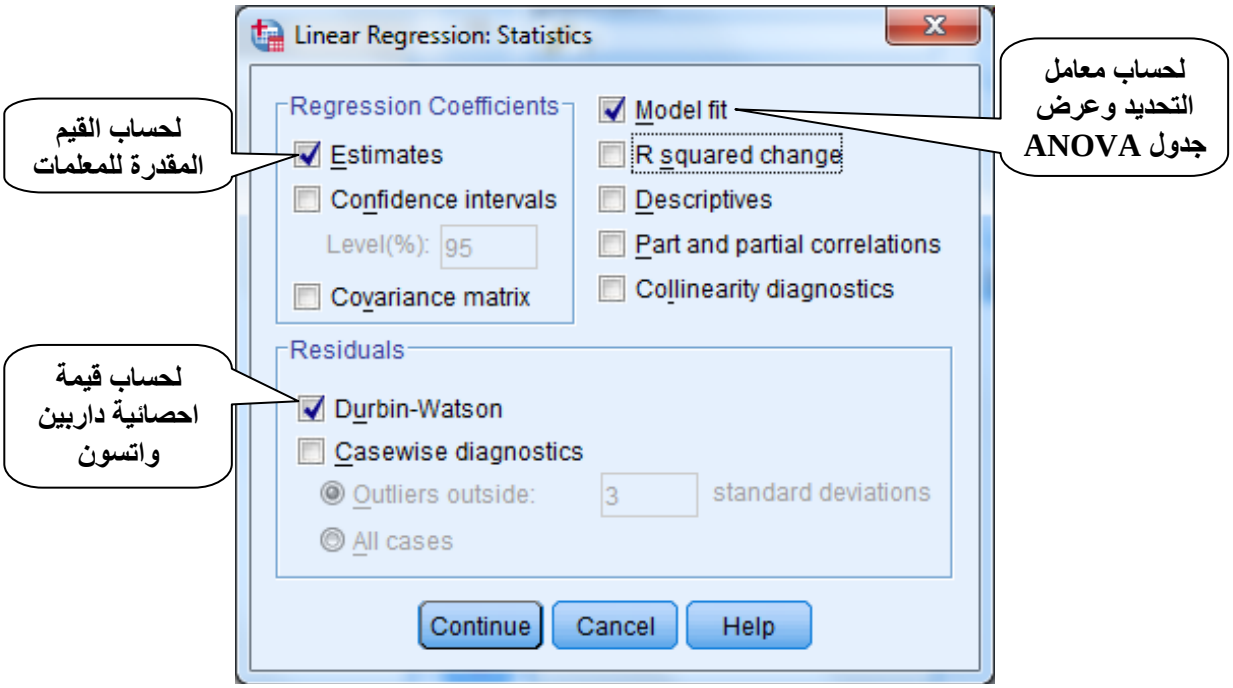
تقدير نموذج انحدار خطي بمتغيرين بطريقة المربعات الصغرى العادية مع اختبار جودة التوفيق والارتباط الذاتي للأخطاء بالإضافة الى اختبار معنوية المعلمات المقدرة باستعمال الاختبارات المناسبة.

للإجابة على المطلوب في هذا المثال بالاستعانة ببرنامج SPSS نتبع الخطوات التالية:

نختار الأمر Regression من قائمة Analyze واختيار Linear فيظهر لنا صندوق الحوار الخاص بـ Linear Regression على الشكل التالي:



بحيث يتم نقل المتغير y الى المجال Dependent كونه متغير تابع، والمتغيرين X_1 وX_2 الى المجال Independent(s) باعتبارها متغيرات مستقلة مفسرة للمتغير التابع. ثم ننقر على الزر Statistics من أجل استخراج قيمة معامل التحديد وقيمة احصائية داربين واستون DW لاختبار الارتباط الذاتي للأخطاء. وذلك بالتأشير على الخانات الموضحة في الصورة التالية:



بالنقر على الزر Continue والعودة الى صندوق الحوار الرئيسي نتابع العملية بالنقر على الزر OK فتظهر لنا النتائج التالية:

الجدول رقم (1): معامل التحديد وإحصائية داربين واتسون:

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	0,877[a]	0,770	0,735	1,05191	2,209

a. Predictors: (Constant), x2, x1

b. Dependent Variable: y

الجدول رقم (2): جدول تحليل التباين:

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	48,151	2	24,076	21,758	0,000[b]
	Residual	14,385	13	1,107		
	Total	62,536	15			

a. Dependent Variable: y

b. Predictors: (Constant), x2, x1

الجدول رقم (3): نتائج تقدير معلمات النموذج:

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	)Constant(	9,724	2,891		3,364	0,005
	x1	-3,779	0,573	-0,995	-6,595	0,000
	x2	2,814	0,948	0,448	2,967	0,011

a. Dependent Variable: y

من خلال نتائج التقدير يمكن كتابة النموذج المقدر على الشكل التالي:

$$Y_i = 9.724 - 3.779X_{1i} + 2.184X_{2i}$$
$$(3.364)\ \ (-6.595)\ \ (2.967)$$
$$R^2 = 0.877; \qquad DW = 2.20$$

تفسير النتائج:

- من خلال الجدول رقم (1)، نلاحظ أن قيمة معامل التحديد R^2=0.877، وهذا يدل على أن القدرة التفسيرية للنموذج كبيرة بحيث أن 87.7% من المتغير التابع مفسرة بالعلاقة الخطية المقدرة، وبالتالي بقاء نسبة ضعيفة من المتغير التابع غير مفسرة (12.3%) وهذا دليل على وجود عوامل

(متغيرات) أخرى تدخل في تفسير المتغير التابع لم يتم ادراجها في النموذج يتم التعبير عليها بالبواقي (الأخطاء العشوائية). كما نلاحظ أن قيمة احصائية داربين واتسون المقدرة بـ 2.20 فهي تدل على عدم وجود ارتباط ذاتي للأخطاء العشوائية كونها أقل من القيمة المجدولة $4-d_u=4-1.75=2.25$.

– من خلال الجدول (2)، نلاحظ أن مجموع مربعات الانحرافات غير المفسرة SSR=14.385 أكبر من مجموع مربعات الانحرافات المفسرة SSE=48,151 وهذا دليل على أن نسبة التباين في المتغير التابع المفسرة بالعلاقة المقدرة كبيرة، كما أن القيمة المحسوبة لاحصائية فيشر المقدرة بـ F=21.578 أكبر من القيمة المجدولة عند مستوى 5% ($F_{(2,12)}=3.89$)، مما يقودنا الى رفض فرضية العدم التي تقر بأن المعلمات المقدرة (من غير الحد الثابت) لا تختلف معنويا عن الصفر، وقبول الفرضية البديلة التي تقر بأن المعلمات المقدرة تختلف معنويا عن الصفر عند مستوى 5%.

– من خلال الجدول (3) الذي يظهر القيم المحسوبة لاختبار Student نلاحظ أن القيمة المحسوبة بالنسبة لمعلمات النموذج المقدرة أكبر من القيمة المجدولة عند مستوى 5% ($t_{12,0.025}=2.17$)، وبالتالي نرفض فرضية العدم ونقبل الفرضية البديلة والإقرار بأن المعلمات المقدرة تختلف معنويا عن الصفر عند مستوى معنوية 5%، وهذا يدل على أنها تساهم في تفسير المتغير التابع، وهذا ما تبين لنا من خلال قيمة معامل التحديد سابقا.

وعموما من خلال هذه النتائج يمكن الاقرار بأن المتغيرين المدرجين في النموذج تساهم إلى حد كبير في تفسير التباين في المتغير التابع، بناءا على مختلف الاختبارات الإحصائية، لكن تشير قيمة معامل التحديد المقدرة (0.877) الى وجود متغيرات أخرى لم يتم إدراجها بحيث تبقى حولى 12.5% من التباين في المتغير التابع ليست مفسرة.

أما من الناحية الاقتصادية فيتم ربط النتائج المتوصل اليها بواقع الظاهرة المدروسة وغالبا ما تفسر معلمات المتغيرات المستقلة في النموذج، بنسبة التغير في المتغير التابع الناتج عن المتغير المستقل المعني، وفي حالات أخرى أين يتم تحويل النموذج باستخدام الدالة اللوغاريتمية فتتحول هذه النسبة الى التغير النسبي والذي يعرف اصطلاحا بالمرونة.

ويمكن للطالب التوسع أكثر باستخدام كل المتغيرات الموجودة في الجدول الخاص بالمثال (4-13) في النموذج واتباع نفس الخطوات المذكورة.

خاتمة:

حاولنا من خلال هذه المطبوعة توضيح بعض المفاهيم والطرق الاحصائية الأكثر استخداما من قبل الطلبة والباحثين في ميدان العلوم الاقتصادية، وبالاستناد إلى البرنامج المقرر من طرف الوزارة الوصية لطلبة السنة الاولى ماستر تخصص اقتصاد وتسيير المؤسسات في مقياس تطبيقات متقدمة في برنامج SPSS، تناولنا بعض تقنيات التحليل الاحصائي للبيانات والتي تعتبر جزء بسيط من الطرق والأساليب الكثيرة التي يمكن تطبيقها باستخدام البرنامج SPSS، والتي تستدعي الالمام الواسع بالجوانب النظرية لأساليب التحليل الاحصائي للبيانات وطرق القياس الاقتصادي. بحيث اعتمدنا على أمثلة بسيطة تساعد الطالب على الفهم وإعادة التدريب عليها بطريقة سهلة، ومن ثم فتح المجال لتعميمها على مختلف البيانات الممكن توفرها في البحوث والدراسات.

وقد اعتمدنا استخدام الاصدار 22 من البرنامج SPSS، كونه يتميز بنوع من التطور في معالجة البيانات مقارنة بالإصدارات السابقة له، لكن نشير الى أنه كل الإصدارت صالحة وتقوم بنفس الوظائف، ويبقى الفرق في بعض الاضافات يغلبها الجانب الشكلي وإضافة بعض التقنيات في تنفيذ الأوامر Command.

ويبقى هذا العمل عبارة عن بداية لأعمال أخرى في طور الانجاز للتوسع أكثر في استعمال البرنامج SPSS وبرامج أخرى في معالجة البيانات وتحليلها بطرق متقدمة. ونرجو من زملائنا وطلبتنا الاعزاء اشعارنا بأي ملاحظات او أخطاء لاحظوها في هذه المطبوعة من أجل العمل على تصحيحها.

قائمة المراجع:

1. جيلالي جلاطو: "**الاحصاء مع تمارين ومسائل محلولة**"، ديوان المطبوعات الجامعية، الجزائر، طبعة 2001.

2. رائد ادريس محمود الخفاجي وعبد الله مجيد حميد العتابي: " **الوسائل الاحصائية في البحوث التربوية والنفسية** (مفهومها، أهميتها وتطبيقاتها باستخدام الحقيبة الاحصائية SPSS)"، دار دجلة للنشر، الأردن، الطبعة الأولى 2015.

3. سعد زغلول بشير: "**دليلك الى البرنامج SPSS الاصدار العاشر**"، **المعهد العربي للتدريب والبحوث** الإحصائية، بغداد، 2003.

4. غيث البحر ، معن التنجي: " **التحليل الإحصائي للاستبيانات باستخدام برنامج IBM SPSS Statistics**"، مركز سبر للدرسات الإحصاييية رالسياسات العامة، تركيا، 2014.

5. محمد بلال الزعبي، عباس الطلافحة: "**النظام الاحصائي SPSS فهم وتحليل البيانات الاحصائية**"، دار وائل للنشر، عمان، الأردن ، الطبعة الثالثة 2012.

6. محمد شامل بهاء الدين فهمي: "**الاحصاء بلا معاناة المفاهيم مع التطبيقات باستخدام البرنامج SPSS**" مركز البحوث بمعهد الادارة العامة، المملكة العربية السعودية، الجزء الأول والثاني، 2005.

7. نبيل جمعة صالح النجار: "**الاحصاء التحليلي مع تطبيقات برمجية SPSS**"، دار الحامد للنشر والتوزيع، الأردن، الطبعة الأولى 2015.

8. Andrew Rutherford : **"Introducing Anova And Ancova A Glm Approach"**, SAGE Publications, First published, London, 2001.

9. Gérald Baillargeon, Fernando ouellet : «Analyse des données avec SPSS pour Windows» ; Les Edition SMG, Canada, 2008.

10. Glenn GAMST, Lawrence S MEYERS et A.J.GUARINO ; "**Analysis of variance designs: A Conceptual and Computational Approach with SPSS and SAS**", Published in the United States of America by Cambridge University Press, New York, 2008.

11.IBM : **"Guide d'utilisation du système central d'IBM SPSS Statistics 22"** , Copyright IBM Corp. 1989, 2013.

12.IBM : "IBM SPSS Statistics 22 - Guide abrégé", Copyright IBM Corp. 1989, 2013.

13. Jacqueline J. Meulman Willem J. Heiser : "**IBM SPSS Categories 20"**, **Copyright IBM Corporation 1989, 2011.**

14. Robert HO. : "**handbook of univariate and multivariate data analysis and interpretation with spss**", Chapman & Hall/CRC is an imprint of Taylor & Francis Group, 2006.