الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي
جامعة الشهيد حمة لخضر-الوادي-
كلية العلوم الاقتصادية والتجارية وعلوم التسيير

مطبوعة في مقياس:

تطبيقات أولية في تحليل المعطيات

مطبوعة دروس مدعمة بتمارين وتطبيقات
موجهة لطلبة السنة الثالثة موارد بشرية

إعداد: د.جديدي موسى

2022/2021

الفهرس	
المحتويات	**الصفحة**
المحور الاول: تقديم برنامج SPSS	
1. التعريف بالنوافذ الرئيسية لبرنامج SPSS	04
2. القوائم الرئيسة لبرنامج SPSS	07
3. إنشاء ملف بيانات جديد في برنامج SPSS	14
المحور الثاني: طريقة تحليل المركبات الاساسية ACP	
1. التقديم النظري لطريقة المركبات الاساسية	23
1.1.التحليل في فضاء المتغيرات	23
2.1.التحليل في فضاء المشاهدات	27
2. خطوات وطريقة الحساب	29
3. خطوات وطريقة المعالجة ياستخدام SPSS	29
4. مثال حسابي	38
5. مثال تدريبي	43
6. مثال تطبيقي للتحليل المركبات الرئيسية باستخدام SPSS	46
7. مثال تطبيقي للتحليل المركبات الرئيسية باستخدام Eviews	60
المحور الثالث: التحليل العاملي المتناظر AFC	
1. مجالات التطبيق	70
2. البيانات المستخدمة	70
3. أهدف التحليل العاملي المتناظر AFC	73
4. التشابه والجمع بين الفئات	75
5. مثال حسابي	76
6. مثال تطبيقي: التحليل العاملي المتناظر AFCباستخدام SPSS	78

المحور الرابع: التحليل العاملي التميزي AFD	
1. اهداف التحليل التميزي	94
2. الدالة التمييزية	95
1.2. قوة دالة التمييز	95
2.2. أنواع دوال التمييز	95
3. تقدير معلمات الدالة التميزية الخطية:	96
4. الاهمية النسبية للمتغيرات	99
5. اختبار قدرة دالة التمييز في التمييز بين مجموعتين	100
6. النقطة الفاصلة	104
7. نسبة الخطأ	104
8. مثال تطبيقي1: التصنيف إلى مجموعتين	107
مثال تطبيقي 2: التصنيف إلى أكثر من مجموعتين	123
المراجع:	133

المحور الاول:

تقديم برنامج SPSS

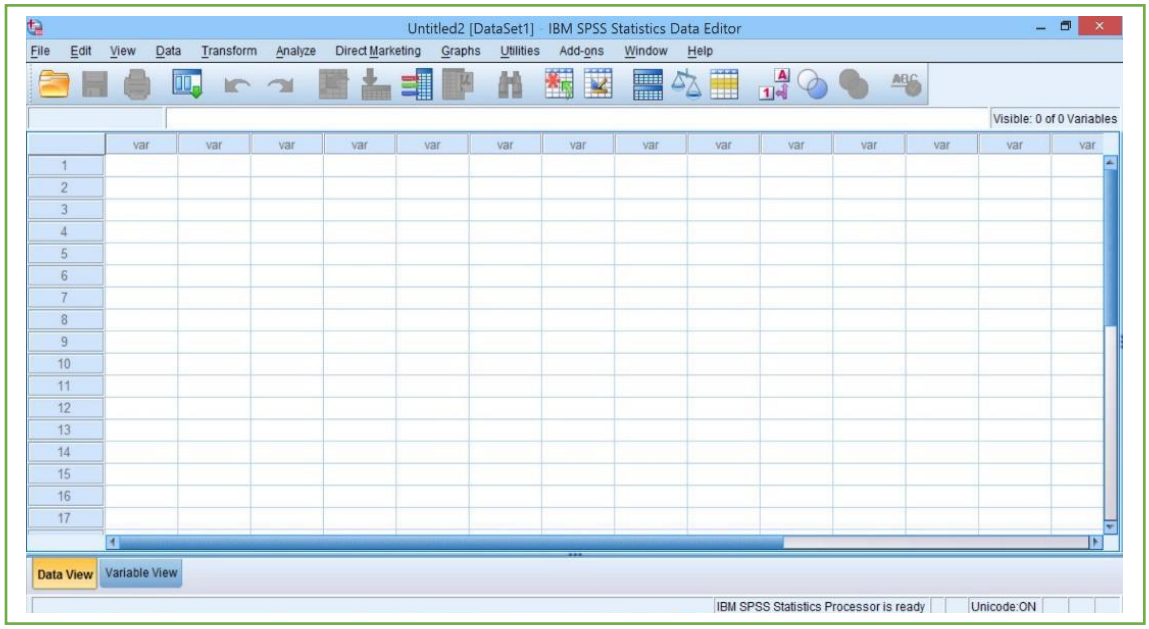
1. التعريف بالنو افذ الرئيسية لبرنامج SPSS

يحتوي نظام SPSS على نوعين من النوافذ هي:

نافذة محرر البيانات Editor Data ونافذة المخرجات Window viewer

1.1. محرر البيانات(Data View):

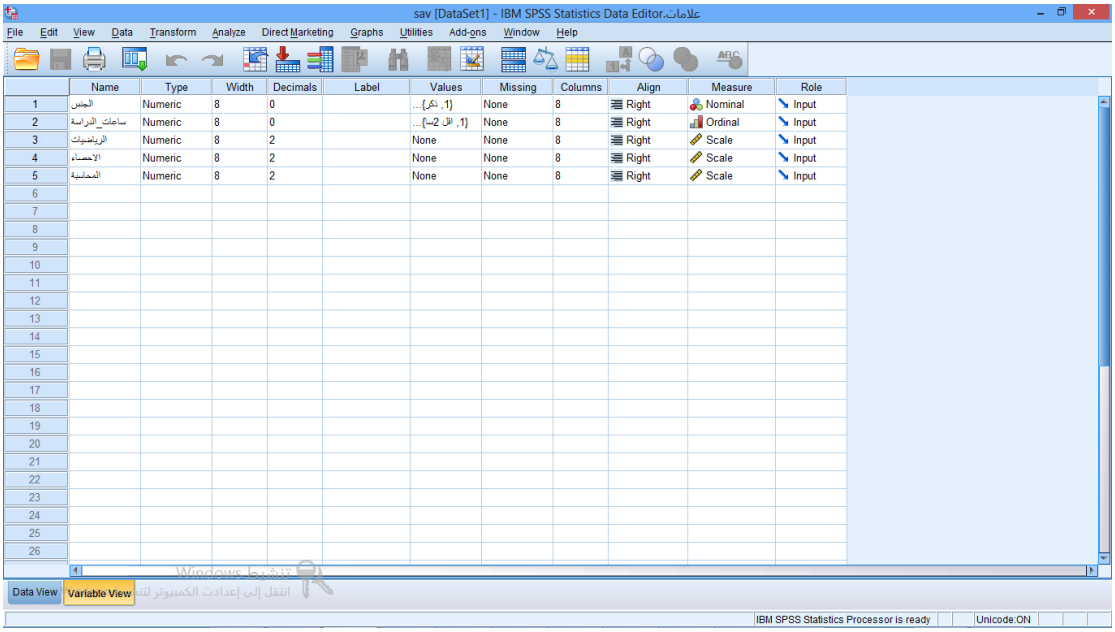
تظهر نافذة محرر البيانات عند فتحSPSSوتعرض مكونات ملف البيانات. تستخد م هذه النافذة لتعريف وإدخال وتحرير وعرض البيانات المراد تحليلها. ويمكن من خلال هذه النافذة أن نقوم بخلق ملفات بياناتجديدة أو تعديل ملفات بيانات موجودة. يمكن ملاحظة شريطين أسفل نافذة محرر البيانات Editor Data وهما شريط عرض البيانات Data View وشريط عرض المتغيرات Variable View. ويمكن التنقل بين الشريطين بالنقر على اسم الشريط المراد تفعيله.



عندما يكون شريط عرض البيانات نشطا فإنه يظهر البيانات المراد تحليلها والتي تم إدخالها ضمن البرنامج، تتضمن الصفوف (الأسطر) ضمن نافذة عرض البيانات الحالات التي تم إجراء القياس عليها أما الأعمدة فتتضمن متغيرات الدراسة.

مثال: في حالة الاستبيان تمثل الصفوف (الأسطر)المستجيبون، والاعمدة تحتوي على عبارات بحيث يحتوي كل عمود على عبارة واحدة من عبارات الاستبيان.

وعند النقر على شريط عرض المتغيرات VariableView تظهر نافذة تعريف المتغيرات التي تستخدم لعرض وتعريف المتغيرات

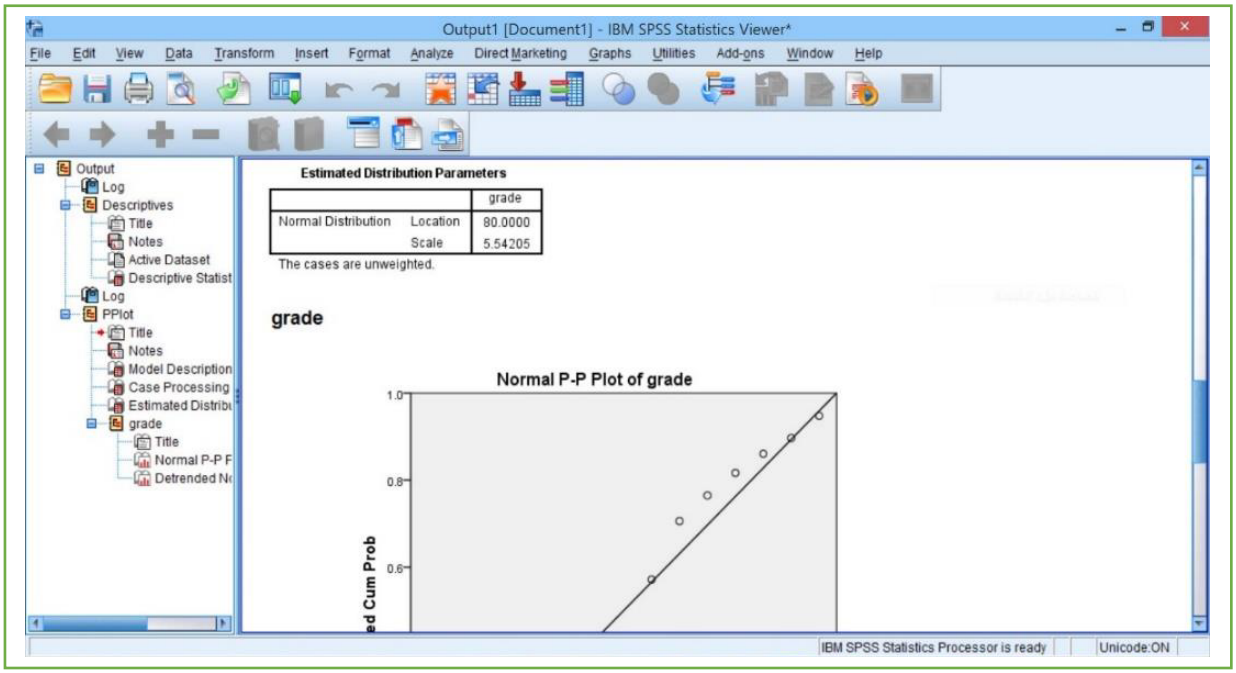


ويتم تعريف المتغيرات بالضغط على العمود مرتين Click او Double بالضغط على View Variable الموجود في أسفل الشاشة لتظهر شاشة أخرى لتعريف المتغيرات بتحديد اسم المتغير النوع، الحجم، العنوان، الترميز. ويتم

الترميـز بالضـغط علـى عـامود Values ومـن ثـم تحديـد قيمـة الرمـز ووصـفه مـع الضغط على مفتاح ADD لإضافة الرمز.

2.1. نافذة المخرجات Output Viewer :

تظهر نافذة المخرجات OutputViewer نتائج التحليلات الإحصائية والرسومات البيانية.

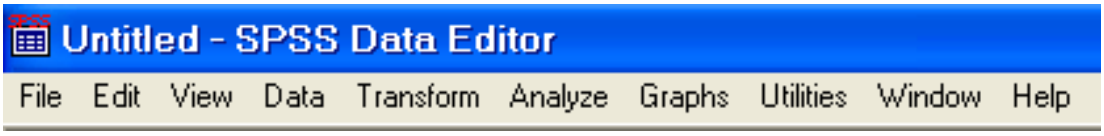


تنقسـم نافـذة المخرجـات الـى قسـمين ،القسـم الأيسـر يحتـوي علـى معلومـات خاصـة بنـوع الإجـراء لـذي تـم تنفيـذه، أمـا القسـم الأيمـن فيحتـوي علـى النتـائج نفسها سواء كانت جداول إحصائية أو رسومات بيانية أو نتائج اختبارات .

2. القوائم الرئيسة لبرنامج SPSS

1.2. شريط التعليمات:

يحتوي برنامج SPSS على مجموعة من القوائم والتي يمكن من خلالها القيام بجميع العمليات المطلوبة من البرنامج. ويوجد في برنامج SPSS على 10 قوائم رئيسة وهي:



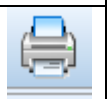
والتي نخلص أهم مهامها في الجدول التالي:

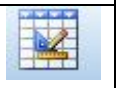
أسم القائمة	مهام القائمة
File Menu	- التحكم بالملفات - إنشاء ملف - فتح ملف أو عرض معلومات عن ملف - أو طباعة ملف - عرض قائمة بآخر الملفات التي تم استخدامها
Edit menu	- تحرير البيانات وتعديل البيانات مثل عمليات النسخ والقص واللصق وعمليات البحث عن المتغيرات
View Menu	- عرض وإخفاء شريط الأدوات وخطوط الشبكة في شاشة محرر البيانات - تعديل الخطوط المستخدمة في البرنامج
Data Menu	- تحديد المتغيرات وقيمها - ترتيب المتغيرات - دمج وفصل الملفات
Transform Menu	- تعديل قيم المتغيرات

	- حساب قيم جديدة للمتغيرات - إعادة ترميز المتغيرات - إنشاء القيم عشوائية
Analyze Menu	- تنفيذ أوامر التحليل الإحصائي المختلفة
Graphs Menu	- تنفيذ أوامر تمثيل البيانات
Utilities Menu	- عرض المعلومات عن المتغيرات والملفات - تحديد المجموعات الجزئية من التغيرات
Windows	- التحكم في النوافذ
Help Menu	تقديم المساعدة للمستخدم

2.2. شريط الأدوات Toolbar

يحوي شريط الأدوات مجموعة من الأيقونات والتي تمكن من الوصول المباشر إلى أحد الأوامر من القوائم المذكورة سابقاً، فعند النقر على إحدى الأيقونات، ينفذ الأمر المرتبط بهذه الأيقونة.

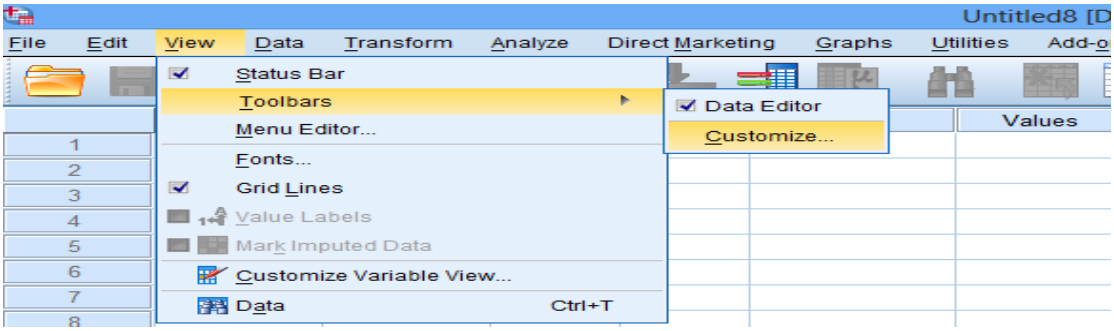
الأيقونة	العنوان	الوظيفة
	open	فتح ملف مخزن
	Save	تخزين ملف
	Print	طباعة ملف
	Dialog Recall	إظهار آخر مجموعة من الإجراءات التي تم استخدامها

	Undo	تراجع عن آخر عملية قمت بها
	Redo	الرجوع عن آخر عملية تراجعت عنها
	Goto Chart	الانتقال إلى تخطيط
	Goto Case	الانتقال إلى حالة (صف)
	Variable	إعطاء معلومات عن المتغير
	Find	بحث عن
	Insert Case	إدراج حالة جديدة إلى الملف
	Insert Variable	إدراج متغير جديد إلى الملف
	Split File	شطر الملف إلى جزأين
	Weight Cases	إعطاء أوزان للحالات
	Select Cases	اختيار مجموعة حالات
	Value Labels	إظهـــــار (أو إخفـــــاء) عنـــــاوين (دلالات) القيم
	Use Sets	استخدام مجموعات من المتغيرات

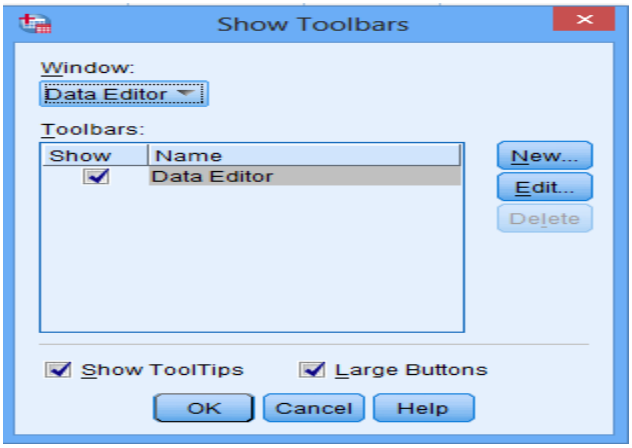
نشـــير إلـــى أنـــه يمكـــن تعـــديل شـــريط الأدوات بحـــذف أو إضـــافة الأيقونـــات التـــي يـــرى المستخدم بأنها مناسبة.

3.2. إضافة أيقونة لشريط الأدوات:

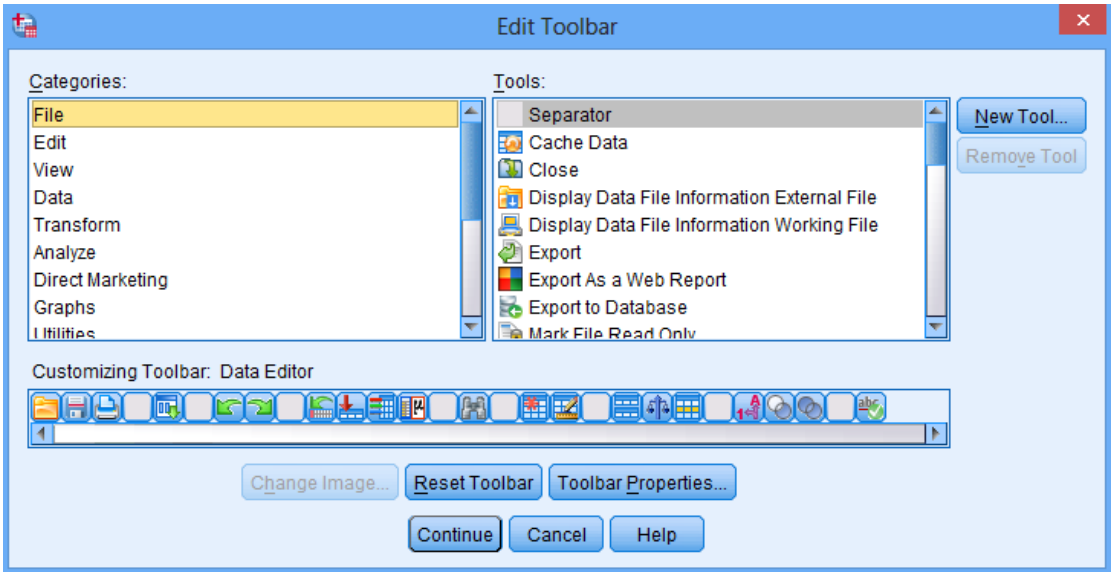
من قائمة View نختار Toolbars ثم Customize، كما هو في الشكل



وبالضغط على Customize، يظهر مربع الحوار التالي:



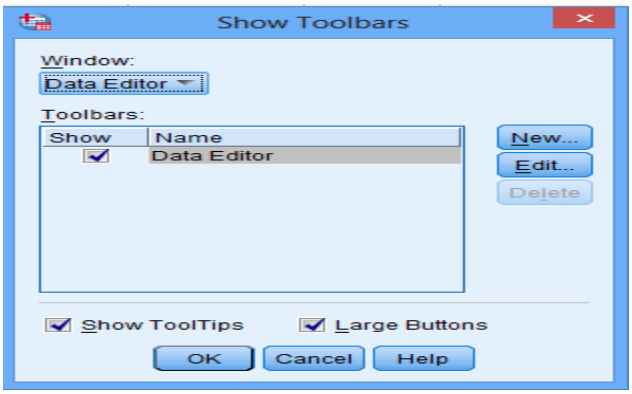
نضغط على Edit يظهر مربع الحوار التالي



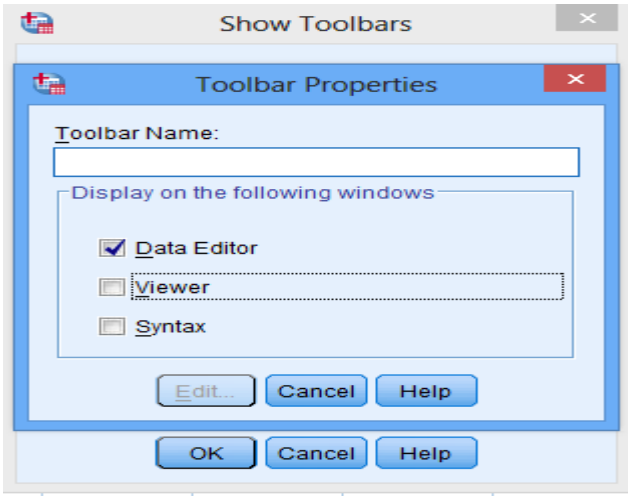
لأضافة أيقونة نختار من المستطيل Categories نوع القائمة التي تحتوي على الإيقونة، ثم من مستطيل Tools نضغط مرتين على الأيقونة التي نريد إضافتها إلى شريط الأدوات ، ثم نضغط على Continue.

4.2. إضافة شريط أدوات:

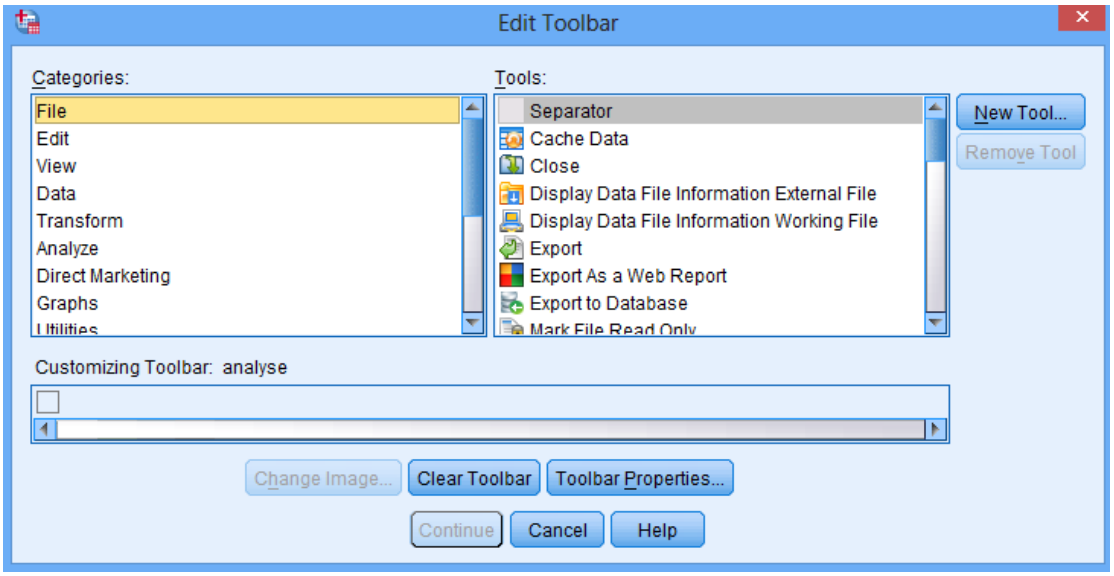
من المربع الحواري التالي نختار New



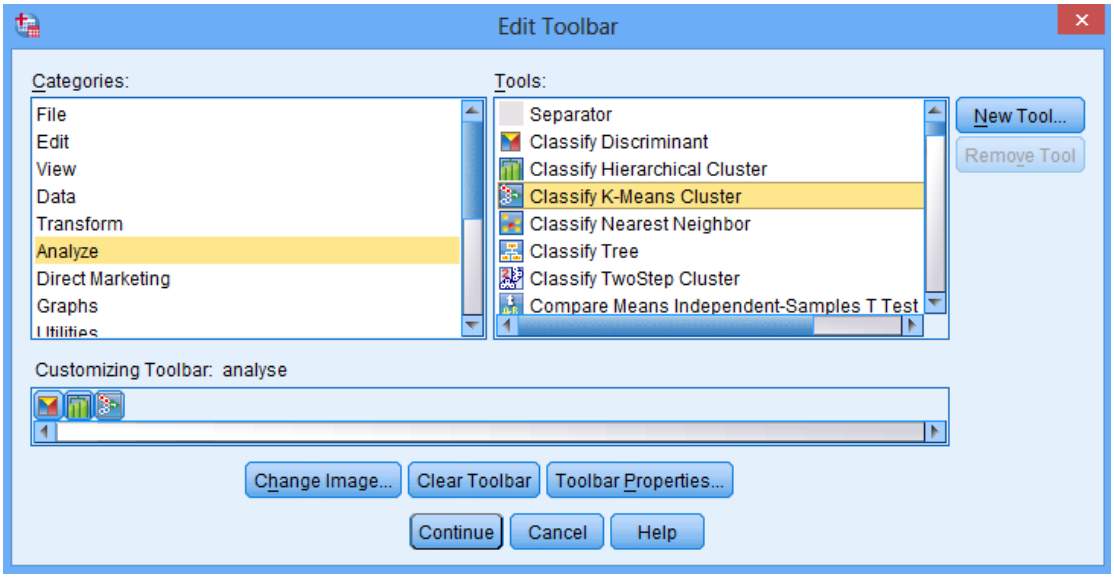
يظهر المربع التالي



نختـــار DataEditor ، و نكتـــب اســـم الشـــريط الجديـــد علـــى ســـبيل analyse، ثـــم نضغط Edit يظهر المربع الحواري التالي:

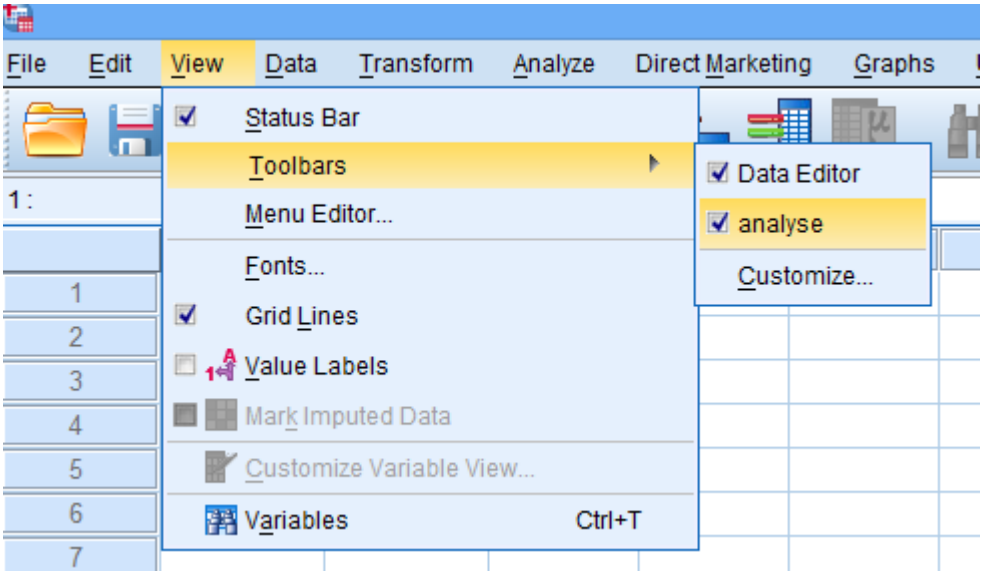


نختـــار مجموعـــة الأيقونـــات التـــي نريـــد إضـــافتها إلـــى شـــريط الأدوات الجديـــد analyse ثم نضغط على Continue. مثل نختار 3ايقونات كالتالي:



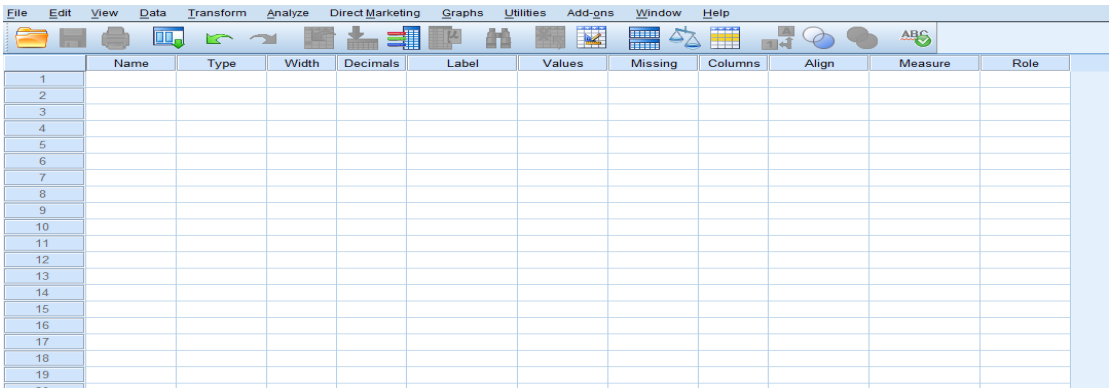
نلاحظ أنه تم إضافة شريط الأدوات الجديد كالتالي

كما أن شريط الأدوات الذي قمنا بإضافته تم إدراجه في قائمة View كالتالي:

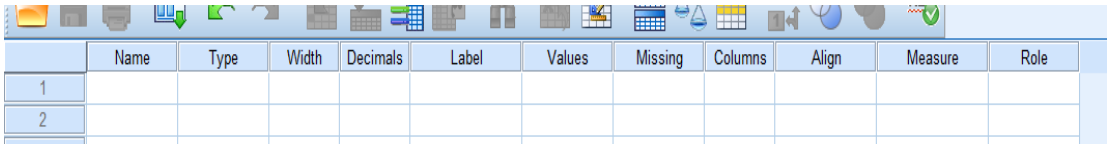


3. إنشاء ملف بيانات في برنامج SPSS

قبل إدخال قيم البيانات في View Data، يجب تعريف المتغيرات Variable View والتي تتطلب التعريف بخصائص المتغيرات.



تحوي VariableView عشرة أعمدة بحيث يحدد كل عامود إحدى خصائص المتغيرات.



والتي يتم تعريفها كالتالي:

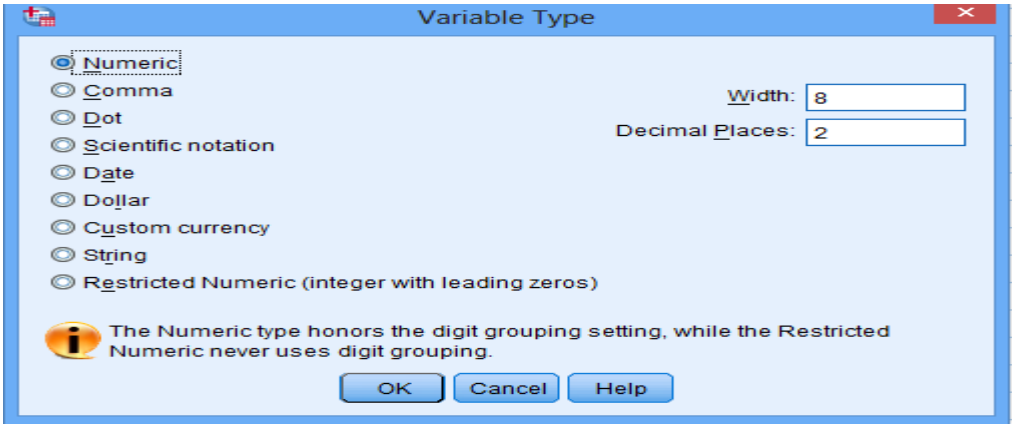
1. أسم المتغير Variable Name:

العمود الأول Name من ورقة View وهو العمود المخصص لكتابة Variable أسماء المتغيرات.

2. نوع المتغير Variable Type:

العمود الثاني Type من ورقة View وهو العمود المخصص Variable لتحديد ما إذا كان المتغير عددي أو غير عددي وكذلك طريقة عرض المتغيرات العددية في ورقة Data View.

عند النقر على الزر Numeric ... يظهر مربع الحوار التالي:



Numeric : نلاحظ أن SPSS يعتبر أن جميع المتغيرات رقمية وعرضها Width 8 أي 8 أرقام وكذلك عدد الأرقام العشرية Decimal Places 2 .

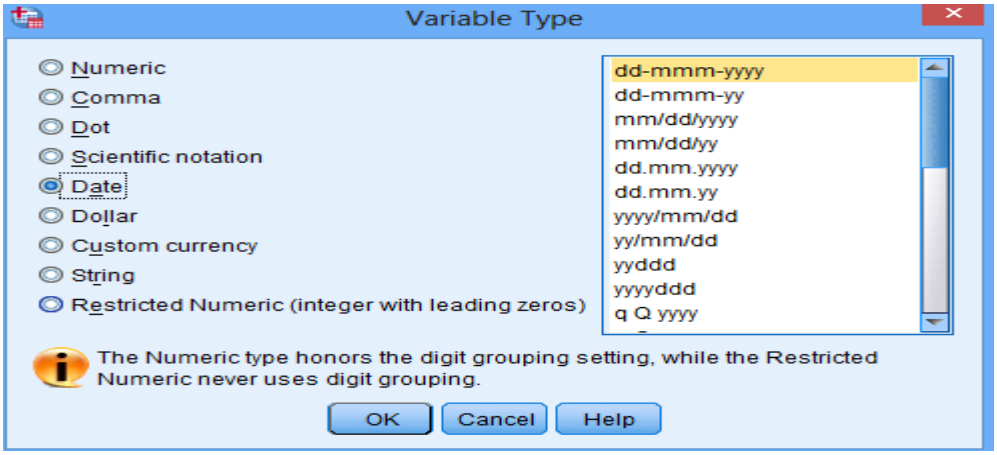
يمكن تغيير عرض أرقام العدد وكذلك عدد الأرقام العشرية بالضغط في خلية المناسبة Width أو Decimal Places.

Comma: تسمح بتعريف متغير رقمي نريد عرض قيمه بحيث تشتمل على فاصلة كل ثلاثة أرقام (للأرقام الأكبر من 1000) مع نقطة لفصل الخانات العشرية. كالعدد 199.786,000.

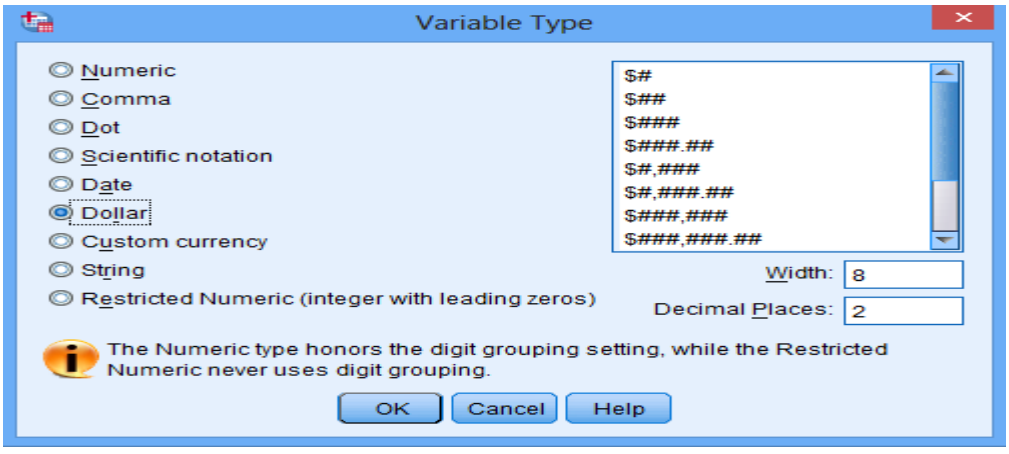
Dot: تسمح بتعريف متغير رقمي نريد عرض قيمه بحيث تشتمل على نقطة كل ثلاثة أرقام (للأرقام الأكبر من 1000) مع فاصلة لفصل الخانات العشرية. كالعدد 199,786.000.

Scientific Notation: تسمح بتعريف متغير رقمي نريد عرض قيمة بشكل تعبير أسي وفي هذا الحالة يستخدم الحرف E ليعبر على 10. يعبر على الرقم 516×10^2 بـ 5.16 E2 .

Date: تسمح بتعريف متغير رقمي نريد عرض قيمة بشكل تاريخ أو تاريخ مع الوقت ، كما يوضحه الشكل التالي :

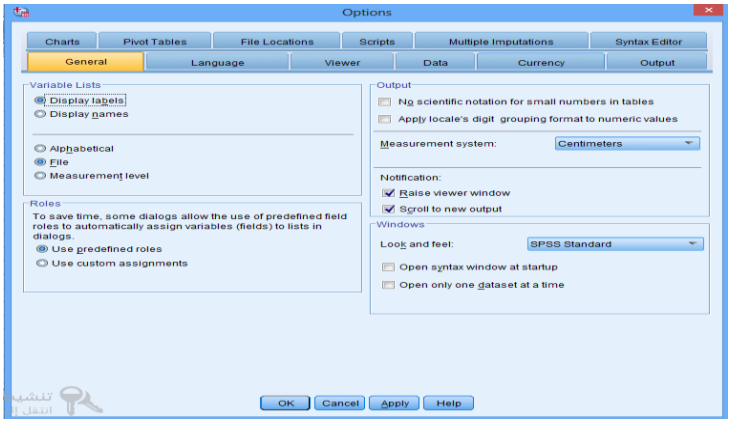


Dollar: تسمح بتعريف متغير رقمي نريد عرض قيمة بحيث تشمل على إشارة الدولار $، ، كما يوضحه الشكل التالي:

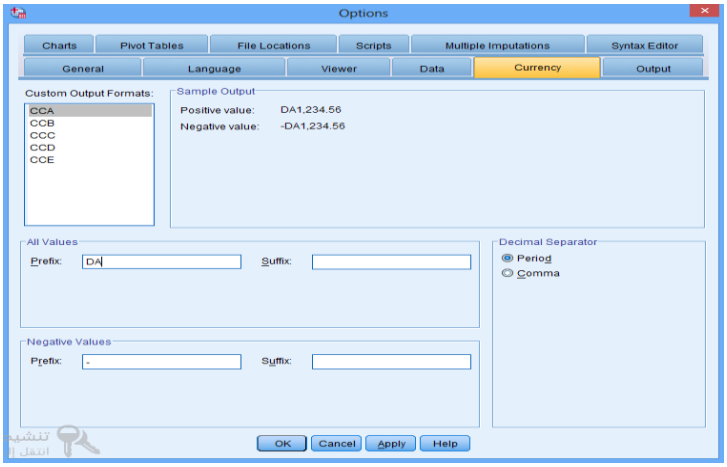


Currencey Custom : تسمح بتعريف متغير رقمي نريد عرض قيمة بحيث تشمل على عملة دولة معينة تم تعريفها . وبالتالي يجب قبل اختيار هذا الخيار تعريف العملة المطلوبة كما يلي:

من قائمة Edit اختار الأمر Options فيظهر مربع الحوار التالي:



اختار النافذة Currency ثم في Values All نكتب في المستطيل المقابل لـ Suffix رمز " DA " لتعبير عن العملة الوطنية.



String : تسمح بتعريف متغير حر قيمه تحتوي على أحرف أو أرقام أو أي رموز أخرى.

3. عرض المتغير Variable width:

العمود الثالث width من ورقة Variable من ورقة View وهو العمود Variable المخصص لتحديد عدد الخانات المستخدمة لعرض قيمة المتغير، ويمكن تحديد عرض المتغير بواسطة الضغط على الأسهم في الخلية المقابلة للمتغير.

4. عدد الخانات العشرية Decimals

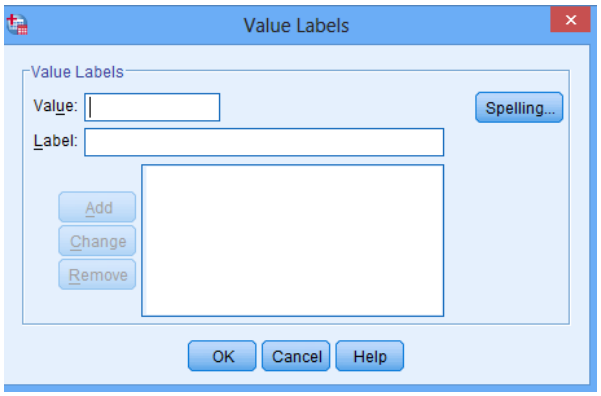
العمود الرابع Decimals من ورقة View وهو العمود المخصص Variable لتحديد المرتبة العشرية لعرض قيمة المتغير، وتحكم بها بالضغط على الأسهم إلى الأعلى وإلى الأسفل.

5. وصف المتغير Variable Label:

العمود الخامس Label من ورقة View وهو العمود Variable المخصص لتقديم وصف للمتغير، وتصل عدد الرموز إلى 256 .

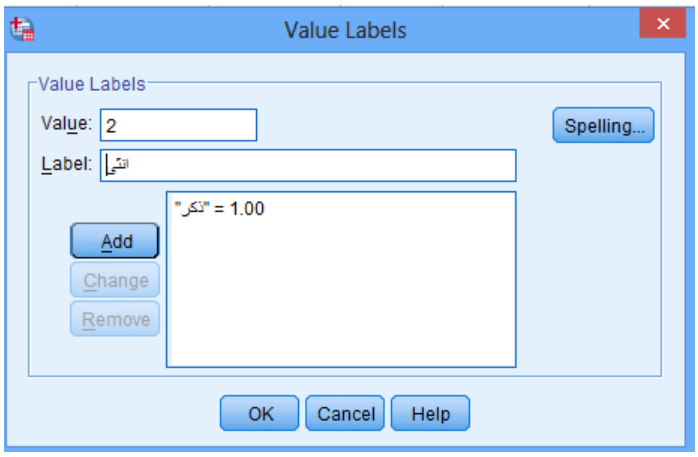
6. وصف القيمة Value Labels

العمود السادس Labels من ورقة View وهو العمود Variable المخصص لوصف القيم المحدد في البيانات وذلك عندما يكون المتغير وصفي، مثال الجنس حيث يعبر عليه عدديا 1 لذكر 2 للأنثى. ويظهر المربع التالي :



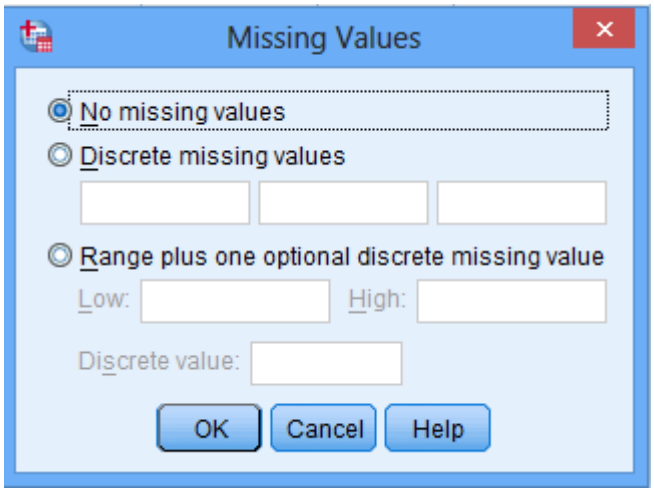
نكتب 1مقابل value، ونكتب ذكر مقابل value label، ثم نضغط add

نكتب 2مقابل value، ونكتب أنثى مقابل value label، ثم نضغط add



7. القيم المفقودة Missing Values

عند رغبتنا في تحديد بعض القيم على أنها قيم مفقودة (أي أن هذه القيم موجودة أصلاً ولكننا لا نرغب إدخالها في التحليل الإحصائي لأي سبب من الأسباب)، فإنه يمكن استخدام مربع الحوار التالي والذي يظهر عند النقر على الخلية التي تقع في العمود الذي يحمل العنوان Missing، وفي حال كانت القيم مفقودة أصلاً نتيجة لعدم وجود مشاهدات في البيانات، ففي هذه الحالة فإن الخلايا تكون فارغة وتحول تلقائياً إلى قيم مفقودة .



8. عرض العمود Column Width:

يمثل عرض العمود عدد الرموز المخصصة للمتغير، ويجب أن يكون عرض العمود أكبر من أو يساوي عرض المتغير المتضمن فيه، ويمكن تغيير عرض العمود لأي متغير بواسطة سحب حدود العمود في ورقة عرض البيانات.

9. محاذاة النصAlignment :

يسمح هذا العمود بضبط محاذاة النص داخل الخلايا لكل متغير، ويتم ذلك بالنقر على الخلية التابعة للمتغير ثم النقر على السهم المتجه للأسفل لاختيار المحاذاة المناسبة. مع العلم بأن المحاذاة الافتراضية هي (Right).

10. القياس Measurement:

العمود العاشر Measurement من ورقة View Variable وهو العمود المخصص لتحديد نوعية البيانات للمتغير والتي يمكن تصنيفها على النحو التالي:

- Scale: ويستخدم هذا التصنيف للبيانات العددية أو لإعطاء دلالة على أن المتغير متغير متصل.
- ordinal: ويستخدم هذا التصنيف لقياس المتغيرات الترتيبية حيث يمكن ترتيب قيم المتغير بحيث تعطي دلالة على أنه يمكن ترتيب القيم تصاعدياً أو تنازليا.
- nominal: ويستخدم هذا التصنيف لقياس المتغيرات الاسمية وهي متغيرات لها عدد من الفئات دون أفضلية لإحداها على الأخرى مثل ذكر وانثى.

المحور الثاني:

طريقة تحليل المركبات

الرئيسية ACP

استعملت هذه الطريقة لأول مرة من طرف pearson karl سنة1901، و أول من ضمها إلى الإحصاء الرياضي هو Hotelling Harold سنة 1933 ، غير أنها لم تصبح واسعة الإستعمال إلا في التسعينات، ذلك لظهور الحاسوب و البرامج الإحصائية التي سهلت العمل بهذه التقنية.

تعتبر طريقة تحليل المركبات الرئيسية أو المكونات الأساسية Analyse en composantes principales إحدى أهم طرق التحليل العاملي حيث تهتم بتقليص أكبر قدر ممكن من المتغيرات أو الأفراد إلى عوامل أو مركبات

رئيسية تحتوي على أكبر قدر من المعلومات أو التباينات الموجودة في المتغيرات أو الأفراد باستخدام التحليل الوصفي، وأيضا تعمل على إسقاطات النقاط في المستوي باستخدام العديد من أساليب تدوير المحاور لأجل الحصول على أحسن تمثيل.

كما يعتبر التحليل في المركبات الرئيسية أحد تقنيات تحليل التي تختص في اختزال البيانات ذات الأبعاد الكبير ، أي اختزال عدد كبير من المتغيرات الخام إلى عدد أقل من المتغيرات الجديدة في شكل مركبات وتكون غالبا أهدادها أقل بكثير من المتغيرات الخام

وعلى خلاف تحليل الإنحدار أو تحليل التباين الذان يعتبران متغير واحد تابع و البقية مستقلة فإن التحليل في المكونات الأساسية كل متغير يقارن ببقية المتغيرات، وإيجاد الجذور المميزة و المتجهات المميزة لمصفوفة التباين و التباين المشترك للمتغيرات التوضيحية، أو إيجاد الجذور المميزة لمصفوفة الارتباط، و هذا يعتمد على طبيعة البيانات.

1. التقديم النظري لطريقة المركبات الرئيسية

1.1. التحليل في فضاء المتغيرات

تشكل مصفوفة البيانات x_{np} سحابة من n نقطة في فضاء بعده p ، أو سحابة من نقاط p في فضاء ذو بعد n. يتكون الرسم التخطيطي الثنائي من إسقاط هذه النقاط على اثنين أو أكثر من الأبعاد المختارة بشكل عشوائي. يتمثل PCA من إسقاط النقاط على خط مستقيم ، مستوى ... مساحة فرعية من أبعاد s (مع p s) تم اختيارها لتحسين معيار معين. بشكل حدسي ، سوف نبحث عن الفضاء الجزئي الذي يعطي أفضل تصور ممكن لسحابة النقاط الخاصة بنا. الاختيار الجيد هو البحث عن أكبر تشتت ممكن أكبر انتشار في الفضاء الجزئي المختار. وبالتالي ، فإننا قادرين على البحث عن دوران لنظامنا الأولي من المحاور (المتغيرات) مما يسمح لنا برؤية السحابة بشكل أفضل. دعونا نحدد u_1 الشعاع المطلوب ($u_1'u_1=1$)؛ والذي يعطي أكبر تشتت للإسقاطات.

لتكن المصفوفة X_{nxp} ؛ كل سطر يمثل ملاحظة. يمثل كل عمود متغيرًا. سنفترض أن كل متغير يتم توسيطه ، أي أننا طرحنا متوسط كل متغير مسبقًا. يتم ذلك بطريقة تجعل مركز ثقل نقطة السحابة متزامنًا مع المركز.

تُعطى إسقاطات الملاحظات n على الشعاع u_1 من خلال:

$$C = Xu_1$$

مجموع مربعات هذه الاسقاطات هو:

$$C'C = u_1'X'Xu_1$$

سنختار u_1 لتعظيم هذه الكمية الأخيرة. لذا فإن المشكلة هي:

تعظيم $u_1'X'Xu_1$ وفقًا لـ $u_1'u_1 = 1$

هذه مشكلة يمكن حلها بطريقة لاغرانج.

نشكل لاغرانج:

$$L = u_1'X'Xu_1 - \lambda (u_1'u_1-1)$$

نشتق المعادلة بالنسبة لكل مكونات p للشعاع u_1 وكذلك بالنسبة إلى مضاعف لاغرانج (λ) ونساوي المشتقات الجزئية لصفر.

$$2(X'Xu_1 - \lambda u_1) = 0$$

$$u_1'u_1 = 1$$

بعد التبسيط نجد:

$$X'X\, u_1 = \lambda u_1$$

$$u_1'u_1 = 1$$

الشكل السابق يمثل معادلة الأشعة الذاتية والقيم الذاتية للمصفوفة X'X. وبالتالي ، فإن الشعاع الذي يعطي الإسقاطات ذات التشتت الأكبر هو الشعاع الذاتي الأول لمصفوفة التباين - التغاير لـ X.

ملاحظات:

- مصفوفة العددية (X'X) هي مصفوفة متماثلة وموجبة. هذا يعني أن القيم الذاتية و الأشعة الذاتية ستكون عددية. بالاضافة الى ذلك ، ستكون قيم القيم الذاتية موجبة أو صفرية.
- الأشعة الذاتية لمصفوفة متماثلة تكون دائمًا متعامدة مع بعضها البعض ، أي $u_1'u_2=0$.
- تحدد الإسقاطات الموجودة على الشعاع u_1 متغيرًا جديدًا وهو عبارة عن مجموعة خطية من المتغيرات الأصلية. يتم إعطاء تباين هذا المتغير الجديد من خلال:

$$1/n\ (u_1'X'Xu_1) = 1/n\ (u_1'\lambda 1u_1) = 1/n\ \lambda_1$$

وبالتالي فإن تباين الإسقاطات يساوي (للعامل 1/n) للقيمة الذاتية. وهكذا يتم الوصول إلى الحد الأقصى مع القيمة الذاتية الأولى. هذا هو السبب في الاحتفاظ بالشعاع الذاتي الأول.

- نعلم أن مجموع القيم الذاتية للمصفوفة يساوي أثر المصفوفة الأصلية. لذلك لا يتغير الحجم الإجمالي للتباين.
- تشير النسبة $\lambda_1/(\Sigma\lambda_i)$ إلى نسبة التباين الكلي الذي يدعمه الشعاع الذاتي الأول.

السؤال 1: ما هي العلاقة بين الأشعة الذاتية والقيم الذاتية لـ X'X وتلك الخاصة بـ X'X / n؟

- **إيجاد الشعاع الثاني u_2:**

نبحث عن الشعاع u_2 الذي يحقق:

تعظيم $u_2'X'Xu_2$

يخضع لـ $u_2'u_1 = 0$

و $u_2'u_2 = 1$

الحل يكمن في استخدام طريقة لاغرانج كما يلي:

$$L = u_2'X'Xu_2 - w\,(u_2'u_1) - \lambda_2\,(u_2'u_2 - 1)$$

نشتق لاغرانج فيما يتعلق بكل مكون من مكونات الشعاع u_2 وفيما يتعلق بـ u_1.

بعد التبسيط ، نجد:

$$X'X\,u_2 = \lambda_2 u_2$$

$$u_2'u_2 = 1$$

$$u_1'u_2 = 0$$

الشعاع الذي تم العثور عليه هو الشعاع الذاتي المرتبط بثاني أكبر قيمة ذاتية لمصفوفة التباين والتغاير. يمكن تعميم هذه النتائج بسهولة على عدة أبعاد ونجد ما يلي:

الفضاء ذو البعد s بحيث s≤p الذي يعطي أفضل تفسير يتم تعريفه بواسطة الأشعة الذاتية لـ X'X المرتبطة بأكبر قيم للقيم الذاتية.

2.1. التحليل في فضاء المشاهدات

لقد درسنا كيف يمكن تطبيق ACP لتحليل سحابة n من المشاهدات في فضاء المتغيرات p. يمكن أيضًا أن نقوم بالتحليل سحابة المتغيرات p في فضاء n من المشاهدات. نحن نبحث عن فضاء فرعي للبعد (p ≤ s) s حيث يكون مجموع مربعات الإسقاطات هو الحد الأقصى.

نطبق نفس التقنية السابقة ونجد أن حل الشعاع الأول يتم الحصول عليه من خلال:

$$XX'\nu_1 = \beta_1\nu_1$$

$$\nu_1'\nu_1 = 1$$

الشعاع الذاتي الأول لـ XX' هو الذي يزيد من تباين الإسقاطات. ، نشكل معادلة تعظيم تباين الإسقاطات من الشعاعين الذاتيين الأول والثاني ، إلى آخره.

نظرية: قيم الذاتية λ_i و β_i متطابقة.

لنضرب المعادلة بـ X':

نجد:

$$X'XX'\nu_i = \beta_i X'\nu_i$$

نلاحظ من خلال هذه المعادلة أن $X'v_i$ هو الشعاع ذاتي لـ $X'X$ مرتبط بقيمة β_i. ولدينا القيم الذاتية لـ $X'X$ تعطى بواسطة λ_i. ومنه نستنتج أن $\beta_i = \lambda_i$

ملاحظة: من خلال مما سبق فإنه ليس من الضروري البحث بشكل مباشر عن القيم الذاتية والأشعة الذاتية لـ XX' ، ويتم ايجاد الاشعة الذاتية u_i بواسطة $X'v_i$ ويمكن توضيح ذلك كما يلي:

$$v_i'XX'v_i = \lambda_i$$

ومنه:

$$u_i = \frac{1}{\sqrt{\lambda_i}} Xv_i$$

وبالضرورة يكون

$$v_i = \frac{1}{\sqrt{\lambda_i}} Xu_i$$

تسمى هذه الصيغة بصيغة الانتقال.

يمكن كتابتها في شكل مصفوفة كالتالي :

$$V = XU\Lambda^{-1/2} \quad U = X'V\Lambda^{-1/2}$$

مثال:

x	3	4	6	6	6	7	7	8	9	9	9	01	11	21	21	31	31	31	31	41	51	71	71	81	02
y	2	01	5	8	01	2	31	9	5	8	41	7	21	01	11	6	41	51	71	7	31	31	71	91	02

نطبق التحليل بالمركبات الرئيسة لمصفوفة التغاير

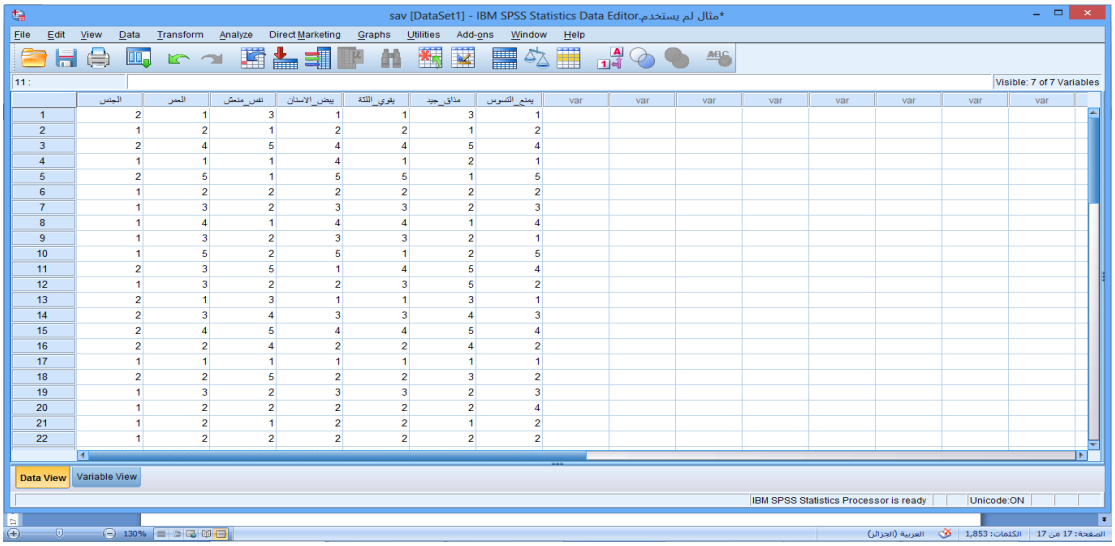
$$\Sigma = \begin{pmatrix} 20.28 & 15.59 \\ 15.59 & 24.06 \end{pmatrix}$$

2. خطوات وطريقة الحساب:

- حساب المتوسط الحسابي والانحراف المعياري لكل متغير
- تحويل الجدول الأصلي إلى جدول ممركز أو جدول معياري ممركز
- حساب القيم الذاتية و الأشعة الذاتية
- حساب التباين الكلي المفسر
- حساب نسبة مساهمة المستوي والمحاور في تمثيل الأفراد
- حساب إحداثيات الأفراد وتمثيلها في المحاور الجديدة
- حساب النوعية النقطية للتمثيل (جودة التمثيل.)
- حساب المساهمة المطلقة للأفراد
- حساب إحداثيات المتغيرات وتمثيلها بيانيا في المستوي
- حساب الاحداثيات الإضافية للمتغيرات والأفراد.

3. خطوات وطريقة المعالجة ياستخدام SPSS:

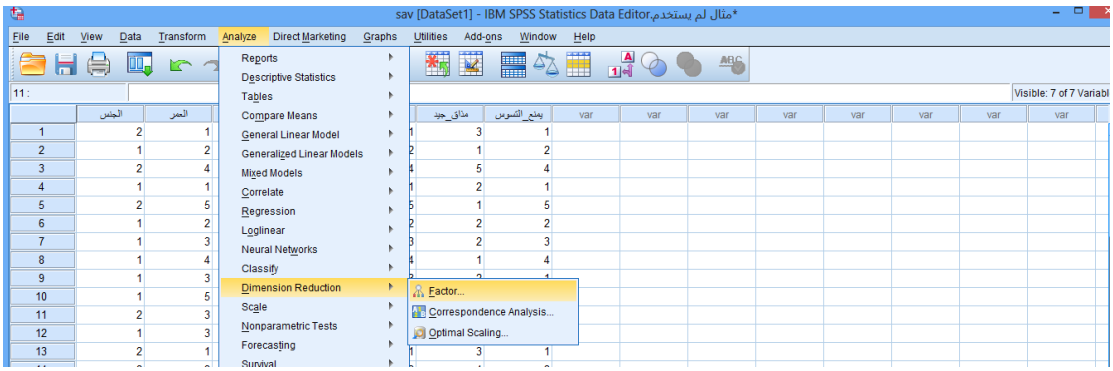
بعد عملية إدخال البيانات في برنامج SPSS كما هو موضح في الشكل التالي:



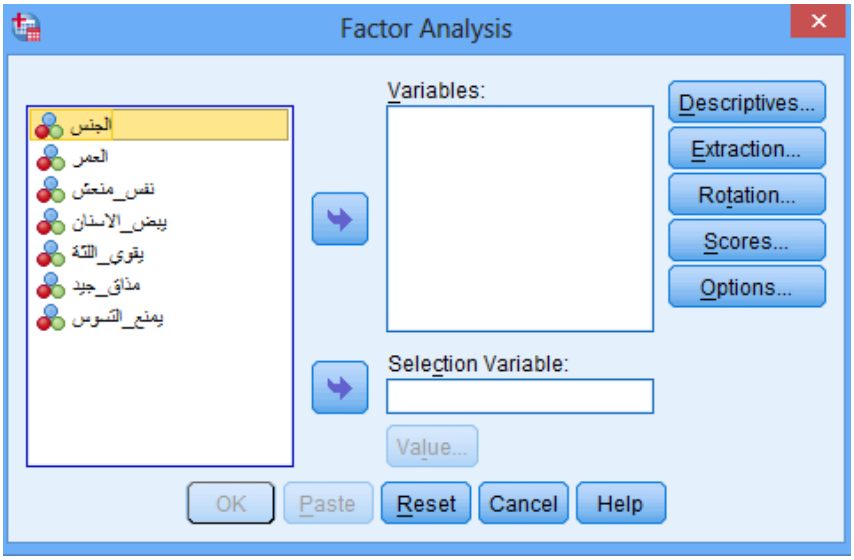
نقوم بعملية التحليل بطريقة المركبات الرئيسية بإتباع الخطوات التالية:

مـــن شـــريط التعليمـــات نختـــار Analyse ثـــم نختـــار Rediction ثـــم Dimension

Factor كما هو موضح في الشكل التالي:

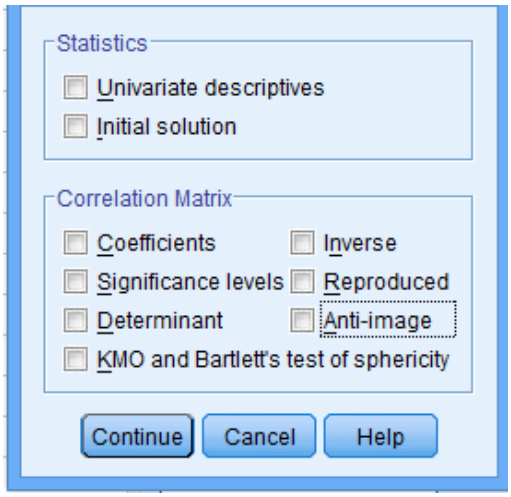


عنـــد ظهـــور مربـــع الحـــوار التـــالي يـــتم إدخـــال المتغيـــرات المـــراد تحليلهـــا كمـــا يحتـــوي على العديد من الاوامر التي نذكرها كالتالي:



- **Descriptves:**

بالضغط على مربع Descriptves يضهر مربع الحوار التالي

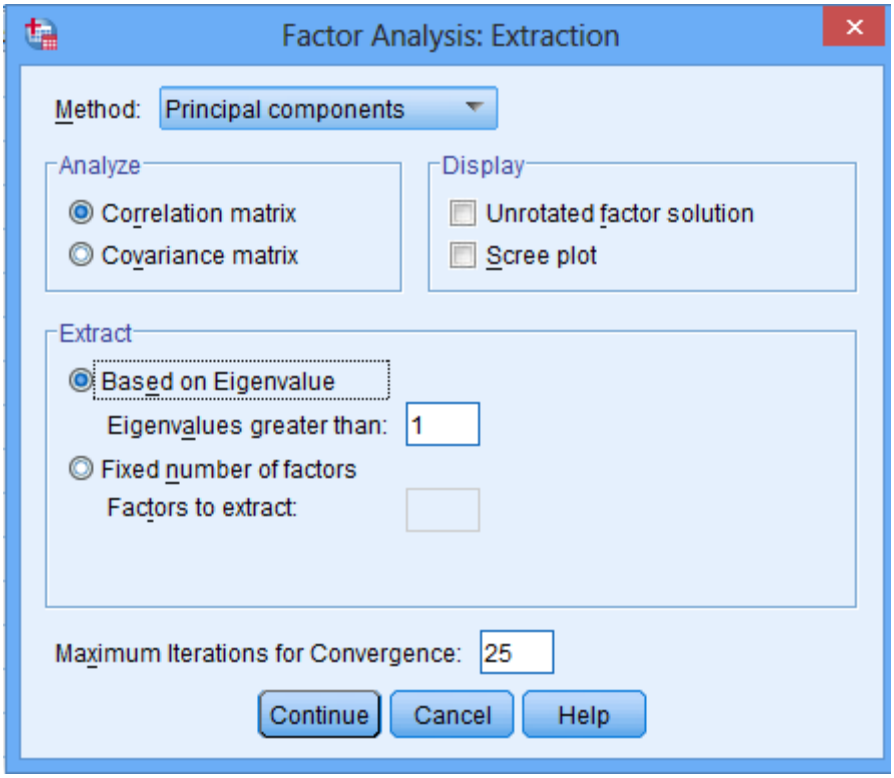


والـــذي مـــن خلالـــه نختـــار مؤشـــرات تحليـــل البيانـــات أحاديـــة البعـــد المتمثلـــة أساسا في الانحراف المعياري و المتوسط الحسابي لوصف المتغيرات.

كما نؤشر على شروط تطبيق التحليل في المكونات الأساسية .

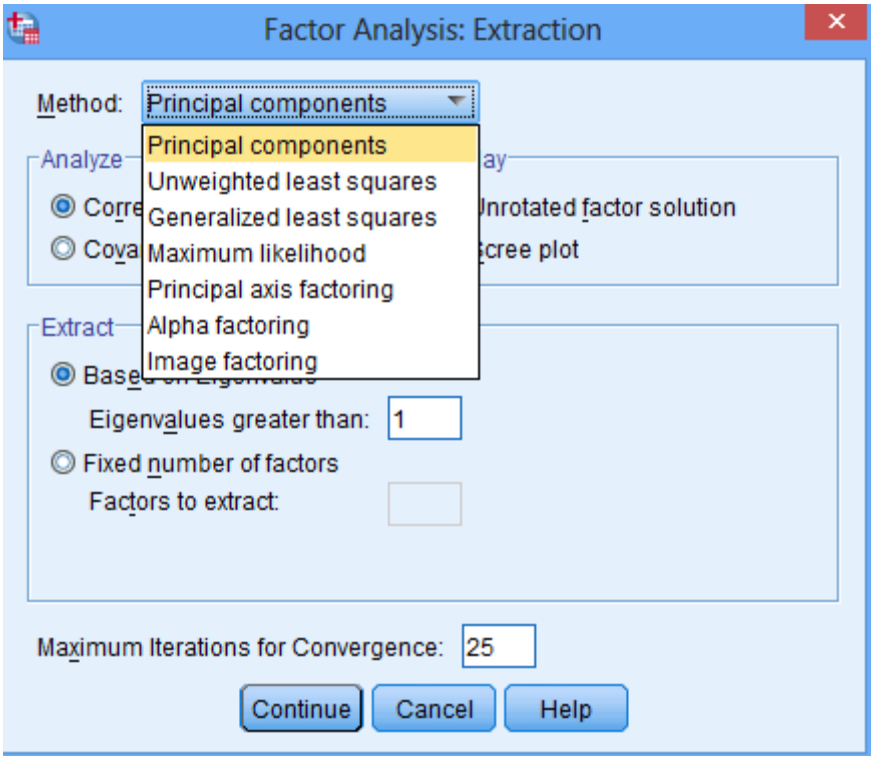
- **Extraction**

بالضغط على مربع Extraction يضهر مربع الحوار التالي:

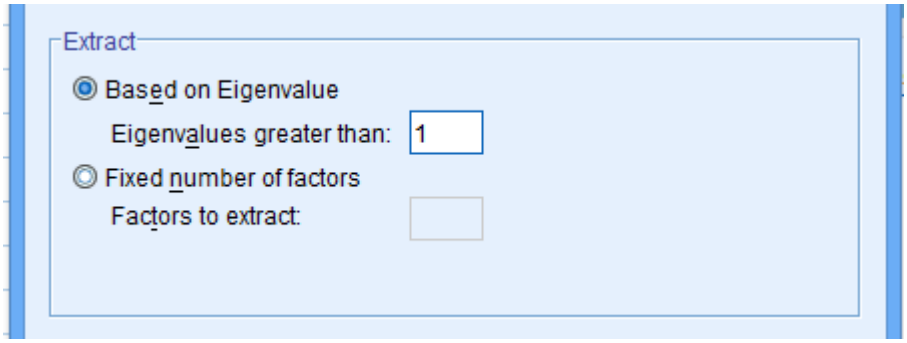


- **Method:**

ويحتوي على العديد من الطرق الخاص بالتحليل العملي، ويتم من خلالها تحديد الطريقة المناسبة لعملية التحليل

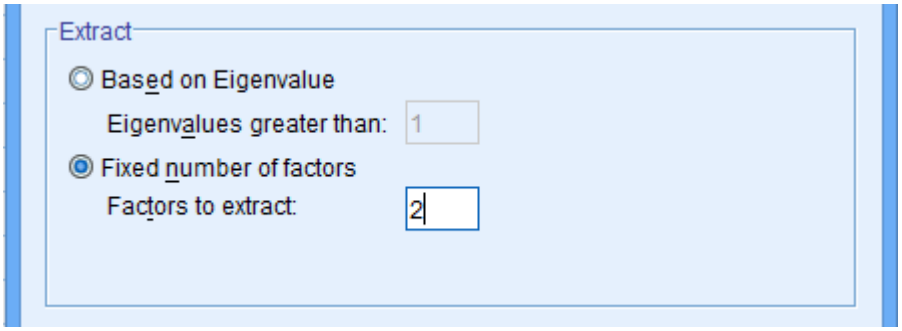


- **Extract:**



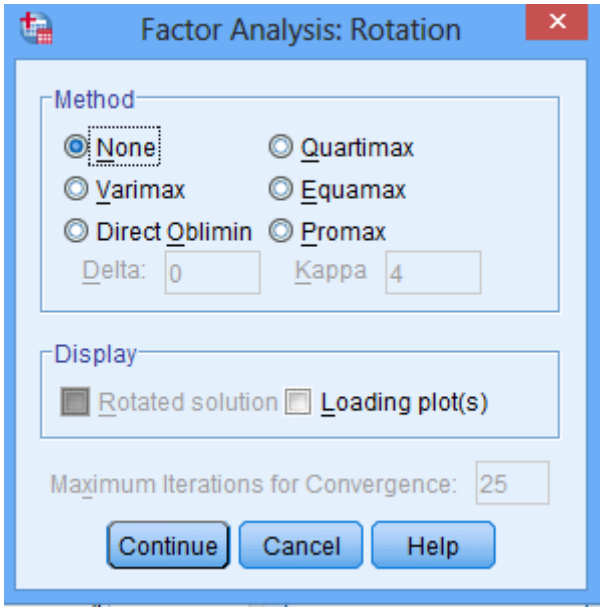
يتم اختيار عدد المكونات المستخرجة إما:

- بالاعتماد على قيمتها، في هذا المثال نستخر القيم الذاتية التي قيمتها أكبر من الواحد.
- بإختيار عددها، مثال نريد إستخراج قيمتين ذاتيتين دون اخذ بعين الاعتبار قيمتيهما، كما في الشكل التالي:

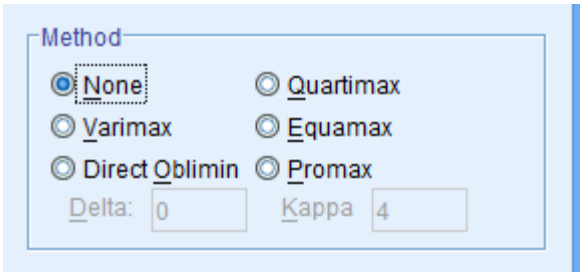


- **Rotation:**

بالضغط على مربع Rotation يضهر مربع الحوار التالي:



يسمح هذا المربع الحواري بإختيار نوع التدوير



التدوير المتعامد:

هنـــاك أكثـــر مـــن طريقـــة تحقـــق التعامـــد بـــين العوامـــل، أشـــهرها طريقـــة تعظـــيم التبـــاين Varimax، ويـــدل التعامـــد علـــى اســـتقلالية المحـــاور العامليـــة، وتســـمح هـــذه الطريقـــة بالاحتفـــاظ بالتعامـــد القـــائم بـــين العوامـــل الأصـــلية. ممـــا يعنـــي أن المحـــور العـــاملي الأول و المحـــور العـــاملي الثـــاني متعامـــدين لا يرتبطـــان مهمـــا تغيـــر موضع المحورين.

التدوير المائل:

يختلـــف هـــذا النـــوع مـــن التـــدوير عـــن الســـابق فـــي كونـــه يفتـــرض ارتبـــاط العوامـــل. هـــذا النـــوع مـــن التـــدوير صـــعب ومعقـــد. ومـــع هـــذا بعـــض البـــاحثين يفضـــلون استخدامه بدلا عن التدوير المتعامد لسببين:

- يتصف بالفاعلية في التعامل مع البيانات كبيرة الحجم
- يســـمح بتفســـير أبســـط تكـــوين ممكـــن للبنيـــة العامليـــة والتوصـــل إلـــى حلـــول يمكن من خلالها تفسير وتعليل العلاقات البينية.

طريقة تعظيم التباين:

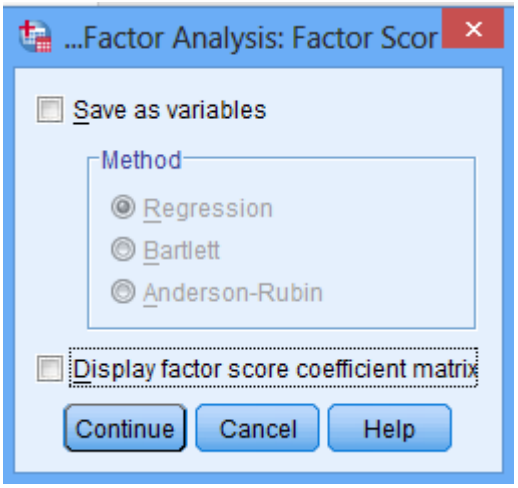
كطريقة التعظيم الرباعي Quartimax وطريقة ايكويماكس Equimax،

تهـــدف هـــذه الطريقـــة الـــى تقليـــل عـــدد المتغيـــرات التـــي تتشـــبع بشـــكل قـــوي بكـــل عامـــل، عـــن طريـــق تـــدوير جميـــع العوامـــل الممكنـــة مثنـــى مثنـــى علـــى حـــدة إلـــى أن يـــتم تعظـــيم مجمـــوع تباينـــات مربعـــات التشـــبعات فـــي أعمـــدة المصـــفوفة العامليـــة، والوصـــول فـــي النهايـــة إلـــى أن يكـــون لكـــل متغيـــر تشـــبع واحـــد عـــال علـــى أحـــد العوامـــل ومـــنخفض علـــى العوامـــل الأخـــرى. وينـــتج مـــن هـــذه العمليـــات إحـــداث تكـــافؤ فـــي

أهمية المتغيرات مما يؤدي إلى التقليل من عدد المتغيرات لكل عامل، وبالتالي يكون العامل أكثر وضوحا ويسهل قراءته وتفسير مكوناته.

Scores

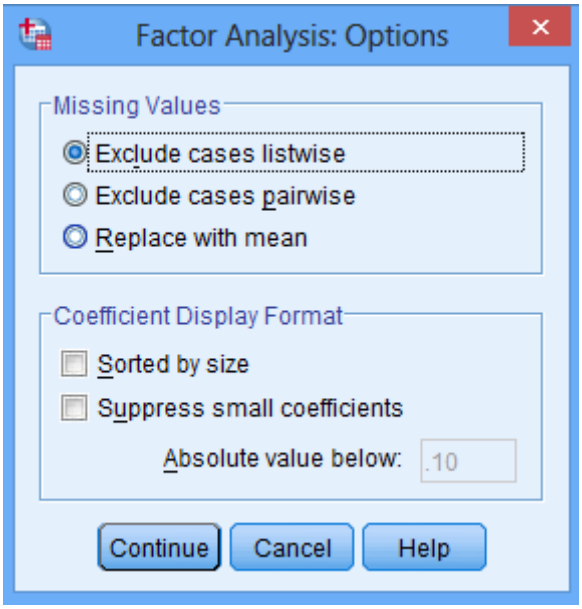
بالضغط على مربع Scores يظهر مربع الحوار التالي:



يسمح هذا المربع الحواري بحفظ احداثيات المركبات المستخرجة في قاعدة البيانات لإستعمالها في التمثيل البياني.

Option

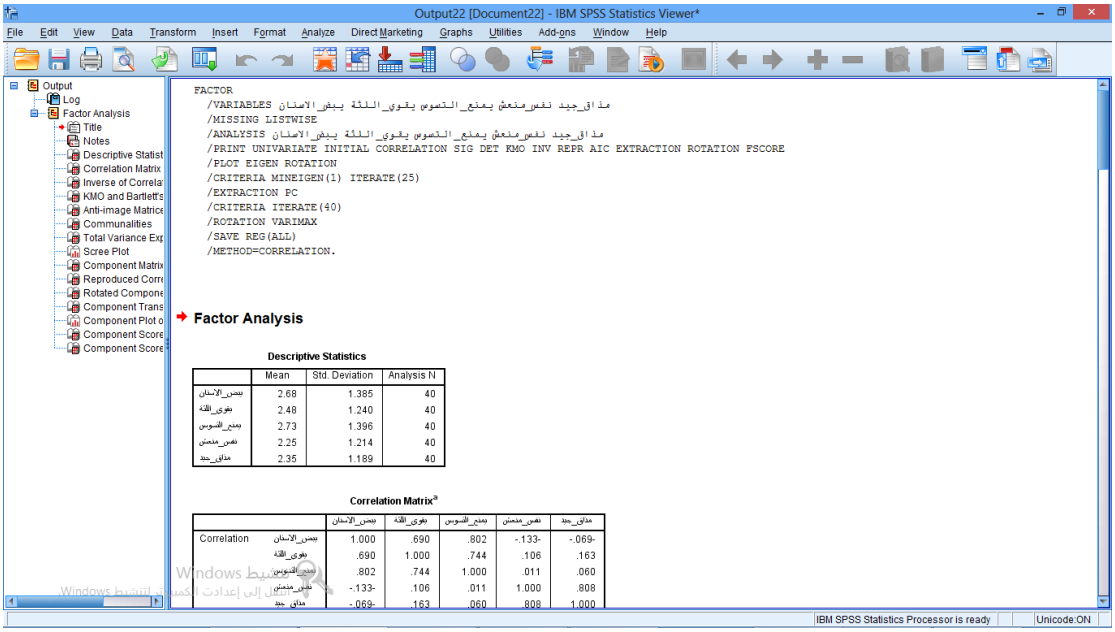
بالضغط على مربع Option يظهر مربع الحوار التالي:



من خلال هذا المربع الحواري تتم معالجة المتغيرات المفقودة إما عن طريق حذف المتغير الذي يحتويها، أو بحذف القيمة المفقودة فقط ويتم تعويضها بالمتوسط الحسابي للمتغير.

أيضا حذف الاحداثيات الضعيفة للمتغيرات من كل محور.

بعد أتمام كل هذه الخطوات نضغط على ok لتظهر النتائج في نافذة المخرجات كالتالي:



وتعـــرض نافـــذة المخرجـــات الجـــداول والبيانـــات والاختبـــارات الاحصـــائية التـــي تـــم اختيارها

4. مثال حسابي

لتكن المصفوفة X بحيث

$$X = \begin{pmatrix} 1 & 1 \\ 1 & 4 \\ 3 & 1 \end{pmatrix}$$

1. احسب المتوسط لكل متغير
2. احسب التباين لكل متغير σ
3. استنتج المصفوفة الممركز والمعممة
4. احسب مصفوفة الارتباط
5. احسب القيم الذاتية لمصفوفة الارتباط
6. احسب الأشعة الذاتية لمصفوفة الارتباط
7. قارن بين $\sum\lambda i$ و $tr(r)$

الحل:

.1

$$\bar{X}_1 = \frac{1}{n}\sum X_{1i} = \frac{1}{3}(1+1+3) = \frac{5}{3}$$

$$\bar{X}_2 = \frac{1}{n}\sum X_{2i} = \frac{1}{3}(1+4+1) = 2$$

.2

$$\sigma_{x_1}^2 = \frac{1}{n}\sum(X_{1i} - \bar{X}_1)^2$$

$$= \frac{1}{3}\left[(1 - \frac{5}{3})^2 + (1 - \frac{5}{3})^2 + (3 - \frac{5}{3})^2\right] = \frac{8}{9}$$

$$\sigma_{x_2}^2 = \frac{1}{n}\sum(X_{2i} - \bar{X}_2)^2$$

$$= \frac{1}{3}[(1 - 2)^2 + (4 - 2)^2 + (1 - 2)^2] = 2$$

.3

نحسب المصفوفة الممركزة والمعممة من خلال العلاقة التالية $Z = \frac{X_i - \bar{X}_i}{\sigma_i}$

$$Z = x_i - \bar{x}_i = \begin{pmatrix} 1 - \frac{5}{3} & 1 - 2 \\ 1 - \frac{5}{3} & 4 - 2 \\ 3 - \frac{5}{3} & 1 - 2 \end{pmatrix} = \begin{pmatrix} \frac{-2}{3} & -1 \\ \frac{-2}{3} & 2 \\ \frac{4}{3} & -1 \end{pmatrix}$$

$$Z = \frac{X_i - \bar{X}_i}{\sigma_i} = \begin{pmatrix} \frac{\frac{-2}{3}}{\frac{2}{3}\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{\frac{-2}{3}}{\frac{2}{3}\sqrt{2}} & \frac{2}{\sqrt{2}} \\ \frac{\frac{4}{3}}{\frac{2}{3}\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix}$$

$$Z = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 & -1 \\ -1 & 2 \\ 2 & -1 \end{pmatrix}$$

.4

نحسب مصفوفة الارتباط من خلال العلاقة التالية $\mathrm{r} = \frac{1}{n} Z'Z$

$$\mathrm{r} = \frac{1}{3} \times \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & -1 & 2 \\ -1 & 2 & -1 \end{pmatrix} \times \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & -1 \\ -1 & 2 \\ 2 & -1 \end{pmatrix}$$

$$\mathrm{r} = \begin{pmatrix} 1 & -0.5 \\ -0.5 & 1 \end{pmatrix}$$

.5

نقوم بحساب القيم الذاتية بتطبيق العلاقة التالية $|r - \lambda \mathrm{I}| = 0$

$$\left| \begin{pmatrix} 1 & -0.5 \\ -0.5 & 1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right| = 0$$

$$\left| \begin{pmatrix} 1 & -0.5 \\ -0.5 & 1 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \right| = 0$$

$$\left| \begin{pmatrix} 1-\lambda & -0.5 \\ -0.5 & 1-\lambda \end{pmatrix} \right| = 0$$

$$(1-\lambda)^2 \times \left(-\tfrac{1}{2}\right)^2 = 0$$

$$\left(1 - \lambda - \tfrac{1}{2}\right) \times \left(1 - \lambda + \tfrac{1}{2}\right) = 0$$

$$\begin{cases} \lambda_1 = \frac{2}{3} = 1.5 \\ \lambda_2 = \frac{1}{2} = 0.5 \end{cases}$$

.6

حساب الشعاع الذاتي الاول u_1

نحسب الشعاع u_1 من خلال العلاقة التالية:

$$(r - \lambda_1 I)u_1 = 0$$

$$\left(\begin{pmatrix} 1 & -0.5 \\ -0.5 & 1 \end{pmatrix} - 1.5\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)\begin{pmatrix} x \\ y \end{pmatrix} = 0$$

$$\begin{cases} -0.5x - 0.5y = 0 \\ -0.5x - 0.5y = 0 \end{cases}$$

حلول المتراجحة السابقة من الشكل y=-x

ومنه $u_1 = \begin{pmatrix} x \\ -x \end{pmatrix}$

$$u_1 = x\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

لدينا $u_1' u_1 = 1$

$$u_1' u_1 = x(1 \quad -1) \times x\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 1$$

بعد حل المعادلة نختار $x = \sqrt{\frac{1}{2}}$

اذا الشعاع الذاتي الأول هو $u_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$

حساب الشعاع الذاتي الثاني u_2

نحسب الشعاع u_2 من خلال العلاقة التالية:

$$(r - \lambda_2 \mathrm{I})\mathrm{u}_2 = 0$$

$$\left(\begin{pmatrix} 1 & -0.5 \\ -0.5 & 1 \end{pmatrix} - 0.5\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)\begin{pmatrix} \mathrm{x} \\ \mathrm{y} \end{pmatrix} = 0$$

$$\begin{cases} 0.5x - 0.5y = 0 \\ -0.5x + 0.5y = 0 \end{cases}$$

حلول المتراجحة السابقة من الشكل $y=x$

ومنه $\mathrm{u}_1 = \begin{pmatrix} x \\ x \end{pmatrix}$

$$\mathrm{u}_1 = x\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

لدينا $\mathrm{u}_1'\mathrm{u}_1 = 1$

$$\mathrm{u}_1'\mathrm{u}_1 = x(1 \quad 1) \times x\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1$$

بعد حل المعادلة نختار $x = \sqrt{\frac{1}{2}}$

اذا الشعاع الذاتي الثاني هو $\mathrm{u}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

.7

$$\mathrm{tr(r)} = \sum a_{ii} = 1+1 = 2$$

$$\sum \lambda_i = 1.5+0.5 = 2$$

نلاحظ أن $\mathrm{tr(r)} = \sum \lambda_i$

5. مثال تدريبي:

مصفوفة البيانات

$$\begin{pmatrix} 21 & 38 & 52 \\ 4 & 51 & 67 \\ 67 & 83 & 0 \\ 67 & 3 & 38 \\ 93 & 5 & 6 \end{pmatrix}$$

المصفوفة الممركز

$$\begin{pmatrix} -29.4 & 2 & 19.4 \\ -46.4 & 15 & 34.4 \\ 16.6 & 47 & 32.6 \\ 16.6 & -33 & 5.4 \\ 42.6 & -31 & -26.6 \end{pmatrix}$$

مصفوفة التباين والتباين المشترك

$$\begin{pmatrix} 1076.6 & -368.6 & -750.24 \\ -368.6 & 897.6 & -66.2 \\ -750.24 & -66.2 & 671.84 \end{pmatrix}$$

القيم الذاتية:

$\lambda_1 = 1732.20$

$\lambda_2 = 906.65$

$\lambda_3 = 7.28$

الاشعة الذاتية:

	u_1	u_2	u_3

x_1	0.79	-0.06	0.62
x_2	-0.30	-0.90	0.31
x_3	-0.54	0.43	0.73

احداثيات المشاهدات

	u_1	u_2	u_3
1	-34.15	8.16	-3.40
2	-59.54	3.84	1.03
3	16.25	-57.29	1.07
4	20.21	31.14	3.94
5	57.24	14.14	-2.65

إحداثيات المتغيرات

	u_1	u_2	u_3
x_1	-32.73	-1.66	1.66
x_2	12.68	-27.13	0.83
x_3	22.37	12.95	1.96

جودة تمثيل المشاهدات

	u_1	u_2	u_3
1	0.94	0.05	0.01
2	1.00	0.00	0.00

3	0.07	0.93	0.00
4	0.29	0.70	0.01
5	0.94	0.06	0.00

جودة تمثيل المتغيرات

	u_1	u_2	u_3
x_1	0.99	0.00	0.00
x_2	0.18	0.82	0.00
x_3	0.74	0.25	0.01

مساهمة المشاهدات

	u_1	u_2	u_3
1	0.13	0.01	0.32
2	0.41	0.00	0.03
3	0.03	0.72	0.03
4	0.05	0.21	0.43
5	0.38	0.04	0.19

مساهمة المتغيرات

	u_1	u_2	u_3

x_1	0.62	0.00	0.38
x_2	0.09	0.81	0.10
x_3	0.29	0.19	0.53

6. مثال تطبيقي للتحليل المركبات الرئيسية باستخدام SPSS

كلف مختبر صيدلاني بإجراء مسح لتحليل معايير اختيار معجون الأسنان. طُلب من 40 شخص تقييم أهمية الخصائص التالية المنسوبة إلى معجون الأسنان:

X1= الجنس

X2= العمر

X3=يعطي نفسا لطيفا

X4=يبيض الأسنان

X5=يقوي اللثة

X6=مذاق جيد

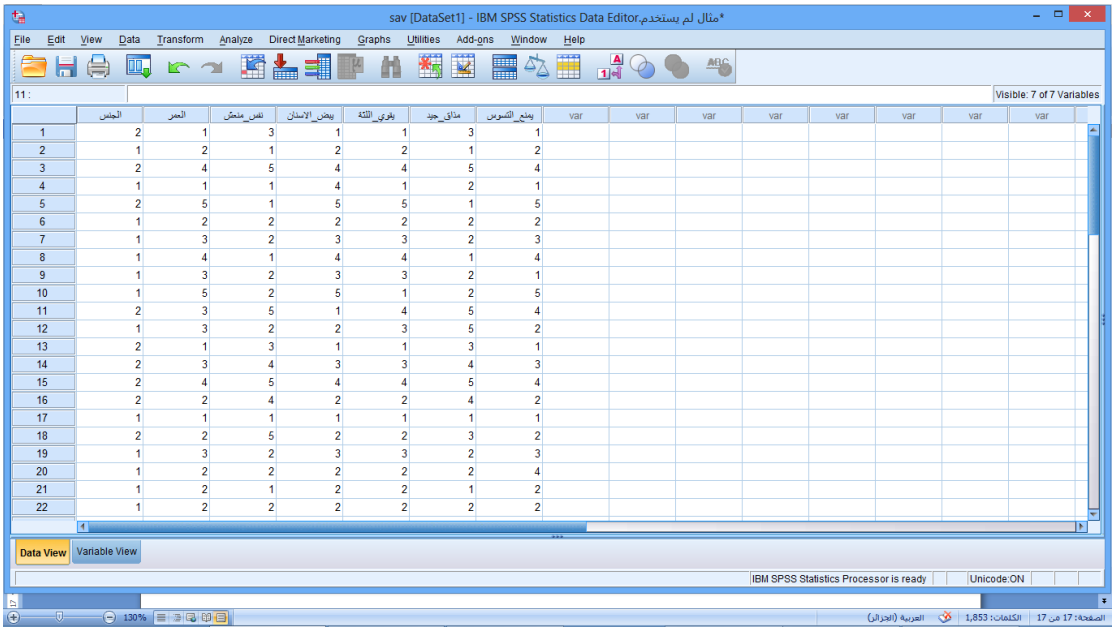
X7=يمنع التسوس

الجدول التالي يبين معطيات المبحوثين:

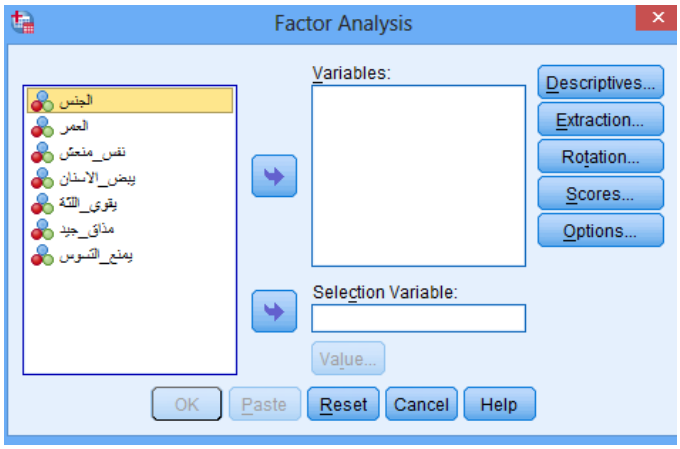
يمنع التسوس	مذاق جيد	يقوي اللثة	يبيض الأسنان	يعطي نفسا لطيفا	العمر	الجنس

2	1	3	1	1	3	1
1	2	1	2	2	1	2
2	4	5	4	4	5	4
1	1	2	1	4	1	1
2	5	1	5	5	1	5
1	2	2	2	2	2	2
1	3	2	3	3	2	3
1	4	1	4	4	1	4
1	3	2	3	3	2	1
1	5	2	5	1	2	5
2	3	5	1	4	5	4
1	3	2	2	3	5	2
2	1	3	1	1	3	1
2	3	4	3	3	4	3
2	4	5	4	4	5	4
2	2	4	2	2	4	2
1	1	1	1	1	1	1
2	2	5	2	2	3	2
1	3	2	3	3	2	3
1	2	2	2	2	2	4
1	2	1	2	2	1	2
1	2	2	2	2	2	2
1	1	2	1	1	1	3
2	2	3	2	2	3	2
1	2	5	2	5	2	5
1	2	1	2	2	1	2
1	2	2	2	2	2	2
1	3	2	5	3	2	5
1	2	1	2	2	1	2
1	3	2	3	3	2	3
1	4	2	4	4	2	4
1	5	1	5	5	1	5
1	4	2	4	4	3	4
2	1	2	1	1	2	1
2	3	1	3	2	3	3
2	2	2	2	2	2	2
1	3	2	1	3	2	1
2	3	5	2	3	5	2
2	1	3	1	1	3	1
1	1	2	1	1	1	1

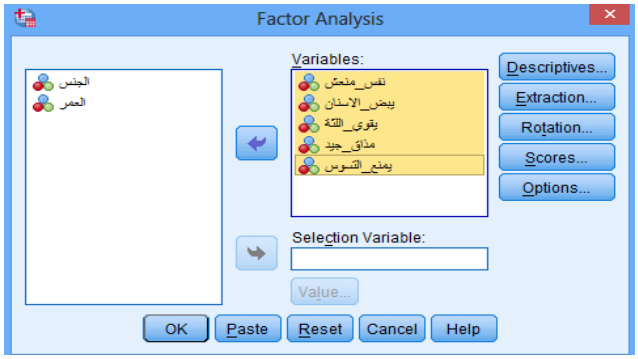
بعد عملية إدخال بيانات الجدول السابق في برنامج SPSS كما هو موضح في الشكل التالي:



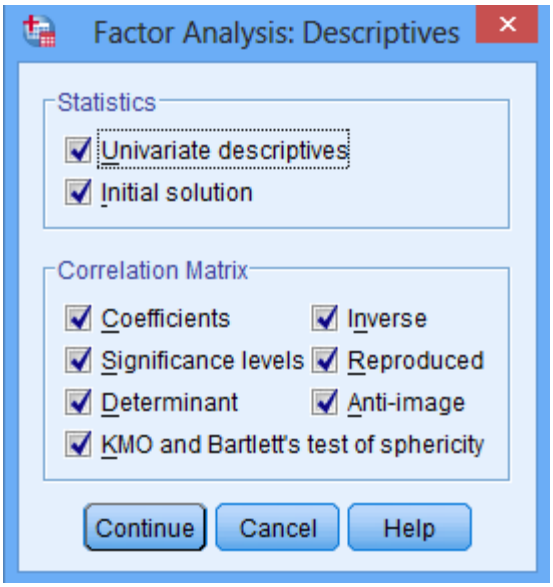
نقوم بعملية التحليل بطريقة المركبات الرئيسية بإتباع الخطوات التالية:



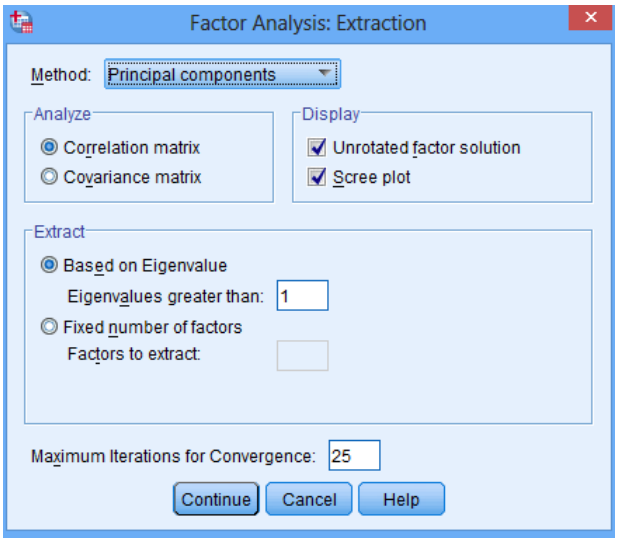
ننقل المتغيرات نفس منعش، بيض الأسنان، يقوي اللثة، مذاق جيد، يمنع التسوس إلى مربع المتغيرات



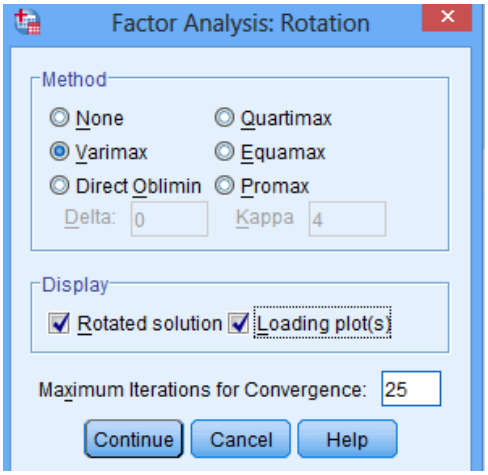
من Descriptves: نؤشر على الاختيارات التالية



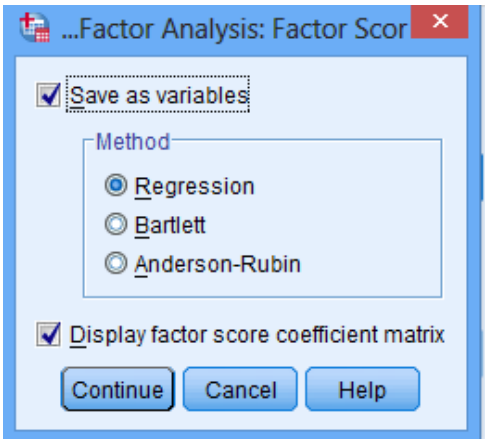
من Extraction نؤشر على الخيارات التالية



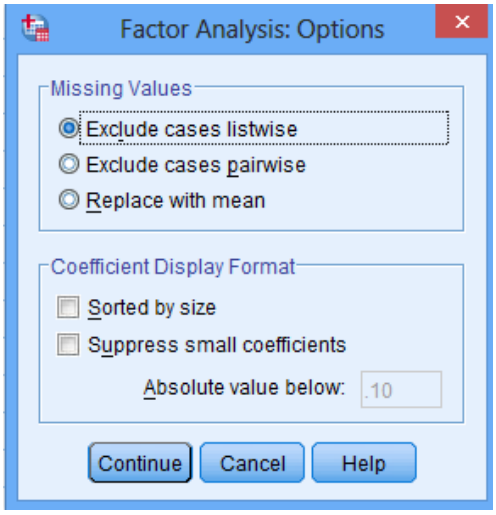
من Rotation نختار ما يلي



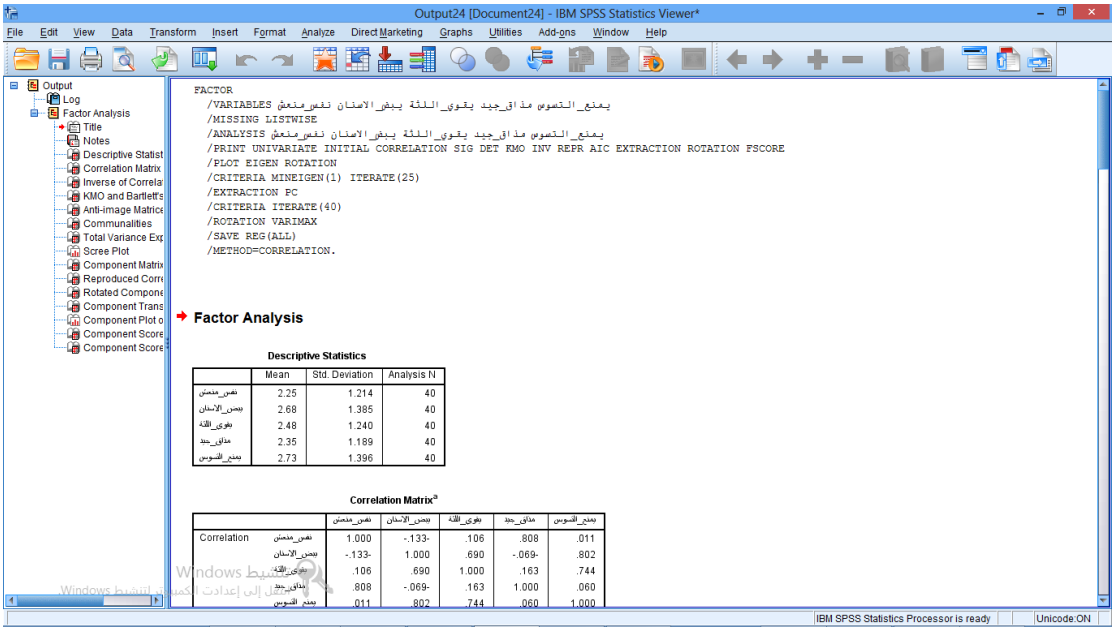
من Scores نختار التالي



من Option نختار الاوامر التالية



بعد إتمام هذه العمليات نضغط ok لنحصل على نافذة المخرجات التالية



والتي تحتوي على الجداول والبيانات التالية:

جدول الإحصاء الوصفي Statistiques descriptives

Descriptive Statistics

	Mean	Std. Deviation	Analysis N
نفس_منعش	2.25	1.214	40
يبيض_الاسنان	2.68	1.385	40
يقوي_اللثة	2.48	1.240	40
مذاق_جيد	2.35	1.189	40
يمنع_التسوس	2.73	1.396	40

يقدم وصف للمتغيرات المستعملة من خلال مؤشرات الاحصاء الوصفي والمتمثلة في المتوسط و الانحراف المعياري. ونلاحظ أن متوسط الافراد الذين يعتقدون أن معجون الاسنان هذا يعطي نفس منعش يساوي 2.25 بإنحراف معياري يساوي 1.214، وايضا الذين يرون أنه يمنع التسوس2.73 بتشتت 1.396.

جدول مصفوفة الإرتباط Matrice de corrélation

Correlation Matrix[a]

		نفس_منعش	بيض_الاسنان	يقوي_اللثة	مذاق_جيد	يمنع_التسوس
Correlation	نفس_منعش	1.000	-.133-	.106	.808	.011
	بيض_الاسنان	-.133-	1.000	.690	-.069-	.802
	يقوي_اللثة	.106	.690	1.000	.163	.744
	مذاق_جيد	.808	-.069-	.163	1.000	.060
	يمنع_التسوس	.011	.802	.744	.060	1.000
Sig. (1-tailed)	نفس_منعش		.206	.257	.000	.472
	بيض_الاسنان	.206		.000	.335	.000
	يقوي_اللثة	.257	.000		.158	.000
	مذاق_جيد	.000	.335	.158		.358
	يمنع_التسوس	.472	.000	.000	.358	

a. Determinant = .046

نـــذكر هنـــا أن مصـــفوفة الارتبـــاط هـــي مصـــفوفة متنـــاظرة وبالتـــالي يمكـــن قـــراءة قيم الارتباط بين المتغيرات إما من المثلث العلوي أو السفلي.

كمـــا أن الجـــدول يعطـــي المعنويـــة لكـــل قيمـــة ارتبـــاط مـــن خـــلال الجـــزء (1- Sig. tailed).

مـــن مصـــفوفة الارتبـــاط نلاحـــظ أن عـــدد قـــيم الارتبـــاط الاكبـــر مـــن 0.5 يســـاوي 5 وهـــو متســـاوي مـــع عـــدد قـــيم الارتبـــاط الاقـــل مـــن 0.5 وبالتـــالي يمكـــن متابعـــة التحليل.

ونلاحـــظ أن قيمـــة الارتبـــاط بـــين (نفـــس مـــنعش) والمتغيـــر (مـــذاق جيـــد) يســـاوي 0.808 وهـــو ارتبـــاط قـــوي حيـــث يجتمـــع أفـــراد العينـــة علـــى أن الـــنفس المـــنعش للمعجـــون لابـــد أن يرافقـــه المـــذاق الجيـــد، كمـــا نلاحـــظ ان قيمـــة Sig.(1-tailed) تســـاوي 0.00 وهـــي أكبـــر مـــن 0.05 مســـتوى المعنويـــة المفتـــرض، وبالتـــالي فإننـــا نقبـــل

الفرضية الصفرية ونرفض الفرضية البديلة،اي أن الارتباط بين المتغير (نفس منعش) والمتغير (مذاق جيد) معنوي.

ونلاحظ أن قيمة الارتباط بين (يبض الاسنان) والمتغير (يقوي اللثة) يساوي 0.690 وهو ارتباط جيد حيث يجتمع أفراد العينة على أن عمل المعجون في تبيض الاسنان من شأنه تقوية اللثة، كما نلاحظ ان قيمة (Sig.(1-tailed تساوي 0.00 وهي أكبر من 0.05 مستوى المعنوية المفترض، وبالتالي فإننا نقبل الفرضية الصفرية ونرفض الفرضية البديلة،اي أن الارتباط بين المتغير (يبض الاسنان) والمتغير (يقوي اللثة) معنوي.

ونلاحظ أن قيمة الارتباط بين المتغير(يمنع التسوس) والمتغير (يبض الاسنان) والمتغير (يقوي اللثة) يساوي 0.802 و0.744 على التوالي وهي ارتباطات جيدة حيث يجتمع أفراد العينة على أن عمل المعجون في القضاء في منع التسوس يساهم في تبيض الاسنان و تقوية اللثة، كما نلاحظ ان قيمة (Sig.(1-tailed تساوي 0.00 وهي أكبر من 0.05 مستوى المعنوية المفترض، وبالتالي فإننا نقبل الفرضية الصفرية ونرفض الفرضية البديلة،اي أن الارتباط بين المتغير (يمنع التسوس) والمتغير (يبيض الاسنان) والمتغير (يقوي اللثة) معنوي.

كما يشير محدد المصفوفة الذي يساوي 0.046 وهوأكبر من 0.00001 إلى عدم وجود مشكل الارتباط الذاتي وجود الحد الادنى من البيانات أي أن المصفوفة ليست مصفوفة فارغة.

Inverse of Correlation Matrix

	منعش_نفس	الاسنان_يبض	اللثة_يقوي	جيد_مذاق	التسوس_يمنع
منعش_نفس	2.960	.450	-.125-	-2.331-	-.162-
الاسنان_يبض	.450	3.222	-.799-	.109	-2.000-
اللثة_يقوي	-.125-	-.799-	2.511	-.291-	-1.209-
جيد_مذاق	-2.331-	.109	-.291-	2.940	-.019-
التسوس_يمنع	-.162-	-2.000-	-1.209-	-.019-	3.506

جدول مؤشر KMO واختبار بارتلت Bartlett's Test

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.654
Bartlett's Test of Sphericity	Approx. Chi-Square	112.262
	df	10
	Sig.	.000

من خلال قياس مؤشر KMO الذي على أساسه نستدل على مدى كفاية عدد أفراد العينة ويجب أن تكون قيمته أكبر من 0.50 حتى تكون العينة كافية وهذا شرط أساسي يجب تحقيقه،

ومن الملاحظ أن مؤشر KMO يساوي0.654 وبتالى فإن حجم العينة كافي لإجراء الدراسة.

اختبار Bartlett لدائرية Sphericity فهو مؤشر للعلاقة بين المتغيرات ويجب أن يكون مستوى الدلالة لهذه العلاقة أقل من 0.05 وذلك حتى نستطيع التأكيد على أن هذه العلاقة دالة إحصائيا. ويظهر من الجدول أن قيمة إختبار Bartlett تساوي 112.262 وهي دالة احصائيا حيث أن .Sig أقل من 0.05 ومنه نرفض الفرضية الصفرية و نقبل الفرضية البديلة أي أنه يوجد معاملات ارتباط غير معدومة.

Anti-image Matrices

		منعش_نفس	الاسنان_يبض	اللثة_يقوي	جيد_مذاق	التسوس_يمنع
Anti-image Covariance	منعش_نفس	.338	.047	-.017-	-.268-	-.016-
	الاسنان_يبض	.047	.310	-.099-	.011	-.177-
	اللثة_يقوي	-.017-	-.099-	.398	-.039-	-.137-
	جيد_مذاق	-.268-	.011	-.039-	.340	-.002-
	التسوس_يمنع	-.016-	-.177-	-.137-	-.002-	.285
Anti-image Correlation	منعش_نفس	.512[a]	.146	-.046-	-.790-	-.050-
	الاسنان_يبض	.146	.715[a]	-.281-	.035	-.595-
	اللثة_يقوي	-.046-	-.281-	.805[a]	-.107-	-.408-
	جيد_مذاق	-.790-	.035	-.107-	.519[a]	-.006-
	التسوس_يمنع	-.050-	-.595-	-.408-	-.006-	.697[a]

a. Measures of Sampling Adequacy(MSA)

جدول جودة التمثيل Communalities

Communalities

	Initial	Extraction
نفس_منعش	1.000	.903
يبض_الاسنان	1.000	.856
يقوي_اللثة	1.000	.807
مذاق_جيد	1.000	.901
يمنع_التسوس	1.000	.871

Extraction Method: Principal Component Analysis.

من جدول نوعية التمثيل فنلاحظ أن تمثيل كل المتغيرات أكبر من 0.4 وبالتالي كل المتغيرات تدخل في الدراسة و لا يتم استبعاد أي متغيرة.

كما نلاحظ أن كل المتغيرات ذات تمثيل عالي حيث كانت نسبة الاستخراج الاصغر 0.807 أي 80.70% للمتغير (يقوي اللثة).

جدول التباين الكلي المشروح Total Variance Explained

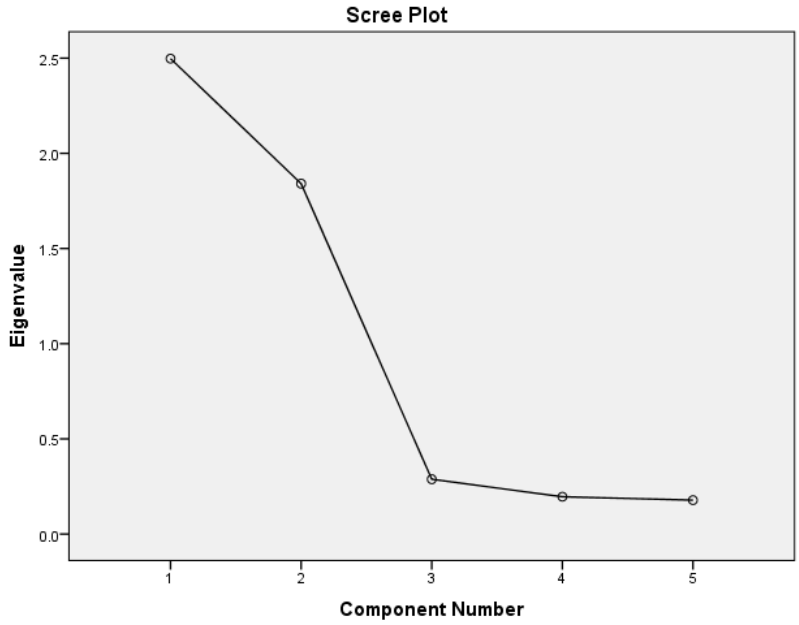
Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.498	49.950	49.950	2.498	49.950	49.950	2.493	49.863	49.863
2	1.841	36.813	86.763	1.841	36.813	86.763	1.845	36.900	86.763
3	.288	5.759	92.522						
4	.196	3.919	96.441						
5	.178	3.559	100.000						

Extraction Method: Principal Component Analysis.

نلاحظ من الجدول أنه توجد قيمتين ذاتيتين أكبر من الواحد، وعليه سيتم اعتماد هاتين القيمتين وذلك حسب محك كايز وإقصاء باقي القيم الذاتية، وبالتالي يتم اختزال البيانات في محورين حيث يختزل الاول نسبة 49.86% من البيانات ويختزل المحور الثاني 36.90% من البيانات ومعا يتحدان في اختزال ما نسبته 86.76% من مجمل البيانات وهي نسبة جد مقبولة حيث 13.34% فقط من البيانات لم نتمكن من تمثيلها.

ويمكن من الشكل الموالي إلقاء نظرة على انحدار القيم الذاتية



حيث نلاحظ أن بعد القيمتين الاولى والثانية ينحدر خط القيم الذاتية بشكل جد معنوي ثم يأخذ الشكل الافقي مع القيم الثلاثة المتبقية، وهو يدعم اعتمادنا في عملية التحليل على القيمتين الاول والثانية.

احداثيات المتغيرات قبل وبعد التدوير:

مصفوفة الاحداثيات قبل التدوير:

Component Matrix[a]

	Component	
	1	2
نفس_منعش		.949
يبض_الاسنان	.900	
يقوي_اللثة	.895	
مذاق_جيد		.941
يمنع_التسوس	.932	

Extraction Method: Principal Component Analysis.

a. 2 components extracted.

المحـــورالاول: نلاحـــظ أن أن المحـــور الاول يحتـــوي علـــى المتغيـــرات يبـــيض الاســـنان والمتغيـــر يقـــوي اللثـــة والمتغيـــر يمنـــع التســـوس، وهـــي متغيـــرات تعكـــس الجانـــب الصحي لمعجون الاسنان وبالتالي يمكن أن نسمي هذا المحور بـ عوامل صحية.

المحـــورالثـــاني: نلاحـــظ أن أن المحـــور الثـــاني يحتـــوي علـــى المتغيـــرات نفـــس مـــنعش، ومـــذاق جيـــد وهـــذه المتغيـــرات تعكـــس الجوانـــب الحســـية وبالتـــالي يمكـــن أن نســـمي الحور بعوامل حسية

Reproduced Correlations

		منعش_نفس	الاسنان_يبض	اللثة_يقوي	جيد_مذاق	التسوس_يمنع
Reproduced Correlation	منعش_نفس	.903[a]	-.151-	.127	.900	.003
	الاسنان_يبض	-.151-	.856[a]	.788	-.090-	.850
	اللثة_يقوي	.127	.788	.807[a]	.186	.830
	جيد_مذاق	.900	-.090-	.186	.901[a]	.065
	التسوس_يمنع	.003	.850	.830	.065	.871[a]
Residual[b]	منعش_نفس		.018	-.020-	-.092-	.009
	الاسنان_يبض	.018		-.099-	.021	-.048-
	اللثة_يقوي	-.020-	-.099-		-.023-	-.085-
	جيد_مذاق	-.092-	.021	-.023-		-.005-
	التسوس_يمنع	.009	-.048-	-.085-	-.005-	

Extraction Method: Principal Component Analysis.

a. Reproduced communalities

b. Residuals are computed between observed and reproduced correlations. There are 3 (30.0%) nonredundant residuals with absolute values greater than 0.05.

مصفوفة الاحداثيات بعد التدوير:

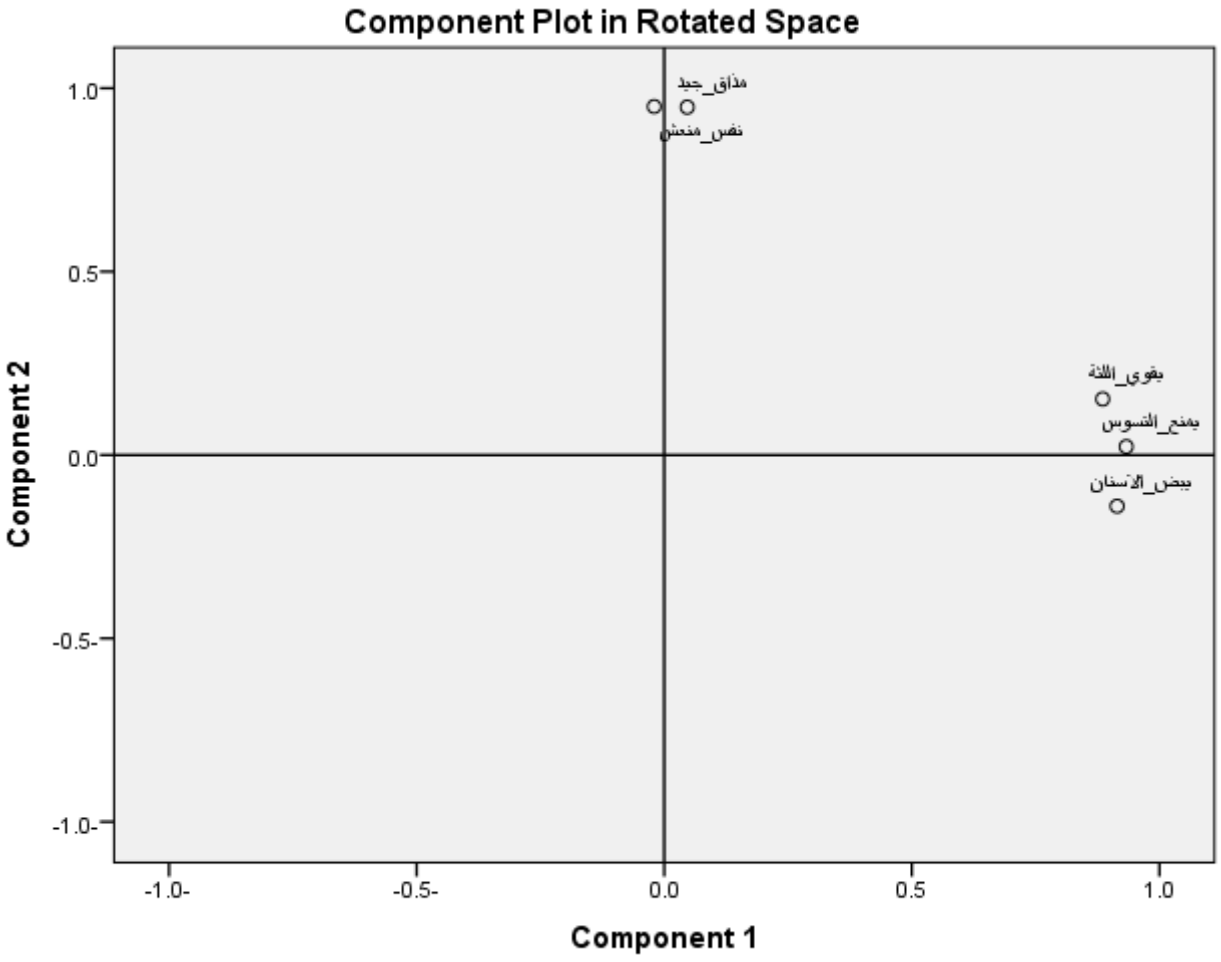
Rotated Component Matrix[a]

	Component	
	1	2
منعش_نفس		.950
الاسنان_يبض	.914	
اللثة_يقوي	.885	
جيد_مذاق		.948
التسوس_يمنع	.933	

Extraction Method: Principal
Component Analysis.
Rotation Method: Varimax with
Kaiser Normalization.
a. Rotation converged in 3 iterations.

تسمح عملية التدوير بإختيار الموقع الانسب لتمثيل المتغيرات، ونلاحظ أنه بعد عملية التدوير تعزز تمثيل المحاور ولو بشكل طفيف

شكل المكونات بعد التدوير



نلاحظ من الشكل السابق أن المتغيرات تنقسم إلى مجوعتين متمايزتين الاولى على المحور الاول والتي تظم ثلاثة متغيرات والمجموعة الثانية ممثلة على المحور الثاني.

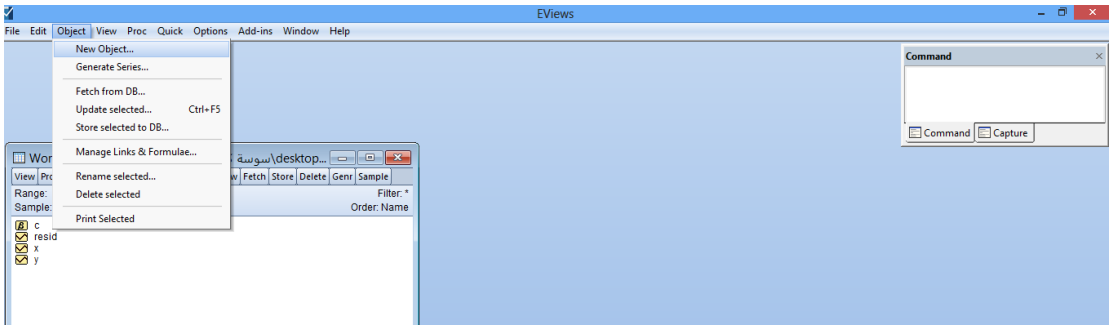
7. مثال تطبيقي للتحليل المركبات الرئيسية باستخدام Eviews

لتوضيح كيفية تطبيق المركبات الرئيسية بإستخدام Eviews ، نستخدم المصفوفة التالية كمثال:

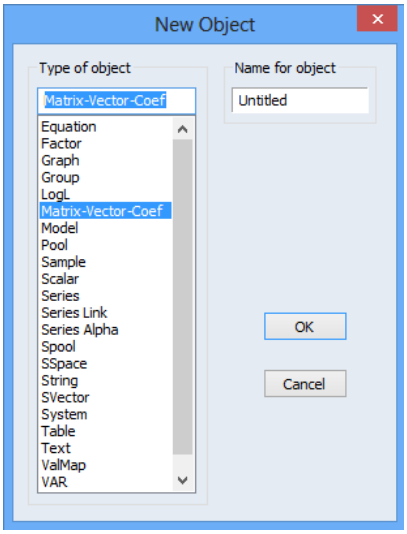
$$A = \begin{pmatrix} 1 & 1 & 3 \\ 1 & 2 & 3 \\ 2 & 5 & 4 \\ 1 & 3 & 1 \\ 3 & 7 & 1 \end{pmatrix}$$

نقوم بإنشاء المصفوفة في برنامج Eviews

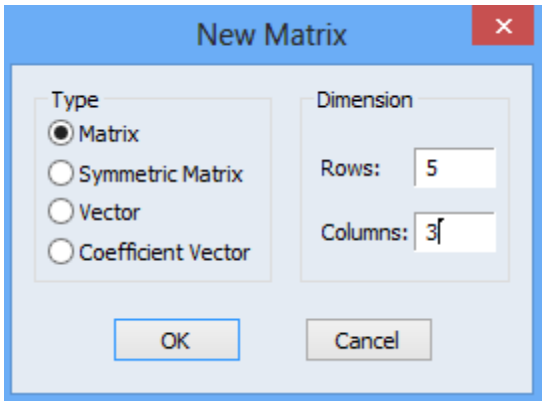
من قائمة objet نختار New objet كما في الشكل الموالي:



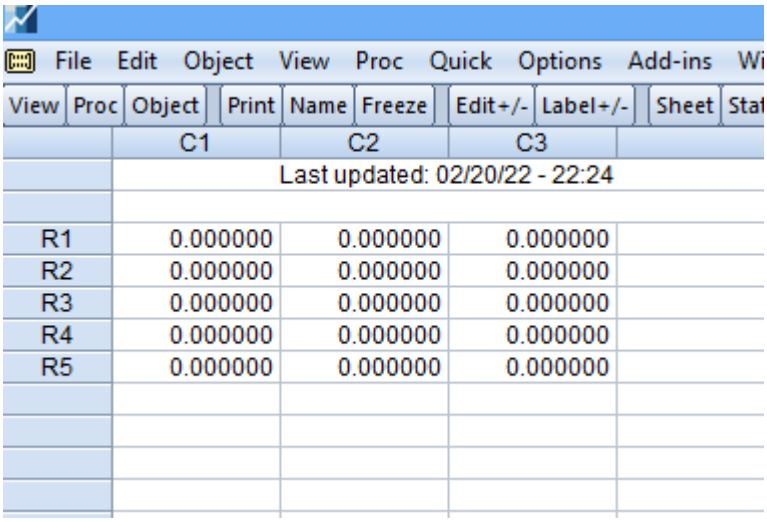
يظهر المربع الحوار التالي نختار منه Matrix vector coefficient



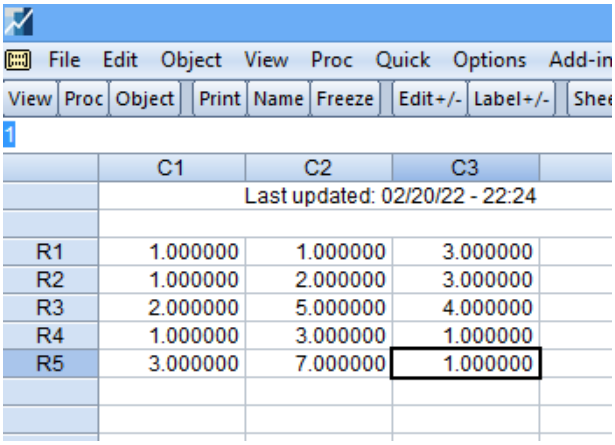
ثـــم نقـــوم بتحديـــد ابعـــاد المصـــفوفة عـــن طريـــق ادخـــال عـــدد الاســـطر Rows والاعمدة Columns



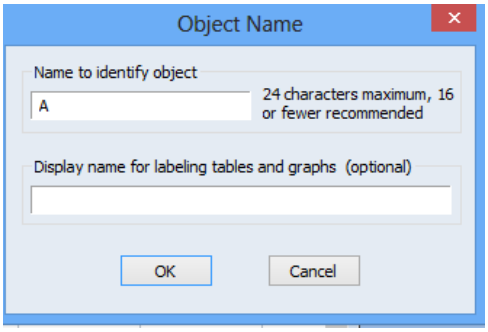
بالضغط على Ok تظهر النافذة التالية:



نقوم بإدخال قيم المصفوفة

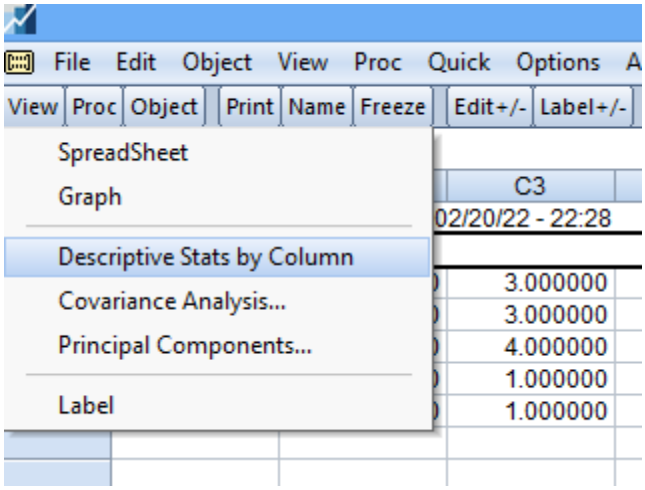


ثم من Name نعطي تسمية للمصفوفة مثال نسميها A كما هو موضح في الشكل التالي

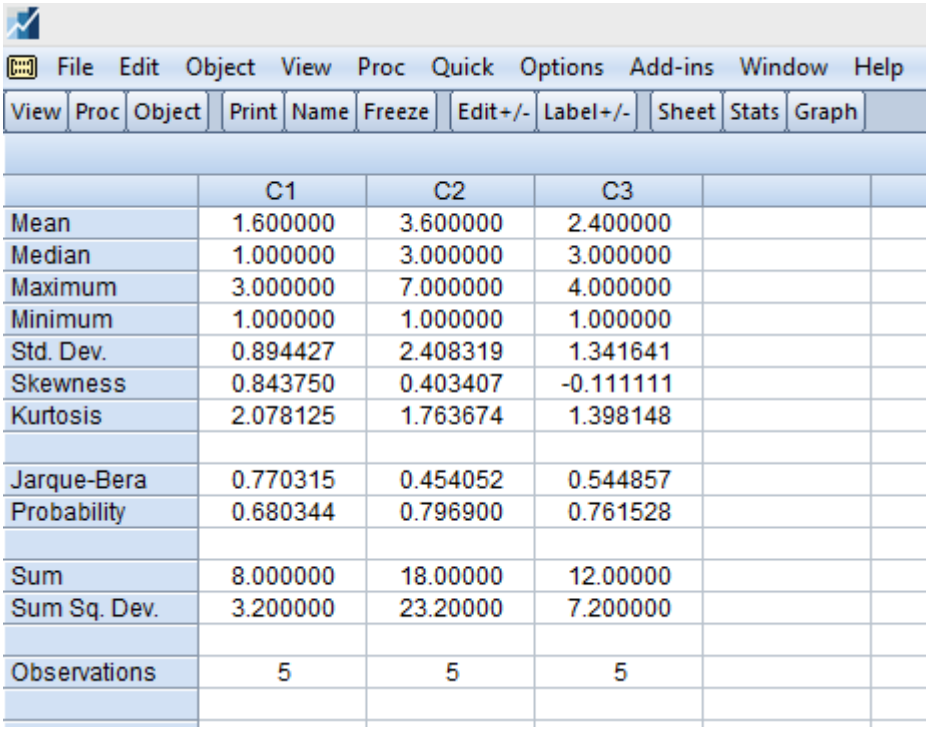


1.7- الاحصاءات الوصفية لمصفوفة المعطيات

للحصول على الاحصاءات الوصفية للمتغيرات المصفوفة من View نختار Descriptive Stats by Column كما في الشكل الموالي:



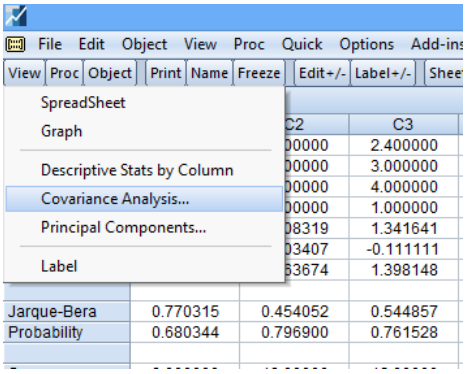
تظهر النتائج التالية



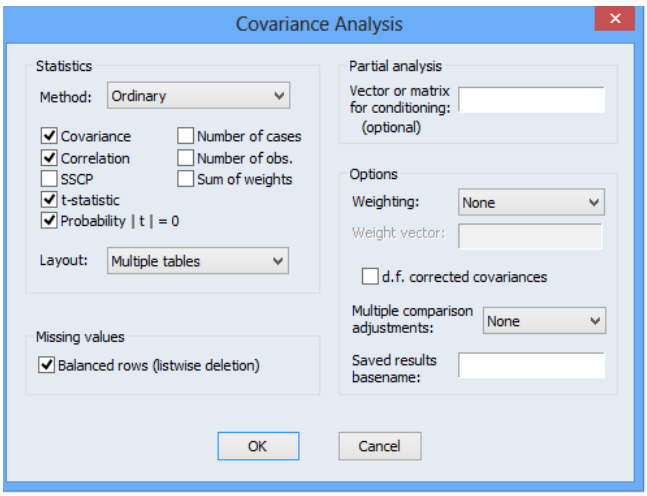
	C1	C2	C3
Mean	1.600000	3.600000	2.400000
Median	1.000000	3.000000	3.000000
Maximum	3.000000	7.000000	4.000000
Minimum	1.000000	1.000000	1.000000
Std. Dev.	0.894427	2.408319	1.341641
Skewness	0.843750	0.403407	-0.111111
Kurtosis	2.078125	1.763674	1.398148
Jarque-Bera	0.770315	0.454052	0.544857
Probability	0.680344	0.796900	0.761528
Sum	8.000000	18.00000	12.00000
Sum Sq. Dev.	3.200000	23.20000	7.200000
Observations	5	5	5

حيث يقدم برنامج Eviews الاحصائيات الوصفية المطلوبة للمتغيرات الثلاثة

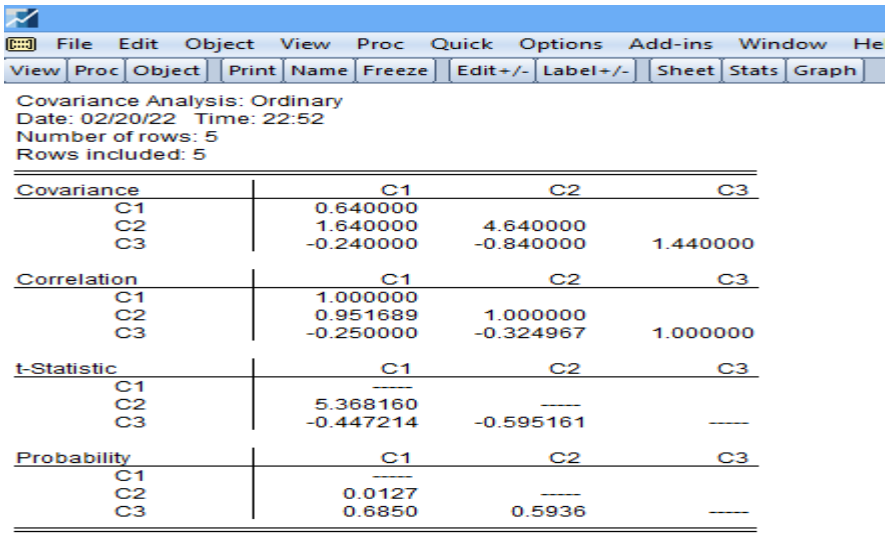
كما نحصل على تحليل التغاير والارتباط من خلال View ثم Covariance Analysis كما هو موضح في الشكل التالي:



يظهر المربع الحواري التالي



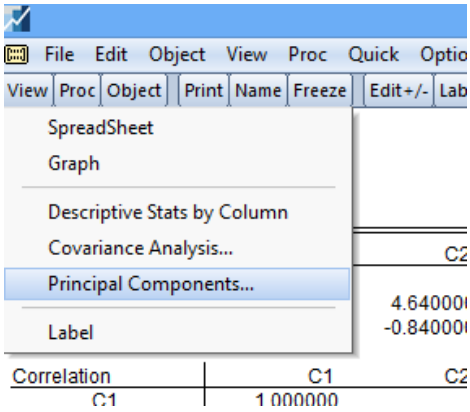
من الربع الحواري السابق نؤشر على الخيارات المطلوبة كما هو موضح ثم نضغط على ok



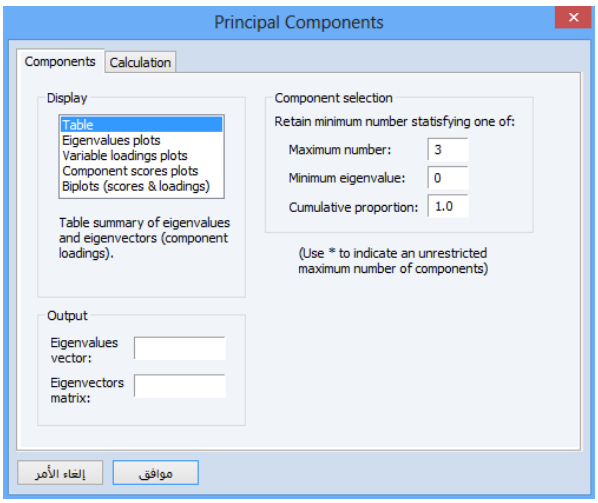
نستنتج من اجدول السابق ان المتغيرين الاول والثاني مرتبطان بشكل قوي حيث أن معمل الارتباط بينهما يساوي 0.95 وهو قوي وموجب، وهو ما تؤكده القيمة الاحصائية لستيودنت والقيمة الاحتمالية. وهو ما يدل على أن المتغيرين الاول والثاني يتشاركان في نفس المحور عاملي.

2.7- استخراج المحاور العاملية وتفسيرها

وللحصول على المحاور العاملية، من قائمة View نختار Principal Components كما هو في الشكل الموالي:



من خلال مربع الحوار التالي يمكن اختيار عدد المحاور العاملية، كما يمكن تحديد العديد من الخيارات



بالضغط على موافق نحصل على النتائج التالية

EVie

File Edit Object View Proc Quick Options Add-ins Window Help

View | Proc | Object | Print | Name | Freeze | Edit+/- | Label+/- | Sheet | Stats | Graph

Principal Components Analysis
A
Date: 02/20/22 Time: 23:07
Number of rows: 5
Rows included: 5
Computed using: Ordinary correlations
Extracting 3 of 3 possible components

Eigenvalues: (Sum = 3, Average = 1)

Number	Value	Difference	Proportion	Cumulative Value	Cumulative Proportion
1	2.101885	1.248845	0.7006	2.101885	0.7006
2	0.853040	0.807965	0.2843	2.954925	0.9850
3	0.045075	---	0.0150	3.000000	1.0000

Eigenvectors (loadings):

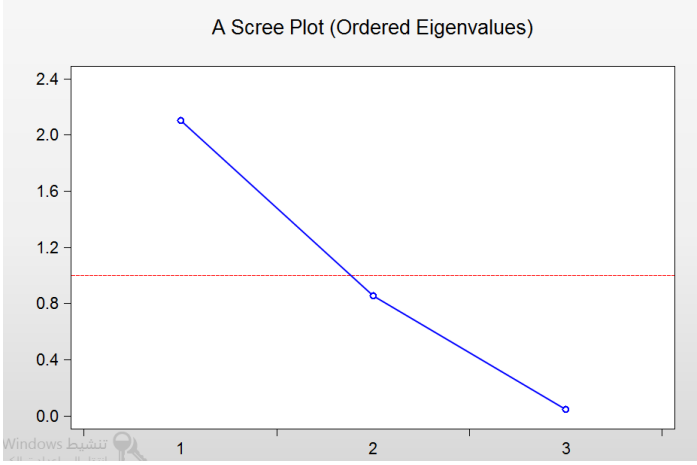
Variable	PC 1	PC 2	PC 3
C1	0.656940	0.288540	-0.696545
C2	0.669590	0.201338	0.714921
C3	-0.346524	0.936060	0.060936

Ordinary correlations:

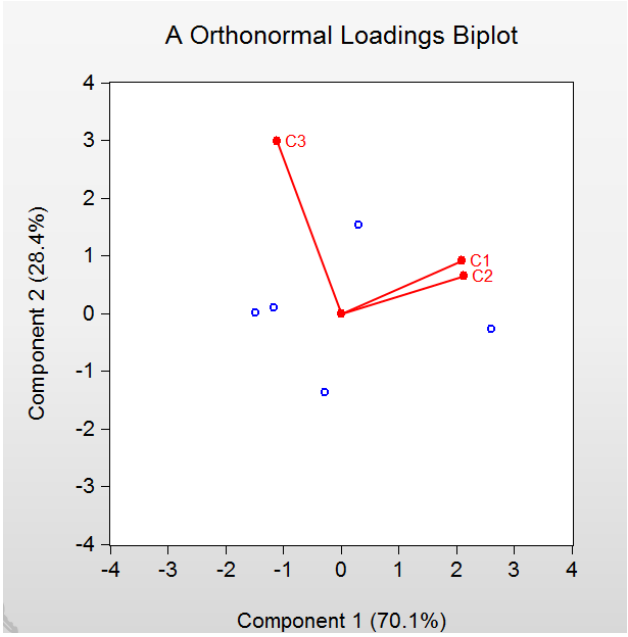
	C1	C2	C3
C1	1.000000		
C2	0.951689	1.000000	
C3	-0.250000	-0.324967	1.000000

من النتائج السابقة نلاحظ أنه يوجد محور عاملي واحد تفوق قيمته الواحد(2.10) ، ويشرح وحده 70% من التباين الكلي المشروح، كما أن المحورين العامليين الاول والثاني يشرحان معا 98.50% من التباين الكلي وهي نسبة جد مقبولة للاعتماد على هذين المحورين فقط.

والشكل الموالي يوضح التمثيل البياني للقيم الذاتية



كما يمكن تمثيل المتغيرات والمشاهدات على المحاور كالتالي



ونلاحظ أن المتغيرين الاول والثاني يتشاركان التمثيل في المحور العاملي الأول، بينما يتم تمثيل المتغير الثالث على المحور العاملي الثاني.

المحور الرابع:

التحليل العاملي المتناظر AFC

يعد التحليل العاملي للمشاهدات من مجموعة التحليل العاملي التي طورت للتعامل مع معطيات تختلف عن طريقة المكونات الرئيسية، غير أن الاختلاف الجوهري يكمن في نوع تغيرات لا في الطريقة، وبالتالي فإنهما لا يختلفان من حيث الهدف في تقديم اختزال للمعلومات في فضاء متعدد الأبعاد بحيث تكون القراءة واضحة ولا تختلف مع المعطيات الاولية عن طريق الحفاظ على أكبر قدر من المعطيات الخام .

يعود أصول الطريقة إلى benzcri paul jean حيث قام في سنة 1961 بدراسة العلاقة بين متغيرين كيفيين باستعمال طريقة المركبات الرئيسية ليتم تطويرها و تعديلها لاحقا من قبله في 1990 عموما تستعمل طريقة أو تقنية AFC في الجداول المزدوجة (الاقتران أو التوافق) ذات المتغيرات الكيفية حيث تسمح بتمثيل نقاط الأسطر و الأعمدة لجدول فئات أو صفات (modalité) المتغيرين انوعيين.

من الواضح أن جدول البيانات أكثر قابلية للإدارة ، من الناحية العملية ، من الجداول المنفصلة الكاملة ، خاصة إذا كان عدد الأفراد كبيرًا.

من الضروري أن نلاحظ الآن أن هناك اختلافين رئيسيين بين ACP و AFC.

- المقياس المستخدم في AFC لتحديد القرب بين صفين أو عمودين هو مقياس Chi-square ، بينما تُستخدم المسافة الإقليدية في PCA.

- يسمح AFC بتمثيل متداخل لصفوف وأعمدة الجدول، بدلاً من تقديم رسمين بيانيين مستقلين ، أحدهما للمتغيرات والآخر للأفراد.

1. مجالات التطبيق

تم استخدام هذا التحليل عمليًا لأنه مصمم لجداول التقابلية وبالتالي يسمح بدراسة الروابط الموجودة بين متغيرين اسميين. وبالتالي فإن مجالات تطبيق AFC تختلف عن تلك الخاصة بـ PCA ، وهي مناسبة لجداول القياسات المتجانسة أو غير المتجانسة.

بالنسبة لهذا التحليل ، يمكننا أيضًا تقديم قائمة طويلة من التخصصات التي وجدت إجابة لمشكلتها من خلال AFC. وهكذا ، فإن علم البيئة، وعلم النفس ، والاقتصاد ، وغير ذلك من العلوم التي قد يكون من المثير للاهتمام فيها بدراسة الروابط بين متغيرين اسميين ، قد قدمت قدرًا كبيرًا من البيانات. المصمم للجداول التقابلية (أي الترددات) ، يمكن تطبيقه على جداول المقاييس المتجانسة (أي نفس نظام الوحدات) ، وجداول الدرجات ، والرتب ، والتفضيلات ، والجداول ذات القيم المنطقية (0 أو 1) ، وايظا على الجداول المأخوذة من استبيانات .

2. البيانات المستخدمة

على عكس ACP فإنه عند استعمال AFC يجب تنظيم البيانات، في جداول تقابلية (تسمى أيضًا جداول التبعية أو الجداول المتقاطعة).

الجدول التقابلية هي جدول للأرقام تم الحصول عليها من خلال تقاطع متغيرين نوعيين محددين على نفس المجموعة المكونة من n من الأفراد.

يمكن أيضًا توسيع CFA ليشمل المتغيرات الكمية المتجانسة عن طريق تحديد بعض الخصائص لهذه المتغيرات. بالامتداد ، و ينطبق أيضًا على جداول المتغيرات الفردية للمتغيرات الكمية المتجانسة ، وفي هذه الحالة يتم اعتبار الأفراد كمتغيرات.

وبالتالي فإن البيانات الأولية تكون منظمة بالطريقة التي سبق شرحها كما هو موضح في الجدول التالي:

	فئات المتغير الثاني
فئات المتغير الأول	$1 \quad \ldots\ldots \quad j \quad \ldots\ldots \quad J$ 1 $\vdots$ $i \quad \ldots\ldots \quad k_{ij} \quad \ldots\ldots$ $\vdots$ I

- يمثل i عدد الصفوف وا مجموعة الصفوف
- يمثل j عدد الأعمدة و J مجموعة الأعمدة
- kij هو عدد الأفراد الذين يمتلكون كل من الصفة i للمتغير الأول والصفة j للمتغير الثاني.

وبالتالي لدينا:

$$\sum_{i \in I} \sum_{j \in J} k_{ij} = n$$

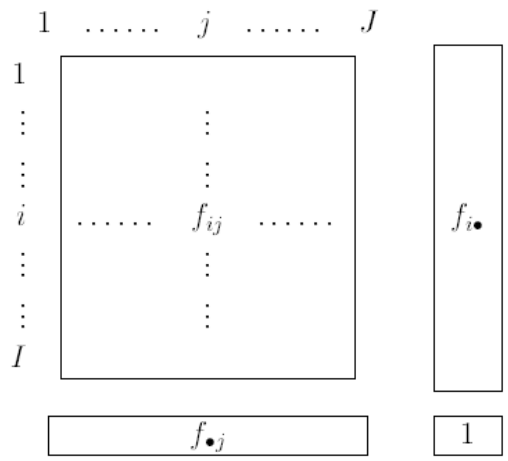
مـــع n العـــدد الإجمـــالي للأفـــراد فـــي الجـــدولالأولي. نلاحـــظ أن الصـــفوف والأعمـــدة فـــي هـــذا النـــوع مـــن الجـــداول تلعـــب دورًا متمـــاثلًا. يمكـــن تعريـــف جـــدول التـــرددات النسبية الذي يؤخذ في الاعتبار ترددات fij تعطى بواسطة:

$$f_{ij} = \frac{k_{ij}}{n}$$

والهوامش بالعلاقة التالية:

$$f_{i\bullet} = \sum_{j \in J} f_{ij}$$

كما هو موضح في جدول التكرار النسبي



بحيث:

$$f_{\bullet j} = \sum_{i \in I} f_{ij}$$

و

$$\sum_{i \in I} f_{i\bullet} = \sum_{j \in J} f_{\bullet j} = \sum_{i \in I} \sum_{j \in J} f_{ij} = 1.$$

3. أهدف التحليل العاملي المتناظر AFC

أهداف AFC هي نفس أهداف ACP بمعنى أن AFC يسعى إلى الحصول على تصنيف الصفوف وتصنيف الأعمدة ، ثم ربط هذين النموذجين.

لذلك من الضروري إجراء تقييم لأوجه التشابه بين السطور (على التوالي الأعمدة) الإجابة على أسئلة من كالتالي:

- ما هي الصفوف المتشابهة (الأعمدة)؟

- أيهما مختلف؟

- هل هناك مجموعات متجانسة من الصفوف (أعمدة على التوالي)؟

- هل من الممكن إبراز تصنيف السطور (على التوالي أعمدة)؟

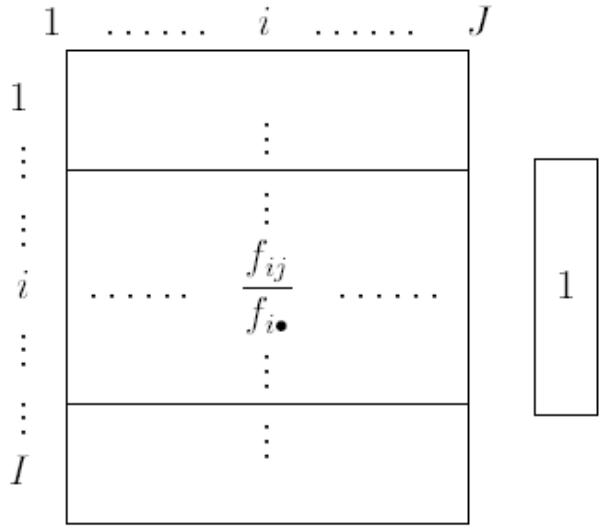
ومع ذلك ، فإن فكرة التشابه بين صفين أو عمودين تختلف عن ACP. في الواقع يكون صفان (على التوالي عمودين) قريبين إذا كانا مرتبطين بنفس الطريقة مع جميع الأعمدة.

لذلك من الضروري البحث عن الصفوف (على التوالي الأعمدة) التي يكون توزيعها أكثر انحرافًا عن البقية ، وتلك التي تتشابه مع بعضها البعض وتلك التي تتعارض . من أجل ربط تصنيف الصفوف بمجموعة الأعمدة ، تتميز كل مجموعة من الصفوف بالأعمدة التي ترتبط بها هذه المجموعة.

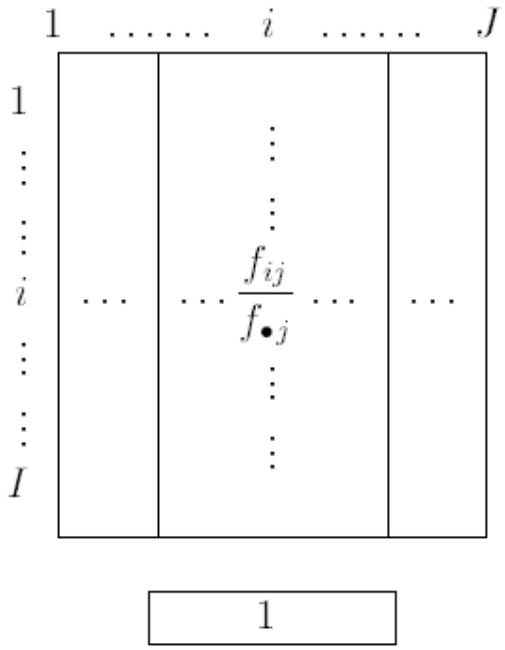
ومن خلال التناظر ، تتميز كل مجموعة من الأعمدة بالصفوف التي ترتبط بها هذه المجموعة. وبالتالي يمكننا تفكيك الرابط بين متغيرين إلى مجموع ميول بسيطة وقابلة للتفسير وقياس أهمية كل منهما.

ستعامل مع الجدول من ناحية كسلسلة من الصفوف ، ثم كسلسلة من الأعمدة. عندما يتم التعامل مع الجدول في الصف ، تتم تسوية البيانات عن طريق القسمة على fi_0.

والغرض من هذا التسوية هو النظر في الروابط بين المتغيرين من خلال الفرق بين النسب المئوية في الصفوف.



ونفس الشيء بالنسبة للصفوف



4. التشابه والجمع بين الفئات

يتم تحديد التشابه بين صفين أو بين عمودين بالمسافة بين الفئات. المسافة المستخدمة هي x^2 ويتم تحديدها بشكل متماثل للصفوف والأعمدة. وبالتالي بين صفين i و i' يتم الحصول عليها بواسطة:

$$d_{\chi^2}(i,i') = \sum_{j\in J} \frac{1}{f_{\bullet j}} \left(\frac{f_{ij}}{f_{i\bullet}} - \frac{f_{i'j}}{f_{i'\bullet}} \right)^2$$

وبين عمودين j و 'jبواسطة:

$$d_{\chi^2}(j,j') = \sum_{i\in I} \frac{1}{f_{i\bullet}} \left(\frac{f_{ij}}{f_{\bullet j}} - \frac{f_{ij'}}{f_{\bullet j'}} \right)^2 .$$

5. مثال حسابي:

	D	E	F	المجموع
A	15	12	3	30
B	10	18	4	32
C	15	5	8	28
المجموع	40	35	15	90

$$V_{11} = \begin{bmatrix} \frac{30}{90} & 0 & 0 \\ 0 & \frac{32}{90} & 0 \\ 0 & 0 & \frac{28}{90} \end{bmatrix}$$

$$V_{22} = \begin{bmatrix} \frac{40}{90} & 0 & 0 \\ 0 & \frac{35}{90} & 0 \\ 0 & 0 & \frac{15}{90} \end{bmatrix}$$

$$V_{12} = \begin{bmatrix} \frac{15}{90} & \frac{12}{90} & \frac{3}{90} \\ \frac{10}{90} & \frac{18}{90} & \frac{4}{90} \\ \frac{15}{90} & \frac{5}{90} & \frac{8}{90} \end{bmatrix}$$

$$(V_{11})^{-1} * V_{12} = \begin{bmatrix} \frac{15}{30} & \frac{12}{30} & \frac{3}{30} \\ \frac{10}{32} & \frac{18}{32} & \frac{4}{32} \\ \frac{15}{28} & \frac{5}{28} & \frac{8}{28} \end{bmatrix}$$

$$V_{12} * (V_{22})^{-1} = \begin{bmatrix} \frac{15}{40} & \frac{12}{35} & \frac{3}{15} \\ \frac{10}{40} & \frac{18}{35} & \frac{4}{15} \\ \frac{15}{40} & \frac{5}{35} & \frac{8}{15} \end{bmatrix}$$

تنشيط ows
انتقل إلى إعدا

$$(V_{11})^{-1} * V_{12} * (V_{22})^{-1} * V_{21} = (V_{11})^{-1} * V_{12} * (V_{22})^{-1} * (V_{12})' = A$$

$$A = \begin{bmatrix} 0.345 & 0.357 & 0.298 \\ 0.335 & 0.400 & 0.265 \\ 0.320 & 0.301 & 0.379 \end{bmatrix}$$

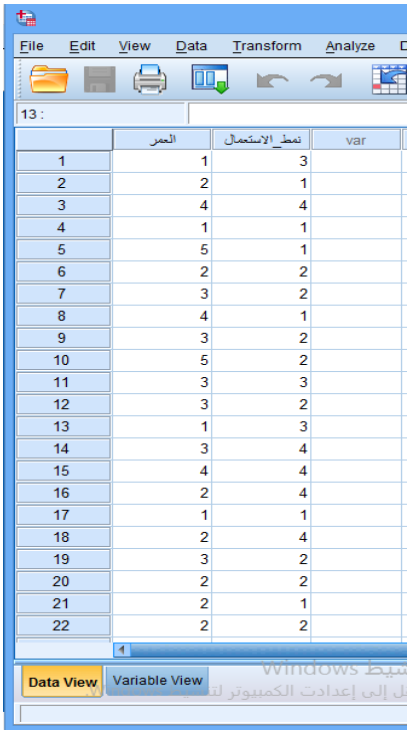
$$(V_{22})^{-1} * V_{21} * (V_{11})^{-1} * V_{12} = (V_{22})^{-1} * (V_{12})' * (V_{11})^{-1} * V_{12} = B$$

$$B = \begin{bmatrix} 0.466 & 0.358 & 0.176 \\ 0.409 & 0.452 & 0.139 \\ 0.469 & 0.325 & 0.206 \end{bmatrix}$$

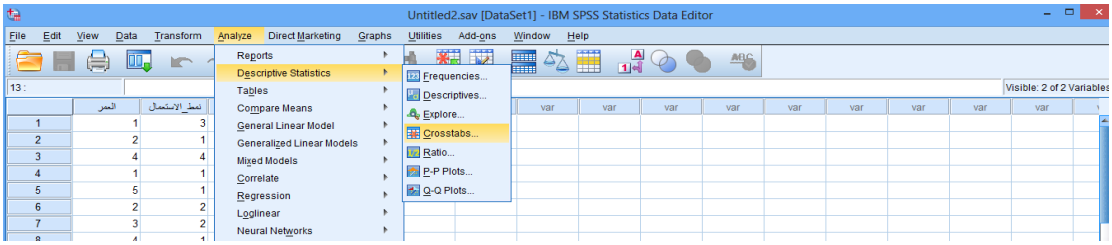
6. مثال تطبيقي: التحليل العاملي المتناظرAFCباستخدام SPSS

تم طرح سؤالين على مجموعة من المنتسبين إلى بريد الجزائر حول طريقة استخدامهم للبطاقة الدفع الالكتروني "الذهبية " وتمحورت الاسئلة حول الفئة العمرية ونمط الاستخدام.

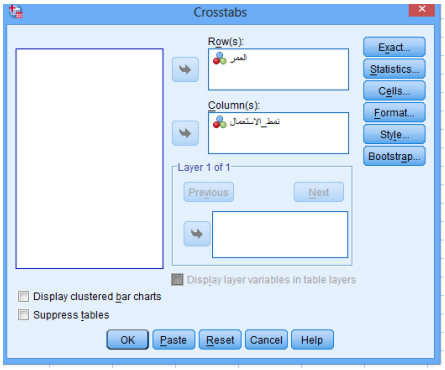
بعد ادخال البيانات في برنامج SPSS كالتالي:



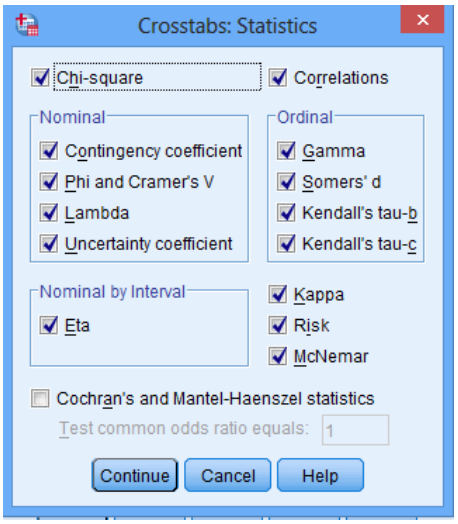
والخطوة الثانية تتمثل في تحديد التكرارات كالتالي:



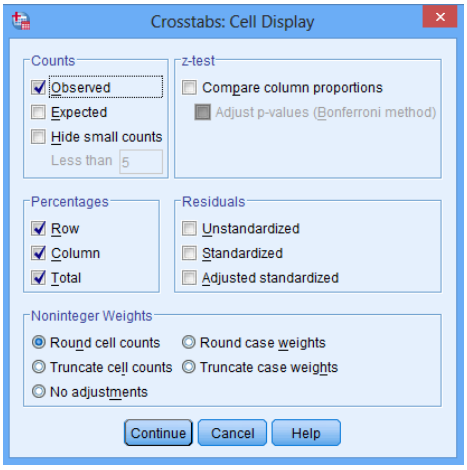
ثم اكمال الجدول الحواري التالي كما يلي



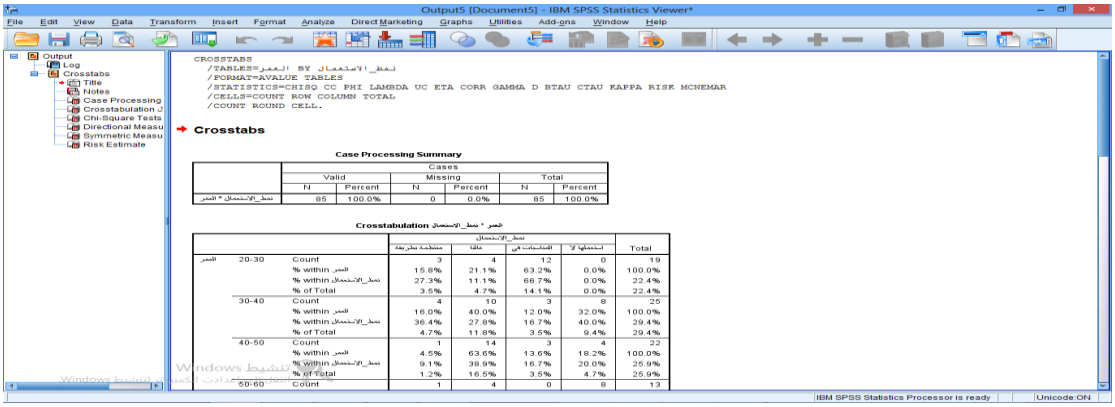
نقوم بالضغط على statistiques والتأشير على الخيارات التالية:



ومن خلال cellules نتمكن من الحصول على النسب كالتالي:



بعد الضغط على ok نحصل على نافذة المخرجات التالية:



نافذة المخرجات تحتوي على الجداول الموالية:

الحالات المعالجة:

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
الاستعمال_نمط * العمر	85	100.0%	0	0.0%	85	100.0%

نلاحظ أنه تم معالجة 85 حالة ولا توجد حالات مفقودة أو غير مدرجة.

جدول اختبار التوافق:

Symmetric Measures

		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Nominal by Nominal	Phi	.733			.000
	Cramer's V	.423			.000
	Contingency Coefficient	.591			.000
Ordinal by Ordinal	Kendall's tau-b	-.011-	.093	-.118-	.906
	Kendall's tau-c	-.011-	.091	-.118-	.906
	Gamma	-.014-	.120	-.118-	.906
	Spearman Correlation	-.016-	.113	-.150-	.881[c]
Interval by Interval	Pearson's R	-.010-	.108	-.090-	.929[c]
Measure of Agreement	Kappa	.050	.063	.865	.387
N of Valid Cases		85			

a. Not assuming the null hypothesis.

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

c. Based on normal approximation.

نشير هنا إلى أنه لا يمكن استعمال معامل الارتباط لبيرسن لأن المتغيرات نوعية وبالتالي نلجأ الي استعمال اختبار التوافق Contingency Coefficient والذي يسمح بإختبار الفرضية التالية:

$$\begin{cases} H_0: \text{لا يوجد ارتباط بين الفئة العمرية و استعمال البطاقة الذهبية} \\ H_1: \text{يوجد ارتباط بين الفئة العمرية و استعمال البطاقة الذهبية} \end{cases}$$

ومن الجدول نجد أن قيمة معامل التوافق تساوي 0.591 وهي قيمة مقبولة، وذات معنوية ذلك لأن مستوي الدلالة يساوي0.00 وهو أقل من 0.05 وبالتالي نرفض الفرضية الصفرية التي مفادها أنه لا يوجد ارتباط بين الفئة العمرية واستعمال البطاقة الذهبية، ونقبل الرضية البديلة التي تؤكد على وجود رتباط بين الفئة العمرية واستعمال البطاقة الذهبية.

الجدول المتقاطع

يوضح الجدول المتقاطع النسب بالنسبة للأسطر والأعمدة والعينة

Crosstabulation الاستعمال_نمط * العمر

			الاستعمال_نمط				
			بطريقة منتظمة	غالبا	في المناسبات	لا استعملها	Total
العمر	20-30	Count	3	4	12	0	19
		% within العمر	15.8%	21.1%	63.2%	0.0%	100.0%
		% within نمط_الاستعمال	27.3%	11.1%	66.7%	0.0%	22.4%
		% of Total	3.5%	4.7%	14.1%	0.0%	22.4%
	30-40	Count	4	10	3	8	25
		% within العمر	16.0%	40.0%	12.0%	32.0%	100.0%
		% within نمط_الاستعمال	36.4%	27.8%	16.7%	40.0%	29.4%

		% of Total	4.7%	11.8%	3.5%	9.4%	29.4%
	40-50	Count	1	14	3	4	22
		% within العمر	4.5%	63.6%	13.6%	18.2%	100.0%
		% within نمط_الاستعمال	9.1%	38.9%	16.7%	20.0%	25.9%
		% of Total	1.2%	16.5%	3.5%	4.7%	25.9%
	50-60	Count	1	4	0	8	13
		% within العمر	7.7%	30.8%	0.0%	61.5%	100.0%
		% within نمط_الاستعمال	9.1%	11.1%	0.0%	40.0%	15.3%
		% of Total	1.2%	4.7%	0.0%	9.4%	15.3%
	60 أكبر من	Count	2	4	0	0	6
		% within العمر	33.3%	66.7%	0.0%	0.0%	100.0%
		% within نمط_الاستعمال	18.2%	11.1%	0.0%	0.0%	7.1%
		% of Total	2.4%	4.7%	0.0%	0.0%	7.1%
Total		Count	11	36	18	20	85
		% within العمر	12.9%	42.4%	21.2%	23.5%	100.0%
		% within نمط_الاستعمال	100.0%	100.0%	100.0%	100.0%	100.0%
		% of Total	12.9%	42.4%	21.2%	23.5%	100.0%

نأخذ كمثال الفئة العمرية التالية

	40-50	Count	1	14	3	4	22
		% within العمر	4.5%	63.6%	13.6%	18.2%	100.0%
		% within نمط_الاستعمال	9.1%	38.9%	16.7%	20.0%	25.9%
		% of Total	1.2%	16.5%	3.5%	4.7%	25.9%
	50-60	Count	1	4	0	8	13

الفئة العمرية تظم الاشخاص الذين تتراوح اعمارهم من 40 إلى 50 سنة

حيث نجد:

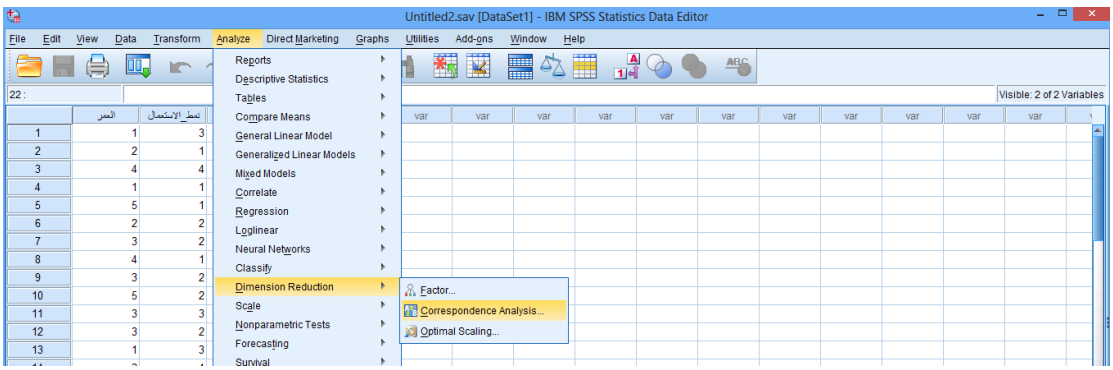
- ✓ استخدام البطاقة الذهبية بصفة منتظمة :
- شخص واحد يستخدم البطاقة الذهبية بصفة منتظمة
- 4.5% من الفئة العمرية 40-50 يستخدمون البطاقة الذهبية بصفة منتظمة

- 9.1% من الاشخاص الذين يستعملون البطاقة الذهبية بصفة منتظمة هم من الفئة العمرية 40-50 سنة.
- 1.2% هم اشخاص من الفئة العمرية 40-50 و يستعملون البطاقة الذهبية بصفة منتظمة.

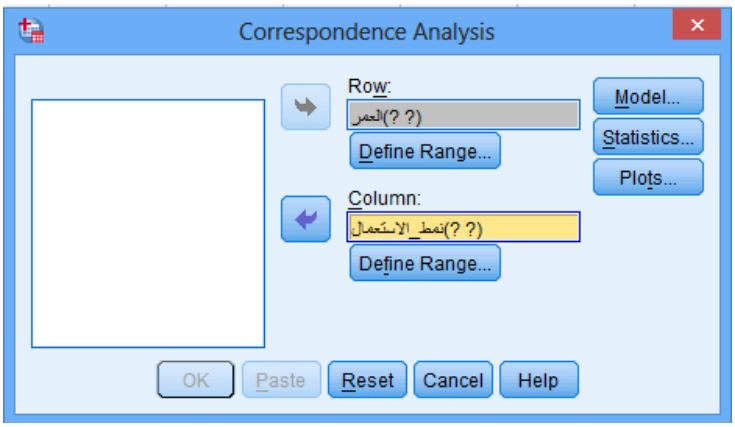
✓ استخدام البطاقة الذهبية في غالب الاحيان:

- 14 شخص يستخدم البطاقة الذهبية في غالب الاحيان
- 63.6% من الفئة العمرية 40-50 يستخدمون البطاقة الذهبية في غالب الاحيان
- 38.9% من الاشخاص الذين يستعملون البطاقة الذهبية في غالب الاحيان هم من الفئة العمرية 40-50 سنة.
- 16.5% هم اشخاص من الفئة العمرية 40-50 و يستعملون البطاقة الذهبية في غالب الاحيان.

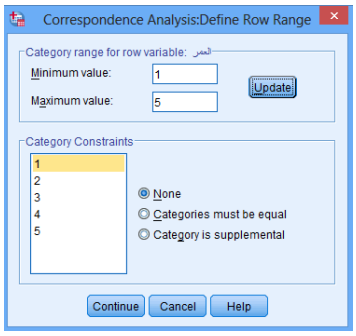
تطبيق AFC بالخطوات التالية:



نقوم بنقل المتغيرات كما هو في الشكل التالي

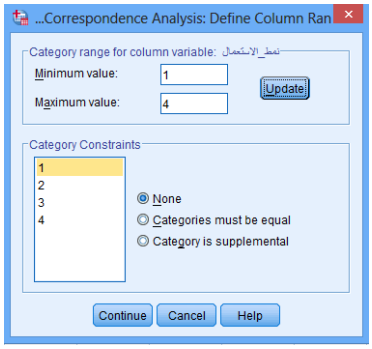


نقوم بإدخال القيم المتعلقة بالفئة العمرية كالتالي



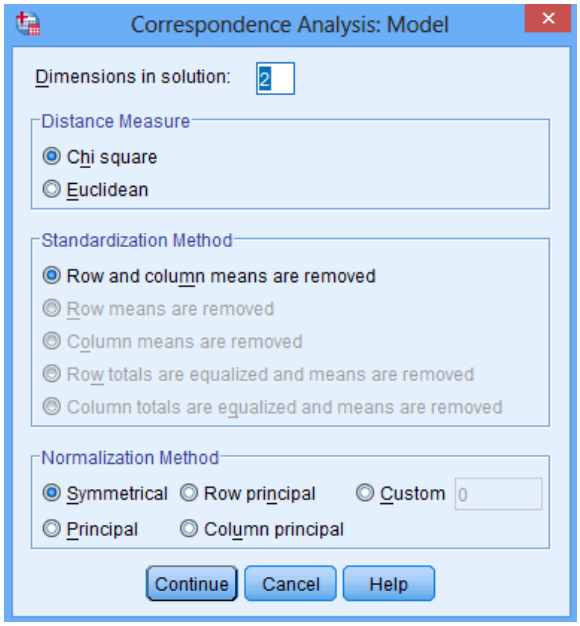
ثم نضغط علىUpdate ثم continue

نقوم بإدخال القيم المتعلقة بطريقة استخدام البطاقة كالتالي



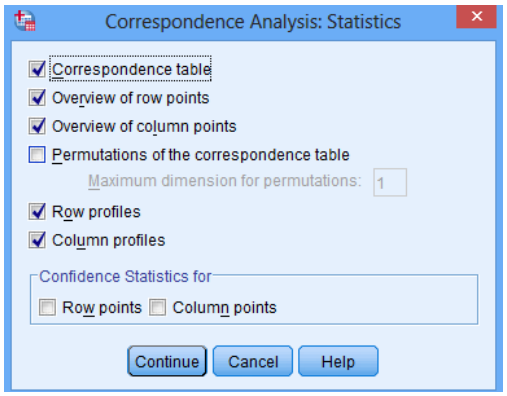
ثم نضغط علىUpdate ثم continue

من Model نؤشر على الخيارات التالية:



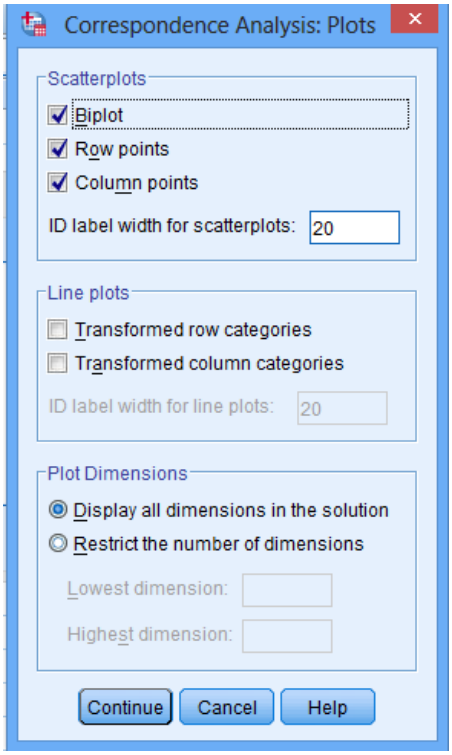
ثم نضغط على continue

نقوم بالضغط على statistiques والتأشير على الخيارات التالية:



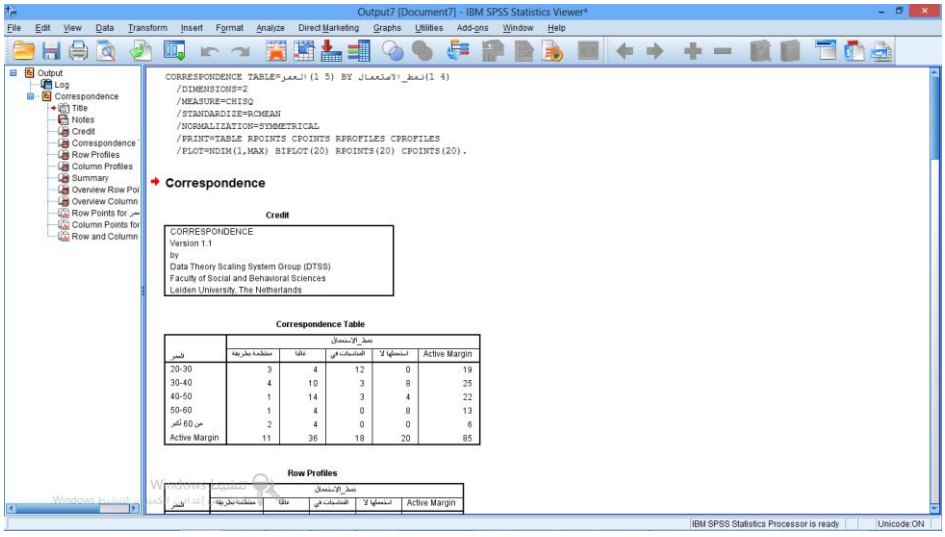
ثم نضغط على continue

نقوم بالضغط على Plots والتأشير على الخيارات التالية:



ثم نضغط على continue

بعد الانتهاء من الخطوات السابقة نضغط ok تظهر نافذة المخرجات التالية



والتي تحوي المخرجات التالية:

الجدول المعطيات الاساسية:

Correspondence Table

العمر	الاستعمال_نمط				
	بطريقة منتظمة	غالبا	في المناسبات	لا استعملها	Active Margin
20-30	3	4	12	0	19
30-40	4	10	3	8	25
40-50	1	14	3	4	22
50-60	1	4	0	8	13
60 أكبر من	2	4	0	0	6
Active Margin	11	36	18	20	85

حيث يوضح مكونات عناصر الاسطر ومكونات عناصر الاعمدة والتكرارات .

جدول التكرارات النسبية للأسطر

Row Profiles

العمر	الاستعمال_نمط				
	بطريقة منتظمة	غالبا	في المناسبات	لا استعملها	Active Margin
20-30	.158	.211	.632	.000	1.000
30-40	.160	.400	.120	.320	1.000
40-50	.045	.636	.136	.182	1.000
50-60	.077	.308	.000	.615	1.000
60 أكبر من	.333	.667	.000	.000	1.000
Mass	.129	.424	.212	.235	

جدول التكرارات النسبية للأعمدة

Column Profiles

العمر	الاستعمال_نمط				
	بطريقة منتظمة	غالبا	في المناسبات	لا استعملها	Mass
20-30	.273	.111	.667	.000	.224
30-40	.364	.278	.167	.400	.294
40-50	.091	.389	.167	.200	.259
50-60	.091	.111	.000	.400	.153
60 أكبر من	.182	.111	.000	.000	.071
Active Margin	1.000	1.000	1.000	1.000	

جدول العوامل المستخرجة:

Summary

Dimension	Singular Value	Inertia	Chi Square	Sig.	Proportion of Inertia		Confidence Singular Value	
					Accounted for	Cumulative	Standard Deviation	Correlation 2
1	.613	.376			.700	.700	.077	.446
2	.343	.118			.219	.918	.090	
3	.209	.044			.082	1.000		
Total		.538	45.689	.000[a]	1.000	1.000		

a. 12 degrees of freedom

حيث نلاحظ أنه تم تلخيص البيانات الاساسية في ثلاثة محاور المحور الاول يختزل 70% من نسبة التشتت أو من التباين الكلي، المحور الثاني يختزل 21.9% والمحور الثالث يختزل 8.2%.

ويوضح الجدول قيمت كاي مربع حيث يقيس وجود علاقة بين العمر واستخدام البطاقة من عدمه ، وقد بلغت قيمة كاي مربع 45.86 وهي قيمة مقبولة، وذات معنوية ذلك لأن مستوي الدلالة يساوي0.00 وهو أقل من 0.05 وبالتالي نرفض الفرضية الصفرية التي مفادها أنه لا يوجد ارتباط بين الفئة العمرية واستعمال البطاقة الذهبية، ونقبل الرضية البديلة التي تؤكد على وجود رتباط بين الفئة العمرية واستعمال البطاقة الذهبية.

جدول خصائص الاسطر

Overview Row Points[a]

		Score in Dimension			Contribution				
					Of Point to Inertia of Dimension		Of Dimension to Inertia of Point		
العمر	Mass	1	2	Inertia	1	2	1	2	Total
20-30	.224	-1.365-	-.334-	.264	.679	.073	.967	.032	1.000
30-40	.294	.311	-.090-	.023	.046	.007	.753	.035	.788
40-50	.259	.190	.508	.052	.015	.195	.110	.441	.551
50-60	.153	1.016	-.846-	.134	.257	.319	.720	.279	1.000
60 أكبر من	.071	.129	1.405	.064	.002	.406	.011	.746	.757
Active Total	1.000			.538	1.000	1.000			

a. Symmetrical normalization

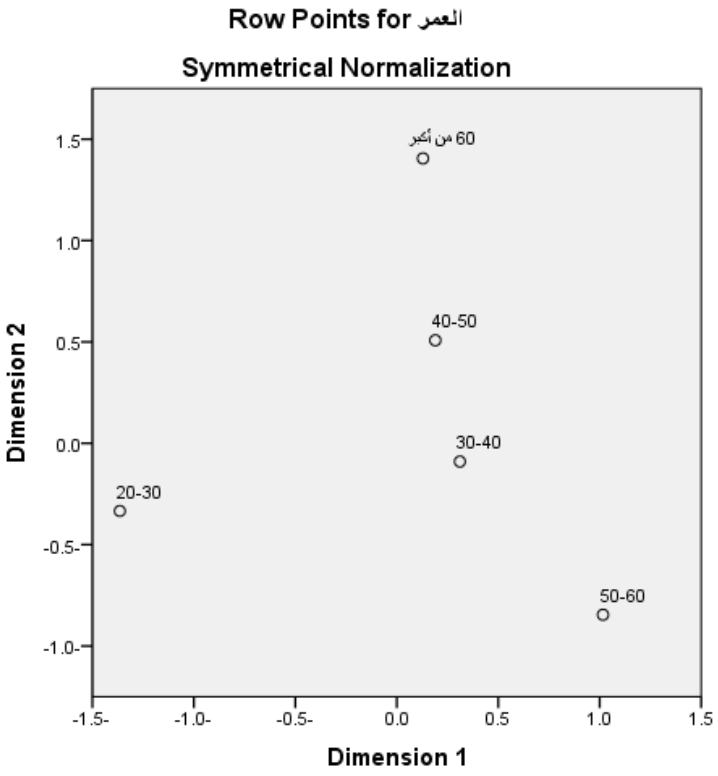
✓ المساهمة في عطالة الابعاد:

- نلاحظ أن أكبر مساهمة للعطالة في المحور الاول هي للفئة العمرية 20-30سنة
- نلاحظ أن أكبر مساهمة للعطالة في المحور الثاني هي للفئة العمرية أكبر من 60سنة

✓ المساهمة في عطالة النقاط:

- يساهم المحور الاول في عطالة 96.7% بالنسبة للفئة العمرية 20-30 و75.3% بالنسبة للفئة العمرية30-40 و72% بالنسبة للفئة العمرية 50-60 سنة.
- يساهم المحور الثاني في عطالة 74.6% بالنسبة للفئة العمرية أكبر من 60 سنة.

كما هو موضح في الشكل التالي:



جدول خصائص الاعمدة

Overview Column Points[a]

نمط_الاستعمال	Mass	Score in Dimension 1	Score in Dimension 2	Inertia	Contribution: Of Point to Inertia of Dimension 1	Contribution: Of Point to Inertia of Dimension 2	Contribution: Of Dimension to Inertia of Point 1	Contribution: Of Dimension to Inertia of Point 2	Contribution: Of Dimension to Inertia of Point Total
بطريقة منتظمة	.129	-.206-	.294	.044	.009	.033	.077	.088	.165
غالبا	.424	.221	.576	.067	.034	.409	.191	.722	.912
فى المناسبات	.212	-1.348-	-.447-	.252	.627	.123	.936	.058	.993
لا استعملها	.235	.927	-.796-	.175	.330	.435	.708	.292	1.000
Active Total	1.000			.538	1.000	1.000			

a. Symmetrical normalization

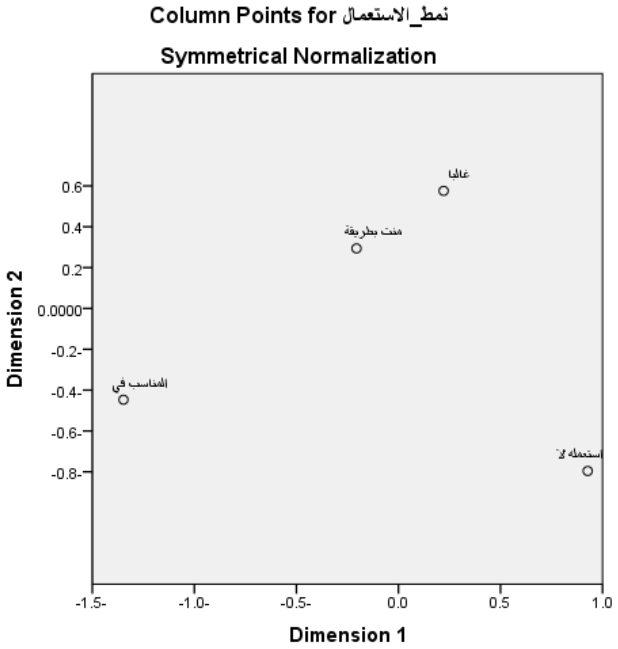
✓ المساهمة في عطالة الابعاد:

- نلاحظ أن أكبر مساهمة للعطالة في المحور الاول هي استخدام البطاقة في المناسبات
- نلاحظ أن أكبر مساهمة للعطالة في المحور الثاني عدم استخدام البطاقة

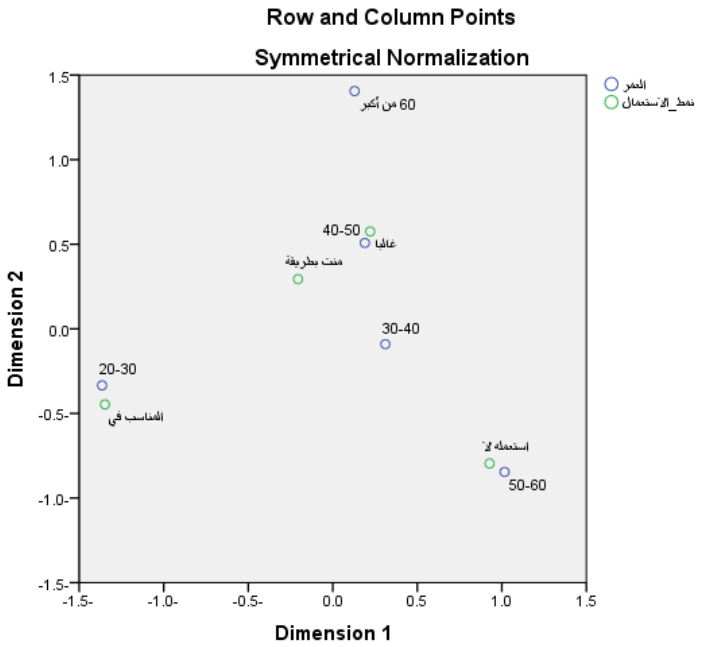
✓ المساهمة في عطالة النقاط:

- يساهم المحور الاول في عطالة 93.6% بالنسبة لاستخدام البطاقة في المناسبات و70.8% بالنسبة لعدم استخدام البطاقة.
- يساهم المحور الثاني في عطالة 72.2% بالنسبة لاستخدام البطاقة غالبا.

كما هو موضح في الشكل التالي:



تموقع نقاط الاسطر والأعمدة:



من الشكل السابق الذي يوضح تموقع نقاط الاسطر والأعمدة نستنتج ثلاثة مجموعات

- المجموعة الاولى: حيث تظم الاشخاص الذين لايستعملون البطاقة والفئة العمرية 50-60.
- المجموع الثانية: تظم الاشخاص الذين يستعملون البطاقة في المناسبات والفئة العمرية 20-30 سنة.
- المجموعة الثالثة: وتضم الاشخاص الذين يستعملون البطاقة في غالب الاحيان أو بطريقة منتظمة والفئتين العمريتين 30-40 و40-50 سنة.

المحور الثالث:

التحليل العاملي التميزي AFD

يعد التحليل التمييزي أحد المناهج الإحصائية المهمة في التحليل الإحصائي متعددة المتغيرات والذي يعني في كيفية التمييز بين مجموعتين أو أكثر من المجتمعات وأن الفكرة الأساسية للتمييز هي التفرقة بين مجتمعات متداخلة أو متشابهة لها الخصائص أو الصفات بمعنى آخر أن التحليل التمييزي هو أسلوب إحصائي يتم بموجبه استعمال مجموعه من المتغيرات للتمييز بين مجموعتين أو أكثر عن طريق دالة تميزية محددة وهي توليفة خطية للمتغيرات التوضيحية و طريقة إيجاد هذه الدالة بإيجاد المعاملات للدالة وفقا للقياسات أو المعايير التي يتم الحصول عليها من المشاهدات.

إن عملية التصنيف وهي العملية اللاحقة لعملية تكوين الدالة التمييزية إذ يتم الاعتماد على هذه الدالة بالتنبؤ وتصنيف المفردة الجديدة لإحدى المجموعات قيد الدراسة بأقل خطأ تصنيف ممكن والتحليل التمييزي أيضا يمكن أستخدمه في مختلف المجالات وفي حالة المجتمعات المتجانسة وغير المتجانسة.

1. اهداف التحليل التميزي

أما الهدف الرئيسي من التحليل التميزي تصنيف المشاهدة او مجموعة من المشاهدات إلى مجاميعها التصنيفية وبأقل خطا تصنيف ممكن.

إن دالة التمييز الخطية والمستندة إلى تركيب خطي للمتغيرات لكي تكون مثالية يجب أن تنتج اصغر احتمال لخطا التصنيف علما بان هناك افتراضات يجب توافرها عند البيانات المستخدمة في التحليل ولكن غالبا ما

نواجه اختراقا لبعض الفرضيات ومن ثم فان تقديرات دالة التمييز الخطية تفقد خواصها فمثلا: عند عدم تساوي مصفوفة التباين والتباين المشترك فان استعمال دالة التمييز التربيعية يكون ضروريا .

2. الدالة التمييزية :

هي دلة يمكن من خلالها التمييز بين المجموعات (الفصل بين المشاهدات) ووضع كل مشاهدة في المجموعة التي تتبع لها بمعنى آخر هي الدالة التي يمكن بواسطتها تمييز (تصنيف) المشاهدات الجديدة (مجهولة الانتماء) الى المجموعة الصحيحة التي يفترض انتمائهم إليها وفقا للمعايير او القياسات التي تم الحصول عليها من المشاهدات المعلومة سابقا.

1.2- قوة دالة التمييز:

تعتمد دالة التمييز علي سلامة توزيعها للمشاهدات (المفردات) علي المجموعات الصحيحة (أي التي تنتمي لها المشاهدات فعلياً) وعليه فإن قوه داله التمييز ترتبط طرديا بالتوزيع السليم للمشاهدات.

2.2- أنواع دوال التمييز:

- ✓ دالة ال تمييز الخطية
- ✓ دالة التمييز غير الخطية
- ✓ دالة التمييز اللوجستية

3. تقدير معلمات الدالة التميزية الخطية:

لتقدير معلمات الدالة التمييزية في حالة وجود مجموعتين نتبع الخطوات الآتية :

1- يجاد متوسط كل متغير في كل مجموعة و كالآتي :

المجموعة الأولى :

$$\bar{X}_1^{(1)} = \sum_{i=1}^{n1} \frac{X_{1i}}{n_1}$$

$$\bar{X}_2^{(1)} = \sum_{i=1}^{n1} \frac{X_{2i}}{n_1}$$

$$\bar{X}_k^{(1)} = \sum_{i=1}^{n1} \frac{X_{ki}}{n_1}$$

حيـــــث $\bar{X}_i^{(1)}$ يمثل متوسط المتغير في المجموعة الأولى .

بنفس الطريقة للمجموعة الثانية :

$$\bar{X}_1^{(2)} = \sum_{i=1}^{n1} \frac{X_{1i}}{n_2}$$

$$\bar{X}_2^{(2)} = \sum_{i=1}^{n2} \frac{X_{2i}}{n_2}$$

$$\vdots$$

$$\bar{X}_k^{(2)} = \sum_{i=1}^{n2} \frac{X_{ki}}{n_2}$$

حيــث $\bar{X}_i^{(2)}$ يمثل متوسط المتغير في المجموعة الثانية :

2- إيجـــاد المســـافة بـــين المتغيـــرينdi مـــن خـــلال إيجـــاد الفـــرق بـــين متوســـطي كـــل متغيرين من المجموعتين .

$$d_i = \bar{X}_{i(1)} - \bar{X}_{i(2)} \, , i = 1,2,3 \ldots . , k$$

$$d_i = \bar{X}_{1(1)} - \bar{X}_{1(2)}$$
$$d_i = \bar{X}_{2(1)} - \bar{X}_{2(2)}$$
$$\vdots$$
$$d_k = \bar{X}_{k(1)} - \bar{X}_{k(2)}$$

يتم وضع هذه الفروق في شكل شعاع عمودي ويرمز له بالرمز d

$$d = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_k \end{bmatrix}$$

3- حساب مجموع المربعات لكل متغير في كل مجموعة و كذلك مجموعة حاصل ضرب كل متغيرين داخل كل مجموعة كالآتي :

$$SS_{x_1}^{(1)} = SS_{11}^{(1)} = \sum_{i=1}^{n1} X_{1i}^2 - \frac{\left(\sum_{i=1}^{n1} X_{1i}\right)^2}{n1}$$

$$\vdots$$

$$SS_{kk}^{(1)} = \sum_{i=1}^{n1} X_{ki}^2 - \frac{\left(\sum_{i=1}^{n1} X_{ki}\right)^2}{n_1}$$

نتابع الحساب بحيث

$$SS_{ii}^{(1)} = \sum_{i=1}^{n1} X_{ii}^2 - \frac{\left(\sum_{i=1}^{n1} X_{ii}\right)^2}{n_1} \quad i = 1,2, \ldots \ldots , n1$$

يمثل مجموع مربعات المتغير (i) في المجموعة الأولي

بنفس الطريقة للمجموعة الثانية بحيث

$$SS_{ii}^{(2)} = \sum_{i=1}^{n2} X_{ii}^2 - \frac{\left(\sum_{i=1}^{n2} X_{ii}\right)^2}{n_2} \quad i = 1,2, \ldots \ldots , n2)$$

يمثل مجموع مربعات المتغير (i) في المجموعة الثانية

ويحسب حاصل ضرب المتغيرين بالعلاقة التالية:

$$S_{ii} = \sum X_{ij} - \frac{(\sum X_i)(\sum X_j)}{n}$$

4- حساب مصفوفة التباين و التباين المشترك للمجموعتين

- حساب التباين

$$V_{ii} = \frac{S_{i(1)} + S_{i(2)}}{n_1 + n_2 - 2}$$

- حساب التغاير

$$V_{ii} = \frac{S_{ij(1)} + S_{ij(2)}}{n_1 + n_2 - 2}$$

- حساب مصفوفة التباين و التباين (S_p^2) المشترك داخل المجموعتين

$$S_p^2 = \frac{[(n_1 - 1)S_{(1)} + (n_2 - 1)S_{(2)}]}{n_1 + n_2 - 2}$$

وعليه من المعادلات الأخيرة فإن مصفوفة التباين و التباين المشترك تكون بالشكل التالي:

$$V = \begin{bmatrix} V_{11} & V_{12} & \cdots & V_{1k} \\ V_{21} & V_{22} & \cdots & V_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ V_{k1} & V_{k2} & \cdots & V_{kk} \end{bmatrix}$$

ومن المعلوم أن مصفوفة التباين والتباين المشترك مربعة و متماثلة قطرها الرئيسي يمثل التباينات و عناصرها الأخرى تمثل التغايرات مشتركة.

5- حساب قيم معاملات الدالة التميزية الخطية

تحسب معاملات الدالة التميزية الخطية من المعادلة التالية

$$\begin{bmatrix} V_{11} & V_{12} & \cdots & V_{1k} \\ V_{21} & V_{22} & \cdots & V_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ V_{k1} & V_{k2} & \cdots & V_{kk} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_k \end{bmatrix}$$

وعليه فإن معاملات الدالة التميزية الخطة تكتب كالتالي:

$$\alpha = V^{-1} d$$

4. الاهمية النسبية للمتغيرات

اهمية التحليل التمييزي هو أنه يسمح بالمقارنة بين المتغيرات المستقلة من حيث أهميتها في عملية التمييز ، و تحسب الاهمية النسبية للمتغيرات من الصيغة التالية :

$$\alpha_j^* = \alpha_j \sqrt{V_{jj}}$$

بحيث: $j = 1,2, \dots \dots \dots \dots \dots \dots, k$

نقوم بمقارنة القيم الناتجة من حساب α_j^* وترتيبهم تنازليا ، وعليه فإن أكبر قيمة من جملة القيمة هو أهم متغير له القدرة على عملية التمييز بين المجموعتين ويليها ثاني أكبر قيمة له القدرة على التمييز إلى اخر قيمة.

5. اختبار قدرة دالة التمييز في التمييز بين مجموعتين :

بعد حساب الدالة التمييزية فأن اختبار قدرتها على التمييز بين مجموعتين تعتبر مرحلة ذات أهمية بالغة في التحليل التمييزي ،ويتم اختبار قدرة الدالة على التمييز بين مجموعتين بإختبار ستيودنت t أو اختبار فيشر F:

1.5- إختبار ستيودنت t لقدرة دالة التمييزية

1- نحسب القيم التمييزية لكل مشاهدة في كل مجموعة و ذلك بتعويض قيم المتغيرات Xi المستقلة .

نحسب القيمة التمييزية للمشاهدة الأولي في المجموعة الأولي :

$$L_{1(1)} = \alpha_1 X_{11} + \alpha_2 X_{21} + \cdots + \alpha_k X_{k1}$$

نحسب القيمة التمييزية للمشاهدة الثانيةفي المجموعة الأولي :

$$L_{2(1)} = \alpha_1 X_{12} + \alpha_2 X_{22} + \cdots + \alpha_k X_{k2}$$

بنفس الطريقة لباقي مشاهدات المجموعة الاولى

2- نقوم بحساب الوسط الحسابي للقيم التمييزية لكل مجموعة :

- الوسط الحسابي للقيم التمييزية للمجموعة الاولى

$$\bar{L}_{(1)} = \frac{\sum_{i=1}^{n1} L_i^{(1)}}{n_1}$$

حيث $i= 1,2,3.....n_1$

- الوسط الحسابي للقيم التمييزية للمجموعة الثانية

$$\bar{L}_{(2)} = \frac{\sum_{i=1}^{n2} L_i^{(2)}}{n_2}$$

حيث $i= 1,2,3.....n_2$

نقوم باختبار الفرق بين متوسط المجموعتين ، ونستخدم لذلك الأختبار الإحصائي (t) للمقارنة بين المتوساطات الحسابية للمجموعات و ذلك لبيان أهمية دالة التصنيف ، ونقوم بإجراء الاختبار الإحصائي كالتالي :

بحيث فرضيات الاختبار تشكل كالتالي:

$$\begin{cases} H_0: \text{الدالة ليس لها القدرة على عملية التمييز} \\ H_1: \text{الدالة لها القدرة على عملية التمييز} \end{cases}$$

ويتم اختبار هذه الفرضية من خلال اختبار الفرق بين متوسطي القيم التمييزية للمجموعتين ، وتكتب الفرضية كالتالي:

$$\begin{cases} H_0: ML_1 = ML_2 \\ H_1: ML_1 \neq ML_2 \end{cases}$$

$$t = \frac{\bar{L}_{(1)} - \bar{L}_{(2)}}{\sqrt{S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

نقارن قيمة t المحسوبة مع الجدولية $t_{(n_1+n_2-2,\frac{\alpha}{2})}$

من الفرضية الصفرية متوسط القيمة التمييزية للمجموعة الأولى لا يختلف إحصائياً عن متوسط القيم التمييزية للمجموعة الثانية مع العلم بأن القيم التمييزية للمجموعتين حسبت من دالة واحدة ، فإذا قبلت هذه الفرضية فذلك يعني أن نمط القيم التمييزية في المجموعتين متشابهة و هذا يعن عدم قدرة الدالة التمييزية على التمييز ، أما إذا رفضة الفرضية الصفرية فهذا يعني قدرة الدالة التمييزية على التميز .

2.5- اختبار فيشر F لقدرة دالة التمييزية:

يتم صياغة الفرضيات كالتالي:

$$\begin{cases} H_0: \text{الدالة ليس لها القدرة على عملية التمييز} \\ H_1: \text{الدالة لها القدرة على عملية التمييز} \end{cases}$$

ويتم اختبار هذه الفرضية من خلال اختبار الفرق بين متوسطي القيم التميزية للمجموعتين ، وتكتب الفرضية كالتالي:

$$\begin{cases} H_0: ML_1 = ML_2 \\ H_1: ML_1 \neq ML_2 \end{cases}$$

نقوم بتكوين جدول تحليل التباين كالتالي:

S.O.V	d.f	S.S	M.S	F
Between X	k-1	SSB	MSB	$\frac{MSB}{MSE}$
Within X	n_1+n_2-k-1	SSE	MSE	
Total	n_1+n_2-2			

ونقوم بالحسابات التالية

- مجموع المربعات داخل المتغيرات

$$SSE = D^2$$

$$D^2 = \alpha_1 d_1 + \alpha_2 d_2 + \cdots + \alpha_k d_k$$

- مجموع المربعات بين المتغيرات المستقلة

$$SSB = \frac{n_1 n_2}{(n_1+n_2)(n_1+n_2-2)} (D^2)^2)$$

- مجموع المربعات الكلي

$$SST = SSB + SSE$$

نجد قيمة F المحسوبة كالآتي

$$F = \frac{n_1+n_2-m-1}{(n_1+n_2-2)m} \quad T^2$$

بحيث m تمثل درجة حرية المتغيرات

أو من

$$F = \frac{MSB}{MSE} = \frac{n_1 n_2}{(n_1 + n_2)(n_1 + n_2 - 2)} D^2$$

بقارنة قيمة F المحسوبة مع الجدولية $F_{(n_1+n_2-k-1,\alpha)}$ فإذا كانت القيمة المحسوبة أكبر من الجدولية نرفض الفرضية الصفرية ونقبل الفرضية البديلة أي أن الدالة التمييزية لها القدرة على التمييز.

6. النقطة الفاصلة

النقطة الفاصلة تمثل الحد الفاصل الذي يفصل بين المجموعتين، وتستخدم لغرض تصنيف المشاهدات إلى المجموعة الأقرب لها ،

فإذا كانت قيمة الدالة بعد تعويض قيمة المشاهدة فيها أكبر من الصفر فالمشاهدة تصنف في المجموعة الأولى ، أما إذا كانت قيمة الدالة أصغر من الصفر فيتم تصنيف المشاهدة في المجموعة الثانية ، ولكن عندما تكون الدالة التمييزية للمجموعة الأولي أكبر من قيمة الدالة التمييزية في المجموعة الثانية فإن المشاهدة تصنف في المجموعة الأولى كما يلي:

$$L > \frac{1}{2}\left(\bar{L}_1^{(1)} + \bar{L}_2^{(2)}\right)$$

تصنف المشاهدة في المجموعة الثانية إذا:

$$L < \frac{1}{2}\left(\bar{L}_1^{(1)} + \bar{L}_2^{(2)}\right)$$

بفرض أن:

$$\alpha_0 = \frac{1}{2}\left(\bar{L}_1^{(1)} + \bar{L}_2^{(2)}\right)$$

للحصول على دالة التصنيف L* نقوم بدمج النقطة الفاصلة α0 مع الدالة التميزية فنحصل على:

$$L^* = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \cdots + \alpha_k x_k$$

يتم تصنيف المشاهدات إلى المجموعات حسب العلاقات التالية:

$$\begin{cases} L^* > \text{تصنف المشاهد في المجموعة الاولى} \\ L^* < \text{تصنف المشاهد في المجموعة الثانية} \\ L^* = 0 \text{ لا يمكن إجراء التصنيف} \end{cases}$$

7. نسبة الخطأ

عند تحديد النقطة الفاصلة بين المجموعتين فإنه قد يكون هناك تصنيف غير صحيح عند استعمال دالة L* التمييز وقد تصنف مشاهدة معينة في المجموعة الاولى بينما تعود في الحقيقة إلى المجموعة الثانية والعكس.وهناك نوعان من أخطاء التصنيف

- خطأ التصنيف الظاهري:

و يحسب من جدول التصنيف التالي:

المجموعة	تابع المجموعة الاولى	تابع المجموعة الثانية	مجموع
الاولى	n_{11}	n_{12}	n_1
الثانية	n_{21}	n_{22}	n_2

n_{11}: عـــدد المشـــاهدات مـــن المجموعـــة الأولـــى والتـــي تـــم تصـــنيفها فـــي نفـــس المجموعة وبالتالي صنفت بطريقة صحيحة

n_{12}: عـــدد المشـــاهدات مـــن المجموعـــة الأولـــى والتـــي تـــم تصـــنيفها خطـــأ فـــي المجموعة الثانية

n_{21}: عـــدد المشـــاهدات التـــي تنتمـــي إلـــى المجموعـــة الثانيـــة وتـــم تصـــنيفها خطـــأ فـــي المجموعة الأولى

n_{22}: عـــدد المشـــاهدات فـــي المجموعـــة الثانيـــة التـــي تـــم تصـــنيفها فـــي نفـــس المجموعة و صنفت بطريقة صحيحة

ونحسب الخطأ الظاهري كما يلي:

P_{12}: نســـبة المشـــاهدات التـــي تنتمـــي للمجموعـــة الأولـــى وصـــنفت خطـــأ فـــي المجموعة الثانية

$$P_{12} = \frac{n_{12}}{n_1}$$

P_{21}: نسبة المشاهدات التي تنتمي للمجموعة الثانية وصنفت خطأ في الأولى

$$P_{21} = \frac{n_{21}}{n_2}$$

ونحسب معدل الخطأ الظاهري كالتالي:

$$\frac{n_{12} + n_{21}}{n_1 + n_2}$$

- خطأ التصنيف الحقيقي

ويعتمد في تحديد احتمال خطأ التصنيف على D^2 إحصائية مهلونوبيس (Mahalnobis)

$$P_1 = P_2 = F\left(-\frac{\sqrt{D^2}}{2}\right)$$

وكلما كانت صغيرةD^2 فهو دليل على صغر خطأ التصنيف والكفاءة دالة التمييز والعكس.

8. مثال تطبيقي1: التصنيف إلى مجموعتين

في هذا المثال سنحاول توضيح مراحل التحليل التمييزي و نتائج باستخدام SPSS و لهذا الغرض نستخدم المثال التالي الذي يمثل عينة من السيارات، ولقد اعتمدنا عل سبعة متغيرات موضحة في الجدول الموالي::

CYL	PUISS	V.MAX	B.V	LONG	LARG	BOIS
870	161	393	1	165	79	1350
1110	177	468	0	160	85	1588
1050	168	424	0	152	68	1294
930	161	412	0	151	59	1222
1105	164	439	1	165	98	1585
1080	169	429	1	160	82	1297
1160	169	449	0	154	79	1796
1010	163	424	0	140	55	1565
1320	173	452	1	180	128	2664
815	157	399	0	140	55	1166
1060	162	428	1	175	109	1570
1160	172	445	1	158	82	1798
1370	169	469	1	160	115	1998
1080	170	438	1	167	98	1993
1129	166	431	0	144	80	1442
1095	165	440	1	165	83	1769
1120	173	459	1	173	100	1979
955	161	404	0	140	68	1294

حيث تمثل المتغيرات:

CYL: سعة الاسطوانة

PUISS: القوة

V.MAX: السرعة القصوى

B.V: نوع مغير السرعة (0: يدوي، 1: أوتوماتيكي)

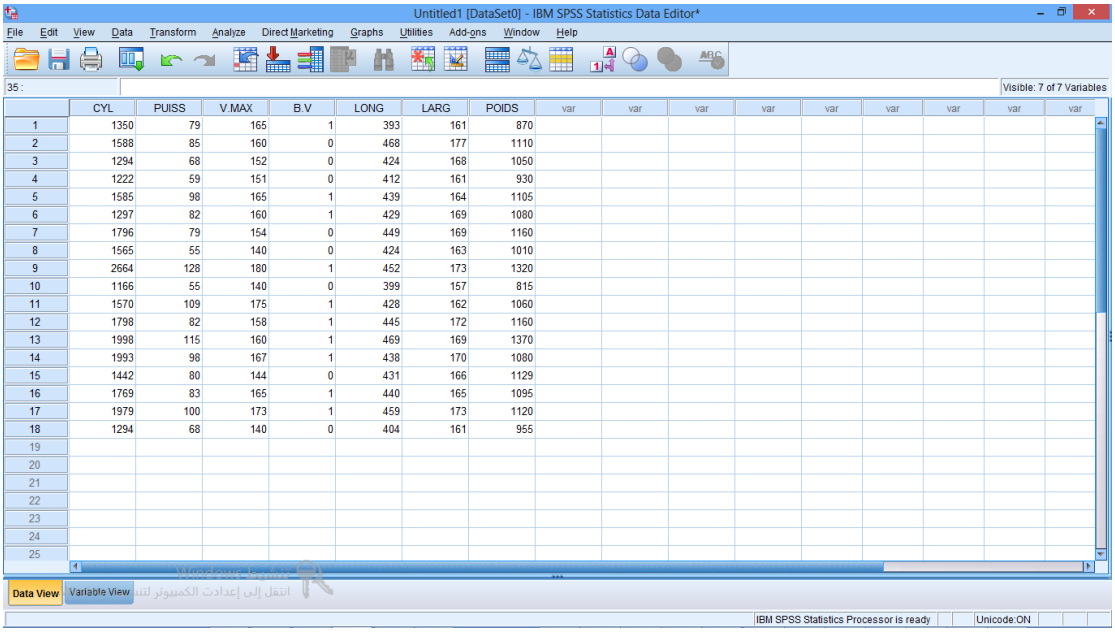
LONG: طول السيارة

LARG: عرض السيارة

BOIS: وزن السيارة

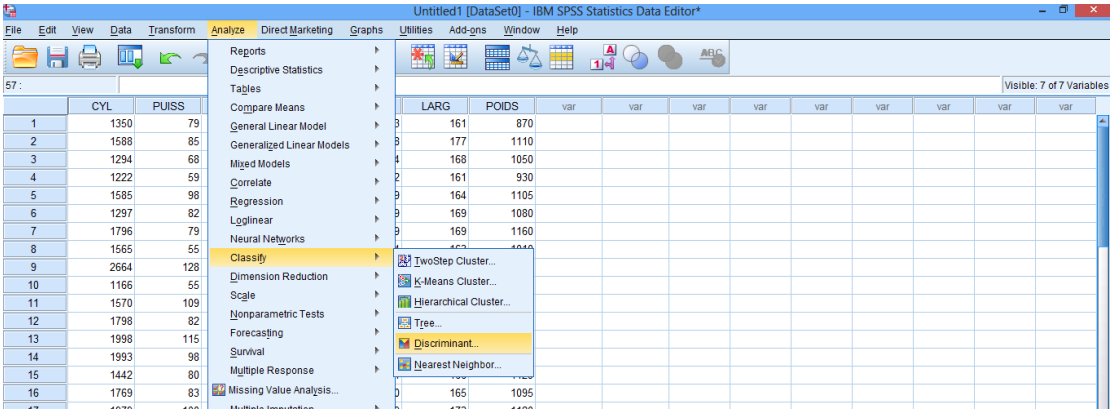
ومن خلال المعطيات الممثلة في الجدول السابق نميز بين نوعين من السيارات نوع بعلبة ناقل السرعات يدوي والأخر بعلبة ناقل للسرعات أوتوماتيكي وهي تتبع المتغيرات المتبقية، وسنحاول توظيح ذلك من خلال التحليل التمييزي بإتباع الخطوات التالية:

بعد ادخال المتغيرات في برنامج SPSSكما هو موضح في الشكل التالي:

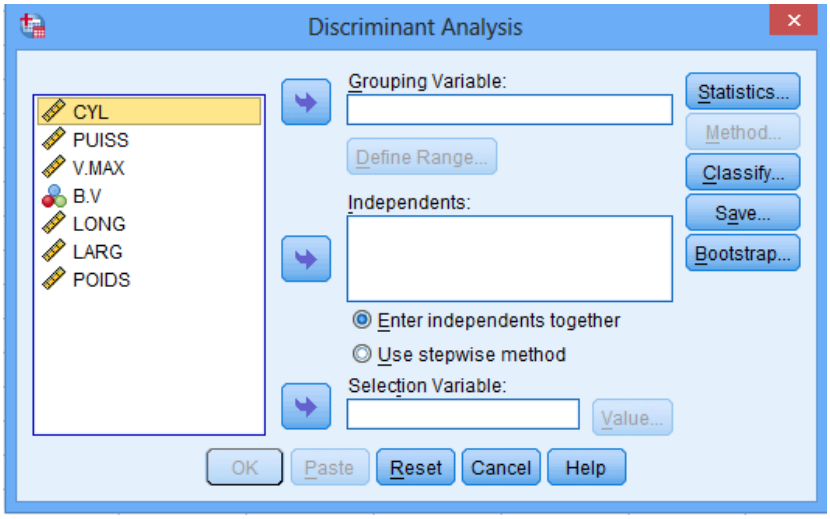


	CYL	PUISS	V.MAX	B.V	LONG	LARG	POIDS
1	1350	79	165	1	393	161	870
2	1588	85	160	0	468	177	1110
3	1294	68	152	0	424	168	1050
4	1222	59	151	0	412	161	930
5	1585	98	165	1	439	164	1105
6	1297	82	160	1	429	169	1080
7	1796	79	154	0	449	169	1160
8	1565	55	140	0	424	163	1010
9	2664	128	180	1	452	173	1320
10	1166	55	140	0	399	157	815
11	1570	109	175	1	428	162	1060
12	1798	82	158	1	445	172	1160
13	1998	115	160	1	469	169	1370
14	1993	98	167	1	438	170	1080
15	1442	80	144	0	431	166	1129
16	1769	83	165	1	440	165	1095
17	1979	100	173	1	459	173	1120
18	1294	68	140	0	404	161	955

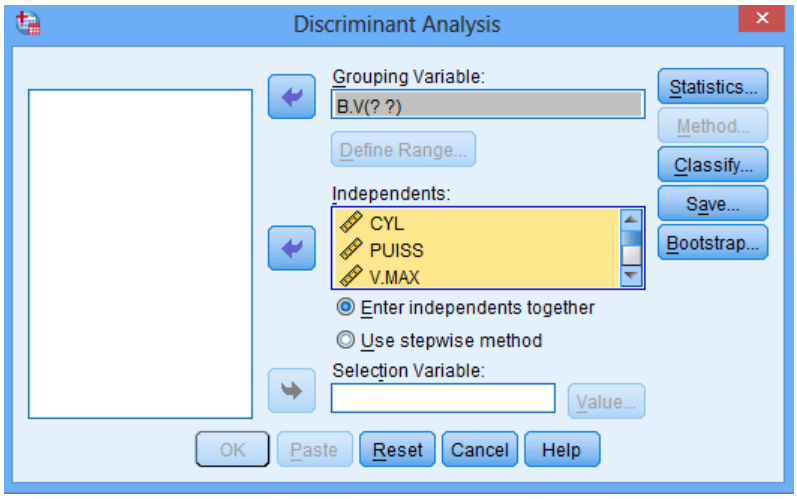
من قائمة Analyse نختار Classify ثم نتختار Disctiminant



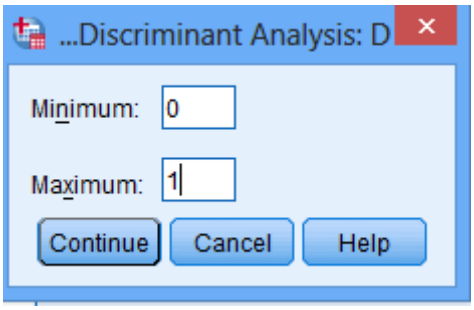
يظهر المربع الحواري التالي



نقوم بتحديد المتغير التابع الذي يمثل المجموعتين يدوي أو أوتوماتيكي B.V (علبة السرعات) ثم ننقله إلى مستطيل Grouping Variable ، وننقل المتغيرات المستقلة(CYL،PUISS ، V.MAX ، LONG،LARG، BOIS) إلى مستطيل Independents، كما هو في الشكل التالي:

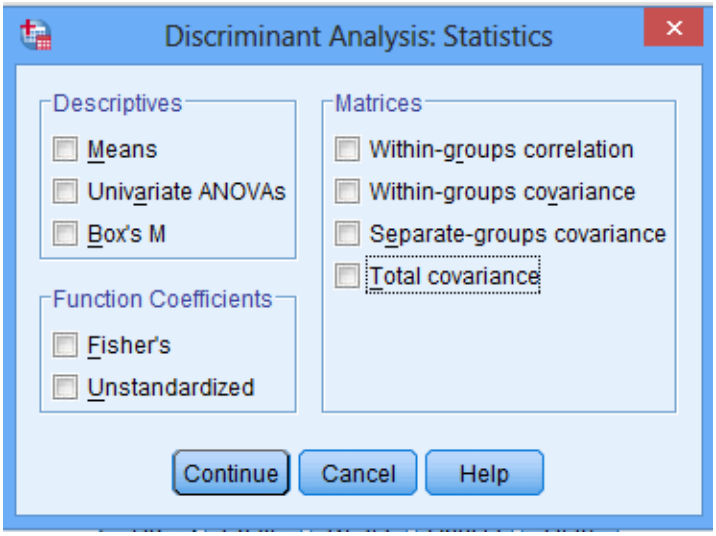


وهنا من الضروري تعريف قيم المتغير التابع وذلك من خلال الضغط على Define Range و تحديد القيمة الدنيا والقيمة العليا للمتغير التابع كما هو في الشكل التالي:



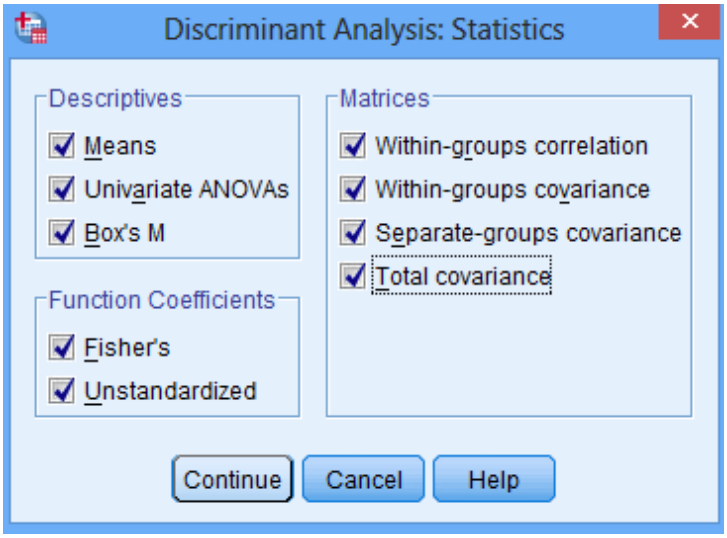
ثم نضغط متابعة

من خلال الضغط على statistiques نحصل على المربع الحواري التالي:

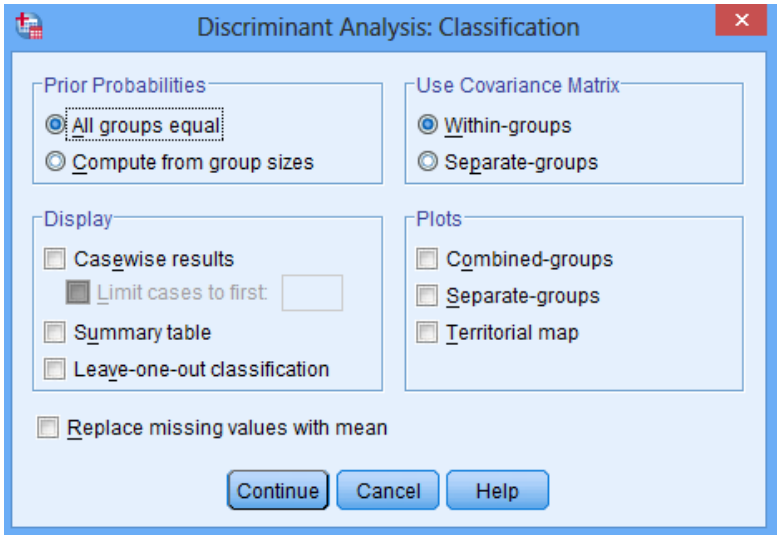


من خلال المربع الحواري السابق يمكن الحصول على احصائيات المجموعات (في هذا المثال مجموعتين فقط) المتمثلة في المتوسطات، و تحليل التباين الاحادي، ومصفوفة الارتباط داخل المجموعات ، مصفوفة التباين داخل المجموعات بالإضافة إلى معاملات الدالة التميزية المعيارية و غير المعيارية.

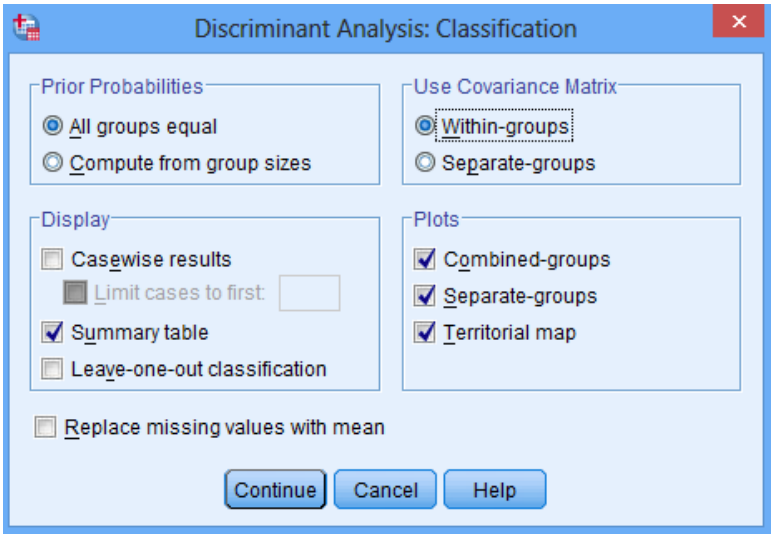
نقوم بالتأشير على الخيارات المطلوبة.



من خلال الضغط على Classify نحصل على المربع الحواري التالي:

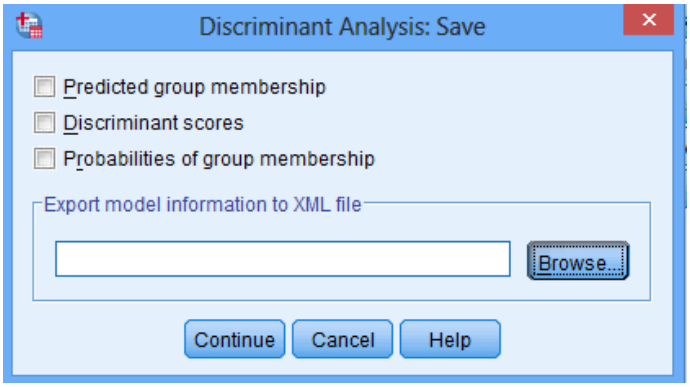


يمكن من خلال المربع الحواري السابق اختيار الاحتمال أن تكون جميع المجموعات متساوية في الاحتمال أو حساب الاحتمال من حجم المجموعات، كما يمكن أختيار التمثيل الباني المناسب، كذلك يطرح هذا المربع الحواري خيار تعويض القيم المفقودة بالمتوسط. نقوم بالتأشير على الخيرات التي من خلالها نجري عملية التحليل كالتالي:

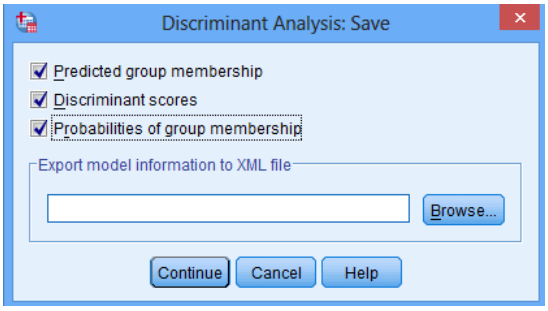


بعد تحديد الخيارات نضغط متابعة.

من خلال الضغط على Save نحصل على المربع الحواري التالي:

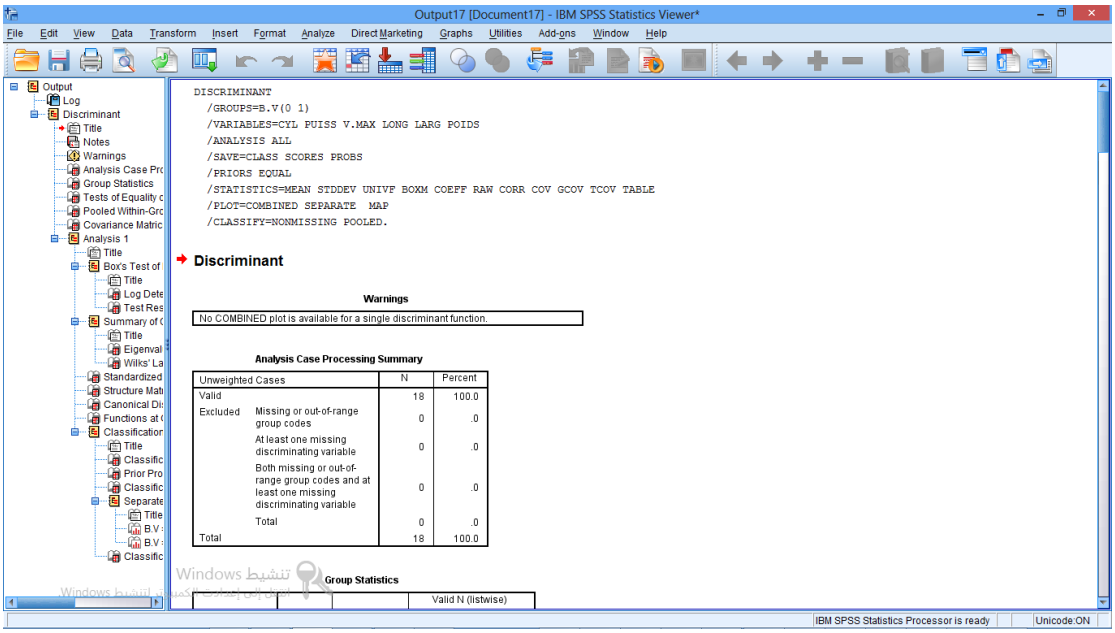


من خلال الربع الحوار السابق يمكن الحصول على المجموعة التي نسبت إليها كل مشاهدة، ويمكن الحصول على الاحداثيات التمييزية لكل مشاهدة، وكذلك احتمالية إنتماء كل مشاهدات إلى المجموعة ويمكن حفظ كل ذلك في قاعدة البيانات الاساسية. نقوم بالتأشير على الخيارات المطلوبة لعملية التحليل كالتالي:



بعد تحديد الخيارات نضغط متابعة.

وبعد الانتهاء من الخطوات السابقة نضغط على ok لنحصل على نافذة المخرجات التالية:



تحليل المخرجات:

الاحصائيات الوصفية للمجموعات:

يمثل الجدول الموالي الاحصائيات الوصفية للمجموعات(في مثالنا هذا مجموعتين)، وكذلك الاحصائيات الوصفية للمعطيات الاساسية الكلية.

Group Statistics

B.V		Mean	Std. Deviation	Valid N (listwise) Unweighted	Weighted
0	CYL	1420.88	215.994	8	8.000
	PUISS	68.63	11.771	8	8.000
	V.MAX	147.63	7.671	8	8.000
	LONG	426.38	23.096	8	8.000
	LARG	165.25	6.205	8	8.000
	POIDS	1019.88	116.454	8	8.000
1	CYL	1800.30	395.894	10	10.000
	PUISS	97.40	16.358	10	10.000
	V.MAX	166.80	7.146	10	10.000
	LONG	439.20	20.666	10	10.000
	LARG	167.80	4.492	10	10.000
	POIDS	1126.00	139.000	10	10.000
Total	CYL	1631.67	373.930	18	18.000
	PUISS	84.61	20.376	18	18.000
	V.MAX	158.28	12.140	18	18.000
	LONG	433.50	22.107	18	18.000
	LARG	166.67	5.314	18	18.000
	POIDS	1078.83	136.958	18	18.000

حيث نلاحظ من الجدول السابق أن المجموعة الاولى تتكون من ثمانية مشاهدات وهي السيارات ذات علبة سرعات يدوية، والمجموعة الثانية مكونة من عشر مشاهدات وهي السيارات ذات علبة السرعات الاوتوماتيكية.

اختبار متغيرات الدراسة:

نقوم هنا بإختبار متغيرات الدراسة وهل من بين هذه المتغيرات توجد متغيرات يتم على أساسها الفصل بين المجموعتين، أي هل هناك متغيرات يتم على اساسها تحديد أن السيارة تكون في المجموعة التي بها علبة سرعات يدوية أو أوتوماتيكية.

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
CYL	.731	5.893	1	16	.027
PUISS	.479	17.429	1	16	.001
V.MAX	.348	30.002	1	16	.000
LONG	.912	1.544	1	16	.232
LARG	.940	1.025	1	16	.326
POIDS	.843	2.979	1	16	.104

نقوم بأجراء الاختبارات التالية :

اختبار المتغير سعة الاسطوانة CYL وفق الفرضيات التالية

$$\begin{cases} H_0: \text{لا توجد فروق بين متوسط المجموعتين ترجع إلى متغير سعة الاسطوانة} \\ H_1: \text{توجد فروق بين متوسط المجموعتين ترجع إلى متغير سعة الاسطوانة} \end{cases}$$

من خلال اختبار فيشر الممثل في الجدول نلاحظ أن مستوى الدلالة للمتغير سعة الاسطوانة CYL هي 0.027 وهي أقل من 0.05، نرفض الفرضية الصفرية و نقبل الفرضية البديلة أي أنه توجد فروق بين متوسط المجموعتين ترجع إلى متغير سعة الاسطوانة، وبالتالي فإن المتغير سعة الاسطوانة CYL يمثل أساس في للفصل بين المجموعتين.

- اختبار المتغير قوة المحرك PUISS وفق الفرضيات التالية

$$\begin{cases} H_0: \text{لا توجد فروق بين متوسط المجموعتين ترجع إلى متغير قوة المحرك} \\ H_1: \text{توجد فروق بين متوسط المجموعتين ترجع إلى متغير قوة المحرك} \end{cases}$$

من خلال اختبار فيشر الممثل في الجدول نلاحظ أن مستوى الدلالة للمتغير قوة المحرك PUISS هي 0.001 وهي أقل من 0.05، وبالتالي نرفض الفرضية الصفرية و نقبل الفرضية البديلة أي أنه توجد فروق بين متوسط المجموعتين ترجع إلى قوة المحرك PUISS ، وبالتالي فإن المتغير قوة المحرك PUISS يمثل أساسا للفصل بين المجموعتين.

- اختبار المتغير السرعة القصوى V.MAXوفق الفرضيات التالية

$$\begin{cases} H_0: \text{لا توجد فروق بين متوسط المجموعتين ترجع إلى متغير السرعة القصوى} \\ H_1: \text{توجد فروق بين متوسط المجموعتين ترجع إلى متغير السرعة القصوى} \end{cases}$$

من خلال اختبار فيشر الممثل في الجدول نلاحظ أن مستوى الدلالة للمتغير السرعة القصوى V.MAX هي 0.000 وهي أقل من 0.05، وبالتالي نرفض الفرضية الصفرية و نقبل الفرضية البديلة أي أنه توجد فروق بين متوسط المجموعتين ترجع إلى السرعة القصوى V.MAX، وبالتالي فإن المتغير السرعة القصوى V.MAX يمثل أساسا للفصل بين المجموعتين.

- اختبار المتغير طول السيارة LONG وفق الفرضيات التالية

$$\begin{cases} H_0: \text{لا توجد فروق بين متوسط المجموعتين ترجع إلى متغير طول السيارة} \\ H_1: \text{توجد فروق بين متوسط المجموعتين ترجع إلى متغير طول السيارة} \end{cases}$$

من خلال اختبار فيشر الممثل في الجدول نلاحظ أن مستوى الدلالة للمتغير طول السيارة LONG هي 0.232 وهي أكبر من 0.05، وبالتالي نقبل الفرضية الصفرية أي أنه لا توجد فروق بين متوسط المجموعتين ترجع إلى طول السيارة LONG ، وبالتالي فإن المتغير طول السيارة LONG لا يمثل أساسا للفصل بين المجموعتين.

- اختبار المتغير عرض السيارة LARG وفق الفرضيات التالية

$$\begin{cases} H_0: \text{لا توجد فروق بين متوسط المجموعتين ترجع إلى متغير عرض السيارة} \\ H_1: \text{توجد فروق بين متوسط المجموعتين ترجع إلى متغير عرض السيارة} \end{cases}$$

من خلال اختبار فيشر الممثل في الجدول نلاحظ أن مستوى الدلالة للمتغير عرض السيارة LARG هي 0.326 وهي أكبر من 0.05، وبالتالي نقبل الفرضية

الصفرية أي أنه لا توجد فروق بين متوسط المجموعتين ترجع إلى عرض السيارة LARG، وبالتالي فإن المتغير عرض السيارة LARG لا يمثل أساسا للفصل بين المجموعتين.

- اختبار المتغير وزن السيارة POIDS وفق الفرضيات التالية

$$\begin{cases} H_0: \text{لا توجد فروق بين متوسط المجموعتين ترجع إلى متغير وزن السيارة} \\ H_1: \text{توجد فروق بين متوسط المجموعتين ترجع إلى متغير وزن السيارة} \end{cases}$$

من خلال اختبار فيشر الممثل في الجدول نلاحظ أن مستوى الدلالة للمتغير وزن السيارة POIDS هي 0.104 وهي أكبر من 0.05، وبالتالي نقبل الفرضية الصفرية أي أنه لا توجد فروق بين متوسط المجموعتين ترجع إلى وزن السيارة POIDS ، وبالتالي فإن المتغير وزن السيارة POIDS لا يمثل أساسا للفصل بين المجموعتين.

من خلال الاختبارات السابقة نستنتج أن المتغيرات الثلاثة سعة الاسطوانة وقوة المحرك والسرعة القصوى تشكل أساسا في الفصل بين المجموعتين.

تحليل الارتباط القانوني:

من الجدول التالي نلاحظ:

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	2.387[a]	100.0	100.0	.839

a. First 1 canonical discriminant functions were used in the analysis.

قيمة الارتباط القانوني بين المتغيرات التابعة والمتمثلة في هذا المثال في مجموعة السيارات ذات علبة السرعات اليدوية ومجموعة السيارات ذات

علبة السرعات الأوتوماتكية يساوي 0.839 وهو إرتباط قوي وموجب وبالتالي فإن المتغيرات المستقل تمثل أساس قوي في الفصل بين المجموعتين.

اختبار كاي مربع x^2:

يسمح هذا الاختبار بمعرفة ما إذا كان وجود المجموعتين مبررا أو لا، وهذا ما يوضحه الجدول التالي:

Wilks' Lambda

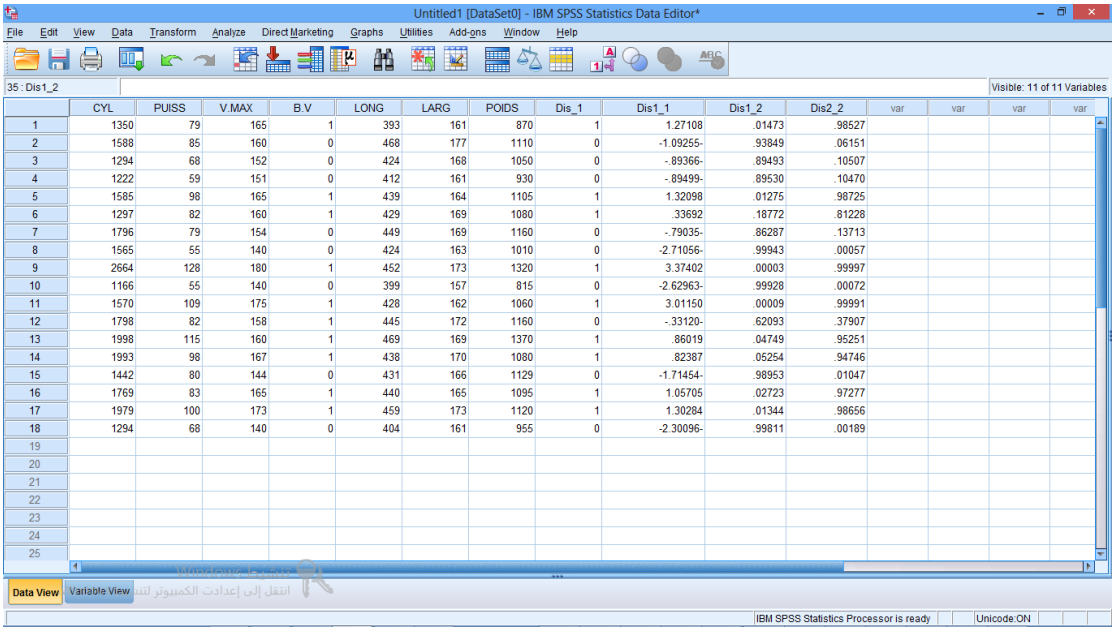
Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.295	15.857	6	.015

نشكل الفرضيات التالية:

$$\begin{cases} H_0: \text{لا يوجد اختلاف بين مراكز الفيئتين} \\ H_1: \text{يوجد اختلاف بين مراكز الفيئتين} \end{cases}$$

من خلال قيمة كاي مربع التي تساوي 15.857 وبمعنوية تساوي 0.015 وهي أقل من 0.05 نرفض الفرضية الصفرية ونقبل الفرضية البديلة أي أنه يوجد اختلاف بين مراكز الفئتين أو المجموعتين وبالتالي فإن للتحليل العاملي التمييزي في هذا التحليل اهمية.

بالعودة الى واجهة البرنامج نلاحظ أنه قد تم أدراج اعمدة جديدة كالتالي:



	CYL	PUISS	V.MAX	B.V	LONG	LARG	POIDS	Dis_1	Dis1_1	Dis1_2	Dis2_2
1	1350	79	165	1	393	161	870	1	1.27108	.01473	.98527
2	1588	85	160	0	468	177	1110	0	-1.09255-	.93849	.06151
3	1294	68	152	0	424	168	1050	0	-.89366-	.89493	.10507
4	1222	59	151	0	412	161	930	0	-.89499-	.89530	.10470
5	1585	98	165	1	439	164	1105	1	1.32098	.01275	.98725
6	1297	82	160	1	429	169	1080	1	.33692	.18772	.81228
7	1796	79	154	0	449	169	1160	0	-.79035-	.86287	.13713
8	1565	55	140	0	424	163	1010	0	-2.71056-	.99943	.00057
9	2664	128	180	1	452	173	1320	1	3.37402	.00003	.99997
10	1166	55	140	0	399	157	815	0	-2.62963-	.99928	.00072
11	1570	109	175	1	428	162	1060	1	3.01150	.00009	.99991
12	1798	82	158	1	445	172	1160	0	-.33120-	.62093	.37907
13	1998	115	160	1	469	169	1370	1	.86019	.04749	.95251
14	1993	98	167	1	438	170	1080	1	.82387	.05254	.94746
15	1442	80	144	0	431	166	1129	0	-1.71454-	.98953	.01047
16	1769	83	165	1	440	165	1095	1	1.05705	.02723	.97277
17	1979	100	173	1	459	173	1120	1	1.30284	.01344	.98656
18	1294	68	140	0	404	161	955	0	-2.30096-	.99811	.00189

ويمثل العمود Dis_1 أعادة التصنيف حسب دالة التمييز، حيث نلاحظ كمثال أن المشاهدة رقم 12 مصنفة ضمن مجموعة السيارات التي بها علبة سرعات أوتوماتيكية ويقترح حسب دالة التمييز أن تكون ضمن مجموعة السيارات بعلبة سرعات يدوية كما هو موضح كالتالي:

10	1166	55	140	0	399	157	815	0	-2.62963-	.99928	.00072
11	1570	109	175	1	428	162	1060	1	3.01150	.00009	.99991
12	1798	82	158	1	445	172	1160	0	-.33120-	.62093	.37907
13	1998	115	160	1	469	169	1370	1	.86019	.04749	.95251
14	1993	98	167	1	438	170	1080	1	.82387	.05254	.94746

ويمثل العمود Dis1_2 احتمال إنتماء المشاهدة إلى المجموع الاولى، والعمود Dis2_2 إنتماء المشاهدة الى المجموعة الثانية كما يلي:

	ARG	BOIS	Dis_1	Dis1_1	Dis1_2	Dis2_2	var
1	161	870	1	1.27108	.01473	.98527	
2	177	1110	0	-1.09255-	.93849	.06151	
3	168	1050	0	-.89366-	.89493	.10507	
4	161	930	0	-.89499-	.89530	.10470	
5	164	1105	1	1.32098	.01275	.98725	
6	169	1080	1	.33692	.18772	.81228	
7	169	1160	0	-.79035-	.86287	.13713	
8	163	1010	0	-2.71056-	.99943	.00057	
9	173	1320	1	3.37402	.00003	.99997	
10	157	815	0	-2.62963-	.99928	.00072	
11	162	1060	1	3.01150	.00009	.99991	
12	172	1160	0	-.33120-	.62093	.37907	
13	169	1370	1	.86019	.04749	.95251	
14	170	1080	1	.82387	.05254	.94746	
15	166	1129	0	-1.71454-	.98953	.01047	
16	165	1095	1	1.05705	.02723	.97277	
17	173	1120	1	1.30284	.01344	.98656	
18	161	955	0	-2.30096-	.99811	.00189	
19							
20							

حيث أن المشاهدة الاولى احتمال إنتمائها إلى المجموعة الاولى (علبة السرعات يدوية) هو 0.1473 واحتمال أن تكون من المجموعة الثانية (علبة السرعات أوتوماتيكية) هو 0.9852 وذلك ما يؤكده التصنيف الاساسي حيث أنها سيارة بعلبة سرعات أوتوماتيكية وأيضا تأكده دالة التمييز حيث تم إبقائها في نفس المجموعة. في حين نلاحظ أن المشاهدة رقم 12(سيارة بعلبة سرعات أوتوماتيكية) احتمال أن تنتمي للمجموعة الأولى(السيارات ذات علبة السرعات اليدوية) هو 0.62 وبالتالي بحسب دالة التمييز يقترح أن تصنف في هذه المجموعة (علبة السرعات اليدوية)، وهنا نستنتج أن هذه السيارة لم يكن من الضروري أن تنتج بعلبة سرعات أوتوماتيكية.

ويمثل الجدول التالي معاملات دالة التميز القانونية:

Standardized Canonical Discriminant Function Coefficients

	Function
	1
CYL	-.160-
PUISS	.015
V.MAX	1.135
LONG	-.457-
LARG	-.429-
BOIS	.745

ويمكن من جدول مصفوفة البناء التالي أن نستنتج معاملات الارتباط بين المتغير التابعة (المجموعتين) وبين المتغيرات المستقلة، ويعطى بالترتيب التنازلي كما هو موضح في الجدول الموالي:

Structure Matrix

	Function
	1
V.MAX	.886
PUISS	.676
CYL	.393
BOIS	.279
LONG	.201
LARG	.164

حيث أن الارتباط بين متغير السرعة القصوى V.MAX ومتغير المجموعات يساوي 0.886 وهو ارتباط قوي.

Classification Function Coefficients

	B.V	
	0	1
CYL	-.056-	-.057-
PUISS	-.785-	-.782-
V.MAX	2.275	2.726
LONG	.016	-.045-

LARG	8.178	7.941
BOIS	-.041-	-.024-
(Constant)	-760.393-	-781.300-

Fisher's linear discriminant functions

Classification Results[a]

			Predicted Group Membership		
		B.V	0	1	Total
Original	Count	0	8	0	8
		1	1	9	10
	%	0	100.0	.0	100.0
		1	10.0	90.0	100.0

a. 94.4% of original grouped cases correctly classified.

نلاحظ أن دالة التمييز نجحت في تصنيف 8 من8 مشاهدات بالنسبة للمجموعة الاولى بنفس التصنيف الأصلي أي 100%، بينما نجحت في تصنيف 9 مشاهدات من أصل 10 في المجموعة الثانية بنفس التصنيف الأصلي أي 90%، وعلى العموم فقد نجحت في تصنيف جميع المشاهدات بنسبة 94.4% وهي نسبة جد مقبولة.

9. مثال تطبيقي 2: التصنيف الى أكثرمن مجموعتين

نستعمل نفس معطيات المثال السابق مع إدراج متغير جديد يمثل عدد اسطوانات السيارة كما يوضحه الجدول التالي:

CYL	PUISS	V.MAX	LONG	LARG	BOIS	N. CYL
870	161	393	165	79	1350	2
1110	177	468	160	85	1588	2
1050	168	424	152	68	1294	2
930	161	412	151	59	1222	2
1105	164	439	165	98	1585	2
1080	169	429	160	82	1297	2
1160	169	449	154	79	1796	2
1010	163	424	140	55	1565	1
1320	173	452	180	128	2664	3
815	157	399	140	55	1166	1
1060	162	428	175	109	1570	3
1160	172	445	158	82	1798	2
1370	169	469	160	115	1998	2
1080	170	438	167	98	1993	2
1129	166	431	144	80	1442	1
1095	165	440	165	83	1769	2
1120	173	459	173	100	1979	3
955	161	404	140	68	1294	1

بعد ادخال المعطيات في برنامج SPSS وإتباع الخطوات نفس الخطوات في المثال السابق نحصل على المخرجات التالية:

الاحصائيات الوصفية للمجموعات:

يمثل الجدول الموالي الاحصائيات الوصفية للمجموعات(في مثالنا هذا ثلاثة مجموعات)، وكذلك الاحصائيات الوصفية للمعطيات الاساسية الكلية.

Group Statistics

N.CYL		Mean	Std. Deviation	Valid N (listwise)	
				Unweighted	Weighted
1	CYL	1366.75	173.742	4	4.000
	PUISS	64.50	12.014	4	4.000
	V.MAX	141.00	2.000	4	4.000
	LONG	414.50	15.416	4	4.000
	LARG	161.75	3.775	4	4.000
	BOIS	977.25	130.283	4	4.000
2	CYL	1608.18	284.516	11	11.000
	PUISS	84.36	15.141	11	11.000
	V.MAX	159.73	5.551	11	11.000
	LONG	436.91	22.363	11	11.000
	LARG	167.73	4.756	11	11.000
	BOIS	1091.82	128.302	11	11.000
3	CYL	2071.00	552.772	3	3.000
	PUISS	112.33	14.295	3	3.000
	V.MAX	176.00	3.606	3	3.000
	LONG	446.33	16.258	3	3.000
	LARG	169.33	6.351	3	3.000
	BOIS	1166.67	136.137	3	3.000
Total	CYL	1631.67	373.930	18	18.000
	PUISS	84.61	20.376	18	18.000
	V.MAX	158.28	12.140	18	18.000
	LONG	433.50	22.107	18	18.000
	LARG	166.67	5.314	18	18.000
	BOIS	1078.83	136.958	18	18.000

حيث نلاحظ من الجدول السابق أن المجموعة الاولى تتكون من اربعة مشاهدات وهي السيارات ذات 3 اسطوانات والمجموعة الثانية مكونة من 11 مشاهدات وهي السيارات بأربعة أسطوانات، والمجموعة الثالثة تتكون من 3 مشاهدات وهي السيارات ذات 6 اسطوانات.

اختبار متغيرات الدراسة:

نقوم هنا بإختبار متغيرات الدراسة وهل من بين هذه المتغيرات توجد متغيرات يتم على أساسها الفصل بين المجموعات الثلاثة، أي هل هناك متغيرات يتم على اساسها تحديد أن السيارة تكون في المجموعة الاولى أم الثانية أم المجموعة الثالثة.

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
CYL	.636	4.297	2	15	.033
PUISS	.444	9.390	2	15	.002
V.MAX	.138	46.784	2	15	.000
LONG	.751	2.482	2	15	.117
LARG	.728	2.798	2	15	.093
BOIS	.792	1.968	2	15	.174

نقوم بأجراء الاختبارات التالية :

اختبار المتغير سعة الاسطوانة CYL وفق الفرضيات التالية

$$\begin{cases} H_0: \text{لا توجد فروق بين متوسط المجموعات ترجع إلى متغير سعة الاسطوانة} \\ H_1: \text{توجد فروق بين متوسط المجموعات ترجع إلى متغير سعة الاسطوانة} \end{cases}$$

من خلال اختبار فيشر الممثل في الجدول نلاحظ أن مستوى الدلالة للمتغير سعة الاسطوانة CYL هي 0.033وهي أقل من 0.05، وبالتالي نرفض الفرضية

الصفرية و نقبل الفرضية البديلة أي أنه توجد فروق بين متوسط المجموعات مردها إلى متغير سعة الاسطوانة، وبالتالي فإن المتغير سعة الاسطوانة CYL يمثل أساس في للفصل بين المجموعتين.

- اختبار المتغير قوة المحرك PUISS وفق الفرضيات التالية

$$\begin{cases} H_0: \text{لا توجد فروق بين متوسط المجموعات ترجع إلى متغير قوة المحرك} \\ H_1: \text{توجد فروق بين متوسط المجموعات ترجع إلى متغير قوة المحرك} \end{cases}$$

من خلال اختبار فيشر الممثل في الجدول نلاحظ أن مستوى الدلالة للمتغير قوة المحرك PUISS هي 0.002 وهي أقل من 0.05، وبالتالي نرفض الفرضية الصفرية و نقبل الفرضية البديلة أي أنه توجد فروق بين متوسط المجموعات ترجع إلى قوة المحرك PUISS ، وبالتالي فإن المتغير قوة المحرك PUISS يمثل أساسا للفصل بين المجموعتين.

- اختبار المتغير السرعة القصوى V.MAXوفق الفرضيات التالية

$$\begin{cases} H_0: \text{لا توجد فروق بين متوسط المجموعات ترجع إلى متغير السرعة القصوى} \\ H_1: \text{توجد فروق بين متوسط المجموعات ترجع إلى متغير السرعة القصوى} \end{cases}$$

من خلال اختبار فيشر الممثل في الجدول نلاحظ أن مستوى الدلالة للمتغير السرعة القصوى V.MAX هي 0.000 وهي أقل من 0.05، وبالتالي نرفض الفرضية الصفرية و نقبل الفرضية البديلة أي أنه توجد فروق بين متوسط المجموعات ترجع إلى السرعة القصوى V.MAX، وبالتالي فإن المتغير السرعة القصوى V.MAX يمثل أساسا للفصل بين المجموعتين.

- اختبار المتغير طول السيارة LONG وفق الفرضيات التالية

$$\begin{cases} H_0: \text{لا توجد فروق بين متوسط المجموعات ترجع إلى متغير طول السيارة} \\ H_1: \text{توجد فروق بين متوسط المجموعات ترجع إلى متغير طول السيارة} \end{cases}$$

من خلال اختبار فيشر الممثل في الجدول نلاحظ أن مستوى الدلالة للمتغير طول السيارة LONG هي 0.117 وهي أكبر من 0.05، وبالتالي نقبل الفرضية الصفرية أي أنه لا توجد فروق بين متوسط المجموعات ترجع إلى طول السيارة LONG ، وبالتالي فإن المتغير طول السيارة LONG لا يمثل أساسا للفصل بين المجموعتين.

- اختبار المتغير عرض السيارة LARG وفق الفرضيات التالية

$$\begin{cases} H_0: \text{لا توجد فروق بين متوسط المجموعات ترجع إلى متغير عرض السيارة} \\ H_1: \text{توجد فروق بين متوسط المجموعات ترجع إلى متغير عرض السيارة} \end{cases}$$

من خلال اختبار فيشر الممثل في الجدول نلاحظ أن مستوى الدلالة للمتغير عرض السيارة LARG هي 0.093 وهي أكبر من 0.05، وبالتالي نقبل الفرضية الصفرية أي أنه لا توجد فروق بين متوسط المجموعات ترجع إلى عرض السيارة LARG، وبالتالي فإن المتغير عرض السيارة LARG لا يمثل أساسا للفصل بين المجموعتين.

- اختبار المتغير وزن السيارة POIDS وفق الفرضيات التالية

$$\begin{cases} H_0: \text{لا توجد فروق بين متوسط المجموعات ترجع إلى متغير وزن السيارة} \\ H_1: \text{توجد فروق بين متوسط المجموعات ترجع إلى متغير وزن السيارة} \end{cases}$$

من خلال اختبار فيشر الممثل في الجدول نلاحظ أن مستوى الدلالة للمتغير وزن السيارة POIDS هي 0.174 وهي أكبر من 0.05، وبالتالي نقبل الفرضية الصفرية أي أنه لا توجد فروق بين متوسط المجموعات ترجع إلى وزن السيارة

POIDS ، وبالتالي فإن المتغير وزن السيارة POIDS لا يمثل أساسا للفصل بين المجموعتين.

من خلال الاختبارات السابقة نستنتج أن المتغيرات الثلاثة سعة الاسطوانة وقوة المحرك والسرعة القصوى تشكل أساسا في الفصل بين المجموعات الثلاثة.

تحليل الارتباط القانوني:

من الجدول التالي نلاحظ:

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	10.794[a]	95.4	95.4	.957
2	.515[a]	4.6	100.0	.583

a. First 2 canonical discriminant functions were used in the analysis.

قيمة الارتباط القانوني بين المتغيرات التابعة والمتمثلة في هذا المثال في مجموعة السيارات ذات 3 أو 4 أو 6 اسطوانات والدالة التمييزية الاولى يساوي 0.957 وهو إرتباط قوي وموجب، بينما قيمة الارتباط القانوني بين المتغير التابعة والدالة التمييزية الثانية هو 0.583 وهو ظعيف مقارنة بالأول.

اختبار كاي مربع x^2:

يسمح هذا الاختبار بمعرفة ما إذا كان وجود المجموعات مبررا أم لا، وهذا ما يوضحه الجدول التالي:

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 2	.056	36.041	12	.000
2	.660	5.196	5	.392

ومن خلال قيمة Wilks' Lambda لدالة الاولى والتي تساوي 0.056 وهي تؤول لصفر والدالة الثانية والتي تساوي 0.660 فإننا نختار المعادلة والتمييزية الاولى.

نشكل الفرضيات التالية:

$$\begin{cases} H_0: \text{لا يوجد اختلاف بين مراكز الفيئتين} \\ H_1: \text{يوجد اختلاف بين مراكز الفيئتين} \end{cases}$$

من خلال قيمة كاي مربع التي تساوي 36.04 وبمعنوية تساوي 0.000 وهي أقل من 0.05 نرفض الفرضية الصفرية ونقبل الفرضية البديلة أي أنه يوجد اختلاف بين مراكز المجموعات وبالتالي فإن للتحليل العاملي التمييزي في هذا التحليل له اهمية.

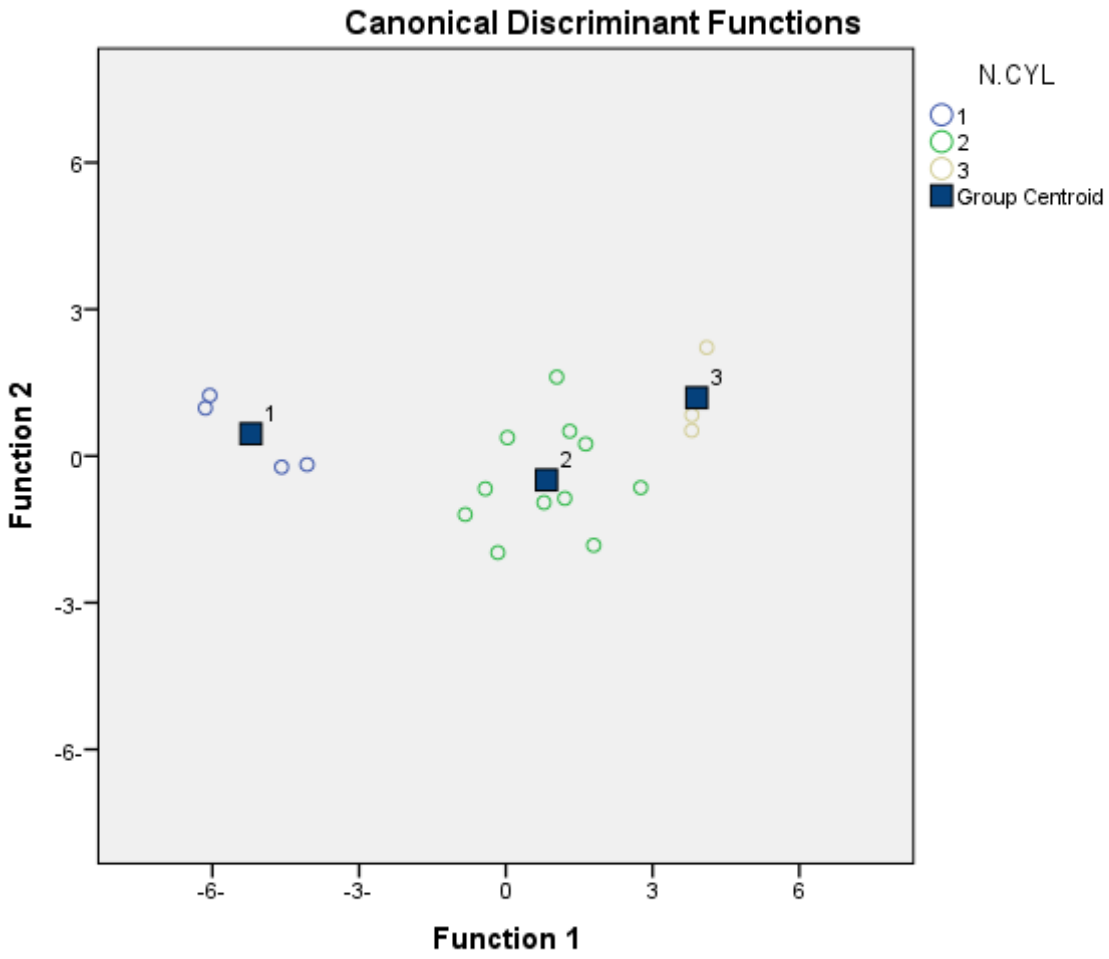
والجدول التالي يقدم معاملات دوال التصنيف لكل مجموعة

Classification Function Coefficients

	N.CYL		
	1	2	3
CYL	-.208-	-.226-	-.228-
PUISS	-7.539-	-8.334-	-8.498-
V.MAX	23.405	25.925	26.940
LONG	.270	.290	.317
LARG	7.730	7.710	7.627
BOIS	.697	.785	.798
(Constant)	-2287.601-	-2676.784-	-2840.835-

Fisher's linear discriminant functions

ويمكن توضيح طريق التصنيف بيانيا من خلال الأشكال التالية:



يوضح الشكل السابق تصنيف المشاهدات إلى ثلاثة مجموعات كما يمثل المربع مركز كل مجموعة.

ونلاحظ أنه تم تصنيف كل المشاهدات وهذا ما يؤكده الجدول التالي:

Classification Results[a]

			Predicted Group Membership			
		N.CYL	1	2	3	Total
Original	Count	1	4	0	0	4
		2	0	11	0	11
		3	0	0	3	3
	%	1	100.0	.0	.0	100.0
		2	.0	100.0	.0	100.0
		3	.0	.0	100.0	100.0

a. 100.0% of original grouped cases correctly classified.

من خلال الجدول السابق نلاحظ أن الدوال التمييزية قد نجحت 100% في تصنيف المشاهدات وذلك أن تصنيف كل المشاهدات متطابق مع التصنيف الاصلي.

المراجع:

1- أبو علام رجاء محمود،" التحليل الإحصائي للبيانات باستخدام برنامج SPSS"، الطبعة الأولى،دار النشر للجامعات، مصر، 2002.

2- الراوي زياد رشاد،"طرق التحليل الاحصائي المتعدد المتغيرات "، منشورات المعهد العربي للتدريب والبحوث الاحصائية الأردن ، 2017.

3- تيغزة أمحمد بوزيان، " التحليل العاملي الاستكشافي والتوكيدي "، الطبعة الأولى دار المسيرة، عمان، الأردن، 2012.

4- جودة محفوظ،" التحليل الاحصائي المتقدم باستخدام spss" ، الطبعة الثانية، دار وائل للنشر عمان، الأردن، 2009.

5- صواليلي صدر الدين، " تحليل المعطيات "، دار هومة،الجزائر،2011.

6- André Bouchier, l'analyse en composantes principales ACP, INRA Montpellier, France, juillier 2005.

7- J.Krzanowski, C.Marriott, Multivariate Analysis: Classification, Covariance Structures and Repeated Measurements, pt.2.1ed edition. Hodder education publisher; 1998.

8- J.F.Hair, et al, Multivariate data analysis, A Global persoective, 7th edition. Peorson Prentice Hall. New Jersey, 2010.

9- R.Feinstein, Multivariable analysis: an introduction,Yale University press-London,1996.

10- R.Tibshirani, G.Walther, T. hastie , Estimating the number of clusters in data set via the gap statistic, Stanford University – U.S.A, 2000.

11- S. Cristins, Surveying Clusters: A comprehensive bibliometric Account, Master' s Dissertation- Faculty of Economics - University of Porto. Portugal, 2007.

12- W. Hardle, L.Simar, Applied Multivariate Statistical analysis, 2ed edition, VerlagBerlin, 2007.