الفصل الثالث : طريقة تحليل المركبات الاساسية "Analyse en composante principale (ACP)"

تعتبر طريقة تحليل المركبات الاساسية ACP من اشهر الطرق واكثرها استعمالا في مختلف البحوث والدراسات العلمية في مجالات عدة (الاقتصاد، الطب، البيولوجيا، التغذية .....)، فهي تعد أداة قوية للغاية لضغط وتوثيق البيانات، وهي مفيدة جدا عندما يكون هناك كم هائل من المعطيات الكمية القابلة للمعالجة لغرض التفسير. لذلك سوف نتطرق في هذا الفصل الى مجموعة من المفاهيم الاساسية المساعدة علا فهم هذه الطريقة , كيفية عمل الطريقة وكذا دراسة حالة مع التطبيق .

III-1 تقديم طريقة المركبات الأساسية "Présentation de le méthode ACP"

يتم تقديم تقديم طريقة المركبات الأساسية عن طريق التطرق الى مجموعة من المفاهيم الاساسية المتعلقة بها وكذا طبيعة بيانات الدراسة.

III-1-1 التعريف بالطريقة

هي طريقة احصائية **وصفية استكشافية ، متعددة الابعاد** تقوم على تخفيض كم هائل من المتغيرات الاصلية (أي الابعاد الاصلية) الى عدد اقل من المتغيرات التمثيلية (variables représentative) (أي المركبات الاساسية او العوامل الرئيسية) باستعمال مجموعة من المفاهيم ذات الطابع الاحصائي الرياضي كالوسط الحسابي ، الانحراف المعياري، معامل الارتباط الخطي، الزوايا التمثيلية، الجبر الخطي....الخ من اجل اعطاء نتائج متنوعة في شكل جداول بسيطة ومخططات تعرف بالمخرجات (outputs). هذه النتائج تكون قادرة على تفسير الظاهرة .

كما تعرف على انها تحليل عاملي، بمعنى انها تقوم بإنتاج عوامل (محاور رئيسية = Axes principaux) وهي تجميعات خطية من المتغيرات الاولية ، الهرمية والمستقلة عن بعضها البعض، تسمى هذه العوامل احيانا ب «الابعاد الكامنة» لأنها تعبر عن عمليات او سيرورات عامة توجه توزيع العديد من الظواهر التي تجد نفسها مرتبطة ببعضها البعض.(Béguin et pu main 2000) .

- **وصفية (Descriptive):** وهذا لكونها تقوم بتحويل المعطيات الاولية الى مخططات واشكال (mapping) وبالتالي تمنحها الطابع الوصفي والمرئي.
- **استكشافية (Exploratoire):** وهذا يتعارض مع الاستدلال ، لأن الهدف هنا هو اظهار الفجوات بين المتغيرات وتشكيل مجموعات من الافراد تشبه بعضها البعض من جهة .كما انها لا تستند على الاحتمالات ولا على الفرضيات او أي نماذج مسبقة من جهة أخرى .ففي حالة الاستبيانات فهي لا تقوم على دراسة سؤال معين : كعلاقة الدخل بالاستهلاك وانما تقوم بدراسة مجموع الاسئلة لا على التعيين.
- **متعدد الابعاد (Multidimensionnelle):** وهذا يتعارض مع احادية البعد فباعتبار انه كل متغيرة تعبر عن بعد معين (متغيرة واحدة تمثل بمحور ، متغيرتين تمثلان بمعلم متعامد متجانس ، 3 متغيرات فأكثر تمثل في فضاء من 3 محاور فأكثر). فهذه الطريقة دائما ما تفترض انه لدينا العديد من المتغيرات على الافراد المعنيين .ففي حالة الاستبيانات مثلا فإنها لا تدرس اجابة واحدة وانما جميع الاجابات المشكلة له باعتبار كل سؤال هو متغير في حد ذاته.

III-1-2 بعض المفاهيم الاساسية :

– **المركبات الرئيسية 'Composantes principales"**

– في طريقة تحليل المركبات الاساسية فإننا نبحث اساسا عن خطوط لتعديل الغيوم النقطية بشكل افضل في مساحة الشعاعية للمتغيرات والافراد، وفقا لمعيار المربعات الصغرى. وبالتالي يتم استخراج المكونات او العوامل الرئيسية التي تعظم من مجموع المربعات في الاسقاطات المتعامدة. لذلك يتم العثور على فراغ شعاعي فرعي ذو بعد اقل ، يمثل المساحة الشعاعية الاصلية . على الرغم من انه يتم استخراج العامل الاول لالتقاط اقصى قدر من التباين، الا انه نادرا ما يمكن التقاط التباين العام. وبالتالي، فان ما تبقى يعاده عامل اخر ثاني، وهكذا ومع ذلك، لا يمكن ان يتجاوز عدد العوامل المستخرجة عدد المتغيرات الاصلية.

– **المتغيرات والمفردات النشطة والاضافية (Variables et individus Actives et supplémentaires)**

ميزة فريدة اخرى لهذه الطريقة وهي انه يمكن تحديد المتغيرات والمفردات النشطة والاضافية. يتم استخدام المتغيرات والملاحظات النشطة بالتوازي مع المكونات الرئيسية، حيث يمكن بعد ذلك عرض المتغيرات والملاحظات الاضافية على المستوى العاملي المحسوب من المتغيرات والمفردات النشطة. هذه المزايا تجعل من طريقة ال ACP اداة قوية لتصنيف وغربلة البيانات.

- **نتائج الطريقة " Résultats attendus"**

تأخذ النتائج شكلين: اوراق البيانات والرسوم البيانية. بينما يمكن استخدام اوراق البيانات في تفسير النتائج ،و توفر الرسوم البيانية المرتبطة مساعدة بصرية لتصنيف المتغيرات والافراد.

– **نتائج عددية:** تنتج ال ACP مجموعة متنوعة من النتائج ، مثل احداثيات المتغيرات والملاحظات ، ومساهمات المتغيرات والملاحظات ، ونتائج العوامل ، ومعاملات نتائج العوامل، ومربعات جيب التمام، والقيم الذاتية ، والاحصاءات الوصفية.

- **الرسومات:** تذكر ان الهدف الرئيسي من ACP هو استعادة مساحة عامل اصغر يمكن ان تسقط عليها النقاط الاصلية المتغيرات او المفردات، بحيث يمكن الكشف عن البنية الاساسية للبيانات و لتسهيل الامر، يمكن انتاج مخططات ثنائية الابعاد للإحداثيات. هذا الخيار متاح للمتغيرات والمفردات. تتوفر ا رسومات ثنائية الابعاد وثلاثية الابعاد للإحصاءات الوصفية. تتوفر العديد من خيارات التمثيل لكل رسم بياني.

III- 1-3 طبيعة بيانات الدراسة " Type de données"

تهتم طريقة المركبات الاساسية بالجداول المستطيلة (rectangulaire) وهي جداول ذات مدخلين (à double entrée Tableau) والتي يطلق عليها عادة اسم المصفوفات بحيث تكون المفردات الاحصائية على الاسطر والمتغيرات على الاعمدة. تتميز هذه المتغيرات بكونها من طبيعة كمية.

– ليكن الجدول التالي X والذي يحوي «k» متغيرو «I» مفردة احصائية :

$$X_{(I,K)} = \begin{array}{c|cccc|} & 1 & 2 & k & K \\ \hline 1 & x_{11} & x_{12} & x_{1k} & x_{1K} \\ 2 & x_{21} & x_{22} & x_{2k} & x_{2K} \\ i & x_{i1} & x_{i2} & x_{ik} & x_{iK} \\ I & x_{I2} & x_{I2} & x_{Ik} & x_{IK} \\ \hline \end{array}$$

حيث:

x_{IK} : القيمة التي تأخذها المفردة الاحصائية I من اجل المتغيرة K.

I : مؤشر السطر، حيث أن i هو عدد المفردات الاحصائية.

K : مؤشر العمود ، حيث ان k هو عدد المتغيرات.

من خلال الجدول X يمكن حساب الوسط الحسابي الموافق لكل متغيرة والذي يعطى بالعلاقة:

$\overline{X_k} = \frac{\sum_{i=1}^{I} x_{ik}}{I}$, كما يمكن حساب الانحراف المعياري بالعلاقة: $\delta_k = \frac{\sum_{i=1}^{I}(x_{ik}-\bar{x}_k)^2}{I}$

و عليه يمكن استنتاج ما يعرف بالمفردة الاحصائية الوسط (L'individu moyenne) والتي تكون احداثياتها هي متوسطات جميع المتغيرات المتواجدة على مستوى الجدول والتي تعطى بالشكل : $I_G\ (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k, \ldots, \bar{x}_K\ )$

مثال : ليكن الجدول التالي والذي يمثل نقاط مجموعة من الطلبة خلال اربعة مقاييس وهي الرياضيات , الفيزياء , الفرنسية والانجليزية :

Etudiant	MATH	PHYS	FRAN	ENGL
Jean	6	6	5	5,5
Alan	8	8	8	8
Anni	6	7	11	9,5
Moni	14,5	14,5	15,5	15
Didi	14	14	12	12,5
Andr	11	10	5,5	7
Pier	5,5	7	14	11,5

Brig	13	12,5	8,5	9,5
Evel	9	9,5	12,5	12

من خلال الجدول نلاحظ ان المفردات هنا يعبر عنها بالطلبة (9 طلبة) اما المتغيرات فهي المقاييس (اربعة مقاييس).

من خلال الجدول يمكن حساب جميع المتوسطات الخاصة بالمقاييس وكذا الانحرافات المعيارية والموضحة في الجدول التالي وهذا على مستوى العمودين الخامس والسادس:

Statistiques descriptives :					
Variable	**Observations**	**Minimum**	**Maximum**	**Moyenne**	**Ecart-type**
MATH	9	5,500	14,500	9,667	3,579
PHYS	9	6,000	14,500	9,833	3,172
FRAN	9	5,000	15,500	10,222	3,684
ENGL	9	5,500	15,000	10,056	2,984

وعليه فان المفردة الاحصائية الوسط ذات الاحداثيات الاتية : (10,056 10,222 9,833 9,667)

4-1-III انواع طرق تحليل المركبات الاساسية "Type de l'ACP"

بالاعتماد على العمليتين الرياضيتين السابقتين واللتان تتمثلان في حساب المتوسط الحسابي والانحراف المعياري يمكن التميز بين نوعين من طرق تحليل المركبات الاساسية حيث نجد : تحليل مركبات اساسية معيارية (ACP nornée) وطريقة تحليل مركبات اساسية غير معيارية او كما تعرف بالمركزة (ACP non nornée ou bien centrée). العملية التي بموجبها يتم تحديد احدى النوعين من طرق تحليل المركبات الاساسية تعرف بالمعيارية (centrage et réduction) حيث ان:

- **المركزة** «**Centrage**» وتعرف ايضا بعبارة التوسيط لارتباطها بمؤشر الوسط الحسابي , ويعني بها نزع المتوسط الحسابي من اجمالي البيانات وبالتالي بدل النظر الى المعطيات الاولية على انها " x_{ij} " فإننا ننظر اليها على اساس " $x_{ij} - \bar{x}_{ij}$ ". حيث تسمح هذه الخطوة بتغيير مبدأ المعلم من المبدأ الاصلي الذي احداثياته (0 0 ...0...0) الى مركز الثقل الذي يرمز له بالرمز (G) «centre de gravite» والذي احداثياته هو متوسطات المتغيرات.

هذه العملية (توسيط البيانات) لا تؤثر على الشكل العام لسحابة النقاط الخاصة ببيانات الدراسة ولهذا نقوم بها دائما أي في جميع الحالات وفي شتى المجالات.

- **التخفيض** « **Réduction** » : وتعرف ايضا بعبارة التوحيد لارتباطها بمؤشر الانحراف المعياري كما انها تكون مصاحبة دائما لعملية التوسيط وتعني بها قسمة مجموع البيانات المركزة على الانحراف المعياري أي : $x_{ik} \rightarrow \frac{x_{ik} - \bar{x}_k}{\delta_k}$

وبالتالي بدل النظر الى ان الجدول علا اساس '' x_{ij} '' فإننا ننظر اليه على اساس " $\frac{x_{ik}-\bar{x}_k}{\delta_k}$ " .هذه العملية تكون مصاحبة دائما للعملية الاولى على خلاف الاولى فهي يمكن ان تستغني عن التخفيض.

وعليه تقتصر «ACP» الممركزة على عملية المركزة فقط اما ال «ACP» المعيارية فتتعدى ذلك الى عملية التخفيض وبالتالي تشمل العمليتين معا. كما تقدم طريقة ACP خيارات لحساب المكونات الرئيسية اما مع مصفوفة الارتباط او مصفوفة التباين. فكون المعطيات تستلزم التعيير او المركزة فان تحليل هذه المعطيات يتم تحقيقه بالاعتماد على مصفوفة الارتباط « matrice de corrélation » وهذا بالنسبة للنوع الثاني « ACP nornée » او مصفوفة التباين والتغاير « matrice de variance covariance » وهذا في حالة ال «ACP centrée».

ملاحظة: عند تحليل مصفوفة التغاير ، سيتأثر التحليل والعوامل المحسوبة منه بالاختلافات في مجموع تباينات المتغيرات النشطة. لذلك ، سيكون التحليل القائم على مصفوفة التغاير مناسبا فقط اذا كان الكشف عن هذه الاختلافات وثيق الصلة بنوع الدراسة التي تجريها. ومع ذلك، في معظم الحالات، لا تكون هذه الاختلافات مثيرة للاهتمام ببساطة لأنه ترجع الى الاختلافات في وحدات القياس. على سبيل المثال، يمكن استخدام قياسين مختلفين لدرجات الحرارة، درجة مئوية وفهرنهايت ، لقياس متغيرين. من الواضح ان ابراز مثل هذا الاختلاف في التحليل سيكون مضيعة للجهد. لذلك فان الحاجة في مثل هذه الحالات هي تحويل البيانات لإزالة تأثير المقياس. تحسب المكونات الرئيسية لاستعادة اكبر قد من التباين على العامل الاول ، ثم على العامل الثاني ، وهكذا هذا يعني انه اذا كانت المتغيرات المعنية غير متجانسة فيما يتعلق بالتباينات ، فان العامل الاكثر اهمية أي الاول غالبا ما يعكس اكبر تباين لبعض المتغيرات. ومع ذلك ، يجب اجراء التحليل باستخدام مصفوفة الارتباط ، والتي ليست سوى مصفوفة التغاير للمتغيرات الموحدة.

- مزايا تطبيق هذه الخطوة

بها مزايا عدة يمكن ان نذكر منها :

- جعل وحدات القياس متماثلة بمعنى اخر فإنها تلغي اثر الاختلاف في وحدات القياس على مجموع المعلومة، لذلك فان تطبيقها حتمي في حالة اختلاف وحدات القياس.
- جعل مركز المعلم من مركز الثقل للسحابة والذي احداثياته من احداثيات المفردة الوسط وليس المبدأ الاصلي . مما يمكن من الحصول على تمثيل جيد لمجموع المفردات أي تشتت السحابة.
- اعطاء نفس الوزن لجميع المتغيرات او جعلها من نفس الاهمية على حد سواء.
- جعل الجداول الابتدائية اكثر وضوح واكثر مقروئية.

تعتبر خطوة ضرورية واساسية تنعكس بالإيجاب على باقي خطوات «ACP»

III-1- 5 الاختبارات القبلية اللازمة قبل عمل ACP (Tests à priori)

هناك بعض الاختبارات القبلية والضرورية (à priori) قبل عمل تحليل للمركبات الاساسية والتي تسمح بتطبيق هذه الطريقة مع ضمان صحة النتائج المحصلة ومن اشهرها : مؤشر كيمو Indice de Kimo ، اختبار بارتلات Bartlett.

الفكرة وراء هذه الاختبارات هو الاجابة على التساؤل التالي : هل من الممكن الحصول على ملخص جيد لإجمالي المعلومة ؟ او بعبارة اخرى هل البيانات التي بحوزتنا صالحة او مساعدة على عمل تحليل للمركبات الاساسية؟

في الواقع يمكننا اعتبار طريقة « ACP » بمثابة ضغط للمعلومات. وهذا ممكن فقط اذا كان هناك بعض التكرار في البيانات . فمن الناحية النظرية كما في التطبيقية ، يتم تطبيق ACP اعتمادا على مصفوفة الارتباط او التغاير غير اننا سوف نركز على مصفوفة الارتباط في الواقع. فهناك حالات عديدة يكون فيها استخدام ACP غير مبرر على الاطلاق لهذا يجب تحقق بعض المعايير المسبقة.

- اختبار « La sphéricité de Bartlett »

حيث يقوم هذا الاختبار اساسا على الفروض التالية : كفرضية صفرية (H0) فان مصفوفة الارتباط تختلف الى حد كبير عن مصفوفة الوحدة مقابل الفرضية البديلة (H1) مصفوفة الارتباط تقترب كثيرا من مصفوفة الوحدة.

وعليه فرفض الفرضية الصفرية يمكن من ضغط المعلومات الى حد كبير، وبذلك يمكن من انشاء ملخص جيد لعموم المعلومة. وهنا يمكن ان نطرح التساؤل: الى أي حد يمكن ان نجد عدد اقل من العوامل عند رفض الفرضية الصفرية ؟. غير ان هذا لا يعني اننا سنجد معلومات مهمة في طريقة ACP المطبقة.

فانطلاقا من مصفوفة الارتباط فإننا يمكن ان نقوم بحساب المحدد (déterminant de la matrice de corrélation) والذي يرمز له بالرمز R تحت الفرض الصفري : " $H0 : \ |R| = 1$ " فإذا كان هناك علاقات خطية متداخلة مثالية ، فسيكون $|R| = 1$.أي جميع المتغيرات مرتبطة ارتباطا تاما وهذا يعني انه هناك محور وحيد او مركبة اساسية واحدة، مما لا يمكن من عمل تحليل للمركبات الاساسية أي استحالة تطبيق الطريقة بحجة اننا امام ما يعرف بالمصفوفة الشاذة وبالتالي لا يمكن عمل مقلوب للمصفوفة. اما اذا كان جميع الارتباطات تقترب الى 0 هذا يعني ان $|R| = 1$ اي اننا امام مصفوفة الوحدة وهذا يعني لا يمكن من ايجاد مركبات جامعة لإجمالي المعلومة وبالتالي لا حاجة لعمل ACP.

ولهذا فإن أحسن حالة يجب ان تكون عليها مصفوفة الارتباط هي أن تجمع الارتباطات بين القوة والاتزان (متوسطة). في الحالة العامة يصعب تحديد قيم العتبة وانما عادة ما يوخذ $|R|$ اقل من 0.00001 .

- مؤشر" kimo "

ويعرف بمؤشر كفاية المعاينة ، ويشترك هذا المؤشر في نفس فكرة الاختبار السابق تحت الفرض الصفري H0 : هل هناك امكانية ايجاد ملخص جيد لإجمالي البيانات.

غير انه يختلف في استراتيجيته حيث : نقطة الانطلاق هي دائما مصفوفة الارتباط. حيث يجب ان تكون المتغيرات مرتبطة جدا او بنسبة اقل من الاساس. الارتباط الخام بين متغيرتين يتأثر ب (K-2) المتغيرات الاخرى. لهذا يستعمل الاختبار الارتباط الجزئي من اجل تقدير العلاقة الصافية بين متغيرتين وهذا بإزالة اثر المتغيرات الاخرى ، ثم يقوم بمقارنة الارتباط الخام مع الارتباط الجزئي. اذا كان هذا الاخير اقل

بكثير(بالقيمة المطلقة) فهذا يعني ان العلاقة تحددها المتغيرات الاخرى، وهذا يتوافق مع فكرة التكرارات للمعلومة، مما يعني امكانية اعداد ملخص جيد لإجمالي البيانات .

ويكفي ان يكون هذا المؤشر >0.5 أي (50 %) لكي تكون مصفوفة الارتباط ملائمة لعمل المركبات الاساسية.

مثال:

اعتمادا على الجدول الابتدائي قم بتحقيق الاختبارين ل بارتلات وكذا كيمو.

Test de sphéricité de Bartlett	
Khi² (Valeur observée)	66,059
Khi² (Valeur critique)	12,592
DDL	6
p-value	< 0,0001
alpha	0,05

نلاحظ من خلال الجدول ان قيمة P المحسوبة أقل بكثير من قيمة مستوى الدلالة أي قيمة P الجدولية (0,05> 0,0001) وعليه فإننا امام رفض الفرض الصفري الدال على عدم وجود ارتباط خطي معنوي مختلف عن الصفر بين متغيرات الدراسة مما يدل على وجود على الاقل احد الارتباطات بين المتغيرات معنوي.

Mesure de précision de l'échantillonnage de Kaiser-Meyer-Olkin	
MATH	0,440
PHYS	0,548
FRAN	0,405
ENGL	0,563
KMO	0,489

من خلال النتيجتين يمكن ان نستنتج انه يمكن القيام بعمل ملخص فعال لمجموع المعلومة .

III- 1 -6 مبدأ طريقة "ACP"

غالبا ما يتم جمع البيانات حول المتغيرات التي ترتبط فقط فيما بينها، ولكنها ايضا تكون متعددة جدا . هذا يجعل تفسيرها والكشف عن هيكلها صعبا جدا. بتحويل المتغيرات الاصلية الى عدد اقل من المتغيرات غير المترابطة . يجعل تحليل المركبات الرئيسية (ACP) هاتين المهمتين أسهل.

تميل طريقة ال ACP الى تحقيق المبادئ التالية:

1- تخفيض عدد المتغيرات الى عدد اقل من المتغيرات التمثيلية الغير مترابطة وبالتالي تعديل الغيوم النقطية بشكل افضل في المساحة الشعاعية ذات الابعاد المنخفضة والتي تمثل المساحة الشعاعية

الاصلية. وهذا بواسطة العوامل او المركبات الرئيسية التي تعظم من حجم التباين الكلي المفسر. فعلى الرغم من انه يتم استخراج العامل الاول الذي يلتقط اقصى قدر ممكن من التباين، الا انه نادرا ما يمكن من التقاط التباين العام. وبالتالي فان ما تبقى يعاده عامل اخر (ثاني) وهكذا.

2- تصنيف المتغيرات والافراد فبالإضافة الى تقليل الابعاد للمساحة الاصلية للمتغيرات ، فانه يمكن ايضا استخدام ACP كتقنية تصنيف ، وبالتالي يمكن تسليط الضوء على العلاقات بين المتغيرات والافراد.

3- تسمح ACP بتحليل البيانات التي تم جعلها حول المتغيرات غير المتجانسة من خلال تقديم خيار لتحليل مصفوفة التغاير او مصفوفة الارتباط.

III-2 البحث عن المركبات الاساسية والمحاور العاملية

ويكون هذا من خلال دراسة جدول البيانات وهذا لاستخراج بعض الحسابات الرياضية بالاعتماد على الجبر الخطي والتي سوف تساعد في الحصول على المركبات الاساسية وكذا اختيار المحاور وهذا من اجل في الاخير اعطاء تخطيط بياني لمجموع المعلومة في صورة خلال بعدين او ثلاث كأقصى تقدير.

III-2-1 دراسة البيانات (دراسة جدول المعطيات)

سوف نفترض في هذا الجزء من الفصل الخاص بطريقة تحليل المركبات الاساسية اننا سوف نقوم بتعيير البيانات او توحيدها أي "ACP nornée" .

دراسة هذا النوع من الجداول الاحصائية يجرنا الى دراسة كل من المفردات الاحصائية وكذا المتغيرات. هذه الدراسة سوف تمكننا في الاخير من الوصول الى تمثيل هذه المعطيات من خلال مجموع الابعاد المحصل عليه ، وبالتالي الحصول على الصورة الاقل تشوها بالنسبة للصورة الحقيقية التي تكون خلال مجموع الابعاد الحقيقية. حيث يمكن اعتبار هذا الجدول حزمة من الاسطر ، كما يمكن اعتبارها حزمة من الاعمدة.(un packet de ligne et de colonne).

أ- دراسة المفردات الاحصائية "étude des individus"

باعتبار الجدول هو حزمة من الاسطر بعبارة اخرى اننا سوف ننظر اليه نظرة افقية أي من وجهة نظر المفردات الاحصائية. فإنه يجب ان نتطرق الى مفهوم التشابه «ressemblance».

نقول ان مفردتين احصائيتين متشابهتين من وجهة نظر المتغيرات اذا كانتا قريبتين من بعضهما البعض من وجهة نظر المتغيرات وهذا هندسيا ، بمعنى آخر اذا كانت لهما نفس الخصائص. هذا التشابه يعبر عنه رياضيا بالمسافة « distance » . أي كلما قلت المسافة بين مفردتين كلما زاد التشابه والعكس صحيح.

اعتمادا على مفهوم التشابه يمكن عمل فحص التشابه "un bilan de ressemblance" لمجموع البيانات مما يمكن من الحصول على تصنيف او تقسيم " partitions" لمجموع المفردات من وجهة نظر المتغيرات وبالتالي نتحصل على مجموعات من الافراد والتي تكون متجانسة داخل المجموعة الواحدة ، ومتنافرة فيما

بينهما « Des groups homogene et des groups hétérogen». في حالة الاستبيانات فانه يمكن تجميع الاسئلة التي تعالج امرا مشتركا او تصب في باب واحد والذي يعرف بالبعد «Dimension » .

تعتبر دراسة المفردات الاحصائية احد مفاتيح دراسة التغير الكلي. ولقد تم تقدير المسافة بين المفردتين عن طريق المسافة الإقليدية او المسافة الإقليدية مربعة والتي تعطى بالعلاقة التالية:

$$d^2(i,l) = \sum_{k=1}^{K} (x_{ik} - x_{lk})^2 \qquad ''\text{Merci Pythagore}''$$

وهي مجموعة مربعات انحرافات القيم التي تاخذها كل مفردة احصائية من اجل كل متغيرة. دراسة المفردات الاحصائية يودنا الى تعريف:

- **الافراد النشطين والاضافيين "Individus actifs-supplémentaires"**

كما هو الحال مع المتغيرات، يمكن اعتبار بعض الافراد كأفراد نشطين والبعض الاخر كأفراد اضافيين. يمكن القيام بذلك باستخدام متغير (نوعي او مجموعة) واستخدام احدى قيمها مثل الرمز الذي يحدد الافراد النشطين. يتم التعامل مع بقية الافراد كأفراد اضافيين. مرة اخرى، كما هو الحال مع المتغيرات النشطة، يشارك فقط الافراد النشطون في حساب المكونات الرئيسية. يتم بعد ذلك اسقاط الافراد الاضافيين في الفضاء الفرعي للمتجه الناتج عن العوامل المحسوبة من المتغيرات النشطة والافراد .قد تنطبق الاستنتاجات التي تستند الى العوامل المحسوبة على افراد اضافيين، حتى لو لم يكونوا مشتركين في الحسابات.

<u>مثال:</u> اعتمادا على الجدول السابق (X)، قم بحساب المسافة بين مجموع الافراد (anni، jean،alan) . ماذا تلاحظ؟

الاجابة: باستعمال علاقة بطاغور نجد:

$$d^2(Jean, Alan) = \sum_{k=1}^{4} (x_{Je\,k} - x_{Al\,k})^2 = (6-8)^2 + (6-8)^2 + (5-8)^2 + (5{,}5-8)^2 = 23{,}5$$

$$d^2(Jean, Anni) = \sum_{k=1}^{4} (x_{Je\,k} - x_{An\,k})^2 = (6-6)^2 + (6-7)^2 + (5-11)^2 + (5{,}5-9{,}5)^2 = 57{,}2$$

- نلاحظ ان jean اقرب ل alan من anni .

وعليه يمكن القول : ان دراسة المفردات الاحصائية تعني :

- دراسة التقارب والتباعد بين مختلف المفردات مثنى مثنى وهذا من وجهة نظر المتغيرات.
- تصنيف هذه المفردات في شكل مجموعات متجانسة واخرى متنافرة.

- المفتاح الاول من مفاتيح دراسة التغير الكلي.

ب- دراسة المتغيرات " étude des variables "

راينا سابقا احد مفاتيح دراسة التغير الكلي، والان سوف ننتقل الى المفتاح الثاني الذي لا يقل اهمية عن سابقه والذي يساهم في دراسته هذا التغير و هي المتغيرات.

باعتبار الجدول هو مجموعة من الاعمدة أي مجموعة من المتغيرات فهذا يؤدي بنا الى التطرق الى مفهوم اخر لا يختلف كثيرا عن التشابه وهو الارتباط ''corrélation'' غير انه يكمن حسابه رياضيا بصفة مباشرة ، ومن بين اهم مؤشرات الارتباط الاكثر شيوعا واستعمالا هو معامل الارتباط الخطي البسيط ل « Pearson».

حيث نقول ان المتغيرتين x،y مرتبطان اذا كانت قيمة معامل الارتباط الخطي تقترب الى الواحد بالقيمة المطلقة، بعبارة اخرى ان المتغيرات مرتبطة اذا كانت تعطى نفس الخصائص للمفردات الاحصائية او انها تحمل بالتقريب نفس المعلومة. وعليه باعتبار الارتباط يمكن عمل ما يعرف بفحص الارتباط « un bilan des corrélation »، حيث يتم تجميع المتغيرات التي تكون مرتبطة فيما بينها. فاذا كان لدينا العديد من المتغيرات فإننا نأخذها جميعها مثنى مثنى، ثم نقوم بحساب جميع معاملات الارتباط الخطي مما يمكننا من الحصول على ما يعرف بمصفوفة الارتباط "matrice de corrélation " والتي تأخذ الشكل التالي:

$$r_{(K\,K)} = \begin{array}{c|cccc|} & 1 & 2 & k & K \\ \hline 1 & 1 & r_{12} & r_{1k} & r_{1K} \\ 2 & r_{21} & 1 & r_{2k} & r_{2K} \\ K & r_{k1} & r_{k2} & 1 & r_{kK} \\ K & r_{l2} & r_{l2} & r_{lk} & 1 \\ \hline \end{array}$$

حيث:

- r_{kK} : هو معامل الارتباط الخطي بين المتغيرة k و R في آن واحد.
- $r_{k2} = r_{2K}$

تتميز هذه المصفوفة بمجموعة من الخصائص :

ت- مصفوفة الارتباط هي مصفوفة مربعة « $r_{(K\,K)}$ » ، فإذا كان لدينا في دراسة ما ، 100 متغيرة هذا يعني اننا سوف نتحصل على مصفوفة ارتباط ذات الابعاد (100 X 100) مما يعني صعوبة دراستها بالعين المجردة لذلك نلجأ إلى إهمال المتغيرات التي لها ارتباطات ضعيفة بباقي المتغيرات وتشكيل مجموعة المتغيرات التي ترتبط ارتباطا قويا فيما بينها.

ث- مصفوفة قطرية حيث ان جميع القيم القطرية = 1 وهذا يدل على الارتباط الذاتي لكل متغيرة.
ج- مصفوفة متجانسة بحيث ان الجزء العلوي أي ما فوق القطر مطابق للجزء السفلي ما تحت القطر وبالتالي يكفي الاحتفاظ بجزء فقط لتسهيل الفحص الكلي للمصفوفة.

حيث يعطى الارتباط بين متغيرتين x و y بالعلاقة التالية:

$$r_{xy} = \frac{\sum_{k=1}^{K}(x_k - \bar{x})(y_k - \bar{y})}{\delta_x \delta_y}$$

بالاعتماد على مفهوم الارتباط يمكن ان نعرف :

- **المتغيرات النشطة و المتغيرات الاضافية " Variables actifs-supplémentaires "**

لنفترض انك مهتم بمتغيرات قليلة فقط لاستخدامها في التحليل، والحفاظ على الاخرين لغرض وحيد هو التفسير، على الرغم من ان المجموعتين تشيران الى نفس السياق وترتبطان في الواقع. مثل التحليل الذي يتم في ACP ، يمكنك التعامل مع المتغيرات كمتغيرات نشطة واضافية. سيتم بعد ذلك حساب المركبات (العوامل) الرئيسية باستخدام المتغيرات النشطة فقط. يمكن اسقاط المتغيرات الاضافية في وقت لاحق في الفضاء الفرعي المتجه الناتج عن العوامل، وبالتالي يتم حسابها، ويمكن اعطاء استنتاجات حول هذه المتغيرات ، حتى لو لم تشارك في التحليل. ومع ذلك، من المهم ان نأخذ في الاعتبار ان مثل هذا التقسيم للمتغيرات ليس الزاميا، وفي حالة معينة يمكن التعامل مع جميع المتغيرات كمتغيرات نشطة.

مثال: في المثال السابق تحصلنا على مصفوفة الارتباط التالية:

	MATH	PHYS	FRAN	ENGL
MATH	**1**	**0,983**	0,227	0,508
PHYS	**0,983**	**1**	0,397	0,652
FRAN	0,227	0,397	**1**	**0,951**
ENGL	0,508	0,652	**0,951**	**1**

من خلال المصفوفة نلاحظ ان هناك ارتباط خطي قوي بين المقاييس الادبية (الفرنسية والانجليزية) كما ان هناك ارتباط خطي موجب بين المقاييس العلمية (الرياضيات والفيزياء) مما يمكن من الحصول على مجموعتين من المتغيرات: المجموعة الاولى (math ،phy) وتعبر عن الاختبار العلمي و المجموعة الثانية (Ang ،Fran) وتعبر عن الاختبار الادبي او اللغوي , كما تجدر الاشارة الا ان الارتباطات الاخرى جاءت ضعيفة عدا التي تتعلق بال(الفيزياء , الانجليزية) و ال(الرياضيات , الانجليزية) فكانت متوسطة.

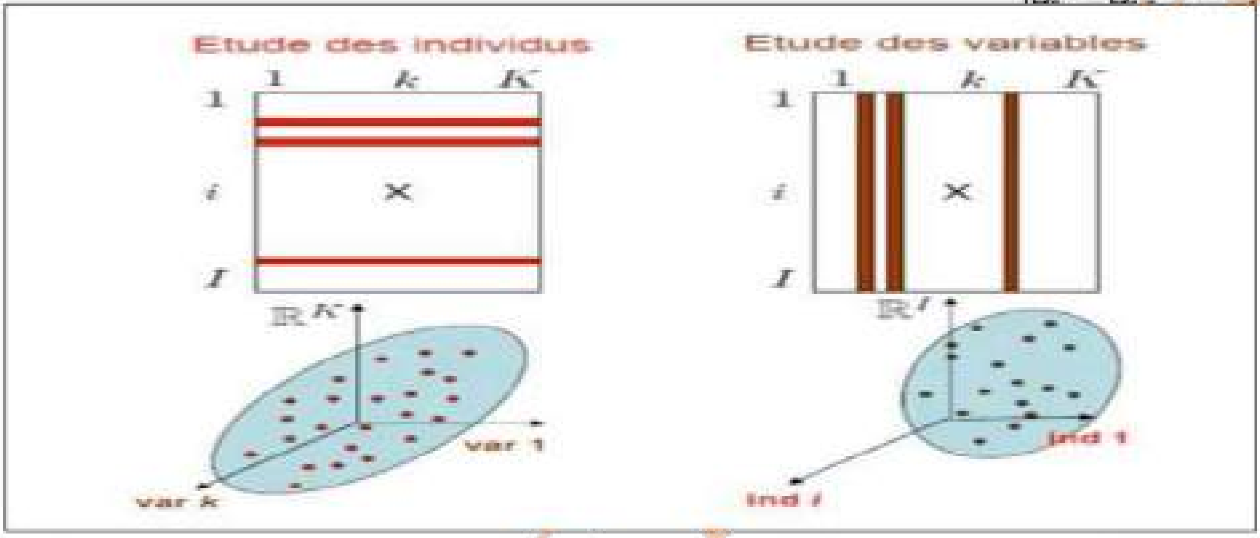
وعليه يمكن القول : ان دراسة المتغيرات تعني :

ح- دراسة الارتباط بين مختلف المتغيرات مثنى مثنى وهذا بالاعتماد على مؤشر الارتباط الخطي البسيط.
خ- تجميع هذه المتغيرات في شكل مجموعات بالنسبة للمحاور.

د- المفتاح الثاني من مفاتيح دراسة التغير الكلي.

III-2-2 تمثيل المعطيات "représentations des données"

كما راينا سابقا ان جدول المعطيات يمكن اعتباره على انه مجموعة من المفردات الاحصائية في تقاطع مع المتغيرات وعليه يمكن التمييز بين نوعين من سحابة النقاط :سحابة نقاط خاصة بالمفردات الاحصائية وهذا في فضاء من k بعد أي « R_K » . وكذا سحابة نقاط خاصة بالمتغيرات وتكون في فضاء « R_I » من I بعد وفقا للتمثيلين التالين:



على اليمين سحابة نقاط الخاصة بالمتغيرات "Nuage de points des variables " أما على اليسار فتجد سحابة النقاط الخاصة بالمفردات ."Nuage de points des individus "

تهدف طريقة تحليل المركبات الاساسية الى الوصول الى تمثيل المفردات الاحصائية من خلال السحابتين 1 و2 وهذا عن طريق التقليل في عدد الابعاد والذي ينتج بتقليص عدد المتغيرات المفسرة للظاهرة وذلك من k متغير الى q متغير حيث ان q اقل بكثير من k (K العدد الابتدائي لمتغيرات الدراسة). وبهذا نكون قد مررنا الى مصفوفة اخرى تحتفظ بنفس عدد المفردات ، وبأقل عدد من المتغيرات ، مع الحفاظ ما امكن من اجمالي المعلومة. هذا الانتقال قد يؤدي بدوره الى حدوث تشوه على مستوى الصورة اثناء اسقاطها على مستوى ذو بعدين او ثلاث.

مع العلم ان مبدأ المعلم الاصلي (origine) لا يعطينا صورة جيدة خصوصا بالنسبة لتشتت المفردات على مستوى سحابة النقاط وهذا لعدم اخذه بعين الاعتبار للمسافة المقدرة بين كل المفردات مثنى مثنى. فاننا نقوم بتغييره الى مبدأ اخر يعبر عنه بمركز الثقل «centre de gravité».

- **تغيير المبدأ " changement de l'origine "**

اعتمادا على المفردة الوسط التي احداثياتها مختلف متوسطات متغيرات الدراسة فانه يمكن للحصول على المركز الجديد O_g وهو مركز الثقل والذي يعطي بالعلاقة التالية:

$$O_G = \acute{X}D1_n = (\frac{1}{I}\sum_{i=1}^{I} x_{i1}, \dots, \frac{1}{I}\sum_{i=1}^{I} x_{ik}, \dots, \frac{1}{I}\sum_{i=1}^{I} x_{iK})$$

حيث:

$\acute{X}_{(K\,I)}$: هي المصفوفة المنقولة من المصفوفة الابتدائية X المتواجدة على مستوى الجدول الابتدائي.

$D_{(I\,I)}$: مصفوفة الثقل للافراد وهي مصفوفة قطرية ، قيم قطرها مساوية الى $1/I$ وهذا بإعطاء نفس الثقل لكل مفرده ، مع العلم انه يمكن اعطاء اثقال مختلفة لمجموع المفردات وذلك حسب الظاهرة

1_I: الشعاع الذاتي.

مثال تطبيقي : اعتماد على المثال السابق ، احسب المبدأ الجديد والممثل بمركز الثقل ؟

$$O_G = (\tfrac{1}{9}\textstyle\sum_{i=1}^{9} x_{i1}, \tfrac{1}{9}\sum_{i=1}^{9} x_{i2}, \tfrac{1}{9}\sum_{i=1}^{9} x_{i3}, \tfrac{1}{9}\sum_{i=1}^{9} x_{i4}) = (9{,}67 \quad 9{,}83 \quad 10{,}22 \quad 10{,}06)$$

نلاحظ ان مركز ثقل السحابة هو متوسط المتغيرات الاربعة.

- **التباين الكلي و جودة الصور "L'inertie et qualité de l'image"**

أحد الاسئلة الهامة التي يجب معالجتها في طريقة تحليل المركبات الاساسية هو عدد المركبات الاساسية التي يمكن ان تمثل بشكل مثالي المجموعة الكاملة من النقاط (المتغيرات والافراد).نظرا لان كل قيمة ذاتية من مصفوفة الارتباط او التغاير تمثل التباين الموضح او المفسر بواسطة المركبة الاساسية، يمكن ان تعطي نسبة من التباين التراكمي المفسرة بعدد معين من المركبات حقيقية الصورة المحصلة . وعليه تعرف جودة الصورة من خلال قيمة التباين الكلي المفسر على مستوى المركبات الرئيسية التي سوف يتم اختيارها من خلال ما يساهم به كل محور (مركبة) في التباين الكلي والمعروف بالمساهمة النسبية للمحاور. ويمكن القول ان الصورة جيدة او ذو جودة اذا :

ج- اذا تمكنت بإخلاص ان تحفظ الشكل العام للسحابة.

ح- تمكنت من اعطاء أحسن تمثيل للتشتت(المفردات) من خلال اظهار او عدم تشويه مختلف المسافات بين مختلف الافراد والتي تحسب بداية خلال k بعد ثم تنتقل بعدها الى q بعد.

خ- التباين الكلي : او كما يعرف بالجمود الكلي «Inertie total» وهي عبارة عن مجموعة التباينات التي يفسرها كل بعد والتي تعطى التباين الاجمالي خلال مجموع الابعاد المختلفة. حيث يمكن حسابه بالعلاقة التالية:

$$I_G = \sum_{k=1}^{K} var\,(V_k) \qquad / \qquad V = \acute{X}_c D\, X = \frac{1}{I}\acute{X}_c\, X_c$$

حيث:

$V_{(K.K)}$: هي مصفوفة التباين والتباين المشترك (التغاير).

$X_{c_{(I.K)}}$: هي المصفوفة الممركزة للمصفوفة X .

مثال: من خلال المثال السابق احسب قيمة التباين الكلي ($Inertie_{\ -total}$)

$$X_c = \begin{pmatrix} -3,67 & -3,83 & -5,22 & -4,64 \\ 1,67 & -1,83 & -2,22 & -2,06 \\ \vdots & \vdots & \vdots & \vdots \\ 1,33 & 0,17 & 4,72 & -3,06 \\ \vdots & \vdots & \vdots & \vdots \\ -0,67 & 0,33 & 2,28 & 1,94 \end{pmatrix}, \ \hat{X}_c \begin{pmatrix} -3,67 & 1,67 & \dots & 1,33 & \cdots & -0,67 \\ -3,83 & -1,83 & \cdots & 0,17 & \cdots & 0,33 \\ -5,22 & -2,22 & \cdots & 4,72 & \cdots & 2,28 \\ -4,64 & -2,06 & \cdots & -3,06 & \cdots & 1,94 \end{pmatrix}$$

وعليه تكون مصفوفة التباين والتغاير على النحو الاتي :

$$V= \begin{array}{c} \text{MATH} \\ \text{PHYS} \\ \text{FRAN} \\ \text{ENGL} \end{array} \begin{pmatrix} 11,39 & 9,92 & 2,66 & 4,82 \\ 9,92 & 8,94 & 4,12 & 5,48 \\ 2,66 & 4,12 & 12,06 & 9,29 \\ 4,82 & 5,48 & 9,29 & 7,91 \end{pmatrix}$$

الحل : من خلال المصفوفة فان التباين الكلي مساوي الى:

$$I_{total} = 11.39 + 8.94 + 12.06 + 7.91 = 40.3$$

نلاحظ ان القيم التي تم جمعها للحصول على التباين الكلي المفسر تمثل قيم القطر الرئيسي للمصفوفة «V». وعليه فان التباين الكلي المفسر لأربعة ابعاد في المثال هو 40.3 .

وسوف نتطرق فيما بعد الى كيفية توزيعه بين مختلف المحاور الرئيسية وهذا بعد التطرق الى طريقة اختيارها.

- **اختيار المحاور "Le choix des axes factoriels"**

تلعب القيم الذاتية «Valeurs propres» لمصفوفة الارتباط او التباين والتغاير للمتغيرات دورا مهما في تحديد المركبات الاساسية. بالإضافة الى تحديد احداثيات العامل للمتغيرات والافراد على تلك المركبات. تحديد المركبات الاساسية او اختيار المحاور يمكن من تحديد التصنيف الذي يسمح من خلاله اقتراح تقليل لحجم المساحة الاصلية للمتغيرات والافراد ، دون فقدان الكثير من المعلومات. استنادا الى القيم الذاتية ، يمكن استخدام العديد من المعايير لتحديد العدد المثالي للعوامل، فنظرا لان مجموع هذه القيم مساوي الى عدد المتغيرات ، فان متوسط القيم الذاتية هو 1 بمعنى اخر انه سوف نقوم بوضع عتبة للتباين والتي عادة ما يتم اختيارها بالاقتراب من الواحد والتي تضمن الضياع القليل للمعلومة المفسرة خلال المستوى التي يتم خلاله رسم مجموع المفردات الاحصائية .

من اجل حساب القيم الذاتية نقوم بتطبيق العلاقة التالية: $\Delta = |V - \lambda I_k| = 0$

حيث:

Δ : هو المحدد.

V : مصفوفة التباين والتغاير.

λ : القيمة الذاتية.

I_k : المصفوفة الذاتية.

مثال: بالاعتماد على مصفوفة التباين والتغاير قم بحساب القيم الذاتية .

الحل:

$$\Delta = \left|\begin{pmatrix} 11,39 & 9,92 & 2,66 & 4,82 \\ 9,92 & 8,94 & 4,12 & 5,48 \\ 2,66 & 4,12 & 12,06 & 9,29 \\ 4,82 & 5,48 & 9,29 & 7,91 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\right| = \mathbf{0}$$

$$= \begin{vmatrix} 11,39-\lambda & . & . & . \\ . & 8,94-\lambda & . & . \\ . & . & 12,06-\lambda & . \\ . & . & . & 7,91-\lambda \end{vmatrix}$$

ومنه نحصل على مجموع القيم الذاتية الاربعة: $\lambda_1 = 28,23$; $\lambda_2 = 12,03$; $\lambda_3 = 0,03$, $\lambda_4 = 0,01$

وعليه وجب الاختيار على المحورين الاول والثاني فقط والموافقات للقيميين الذاتيتين λ_1 و λ_2 باعتبارهما اكبر من الواحد ، اما المحورين 3 و 4 فسوف نقوم بإهمالهما وهذا لان λ_3 و $\lambda_4 < 1$

.من خلال ما سبق نلاحظ ان:

$$I_G = \sum_{K=1}^{K} \lambda_k$$

$$I_G = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 28,23 + 12,03 + 0,03 + 0,01 = 40,3$$

- **المساهمة النسبية للمحاور:**

لمعرفة مساهمة كل محور في التباين الكلي يكفي ان نطبق العلاقة التالية :

$$\mathrm{Cr}\,(\Delta_k / I_G) = \frac{\lambda_k}{\sum_{k=1}^{K} \lambda_k}$$

حيث:

λ_k : هو القيمة الذاتية k.

Δ_k : هو المحور الموافق للقيمة الذاتية λ_k .

مثال : ماهي مساهمة المحاور الاربعة في تشكيل التباين الكلي ؟

	F1	F2	F3	F4
Valeur propre	2,876	1,120	0,004	0,001
Variabilité (%)	71,892	27,992	0,089	0,026

% cumulé	71,892	99,884	99,974	100,000

بتطبيق العلاقة السابقة نجد أن مساهمة المحور الاول:

$$Cr(\Delta_1/I_G) = \frac{\lambda_1}{\lambda_1+\lambda_2+\lambda_3+\lambda_4} = \frac{28,23}{40,3} = 0,7189 = 71,89\ \%$$

– مساهمة المحاور الاخرى هي على الترتيب : 0,27992 ,0,0026 ,0,0089 .

نلاحظ ان اكبر مساهمة هي مساهمة المحور الاول ب 70 % اما الباقي فقد كانت من نصيب المحور الثاني نسبة 30 % من اجمالي التباين اما المحورين 3 و 4 فلم يساهما باي نسبة وهذا يعكس القيمتين الذاتين المقابلتين لهما.

– **التمثيل البياني للمفردات خلال المستوى الاول(بمساعدة المحورين الاول والثاني):**

ويتعلق الامر بالرسومات المرتبطة بالأفراد ، حيث يتم تحديد زوج من محاور العوامل من مجموعة المحاور الاجمالية والذي يساوي عدد المتغيرات ، ثم يتم اسقاط نقاط الفضاء العاملي المولدة من قبل الافراد على ذلك المستوى .حيث يمكن استخدام هذه الرسوم البيانية لتصنيف المفردات الى فئات .هذا التصنيف يتم حسب احداثياتهم المقابلة على محاور العوامل. لهذا وجب حساب احداثيات المفردات على مستوى محاور العوامل المختارة والتي تعطى بالعبارة التالية :

$$y_i = \hat{A}\,.\hat{X}_c$$

حيث:

$\hat{A}$: تمثل المصفوفة المنقولة لمصفوفة الاشعة الذاتية المقابلة للقيم الذاتية λ_1 و λ_2 . هذا اذا اقتصرنا فقط على المحورين العاملين الاول والثاني فقط. أي (F_2 , F_1) حيث يمكن ان نلجأ الى المستوى الثاني والمشكل من (F_1 , F_3) وهنا نأخذ الشعاع الذاتي المقابل ل λ_3 بعين الاعتبار.

$\hat{X}_c$: هي المصفوفة المنقولة للمصفوفة الممركزة للمصفوفة X .

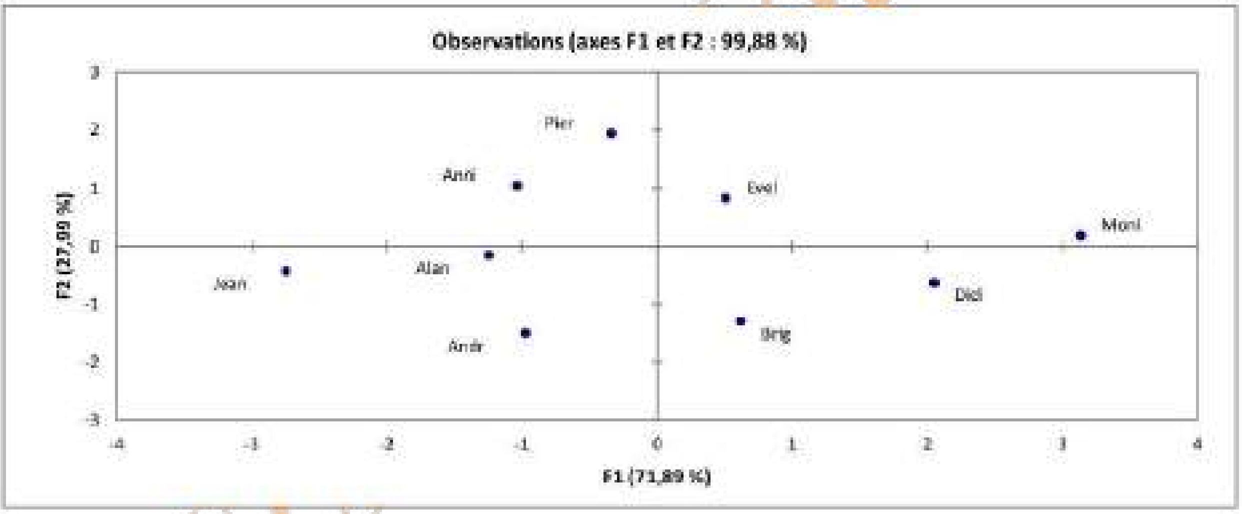
مثال : باستخدام العبارة السابقة واعتمادا على المثال السابق اعط ، مركبات الافراد على مستوى المستوى العاملي الاول ؟ ثم قم بإسقاطها على المحورين (F_2 , F_1) .

Coordonnées des observations		
Individus	**F1**	**F2**
Jean	-2,743	-0,427
Alan	-1,241	-0,153
Anni	-1,031	1,049
Moni	3,138	0,186

Didi	2,051	-0,628
Andr	-0,971	-1,498
Pier	-0,335	1,937
Brig	0,620	-1,291
Evel	0,510	0,824

– **الاحداثيات العاملية للأفراد:**

الاحداثيات العاملية للأفراد ليست ارتباطات كما هو الحال بالنسبة للمتغيرات (عندما نقوم بتحليل مصفوفة الارتباطات) هي ببساطة النقاط التي يتم اسقاطها على الخطوط التي تعبر سحابة النقاط متعددة الابعاد(في ما معنى المربعات الصغرى) للمساحة التي تنتجها اشعة الافراد. الافراد الذين يساهمون اكثر في عامل معين (على عكس اولئك الذين لديهم متوسط مساهمة) هم الاكثر تمثيلا للمفهوم الذي يمثل العامل المشكل. على سبيل المثال، اذا تمكنا بشكل واضح من تصنيف عامل في التحليل كعامل «بناء الجسم»، فان الافراد الذين يساهمون اكثر في هذا العامل سيكونون من لديهم اعلى بناء ومن هم اقل بناء مقارنة للأفراد من البناء المتوسط.



ان هذه العملية تسبقها عملية اساسية وهي دراسة المتغيرات وتمثيلها في المحاور العاملية الجديدة ، وهذه الادلة لا يمكن تفسير المفردات الاحصائية وتصنيفها الا باعتماد على المتغيرات فهي التي تعطيها مجموع الخصائص الاساسية التي على اساسها يتم ضبط مفهوم التشابه بالإضافة للمسافة الإقليدية ، وعليه يمكن الوصول الى تشكيلات متجانسة من الافراد ، والذي يمثل الغرض النهائي من طريقة ACP .

– **تمثيل المتغيرات "Représentation des variables"**

والمقصود به هو اسقاط المتغيرات على المحاور العاملية المختارة ، وذلك من خلال احداثياتها المتمثلة في شكل ارتباطات هذه المتغيرات وتلك المحاور اذا كان التحليل الحالي يعتمد على مصفوفة الارتباط ، ويكون هذا في تمثيل رسومي ثنائي الابعاد ، داخل دائرة تسمى بدائرة الارتباط (Cercle de corrélation) مع زوج من محاور العوامل الخاصة . توفر المتغيرات، عند عرضها في هذه الدائرة الكثير من المعلومات. على

سبيل المثال ، كلما كانت النقطة المتغيرة اقرب الى محيط الدائرة ، كلما كان تمثيلها جيدا وكان ارتباطها مع محور العامل المقابل افضل وهكذا يتم تحديد المتغيرات المرتبطة بمحور معين ، وبالتالي المشاركة في المعلومات المقدمة من المتغيرات التي توضح العامل المحدد .

وبالمثل : فان موقع الاحداثيات العاملية لمتغير على طول محاور العوامل بتصنيفها في فئة او اخرى .على سبيل المثال ، بدءا من المحور العاملي الاول يمكن تصنيف المتغيرات الى فئتين ، اعتمادا على ما اذا كانت الاحداثيات العاملية للمتغيرات مرتبة على جانب او اخر من المحور العاملي ، وبعبارة اخرى يتم تصنيف المتغيرات وفقا لإشارة احداثيات العامل .

لتحقيق هذا وجب اولا حساب احداثيات هذه المتغيرات على مستوى المحاور المختارة والغني يكون وفقا للعلاقة التالية:

$$R_{ZV} = \sqrt{\lambda_j}\, I_j\, A\, \frac{1}{\delta_j}\, I_j$$

حيث ان:

R_{ZV} : هي مصفوفة احداثيات المتغيرات على المحاور .

$\sqrt{\lambda_j}$: هو شعاع القيم الذاتية في الفضاء R^K .

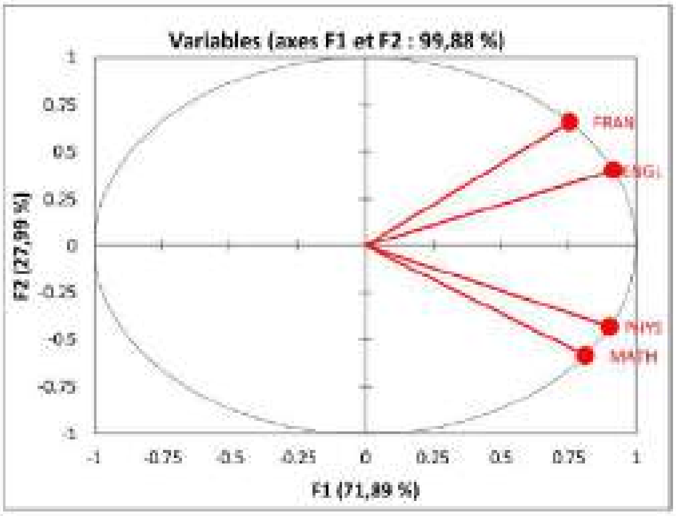
$\frac{1}{\delta_j}$: هو شعاع مقلوب الانحرافات المعيارية في الفضاء R^K.

مثال: بالاعتماد على النتائج السابقة. كيف يكون تمثيل المتغير خلال محاور العوامل المختارة ؟

Coordonnées des variables (Corrélations entre les variables et les facteurs)		
	F1	**F2**
MATH	0,811	-0,584
PHYS	0,902	-0,430
FRAN	0,753	0,657
ENGL	0,915	0,401

– **الاحداثيات العاملية للمتغيرات:**

في مفاهيم التحليل العاملي، تسمى الاحداثيات العاملية « اوزان العوامل ». من الناحية الرياضية، فان المكون الرئيسي (المركبة الرئيسية) هو مزيج خطي من المتغيرات الاكثر ارتباطا به. هذا يعني ضمنيا ان احداثيات عامل متغير هي الارتباطات بين المتغير ومحور العامل. وبالتالي، يجب ان يتم تفسير المكونات الرئيسية من حيث الارتباطات بين المتغير ومحور العامل. وبالتالي، يجب ان يتم تفسير المكونات الرئيسية من حيث الارتباط. مع وضع هذه الحقيقة وهدف تفسير العوامل في الاعتبار، بالنظر الى مجموعة من المتغيرات ، سنبحث بشكل طبيعي عن المتغيرات ذات القيم الاعلى بالقيمة المطلقة لإحداثيات العوامل للعوامل المعينة. الاحصاءات المفيدة الاخرى في المساعدة التفسيرية هي : المساهمة النسبية للمحور العاملي في انحراف المتغيرات ، والمساهمة النسبية في تباين المحور العاملي. تعتمد هذه الاحصائيات على سعة احداثيات العامل.



- **مساهمة المتغيرات والأفراد:**

ان مساهمة المتغير هي في الواقع المساهمة النسبية للمتغير في تباين المحور. ان قيم هذه الاحصائية تجعل من الممكن تحديد المتغيرات المراد تفسيرها فيما يتعلق بالإحداثيات العاملية، أي ارتباطها بالمحور العاملي. بطبيعة الحال، من الضروري دراسة المتغيرات التي تفسر اكثر التباين نسبيا على المحور العاملي.

كما هو الحال في المتغيرات، فان مساهمة الفرد هي ايضا المساهمة النسبية لذلك الفرد في تباين المحور العاملي، وهكذا ، بطريقة أخرى فان مساهمة فرد ما هو مقياس لأهمية ذلك الفرد بالنسبة لذلك العامل. فكلما كانت مساهمة احد الافراد كبيرة كلما كان وزنه على ذلك العامل كبير. ونتيجة لذلك عند تفسير المركبات الاساسية , فأنه وجب البدأ بالأفراد الذين تكون مساهماتهم هي الاكثر اهمية. بالمعنى الدقيق للكلمة، لا يجب تفسير هذه الاحصائية للأفراد الاضافيين حيث ان الافراد النشطين فقط هم الذين يساهمون في المحاور.

III-3 تقنيات مساعدة على تفسير النتائج " Aides à l'interprétation"

ان الهدف الرئيسي من استخدام هذا النوع من طرق الاستكشاف هو الوصول الى ترجمة احصائية دقيقة لجداول المعطيات ثم الى تمثيلات هندسية تلخص لنا المعلومات الاجمالية مع اعطاء قراءة شاملة وواضحة لمجموع البيانات قصد تحليلها وتفسيرها. هذه التمثيلات هي بمثابة مخرجات نهائية لطريقة ACP والتي تتمثل اساسا في سحابة النقاط الخاصة بالمفردات والتي تسمى عادة «خريطة المفردات» والتي تكون مصاحبة في الغالب لسحابة المتغيرات والمعروفة باسم «دائرة الارتباطات».

إن تفسير خريطة المفردات يرتكز اساسا على دائرة الارتباطات لأنه وبكل بساطة لا يمكن النظر للمفردات الا من وجهة نظر المتغيرات. ولهذا قمنا بتجميع بعض التقنيات كمساعدة على تفسير النتائج حيث نعتمد في ذلك على الارتباطات بين المتغيرات فيما بينهما مثنى مثنى، وكذا ارتباطهما بالمحاور (المركبات الرئيسية)، كما

ان ما ينطبق على المتغيرات ينطبق على المفردات الاحصائية. بالإضافة الى ما يتعلق بجودة التمثيل (الصورة) سواء العامة او الجزئية.

تفسير الاحداثيات العاملية للمتغيرات:

كما ذكرنا سابقا، فان احداثيات العامل ليست اكثر من الارتباط بين المتغير ومحور العامل. فكلما زادت القيمة المطلقة لوزن متغير على عامل معين، زاد ارتباط المتغير بهذا العامل. وبعبارة اخرى ، كلما زاد التنسيق العاملي للمتغير، زاد مساهمة المتغير في المفهوم الذي يمثله هذا العامل. على سبيل المثال ، يمكن تفسير عامل ذو اوزان عالية لثلاثة مقاييس لجسم الشخص، مثل الوزن بالبنية، والارتفاع بالسنتمتر، ومحيط الصدر بالسنتمتر، على انه يمثل حجم الجسم. أي ان هذه المتغيرات الثلاثة تساهم بقوة اكبر في هذا المحور.

تفسير الاحداثيات العاملية للأفراد:

يتم تفسير الاحداثيات العاملية للأفراد فيما يتعلق بمساهمتها في التباين. في الخطوة الاولى ، نبحث عن الافراد الذين لديهم اعلى مساهمات للعامل المختار. يمكننا بعد ذلك اختيار مجموعة فرعية من هؤلاء الافراد والبحث عن مساهمات اكبر من متوسط المساهمة. ثم يتم تقسيم المجموعة الفرعية من هذه النقاط الى مجموعتين: الاولى ذات الاحداثيات السلبية، والثانية باحداثيات ايجابية. هذا التقسيم يجعل من الممكن تسليط الضوء على الاختلافات الموجودة بين الافراد، وبالتالي الكشف عن هيكل البيانات المخفية في الافراد.

تفسير الرسومات المرتبطة بالمتغيرات:

الاحداثيات العاملية للمتغيرات في ارتباطات المتغيرات والمحور العاملي اذا كان التحليل الحالي يعتمد على مصفوفة الارتباط. في تمثيل رسومي ثنائي الابعاد، يتم العثور عليها داخل دائرة، تسمى دائرة الارتباط، مع زوج من محاور العوامل. توفر المتغيرات، عند عرضها في هذه الدائرة، الكثير من المعلومات. على سبيل المثال ، كلما كانت النقطة اقرب الى الدائرة، كلما كان ارتباط المتغير المقابل مع محاور العامل افضل. وهكذا يمكن تحديد المتغيرات المرتبطة بمحور معين، وبالتالي المشاركة في المعلومات المقدمة من المتغيرات التي توضح العامل المحدد. وبالمثل، فان موقع الاحداثيات العاملية لمتغير على طول محاور العوامل يصنفها في فئة او اخرى. على سبيل المثال، بدءا من المحور العاملي الاول، يمكن تصنيف المتغيرات الى فئتين، اعتمادا على ما اذا كانت الاحداثيات العاملية للمتغيرات مرتبة على جانب او اخر من المحور العاملي. وبعبارة اخرى، يتم تصنيف المتغيرات وفقا لعلامة احداثيات العامل . يمكن الحصول على المزيد من معلومات التصنيف الاساسية من موامرة احداثيات العامل عن طريق تكرار هذا التمرين لمحاور عوامل اخرى.

تفسير الرسومات المرتبطة بالأفراد:

بالنسبة للرسومات البيانية المرتبطة بالأفراد، يتم تحديد زوج من محاور العوامل من مجموعة محاور العوامل. ثم يتم اسقاط نقاط الفضاء العاملي المولدة من قبل الافراد على المستوى العاملي الناتج عن زوج من المحاور المختارة . يمكن استخدام هذه الرسوم البيانية لتصنيف الملاحظات الفردية الافراد الى فئات. يتم تصنيف الافراد حسب احداثياتهم المقابلة على محاور العوامل. من خلال النظر في ازواج مختلفة من المحاور بين العوامل المحسوبة، يمكن تمييز المزيد من فئات الافراد.

بالاعتماد على الزوايا المثلثية الشهيرة:

باعتبار ان جب تمام الزاوية بين متغيرين في «R» يتوافق مع معامل الارتباط بين المتغيرين اللذان يتشاركان نفس الزاوية. وعليه فانه :

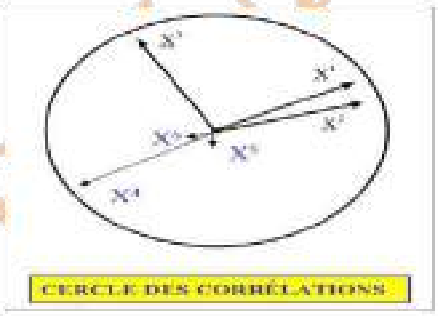
- اذا كان المتغيران قريبان جدا من بعضهما البعض، فانه تكون الزاوية بينهما قريبة من 0 درجة وهذا يعني ان جب (0°)=1، لذلك فان معامل الارتباط يكون قريبا من 1 ، وتكون المتغيرتين مرتبطتين ارتباطا قويا موجبا.
- اذا كان المتغيران متباعدان بحيث يشكلان زاوية قائمة اي 90° درجة وعليه فان : جب (90°)=0 = r، لذلك يكون المتغيرتين مستقلان.
- اذا كانت الزاوية المشكلة بواسطة هذين المتغيرين مساوية الى 180 درجة فهذا يعني : جب (180°) =-1. وعليه يكون الارتباط بين المتغيرتين ارتباط قوي سالب.

وهذه القاعدة يمكن تطبيقها بالنسبة لارتباط المتغير مع المحاور. غير انه يجب ان نحذر لان هذه القاعدة ليست صحيحة دائما لأنه يجب ان يكون المتغيرات ممثلة تمثيلا جيدا على المستوى العاملي. ولكي نتعرف على جودة تمثيل هذه المتغيرات يكفي ان ننظر الى دائرة المتغيرات حيث انه كلما اقترب تموضع المتغيرات الى حدود الدائرة كلما كان تمثيلها جيدا. اما اذا اقتربت الى المركز كان التمثيل سيئا. وعليه يمكن اهمال المتغيرات ذات التمثيل السيئ في حين نحتفظ فقط بالمتغيرات ذات التمثيل الجيد.

بشكل عام تجدر الاشارة الى ان انشاء دائرة الارتباطات ليس من اجل تفسير الارتباطات بين المتغيرات الابتدائية للدراسة، وانما تم انشاؤها من اجل تفسير محاور العطالة المساهمة في التباين الكلي.(Inertie total).

اما فيما يخص جودة التمثيل الكلي فان قيمة التباين الكلي هي التي تدل على ما اذا كان التمثيل جيدا من قبل المحاور الرئيسية وهذا كحد ادنى لعتبة التباين وهذا للتقليل من حجم ضياع المعلومة . هذه النسبة تتوزع ما بين المركبات الرئيسية ، كنسبة اكبر للمركبة الاولى ونسبة اقل لباقي المركبات الاخرى.

مثال تطبيقي:



- ماذا يمثل الشكل؟
- ماهي التقنيات التي يعتمد عليها في تفسير هذا النوع من النتائج ؟

- يمثل هذا الشكل دائرة الارتباطات او سحابة النقاط الخاصة بالمتغيرات. التقنيات التي يعتمد عليها في تفسير هذا النوع من النتائج هي:

. البعد والقرب عن محيط الدائرة والذي يفسر التمثيل الجيد و السيء للمتغيرات (المتغيرات التي يتم تمثيلها بشكل جيد ، أي تلك القريبة من محيط دائرة الارتباط والعكس صحيح) وهذا يكون بواسطة طول الشعاع.

. احداثيات المتغيرات بالنسبة للمحاور لتحديد المتغيرات التي تصنع المحاور. وبذلك نكون قادرين على تسمية المحاورة وفقًا لتلك للمتغيرات

. تموضع المتغيرات بالنسبة لبعضها البعض (الزاوية بين متغيرين)، وبما أن معامل الارتباط بين متغيرين هو جيب تمام الزاوية التي شكلتها المتجهات فإننا نجد: $\cos(\text{angle}) = r(A, B)$ نستنتج أن:

- هناك متغيرين متقاربين أو متضايقين (زاوية 0 °) مرتبطان بإيجابية (معامل ارتباط قريب من 1) وهذا في حالة المتغيرتين $\cos(\text{angle}) = \cos(0°) = r(X1:X2) = 1$.

- هناك متغيران متلازمان متعاكسان، متقابلان (يشكلان زاوية 180 °) مترابطين سلبا (معامل الارتباط قريب من -1) وهذا في حالة المتغيرتين $\cos(\text{angle}) = r(X2, X4) = r(X1, X4) = -1$

- هناك متغيرين لا يرتبطان على الإطلاق وهذا بالنسبة للزوايا القائمة (زاوية 90 درجة). وهنا نجد كل من X1وX3 او X3و X2 او X4و X3 (معامل الارتباط يساوي 0)

- جميع المتغيرات جاء تمثيلها جيد وهذا لقربها من محيط الدائرة عدا المتغيرتين X5 و X6 اللذان يقتربان من مركز الدائرة والذي هو عبارة عن مركز الثقل لذلك جاء تمثيلهما سيئا.

III-4 دراسة حالة

يقدم الجدول التالي هيكل الميزانية العمومية لمجموعة النفط من 1969 إلى 1984

Année	NET	INT	SUB	LMT	DCT	IMM	EXP	VRD
1969	17,93	3,96	0,88	7,38	19,86	25,45	5,34	19,21
1970	16,21	3,93	0,94	9,82	19,11	26,58	5,01	18,4
1971	19,01	3,56	1,91	9,43	17,87	25,94	5,4	16,88
1972	18,05	3,33	1,73	9,72	18,83	26,05	5,08	17,21
1973	16,56	3,1	2,14	9,39	20,36	23,95	6,19	18,31
1974	13,09	2,64	2,44	8,1	25,05	19,48	11,61	17,59
1975	13,43	2,42	2,45	10,83	22,07	22,13	11,17	15,49
1976	9,83	2,46	1,79	11,81	24,1	22,39	11,31	16,3
1977	9,46	2,33	2,3	11,46	24,45	23,07	11,16	15,77
1978	10,93	2,95	2,25	10,72	23,16	24,17	9,64	16,2
1979	13,02	3,74	2,21	7,99	23,04	19,53	12,6	17,87
1980	13,43	3,6	2,29	7,09	23,59	17,61	16,67	15,72
1981	13,37	3,35	2,58	6,76	23,94	18,04	15,42	16,54
1982	11,75	2,74	3,11	7,37	25,04	18,11	14,71	17,18
1983	12,95	3,05	3,85	7,12	23,4	19,17	11,86	18,97

1984	13	3	4	7	24	20	12	17

تمثلت بنود هذه الميزانية في العناصر التالية:

NET: صافي الاسهم ويمثل جميع حقوق ملكية الشركة.
INT: سعر الفائدة وتمثل جميع التكاليف المالية التي تتحملها الشركة
SUB: الاعانات وتمثل المبلغ الاجمالي لمنح الدولة
LMT: ديون طويلة ومتوسطة الاجل
DCT: ديون قصيرة الاجل
IMM: الاصول الثابتة وتمثل جميع الاراضي الشركة والمعدات
EXP: قيمة الاستخدامات
VRD: القيم القابلة للتحقيق والمتاحة وتتمثل في جميع مستحات الشركة القصيرة الاجل

تم تقسيم البيانات كنسبة مئوية بحلول العام , حيث ان مجموع قيم السطر الواحد تساوي 100 وذلك لتجنب الاثار الناتجة عن التضخم. حيث نسعى من خلال تحليل هذه البيانات للإجابة على:

1- ما هو تطور هيكل الميزانية العمومية على مدى 15 عاما؟

2- هل يمكننا تسليط الضوء على عدة فترات ؟ اذا كان الامر كذلك , كيف تتميز؟

- الجداول من 1 الى 10 تم الحصول عليها باستخدام طريقة تحليل المركبات الاساسية وهذا بواسطة برنامج Excel Stat .
- **المطلوب:** قم بتفسير مجموع النتائج بدءا بمصفوفة الارتباط الى الاختبارات القبلية ثم النتائج المتعلقة بالرسومات المختلفة.

الجدول الاول: الاحصاءات الوصفية (Statistiques descriptives)

Variable	Observations	Minimum	Maximum	Moyenne	Ecart-type
NET	16	9,460	19,010	13,876	2,893
INT	16	2,330	3,960	3,135	0,531
SUB	16	0,880	4,000	2,304	0,845
LMT	16	6,760	11,810	8,874	1,729
DCT	16	17,870	25,050	22,367	2,369
IMM	16	17,610	26,580	21,979	3,154
EXP	16	5,010	16,670	10,323	3,849
VRD	16	15,490	19,210	17,165	1,154

- من خلال جدول الاحصاءات الوصفية فانه يظهر لنا مختلف الخصائص الوصفية لعموم المتغيرات حيث انه لا وجود للقيم المفقودة كما انه يوجد اختلاف كبير بين المتوسطات الحسابية وكذا تشتت البيانات وهذا راجع لطبيعة المؤشرات في حد ذاتها التي تم الاعتماد عليها في تشخيص الحالة الصحية للمركب البترولي.

الجدول الثاني: مصفوفة الارتباط الخطي البسيط ((Matrice de corrélation (Pearson (n))								
Variables	**NET**	**INT**	**SUB**	**LMT**	**DCT**	**IMM**	**EXP**	**VRD**
NET	**1**	**0,687**	-0,448	-0,213	**-0,890**	**0,548**	**-0,703**	0,493
INT	**0,687**	**1**	-0,449	-0,460	**-0,601**	0,246	-0,340	**0,530**
SUB	-0,448	-0,449	**1**	-0,409	**0,613**	**-0,693**	**0,608**	-0,142
LMT	-0,213	-0,460	-0,409	**1**	-0,188	**0,598**	-0,390	-0,419
DCT	**-0,890**	**-0,601**	**0,613**	-0,188	**1**	**-0,817**	**0,864**	-0,353
IMM	**0,548**	0,246	**-0,693**	**0,598**	**-0,817**	**1**	**-0,945**	0,202
EXP	**-0,703**	-0,340	**0,608**	-0,390	**0,864**	**-0,945**	**1**	-0,461
VRD	0,493	**0,530**	-0,142	-0,419	-0,353	0,202	-0,461	**1**
Les valeurs en gras sont différentes de 0 à un niveau de signification alpha=0,05								

- من خلال مصفوفة الارتباط نلاحظ ان أقوى الارتباطات هي تلك المقابلة لمعاملات الارتباط التي تقترب من 1 أو -1. في هذا المثال نجد ان ، معامل الارتباط الذي تكون قيمته المطلقة أقرب إلى 1 هو الذي يتعلق بالمتغيرات (DCTو NET) , (IMMو DCT) , (DCTو EXP) , (EXPو NET) و (IMMو EXP). أما بقي الارتباطات فكانت معظمها متوسطة باختلاف اشاراتها. كما جاءت معظم معاملات الارتباط سالبة مما يدل على ان هناك تدهور الحالة العامة لميزانية المركب البترولي .

الجدول الثالث : اختبار بارتلات (Test de sphéricité de Bartlett)	
Khi² (Valeur observée)	205,958
Khi² (Valeur critique)	41,337
DDL	28
p-value	< 0,0001
Alpha	0,05

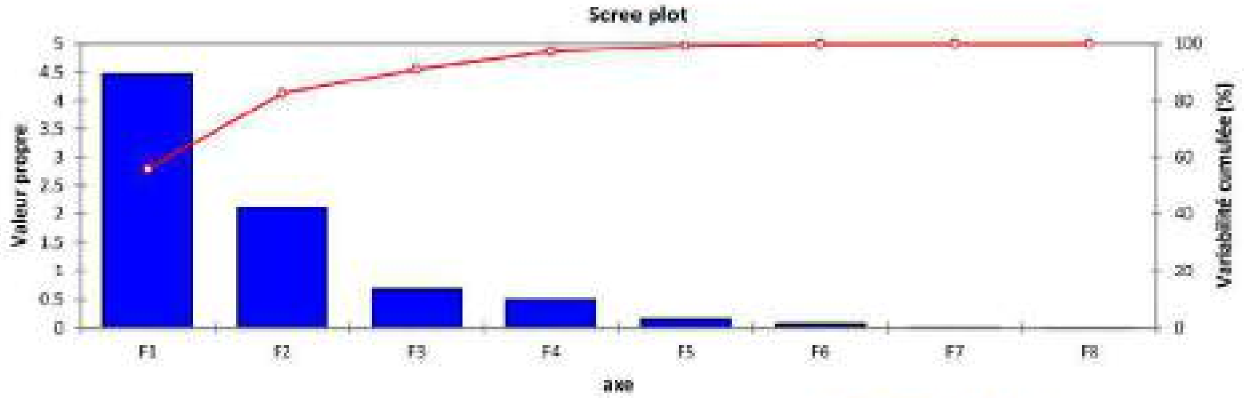
- من خلال اختبار بارتلات يظهر لنا ان هناك امكانية ضغط المعلومات الى حد كبير، وبذلك يمكن من انشاء ملخص جيد لعموم المعلومة .

الجدول الرابع: القيم الذاتية (Valeurs propres)								
	F1	**F2**	**F3**	**F4**	**F5**	**F6**	**F7**	**F8**
Valeur propre	4,466	2,123	0,679	0,501	0,157	0,065	0,009	0,000
Variabilité (%)	55,827	26,537	8,487	6,267	1,960	0,807	0,113	0,001
% cumulé	55,827	82,364	90,852	97,119	99,079	99,886	99,999	100,000

- من خلال جدول القيم الذاتية يتضح أن مجموع القيم الذاتية يساوي الى 8. في حالة ACP المعياري مثل ما هوهنا ، فإن هذه القيمة تساوي عدد المتغيرات. هذه القيمة تتوافق أيضا مع التباين الكلي لسحابة الأفراد (Inertie totale). وعليه يمكننا أن نختار التركيز فقط على القيم الذاتية التي تكون مساهمتها في التباين فوق المتوسط. هذا يؤدي لدراسة المكونات الرئيسية (المحاور الاساسية) المقابلة للقيم

الذاتية الأكبر من 1. ومع ذلك، لا تتحقق سوى القيمتين الأوليتين هذه الخاصية في هذا المثال ($\lambda_1 = 4,466, \lambda_2 = 2,123$). أما باقي المحاور فيتم اهماله نظرا لقيمة لا تدا التي تقل بكثير من 1.

المخطط 1 : منحنى القيم الذاتية



– نفس الملاحظة تظهرها لنا الرسم البياني للقيم الذاتية.

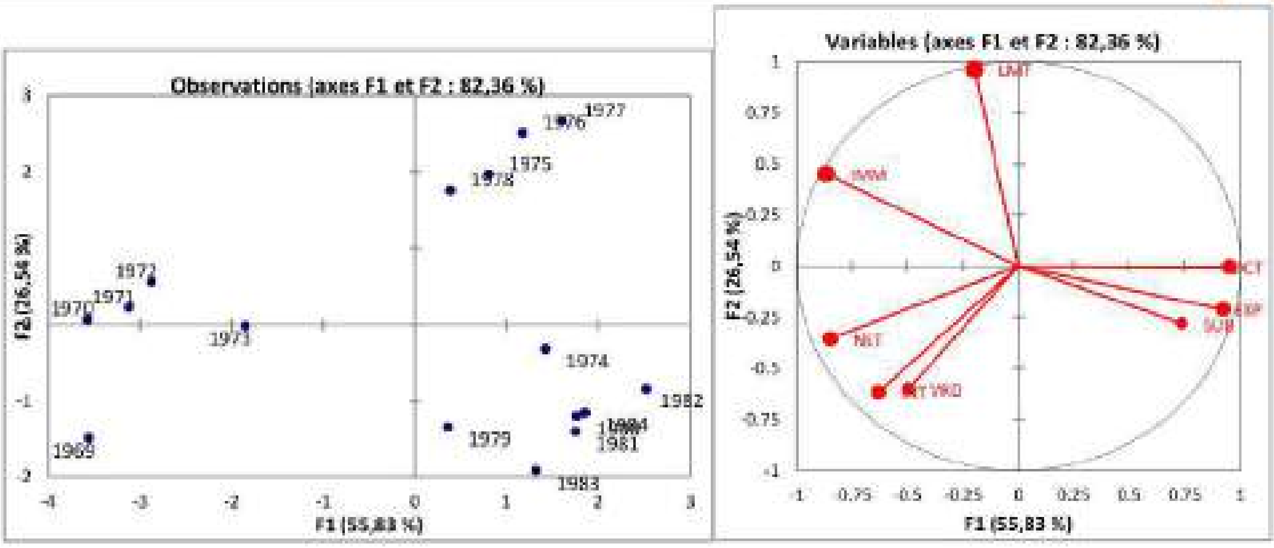
الجدول الخامس والسادس : احداثيات كل من المتغيرات والمفردات الاحصائية بالنسبة للمحورين الاول والثاني

(Coordonnées des variables et des observations sur les axes)

Coordonnées des observations		
Observation	**F1**	**F2**
1969	-3,559	-1,492
1970	-3,574	0,059
1971	-3,118	0,220
1972	-2,872	0,550
1973	-1,849	-0,022
1974	1,427	-0,325
1975	0,804	1,966
1976	1,171	2,508
1977	1,608	2,659
1978	0,387	1,754
1979	0,362	-1,349
1980	1,764	-1,203
1981	1,752	-1,402
1982	2,520	-0,846
1983	1,324	-1,919
1984	1,852	-1,159

Corrélations entre les variables et les facteurs (coordonnées des variables sur les axes)		
	F1	**F2**
NET	-0,849	-0,356
INT	-0,631	-0,618
SUB	0,739	-0,281
LMT	-0,197	0,962
DCT	0,954	-0,005
IMM	-0,867	0,451
EXP	0,926	-0,212
VRD	-0,495	-0,601

- الجدولين الخامس والسادس يساعدان في الحصول على المخططين 2 و 3 واللذان يمثلان سحابة النقط للمفردات وكذا سحابة النقاط للمتغيرات.



المخطط : 2 و 3: يمثل كل من سحابة النقاط الخاصة بالمتغيرات والأفراد

- من خلال دائرة المتغيرات يظهر لنا ان جميع المتغيرات مرتبطة ارتباطا قويا بالمحور الرئيسي الاول عدا المتغيرة LMT التي ارتباطها ضعيف جدا , وعليه يمكن اعتبار المحور الاول كمحور التناظر الكلي خصوصا من حيث انه يفسر لنا نسبة 56% من اجمالي المعلومة. حيث ومن خلال المخطط يظهر لنا ان المتغيرات (DCT, EXP وSUB) متناظرة مع (NET ,IMM). كما يناظر المحور الثاني بين المتغيرتين اللتان ترتبطان معه ارتباطا قويا اي VRD وال LMT وبالتالي يعتبر محور التناظر الجزئي بحيث أنه تمكن من تفسير %5, 26 من باقي المعلومة الغير مفسرة .
- من خلال سحابة النقاط الخاصة بالمفردات فانه وباتباع نفس الفكرة السابقة وبالاعتماد على المتغيرات فانه يمكن التمييز بين ثلاث انواع من المجموعات المختلفة وهي المجموعة الاولى وتظم السنوات (70, 71,72,و 69) , المجموعة الثانية وتحوي السنوات (74, 79,80, 81,82,83,84) , المجموعة الثالثة وتتمثل في (75, 76, 77, 78).

يمكن تأكيد نتائج هذه الطريقة عن طريق اسلوب التصنيف التسلسلي الذي سوف يتم دراسته لاحقا حيث أعطى النتائج التالية (انظر المخطط 4).

- من المعروف ان المتغيرات التي تساهم في تشكيل المحور الاول هي المتغيرات التي تكون احداثياتها ربية من الواحد بالقيمة المطلقة على ذلك المحور. من اجل تحديد هذه المتغيرات فإننا سوف نعتمد

الجدول السابع حيثان الاشارة تدل على الجهة التي يساهم فيها حيث نحصل على : DCT SUB NET EXP IMM
كلها تساهم في تشكيل المحور الاول بنسب متفاوتة اما بالنسبة للمحور الثاني فنجد VRD LMT INT

الجدول السابع: مساهمة المتغيرات في تشكيل المحاور((%) Contributions des variables)		
	F1	**F2**
NET	**16,121**	**5,962**
INT	8,913	**17,970**
SUB	**12,235**	3,725
LMT	0,869	**43,606**
DCT	**20,381**	0,001
IMM	**16,814**	9,569
EXP	**19,190**	2,126
VRD	5,477	**17,040**

- -يمكن اسقاط نفس الفكرة على جدول مساهمة الافراد لمعرفة الافراد التي تشارك في تشكيل المحاور

الجدول الثامن: مساهمة المفردات في تشكيل المحاور((%) Contributions des observations)		
	F1	F2
1969	17,724	6,553
1970	17,872	0,010
1971	13,602	0,142
1972	11,546	0,892
1973	4,782	0,001
1974	2,849	0,312
1975	0,904	11,383
1976	1,919	18,522
1977	3,619	20,820
1978	0,210	9,059
1979	0,183	5,357
1980	4,352	4,263
1981	4,296	5,785
1982	8,889	2,105
1983	2,452	10,840
1984	4,801	3,956

- الجدولين التاسع والعاشر يبين جودة التمثيل لكل من المتغيرات والافراد في المخطط الاول (المركبتين الرئيسيتين الاولى والثانية) من خلال " Cosinus carrés " للزاوية المحصورة بين مختلف هذه المتغيرات والافراد مع المحاور.

- بالنسبة للجدول المتعلق بالأفراد: المفردات الاحصائية الأفضل تمثيل هي السنوات **1969**, **1970**, **1971**, **1972,1973** , **1974**، **1980**, **1981** و **1982** هذا على مستوى المحور الاول اما بالنسبة للمحور الثاني فنجد **1975,1976** , **1977** و **1978** و **1979** مع اعلى جودة تمثيل 0,931]. أما باقي المفردات الاحصائية فكان تمثيلها سيئا بالنسبة للمركبتين الاولى والثانية (**الجدول التاسع**).
- بالنسبة للجدول المتعلق بالمتغيرات: جميع المتغيرات جاء تمثيلها جيدا **NET** , , **SUB** , **DCT**و **IMM** على مستوى المحور الاول عدا **INT**و **LMT** اما بالنسبة للمحور الثاني فنجد أن جميع المتغيرات كان تمثيلها سيئا عدا **LMT** كان تمثيلها جيدا على مستوى هذا المحور (**الجدول العاشر**).

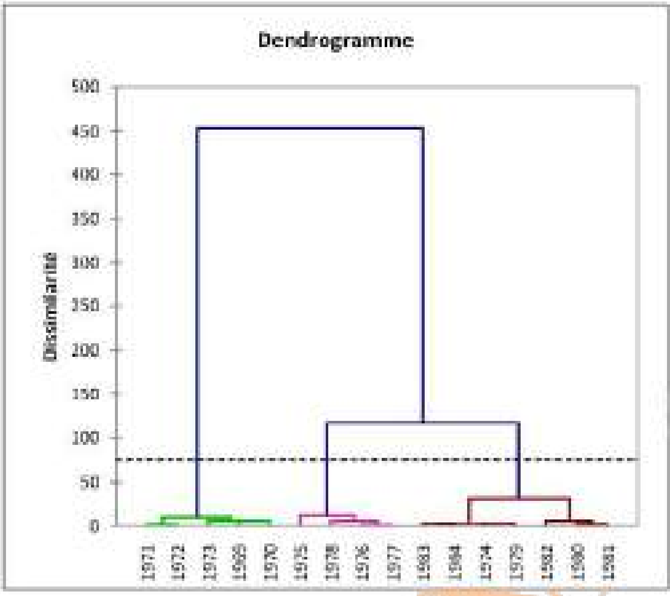
الجدول التاسع: مربع جيب (Cosinus carrés des variables)		
	F1	F2
NET	0,720	0,127
INT	0,398	0,382
SUB	0,546	0,079
LMT	0,039	0,926
DCT	0,910	0,000
IMM	0,751	0,203
EXP	0,857	0,045
VRD	0,245	0,362
Les valeurs en gras correspondent pour chaque variable au facteur pour lequel le cosinus carré est le plus grand		

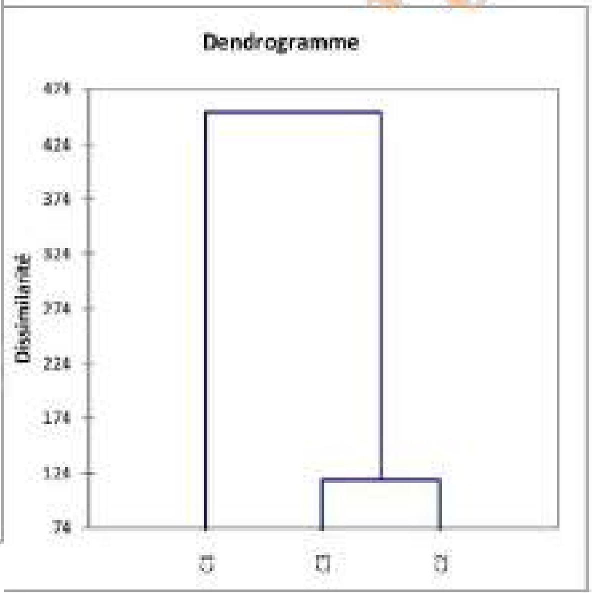
الجدول التاسع: مربع جيب (Cosinus carrés des observations)		
	F1	F2
1969	**0,786**	0,138
1970	**0,931**	0,000
1971	**0,830**	0,004
1972	**0,892**	0,033
1973	**0,756**	0,000
1974	**0,574**	0,030
1975	0,114	**0,682**
1976	0,161	**0,737**
1977	0,262	**0,716**
1978	0,039	**0,795**
1979	0,040	**0,557**
1980	**0,350**	0,163
1981	**0,492**	0,315

1982	**0,871**	0,098
1983	0,203	**0,427**
1984	0,500	0,196
Les valeurs en gras correspondent pour chaque observation au facteur pour lequel le cosinus carré est le plus grand		

- نتائج طريقة التصنيف التسلسلي التصاعدي:

من خلال النتائج المحصلة عن طريق التصنيف التسلسلي التصاعدي نلاحظ ان هناك ثلاث مجموعات وهذا ما يتوافق مع النتائج المقدمة عن طريق التحليل المركبات الاساسية.





Observation	Classe	Observation	Classe	Observation	Classe
1969	1	1974	2	1975	3
1970	1	1979	2	1976	3
1971	1	1980	2	1977	3
1972	1	1981	2	1978	3
1973	1	1982	2		
		1983	2		
		1984	2		

قائمة المراجع

-ESCOFIER B., & PAGES. J, 2008, « Analyses factorielles simples et multiples; objectifs, méthodes et interprétation », 4e édition, Dunod, Paris, 318 p.

-HUSSON . F., Pagès J., 2016, « Analyse des données avec R », 2e Edition, Presses Universitaires de Rennes, 240 p.

- **BACCINI A., SABATIER. P.**, 2010, « Statistique Descriptive Multidimensionnelle (pour les nuls) », Institut de Mathématiques de Toulouse, UMR CNRS 5219, Toulouse cedex 9, France.

- **CARPENTIER F-G.**, 2013/2014, « Analyse multidimensionnelle des données - Master 2ème année - Psychologie Sociale des Représentations », Réf : PSR92C – (polycopié et fichiers de données utilisés) / URL / http://geai.univ-brest.fr/~carpenti/

- **PHILIPPE C.**, 2006, « Principe de l'analyse factorielle », l'université de Versailles – St-Quentin Version, p 34.