

## **L'ANALYSE DE DONNEES : POUR QUOI FAIRE ? Familiarisation au commentaire et à l'interprétation des résultats d'une Analyse en Composantes Principales normées (ACP normée)**

**Myriam BARON, Université Paris 7 – Denis Diderot  
Christine ZANIN, Université Paris 7 – Denis Diderot**

**Type :** TD, enseignements intégrés

**Niveau :** 3<sup>e</sup> année de Licence et au-delà

**Durée :** 1 séance de 3 heures, 2 séances de 2 heures

**Thèmes :** Formulation de questions de recherche, de problématiques  
Utilisation de l'analyse de données pour y répondre  
Géographie de la population  
Etudes de cas  
Cartographie thématique

### **Objectifs :**

Cette ou ces séances de travaux dirigés constitue(nt) soit une introduction, soit un exemple de commentaire et d'interprétation de résultats d'analyse de données utilisées en géographie à partir d'un exemple d'Analyse en Composantes Principales normées (ACP normée). Elle(s) s'adresse(nt) à des étudiants qui ont pu déjà acquérir des notions de statistiques uni- et bivariées au cours des deux premières années du cursus de géographie. On insiste, lors des commentaires à proprement parler, sur le fait qu'il existe deux types de résultats dans une ACP normée (comme dans la plupart des méthodes d'analyse de données) : d'une part ceux qui rendent compte de l'efficacité de l'analyse effectuée ; d'autre part ceux qui sont détaillés et se rapportent aux variables étudiées dans le travail mais aussi aux unités spatiales sur lesquelles ces phénomènes sont observés. Comme dans le cas de l'utilisation des statistiques uni- et bivariées, la démarche proposée vise à insister sur la relation qui doit exister entre traitements statistiques et questionnements, interrogations préalables : pour faire bref, l'outil ou les outils ne doivent pas être déconnectés d'une problématique, d'une thématique. Enfin, on propose de travailler sur les différentes formes graphiques et cartographiques des résultats d'une ACP normée.

Cette séance peut être une première entrée dans l'univers des méthodes d'analyse des données. Selon les publics concernés, elle pourra être envisagée uniquement sur supports papier ou combinée avec la prise en main d'un logiciel qui permet de réaliser des analyses de données. Ici, il s'agit de résultats d'ACP obtenus sous le logiciel Xlstat.

**Déroulement :**

La séance de travail est surtout basée sur la familiarisation au commentaire et à l'interprétation des résultats d'ACP normée par les étudiants. L'ensemble des documents est distribué au début de la séance. On travaille à partir des résultats d'une ACP normée menée sur un tableau élémentaire de données qui concerne la situation démographique des 51 états des Etats-Unis d'Amérique appréhendée à partir de résultats provenant des 3 derniers recensements de population (1980, 1990 et 2000). Cette séance comporte trois temps :

- **le premier temps** est celui de **la familiarisation avec le vocabulaire propre à l'ACP normée**, de son appropriation à partir des résultats généraux qui concernent la qualité, l'efficacité de l'analyse factorielle employée.
- **le deuxième temps** est **consacré au commentaire des résultats détaillés de l'ACP normée**. Ces résultats concernent d'une part l'organisation de l'information en fonction des variables, d'autre part l'organisation de cette même information en fonction des unités spatiales, en l'occurrence ici les 51 états américains. Il vise à bien vérifier que les étudiants sont capables de relier les résultats statistiques à des contenus thématiques, mais aussi de formuler clairement des conclusions à partir de ces premiers résultats. Au cours de cette partie, on insiste également sur les différents types de valorisation graphique des résultats détaillés de l'ACP normée.
- Enfin, **le troisième temps** est consacré aux **différentes façons de valoriser cartographiquement les résultats d'une ACP normée**. L'accent est mis sur les bonnes relations entre variables statistiques et variables visuelles mais aussi sur la mise en page de la carte et le soin extrême qui doit être porté à l'habillage du document cartographique ; enfin sur la clarté de la légende et donc sa grande lisibilité.

## L'ANALYSE DE DONNEES : POUR QUOI FAIRE ?

### 1. Réorganiser et hiérarchiser un ensemble d'informations peut-être...

---

**Documents :** Premiers tableau et graphique des résultats de l'ACP normée, intitulés « 1. Les valeurs propres »

**Sources :** New York Times Almanac, 2001, 2002; census 1980, 1990, 2000

---

Comme il peut s'agir d'un premier commentaire d'une ACP normée, on ne peut éviter de faire des rappels de cours relativement importants pour insister sur la spécificité de l'analyse de données appliquée à la géographie et en particulier des analyses factorielles, dont un exemple est l'analyse en composantes principales normée. On insiste bien sur le fait que l'ACP doit permettre d'extraire le maximum d'informations sous une forme simple et cohérente à partir d'un ensemble très important de données. Elle sert à mettre en évidence les interrelations entre les variables, et les ressemblances et oppositions entre les unités géographiques analysées. On se référera à la Feuilles de Géographie intitulée « *Un tableau de données pour quoi faire ?* » pour bien remettre en perspective le travail effectué dans cette feuille.

Les premiers tableau et graphique de résultats de l'ACP normée menée sur le tableau élémentaire de données, qui décrit chacun des 51 états des Etats-Unis d'Amérique en fonction des caractéristiques des populations résidentes ainsi que leurs dynamiques démographiques, permettent de rendre compte de l'efficacité et donc de la pertinence de ce type d'analyse. Il s'agit d'un tableau qui, pour chacune des composantes (ou facteurs ou axes factoriels), rend compte à la fois de la valeur propre qui la(le) caractérise, de la part de variance ou inertie ou information totale prise en compte par ladite composante ainsi que de l'inertie ou variance ou information cumulée à laquelle on parvient en prenant en compte chacun des facteurs résultats de l'ACP normée. C'est l'occasion de rappeler qu'une autre particularité de l'ACP normée réside dans la présentation de ses résultats. Ces derniers sont présentés sous forme de *composantes dites aussi facteurs ou axes factoriels*, qui sont des combinaisons linéaires de toutes les variables présentes dans le tableau élémentaire de données que l'on veut analyser.

Le but de l'ACP est de réorganiser, de hiérarchiser l'information en présentant tous les axes factoriels (ou composantes ou facteurs) selon le degré de différenciation qu'ils introduisent entre les individus et donc dans l'espace géographique. L'analyse revient à effectuer un changement de repères, puisque ce ne sont plus les variables qui définissent chacun des axes mais ces derniers sont définis comme une combinaison linéaire unique de l'ensemble des variables. Le premier axe différencie le plus les unités spatiales, le deuxième axe un peu moins et le troisième encore moins. On n'étudie donc que les premiers axes factoriels qui rendent compte de la majeure partie de l'information contenue dans le tableau élémentaire de données : on en obtient alors une représentation très simplifiée. Le tableau comprend donc les valeurs propres associées à chacun des axes. En outre, l'analyse des parts d'information associées aux axes permet de cerner la structure des données. Ce résultat est présenté en pourcentage : % d'inertie = % de variance = % d'information. Il indique le rôle de chacun des axes dans la différenciation des individus

analysés. Enfin, est associé à ce tableau de synthèse un histogramme des valeurs propres qui est un diagramme en bâtons des pourcentages d'inertie. En abscisse les valeurs propres et en ordonnée leur pourcentage d'inertie. Il ne fait que visualiser les informations contenues dans le tableau. Le relief de l'histogramme est la principale information à observer :

→ Lorsqu'il est très marqué, cela rend compte de la hiérarchisation nette de notre information : ce qui se traduit par le fait qu'un, deux ou trois oppositions importantes se dégagent clairement. On a, dans ce cas, des variables avec des relations significatives et facilement structurables et résumables.

→ Si l'histogramme est plus compact, si moins de pics se distinguent parmi les premières valeurs propres, le nuage de points est plutôt arrondi sans axe d'allongement marqué. On peut en déduire que les interrelations entre variables sont sans doute faibles et qu'il ne se dégage pas de combinaison géographique simple. L'information n'a pas de structure, d'oppositions évidentes. Il est alors difficile de l'interpréter et il vaut mieux revenir à la problématique ou au choix des variables. Ces éléments permettent facilement de dire si l'ACP réalisée fournit un résumé efficace de l'information initiale.

Dans le travail sur les différenciations des états américains en fonction de leurs caractéristiques et dynamiques démographiques, le tableau des valeurs propres montre que les 5 premiers axes sont bien hiérarchisés, avec une quasi-équivalence des 2 premiers facteurs (5,04 et 4,26). Ce qui indique que ces 2 premiers facteurs sont ceux qui introduisent le plus de différences entre les axes. Traduit en % d'inertie (Total des axes = 100 % de l'information), les deux premières composantes résument plus de 51% de la variance totale (%cumulé égal à 51,65%), les trois premiers axes expliquent 63 % de la variance totale (de même %cumulé égal à 63,61%) : ce qui représente une large majorité de l'information initiale. En effet, au départ il y avait 18 variables principales, il y a donc 18 facteurs ou axes factoriels ou composantes. Or, les 5 premiers facteurs sont caractérisés chacun par des valeurs propres supérieures à 1 et résument près de 80% de l'information totale. Ainsi, en retenant moins d'un tiers de l'ensemble des facteurs, on parvient à rendre compte des 4/5<sup>e</sup> de l'inertie totale du nuage de points. Enfin, l'histogramme des valeurs propres est assez marqué, il indique qu'il existe au moins 2 oppositions majeures qui permettront de mettre en valeur une structuration claire de l'information initiale du tableau de données sur les caractéristiques démographiques des états américains.

## **Bibliographie**

### **Ouvrages indispensables**

\*\*\*\*Sanders L., 1989, *L'analyse des données appliquée à la géographie*, Montpellier: Reclus, collection Alidade, 268 p.

Cet ouvrage est un aperçu très détaillé des principales notions d'analyse de données. Il présente l'avantage d'être rédigé par une géographe et de comporter de nombreux exemples géographiques pour illustrer les différentes façons d'utiliser les analyses factorielles et les classifications.

\*\*Taylor P.J., 1977, *Quantitative methods in Geography. An introduction to spatial analysis*, Illinois: Waveland Press, 386 p.

Cet ouvrage présente l'avantage d'aborder les notions liées à l'analyse de données, de présenter des exemples qui ne se recoupent pas avec ceux du livre de Lena Sanders et d'être en langue anglaise !

## Ouvrages de références

\*\*Béguin H., 1979, *Méthodes d'analyse géographique quantitative*, Paris: Litec, collection Droit, 252 p.

Ouvrage très complet. Il présente un large éventail des méthodes et outils statistiques utilisés en géographie. Il comporte des éléments vus en DEUG.

\*Benzécri J., 1973, *L'analyse des données*, Paris, DUNOD, 2 volumes

Il s'agit de l'ouvrage de référence sur l'analyse des données. Il n'est pas du tout géographique et plutôt très théorique. Il est d'abord destiné à ceux qui veulent comprendre quels sont les fondements mathématiques de l'analyse des données et leurs caractères relativement universels pour ce qui est des applications.

\*\*\*Groupe Chadule, 1994, *Initiation aux pratiques statistiques en Géographie*, Paris: Masson, collection Géographie, 4ème édition, 189 p.

Cet ouvrage fait partie des références connues depuis le DEUG. Il permet donc de réviser efficacement les notions supposées acquises en statistique uni- et bivariée. Il permet aussi d'entrevoir les liens possibles entre méthodes d'analyse statistique et méthodes d'analyse de données.

\*Charre J., 1995, *Statistique et Territoire*, Montpellier, RECLUS, Espaces Modes d'Emploi, 119 p.

Il s'agit d'un ouvrage de synthèse, d'une réflexion concernant l'utilisation des statistiques (données mais aussi méthodes de traitement) en géographie. Il est surtout destiné à des personnes qui souhaitent avoir une vue d'ensemble mais aussi qui veulent parvenir à formaliser d'éventuelles critiques quant aux abus d'un tel recours. Il suppose donc d'avoir pris un certain recul vis à vis de cette approche. On retiendra les chapitres 5 et 6 qui sont plus centrés sur les méthodes d'analyse factorielle et de classification.

## L'ANALYSE DE DONNEES : POUR QUOI FAIRE ?

### 2. En savoir plus sur les combinaisons entre caractéristiques démographiques et dynamiques des états américains...

---

**Documents :** - Tableaux des résultats détaillés de l'ACP normée se rapportant aux phénomènes démographiques retenus (2. Les variables ; 3. Les états américains)  
- Graphiques 4, 5 et 6 des 3 premiers plans factoriels

**Sources :** New York Times Almanac, 2001, 2002; census 1980, 1990, 2000

---

Les résultats détaillés concernant les variables spécifient pour chacune d'elles sa coordonnée sur chacun des facteurs, sa qualité de représentation et sa contribution à la structuration de chacun des axes factoriels (ici employés comme synonymes de facteurs).

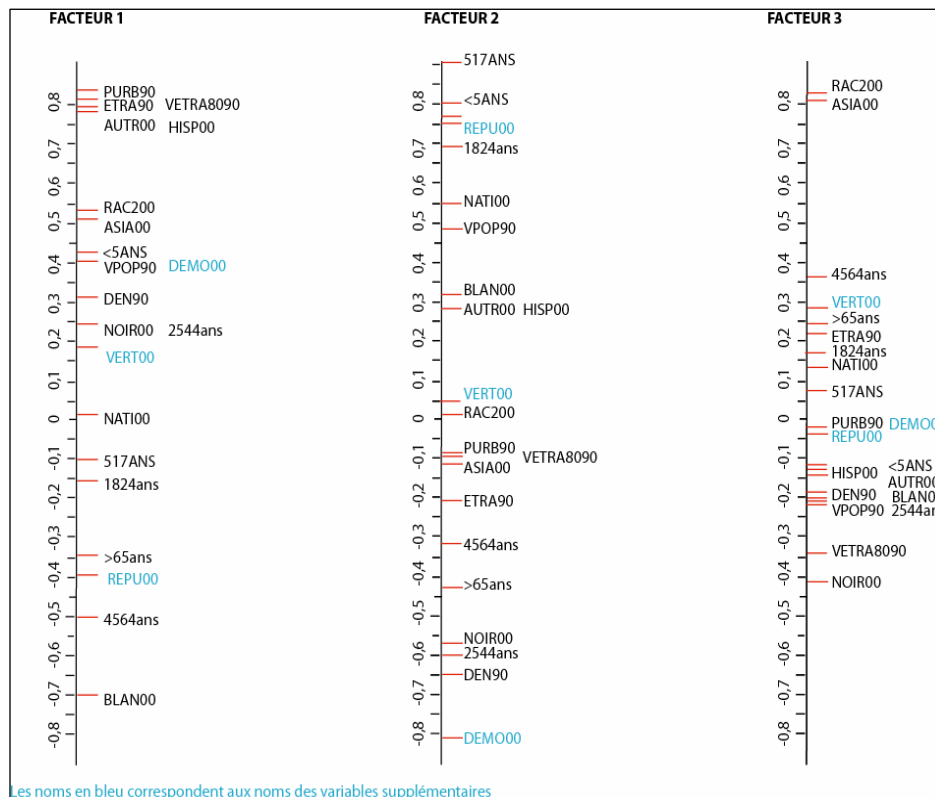
La coordonnée d'une variable sur un axe factoriel est représentée par la projection de cette variable sur l'axe. Cette projection correspond au coefficient de corrélation entre la variable et l'axe. La corrélation traduit la proximité géométrique entre les vecteurs caractérisant les variables et l'axe. Les coordonnées varient donc entre  $-1$  et  $+1$ . Dans la lecture du tableau, on s'intéressera aux associations de variables (c'est - à - dire par exemple les variables qui ont des coordonnées très positives et qui donc se projettent d'un même côté de l'axe) mais aussi aux oppositions de variables les plus fortes (c'est - à - dire celles qui ont les coefficients de corrélation non seulement les plus forts en valeurs absolues mais aussi de signes contraires  $+$  et  $-$ ). C'est ainsi que le premier facteur, qui rend compte de l'opposition majeure entre les variables démographiques sélectionnées, oppose des états où la population blanche dite aussi White Anglo-Saxon Protestant (WASP) (coordonnée égale à  $-0,72$ ), celle âgée de plus de 45 ans ( $-0,51$ ) sont sur représentées à des états dans lesquels la population urbaine est importante ( $+0,82$ ), qui sont caractérisés par une forte population étrangère ( $+0,80$ ) ayant une forte évolution entre 1980 et 90 ( $+0,77$ ) plutôt d'origine hispanique ( $+0,76$ ), ou de la catégorie autres origines étrangères ( $+0,76$ ). Ce qui peut être résumé par le titre suivant : « **Population blanche âgée contre population étrangère jeune** » (cf. graphique 1). Tandis que le deuxième facteur oppose des états où la population est jeune (moins de 25 ans) ( $+0,91$  pour les 5-17 ans,  $+0,80$  pour les moins de 5 ans et  $+0,69$  pour les 18-24 ans)) et les Amérindiens très présents ( $+0,55$ ), à des états denses ( $-0,65$ ) où la population active jeune (25-44 ans) est forte ( $-0,60$ ) et à forte composante afro-américaine ( $-0,57$ ). Ce qui peut être résumé par le titre : « **Forte variation de la population avec sur représentation des amérindiens contre densité d'une population adulte et plutôt afro-américaine** ». Enfin, le troisième facteur oppose des états où les population d'origine asiatique ( $+0,81$ ) ou déclarant appartenir à deux races ( $+0,83$ ) sont sur représentées à des états où la population est plutôt afro-américaine ( $-0,41$ ). Ce qui peut être résumé par : « **Population asiatique ou de double origine contre Afro-américains** ». C'est l'occasion d'insister sur un point important des résultats détaillés d'analyse factorielle quelle qu'elle soit (ACP normée, Analyse Factorielle des

Correspondances etc.), les coordonnées positives de variables sur un facteur ne correspondent pas seulement à des sur-représentations ; les coordonnées négatives de variables ne correspondent pas

directement à des sous-représentations : elles rendent compte essentiellement de la structuration de chacun des facteurs ou axes factoriels ou composantes.

A l'issue de ce travail très analytique, axe par axe, il reste l'image d'une différenciation des états américains essentiellement basée sur les communautés « ethniques ». L'opposition majeure s'organise autour d'un axe populations WASP contre toutes les communautés étrangères et particulièrement les hispaniques. Puis les oppositions secondaires viennent nuancer le « bloc » des communautés étrangères, en soulignant le positionnement particulier des afro-américains (communauté d'origine étrangère la plus importante jusqu'à une date récente aux USA), puis celui des amérindiens (dits aussi « natives », très peu nombreux actuellement), enfin celui des asiatiques et des personnes déclarant appartenir à au moins deux races (question nouvelle et propre au recensement de 2000). La composition par âge ne joue que de manière secondaire dans cette analyse.

**Graphique 1 : les 3 premiers facteurs de l'ACP normée**

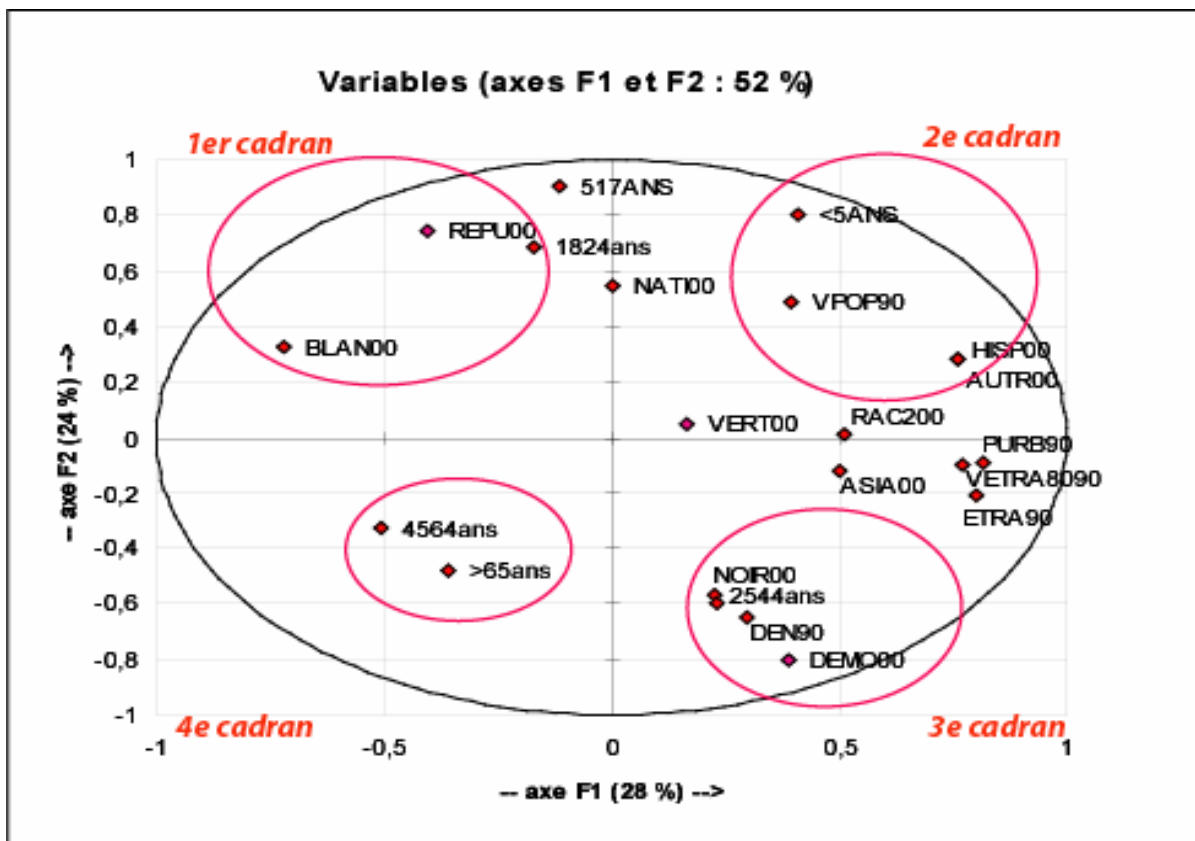


Quant à l'étude par plan factoriel, elle permet de synthétiser encore davantage l'information concernant les individus et les variables. L'information, qui est donnée, concerne simultanément la position d'un individu ou d'une variable sur deux axes (les 1<sup>er</sup> et 2<sup>e</sup> facteurs constituent le

premier plan factoriel, les 1<sup>er</sup> et 3<sup>e</sup> facteurs constituent le troisième plan factoriel enfin les 2<sup>e</sup> et 3<sup>e</sup> facteurs constituent le deuxième plan factoriel). Si on analyse le premier plan factoriel qui est constitué des deux premiers facteurs et qui résume 52 % de l'information totale, on considère cette figure selon ses 4 cadrans. Pour chacun d'eux, on identifie les variables qui sont les plus éloignées du centre de gravité du nuage de points. Ce sont ces variables qui vont permettre de caractériser le contenu de chacun des 4 cadrans (cf. graphique 2). Le premier cadran (en haut à gauche) est marqué par la présence des populations blanches (WASP) et jeunes (5-17 ans et 18-24 ans). Le deuxième cadran (en haut à droite) correspond aux projections de la variation de la population, des moins de 5 ans, de la population hispanique et des autres populations étrangères. Quant au troisième cadran (en bas à droite), il regroupe la densité de population, les jeunes actifs de 25 à 44 ans et les afro-américains. Enfin le quatrième et dernier cadran (en bas à gauche) est marqué par les seules présences des populations actives âgées et des plus de 65 ans.

Les autres variables restent proches de l'un ou l'autre facteur (5-17 ans, Amérindiens sur l'axe 2 et appartenant à au moins deux races, étrangers d'origine asiatique, part des étrangers dans la population totale, part de la population urbaine dans la population totale sur l'axe 1), et ne sont donc pas bien représentées sur ce plan factoriel. L'analyse du plan doit se faire en fonction des variables les plus représentatives.

*Graphique 2 : les premier plan factoriel de l'ACP normée*





A l'issue de ce premier travail, on est en mesure de souligner le fait que les caractéristiques et les dynamiques des populations révèlent les différences entre les 51 états américains et donc l'hétérogénéité du territoire des Etats-Unis d'Amérique. Les états où la population est majoritairement blanche sont plutôt des états âgés, les états denses sont plutôt des états où la population est étrangère, jeune, adulte et à forte composante afro-américaine. Enfin c'est dans les états qui présentent de fortes variations de populations que la population amérindienne est importante.

Il reste à voir un dernier point qui concerne la manière dont cette situation démographique complexe et multi-dimensionnelle renvoie à des comportements politiques. Pour ce faire, on a sélectionné les résultats des trois principales formations politiques lors des élections présidentielles de 2000. C'est alors l'occasion de rappeler voire d'expliquer ce que l'on entend par variable supplémentaire. Une variable supplémentaire est une variable qui ne contribue pas à la formation des axes mais qui est projetée comme si elle faisait partie de l'analyse afin de pouvoir connaître sa position par rapport au nuage de points. Son utilisation peut relever de 2 logiques :

- Mettre à l'écart un élément marginal ou aberrant par rapport à l'ensemble des autres. L'examen des contributions permet de repérer les variables dont le rôle paraît excessivement fort dans la formation des premiers axes.

- Mettre en relation un autre groupe de variables. On effectue une analyse factorielle sur un groupe homogène de variables et il est intéressant de connaître la position d'une ou plusieurs autres variables par rapport à ce groupe. Il est alors utile de connaître la position de ces variables sur les axes factoriels formés par l'ensemble des variables étudiées, sans pour autant que ces variables interviennent dans les calculs. Dans le cas de notre travail sur les caractéristiques démographiques des USA, c'est cette dernière utilisation qui est privilégiée.

Les trois principales formations politiques lors des élections présidentielles de 2000 à savoir les républicains, les démocrates et les écologistes (verts), sont donc les variables supplémentaires dans ce travail. La question que l'on se pose alors est de savoir s'il peut exister une relative coïncidence entre la projection d'une des trois principales formations politiques américaines et certaines caractéristiques des populations mises en évidence. De plus, compte-tenu du fait qu'il n'y a que 3 formations, il s'agit d'identifier les populations qui n'ont pas de représentation politique majeure. Sur le facteur 1, les démocrates se situent plutôt dans les états présentant une majorité de populations urbaine et/ou étrangère alors que les républicains se projettent plutôt là où la population est blanche et âgée. La formation politique des verts n'est pas représentée correctement sur ce premier axe (coordonnées = 0,17). Sur le facteur 2, les républicains « coïncident » avec les états dans lesquels les jeunes adultes sont nombreux et où il y a une forte variation de la population alors que les démocrates sont associés aux états denses et dans lesquels il y a une forte présence d'afro-américains. Les verts restent non représentés sur cet axe. Sur le facteur 3 les verts sont représentés plutôt dans les états où la population asiatique ou déclarant une double origine est plus nombreuse que dans l'ensemble du territoire étasunien. Démocrates et républicains ne sont pas représentés sur cet axe. Si on observe la projection de ces trois formations politiques dans le premier plan factoriel, on constate que les républicains se situent

dans le 1<sup>er</sup> cadran (WASP, jeunes) ; les démocrates dans le 3<sup>e</sup> cadran avec les minorités afro-américaines, la densité de population et les étrangers ; les verts dans le 2<sup>e</sup> cadran mais proche du

centre et donc mal représentés. Les verts se positionnent plus clairement dans le 2<sup>e</sup> graphique factoriel (facteur 2 et 3) mais sont encore très proches du centre de l'axe. Enfin, les états dans lesquels les populations actives âgées sont nombreuses ainsi que les personnes âgées ne bénéficient d'aucune représentation politique majeure. Il est intéressant de souligner que les trois principaux candidats aux élections présidentielles de 2000 représentent plusieurs dimensions des populations américaines (combinaisons des structures par âge mais aussi des origines ethniques).

## L'ANALYSE DE DONNEES : POUR QUOI FAIRE ?

### 3. Mettre toute cette complexité et cette richesse en cartes...

---

**Documents :** - Fond de carte de localisation de 49 états américains  
- Tableaux des résultats détaillés de l'ACP normée se rapportant aux différents états des USA

**Sources :** New-York Times Almanac, 2001, 2002; census 1980, 1990, 2000

---

Quoi de plus satisfaisant que de finir une séance de travail par la réalisation d'une voire de plusieurs cartes ... encore faut-il identifier clairement la nature des variables statistiques à cartographier.

3 modes de représentation cartographiques des résultats d'une ACP sont possibles selon l'interprétation donnée au facteur :

- Si le facteur a été identifié comme **une simple échelle de variation**, comme un axe d'intensité qui ordonne toutes les unités géographiques selon une progression unique des valeurs, des plus faibles aux plus élevées, on peut le considérer comme un caractère quantitatif ordinaire. On peut alors le cartographier :

- par une progression de valeur simple et ordonnée, du noir au blanc ou en couleurs ordonnées.

- en signes proportionnels pour un facteur de taille

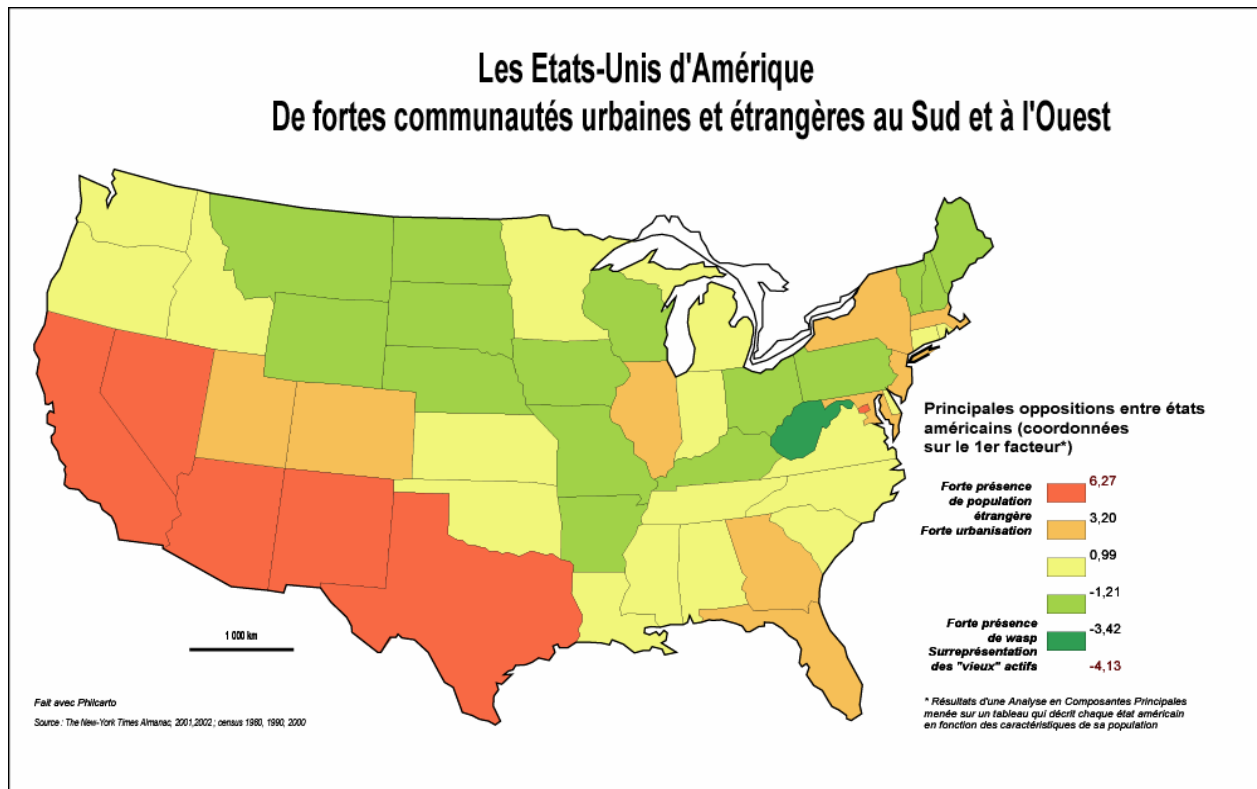
Dans ce cas le titre de la carte sera celui que l'on a donné à l'axe factoriel.

La légende pose un problème plus complexe. Bien que les coordonnées des lieux sur les facteurs soient précisément mesurées, elles se rapportent à une combinaison linéaire de toutes les variables comportant de nombreux coefficients, et leurs valeurs ne sont pas facilement interprétables. On peut donc préférer porter en légende une indication qualitative ordonnée, moins précise mais plus expressive.

- Si le facteur est interprété comme l'expression **d'une opposition de structures entre 2 sous-ensembles** d'unités géographiques qualitativement différentes, il ne s'agit plus de visualiser un classement mais de montrer qu'il existe 2 situations extrêmes différentes, de part et d'autre de situations moyennes décrites par les coordonnées faibles des lieux proches du centre de gravité du nuage de points. Le facteur est alors bipolaire (ex : oppositions urbain/rural, textile/sidérurgique, qui traduisent 2 types de combinaisons géographiques qualitativement différentes). On utilise alors souvent une double progression visuelle pour souligner cette dualité de signification du facteur. C'est ce qui est proposé avec la carte 1. Sur cette dernière se dessine une organisation en bandes successives d'orientation générale Sud-Est/Nord-Ouest. On montre qu'il y a une relative régionalisation des phénomènes démographiques étudiés. Le Sud-Ouest du Texas à la Californie constitue un ensemble d'états caractérisés par une forte présence des

communautés étrangères. Alors qu'une partie du Nord des Grandes Plaines, une partie des Appalaches et les états de l'extrême Nord-Est sont caractérisés par une forte présence WASP. Entre ces deux situations démographiques fortement différenciées, se trouve une bande d'états en situations intermédiaires. On peut bien évidemment faire de même pour les facteurs suivants.

*Carte 1 : exemple de représentation cartographique des résultats d'ACP par facteur*



- On peut également choisir de cartographier **un plan factoriel** :
  - le plan concentrant le plus d'informations, c'est-à-dire celui formé par les deux premiers axes factoriels
  - un plan apparaissant comme particulièrement intéressant au niveau de l'interprétation.

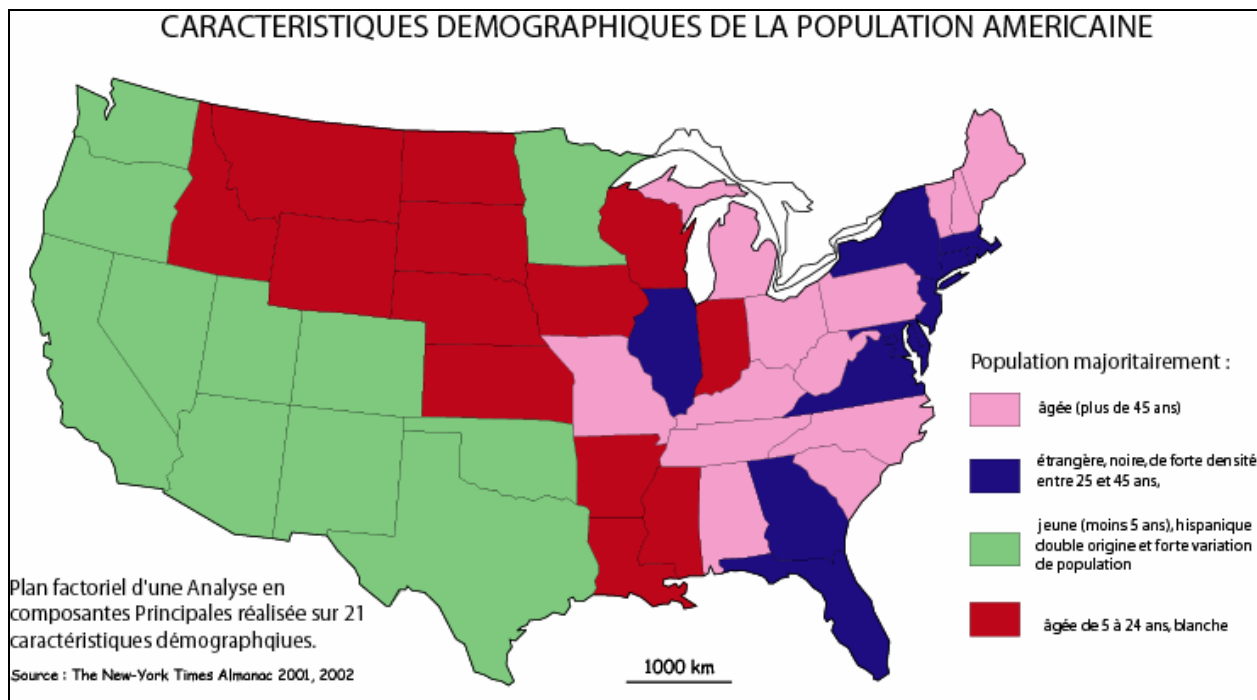
**1<sup>er</sup> cas** : On choisit un figuré particulier ou une couleur différenciée pour chaque quart du plan factoriel, et on le reporte sur la carte. La légende est alors obligatoirement qualitative.

*NB : Du fait de l'interprétation particulière des plans factoriels d'une AFC, il est souvent préférable de cartographier ces plans factoriels.*

Le territoire des Etats-Unis présente alors un tout autre aspect. L'organisation en bandes disparaît au profit d'une opposition est-ouest assez marquée. Á l'est d'une ligne allant du Michigan à l'Alabama, une Amérique « melting-pot » apparaît : plutôt âgée ou à population active dominante. La Floride et la Géorgie ou encore les états très fortement peuplés comme le New Jersey, le Connecticut ou la Virginie montrent une sur représentation des communautés étrangères plutôt afro-américaines.

L'Ouest américain se divise en 2 zones assez distinctes avec, au Nord, une population WASP jeune et au sud, une forte variation de la population plutôt hispanique ou de double origine. L'intérêt de cette représentation réside dans la possibilité d'établir des oppositions claires entre les zones : une typologie nette organise ainsi le territoire.

*Carte 2 : exemples de représentation cartographique des résultats d'ACP par plan factoriel*

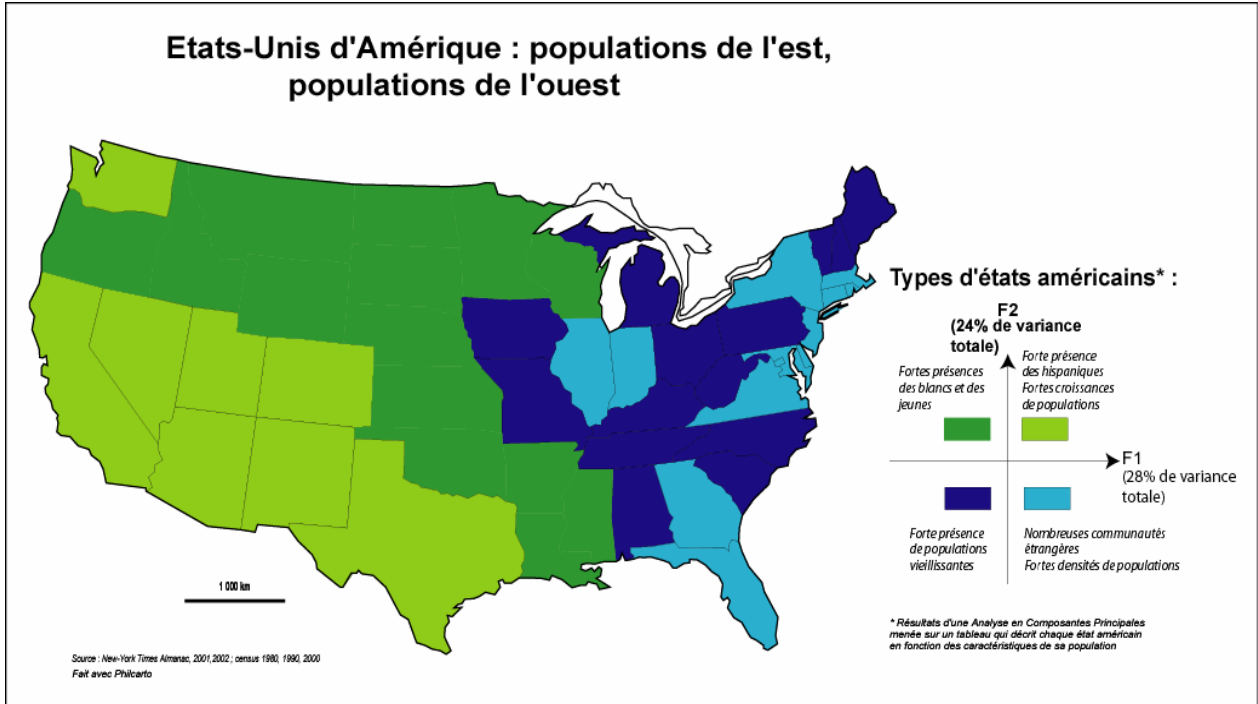


**2<sup>ème</sup> cas** : On cherche à mettre en relation les différents cadrans du plan factoriel par le jeu du choix des trames et des couleurs de la légende.

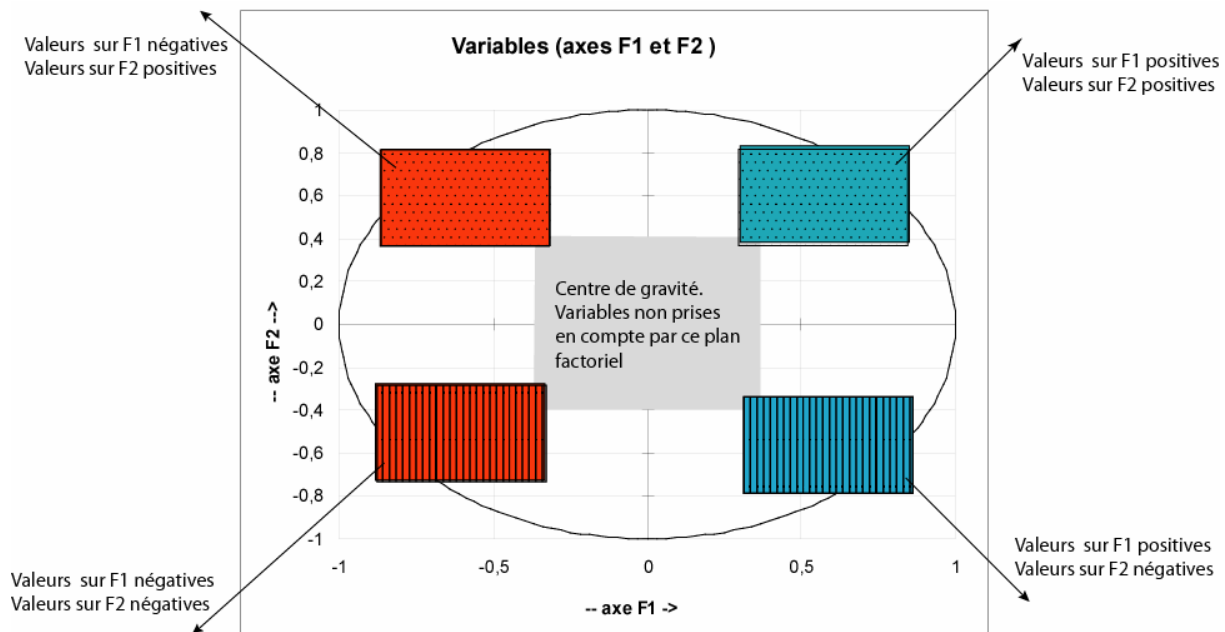
Par exemple : on peut adopter la valeur pour les variations sur le facteur 1 et la couleur pour les variations du facteur 2 : c'est ce qui est proposé sur la carte 3. La régionalisation des Etats-Unis qui ressort dans ce cas souligne, comme dans le cas de la carte 2, la coupure Est-Ouest de part et d'autre d'une ligne qui correspond au cours du Mississippi. A l'Est du fleuve, les états sont caractérisés soit par des populations vieillissantes, soit par de fortes densités de population et une forte présence de communautés étrangères. Au sein de ce premier ensemble d'états, il est très difficile de dégager des règles d'organisation. A l'Ouest du fleuve en revanche, se dégage une organisation spatiale plutôt nette selon un axe Sud-Est (à partir de l'état du Mississippi) Nord-Ouest (jusqu'à l'état de l'Oregon). Au Nord de cette ligne, on trouve surtout des états où les WASP et les jeunes sont sur représentés. Au Sud, on trouve surtout des états dont les populations croissent rapidement et dans lesquels il y a de fortes communautés hispaniques.

La confrontation des cartes 1 et 3 met bien en évidence les configurations régionales différentes qui émergent selon que l'on cartographie les positions des états sur un facteur ou sur un plan factoriel.

**Carte 3 : exemple de représentation cartographique des résultats d'ACP par plan factoriel**



On peut également adopter la couleur pour les variations sur le facteur 1 et la texture-structure pour les variations du facteur 2



Dans la légende, le titre doit être toujours explicite sur le contenu du facteur, même si le titre est long. Le titre et la légende varient selon le public auquel on s'adresse :

→ quand le grand public est visé, le titre doit être lisible et compréhensible immédiatement, et la légende ordonnée qualitativement. Indiquer toujours le label des variables en clair, et non les codes.

→ pour les spécialistes, on peut rajouter :

⇒ **un diagramme en bâtons, disposés horizontalement, correspondant à chacune des variables entrant dans l'analyse, et dont la longueur est proportionnelle à la corrélation entre cette variable et le facteur considéré (coordonnée).**

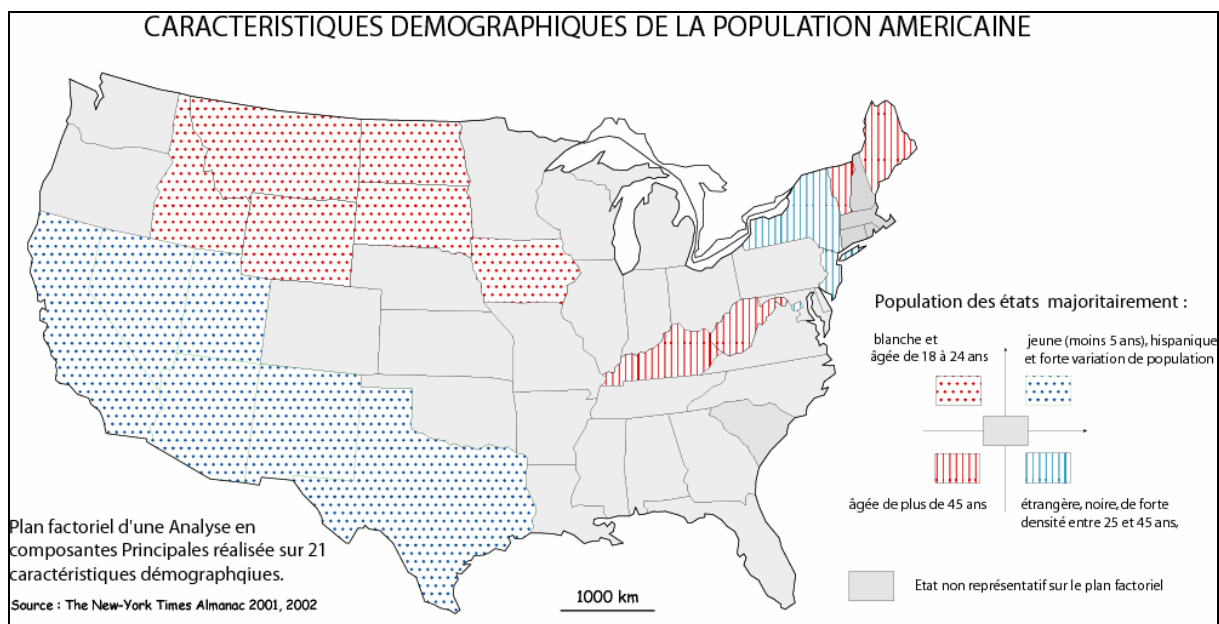
⇒ l'orientation des barres de part et d'autre d'une ligne verticale figurant une corrélation nulle permet de distinguer les corrélations négatives à gauche et positives à droite.

⇒ On peut réordonner les variables selon l'importance de leur corrélation, avec le premier facteur qui est le plus important. Mais attention, il faut garder le même ordre de variable selon les facteurs. Il suffit ainsi de décrypter une seule fois la liste des variables pour analyser la signification de chacun des autres facteurs.

⇒ On rajoute parfois (mais très rarement) un petit diagramme à secteurs, qui visualisent la part de variance prise par chacun des facteurs dans l'ensemble des données analysées.

On obtient alors une carte qui ressemble à la carte 4. Ce dernier exemple de représentation cartographique fait ressortir deux nouvelles informations par rapport à la carte 3. Tout d'abord, la majorité des états ont des caractéristiques démographiques assez moyennes (proches du centre de gravité du nuage de points). Ces états se répartissent sur l'ensemble du territoire américain du Massachusetts à l'Oregon en passant par la Floride, l'Oklahoma ou encore le Colorado. De plus, les situations exceptionnelles concernent plus les états de la partie occidentale du territoire.

**Carte 4 : exemple de représentation cartographique des résultats d'ACP par plan factoriel**



A l'issu de ces différents travaux cartographiques, on peut proposer aux étudiants de réaliser une synthèse des organisations territoriales dominantes selon les résultats cartographiés. Pour cela on peut se référer à l'approche chorématique de Roger Brunet (cf. Document 1). Après une présentation brève de l'analogie avec les phonèmes en linguistique, on fait repérer dans la table des chorèmes les plus petites organisations territoriales qui peuvent être identifiées sur les cartes précédemment réalisées. En soulignant que ce qui est représenté a à voir d'une part avec les notions de contact et de tropisme; d'autre part avec des représentations linéaires ou zonales. Au final, sur les 28 chorèmes de base, seulement 3 semblent pouvoir rendre compte des organisations spatiales qui s'imposent sur les différentes cartes réalisées : à savoir la rupture ou l'interface, les aires en contact et la ligne de partage.

*Document 1*

|                        | POINT                         | LIGNE                          | AIRE                                      | RESEAU                            |
|------------------------|-------------------------------|--------------------------------|---|-----------------------------------|
| maillage               | <br>chef-lieu                 | <br>limite administrative      | <br>Etat, région...                       | <br>centres, limites et polygones |
| quadrillage            | <br>tête de réseau carrefour  | <br>voies de communication     | <br>aire de desserte irrigation, drainage | <br>réseau                        |
| attraction             | <br>points attirés satellites | <br>lignes d'isotropie orbites | <br>aire d'attraction                     | <br>liaisons préférentielles      |
| contact                | <br>point de passage          | <br>rupture, interface         | <br>aires en contact                      | <br>base tête de pont             |
| tropisme               | <br>flux directionnel         | <br>ligne de partage           | <br>surfaces de tendance                  | <br>dissymétries                  |
| dynamique territoriale | <br>évolutions ponctuelles    | <br>axes de propagation        | <br>aires d'extension                     | <br>tissu du changement           |
| hiérarchie             | <br>semis urbain              | <br>limites administratives    | <br>sous-ensemble                         | <br>réseau maillé                 |

**64.1. La table des chorèmes**  
Ou les sept fois quatre = vingt-huit figures de base de toute distribution spatiale, et donc de toute carte.

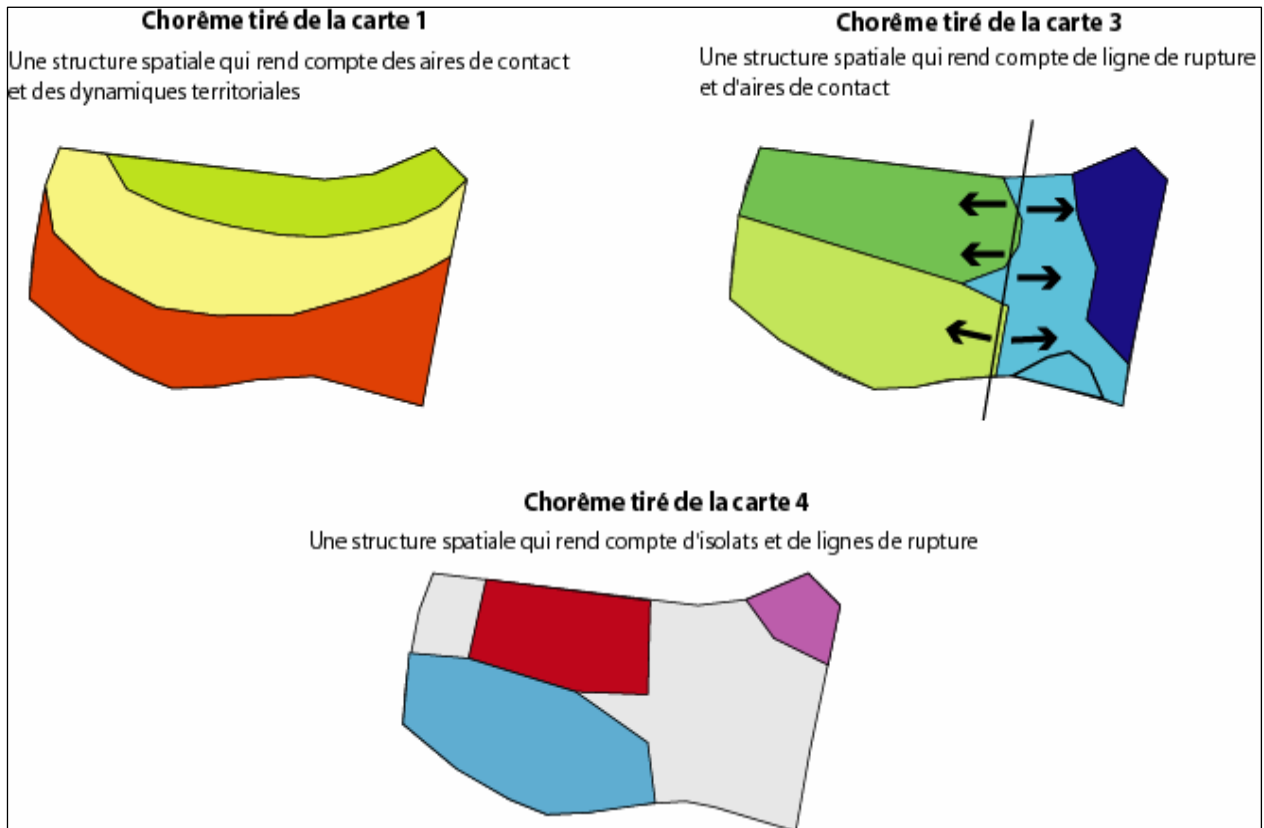
Source : Brunet R., 1987, *La carte, mode d'emploi*, Paris: Fayard/Reclus

Les chorèmes réalisés à partir de 3 cartes de résultats sélectionnées (cartes 1., 3. et 4.) soulignent bien le pouvoir que la carte a d'ancrer une image pour un phénomène étudié. Si le message à faire passer est que les dynamismes démographiques ne concernent que les états de la Sun Belt alors il est clair qu'il faut privilégier uniquement la première carte. En revanche, si la démonstration vise à montrer que la manière dont les Etats-Unis se sont construits territorialement pèse encore



aujourd'hui sur les populations alors il est clair qu'il faut opter pour la carte 3. Il est important de souligner que ces 3 cartes et les chorèmes qui leur sont associés sont quand même 3 images différentes d'une même réalité.

### **Document 2 : CHOREMES DE SYNTHÈSE**



## **Bibliographie**

### **Ouvrages indispensables**

\*\*\*\*Béguin M. et Pumain D., 1994, *La représentation des données géographiques. Statistique et cartographie*, Paris: Armand Colin, Cursus, 192 p.

Cet ouvrage a déjà été recommandé dans les enseignements de DEUG relatifs aux méthodes et aux outils de la géographie. Il est surtout centré sur les questions relatives aux représentations cartographiques des résultats statistiques. A regarder de nouveau pour se remémorer les règles de base de la sémiologie graphique.

\*\*\*\*Zanin C., Trémelo M.-L., 2003, *Savoir faire une carte. Aide à la conception et à la réalisation d'une carte thématique univariée*, Paris : Belin, collection Sup Géographie, 200 p.

### **Ouvrages de références**

\*\*\*Brunet R., 1987, *La carte, mode d'emploi*, Paris: Fayard/Reclus, 270 p.

Plus théorique que l'ouvrage de M. Béguin et D. Pumain, *La carte, mode d'emploi* offre une vue plus synthétique de la démarche cartographique et de l'ensemble des questions que l'on peut se poser face à un document cartographique. A consulter pour remettre son travail cartographique en

perspective.

## **Prolongements possibles**

Dans le cadre d'enseignements intégrés ou de travaux dirigés qui peuvent être articulés sur plusieurs séances et dont un des objectifs est de rendre compte de ce que peut être une démarche complète en Analyse de Données, on peut imaginer au moins deux types de prolongements qui supposent d'engager de nouveaux traitements de l'information sur le tableau élémentaire de données déjà existant. Et ce en ne développant pas le fait que l'on puisse réaliser une Classification Ascendante Hiérarchique (CAH) sur ces mêmes données et confronter graphiquement et cartographiquement les résultats de l'ACP déjà obtenue avec ceux de la CAH.

**Premier type de prolongement possible : celui sous forme de vérifications.** Il s'inscrit dans une démarche de vérification de la structuration de l'information, de sa stabilité. Au cours de ce travail de commentaire de résultats d'une ACP normée, on a pu insister sur le fait qu'un état comme la Californie contribuait à plus de 15% à la structuration du premier facteur. De la même manière, le District of Columbia, autrement dit la capitale fédérale Washington, contribuait à structurer le deuxième facteur à plus de 30%. Pour vérifier si l'importance de ces états dans la construction du premier plan factoriel ne masque pas d'autres phénomènes démographiques, on peut être tenté de les déclarer en « individus supplémentaires » (cf. Documents supplémentaires n°1). On explique alors aux étudiants que ces états seront toujours regroupés en fin de tableau et qu'il convient de vérifier que leurs contributions à la structuration des différents facteurs sont toujours nulles. Pour vérifier si le fait qu'ils ne jouent cette fois aucun rôle dans l'ACP normée modifie ou non la structuration de l'information, on mesure les corrélations des coordonnées des variables sur chacun des premiers facteurs de l'ACP normée qui comprend les 51 états états-uniens et de celle qui vient d'être effectuée et qui traite à part les états de Californie et du District of Columbia. On est alors en mesure de conclure que les relations sont toutes positives et extrêmement significatives sur les 4 premiers facteurs. On fait de même pour les 49 états qui jouent un rôle actif dans les deux ACP normées. On est alors en mesure de conclure qu'il s'agit bel et bien d'une opération de vérification : il n'y a pas de changement important dans la structuration de l'information ; cette dernière est juste plus aérée, plus lisible.

**Deuxième type de prolongement possible : celui sous forme d'approfondissement.** On s'intéresse dans ce cas aux projections des principaux résultats aux trois derniers scrutins présidentiels (cf. Documents supplémentaires n°2). Par rapport à ce qui a été mis à jour dans la représentation de certains groupes de population par un des grands partis politiques présents lors des élections présidentielles de 2000, on veut savoir si les mêmes groupes ont toujours été

représentés par les mêmes familles politiques ou s'il y a eu des changements au cours des trois derniers scrutins présidentiels. On étoffe ainsi la notion de variable supplémentaire en introduisant la dimension temporelle. En effet, on regarde la « projection des résultats des élections de 1996, 2000 et 2004 : ce qui revient à identifier les « trajectoires » des principaux partis qui animent la vie politique américaine. On voit alors que les deux principaux partis que sont les Républicains et les Démocrates ont très peu changé de position à l'intérieur d'un même cadran durant cette période. A l'opposé, la troisième force politique qui peut être soit écologiste soit réformiste s'est positionnée très différemment selon les scrutins, tendant à représenter des



## **Familiarisation avec les commentaires de résultats d'une Analyse en Composantes Principales**

### *Travaux autour des résultats d'une ACP concernant les situations démographiques et politiques des états américains*

---

**Documents :**

- Tableau descriptif des variables sélectionnées pour l'Analyse
- Tableaux des principaux résultats d'une Analyse en Composantes Principales (ACP)
- Graphiques

**Document à rendre :** - 1 représentation cartographique des résultats du premier facteur de l'ACP concernant les différenciations démographiques entre états américains

---

On étudie les caractéristiques de la répartition des populations dans les 51 états des Etats - Unis d'Amérique. Pour ce faire, on dispose d'un tableau qui pour chacun des états américains indique la répartition des populations par tranche d'âge, par « races », par origine étrangère, ainsi que les différentes dynamiques de population. Enfin, pour chacun des états, on dispose des résultats aux dernières élections présidentielles de 2004.

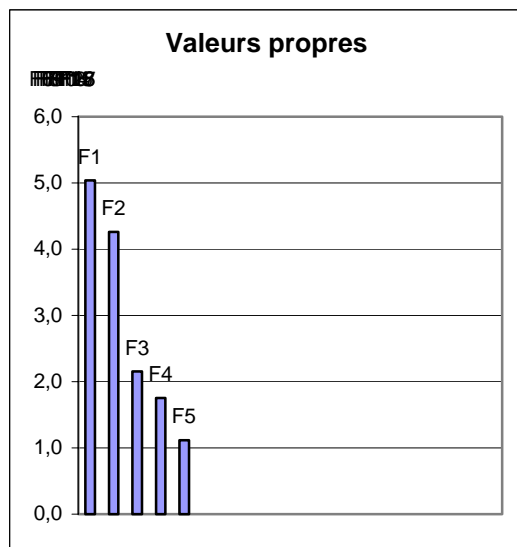
1. A partir des premiers tableaux et des graphiques donnant les principaux résultats d'une Analyse en Composantes Principales (ACP), vous indiquerez, en ayant soin de bien justifier votre réponse, si cette analyse fournit un résumé efficace de l'information initiale.
2. A partir des tableaux suivants et des graphiques donnant les principaux résultats d'une Analyse en Composantes Principales, vous indiquerez les oppositions majeures entre les variables mises en évidence et leurs traductions spatiales. Pour ce faire, vous construirez graphiquement les trois premiers facteurs et leur donnerez à chacun un titre ; vous donnerez également un titre à chacun des cadrans du premier plan factoriel. Que pouvez-vous en conclure sur les caractéristiques et les dynamiques des populations américaines à la fin du XXème siècle ?
3. En fonction de ces caractéristiques démographiques, on tente de comprendre les votes lors des dernières élections présidentielles américaines. Pour ce faire, on a projeté les résultats des candidats des trois principales formations politiques en variables supplémentaires.
  - a. Vous définirez précisément ce que l'on entend par variable(s) supplémentaire(s) dans une ACP.
  - b. Vous caractériserez le plus précisément possible les projections des 3 grandes formations politiques en fonction des caractéristiques démographiques des états américains. Quelles interprétations proposez-vous ? Quelle(s) est(sont) votre(s) principale(s) conclusion(s) sur cette étude ?

4. A partir des résultats contenus dans les seconds tableaux, vous proposerez :
- c. Et réaliserez au moins 2 représentations cartographiques des résultats de l'ACP, en ayant soin de justifier vos choix de variables visuelles, de présentation de la légende et plus généralement d'habillage de votre document.
  - d. de nouvelles démarches ou de nouveaux traitements possibles en fonction des résultats que vous aurez commentés.

### Description des variables retenues pour l'Analyse sur les Etats-Unis d'Amérique

|   |                  |   |
|---|------------------|---|
| <b>Elections présidentielles 2000</b>           | <b>REPU00</b>    | Part des votes en faveur des républicains aux élections de 2000   |
|   | <b>DEMO00</b>    | Part des votes en faveur des démocrates aux élections de 2000   |
|   | <b>VERT00</b>    | Part des votes en faveur des verts aux élections de 2000  |
| <b>Population</b>                               | <b>POPTOT</b>    | Population totale en 2000 (habitants)   |
|   | <b>VPOP90</b>    | Variation de la population de l'Etat (%/1990)   |
|   | <b>DEN90</b>     | Densité de population de l'Etat (hab/miles <sup>2</sup> )   |
| <b>Rural/Urban</b>                              | <b>PURB90</b>    | Part de la population urbaine dans la population totale en 1990 (%)                                       |
| <b>Population résidente par race et origine</b> | <b>BLAN00</b>    | Part de la population blanche dans l'ensemble de la population de l'Etat (%)                              |
|   | <b>NOIR00</b>    | Part de la population noire dans l'ensemble de la population de l'Etat (%)                                |
|   | <b>NATI00</b>    | Part de la population amérindienne dans l'ensemble de la population de l'Etat (%)                         |
|   | <b>ASIA00</b>    | Part de la population d'origine asiatique dans l'ensemble de la population de l'Etat (%)                  |
|   | <b>AUTR00</b>    | Part des autres populations dans l'ensemble de la population de l'Etat (%)                                |
|   | <b>RAC200</b>    | Part de la population se déclarant appartenir à deux races dans l'ensemble de la population de l'Etat (%) |
|   | <b>HISP00</b>    | Part de la population d'origine hispanique dans l'ensemble de la population de l'Etat (%)                 |
| <b>Population par classes d'âge</b>             | <b>&lt;5ANS</b>  | Part des moins de 5 ans dans la population totale de l'Etat (%)   |
|   | <b>517ANS</b>    | Part des 5-17 ans dans la population totale de l'Etat (%)   |
|   | <b>1824ans</b>   | Part des 18-24 ans dans la population totale de l'Etat (%)  |
|   | <b>2544ans</b>   | Part des 25-44 ans dans la population totale de l'Etat (%)  |
|   | <b>4564ans</b>   | Part des 45-64 ans dans la population totale de l'Etat (%)  |
|   | <b>&gt;65ans</b> | Part des plus de 65 ans dans la population totale de l'Etat (%)   |
| <b>Population d'origine étrangère</b>           | <b>ETRA90</b>    | Part des personnes d'origine étrangère dans la population totale en 1980 (%)                              |
|   | <b>VETRA8090</b> | Nombre de personnes d'origine étrangère en 1980   |

## 1. Les valeurs propres



| Qualité de l'analyse | F1    | F2    | F3    | F4    | F5    |
|----------------------|-------|-------|-------|-------|-------|
| Valeur propre        | 5,04  | 4,26  | 2,15  | 1,75  | 1,11  |
| % variance           | 27,98 | 23,67 | 11,96 | 9,73  | 6,19  |
| % cumulé             | 27,98 | 51,65 | 63,61 | 73,34 | 79,54 |

## 2. Les variables.

### a. Coordonnées sur les différents facteurs

| Codes des Variables | F1    | F2    | F3    | F4    | F5    |
|---------------------|-------|-------|-------|-------|-------|
| VPOP90              | 0,39  | 0,49  | -0,22 | -0,20 | -0,16 |
| DEN90               | 0,30  | -0,65 | -0,20 | 0,27  | 0,39  |
| PURB90              | 0,82  | -0,09 | -0,02 | -0,09 | -0,05 |
| BLAN00              | -0,72 | 0,32  | -0,22 | -0,48 | -0,16 |
| NOIR00              | 0,23  | -0,57 | -0,41 | 0,56  | 0,09  |
| NATI00              | 0,00  | 0,55  | 0,14  | 0,15  | 0,65  |
| ASIA00              | 0,50  | -0,12 | 0,81  | 0,13  | -0,21 |
| AUTR00              | 0,76  | 0,28  | -0,19 | -0,40 | 0,22  |
| RAC200              | 0,51  | 0,01  | 0,83  | 0,13  | -0,04 |
| HISP00              | 0,76  | 0,28  | -0,13 | -0,42 | 0,25  |
| ETRA90              | 0,80  | -0,21 | 0,22  | -0,29 | -0,05 |
| VETRA8090           | 0,77  | -0,10 | -0,34 | 0,11  | 0,05  |
| <5ANS               | 0,41  | 0,80  | -0,13 | 0,25  | -0,20 |
| 517ANS              | -0,12 | 0,91  | 0,06  | 0,17  | -0,07 |
| 1824ans             | -0,17 | 0,69  | 0,12  | 0,51  | 0,08  |
| 2544ans             | 0,23  | -0,60 | -0,21 | 0,15  | -0,39 |
| 4564ans             | -0,51 | -0,32 | 0,36  | -0,08 | 0,32  |
| >65ans              | -0,36 | -0,48 | 0,24  | -0,46 | 0,11  |
| REPU00              | -0,40 | 0,75  | -0,04 | 0,12  | -0,03 |
| DEMO00              | 0,39  | -0,81 | -0,02 | -0,13 | 0,00  |
| VERT00              | 0,17  | 0,05  | 0,29  | -0,03 | 0,10  |

**2. Les variables.**

**b. Contributions aux différents facteurs (%)**

| <b>Codes des variables</b> | <b>F1</b> | <b>F2</b> | <b>F3</b> | <b>F4</b> | <b>F5</b> |
|----------------------------|-----------|-----------|-----------|-----------|-----------|
| <b>VPOP90</b>              | 3,06      | 5,61      | 2,33      | 2,34      | 2,16      |
| <b>DEN90</b>               | 1,76      | 9,97      | 1,92      | 4,27      | 13,78     |
| <b>PURB90</b>              | 13,22     | 0,20      | 0,01      | 0,43      | 0,19      |
| <b>BLAN00</b>              | 10,24     | 2,47      | 2,22      | 12,98     | 2,18      |
| <b>NOIR00</b>              | 1,03      | 7,60      | 7,65      | 18,00     | 0,70      |
| <b>NATI00</b>              | 0,00      | 7,06      | 0,97      | 1,23      | 38,27     |
| <b>ASIA00</b>              | 5,01      | 0,32      | 30,50     | 1,01      | 3,85      |
| <b>AUTR00</b>              | 11,42     | 1,91      | 1,61      | 9,24      | 4,34      |
| <b>RAC200</b>              | 5,15      | 0,00      | 31,95     | 0,94      | 0,14      |
| <b>HISP00</b>              | 11,53     | 1,87      | 0,84      | 10,05     | 5,59      |
| <b>ETRA90</b>              | 12,81     | 1,03      | 2,18      | 4,79      | 0,22      |
| <b>VETRA8090</b>           | 11,86     | 0,22      | 5,48      | 0,69      | 0,21      |
| <b>&lt;5ANS</b>            | 3,33      | 15,17     | 0,75      | 3,52      | 3,61      |
| <b>517ANS</b>              | 0,27      | 19,33     | 0,17      | 1,60      | 0,38      |
| <b>1824ans</b>             | 0,56      | 11,02     | 0,65      | 14,94     | 0,51      |
| <b>2544ans</b>             | 1,08      | 8,38      | 2,04      | 1,31      | 13,43     |
| <b>4564ans</b>             | 5,14      | 2,47      | 6,07      | 0,38      | 9,44      |
| <b>&gt;65ans</b>           | 2,54      | 5,37      | 2,67      | 12,28     | 1,01      |
| <b>REPU00</b>              | 0,00      | 0,00      | 0,00      | 0,00      | 0,00      |
| <b>DEMO00</b>              | 0,00      | 0,00      | 0,00      | 0,00      | 0,00      |
| <b>VERT00</b>              | 0,00      | 0,00      | 0,00      | 0,00      | 0,00      |

### 3. Les états américains.

#### a. Coordonnées sur les différents facteurs

| Etats                | Codes des états | F1    | F2    | F3    | F4    | F5    |
|----------------------|-----------------|-------|-------|-------|-------|-------|
| Alabama              | AL              | -1,10 | -0,97 | -0,55 | 1,26  | -0,06 |
| Alaska               | AK              | 0,58  | 4,88  | 1,17  | 3,49  | 2,03  |
| Arizona              | AZ              | 3,22  | 2,54  | -0,93 | -1,89 | 0,86  |
| Arkansas             | AR              | -1,39 | 0,12  | -0,30 | 0,15  | 0,22  |
| California           | CA              | 6,27  | 0,86  | 0,17  | -1,34 | -0,40 |
| Colorado             | CO              | 1,17  | 1,04  | -0,41 | -1,00 | 0,06  |
| Connecticut          | CT              | 0,37  | -1,44 | 0,01  | -1,39 | -0,28 |
| Delaware             | DE              | 0,11  | -1,31 | -0,80 | 0,26  | -0,93 |
| District of Columbia | DC              | 3,97  | -8,44 | -2,08 | 3,02  | 3,34  |
| Florida              | FL              | 1,22  | -1,64 | 0,08  | -2,74 | 0,48  |
| Georgia              | GA              | 1,53  | -0,12 | -2,02 | 2,02  | -1,18 |
| Hawaii               | HI              | 4,77  | -1,56 | 8,96  | 1,74  | -1,24 |
| Idaho                | ID              | -0,54 | 2,77  | -0,36 | 0,08  | -0,08 |
| Illinois             | IL              | 1,47  | -0,15 | -0,42 | -0,12 | -0,48 |
| Indiana              | IN              | -0,90 | 0,01  | -0,29 | 0,01  | -0,72 |
| Iowa                 | IA              | -2,19 | -0,04 | 0,32  | -0,82 | 0,05  |
| Kansas               | KS              | -0,36 | 0,72  | -0,26 | -0,01 | -0,28 |
| Kentucky             | KY              | -2,14 | -0,29 | -0,12 | 0,22  | -0,32 |
| Louisiana            | LA              | -0,52 | 0,36  | -0,55 | 2,47  | -0,04 |
| Maine                | ME              | -3,21 | -1,52 | 0,54  | -1,22 | -0,22 |
| Maryland             | MD              | 1,54  | -1,82 | -0,95 | 1,24  | -0,91 |
| Massachussets        | MA              | 1,05  | -1,36 | -0,38 | -0,38 | -1,00 |
| Michigan             | MI              | -0,99 | -0,33 | 0,07  | 0,44  | -0,64 |
| Minnesota            | MN              | -0,94 | 0,46  | -0,09 | -0,07 | -0,89 |
| Mississippi          | MS              | -1,19 | 0,38  | -0,78 | 2,60  | -0,27 |
| Missouri             | MO              | -1,41 | -0,26 | -0,11 | 0,02  | -0,51 |
| Montana              | MT              | -3,21 | 0,91  | 1,39  | -0,55 | 2,05  |
| Nebraska             | NE              | -1,41 | 0,96  | 0,06  | -0,36 | -0,18 |
| Nevada               | NV              | 3,48  | 1,53  | -0,95 | -1,98 | -0,61 |
| New Hampshire        | NH              | -1,73 | -0,81 | -0,50 | -0,75 | -1,75 |
| New Jersey           | NJ              | 2,03  | -1,80 | 0,09  | -1,06 | -0,03 |
| New Mexico           | NM              | 3,50  | 3,43  | -0,53 | -1,64 | 3,28  |
| New York             | NY              | 2,38  | -1,56 | 0,39  | -1,07 | 0,06  |
| North Carolina       | NC              | -0,17 | -0,50 | -1,12 | 0,68  | -0,25 |
| North Dakota         | ND              | -3,17 | 0,82  | 0,98  | -0,11 | 0,93  |
| Ohio                 | OH              | -1,45 | -0,59 | 0,06  | 0,11  | -0,52 |
| Oklahoma             | OK              | -0,65 | 1,18  | 0,96  | 0,51  | 1,77  |
| Oregon               | OR              | -0,21 | 0,10  | 0,47  | -1,13 | 0,31  |
| Pennsylvania         | PA              | -1,69 | -1,71 | 0,32  | -1,16 | 0,04  |
| Rhode Island         | RI              | 0,81  | -1,58 | -0,09 | -1,78 | -0,33 |
| South Carolina       | SC              | -1,10 | -0,97 | -0,64 | 1,44  | -0,22 |



|               |    |       |       |       |       |       |
|---------------|----|-------|-------|-------|-------|-------|
| South Dakota  | SD | -2,52 | 1,99  | 0,63  | 0,19  | 1,34  |
| Tennessee     | TN | -1,12 | -0,81 | -0,59 | 0,38  | -0,37 |
| Texas         | TX | 3,92  | 2,25  | -1,48 | -0,43 | -0,02 |
| Utah          | UT | 1,38  | 6,29  | -1,18 | 2,15  | -1,84 |
| Vermont       | VT | -3,25 | -1,30 | 0,37  | -0,87 | -0,50 |
| Virginia      | VA | 0,91  | -1,24 | -0,89 | 1,05  | -0,83 |
| Washington    | WA | 0,69  | 0,37  | 0,40  | -0,24 | -0,65 |
| West Virginia | WV | -4,13 | -1,29 | 1,14  | -0,96 | 1,02  |
| Wisconsin     | WI | -1,50 | 0,02  | 0,01  | -0,30 | -0,40 |
| Wyoming       | WY | -2,20 | 1,43  | 0,82  | -0,16 | 1,13  |

### 3. Les états américains.

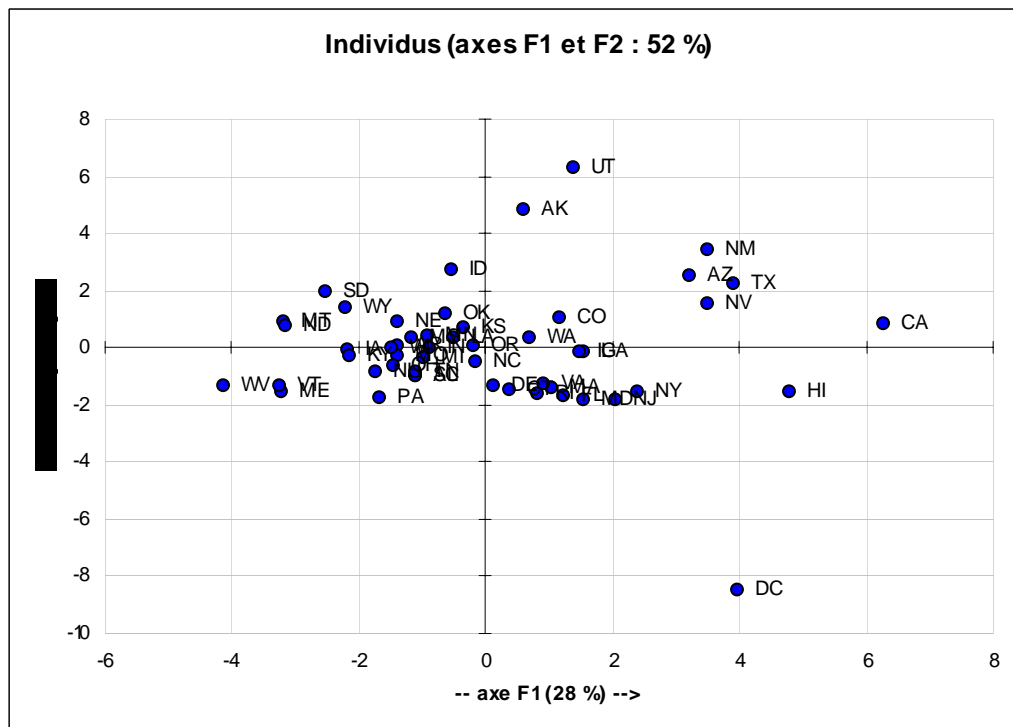
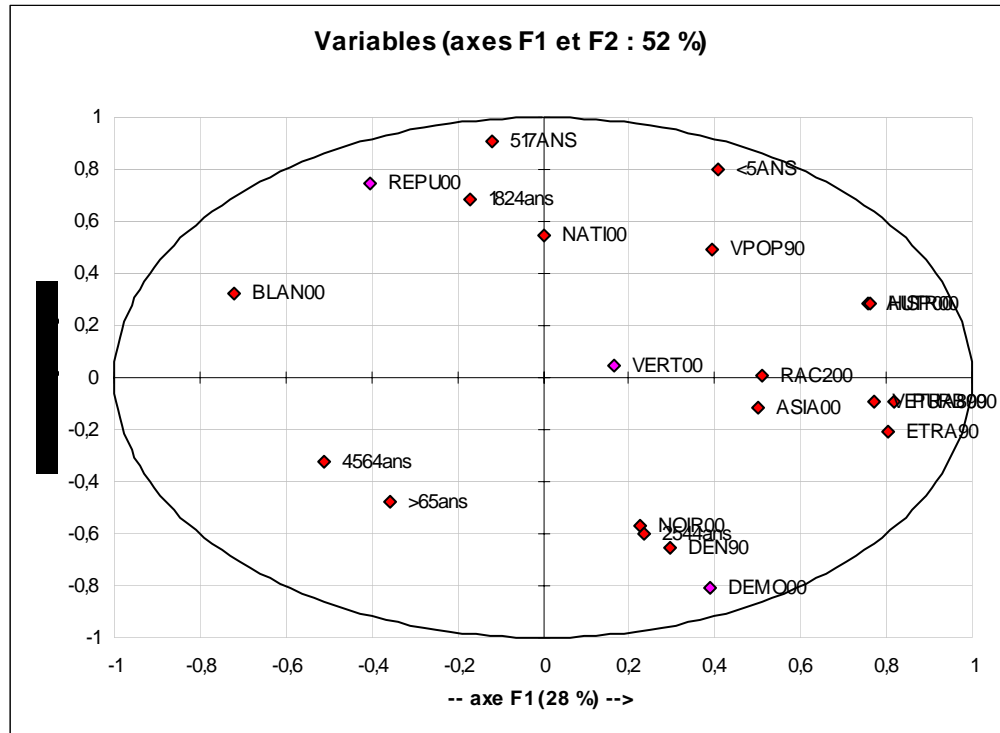
#### b. Contributions aux différents facteurs (%)

| Etats                | Codes des états | F1    | F2    | F3    | F4    | F5    |
|----------------------|-----------------|-------|-------|-------|-------|-------|
| Alabama              | AL              | 0,47  | 0,43  | 0,28  | 1,77  | 0,01  |
| Alaska               | AK              | 0,13  | 10,97 | 1,25  | 13,65 | 7,23  |
| Arizona              | AZ              | 4,04  | 2,98  | 0,79  | 3,99  | 1,30  |
| Arkansas             | AR              | 0,75  | 0,01  | 0,08  | 0,03  | 0,09  |
| California           | CA              | 15,33 | 0,34  | 0,03  | 2,00  | 0,28  |
| Colorado             | CO              | 0,53  | 0,50  | 0,16  | 1,12  | 0,01  |
| Connecticut          | CT              | 0,05  | 0,96  | 0,00  | 2,15  | 0,14  |
| Delaware             | DE              | 0,00  | 0,79  | 0,58  | 0,08  | 1,51  |
| District of Columbia | DC              | 6,14  | 32,80 | 3,94  | 10,23 | 19,58 |
| Florida              | FL              | 0,58  | 1,24  | 0,01  | 8,38  | 0,41  |
| Georgia              | GA              | 0,91  | 0,01  | 3,72  | 4,57  | 2,46  |
| Hawaii               | HI              | 8,87  | 1,11  | 73,09 | 3,40  | 2,73  |
| Idaho                | ID              | 0,11  | 3,54  | 0,12  | 0,01  | 0,01  |
| Illinois             | IL              | 0,84  | 0,01  | 0,16  | 0,02  | 0,41  |
| Indiana              | IN              | 0,31  | 0,00  | 0,08  | 0,00  | 0,92  |
| Iowa                 | IA              | 1,86  | 0,00  | 0,09  | 0,76  | 0,00  |
| Kansas               | KS              | 0,05  | 0,24  | 0,06  | 0,00  | 0,14  |
| Kentucky             | KY              | 1,78  | 0,04  | 0,01  | 0,05  | 0,18  |
| Louisiana            | LA              | 0,10  | 0,06  | 0,28  | 6,83  | 0,00  |
| Maine                | ME              | 4,01  | 1,07  | 0,26  | 1,66  | 0,09  |
| Maryland             | MD              | 0,92  | 1,52  | 0,83  | 1,71  | 1,45  |
| Massachussets        | MA              | 0,43  | 0,85  | 0,13  | 0,16  | 1,75  |
| Michigan             | MI              | 0,38  | 0,05  | 0,00  | 0,22  | 0,73  |
| Minnesota            | MN              | 0,34  | 0,10  | 0,01  | 0,01  | 1,39  |
| Mississippi          | MS              | 0,55  | 0,07  | 0,56  | 7,56  | 0,13  |
| Missouri             | MO              | 0,77  | 0,03  | 0,01  | 0,00  | 0,47  |
| Montana              | MT              | 4,01  | 0,39  | 1,77  | 0,34  | 7,38  |
| Nebraska             | NE              | 0,77  | 0,42  | 0,00  | 0,14  | 0,06  |
| Nevada               | NV              | 4,72  | 1,07  | 0,82  | 4,39  | 0,66  |
| New Hampshire        | NH              | 1,17  | 0,30  | 0,23  | 0,63  | 5,41  |
| New Jersey           | NJ              | 1,60  | 1,50  | 0,01  | 1,27  | 0,00  |

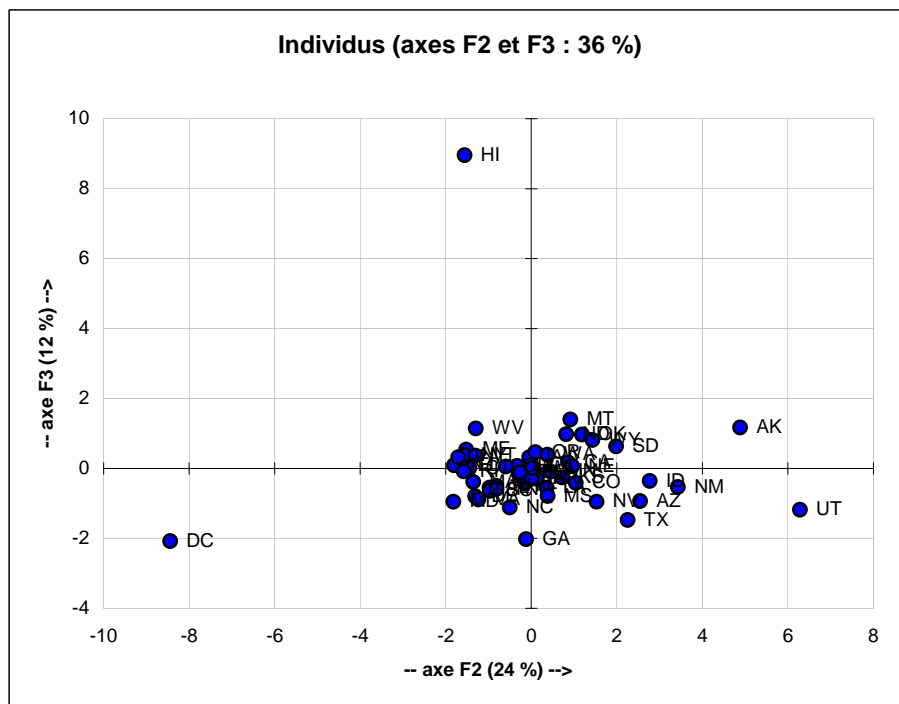
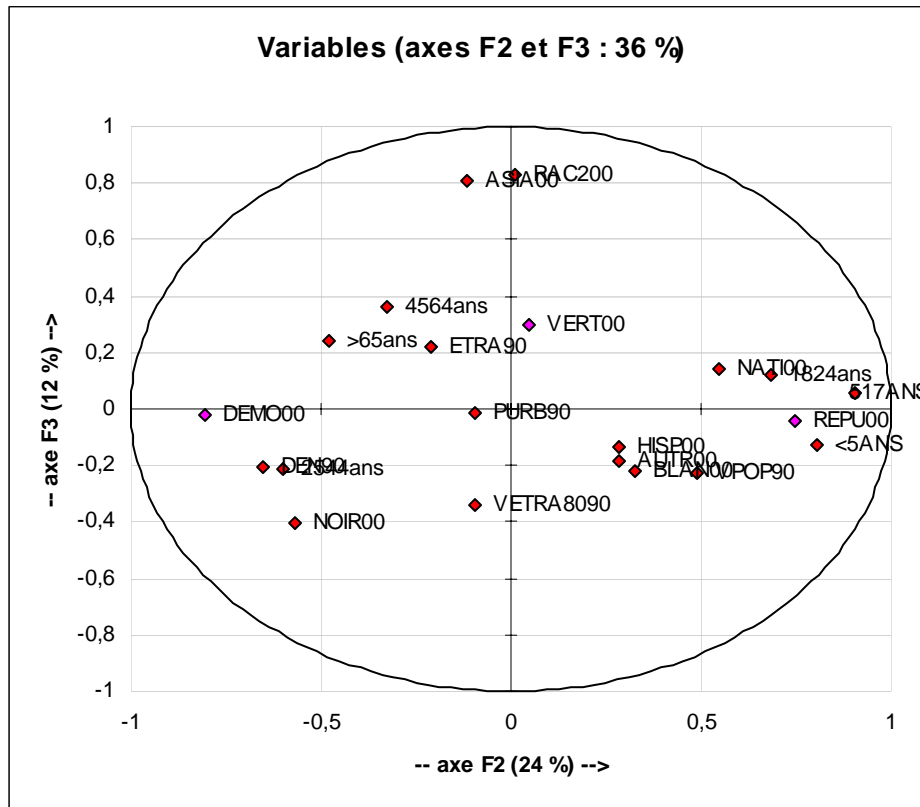
**Myriam BARON, Christine ZANIN**, *Familiarisation au commentaire et à l'interprétation des résultats d'une Analyse en Composantes Principales normées (ACP normée)*, Feuilles de Géographie, I-2005, Feuilles n° 57, 32 pages

|                       |           |      |       |      |      |       |
|-----------------------|-----------|------|-------|------|------|-------|
| <b>New Mexico</b>     | <b>NM</b> | 4,78 | 5,42  | 0,26 | 3,02 | 18,89 |
| <b>New York</b>       | <b>NY</b> | 2,21 | 1,11  | 0,14 | 1,29 | 0,01  |
| <b>North Carolina</b> | <b>NC</b> | 0,01 | 0,12  | 1,14 | 0,52 | 0,11  |
| <b>North Dakota</b>   | <b>ND</b> | 3,92 | 0,31  | 0,87 | 0,01 | 1,53  |
| <b>Ohio</b>           | <b>OH</b> | 0,82 | 0,16  | 0,00 | 0,01 | 0,47  |
| <b>Oklahoma</b>       | <b>OK</b> | 0,16 | 0,64  | 0,85 | 0,30 | 5,52  |
| <b>Oregon</b>         | <b>OR</b> | 0,02 | 0,00  | 0,20 | 1,44 | 0,17  |
| <b>Pennsylvania</b>   | <b>PA</b> | 1,11 | 1,34  | 0,09 | 1,49 | 0,00  |
| <b>Rhode Island</b>   | <b>RI</b> | 0,26 | 1,15  | 0,01 | 3,55 | 0,19  |
| <b>South Carolina</b> | <b>SC</b> | 0,47 | 0,43  | 0,37 | 2,32 | 0,09  |
| <b>South Dakota</b>   | <b>SD</b> | 2,48 | 1,82  | 0,36 | 0,04 | 3,17  |
| <b>Tennessee</b>      | <b>TN</b> | 0,49 | 0,30  | 0,32 | 0,16 | 0,24  |
| <b>Texas</b>          | <b>TX</b> | 5,97 | 2,34  | 1,99 | 0,21 | 0,00  |
| <b>Utah</b>           | <b>UT</b> | 0,74 | 18,19 | 1,27 | 5,18 | 5,96  |
| <b>Vermont</b>        | <b>VT</b> | 4,11 | 0,78  | 0,12 | 0,85 | 0,43  |
| <b>Virginia</b>       | <b>VA</b> | 0,32 | 0,71  | 0,72 | 1,24 | 1,22  |
| <b>Washington</b>     | <b>WA</b> | 0,19 | 0,06  | 0,14 | 0,06 | 0,73  |
| <b>West Virginia</b>  | <b>WV</b> | 6,63 | 0,77  | 1,19 | 1,04 | 1,83  |
| <b>Wisconsin</b>      | <b>WI</b> | 0,87 | 0,00  | 0,00 | 0,10 | 0,29  |
| <b>Wyoming</b>        | <b>WY</b> | 1,89 | 0,94  | 0,61 | 0,03 | 2,26  |

#### 4. Graphiques concernant le premier plan factoriel



### 5. Graphiques concernant le deuxième plan factoriel





Documents supplémentaires n°1

Tableau 1 : les coordonnées des variables sur les premiers facteurs des deux ACP normées

| Variables | F1sans<br>Californie et<br>Washington |                   | F2sans<br>Californie et<br>Washington |                   | F3sans<br>Californie et<br>Washington |                   | F4sans<br>Californie et<br>Washington |                   |
|-----------|---------------------------------------|-------------------|---------------------------------------|-------------------|---------------------------------------|-------------------|---------------------------------------|-------------------|
|           | DC                                    | F1 tous les états | DC                                    | F2 tous les états | DC                                    | F3 tous les états | DC                                    | F4 tous les états |
| PUSA      | 0,52                                  | 0,64              | -0,34                                 | -0,04             | 0,16                                  | -0,19             | -0,07                                 | -0,45             |
| VPOP90    | 0,55                                  | 0,40              | 0,40                                  | 0,50              | 0,01                                  | 0,12              | -0,21                                 | 0,10              |
| DEN90     | 0,36                                  | 0,26              | -0,66                                 | -0,69             | -0,09                                 | 0,22              | 0,36                                  | 0,43              |
| PURB90    | 0,82                                  | 0,80              | -0,06                                 | -0,08             | -0,10                                 | -0,01             | 0,32                                  | 0,05              |
| BLAN00    | -0,61                                 | -0,65             | 0,11                                  | 0,42              | -0,52                                 | -0,44             | 0,45                                  | -0,30             |
| NOIR00    | 0,17                                  | 0,22              | -0,33                                 | -0,64             | 0,80                                  | 0,64              | -0,36                                 | 0,16              |
| NATI00    | 0,01                                  | -0,03             | 0,59                                  | 0,50              | -0,44                                 | -0,21             | -0,31                                 | 0,68              |
| ASIA00    | 0,73                                  | 0,81              | -0,39                                 | -0,12             | -0,07                                 | -0,23             | 0,23                                  | -0,34             |
| AUTR00    | 0,80                                  | 0,84              | 0,32                                  | 0,32              | -0,31                                 | -0,21             | -0,15                                 | 0,17              |
| RAC200    | 0,73                                  | 0,77              | 0,17                                  | 0,21              | -0,47                                 | -0,34             | -0,07                                 | 0,24              |
| HISP00    | 0,78                                  | 0,80              | 0,30                                  | 0,33              | -0,31                                 | -0,22             | -0,22                                 | 0,22              |
| ETRA90    | 0,77                                  | 0,84              | -0,39                                 | -0,19             | -0,31                                 | -0,36             | 0,15                                  | -0,14             |
| VETRA8090 | 0,78                                  | 0,79              | -0,05                                 | -0,15             | 0,26                                  | 0,26              | -0,29                                 | 0,11              |
| <5ANS     | 0,54                                  | 0,41              | 0,67                                  | 0,75              | 0,40                                  | 0,42              | 0,20                                  | -0,15             |
| 517ANS    | -0,01                                 | -0,11             | 0,83                                  | 0,89              | 0,16                                  | 0,11              | 0,33                                  | -0,23             |
| 1824ans   | -0,27                                 | -0,23             | 0,77                                  | 0,65              | 0,34                                  | 0,46              | 0,13                                  | 0,02              |
| 2544ans   | 0,13                                  | 0,28              | -0,70                                 | -0,66             | 0,33                                  | 0,10              | 0,19                                  | -0,46             |
| 4564ans   | -0,52                                 | -0,52             | -0,33                                 | -0,33             | -0,30                                 | -0,38             | -0,45                                 | 0,13              |
| >65ans    | -0,42                                 | -0,41             | -0,37                                 | -0,33             | -0,44                                 | -0,53             | -0,16                                 | 0,18              |

Tableau 2 : matrice de corrélation concernant les coordonnées des variables sur les premiers facteurs des deux ACP normées

|   | F1sans<br>Californie et<br>Washington |                   | F2sans<br>Californie et<br>Washington |                   | F3sans<br>Californie et<br>Washington |                   | F4sans<br>Californie et<br>Washington |                   |
|---|---------------------------------------|-------------------|---------------------------------------|-------------------|---------------------------------------|-------------------|---------------------------------------|-------------------|
|   | DC                                    | F1 tous les états | DC                                    | F2 tous les états | DC                                    | F3 tous les états | DC                                    | F4 tous les états |
| F1sans<br>Californie et<br>Washington<br>DC |                                       |                   |                                       |                   |                                       |                   |                                       |                   |
| F1 tous les états                           |                                       | <b>0,987</b>      |                                       |                   |                                       |                   |                                       |                   |
| F2sans<br>Californie et<br>Washington<br>DC | -0,039                                | -0,122            |                                       |                   |                                       |                   |                                       |                   |
| F2 tous les états                           | -0,017                                | -0,094            | <b>0,949</b>                          |                   |                                       |                   |                                       |                   |
| F3sans<br>Californie et<br>Washington<br>DC | 0,088                                 | 0,097             | 0,008                                 | -0,122            |                                       |                   |                                       |                   |
| F3 tous les états                           | 0,123                                 | 0,083             | 0,209                                 | 0,020             | <b>0,899</b>                          |                   |                                       |                   |
| F4sans<br>Californie et<br>Washington<br>DC | -0,008                                | -0,028            | -0,001                                | 0,144             | 0,002                                 | 0,055             |                                       |                   |
| F4 tous les états                           | 0,013                                 | -0,049            | 0,193                                 | -0,012            | -0,310                                | 0,010             | <b>-0,487</b>                         |                   |

En gras, valeurs significatives (hors diagonale) au seuil  $\alpha=0,050$  (test bilatéral)

**Tableau 3 : matrice de corrélation concernant les coordonnées des états sur les premiers facteurs des deux ACP normées**

|   | F1sans<br>Californie et<br>Washington<br>DC | F1 tous les<br>états | F2sans<br>Californie et<br>Washington<br>DC | F2 tous les<br>états | F3sans<br>Californie et<br>Washington<br>DC | F3 tous les<br>états | F4sans<br>Californie et<br>Washington<br>DC | F4 tous les<br>états |
|---|---|----------------------|---|----------------------|---|----------------------|---|----------------------|
| F1sans<br>Californie et<br>Washington<br>DC | 1   |                      |   |                      |   |                      |   |                      |
| F1 tous les<br>états                        | <b>0,999</b>                                | 1                    |   |                      |   |                      |   |                      |
| F2sans<br>Californie et<br>Washington<br>DC | 0,000                                       | -0,009               | 1   |                      |   |                      |   |                      |
| F2 tous les<br>états                        | 0,171                                       | 0,159                | <b>0,969</b>                                | 1                    |   |                      |   |                      |
| F3sans<br>Californie et<br>Washington<br>DC | 0,000                                       | 0,009                | 0,000                                       | -0,114               | 1   |                      |   |                      |
| F3 tous les<br>états                        | 0,059                                       | 0,061                | <b>0,336</b>                                | 0,234                | <b>0,926</b>                                | 1                    |   |                      |
| F4sans<br>Californie et<br>Washington<br>DC | 0,000                                       | -0,015               | 0,000                                       | 0,120                | 0,000                                       | 0,056                | 1   |                      |
| F4 tous les<br>états                        | 0,072                                       | 0,059                | <b>0,519</b>                                | <b>0,488</b>         | <b>-0,455</b>                               | -0,235               | <b>-0,564</b>                               | 1                    |

*En gras, valeurs significatives (hors diagonale) au seuil alpha=0,050 (test bilatéral)*

Documents supplémentaires n°2

Graphique 1 : Positionnement des principaux candidats aux élections présidentielles

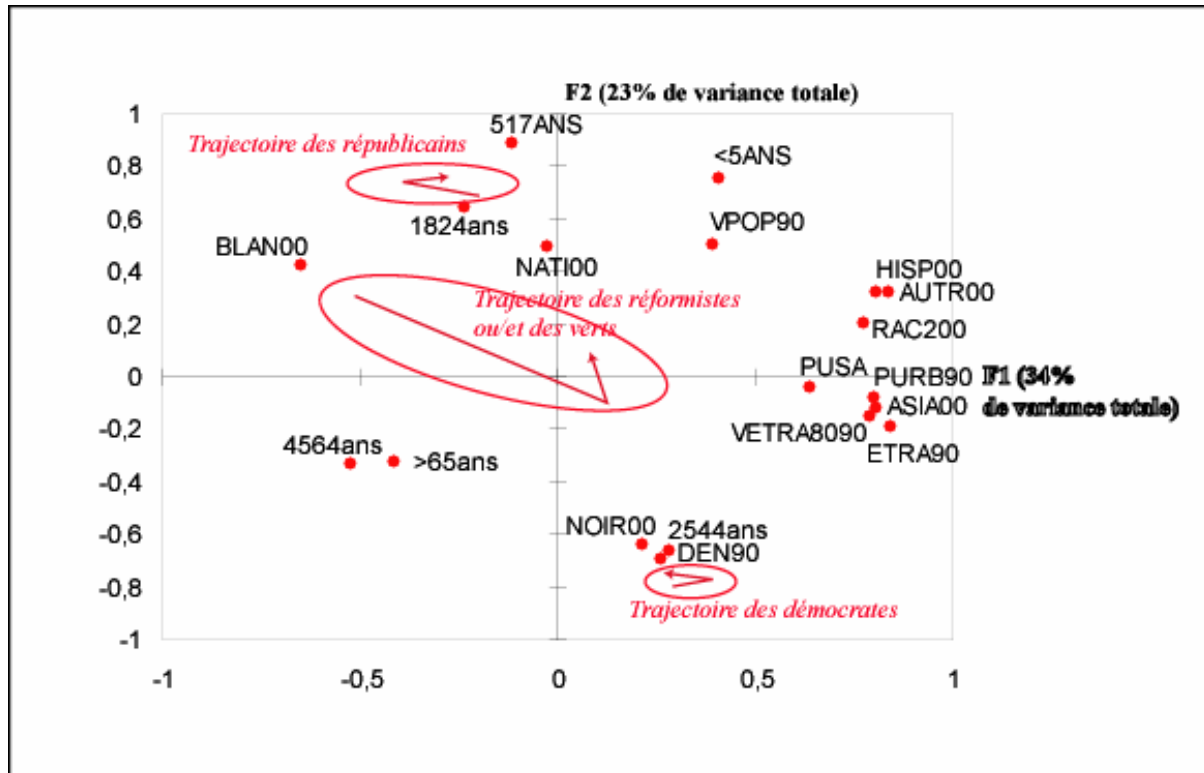


Tableau 2 : matrice de corrélation concernant les coordonnées des candidats sur les premiers facteurs de l'ACP normée

|                           | Candidat républicain 1996 | Candidat républicain 2000 | Candidat républicain 2004 | Candidat démocrate 1996 | Candidat démocrate 2000 | Candidat démocrate 2004 | Candidat réformateur 1996 | Candidat vert 2000 | Candidat réformateur 2004 |
|---------------------------|---------------------------|---------------------------|---------------------------|-------------------------|-------------------------|-------------------------|---------------------------|--------------------|---------------------------|
| Candidat républicain 1996 | 1                         |                           |                           |                         |                         |                         |                           |                    |                           |
| Candidat républicain 2000 | <b>0,956</b>              | 1                         |                           |                         |                         |                         |                           |                    |                           |
| Candidat républicain 2004 | <b>0,965</b>              | <b>0,977</b>              | 1                         |                         |                         |                         |                           |                    |                           |
| Candidat démocrate 1996   | <b>-0,952</b>             | <b>-0,958</b>             | <b>-0,951</b>             | 1                       |                         |                         |                           |                    |                           |
| Candidat démocrate 2000   | <b>-0,920</b>             | <b>-0,985</b>             | <b>-0,953</b>             | <b>0,963</b>            | 1                       |                         |                           |                    |                           |
| Candidat démocrate 2004   | <b>-0,961</b>             | <b>-0,979</b>             | <b>-0,999</b>             | <b>0,956</b>            | <b>0,963</b>            | 1                       |                           |                    |                           |
| Candidat réformateur 1996 | 0,305                     | <b>0,475</b>              | 0,419                     | <b>-0,563</b>           | <b>-0,559</b>           | -0,440                  | 1                         |                    |                           |
| Candidat vert 2000        | <b>-0,535</b>             | -0,453                    | <b>-0,486</b>             | 0,324                   | 0,298                   | 0,448                   | 0,252                     | 1                  |                           |
| Candidat réformateur 2004 | -0,157                    | -0,108                    | -0,125                    | 0,016                   | -0,010                  | 0,078                   | 0,266                     | <b>0,635</b>       | 1                         |

En gras, valeurs significatives (hors diagonale) au seuil alpha=0,050 (test bilatéral)