

Le rôle des méthodes d'analyse des données dans la prise de décision en gestion du territoire

Dr . Harrar sabeha / Bouhanna keltoume / Oubekhti nassira
Université Tlemcen - Alger

Introduction:

L'utilisation des méthodes statistiques touche actuellement tous les secteurs de l'activité humaine. Les méthodes statistiques font partie des connaissances de base de l'ingénieur, du gestionnaire, de l'économiste, du chercheur, etc. Parmi les nombreuses méthodes statistiques, l'Analyse Exploratoire des Données prend un intérêt majeur et connaît une large utilisation de la part des chercheurs dans tous les domaines notamment en économie spatiale.

La gestion du territoire implique la prise en compte et l'analyse du territoire dans toutes ses dimensions (économiques, sociales, démographiques, culturelles, etc.) et repose donc sur un large éventail d'informations. La présentation spatiale des informations implique et mobilise des méthodes statistiques, des modèles et des outils de simulation capables de produire des argumentaires propres à chaque dimension. D'où l'importance des méthodes d'analyse factorielle pour le dépouillement, le traitement et l'interprétation des données.

L'analyse factorielle constitue une famille de méthodes du groupe de techniques de l'analyse des données. Elle s'est développée avec l'amélioration continue de l'outil informatique.

Aujourd'hui, des vastes données d'enquêtes sont dépouillées et, fournissent de grands tableaux qui se prêtent aisément à l'interprétation. Des données issues d'investigations spécifiques sont rassemblées et constituent une masse

importante et apparemment indéchiffrable d'informations mais, qu'on peut désormais traiter sans difficultés.

Le présent article se propose de démontrer le rôle des méthodes d'analyse factorielle dans la prise de décision dans le domaine de la gestion du territoire, le questionnement central de ce papier consiste à identifier les principales méthodes d'analyse des données et le champ d'utilisation de chacune, dans ce sens nous tenterons de répondre à la question suivante :

- En quoi l'analyse de données constitue un potentiel et un intérêt indiscutables dans tous les domaines d'intervention liés à l'espace, de l'aménagement jusqu'à la gestion du territoire?

Pour ce faire notre papier est composé de deux grandes parties, la première qui constitue le cadre théorique de l'article sera consacrée à la problématique de gestion et de planification territoriale, où nous allons apporter quelques précisions sur les concepts y afférent à la gestion locale, et dans un deuxième temps nous présentons les instruments et les outils locaux de gestion. La deuxième partie consacrée à la présentation des différentes méthodes d'analyse factorielle sera accompagnée d'un exemple de l'utilisation de l'Analyse en Composantes Principales pour l'exploitation, l'analyse et l'interprétation du tableau de bord du Schéma Régional d'Aménagement du Territoire (SRAT).

I- la problématique de la gestion du territoire :

Le recours au local devient incontournable dans un contexte de réorganisation du territoire qui vise notamment le contrôle de l'étalement urbain et l'atteinte d'une meilleure équité fiscale au niveau des agglomérations. En effet le local constitue le meilleur échelon pour résoudre les problèmes principaux du territoire (P. AYDALOT 1985, P 68).

Le niveau local, est une communauté d'acteurs publics et privés offre un potentiel de ressources humaines, financières et physiques, d'infrastructures éducatives et institutionnelles dont la mobilisation et la valorisation engendrent des idées et des projets de développements (P.PREVOT .2003, P 77).

Le terme de gestion renvoie à la prise en charge, l'exercice d'une responsabilité sur, l'administration et l'utilisation des moyens pour parvenir à une fin. La définition de la gestion locale prend son origine de là, et signifie ainsi « tirer le meilleur parti du territoire en fonction d'objectifs que l'on s'est assigné ». Dans ce sens la gestion locale se rapproche de l'aménagement du territoire et de la planification spatiale (E. JOS, 2007, p24).

Le développement local s'opère à travers la planification, la programmation, la budgétisation, la mise en œuvre des actions et opérations programmées ainsi que le suivi et l'évaluation. Et c'est au niveau du territoire que s'exécutent toutes ces activités, dont la parfaite connaissance se révèle comme une nécessité tant au niveau de la conception, de l'opérationnalisation, du suivi de la gestion.

L'Aménagement du territoire et la gestion de l'espace en Algérie:

« L'aménagement du territoire est l'art ou la technique (plutôt que la science) de disposer avec ordre, à travers l'espace d'un pays et dans une vision prospective, les hommes et leurs activités, les équipements et les moyens de communication qu'il peuvent utiliser, en prenant en compte les contraintes naturelles, humaines, vois stratégiques » (J.LEVY FRANCOIS GOLAY, 2004, P 103).

La question de l'aménagement du territoire est toujours liée à une volonté de politique c'est-à-dire une vision à long terme pour une planification et une programmation de l'espace. Cette politique selon Eugène Claudius-Petit est « la recherche dans le cadre géographique d'une meilleure répartition des hommes en fonction des ressources naturelles et de l'activité économique » (E. Claudius-Petit, 2003, p 47).

La stratégie spatiale du gouvernement cherche plusieurs objectifs à moyen et long termes. Elle vise à organiser l'espace en étudiant les possibilités de freiner la forte concentration démographique dans les régions littorales déjà surpeuplées, elle oriente aussi le mouvement d'urbanisation et du développement. La politique d'aménagement du territoire cherche la stabilité sociale en mettant de nouveaux instruments pour établir une

hiérarchie urbaine afin de renforcer l'intégration des grandes métropoles nationales dans leur environnement local et régional (M.DAHMANI, 1984, P 45).

La nouvelle politique d'aménagement du territoire pour assurer une gestion de l'espace harmonieuse et efficace s'articule autour des instruments définis par la loi n°01-20 du 12 décembre 2001 relative à l'aménagement de territoire et au développement durable de territoire. Ces instruments sont :

- **Le schéma national d'aménagement du territoire (SNAT) :** il définit la politique générale d'aménagement tout en faisant l'état des lieux pour préciser les orientations globales en fonction des vocations spatiales.
- **Le schéma régional d'aménagement du territoire (SRAT) :** il est élaboré à partir des orientations du SNAT en prenant en considération les spécificités régionales afin de faire participer les acteurs locaux aux actions à initier.
- **Le schéma directeur régional du littoral :** cet instrument souligne la nécessité de protéger et de préserver les espaces sensibles du littoral algérien.
- **Le schéma directeur de protection des espaces agricoles et de lutte contre la désertification :** il évoque les possibilités du développement rural.
- **Le schéma d'aménagement de la wilaya (PAW) :** il précise l'organisation des services publics et les modalités de protection de l'environnement.

La politique de l'aménagement du territoire permet la reconnaissance des spécificités locales, en mettant en place des outils locaux et régionaux suivants :

☞ **Le Conseil National d'Aménagement du Territoire:**

Il oriente la stratégie globale en matière d'aménagement et du développement durable.

☞ **L'Observatoire National d'Aménagement du Territoire:**

Il recueille et diffuse l'information, établit les différentes variantes et scénarios et établit les outils d'aide à la décision. Il évalue l'action régionale.

Les structures régionales:

La conférence régionale d'aménagement et du développement durable du territoire est associée à l'élaboration, la mise en œuvre du S.R.A.T et à l'installation des délégations régionales à l'aménagement du territoire (D.R.A.T).

Le décret du 27 juin 2004, définit les zones de montagnes et des massifs montagneux.

L'importance de la gestion locale dans le développement spatial

C'est au niveau local que sont prises aujourd'hui de nombreuses initiatives prouvant qu'un autre développement est possible.

La pratique de la planification acquiert une importance majeure dans l'organisation du territoire, Les défis de gestion que nous rencontrons aujourd'hui s'articulent autour de la diffusion d'un développement spatial équilibré à travers la redistribution des investissements publics sur tout le territoire (P. Veltz 1996, P 198). Ceci contribue au renforcement du développement, l'amélioration du niveau de vie, la valorisation des potentialités des zones rurales et la mise en œuvre d'un environnement économique et social attractif ainsi que l'atteinte d'une meilleure équité sociale.

II- L'analyse exploratoire des données :

La statistique descriptive s'est enrichie ces dernières années de nombreuses techniques de visualisation de données multidimensionnelles c'est l'analyse des données qui est une des branches les plus vivantes de la discipline statistique. L'analyse de données s'inscrit dans le cadre de la statistique exploratoire multidimensionnelle.

L'analyse des données fondée sur l'usage de l'ordinateur constitue une nouvelle méthodologie que la statistique apporte à la science et notamment aux sciences de l'homme.

Définition:

Selon J-P. Fenelon. "l'analyse des données est un ensemble de techniques pour découvrir la structure, éventuellement compliquée, d'un tableau de nombres à plusieurs

dimensions et de traduire par une structure plus simple et qui la résume au mieux. Cette structure peut le plus souvent, être représentée graphiquement”

Le but principal de l'analyse des données, qui est essentiellement une méthode descriptive, est de décrire, de réduire, de classer et de clarifier les données en tenant compte de nombreux points de vue et d'étudier, en dégagant les grands traits, les liaisons, les ressemblances ou les différences entre les variables ou groupes de variables(L. Lebart , A. Morineau , M. Piron . 2000, P 17).

Les techniques de l'analyse des données:

Le développement de ces techniques procède de l'effet conjugué de plusieurs facteurs : la possibilité de traiter des tableaux de données complexes et de grandes tailles, le regain d'intérêt suscité par de tels tableaux, la possibilité d'exécuter des algorithmes complexes et le désir de s'affranchir d'hypothèses techniques de commodité souvent irréalistes (X. BRY 1995, p 154).

L'analyse des données recouvre deux grands groupes de méthodes :

➤ ***Les Techniques Factorielles***

Elles relèvent de la géométrie euclidienne et conduisent à l'extraction de valeurs et de vecteurs propres. L'analyse factorielle porte sur des *nuages de points* dont on cherche à trouver les directions d'allongement maximal (*axes factoriels*). Elle permet de procéder à des transformations ou à des approximations pour en obtenir une représentation plane.

Les techniques factorielles sont utilisées soit seules, soit conjointement avec la classification automatique.

➤ ***La Classification Automatique.***

Elle porte sur des ensembles d'individus qu'il s'agit de regrouper en catégories homogènes.

Ces méthodes sont caractérisées par le choix d'un indice de proximité et d'un algorithme d'agrégation ou de désagrégation qui permettent d'obtenir une partition ou arbre de classification.

Un bref aperçu historique:

Les méthodes d'analyse de données ont commencées à être développées dans les années 50 poussées par le développement de l'informatique et du stockage des données qui depuis n'a cessé de croître.

Les principes des méthodes d'analyse des données sont anciens, leurs techniques se trouvent dans les travaux de Ch. Spearman (1904) où le concept de « facteur » fut introduit pour la première fois. Appelé « facteur général d'aptitude », ce concept renvoie à une variable explicative cachée.

D'autres travaux fondés sur le même principe furent apparus vers les années 30 (travaux de C. Burt et de L.L. Thurstone) et ont posé le problème de la recherche de plusieurs facteurs.

Il s'agissait déjà de résumer à l'aide d'un petit nombre de facteurs une information multidimensionnelle.

L'analyse en composantes principales est Conçue par Karl Pearson en 1901, et intégrée à la statistique mathématique par Harold Hotelling en 1933.

Les principes de la méthode furent posés par Pearson bien qu'il ne s'agissait pas de l'analyse en composantes principales telle que nous la présentons, mais les idées essentielles de la méthode étaient déjà présentées par cet auteur. Pour Pearson, « les individus colonnes du tableau à analyser étant considérés comme des vecteurs d'un espace à dimensions, on proposait de réduire la dimension de l'espace en projetant le nuage des points individus sur le sous-espace de dimension p_k (k petit fixé) permettant d'ajuster au mieux le nuage ».

D'un point de vue plus récent écrit L. Lebart, l'analyse au composantes principales est «une technique de représentation des données, ayant un caractère optimal selon certains critères algébriques et géométriques spécifiés et que l'on utilise en général sans référence à des hypothèses de nature statistique ou à un modèle particulier » (L. Lebart, A. Morineau, M. Piron. 2000, P 69).

Quant à l'analyse factorielle des correspondances, elle fut introduite par J.P Benzécri (1962). Selon l'auteur « l'analyse des correspondances telle qu'on la pratique en 1977 ne se borne pas à extraire des facteurs de tout tableau de nombres positifs. Elle donne pour la préparation des données des règles telles que le *codage sous-forme disjonctive complète* ; aide à critiquer la validité des résultats, principalement par des calculs de contribution ; fournit des procédés efficaces de discrimination et de régression ; se conjugue harmonieusement avec la classification automatique ».

Les diverses méthodes d'analyse factorielle :

Les méthodes d'analyse factorielle les plus courantes sont l'analyse en composantes principales(ACP), l'analyse factorielle des correspondances (AFC), l'analyse factorielle multiple (AFM), l'analyse factorielle discriminante (AFD).

Nous présenterons dans ce qui suit l'analyse en composantes principales en détailles autres seront présentées brièvement.

L'Analyse en Composantes Principales:

C'est l'une des méthodes les plus employées. Elle est particulièrement adaptée aux variables quantitatives, continues, a priori corrélées entre elles.

L'ACP sert à décrire des tableaux "individus-variables quantitatives" de grande dimension Sa première fonction première en est la réduction des données.

Le tableau se présente sous la forme:

	X_1	X_j	X_p
Individu 1	x_{11}		x_{1j}		x_{1p}
Individui	x_{i1}		x_{ij}		x_{ip}
Individun	x_{n1}		x_{nj}		x_{np}

L'ACP est une transformation des variables de la matrice des données d'origine en un nombre plus restreint de variables significatives appelées facteurs; il s'agit en fait de la factorisation ou de "l'agrégation" qui admet de réduire l'information redondante.

Description sommaire de la technique

L'idée à la base de l'analyse en composantes principales est le remplacement des anciens axes (donc les anciennes variables x_k) par de nouveaux axes (donc par des variables nouvelles C_k). Ces nouvelles variables C_k sont appelées *composantes principales*; elles s'expriment comme combinaisons linéaires des anciennes variables x_1, \dots, x_m .

$$C_k = a_{k1}x_1 + a_{k2}x_2 + \dots + a_{km}x_m$$

Le principe de l'APC consiste donc à réduire le nombre de données tout en rendant compte de la variance observée dans la masse de données.

Ainsi la première composante extraite (C_1) correspondra à la plus grande variance présente dans les variables initiales. Idéalement, 80% ou 70% de la variance expliquée à l'aide de la première composante serait résultat très apprécié du chercheur.

$$C_1 = \hat{a}_1 \text{ var1} + \hat{a}_2 \text{ var2} + \hat{a}_3 \text{ var3} \dots + \hat{a}_k \text{ vark}$$

La variance restante inexpliquée par la première composante est soumise à son tour au même processus d'extraction des composantes. Cette deuxième composante est indépendante de la première et explique la plus grande proportion de variance possible parmi la variance laissée inexpliquée par la première composante.

Le même processus se s'applique pour obtenir plusieurs composante et qui expliquent chacune une proportion de variance de moins en moins importante. La variance totale cumulée des composantes atteindra 100% mais étant donné que l'objectif de l'ACP est la réduction des données, le chercheur doit se décider du nombre de composantes à retenir. Pour ce faire le chercheur dispose de plusieurs critères pour déterminer le nombre de composantes opportun à extraire.

D'une manière générale, on procède aux étapes suivantes :

Etape 1: Sélection des axes et des plans retenus principalement par rapport aux valeurs propres.

Etape 2: Projection des variables et individus dans un plan donné (F1F2 en premier)

- Examen des qualités des représentations (*qlt*) dans le plan pour éliminer les individus mal représentés
- Bilan des contributions (*ctr*) pour un axe afin de donner un sens à cet axe (opposition, tendance ...)
- Topographie des variables et individus afin d'identifier des groupes, des oppositions, des tendances notamment à l'aide de la fonction *s.class*
- Utiliser ses connaissances sur le sujet pour proposer des explications sur les résultats de l'analyse
- Utiliser des individus ou variables supplémentaires ou des profils type (moyenne des H et des F par exemple).

Le nombre d'axes à retenir :

Pour définir le nombre d'axes étudiés, on étudie les valeurs propres obtenues. Chaque valeur propre correspond à la part d'inertie projetée sur un axe donné.

On ne retient donc que les axes avec les plus fortes valeurs propres. Le choix des axes retenus est un peu délicat. On peut donner quelques règles :

- Règle de Kaiser en ACP normée: on ne s'intéresse qu'aux axes avec une valeur propre supérieure à 1 (= inertie d'une variable initiale).
- Règle de l'inertie minimale : On sélectionne les premiers axes afin d'atteindre un % donné d'inertie expliquée (70% par exemple).
- Règle du coude : On observe souvent de fortes valeurs propres au départ puis ensuite de faibles valeurs avec un décrochage dans le diagramme. On retient les axes avant le décrochage.
- Règle de bon sens : On analyse les plans et axes et on ne retient que ceux interprétables.

L'Analyse Factorielle des Correspondances

L'analyse factorielle des correspondances (AFC), ou analyse des correspondances simples, est une méthode exploratoire d'analyse adaptée au traitement de certains types de tableaux rectangulaires de données: les *tableaux de contingence*,

ceux-ci sont obtenus par croisement de variables qualitatives. Cette analyse permet donc de traiter des variables qualitatives et est surtout adaptée à ce type de variables.

Elle a été développée essentiellement par J.-P. Benzecri durant la période 1970-1990. C'est une forme particulière de l'ACP, qui vise à rassembler en un nombre réduit de dimensions la plus grande partie de l'information initiale en s'attachant non pas aux valeurs absolues mais aux correspondances entre les variables, c'est-à-dire aux valeurs relatives. Cette réduction est d'autant plus utile que le nombre de dimensions initial est élevé.

La notion de "réduction" est commune à toutes les techniques factorielles – c'est-à-dire où l'on extrait des facteurs – l'AFC offre la particularité (contrairement aux ACP) de fournir un espace de représentation commun aux variables et aux individus.

L'analyse des correspondances multiples (MCA)

C'est une extension de l'AFC s'applique à des situations où les individus statistiques sont décrits par deux variables nominales. C'est une analyse factorielle appliquée à des tableaux disjonctifs complets; le tableau de contingence à deux critères est remplacé par un tableau à plusieurs critères. L'exemple typique est l'analyse d'un tableau présentant des personnes devant répondre à un ensemble de questions à choix multiples.

L'analyse factorielle discriminante (AFD)

L'analyse factorielle discriminante (AFD) est une méthode descriptive et explicative, apparentée à l'analyse en composantes principales (ACP), s'appliquant à des données quantitatives sur lesquelles est déjà définie une typologie. Par exemple des indicateurs associés à des clients d'une banque classés comme "bons payeurs", "mauvais payeurs" ou "ayant fait faillite".

Le but de la méthode, comme en ACP, est de réduire le nombre de dimensions des données, en recherchant celles suivant lesquelles les classes se séparent le mieux. Les directions factorielles discriminantes successives sont déterminées, tandis que des graphiques factoriels plans

permettent ici encore de visualiser les individus ou les variables. Divers indicateurs et tests sont également calculés, qui permettent de juger de l'intérêt et de la pertinence des résultats obtenus.

III- L'apport de l'analyse factorielle dans l'élaboration des plans de développement

Dans la planification locale la plupart des études des relations entre la structure du territoire et le développement local est à la base de l'élaboration des plans de développement local. Les méthodes usuellement utilisées pour déterminer les principaux axes de développement reposent principalement sur l'analyse factorielle.

En effet le domaine de la gestion locale, connaît de plus en plus de recours à l'utilisation des méthodes d'analyse exploratoire des données rendant ainsi l'exploitation scientifique et statistique des données plus pratique la tâche de planification plus facile. L'intérêt de ces méthodes réside dans le fait que leur principe général est la synthétisation de l'information,

La qualité de toute planification spatiale dépend en grande partie des données de la recherche donc des méthodes mises en œuvre. En effet pour traiter les problèmes complexes impliqués par la planification régionale, le recours à l'analyse exploratoire multidimensionnelle s'impose du fait de l'abondance des informations mobilisées dans l'élaboration des plans et des schémas de planification locale.

Exemple du tableau de bord du SRAT de la région nord ouest :

A l'instar des instruments de gestion du territoire, le SRAT constitue un outil pertinent de suivi et d'observation des situations territoriales. En effet le SRAT tel qu'il est défini par la loi est un outil dont l'Etat se dote pour assurer, sur le moyen et le long terme, la cohérence spatiale de ses différentes actions entre elles. Il donne les grandes orientations pour l'aménagement et la planification des grandes régions, il est conforme au schéma national d'aménagement du territoire.

Pour être représentatif des différentes composantes de l'espace régional et pour qu'il touche à tous les problèmes du territoire, le SRAT devrait s'appuyer sur un tableau de bord

synthétique composé de données fiables, représentatives et exploitables. Cette base de données permet ainsi, grâce à des méthodes de classement statistique, de dégager une typologie des situations locales. L'objectif de la détermination des situations locale est donc de mettre en place un système de vigilance et d'alerte territoriale.

Dans cette partie nous allons démontrer l'utilité et l'apport de l'analyse factorielle dans le traitement, l'exploitation et l'analyse des données contenues dans les tableaux de bord des régions programmes.

1. Elaboration du tableau de bord :

Le tableau de bord comme nous l'avons défini précédemment recueille et diffuse l'information, établit les différentes variantes et scénarios et établit les outils d'aide à la décision. Il évalue l'action régionale.

La finalité de cet instrument est de fournir une base de données originale et précieuse permettant les responsables locaux de saisir les données en temps réel et qui reflètent les différentes facettes de la réalité et des contextes locaux.

Les données sont sous forme d'indicateurs, le choix de ces derniers obéit aux considérations suivantes :

- D'abord l'échelle territoriale de l'observation : le tableau de bord est établie à partir d'un niveau de base qui se situe entre la wilaya ou la commune ou la daïra.
- La pertinence des données : il n'est pas nécessaire de multiplier les indicateurs ; le plus important est qu'ils soient synthétiques et exploitables.

Les indicateurs utilisés :

Le tableau de bord du SRAT est constitué d'un nombre très important d'indicateurs, ceux-ci sont choisis d'une manière à ce qu'ils soient représentatifs de la réalité du territoire et ils touchent donc à tous les aspects du territoire à savoir :

- **démographiques et sociétaux**: cet ensemble regroupe les indicateurs suivants :

Le taux de croissance démographique, la densité de la population, les taux d'urbanisation, la taille moyenne des ménages, indice synthétique de fécondité, taux comparatif de mortalité,...

- **sociaux et d'équipement:** les indicateurs qui peuvent être utilisés dans ce groupe sont : parc de logement, taux d'occupation des logements, taux d'électrification, taux d'alimentation en eau potable, taux de scolarisation, taux d'occupation des classes, nombre de médecins par 1000 habitants ...
- **indicateurs d'infrastructures de transport :** ils concernent les variables suivantes :
Linéaire total des routes nationales, linéaire des chemins de wilaya et des chemins communaux, densité routière...
- **indicateurs économiques :** ceux-ci concernent le taux d'activité, le taux de chômage, l'emploi agricole, l'emploi industriel, nombre de PME, production industrielle,...
- **indicateurs naturels:** cet ensemble regroupe les données relatives à la structure physique du territoire comme le taux de forêts, la superficie de steppe, les terres touchées par la désertification ou l'érosion...

La nécessité du recours aux méthodes factorielles

La masse de données contenue dans le tableau de bord est très abondante. On peut se retrouver avec un tableau constitué de plus de 100 (chaque indicateur contient 30 à 35 données) colonnes et 262 lignes si les observations sont prises par commune ou 91 lignes si les observations sont prises par Daira.

L'analyse de chaque variable individuellement en fonction de toutes les wilayas reste insuffisante car ceci ne permet pas d'élaborer un examen complet qui autorise de constituer une typologie selon les critères de chaque wilaya en fonction des paramètres statistiques.

Pour remédier à ce problème il faut s'approcher davantage des techniques de l'analyse multi variée qui permettent d'aborder la complexité des problèmes géographiques.

En effet, cette masse de données ne peut être dépouillée qu'avec l'analyse factorielle. Cette méthode fournit une vision synthétique tout en gardant l'essentiel de l'information contenue dans le tableau d'origine. Ainsi en employant cette méthode le planificateur ou l'aménageur sera en mesure de dégager une typologie des situations locales et de détecter les principales distorsions susceptibles de se développer entre les différentes composantes de la croissance.

Méthodologie :

Dans cette partie nous avons proposé d'étudier la structure du territoire régional en se basant sur 18 indicateurs pris sur l'ensemble des daïras de la région nord ouest.

Il faut rappeler ici que la région nord ouest est un Espace de Programmation Territoriale défini par la loi 01-20 du 12/12/2001, relative à l'aménagement et au développement durable du territoire. La région nord ouest est composée de 07 wilayets à savoir : Tlemcen, Ain Témouchent, Oran, Sidi Bel Abbés, Mostaganem, Mascara et Relizane.

Elle s'étale sur une superficie de 35737 Km², qui représente 1,6% de la superficie de l'Algérie.

Par rapport aux autres régions, la région nord ouest est limitée par le sud par la région des hauts plateaux et du côté de l'est par la région centre nord. Elle occupe une place stratégique, avec ses frontières avec le Maroc par l'ouest, et ses limites par la méditerranée par le nord qui lui offrent une bande littorale de la wilaya de Tlemcen à celle de Mostaganem de 340 kms.

Les indicateurs de l'étude

POP : population

DEN POP : densité mesurée par le nombre d'habitants par km²

PARC LOG : nombre de logements

TOL : taux d'occupation des logements

L R D'AEP : la longueur totale du réseau d'alimentation en eau

T R AEP : le taux de raccordement en eau potable

T C E : taux de couverture en électricité

TOC 1 ET 2 : taux d'occupation des classes en primaire et en moyen

TOC 3 : taux d'occupation des classes en secondaire

CC : linéaire des chemins communaux

CW : linéaire des chemins de wilaya

RN : linéaire des routes nationales

POP ACTIVE : population active

POP OCCUPEE : population occupée

T C : taux de chômage

PME : effectif des petites et moyennes entreprises

SAU TOTALE : superficie agricole utile

SAU IRRIGUEE : superficie agricole utile irriguée.

Les wilayas concernées par notre étude sont les suivantes :

Mostaganem avec ses 10daïras, Oran avec 09 daïras, Ain Témouchent avec 08daïras et Sidi Bel Abbes avec 15 daïras.

Les données réunies sous forme de matrice de synthèse comportent en ligne 42 villes et en colonnes 18 variables, ce qui correspond à 756 données. L'analyse à l'œil nu de toutes ses données est quasiment impossible sans l'application de la méthode suggérée. Dans cet ordre d'idée l'utilisation de l'analyse des données et plus particulièrement de l'analyse en composantes principales est indiquée pour dépasser les insuffisances et les erreurs de visualisation et d'interprétation de ce stock de données.

Résultats:

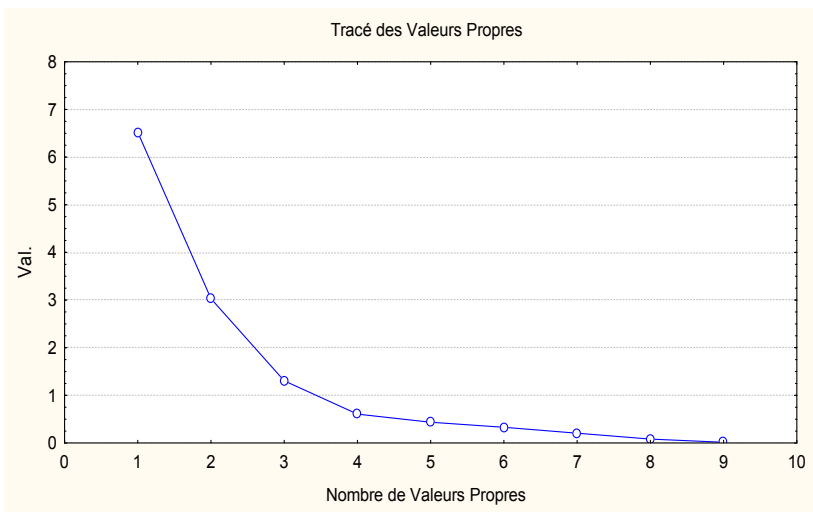
La matrice des données est exploitée à l'aide du logiciel Statistique 07 ; qui nous a fourni les résultats suivants :

1- identification des composantes principales :

Une ACP a permis de synthétiser les 18 items retenus pour représenter la structure de la région nord ouest en 4 facteurs seulement qui expliquent 60,30% de la variance totale ce qui est largement significatif dans un cadre exploratoire.

Tableau 1 : valeurs propres

ANALYSE FACTOR.		Extraction: Factoris. Axe Princ.		
Valeur	ValPropr	% Total Variance	Cumul ValPropr	Cumul %age
1	6,520640	36,22578	6,52064	36,22578
2	3,028834	16,82685	9,54947	53,05263
3	1,305775	7,25431	10,85525	60,30694



- Le premier axe factoriel explique à lui seul 36,22% de la variance totale.
- Le deuxième axe véhicule 16,82% de l'information initiale.
- Quant au troisième axe, il explique 7,25.

Composition et interprétation des axes factoriels :

L'examen du tableau des contributions des variables nous permet d'identifier les variables qui ont un rôle dominant dans la formation d'un axe factoriel.

ANALYSE FACTOR.			
Extraction: Factoris. Axe Princ. (Poids marqués > ,700000)			
Variable	Fact. 1	Fact. 2	Fact. 3
POP	,999552	,007741	,033095
DENPOP	,879013	,042362	,321118
PARCLOG	,998735	-,066011	-,023094
TOL	,140382	,850776	,380738
CC	,033530	,588749	-,065207
CW	-,102458	,519908	,087430
RN	,114589	-,042590	,345266
LRAEP	,750328	,142258	-,454128
TRAEP	-,003676	,248973	,017892
POPACTIV	,938368	-,138326	-,146728
POPOCCUP	,940791	-,138570	-,149566
TC	,132789	-,055649	-,171822
TOC1ET2	,373480	,215183	,548932
TOC3	,385435	,295841	,344904
TCE	,029753	-,820596	,356186
PME	,886963	-,096969	-,086923
SAU	-,530382	-,131357	,041070
SAUIRRIG	-,061151	,850023	-,297647
Var Expl	6,520640	3,028834	1,305775
Prp. Tot	,362258	,168269	,072543

Le premier axe factoriel dégagé exprime lui seul 36,22% de la variance totale et est représenté par les items suivants : population, densité de la population, le parc de logement, longueur du réseau d'alimentation en eau potable, population active, population occupée, et le nombre de PME.

D'après les variables qui lui sont très fortement corrélées, cet axe reflète le potentiel socio économique de la région en termes de population, et d'activité.

On remarque aussi une corrélation positive plus au moins forte des variables relatives aux taux d'occupation des classes tout cycle confondu, ce qui rend compte de la pression en classe. Par contre la variable SAU est corrélée négativement avec l'axe, ce qui nous amène à déduire de la faible présence de l'agriculture dans cet axe.

Cet axe est résumé dans le facteur «Potentiel démographique et dynamique économique».

Le deuxième axe factoriel qui représente 16,82% des nuages de points, regroupe les items

Relatifs à : taux d'occupation de logements, taux de couverture en électricité, et la superficie agricole utile irriguée.

Cet axe est corrélé positivement avec CC ET CW et négativement avec le taux de couverture d'électricité ; ceci peut traduire le caractère rural de l'axe où l'on remarque une forte présence des surfaces agricoles.

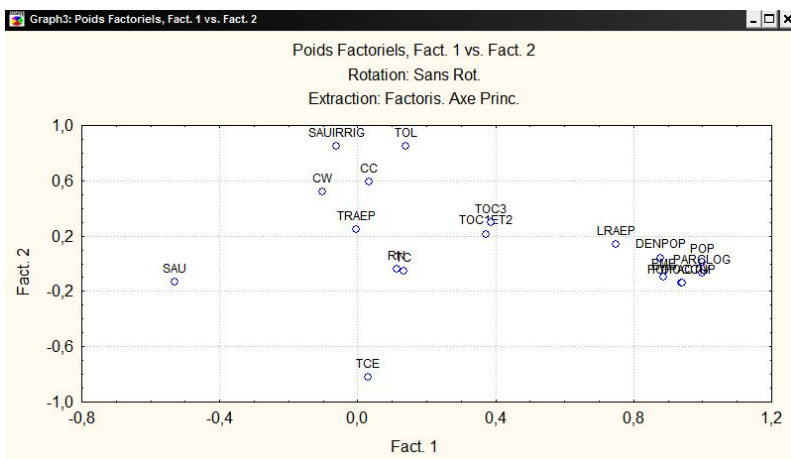
Nous l'avons donc désigné par le facteur «potentiel agricole».

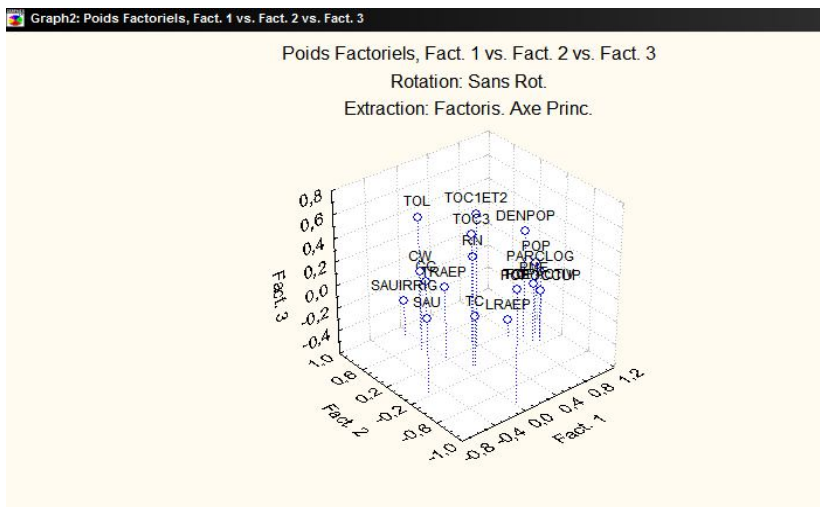
Le troisième axe, représente 7,25% de l'inertie totale, regroupe quant à lui les items suivants :

Routes nationales, taux d'occupation du logement et le taux d'occupation des classes. Ces indicateurs sont corrélés positivement avec l'axe, ils rendent compte de la pression démographique et du fort potentiel d'infrastructures de transport.

Il a été qualifié par le facteur «pression démographique et potentiel infrastructurel»

Pour mieux visualiser l'interprétation des différents axes, nous les représentons graphiquement (figure ci-dessus) en fonction des variables initiales déterminantes.





Interprétation de la distribution des individus par axe factoriel

Pour visualiser la distribution des individus qui sont ici les daïras de la région, nous utilisons le tableau des coordonnées de chaque individu sur l'axe.

Le premier axe :

D'après le tableau on remarque qu'il y a une seule ville bien présentée dans l'axe N°1 à savoir la ville d'Oran. En suite la ville de Sidi Bel Abbés vient en deuxième position avec une contribution moyenne. Les chefs lieux des autres wilayas (Mostaganem et Ain Témouchent) sont très faiblement présentés par cet axe.

On peut conclure que la ville d'Oran se distingue clairement dans l'axe 1, celui qui reflète à la fois son potentiel économique et sa forte demande sociale.