

Analyse de Données

Table des matières

INTRODUCTION	1
Analyse du sujet	1
Mise en oeuvre	1
DÉFINITION DES VARIABLES	1
DÉFINITIONS DES INDIVIDUS	2
LA MÉTHODE	3
Choix d'une ACP	3
Choix de l'ACP Centrée Réduite	3
LA MATRICE CENTRÉ RÉDUITE	5
Calcul	5
LA MATRICE DES CORRÉLATIONS	6
Calcul	6
LA MATRICE DES CORRÉLATIONS	6
Interprétation	7
VALEUR PROPRE	8
Rappel	8
VALEURS PROPRES	8
Interprétation	8
LA MATRICE DE SATURATION	9
Calcul	9

LA MATRICE DE SATURATION	9
Interprétation	9
LE CERCLE DES CORRÉLATIONS	10
Rappel	10
LE CERCLE DES CORRÉLATIONS SELON LES AXES F1 ET F2	10
Interprétation	10
ANALYSE DES INDIVIDUS	12
Calcul	12
COORDONNÉES DES OBSERVATIONS ET QUALITÉ PAR RAPPORT À L'AXE	12
Interprétation axe F1	14
Interprétation axe F2	15
Interprétation générale	15
CONCLUSION	16
Problématique	16
Les limites	16
Ouverture	16
Un élément de Réponse ?	16
SYNTHÈSE PERSONNEL	18
Jérôme	18
Lionel	18
Loan	18
Solène	18
Yann	18
Équipe	18

1.Introduction

Analyse du sujet

Passionné de la géopolitique, notre groupe souhaitons nous intéresser sur les nouvelles puissances mondiales face aux anciennes. Étudiant de l'EFREI notre étude c'est particulièrement établie sur la répartition des technologies de communication de chaque pays. Au premier abord nous pouvons penser, que le développement d'un pays peut être calculé par rapport aux technologies de l'informations exploité dans celui-ci.

Dès lors demandons nous si les ressources mis en oeuvres dans les moyens de communications d'un pays reflète sa croissance ?

Mise en oeuvre

Définition des variables

Nom de la variable	Unité	Commentaire
Les journaux	à l'unité	Le nombre de quotidiens régionaux et nationaux disponible par pays
Les ventes de journaux	en million	Le nombre de ventes des journaux par mois
Les stations de radio	à l'unité	Le nombre de station de radio disponible par pays
Le téléphone	en million	Le nombre d'utilisateur de téléphone
Le téléphone portable	en million	Le nombre d'utilisateur de téléphone portable
Le nombre d'opérateur	à l'unité	Le nombre d'opérateur téléphonique par pays
La télévision	en million	Le nombre de poste de télévisions par pays (chiffre de 2006)
Les stations de télévisions	à l'unité	Le nombre de chaînes proposées par pays
Internet	en million	Le nombre d'utilisateur d'internet par pays
Le nombre de fournisseurs internet	à l'unité	Le nombre de fournisseur d'accès internet (FAI) par pays.

Définitions des individus

Le globe possède 196 pays, la première étape de notre démarche est de lister l'ensemble des pays. Néanmoins certains pays comme la Corée du Nord, la Mauritanie ou encore l'Albanie ne recense aucune informations par négligence ou encore par confidentialité (ndlr Corée du Nord). Cependant la sélection des 52 pays opéré (tableau ci-joint) représente 95,2 % de la population mondiale répartie sur les 5 continents. Les pays ont été également sélectionné pour leurs diversités économiques. Nous pouvons considérer cette étude mondiale.

Afrique	Amérique	Asie	Europe	Océanie
Afrique du Sud	Argentine	Afghanistan	Allemagne	Australie
Algérie	Bolovie	Arabie Saoudite	Autriche	
Angola	Brésil	Arménie	Biélorussie	
Bénin	Canada	Birmanie	Bulgarie	
Burkina Faso	Chili	Brunei	Croatie	
Cameroun	Colombe	Cambodge	Danemark	
Congo	Cuba	Chine	Espagnie	
Cote d'Ivoire	Equateur	Corée du Sud	France	
Egypte	Etats-Unis	Inde	Grèce	
Ethiopie	Mexique	Irak	Irlande	
Maroc		Iran	Portugal	
		Israël	Royaume-Uni	
		Japon		
		Koweït		
		Liban		
		Russie		
		Turquie		

2. La Méthode

Choix d'une ACP

Une ACP permet d'explorer les liaisons entre variables et individus. Elle réduit l'information en un nombre de composantes plus limitées que le nombre initial de variables afin de rendre l'analyse plus simple. Le but est de faire disparaître les «redondances» entre les variables.

Notre analyse comporte 52 individus observés sur 10 variables quantitatives, nous avons donc beaucoup de variables à traiter. **L'emploi d'une ACP** nous a paru la plus convaincante pour définir un lien entre les pays et les technologies de communication.

Nous utilisons **le logiciel XLSTAT**, un module excel de la gamme Office de Microsoft, spécialisé dans le développement d'une ACP.

Choix de l'ACP Centrée Réduite

Il s'est ensuite posé la question de déterminer si nous allions faire une ACP centrée ou une ACP centrée réduite. Nous avons donc, dans un premier temps comparé la moyenne et l'écart-type pour chaque variable.

Rappel des formules

Calcul de la moyenne

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Calcul de la variance

$$V(X) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2$$

Calcul de l'écart type (covariance)

$$\sigma_x = \sqrt{V(X)}$$

Variable	Observations	Obs. avec données manquantes	Obs. sans données manquantes	Minimum	Maximum	Moyenne	Ecart-type
téléphone	52	0	52	0,015	368,000	20,944	55,941
TV	52	0	52	0,010	312,000	21,044	52,796
téléphone portable	52	0	52	0,044	461,000	39,859	77,969
Internet	52	0	52	0,031	231,000	18,672	40,678
Journaux	52	0	52	3,000	142,000	28,750	26,696
Stations de Radio	52	0	52	2,000	13804,000	701,673	2036,974
Stations de Télé	52	0	52	1,000	10500,000	507,327	1785,980

Variable	Observations	Obs. avec données manquantes	Obs. sans données manquantes	Minimum	Maximum	Moyenne	Ecart-type
Nb d'opérateur	52	0	52	1,000	47,000	6,115	6,839
Nb de fournisseur internet	52	0	52	1,000	357,000	20,115	49,938
Vente de journaux	52	0	52	0,120	17,300	1,627	2,814

On note une grande différence entre la moyenne et l'écart type pour chaque variable. Pour la variable «Station de Télé» par exemple, la différence est de 1335,3. Ensuite, la différence entre les valeurs maximales et minimales est de même norme, toujours pour la même variable, on note une différence maximum-minimum de 13802. Pour finir, les variables sont exprimées avec des unités différentes (million ou à l'unité).

C'est pourquoi le choix d'effectuer une ACP centrée réduite s'impose de lui même.

3. La Matrice Centrée Réduite

Calcul

Une matrice centrée réduite est une matrice permettant de recentrer et rendre homogène les valeurs de la matrice A.

Elle est justifiée par l'usage des unités différentes et est calculé à travers le centrage de la matrice (établie à partir de la moyenne des variable) et réduite pour éviter les fortes dispersions pouvant influencer sur notre étude et donner à certaines variables une importance trop importante (cf ACP centrée réduite).

Formule :

$$R = \frac{1}{n} A^t \cdot A$$

4. La Matrice Des Corrélations

Calcul

Calcul de la matrice de corrélation :

$$R = \frac{1}{n} {}^t A^s \cdot A^s$$

La matrice R est carrée et symétrique dont l'ordre est égal au nombre de variables (ici 10). Ceci consiste un premier élément de vérification. Les coefficients sont compris entre -1 et 1. La variable de corrélation permet de définir des liens entre chaque variable, ces liens peuvent être de deux natures soit connoté positivement (corrélation) ou alors antagoniste (anti-corrélation).

La matrice des corrélations

Variables	téléphone	Poste TV	téléphone portable	Internet	Journaux	Stations de Radio	Stations de Télé	Opérateur	FAI	Vente de journaux
téléphone	1	0,976	0,958	0,848	0,558	0,409	0,564	0,483	0,517	0,960
Poste TV	0,976	1	0,951	0,905	0,681	0,555	0,658	0,596	0,654	0,953
téléphone portable	0,958	0,951	1	0,878	0,649	0,475	0,663	0,557	0,577	0,969
Internet	0,848	0,905	0,878	1	0,772	0,748	0,753	0,825	0,861	0,872
Journaux	0,558	0,681	0,649	0,772	1	0,684	0,624	0,761	0,755	0,625
Stations de Radio	0,409	0,555	0,475	0,748	0,684	1	0,758	0,834	0,935	0,452
Station de Télé	0,564	0,658	0,663	0,753	0,624	0,758	1	0,727	0,837	0,600
Opérateur	0,483	0,596	0,557	0,825	0,761	0,834	0,727	1	0,930	0,545
FAI	0,517	0,654	0,577	0,861	0,755	0,935	0,837	0,930	1	0,561
Vente de journaux	0,960	0,953	0,969	0,872	0,625	0,452	0,600	0,545	0,561	1

Nous pouvons remarquer que la matrice de corrélation ne possède aucune variable n'ayant de résultats négatives. De plus le minimum est atteint par la corrélation [téléphone, Station de Radio] est égal à 0.409, ce résultat est surprenant de faite par un minimum élevé. Dès lors on peut dénombrer trois type de corrélation : «*inexistante*», faible (par rapport aux autres) et *forte*».

Interprétation

- Le téléphone, la télévision, le téléphone portable, la vente de journaux ont une forte corrélation, ces derniers expriment un domaine connu depuis des millénaires qu'est la communication. Avec l'ère pré-informatique se sont ses quatre grands moyens usités par les consommateurs.

On peut signaler que la bulle internet est la grande invention post informatique en matière de communication pouvant ainsi expliquer qu'il ne soit pas corrélé avec ses moyens d'une autres générations.

- La forte corrélation télévision, téléphone portable, Internet et la vente de journaux sont les moyens de communications utilisés également par l'homme vivant dans la révolution informatique.

Le développement des packs box Haut débit par les fournisseur d'accès internet, téléphone, internet, télévision peuvent aussi incité le consommateur à posséder les trois services.

- La forte corrélation fournisseurs d'accès internet (FAI) et le nombre d'opérateurs est explicable aisément du faite que les opérateurs téléphoniques et les FAI sont les mêmes entreprises.
- La forte corrélation entre fournisseurs d'accès internet et les stations de radio semble abstraite. Néanmoins leur valeur de corrélation est de plus de 0,935 soit l'une des plus élevée !

Le développement des radios sur internet expliquerait un nombre croissant de radio.

- On peut noter que le nombre d'opérateurs ne dépend pas du nombre d'accès aux téléphones, contrairement aux fournisseurs d'accès internet (FAI) qui possède une bonne corrélation avec l'accès à Internet.

A signaler que le téléphone est une invention du XIXème siècle, alors qu'internet date de la fin du XXème siècle. La confrontation des époques peut probablement expliquer l'opposition entre les FAI et internet contre le téléphone et les opérateurs. Le produit téléphones existant depuis un siècle, la concurrence c'est mutuellement détruite.

- Le nombre de chaînes (station de télé) ne dépend pas fortement de la distribution de postes de télévisions.

Les habitants de pays riches peuvent posséder plusieurs postes de télévisions.

- Il n'y a pas de forte corrélation entre le nombre de journaux d'un pays et la vente de journaux total de ce pays. Le financement de ces journaux ne dépendrait-il pas de leurs ventes?

La presse gratuite dévalorise les chiffres de la vente de la presse libre.

- Les variables journaux et station de TV ne possède pas de corrélation très forte avec les autres variables. On peut s'attendre à une indépendance de ces deux variables.

5. Valeur Propre

Rappel

R est une matrice symétrique, il est possible de la diagonaliser et ainsi définir les valeurs propres de chaque variable.

Formule à résoudre : $\det (R - \lambda I) = 0$

Nous obtenons des valeurs propres toutes positives. **Cette étape permet de déterminer la quantité d'information portée par chaque axe.**

Nous avons ici un deuxième élément de vérification : la somme des valeurs doit être égale à 1 et le nombre de valeurs propres égales à la trace de la matrice R soit le nombre des variables ici 10.

Valeurs propres

	F1	F2	F3	F4	F5	F6	F7	F8	F9
Valeur propre	7,517	1,550	0,395	0,245	0,158	0,059	0,039	0,028	0,007
Variabilité (%)	75,175	15,500	3,948	2,448	1,576	0,591	0,393	0,283	0,066
% cumulé	75,175	90,675	94,622	97,070	98,646	99,236	99,630	99,912	99,979

Interprétation

Chaque valeur propre est représentée – par rapport à chaque axe – par la Qualité Globale d'Explication (QGE). Notre axe F1 représente environ 75% de l'information totale de notre étude (avec une valeur propre $\lambda_1 = 7,5175$). **Il est plus que représentatif** – et pourrait quasiment suffire à lui tout seul à faire notre analyse – de l'information générale. L'axe F2, représente 15,5% de l'information. On peut supposer qu'elle représente approximativement les 2 variables qui ont des corrélations moins intéressantes avec les autres (Stations de Télé et Journaux), évoquées dans la matrice de corrélation.

Les 8 valeurs propres qui suivent sont trop peu importantes pour être prises en compte. **En effet, le pourcentage cumulé de ces 2 variables est d'environ 90,6%**, ce qui est plus que suffisant pour faire une analyse complète de nos données.

6. La Matrice De Saturation

Calcul

La matrice des saturations permet de définir les corrélations entre les variables et les facteurs.

Formule :

$$S = \frac{1}{n} {}^t A^s \cdot F \cdot \Delta^{-1/2}$$

De ce fait, elle constitue un indice important de la proportion des informations des facteurs sur les variables.

La matrice de saturation

	F1	F2
téléphone	0,847	-0,511
TV	0,921	-0,353
téléphone portable	0,893	-0,417
Internet	0,979	-0,028
Journaux	0,818	0,191
Stations de Radio	0,784	0,532
Stations de Télé	0,826	0,254
Nb d'opérateur	0,833	0,441
Nb de fournisseur internet	0,876	0,458
Vente de journaux	0,877	-0,449

Interprétation

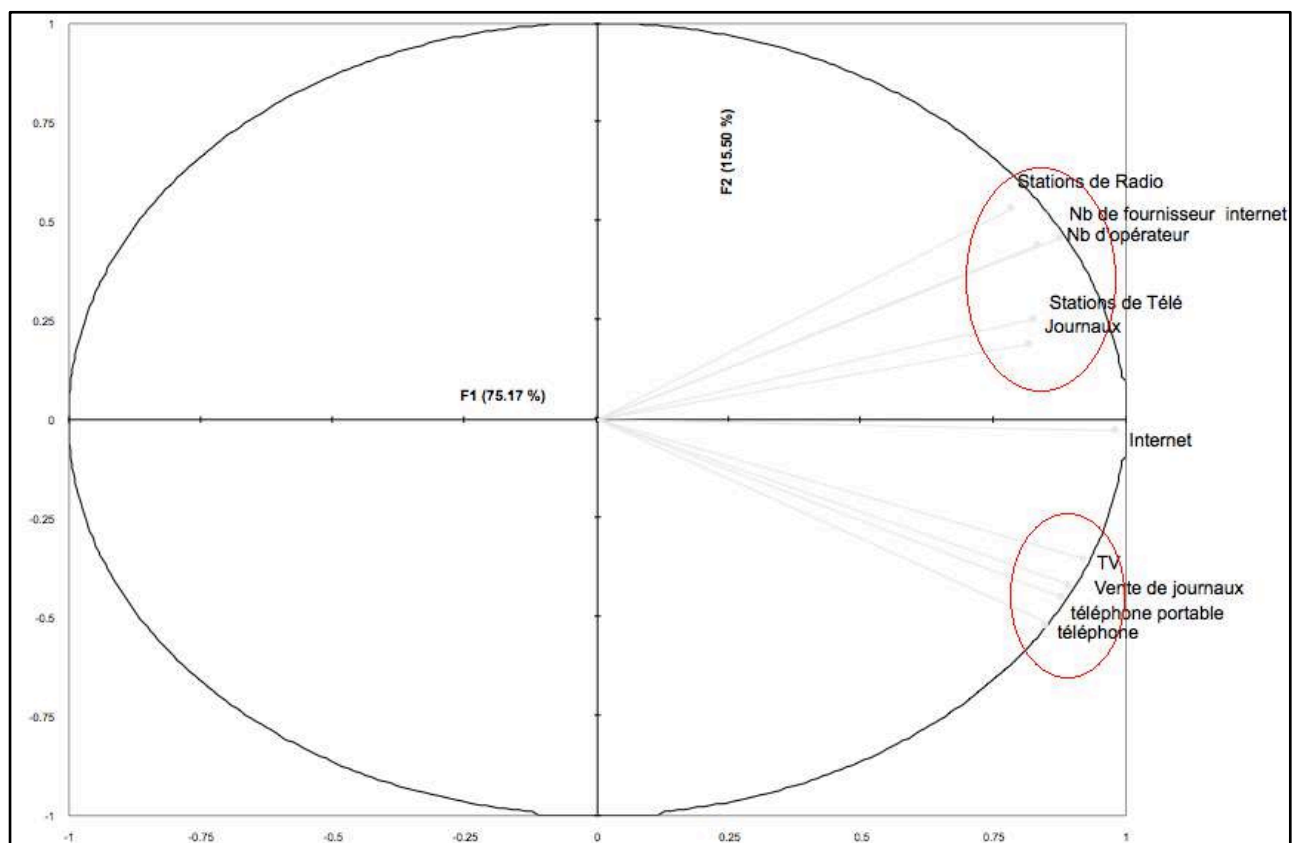
Une analyse rapide de cette matrice (comprise entre 1 et -1) nous fait remarquer que la totalité des variables ont en forte corrélation avec l'axe F1 (qui possède une QGE très importante), seul les stations de radio se distinguent : elle est la moins forte sur l'axe F1 et la plus importante sur l'axe F2.

7. Le Cercle Des Corrélations

Rappel

L'étude graphique ne doit porter que sur les variables se trouvant proches du cercle des corrélations, c'est-à-dire celles qui sont suffisamment représentées. Dans notre cas, elles sont toutes proches du cercle. Seul les variables « station de télévision » et « journaux » sont légèrement plus éloignées du cercle étant donné leur indépendance déjà commentée dans les valeurs propres.

Le cercle des corrélations selon les axes F1 et F2



Interprétation

On remarque une opposition entre les variables {« Station de Radio », « Nb de fournisseur internet », « Nb d'opérateur », « Station de Télé », « Journaux »} et {« TV », « Téléphone portable », « Vente de journaux », « Téléphone »}.

- Le groupe au dessous de l'axe F1 représente l'utilisation des services (le matériel) et celui du dessus, les offres de services.

On voit, également, que la diffusion internet est neutre (séparée des deux groupes), on ne peut la regrouper ni dans les offres, ni dans le « matériel ». Cette neutralité s'expliquait déjà dans la matrice de corrélations dont la valeur moyenne était proche de 0,8, proche également de l'ensemble des valeurs des variables.

Cela peut s'expliquer par le fait qu'internet n'est ni un service ni un matériel, on ne peut pas le caractériser dans l'un ou dans l'autre.

► **Toutes les variables sont corrélées avec l'axe F1, ce qui montre le développement des médias dans le monde.**

8. Analyse des Individus

Calcul

Cette matrice correspond à la qualité de représentation pour chaque pays pour chaque axe sélectionné. Il s'agit en fait de des coordonnées du pays par rapport à chaque axe.

$$\cos^2(m; axe) = \frac{coord^2(m, axe)}{\sum coord^2(m)}$$

La contribution d'un individu par rapport à son axe est donnée par la formule

$$F = A^s \times U$$

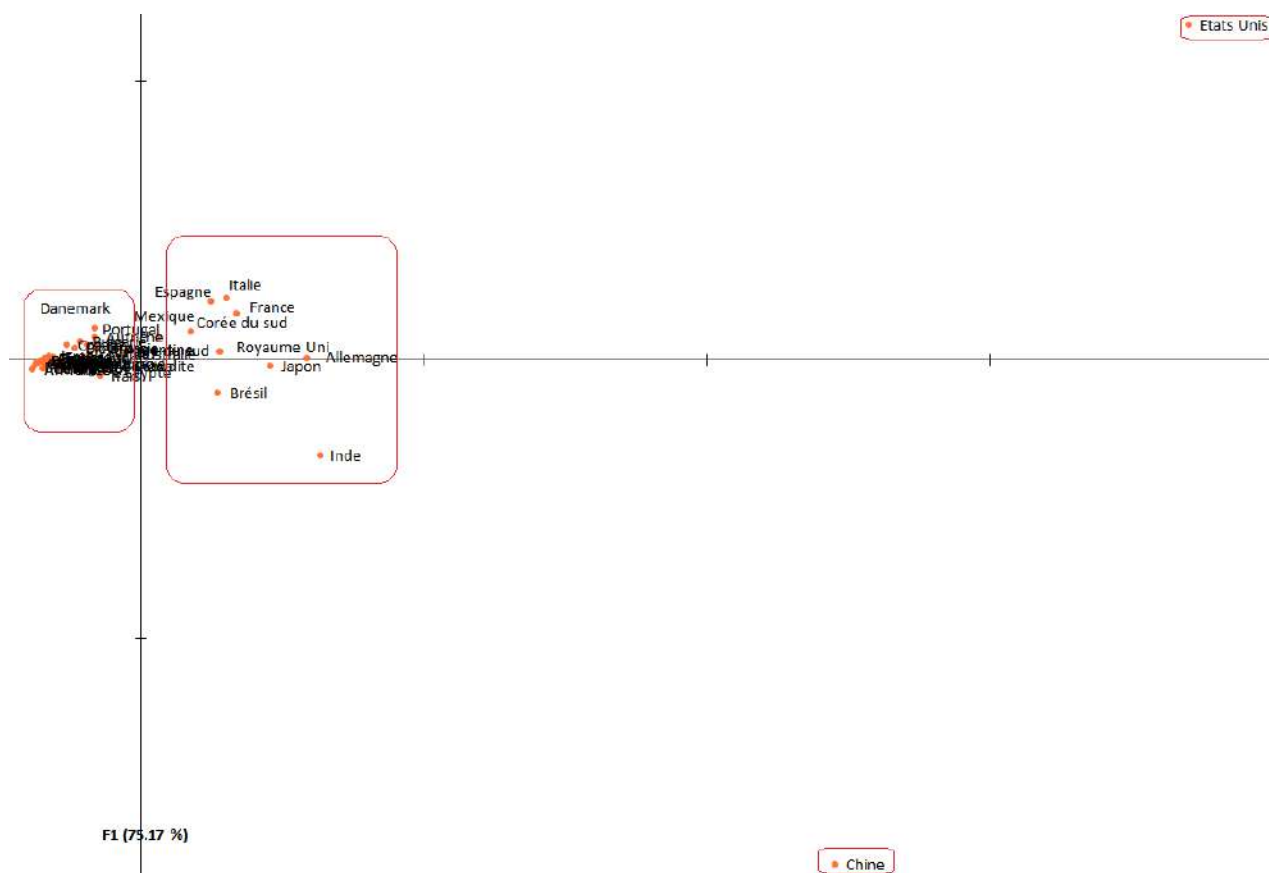
A^s la matrice centrée réduite

U la matrice de changement de base

Coordonnées des observations et qualité par rapport à l'axe

Observation	F1	Qualité	F2	Qualité
Afghanistan	-1,373	0,482	0,025	0,001
Afrique du sud	-0,636	0,103	0,151	0,028
Algérie	-1,106	0,313	-0,100	0,012
Allemagne	2,339	1,399	0,027	0,001
Angola	-1,377	0,485	-0,043	0,002
Arabie Saoudite	-0,946	0,229	-0,080	0,008
Argentine	-0,376	0,036	0,152	0,029
Arménie	-1,538	0,605	-0,133	0,022
Australie	-0,247	0,016	0,092	0,010
Autriche	-0,659	0,111	0,339	0,143
Bénin	-1,389	0,493	-0,117	0,017
Biélorussie	-0,949	0,230	0,170	0,036
Birmanie	-1,535	0,602	-0,104	0,013
Bolivie	-1,209	0,374	-0,009	0,000
Brésil	1,079	0,298	-0,459	0,261
Brunei	-1,436	0,527	-0,003	0,000
Bulgarie	-0,870	0,194	0,269	0,090
Burkina Faso	-1,220	0,381	-0,053	0,004
Cambodge	-1,504	0,579	-0,045	0,003
Cameroun	-1,314	0,442	-0,012	0,000

Observation	F1	Qualité	F2	Qualité
Canada	-0,434	0,048	-0,073	0,007
Chili	-0,742	0,141	0,071	0,006
Chine	9,801	24,574	-7,227	64,796
Colombie	-0,736	0,139	-0,044	0,002
Congo	-1,339	0,459	0,032	0,001
Corée du sud	0,695	0,124	0,413	0,212
Cote d'Ivoire	-1,407	0,507	-0,113	0,016
Croatie	-1,054	0,284	0,222	0,061
Cuba	-1,350	0,466	0,016	0,000
Danemark	-0,669	0,115	0,457	0,259
L'Egypte	-0,557	0,079	-0,174	0,038
Equateur	-1,243	0,396	0,043	0,002
Espagne	0,985	0,248	0,843	0,882
Etats Unis	14,801	56,041	4,806	28,660
Ethiopie	-1,466	0,549	-0,041	0,002
France	1,347	0,464	0,671	0,559
Grèce	-0,640	0,105	0,183	0,041
Inde	2,526	1,632	-1,368	2,321
Irak	-1,141	0,333	-0,006	0,000
Iran	-0,579	0,086	-0,232	0,067
Irlande	-1,244	0,396	0,017	0,000
Israël	-1,306	0,436	0,064	0,005
Italie	1,200	0,368	0,896	0,997
Japon	1,823	0,851	-0,079	0,008
Koweït	-1,437	0,528	-0,020	0,000
Liban	-1,189	0,362	-0,024	0,001
Maroc	-1,120	0,321	-0,133	0,022
Mexique	0,240	0,015	0,339	0,142
Portugal	-0,770	0,152	0,223	0,061
Royaume Uni	1,113	0,317	0,125	0,019
Russie	2,457	1,544	0,250	0,078



Interprétation axe F1

Le premier constat visible après lecture de ce graphique est l'opposition majeure entre chine-état unis et le reste du monde sur l'axe F1.

Cette forte opposition s'explique par la sur représentation des données globales d'information des deux pays. En effet la qualité de chaque pays contribue largement à la définition de l'axe. **La Chine et les États-Unis ont un taux de contribution dans les observations sont presque 55 fois supérieures à la majorité des pays étudiés !** La très grande majorité des pays étudiés sont ici en légère sous représentation, cela s'explique par la forte sur représentation du duo chine/état unis qui faut hausser les statistiques mondiales.

Cette sur représentation peut à la fois s'expliquer par la population très importante de ces deux pays pouvant être proportionnel au media, mais pas seulement (l'Inde, second foyer mondial n'est pas autant représenté sur l'axe par exemple). La seconde raison est probablement liée à leurs influences respectives dans le monde (premier et troisième puissance mondiale). De plus on peut noter que la majorité des pays étant représenté sur l'axe F1 (abscisse > 0) sont à la tête des puissances mondiales (les 15 premières) et font partie des pays développés (Europe) ou en voie de développement (Mexique, Brésil, Inde). **Il apparaît donc que les pays important économiquement ont une proportion à être sur représenté mondialement au niveau de l'accès à l'information (axe F1) pouvant nous rapproché dans une étude économique.**

Le point d'axe (0,0) est la représentation moyenne au sein du monde. **L'Australie est donc le pays symbolisant l'équilibre vis à vis de l'accès à l'information dans le monde.**

Interprétation axe F2

On remarque que la Chine et les États-Unis sont fortement opposés, il y a une réelle discordance entre ces deux pays, et plus généralement entre les pays développés et ceux en développement.

Rappel : L'axe F2 symbolise le rapport entre l'offre disponible et la quantité de matériels pour cette utilisation.

La Chine apparaît donc énormément équipée en matériel et les États-Unis disposant d'une offre très abondante et diversifiée.

Cette répartition semble s'expliquer par la politique mise en vigueur dans chaque pays.

La Chine est un pays d'idéologie dictatoriale communiste où le partage matériel est une figure cruciale et où le contrôle de l'information est constante. **Une proportion de matériel important face à la quantité d'offre proposée semble donc logique.**

A l'inverse, les États-Unis est un pays capitaliste où l'offre prime et où la liberté d'entreprendre et de diffuser l'information est essentielle. **Elle est donc surexposée face à la quantité de matériel disponible.**

D'une manière plus générale, la plupart des pays situés en dessous de l'axe F1 sont des pays en développement favorisant le matériel plus que l'offre (bien souvent pour des raisons politiques) et les pays situés au dessus sont des pays développés où l'offre est primordiale.

L'axe F2 semble représenté des régimes politiques différents :

- ordonnée > 0 tend vers un régime capitaliste, propriété individuelle
- ordonnée < 0 tend vers un régime communiste, nationalisation

On pourrait penser que la répartition de la population de chaque pays influe sur la position de l'axe F2. Néanmoins une simple logique peut nous indiquer le manque d'implication de ce critère. La Chine avec ses 1,3 milliards de personnes ont un fort taux d'équipement, on aurait pu croire que ce vaste marché de consommateur aurait influencé le développement des offres cependant l'inverse se produit. **On peut donc considérer que seul la politique du pays peut limiter ou accroître la profusion d'offres.**

Interprétation générale

L'analyse mondiale de l'information nous permet donc de penser que la diffusion de l'information est étroitement liée économiquement et politiquement et ce dans chaque pays. Ainsi, les pays situés à droite des axes sont les principaux acteurs de l'ONU possédant une importance économique au sein du monde.

Conclusion

Problématique

Les ressources mis en oeuvres dans les moyens de communications d'un pays reflète sa croissance ?

Cette problématique définie d'après un raisonnement purement instinctif semble corroborer à la vérité. En effet nous avons pu voir que les pays possédant un fort taux de développement des ressources communicatives sont les pays influençant et dirigeant les grandes instances mondiales. Néanmoins cette analyse a permis également de faire valoir un concept plus camouflé : **le développement des médias selon le régime politique**. On peut ainsi ressentir la gouvernance du pays selon le rapport offre de services et utilisations de services.

Les limites

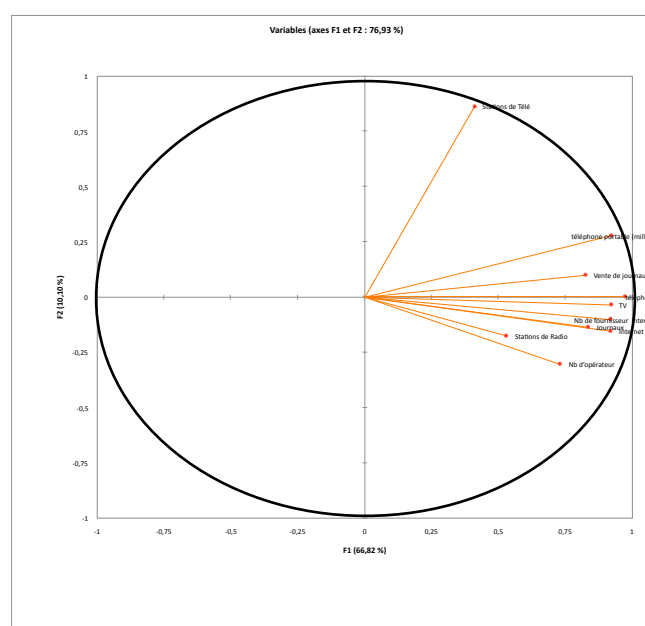
- Il est évident que la sur représentation de la Chine et des États-Unis sur chacun des axes a pu dénaturé l'information.
- La corrélations et l'absence d'anti corrélation entre toutes les variables peut induire de graves erreurs sur l'interprétation. Ainsi la population de chaque pays aurait pu être prise en compte pour réellement vérifié les dires.

Ouverture

Il nous semblerait intéressant de perfectionné l'ACP sans les individus sur représenté qu'est la Chine et les États-Unis. Une ACP pour chaque continent serait également intéressant permettant ainsi d'accentuer ou de réfuter les hypothèses émisent tout au long de l'analyse.

Un élément de Réponse ?

Une ACP sans les individus sur représenté

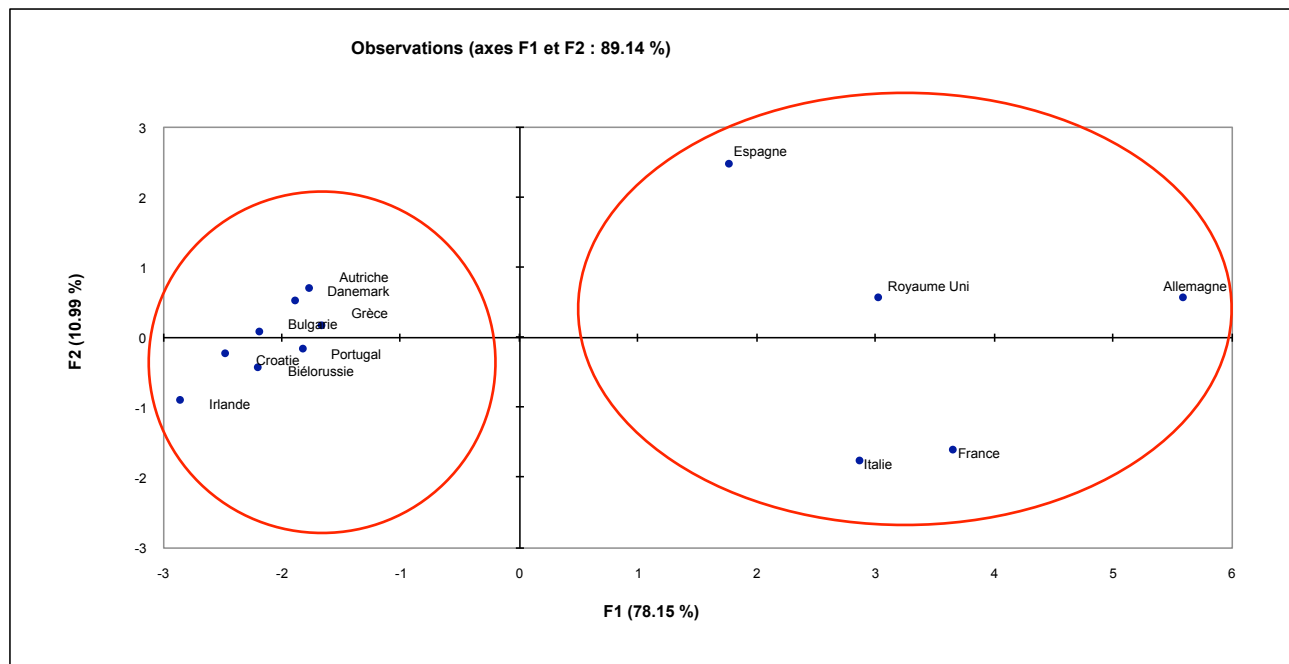


Le cercle des corrélations de cette ACP montrerait une interprétation différente par rapport à l'ACP avec les États-Unis et la Chine. Ainsi l'opposition matériel et offre de service ne serait pas aussi prononcé.

Cette ACP tendrait à réfuter la corrélation démontré précédemment.

Une ACP par continent

Pour cette exemple nous avons choisis l'Europe, pour nos acquis sur son développement, son histoire, ses pays... Le cercle des corrélations montre une interprétation quasi identique pour le développement des médias dans un pays (à savoir corrélation de toutes les variables sur le même axe).



On constate qu'il y a deux groupes qui se distinguent. Ces deux groupes sont identifiables sur le graphique mais également dans l'Union Européenne et l'Europe en générale. En effet le groupe composé de la France, Allemagne représente les moteurs, les puissances de cette Europe. On peut également signalé les positions mondiales respectifs de ces pays (Allemagne 4ème puissance mondiale, France 5ème, Royaume-Unis 6ème).

Ainsi cette ACP tendrait à vérifier la corrélation démontré précédemment.

Synthèse Personnel

Jérôme

Réussir à trouver un nombre d'informations importants sur 52 pays ne fut pas une chose aisée, pour certain pays, les informations ont réclamé un travail de recherche approfondit.

Lionel

Le contexte du sujet fut très intéressant, seulement l'interprétation des résultats nous poussait parfois à une interprétation sortie du contexte analyse de donnée vers celle d'une analyse géopolitique.

Loan

Le grand nombre de variables fortement corrélées entre elles n'a pas simplifié les déductions. D'ailleurs le fait de ne trouver aucune anti-corrélation et des résultats aussi proches les uns des autres le prouve et le choix de ce qui déterminait les axes le prouve.

Solène

Le sujet choisi donnait des résultats durs à analyser, les données étaient importantes (52 individus) et les résultats parfois peu évidents à interpréter.

Yann

Les difficultés que j'ai rencontré sont à l'interprétation des données statistiques au niveau de l'espace géographique : les données étant majoritairement toutes corrélées à un seul axe, il fut peu évident de dégager d'autres informations avec un second axe.

Équipe

Le sujet de par son étude vaste nous a aussi dispersé vers une étude par groupe de pays, difficile à gérer en termes d'horaires.