

ANALYSE GÉOMÉTRIQUE DE DONNÉES BAROMÉTRIQUES

Frédéric Cassor⁽¹⁾ & Brigitte Le Roux^(1,2)

¹ CEVIPOF Sciences-Po, 98 Rue de l'Université 75007 Paris
frederik.cassor@sciencespo.fr

² MAP5, Université Paris Descartes 45 rue des Saints Pères 75270 Paris Cedex 06
Brigitte.LeRoux@mi.parisdescartes.fr

Résumé. Les données d'une *enquête barométrique* comportent plusieurs groupes d'individus représentatifs d'une même population. Chaque groupe est interrogé à des périodes différentes. Chaque groupe répond à la même batterie de questions. Nous comparons, en utilisant les méthodes d'analyse géométrique des données, deux vagues d'enquête du « Baromètre de la confiance politique » initié par le CEVIPOF, qui ont eu lieu en décembre 2017 et en décembre 2018.

L'étude est menée à partir de 29 questions bipolaires qui relèvent de 4 composantes de confiance (politique, institutionnelle, économique et interpersonnelle). Nous avons construit un *espace de la confiance* en effectuant une analyse des correspondances après avoir dédoublé les questions et les avoir pondérées selon leur thème. Les individus interrogés en 2017 (vague 9) ont été pris comme ensemble de référence, ceux interrogés en 2018 (vague 10) ont été mis en éléments supplémentaires. Pour étudier l'*évolution* des réponses aux questions, nous avons appliqué les formules de transition aux individus supplémentaires, ce qui permet de visualiser les écarts entre les modalités des questions des deux vagues. Par ailleurs, nous avons effectué une classification des individus interrogés en 2017 (méthode de Ward). Nous avons identifié 5 classes d'individus au regard de la confiance ("hyperconfiants", "confiants modérés", "défiant altruistes", "défiant autoritaires", "hyperdéfiant"). Nous avons ensuite affecté les individus interrogés en 2018 aux classes en prenant un critère d'affectation basé sur la distance de Mahalanobis associée à chaque classe afin de tenir compte de la forme des classes ce qui nous a permis d'étudier les *transferts entre classes*.

Mots-clés. Étude de cas, , enquête barométrique, analyse géométrique des données, codage par dédoublement.

Abstract. Barometer data include several groups of individuals that are representative of the same population. Those groups are interviewed at different periods. They answer the same set of questions. We study changes by using Geometric Data Analysis. Data stem from "Le baromètre de la confiance politique" ("Barometer of political trust"), a survey carried out by CEVIPOF (Center for political studies in Sciences Po, Paris) that studies trust of the French citizens registered to vote. More precisely, we compare survey data that are collected in December 2017 and in December 2018, by using methods of Geometric Data Analysis.

Twenty-nine bipolar questions, which are relevant to four themes (political, institutional, economic, interpersonal trust), are selected for this statistical analysis. For constructing the “trust space”, we perform correspondence analysis (CA) after doubling and weighting questions. Individuals of Wave 9 (2017) are put as active individuals (reference wave), the ones of Wave 10 (2018) are put as supplementary elements.

Then, for studying how the *responses to questions change over time*, we apply transition formulas to supplementary individuals. Moreover, a Euclidean hierarchical clustering (Ward’s method) is applied to active individuals. Five clusters of individuals as regards trust are distinguished ("hyper-trusters", "moderate trusters", "moderate distrusters", "authoritarian distrusters", "hyper-distrusters"). Then, using the Mahalanobis distance related to each cluster, we assign individuals of Wave 10 (2018) to the five clusters, thus we study the *transfers between clusters over time*.

Keywords. Case study, Barometer Survey Data, Geometric Data Analysis, Doubling coding.

1 Les données

Les données proviennent d’une enquête barométrique conduite par le CEVIPOF¹ qui est réalisée chaque fin d’année depuis 2009. Les questions portent sur la confiance chez l’électeur français [4]. Les échantillons (d’environ 2000 personnes) sont représentatifs des électeurs français².

Les objectifs de cette étude sont d’apporter des réponses aux questions suivantes :

1. Comment s’organisent les différentes composantes de la confiance ? Comment les individus se distribuent-ils dans cet espace de la confiance ?
2. Comment a évolué la confiance entre 2017 et 2018 ?

2 Espace de la confiance

Pour construire l’espace de la confiance et établir une typologie des électeurs français au regard de la confiance, nous avons retenu quatre composantes de la confiance mesurées par 29 questions (27 questions sont codées selon une échelle de type Likert avec 4 ou 5 échelons, deux sont dichotomiques).

1. Pour les détails concernant les différentes vagues d’enquête, voir le site internet du CEVIPOF <http://www.sciencespo.fr/cevipof/fr/content/le-barometre-de-la-confiance-politique>

2. Ils ont été constitués par la méthode des quotas, au regard des critères de sexe, d’âge, de catégorie socio-professionnelle, après stratification par région de résidence et de taille de commune. La méthode de recueil des données est l’enquête en ligne en utilisant le système CAWI (Computer Assisted Web Interview) ; les enquêtes ont été réalisées par OpinionWay.

- *Confiance politique* : confiance dans toute une série de rôles et d'institutions politiques (président, premier ministre, député, maire, conseiller général, partis politiques, députés européens, Union européenne).
- *Confiance institutionnelle* : confiance dans diverses grandes institutions publiques et privées (les hôpitaux, la police, l'armée, la justice, les médias, les syndicats, la sécurité sociale, les grandes conférences internationales comme le G20).
- *Confiance économique* : confiance envers des organisations du monde économique (les banques, les grandes entreprises publiques, les grandes entreprises privées, les PME, confiance/contrôle des entreprises par l'État).
- *Confiance interindividuelle* : envers les voisins, les gens rencontrés pour la 1^{ère} fois, ceux ayant une opinion religieuse différente, ceux d'une autre nationalité, confiance/prudence vis-à-vis des gens; les gens peuvent changer la société par leurs choix et leurs actions, les gens de mon pays ont la possibilité de choisir leur propre vie; optimisme sur mon propre avenir.

Après avoir procédé à un codage par dédoublement ([2], [5]) et à la pondération des questions par l'inverse du nombre de questions de la composante de confiance, nous avons effectué une analyse des correspondances en prenant les 1820 individus interrogés en 2017 (vague 9) comme référence (individus actifs), et en mettant ceux interrogés en 2018 en éléments supplémentaires. Nous interpréterons les deux premiers axes.

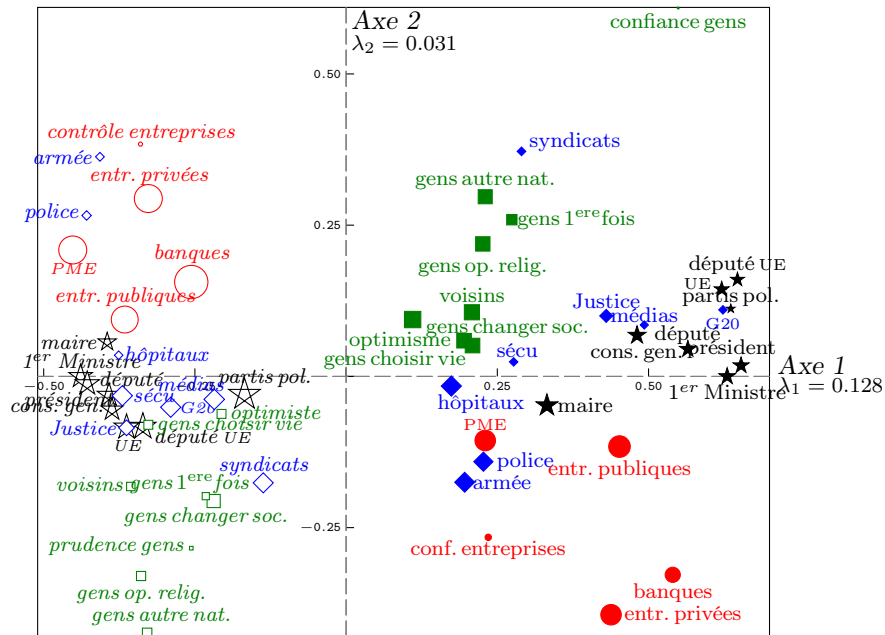


FIGURE 1 – Plan 1-2 : nuage des poles (+) et (- en italiques) des questions.

Le 1^{er} axe ($\lambda_1 = 0.128$, $\tau_1 = 32\%$) est un axe général qui mesure le *niveau global de confiance*. Le 2^{ème} axe ($\lambda_2 = 0.031$, $\tau_2 = 8\%$) montre une opposition entre une « confiance

d'en haut » — les grandes entreprises privées, les banques — et une « confiance d'en bas » à dimension plus « société civile » — les voisins, les gens d'une autre nationalité, les autres en général — ainsi que dans les syndicats.

3 Évolution de la confiance

L'évolution de la confiance est étudiée, d'une part en construisant les trajectoires des questions dans l'espace de la confiance, d'autre part en effectuant une classification des individus de référence (vague 9) et en affectant ceux de la vague 10 aux classes.

3.1 Trajectoires des questions

Tout d'abord, les électeurs interviewés en 2018 ont été mis en éléments supplémentaires dans l'AC. Puis, en utilisant les propriétés de l'AC après dédoublement des questions (formules de transition) nous représentons, dans l'espace des questions, les pôles des questions à partir des réponses des individus supplémentaires³. Pour chaque pôle des questions, les écarts entre les deux périodes sont représentés par des vecteurs (2017 → 2018). Le graphique 2 montre les évolutions pour 6 questions.

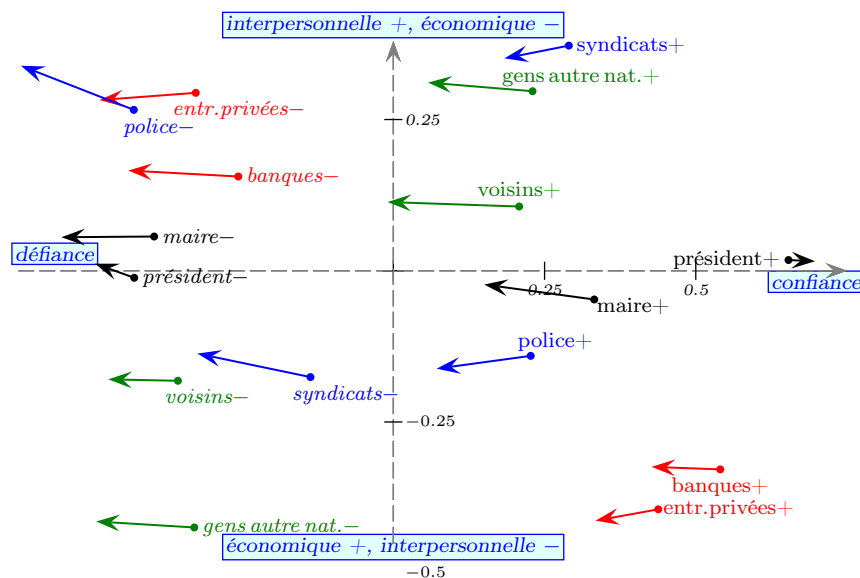


FIGURE 2 – Trajectoires de 8 questions dans le plan 1-2.

3. Pour chaque échelon et pour chaque vague, on calcule le point moyen des individus ayant choisi les modalités --, -, =, +, ++, puis les barycentres de ces points moyens pondérés par le produit de l'effectif de la modalité par le codage numérique de l'échelon de l'échelle (respectivement 4,3,2,1,0 pour le pôle q^- et 0,1,2,3,4 pour le pôle q^+). Enfin, on procède à un produit d'affinités orthogonales selon les axes principaux, afin de représenter le point représentant chaque pôle dans l'espace des questions.

On voit que les écarts sont faibles pour "Le président" ; pour les autres questions, les écarts vont tous dans le sens d'une plus grande défiance.

3.2 Classification ascendante hiérarchique

Nous avons effectué une classification euclidienne des individus interviewés en 2017 (méthode de Ward). Les indices de niveau décroissent lentement à partir du 5^e, on interprétera donc une partition en 5 classes.

1. Les « *hyperconfiants* » ($n = 184$, 10%).
Les modalités surreprésentées sont celles d'une grande (+), voire très grande (++) confiance politique, institutionnelle, économique et interindividuelle.
2. Les « *confiants modérés* » ($n = 518$, 28%).
Les modalités surreprésentées sont celles d'une confiance politique modérée (+) (sauf vis-à-vis des partis politiques), d'une confiance institutionnelle (sauf pour les médias et les syndicats) et sociétale et celles d'une confiance économique vis-à-vis des entreprises (+) mais avec une certaine méfiance vis-à-vis des banques (-).
3. Les « *défiants altruistes* » ($n = 163$, 9%)
Dans cette classe, on trouve une défiance dans les domaines du politique et de l'économie (sauf pour les PME), dans les institutions qui assurent un maintien de l'ordre et la sécurité (armée, police), mais aussi une bonne confiance (+ et ++) dans les syndicats, la sécurité sociale et une confiance interpersonnelle certaine (confiance d'en bas).
4. Les « *défiants autoritaires* » ($n = 563$, 31%).
La classe est faite de défiance politique avec une forte défiance vis-à-vis des acteurs politiques nationaux et une forte défiance interpersonnelle. Par ailleurs, il y a une certaine confiance dans l'économie, et s'il y a une forte défiance vis-à-vis des syndicats, des grandes conférences internationales comme le G20 et des médias, on trouve une confiance (+) dans l'armée et la police.
5. Les « *hyperdéfiants* » ($n = 392$, 22%)
Dans cette classe, on trouve une défiance profonde (--) tous azimuts.

Les ellipses de concentration de 5 classes sont représentées dans le plan principal 1-2 (voir figure 3).

Evolution des classes

Pour compléter cette étude, nous avons affecté les individus interrogés en 2018 aux classes de cette partition, selon la procédure de classement présentée dans Le Roux et Cassor (2015)⁴ et Benzécri (1977). Les pourcentages d'individus dans chaque classe sont donnés dans le tableau suivant.

4. Un script R est disponible auprès des auteurs.

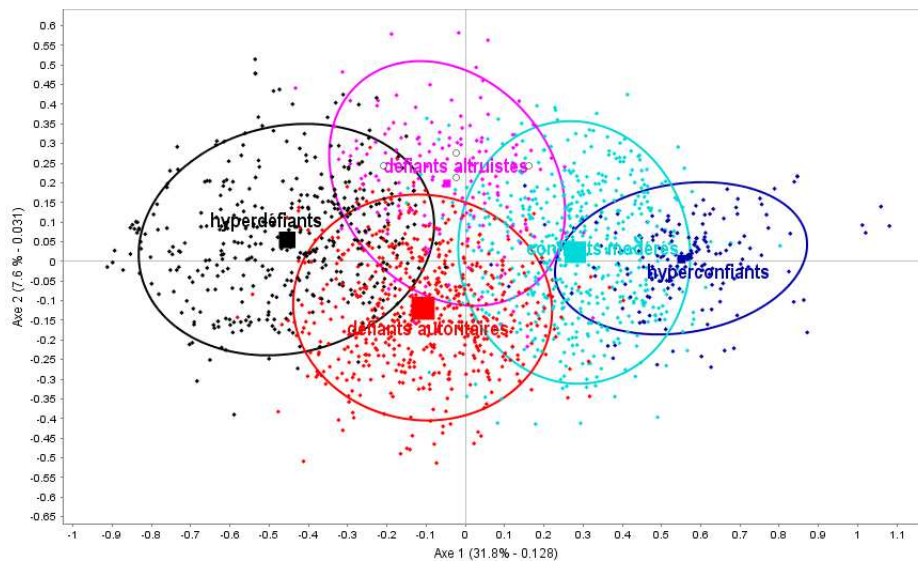


FIGURE 3 – Plan 1-2 : ellipses de concentration des 5 classes de la CAH.

classes	Déc 2017	Déc 2018	Évolution
hyperconfiants	10.1	6.2	-3.9
confiants modérés	28.5	21.3	-7.2
défiant altruistes	9.0	7.7	-1.3
défiant autoritaires	30.9	39.0	+8.1
hyperdéfiant	21.5	25.8	+4.3

Les classes de confiance ont diminué en proportion au profit des « défiant autoritaires » et aussi des « hyperdéfiant ».

Bibliographie

- [1] Benzécri, J-P. (1977), Analyse discriminante et analyse factorielle, *Les cahiers de l'analyse des données*, 2 (4), 369-406.
- [2] Le Roux, B. (2014), *Analyse géométrique des données multidimensionnelles*, Dunod, Paris.
- [3] Le Roux, B. et Cassor F. (2015), “Assigning Objects to Classes of a Euclidean Ascending Clustering”, dans *Statistical Learning and Data Sciences* (Third International Symposium, SLDS 2015, Egham, UK, April 20-23, 2015, Proceedings), Gammerman, A., Vovk, V., Papadopoulos H. (eds), pp 389-396, Springer.
- [4] Le Roux, B. et Perrineau, P. (2011), Les différents types d'électeurs au regard de différents types de confiance, *Les cahiers du CEVIPOF*, [http ://www.cevipof.com/fr/les-publications/les-cahiers-du-cevipof/](http://www.cevipof.com/fr/les-publications/les-cahiers-du-cevipof/), 54, 5-35.
- [5] Murtagh, F. (2005). *Correspondence Analysis and Data coding with Java and R*, Chapman and Hall, London.