

ANALYSE STATISTIQUE DE DONNÉES BAROMÉTRIQUES: APPLICATION AU "BAROMÈTRE DE LA CONFIANCE POLITIQUE"

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. Le principal objectif des données barométriques est d'étudier les changements qui se produisent au cours du temps. Nous étudierons ce problème à l'aide des méthodes d'analyse géométrique des données.

Les données sont issues de l'enquête "Baromètre de la confiance politique" initiée par le CEVIPOF portant sur la confiance des électeurs français. Pour l'analyse statistique, on a retenu 19 questions qui relèvent de 4 thèmes (politique, institutionnel, économique et interpersonnel). Nous avons construit un espace de la confiance en effectuant une analyse des correspondances après dédoublement des questions et pondération par thème. Les individus interrogés en 2016 (vague 8) ont été pris comme ensemble de référence, ceux des années 2009 puis 2012 à 2015 ont été mis en éléments supplémentaires.

Pour étudier l'évolution des réponses aux questions, nous avons appliqué les formules de transition aux individus supplémentaires. Par ailleurs, nous avons effectué une classification des individus de l'année 2016 (méthode de Ward). Nous avons identifié 4 classes d'individus au regard de la confiance ("hyperconfiants", "confiants", "défiant", "hyperdéfiant"). Puis, nous avons affecté les individus des autres années aux classes en prenant un critère d'affectation basé sur la distance de Mahalanobis associée à chaque classe afin d'étudier les *transferts entre classes*.

Mots-clés. Analyse géométrique des données, codage par dédoublement, étude de cas, données d'enquête barométrique, données temporelles.

Abstract. The main objective of barometer data is to analyze the change of questions over time. We study the change by using Geometric Data Analysis. The data source stems from the survey entitled "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. The data have been collected since 2009 on a yearly basis.

Nineteen questions that are relevant to four themes (political, non-political and economic institutions, interpersonal trust) are selected for this statistical analysis. For constructing the "trust space", we perform correspondence analysis (CA) after doubling the

data and we weight the questions. Individuals of Wave 8 (2016) are put as active individuals (reference wave), the ones from the other waves (2009 then 2012 to 2015) 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. Four classes of individuals as regards trust are distinguished ("hyper-trusters", "moderate trusters", "moderate distrusters", "hyper-distrusters"). Then we assign individuals from the other waves to the four classes in order to observe some *transfers between classes over time*.

Keywords. Geometric Data Analysis, Doubling coding, Case Study, Barometer Survey Data, Temporal Data.

1 Les données

Cette étude est consacrée à l'analyse de la confiance chez les électeurs français. La confiance est un phénomène complexe comprenant plusieurs composantes : on peut distinguer habituellement la confiance que nous accordons à des personnes de celle que nous prêtons à des institutions ([4]).

Les données analysées proviennent d'enquêtes initiées par le CEVIPOF qui prennent en compte les composantes diverses et parfois hétérogènes de la confiance. Les échantillons (d'environ 2000 personnes) sont représentatifs des électeurs français; 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.

Les objectifs de cette étude sont les suivants :

1. Comment s'organise les différentes composantes de la confiance ?
2. Comment les individus se distribuent-ils dans cet espace de la confiance ?
3. Comment évolue la confiance depuis 2009 ?

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 19 questions (18 sont codées selon une échelle de Likert à quatre échelons, et 1 question est dichotomique).

- *Confiance politique*: confiance dans toute une série de rôles et d'institutions politiques (l'institution présidentielle; le premier ministre; votre député; le maire de votre commune; votre conseiller général; les partis politiques; l'Union européenne).

- *Confiance institutionnelle*: confiance dans diverses grandes institutions publiques et privées (les hôpitaux; la police; les médias; les syndicats; l'OMC)
- *Confiance économique*: confiance envers des organisations du monde économique (les banques; les grandes entreprises publiques; les grandes entreprises privées; confiance/contrôle des entreprises par l'État).
- *Confiance interindividuelle*: confiance avec les voisins, les gens rencontrés pour la 1^{ère} fois, les gens d'une autre nationalité.

Après un codage par dédoublement ([2], [5]) et 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 1963 individus de la vague 8 comme référence (individus actifs), et en mettant ceux des autres vagues en éléments supplémentaires. Nous interpréterons les deux premiers axes.

Le 1^{er} axe traduit un *niveau global de confiance*. On voit que les questions politiques sont les plus clivantes en ce sens qu'elles réunissent un petit nombre d'individus aux opinions très tranchées du point de vue de la confiance, d'où leur position à l'extrême droite du graphique.

Le 2^{ème} axe est un axe qui oppose la confiance interpersonnelle et la défiance économique à la défiance interpersonnelle et la confiance économique.

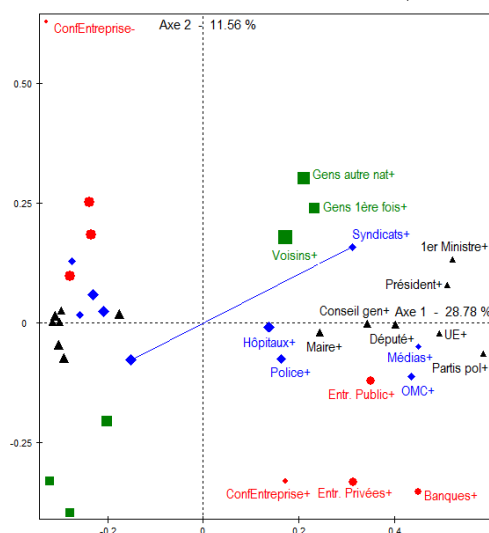


Figure 1. Nuage des questions avec les pôles (+ et -) dans le 1^{er} plan principal: les symboles sont proportionnels au poids des questions.

3 Confiance politique: évolution

Pour montrer cette évolution, nous procédons en deux temps. Tout d'abord, les individus des autres vagues sont mis en éléments supplémentaires dans l'AC. Puis, utilisant les propriétés de l'AC dédoublee, on construit les points associés aux deux pôles des questions en utilisant les formules de transition, ce qui revient, pour chaque question et pour chaque vague, à calculer la moyenne pondérée des coordonnées des individus, puis à opérer un produit d'affinités orthogonales selon les axes principaux. Ensuite, on compare leurs positions par rapport à celles de 2016 (figure suivante).

La confiance dans les différentes institutions politiques (sauf pour l'institution présidentielle et le premier ministre) évolue selon une trajectoire similaire. De 2012 à 2013, la confiance dans ces institutions est plus consensuelle et glisse vers la dimension économique. En

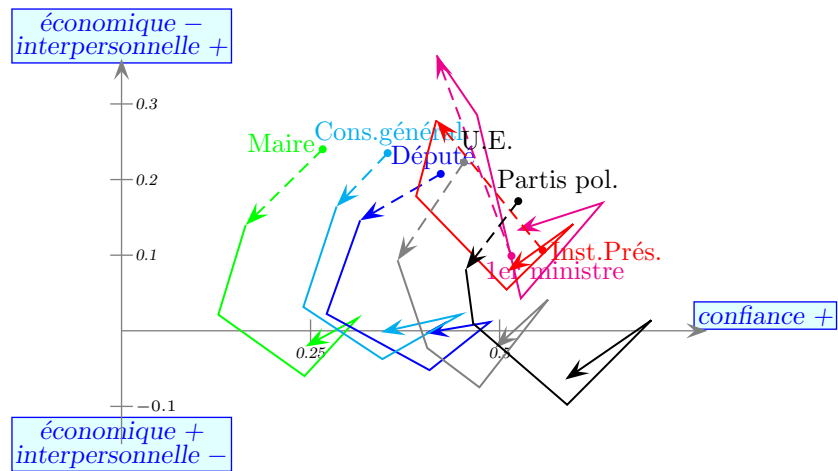
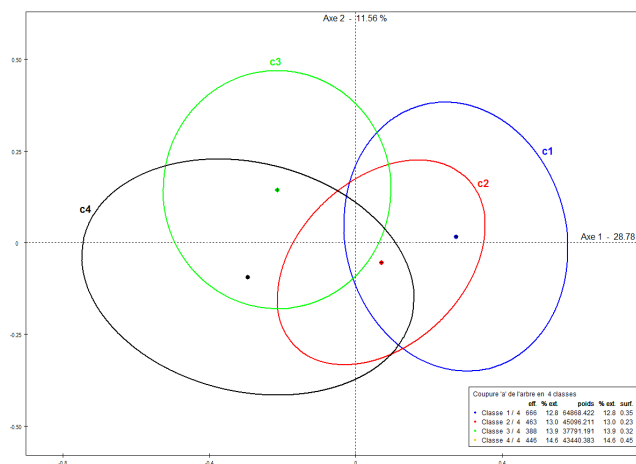


Figure 1: Trajectoires des questions du thème politique dans le plan principal 1-2 (En pointillé, évolution entre 2009 et 2012).

2015, on note un regain du niveau de confiance pour toutes les institutions, marqué par leur position à l'extrême droite de l'axe 1, avant de revenir à la situation de 2014.

4 Typologie des individus

Nous avons effectué une classification euclidienne des individus de la vague 8 (méthode de Ward). Les indices de niveau décroissent lentement à partir du 4^e: on interprétera une partition en 4 classes. Les ellipses de concentration de ces 4 classes dans le plan principal 1-2 sont représentés sur la figure suivante.



1) Les "hyperconfiants" (classe *c1*, $n_{c1} = 666$, 34%).

Dans cette classe, les modalités surreprésentées sont celles d'une grande confiance pour les 4 composantes: politique, institutionnelle, économique, interindividuelle.

2) Les "confiants modérés" (classe *c2*, $n_{c2} = 463$, 24%).

Les modalités surreprésentées sont celles d'une bonne confiance dans les institutions économiques, plus contrastée dans les institutions publiques et privées, modérée dans les rapports interindividuels et d'une défiance pour le politique.

3) Les "défiants" (classe *c3*, $n_{c3} = 388$, 20%).

Les modalités surreprésentées correspondent à de la défiance forte à modérée en ce qui concerne les quatre composantes de la confiance.

4) Les "hyperdéfiants" (classe *c4*, $n_{c4} = 446$, 23%).

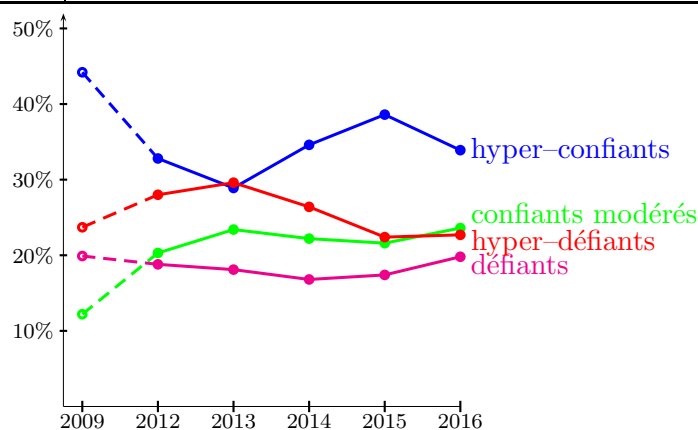
Dans cette classe, les modalités surreprésentées sont celles d'une très forte défiance dans tous les domaines.

Evolution des classes

Pour compléter cette étude, nous avons affecté les individus des vagues 2009 puis 2012 à 2015 aux classes, selon la procédure de classement présentée en [1, 3].

Les pourcentages d'individus dans chaque classe sont donnés dans le tableau suivant. La figure ci-après montre les évolutions.

classes	Déc. 2009	Déc 2012	Déc 2013	Déc 2014	Déc 2015	Déc 2016
<i>c1</i> hyperconfiants	44	33	29	35	39	34
<i>c2</i> confiants modérés	12	20	23	22	22	24
<i>c3</i> défiants	20	19	18	17	17	20
<i>c4</i> hyperdéfiants	24	28	30	26	22	23
<i>Ensemble</i>	100	100	100	100	100	100



Les classes intermédiaires sont à peu près stables, alors que les classes extrêmes évoluent dans le sens d'une plus grande confiance de 2013 à 2015.

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/>, 54, 5-35.
- [5] Murtagh, F. (2005). *Correspondence Analysis and Data coding with Java and R*, Chapman and Hall, London