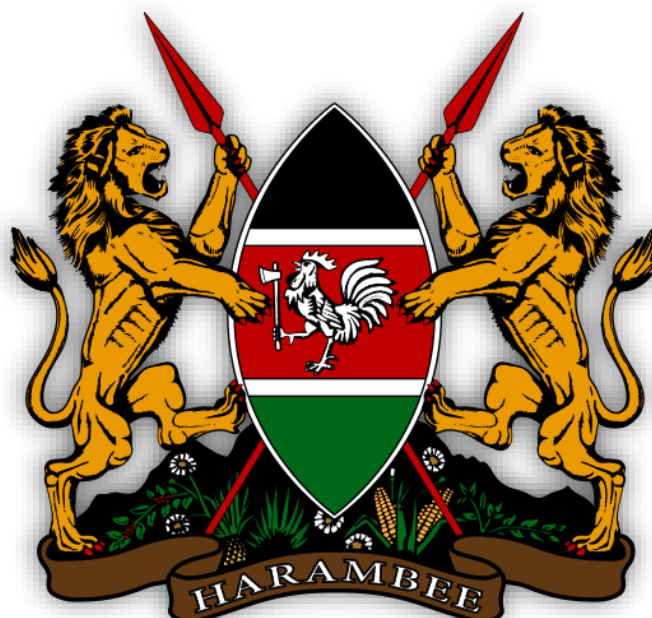


Enquête KDHS 2008-2009



Crédit : Ashiriz – Creative Commons Licence

Étude du lien entre l'éducation et le niveau de vie des ménages au Kenya en 2009

MECHOUK Lokman - ADJIMAN Clément – EDOHR Marvin
Projet encadré par : Mme M-C. Roubaud et Mme P. Rouaud
M1 MASS – Aix-Marseille Université – 2017-2018

Remerciements

Avant tout nous remercions nos professeurs Madame Roubaud et Madame Rouaud qui, tout au long du semestre, nous ont fourni les connaissances nécessaires à l'élaboration de ce rapport.

Nous souhaitons aussi remercier notre professeur Madame Gastineau pour l'aide qu'elle nous a apporté durant l'étape de réflexion en sa qualité de Démographe.

Sommaire

I. Introduction

- A. Présentation du Kenya
- B. Présentation de l'enquête KDHS 2008-09
- C. Présentation du projet

II. Analyse préliminaire

- A. Réflexion préalable à l'étude
 - B. Étude univariée
 - C. Étude bivariée

III. Analyse des Correspondances Multiple

- A. Méthodologie
- B. Résultats de l'ACM

IV. Conclusion

V. Bibliographie

VI. Table des matières

VII. Annexes

I. Introduction

A. Présentation du Kenya

La République du Kenya est située en Afrique de l'est. C'est un pays de 580 000 km² avec en son sein actuellement 48 millions d'habitants. Elle est bordée à l'est par l'océan Indien dont la côte de 500 km contient des marécages de mangroves ainsi que le port de Mombasa (qui permet au pays une ouverture économique et commerciale importante).

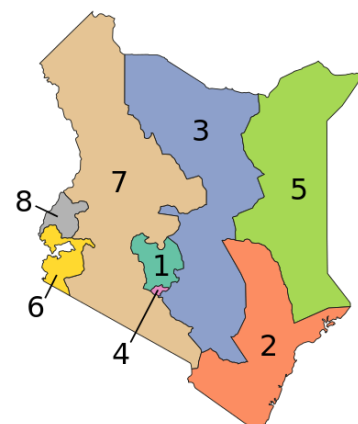


Figure 1 : Les huit régions du Kenya
Crédits : Public domain natural earth data

La majeure partie de la population est concentrée soit à l'ouest le long de la rive du lac Victoria, soit dans la région de la capitale Nairobi ou encore au sud-est le long de la côte de l'océan Indien. Le pays étant membre du Commonwealth, les langues officielles sont l'anglais et le swahili. De ce fait, le système légal en place est un mix entre la *common law* anglaise, la loi islamique et le droit coutumier.

En 2009, l'économie Kenyane est prédominée par l'agriculture avec une forte base industrielle. Le café, le thé et l'horticulture (fleurs, fruits et légumes) sont les principaux produits d'exportation agricoles. En 2008, ils ont contribué à 45% de la recette d'exportation totale d'après l'agence nationale des statistiques du Kenya (KNBS).

D'après la même source, depuis les années 2000 les performances de croissance sont médiocres et n'ont pas suivi le rythme de la croissance démographique. Ceci a été causée par des chocs externes et des problèmes structurels internes : sécheresse, faiblesse de prix des produits de base, récession mondiale, intempéries ainsi que de mauvaises infrastructures. Cette faible croissance a contribué à la détérioration du bien-être général de la population kenyane. De plus, le pays a été incapable de créer des emplois à un rythme correspondant à l'augmentation de la population active.

Ces soucis s'ajoutent au fait que le Kenya connaît en cette période une épidémie de VIH importante et géographiquement hétérogène qui affecte tous les secteurs économiques du pays.

L'enquête KDHS 2008-09 arrive donc à un moment crucial pour fournir des données destinées à implémenter de nouveaux programmes politiques, à suivre et évaluer les programmes existants et à concevoir de nouvelles stratégies démographiques, sociales et de santé au Kenya.

B. Présentation de l'enquête KDHS 2008-09

Les enquêtes KDHS (*Kenya Demographic and Health Survey*) sont des études longitudinales menées tous les cinq ans depuis 1989 afin d'évaluer l'état démographique et sanitaire du Kenya et ainsi mettre en place des programmes politiques visant l'amélioration du niveau de vie de la population.

Dans celle de 2008-2009, contrairement aux enquêtes précédentes, des informations ont été collectées sur : les niveaux de fécondité, les mariages, les activités sexuelles, les préférences de fertilité, la sensibilisation et l'utilisation des méthodes de planification familiale, les pratiques d'allaitement maternel, l'état nutritionnel des femmes et des jeunes enfants, la mortalité infantile et maternelle, la santé maternelle et infantile ainsi que la sensibilisation et le comportement adopté vis-

à-vis du VIH et d'autres infections sexuellement transmissibles. Cette enquête a aussi servi à recueillir des données sur le paludisme, l'utilisation des moustiquaires, la violence domestique et le dépistage du VIH chez les adultes.

L'enquête KDHS de 2008-09 a été conçue pour couvrir l'ensemble du pays et en particulier les zones du nord du pays qui n'étaient pas couvertes dans les enquêtes KDHS antérieures. Pour l'élaboration de cette étude, trois questionnaires ont été administrés à : un échantillon de femmes en âge de procréer de 15 à 49 ans et un échantillon d'hommes âgés de 15 à 54 ans dans un sous-échantillon de ménages sur deux. Les 9050 ménages ont été sélectionnés selon un plan de sondage stratifié sur la variable région. Cependant on ne prendra pas en compte les poids des individus dans notre analyse de données, cela signifie que les populations du nord-est (675 ménages) seront sous-représentées par rapport aux autres régions (environ 1300 ménages pour chaque). L'échantillon ne sera donc pas représentatif de la population kenyane lors de notre étude.

Dans le cadre de notre projet, nous nous baserons exclusivement sur la base de données tirée du questionnaire ménage.

C. Présentation du projet

L'éducation est un de nos soucis les plus importants. Selon nous, aucune variable ne devrait influencer sur la possibilité qu'un individu ait accès aux études supérieures et surement pas le niveau de richesse. Nous voulions donc au début étudier le lien entre l'éducation des membres du ménage et le niveau de richesse dudit ménage. Cependant, suite à une discussion avec Mme Gastineau, démographe au LPED et également notre professeur d'analyse démographique, nous a appris qu'il n'était pas possible ni intéressant d'agréger simplement les variables éducation de chacun des membres du ménage. Elle nous a ainsi conseillé de ne prendre en compte que l'éducation du chef de ménage qui est en outre moins biaisé par rapport à l'âge. En effet, si tous les membres d'un ménage ont moins de 16 ans, leur niveau d'éducation ne sera jamais plus élevé que le secondaire. Cela ne reflèterait pas le niveau d'éducation qu'ils peuvent atteindre et pourrait de ce fait fausser notre conclusion, ce qui n'est pas le cas si on prend seulement le niveau d'éducation du chef de ménage.

La problématique à laquelle nous répondrons donc à l'issue de notre étude est donc : **existe-t-il un lien entre la richesse d'un ménage et le niveau d'étude du chef de ménage ?**

Pour cela nous mettrons en relation la variable « niveau d'éducation du chef de ménage » avec les variables « score de richesse » et « facteur de richesse ». Nous mettrons aussi la variable d'éducation en relation avec des variables d'équipements des ménages.

II. Analyse préliminaire

A. Réflexion préalable à l'étude

Au début nous avons une table complète contenant les 9057 individus enquêtés et les 2191 variables correspondant à leurs réponses aux questions.

Après traitement, on passe à une table contenant toujours 9057 individus mais nous avons gardé seulement 31 variables : *identification, taille du ménage, lieu de vie (région, type, place), confort du lieu de vie (24 variables), chef de ménage (Sexe et âge)*

On y a ajouté ensuite les variables *indice de richesse, score de richesse et niveau d'éducation du chef de ménage*. Elles permettront de répondre à notre problématique.

Nous avons effectué des tris à plat (Annexe 1) qui ont permis de mettre en évidence plusieurs soucis :

- 1- Certaines variables contiennent un nombre beaucoup trop important de modalités. Par exemple « Matériau principal des murs » en contient 18. Il est difficile d'en analyser autant. Pour pallier à ce souci, nous les avons regroupés en modalités de qualité. Par exemple : sans mur, boue, fumier et toutes les constructions à base de boue seront regroupés dans la modalité « mauvaise qualité » tandis que les constructions de briques, argiles et ciments seront regroupés dans la modalité « Bonne qualité ». Pour les sources d'eau, nous prendrons en compte la qualité de l'eau ainsi que la sécurité relative à l'hygiène. Une eau provenant d'un lac sera donc comprise dans la modalité « eau risquée et source insalubre » tandis qu'une eau provenant d'un puit protégé sera comprise dans la modalité « eau de bonne qualité et source sûre ». Nous avons fait cela pour toutes les variables qui le requièrent (voir Annexe 2)
- 2- Certaines modalités sont inutilisables comme « Non applicable », « Manquante » ou « Autre ». Pour certaines variables, la modalité « Non applicable » était en fait utilisable, par exemple pour « Partage les toilettes avec d'autres ménages » les 1400 individus nous avons simplement changé la modalité « Non applicable » par « Pas de toilettes » car en effet le questionnaire demande aux ménages qui ont répondu « Pas de toilettes » à la question « Quel genre de toilettes les membres de votre ménage utilisent-ils habituellement ? » de sauter la question « Partagez-vous vos toilettes avec d'autres ménages ? ». Cela est notifié par la valeur entourée en rouge.

108	What kind of toilet facility do members of your household usually use?	FLUSH OR POUR FLUSH TOILET FLUSH TO PIPED SEWER SYSTEM... 11 FLUSH TO SEPTIC TANK 12 FLUSH TO PIT LATRINE 13 FLUSH TO SOMEWHERE ELSE 14 FLUSH, DON'T KNOW WHERE 15 PIT LATRINE VENTILATED IMPROVED PIT LATRINE 21 PIT LATRINE WITH SLAB 22 PIT LATRINE WITHOUT SLAB/ OPEN PIT 23 COMPOSTING TOILET 31 BUCKET TOILET 41 HANGING TOILET/HANGING LATRINE . 51 NO FACILITY/BUSH/FIELD 61 OTHER 96 (SPECIFY)	→ 111
109	Do you share this toilet facility with other households?	YES 1 NO 2	→ 111

Figure 2 : Protocole du questionnaire ménage

Nous avons donc changé les modalités qui le nécessitaient, nous avons aussi supprimé les modalités « Manquante » et « Autres » car elles n'étaient pas utilisables. On s'est basé sur la construction du questionnaire pour tous nos choix.

Après avoir arrangé ces soucis, nous sommes passés de 9050 individus à 8938. Nous avons donc perdu 112 individus ce qui représente 1% de toute la table. Nous considérons que c'est une perte acceptable au vu de la facilité de traitement que cela offre en contrepartie. De plus, on remarque en comparant les tables avant et après traitement ci-dessous que les statistiques (ici des variables quantitatives) ne sont pas très différentes, nous n'avons donc pas biaisé l'échantillon en supprimant les individus aux modalités problématiques :

Table ménage avant traitement des modalités						
La procédure MEANS						
Variable	Libellé	N	Moyenne	Ec-type	Minimum	Maximum
HV009	Number of household members	9057	4.2525119	2.4842593	1.0000000	19.0000000
HV014	Number of children 5 and under	9057	0.7869052	0.9783843	0	6.0000000
HV204	Time to get to water source	9047	412.9828672	470.3324524	1.0000000	998.0000000
HV216	Rooms used for sleeping	9050	1.8290608	1.0485273	0	16.0000000
HV220	Age of head of household	9051	43.8404596	15.9940998	13.0000000	98.0000000
HV271	Wealth index factor score (5 decimals)	9057	0.000552059	100000.00	-209666.00	262456.00

Table ménage après traitement des modalités						
La procédure MEANS						
Variable	Libellé	N	Moyenne	Ec-type	Minimum	Maximum
HV009	Number of household members	8938	4.2584471	2.4855178	1.0000000	19.0000000
HV014	Number of children 5 and under	8938	0.7870888	0.9779847	0	6.0000000
HV204	Time to get to water source	8928	412.2062052	470.2669460	1.0000000	998.0000000
HV216	Rooms used for sleeping	8932	1.8311688	1.0501586	0	16.0000000
HV220	Age of head of household	8932	43.8048589	15.9647077	13.0000000	98.0000000
HV271	Wealth index factor score (5 decimals)	8938	105.6072947	99946.89	-204548.00	262456.00

Tableaux 1 : Table ménage avant et après traitement

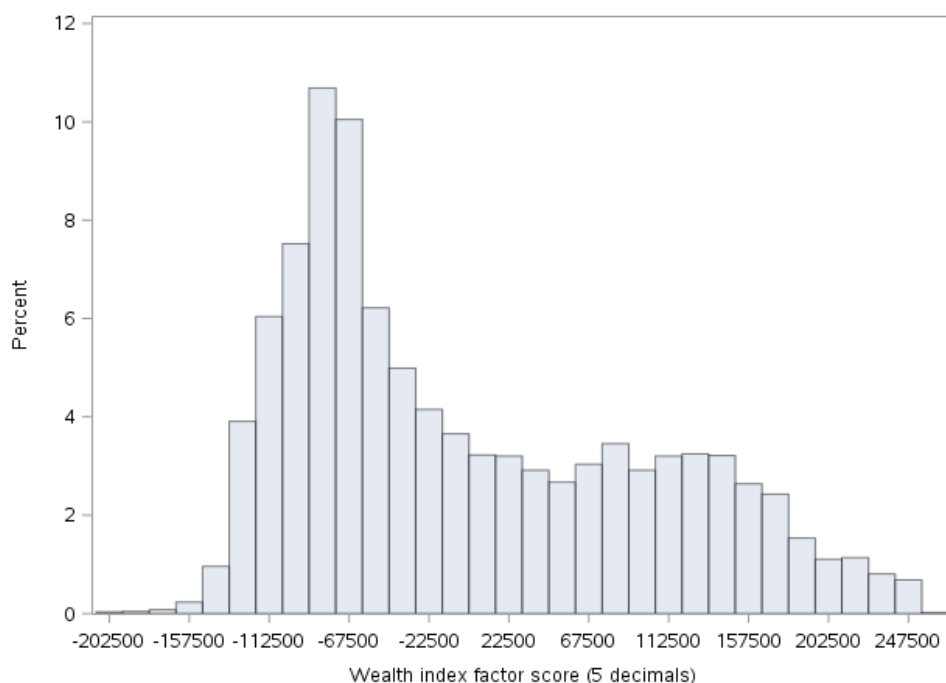
B. Étude univariée

Pour commencer, nous ferons une rapide visualisation de l'ensemble des variables de la table ménage afin de se familiariser avec l'échantillon. Ensuite nous nous focaliserons plus principalement sur nos variables d'intérêt, c'est-à-dire l'index de richesse, le facteur score de richesse ainsi que le niveau d'éducation du chef de ménage.

Pour commencer, regardons attentivement les statistiques précédentes. Nous pouvons voir qu'un ménage moyen contient quatre membres avec un enfant de moins de cinq ans (au maximum ils sont six). La plupart des ménages n'ont qu'une ou deux chambres à coucher. L'âge moyen du chef de ménage est de 44 ans, mais cet âge varie entre 13 et 98 ans.

Les tris à plat des variables qualitatives (voir annexe 1) nous donnent des informations supplémentaires : les régions sont toutes pareillement représentées (environ 13% de la population pour chaque) sauf le nord-est qui ne contient que 7% de la population kenyane. La grande majorité (68%) vit dans un espace rural. En ce qui concerne la qualité des équipements des ménages, elle est assez problématique. En effet, 60% des ménages sont alimentés par une eau (potable ou non) de qualité médiocre ou risquée. Pour les toilettes, 66% des ménages utilisent des latrines à fosse simple ce qui pose des risques de contamination élevé ; pire encore 17% des ménages ne possèdent pas d'installation de toilettes et font leur besoin dans la nature ce qui est encore plus risqué. En ce qui concerne les équipements, les trois quarts des ménages n'ont pas accès à l'électricité et la même proportion vit dans un environnement précaire en ce qui concerne le matériau de leurs habitations.

Intéressons-nous à présent aux variables que nous avons choisi. Commençons par le facteur de richesse.

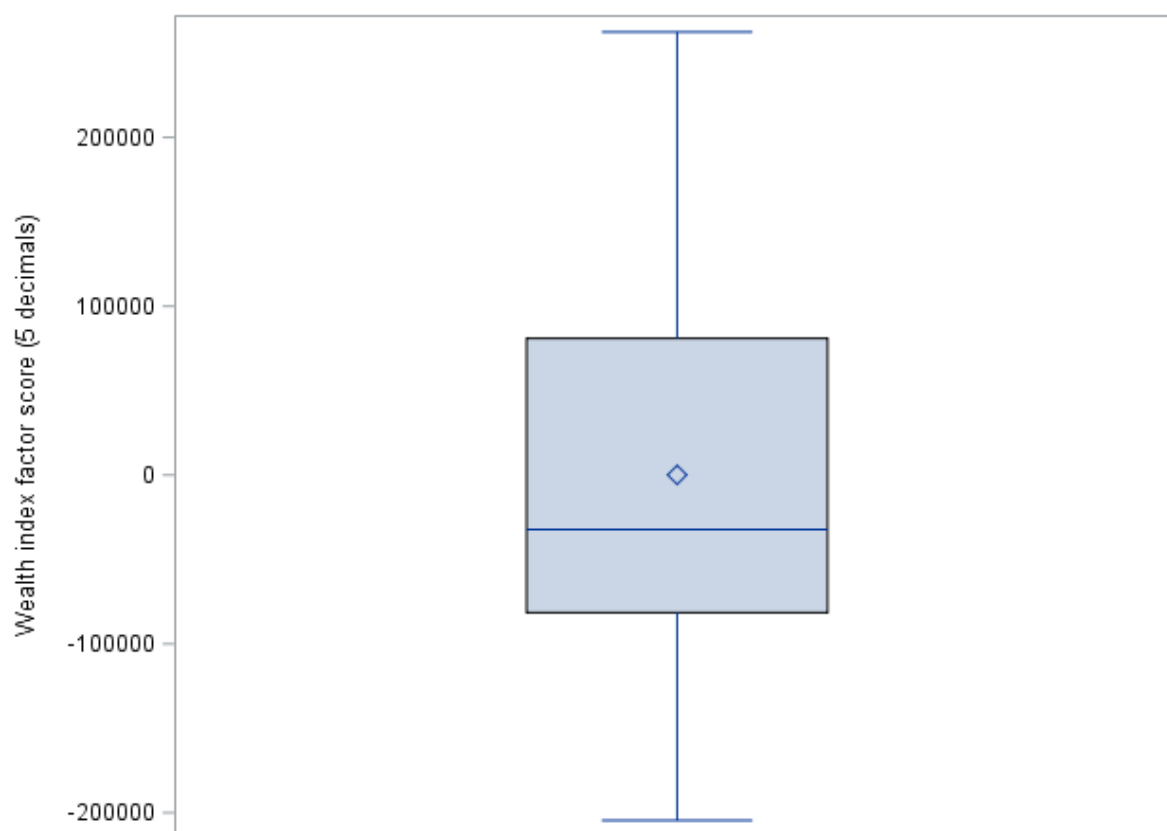


Graphique 1 : Histogramme de la distribution du facteur de richesse

Le facteur de richesse est un score calculé sur la base des possessions matérielles (comme un réfrigérateur ou un lit) ou non matérielles (comme l'éducation, le sommeil ou le temps nécessaire pour avoir accès à de l'eau) d'un ménage. Chaque possession ou non possession se voit attribué un nombre

positif ou négatif généré grâce à une analyse en composante principale. Ces valeurs ont été additionnées pour obtenir le facteur de richesse.

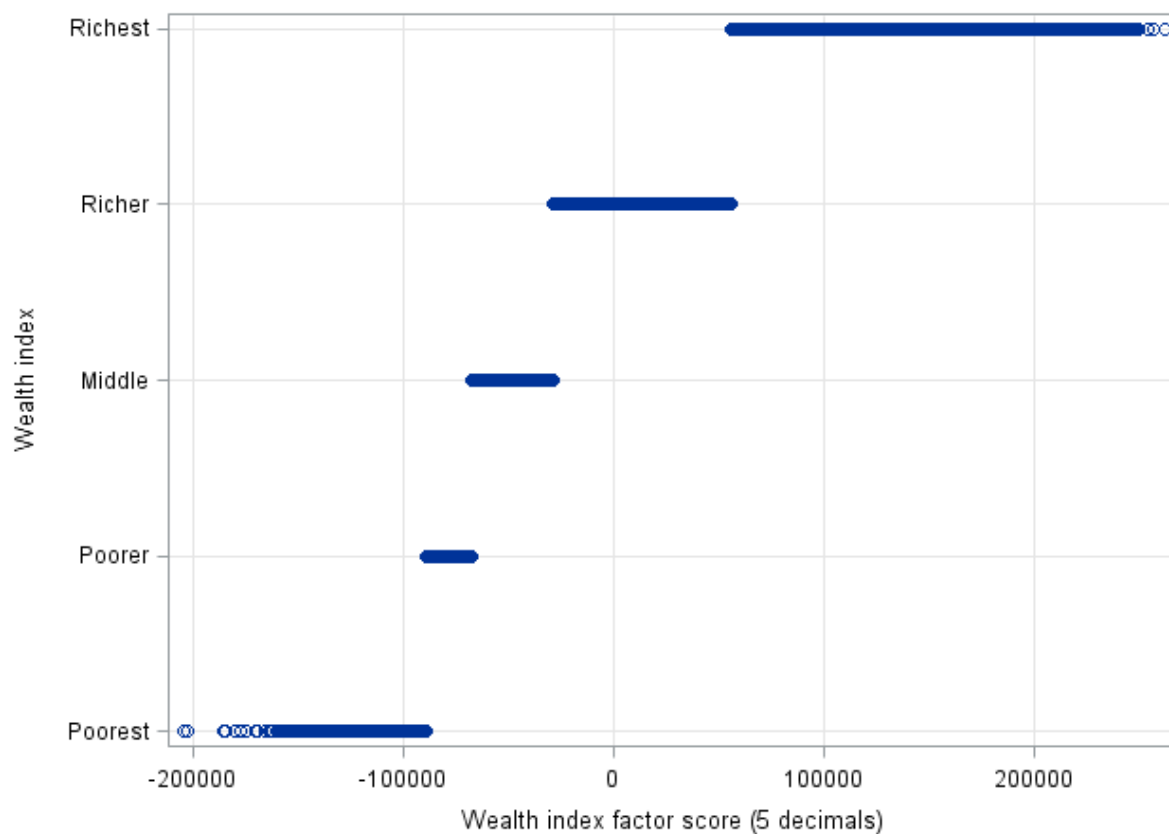
Le diagramme en barres verticales précédent nous permet de voir la répartition de la richesse de notre échantillon, on observe d'ailleurs un pic entre -112500 et -67500. Cela signifie qu'un nombre important de Kenyans est concentré autour de ces valeurs et ne possède pas de matériel nécessaires à une vie qualifiée de moyenne (score de 0). Il est aussi intéressant de visualiser cette variable dans un box-plot :



Graphique 2 : Box-plot du facteur de richesse

Il nous montre en plus que l'échantillon est hétérogène. En effet, bien que la moyenne est proche de zéro, les 95% de la population possèdent un score de richesse allant de -100000 à 100000.

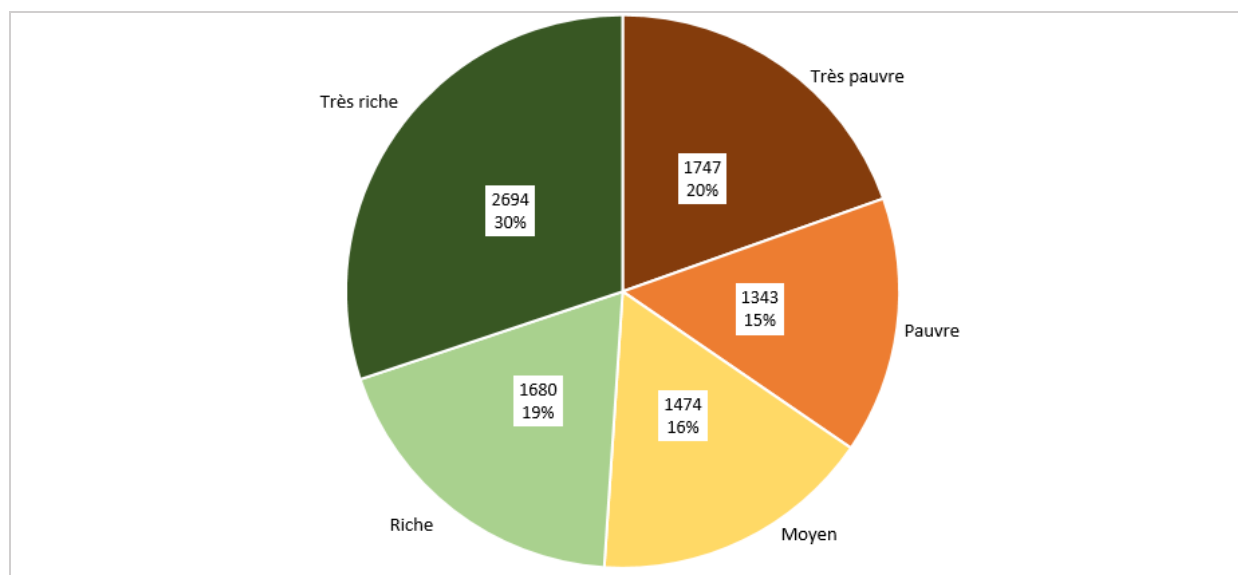
A partir du facteur de richesse, un index de richesse a été calculé. Il est plus simple à utiliser car il ne possède que cinq modalités : « les plus pauvres », « pauvres », « moyens », « riches » et « les plus riches ».



Graphique 3 : Nuage de point de l'indice de richesse en fonction du facteur de richesse

Ce nuage de point montre comment la variable « index de richesse » a été construite. On ne peut pas analyser ce graphique car les points se superpose mais il nous permet de savoir quelles valeurs du score de richesse équivalent à quel index de richesse.

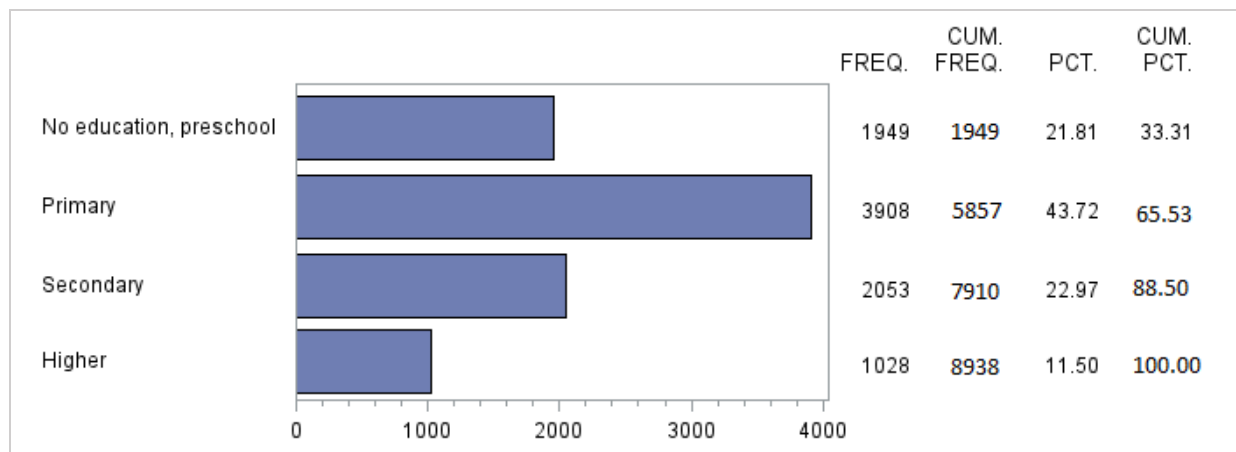
Pour analyser la variable « index de richesse » nous avons créé un diagramme circulaire :



Graphique 4 : Diagramme de la répartition de l'indice de richesse dans l'échantillon

Cette illustration montre que 20% de la population est extrêmement pauvre et 35% est en dessous du seuil moyen en terme de caractéristiques du niveau de vie à long terme des ménages.

Regardons maintenant les caractéristiques du chef de ménage en terme d'éducation. C'est en effet cette variable qui nous intéresse le plus :



Graphique 5 : Histogramme de la distribution du niveau d'éducation du chef de ménage

On obtient alors des informations intéressantes. 44% des chefs de ménage de notre échantillon n'ont eu qu'une brève éducation et ont, au mieux, complété le cycle primaire. 23% ont terminé le cycle secondaire et seulement 12% l'ont dépassé.

Ce qui nous intéresserait maintenant, c'est de savoir si ces inégalités d'éducation des chefs de ménage ont un lien avec les niveaux de richesse des ménages.

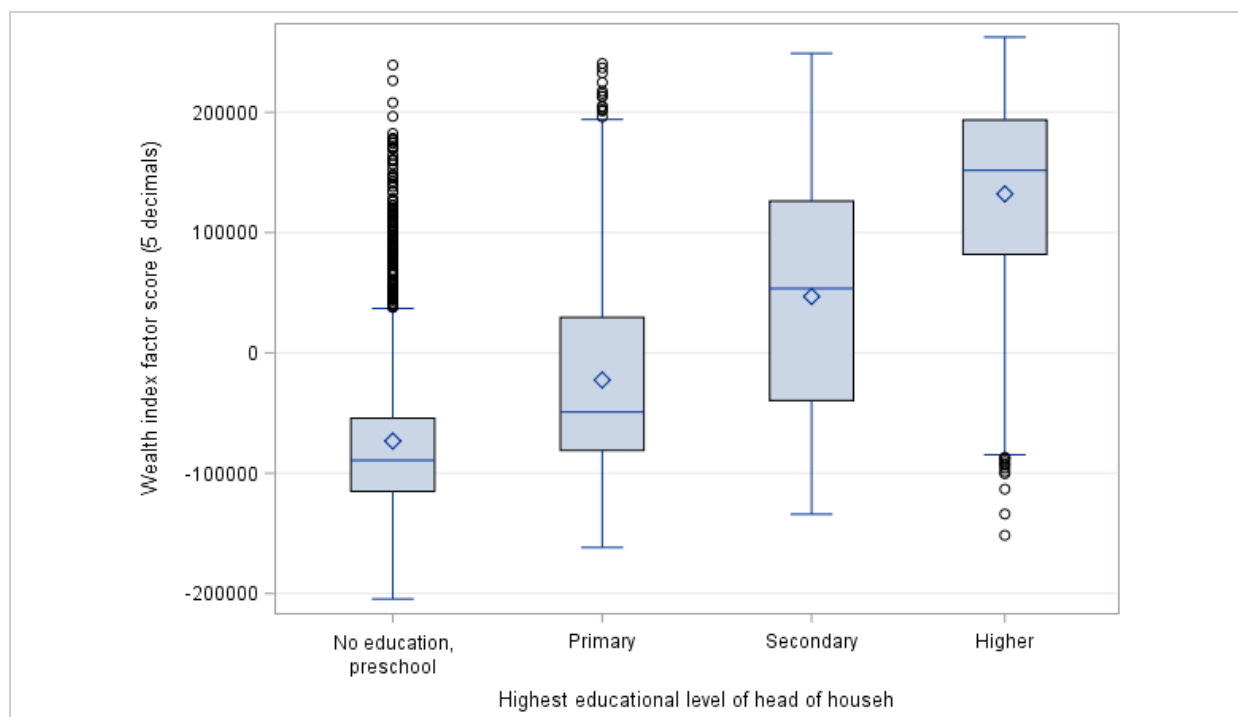
C. Étude bivariable

Pour répondre à notre problématique, nous allons croiser la variable « Niveau d'éducation du chef de ménage » avec les variables relatives à la richesse ainsi qu'au confort des ménages :

- Place de résidence (d'après le rapport complet KDHS 2008-09, le fait de vivre dans un milieu rural ou urbain est un indicateur de richesse)
- Source d'eau potable (nous n'avons pas besoin de prendre aussi la source d'eau non-potable car, on le voit dans les tris à plat à l'annexe 1, les deux types d'eau proviennent de la même source)
- Type d'installation de toilettes
- Accès à l'électricité
- Possède un réfrigérateur
- Qualité du sol, des murs et du plafond
- Possède une ou des moustiquaires

Nous croiserons tout d'abord le niveau d'éducation avec : « Score de richesse » et « Index de richesse » qui sont, comme il a été dit précédemment, des agrégations de toutes les possessions matérielles ou non des ménages. La variable du niveau d'éducation est qualitative tout comme l'index de richesse mais celle du score de richesse est quantitative. Il est utile d'avoir des données quantitatives et qualitatives pour la richesse car cela nous permettra différentes méthodes et visualisations statistiques.

Pour commencer nous avons réalisé un box-plot afin d'avoir une image de la correspondance entre l'éducation du chef de ménage et les niveaux de richesse :



Graphique 6 : Box-plot des scores de richesse groupés par les niveaux d'éducation du chef de ménage

Il y a clairement un lien statistique qui relie ces deux variables. En effet, plus l'éducation du chef de ménage est élevée, plus la richesse du ménage est élevée. On ne peut cependant pas conclure avec notre étude si c'est la richesse qui influe sur l'éducation ou l'inverse.

Mettons maintenant en relation le niveau d'éducation du chef de ménage avec les variables de la table ménage choisies précédemment.

➤ Tests d'indépendance

On veut vérifier ici que la variable « Niveau d'éducation du chef de ménage » dépend bien des variables de la table ménage que nous avons sélectionné. Nous allons donc effectuer un test du Chi².

L'hypothèse nulle H₀ est donc que la variable X et la variable Y sont indépendantes.

L'hypothèse alternative H₁ est que ces deux variables sont dépendantes.

$$\chi^2 = \sum_{i,j} \frac{(n_{ij} - n_{ij}^*)^2}{n_{ij}^*}$$

La statistique du test utilisée est celle du Chi² : c'est une mesure de la distance de l'indépendance entre deux variables. Quand la distance est grande, cela signifie que les variables sont dépendantes l'une de l'autre. Dans la formule, n_{ij} est l'effectif observé et n_{ij}^{*} est l'effectif que l'on aurait observé si les variables étaient indépendantes.

La règle de décision est que si la p-valeur est inférieure au niveau de risque (on prend ici 5%) on rejette l'hypothèse nulle.

Commençons donc de tester l'indépendance entre la variable « Niveau d'éducation » et la variable « Place de résidence ». On obtient tout d'abord le tableau croisé que l'on peut regarder pour se faire une vague idée de la relation entre ces deux variables.

Fréquence Pourcentage Pct de ligne Pct de col.	Table of cHV106_01 by cHV026					
	cHV106_01(Highest educational level of head of househ)	cHV026(Place of residence)				
		Capital or large city	Small city	Town	Countryside	Total
	No education, preschool	62	2	182	1703	1949
		0.69	0.02	2.04	19.05	21.81
		3.18	0.10	9.34	87.38	
		4.13	6.67	13.63	28.05	
	Primary	346	12	505	3045	3908
		3.87	0.13	5.65	34.07	43.72
		8.85	0.31	12.92	77.92	
		23.05	40.00	37.83	50.15	
	Secondary	538	12	440	1063	2053
		6.02	0.13	4.92	11.89	22.97
		26.21	0.58	21.43	51.78	
		35.84	40.00	32.96	17.51	
	Higher	555	4	208	261	1028
		6.21	0.04	2.33	2.92	11.50
		53.99	0.39	20.23	25.39	
		36.98	13.33	15.58	4.30	
	Total	1501	30	1335	6072	8938
		16.79	0.34	14.94	67.93	100.00

Tableau 2 : Tableau croisé des variables "niveau d'éducation du chef de ménage" et "place de résidence du ménage"

On voit rapidement que 87% des chefs de ménage qui n'ont pas eu d'éducation ou seulement au préscolaire (maternelle) vivent dans la campagne du Kenya. Alors que 80% de ceux qui ont dépassé le secondaire vivent maintenant dans la capital Nairobi ou dans une grande ville. On voit déjà bien qu'il y a un lien entre le niveau d'éducation du chef de ménage et la place de résidence du ménage. Mais ce qui le confirme (avec un risque 5%) est le test du χ^2 que nous avons réalisé pour ces deux variables.

Statistiques pour la table de CHV106_01 par CHV026			
Statistique	DDL	Valeur	Prob
Khi-2	9	1971.4474	<.0001
Test du rapport de vraisemblance	9	1861.2083	<.0001
Khi-2 de Mantel-Haenszel	1	1737.5712	<.0001
Coefficient Phi		0.4696	
Coefficient de contingence		0.4251	
V de Cramer		0.2712	

Tableau 3 : Statistiques des tests d'indépendance pour la table "niveau d'éducation" par "place de résidence"

On voit bien ici que le niveau d'éducation et la place de résidence d'un ménage sont liées. De plus, grâce au tableau croisé précédent on peut en conclure que c'est le niveau d'éducation du chef de ménage qui influe sur la place de résidence d'un ménage.

Nous avons effectué ce test pour toutes les autres variables que nous avons choisi dans la table ménage toujours en fonction du niveau d'éducation du chef de ménage et voici le résultat :

Variable X	Variable Y	DDL	Valeur du Chi ²	Prob.
Niveau d'éducation du chef de ménage	Place de résidence	9	1971.44	< 0,0001
	Source d'eau potable	6	758.15	< 0,0001
	Installation de toilettes	6	3360.80	< 0,0001
	Accès à l'électricité	3	2470.36	< 0,0001
	Accès à un réfrigérateur	3	1993.06	< 0,0001
	Qualité du sol	6	2076.28	< 0,0001
	Qualité des murs	6	1654.65	< 0,0001
	Qualité du plafond	6	2877.76	< 0,0001
	Moustiquaire	3	114.74	< 0,0001

Tableau 4 : Statistiques du test chi² entre la variable "niveau d'éducation" par rapport à neuf variables d'équipements des ménages

On voit donc bien que le niveau d'éducation du chef de ménage a un lien statistique avec le confort des ménages. En effet, tous les tests d'indépendance aboutissent à un rejet de l'hypothèse nulle.

III. Analyse des Correspondances Multiple

A. Méthodologie

Nous avons tout d'abord commencé par recoder les modalités :

Variable	Nb de modalités	Code des modalités				
Source of drinking water	3 modalités	Eau++	Eau+	Eau-		
Type of toilet facility	3 modalités	T++	T+	T-		
Has electricity	2 modalités		Ele+	Ele-		
Has radio	2 modalités		Rad+	Rad-		
Has television	2 modalités		Tv+	Tv-		
Has refrigerator	2 modalités		Ref+	Ref-		
Has telephone	2 modalités		Tel+	Tel-		
Main floor material	3 modalités	Sol++	Sol+	Sol-		
Type of cooking fuel	3 modalités	C++	C+	C-		
Have bednet for sleeping	2 modalités		M+	M-		
Wealth index	5 modalités	r--	r-	r	r+	r++
Highest educational level	4 modalités	E-	E	E+	E++	

Tableau 5 : Codes des modalités pour les 12 variables d'intérêt

Pour certaines variables il y a des modalités qui posent problème (surlignées en jaune dans le tableau ci-dessus). Soit il y avait trop de modalités (Wealth index) soit certaines modalités étaient de trop faibles effectifs (qualité du sol, énergie de cuisson). FI+ qui correspond à une qualité de sol rudimentaire a été supprimée et les 39 individus correspondants sont passés en FI- (car c'était ce qui correspondait le plus). Pour le type d'énergie de cuisson les individus en C- (qualité précaire) sont passés en C+ (qualité rudimentaire).

Nous avons alors notre tableau de codage condensé dont voici une partie (pour montrer brièvement à quoi il ressemble) :

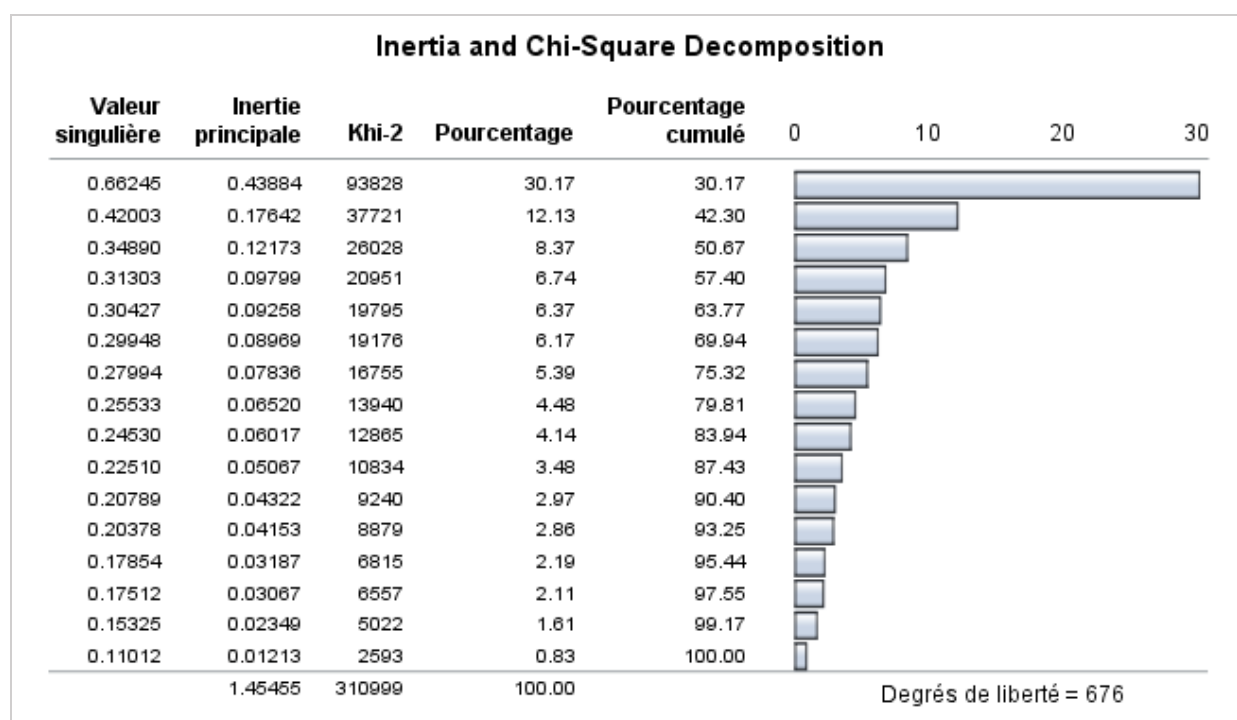
Obs.	cHV201	cHV205	cHV206	cHV207	cHV208	cHV209	cHV213	cHV221	cHV226	cHV227	cHV270	cHV106_01
1	Eau-	T+	Ele-	Rad+	Tv-	Ref-	Sol-	Tel-	C+	M+	r-	E-
2	Eau++	T+	Ele-	Rad-	Tv-	Ref-	Sol-	Tel-	C+	M-	r-	E-
3	Eau-	T+	Ele-	Rad-	Tv-	Ref-	Sol-	Tel-	C+	M+	r-	E
4	Eau-	T-	Ele-	Rad+	Tv-	Ref-	Sol-	Tel-	C+	M+	r-	E
5	Eau-	T-	Ele-	Rad-	Tv-	Ref-	Sol-	Tel-	C+	M-	r-	E-
6	Eau-	T-	Ele-	Rad-	Tv-	Ref-	Sol-	Tel-	C+	M+	r-	E
7	Eau++	T-	Ele-	Rad-	Tv-	Ref-	Sol-	Tel-	C+	M-	r-	E
8	Eau-	T-	Ele-	Rad-	Tv-	Ref-	Sol-	Tel-	C+	M+	r-	E

Tableau 6 : Échantillon du tableau de codage condensé littéral

A partir de cela, on effectue l'ACM utilisant la table de Burt associée. Dans cette analyse, nous avons mis les variables associées à la qualité des équipements en active et la variable d'éducation en supplémentaire. Nous présenterons dans la suite du dossier les résultats de cette ACM.

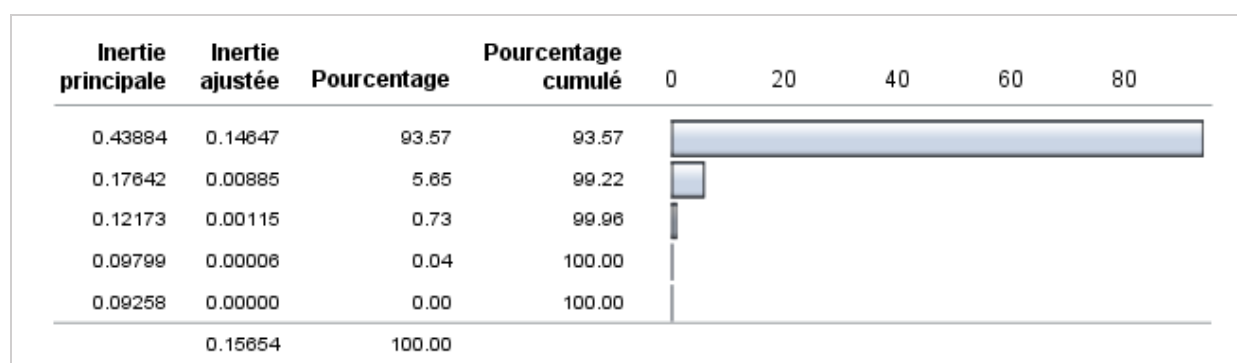
B. Résultats de l'ACM

Commençons par la décomposition de l'inertie :



Graphique 7 : Décomposition de l'inertie et statistiques du Chi²

On peut voir tout d'abord que 16 axes factoriels ont été représentés (il y en a normalement 19 mais ceux dont la valeur propre était trop faible n'ont pas été représentés). Le pourcentage d'inertie expliqué par le premier plan factoriel est de 42% on ne peut pas dire si ce pourcentage est élevé ou faible car il y a un nombre de modalités très important. Nous pouvons cependant regarder les taux d'inertie modifiés par la formule de Benzecri :



Graphique 8 : Décomposition de l'inertie modifiée (méthode de Benzecri)

On voit ici alors que le premier axe explique 95% de l'inertie totale et le second axe 4%. En tout, le premier plan factoriel explique 99% de l'inertie totale et on peut dire que le premier est le plus intéressant à étudier car il nous donne un nombre d'informations plus important.

Ensuite on regarde les contributions des modalités aux différents axes que l'on résume dans les deux tableaux suivants.

Contributions des modalités à l'axe 1 :

Coordonnées des colonnes		
	Dim1	Dim2
Eau+	0,6535	-0,593
Eau++	0,3183	0,1159
Eau-	-0,525	0,1431
T+	-0,22	-0,509
T++	1,7609	0,5657
T-	-0,857	1,4256
Ele+	1,4592	0,1383
Ele-	-0,486	-0,046
Rad+	0,249	-0,317
Rad-	-0,62	0,7899
Tv+	1,1829	-0,14
Tv-	-0,477	0,0566
Ref+	2,307	1,1491
Ref-	-0,211	-0,105
Sol++	0,8083	-0,318
Sol-	-0,661	0,2598
Tel+	2,5501	1,6571
Tel-	-0,09	-0,058
C+	-0,227	-0,103
C++	2,2292	1,0089
M+	0,0628	-0,076
M-	-0,114	0,1381
r	-0,495	-0,542
r+	-0,092	-1,004
r++	1,2403	0,0653
r-	-0,678	-0,074
r--	-0,885	1,3792

Contribution positive

T++
Ele+
Ref+
Tel+
C++

Contribution négative

T-
r--
r-

Contributions des modalités à l'axe 2 :

Contribution positive

T-
Ref+
Tel+
r--

Contribution négative

r+
Eau+

On remarque que les modalités qui font référence à une qualité convenable d'équipement (toilettes avec chasse d'eau, énergie de cuisson sûre et de bonne qualité, possession de télévision, réfrigérateur et téléphone) contribuent positivement à l'axe 1.

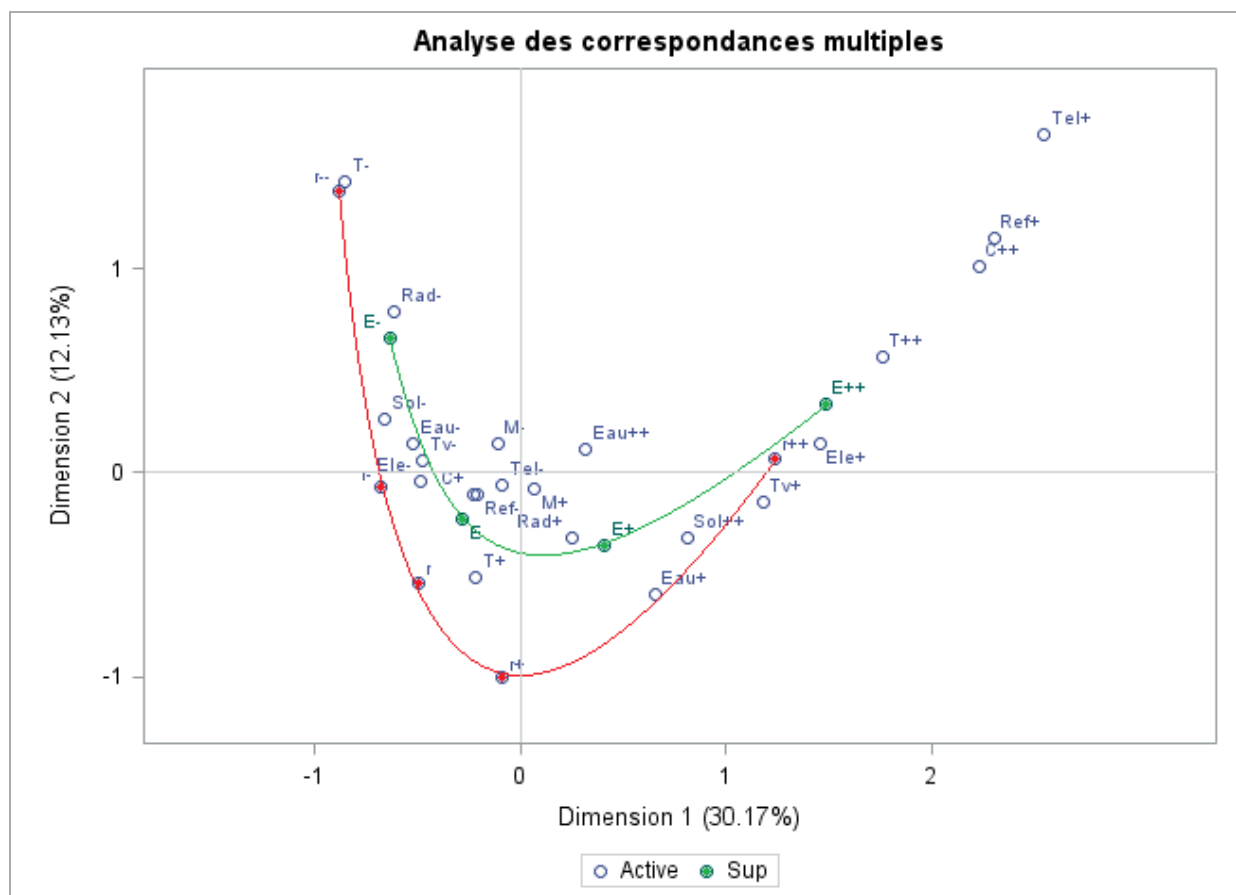
Le fait de ne pas posséder de toilette ou de faire partie d'un ménage globalement très pauvre ou pauvre contribue au contraire négativement à l'axe 1.

Tableau 7 : Heatmap des coordonnées des modalités sur le premier plan factoriel

En ce qui concerne l'axe 2 c'est le fait de posséder un téléphone et un réfrigérateur ainsi que le fait de ne pas posséder de toilettes et de faire partie d'un ménage globalement pauvre qui y contribuent positivement.

Et le fait d'avoir accès à une eau potable de moyenne qualité ainsi que de faire partie d'un ménage riche contribuent négativement à l'axe 2.

Regardons maintenant le graphe de l'analyse des correspondances multiples de l'axe 1 et 2. Les modalités relatives à l'éducation du chef de ménages sont en supplémentaires.



Graphique 9 : Cartographie des correspondances multiples (modalités Educations en supplémentaires)

La première chose qu'il faut examiner, c'est l'allure générale du nuage de point des modalités. Ici on voit clairement que l'axe 1 discrimine les modalités de ménages en deux parties opposées. D'une part on observe des équipements de ménages qualifiés de bonne qualité (comme des toilettes à chasse d'eau (T++) l'utilisation d'énergie de cuisson supérieure) et la possession d'équipements comme le téléphone ou un réfrigérateur. De l'autre part de l'axe, on a des équipements de mauvaise qualité ou la non-possession de ces équipements. De plus on voit clairement, grâce aux modalités r-- r- et r, que du côté négatif de l'axe 1 ce sont des ménages globalement pauvres alors que du côté positif ce sont des ménages très riches (ce qui est logique au vu des modalités d'équipements des ménages).

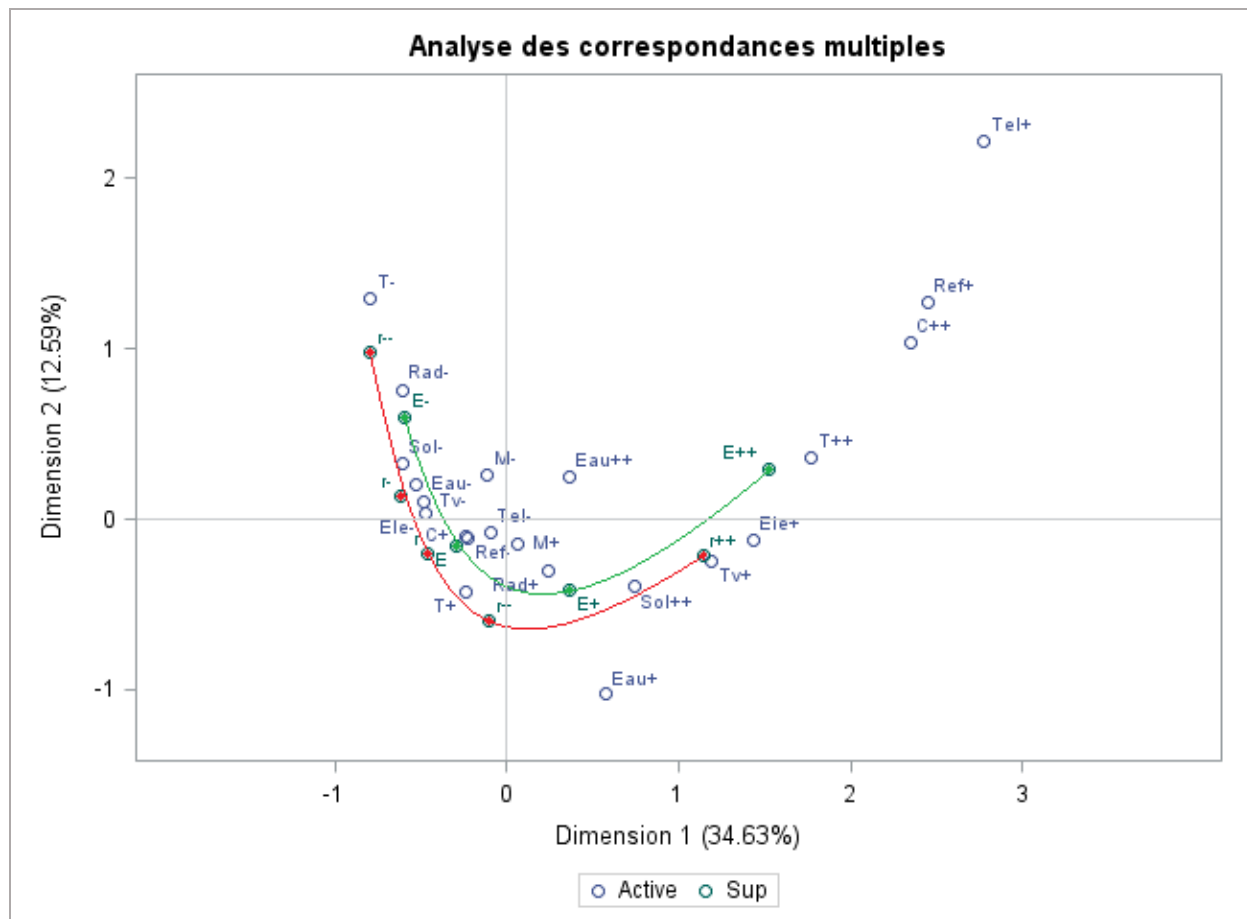
De plus, la non-possession d'équipements ainsi que la possession d'équipements de mauvaise qualité voire dangereuse est la norme pour l'ensemble des ménages enquêtés (sauf pour la radio, la majorité en possède au moins une). On voit cela de la même façon, ces modalités se regroupent à proximité de l'origine.

Il est remarquable de constater que le fait de posséder un téléphone, un réfrigérateur ou une énergie de cuisson de qualité sont des luxes que même les ménages les plus riches peinent à s'accorder. En effet, les modalités correspondantes se retrouvent à une distance très élevée de l'origine mais aussi très élevée de la modalité r++.

Mais ce qui est en lien avec notre problématique ici, c'est que l'on voit que les ménages pauvres (vivant dans un environnement précaire voire dangereux) correspondent à des ménages où le chef n'a pas dépassé le préscolaire. Les ménages dont le chef de ménage a dépassé le secondaire correspondent plutôt à des ménages suréquipés (proximité entre E++ et r++). Un fait intéressant à voir aussi est que la majorité des chefs de ménage n'ont pas dépassés le secondaire. En effet, les modalités E et E+ sont les plus proches de l'origine.

On remarque notamment grâce à la courbe tracée en vert sur le plan factoriel ci-dessus une structure du nuage de points intéressante car elle permet de voir pour chaque niveau d'éducation du chef de ménage la proximité avec les équipements des ménages.

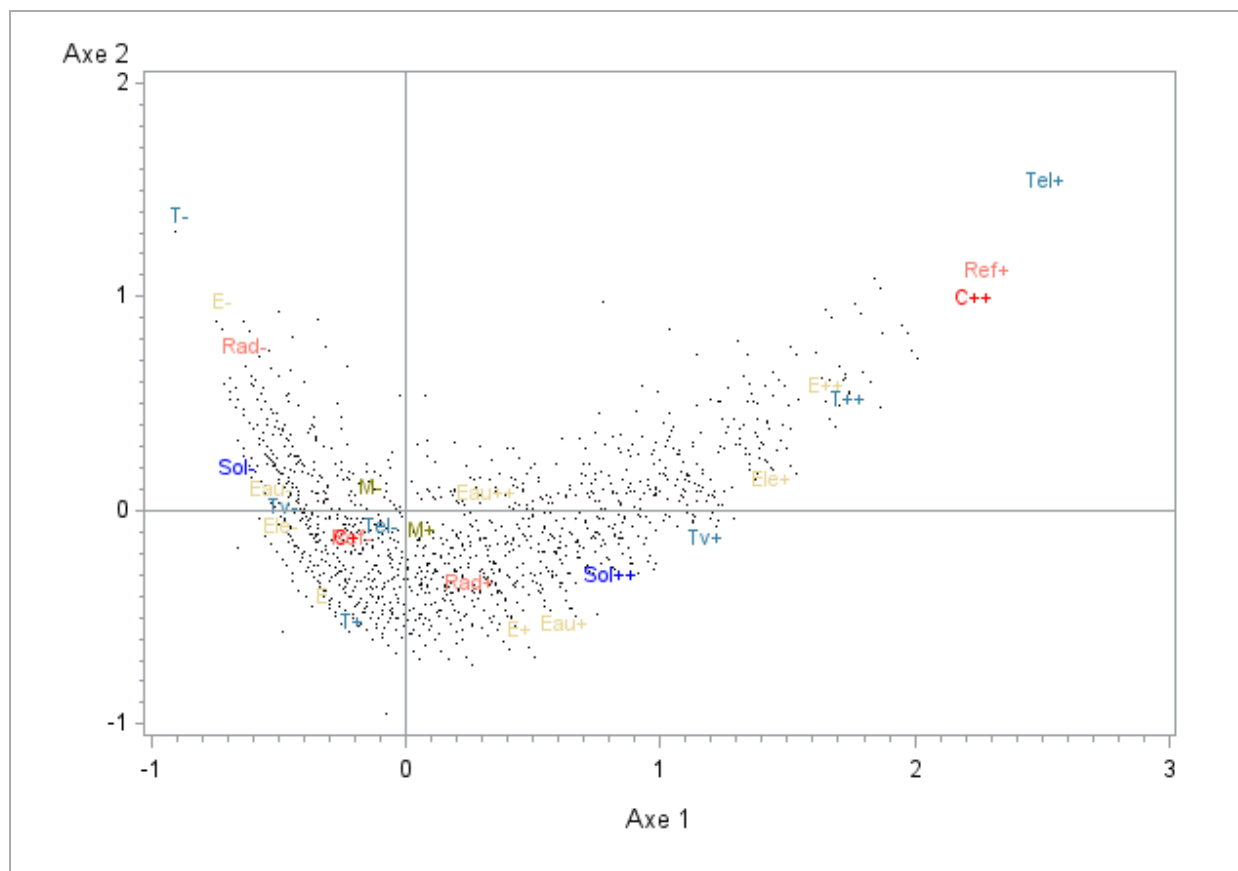
Cela est encore plus visible quand on effectue l'analyse en composante principale en assignant les modalités de richesses en supplémentaires :



Graphique 10 : Cartographie des correspondances multiples (modalités Educations et Richesses en supplémentaires)

On voit alors que la courbe des répartitions des richesses (en rouge) se superpose presque parfaitement avec la courbe des répartitions d'éducation des chefs de ménage avec en bonus un effet Guttman caractéristique des modalités hiérarchiques (bonne qualité, qualité moyenne, mauvaise qualité).

Cela est en fait assez visible sur le graphique du nuage des profils de ménages sur le premier plan factoriel :



Graphique 11 : Cartographie des correspondances multiples (profils ménages et modalités)

IV. Conclusion

Notre étude a très clairement mis en lumière un lien entre l'éducation du chef de ménage et le niveau de richesse ainsi que le confort des ménages en question.

Les ménages dont les chefs de famille ont eu droit à une éducation dans le supérieur ont accès à des équipements secondaires tels que la télévision ou le téléphone mais c'est aussi eux qui ont le plus droit à des équipements vitaux tels que l'électricité, des moyens de cuissons sans danger ou un cadre de vie confortable (Sol++).

Les ménages dont les chefs de famille n'ont pas dépassé le secondaire sont plutôt caractérisés par un manque d'équipements vitaux tels que des toilettes ou une eau potable saine, aucun équipement de loisir ou d'informations même pas la radio, et un cadre de vie très inconfortable (sols en terre/boue/sable).

L'étude n'a cependant pas permis de décrire un effet de causalité, il serait intéressant d'étudier cela avec peut-être un modèle de régression afin de savoir s'il faut améliorer les infrastructures d'éducation afin ou si le fait d'améliorer l'accès à certains types d'équipement des Kenyans aura pour effet d'améliorer le niveau d'éducation. En effet il est possible que ce soit le niveau de richesse faible de certains ménages qui oblige leurs membres de se lancer directement dans un travail manuel sans avoir la possibilité d'accéder à des études. Il est possible aussi que ce soit le fait de ne pas se consacrer à des études qui ont pour effet de ne pas avoir accès à un cadre de vie confortable. On ne peut pas, avec cette étude, conclure sur ce point ; c'est là la limite d'une analyse factorielle.

De plus, toute l'analyse a été réalisée sans avoir pris en compte les poids des individus. Nous avons donc décrit l'échantillon des individus mais nous ne pouvons pas généraliser cela à toute la population kenyane de 2009. Il serait plus correct de refaire l'analyse en prenant en compte ce point non négligeable.

V. Bibliographie

- Central Intelligence Agency. (2017, Décembre). **Overview of Kenya - Africa**. Récupéré sur The world factbook: <https://www.cia.gov/library/publications/the-world-factbook/geos/ke.html>
- Kenya National Bureau of Statistics. (2008). **KDHS Survey Household questionnaire**. Nairobi.
- Kenya National Bureau of Statistics. (2008-09). **Kenya Demographic and Health Survey**. Nairobi.
- SAS Entreprise. (2012). **Les analyses en correspondances multiples dans SAS**. France.

VI. Table des matières

I.	INTRODUCTION.....	2
A.	PRESENTATION DU KENYA.....	2
	FIGURE 1 : LES HUIT REGIONS DU KENYA	2
B.	PRESENTATION DE L'ENQUETE KDHS 2008-09	2
C.	PRESENTATION DU PROJET	3
II.	ANALYSE PRELIMINAIRE.....	4
A.	REFLEXION PREALABLE A L'ETUDE	4
	FIGURE 2 : PROTOCOLE DU QUESTIONNAIRE MENAGE	5
	TABLEAUX 1 : TABLE MENAGE AVANT ET APRES TRAITEMENT	5
B.	ETUDE UNIVARIEE	6
	GRAPHIQUE 1 : HISTOGRAMME DE LA DISTRIBUTION DU FACTEUR DE RICHESSE	6
	GRAPHIQUE 2 : BOX-PLOT DU FACTEUR DE RICHESSE	7
	GRAPHIQUE 3 : NUAGE DE POINT DE L'INDICE DE RICHESSE EN FONCTION DU FACTEUR DE RICHESSE	8
	GRAPHIQUE 4 : DIAGRAMME DE LA REPARTITION DE L'INDICE DE RICHESSE DANS L'ECHANTILLON	8
	GRAPHIQUE 5 : HISTOGRAMME DE LA DISTRIBUTION DU NIVEAU D'EDUCATION DU CHEF DE MENAGE	9
C.	ETUDE BIVARIEE	9
	GRAPHIQUE 6 : BOX-PLOT DES SCORES DE RICHESSE GROUPES PAR LES NIVEAUX D'EDUCATION DU CHEF DE MENAGE.....	10
	TABLEAU 2 : TABLEAU CROISE DE "NIVEAU D'EDUCATION DU CHEF DE MENAGE" PAR "PLACE DE RESIDENCE DU MENAGE"	11
	TABLEAU 3 : STATISTIQUES DES TESTS D'INDEPENDANCE POUR LA TABLE "NIVEAU D'EDUCATION" PAR "PLACE DE RESIDENCE"	12
	TABLEAU 4 : TEST χ^2 ENTRE "NIVEAU D'EDUCATION" PAR RAPPORT A NEUF VARIABLES D'EQUIPEMENTS DES MENAGES.....	12
III.	ANALYSE DES CORRESPONDANCES MULTIPLE	13
A.	METHODOLOGIE.....	13
	TABLEAU 5 : CODES DES MODALITES POUR LES 12 VARIABLES D'INTERET.....	13
	TABLEAU 6 : ÉCHANTILLON DU TABLEAU DE CODAGE CONDENSE LITTERAL	13
B.	RESULTATS DE L'ACM	14
	GRAPHIQUE 7 : DECOMPOSITION DE L'INERTIE ET STATISTIQUES DU χ^2	14
	GRAPHIQUE 8 : DECOMPOSITION DE L'INERTIE MODIFIEE (METHODE DE BENZECRI)	14
	TABLEAU 7 : HEATMAP DES COORDONNEES DES MODALITES SUR LE PREMIER PLAN FACTORIEL.....	15
	GRAPHIQUE 9 : CARTOGRAPHIE DES CORRESPONDANCES MULTIPLES (EDUCATIONS EN SUPPLEMENTAIRES)	16
	GRAPHIQUE 10 : CARTOGRAPHIE DES CORRESPONDANCES MULTIPLES (EDUCATIONS ET RICHESSES EN SUPPLEMENTAIRES).....	17
	GRAPHIQUE 11 : CARTOGRAPHIE DES CORRESPONDANCES MULTIPLES (PROFILS MENAGES ET MODALITES)	18
IV.	CONCLUSION	19
V.	BIBLIOGRAPHIE.....	20
VI.	TABLE DES MATIERES.....	21
VII.	ANNEXES	22
	ANNEXE 1 : TRIS A PLAT AVANT TRAITEMENT DES MODALITES	
	ANNEXE 2 : TRIS A PLAT APRES TRAITEMENT DES MODALITES	
	ANNEXE 4 : GRAPHIQUES DESCRIPTIFS.....	
	ANNEXE 5 : TABLEAUX CROISES ET TESTS D'INDEPENDANCE	
	ANNEXE 6 : ANALYSE DES CORRESPONDANCES MULTIPLES.....	
	ANNEXE 7 : DICTIONNAIRE DES VARIABLES D'INTERETS ET DES FORMATS SAS	