

Master 1 MAS - DS Données catégorielles

Compte rendu du mini projet

Réalisé par :

 Saïda GUEZOU

 Benoit GILLES

Encadré par :

 Laurence REBOUL

 Nicolas PECH

2020/2021

Table des matières

Remerciements	3
Résumé.....	4
Abstract.....	4
Liste des figures	5
Liste des sigles et abbréviations	6
I. Introduction	7
II. Le Kenya	8
III. Enquête démographique et socio-économique de 2014	10
IV. Problématique	12
V. Méthodologie	13
VI. Analyses statistiques.....	16
A. Etude univariée (Tris à plat).....	16
1. Représentation graphique.....	16
B. Etude bivariable (Tris croisés).....	25
1. Création des variables pauvreté extrême et pauvreté matérielle	25
2. Croisement de couple de variables pertinentes	26
C. ACM.....	33
C.1 Contribution et coordonnées des modalités sur l'axe 1 et 2	33
C.2 Qualités de représentation des modalités	34
C.3 Inertie et taux d'inertie modifiés	35
C.4 ACM à partir du tableau de Burt.....	36
C.5 ACM sur les modalités à partir du tableau disjonctif complet	37
D. ACM avec les variables région et lieu de résidence	38
D.1 Qualités de représentation.....	38
D.2 Coordonnées et contributions	39
D.3 ACM sur les modalités à partir du tableau de Burt	40
D.4 ACM avec les variables de localisation en supplémentaire	41
Synthèse, critique et conclusion	42
Bibliographie	44
Annexe 1 : Recodage des variables.....	45

Remerciements

Avant de commencer, nous voudrions remercier madame Laurence Reboul et monsieur Nicolas Pech : tout d'abord, pour nous avoir donné l'opportunité de réaliser ce travail enrichissant, puis, pour toute leur aide et encadrement tout au long de la réalisation de ce projet.

Résumé

Dans ce rapport, nous souhaitons répondre à la question suivante : Comment la pauvreté est-elle répartie au sein du Kenya ? Nous nous basons sur une enquête, la Kenya Demographic and Health Survey (KDHS), effectuée en 2014. Les données recueillies nous serviront à analyser la proportion de biens matériels ainsi que de services au sein des foyers interrogés lors de l'enquête.

Afin de pouvoir qualifier un foyer comme pauvre ou riche, nous avons décidé de créer des indicateurs de richesse à partir des variables dont nous disposons. Ces indicateurs seront expliqués pendant l'étude. Nous travaillerons à partir de la base de données ménage.

Un regret que nous partageons est de ne pas avoir eu accès aux salaires des foyers, ce qui nous aurait aidé lors de notre étude. Nous avons pu néanmoins dégager plusieurs réponses et valider plusieurs hypothèses quant à nos questionnements.

Abstract

In this report, we wish to answer the following question: How is poverty distributed within Kenya? We are based on a survey, the Kenya Demographic and Health Survey (KDHS), conducted in 2014. The data collected will be used to analyze the proportion of material goods as well as services within the households surveyed in the survey.

To be able to qualify a household as poor or rich, we decided to create wealth indicators based on the variables available to us. These indicators will be explained during the study. We will work from the household database.

One regret we share is that we did not have access to household wages, which would have helped us in our study. Nevertheless, we were able to identify several answers and validate several hypotheses regarding our questions.

Liste des figures

Figure 1 : Carte du Kenya.....	8
Figure 2 : Carte sur la diversité des communautés au Kenya	9
Figure 3 : Type de source d'eau potable par ménage	16
Figure 4 : Effectifs et fréquences de l'accès à l'eau potable	17
Figure 5 : Accès à l'électricité	17
Figure 6 : Effectifs et fréquences de l'accès à l'électricité	18
Figure 7 : Conditions matérielles des murs	18
Figure 8 : Conditions matérielles du toit	18
Figure 9 : Types de toilettes	19
Figure 10 : Possession de la télévision	20
Figure 11 Effectifs et fréquences de l'accès à la télévision	20
Figure 12 : Accès à la radio	20
Figure 13 : Effectifs et fréquences de la possession de la radio.....	21
Figure 14 : Type de cuisine	21
Figure 15 : Effectifs et fréquences de la possession de voiture	21
Figure 16 : Possession d'un réfrigérateur.....	22
Figure 17 : Possession de la voiture	22
Figure 18 : Possession de la moto	22
Figure 19 : Possession d'un téléphone	22
Figure 20 : Effectifs et fréquences de la variable région	22
Figure 21 : Nombre de ménage par région	23
Figure 22 : Effectifs et fréquences de la variable lieu de résidence	23
Figure 23 : Nombre de ménages par lieu de résidence.....	24
Figure 24 : répartition du milieu urbain et rural dans chaque région.....	26
Figure 25 : Pourcentage en colonne de la répartition de l'urbanisme dans chaque région	26
Figure 26 : Contribution des régions à la statistique du chi2	27
Figure 27 : Pauvreté extrême selon le lieu de résidence	27
Figure 28 : Pourcentage en colonne de la pauvreté extrême selon le lieu de résidence	28
Figure 29 : Pauvreté matérielle selon le lieu de résidence	29
Figure 30 : pourcentage en colonne de la pauvreté matérielle et lieu de résidence	29
Figure 31 : Pauvreté matérielle selon les régions	30
Figure 32 : Pourcentage en colonne de la pauvreté matérielle selon les différentes régions.....	30
Figure 33 Contribution des colonnes à la statistique du chi2	31
Figure 34 : Pauvreté extrême selon les régions.....	31
Figure 35 : pourcentage en colonne de la pauvreté extrême dans les différentes régions.....	31
Figure 36 : Coordonnées	33
Figure 37 : Contributions.....	33
Figure 38 : Qualités de représentation des modalités sur le 1er plan factoriel	34
Figure 39 : Inertie expliquée par les différents axes	35
Figure 40 : Inertie restitué par chaque axes.....	35
Figure 41 : ACM sur les modalités à partir du tableau de Burt	36
Figure 42 : ACM sur les modalités à partir du tableau disjonctif complet	37
Figure 43 : Carte des individus	37
Figure 44 : qualités de représentation sur le 1 ^{er} plan factoriel avec les deux variables de localisation	38
Figure 45 : Coordonnée des modalités avec les variables de localisation	39
Figure 46 : Contributions des modalités avec les variables de localisation	39
Figure 47 : ACM à partir du tableau de Burt avec les variables de localisation	40
Figure 48 : ACM avec les variables de localisation en supplémentaire.....	41
Figure 49 : Valeurs test des variables de localisation.....	41

Liste des sigles et abbréviations

USAID : United States Agency for International Development

UNFPA : United Nations Fund for Population Activities

DfID : Department for International Development

DANIDA : Danish International Development Agency

UNICEF : United Nations Children's Fund

KfW : German Development Bank

CHAI : Clinton Health Access Initiative

WFP : World Food Programme

MI : Micronutrient Initiative

I. Introduction

L'année 2020 a été particulière. L'ensemble de l'humanité a été frappé par une crise sanitaire de grande ampleur. Les conséquences de cette crise sont intéressantes à bien des égards, mais ce qui est particulièrement frappant est que cette pandémie a révélé les nombreuses inégalités qu'il y a entre les pays et dans les pays. C'est le cas de notre sujet d'étude : le Kenya qui connaît de nombreuses difficultés face à cette crise.

Pourtant le Kenya n'a pas une économie en détresse, bien au contraire. Depuis une décennie, la croissance de son PIB est en constante augmentation (variant autour de 5-6% par année) soutenue en grande partie par le développement des infrastructures urbaines. Malgré tout, la crise sanitaire de la COVID-19 a apporté des difficultés économiques et financières : d'après la dernière analyse économique de la Banque mondiale, le Kenya verra sa croissance économique se contracter de 1 à 1,5 % en 2020. De plus, la Commission Européenne ajoute que la baisse du PIB s'accompagne d'une dépréciation du shilling kenyan (KSh), d'une réduction des investissements intérieurs et d'une augmentation du déficit public de 17,2 milliards de KSh.

Mais les difficultés ne sont pas seulement d'ordre économique : toujours d'après la banque mondiale, la COVID-19 a considérablement réduit les progrès dans la réduction de la pauvreté au Kenya plongeant plus de deux millions de citoyens dans la pauvreté. La crise causée par la COVID-19 a révélé une fragilité du pays et les conditions de vie des citoyens ; malgré des actions de grandes ampleurs menées par l'Etat kenyan, la situation est encore précaire.

Les enquêtes et les analyses statistiques sont des outils essentiels pour diriger les actions publiques. Elles permettent de quantifier une situation mais aussi de révéler des structures et leur dynamique au sein d'un pays. Pour pouvoir lutter efficacement contre la pauvreté, il est incontournable d'avoir une idée précise de son ampleur et de sa répartition au sein du territoire. C'est toute la motivation de notre projet : mener une analyse et une réflexion complète sur la pauvreté au Kenya en se basant sur une enquête de grande ampleur.

Pour commencer, nous donnerons des informations générales sur le Kenya : sa situation géographique, son fonctionnement général administratif, son histoire moderne, son contexte démographique et socio-économique. Puis, nous présenterons en détail l'enquête nationale démographique et socio-économique réalisée en 2014 par le « Kenya National Bureau of Statistics » qui est le point de départ de notre travail. Dans une troisième partie, nous énoncerons la problématique qui motive notre projet. Nous expliquerons ensuite la méthodologie pour arriver aux analyses statistiques et leurs résultats. Enfin, nous finirons avec une synthèse qui mettra en avant les résultats que nous considérons comme les plus importants de notre travail.

II. Le Kenya

Présentation géographique

Le Kenya est un pays d'Afrique de l'Est bordant l'océan Indien. Il est entouré par la Tanzanie au sud-est, l'Ouganda à l'ouest, le sud du Soudan et l'Éthiopie au nord, et la Somalie à l'est. Du sud-est à l'ouest du territoire, l'altitude augmente progressivement à un peu plus de 1 000 m au-dessus du niveau de la mer avant de s'aplatir avec le plateau de Yatta. Le centre du pays est occupé par le mont Kenya : le plus haut sommet du Kenya et la seconde d'Afrique – le point culminant atteint les 5 199 mètres.

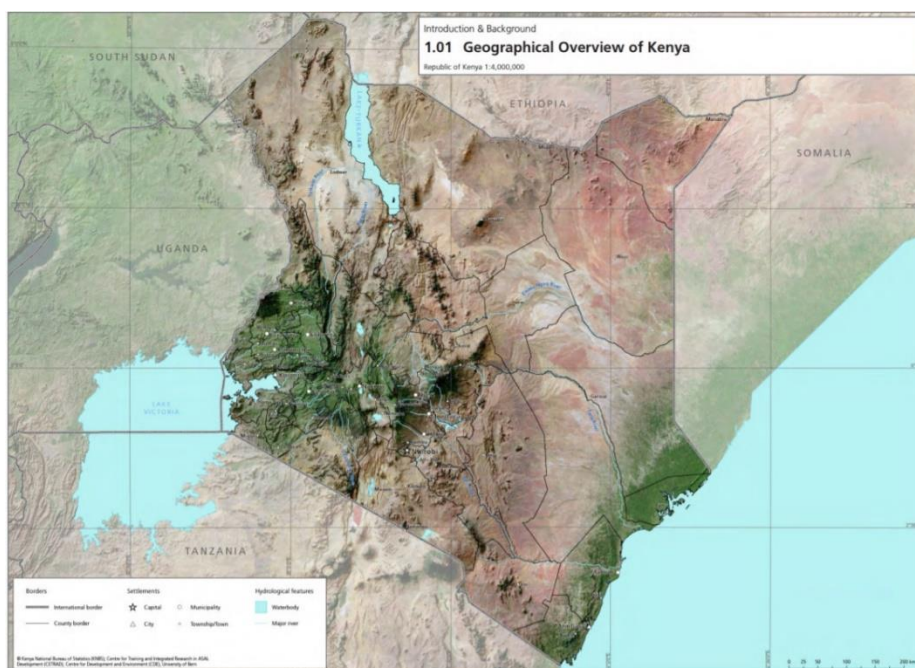


Figure 1 : Carte du Kenya

Situation politique

La capitale économique et politique du Kenya est Nairobi qui est aussi la ville la plus peuplée de ce pays. Le régime politique du pays est une république dont le chef de l'Etat a le titre de président.

Administrativement, le pays est découpé en 47 comtés ayant leur propre administration locale gérée par leur gouverneur.

Brève histoire du Kenya

Tout au long de son histoire, le territoire du Kenya a été une scène de conquête et de migration. La colonisation de ce territoire commença au début du VIII^{ème} siècle avec l'établissement de colonie arabe et perse. Elle se poursuit au XVI^{ème} siècle avec l'annexion de la côte maritime du Zanguebar (ancienne appellation d'une partie de la côte Est de l'Afrique) par des sultans d'Oman¹. A la fin du XIX^{ème} siècle, l'ensemble du territoire est sous le pouvoir du Royaume-Uni. Son histoire coloniale s'arrête officiellement le 12 décembre 1963 avec la fin de l'emprise coloniale anglaise.

¹ Pays situé au sud de la péninsule arabique, sur les bords du golfe d'Oman et de la mer d'Arabie

Diversité ethnique et culturelle

Le Kenya est aujourd'hui le résultat d'une histoire complexe et atypique, cela se traduit par une diversité culturelle, sociale et économique sur l'ensemble du territoire. Dans cette région africaine d'une superficie de 580 367 km² (comparable à l'Espagne avec ses 510 000 km²), plus de 68 langues sont parlées, dont deux reconnues : l'anglais et le kiswahili. De nombreuses religions sont pratiquées dont le christianisme (anglicanisme, protestantisme et catholicisme), l'islam (sunnites, chiites et ahmadis) ou encore certaines religions traditionnelles africaines. La population est aussi très hétérogène : il existe de nombreuses ethnies qui cohabitent.

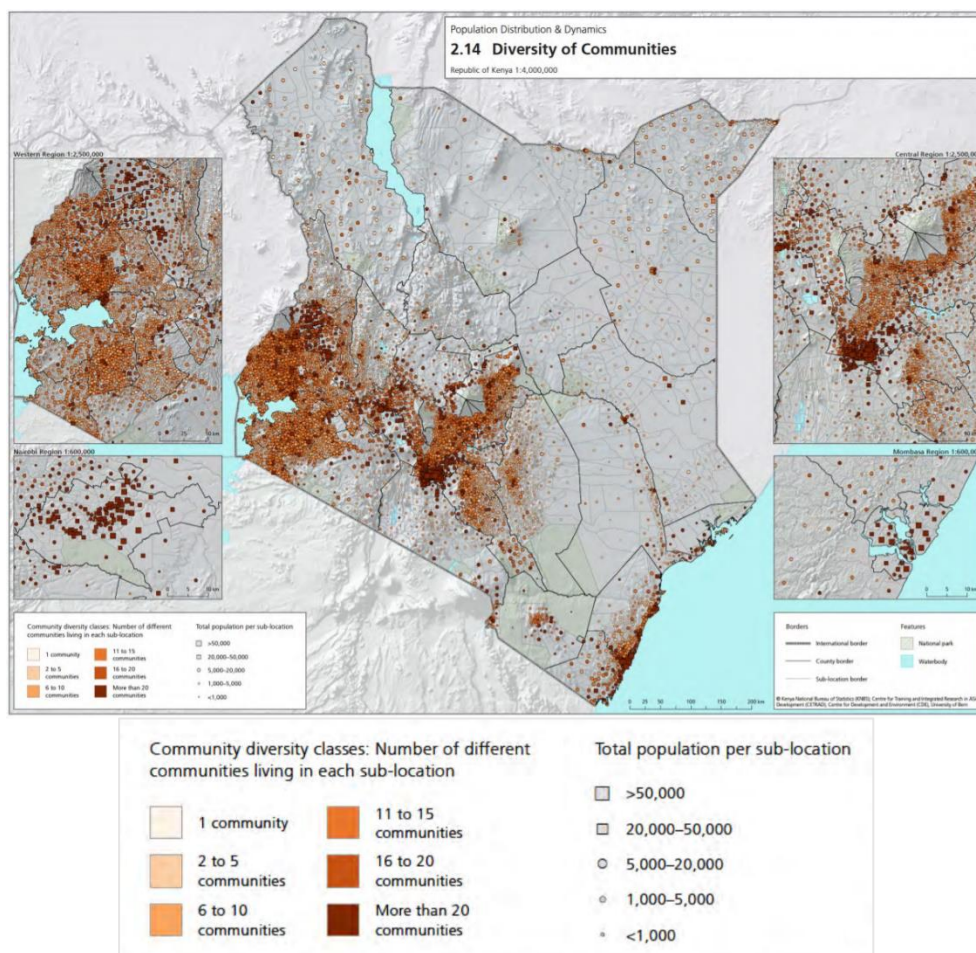


Figure 2 : Carte sur la diversité des communautés au Kenya

Contexte démographique et socio-économique

Depuis les années 1970, le pays connaît un essor économique et démographique : la population a presque été multipliée par 4 durant ces 50 dernières années, passant d'une population d'environ 12 millions en 1970 à 47 millions d'habitants en 2019, en parallèle, la croissance du PIB dépasse souvent les 2 points par année. Son développement économique repose sur différents secteurs comme l'agriculture, le tourisme ou encore l'exploitation des énergies fossiles. Néanmoins, malgré un pays ayant des terres riches et une population jeune, la situation socio-économique du pays est compliquée à bien des égards : la balance commerciale du pays est déficitaire et le PIB par habitant ne dépasse pas les 888 USD. Les pouvoirs publics du pays cherchent à mener des actions politiques de grandes envergures pour endiguer la situation précaire de la population. Pour éclairer ces actions politiques, des enquêtes auprès de la population sont alors menées.

III. Enquête démographique et socio-économique de 2014

L'enquête et son contexte

En 2014, une enquête nationale du nom de KDHS2014 a été réalisée entre mai et octobre par le Kenya National Bureau of Statistics (KNBS) en partenariat avec différentes organisations nationales telles que le National AIDS Control Council, le National Council for Population and Development ou encore le Kenya Medical Research Institute. Cette enquête a été financée d'une part par le Gouvernement du Kenya, mais également par d'autres partenaires à l'échelle mondiale. On retrouve des organismes comme USAID, UNFPA, DfID, DANIDA, UNICEF, KfW, CHAI, WFP, et MI. Cette enquête est dans la continuité d'une suite d'enquêtes effectuées tous les 5 ans au Kenya, dont la première date de 1989. Elle consiste à fournir des informations générales sur la population des différents comtés et de leurs conditions de vie à partir d'un échantillon de 40 300 ménages. Les objectifs spécifiques de l'enquête KDHS2014 sont les suivants :

- Estimer la fertilité et le taux de mortalité selon différentes tranches d'âge
- Mesurer les changements dans la fertilité et la prévalence de la contraception
- Examiner les indicateurs de base de la santé maternelle et infantile
- Collecter des données anthropométriques chez les enfants et les femmes
- Décrire les modèles que suivent la transmission du SIDA et autres maladie sexuellement transmissibles
- Déterminer l'ampleur et les tendances de la violence conjugale et des excisions génitales féminines

Plan de sondage

L'échantillon pour la KDHS 2014 a été tiré d'une base de sondage principale, le cinquième programme national d'enquête et d'évaluation par sondage (connu aussi sous le nom de NASSEP V). C'est un cadre que le KNBS exploite actuellement pour mener des enquêtes auprès des ménages à travers le Kenya. Le développement de la base de sondage a commencé en 2012 et contient un total de 5 360 grappes réparties en quatre sous-échantillons égaux. Ces grappes ont été dessinées avec une probabilité stratifiée proportionnelle à la méthodologie d'échantillonnage de taille à partir de 96 251 zones de dénombrement dans le recensement de la population et du logement au Kenya de 2009. La KDHS 2014 a utilisé deux sous-échantillons de la base NASSEP V qui ont été développés en 2013. Environ la moitié des clusters de ces deux sous-échantillons ont été mis à jour entre novembre 2013 et septembre 2014.

La KDHS 2014 a été conçu pour produire des estimations représentatives pour la plupart des indicateurs d'enquête au niveau national, pour les zones urbaines et rurales séparément, au niveau régional (anciennement provincial) et pour certains indicateurs au niveau du comté. Afin d'atteindre ces objectifs, l'échantillon a été conçu pour comprendre 40 300 ménages de 1 612 grappes réparties à travers le pays, avec 995 grappes en milieu rural et 617 en milieu urbain. Les échantillons ont été sélectionnés indépendamment dans chaque strate, en utilisant un plan d'échantillonnage à deux degrés. Dans la première étape, les 1 612 clusters ont été sélectionnés avec une probabilité égale à partir de la trame NASSEP V. Les ménages issus des opérations de listage ont servi de base de sondage

pour la deuxième étape de sélection, dans laquelle 25 ménages ont été sélectionnés dans chaque grappe. Les enquêteurs n'ont visité que les ménages présélectionnés et aucun remplacement des ménages n'a été autorisé pendant la collecte des données. En raison de l'attribution non proportionnelle aux strates d'échantillonnage et de la taille fixe de l'échantillon par grappe, l'enquête n'était pas auto-pondérée. Les données obtenues ont donc été pondérées pour être représentatives aux niveaux national, régional et des comtés.

Organisation du questionnaire

La KDHS2014 a utilisé 3 questionnaires ciblant 3 publics différents, afin de récolter des informations précises sur les foyers. On compte un questionnaire pour les ménages, un questionnaire pour les femmes âgées de 15 à 49 ans et un pour les hommes âgés de 15 à 54 ans. Néanmoins, le nombre conséquent de foyers interrogés a nécessité la création de questionnaires supplémentaires : un questionnaire court pour les femmes, et de même pour les ménages afin de privilégier le bon déroulement de l'enquête.

Au total, cinq questionnaires ont été utilisés dans la KDHS2014 : Un questionnaire long pour les hommes, 2 questionnaires (un court et un long) pour les femmes, et de même pour les ménages. L'échantillon interrogé a été divisé en deux parties. Dans la moitié des cas, les ménages ont reçu le questionnaire long pour le ménage, le questionnaire long pour la femme, et le questionnaire pour l'homme. Dans l'autre moitié, les ménages ont répondu au questionnaire du ménage et au questionnaire court pour la femme. Finalement, un ménage sur deux a été sélectionné pour répondre à l'intégralité des questionnaires et les ménages restants ont été sélectionnés pour les questionnaires courts.

Contenu des questionnaires :

Le questionnaire portant sur le ménage contient des questions pour déterminer le nombre de personnes résidant dans le ménage. Il permettait également d'identifier les femmes susceptibles de pouvoir répondre au questionnaire les concernant. Il contient également des questions afin de déterminer les conditions de vie du ménage, telles que des questions portant sur l'accès à l'eau potable ou le type de matériau des murs de l'habitation, etc.

Le questionnaire destiné aux femmes regroupe des questions portant sur les violences domestiques, certaines maladies, et sur leurs activités sexuelles, couvrant ainsi la période maternelle, les moyens de contraception ou encore les morts infantiles.

Enfin, le questionnaire destiné aux hommes est similaire à celui destiné aux femmes, mais omet toute la partie sur la maternité ainsi que d'autres questions uniquement destinées aux femmes.

IV. Problématique

Considérations économiques

Il est avéré que les conditions de vie au sein d'une population ont un impact non négligeable sur le développement économique et social d'un pays. La croissance économique est l'augmentation durable de la richesse produite ou l'accroissement de la production globale dans un territoire. Ce phénomène peut entraîner une augmentation des conditions de vie des ménages. Cette implication entre développement économique et amélioration des conditions de vie n'est pas forcément automatique : il est souvent nécessaire que l'administrateur du territoire mène un programme économique.

Mais l'inverse est aussi vrai : une bonne condition de vie est un facteur positif à la croissance économique.

Situation du Kenya et motivation du projet

Les pouvoirs publics ont bien conscience de cette réalité. En 2008, le président Mwai Kibaki lance un programme de développement du pays de 2008 à 2030. Son objectif est d'aider le Kenya à devenir un « **Pays nouvellement industrialisé à revenu intermédiaire offrant une qualité de vie élevée à tous ses citoyens d'ici 2030 dans un environnement propre et sûr** ». Ce programme se fonde sur 3 piliers (économique, social et politique) et s'organise sur un calendrier composé de cinq quinquennaux.

Problématique : conditions de vie des kenyans

L'objectif de notre projet est de produire une appréciation des conditions de vie des kenyans et d'analyser une corrélation entre les conditions et la situation géographique. Ainsi, notre étude s'articule autour de deux questions sur la pauvreté :

1. Quels sont les types de condition de vie des kenyans et leurs répartitions au sein de cette population ?
2. Existe-t-il un lien entre les conditions de vie d'un ménage et sa situation géographique au Kenya?

Comment définir la pauvreté ?

Dans la version de l'enquête KDHS2014, aucun indicateur mesure directement la pauvreté au sein d'un ménage. Néanmoins, la pauvreté est un phénomène multidimensionnel qui peut alors se définir de différentes façons : en termes monétaires (critère de revenu), en termes de conditions de vie (accès à un certain nombre de biens et services considérés comme essentiels) ou en termes subjectifs (sentiment d'être pauvre). L'enquête nous donne des indications sur des données qui influencent les conditions de vie d'un ménage, on décide alors de définir la pauvreté en termes de conditions de vie. Ce choix est expliqué dans la partie méthodologie.

Notion de la pauvreté

Nous rappelons que nous définissons la notion de pauvreté en termes de conditions de vie. En effet, l'indicateur de revenus d'une personne ou d'un ménage n'est malheureusement pas fourni dans notre base de données, nous allons donc nous focaliser plutôt sur les critères de possessions de biens ou de services.

Dans un premier temps, nous allons séparer la notion de pauvreté en deux catégories :

- Pauvreté extrême : le ménage n'a pas accès à des biens essentiels.
- Pauvreté matérielle : le ménage ne possède pas des biens de confort.

Pour qu'un ménage soit dans la catégorie pauvreté extrême, il doit répondre à l'un des trois critères suivants :

- Le ménage n'a pas accès à l'eau potable
- Le ménage n'a pas accès à l'électricité
- L'habitat du ménage est insalubre (état des murs et du toit)

Pour qu'un ménage soit dans la catégorie pauvreté matérielle, il doit répondre à l'un des quatre critères suivants :

- Ne pas avoir accès à des conditions sanitaires minimales (type de toilettes)
- Ne pas avoir un moyen pour s'informer (télévision et radio)
- Ne pas avoir de moyen de transport (moto et voiture)
- Ne pas avoir de cuisine ou réfrigérateur

Nous avons choisi des variables de qualité matérielle telle que la radio, mais rien ne nous empêchait d'en choisir une autre. Cependant après l'étude univariée, nous avons constaté que les variables : possession du téléphone, de la voiture, de la moto et du réfrigérateur ne permettront pas de bien déterminer la répartition des biens matériels car la plupart des ménages n'y ont pas accès. Nous avons donc décidé de ne pas les prendre en compte dans la création de la variable pauvreté matérielle.

Nous avons décidé d'enlever les observations pour lesquelles au moins une des réponses avait pour code 96 correspondant à une réponse « autre », que nous ne pouvons pas analyser car c'est une modalité qu'on ne peut pas classer selon nos critères, autrement dit, on ne peut dire que ce code désigne un bon matériel de mur ou le contraire dans l'exemple de la variable HV214 représentant le matériau des murs d'un ménage.

Ensuite, pour chacune des variables, nous avons analysé les réponses données à l'aide du fichier KEHR70FL. Si la réponse donnée par le foyer indique une forme de richesse, alors celle-ci sera recodée en 1, si la réponse indique une forme de pauvreté, elle sera recodée en 0.

Le choix de recodage est arbitraire. Prenons l'exemple de HV201, cette variable qui caractérise l'accès à l'eau potable au sein d'un foyer, comporte différentes modalités. Et en fonction du point de vue du lecteur, le choix du codage pourrait être remis en question.

Désormais, on peut regrouper nos variables selon 3 catégories :

1. Pauvreté matérielle :

Pauvreté matérielle	
Variable	Signification
HV209	Type de cuisine
HV212	Possession d'une voiture
HV221	Possession d'un téléphone
HV207	Possession d'un réfrigérateur

2. Pauvreté extrême :

Pauvreté extrême	
Variable	Signification
HV201	Type de source d'eau potable
HV205	Type de toilettes
HV206	Accès à l'électricité
HV207	Possession de la télévision
HV208	Possession d'une radio
HV211	Possession d'une moto
HV214	Matériau des murs
HV215	Matériau du toit

3. Localisation géographique :

Localisation	
Variable	Signification
HV024	Région
HV025	Lieu de résidence

Création de score

Après avoir recodé les modalités de nos variables, nous pouvons créer un score pour les deux variables : pauvreté extrême et matérielle, en supposant que toutes les variables associées à chaque pauvreté ont le même poids. En effet, pour la pauvreté extrême, on définit la variable ``p_extreme`` dont la valeur sera égale à la somme des valeurs prises par les quatre variables associées à la pauvreté extrême.

De même, ``p_materielle`` est définie comme étant la somme des valeurs prises par les quatre variables de pauvreté matérielle. ``p_extreme`` et ``p_materielle`` sont donc nos variables de score, prenant des valeurs entre 0 et 4. Un score proche de 0 indiquera une pauvreté importante, tandis qu'un score proche de 4 indiquera une richesse.

Plus précisément, les scores de la pauvreté extrême et matérielle sont définis dans le tableau ci-dessous :

Score de la pauvreté extrême	Signification	Score de la pauvreté matérielle	Signification
0	Extrêmement pauvre	0	Aucun bien
1	Pauvre	1	1 bien essentiel
2	Moyen	2	2 biens essentiels
3	aisé	3	3 biens essentiels
4	Riche	4	Tous les biens essentiels

Bien qu'il existe une variable concernant l'indice de pauvreté de la variable HV270, faisant appel aux variables concernant l'habitat, les biens d'équipement, et divers services, nous avons décidé de ne pas tenir compte de cette variable, car nous ne savons pas sur quels indicateurs celle-ci se basait, et nous ne pouvions donc pas interpréter correctement nos résultats. C'est pour cela que nous avons envisagé de créer nos deux indices grâce à la création des deux variables pauvreté extrême et pauvreté matérielle.

VI. Analyses statistiques

A. Etude univariée (Tris à plat)

1. Représentation graphique

On peut classer les variables de la table *menage* selon trois catégories :

- Variables ordinales : HV206, HV207, HV208, HV209, HV210, HV211, HV212, HV221, HV225, HV227 et HV228.
- Variables nominales : HHID, HV001, HV024, HV025, HV026, HV201, HV202, HV205, HV213, HV214, HV215, HV217, HV219 et HV226.
- Variables quantitatives : HV005, HV009, HV014, HV204, HV216, HV218 et HV220.

A part l'identifiant, elles sont toutes de type numérique dans la table *menage*. Les variables qualitatives (nominales et ordinales) doivent être passées en « *factor* » ou « *character* » pour certains traitements, c'est pour cela qu'on change le type de numérique « *factor* » (on crée de nouvelles variables en rajoutant le préfixe « c »).

Ensuite, nous calculons les effectifs et les fréquences des variables qualitatives choisies (données en sortie Tris à plat) et nous affichons les histogrammes ou les diagrammes circulaires de nos variables afin de mieux visualiser nos données.

a. Pauvreté extrême

Histogramme de la variable HV201 (source d'eau potable)

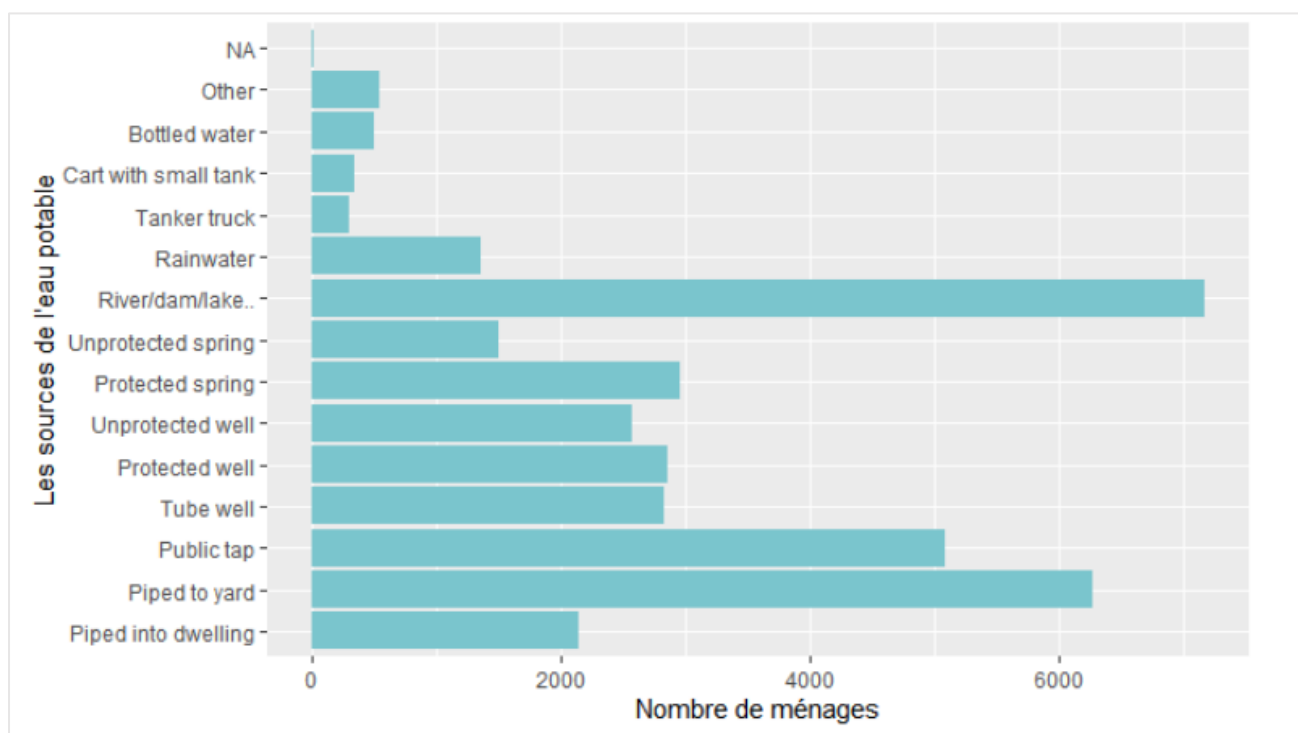


Figure 3 : Type de source d'eau potable par ménage

Concernant l'accès à l'eau potable, on peut séparer la population kenyane en 4 catégories. On remarque qu'environ 35% des kenyans ont accès à l'eau potable via des canalisations reliées soit directement à leur domicile, soit à une fontaine publique, destinée à plusieurs foyers. 20% des kenyans utilisent un puit afin d'accéder à l'eau potable. Un tiers utilise l'eau que l'on peut trouver naturellement, soit en provenance de sources protégées ou non soit via des lacs, des rivières, des étangs et autres sources de ce type. Le reste de la population utilise d'autres moyens d'accès à l'eau potable, tels que l'eau en bouteille ou l'eau de pluie. On remarque une hétérogénéité au niveau des réponses des foyers, mais on constate également que beaucoup de foyers n'ont pas accès à l'eau potable tels qu'on pourrait l'avoir au sein d'un foyer Français.

	Effectif	Effectif cumulé	Fréquence	Fréquence cumulée
Piped.into.dwelling	2144	2144	5.88525940	5.885259
Piped.to.yard	6271	8415	17.21383475	23.099094
Public.tap	5081	13496	13.94729618	37.046390
Tube.well	2829	16325	7.76557782	44.811968
Protected.well	2851	19176	7.82596761	52.637936
Unprotected.well	2572	21748	7.06011529	59.698051
Protected.spring	2953	24701	8.10595663	67.804008
Unprotected.spring	1496	26197	4.10650563	71.910513
River.dam.lake..	7176	33373	19.69805106	91.608564
Rainwater	1361	34734	3.73593192	95.344496
Tanker.truck	298	35032	0.81800714	96.162503
Cart.with.small.tank	348	35380	0.95525666	97.117760
Bottled.water	505	35885	1.38622015	98.503980
Other	538	36423	1.47680483	99.980785
NA.	7	36430	0.01921493	100.000000

Figure 4 : Effectifs et fréquences de l'accès à l'eau potable

Histogramme de la variable HV206 (accès à l'électricité)

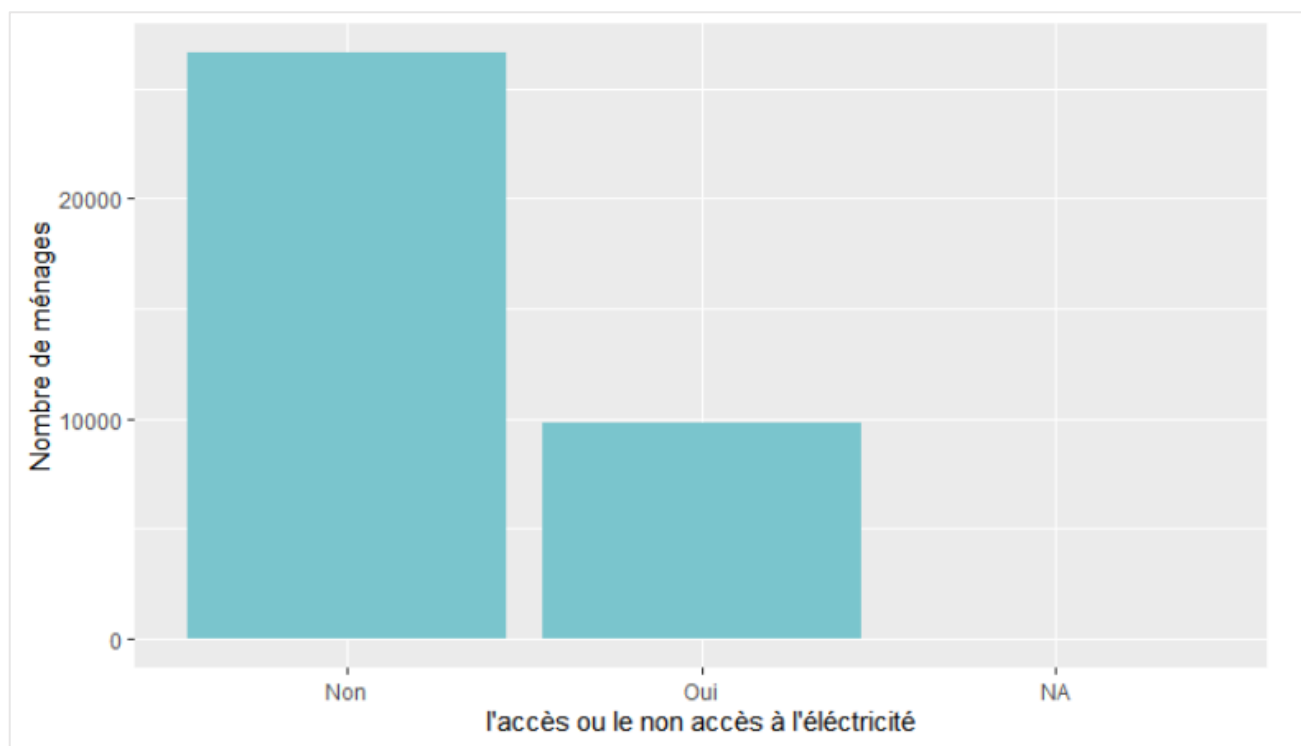


Figure 5 : Accès à l'électricité

On remarque qu'environ 73% de la population n'a pas accès à l'électricité tandis qu'uniquement 27% y ont accès comme on le voit à l'aide de l'histogramme ci-dessus et dans la colonne fréquence de ce tableau. L'électricité semble être un bien assez rare au sein des foyers Kenyans, et l'on pourra s'appuyer sur cette variable et l'utiliser en tant qu'indicateur fondamental de la pauvreté au Kenya par la suite.

	Effectif	Effectif cumulé	Fréquence	Fréquence cumulée
Non	26606	26606	73.0332144	73.03321
Oui	9803	36409	26.9091408	99.94236
NA.	21	36430	0.0576448	100.00000

Figure 6 : Effectifs et fréquences de l'accès à l'électricité

Histogramme des conditions matérielles des murs et du toit d'un ménage (HV214/HV215)

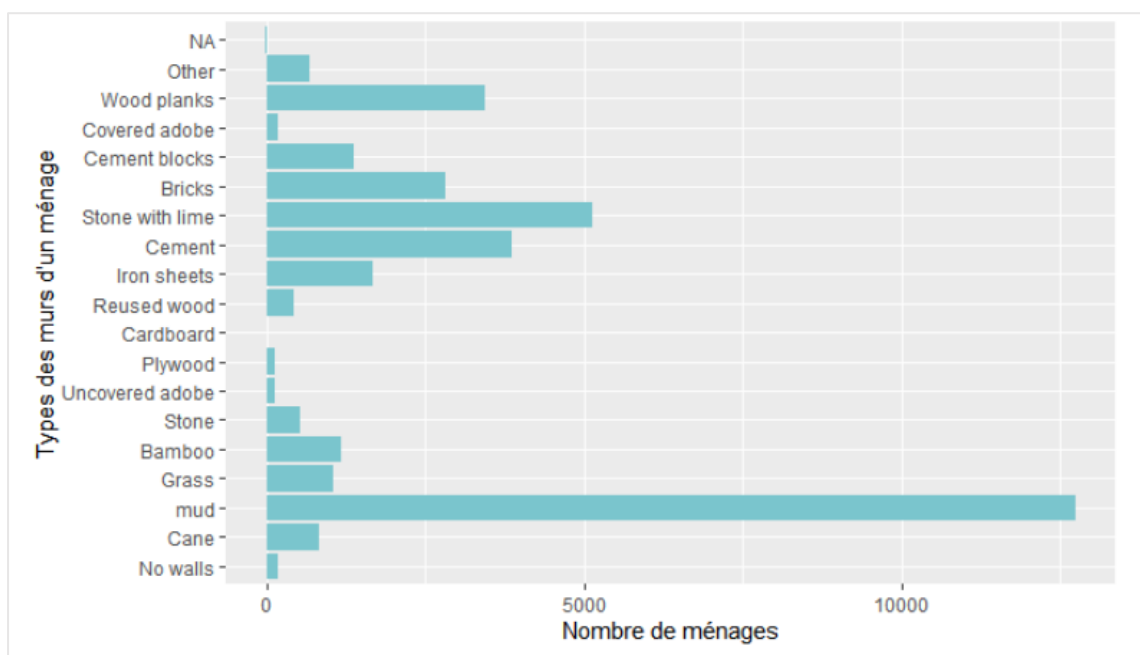


Figure 7 : Conditions matérielles des murs

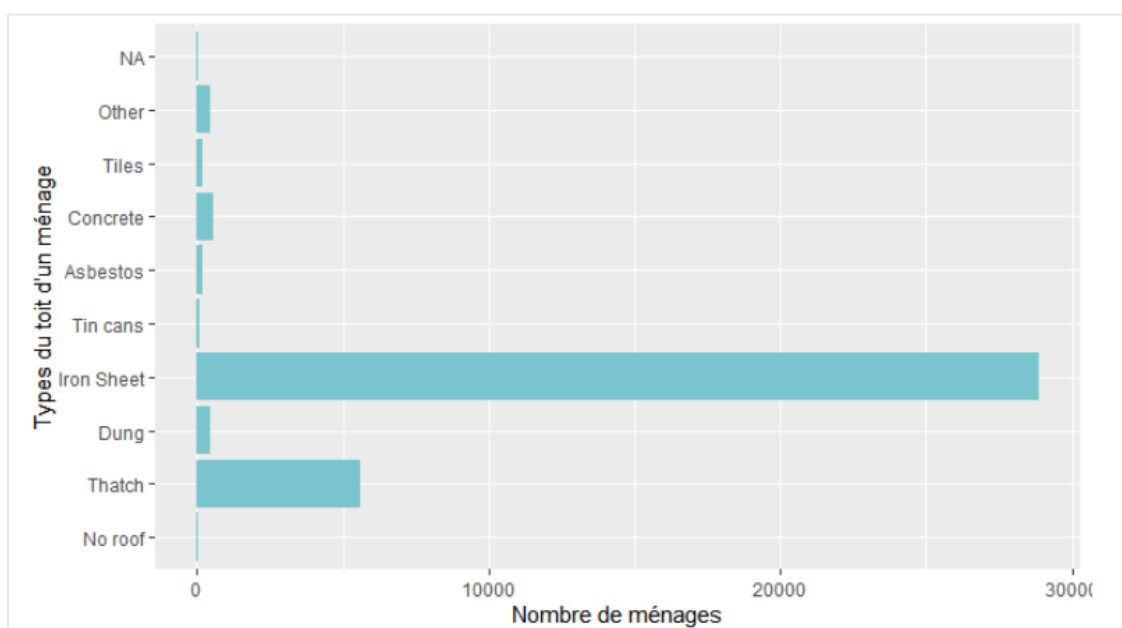


Figure 8 : Conditions matérielles du toit

D'après la figure 7, on remarque une haute proportion de foyers (plus du tiers) a des habitations dont les murs sont faits en boue et en gazon, ce qui ne constitue pas une habitation saine et sécurisée. Un autre tiers des foyers interrogés quant à eux possèdent une habitation dont les murs sont faits en pierres, briques et ciment. Dans le dernier tiers des foyers interrogés on retrouve des foyers dont les habitations ont des murs en feuille de fer, en bois, en carton, mais également des foyers n'ayant pas de murs dans leur habitation. Les résultats sont encore ici très hétérogènes.

D'après l'histogramme de la figure 8, il y a environ 80% des foyers interrogés utilisant un toit composé de feuilles de fer, tandis que 15% utilisent des toits en chaume ou en herbe, nettement moins robuste et moins isolant. Les 5 derniers pourcents sont des modalités très faiblement représentées.

b. Pauvreté matérielle

Histogramme de la variable HV205 (Types de toilettes dans un ménage)

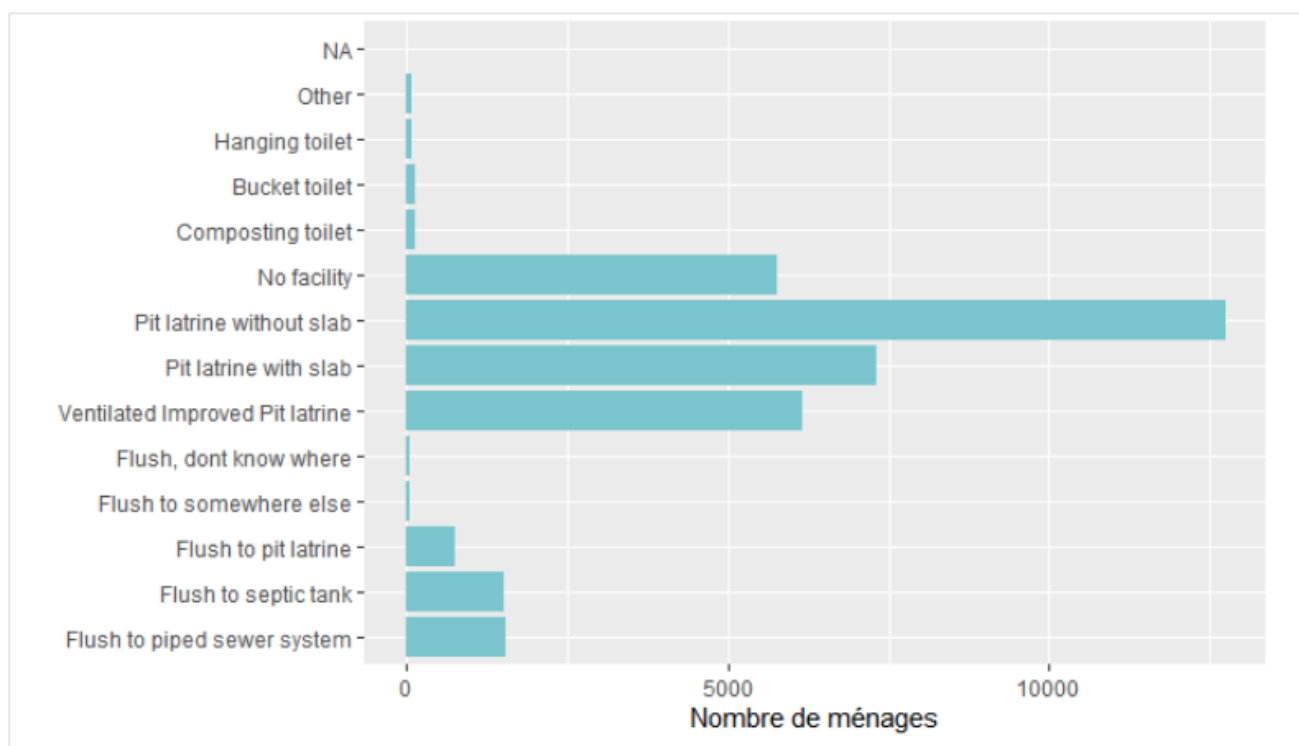


Figure 9 : Types de toilettes

L'histogramme ci-dessus correspond au nombre de ménages selon le type de toilettes. Il montre qu'il y a 4 catégories remarquables. En effet, il y a environ 6% des kenyans utilisant soit un système de latrines relié aux égouts ou à la fosse septique ou des latrines à chasse d'eau. Environ 16% de la population qui n'a pas d'installation de toilettes dans un ménage. Il y a 70% des kenyans ayant des toilettes qui ne sont bien équipées et concernant les autres types de toilettes, le pourcentage des ménages qui les possèdent est quasiment nul.

Histogramme de la variable HV208 (possession de la télévision)

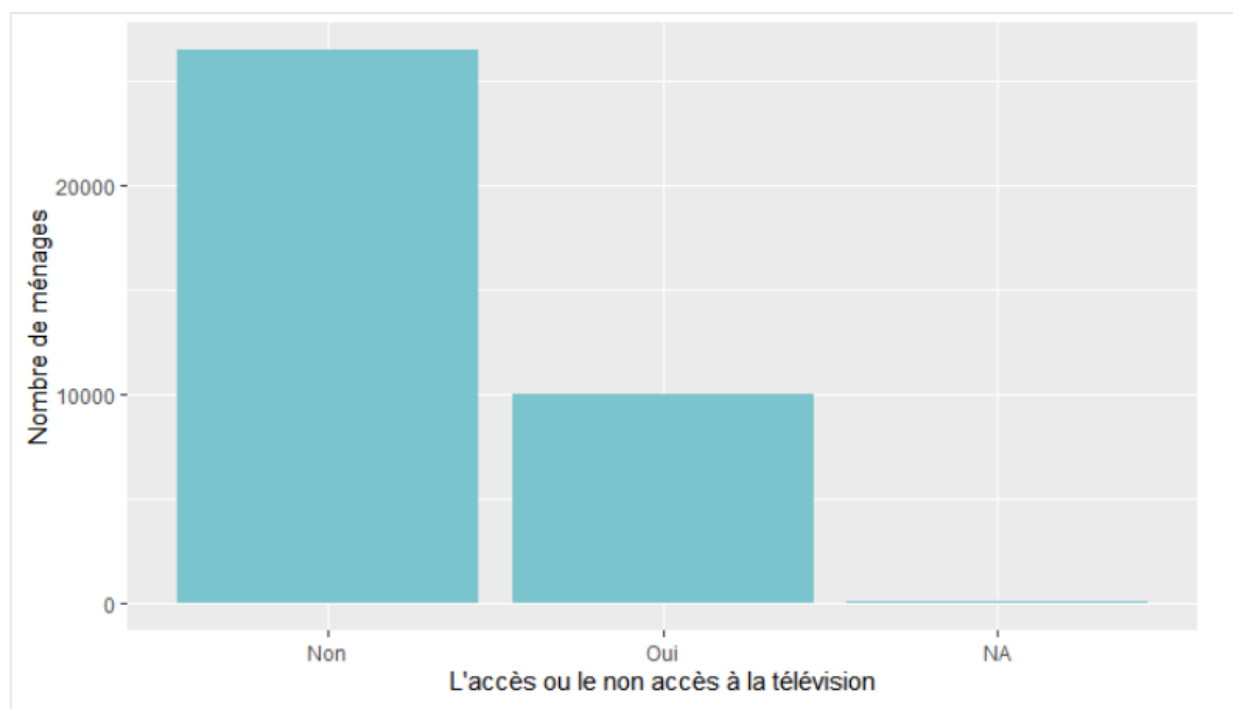


Figure 10 : Possession de la télévision

En se basant sur la figure 10 et le tableau des fréquences (figure 11), on en déduit qu'environ 73% de la population interrogée n'ont pas accès à la télévision tandis qu'uniquement 27% y ont accès. La télévision ne semble pas être un bien très répandu au Kenya.

	Effectif	Effectif cumulé	Fréquence	Fréquence cumulée
Non	26435	26435	72.56382103	72.56382
Oui	9963	36398	27.34833928	99.91216
NA.	32	36430	0.08783969	100.00000

Figure 11 Effectifs et fréquences de l'accès à la télévision

Histogramme de la variable HV207 (possession de la radio)

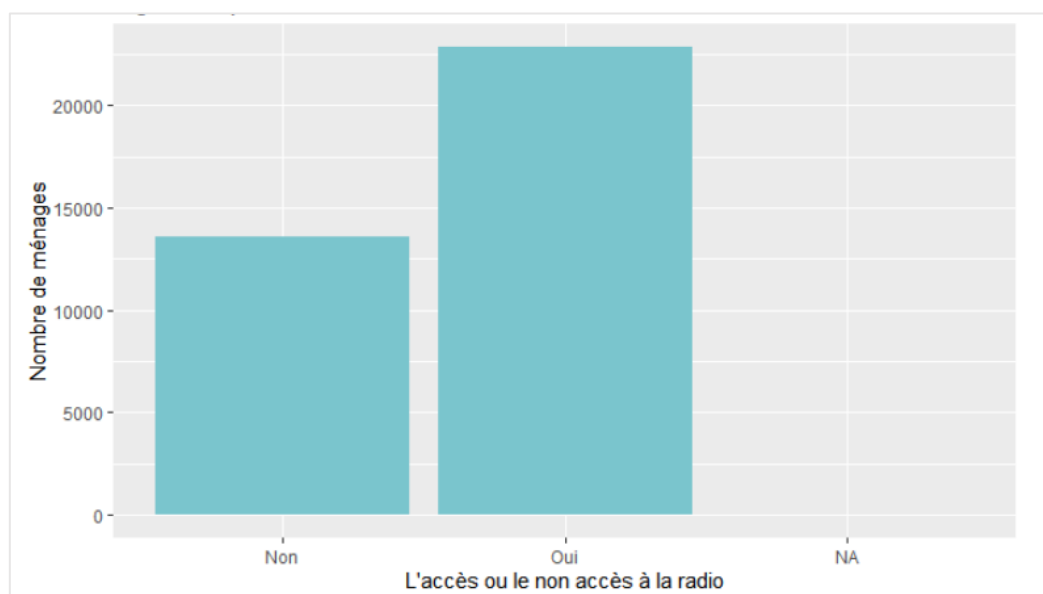


Figure 12 : Accès à la radio

Concernant la radio, il y a environ 63% des kenyans la possédant et 37% qui n'ont en pas. Ce qui signifie qu'au Kenya, les ménages ont plus facilement accès à la radio qu'à la télévision. Malgré une faible proportion de télévision dû au manque d'électricité, les foyers ont tout de même un accès à l'information.

	Effectif	Effectif cumulé	Fréquence	Fréquence cumulée
Non	13568	13568	37.24402965	37.24403
Oui	22848	36416	62.71754049	99.96157
NA.	14	36430	0.03842987	100.00000

Figure 13 : Effectifs et fréquences de la possession de la radio

Histogramme de la variable HV226 (Types de cuisine dans un ménage)

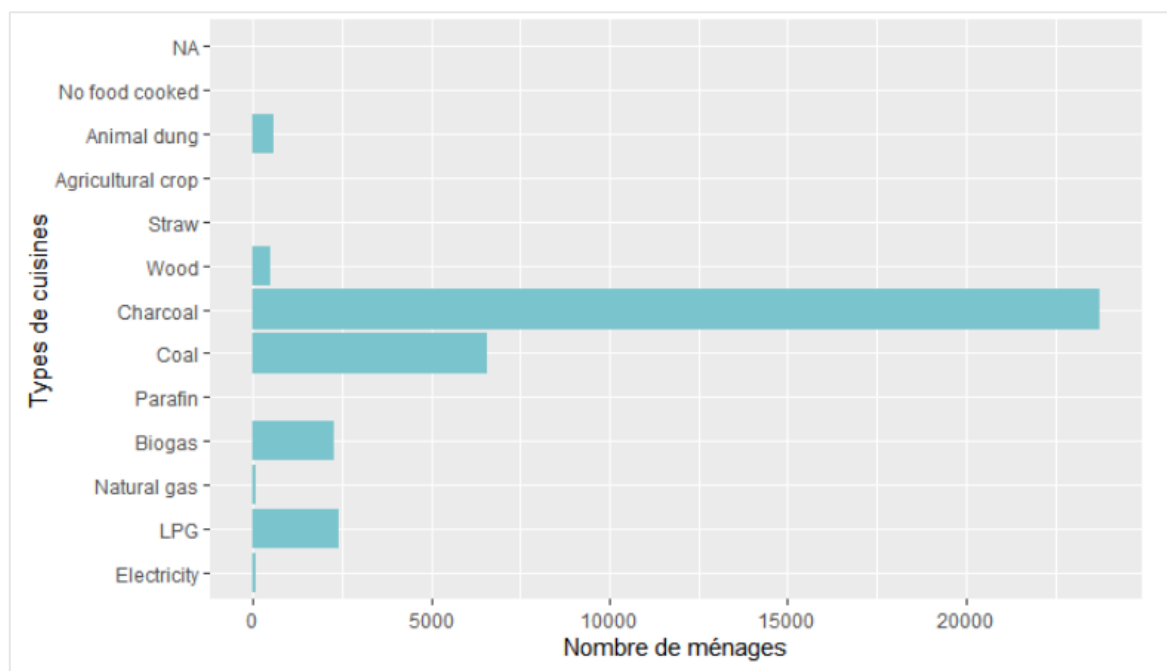


Figure 14 : Type de cuisine

L'histogramme ci-dessus représente le nombre de ménage selon les types de cuisine. On remarque qu'il y a environ 83% de la population utilisant soit le Charbon ou le bois. Environ 13% des kenyans qui choisissent de cuisiner à l'aide du GPL ou du paraffine/ Kérosène.

Remarque :

Concernant la possession du téléphone, de la voiture, de la moto et du réfrigérateur, on voit bien que la majorité des ménages ne possèdent pas ces biens d'après les diagrammes circulaires ci-dessous, ce qui signifie que ces 4 variables ne peuvent pas être pertinentes pour notre étude car elles ne représentent pas des biens nécessaires ou indispensables.

En effet, d'après le tableau de la figure 15 représentant les effectifs et fréquences de la possession de la voiture, on voit qu'il y a uniquement 4 % des ménages interrogés qui possèdent ce bien, ce qui confirme le résultat de la figure 18 ci-dessous. On retrouve également la même chose pour les trois autres variables.

	Effectif	Effectif cumulé	Fréquence	Fréquence cumulée
Non	34959	34959	96.02538	96.02538
Oui	1447	36406	3.97462	100.00000

Figure 15 : Effectifs et fréquences de la possession de voiture

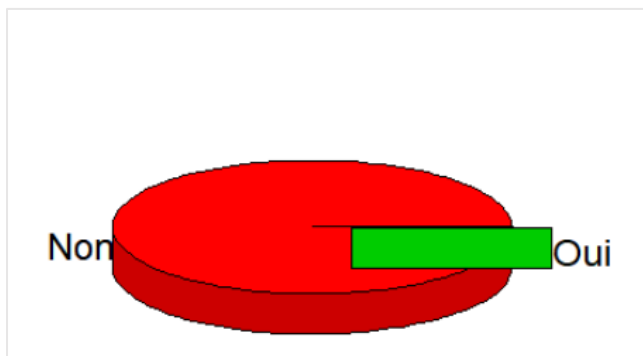


Figure 19 : Possession d'un téléphone

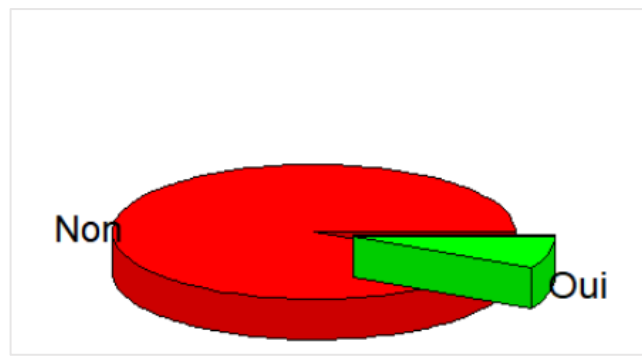


Figure 18 : Possession de la moto

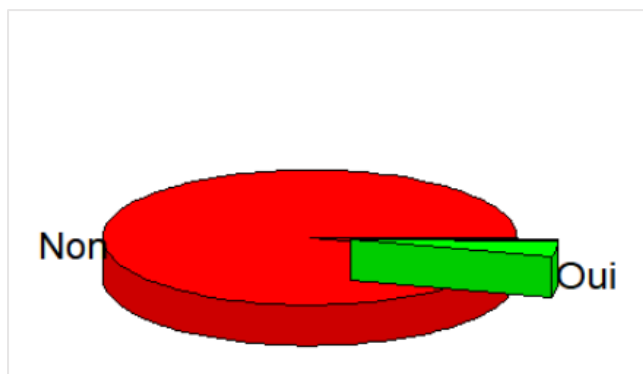


Figure 17 : Possession de la voiture

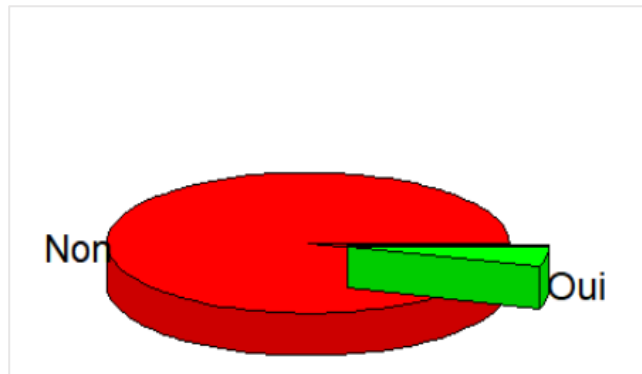


Figure 16 : Possession d'un réfrigérateur

c. Répartition géographique des ménages :

Après avoir analysé les variables qui nous intéressent, nous souhaitons visualiser la répartition des régions au Kenya ainsi que celle de l'urbanisme. Cela va nous aider à former quelques hypothèses concernant la répartition de la pauvreté au sein du Kenya. On s'intéresse donc à l'étude des variables Région (HV024) et lieu de résidence (HV025).

	Effectif	Effectif cumulé	Fréquence	Fréquence cumulée
Coast	4476	4476	12.286577	12.28658
North Eastern	1857	6333	5.097447	17.38402
Eastern	6261	12594	17.186385	34.57041
Central	4041	16635	11.092506	45.66292
Rift Valley	10534	27169	28.915729	74.57864
Western	3220	30389	8.838869	83.41751
Nyanza	4801	35190	13.178699	96.59621
Nairobi	1240	36430	3.403788	100.00000

Figure 20 : Effectifs et fréquences de la variable région

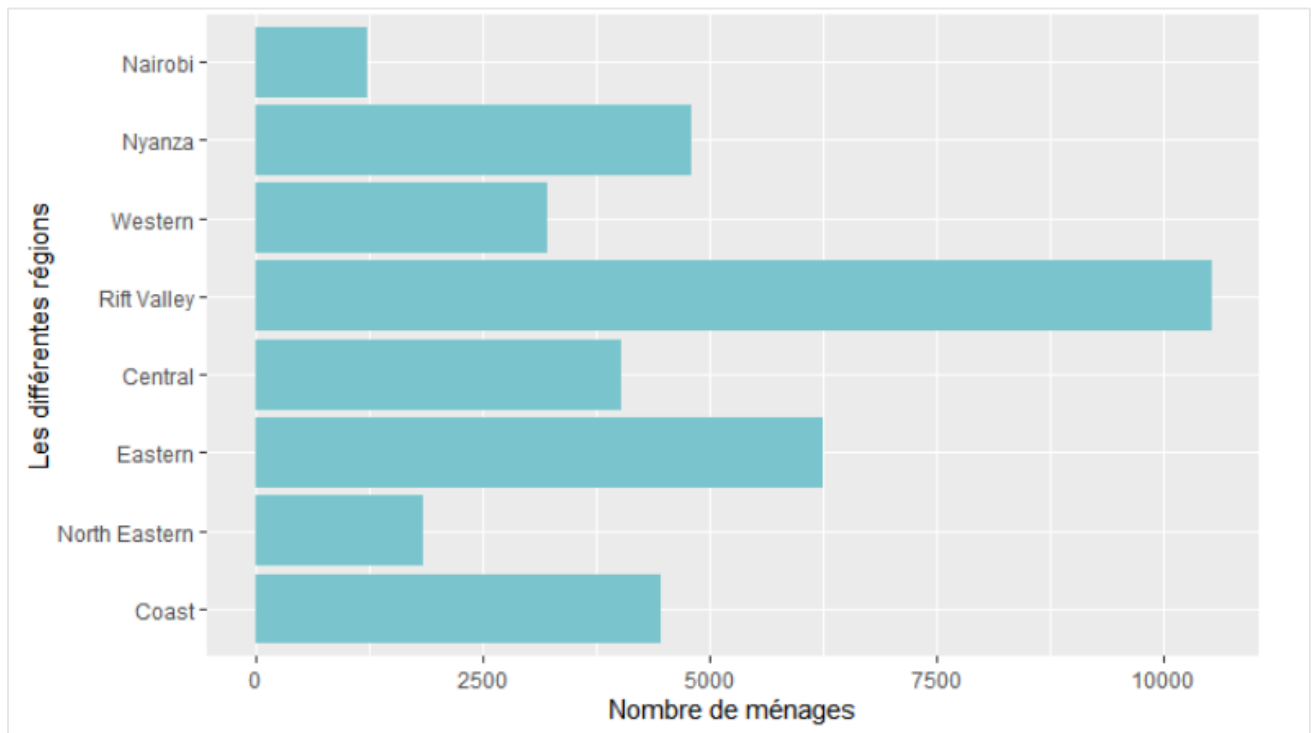


Figure 21 : Nombre de ménage par région

En faisant une analyse similaire de la figure (21) correspondant à la répartition des ménages selon les différentes régions et du tableau des effectifs et fréquences de la figure (20), on remarque qu'environ 29% de la population habitent dans la région Rift Valley. Environ 42% des kenyans se situent soit dans la région Eastern, Nyanza ou Coast. En revanche, il n'y a que 3% de population qui sont dans la capitale Nairobi.

Concernant la répartition des ménages interrogés entre le milieu urbain et rural, on voit qu'environ 62 % de la population kenyane se situe dans le milieu rural et uniquement 38% dans le milieu urbain comme le montrent les figures (22) et (23).

	Effectif	Effectif cumulé	Fréquence	Fréquence cumulée
Urban	13914	13914	38.1938	38.1938
Rural	22516	36430	61.8062	100.0000

Figure 22 : Effectifs et fréquences de la variable lieu de résidence

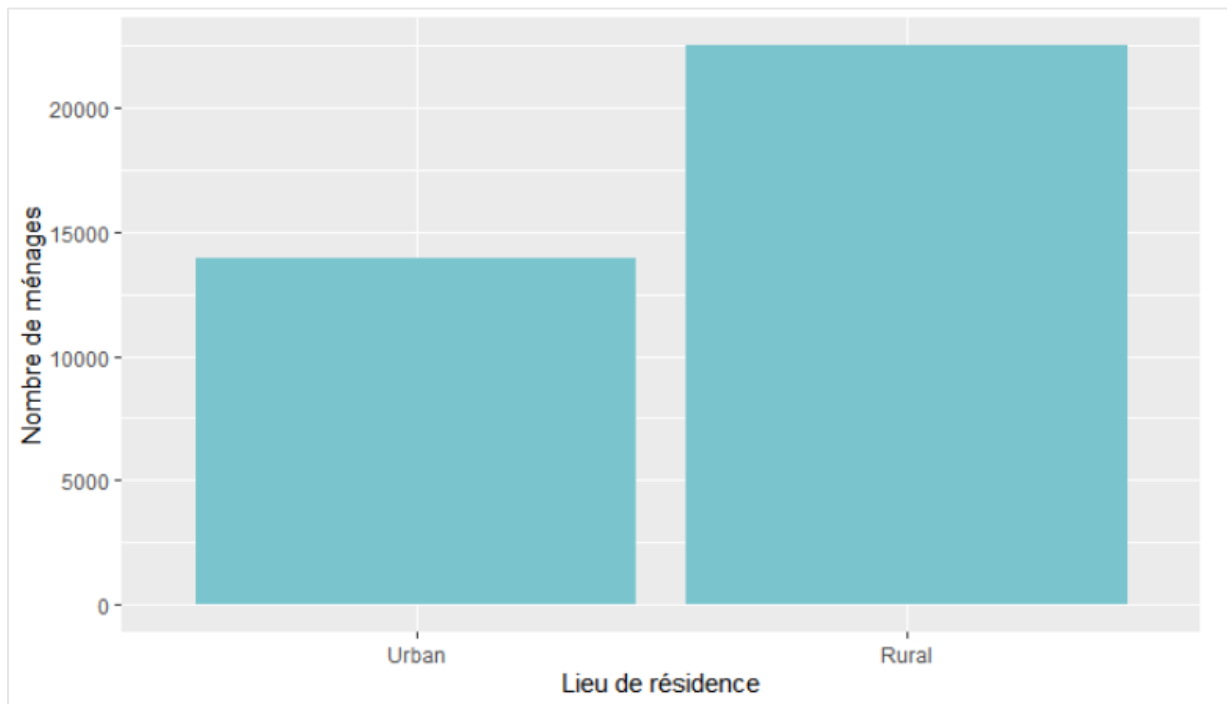


Figure 23 : Nombre de ménages par lieu de résidence

Conclusion

D'après cette première étude univarié, on peut remarquer que beaucoup de ménages répondent aux critères de pauvreté extrême et matérielle. Les biens de type moto ou vélo peuvent être considérés comme luxueux au vu de la faible proportion de foyers qui en possèdent. Pour la suite, on décide d'analyser en profondeur les variables qualifiant la pauvreté extrême et matérielle, ainsi que les autres variables région et le lieu de résidence afin d'expliquer la répartition de la pauvreté au sein du Kenya.

Concernant les données manquantes, on voit qu'elles sont sous représentées lors de cette première étude. En revanche, on les prendra en compte pour avoir des résultats plus en moins précis car elles signifient que la question a été posée et nous n'avons pas la réponse.

B. Etude bivariable (Tris croisés)

1. Création des variables pauvreté extrême et pauvreté matérielle

Lors de cette étape, nous voulons étudier des couples de variables pertinentes permettant de répondre à notre problématique. Nous avons évoqué précédemment les termes de pauvreté matérielle et extrême. Nous allons maintenant définir une méthode afin de déterminer si un foyer appartient ou pas à chacune de ces catégories.

A partir de la table ménage, nous avons créé une table nommée "*tab_vb_int*" contenant les variables que nous avons jugées importantes pour cette partie. On rappelle qu'il s'agit des variables de localisation HV024 et HV025, permettant de répartir la pauvreté des foyers d'une manière géographique, ainsi que les différentes variables de pauvreté : HV201, HV205, HV206, HV207, HV208, HV214, HV215, HV226.

Nous rappelons qu'on regroupe les variables HV201, HV206, HV214 et HV215 pour le critère de pauvreté extrême d'un ménage. Concernant la pauvreté matérielle, on la caractérise par les variables HV205, HV207, HV208 et HV226.

Il est également important de recoder les modalités des variables dans le cas où elles sont supérieures à deux comme les variables HV201, HV205, HV214, HV215 et HV226.

✓ Recodage

Prenons l'exemple de la variable HV214 correspondant au matériel des murs de l'habitat. Elle contient les modalités suivantes :

11	No walls	24	Plywood	33	Bricks
12	Cane	25	Cardboard	34	Cement blocks
13	mud	26	Reused wood	35	Covered adobe
14	Grass	27	Iron sheets	36	Wood planks
21	Bamboo	31	Cement	96	Other
22	Stone	32	Stone with lime	99	missing
23	Uncovered adobe	25	Cardboard		

Toutes les réponses dont le code était au-dessus de 20, correspondant à un bon matériau des murs, ont été recodées en 1, et 0 sinon. Les réponses manquantes, codées naturellement (NA) restent inchangées. Le détail de recodage de toutes les variables se trouvent en annexe 1.

Nous rappelons qu'après le recodage des modalités des variables et la création des deux variables pauvreté matérielle et extrême, nous retrouvons alors avec 4 variables, p_extreme, p_materielle, HV024 (région) et HV025 (lieu de résidence). Nous avons donc décidé de croiser les 5 couples de variables suivants : (HV024 et HV025), (p_extreme et HV025), (p_materielle et HV025), (p_materielle

et HV024), (p_{extreme} et HV024) afin d'effectuer un test de χ^2 et avoir une idée plus précise à propos de la dépendance des deux variables dans chaque couple.

2. Croisement de couple de variables pertinentes

Nous décidons dans cette partie d'interpréter les profils colonnes car il nous semble qu'ils sont plus intéressants que les profils lignes. En effet, dans l'exemple du croisement entre la variable région et lieu de résidence, c'est mieux d'analyser la répartition de l'urbanisme dans chaque région que de regarder la répartition des régions dans le milieu urbain et rural.

a. Croisement entre les variables région et lieu de résidence

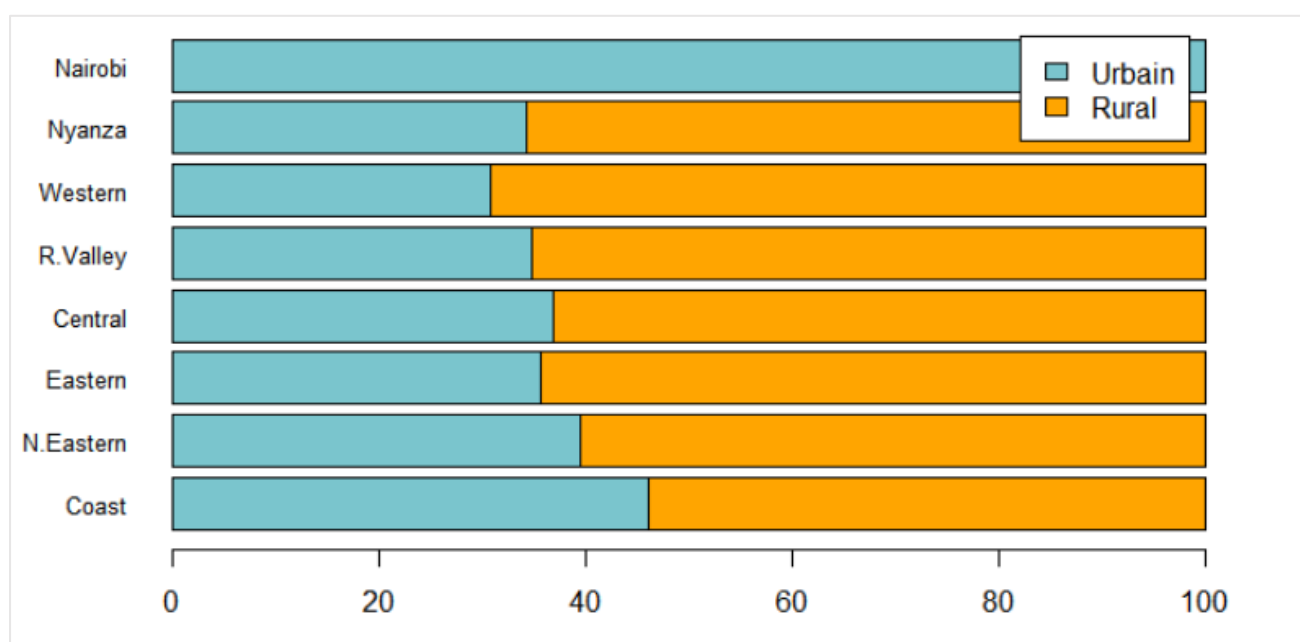


Figure 24 : répartition du milieu urbain et rural dans chaque région

	Coast	N.Eastern	Eastern	Central	R.Valley	Western	Nyanza	Nairobi
Urban	46.08071	39.50000	35.62632	36.83013	34.84575	30.79105	34.19570	100.00000
Rural	53.91929	60.50000	64.37368	63.16987	65.15425	69.20895	65.80430	0.00000

Figure 25 : Pourcentage en colonne de la répartition de l'urbanisme dans chaque région

D'après le graphique de la figure 24 et le tableau de la figure 25 correspondant au pourcentage des ménages interrogés entre le milieu urbain et rural dans les régions au Kenya, on en déduit qu'environ 46% de ces ménages de la région *Coast* sont localisés dans le milieu urbain et moins de 40% pour les régions *N.Eastern*, *Eastern*, *Central*, *R.Valley*, *Western* et *Nyanza*. En revanche, la capitale *Nairobi* est caractérisée uniquement par le milieu urbain.

Nous voulons maintenant tester les deux hypothèses du test (H_0) et (H_1) :

(H_0) : les variables région et lieu de résidence sont indépendantes

(H_1) : il existe un lien entre les région et lieu de résidence

Soit d_0^2 la valeur de X-squared dans le tableau suivant, cette valeur peut être considérée comme la réalisation d'une variable aléatoire D^2 qui suit approximativement une loi de χ^2 à 7 degré de liberté. En effet, ici les conditions d'application de ce test sont vérifiées car l'effectif théorique est supérieur ou égal à 5% dans toutes les cases du croisement des modalités (résultat donné en sortie Tris croisés).

Croisement entre région et lieu de résidence		
X-squared	df	P-value
2180.4	7	< 2.2 e-16

On définit la p-valeur = $IP_{(H_0)}(\text{rejeter } (H_0)) = IP_{(H_0)}(D^2 \geq d_0^2) = IP(X^2(7) \geq 2180.4)$ avec X^2 suit une loi de χ^2 de ddl = 7.

Règle de décision : On choisit un risque de première espèce $\alpha = 5\%$. Si la p-valeur $\leq \alpha$ alors on rejette (H_0) au risque de 5% de se tromper.

D'après le résultat de ce test, la p-valeur est inférieure à 2.2 e-16, c'est une valeur très proche de 0, ce qui signifie que c'est largement inférieur à 5%. Dans ce cas, on peut rejeter l'hypothèse (H_0) et affirmer qu'il existe bien un lien significatif entre les variables région et lieu de résidence.

On en conclut ainsi que c'est pertinent de faire un croisement entre les variables (pauvreté extrême et matérielles) à la fois avec la variable région et la variable lieu de résidence. Cela nous permettra d'analyser au mieux la répartition de la pauvreté selon les milieux urbain et rural dans les différentes régions au Kenya.

Concernant la contribution des colonnes à la statistique de χ^2 (figure 11. 2), on peut remarquer que Nairobi est la modalité contribuant le plus à la statistique du χ^2 . En effet, elle se caractérise par un taux d'urbanisation égal à 100% ce qui la rend atypique.

Coast	N. Eastern	Eastern	Central	R. Valley	western	Nyanza	Nairobi
102.779	0.663	20.730	5.030	58.318	80.954	38.636	1873.329

Figure 26 : Contribution des régions à la statistique du χ^2

Figure 11. 2 Contribution des colonnes à la statistique du χ^2

b. Croisement entre les variables pauvreté extrême et lieu de résidence

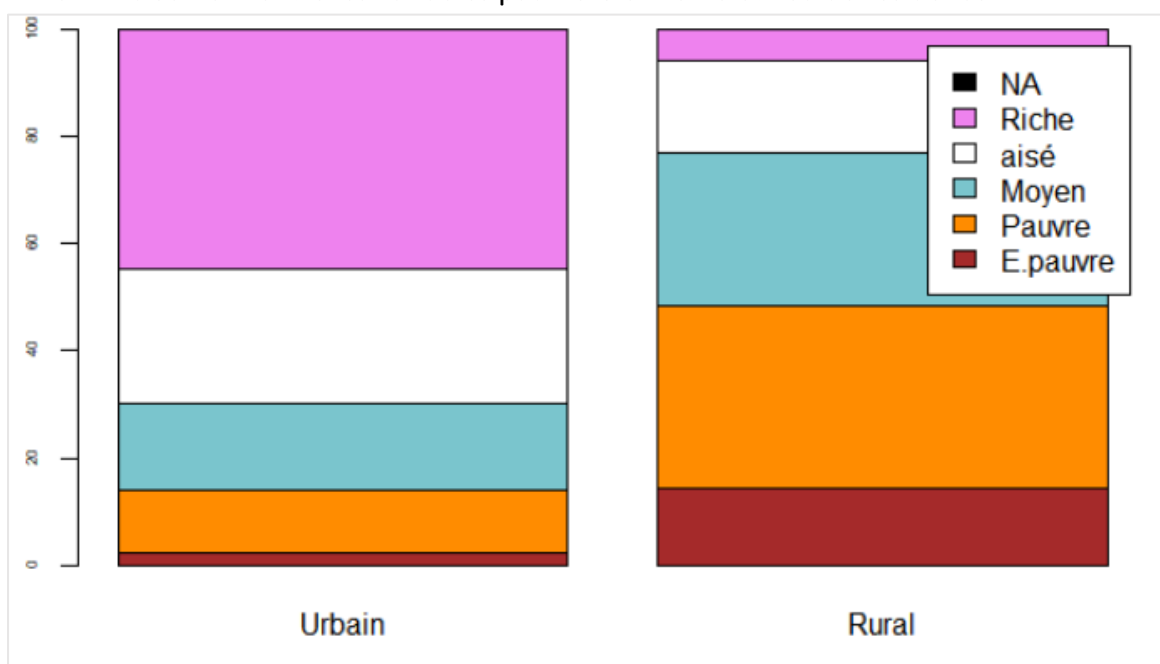


Figure 27 : Pauvreté extrême selon le lieu de résidence

	Urban	Rural
Extrêmement pauvre	2.3046614	14.2530452
Pauvre	11.5233068	34.2745135
Moyen	16.1698015	28.3660802
aisé	25.2546279	17.2399309
Riches	44.6212178	5.7077519
MissingData	0.1263847	0.1586783

Figure 28 : Pourcentage en colonne de la pauvreté extrême selon le lieu de résidence

On cherche à identifier quel est le milieu de vie pour lequel la pauvreté extrême est la plus importante. A l'aide des figures 27 et 28, on s'aperçoit qu'environ 45% des foyers interrogés vivant en milieu urbain sont dans la catégorie Riche, environ 41% dans les catégories Moyen et aisé et seulement 14% des ménages interrogés sont pauvre ou extrêmement pauvre.

En revanche, concernant les ménages interrogés en milieu rural, seulement 23% sont dans les catégories riche et aisé, environ 28% dans la catégorie moyen, tandis que 48% de ces ménages sont pauvre ou extrêmement pauvre.

De plus, avec le même raisonnement que précédemment, on définit la p-valeur = $IP_{(H_0)}(\text{rejeter } (H_0)) = IP_{(H_0)}(D^2 \geq d_0^2) = IP(X^2(5) \geq 9722.2)$ avec X^2 suit une loi de chi2 de ddl = $(6-1)(2-1) = 5$.

Ici, nous avons également que les effectifs théoriques sont inférieurs à 5% donc le test de chi2 peut être réalisé.

Croisement entre la pauvreté extrême et lieu de résidence		
X-squared	df	P-value
9722.2	5	< 2.2 e-16

D'après le résultat de ce test, la p-valeur est inférieure à 2.2 e-16, ce qui signifie que c'est largement inférieur à 5%. Dans ce cas, on rejette l'hypothèse (H_0) : les variables pauvreté extrême et lieu de résidence sont indépendantes et affirmer qu'il existe bien un lien significatif entre ces deux variables (H_1).

On rappelle aussi que la modalité Urban est celle qui contribue le plus à la statistique du chi2 avec une valeur de 5972.765, par conséquent c'est la modalité la plus atypique.

On en conclut donc qu'un foyer se trouvant dans un milieu rural aura plus de chances d'être dans une situation de pauvreté extrême qu'un foyer en milieu urbain.

c. Croisement entre les variables pauvreté matérielle et lieu de résidence

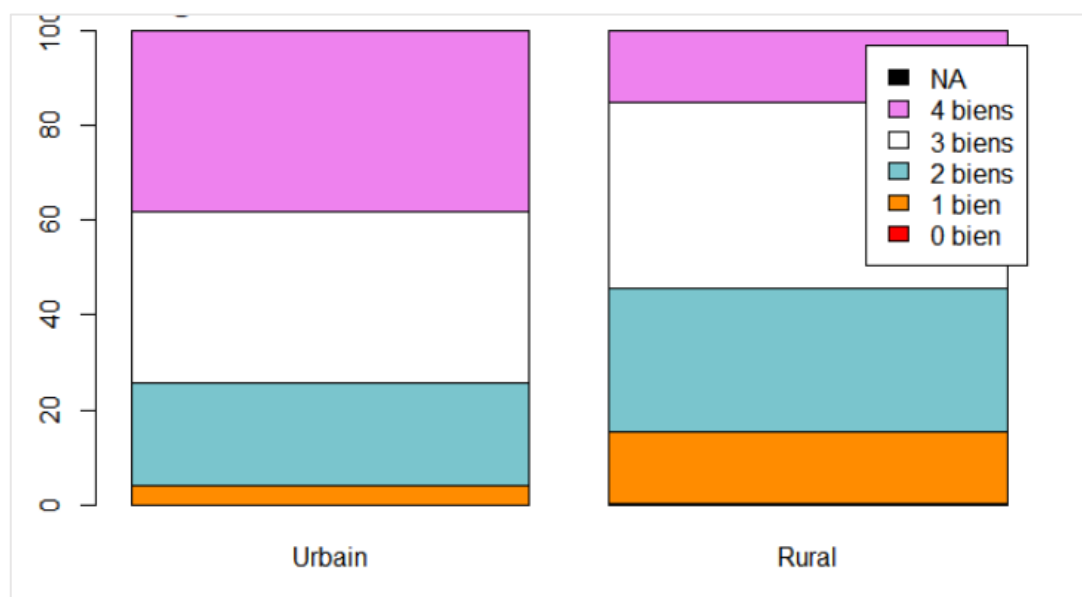


Figure 29 : Pauvreté matérielle selon le lieu de résidence

	Urban	Rural
0 bien	0.04460635	0.15867830
1 bien	4.01457141	15.05110375
2 biens	21.55973534	30.43356513
3 biens	36.14601145	39.30554907
4 biens	38.04178128	14.88309143
NA	0.19329418	0.16801232

Figure 30 : pourcentage en colonne de la pauvreté matérielle et lieu de résidence

Concernant le croisement entre les variables pauvreté matérielle et lieu de résidence, on cherche à déterminer la répartition de la pauvreté matérielle entre le milieu urbain et rural. En effet, on remarque que plus de 74% des ménages interrogés vivant dans le milieu urbain ont plus de 2 biens matériels contre environ 54% dans le milieu rural. Cela signifie qu'il y a un peu plus de ménages ayant accès à des biens matériels dans le milieu urbain que rural. En revanche, il y a environ 4 % des ménages interrogés n'ayant accès à aucun bien ou uniquement qu'à 1 seul dans le milieu urbain contre 15% dans le milieu rural. Il y a très peu de ménages résidant en milieu rural ayant tous les biens matériels (environ 15%), cela pourrait être dû au fait qu'un ménage peut avoir une télévision ou la radio mais pas les 2 à la fois donc ce résultat n'est pas significatif.

De la même façon, on définit la p-valeur = $P(X^2(5) \geq 3065.7)$ avec X^2 suit une loi de χ^2 de ddl = 5.

Ici, nous avons également que les effectifs théoriques sont inférieurs à 5% donc le test de χ^2 peut être réalisé.

D'après le résultat de ce test, la p-valeur est inférieure à 2.2×10^{-16} , ce qui signifie que c'est largement inférieur à 5%. Dans ce cas, on rejette l'hypothèse (H_0) : les variables pauvreté matérielle et lieu de résidence sont indépendantes et affirmer qu'il existe bien un lien significatif entre ces deux variables (H_1). On en conclut qu'il y a plus de facilité à accéder à des biens matériels dans un milieu urbain.

Croisement entre la pauvreté matérielle et lieu de résidence		
X-squared	df	P-value
3065.7	5	< 2.2 e-16

d. Croisement entre les variables pauvreté matérielle et région

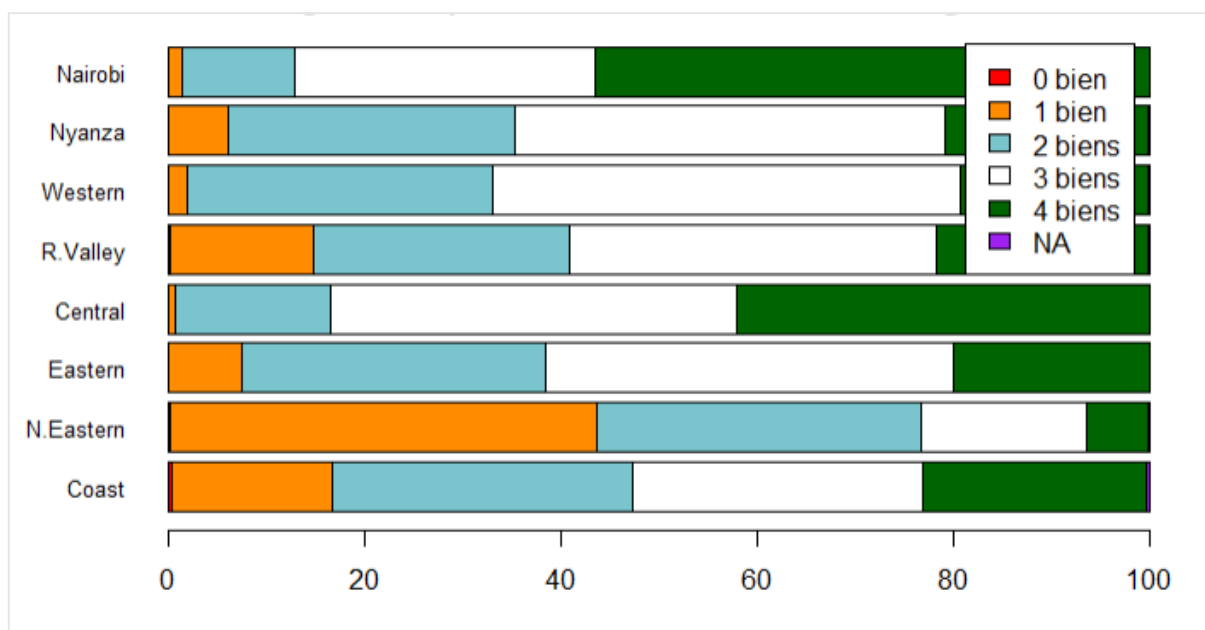


Figure 32 : Pauvreté matérielle selon les régions

	Coast	N.Eastern	Eastern	Central	R.Valley	Western	Nyanza	Nairobi
Aucun bien	0.34786642	0.22222222	0.07037298	0.00000000	0.13020833	0.06303183	0.04172752	0.00000000
1 bien essentiel	16.28014842	43.38888889	7.33638283	0.78361982	14.69350962	1.85943902	6.07135406	1.44557823
2 biens essentiels	30.63543599	33.05555556	31.10485574	15.69767442	26.01161859	31.07469272	29.18839975	11.39455782
3 biens essentiels	29.49907236	16.88888889	41.53764954	41.43073812	37.39983974	47.74661204	43.77216774	30.69727891
Tous les biens essentiels	22.79684601	6.22222222	19.93314567	42.06268959	21.59455128	19.03561298	20.65512205	56.46258503
MissingData	0.44063080	0.22222222	0.01759324	0.02527806	0.17027244	0.22061141	0.27122888	0.00000000

Figure 31 : Pourcentage en colonne de la pauvreté matérielle selon les différentes régions

Dans la figure 31, on retrouve le graphe des profils colonnes du couple (pauvreté matérielle et région). Sur ce graphe, nous pouvons distinguer des profils très différents. En effet, on retrouve tout d'abord la capitale Nairobi, pour laquelle la plupart des foyers interrogés ont accès à un tous les biens (56%), suivi de la région Central (42%).

Les quatre régions Western, Nyanza, Central et Nairobi présentent un faible nombre de ménages n'ayant accès à aucun ou qu'à un seul bien, ce qui signifie qu'elles sont des régions où les foyers sont plutôt bien équipés.

En contrepartie, les régions Coast, N.Eastern et R.Valley, ont plus de 19% des ménages interrogés ayant accès à un seul bien. Il y a plus de 30% des ménages ayant accès à 3 biens matériels dans toutes les régions à l'exception de la région N.Eastern (17%) qu'on peut qualifier d'une région assez pauvre en termes de biens matériels selon nos critères. De plus, c'est la région qui contribue le plus à la statistique du chi2 ainsi que la région Central, nous pouvons donc les considérer comme des points atypiques.

En faisant le test du chi2, on s'aperçoit que :

p-valeur = $IP(X^2(35) \geq 5248.9) < 2.2 \text{ e-}16$ avec X^2 suit une loi de chi2 de ddl = 35. Ce qui signifie que c'est largement inférieur à 5%. Dans ce cas, on rejette l'hypothèse (H_0) : les variables pauvreté matérielle et région sont indépendantes et affirmer qu'il existe bien un lien significatif entre ces deux variables (H_1).

Croisement entre la pauvreté matérielle et région		
X-squared	df	P-value
5248.9	35	< 2.2 e-16

Coast	N.Eastern	Eastern	Central	R.Valley	Western	Nyanza	Nairobi
263.740	2244.138	161.141	1129.248	166.418	363.304	172.814	748.065

Figure 33 Contribution des colonnes à la statistique du chi2

On peut remarquer d'après la figure 33, que les modalités N.Eastern, Central et Nairobi sont les modalités contribuant le plus à la statistique du chi2. Ce sont des régions dont les foyers interrogés s'écartent de la moyenne, ils ont soit beaucoup de biens matériels qui est le cas de Central et Nairobi soit peu de biens matériels pour N.Eastern.

Nous allons maintenant faire le croisement entre la pauvreté extrême et la région afin de déterminer plus précisément la répartition de la pauvreté avec ses deux dimension extrême et matérielle dans les différentes région au Kenya.

e. Croisement entre les variables pauvreté extrême et région

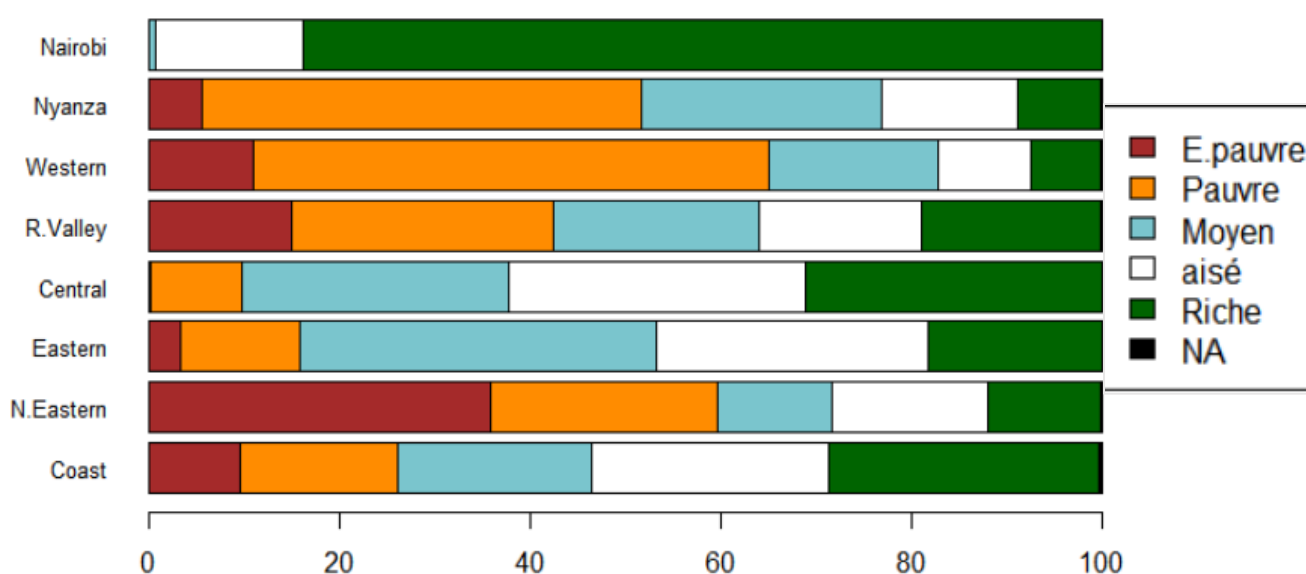


Figure 34 : Pauvreté extrême selon les région

	Coast	N.Eastern	Eastern	Central	R.Valley	Western	Nyanza	Nairobi
Extrêmement pauvre	9.62430427	35.83333333	3.36030964	0.10111223	14.99399038	10.87299086	5.57062383	0.00000000
Pauvre	16.48886827	23.83333333	12.50879662	9.65621840	27.40384615	54.14434289	46.04631755	0.00000000
Moyen	20.24582560	12.00000000	37.38564391	28.03336704	21.53445513	17.77497636	25.18255790	0.76530612
aisé	24.95361781	16.27777778	28.41308937	31.09201213	17.00721154	9.73841790	14.37513040	15.39115646
Riche	28.29313544	11.77777778	18.29697396	31.04145602	18.95032051	7.34320832	8.63759649	83.84353741
MissingData	0.39424861	0.27777778	0.03518649	0.07583418	0.11017628	0.12606366	0.18777384	0.00000000

Figure 35 : pourcentage en colonne de la pauvreté extrême dans les différentes régions

La figure 34 correspondant à la répartition de la pauvreté extrême dans les différentes régions montre que 84% des foyers interrogés résidant dans la capitale Nairobi se retrouvent dans la catégorie riche et 15% dans la catégorie moyen.

On retrouve les régions North Eastern et Western comme étant les régions les plus pauvres, pour lesquelles presque 60% de la population interrogée vit dans un logement insalubre, ils sont donc soit pauvre ou extrêmement pauvres. Ensuite, on retrouve les régions Rift Valley et Nyanza avec un pourcentage de plus de 40% des ménages interrogés pauvres ou extrêmement pauvres.

La région Central reste quant à elle une région pour laquelle les foyers interrogés sont bien équipés, et à un profil similaire à la région Eastern. Ici, la région Coast ne présente pas de signes de pauvreté extrême trop importants.

De plus, après avoir effectué le test de chi2 dont le résultat est donné par ce tableau, on remarque que la p-valeur = $IP(X^2(35) \geq 5248.9) < 2.2 \text{ e-}16$ avec X^2 suit une loi de chi2 de ddl = 35. Ce qui signifie que c'est largement inférieur à 5%. Dans ce cas, on rejette l'hypothèse (H_0) : les variables pauvreté extrême et région sont indépendantes et affirmer qu'il existe bien un lien significatif entre ces deux variables (H_1).

Croisement entre la pauvreté extrême et région		
X-squared	df	P-value
10479	35	< 2.2 e-16

Conclusion

Pour synthétiser, les régions Nairobi, Central sont les régions les plus riches et mieux équipées en termes de conditions matérielles tandis que les régions R.Valley, N. Eastern et Nyanza, sont les régions les plus pauvres et moins équipées. Concernant, les autres régions, on peut les classer dans la classe moyenne où on trouve une hétérogénéité.

Afin de déterminer plus en détail la répartition de la pauvreté au Kenya, nous allons faire une Analyse en Composantes Multiples (ACM) qui permet de croiser cette fois-ci toutes les modalités des variables pour avoir plus de précision sur les équipements de chaque région et de chaque lieu de résidence. En effet, grâce à l'étape tris croisés, nous avons pu voir les nombres de biens de chaque région mais à l'aide de l'ACM, nous allons déterminer quels sont ces biens et les caractéristiques de chaque région.

Pour cela, nous allons recoder à nouveau les modalités de nos variables en mettant un nom représentatif de la modalité. Prenons l'exemple de la variable HV214 correspondant au matériau des murs dans un ménage dont on a détaillé le codage précédemment, son nouveau recodage est défini donc ainsi :

$HV214 < 20 \rightarrow \text{"Mur-"}$

$HV214 > 20 \rightarrow \text{"Mur+"}$

Cela signifie que "Mur+" correspond à un bon matériau et "Mur-" désigne un matériau caractérisant un signe de pauvreté.

C. ACM

C.1 Contribution et coordonnées des modalités sur l'axe 1 et 2

	Dim 1	Dim 2
Eau-	3.781	0.083
Eau+	0.552	1.651
Eau++	6.639	0.056
WC-	4.353	17.544
WC+	0.161	9.331
WC++	12.300	8.323
Elc-	5.796	0.000
Elc+	15.298	0.001
R-	2.795	8.601
R+	1.573	4.840
TV-	4.965	0.111
TV+	12.790	0.287
Fr-	0.443	0.632
Fr+	8.850	12.624
Sol-	6.750	1.820
Sol+	6.058	7.525
Sol++	5.078	17.740
Mur+	0.000	0.000
T+	0.000	0.000
Tel-	0.003	0.009
Tel+	0.608	2.155
Cui-	0.045	0.044
Cui+	0.203	0.171
Cui++	0.001	0.006
Mos-	0.624	4.194
Mos+	0.335	2.253

Figure 37 : Contributions

	Dim 1	Dim 2
Eau-	-0.250	-0.016
Eau+	-0.250	0.189
Eau++	0.383	-0.015
WC-	-0.494	0.434
WC+	-0.044	-0.146
WC++	0.985	0.355
Elc-	-0.264	0.001
Elc+	0.698	-0.002
R-	-0.261	0.200
R+	0.147	-0.113
TV-	-0.246	0.016
TV+	0.632	-0.041
Fr-	-0.064	-0.033
Fr+	1.274	0.666
Sol-	-0.323	0.073
Sol+	0.362	-0.177
Sol++	1.218	0.996
Mur+	0.000	0.000
T+	0.000	0.000
Tel-	-0.005	-0.004
Tel+	1.101	0.907
Cui-	0.155	0.067
Cui+	-0.355	0.143
Cui++	0.003	-0.003
Mos-	-0.125	0.142
Mos+	0.067	-0.076

Figure 36 : Coordonnées

D'après la figure 37 correspondant aux contributions des modalités à l'inertie de l'axe 1 et 2, on remarque que les modalités WC++, Elc+ et TV+ sont les modalités qui contribuent le plus à la formation de l'axe 1, tandis que Sol++, WC- et Fr+ sont celles qui contribuent le plus à la formation de l'axe 2.

De plus, en analysant les coordonnées des modalités sur les 2 premiers axes, contenues dans la Figure 36, on peut déduire que l'axe 1 oppose les modalités (+) ou (++) du côté positif aux modalités (-) du côté négatif. Sur la représentation des modalités sur le premier plan factoriel de la figure 41 ci-dessous, on s'attend à avoir les modalités traduisant une richesse, ou un confort à droite de la carte,

tandis que les modalités traduisant un manque de bien ou des conditions de vie mauvaises à gauche. Un individu à droite de notre carte aura donc plus de biens qu'un individu situé à gauche.

L'axe 2 quant à lui oppose les modalités qualifiant une grande richesse / pauvreté aux modalités qualifiant des conditions de vie moyenne.

C.2 Qualités de représentation des modalités

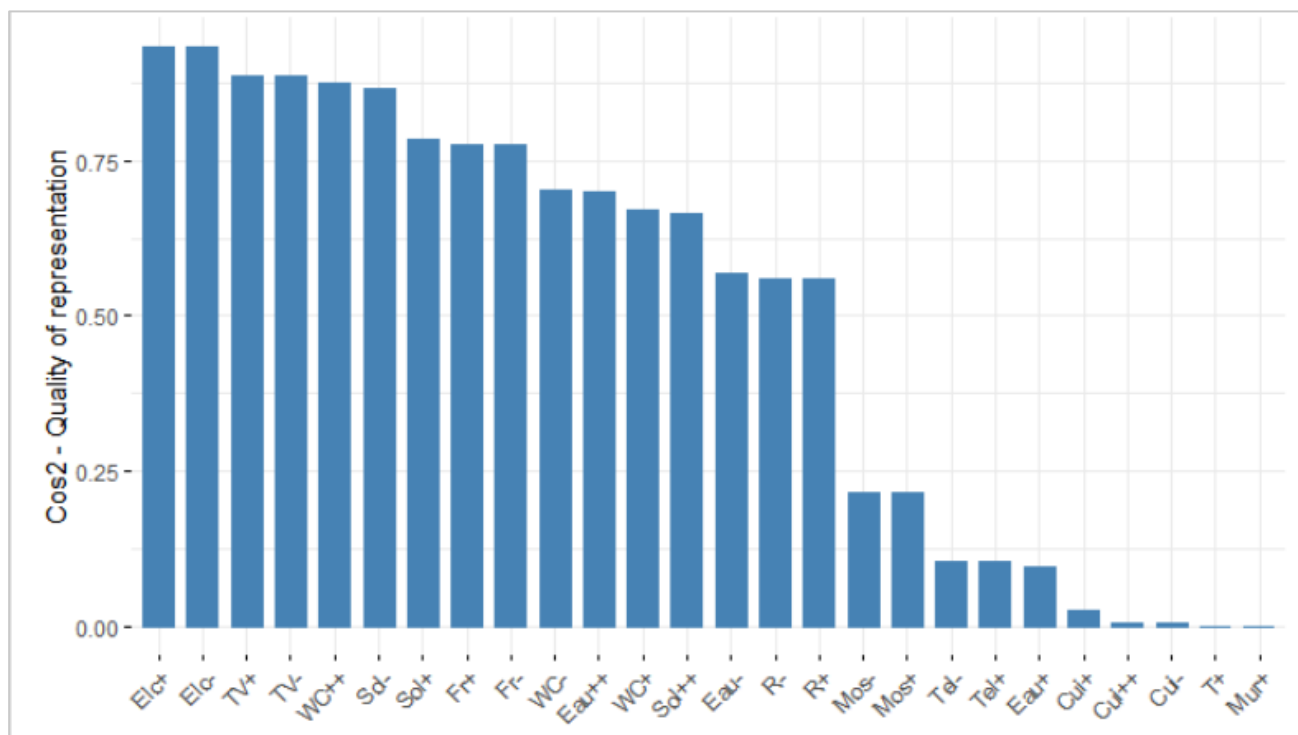


Figure 38 : Qualités de représentation des modalités sur le 1er plan factoriel

Pour étudier la qualité de représentation des individus sur les axes, on calcule le cos2 pour toutes nos modalités et on les représente dans la Figure 38. Plus cette valeur est proche de 1 plus la qualité de représentation est meilleure.

On s'aperçoit ici que les modalités Cui+, Cui++, Cui-, T+ et Mur+ sont très mal représentées dans le premier plan factoriel. On ne peut donc rien dire quant à ces modalités, ce qui explique certains résultats incohérents au niveau des coordonnées dans la Figure 36. On en déduit également que les modalités Mos+, Mos-, Tel-, Tel+ et Eau+, sont moins bien représentées avec un cos2 inférieur à 0.25. En revanche, toutes les autres modalités sont très bien représentées avec un cos2 supérieur à 0.5.

C.3 Inertie et taux d'inertie modifiés

	eigenvalue	percentage of variance	cumulative percentage of variance
dim 1	0.073	51.143	51.143
dim 2	0.014	9.800	60.943
dim 3	0.010	6.724	67.667
dim 4	0.008	5.287	72.953
dim 5	0.007	4.712	77.665
dim 6	0.006	4.481	82.146
dim 7	0.006	4.268	86.413
dim 8	0.005	3.775	90.189
dim 9	0.004	2.912	93.101
dim 10	0.003	2.109	95.210
dim 11	0.003	1.790	97.000
dim 12	0.002	1.333	98.333
dim 13	0.002	1.183	99.516
dim 14	0.001	0.484	100.000

Figure 39 : Inertie expliquée par les différents axes

Nous rappelons que le nombre total d'axes vaut $\min(M-p, n-1) = \min(26-12, n-1) = 14$.

D'après la figure 39, nous remarquons que l'axe 1 permet de restituer 51% de l'inertie totale et l'axe 2 environ 10%. Ainsi, le premier plan factoriel permet donc de restituer 61 % de l'inertie totale. Nous pouvons donc retenir que 2 axes.

En supposant que nous avons choisi deux axes, d'après le critère des taux qu'on a calculé (tableau à coté) et représenté dans la figure 40, nous pouvons garder 2 axes comme il est aussi facile de lire en 2 qu'en 1 dimension, on garde 2 axes.

dim 1	dim 2
0.02219078	0.97780922

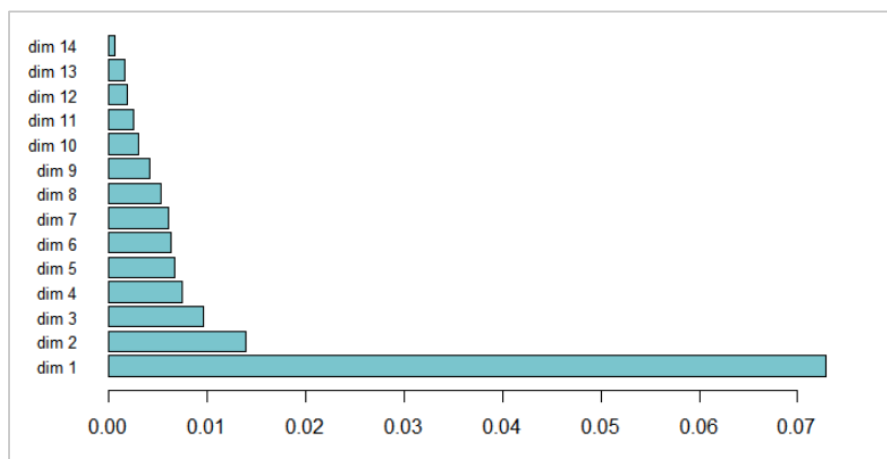


Figure 40 : Inertie restitué par chaque axes

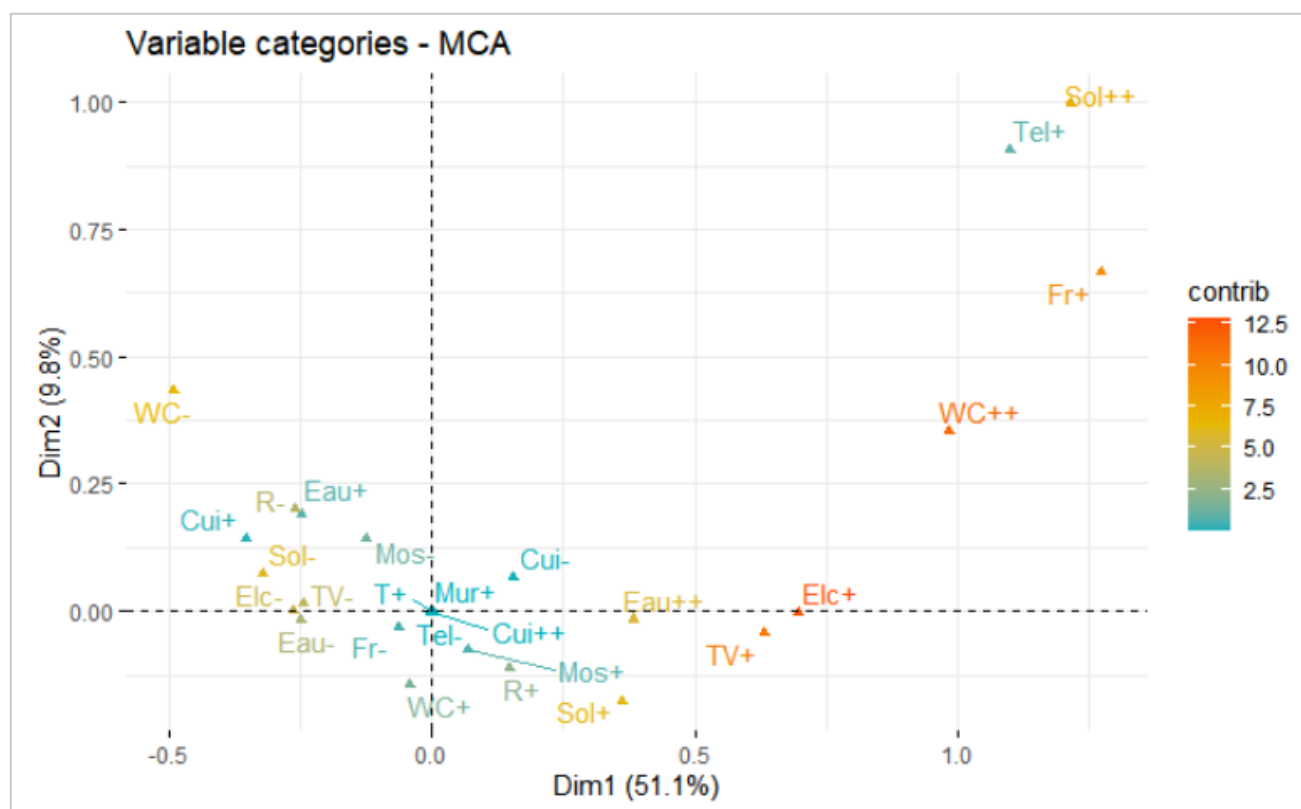


Figure 41 : ACM sur les modalités à partir du tableau de Burt

En regardant la forme de l'ACM obtenue, on remarque l'apparition de l'effet Guttman, se traduisant par une représentation relativement parabolique, passant par l'origine.

Avant d'analyser la carte des modalités, on rappelle que l'axe 1 oppose les extrêmes, avec à droite des foyers caractérisés par la richesse et à gauche des foyers caractérisés par la pauvreté. De plus l'axe 2 est un axe d'opposition entre extrêmes et moyens. On retrouvera les foyers très riches ou très pauvres en haut de la carte tandis que les foyers avec des conditions de vie moyennes en bas de la carte.

On voit donc apparaître 4 principales catégories. Les foyers très riches seront situés en haut à droite de la carte, caractérisés principalement par la possession de nombreux biens, comptant parmi eux un réfrigérateur, un sol et des toilettes de bonne qualité pour leur habitation. En bas à droite de la carte seront situés les foyers riches caractérisés par un accès à l'électricité, à l'eau potable, à la télévision et possédant un sol de qualité moyenne.

Les foyers situés en bas à gauche de la carte seront des foyers pauvres, caractérisés par l'absence de réfrigérateur, un mauvais accès à l'eau potable et généralement peu de biens et de services. Enfin, dans la partie en haut à gauche de la carte, seront situés les foyers très pauvres caractérisés par peu de biens matériels et une qualité de vie très basse, vivant sans électricité, avec un sol et des toilettes de mauvaise qualité.

Il est difficile d'analyser les autres modalités (Cui++, Cui-, Cui+, T+, Tel+ et Tel-) car celles-ci sont mal représentées dans le premier plan factoriel, comme le montre la figure 38. Mais nous pouvons apercevoir des profils de richesse et de pauvreté apparaître. Nous allons désormais nous intéresser à la caractérisation de la pauvreté au sein des régions.

C.5 ACM sur les modalités à partir du tableau disjonctif complet

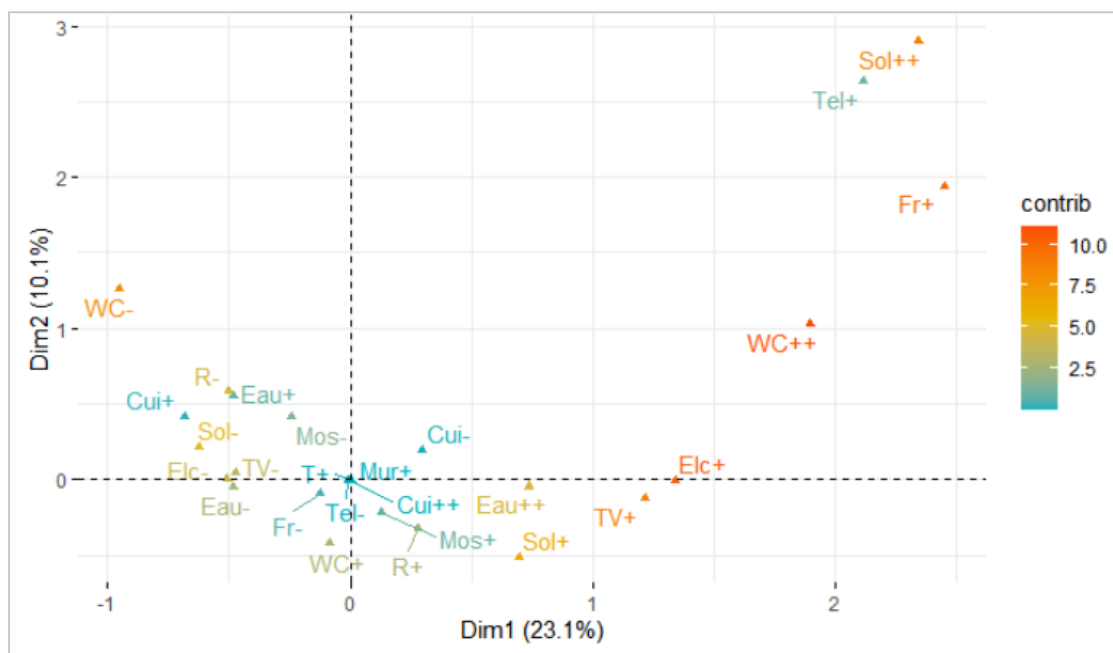


Figure 42 : ACM sur les modalités à partir du tableau disjonctif complet

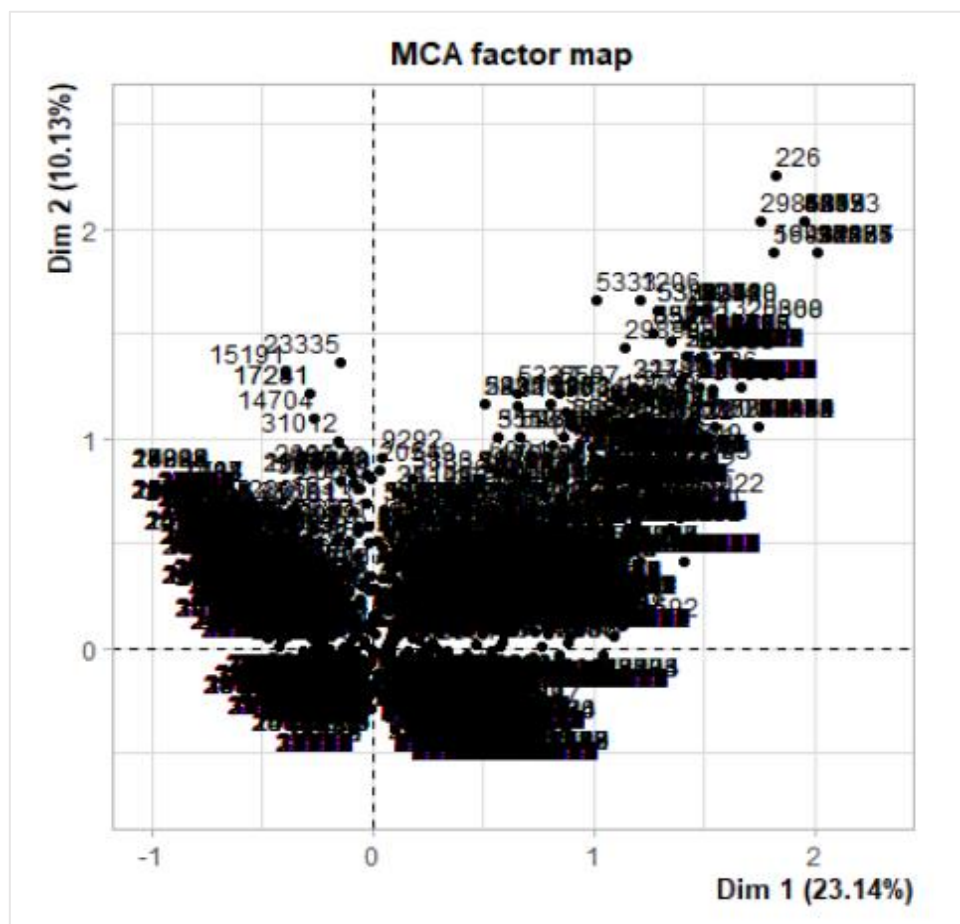


Figure 43 : Carte des individus

Nous retrouvons des résultats similaires avec l'ACM à partir du tableau de Burt. Les différences observées sur les graphiques sont juste que l'ACM de Burt ne permet pas une représentation des individus alors que les individus sont représentés sur l'ACM du tableau disjonctif. Dans notre cas, cette représentation n'est pas indispensable car les individus ne sont pas nommés.

Néanmoins, on constate une forte concentration de ménages situé vers l'origine au niveau de la carte des individus, qui représentent une population « moyenne », ni trop riche, ni trop pauvre. A contrario, les ménages situés aux extrémités des axes sont des foyers très riches, ou très pauvres. Les individus suivent eux aussi la courbe typique de Guttman de façon plus grossière. Des individus se démarquent aux extrêmes des axes et sont donc des ménages sous-équipés ou suréquipés comme vu sur l'ACM précédente.

D. ACM avec les variables région et lieu de résidence

D.1 Qualités de représentation

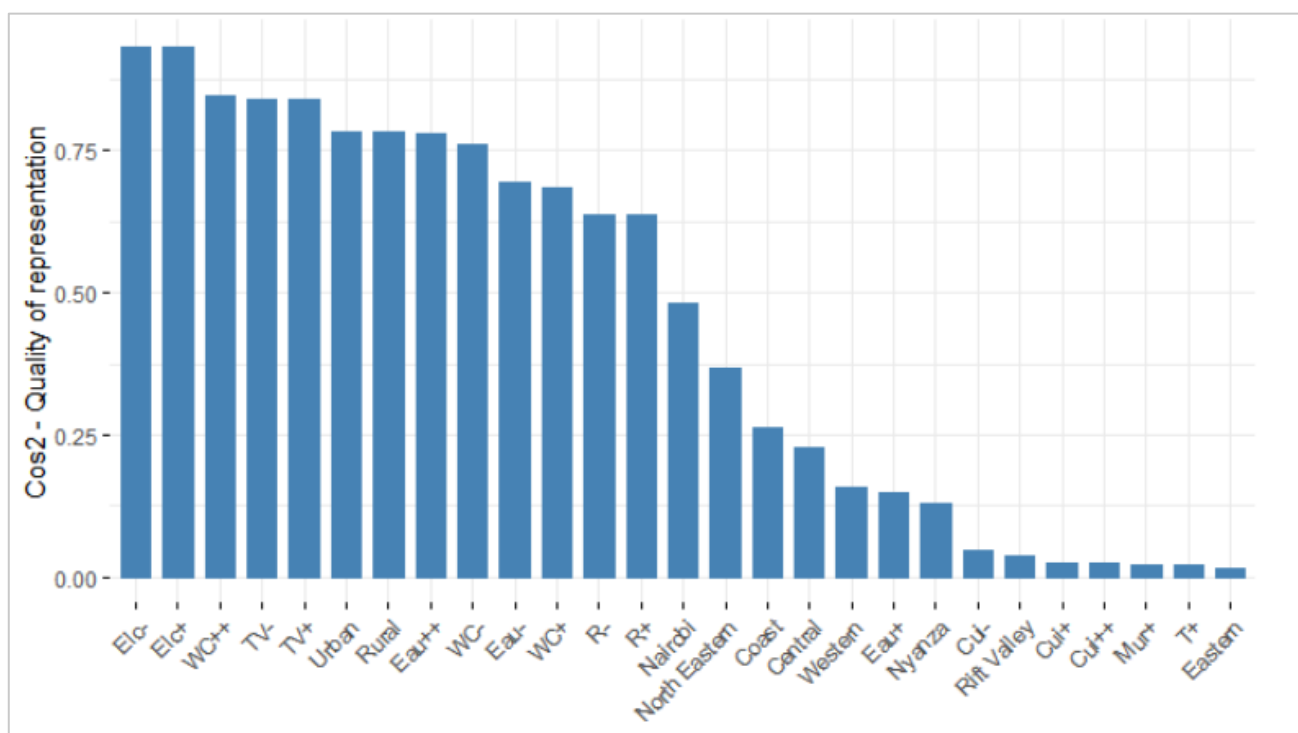


Figure 44 : qualités de représentation sur le 1^{er} plan factoriel avec les deux variables de localisation

De même que pour l'ACM faite à partir du tableau de Burt lors de la partie C, on remarque que certaines modalités restent mal représentées comme les modalités de type de cuisine, Mur+ et T+ ainsi que les modalités Eastern et Rift Valley, caractérisant des régions ayant un \cos^2 inférieur à 0.05. On pourra tout de même étudier les régions Nairobi, North Eastern, Coast, Central et Western ainsi que les deux modalités du lieu de résidence Urban et Rural.

D.2 Coordonnées et contributions

	Dim 1	Dim 2
Central	0.262	-0.350
Coast	0.203	0.404
Eastern	-0.063	-0.061
Nairobi	1.282	0.404
North Eastern	-0.352	0.795
Nyanza	-0.174	-0.238
Rift Valley	-0.091	0.036
Western	-0.213	-0.349
Rural	-0.289	-0.041
Urban	0.462	0.065
Eau-	-0.294	-0.103
Eau+	-0.284	0.324
Eau++	0.449	0.074
WC-	-0.519	0.557
WC+	-0.051	-0.166
WC++	1.068	0.312
Elc-	-0.281	-0.010
Elc+	0.743	0.028
R-	-0.249	0.286
R+	0.140	-0.161
TV-	-0.248	0.036
TV+	0.639	-0.093
Mur+	0.000	0.000
T+	0.000	0.000
Cui-	0.262	0.465
Cui+	-0.421	0.056
Cui++	0.002	-0.009

Figure 45 : Coordonnée des modalités avec les variables supplémentaires

	Dim 1	Dim 2
Central	0.841	5.747
Coast	0.545	8.207
Eastern	0.070	0.249
Nairobi	5.985	2.271
North Eastern	0.691	13.403
Nyanza	0.449	3.196
Rift Valley	0.257	0.150
Western	0.446	4.559
Rural	5.533	0.422
Urban	8.848	0.674
Eau-	4.924	2.314
Eau+	0.671	3.351
Eau++	8.574	0.899
WC-	4.541	19.939
WC+	0.204	8.294
WC++	13.630	4.440
Elc-	6.197	0.032
Elc+	16.358	0.086
R-	2.415	12.092
R+	1.359	6.805
TV-	4.774	0.383
TV+	12.299	0.988
Mur+	0.000	0.000
T+	0.000	0.000
Cui-	0.121	1.451
Cui+	0.269	0.018
Cui++	0.000	0.030

Figure 46 : Contributions des modalités avec les variables supplémentaires

D'après la figure 46 correspondant aux contributions des modalités à l'inertie de l'axe 1 et 2, on remarque que les modalités Wc++, Tv+, Elc+, Eau++, Nairobi et Urban contribuent le plus à la formation de l'axe 1, tandis que les modalités R-, Wc-, N. Eastern, Coast et Central contribuent le plus à la formation de l'axe 2.

Concernant les coordonnées de la figure 45, on retrouve le même résultat que l'ACM précédente à partir du tableau de Burt sauf que dans ce cas, l'axe 1 oppose aussi les régions Nairobi, Coast et Central du côté positif aux régions Estern, North Estern, Rift Valley, Western et Nyanza du côté négatif.

A propos de l'axe 2, on retrouvera en haut la région riche Nairobi, ainsi que North Eastern qui quant à elle est très pauvre. En bas on retrouve des régions de pauvreté moyenne telles que Nyanza, Western ou Central. On peut également retrouver le lien de corrélation entre la qualité de vie d'un foyer et son lieu de résidence. On remarque que la modalité Urban se situe à droite de la carte, et traduit donc qu'un foyer habitant en ville aura plus d'équipements et des conditions de vie supérieures à un foyer vivant en milieu rural, dont la modalité se trouve à gauche de la carte.

D.3 ACM sur les modalités à partir du tableau de Burt

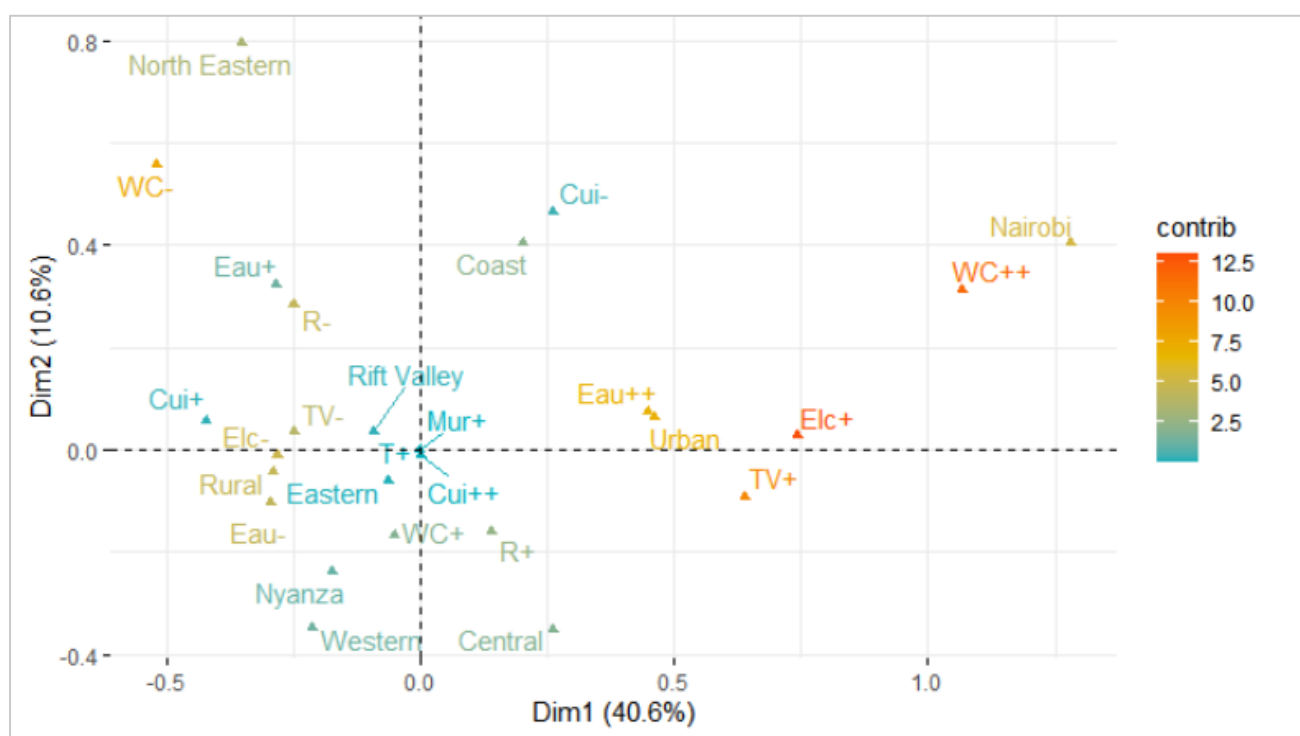


Figure 47 : ACM à partir du tableau de Burt avec les variables supplémentaires

L'allure de notre ACM change comparée aux précédentes, l'effet Guttman est moins présent, mais l'interprétation des axes reste semblable. Sur l'axe 1, on retrouve les modalités symboliques de richesse et de pauvreté, avec à droite les modalités riches, et à gauche les modalités pauvres. L'axe 2 oppose toujours les extrêmes au moyens. Rajoutons maintenant l'interprétation avec les régions et le lieu de résidence. On peut déjà remarquer un phénomène intéressant : la région Coast est située dans la partie haute de la carte, au milieu de celle-ci et en faisant le lien avec le résultat des tris croisés, on peut s'apercevoir que Coast est une région très hétérogène, avec des foyers très pauvres et d'autres très riches, selon nos critères de pauvreté matérielle et extrême. Ainsi, il est difficile de définir un profil pour cette région, car sa population est très diversifiée.

On remarque également que Nairobi et Urban sont du côté droit de la carte, faisant de ces modalités des caractéristiques géographiques d'un foyer possédant une bonne qualité de vie. North Eastern est située à gauche de la carte, accompagnée de Rural, Nyanza et d'autres modalités de pauvreté. On en déduit que Nairobi compte parmi ses habitants beaucoup de foyers bien équipés, et que de manière

générale, les foyers interrogés en zone urbaine vivent mieux que les foyers interrogés en zone rurale. De plus, on confirme que la région North Eastern est très pauvre, caractérisée par le nonaccès à la radio et la possession des toilettes de mauvaises qualités, dont les foyers vivent dans de très mauvaises conditions de vie. Concernant la région Central, elle est caractérisée par la possession de la radio et des toilettes de bonne qualité, c'est une région un peu riche. Enfin, les régions Western et Nyanza sont des régions pauvres, caractérisées par un accès précaire à l'eau potable, la non-possession de télévision et le nonaccès à l'électricité mais des toilettes de bonnes qualités.

D.4 ACM avec les variables de localisation en supplémentaire

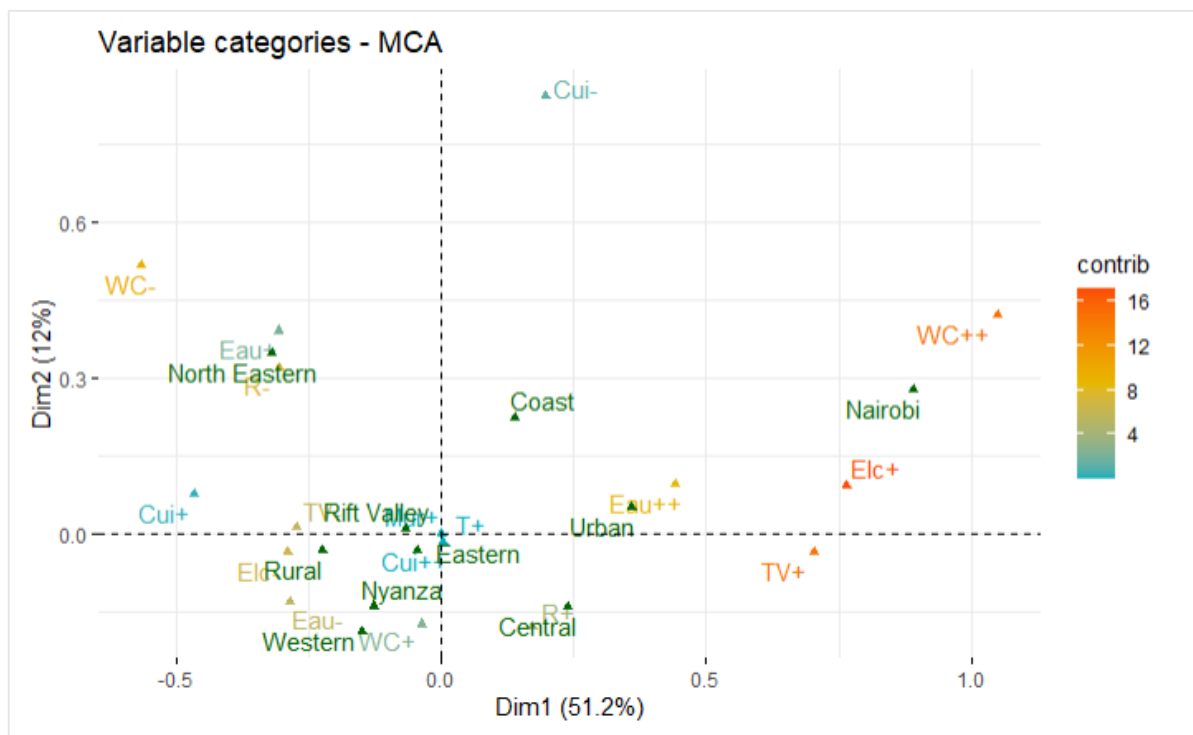


Figure 48 : ACM avec les variables de localisation en supplémentaire

En mettant les variables région et lieu de résidence en supplémentaires, on peut remarquer quelques changements. Les modalités des régions et des lieux de résidence ont l'air de se rapprocher du centre de la carte, sans pour autant vraiment changer les interprétations que nous avons fournies auparavant.

Une valeur test mesure l'écart en moyenne de la modalité au centre de gravité du nuage. Des valeurs tests supérieures à 2 en valeur absolue d'une modalité, signifie que les individus qui l'a possède a tendance à s'éloigner de la moyenne. En effet, d'après la figure 49, on voit que Nairobi, Central et Urbain a une très forte valeur test (largement supérieures à 2) sur l'axe 1, ce qui signifie que les ménages situés dans le milieu urbain à Nairobi et Central sont des ménages plutôt riche donc supérieur à la moyenne, tandis que les ménages résidant dans le milieu rural en N. Estern, Rift Vlley, Nyanza et Western sont plutôt pauvres (valeurs test inférieures à -2).

	Dim 1	Dim 2
Central	15.90965	-9.33039
Coast	9.62070	15.54692
Eastern	-3.69545	-2.59622
Nairobi	30.99918	9.68002
North Eastern	-13.90431	15.06687
Nyanza	-9.47200	-10.31086
Rift Valley	-7.89392	1.05965
Western	-8.79782	-11.05159
Rural	-52.80962	-7.59128
Urban	52.80962	7.59128

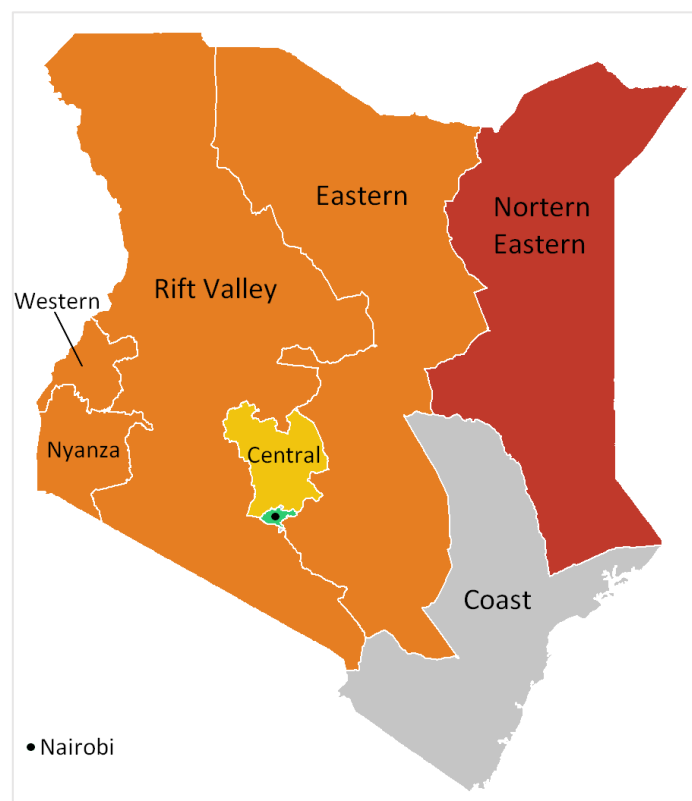
Figure 49 : Valeurs test pour région et lieu de résidence

Synthèse, critique et conclusion

L'un des premiers résultats de notre analyse statistique est qu'une grande partie des ménages au Kenya vivent dans des conditions de pauvreté : 13.82% des ménages urbains sont considérés pauvres (pauvreté extrême et pauvreté) dans notre étude et 48.52% pour les ménages ruraux. Le second résultat est que la majorité des ménages considérés comme pauvres sont dans des milieux ruraux.

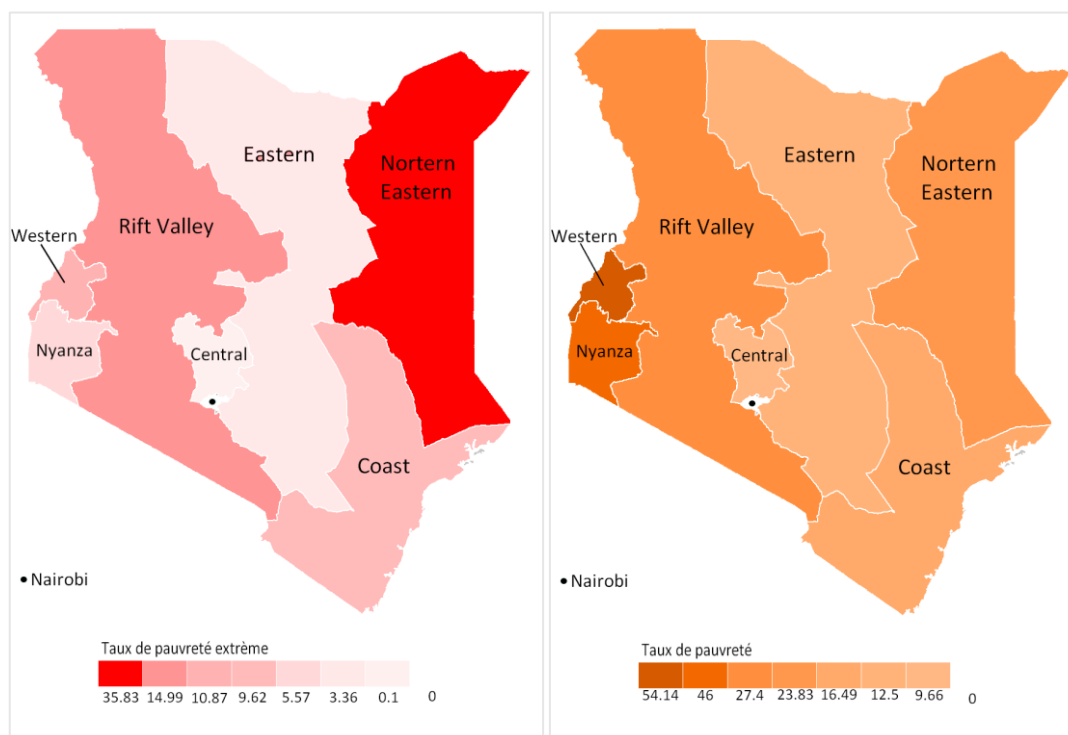
D'après nos deux variables de pauvreté, la région de Nairobi est sans conteste la région la plus riche (en vert sur la carte ci-dessous), vient ensuite la région Central qui est considérée comme modeste (en jaune). Les régions Rift Valley, Eastern, Nyanza et Western sont quant à elles identifiées comme pauvres. La région Northern Eastern (en rouge) est considérée comme extrêmement pauvre. Cependant, Coast est particulière car c'est une région où il y a des ménages à la fois riches et très pauvres.

Il semblerait qu'au plus on s'éloigne de la capitale, au plus la pauvreté grandit. De plus, nous avons pu voir qu'habiter en zone urbaine favorisait l'accès à de meilleures conditions de vie pour les ménages interrogés.



De cette carte, nous pouvons donc constater que les régions de Nairobi, Central et N.Eastern se distinguent par rapport aux autres régions : elles représentent des extrêmes.

Pour affiner notre analyse et apporter des distinctions entre les quatre régions : Rift Valley, Nyanza, Western et Eastern. Nous allons représenter le taux de pauvreté et de pauvreté extrême sur la carte du Kenya.



On peut donc en conclure qu'il existe tout de même des différences entre ces quatre régions : Rift Valley est la région où il y a le plus de ménages dans des conditions de pauvreté et Eastern est la région qui est le moins touché par la pauvreté.

Au-delà de ces conclusions, nous avons rencontré des difficultés et des remises en question lors de cette étude. Le fait de ne pas avoir accès au salaire des ménages interrogés nous a pénalisé, et nous aurait permis de mieux quantifier la pauvreté. De plus, nous n'avons pas pu expliquer comment qualifier la région Coast. Ceci est principalement dû au fait que les foyers interrogés étaient très hétérogènes au niveau de leur réponses. On voit apparaître alors l'idée d'effectuer l'étude non pas au niveau des régions, mais au niveau des comtés, afin d'avoir une vision plus précise de la répartition de la pauvreté au sein du Kenya.

Enfin, certaines modalités de notre ACM étaient mal représentées, et n'ont pas pu rentrer dans la caractérisation de la pauvreté au sein des régions. Il aurait été intéressant d'analyser uniquement certaines variables afin de tenir compte de ces modalités.

Pour finir, cette étude permet d'avoir une photographie du Kenya en 2014. On peut en conclure que la pauvreté est encore aujourd'hui assez importante dans ce pays et qu'il est possible d'observer une répartition particulière de la pauvreté. Cette étude pourrait par exemple servir à prévoir les régions les plus faibles face à des crises économiques ou encore sanitaire. En 2020, l'ensemble des pays du monde ont dû faire face à la crise de la COVID-19, le Kenya a vu 2 millions de personnes de sa population vivre sous le seuil de pauvreté, il est très probable que la grande partie de cette population se situe dans les régions majoritairement rurales.

1. JRC Technical Report, Comision Européenne. (2020). *COVID-19 impacts and short-term*.
2. (2007). *Kenya Vision 2030 - A Globally Competitive and Prosperous Kenya* .
3. Ménard, S. (2020). *Kenya – Situation économique et financière*.
4. The World Bank. (2020). *Kenya Economic Update: COVID-19 Erodes Progress in Poverty Reduction in Kenya, Increases Number of Poor Citizens*.
5. The World Bank. (2020). *Kenya's GDP Contracts Under Weight of COVID-19, Impacting Lives and Livelihoods*.
6. Urs Wiesmann (CDE), Boniface Kiteme (CETRAD), and Zachary Mwangi (KNBS). (2016). *Socio-Economic Atlas of Kenya*

Annexe 1 : Recodage des variables

1/ Recodage pour la création des variables de pauvreté matérielle et pauvreté extrême

Source d'eau potable HV201 :

$HV201 < 30$ OU $HV201 = 71 \rightarrow 1$

$HV201 \Rightarrow 30$ ET $HV201 \neq 71 \rightarrow 0$

Type de toilettes HV205 :

$HV205 < 30 \rightarrow 1$

$HV205 \Rightarrow 30 \rightarrow 0$

Matériel des murs HV214 :

$HV214 < 20 \rightarrow 0$

$HV214 \Rightarrow 20 \rightarrow 1$

Matériel du toit HV215 :

$HV215 < 20 \rightarrow 0$

$HV215 \Rightarrow 20 \rightarrow 1$

Type de cuisine HV226 :

$HV226 < 11 \rightarrow 0$

$HV226 = 11$ OU $HV226 = 95 \rightarrow 1$

2/ Recodage pour la création de modalités réduites pour l'ACM

Source d'eau potable HV201 :

$(HV201 \geq 30$ ET $HV201 \leq 32)$ OU $(HV201 \geq 40$ ET $HV201 < 71) \rightarrow \text{"Eau-"}$

$HV201 = 20$ OU $HV201 = 21 \rightarrow \text{"Eau+"}$

$HV201 \geq 10$ ET $HV201 \leq 13$ OU $HV201 = 71 \rightarrow \text{"Eau++"}$

Type de toilettes HV205 :

$HV205 \geq 30$ ET $HV205 \leq 43 \rightarrow \text{"WC-"}$

$HV205 \geq 20$ ET $HV205 \leq 23 \rightarrow \text{"WC+"}$

$HV205 \geq 10$ ET $HV205 \leq 15 \rightarrow \text{"WC++"}$

Accès à l'électricité HV206 :

$HV206 = 1 \rightarrow \text{"Elc+"}$

$HV206 = 0 \rightarrow \text{"Elc-"}$

Accès à la radio HV207 :

HV207 = 1 → "R+"

HV207 = 0 → "R-"

Accès à la télévision HV208 :

HV208 = 1 → "TV+"

HV208 = 0 → "TV-"

Accès à un réfrigérateur HV209 :

HV209 = 1 → "Fr+"

HV209 = 0 → "Fr-"

Matériel du sol HV213 :

HV213 = 11 OU HV213 = 12 OU HV213 = 35 → "Sol-"

HV213 = 21 OU HV213 = 22 OU HV213 = 34 → "Sol+"

HV213 >= 31 ET HV213 <= 33 → "Sol++"

Matériel du murs HV214 :

HV214 < 20 → "Mur-"

HV214 > 20 → "Mur+"

Matériel du toit HV215 :

HV215 < 20 → "T-"

HV215 > 20 → "T+"

Accès à un téléphone HV221 :

HV221 = 1 → "Tel+"

HV221 = 0 → "Tel-"

Type de cuisine HV226 :

HV226 = 95 → "Cui-"

HV226 >= 9 ET HV226 <= 11 → "Cui+"

HV226 >= 1 ET HV226 <= 8] → "Cui++"

Avoir une moustiquaire HV227 :

HV227 = 1 → "Mos+"

HV227 = 0 → "Mos-"

Région HV024 :

HV024 = 1 → "Coast"

HV024 = 2 → "North Eastern"

HV024 = 3 → "Eastern"

HV024 = 4 → "Central"

HV024 = 5 → "Rift Valley"

HV024 = 7 → "Western"

HV024 = 8 → "Nyanza"

HV024 = 9 → "Nairobi"

Lieu de résidence HV025 :

HV025 = 1 → "Urban"

HV025 = 2 → "Rural"