|  |
| --- |
| **REPUBLIQUE DU SENEGAL** |
|  |
| **\*\*\*\*\*\*\*\*\*\*** |
| ***Un Peuple - Un But - Une Foi*** |
| **\*\*\*\*\*\*\*\*\*\*** |
| **Agence nationale de la Statistique et de la démographie** |
|  |
| **\*\*\*\*\*\*\*\*\*\*** |
| **Ecole nationale de la Statistique et de l'Analyse économique Pierre Ndiaye** |
|  |

##### Projet statistique sur R : Evaluation

|  |  |
| --- | --- |
| **Rédigé par** | **Sous la supervision de** |
| SOMA Ben Idriss | M. Aboubacar HEMA |
| *Élève Ingénieur Statisticien Économiste* | *Data-scientist* |
| **Année académique 2024-2025** | |

###### Sommaire

[I. Introduction 1](#_Toc197988152)

[II. I. Analyse de consistence des bases de données 1](#_Toc197988153)

[1. Nettoyage des noms des bases 1](#_Toc197988154)

[2. Nombre de valeurs manquantes 1](#_Toc197988155)

[a. Présence de doublons 1](#_Toc197988156)

[3. liens entre les deux bases 1](#_Toc197988157)

[III. II. Analyse socio-economique des deplaces internes du Sudan en 2023 1](#_Toc197988158)

[4. Pyramide des ages des individus 1](#_Toc197988159)

[5. Caracteristiques du chef de menage 1](#_Toc197988160)

[6. 3) Crowding Index ou l’indice d’affluence 1](#_Toc197988161)

[b. le nombre d’individu par menage 1](#_Toc197988162)

[c. Calcul de l’indice d’affluence 1](#_Toc197988163)

[d. La proportion de menage dont le crowding index est: <1, 1<= crowding\_index<2, 2<= crowding\_index<3 et >=3 1](#_Toc197988164)

[e. statistiques suivant que le menage soit un refugie ou un member de la communaute d’acceuil (host community North) 1](#_Toc197988165)

[IV. Analyse de la securite alimentaire des deplaces internes 1](#_Toc197988166)

[f. Score de consommation alimentaire (SCA): 1](#_Toc197988167)

[g. score de consommation alimentaire 1](#_Toc197988168)

[h. Tableau des poids (total = 16) 1](#_Toc197988169)

[i. Répresentation spatiale (région et département) du SCA et de ses différentes catégorisations. 1](#_Toc197988170)

[7. L’indice réduit des stratégies de survie (rCSI): 1](#_Toc197988171)

[j. Statistiques des variables rcsi 1](#_Toc197988172)

[k. Indice réduit des stratégies de survie en utilisant les Nouvelles variables 1](#_Toc197988173)

[l. statistiques descriptives sur les Nouvelles variables 1](#_Toc197988174)

[m. l’indice rCSI (pondéré) 1](#_Toc197988175)

[n. Tableau des poids 1](#_Toc197988176)

[o. Répresentation spatiale selon le milieu de residence du rCSI 1](#_Toc197988177)

[V. conclusion 1](#_Toc197988178)

[VI. Remerciements 1](#_Toc197988179)

# Introduction

L’Afrique traverse aujourd’hui une période charnière marquée à la fois par des défis socio-économiques majeurs et par une profonde mutation démographique, politique et environnementale. Dans ce contexte, les déplacements forcés de population constituent un phénomène de plus en plus préoccupant, en particulier dans des pays en crise comme le Soudan du Sud. En 2023, ce pays a connu une recrudescence des conflits internes et des catastrophes naturelles, entraînant des vagues massives de déplacements internes. Ces situations affectent gravement la stabilité des ménages, l’accès à l’alimentation, les conditions de logement et la sécurité.

Le présent rapport s’inscrit dans le cadre d’un projet d’analyse statistique appliquée sous R. Il vise à fournir une lecture quantitative et critique de l’enquête menée sur les déplacés internes du Soudan du Sud en 2023, en exploitant les outils de traitement et d’analyse de données offerts par le logiciel R. À travers une démarche structurée, nous aborderons des aspects fondamentaux de la situation humanitaire des déplacés, allant de la cohérence des bases de données à l’étude détaillée de leurs caractéristiques socio-économiques, en passant par des indicateurs essentiels comme l’indice de surpeuplement, le score de consommation alimentaire (SCA) et l’indice réduit des stratégies de survie (rCSI).

L’objectif de ce travail est double : d’une part, mobiliser les compétences acquises dans le cadre du cours de statistique appliquée pour traiter un cas réel, et d’autre part, mettre en lumière des dynamiques sociales, économiques et nutritionnelles cruciales dans un contexte de déplacement. L’intégralité de l’analyse a été réalisée à l’aide de R Markdown afin de produire un document reproductible, structuré et transparent, conformément aux exigences pédagogiques de cet exercice.

# I. Analyse de consistence des bases de données

[1] 3058 1312

la base principale est consttuée de 3058 observations et de 1312 variables .

[1] 22092 144

la base individu est constituée de 144 variables et de 22092 observations

## Nettoyage des noms des bases

On va d’abord standariser les noms de toutes les variables des deux bases en minuscule pour une meilleure analyse .

Nous allons afficher les noms des 10 prémières variables pour les deux bases

Pour la base principale :

| Variable |
| --- |
| intro\_07\_1 |
| start |
| end |
| admin0 |
| admin1 |
| admin2 |
| admin3 |
| final\_01 |
| id |
| positionof\_hh |
| Source: base principale |

Pour la base individu:

| Variable |
| --- |
| intro\_07\_1 |
| start |
| end |
| admin0 |
| admin1 |
| admin2 |
| admin3 |
| final\_01 |
| id |
| rosterposition |
| Source: base individus |

## Nombre de valeurs manquantes

Nombre de valeurs manquantes de la base principale :

[1] 2629978

[1] 2056475

Il ya 2629978 Valeurs manquantes dans la base principale et 2056475 dans la base individus

Nous allons afficher les colonnes ou les valeurs manquantes dépassent 70% pour la base principale. Voici une répresentation des valeurs manquantes en fonction des 10 premières valeurs :

| Variable | Nb\_Valeurs\_Manquantes | Pourcentage |
| --- | --- | --- |
| hh\_lang03b | 2,464 | 80.6 |
| hh\_31b1 | 3,056 | 99.9 |
| hh\_31b2 | 3,056 | 99.9 |
| hh\_31b3 | 3,056 | 99.9 |
| hh\_31b4 | 3,056 | 99.9 |
| hh\_31b5 | 3,056 | 99.9 |
| hh\_31b6 | 3,056 | 99.9 |
| hh\_31b7 | 3,056 | 99.9 |
| hh\_31b8 | 3,056 | 99.9 |
| hh\_31b9 | 3,056 | 99.9 |
| Source : base principale . | | |

[1] 807 3

Une analyse des valeurs manquantes révèle que 807 variables présentent plus de 70 % de données manquantes dans la base principale. Ces variables, majoritairement incomplètes, risquent de biaiser les résultats sans apporter une valeur ajoutée significative à l’analyse. Par conséquent, elles seront supprimées afin d’améliorer la qualité et la robustesse des analyses statistiques à venir.

Nous allons afficher les colonnes ou les valeurs manquantes dépassent 70% pour la base individus :

| Variable | Nb\_Valeurs\_Manquantes | Pourcentage |
| --- | --- | --- |
| hh\_00a\_year | 20,282 | 91.8 |
| hh\_00a\_month | 20,268 | 91.7 |
| hh\_00b\_year | 22,077 | 99.9 |
| hh\_00b\_month | 22,089 | 100.0 |
| hh\_11b | 22,025 | 99.7 |
| id\_01a | 17,885 | 81.0 |
| id\_01b | 15,815 | 71.6 |
| id\_02 | 15,851 | 71.7 |
| id\_03 | 16,137 | 73.0 |
| id\_04 | 21,852 | 98.9 |
| Source : base individu . | | |

Nous allons à présent supprimer les variables ou il ya 70% de na dans la base principale et dans la base individus

### Présence de doublons

[1] 0

Il ya 0 doublons dans la base principale

[1] 0

Il ya 0 doublons dans la base individus

Après ca nous allons supprimer les lignes et les colonnes entièrement vides dans les deux bases

Les données disponibles se composent de deux bases complémentaires : une base individuelle renseignant les caractéristiques sociodémographiques des personnes enquêtées, et une base ménage fournissant des informations sur leur environnement de vie. Ces deux bases sont reliées par un identifiant commun du ménage, ce qui permet d’associer chaque individu à son ménage et de croiser les informations pour une analyse plus approfondie.

## liens entre les deux bases

[1] 3058

On voit qu’il y a 3058 ID qui se répètent dans les deux bases ce qui correspond au nombre d’observations de la base principale .

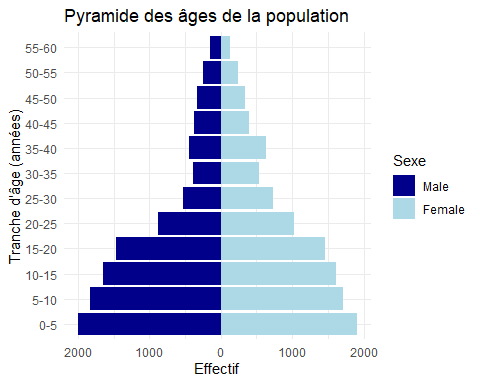
Nous allons voir les variables communes entre les deux bases

[1] "intro\_07\_1" "start" "end"   
 [4] "admin0" "admin1" "admin2"   
 [7] "admin3" "final\_01" "id"   
[10] "intro\_camp\_label" "wgh\_samp\_resc\_pop" "wgh\_samp\_resc\_str"

Les variables communes entre les deux bases sont : intro\_07\_1, start, end, admin0, admin1, admin2, admin3, final\_01, id, intro\_camp\_label, wgh\_samp\_resc\_pop, et wgh\_samp\_resc\_str. Ces variables, notamment l’identifiant unique id et les informations géographiques (admin0 à admin3), permettront de lier les informations relatives aux individus à celles des ménages pour une analyse croisée et cohérente des données.

# II. Analyse socio-economique des deplaces internes du Sudan en 2023

## Pyramide des ages des individus



La pyramide des âges présente une structure équilibrée entre hommes et femmes, avec une base légèrement plus large indiquant une population jeune. Les tranches d’âge actives (20-50 ans) sont bien représentées, suggérant un potentiel économique important. La proportion réduite des 65+ reflète une espérance de vie modérée. Cette distribution appelle des politiques ciblées sur l’emploi des jeunes et la couverture santé des seniors.

## Caracteristiques du chef de menage

| **Caractéristiques du chef de ménage** | **N** | **Refugees** N = 2,040*1* | **Host community North** N = 987*1* | **p-value***2* |
| --- | --- | --- | --- | --- |
| **ageYears** | 3,027 | 36 (1, 59) | 37 (2, 59) | 0.2 |
| **Sexe** | 3,027 |  |  | 0.4 |
| Male |  | 45% | 47% |  |
| Female |  | 55% | 53% |  |
| **What is [your/name's] present marital status?** | 3,011 |  |  | <0.001 |
| monogamous/married |  | 61% | 76% |  |
| polygamous/married |  | 11% | 12% |  |
| non-formal union |  | 0.6% | 0.3% |  |
| separated |  | 5.3% | 1.7% |  |
| divorced |  | 3.1% | 1.2% |  |
| widow or widower |  | 5.4% | 4.8% |  |
| never married |  | 13% | 3.8% |  |
| *1*Mean (Min, Max); % | | | | |
| *2*Wilcoxon rank sum test; Pearson's Chi-squared test | | | | |

## 3) Crowding Index ou l’indice d’affluence

### le nombre d’individu par menage

| **Characteristic** | **N = 3,058**1 |
| --- | --- |
| nombre\_personnes | 6.9 (3.1) | Médiane: 7.0 | Min: 0.0 | Max: 17.0 |
| 1Mean (SD) | Médiane: Median | Min: Min | Max: Max | |
| Source: Base principale. | |

Le tableau ci-dessus présente les statistiques descriptives de la variable nombre\_personnes (le nombre de personnes dans chaque ménage), avec un total de 3 058 ménages inclus dans l’analyse (N = 3,058).

Moyenne (écart-type) : La moyenne du nombre de personnes dans un ménage est de 6,9 avec un écart-type de 3,1. Cela indique une légère variation autour de la moyenne, avec certains ménages ayant beaucoup plus de personnes que la moyenne.  
  
Médiane : La médiane du nombre de personnes dans un ménage est de 7,0. Cela signifie que la moitié des ménages ont un nombre de personnes inférieur ou égal à 7, et l’autre moitié a un nombre supérieur ou égal à 7.  
  
Min - Max : Le nombre de personnes dans les ménages varie de 0 (un ménage sans personnes, potentiellement une valeur aberrante ou une erreur) à 17. Ce large écart montre qu’il existe des ménages de tailles variées.

Ces informations mettent en lumière une certaine hétérogénéité dans la taille des ménages observés.

| Nombre de pièces | **N = 3,058** |
| --- | --- |
| How many separate rooms do the members of your HH occupy? |  |
| Mean ± SD | 2.39 ± 1.32 |
| Median | 2.00 |
| Min - Max | 1.00 - 21.00 |
| Source: Base principale. | |

Le tableau ci-dessus présente les statistiques descriptives de la variable Nombre de pièces (nombre de pièces occupées par les membres du ménage), avec un total de 3 058 ménages inclus dans l’analyse (N = 3,058).

Moyenne (écart-type) : La moyenne du nombre de pièces par ménage est de 2,39 avec un écart-type de 1,32. Cela indique que, en moyenne, chaque ménage occupe environ 2 à 3 pièces, avec une variation notable (écart-type élevé) montrant des différences significatives entre les ménages.

Médiane : La médiane du nombre de pièces est de 2,00. Cela signifie que la moitié des ménages occupent 2 pièces ou moins, et l’autre moitié occupe 2 pièces ou plus.

Min - Max : Le nombre de pièces dans les ménages varie de 1 à 21, avec des valeurs extrêmes qui indiquent une large variabilité dans la taille des logements.

Ces statistiques suggèrent que la plupart des ménages occupent un petit nombre de pièces, bien que certains ménages occupent un nombre beaucoup plus élevé.

L’indice d’affluence (ou “crowding index”) est une mesure qui permet d’évaluer l’intensité de l’occupation d’un logement. Il est calculé en divisant le nombre d’individus d’un ménage par le nombre de pièces du logement, en excluant la cuisine et les couloirs. La formule de l’indice est la suivante :

Crowding Index = Nombre d’individus par ménage / Nombre de pièces du logement Interprétation : - Si l’indice est inférieur à 1, cela signifie qu’il y a moins d’individus que de pièces dans le logement, ce qui peut être un indicateur d’un logement relativement spacieux. - Si l’indice est égal à 1, cela signifie qu’il y a un nombre égal d’individus et de pièces, ce qui peut être considéré comme une occupation standard. - Si l’indice est supérieur à 1, cela indique qu’il y a plus d’individus que de pièces, signalant potentiellement un problème de surpeuplement dans le logement.

En calculant cet indice, on cherche à mieux comprendre la densité d’occupation des ménages. Une valeur plus élevée pourrait indiquer une surpopulation, tandis qu’une valeur plus basse pourrait signifier un logement plus spacieux.

### Calcul de l’indice d’affluence

nous allons calculer l’indicateur pour chaque ménage.

### La proportion de menage dont le crowding index est: <1, 1<= crowding\_index<2, 2<= crowding\_index<3 et >=3

| crowding\_cat | n | proportion |
| --- | --- | --- |
| 1-2 | 507 | 16.6 |
| 2-3 | 740 | 24.2 |
| <1 | 123 | 4.0 |
| >=3 | 1,682 | 55.1 |
| Source: Base principale. | | |

Le tableau montre la répartition des ménages selon l’indice d’affluence. Environ 4,0% des ménages ont un indice inférieur à 1, indiquant une faible densité de personnes par pièce. 16,6% des ménages ont un indice entre 1 et 2, tandis que 24,2% se situent entre 2 et 3, suggérant une densité moyenne. La majorité des ménages, soit 55,1%, ont un indice supérieur ou égal à 3, ce qui indique une forte densité et potentiellement des conditions de surpeuplement. Ces données peuvent éclairer les politiques sur le logement et l’aide aux populations vulnérables.

### statistiques suivant que le menage soit un refugie ou un member de la communaute d’acceuil (host community North)

| Population group | mean\_crowding\_index | median\_crowding\_index | sd\_crowding\_index | min\_crowding\_index | max\_crowding\_index | count |
| --- | --- | --- | --- | --- | --- | --- |
| Refugees | 3.292494 | 3 | 1.994662 | 0 | 14 | 2,064 |
| Host community North | 3.690677 | 3 | 2.217085 | 0 | 13 | 988 |
| Source: Base principale. | | | | | | |

Les statistiques sur l’indice d’affluence (crowding index) révèlent des différences notables entre les réfugiés et les membres de la communauté d’accueil du Nord. Pour les réfugiés, l’indice moyen est de 3,29, avec une médiane de 3, ce qui indique une densité relativement élevée de personnes par pièce, bien que l’écart-type de 1,99 montre une certaine variabilité. L’indice d’affluence varie de 0 à 14, avec un total de 2 064 ménages dans ce groupe.

Pour la communauté d’accueil du Nord, l’indice moyen est plus élevé à 3,69, avec une médiane de 3, et un écart-type de 2,22, suggérant également une densité élevée mais avec une plus grande dispersion. Les indices varient de 0 à 13, et il y a 988 ménages dans ce groupe.

Ces résultats montrent que les deux groupes vivent dans des conditions de densité relativement élevées, avec une légère différence entre les deux en termes de moyenne et de variabilité.

# Analyse de la securite alimentaire des deplaces internes

### Score de consommation alimentaire (SCA):

Score de consommation alimentaire (SCA)

| Score de consommation alimentaire (SCA) | Score de consommation alimentaire (SCA)1 |
| --- | --- |
| Days most hh members ate cereals, grains, roots and tubers Rice(Past 7days) |  |
| 0 | 674 (22%) |
| 1 | 450 (15%) |
| 2 | 252 (8.3%) |
| 3 | 365 (12%) |
| 4 | 230 (7.5%) |
| 5 | 297 (9.7%) |
| 6 | 92 (3.0%) |
| 7 | 689 (23%) |
| Unknown | 3 |
| Food\_div1: How was this food acquired? Main source of food (Past 7 days) |  |
| own production (crops, livestock, fishing / hunting, gathering) | 397 (17%) |
| against money on market | 1,441 (61%) |
| barter trade | 11 (0.5%) |
| aid from ngos / international organizations | 435 (18%) |
| aid from peers (family relatives, friends, neighbours) | 68 (2.9%) |
| other: specify | 19 (0.8%) |
| don't know | 2 (<0.1%) |
| refuse to answer | 2 (<0.1%) |
| Unknown | 677 |
| Food\_div1 : Quantity eaten was sufficient to cover the HH needs? (Past 7 d) |  |
| sufficient | 486 (20%) |
| insufficient | 1,874 (79%) |
| don't know | 12 (0.5%) |
| Unknown | 680 |
| Days most hh members ate ulses/ legumes / nuts: beans, cowpeas (Past 7days) |  |
| 0 | 1,896 (62%) |
| 1 | 316 (10%) |
| 2 | 276 (9.1%) |
| 3 | 228 (7.5%) |
| 4 | 89 (2.9%) |
| 5 | 86 (2.8%) |
| 6 | 43 (1.4%) |
| 7 | 112 (3.7%) |
| Unknown | 6 |
| Food\_div2: How was this food acquired? Main source of food (Past 7 days) |  |
| own production (crops, livestock, fishing / hunting, gathering) | 227 (20%) |
| against money on market | 690 (60%) |
| barter trade | 2 (0.2%) |
| aid from ngos / international organizations | 200 (17%) |
| aid from peers (family relatives, friends, neighbours) | 27 (2.3%) |
| other: specify | 4 (0.3%) |
| don't know | 1 (<0.1%) |
| refuse to answer | 1 (<0.1%) |
| Unknown | 1,900 |
| Food\_div2 : Quantity eaten was sufficient to cover the HH needs? (Past 7 d) |  |
| sufficient | 317 (28%) |
| insufficient | 831 (72%) |
| don't know | 4 (0.3%) |
| Unknown | 1,900 |
| Days most hh members ate milk and other dairy products...(Past 7days) |  |
| 0 | 2,474 (81%) |
| 1 | 227 (7.5%) |
| 2 | 113 (3.7%) |
| 3 | 79 (2.6%) |
| 4 | 27 (0.9%) |
| 5 | 38 (1.2%) |
| 6 | 15 (0.5%) |
| 7 | 72 (2.4%) |
| Unknown | 7 |
| Days most hh members ate meat, fish and eggs...(Past 7days) |  |
| 0 | 2,176 (72%) |
| 1 | 397 (13%) |
| 2 | 230 (7.6%) |
| 3 | 122 (4.0%) |
| 4 | 34 (1.1%) |
| 5 | 31 (1.0%) |
| 6 | 12 (0.4%) |
| 7 | 40 (1.3%) |
| Unknown | 10 |
| Days most hh members ate vegetables and leaves: spinach, onion (Past 7days) |  |
| 0 | 2,097 (69%) |
| 1 | 244 (8.0%) |
| 2 | 243 (8.0%) |
| 3 | 175 (5.8%) |
| 4 | 91 (3.0%) |
| 5 | 64 (2.1%) |
| 6 | 14 (0.5%) |
| 7 | 111 (3.7%) |
| Unknown | 13 |
| Food\_div5: How was this food acquired? Main source of food (Past 7 days) |  |
| own production (crops, livestock, fishing / hunting, gathering) | 222 (23%) |
| against money on market | 665 (70%) |
| barter trade | 5 (0.5%) |
| aid from ngos / international organizations | 8 (0.8%) |
| aid from peers (family relatives, friends, neighbours) | 10 (1.1%) |
| other: specify | 33 (3.5%) |
| don't know | 2 (0.2%) |
| refuse to answer | 1 (0.1%) |
| Unknown | 2,106 |
| Food\_div5 : Quantity eaten was sufficient to cover the HH needs? (Past 7 d) |  |
| sufficient | 262 (28%) |
| insufficient | 673 (71%) |
| don't know | 11 (1.2%) |
| Unknown | 2,106 |
| Days most hh members ate orange vegetables: carrot, red pepper (Past 7days) |  |
| 0 | 717 (76%) |
| 1 | 50 (5.3%) |
| 2 | 63 (6.7%) |
| 3 | 31 (3.3%) |
| 4 | 31 (3.3%) |
| 5 | 32 (3.4%) |
| 6 | 4 (0.4%) |
| 7 | 14 (1.5%) |
| Unknown | 2,110 |
| Days most hh members ate green leafy vegetables: spinach,… (Past 7days) |  |
| 0 | 405 (43%) |
| 1 | 41 (4.4%) |
| 2 | 88 (9.3%) |
| 3 | 121 (13%) |
| 4 | 97 (10%) |
| 5 | 97 (10%) |
| 6 | 32 (3.4%) |
| 7 | 61 (6.5%) |
| Unknown | 2,110 |
| Days most hh members ate fruits (Past7 days) |  |
| 0 | 2,773 (91%) |
| 1 | 55 (1.8%) |
| 2 | 55 (1.8%) |
| 3 | 47 (1.5%) |
| 4 | 25 (0.8%) |
| 5 | 38 (1.3%) |
| 6 | 3 (<0.1%) |
| 7 | 42 (1.4%) |
| Unknown | 14 |
| Days most hh members ate oil/fat/butter: vegetable oil, palm oil(Past 7days) |  |
| 0 | 2,141 (70%) |
| 1 | 152 (5.0%) |
| 2 | 198 (6.5%) |
| 3 | 199 (6.5%) |
| 4 | 109 (3.6%) |
| 5 | 70 (2.3%) |
| 6 | 40 (1.3%) |
| 7 | 133 (4.4%) |
| Unknown | 10 |
| Days most hh members ate sugar or sweet: sugar, honey, jam, cakes(Past 7days) |  |
| 0 | 2,277 (75%) |
| 1 | 193 (6.4%) |
| 2 | 161 (5.3%) |
| 3 | 129 (4.2%) |
| 4 | 73 (2.4%) |
| 5 | 66 (2.2%) |
| 6 | 33 (1.1%) |
| 7 | 105 (3.5%) |
| Unknown | 15 |
| Days most hh members ate condiments: tea, coffee/cocoa, salt...(Past 7days) |  |
| 0 | 2,238 (74%) |
| 1 | 131 (4.3%) |
| 2 | 173 (5.7%) |
| 3 | 173 (5.7%) |
| 4 | 70 (2.3%) |
| 5 | 82 (2.7%) |
| 6 | 36 (1.2%) |
| 7 | 138 (4.5%) |
| Unknown | 11 |
| 1n (%) | |
| Source: Base princpale | |

Les données révèlent une insécurité alimentaire aiguë au sein des ménages enquêtés. Près de 84 % ont manqué de nourriture ou d’argent pour se nourrir au cours des 7 derniers jours, et 89 % ont exprimé une inquiétude à ce sujet. Une majorité significative a été contrainte d’adopter des stratégies de survie : 83 % ont réduit la diversité ou la qualité de leur alimentation, 85 % ont mangé moins que nécessaire, et 70 % des adultes ont sacrifié leurs repas au profit des enfants. Le recours à l’aide extérieure ou à des moyens extrêmes comme emprunter, vendre des biens ou mendier reste fréquent. En somme, les indicateurs pointent vers une crise alimentaire grave, marquée par des privations fréquentes et des mécanismes d’adaptation à court terme, souvent non durables.

### score de consommation alimentaire

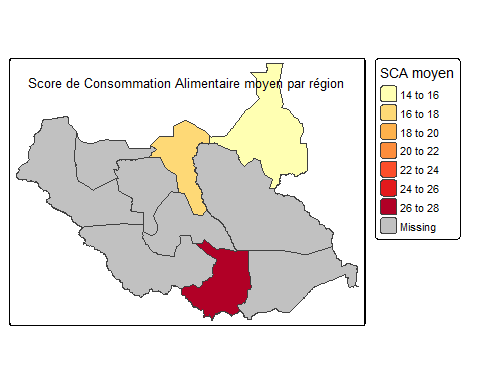
| **Groupe.alimentaire** | **Variable** | **Poids** |
| --- | --- | --- |
| Céréales et tubercules | food\_div1 | 2.0 |
| Légumineuses | food\_div2 | 3.0 |
| Lait et produits laitiers | food\_div3 | 4.0 |
| Viande, poisson et œufs | food\_div4 | 4.0 |
| Légumes | food\_div5 | 1.0 |
| Fruits | food\_div6 | 1.0 |
| Fruits riches en vitamine A | food\_div61 | 1.0 |
| Huiles et graisses | food\_div7 | 0.5 |
| Sucres | food\_div8 | 0.5 |
| Condiments/épices | food\_div9 | 0.0 |

Nous allons calculer le SCA en fonction de ca .

### Tableau des poids (total = 16)

| Groupe | Variable | Poids |
| --- | --- | --- |
| Céréales | Food\_div1 | 2.0 |
| Légumineuses | Food\_div2 | 3.0 |
| Légumes | Food\_div3 | 1.0 |
| Fruits | Food\_div4 | 1.0 |
| Viande/Poisson/Œufs | Food\_div5 | 4.0 |
| Produits laitiers | Food\_div6 | 4.0 |
| Tubercules | Food\_div61 | 0.5 |
| Sucre | Food\_div7 | 0.5 |
| Matières grasses | Food\_div8 | 0.5 |
| Épices/condiments | Food\_div9 | 0.0 |

### Répresentation spatiale (région et département) du SCA et de ses différentes catégorisations.



Le document présente le SCA moyen par région, classé en intervalles allant de 14 à 28, avec une catégorie “Missing” pour les données manquantes.

Un SCA plus élevé indique une meilleure consommation alimentaire :  
  
 14–20 : Consommation faible à limite, reflétant une insécurité alimentaire possible.  
  
 20–28 : Consommation acceptable à bonne, suggérant une situation alimentaire plus stable.  
  
 Missing : Données non disponibles, nécessitant une collecte supplémentaire pour une analyse complète.

## L’indice réduit des stratégies de survie (rCSI):

### Statistiques des variables rcsi

Nous allons commencer par faire une analyse descriptive des variables composant le rCSI, qui peuvent être binaires (avec les modalités “yes” et “no”). Ensuite, nous générerons des variables numériques selon les conditions spécifiées (si la modalité est “yes”, attribuer une valeur aléatoire entre 1 et 7, sinon 0).

| Statistiques sur les variables RCSI | **N = 3,052**1 |
| --- | --- |
| A member ate less preferred food and less expensive food [Past 30/7 days] | 2,488 (82%) |
| Valeurs Manquantes | 1 |
| Any member borrowed food/relied on help to get more food...[Past 30/7 days] | 1,827 (60%) |
| Valeurs Manquantes | 5 |
| Any member ate a smaller meal than you felt you needed [Past 30/7 days] | 2,594 (85%) |
| Valeurs Manquantes | 4 |
| Adults ate less to have more food for children under 5? [Past 30/7 days] | 2,140 (70%) |
| Valeurs Manquantes | 6 |
| Any member ate fewer meals in a day [Past 30/7 days] | 2,541 (83%) |
| Valeurs Manquantes | 1 |
| 1n (%) | |
| Source: Base principale | |

Les données montrent que 82% des membres ont consommé des aliments moins préférés et moins chers au cours des 30 derniers jours, 60% ont emprunté de la nourriture ou reçu de l’aide pour en obtenir davantage, et 85% ont mangé des repas plus petits que ce qu’ils jugeaient nécessaire. De plus, 70% des adultes ont sacrifié leur alimentation pour garantir plus de nourriture aux enfants de moins de 5 ans, et 83% ont réduit le nombre de repas quotidiens. Les valeurs manquantes sont faibles, avec un maximum de 7 pour certaines variables, indiquant des données relativement complètes. Ces résultats reflètent des comportements alimentaires contraints et un stress alimentaire au sein des ménages.

On s’interesse maintenant aux valeurs numeriques comprisent entre 0 et 7. Pour chacune des variables qui composent le rCSI, creez des nouvelles variables pour chacune des variables du rCSI selon les conditions suivantes: Si la modalite de la variable est yes on génère aleatoirement une valeur comprise entre 1 et 7 attribue a la variable, dans le cas contraire la nouvelle variable prend 0.

### Indice réduit des stratégies de survie en utilisant les Nouvelles variables

### statistiques descriptives sur les Nouvelles variables

| statistiques descriptives sur les nouvelles variables | **N = 3,052**1 |
| --- | --- |
| food02a\_score |  |
| 0 | 563 (18%) |
| 1 | 360 (12%) |
| 2 | 350 (11%) |
| 3 | 365 (12%) |
| 4 | 336 (11%) |
| 5 | 350 (11%) |
| 6 | 349 (11%) |
| 7 | 378 (12%) |
| Valeurs manquantes | 1 |
| food05a\_score |  |
| 0 | 1,220 (40%) |
| 1 | 259 (8.5%) |
| 2 | 306 (10%) |
| 3 | 243 (8.0%) |
| 4 | 258 (8.5%) |
| 5 | 260 (8.5%) |
| 6 | 258 (8.5%) |
| 7 | 243 (8.0%) |
| Valeurs manquantes | 5 |
| food06a\_score |  |
| 0 | 454 (15%) |
| 1 | 350 (11%) |
| 2 | 345 (11%) |
| 3 | 377 (12%) |
| 4 | 395 (13%) |
| 5 | 362 (12%) |
| 6 | 351 (12%) |
| 7 | 414 (14%) |
| Valeurs manquantes | 4 |
| food07a\_score |  |
| 0 | 906 (30%) |
| 1 | 328 (11%) |
| 2 | 303 (9.9%) |
| 3 | 319 (10%) |
| 4 | 249 (8.2%) |
| 5 | 339 (11%) |
| 6 | 324 (11%) |
| 7 | 278 (9.1%) |
| Valeurs manquantes | 6 |
| food08a\_score |  |
| 0 | 510 (17%) |
| 1 | 374 (12%) |
| 2 | 377 (12%) |
| 3 | 352 (12%) |
| 4 | 335 (11%) |
| 5 | 352 (12%) |
| 6 | 371 (12%) |
| 7 | 380 (12%) |
| Valeurs manquantes | 1 |
| 1n (%) | |
| Source: Calculs de l'auteur basés sur les données de l'enquête. | |

### l’indice rCSI (pondéré)

Les poids standard sont généralement :

food02a (5), food05a (4), food06a (3), food07a (2), food08a (7)  
  
Total des poids = 21

### Tableau des poids

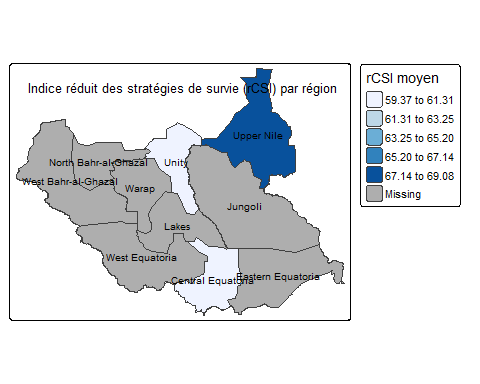
Poids attribués aux stratégies de survie (rCSI)

L’indice réduit des stratégies de survie (rCSI) repose sur cinq comportements alimentaires fréquemment adoptés par les ménages confrontés à une insécurité alimentaire. À chacune de ces stratégies est attribué un poids reflétant sa gravité ou son impact potentiel sur la sécurité alimentaire. Le tableau ci-dessous présente les poids utilisés dans le calcul du rCSI. La somme totale des poids est égale à 21, conformément à la méthodologie standard :

Poids attribués aux stratégies de survie (rCSI)

| Variable | Poids |
| --- | --- |
| food02a | 5 |
| food05a | 4 |
| food06a | 3 |
| food07a | 2 |
| food08a | 7 |

### Répresentation spatiale selon le milieu de residence du rCSI



Ce graphique présente l’Indice réduit des stratégies de survie (rCSI) moyen, classé en intervalles de 59.37 à 69.08, ainsi que huit régions (ex. North Bahr-al-Ghazal, Unity). Un rCSI élevé indique une insécurité alimentaire modérée à sévère, avec des ménages recourant fréquemment à des stratégies de survie (ex. sauter des repas). Les scores les plus critiques (65.20–69.08) reflètent des besoins urgents, tandis que les données manquantes limitent l’analyse complète. Action recommandée : Prioriser les régions aux scores >65 et améliorer la collecte de données pour les zones sans information.

# conclusion

Ce projet d’analyse statistique, réalisé sous R, a permis d’explorer en profondeur les conditions socio-économiques et la sécurité alimentaire des déplacés internes au Soudan du Sud en 2023. Les résultats révèlent des défis majeurs :

Conditions de vie précaires :

L’indice d’affluence moyen élevé (3,29 pour les réfugiés, 3,69 pour les communautés hôtes) témoigne d’un surpeuplement critique, avec 55% des ménages ayant un indice ≥3.

La taille moyenne des ménages (6,9 personnes) et le nombre limité de pièces (2,4 en moyenne) exacerbent ces conditions.

Insécurité alimentaire aiguë :

Le Score de Consommation Alimentaire (SCA) montre que 84% des ménages ont réduit leur diversité alimentaire, et 85% ont consommé des portions insuffisantes.

L’indice rCSI (59-69) confirme des stratégies de survie extrêmes (70% des adultes sautant des repas pour les enfants).

Disparités géographiques :

Les représentations spatiales mettent en lumière des régions critiques (ex : scores rCSI >65) nécessitant une intervention urgente.

Les données manquantes (“Missing”) limitent l’analyse pour certaines zones, soulignant le besoin d’améliorer la collecte.

Recommandations :

Prioriser les zones à SCA <20 et rCSI >65 pour l’aide alimentaire.

Améliorer les conditions de logement via des programmes ciblés.

Renforcer les systèmes de suivi pour combler les lacunes de données.

Méthodologie : L’utilisation de R Markdown a assuré une analyse reproductible, intégrant tests statistiques (Chi², Wilcoxon) et visualisations (pyramides des âges, cartes). Les codes commentés et la gestion rigoureuse des valeurs manquantes ont renforcé la robustesse des résultats.

Perspective : Croiser ces indicateurs avec des données climatiques ou de conflits pourrait affiner l’analyse des causes profondes des déplacements.

SOMA Ben Idriss

Élève Ingénieur Statisticien Économiste

# Remerciements

À M. Aboubacar Hema pour son encadrement