|  |
| --- |
| **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** |
| Paul BALAFAI | M. Aboubacar HEMA |
| *Élève Ingénieur Statisticien Économiste* | *Data-scientist* |

|  |
| --- |
| **Année académique 2024-2025** |

###### Sommaire

N’est pas généré automatiquement donc le faire directement avec word.

###### Liste des figures

N’est pas généré automatiquement donc le faire directement avec word.

###### Liste des tableaux

N’est pas généré automatiquement donc le faire directement avec word.

###### Introduction

Le Sud-Soudan, plus jeune nation du monde ayant accédé à l’indépendance en 2011, demeure confronté à une crise humanitaire parmi les plus graves à l’échelle mondiale. L’année 2023 marque une période particulièrement critique dans l’histoire des déplacements forcés de population dans ce pays, où les dynamiques complexes de violence, d’instabilité politique, de catastrophes naturelles et d’insécurité alimentaire continuent d’engendrer des mouvements massifs de populations.  
Cette analyse se propose d’examiner en profondeur les données recueillies lors de l’Enquête sur les déplacements forcés au Sud-Soudan en 2023, afin de dresser un portrait précis de cette crise humanitaire. Notre étude vise à comprendre les causes profondes, l’ampleur et les caractéristiques des déplacements, ainsi que leurs conséquences sur les populations affectées. Le contexte sud-soudanais reste marqué par les séquelles d’une guerre civile dévastatrice (2013-2018) et par des épisodes récurrents de violence intercommunautaire. Ces facteurs, combinés à des inondations catastrophiques, à une sécheresse persistante dans certaines régions et à une inflation galopante, ont créé un environnement propice aux déplacements forcés à grande échelle. En 2023, cette situation s’est encore détériorée avec l’éclatement de nouveaux conflits dans certaines régions et l’aggravation des conditions climatiques extrêmes.  
Notre analyse se fondera sur des données quantitatives et qualitatives collectées auprès des populations déplacées, des communautés d’accueil et des acteurs humanitaires intervenant dans les différents États du Sud-Soudan. Elle explorera les tendances géographiques des déplacements, les profils démographiques des personnes déplacées, les besoins prioritaires identifiés et les mécanismes d’adaptation développés par ces populations vulnérables.  
À travers cette étude, nous espérons contribuer à une meilleure compréhension de la situation actuelle, afin d’informer les décisions politiques et les interventions humanitaires visant à protéger et assister les personnes déplacées au Sud-Soudan. L’analyse des données de 2023 révélera également l’évolution de cette crise par rapport aux années précédentes, permettant ainsi d’identifier les tendances émergentes et les défis à venir dans la gestion des déplacements forcés dans ce pays.

# I. Analyse de la consistance des bases de données

Dans cette section, nous présentons successivement :

* **La dimension des jeux de données**
* **La cohérence des données entre les différentes bases**
* **La visualisation des premières observations**
* **La détection et suppression des doublons**
* **L’analyse des valeurs manquantes**
* **La suppression des colonnes entièrement vides**

## Dimension des bases de données

La base principale contient 3058 observations et 1312 variables.

La base individus contient 22092 observations et 144 variables.

## Vérification de la cohérence des liens entre bases

Nombre d'individus sans correspondance dans la base principale : 0

Nombre d'ID dans la base principale sans individu associé : 0

Le d’individus sans correspondance dans la base principale est nul, cela signifie que tous les individus de la base individus appartiennent à un ID présent dans la base principale. Ce qui montre que les bases sont cohérentes.

## Affichage des 8 premières lignes

Nous allons visualiser les premières lignes de la base base\_individus pour voir leur structure.

Premières observations de la base base\_individus

| Population group | interview's start day and time | interview's end day and time | Host country | State | County | Payam | Final outcome of the interview |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Refugees | 2023-04-29 | 2023-04-29 | SSD | Central Equatoria | Juba | Not applicable | Complete Interview |
| Refugees | 2023-04-29 | 2023-04-29 | SSD | Central Equatoria | Juba | Not applicable | Complete Interview |
| Refugees | 2023-04-29 | 2023-04-29 | SSD | Central Equatoria | Juba | Not applicable | Complete Interview |
| Refugees | 2023-04-29 | 2023-04-29 | SSD | Central Equatoria | Juba | Not applicable | Complete Interview |
| Refugees | 2023-04-29 | 2023-04-29 | SSD | Central Equatoria | Juba | Not applicable | Complete Interview |
| Refugees | 2023-04-29 | 2023-04-29 | SSD | Central Equatoria | Juba | Not applicable | Complete Interview |

ous allons également visualiser les premières lignes de la base base\_principale pour voir leur structure.

Premières observations de la base base\_individus

| Population group | interview's start day and time | interview's end day and time | Host country | State | County | Payam | Final outcome of the interview |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Refugees | 2023-04-29 | 2023-04-29 | SSD | Central Equatoria | Juba | Not applicable | Complete Interview |
| Refugees | 2023-07-21 | 2023-07-21 | SSD | Upper Nile | Maban | Jinkuata | Complete Interview |
| Host community North | 2023-05-22 | 2023-05-22 | SSD | Upper Nile | Maban | Boung | Complete Interview |
| Host community North | 2023-07-26 | 2023-07-26 | SSD | Upper Nile | Maban | Boung | Complete Interview |
| Host community North | 2023-05-29 | 2023-05-29 | SSD | Unity | Pariang | Aliiny | Complete Interview |
| Host community North | 2023-05-12 | 2023-05-12 | SSD | Upper Nile | Maban | Banashowa | Complete Interview |

## Présence de doublons dans les bases de données

Commencons l’analyse des doublons par la base base\_individus.

La base base\_individus contient 0 doublons.

La base base\_individus contient 0 doublons.

La base principale et la base individus ne contiennet aucun doublon.

## Présence de valeurs manquantes

Pourcentage de valeurs manquantes par variable dans la base base\_individus.

Table 2 – Variables avec valeurs manquantes dans base\_individus

| \*\*Variable\*\* | \*\*N missings\*\* | \*\*% Missing\*\* |
| --- | --- | --- |
| start | 6 | 0.0% |
| admin1 | 896 | 4.1% |
| admin2 | 896 | 4.1% |
| admin3 | 896 | 4.1% |
| HH\_03 | 30 | 0.1% |
| HH\_02 | 20 | 0.1% |
| ageYears | 295 | 1.3% |
| HH\_00 | 10 | 0.0% |
| HH\_00a\_year | 20,282 | 91.8% |
| HH\_00a\_month | 20,268 | 91.7% |

A present Visualisonsles valeurs manquantes dans la base base\_principale.

Table 1 – Variables avec valeurs manquantes dans base\_principale

| \*\*Variable\*\* | \*\*N missing\*\* | \*\*% missing\*\* |
| --- | --- | --- |
| start | 1 | 0.0% |
| admin1 | 94 | 3.1% |
| admin2 | 94 | 3.1% |
| admin3 | 94 | 3.1% |
| Intro\_camp\_label | 94 | 3.1% |
| Intro\_08 | 4 | 0.1% |
| Intro\_09 | 2 | 0.1% |
| origincountry | 77 | 2.5% |
| HH\_Lang01 | 200 | 6.5% |
| HH\_Lang02a | 399 | 13.0% |

## Suppression des colonnes vides

Grace à la fonction remove\_empty de Janitor, nous allons supprimer les colonnes vides.

Après application, la nouvelle dimension de la base base\_principale est de:

[1] 0 68

Ainsi nous notons que 68 colonnes ont été supprimées dans la base principale.

Faisons de meme pour la base\_individus:

La nouvelle dimension de la base est de:

[1] 0 0

Ici aussi, aucune colonne n’a été supprimée.

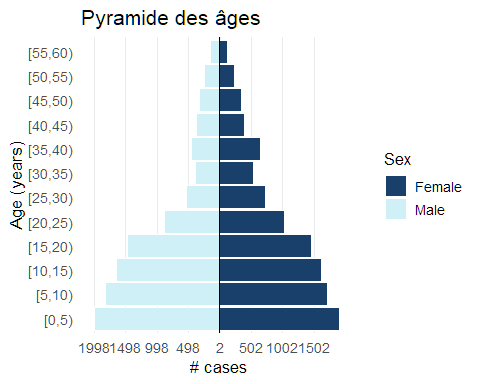
## nettoyage du nom des variables

# II. Analyse des données et calcul d’indicateurs

Cette section couvre l’analyse socio-démographique des ménages ainsi que le calcul d’indicateurs.

### Pyramide des âges

Pour mieux visualiser la structure démographique de la population étudiée, nous allons créer une pyramide des âges en utilisant les variables age\_years et hh\_02 de la base individus.



Pyramide des âges de la population étudiée

La pyramide des âges ci-dessus permet d’observer la distribution par âge et par sexe de la population étudiée, avec les hommes à gauche (bleu clair) et les femmes à droite (bleu foncé), comme dans l’exemple fourni.

### Caracteristiques du chef de menage

| id | n\_individu |
| --- | --- |
| 1 | 15 |
| 1 | 15 |
| 1 | 15 |
| 1 | 15 |
| 1 | 15 |
| 1 | 15 |

| id | n\_individu |
| --- | --- |
| 3077 | 4 |
| 3077 | 4 |
| 3077 | 4 |
| 3077 | 4 |
| 3078 | 2 |
| 3078 | 2 |

#### statistiques descriptives sur la variable

Statistiques descriptives du nombre d'individus par ménage

| Moyenne | Mediane | Min | Max | Ecart\_type |
| --- | --- | --- | --- | --- |
| 8.623484 | 8 | 1 | 17 | 3.208775 |

En moyenne, les menages sont constitués de 8 individus. On note egalement que la moitié des ménages compte plus de 8 individus, le menage le plus grand comporte 17 individus.

Diviser la colonne en deux partie

# III Analyse de la securite alimentaire des deplaces internes

Cette consiste en l’analyse de la securite alimentaire afin de mesurer l’impact du conflit sur la securite alimentaire des deplaces internes

#### Calcul du score de consommation alimentaire

Le score de consommation alimentaire (SCA) est un score composite basé sur la diversité alimentaire des ménages, la fréquence des repas et l’importance nutritionnelle relative des aliments consommés. Le module SCA demande aux répondants combien de jours, au cours des 7 derniers jours, la plupart des membres de leur ménage (+50%) ont consommé chacun des produits/groupes alimentaires, à l’intérieur ou à l’extérieur de leur domicile.

#### analyse descriptives des variables du SCA

Mettre la page en paysage

suite texte suite texte suite texte suite texte suite texte suite texte suite texte suite texte suite texte suite texte suite texte, note suite texte suite texte suite texte suite texte suite texte suite texte suite texte suite texte suite texte suite texte suite texte suite texte

# Table des matières

###### Conclusion