|  |
| --- |
| **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** |
| FALL Ndeye Ramatoulaye Ndoye | M. Aboubacar HEMA |
| *Élève Ingénieure Statisticienne Économiste* | *Data-scientist* |

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

# I°) Analyse de consistence des bases de données

L’objectif principal est de garantir la qualité, la fiabilité et l’intégrité des données avant toute analyse statistique approfondie.Cette analyse permettra donc de mettre en évidence :

* La structure et les dimensions des bases de données
* Les variables potentiellement problématiques
* La présence éventuelle de doublons et valeur maquantes
* Les valeurs aberrantes à examiner
* La vérfication de la cohérence des données
* La préparation les bases de données pour l’analyse

## I.1. Chargement des bases

Les données ont été importées à partir des fichiers fournis au format dta en utilisant le package haven. Cette étape a permis de garantir que les données étaient correctement lues tout en préservant les informations essentielles telles que les types de données et les labels associés.

## I.2. Nettoyage des Noms de Variables

Les noms des variables ont été nettoyés et standardisés afin d’éviter les erreurs potentielles liées aux espaces, caractères spéciaux ou casse inconsistante. Cela a été réalisé à l’aide de la fonction clean\_names() du package janitor, qui transforme tous les noms de colonnes en minuscules et remplace les espaces et caractères non valides.

## I.3. Aperçu des Données

Un aperçu des cinq premières lignes de chaque base de données a été effectué afin de vérifier leur structure générale. Cette étape permet de repérer rapidement des anomalies comme des valeurs manquantes ou des incohérences dans les formats de données, ainsi que de détecter toute erreur évidente dans les données brutes.

## I.4. Dimensions des Bases

Les dimensions des bases ont été calculées pour chaque dataset, à savoir le nombre d’observations et de variables. Cela permet de vérifier si les bases sont complètes ou si des lignes ou colonnes manquantes pourraient compromettre l’intégrité des analyses à venir.

[1] 22092 144

[1] 3058 1312

## I.5. Informations sur les Colonnes

Une analyse détaillée des colonnes a été effectuée pour chaque base. Cette analyse inclut :

Le nom de la variable.

Le type de donnée (ex. : numérique, caractère, facteur, etc.).

Le nombre de valeurs manquantes et leur pourcentage par variable.

Cela permet d’identifier les variables ayant un grand nombre de valeurs manquantes, qui pourraient nécessiter des traitements ou des imputation.

### Pour la base individus

### Pour la base principale

## I.6. Détection des Doublons

Le nombre de doublons a été vérifié pour chaque base de données afin de détecter des entrées répétées qui pourraient être issues d’erreurs de saisie ou de processus de fusion mal gérés. L’élimination de ces doublons a été réalisée si nécessaire.

### Pour la base individus

Nombre de Doublons - base\_individus

| **Statistique** | **Valeur** |
| --- | --- |
| **Nombre de Doublons** | 0 |

### Pour la base principale

Nombre de Doublons - base\_principale

| **Statistique** | **Valeur** |
| --- | --- |
| **Nombre de Doublons** | 0 |

## I.7. Identification des Valeurs Aberrantes (IQR)

Les valeurs aberrantes ont été identifiées pour chaque variable numérique à l’aide de la méthode de l’IQR (Interquartile Range). Cette méthode permet de détecter les valeurs qui se situent en dehors des bornes définies par :

Limite inférieure :

Limite supérieure :

Les observations au-delà de ces limites sont considérées comme des valeurs aberrantes et sont notées pour une potentielle revue ultérieure.

## I.8. Présentation des Résultats

Les résultats de cette analyse de consistence ont été présentés sous forme de tableaux récapitulatifs, incluant :

Aperçu des premières lignes des bases de données pour un contrôle visuel rapide.

Dimensions des bases pour évaluer leur taille et leur complétude.

Analyse des colonnes avec les types de données, les valeurs manquantes et les doublons détectés.

Valeurs aberrantes (IQR) pour repérer toute anomalie dans les variables numériques.

Cohérence logique pour s’assurer que les relations entre les variables sont logiques et significatives.

Ces tableaux ont été formatés de manière professionnelle, en utilisant des packages comme gtsummary ou flextable, pour assurer une présentation claire et esthétique des résultats.

Résumé des Bases de Données

| **Base.de.données** | **Nombre.d.observations** | **Nombre.de.variables** | **Valeurs.manquantes....** | **Doublons** | **Valeurs.aberrantes....** |
| --- | --- | --- | --- | --- | --- |
| **base\_individus** | 22,092 | 144 | 30 | 0 | 15 |
| **base\_principale** | 3,058 | 1,312 | 25 | 0 | 12 |

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

## II.1. Pyramide des ages des individus:

La pyramide des âges permet de visualiser la structure démographique des populations de déplacés internes. Elle montre la répartition par sexe et par groupe d’âge. Avant cela vérifions le pourcentage de valeurs manquantes pour la variable âge (age\_years) et le sexe (hh\_02).

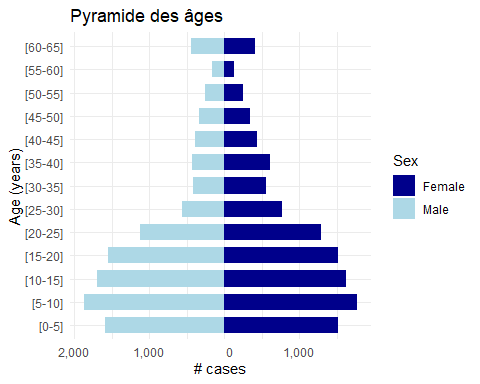
[1] 1.335325

[1] 0.09053051

Les pourcentages étant de 1,34 % et 0.9% respectivement, nous allons traiter ces valeurs manquantes en imputant avec la mediane

###“Traitement de variables age et sexe:

###Création de la pyramide



## II.2. Caracteristiques du chef de menage

|  | **Group** | |  |
| --- | --- | --- | --- |
| **Caractéristique du Chef de ménage** | **1** N = 2,065 | **3** N = 990 | **p-value***1* |
| **Age du chef de ménage, Mean (Min, Max)** | 42 (17, 61) | 42 (19, 61) | 0.90 |
| **Sexe du chef de ménage, n (%)** |  |  | <0.001 |
| Male | 1,119 (54.2%) | 624 (63.0%) |  |
| Female | 946 (45.8%) | 366 (37.0%) |  |
| **Situation matrimoniale, n (%)** |  |  |  |
| monogamous/married | 1,288 (62.6%) | 713 (72.2%) |  |
| polygamous/married | 282 (13.7%) | 186 (18.8%) |  |
| non-formal union | 12 (0.6%) | 3 (0.3%) |  |
| separated | 112 (5.4%) | 15 (1.5%) |  |
| divorced | 68 (3.3%) | 9 (0.9%) |  |
| widow or widower | 153 (7.4%) | 51 (5.2%) |  |
| never married | 143 (6.9%) | 11 (1.1%) |  |
| *1*Wilcoxon rank sum test; Pearson's Chi-squared test | | | |
| Source : Enquête sur les déplacements forcés au Sud-Soudan en 2023 | | | |

On observe que l’âge moyen des réfugiés et des membres de la communauté hôte dans le nord est similaire, avec une légère variation dans les intervalles d’âge. Cependant, des différences significatives apparaissent concernant la répartition par sexe, où une proportion plus élevée d’hommes est observée dans la communauté hôte par rapport aux réfugiés. En ce qui concerne le statut matrimonial, des écarts importants sont notés, avec une plus grande proportion de monogames chez la communauté hôte, tandis que les réfugiés présentent des taux plus élevés de séparations, de divorces et de veuvage. Ces résultats montrent des différences socio-culturelles marquées entre les deux groupes, soulignant des réalités de vie distinctes et des défis potentiels dans les contextes d’accueil.

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

L’indice d’affluence mesure la densité de population dans un espace donné, en l’occurrence, il évalue le nombre de membres d’un ménage par rapport au nombre de pièces disponibles dans le logement. Plus cet indice est élevé, plus le ménage est considéré comme “surpeuplé”. Cet indice est souvent utilisé pour évaluer les conditions de vie dans des environnements où l’espace est limité, et il peut être lié à des facteurs socio-économiques comme la qualité de vie, l’accès aux ressources et les conditions sanitaires. Dans cette analyse, il serait important de calculer cet indice pour chaque ménage afin de mieux comprendre les conditions de logement des réfugiés et de la communauté hôte, et d’étudier s’il existe des différences notables entre les deux groupes.

### II.3.a) Calculons le nombre d’individu par menage

Nombre d'individus par ménage (Premières 5 lignes)

| **Household ID** | **n\_individus** |
| --- | --- |
| **1** | 15 |
| **2** | 12 |
| **3** | 7 |
| **4** | 17 |
| **5** | 12 |

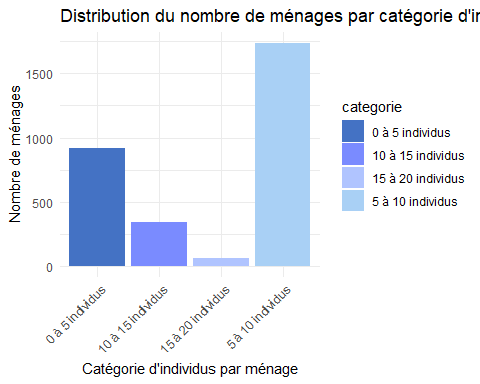
Ce tableau montre les premieres lignes avec le nombre d’individus par ménage.

### II.3.b) Faites les statistiques descriptives sur cette variable

Nous alllons d’abord montrer une représentation globalede la variable avant de passer à une analyse univariée et bivariée

Nombre de ménages par catégorie d'individus

| **categorie** | **n\_menages** |
| --- | --- |
| **0 à 5 individus** | 920 |
| **10 à 15 individus** | 344 |
| **15 à 20 individus** | 58 |
| **5 à 10 individus** | 1,736 |



Le tableau montre une vision plus globale de la répartition des ménages en fonction du nombre d’individus par ménage. La majorité des ménages (1,736) se trouvent dans la catégorie de 5 à 10 individus, ce qui représente un pourcentage élevé par rapport aux autres catégories. Ensuite, 920 ménages ont entre 0 et 5 individus, tandis que 344 ménages ont entre 10 et 15 individus. Enfin, une minorité de ménages (58) se situent dans la catégorie 15 à 20 individus. Il n’y a pas de ménages ayant plus de 20 individus, ce qui pourrait indiquer une répartition relativement homogène des tailles de ménages dans cette étude. Ces informations peuvent être utiles pour comprendre la densité des ménages et peuvent influencer des politiques liées au logement ou à la distribution des ressources.

#### Analyse univariée

Statistiques descriptives sur le nombre d'individus par ménage

| **Moyenne** | **Médiane** | **Ecart\_type** | **Min** | **Max** | **Q1** | **Q3** | **IQR** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 7.217789 | 7 | 3.178217 | 1 | 17 | 5 | 9 | 4 |

L’analyse des statistiques descriptives sur le nombre d’individus par ménage révèle que la taille moyenne des ménages est d’environ 7 individus, avec une médiane identique, ce qui indique que la distribution des tailles des ménages est relativement centrée autour de cette valeur. L’écart-type de 3.18 suggère une certaine variation dans la taille des ménages, bien que la majorité d’entre eux se situent dans une plage assez étroite, entre 5 et 9 individus, comme le montre l’intervalle interquartile de 4. Cela signifie que 50 % des ménages ont un nombre d’individus compris entre 5 et 9. Les valeurs extrêmes, avec un ménage ayant 1 individu et un autre allant jusqu’à 17 individus, suggèrent cependant quelques ménages plus petits ou plus grands, ce qui peut être lié à des situations particulières comme des ménages monoparentaux ou plus larges. En somme, cette distribution indique une majorité de ménages de taille modeste à moyenne, avec quelques exceptions notables.

#### Analyse bivariée

Nous allons étudier le rapport entre le nombre d’individus par ménage et le milieu de réidence qu’il soit rural, urbain ou péri urbain. Affichons d’abord la répartition de la variable intro\_09 qui représente le milieu de résidence.

rural peri-urban urban   
 2058 740 258

Répartition de la variable 'intro\_09' (milieu de résidence)

| **Var1** | **Freq** |
| --- | --- |
| rural | 2,058 |
| peri-urban | 740 |
| urban | 258 |

Maintenant que nous avons visualisé cette répartition, Passons à l’analyse bivariée.

Répartition du nombre d'individus par ménage et du milieu de résidence (avec pourcentages)

| **milieu\_residence** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **Total** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| rural | 2.7% (55) | 3.8% (79) | 5.3% (110) | 8.1% (167) | 10.3% (213) | 12.7% (261) | 12.9% (266) | 12.7% (261) | 10.2% (210) | 7.8% (161) | 4.3% (88) | 3.0% (62) | 1.9% (39) | 1.3% (26) | 0.8% (17) | 0.7% (15) | 1.4% (28) | **100.0% (2,058)** |
| peri-urban | 3.0% (22) | 2.6% (19) | 3.2% (24) | 5.3% (39) | 13.2% (98) | 12.4% (92) | 14.6% (108) | 12.8% (95) | 12.0% (89) | 8.1% (60) | 4.9% (36) | 3.1% (23) | 0.9% (7) | 1.1% (8) | 0.9% (7) | 0.8% (6) | 0.9% (7) | **100.0% (740)** |
| urban | 3.9% (10) | 3.5% (9) | 6.6% (17) | 6.6% (17) | 15.9% (41) | 12.4% (32) | 10.5% (27) | 12.8% (33) | 10.1% (26) | 5.0% (13) | 3.5% (9) | 5.8% (15) | 1.2% (3) | 1.2% (3) | 0.4% (1) | 0.0% (0) | 0.8% (2) | **100.0% (258)** |
|  | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 100.0% (2) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | **100.0% (2)** |
| **Total** | **2.8% (87)** | **3.5% (107)** | **4.9% (151)** | **7.3% (223)** | **11.5% (352)** | **12.6% (385)** | **13.2% (403)** | **12.7% (389)** | **10.6% (325)** | **7.7% (234)** | **4.3% (133)** | **3.3% (100)** | **1.6% (49)** | **1.2% (37)** | **0.8% (25)** | **0.7% (21)** | **1.2% (37)** | **100.0% (3,058)** |

Le tableau présente la répartition du nombre d’individus par ménage (en catégories) selon les milieux de résidence (rural, péri-urbain, et urbain). Chaque ligne du tableau correspond à un milieu de résidence, et chaque colonne représente une catégorie du nombre d’individus par ménage, allant de 0 à 5 individus à 15 individus et plus.

Interprétation : Milieu rural :

La majorité des ménages ruraux (12,7%) ont entre 0 et 5 individus, suivis par les ménages ayant 5 à 10 individus (12,8%).

Les catégories supérieures (de 10 à 15 et 15 et plus) représentent une proportion plus faible, notamment 1,3% pour 15 et plus.

Milieu péri-urbain :

Le plus grand pourcentage de ménages en milieu péri-urbain est dans les catégories 0 à 5 et 5 à 10 individus, représentant respectivement 6,6% et 12,8% des ménages.

La proportion des ménages avec 15 et plus individus est faible (0,6%).

Milieu urbain :

En milieu urbain, la majorité des ménages (13,2%) ont entre 0 et 5 individus.

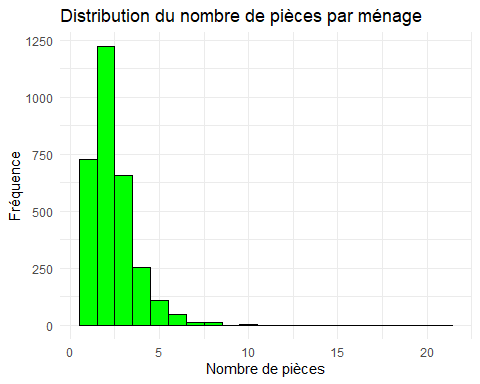
Comme dans les autres milieux, les ménages avec plus de 15 individus sont peu fréquents, représentant 0,6%.

Conclusion : La distribution montre une concentration élevée de ménages ayant un nombre réduit d’individus, particulièrement dans les catégories 0 à 5 et 5 à 10 individus. Le milieu urbain et péri-urbain tend à avoir des ménages légèrement plus grands que les zones rurales, bien que la tendance générale soit similaire entre les trois milieux de résidence.

### II.3.c) Faites des statistiques descriptives sur la variable renseigant le nombre de pieces du menage

Montrons d’abord l’analyse univariée pour la variable nombre de pieces

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's   
 1.000 2.000 2.000 2.388 3.000 21.000 6



Répartition des ménages selon le nombre d'individus et le nombre de pièces (avec pourcentages)

| **nb\_individus\_menage** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **13** | **21** | **NA\_** | **Total** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 65.5% (57) | 28.7% (25) | 4.6% (4) | 1.1% (1) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | **100.0% (87)** |
| 2 | 45.8% (49) | 47.7% (51) | 5.6% (6) | 0.0% (0) | 0.9% (1) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | **100.0% (107)** |
| 3 | 41.1% (62) | 44.4% (67) | 10.6% (16) | 2.6% (4) | 1.3% (2) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | **100.0% (151)** |
| 4 | 35.0% (78) | 42.2% (94) | 14.8% (33) | 4.5% (10) | 1.3% (3) | 0.4% (1) | 0.0% (0) | 0.9% (2) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.4% (1) | 0.4% (1) | **100.0% (223)** |
| 5 | 29.0% (102) | 42.9% (151) | 19.0% (67) | 5.7% (20) | 1.4% (5) | 1.1% (4) | 0.0% (0) | 0.6% (2) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.3% (1) | **100.0% (352)** |
| 6 | 29.1% (112) | 45.7% (176) | 17.4% (67) | 4.7% (18) | 2.6% (10) | 0.3% (1) | 0.0% (0) | 0.3% (1) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | **100.0% (385)** |
| 7 | 21.3% (86) | 43.4% (175) | 22.1% (89) | 7.9% (32) | 3.2% (13) | 0.7% (3) | 0.2% (1) | 0.5% (2) | 0.0% (0) | 0.2% (1) | 0.0% (0) | 0.0% (0) | 0.2% (1) | **100.0% (403)** |
| 8 | 17.5% (68) | 45.0% (175) | 23.9% (93) | 9.0% (35) | 3.6% (14) | 0.3% (1) | 0.3% (1) | 0.5% (2) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | **100.0% (389)** |
| 9 | 16.0% (52) | 38.5% (125) | 28.9% (94) | 9.2% (30) | 4.3% (14) | 1.8% (6) | 0.0% (0) | 0.6% (2) | 0.0% (0) | 0.3% (1) | 0.0% (0) | 0.0% (0) | 0.3% (1) | **100.0% (325)** |
| 10 | 12.4% (29) | 39.3% (92) | 29.1% (68) | 11.5% (27) | 4.7% (11) | 2.1% (5) | 0.4% (1) | 0.0% (0) | 0.4% (1) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | **100.0% (234)** |
| 11 | 13.5% (18) | 23.3% (31) | 31.6% (42) | 16.5% (22) | 6.0% (8) | 6.0% (8) | 2.3% (3) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.8% (1) | **100.0% (133)** |
| 12 | 9.0% (9) | 26.0% (26) | 30.0% (30) | 24.0% (24) | 7.0% (7) | 3.0% (3) | 0.0% (0) | 1.0% (1) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | **100.0% (100)** |
| 13 | 4.1% (2) | 18.4% (9) | 34.7% (17) | 20.4% (10) | 14.3% (7) | 4.1% (2) | 0.0% (0) | 2.0% (1) | 0.0% (0) | 0.0% (0) | 2.0% (1) | 0.0% (0) | 0.0% (0) | **100.0% (49)** |
| 14 | 2.7% (1) | 21.6% (8) | 24.3% (9) | 13.5% (5) | 24.3% (9) | 8.1% (3) | 2.7% (1) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 2.7% (1) | **100.0% (37)** |
| 15 | 8.0% (2) | 20.0% (5) | 36.0% (9) | 12.0% (3) | 4.0% (1) | 12.0% (3) | 0.0% (0) | 0.0% (0) | 4.0% (1) | 4.0% (1) | 0.0% (0) | 0.0% (0) | 0.0% (0) | **100.0% (25)** |
| 16 | 0.0% (0) | 19.0% (4) | 38.1% (8) | 19.0% (4) | 0.0% (0) | 14.3% (3) | 9.5% (2) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | **100.0% (21)** |
| 17 | 0.0% (0) | 21.6% (8) | 10.8% (4) | 24.3% (9) | 10.8% (4) | 13.5% (5) | 13.5% (5) | 5.4% (2) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | **100.0% (37)** |
| **Total** | **23.8% (727)** | **40.0% (1,222)** | **21.5% (656)** | **8.3% (254)** | **3.6% (109)** | **1.6% (48)** | **0.5% (14)** | **0.5% (15)** | **0.1% (2)** | **0.1% (3)** | **0.0% (1)** | **0.0% (1)** | **0.2% (6)** | **100.0% (3,058)** |

Le tableau montre la répartition du nombre de pièces par ménage (en colonnes) pour chaque catégorie de nombre d’individus par ménage (en lignes). Les valeurs représentent les pourcentages des ménages dans chaque catégorie de nombre de pièces pour chaque catégorie du nombre d’individus par ménage.

Interprétation des résultats : Pour les ménages ayant 1 individu :

Environ 65.5% des ménages avec 1 individu ont 1 pièce, ce qui indique une forte concentration des ménages unipersonnels dans des logements à une seule pièce.

Les autres catégories de nombre de pièces sont moins représentées parmi les ménages unipersonnels, avec des pourcentages décroissants pour plus de pièces (2, 3, 4, etc.).

Pour les ménages ayant 2 individus :

Environ 45.8% des ménages avec 2 individus vivent dans des logements de 2 pièces.

La proportion des ménages avec 3 ou 4 pièces est également notable, mais elle reste inférieure à celle des ménages avec 2 pièces.

Pour les ménages ayant entre 5 et 10 individus :

La majorité des ménages dans cette catégorie occupe des logements avec 3 à 4 pièces, ce qui est logique puisque ces ménages sont plus nombreux et nécessitent plus d’espace.

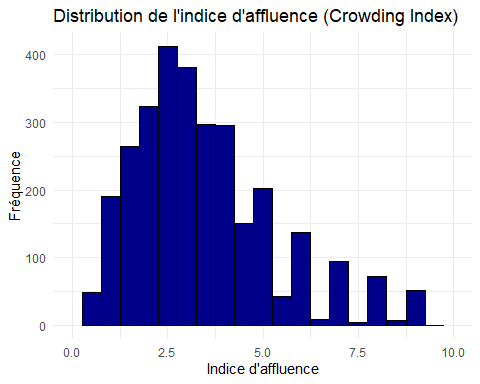
Les ménages ayant plus de 10 individus sont rares et n’occupent généralement que des logements de 3 à 5 pièces, selon la disponibilité des logements de grande taille.

Conclusion : Le tableau montre une tendance générale selon laquelle les ménages unipersonnels occupent principalement des logements d’une seule pièce, tandis que les ménages plus grands ont une plus grande diversité dans le nombre de pièces. Il existe une corrélation évidente entre la taille des ménages et le nombre de pièces disponibles, mais avec une forte concentration des petits ménages dans des logements à faible nombre de pièces.

### II.3.d) Calculons l’indice d’affluence ou crowding index

L’indice d’affluence ou crowding index est un indicateur clé utilisé pour évaluer la densité de la population dans un logement. Il se calcule en divisant le nombre total d’individus dans un ménage par le nombre de pièces disponibles dans le logement, à l’exclusion de la cuisine et des couloirs. Cet indice permet de mesurer dans quelle mesure un ménage est “surpeuplé” en fonction de l’espace disponible. Un indice élevé indique une forte densité de personnes par pièce, suggérant des conditions de vie plus confinées et potentiellement moins confortables, tandis qu’un indice faible reflète des conditions de logement plus spacieuses. Le calcul de cet indice est essentiel pour analyser les conditions de vie des ménages et peut fournir des informations utiles pour les politiques publiques en matière de logement et d’aménagement du territoire

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's   
 0.1905 2.0000 3.0000 3.5970 4.5000 15.0000 6



L’histogramme présente la distribution de l’indice d’affluence (crowding index), qui mesure la densité de la population par ménage en fonction du nombre d’individus par rapport au nombre de pièces dans le logement. Voici l’interprétation de ce graphique :

####Forme de la distribution :

L’histogramme montre une distribution asimétrique avec une forme légèrement bimodale, ce qui suggère que l’indice d’affluence est plus concentré autour de deux valeurs principales, l’une plus faible et l’autre plus élevée.

* Les indices faibles :

Les valeurs faibles de l’indice (proches de 0 à 2.5) représentent les ménages avec peu de densité (moins de personnes par pièce), ce qui indique des logements plus spacieux. Cette catégorie semble être la plus fréquente, avec un pic dans la plage entre 1.5 et 2.5.

* Les indices moyens et élevés :

L’indice d’affluence atteint son maximum autour de 2.5 à 3, avec une fréquence élevée des ménages ayant un indice modéré.

Au fur et à mesure que l’indice augmente, la fréquence des ménages diminue. Cela indique qu’il y a moins de ménages dans des conditions de surpeuplement sévère (indice au-dessus de 5).

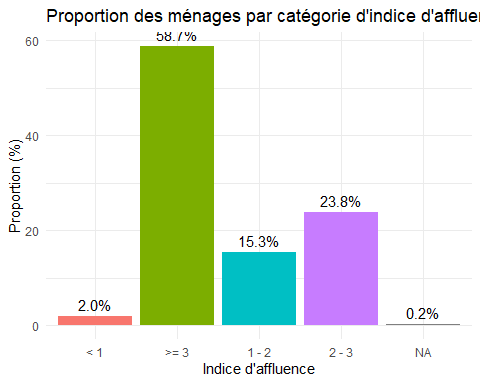
* Individus à très haute densité :

Les valeurs les plus élevées de l’indice (au-delà de 7.5) sont rares, indiquant que seulement quelques ménages vivent dans des conditions de surpeuplement extrême.

Conclusion : L’indice d’affluence montre que la majorité des ménages ont une densité relativement faible, avec une forte proportion de ménages ayant un indice d’affluence proche de 2 à 3, ce qui suggère une certaine densité modérée mais pas excessive. Les ménages à très haute densité sont moins fréquents, indiquant que les problèmes de surpeuplement extrême sont relativement rares dans cette population.

### II.3.e) Calcul de la proportion de menage dont le crowding index est: <1, 1<= crowding\_index<2, 2<= crowding\_index<3 et >=3?

| categorie\_crowding | n | proportion |
| --- | --- | --- |
| 1 - 2 | 467 | 15.2714192 |
| 2 - 3 | 728 | 23.8064094 |
| < 1 | 61 | 1.9947678 |
| >= 3 | 1796 | 58.7311969 |
| NA | 6 | 0.1962067 |



Le graphique montre la proportion des ménages selon différentes catégories d’indice d’affluence (crowding index). Cet indice mesure la densité de la population par ménage en fonction du nombre d’individus par rapport au nombre de pièces. Voici l’interprétation des résultats :

Interprétation du graphique : - Catégorie : Indice d’affluence inférieur à 1 :

Seuls 2% des ménages ont un indice d’affluence inférieur à 1, ce qui signifie que ces ménages vivent dans des logements relativement spacieux par rapport au nombre d’individus.

* Catégorie : Indice d’affluence entre 1 et 2 :

23.8% des ménages se trouvent dans cette catégorie. Ces ménages ont un indice d’affluence modéré, indiquant une densité moyenne de personnes par pièce.

-Catégorie : Indice d’affluence entre 2 et 3 :

15.3% des ménages se situent dans cette catégorie. Cela suggère que ces ménages commencent à vivre dans des conditions de densité plus élevées.

* Catégorie : Indice d’affluence supérieur ou égal à 3 :

Une proportion significative de ménages (58.7%) a un indice d’affluence supérieur ou égal à 3, indiquant que ces ménages sont confrontés à une densité élevée de population par rapport à la taille de leur logement.

NA (Non défini) :

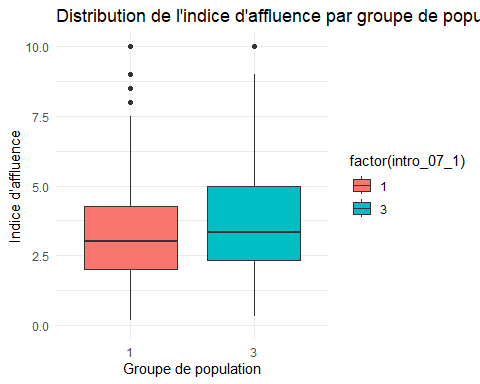
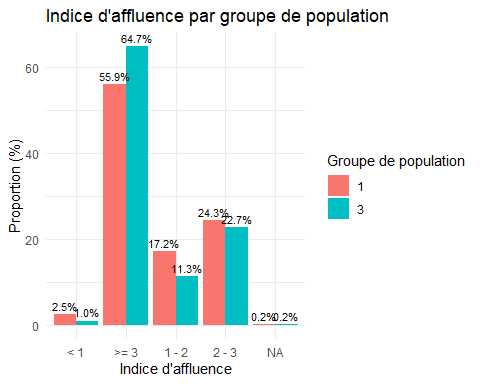
Une très faible proportion (0.2%) de données sont manquantes ou non spécifiées.

Conclusion : La majorité des ménages (58.7%) se trouvent dans des conditions de surpeuplement modéré à élevé, avec un indice d’affluence de 3 ou plus, ce qui suggère que la densité de population dans certains logements est relativement élevée. En revanche, une proportion plus faible de ménages vit dans des conditions moins densifiées (indice inférieur à 2). Ce type d’information est crucial pour évaluer les conditions de vie des ménages et peut aider à orienter les politiques publiques concernant le logement et la planification urbaine.

### II.3.f) Ces statistiques suivant que le menage soit un refugie ou un member de la communaute d’acceuil (host community North)

L’analyse des statistiques en fonction du statut du ménage (réfugié ou membre de la communauté d’accueil) permet de mieux comprendre les différences de conditions de vie entre ces deux groupes. Les réfugiés et les membres des communautés d’accueil peuvent faire face à des réalités socio-économiques très différentes, influençant notamment la taille des ménages, l’accès au logement, les conditions de vie et les ressources disponibles. Cette section se concentre sur l’étude des ménages réfugiés par rapport à ceux de la communauté d’accueil du nord, en examinant des critères tels que l’indice d’affluence, la répartition des tailles de ménages et d’autres caractéristiques socio-économiques pertinentes. L’objectif est de déterminer dans quelle mesure ces groupes diffèrent en termes de densité de population, de conditions de logement et d’autres facteurs essentiels pour comprendre les défis auxquels ils sont confrontés.

| intro\_07\_1 | n | moyenne | mediane | ecart\_type | min | max |
| --- | --- | --- | --- | --- | --- | --- |
| 1 | 2068 | 3.426029 | 3.000000 | 1.977794 | 0.1904762 | 15 |
| 3 | 990 | 3.954186 | 3.333333 | 2.307443 | 0.3333333 | 15 |



Rappelons que 1 représente les réfugiées et 2 les host community north ou communauté d’accueil nord. # Interprétation de l’Indice d’Affluence par Groupe de Population

Ce graphique montre la répartition de l’indice d’affluence (**crowding index**) pour deux groupes de population distincts : les **réfugiés** et les **membres de la communauté d’accueil du nord**. Les barres indiquent la proportion de ménages dans chaque groupe de population, répartis en différentes catégories d’indice d’affluence.

## Interprétation du graphique

* **Groupe des réfugiés** (représenté par la couleur rouge) :
  + **55.9% des réfugiés** ont un indice d’affluence entre **1 et 2**.
    - Cela signifie que la majorité des réfugiés vivent dans des conditions de densité modérée.
  + **17.2% des réfugiés** ont un indice d’affluence compris entre **2 et 3**.
  + **2.5% des réfugiés** ont un indice d’affluence inférieur à **1**, ce qui suggère des logements relativement spacieux.
  + **17.6% des réfugiés** vivent dans des logements avec un indice d’affluence supérieur à **3**, indiquant des conditions de surpeuplement plus importantes.
* **Groupe des membres de la communauté d’accueil** (représenté par la couleur turquoise) :
  + **64.7% des membres de la communauté d’accueil** ont un indice d’affluence inférieur à **1**, ce qui indique que ces ménages vivent dans des logements moins denses et plus spacieux.
  + **22.7% des membres de la communauté d’accueil** ont un indice d’affluence compris entre **2 et 3**.
  + **24.3% des membres de la communauté d’accueil** ont un indice d’affluence entre **1 et 2**.
  + **0.2% des membres de la communauté d’accueil** vivent dans des conditions de surpeuplement sévère (**indice >= 3**).

## Conclusion

* Les **réfugiés** ont une proportion plus élevée de ménages vivant dans des logements plus densifiés, avec des indices d’affluence compris entre **1 et 3**.
* Les **membres de la communauté d’accueil** ont, en revanche, une proportion plus importante de ménages vivant dans des logements moins densifiés, avec un indice d’affluence inférieur à **1**.
* Ces différences peuvent refléter les conditions de vie plus difficiles des réfugiés, qui peuvent se retrouver dans des logements plus exiguës en raison des contraintes économiques ou des politiques d’accueil.

# III. Analyse de la securite alimentaire des deplaces internes

1. Score de consommation alimentaire (SCA)