***Examen de projet statistique sur R, ISE1 2024-2025***

***Enseignant: Aboubacar Hema***

**Instructions :**

Vous avez 4 heures à partir du moment où vous recevez ce test pour envoyer vos réponses dans le compte GitHub du cours ([Abson-dev/Projet-statistique-sous-R](https://github.com/Abson-dev/Projet-statistique-sous-R)). **Les soumissions tardives ne seront pas prises en compte.**

Vous devez compléter ce test **seul.e, sans l’aide d’une tierce personne**. En revanche, l’utilisation de **ChatGPT, Grock, Deepseek, etc** sont **autorisés pour effectuer des recherches documentaires ou consulter le support R**.

Les codes que vous utiliserez pour répondre aux questions seront à intégrer dans le corps de votre rapport. C’est pourquoi l’utilisation de R Markdown est à privilégier. En ce qui concerne l’organisation du travail à rendre, vous vous inspirerez de la façon dont est organisé le sujet de l’examen (vous n’êtes pas obliger de recopier l’énoncé). Le barème est indiqué en face de chaque question. Il tient compte du code proposé et des commentaires sur les résultats obtenus, nous tiendrons compte des commentaires que vous aurez fait dans la notation.

**IMPORTANT:** **il est demandé de commenter vos lignes de code. Vous pourriez être pénalisés si vous ne commentez pas vos codes. (+/- 10 points)**

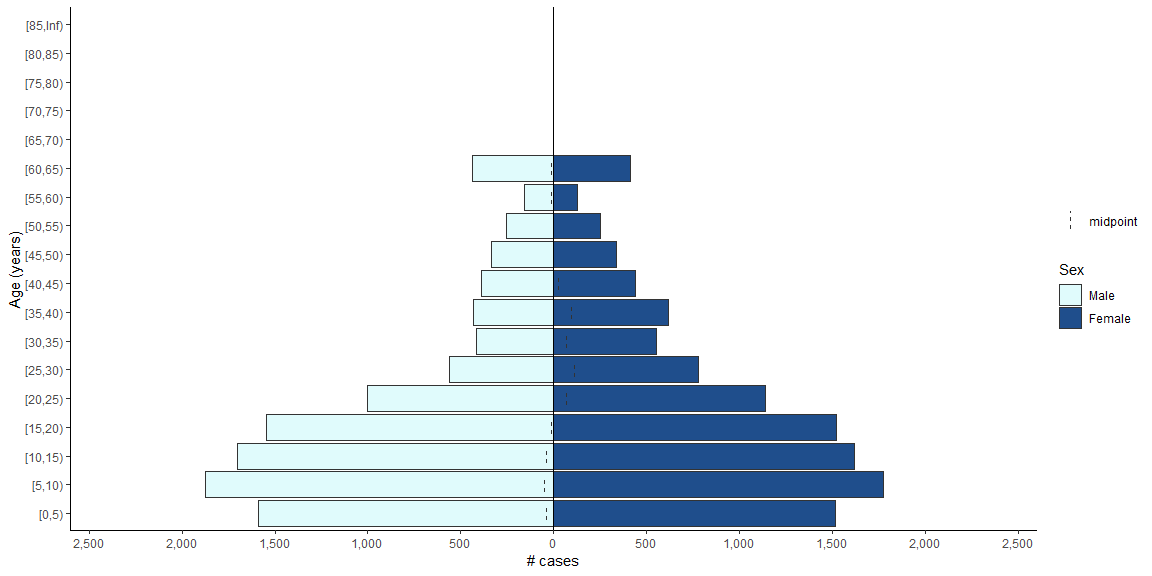
Cette etude est un cas reel, portant sur le Sudan. Il s’agit de l’***Enquête sur les déplacements forcés au Sud-Soudan en 2023***.

1. **Faites une analyse de consistence des bases de données: 5 points**

Toutes les étapes doivent etre précisées et justifiées.

1. **Analyse socio-economique des deplaces internes du Sudan en 2023: 80 points**
2. **Pyramide des ages des individus: 5 points**

**Proposez une pyramide des ages des individus.**



1. **Caracteristiques du chef de menage**

**Produisez un tableau similaire**

| Caracteristiques du chef de menage | **Refugees**, N = 2,0681 | **Host community North**, N = 9901 | **p-value**2 |
| --- | --- | --- | --- |
| ageYears | 42(17,61) | 42(19,61) | >0.9 |
| Sex |  |  | <0.001 |
| Male | 54 | 63 |  |
| Female | 46 | 37 |  |
| What is [your/name's] present marital status? |  |  |  |
| monogamous/married | 63 | 72 |  |
| polygamous/married | 14 | 19 |  |
| non-formal union | 0.6 | 0.3 |  |
| separated | 5.4 | 1.5 |  |
| divorced | 3.3 | 0.9 |  |
| widow or widower | 7.4 | 5.2 |  |
| never married | 6.9 | 1.1 |  |
| 1Mean(Minimum,Maximum); % | | | |
| 2Wilcoxon rank sum test; Pearson's Chi-squared test; Fisher's exact test | | | |

1. **Crowding Index ou l’indice d’affluence**

L'indice d'affluence est le nombre de membres du menage divise par le nombre de pièces (à l'exclusion de la cuisine et des couloirs).

* 1. Calculez le nombre d’individu par menage
  2. Faites les statistiques descriptives sur cette variable
  3. Faites des statistiques descriptives sur la variable renseigant le nombre de pieces du menage
  4. Calculez l’indice d’affluence ou crowding index
  5. Quelle est la proportion de menage dont le crowding index est: <1, 1<= crowding\_index<2, 2<= crowding\_index<3 et >=3?
  6. Faites ces statistiques suivant que le menage soit un refugie ou un member de la communaute d’acceuil (host community North)

1. **Analyse de la securite alimentaire des deplaces internes**
2. **Score de consommation alimentaire (SCA): 10 points**

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.

* 1. Faites une analyse descriptive des variables qui composent le SCA
  2. Calculer le score de consommation alimentaire
  3. Faites un tableau illustrant le poids attribue a chaque groupe alimentaire pour le calcul du SCA (la somme totale des poids doit etre egale a 16)
  4. Categoriser le SCA selon les seuil 21/35 et 28/42
  5. Faites une répresentation spatiale (région et département) du SCA et de ses différentes catégorisations.

1. **L’indice réduit des stratégies de survie (rCSI):5 points**

L'indice réduit des stratégies de survie, basé sur la consommation, est utilisé pour évaluer le niveau de stress d'un ménage dû à une pénurie alimentaire. Il mesure les stratégies comportementales que les gens appliquent lorsqu'ils n'ont pas accès à suffisamment de nourriture ou lorsqu'ils prévoient une diminution de la sécurité alimentaire.

* 1. Faites une analyse descriptive des variables qui composent le rCSI

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** generez aleatoirement une valeur comprise entre 1 et 7 attribue a la variable, dans le cas contraire la nouvelle variable prend 0.

* 1. Faites des statistiques descriptives sur les Nouvelles variables
  2. Calucler l’indice réduit des stratégies de survie en utilisant les Nouvelles variables
  3. Faites un tableau illustrant le poids attribue a chaque variable pour le calcul du rCSI (la somme totale des poids doit etre egale a 21)
  4. Faites une representation spatiale selon le milieu de residence du rCSI