|  |
| --- |
| **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 avec R* |

# EVALUATION

|  |  |
| --- | --- |
| **Fait par** | **Sous la supervision de** |
| Fogwoung Djoufack Sarah-Laure | M. Aboubacar HEMA |
| *Élève Ingénieure Statisticienne Économiste* | *Research analyst* |

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

# Sommaire

# Liste des figures

# Liste des tableaux

# Introduction

Dans le cadre de l’évaluation faisant suite aux 20 heures de cours, ce document présente une application concrète des notions abordées. Il s’agit ici de (…) — [à compléter avec la consigne précise] — afin de mettre en pratique les compétences acquises tout au long de la formation.

# Préliminaires

## Installation des packages necessaires

Dans le cadre de cette évaluation, plusieurs packages R ont été mobilisés pour assurer à la fois la manipulation des données, leur traitement statistique et la présentation des résultats. Les packages du tidyverse ont permis d’effectuer des opérations essentielles telles que l’importation, le nettoyage, la transformation et la visualisation des données. Des outils complémentaires tels que janitor et labelled ont été employés respectivement pour automatiser le nettoyage des variables (notamment la normalisation des noms de colonnes) et pour gérer efficacement les étiquettes de variables et de modalités. Par ailleurs, le package haven a permis de lire des bases stata, tandis que sf a servi à la gestion de données spatiales. Pour l’aspect la mise en forme des résultats, des packages tels que gt, gtsummary, flextable, officer et officedown ont été utilisés afin de générer des tableaux statistiques clairs et adaptés à un rendu professionnel, notamment dans Word. glue a facilité la génération dynamique de texte, et questionr (avec ireca été utile pour des analyses rapides et descriptives sur les données à travers l’interface offerte.

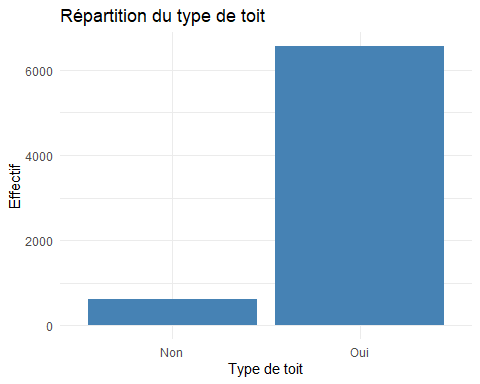
## Chargement des bases de données

Premieres observations

| Region ID | Commune ID | Village ID | Household ID | Survey round : Baseline, Endline | eating occasion | Meal consumed? Y/N | Who consummed? |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 2 | 25 | 1,000 | 4948484848535052 | Baseline | Breakfast | Yes | mother |
| 2 | 25 | 1,000 | 4948484848535052 | Baseline | Lunch | No | mother |
| 2 | 25 | 1,000 | 4948484848535052 | Baseline | Dinner | Yes | mother |
| 2 | 25 | 1,000 | 4948484848535052 | Baseline | Snacks | Yes | mother |
| 2 | 25 | 1,000 | 4948484848535053 | Baseline | Breakfast | Yes | mother |
| 2 | 25 | 1,000 | 4948484848535053 | Baseline | Lunch | No | mother |

Table : Nombre de doublons

| **Variable** | **Nombre\_doublons** |
| --- | --- |
| **enquete** | 7,175 |
| **logem** | 7,172 |
| **hhid** | 0 |



**Figure** **:**  Répartition du type de toit Sarah