

# Rapport de traitement des données: répartition des activités agricoles et d'élevage - Sénégal

Kadidja GUEBEDIAN G A NKEN-Isep2

Année académique 2024-2025

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Chargement des bibliothèques</b>	<b>2</b>
<b>3</b>	<b>Importation des données</b>	<b>2</b>
<b>4</b>	<b>Préparation des données</b>	<b>3</b>
<b>5</b>	<b>Agrégation des données</b>	<b>3</b>
5.1	Agrégation par ménage . . . . .	3
5.2	Agrégation par région . . . . .	3
<b>6</b>	<b>Conclusion</b>	<b>5</b>

# 1 Introduction

Ce document a pour objectif de présenter les traitements faits aux données de l'enquête EHCVM 2018 faite au Sénégalais, traitements permettant de mieux entamer l'analyse de la distribution des activités agricoles et pastorales au Sénégal.

## 2 Chargement des bibliothèques

Dans ce projet, plusieurs bibliothèques R ont été mobilisées :

`sf` : pour lire et manipuler les fichiers géographiques (.shp) et produire des cartes.

`haven` : pour importer les fichiers de données d'enquête au format Stata (.dta).

`dplyr` et `tidyr` : pour nettoyer, transformer, regrouper et structurer les données.

`tmap` : pour créer des cartes thématiques représentant les indicateurs par région.

`ggplot2` : pour réaliser des graphiques personnalisés (barres, histogrammes, etc.).

`patchwork` : pour combiner plusieurs graphiques sur une même figure.

```
library(sf)
library(haven)
library(dplyr)
library(tidyr)
library(tmap)
library(ggplot2)
```

## 3 Importation des données

Pour ce rapport, nous avons eu besoin d'utiliser de 05 bases à savoir -la base s16a, données pour l'élevage ;

-la base s16b, données pour l'élevage ;

-la base s17, données pour l'agriculture ;

-la base s00, données pour les coordonnées géographiques ;

-la base s16c, données pour l'élevage.

```
s16a <- read_dta("C:/Users/LENOVO/Desktop/second semestre/ISEP2025/ISEP2025/projet_cartographie_Shiny/s16a.dta")
s16b <- read_dta("C:/Users/LENOVO/Desktop/second semestre/ISEP2025/ISEP2025/projet_cartographie_Shiny/s16b.dta")
s17 <- read_dta("C:/Users/LENOVO/Desktop/second semestre/ISEP2025/ISEP2025/projet_cartographie_Shiny/s17.dta")
s00 <- read_dta("C:/Users/LENOVO/Desktop/second semestre/ISEP2025/ISEP2025/projet_cartographie_Shiny/s00.dta")
s16c <- read_dta("C:/Users/LENOVO/Desktop/second semestre/ISEP2025/ISEP2025/projet_cartographie_Shiny/s16c.dta")
```

## 4 Préparation des données

Pour la préparation des données, nous sommes passés par le renommage des variables que nous utiliserons. Ensuite, nous avons extrait les variables dont nous avons besoins en créant de nouveaux dataframes ne contenant que ces variables.

```
s00 <- s00 %>%
  rename(region = s00q01, departement = s00q02, milieu = s00q04) %>%
  select(grappe, menage, region, departement, milieu)

s16a <- s16a %>%
  rename(culture = s16aq08, superficie = s16aq09a)

s17 <- s17 %>%
  rename(espece = s17q02, nb_animaux = s17q05)
```

## 5 Agrégation des données

L'agrégation des données passent par la création d'une base de données consolidée à l'échelle du ménage et à l'échelle régionale afin d'analyser simultanément les activités agricoles et d'élevage . Il repose sur plusieurs étapes clés :

### 5.1 Agrégation par ménage

Cette étape compte le nombre de parcelles cultivées par ménage (nb\_parcelles). Chaque observation dans s16a correspond à une parcelle. En regroupant par grappe et menage, on obtient le total de parcelles pour chaque ménage. Ensuite, on totalise le nombre d'animaux détenus par chaque ménage (nb\_animaux). Nous aboutissons alors à une base commune par ménage, contenant à la fois le nombre de parcelles et le nombre d'animaux. les valeurs manquantes dans nb\_parcelles et nb\_animaux seront remplacées par 0.

```
agriculture <- s16a %>%
  group_by(grappe, menage) %>%
  summarise(nb_parcelles = n(), .groups = "drop")

elevage <- s17 %>%
  group_by(grappe, menage) %>%
  summarise(nb_animaux = sum(nb_animaux, na.rm = TRUE), .groups = "drop")

merged_data <- full_join(agriculture, elevage, by = c("grappe", "menage"))
merged_geo <- left_join(merged_data, s00, by = c("grappe", "menage"))

merged_geo <- merged_geo %>%
  mutate(nb_parcelles = replace_na(nb_parcelles, 0),
         nb_animaux = replace_na(nb_animaux, 0))
```

### 5.2 Agrégation par région

Cette partie permet de résumer les activités agricoles et pastorales à l'échelle régionale, à partir des données préalablement agrégées par ménage. Ainsi on calcule pour chaque région, le nombre total de parcelles agricoles et le nombre total d'animaux détenus

```

region_data <- merged_geo %>%
  group_by(region) %>%
  summarise(
    total_parcelles = sum(nb_parcelles, na.rm = TRUE),
    total_animaux = sum(nb_animaux, na.rm = TRUE)
  )

region_data <- region_data %>% mutate(region = toupper(region))

correspondance <- data.frame(
  region = as.character(1:14),
  NOMREG = c("DAKAR", "ZIGUINCHOR", "DIOURBEL", "SAINT LOUIS", "TAMBACOUNDA",
             "KAOLACK", "THIES", "LOUGA", "FATICK", "KOLDA",
             "MATAM", "KAFFRINE", "KEDOUGOU", "SEDHIOU")
)

region_data <- left_join(region_data, correspondance, by = "region")

sen_map <- st_read("C:/Users/LENOVO/Desktop/second semestre/ISEP2025/ISEP2025/projet_cartographie_Shiny/

## Reading layer `Limite_Région' from data source
##   `C:\Users\LENOVO\Desktop\second semestre\ISEP2025\ISEP2025\projet_cartographie_Shiny\Limite_Région'
##   using driver `ESRI Shapefile'
## Simple feature collection with 14 features and 8 fields
## Geometry type: POLYGON
## Dimension:      XY
## Bounding box:   xmin: 227586.3 ymin: 1362012 xmax: 897104.7 ymax: 1845672
## CRS:            NA

map_data <- left_join(sen_map, region_data, by = "NOMREG")

```

## 6 Conclusion

Ce prétraitement a permis de structurer efficacement les données issues de l'enquête EHCVM 2018 en les agréant à l'échelle régionale. En croisant les informations sur les ménages agricoles et d'élevage, puis en les reliant à leur localisation géographique, nous avons obtenu une base consolidée. Le renommage des variables et l'harmonisation des noms de régions ont facilité l'intégration avec les données spatiales (shapefile), ouvrant ainsi la voie à une visualisation cartographique cohérente et exploitable des dynamiques agricoles et pastorales sur l'ensemble du territoire sénégalais.