TP\_2

ZINABA\_Albert

2025-01-23

Table of Contents

# I-Telechargements des packages et importation des bases de données

## I-1-Telechargement des packages

library(utils)  
library(dplyr)

## I-2 Importation de la base Individu

Individu<- utils::read.csv("../Données/ehcvm\_individu\_bfa2021.csv")  
Individu  
  
  
#La syntaxe ci-dessus permet a travers la fonction 'read.csv()' permet de charger #le fichier 'ehcvm\_individu\_bfa2021.csv' dans un objet 'Individu' sous #forme de dataframe manipulable dans R.

## I-3-Importation de la base Menage

Menage<- utils::read.csv("../Données/ehcvm\_menage\_bfa2021.csv")  
Menage  
  
#Cette syntaxe aussi permet de charger le #fichier 'ehcvm\_menage\_bfa2021.csv' dans un objet 'Menage' sous forme de #dataframe #manipulable dans R.

## I-4 Structure des deux bases

Nous allons utiliser la fonction “str” du package “utils” pour afficher la structure des deux bases.

utils::str(Individu)  
utils::str(Menage)

# II-Jointure des deux bases

Ici nous utilisons la fonction “left\_join” du package dplyr pour fusioner les deux bases. On obtient une nouvelle base de données tout en gardeant les identifiants des individus.

Base\_merge<- dplyr:: left\_join(Individu,Menage,by="hhid")  
Base\_merge

## II-1 Creation d’une nouvelle base de données avec un identifiant unique pour chaque individu

Cette nouvelle base a été crée en ajoutant la variable “identifiant” crée par combinaison de deux vriables a savoir “hhid” et “pid”, issus de la base “base\_merge”. Cette nouvelle base nommée “Base\_Id” contient a present 96 variables.

Base\_Id <- Base\_merge %>%  
 dplyr:: mutate(identifiant = hhid \* 100 + pid) %>%  
 select(country.x,year.x,vague.x, hhid, grappe.x, menage.x, pid, identifiant, everything()) %>%  
 arrange(identifiant)  
Base\_Id

## II-2 Filtrage de données par l’identifiant de l’individu

A travers la fonction “filter et select” de dplyr nous filtrons certaines informations d’un individu dans la Base\_Id en inserant comme argument son identifiant crée. Par exemple avec la syntaxe suivante on filtre les informations de l’individu “58600507” comme son sexe, son age, son logement et lien avec le chef de menage, sa religion, sa region entre autres.

Donnee\_filtre <- Base\_Id %>%  
 dplyr:: filter(identifiant == "58600507") %>%  
 select(sexe, age, logem, lien, religion, region)  
Donnee\_filtre