Merge de bases welfare

ILLY Jacques

2025-04-01

Ce travail consiste consiste à merger les bases welfare des EHCVM 2018 ET 2021 en une seule base. Pour cela, nous prenons pour le Burkina Faso.

# **I-Installation des outils**

## 1-Installation et Telechargement des packages nécessaires

# Nous donnons à packages le vecteurs des packages  
packages <- c("dplyr", "haven", "readr", "gt", "labelled", "questionr", "utils","cardx" , "survey", "gtsummary")  
#Installation   
for(pack in packages){  
 if(!requireNamespace(pack, quietly = TRUE)){  
 install.packages(pack)  
 }  
 #telechargement  
 library(pack, character.only = TRUE)  
}

##   
## Attachement du package : 'dplyr'

## Les objets suivants sont masqués depuis 'package:stats':  
##   
## filter, lag

## Les objets suivants sont masqués depuis 'package:base':  
##   
## intersect, setdiff, setequal, union

## Warning: le package 'questionr' a été compilé avec la version R 4.4.3

## Le chargement a nécessité le package : grid

## Le chargement a nécessité le package : Matrix

## Le chargement a nécessité le package : survival

##   
## Attachement du package : 'survey'

## L'objet suivant est masqué depuis 'package:graphics':  
##   
## dotchart

## 2-Installation des bases

#installation de welfare 2018  
welfare18 <- read\_dta("../Données/ehcvm\_welfare\_bfa2018.dta")  
#installation de welfare 2021  
welfare21 <- read\_dta("../Données/ehcvm\_welfare\_bfa2021.dta")

# **II- Inspection des donnés**

## 1-Aperçu des premières et dernières observations

#base welfare18  
head(welfare18)

## # A tibble: 6 × 34  
## country year hhid grappe menage vague zae region milieu hhweight  
## <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl+lbl> <dbl+l> <dbl+l> <dbl>  
## 1 BFA 2018 1009 1 9 2 4 [Grand-sah… 5 [Cen… 2 [Rur… 836.  
## 2 BFA 2018 1015 1 15 2 4 [Grand-sah… 5 [Cen… 2 [Rur… 836.  
## 3 BFA 2018 1036 1 36 2 4 [Grand-sah… 5 [Cen… 2 [Rur… 836.  
## 4 BFA 2018 1061 1 61 2 4 [Grand-sah… 5 [Cen… 2 [Rur… 836.  
## 5 BFA 2018 1121 1 121 2 4 [Grand-sah… 5 [Cen… 2 [Rur… 836.  
## 6 BFA 2018 1141 1 141 2 4 [Grand-sah… 5 [Cen… 2 [Rur… 836.  
## # ℹ 24 more variables: hhsize <dbl>, eqadu1 <dbl>, eqadu2 <dbl>,  
## # hgender <dbl+lbl>, hage <dbl>, hmstat <dbl+lbl>, hreligion <dbl+lbl>,  
## # hnation <dbl+lbl>, halfab <dbl+lbl>, heduc <dbl+lbl>, hdiploma <dbl+lbl>,  
## # hhandig <dbl+lbl>, hactiv7j <dbl+lbl>, hactiv12m <dbl+lbl>,  
## # hbranch <dbl+lbl>, hsectins <dbl+lbl>, hcsp <dbl+lbl>, dali <dbl>,  
## # dnal <dbl>, dtot <dbl>, pcexp <dbl>, zref <dbl>, def\_spa <dbl>,  
## # def\_temp <dbl>

tail(welfare18)

## # A tibble: 6 × 34  
## country year hhid grappe menage vague zae region milieu hhweight  
## <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl+lbl> <dbl+l> <dbl+l> <dbl>  
## 1 BFA 2018 585046 585 46 2 2 [Grand-es… 8 [Est] 2 [Rur… 378.  
## 2 BFA 2018 585052 585 52 2 2 [Grand-es… 8 [Est] 2 [Rur… 378.  
## 3 BFA 2018 585057 585 57 2 2 [Grand-es… 8 [Est] 2 [Rur… 378.  
## 4 BFA 2018 585063 585 63 2 2 [Grand-es… 8 [Est] 2 [Rur… 378.  
## 5 BFA 2018 585091 585 91 2 2 [Grand-es… 8 [Est] 2 [Rur… 378.  
## 6 BFA 2018 585102 585 102 2 2 [Grand-es… 8 [Est] 2 [Rur… 378.  
## # ℹ 24 more variables: hhsize <dbl>, eqadu1 <dbl>, eqadu2 <dbl>,  
## # hgender <dbl+lbl>, hage <dbl>, hmstat <dbl+lbl>, hreligion <dbl+lbl>,  
## # hnation <dbl+lbl>, halfab <dbl+lbl>, heduc <dbl+lbl>, hdiploma <dbl+lbl>,  
## # hhandig <dbl+lbl>, hactiv7j <dbl+lbl>, hactiv12m <dbl+lbl>,  
## # hbranch <dbl+lbl>, hsectins <dbl+lbl>, hcsp <dbl+lbl>, dali <dbl>,  
## # dnal <dbl>, dtot <dbl>, pcexp <dbl>, zref <dbl>, def\_spa <dbl>,  
## # def\_temp <dbl>

#base welfare21  
head(welfare21)

## # A tibble: 6 × 45  
## country year hhid grappe menage vague zae region milieu hhweight  
## <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl+lbl> <dbl+l> <dbl+l> <dbl>  
## 1 bfa 2021 586005 586 5 2 1 [Grand-ou… 1 [Bou… 2 [Rur… 439  
## 2 bfa 2021 586028 586 28 2 1 [Grand-ou… 1 [Bou… 2 [Rur… 439  
## 3 bfa 2021 586043 586 43 2 1 [Grand-ou… 1 [Bou… 2 [Rur… 439  
## 4 bfa 2021 586044 586 44 2 1 [Grand-ou… 1 [Bou… 2 [Rur… 439  
## 5 bfa 2021 586052 586 52 2 1 [Grand-ou… 1 [Bou… 2 [Rur… 439  
## 6 bfa 2021 586082 586 82 2 1 [Grand-ou… 1 [Bou… 2 [Rur… 439  
## # ℹ 35 more variables: hhsize <dbl>, eqadu1 <dbl>, eqadu2 <dbl>,  
## # hgender <dbl+lbl>, hage <dbl>, hmstat <dbl+lbl>, hreligion <dbl+lbl>,  
## # hnation <dbl+lbl>, hethnie <dbl+lbl>, halfa <dbl+lbl>, halfa2 <dbl+lbl>,  
## # heduc <dbl+lbl>, hdiploma <dbl+lbl>, hhandig <dbl+lbl>, hactiv7j <dbl+lbl>,  
## # hactiv12m <dbl+lbl>, hbranch <dbl+lbl>, hsectins <dbl+lbl>, hcsp <dbl+lbl>,  
## # dali <dbl>, dnal <dbl>, dtot <dbl>, dtet <dbl>, pcexp <dbl>, zref <dbl>,  
## # def\_spa <dbl>, def\_temp <dbl>, zaemil <dbl>, dif <dbl>, p0 <dbl>, …

tail(welfare21)

## # A tibble: 6 × 45  
## country year hhid grappe menage vague zae region milieu hhweight  
## <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl+lbl> <dbl+l> <dbl+l> <dbl>  
## 1 bfa 2021 1185024 1185 24 1 2 [Grand-e… 8 [Est] 2 [Rur… 821  
## 2 bfa 2021 1185033 1185 33 1 2 [Grand-e… 8 [Est] 2 [Rur… 821  
## 3 bfa 2021 1185034 1185 34 1 2 [Grand-e… 8 [Est] 2 [Rur… 821  
## 4 bfa 2021 1185035 1185 35 1 2 [Grand-e… 8 [Est] 2 [Rur… 821  
## 5 bfa 2021 1185047 1185 47 1 2 [Grand-e… 8 [Est] 2 [Rur… 821  
## 6 bfa 2021 1185051 1185 51 1 2 [Grand-e… 8 [Est] 2 [Rur… 821  
## # ℹ 35 more variables: hhsize <dbl>, eqadu1 <dbl>, eqadu2 <dbl>,  
## # hgender <dbl+lbl>, hage <dbl>, hmstat <dbl+lbl>, hreligion <dbl+lbl>,  
## # hnation <dbl+lbl>, hethnie <dbl+lbl>, halfa <dbl+lbl>, halfa2 <dbl+lbl>,  
## # heduc <dbl+lbl>, hdiploma <dbl+lbl>, hhandig <dbl+lbl>, hactiv7j <dbl+lbl>,  
## # hactiv12m <dbl+lbl>, hbranch <dbl+lbl>, hsectins <dbl+lbl>, hcsp <dbl+lbl>,  
## # dali <dbl>, dnal <dbl>, dtot <dbl>, dtet <dbl>, pcexp <dbl>, zref <dbl>,  
## # def\_spa <dbl>, def\_temp <dbl>, zaemil <dbl>, dif <dbl>, p0 <dbl>, …

## 2-Structure des bases

#base welfare18  
str(welfare18)

## tibble [7,010 × 34] (S3: tbl\_df/tbl/data.frame)  
## $ country : chr [1:7010] "BFA" "BFA" "BFA" "BFA" ...  
## ..- attr(\*, "label")= chr "Pays"  
## ..- attr(\*, "format.stata")= chr "%9s"  
## $ year : num [1:7010] 2018 2018 2018 2018 2018 ...  
## ..- attr(\*, "label")= chr "Annee enquete"  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ hhid : num [1:7010] 1009 1015 1036 1061 1121 ...  
## ..- attr(\*, "label")= chr "Idenfiant menage"  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ grappe : num [1:7010] 1 1 1 1 1 1 1 1 1 1 ...  
## ..- attr(\*, "label")= chr "Numero grappe"  
## ..- attr(\*, "format.stata")= chr "%6.2f"  
## $ menage : num [1:7010] 9 15 36 61 121 141 168 210 215 235 ...  
## ..- attr(\*, "label")= chr "Numero menage"  
## ..- attr(\*, "format.stata")= chr "%6.2f"  
## $ vague : num [1:7010] 2 2 2 2 2 2 2 2 2 2 ...  
## ..- attr(\*, "label")= chr "Vague"  
## ..- attr(\*, "format.stata")= chr "%6.2f"  
## $ zae : dbl+lbl [1:7010] 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4,...  
## ..@ label : chr "Zone agroecologique"  
## ..@ format.stata: chr "%12.0g"  
## ..@ labels : Named num [1:5] 1 2 3 4 5  
## .. ..- attr(\*, "names")= chr [1:5] "Grand-ouest" "Grand-est" "Grand-centre" "Grand-sahel" ...  
## $ region : dbl+lbl [1:7010] 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 12, 1...  
## ..@ label : chr "Region residence"  
## ..@ format.stata: chr "%17.0g"  
## ..@ labels : Named num [1:13] 1 2 3 4 5 6 7 8 9 10 ...  
## .. ..- attr(\*, "names")= chr [1:13] "Boucle du Mouhoun" "Cascades" "Centre" "Centre-Est" ...  
## $ milieu : dbl+lbl [1:7010] 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1,...  
## ..@ label : chr "Milieu residence"  
## ..@ format.stata: chr "%9.0g"  
## ..@ labels : Named num [1:2] 1 2  
## .. ..- attr(\*, "names")= chr [1:2] "Urbain" "Rural"  
## $ hhweight : num [1:7010] 836 836 836 836 836 ...  
## ..- attr(\*, "label")= chr "Ponderation menage"  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ hhsize : num [1:7010] 9 5 3 8 5 6 9 4 6 6 ...  
## ..- attr(\*, "label")= chr "Taille menage"  
## ..- attr(\*, "format.stata")= chr "%6.2f"  
## $ eqadu1 : num [1:7010] 5.95 3.32 2.21 5.35 3.52 ...  
## ..- attr(\*, "label")= chr "Nbr adultes-equiv. FAO"  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ eqadu2 : num [1:7010] 4.56 2.85 2.03 4.18 2.85 ...  
## ..- attr(\*, "label")= chr "Nbr adultes-equiv. alt."  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ hgender : dbl+lbl [1:7010] 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,...  
## ..@ label : chr "Genre du CM"  
## ..@ format.stata: chr "%6.2f"  
## ..@ labels : Named num [1:2] 1 2  
## .. ..- attr(\*, "names")= chr [1:2] "Masculin" "Féminin"  
## $ hage : num [1:7010] 30 38 23 37 40 32 50 27 44 50 ...  
## ..- attr(\*, "label")= chr "Age du CM"  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ hmstat : dbl+lbl [1:7010] 2, 2, 2, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2, 3, 6, 2, 2, 2,...  
## ..@ label : chr "Situation famille du CM"  
## ..@ format.stata: chr "%6.2f"  
## ..@ labels : Named num [1:7] 1 2 3 4 5 6 7  
## .. ..- attr(\*, "names")= chr [1:7] "Célibataire" "Marié(e) monogame" "Marié(e) polygame" "Union libre" ...  
## $ hreligion: dbl+lbl [1:7010] 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,...  
## ..@ label : chr "Religion du CM"  
## ..@ format.stata: chr "%6.2f"  
## ..@ labels : Named num [1:5] 1 2 3 4 5  
## .. ..- attr(\*, "names")= chr [1:5] "Musulman" "Chrétien" "Animiste" "Autre Réligion" ...  
## $ hnation : dbl+lbl [1:7010] 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,...  
## ..@ label : chr "Nationalite du CM"  
## ..@ format.stata: chr "%6.2f"  
## ..@ labels : Named num [1:12] 1 2 3 4 5 6 7 8 9 10 ...  
## .. ..- attr(\*, "names")= chr [1:12] "Benin" "Burkina Faso" "Côte d'Ivoire" "Guinée Bissau" ...  
## $ halfab : dbl+lbl [1:7010] 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 1,...  
## ..@ label : chr "Alphabetisation du CM"  
## ..@ format.stata: chr "%9.0g"  
## ..@ labels : Named num [1:2] 0 1  
## .. ..- attr(\*, "names")= chr [1:2] "Non" "Oui"  
## $ heduc : dbl+lbl [1:7010] 1, 1, 1, 1, 1, 1, 1, 3, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,...  
## ..@ label : chr "Education du CM"  
## ..@ format.stata: chr "%15.0g"  
## ..@ labels : Named num [1:9] 1 2 3 4 5 6 7 8 9  
## .. ..- attr(\*, "names")= chr [1:9] "Aucun" "Maternelle" "Primaire" "Second. gl 1" ...  
## $ hdiploma : dbl+lbl [1:7010] 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,...  
## ..@ label : chr "Diplome du CM"  
## ..@ format.stata: chr "%6.2f"  
## ..@ labels : Named num [1:11] 0 1 2 3 4 5 6 7 8 9 ...  
## .. ..- attr(\*, "names")= chr [1:11] "Aucun" "CEPE" "BEPC" "CAP" ...  
## $ hhandig : dbl+lbl [1:7010] 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,...  
## ..@ label : chr "Handicap majeur CM"  
## ..@ format.stata: chr "%9.0g"  
## ..@ labels : Named num [1:2] 0 1  
## .. ..- attr(\*, "names")= chr [1:2] "Non" "Oui"  
## $ hactiv7j : dbl+lbl [1:7010] 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,...  
## ..@ label : chr "Activite 7 jours du CM"  
## ..@ format.stata: chr "%19.0g"  
## ..@ labels : Named num [1:6] 1 2 3 4 5 6  
## .. ..- attr(\*, "names")= chr [1:6] "Occupe" "Chomeur" "TF cherchant emploi" "TF cherchant pas" ...  
## $ hactiv12m: dbl+lbl [1:7010] 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,...  
## ..@ label : chr "Activite 12 mois du CM"  
## ..@ format.stata: chr "%14.0g"  
## ..@ labels : Named num [1:4] 1 2 3 4  
## .. ..- attr(\*, "names")= chr [1:4] "Occupe" "Trav. fam." "Non occupe" "Moins de 5 ans"  
## $ hbranch : dbl+lbl [1:7010] 2, 1, 3, 4, 3, 1, 1, 1, 1, 1, 2, 2, 2, ...  
## ..@ label : chr "Branche activite du CM"  
## ..@ format.stata: chr "%16.0g"  
## ..@ labels : Named num [1:11] 1 2 3 4 5 6 7 8 9 10 ...  
## .. ..- attr(\*, "names")= chr [1:11] "Agriculture" "Elevage/peche" "Indust. extr." "Autr. indust." ...  
## $ hsectins : dbl+lbl [1:7010] 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, ...  
## ..@ label : chr "Secteur instit. du CM"  
## ..@ format.stata: chr "%6.2f"  
## ..@ labels : Named num [1:6] 1 2 3 4 5 6  
## .. ..- attr(\*, "names")= chr [1:6] "Etat/Collectivités locales" "Entreprise publique/ parapublique" "Entreprise Privée" "Entreprise associative" ...  
## $ hcsp : dbl+lbl [1:7010] 9, 9, 4, 9, 4, 9, 9, 9, 9, 9, 9, 9, 9, ...  
## ..@ label : chr "CSP du CM"  
## ..@ format.stata: chr "%6.2f"  
## ..@ labels : Named num [1:10] 1 2 3 4 5 6 7 8 9 10  
## .. ..- attr(\*, "names")= chr [1:10] "Cadre supérieur" "Cadre moyen/agent de maîtrise" "Ouvrier ou employé qualifié" "Ouvrier ou employé non qualifié" ...  
## $ dali : num [1:7010] 436192 315736 491038 1735234 1207729 ...  
## ..- attr(\*, "label")= chr "Conso annuelle alim. menage"  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ dnal : num [1:7010] 299152 498125 316298 939805 777997 ...  
## ..- attr(\*, "label")= chr "Conso annuelle non alim. menage"  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ dtot : num [1:7010] 735343 813861 807337 2675039 1985726 ...  
## ..- attr(\*, "label")= chr "Conso annuelle totale menage"  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ pcexp : num [1:7010] 90573 180440 298322 370674 440252 ...  
## ..- attr(\*, "label")= chr "Indicateur de bien-être"  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ zref : num [1:7010] 194629 194629 194629 194629 194629 ...  
## ..- attr(\*, "label")= chr "Seuil pauvrete national"  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ def\_spa : num [1:7010] 0.902 0.902 0.902 0.902 0.902 ...  
## ..- attr(\*, "label")= chr "Deflateur spatial"  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ def\_temp : num [1:7010] 0.991 0.991 0.991 0.991 0.991 ...  
## ..- attr(\*, "label")= chr "Deflateur temporel"  
## ..- attr(\*, "format.stata")= chr "%9.0g"

#base welfare21  
str(welfare21)

## tibble [7,176 × 45] (S3: tbl\_df/tbl/data.frame)  
## $ country : chr [1:7176] "bfa" "bfa" "bfa" "bfa" ...  
## ..- attr(\*, "label")= chr "Pays"  
## ..- attr(\*, "format.stata")= chr "%9s"  
## $ year : num [1:7176] 2021 2021 2021 2021 2021 ...  
## ..- attr(\*, "label")= chr "Annee enquete"  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ hhid : num [1:7176] 586005 586028 586043 586044 586052 ...  
## ..- attr(\*, "label")= chr "Idenfiant menage"  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ grappe : num [1:7176] 586 586 586 586 586 586 586 586 586 586 ...  
## ..- attr(\*, "label")= chr "Numero grappe"  
## ..- attr(\*, "format.stata")= chr "%6.0f"  
## $ menage : num [1:7176] 5 28 43 44 52 82 83 99 109 111 ...  
## ..- attr(\*, "label")= chr "Numero menage"  
## ..- attr(\*, "format.stata")= chr "%6.0f"  
## $ vague : num [1:7176] 2 2 2 2 2 2 2 2 2 2 ...  
## ..- attr(\*, "label")= chr "VAGUE"  
## ..- attr(\*, "format.stata")= chr "%6.0f"  
## $ zae : dbl+lbl [1:7176] 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,...  
## ..@ label : chr "Zone agroecologique"  
## ..@ format.stata: chr "%12.0g"  
## ..@ labels : Named num [1:5] 1 2 3 4 5  
## .. ..- attr(\*, "names")= chr [1:5] "Grand-ouest" "Grand-est" "Grand-centre" "Grand-sahel" ...  
## $ region : dbl+lbl [1:7176] 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,...  
## ..@ label : chr "Region residence"  
## ..@ format.stata: chr "%17.0f"  
## ..@ labels : Named num [1:13] 1 2 3 4 5 6 7 8 9 10 ...  
## .. ..- attr(\*, "names")= chr [1:13] "Boucle du Mouhoum" "Cascades" "Centre" "Centre-Est" ...  
## $ milieu : dbl+lbl [1:7176] 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,...  
## ..@ label : chr "Milieu residence"  
## ..@ format.stata: chr "%6.0f"  
## ..@ labels : Named num [1:2] 1 2  
## .. ..- attr(\*, "names")= chr [1:2] "Urbain" "Rural"  
## $ hhweight : num [1:7176] 439 439 439 439 439 439 439 439 439 439 ...  
## ..- attr(\*, "label")= chr "Ponderation menage"  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ hhsize : num [1:7176] 8 33 4 12 12 6 9 9 9 15 ...  
## ..- attr(\*, "label")= chr "Taille menage"  
## ..- attr(\*, "format.stata")= chr "%6.0f"  
## $ eqadu1 : num [1:7176] 6.72 23.76 2.83 8.52 9.21 ...  
## ..- attr(\*, "label")= chr "Nbr adultes-equiv. FAO"  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ eqadu2 : num [1:7176] 4.18 14.02 2.44 5.98 6.28 ...  
## ..- attr(\*, "label")= chr "Nbr adultes-equiv. alt."  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ hgender : dbl+lbl [1:7176] 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 1, 1, 1, 2, 1, 2, 1,...  
## ..@ label : chr "Genre du CM"  
## ..@ format.stata: chr "%6.0f"  
## ..@ labels : Named num [1:2] 1 2  
## .. ..- attr(\*, "names")= chr [1:2] "Masculin" "Féminin"  
## $ hage : num [1:7176] 45 45 31 54 40 57 46 37 66 41 ...  
## ..- attr(\*, "label")= chr "Age du CM"  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ hmstat : dbl+lbl [1:7176] 2, 3, 2, 2, 3, 2, 2, 2, 3, 3, 5, 2, 2, 3, 5, 2, 5, 2,...  
## ..@ label : chr "Situation famille du CM"  
## ..@ format.stata: chr "%6.0f"  
## ..@ labels : Named num [1:7] 1 2 3 4 5 6 7  
## .. ..- attr(\*, "names")= chr [1:7] "Célibataire" "Marié(e) monogame" "Marié(e) polygame" "Union libre" ...  
## $ hreligion: dbl+lbl [1:7176] 1, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 3, 2, 2, 3, 3, 3,...  
## ..@ label : chr "Religion du CM"  
## ..@ format.stata: chr "%6.0f"  
## ..@ labels : Named num [1:5] 1 2 3 4 5  
## .. ..- attr(\*, "names")= chr [1:5] "Musulman" "Chrétien" "Animiste" "Autre Réligion" ...  
## $ hnation : dbl+lbl [1:7176] 2, 2, 2, 2, 2, 2, 2, 10, 2, 2, 2, 2, 2, ...  
## ..@ label : chr "Nationalite du CM"  
## ..@ format.stata: chr "%6.0f"  
## ..@ labels : Named num [1:17] 1 2 3 4 5 6 7 8 9 10 ...  
## .. ..- attr(\*, "names")= chr [1:17] "Bénin" "Burkina Faso" "Cape-vert" "Cote d'ivoire" ...  
## $ hethnie : dbl+lbl [1:7176] 1, 1, 10, 11, 11, 1, 1, NA, 11, 11, 11, 1, 10, 1...  
## ..@ label : chr "Ethnie du CM"  
## ..@ format.stata: chr "%6.0f"  
## ..@ labels : Named num [1:12] 1 2 3 4 5 6 7 8 9 10 ...  
## .. ..- attr(\*, "names")= chr [1:12] "Mossi" "Peul" "Gourmantché" "Gourounsi" ...  
## $ halfa : dbl+lbl [1:7176] 0, 0, 0, 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0,...  
## ..@ label : chr "Alpha. lire/ecr. CM"  
## ..@ format.stata: chr "%9.0g"  
## ..@ labels : Named num [1:2] 0 1  
## .. ..- attr(\*, "names")= chr [1:2] "Non" "Oui"  
## $ halfa2 : dbl+lbl [1:7176] 0, 0, 0, 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0,...  
## ..@ label : chr "Alpha. lire/ecr./comp. CM"  
## ..@ format.stata: chr "%9.0g"  
## ..@ labels : Named num [1:2] 0 1  
## .. ..- attr(\*, "names")= chr [1:2] "Non" "Oui"  
## $ heduc : dbl+lbl [1:7176] 1, 1, 1, 1, 1, 1, 1, 7, 3, 1, 1, 9, 1, 1, 1, 1, 1, 1,...  
## ..@ label : chr "Education du CM"  
## ..@ format.stata: chr "%23.0g"  
## ..@ labels : Named num [1:9] 1 2 3 4 5 6 7 8 9  
## .. ..- attr(\*, "names")= chr [1:9] "Aucun" "Maternelle" "Primaire" "Postprimaire général" ...  
## $ hdiploma : dbl+lbl [1:7176] 0, 0, 0, 0, 0, 0, 0, 4, 0, 0, 0, 6, 0, 0, 0, 0, 0, 0,...  
## ..@ label : chr "Diplome du CM"  
## ..@ format.stata: chr "%6.0f"  
## ..@ labels : Named num [1:11] 0 1 2 3 4 5 6 7 8 9 ...  
## .. ..- attr(\*, "names")= chr [1:11] "Aucun" "CEPE" "BEPC" "CAP" ...  
## $ hhandig : dbl+lbl [1:7176] 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,...  
## ..@ label : chr "Handicap majeur CM"  
## ..@ format.stata: chr "%9.0g"  
## ..@ labels : Named num [1:2] 0 1  
## .. ..- attr(\*, "names")= chr [1:2] "Non" "Oui"  
## $ hactiv7j : dbl+lbl [1:7176] 1, 1, 1, 4, 1, 1, 1, 1, 5, 1, 5, 1, 1, 1, 1, 1, 1, 1,...  
## ..@ label : chr "Activite 7 jours du CM"  
## ..@ format.stata: chr "%19.0g"  
## ..@ labels : Named num [1:6] 1 2 3 4 5 6  
## .. ..- attr(\*, "names")= chr [1:6] "Occupe" "TF cherchant emploi" "TF cherchant pas" "Chomeur" ...  
## $ hactiv12m: dbl+lbl [1:7176] 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 3, 1, 1, 1, 1, 1, 1, 1,...  
## ..@ label : chr "Activite 12 mois du CM"  
## ..@ format.stata: chr "%14.0g"  
## ..@ labels : Named num [1:4] 1 2 3 4  
## .. ..- attr(\*, "names")= chr [1:4] "Occupe" "Trav. fam." "Non occupe" "Moins de 5 ans"  
## $ hbranch : dbl+lbl [1:7176] 10, 10, 6, 1, 11, 10, 11, 11, 1, 1, NA, 9, 1, ...  
## ..@ label : chr "Branche activite du CM"  
## ..@ format.stata: chr "%18.0g"  
## ..@ labels : Named num [1:11] 1 2 3 4 5 6 7 8 9 10 ...  
## .. ..- attr(\*, "names")= chr [1:11] "Agriculture" "Elevage/syl./peche" "Indust. extr." "Autr. indust." ...  
## $ hsectins : dbl+lbl [1:7176] 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, NA, 1, 3, ...  
## ..@ label : chr "Secteur instit. du CM"  
## ..@ format.stata: chr "%6.0f"  
## ..@ labels : Named num [1:6] 1 2 3 4 5 6  
## .. ..- attr(\*, "names")= chr [1:6] "Etat/Collectivités locales" "Entreprise publique/ parapublique" "Entreprise Privée" "Entreprise associative" ...  
## $ hcsp : dbl+lbl [1:7176] 9, 10, 9, 9, 9, 9, 9, 9, 9, 9, NA, 3, 9, ...  
## ..@ label : chr "CSP du CM"  
## ..@ format.stata: chr "%6.0f"  
## ..@ labels : Named num [1:10] 1 2 3 4 5 6 7 8 9 10  
## .. ..- attr(\*, "names")= chr [1:10] "Cadre supérieur" "Cadre moyen/agent de maîtrise" "Ouvrier ou employé qualifié" "Ouvrier ou employé non qualifié" ...  
## $ dali : num [1:7176] 1224618 5900909 524436 564769 502364 ...  
## ..- attr(\*, "label")= chr "Conso annuelle alim. menage"  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ dnal : num [1:7176] 1275733 5590681 296808 435736 461143 ...  
## ..- attr(\*, "label")= chr "Conso annuelle non alim. menage"  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ dtot : num [1:7176] 2500352 11491589 821244 1000505 963507 ...  
## ..- attr(\*, "label")= chr "Conso annuelle totale menage"  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ dtet : num [1:7176] 312544 348230 205311 83375 80292 ...  
## ..- attr(\*, "label")= chr "Conso annuelle par personne"  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ pcexp : num [1:7176] 330982 368773 217423 88294 85029 ...  
## ..- attr(\*, "label")= chr "Indicateur de bien-être"  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ zref : num [1:7176] 247806 247806 247806 247806 247806 ...  
## ..- attr(\*, "label")= chr "Seuil pauvrete national"  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ def\_spa : num [1:7176] 0.944 0.944 0.944 0.944 0.944 ...  
## ..- attr(\*, "label")= chr "Deflateur spatial"  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ def\_temp : num [1:7176] 1.06 1.06 1.06 1.06 1.06 ...  
## ..- attr(\*, "label")= chr "Deflateur temporel"  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ zaemil : num [1:7176] 2 2 2 2 2 2 2 2 2 2 ...  
## ..- attr(\*, "label")= chr "zae selon le milieu de résidence"  
## ..- attr(\*, "format.stata")= chr "%21.0g"  
## $ dif : num [1:7176] -83176 -120967 30383 159512 162777 ...  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ p0 : num [1:7176] 0 0 100 100 100 100 100 0 100 100 ...  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ p1 : num [1:7176] 0 0 12.3 64.4 65.7 ...  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ p2 : num [1:7176] 0 0 1.5 41.4 43.1 ...  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ pauv : num [1:7176] 0 0 1 1 1 1 1 0 1 1 ...  
## ..- attr(\*, "format.stata")= chr "%9.0g"  
## $ quintiles: num [1:7176] 4 4 2 1 1 2 1 5 1 1 ...  
## ..- attr(\*, "label")= chr "5 quantiles of pcexp"  
## ..- attr(\*, "format.stata")= chr "%8.0g"  
## $ deciles : num [1:7176] 7 7 4 1 1 3 1 10 1 1 ...  
## ..- attr(\*, "label")= chr "10 quantiles of pcexp"  
## ..- attr(\*, "format.stata")= chr "%8.0g"

#base welfare18  
summary(welfare18)

## country year hhid grappe   
## Length:7010 Min. :2018 Min. : 1009 Min. : 1   
## Class :character 1st Qu.:2018 1st Qu.:147031 1st Qu.:147   
## Mode :character Median :2018 Median :293098 Median :293   
## Mean :2018 Mean :293038 Mean :293   
## 3rd Qu.:2018 3rd Qu.:439136 3rd Qu.:439   
## Max. :2018 Max. :585102 Max. :585   
##   
## menage vague zae region   
## Min. : 1.00 Min. :1.0 Min. :1.000 Min. : 1.000   
## 1st Qu.: 36.00 1st Qu.:1.0 1st Qu.:1.000 1st Qu.: 3.000   
## Median : 73.00 Median :1.0 Median :2.000 Median : 7.000   
## Mean : 85.67 Mean :1.5 Mean :2.476 Mean : 6.746   
## 3rd Qu.:119.00 3rd Qu.:2.0 3rd Qu.:4.000 3rd Qu.:10.000   
## Max. :527.00 Max. :2.0 Max. :5.000 Max. :13.000   
##   
## milieu hhweight hhsize eqadu1   
## Min. :1.000 Min. : 21.4 Min. : 1.000 Min. : 0.660   
## 1st Qu.:1.000 1st Qu.: 129.1 1st Qu.: 4.000 1st Qu.: 2.830   
## Median :2.000 Median : 336.4 Median : 6.000 Median : 4.210   
## Mean :1.551 Mean : 443.3 Mean : 6.507 Mean : 4.818   
## 3rd Qu.:2.000 3rd Qu.: 614.3 3rd Qu.: 8.000 3rd Qu.: 5.990   
## Max. :2.000 Max. :3355.6 Max. :51.000 Max. :37.205   
##   
## eqadu2 hgender hage hmstat   
## Min. : 1.000 Min. :1.000 Min. : 16.00 Min. :1.000   
## 1st Qu.: 2.445 1st Qu.:1.000 1st Qu.: 35.00 1st Qu.:2.000   
## Median : 3.246 Median :1.000 Median : 45.00 Median :2.000   
## Mean : 3.568 Mean :1.149 Mean : 46.69 Mean :2.566   
## 3rd Qu.: 4.333 3rd Qu.:1.000 3rd Qu.: 57.00 3rd Qu.:3.000   
## Max. :21.990 Max. :2.000 Max. :100.00 Max. :7.000   
##   
## hreligion hnation halfab heduc   
## Min. :1.000 Min. : 1.000 Min. :0.0000 Min. :1.000   
## 1st Qu.:1.000 1st Qu.: 2.000 1st Qu.:0.0000 1st Qu.:1.000   
## Median :1.000 Median : 2.000 Median :0.0000 Median :1.000   
## Mean :1.488 Mean : 2.021 Mean :0.4006 Mean :2.091   
## 3rd Qu.:2.000 3rd Qu.: 2.000 3rd Qu.:1.0000 3rd Qu.:3.000   
## Max. :5.000 Max. :12.000 Max. :1.0000 Max. :9.000   
##   
## hdiploma hhandig hactiv7j hactiv12m   
## Min. : 0.0000 Min. :0.00000 Min. :1.000 Min. :1.000   
## 1st Qu.: 0.0000 1st Qu.:0.00000 1st Qu.:1.000 1st Qu.:1.000   
## Median : 0.0000 Median :0.00000 Median :1.000 Median :1.000   
## Mean : 0.5427 Mean :0.05835 Mean :1.457 Mean :1.168   
## 3rd Qu.: 0.0000 3rd Qu.:0.00000 3rd Qu.:1.000 3rd Qu.:1.000   
## Max. :10.0000 Max. :1.00000 Max. :5.000 Max. :3.000   
##   
## hbranch hsectins hcsp dali   
## Min. : 1.00 Min. :1.000 Min. : 1.000 Min. : 17857   
## 1st Qu.: 1.00 1st Qu.:3.000 1st Qu.: 9.000 1st Qu.: 467192   
## Median : 1.00 Median :3.000 Median : 9.000 Median : 732191   
## Mean : 3.57 Mean :2.894 Mean : 7.892 Mean : 942295   
## 3rd Qu.: 6.00 3rd Qu.:3.000 3rd Qu.: 9.000 3rd Qu.: 1171909   
## Max. :11.00 Max. :6.000 Max. :10.000 Max. :12597247   
## NA's :549 NA's :549 NA's :549   
## dnal dtot pcexp zref   
## Min. : 58281 Min. : 83213 Min. : 40675 Min. :194629   
## 1st Qu.: 501631 1st Qu.: 1029927 1st Qu.: 179173 1st Qu.:194629   
## Median : 818381 Median : 1586158 Median : 274952 Median :194629   
## Mean : 1104332 Mean : 2046627 Mean : 387040 Mean :194629   
## 3rd Qu.: 1323886 3rd Qu.: 2517495 3rd Qu.: 445358 3rd Qu.:194629   
## Max. :10083404 Max. :18610512 Max. :10279718 Max. :194629   
##   
## def\_spa def\_temp   
## Min. :0.9021 Min. :0.9809   
## 1st Qu.:0.9266 1st Qu.:0.9912   
## Median :0.9650 Median :1.0028   
## Mean :1.0242 Mean :1.0019   
## 3rd Qu.:1.1300 3rd Qu.:1.0182   
## Max. :1.2763 Max. :1.0297   
##

#base welfare21  
summary(welfare21)

## country year hhid grappe   
## Length:7176 Min. :2021 Min. : 586005 Min. : 586.0   
## Class :character 1st Qu.:2021 1st Qu.: 735276 1st Qu.: 735.0   
## Mode :character Median :2021 Median : 885067 Median : 885.0   
## Mean :2021 Mean : 885191 Mean : 885.1   
## 3rd Qu.:2021 3rd Qu.:1035042 3rd Qu.:1035.0   
## Max. :2021 Max. :1185051 Max. :1185.0   
##   
## menage vague zae region milieu   
## Min. : 1.00 Min. :1.0 Min. :1.00 Min. : 1.000 Min. :1.000   
## 1st Qu.: 28.00 1st Qu.:1.0 1st Qu.:1.00 1st Qu.: 3.000 1st Qu.:1.000   
## Median : 58.00 Median :2.0 Median :2.00 Median : 6.000 Median :2.000   
## Mean : 73.15 Mean :1.5 Mean :2.41 Mean : 6.607 Mean :1.535   
## 3rd Qu.: 99.00 3rd Qu.:2.0 3rd Qu.:4.00 3rd Qu.:10.000 3rd Qu.:2.000   
## Max. :653.00 Max. :2.0 Max. :5.00 Max. :13.000 Max. :2.000   
##   
## hhweight hhsize eqadu1 eqadu2   
## Min. : 16.0 Min. : 1.000 Min. : 0.660 Min. : 1.000   
## 1st Qu.: 134.0 1st Qu.: 4.000 1st Qu.: 2.830 1st Qu.: 2.445   
## Median : 343.0 Median : 6.000 Median : 4.210 Median : 3.246   
## Mean : 450.5 Mean : 6.437 Mean : 4.785 Mean : 3.543   
## 3rd Qu.: 641.0 3rd Qu.: 8.000 3rd Qu.: 5.949 3rd Qu.: 4.257   
## Max. :3154.0 Max. :43.000 Max. :32.060 Max. :19.030   
##   
## hgender hage hmstat hreligion   
## Min. :1.00 Min. : 16.00 Min. :1.000 Min. :1.000   
## 1st Qu.:1.00 1st Qu.: 34.00 1st Qu.:2.000 1st Qu.:1.000   
## Median :1.00 Median : 43.00 Median :2.000 Median :1.000   
## Mean :1.15 Mean : 45.68 Mean :2.551 Mean :1.437   
## 3rd Qu.:1.00 3rd Qu.: 56.00 3rd Qu.:3.000 3rd Qu.:2.000   
## Max. :2.00 Max. :114.00 Max. :7.000 Max. :5.000   
##   
## hnation hethnie halfa halfa2   
## Min. : 1.000 Min. : 1.000 Min. :0.000 Min. :0.0000   
## 1st Qu.: 2.000 1st Qu.: 1.000 1st Qu.:0.000 1st Qu.:0.0000   
## Median : 2.000 Median : 1.000 Median :0.000 Median :0.0000   
## Mean : 2.055 Mean : 3.571 Mean :0.436 Mean :0.4299   
## 3rd Qu.: 2.000 3rd Qu.: 6.000 3rd Qu.:1.000 3rd Qu.:1.0000   
## Max. :18.000 Max. :12.000 Max. :1.000 Max. :1.0000   
## NA's :66   
## heduc hdiploma hhandig hactiv7j   
## Min. :1.000 Min. : 0.0000 Min. :0.00000 Min. :1.000   
## 1st Qu.:1.000 1st Qu.: 0.0000 1st Qu.:0.00000 1st Qu.:1.000   
## Median :1.000 Median : 0.0000 Median :0.00000 Median :1.000   
## Mean :2.255 Mean : 0.6361 Mean :0.05755 Mean :1.638   
## 3rd Qu.:3.000 3rd Qu.: 0.0000 3rd Qu.:0.00000 3rd Qu.:1.000   
## Max. :9.000 Max. :10.0000 Max. :1.00000 Max. :5.000   
##   
## hactiv12m hbranch hsectins hcsp   
## Min. :1.000 Min. : 1.000 Min. :1.000 Min. : 1.000   
## 1st Qu.:1.000 1st Qu.: 1.000 1st Qu.:3.000 1st Qu.: 9.000   
## Median :1.000 Median : 3.000 Median :3.000 Median : 9.000   
## Mean :1.264 Mean : 4.065 Mean :2.884 Mean : 7.625   
## 3rd Qu.:1.000 3rd Qu.: 6.000 3rd Qu.:3.000 3rd Qu.: 9.000   
## Max. :3.000 Max. :11.000 Max. :6.000 Max. :10.000   
## NA's :939 NA's :938 NA's :938   
## dali dnal dtot dtet   
## Min. : 40986 Min. : 34233 Min. : 155415 Min. : 42564   
## 1st Qu.: 601549 1st Qu.: 563463 1st Qu.: 1246253 1st Qu.: 218658   
## Median : 914039 Median : 947706 Median : 1912858 Median : 332757   
## Mean : 1127436 Mean : 1252953 Mean : 2380389 Mean : 468194   
## 3rd Qu.: 1403725 3rd Qu.: 1543807 3rd Qu.: 2907424 3rd Qu.: 554304   
## Max. :10465574 Max. :20629712 Max. :29525314 Max. :7242476   
##   
## pcexp zref def\_spa def\_temp   
## Min. : 41298 Min. :247806 Min. :0.9443 Min. :0.8757   
## 1st Qu.: 225605 1st Qu.:247806 1st Qu.:0.9517 1st Qu.:0.9360   
## Median : 341007 Median :247806 Median :0.9693 Median :0.9885   
## Mean : 475346 Mean :247806 Mean :0.9795 Mean :0.9825   
## 3rd Qu.: 562288 3rd Qu.:247806 3rd Qu.:0.9829 3rd Qu.:1.0369   
## Max. :7508059 Max. :247806 Max. :1.0870 Max. :1.0734   
##   
## zaemil dif p0 p1   
## Min. :1.000 Min. :-7260254 Min. : 0.00 Min. : 0.000   
## 1st Qu.:2.000 1st Qu.: -314482 1st Qu.: 0.00 1st Qu.: 0.000   
## Median :4.000 Median : -93201 Median : 0.00 Median : 0.000   
## Mean :4.354 Mean : -227540 Mean : 29.91 Mean : 8.621   
## 3rd Qu.:7.000 3rd Qu.: 22201 3rd Qu.:100.00 3rd Qu.: 8.959   
## Max. :9.000 Max. : 206508 Max. :100.00 Max. :83.335   
##   
## p2 pauv quintiles deciles   
## Min. : 0.0000 Min. :0.0000 Min. :1.000 Min. : 1.000   
## 1st Qu.: 0.0000 1st Qu.:0.0000 1st Qu.:2.000 1st Qu.: 4.000   
## Median : 0.0000 Median :0.0000 Median :4.000 Median : 7.000   
## Mean : 3.5056 Mean :0.2991 Mean :3.481 Mean : 6.506   
## 3rd Qu.: 0.8026 3rd Qu.:1.0000 3rd Qu.:5.000 3rd Qu.: 9.000   
## Max. :69.4466 Max. :1.0000 Max. :5.000 Max. :10.000   
##

## 3-Les dimensions des bases et les noms des variables

### Dimension

#base welfare18  
print("La base welfare18 a respestivement d'observation et de variable")

## [1] "La base welfare18 a respestivement d'observation et de variable"

dim(welfare18)

## [1] 7010 34

#base welfare21  
print("La base welfare21 a respestivement d'observation et de variable")

## [1] "La base welfare21 a respestivement d'observation et de variable"

dim(welfare21)

## [1] 7176 45

### Noms des variables

print("Les noms des variables de la base welfare18 sont :")

## [1] "Les noms des variables de la base welfare18 sont :"

#base welfare18  
names(welfare18)

## [1] "country" "year" "hhid" "grappe" "menage" "vague"   
## [7] "zae" "region" "milieu" "hhweight" "hhsize" "eqadu1"   
## [13] "eqadu2" "hgender" "hage" "hmstat" "hreligion" "hnation"   
## [19] "halfab" "heduc" "hdiploma" "hhandig" "hactiv7j" "hactiv12m"  
## [25] "hbranch" "hsectins" "hcsp" "dali" "dnal" "dtot"   
## [31] "pcexp" "zref" "def\_spa" "def\_temp"

print("Les noms des variables de la base welfare21 sont :")

## [1] "Les noms des variables de la base welfare21 sont :"

#base welfare21  
names(welfare21)

## [1] "country" "year" "hhid" "grappe" "menage" "vague"   
## [7] "zae" "region" "milieu" "hhweight" "hhsize" "eqadu1"   
## [13] "eqadu2" "hgender" "hage" "hmstat" "hreligion" "hnation"   
## [19] "hethnie" "halfa" "halfa2" "heduc" "hdiploma" "hhandig"   
## [25] "hactiv7j" "hactiv12m" "hbranch" "hsectins" "hcsp" "dali"   
## [31] "dnal" "dtot" "dtet" "pcexp" "zref" "def\_spa"   
## [37] "def\_temp" "zaemil" "dif" "p0" "p1" "p2"   
## [43] "pauv" "quintiles" "deciles"

### Les variables dont leurs noms se diffèrent dans les deux bases

#les variables qui ne sont pas commun aux deux bases  
nam\_welfare18 <- colnames(welfare18) # extraction des noms des variables dans chaque bases  
nam\_welfare21 <- colnames(welfare21)  
# Affichages des variables non commun  
  
#comparaisons des colonnes   
if (identical(sort(nam\_welfare21)#la fonction "sort" est utilisé pour trier vecteur nam\_welfare par ordre croissante(alphab) pour faciliter la comparaison,   
 ,sort(nam\_welfare18))){  
 print("Les deux bases ont les mêmes colonnes.")  
}else{  
 print("Les deux bases n'ont pas les mêmes colonnes.")  
 #Les variables présente dans welfare18 et non dans welfare21  
 var\_non\_commun18 <- setdiff(nam\_welfare18, nam\_welfare21)  
 #Les variables présente dans welfare21 et non dans welfare18  
 var\_non\_commun21 <- setdiff(nam\_welfare21, nam\_welfare18)  
   
 cat("Les variables présente dans welfare18 et non dans welfare21: ", var\_non\_commun18)# Affichage des variables présente dans welfare18 et non dans welfare21  
   
 cat("\n Les variables présente dans welfare21 et non dans welfare18: ", var\_non\_commun21)# Affichage des variables présente dans welfare21 et non dans welfare18  
}

## [1] "Les deux bases n'ont pas les mêmes colonnes."  
## Les variables présente dans welfare18 et non dans welfare21: halfab  
## Les variables présente dans welfare21 et non dans welfare18: hethnie halfa halfa2 dtet zaemil dif p0 p1 p2 pauv quintiles deciles

## 4- Les labelles

### Les labelles des variables

print("Les labelles des variables de la base welfare18")

## [1] "Les labelles des variables de la base welfare18"

var\_label(welfare18)

## $country  
## [1] "Pays"  
##   
## $year  
## [1] "Annee enquete"  
##   
## $hhid  
## [1] "Idenfiant menage"  
##   
## $grappe  
## [1] "Numero grappe"  
##   
## $menage  
## [1] "Numero menage"  
##   
## $vague  
## [1] "Vague"  
##   
## $zae  
## [1] "Zone agroecologique"  
##   
## $region  
## [1] "Region residence"  
##   
## $milieu  
## [1] "Milieu residence"  
##   
## $hhweight  
## [1] "Ponderation menage"  
##   
## $hhsize  
## [1] "Taille menage"  
##   
## $eqadu1  
## [1] "Nbr adultes-equiv. FAO"  
##   
## $eqadu2  
## [1] "Nbr adultes-equiv. alt."  
##   
## $hgender  
## [1] "Genre du CM"  
##   
## $hage  
## [1] "Age du CM"  
##   
## $hmstat  
## [1] "Situation famille du CM"  
##   
## $hreligion  
## [1] "Religion du CM"  
##   
## $hnation  
## [1] "Nationalite du CM"  
##   
## $halfab  
## [1] "Alphabetisation du CM"  
##   
## $heduc  
## [1] "Education du CM"  
##   
## $hdiploma  
## [1] "Diplome du CM"  
##   
## $hhandig  
## [1] "Handicap majeur CM"  
##   
## $hactiv7j  
## [1] "Activite 7 jours du CM"  
##   
## $hactiv12m  
## [1] "Activite 12 mois du CM"  
##   
## $hbranch  
## [1] "Branche activite du CM"  
##   
## $hsectins  
## [1] "Secteur instit. du CM"  
##   
## $hcsp  
## [1] "CSP du CM"  
##   
## $dali  
## [1] "Conso annuelle alim. menage"  
##   
## $dnal  
## [1] "Conso annuelle non alim. menage"  
##   
## $dtot  
## [1] "Conso annuelle totale menage"  
##   
## $pcexp  
## [1] "Indicateur de bien-être"  
##   
## $zref  
## [1] "Seuil pauvrete national"  
##   
## $def\_spa  
## [1] "Deflateur spatial"  
##   
## $def\_temp  
## [1] "Deflateur temporel"

print("Les labelles des variables de la base welfare21")

## [1] "Les labelles des variables de la base welfare21"

var\_label(welfare21)

## $country  
## [1] "Pays"  
##   
## $year  
## [1] "Annee enquete"  
##   
## $hhid  
## [1] "Idenfiant menage"  
##   
## $grappe  
## [1] "Numero grappe"  
##   
## $menage  
## [1] "Numero menage"  
##   
## $vague  
## [1] "VAGUE"  
##   
## $zae  
## [1] "Zone agroecologique"  
##   
## $region  
## [1] "Region residence"  
##   
## $milieu  
## [1] "Milieu residence"  
##   
## $hhweight  
## [1] "Ponderation menage"  
##   
## $hhsize  
## [1] "Taille menage"  
##   
## $eqadu1  
## [1] "Nbr adultes-equiv. FAO"  
##   
## $eqadu2  
## [1] "Nbr adultes-equiv. alt."  
##   
## $hgender  
## [1] "Genre du CM"  
##   
## $hage  
## [1] "Age du CM"  
##   
## $hmstat  
## [1] "Situation famille du CM"  
##   
## $hreligion  
## [1] "Religion du CM"  
##   
## $hnation  
## [1] "Nationalite du CM"  
##   
## $hethnie  
## [1] "Ethnie du CM"  
##   
## $halfa  
## [1] "Alpha. lire/ecr. CM"  
##   
## $halfa2  
## [1] "Alpha. lire/ecr./comp. CM"  
##   
## $heduc  
## [1] "Education du CM"  
##   
## $hdiploma  
## [1] "Diplome du CM"  
##   
## $hhandig  
## [1] "Handicap majeur CM"  
##   
## $hactiv7j  
## [1] "Activite 7 jours du CM"  
##   
## $hactiv12m  
## [1] "Activite 12 mois du CM"  
##   
## $hbranch  
## [1] "Branche activite du CM"  
##   
## $hsectins  
## [1] "Secteur instit. du CM"  
##   
## $hcsp  
## [1] "CSP du CM"  
##   
## $dali  
## [1] "Conso annuelle alim. menage"  
##   
## $dnal  
## [1] "Conso annuelle non alim. menage"  
##   
## $dtot  
## [1] "Conso annuelle totale menage"  
##   
## $dtet  
## [1] "Conso annuelle par personne"  
##   
## $pcexp  
## [1] "Indicateur de bien-être"  
##   
## $zref  
## [1] "Seuil pauvrete national"  
##   
## $def\_spa  
## [1] "Deflateur spatial"  
##   
## $def\_temp  
## [1] "Deflateur temporel"  
##   
## $zaemil  
## [1] "zae selon le milieu de résidence"  
##   
## $dif  
## NULL  
##   
## $p0  
## NULL  
##   
## $p1  
## NULL  
##   
## $p2  
## NULL  
##   
## $pauv  
## NULL  
##   
## $quintiles  
## [1] "5 quantiles of pcexp"  
##   
## $deciles  
## [1] "10 quantiles of pcexp"

### Les labelles des valeurs

print("Les labelles des valeurs de la base welfare18")

## [1] "Les labelles des valeurs de la base welfare18"

val\_labels(welfare18)

## $country  
## NULL  
##   
## $year  
## NULL  
##   
## $hhid  
## NULL  
##   
## $grappe  
## NULL  
##   
## $menage  
## NULL  
##   
## $vague  
## NULL  
##   
## $zae  
## Grand-ouest Grand-est Grand-centre Grand-sahel Ouaga   
## 1 2 3 4 5   
##   
## $region  
## Boucle du Mouhoun Cascades Centre Centre-Est   
## 1 2 3 4   
## Centre-Nord Centre-Ouest Centre-Sud Est   
## 5 6 7 8   
## Hauts Bassins Nord Plateau-Central Sahel   
## 9 10 11 12   
## Sud-Ouest   
## 13   
##   
## $milieu  
## Urbain Rural   
## 1 2   
##   
## $hhweight  
## NULL  
##   
## $hhsize  
## NULL  
##   
## $eqadu1  
## NULL  
##   
## $eqadu2  
## NULL  
##   
## $hgender  
## Masculin Féminin   
## 1 2   
##   
## $hage  
## NULL  
##   
## $hmstat  
## Célibataire Marié(e) monogame Marié(e) polygame Union libre   
## 1 2 3 4   
## Veuf(ve) Divorcé(e) Séparé(e)   
## 5 6 7   
##   
## $hreligion  
## Musulman Chrétien Animiste Autre Réligion Sans Réligion   
## 1 2 3 4 5   
##   
## $hnation  
## Benin Burkina Faso Côte d'Ivoire   
## 1 2 3   
## Guinée Bissau Mali Niger   
## 4 5 6   
## Sénégal Togo Nigéria   
## 7 8 9   
## Autre CEDEAO Autre Afrique Autre pays hors Afrique   
## 10 11 12   
##   
## $halfab  
## Non Oui   
## 0 1   
##   
## $heduc  
## Aucun Maternelle Primaire Second. gl 1 Second. tech. 1   
## 1 2 3 4 5   
## Second. gl 2 Second. tech. 2 Postsecondaire Superieur   
## 6 7 8 9   
##   
## $hdiploma  
## Aucun CEPE BEPC CAP BT   
## 0 1 2 3 4   
## BAC DEUG, DUT, BTS Licence Maitrise Master/DEA/DESS   
## 5 6 7 8 9   
## Doctorat/Phd   
## 10   
##   
## $hhandig  
## Non Oui   
## 0 1   
##   
## $hactiv7j  
## Occupe Chomeur TF cherchant emploi TF cherchant pas   
## 1 2 3 4   
## Inactif Moins de 5 ans   
## 5 6   
##   
## $hactiv12m  
## Occupe Trav. fam. Non occupe Moins de 5 ans   
## 1 2 3 4   
##   
## $hbranch  
## Agriculture Elevage/peche Indust. extr. Autr. indust.   
## 1 2 3 4   
## BTP Commerce Restaurant/Hotel Trans./Comm.   
## 5 6 7 8   
## Education/Sante Services perso. Aut. services   
## 9 10 11   
##   
## $hsectins  
## Etat/Collectivités locales   
## 1   
## Entreprise publique/ parapublique   
## 2   
## Entreprise Privée   
## 3   
## Entreprise associative   
## 4   
## Ménage comme employeur de personnel domestique   
## 5   
## Organisme international /Ambassade   
## 6   
##   
## $hcsp  
## Cadre supérieur Cadre moyen/agent de maîtrise   
## 1 2   
## Ouvrier ou employé qualifié Ouvrier ou employé non qualifié   
## 3 4   
## Manœuvre, aide ménagère Stagiaire ou Apprenti rénuméré   
## 5 6   
## Stagiaire ou Apprenti non rénuméré Aide familial   
## 7 8   
## Travailleur pour compte propre Patron   
## 9 10   
##   
## $dali  
## NULL  
##   
## $dnal  
## NULL  
##   
## $dtot  
## NULL  
##   
## $pcexp  
## NULL  
##   
## $zref  
## NULL  
##   
## $def\_spa  
## NULL  
##   
## $def\_temp  
## NULL

print("Les labelles des valeurs de la base welfare21")

## [1] "Les labelles des valeurs de la base welfare21"

val\_labels(welfare21)

## $country  
## NULL  
##   
## $year  
## NULL  
##   
## $hhid  
## NULL  
##   
## $grappe  
## NULL  
##   
## $menage  
## NULL  
##   
## $vague  
## NULL  
##   
## $zae  
## Grand-ouest Grand-est Grand-centre Grand-sahel Ouaga   
## 1 2 3 4 5   
##   
## $region  
## Boucle du Mouhoum Cascades Centre Centre-Est   
## 1 2 3 4   
## Centre-Nord Centre-Ouest Centre-Sud Est   
## 5 6 7 8   
## Hauts-Bassins Nord Plateau-Central Sahel   
## 9 10 11 12   
## Sud-Ouest   
## 13   
##   
## $milieu  
## Urbain Rural   
## 1 2   
##   
## $hhweight  
## NULL  
##   
## $hhsize  
## NULL  
##   
## $eqadu1  
## NULL  
##   
## $eqadu2  
## NULL  
##   
## $hgender  
## Masculin Féminin   
## 1 2   
##   
## $hage  
## NULL  
##   
## $hmstat  
## Célibataire Marié(e) monogame Marié(e) polygame Union libre   
## 1 2 3 4   
## Veuf(ve) Divorcé(e) Séparé(e)   
## 5 6 7   
##   
## $hreligion  
## Musulman Chrétien Animiste Autre Réligion Sans Réligion   
## 1 2 3 4 5   
##   
## $hnation  
## Bénin Burkina Faso Cape-vert   
## 1 2 3   
## Cote d'ivoire Gambie Ghana   
## 4 5 6   
## Guinee Guinée Bissau Liberia   
## 7 8 9   
## Mali Niger Nigeria   
## 10 11 12   
## Sénégal Serra-Leonne Togo   
## 13 14 15   
## Autre Afrique Autre pays hors Afrique   
## 17 18   
##   
## $hethnie  
## Mossi Peul Gourmantché Gourounsi Bissa   
## 1 2 3 4 5   
## Lobi Senoufo Samo Marka/Dafi Bobo   
## 6 7 8 9 10   
## Autres ethnies Naturalisé   
## 11 12   
##   
## $halfa  
## Non Oui   
## 0 1   
##   
## $halfa2  
## Non Oui   
## 0 1   
##   
## $heduc  
## Aucun Maternelle Primaire   
## 1 2 3   
## Postprimaire général Postprimaire technique Secondaire général   
## 4 5 6   
## Secondaire technique Postsecondaire Superieur   
## 7 8 9   
##   
## $hdiploma  
## Aucun CEPE BEPC CAP BT   
## 0 1 2 3 4   
## BAC DEUG, DUT, BTS Licence Maitrise Master/DEA/DESS   
## 5 6 7 8 9   
## Doctorat/Phd   
## 10   
##   
## $hhandig  
## Non Oui   
## 0 1   
##   
## $hactiv7j  
## Occupe TF cherchant emploi TF cherchant pas Chomeur   
## 1 2 3 4   
## Inactif Moins de 5 ans   
## 5 6   
##   
## $hactiv12m  
## Occupe Trav. fam. Non occupe Moins de 5 ans   
## 1 2 3 4   
##   
## $hbranch  
## Agriculture Elevage/syl./peche Indust. extr. Autr. indust.   
## 1 2 3 4   
## BTP Commerce Restaurant/Hotel Trans./Comm.   
## 5 6 7 8   
## Education/Sante Services perso. Aut. services   
## 9 10 11   
##   
## $hsectins  
## Etat/Collectivités locales   
## 1   
## Entreprise publique/ parapublique   
## 2   
## Entreprise Privée   
## 3   
## Entreprise associative   
## 4   
## Ménage comme employeur de personnel domestique   
## 5   
## Organisme international /Ambassade   
## 6   
##   
## $hcsp  
## Cadre supérieur   
## 1   
## Cadre moyen/agent de maîtrise   
## 2   
## Ouvrier ou employé qualifié   
## 3   
## Ouvrier ou employé non qualifié   
## 4   
## Manœuvre, aide ménagère   
## 5   
## Stagiaire ou Apprenti rénuméré   
## 6   
## Stagiaire ou Apprenti non rénuméré   
## 7   
## Travailleur Familial contribuant pour une entreprise familial   
## 8   
## Travailleur pour compte propre   
## 9   
## Patron   
## 10   
##   
## $dali  
## NULL  
##   
## $dnal  
## NULL  
##   
## $dtot  
## NULL  
##   
## $dtet  
## NULL  
##   
## $pcexp  
## NULL  
##   
## $zref  
## NULL  
##   
## $def\_spa  
## NULL  
##   
## $def\_temp  
## NULL  
##   
## $zaemil  
## NULL  
##   
## $dif  
## NULL  
##   
## $p0  
## NULL  
##   
## $p1  
## NULL  
##   
## $p2  
## NULL  
##   
## $pauv  
## NULL  
##   
## $quintiles  
## NULL  
##   
## $deciles  
## NULL

### Verification de l’identité des labelle des variables.

# Extraction des labelles des deux bases  
labels\_welfare18 <- sapply(welfare18, function(x) attr(x, "label"))  
labels\_welfare21 <- sapply(welfare21, function(x) attr(x, "label"))  
  
# Comparer les labels des variables entre welfare18 et welfare21  
if(identical(labels\_welfare18, labels\_welfare21)){  
 print("les labelles des variables sont identiques")  
}else{  
 print("les labelles des variables ne sont pas identiques")  
}

## [1] "les labelles des variables ne sont pas identiques"

### Dectecter les variables de labelles différente dans les deux bases

#Fonction qui Compare les labels des valeurs entre les deux bases  
compare\_labels\_variables <- function(df1, df2) {  
 identical\_labels <- list() # a defaut, on attribut list vide à identical\_labels  
   
 # Comparer les labels des valeurs pour chaque variable  
 for (var in names(df1)) {  
 # Vérifier si les deux variables existent dans les deux bases  
 if (var %in% names(df2)) {  
 label1 <- attr(df1[[var]], "label")# on attribut le labelle de var dans df1 à label1  
 label2 <- attr(df2[[var]], "label")# on attribut le labelle de var dans df2 à label2  
   
 # Comparer les labels des valeurs  
 if (!identical(label1, label2)) {  
 identical\_labels <- append(identical\_labels, var)  
   
 }  
 }  
 }  
 if(length(identical\_labels)==0){  
 print("les labelles sont identiques pour toute les variables")  
 }else{  
 print(paste("Les labels diffèrents pour la variable:", identical\_labels))  
 }  
   
   
   
 #return(identical\_labels)  
}  
  
# Appel de la fonction  
compare\_labels\_variables(welfare18, welfare21)

## [1] "Les labels diffèrents pour la variable: vague"

### Dectection des variables dont leur valeurs ont des labelles différentes dans les deux bases

#Fonction qui Compare les labels des valeurs entre les deux bases  
compare\_labels\_values <- function(df1, df2){   
# Extraction des variables de noms communs  
variabl\_commun <- intersect(names(df1), names(df2))  
  
  
variable\_label\_diff <- list() #créer une liste vide   
  
for (variable in variabl\_commun) { #parcourir les variables en communs dans les deux bases  
   
 if(labelled::is.labelled(df2[[variable]])){ #vérifier si la variable en 2018 est labéliser  
   
 value\_label18 <- labelled::val\_labels(df2[[variable]]) #recupérer les labels de la variable en 2018  
   
 }else{  
 value\_label18 <- NULL #Mettre vide dans le cas ou la variable en 2018 n'est pas labéliser  
 }  
   
 if(labelled::is.labelled(df1[[variable]])){ #vérifier si la variable en 2021 est labéliser  
   
 value\_label21 <- labelled::val\_labels(df1[[variable]]) #recupérer les labels de la variable en 2021  
   
 }else{  
   
 value\_label21 <- NULL #Mettre vide dans le cas ou la variable en 2021 n'est pas labéliser  
 }  
   
 if(!identical(value\_label18, value\_label21)){ #Vérifier si les labels de la variable sont identiques entre 2018 et 2021  
 variable\_label\_diff <- append(variable\_label\_diff, variable) #Si les labels diffèrent, alors ajouter le nom de la variable à la liste créer  
 print(variable)  
 }  
}  
  
if(length(variable\_label\_diff)==0){  
 print("les labelles des modalité sont identiques pour toute")  
 }else{  
 print("Les labels des modalité diffèrents pour ces variables")   
 }  
   
  
}  
  
compare\_labels\_values(welfare18, welfare21)

## [1] "region"  
## [1] "hnation"  
## [1] "heduc"  
## [1] "hactiv7j"  
## [1] "hbranch"  
## [1] "hcsp"  
## [1] "Les labels des modalité diffèrents pour ces variables"

### 4- Comparaison des types de variables commun

# Extraction des colonnes communes  
colonn\_com <- intersect(names(welfare18), names(welfare21))  
  
# Initialiser les listes pour les colonnes avec types identiques et différents  
type\_com <- list()  
type\_diff <- list()  
  
# Comparaison des types des colonnes communes  
for (variabl in colonn\_com) {#parcour toutes les variables  
 # Comparaison des classes des variables  
 if (identical(class(welfare18[[variabl]]), class(welfare21[[variabl]]))) {  
 type\_com<- append(type\_com, variabl)# On affect la variabl a la liste  
 } else {  
 type\_diff <- append(type\_diff, variabl) # Ajouter la variable à la liste  
  
 #pour stocker les types (classes) des colonnes communes entre les deux dataframes   
 }  
}  
  
# Afficher les résultats  
if (length(type\_diff) == 0) {#mesure la longueur de la liste  
 print("Toutes les colonnes communes ont les mêmes classes.")  
} else {  
 print("Les colonnes suivantes ont des classes différentes entre les deux dataframes :")  
 print(type\_diff)  
}

## [1] "Toutes les colonnes communes ont les mêmes classes."

# **III-Netoyage des bases**

Dans les parties précédentes nous avons vérifier que les bases ont des *dimensions* différentes ainsi que des variales communes de labelles et de noms différents. Dans cette partie, nous allons apporter des corrections (recodage des labelles et noms), et gerer les cas des *doublons* ainsi qu’autres *incohérences*.

## 1- Les doublons

# Verification de la présence des doublons dans chaque base welfare18  
if (sum(duplicated(welfare18))==0){# Verifie le nombre de doublons dans welfare18  
 print("La base welfare 18 n'a pas de doublons")  
}else{  
 print("ll y a de doublons au nombre de")   
 print( sum(any(duplicated(welfare18))))#Renvoi le nombre de doublons  
}

## [1] "La base welfare 18 n'a pas de doublons"

# Verification de la présence des doublons dans chaque base welfare21  
if (sum(duplicated(welfare21))==0){# Verifie le nombre de doublons dans welfare21  
 print("La base welfare 21 n'a pas de doublons")  
}else{  
 print("ll y a de doublons au nombre de")   
 print( sum((duplicated(welfare21))))# Renvoi le nombre de doublons  
}

## [1] "La base welfare 21 n'a pas de doublons"

## 2-Les valeurs manquantes

#Les valeurs les valeur manquantes dans chaque colonne de la base welfare18  
print("Les valeurs manques dans la base welfare18 se repartissent par colonne comme suite:")

## [1] "Les valeurs manques dans la base welfare18 se repartissent par colonne comme suite:"

print(colSums(is.na(welfare18)))#Affiche le nombre des valeurs manquantes par colonnes

## country year hhid grappe menage vague zae region   
## 0 0 0 0 0 0 0 0   
## milieu hhweight hhsize eqadu1 eqadu2 hgender hage hmstat   
## 0 0 0 0 0 0 0 0   
## hreligion hnation halfab heduc hdiploma hhandig hactiv7j hactiv12m   
## 0 0 0 0 0 0 0 0   
## hbranch hsectins hcsp dali dnal dtot pcexp zref   
## 549 549 549 0 0 0 0 0   
## def\_spa def\_temp   
## 0 0

#Les valeurs les valeur manquantes dans chaque colonne de la base welfare21  
print("Les valeurs manques dans la base welfare21 se repartissent par colonne comme suite:")

## [1] "Les valeurs manques dans la base welfare21 se repartissent par colonne comme suite:"

print(colSums(is.na(welfare21)))#Affiche le nombre des valeurs manquantes par colonnes

## country year hhid grappe menage vague zae region   
## 0 0 0 0 0 0 0 0   
## milieu hhweight hhsize eqadu1 eqadu2 hgender hage hmstat   
## 0 0 0 0 0 0 0 0   
## hreligion hnation hethnie halfa halfa2 heduc hdiploma hhandig   
## 0 0 66 0 0 0 0 0   
## hactiv7j hactiv12m hbranch hsectins hcsp dali dnal dtot   
## 0 0 939 938 938 0 0 0   
## dtet pcexp zref def\_spa def\_temp zaemil dif p0   
## 0 0 0 0 0 0 0 0   
## p1 p2 pauv quintiles deciles   
## 0 0 0 0 0

## 3-Suppression des espaces dans les noms des variables

if (any(grepl(" ",names(welfare18))))#verification de presence d'espace dans la base welfare18  
 {  
names(welfare18) <- gsub(" ","",names(welfare18))#Remplace les vides par les lettres suivante dans le noms de la variables  
print("Espaces supprimer avec succè dans la base walfare18")  
}else {  
 print("Aucun espace trouvé la base welfare18.")  
}

## [1] "Aucun espace trouvé la base welfare18."

if (any(grepl(" ",names(welfare21))))#verification de presence d'espace dans la base welfare21  
 {  
 names(welfare21) <- gsub(" ","",names(welfare21))#Remplace les vides par les lettres suivante dans le noms de la variables  
print("Espaces supprimer avec succè dans la base walfare21")  
}else {  
 print("Aucun espace trouvé la base welfare21.")  
}

## [1] "Aucun espace trouvé la base welfare21."

## 4-Renommination des noms des variables et le recodage des labelles

### Renommination des noms des variables

Nous avons constaté en II) que la variables *halfa* et *halfab* contiennent les mêmes informations dans leurs bases respectives. Pour cela, nous allons renommer *halfa* à *halfab*

colnames(welfare18)[colnames(welfare18)=="halfab"] <- "halfa"

### Recodage des labelles

#### *Les variables*

On a constater dans la partie II que la variables *vague* et *halfa* est communes aux deux bases mais de labelles différents.

#Renommination du labelle de vague  
var\_label(welfare21$vague) <- "Vague"  
#Renommination du labelle de halfa  
var\_label(welfare21$halfa) <- "Alphabetisation du CM"

#### *Les valeurs(modalité)*

#### *Region*

# On renommes les valeurs(modalités) qui ne sont pas du même labelles   
val\_label(welfare21$region, 1) <- c("Boucle du Mouhoun")  
val\_label(welfare21$region, 9) <- c("Hauts Bassins")  
#Une autre façon de faire est de considerer tous les variables de welfare18 et les affecter à welfare21 val\_labels(welfare21$region) <-val\_labels(welfare18$region). La base welfare21 aura desormain tous les labelles de welfare18

#### *hnation*

Pour cette variable, on va recoder et regrouper les pays de la CEDEAO dans la base welfare21 et ensuite affecter les labelles de welfare18 à welfare21.

# Recodages des modalités  
welfare21 <- welfare21 %>% mutate(hnation=dplyr::recode(hnation,   
 '4'=3, # on renmplacer la valeur à gauche par celle de la droite  
 '8'=4,  
 '11'=6,  
 '10'=5,  
 '15'=8,  
 '13'=7,  
 '17'=11,  
 '18'=12,  
 '12'=9,  
 '3'=10, '5'=10, '6'=10, '7'=10, '14'=10, '9'=10  
 ))  
  
  
#On remet à jour la labellisation   
val\_labels(welfare21$hnation) <-val\_labels(welfare18$hnation)

#### *heduc*

Pour cette variables, on a juste à renommer les étiquettes (labelles) des modalités de welfare18

#Nous allons garder les labelles de welfare21  
  
val\_labels(welfare18$heduc) <- val\_labels(welfare21$heduc)

#### *hactiv7j*

Pour cette variable, on va recoder les labelles de welfare21

#Recodage des labelles  
welfare21 <- welfare21 %>% mutate(hactiv7j=dplyr::recode(hactiv7j,   
 '2'=3, # on renmplacer la valeur à gauche par celle de la droite  
 '3'=4,  
 '4'=1  
 ))  
  
#Remise à jour de la labellisation   
val\_labels(welfare21$hactiv7j) <-val\_labels(welfare18$hactiv7j)# Après avoir recoder, on donne tous les labelles des modalité de welfare18 à ceux de welfare21

#### *hbranch* et

Ici également, nous allons juste renommer les labelles de welfare21. A vu d’oiel, c’est une seule labellisation de différence, pour ne pas prendre de risque, nous allons affecter tous les labelles de walfare18 à welfare21

val\_labels(welfare21$hbranch) <-val\_labels(welfare18$hbranch)#on donne tous les labelles des modalité de welfare18 à ceux de welfare21

#### *hcsp*

val\_labels(welfare21$hcsp) <-val\_labels(welfare18$hcsp)#on donne tous les labelles des modalité de welfare18 à ceux de welfare21

# **IV-Reimportation des base et le Merge**

## 1- Sauvegarde des bases

Nous allons exporter les bases nettoyées dans le dossier donnée\_nettoyée

# On exporte les welfare dans le dossier donnée\_nettoyée en leur donnant le nom welfare final  
write\_dta(welfare18, "../Données/Donnée\_nettoyées/welfare2018\_final")  
write\_dta(welfare21, "../Données/Donnée\_nettoyées/welfare2021\_final")

## 2- réimportation des bases

Nous réimportons les bases néttoyée

# importation des bases  
welfare\_2018 <- read\_dta("../Données/Donnée\_nettoyées/welfare2018\_final")  
welfare\_2021 <- read\_dta("../Données/Donnée\_nettoyées/welfare2021\_final")

## 3-Verification

Nous allons verifier quelques corrections

#Verification de l'identité des labelles des variables avec la fonction compare\_labels\_variable  
compare\_labels\_variables(welfare\_2018, welfare\_2021)

## [1] "les labelles sont identiques pour toute les variables"

#Verification de l'identité des labelles des valeurs avec la fonction compare\_labels\_value  
compare\_labels\_values(welfare\_2018, welfare\_2021)

## [1] "les labelles des modalité sont identiques pour toute"

## 4-Le merge

Nous allons empiler les deux bases. Le merge ici consiste à un ajout d’observation car les individus sont différent pour ces deux bases.

# Nous allons fair l'empilement avec bind\_rows et nous allons appeler la nouvelle base simplement welfare  
  
welfare <- bind\_rows(welfare\_2018, welfare\_2021)

Sauvegarde la base welfare dans le dossier Donnée\_nettoyées

# On exporte les welfare dans le dossier donnée\_nettoyée en leur donnant le nom welfare final  
welfare <- read\_dta("../Données/Donnée\_nettoyées/welfare")

Réimportation de la base welfare

# On exporte les welfare dans le dossier donnée\_nettoyée en leur donnant le nom welfare final  
write\_dta(welfare, "../Données/Donnée\_nettoyées/welfare")

# **V-Statistique descriptive**

Dans cette partie nous allons faire de staistique descriptive sur quelques variables.

## 1-Milieu de residence, religion et emploi

tableau\_welfare1 <- welfare %>% labelled::to\_factor()%>%#conversion en facteur de la base welfare  
 select(hreligion, region, milieu, hcsp, hactiv7j, hactiv12m, hbranch, zae) %>%#On selection les variables à étudier  
 tbl\_summary(#Formation du tableau  
 label=list(# Liste des valeur à labelliser  
 hreligion ~ "la religion du chef de menage(CM)",   
 region ~ "La région du chef de menage",  
 milieu ~"Le milieu de résidence" ,  
 hcsp ~ "csp du CM",  
 hactiv7j ~ "l'activité du chef de menage en 7 jours",  
 hactiv12m ~ "l'activité du chef de menage en 12 mois",  
 hbranch ~ "la branche d'activité",  
 zae ~ "Zone agroécologique"  
 ),# On a labellisé les valeur à étudier  
 missing = "always",# Calcul en tenant compte des valeur manquante  
 missing\_text="valeur manquant"#Affiche un message pour signifier les valeur manquante  
) %>% modify\_header(#Modifier l'entête du tableau  
 label="statistiques descriptives") %>%as\_gt() %>%#Rend le tableau compatible à la modification, permettant d'associer des couleurs ect.  
 tab\_header(title="Statistique de milieu et emploi") %>%  
 #Titre du tableau  
 tab\_style(#Mettons la bordure en bleu pour toutes les colonnes  
 style=cell\_borders(color="blue", weight = px(4)), locations = cells\_body(columns=everything())) %>%  
 tab\_style(  
 #Mettons le fonf en bleuclaire pour toute les cellule  
 style=cell\_fill(color="lightblue"), locations = cells\_column\_labels(columns=everything())) %>%   
 tab\_style(  
 #Mettons le fonf en bleuclaire pour toute les cellule  
 style=cell\_fill(color="lightblue"), locations = cells\_body(columns=everything())) %>%   
 tab\_style(  
 #mettons le texte en marron et en gras  
 style=cell\_text(color="black", weight="bold" ), locations = cells\_body(columns=everything()))%>%  
 #Ajouter une source de donnée au tableau  
 gt:: tab\_source\_note(source\_note="souce: Donné de EHCVM 2018-2021 Burkina Faso")   
  
# Affichage tableau  
tableau\_welfare1

Table 1: Statistique de milieu et emploi

| statistiques descriptives | **N = 14,186***1* |
| --- | --- |
| **la religion du chef de menage(CM)** |  |
| **Musulman** | **8,827 (62%)** |
| **Chrétien** | **4,284 (30%)** |
| **Animiste** | **1,009 (7.1%)** |
| **Autre Réligion** | **11 (<0.1%)** |
| **Sans Réligion** | **55 (0.4%)** |
| **valeur manquant** | **0** |
| **La région du chef de menage** |  |
| **Boucle du Mouhoun** | **1,187 (8.4%)** |
| **Cascades** | **982 (6.9%)** |
| **Centre** | **1,615 (11%)** |
| **Centre-Est** | **1,224 (8.6%)** |
| **Centre-Nord** | **1,020 (7.2%)** |
| **Centre-Ouest** | **1,068 (7.5%)** |
| **Centre-Sud** | **947 (6.7%)** |
| **Est** | **1,040 (7.3%)** |
| **Hauts Bassins** | **1,451 (10%)** |
| **Nord** | **897 (6.3%)** |
| **Plateau-Central** | **791 (5.6%)** |
| **Sahel** | **909 (6.4%)** |
| **Sud-Ouest** | **1,055 (7.4%)** |
| **valeur manquant** | **0** |
| **Le milieu de résidence** |  |
| **Urbain** | **6,487 (46%)** |
| **Rural** | **7,699 (54%)** |
| **valeur manquant** | **0** |
| **csp du CM** |  |
| **Cadre supérieur** | **182 (1.4%)** |
| **Cadre moyen/agent de maîtrise** | **751 (5.9%)** |
| **Ouvrier ou employé qualifié** | **642 (5.1%)** |
| **Ouvrier ou employé non qualifié** | **879 (6.9%)** |
| **Manœuvre, aide ménagère** | **139 (1.1%)** |
| **Stagiaire ou Apprenti rénuméré** | **119 (0.9%)** |
| **Stagiaire ou Apprenti non rénuméré** | **38 (0.3%)** |
| **Aide familial** | **73 (0.6%)** |
| **Travailleur pour compte propre** | **9,589 (76%)** |
| **Patron** | **287 (2.3%)** |
| **valeur manquant** | **1,487** |
| **l'activité du chef de menage en 7 jours** |  |
| **Occupe** | **12,232 (86%)** |
| **Chomeur** | **35 (0.2%)** |
| **TF cherchant emploi** | **25 (0.2%)** |
| **TF cherchant pas** | **74 (0.5%)** |
| **Inactif** | **1,820 (13%)** |
| **Moins de 5 ans** | **0 (0%)** |
| **valeur manquant** | **0** |
| **l'activité du chef de menage en 12 mois** |  |
| **Occupe** | **12,600 (89%)** |
| **Trav. fam.** | **99 (0.7%)** |
| **Non occupe** | **1,487 (10%)** |
| **Moins de 5 ans** | **0 (0%)** |
| **valeur manquant** | **0** |
| **la branche d'activité** |  |
| **Agriculture** | **6,453 (51%)** |
| **Elevage/peche** | **402 (3.2%)** |
| **Indust. extr.** | **408 (3.2%)** |
| **Autr. indust.** | **746 (5.9%)** |
| **BTP** | **566 (4.5%)** |
| **Commerce** | **1,515 (12%)** |
| **Restaurant/Hotel** | **156 (1.2%)** |
| **Trans./Comm.** | **280 (2.2%)** |
| **Education/Sante** | **724 (5.7%)** |
| **Services perso.** | **740 (5.8%)** |
| **Aut. services** | **708 (5.6%)** |
| **valeur manquant** | **1,488** |
| **Zone agroécologique** |  |
| **Grand-ouest** | **4,675 (33%)** |
| **Grand-est** | **3,211 (23%)** |
| **Grand-centre** | **2,562 (18%)** |
| **Grand-sahel** | **2,826 (20%)** |
| **Ouaga** | **912 (6.4%)** |
| **valeur manquant** | **0** |
| *1*n (%) | |
| souce: Donné de EHCVM 2018-2021 Burkina Faso | |

On constate que les chefs de menage sont majoritairement musulmans, Soit **62%** de musulman,**30%** de chrétien. La région de centre occupe **11%** et les Hauts bassins **%10%** des menage recensés. C’est le fait que ces zones sont des zonnes urbaines. Par contre, les zones urbaines, en generale, occupe **46%** contre **54%** en milieu **rurale**. L’information intéressante est que les **76%** travaille pour leur compte propre. Mais il y a moin de cadre supérieure, soit seulement **1.4%**. On a également seulement **0.2%** de chomeur et **86%** de chef de ménage occupé. Il y a de quant même plus de **13%** de personne inactif. *51* de ces chef de ménage sont des agriculteurs et l’agricuture est **33%** pratiqué dans la **Grand Ouest**

## Niveau d’etude et identé du chef de ménage

tableau\_welfare2 <- welfare %>% labelled::to\_factor()%>%#conversion en facteur de la base welfare  
 select(hhsize, halfa, heduc, hdiploma, hgender, hmstat, hage ) %>%#On selection les variables à étudier  
 tbl\_summary(#Formation du tableau  
 label=list(# Liste des valeur à labelliser  
 hhsize~ "la taille du chef de menage(CM)",   
 halfa ~ "Le niveau d'alphabetisation du chef de menage",  
 heduc ~"Le mniveau d'étude du CM" ,  
 hdiploma ~ "Le diplome du CM",  
 hgender ~ "Le sexe du chef de menage ",  
 hmstat ~ "La statut matrimoniale du chef de menage",  
 hage ~ "L'âge duchef de melange"  
 ),# On a labellisé les valeur à étudier  
 statistic = list(hage~ "{mean}, {sd} ,{median}"),#statistique sur l'age  
 missing = "always",# Calcul en tenant compte des valeur manquante  
 missing\_text="valeur manquant"#Affiche un message pour signifier les valeur manquante  
) %>% modify\_header(#Modifier l'entête du tableau  
 label="statistiques descriptives") %>%as\_gt() %>%#Rend le tableau compatible à la modification, permettant d'associer des couleurs ect.  
 tab\_header(title="Statistique du niveau d'étude et identité") %>%  
 #Titre du tableau  
 tab\_style(#Mettons la bordure en bleu pour toutes les colonnes  
 style=cell\_borders(color="orange", weight = px(4)), locations = cells\_body(columns=everything())) %>%  
 tab\_style(  
 #Mettons le fonf en bleuclaire pour toute les cellule  
 style=cell\_fill(color="lightgreen"), locations = cells\_column\_labels(columns=everything())) %>%   
 tab\_style(  
 #Mettons le fonf en bleuclaire pour toute les cellule  
 style=cell\_fill(color="lightgreen"), locations = cells\_body(columns=everything())) %>%   
 tab\_style(  
 #mettons le texte en marron et en gras  
 style=cell\_text(color="black", weight="bold" ), locations = cells\_body(columns=everything()))%>%  
 #Ajouter une source de donnée au tableau  
 gt:: tab\_source\_note(source\_note="souce: Donné de EHCVM 2018-2021 Burkina Faso")   
  
# Affichage tableau  
tableau\_welfare2

Table 1: Statistique du niveau d'étude et identité

| statistiques descriptives | **N = 14,186***1* |
| --- | --- |
| **la taille du chef de menage(CM)** | **6.0 (4.0, 8.0)** |
| **valeur manquant** | **0** |
| **Le niveau d'alphabetisation du chef de menage** |  |
| **Non** | **8,249 (58%)** |
| **Oui** | **5,937 (42%)** |
| **valeur manquant** | **0** |
| **Le mniveau d'étude du CM** |  |
| **Aucun** | **9,576 (68%)** |
| **Maternelle** | **1 (<0.1%)** |
| **Primaire** | **2,017 (14%)** |
| **Postprimaire général** | **1,132 (8.0%)** |
| **Postprimaire technique** | **22 (0.2%)** |
| **Secondaire général** | **709 (5.0%)** |
| **Secondaire technique** | **85 (0.6%)** |
| **Postsecondaire** | **70 (0.5%)** |
| **Superieur** | **574 (4.0%)** |
| **valeur manquant** | **0** |
| **Le diplome du CM** |  |
| **Aucun** | **11,181 (79%)** |
| **CEPE** | **1,360 (9.6%)** |
| **BEPC** | **760 (5.4%)** |
| **CAP** | **43 (0.3%)** |
| **BT** | **10 (<0.1%)** |
| **BAC** | **337 (2.4%)** |
| **DEUG, DUT, BTS** | **104 (0.7%)** |
| **Licence** | **203 (1.4%)** |
| **Maitrise** | **111 (0.8%)** |
| **Master/DEA/DESS** | **68 (0.5%)** |
| **Doctorat/Phd** | **9 (<0.1%)** |
| **valeur manquant** | **0** |
| **Le sexe du chef de menage** |  |
| **Masculin** | **12,065 (85%)** |
| **Féminin** | **2,121 (15%)** |
| **valeur manquant** | **0** |
| **La statut matrimoniale du chef de menage** |  |
| **Célibataire** | **653 (4.6%)** |
| **Marié(e) monogame** | **8,587 (61%)** |
| **Marié(e) polygame** | **3,172 (22%)** |
| **Union libre** | **320 (2.3%)** |
| **Veuf(ve)** | **1,201 (8.5%)** |
| **Divorcé(e)** | **105 (0.7%)** |
| **Séparé(e)** | **148 (1.0%)** |
| **valeur manquant** | **0** |
| **L'âge duchef de melange** | **46, 15 ,44** |
| **valeur manquant** | **0** |
| *1*Median (Q1, Q3); n (%); Mean, SD ,Median | |
| souce: Donné de EHCVM 2018-2021 Burkina Faso | |

Une grande partie des chefs de ménage ne sont pas alphabathétisés, soit **58%**. Plus de **68%** n’ont pas fait l’école; seulement **4%** a un niveau supérieur. En terme de diplôme, les **79%** n’ont aucun diplôme; seulement **1.4%** ont la licence. On constace églement que **85%** des chef de ménage sont des hommes; avec la majorité marié monogame.