

Statistiques descriptives (TP4)

Aissatou Sega DIALLO

2025-02-24

Installation et chargement des packages

Cette section assure la disponibilité des bibliothèques nécessaires.

```
Packages <- c("readr", "haven", "utils", "dplyr", "labelled", "gtsummary")

for (x in Packages) {
  if (!requireNamespace(x, quietly = TRUE)) {
    install.packages(x)
  }
  library(x, character.only = TRUE)
}
```

```
## Warning: le package 'haven' a été compilé avec la version R 4.3.3
```

```
## Warning: le package 'dplyr' a été compilé avec la version R 4.3.3
```

```
## Warning: le package 'gtsummary' a été compilé avec la version R 4.3.3
```

```
##Travail avec la base ménage
```

Importation des données

Chargement de la base principale

```
## Importation de la Base ménage
```

```
Base_men <- readr::read_csv("../Données/ehcvm_menage_bfa2021.csv")
```

```
## Rows: 7176 Columns: 39
```

```
## -- Column specification -----
```

```
## Delimiter: ","
```

```
## chr (28): country, logem, mur, toit, sol, eauboi_ss, eauboi_sp, elec_ac, ele...
```

```
## dbl (11): hhid, year, grappe, menage, vague, superf, grosrum, petitrum, porc...
```

```
##
```

```
## i Use 'spec()' to retrieve the full column specification for this data.
```

```
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
Base_men
```

```
## # A tibble: 7,176 x 39
##   country   hhid year grappe menage vague logem   mur   toit   sol   eauboi_ss
##   <chr>     <dbl> <dbl> <dbl> <dbl> <dbl> <chr>   <chr> <chr> <chr> <chr>
## 1 bfa      586005 2021   586     5     2 Proprie~ Oui   Oui   Oui   Non
## 2 bfa      586028 2021   586    28     2 Proprie~ Oui   Oui   Oui   Oui
## 3 bfa      586043 2021   586    43     2 Proprie~ Oui   Oui   Oui   Non
## 4 bfa      586044 2021   586    44     2 Locatai~ Oui   Oui   Oui   Non
## 5 bfa      586052 2021   586    52     2 Locatai~ Oui   Oui   Oui   Non
## 6 bfa      586082 2021   586    82     2 Proprie~ Oui   Oui   Oui   Non
## 7 bfa      586083 2021   586    83     2 Proprie~ Oui   Oui   Oui   Non
## 8 bfa      586099 2021   586    99     2 Locatai~ Oui   Oui   Oui   Oui
## 9 bfa      586109 2021   586   109     2 Proprie~ Non   Non   Non   Non
## 10 bfa     586111 2021   586   111     2 Locatai~ Oui   Oui   Oui   Non
## # i 7,166 more rows
## # i 28 more variables: eauboi_sp <chr>, elec_ac <chr>, elec_ur <chr>,
## #   elec_ua <chr>, ordure <chr>, toilet <chr>, eva_toi <chr>, eva_eau <chr>,
## #   tv <chr>, fer <chr>, frigo <chr>, cuisin <chr>, ordin <chr>, decod <chr>,
## #   car <chr>, superf <dbl>, grosrum <dbl>, petitrum <dbl>, porc <dbl>,
## #   lapin <dbl>, volail <dbl>, sh_id_demo <chr>, sh_co_natu <chr>,
## #   sh_co_eco <chr>, sh_id_eco <chr>, sh_co_vio <chr>, sh_co_oth <chr>, ...
```

Aperçu de la structure des données

```
#Structure base
utils::str(Base_men)
```

```
## spc_tbl_ [7,176 x 39] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
## $ country   : chr [1:7176] "bfa" "bfa" "bfa" "bfa" ...
## $ hhid      : num [1:7176] 586005 586028 586043 586044 586052 ...
## $ year      : num [1:7176] 2021 2021 2021 2021 2021 ...
## $ grappe    : num [1:7176] 586 586 586 586 586 586 586 586 586 ...
## $ menage    : num [1:7176] 5 28 43 44 52 82 83 99 109 111 ...
## $ vague     : num [1:7176] 2 2 2 2 2 2 2 2 2 2 ...
## $ logem     : chr [1:7176] "Proprietaire titre" "Proprietaire titre" "Proprietaire sans titre" "Loc
## $ mur       : chr [1:7176] "Oui" "Oui" "Oui" "Oui" ...
## $ toit      : chr [1:7176] "Oui" "Oui" "Oui" "Oui" ...
## $ sol       : chr [1:7176] "Oui" "Oui" "Oui" "Oui" ...
## $ eauboi_ss : chr [1:7176] "Non" "Oui" "Non" "Non" ...
## $ eauboi_sp : chr [1:7176] "Non" "Oui" "Non" "Non" ...
## $ elec_ac   : chr [1:7176] "Oui" "Oui" "Non" "Non" ...
## $ elec_ur   : chr [1:7176] "Oui" "Oui" "Non" "Non" ...
## $ elec_ua   : chr [1:7176] "Non" "Non" "Oui" "Oui" ...
## $ ordure    : chr [1:7176] "Non" "Non" "Non" "Non" ...
## $ toilet    : chr [1:7176] "Oui" "Oui" "Non" "Non" ...
## $ eva_toi   : chr [1:7176] "Oui" "Oui" "Non" "Non" ...
## $ eva_eau   : chr [1:7176] "Oui" "Oui" "Non" "Non" ...
## $ tv        : chr [1:7176] "Oui" "Oui" "Non" "Non" ...
## $ fer       : chr [1:7176] "Non" "Non" "Non" "Non" ...
## $ frigo     : chr [1:7176] "Non" "Oui" "Non" "Non" ...
## $ cuisin    : chr [1:7176] "Non" "Non" "Non" "Non" ...
## $ ordin     : chr [1:7176] "Non" "Non" "Non" "Non" ...
```

```

## $ decod      : chr [1:7176] "Non" "Oui" "Non" "Non" ...
## $ car       : chr [1:7176] "Non" "Oui" "Non" "Non" ...
## $ superf    : num [1:7176] 1.25 NA NA 0.932 NA ...
## $ grosrum   : num [1:7176] 0 20 0 1 0 1 0 0 0 1 ...
## $ petitrum  : num [1:7176] 0 6 4 4 0 0 7 0 0 0 ...
## $ porc      : num [1:7176] 0 0 0 0 0 0 0 0 0 0 ...
## $ lapin     : num [1:7176] 0 0 0 0 0 0 0 0 0 0 ...
## $ volail    : num [1:7176] 0 0 15 11 0 0 20 0 0 5 ...
## $ sh_id_demo: chr [1:7176] "Non" "Non" "Non" "Non" ...
## $ sh_co_natu: chr [1:7176] "Non" "Non" "Oui" "Oui" ...
## $ sh_co_eco  : chr [1:7176] "Non" "Non" "Oui" "Oui" ...
## $ sh_id_eco  : chr [1:7176] "Non" "Oui" "Non" "Non" ...
## $ sh_co_vio  : chr [1:7176] "Oui" "Oui" "Non" "Non" ...
## $ sh_co_oth  : chr [1:7176] "Non" "Non" "Non" "Non" ...
## $ enquete    : chr [1:7176] "grappes de l'enquête transversale" "grappes de l'enquête transversale"
## - attr(*, "spec")=
## .. cols(
## ..   country = col_character(),
## ..   hhid = col_double(),
## ..   year = col_double(),
## ..   grappe = col_double(),
## ..   menage = col_double(),
## ..   vague = col_double(),
## ..   logem = col_character(),
## ..   mur = col_character(),
## ..   toit = col_character(),
## ..   sol = col_character(),
## ..   eauboi_ss = col_character(),
## ..   eauboi_sp = col_character(),
## ..   elec_ac = col_character(),
## ..   elec_ur = col_character(),
## ..   elec_ua = col_character(),
## ..   ordure = col_character(),
## ..   toilet = col_character(),
## ..   eva_toi = col_character(),
## ..   eva_eau = col_character(),
## ..   tv = col_character(),
## ..   fer = col_character(),
## ..   frigo = col_character(),
## ..   cuisin = col_character(),
## ..   ordin = col_character(),
## ..   decod = col_character(),
## ..   car = col_character(),
## ..   superf = col_double(),
## ..   grosrum = col_double(),
## ..   petitrum = col_double(),
## ..   porc = col_double(),
## ..   lapin = col_double(),
## ..   volail = col_double(),
## ..   sh_id_demo = col_character(),
## ..   sh_co_natu = col_character(),
## ..   sh_co_eco = col_character(),
## ..   sh_id_eco = col_character(),
## ..   sh_co_vio = col_character(),

```

Characteristic	N = 7,176 ^I
logem	
Autre	812 (11%)
Locataire	962 (13%)
Proprietaire sans titre	3,733 (52%)
Proprietaire titre	1,669 (23%)
toit	
Non	613 (8.5%)
Oui	6,563 (91%)
sol	
Non	1,383 (19%)
Oui	5,793 (81%)

^In (%)

```
##    ..  sh_co_oth = col_character(),
##    ..  enquete = col_character()
##    .. )
## - attr(*, "problems")=<externalptr>
```

Affichage des noms des variables

```
#Noms variables
colnames(Base_men)
```

```
## [1] "country"    "hhid"       "year"       "grappe"     "menage"
## [6] "vague"      "logem"      "mur"        "toit"       "sol"
## [11] "eauboi_ss"  "eauboi_sp"  "elec_ac"    "elec_ur"    "elec_ua"
## [16] "ordure"     "toilet"     "eva_toi"    "eva_eau"    "tv"
## [21] "fer"        "frigo"      "cuisin"     "ordin"      "decod"
## [26] "car"        "superf"     "grosum"     "petitrum"   "porc"
## [31] "lapin"      "volail"     "sh_id_demo" "sh_co_natu" "sh_co_eco"
## [36] "sh_id_eco"  "sh_co_vio"  "sh_co_oth"  "enquete"
```

##Statistiques et sorties de tableaux

Sorties de tableau avec labélisation des variables

```
#Activation de la labélisation
Base_men %>% labelled::to_factor() %>% select("logem","toit","sol") %>% tbl_summary()
```

```
#Labélisation des variables
Base_men %>% labelled::to_factor() %>% select("logem","toit","sol") %>% tbl_summary(label= list(logem ~
```

Sorties de tableaux en changeant l'entete

Characteristic	N = 7,176 ^I
Type de logement occupé	
Autre	812 (11%)
Locataire	962 (13%)
Propriétaire sans titre	3,733 (52%)
Propriétaire titre	1,669 (23%)
Type de toit	
Non	613 (8.5%)
Oui	6,563 (91%)
Type de sol	
Non	1,383 (19%)
Oui	5,793 (81%)
^I n (%)	

Description du logement	N = 7,176 ^I
Type de logement occupé	
Autre	812 (11%)
Locataire	962 (13%)
Propriétaire sans titre	3,733 (52%)
Propriétaire titre	1,669 (23%)
Type de toit	
Non	613 (8.5%)
Oui	6,563 (91%)
Type de sol	
Non	1,383 (19%)
Oui	5,793 (81%)
^I n (%)	

#Modification de l'entete

```
Base_men %>% labelled::to_factor() %>% select("logem","toit","sol") %>% tbl_summary(label= list(logem ~
```

Statistiques

```
Base_men %>% labelled::to_factor() %>% select("logem","toit","sol","superf","grosrum","petitrum") %>% t
  grosrum ~ "{mean} ({sd})",
  petitrum ~ "{mean} ({sd}")) %>% modify_header(label="Description du logement") %>%
```

Gestion des valeurs manquantes

```
Base_men %>% labelled::to_factor() %>% select("logem","toit","sol","superf","grosrum","petitrum") %>% t
  grosrum ~ "{mean} ({sd})",
  petitrum ~ "{mean} ({sd})",missing = "always",
  missing_text = "valeurs manquantes") %>% modify_header(label="Description du logement
```

Description du logement	N	N = 7,176 ^I
Type de logement occupé	7,176	
Autre		812 (11%)
Locataire		962 (13%)
Propriétaire sans titre		3,733 (52%)
Propriétaire titre		1,669 (23%)
Type de toit	7,176	
Non		613 (8.5%)
Oui		6,563 (91%)
Type de sol	7,176	
Non		1,383 (19%)
Oui		5,793 (81%)
superf	4,359	3.46 (3.84)
Unknown		2,817
grosum	7,176	1.74 (6.81)
petitrum	7,176	6 (10)

^In (%); Mean (SD)

Description du logement	N	N = 7,176 ^I
Type de logement occupé	7,176	
Autre		812 (11%)
Locataire		962 (13%)
Propriétaire sans titre		3,733 (52%)
Propriétaire titre		1,669 (23%)
valeurs manquantes		0
Type de toit	7,176	
Non		613 (8.5%)
Oui		6,563 (91%)
valeurs manquantes		0
Type de sol	7,176	
Non		1,383 (19%)
Oui		5,793 (81%)
valeurs manquantes		0
superf	4,359	3.46 (3.84)
valeurs manquantes		2,817
grosum	7,176	1.74 (6.81)
valeurs manquantes		0
petitrum	7,176	6 (10)
valeurs manquantes		0

^In (%); Mean (SD)

```
#Travail avec la base welfare
```

Chargement de la base

```
Base_wel <- readr::read_csv("../Données/ehcvm_welfare_2b_bfa2021.csv")
```

```
## Rows: 7176 Columns: 45
## -- Column specification -----
## Delimiter: ","
## chr (19): country, zae, region, milieu, hgender, hmstat, hreligion, hnation,...
## dbl (26): year, hhid, grappe, menage, vague, hhweight, hhsize, eqadu1, eqadu...
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
Base_wel
```

```
## # A tibble: 7,176 x 45
##   country year   hhid grappe menage vague zae   region milieu hhweight hhsize
##   <chr>   <dbl> <dbl> <dbl> <dbl> <dbl> <chr> <chr> <chr>   <dbl> <dbl>
## 1 bfa     2021 586005   586     5     2 Grand~ Boucl~ Rural     439     8
## 2 bfa     2021 586028   586    28     2 Grand~ Boucl~ Rural     439    33
## 3 bfa     2021 586043   586    43     2 Grand~ Boucl~ Rural     439     4
## 4 bfa     2021 586044   586    44     2 Grand~ Boucl~ Rural     439    12
## 5 bfa     2021 586052   586    52     2 Grand~ Boucl~ Rural     439    12
## 6 bfa     2021 586082   586    82     2 Grand~ Boucl~ Rural     439     6
## 7 bfa     2021 586083   586    83     2 Grand~ Boucl~ Rural     439     9
## 8 bfa     2021 586099   586    99     2 Grand~ Boucl~ Rural     439     9
## 9 bfa     2021 586109   586   109     2 Grand~ Boucl~ Rural     439     9
## 10 bfa    2021 586111   586   111     2 Grand~ Boucl~ Rural     439    15
## # i 7,166 more rows
## # i 34 more variables: eqadu1 <dbl>, eqadu2 <dbl>, hgender <chr>, hage <dbl>,
## #   hmstat <chr>, hreligion <chr>, hnation <chr>, hethnie <chr>, halfa <chr>,
## #   halfa2 <chr>, heduc <chr>, hdiploma <chr>, hhandig <chr>, hactiv7j <chr>,
## #   hactiv12m <chr>, hbranch <chr>, hsectins <chr>, hcsp <chr>, dali <dbl>,
## #   dnal <dbl>, dtot <dbl>, dtet <dbl>, pcexp <dbl>, zref <dbl>, def_spa <dbl>,
## #   def_temp <dbl>, zaemil <dbl>, dif <dbl>, p0 <dbl>, p1 <dbl>, p2 <dbl>, ...
```

Aperçu de la base

```
utils::str(Base_wel)
```

```
## spc_tbl_ [7,176 x 45] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
## $ country : chr [1:7176] "bfa" "bfa" "bfa" "bfa" ...
## $ year : num [1:7176] 2021 2021 2021 2021 2021 ...
## $ hhid : num [1:7176] 586005 586028 586043 586044 586052 ...
## $ grappe : num [1:7176] 586 586 586 586 586 586 586 586 586 586 ...
## $ menage : num [1:7176] 5 28 43 44 52 82 83 99 109 111 ...
## $ vague : num [1:7176] 2 2 2 2 2 2 2 2 2 2 ...
## $ zae : chr [1:7176] "Grand-ouest" "Grand-ouest" "Grand-ouest" "Grand-ouest" ...
## $ region : chr [1:7176] "Boucle du Mouhoum" "Boucle du Mouhoum" "Boucle du Mouhoum" "Boucle du Mouhoum" ...
## $ milieu : chr [1:7176] "Rural" "Rural" "Rural" "Rural" ...
```

```

## $ hhweight : num [1:7176] 439 439 439 439 439 439 439 439 439 439 ...
## $ hhsiz  : num [1:7176] 8 33 4 12 12 6 9 9 9 15 ...
## $ eqadu1 : num [1:7176] 6.72 23.76 2.83 8.52 9.21 ...
## $ eqadu2 : num [1:7176] 4.18 14.02 2.44 5.98 6.28 ...
## $ hgender : chr [1:7176] "Masculin" "Masculin" "Masculin" "Masculin" ...
## $ hage    : num [1:7176] 45 45 31 54 40 57 46 37 66 41 ...
## $ hmstat  : chr [1:7176] "Marié(e) monogame" "Marié(e) polygame" "Marié(e) monogame" "Marié(e) monogame" ...
## $ hreligion: chr [1:7176] "Musulman" "Musulman" "Chrétien" "Musulman" ...
## $ hnation : chr [1:7176] "Burkina Faso" "Burkina Faso" "Burkina Faso" "Burkina Faso" ...
## $ hethnie : chr [1:7176] "Mossi" "Mossi" "Bobo" "Autres ethnies" ...
## $ halfa   : chr [1:7176] "Non" "Non" "Non" "Oui" ...
## $ halfa2  : chr [1:7176] "Non" "Non" "Non" "Oui" ...
## $ heduc   : chr [1:7176] "Aucun" "Aucun" "Aucun" "Aucun" ...
## $ hdiploma : chr [1:7176] "Aucun" "Aucun" "Aucun" "Aucun" ...
## $ hhandig : chr [1:7176] "Non" "Non" "Non" "Non" ...
## $ hactiv7j : chr [1:7176] "Occupe" "Occupe" "Occupe" "Chomeur" ...
## $ hactiv12m : chr [1:7176] "Occupe" "Occupe" "Occupe" "Occupe" ...
## $ hbranch  : chr [1:7176] "Services perso." "Services perso." "Commerce" "Agriculture" ...
## $ hsectins : chr [1:7176] "Entreprise Privée" "Entreprise Privée" "Entreprise Privée" "Entreprise Privée" ...
## $ hcsp     : chr [1:7176] "Travailleur pour compte propre" "Patron" "Travailleur pour compte propre" ...
## $ dali     : num [1:7176] 1224618 5900909 524436 564769 502364 ...
## $ dnal     : num [1:7176] 1275733 5590681 296808 435736 461143 ...
## $ dtot     : num [1:7176] 2500352 11491589 821244 1000505 963507 ...
## $ dtet     : num [1:7176] 312544 348230 205311 83375 80292 ...
## $ pcexp    : num [1:7176] 330982 368773 217423 88294 85029 ...
## $ zref     : num [1:7176] 247806 247806 247806 247806 247806 ...
## $ def_spa  : num [1:7176] 0.944 0.944 0.944 0.944 0.944 ...
## $ def_temp : num [1:7176] 1.06 1.06 1.06 1.06 1.06 ...
## $ zaemil   : num [1:7176] 2 2 2 2 2 2 2 2 2 2 ...
## $ dif      : num [1:7176] -83176 -120967 30383 159512 162777 ...
## $ p0       : num [1:7176] 0 0 100 100 100 100 100 0 100 100 ...
## $ p1       : num [1:7176] 0 0 12.3 64.4 65.7 ...
## $ p2       : num [1:7176] 0 0 1.5 41.4 43.1 ...
## $ pauv     : num [1:7176] 0 0 1 1 1 1 1 0 1 1 ...
## $ quintiles : num [1:7176] 4 4 2 1 1 2 1 5 1 1 ...
## $ deciles   : num [1:7176] 7 7 4 1 1 3 1 10 1 1 ...
## - attr(*, "spec")=
## .. cols(
## ..   country = col_character(),
## ..   year = col_double(),
## ..   hhid = col_double(),
## ..   grappe = col_double(),
## ..   menage = col_double(),
## ..   vague = col_double(),
## ..   zae = col_character(),
## ..   region = col_character(),
## ..   milieu = col_character(),
## ..   hhweight = col_double(),
## ..   hhsiz  = col_double(),
## ..   eqadu1 = col_double(),
## ..   eqadu2 = col_double(),
## ..   hgender = col_character(),
## ..   hage = col_double(),
## ..   hmstat = col_character(),

```



```
## .. hreligion = col_character(),
## .. hnation = col_character(),
## .. hethnie = col_character(),
## .. halfa = col_character(),
## .. halfa2 = col_character(),
## .. heduc = col_character(),
## .. hdiploma = col_character(),
## .. hhandig = col_character(),
## .. hactiv7j = col_character(),
## .. hactiv12m = col_character(),
## .. hbranch = col_character(),
## .. hsectins = col_character(),
## .. hcsp = col_character(),
## .. dali = col_double(),
## .. dnal = col_double(),
## .. dtot = col_double(),
## .. dtet = col_double(),
## .. pcexp = col_double(),
## .. zref = col_double(),
## .. def_spa = col_double(),
## .. def_temp = col_double(),
## .. zaemil = col_double(),
## .. dif = col_double(),
## .. p0 = col_double(),
## .. p1 = col_double(),
## .. p2 = col_double(),
## .. pauv = col_double(),
## .. quintiles = col_double(),
## .. deciles = col_double()
## .. )
## - attr(*, "problems")=<externalptr>
```

Affichage des noms des variables

```
colnames(Base_wel)
```

```
## [1] "country" "year" "hhid" "grappe" "menage" "vague"
## [7] "zae" "region" "milieu" "hhweight" "hysize" "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"
```

Sorties de tableaux

```
Base_wel %>% labelled::to_factor() %>% select("hgender", "hage", "hmstat", "heduc", "hdiploma") %>% tbl_summary(
  digits = list(hage ~ 2),
  #digits = everything()~c(0,0,0)
  missing = "always",
  missing_text = "valeurs manquantes") %>% modify_header(label="Caractéristiques du Ch")
```

Caractéristiques du Chef de ménage	N	N = 7,176 ^I
Sexe du Chef de ménage	7,176	
Féminin		1,075 (15%)
Masculin		6,101 (85%)
valeurs manquantes		0
Age du Chef de ménage	7,176	45.68 (14.67)
valeurs manquantes		0
Situation matrimoniale	7,176	
Célibataire		349 (4.9%)
Divorcé(e)		50 (0.7%)
Marié(e) monogame		4,352 (61%)
Marié(e) polygame		1,572 (22%)
Séparé(e)		86 (1.2%)
Union libre		199 (2.8%)
Veuf(ve)		568 (7.9%)
valeurs manquantes		0
Niveau d'éducation	7,176	
Aucun		4,706 (66%)
Postprimaire technique		8 (0.1%)
Postprimaire général		622 (8.7%)
Postsecondaire		44 (0.6%)
Primaire		1,051 (15%)
Secondaire général		395 (5.5%)
Secondaire technique		39 (0.5%)
Supérieur		311 (4.3%)
valeurs manquantes		0
diplome le plus élevé	7,176	
Aucun		5,551 (77%)
BAC		229 (3.2%)
BEPC		394 (5.5%)
BT		5 (<0.1%)
CAP		20 (0.3%)
CEPE		726 (10%)
DEUG, DUT, BTS		58 (0.8%)
Doctorat/Phd		4 (<0.1%)
Licence		110 (1.5%)
Maitrise		43 (0.6%)
Master/DEA/DESS		36 (0.5%)
valeurs manquantes		0

^In (%); Mean (SD)