Untitled

Amadou SAWADOGO & Abdou NIANG

2022-11-06

Commencons par charger le tableau de données brutes

```
setwd("C:/Users/abdou/Documents/cours seminaire integration/ODD_CSV_V2/")
getwd()
```

```
## [1] "C:/Users/abdou/Documents/cours seminaire integration/ODD_CSV_V2"
```

```
data=read.table('ODD_DEP.csv',sep=";",header=TRUE)
```

Début du nettoyage

Nous avons créé un tableau de données appelé dataINSEE qui regroupe les 101 départements ainsi que l'ensemble des codes des variables

Nous avons ensuite renommé les lignes en nom de departements sans repetition

Cat = categorie

###les deux dernieres lignes de codes servant a selectionner les sous-champs de notre tableau de données et enfin l'affichage pour le departement de la Meuse a titre d'illustration.

```
dataINSEE = data.frame(matrix(nrow=101,ncol=849))
row.names(dataINSEE) = unique(data$libgeo)
cat<-data$sous_champ
cat[data$libgeo=="Meuse"]
```

```
##   [1] "total"          "moins30"        "30_39"
##   [4] "40_49"          "50_59"          "60_74"
##   [7] "75plus"         ""               ""
##  [10] ""               ""               ""
##  [13] ""               ""               ""
##  [16] ""               ""               ""
##  [19] ""               ""               ""
##  [22] ""               "aperm"          "arables"
##  [25] "arom"           "autre"          "betind"
##  [28] "bois"           "cerea"          "choux"
##  [31] "fibr"           "fleur"          "fouran"
##  [34] "fruit"          "inddiv"         "jach"
##  [37] "jfame"          "jfamh"          "lande"
##  [40] "lfrais"         "lsec"           "olea"
##  [43] "pepin"          "perm"           "pote"
##  [46] "prairartif"     "prot"           "saue"
##  [49] "saut"           "semen"          "solartif"
##  [52] "sthe"           "sthh"           "stot"
##  [55] "sucdom"         "tdom"           "vigne"
##  [58] ""               ""               ""
##  [61] ""               ""               "homme"
##  [64] "femme"          "homme"          "femme"
##  [67] ""               ""               ""
##  [70] ""               "homme"          "femme"
##  [73] ""               ""               ""
##  [76] ""               "homme"          "femme"
##  [79] ""               ""               ""
##  [82] ""               ""               ""
##  [85] ""               ""               ""
##  [88] "age_1"          "age_2"          "age_3"
##  [91] "cs_3"           "cs_4"           "cs_5"
##  [94] "cs_6"           "ens"            "age_1"
##  [97] "age_2"          "age_3"          "cs_3"
## [100] "cs_4"           "cs_5"           "cs_6"
```

```
## [103] "ens"          "age_1"        "age_2"
## [106] "age_3"        "cs_3"         "cs_4"
## [109] "cs_5"         "cs_6"         "ens"
## [112] ""             ""             ""
## [115] "f_15_24"      "f_25_54"      "f_55_64"
## [118] "f"            "h_15_24"      "h_25_54"
## [121] "h_55_64"      "h"            "f_15_24"
## [124] "f_25_54"      "f_55_64"      "f"
## [127] "h_15_24"      "h_25_54"      "h_55_64"
## [130] "h"            "f_15_24"      "f_25_54"
## [133] "f_55_64"      "h_15_24"      "h_25_54"
## [136] "h_55_64"      "f"            "h"
## [139] "total"        "15_24"        "25_54"
## [142] "55_64"        ""             ""
## [145] ""             ""             ""
## [148] "supe_aep"     "supe_ind"     "supe_irr"
## [151] "supe_elec"    "supe_can"     "supe_hydro"
## [154] "sout_aep"     "sout_ind"     "sout_irr"
## [157] "sout_elec"    "sout_can"     "saum_ind"
## [160] "saum_elec"    "saum_can"     "saum_aep"
## [163] "saum_irr"     "saum_hydro"   "total"
## [166] ""             ""             ""
## [169] ""             ""             ""
## [172] ""             ""             ""
## [175] ""             ""             ""
## [178] ""             ""             ""
## [181] ""             ""             ""
## [184] ""             ""             ""
## [187] ""             ""             ""
## [190] ""             ""             ""
## [193] ""             ""             "noncl"
## [196] "excel"        "bonne"        "suffi"
## [199] "insuf"        "total"        "noncl"
## [202] "excel"        "bonne"        "suffi"
## [205] "insuf"        "noncl"        "excel"
## [208] "bonne"        "suffi"        "insuf"
## [211] "total"        "noncl"        "excel"
## [214] "bonne"        "suffi"        "insuf"
## [217] ""             ""             ""
## [220] ""             ""             ""
## [223] ""             ""             "total"
## [226] "homme"        "femme"        "15_24"
## [229] "25_49"        "plus50"       "f_15_24"
## [232] "f_25_54"      "f_55_64"      "h_15_24"
## [235] "h_25_54"      "h_55_64"      "15_24"
## [238] "25_54"        "55_64"        "h"
## [241] "f"            "total"        "f_15_24"
## [244] "f_25_54"      "f_55_64"      "h_15_24"
## [247] "h_25_54"      "h_55_64"      "15_24"
## [250] "25_54"        "55_64"        "h"
## [253] "f"            "total"        "f_15_24"
## [256] "f_25_54"      "f_55_64"      "h_15_24"
## [259] "h_25_54"      "h_55_64"      "f"
## [262] "h"            "15_24"        "25_54"
## [265] "55_64"        "total"        "f_15_24"
## [268] "f_25_54"      "f_55_64"      "h_15_24"
## [271] "h_25_54"      "h_55_64"      "15_24"
## [274] "25_54"        "55_64"        "h"
## [277] "f"            "total"        "f_15_24"
## [280] "f_25_54"      "f_55_64"      "h_15_24"
## [283] "h_25_54"      "h_55_64"      "15_24"
## [286] "25_54"        "55_64"        "h"
## [289] "f"            "f_15_24"      "f_25_54"
## [292] "f_55_64"      "h_15_24"      "h_25_54"
## [295] "h_55_64"      "total"        "f"
## [298] "h"            "15_24"        "25_54"
## [301] "55_64"        "total"        ""
## [304] "3600z"        "3700z"        "3811z"
## [307] "3812z"        "3821z"        "3822z"
## [310] "3831z"        "3832z"        "3900z"
## [313] "4329a"        "9104z"        "total"
## [316] ""             "3600z"        "3700z"
## [319] "3811z"        "3812z"        "3821z"
## [322] "3822z"        "3831z"        "3832z"
## [325] "3900z"        "4329a"        "9104z"
## [328] "total"        ""             ""
## [331] ""             ""             ""
## [334] ""             ""             ""
```

```
## [337] ""                ""                ""
## [340] ""                ""                ""
## [343] ""                "coop"            "mutu"
## [346] "asso"            "fond"            "coop"
## [349] "mutu"            "asso"            "fond"
## [352] ""                "coop"            "mutu"
## [355] "asso"            "fond"            ""
## [358] ""                ""                "as"
## [361] "astt"            "baptvcp"         "btp"
## [364] "ccp"             "cna"             "m"
## [367] "sca"             "tegelc"          "total"
## [370] "as"              "astt"            "baptvcp"
## [373] "btp"             "ccp"             "cna"
## [376] "m"               "sca"             "tegelc"
## [379] "total"           "acc_trajet_ens"  "acc_trav_ens"
## [382] "maladie_prof_ens" "act"            "nav"
## [385] "act"             "nav"             "act"
## [388] "nav"             "act"             "nav"
## [391] "tec"             "voiture"         "deux_roues"
## [394] "velo"            "deux_roues_mot"  "pied"
## [397] "sans_transp"     "tec"             "voiture"
## [400] "deux_roues"      "deux_roues_mot"  "velo"
## [403] "pied"            "sans_transp"     ""
## [406] ""                ""                "Ami"
## [409] "Bar"             "Eng"             "Ind"
## [412] "Min"             "Nuc"             "Tmd"
## [415] "sb"              "sh"              "total"
## [418] ""                "comappr"         "comapprev"
## [421] "compres"         ""                "bio"
## [424] "nonbio"          "total"           ""
## [427] ""                ""                ""
## [430] ""                ""                ""
## [433] ""                ""                ""
## [436] ""                ""                ""
## [439] ""                ""                ""
## [442] ""                ""                ""
## [445] ""                ""                ""
## [448] ""                ""                ""
## [451] ""                ""                ""
## [454] ""                ""                ""
## [457] ""                ""                ""
## [460] ""                ""                "cet2"
## [463] "compost"         "inciavec"        "incisans"
## [466] "metha"           "tri"             "total"
## [469] ""                ""                ""
## [472] ""                ""                ""
## [475] ""                ""                ""
## [478] ""                ""                ""
## [481] ""                ""                ""
## [484] ""                ""                ""
## [487] ""                ""                ""
## [490] ""                "occas"           "vac"
## [493] "res_princ"       "res_sec"         "occas"
## [496] "vac"             "res_princ"       "res_sec"
## [499] "equip_prox"      "equip_inter"     "equip_sup"
## [502] ""                "equip_prox"      "equip_inter"
## [505] "equip_sup"       ""                ""
## [508] ""                ""                ""
## [511] ""                ""                "org"
## [514] "rec"             "cet"             "inca"
## [517] "incs"            "nd"              "total"
## [520] "Air_HCN"         "Air_NH3"         "Air_As"
## [523] "Eau_As"          "Eau_N"           "Air_Cd"
## [526] "Eau_Cd"          "Air_HCl"         "Air_CVM"
## [529] "Eau_CVM"         "Air_Cr"          "Eau_Cr"
## [532] "Eau_Cr6"         "Air_COVNM"       "Eau_AOX"
## [535] "Air_Cu"          "Eau_Cu"          "Eau_CN"
## [538] "Eau_DBO5"        "Eau_DCO"         "Air_DCM"
## [541] "Eau_DCM"         "Air_CO2"         "Eau_Sn"
## [544] "Eau_Fe"          "Air_HF"          "Eau_F"
## [547] "Eau_HC"          "Air_Mn"          "Eau_Mn"
## [550] "Eau_MES"         "Air_Hg"          "Eau_Hg"
## [553] "Air_CO"          "Air_Ni"          "Eau_Ni"
## [556] "Air_NOx"         "Air_SOx"         "Air_C"
## [559] "Eau_C"           "Eau_P"           "Air_Pb"
## [562] "Eau_Pb"          "Air_N2O"         "Air_TSP"
## [565] "Eau_sulfates"    "Air_H2S"         "Eau_Ti"
## [568] "Air_Zn"          "Eau_Zn"          "Air_HCN"
```

```
## [571] "Air_NH3"       "Air_As"        "Eau_As"
## [574] "Eau_N"         "Air_Cd"        "Eau_Cd"
## [577] "Air_CVM"       "Eau_CVM"       "Air_Cr"
## [580] "Eau_Cr"        "Eau_Cr6"       "Air_COVNM"
## [583] "Eau_AOX"       "Air_Cu"        "Eau_Cu"
## [586] "Eau_CN"        "Eau_DBO5"      "Eau_DCO"
## [589] "Air_DCM"       "Eau_DCM"       "Air_CO2"
## [592] "Eau_Sn"        "Eau_Fe"        "Air_HF"
## [595] "Eau_F"         "Eau_HC"        "Air_Mn"
## [598] "Eau_Mn"        "Eau_MES"       "Air_Hg"
## [601] "Eau_Hg"        "Air_CO"        "Air_Ni"
## [604] "Eau_Ni"        "Air_NOx"       "Air_SOx"
## [607] "Air_C"         "Eau_C"         "Eau_P"
## [610] "Air_Pb"        "Eau_Pb"        "Air_N2O"
## [613] "Air_TSP"       "Eau_sulfates" "Air_H2S"
## [616] "Eau_Ti"        "Air_Zn"        "Eau_Zn"
## [619] "atm"           "ava"           "temp82"
## [622] "temp99"        "inoi"          "inom"
## [625] "mvth"          "mvts"          "sei"
## [628] "volc"          "trop"          "atm"
## [631] "ava"           "temp82"        "temp99"
## [634] "inoi"          "inom"          "mvth"
## [637] "mvts"          "sei"           "volc"
## [640] "trop"          "atm"           "ava"
## [643] "feu"           "ino"           "mvt"
## [646] "sei"           "volc"          "atm"
## [649] "ava"           "feu"           "ino"
## [652] "mvt"           "sei"           "volc"
## [655] "atm"           "ava"           "feu"
## [658] "ino"           "mvt"           "sei"
## [661] "volc"          ""              "ch4"
## [664] "co2nonbio"     "hfc"           "n2o"
## [667] "pfc"           "sf6"           ""
## [670] ""              ""              ""
## [673] ""              ""              "ch4"
## [676] "co2bio"        "co2nonbio"     "hfc"
## [679] "n2o"           "pfc"           "sf6"
## [682] "lt"            "lr"            "lt"
## [685] "lr"            "lt"            "lr"
## [688] ""              ""              "Atm"
## [691] "Ava"           "Feu"           "Ino"
## [694] "Mvt"           "Rad"           "Sei"
## [697] "Vol"           ""              ""
## [700] ""              ""              ""
## [703] ""              ""              "classe_A"
## [706] "classe_B"      "classe_C"      "tot"
## [709] "tresfaible"    "faible"        "modere"
## [712] "moyen"         "fort"          ""
## [715] ""              ""              ""
## [718] ""              ""              ""
## [721] "FOR_PRO"       "TOU_PRO"       "ZNIEFF1"
## [724] "ZNIEFF2"       "ZZZ"           "SIC"
## [727] "ZPS"           "NATURA"        "PNC"
## [730] "RNR"           "APB"           "RNCFS"
## [733] "RBFD"          "RN"            "PNP"
## [736] "PNR"           "RAMSAR"        "BIO"
## [739] "CELRL"         "bonne"         "excel"
## [742] "mauv"          "medio"         "tot"
## [745] "trmau"         ""              "1"
## [748] "11"            "12"            "13"
## [751] "14"            "2"             "21"
## [754] "22"            "23"            "24"
## [757] "3"             "31"            "32"
## [760] "33"            "4"             "41"
## [763] "42"            "5"             "51"
## [766] "52"            "9"             "99"
## [769] ""              "1"             "11"
## [772] "12"            "13"            "14"
## [775] "2"             "21"            "22"
## [778] "23"            "24"            "3"
## [781] "31"            "32"            "33"
## [784] "4"             "41"            "42"
## [787] "5"             "51"            "52"
## [790] "9"             "99"            "1_X"
## [793] "1_2"           "1_3"           "1_4"
## [796] "1_5"           "X_1"           "X_2"
## [799] "X_3"           "X_4"           "X_5"
## [802] ""              ""              ""
```

```
## [802]
## [805] ""                ""                ""
## [808] ""                ""                ""
## [811] ""                "asso"            "ent"
## [814] "epci"            "epub"            ""
## [817] "agri"            "artif"           "nat"
## [820] "critair1_gaz"        "critair1_essence"     "critair1_esshybride"
## [823] "critair1_gazolhybride" "critair2_gazol"      "critair2_essence"
## [826] "critair3_gazol"      "critair3_essence"     "critair4_gazol"
## [829] "critair5_gazol"      "critairE_elechydro"   "critairnd_nd"
## [832] "critairnd_gazol"     "critairnd_essence"    "critairnc_gazol"
## [835] "critairnc_essence"   ""                "pdc_nbtot"
## [838] "irve_nb"          "gratuit"          "gratuit_client"
## [841] "payant"          "gratuit_1214_1921"  "H_24_7"
## [844] "H_8_5"           "H_12_6"           "H_var"
## [847] "H_na"            ""                ""
```

```r
colnames(dataINSEE) = cat[data$libgeo=="Meuse"]

nbr_rep_row = rep(0,time=101)

for(k in 0:101){
  nbr_rep_row[k] = as.character(unique(data$libgeo)[k])
}

nbr_rep_col=rep(0,time=849)

for(i in 0:101){
  nbr_rep_col = data$libgeo==nbr_rep_row[i]
  for(j in 0:848){
   if(colnames(dataINSEE)[j+1]==""){
     colnames(dataINSEE)[j+1] = data$variable[data$libgeo=="Meuse"][j+1]
   } else{
     colnames(dataINSEE)[j+1] = paste(data$variable[data$libgeo=="Meuse"][j+1]
                      ,data$sous_champ[data$libgeo=="Meuse"][j+1],sep=" : ")
   }

   dataINSEE[i,] = data$A2018[nbr_rep_col]
  }

}
```

Choix des variables a partir de

```r
library(tidyverse)
```

```
## Warning: le package 'tidyverse' a été compilé avec la version R 4.1.3
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.3.6     v purrr   0.3.5
## v tibble  3.1.8     v dplyr   1.0.10
## v tidyr   1.2.1     v stringr 1.4.1
## v readr   2.1.3     v forcats 0.5.2
```

```
## Warning: le package 'ggplot2' a été compilé avec la version R 4.1.3
```

```
## Warning: le package 'tibble' a été compilé avec la version R 4.1.3
```

```
## Warning: le package 'tidyr' a été compilé avec la version R 4.1.3
```

```
## Warning: le package 'readr' a été compilé avec la version R 4.1.3
```

```
## Warning: le package 'purrr' a été compilé avec la version R 4.1.3
```

```
## Warning: le package 'dplyr' a été compilé avec la version R 4.1.3
```

```
## Warning: le package 'stringr' a été compilé avec la version R 4.1.3
```

```
## -- Conflicts ------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
Tab_fin = select(dataINSEE, `taux_pvt : total`,`taux_pvt : moins30`,`taux_pvt : 30_39`,
`taux_pvt : 40_49`,`taux_pvt : 50_59`,`taux_pvt : 60_74`,
`taux_pvt : 75plus`,`part_actifs_stables : `,`part_depl_dom_trav : tec`, `esper_vie : homme`,
`parc_vp : critair1_essence`,`moy_dist_km : act`,`moy_dist_min_hp : act`,`ges : ch4`,
`nb_pop_hlm : `,`part_foy_fisc_impos : `,`apl_medgen_moins65 : `,`revenu_decl_median : `,
`esper_vie : femme` )
```

Renommer les variables

```
names(Tab_fin)[names(Tab_fin) == "taux_pvt : total"] =  "Taux de pauvrete total"
names(Tab_fin)[names(Tab_fin) == "taux_pvt : moins30"] = "Taux de pauvrete > 30 "
names(Tab_fin)[names(Tab_fin) == "taux_pvt : 30_39"] = "Taux de pauvrete 30 - 39"
names(Tab_fin)[names(Tab_fin) == "taux_pvt : 40_49"] = "Taux de pauvrete 40 - 49"
names(Tab_fin)[names(Tab_fin) == "taux_pvt : 50_59"] = "Taux de pauvrete 50 -59"
names(Tab_fin)[names(Tab_fin) == "taux_pvt : 60_74"] = "Taux de pauvrete 60 -74"
names(Tab_fin)[names(Tab_fin) == "taux_pvt : 75plus"] = "Taux de pauvrete => 75"
names(Tab_fin)[names(Tab_fin) == "part_actifs_stables :"] = "Part des actifs stables :"
names(Tab_fin)[names(Tab_fin) == "part_depl_dom_trav : tec"] =  "PDDT tec"
names(Tab_fin)[names(Tab_fin) == "esper_vie : homme"] = "esperance de vie homme"
names(Tab_fin)[names(Tab_fin) == "parc_vp : critair1_essence"] = " part des vehicules criticair 1 essence"
names(Tab_fin)[names(Tab_fin) == "moy_dist_km : act"] = "Moy_dist_km_actif"
names(Tab_fin)[names(Tab_fin) == "moy_dist_min_hp : act"] = "Moy_dist_min_actif"
names(Tab_fin)[names(Tab_fin) == "ges : ch4"] = " Emission en equivalent CO2 pour le methane"
names(Tab_fin)[names(Tab_fin) == "nb_pop_hlm : "] = "nombre de personnes en logement social"

names(Tab_fin)[names(Tab_fin) == "part_foy_fisc_impos : "] = "part des foyers fiscaux imposés"
names(Tab_fin)[names(Tab_fin) == "apl_medgen_moins65 : "] = "accessiblité potentielle au medecins generalistes"
names(Tab_fin)[names(Tab_fin) == "revenu-decl_median : "] = "revenu fiscal median declaré par unité"
names(Tab_fin)[names(Tab_fin) == "esper_vie : femme"] = "esperance de vie des femme"
colnames(Tab_fin)
```

```
##  [1] "Taux de pauvrete total"
##  [2] "Taux de pauvrete > 30 "
##  [3] "Taux de pauvrete 30 - 39"
##  [4] "Taux de pauvrete 40 - 49"
##  [5] "Taux de pauvrete 50 -59"
##  [6] "Taux de pauvrete 60 -74"
##  [7] "Taux de pauvrete => 75"
##  [8] "part_actifs_stables : "
##  [9] "PDDT tec"
## [10] "esperance de vie homme"
## [11] " part des vehicules criticair 1 essence"
## [12] "Moy_dist_km_actif"
## [13] "Moy_dist_min_actif"
## [14] " Emission en equivalent CO2 pour le methane"
## [15] "nombre de personnes en logement social"
## [16] "part des foyers fiscaux imposés"
## [17] "accessiblité potentielle au medecins generalistes"
## [18] "revenu fiscal median declaré par unité"
## [19] "esperance de vie des femme"
```

Descriptif des variables

Taux de pauvreté total : Le taux de pauvreté est la proportion d'individus appartenant à des ménages dont le niveau de vie, c'est-à-dire le revenu disponible (après transferts, impôts et prestations sociales) par unité de consommation (UC) est inférieur au seuil de 60 % de la médiane du revenu disponible par UC de l'ensemble de la population. En 2017, ce seuil est de 1 041 euros par mois en France

Taux de pauvreté > 30 : Le taux de pauvreté pour la tranche d'âge de mois de 30 ans

Taux de pauvreté 30 - 39 : Le taux de pauvreté pour la tranche d'âge entre 30 ans et 39 ans

Taux de pauvreté 40 - 49 : Le taux de pauvreté pour la tranche d'âge entre 40 ans et 49 ans

Taux de pauvreté 50 -59 : Le taux de pauvreté pour la tranche d'âge entre 50 ans et 59 ans

Taux de pauvreté 60 -74 : Le taux de pauvreté pour la tranche d'âge entre 60 ans 74 ans

Taux de pauvreté =>75 : Le taux de pauvreté pour la tranche d'âge de 75 ans et plus

Part des actifs stables : Part des actifs stables parmi les actifs ayant un emploi

PDDT_Transport en commun : parts modales des déplacements domicile-travail en transports en commun en %

nb_pop_hlm: Nombres de personnes en logement social

revenu_decl_median : revenu fiscal median declaré par unité de consommationn : Egalement appelé "niveau de vie", c'est le revenu disponible par "équivalent adulte". Il est calculé en rapportant le revenu disponible du ménage au nombre d'unités de consommation qui le composent.

parc_vp : critair1_essence: part des vehicules critair1 essence(Voitures essence à partir de janvier 2011) dans l'ensemble du parc automobile

ges : ch4 Emission en equivalent CO2 pour le methane

Moy_dist_km_actif : Distance moyenne des navettes domicile-Travail pour les actifs occupés, (en km)

Moy_dist_min_actif : Durée moyenne des navettes domicile-Travail pour les actifs occupés (en minute)

apl_medgen_moins65 : mesure l'adéquation spatiale entre l'offre et la demande de soins de premier recours dans une localité

Esperance de vie des hommes : Espérance de vie à la naissance pour les Hommes

Esperance de vie des femmes : Espérance de vie à la naissance pour les Hommes

part_foy_fisc_impos nombre de foyers soumis a l'impot

```
write.csv(Tab_fin,"C:/Users/abdou/Documents/cours seminaire integration/tableauFinal.csv", row.names = FALSE)
```