Projet final du logiciel statistique R

Sous la supervision de M. HEMA ABOUBACAR

Jean Pierre Adiouma NDIAYE

2023-07-24

Le présent document regroupe les deux premières parties du projet. La dernière partie concernant R chiny est jointe à travers un autre fichier nommé "app.R" et une base de données nommée "ACLED-Western\_Africa" qui permet d'ouvrir l'application. Egalement, nous avons proposé un code à la fin du document RmarkDown qui permet de faciliter l'accès.

# Partie I

## Préparation des données

### Description

#### Importation et mise en forme

Nous allons importer la base de données dans un objet de type data.frame nommé projet.

library(readxl)  
projet <- read\_excel("~/ISEP3/R/Base\_Partie 1.xlsx")  
# On vérifie bien qu'il s'agit d'un dataFrame  
class(projet)

## [1] "tbl\_df" "tbl" "data.frame"

#### Sélection des variables mentionnées dans la section description.

# Renommer la première variable en "key"  
# names(projet)[0] <- "key"  
projet<- projet[c("key","q1","q2","q23","q24","q24a\_1","q24a\_2","q24a\_3","q24a\_4","q24a\_5","q24a\_6","q24a\_7","q24a\_9","q24a\_10","q25","q26","q12","q14b","q16","q17","q19","q20","filiere\_1","filiere\_2","filiere\_3","filiere\_4","q8","q81","gps\_menlatitude","gps\_menlongitude","submissiondate","start","today")]  
variables <- names(projet)

#### Faisons un tableau qui résume les valeurs manquantes par variable

library(tidyverse)  
library(flextable)  
library(knitr)  
na\_var\_eff <- colSums(is.na(projet))  
na\_var\_pct <- colSums(is.na(projet)\* 100/nrow(projet))  
## Créer un tableau récapitulatif des valeurs manquantes  
na\_var <- base::data.frame(Variables = variables, Effectifs = na\_var\_eff,  
 Pourcentage = na\_var\_pct)  
## Afficher les résultats  
flextable(na\_var) %>% bold(part = "header")

| **Variables** | **Effectifs** | **Pourcentage** |
| --- | --- | --- |
| key | 0 | 0.0 |
| q1 | 0 | 0.0 |
| q2 | 0 | 0.0 |
| q23 | 0 | 0.0 |
| q24 | 0 | 0.0 |
| q24a\_1 | 0 | 0.0 |
| q24a\_2 | 0 | 0.0 |
| q24a\_3 | 0 | 0.0 |
| q24a\_4 | 0 | 0.0 |
| q24a\_5 | 0 | 0.0 |
| q24a\_6 | 0 | 0.0 |
| q24a\_7 | 0 | 0.0 |
| q24a\_9 | 0 | 0.0 |
| q24a\_10 | 0 | 0.0 |
| q25 | 0 | 0.0 |
| q26 | 0 | 0.0 |
| q12 | 0 | 0.0 |
| q14b | 1 | 0.4 |
| q16 | 1 | 0.4 |
| q17 | 131 | 52.4 |
| q19 | 120 | 48.0 |
| q20 | 0 | 0.0 |
| filiere\_1 | 0 | 0.0 |
| filiere\_2 | 0 | 0.0 |
| filiere\_3 | 0 | 0.0 |
| filiere\_4 | 0 | 0.0 |
| q8 | 0 | 0.0 |
| q81 | 0 | 0.0 |
| gps\_menlatitude | 0 | 0.0 |
| gps\_menlongitude | 0 | 0.0 |
| submissiondate | 0 | 0.0 |
| start | 0 | 0.0 |
| today | 0 | 0.0 |

#### Vérifier s’il y a des valeurs manquantes pour la variable key dans la base projet. Si oui, identifier la (ou les) PME concernée(s).

# Vérifier les valeurs manquantes pour la variable "key"  
valeurs\_manquantes\_key <- is.na(projet$key)  
View(valeurs\_manquantes\_key)  
# Identifier les PME concernées (celles avec des valeurs manquantes pour la variable "key")  
pme\_concernees <- subset(projet, valeurs\_manquantes\_key)$key  
# Afficher les PME concernées  
print(pme\_concernees)

## character(0)

### Création de variables

Dans cette partie, nous répondrons simultanément à certaines questions.

#### Renommons la variable q1 en region q2 en departement et q23 en sexe

library(dplyr)  
names(projet)[names(projet) == "q1"] <- "region"  
names(projet)[names(projet) == "q2"] <- "departement"  
names(projet)[names(projet) == "q23"] <- "sexe"

#### Créons la variable sexe\_2 qui vaut 1 si sexe égale à Femme et 0 sinon.

projet$sexe\_2 <- ifelse(projet$sexe == "Femme" , 1, 0)

#### Créons un data.frame nommé langues qui prend les variables key et les variables correspondantes décrites plus haut.

langues <- projet %>% select("key", starts\_with("q24a\_"))

#### Créons une variable parle qui est égale au nombre de langue parlée par le dirigeant de la PME.

langues$parle <- rowSums(projet[,names(langues)[-1]],dims = 1)

#### Sélectionnez uniquement les variables key et parle, l’objet de retour sera langues.

# avec la fonction select je règle mon problème  
langues<- dplyr::select(langues,key, parle)  
langues <- as.data.frame(langues)

#### Merger les data.frame projet et langues:

# Fusionner les data.frames "projet" et "langues"  
projet\_merge <- merge(projet, langues, by = "key")

## 2 Analyses descriptives (10 pts)

# appel au package gtsummary  
#install.packages("gtsummary")  
library(gtsummary)  
library(dplyr)  
#installed.packages(flextable)  
library(flextable)

### creation de tableau

#### repartition des PME suivant le sexe du responsable

# la fonction tbl\_summary nous vient en aide en selectionnant la variable q23   
t\_sexe <- projet\_merge[c("sexe")] %>%  
 tbl\_summary(   
 label = (sexe ~ " repartition des PME selon le sexe de leur responsable")) %>%  
 bold\_labels() %>%  
 italicize\_labels()  
t\_sexe

| **Characteristic** | **N = 250**1 |
| --- | --- |
| ***repartition des PME selon le sexe de leur responsable*** |  |
| Femme | 191 (76%) |
| Homme | 59 (24%) |
| 1n (%) | |

#### repartition selon le niveau d’instruction

# de même que precedamment , mais cette fois avec   
t\_instruction <- projet\_merge[c("q25")] %>%  
 tbl\_summary(   
 label = (q25 ~ " repartition des responsables des PME selon le niveau d'instruction")) %>%bold\_labels() %>%italicize\_labels()  
t\_instruction

| **Characteristic** | **N = 250**1 |
| --- | --- |
| ***repartition des responsables des PME selon le niveau d'instruction*** |  |
| Aucun niveau | 79 (32%) |
| Niveau primaire | 56 (22%) |
| Niveau secondaire | 74 (30%) |
| Niveau Superieur | 41 (16%) |
| 1n (%) | |

#### repartition des PME selon le statut juridique

# que faire sans tbl\_summary ? je l'applique a q12  
t\_juridique <- projet\_merge[c("q12")] %>%  
 tbl\_summary(   
 label = (q12 ~ " repartition des PME selon leur statut juridique")  
 ) %>%  
 bold\_labels() %>%  
 italicize\_labels()  
t\_juridique

| **Characteristic** | **N = 250**1 |
| --- | --- |
| ***repartition des PME selon leur statut juridique*** |  |
| Association | 6 (2.4%) |
| GIE | 179 (72%) |
| Informel | 38 (15%) |
| SA | 7 (2.8%) |
| SARL | 13 (5.2%) |
| SUARL | 7 (2.8%) |
| 1n (%) | |

#### repartition des PME selon que cela soit dirige par un proprietaire/ locataire

# que faire sans tbl\_summary ? je l'applique a q81  
t\_loc <- projet\_merge[c("q81")] %>%  
 tbl\_summary(   
 label = (q81 ~ " repartition des PME selon leur proprietaire/locataire")) %>%  
 bold\_labels() %>%  
 italicize\_labels()  
t\_loc

| **Characteristic** | **N = 250**1 |
| --- | --- |
| ***repartition des PME selon leur proprietaire/locataire*** |  |
| Locataire | 24 (9.6%) |
| Propriétaire | 226 (90%) |
| 1n (%) | |

#### repartition des PME selon le sexe et le niveau d’instruction des responsables

t\_sexe\_ins <- projet\_merge %>%  
 tbl\_cross(  
 row= q25,  
 col=sexe,  
 percent = "column",  
 label = list(sexe ~ " repartition des PME selon le sexe de leur responsable" , q25 ~ " repartition des responsables des PME selon le niveau d'instruction"))  
  
t\_sexe\_ins

|  | repartition des PME selon le sexe de leur responsable | |  |
| --- | --- | --- | --- |
|  | Femme | Homme | Total |
| repartition des responsables des PME selon le niveau d'instruction |  |  |  |
| Aucun niveau | 70 (37%) | 9 (15%) | 79 (32%) |
| Niveau primaire | 48 (25%) | 8 (14%) | 56 (22%) |
| Niveau secondaire | 56 (29%) | 18 (31%) | 74 (30%) |
| Niveau Superieur | 17 (8.9%) | 24 (41%) | 41 (16%) |
| Total | 191 (100%) | 59 (100%) | 250 (100%) |

t\_sexe\_instr <- projet\_merge[c("sexe","q25")] %>%  
 tbl\_summary()  
t\_sexe\_instr

## Table printed with {flextable}, not {gt}. Learn why at  
## https://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html  
## To suppress this message, include `message = FALSE` in the code chunk header.

| **Characteristic** | **N = 250**1 |
| --- | --- |
| sexe |  |
| Femme | 191 (76%) |
| Homme | 59 (24%) |
| q25 |  |
| Aucun niveau | 79 (32%) |
| Niveau primaire | 56 (22%) |
| Niveau secondaire | 74 (30%) |
| Niveau Superieur | 41 (16%) |
| 1n (%) | |

#### Etudions la répartition selon le statut juridique et le sexe

t\_sexe\_jur <- projet\_merge %>%  
 tbl\_cross(  
 row= q12,  
 col=sexe,  
 percent = "column",  
 label = list(sexe ~ " repartition des PME selon le sexe de leur responsable" , q12 ~ " repartition des responsables des PME selon le statut juridique")) %>%  
 add\_p(source\_note = TRUE)  
t\_sexe\_jur

|  | repartition des PME selon le sexe de leur responsable | |  |
| --- | --- | --- | --- |
|  | Femme | Homme | Total |
| repartition des responsables des PME selon le statut juridique |  |  |  |
| Association | 3 (1.6%) | 3 (5.1%) | 6 (2.4%) |
| GIE | 149 (78%) | 30 (51%) | 179 (72%) |
| Informel | 32 (17%) | 6 (10%) | 38 (15%) |
| SA | 1 (0.5%) | 6 (10%) | 7 (2.8%) |
| SARL | 2 (1.0%) | 11 (19%) | 13 (5.2%) |
| SUARL | 4 (2.1%) | 3 (5.1%) | 7 (2.8%) |
| Total | 191 (100%) | 59 (100%) | 250 (100%) |
| Fisher's exact test, p<0.001 | | | |

#### Tableau général

# Et si je faisais un tableau ou je combine tous ?  
table\_gene <- projet\_merge %>%   
 tbl\_summary(  
 include = c(q12, q25, q81, sexe), # différentes variables à inclure dans le tableau  
 by = sexe, # regroupement par sexe  
 label = list(  
 sexe ~ "Répartition des PME selon le sexe de leur responsable",  
 q12 ~ "Répartition des responsables des PME selon le statut juridique",  
 q81 ~ "Répartition des PME selon leur propriétaire/locataire",  
 q25 ~ "Répartition des responsables des PME selon le niveau d'instruction")) %>%  
 bold\_labels() %>% # ajout des styles aux libellés  
 italicize\_labels() %>%  
 modify\_header(  
 label = "\_\_variable\_\_",  
 stat\_1 = "\*\*Femme\*\*, N = 191",  
 stat\_2 = "\*\*Homme\*\*, N = 59")  
table\_gene

| variable | **Femme**, N = 1911 | **Homme**, N = 591 |
| --- | --- | --- |
| ***Répartition des responsables des PME selon le statut juridique*** |  |  |
| Association | 3 (1.6%) | 3 (5.1%) |
| GIE | 149 (78%) | 30 (51%) |
| Informel | 32 (17%) | 6 (10%) |
| SA | 1 (0.5%) | 6 (10%) |
| SARL | 2 (1.0%) | 11 (19%) |
| SUARL | 4 (2.1%) | 3 (5.1%) |
| ***Répartition des responsables des PME selon le niveau d'instruction*** |  |  |
| Aucun niveau | 70 (37%) | 9 (15%) |
| Niveau primaire | 48 (25%) | 8 (14%) |
| Niveau secondaire | 56 (29%) | 18 (31%) |
| Niveau Superieur | 17 (8.9%) | 24 (41%) |
| ***Répartition des PME selon leur propriétaire/locataire*** |  |  |
| Locataire | 16 (8.4%) | 8 (14%) |
| Propriétaire | 175 (92%) | 51 (86%) |
| 1n (%) | | |

# j'ai quelque chose de magnifique

#### A notre imagination

##### Répartition selon le sexe par filière

t1<-subset(projet\_merge,filiere\_1==1)%>%  
 dplyr::select(sexe, q25, q12, q81,filiere\_1) %>%  
 gtsummary::tbl\_summary(  
 by = filiere\_1,  
 statistic = list(  
 all\_categorical() ~ "{n} / {N} ({p}%)"  
 ),  
 percent = "column"   
 ) %>%  
 modify\_header(label ~ "\*\*Variable\*\*")   
  
t2<-subset(projet\_merge,filiere\_2==1)%>%  
 dplyr::select(sexe, q25, q12, q81,filiere\_2) %>%  
 gtsummary::tbl\_summary(  
 by = filiere\_2,  
 statistic = list(  
 all\_categorical() ~ "{n} / {N} ({p}%)"  
 ),   
 missing = "no",  
 percent = "column"   
 ) %>%  
 modify\_header(label ~ "\*\*Variable\*\*") %>%  
 bold\_labels()   
  
t3<-subset(projet\_merge,filiere\_3==1)%>%  
 dplyr::select(sexe, q25, q12, q81,filiere\_3) %>%  
 gtsummary::tbl\_summary(  
 by = filiere\_3,  
 statistic = list(  
 all\_categorical() ~ "{n} / {N} ({p}%)"  
 ),   
 missing = "no",  
 percent = "column"   
 ) %>%  
 modify\_header(label ~ "\*\*Variable\*\*") %>%  
 bold\_labels()  
  
t4<-subset(projet\_merge,filiere\_4==1)%>%  
 dplyr::select(sexe, q25, q12, q81,filiere\_4) %>%  
 gtsummary::tbl\_summary(  
 by = filiere\_4,  
 statistic = list(  
 all\_categorical() ~ "{n} / {N} ({p}%)"  
 ),   
 missing = "no",  
 percent = "column"   
 ) %>%  
 modify\_header(label ~ "\*\*Variable\*\*") %>%  
 bold\_labels()

Nous allons fusionner les quatres bases

gtsummary::tbl\_merge(list(t1,t2,t3,t4),  
 tab\_spanner = c("Arachide", "Anacarde","Mangue","Riz"))

|  | Arachide | Anacarde | Mangue | Riz |
| --- | --- | --- | --- | --- |
| **Variable** | **1**, N = 1081 | **1**, N = 611 | **1**, N = 891 | **1**, N = 921 |
| **sexe** |  |  |  |  |
| Femme | 93 / 108 (86%) | 40 / 61 (66%) | 68 / 89 (76%) | 77 / 92 (84%) |
| Homme | 15 / 108 (14%) | 21 / 61 (34%) | 21 / 89 (24%) | 15 / 92 (16%) |
| **q25** |  |  |  |  |
| Aucun niveau | 43 / 108 (40%) | 13 / 61 (21%) | 26 / 89 (29%) | 11 / 92 (12%) |
| Niveau primaire | 23 / 108 (21%) | 17 / 61 (28%) | 24 / 89 (27%) | 26 / 92 (28%) |
| Niveau secondaire | 34 / 108 (31%) | 15 / 61 (25%) | 25 / 89 (28%) | 32 / 92 (35%) |
| Niveau Superieur | 8 / 108 (7.4%) | 16 / 61 (26%) | 14 / 89 (16%) | 23 / 92 (25%) |
| **q12** |  |  |  |  |
| Association | 2 / 108 (1.9%) | 3 / 61 (4.9%) |  | 2 / 92 (2.2%) |
| GIE | 79 / 108 (73%) | 35 / 61 (57%) | 73 / 89 (82%) | 77 / 92 (84%) |
| Informel | 23 / 108 (21%) | 12 / 61 (20%) | 5 / 89 (5.6%) | 3 / 92 (3.3%) |
| SA | 2 / 108 (1.9%) | 2 / 61 (3.3%) | 3 / 89 (3.4%) | 3 / 92 (3.3%) |
| SARL | 1 / 108 (0.9%) | 6 / 61 (9.8%) | 6 / 89 (6.7%) | 5 / 92 (5.4%) |
| SUARL | 1 / 108 (0.9%) | 3 / 61 (4.9%) | 2 / 89 (2.2%) | 2 / 92 (2.2%) |
| **q81** |  |  |  |  |
| Locataire | 12 / 108 (11%) | 7 / 61 (11%) | 11 / 89 (12%) | 9 / 92 (9.8%) |
| Propriétaire | 96 / 108 (89%) | 54 / 61 (89%) | 78 / 89 (88%) | 83 / 92 (90%) |
| 1n / N (%) | | | | |

#### croisement avec date

#### huxtable

#### recoder la variable date\_enquete

# je renomme la variable date pour ceer les statistique   
# creer un intervalle   
intervalle <- c(0,1,10,Inf)  
  
projet\_mod$duree <- cut(projet\_mod$duree\_enquetes, breaks = intervalle, labels = c("inferieur a 1 jour", "entre 1 et 10 jours", "Plus de 10"))  
projet\_mod

## key region departement  
## 1 uuid:004b9117-d180-4031-a6af-6b4efabb5f53 Diourbel Bambey  
## 2 uuid:007d8eb4-45eb-44f4-aeac-722adc60aec8 Thiès Tivaouane  
## 3 uuid:030ada55-8dd2-4f57-b1b7-aaccd707c118 Saint-Louis Dagana  
## 4 uuid:04b4cd8d-0297-4dc0-9715-9c07120bab23 Diourbel Mbacké  
## 5 uuid:080548bf-e68a-49b8-9f04-f920b44244aa Saint-Louis Dagana

* statistiques descriptives

table\_date<-projet\_mod %>%   
 tbl\_summary(  
 include = c("region", "departement", "duree", "q12", "q16", "q19", "q17"), # différentes variables à inclure dans le tableau  
 by = "duree", # regroupement par sexe  
 percent="col",  
 label = list( departement ~ "Répartition des PME selon le departement ",  
 region ~ "Répartition des PME selon la region ", q12 ~ "Répartition des responsables des PME selon le statut juridique", q16 ~ "L’entreprise est-elle désservie par une route bitumée ? ", q17 ~ "Etat de la route bitumée",  
 q19 ~ "repartition des PME suivant l'etat de la piste qui mene a l'entreprise", duree~ "durre des enquete en annee"  
   
   
 )  
 ) %>%  
 bold\_labels() %>% # ajout des styles aux libellés  
 italicize\_labels()%>%  
 modify\_header(  
 label = "\_\_variable\_\_",  
 stat\_1 = "\*\*inferieur a un jour \*\*, ({n}/{N})",  
 stat\_2 = "\*\*entre 1 et 10 jours\*\*, ({n}/{N})",  
 stat\_3 = "\*\*plus de 10 jours\*\*, ({n}/{N})"  
   
   
   
)%>% as\_flex\_table()   
  
table\_date

| Variable | **inferieur a un jour** , (57/250)1 | **entre 1 et 10 jours**, (140/250)1 | **plus de 10 jours**, (53/250)1 |
| --- | --- | --- | --- |
| ***Répartition des PME selon la region*** |  |  |  |
| Dakar | 0 (0%) | 0 (0%) | 1 (1.9%) |
| Diourbel | 21 (37%) | 12 (8.6%) | 1 (1.9%) |
| Fatick | 6 (11%) | 18 (13%) | 6 (11%) |
| Kaffrine | 3 (5.3%) | 5 (3.6%) | 0 (0%) |
| Kaolack | 0 (0%) | 15 (11%) | 6 (11%) |
| Kolda | 0 (0%) | 8 (5.7%) | 1 (1.9%) |
| Saint-Louis | 0 (0%) | 14 (10%) | 28 (53%) |
| Sédhiou | 0 (0%) | 4 (2.9%) | 0 (0%) |
| Thiès | 26 (46%) | 21 (15%) | 4 (7.5%) |
| Ziguinchor | 1 (1.8%) | 43 (31%) | 6 (11%) |
| ***Répartition des PME selon le departement*** |  |  |  |
| Bambey | 10 (18%) | 9 (6.4%) | 1 (1.9%) |
| Bignona | 0 (0%) | 11 (7.9%) | 2 (3.8%) |
| Birkelane | 0 (0%) | 1 (0.7%) | 0 (0%) |
| Dagana | 0 (0%) | 12 (8.6%) | 27 (51%) |
| Diourbel | 9 (16%) | 3 (2.1%) | 0 (0%) |
| Fatick | 2 (3.5%) | 13 (9.3%) | 0 (0%) |
| Foundiougne | 4 (7.0%) | 5 (3.6%) | 6 (11%) |
| Goudomp | 0 (0%) | 3 (2.1%) | 0 (0%) |
| Kaffrine | 3 (5.3%) | 1 (0.7%) | 0 (0%) |
| Kaolack | 0 (0%) | 14 (10%) | 2 (3.8%) |
| Kolda | 0 (0%) | 7 (5.0%) | 0 (0%) |
| Koungheul | 0 (0%) | 3 (2.1%) | 0 (0%) |
| Mbacké | 2 (3.5%) | 0 (0%) | 0 (0%) |
| Mbour | 14 (25%) | 8 (5.7%) | 0 (0%) |
| Médina Yoro Foula | 0 (0%) | 1 (0.7%) | 0 (0%) |
| Nioro | 0 (0%) | 1 (0.7%) | 4 (7.5%) |
| Oussouye | 0 (0%) | 2 (1.4%) | 0 (0%) |
| Podor | 0 (0%) | 0 (0%) | 1 (1.9%) |
| Rufisque | 0 (0%) | 0 (0%) | 1 (1.9%) |
| Saint-Louis | 0 (0%) | 2 (1.4%) | 0 (0%) |
| Sédhiou | 0 (0%) | 1 (0.7%) | 0 (0%) |
| Thiès | 11 (19%) | 9 (6.4%) | 3 (5.7%) |
| Tivaouane | 1 (1.8%) | 4 (2.9%) | 1 (1.9%) |
| Velingara | 0 (0%) | 0 (0%) | 1 (1.9%) |
| Ziguinchor | 1 (1.8%) | 30 (21%) | 4 (7.5%) |
| ***Répartition des responsables des PME selon le statut juridique*** |  |  |  |
| Association | 0 (0%) | 4 (2.9%) | 2 (3.8%) |
| GIE | 42 (74%) | 101 (72%) | 36 (68%) |
| Informel | 14 (25%) | 20 (14%) | 4 (7.5%) |
| SA | 0 (0%) | 2 (1.4%) | 5 (9.4%) |
| SARL | 1 (1.8%) | 7 (5.0%) | 5 (9.4%) |
| SUARL | 0 (0%) | 6 (4.3%) | 1 (1.9%) |
| ***L’entreprise est-elle désservie par une route bitumée ?*** |  |  |  |
| Non | 37 (65%) | 76 (54%) | 17 (33%) |
| Oui | 20 (35%) | 64 (46%) | 35 (67%) |
| Unknown | 0 | 0 | 1 |
| ***repartition des PME suivant l'etat de la piste qui mene a l'entreprise*** |  |  |  |
| Bon état | 0 (0%) | 2 (2.6%) | 0 (0%) |
| Etat moyen | 27 (73%) | 39 (51%) | 7 (41%) |
| Mauvais état | 10 (27%) | 35 (46%) | 10 (59%) |
| Unknown | 20 | 64 | 36 |
| ***Etat de la route bitumée*** |  |  |  |
| Bon état | 13 (65%) | 46 (72%) | 33 (94%) |
| Etat moyen | 7 (35%) | 15 (23%) | 2 (5.7%) |
| Mauvais état | 0 (0%) | 3 (4.7%) | 0 (0%) |
| Unknown | 37 | 76 | 18 |
| 1n (%) | | | |

## 3 Un peu de cartographie

### Importation des packages nécessaires

library(ggplot2)  
library(sf)

#install.packages(st)  
library(st)

library(readxl)

### Importation de la base de données

Base\_Partie\_2 <- read\_excel("~/ISEP3/R/Base\_Partie 2.xlsx")

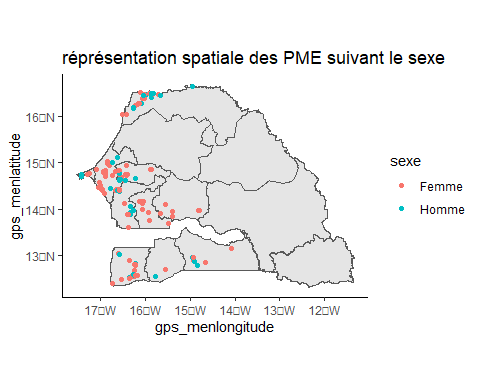
### Transformer le data.frame en données géographiques dont l’objet sera nommé projet\_map.

# Charger le package sf  
library(sf)  
# Transformer le data.frame en objet spatial  
projet\_map <- st\_as\_sf(projet, coords = c("gps\_menlongitude", "gps\_menlatitude"), crs = 4326)  
senegal <- st\_read("C:\\Users\\LENOVO T470\\Documents\\ISEP3\\R\\gadm41\_SEN\_1.shp")

## Reading layer `gadm41\_SEN\_1' from data source   
## `C:\Users\LENOVO T470\Documents\ISEP3\R\gadm41\_SEN\_1.shp' using driver `ESRI Shapefile'  
## Simple feature collection with 14 features and 11 fields  
## Geometry type: MULTIPOLYGON  
## Dimension: XY  
## Bounding box: xmin: -17.54319 ymin: 12.30786 xmax: -11.34247 ymax: 16.69207  
## Geodetic CRS: WGS 84

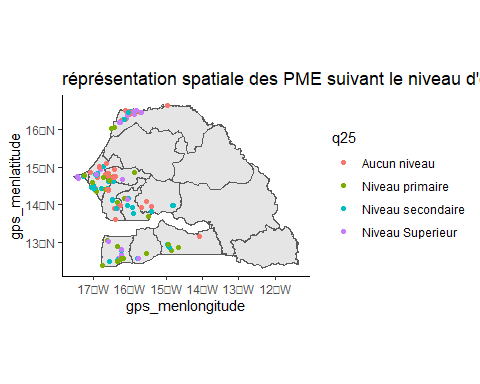
### Réprésentation spatiale des PME suivant le sexe

ggplot(projet)+geom\_sf(data=senegal)+geom\_point(aes(x=gps\_menlongitude, y=gps\_menlatitude, color= sexe))+ labs (title= "réprésentation spatiale des PME suivant le sexe", x="gps\_menlongitude", y="gps\_menlatitude" )+ theme\_classic()



### réprésentation spatiale des PME suivant le niveau d’ education

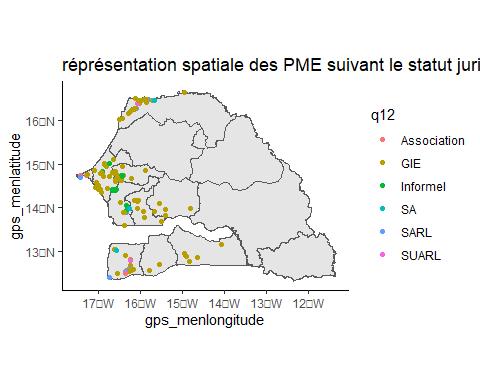
ggplot(projet)+geom\_sf(data=senegal)+geom\_point(aes(x=gps\_menlongitude, y=gps\_menlatitude, color= q25))+ labs (title= "réprésentation spatiale des PME suivant le niveau d'education ", x="gps\_menlongitude", y="gps\_menlatitude" )+ theme\_classic()



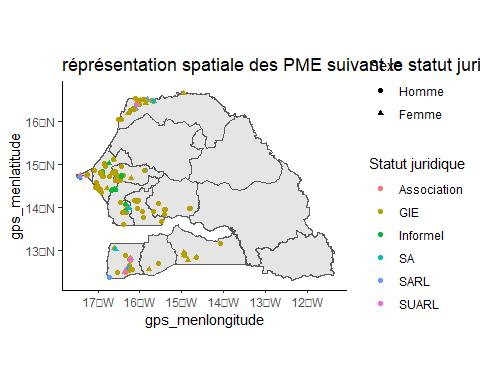
### Faisons une analyse spatiale de votre choix

#### Selon le statut juridique

ggplot(projet)+geom\_sf(data=senegal)+geom\_point(aes(x=gps\_menlongitude, y=gps\_menlatitude, color= q12))+ labs (title= "réprésentation spatiale des PME suivant le statut juridique ", x="gps\_menlongitude", y="gps\_menlatitude" )+ theme\_classic()



ggplot(projet)+geom\_sf(data=senegal)+geom\_point(aes(x=gps\_menlongitude, y=gps\_menlatitude, color= q12, shape=sexe))+ labs (title= "réprésentation spatiale des PME suivant le statut juridique et le sexe du responsable", x="gps\_menlongitude", y="gps\_menlatitude" )+ scale\_color\_discrete(name = "Statut juridique") +  
 scale\_shape\_discrete(name = "Sexe", labels = c("Homme", "Femme")) +  
 theme\_classic()



## Partie 3

Nous avons mis dans le compte github les codes et bases concernées pour ce qui concerne la partie 3. A travers un fichier nommé "app.R" et une base de données nommée "ACLED-Western\_Africa" on peut ouvrir l'application.