

République du Sénégal

Un Peuple-Un But-Une Fois



MINISTERE DE L'ECONOMIE DU PLAN ET DE LA COOPERATION

Agence nationale de la Statistique et de la Démographie



Ecole nationale de la Statistique et de l'Analyse économique



Statistiques Exploratoires Spatiales avec R

Analyse de la Malaria au Cameroun

Rédigé par :

Jeanne De La Flèche ONANENA AMANA

Khadidiatou COULIBALY

Samba DIENG

Tamsir NDONG

Elèves en ISE1 Cycle

Sous la supervision de :

M. Aboubacar HEMA

Année Académique

2024- 2025

CHAPTER 1

LE TP PROPREMENT DIT

1.1 CONSIGNE

1. Importer et visualiser le raster population
2. Calculer le nombre de personnes par admin et exporter sous format .csv
3. Ramener la taille des pixels à 5km
4. Visualiser le nouveau raster de la population
5. Calculer un nouveau raster d'enfants de 0 à 12ans (0.1%)
6. Créer 3 nouveaux rasters binarisés
7. Multiplier chacun d'eux par celui de la population
8. Calculer nombre d'enfants atteints de la malaria par admin et exporter
9. Quel est le taux d'enfants atteints de malaria par admin ?

1.2 Chargement des packages

```
library(stars)          ## Manipulation des données raster et vecteur
library(sf)              ## Manipulation des objets géospatiaux
library(ggplot2)         ## Visualisations graphiques
library(ggspatial)       ## Ajout d'éléments cartographiques (flèche du nord, échelle)
library(raster)          ## Manipulation des données raster
library(cowplot)         ## Extraction et affichage de la légende
library(leaflet)         ## Création de cartes interactives avec limites administratives
library(viridis)          ## Palette de couleurs viridis
##library(units)          ## Gestion des unités dans les données géospatiales
library(dplyr)            ## Gestion des bases de données
library(exactextractr)   ##Une sorte de "reducer"
library(kableExtra)
library(knitr)

## Setting the work directory
knitr::opts_knit$set(root.dir =
"C:/Users/DELL/Documents/ISEP3_2025/Stats_spatiale/Statistique-Exploratoire-Spatiale/")

knitr::opts_chunk$set(fig.align = "center", fig.width = 6)
```

```

## Charger les rasters

cameroun<- st_read("cmr_admbnda_adm0_inc_20180104.shp", quiet = TRUE)
region <- st_read("cmr_admbnda_adm1_inc_20180104.shp", quiet = TRUE)
departement <- st_read("cmr_admbnda_adm2_inc_20180104.shp", quiet = TRUE)
arrondissement <- st_read("cmr_admbnda_adm3_inc_20180104.shp", quiet = TRUE)

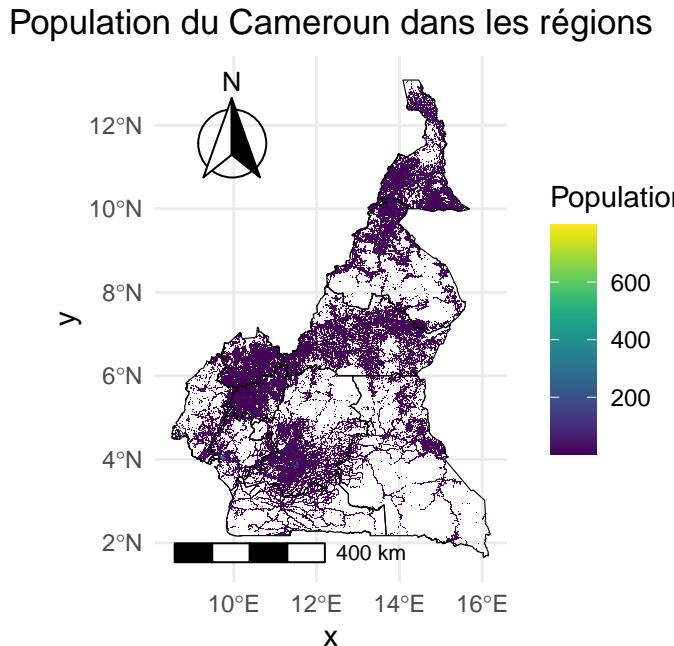
## Charger les rasters en faisant un crop et un mask
malaria_2022_CMR <-raster("Indic2022_taux_maliaria_3niveaux.tif") %>%
  crop(cameroun) %>%
  mask(cameroun)

taux_2022 <- raster("202406_Global_Pf_Parasite_Rate_CMR_2022.tiff") %>%
  crop(cameroun) %>%
  mask(cameroun)

pop <- raster("CMR_population_v1_0_gridded.tif") %>%
  crop(cameroun) %>%
  mask(cameroun)

pop_df <- as.data.frame(rasterToPoints(pop), stringsAsFactors = FALSE)
colnames(pop_df) <- c("x", "y", "value")
ggplot() +
  geom_tile(data = pop_df, aes(x = x, y = y, fill = value)) +
  geom_sf(data = region, fill = NA, color = "black", size = 0.5) +
  ## Ajouter les contours des régions
  scale_fill_viridis(option = "viridis", na.value = "transparent") +
  labs(title = "Population du Cameroun dans les régions",
       fill = "Population") +
  theme_minimal() +
  theme(
    plot.title = element_text(hjust = 0.5),
    legend.position = "right"
  ) +
  annotation_north_arrow(location = "tl",
                         which_north = "true", pad_x = unit(0.1, "in"),
                         pad_y = unit(0.1, "in"),
                         style = north_arrow_fancy_orienteering()) +
  annotation_scale(location = "bl", width_hint = 0.5)

```



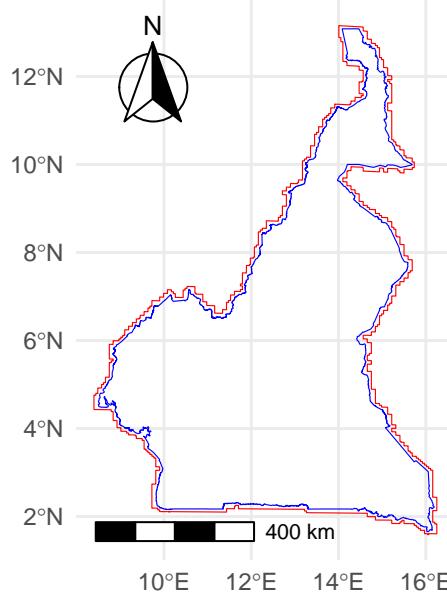
On étend artificiellement notre shapefile du pays

```
cameroun_1 <- st_buffer(cameroun, dist = 5000)
```

Représentons

```
ggplot() +
  geom_sf(data = cameroun_1, fill= NA, color = "red", size = 0.5) + # Ajout du shapefile étendu
  geom_sf(data = cameroun, fill= NA, color = "blue", size = 0.5) + # Ajout du shp original
  scale_fill_viridis(option = "viridis", na.value = "transparent") +
  labs(title = "Cameroun avec et sans buffer") +
  theme_minimal() +
  theme(
    plot.title = element_text(hjust = 0.5),
    legend.position = "right"
  ) +
  annotation_north_arrow(location = "tl",
    which_north = "true", pad_x = unit(0.1, "in"),
    pad_y = unit(0.1, "in"), style = north_arrow_fancy_orienteering()) +
  annotation_scale(location = "bl", width_hint = 0.5)
```

Cameroun avec et sans buffer



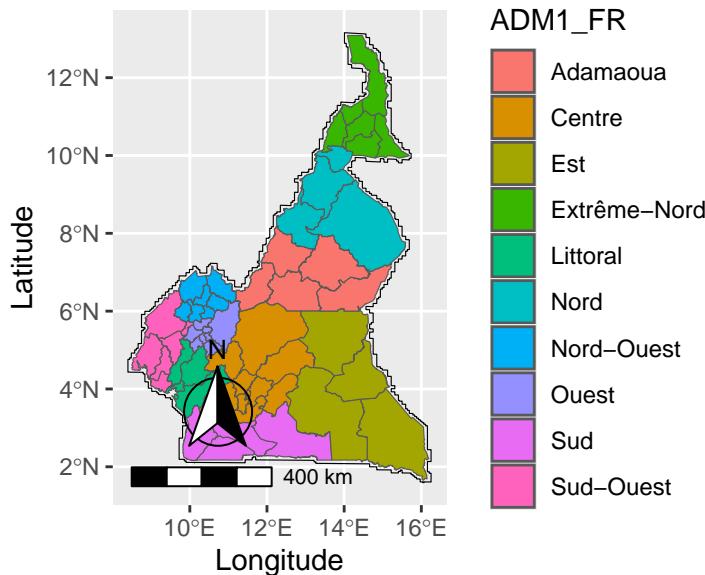
On va maintenant faire une “intersection”

```
region <- suppressWarnings(st_intersection(region, cameroun_1))
departement <- suppressWarnings(st_intersection(departement, cameroun_1))
arrondissement <- suppressWarnings(st_intersection(arrondissement, cameroun_1))
```

Il semble que cela redonne les mêmes shp... Visualisons

```
# 
ggplot() +
  geom_sf(data = departement, mapping = aes(fill = ADM1_FR)) +
  geom_sf(data = cameroun_1, fill= NA, color = "black", size = 0.5) + # Ajout du shp
  ggttitle("Départements") + # Titre de la carte
  theme(plot.title = element_text(hjust = 0.5, size = 25)) +
  xlab("Longitude") +
  ylab("Latitude") +
  annotation_north_arrow(location = "bl", which_north = "true", pad_x = unit(0.25, "in"))
  annotation_scale(location = "bl", width_hint = 0.5)
```

Départements



1.4 Calcul du nombre de personnes par admin (0-3) et exportation sous format .csv

```
## Calculer la population pour chaque niveau administratif
pop_admin0 <- exact_extract(pop, cameroun, fun = "sum",
                             progress = FALSE)
cameroun$population <- pop_admin0

pop_cameroun <- cameroun %>%
  st_drop_geometry() %>%
  dplyr::select(ADM0_FR, population)
```

```
pop_cameroun <- data.frame(pop_cameroun)
```

```
kable(pop_cameroun)
```

ADM0_FR	population
Cameroun (le)	28611876

```
pop_admin1 <- exact_extract(pop, region, fun = "sum",
                             progress = FALSE)
region$population <- pop_admin1

pop_region <- region %>%
  st_drop_geometry() %>%
  dplyr::select(ADM1_FR, population)

pop_region <- data.frame(pop_region)

kable(head(pop_region, 5))
```

ADM1_FR	population
Adamaoua	1649247
Centre	5207611
Est	1508061
Extrême-Nord	5346909
Littoral	4456166

```
pop_admin2 <- exact_extract(pop, departement, fun = "sum",
                             progress = FALSE)
departement$population <- pop_admin2

pop_departement <- departement %>%
  st_drop_geometry() %>%
  select(ADM2_FR, population)

pop_departement <- data.frame(pop_departement)

kable(head(pop_departement, 5))
```

ADM2_FR	population
Bamboutos	385835.3
Bénoué	1660537.5
Boumba-et-Ngoko	157055.0
Boyo	135210.3
Bui	261324.8

```
pop_admin3 <- exact_extract(pop, arrondissement, fun = "sum", progress = FALSE)
arrondissement$population <- pop_admin3

pop_arrondissement <- arrondissement %>%
  st_drop_geometry() %>%
  select(ADM3_FR, population)

pop_arrondissement <- data.frame(pop_arrondissement)

kable(head(pop_arrondissement, 5))
```

ADM3_FR	population
Abong-Mbang	40934.281
Afanloum	3442.716
Ako	33204.461
Akoeman	8750.636
Akom 2	13346.139

```
# Vérifications
t <- data.frame(sum_region = sum(pop_region$population),
                 sum_departements=sum(pop_departement$population),
                 sum_arrondissements=sum(pop_arrondissement$population))
```

```
kable(t)
```

sum_region	sum_departements	sum_arrondissements
28611876	28611876	28611876

```
output_folder <- "Outputs/"
dir.create(output_folder, showWarnings = FALSE)

## Exporter chaque data frame sous format CSV
write.csv(pop_cameroun, file.path(output_folder,
                                    "pop_cameroun.csv"), row.names = FALSE)
write.csv(pop_region, file.path(output_folder,
                                 "pop_region.csv"), row.names = FALSE)
write.csv(pop_departement, file.path(output_folder,
                                       "pop_departement.csv"), row.names = FALSE)
write.csv(pop_arrondissement, file.path(output_folder,
                                         "pop_arondissement.csv"), row.names = FALSE)
```

1.5 Résolution à 5km par aggrégation (somme)

```
# Voyons la taille des pixels en degrés

res(pop)

## [1] 0.0008333333 0.0008333333

pop_newResw <- aggregate(pop, fact = 50, fun = sum,
                           filename = "CMR_population_aggregated_5km.tif",
                           overwrite = TRUE)

# facteur d'agrégation : 5000m/100m
# Agrégation des valeurs : somme des valeurs des pixels ('sum')
```

1.6 Visualisation du nouveau raster

Faisons d'abord une petite vérification...

```
# Vérifications
pop_a <- exact_extract(pop_newResw %>%
                           crop(cameroun_1) %>%
                           mask(cameroun_1),
                           cameroun_1, fun = "sum", progress = FALSE)

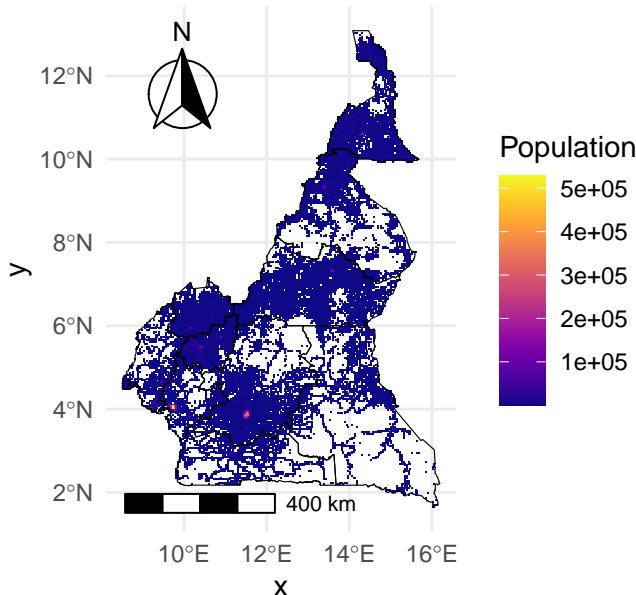
print( "La population totale avec le nouveau raster est")

## [1] "La population totale avec le nouveau raster est"
pop_a

## [1] 28614930
```

```
pop_df <- as.data.frame(rasterToPoints(pop_newResw), stringsAsFactors = FALSE)
colnames(pop_df) <- c("x", "y", "value")
ggplot() +
  geom_tile(data = pop_df, aes(x = x, y = y, fill = value)) +
  geom_sf(data = region, fill = NA, color = "black", size = 0.5) +
  scale_fill_viridis(option = "plasma", na.value = "transparent") +
  labs(title = "Population du Cameroun dans les régions",
       fill = "Population") +
  theme_minimal() +
  theme(
    plot.title = element_text(hjust = 0.5),
    legend.position = "right"
  ) +
  annotation_north_arrow(location = "tl", which_north = "true",
                          pad_x = unit(0.1, "in"), pad_y = unit(0.1, "in"), style = north_arrow_f
annotation_scale(location = "bl", width_hint = 0.5)
```

Population du Cameroun dans les régions



1.7 Raster des enfants de 0 à 12 ans (0.1% de la population)

```
## Multiplication
pop_child <- pop_newResw*0.001
```

Vérifions encore

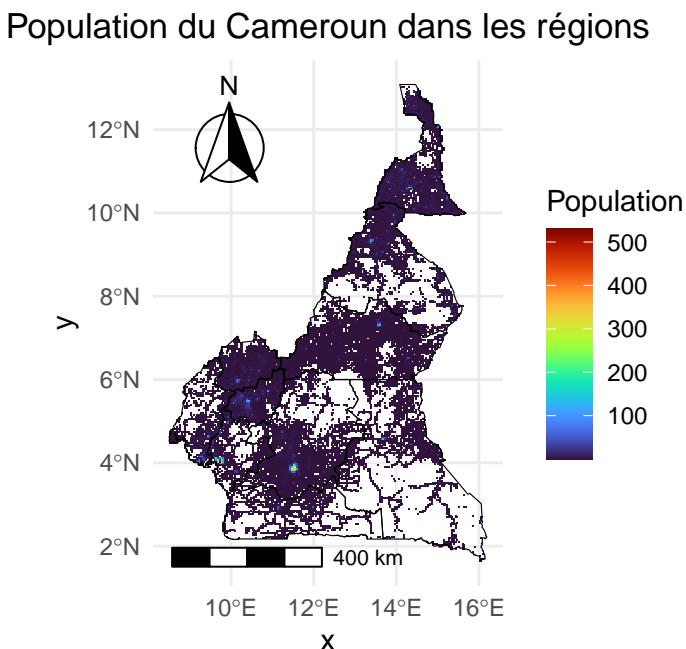
```
pop_b <- exact_extract(pop_child %>%
                           crop(cameroun_1) %>%
                           mask(cameroun_1), cameroun_1, fun = "sum",
                           progress = FALSE)
```

```
# Population d'enfants :
pop_b
```

```
## [1] 28614.93  
# Ration  
(pop_b/pop_a)*100  
  
## [1] 0.1
```

1.8 Visualisation du nouveau raster

```
pop_df <- data.frame(rasterToPoints(pop_child), stringsAsFactors = FALSE)
colnames(pop_df) <- c("x", "y", "value")
ggplot() +
  geom_tile(data = pop_df, aes(x = x, y = y, fill = value)) +
  geom_sf(data = region, fill = NA, color = "black", size = 0.5) + ## Ajouter les co
  scale_fill_viridis(option = "turbo", na.value = "transparent") +
  labs(title = "Population du Cameroun dans les régions",
       fill = "Population") +
  theme_minimal() +
  theme(
    plot.title = element_text(hjust = 0.5),
    legend.position = "right"
  ) +
  annotation_north_arrow(location = "tl", which_north = "true", pad_x = unit(0.1, "in")
  annotation_scale(location = "bl", width_hint = 0.5)
```



1.9 Rasters binarisés à partir de celui du taux de malaria

Utilisation du resampling... avec la méthode des plus proches voisins (nrb). On aligne d'abord les rasters suivant celui de la population

```
# Raster des taux en 2022
taux_2022_1 <- resample(taux_2022, pop_child, method = "ngb")

# Nombre d'enfants "par situation" (atteints ou non)
aucun <- calc(malaria_2022_CMR_1 == 1,
  fun = function(x) { ifelse(x, 1, 0)} *pop_child
moyen <- calc(malaria_2022_CMR_1 == 2,
  fun = function(x) { ifelse(x, 1, 0)} *pop_child
grave <- calc(malaria_2022_CMR_1 == 3,
  fun = function(x) { ifelse(x, 1, 0)} *pop_child
```

Calculons le nombre d'enfants atteints par situation.

```
# Enfants d'enfants atteints par situation
pop_aucun <- aucun*taux_2022_1
pop_moyen <- moyen*taux_2022_1
pop_grave <- grave*taux_2022_1
```

```
# Calculons le nombre d'enfants atteints par situation
A <- sum(values(pop_aucun), na.rm = TRUE)
B <- sum(values(pop_moyen), na.rm = TRUE)
C <- sum(values(pop_grave), na.rm = TRUE)

# Affichage
t1 <- data.frame(
  situation_aucun = c(A, A*100/(A+B+C)),
  situation_moyen = c(B, B*100/(A+B+C)),
  situation_grave = c(C, C*100/(A+B+C)),
  somme = c( A+B+C, A*100/(A+B+C)+ B*100/(A+B+C)+ C*100/(A+B+C))
)

kable(t1)
```

situation_aucun	situation_moyen	situation_grave	somme
5106.9328	506.40287	306.064302	5919.4
86.2745	8.55497	5.170529	100.0

Vérifications...

```
## Vérifications...
```

```
# Raster ayant toute la population malade
pop_malaria <- pop_child* resample(taux_2022, pop_child, method = "ngb")
enf_malades <- sum(values(pop_malaria), na.rm = TRUE)

t_1 <- data.frame(total_enfants_malades = enf_malades ) %>% t()

kable(t_1)
```

total_enfants_malades	5933.623
-----------------------	----------

1.10 Nombre d'enfants atteints par admin et enregistrer le résultat en fichier csv

```
pop_cameroun$pop_malade <- exact_extract(pop_malaria,
                                             cameroun, fun = "sum", progress = FALSE)

pop_region$pop_malade <- exact_extract(pop_malaria,
                                           region, fun = "sum", progress = FALSE)

pop_departement$pop_malade <- exact_extract(pop_malaria,
                                               departement, fun = "sum", progress = FALSE)

pop_arrondissement$pop_malade <- exact_extract(pop_malaria,
                                                 arrondissement, fun = "sum", progress = FALSE)
```

Exportons.

```
## Exporter chaque data frame sous format CSV
write.csv(pop_cameroun,
          file.path(output_folder, "cameroun.csv"), row.names = FALSE)
write.csv(pop_region,
          file.path(output_folder, "region_cmr.csv"), row.names = FALSE)
write.csv(pop_departement,
          file.path(output_folder, "departement_cmr.csv"), row.names = FALSE)
write.csv(pop_arrondissement,
          file.path(output_folder, "arondissement_cmr.csv"), row.names = FALSE)
```

Petites vérifications

Vérifications

```
t2 <- data.frame(
  enfts_malades_cameroun = sum(pop_cameroun$pop_malade),
  enfts_malades_region = sum(pop_region$pop_malade),
  enfts_malades_departements = sum(pop_departement$pop_malade),
  enfts_malades_arondissements = sum(pop_arrondissement$pop_malade)) %>% t()

kable(t2)
```

enfts_malades_cameroun	5909.672
enfts_malades_region	5909.673
enfts_malades_departements	5909.673
enfts_malades_arondissements	5909.673

Petite visualisation...

kable(head(pop_arrondissement, 10))

ADM3_FR	population	pop_malade
Abong-Mbang	40934.281	11.730131
Afanloum	3442.716	1.559192
Ako	33204.461	9.463091

ADM3_FR	population	pop_malade
Akoeman	8750.636	3.749743
Akom 2	13346.139	5.935190
Akono	13826.283	5.929521
Akonolinga	77040.172	26.515780
Akwaya	35409.289	9.439877
Alou	17685.900	2.812248
Ambam	63228.043	25.558369

1.11 Taux d'enfants atteints par admin

On prend la population d'enfants malades sur celle totale d'enfants

```
pop_cameroun$taux_malade <- pop_cameroun$pop_malade/(pop_cameroun$population*0.001)
```

```
pop_region$taux_malade <- pop_region$pop_malade/(pop_region$population*0.001)
```

```
pop_departement$taux_malade <- pop_departement$pop_malade/(pop_departement$population*0.001)
```

```
pop_arrondissement$taux_malade <- pop_arrondissement$pop_malade/(pop_arrondissement$population*0.001)
```

Petite visualisation...

```
kable(head(pop_arrondissement, 10))
```

ADM3_FR	population	pop_malade	taux_malade
Abong-Mbang	40934.281	11.730131	0.2865601
Afanloum	3442.716	1.559192	0.4528957
Ako	33204.461	9.463091	0.2849946
Akoeman	8750.636	3.749743	0.4285109
Akom 2	13346.139	5.935190	0.4447121
Akono	13826.283	5.929521	0.4288586
Akonolinga	77040.172	26.515780	0.3441812
Akwaya	35409.289	9.439877	0.2665933
Alou	17685.900	2.812248	0.1590107
Ambam	63228.043	25.558369	0.4042252