University of Ouaga 1 prof. Joseph Ki-Zerbo
--
Doctorate School of Informatique and climate change
--
Master Research Program

Course Title:

Spatial Data Analysis

Practical exercise

Student name:

KITEGI M. Simon Pierre

simkitegi@gmail.com

Lectures:

Steven Hill (steven.hill@uni-wuerzburg.de)

Dr. Sarah Schönbrodt-Stitt (sarah.schoenbrodt-stitt@uni-wuerzburg.de)

September 2020

Outline

The R language version 4 was used to do the exercises.

Exercise 1:

Import the fires dataset (fires.csv)

```
fires <- read.csv("data/fire.csv")
```

1.1 How many fires were detected in total?

```
number_of_fires <- nrow(fires)
number_of_fires

## [1] 21460
```

1.2 During which time period were the fires detected?

```
timePeriod <- c(head(fires$ACQ_DATE,1), tail(fires$ACQ_DATE,1))

timePeriod

## [1] "2019-01-01 - 2019-01-31"
```

1.3 Identify the ten brightest fires during this period. Where are they located?

```
library(dplyr)

ten_brightest_fires <- fires %>% arrange(desc(BRIGHTNESS)) %>% head(10)
ten_brightest_fires

##    LATITUDE LONGITUDE BRIGHTNESS SCAN TRACK   ACQ_DATE ACQ_TIME SATELL
ITE
## 1    9.6760   10.6130      422.9
## 2   10.5655    6.7566      421.8
## 3    8.1329   11.1870      419.7
## 4    7.6403    8.1262      418.0
## 5   12.4460    5.7779      413.4
## 6    7.6324    8.3984      405.4
## 7    8.6089   10.5270      399.9
## 8    9.6863   10.6208      397.0
## 9   11.3494   13.5145      396.7
```

```
## 10  10.0935    3.8743      395.2
```

The locations of the ten brightest fires:

```
library(leaflet)
leaflet(ten_brightest_fires) %>% addProviderTiles("Esri") %>%
addCircleMarkers(~LONGITUDE, ~LATITUDE, col = "red")
```

Figure 1: Ten brightest fires in Nigeria

1.4 Extract all fires detected with a confidence higher than 70 percent. How much are these? Show in a plot.

```
fires_higher_confidence <- fires %>% filter(CONFIDENCE > 70) %>% arrange(desc(CONFID
ENCE))
number_fire_higher_confidence<- nrow(fires_higher_confidence)
number_fire_higher_confidence

## [1] 7285

#Plot
library(ggplot2)
fires %>% ggplot(aes(x=LONGITUDE, y=LATITUDE, group = CONFIDENCE)) + geom_point(
```

```
aes(color = CONFIDENCE > 70)) + scale_color_manual(values=c("black", "#00AFBB"))+ ggtitle("Fires higher confidence") + theme_light()
```

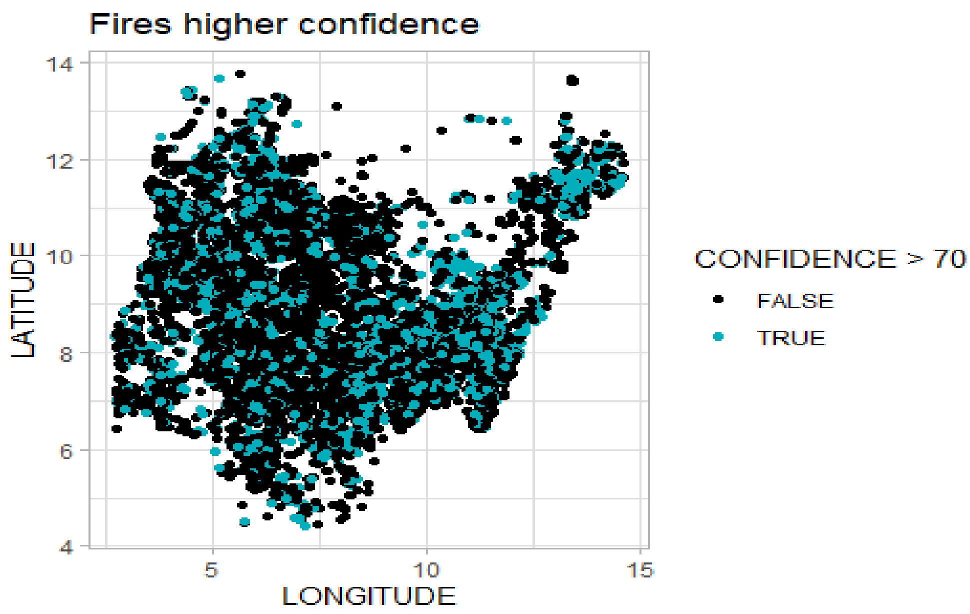


Figure 2: Higher confidence fires map

1.5 Create a histogram showing the distribution of the fire radiative power. Clearly indicate the description of the y-axis and x-axis

```
fires%>% ggplot(aes(FRP)) + geom_histogram(bins = 30, fill = "#00AFBB") + labs(title = "Histogram of the fire radiative power")
```

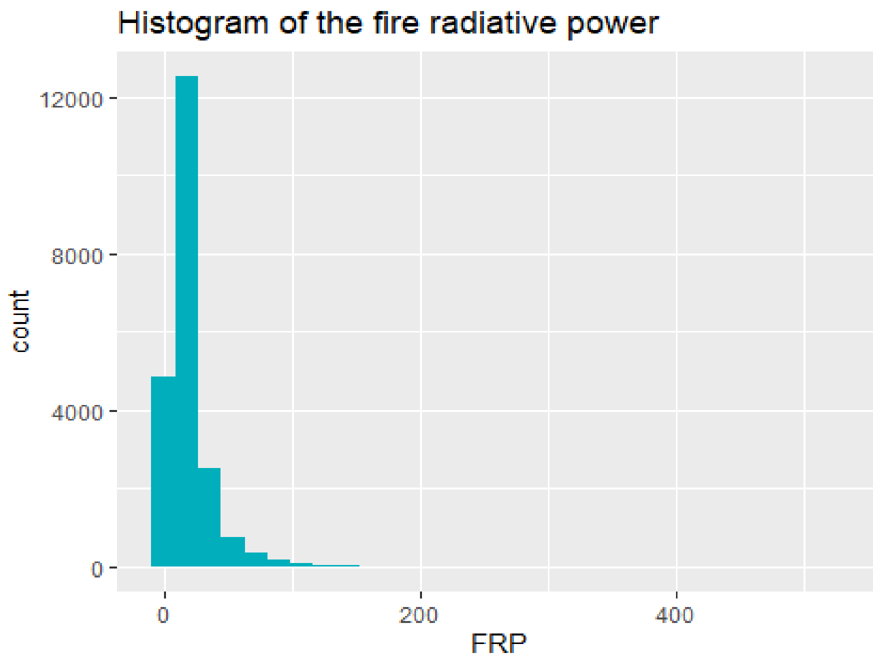


Figure 3: Histogram of the distribution of the radiative power

Description of the axis

The x-axis shows the majors intervals of the fire raiative power and the y-axis counts the number of stations according to x-axis intervals

1.6 Create a plot showing the numbers of fires for each day. Convert the dataset into a spatial "geopandas dataframe

I'm using R so i will convert into a spatial dataframe

```
library(lubridate)

Day_number_fires <- fires %>% group_by(day(ACQ_DATE)) %>% count()
colnames(Day_number_fires)<- c("Days", "Number_of_Fires")
Day_number_fires %>% ggplot(aes(Days, Number_of_Fires)) +  scale_x_continuous(breaks
= c(1:31)) +
  geom_bar(stat = "identity", fill = "#00AFBB") + theme_light()+  ggtitle("Number of fires per da
y")
```

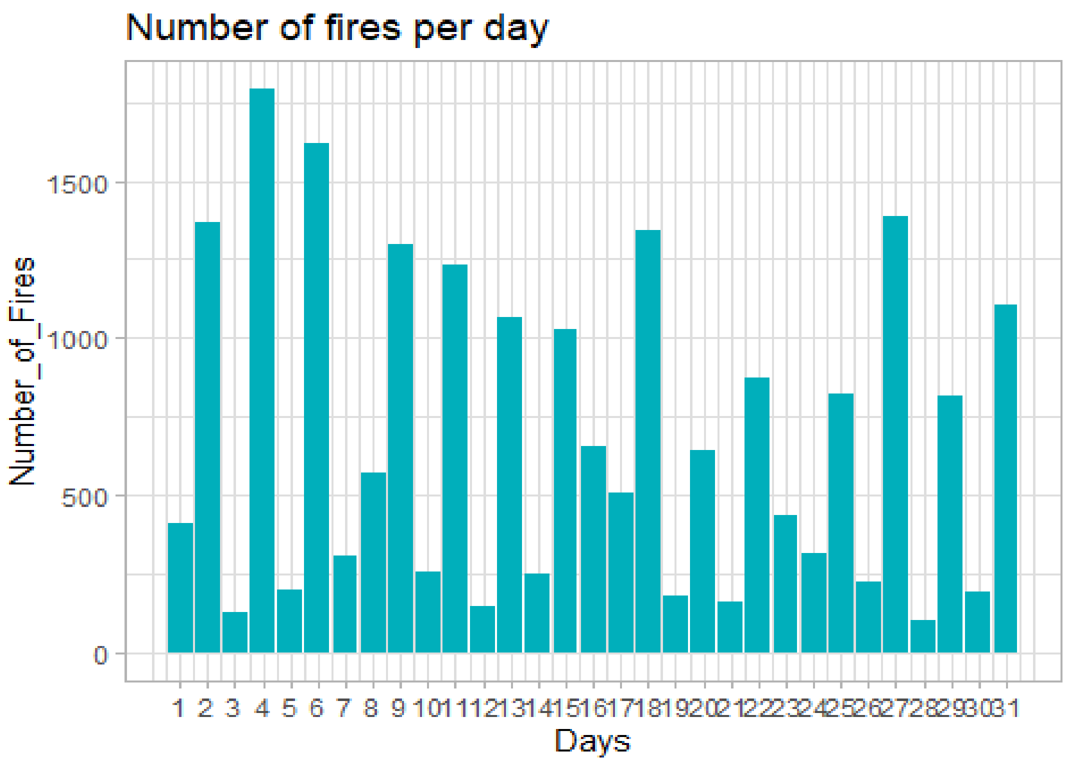


Figure 4: Number of fires per day

```
#convert the data to spatial dataframe
library(sf)library(spbabel)
fires_to_spatial_data <- st_as_sf(x = fires,
                   coords = c("LONGITUDE", "LATITUDE"), crs = 3850 )
```

1.7 Create a heat map showing the density of fires in Nigeria

```
library(leaflet)

library(leaflet.extras)
leaflet(fires) %>% addProviderTiles("Esri") %>%
addHeatmap(lng=~LONGITUDE, lat=~LATITUDE, intensity = ~BRIGHTNESS, radius=8)
```

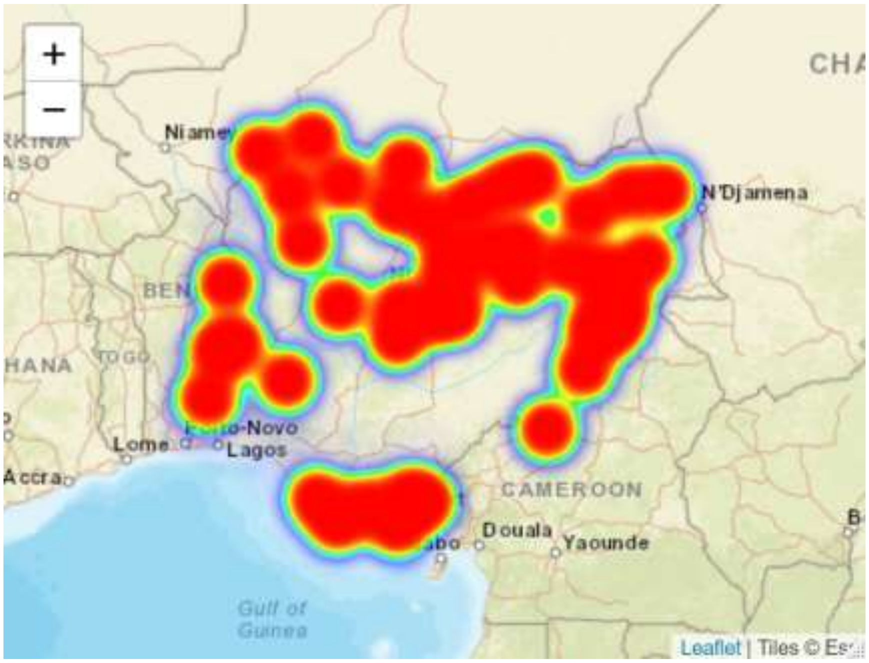


Figure 5: Heatmap of the fire's density in Nigeria

1.8 Investigate if the fires detected in Nigeria are distributed rather regular or if the fires appear rather clustered (if so try to identify the main clusters and explain)

The centred version of the K-function is K(r) − Kpois(r) = K(r) − πr2

This can be useful for classifying a point pattern as random, clustered, or regular, because the function is zero if the point pattern is completely random. It is less useful for other purposes. A commonly used transformation of K proposed by Besag [103] is the L-function L(r) = r K(r) π.

```
library(spatstat)

fires_ppp<- as.ppp(fires_to_spatial_data)

fires_ppp_unmark<- unmark.ppp(fires_ppp)

lest <- Lest(fires_ppp_unmark, correction = "isotropic")
plot(lest, main = "")
```

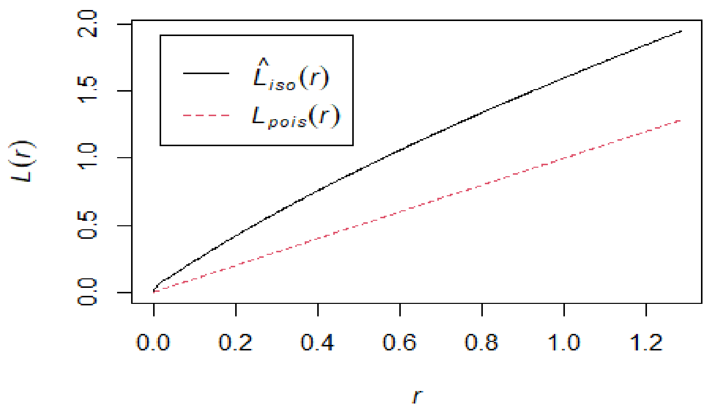


Figure6: Result from the lest test

Interpretations

The Lpois is lower than the Liso. It means that the data is clustered.

Identification of the main clusters:

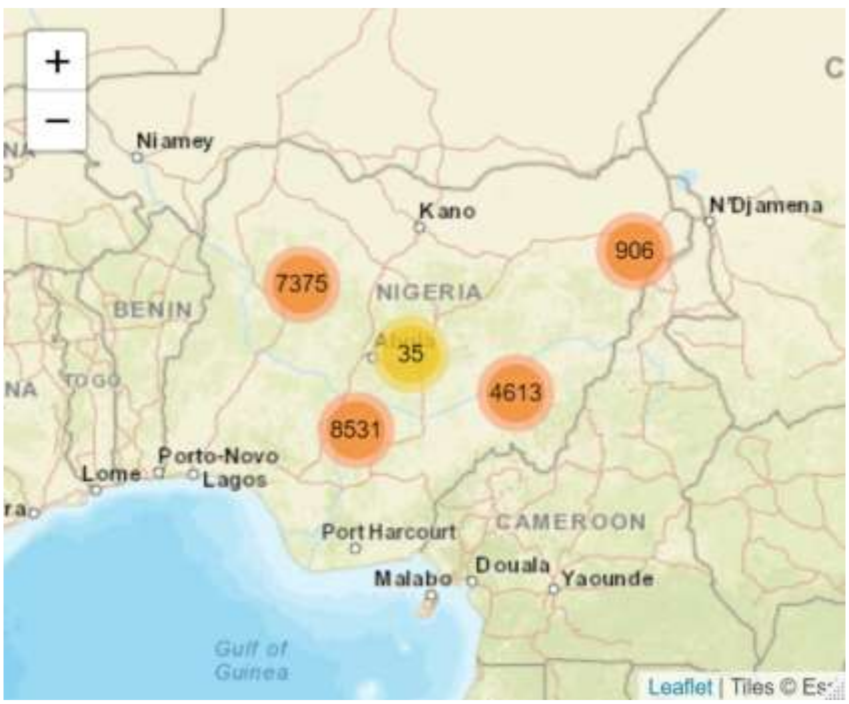


Figure 7: Clusters

Import the local area government dataset (nigeria_lga.shp).

```
library(sf)
nigeria_lga <- st_read("data/nigeria_lga.shp")

## Reading layer `nigeria_lga' from data source `C:\Users\Simon Pierre KITEGI\Documents\Proj
ect-SDA\data\nigeria_lga.shp' using driver `ESRI Shapefile'
## Simple feature collection with 774 features and 12 fields
## geometry type:  MULTIPOLYGON
## dimension:      XY
## bbox:           xmin: 299727.7 ymin: 472872.8 xmax: 1633944 ymax: 1550860
## projected CRS:  WGS 84 / World Mercator
```

1.9 Count the fires occurring in the different local government areas in Nigeria. How much are they?

```
#
nigeria_lga<- st_transform(nigeria_lga, 4326)
fires_sp<- st_set_crs(fires_to_spatial_data, 4326)

intersection <- st_intersection(x = nigeria_lga, y = fires_sp)

int_result <- intersection %>%
  group_by(LGAName) %>%
  count()

#we are displaying the two first result from the table


head(int_result,2)

## Simple feature collection with 2 features and 2 fields
## geometry type:  MULTIPOINT
## dimension:      XY
## bbox:           xmin: 6.7885 ymin: 8.4991 xmax: 13.4692 ymax: 13.6603
```

```
## geographic CRS: WGS 84
##   LGAName  n                  geometry
## 1  Abadam 12 MULTIPOINT ((13.3802 13.660...
## 2   Abaji 24 MULTIPOINT ((6.7885 8.6158)...
```

1.10 Create a map displaying the number of fires for each local government area

```
a<-st_join(nigeria_lga, int_result)

library(ggplot2)
library(tidyverse)

ggplot(data= a) +
  geom_sf(color = "#00AFBB") +
  geom_sf_text(aes(label = n), col = "black", size = 2)
```

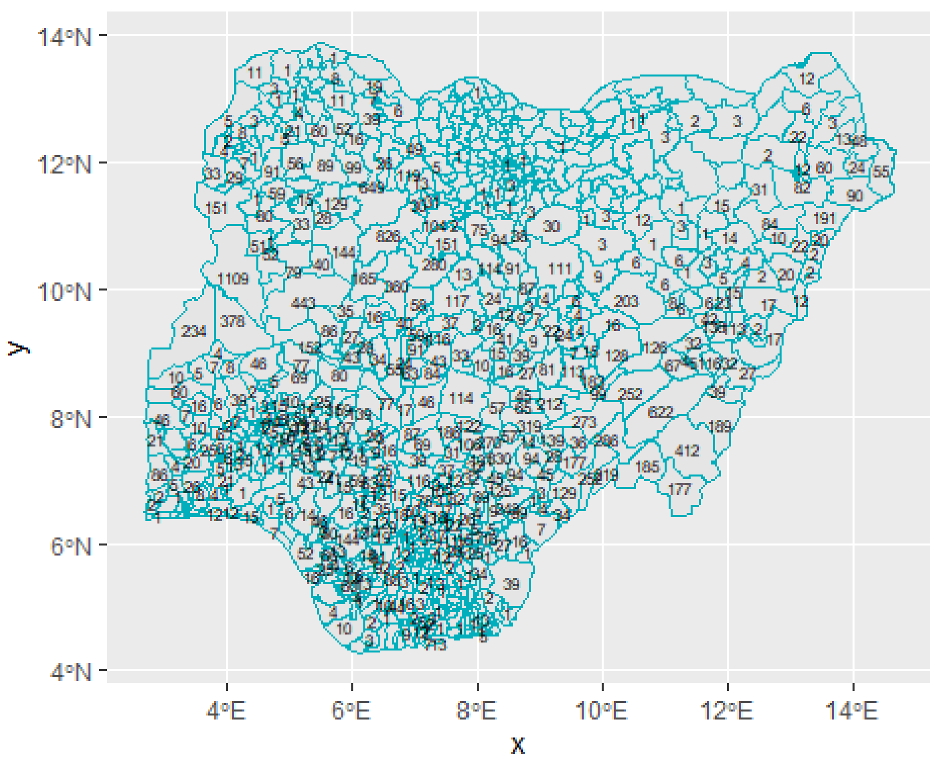


Figure 8: Number of fires per location

1.11 Investigate the spatial correlation between the numbers of fires in the different local government areas

```
library(spdep)

nb <- poly2nb(a, queen=TRUE)

nb[[1]]

##  [1]   4   5   6  10  11 403 410 553 563 571

lw <- nb2listw(nb, style = "W",  zero.policy=T)

a$n[is.na(a$n)]<- 0
moran.test(a$n,lw, zero.policy = T)

## Result
##  Moran I test under randomisation
##
## data:  a$n
## weights: lw  n reduced by no-neighbour observations
##
##
## Moran I statistic standard deviate = 20.957, p-value < 2.2e-16
## alternative hypothesis: greater
## sample estimates:
## Moran I statistic       Expectation          Variance
##      0.4334076495     -0.0012970169      0.0004302594
```

Conclusion: The moran test is positive. That means from the Moran test there is a correlation between the numbers of fires in the different local government areas. There are some clusters in the data.

Exercice II

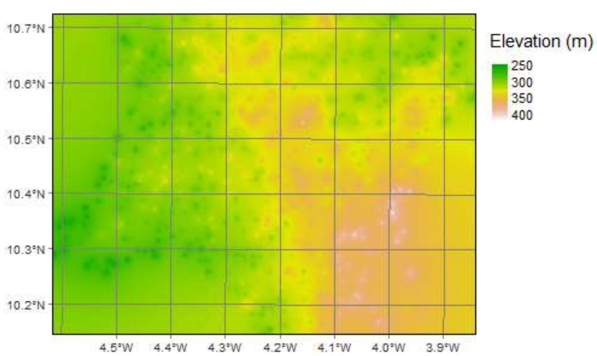
Import the elevation dataset

2.1 Choose a spatial interpolation (deterministic or stochastic spatial interpolation technique) method and create a continuous elevation raster in a spatial resolution (cell size) of 20 m

Using the new data sent by Dr. Sarah the cascade comoe data

```
#Import the elevation dataset  (cca_elevation.shp)
library(sf)

## Linking to GEOS 3.8.0, GDAL 3.0.4, PROJ 6.3.1

library(sp)
options(warn=0)
ccashp<-st_read("data/CascadeComoe_elevation_epsg32630.shx")
## Simple feature collection with 1002 features and 4 fields
## geometry type:  POINT
## dimension:      XY
## bbox:           xmin: 322835 ymin: 1121803 xmax: 407795 ymax: 1185853
## projected CRS:  WGS 84 / UTM zone 30N

cca<-st_as_sf(ccashp)
cca<-as(cca, "Spatial")
class(cca)

## [1] "SpatialPointsDataFrame"
## attr(,"package")
## [1] "sp"

library(sp)
library(gstat)
options(warn=0)
#idw method
x.range <- as.integer(c(322835,407795))
y.range <- as.integer(c(1121803,1185853))

#create a grid
grd <- expand.grid(x=seq(from=x.range[1], to=x.range[2], by=20), y=seq(from=y.range[1], to=y
.range[2], by=20))

## we convert the grid to SpatialPixel class
coordinates(grd) <- c("x","y")
gridded(grd)<-TRUE
proj4string(grd) <- CRS("+proj=utm +zone=30 +datum=WGS84 +units=m +no_defs")
```

```
library(gstat)
data.idw <- idw(Elev_m ~1, cca, grd)

[inverse distance weighted interpolation]

library(raster)
dataRaster<-raster("data/SRTM30mCascadeComoe_epsg32630.tif")
tm_shape(dataRaster) +
 tm_raster(title = "Elevation (m)",
          palette = terrain.colors(64), style = "cont")+
 tm_legend(outside = TRUE)

library(raster)
r <- raster(data.idw)
library(tmap)
tm_shape(r) +tm_raster(style = "cont", title = "Elevation (m)",
          palette = terrain.colors(64))+
 tm_legend(outside = TRUE)
```



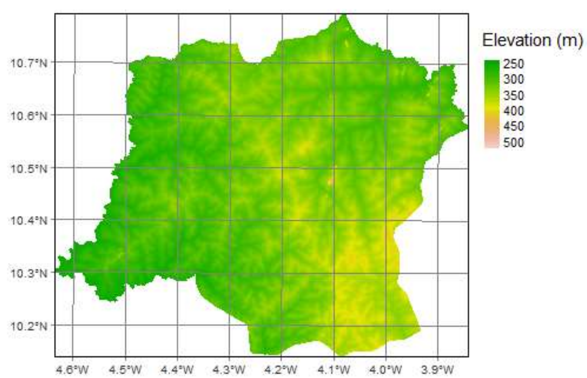


Figure 9: idw interpolation plot

Figure 10: The digital elevation model

There is a difference in terms of the highest level of the interpolation and the digital elevation model. We are observing that the interpolation is quietly looking like the digital elevation model. We select this model of interpolation due to the size of the memory of our computer. We were about to use the kriging model which take in account all the point for the interpolation.

Exercise III

3.1 Evaluate which datasets you will need to investigate the forest cover in Nigeria between 2000 and 2010.

We need four datasets:

- Tree canopy cover for year 2000
- Global forest cover loss 2000–2014
- Global forest cover gain 2000–2012
- Year of gross forest cover loss event

3.2 Create a plot showing the average forest cover for each protected area in the year 2000

Import the protected areas dataset (nigeria_pa.shp).

```
library(raster)
library(sf)
library(dplyr)
library(ggplot2)
nigeria<-st_read("data/nigeria_pa.shp")
## Reading layer `nigeria_pa' from data source `C:\Users\Simon Pierre KITEGI\Documents\Proj
ect-SDA\data\nigeria_pa.shp' using driver `ESRI Shapefile'
## Simple feature collection with 951 features and 33 fields
## geometry type:  MULTIPOLYGON
## dimension:      XY
## bbox:           xmin: 2.71982 ymin: 4.56824 xmax: 14.28409 ymax: 13.75343
## geographic CRS: WGS 84

#import the forest cover for the year 2000
f2000<-raster("data/DataExercise3/forestcover_2000.tif")
area<-mask(f2000,nigeria)
values<-extract(area,as(nigeria, "Spatial"),fun = mean)
new_data<-mutate(nigeria,values)
ggplot(new_data) +
  geom_sf(aes(fill = values)) +
  scale_fill_gradient(low = "#edf8e9", high = "#005a32")
```

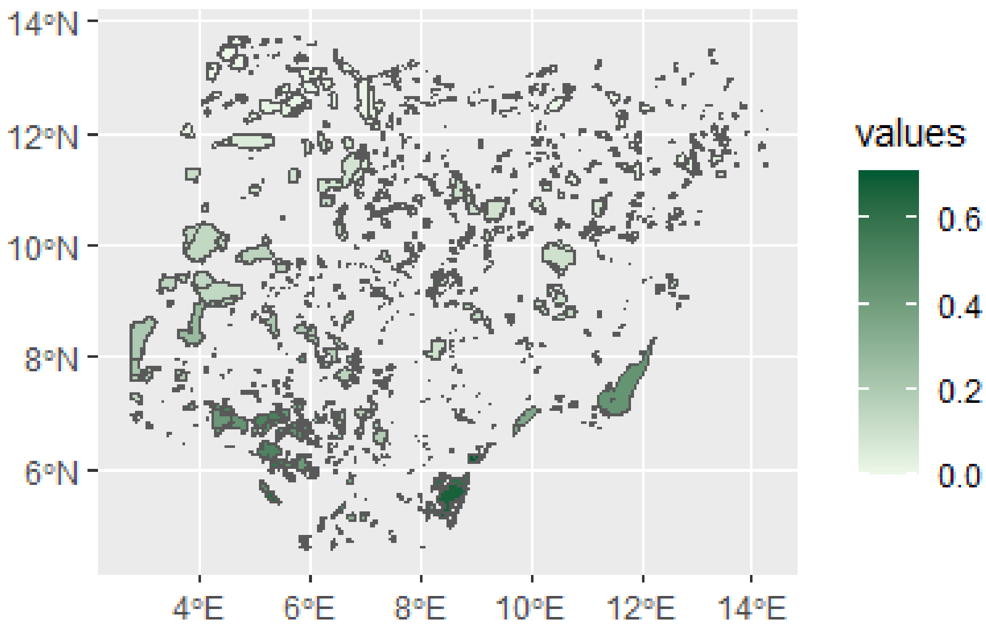


Figure 11: Forest cover in 2000

3.3 Calculate the forest cover in each protected area in Nigeria for each year between 2000 and 2010 (Note: Use the tree cover 2000 dataset as baseline)

```
#calculate the forest cover per year per protected area
values2000<-extract(area,as(nigeria, "Spatial"),fun = sum)
#import the forest lossyear data
flossyear<-raster("data/DataExercise3/forestcover_lossyear.tif")
areaflossyear<-mask(f2000,nigeria)
valuesflossyear<-extract(area,as(nigeria, "Spatial"),fun = mean)
#Computation of the forest cover during the period
datax<-mutate(nigeria,values2000,valuesflossyear)
datafcp<-datax%>%mutate(fcp=values2000/100,valuesflossyear)
#the map
ggplot(datafcp) +
  geom_sf(aes(fill = fcp)) +
  scale_fill_gradient(low = "#edf8e9", high = "#005a32")
```

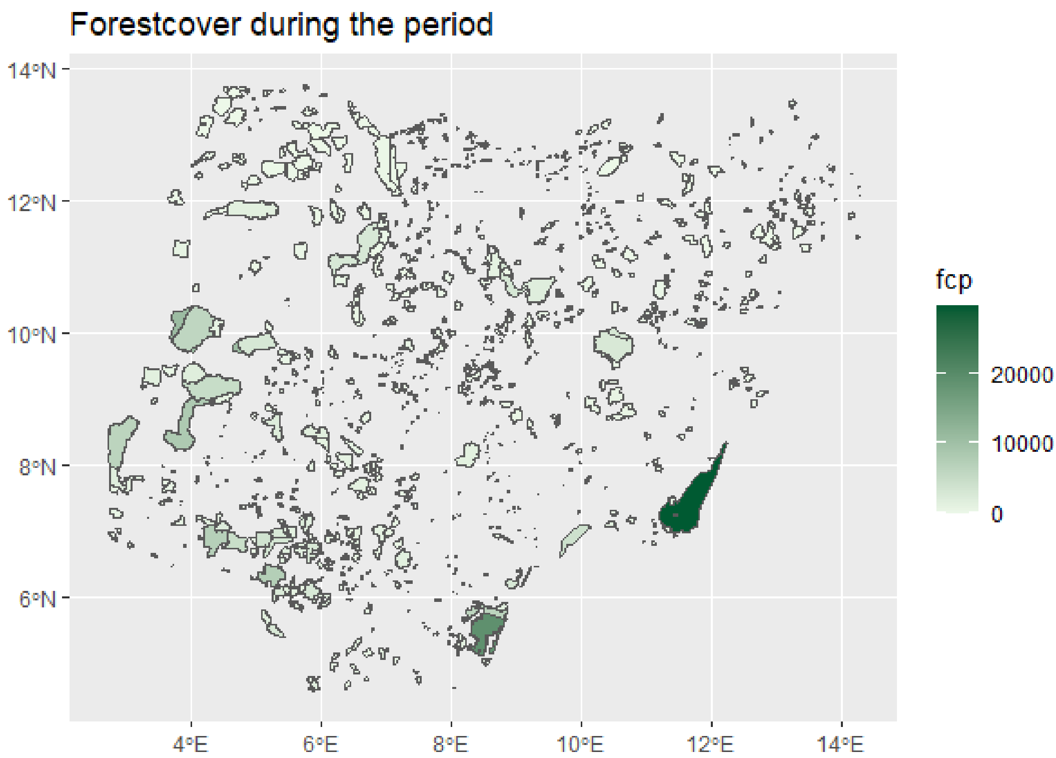


Figure 12: Forest cover during the period

3.4 Identify the protected area with the highest forest loss and create a map showing the forest loss/gain between 2000 and 2010 in this area.

```
library(raster)
#import the forest cover gain and loss
fgain<-raster("data/DataExercise3/forestcover_gain.tif")
areagain<-mask(fgain,nigeria)
vgain<-extract(areagain,as(nigeria, "Spatial"), fun = sum)
datagain<-mutate(nigeria,vgain)
floss<-raster("data/DataExercise3/forestcover_loss.tif")
areafloss<-mask(floss,nigeria)
vloss<-extract(areafloss,as(nigeria, "Spatial"), fun = sum)
dataloss<-mutate(nigeria,vloss)
#area with the most loss
Mostloss<-dataloss%>% select(NAME, vloss)%>%arrange(desc(vloss))%>%head(1)
NAME   vloss                     geometry
```

1 Okpara 2774.29 MULTIPOLYGON (((3.09027 8.7...

```
#the gain of the area
Mostlossgain<-datagain %>% select(NAME,vgain)%>%filter(NAME == "Okpara")  library(raster)
Mostgain<-datagain %>% select(NAME,vgain)%>%filter(NAME == "Okpara")
#the map
Name<-c("vgain", "vloss")
value<-c(0.88,2774.29)
okpara<-data.frame(Name,value)
library(ggplot2)
ggplot(okpara,aes(fill = Name, x= Name, y=value))+ geom_bar(stat = "identity", position = "stack", col = "#228B22")
```

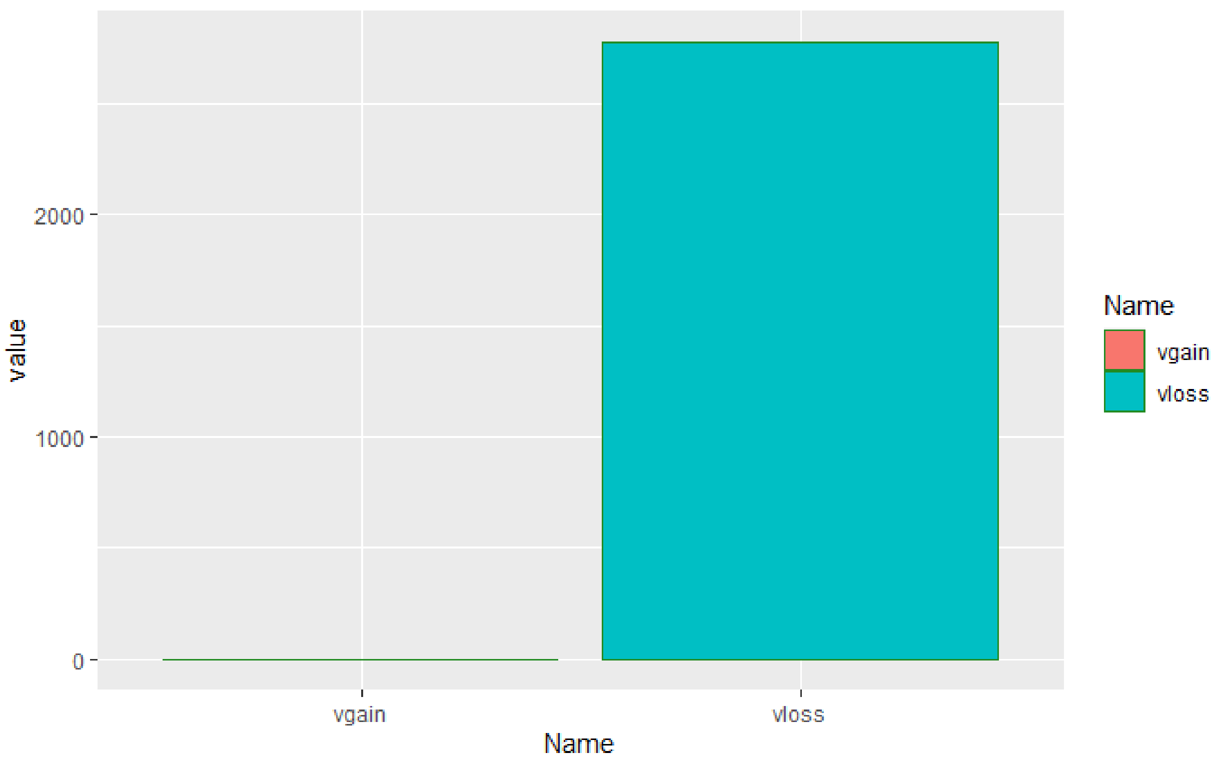


Figure 13: Forest cover gain and loss in Okpara