```
############################################################
#### Mapping the SOC stocks of Bangladesh
############################################################


# Table of Contents:


# 1. Import the data
# 2. Extract from HWSD database the data from Bangladesh
# 3. Sinusoidal projection
# 4. The SOC stocks of the topsoil and subsoil for the total MU
# 5. Tables of SOC per soil type
# 6. Tables of SOC per major soil type
# 7. The mean total SOC stocks per MU
# 8. Allocate to the spatial dataframe the SOC values
# 9. SOC per district
# 10.SOC per division
# 10. Map of SOC stocks
# 11. Graphical analysis

# Erase memory
rm(list=ls())


# Define your directory
setwd("C:/Users/Rashed/Desktop/temp/District&Divisions/")


# Install packages
require(maps)
require(mapdata)
require(maptools)
library(rgdal)
library(raster)
library(sp)
library(rgeos)
library(ggplot2)
library(tmap)
require(tmap)





#---------------------------------------------------------------
# 1. Import the data
# ---------------------------------------------------------------


# The Global HWSD database
HWSD <- read.csv("HWSD_DATA.csv")


# The glossary of the HWSD database
GLO <- read.csv("GLO.csv")


# Read the intersection map of soil and boundary for Bangladesh
BD <- readOGR(dsn = "C:/Users/Rashed/Desktop/temp/District&Divisions", layer = "HWSD_Bgd_District_B


#Plot the intersection map with the boundary districts and soil data
plot(BD)


#Check the names of the columns of the BD database
names(BD)
dim(BD)
```

```
#Rename the column 3: MU_GLOBAL
names(BD) [3] <- "MU_GLOBAL"


# Calculate the he total area (ha) of the mapping units
totar <- tapply(BD$Area_ha, BD$MU_GLOBAL, sum)
sum(totar) # 14,441,574


#Create a database with the total area of each mapping unit
totar <- data.frame(totar)


# The number of soil map units in the spatial database BD
length(unique(BD$MU_GLOBAL)) # There are 28 MU in the BD spatial database
unique(BD$MU_GLOBAL)


#--------------------------------------------------------------
# 2. Extract from HWSD database the data from Bangladesh
# --------------------------------------------------------------

# Extract from the global database HWSD the relevant mapping units for Bangladesh
SBD <- HWSD[HWSD$MU_GLOBAL %in% unique(BD$MU_GLOBAL) , ]


# Check the SBD database
names(SBD)
dim(SBD) # 94 observation and 57 variables
str(SBD)



#The number of the unique soil map units in the SBD and BD database
length(unique(SBD$MU_GLOBAL)) # 27 MU
length(unique(BD$MU_GLOBAL)) # 28 MU
# In the BD database there is one MU that is zero.


#-----------------------------------------------------------------
# 3. Sinusoidal projection
# -----------------------------------------------------------------

#Bangladesh Transverse Mercator (BTM)

proj_BTM<- "+proj=tmerc +lon_0=90e +x_0=500000 +y_0=-2000000 +ellps=evrst30"

BD_BTM <-spTransform(BD, CRS=proj_BTM)
plot(BD_BTM)


#-----------------------------------------------------------------------
# 4. The SOC stocks of the topsoil and subsoil for the total MU
#-----------------------------------------------------------------------

#SOC (kg m-2) = V*(1-(GRAVEL(Gravel Content %vol))*REF_BULK_DENSITY (Reference Bulk Density kg/dm3 )
#OC (Organic Carbon % weight)


### Calculate the SOC stock of the topsoil (0-30 cm)

# Replace NA with zero
SBD$T_OC[is.na(SBD$T_OC)]<-0
SBD$T_REF_BULK_DENSITY[is.na(SBD$T_REF_BULK_DENSITY)]<-0
SBD$T_GRAVEL[is.na(SBD$T_GRAVEL)]<-0

SBD$T_SOC <- (0.3*(1-SBD$T_GRAVEL/100)*SBD$T_REF_BULK_DENSITY*1000*
```

```
                SBD$T_OC/100)

mean(SBD$T_OC)# 2.95617 %
mean(SBD$T_REF_BULK_DENSITY)# 1.370638 kg/dm3
mean(SBD$T_GRAVEL) # 7.542553 %
mean(SBD$T_SOC)# 10.63589 kg/m2
sd(SBD$T_SOC) # 24.73698
#The mean SOC stock of the topsoil for the total MU of Bangladesh is 10.63589 kg/m2.


##Calculate the SOC stocks (kg/m2) of the subsoil (30-100 cm)

# Replace NA values with zero
SBD$S_OC[is.na(SBD$S_OC)]<-0
SBD$S_REF_BULK_DENSITY[is.na(SBD$S_REF_BULK_DENSITY)]<-0
SBD$S_GRAVEL[is.na(SBD$S_GRAVEL)]<-0

SBD$S_SOC <- (0.7*(1-SBD$S_GRAVEL/100)*SBD$S_REF_BULK_DENSITY*1000*
                SBD$S_OC/100)

mean(SBD$S_OC)# 2.274362 %
mean(SBD$S_REF_BULK_DENSITY)# 1.334681 kg/dm3
mean(SBD$S_GRAVEL) # 8.212766 %
mean(SBD$S_SOC) # 20.59767 kg/m2.
sd(SBD$S_SOC) # 68.93182


## Calculate the total SOC of the topsoil and subsoil (0-100 cm)
SBD$Tot_SOC <- SBD$T_SOC + SBD$S_SOC
mean(SBD$Tot_SOC) # 31.23356 kg/m2
sd(SBD$Tot_SOC) # 93.58044


#------------------------------------------------------------------------
# 5.  Tables of SOC per soil type
#------------------------------------------------------------------------

# According to HWSD the soil classification system that is used in Bangladesh is
# FAO - 74 classification system

SBD$SU_SYM74 # Soil unit symbol (FAO-74)
SBD$SU_CODE74 # Soil unit code (FAO-74)


# Load the file with the soil codes, names and symbols according to the FAO-74 classification system
D_SYMBOL74 <-read.csv("D_SYMBOL74.csv")


#Check the names of the columns of the "D_SYMBOL74" file
names(D_SYMBOL74)


# In the file D_SYMBOL74 we rename the columns
# 1,2,3, to make the link with the SBD

names(D_SYMBOL74) [1] <- "SU_CODE74"
names(D_SYMBOL74) [2] <- "FAO_74_Name"
names(D_SYMBOL74) [3] <-  "SU_SYM74"
names(D_SYMBOL74)
names(SBD)



# Load the file with the  FAO-74 soil classification and the international soil
# classification system (WRB). This file will help us to pass from  FAO -74 classification
# to WRB classification.We will identify which soil types based on the FAO -74 classification
# correspond to the WRB classification
clasf<-read.csv("WRB_FAO74.csv")
```

```
#Create a new column in the SBD database to add the soil names that
#correspond to each soil code SU_CODE74
SBD$FAO_74_Name <- D_SYMBOL74$FAO_74_Name[
  match(SBD$SU_CODE74,D_SYMBOL74$SU_CODE74)]


#Create a new column in the SBD database to add the names of the
#major soil types that correspond to each soil code SU_CODE74
SBD$Major_74 <- clasf$Major_74[
  match(SBD$SU_CODE74, clasf$SU_CODE74)]



# Create a table of SOC stocks per soil type for the topsoil
SBD_74 <- data.frame(T_SOC= SBD$T_SOC, FAO_74_Name= SBD$FAO_74_Name)

mean_T_SOC_74<- tapply(SBD_74$T_SOC, SBD_74$FAO_74_Name, mean)
sd_T_SOC_74<- tapply(SBD_74$T_SOC, SBD_74$FAO_74_Name, sd)
min_T_SOC_74<- tapply(SBD_74$T_SOC, SBD_74$FAO_74_Name, min)
max_T_SOC_74<- tapply(SBD_74$T_SOC, SBD_74$FAO_74_Name, max)
n_T_SOC_74 <- tapply(SBD_74$T_SOC, SBD_74$FAO_74_Name, length)

table_T_SOC_74 <- data.frame(mean=mean_T_SOC_74, sd=sd_T_SOC_74,
                             n=n_T_SOC_74, min=min_T_SOC_74,
                             max=max_T_SOC_74)


# Create a subset of the table_T_SOC_74 to keep the rows where n is not NA
subtab <- subset(table_T_SOC_74, !is.na(n))


# Create a csv table of the above table
write.csv(subtab, file = "table_T_SOC_74.csv" )


# Create a table of SOC stocks per soil type for the subsoil
SBD_74 <- data.frame(S_SOC= SBD$S_SOC, FAO_74_Name= SBD$FAO_74_Name)

mean_S_SOC_74<- tapply(SBD_74$S_SOC, SBD_74$FAO_74_Name, mean)
sd_S_SOC_74<- tapply(SBD_74$S_SOC, SBD_74$FAO_74_Name, sd)
min_S_SOC_74<- tapply(SBD_74$S_SOC, SBD_74$FAO_74_Name, min)
max_S_SOC_74<- tapply(SBD_74$S_SOC, SBD_74$FAO_74_Name, max)
n_S_SOC_74 <- tapply(SBD_74$S_SOC, SBD_74$FAO_74_Name, length)

table_S_SOC_74 <- data.frame(mean=mean_S_SOC_74, sd=sd_S_SOC_74,
                             n=n_S_SOC_74, min=min_S_SOC_74,
                             max=max_S_SOC_74)

# Create a subset of the table_S_SOC_74 to keep the rows where n is not NA
subtab2 <- subset(table_S_SOC_74, !is.na(n))


# Create a csv table the above table
write.csv(subtab2, file = "table_S_SOC_74.csv" )


#-----------------------------------------------------------------------
# 6. Tables of SOC per major soil type
#-----------------------------------------------------------------------
names(SBD)


## Where FAO_74_Name==Calcaric Gleysols, and Major_74== NA
# Replace NA with Gleysols
SBD[SBD$FAO_74_Name == "Calcaric Gleysols" & is.na(SBD$Major_74), "Major_74"] <- "Gleysols"


## Where FAO_74_Name==Humic Cambisols, and Major_74== NA
```

```
## Replace NA with Cambisols
SBD[SBD$FAO_74_Name == "Humic Cambisols" & is.na(SBD$Major_74), "Major_74"] <- "Cambisols"


## Where FAO_74_Name==Eutric Histosols, and Major_74== NA
## Replace NA with Histosols
SBD[SBD$FAO_74_Name == "Eutric Histosols" & is.na(SBD$Major_74), "Major_74"] <- "Histosols"


## Where FAO_74_Name==Humic Acrisols, and Major_74== NA
## Replace NA with Acrisols
SBD[SBD$FAO_74_Name == "Humic Acrisols" & is.na(SBD$Major_74), "Major_74"] <- "Acrisols"


## Where FAO_74_Name==Gleyic Acrisols, and Major_74== NA
## Replace NA with Acrisols
SBD[SBD$FAO_74_Name == "Gleyic Acrisols" & is.na(SBD$Major_74), "Major_74"] <- "Acrisols"


## Where FAO_74_Name==Orthic Luvisols, and Major_74== NA
## Replace NA with Luvisols
SBD[SBD$FAO_74_Name == "Orthic Luvisols" & is.na(SBD$Major_74), "Major_74"] <- "Luvisols"


## Where FAO_74_Name==Histosols, and Major_74== NA
## Replace NA with Histosols
SBD[SBD$FAO_74_Name == "Histosols" & is.na(SBD$Major_74), "Major_74"] <- "Histosols"


## Where FAO_74_Name==Histosols, and Major_74== NA
## Replace NA with Histosols
SBD[SBD$FAO_74_Name == "Histosols" & is.na(SBD$Major_74), "Major_74"] <- "Histosols"


## Where FAO_74_Name==Eutric Planosols, and Major_74== NA
## Replace NA with Planosols
SBD[SBD$FAO_74_Name == "Eutric Planosols" & is.na(SBD$Major_74), "Major_74"] <- "Planosols"



# Create a table of SOC stocks per major soil type for the topsoil
SBD_74 <- data.frame(T_SOC= SBD$T_SOC, Major_74= SBD$Major_74)

mean_T_SOC_M <- tapply(SBD_74$T_SOC, SBD_74$Major_74, mean)
sd_T_SOC_M<- tapply(SBD_74$T_SOC, SBD_74$Major_74, sd)
n_T_SOC_M<- tapply(SBD_74$T_SOC, SBD_74$Major_74, length)
min_T_SOC_M<- tapply(SBD_74$T_SOC, SBD_74$Major_74, min)
max_T_SOC_M<- tapply(SBD_74$T_SOC, SBD_74$Major_74, max)

table_T_SOC_M<- data.frame(mean=mean_T_SOC_M, sd=sd_T_SOC_M,
                           n=n_T_SOC_M, min=min_T_SOC_M,
                           max=max_T_SOC_M)


# Create a subset of the table_T_SOC_M to remove the rows when n=NA
subtab3 <- subset(table_T_SOC_M, !is.na(n))


# Create a csv table of the above table
write.csv(subtab3, file = "table_T_SOC_M.csv" )


## Create a table of SOC stocks per major soil type for the subsoil
SBD_74 <- data.frame(S_SOC= SBD$S_SOC, Major_74= SBD$Major_74)

mean_S_SOC_M <- tapply(SBD_74$S_SOC, SBD_74$Major_74, mean)
sd_S_SOC_M<- tapply(SBD_74$S_SOC, SBD_74$Major_74, sd)
n_S_SOC_M<- tapply(SBD_74$S_SOC, SBD_74$Major_74, length)
min_S_SOC_M<- tapply(SBD_74$S_SOC, SBD_74$Major_74, min)
```

```
max_S_SOC_M<- tapply(SBD_74$S_SOC, SBD_74$Major_74, max)

table_S_SOC_M<- data.frame(mean=mean_S_SOC_M, sd=sd_S_SOC_M,
                           n=n_S_SOC_M, min=min_S_SOC_M,
                           max=max_S_SOC_M)

# Create a subset of the table_T_SOC_M to remove the rows when n=NA
subtab4 <- subset(table_S_SOC_M, !is.na(n))


# Create a csv table of the above table
write.csv(subtab4, file = "table_S_SOC_M.csv" )




## Create a table of SOC per major soil type for the topsoil and subsoil
SBD_74 <- data.frame(Tot_SOC= SBD$T_SOC+SBD$S_SOC, Major_74= SBD$Major_74)

mean_Tot_SOC_M <- tapply(SBD_74$Tot_SOC, SBD_74$Major_74, mean)
sd_Tot_SOC_M <- tapply(SBD_74$Tot_SOC, SBD_74$Major_74, sd)
n_Tot_SOC_M <- tapply(SBD_74$Tot_SOC, SBD_74$Major_74, length)
min_Tot_SOC_M <- tapply(SBD_74$Tot_SOC, SBD_74$Major_74, min)
max_Tot_SOC_M <- tapply(SBD_74$Tot_SOC, SBD_74$Major_74, max)

table_Tot_SOC_M <- data.frame(mean=mean_Tot_SOC_M, sd=sd_Tot_SOC_M,
                              n=n_Tot_SOC_M, min=min_Tot_SOC_M,
                              max=max_Tot_SOC_M)


# Create a subset of the table_Tot_SOC_74 to remove the rows when n=NA
subtab7 <- subset(table_Tot_SOC_M, !is.na(n))


# Create a csv table of SOC stocks for the Tot_SOC
write.csv(subtab7, file = "table_Tot_SOC_M.csv" )


#-----------------------------------------------------------------------
# 7.The mean total SOC stocks per MU
#-----------------------------------------------------------------------


# we calcualte the total SOC per soil types per MU
#considering that in one MU there is more than one soil type
# For each soil unit of the map unit
SBD$T_SOC2 <- SBD$T_SOC*SBD$SHARE/100
SBD$S_SOC2 <- SBD$S_SOC*SBD$SHARE/100


# Calculate the sum of SOC stocks (kg/m2) for each MU
# Topsoil
T_SOC <- tapply(SBD$T_SOC2, SBD$MU_GLOBAL, sum)
T_SOC2 <- data.frame(MU_GLOBAL=names(T_SOC), T_SOC=T_SOC)
dim(T_SOC2)


# Subsoil
S_SOC <- tapply(SBD$S_SOC2, SBD$MU_GLOBAL, sum)
S_SOC2 <- data.frame(MU_GLOBAL=names(S_SOC), S_SOC=S_SOC)
dim(S_SOC2)

# Total
# Replace NA with zero
SBD$S_SOC2[is.na(SBD$S_SOC2)]<-0
SBD$T_SOC2[is.na(SBD$T_SOC2)]<-0
SBD$Tot_SOC2<- SBD$T_SOC2 + SBD$S_SOC2

Tot_SOC <- tapply(SBD$Tot_SOC2, SBD$MU_GLOBAL, sum)
```

```
Tot_SOC2 <- data.frame(MU_GLOBAL=names(Tot_SOC), Tot_SOC=Tot_SOC)
dim(Tot_SOC2$Tot_SOC)

MU <- data.frame(MU_GLOBAL=names(Tot_SOC), T_SOC=T_SOC, S_SOC=S_SOC, Tot_SOC=Tot_SOC)

write.csv(MU, file="MU.csv")


#------------------------------------------------------------------------
# 8. Allocate to the spatial dataframe the SOC values
#------------------------------------------------------------------------

# We link the spatial dataframe (BD_BTM) and the data.frames created above (T_SOC2, S_SOC2, Tot_SOC2
#with the T_SOC, S_SOC, and Tot_SOC values

BD_BTM$T_SOC <- T_SOC2$T_SOC[
  match(BD_BTM$MU_GLOBAL, T_SOC2$MU_GLOBAL)]

BD_BTM$S_SOC <- S_SOC2$S_SOC[
  match(BD_BTM$MU_GLOBAL, S_SOC2$MU_GLOBAL)]

BD_BTM$Tot_SOC <- Tot_SOC2$Tot_SOC[
  match(BD_BTM$MU_GLOBAL, Tot_SOC2$MU_GLOBAL)]

# quick look on the new values in the spatial data frame
names(BD_BTM)
head(BD_BTM)

# Several polygons do not have MU values
qc<-subset(BD_BTM,BD_BTM$Tot_SOC=="0")
length(unique(qc$MU_GLOBAL))

# WHEN MU = 6997, there is not SOC values (NA)
# MU=6997 is present in the SBD database and in the special database BD.
# However it is classiffied as a nonsoil unit (ISSOIL=0).


# Calculation of the land area for each MU

BD_BTM$S_SOC[is.na(BD_BTM$S_SOC)]<-0
BD_BTM$T_SOC[is.na(BD_BTM$T_SOC)]<-0
BD_BTM$Tot_SOC[is.na(BD_BTM$Tot_SOC)]<-0


# We allocate to the spatial data frame the soil names and major soil names
BD_BTM$FAO_74_Name <- SBD$FAO_74_Name[
  match(BD_BTM$MU_GLOBAL, SBD$MU_GLOBAL)]

BD_BTM$Major_74 <- SBD$Major_74[
  match(BD_BTM$MU_GLOBAL, SBD$MU_GLOBAL)]



#-------------------------------------------------------------------
# 9. SOC per district
#-------------------------------------------------------------------

# BD_BTM is the database
names(BD_BTM)
length(unique(BD_BTM$District)) # 64 Districts

# Creat a table with the total SOC for the topsoil for the different districts
SBD_DI <- data.frame(DIS= BD_BTM$District, T_SOC= BD_BTM$T_SOC,
                     Major_74= BD_BTM$Major_74, area=BD_BTM$Area_ha)

mean_T_SOC_D<- tapply(BD_BTM$T_SOC, BD_BTM$District, mean)
sd_T_SOC_D<- tapply(BD_BTM$T_SOC, BD_BTM$District, sd)
min_T_SOC_D<- tapply(BD_BTM$T_SOC, BD_BTM$District, min)
max_T_SOC_D<- tapply(BD_BTM$T_SOC, BD_BTM$District, max)
```

```
n_T_SOC_D<- tapply(BD_BTM$T_SOC, BD_BTM$District, length)
area_D <- tapply(BD_BTM$Area_ha, BD_BTM$District, sum)


table_T_SOC_D <- data.frame(mean=mean_T_SOC_D, sd=sd_T_SOC_D,
                            min=min_T_SOC_D,  max=max_T_SOC_D,
                            n=n_T_SOC_D, area_D=area_D)

write.csv(table_T_SOC_D, "table_T_SOC_D.csv")


# A table of SOC for the subsoil per district
SBD_DI <- data.frame(DIS= BD_BTM$District, S_SOC= BD_BTM$S_SOC,
                     Major_74= BD_BTM$Major_74, area=BD_BTM$Area_ha)

mean_S_SOC_D<- tapply(BD_BTM$S_SOC, BD_BTM$District, mean)
sd_S_SOC_D<- tapply(BD_BTM$S_SOC, BD_BTM$District, sd)
min_S_SOC_D<- tapply(BD_BTM$S_SOC, BD_BTM$District, min)
max_S_SOC_D<- tapply(BD_BTM$S_SOC, BD_BTM$District, max)
n_S_SOC_D<- tapply(BD_BTM$S_SOC, BD_BTM$District, length)
area_D <- tapply(BD_BTM$Area_ha, BD_BTM$District, sum)


table_S_SOC_D <- data.frame(mean=mean_S_SOC_D, sd=sd_S_SOC_D,
                            min=min_S_SOC_D,  max=max_S_SOC_D,
                            n=n_S_SOC_D, area_D=area_D)


write.csv(table_S_SOC_D, "table_S_SOC_D.csv")



# A table of SOC for the topsoil and subsoil per distrcit
SBD_DI <- data.frame(DIS= BD_BTM$District, Tot_SOC= BD_BTM$Tot_SOC,
                     Major_74= BD_BTM$Major_74, area=BD_BTM$Area_ha)

mean_Tot_SOC_D<- tapply(BD_BTM$Tot_SOC, BD_BTM$District, mean)
sd_Tot_SOC_D<- tapply(BD_BTM$Tot_SOC, BD_BTM$District, sd)
min_Tot_SOC_D<- tapply(BD_BTM$Tot_SOC, BD_BTM$District, min)
max_Tot_SOC_D<- tapply(BD_BTM$Tot_SOC, BD_BTM$District, max)
n_Tot_SOC_D<- tapply(BD_BTM$Tot_SOC, BD_BTM$District, length)
area_D <- tapply(BD_BTM$Area_ha, BD_BTM$District, sum)


table_Tot_SOC_D <- data.frame(mean=mean_Tot_SOC_D, sd=sd_Tot_SOC_D,
                              min=min_Tot_SOC_D,  max=max_Tot_SOC_D,
                              n=n_Tot_SOC_D, area_D=area_D)


write.csv(table_Tot_SOC_D, "table_Tot_SOC_D.csv")



#Find the smallest and biggest mean from table_Tot_SOC_D
min(table_Tot_SOC_D$mean, na.rm = TRUE) #0.5485549 kgC/m2
max(table_Tot_SOC_D$mean, na.rm = TRUE) #198.6923 kgC/m2




#---------------------------------------------------------------------------
# 10. SOC stocks per division of Bangladesh
#---------------------------------------------------------------------------


# BD_BTM is the database
names(BD_BTM)
length(unique(BD_BTM$Division)) # 8 Divisions
```

```
# Creat a table with the SOC for the topsoil for the different Divisions
SBD_DIV <- data.frame(DIV= BD_BTM$Division, T_SOC= BD_BTM$T_SOC,
                      Major_74= BD_BTM$Major_74, area=BD_BTM$Area_ha)

mean_T_SOC_DIV <- tapply(BD_BTM$T_SOC, BD_BTM$Division, mean)
sd_T_SOC_DIV <- tapply(BD_BTM$T_SOC, BD_BTM$Division, sd)
min_T_SOC_DIV <- tapply(BD_BTM$T_SOC, BD_BTM$Division, min)
max_T_SOC_DIV <- tapply(BD_BTM$T_SOC, BD_BTM$Division, max)
n_T_SOC_DIV <- tapply(BD_BTM$T_SOC, BD_BTM$Division, length)
area_DIV <- tapply(BD_BTM$Area_ha, BD_BTM$Division, sum)


table_T_SOC_DIV <- data.frame(mean=mean_T_SOC_DIV, sd=sd_T_SOC_DIV,
                             min=min_T_SOC_DIV,  max=max_T_SOC_DIV,
                             n=n_T_SOC_DIV, area_DIV=area_DIV)

write.csv(table_T_SOC_DIV, "table_T_SOC_DIV.csv")



# Creat a table with the SOC for the subsoil for the different Divisions
SBD_DIV <- data.frame(DIV= BD_BTM$Division, S_SOC= BD_BTM$S_SOC,
                       Major_74= BD_BTM$Major_74, area=BD_BTM$Area_ha)

mean_S_SOC_DIV <- tapply(BD_BTM$S_SOC, BD_BTM$Division, mean)
sd_S_SOC_DIV <- tapply(BD_BTM$S_SOC, BD_BTM$Division, sd)
min_S_SOC_DIV <- tapply(BD_BTM$S_SOC, BD_BTM$Division, min)
max_S_SOC_DIV <- tapply(BD_BTM$S_SOC, BD_BTM$Division, max)
n_S_SOC_DIV <- tapply(BD_BTM$S_SOC, BD_BTM$Division, length)
area_DIV <- tapply(BD_BTM$Area_ha, BD_BTM$Division, sum)


table_S_SOC_DIV <- data.frame(mean=mean_S_SOC_DIV, sd=sd_S_SOC_DIV,
                               min=min_S_SOC_DIV,  max=max_S_SOC_DIV,
                               n=n_S_SOC_DIV, area_DIV=area_DIV)

write.csv(table_S_SOC_DIV, "table_S_SOC_DIV.csv")


# A table of SOC for the topsoil and subsoil per Division
SBD_DIV <- data.frame(DIIV= BD_BTM$Division, Tot_SOC= BD_BTM$Tot_SOC,
                      Major_74= BD_BTM$Major_74, area=BD_BTM$Area_ha)

mean_Tot_SOC_DIV <- tapply(BD_BTM$Tot_SOC, BD_BTM$Division, mean)
sd_Tot_SOC_DIV <- tapply(BD_BTM$Tot_SOC, BD_BTM$Division, sd)
min_Tot_SOC_DIV <- tapply(BD_BTM$Tot_SOC, BD_BTM$Division, min)
max_Tot_SOC_DIV <- tapply(BD_BTM$Tot_SOC, BD_BTM$Division, max)
n_Tot_SOC_DIV <- tapply(BD_BTM$Tot_SOC, BD_BTM$Division, length)
area_DIV <- tapply(BD_BTM$Area_ha, BD_BTM$Division, sum)


table_Tot_SOC_DIV <- data.frame(mean=mean_Tot_SOC_DIV, sd=sd_Tot_SOC_DIV,
                               min=min_Tot_SOC_DIV,  max=max_Tot_SOC_DIV,
                               n=n_Tot_SOC_DIV, area_D=area_DIV)


write.csv(table_Tot_SOC_DIV, "table_Tot_SOC_DIV.csv")



#-------------------------------------------------------------------------
# 11. Map of SOC stocks
#-------------------------------------------------------------------------

# Map representing the different soil types
proj_BTM_m<- "+proj=tmerc +lon_0=90e +x_0=500000 +y_0=-2000000 +ellps=evrst30 +units=m"
tm_shape(BD_BTM, projection = proj_BTM_m) +
```

```
  tm_fill("Major_74", style="kmeans",
          labels=levels(BD_BTM$Major_74),
          title="Soil types")+
  tm_borders() +
  tm_layout(legend.outside = TRUE)+
  tm_scale_bar(size = 0.3, position = c("left", "bottom"))+
  tm_compass(size= 0.3, position = c("right", "top"))



## Map representing the topsoil organic carbon stock


png("temp.png",width = 12,height = 16,units = "in",res = 300)


tm_shape(BD_BTM, projection = proj_BTM_m) +
  tm_fill("T_SOC", style="kmeans",
          labels=levels(BD_BTM$T_SOC),
          title="SOC 0-30cm")+
  tm_borders() +
  tm_layout(legend.outside = TRUE)+
  tm_scale_bar(size = 0.3, position = c("left", "bottom"))+
  tm_compass(size= 0.3, position = c("right", "top"))

dev.off()

# Map representing the SOC of the subsoil
tm_shape(BD_BTM) +
  tm_fill("S_SOC", style="kmeans",
          labels=levels(BD_BTM$S_SOC),
          title="SOC 30-100cm")+
  tm_borders()+
  tm_layout(legend.outside = TRUE)+
  tm_scale_bar(size = 0.3, position = c("left", "bottom"))+
  tm_compass(size= 0.3, position = c("right", "top"))


# Map representing the SOC for the 0-100 cm soil layer
tm_shape(BD_BTM) +
  tm_fill("Tot_SOC", style="kmeans",
          labels=levels(BD_BTM$Tot_SOC),
          title="SOC 0-100cm")+
  tm_borders()+
  tm_layout(legend.outside = TRUE)+
  tm_scale_bar(size = 0.3, position = c("left", "bottom"))+
  tm_compass(size= 0.3, position = c("right", "top"))



#------------------------------------------------------------------------
# 11. Graphic analysis
#------------------------------------------------------------------------

names(BD_BTM)
dim(BD_BTM)

BD_BTM2<-data.frame(BD_BTM)

summary(BD_BTM2)
par(mfrow=c(1,1))
names(BD_BTM2)


# Make a  histogram of the T_SOC per soil type
hist(as.numeric(BD_BTM2$T_SOC), freq=FALSE, breaks=100)
hist(as.numeric(BD_BTM2$T_SOC), freq=TRUE, breaks=100, col="blue")
```

```
# Make a  histogram of the S_SOC per soil type
hist(as.numeric(BD_BTM2$S_SOC), freq=FALSE, breaks=100)
hist(as.numeric(BD_BTM2$S_SOC), freq=TRUE, breaks=100, col="blue")


# Make a  histogram of the Tot_SOC per soil type
hist(as.numeric(BD_BTM2$Tot_SOC), freq=FALSE, breaks=100)
hist(as.numeric(BD_BTM2$Tot_SOC), freq=TRUE, breaks=100, col="blue")


### Identify values for your intervals
interv <- seq(0, 250, 50)
hist <- hist(BD_BTM2$T_SOC, breaks = interv, plot = FALSE)


## The graphical relationship between T_SOC and S_SOC for the different soil types
ggplot(BD_BTM2)+
  geom_point(aes(x=T_SOC,y=S_SOC,colour=Major_74))+
  scale_y_log10()+
  facet_wrap(~Major_74)+
  ggtitle("Topsoil versus subsoil")



## The graphical relationship between Tot_SOC and area for the different soil types
ggplot(BD_BTM2)+
  geom_point(aes(x=Tot_SOC,y=area, colour=Major_74))+
  facet_wrap(~Major_74)+
  ggtitle("Total carbon per area")


## Frenquency of the T_SOC values in MU
ggplot(BD_BTM2, aes(x=T_SOC, colour=Major_74)) + geom_density()


# Frenquency of the S_SOC values in MU
ggplot(BD_BTM2, aes(x=S_SOC, colour=Major_74)) + geom_density()


# Frenquency of the Tot_SOC values in MU
ggplot(BD_BTM2, aes(x=Tot_SOC, colour=Major_74)) + geom_density()


# Create a new column in the database BD_BTM2 to add
# the acronyms of the major soil types. Use 2 letters for the
# abbreviation

BD_BTM2$MJ74 <- abbreviate(BD_BTM2$Major_74, 2)
# we need to create a new table for this graph
# [ID], [MU], [MJ74], [SOC], [LAYER]


#Create a database for the subsoil with ID, MU, MJ74, S_SOC_, soil layer
df1<-data.frame(MU=BD_BTM2$MU_GLOBAL, MJ74= BD_BTM2$MJ74, SOC= BD_BTM2$S_SOC, LAYER="S")

#Add another column with the id
id <- rownames(df1)


# Combine the id wih the df1 database
df1 <- cbind(id=id, df1)


#Create a database for the topsoil with ID, MU, MJ74, T_SOC, soil layer
df2<-data.frame(MU=BD_BTM2$MU_GLOBAL, MJ74= BD_BTM2$MJ74, SOC= BD_BTM2$T_SOC, LAYER="T")


#Add another column with the id
id <- rownames(df2)
```

```
# Combine the id wih the df2 database
df2 <- cbind(id=id, df2)


# Combine by rows the df1 and df2 database created above
df3 <- rbind(df1, df2)


#Remove rows when MJ74=NA
df4 <- df3[complete.cases(df3),]


## Make a graph with the SOC for the topsoil and subsoil per soil type
ggplot(df4,aes(MJ74,SOC,fill=LAYER))+
  geom_bar(stat="identity",position='dodge')


## Prepare a boxplot to look at the distribution of SOC values for the major soil types
# for the subsoil df1, topsoil df2
par(mfrow=c(1,2))

ggplot(df1, aes(x=MJ74, y=SOC)) + geom_boxplot() +
  stat_summary(fun.y=mean, geom="point", shape=5, size=4)

ggplot(df2, aes(x=MJ74, y=SOC)) + geom_boxplot() +
  stat_summary(fun.y=mean, geom="point", shape=5, size=4)


# Relationship between T_clay and T_SOC. The data are in SBD
p <- ggplot(SBD, aes(x=T_CLAY, y=T_SOC, pch=Major_74, colour=Major_74))
p <- p + geom_point(cex=3)
p <- p + scale_shape_manual(values = 0:length(unique(SBD$Major_74)))
p


# Relationship between S_clay and S_SOC
p <- ggplot(SBD, aes(x=S_CLAY, y=S_SOC, pch=Major_74, colour=Major_74))
p <- p + geom_point(cex=3)
p <- p + scale_shape_manual(values = 0:length(unique(SBD$Major_74)))
p


# Relationship between T_bulk density and T_SOC
p <- ggplot(SBD, aes(x=T_BULK_DENSITY, y=T_SOC, pch=Major_74, colour=Major_74))
p <- p + geom_point(cex=3)
p <- p + scale_shape_manual(values = 0:length(unique(SBD$Major_74)))
p


# Relationship between S_bulk density and S_SOC
p <- ggplot(SBD, aes(x=S_BULK_DENSITY, y=S_SOC, pch=Major_74, colour=Major_74))
p <- p + geom_point(cex=3)
p <- p + scale_shape_manual(values = 0:length(unique(SBD$Major_74)))
p

###Figure presenting the Total SOC and the different Districts
# BD_BTM is the database
names(BD_BTM)
BD_BTM2<-data.frame(BD_BTM)


#Create a database for the topsoil with ID, MU, MJ74, T_SOC, soil layer
#df2<-data.frame(MU=BD_BTM2$MU_GLOBAL, MJ74= BD_BTM2$MJ74, SOC= BD_BTM2$T_SOC, LAYER="T")

#Create a database for the total SOC with
#ID, MU, Districts, total SOC

DF<- data.frame(MU=BD_BTM2$MU_GLOBAL, DIS=BD_BTM2$District,
```

```
                T_SOC=BD_BTM2$T_SOC)


#Add another column with the id
ID<- rownames(DF)


# Combine the ID wih the DF database
DF2 <- cbind(ID=ID, DF)


ggplot(DF2, aes(x=DIS, y=T_SOC)) + geom_boxplot() +
  stat_summary(fun.y=mean, geom="point", shape=5, size=4)+
  theme(axis.text.x=element_text(angle = 90, vjust = 0.5))+
  labs(x=NULL, y="SOC (kg/m2)")



#Create a database for the total SOC with 0-100 cm
#ID, MU, Divisions, total SOC

DIV<- data.frame(MU=BD_BTM2$MU_GLOBAL, DIV=BD_BTM2$Division,
                Tot_SOC=BD_BTM2$Tot_SOC)

names(DIV)


#Add another column with the id
ID <- rownames(DIV)


# Combine the ID wih the DIV database
DIV2 <- cbind(ID=ID, DIV)

names(DIV2)

ggplot(DIV2, aes(x=DIV, y=Tot_SOC)) + geom_boxplot() +
  stat_summary(fun.y=mean, geom="point", shape=5, size=4)+
  theme(axis.text.x=element_text(angle = 90, vjust = 0.5))+
  labs(x=NULL, y="SOC (kg/m2)")+
  scale_y_continuous( limits=c(0, 100))
```