

Root necrosis

2023-10-23

#####In the initial dataset, data for most checks was not captured during image capture. However, when we consolidated the data at the plant level, both root and image numbers from the missing checks were assigned as zero . This inadvertently introduced a level (“zero”) that was not supposed to exist in the dataset.

#To rectify this issue, all checks with a zero level were subsequently removed from the dataset. The mixed model analysis was then re-executed using this refined data. Importantly, this modification should not impact the GWAS portion of the paper, as these checks did not possess genotype data and were originally excluded from the analysis.

##plan is do not combine plant and image data files

#analyze them seperatly until you get Blups to merge files

#required packages

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.3      v readr      2.1.4
## v forcats    1.0.0      v stringr    1.5.0
## v ggplot2     3.4.4      v tibble     3.2.1
## v lubridate  1.9.3      v tidyr      1.3.0
## v purrr      1.0.2
```

```
## -- Conflicts ----- tidyverse_conflicts() --
```

```
## x dplyr::filter() masks stats::filter()
```

```
## x dplyr::lag()      masks stats::lag()
```

```
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(dplyr)
```

```
library(lme4)
```

```
## Loading required package: Matrix
```

```
##
```

```
## Attaching package: 'Matrix'
```

```
##
```

```
## The following objects are masked from 'package:tidyr':
```

```
##
```

```
##      expand, pack, unpack
```

```
library(tidyr)
```

```
library(readr)
```

```
library(ggpubr)
```

```
library(minque)
```

```
#####Namulonge first year data 2020/2021
Namulonge_2020_formatted <- read.csv("~/Desktop/QPCR_PAPER/Rootnecrosis_paper/Data/Namulonge_2020_forma
dim(Namulonge_2020_formatted ) #19532    29
```

```
## [1] 19532    29
```

```
library(data.table)
```

```
##
```

```
## Attaching package: 'data.table'
```

```
## The following objects are masked from 'package:lubridate':
```

```
##
```

```
##      hour, isoweek, mday, minute, month, quarter, second, wday, week,
```

```
##      yday, year
```

```
## The following objects are masked from 'package:dplyr':
```

```
##
```

```
##      between, first, last
```

```
## The following object is masked from 'package:purrr':
```

```
##
```

```
##      transpose
```

```
#Namulonge_2020_for <- separate(Namulonge_2020_formatted, Image.ID, into = "Clone_Id", sep =30, remove .
```

```
Namulonge_2020_for1 <- data.frame(do.call("rbind", strsplit(sub('(^[_]+_[^_]+_[^_]+_[^_]+_[^_]+_[^_]+_
```

```
#Naming the columns
```

```
names(Namulonge_2020_for1)[1] <- "ID"
```

```
names(Namulonge_2020_for1)[2] <- "Other"
```

```
dim(Namulonge_2020_for1) #19532    2
```

```
## [1] 19532    2
```

```
Names_ids <- Namulonge_2020_formatted[ , c(1,2)]
```

```
data_table_1 = data.table(Namulonge_2020_for1)
```

```
data_table_2 = data.table(Names_ids)
```

```
DATA2 <-cbind(data_table_1, data_table_2)
```

```
Namulonge_2020_for1.1 <- data.frame(do.call("rbind", strsplit(sub('(^[_]+_[^_]+_[^_]+_[^_]+_[^_]+_[^_]+_
```

```

#Naming the columns
names(Namulonge_2020_for1.1)[1] <- "Plot_id"
names(Namulonge_2020_for1.1)[2] <- "Plant_Num"
dim(Namulonge_2020_for1.1) #19532      2

## [1] 19532      2

Names_ids.1 <- Namulonge_2020_formatted[, c(1,2)]

data_table_1.1 = data.table(Namulonge_2020_for1.1)
data_table_2.1 = data.table(Names_ids.1)

DATA2.1 <- cbind(data_table_1.1, data_table_2.1)

Namulonge_2020_for2 <- data.frame(do.call("rbind", strsplit(sub('^([_]+)(.*)$', '\\1 \\2', Namulonge_2020_for1.1), '\\1 \\2')))
names(Namulonge_2020_for2)[1] <- "Root_number"
names(Namulonge_2020_for2)[2] <- "Other1"
data_table_3 = data.table(Namulonge_2020_for2)
DATA3 <- cbind(DATA2.1, data_table_3)

Namulonge_2020_for3 <- data.frame(do.call("rbind", strsplit(sub('^([_]+)(.*)$', '\\1 \\2', Namulonge_2020_for2), '\\1 \\2')))
names(Namulonge_2020_for3)[1] <- "Image_number"
names(Namulonge_2020_for3)[2] <- "Other2"
data_table_4 = data.table(Namulonge_2020_for3)
DATA4 <- cbind(DATA3, data_table_4)
DATA5 <- cbind(DATA4, DATA2.1)
write_csv(DATA5, file = "Namulonge_2020_map.csv")
names(DATA5)[11] <- "Xs"
names(DATA5)[12] <- "dropped"
DATA5 <- subset(DATA5, select = -c(X,Xs,dropped))

Root_necrosis_Namulonge_2020 <- merge(Namulonge_2020_formatted, DATA5, by="Image.ID")
dim(Root_necrosis_Namulonge_2020) #19532      37

## [1] 19532      37

####add year location in the file

length_of_column <- length(Root_necrosis_Namulonge_2020$Image.ID)

Root_necrosis_Namulonge_2020$Location <- rep('Namulonge', length_out = length_of_column)

Root_necrosis_Namulonge_2020$Year <- rep('2019', length_out = length_of_column)

#second year Namulonge

#Namulonge_2021
Namulonge_2021_formatted <- read.csv("~/Desktop/QPCR_PAPER/Rootnecrosis_paper/Data/Namulonge_2021_formatted.csv")
dim(Namulonge_2021_formatted) #15240      29

```

```
## [1] 15240      29
```

```
Namulonge_2021_for1 <- data.frame(do.call("rbind", strsplit(sub('^[_]+[_]+[_]+[_]+[_]+[_]+[_]+',  
#Naming the columns  
names(Namulonge_2021_for1)[1] <- "ID"  
names(Namulonge_2021_for1)[2] <- "Other"  
dim(Namulonge_2021_for1) #15240      2
```

```
## [1] 15240      2
```

```
Names_ids_2021 <- Namulonge_2021_formatted[ , c(1,2)]
```

```
data_table_1_2021 = data.table(Namulonge_2021_for1)  
data_table_2_2021 = data.table(Names_ids_2021)
```

```
DATA2_2021 <- cbind(data_table_1_2021, data_table_2_2021)
```

```
Namulonge_2021_for1.1 <- data.frame(do.call("rbind", strsplit(sub('^[_]+[_]+[_]+[_]+[_]+[_]+[_]+',  
#Naming the columns  
names(Namulonge_2021_for1.1)[1] <- "Plot_id"  
names(Namulonge_2021_for1.1)[2] <- "Plant_Num"  
dim(Namulonge_2021_for1.1) #15240      2
```

```
## [1] 15240      2
```

```
Names_ids.1 <- Namulonge_2021_formatted[ , c(1,2)]
```

```
data_table_1.1 = data.table(Namulonge_2021_for1.1)  
data_table_2.1 = data.table(Names_ids.1)
```

```
DATA2.11 <- cbind(data_table_1.1, data_table_2.1)
```

```
Namulonge_2021_for2 <- data.frame(do.call("rbind", strsplit(sub('^[_]+_(_.*)$', '\\1 \\2', Namulonge_  
names(Namulonge_2021_for2)[1] <- "Root_number"  
names(Namulonge_2021_for2)[2] <- "Other1"  
data_table_3_2021 = data.table(Namulonge_2021_for2)  
DATA3_2021 <- cbind(DATA2_2021, data_table_3_2021)
```

```
Namulonge_2021_for3 <- data.frame(do.call("rbind", strsplit(sub('^[_]+_(_.*)$', '\\1 \\2', Namulonge_  
names(Namulonge_2021_for3)[1] <- "Image_number"  
names(Namulonge_2021_for3)[2] <- "Other2"
```

```
data_table_4_2021 = data.table(Namulonge_2021_for3)
DATA4_2021 <- cbind(DATA3_2021, data_table_4_2021)
```

```
DATA5_2021 <- cbind(DATA4_2021, DATA2.11)
dim(DATA5_2021) # 15240 12
```

```
## [1] 15240 12
```

```
write_csv(DATA5_2021, file = "Namulonge_2021_map.csv")
```

```
names(DATA5_2021)[11] <- "Xs"
names(DATA5_2021)[12] <- "dropped"
DATA5_2021 <- subset(DATA5_2021, select = -c(X,Xs,dropped))
```

```
Root_necrosis_Namulonge_2021 <- merge(Namulonge_2021_formatted, DATA5_2021, by="Image.ID")
dim(Root_necrosis_Namulonge_2021) #15240 37
```

```
## [1] 15240 37
```

```
####add year location in the file
```

```
length_of_column <- length(Root_necrosis_Namulonge_2021$Image.ID)

Root_necrosis_Namulonge_2021$Location <- rep('Namulonge', length_out = length_of_column)

Root_necrosis_Namulonge_2021$Year <- rep('2020', length_out = length_of_column)
```

```
# extracting the namulonge code for 2019 to get the gerplasm name that will be used in creating the code
Namulonge_2019 <- read.csv("~/Desktop/QPCR_PAPER/Rootnecrosis_paper/Data/Namulonge_2019.csv")
dim(Namulonge_2019) #3592 33
```

```
## [1] 3592 33
```

```
Name_codes2019 <- Namulonge_2019 %>% dplyr::select(germplasmName , observationUnitName)
```

```
names(Root_necrosis_Namulonge_2021)[names(Root_necrosis_Namulonge_2021) == "ID"] <- "observationUnitName"
```

```
Root_necrosis_Namulonge_2021.1 <- merge(Root_necrosis_Namulonge_2021, Name_codes2019, by="observationUnitName")
```

```
Root_necrosis_Namulonge_2021.2 <- subset (Root_necrosis_Namulonge_2021.1, select = -observationUnitName)
```

```
dim(Root_necrosis_Namulonge_2021.2) #15240 39
```

```
## [1] 15240 39
```

```
####add year location in the file
```

```
length_of_column <- length(Root_necrosis_Namulonge_2021.2$Image.ID)
```

```
Root_necrosis_Namulonge_2021.2$Location <- rep('Namulonge', length_out = length_of_column)

Root_necrosis_Namulonge_2021.2$Year <- rep('2020', length_out = length_of_column)
```

```
#serere
Serere_2020_formatted <- read.csv("~/Desktop/QPCR_PAPER/Rootnecrosis_paper/Data/Serere_2020_formatted.csv")
dim(Serere_2020_formatted) #9949 29
```

```
## [1] 9949 29
```

```
Serere_2020_for1 <- data.frame(do.call("rbind", strsplit(sub('^[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_'+
```

```
#Naming the columns
names(Serere_2020_for1)[1] <- "ID"
names(Serere_2020_for1)[2] <- "Other"
dim(Serere_2020_for1) #9949 2
```

```
## [1] 9949 2
```

```
Names_ids_serere <- Serere_2020_formatted[, c(1,2)]

data_table_1_serere = data.table(Serere_2020_for1)
data_table_2_serere = data.table(Names_ids_serere)
```

```
DATA2_serere <- cbind(data_table_1_serere, data_table_2_serere)
```

```
Serere_2020_for1.1 <- data.frame(do.call("rbind", strsplit(sub('^[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_'+
```

```
#Naming the columns
names(Serere_2020_for1.1)[1] <- "Plot_id"
names(Serere_2020_for1.1)[2] <- "Plant_Num"
dim(Serere_2020_for1.1) #9949 2
```

```
## [1] 9949 2
```

```
Names_ids_serere.1 <- Serere_2020_formatted[, c(1,2)]
```

```
data_table_1.1 = data.table(Serere_2020_for1.1)
data_table_2.1 = data.table(Names_ids_serere.1)
```

```
DATA2.11 <- cbind(data_table_1.1, data_table_2.1)
```

```
dim(DATA2.11) #9949 4
```

```
## [1] 9949 4
```

```

Serere_2020_for2 <- data.frame(do.call("rbind", strsplit(sub('^[_]+)(.*)$', '\\1 \\2', Serere_2020_f
names(Serere_2020_for2)[1] <- "Root_number"
names(Serere_2020_for2)[2] <- "Other1"
data_table_3_serere = data.table(Serere_2020_for2)
DATA3_serere <- cbind(DATA2_serere, data_table_3_serere)

Serere_2020_for3 <- data.frame(do.call("rbind", strsplit(sub('^[_]+)(.*)$', '\\1 \\2', Serere_2020_f
names(Serere_2020_for3)[1] <- "Image_number"
names(Serere_2020_for3)[2] <- "Other2"
data_table_4_serere = data.table(Serere_2020_for3)

DATA4_serere <- cbind(DATA3_serere, data_table_4_serere)

DATA5_serere <- cbind(DATA4_serere, DATA2.11)

write_csv(DATA5_serere, file = "Serere_2020_map.csv")

names(DATA5_serere)[11] <- "Xs"
names(DATA5_serere)[12] <- "dropped"
DATA5_serere <- subset(DATA5_serere, select = -c(X,Xs,dropped))

Root_necrosis_Serere_2020 <- merge(Serere_2020_formatted, DATA5_serere, by="Image.ID")

dim(Root_necrosis_Serere_2020) #9949 37

## [1] 9949 37

####add year location in the file

length_of_column <- length(Root_necrosis_Serere_2020 $Image.ID)

Root_necrosis_Serere_2020 $Location <- rep('Serere', length_out = length_of_column)
Root_necrosis_Serere_2020 $Year <- rep('2019', length_out = length_of_column)

# Serere_2021_data -----

Serere_2021_formatted <- read.csv("~/Desktop/QPCR_PAPER/Rootnecrosis_paper/Data/Serere_2021_formatted.c
dim(Serere_2021_formatted) #5131 29

## [1] 5131 29

```

```
Serere_2021_2_formatted <- read.csv("~/Desktop/QPCR_PAPER/Rootnecrosis_paper/Data/Serere_2021_2_formatted.csv")
dim(Serere_2021_2_formatted) #1522 29
```

```
## [1] 1522 29
```

```
Serere_2021_2_for1 <- data.frame(do.call("rbind", strsplit(sub('^([_]+)[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_'+
```

```
#Naming the columns
names(Serere_2021_2_for1)[1] <- "ID"
names(Serere_2021_2_for1)[2] <- "Other"
dim(Serere_2021_2_for1) #1522 2
```

```
## [1] 1522 2
```

```
Names_ids <- Serere_2021_2_formatted[,c(1,2)]

data_table_1 = data.table(Serere_2021_2_for1)
data_table_2 = data.table(Names_ids)
```

```
DATA2 <-cbind(data_table_1, data_table_2)
```

```
Serere_2021_2_for1.1 <- data.frame(do.call("rbind", strsplit(sub('^([_]+)[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_'+
```

```
df1_separated <- Serere_2021_2_for1.1%>%
  separate(X2, into = c("Plot_id", "Plant_Num"), sep = "_")

dim(df1_separated)
```

```
## [1] 1522 3
```

```
#Naming the columns
#names(Serere_2021_2_for1.1)[1] <- "Plot_id"
#names(Serere_2021_2_for1.1)[2] <- "Plant_Num"
#dim(Serere_2021_2_for1.1) #1522 2

#Names_ids.1 <- Serere_2021_2_formatted[,c(1,2)]

#data_table_1.1 = data.table(Serere_2021_2_for1.1)
#data_table_2.1 = data.table(Names_ids.1)

#DATA2.1 <-cbind(data_table_1.1, data_table_2.1)

Serere_2021_2_for2 <- data.frame(do.call("rbind", strsplit(sub('^([_]+)[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_]+[_'+
```

```
names(Serere_2021_2_for2)[1] <- "Root_number"
```



```

names(Serere_2021_2_for2)[2] <- "Other1"
data_table_3 = data.table(Serere_2021_2_for2)
DATA3 <-cbind(DATA2, data_table_3)

Serere_2021_2_for3 <- data.frame(do.call("rbind", strsplit(sub('^[_]+)(.*)$', '\\1 \\2', Serere_2021_2_for2),
names(Serere_2021_2_for3)[1] <- "Image_number"
names(Serere_2021_2_for3)[2] <- "Other2"
data_table_4 = data.table(Serere_2021_2_for3)
DATA4 <-cbind(DATA3, data_table_4)

DATA5 <-cbind(DATA4, df1_separated)

dim(DATA5)

```

```
## [1] 1522 11
```

```
write_csv(DATA4, file = "Serere_2021_2_map.csv")
```

```

dataserere_2021_necrosis <-merge(Serere_2021_2_formatted, DATA5, by="Image.ID")
dim(dataserere_2021_necrosis) #1522 39

```

```
## [1] 1522 39
```

```

dataserere_2021_necrosis <- subset(dataserere_2021_necrosis, select = -X.y)

dataserere_2021_necrosis <- dataserere_2021_necrosis %>% rename(X = X.x)
dim(dataserere_2021_necrosis) #1522 35

```

```
## [1] 1522 38
```

```

#####location year effects

length_of_column <- length(dataserere_2021_necrosis$Image.ID)

dataserere_2021_necrosis$Location <- rep('Serere', length_out = length_of_column)
dataserere_2021_necrosis$Year <- rep('2020', length_out = length_of_column)

dim(dataserere_2021_necrosis) #1522 40

```

```
## [1] 1522 40
```

```

#second file
#this is file observation ids were used
dim(Serere_2021_formatted) #5131 29

```

```
## [1] 5131 29
```

```
Serere_2021_for1 <- data.frame(do.call("rbind", strsplit(sub('^([_]+)(.*)$', '\\1 \\2', Serere_2021_f

#Naming the columns
names(Serere_2021_for1)[1] <- "ID"
names(Serere_2021_for1)[2] <- "Other"
dim(Serere_2021_for1) #5131 2
```

```
## [1] 5131 2
```

```
Names_ids <- Serere_2021_formatted[,c(1,2)]
```

```
data_table_1 = data.table(Serere_2021_for1)
data_table_2 = data.table(Names_ids)
```

```
DATA2_2021_serere <- cbind(data_table_1, data_table_2)
```

```
Serere_2021_for2 <- data.frame(do.call("rbind", strsplit(sub('^([_]+)(.*)$', '\\1 \\2', Serere_2021_f
names(Serere_2021_for2)[1] <- "Root_number"
names(Serere_2021_for2)[2] <- "Other1"
data_table_3 = data.table(Serere_2021_for2)
DATA3 <- cbind(DATA2_2021_serere, data_table_3)
```

```
Serere_2021_for3 <- data.frame(do.call("rbind", strsplit(sub('^([_]+)(.*)$', '\\1 \\2', Serere_2021_2
names(Serere_2021_for3)[1] <- "Image_number"
names(Serere_2021_for3)[2] <- "Other2"
data_table_4 = data.table(Serere_2021_for3)
DATA4 <- cbind(DATA3, data_table_4)
```

```
## Warning in as.data.table.list(x, keep.rownames = keep.rownames, check.names =
## check.names, : Item 2 has 1522 rows but longest item has 5131; recycled with
## remainder.
```

```
dim(DATA4) #5131 8
```

```
## [1] 5131 8
```

```
write_csv(DATA4, file = "Serere_2021_version2_map.csv")
```

```
#getting the plot id
library(dplyr)
Serere_2020 <- read.csv("~/Desktop/QPCR_PAPER/Rootnecrosis_paper/Data/Serere_2020.csv")
dim(Serere_2020) #3018 33
```

```
## [1] 3018 33
```

```
Namings_id <- Serere_2020 %>% dplyr::select(observationUnitDbId , observationUnitName)

names(DATA4)[1] <- "observationUnitDbId"

DATA4$observationUnitDbId <- as.numeric(DATA4$observationUnitDbId )
dim(DATA4) #5131 8
```

```
## [1] 5131 8
```

```
Dataset_serere1 <- merge(DATA4,Namings_id, by= "observationUnitDbId", all.x = TRUE)
dim(Dataset_serere1) # 5131 9
```

```
## [1] 5131 9
```

```
Serere_necrosis_2021_part1 <- merge(Dataset_serere1,Serere_2021_formatted,by="Image.ID")
dim(Serere_necrosis_2021_part1) #5131 37
```

```
## [1] 5131 37
```

```
#####BECAUSE THERE WERE NO PLOTS ATTACHED USE THE OBSERVATION UNIT NAME TO EXTRACT THE PLOT AND PLANT
```

```
Serere_2021_part_1.2 <- data.frame(do.call("rbind", strsplit(sub('(^[_]+_[^_]+_[^_]+_[^_]+_[^_]+)(.*)$'),
```

```
df_separated <- Serere_2021_part_1.2 %>%
  separate(X2, into = c("Plot_id", "Plant_Num"), sep = "_")
```

```
DATA6_serere <- subset(df_separated, select = - X1)
```

```
DATA6_serere$observationUnitName <-Serere_necrosis_2021_part1$observationUnitName
```

```
Serere_2021_2 <-bind_cols(Serere_necrosis_2021_part1, DATA6_serere)
```

```
## New names:
```

```
## * 'observationUnitName' -> 'observationUnitName...9'
## * 'observationUnitName' -> 'observationUnitName...40'
```

```
dim(Serere_2021_2)
```

```
## [1] 5131 40
```

```
length_of_column <- length(Serere_2021_2$Image.ID)
```

```
Serere_2021_2 $Location <- rep('Serere', length_out = length_of_column)
```

```
Serere_2021_2 $Year <- rep('2020', length_out = length_of_column)
```

```
#####merge final dataset
```

```
dim(Root_necrosis_Namulonge_2020) #19532 39
```

```
## [1] 19532 39
```

```
dim(Root_necrosis_Namulonge_2021.2) #15240 39
```

```
## [1] 15240 39
```

```
dim(Root_necrosis_Serere_2020) #9949 39
```

```
## [1] 9949 39
```

```
dim(dataserere_2021_necrosis) #1522 40
```

```
## [1] 1522 40
```

```
dim(Serere_2021_2) #5131 42
```

```
## [1] 5131 42
```

```
Root1 <- merge(Root_necrosis_Namulonge_2020,Root_necrosis_Namulonge_2021.2, all = TRUE)  
dim(Root1)
```

```
## [1] 34772 40
```

```
Root2 <- merge(Root1,Root_necrosis_Serere_2020, all = TRUE)  
dim(Root2) #44721
```

```
## [1] 44721 40
```

```
Root3 <- merge(Root2,dataserere_2021_necrosis, all = TRUE)  
dim(Root3) #46243 41
```

```
## [1] 46243 41
```

```
Root4 <- merge(Root3,Serere_2021_2, all = TRUE)  
dim(Root4) # 51374 46
```

```
## [1] 51374 46
```

```
write.csv(Root4 , file = "root_necrosis_2023.csv")
```

```
dim(Namulonge_2020_formatted) #19532 29
```

```
## [1] 19532 29
```

```
dim(Namulonge_2021_formatted) #15240 29
```

```
## [1] 15240 29
```

```
dim(Serere_2020_formatted) #9949 29
```

```
## [1] 9949 29
```

```
dim(Serere_2021_formatted) #5131 29
```

```
## [1] 5131 29
```

```
dim(Serere_2021_2_formatted )#1522 29
```

```
## [1] 1522 29
```

```
#total <- 51374
```