Distribution diamétrique (INRAE)

2024-04-04

Ce script est créé dans le but d’observer les distributions diamétriques des espèces que l’on va étudier dans la deuxième partie. La deuxième partie sera réalisé sur les données de l’INRAE qui comprend 29espèces d’intérêt.

#Mise en place

setwd("~/cours\_r/Stage\_M2")

library(tidyverse) #pour toutes les fonctions associées (dplr, ggplot2,...)

## Warning: le package 'tidyverse' a été compilé avec la version R 4.3.3

## Warning: le package 'ggplot2' a été compilé avec la version R 4.3.3

## Warning: le package 'tibble' a été compilé avec la version R 4.3.3

## Warning: le package 'tidyr' a été compilé avec la version R 4.3.3

## Warning: le package 'readr' a été compilé avec la version R 4.3.3

## Warning: le package 'purrr' a été compilé avec la version R 4.3.3

## Warning: le package 'dplyr' a été compilé avec la version R 4.3.3

## Warning: le package 'stringr' a été compilé avec la version R 4.3.3

## Warning: le package 'forcats' a été compilé avec la version R 4.3.3

## Warning: le package 'lubridate' a été compilé avec la version R 4.3.3

## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
## ✔ dplyr 1.1.4 ✔ readr 2.1.5  
## ✔ forcats 1.0.0 ✔ stringr 1.5.1  
## ✔ ggplot2 3.5.0 ✔ tibble 3.2.1  
## ✔ lubridate 1.9.3 ✔ tidyr 1.3.1  
## ✔ purrr 1.0.2   
## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()  
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

Nous allons nous servir de l’inventaire floristique du projet ALT réalisé sur 5ha de la P16 à Paracou (Guyane fr). C’est un inventaire exhaustive des arbres à un diamètre à hauteur de poitrine (DBH) à partir de 1cm. Il a été réalisé à partir de 2018 sur 3ans.

#ALT  
paracou\_9ha <- read.csv("~/cours\_r/Stage\_M2/understory\_paracou\_9ha\_sp\_20240126.csv")  
paracou\_G <- paracou\_9ha[,c(2:25)] #TreeID jusqu'à scientificName car erreur  
  
#INRAE (Regeneration plot)  
Paracou\_Juveniles <- read.csv("~/cours\_r/Stage\_M2/2024ParacouJuveniles(v2).csv", sep=";", comment.char="#")  
  
#Création d'une colonne scientificName  
Paracou\_Juveniles <- Paracou\_Juveniles %>%  
 mutate(scientificName = paste(Genus, Species, sep = " "))

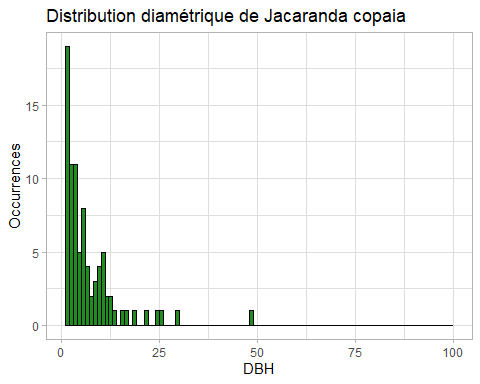
Nous avons extraits la liste des espèces d’intérêt de l’INRAE.

INRAE <- Paracou\_Juveniles %>%  
 filter(Project == "ParacouRegeneration") %>%  
 distinct(scientificName) %>%  
 pull()

#Distribution simple

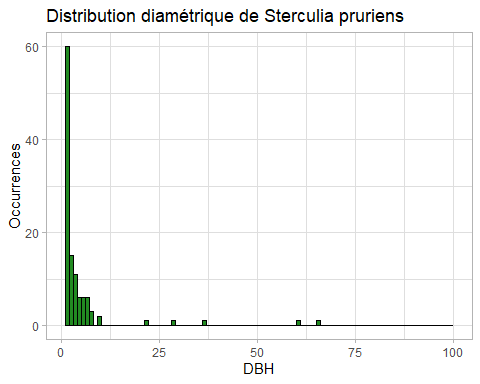
#Création de l'objet comprenant l'ensemble des graphes  
INRAE\_c <- list()  
  
#Création de la boucle  
for(k in 1:length(INRAE)) {  
 plot.data <- paracou\_G %>%  
 filter (scientificName == INRAE[[k]]) %>%  
 group\_by(DBH) %>%   
 arrange (DBH) %>%   
 ggplot(aes (x = DBH)) +  
 geom\_histogram(breaks = seq(1, 100, by = 1 ),fill = "forestgreen", color = "black") + #début à 1  
 theme\_light() +  
 labs(x = "DBH", y = "Occurrences", title = paste("Distribution diamétrique de", INRAE[k]))  
 #ajout du graphique à la liste  
 INRAE\_c[[k]] <- plot.data  
}  
  
#Affichage des graphes contenus dans la liste  
print(INRAE\_c)

## [[1]]

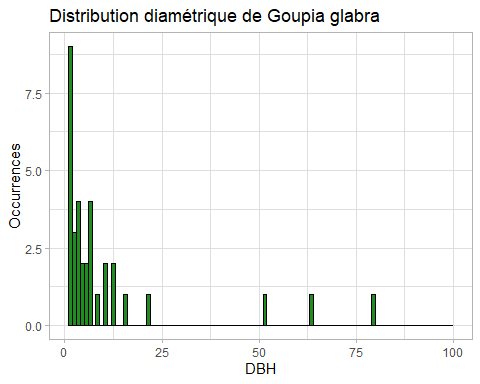


##   
## [[2]]

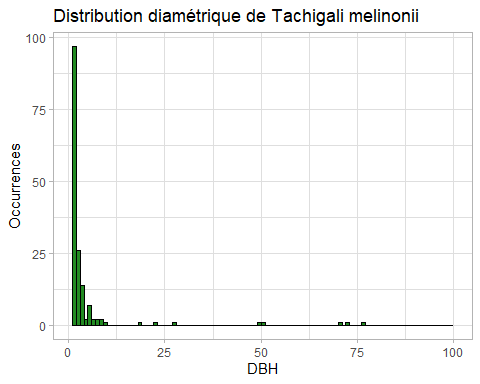
## Warning: Removed 1 row containing non-finite outside the scale range  
## (`stat\_bin()`).



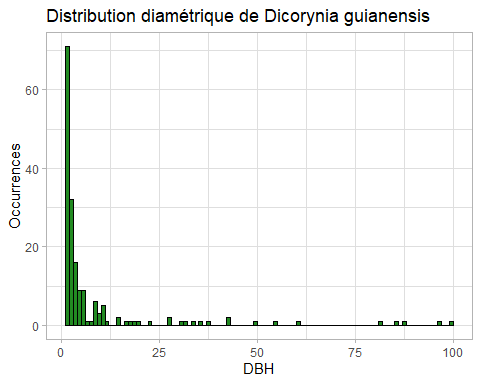
##   
## [[3]]



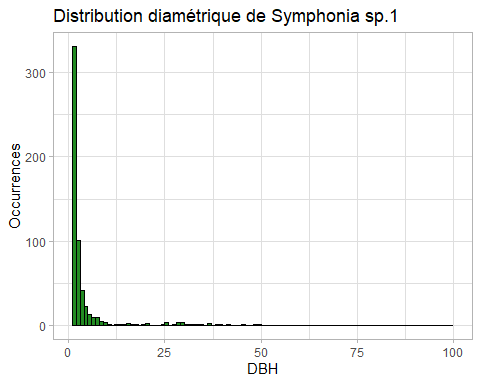
##   
## [[4]]



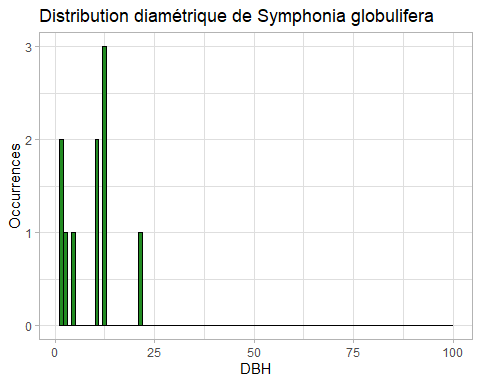
##   
## [[5]]



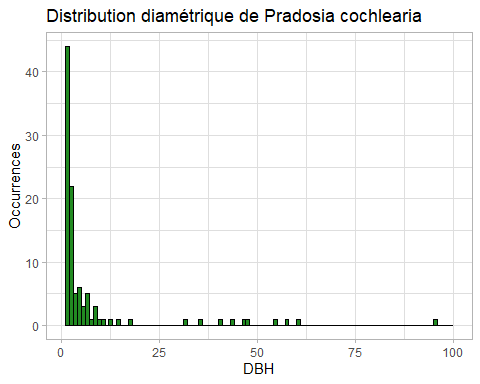
##   
## [[6]]



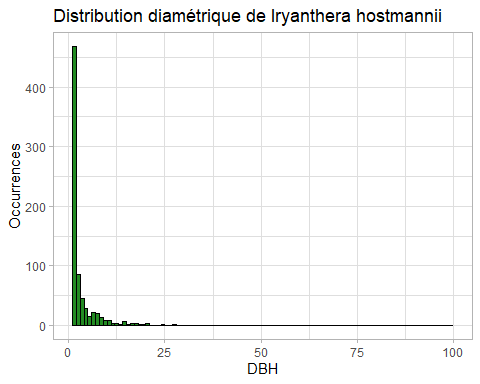
##   
## [[7]]



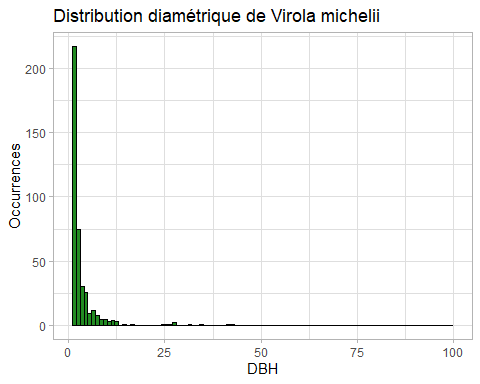
##   
## [[8]]



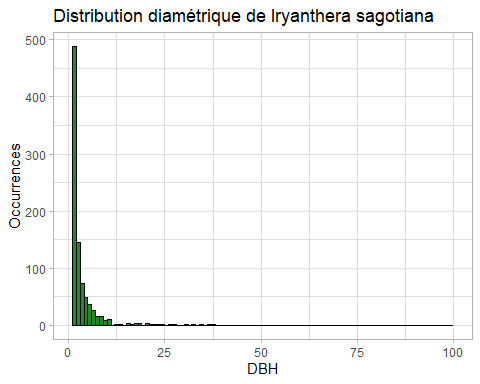
##   
## [[9]]



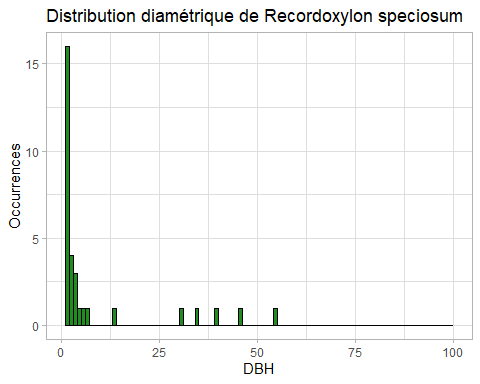
##   
## [[10]]



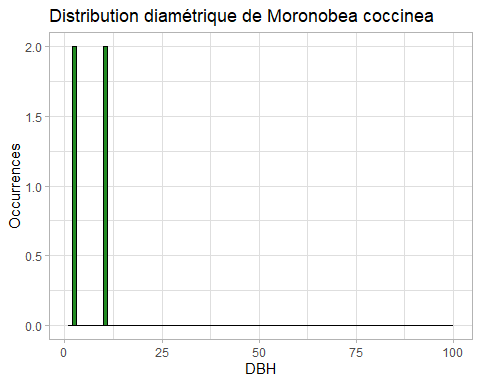
##   
## [[11]]



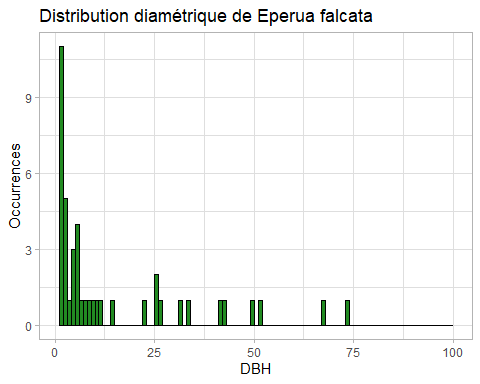
##   
## [[12]]



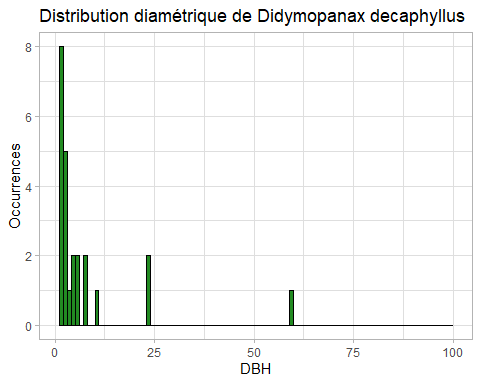
##   
## [[13]]



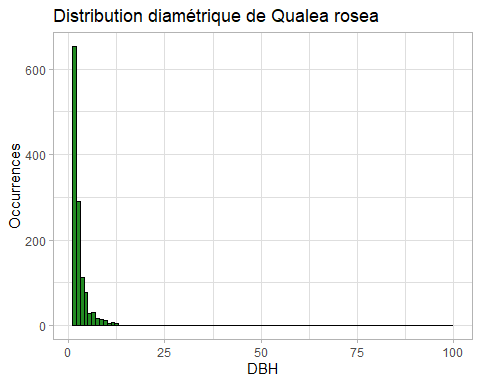
##   
## [[14]]



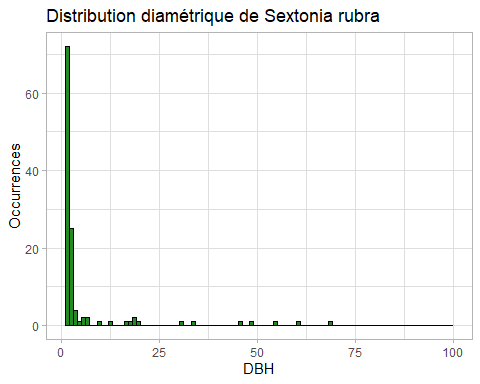
##   
## [[15]]



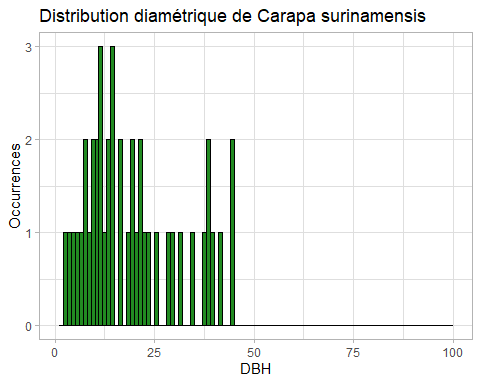
##   
## [[16]]



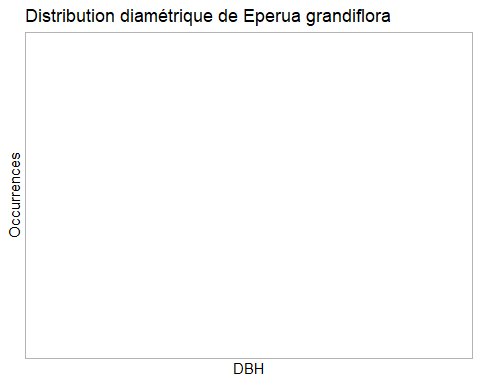
##   
## [[17]]



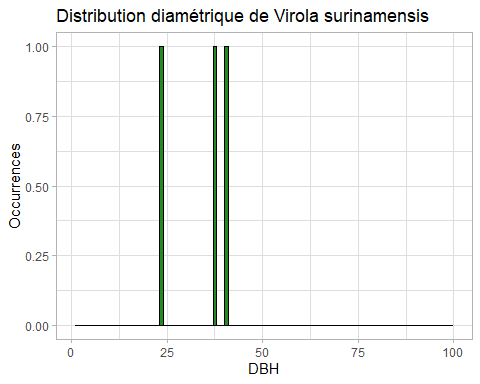
##   
## [[18]]



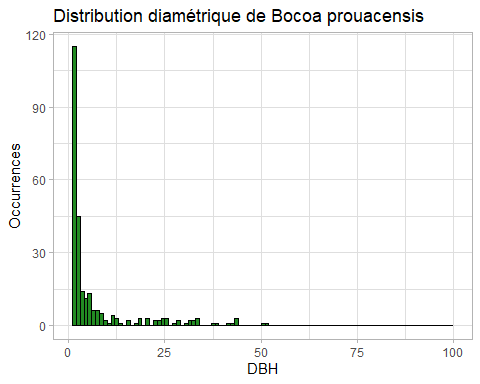
##   
## [[19]]



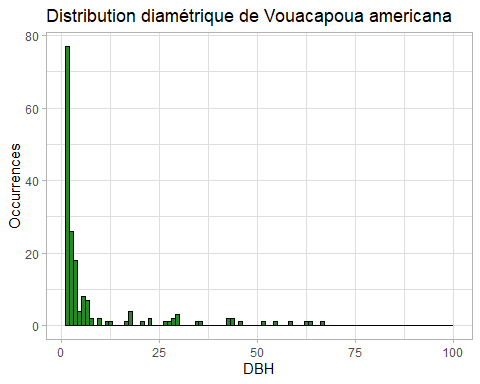
##   
## [[20]]



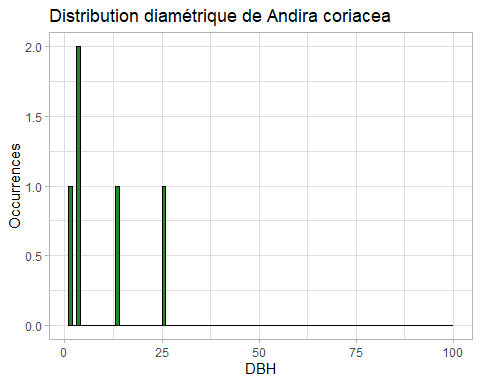
##   
## [[21]]



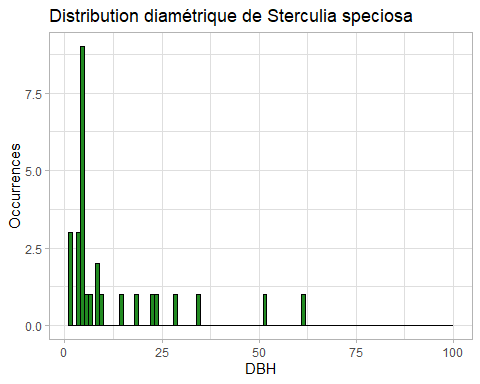
##   
## [[22]]



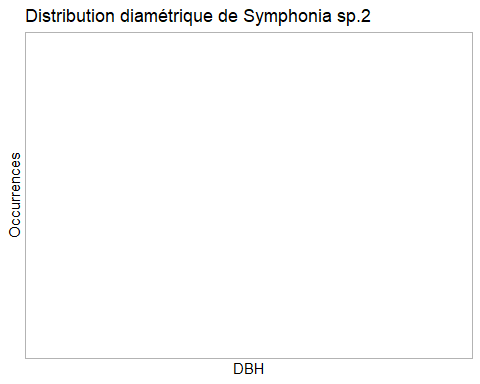
##   
## [[23]]



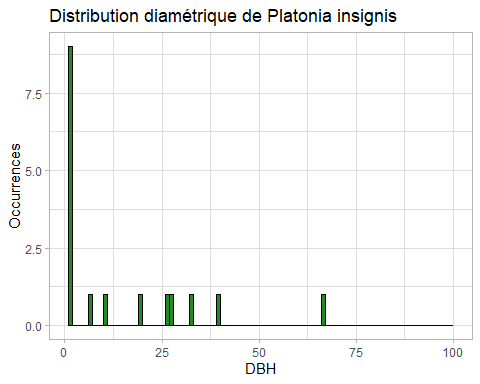
##   
## [[24]]



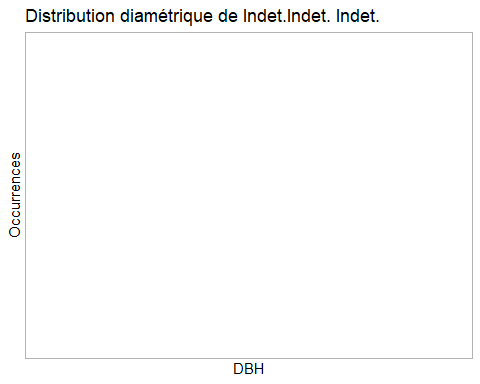
##   
## [[25]]



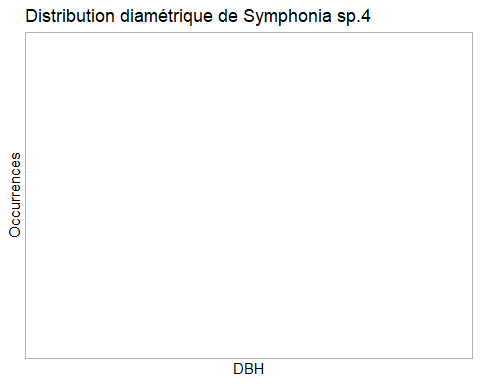
##   
## [[26]]



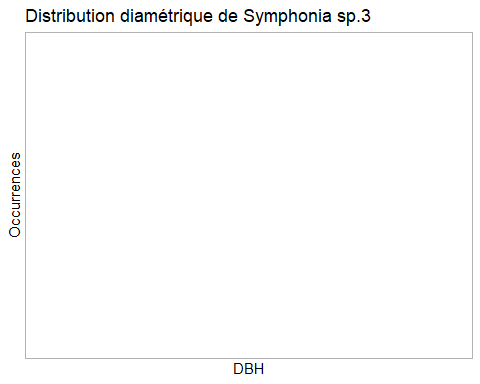
##   
## [[27]]



##   
## [[28]]



##   
## [[29]]

 De visuel, les distributions diamétrique d’une majorité d’entre elles sont exploitable. Nous nottons que les *Symphonia spp.* ainsi que *Eperua grandiflora* et *Ind* (indéterminé) n’ont pas de correspondance dans notre jeu de donnée et donc pas de données associée. Parmis les faibles occurences dans les distributions nous avons : \* [0;10] *Goupia Glabra*, *Symphonia globulifera*, *Moronobea coccinea* (4), *Didymlipanaw decaphyllus*, *Carapa surinamensis*, *Virola surinamensis* (3), *Andira coriacea* (5), *Sterculia speciosa*, *Platonia insignis* \* [10;30] *Recordoxylon speciosum*, *Eperua falcata*, *Jacaranda copaia*

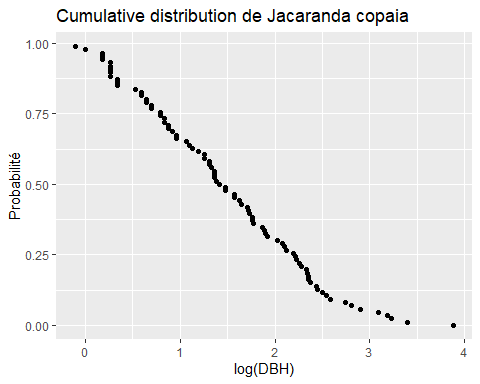
Finalement dans la liste des 29 il nous reste 21 espèces où l’on pourrait potentiellement ressortir quelque chose de bien. NOus allons retirer dèsaprésent les espèces avec trop peu de relevés.

#Retrait de la liste des espèces non-utilisable  
not <- c("Moronobea coccinea","Virola surinamensis", "Andira coriacea","Eperua grandiflora","Symphonia sp.3", "Symphonia sp.4", "Symphonia sp.2","Indet.Indet. Indet.")  
  
INRAE\_2 <- setdiff(INRAE, not)

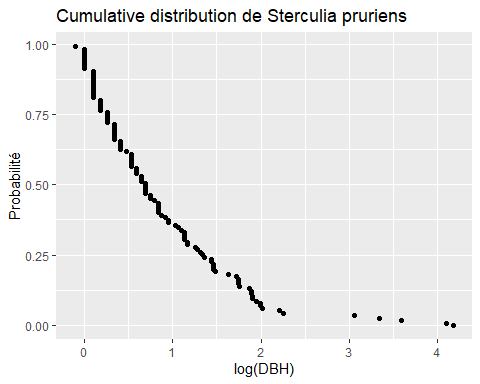
#Cumulative distribution function La cumulative distribution est une fonction de probabilité, je ne sais plus tout s’est effacé.

#Création de la liste  
INRAE\_cdf <- list()  
  
#Création des courbes  
for(k in 1:length (INRAE\_2)) {  
 plot.cdf <- paracou\_G %>%  
 filter(scientificName == INRAE\_2[[k]]) %>%  
 select(DBH) %>%  
 na.omit() %>%  
 mutate(DBH = sort(DBH),  
 ord = order(DBH),  
 cdf = 1 - (ord / max(ord))) %>%  
 ggplot(aes(x = log(DBH), y = cdf)) +  
 geom\_point() +  
 labs(x = "log(DBH)", y = "Probabilité", title = paste("Cumulative distribution de", INRAE\_2[k]))  
#ajout du graphique à la liste   
 INRAE\_cdf[[k]] <- plot.cdf  
}  
  
#Affichage des graphes  
print(INRAE\_cdf)

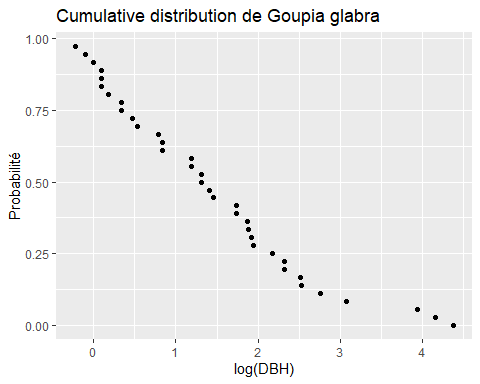
## [[1]]



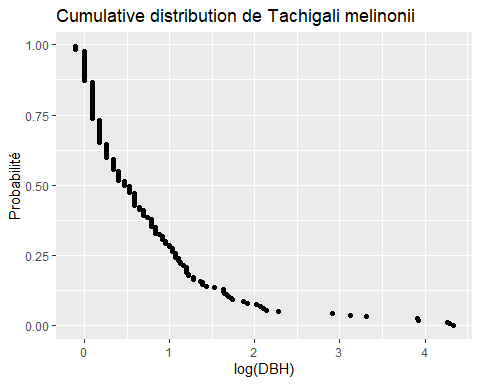
##   
## [[2]]



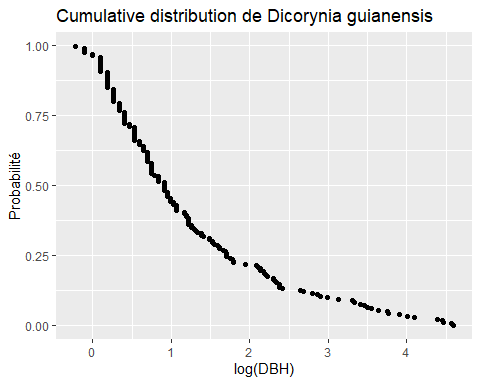
##   
## [[3]]



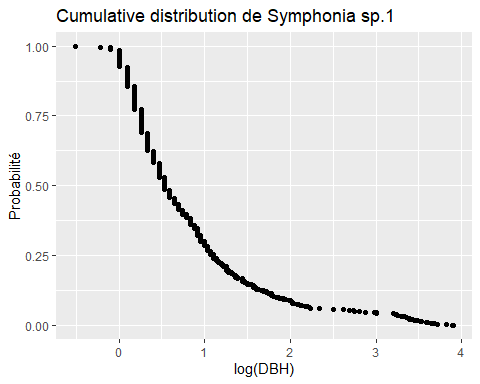
##   
## [[4]]



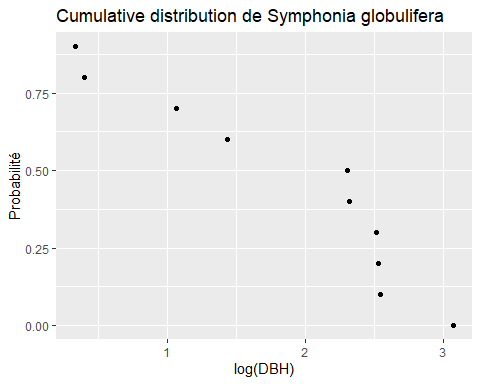
##   
## [[5]]



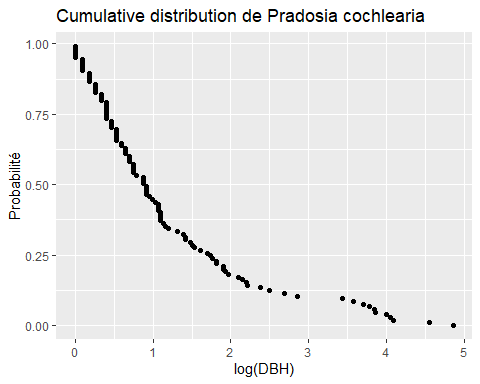
##   
## [[6]]



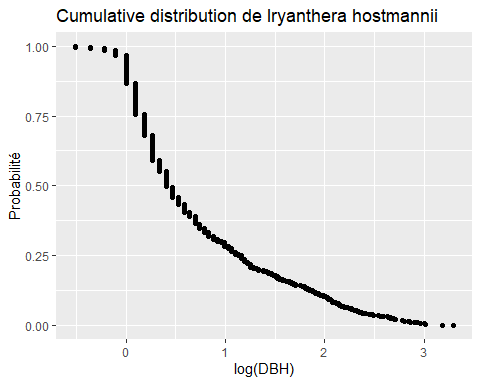
##   
## [[7]]



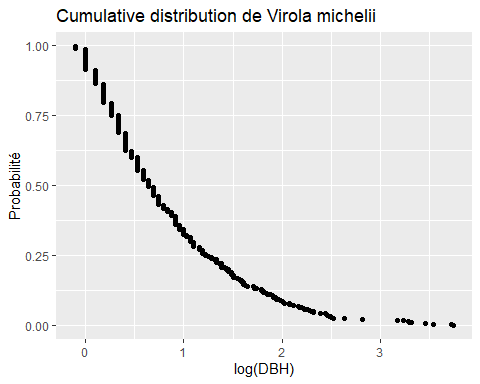
##   
## [[8]]



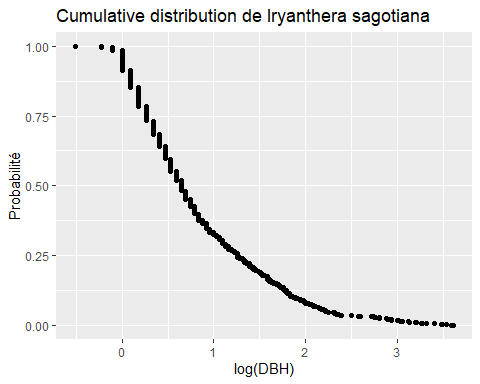
##   
## [[9]]



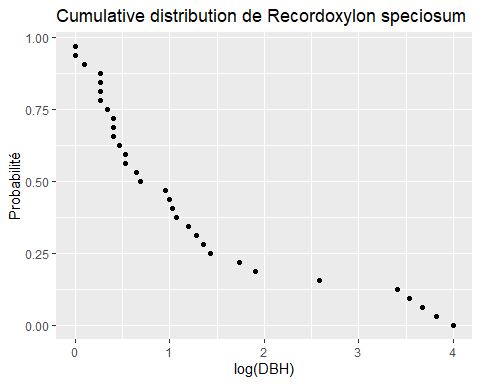
##   
## [[10]]



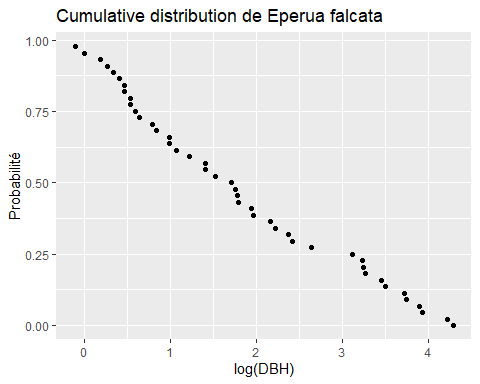
##   
## [[11]]



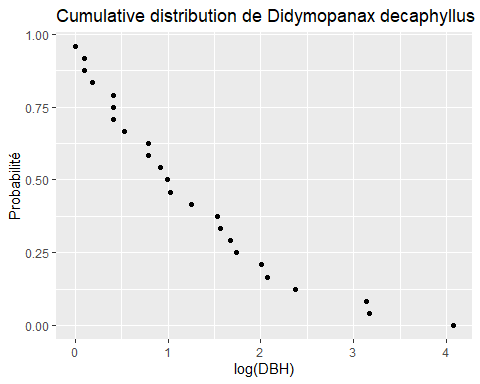
##   
## [[12]]



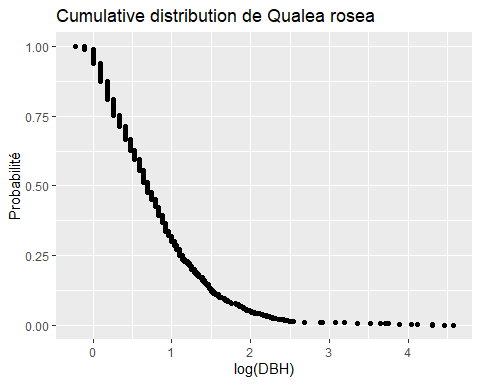
##   
## [[13]]



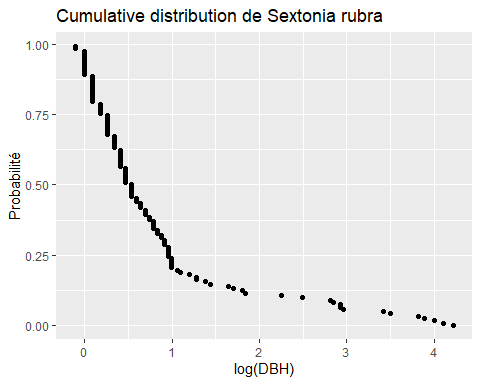
##   
## [[14]]



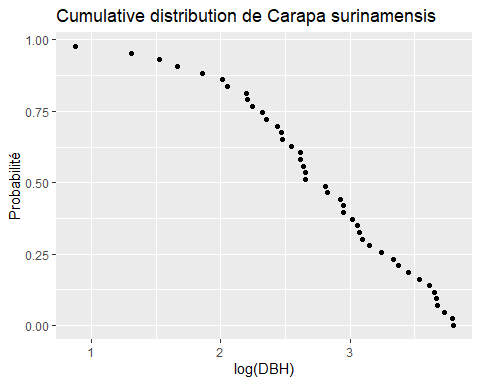
##   
## [[15]]



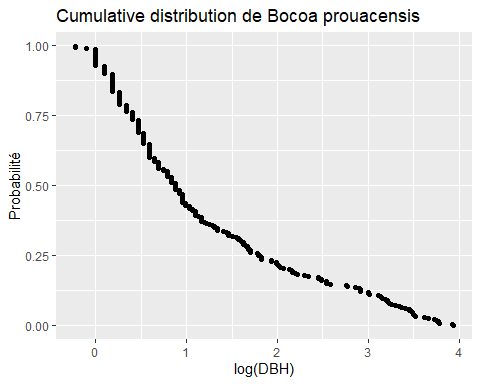
##   
## [[16]]



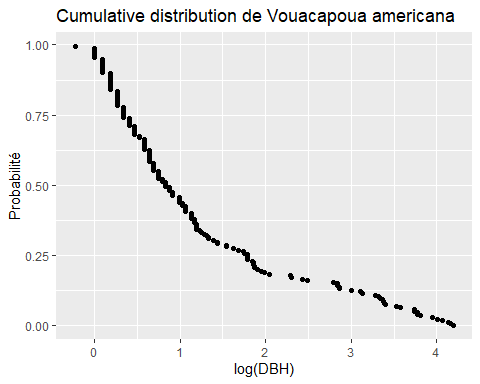
##   
## [[17]]



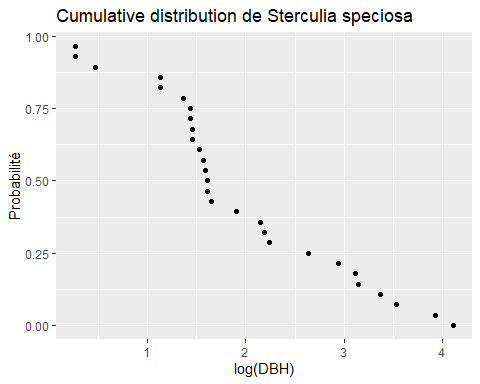
##   
## [[18]]



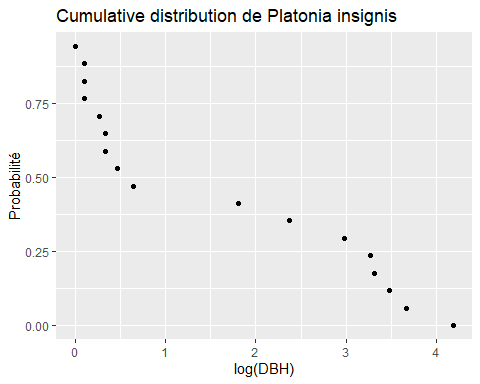
##   
## [[19]]



##   
## [[20]]



##   
## [[21]]

 Maintenant que l’on dispose du tracé de ces courbes nous allons réaliser une comparaison de cou

D’après la comparaison des espèces que nous avais données Molino nous avons les espèces :

#Présence d’adulte Nous aimerions faire un tableau qui indique si oui ou non, les espèces présentent des individus “adultes”. Dans le cas où c’est non nous irons voir au niveau de leur autécologie ce qu’il est dit.