```r
#qusetion 2

# part a -----------------

my_data <- read.table("forest.txt",header = T)
plot(my_data)

#part b -----------------

my_N <- 1138
my_n <- 20              #get n and N from qusetion
c <- 1.96
y_err <- sd(my_data$Age) ^ 2
y_bar <- mean(my_data$Age)
y_L <- y_bar - c*sqrt((1-my_n/my_N) * y_err / my_n)
y_U <- y_bar + c*sqrt((1-my_n/my_N) * y_err / my_n)
y_L                                                  # 94.94916
y_U                                                  # 119.8508

#part c -----------------

x_mu <- 10.3            #from qusetion
x_bar <- mean(my_data$Diameter)
theta <- y_bar / x_bar
mu_rat <- theta * x_mu
rat_err <- 1 / (my_n - 1) * sum((my_data$Age - theta * my_data$Diameter)^2)
rat_L <- mu_rat - c*sqrt((1-my_n/my_N) * rat_err / my_n)
rat_U <- mu_rat + c*sqrt((1-my_n/my_N) * rat_err / my_n)
rat_L                                                # 109.8262
rat_U                                                # 125.4147

# It justified because from the graph we could find that there is a linear relationship
# between x and y, and both of them are random variables and countinus, and:
moe_sps <- 1.96* sqrt((1 - my_n/my_N) * y_err/my_n)
moe_rat <- 1.96* sqrt((1 - my_n/my_N) * rat_err/my_n)
moe_rat               # 7.794237
moe_sps               # 12.45084
# moe of ratio is better than sps, yes, this is justified

# part d -----------------

M <- lm(my_data$Age ~ my_data$Diameter)
M <- M$coefficients[2]
mu_reg <- y_bar + M * (x_mu - x_bar)
reg_err <- 1 / (my_n - 1) * sum(my_data$Age - y_bar - M * (my_data$Diameter - x_bar)^2)
reg_L <- mu_reg - c*sqrt((1-my_n/my_N) * -reg_err / my_n)
reg_U <- mu_reg + c*sqrt((1-my_n/my_N) * -reg_err / my_n)
reg_L                                                # 115.5829
reg_U                                                # 121.144
# It justified because from the graph we could find that there is a linear relationship
# between x and y, and both of them are random variables and countinus, and:
moe_sps <- 1.96* sqrt((1 - my_n/my_N) * y_err/my_n)
moe_reg <- 1.96* sqrt((1 - my_n/my_N) * -reg_err/my_n)
moe_reg               # 2.78056
moe_sps               # 12.45084
# moe of regression is better than sps, yes, this is justified.

# part e -----------------

theta_err <- 1/ x_bar^2 * (1 - my_n / 2 ) * rat_err / my_n
theta_L <- theta - c*sqrt((1-my_n/my_N) * -theta_err / my_n)
theta_U <- theta + c*sqrt((1-my_n/my_N) * -theta_err / my_n)
theta_L                                              # 10.86353
theta_U                                              # 11.97539

# part f -----------------

cou <- my_data$Age / my_data$Diameter
cou_mu <- mean(cou)
cou_err <- sd(cou)
cou_L <- theta - c*sqrt((1-my_n/my_N) * cou_err / my_n)
cou_U <- theta + c*sqrt((1-my_n/my_N) * cou_err / my_n)
cou_L                                                # 10.83303
cou_U                                                # 12.00589
```