Assignment1

Zimeng Ming V00844078

2019/1/22

Question2:

a):

```
Question1_data<-read.table("Question1.txt",header = T)

Question1_data
```

```
##    growth trt plot pot
## 1    14.6   1    1   1
## 2    15.2   1    1   1
## 3    13.2   1    1   2
## 4    12.9   1    1   2
## 5    16.4   1    1   3
## 6    12.2   1    1   3
## 7     7.1   2    2   4
## 8     7.7   2    2   4
## 9     6.8   2    2   5
## 10    6.0   2    2   5
## 11   10.0   2    2   6
## 12    8.3   2    2   6
## 13   18.5   1    3   7
## 14   16.7   1    3   7
## 15   22.2   1    3   8
## 16   18.8   1    3   8
## 17   24.7   1    3   9
## 18   20.3   1    3   9
## 19    9.7   2    4  10
## 20    8.8   2    4  10
## 21    6.8   2    4  11
## 22    9.0   2    4  11
## 23   10.4   2    4  12
## 24   11.3   2    4  12
```

b): Sort the data by plant growth

```
attach(Question1_data)

Question1_data_sorted <- cbind(growth[order(Question1_data$growth)],trt[order(Questio
n1_data$growth)],plot[order(Question1_data$growth)],pot[order(Question1_data$growth)]
)

Question1_data_sorted
```

```
##       [,1] [,2] [,3] [,4]
##  [1,]  6.0    2    2    5
##  [2,]  6.8    2    2    5
##  [3,]  6.8    2    4   11
##  [4,]  7.1    2    2    4
##  [5,]  7.7    2    2    4
##  [6,]  8.3    2    2    6
##  [7,]  8.8    2    4   10
##  [8,]  9.0    2    4   11
##  [9,]  9.7    2    4   10
## [10,] 10.0    2    2    6
## [11,] 10.4    2    4   12
## [12,] 11.3    2    4   12
## [13,] 12.2    1    1    3
## [14,] 12.9    1    1    2
## [15,] 13.2    1    1    2
## [16,] 14.6    1    1    1
## [17,] 15.2    1    1    1
## [18,] 16.4    1    1    3
## [19,] 16.7    1    3    7
## [20,] 18.5    1    3    7
## [21,] 18.8    1    3    8
## [22,] 20.3    1    3    9
## [23,] 22.2    1    3    8
## [24,] 24.7    1    3    9
```

c)Mean and deviation:

```
print("The mean for the growth is: ")
```

```
## [1] "The mean for the growth is: "
```

```
mean(growth)
```

```
## [1] 12.81667
```

```
print("The Standard Deviation of the sample is :")
```

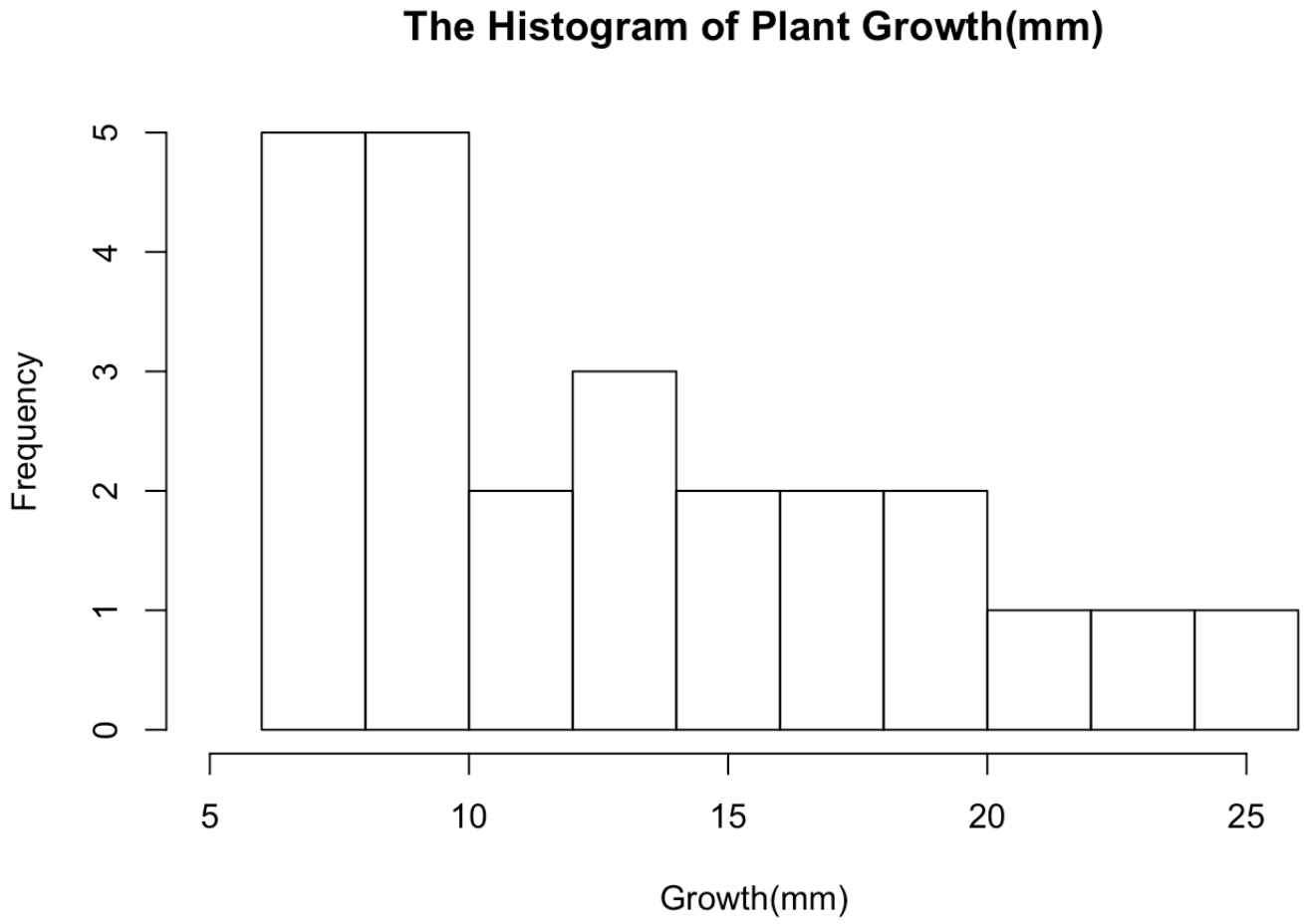
```
## [1] "The Standard Deviation of the sample is :"
```

```
sqrt(var(growth))
```

```
## [1] 5.296813
```

d. Plot the data

```
hist(growth,main="The Histogram of Plant Growth(mm)",xlab = "Growth(mm)",breaks = 10,
xlim = c(5,26) )
```



Question3: The sum of the squares of a vector

```
sum_of_squares<-function(y){

  y_square=0
  #using for loop to calculate the sum square of (Y[i]-mean)
  for(i in 1:length(y)){
    y_square=(y[i])^2+y_square
  }
  #using the sqyares formula.
  y_square-(length(y)*((sum(y)/length(y)))^2)

}

#Test the data
y<-c(11,11,10,8,11,3,15,11,7,6)
sum_of_squares(y)
```

```
## [1] 102.1
```

However, for this question, I have different answers for the code using to calculate the variance. sum_of_squares<-function(y){ y_square=0 #using for loop to calculate the sum square of (Y[i]-mean) for(i in 1:length(y)){ y_square=(y[i]-mean(y))^2+y_square } #using the sqyares formula. y_square/(length(y)-1) } #Test the data y<-c(11,11,10,8,11,3,15,11,7,6) sum_of_squares(y) ## [1] 11.34444

But I do not know what is wrong with my code. I could not find a way that using the formula of sum of sequares that can get accurate variance.

Question4:

```
power.t.test(delta = 6,sd=3,sig.level = 0.05,power = 0.8,type = "one.sample")
```

```
##
##      One-sample t test power calculation
##
##               n = 4.220731
##           delta = 6
##              sd = 3
##       sig.level = 0.05
##           power = 0.8
##     alternative = two.sided
```

so the sample size required is 5.

Question5:

```
#for the function, we use n=z^2*s^2/d^2 where z is confidence interval, s is standard deviation and d is margin of error.

Population_via_CI<-function(CI, margin_of_error,standard_deviation){
  #set the value for Confidence Interval
  z<-qt(CI,df = Inf)

  #set margin of error
  d<-margin_of_error

  #set the standard deviation
  s<-standard_deviation

  #calculate the population
  ((z/d)^2)*(s^2)


}
```

Question6:

```
Population_via_CI(0.95,2,3)
```

```
## [1] 6.087473
```

So the answer is n=7.