

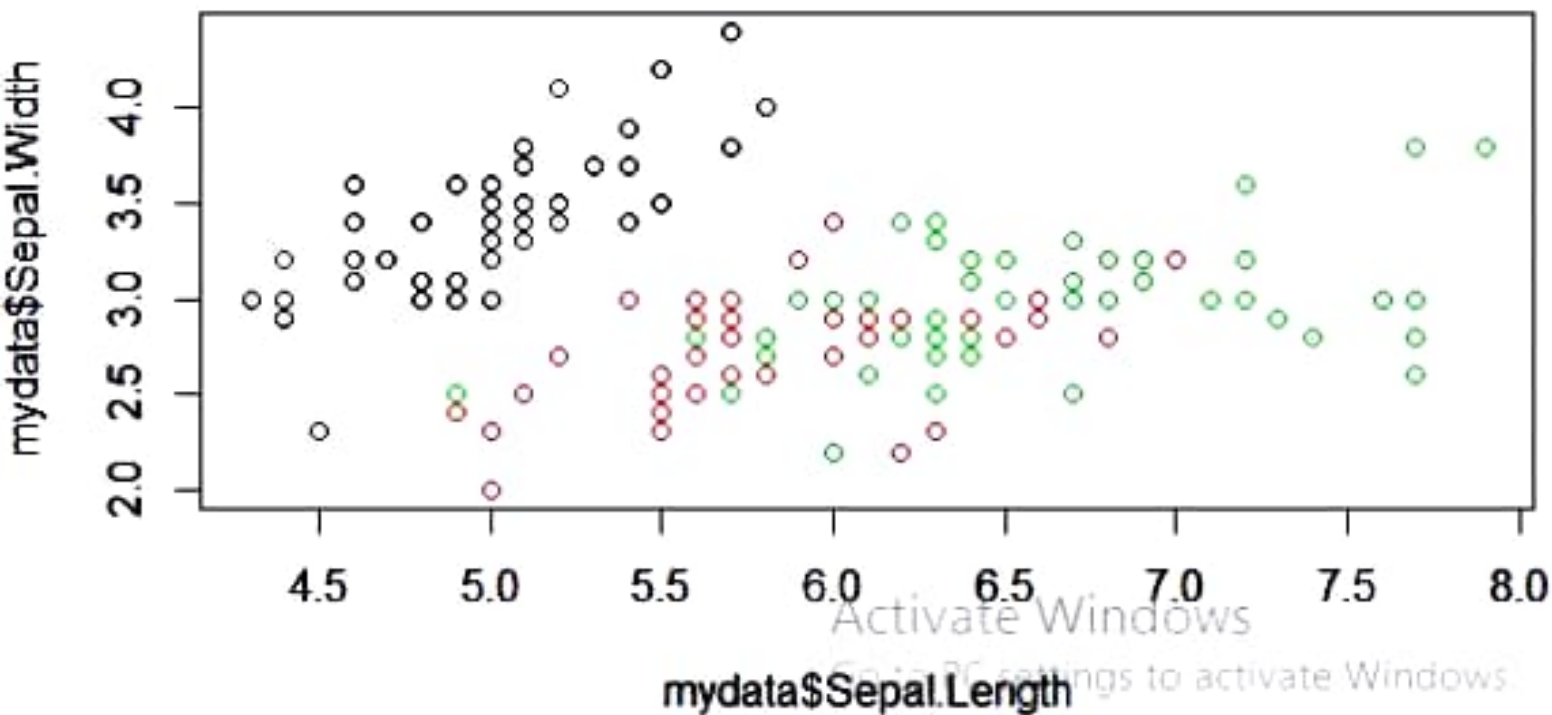
```
else { $conn -> error; }
```

```
?>
```

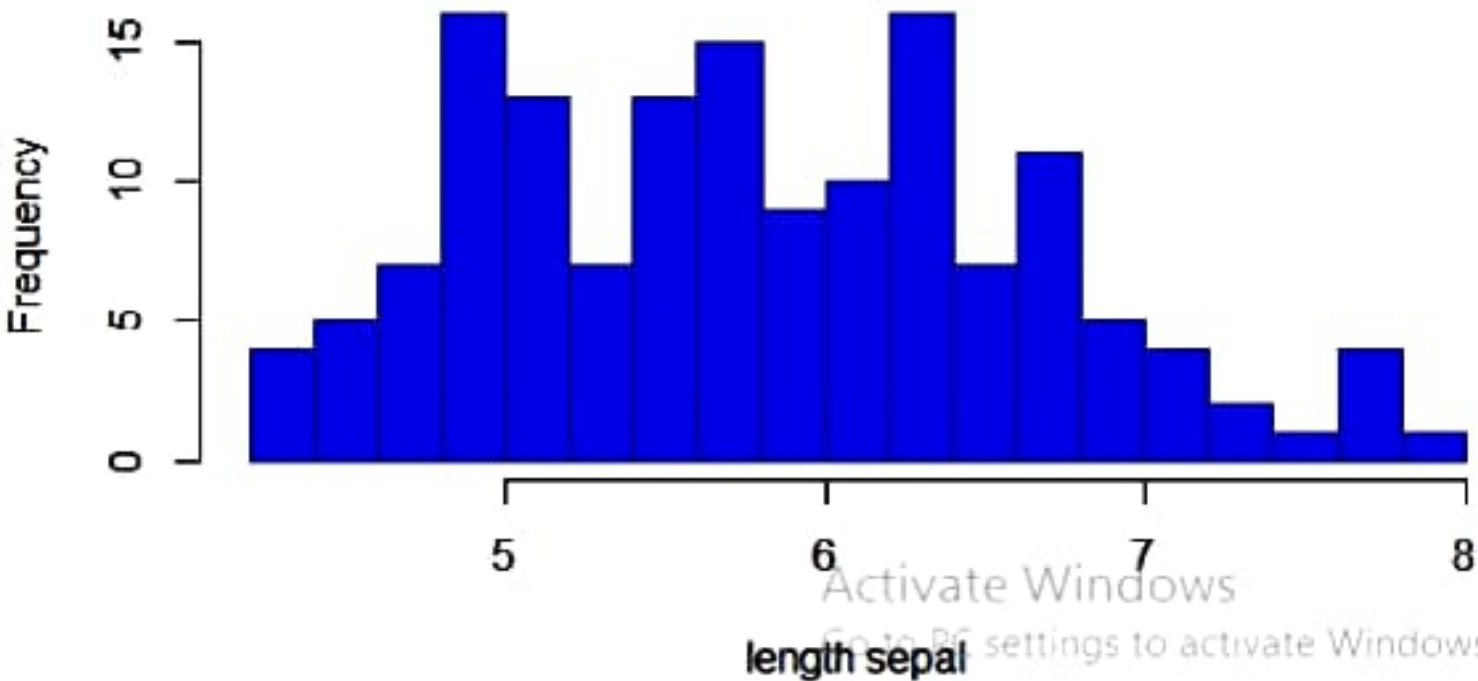
R.

Ques 3. library(dplyr)
setwd("D:/rp")
mydata <- ~~read.csv~~ iris
dim(mydata)
summarize(mydata)
names(mydata)
head(mydata, 5)
tail(mydata, 5)
mysubdata <- select(mydata, Sepal.length, Species)
mysubdata
describe(mysubdata)
plot(mydata\$Sepal.length, mydata\$Sepal.width, col = mydata\$Species)
hist(mydata\$Sepal.length, xlab = "Sepal length", col = "blue", border = "red",
breaks = 20)

```
~/
> summary(mydata)
> dim(mydata)
[1] 150 5
> summary(mydata)
  Sepal.Length  Sepal.width  Petal.Length  Petal.width  Species
Min.   :4.300   Min.   :2.000   Min.   :1.000   Min.   :0.100   setosa    :50
1st Qu.:5.100   1st Qu.:2.800   1st Qu.:1.600   1st Qu.:0.300   versicolor:50
Median :5.800   Median :3.000   Median :4.350   Median :1.300   virginica :50
Mean   :5.843   Mean   :3.057   Mean   :3.758   Mean   :1.199
3rd Qu.:6.400   3rd Qu.:3.300   3rd Qu.:5.100   3rd Qu.:1.800
Max.   :7.900   Max.   :4.400   Max.   :6.900   Max.   :2.500
> names(mydata)
> head(mydata,5)
  Sepal.Length Sepal.width Petal.Length Petal.width Species
1           5.1         3.5         1.4         0.2   setosa
2           4.9         3.0         1.4         0.2   setosa
3           4.7         3.2         1.3         0.2   setosa
4           4.6         3.1         1.5         0.2   setosa
5           5.0         3.6         1.4         0.2   setosa
> tail(mydata,5)
  Sepal.Length Sepal.width Petal.Length Petal.width Species
146          6.7         3.0         5.2         2.3 virginica
147          6.3         2.5         5.0         1.9 virginica
148          6.5         3.0         5.2         2.0 virginica
149          6.2         3.4         5.4         2.3 virginica
150          5.9         3.0         5.1         1.8 virginica
> describe(mydata)
```



Histogram of mydata\$Sepal.Length



Ques 1. ~~summary (mydata)~~
~~min (mydata)~~ \$
~~max~~

Ques 4. summary (mydata)
min (mydata \$ Sepal.length)
max (mydata \$ Sepal.length)
mean (mydata \$ Sepal.length)
sd (mydata \$ Sepal.length) // for standard deviation
var (mydata \$ Sepal.length) // for variance
cv = sd (mydata \$ Sepal.length) / mean (mydata \$ Sepal.length)
// for coefficient of variation

~~barplot (table (mydata \$ size))~~

barplot (mydata \$ Sepal.length , xlab = "species" , ylab = "Sepal.length" ,
names.arg = mydata \$ species)

plot (mydata \$ Sepal.length , mydata \$ species) // scatter plot
ggplot (mydata) + aes (x = Sepal.length) + geom_histogram ()

Inferential

for z test

```
> z.test = function ( a , mu , var ) {  
  zeta = ( mean ( a ) - mu ) / ( sqrt ( var / length ( a ) ) )  
  return ( zeta )  
}
```

~~a~~ mydata\$Sepal.length

$z \leftarrow z.test(a, mean(mydata$Sepal.Length), var(mydata$Sepal.Length))$

$p\text{-value} \leftarrow 2 * pnorm(-abs(z))$

~~f~~-test

$aoV(formula = Sepal.length \sim Species, data = mydata)$

```
> aov(formula=Sepal.Length~Species,data=iris)
```

```
Call:
```

```
  aov(formula = Sepal.Length ~ Species, data = iris)
```

```
Terms:
```

	Species	Residuals
Sum of Squares	63.21213	38.95620
Deg. of Freedom	2	147

```
Residual standard error: 0.5147894
```

```
Estimated effects may be unbalanced
```

```
> |
```

```
> z.test1=function(a,mu,var){  
+   zeta=(mean(a)-mu)/(sqrt(var/length(a)))  
+   return(zeta)  
+ }  
> a<-mydata$Sepal.Length  
> z<-z.test1(a,mean(mydata$Sepal.Length),var(mydata$Sepal.Length))  
> p_value<-2*pnorm(-abs(z))  
> p_value  
[1] 1  
> |
```