Name: Alpanshu Kataria

Reg. No.: 18BCE1267

Subject: Data Visualization

Lab 1 (03/02/2021)

Code for normal plots:

```
rm(list=ls())


library(MASS)

newsurvey<- na.omit(survey)


#1.  Plot a bar graph for the number of male and female participants in the survey. Provide the title as "Male and Female participants" and specify the colours for the bars.

gender<- newsurvey$Sex

gender_freq<- table(gender)

barplot(gender_freq,main="Male and Female participants",ylab="frequency",col=c('red',"blue"))


#2.  Plot a bar graph for the number of left handers and right handers in the survey. Provide the title as "Left Handers and Right Hnaders" and specify the colours for the bars.

write_hand<- newsurvey$W.Hnd

write_hand_freq<- table(write_hand)

barplot(write_hand_freq,main="Left Handers and Right Handers",ylab="count",col=c('red',"blue"))
```

#3. Plot the distribution between male left handers and female left handers using bar chart. Provide the title as "Female Left Handers and Male Left Handers" and specify the colours for the bars.

```
left_hand=write_hand=="Left"

left_hand_data=newsurvey[left_hand,]

g <- left_hand_data$Sex

g_freq<- table(g)

barplot(g_freq,main="Female Left Handers and Male Left Handers",ylab="count",col=c('cyan',"green"))
```

#4. Draw the distribution of smoking habits of male left handers using pie chart.

```
male_left_hand<- left_hand_data$Sex=="Male"

male_left_hand_data<- left_hand_data[male_left_hand,]

smk<- male_left_hand_data$Smoke

smk_freq<- table(smk)

pie(smk_freq,main="Distribution of smoking habits of male left handers",col=c('red','magenta','cyan',"green"))
```

#5. Draw the histogram of age distribution with the title as 'Age distribution' and xlabel as 'Age range' and ylabel as 'frequency'.

```
age_rg<- newsurvey$Age

hist(age_rg,main='Age distribution',right=FALSE,xlab='Age range',ylab='frequency')
```

#6. Reveal the relationship between writing hand span and the age using scatter plot.

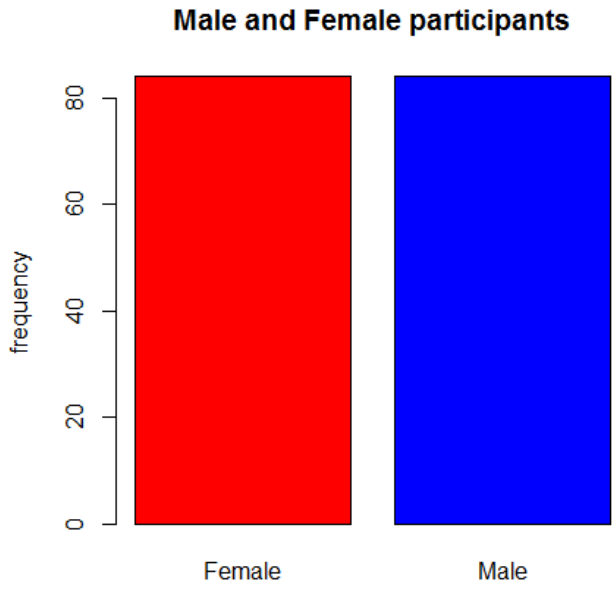
```
write_hand_span<- newsurvey$Wr.Hnd

plot(write_hand_span,age_rg,main="Relationship between writing hand span and the age")
```

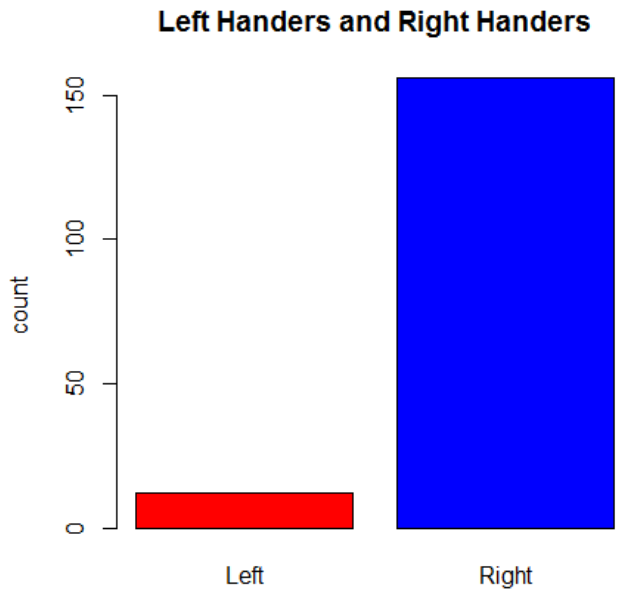
#7. Draw the boxplot for pulse rate and analyse the five summary statistics.

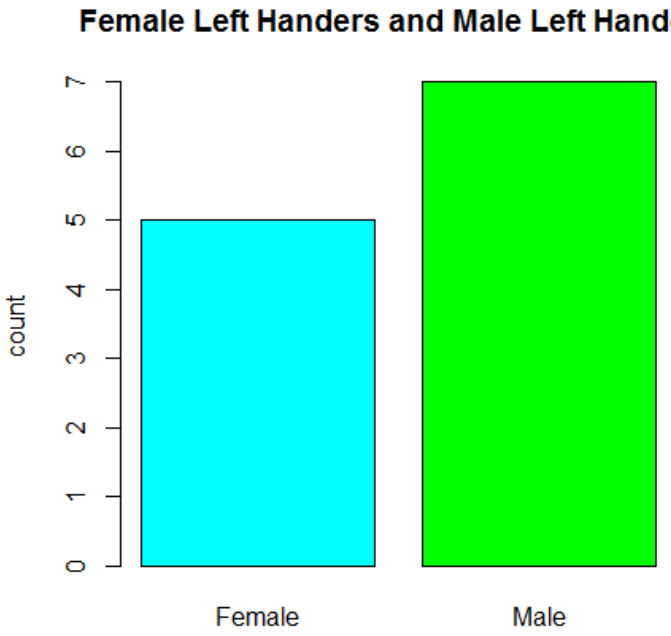
```
pr<- newsurvey$Pulse
```

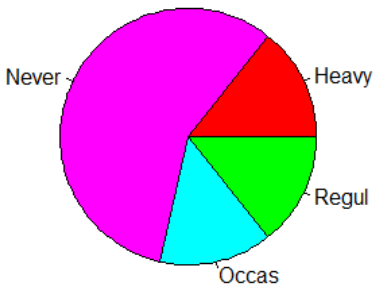
```
boxplot(pr,main="Box plot for Pulse rate",ylab="Pulse rate")
```

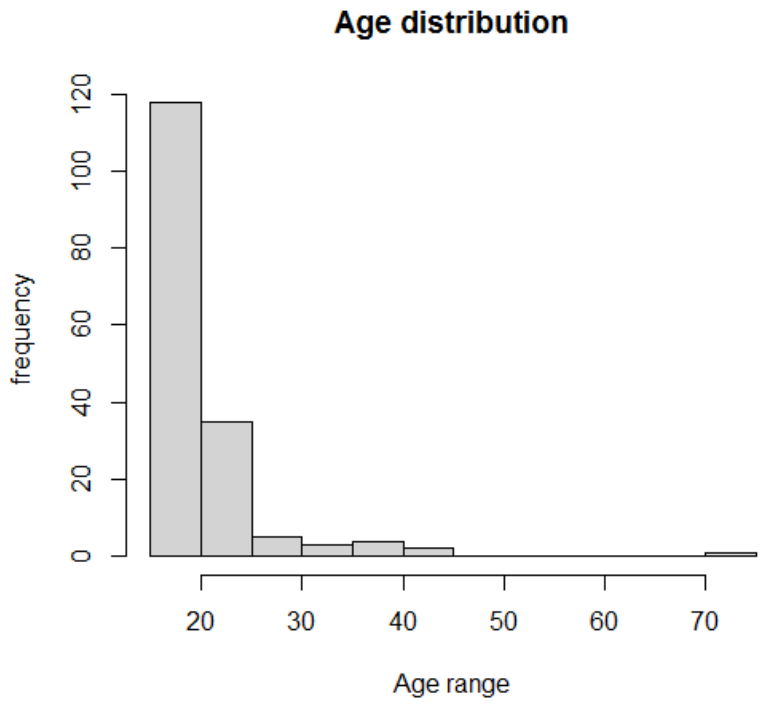
Outputs:

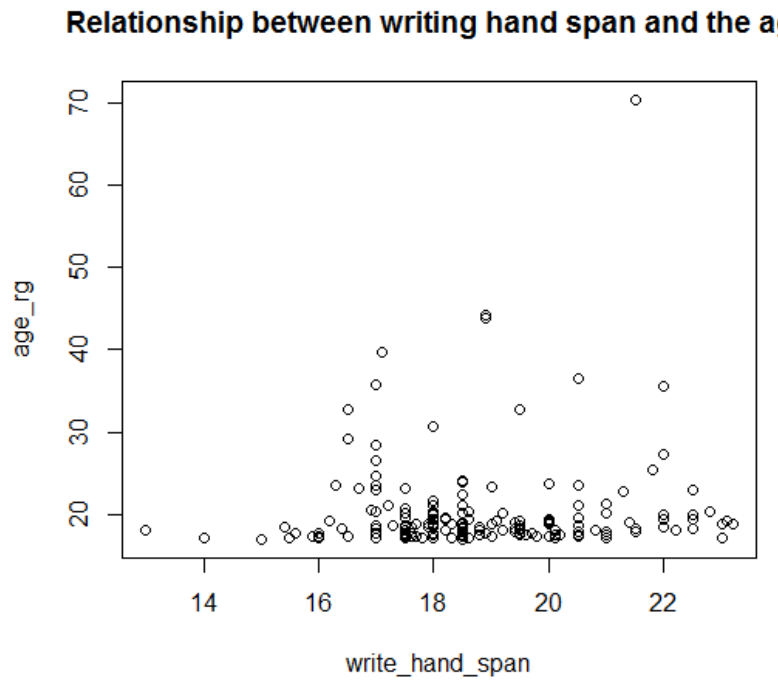


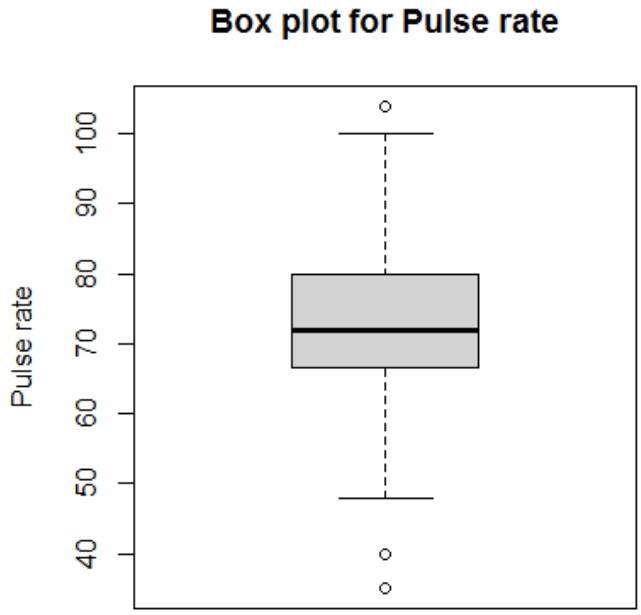



Female Left Handers and Male Left Handers
count
0
1
2
3
4
5
6
7
Female
Male


Distribution of smoking habits of male left handers
Never
Heavy
Regul
Occas


Age distribution
frequency
0
20
40
60
80
100
120
20
30
40
50
60
70
Age range


Relationship between writing hand span and the age
age_rg
20
30
40
50
60
70
14
16
18
20
22
write_hand_span



Code for ggplot2 plots:

```
rm(list=ls())


library(MASS)

library(ggplot2)

newsurvey<- na.omit(survey)


#1. Plot a bar graph for the number of male and female participants in the survey. Provide the title as "Male and Female participants" and specify the colours for the bars.

gender<- newsurvey$Sex

gender_freq<- table(gender)

ggplot(data=newsurvey,mapping=aes(x=Sex,fill=Sex))+geom_bar()+ggtitle("Male and Female participants")
```

#2. Plot a bar graph for the number of left handers and right handers in the survey. Provide the title as "Left Handers and Right Hnaders" and specify the colours for the bars.

```
write_hand<- newsurvey$W.Hnd

ggplot(data=newsurvey,mapping=aes(x=W.Hnd,fill=W.Hnd))+geom_bar()+scale_fill_manual(values = c("red", "green") )+ggtitle("Left Handers and Right Hnaders")
```

#3. Plot the distribution between male left handers and female left handers using bar chart. Provide the title as "Female Left Handers and Male Left Handers" and specify the colours for the bars.

```
left_hand=write_hand=="Left"

left_hand_data=newsurvey[left_hand,]

ggplot(data=left_hand_data,mapping=aes(x=Sex,fill=Sex))+geom_bar()+scale_fill_manual(values = c("red", "green") )+ggtitle("Female Left Handers and Male Left Handers")
```

#4. Draw the distribution of smoking habits of male left handers using pie chart.

```
male_left_hand<- left_hand_data$Sex=="Male"

male_left_hand_data<- left_hand_data[male_left_hand,]

ggplot(data=male_left_hand_data,aes(x="",y=Sex,fill=Smoke))+geom_bar(stat="identity", width=1, color="White")+coord_polar("y", start=0)+ggtitle("smoking habits of male left handers")+theme_void()
```

#5. Draw the histogram of age distribution with the title as 'Age distribution' and xlabel as 'Age range' and ylabel as 'frequency'.

```
ggplot(newsurvey, aes(x=Age)) + geom_histogram(binwidth=5, fill="#69b3a2", color="#e9ecef", alpha=0.9)+labs(y= "Frequency", x = "Age Range")+ggtitle("Age distribution")
```

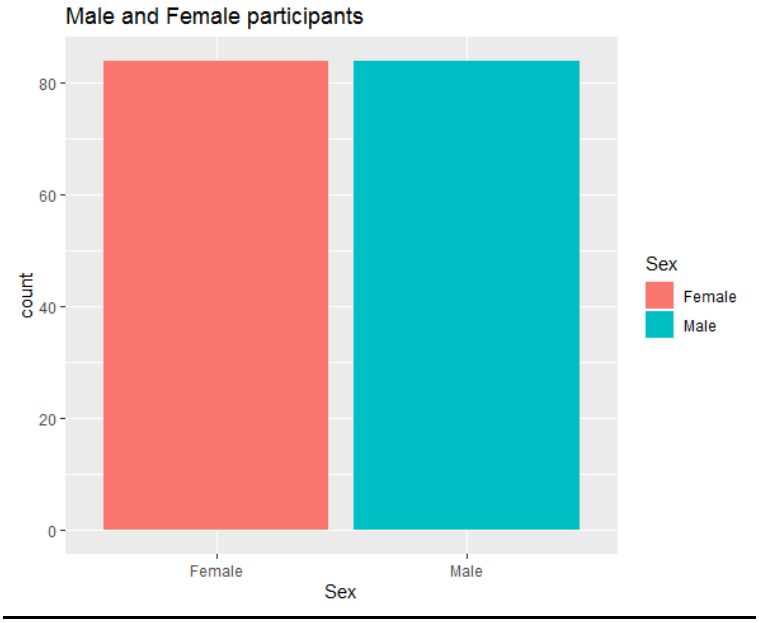
#6. Reveal the relationship between writing hand span and the age using scatter plot.

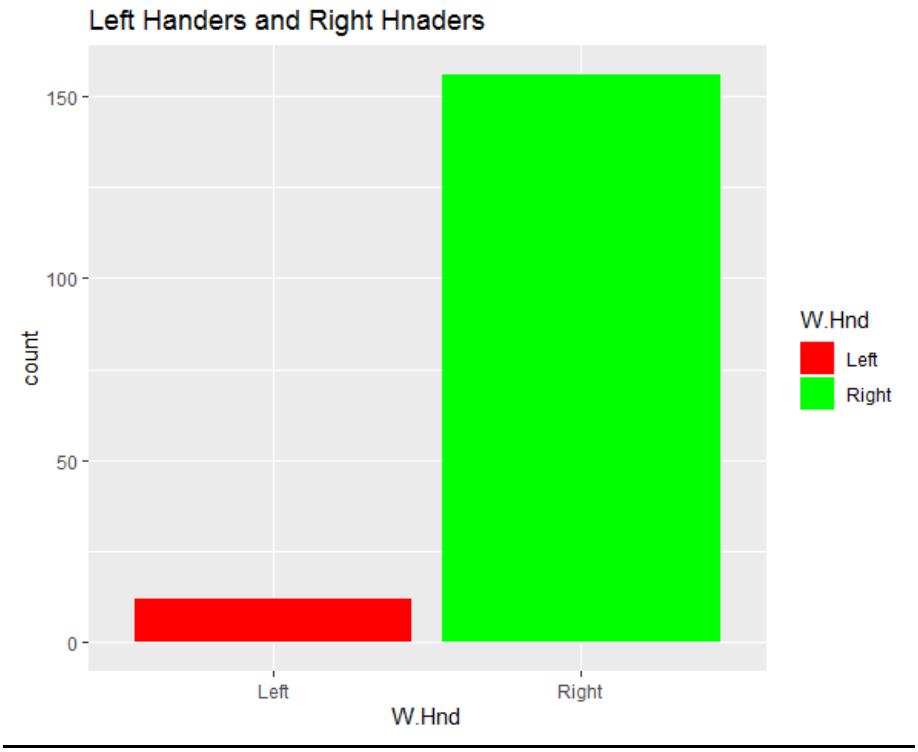
```
write_hand_span<- newsurvey$Wr.Hnd

ggplot(newsurvey, aes(x=Wr.Hnd, y=Age)) + geom_point()
```

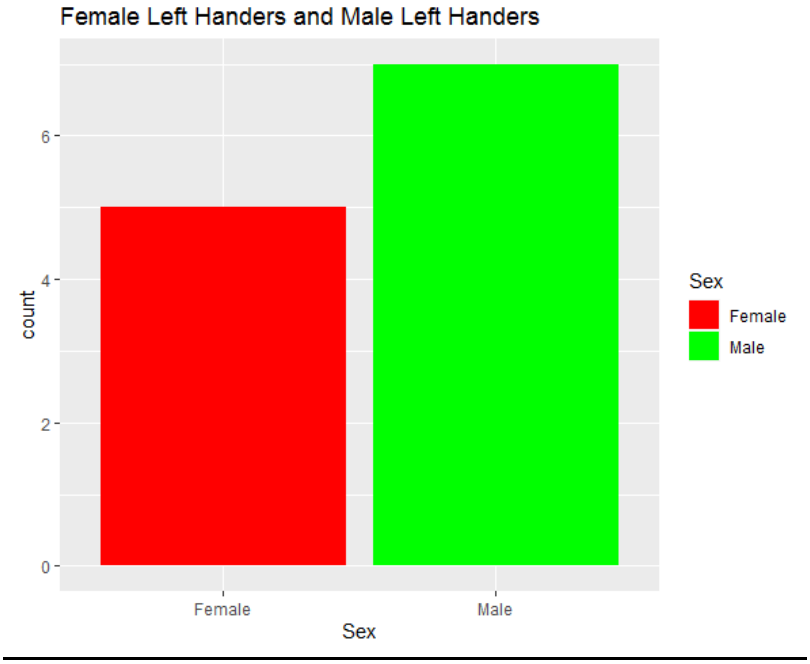
#7. Draw the boxplot for pulse rate and analyse the five summary statistics.

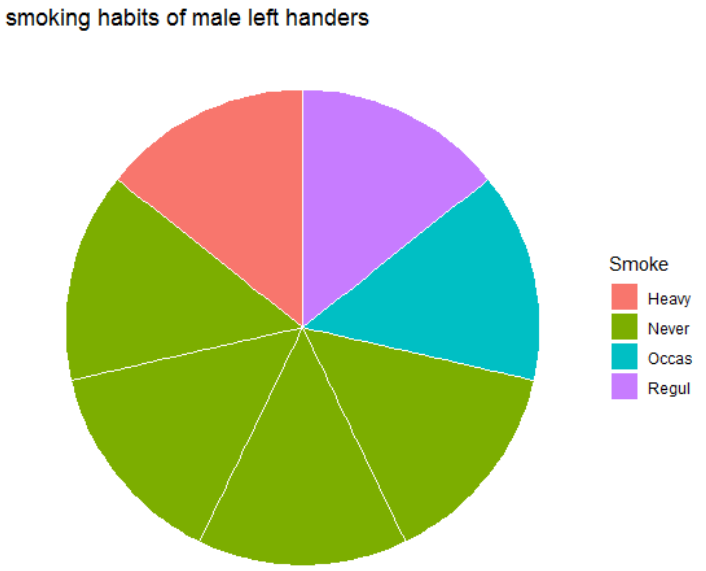
```
pr<- newsurvey$Pulse

ggplot(newsurvey, aes(y=Pulse, fill=Pulse)) + geom_boxplot(fill="slateblue", alpha=0.2) + ylab("Pulse Rate")
```

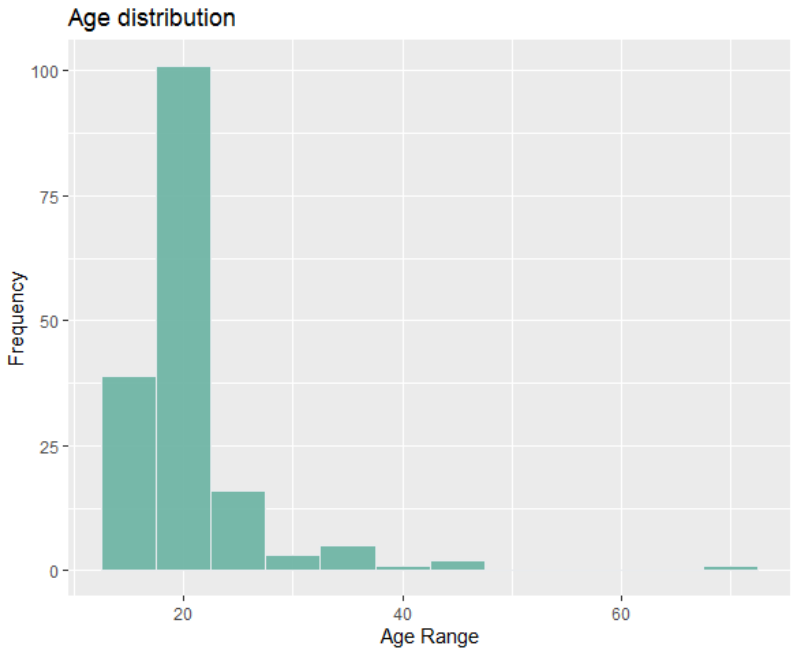
Output:

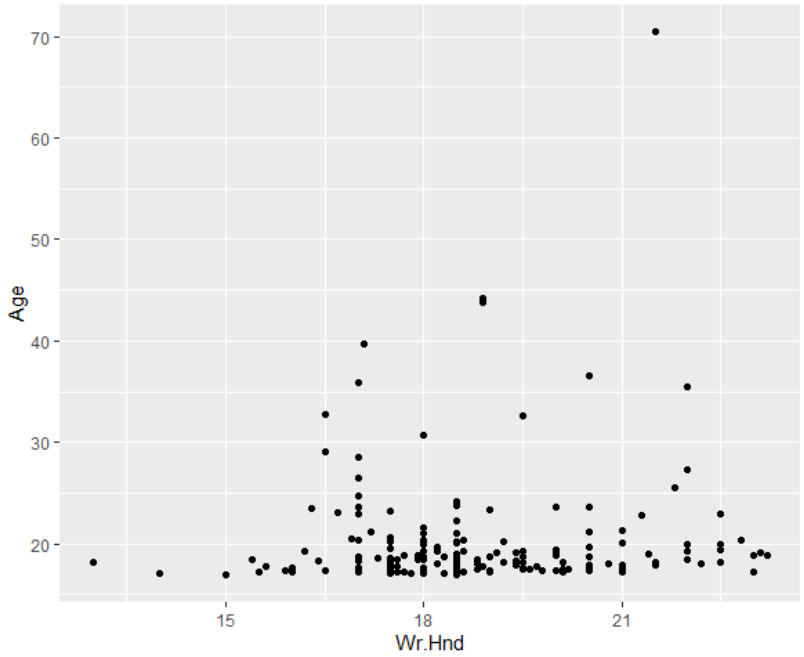


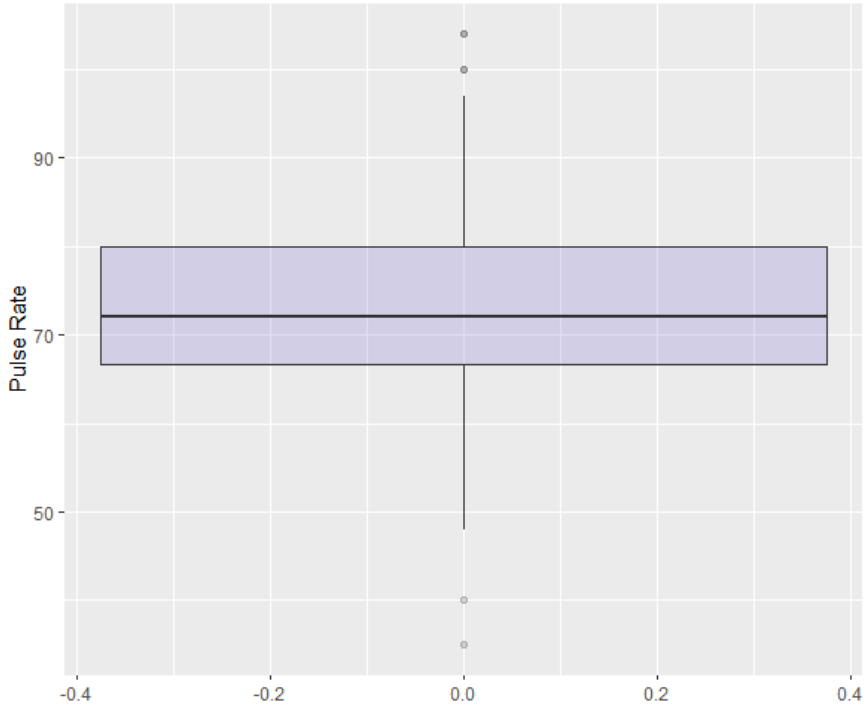

Left Handers and Right Hnaders
150
100
50
0
count
Left
Right
W.Hnd
W.Hnd
Left
Right


Female Left Handers and Male Left Handers
6
4
2
0
count
Female
Male
Sex
Sex
Female
Male


smoking habits of male left handers
Smoke
Heavy
Never
Occas
Regul


Age distribution
100
75
50
25
0
Frequency
20
40
60
Age Range


70
60
50
40
30
20
Age
15
18
21
Wr.Hnd

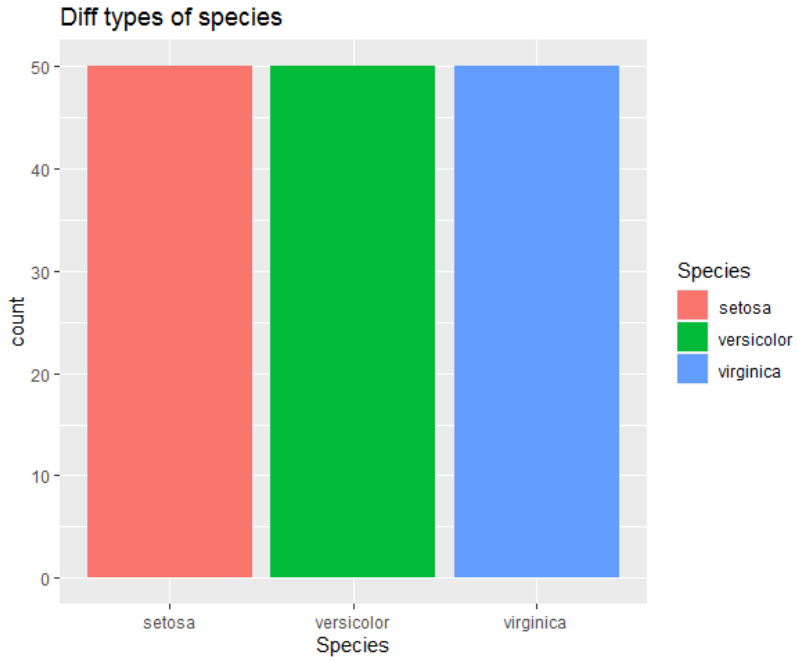

90
70
50
Pulse Rate
-0.4
-0.2
0.0
0.2
0.4

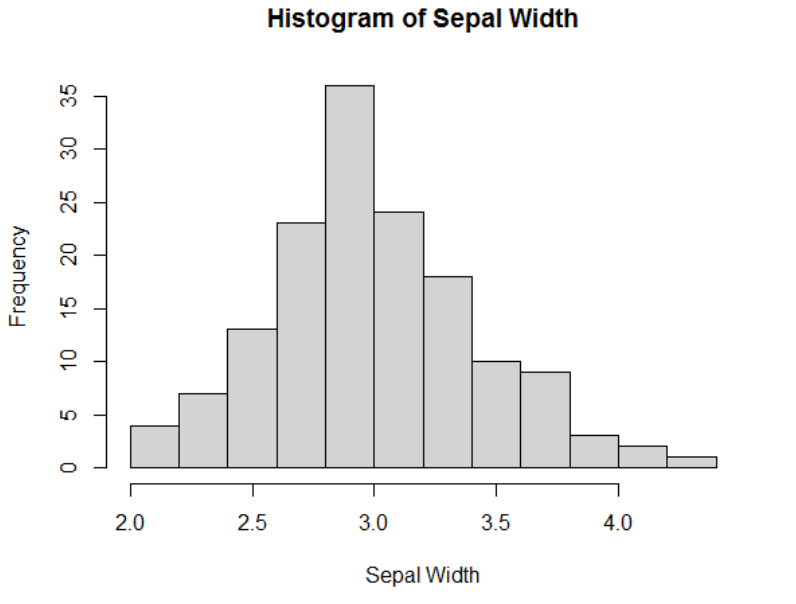
Code for plotting a different dataset(Iris dataset):

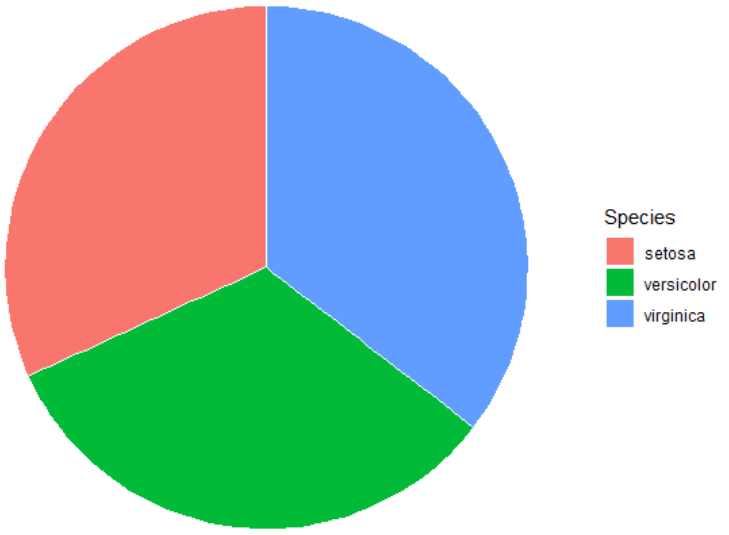
```
rm(list=ls())


library(MASS)

library(ggplot2)

newdata<- na.omit(iris)


ggplot(data=newdata,mapping=aes(x=Species,fill=Species))+geom_bar()+ggtitle("Diff types of species")



hist(iris$Sepal.Width, freq=NULL, density=NULL, breaks=12,xlab="Sepal Width", ylab="Frequency", main="Histogram of Sepal Width")


set.seed(1234)

iris1 <- iris[sample(1:nrow(iris), 110), ]

pie <- ggplot(iris1, aes(x=factor(1), fill=Species)) +geom_bar( width=1, color="White") + labs(x="", y="")+theme_void()

pie + coord_polar(theta="y")


scatter <- ggplot(data=iris, aes(x = Sepal.Length, y = Sepal.Width))

scatter + geom_point(aes(color=Species, shape=Species)) +xlab("Sepal Length") +  ylab("Sepal Width") +ggtitle("Sepal Length-Width")
```

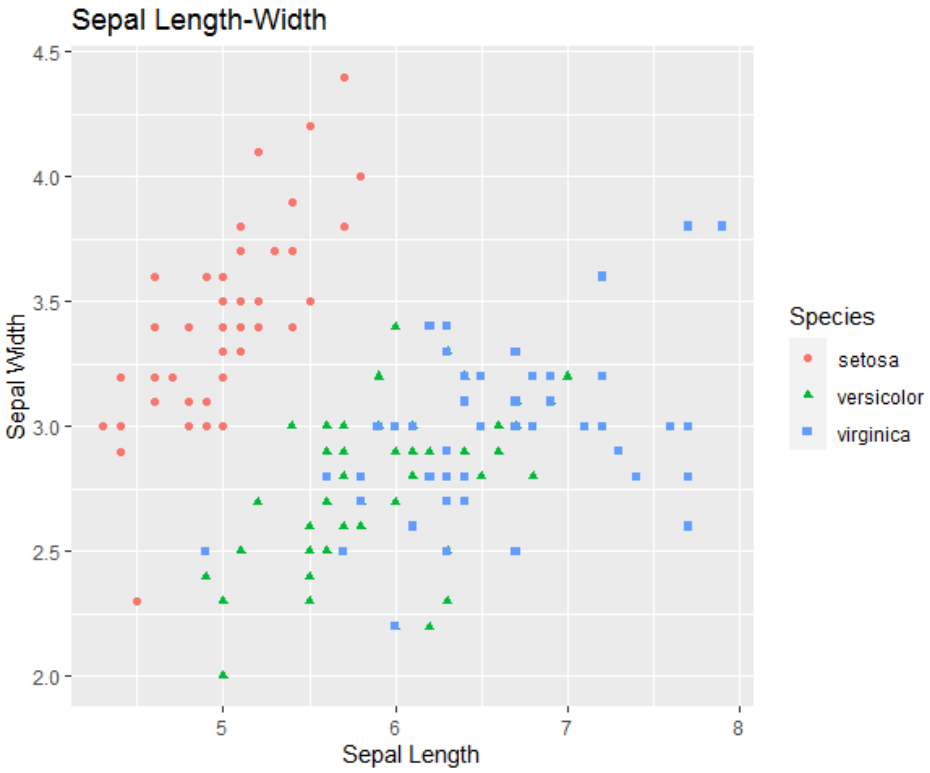
boxplot(Sepal.Length~Species,data=iris, xlab="Species", ylab="Sepal Length", main="Iris Boxplot")

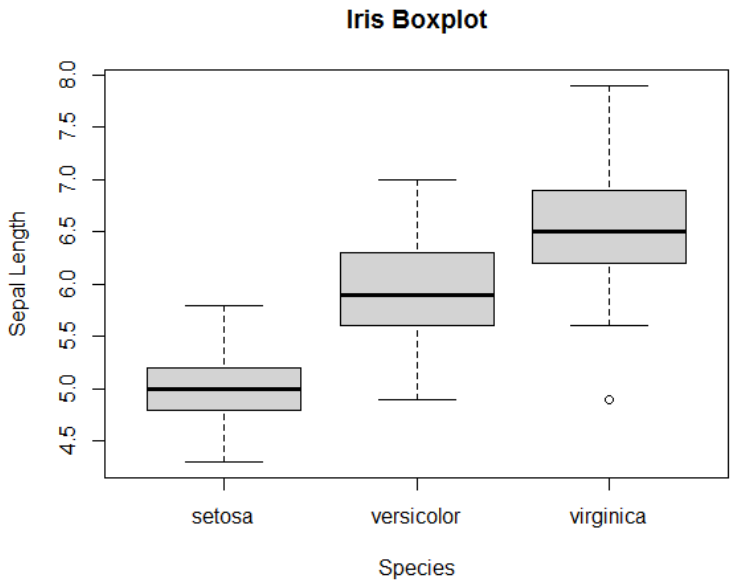
Output:






Species
setosa
versicolor
virginica


Sepal Length-Width
Sepal Width
Sepal Length
4.5
4.0
3.5
3.0
2.5
2.0
5
6
7
8
Species
setosa
versicolor
virginica


Iris Boxplot
Sepal Length
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
setosa
versicolor
virginica
Species