Lab Ex5 - Visualizing Data

Reg. No: 18BCE1357

Name: Akash Manish Lad

Setup

```
library(help = "graphics")
library(MASS)
library(surveydata)
```

```
##
## Attaching package: 'surveydata'
```

```
## The following object is masked from 'package:stats':
##
##     filter
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following object is masked from 'package:MASS':
##
##     select
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

Cleaning and Storing Data

```
df <- survey
df <- na.omit(df)
str(df)
```

```
## 'data.frame':    168 obs. of  12 variables:
##  $ Sex   : Factor w/ 2 levels "Female","Male": 1 2 2 1 2 1 2 2 1 1 ...
##  $ Wr.Hnd: num  18.5 19.5 20 18 17.7 17 20 18.5 17 19.5 ...
##  $ NW.Hnd: num  18 20.5 20 17.7 17.7 17.3 19.5 18.5 17.2 20.2 ...
##  $ W.Hnd : Factor w/ 2 levels "Left","Right": 2 1 2 2 2 2 2 2 2 2 ...
##  $ Fold  : Factor w/ 3 levels "L on R","Neither",..: 3 3 2 1 1 3 3 3 1 1 ...
##  $ Pulse : int  92 104 35 64 83 74 72 90 80 66 ...
##  $ Clap  : Factor w/ 3 levels "Left","Neither",..: 1 1 3 3 3 3 3 3 3 2 ...
##  $ Exer  : Factor w/ 3 levels "Freq","None",..: 3 2 3 3 1 1 3 3 1 3 ...
##  $ Smoke : Factor w/ 4 levels "Heavy","Never",..: 2 4 2 2 2 2 2 2 2 2 ...
##  $ Height: num  173 178 165 173 183 ...
##  $ M.I   : Factor w/ 2 levels "Imperial","Metric": 2 1 2 1 1 2 2 2 1 2 ...
##  $ Age   : num  18.2 17.6 23.7 21 18.8 ...
##  - attr(*, "na.action")= 'omit' Named int [1:69] 3 4 12 13 15 16 19 25 26 29 ...
##   ..- attr(*, "names")= chr [1:69] "3" "4" "12" "13" ...
```

Viewing data

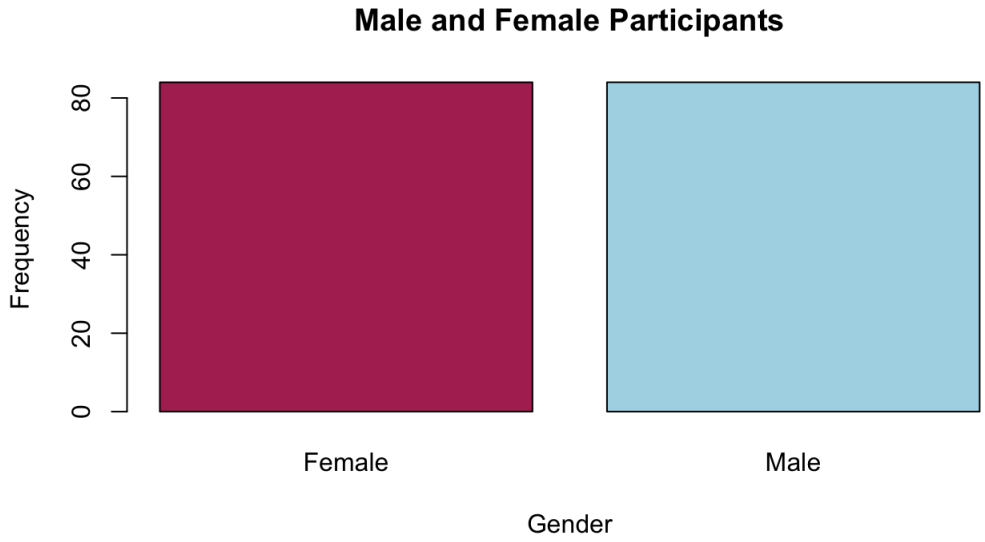
```
names(df)
```

```
##  [1] "Sex"    "Wr.Hnd" "NW.Hnd" "W.Hnd"  "Fold"   "Pulse"  "Clap"   "Exer"
##  [9] "Smoke"  "Height" "M.I"    "Age"
```

Ex-5a Visualization using basic graphics

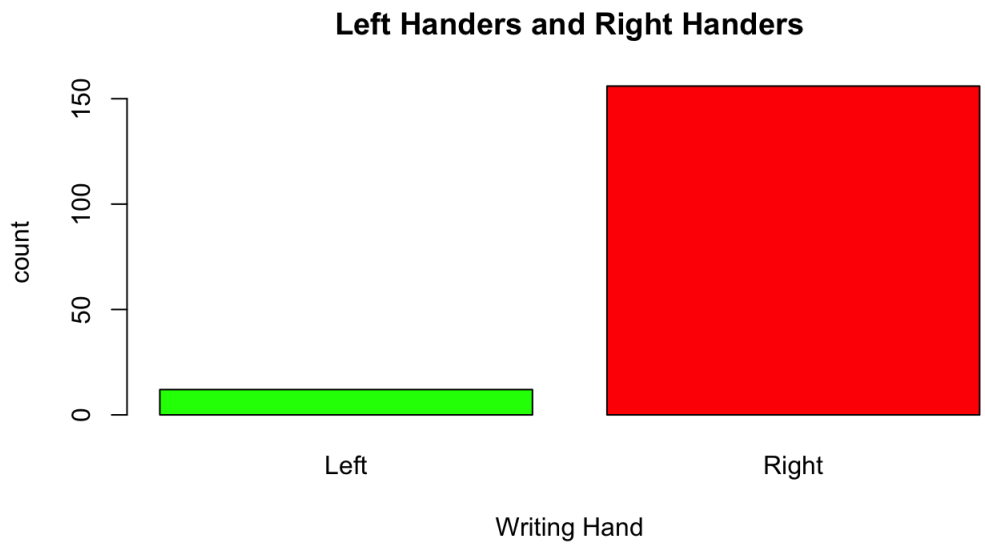
Q1) Plot a bar graph for the number of male and female participants in the survey. Provide the title as “Male and Female participants”, y-axis label as “frequency” and specify the colours for the bars.

```
table = table(df$Sex)
barplot(table, xlab = 'Gender',ylab ='Frequency' ,main = 'Male and Female Participants',
        col=c('maroon','light blue'))
```



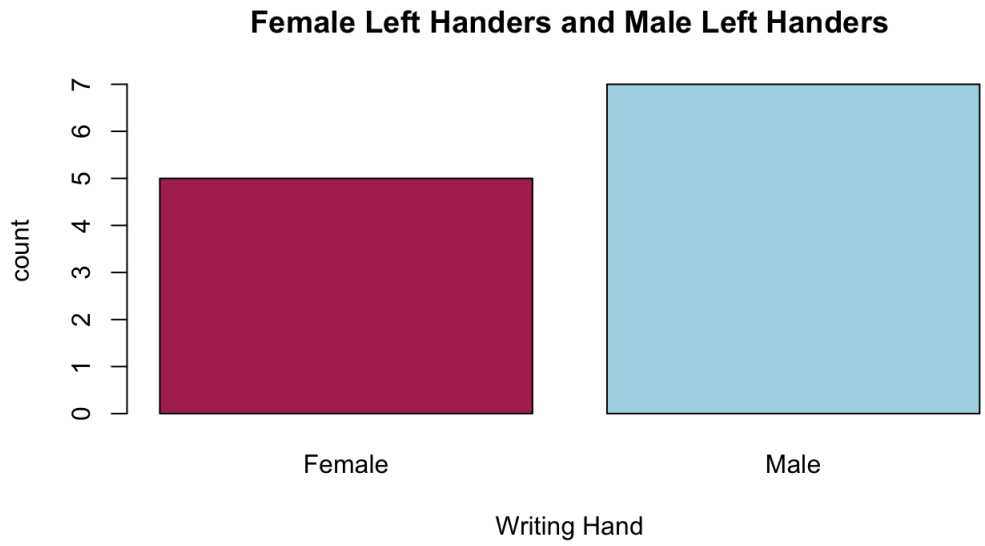
Q2) Plot a bar graph for the number of left handers and right handers in the survey. Provide the title as “Left Handers and Right Handers”, y-axis label as “count” and specify the colours for the bars

```
a=table(df$W.Hnd)
barplot(a,main="Left Handers and Right Handers",ylab="count",xlab="Writing Hand",col=c("Green","red"))
```



Q3) 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 , y-axis label as “count” and specify the colours for the bars.

```
temp<-df%>%
  filter(W.Hnd=='Left')
a=table(temp$Sex)
barplot(a,main="Female Left Handers and Male Left Handers",ylab="count",xlab="Writing Hand",col=c("maroon","li
```

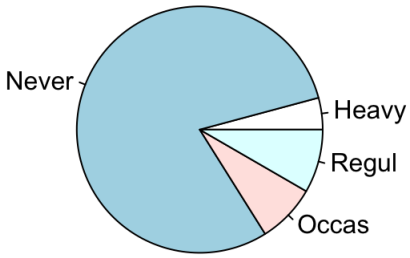


Q4) Draw the distribution of smoking habits of male left handers using pie chart.

```
df%>%
  filter(df$W.Hnd=='Left',df$Sex=='Male')
```

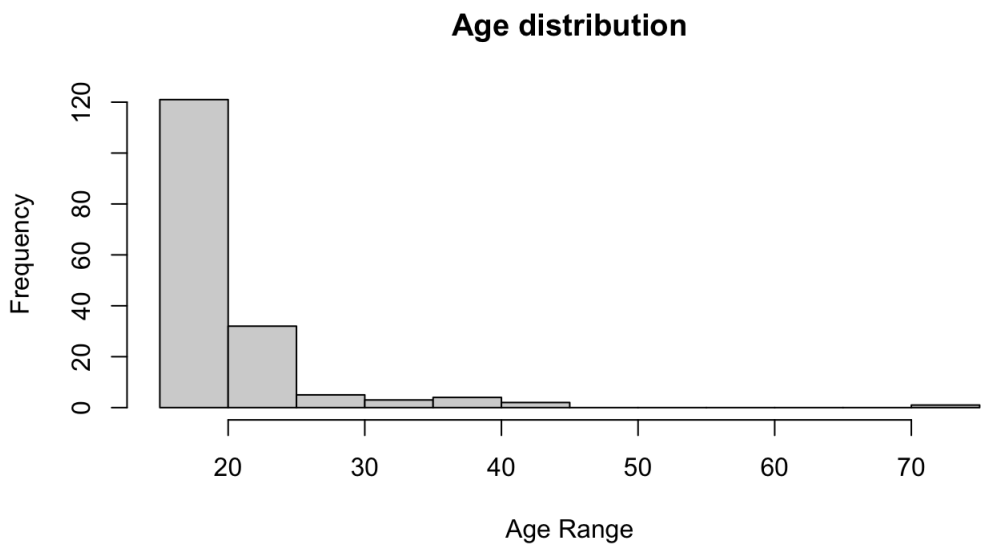
```
##    Sex Wr.Hnd NW.Hnd W.Hnd   Fold Pulse    Clap Exer Smoke Height      M.I
## 1 Male   19.5   20.5  Left R on L   104    Left None Regul 177.80 Imperial
## 2 Male   19.4   19.2  Left R on L    74   Right Some Never 182.88 Imperial
## 3 Male   22.0   21.5  Left R on L    55    Left Freq Never 200.00   Metric
## 4 Male   23.0   22.0  Left L on R    83    Left Some Heavy 193.04 Imperial
## 5 Male   19.8   20.0  Left L on R    59   Right Freq Never 180.00   Metric
## 6 Male   20.5   19.5  Left L on R    80   Right Some Occas 182.88 Imperial
## 7 Male   17.5   17.0  Left L on R    97 Neither None Never 165.00   Metric
##       Age
## 1 17.583
## 2 18.333
## 3 18.500
## 4 18.917
## 5 17.417
## 6 18.667
## 7 19.500
```

```r
a=table(df$Smoke)
pie(a)
```



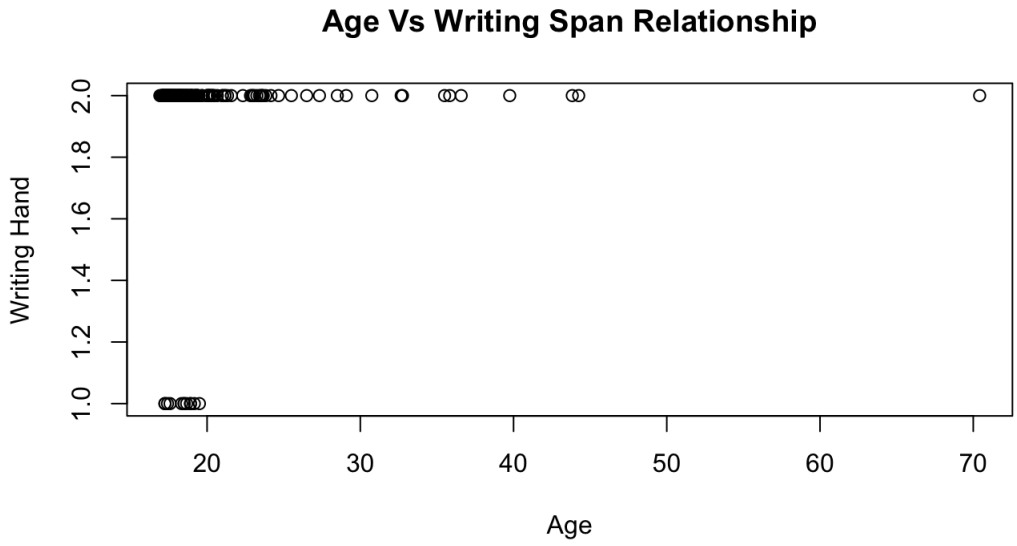
Q5) Draw the histogram of age distribution with the title as ‘Age distribution’ and xlabel as ‘Age range’ and ylabel as ‘frequency’.

```r
hist(df$Age,main='Age distribution',xlab ='Age Range' ,ylab = 'Frequency')
```



Q6) Reveal the relationship between the age and writing hand span using scatter plot

```
plot(df$Age,df$W.Hnd,main = "Age Vs Writing Span Relationship",xlab = "Age",ylab = "Writing Hand")
```

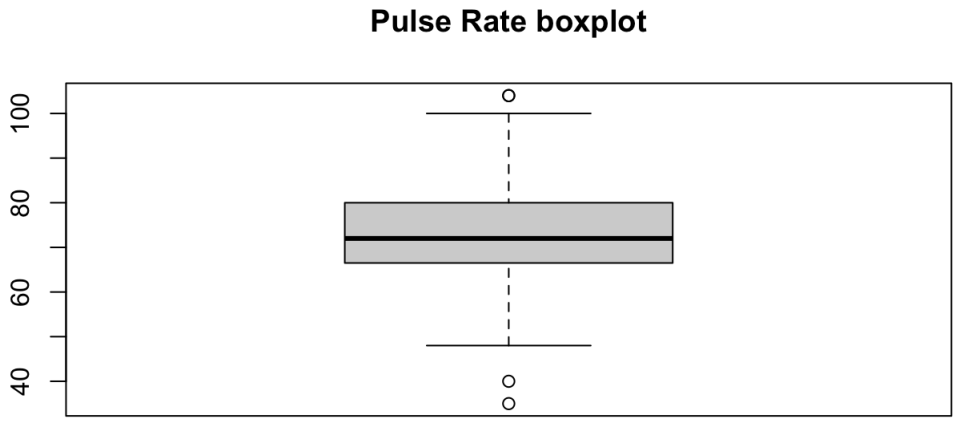


Q7) Draw the boxplot for pulse rate to analyse the five summary statistics. Provide appropriate title and label.

```
summary(df$Pulse)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   35.00   66.75   72.00   74.02   80.00  104.00
```

```
boxplot(df$Pulse,main='Pulse Rate boxplot')
```



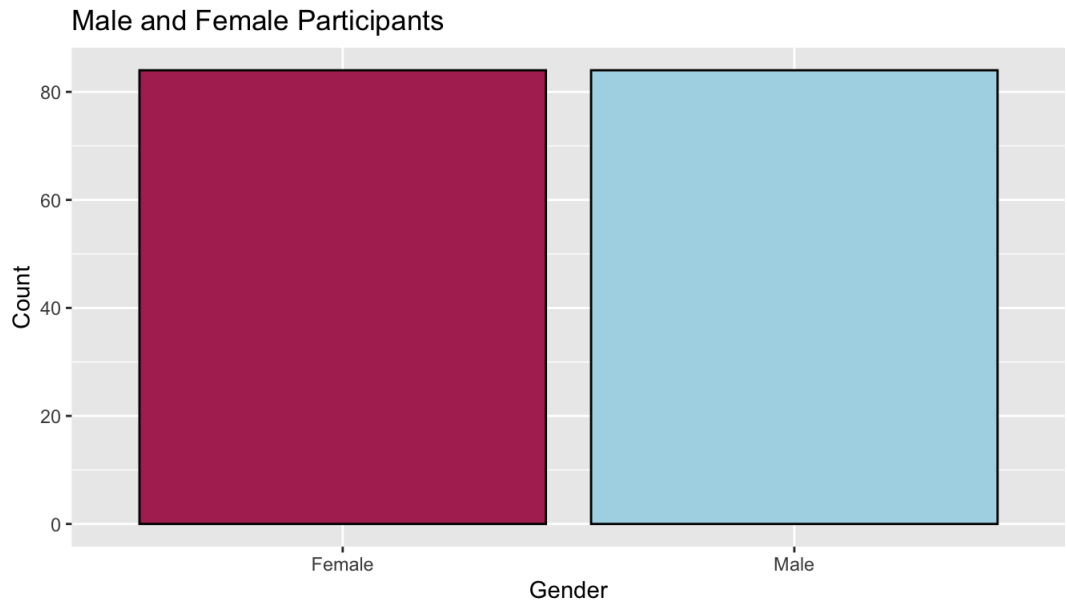
Ex-5b Visualization using grammar of graphics

Q1) Install the package ggplot2 and import it.

```
#install.packages("ggplot2")
library(ggplot2)
```

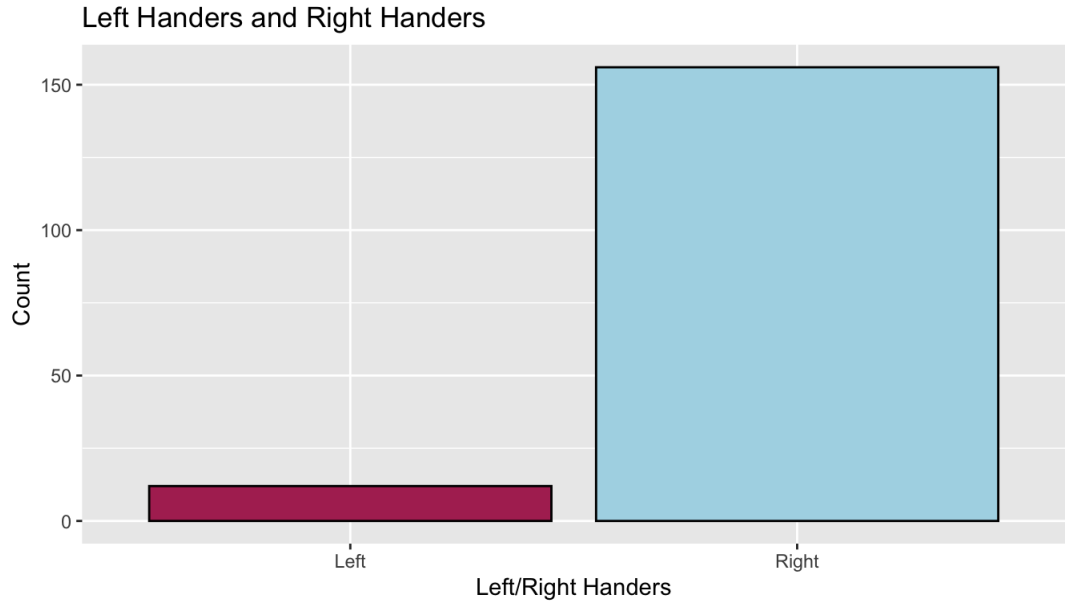
Q2) Plot a bar graph for the number of male and female participants in the survey. Provide the title as “Male and Female participants” and specify the colours for the bars.

```
ggplot(data = df,aes(x=Sex))+(labs(title = 'Male and Female Participants',x='Gender',y='Count'))+geom_bar(colo
```



Q3) Plot a bar graph for the number of left handers and right handers in the survey. Provide the title as “Left Handers and Right Handers” and specify the colours for the bars.

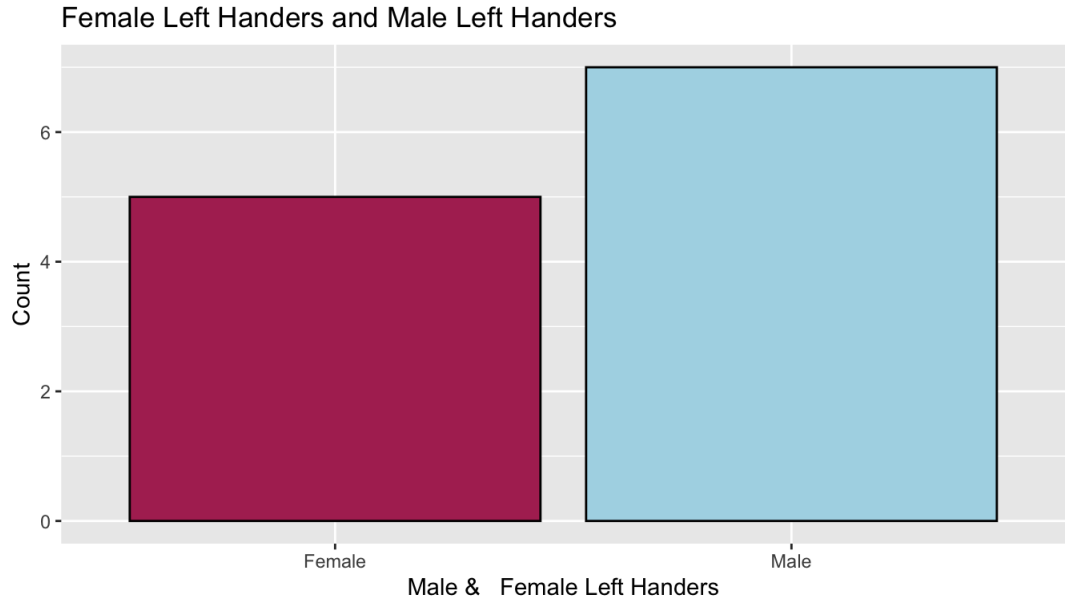
```
ggplot(data = df,aes(x=df$W.Hnd))+(labs(title = 'Left Handers and Right Handers',x='Left/Right Handers',y='Co
```



Q4) 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

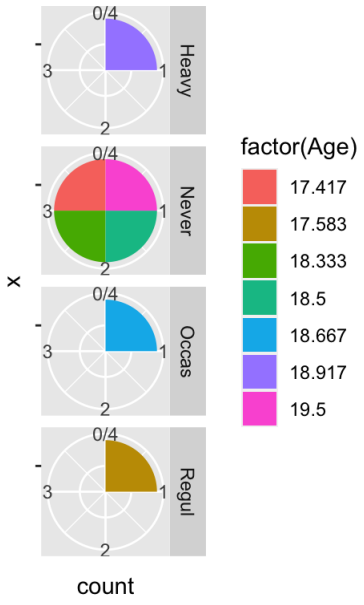
```
tmp <- df%>%
  filter(df$W.Hnd=='Left')

ggplot(data = tmp,aes(x=Sex))+(labs(title = 'Female Left Handers and Male Left Handers',x='Male &   Female Lef
```



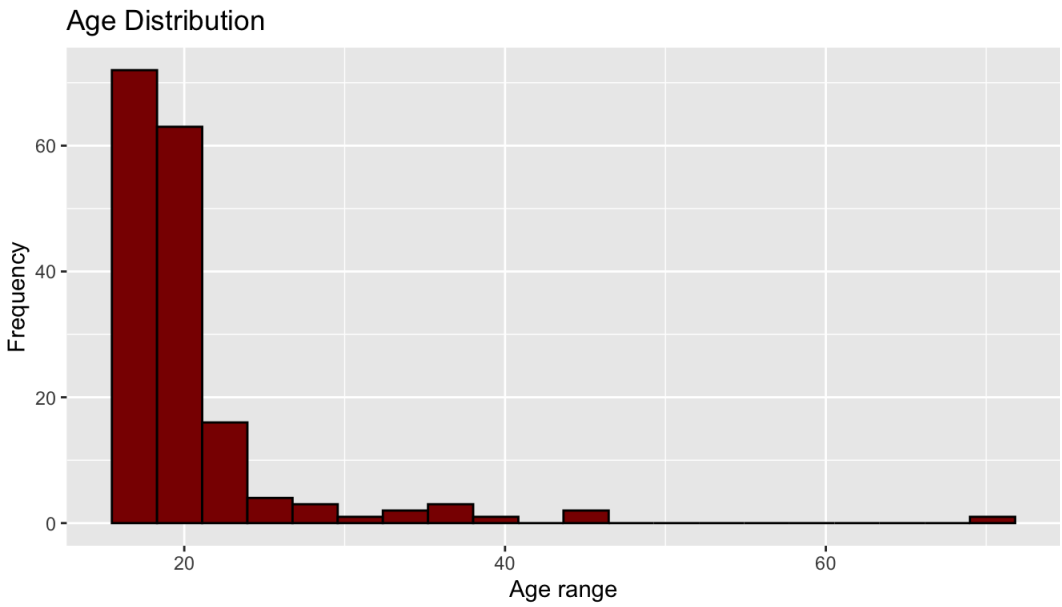
Q5) Draw the distribution of smoking habits of male left handers based on age using pie chart.

```
temp<-df%>%
  filter(Sex=='Male' & W.Hnd=='Left')
ggplot(temp,aes(x="",fill=factor(Age)))+
  geom_bar()+coord_polar("y",start=0)+facet_grid(Smoke~.)
```



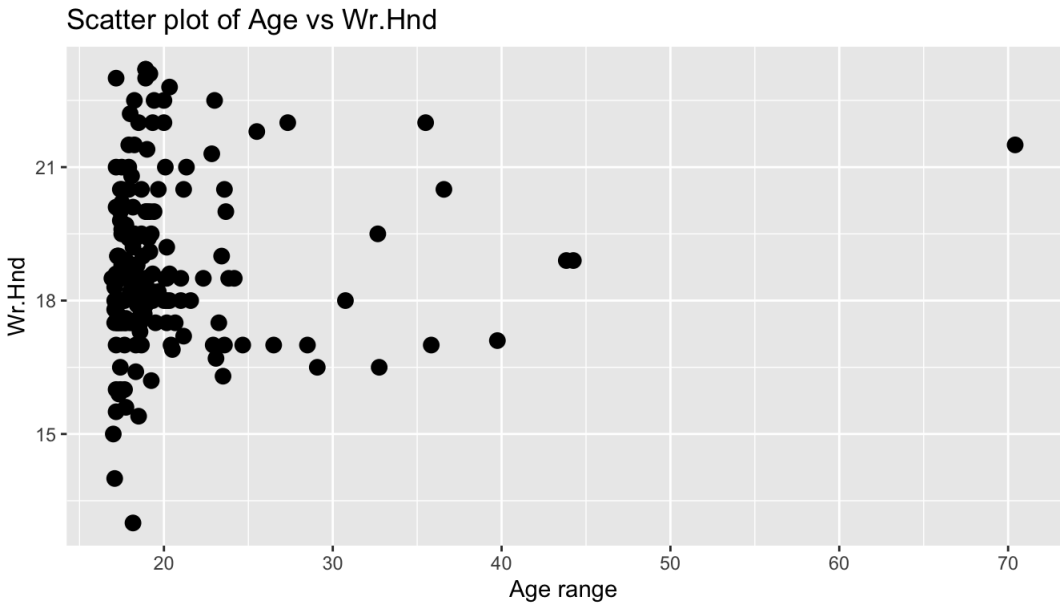
Q6) Draw the histogram of age distribution with the title as ‘Age distribution’ and xlabel as ‘Age range’ and ylabel as ‘frequency’.

```
ggplot(df,aes(x=Age))+
  geom_histogram(stat="bin",bins="20",color="black",fill="Darkred")+labs(x="Age range",y="Frequency",title="Ag
```



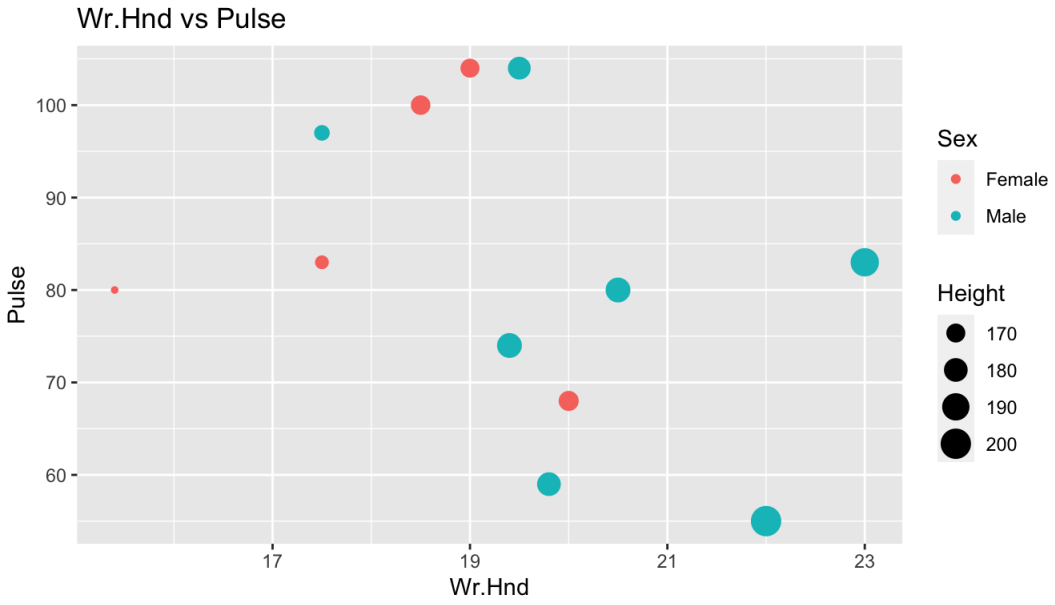
Q7) Reveal the relationship between the age and writing hand span using scatter plot.

```
ggplot(df,aes(x=Age,y=Wr.Hnd))+
  geom_point(color="black",size=3)+labs(x="Age range",y="Wr.Hnd",title="Scatter plot of Age vs Wr.Hnd")
```



Q8) Plot the distribution of writing hand span vs. pulse rate of left handers. Provide colour based on gender and vary the size of the point based on height of the student

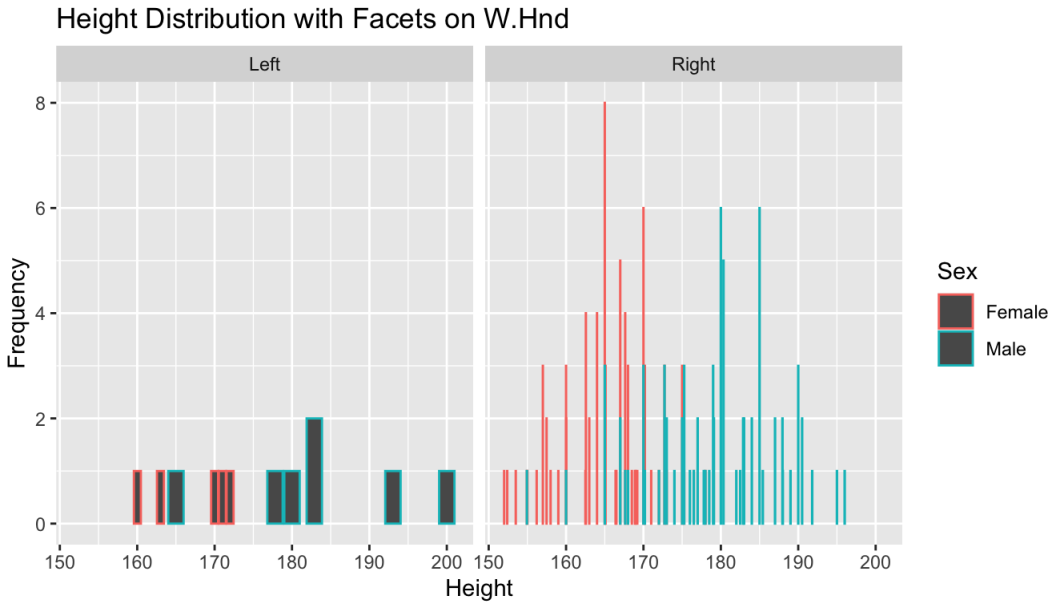
```
temp<-df%>%
  filter(W.Hnd=='Left')
ggplot(temp,aes(x=Wr.Hnd,y=Pulse))+
  geom_point(aes(size=Height,color=Sex))+labs(x="Wr.Hnd",y="Pulse",title="Wr.Hnd vs Pulse")
```



Q9) Plot the distribution of height of the students with filled colour based on gender with facet based on left and write handers.

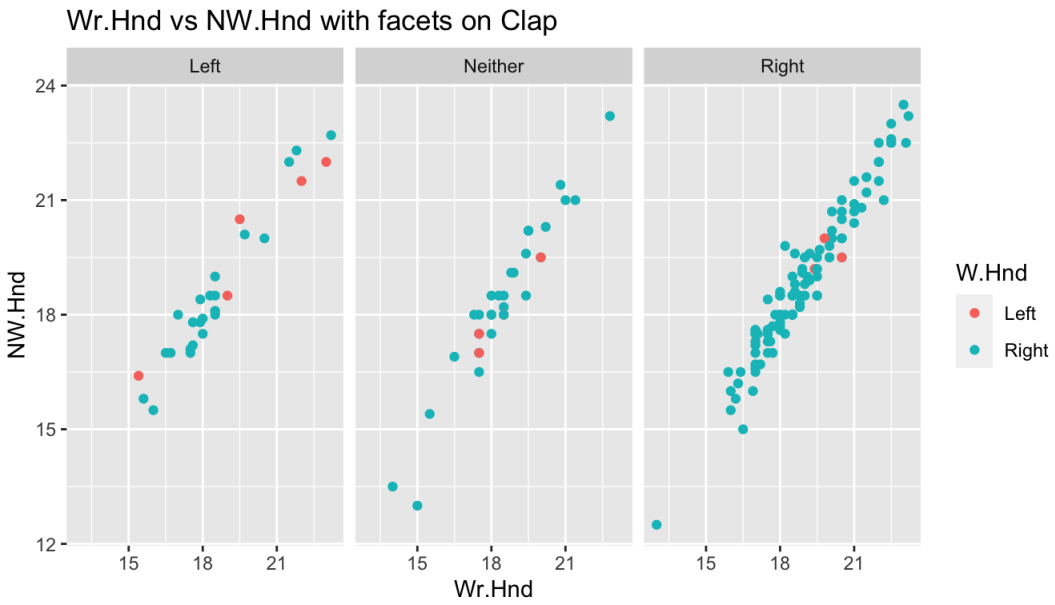
```
ggplot(df,aes(x=Height,color=Sex))+
  geom_bar()+labs(x="Height",y="Frequency",title="Height Distribution with Facets on W.Hnd")+facet_grid(.~W.Hr
```

```
## Warning: position_stack requires non-overlapping x intervals
```



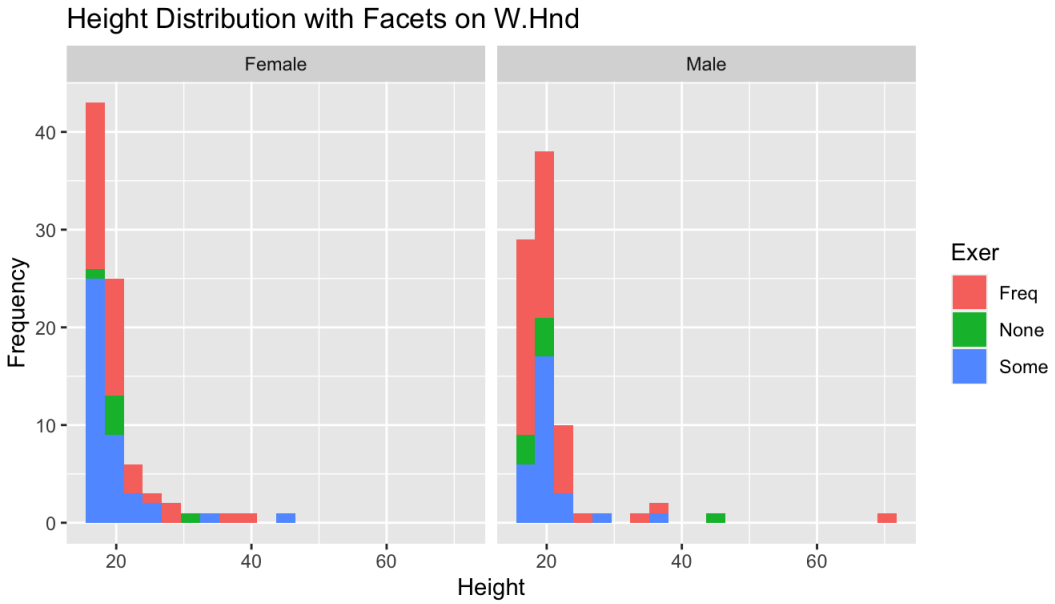
Q10) Plot the trend of span of writing hand vs. non-writing hand coloured and grouped based on left and right handers with facet label based on clap.

```
ggplot(df,aes(x=Wr.Hnd,y=NW.Hnd,group=W.Hnd))+
  geom_point(aes(color=W.Hnd))+labs(x="Wr.Hnd",y="NW.Hnd",title="Wr.Hnd vs NW.Hnd with facets on Clap") + face
```



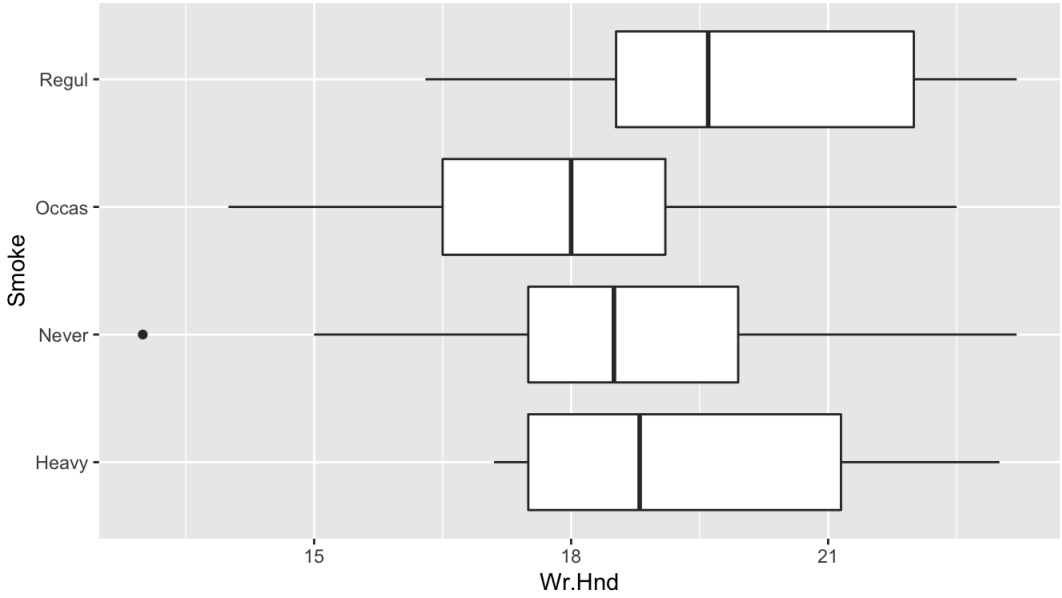
Q11) Plot the distribution of age of students based on categories under exercise with facet wrap based on gender

```
ggplot(df,aes(x=Age))+
  geom_bar(stat="bin",bins='20',aes(fill=Exer))+labs(x="Height",y="Frequency",title="Height Distribution with
```



Q12) Plot the box plot of writing hand span with respect to smoking habits of students

```
ggplot(df,aes(x=Wr.Hnd,y=Smoke))+
  geom_boxplot()
```


Regul
Occas
Never
Heavy
Smoke
15
18
21
Wr.Hnd