

10Oct

Devanshi

2023-10-10

#Stem and leaf plot is a technique of displaying the frequencies with which some classes of values may occur. # It is basically the method of representing the quantitative data in the graphical format. # The term stem and leaf is so because the plot is given in a tabular format where each numeric value or data item is split into a stem # Sometimes decimal places are rounded up and altering the scale can provide a better picture of the data # In general, the default stem() command tries to provide a clear picture of the distribution, which sometimes lead to loss of decimal places # Q. On world Obesity Day, suppose in a school a teacher decided to measure the weight of any 10 students whom she feels may be obese. So she records the weight of 10 students as: 54, 43, 67, 76, 45, 59, 66, 78, 80, 92

```
data= c(54, 43, 67, 76, 45, 59, 66, 78, 80, 92)
table(data)
```

```
## data
## 43 45 54 59 66 67 76 78 80 92
##  1  1  1  1  1  1  1  1  1  1
```

```
stem(data)
```

```
##
##  The decimal point is 1 digit(s) to the right of the |
##
##  4 | 35
##  5 | 49
##  6 | 67
##  7 | 68
##  8 | 0
##  9 | 2
```

```
stem(data, scale=2)
```

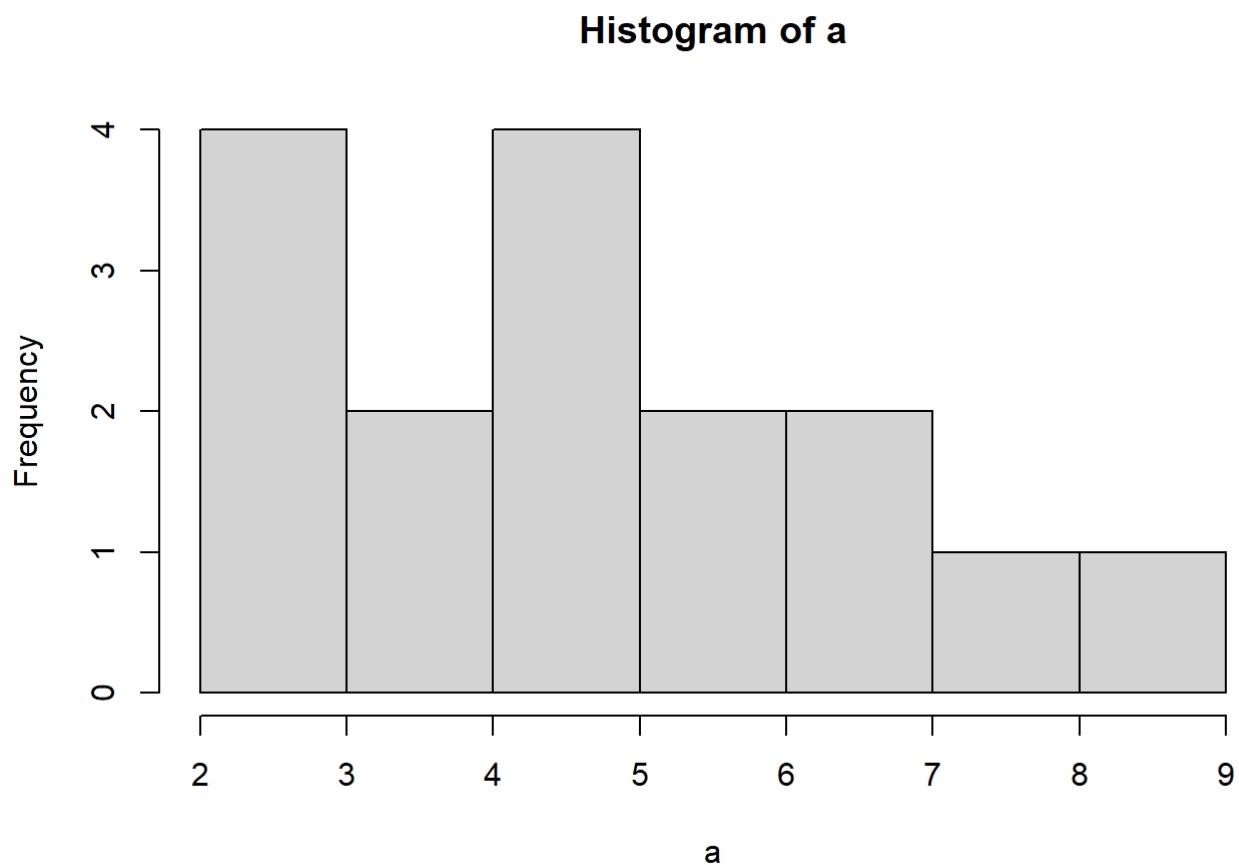
```
##
##  The decimal point is 1 digit(s) to the right of the |
##
##  4 | 3
##  4 | 5
##  5 | 4
##  5 | 9
##  6 |
##  6 | 67
##  7 |
##  7 | 68
##  8 | 0
##  8 |
##  9 | 2
```

#Histograms #A histogram is a type of bar chart which shows the frequency of the number of values which are compared with a set of value ranges #The histogram is used for the distribution, whereas bar chart is used for comparing different entities. #In histogram, each bar represents the height of values present in the given range. #The frequencies are represented on the y-axis #The x-axis shows the values separated into various bins

```
a=c(3,5,7, 5, 3, 2, 6, 8, 5,6, 9, 4, 5,7, 3, 4)
table(a)
```

```
## a
## 2 3 4 5 6 7 8 9
## 1 3 2 4 2 2 1 1
```

```
hist(a)
```



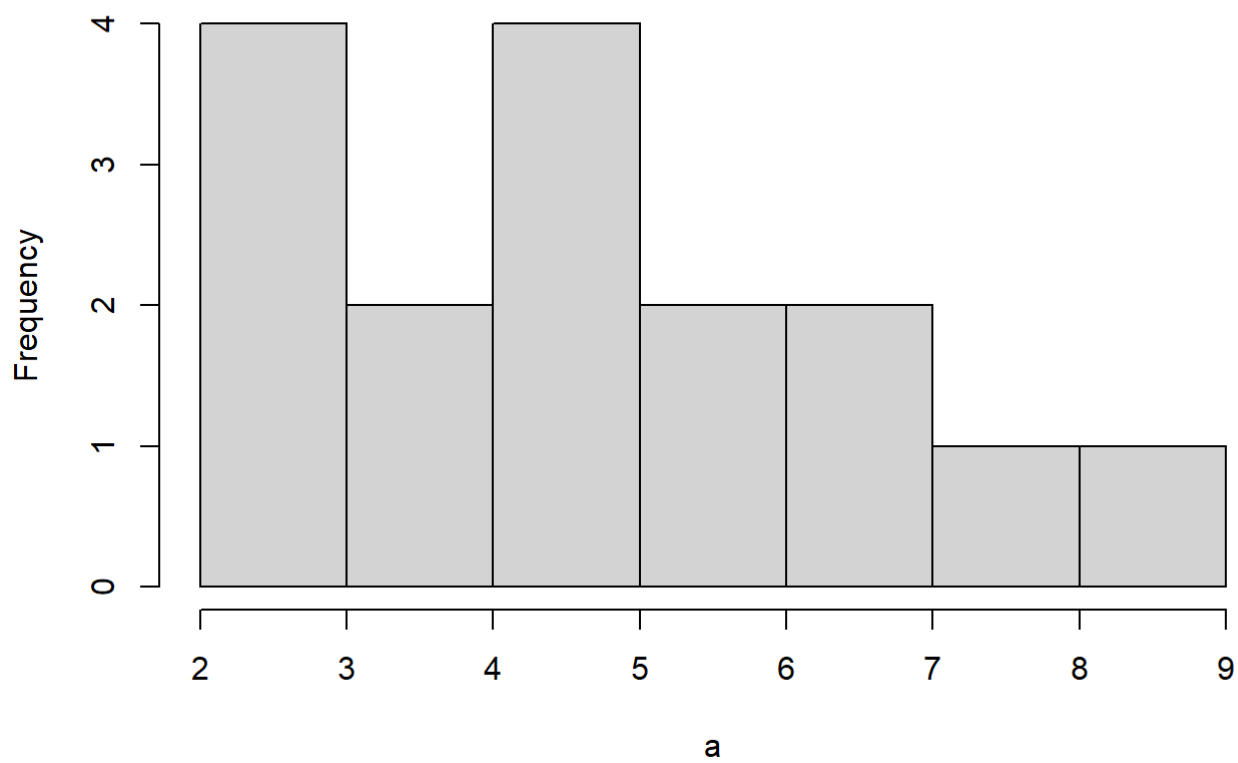
#breaks= instruction is used to alter the number of columns that are displayed Specify the number of breaks as a simple number or a range of numbers

```
a=c(3,5,7, 5, 3, 2, 6, 8, 5,6, 9, 4, 5,7, 3, 4)
table(a)
```

```
## a
## 2 3 4 5 6 7 8 9
## 1 3 2 4 2 2 1 1
```

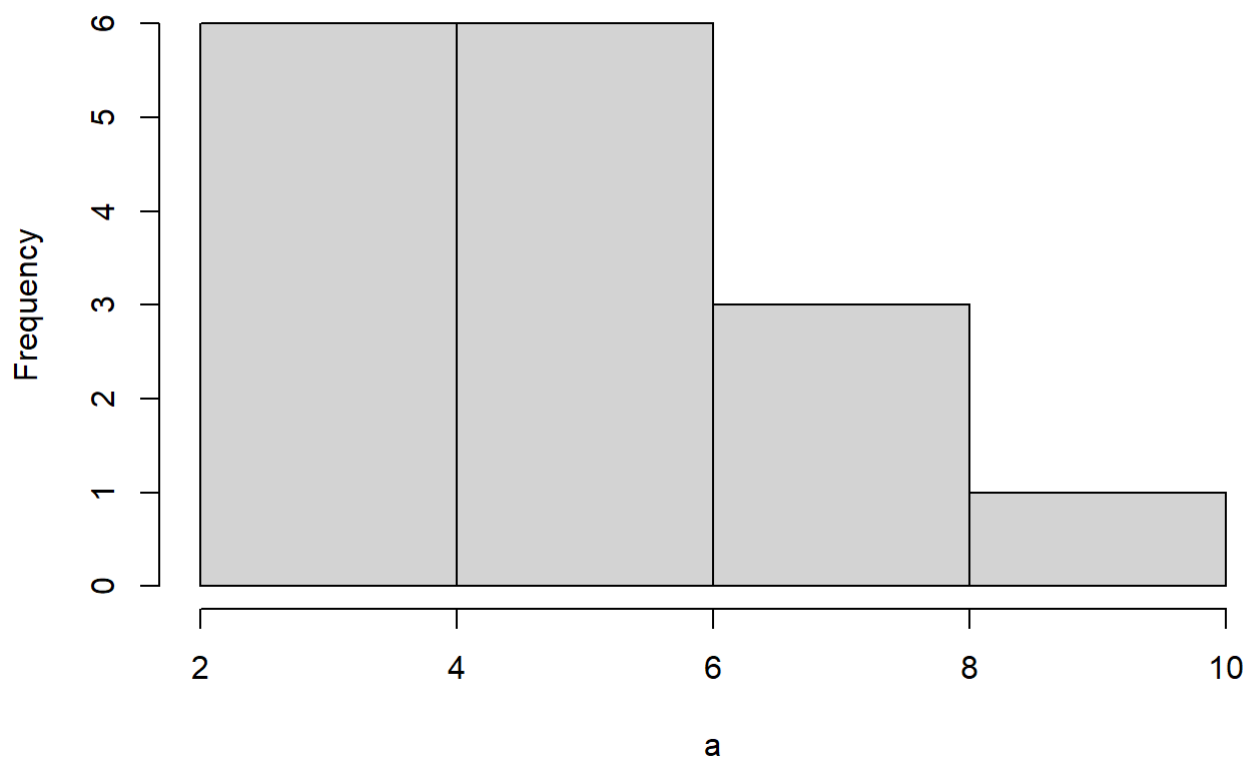
```
hist(a, breaks="st")
```

Histogram of a



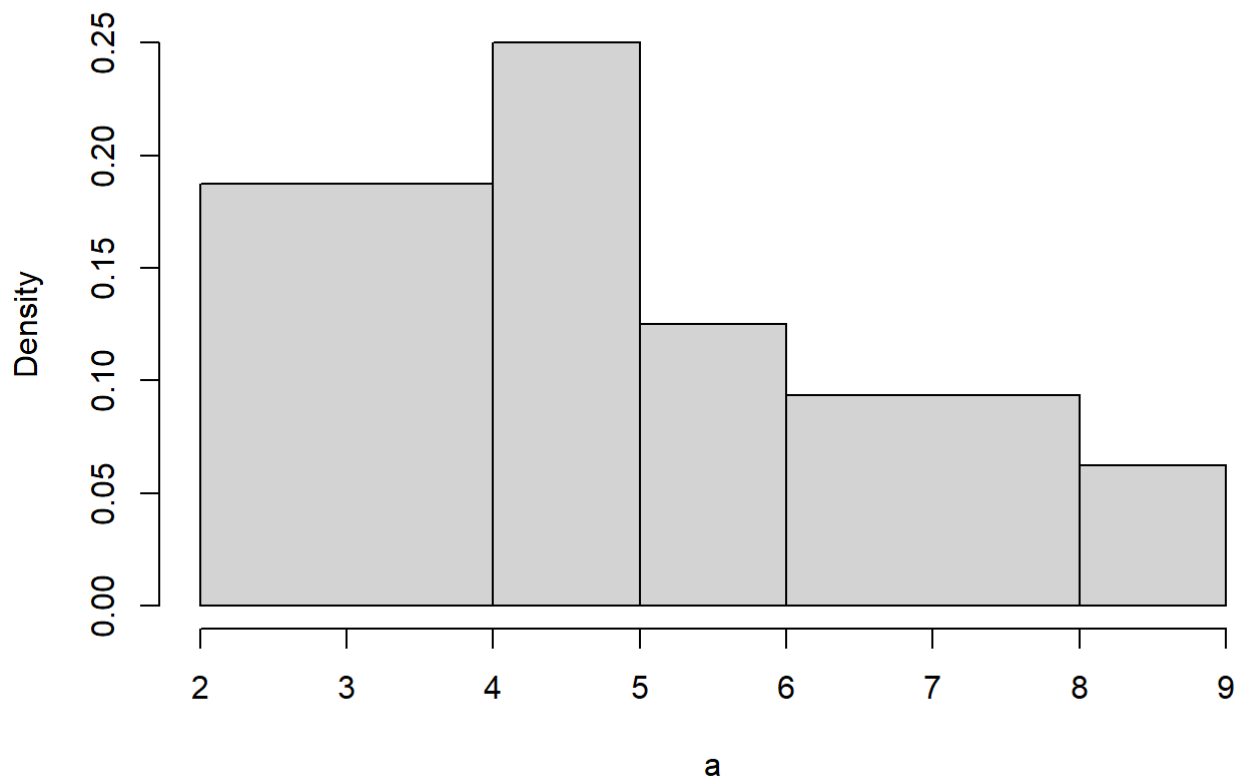
```
hist(a, breaks="sc")  
hist(a, breaks="fr")
```

Histogram of a



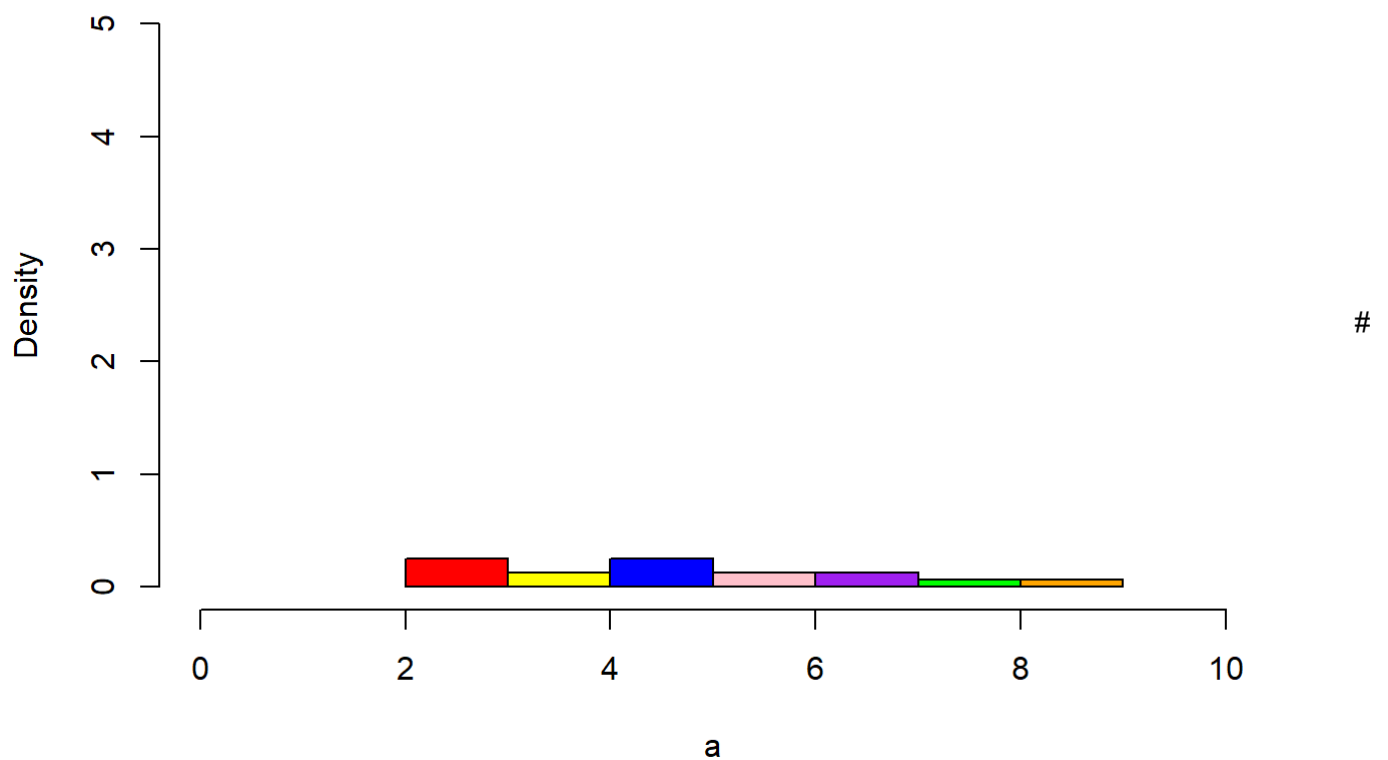
```
hist(a, breaks= c(2,4,5,6,8,9))
```

Histogram of a



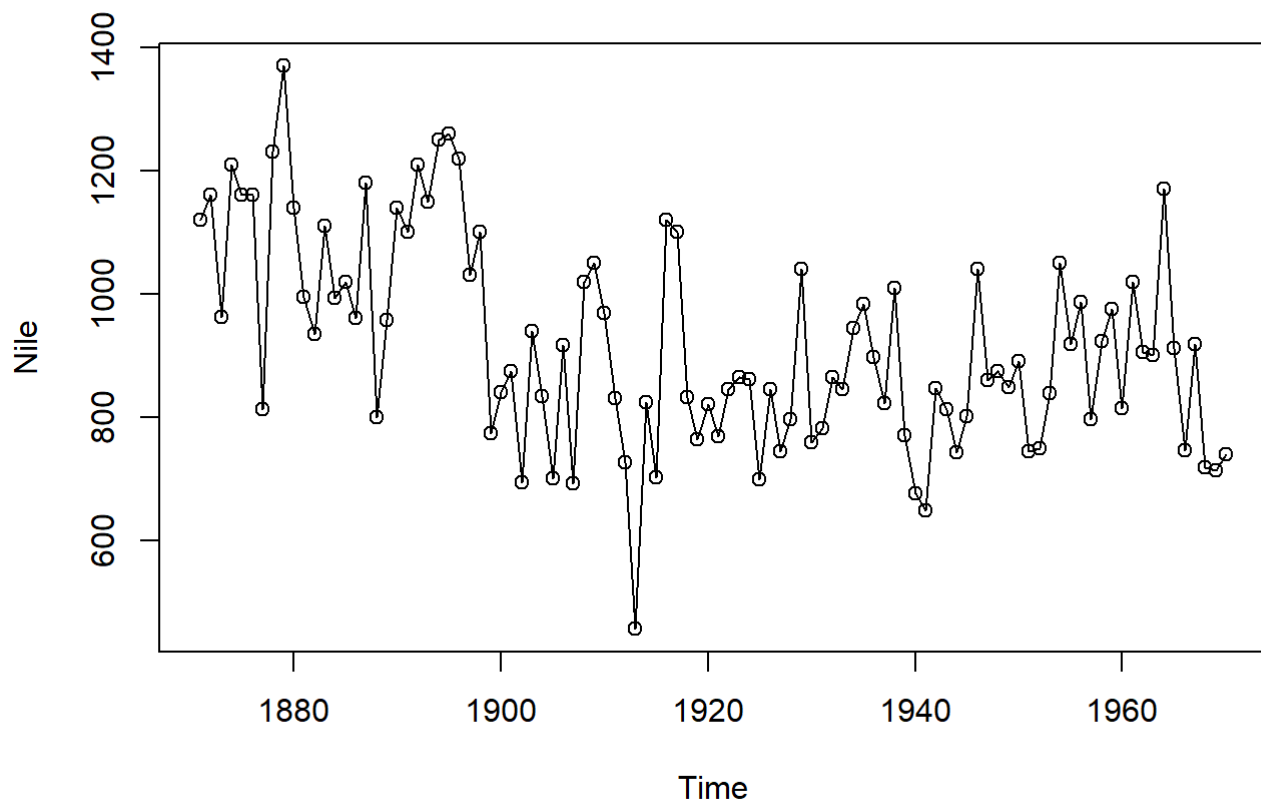
Additional functions: -col= -main= -ylab= -xlab= -Range of X-axis xlim= c(a,b) -Range of Y-axis ylim= c(a,b) -main=NULL removes the heading -freq= F gives density

```
a=c(3,5,7, 5, 3, 2, 6, 8, 5,6, 9, 4, 5,7, 3, 4)
hist(a, col=c("red", "yellow", "blue", "pink", "purple", "green", "orange"), ylim=c(0,5), xlim=c(0,10), main=NULL, freq = F)
```



Line Charts type= Instruction can alter the way data is drawn on the plot area type= "P" means points only type="b" means points with line segments type= "l" means only lines (default) type= "o" means lines overplotted with points, that is no gap between line segments type= "c" means line segment with small gaps where the points will be type= "n" gives no plot

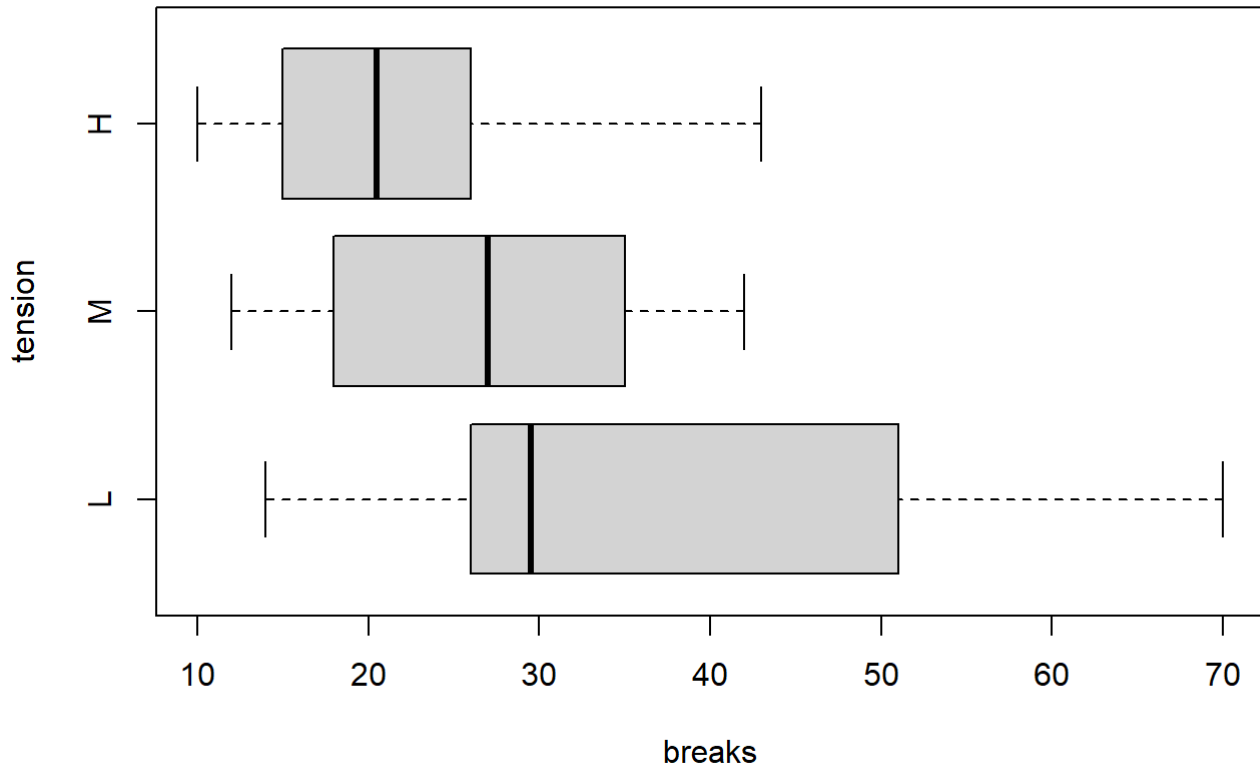
```
plot(Nile, type="o")
```



Q. Look at the warpbreaks data that comes built into R. Create a box-whisker plot of the number of breaks for the different tensions. Make the plot using horizontal plot and display the whiskers to the extreme range of the data. How can you draw this graph to display only a single type of wool?

```
attach(warpbreaks)
boxplot(breaks~tension, horizontal= T, range=0)
title("Tension")
```

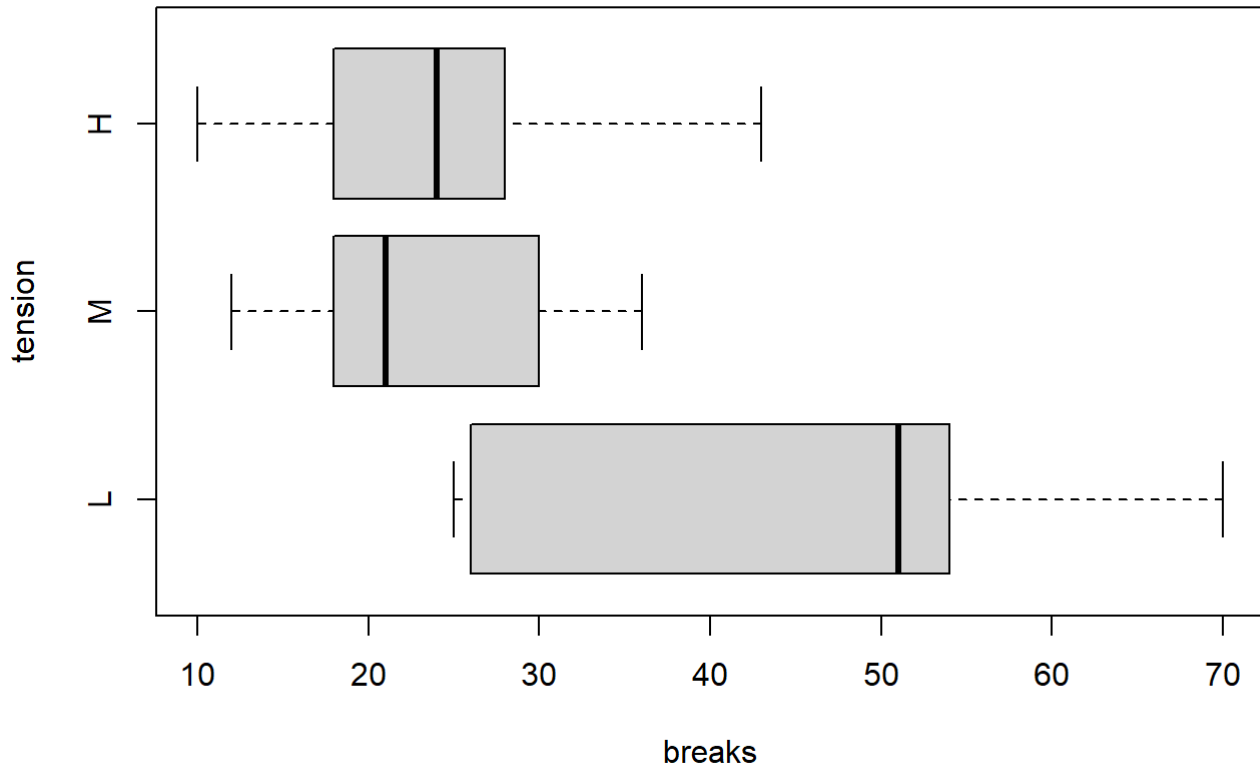
Tension



```
detach(warpbreaks)
```

```
attach(warpbreaks)
d=subset(warpbreaks, wool=="A")
boxplot(breaks~tension, d, horizontal= T, range=0)
title("Tension")
```

Tension

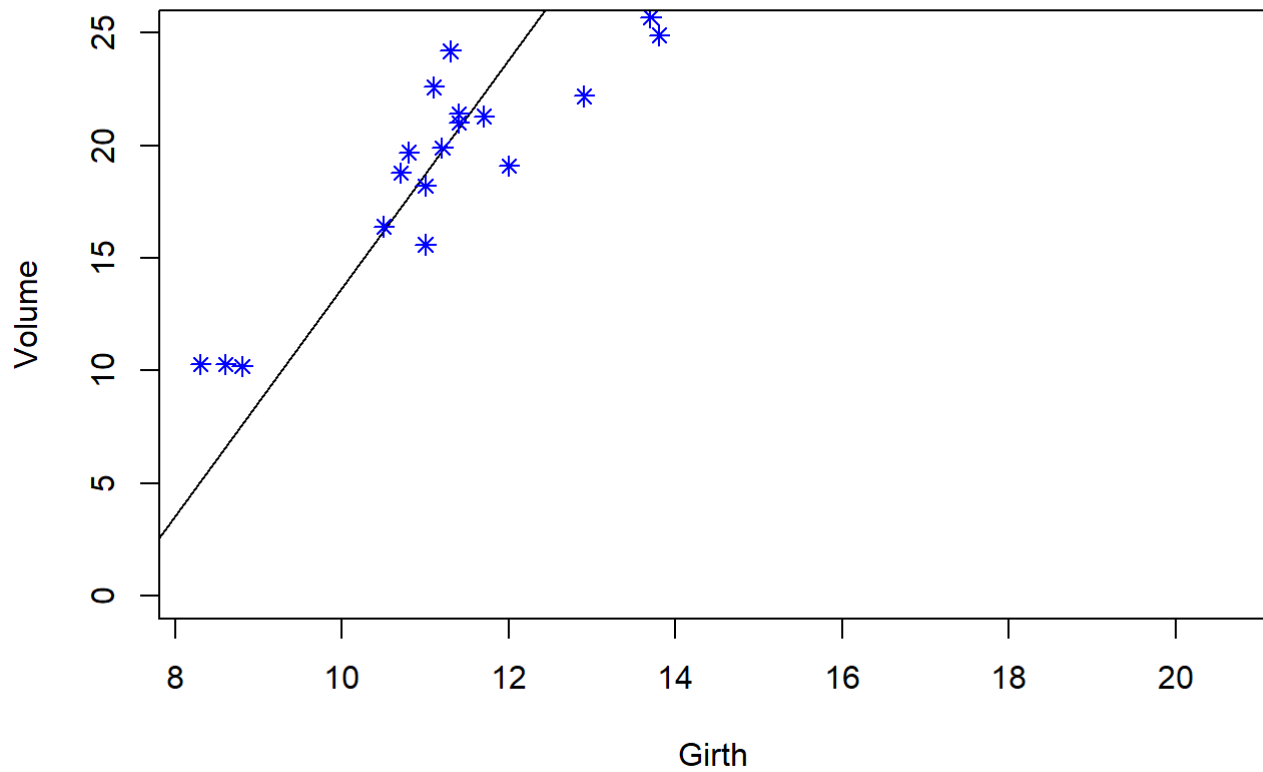


```
detach(warpbreaks)
```

Q. The trees data come as part of R. The data is composed of three columns: the Girth of black cherry trees (in inches), the Height (in feet). and the Volume of wood (in cubic feet). How can you make a scatter plot of girth versus volume and display a line of best-fit? Modify the axes so that the intercept is seen clearly. Use an appropriate plotting symbol and colours to help make the plot chart more "interesting".

```
View (trees)
attach(trees)
plot(Girth, Volume, col="blue", pch=8, main="trees", abline(lm(Volume~Girth)), ylim=c(0,25))
```


trees



```
cor(Girth, Volume) #corelation lies between -1 and 1 if it's close to 0, there is no correlation
```

```
## [1] 0.9671194
```

```
detach (trees)
```