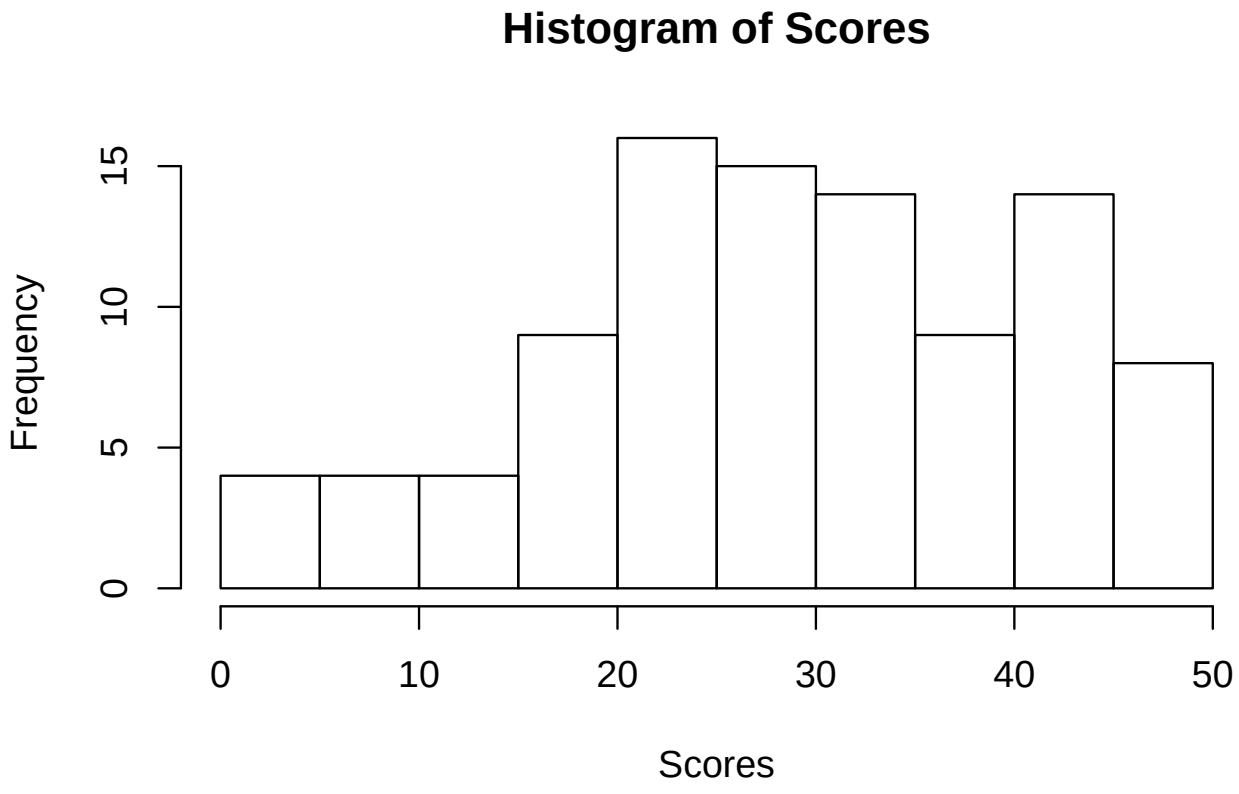
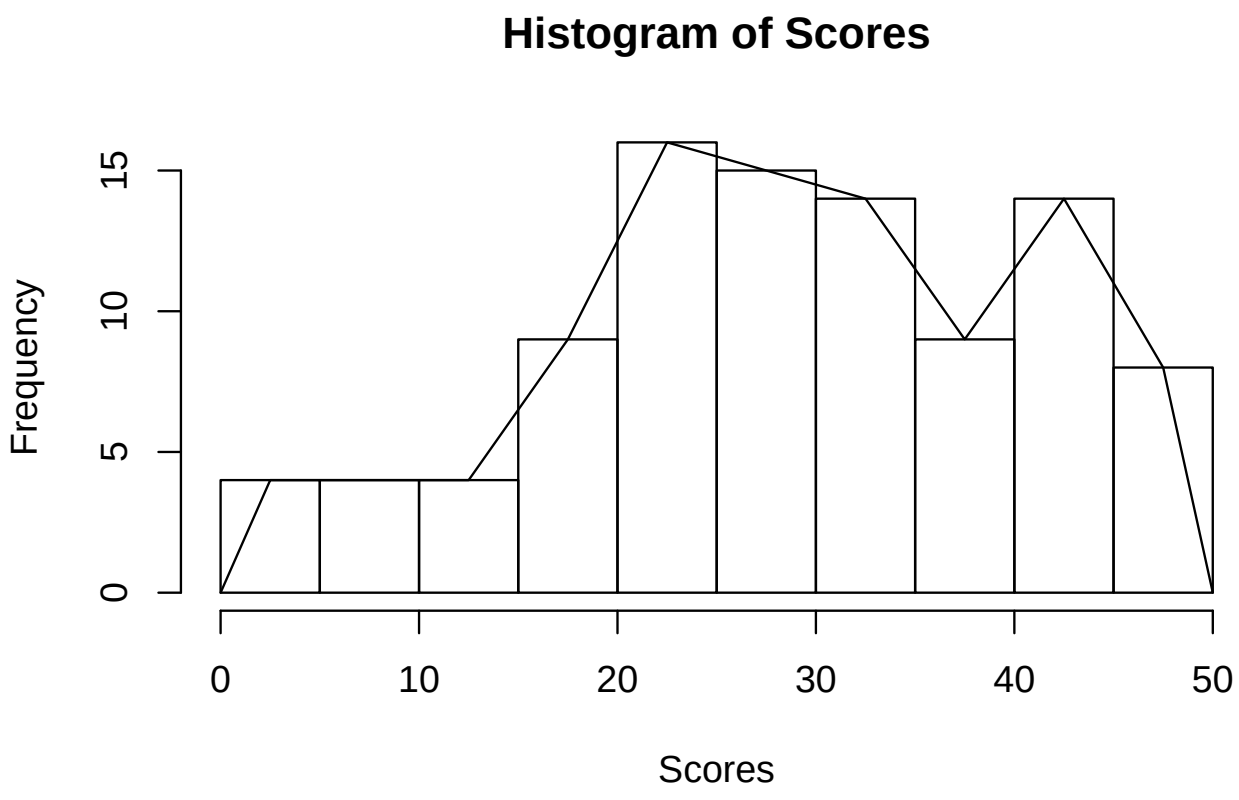
Histogram-Wrap up

```
> Scores = scan("Scores")

> hist(Scores)
```



Histogram-Wrap up

```
> Score1=hist(Scores)
> lines(c(min(Score1$breaks),Score1$mids,max(Score1$breaks)),c(0,Score1$counts,0),type="l")
```



Histogram-Wrap up

```
> bins = seq(0,50, by=5)
> intervaldata = cut(Scores, breaks=bins)
> table(intervaldata)

intervaldata
  (0,5]  (5,10] (10,15] (15,20] (20,25] (25,30] (30,35] (35,40] (40,45] (45,50]
      2       4       4       9      16      15      14       9      14       8
```

Density Plot

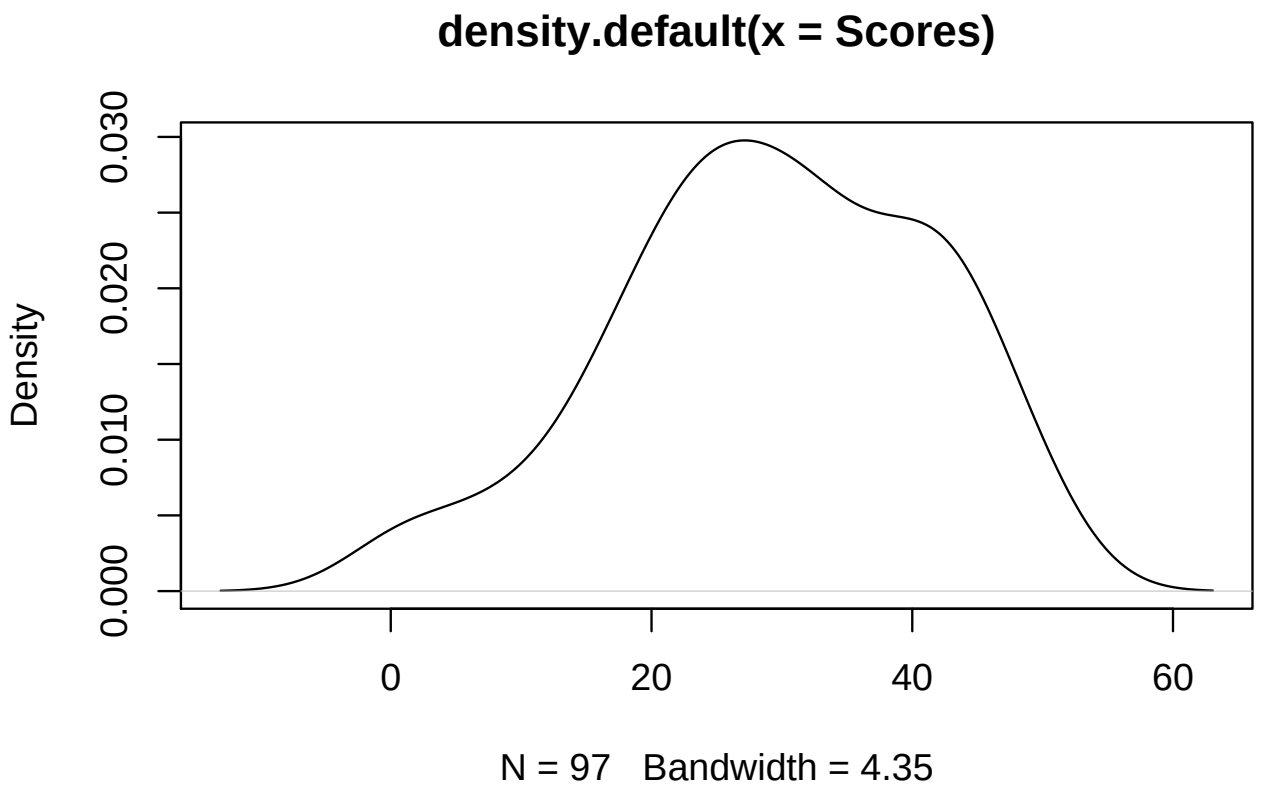
- The density() function has means to do automatic selection of bandwidth.
- If we use the default choice it is easy to add a density plot to a histogram.
- We just call the lines function with the result from density (or plot if it is the first graph).
- The basic idea is for each point to take some kind of average for the points nearby and based on this give an estimate for the density.

Density Plot

```
> density(Scores)

Call:
        density.default(x = Scores)

Data: Scores (97 obs.);        Bandwidth 'bw' = 4.35

       x                    y
 Min.   :-13.051   Min.   :2.477e-05
 1st Qu.:  5.974   1st Qu.:2.830e-03
 Median : 25.000   Median :1.026e-02
 Mean   : 25.000   Mean   :1.313e-02
 3rd Qu.: 44.026   3rd Qu.:2.459e-02
 Max.   : 63.051   Max.   :2.977e-02
```

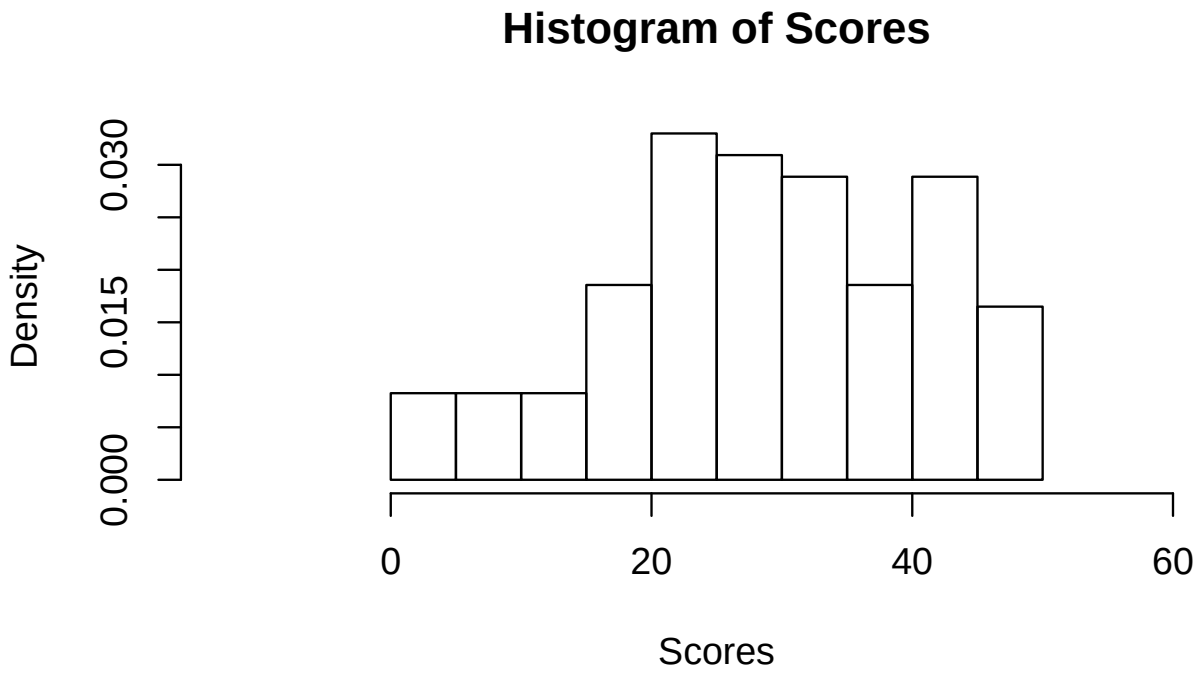
Density Plot

```
> plot(density(Scores))
```



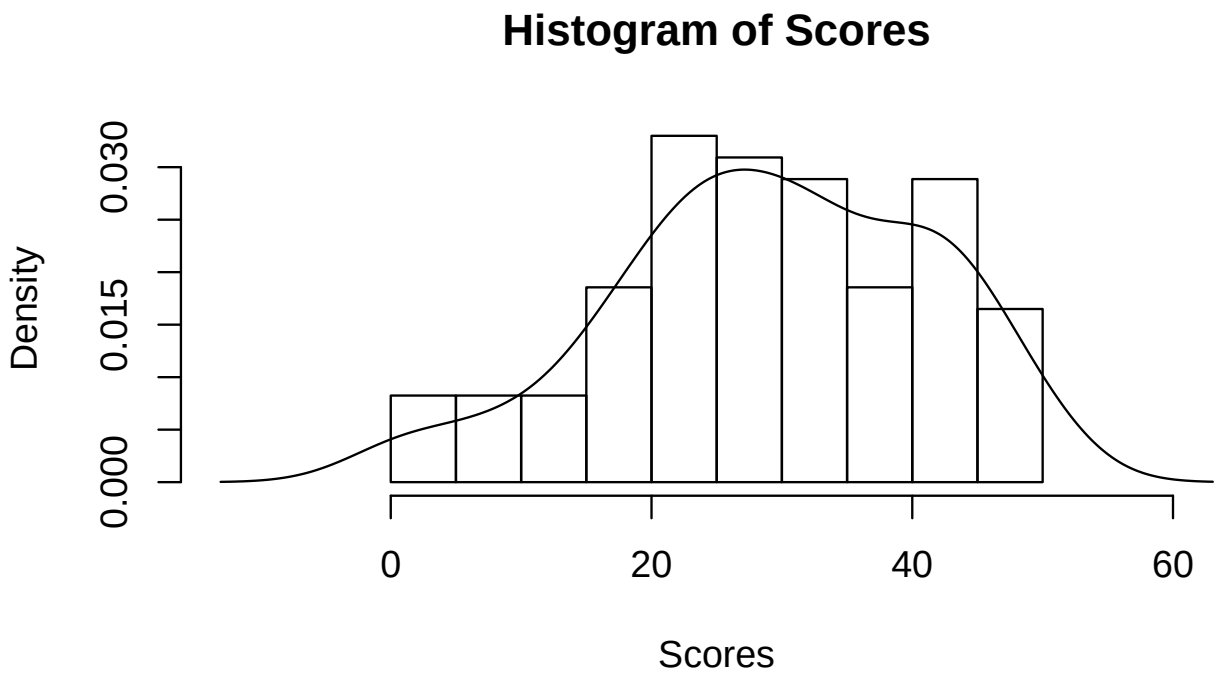
Histogram and Density Plot together

```
> Scores_hd = hist(Scores, plot=FALSE)
> Scores_dd = density(Scores)
> hist(Scores, probability=TRUE,
+ xlim = range(c(Scores_hd$breaks, Scores_dd$x)),
+ ylim = range(c(Scores_hd$density, Scores_dd$y)))
```



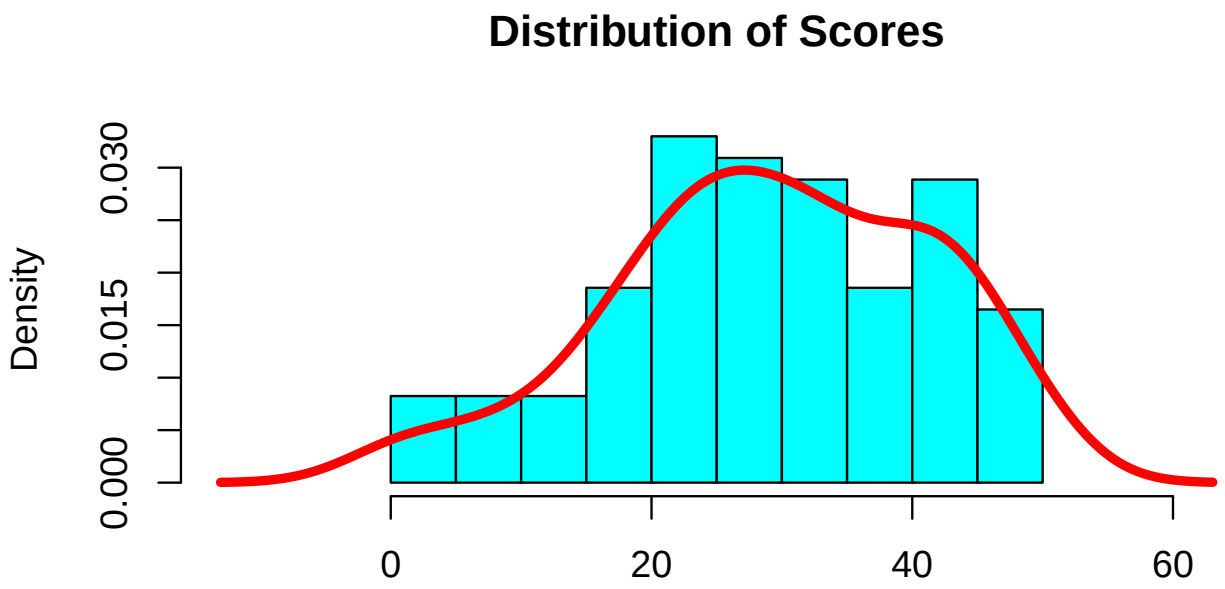
Histogram and Density Plot together

```
> Scores_hd = hist(Scores, plot=FALSE)
> Scores_dd = density(Scores)
> hist(Scores, probability=TRUE,
+ xlim = range(c(Scores_hd$breaks, Scores_dd$x)),
+ ylim = range(c(Scores_hd$density, Scores_dd$y)))
> lines(Scores_dd)
```



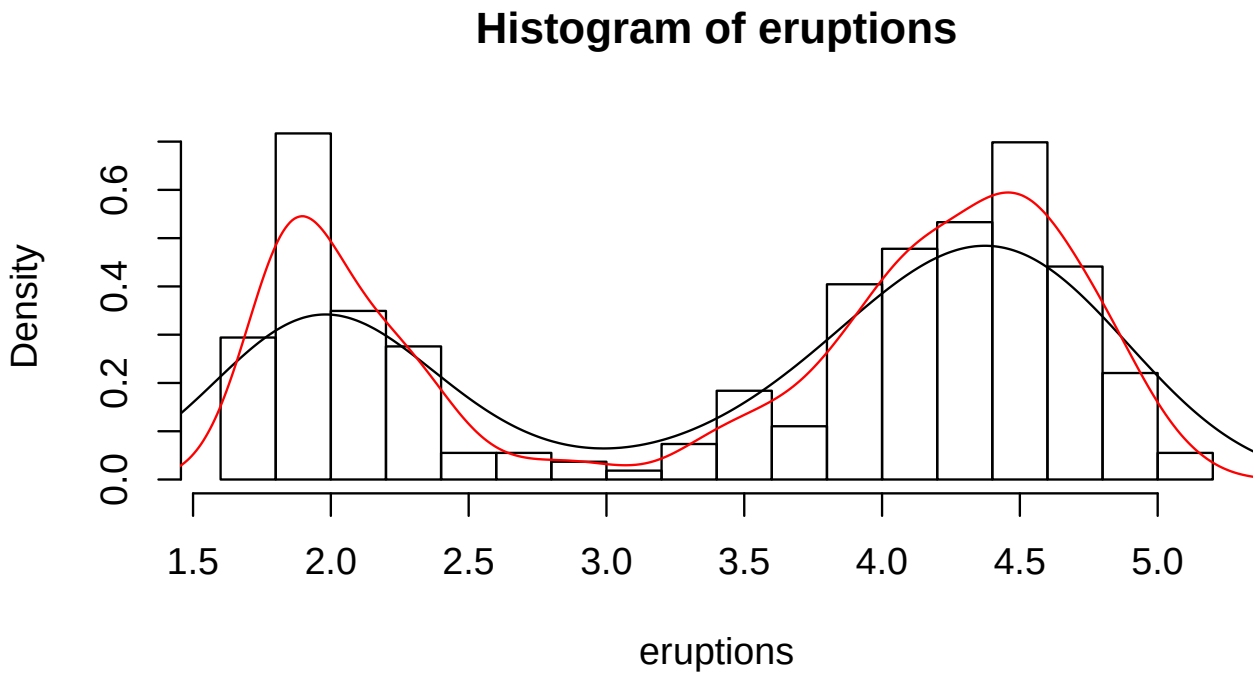
Histogram and Density Plot together

```
> Scores_hd = hist(Scores, plot=FALSE)
> Scores_dd = density(Scores)
> hist(Scores, main = "Distribution of Scores",
+ probability=TRUE,
+ xlim = range(c(Scores_hd$breaks, Scores_dd$x)),
+ ylim = range(c(Scores_hd$density, Scores_dd$y)), col="cyan")
> lines(Scores_dd, col="red", lwd=4)
```



Histogram and Density Plot together

```
> data(faithful)
> attach(faithful)
> hist(eruptions,15,prob=T)
> lines(density(eruptions))
> lines(density(eruptions, bw="SJ"), col="red")
```



Stem-Leaf Plot

- There are a range of graphical summaries of data.
- If the data set is relatively small, the stem-and-leaf diagram is very useful for seeing the shape of the distribution and the values.

  cd
- The number on the left of the bar is the stem, the number on the right the digit.
- One can read of the observations by reading the stem along with each (respective) leaf.

Stem-Leaf Plot

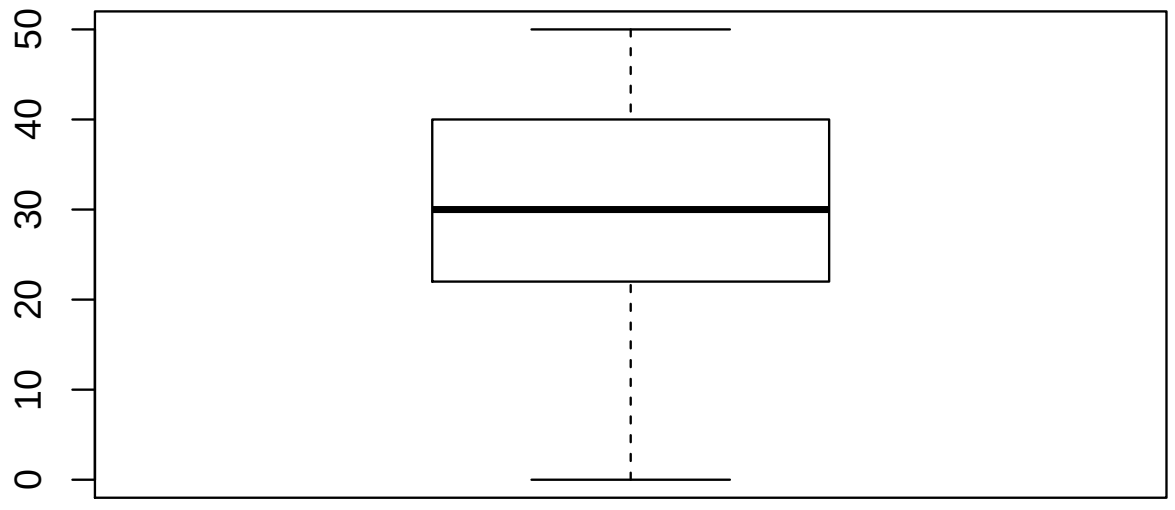
```
> stem(Scores)

  The decimal point is 1 digit(s) to the right of the |

  0 | 00238899
  1 | 44456777889
  2 | 00112222233444555566677788889
  3 | 00000112223333455556778
  4 | 00001122222233444446678 99
  5 | 00

>
```

Box-Plot

- The boxplot is used to summarize data using the five number summary.
- From the display one can check easily if the data is symmetric or has suspected outliers.
- Its simplicity is its feature.

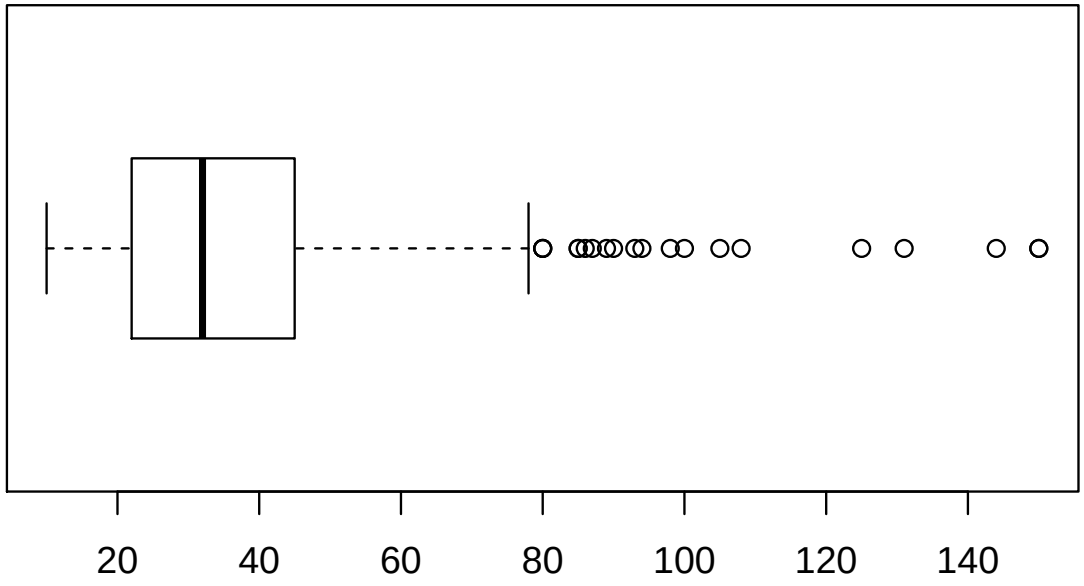
Box-Plot

```
> boxplot(Scores)
```



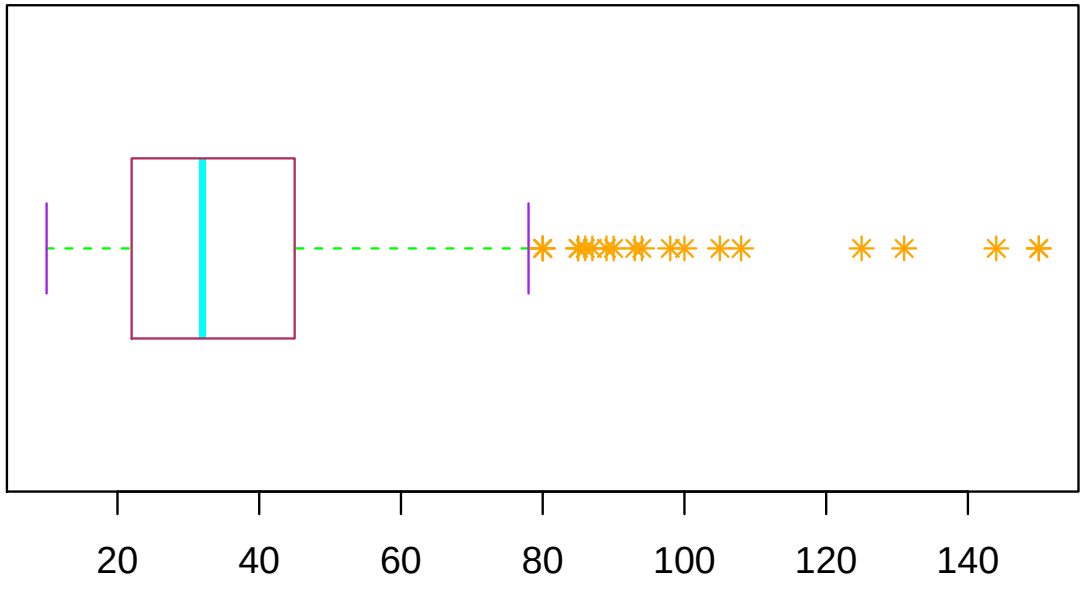
Box-Plot

```
> library(UsingR)
> boxplot(kid.weights$weight, horizontal=TRUE)
```



Box-Plot

```
> library(UsingR)
> boxplot(kid.weights$weight, horizontal=TRUE,
+ whiskcol="green",
+ boxcol="maroon",
+ staplecol="purple",
+ outcol="orange",
+ outpch=8,
+ outlwd=1,
+ medcol="cyan")
```



Box-Plot

Arguments for the appearance of line/symbol elements (for more information see `?par` in R):

`lty`: line style (e.g.,= 0 to remove the line,= 1 for solid line)

`lwd`: line width

`col`: line color

`pch`: symbol style

`lwd`: line width

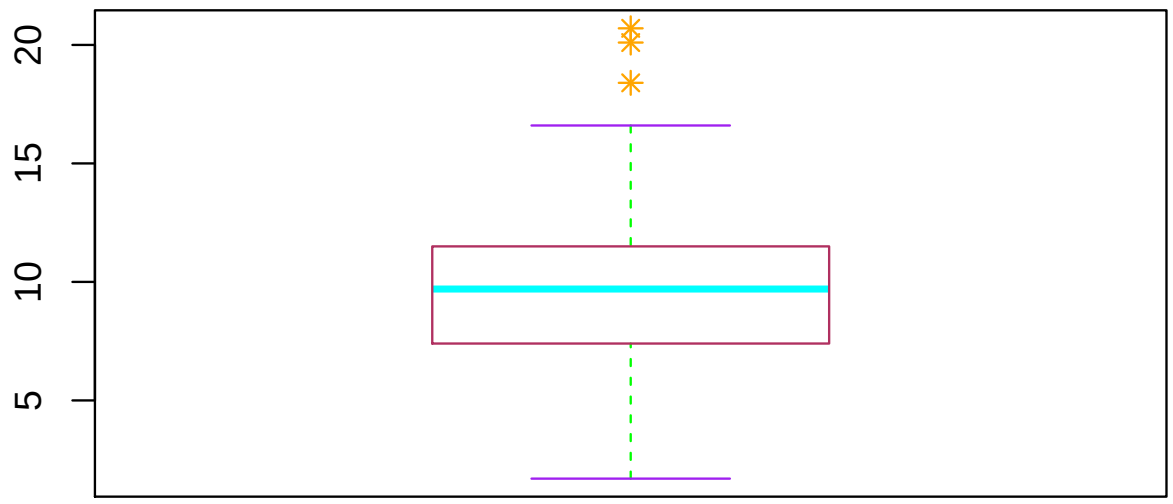
`cex`: size of symbol

`col`: color

`bg`: background color

Box-Plot

```
> boxplot(airquality$Wind, whiskcol="green", boxcol="maroon",
+ staplecol="purple", outcol="orange", outpch=8,outlwd=1,medcol="cyan"
+ )
```



Box-Plot

```
> boxplot(airquality, whiskcol="green", boxcol="maroon",
+ staplecol="purple", outcol="orange", outpch=8,outlwd=1,medcol="cyan"
+ )
```