

Finding Basic Statistics Using R

1. Open the R program. The first greater-than sign (>) appears automatically. If you have written a complete command, when the "ENTER" key is pressed, R will run the command and then a greater than sign will appear on the next line. If the command is not complete and the "Enter" key is pressed, a plus sign (+) will appear on the next line.

2. Assign a name to your data values. Use commas between the numbers.

```
> data <- c(list of numbers)
```

3. Find the mean of your data.

```
> mean(data)
```

4. Find the minimum of your data.

```
> min(data)
```

5. Find the first quartile of data.

```
> quantile(data, 0.25)
```

6. Find the median of your data.

```
> median(data)
```

7. Find the third quartile of your data.

```
> quantile(data, 0.75)
```

8. Find the maximum of your data.

```
> max(data)
```

9. Find the midrange of your data.

```
> (min(data) + max(data))/2
```

10. Find the range of your data.

```
> max(data) – min(data)
```

11. Find the sample standard deviation of your data

```
> sd(data)
```

12. Find the interquartile range of your data.

```
> IQR(data)
```

13. Find the standard error of your data.

```
> sd(data)/sqrt(length(data))
```

14. Find the mode of your data. Sadly, this formula and the next are fairly complicated.

```
> as.numeric(names(which(table(data)==max(table(data))))))
```

15. Find the skewness measure of your data. This formula differs slightly from the one used by Minitab, but it seems to be used more widely than the Minitab formula.

```
> skew -> function(x){m3<-sum((x - mean(x))^3)/length(x)
+ s3<-(sd(x))^3; m3/s3}
> skew(data)
```

Example Triola, *Elementary Statistics*, 2014, p. 751. The data in this example are ages of female actors who won the Oscar for their work in a leading role.

26	25	33	35	35	28	30	29	61	32	33	45
----	----	----	----	----	----	----	----	----	----	----	----

Open the R program and type in the following. R ignores statements following the number symbol (#).

```
> ages <- c(26, 25, 33, 35, 35, 28, 30, 29, 61, 32, 33, 45) #Input the data.
> mean(ages) #Find the mean of the ages.
[1] 34.33333
> min(ages) #Find the minimum of the ages.
[1] 25
> quantile(ages, 0.25) #Find the first quartile of the ages.
25%
28.75
> median(ages) #Find the median of the ages.
[1] 32.5
> quantile(ages, 0.75) #Find the third quartile of the ages.
75%
35
> max(ages) #Find the maximum of the ages.
[1] 61
> (min(ages) + max(ages))/2 #Find the midrange of the ages.
[1] 43
> max(ages) - min(ages) #Find the range of the ages.
[1] 36
> sd(ages) #Find the sample standard deviation of the ages.
[1] 9.902556
> IQR(ages) #Find the IQR of the ages.
[1] 6.25
> sd(ages)/sqrt(length(ages)) #Find the standard error of the ages.
[1] 2.858622
> as.numeric(names(which(table(ages)==max(table(ages))))) #Find the mode.
[1] 33 35
> skew<-function(x){m3<-sum((x-mean(x))^3)/length(x)
+ s3<-(sd(x))^3
+ m3/s3} #Define the skewness function.
> skew(ages) #Find the skewness of the ages.
[1] 1.568824
```