6-14

Data: 2.13, 2.96, 3.02, 1.82, 1.15, 1.37, 2.04, 2.47, 2.60. (n=9)

a

Calculate sample mean and standard deviation by hand

$$\bar{x} = \frac{2.13+2.96+3.02+1.82+1.15+1.37+2.04+2.47+2.6}{9} = \frac{5.09+4.84+1.15+3.41+5.07}{9} = \frac{8.5+4.84+1.15+5.07}{9} = \frac{13.57+5.99}{9} = \frac{19.56}{9} = 2\frac{1.56}{9} = 2.1733$$

$$s^2 = \frac{(2.13-2.1733)^2+(2.96-2.1733)^2+(3.02-2.1733)^2+(1.82-2.1733)^2+(1.15-2.1733^2)+(1.37-2.1733)^2+(2.04-2.1733)^2+(2.47-2.17}{8}$$

$$\frac{(-0.043333)^2+(0.786667)^2+(0.846667)^2+(-0.353333)^2+(-1.023333)^2+(-0.803333)^2+(-0.133333)^2+(0.296667)^2+(0.426667)^2}{8} =$$

$$\frac{0.001877778+0.61884444+0.71684444+0.124844442+1.0472111+0.64534444+0.01777777+0.08801111+0.18204444}{8} =$$

$$\frac{0.00187489+0.61889689+0.7169+0.12482+1.04714+0.64529+0.0177689+0.08803+0.18207}{8} = 0.43035$$

$$s^2 = 0.43035$$

$$s = \sqrt{0.43035} = 0.65601$$

b

Calculate sample meadian by hand

2.13, 2.96, 3.02, 1.82, 1.15, 1.37, 2.04, 2.47, 2.60 -> 1.15, 1.37, 1.82, 2.04, 2.13, 2.47, 2.6, 2.96, 3.02 -> 1.37, 1.82, 2.04, 2.13, 2.47, 2.6, 2.96 -> 1.82, 2.04, 2.13, 2.47, 2.6, -> 2.04, 2.13, 2.47, -> 2.13

c

Repeat above using R

```r
data <- c(2.13, 2.96, 3.02, 1.82, 1.15, 1.37, 2.04, 2.47, 2.60)

print(paste("Sample Mean:", mean(data)))
```

```
[1] "Sample Mean: 2.17333333333333"
```

```r
print(paste("Sample Standard Deviation:", sd(data)))
```

```
[1] "Sample Standard Deviation: 0.656010670644922"
```

```r
print(paste("Sample Median:", median(data)))
```

```
[1] "Sample Median: 2.13"
```

6-44

a

Comment on the shape of the distribution

b

Comment on the outliers of the data (DO NOT USE 1.5 IQR Rule)

c

Which do you think has a higher value, sample meann or meadian? (EXPLAIN)

d

Do you think the sample standard deviation is big or small? (EXPLAIN)

e

Find the 3rd quartile and 80th percentile by hand

f

Repeat part e using R

```r
data <- c(450, 450, 473, 507, 457, 452, 453, 1215, 1256, 1145, 1085, 1066, 1111, 1364, 1254,
1575, 1617, 1733, 2753, 3186, 3227, 3469, 1911, 2588, 2635, 2725)

quantile(data, probs=c(0.75,0.8))
```

```
   75%    80%
2249.5 2625.6
```

6-42

a

Use R to find the 5 number summary

```r
str = "680 669 719 699 670 710 722 663 658 634 720 690 677 669 700 718 690 681 702 696 692 69
data = c(as.numeric(strsplit(str, " ")[[1]]))
summary(data)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  634.0   667.8   683.0   686.8   703.2   763.0
```

b

Identify any outliers by hand using the 1.5 IQR Rule

c

Construct a box plot by hand based on your results in pars a and b.

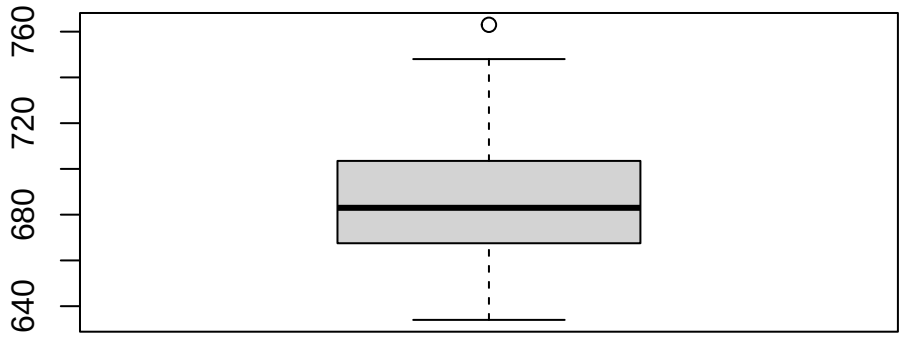
d

Describe the shape of the data distribution based on the boxplot that you created in part c

e

Repeat part c using R

```
boxplot(data)
```



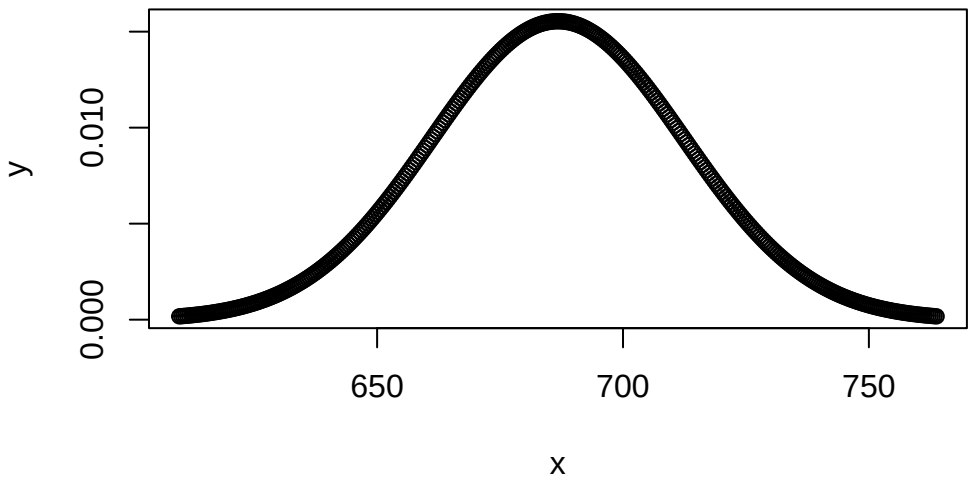
f

Construct a normal probability plot for the data using R

```
mean = mean(data)
std = sd(data)

x = seq(mean-3*std,mean+3*std, by=0.4)
y = dnorm(x, mean, std)

plot(x,y)
```



g

Is it reasonable to assume the data is normally distributed? Why or why not?