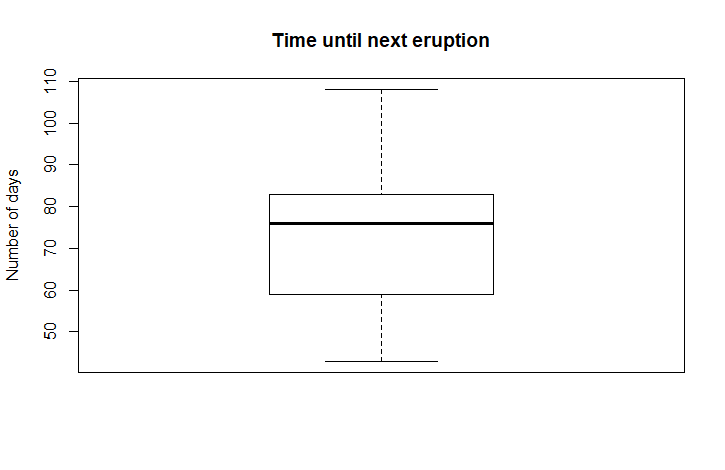
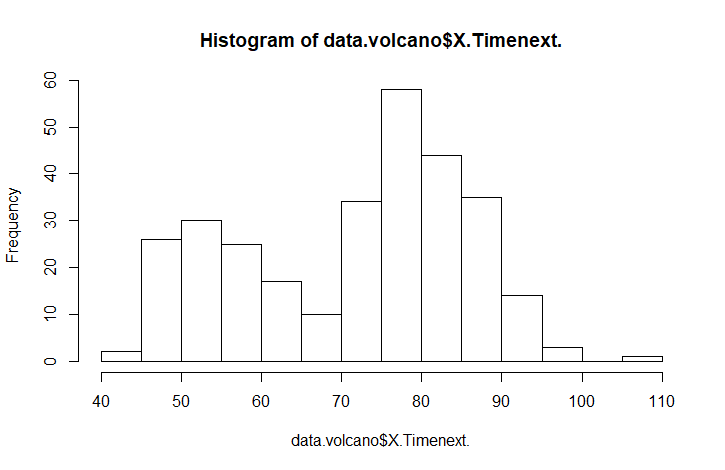
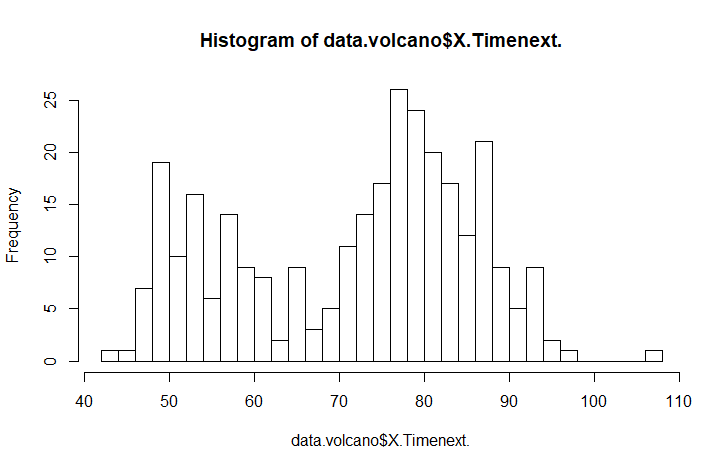
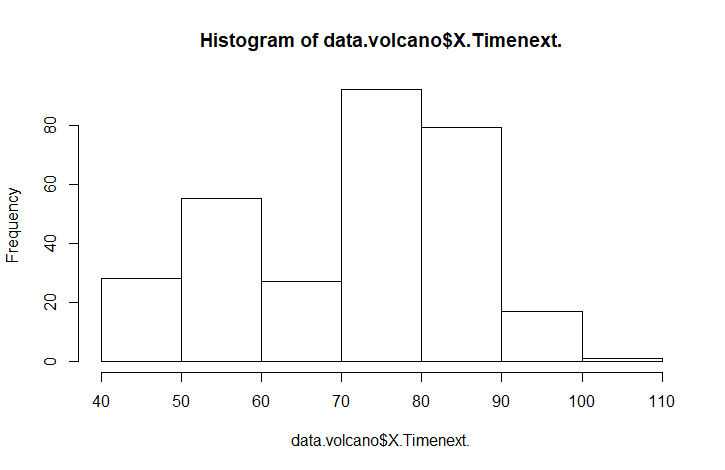
Q6.

a)

b)

By analyzing the data using the boxplot, we can see that it has two peaks, something that we weren’t able to observe in the boxplot.

c)

Here we can notice that with the first, smaller peak that we observed in the initial histogram is not well defined, and there are several individual peaks that create it. The same would be argued for the global peak, which again peaks at around 75, goes down, only to have a second peak near 85.

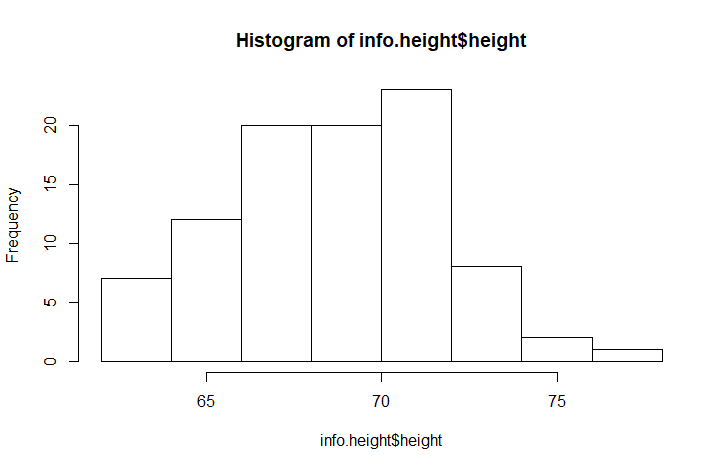
Q7.

a)Gender: Female Male

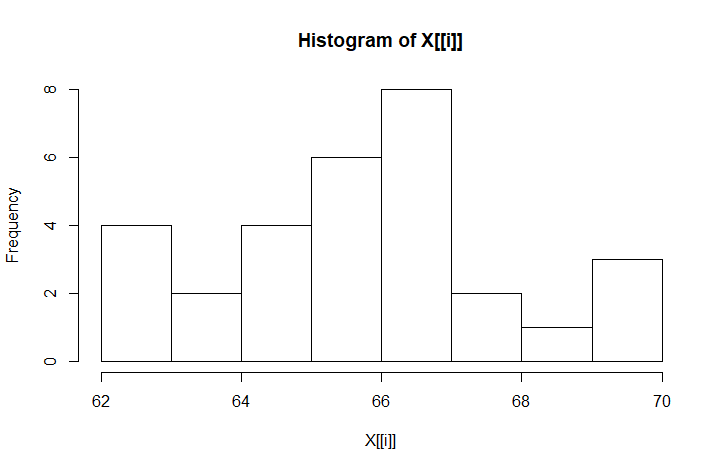
Mean: 66.16333 70.52000

b)

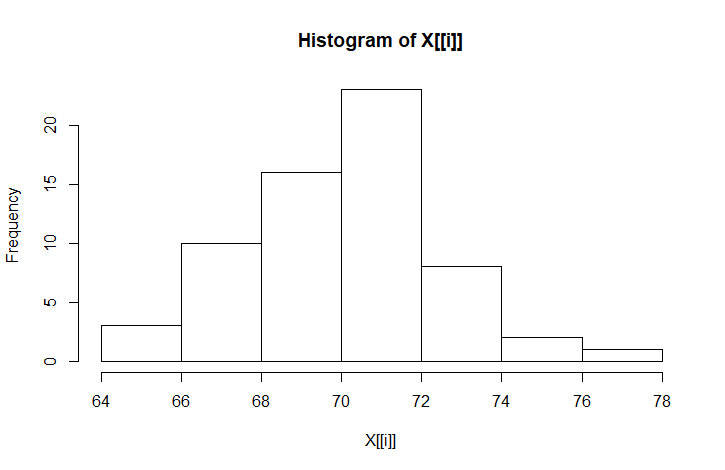
* Std Male = 2.528235
* Std Female = 2.152862
* Std Total = 3.156163

c)

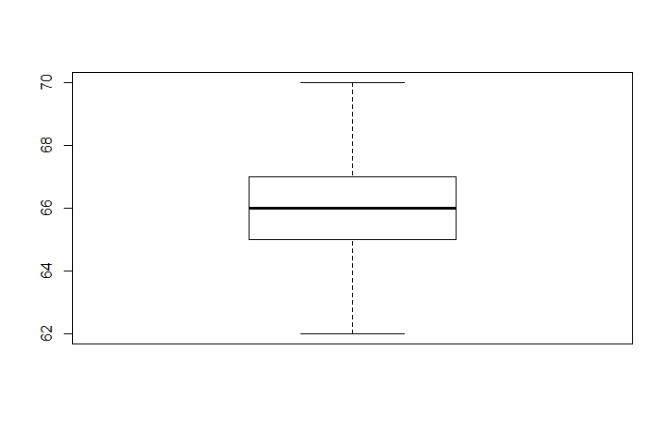
Histogram – Total height



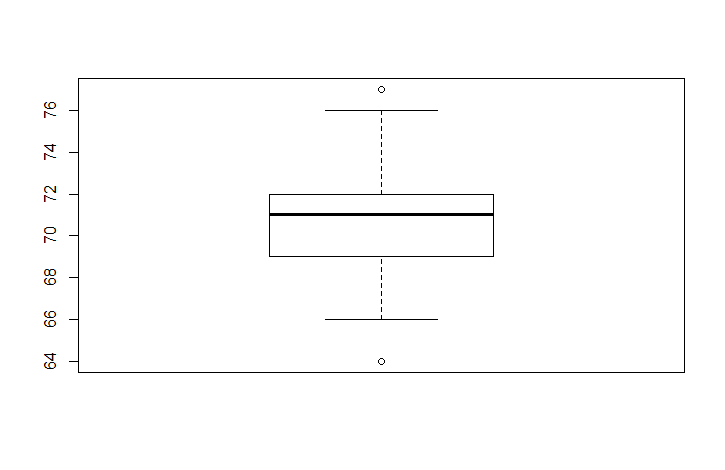
Histogram – Female height

Histogram – Male height

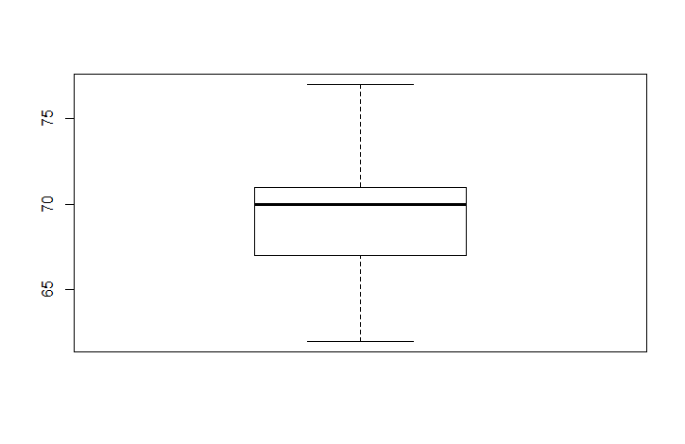
Each plot has something that resembles an normal distribution with a peak and tails. However, some of the plots tend to exhibit higher skewness than others.

d)

Box plot -Females



Box plot Males

Boxplot all.

We can notice that the outliers are determined by the males in the class.

Code for generating graphs:

Q6.

data.volcano = read.delim("oldfaithful.txt", header = TRUE, sep = "\t")

View(data.volcano)

boxplot(data.volcano$X.Timenext., main ="Time until next eruption", ylab = "Number of days")

hist(data.volcano$X.Timenext.)

hist(data.volcano$X.Timenext., breaks=30)

hist(data.volcano$X.Timenext., breaks = 5)

Q7.

info.height <- read.delim("stat431height.txt", sep = "\t")

View(info.height)

hist(info.height$height)

tapply(info.height$height, info.height$gender, mean)

tapply(info.height$height, info.height$gender, sd)

tallness = info.height$height

SD(info.height$height, na.rm = FALSE)

tapply(info.height$height, info.height$gender, hist)

tapply(info.height$height, info.height$gender, boxplot)

boxplot(info.height$height)