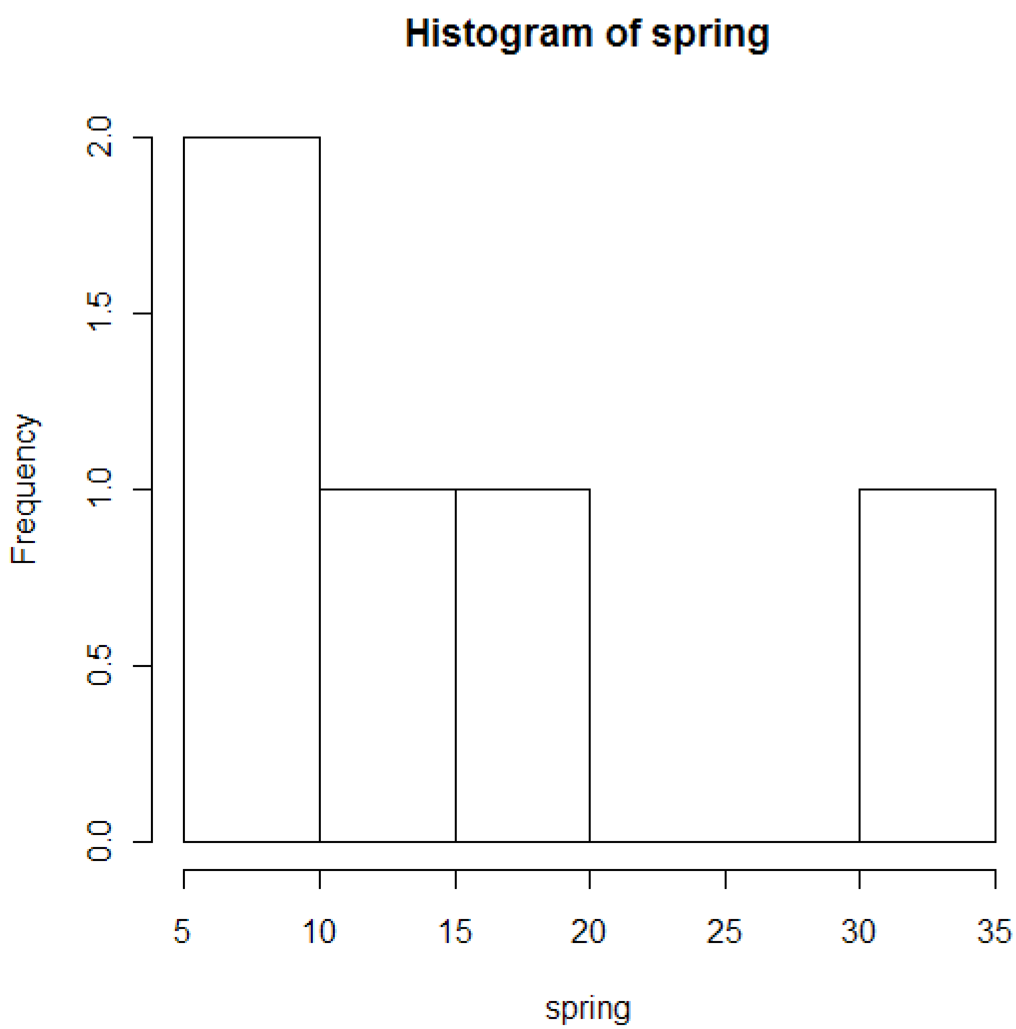
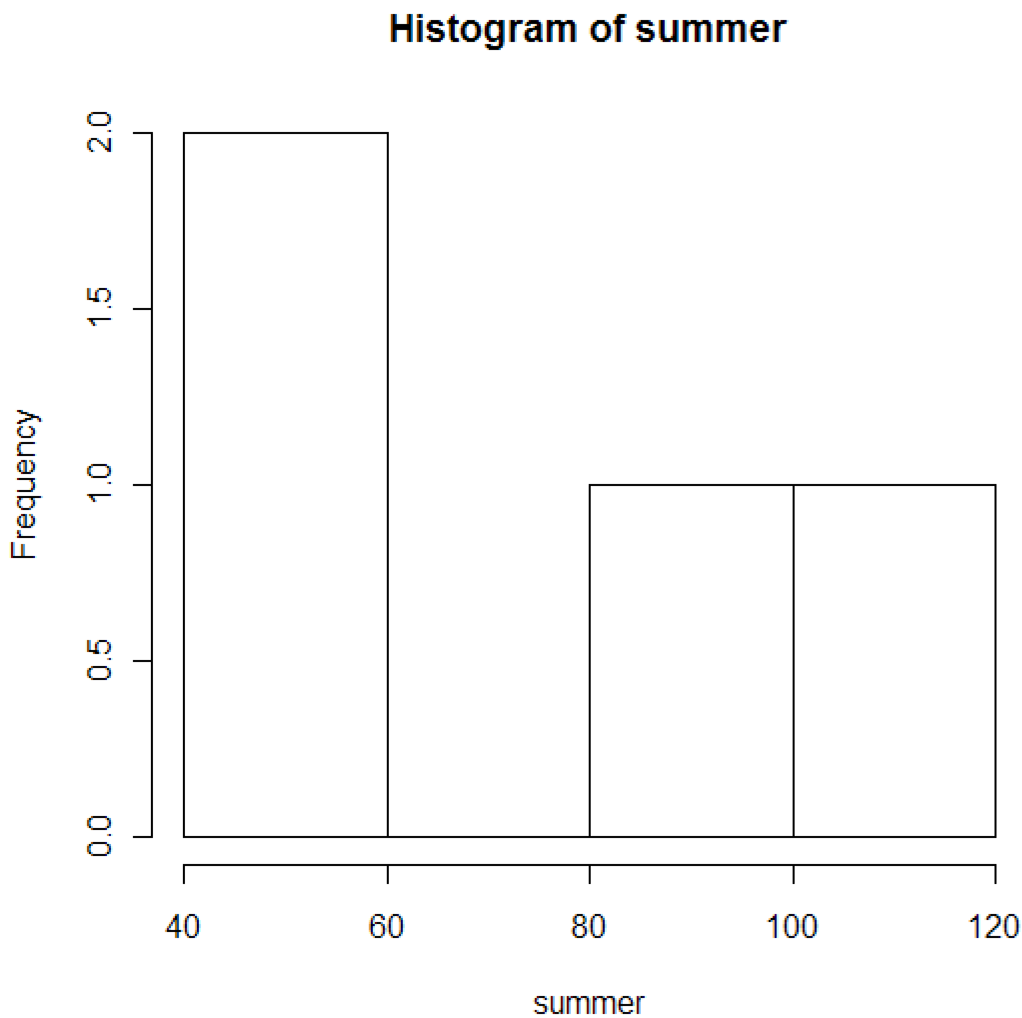
Analyzing alligator data using R
(Exercise 8.91)

```
> spring=c(8,12.1,8.1,18.2,31.7)
> summer=c(102,81.7,54.7,50.7)

#####################
# Stats for spring
> mean(spring)
[1] 15.62

> sd(spring)
[1] 9.902373

>  hist(spring)
```




```
######################
# 95% CI for mean range in spring
```

```
# 1st Check out the t critical value:
> qt(.975,4)
[1] 2.776445

# Here’s the 95% CI
> mean(spring)+c(-1,1)*qt(.975,4)*sd(spring)/sqrt(5)
[1]  3.324579 27.915421

# Of course, R can do all of this for you
> t.test(spring)

        One Sample t-test

data:  spring
t = 3.5272, df = 4, p-value = 0.02429
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
  3.324579 27.915421
sample estimates:
mean of x
    15.62

##############################
# Stats for summer
> mean(summer)
[1] 72.275

> sd(summer)
[1] 24.12998

> hist(summer)
```




```
################################
# 95% CI for mean range in summer
# 1st Check out the t critical value:
> qt(.975,3)
[1] 3.182446

# 95% CI by hand
> mean(summer)+c(-1,1)*qt(.975,3)*sd(summer)/sqrt(4)
[1]  33.87882 110.67118

# 95% CI using R:
> t.test(summer)

        One Sample t-test

data:  summer
t = 5.9905, df = 3, p-value = 0.009314
```

```
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
  33.87882 110.67118
sample estimates:
mean of x
  72.275

##################
# Let's compare the two samples

# Difference in sample means summer – spring is
> mean(summer)-mean(spring)
[1] 56.655

# The pooled standard deviation:
> sqrt((4*9.902373^2 + 3*24.12998^2)/7)
[1] 17.48058

# The critical value for a 95% CI:
> qt(.975,7)
[1] 2.364624

# Here's the 95% CI for mean (summer- spring):
> 56.655 + c(-1,1)*qt(.975,7)*17.48*sqrt(1/5 + 1/4)
[1] 28.92756 84.38244

# Now we'll get R to do this for us:
> t.test(summer,spring,var.equal=TRUE)

        Two Sample t-test

data:  summer and spring
t = 4.8314, df = 7, p-value = 0.001896
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 28.92663 84.38337
sample estimates:
mean of x mean of y
  72.275    15.620

#########################################
# Checking the assumptions:
```

1. These small sample techniques assume that the data are normal for the summer alligators and for the # spring alligators. Do the histograms of the data support this?

2. This two sample technique requires that the spring alligators were independent of each other, that the summer alligators were independent of each other, and that the spring and summer alligators are also independent of each other (e.g. these are different alligators).

3. Lastly, we assumed that the true standard deviation of the spring alligators was equal to the true standard deviation for the summer alligators.   Do the data support this assumption?

Recall that F-statistic = S^2_{spring}/S^2_{summer} (when we assume that the true standard deviations are equal) has an F distribution ONLY when the data are normal!

Let’s check in R:

```
# Calculate the standard deviaitons
> sd(spring)
[1] 9.902373
> sd(summer)
[1] 24.12998

# Calculate the F-statistic
> F=var(summer)/var(spring)
> F
[1] 5.937932

# How likely is it to get a variance that is about 6 times the other or more when the true standard
# deviations are equal?
> 1-pf(F,4,3)   % Calculate an upper tail prob
[1] 0.08756749
```