

T tests and Normality

Lynn Waterhouse

March 22, 2017

Checking for Normality

There are 5 major tests used: -Shapiro-Wilk W test -Anderson-Darling test -Martinez-Iglewicz test -Kolmogorov-Smirnov test -D'Agostino Omnibus test

Note: Power of all is weak if $N < 10$

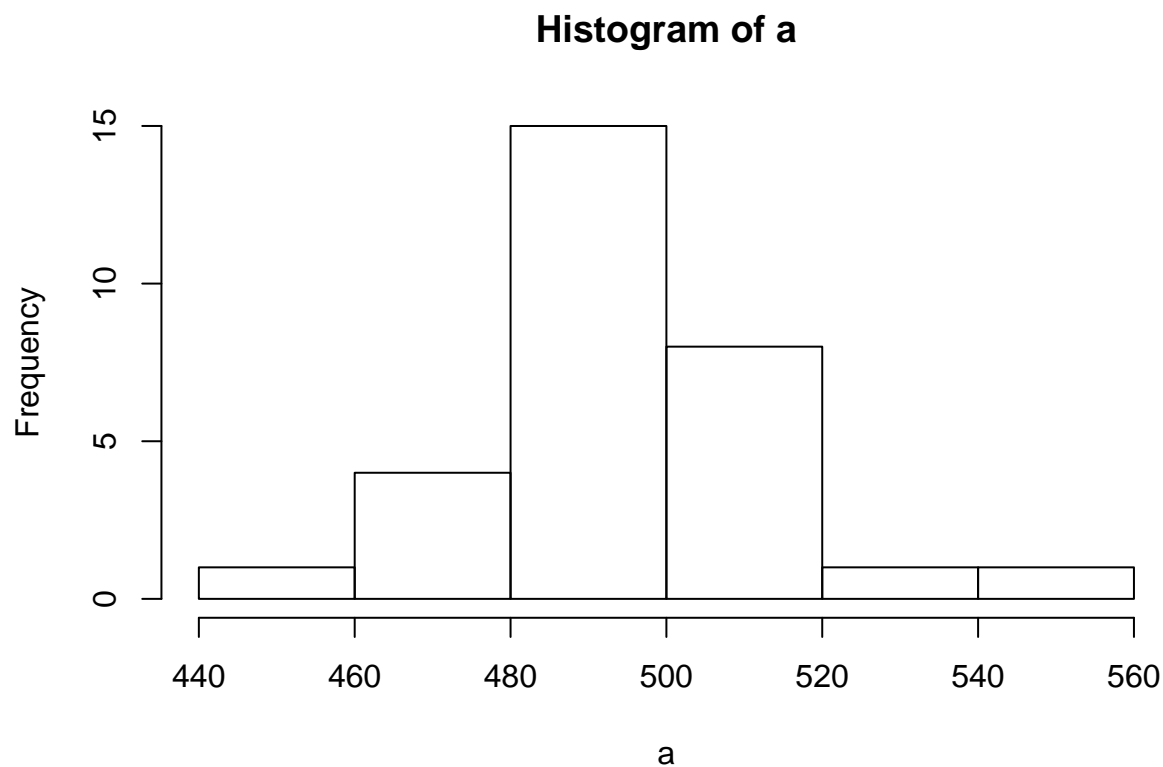
Shapiro Wilk W Test

Recall we fail to reject the null hypothesis that the distribution is normal if the calculated p-value is > 0.05 .

```
set.seed(1234)
a<-floor(rnorm(30,500,20))
a
```

```
## [1] 475 505 521 453 508 510 488 489 488 482 490 480 484 501 519 497 489
## [18] 481 483 548 502 490 491 509 486 471 511 479 499 481
```

```
hist(a)
```

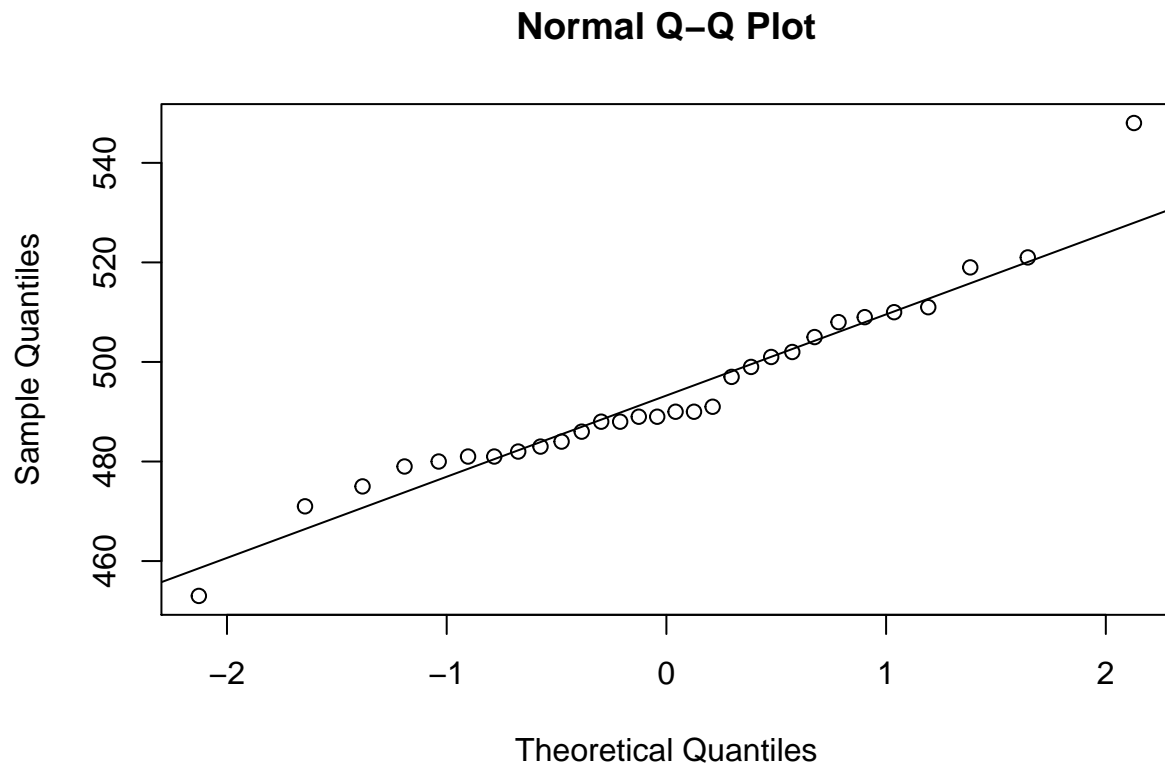


```
shapiro.test(a)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  a  
## W = 0.94795, p-value = 0.1489
```

```
qqnorm(a)
```

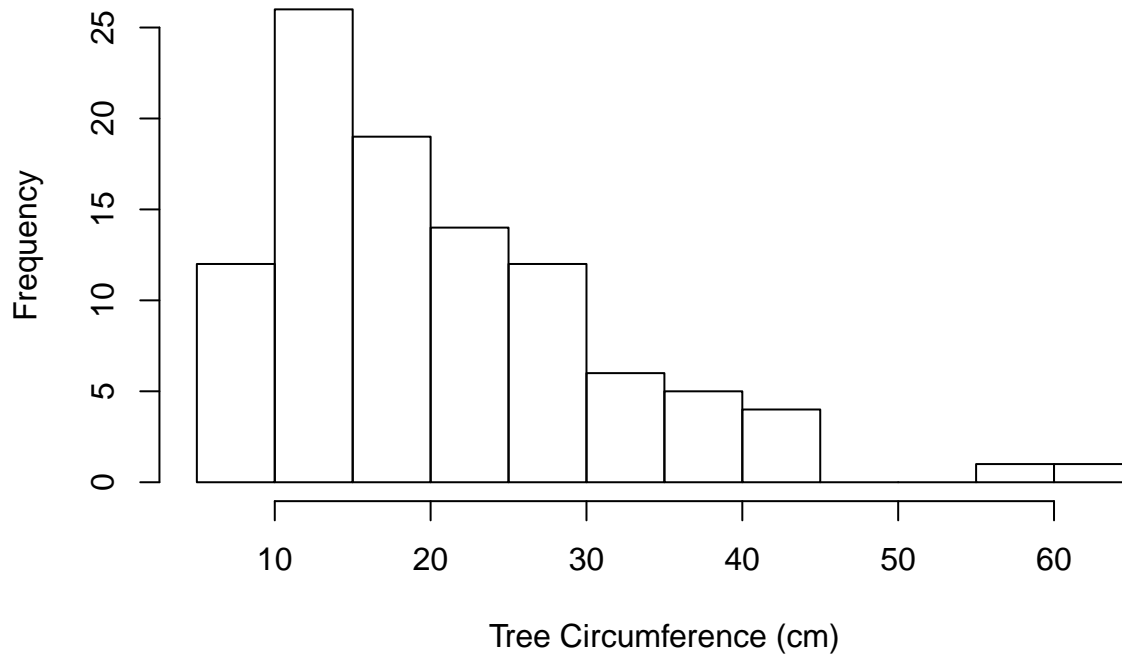
```
qqline(a)
```



Log and Square Root Transformations

Imagine we have collected measurements on the circumference of trees in a forest. There are many young trees, so we have small circumferences. There are a few older trees, with larger circumferences. If we look at the data in R we can look at a histogram of the circumferences.

Histogram



```
## [1] 21.32
## [1] "numeric"
## [1] 19
## [1] 10.72134
```

We can also look at the mean, mode, median, and standard deviation.

```
mean(dat1)
```

```
## [1] 21.32
```

```
mode(dat1)
```

```
## [1] "numeric"
```

```
median(dat1)
```

```
## [1] 19
```

```
sd(dat1)
```

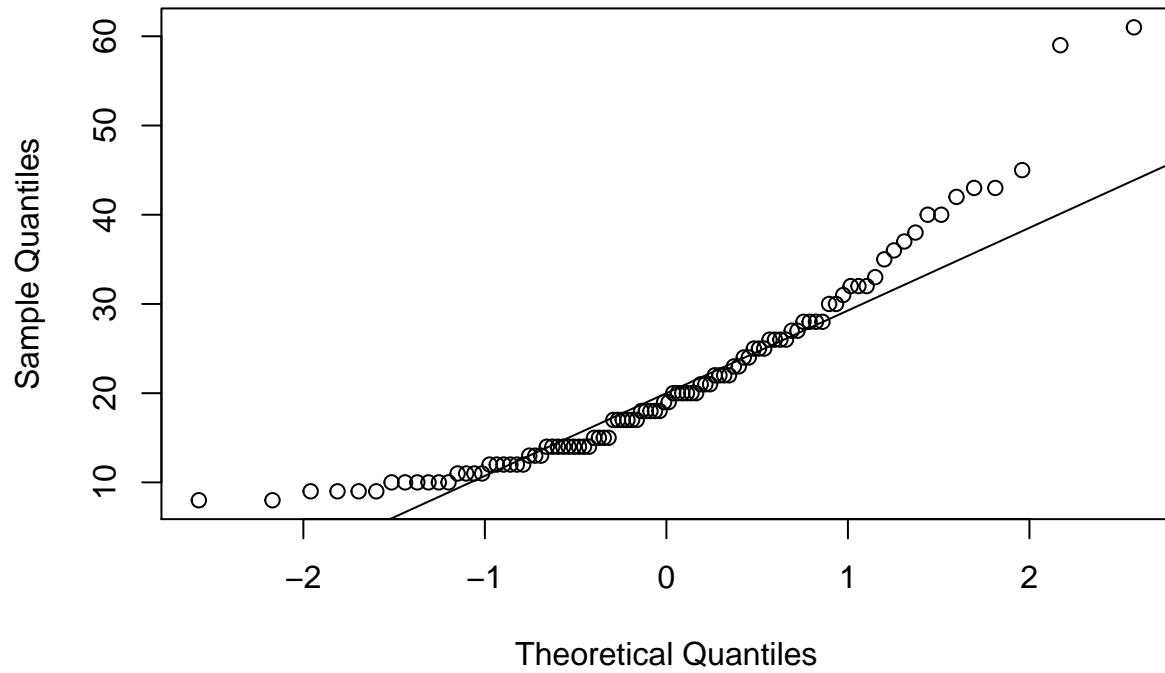
```
## [1] 10.72134
```

We can visually (qualitatively) evaluate normality by looking at the QQ plot.

```
qqnorm(dat1)
```

```
qqline(dat1)
```

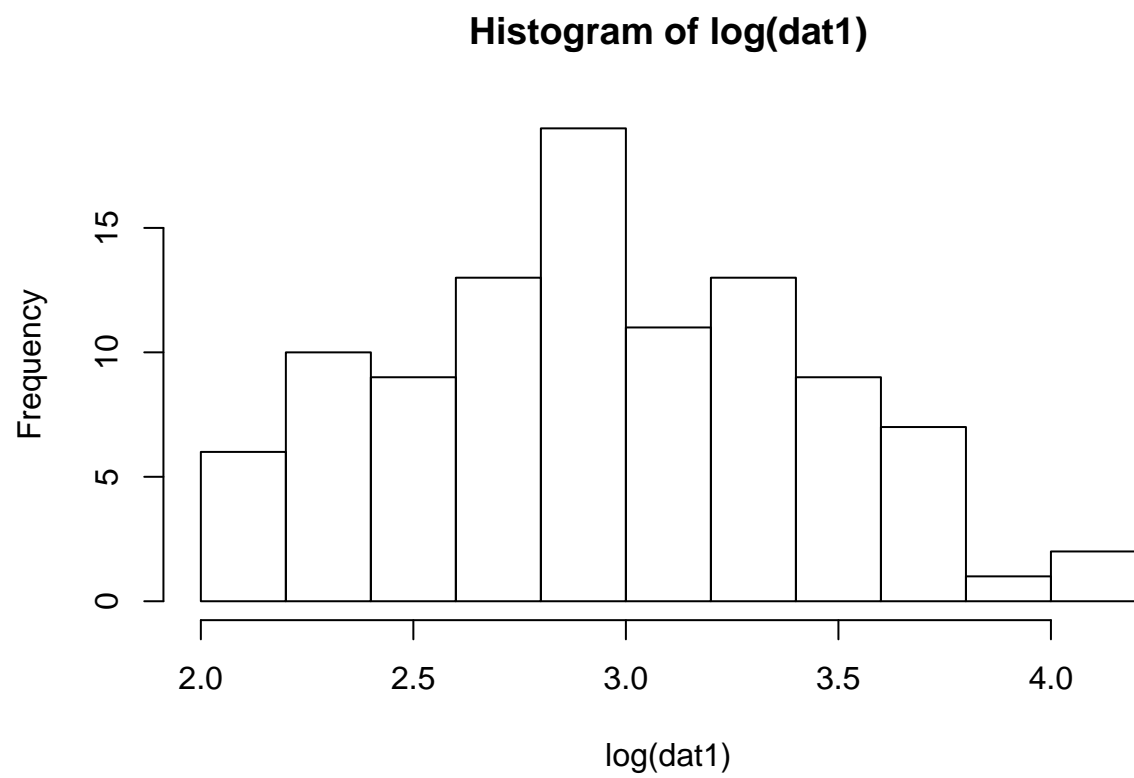
Normal Q-Q Plot



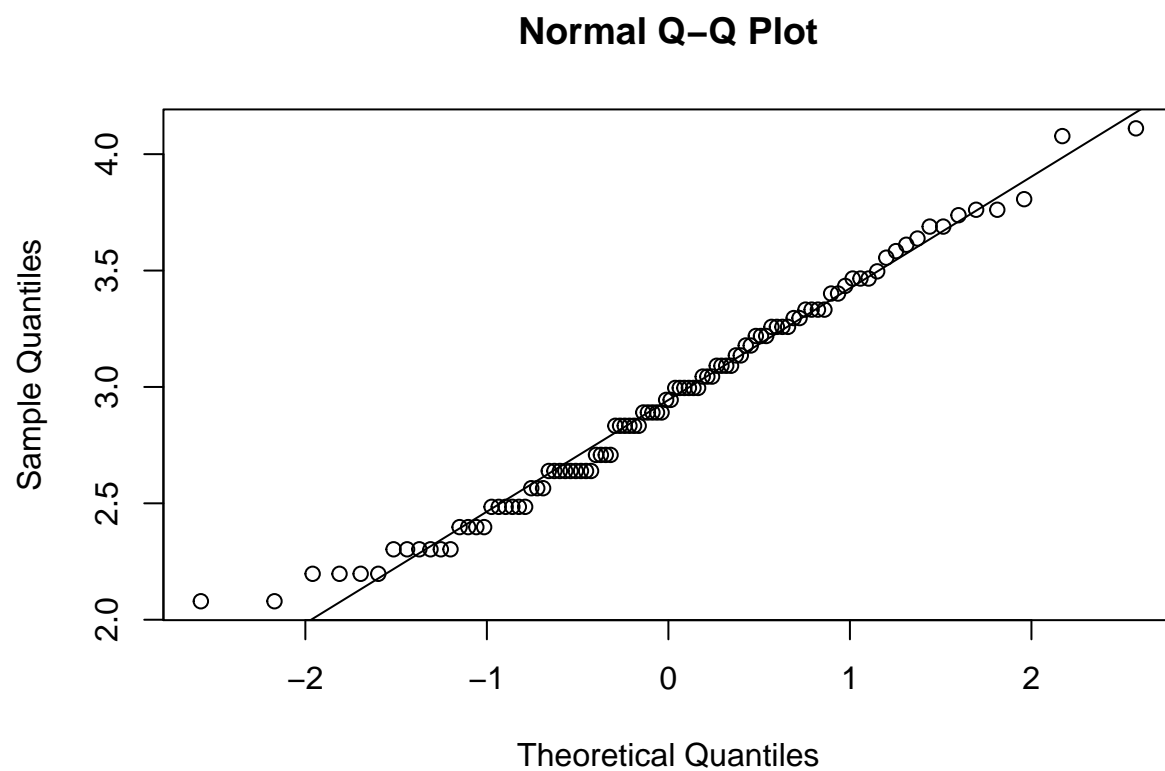
What do you think? Doesn't look so great, right?

So we can try a log transformation.

```
hist(log(dat1))
```



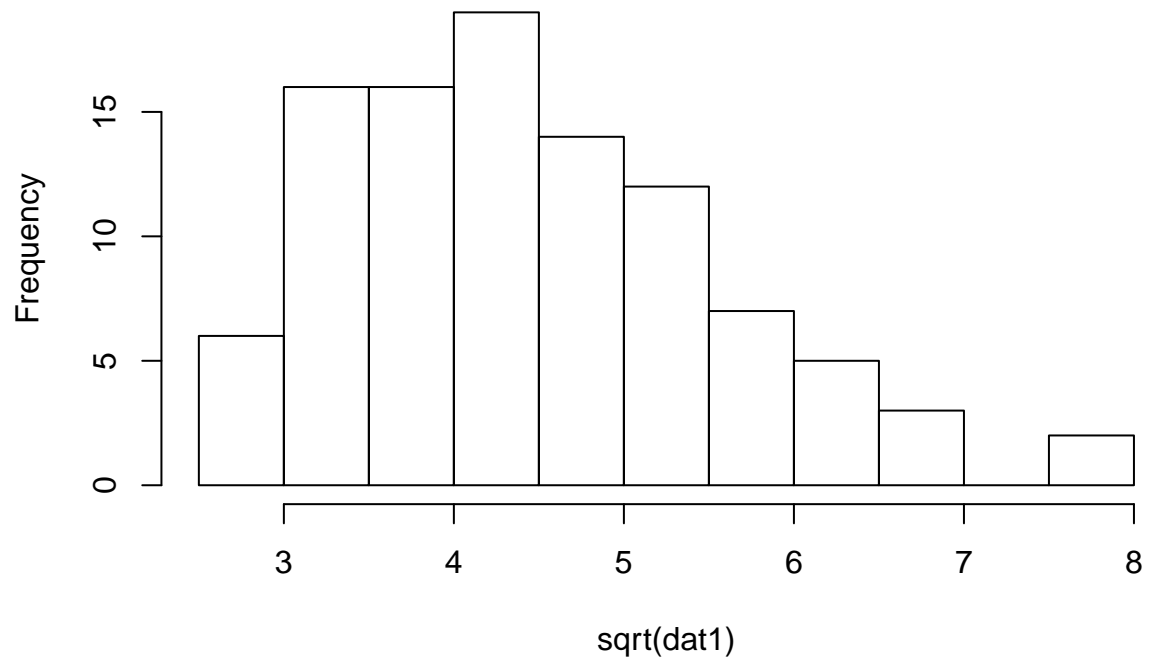
```
qqnorm(log(dat1))  
qqline(log(dat1))
```



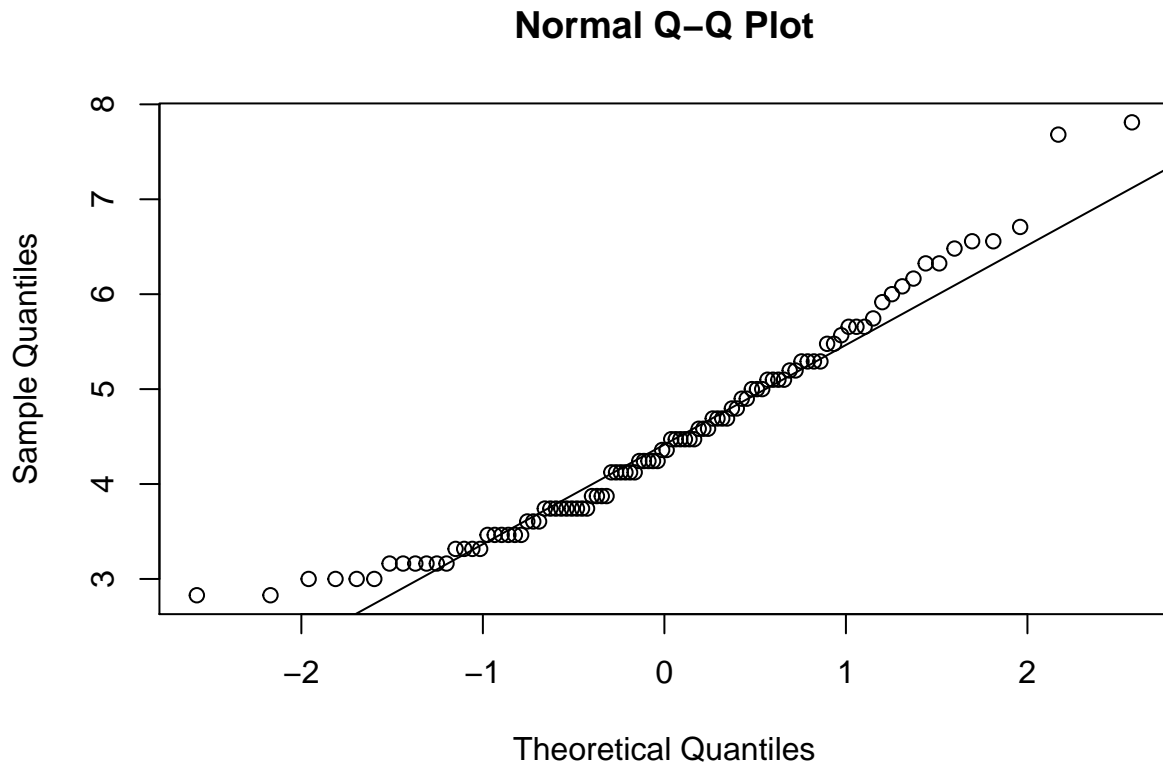
We can also try a square root transformation.

```
hist(sqrt(dat1))
```

Histogram of sqrt(dat1)



```
qqnorm(sqrt(dat1))  
qqline(sqrt(dat1))
```



Which did you think looks better? Why?

In fisheries, log transformations and square root transformations tend to be the most common transformations. This is because with catch data we often have lots of low catch days and then some very high catches.

Cautionary warning when using normality tests

Null is that distribution is normal.

When the sample size is small, even big departures from normality are not detected, and when your sample size is large, even the smallest deviation from normality will lead to a rejected null.

Look at a small sample size, non-normal.

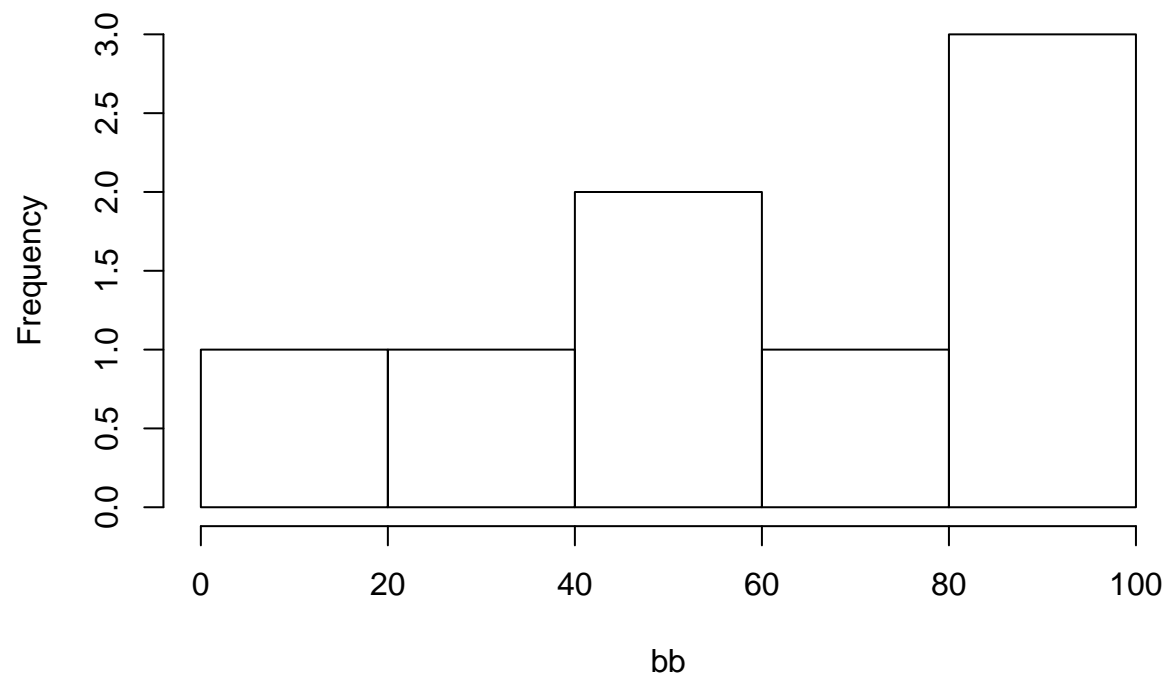
```
set.seed(123)
bb<-runif(8,1,100)
```

```
#look at shapiro.test
shapiro.test(bb)
```

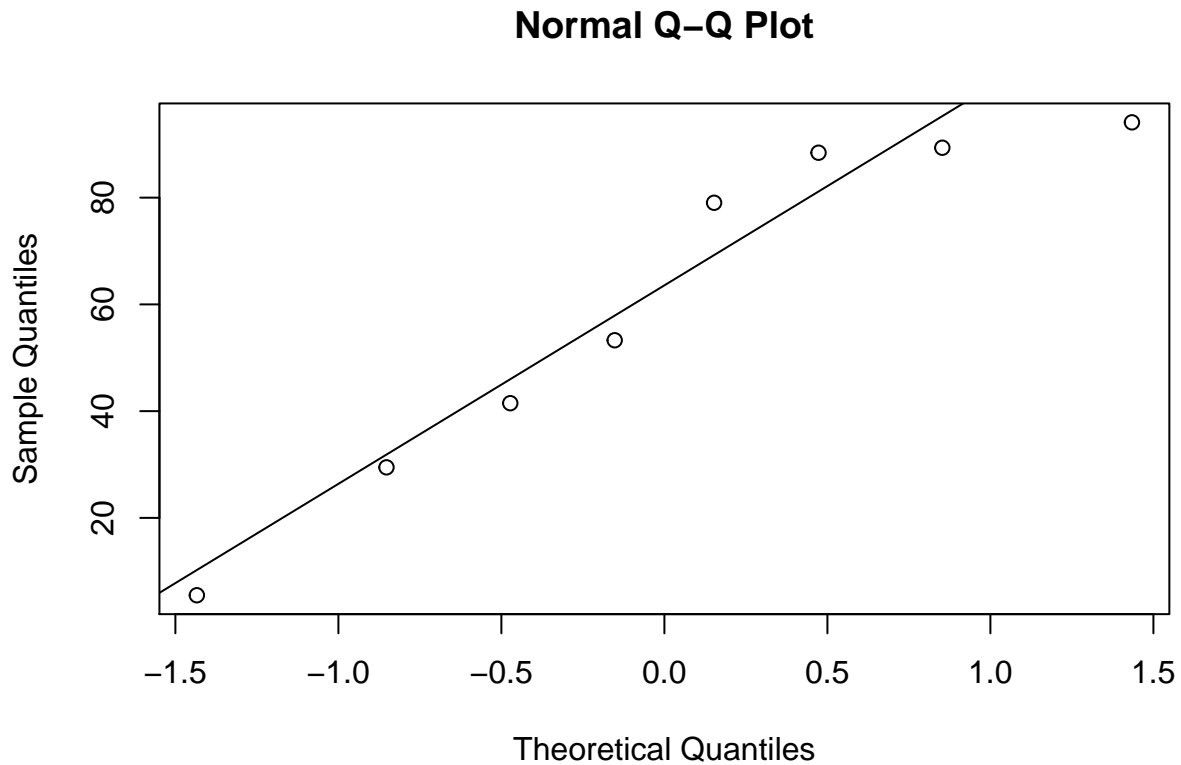
```
##
##  Shapiro-Wilk normality test
##
## data:  bb
## W = 0.90328, p-value = 0.3091
```

```
#look at histogram
hist(bb)
```


Histogram of bb



```
#check out the qq-plots  
qqnorm(bb)  
qqline(bb)
```



T-tests

One Sample T-test

A one sample t-test can be used to compare a sample average to some constant.

The mean weight of abalone sold at market in a package is listed as 500g, you want to check if the abalone sold in the market have a mean of 500g. You have a sample of 30 abalone. The mean is 493.6667g and the sample standard deviation 18.02552g.

Now in R look at the data. We will look at the mean and standard deviation. We will look at a histogram to qualitatively evaluate normality.

```
set.seed(1234)
a<-floor(rnorm(30,500,20))
a

## [1] 475 505 521 453 508 510 488 489 488 482 490 480 484 501 519 497 489
## [18] 481 483 548 502 490 491 509 486 471 511 479 499 481

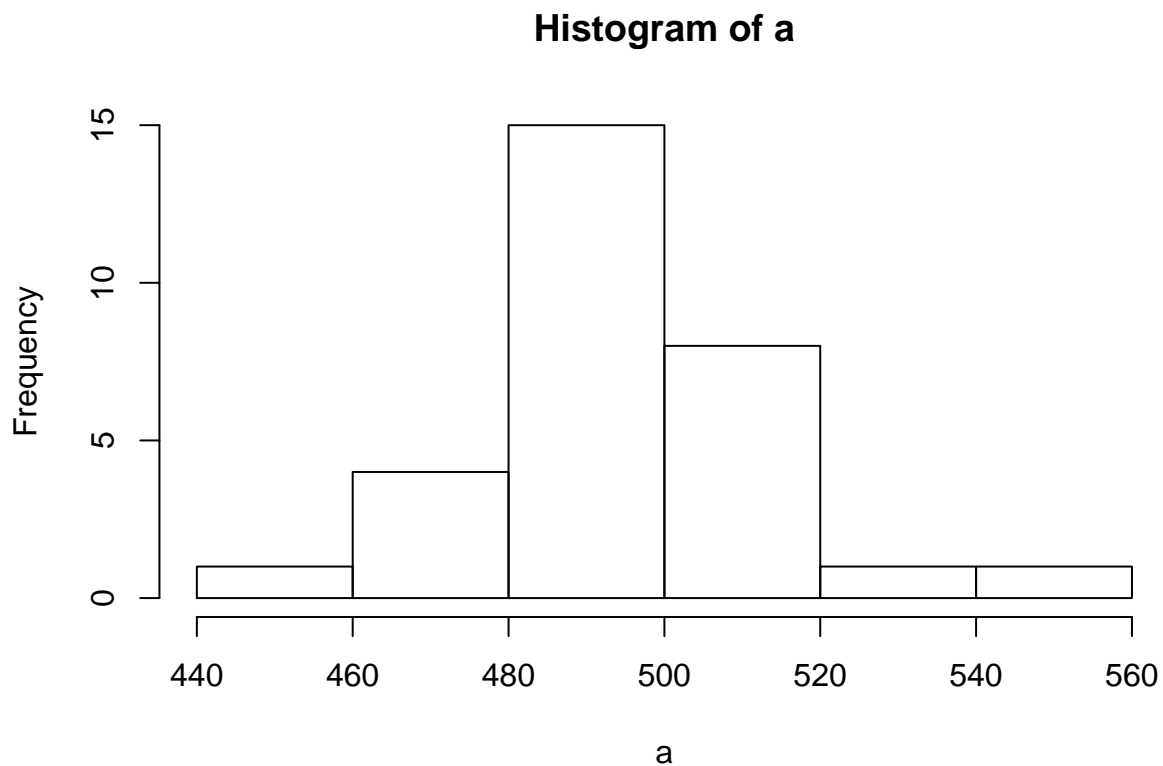
mean(a)

## [1] 493.6667

sd(a)

## [1] 18.02552
```

```
hist(a)
```



What are the hypotheses?

Null: average weight = 500 Alternative: average weight not equal to 500

Conduct the T-test

```
t.test(a,mu=500)
```

```
##
##  One Sample t-test
##
## data:  a
## t = -1.9244, df = 29, p-value = 0.06416
## alternative hypothesis: true mean is not equal to 500
## 95 percent confidence interval:
##  486.9358 500.3975
## sample estimates:
## mean of x
##  493.6667
```

Say I am worried about being ripped off by the store so I go and buy 1000 bags of abalone. I am only concerned with if the store is giving me less than 500g of abalone.

H0-Null : mean \geq 500 Ha- Alternative: mean $<$ 500

```
t.test(a,mu=500,alternative="less")
```

```
##
## One Sample t-test
##
## data: a
## t = -1.9244, df = 29, p-value = 0.03208
## alternative hypothesis: true mean is less than 500
## 95 percent confidence interval:
##      -Inf 499.2585
## sample estimates:
## mean of x
## 493.6667
```

Calculating the 95% Confidence interval by hand

Using the mean, standard deviation, and sample size we can also calculate the 95% Confidence interval by hand. This is for the 2 sided hypotheses. Null: mean = 500 Alternative: mean not equal 500

```
## [1] 486.9358
## [1] 500.3975
## [1] 487.0847
## [1] 500.2487
```

If the confidence interval contains 500 it is like having a p-value >0.05.

If the confidence interval does not contain 500 it is like having a p-value <=0.05

Paired T-tests

We are curious if white seabass have a deep tail. We can tell measure the “deepness” of the tail by looking at the difference between total length of the fish and standard length. We consider a tail deep if it is more than 1/6 of the body length of the fish.

Look at the data

```
##      wsb.TL wsb.SL
## 1      317.8 284.8
## 2      488.0 434.4
## 3      359.1 312.1
## 4      520.2 457.5
## 5      539.8 433.2
## 6      235.5 199.7
## 7      399.6 350.0
## 8      523.4 430.0
## 9      407.5 339.0
## 10     375.2 308.9
## 11     545.3 444.0
## 12     374.1 314.8
## 13     450.4 379.0
## 14     414.7 347.1
## 15     255.0 207.9
## 16     525.9 428.0
```

```
## 17 303.7 250.0
## 18 234.3 198.4
## 19 331.5 274.0
## 20 544.5 482.3
## 21 522.4 420.3
## 22 455.6 384.6
## 23 437.8 385.2
## 24 558.1 453.3
## 25 442.9 379.2
## 26 460.9 378.2
## 27 405.0 329.2
## 28 422.0 369.4
## 29 318.3 283.1
## 30 270.0 226.1
```

```
#read in the dataset
#read in wsb.dat.csv
```

```
#we are interested in if the tail is longer than 1/6
#multiplied by the average body length (deep tail)
```

```
# null hypothesis- difference <= 1/6*avg body length
# alternative hyp- difference > 1/6*avg body length
#set working directory
wsb.dat<-read.csv("wsb.dat.csv") #read in the data
head(wsb.dat) #look at first 6 rows
```

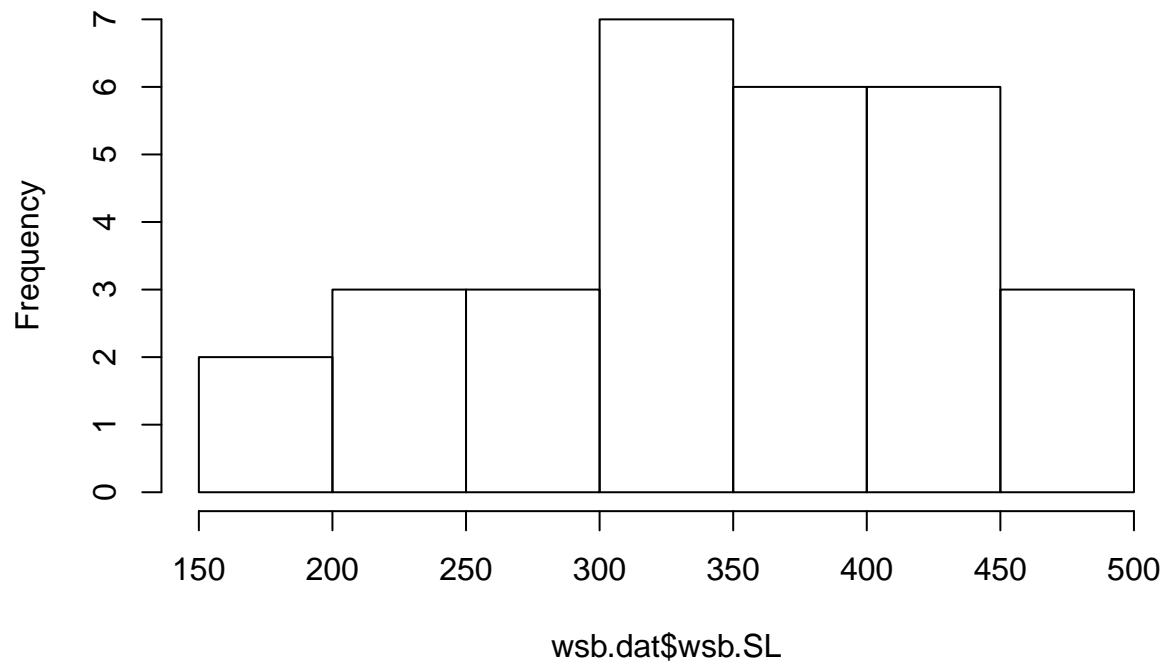
```
##      wsb.TL wsb.SL
## 1  317.8  284.8
## 2  488.0  434.4
## 3  359.1  312.1
## 4  520.2  457.5
## 5  539.8  433.2
## 6  235.5  199.7
```

```
summary(wsb.dat)
```

```
##      wsb.TL      wsb.SL
## Min.   :234.3  Min.   :198.4
## 1st Qu.:338.4  1st Qu.:290.8
## Median :418.4  Median :359.7
## Mean   :414.6  Mean   :349.5
## 3rd Qu.:512.1  3rd Qu.:426.1
## Max.   :558.1  Max.   :482.3
```

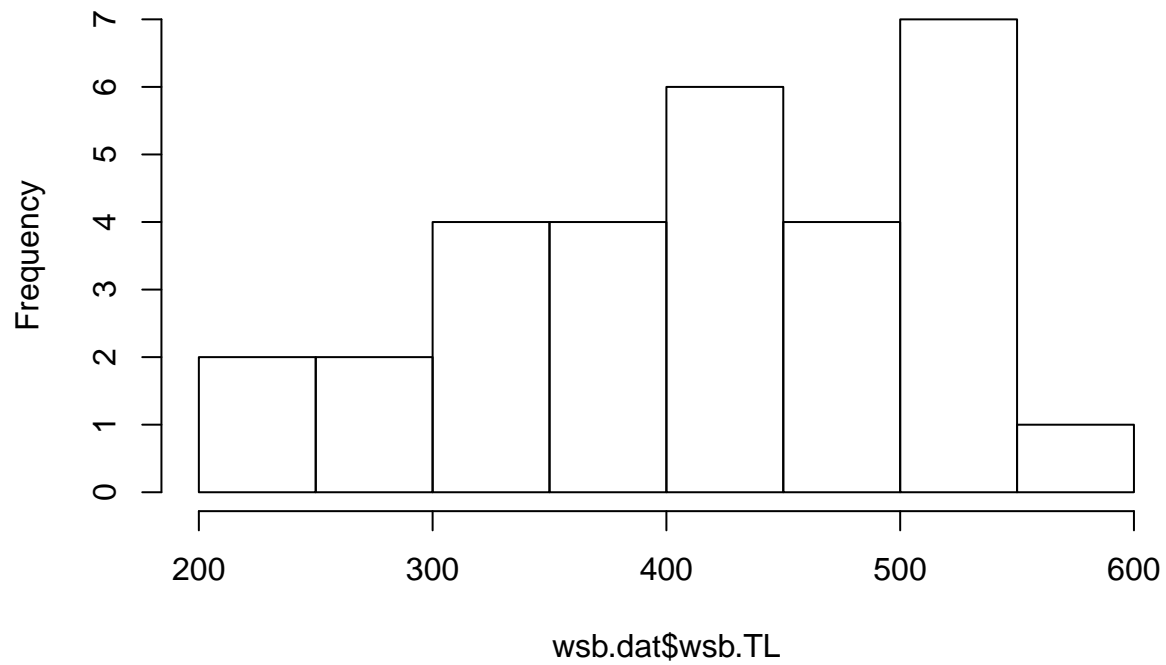
```
hist(wsb.dat$wsb.SL)
```

Histogram of wsb.dat\$wsb.SL



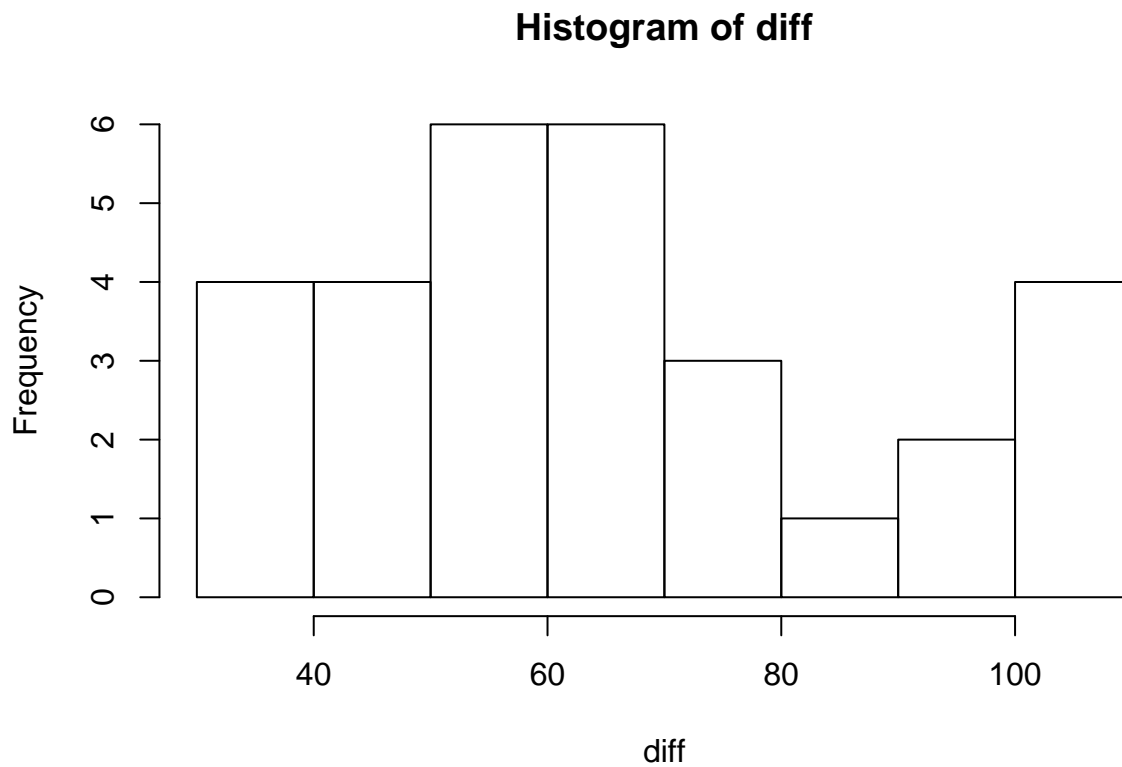
```
hist(wsb.dat$wsb.TL)
```

Histogram of wsb.dat\$wsb.TL



Make a new variable for the difference between Total length and Standard length. Look at a histogram of the data

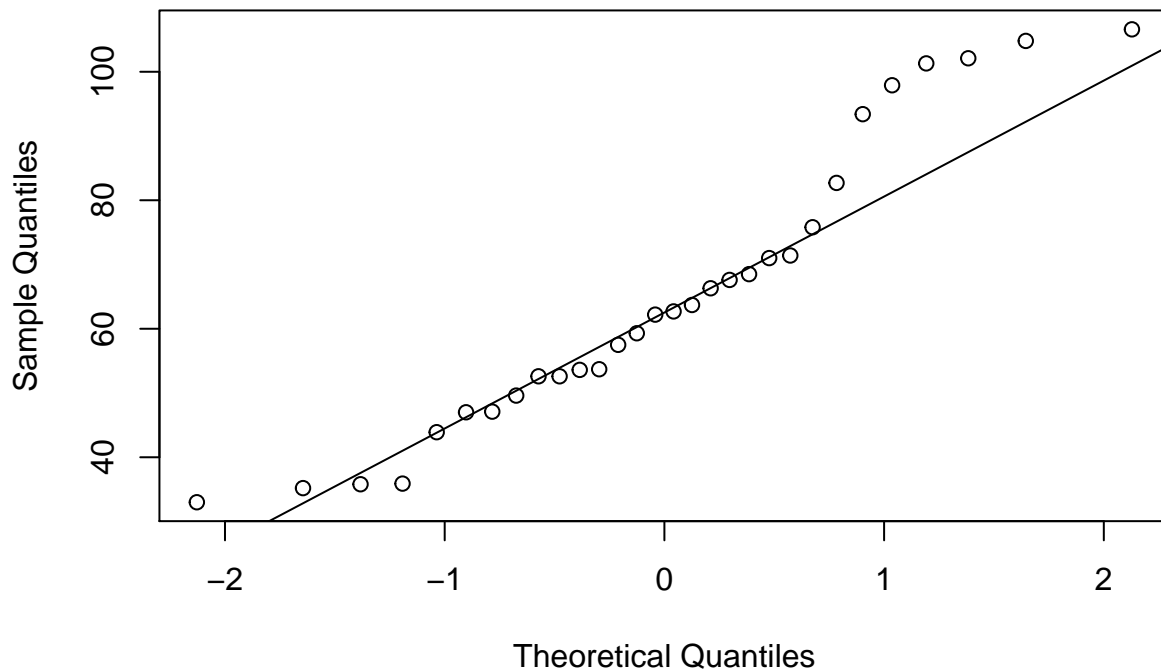
```
diff<-wsb.dat$wsb.TL-wsb.dat$wsb.SL  
hist(diff)
```



In order to use the paired t-test the data (differences) need to be normal. Let's evaluate the normality. We can do this 2 ways. Using the qqnorm plot and also with shapiro.wilk.

```
qqnorm(diff)
qqline(diff)
```


Normal Q-Q Plot



```
shapiro.test(diff)
```

```
##  
##  Shapiro-Wilk normality test  
##  
## data:  diff  
## W = 0.93054, p-value = 0.05075
```

Ok, so what do we conclude?

Now, let's do the t-test. We treat it the same as the one sample t-test. We are interested in if the differences is more than $1/6$ the body length. In order to keep things simple we will calculate $1/6$ body length using the mean total length times $1/6$. This means the null hypothesis will be: that $\text{diff} < 1/6 * \text{mean}(\text{TL})$.

```
null.diff <- (1/6) * mean(wsb.dat$wsb.TL)
```

```
t.test(diff, mu=null.diff, alternative="greater")
```

```
##  
##  One Sample t-test  
##  
## data:  diff  
## t = -0.98177, df = 29, p-value = 0.8328  
## alternative hypothesis: true mean is greater than 69.10278  
## 95 percent confidence interval:  
##  58.33633      Inf  
## sample estimates:  
## mean of x
```

```
##      65.16
```

What do we conclude?

We fail to reject the null hypothesis. We do not have enough statistical evidence to claim that the white seabass have a “deep” tail.

Two sample T-test

There are two types of two sample t-tests. You can easily choose which one is appropriate by looking at the variances of the two samples. If the variances are homogeneous, you can use the two sample t-test with homogeneous variances. If the variances are not homogeneous, you use the two sample t-test for heterogeneous variances.

Testing homogeneity of variances

We test for homogeneity of variances in R using ‘var.test(x,y)’, where x and y are the two samples. It tests the null hypothesis that they are equal (actually that the ratio of the two variances is equal to 1).

Let’s look at an example where we know we have equal variances.

```
x<-rnorm(40,5,2)
y<-rnorm(40,8,2)

var.test(x,y)
```

```
##
## F test to compare two variances
##
## data:  x and y
## F = 0.87515, num df = 39, denom df = 39, p-value = 0.6792
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.4628643 1.6546557
## sample estimates:
## ratio of variances
##      0.8751463
```

One where we do not have equal variances.

```
x1<-rnorm(40,5,2)
y1<-rnorm(40,8,4)

var.test(x1,y1)
```

```
##
## F test to compare two variances
##
## data:  x1 and y1
## F = 0.34158, num df = 39, denom df = 39, p-value = 0.001123
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.180664 0.645841
## sample estimates:
## ratio of variances
##      0.3415849
```

Two Sample T-test with Homogeneous Variances

We want to compare the weights of rays in two sites, B and G. We have the weights recorded in kg in the file "ray.weightskg.csv".

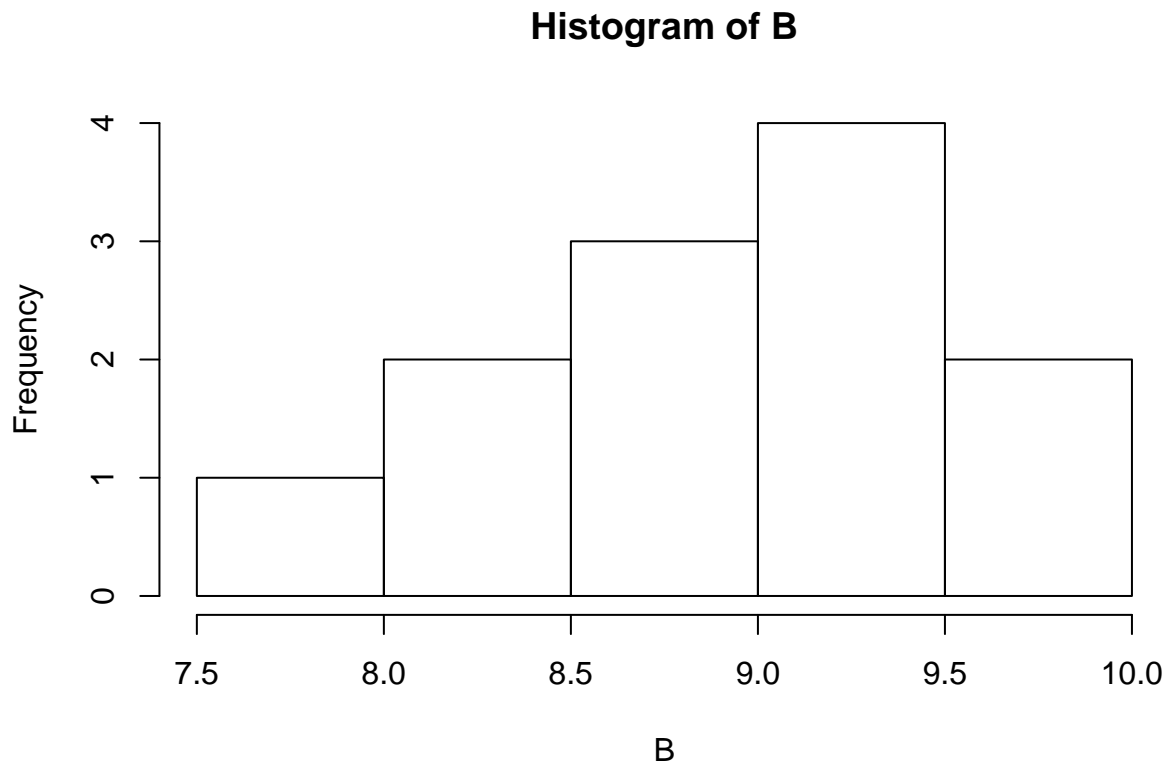
```
ray.weightskg<-read.csv("ray.weightskg.csv")  
head(ray.weightskg)
```

```
##      B      G  
## 1 8.8  9.9  
## 2 8.4  9.0  
## 3 7.9 11.1  
## 4 8.7  9.6  
## 5 9.1  8.7  
## 6 9.6 10.4
```

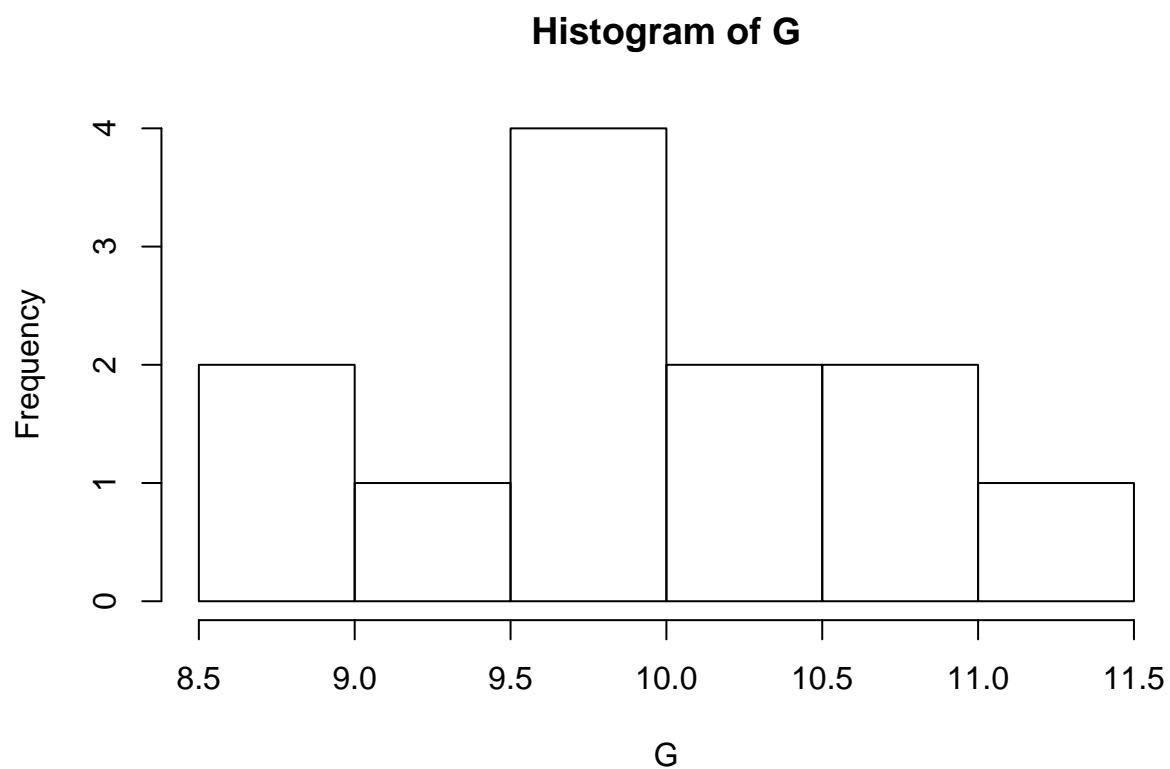
```
B<-ray.weightskg$B  
G<-ray.weightskg$G
```

Look at the data and evaluate normality.

```
hist(B)
```



```
hist(G)
```



```
shapiro.test(B)
```

```
##  
##  Shapiro-Wilk normality test  
##  
## data:  B  
## W = 0.98603, p-value = 0.9977
```

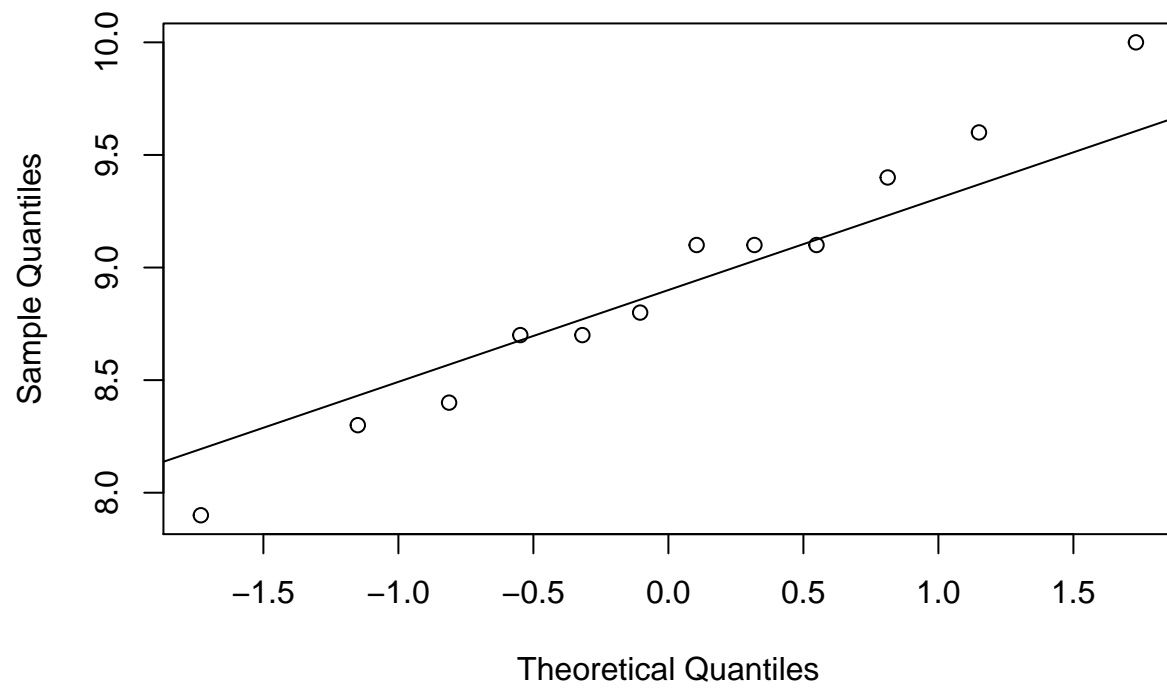
```
shapiro.test(G)
```

```
##  
##  Shapiro-Wilk normality test  
##  
## data:  G  
## W = 0.95376, p-value = 0.6925
```

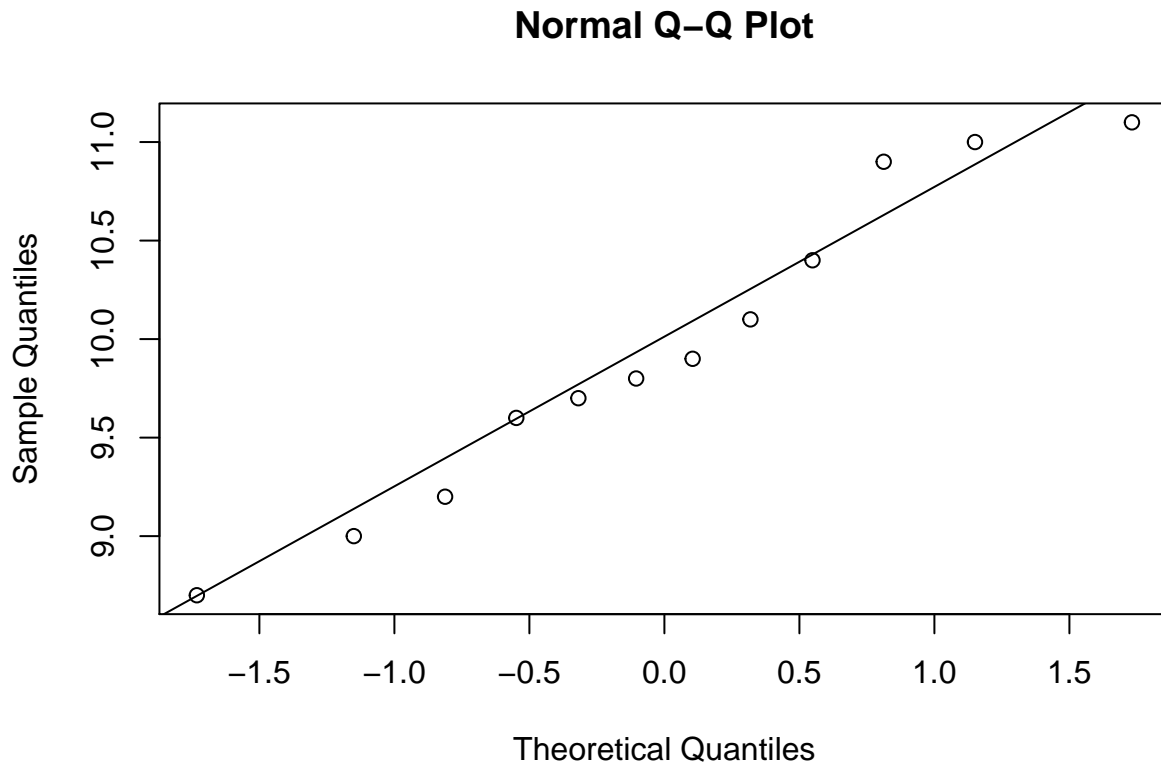
```
qqnorm(B)
```

```
qqline(B)
```

Normal Q-Q Plot



```
qqnorm(G)  
qqline(G)
```



Check if we have homogeneity of variances.

```
var.test(B,G)
```

```
##
##  F test to compare two variances
##
## data:  B and G
## F = 0.55412, num df = 11, denom df = 11, p-value = 0.3419
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.1595198 1.9248590
## sample estimates:
## ratio of variances
##      0.5541237
```

Now, let's do the two sample t-test with homogeneity of variances.

```
t.test(B,G, var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  B and G
## t = -3.6252, df = 22, p-value = 0.001498
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -1.6113715 -0.4386285
```

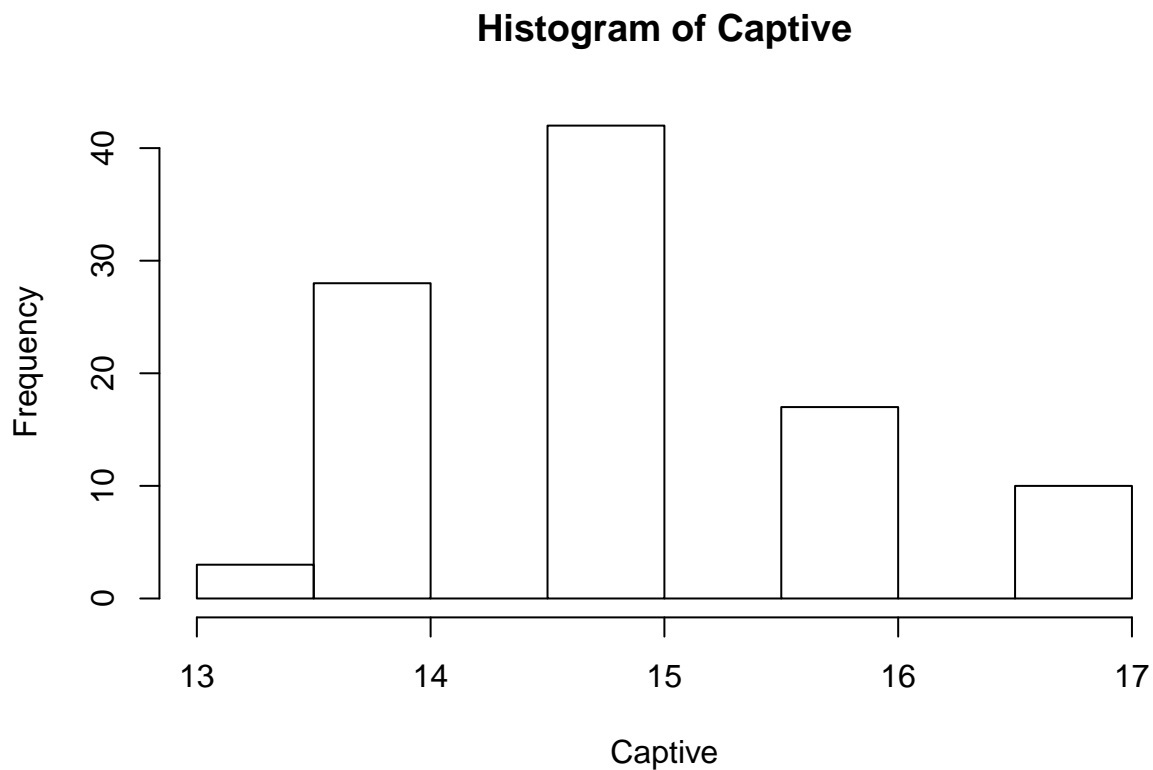
```
## sample estimates:  
## mean of x mean of y  
##      8.925      9.950
```

So what do we conclude?

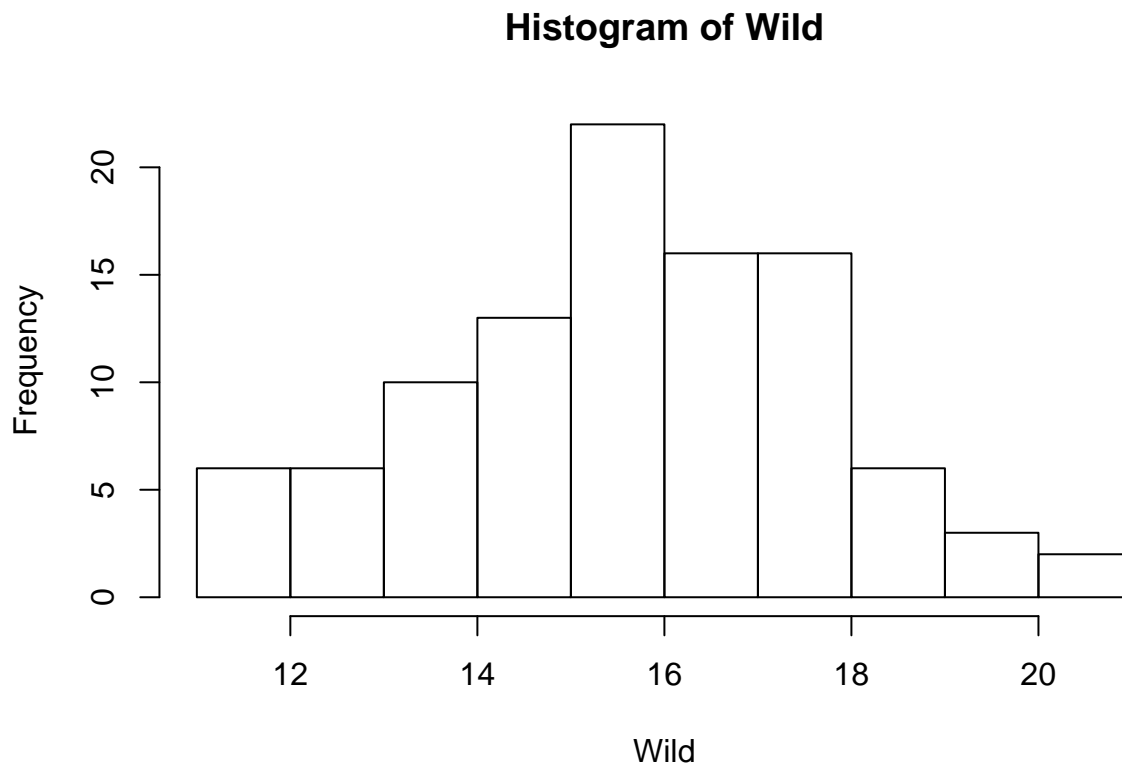
Two Sample T-test with Non-Homogeneous Variances

We want to see if wild caught salmon weigh more than captive salmon. We catch 100 salmon of age 10 from the wild and 100 that are captive, and weigh them all in kg. The data are in “salmon.weight.csv”. Read in the data and evaluate the normality.

```
salmon<-read.csv("salmon.weight.csv")  
Captive<-salmon$Captive  
Wild<-salmon$Wild  
  
hist(Captive)
```



```
hist(Wild)
```



```
shapiro.test(Captive)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: Captive  
## W = 0.88736, p-value = 3.767e-07
```

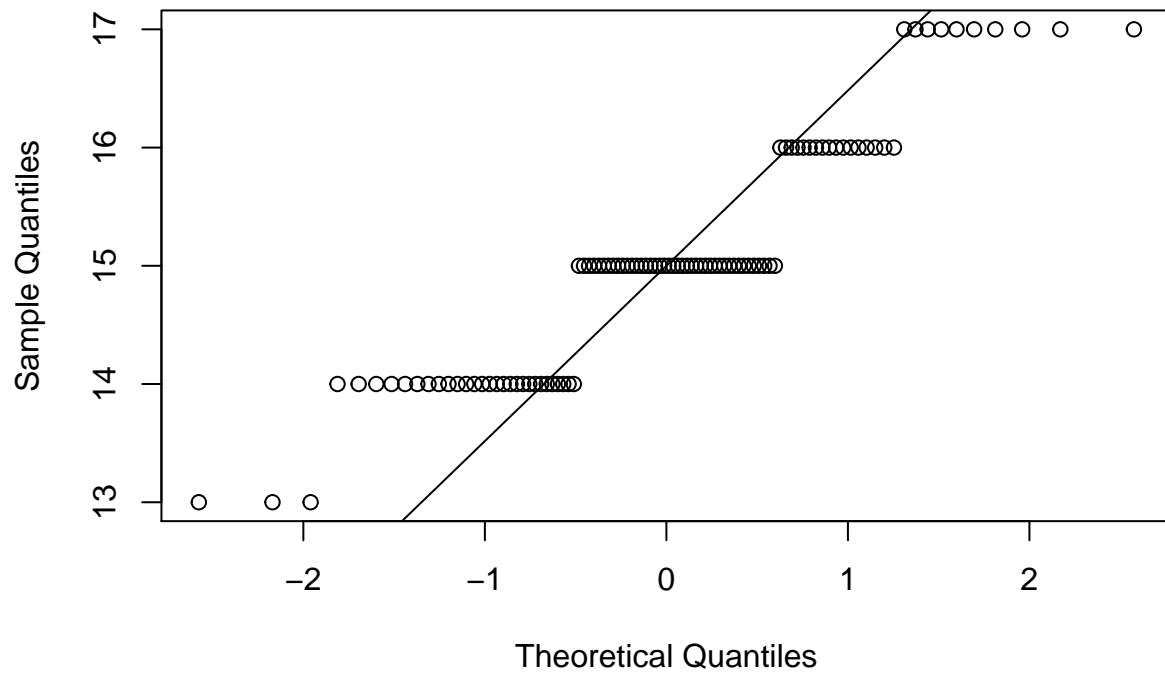
```
shapiro.test(Wild)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: Wild  
## W = 0.97456, p-value = 0.04986
```

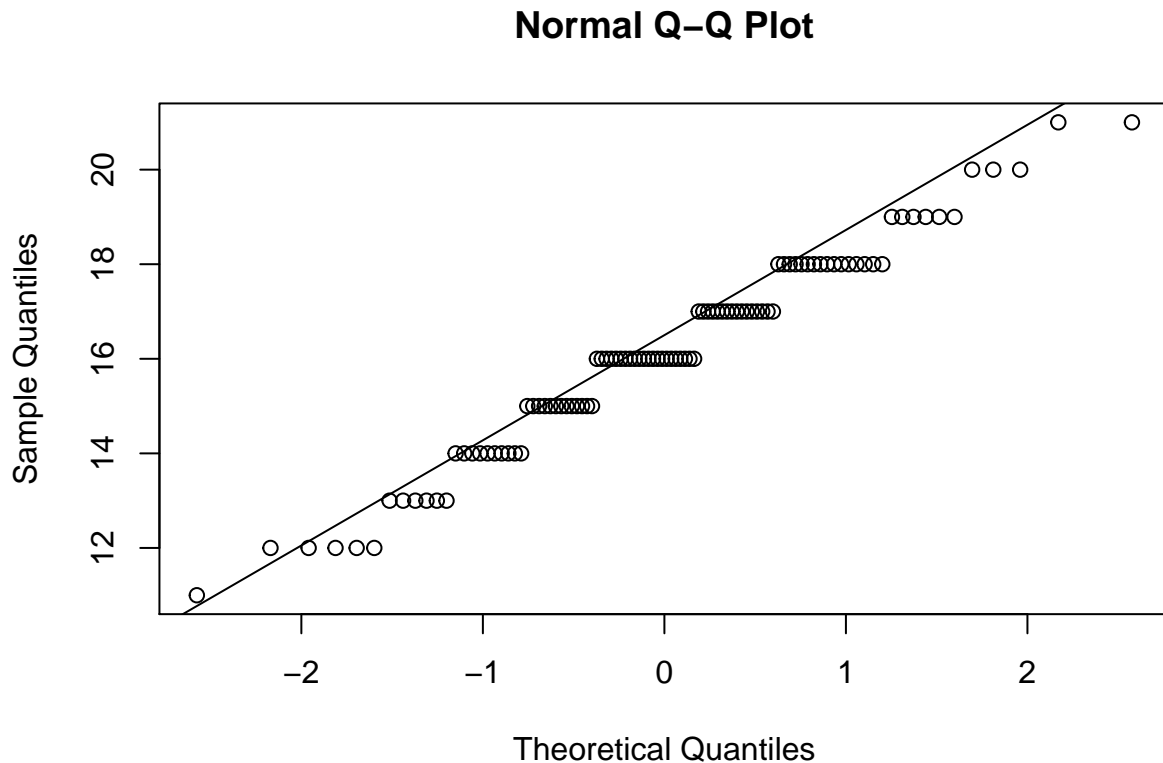
```
qqnorm(Captive)
```

```
qqline(Captive)
```


Normal Q-Q Plot



```
qqnorm(Wild)  
qqline(Wild)
```



What are the Hypotheses? Null: wild \leq captive Alternative: wild $>$ captive

Evaluate the homogeneity of the variances.

```
var.test(Captive,Wild)
```

```
##
##  F test to compare two variances
##
## data:  Captive and Wild
## F = 0.21898, num df = 99, denom df = 99, p-value = 5.351e-13
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.1473361 0.3254495
## sample estimates:
## ratio of variances
##      0.218976
```

The p-value < 0.05 , so now we can perform the two-sample t-test for non-homogeneous (or heterogeneous) variances.

Conduct the t-test. What do you conclude?

```
t.test(Wild, Captive, var.equal=F, alternative="greater")
```

```
##
##  Welch Two Sample t-test
##
## data:  Wild and Captive
```

```
## t = 4.6694, df = 140.37, p-value = 3.495e-06
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  0.7034827      Inf
## sample estimates:
## mean of x mean of y
##    16.12    15.03
```

The p-value is <0.05 , so we reject the null hypothesis.

We conclude that wild salmon weigh a statistically significant more amount than captive salmon.