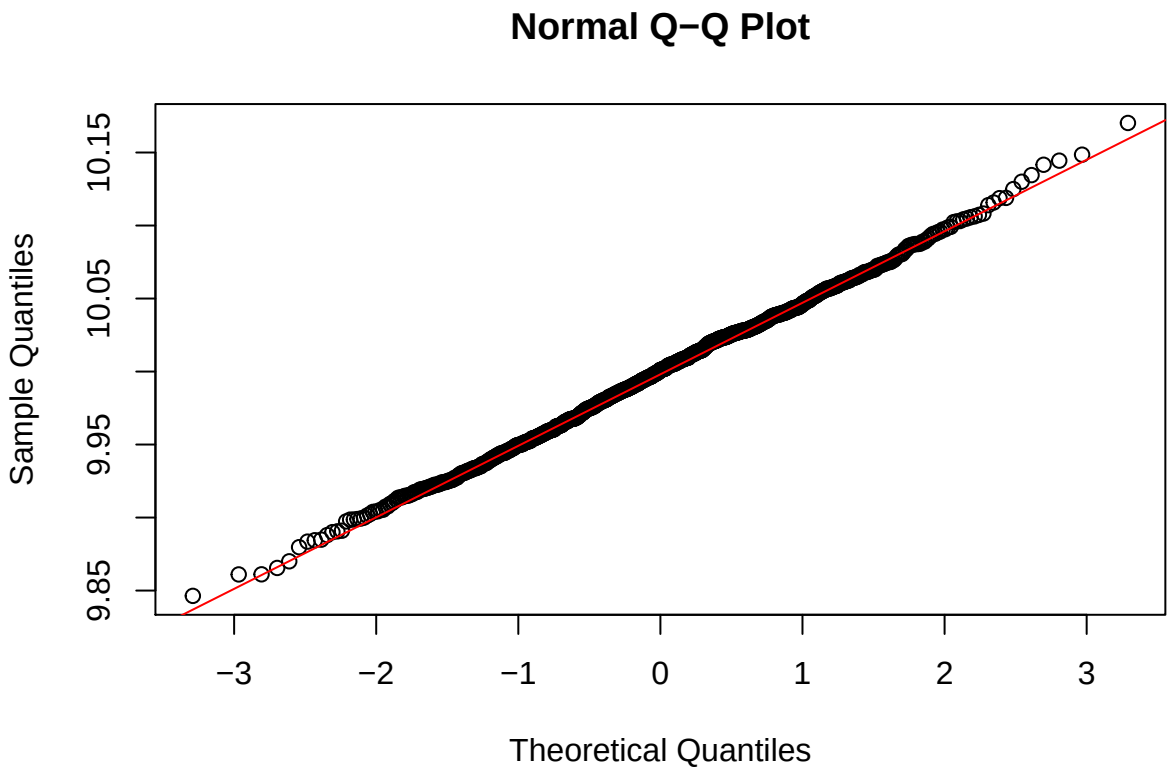
Statistic Tests

Raha

```
#One-Sample t-test
mydata<- read.csv(file="https://raw.githubusercontent.com/RahaSoleymanzadeh/Learning-R/master/Data/scre
# check normalty of data
qqnorm(mydata$Screen_size.in.cm.)
qqline(mydata$Screen_size.in.cm., col="red")
```




```
# Conduct a one-sample t-test
# Let's compute it manually:
(mean(mydata$Screen_size.in.cm.)- 10 ) / ((sd(mydata$Screen_size.in.cm.))/sqrt((length(mydata$Screen_si
```

```
## [1] -0.3954815
```

```
# Null Hypothesis: Mean screensize of sample does not differ from 10cm.
t.test(mydata$Screen_size.in.cm., mu=10)
```

```
##
##  One Sample t-test
##
## data:  mydata$Screen_size.in.cm.
## t = -0.39548, df = 999, p-value = 0.6926
## alternative hypothesis: true mean is not equal to 10
## 95 percent confidence interval:
##   9.996361 10.002418
## sample estimates:
## mean of x
##   9.99939
```

The p-value = 0.6926, which is greater than 0.05. Therefore, we fail to reject the null hypothesis at a 95% confidence interval.

```
#One-Sample t-test
results = c(65, 78, 88, 55, 48, 95, 66, 57, 79, 81)
# Let's compute it manually:
(mean(results)- 75 ) / (sd(results)/sqrt(10))
```

```
## [1] -0.7830291
```

```
t.test (results, mu=75)
```

```
##
##  One Sample t-test
##
## data:  results
## t = -0.78303, df = 9, p-value = 0.4537
## alternative hypothesis: true mean is not equal to 75
## 95 percent confidence interval:
##  60.22187 82.17813
## sample estimates:
## mean of x
##      71.2
```

The p-value = 0.4537, which is greater than 0.05. Therefore, we fail to reject the null hypothesis at a 95% confidence interval.

```
#Independent Two-Sample t-test
mydata2<- read.csv(file = "https://raw.githubusercontent.com/RahaSoleymanzadeh/Learning-R/master/Data/i
dim(mydata2)
```

```
## [1] 1000    2
```

```
var(mydata2$screensize_sample1)
```

```
## [1] 0.00238283
```

```r
var(mydata2$screensize_sample2)
```

```
## [1] 0.002353585
```

```r
#Almost equal Variance
#Null Hypothesis: There is no difference between the mean of two samples.
t.test(mydata2$screensize_sample1, mydata2$screensize_sample2, var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  mydata2$screensize_sample1 and mydata2$screensize_sample2
## t = 1.3072, df = 1998, p-value = 0.1913
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.001423145  0.007113085
## sample estimates:
## mean of x mean of y
## 10.000976  9.998131
```

The p-value = 0.1913, which is greater than 0.05. Therefore, we fail to reject the null hypothesis at a 95% confidence interval.

```r
#Paired Sample t-test
mydata3<- read.csv(file="https://raw.githubusercontent.com/RahaSoleymanzadeh/Learning-R/master/Data/pai
dim(mydata3)
```

```
## [1] 25  2
```

```r
#Null Hypothesis: There is no difference between the means of tyres before and after.
t.test(mydata3$tyre_1,mydata3$tyre_2, paired = TRUE)
```

```
##
##  Paired t-test
##
## data:  mydata3$tyre_1 and mydata3$tyre_2
## t = -5.2662, df = 24, p-value = 2.121e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -2201.6929  -961.8515
## sample estimates:
## mean of the differences
##               -1581.772
```

The p-value = 2.121e-05, which is less than 0.05. Therefore, we reject the null hypothesis at a 95% confidence interval. The negative mean in the difference(-1581.772)depicts that the average kilometers covered by tyre 2 are more than by tyre 1.

Sometimes, the population variance is the primary objective in an experimental investigation.

```r
library(EnvStats)
```

```
## Warning: package 'EnvStats' was built under R version 3.6.3
```

```
##
## Attaching package: 'EnvStats'
```

```
## The following objects are masked from 'package:stats':
##
##     predict, predict.lm
```

```
## The following object is masked from 'package:base':
##
##     print.default
```

```r
grades <- c(72, 71, 76, 77, 78, 68, 73, 71, 78, 78)
samplevariance <- var(grades)
popvariance <- 4
varTest(grades, alternative="greater", sigma.squared = popvariance)
```

```
##
##  Chi-Squared Test on Variance
##
## data:  grades
## Chi-Squared = 29.9, df = 9, p-value = 0.0004562
## alternative hypothesis: true variance is greater than 4
## 95 percent confidence interval:
##  7.068985      Inf
## sample estimates:
## variance
## 13.28889
```

The p-value = 0.0004562, which is less than 0.05. Therefore, we reject the null hypothesis