Testing.R

```
# install readxl package first
library(readxl)

# get data - see excel file for data description information
Congress <- read_excel("Piracy.xlsx", col_names = TRUE)
# two-tailed test of H0: mu(Years) = 10
t.test(Congress$years, alternative = "two.sided", mu = 10)

##
##  One Sample t-test
##
## data:  Congress$years
## t = 4.1597, df = 533, p-value = 3.713e-05
## alternative hypothesis: true mean is not equal to 10
## 95 percent confidence interval:
##  10.9290 12.5916
## sample estimates:
## mean of x
##   11.7603

# one-tailed test of H0: mu(money_pro) <= 25,000
t.test(Congress$money_pro, alternative = "greater", mu = 25000)

##
##  One Sample t-test
##
## data:  Congress$money_pro
## t = 0.27262, df = 533, p-value = 0.3926
## alternative hypothesis: true mean is greater than 25000
## 95 percent confidence interval:
##  21792.3     Inf
## sample estimates:
## mean of x
##  25635.94
```

```r
# one-tailed test of H0: mu(money_con) >= 25,000
t.test(Congress$money_con, alternative = "less", mu = 25000)
```

```
##
##  One Sample t-test
##
## data:  Congress$money_con
## t = -1.7363, df = 533, p-value = 0.04155
## alternative hypothesis: true mean is less than 25000
## 95 percent confidence interval:
##      -Inf 24830.33
## sample estimates:
## mean of x
##  21672.77
```

```r
# two-tailed test for H0: (proportion with stance = "yes") = .10
table(Congress$stance)
```

```
##
## leaning no         no  undecided    unknown        yes
##         44        122         11        294         63
```

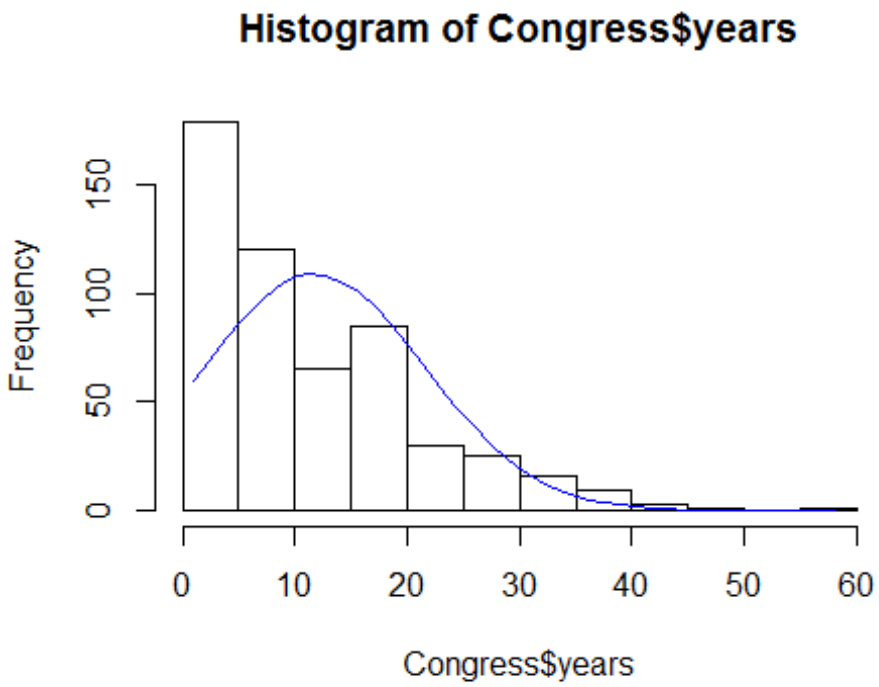
```r
p_hat <- 63/534
  # exact binomial estimation method
binom.test(63, 534, p = .10, alternative = "two.sided")
```

```
##
##  Exact binomial test
##
## data:  63 and 534
## number of successes = 63, number of trials = 534, p-value = 0.1702
## alternative hypothesis: true probability of success is not equal to 0.1
## 95 percent confidence interval:
##  0.09185852 0.14841005
## sample estimates:
## probability of success
##              0.1179775
```

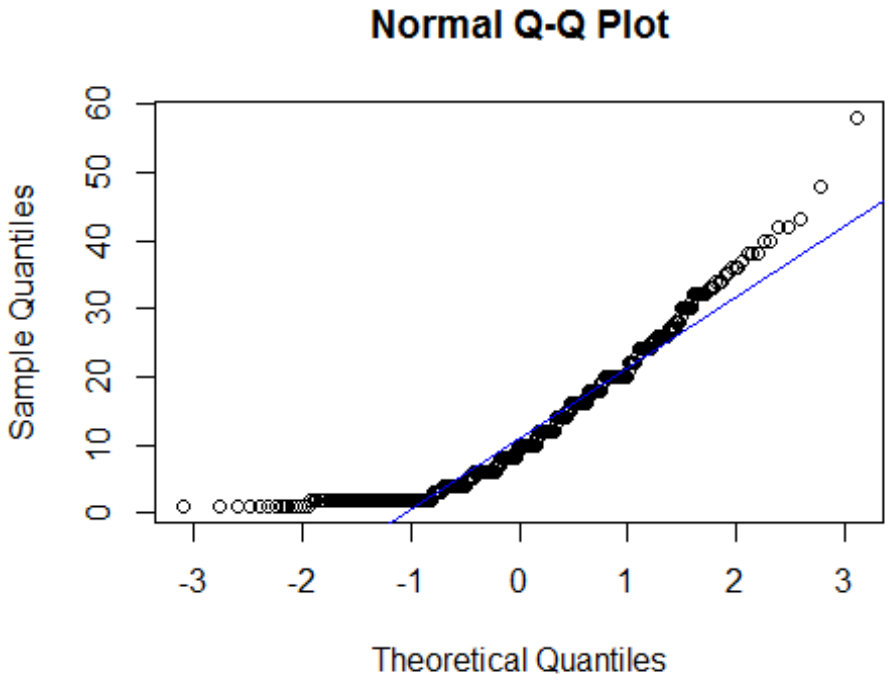
```r
  # using a Wilson score interval method that is generally preferred
prop.test(63, 534, p = .10, alternative = "two.sided")
```

```
##
##  1-sample proportions test with continuity correction
##
## data:  63 out of 534, null probability 0.1
## X-squared = 1.7231, df = 1, p-value = 0.1893
## alternative hypothesis: true p is not equal to 0.1
## 95 percent confidence interval:
##  0.09247091 0.14912640
## sample estimates:
##         p
## 0.1179775
```

```r
# Checking normality of Years distribution
  # histogram: years, with normal curve
h <- hist(Congress$years)
  # code to add normal curve
x <- Congress$years
xfit <- seq(min(x), max(x), length = 40)
yfit <- dnorm(xfit, mean = mean(x), sd = sd(x))
yfit <- yfit*diff(h$mids[1:2])*length(x)
lines(xfit, yfit, col="blue")
```



```r
  # probability plot
qqnorm(x)
qqline(x, col = "blue")
```




```r
  # goodness of fit test of H0: normal
  # but interpret with caution for large data sets
shapiro.test(x)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  x
## W = 0.8852, p-value < 2.2e-16
```