STAT350 Lab5
Tian Qiu
Instructor: Womble

1. (10 points) GROUP PART: This is a group assignment and is due on Blackboard at Midnight on FRIDAY, March 4. Be sure that the names and sections of each person are at the top of the page. Combine your data with 3 or 4 other students (in any of your instructor's sections) and answer the following questions (no data is required for this part):

a. Is the number of intervals that contain the mean what you would expect for the combined data? Please explain your answer.

b. Are the results from part 4a (the group part) more consistent with the theory than part 3 (the individual part)? Is this what you expected? Please explain.

A. 1.

```r
# A
i=0
while(i < 30) {
  data.vec <- rnorm(40,mean=10,sd=2)
  data.mat <- matrix(data.vec, ncol = n)
  avg <- apply(data.mat, 1, mean)

  m = mean(avg)
  std = sd(avg)
  print(z.test(data.vec, conf.level = 0.80, alternative="two.sided", sd=std))
  i<- i+1
}
```

A. 2.

One Sample z-test

data: data.vec
z = 138.48, n = 40.000000, Std. Dev. = 0.455070, Std. Dev. of the sample mean = 0.071953, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
 9.871684 10.056108
sample estimates:
mean of data.vec
 9.963896

One Sample z-test

data: data.vec
z = 316.67, n = 40.00000, Std. Dev. = 0.20833, Std. Dev. of the sample mean = 0.03294, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
 10.38871 10.47314
sample estimates:
mean of data.vec
 10.43093

One Sample z-test

data: data.vec
z = 194.78, n = 40.000000, Std. Dev. = 0.330230, Std. Dev. of the sample mean = 0.052214, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
10.10344 10.23727
sample estimates:
mean of data.vec
10.17035

One Sample z-test

data: data.vec
z = 52.202, n = 40.00000, Std. Dev. = 1.22650, Std. Dev. of the sample mean = 0.19393, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
9.87479 10.37185
sample estimates:
mean of data.vec
10.12332

One Sample z-test

data: data.vec
z = 93.652, n = 40.00000, Std. Dev. = 0.68529, Std. Dev. of the sample mean = 0.10835, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
10.00879 10.28651
sample estimates:
mean of data.vec
10.14765

One Sample z-test

```
data:  data.vec
z = 207.54, n = 40.000000, Std. Dev. = 0.311880, Std. Dev. of the sample mean =
0.049313, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
 10.17148 10.29787
sample estimates:
mean of data.vec
        10.23468


    One Sample z-test

data:  data.vec
z = 127.34, n = 40.000000, Std. Dev. = 0.487950, Std. Dev. of the sample mean =
0.077151, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
 9.725311 9.923057
sample estimates:
mean of data.vec
        9.824184


    One Sample z-test

data:  data.vec
z = 79.571, n = 40.00000, Std. Dev. = 0.83597, Std. Dev. of the sample mean =
0.13218, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
 10.34827 10.68706
sample estimates:
mean of data.vec
        10.51766


    One Sample z-test

data:  data.vec
z = 206.88, n = 40.000000, Std. Dev. = 0.295750, Std. Dev. of the sample mean =
```

0.046762, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
 9.614249 9.734104
sample estimates:
mean of data.vec
 9.674177

One Sample z-test

data: data.vec
z = 183.83, n = 40.000000, Std. Dev. = 0.335950, Std. Dev. of the sample mean = 0.053118, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
 9.696544 9.832691
sample estimates:
mean of data.vec
 9.764617

One Sample z-test

data: data.vec
z = 42.817, n = 40.00000, Std. Dev. = 1.43210, Std. Dev. of the sample mean = 0.22644, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
 9.405435 9.985828
sample estimates:
mean of data.vec
 9.695632

One Sample z-test

data: data.vec
z = 266.83, n = 40.000000, Std. Dev. = 0.227220, Std. Dev. of the sample mean = 0.035927, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0

80 percent confidence interval:
 9.540203 9.632287
sample estimates:
mean of data.vec
 9.586245

One Sample z-test

data: data.vec
z = 141.32, n = 40.000000, Std. Dev. = 0.468100, Std. Dev. of the sample mean = 0.074014, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
 10.36509 10.55480
sample estimates:
mean of data.vec
 10.45995

One Sample z-test

data: data.vec
z = 142.4, n = 40.000000, Std. Dev. = 0.429010, Std. Dev. of the sample mean = 0.067832, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
 9.572130 9.745992
sample estimates:
mean of data.vec
 9.659061

One Sample z-test

data: data.vec
z = 103.34, n = 40.000000, Std. Dev. = 0.610170, Std. Dev. of the sample mean = 0.096477, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
 9.846288 10.093568

sample estimates:
mean of data.vec
9.969928

One Sample z-test

data: data.vec
z = 71.359, n = 40.00000, Std. Dev. = 0.92940, Std. Dev. of the sample mean = 0.14695, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
10.29783 10.67448
sample estimates:
mean of data.vec
10.48615

One Sample z-test

data: data.vec
z = 116.74, n = 40.000000, Std. Dev. = 0.530360, Std. Dev. of the sample mean = 0.083858, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
9.681850 9.896786
sample estimates:
mean of data.vec
9.789318

One Sample z-test

data: data.vec
z = 119.45, n = 40.000000, Std. Dev. = 0.522000, Std. Dev. of the sample mean = 0.082535, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
9.753376 9.964923
sample estimates:
mean of data.vec

9.859149

One Sample z-test

data: data.vec
z = 117.42, n = 40.000000, Std. Dev. = 0.530520, Std. Dev. of the sample mean = 0.083882, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
9.742195 9.957194
sample estimates:
mean of data.vec
9.849694

One Sample z-test

data: data.vec
z = 158, n = 40.000000, Std. Dev. = 0.394660, Std. Dev. of the sample mean = 0.062401, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
9.779286 9.939227
sample estimates:
mean of data.vec
9.859257

One Sample z-test

data: data.vec
z = 66.264, n = 40.00000, Std. Dev. = 0.96336, Std. Dev. of the sample mean = 0.15232, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
9.898161 10.288574
sample estimates:
mean of data.vec
10.09337

One Sample z-test

data: data.vec
z = 90.498, n = 40.00000, Std. Dev. = 0.67441, Std. Dev. of the sample mean = 0.10663, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
9.513522 9.786836
sample estimates:
mean of data.vec
9.650179

One Sample z-test

data: data.vec
z = 87.328, n = 40.00000, Std. Dev. = 0.72398, Std. Dev. of the sample mean = 0.11447, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
9.849939 10.143343
sample estimates:
mean of data.vec
9.996641

One Sample z-test

data: data.vec
z = 101.08, n = 40.000000, Std. Dev. = 0.608930, Std. Dev. of the sample mean = 0.096281, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
9.608819 9.855597
sample estimates:
mean of data.vec
9.732208

One Sample z-test

data: data.vec
z = 132.42, n = 40.000000, Std. Dev. = 0.490630, Std. Dev. of the sample mean = 0.077576, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
 10.17299 10.37183
sample estimates:
mean of data.vec
 10.27241

One Sample z-test

data: data.vec
z = 61.505, n = 40.00000, Std. Dev. = 0.97596, Std. Dev. of the sample mean = 0.15431, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
 9.293369 9.688891
sample estimates:
mean of data.vec
 9.49113

One Sample z-test

data: data.vec
z = 136.37, n = 40.00000, Std. Dev. = 0.49009, Std. Dev. of the sample mean = 0.07749, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
 10.46826 10.66687
sample estimates:
mean of data.vec
 10.56757

One Sample z-test

data: data.vec

z = 98.675, n = 40.000000, Std. Dev. = 0.602510, Std. Dev. of the sample mean = 0.095265, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
 9.278205 9.522378
sample estimates:
mean of data.vec
 9.400292

One Sample z-test

data: data.vec
z = 209.85, n = 40.000000, Std. Dev. = 0.288980, Std. Dev. of the sample mean = 0.045691, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
 9.529825 9.646936
sample estimates:
mean of data.vec
 9.58838

One Sample z-test

data: data.vec
z = 77.165, n = 40.00000, Std. Dev. = 0.78760, Std. Dev. of the sample mean = 0.12453, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
80 percent confidence interval:
 9.449834 9.769018
sample estimates:
mean of data.vec
 9.609426

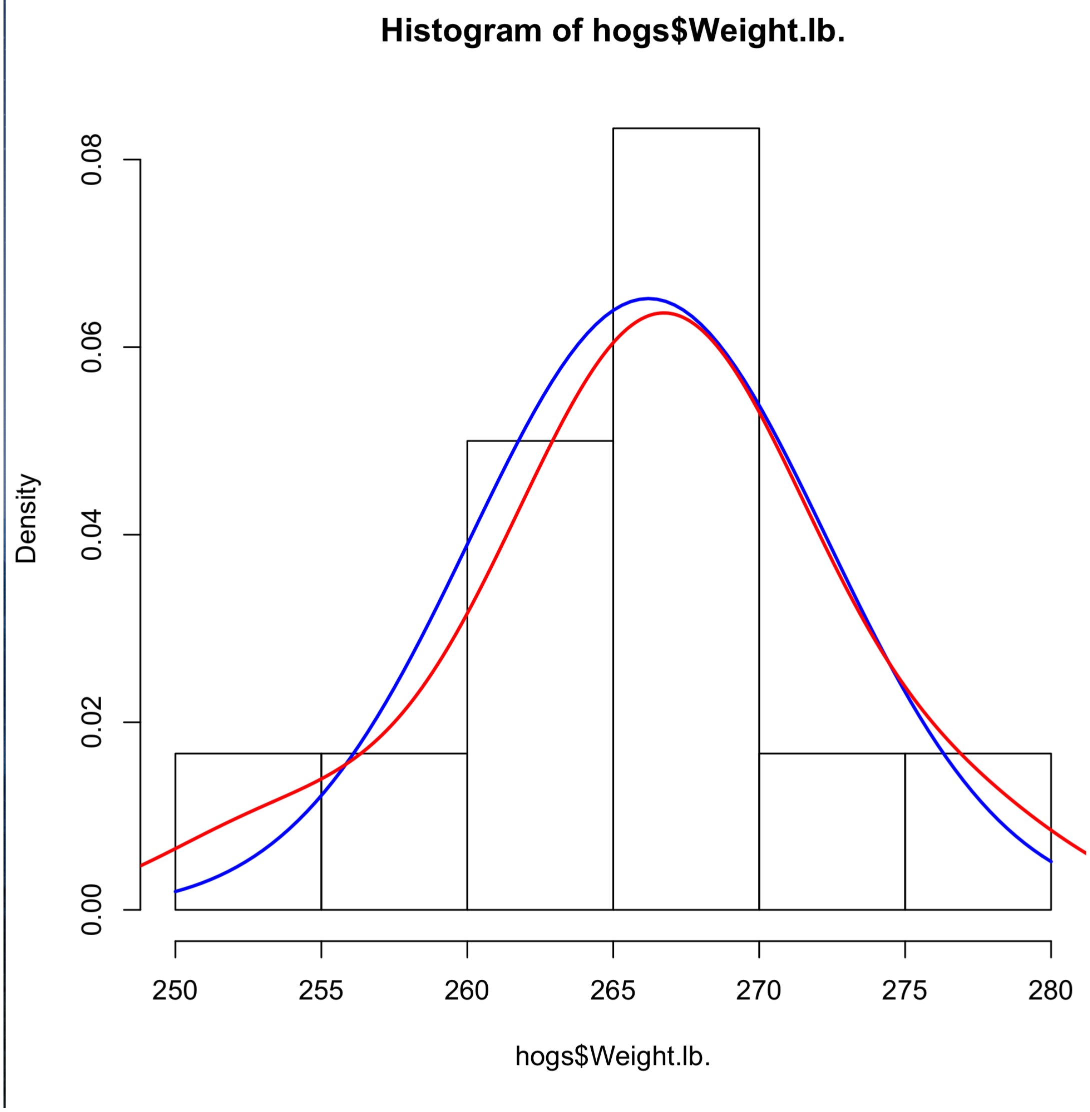
A. 3.

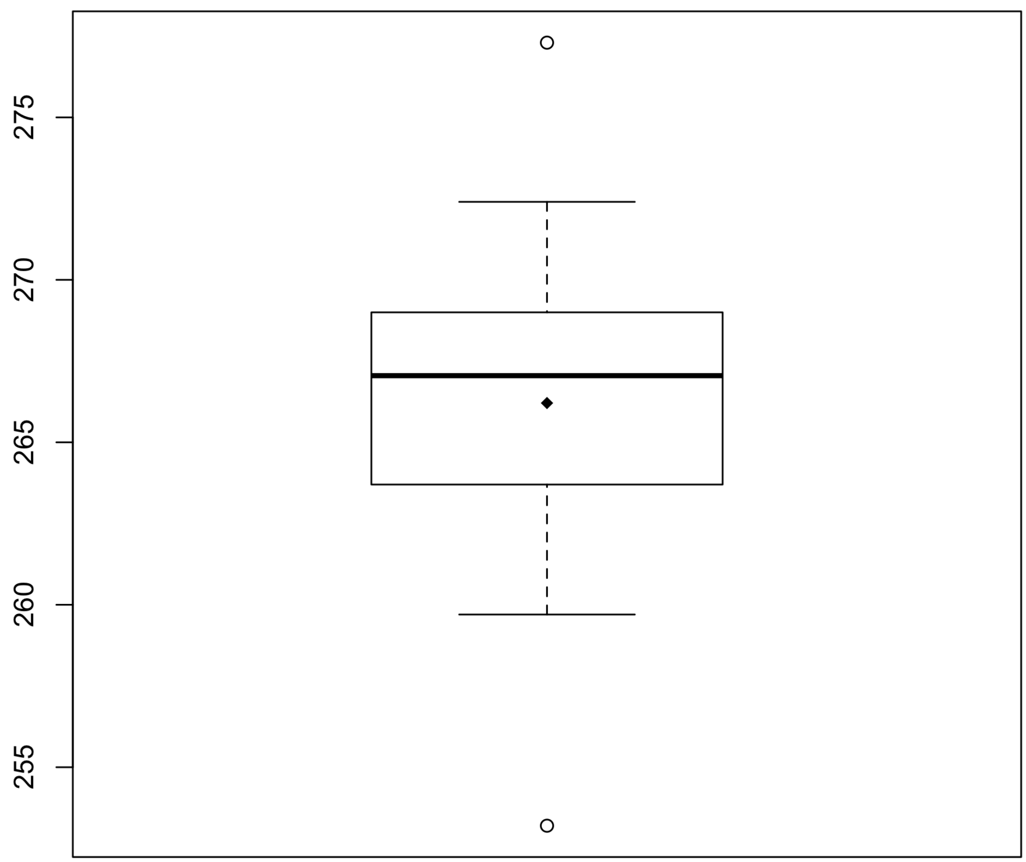
Four of these intervals contain the population mean u=10, which is highlighted with yellow. It is what I have expected, because 40 random normal distribution number is not enough to limit the interval precisely.

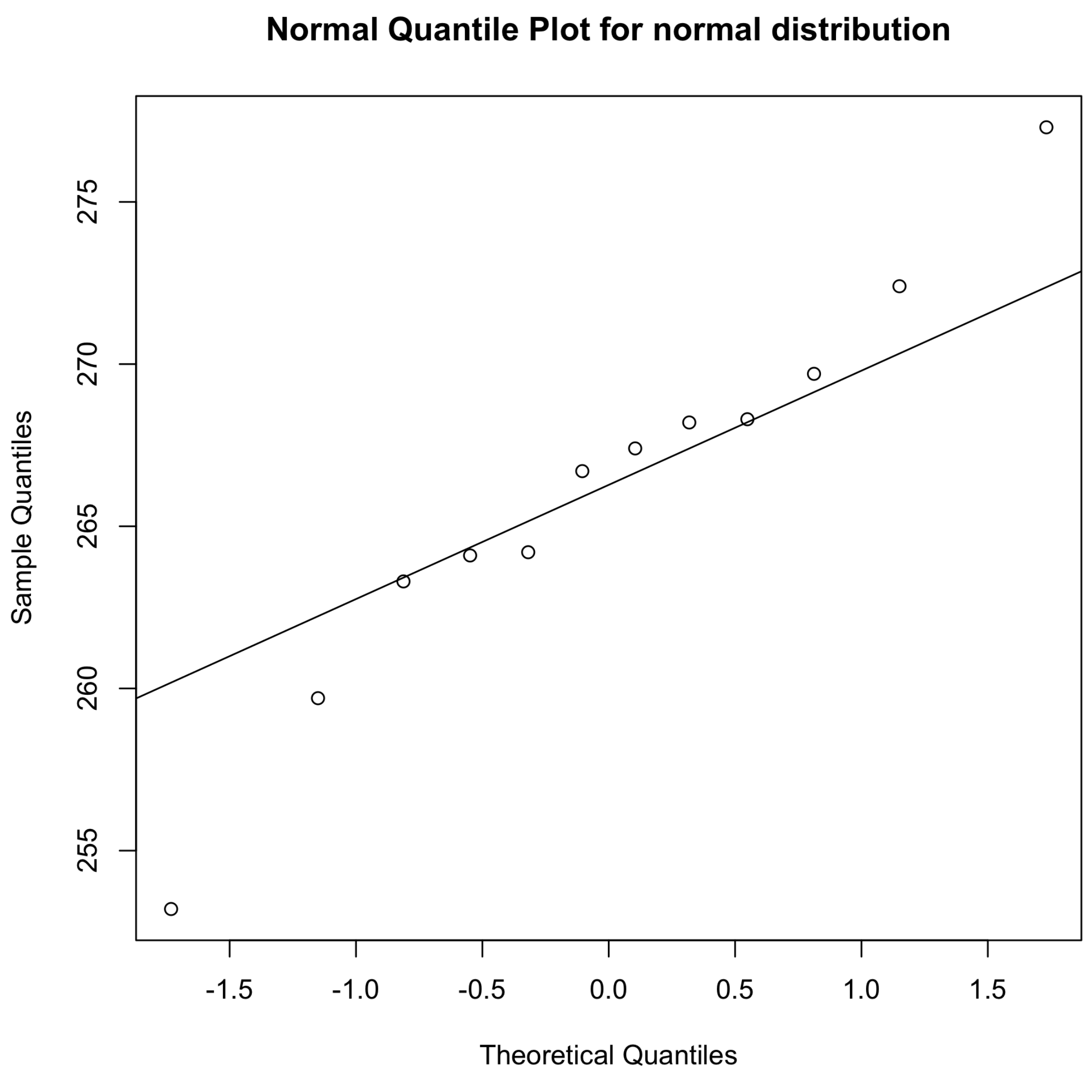
B.1.

```r
# B
hogs=read.table(file="hogs.txt",header=T)

print(hogs$Weight.lb.)
quartz()
hist(hogs$Weight.lb.,freq = FALSE)          # frequency should be false to get density curve.....
means <- mean(hogs$Weight.lb.)
print(means)
std <- sd(hogs$Weight.lb.)
curve(dnorm(x, mean=means, sd=std), col="blue", lwd=2, add=TRUE)   # normal distribution line
lines(density(hogs$Weight.lb., adjust=2),col = "red", lwd=2)
mean(hogs$Weight.lb.)
sd(hogs$Weight.lb.)

# boxplot
quartz()
boxplot(hogs$Weight.lb.)
points(means, pch = 18)



#  b) Make a Normal quantile plot to confirm that there are no systematic departures from Normality.
quartz()
qqnorm(hogs$Weight.lb.,main="Normal Quantile Plot for normal distribution")
qqline(hogs$Weight.lb.)

t.test(hogs$Weight.lb., conf.level=0.95, alternative = "two.sided")
library(TeachingDemos)
stdev = sd(hogs$Weight.lb.)
stdev
z.test(hogs$Weight.lb., conf.level = 0.95, alternative="two.sided", sd=stdev)
```

B.2


Histogram of hogs$Weight.lb.
Density
0.00
0.02
0.04
0.06
0.08
250
255
260
265
270
275
280
hogs$Weight.lb.


255
260
265
270
275


Normal Quantile Plot for normal distribution
Sample Quantiles
255
260
265
270
275
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
Theoretical Quantiles

B.3.

The distribution is slightly skewed to the left but totally symmetric. Also, it has few outliers hence the graph looks pretty normally distributed.

B.4.

It is appropriate to analyze these data using the t procedures because the question only offered the population mean. We do not have any information about population variance and we do know the sample mean and the sample variance. Hence, using t test is appropriate.

B.5.

```
> z.test(hogs$Weight.lb., conf.level = 0.95, alternative="two.sided", sd=stdev)

	One Sample z-test

data:  hogs$Weight.lb.
z = 150.67, n = 12.0000, Std. Dev. = 6.1205, Std. Dev. of the sample mean = 1.7668, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 262.7454 269.6713
sample estimates:
mean of hogs$Weight.lb.
               266.2083
```

B.6.

```
> t.test(hogs$Weight.lb., conf.level=0.95, alternative = "two.sided")

        One Sample t-test

data:  hogs$Weight.lb.
t = 150.67, df = 11, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 262.3195 270.0971
sample estimates:
mean of x
 266.2083
```

B.7.

These two intervals are different because we use different test procedures. In z-test we do not consider degree of freedom while in t-test we do consider DF.

B.8.

We do have evidence to reject the claim that μ = 275.4 because 275.4 is not included in t-test confidence interval we calculated.