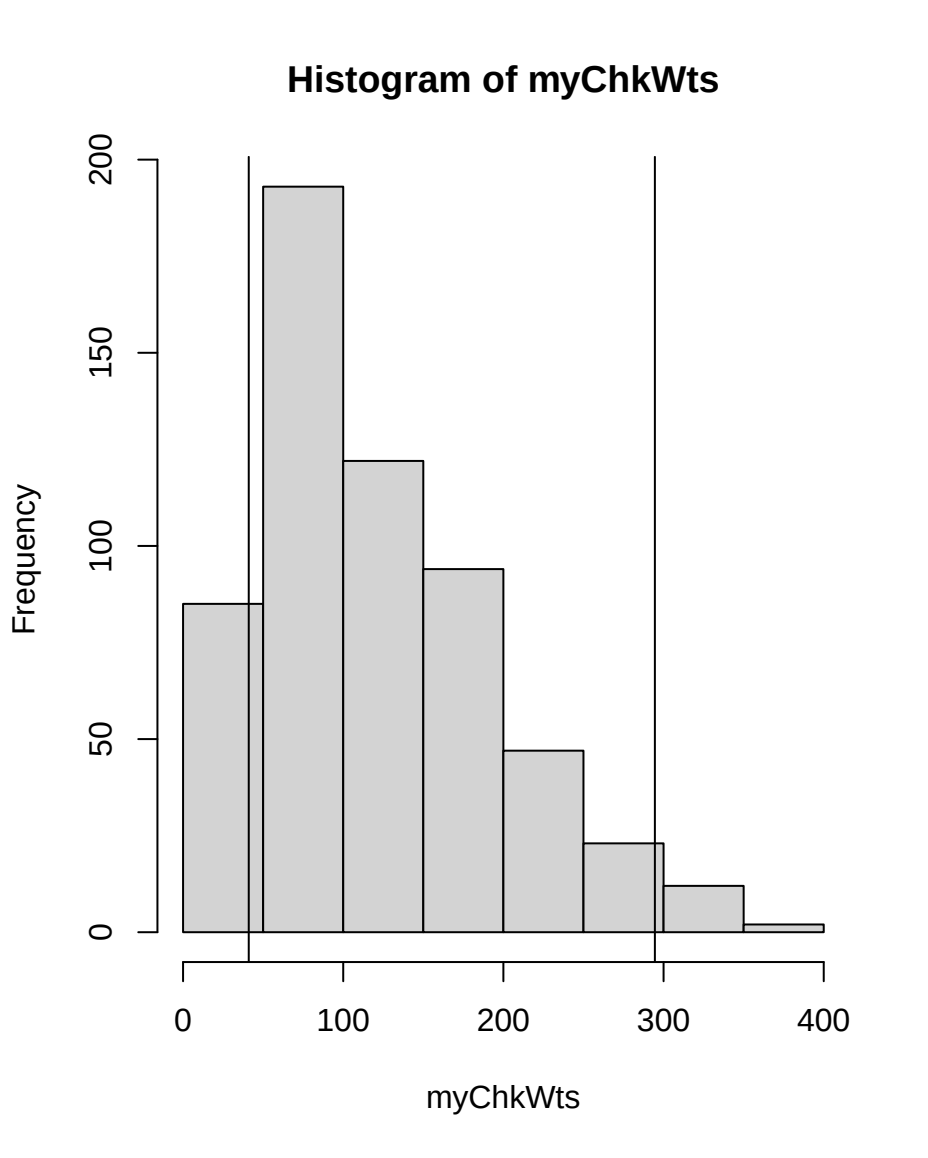
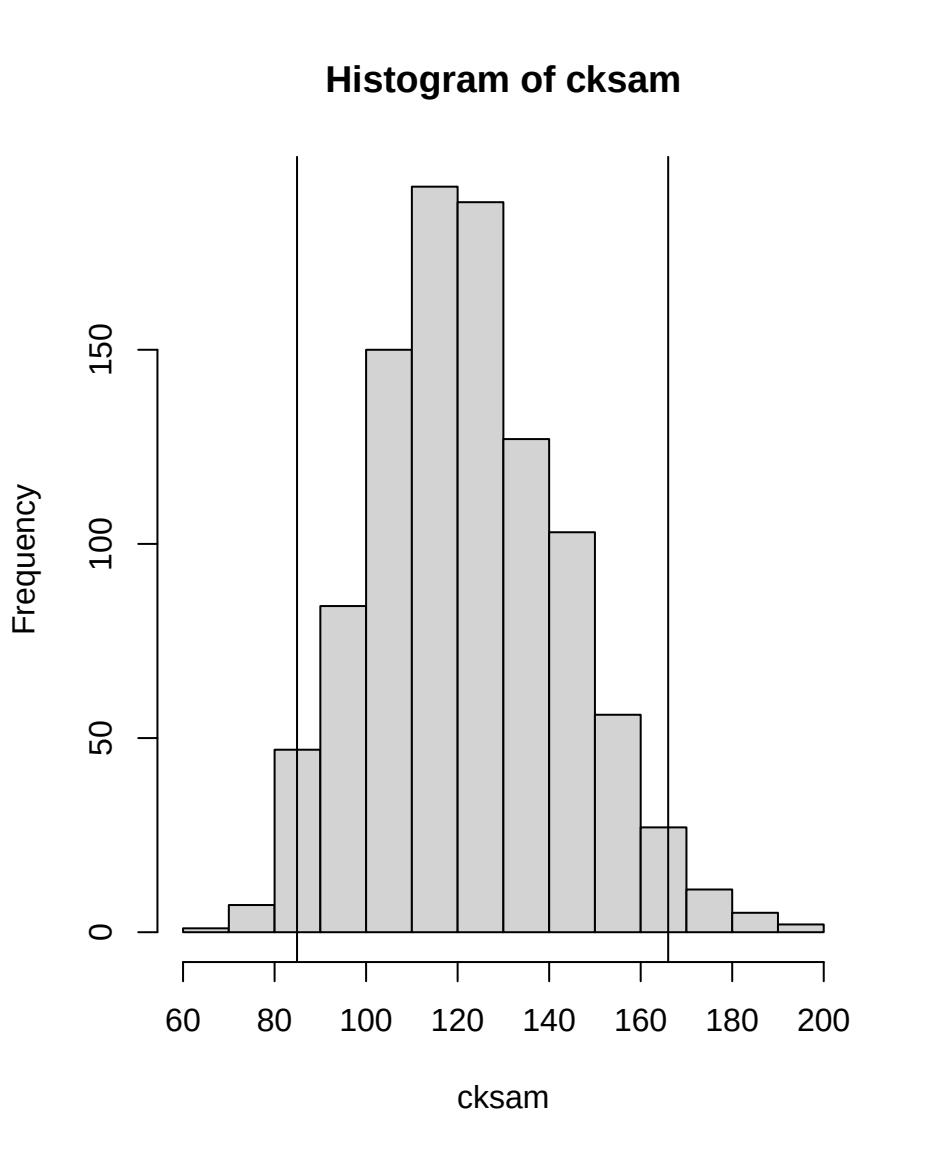
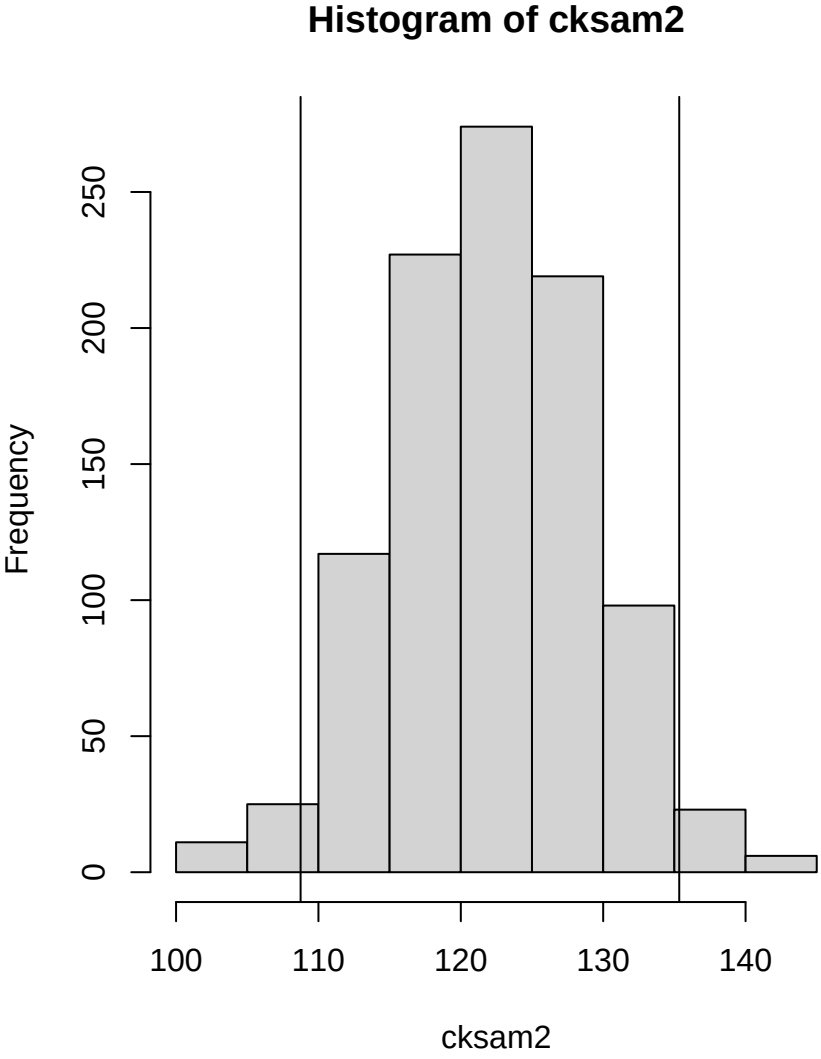
```
#########################################################################
#
# Author: DeAndre Thomas
# Professor: Dr. Anastasopoulos
# Purpose: Homework 3
# Class: IST 772
# Date: 04/29/2022
#
#########################################################################

# Question 2

# importing ChickWeight data file and renaming as "data" per HW instruction
data = ChickWeight
# View(data)

summary(data)
# The four different variables are 'weight', 'time', 'chick', and 'diet'.

dim(data)
# The first number, 578, is the number of rows in the data set.The rows of
data
# are unique to the chicks that were observed to create this dat set.

#########################################################################

# Question 3

summary(data$weight)
# This line of code gives maximum value, minimum value and quantiles of the
# weight column of this dataframe.

head(data$weight)
# This line provides the first 5 data points in the weight column of this
# frame.

mean(data$weight)
# # This line provides the mean or average of all of the data points in the
# weight column of this data frame.

myChkWts <- data$weight
# This line creates a subset of the ChickWeight data frame, that only consist
# of the weight column.

quantile(myChkWts,0.50)
# This line provides the median/50 percentile/2nd quartile of the weight
column
# of the data frame.

#########################################################################

# Question 4

# Creating a histogram of myChkWts
hist(myChkWts)
abline(v=quantile(myChkWts, 0.025))
```

```
abline(v=quantile(myChkWts, 0.975))

# The shape of this histogram is right-skewed, enabling the mean and median to
# be to the right of the most frequeny values (between 50 and 100).
# We are able to see that more data falls under the 2.5 percentile than it
does
# the 97.5 percentile.

#########################################################################

# Question 5

cksam<- (replicate(1000,mean(sample(myChkWts,size=11,replace=TRUE))))
hist(cksam)
abline(v=quantile(cksam, 0.025))
abline(v=quantile(cksam, 0.975))

#########################################################################

# Question 6

# The difference between distribution of raw data and that of sample data is
# that of sampling means will demonstrate normal distribution and that of raw
# data can vary in shape. Also, the quantiles of sampling means are
# similar to that of the original data.

#########################################################################

# Question 7

cksam2<- (replicate(1000,mean(sample(myChkWts,size=100,replace=TRUE))))
hist(cksam2)
abline(v=quantile(cksam2, 0.025))
abline(v=quantile(cksam2, 0.975))

# This data is better because we are using more observations. Thus, this
# example is providing a better, more accurate represensation of the data.
```


Histogram of myChkWts
Frequency
0
50
100
150
200
0
100
200
300
400
myChkWts


Histogram of cksam
Frequency
0
50
100
150
60
80
100
120
140
160
180
200
cksam


Histogram of cksam2
Frequency
0
50
100
150
200
250
100
110
120
130
140
cksam2