**Instructions:** This lab is due on Wednesday, 1/30 during the first half of class. You should hand it in when you turn in your homework. Please show work in a neat and organized manner.

The following are a simple random sample of number of minutes (on average during a week) that students exercised in minutes, as reported by students at Foothill college. The sample is derived from data collected by Professor Butterworth.

60	120	120	120	120	120	0	180	120	180	180	180
240	120	120	60	180	360	240	420	420	420	300	240
300	180	0	150	300	0	180	300				

1. Use your TI-83/84 to sort the data. Create a stem-and-leaf diagram with a stem unit of "x100" and a leaf unit of "x10".

Stem (x100)	Leaves (x10)
0	0 0 0 6 6
1	2 2 2 2 2 2 2 5
2	
3	
4	
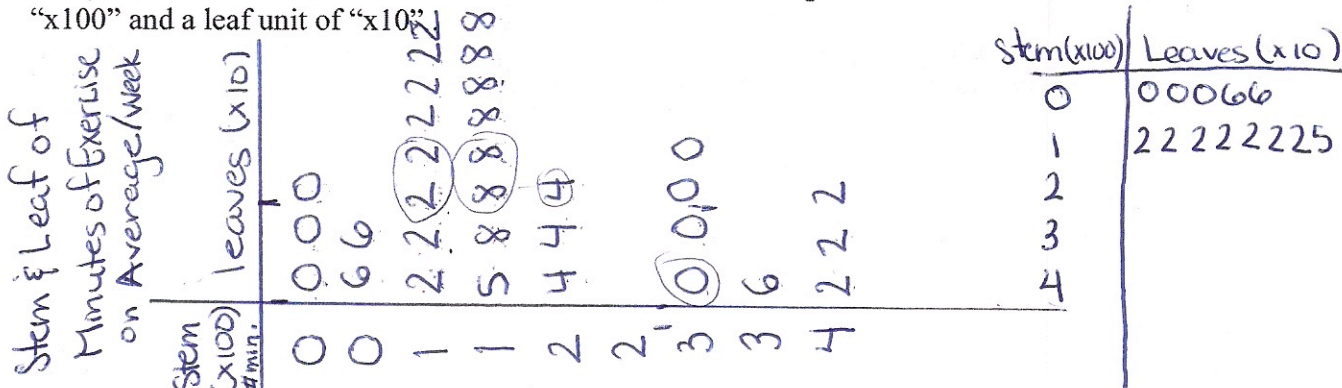
2. What do you notice about the shape of the data based upon the stem & leaf? (Write a complete sentence using words like symmetric, left or right skewed.)

The data piles up on the left and trickles out toward the right indicating right skew.

3. a) What is the minimum? **0**   b) What is the maximum? **420**

c) What is the range of the data? (show work) $420 - 0 = \boxed{420}$

4. Find the class width if there are to be 5 classes. Show the work for computing it here. Round up to the next whole number no matter what!

Class Width = $\frac{420}{5} = 84 \uparrow 85$

5. Find **all** the lower class limits for a frequency distribution/table with 5 classes. Give the 6th lower class limit too. Show the work here.

$\underset{1st}{0} + 85 = \underset{2nd}{85} + 85 = \underset{3rd}{170} + 85 = \underset{4th}{255} + 85 = \underset{5th}{340} + 85 = \underset{6th}{425}$

6. Create a frequency table/distribution for the data using the 5 classes created by using the lower limits in question #5.

Frequency Distribution of Minutes of Ave Excerise/week

Classes (in Min)	Frequency
0 – 84	5
85 – 169	9
170 – 254	10
255 – 339	4
340 – 424	4
425	32

7. Create a frequency histogram for the distribution in question #6. Make sure that you use the boundaries to create the bars. Label correctly!

Frequency Histogram of Ave Excercise/week
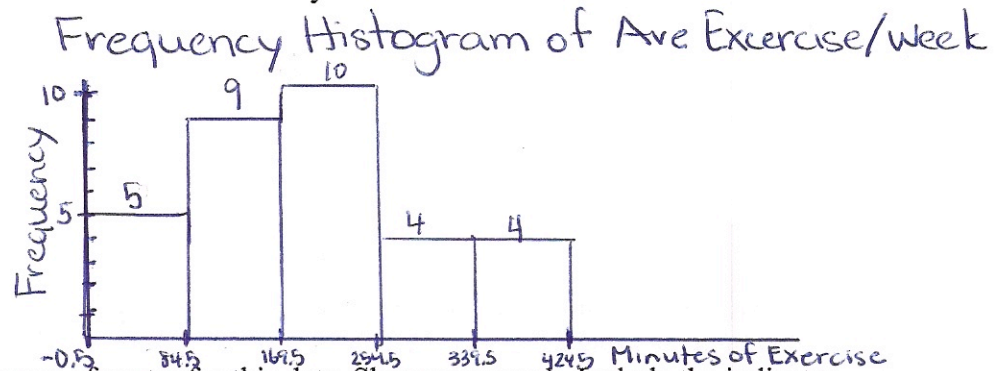
8. Correctly find the measures of center for this data. Show your work. Include the indicator function for the median. Use correct notation.

a) Mean: $\bar{X} = \frac{\Sigma x}{n} = \frac{0+0+\ldots+420}{32} = \frac{6030}{32} = 188.4375 \approx \boxed{188.4 \text{ min}}$

b) Median: $L_{50} = \frac{1}{2} \cdot 32 = 16 \text{\&} 17$ $\boxed{\tilde{x} = \frac{180+180}{2} = 180 \text{ min.}}$

c) Mode: $\boxed{120 \text{min.}}$

9. Of the three measures of center, which is the most appropriate for this data? Why?

The median is most appropriate because the data is skewed.

10. Give the variance for the data. Show your work using input of calculator values into the appropriate formula. Notate correctly! Round to 1 more decimal than original data.

$$S^2 = \frac{n\Sigma x^2 - (\Sigma x)^2}{n(n-1)} = \frac{32(1563300) - (6030)^2}{32(31)}$$

$$= \frac{50025600 - 36360900}{992} = \frac{13664700}{992}$$

$$= 13774.89919 \approx 13774.9 \text{ min}^2$$

11. Use the exact variance to find the standard deviation. Show your work. Notate correctly. Round correctly.

$$S = \sqrt{\frac{1366470}{992}} = 117.3665165 \approx 117.4 \text{ min}$$

12. Why is the standard deviation easier for us to use than the variance?

The standard deviation has the same units a the data and makes it easier to interpret than variance which has squared units.

13. Why can't we use the reported standard deviation to find the variance, even though in theory the variance is the square of the standard deviation?

The reported standard deviation would be rounded and therefore round-off error make the squared std.dev. unequal to the variance.

14. What is the indicator function for the first quartile, Q_1? Show your work here.

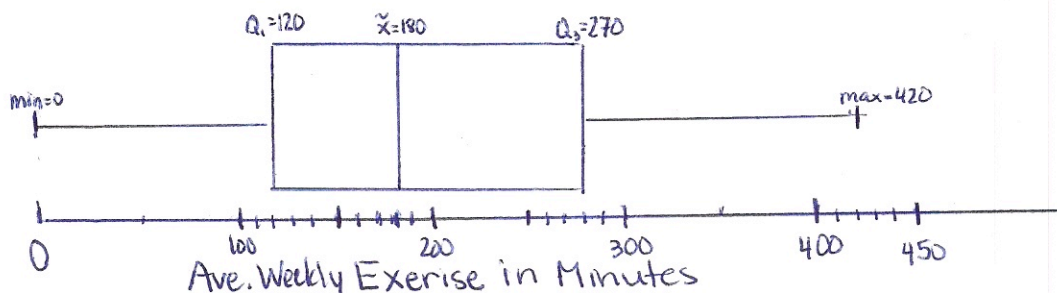$$L_{25} = \frac{1}{4} \cdot 32 = 8 \text{ & } 9^{th}$$

Nothing from here counts

15. Give the five number summary. Do not use your calculator's values to find quartiles. Use the indicator functions.

$$\text{min} = 0 \quad Q_1 = \frac{120+120}{2} = 120 \quad \tilde{x} = 180 \quad Q_3 = \frac{240+300}{2} = 270 \quad \text{max} = 420$$

all in minutes

16. Draw a box-plot for the data. Make sure that it is a **scaled** drawing. Make the scale separately!

17. Give the IQR of the data. Show your work here.

$$\text{IQR} = Q_3 - Q_1$$
$$= 270 - 120$$
$$= 150 \text{min}.$$

18. Using the IQR "fences" to gauge, are there any outliers in this data? Show your work. a)
  a) Give the "fences".
  b) Name the outliers if there are any.

a) $1.5 \, IQR = 1.5(150) = 225$
$Q_1 - 1.5 \, IQR \Rightarrow 120 - 225 = -105$ $Q_3 + 1.5 \, IQR \Rightarrow 270 + 225 = 495$

b) There are no outliers b/c all ~~outliers~~ data points
   are within -105 and 495

19. Which is a better way to judge the percentage of data that lies within 2 standard
  deviations of the mean for this data set:

        Empirical Rule   or   (Chebyshev's Theorem)
                                                  ↳ b/c data is skewed

20. Use the **mean, median and mode** to comment on the shape that you described in
  question #2. Is your original determination upheld or refuted?

The mean and median are to the right of
the mode showing that the data is right
skewed. This is consistent with the original
determination after looking ~~at~~ the stem-and-
leaf.

EC1: Find the z-score of 180 minutes.

$$z = \frac{180 - 188.44}{117.37} = -0.0719 \approx \boxed{-0.07}$$

EC2: Find the percentile of 150 inches.

$$P_? = \frac{13}{32} \cdot 100\% = \boxed{40.625\%\text{tile}}$$

$$P_{40} = 150$$