

# Analyzing Data Activity #1

Names \_\_\_\_\_

*Key*

## 1<sup>st</sup> hour Chapter Test Scores

23 15 38 28 36 38 35 38 30 35 32 29 36 21 30  
26 28 30 38 30 29 29 28 38 31 38 36

## 4<sup>th</sup> hour Chapter Test Scores

33 38 6 32 26 38 30 33 28 33 34 14 39 28 37  
38 18 18 26 36 28 27

## 5<sup>th</sup> hour Chapter Test Scores

38 26 27 34 21 38 28 32 35 16 20 28 30 30 17  
36 25 21 38 17 33 24 31 31

## 6<sup>th</sup> hour Chapter Test Scores

30 13 33 30 32 27 37 38 25 38 24 27 18 18 36  
37 30 25 38 30 39 32 38

1. Use the data above to answer the following information:

1<sup>st</sup> hour

4<sup>th</sup> hour

5<sup>th</sup> hour

6<sup>th</sup> hour

n = 27

n = 22

n = 24

n = 23

max = 38

max = 39

max = 38

max = 39

min = 15

min = 6

min = 16

min = 13

mean = 31.3

mean = 29.1

mean = 28.2

mean = 30.2

median = 30

median = 31

median = 29

median = 30

mode = 38

mode = 28, 33, 38

mode = 38

mode = 30, 38

$\sigma$  = 5.7

$\sigma$  = 8.4

$\sigma$  = 6.7

$\sigma$  = 7.1

range = 23

range = 33

range = 22

range = 26

Q1 = 28

Q1 = 26

Q1 = 22.5

Q1 = 25

Q3 = 36

Q3 = 36

Q3 = 33.5

Q3 = 37

IQR = 8

IQR = 10

IQR = 11

IQR = 12

Outliers = 15

Outliers = 6

Outliers = None

Outliers = None

$28 - (1.5 \times 8) = 16$

$26 - (1.5 \times 10) = 11$

$22.5 - (1.5 \times 11) = 6$

$25 - (1.5 \times 12) = 13$

$36 + (1.5 \times 8) = 48$

$36 + (1.5 \times 10) = 51$

$33.5 + (1.5 \times 11) = 50$

$37 + (1.5 \times 12) = 49$

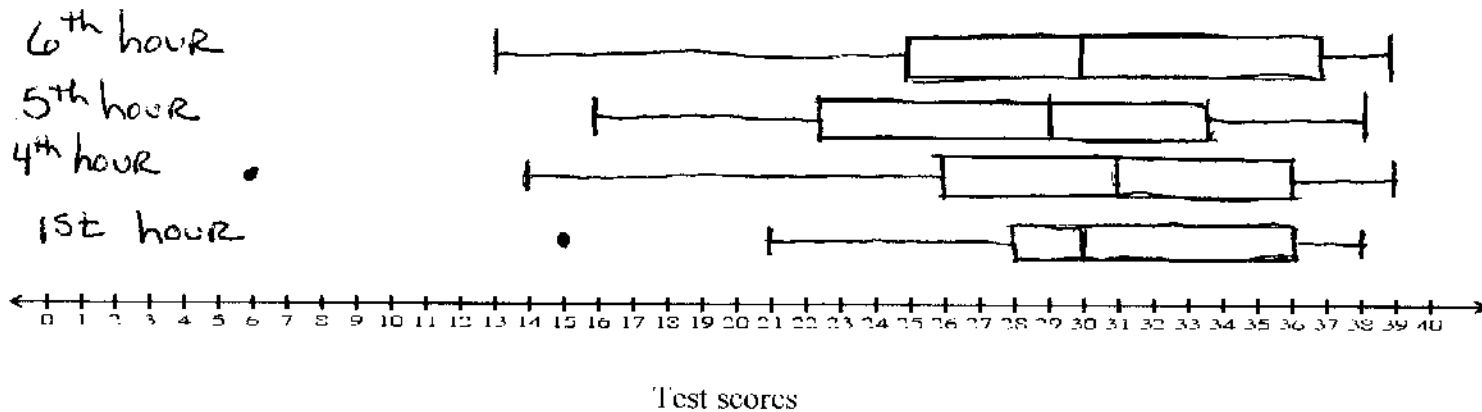
2. For each hour, draw a box and whiskers plot. Plot all 4 on one graph using the same scale.

1<sup>st</sup> hour

4<sup>th</sup> hour

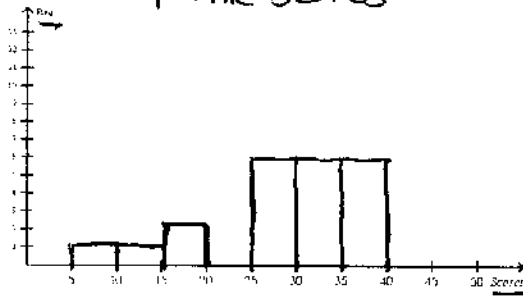
5<sup>th</sup> hour

6<sup>th</sup> hour

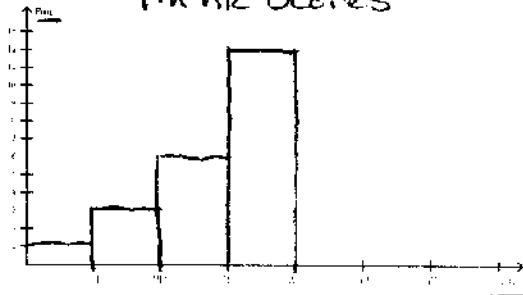


3. Draw a histogram for 4<sup>th</sup> hour's scores using a scale of 5. Draw a histogram for 4<sup>th</sup> hour's scores using a scale of 10.

4<sup>th</sup> hr scores



4<sup>th</sup> hr Scores



4. Using the mean, median, or mode make a case for 1<sup>st</sup> hour's test scores being better than 6<sup>th</sup> hour's test scores. Make a case for 4<sup>th</sup> hour's test scores being better than 5<sup>th</sup> hour's test scores. Which hour do you think has the best overall test score and what did you use (mean, median, or mode) to determine it?

**The mean shows higher test score averages for 1<sup>st</sup> hour than 6<sup>th</sup> hour. Both the mean and median show 4<sup>th</sup> hour's test score average being higher than 5<sup>th</sup> hours. The mode is not a good indicator of the average in this scenario.**

5. Which do you consider (mean, median, or mode) to be the best indicator of the “average” test score and why?

**The mode is not a good indicator of the average because of the scenario and the multiple modes in 4<sup>th</sup> and 6<sup>th</sup> hours. The mean is the best indicator of the average for the test scores. The outliers in 1<sup>st</sup> and 4<sup>th</sup> hours do lower the mean because the data is skewed left, in which case, the median might be considered best.**

6. How can the use of mean, median, and mode be used to mislead or misinform?

**The mean, median, and mode can be used as an “average,” but have much different values for the center of the data set. One can be used to falsely inflate the average.**

7. How does the outlier affect the mean, median, or mode?

**The outlier affects the mean by skewing the data to the left or right which either lowers or raises the mean. The outlier does not affect the median or the mode.**

8. Does considering outliers change your stance for question #4? Why or why not?

**Not really because the mean for 1<sup>st</sup> hour is still be higher than 6<sup>th</sup> hours. If this was not the case, throwing the outlier out of the data could be considered. The median would show the averages for the two hours to be the same.**

9. What is the relationship between percentiles and the median and quartiles?

**The median is the 50<sup>th</sup> percentile. The first quartile is the 25<sup>th</sup> percentile and the third quartile is the 75<sup>th</sup> percentile.**

10. What does it mean when a box and whiskers graph has long whiskers?

**The range of the data has a wide spread which could mean less consistency.**

11. What is the difference between a bar graph and a histogram?

**A bar graph does not have a scale (only labels) on one of the axes. The histogram has a scale on the x-axis for the data values and a scale for the frequency of the data values on the y-axis. In a histogram, changing the scale will change the number and width of the bars. This does not happen on a bar graph**

12. How does changing the x scale on a histogram affect the graph? How can this be used to mislead or misinform?

**Changing the scale can change the look of the graph. Scale changes will affect the number of bars, their widths and their heights. This can lead people to believe there is more of an increase/decrease than there actually is. Refer to problem #3.**