

Pick up the Warm Up/Classwork handout  
You will need your Graphing Calculator

1. Quickly enter the data and find the mean (no work necessary)

123, 234, 95, 45, 199, 234

STAT → CALC → 1-Variable Stat

$$\begin{aligned}\bar{x} &= 139.2857 \\ &= 139\end{aligned}$$

2. Find the weighted average the average number of ice cream sold. Show the formula for the mean, all critical totals and then the answer. [Do so by entering the data in your GDC and quickly calculating the weighted average. GDC = Graphing Display Calculator]

Ice creams sold (summer)	Frequency
20	1
40	2
60	4
80	6
100	8
120	7

$$\bar{x} = \frac{\sum fx}{n} = \frac{\text{critical totals}}{\text{answer}}$$

formula                      critical totals                      answer

$$\text{or } \bar{x} = \frac{\sum fx}{\sum f}$$

2. Find the weighted average the average number of ice cream sold. Show the formula for the mean, all critical totals and then the answer. [Do so by entering the data in your GDC and quickly calculating the weighted average. GDC = Graphing Display Calculator]

Ice creams sold (summer)	Frequency
20	1
40	2
60	4
80	6
100	8
120	7

$$\bar{x} = \frac{\sum fx}{n} = \frac{\text{critical totals}}{\text{answer}}$$

formula                      critical totals                      answer

$$\text{or } \bar{x} = \frac{\sum fx}{\sum f}$$

2. Find the weighted average the average number of ice cream sold. Show the formula for the mean, all critical totals and then the answer. [Do so by entering the data in your GDC and quickly calculating the weighted average. GDC = Graphing Display Calculator]

Ice creams sold (summer)	Frequency
20	1
40	2
60	4
80	6
100	8
120	7

$$\bar{x} = \frac{\sum fx}{n} = \frac{2460}{28} = 87.9 \text{ ice creams}$$

formula
critical totals
answer

$$\text{or } \bar{x} = \frac{\sum fx}{\sum f}$$

1-Variab Stat L<sub>1</sub>, L<sub>2</sub>

3

The heights of 25 hockey players (to the nearest cm):

Height (h cm)	[120, 130)	[130, 140)	[140, 150)	[150, 160)	[160, 170)
Frequency	1	2	7	14	1

$$\frac{25+1}{2}$$

13<sup>th</sup>

- a) Determine the modal height interval [150, 160)  
150 to 160
- b) Determine the median height interval  
[150, 160)
- c) Is the data categorical, quantitative discrete, or quantitative continuous data ?

- 4 Below is data that has been collected and grouped. The actual data is not available. Estimate the mean score, showing the correct formula and critical values. Practice using your GDC lists as a spreadsheet.

mid-interval	Score (x)	Frequency (f)
3	1-5	7
8	6-10	12
13	11-15	15
18	16-20	10
23	21-25	11

$$\bar{x} = \frac{\sum fx}{\sum f} = \text{---} =$$

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{745}{55} = 13.54545... = 13.5$$

## Questions on the assignment ?

When going over HW, only pens of a different color as you add, edit, mark your work.

You should have already checked the answers from the solutions posted.

p. 133  
11

$$14 \text{ matches} \times 16.5 \frac{\text{goals}}{\text{match}} = 231 \text{ total}$$

$$\frac{231 + 21 + 24}{16} =$$

Goal for the Day:

**Use Quartiles and Box Plots to help see the distribution of data and to analyze it.**

**Calculate Outliers**

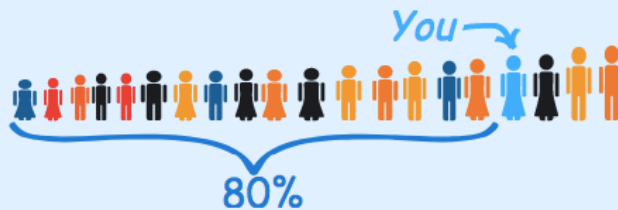
# First some Visuals

## Percentiles

Percentile: the value below which a percentage of data falls.

Example: You are the fourth tallest person in a group of 20

80% of people are shorter than you:



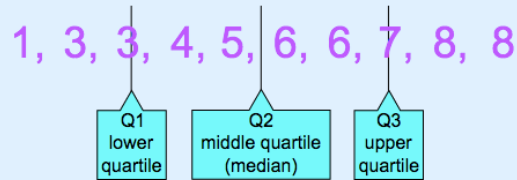
That means you are at the **80th percentile**.

If your height is 1.85m then "1.85m" is the 80th percentile height in that group. ←

# Some Special Percentiles

Example: 1, 3, 3, 4, 5, 6, 6, 7, 8, 8

The numbers are in order. Cut the list into quarters:



In this case Quartile 2 is half way between 5 and 6:

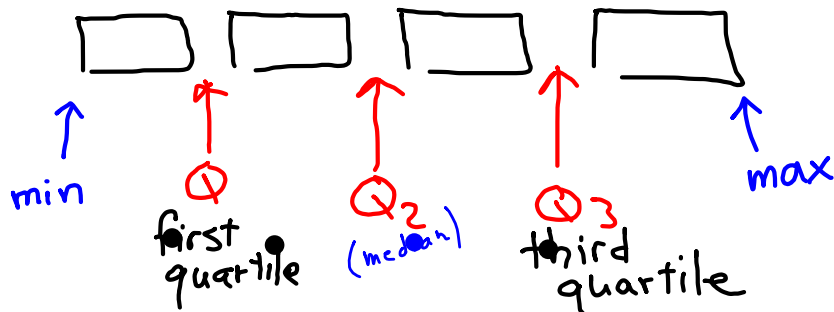
$$Q2 = (5+6)/2 = 5.5$$

And the result is:

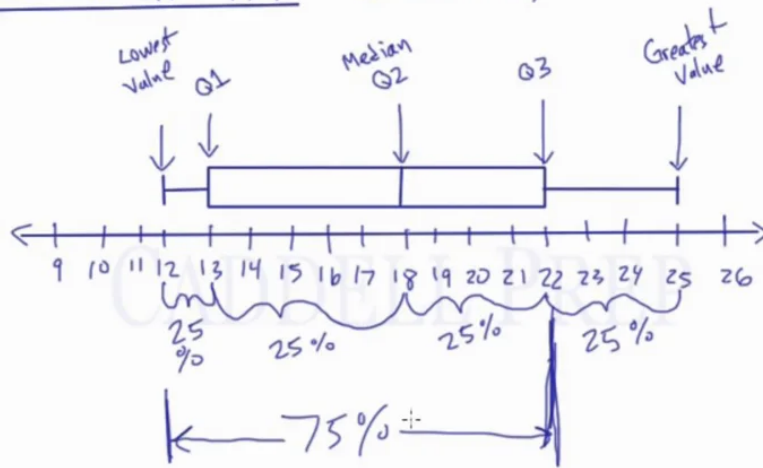
- Quartile 1 (Q1) = 3
- Quartile 2 (Q2) = 5.5
- Quartile 3 (Q3) = 7

The Quartiles also divide the data into divisions of 25%, so:

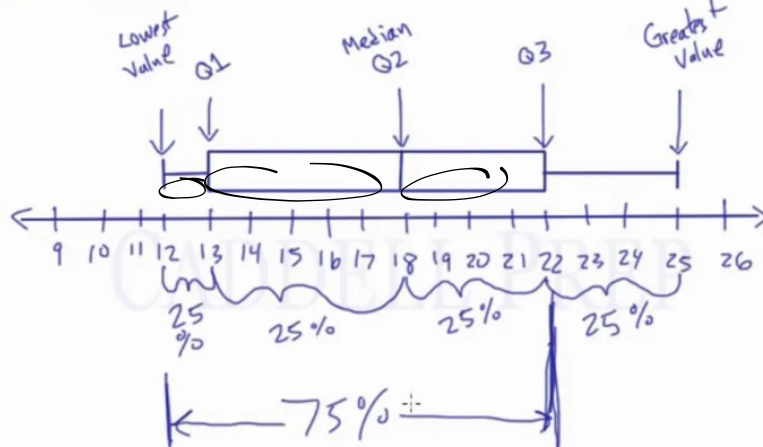
- Quartile 1 (Q1) can be called the **25th percentile**
- Quartile 2 (Q2) can be called the **50th percentile**
- Quartile 3 (Q3) can be called the **75th percentile**



## Box & Whisker Plot (Quartiles)

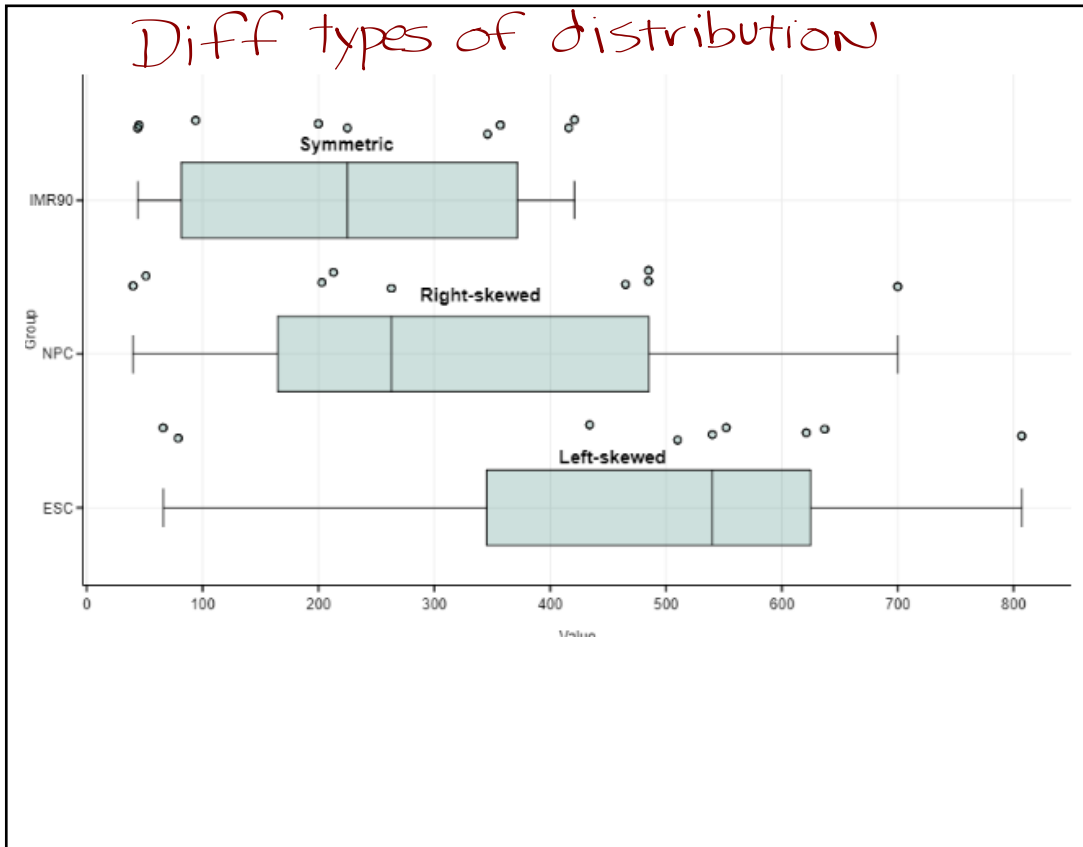
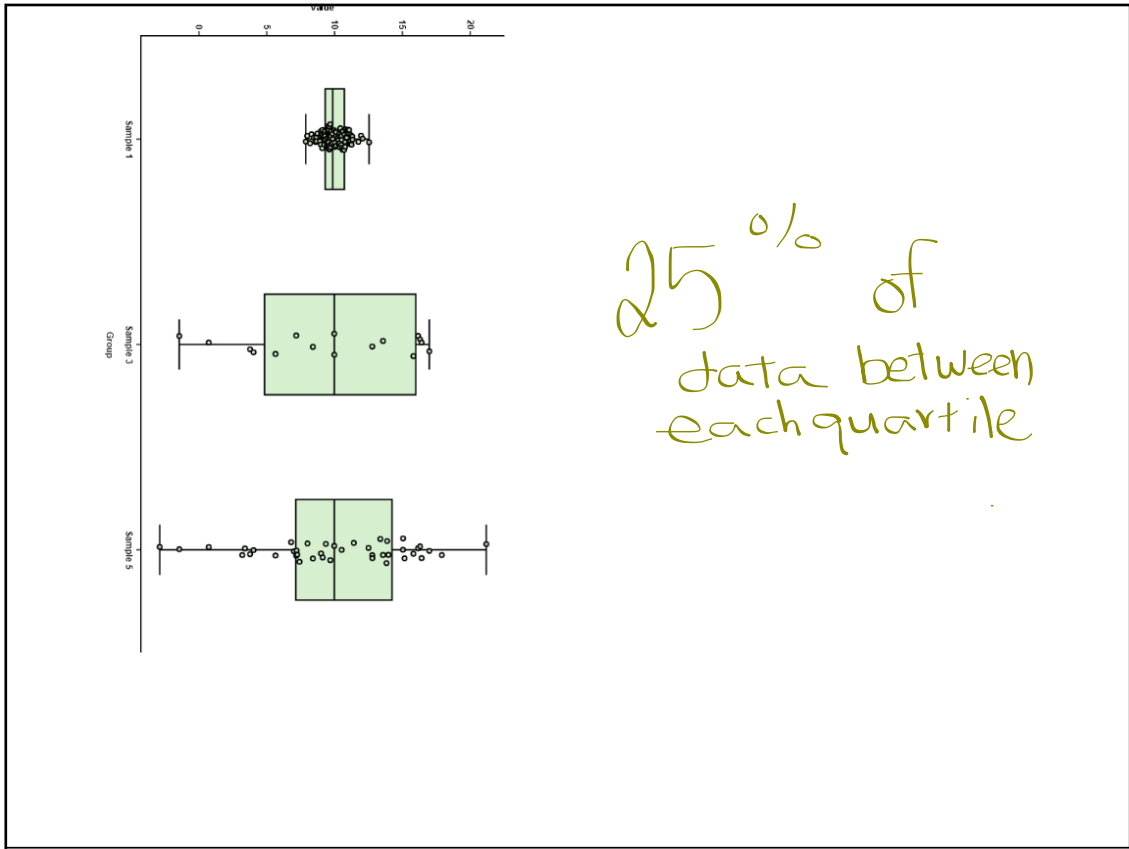


## Box & Whisker Plot (Quartiles)

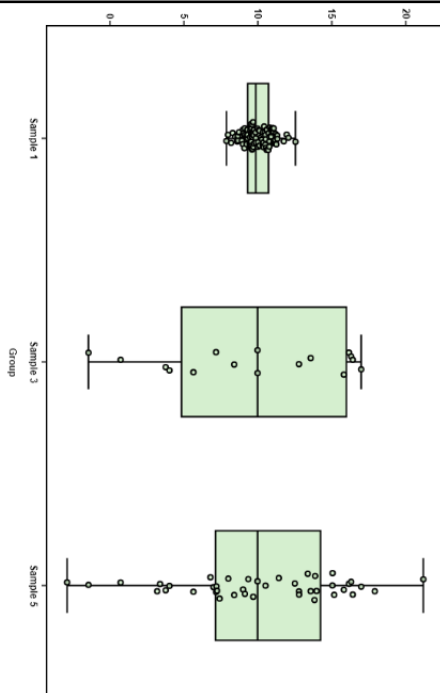
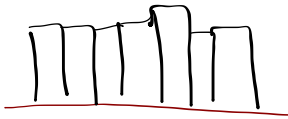


Chunks  
of  
25%





Approximately  
Uniform



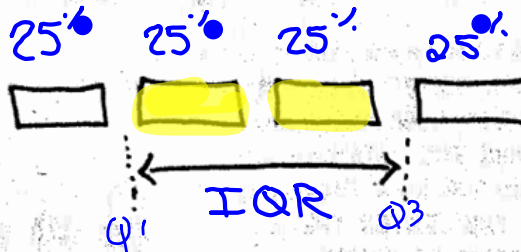
The whiskers are not as reliable for measuring variability in the data

Single extreme values can completely change the whisker.

AGAIN, THERE'S MORE THAN ONE WAY TO MEASURE A SPREAD. ONE WAY IS

## INTERQUARTILE RANGE

THE IDEA IS TO DIVIDE  
THE DATA INTO FOUR  
EQUAL GROUPS AND SEE  
HOW FAR APART THE  
EXTREME GROUPS ARE.



$$IQR = Q3 - Q1$$

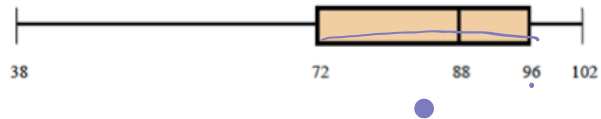
I have a hunch you know at least a little about Box Plots. Let's see what you know.

Pick up the hand out "Interpreting a Box & Whisker Plot

Work through as many as you can. Check in with each other for answers

Refer to the box & whisker graph below which shows the test results of a math class

Test Scores (as %) for 9<sup>th</sup> Period

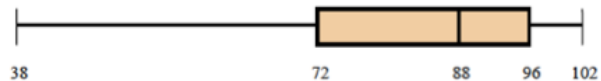


- 102 a. What was the high score on the test?  
75% b. What percent of the class scored above a 72?  
88 c. What was the median score on the test?  
25% d. What percent of the class scored between 88 & 96?  
cbd e. How many students scored a 38?  
24 f. What is the range of scores that the middle 50% of the sophomores. In other words, what is the IQR, *interquartile range*?

A typical interpretation of the IQR would go like this. "The middle 50 % of the scores has a variation of 24 points."

Refer to the box & whisker graph below which shows the test results of a math class

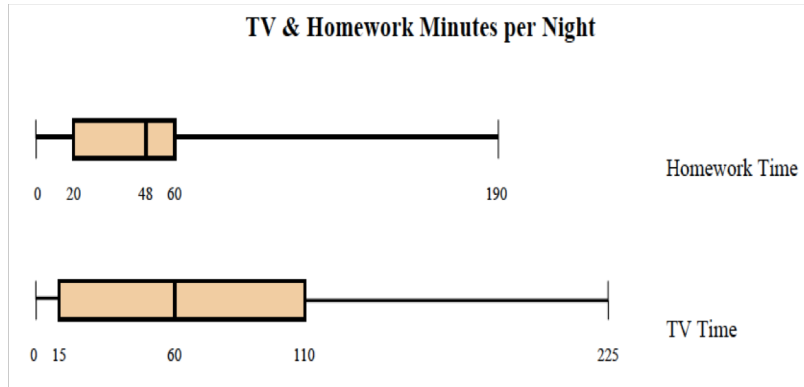
Test Scores (as %) for 9<sup>th</sup> Period



- 102 a. What was the high score on the test?  
75 b. What percent of the class scored above a 72?  
88 c. What was the median score on the test?  
25% d. What percent of the class scored between 88 & 96?  
cbd e. How many students scored a 38?  
24 f. What is the range of scores that the middle 50% of the sophomores. In other words, what is the IQR, *interquartile range*?

A typical interpretation of the IQR would go like this. "The middle 50 % of the scores has a variation of 24 points."

2.



**Identify if each statement is true, false, or cannot be determined (cbd)**

- \_\_\_\_\_ a. Some sophomores didn't watch TV that month.
- \_\_\_\_\_ b. The TV box & whisker graph contains more data than the homework graph.
- \_\_\_\_\_ c. 25% of the sophomores spend between 48 & 60 minutes per night on homework.
- \_\_\_\_\_ d. 15% of the sophomores didn't watch TV that month.
- \_\_\_\_\_ e. In general, these sophomores spend more time watching TV than doing homework.
- \_\_\_\_\_ f. The TV data is has more variation than the homework data.
- \_\_\_\_\_ g. 225 sophomores watch TV.
- \_\_\_\_\_ h. The 3<sup>rd</sup> quartile of the TV data is from 110 to 225 minutes.

identify if each statement is true, false, or cannot be determined (cbd)

- T a. Some sophomores didn't watch TV that month.
- F ~~cbd~~ b. The TV box & whisker graph contains more data than the homework graph.
- T c. 25% of the sophomores spend between 48 & 60 minutes per night on homework.
- cbd d. 15% of the sophomores didn't watch TV that month.
- T e. In general, these sophomores spend more time watching TV than doing homework.
- T f. The TV data is has more variation than the homework data.
- cbd g. 225 sophomores watch TV.
- F h. The 3<sup>rd</sup> quartile of the TV data is from 110 to 225 minutes.

## How many concerts have you been to ?

Record you values. We will use this data to investigate outliers and boxplots.

2	0	6	2	10
2	2	10	2	3
1	1	3	5	4
<del>1</del>	0	3	1	40
1	4	2	6	5
0	2	2	7	49
0	1	30	11	
0	0	0	35	

# How many concerts have you been to?



# JAY-Z



**How many concerts have you been to? We will use this data to investigate outliers and boxplots.**

1. Enter the data in list 1 ( $L_1$ ) of your graphing calculator, but out any zeros. While you are waiting for the others to finish, see if you can figure out how to have your calculator produce the 5-number summary.
2. Now calculate the 5-number summary.

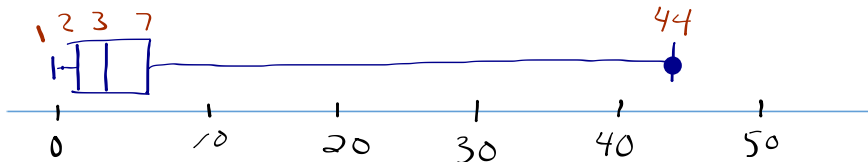
Minimum = 1      Q1 = 2      Median = 3      Q3 = 7      Maximum = 44

↑  
Same as  
Q<sub>2</sub>

2. Now calculate the 5-number summary.

Minimum = 1      Q1 = 2      Median = 3      Q3 = 7      Maximum = 44

3. At this point, we can make a box plot. *Make box plots above a single number line, labeled with an easy to read values on the scale.*



# of concerts attended

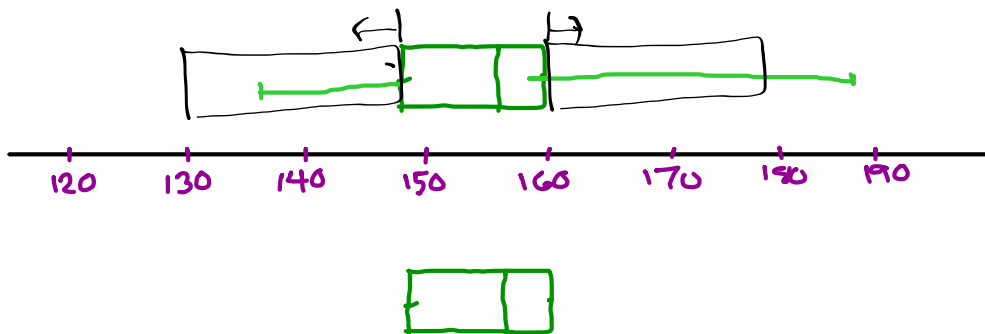
4. We will now add an extreme value to our data set (if needed). Determine if there are outliers.

a. Find IQR:  $5$  concerts Interpret the IQR *The middle half of #concerts attended varies by 5 concerts*  
 $7-2$

b. Calculate  $Q1 - (1.5 \times IQR)$ :  $2 - 1.5(5) = -5.5$   
Any number less than this is a low outlier. Any low outliers? no

c. Calculate  $Q3 + (1.5 \times IQR)$ :  $7 + 1.5(5) = 14.5$   
Any number more than this is a high outlier. Any high outliers? 30 35 40 44

A visual understanding of an outlier



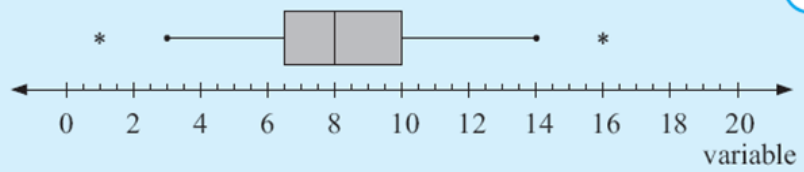




- **The upper boundary = upper quartile + 1.5 × IQR.**  
Any data larger than the upper boundary is an outlier.
- **The lower boundary = lower quartile – 1.5 × IQR.**  
Any data smaller than the lower boundary is an outlier.

As 16 is above the upper boundary it is an outlier.  
As 1 is below the lower boundary it is an outlier.

So, the boxplot is:

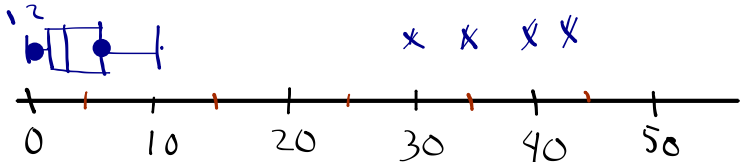


Notice that the whisker is drawn to the last value that is not an outlier.





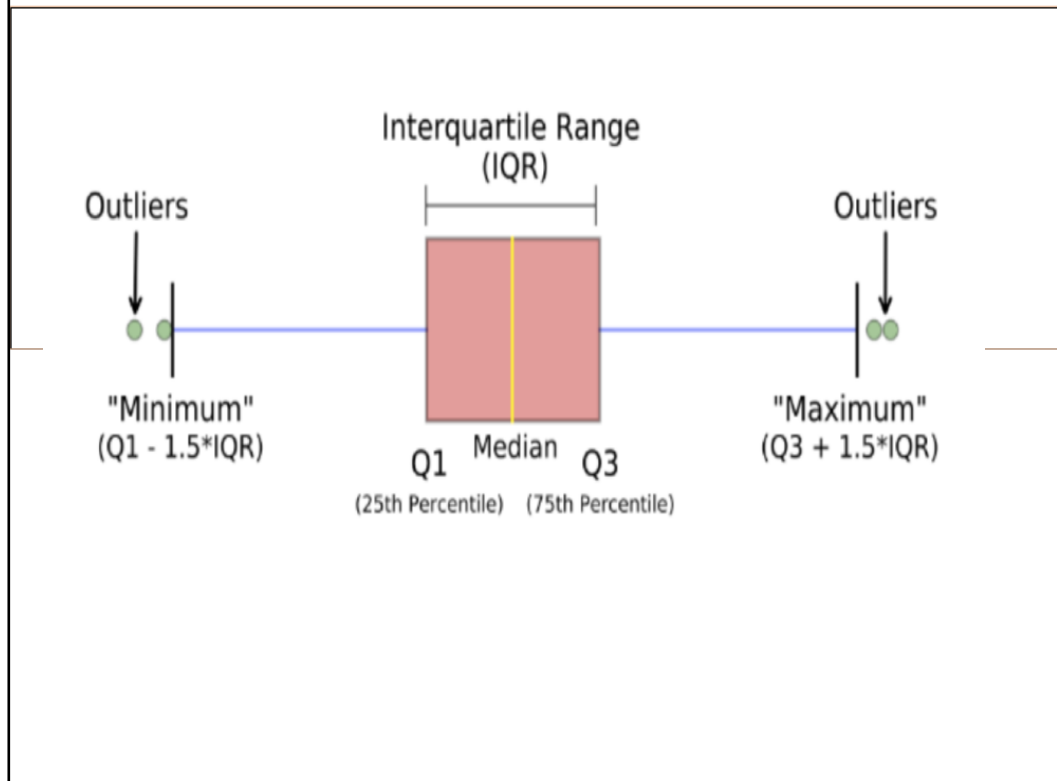
5. Now make a new Box Plot that shows outliers. Same scale as above.



## Summary: Boxplots and Outliers



## Summary: Boxplots and Outliers



### Application - Which is best in reducing stress?

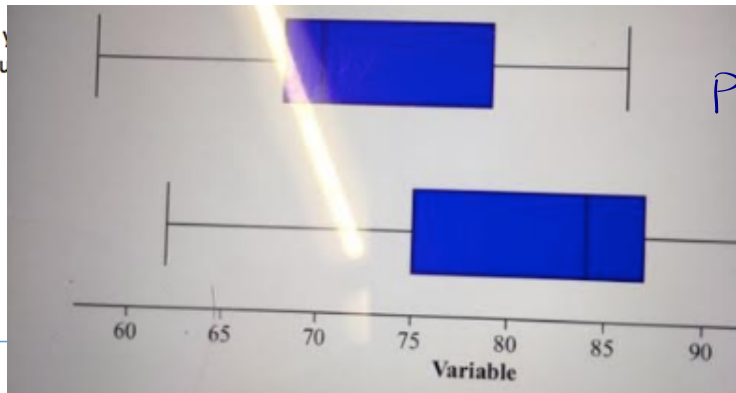
If you are a dog lover, having your dog with you may reduce your stress level. Does having a friend with you reduce stress? To examine the effect of pets and friends in stressful situations, researchers recruited 30 women who said they were dog lovers. Fifteen women were assigned at random to two groups: to do a stressful task alone or with a with their dog present. The stressful task was to count backward by 13s or 17s. The woman's average heart rate during the task was one measure of the effect of stress. The table below shows the data.

Pet 58 64 65 68 69 69 69 70 70 72 76 79 85 86 99

Alone 62 70 73 75 77 80 84 84 84 87 87 87 90 91 99

1. Use the rule to show that 99 is an outlier for the Pet group. Actually, write down the entire calculations

2. Use your group



Heart rates in the three

Pet

Alone

3. **Interpretation:** Based on the boxplots, does it appear that the presence of a pet or friend reduces heart rate during a stressful task? Justify your answer.

## Assignment

Finish the handout from class and....

### **HH Ch. 6 Packet**

p.169....3

p.174.... 12

p.175...1

p.182.... 4

p.184....2

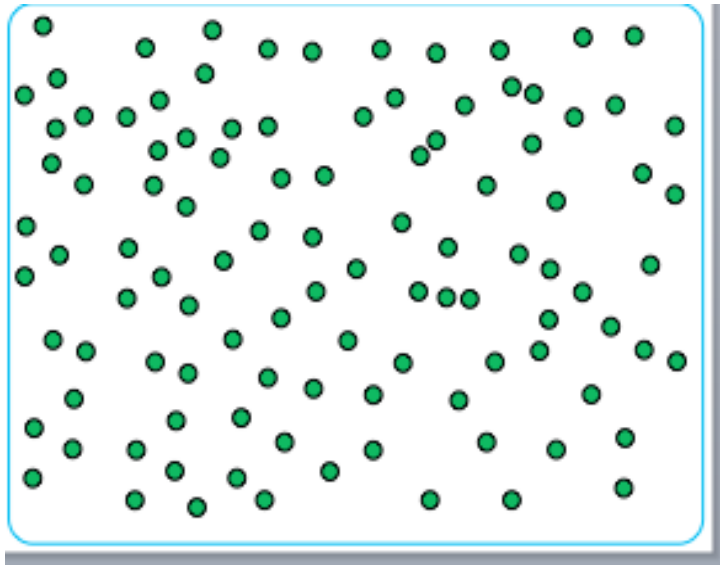
p.188....2

How many dots?

On the next slide there will be a number of dots, the slide will be shown for approximately 1 second.

Estimate how many dots do you think there are?

No cheating, just write down your first estimate – do not share this with anyone at this stage.



Write down your answer

Do not change or share at this point

let's enter to make a box plot

so we can see

----- the distribution of the guesses

----- the variation of guesses

Before I tell you the answer,  
you get to the opportunity  
re-guess

From the re-guesses we'll  
make a new plot

## Phenomena

Wisdom of a crowd acting  
independently is often more  
accurate than

GUESS 1  
mean

Crowd psychology where  
there is interaction.

GUESS 2  
mean



Answer

104 dots

