

(A) Lesson Context

BIG PICTURE of this UNIT:	<ul style="list-style-type: none"> • How do we analyze and then make conclusions from a data set? • How do I present my data and the outcomes of my analysis? • How do I use data & statistics to make decisions? • How do I decide on the validity/reliability of my data? Of my analysis? Of my conclusions? Of my decision? 		
CONTEXT of this LESSON:	<p>Where we've been</p> <p>Determining the central tendency of a data set, given an assortment of data presentations</p>	<p>Where we are</p> <p>Measuring the dispersion of a data set, using range & quartiles and visualizing the dispersion using a box and whisker plot</p>	<p>Where we are heading</p> <p>How do I analyze and make conclusions from a data set, in whatever way this data gets presented?</p>

(B) Lesson Objectives:

- a. Calculate the range of a data set
- b. Calculate the quartiles of a data set
- c. Present the quartiles using a visual → a box & whisker plot

(C) Opening Activity → Which Class did “better”?

<p>Central Tendencies: Results Summary from Class A</p>	<p>Central Tendencies: Results Summary from Class B</p>
<p>—Mean grade → 63%</p> <p>—Median grade → 60%</p> <p>—Modal Grade → 60%, 80%</p>	<p>—Mean grade → 63%</p> <p>—Median grade → 60%</p> <p>—Modal Grade → 60%</p>
<p>Can we determine which class did “better” if all we know are the central tendencies?</p>	

(D) Opening Activity → Which Class did “better” – More Information added?

Central Tendencies: Results Summary from Class A	Central Tendencies: Results Summary from Class B
—Mean grade → 63% —Median grade → 60% —Modal Grade → 60% —Range → 70%	—Mean grade → 63% —Median grade → 60% —Modal Grade → 60% —Range → 70%
Can we determine which class did “better” after we have now added the range to the central tendencies?	

(E) Measures of Spread/Dispersion/Variance – Range

So what else do we need? → we need some way to come up with some other “summary numbers” that allow us to get a picture of the “spread” or “dispersion” or “distribution” or “variance” in the data

One such number is the RANGE (as we have discussed before)

The range is simply the difference between the largest and the smallest observed values in a data set. Thus, range, including any outliers, is the actual spread of data. A great deal of information is ignored when computing the range, since only the largest and smallest data values are considered.

(F) Measures of Spread/Dispersion/Variance – Quartiles

Quartiles simply split data into quarters:

The first quartile or the lower quartile is $\frac{1}{4}$ of the way into an ordered data set → cuts off lowest 25% of data = **25th percentile**

The third quartile or the upper quartile is $\frac{3}{4}$ of the way into an ordered data set → **third quartile** (designated Q_3) = **upper quartile** = cuts off highest 25% of data, or lowest 75% = **75th percentile**

The second quartile or the median is $\frac{2}{4}$ (or $\frac{1}{2}$) of the way into an ordered data set → **second quartile** (designated Q_2) = **median** = cuts off data set in half (50% of data) = **50th percentile**

(G) Example #1 – Analyzing Quiz Scores → 5 Number Summary

Here are QUIZ results from BOYS in our IM2 classes. The scores have already been sorted/ordered.

3, 3, 3, 3, 4, 4, 4, 5, 6, 6, 7, 7, 8, 8, 8, 8, 8, 9, 9, 9, 9, 10, 10

- (a) Determine the minimum score →
- (b) Determine the maximum score →
- (c) Calculate the range of the scores →
- (d) Determine the lower quartile score →
- (e) Determine the median score →
- (f) Determine the upper quartile score →

Here are QUIZ results from GIRLS in our IM2 classes. The scores have already been sorted/ordered.

4, 4, 5, 5, 5, 6, 6, 6, 7, 8, 8, 8, 9, 9, 9, 9, 9, 9, 10, 10, 10, 10

- (a) Determine the minimum score →
- (b) Determine the maximum score →
- (c) Calculate the range of the scores →
- (d) Determine the lower quartile score →
- (e) Determine the median score →
- (f) Determine the upper quartile score →

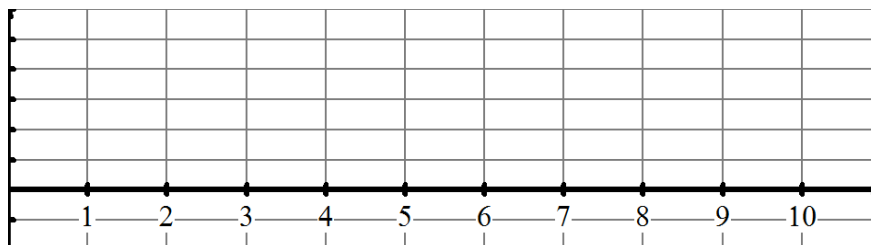
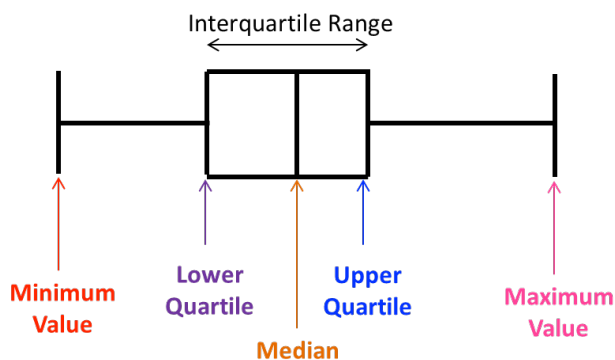
What observations & conclusions can we make from looking at the results of the data calculations?

(H) Example #1 – Analyzing Quiz Scores → Visualizations with Box & Whisker Plots

A box plot is a way of illustrating key information about a set of data

They are also very useful for comparing the distribution of two sets of data

To draw a box plot, you need FIVE pieces of information: (i) The minimum value, (ii) the lower quartile (iii) the median, (iv) the upper quartile (v) the maximum value



Example #2 – Comparing Golf Clubs

Below are the ages of 15 members of a golf club.

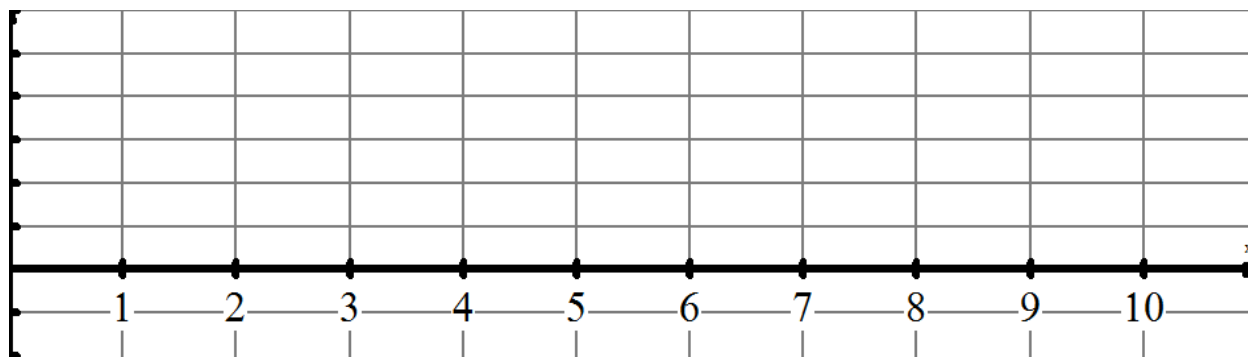
20, 20, 21, 24, 27, 29, 33, 33, 36, 39, 50, 57, 60, 65, 65

Below are the ages of members of a second golf club.

18, 23, 25, 27, 27, 37, 38, 45, 47, 49, 49, 50, 50, 58, 61, 66, 69

- (a) Determine the minimum age →
- (b) Determine the maximum age →
- (c) Calculate the range of the age →
- (d) Determine the lower quartile age →
- (e) Determine the median age →
- (f) Determine the upper quartile age →

- (g) Determine the minimum age →
- (h) Determine the maximum age →
- (i) Calculate the range of the age →
- (j) Determine the lower quartile age →
- (k) Determine the median age →
- (l) Determine the upper quartile age →



- Which golf club has more younger members?
- Which golf club has more older members?
- What can you say about the age profile of each golf club in general?

(I) Example #3

1. Here are the marks scored by 15 students in their algebra test.

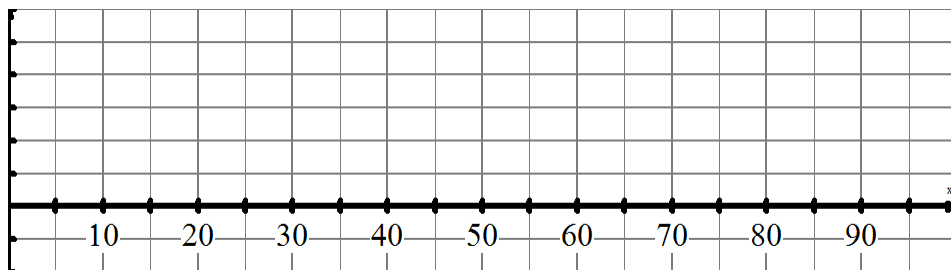
75 82 69 82 95 80 74 77 60 51 68 84 90 100 45

a) Find the median, quartiles and interquartile range for these data

b) In a probability test the following summary data was collected

Minimum	Lower Quartile	Median	Upper Quartile	Maximum
60	72	78	84	98

c) On the same axis draw box plots for both tests.



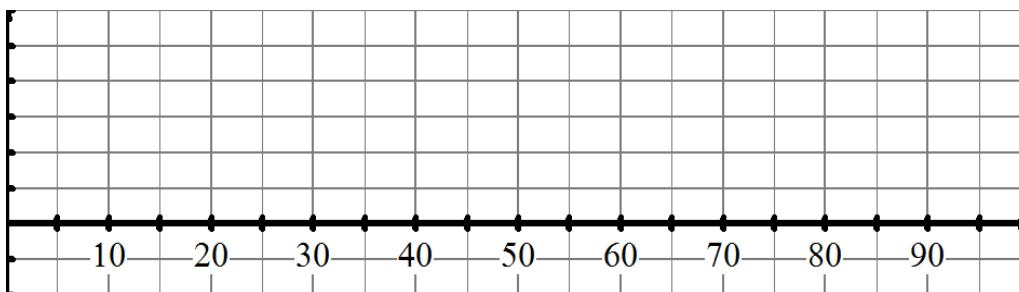
d) Compare and contrast the results from the two tests. (Which test did they do better? Which test had more varied results?).

2. This back-to-back stem and leaf diagram shows heart rates of two groups of students before and after they walk up the stairs

Before							After						
	8	8	4	3	2	6	8						
9	9	7	5	4	3	7	0	1	5	6	6	7	9
			3	2	0	8	0	0	1	4	8		
					0	9	2						
						10	0						

legend | 6 | 7 = 67 beats per minute

- a) Find the median and quartiles for these data.
- b) Draw a box plot for each on the same scale



- c) Compare the heart rates before and after the walk up the stairs.

3. The data below shows the price of petrol (to the nearest pence) at different locations for Shell and BP.

SHELL: 130, 129, 132, 133, 136, 130, 129, 130, 130, 131, 133, 134, 140

BP: 131, 133, 133, 133, 132, 134, 138, 130, 133, 134, 135, 136, 137, 132, 135

- a) Calculate the quartiles, median and interquartile range for each subject
- b) Use your data to draw box plots for each subject
- c) Comment on at least two aspects of your box plots to compare the IQs of the graduates.

4. This summary data has been collected on times taken, in minutes, to run 5km in two different fun runs in Victoria Park.

Run 1	
Median	35
Minimum	23
Range	40
Interquartile range	12
Upper Quartile	42

Run 2	
Median	36
Maximum	55
Range	30
Interquartile range	15
Lower Quartile	29

- a) Draw box plots for both runs using the same scale
- b) Write down two comparisons between the races
- c) Emma says that Race 1 attracted more serious runners. Use your box plots to comment on what Emma has said.
5. The data below shows the IQ of 11 Maths and 11 Geography university graduates.
- MATHS: 98, 103, 105, 99, 110, 94, 98, 100, 120, 117, 113
 GEOGRAPHY: 93, 99, 110, 111, 95, 97, 90, 99, 92, 102, 103
- a) Calculate the quartiles, median and interquartile range for each subject
- b) Use your data to draw box plots for each subject
- c) Comment on at least two aspects of your box plots to compare the IQs of the graduates.