

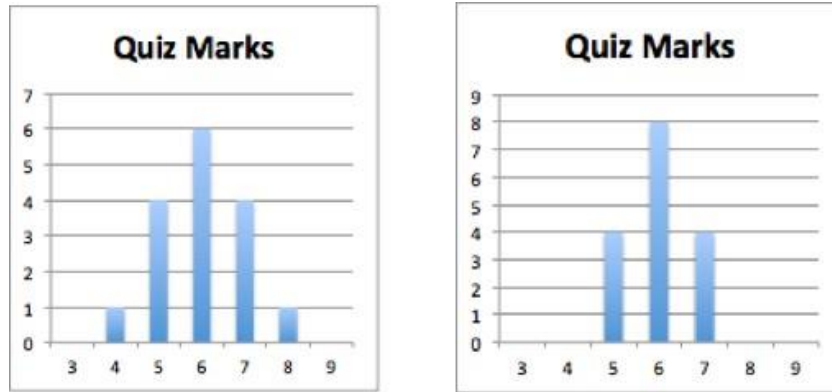
## Section 3.3 – Measures of Spread

MDM4U

David Chen

### Part 1: Minds On

Describe the similarities and differences between the following two graphs:

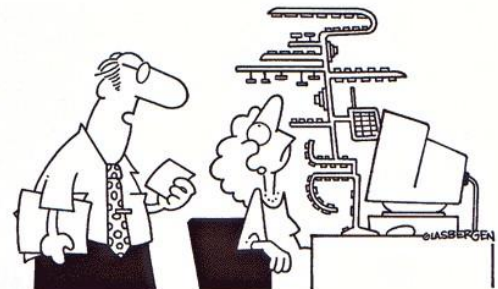


- they both have 16 elements
- mean, median, and mode are all 6
- the data is more spread out on the left
- the data on the right graph is more tightly clustered around the mean.

In the previous section we learned how to describe a set of data using measures of central tendency. Just as there are several measures of central tendency, there are also different measures of spread for a set of data:

- **Interquartile Ranges**
- **Standard Deviation**
- **Percentiles**

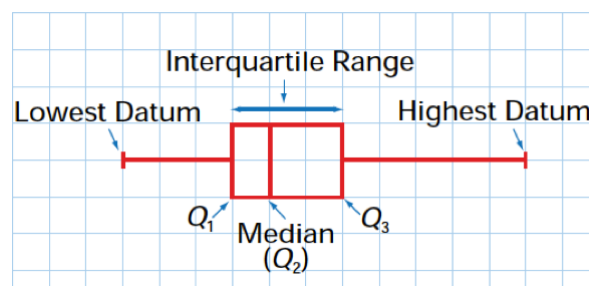
Copyright 2002 by Randy Glasbergen.  
www.glasbergen.com



"It's the new keyboard for the statistics lab. Once you learn how to use it, it will make computation of the standard deviation easier."

### Part 2: Interquartile Range (IQR)

A **boxplot** is a visual representation of data divided into four groups (quartiles) with equal numbers of values in each quartile.



The three dividing points are:

$Q_1$  – median of the lower half of the data

$Q_2$  – median of all of the data (must calculate this first)

$Q_3$  – median of the upper half of the data

The interquartile range is the range of the middle half of the data ( $= Q_3 - Q_1$ ). The larger the interquartile range, the larger the spread of the central half of the data. This is a good measure of spread (dispersion) but can be tedious to calculate.

**Remember:** The median is the middle value in an ordered data set. If the set has an even number of data points, then the median is halfway between the two middle-most values.

**Example 1:** Here are the final grades in a grade 9 physical education class:

88, 56, 72, 67, 59, 48, 81, 62  
90, 75, 75, 43, 71, 64, 78, 84

a) First put the data into numerical order

43, 48, 56, 59, 62, 64, 67, 71, 72, 75, 75, 78, 81, 84, 88, 90

This is always the first step  
when determining the quartiles

b) Calculate the following statistics:

Range = 47

Median =  $Q_2$  = 71.5

$Q_1$  = 60.5

$Q_3$  = 79.5

IQR =  $Q_3 - Q_1 = 79.5 - 60.5 = 19$

**Note:** although a quartile is, strictly speaking, a single value, people sometimes speak of datum being within a quartile. What they really mean is that the datum is in the quarter whose upper boundary is the quartile. Example: if a value  $x$  is within the second quartile, then  $Q_1 < x \leq Q_2$

c) If your final mark was 75%, what quartile were you within?

That is between  $Q_2$  and  $Q_3$ ; therefore it is within the 3<sup>rd</sup> quartile.

### Part 3: Standard Deviation

The deviation of a piece of data is the **distance** it is from the **mean** of the set of data. If you were to take all the deviations for an entire set of data, square each one of them, and then find the average, you would have what is called the variance. The square root of the variance is called the **standard deviation**.

Population Standard Deviation	Sample Standard Deviation
$\sigma = \sqrt{\frac{\sum(x - \mu)^2}{N}}$ <p><math>\sigma</math> – lower case sigma; population standard deviation</p> <p><math>\sum</math> --- capital sigma; sum of</p> <p><math>\mu</math> --- mu; population mean</p> <p><math>N</math> --- number of items in population</p>	$s = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}}$ <p><math>s</math> – sample standard deviation</p> <p><math>\sum</math> --- capital sigma; sum of</p> <p><math>\bar{x}</math> – x-bar; sample mean</p> <p><math>n</math> --- number of items in sample</p>

**Example 2:** The heights of all of the players on a basketball team are shown in the table below. Calculate the standard deviation of the population.

Start by calculating the mean:

$$\mu = \frac{183 + 165 + \dots + 154}{7} = \frac{1155}{7} = 165$$

Player	Height	Deviation, $x - \mu$	$(x - \mu)^2$
Laura	183	18	324
Jamie	165	0	0
Deepa	148	-17	289
Colleen	146	-19	361
Ingrid	181	16	256
Justiss	178	13	169
Sheila	154	-11	121
			$\sum(x - \mu)^2 = 1520$

$$\sigma = \sqrt{\frac{\sum(x - \mu)^2}{N}} = \sqrt{\frac{1520}{7}} = 14.7$$

Therefore, the mean height of the team is 165 cm with a standard deviation of 14.7 cm.

## Part 4: Comparing Standard Deviations

**Example 3:** Felix and Melanie have a job laying patio stones. Their boss is interested in who the better worker is so randomly throughout the week he chooses a few hours to record how many stones each of the workers lays. The data is recorded in the table below:

<b>Felix</b>	34	41	40	38	38	45
<b>Melanie</b>	51	28	36	44	41	46

Calculate the mean and standard deviation of each sample and compare use them to compare the two workers.

Using the Ti-84 to calculate these statistics:

- Input Felix's data in to L1: STAT → EDIT
- Calculate the statistics for Felix: STAT → CALC → 1-VARSTATS → List: L1 → Calculate

NORMAL FLOAT AUTO REAL RADIAN MP	NORMAL FLOAT AUTO REAL RADIAN MP	NORMAL FLOAT AUTO REAL RADIAN MP
<b>L1</b>	<b>1-Var Stats</b>	<b>1-Var Stats</b>
34	List:L1	$\bar{x}=39.33333333$
41	FreqList:	$\Sigma x=236$
40	Calculate	$\Sigma x^2=9350$
38		$Sx=3.669695719$
38		$\sigma x=3.34995854$
45		$n=6$
		$\min X=34$
		$\downarrow Q1=38$
L1(?)=		

Statistics for Felix:

$$\bar{x} = 39.33$$

$$s = 3.67$$

- Input Melanie's data in to L2: STAT → EDIT
- Calculate the statistics for Melanie: STAT → CALC → 1-VARSTATS → List: L2 → Calculate

NORMAL FLOAT AUTO REAL RADIAN MP	NORMAL FLOAT AUTO REAL RADIAN MP	NORMAL FLOAT AUTO REAL RADIAN MP
<b>L2</b>	<b>1-Var Stats</b>	<b>1-Var Stats</b>
34	List:L2	$\bar{x}=41$
41	FreqList:	$\Sigma x=246$
40	Calculate	$\Sigma x^2=10414$
38		$Sx=8.099382693$
38		$\sigma x=7.393691004$
45		$n=6$
		$\min X=28$
		$\downarrow Q1=36$
L2(?)=		

Statistics for Melanie:

$$\bar{x} = 41$$

$$s = 8.10$$

On average, it appears that Melanie lays more patio stones per hour. However, her standard deviation is larger which means that her data is more widely spread indicating that she is a less consistent worker.

## Part 5: Standard Deviation for Data with Weights/Frequencies

A railway line gives out small bags of peanuts to its travellers, and each bag doesn't always contain the same number of peanuts. The following table represents a sample of 31 bags showing the number of peanuts per bag. Calculate the sample standard deviation for the number of peanuts in each bag.

<b>Number of Peanuts</b>	28	29	30	31	32	33
<b>Frequency</b>	2	5	10	9	4	1

**Solution:**

- Input number of peanuts in to L1: STAT → EDIT
- Input frequencies in to L2
- Calculate statistics: STAT → CALC → 1-VARSTATS → List: L1 → Frequencies: L2 → CALCULATE

L1	L2	L3	L4	L5	2
28	2	---	---	---	
29	5				
30	10				
31	9				
32	4				
33	1				
---	---				

1-Var Stats

List:L1  
FreqList:L2  
Calculate

1-Var Stats

$\bar{x}=30.35483871$   
 $\Sigma x=941$   
 $\Sigma x^2=28607$   
 $Sx=1.198565451$   
 $\sigma x=1.179075283$   
 $n=31$   
 $\min X=28$   
 $Q1=30$

On average there are about 30.355 peanuts per bag with a standard deviation of 1.199 peanuts.