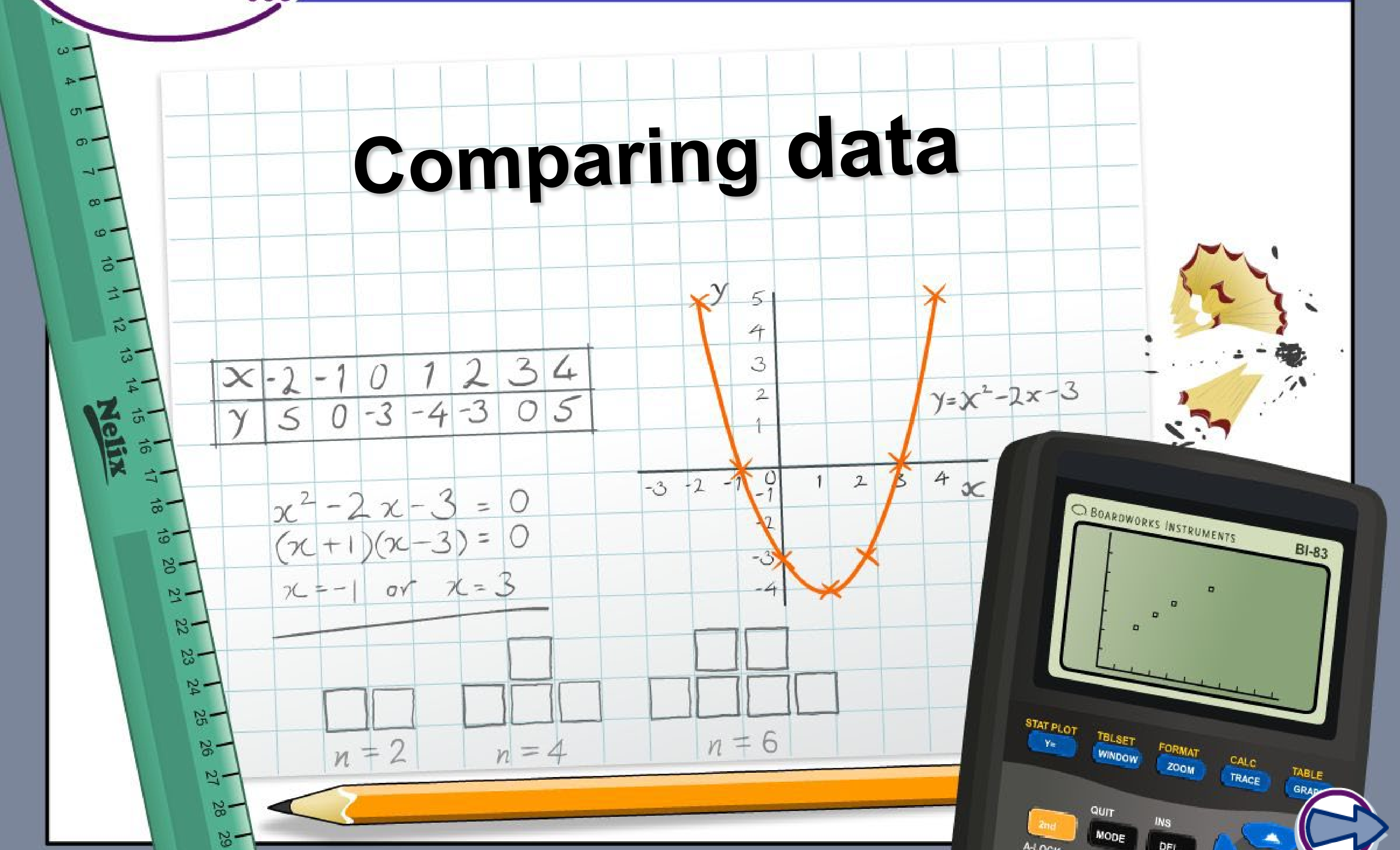
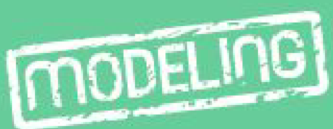
# Comparing data



# Information

## Common core icons

This icon indicates a slide where the Standards for Mathematical Practice are being developed. Details of these are given in the Notes field.



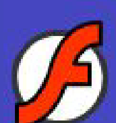

Slides containing examples of mathematical modeling are marked with this stamp.

This icon indicates an opportunity for discussion or group work.

The **Standards for Mathematical Practice** outlined in the Common Core State Standards for Mathematics describe varieties of expertise that mathematics educators at all levels should seek to develop in their students.

They are:

1) **Make sense of problems and persevere in solving them.**
2) **Reason abstractly and quantitatively.**
3) **Construct viable arguments and critique the reasoning of others.**
4) **Model with mathematics.**
5) **Use appropriate tools strategically.**
6) **Attend to precision.**
7) **Look for and make use of structure.**
8) **Look for and express regularity in repeated reasoning.**



This icon indicates that the slide contains activities created in Flash. These activities are not editable.

This icon indicates teacher's notes in the Notes field.

# Measures of central tendency

Sets of data can be compared and interpreted using different **measures of central tendency**.

Measures of central tendency provide a typical value for a set of data. There are three commonly used types:

**MODE**
most common value

**MEAN**

$$\frac{\text{sum of values}}{\text{number of values}}$$

**MEDIAN**
middle value

The **range** is not a measure of central tendency, but it tells you how the data is **spread**:

**RANGE**
largest value – smallest value

# Which measure of central tendency?

MEAN | MEDIAN | MODE

It is important to be able to choose the measure of central tendency that is most appropriate to describe a set of data.

Press on each of the tabs to see when each of the measures of central tendency should be used.

# Mean, median or mode?

When we use the word 'average' we are generalizing and we could be using any of the measures of central tendency.

**Discuss which measure of central tendency might have been used to produce each of these statements.**

a) The average man is 46 years old.

b) The average goals per soccer game is 2.72.

c) The average wage is $36,450.

d) The average age of a 7th grade class is 12.6.

e) The average age of the same class is 13.

f) The average American drives a Ford.

g) The average family has 0.8 pets.

h) The average grade 8 student's favorite colour is red.

# Outliers affecting median and mean

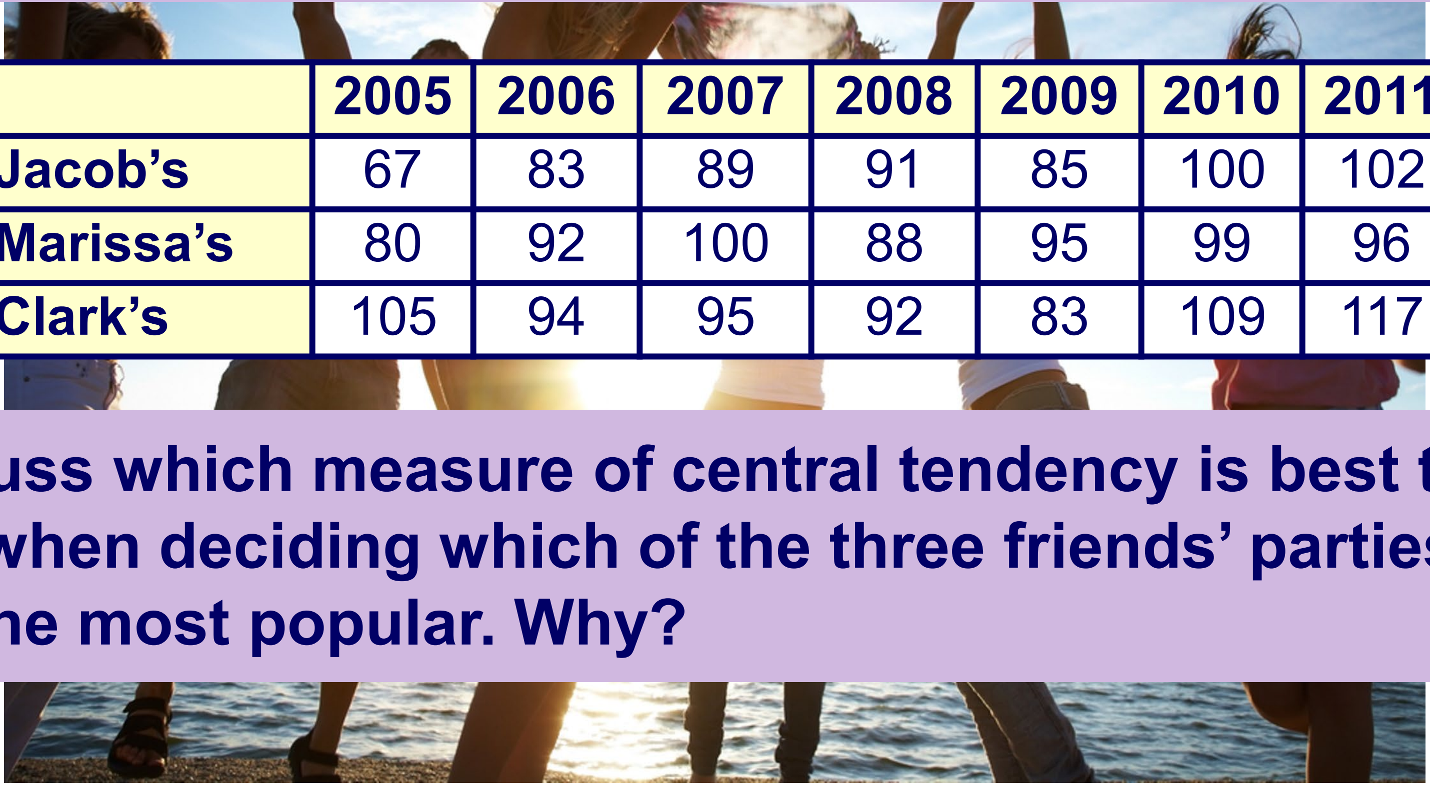
**Work out the mean and median for these data sets, then decide which is the more appropriate average.**

| data set | mean | median | best measure |
|---|---|---|---|
| –16, 4, 5, 5, 5, 5, 5, 6, 6 | 2.77 | 5 | **median** |
| 10, 15, 15, 19, 22, 24, 24, 25 | 19.25 | 20.5 | **mean** |
| 12, 34, 36, 37, 37, 38, 39, 40 | 34.13 | 37 | **median** |
| 72, 97, 99, 101, 112, 130, 134 | 106.43 | 101 | **mean** |
| 275, 277, 278, 279, 291, 305, 315 | 288.57 | 279 | **mean** |
| 22, 26, 29, 29, 30, 36, 38, 40, 99 | 38.78 | 30 | **median** |
| 31, 42, 56, 72, 89, 90, 91, 91, 92 | 72.67 | 89 | **mean** |

# Summer parties

Every year, three friends each compete to throw the best summer party.

**What are the mean and median for these attendance figures for the three parties?**



| | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 |
|---|---|---|---|---|---|---|---|
| **Jacob's** | 67 | 83 | 89 | 91 | 85 | 100 | 102 |
| **Marissa's** | 80 | 92 | 100 | 88 | 95 | 99 | 96 |
| **Clark's** | 105 | 94 | 95 | 92 | 83 | 109 | 117 |

**Discuss which measure of central tendency is best to use when deciding which of the three friends' parties are the most popular. Why?**

# Card-throwing contest

Here are Max's results (in feet) from a card-throwing contest:

159.4   137.2   20.1   139.8   121.5   149.6   114.9   122.3

**Discuss: What is the mean result? Is this a fair representation of Max's ability? Explain.**

**What would be a fair way for the result of this competition to be decided?**

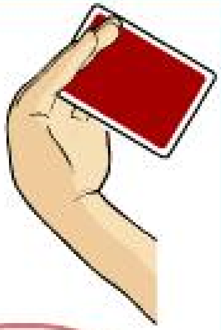
A value that is significantly higher or lower than the other values in a data set is called an **outlier**.

Outliers can increase or reduce the mean dramatically, making it a less accurate measure of the data.

It may be appropriate in research or experiment to remove an outlier before analyzing results.

# Card-throwing contest

## Who is the best at throwing cards?

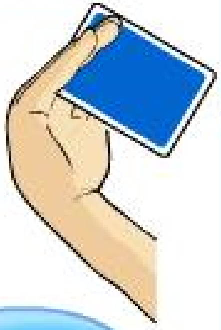


| 1 | |
|---|---|
| | |

Press **Throw** to make each player throw their cards in the card-throwing contest. The distances (in feet) will be recorded in the table. Discuss whether there are any outliers, and how appropriate the mean, median and mode would be to describe each set of data. Which player did the best? Why?

start

Throw



| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|
| | | | | | | | |

Throw

# The average wage?

Mary owns a small company. The wages for each employee in the company are:

| | |
|---|---|
| Mary | $180,000 |
| Her assistant | $40,000 |
| 2 salesmen | $36,000 each |
| Administrator | $37,000 |

**Find the mean, median and mode of the wages of everyone in the company.**

**Which of these would you use to:**

**a) argue for a pay raise for one of Mary's employees**

**b) show that Mary does not pay very well**

**c) represent most of the wages of the workers?**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 80 | 134 | –31 | 267 | 108 | –18 | 295 | 235 | 236 |

Mean:

Median:

Range:

next

# Gymnastics scores

Here is a summary of Chris and Anna's scores (out of 10) in 10 gymnastics contests.

| | Chris | Anna |
|---|---|---|
| **Mean** | 6.7 | 6.4 |
| **Range** | 4 | 5 |

**Discuss which of these conclusions are correct:**

- **Chris is more reliable.**
- **Anna is the best because her mean is lower.**
- **Anna is better because her range is higher.**
- **Chris must have gotten a higher best score.**
- **On average, Chris is better and more consistent.**

# Comparing gymnastics scores

| | Chris | Anna |
|---|---|---|
| Mean | 6.7 | 6.4 |
| Range | 4 | 5 |


Here are Chris and Anna's scores:

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Data A | 7 | 5 | 9 | 6 | 7 | 5 | 5 | 8 | 8 | 7 |
| Data B | 8 | 7 | 4 | 6 | 8 | 7 | 6 | 7 | 8 | 3 |

**Use the summary table above to decide which data set is Chris's and which is Anna's.**

**Chris thinks that the interquartile range still shows him to be the most consistent. Is he correct? Justify your answer.**

# Comparing remote car racing times

| 2010 | 2011 |
|---|---|
| 12.1 | 12.3 |
| 14.0 | 13.7 |
| 15.3 | 15.5 |
| 15.4 | 15.5 |
| 15.4 | 15.6 |
| 15.6 | 15.9 |
| 15.7 | 16.0 |
| 15.7 | 16.1 |
| 16.1 | 16.1 |
| 16.7 | 17.1 |
| 17.0 | 22.9 |

Here are the times (in minutes) from a remote control car race in both 2010 and 2011.

**Find the mean and range for each year.**

| | 2010 | 2011 |
|---|---|---|
| **Mean** | 15.4 | 16.1 |
| **Range** | 4.9 | 10.6 |

**Which year was better and why?**

**Why might this comparison be unfair?**

The **interquartile range** is a better measure of spread when the data contains an outlier.

# Locating the upper and lower quartiles

| 2010 | 2011 |
|---|---|
| 12.1 | 12.3 |
| 14.0 | 13.7 |
| 15.3 | 15.5 |
| 15.4 | 15.5 |
| 15.4 | 15.6 |
| 15.6 | 15.9 |
| 15.7 | 16.0 |
| 15.7 | 16.1 |
| 16.1 | 16.1 |
| 16.7 | 17.1 |
| 17.0 | 22.9 |

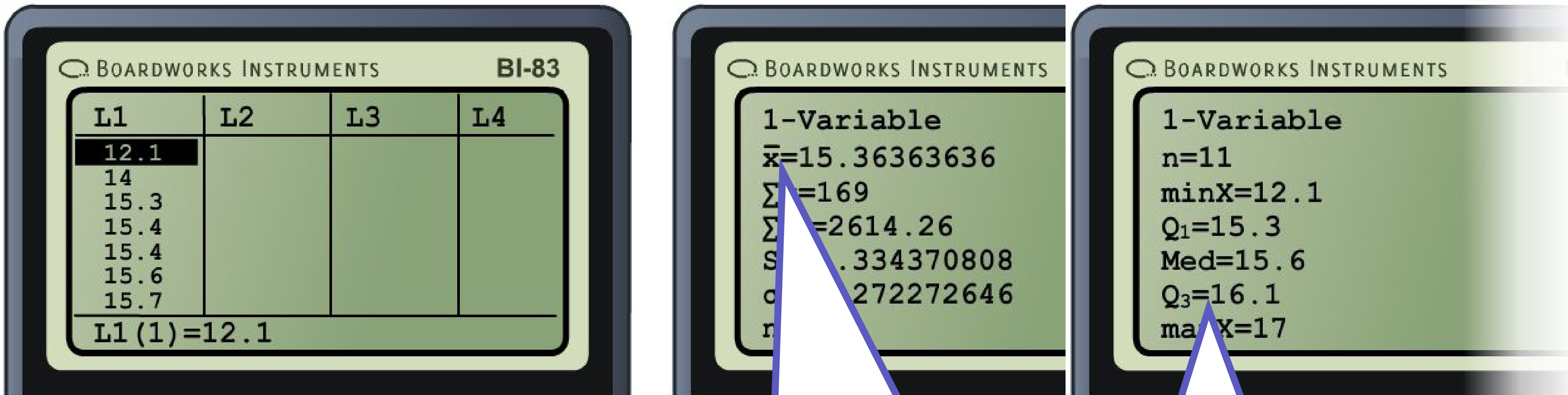
**Compare the 2010 and 2011 remote control car racing results (in minutes) by finding the median and compare the interquartile ranges with the ranges.**

| | 2010 | 2011 |
|---|---|---|
| **Median** | 15.6 | 15.9 |
| **Range** | 4.9 | 10.6 |
| **IQ Range** | 0.8 | 0.6 |

interquartile range for 2010: 16.1 – 15.3 = **0.8 mins**

interquartile range for 2011: 16.1 – 15.5 = **0.6 mins**

Enter the data from 2010 into a graphing calculator using the STAT feature to analyze it.




Choose 'CALC' and then '1-VAR' (one variable) to see the statistics screen.

The mean, $\bar{x}$, is 15.36.

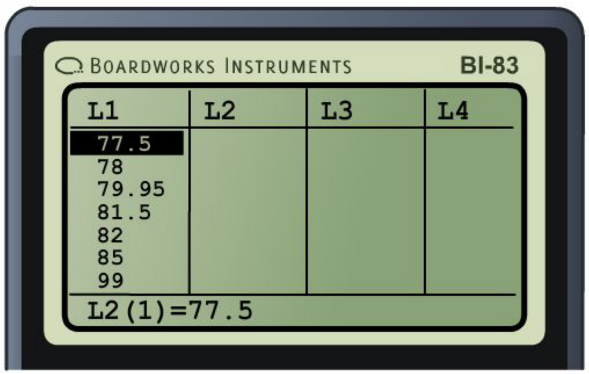
Use the up and down arrows to scroll down and see $Q_1$, the median and $Q_3$.

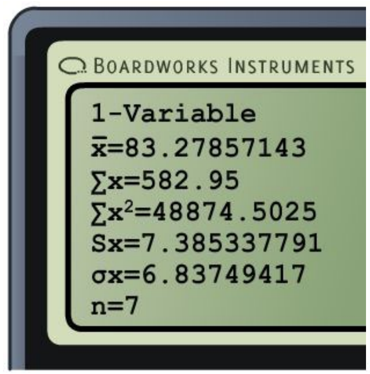
# Store prices

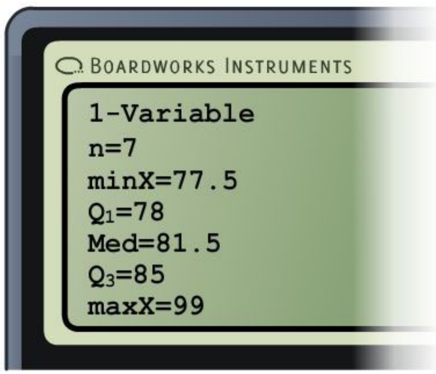
The students in Mrs. Andrew's algebra class surveyed seven department stores within a certain area for prices of graphing calculators. The table below shows the prices.

| Store | Price |
|---|---|
| A | $77.50 |
| B | $78.00 |
| C | $79.95 |
| D | $81.50 |
| E | $82.00 |
| F | $85.00 |
| G | $99.00 |

**Use your graphing calculator to find the mean, the median, $Q_1$, $Q_3$ and the range. What is the interquartile range?**




mean ($\bar{x}$) = $83.28, median = $81.50, $Q_1$ = $78, $Q_3$ = $85 and range = maxX – minX = $21.50.

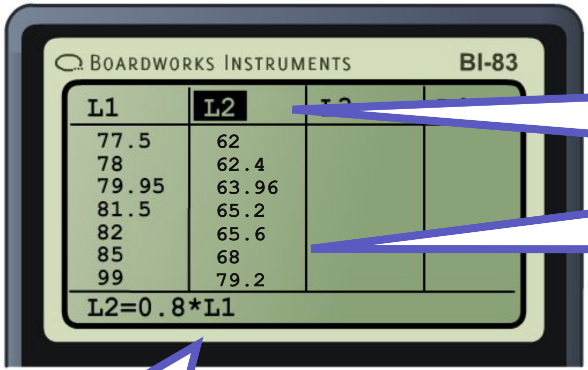
Interquartile range = $Q_3 - Q_1$ = $7.

There is a back-to-school sale in all of the stores and all calculator prices are reduced by 20%. Use your calculator to determine the new mean calculator price.

If the price is reduced by 20% the new price will be 80% of the original price.

To put the new prices in L2, arrow up to the **top** of list 2, then type 0.80*L1.

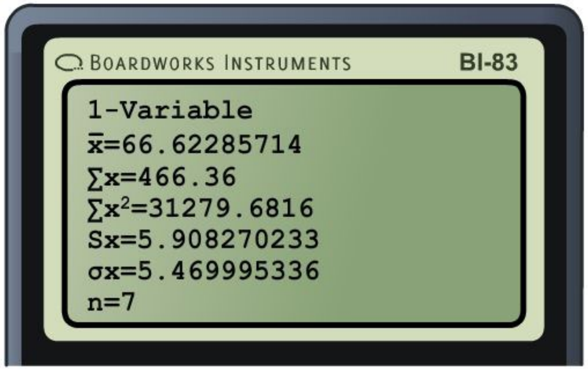
Next, select ‘CALC’ then ‘1-VAR’ to see the statistics screen.




Select ‘L2’ up here

New prices appear in L2

Typing occurs here, then press ‘enter’

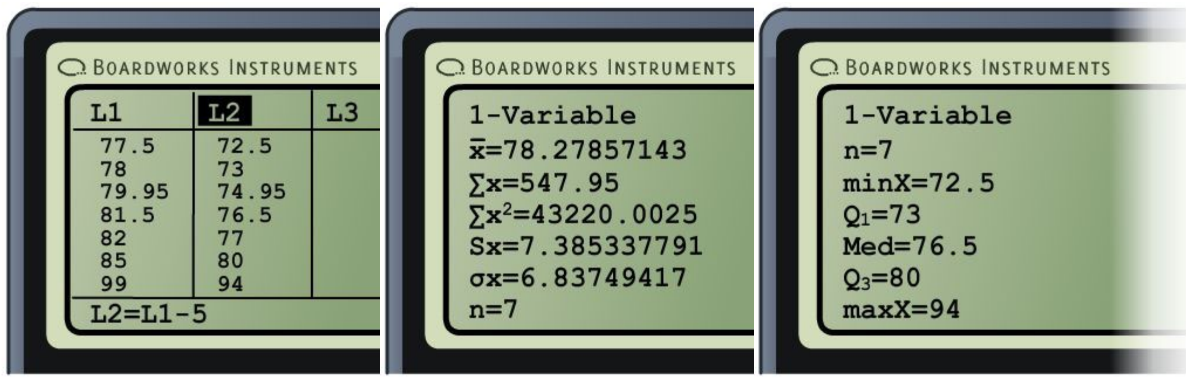



The new mean price is **$66.62**.

# Changing the data

**For the back-to-school sale, the stores all decide to lower the original price of their calculators by $5.00. Use the STAT feature of your graphing calculator to find out how this affects the mean, median and range.**

Create list 2 with the formula L1 – 5, then select 'CALC' and '1-VAR'.




The mean ($78.28) and median ($76.50) are lower.

The range (maxX – minX) = $94 – $72.50 = $21.5 is the same.

Read the following questions about the mean, median, mode and range and discuss them either in groups or together as a class.

Press **start** to begin.

start