Mean, Mode, Median

Distributions

Measures of Central Tendency

Where is the middle?

What is the most representative value?

How can we summarize a data set in one number?

Measures of Central Tendency

Median

- Middle-most value.

Mode

- Most frequently occurring value.
- Highest bar on bar graph.

Mean

- Add up the values and divide by the number

1 Find the median for each set of data.

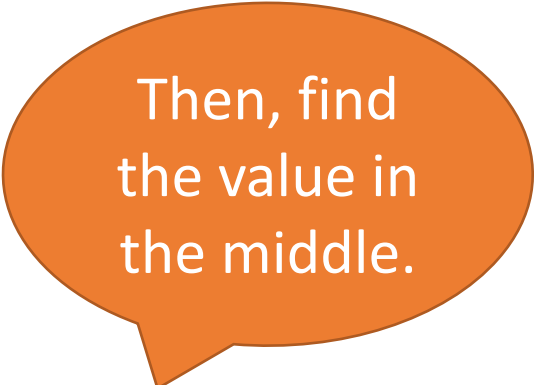A. 49, 32, 67, 55, 58
B. 3.1, 5.2, 4.4, 5.0, 3.8, 2.6, 4.7
C. 29, 12, 30, 22, 7, 23, 36, 15, 18, 9
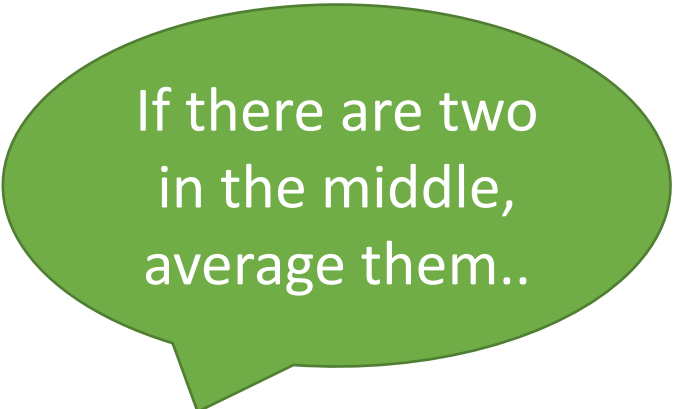D. 81.6, 83.7, 78.5, 82.8, 81.2, 76.3, 83.5, 78.9
E. 110, 115, 109, 110, 116, 113, 112, 116, 110, 106, 113

First, sort.

Then, find the value in the middle.

If there are two in the middle, average them..

(2) The weekly salaries for 5 people who work at a Las Vegas hotel are given in the table.

A. What is the mean salary?
B. What is the median salary?

Star of stage show	$45,000
Hotel manager	2,000
Chief chef	1,500
Publicity director	1,100
Lifeguard	400

For mean, add up the values.

Then, divide by the count.

3 Find the mode (or modes) for each set of data.

A.

Suit Sizes		
36	39	40
37	39	41
37	39	41
38	40	41
38	40	42
38	40	44
39	40	44

B.

Times in 50-m Dash (s)			
5.7	6.3	6.7	6.9
5.9	6.3	6.7	7.0
6.0	6.4	6.8	7.2
6.0	6.5	6.9	7.3
6.2	6.5	6.9	7.3
6.3	6.7	6.9	7.5

Measures of Central Tendency

Median

- Need to sort data, that's a little hard to do.
- Not influenced by outliers.

Mode

- Sometimes not possible in the data set
- Useful for a small range or in bar graphs.
- Can be more than one
- Eg. Clothing sizes

Mean

- Easy to calculate
- Works for all data
- Influenced by outliers.

Example 6 Measuring Central Tendency

Describe the central tendency of each of the following monthly incomes for six salespeople working on commission using the most appropriate measure.

(a) January: $1241, $1499, $2020, $1371, $1622, $1853

Solution

(a) **Mean** $\dfrac{\$1241 + \$1499 + \$2020 + \$1371 + \$1622 + \$1853}{6} = \$1601$

Median $\dfrac{\$1499 + \$1622}{2} = \$1560.50$

Mode None as no repeated data

The values in this set are evenly distributed. Both the mean and median provide a good measure of the central tendency.

Complete each table. Write the letter for each frequency in the box above the corresponding value at the bottom of the page. Make a histogram for each set of data.

Test Scores for 40 Students

80	89	66	73	60	97	53	79	70	58
71	99	44	88	80	69	72	83	86	76
91	77	61	83	57	88	49	77	75	95
64	85	75	72	94	66	84	77	86	82

Score	Tally	Frequency
41-50		U
51-60		T
61-70		E
71-80		A
81-90		W
91-100		G
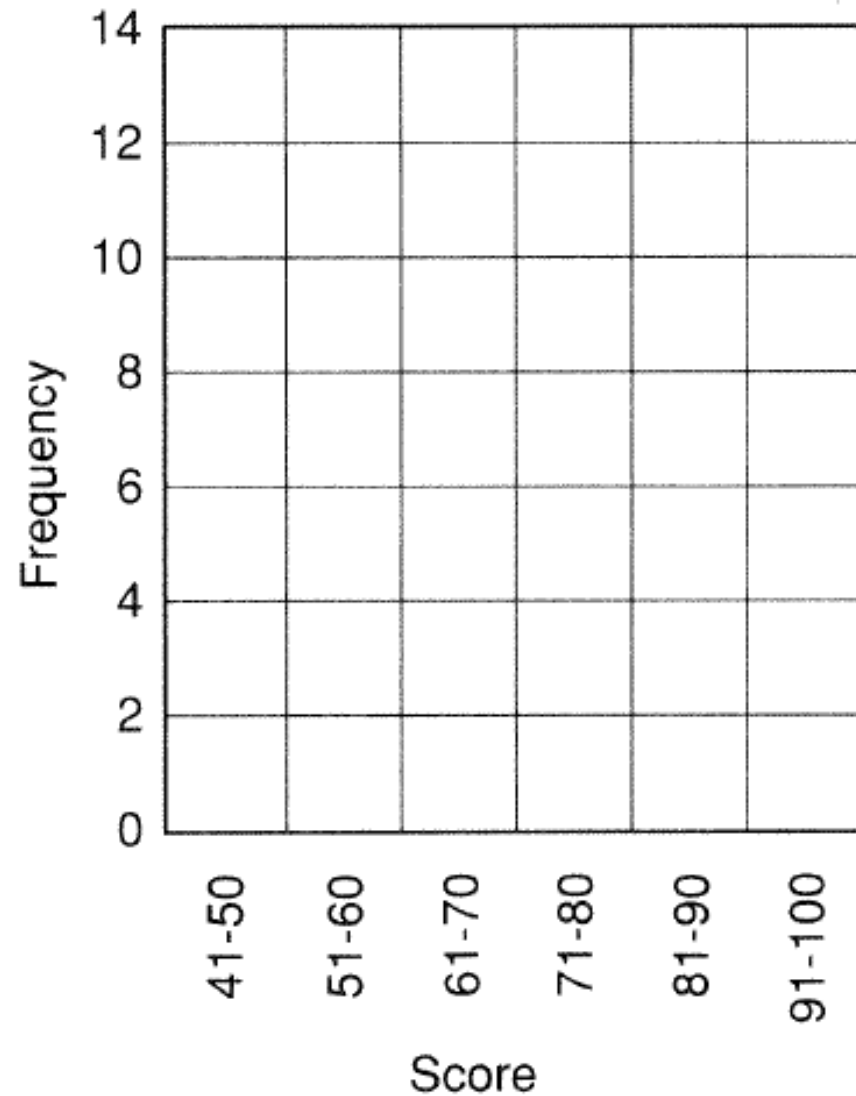
U-SHAPED DISTRIBUTION

A U-shaped distribution occurs when there are peaks at either end of the range. It may also be described as a **bimodal** distribution. The scores from the game of spider solitaire form this type of distribution.

Spider Solitaire Scores

Spider Solitaire

Score	400–449	450–499	500–549	550–599	600–649	650–699	700–749	750–799	800–849
Frequency	15	10	6	4	0	4	6	10	15
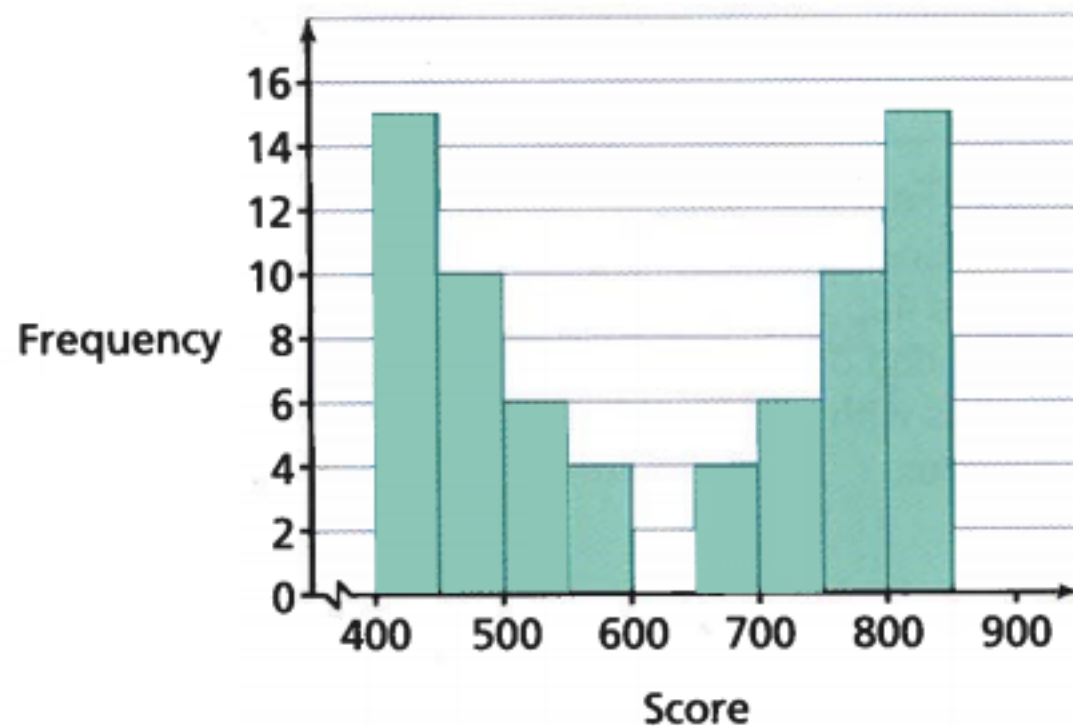
UNIFORM DISTRIBUTION

When each outcome has a similar frequency, it is called a uniform distribution. The height of each bar is roughly equal. This is the distribution you would expect from an experiment such as rolling a die.

The Die-Rolling Game

Score	1	2	3	4	5	6
Frequency	18	17	18	18	16	19

The Die-Rolling Game
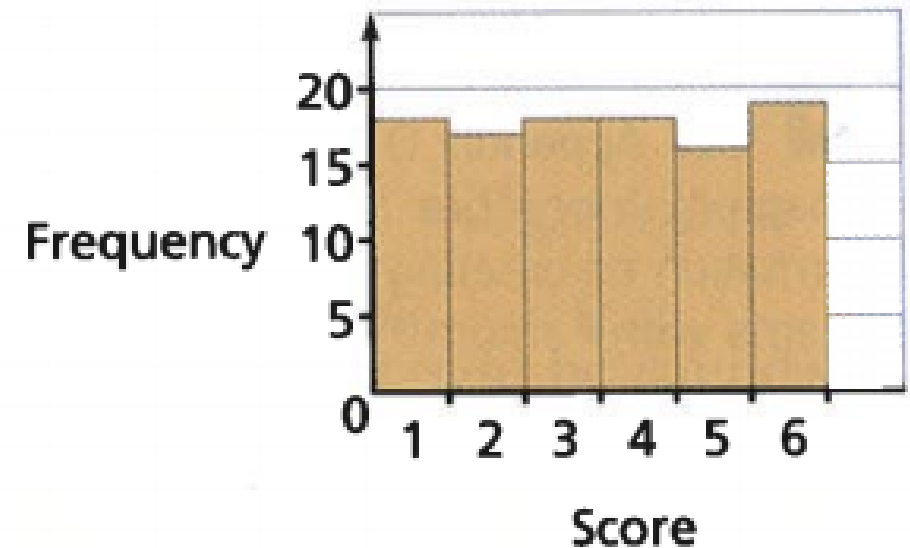
MOUND-SHAPED DISTRIBUTIONS

In this distribution, there is an interval with the greatest frequency, and the frequencies of all other intervals decrease on either side of that. The frequency distribution then takes on a mound shape. Rolling a pair of dice and recording the sum results in this type of distribution.

Dice Sums

Frequency vs Sum histogram

Dice Sums

Sum	2	3	4	5	6	7	8	9	10	11	12
Frequency	1	2	3	4	5	8	6	3	2	1	0

Each of there three distributions is **symmetric**.
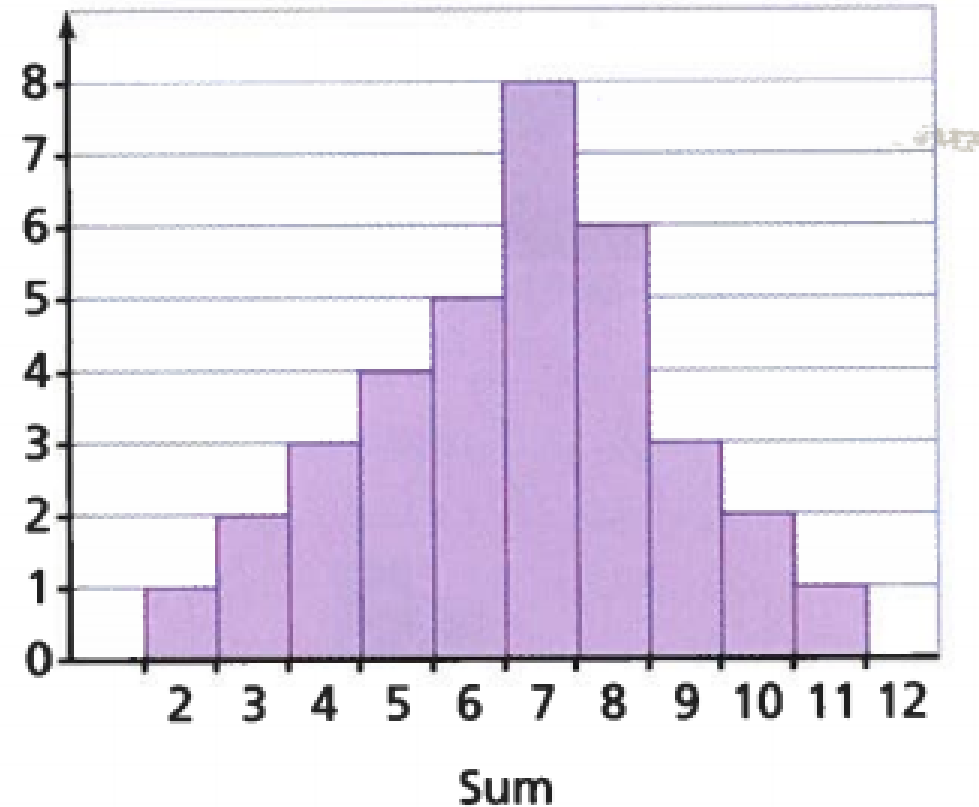
SKEWED DISTRIBUTIONS

In a skewed distribution, the interval or group of intervals that contains the greatest frequencies is near one end of the histogram. As a result, these distributions seem to tail off to the left or right. The direction of the skew is determined by the direction the mean has shifted. This will be discussed in detail in Section 3.2.) Scores from a game of solitaire produce a distribution that is skewed to the right (right-skewed).

Solitaire

Score	0–19	20–39	40–59	60–79	80–99	100–119	120–139	140–159
Frequency	20	15	12	7	5	2	1	1

Solitaire Scores
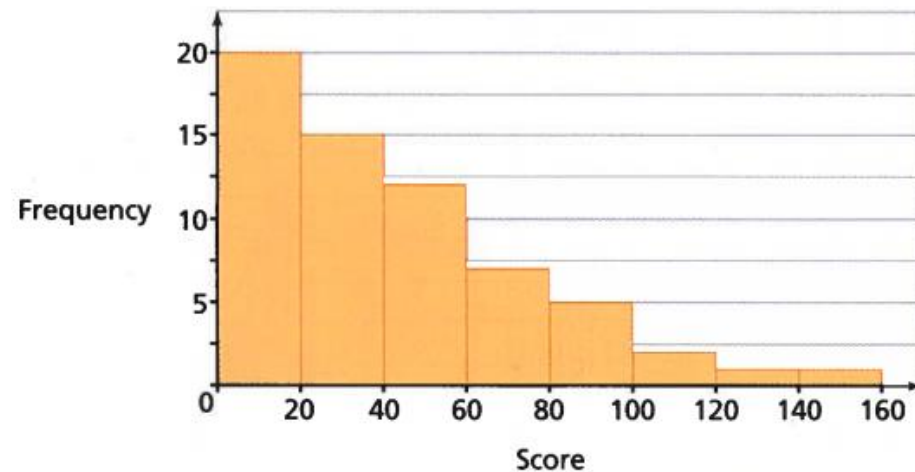
Example 5 Exploring Distributions and Central Tendency

Compare the following data sets. What is the relationship between the shape of the distribution and the mean, median, and mode?

(a)
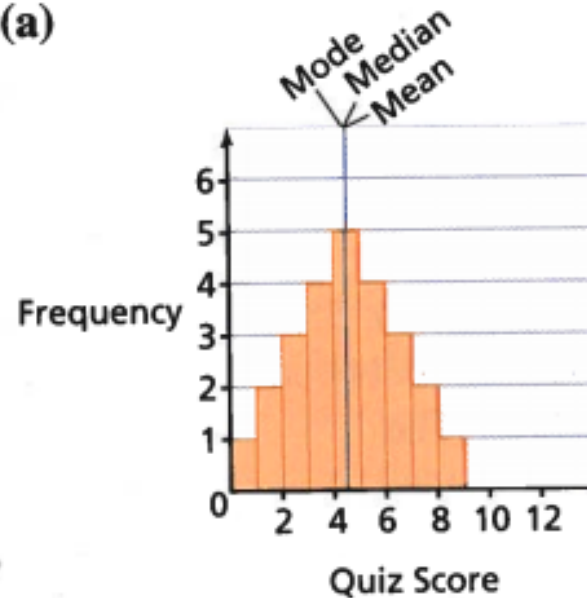
(b)
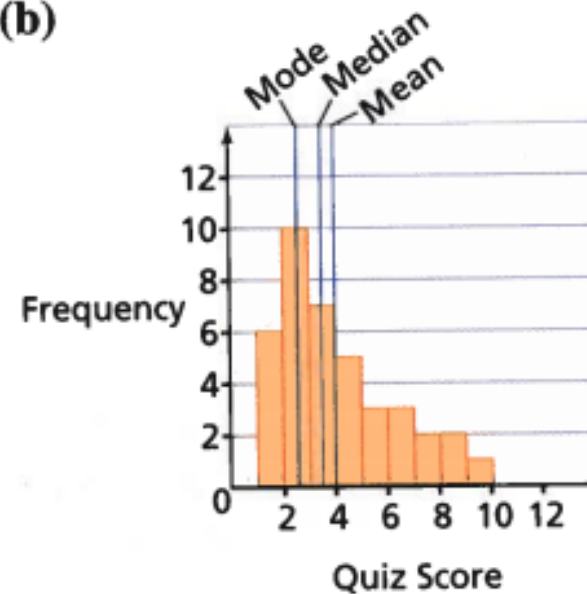
Solution

In part (a), the distribution is mound-shaped and symmetric. The mean, median, and mode are all equal. In part (b), the distribution is skewed right. Notice that mode < median < mean. When the distribution is skewed left, you would find that mean < median < mode. The outliers affect the mean more than they do the median and mode. In addition, notice that the median is always between the mean and mode for non-symmetric distributions.

DIFFERENT DISTRIBUTIONS

For distributions that are **symmetric** about the centre, the mean and median will be approximately equal.

mean and median
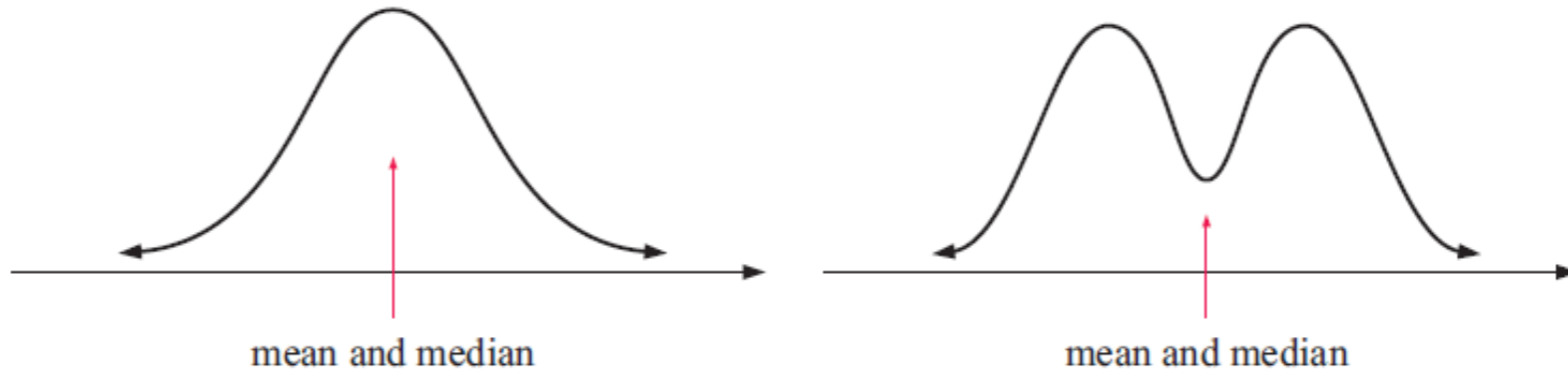
mean and median

If the data set has symmetry, both the mean and the median should accurately measure the centre of the distribution.

If the data set is not symmetric, it may be positively or negatively skewed:

positively skewed distribution

negatively skewed distribution

mode

median mean

mode

mean median
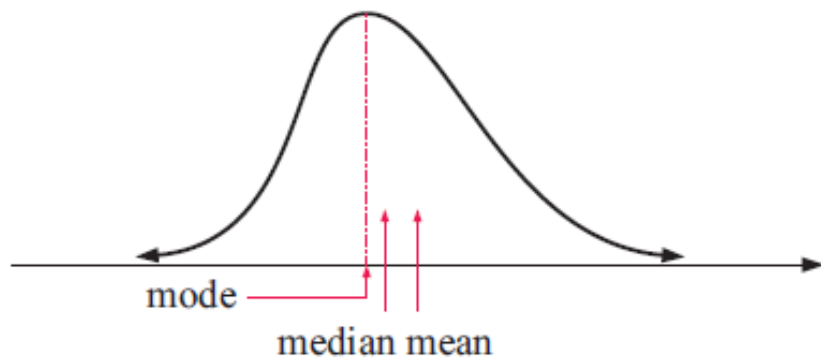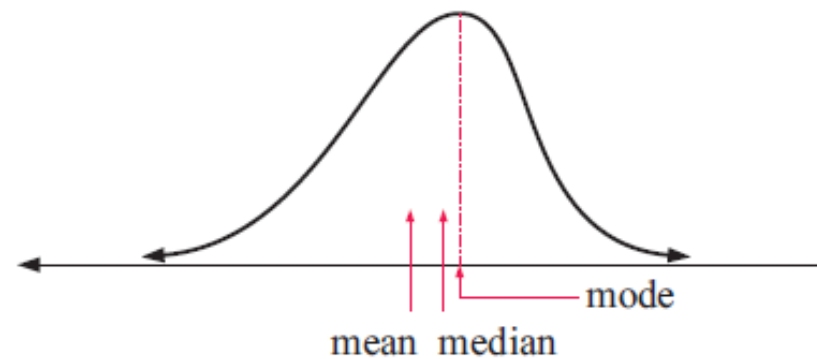
Notice that the mean and median are clearly different for these skewed distributions.

Mean lies between the two modes
Mode Mode
A
Bimodal Distribution

Mean is approximately equal to the median
B
Flat Distribution

Mean lies to the right of the median
C
Right-Skewed Distribution

Mean lies to the left of the median
D
Left-Skewed Distribution
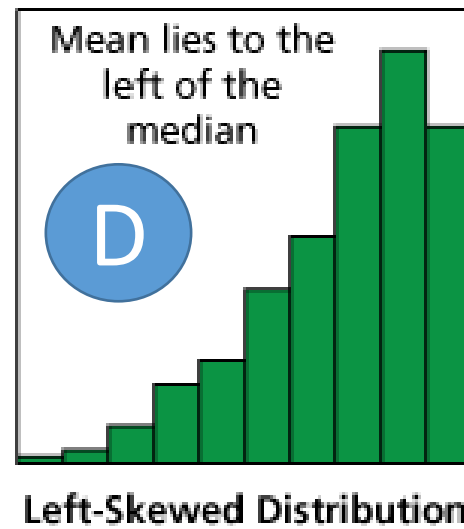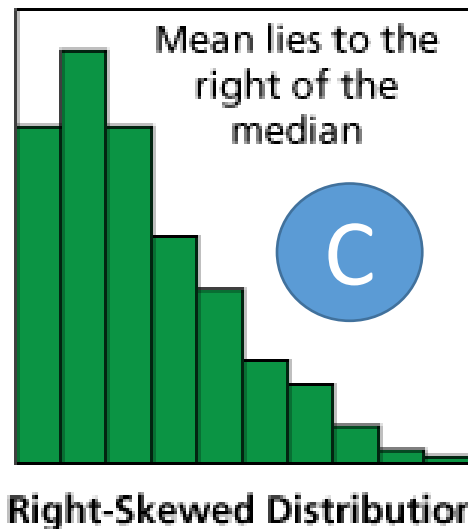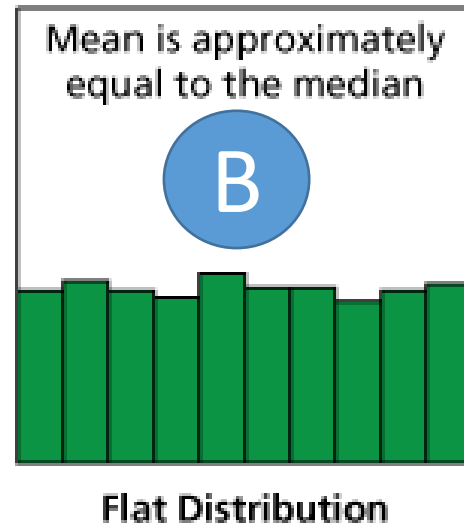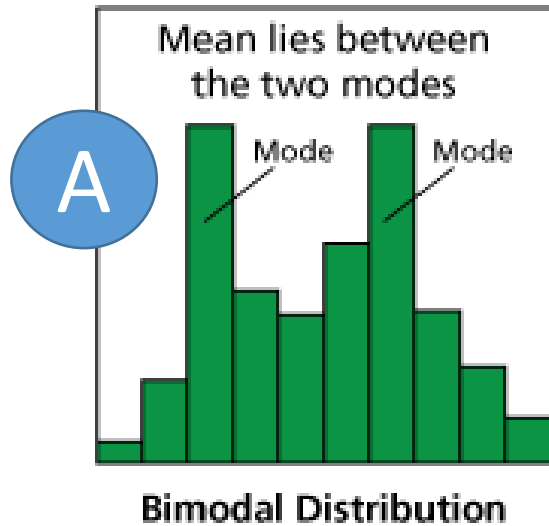
Mark Calculation:
When you calculate marks, you want to find the grade that is most representative of the student.

If marks are required to be calculated by average, which students benefit and which do not?

If marks are required to be calculated by mode, which students benefit and which do not?

Which graph matches each description?
(i) mound shaped (ii) bell-curve   (iii) left skew
(iv) bi-modal   (v) right skew   (vi) u-shaped
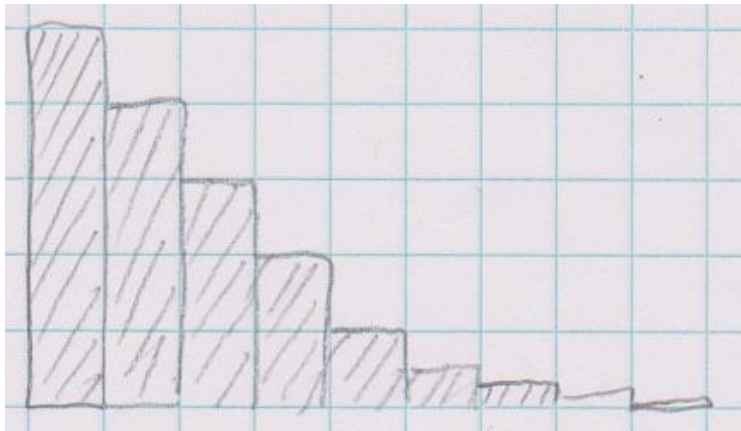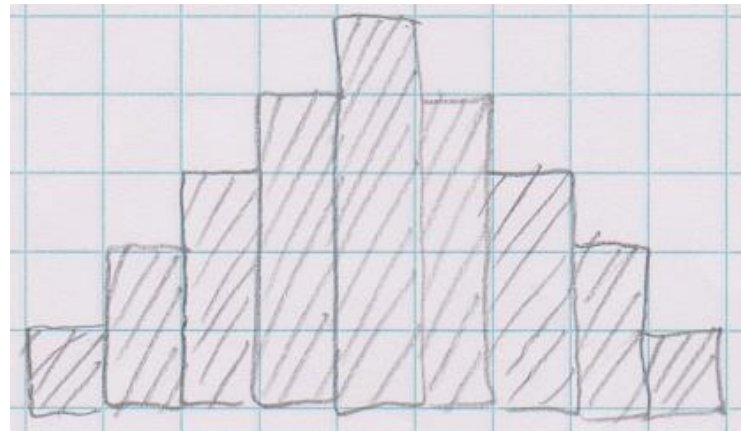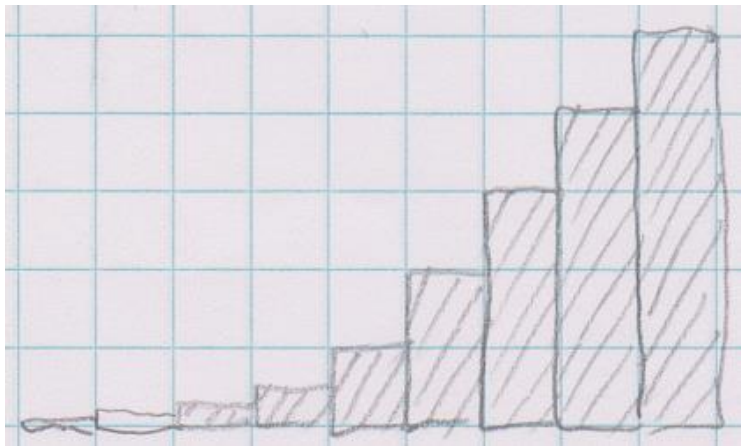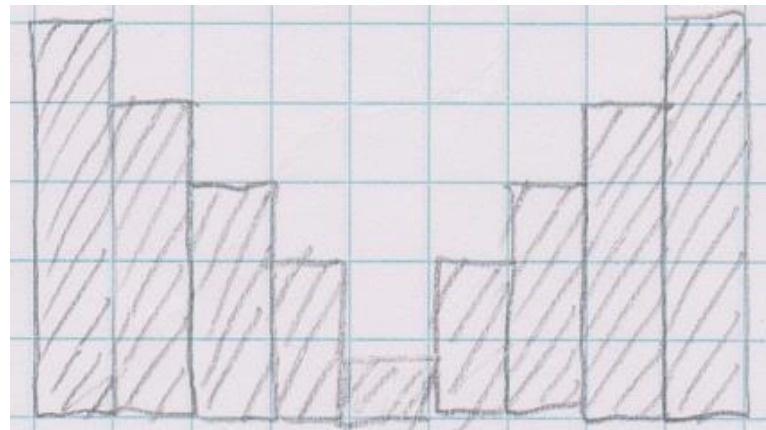
Which graph could have:
A mean of 16, a median of 16 and modes of 2 and 22?
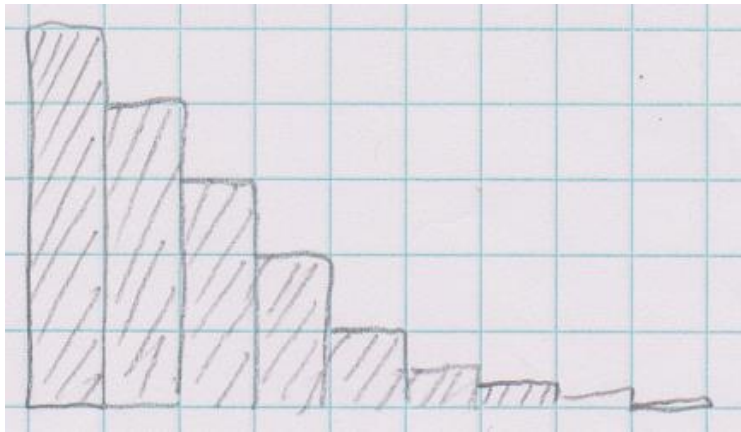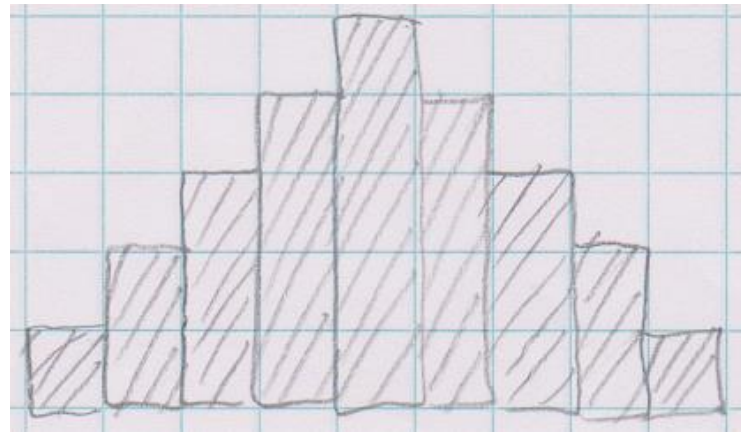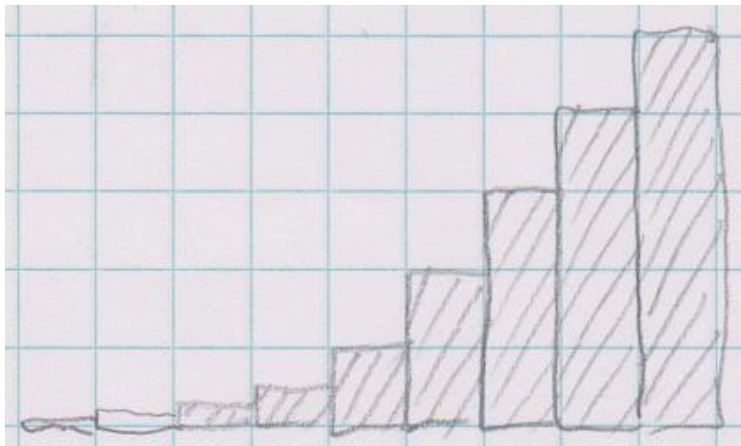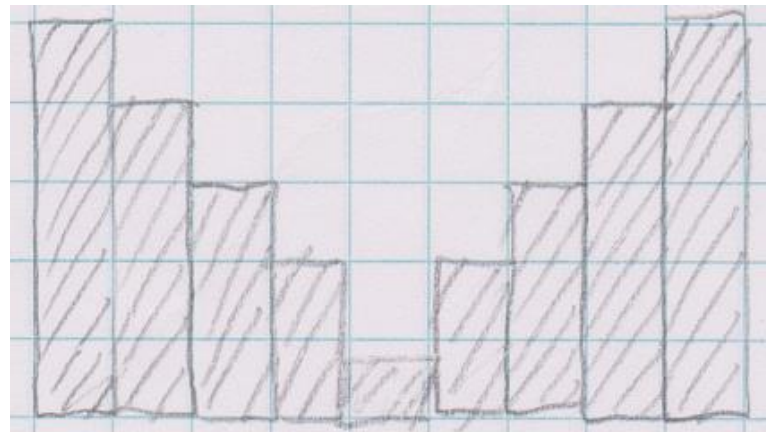
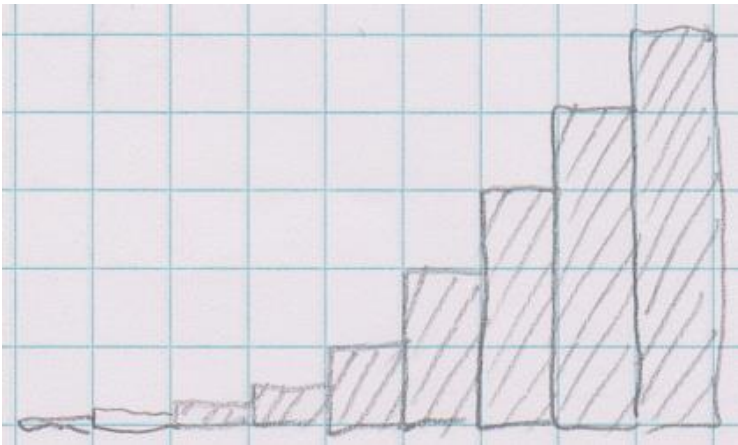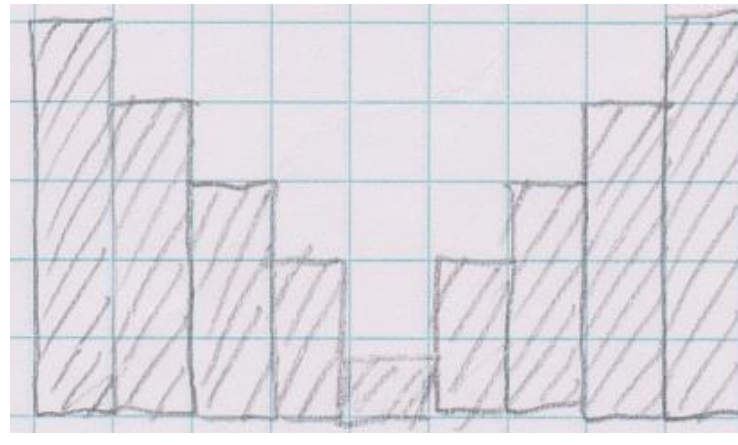# Which graph could have:
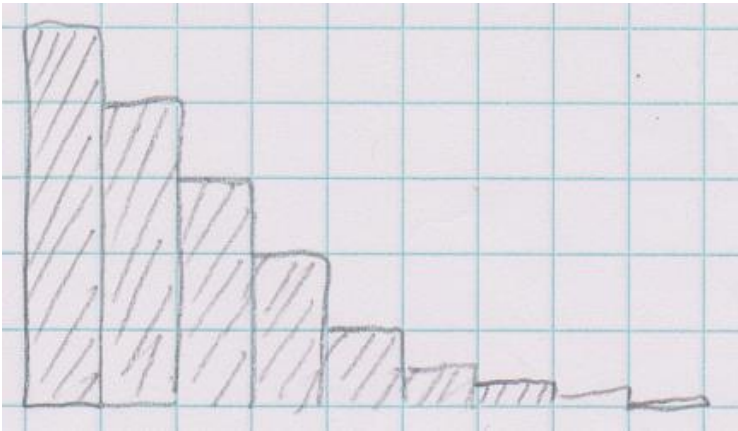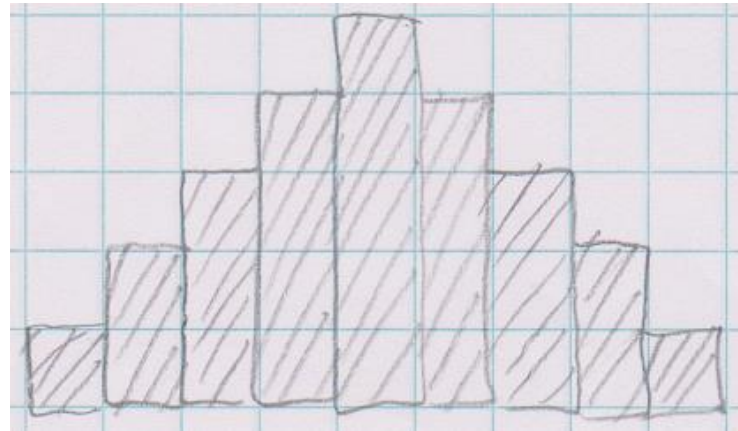## A mode of 4, median of 8 and a mean of 12?
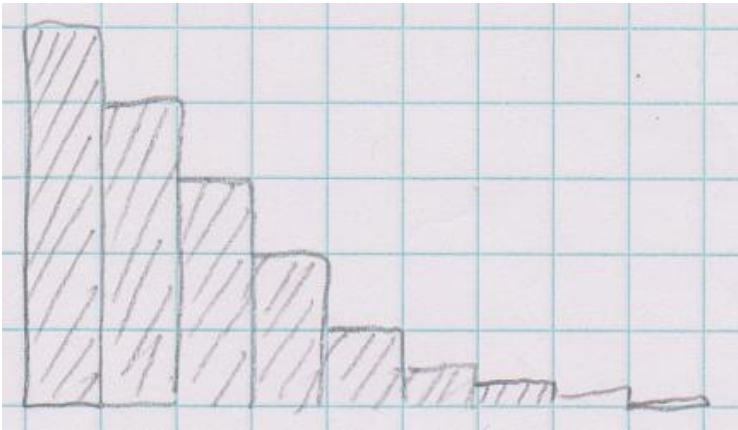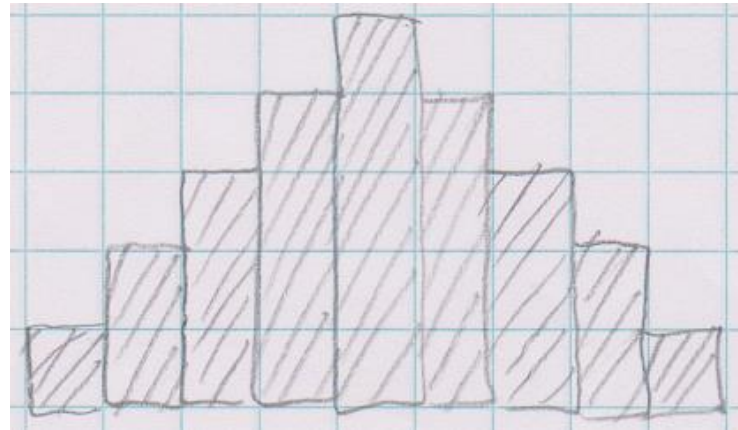
Which graph could have:
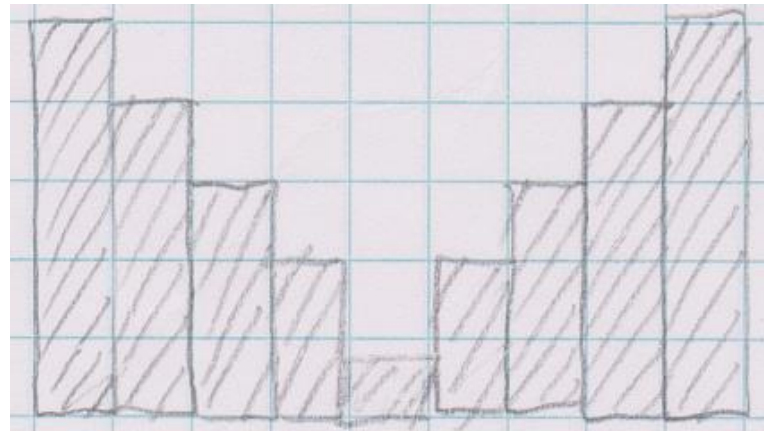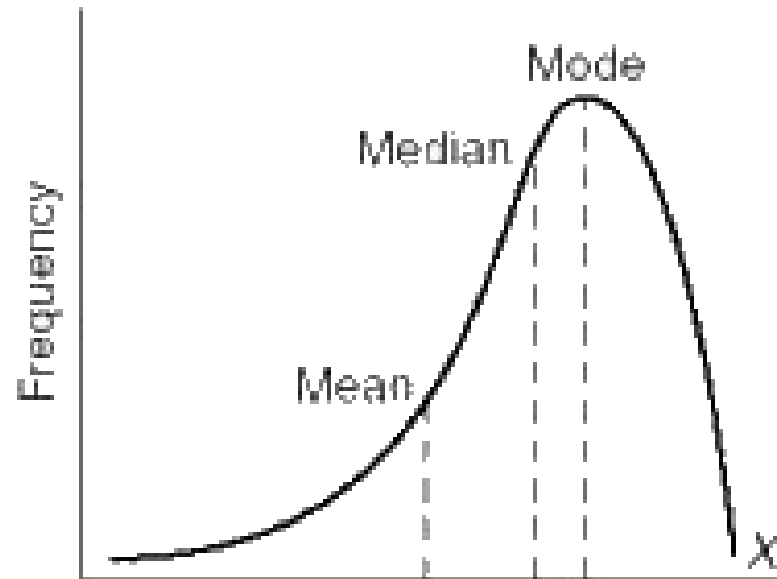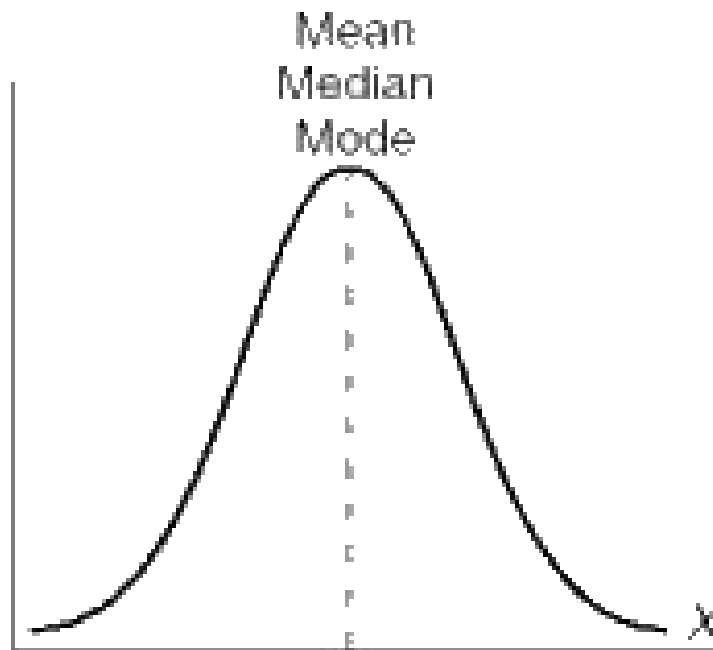## A mode of 34, median of 34 and a mean of 34?

(a) Negatively skewed **(b) Normal (no skew)** **(c) Positively skewed**
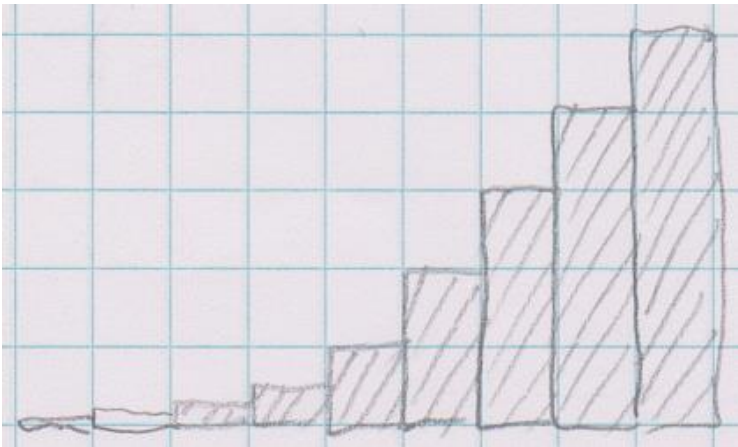
Mode
Median
Mean
Frequency
Negative Direction

Mean
Median
Mode
Perfectly Symmetrical
Distribution
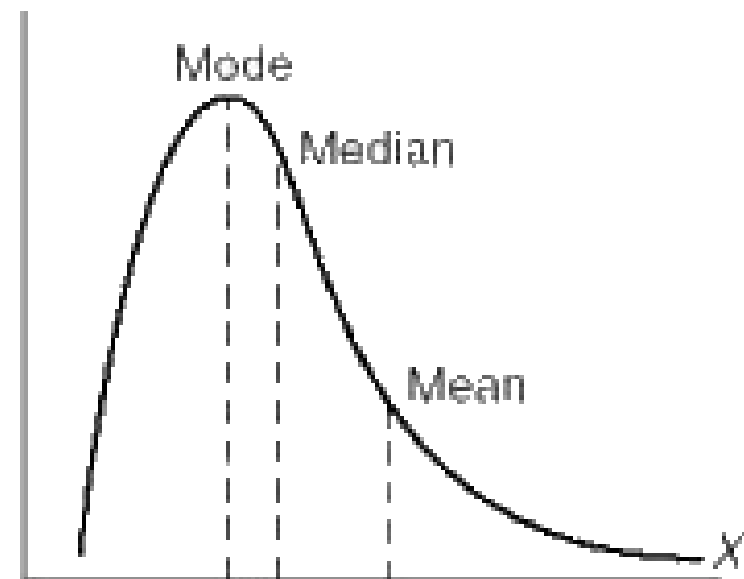
Mode
Median
Mean
Positive Direction

Histograms for IQ Test Components