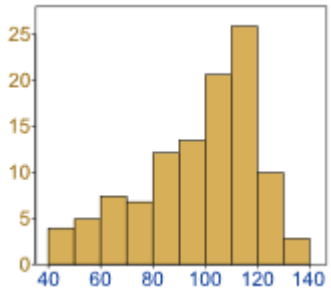
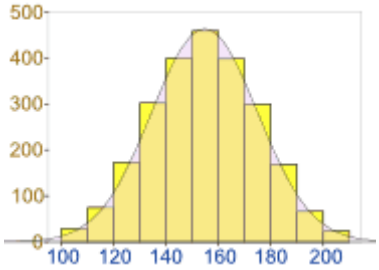
SKEWNESS

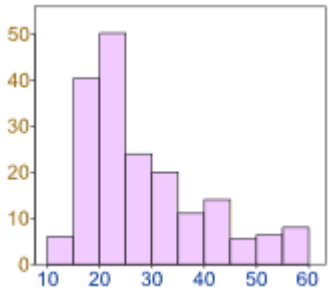
Data can be "skewed", meaning it tends to have a **long tail** on one side or the other:



Negative Skew



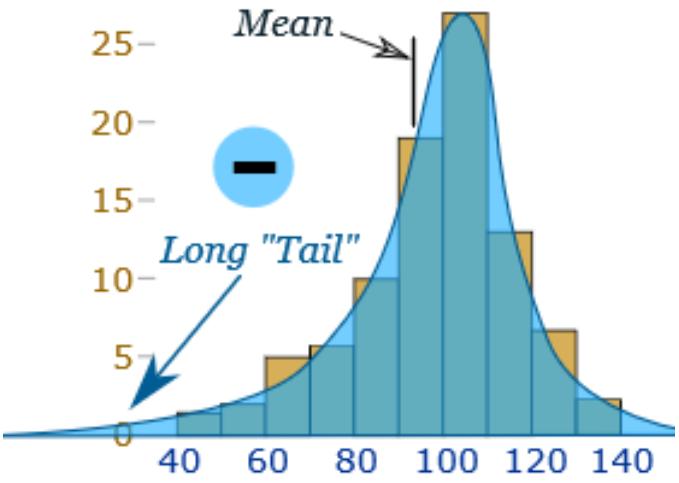
No Skew



Positive Skew

Negative Skew?

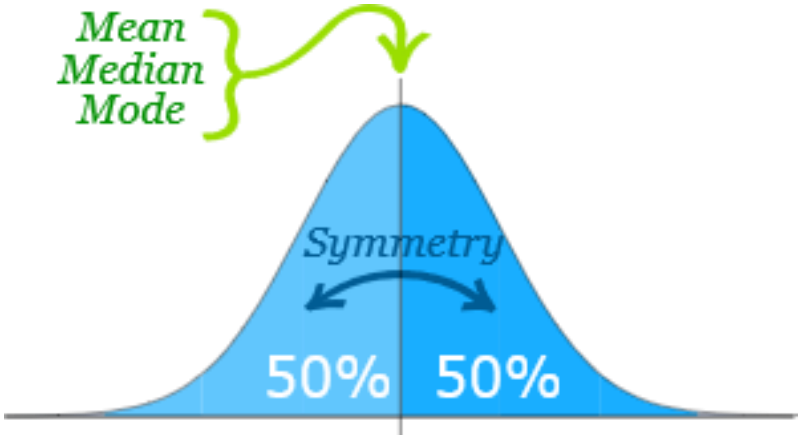
Why is it called **negative** skew? Because the long "tail" is on the negative side of the peak.



People sometimes say it is "skewed to the left" (the long tail is on the left hand side)

The mean is also on the left of the peak.

The Normal Distribution has No Skew

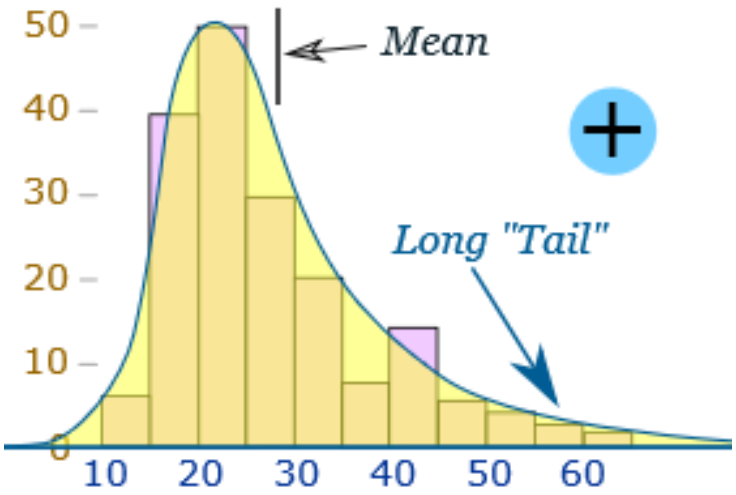


It is perfectly symmetrical.

And the Mean is exactly at the peak.

Positive Skew

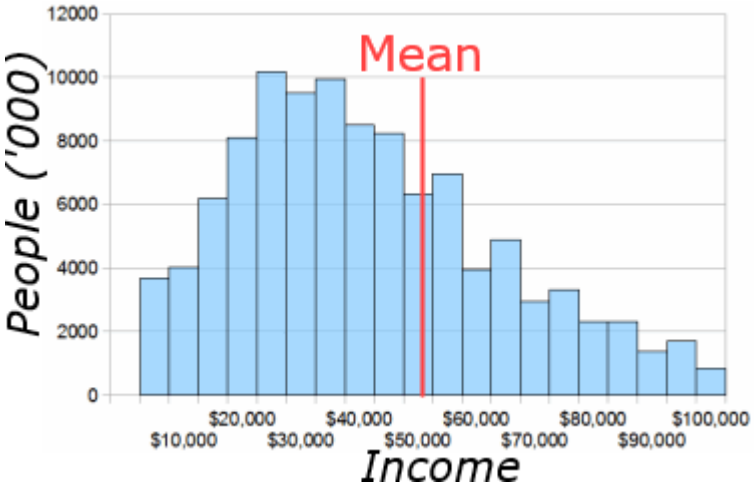
And **positive** skew is when the long tail is on the positive side of the peak, and some people say it is "skewed to the right".



The mean is on the right of the peak value.

Example: Income Distribution

Here is some data extracted from a recent Census.



As you can see it is **positively skewed** ... in fact the tail continues way past $100,000

Question 1

The Cyhelsky skewness coefficient is defined by:

$$\frac{(\text{Number of observations below the mean} - \text{Number of observations above the mean})}{\text{Total number of observations}}$$

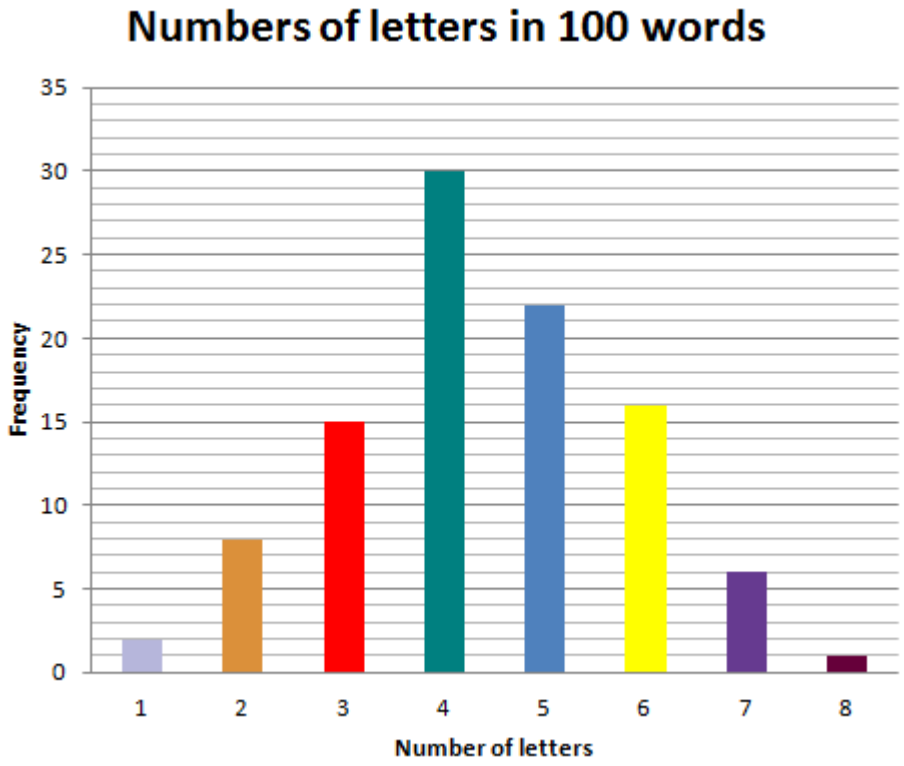
Calculate the Cyhelsky skewness coefficient for the set of numbers:

$$11, 14, 17, 18, 27, 27, 29, 31, 38, 39$$

A -0.4 B -0.2 C 0.2

Question 2

Olivia chose a 100 word passage and recorded the number of letters in each word. Her results are shown in the following bar graph:



The Cyhelsky skewness coefficient is defined by:

$$\frac{(\text{Number of observations below the mean} - \text{Number of observations above the mean})}{\text{Total number of observations}}$$

Calculate the Cyhelsky skewness coefficient for Olivia's data.

A -0.1 B 0.1 C 0.54

Question 3

The Pearson first skewness coefficient is defined by $\frac{\text{Mean} - \text{Mode}}{\text{Standard deviation}}$

Calculate the Pearson first skewness coefficient for the set of numbers:

$$2, 5, 7, 7, 11, 12, 14, 15, 17, 20$$

A -0.09 B 0.56 C 0.74 D 0.93

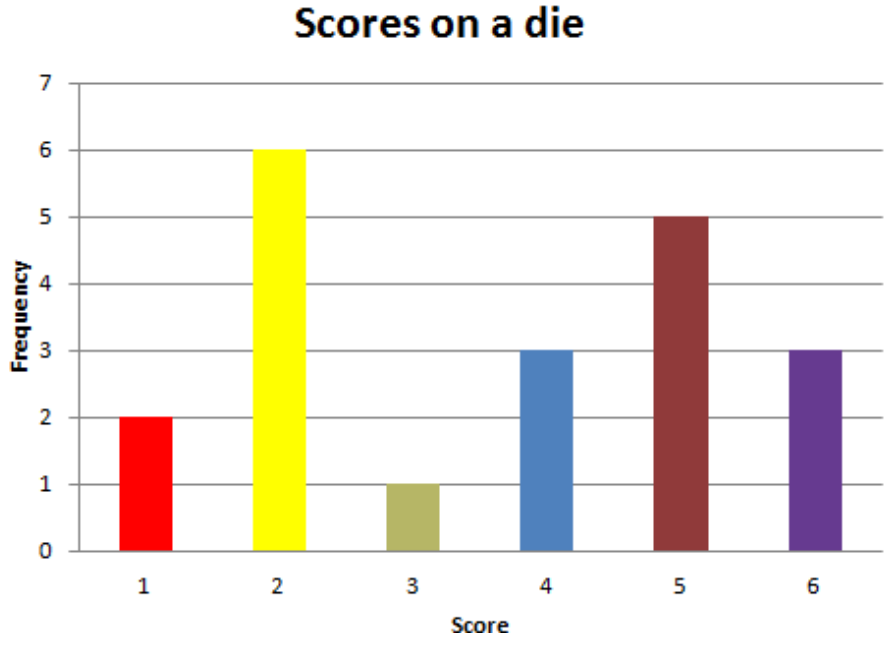
Question 4

Calculate the Pearson first skewness coefficient for the set of numbers:

180, 176, 154, 185, 169, 185, 166, 173, 129, 168

A -1.04 B -0.97 C -0.88 D 1.04

Question 5

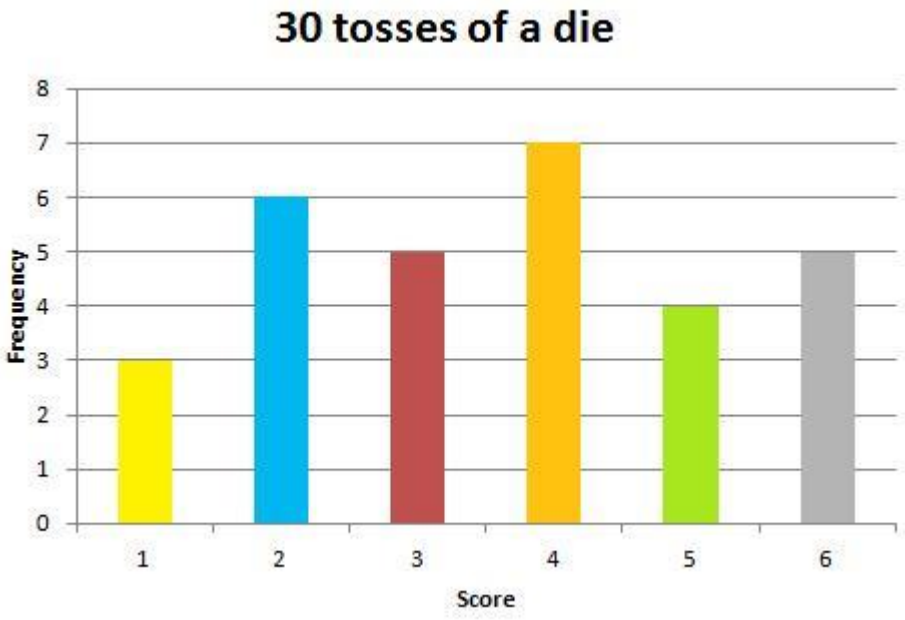


The bar graph shows the scores obtained from 20 throws of a die.
Calculate the Pearson first skewness coefficient for the data in the bar graph.

A -1.42 B -0.95 C 0.95 D 1.42

Question 6

Owen tossed a die 30 times, and drew a bar graph of his results:



Calculate the Pearson first skewness coefficient for the data in the bar graph.

A -2.15 B -0.25 C 0.25 D 2.15

Question 7

The Pearson second skewness coefficient is defined by $\dfrac{3(\text{Mean} - \text{Median})}{\text{Standard deviation}}$

Calculate the Pearson second skewness coefficient for the set of numbers:

$$3, 6, 9, 11, 15, 26, 27, 30, 30, 31, 32$$

A -1.68 B -0.05 C 0.05 D 1.68

Question 8

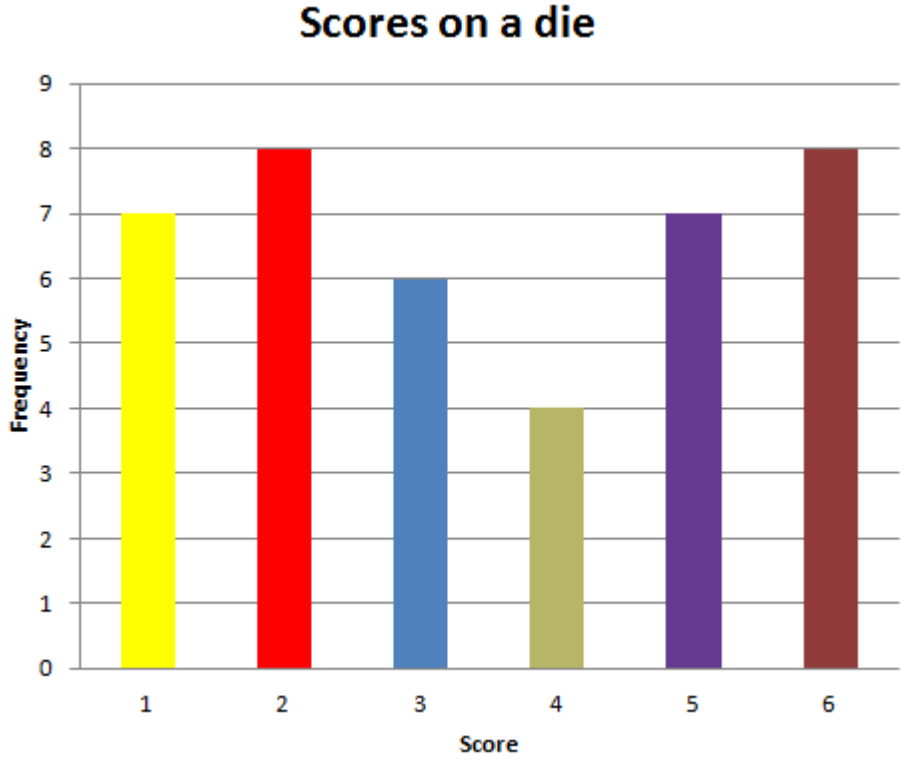
Calculate the Pearson second skewness coefficient for the set of numbers:

$$87,\ 54,\ 39,\ 62,\ 27,\ 76,\ 13,\ 59,\ 28,\ 55$$

A 0.62 B 0.21 C -0.21 D -0.62

Question 9

Emma rolled a die a number of times and recorded her results in a bar graph, as follows:

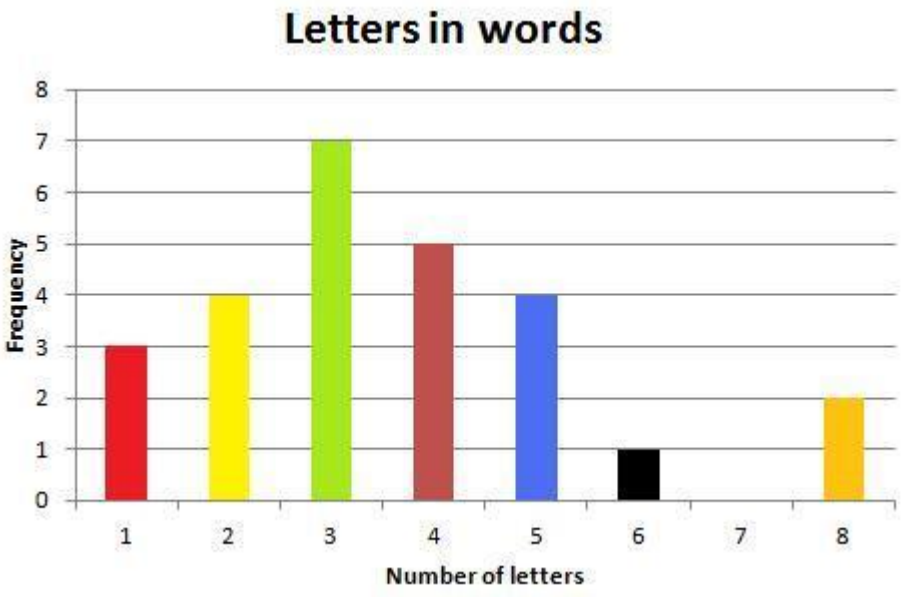


Calculate the Pearson second skewness coefficient for Emma's data.

A -2.50 B -0.83 C 0.83 D 2.50

Question 10

Michael recorded the number of letters in each word of a sentence, and drew a bar graph of his results:



The Pearson second skewness coefficient is defined by $\dfrac{3(\text{Mean} - \text{Median})}{\text{Standard deviation}}$

Calculate the Pearson second skewness coefficient for Michael's data.

A -1.01 B -0.63 C 0.19 D 1.01