8.5 Cumulative Frequency

To calculate the cumulative frequency add up the frequencies of the data values as you go along.

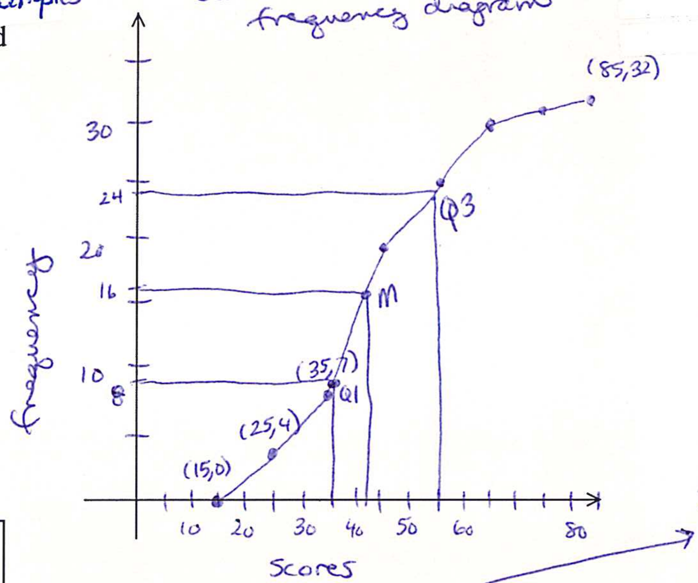
→ (ō – jive)

A cumulative frequency diagram (Cumulative frequency graph) or **ogive** is most useful when trying to calculate the median, quartiles and percentiles of a large set of grouped or continuous data.

Using the NFL data from the previous examples Draw a cumulative frequency diagram and find median and interquartile range.

Scores (s)	f	*Cumulative Frequency*
$15 \leq s < 25$	4	4
$25 \leq s < 35$	3	7
$35 \leq s < 45$	13	20
$45 \leq s < 55$	6	26
$55 \leq s < 65$	3	29
$65 \leq s < 75$	2	31
$75 \leq s < 85$	1	32

Cumulative frequency diagram



Cumulative frequency

median

$\frac{32}{2} = 16$

Q1

$\frac{1}{4}(32) = 8$

Q3

$\frac{3}{4}(32) = 24$

Make sure to label your graph properly.

Find Q1, mid, Q3 and draw lines coming from its location on the frequency line to the graph + then drop it down to the x-axis to approximate the values of median & IQR

$\boxed{\text{median} = 42.5 \text{ pts} \quad IQR = 55 - 35 = 20 \text{pts}}$

Exercise 8G

8.6 Variance and Standard Deviation

The range and interquartile range are good measures of spread but each one is calculated from only two data values.

The variance combines all the values in a data set to produce a measure of spread.

- It is the arithmetic mean of the squared difference between each value and the mean value.

If you want to know why there are advantages to squaring the above difference read page 276 of your book.

* not really used in application

Because the differences are squared, the units of variance are not the same as the units of the data.

The standard deviation is the square root of the variance and has the same units as the data.

The formulae for the variance and standard deviation are:

$$\sigma^2 = \text{Population Variance} = \frac{\sum_{i=1}^{n}(x-\mu)^2}{n}$$

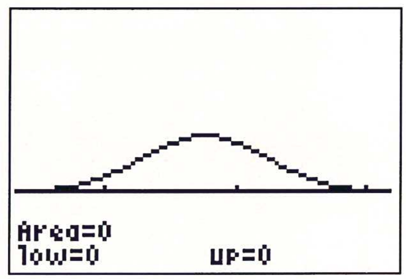
$$\sigma = \text{Population Standard Deviation} = \sqrt{\frac{\sum_{i=1}^{n}(x-\mu)^2}{n}}$$

Example: A bag of M&Ms is supposed to weight 1.69 oz. Here is a list of 10 M&M bags: ____________

Activity

Either give students 10 weights or do the example in the other file. Find the mean and standard deviation.

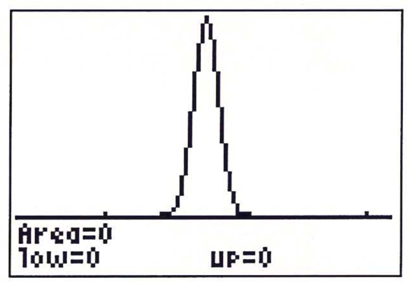
* On the GDC, use the value σx for standard deviation not sx. *



The standard deviation shows how much variation there is from the mean and gives an idea of the shape of the distribution.

- Low standard deviations (bottom picture) shows the data points tend to be very close to the mean.

- High standard deviation (top picture) indicates that the data is spread out over a large range of values.



Properties of standard deviation

- Standard deviation is only used to measure spread or dispersion around the mean of a data set.
- Standard deviation is never negative
- Standard deviation is sensitive to outliers. A single outlier can increase the stand deviation and in turn, distort the representation of spread.
- For data with approximately the same mean, the greater the spread, the greater the standard deviation.
- If all values of a data set are the same, the standard deviation is zero because each value is equal to the mean.

Hmwk 8H