

More Descriptive Statistics!

See page 471 #4

Number of Children	Frequencies
1	5
2	28
3	15
4	8
5	2
6	1
Total	59

Mean =

Mode =

Median =

TI stuff:

1-Var Stats $\bar{x}=2.610169492$ $\sum x=154$ $\sum x^2=466$ $s_x=1.050729618$ $\sigma_x=1.041787075$ $\downarrow n=59$ ■	1-Var Stats $\uparrow n=59$ $\min X=1$ $Q_1=2$ $\text{Med}=2$ $Q_3=3$ $\max X=6$
---	--



Australian family [above] from Google Images

Average Australian family has 2.2 children:

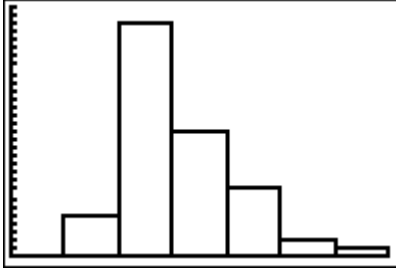
Skewness?

Remember: If mean $>$ median, then positively skewed

If mean $<$ median, then negatively skewed

Central Tendencies?

Histogram:



Number of colleges	Frequency
0	
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	

Here is an interesting set of statistics:

http://www.epinions.com/content_4821524612

#11 page 472

Towards the end of the season, a netballer had played 14 matches and had an average of 16.5 goals per game. In the final two matches of the season the netballer threw 21 goals and 24 goals. Find the netballer's new average.

<http://www.netball.org>

Statistics with Grouped Data

If you are given data that is in interval form, then use the midpoint of the interval to represent all scores within that interval.

Let's look at Example 6 on page 473



Image from Google

Find the approximate mean of the ages of the bus drivers data, to the nearest year.

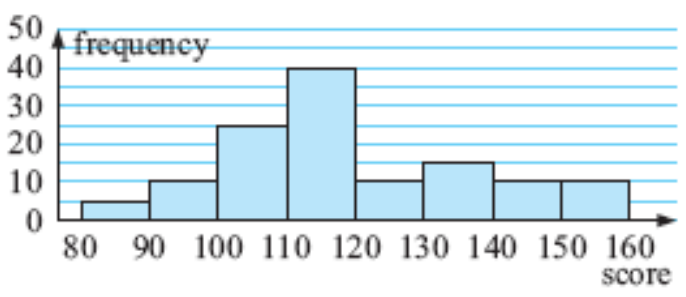
Find the approximate mean of the following *ages of bus drivers* data, to the nearest year:

<i>age (yrs)</i>	21-25	26-30	31-35	36-40	41-45	46-50	51-55
<i>frequency</i>	11	14	32	27	29	17	7

L1	L2	L3	2	1-Var Stats	1-Var Stats
23	11	-----		$\bar{x}=37.67153285$	$\uparrow n=137$
28	14			$\sum x=5161$	$\text{min}X=23$
33	32			$\sum x^2=203083$	$Q_1=33$
38	27			$Sx=7.97985468$	$\text{Med}=38$
43	29			$\sigma x=7.950677783$	$Q_3=43$
48	17			$\downarrow n=137$	$\text{max}X=53$
53	7				
L2(1)=11					

Let's try a problem that has the data already graphed in groups. See page 474 #3

AMERICA'S MOST SMARTEST MODEL



What do we need to do first?

Midpoint (x)	Frequency (f)	$f x$
85		
95		
105		
115		

125		
135		
145		
155		
	Σf	$\Sigma f x$

$$\bar{x} = \frac{\Sigma f x}{\Sigma f}$$

Now verify with your TI!

Cumulative Frequency and Cumulative Data

Our textbook says that sometimes we would like to know the number of scores that are above or below a certain value. When we study normal distributions, we will be looking for scores above and below a given “standard deviation” range.

Sometimes the cumulative frequency graph is a smooth curve and other times the graph is a collection of connected line segments. The CFG is sometimes called an ogive.

Here is a very simple cumulative frequency graph.

<http://www.cliffsnotes.com/WileyCDA/CliffsReviewTopic/Ogive-Cumulative-Line-Graph-topicArticleId-25951,articleId-25896.html>

For example, if you saved \$300 in both January and April and \$100 in each of February, March, May, and June, an ogive would look like Figure 1 .



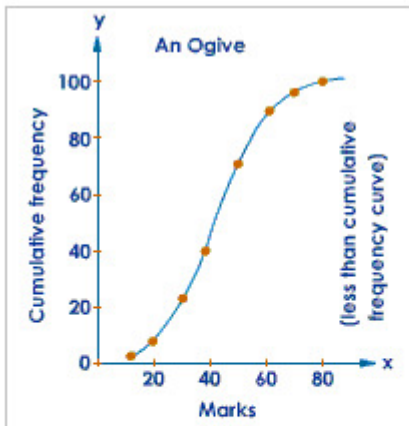
Figure 1 Ogive of accumulated savings for one year.

Consider the points on the graph:

(0, 0), (Jan, 300), (Feb, 400), (Mar, 500), (Apr, 800), (May, 900), and (Jun, 1000)

Here is an example of a smooth cumulative frequency graph:

<http://www.tutorvista.com/content/math/statistics-and-probability/statistics-graphical-representation/ogive.php>



Example 6:

Using the data given below, construct a 'more than' cumulative frequency table and draw the Ogive.

Marks	1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80
Frequency	3	8	12	14	10	6	5	2

Let's look at Example 7 on page 474

Image from Google Image



<i>Weight (w kg)</i>	<i>frequency</i>
$55 \leq w < 60$	2
$60 \leq w < 65$	3
$65 \leq w < 70$	12
$70 \leq w < 75$	14
$75 \leq w < 80$	19
$80 \leq w < 85$	37
$85 \leq w < 90$	22
$90 \leq w < 95$	8
$95 \leq w < 100$	2
$100 \leq w < 105$	1

Here is our data

Our textbook first makes a cumulative frequency table. The cumulative frequency is just a running total – think of the money saved in our first graph.

a

Weight (w kg)	frequency	cumulative frequency
$55 \leq w < 60$	2	2
$60 \leq w < 65$	3	5
$65 \leq w < 70$	12	17
$70 \leq w < 75$	14	31
$75 \leq w < 80$	19	50
$80 \leq w < 85$	37	87
$85 \leq w < 90$	22	109
$90 \leq w < 95$	8	117
$95 \leq w < 100$	2	119
$100 \leq w < 105$	1	120

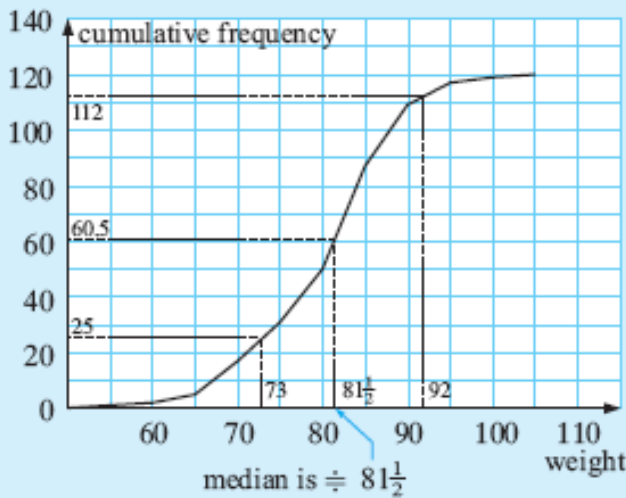
this is $2 + 3$
 this is $2 + 3 + 12$, etc.
 this 50 means that there are 50 players who weigh less than 80 kg

Note: The cumulative frequency gives a *running total* of the number of players up to certain weights.

Now let's look at the graph. What do you notice about the graph? Note: this graph is small and hard to read!

b

Cumulative frequency polygon of footballers' weights



- c i The median is the average of the 60th and 61st weights. Call it 60.5. Reading from the graph, the median $\hat{=} 81.5$.
- ii There are 25 men who weigh less than 73 kg.
- iii There are $120 - 112 = 8$ men who weigh more than 92 kg.

Things to consider:

We need to find the boundaries for each interval.

We need to find the cumulative frequency.

We need to plot the correct points, then make the line segments.

Let's look at an example from the Alan Wicks IBSL textbook.

Here is a table of values with the boundaries already given.

This data gives us an upper class boundary for the heights of a set of employees.

Boundary <i>L1</i>	Frequency	Cum. Freq. <i>L2</i>
≤ 154.5	0	0
≤ 159.5	5	5
≤ 164.5	8	13
≤ 169.5	8	21
≤ 174.5	10	31
≤ 179.5	10	41
≤ 184.5	6	47
≤ 189.5	3	50
≤ 194.5	2	52
≤ 199.5	1	53

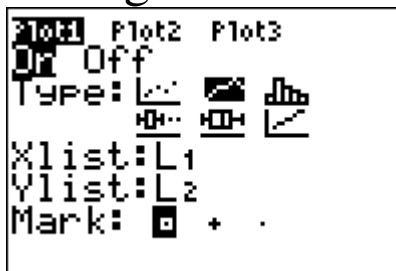
We can use our TI to make a cumulative frequency graph.

Put the values of the boundaries in L1 and the cumulative frequencies in L2.

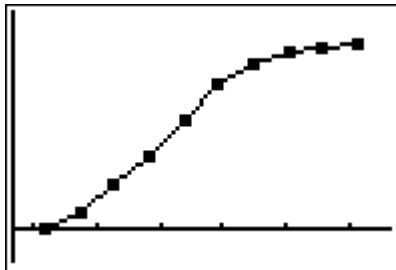
L1	L2	L3	2
154.5	0	-----	
159.5	5		
164.5	13		
169.5	21		
174.5	31		
179.5	41		
184.5	47		

L2 = {0, 5, 13, 21, 31, 41, 47}

Now go to STATPLOT and make the following choices:



Now choose ZOOMSTAT to get our beautiful graph!



Check your points with TRACE

So far, the IB exams usually have you read a cumulative frequency graph, rather than make one. To make one is very time consuming with a TI.

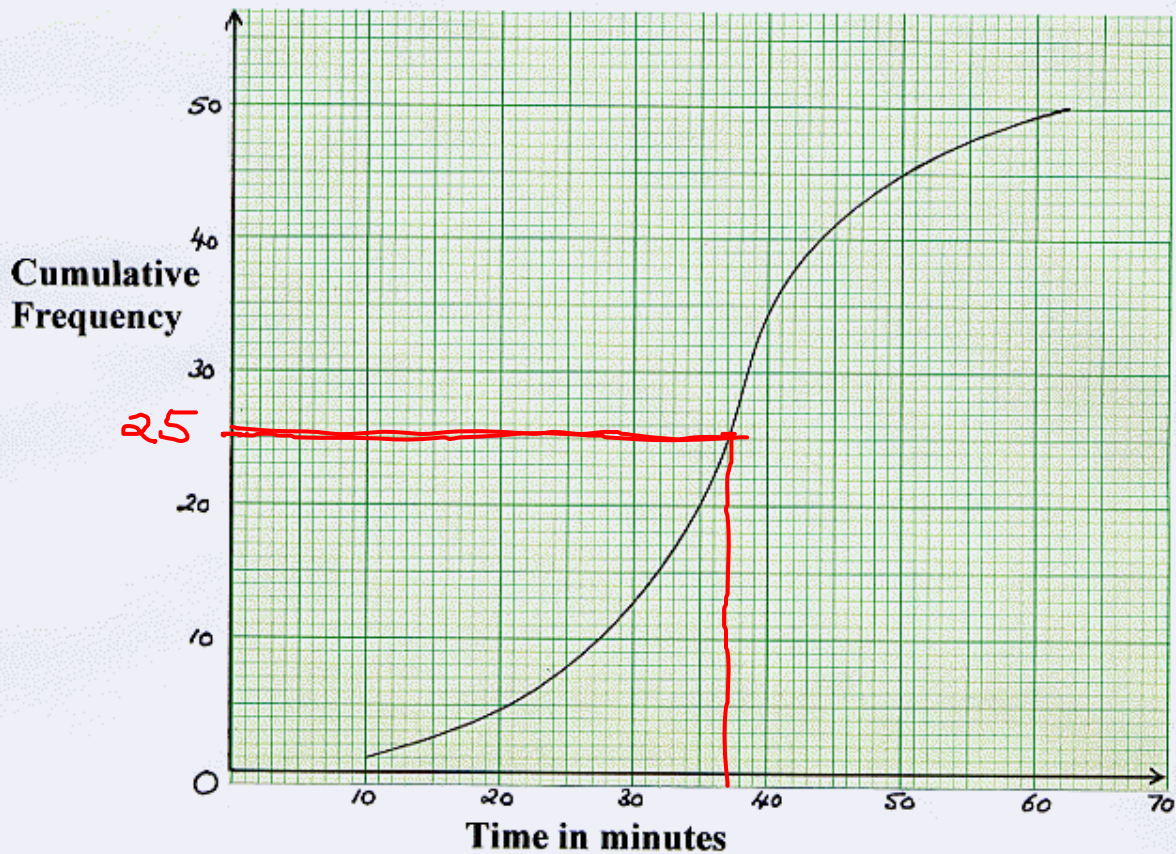
So here is a problem that I found that looks like a typical IB problem:

From: <http://www.gcsemathspastpapers.com/images/p5j04q13.htm>

From: <http://www.gcsemathspastpapers.com/images/p5j04q13.htm>

50 girls were each asked to record how many minutes it took them to complete a piece of maths homework.

The cumulative frequency graph below gives information about the times it took them to do this piece of homework.



Use the cumulative frequency graph to estimate the median time.

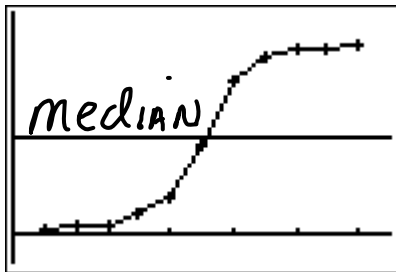
Let's try #1 on page 475

The following frequency distribution was obtained by asking 50 randomly selected people the size of their shoes.

Shoe size	5	$5\frac{1}{2}$	6	$6\frac{1}{2}$	7	$7\frac{1}{2}$	8	$8\frac{1}{2}$	9	$9\frac{1}{2}$	10
frequency	1	1	0	3	5	13	17	7	2	0	1

Draw an ogive for the data and use it to find:

- a the median shoe size
- b how many people had a shoe size of:
 - i $7\frac{1}{2}$ or more $\sim n = 40$
 - ii 8 or less. $\sim n = 40$



median ≈ 7.6

The table below gives the distribution of the life of electric light globes.

Draw an ogive for the data and use it to estimate:

- a the median life of a globe
- b the percentage of globes which had a life of 2700 hours or less
- c the number of globes which had a life between 1500 and 2500 hours.

Life (hours)	Number of globes
$0 \leq l < 500$	5
$500 \leq l < 1000$	17
$1000 \leq l < 2000$	46
$2000 \leq l < 3000$	79
$3000 \leq l < 4000$	27
$4000 \leq l < 5000$	4



For 1500 CF ≈ 45
 For 2500 CF ≈ 108

Median ≈ 2300

For a life of 2700 or less

