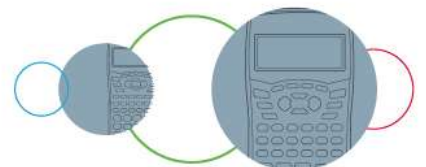


# SHARP

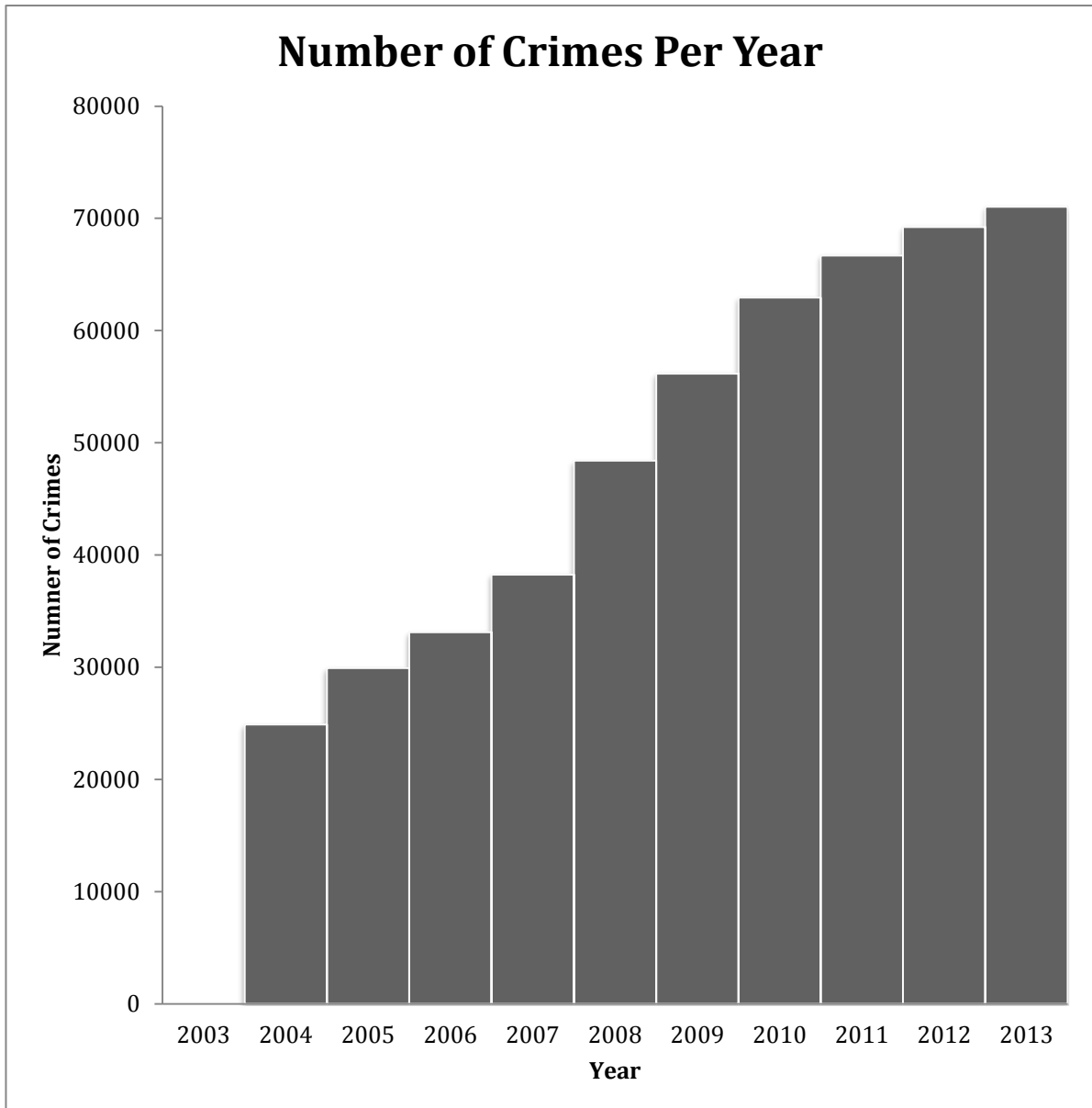
## Worksheet 11 Memorandum: Statistics

### Grade 11 Mathematics

1.
  - a) mean - the average of the data
  - b) variance – the average of the squared differences from the mean
  - c) histogram – a graph that represents continuous data and has rectangles drawn next to each other to represent the frequency of observations
  - d) bar graph – a graph that represents discrete data and has rectangles drawn with gaps between them to represent the frequency of observations
  - e) ogive – a graph that plots the cumulative frequency against the intervals for the data given.
  - f) symmetric – the graph the first half of the graph (e.g. a histogram) mirrors the second half of the graph around an approximate middle point.
  - g) skewed – when the graphs taller columns lie more to the left or the right and lower columns lie in the opposite direction.
  - h) outlier – a value that does not seem to fit in with the other data and may be due either to a strange occurrence or because the data was written down incorrectly.
  
2.
  - a)  $x$ -axis: the year from 2004 to 2013 with a + 1.5cm gap between each year  
 $y$ -axis: the number of crimes in thousands with a scale of 10 000 per 1.5cm



b)



c)  $\bar{x} = \frac{\text{sum of all observations}}{\text{number of observations}}$

$$\bar{x} = \frac{24\,886 + 29\,927 + 33\,116 + 38\,261 + 48\,405 + 56\,165 + 62\,939 + 66\,697 + 69\,232 + 71\,029}{10}$$

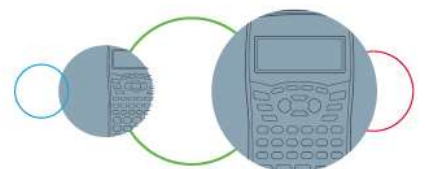
$$\bar{x} = \frac{500\,657}{10}$$

$$\bar{x} = 50\,066$$

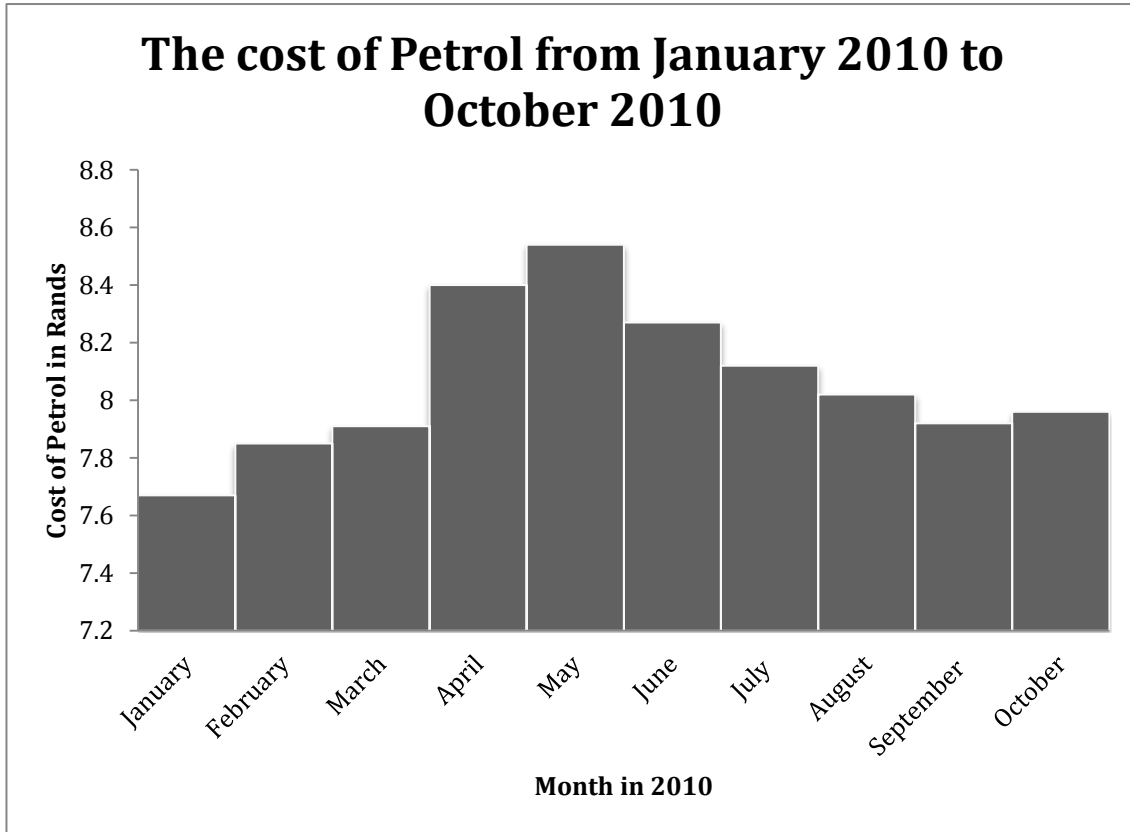
$$\sigma = 16\,599.1451$$

d) The general trend is that the number of people driving under the influence is increasing every year. This can be seen on the graph as every column is taller than the previous column.

e) The data is skewed to the left.



3. a)



b)  $\bar{x} = \frac{\text{sum of all observations}}{\text{number of observations}}$

$$\bar{x} = \frac{7.67+7.85+7.91+8.40+8.54+8.27+8.12+8.02+7.92+7.96}{10}$$

$$\bar{x} = \frac{80.66}{10}$$

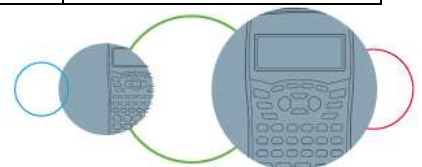
$$\bar{x} = 8.07$$

$$\sigma = 0.2536$$

c) The histogram is symmetric

4. a)

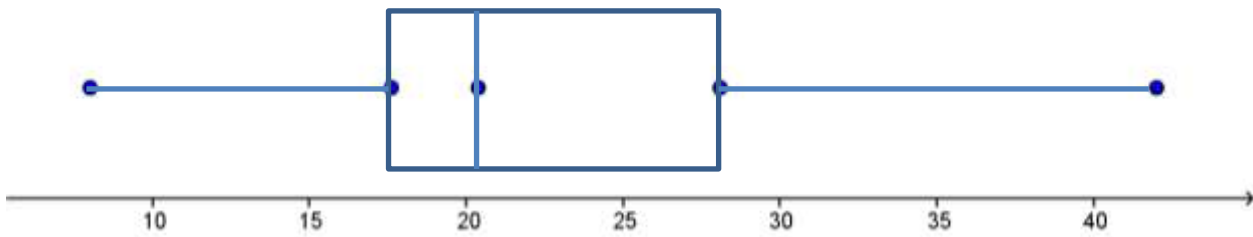
Country	1993 - 2000	2005-2008	Difference
Angola	37%	15.6%	21.4%
Botswana	15.1%	11.2%	3.9%
Cameroon	17.8%	16.6%	1.2%
Central African Republic	20.4%	28%	-7.6%
Democratic Republic of Congo	30.7%	28.2%	2.5%



Ethiopia	42%	34.6%	7.4%
Ghana	25.1%	13.9%	11.2%
Kenya	17.6%	18.4%	-0.8%
Lesotho	18.9%	16.6%	2.3%
Malawi	26.5%	15.5%	11%
Mozambique	28.1%	18.3%	9.8%
South Africa	8.0%	8.7%	-0.7%
Tanzania	19.6%	14.9%	4.7%
Zambia	26.9%	16.7%	10.2%
Zimbabwe	11.5%	14.0%	-2.5%

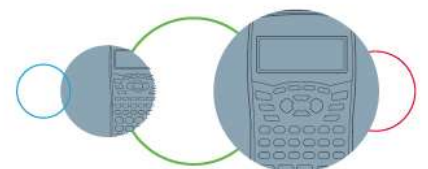
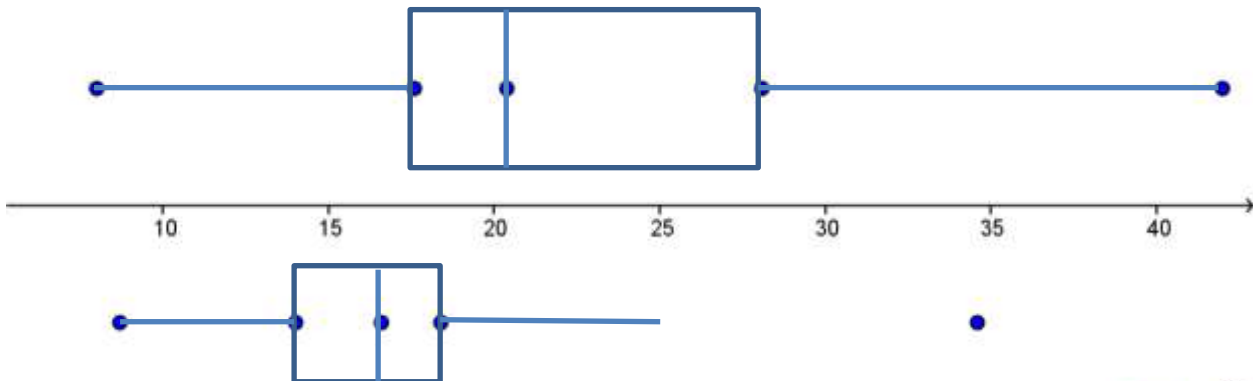
b) 8.0%      11.5%      15.1%      17.6%      17.8%      18.9%  
19.6%      20.4%      25.1%      26.5%      26.9%      28.1%  
30.7%      37.0%      42%

Min = 8.0%                      Max = 42%                      Q1 = 17.6%                      Q3 = 28.1  
Median = 20.4



c) 8.7%      11.2%      13.9%      14.0%      14.9%      15.5%  
15.6%      16.6%      16.6%      16.7%      18.3%      18.4%  
28.0%      28.2%      34.6%

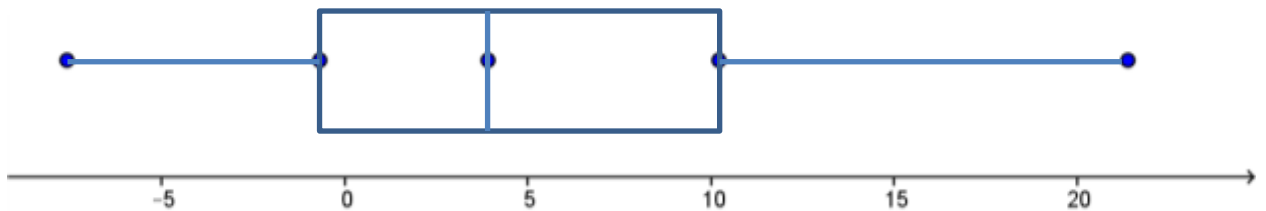
Min = 8.7%                      Max = 34.6%                      Q1 = 14.0%                      Q3 = 18.4%  
Median = 16.6%



d) The second box and whisker plot (2005 – 2008) has a much smaller spread of data than the first box and whisker plot (1993 – 2000). More than 50% of the second set of data lies within the first 25% of the first set of data. The second set of data has one outlier, which is still less than the maximum of the first set of data.

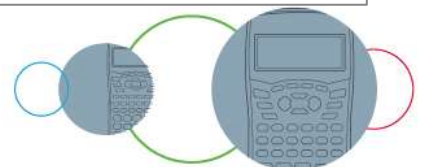
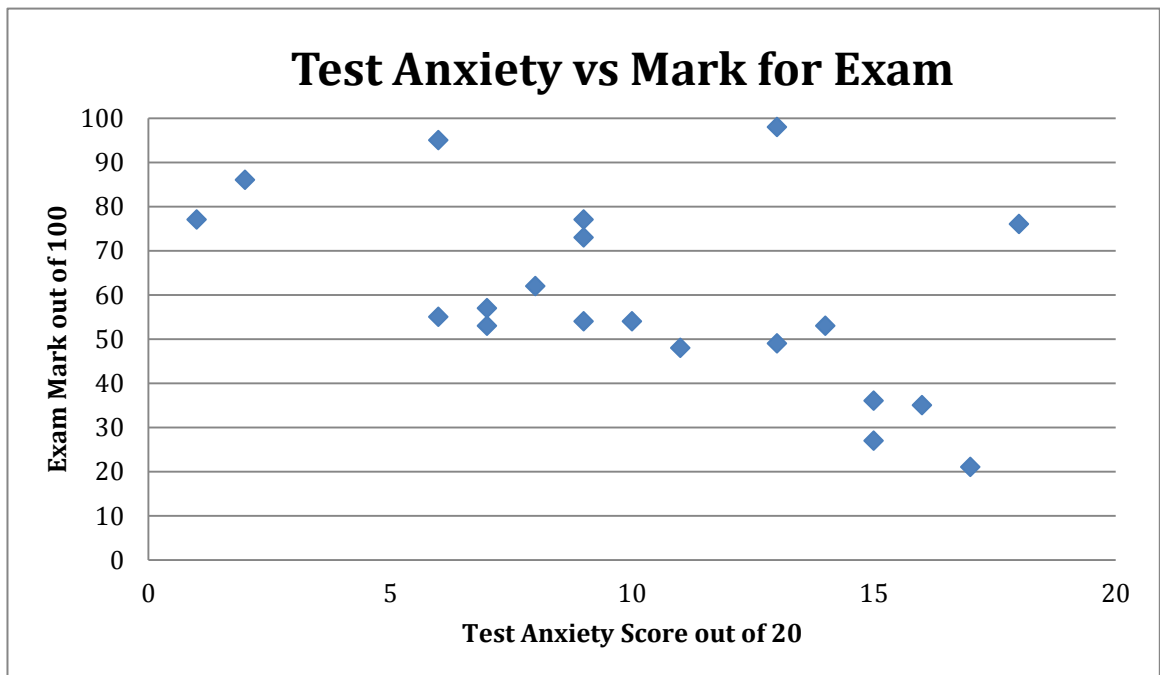
e) Yes, in the 2005 – 2008 set of data there is one outlier – 34.6

f)	-7.6%	-2.5%	-0.8%	-0.7%	1.2%	2.3%
	2.5%	3.9%	4.7%	7.4%	9.8%	10.2%
	11%	11.2%	21.4%			
	Min = -7.6%		Max = 21.4%	Q1 = -0.7%		Q3 = 10.2
	Median = 3.9%					



g) These box and whisker plots show that generally there has been a decrease in the number of children under the age of 5 that are undernourished between 1993 and 2008. Only 25% of the countries shown have an increased number of children who are undernourished, while 75% of the countries improved their undernourishment percentages.

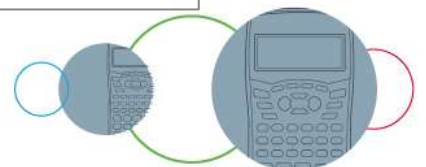
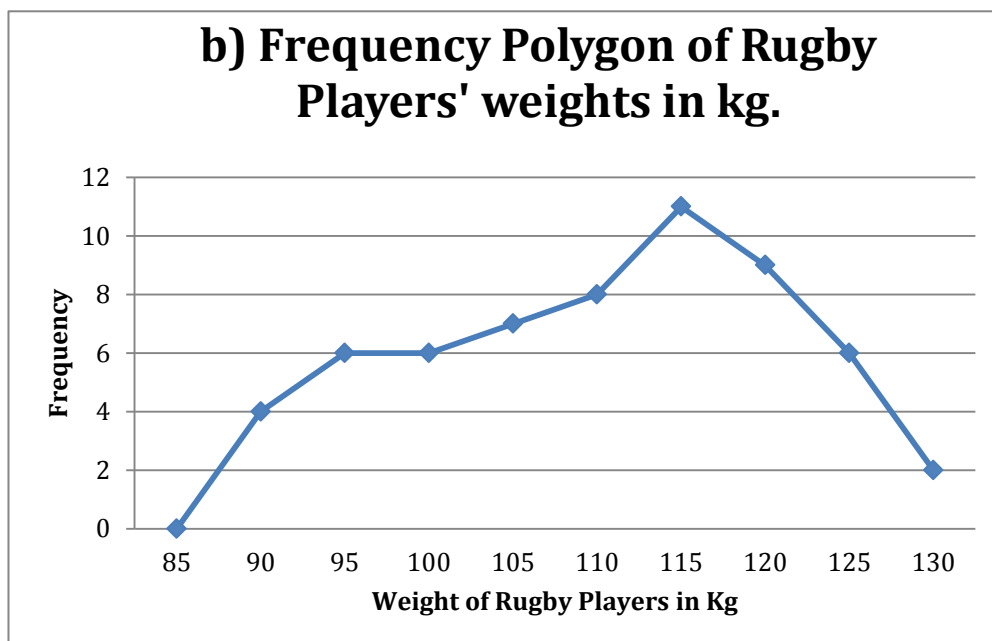
5. a)



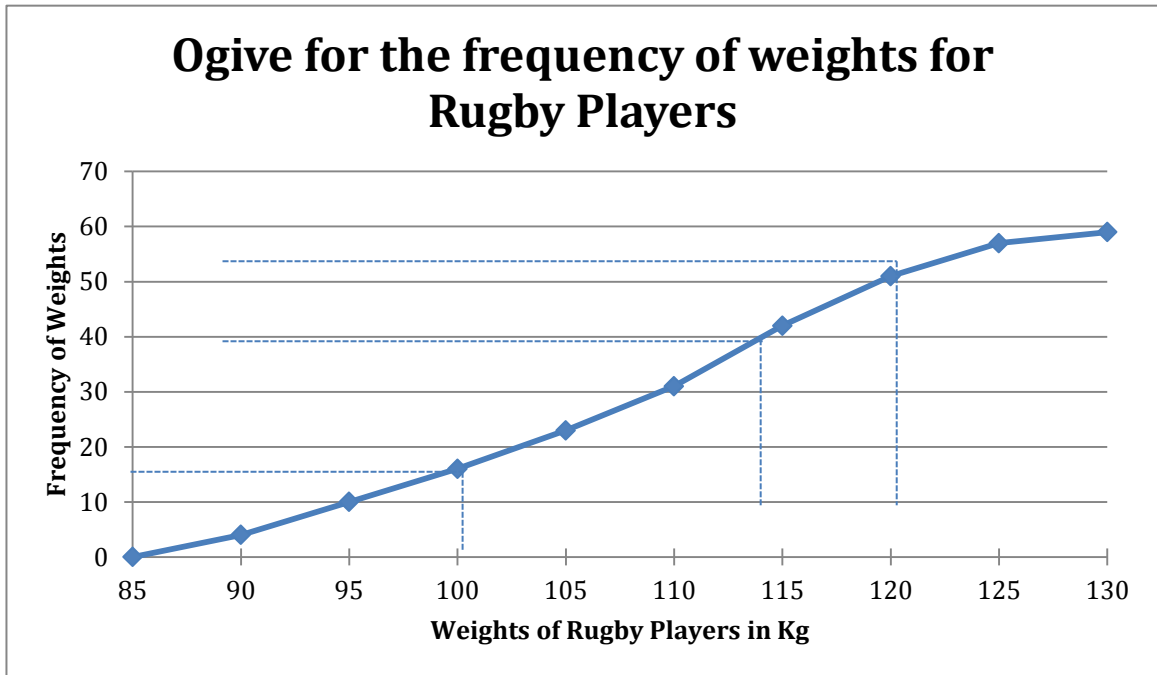
- b) Yes there are outliers. Any valid reason can be accepted some examples could be  
The student does well because they worry about the test and so he or she studies for it more and so does better than if he or she hadn't worried about it.  
The mark or test anxiety score could have been recorded incorrectly.
- c) Yes, there is a negative linear relationship for the data.
- d) Yes, the researcher does seem to have proved his theory as there is a general trend that as the students' test anxiety scores increase their marks for the exam go down.

6. a)

Weight (in Kg)	Frequency	Cumulative Frequency
$85 \leq x < 90$	4	4
$90 \leq x < 95$	6	10
$95 \leq x < 100$	6	16
$100 \leq x < 105$	7	23
$105 \leq x < 110$	8	31
$110 \leq x < 115$	11	42
$115 \leq x < 120$	9	51
$120 \leq x < 125$	6	57
$125 \leq x < 130$	2	59



c)



- d) First Quartile = ± 100  
 Median = ± 109  
 Third Quartile = ± 116

- e) How many rugby players way more than 116?  
 Thus  $59 - 42 = 17$  rugby players are in the top 25% of weights.

7. a) 19° 20° 20° 21° 22° 24° 25° 25° 25° 25° 26° 26°  
 26° 27° 27° 27° 28° 28° 29° 29° 29° 29° 30° 30°  
 30° 30° 31° 31° 32° 32°

$$\bar{x} = \frac{\text{sum of all observations}}{\text{total number of obseravtions}}$$

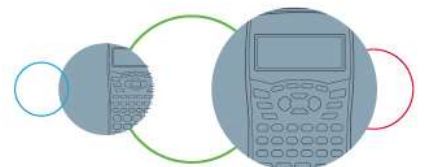
$$\bar{x} = \frac{19+20+20+21+22+24+25+25+25+25+26+26+27+27+27+28+28+29+29+29+29+30+30+30+30+30+31+31+32+32}{30}$$

$$\bar{x} = \frac{803}{30}$$

$$\bar{x} = 26.77$$

$$\text{Median} = (27 + 27) \div 2 = 27$$

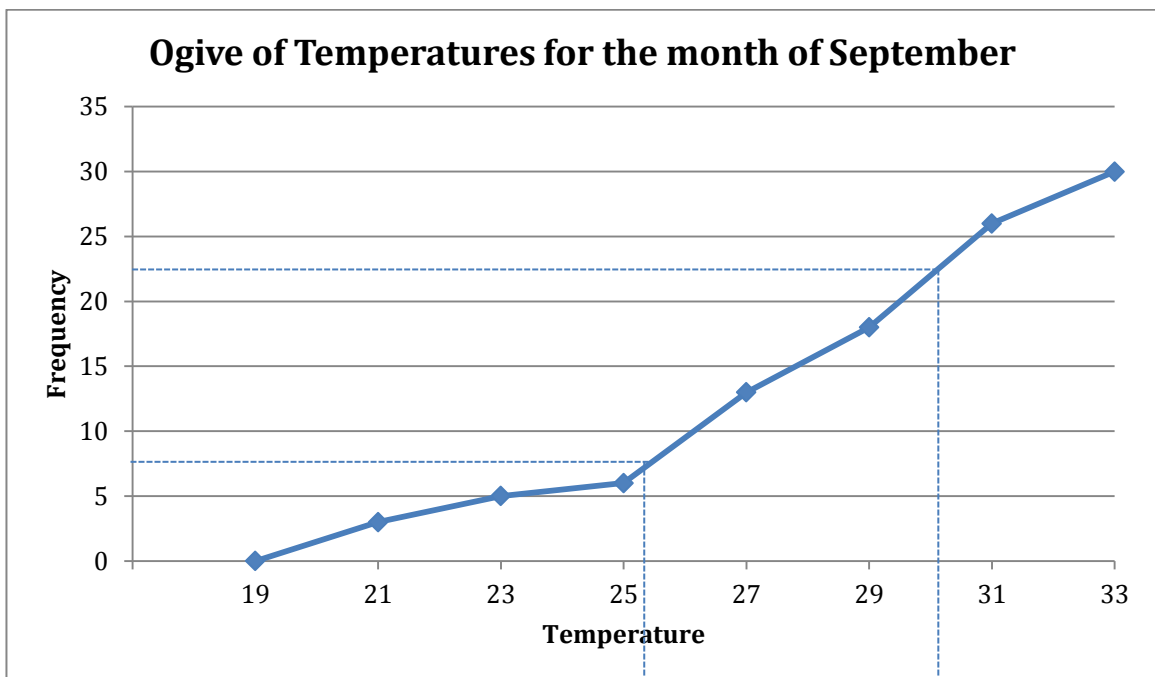
$$\text{Mode} = 25, 29, 30$$



b)  $\sigma = 3.58$

c)

Group	Frequency	Cumulative Frequency
$19 \leq x < 21$	3	3
$21 \leq x < 23$	2	5
$23 \leq x < 25$	1	6
$25 \leq x < 27$	7	13
$27 \leq x < 29$	5	18
$29 \leq x < 31$	8	26
$31 \leq x < 33$	4	30



d)

A

B

e) Labelled on the graph above A – First Quartile =  $\pm 25$   
 B – Third Quartile =  $\pm 30$

f) Interquartile Range =  $30 - 25 = 5$

