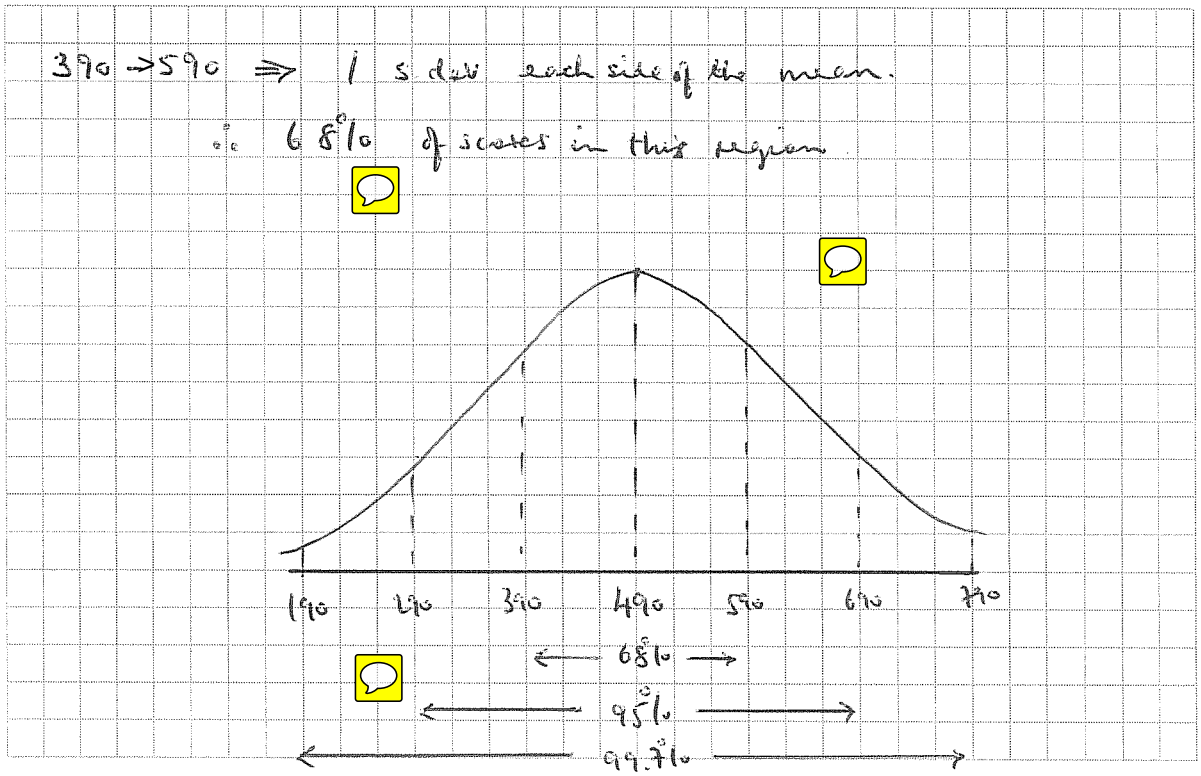
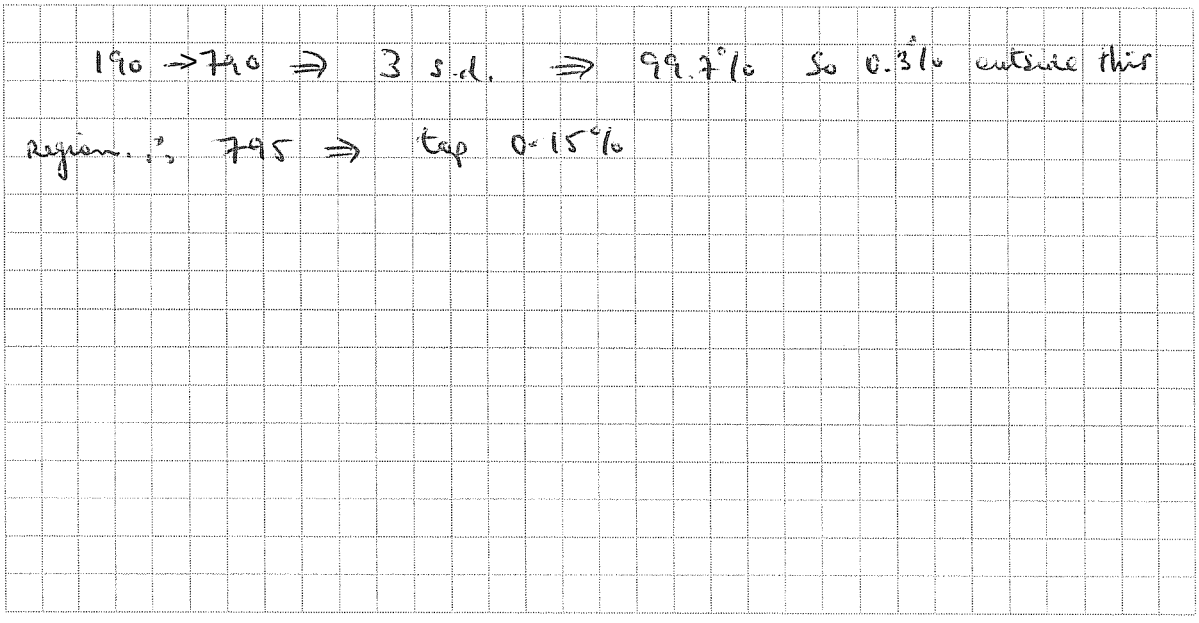
Question LCHL: Descriptive Statistics

To enter a particular college course, candidates must complete an aptitude test. In 2010 the mean score was 490 with a standard deviation of 100. The distribution of the scores on the aptitude test is a normal distribution.

(a) What percentage of candidates scored between 390 and 590 on this aptitude test?



(b) One student scored 795 on this test. How does this student's score compare to the rest of the scores?



(c) The college admits only students who were among the highest 16% of the scores on this test. What score would a student need on this test to be qualified for admission to this college? Explain your answer.

Highest 16% ⇒ 32% on the "tails".

∴ 68% Region ⇒ 1 standard deviation from mean

Thus highest 16% is above 590 on test.

The College will except students who score higher than 590

(d) Alice is preparing to sit the aptitude test in 2011. She heard that a score of over 650 would guarantee her a place on the course. She knew 20 people who were going to take the test. Based on the mean and standard deviation in 2010, approximately how many of the people Alice knew were likely to get a score of above 650 and secure a place on the course? Justify your answer.

$\frac{650-490}{100} = 1.6$ Z score of 1.6 ⇒ 0.9452

$1 - 0.9452 = 0.0548$

$20 \times 0.0548 = 1.096$

Thus approximately 1 person out of the 20 is likely to score above 650 and secure a place on the course.

LCOL: Descriptive Statistics.

David noticed that, when he drank a bottle of sports drink before going out for a run one day, his performance time improved. He set about doing an experiment to see whether drinking the sports drink increases performance when running.

He recorded the times of people in his running club to complete a 5km run without drinking the sports drink and then on another day he recorded the time it took the same people to complete 5km having taken the sports drink.

He recorded the information in a back-to-back stem and leaf plot:

Without taking the sports drink.		**Having taken the sports drink.**
5	20	3 4
1 1 1	21	3 4 7
8 8 4 3 2	22	
	23	1 2 2
	24	0
	25	8
	26	1
	27	
	28	2 3 6 6 7 7
	29	2 4 4 5 5 5 8 9 9
	30	1 3 4 5 6 7 8 8 9
5	31	
6 4 4 3 0 0	32	1 1 4 9
9 9 6 5 4 4 3 3 2 1	33	3 3 3 2
7 7 5 5 6 6 6 1 0	34	5
8 8 8 3 3	35	0 0
7 3 2	36	1
1	37	2
	38	3 5
2 2	39	
4 4 2 0	40	

Key: 32 | 1 means 32.1 minutes

(i) Based on the diagrams, approximate the median speed without taking the sports drink and the median speed having taken the sports drink. What does this information tell you?

Median times: No sports drink → 33.9 mins
with sports drink → 29.9 mins

median speeds: No sports drink → $\frac{5}{33.9} = 0.147$ km/min

with sports drink → $\frac{5}{29.9} = 0.167$ km/min

So the median speed increased when they ran 5 km after taking the sports drink.

(ii) Compare the distributions of each of the data sets above.

The Range of times without the sports drink is 30.5 - 40.4 mins
The Range of times after taking the sports drink is 30.3 - 38.5 mins

The distribution of times without the sports drink is more skewed than the times after taking the sports drink, which is more symmetrical

For the data without the sports drink, the data are clustered around 32 - 36 minutes, whereas for the data where the Runners had taken the sports drink the times are clustered around 28 - 30 minutes.

(iii) Is there evidence from the diagram to suggest that taking the sports drink improves performance? Justify your conclusions.

There is evidence to suggest that performance improves after taking the sports drink. The range of times is smaller after taking the drink before running the 5 km. Without the drink only 20% of the runners took less than 32 mins to run the 5 km. after taking the drink, 70% of the runners completed the 5 km in less than 32 mins.

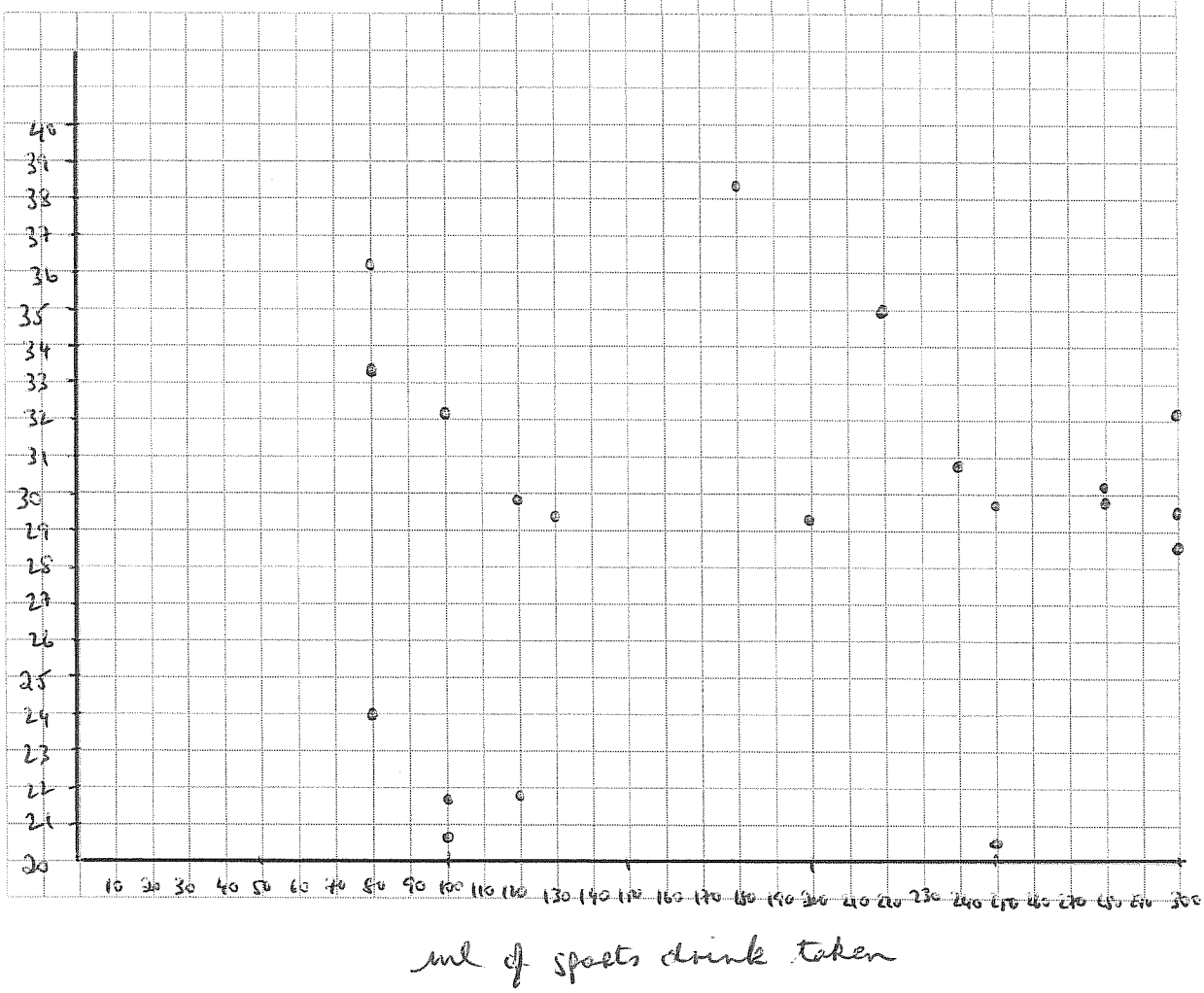
(iv) Make an argument, based on the two data sets, that taking the sports drink does not improve performance.

18% of the runners took less than 23 minutes to run the 5 km without taking the sports drink. Only 10% of the runners ran the 5 km, after taking the drink, in less than 23 minutes

(v) After completing the experiment, David wondered how accurate his study was. He realised that he had not specified how much of the sports drink the runners should take. He asked 20 of the runners approximately how many millilitres of sports drink they had taken and recorded this alongside their time. The results are as follows:

Time (mins)	Sports drink (ml)
20.3	250
21.7	100
21.8	120
24	80
28.6	300
29.4	130
29.5	300
29.9	280
32.1	300
32.1	100
33.2	80
35	220
38.3	180
20.6	100
29.2	200
29.8	250
36.1	80
29.9	120
30.9	240
30.1	280

Display the data in a way that allows you to examine the relationship between the two data sets.



(vi) Is there evidence to suggest that there is a relationship between the time taken to complete 5km and the amount of sports drink taken before the race?

No, the scatterplot doesn't show a noticeable relationship between the time taken to Run the 5 km and the amount of sports drink taken before the Run.

(vii) The correlation coefficient for data in part **(v)** above is one of the following. Circle the correct correlation coefficient, based on your graph.

A -0.82 **B** 0.13 **C** 0.95 **D** 0.6

JCOL: Descriptive Statistics

A group of students were asked "*Do you get worried about your exams*?" They were asked to circle one of following to answer the question: Never, Rarely, Sometimes, Frequently.

The data below shows the answers from a sample of boys and girls.

Boys
Frequently
Never
Never
Sometimes
Sometimes
Rarely
Sometimes
Sometimes
Frequently
Never
Sometimes
Rarely
Rarely
Frequently
Never
Rarely
Rarely
Frequently
Never
Frequently
Never
Sometimes
Never
Frequently
Rarely
Sometimes
Rarely
Never
Sometimes
Rarely

Girls
Never
Sometimes
Sometimes
Rarely
Never
Frequently
Frequently
Never
Sometimes
Rarely
Frequently
Rarely
Sometimes
Frequently
Frequently
Frequently
Frequently
Rarely
Frequently
Frequently
Frequently
Sometimes
Sometimes
Sometimes
Never
Rarely
Frequently
Frequently
Never
Never
Frequently

(a) How many students were in each sample?

30

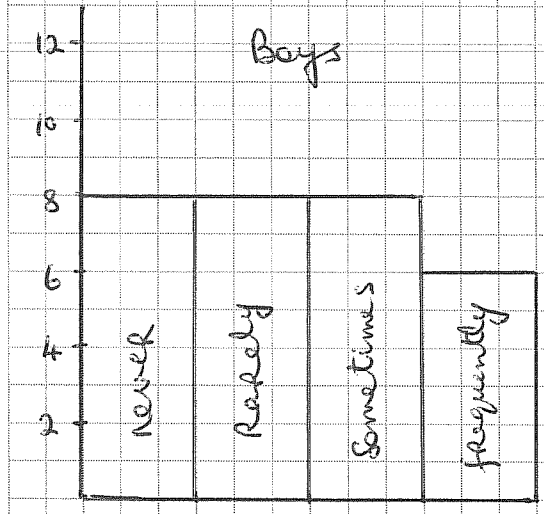
(b) Display the data in a way which allows you to compare the two samples.

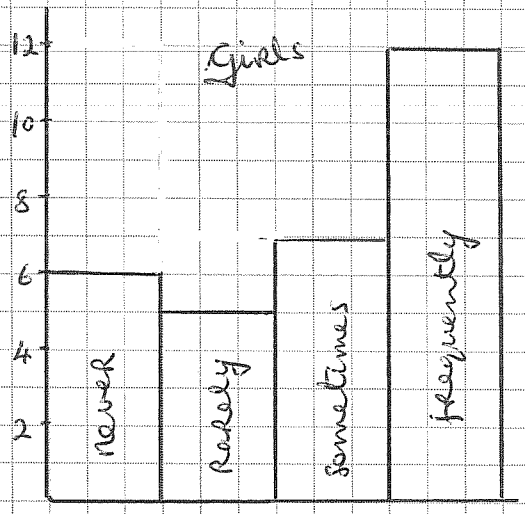
Boys

Frequently	6
Sometimes	8
Rarely	8
never	8

Girls

Frequently	12
Sometimes	7
Rarely	5
never	6





(c) Compare the two sets based on your display.

More girls than boys said they worried frequently (twice as many).

More boys than girls said they Rarely or never worry about these exams (16 compared to 11.