



General Certificate of Secondary Education
2023

Centre Number

--	--	--	--	--

Candidate Number

--	--	--	--	--

Statistics

Unit 2 (With Calculator)

Higher Tier

[GST22]



GST22

MONDAY 19 JUNE, AFTERNOON

TIME

2 hours.

INSTRUCTIONS TO CANDIDATES

Write your Centre Number and Candidate Number in the spaces provided at the top of this page.

You must answer the questions in the spaces provided.

Do not write outside the boxed area on each page or on blank pages.

Complete in black ink only. **Do not write with a gel pen.**

Any working **must** be clearly shown in the spaces provided. Marks may be awarded for partially correct solutions.

Answer **all ten** questions.

INFORMATION FOR CANDIDATES

The total mark for this paper is 100.

Figures in brackets printed down the right-hand side of pages indicate the marks awarded to each question or part question.

You should have a calculator, ruler, compasses and protractor.

The formula sheet is on page 2.



HIGHER TIER FORMULA SHEET

$$\text{Standard deviation} = \sqrt{\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2}$$

Spearman's Rank Correlation Coefficient

$$r_s = 1 - \left(\frac{6 \sum d^2}{n(n^2 - 1)}\right)$$





BLANK PAGE
DO NOT WRITE ON THIS PAGE
(Questions start overleaf)

13832.06 R

[Turn over



32GST2203

Answer all questions

- 1 The table below shows the crop areas, in hectares, in June for the period 2016 to 2020

Crop areas in Northern Ireland, 2016–2020 (hectares)

	2016	2017	2018	2019	2020
Grass					
Under 5 years old	148 166	144 354	144 454	148 818	141 424
Over 5 years old	652 613	660 598	663 166	659 927	674 833
Total grass	800 780	804 952	807 620	808 745	816 257
Total hill or rough land	137 191	140 400	143 156	143 609	140 581
Cereals					
Oats	2257	2252	2013	1868	1879
Wheat	8616	8730	6845	8104	7132
Barley (winter)	7628	7114	5809	7739	7772
Barley (spring)	14 708	14 031	14 894	11 948	12 564
Mixed corn	230	170	101	161	113
Total cereals	33 439	32 298	29 662	29 820	29 460
Other field crops					
Potatoes	3737	4053	3601	3777	3715
Arable crop silage	3324	3605	4272	3924	4187
Forage maize	1294	1334	1558	1618	1874
SRC willow	593	644	655	586	533
Other crops	2086	2321	2391	2392	3114
Total other field crops	11 033	11 957	12 477	12 296	13 422
Horticultural crops					
Fruit	1501	1522	1474	1441	1330
Vegetables	1234	1262	1125	1163	1178
Ornamentals	182	188	199	179	172
Total horticultural crops	2917	2973	2798	2782	2680
Total crops	47 388	47 227	44 937	44 898	45 562
Other land					
Woods and plantations	16 133	15 786	16 307	15 669	17 276
Other land	10 486	10 060	10 376	10 243	10 145
Total other land	26 618	25 846	26 683	25 912	27 421

13832.06 R



32GST2204

(a) Describe the trend in the total number of hectares of grass from 2016 to 2020

_____ [1]

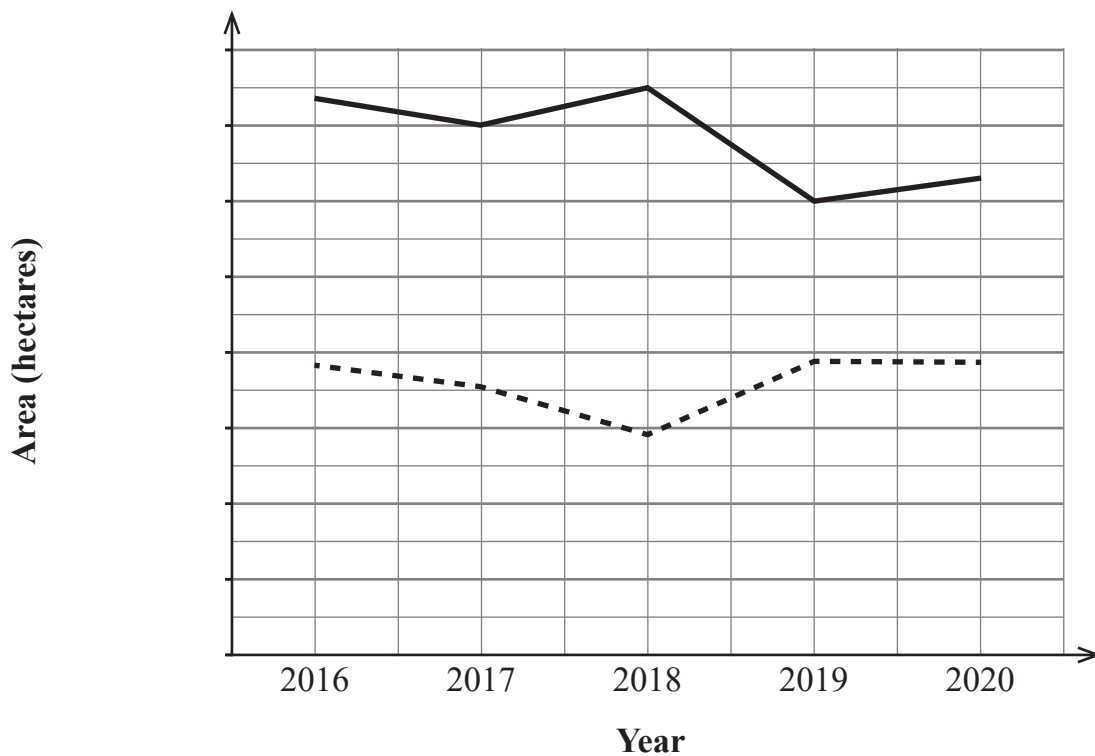
According to the table, the total area of land used for horticultural crops was 2782 hectares in 2019

However, $1441 + 1163 + 179 = 2783$

(b) Suggest a reason for this difference.

_____ [1]

The line graph below shows the crop areas of land used for two types of cereal between 2016 and 2020



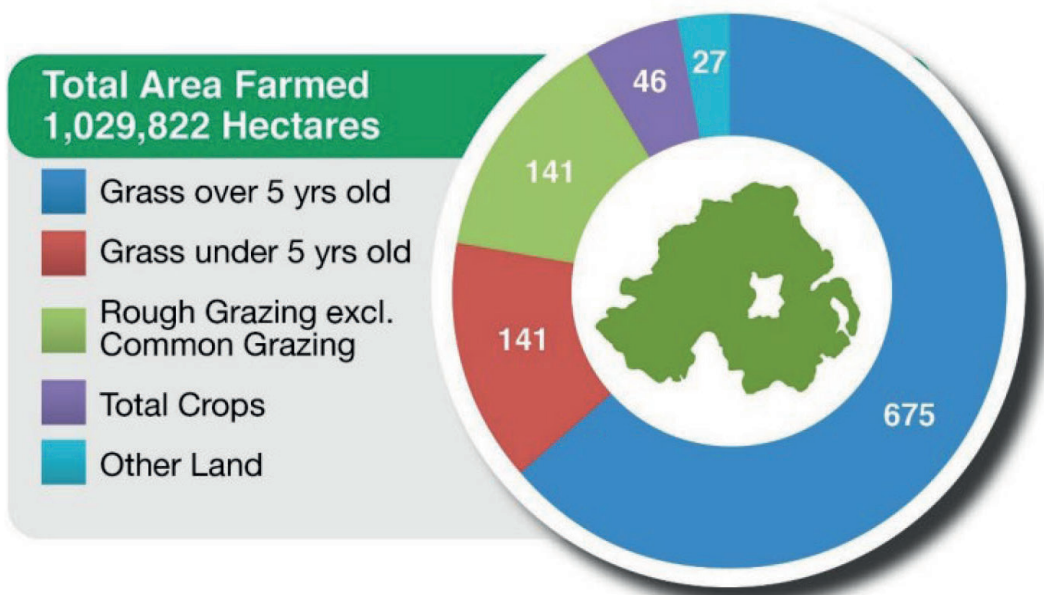
(c) Using figures from the table, complete the vertical axis and label each line. [2]

[Turn over



The chart below shows the land use by agricultural production type for 2020

Land use by agricultural production type, 2020



Source: © Crown Copyright / NISRA

(d) Explain how the numbers in each category in this chart could be misleading.

[1]





BLANK PAGE

DO NOT WRITE ON THIS PAGE

(Questions continue overleaf)

13832.06 R

[Turn over



32GST2207

- 2 The table below shows the total agricultural labour force by farm business size in Northern Ireland for 2020

Farm business size				
Very small	Small	Medium	Large	Total
34 138	6727	3281	7155	51 301

Aidan is planning an investigation about working conditions among farm workers.

- (a) What is the population for Aidan's investigation?

_____ [1]

- (b) Comment on the availability of a sampling frame for Aidan's investigation.

_____ [1]

As part of his investigation, Aidan wishes to find out if workers on large farms work more hours than those on small farms.

- (c) Write a hypothesis for Aidan's investigation.

_____ [1]

- (d) For this part of his investigation, write down the names of two variables for which Aidan will need to collect data.

1. _____ [1]

2. _____ [1]



- (e) Could the product moment correlation coefficient be used with the variables you identified in part (d)?

Explain your answer.

[2]

Aidan plans to distribute a questionnaire to farm workers.

He designs a question to determine how many hours each of them worked:

How many hours do you work?

Number of hours: _____

- (f) Suggest a more suitable question and response section.

[2]

[Turn over



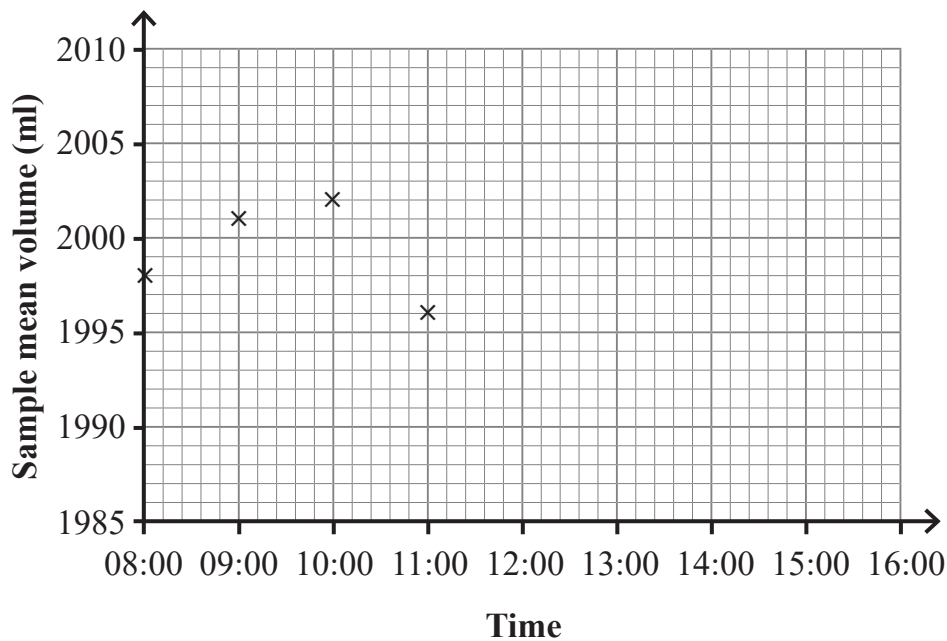
3 A machine is programmed to fill bottles with 2000 ml of lemonade.

Samples are taken every hour and the sample mean is recorded.

The sample means for one day are given in the table below.

Time	Sample mean volume (ml)
08:00	1998
09:00	2001
10:00	2002
11:00	1996
12:00	2003
13:00	2003
14:00	2004
15:00	2007

(a) Plot the last four sample means on the control chart below.



[2]



The production manager has decided that warning lines should go onto the control chart at 5 ml above and below the target volume of 2000 ml.

Action lines should be added at 8 ml above and below the target volume.

(b) Draw and label the target, warning and action lines on the control chart. [3]

(c) What advice would you give to the production manager based on the control chart?

[3]

[Turn over



The general manager and the production manager have a meeting to discuss how the quality assurance system might be improved.

The general manager states:

“We should take samples every half an hour.”

(d) Give an advantage and a disadvantage of this suggestion.

Advantage _____
_____ [1]

Disadvantage _____
_____ [1]

The production manager states:

“We should reduce our warning and action limits to 3 ml and 5 ml outside the target volume.”

(e) Give an advantage and a disadvantage of this suggestion.

Advantage _____
_____ [1]

Disadvantage _____
_____ [1]

(f) Why would it be useful to calculate a sample range as well as a sample mean as part of the quality assurance process?

_____ [2]





BLANK PAGE

DO NOT WRITE ON THIS PAGE

(Questions continue overleaf)

[Turn over

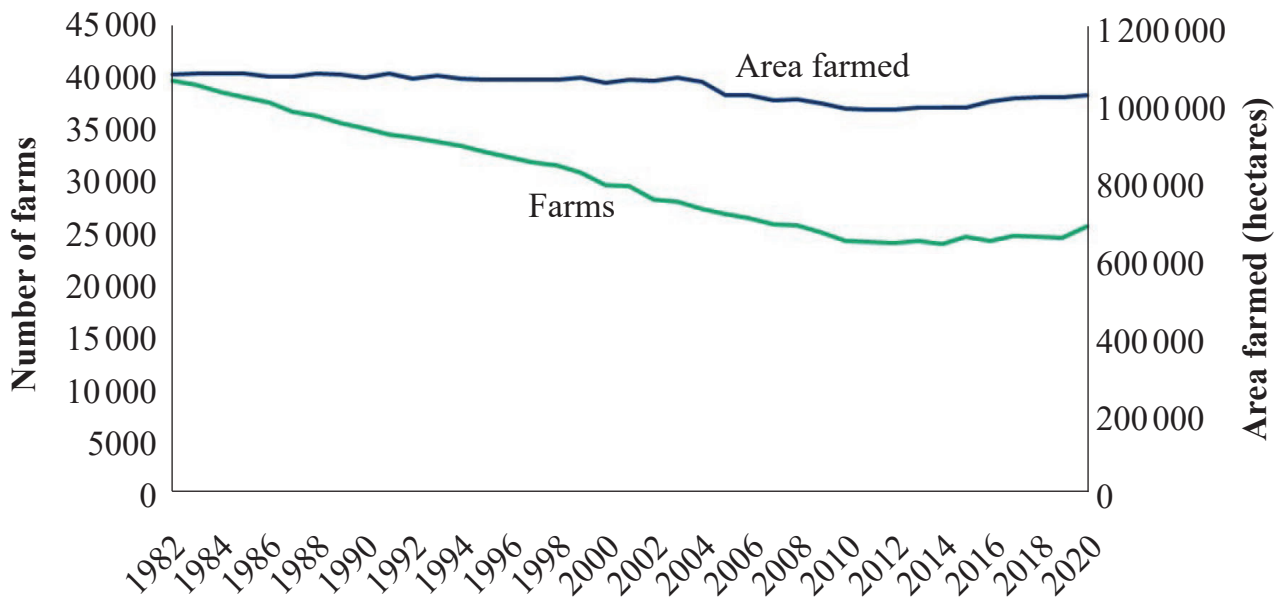
13832.06 R



32GST2213

- 4 The diagram below shows the number of farms and area farmed in Northern Ireland between 1982 and 2020

Number of farms and area farmed, 1982–2020



Source: © Crown Copyright / NISRA

- (a) Calculate the approximate mean area farmed on a farm in Northern Ireland in 2020

_____ hectares [3]

- (b) Compare the trend in the number of farms with the trend in the area farmed.

 _____ [2]



(c) Suggest a reason for the differences in these trends.

[1]

[Turn over

13832.06 R



32GST2215

The table below shows the total number of pigs and the number of large farms in Northern Ireland from 2010 to 2019

Year	Total number of pigs	Number of large farms
2010	424 604	1493
2011	425 268	1518
2012	426 924	1558
2013	480 317	1534
2014	517 075	1546
2015	569 738	1629
2016	601 101	1701
2017	649 120	1722
2018	633 644	1693
2019	674 428	1680

(d) Calculate the product moment correlation coefficient for the data in the table.

_____ [2]

(e) (i) State the type of correlation found between the number of large farms and the total number of pigs.

_____ [1]

(ii) Give an interpretation of what the value calculated in part (d) means in context.

 _____ [1]





BLANK PAGE

DO NOT WRITE ON THIS PAGE

(Questions continue overleaf)

[Turn over

13832.06 R



32GST2217

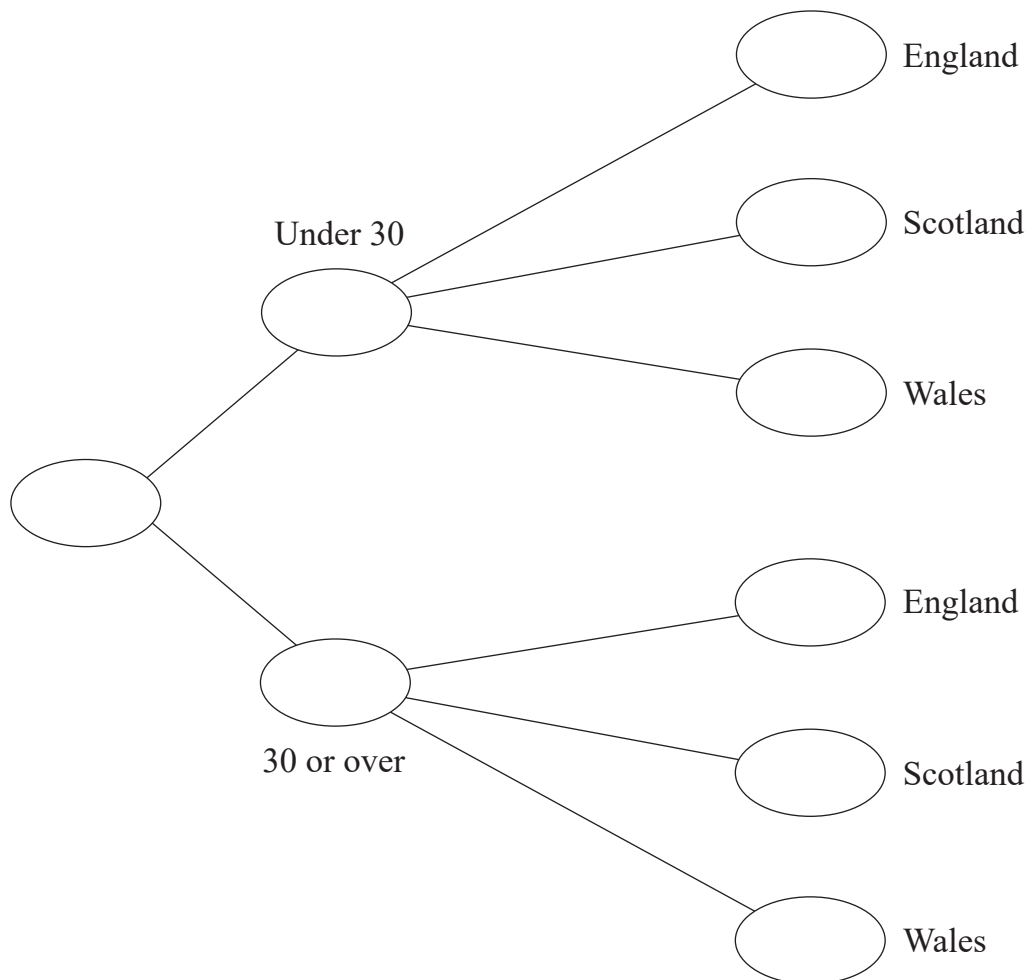
5 Janet collected data on holiday destinations for some people from Northern Ireland in July 2021

She grouped the results by the age of each respondent.

The table below shows her results.

Holiday destination	Age (years)	
	Under 30	30 or over
England	26	34
Scotland	19	33
Wales	40	53

(a) Use the information in the two-way table to complete the frequency tree below.



[3]



- (b) Using the data opposite, calculate the probability that a person, chosen at random,
- (i) was under 30 years old and went on holiday to Scotland;

Answer _____ [2]

- (ii) went on holiday to England.

Answer _____ [2]

One person aged under 30 years old was selected at random.

- (c) What is the probability that this person did not go on holiday to Wales?

Answer _____ [2]

Janet thinks that those 30 years old or over are more likely to go on holiday in Wales than those under 30 years old.

- (d) Is she correct? Give a reason for your answer.

[3]

- (e) Give a reason why the frequency tree was a better way to display the data than the table.

[1]

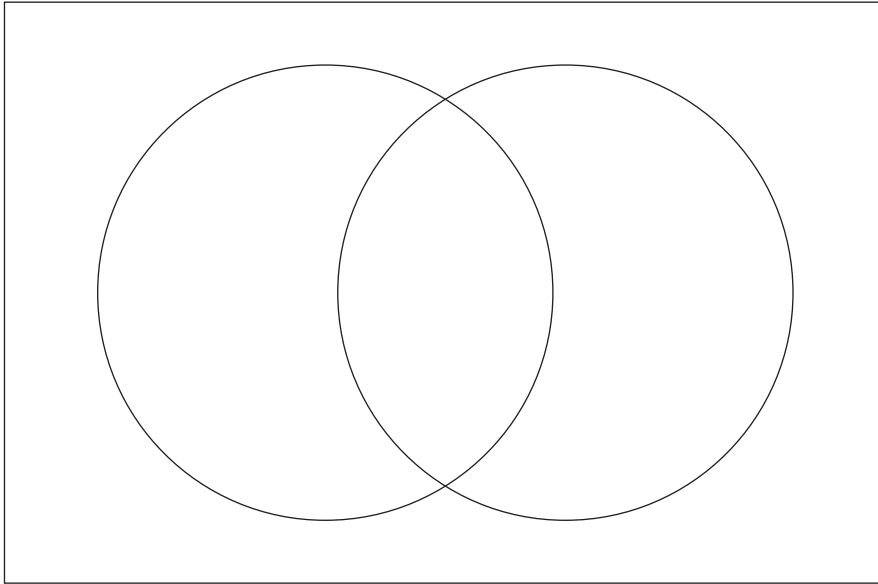
[Turn over



6 Customers at an ice-cream shop can have toppings and/or sauce on their ice-cream cone.

- Two-fifths of the customers take sauce only.
- 15% of the customers take sauce and toppings.
- One in five customers do not take sauce or toppings.

(a) Complete the Venn diagram below to illustrate this information.



[2]

(b) Find the probability that a customer, selected at random, takes toppings or sauce, but not both, on their ice-cream cone.

Answer _____ [2]



(c) Given that a customer, selected at random, does not take a topping on their ice-cream cone, what is the probability that they do not take sauce either?

Answer _____ [2]

(d) During one Sunday, 80 ice-cream cones are sold.

(i) How many of these would you expect to be with sauce or toppings or both?

Answer _____ [1]

(ii) How might the information in part (d)(i) be useful for the manager of the ice-cream shop?

[1]

[Turn over



7 Imogen is comparing the journey times of two groups of passengers.

The journey times, in minutes, for group A are recorded in the table below.

Journey time, t (minutes)	Frequency	Cumulative frequency
$0 < t \leq 60$	2	
$60 < t \leq 120$	3	
$120 < t \leq 180$	7	
$180 < t \leq 210$	8	
$210 < t \leq 240$	6	
$240 < t \leq 300$	4	

(a) Complete the cumulative frequency column in the table above.

[2]

(b) How many passengers in group A had a journey time of 2 hours or less?

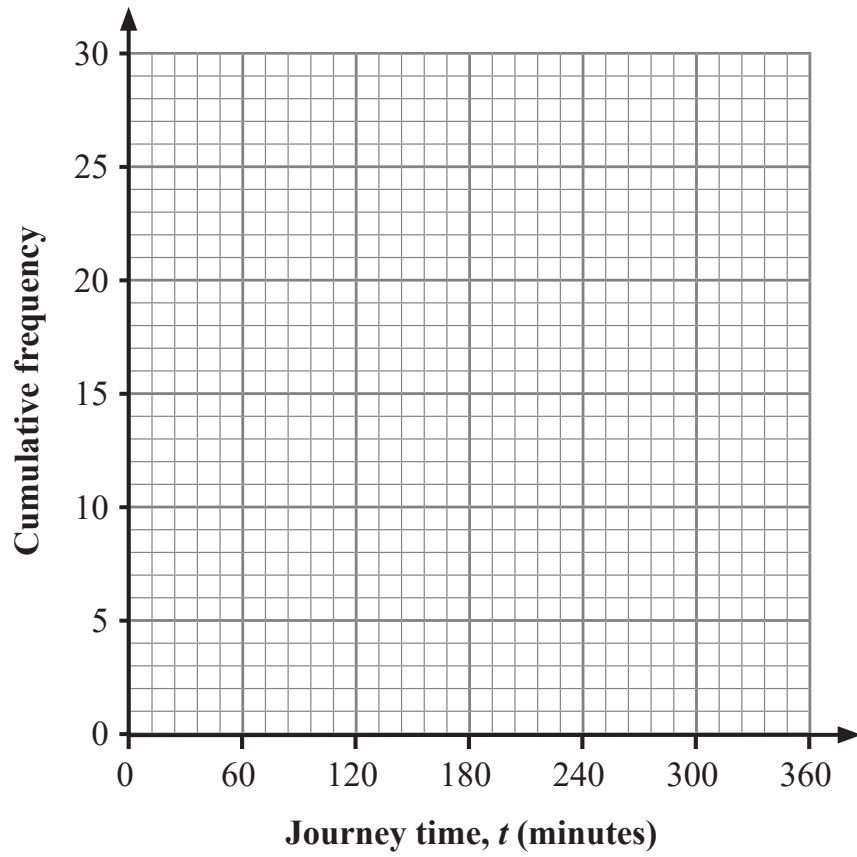
Answer _____ [1]

(c) Write down the class interval which contains the median.

Answer _____ [1]



(d) Using the axes below, draw a cumulative frequency diagram for the data in the table.

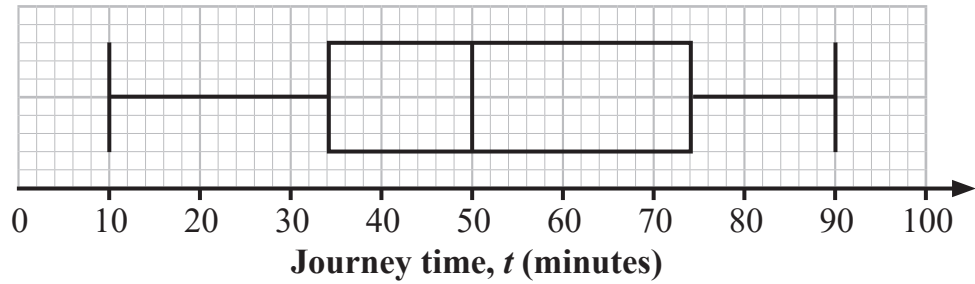


[3]

[Turn over



The journey times for those in group B are summarised in the box plot below.



(e) Complete the table below.

	Median (minutes)	Interquartile range (minutes)
Group A		
Group B		

[4]

(f) Using the information in the table in part (e), compare the journey times for the passengers in the two groups.

[4]





BLANK PAGE
DO NOT WRITE ON THIS PAGE
(Questions continue overleaf)

13832.06 R

[Turn over



32GST2225

8 Kate sometimes takes sandwiches to school for lunch.

The probability that she takes sandwiches to school for lunch on a particular day is 0.37

- (a) Write down the probability that Kate does not take sandwiches to school on a particular day.

Answer _____ [1]

Kate wishes to use the Binomial Distribution with $(p + q)^n$ as a model for the number of days she takes sandwiches to school for lunch during a 5-day school week.

- (b) For Kate's model, write down the values of n and p .

$n =$ _____ [1]

$p =$ _____ [1]



In part (c) you may use

$$(p + q)^5 = p^5 + 5p^4q + 10p^3q^2 + 10p^2q^3 + 5pq^4 + q^5$$

(c) During a 5-day school week, calculate the probability that Kate takes sandwiches to school for lunch:

(i) only once;

Answer _____ [2]

(ii) at least once.

Answer _____ [3]



- 9 The table below shows the crop area for wheat, in hectares, in Northern Ireland from 2016 to 2020 in June each year.

The associated chain base index numbers are shown.

The base year is 2015

Year	2016	2017	2018	2019	2020
Crop area for wheat	8616	8730	6845	8104	7132
Chain base index number	108.0	101.3	78.4	118.4	

- (a) Calculate the crop area in hectares for wheat in Northern Ireland in 2015

Show your working and give your answer to the nearest whole number.

Answer _____ hectares [3]

- (b) Calculate the chain base index number for the crop area for wheat in Northern Ireland in 2020

Show your working.

Answer _____ [2]



(c) Give an interpretation of your answer to part (b).

[3]

[Turn over

13832.06 R



32GST2229

10 Members of an all-female choir sing in either the soprano or alto section.

The ages, in years, of the members of the alto section of the choir are as follows:

42 23 34 18 22 55 20 36 27 43

(a) Show that the mean age of these members of the choir is 32 years.

[1]

(b) Calculate the standard deviation of the ages of these members of the choir.

Answer _____ [3]



For those in the soprano section of the choir, Jan calculates the mean age as 29.5 years and the standard deviation as 8.3 years.

One member of the choir is selected at random and her age is 28

- (c) Use standardised scores to decide if this person is more likely to belong to the soprano or alto section of the choir.

You may assume that the ages of members of the soprano and alto sections of the choir are both normally distributed.

[5]

- (d) Identify one limitation of your calculation in part (c).

[1]

THIS IS THE END OF THE QUESTION PAPER



DO NOT WRITE ON THIS PAGE

For Examiner's use only	
Question Number	Marks
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	

Total Marks	
--------------------	--

Examiner Number

--

Permission to reproduce all copyright material has been applied for.
In some cases, efforts to contact copyright holders may have been unsuccessful and CCEA will be happy to rectify any omissions of acknowledgement in future if notified.

13832.06 R



32GST2232