Name: _____

GCSE Statistics

Planning Investigations

Total marks available: 55

Total marks achieved: _____

Instructions
• Use black ink or ball-point pen.
• Fill in the boxes at the top of this page with your name, centre number and candidate number.
• Answer all questions.
• Answer the questions in the spaces provided – there may be more space than you need.
• Scientific calculators may be used.
• You must show all your working out with your answer clearly identified at the end of your solution.

Information
• The marks for each question are shown in brackets – use this as a guide as to how much time to spend on each question.

Advice
• Read each question carefully before you start to answer it.
• Try to answer every question.
• Check your answers if you have time at the end.

Q1.

Youssef is investigating the amount of time students spend on social media.

Here is what Youssef has written down for his hypothesis and how he plans to collect, process and present his data.

Hypothesis

- Do girls spend more time than boys on social media?

Data to collect

- Gender of student
- Time spent on social media in one week (using categories: 1–2 hrs, 3–4 hrs, 5–6 hrs)
- Name of student

Processing and presenting data

- Work out an estimate of the mean weekly time boys spend on social media
- Work out an estimate of the mean weekly time girls spend on social media
- Draw a scatter graph with gender on the horizontal axis and time spent on social media on the vertical axis
- Draw a time series graph showing how the amount of time spent on social media has changed over the years

Discuss whether Youssef's hypothesis and his plans for collecting, processing and presenting his data are appropriate.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(Total for question = 5 marks)

Q2.

Gary is going to investigate the amounts of time students spend watching TV.

He is going to write a plan for this investigation.

His hypothesis is

"The amount of time that boys spend watching TV is greater than the amount of time that girls spend watching TV".

Write down three other things he should include in his plan.

Explain why each of these things is appropriate.

You must refer to more than one stage of the statistical enquiry cycle.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(Total for question = 6 marks)

Q3.

Tariq wants to investigate the distance people travel to get to their holiday destination and the price of the holiday.

He is going to write a plan for this investigation.

Write down **one** thing that he should include in his plan for each of

- data collection
- presenting and interpreting diagrams
- presenting and interpreting calculations

Explain why each of your choices is appropriate.

You may use the blank space below to plan your answer and write your answer on the lines provided on the opposite page.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(Total for question = 6 marks)

Q4.

A vet wants to estimate the average weight of adult female shorthair cats.

She took a sample of 8 adult female shorthair cats that were brought to her clinic.

Here are the cats' weights in kilograms.

4.0 4.5 9.8 5.1 4.1 4.5 4.2 3.9

(a) Identify the outlier.

..... kg

(1)

The vet plans to estimate the average weight of adult female shorthair cats using this data.

(b) Give her advice.

You should consider the data she plans to use and which average she should choose.
Give reasons for your answer.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(5)

(Total for question = 6 marks)

Q5.

A database contains information about the 2567 people registered at a medical centre.

Here is the information that the database contains on each person.

Name	Date of Registration	Age	Phone Number	Address	Postcode	Number of Appointments

Rebecca wants to see if there is a relationship between the age of a person registered at the medical centre and the number of appointments this person has had at the medical centre.

Describe how Rebecca can plan to use technology to help her process and present the data for this investigation.

You should refer to the advantages of using technology to process and present data compared to doing so by hand.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(Total for question = 6 marks)

Q6.

Barbara found the following information about the average gross pay, in euros (€), for males and for females in ten countries in 2010

Country	Average gross pay (€)	
	Male	Female
Denmark	62 120	49 254
Germany	44 465	34 740
Ireland	48 459	19 177
Spain	29 009	25 101
France	37 627	30 406
Netherlands	47 373	36 696
Portugal	19 424	15 299
Sweden	41 311	33 305
United Kingdom	42 710	31 115
Switzerland	60 135	48 972
Mean	43 263	32 407
Product moment correlation coefficient	0.832	

(Source: *Eurostat*)

Barbara investigates this information using a suitable graph.

Using the information in the table and her graph, she reaches the following conclusion:

There is strong evidence that, for every €10 000 increase in gross pay for males, gross pay for females increases by €8000

(a) Explain, giving reasons, what Barbara is most likely to have done to reach this conclusion.

You should consider

- which type of graph she will have used and why
- how she will have used the graph and the information in the table

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(5)

(b) Explain why Barbara's conclusion, based on using the data in the table, may be unreliable.

.....

.....

.....

(1)

(Total for question = 6 marks)

Q7.

Anne is investigating the effect of caffeine on reaction times.

Here is what Anne has written down for her hypothesis and how she plans to collect, process and present her data.

Hypothesis

- Does caffeine reduce reaction times?

Data collection

- Pre-test the reaction time test
- Each person to complete the reaction time test, then drink some coffee, wait a while and then complete the reaction time test a second time
- Each person to record their reaction times in the spreadsheet

Processing and presenting data

- Look for outliers in the data and remove them if found
- Calculate the median and quartiles
- Draw a box plot for the reaction times after coffee has been drunk

Discuss whether Anne's hypothesis and her plans for collecting, processing and presenting data are appropriate.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(Total for question = 5 marks)

Q8.

Kate is investigating the lengths of earthworms.
She wants to find out if the lengths of common earthworms can be modelled by a normal distribution.
She has found out that the greatest length of the common earthworm is 36 cm.

(Source: *www.nationalgeographic.com*)

Kate plans to group the data that she will collect in the table shown below.

Length (l cm)	Frequency
$0 < l \leqslant 5$	
$5 < l \leqslant 10$	
$10 < l \leqslant 20$	
$20 < l \leqslant 30$	
$30 < l \leqslant 35$	
$35 < l \leqslant 40$	

(a) Comment on whether Kate's choice of class intervals in the grouped frequency table is appropriate.

.....

.....

.....

(2)

Bien is researching the lengths of time, in days, lived by flies.
He wants to know if the lengths of time lived by flies can be modelled by a normal distribution.

He has collected data for 80 flies.

The table gives information about the length of time lived by each fly.

Length of time lived (days)	Frequency
0–5	5
6–10	8
11–15	12
16–17	13
18–19	18
20–21	15
22–25	9

Here is Bien's plan to test whether the lengths of time lived by flies can be modelled by a normal distribution.

> Using the data collected I will draw a histogram. To do this I will need to calculate the frequency densities. Here is an example of a frequency density calculation for my data:
>
> Class interval 6–10
> Frequency 8
>
> $$\text{Frequency density} = \frac{8}{10-6} = 2$$
>
> I will then use the mean and standard deviation of the data to work out the amount of data within 1 standard deviation of the mean and the amount of data within 2 standard deviations of the mean.

(b) Comment on whether Bien's plan is appropriate.

.....

.....

.....

.....

.....

.....

.....

(4)

(Total for question = 6 marks)

Q9.

Mehmet wants to conduct an investigation about school meals at his school.
He wants to take a sample of students from his school.

(a) Write down a sampling frame he could use.

.....

.....

(1)

Mehmet has two hypotheses:

- Younger students have school meals, rather than packed lunches, more often than older students do.
- More female students than male students are vegetarian.

Here are sections A, B and C of his plan for his investigation.

A. Sampling method:

Take a sample of 10 students from each of the year groups at school.

B. Questions:

1. What is your gender? Please tick (✓)
   Male ☐ Female ☐
2. What is your year group? Please tick (✓)
   Year 7 ☐ Year 8 ☐ Year 9 ☐ Year 10 ☐ Year 11 ☐ Year 12 ☐ Year 13 ☐
3. What is your favourite meal?.....
4. Do you think eating meat is cruel to animals? Please tick (✓)
   Yes ☐ No ☐
5. How often do you eat school meals? Please tick (✓)
   Once ☐ Two times ☐ Three times ☐ Four times ☐ Five times ☐

C. Presenting data:

Draw a bar chart with age on horizontal axis and total number of school meals on vertical axis (to see if younger students have taller bars).

Draw a pictogram showing number of vegetarian females and number of vegetarian males.

(b) Name the sampling method that Mehmet plans to use.

.....

(1)

(c) Discuss whether or not Mehmet's plans are appropriate.
You must comment on each of the three sections A, B and C.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(5)

Mehmet also collects information about the amounts that the students pay for their school meals.

He sorts the costs of the meals into classes with equal widths of 5p.

The table gives some of his results.

Cost of meal	Number of meals
£1.50 – £1.54	2
£1.55 – £1.59	3
£1.60 – £1.64	0
£1.65 – £1.69	1
£1.70 – £1.74	3

Mehmet decides to use classes with equal widths of 10p instead of 5p.

(d) Write down one advantage and one disadvantage of doing this.

Advantage: .....

.....

Disadvantage: .....

.....

(2)

(Total for question = 9 marks)