

QQQ – Statistics Year 1 – Chapter 1 – Data Collection (Version 2)

Total Marks: 23

(23 = Platinum, 21 = Gold, 18 = Silver, 16 = Bronze)

1. Lenny is one of a team of people interviewing shoppers in a town centre.

He is asked to survey 50 women between the ages of 18 and 29

Identify the name of this type of sampling.

Circle your answer.

[1 mark]

- simple random
- stratified
- quota
- systematic

(Total 1 marks)

AQA June 2019 Paper 3 Q11

2. The mass of a bag of nuts produced by a company is known to have a mean of 40 grams and a standard deviation of 3 grams.

The company produces five different flavours of nuts.

The bags of nuts are packed in large boxes.

Given the information above, identify the continuous variable from the options below.

Tick (✓) **one** box.

[1 mark]

- The flavours of the bags of nuts
- The known standard deviation of the mass of a bag of nuts
- The mass of an individual bag of nuts
- The number of bags of nuts in a large box

(Total 1 marks)

AQA June 2023 Paper 2 Q12

3. The table below shows the total monthly rainfall (in mm) in England and Wales in a sample of six years.

The sample of six years was taken from a data set covering every year from 1768 to 2018.

	Jan	Feb	Mar	Apr	Ma y	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1768	109. 2	129. 1	12.8	85.6	46.1	148. 7	121. 9	91.6	136. 8	119. 4	142. 5	103. 6
1818	98.0	65.8	134. 7	135. 6	55.9	31.2	50.4	21.0	115. 6	75.8	112. 0	46.8
1868	99.9	62.2	71.1	61.4	36.7	16.5	20.0	106. 7	90.2	95.6	61.4	185. 6
1918	91.2	61.6	36.7	63.3	58.5	30.9	110. 0	62.9	189. 5	69.1	66.3	122. 5
1968	85.8	47.6	59.5	68.8	78.7	94.0	107. 8	72.2	148. 1	99.0	69.6	84.2
2018	104. 5	52.8	115.1	91.4	51.9	16.5	39.6	76.7	67.0	75.8	104. 9	116. 0

Deduce the sampling method **most likely** to have been used to collect this sample.

Circle your answer.

[1 mark]

- Opportunity
- Simple Random
- Stratified
- Systematic

(Total 1 marks)

4. Jo is investigating the popularity of a certain band amongst students at her school. She decides to survey a sample of 100 students.

- (i) State an advantage of using a stratified sample rather than a simple random sample. [1]
- (ii) Explain whether it would be reasonable for Jo to use her results to draw conclusions about all students in the UK. [1]

(Total 2 marks)

OCR A June 2018 Paper 1 Q9

5. Before introducing a new rule the secretary of a golf club decided to find out how members might react to this rule.

- (a) Explain why the secretary decided to take a random sample of club members rather than ask all the members. (1)
- (b) Suggest a suitable sampling frame. (1)
- (c) Identify the sampling units. (1)

(Total 3 marks)

Edexcel S2 June 2006 Q1

6. Keith is studying the variable Daily Mean Wind Direction, in degrees, from the large data set.

Keith summarised the data for Camborne from 1987 into 4 directions A , B , C and D representing North, South, East and West in some order.

Direction	A	B	C	D
Frequency	22	48	56	58

- (a) Using your knowledge of the large data set state, giving a reason, which direction A represents. (1)

The entry for Hurn on 27th September 1987 was 999

- (b) State, giving a reason, what Keith should do with this value. (2)

(Total 3 marks)

Edexcel June 2024 Statistics Q2

7. The manager of a factory wants to introduce a bonus scheme.

The factory has 65 employees who work in production and 28 employees who work in the office.

The manager decides to collect the opinions of a sample of these 93 employees.

(a) Explain how the manager could collect a simple random sample of 20 employees.

[3 marks]

(b) The manager collected a simple random sample of 20 employees.

The manager noticed that all 20 of the employees in the sample worked in production and therefore the sample was not representative.

State a different method of sampling that would give a representative sample. **[1 mark]**

(Total 4 marks)

AQA June 2023 Paper 2 Q14

8. A teacher is investigating how pupils travel to and from school each day. Pupils can either travel by bus, train, car, bicycle or walk.

The teacher decides to collect a sample of size 60 for the investigation.

(a) The teacher lives in a village 10 miles away from the school.

Explain how collecting a sample which just consists of pupils who live in the same village as the teacher might introduce bias. **[1]**

The table below shows how many students there are in each year.

Year 7	Year 8	Year 9	Year 10	Year 11
86	105	107	101	101

(b) The teacher decides to use the method of proportional stratified sampling.

Calculate the number of pupils in the sample who are in Year 9. **[2]**

The teacher generates a sample of 10 pupils from the 86 in Year 7 by listing them in alphabetical order and selecting the first name on the list and every ninth name thereafter.

(c) Explain whether this method will generate a simple random sample of the pupils who travel in Year 7. **[1]**

(Total 4 marks)

OCR MEI June 2024 Paper 2 Q9

9. Describe one advantage and one disadvantage of

(a) quota sampling, **(2)**

(b) simple random sampling. **(2)**

(Total 4 marks)

Solutions

1.

Scheme	Marks	AO
Circles correct answer	B1	1.2
	(1 mark)	

Typical solution

Quota

2.

Scheme	Marks	AO
Ticks the correct box	B1	1.2
	(1 mark)	

Typical solution

The mass of an individual bag of nuts

3.

Scheme	Marks	AO
Circles the correct answer	B1	2.2b
	(1 mark)	

Typical solution

Systematic

4.

Mark scheme (i)

Scheme	Marks	AO
Any mention of diff categories or types, eg Includes students in all years (or classes)	E1	2.4

More representative of diff ages Prevents obtaining too many in one year gp		
	[1]	

Notes

or any mention of proportions, eg

Uses the right proportions of students

Ignore all else

NOT eg:

Wider variety results

More representative of pop

Mark scheme (ii)

Scheme	Marks	AO
Must include <u>reason</u> why not rep've, eg Her school may be biased Students' friends may be in local band Opinions at one school not indep of one another Ignore all else	E1	2.3
	[1]	

Notes

except allow Small sample or Only 100 students in sample

Student opinion elsewhere may differ

Diff conditions, or diff types of school, in diff areas,

Diff bands liked in diff areas

NOT: Not representative of UK

"No" may be implied

5.

Mark scheme (a)

Scheme	Marks
Saves time / cheaper / easier or	B1

<u>A census/asking all members</u> takes a long time or is expensive or difficult to carry out	
	(1)

Notes

B1 any one

Mark scheme (b)

Scheme	Marks
<u>List, register or database of all club members/golfers</u> or <u>Full membership list</u>	B1
	(1)

Mark scheme (c)

Scheme	Marks
Club <u>member(s)</u> .	B1
	(1)
	(3 marks)

6.

Mark scheme (a)

Scheme	Marks	AO
East Since : The prevailing winds in Camborne/UK are from the South or West or North <u>or</u> East since lowest frequency <u>or</u> Camborne is in South/West (and so wind is less likely to come from East)	B1	2.2b
	(1)	

Notes

B1 for East **and** a suitable reason based on: general winds in UK / Camborne
allow East and Camborne indicated on a sketch of map
contradictory comment e.g. 'Camborne is in the North, so wind is less likely to come from the East' is B0

Mark scheme (b)

Scheme	Marks	AO
e.g. 'wind direction should be in the range 0 to <u>360</u> '	M1	2.4/1.2
e.g. 'so he should <u>ignore</u> / <u>remove</u> the value'	A1	1.1b
	(2)	
	(3 marks)	

Notes

M1 for giving a reason which correctly uses 360
for reason allow e.g. 'only goes up to 360' also condone '10 to 360'
this mark may be implied by e.g. 'subtract 720'

A1 for stating appropriate action to take provided M1 clearly scored
ignore o.e. or remove o.e. allow e.g. 'reject', 'discard', etc.
must be suggesting what to do with the value, not just stating that it is an anomaly
allow change it to 279

7.

Mark scheme (a)

Scheme	Marks	AO
Sets up the numbering of the population (allow 00 to 92) Allow 01 to 93	E1	2.4
Explains how the calculator (random number generator) will be used to generate random numbers (two-digit PI by between 1 and 93 or in this range or ignore numbers out of range or using these numbers) If using random number tables must state 2-digit numbers OE	E1	2.4
Explains how to deal with repeats (PI) and identify 20 different employees/numbers	E1	2.4

	(3)	
--	------------	--

Typical solution

Number the employees from 1 to 93

Choose random numbers from the calculator in this range

Ignore repeats and continue until 20 different employees have been selected

Mark scheme (b)

Scheme	Marks	AO
States either Stratified or Quota	B1	1.2
	(1)	
	(4 marks)	

Typical solution

Stratified sampling

8.

Mark scheme (a)

Scheme	Marks	AO
(probably) wouldn't include any pupils who eg cycle or eg walk to school (and hence biased towards certain methods of transport) oe	B1	2.2b
	[1]	

Notes

B1: must refer to at least one of the given modes of transport

Mark scheme (b)

Scheme	Marks	AO
$\frac{107}{500} \times 60$	M1	1.1
13	A1	1.1
	[2]	

Notes

M1: allow slip in calculation of 500 if clearly their sum of all pupils; may be implied by 12.84 or 12.8

A1: B2 for 13 unsupported

Mark scheme (c)

Scheme	Marks	AO
not simple random sampling because every possible sample does not have an equal probability of being selected oe	B1	2.4
	[1]	

Notes

B1: allow

eg **no** because not possible to select every possible **sample**

eg **not** simple random sampling because every pupil not equally likely to be selected

do **not** allow

eg **not** simple random sampling because it's systematic sampling

9.

Mark scheme (a)

Scheme	Marks
<u>Advantages:</u> – does not require the existence of a population list (sampling frame) – <u>field work can be done quickly</u> as representative sample can be achieved with a small sample size – costs kept to a minimum (<u>cheaply</u>) – administration relatively <u>easy</u> – non-response not an issue any one	B1
<u>Disadvantages:</u> – not possible to estimate sampling errors – interviewer choice and may not be able to judge easily / <u>may lead to bias</u> – non-response not recorded – non-random process any one	B1
	(2)

Mark scheme (b)

Scheme	Marks
<u>Advantages:</u> – <u>random process</u> so possible to <u>estimate sampling errors</u> – free from <u>bias</u> any one	B1
<u>Disadvantages:</u> – not suitable when sample size is large – <u>sampling frame required</u> which <u>may not exist</u> or may be difficult to construct for a large population. any one	B1
	(2)
	(4 marks)

Notes

NO REPETITION / OPPOSITES