



S07-30min Collecting Data: 30-minute Pack

EDEXCEL 1ST0 · Calculator allowed · about 30 minutes

Total Marks

Name: _____ Date: ____ / ____ / ____

Answer ALL questions. Show all your working.

- 1** A student measures the wingspan, in mm, of 30 butterflies for a project. State whether this data is discrete or continuous.

(Total for Question 1 is 1 mark)

- 2** For each statement about collecting data, write down whether it is True or False.

(i) A pilot survey is a small-scale trial run of a survey, carried out before the main survey takes place. (1)

(ii) Data collection sheet class intervals such as '10 - 20' and '20 - 30' are well designed, because they give the respondent more choice. (1)

(iii) Data collected by a researcher carrying out her own experiment is called secondary data. (1)

(Total for Question 2 is 3 marks)

- 3** Write down whether the following statement is True or False.
A response bias occurs when the wording of a question influences the answer a respondent gives.

(Total for Question 3 is 1 mark)

- 4 A survey question reads: 'Don't you agree that the new park is a great improvement?' Give one reason why this is a leading question.

(Total for Question 4 is 2 marks)

- 5 A tally chart records the number of goals scored by a football team in each of 20 matches: 4 matches with 0 goals, 7 matches with 1 goal, 6 matches with 2 goals, and 3 matches with 3 goals. Work out the total number of goals scored across all 20 matches.

(Total for Question 5 is 2 marks)

- 6 A researcher wants to record, for each pupil in a sample, both their year group (Year 9, Year 10 or Year 11) and how they usually travel to school (Walk, Cycle or Bus). Design a two-way data collection sheet that could be used to record this information for a sample of pupils.

(Total for Question 6 is 3 marks)

- 7 A market researcher wants to survey a sample of 200 shoppers about their spending habits before running a full survey of 5000 shoppers. Before starting, she runs a smaller trial with 25 shoppers.
- (a) Name this smaller trial.
 - (b) Give two reasons why it is useful to carry this out before the main survey.

(Total for Question 7 is 3 marks)

- 8** A student is recording how long, in minutes, 40 people take to travel to work, using the data collection sheet below. Two features of this data collection sheet make it unsuitable.

Time taken (minutes)	Tally	Frequency
0 - 10		
10 - 20		
20 - 30		

- (a) Explain two reasons why the class intervals shown are not suitable for recording this data. (2)
- (b) Rewrite the class intervals so that the data collection sheet is suitable, using inequality notation. (2)

(Total for Question 8 is 4 marks)

-
- 9** A gardener wants to find out how the temperature inside a greenhouse changes during the day. She records the temperature every hour from 8 am to 6 pm.

- (a) Give one advantage of using an electronic data logger, rather than a person taking readings by hand, to record the temperature. (1)
- (b) Design a data collection sheet the gardener could use to record the readings by hand if a data logger were not available. Include suitable column headings and enough rows for all the readings described above. (2)
- (c) The gardener wants to see how the temperature changes during the day. Suggest a suitable way to display this data once it has been collected, and give a reason for your choice. (1)

(Total for Question 9 is 4 marks)

- 10** A student wants to investigate whether pupils who walk to school arrive earlier than pupils who travel by car, using data from 60 pupils at her school. She plans to collect the data using a questionnaire given to a sample of 60 pupils chosen from a list, numbered 1 to 900, using every 15th pupil on the list.
- (a) State the name of the sampling method used.
 - (b) Give one reason why a questionnaire, rather than direct observation, might not give reliable data for this investigation.

(Total for Question 10 is 3 marks)

- 11** A researcher wants to test the hypothesis: 'People who commute by bicycle sleep for longer, on average, than people who commute by car.' She plans to give a questionnaire to 40 cyclists and 40 car drivers, asking each how many hours they sleep on a typical night.
- (a) Write down one question, with a set of response options, that could be used to collect this data.
 - (b) Give one way the results could be used to test the hypothesis.

(Total for Question 11 is 3 marks)

Mark scheme · S07-30min Collecting Data: 30-minute Pack

Question 1

- **B1** Continuous cao
- **Answer: Continuous**

Question 2

- (i) **B1** True cao
- **(i) Answer: True**
- (ii) **B1** False cao
- **(ii) Answer: False**
- (iii) **B1** False cao
- **(iii) Answer: False**

Question 3

- **B1** True cao
- **Answer: True**

Question 4

- **B1** identifies that the question suggests or prompts a particular answer
- **B1** correct supporting detail, e.g. the phrase 'don't you agree' pressures the respondent to say yes oe
- **Answer: It is leading because the wording ("don't you agree") pushes the respondent towards answering yes, rather than giving their own honest opinion.**

Question 5

- **M1** correct method, $(4 \times 0) + (7 \times 1) + (6 \times 2) + (3 \times 3)$
- **A1** 28 cao
- **Answer: 28**

Question 6

- **B1** a two-way grid drawn with year groups as the headings for one set of rows/columns and methods of travel as the headings for the other
- **B1** cells provided for every combination of year group and method of travel, ready for tally marks
- **B1** row and column totals included, clearly labelled
- **Answer: A two-way table with Year 9/10/11 as rows, Walk/Cycle/Bus as columns, a tally/frequency cell for each combination, and a Total row and column.**

Question 7

- **B1** Pilot survey (pilot study) cao
- **B1** first valid reason, e.g. it checks that questions are clear and easy to understand oe
- **B1** second valid reason, e.g. it can reveal how long the survey takes, or whether the response categories are suitable oe
- **Answer: (a) Pilot survey. (b) It checks that questions are clear and not ambiguous, and it can show whether the response categories or timing are suitable before the full survey is run.**

Question 8

- (a) **B1** the intervals overlap (e.g. a time of 10 minutes could be tallied in either '0 - 10' or '10 - 20'), so it is not clear where to record some values
- (a) **B1** there is no interval for times of 30 minutes or more, so some people's travel times could not be recorded at all
- **(a) Answer: The intervals overlap at their boundaries, and there is no interval to cover times of 30 minutes or more.**
- (b) **B1** non-overlapping intervals using correct inequality notation, e.g. $0 \leq t < 10$, $10 \leq t < 20$, $20 \leq t < 30$
- (b) **B1** an interval included to cover all remaining times, e.g. $t \geq 30$ (or '30 and over')
- **(b) Answer: e.g. $0 \leq t < 10$, $10 \leq t < 20$, $20 \leq t < 30$, $t \geq 30$ (t = time in minutes).**

Question 9

- (a) **B1** e.g. a data logger can take readings automatically at exact times without anyone needing to be present, reducing the risk of missed or inaccurate readings
- (a) **Answer: It records readings automatically and accurately without needing a person present at every hour.**
- (b) **B1** columns for Time and Temperature clearly headed, with units (deg C) shown
- (b) **B1** 11 rows provided, one for each hour from 8 am to 6 pm inclusive
- (b) **Answer: Table with columns Time | Temperature (deg C), and 11 rows for 8 am, 9 am, 10 am, 11 am, 12 pm, 1 pm, 2 pm, 3 pm, 4 pm, 5 pm, 6 pm.**
- (c) **B1** a line graph/time series graph, because it shows the trend/change in temperature over time
- (c) **Answer: A line graph (time series graph), because it clearly shows how temperature changes over time.**

Question 10

- **B1** Systematic sampling
- **B1** valid reason a questionnaire may be unreliable, e.g. pupils may not remember their exact arrival time accurately
- **B1** correct supporting detail linking the reason to this investigation, e.g. this could make the reported times inaccurate compared with observing pupils directly
- **Answer: (a) Systematic sampling. (b) Pupils may not remember or report their exact arrival time accurately, so questionnaire answers could be less accurate than direct observation.**

Question 11

- **B1** a closed question about hours of sleep with non-overlapping response options, e.g. 'How many hours do you usually sleep? Tick one: fewer than 6 / 6 to 8 / more than 8'
- **B1** options do not overlap and cover all realistic values
- **B1** valid way to test the hypothesis, e.g. compare the mean number of hours slept by cyclists with the mean for car drivers
- **Answer: (a) e.g. 'How many hours do you usually sleep on a school night? Tick one: fewer than 6 / 6 to 8 / more than 8'. (b) Compare the mean (or median) sleep time of the cyclist group with the car-driver group to see which is greater.**