



S07-test Collecting Data: Test Pack

EDEXCEL 1ST0 · Calculator allowed · about 25 minutes

Total Marks

Name: _____ Date: ____ / ____ / ____

Answer ALL questions. Show all your working.

- 1** For each variable below, write down whether it is best described as categorical, discrete or continuous data.

(a) The eye colour of a group of pupils. (1)

.....

(b) The number of pets owned by a family. (1)

.....

(c) The height, in cm, of a sunflower plant. (1)

.....

(d) The number of goals scored by a team in a match. (1)

.....

(Total for Question 1 is 4 marks)

-
- 2** State whether the height of a plant, measured in cm, is discrete or continuous data.

(Total for Question 2 is 1 mark)

3 A student wants to find out how pupils in her class travel to school on a normal school day.

(a) Suggest four suitable, non-overlapping categories she could use on a data collection sheet, making sure every pupil's response can be recorded. (2)

(b) Design the data collection sheet, using your four categories from part (a), with columns for tally and frequency. (2)

(Total for Question 3 is 4 marks)

4 A data collection sheet for recording vehicle colours passing a school has the categories: Red, Blue, Black, White. Give one reason why this data collection sheet is not suitable.

(Total for Question 4 is 2 marks)

5 A researcher wants to test a new 20-question survey before using it on the full sample of 500 people. She gives the survey to 15 people first, to check the questions are clear. Name this stage of the survey process.

(Total for Question 5 is 1 mark)

6 Before carrying out a full survey of 200 shoppers about their spending habits, a market researcher plans to carry out a pilot survey of 15 shoppers first.

- (a)** Give one reason why the researcher should carry out a pilot survey before the full survey. **(1)**
- (b)** State one thing the researcher should check for during the pilot survey. **(1)**
- (c)** One question on the pilot survey reads: 'Don't you agree that you spend too much money on impulse purchases?' Identify one flaw in this question, and rewrite it as a better, unbiased question. **(2)**

(Total for Question 6 is 4 marks)

7 A data collection sheet is used to record the number of text messages sent by 45 pupils in one day. The tally totals recorded were: 0-9 messages: 11 pupils, 10-19 messages: 18 pupils, 20-29 messages: 9 pupils, and the rest sent 30 or more messages. Work out the number of pupils who sent 30 or more messages, and state whether the class intervals used are suitable for continuous or discrete data.

(Total for Question 7 is 3 marks)

8 A newspaper wants to find out what proportion of adults in a city support a new cycle lane. A reporter plans to stand outside a cycling shop on a Saturday morning and ask the first 50 customers who leave the shop whether they 'Support' or 'Do not support' the cycle lane.

- (a) Identify one way in which this data collection method could produce biased results, and explain why. (2)

- (b) Suggest one improvement the reporter could make to the data collection method to reduce this bias. (2)

- (c) Suggest one additional response category that should be added to the data collection sheet, and explain why it is needed. (2)

(Total for Question 8 is 6 marks)

Mark scheme · S07-test Collecting Data: Test Pack

Question 1

- (a) **B1** categorical cao
- (a) **Answer: Categorical**
- (b) **B1** discrete cao
- (b) **Answer: Discrete**
- (c) **B1** continuous cao
- (c) **Answer: Continuous**
- (d) **B1** discrete cao
- (d) **Answer: Discrete**

Question 2

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

Question 3

- (a) **B1** three sensible, mutually exclusive methods of travel given (e.g. Walk, Cycle, Car, Bus)
- (a) **B1** an 'Other' or similar catch-all category included, so every response can be recorded
- (a) **Answer: e.g. Walk, Cycle, Bus, Other (accept any 3 sensible non-overlapping methods plus a catch-all category).**
- (b) **B1** categories from (a) listed in one column, ft
- (b) **B1** tally and frequency columns both included and clearly headed
- (b) **Answer: Table with columns Method of travel | Tally | Frequency, using the four categories from (a).**

Question 4

- **B1** identifies the sheet is missing an 'Other' category
- **B1** correct supporting detail, e.g. a vehicle that is silver or green could not be recorded
- **Answer: It has no 'Other' category, so a vehicle of a colour not listed (e.g. silver or green) could not be recorded.**

Question 5

- **B1** Pilot survey (pilot study) cao
- **Answer: Pilot survey**

Question 6

- (a) **B1** e.g. to check that the questions are clear and easy to understand, and that the data collection sheet works well, so problems can be corrected before the full survey
- (a) **Answer: To check the questions and data collection sheet work well, so any problems can be fixed before the full survey.**
- (b) **B1** e.g. whether the questions are easy to understand / whether there are enough response categories / whether the survey takes a sensible amount of time to complete
- (b) **Answer: Whether the questions are clear and easy for shoppers to understand.**
- (c) **B1** identifies that the question is leading/biased, as it pressures the respondent to agree
- (c) **B1** gives an improved, neutral question with a clear response format, e.g. 'How much money, if any, do you spend on impulse purchases in a typical week? Please tick a box: £0, £1-£10, £11-£25, more than £25'
- (c) **Answer: The original question is leading; an improved version asks neutrally, e.g. 'How much do you spend on impulse purchases each week?' with response boxes.**

Question 7

- **M1** correct method, 45 minus (11+18+9)
- **A1** 7 cao
- **B1** Discrete cao, with valid reason, e.g. the number of messages can only take whole-number values
- **Answer: 7 pupils sent 30 or more messages. The class intervals are for discrete data, since the number of text messages can only be a whole number.**

Question 8

- (a) **B1** identifies that only customers of a cycling shop are surveyed
- (a) **B1** explains that these customers are more likely to already support cycling infrastructure than adults in the city generally, so the results would not be representative
- **(a) Answer: Only cycling-shop customers are surveyed, who are more likely to support a cycle lane than adults generally, so results would be biased.**
- (b) **B1** a valid improvement given, e.g. survey people at several different locations and times across the city (not only outside one shop)
- (b) **B1** brief justification that this would give a more representative range of adults' opinions
- **(b) Answer: Survey people at several different locations and times across the city, giving a more representative sample of adults.**
- (c) **B1** an 'Undecided/No opinion' (or similar) category suggested
- (c) **B1** explanation: because not every respondent will have a definite view either way, and forcing a choice would misrepresent their actual opinion
- **(c) Answer: Add an 'Undecided / No opinion' category, since some people will not have a strong view either way.**