



## S04-1hour Types of Data: 1-hour Pack

EDEXCEL 1ST0 · Calculator allowed · about 60 minutes

Total Marks

Name: \_\_\_\_\_ Date: \_\_\_\_ / \_\_\_\_ / \_\_\_\_

**Answer ALL questions. Show all your working.**

**1** Data can also be classified by where it comes from: primary data or secondary data.

(a) Define what is meant by primary data. (1)

(b) Define what is meant by secondary data. (1)

(c) For each scenario, write down whether primary or secondary data is being used. (4)

| Scenario                                                                              | Primary or secondary |
|---------------------------------------------------------------------------------------|----------------------|
| A student times her own class doing a memory test.                                    |                      |
| A researcher uses population figures published by the Office for National Statistics. |                      |
| A shop owner counts how many customers enter her shop each day.                       |                      |
| A student uses last year's weather records from the Met Office website.               |                      |

.....  
(Total for Question 1 is 6 marks)

**2** State whether this is True or False: Continuous data can only take whole-number values.

(Total for Question 2 is 1 mark)

**3** For each variable, write down whether it is qualitative or quantitative.

- (a) Eye colour.
- (b) The time taken to bake a cake, in minutes.

(Total for Question 3 is 2 marks)

---

**4** State whether each of the following uses primary or secondary data.

- (a) A student stands outside her house and counts the cars that pass, herself.
- (b) A researcher uses population figures published by the Office for National Statistics.

(Total for Question 4 is 2 marks)

---

**5** Complete each sentence with the correct word.

- (a) Categorical data whose categories have no natural order is called \_\_\_\_\_ data.
- (b) Categorical data whose categories have a natural order or ranking is called \_\_\_\_\_ data.

(Total for Question 5 is 2 marks)

**6** Categorical (qualitative) data can be further divided into nominal data and ordinal data.

**(a)** Define what is meant by nominal data. (1)

**(b)** Define what is meant by ordinal data. (1)

**(c)** For each variable, write down whether it is nominal or ordinal. (5)

| Variable                                 | Nominal or ordinal |
|------------------------------------------|--------------------|
| Eye colour                               |                    |
| T-shirt size (S, M, L, XL)               |                    |
| Exam grade achieved (9, 8, 7, ..., 1, U) |                    |
| Postcode                                 |                    |
| Favourite type of music                  |                    |

**(d)** Explain why exam grade is an example of ordinal data rather than nominal data. (1)

**(Total for Question 6 is 8 marks)**

**7** A geography student wants to compare rainfall this year with rainfall recorded at the same weather station 50 years ago.

**(a)** State whether the student needs primary or secondary data to find the rainfall figures from 50 years ago. (1)

**(b)** Give two reasons to justify your answer to part (a). (2)

**(Total for Question 7 is 3 marks)**

**8** Data can be recorded as raw (ungrouped) data, or it can be organised into grouped data.

**(a)** Define what is meant by raw (ungrouped) data. (1)

**(b)** Define what is meant by grouped data. (1)

**(c)** The raw ages, in years, of 15 people at a gym class are shown below. (3)

23, 45, 31, 29, 52, 38, 41, 27, 33, 49, 36, 44, 28, 55, 32

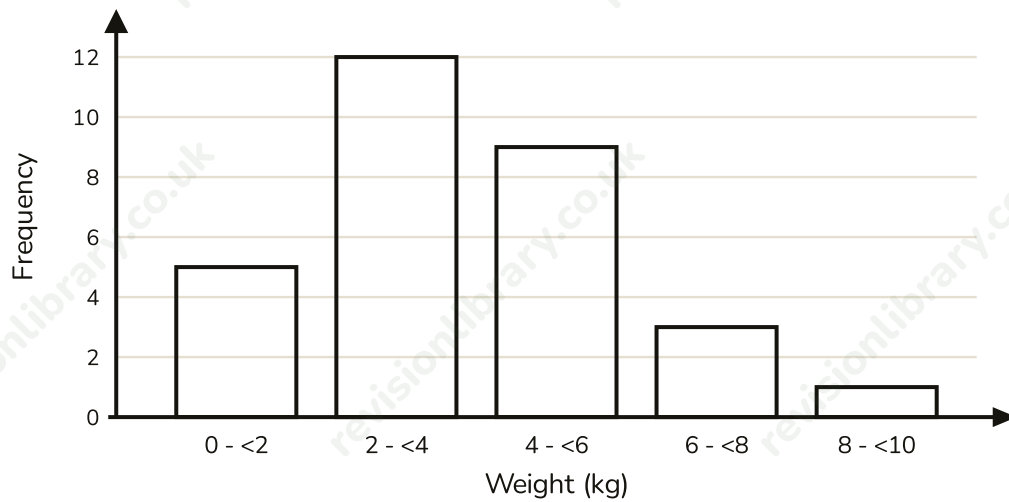
Complete the grouped frequency table.

| Age (years) | Frequency |
|-------------|-----------|
| 20 - 29     |           |
| 30 - 39     |           |
| 40 - 49     |           |
| 50 - 59     |           |

**(d)** Give one advantage and one disadvantage of grouping this data. (2)

**(Total for Question 8 is 7 marks)**

- 9** A courier company grouped the weights of 30 parcels delivered in one day. The bar chart shows the results.



- (a) Use the bar chart to write down the number of parcels with weight in the interval 4 - <6 kg. (1)
- .....
- (b) Explain why the horizontal axis is labelled using class intervals (such as 2 - <4 kg) rather than the exact weight of each parcel. (2)

(Total for Question 9 is 3 marks)

- 10** A student is carrying out a statistical investigation into the question: 'Do students who spend more time revising get higher test scores?' She plans to follow the statistical enquiry cycle: plan, collect, process, interpret.
- (a) At the plan stage, state two variables the student should record for each student in her sample. (2)
- (b) Classify the data type of each variable named in part (a). (2)
- (c) At the collect stage, state whether the student should use primary or secondary data, and give a reason for your answer. (2)
- (d) At the interpret stage, the student's sample only includes students from one class in her year group. Explain one way this could limit the conclusions she can draw. (1)

(Total for Question 10 is 7 marks)

- 11** The table shows 8 variables collected in a school survey about travel to school and lifestyle. For each variable, write down whether it is qualitative, discrete quantitative, or continuous quantitative.

| Variable                                 | Data type |
|------------------------------------------|-----------|
| Mode of transport to school              |           |
| Distance travelled to school (km)        |           |
| Number of school bags carried            |           |
| Time taken to travel to school (minutes) |           |
| T-shirt size worn (S, M, L, XL)          |           |
| Favourite school subject                 |           |
| Height of each pupil (cm)                |           |
| Number of days absent this term          |           |

(Total for Question 11 is 8 marks)

**12** A supermarket manager wants to decide whether to start stocking a new brand of oat milk. She is considering two options: (1) survey her own customers in-store, or (2) look at national sales reports for oat milk published by a market research company.

- (a) Which option, (1) or (2), would produce primary data? Give a reason. (2)
- (b) Give one advantage of the manager using option (2) instead of option (1). (1)
- (c) Give one disadvantage of the manager using option (2) instead of option (1). (1)
- (d) State which option you would recommend, with a reason. (1)

(Total for Question 12 is 5 marks)

---

**13** In a survey, students are asked which of three school clubs they attend. Their answers are recorded using codes: 1 = Drama club, 2 = Sports club, 3 = Art club.

- (a) Explain why this data is qualitative, even though it has been recorded using numbers. (2)
- (b) Is this an example of nominal or ordinal data? Give a reason. (1)

(Total for Question 13 is 3 marks)

# Mark scheme · S04-1hour Types of Data: 1-hour Pack

---

## Question 1

- (a) **B1** data collected first-hand by the person or organisation using it, for their own specific purpose oe
- (a) **Answer: Primary data is collected first-hand by the person or organisation that will use it, for their own specific purpose.**
- (b) **B1** data that already exists, collected by someone else, usually for a different purpose oe
- (b) **Answer: Secondary data already exists and was collected by someone else, usually for a different purpose.**
- (c) **B1** Timing her own class: primary
- (c) **B1** Using ONS population figures: secondary
- (c) **B1** Counting her own customers: primary
- (c) **B1** Using Met Office records: secondary
- (c) **Answer: Timing her own class - primary; ONS population figures - secondary; Counting her own customers - primary; Met Office records - secondary.**

## Question 2

- **B1** False cao
- **Answer: False**

## Question 3

- **B1** (a) Qualitative cao
- **B1** (b) Quantitative cao
- **Answer: (a) Qualitative (b) Quantitative**

## Question 4

- **B1** (a) Primary cao
- **B1** (b) Secondary cao
- **Answer: (a) Primary (b) Secondary**

## Question 5

- **B1** (a) nominal
- **B1** (b) ordinal
- **Answer: (a) nominal (b) ordinal**

## Question 6

- (a) **B1** categorical data with no natural order between the categories oe
- (a) **Answer: Nominal data is categorical data whose categories have no natural order (for example, eye colour).**
- (b) **B1** categorical data that has a natural order or ranking between the categories oe
- (b) **Answer: Ordinal data is categorical data whose categories have a natural order or ranking (for example, T-shirt size).**
- (c) **B1** Eye colour: nominal
- (c) **B1** T-shirt size: ordinal
- (c) **B1** Exam grade achieved: ordinal
- (c) **B1** Postcode: nominal
- (c) **B1** Favourite type of music: nominal
- (c) **Answer: Eye colour - nominal; T-shirt size - ordinal; Exam grade achieved - ordinal; Postcode - nominal; Favourite type of music - nominal.**
- (d) **B1** the grades have a natural order or ranking (a grade 9 is higher than a grade 8, and so on), unlike a nominal variable such as eye colour where no category is 'higher' than another oe
- (d) **Answer: Exam grades have a clear order or ranking (grade 9 is higher than grade 8, which is higher than grade 7, and so on), whereas nominal categories such as eye colour have no such order.**

### Question 7

- (a) **B1** Secondary data
- **(a) Answer: Secondary data.**
- (b) **B1** it is impossible to measure rainfall from 50 years ago herself, since that time has already passed, oe
- (b) **B1** reliable historical rainfall records already exist (e.g. from the Met Office), so there is no need to collect it again, oe
- **(b) Answer: She cannot measure rainfall from 50 years ago herself, since that time has passed, and reliable historical records already exist (e.g. from the Met Office) for her to use instead.**

### Question 8

- (a) **B1** individual data values listed separately, not organised into class intervals or groups oe
- **(a) Answer: Raw (ungrouped) data is a list of individual data values, with each one recorded separately.**
- (b) **B1** data organised into class intervals (groups of values), usually shown with a frequency for each interval oe
- **(b) Answer: Grouped data is data that has been organised into class intervals, usually shown with a frequency for each interval.**
- (c) **B1**  $20 - 29 = 4$  and  $30 - 39 = 5$ , both correct
- (c) **B1**  $40 - 49 = 4$
- (c) **B1**  $50 - 59 = 2$ , with all four frequencies summing to 15
- **(c) Answer: 20 - 29: 4, 30 - 39: 5, 40 - 49: 4, 50 - 59: 2**
- (d) **B1** advantage: e.g. it is easier to see the overall pattern, and simpler to display in a table or chart oe
- (d) **B1** disadvantage: e.g. the original, exact individual ages are lost, so the exact mean or exact range can no longer be found oe
- **(d) Answer: Advantage: it is easier to see the overall pattern of ages and to display the data in a chart. Disadvantage: the exact individual ages are lost, so an exact mean cannot be calculated from the grouped table alone.**

### Question 9

- (a) **B1** 9 cao
- **(a) Answer: 9**
- (b) **B1** weight (mass) is continuous data, so parcels can take any value and are very unlikely to share exactly the same weight oe
- (b) **B1** grouping into class intervals lets the data be organised and displayed sensibly, rather than having up to 30 separate individual values oe
- **(b) Answer: Weight is continuous data, so almost every parcel could have a slightly different exact weight. Grouping the weights into class intervals allows the data to be organised and displayed sensibly, instead of showing up to 30 separate individual values.**

### Question 10

- (a) **B1** time spent revising (e.g. in hours)
- (a) **B1** test score (e.g. out of 100)
- **(a) Answer: Time spent revising, and test score.**
- (b) **B1** time spent revising: continuous quantitative data
- (b) **B1** test score: discrete quantitative data (assuming whole-number marks) oe
- **(b) Answer: Time spent revising is continuous quantitative data; test score is discrete quantitative data.**
- (c) **B1** primary
- (c) **B1** because she needs data specific to her own sample of students and no existing data set matches her exact question oe
- **(c) Answer: Primary data - she needs to collect data specific to her own sample of students, and no existing data set matches her exact research question.**
- (d) **B1** the sample may not be representative of all students (e.g. in other classes or schools), so the conclusion may not generalise oe
- **(d) Answer: A sample from only one class may not be representative of all students, so her conclusion might not apply more widely.**

### Question 11

- **B1** Mode of transport to school: qualitative
- **B1** Distance travelled to school (km): continuous quantitative
- **B1** Number of school bags carried: discrete quantitative
- **B1** Time taken to travel to school (minutes): continuous quantitative
- **B1** T-shirt size worn: qualitative
- **B1** Favourite school subject: qualitative
- **B1** Height of each pupil (cm): continuous quantitative
- **B1** Number of days absent this term: discrete quantitative
- **Answer: Mode of transport - qualitative; Distance travelled - continuous quantitative; Number of school bags - discrete quantitative; Time taken to travel - continuous quantitative; T-shirt size worn - qualitative; Favourite school subject - qualitative; Height of each pupil - continuous quantitative; Number of days absent - discrete quantitative.**

### Question 12

- (a) **B1** (1)
- (a) **B1** because the manager would be collecting the data herself, first-hand, for her own specific purpose
- **(a) Answer: Option (1), because the manager collects the data herself, first-hand, for her own specific purpose.**
- (b) **B1** e.g. it is quicker and cheaper than running her own survey, and covers a much larger, national sample
- **(b) Answer: It is quicker and cheaper than running her own survey, and gives information from a much larger, national sample.**
- (c) **B1** e.g. the national report may not reflect the preferences of her own local customers
- **(c) Answer: The national report may not reflect the preferences of her own local customers.**
- (d) **B1** either option accepted, provided a valid supporting reason relevant to cost, time or relevance to her own customers is given
- **(d) Answer: Either option can be recommended, provided the reason given is a valid comparison of cost, time or relevance to her own customers.**

### Question 13

- (a) **B1** the numbers are just labels/codes for categories (Drama, Sports, Art), not numerical quantities that can be measured or counted
- (a) **B1** it would not make sense to carry out arithmetic with the codes, e.g. finding an 'average club code' would be meaningless
- **(a) Answer: The numbers 1, 2 and 3 are just labels for the categories Drama, Sports and Art - they do not represent a measurable quantity. It would not make sense to calculate something like an 'average club code', which shows the data is really qualitative.**
- (b) **B1** nominal, because the three clubs have no natural order or ranking
- **(b) Answer: Nominal, because there is no natural order or ranking between Drama club, Sports club and Art club.**