



Questionnaires, experiments, simulation and cleaning data

About 90 minutes Calculator allowed 91 marks

Date _____

A Do Now

The first two are what this chapter needs. The rest bring back earlier chapters so nothing is forgotten.

A1 Needed today [2 marks]

Isla writes this hypothesis: "Heavier cars and their fuel economy in miles per gallon are interesting." Explain why it is not a suitable hypothesis, and write a better one.

A2 Needed today [2 marks]

Maya uses house prices from a newspaper article for an investigation. Is this primary or secondary data for Maya? Give one disadvantage of using it.

A3 From chapter 2 [2 marks]

A quota sample of 400 people matches the area: 55% are men, and 35% of the men are over 60. How many men who are over 60 are in the sample?

Answer

A4 Maths skill [2 marks]

Ali has £120. He spends $\frac{3}{5}$ of it on a coat and $\frac{1}{4}$ of what is left on shoes. How much money does he have left?

Answer

A5 Maths skill [2 marks]

A spinner can land on A, B, C or D. $P(A) = 0.3$, $P(B) = 0.05$ and $P(C) = 0.1$. It is spun 400 times. How many times would you expect it to land on D?

Answer

B Example, then your turn

Your teacher works through each example with you. Then try the one beside it.

EXAMPLE 1

600 drivers are asked, "Have you driven above the speed limit in the last month?" Each driver rolls a fair dice in private. If it shows 1, they tick Yes. Otherwise they answer truthfully. 245 drivers tick Yes.

Estimate how many of the 600 drivers have driven above the speed limit. State one assumption you have made.

Hint Work out how many of the 600 are expected to roll a 1.

B1 400 students are asked, "Have you ever copied someone's homework?" Each spins a fair spinner with 4 equal sections in private. If it lands on the section marked Yes, they tick Yes. Otherwise they answer truthfully. 172 students tick Yes.

(a) Estimate how many of the 400 students have copied homework. [3 marks]

Answer

(b) State one assumption you have made. [1 mark]

Total for B1: 4 marks

EXAMPLE 2

A shop owner will use this sheet to record how 50 customers pay.

Method	Tally	Frequency
Cash		
Card		

Hint Look at the sheet itself: the options and the columns.

EXAMPLE 3

A school tests a new reading scheme on 60 pupils. Explain why a control group is needed, and how the pupils should be put into the two groups.

Hint What would the scheme's results be compared with?

EXAMPLE 4

A goalkeeper saves a penalty with probability 0.35. Use two-digit random numbers to simulate penalties until the keeper saves two in a row. How many penalties are taken?

71 22 48 09 66 30 12 85

Hint Use 35 of the 100 numbers 00 to 99 for a save.

B2 A student will use this sheet [3 marks]
to record how long 40 students slept last night. Is the sheet appropriate? Give reasons.

Sleep	Tally	Frequency
0 to 5 hours		
5 to 7 hours		
7 to 9 hours		

B3 A vet tests a new flea treatment on 80 dogs. [2 marks]
Explain why she needs a control group, and how she should decide which dogs get the treatment.

B4 A basketball player scores [3 marks]
with probability 0.6. Describe how to use the digits 0 to 9 to simulate her shots. Then use these digits to simulate shots until she misses twice in a row. How many shots does she take?

3 7 1 8 9 2 6 6

Answer

EXAMPLE 5

A council wants to know whether the residents of a town support a new cycle lane. Ade asks 15 cyclists in the town centre. Comment on the reliability and the validity of his results.

Hint Reliability is about the size of the sample. Validity is about whether he finds out what the council wants to know.

B5 Beth wants to know how much time the students at her school spend reading for fun each week. She asks 200 students in the school library one lunchtime. Comment on the reliability and the validity of her results. [2 marks]

EXAMPLE 6

For a set of 5 km running times, the lower quartile is 23 minutes and the upper quartile is 30 minutes. The fastest time is 14 minutes. Show that 14 minutes is not an outlier.

Hint Work out the lower limit, $Q_1 - 1.5 \times \text{IQR}$.

B6 A farm's eggs have mean mass 58 g and standard deviation 4.5 g. The heaviest egg is 71 g. Using the mean ± 3 standard deviations rule, show that 71 g is not an outlier. [2 marks]

C Practice

The questions come from all four lessons, in a mixed order.

C1 A survey asks, "How often do you exercise?" with the boxes Rarely, Sometimes and Often. Give two criticisms of the question and its boxes. Then write a better question with response boxes. [3 marks]

C2 The numbers of emails 15 office workers received in one day: [3 marks]
 3, 11, 14, 15, 17, 18, 19, 20, 21, 22, 23, 25, 26, 28, 44
 Which of these values are outliers by the $1.5 \times \text{IQR}$ rule?

Answer

- C3** Tara investigates whether the amount of salt in water affects the temperature at which the water freezes. Write down the explanatory variable, the response variable and one variable she should keep the same. [3 marks]

- C4** Seven runners' times for 1 km were meant to be recorded in seconds: [3 marks]
285, 4:10, 302, 3:55, 268, 0, 291
Clean the data and work out the median time in seconds.

Answer

- C5** A college sends a questionnaire by post to 2000 former students. Give one advantage and one disadvantage of a postal questionnaire compared with interviews. [2 marks]

- C6** Kai counts his pulse for 15 seconds once after a run and multiplies by 4. Give two ways he could make his result more reliable. [2 marks]

- C7** A researcher compares house prices in a town before and after its only factory closed. Explain why this is a natural experiment, and give one disadvantage of a natural experiment. [2 marks]

D Exam-style questions

Answer in the context of the question. Where a question asks a question, finish with your conclusion.

- D1** At a cinema snack counter, 45% of customers buy popcorn, 35% buy nachos and 20% buy sweets. Ali simulates the customers with a dice: 1 or 2 is popcorn, 3 or 4 is nachos and 5 or 6 is sweets. Ali says, "My simulation is a good model of the snack counter." [2 marks]

Is Ali correct? Give a reason for your answer.

- D2** A council lowers the speed limit on Mill Road from 30 mph to 20 mph. One month later a camera records the speeds, in mph, of 15 cars:
14, 17, 18, 19, 19, 20, 21, 21, 22, 23, 24, 25, 26, 27, 41

- (a) Write down the explanatory variable and the response variable in the council's study. [2 marks]

- (b) Using the $1.5 \times \text{IQR}$ rule, show that 41 mph is an outlier and that 14 mph is not. [3 marks]

- (c) The council says the 41 mph should be removed because it is an outlier. Do you agree? Give a reason. [1 mark]

Total for D2: 6 marks

D3 Ivy wants to find out how people heard about a music festival. She will use this [3 marks]
 sheet to record the answers of 80 people at the gate. Is her data collection
 sheet appropriate? Give reasons for your answer.

How you heard	Tally	Frequency
Poster		
Radio		
Friend		

D4 Rosa wants to test whether the height a tennis ball is dropped from affects how [4 marks]
 high it bounces. Describe how she should carry out the experiment. In your
 answer, name the explanatory and response variables, give two variables she
 should keep the same, and say how she can make her results reliable.

D5 The probability that Ella's bus is late is 0.15.

(a) Explain how Ella can use two-digit random numbers to simulate whether the [1 mark]
 bus is late on one day.

(b) Ella uses these numbers to simulate 20 days. [2 marks]

47 08 93 61 12 85 30 76 54 21 99 03 68 40 17 82 11 59 36 74

Use them to estimate the probability that the bus is late.

Answer

- (c) Ella says, "This shows the bus is late more often than 0.15." Is she correct? [1 mark]
Give a reason.

Total for D5: 4 marks

D6 A dentist tests a new mouthwash. 120 people use the mouthwash and 120 people use a placebo. After 6 months, 24 of the mouthwash group and 36 of the placebo group need a filling.

- (a) Work out the percentage of each group that need a filling. [2 marks]

Answer

- (b) Does the mouthwash appear to work? Give a reason. [1 mark]

Total for D6: 3 marks

D7 Ravi wants to find out how many hours a week the adults in his town spend exercising. He asks 30 people at a gym on a Saturday morning, "How many hours did you exercise last week?" Comment on the validity of his results. [2 marks]

D8 Noah wants to find out the favourite sport of the 900 students at his school. His plan: [4 marks]

- give a questionnaire to the 30 students in his own PE class;
- ask, "What is your favourite sport? Football / Rugby";
- record the answers in a table.

Assess Noah's plan.

D9 The heights of 15 adult women have lower quartile 1.58 m and upper quartile 1.70 m. The tallest is 1.86 m and the shortest is 1.15 m.

(a) Use the $1.5 \times \text{IQR}$ rule to decide which of these two heights is an outlier. [2 marks]

Answer

(b) Should the outlier be removed? Give a reason. [1 mark]

.....

.....

Total for D9: 3 marks

D10 A teacher records ten test marks out of 50:

34, 41, 28, 0, 45, 38, 52, 36, 34, 40

The student who scored 0 was absent.

(a) Identify two values that need cleaning, and say why. [2 marks]

.....

(b) Clean the data and work out the mean mark. [2 marks]

Answer

Total for D10: 4 marks

E Extension

No route is given. The hints are at the end of the sheet. Use one at a time.

E1 960 people answer a sensitive question. Each flips a coin in private. Heads: [3 marks]
answer truthfully. Tails: roll a dice, and tick Yes if it shows 6 and No
otherwise. 272 people tick Yes. Estimate how many of the 960 people would
truthfully answer Yes.

Answer

E2 A spinner lands on red with probability $\frac{2}{7}$ and on blue otherwise. Two-digit random numbers are used, with 00 to 19 for red, 20 to 69 for blue, and 70 to 99 ignored. These numbers simulate spins until red has come up twice: [3 marks]

83 45 12 77 66 91 30 05

How many spins did the simulation take?

Answer

E3 20 values have mean 50. One value is found to be an error and removed. [3 marks]
The other 19 values have mean 48 and standard deviation 5. Find the value that was removed, and show that it is an outlier of the other 19 values using the mean ± 3 standard deviations rule.

Answer

E4 In a random response survey, each person flips a coin: heads, tick Yes; tails, [3 marks]
answer the question truthfully. 35% of the 600 people would truthfully answer Yes. What proportion of the people who tick Yes were answering the question, rather than following the coin? Give your answer to 3 decimal places.

Answer

F How did it go?

Colour one circle on each line.

	Not yet	Nearly	Yes
I can criticise and improve a question or a data collection sheet, and use the random response method.	☐	☐	☐
I can plan a fair experiment with a control group and name its variables.	☐	☐	☐
I can comment on reliability and validity, and run a simulation with random numbers.	☐	☐	☐
I can clean data, find outliers with either rule, and say whether to remove them.	☐	☐	☐

Extension hints

Use one hint at a time. Cover the ones below it with a sheet of paper.

-
- | | | |
|-----------|---------------|--|
| E1 | Hint 1 | About half the people answer truthfully. Of the other half, how many are expected to roll a 6? |
| | Hint 2 | $480 \div 6 = 80$ are expected to tick Yes because of the dice. |
| | Hint 3 | So $272 - 80 = 192$ of the 480 truthful answers are Yes: $192 \div 480 = 0.4$. |
-
- | | | |
|-----------|---------------|--|
| E2 | Hint 1 | Why are 70 to 99 ignored? Check that 20 numbers out of 70 is $\frac{2}{7}$. |
| | Hint 2 | Cross out every number from 70 to 99: it is not a spin. |
| | Hint 3 | The spins are 45, 12, 66, 30, 05: red at 12 and at 05. |
-
- | | | |
|-----------|---------------|---|
| E3 | Hint 1 | Work out the total of the 20 values and the total of the 19 values. |
| | Hint 2 | $1000 - 912 = 88$. |
| | Hint 3 | The upper limit for the 19 values is $48 + 3 \times 5 = 63$. |
-
- | | | |
|-----------|---------------|--|
| E4 | Hint 1 | About 300 tick Yes because of the coin. |
| | Hint 2 | Of the 300 who answer truthfully, $0.35 \times 300 = 105$ tick Yes. |
| | Hint 3 | Of the $300 + 105 = 405$ who tick Yes, only those 105 were answering the question. The 300 who threw heads ticked Yes whatever their answer. |
-