

Name _____

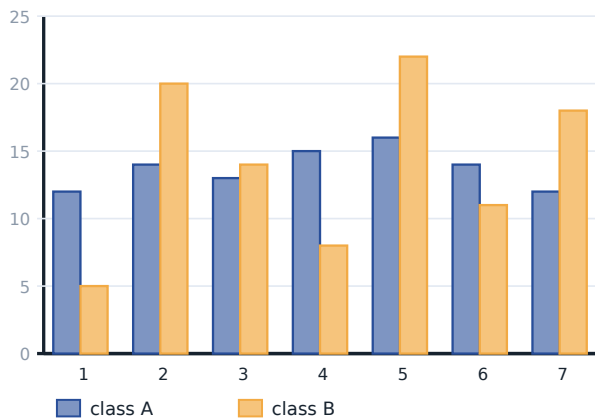
Date _____

16.5 Comparing data sets and drawing conclusions — worked answers

Compare two data sets using an average and a range, identify patterns and trends, and discuss conclusions and the sources of variation. · 46 questions

A · Fluency — Averages, ranges and what a statement tells you.

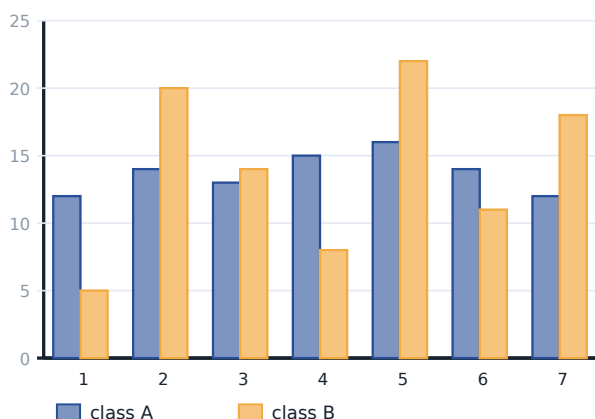
- 1 What is the mean for **class A**? Give the answer to one decimal place.



Answer: 13.7

$12 + 14 + 13 + 15 + 16 + 14 + 12 = 96$, over 7 values: $96 \div 7 = 13.7$.

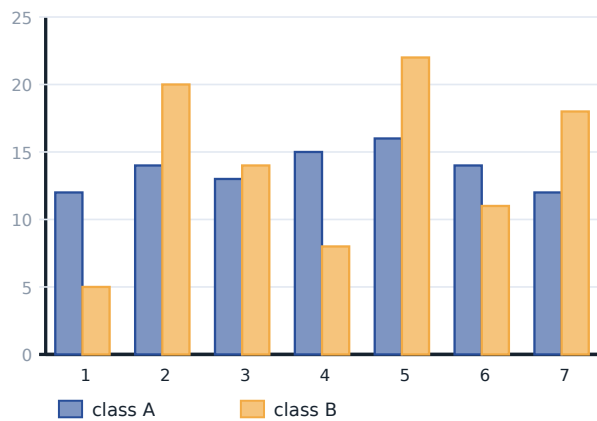
- 2 What is the mean for **class B**? Give the answer to one decimal place.



Answer: 14.0

$5 + 20 + 14 + 8 + 22 + 11 + 18 = 98$, over 7 values: $98 \div 7 = 14.0$.

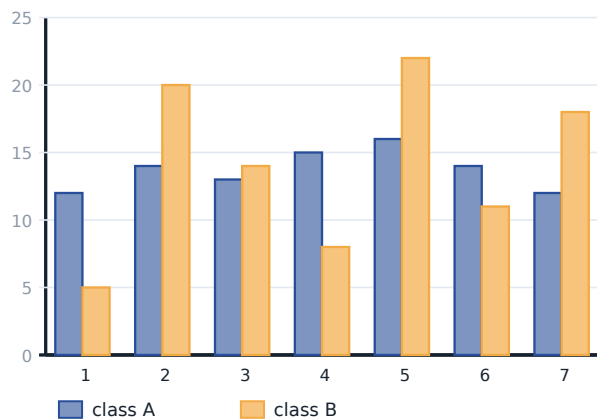
3 What is the range for **class A**?



Answer: 4

$$16 - 12 = 4.$$

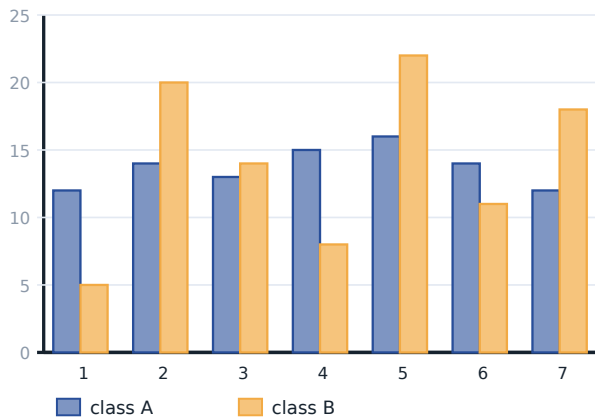
4 What is the range for **class B**?



Answer: 17

$$22 - 5 = 17.$$

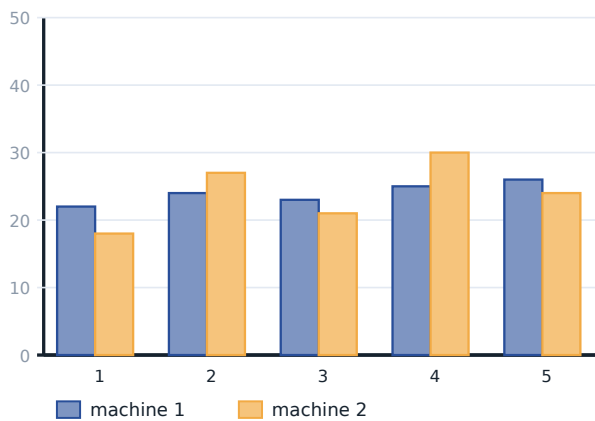
- 5 Which is the more **consistent**? Give the number, 1 for class A and 2 for class B.



Answer: 1

The ranges are 4 and 17. The smaller range belongs to **class A**, so that is the more consistent — its values vary less.

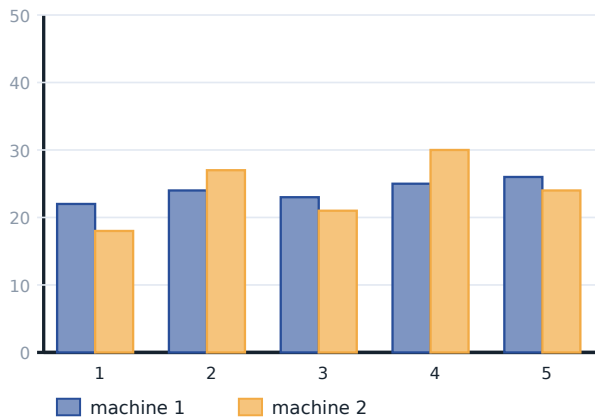
- 6 What is the mean for **machine 1**? Give the answer to one decimal place.



Answer: 24.0

$22 + 24 + 23 + 25 + 26 = 120$, over 5 values: $120 \div 5 = 24.0$.

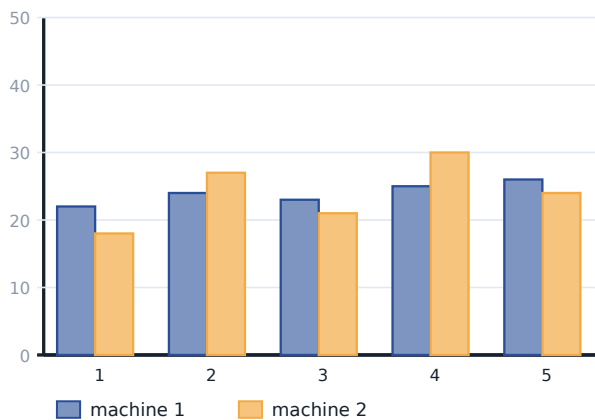
7 What is the mean for **machine 2**? Give the answer to one decimal place.



Answer: 24.0

$18 + 27 + 21 + 30 + 24 = 120$, over 5 values: $120 \div 5 = 24.0$.

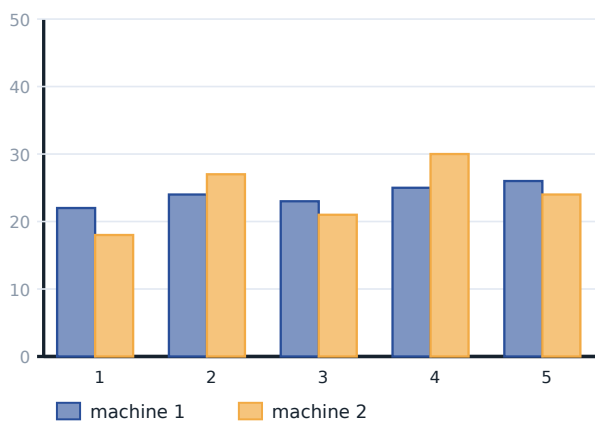
8 What is the range for **machine 1**?



Answer: 4

$26 - 22 = 4$.

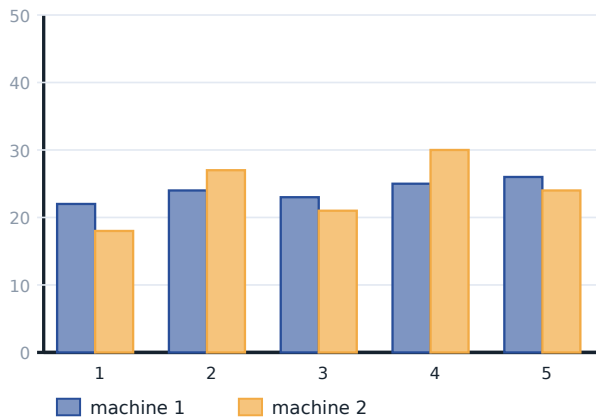
9 What is the range for **machine 2**?



Answer: 12

$30 - 18 = 12$.

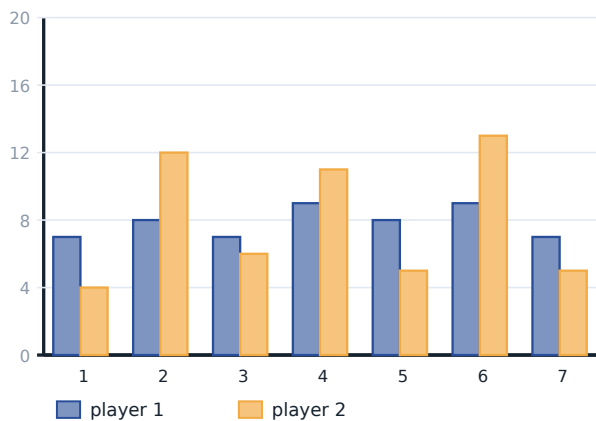
- 10** Which is the more **consistent**? Give the number, 1 for machine 1 and 2 for machine 2.



Answer: 1

The ranges are 4 and 12. The smaller range belongs to **machine 1**, so that is the more consistent — its values vary less.

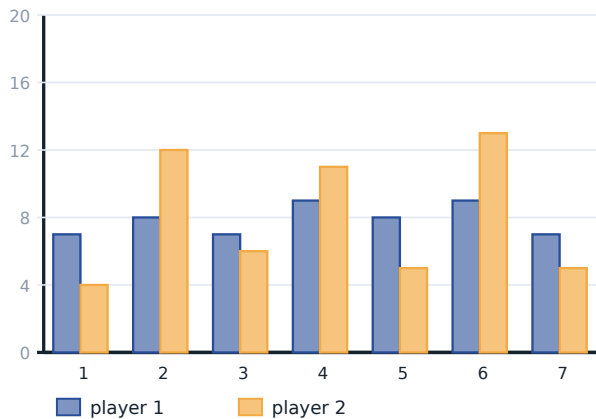
- 11** What is the mean for **player 1**? Give the answer to one decimal place.



Answer: 7.9

$7 + 8 + 7 + 9 + 8 + 9 + 7 = 55$, over 7 values: $55 \div 7 = 7.9$.

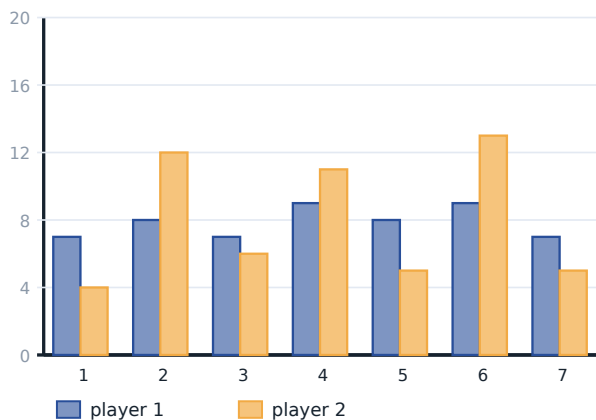
12 What is the mean for **player 2**? Give the answer to one decimal place.



Answer: 8.0

$4 + 12 + 6 + 11 + 5 + 13 + 5 = 56$, over 7 values: $56 \div 7 = 8.0$.

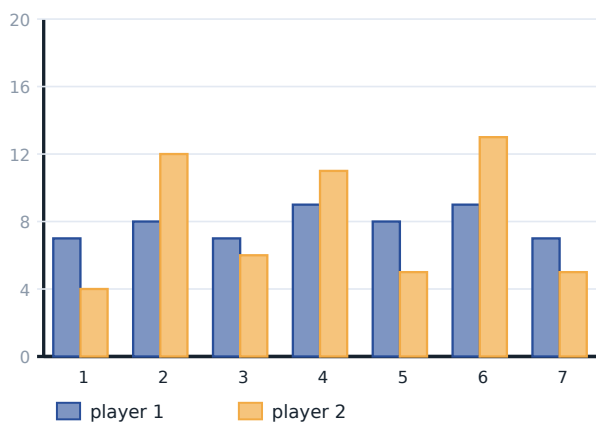
13 What is the range for **player 1**?



Answer: 2

$9 - 7 = 2$.

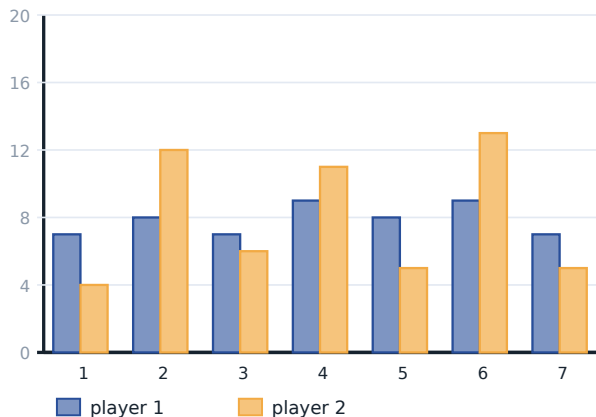
14 What is the range for **player 2**?



Answer: 9

$13 - 4 = 9$.

- 15** Which is the more **consistent**? Give the number, 1 for player 1 and 2 for player 2.



Answer: 1

The ranges are 2 and 9. The smaller range belongs to **player 1**, so that is the more consistent — its values vary less.

- 16** The mean of class A is higher than the mean of class B. What does that tell you?
- ☐ nothing useful
 - ☐ A typically scores better
 - ☐ the data is wrong
 - ☐ the chart is misleading

Answer: A typically scores better

Read it as a sentence about the situation, not just a number: A typically scores better.

- 17** The range of class A is smaller than the range of class B. What does that tell you?
- ☐ nothing useful
 - ☐ A is more consistent
 - ☐ the data is wrong
 - ☐ the chart is misleading

Answer: A is more consistent

Read it as a sentence about the situation, not just a number: A is more consistent.

- 18** Both classes have the same mean but different ranges. What does that tell you?
- ☐ nothing useful
 - ☐ they are typically the same, but one varies more
 - ☐ the data is wrong
 - ☐ the chart is misleading

Answer: they are typically the same, but one varies more

Read it as a sentence about the situation, not just a number: they are typically the same, but one varies more.

- 19** One value is far above all the others. What does that tell you?
- ☐ nothing useful
 - ☐ there is an outlier, which will have pulled the mean up
 - ☐ the data is wrong
 - ☐ the chart is misleading

Answer: there is an outlier, which will have pulled the mean up

Read it as a sentence about the situation, not just a number: there is an outlier, which will have pulled the mean up.

20 The line graph rises every year. What does that tell you?

- ☐ nothing useful
- ☐ there is an upward trend
- ☐ the data is wrong
- ☐ the chart is misleading

Answer: there is an upward trend

Read it as a sentence about the situation, not just a number: there is an upward trend.

21 The sample was only 5 people. What does that tell you?

- ☐ nothing useful
- ☐ the sample is too small to be confident
- ☐ the data is wrong
- ☐ the chart is misleading

Answer: the sample is too small to be confident

Read it as a sentence about the situation, not just a number: the sample is too small to be confident.

B · Reasoning — What makes a comparison, a trend and a fair sample.

22 To compare two data sets properly you should quote ...

- ☐ an average only
- ☐ a range only
- ☐ an average and a range
- ☐ the largest values

Answer: an average and a range

One says what is typical, the other how much it varies. Either alone is half a comparison.

23 Two classes have the same mean but different ranges. What can you say?

- ☐ They are identical
- ☐ They typically score the same, but one is more consistent
- ☐ One is better
- ☐ Nothing

Answer: They typically score the same, but one is more consistent

This is the standard two-mark comparison sentence.

24 A smaller range means ...

- ☐ a lower average
- ☐ more consistent data
- ☐ less data
- ☐ a mistake

Answer: more consistent data

The values are packed closer together.

25 Why is a conclusion from 5 people unreliable?

- ☐ Five is an odd number
- ☐ The sample is too small to represent the whole population
- ☐ It is not unreliable
- ☐ Because of the mean

Answer: The sample is too small to represent the whole population

This is Unit 6's sample-size idea coming back to be used.

26 What is a **trend**?

- ☐ A single high value
- ☐ The general direction the data is moving in
- ☐ The mode
- ☐ The range

Answer: The general direction the data is moving in

One dip does not destroy a rising trend.

27 Does a rising trend prove that the rise will continue?

- ☐ Yes
- ☐ No

Answer: No

It describes what has happened, not what must happen next. Predicting far beyond the data is a real risk.

28 A survey asked only people leaving a gym how much they exercise. The problem is ...

- ☐ the sample is too small
- ☐ the sample is biased — those people are not typical
- ☐ the question is unclear
- ☐ there is no problem

Answer: the sample is biased — those people are not typical

How the sample was chosen matters as much as its size.

29 Which of these is a source of variation worth mentioning?

- ☐ The weather on the day of an outdoor experiment
- ☐ The colour of the paper
- ☐ The name of the school
- ☐ Nothing

Answer: The weather on the day of an outdoor experiment

Anything that could reasonably have affected the readings belongs in the discussion.

30 Should a conclusion mention how the data was collected?

- ☐ Yes
- ☐ No

Answer: Yes

A perfect calculation on a biased sample still gives a wrong conclusion.

31 A good comparison sentence ...

- ☐ quotes only the numbers
- ☐ quotes the numbers and says what they mean in the context
- ☐ avoids numbers
- ☐ is as short as possible

Answer: quotes the numbers and says what they mean in the context

"A's mean is 13.7 against B's 14.0, so they score about the same, but A's range of 4 against B's 17 shows A is far more consistent."

C · Challenge — Differences between summaries, and predicting carefully.

32 Class A scores 12, 14, 13, 15, 16, 14, 12. What is its range?

Answer: 4

$16 - 12 = 4$.

- 33** Class B scores 5, 20, 14, 8, 22, 11, 18. What is its range?

Answer: 17

$$22 - 5 = 17.$$

- 34** Class A totals 96 over 7 learners and class B totals 98 over 7. What is the difference between their means, to one decimal place?

Answer: 0.3

A's mean is $96 \div 7 = 13.7$ and B's is $98 \div 7 = 14.0$, both to one decimal place, so the difference is 0.3.

- 35** Two shops have mean sales of 40. Shop A's range is 5, shop B's is 35. How much larger is B's range?

Answer: 30

$35 - 5 = 30$. Same typical sales, very different reliability.

- 36** A survey of 8 people found a mean of 6. A survey of 800 found a mean of 9. Which mean is more reliable? Answer 1 for the first and 2 for the second.

Answer: 2

The larger sample. A sample of 8 could easily be unrepresentative by chance, exactly as Unit 6 showed.

- 37** Sales rose 20, 24, 28, 32 over four years. If the trend continued, what would year 5 be?

Answer: 36

The rise is 4 each year: $32 + 4 = 36$. But a trend is a description, not a promise — say so alongside the prediction.

- 38** Two data sets have means of 25 and 25 and ranges of 3 and 18. How many of the two summary statistics differ?

Answer: 1

The means are equal, so only the ranges differ — one of the two. That is precisely why both must be quoted.

D · Word problems — Machines, players, shops and surveys.

- 39** Machine 1 fills bottles to 22, 24, 23, 25, 26 ml. What is its range, in millilitres?

Answer: 4

$$26 - 22 = 4 \text{ ml.}$$

- 40** Machine 2 fills bottles to 18, 27, 21, 30, 24 ml. What is its range, in millilitres?

Answer: 12

$$30 - 18 = 12 \text{ ml.}$$

- 41** Both machines have the same mean fill of 24 ml. Which should the factory keep? Answer 1 or 2.

Answer: 1

Machine 1. The means are equal, but its range of 4 ml against 12 ml means it is far more consistent — and consistency is what a bottling line needs.

- 42** A player scores 7, 8, 7, 9, 8, 9, 7 goals; another scores 4, 12, 6, 11, 5, 13, 5. What is the first player's range?

Answer: 2

$$9 - 7 = 2.$$

- 43** For those two players, what is the second player's range?

Answer: 9

$13 - 4 = 9$. The second is far less predictable, even though the totals are similar.

- 44** A shop's sales rose each month for six months. In the seventh they fell. Does that end the upward trend? Answer 1 for yes and 2 for no.

Answer: 2

No — one month against six is not enough to say the trend has changed. Watch a few more months before concluding anything.

- 45** A class of 28 was surveyed about the whole school of 900. What percentage of the school was asked, to the nearest whole percent?

Answer: 3

$28 \div 900 = 0.0311\dots$, which is 3% to the nearest whole percent. A small sample, and all from one class — so it may also be biased.

- 46** Rainfall was measured on 3 days and on 300 days. Which gives the better estimate of the yearly mean? Answer 1 for 3 days and 2 for 300 days.

Answer: 2

300 days. More readings smooth out the day-to-day variation, exactly as more trials did in Unit 13.