

# Data, Fit, and Intro Statistics

Trend, center, spread, and basic model fit reasoning.

Name \_\_\_\_\_ Date \_\_\_\_\_

32 main 2-up grid 13 pages visible side quests

## Completion Reward



Shown here as a small pack artifact, not a preview destination.

### 1. What does the mean of a data set measure?

- A. A typical central value based on equal sharing
- B. The largest value only
- C. The number of data points
- D. The middle value after sorting no matter the context

### 1.1. The mean of a data set measures:

- A. a typical balance point
- B. the middle value only
- C. the largest spread
- D. the steepness of a graph

### 1.2. The median of a data set is:

- A. the average of all values
- B. the middle value when the data are ordered
- C. the largest value
- D. the difference between max and min

### 1.3. An outlier usually affects which measure more?

- A. median
- B. mean
- C. neither
- D. mode only

### 1.4. When is median often more useful than mean?

- A. when the data have an outlier
- B. when all values are equal
- C. when the sample is tiny
- D. when the graph is linear

### 1.5. What is the median of 2, 4, 5, 9, 12?

- A. 4
- B. 5
- C. 6
- D. 12

### 2. What does the range of a data set measure?

- A. The average of all the data
- B. The middle value only
- C. How far the data spread from the minimum to the maximum
- D. The exact trend line for the data

### 2.1. The range of a data set measures:

- A. the middle value
- B. the spread from smallest to largest
- C. the average only
- D. the line of fit

### 2.2. What is the range of 12, 15, 18, and 19?

- A. 4
- B. 5
- C. 6
- D. 7

### 2.3. Which example is univariate data?

- A. height and weight for each student
- B. time studied and quiz score
- C. just the quiz scores of one class
- D. temperature and ice cream sales

### 2.4. Which example is bivariate data?

- A. shoe sizes only
- B. test scores only
- C. height and arm span for each student
- D. the days of the week

### 2.5. A larger range usually means:

- A. less spread
- B. more spread
- C. the same mean
- D. a stronger association

### 3. If $y$ tends to decrease as $x$ increases, what kind of association is shown?

- A. Negative association
- B. Positive association
- C. No association
- D. Constant association

### 3.1. If $y$ tends to increase as $x$ increases, the association is:

- A. negative
- B. positive
- C. none
- D. impossible

### 3.2. If $y$ tends to decrease as $x$ increases, the association is:

- A. negative
- B. positive
- C. none
- D. circular

### 3.3. Which description suggests a weak association?

- A. points tightly hug a line
- B. points loosely trend upward with lots of scatter
- C. every point is identical
- D. there is no graph

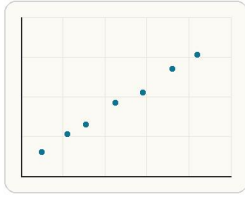
### 3.4. A scatterplot mainly helps show:

- A. the relationship between two variables
- B. the exact median only
- C. a proof of causation
- D. one variable only

### 3.5. Why does an upward-trending scatterplot not prove causation?

- A. because association does not by itself prove one variable causes the other
- B. because positive trends are always fake
- C. because scatterplots ignore data
- D. because slope cannot be positive

4. What does a scatterplot help show?



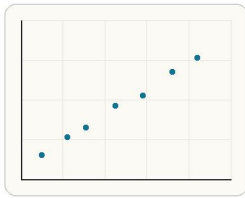
Each point represents one paired observation, so the graph helps you see how two quantitative variables move together.

- A. Only one value at a time
- B. The exact answer to a system
- C. How two quantitative variables are related
- D. A guaranteed cause-and-effect rule

4.3. An outlier is a data point that:

- A. fits the pattern perfectly
- B. lies unusually far from the rest
- C. is always the largest
- D. must be removed

5. If one variable tends to increase as the other increases, what kind of association is this?



As x increases, the plotted y-values tend to rise too, which is the visual pattern of positive association.

- A. Negative
- B. None
- C. Impossible
- D. Positive

5.3. Which description suggests a weak association?

- A. points tightly hug a line
- B. points loosely trend upward with lots of scatter
- C. every point is identical
- D. there is no graph

4.1. The mean of 10, 12, and 14 is:

- A. 11
- B. 12
- C. 13
- D. 14

4.4. If y tends to decrease as x increases, the association is:

- A. positive
- B. negative
- C. none
- D. vertical

5.1. If y tends to increase as x increases, the association is:

- A. negative
- B. positive
- C. none
- D. impossible

5.4. A scatterplot mainly helps show:

- A. the relationship between two variables
- B. the exact median only
- C. a proof of causation
- D. one variable only

4.2. The median of 3, 8, 9, 20, 30 is:

- A. 8
- B. 9
- C. 20
- D. 14

4.5. Why doesn't correlation prove causation?

- A. A trend alone does not show one variable causes the other
- B. Correlated data are always random
- C. Causation is only for lines
- D. Scatter plots cannot compare variables

5.2. If y tends to decrease as x increases, the association is:

- A. negative
- B. positive
- C. none
- D. circular

5.5. Why does an upward-trending scatterplot not prove causation?

- A. because association does not by itself prove one variable causes the other
- B. because positive trends are always fake
- C. because scatterplots ignore data
- D. because slope cannot be positive

**6. When is the median usually more helpful than the mean?**

- A. When every value is the same
- B. When there are exactly two data points
- C. When you want to multiply the values
- D. When the data set has an outlier

**6.3. An outlier usually affects which measure more?**

- A. median
- B. mean
- C. neither
- D. mode only

**7. Which example is bivariate data?**

- A. A list of quiz scores only
- B. Each student's study time and quiz score
- C. A list of shoe sizes only
- D. A list of one class's ages only

**7.3. Which example is univariate data?**

- A. height and weight for each student
- B. time studied and quiz score
- C. just the quiz scores of one class
- D. temperature and ice cream sales

**8. When might median be more helpful than mean?**

- A. When the data are perfectly balanced
- B. When an outlier would pull the mean away from the center
- C. When you want the largest value
- D. Never

**8.3. An outlier usually affects which measure more?**

- A. median
- B. mean
- C. neither
- D. mode only

**6.1. The mean of a data set measures:**

- A. a typical balance point
- B. the middle value only
- C. the largest spread
- D. the steepness of a graph

**6.4. When is median often more useful than mean?**

- A. when the data have an outlier
- B. when all values are equal
- C. when the sample is tiny
- D. when the graph is linear

**7.1. The range of a data set measures:**

- A. the middle value
- B. the spread from smallest to largest
- C. the average only
- D. the line of fit

**7.4. Which example is bivariate data?**

- A. shoe sizes only
- B. test scores only
- C. height and arm span for each student
- D. the days of the week

**8.1. The mean of a data set measures:**

- A. a typical balance point
- B. the middle value only
- C. the largest spread
- D. the steepness of a graph

**8.4. When is median often more useful than mean?**

- A. when the data have an outlier
- B. when all values are equal
- C. when the sample is tiny
- D. when the graph is linear

**6.2. The median of a data set is:**

- A. the average of all values
- B. the middle value when the data are ordered
- C. the largest value
- D. the difference between max and min

**6.5. What is the median of 2, 4, 5, 9, 12?**

- A. 4
- B. 5
- C. 6
- D. 12

**7.2. What is the range of 12, 15, 18, and 19?**

- A. 4
- B. 5
- C. 6
- D. 7

**7.5. A larger range usually means:**

- A. less spread
- B. more spread
- C. the same mean
- D. a stronger association

**8.2. The median of a data set is:**

- A. the average of all values
- B. the middle value when the data are ordered
- C. the largest value
- D. the difference between max and min

**8.5. What is the median of 2, 4, 5, 9, 12?**

- A. 4
- B. 5
- C. 6
- D. 12

**9. If values of  $y$  tend to increase when values of  $x$  increase, what kind of association is shown?**

- A. Negative association
- B. No association
- C. Quadratic association only
- D. Positive association

**9.1. If  $y$  tends to increase as  $x$  increases, the association is:**

- A. negative
- B. positive
- C. none
- D. impossible

**9.2. If  $y$  tends to decrease as  $x$  increases, the association is:**

- A. negative
- B. positive
- C. none
- D. circular

**9.3. Which description suggests a weak association?**

- A. points tightly hug a line
- B. points loosely trend upward with lots of scatter
- C. every point is identical
- D. there is no graph

**9.4. A scatterplot mainly helps show:**

- A. the relationship between two variables
- B. the exact median only
- C. a proof of causation
- D. one variable only

**9.5. Why does an upward-trending scatterplot not prove causation?**

- A. because association does not by itself prove one variable causes the other
- B. because positive trends are always fake
- C. because scatterplots ignore data
- D. because slope cannot be positive

**10. A student says a very large outlier never changes the mean. What is wrong?**

- A. The mean ignores all extreme values automatically.
- B. Outliers only affect the median, never the mean.
- C. An outlier can pull the mean noticeably upward or downward.
- D. Outliers are removed before any average is found.

**10.1. The mean of a data set measures:**

- A. a typical balance point
- B. the middle value only
- C. the largest spread
- D. the steepness of a graph

**10.2. The median of a data set is:**

- A. the average of all values
- B. the middle value when the data are ordered
- C. the largest value
- D. the difference between max and min

**10.3. An outlier usually affects which measure more?**

- A. median
- B. mean
- C. neither
- D. mode only

**10.4. When is median often more useful than mean?**

- A. when the data have an outlier
- B. when all values are equal
- C. when the sample is tiny
- D. when the graph is linear

**10.5. What is the median of 2, 4, 5, 9, 12?**

- A. 4
- B. 5
- C. 6
- D. 12

**11. Find the mean of 4, 6, 8, and 10. Answer with a number.**

**11.1. The mean of a data set measures:**

- A. a typical balance point
- B. the middle value only
- C. the largest spread
- D. the steepness of a graph

**11.2. The median of a data set is:**

- A. the average of all values
- B. the middle value when the data are ordered
- C. the largest value
- D. the difference between max and min

**11.3. An outlier usually affects which measure more?**

- A. median
- B. mean
- C. neither
- D. mode only

**11.4. When is median often more useful than mean?**

- A. when the data have an outlier
- B. when all values are equal
- C. when the sample is tiny
- D. when the graph is linear

**11.5. What is the median of 2, 4, 5, 9, 12?**

- A. 4
- B. 5
- C. 6
- D. 12

**12. Find the median of 3, 5, 7, 8, and 20. Answer with a number.**

12.1. The mean of a data set measures:

- A. a typical balance point
- B. the middle value only
- C. the largest spread
- D. the steepness of a graph

12.2. The median of a data set is:

- A. the average of all values
- B. the middle value when the data are ordered
- C. the largest value
- D. the difference between max and min

12.3. An outlier usually affects which measure more?

- A. median
- B. mean
- C. neither
- D. mode only

12.4. When is median often more useful than mean?

- A. when the data have an outlier
- B. when all values are equal
- C. when the sample is tiny
- D. when the graph is linear

12.5. What is the median of 2, 4, 5, 9, 12?

- A. 4
- B. 5
- C. 6
- D. 12

**13. Find the range of 12, 15, 18, and 19. Answer with a number.**

13.1. The range of a data set measures:

- A. the middle value
- B. the spread from smallest to largest
- C. the average only
- D. the line of fit

13.2. What is the range of 12, 15, 18, and 19?

- A. 4
- B. 5
- C. 6
- D. 7

13.3. Which example is univariate data?

- A. height and weight for each student
- B. time studied and quiz score
- C. just the quiz scores of one class
- D. temperature and ice cream sales

13.4. Which example is bivariate data?

- A. shoe sizes only
- B. test scores only
- C. height and arm span for each student
- D. the days of the week

13.5. A larger range usually means:

- A. less spread
- B. more spread
- C. the same mean
- D. a stronger association

**14. Find the mean of 5, 5, 7, 9, and 14. Answer with a number.**

14.1. The mean of a data set measures:

- A. a typical balance point
- B. the middle value only
- C. the largest spread
- D. the steepness of a graph

14.2. The median of a data set is:

- A. the average of all values
- B. the middle value when the data are ordered
- C. the largest value
- D. the difference between max and min

14.3. An outlier usually affects which measure more?

- A. median
- B. mean
- C. neither
- D. mode only

14.4. When is median often more useful than mean?

- A. when the data have an outlier
- B. when all values are equal
- C. when the sample is tiny
- D. when the graph is linear

14.5. What is the median of 2, 4, 5, 9, 12?

- A. 4
- B. 5
- C. 6
- D. 12

**15. What does a residual tell you?**

- A. The median of the data set
- B. The exact slope of every point
- C. Whether the variables are automatically causal
- D. How far an actual data value is from the predicted value

15.3. A line of fit is  $y = 2x + 5$ . What is the predicted y-value when  $x = 6$ ?

- A. 12
- B. 15
- C. 17
- D. 19

**16. When is the median usually more useful than the mean?**

- A. When the data are perfectly balanced
- B. When an outlier pulls the data far to one side
- C. When every value is identical
- D. When you want to multiply the data

16.3. An outlier usually affects which measure more?

- A. median
- B. mean
- C. neither
- D. mode only

**17. Which description suggests a weak association?**

- A. The points cluster tightly around an upward trend.
- B. The points are widely scattered with no clear trend line.
- C. Every increase in  $x$  matches a steady increase in  $y$ .
- D. The data fit a clear straight line closely.

17.3. Which description suggests a weak association?

- A. points tightly hug a line
- B. points loosely trend upward with lots of scatter
- C. every point is identical
- D. there is no graph

15.1. A residual tells you:

- A. how far the actual value is from the predicted value
- B. the slope of the fit line
- C. the average of the data
- D. the number of outliers

15.4. A line of fit predicts 42, but the actual value is 46. What is the residual?

- A. -4
- B. 0
- C. 4
- D. 88

16.1. The mean of a data set measures:

- A. a typical balance point
- B. the middle value only
- C. the largest spread
- D. the steepness of a graph

16.4. When is median often more useful than mean?

- A. when the data have an outlier
- B. when all values are equal
- C. when the sample is tiny
- D. when the graph is linear

17.1. If  $y$  tends to increase as  $x$  increases, the association is:

- A. negative
- B. positive
- C. none
- D. impossible

17.4. A scatterplot mainly helps show:

- A. the relationship between two variables
- B. the exact median only
- C. a proof of causation
- D. one variable only

15.2. If a point lies above the line of fit, its residual is usually:

- A. positive
- B. negative
- C. zero
- D. undefined

15.5. A reasonable line of fit should:

- A. hit every point exactly
- B. balance the points overall
- C. always pass through the origin
- D. have slope 1

16.2. The median of a data set is:

- A. the average of all values
- B. the middle value when the data are ordered
- C. the largest value
- D. the difference between max and min

16.5. What is the median of 2, 4, 5, 9, 12?

- A. 4
- B. 5
- C. 6
- D. 12

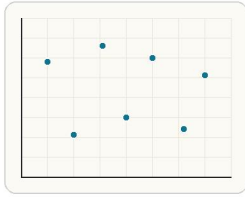
17.2. If  $y$  tends to decrease as  $x$  increases, the association is:

- A. negative
- B. positive
- C. none
- D. circular

17.5. Why does an upward-trending scatterplot not prove causation?

- A. because association does not by itself prove one variable causes the other
- B. because positive trends are always fake
- C. because scatterplots ignore data
- D. because slope cannot be positive

18. What type of association does the scatterplot show?



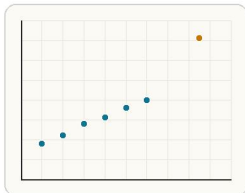
The points do not form a steady upward or downward pattern, so the visual evidence does not support a strong association.

- A. Positive association
- B. Negative association
- C. No clear association
- D. Perfect linear association

18.3. Which description suggests a weak association?

- A. points tightly hug a line
- B. points loosely trend upward with lots of scatter
- C. every point is identical
- D. there is no graph

19. Which statement best describes point B?



Most points stay together in one band while point B sits far from the group, making it visually stand out as an outlier.

- A. It lies exactly on the main trend line
- B. It proves the association is negative
- C. It is an outlier far from the overall cluster
- D. It makes the scatterplot a function

19.3. An outlier usually affects which measure more?

- A. median
- B. mean
- C. neither
- D. mode only

18.1. If  $y$  tends to increase as  $x$  increases, the association is:

- A. negative
- B. positive
- C. none
- D. impossible

18.4. A scatterplot mainly helps show:

- A. the relationship between two variables
- B. the exact median only
- C. a proof of causation
- D. one variable only

19.1. The mean of a data set measures:

- A. a typical balance point
- B. the middle value only
- C. the largest spread
- D. the steepness of a graph

19.4. When is median often more useful than mean?

- A. when the data have an outlier
- B. when all values are equal
- C. when the sample is tiny
- D. when the graph is linear

18.2. If  $y$  tends to decrease as  $x$  increases, the association is:

- A. negative
- B. positive
- C. none
- D. circular

18.5. Why does an upward-trending scatterplot not prove causation?

- A. because association does not by itself prove one variable causes the other
- B. because positive trends are always fake
- C. because scatterplots ignore data
- D. because slope cannot be positive

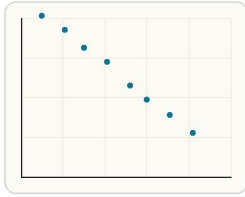
19.2. The median of a data set is:

- A. the average of all values
- B. the middle value when the data are ordered
- C. the largest value
- D. the difference between max and min

19.5. What is the median of 2, 4, 5, 9, 12?

- A. 4
- B. 5
- C. 6
- D. 12

20. What kind of association does the scatterplot show?



As  $x$  increases, the cloud of points trends downward, which is the visual signal for negative association.

- A. Positive
- B. None
- C. Cannot tell
- D. Negative

20.1. If  $y$  tends to increase as  $x$  increases, the association is:

- A. negative
- B. positive
- C. none
- D. impossible

20.2. If  $y$  tends to decrease as  $x$  increases, the association is:

- A. negative
- B. positive
- C. none
- D. circular

20.3. Which description suggests a weak association?

- A. points tightly hug a line
- B. points loosely trend upward with lots of scatter
- C. every point is identical
- D. there is no graph

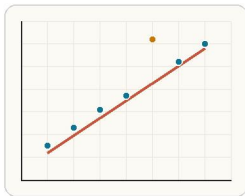
20.4. A scatterplot mainly helps show:

- A. the relationship between two variables
- B. the exact median only
- C. a proof of causation
- D. one variable only

20.5. Why does an upward-trending scatterplot not prove causation?

- A. because association does not by itself prove one variable causes the other
- B. because positive trends are always fake
- C. because scatterplots ignore data
- D. because slope cannot be positive

21. Point A is above the line of fit. What does that mean?



A point above the model line has an actual  $y$ -value that is greater than the line predicted at that  $x$ -value.

- A. The actual  $y$ -value is less than the predicted  $y$ -value
- B. The residual is zero
- C. There is no association
- D. The actual  $y$ -value is greater than the predicted  $y$ -value

21.1. A residual tells you:

- A. how far the actual value is from the predicted value
- B. the slope of the fit line
- C. the average of the data
- D. the number of outliers

21.2. If a point lies above the line of fit, its residual is usually:

- A. positive
- B. negative
- C. zero
- D. undefined

21.3. A line of fit is  $y = 2x + 5$ . What is the predicted  $y$ -value when  $x = 6$ ?

- A. 12
- B. 15
- C. 17
- D. 19

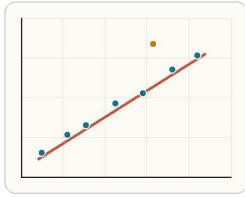
21.4. A line of fit predicts 42, but the actual value is 46. What is the residual?

- A. -4
- B. 0
- C. 4
- D. 88

21.5. A reasonable line of fit should:

- A. hit every point exactly
- B. balance the points overall
- C. always pass through the origin
- D. have slope 1

22. If a point lies above the line of fit, what does that mean about its residual?



When a point is above the line of fit, the residual is positive because actual exceeds predicted.

- A. The residual is negative
- B. The residual must be zero
- C. Residuals do not apply
- D. The residual is positive

22.3. A line of fit is  $y = 2x + 5$ . What is the predicted  $y$ -value when  $x = 6$ ?

- A. 12
- B. 15
- C. 17
- D. 19

22.1. A residual tells you:

- A. how far the actual value is from the predicted value
- B. the slope of the fit line
- C. the average of the data
- D. the number of outliers

22.2. If a point lies above the line of fit, its residual is usually:

- A. positive
- B. negative
- C. zero
- D. undefined

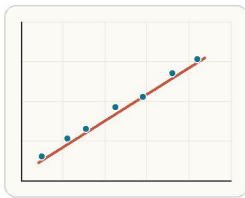
22.4. A line of fit predicts 42, but the actual value is 46. What is the residual?

- A. -4
- B. 0
- C. 4
- D. 88

22.5. A reasonable line of fit should:

- A. hit every point exactly
- B. balance the points overall
- C. always pass through the origin
- D. have slope 1

23. What does the slope of a line of fit represent in context?



The slope of the fit line estimates how much the predicted output changes when the input increases by 1 unit.

- A. The predicted change in  $y$  for each one-unit change in  $x$
- B. The exact cause of the trend
- C. The number of points on the graph
- D. The mean of the data set

23.1. A residual tells you:

- A. how far the actual value is from the predicted value
- B. the slope of the fit line
- C. the average of the data
- D. the number of outliers

23.2. If a point lies above the line of fit, its residual is usually:

- A. positive
- B. negative
- C. zero
- D. undefined

23.3. A line of fit is  $y = 2x + 5$ . What is the predicted  $y$ -value when  $x = 6$ ?

- A. 12
- B. 15
- C. 17
- D. 19

23.4. A line of fit predicts 42, but the actual value is 46. What is the residual?

- A. -4
- B. 0
- C. 4
- D. 88

23.5. A reasonable line of fit should:

- A. hit every point exactly
- B. balance the points overall
- C. always pass through the origin
- D. have slope 1

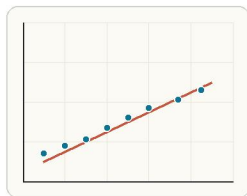
**24. A student says a scatterplot showing that study time and quiz score rise together proves studying more definitely causes higher scores. What is wrong?**

- A. Positive association always means one variable causes the other.
- B. Association does not by itself prove causation.
- C. Scatterplots cannot show any relationship at all.
- D. Cause can only be studied with quadratic models.

**24.3. Which description suggests a weak association?**

- A. points tightly hug a line
- B. points loosely trend upward with lots of scatter
- C. every point is identical
- D. there is no graph

**25. A student says "study time causes high scores because the scatterplot slopes upward." What is the issue?**



An upward trend can show a relationship in the data without proving that one variable directly causes the other.

- A. An upward trend means x and y are equal
- B. Association alone does not prove causation
- C. Scatterplots can only show negative trends
- D. Nothing is wrong

**25.3. Which description suggests a weak association?**

- A. points tightly hug a line
- B. points loosely trend upward with lots of scatter
- C. every point is identical
- D. there is no graph

**24.1. If y tends to increase as x increases, the association is:**

- A. negative
- B. positive
- C. none
- D. impossible

**24.4. A scatterplot mainly helps show:**

- A. the relationship between two variables
- B. the exact median only
- C. a proof of causation
- D. one variable only

**25.1. If y tends to increase as x increases, the association is:**

- A. negative
- B. positive
- C. none
- D. impossible

**24.2. If y tends to decrease as x increases, the association is:**

- A. negative
- B. positive
- C. none
- D. circular

**24.5. Why does an upward-trending scatterplot not prove causation?**

- A. because association does not by itself prove one variable causes the other
- B. because positive trends are always fake
- C. because scatterplots ignore data
- D. because slope cannot be positive

**25.2. If y tends to decrease as x increases, the association is:**

- A. negative
- B. positive
- C. none
- D. circular

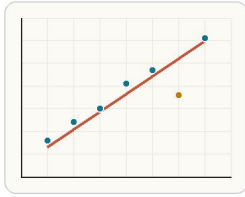
**25.4. A scatterplot mainly helps show:**

- A. the relationship between two variables
- B. the exact median only
- C. a proof of causation
- D. one variable only

**25.5. Why does an upward-trending scatterplot not prove causation?**

- A. because association does not by itself prove one variable causes the other
- B. because positive trends are always fake
- C. because scatterplots ignore data
- D. because slope cannot be positive

26. A student says point C below the line means the model predicted too low. What is wrong?



A point below the line means the model predicted higher than the actual data value at that x-value.

- A. A point below the line means the model predicted too high
- B. A point below the line means the residual is positive
- C. A point below the line proves causation
- D. Nothing is wrong

26.3. A line of fit is  $y = 2x + 5$ . What is the predicted y-value when  $x = 6$ ?

- A. 12
- B. 15
- C. 17
- D. 19

26.1. A residual tells you:

- A. how far the actual value is from the predicted value
- B. the slope of the fit line
- C. the average of the data
- D. the number of outliers

26.2. If a point lies above the line of fit, its residual is usually:

- A. positive
- B. negative
- C. zero
- D. undefined

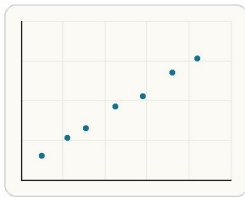
26.4. A line of fit predicts 42, but the actual value is 46. What is the residual?

- A. -4
- B. 0
- C. 4
- D. 88

26.5. A reasonable line of fit should:

- A. hit every point exactly
- B. balance the points overall
- C. always pass through the origin
- D. have slope 1

27. A student says a scatterplot proving higher study time is linked to higher scores means studying more definitely causes higher scores. What is wrong?



A trend in the scatterplot can support association, but the graph alone does not establish a cause-and-effect claim.

- A. Association alone does not prove causation
- B. Scatterplots never show any relationship
- C. Study time can only be on the y-axis
- D. Nothing is wrong

27.1. If y tends to increase as x increases, the association is:

- A. negative
- B. positive
- C. none
- D. impossible

27.2. If y tends to decrease as x increases, the association is:

- A. negative
- B. positive
- C. none
- D. circular

27.3. Which description suggests a weak association?

- A. points tightly hug a line
- B. points loosely trend upward with lots of scatter
- C. every point is identical
- D. there is no graph

27.4. A scatterplot mainly helps show:

- A. the relationship between two variables
- B. the exact median only
- C. a proof of causation
- D. one variable only

27.5. Why does an upward-trending scatterplot not prove causation?

- A. because association does not by itself prove one variable causes the other
- B. because positive trends are always fake
- C. because scatterplots ignore data
- D. because slope cannot be positive

**28. A line of fit is  $y = 2x + 5$ . Predict  $y$  when  $x = 6$ . Answer with a number.**

28.3. A line of fit is  $y = 2x + 5$ . What is the predicted  $y$ -value when  $x = 6$ ?

- A. 12
- B. 15
- C. 17
- D. 19

**29. A line of fit predicts 42, but the actual value is 46. What is the residual? Answer with a number.**

29.3. A line of fit is  $y = 2x + 5$ . What is the predicted  $y$ -value when  $x = 6$ ?

- A. 12
- B. 15
- C. 17
- D. 19

**30. Which description sounds most like a good line of fit?**

- A. It touches only the highest point.
- B. It runs through the middle of the trend with points balanced above and below it.
- C. It must pass through every point exactly.
- D. It is always horizontal.

30.3. A line of fit is  $y = 2x + 5$ . What is the predicted  $y$ -value when  $x = 6$ ?

- A. 12
- B. 15
- C. 17
- D. 19

28.1. A residual tells you:

- A. how far the actual value is from the predicted value
- B. the slope of the fit line
- C. the average of the data
- D. the number of outliers

28.4. A line of fit predicts 42, but the actual value is 46. What is the residual?

- A. -4
- B. 0
- C. 4
- D. 88

29.1. A residual tells you:

- A. how far the actual value is from the predicted value
- B. the slope of the fit line
- C. the average of the data
- D. the number of outliers

29.4. A line of fit predicts 42, but the actual value is 46. What is the residual?

- A. -4
- B. 0
- C. 4
- D. 88

30.1. A residual tells you:

- A. how far the actual value is from the predicted value
- B. the slope of the fit line
- C. the average of the data
- D. the number of outliers

30.4. A line of fit predicts 42, but the actual value is 46. What is the residual?

- A. -4
- B. 0
- C. 4
- D. 88

28.2. If a point lies above the line of fit, its residual is usually:

- A. positive
- B. negative
- C. zero
- D. undefined

28.5. A reasonable line of fit should:

- A. hit every point exactly
- B. balance the points overall
- C. always pass through the origin
- D. have slope 1

29.2. If a point lies above the line of fit, its residual is usually:

- A. positive
- B. negative
- C. zero
- D. undefined

29.5. A reasonable line of fit should:

- A. hit every point exactly
- B. balance the points overall
- C. always pass through the origin
- D. have slope 1

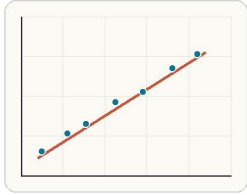
30.2. If a point lies above the line of fit, its residual is usually:

- A. positive
- B. negative
- C. zero
- D. undefined

30.5. A reasonable line of fit should:

- A. hit every point exactly
- B. balance the points overall
- C. always pass through the origin
- D. have slope 1

31. Which line of fit is more reasonable for a scatterplot: one that balances the points overall or one that must pass through every point?



A good fit line follows the overall trend and balances the points rather than trying to pass through every observation.

- A. The line that passes through every point
- B. The line that balances the points overall
- C. No line can ever be used
- D. Both are equally required

31.3. A line of fit is  $y = 2x + 5$ . What is the predicted y-value when  $x = 6$ ?

- A. 12
- B. 15
- C. 17
- D. 19

31.1. A residual tells you:

- A. how far the actual value is from the predicted value
- B. the slope of the fit line
- C. the average of the data
- D. the number of outliers

31.2. If a point lies above the line of fit, its residual is usually:

- A. positive
- B. negative
- C. zero
- D. undefined

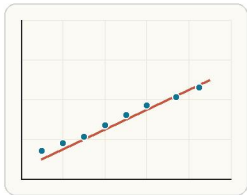
31.4. A line of fit predicts 42, but the actual value is 46. What is the residual?

- A. -4
- B. 0
- C. 4
- D. 88

31.5. A reasonable line of fit should:

- A. hit every point exactly
- B. balance the points overall
- C. always pass through the origin
- D. have slope 1

32. Using the line of fit, what is a reasonable estimate for y when  $x = 7$ ?



Use the model line to read a reasonable approximate y-value at the given x instead of expecting an exact data point.

- A. 4
- B. 2
- C. 7
- D. 10

32.1. A residual tells you:

- A. how far the actual value is from the predicted value
- B. the slope of the fit line
- C. the average of the data
- D. the number of outliers

32.2. If a point lies above the line of fit, its residual is usually:

- A. positive
- B. negative
- C. zero
- D. undefined

32.3. A line of fit is  $y = 2x + 5$ . What is the predicted y-value when  $x = 6$ ?

- A. 12
- B. 15
- C. 17
- D. 19

32.4. A line of fit predicts 42, but the actual value is 46. What is the residual?

- A. -4
- B. 0
- C. 4
- D. 88

32.5. A reasonable line of fit should:

- A. hit every point exactly
- B. balance the points overall
- C. always pass through the origin
- D. have slope 1