Data, Fit, and Intro Statistics

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

Name _____ Date _____

32 main · 2-up grid · 3 pages

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

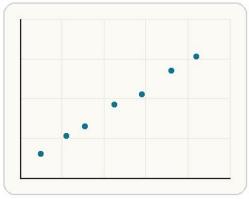
2. What does the range of a data set measure?

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

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

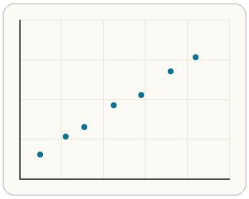
4. What does a scatterplot help show?



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

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

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

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

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

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

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.

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

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

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

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

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

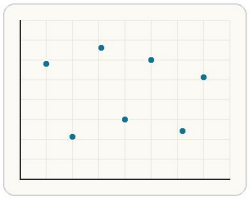
16. When is the median usually more useful than the mean?

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

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.

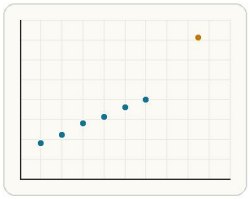
18. What type of association does the scatterplot show?



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

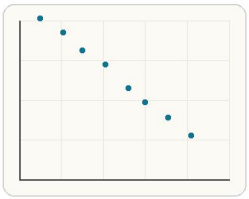
19. Which statement best describes point B?



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

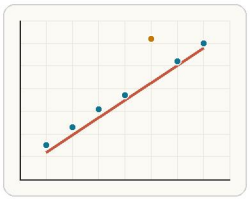
20. What kind of association does the scatterplot show?



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

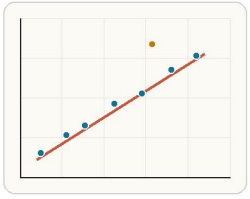
21. Point A is above the line of fit. What does that mean?



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

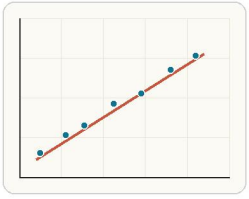
22. If a point lies above the line of fit, what does that mean about its residual?



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

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



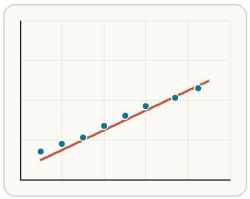
The slope of the fit line estimates how much the predicted output changes when the input increases by 1 unit.

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

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.

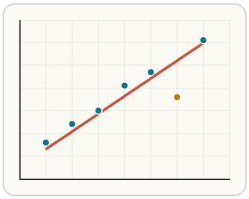
25. A student says "study time causes high scores because the scatterplot slopes upward." What is the issue?



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

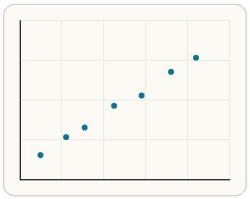
26. A student says point C below the line means the model predicted too low. What is wrong?



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

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

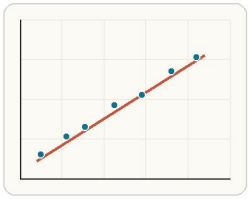
28. A line of fit is y = 2x + 5. Predict y when x = 6. Answer with a number.

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

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.

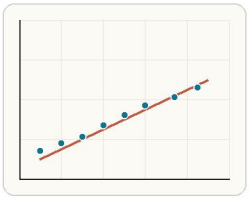
31. Which line of fit is more reasonable for a scatterplot: one that balances the points overall or one that must pass through every point?



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

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