

Data Detectives

Summarize data and use models to make predictions.

HIGH SCHOOL • MA.912.DP.2

MA.912.DP.2 "Solve problems involving univariate and bivariate numerical data."

NAME _____

DATE _____

DESCRIBE AND PREDICT

For one-variable data, use the mean or median to describe a typical value and use range or interquartile range to describe spread. The interquartile range is the upper quartile minus the lower quartile. For two-variable data, use $y = mx + b$ to predict y from x , and find a residual by subtracting predicted y from actual y .

WORKED EXAMPLE

Find the mean and range of 31, 35, 39.

1. Add the values: $31 + 35 + 39 = 105$.
2. Divide by the number of values: $105 / 3 = 35$.
3. Subtract the least value from the greatest value: $39 - 31 = 8$.

Answer: Mean = 35; range = 8

PRACTICE 6 problems

- 1** Part 1 PRACTICE, 1. Find the mean of 11, 13, 17, and 19.

ANSWER _____

- 2** 2. Find the interquartile range of 2, 6, 10, 14, 18. Use the median-of-halves method.

ANSWER _____

- 3** 3. Which set has the greater range? Set A: 4, 9, 11, 16. Set B: 5, 7, 13, 15.

ANSWER _____

- 4** 4. A model is $y = 2x + 5$. Find the predicted value of y when $x = 7$.

ANSWER _____

- 5** 5. A model is $y = 4x + 1$. When $x = 6$, the actual value of y is 29. Find the residual.

ANSWER _____

- 6** 6. Model A is $y = 5x + 2$. Model B is $y = 4x + 7$. When $x = 5$, which model predicts the greater y -value?

ANSWER _____

APPLY IT Show your work

Part 2 APPLY, 1. A student records these lap times, in minutes, during track practice: 41, 43, 45, 47.

- 1** Find the mean time. Then a slower lap time of 71 minutes is added. Find the new mean and decide whether the mean or median better describes a typical lap time for all five laps.

WORK SPACE

ANSWER _____

2. A student council models ticket sales y from the number of social-media posts x with $y = 7x + 9$.

- 2** After 11 posts, the council sells 90 tickets. Find the predicted number of tickets, the residual, and what the residual means.

WORK SPACE

ANSWER _____

CHALLENGE One step up

CHALLENGE

Part 3 CHALLENGE. A trend line passes through the points (1, 10) and (5, 26). Find the equation of the line. Use it to predict y when $x = 7$. If the actual value is 36, find the residual and state whether the model overpredicts or underpredicts.

PART 1 • PRACTICE

#	PROBLEM	ANSWER
1	Part 1 PRACTICE, 1. Find the mean of 11, 13, 17, and 19.	15
2	2. Find the interquartile range of 2, 6, 10, 14, 18. Use the median-of-halves method.	12
3	3. Which set has the greater range? Set A: 4, 9, 11, 16. Set B: 5, 7, 13, 15.	Set A, with a range of 12
4	4. A model is $y = 2x + 5$. Find the predicted value of y when $x = 7$.	19
5	5. A model is $y = 4x + 1$. When $x = 6$, the actual value of y is 29. Find the residual.	4
6	6. Model A is $y = 5x + 2$. Model B is $y = 4x + 7$. When $x = 5$, which model predicts the greater y -value?	Neither model. Both predict 27.

PART 2 • APPLY IT

- Original mean: 44 minutes. New mean: 49.4 minutes. The median, 45 minutes, better describes a typical lap time.** The original total is 176, so $176 / 4 = 44$. With 71 included, the total is 247 and $247 / 5 = 49.4$; the high time pulls the mean upward, so use the median.
- Predicted sales: 86 tickets. Residual: 4 tickets. Actual sales were 4 tickets greater than predicted.**
Substitute 11 for x : $7(11) + 9 = 86$. The residual is actual minus predicted, so $90 - 86 = 4$.

CHALLENGE

$y = 4x + 6$; predicted $y = 34$; residual = 2; the model underpredicts.

Accept equivalent wording for comparisons and interpretations. For residuals, use actual value minus predicted value; for the interquartile range item, use the median-of-halves convention that excludes the overall median.