Name: ___________________________________ Date: ________________

RESIDUALS
COMMON CORE ALGEBRA I

In the last lesson, we saw how the **correlation coefficient** (or *r*-value) measures the **predictability** of the model (or how well it will do in its predictions). Although the *r*-value is an excellent measure, it does not tell us whether the model is **appropriate** only whether it does a good job at **predicting**. Today we will examine what are known as **residuals and residual graphs** to determine if a linear model makes sense.

***Exercise* #1:** A skydiver jumps from an airplane and an attached micro-computer records the time and speed of the diver for the first 12 seconds of the diver's freefall. The data is shown in the table below.

Time (sec)	0	2	4	6	8	10	12
Speed (ft/sec)	0	25	46	60	68	72	74

(a) Find the equation for the line of best fit for this data set. Round both coefficients to the nearest *tenth*. As well, determine the correlation coefficient and round it to the nearest *hundredth*. Based on the correlation coefficient, characterize the fit as positive or negative and how strong of a fit it is. Create a plot with both the data and line of best fit shown on it. You do not need to reproduce the plot below.

(b) The **residual** of a data point is defined as the **difference between the observed *y*-value and the predicted *y*-value**. Using tables on your calculator fill in the table below with the predicted values (rounded to the nearest integer) and the residuals for each data point.

Time (sec)	0	2	4	6	8	10	12
Speed (ft/sec)	0	25	46	60	68	72	74
Prediction (ft/sec)							
Residual (ft/sec)							

(c) Sketch a plot of the residuals below. Your teacher will need to show you how to do this on your graphing calculator. Make sure all other scatter plots and equations are off. Do the residuals show any distinct pattern?

Generally, we do not want residuals to fall along a curve or make a distinct pattern. If so, then it is likely that a **linear model** is not appropriate to fit the data and perhaps an **exponential or quadratic** model is better.

***Exercise* #2:** A school district was attempting to correlate the number of hours a student studies in a given week with their grade point average. They surveyed 8 students and found the following data.

Hours Studying	3	7	2	11	8	16	5	9
GPA	78	80	75	94	89	92	80	84

(a) Find the equation for the line of best fit and the associated *r*-value. Round the linear coefficients to the nearest *tenth* and the *r*-value to the nearest *hundredth*. Then create a scatter plot with both the data and the line graphed. You do not need to reproduce that graph here unless your teacher asks you to.

(b) What is the value of the residual associated with the data point $(11, 94)$? Show the calculation that leads to your answer.

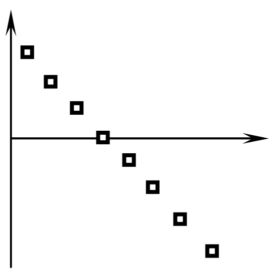
(c) Produce, using your calculator, the residual graph. It does not need to be exact, but show your **WINDOW** and the correct general location of the residuals.

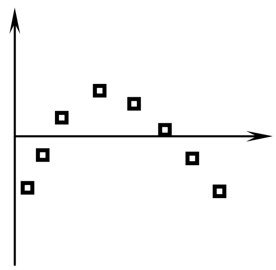
(d) Why does this residual plot show a more appropriate linear model than the one in Exercise #1, even though the *r*-value is worse?

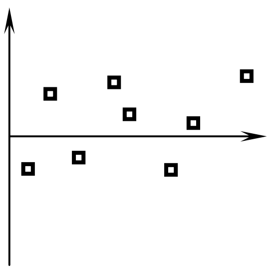
Name: ________________________________ Date: _______________

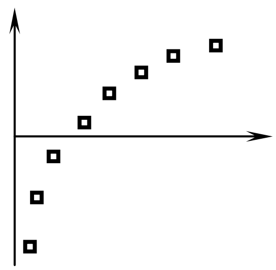
RESIDUALS
COMMON CORE ALGEBRA I HOMEWORK

1. Which of the following residual plots indicates a model that is most appropriate?

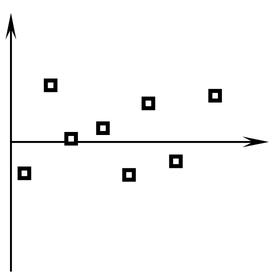
(1)


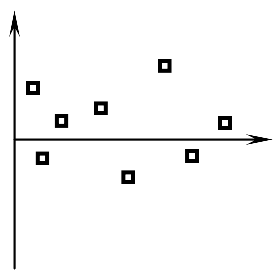
(3)


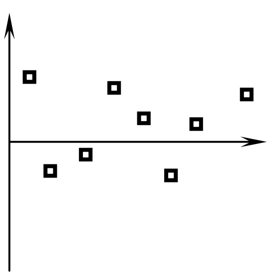
(2)


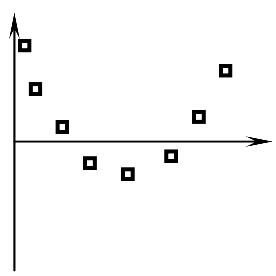
(4)


2. Which of the following residual plots would indicate the linear model used to produce it was an inappropriate choice?

(1)


(3)


(2)


(4)


3. A set of data is fit with linear regression. The equation for the best fit line was $y=5.2x+18$. If the observed value when $x=10$ was $y=62$, then which of the following represents the value of the residual for this data point?

(1) 52 (3) 8

(2) −8 (4) −10

4. Physics students are performing a lab where they allow a ball to roll down a ramp and record the distance that it has rolled versus the time it has been rolling. The data for one such experiment are shown below.

Time (sec)	0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
Distance (ft)	0	0.4	1.5	3.2	5.6	8.5	12.6	17.2	22.8

(a) Determine the equation for the line of best fit. Round your coefficients to the nearest *tenth*. Also, determine the correlation coefficient. Round it to the nearest *hundredth*.

(b) Using your calculator, produce a scatter plot of this data set along with the line of best fit. Do your best job to sketch it below. Label the **WINDOW** you used.

(c) Calculate the residual for the data point $(2.0, 5.6)$. Show your calculation below.

(d) Create a graph of the residuals using your calculator. Draw a sketch of the graph below, showing the **WINDOW** and the approximate location of the points. Make sure to turn off your other scatter plot and the line of best fit. Sketch the residual plot to the right.

(e) Is the linear model appropriate for this data set given the residual plot? Explain below.