Residuals in the Common Core

Sharon Taylor
Georgia Southern University

Kathleen Cage Mittag
University of Texas-San Antonio (retired)

taylors@georgiasouthern.edu
kcagemittag@yahoo.com

NCTM Regional
Las Vegas
October 24, 2013

Below is a data set for homework averages and test grades for 10 students.

Homework Average	Test 1 Grade
56	42
75	75
92	90
91	97
78	74
77	59
69	74
80	83
80	87
55	45

One way to analyze this data is to use a scatterplot which is given for you on the next page.

The next procedure is adapted from "Spaghetti Regression" in Unit 4 of Coordinate Algebra from georgiastandards.org

Visualize what you think is a good line that fits the data. Place your spaghetti on the scatterplot so that the spaghetti represents your line. Tape the spaghetti to the scatterplot so that it does not roll around on your paper. Estimate an equation for your line.

y =

Now measure the vertical distance from each point on the scatterplot to the spaghetti line.
(This differs from the task. The task says students can measure horizontally, vertically, or perpendicular as long as the group decides. Residuals are defined by the vertical distances from a point to the line, which is what we are using here.)

Once you measure your vertical distances, you will notice that some will be above the line (call these values positive) and some will be below the line (call these values negative). Add the distances and see what happens.
(This differs from the task. The task asks students to measure with spaghetti pieces and then tape all the pieces together to see which piece is shortest. Residual analysis is based on observing that the total vertical distance of the "best fit" line is 0.)

positive values = _____ negative values = _____ sum = _____

Now explore residuals using the calculator. You will mimic the process you used by hand.

You need to find the distances from the point to the line (y – value of each point minus the y – coordinate of your line). You will then use the calculator to add those distances. You can also have the calculator plot the distances so that you can look for a pattern. What do you notice?

You should also use the calculator to find the line of best fit. How close is your line to the calculator line?

Line of best fit: y =

Now explore the residuals of **the** line of best fit instead of your line and see what happens.

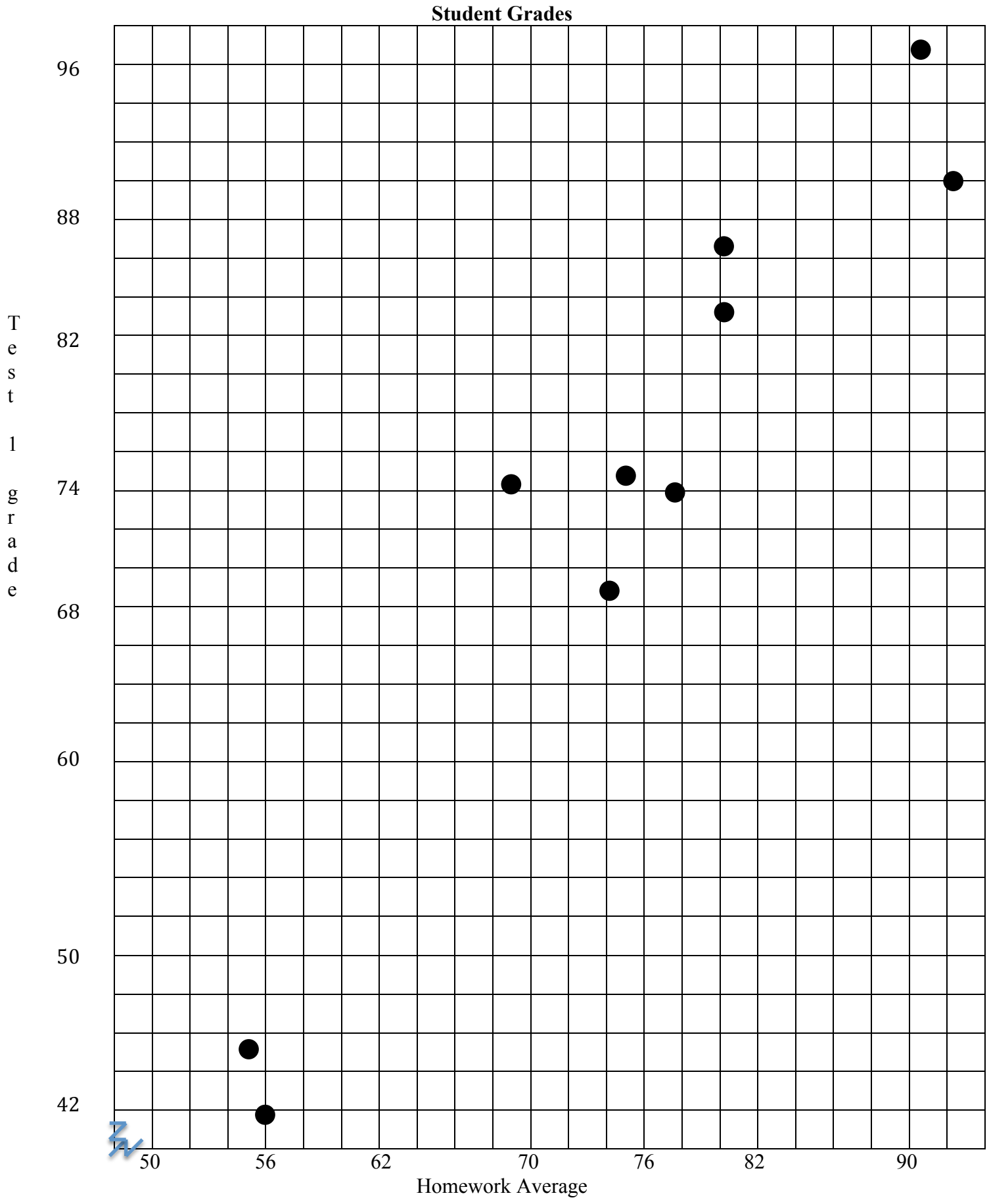
x – value	*y* – value	predicted y – value (plug *x* into regression line)	difference (actual – predicted)

What happens when you add all the differences in the last column?

The calculator should have also given you *r* and r^2 values. What do these values tell you about the goodness of fit for the best fit line? How does that compare with what you know about the residuals?

r =

r^2 =


Student Grades
96
88
82
74
68
60
50
42
Test 1 grade
50
56
62
70
76
82
90
Homework Average