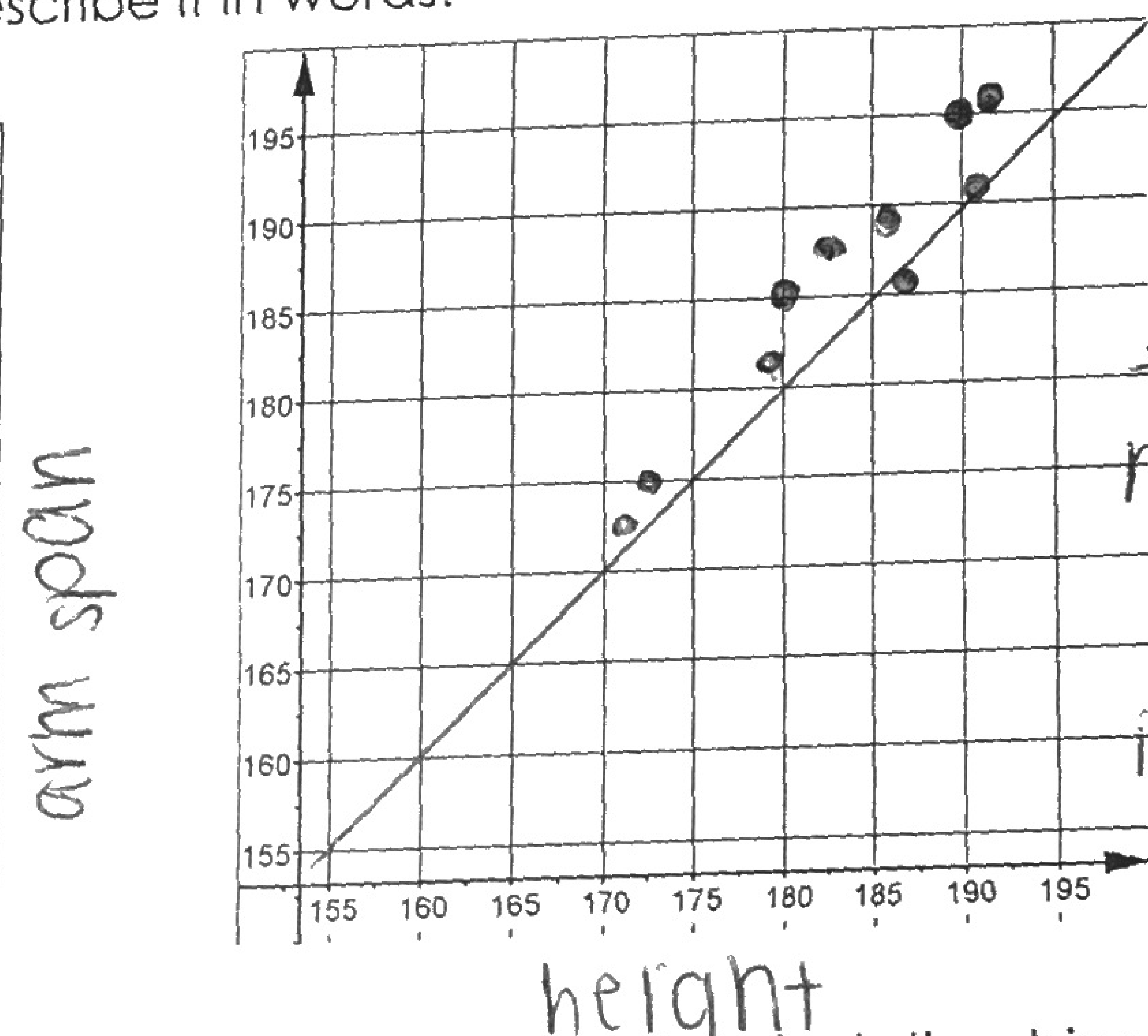
Name: Key **Date:** ______ **Period:** ______

One factor that talent scouts look for in potentially competitive swimmers is the ratio of their height to their arm span. For most people, arm span is generally equal to height. Consider U.S. Olympic swimmer Michael Phelps, who is 6 feet, 4 inches (193 centimeters) tall with an arm span of 6 feet, 7 inches (200 centimeters). In fact, the U.S. swim team found that its male swimmers have an average height of 187.1 centimeters and an average arm span of 192.9 centimeters. Of course, other factors influence the success of a swimmer, but coaches often look at a swimmer's physical attributes, including arm span, to determine which strokes he or she should focus on.

At a local competitive swim club, the coach measured the height and arm span of his top 10 swimmers. The data are shown in the table below.

1. Make a scatterplot of the data and describe it in words.

Height (cm)	Arm Span (cm)
172	173
173	175
179	182
180	185
183	187
186	189
187	186
190	195
191	191
192	196



The data appears to be relatively linear & is increasing.

This model represents a **directly proportional** situation. Directly proportional relationships always have linear graphs that pass through the origin, and the **constant of proportionality** is the slope of that line.

What does it mean for two quantities to be directly proportional?

if one quantity increases, so does the other & vice versa.

Think about a person with a height of 0 centimeters. Theoretically, what is his/her arm span?

around 0 cm.

The general function rule that describes a directly proportional relationship is $\boxed{y = kx.}$

2. Use what you know about the situation and the data to find a function model for this data set. Explain your reasoning.

since the average height is 187.1 cm and arm span is 192.9, the ratio of $\frac{\text{arm span}}{\text{height}} = \frac{192.9}{187.1} = 1.031$ is the constant of proportionality (slope). you can assume it's a directly proportional relationship so $\boxed{y = 1x}$

When two quantities are related linearly, there is a correlation between the two quantities. This correlation can actually be calculated numerically by comparing the values in the data set to the mean and standard deviation of the data.

This numerical value is called the **correlation coefficient** and is usually expressed as *r*. The correlation coefficient (*r*) is a number between – 1 and 1 and indicates the direction and strength of the linear relationship between the two quantities.

L1 = height (x) L2 = arm span (y)

3. Use your **graphing calculator** to compute a regression analysis of the swimmers' arm spans in relation to their height.

stat → calc → 4: Lin Reg ax+b (L1, L2

What does the information from the calculator tell you?

$y = 1.022x - 1.429$

How does the equation given by the calculator compare to the function you found in Question 2? the equations are very similar.

What does the *r*-value on your calculator tell you? * 2nd → "0" (catalog) → diagnostic on $\boxed{\text{enter}}$

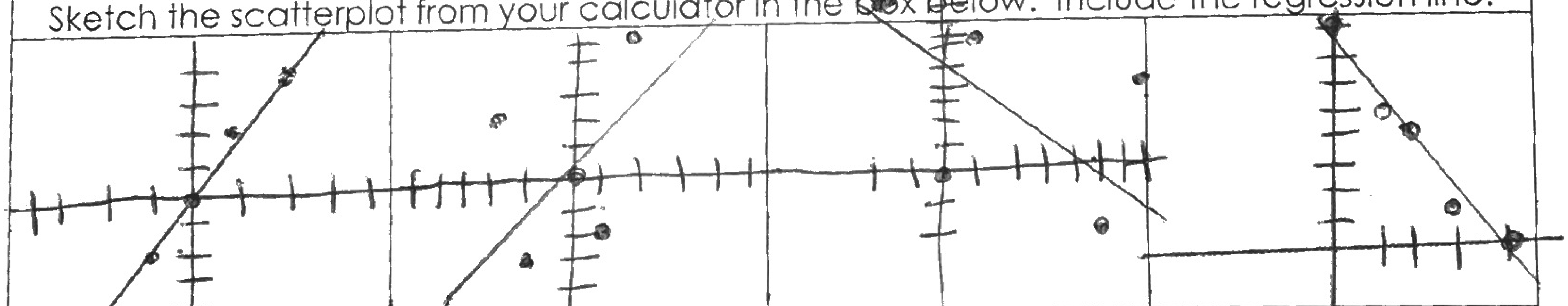
$r = 0.963$

r^2 is the coeff. of determination (% of data closest to the line of best fit)

4. Work in a group of four students. Each group member enters one of the data sets below into a graphing calculator, makes a scatterplot, and performs a linear regression analysis. Compare the graphs and the values of the correlation coefficients (*r*). Record an observation about how the value of *r* describes the strength and direction of the relationship between the variables. scatter plot → . 2nd → stat plot → turn on

X	Y	X	Y	X	Y	X	Y
-2	-4	-2	2	-2	9	0	8
-1	-2	-1	-3	0	0	1	5
0	0	0	0	1	7	2	4
1	2	1	-2	5	-2	3	1
2	4	2	5	7	4	4	0

Sketch the scatterplot from your calculator in the box below. Include the regression line.



r-value: $r = 1$	r-value: $r = 0.345$	r-value: $r = -0.477$	r-value: $r = -0.985$
equation: $y = 2x$	equation: $y = 0.7x + 0.4$	equation: $y = -0.6x + 4.91$	equation: $y = -2x + 7.6$

no correlation

strong negative ← -1 0.5 moderate 0 0.5 1 → strong positive

* closer r is to 0 the weaker the correlation.

* r ⊕ data increasing; r ⊖ decreasing

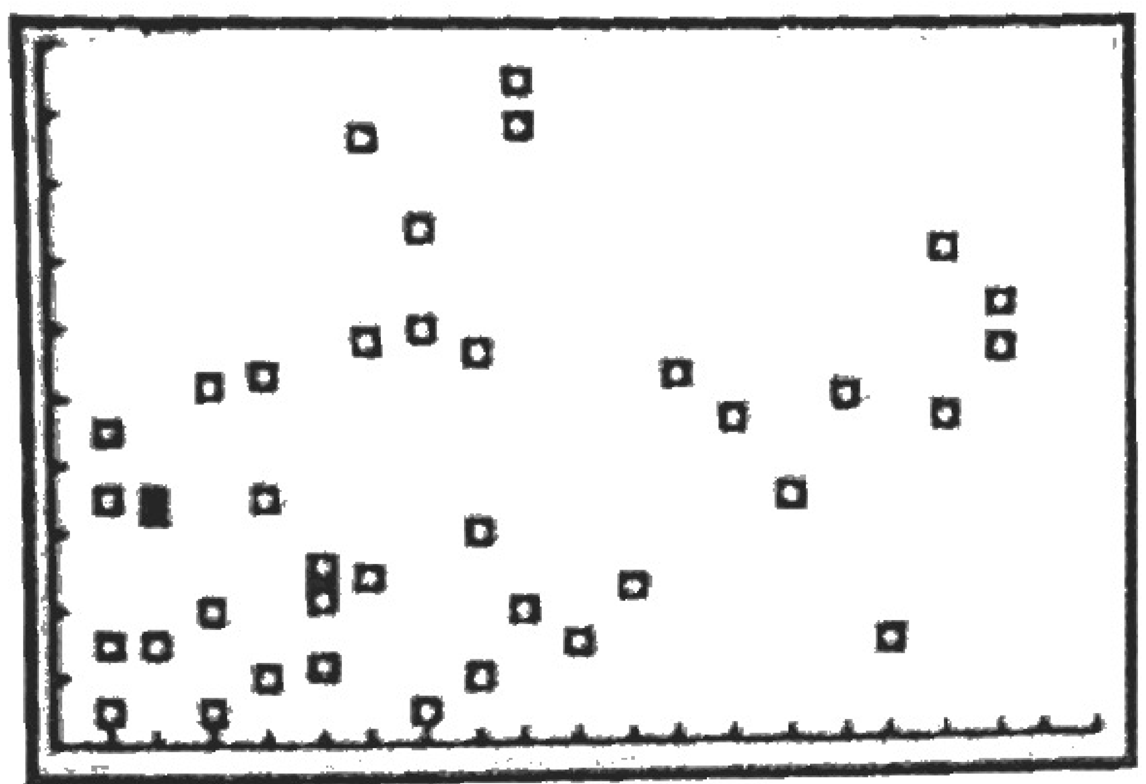
AMDM Unit 5 – Regression SAS 1: Analyzing Linear Regression Equations

Circle the bold word that makes the statement true: When r gets closer to ± 1, the data have a **low** **high** correlation and the line is a **good** **bad** fit for the data.

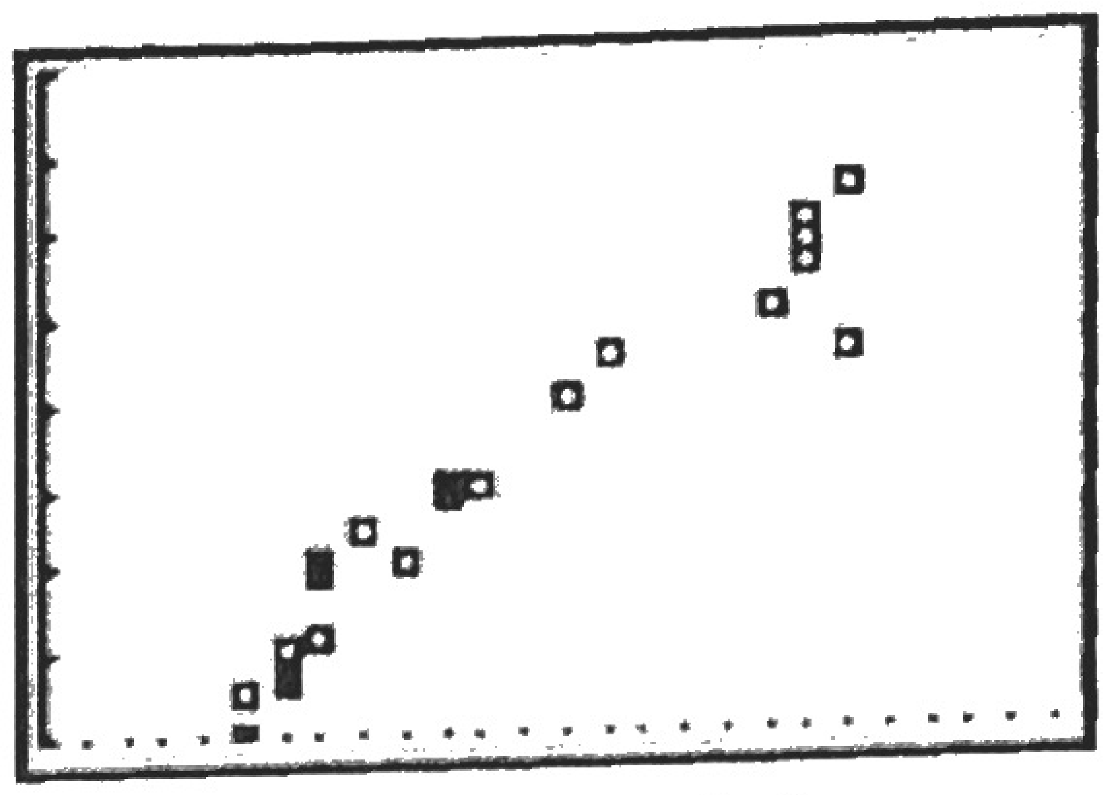
Finish this statement: When ***r*** is closer to 0, the data:
there is a low correlation & the line is a poor fit.

5. Consider each scatterplot below. Write the letter to match each *r*-value to a scatterplot.

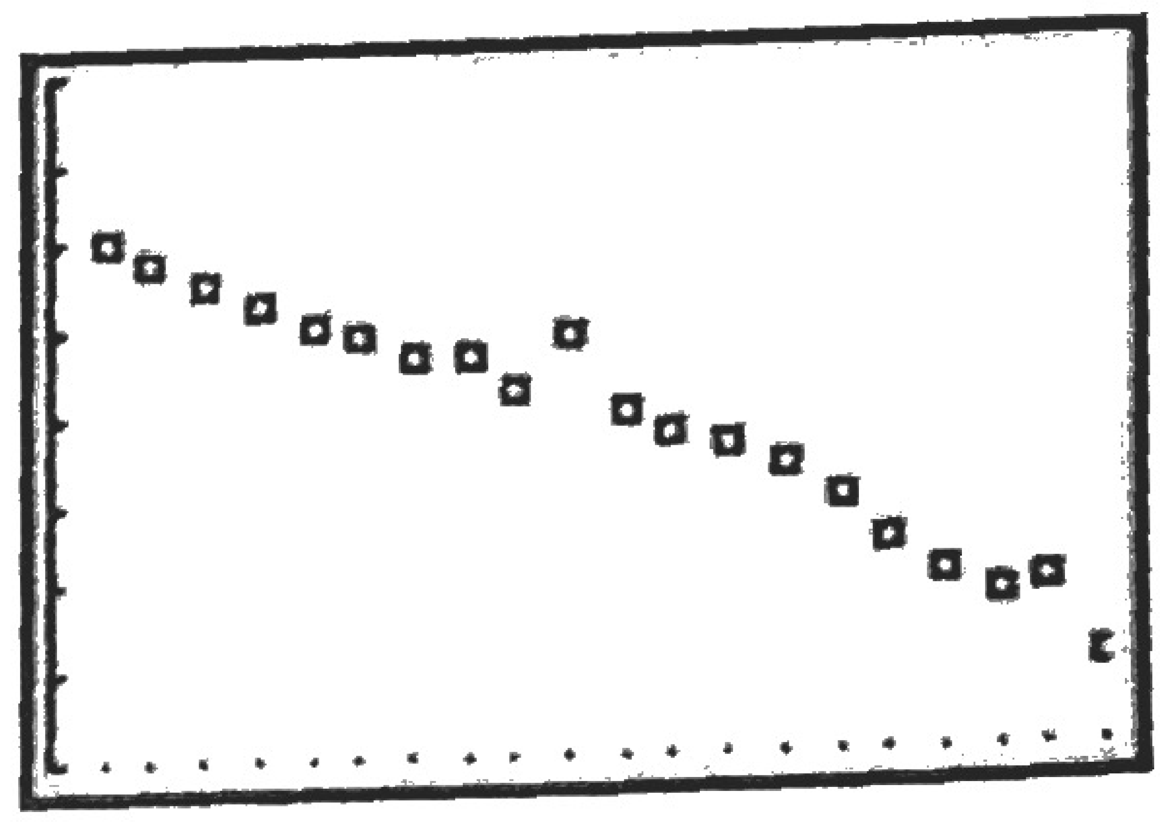
A) r = 0.972 B) r = 0.33 C) r = – 0.976



B) r = 0.33



A) r = 0.972



C) r = 0.976

6. REFLECTION: Does a strong correlation indicate a cause-and-effect relationship between variables? Give examples to justify your response.

A strong correlation does not indicate a cause & effect relationship between the variables. for example, the height & arm span data. just because a person has arm span greater than his/her height, this does not mean they are an olympic swimmer. strong correlation but not causation.