STAT 1350, 3/10 Discussion Questions

1. You calculate the correlation between height and weight for a simple random sample of 50 students from your college. Another student does the same for a simple random sample of 200 students from the college. The other student should typically get a higher, lower or the same correlation?

about the same

2. What does a scatterplot of data show us? How is a scatterplot different from a line graph?

relationship between 2 variables; points are not connected

3. The correlation between the heights of fathers and the heights of their (adult) sons is $r = 0.52$. This tells us that the relationship is strong moderate or weak? Positive, negative or zero?

positive, moderate

4. A study of grade school children finds that the correlation between hours of television watched per week during a school year and reading scores is $r = -0.63$. This tells us that the relationship is strong moderate or weak? Positive, negative or zero? As the number of hours of television watched increases, the reading score tends to do what?

moderate, negative

reading scores go down

For #3 and #4, what variable is the explanatory variable and which is the response variable?

3. height of fathers: explanatory; height of sons: response
4. hours watching television: explanatory; reading scores: response

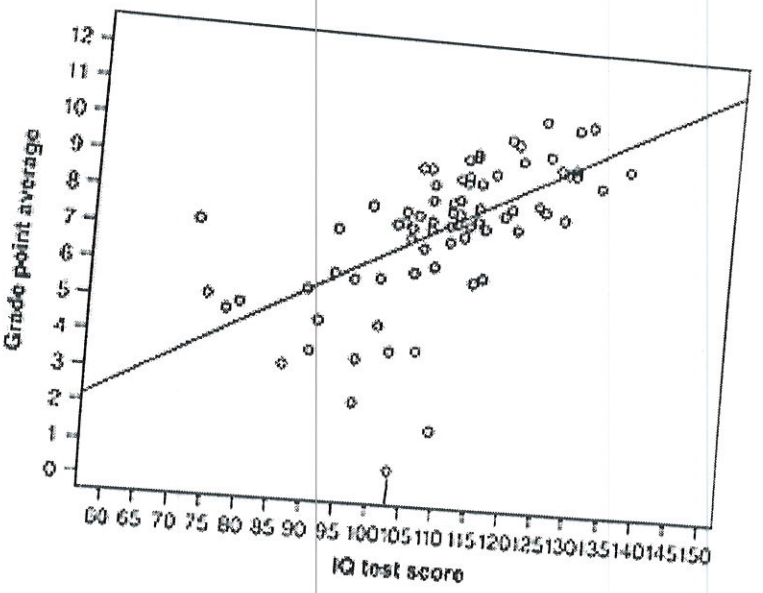
A study of home heating costs collects data on the size of houses and the monthly cost to heat the houses with natural gas. Here is the data:

Size of House	Heating Cost
1,200 sq ft	$150
2,300 sq ft	$375
1,800 sq ft	$270
2,000 sq ft	$315

5. Just by looking at the data (*don't* do a calculation) you can see that the correlation between house size and heating cost is positive, negative or near zero?

positive

An education researcher measured the IQ test scores of 78 seventh-grade students in a rural school, and also their school grade point average (GPA) measured on a 12-point scale. Here is a graph of GPA versus IQ for these students:



6. The name for this kind of graph is what?

Scatterplot

7. The IQ score of the student who has the lowest GPA is approximately what?

103

8. The graph shows the relationship between which two variables? Which is the explanatory variable and which is the response variable?

explanatory: IQ

response: GPA

9. Tall men tend to marry women who are taller than average, but the degree of association between the height of a husband and the height of his wife isn't very big. Give an example of a weak-to-moderate correlation value that would mean that taller-than-average men tend to marry taller-than-average women.

$r = .35$

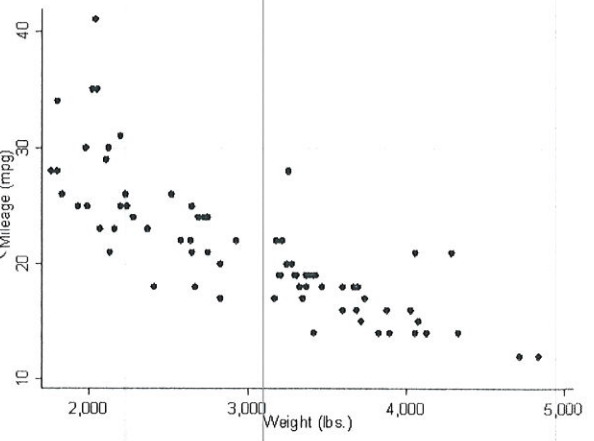
10. Here are the heights of a young girl at several ages, from a pediatrician's records:

Age in Months	36	51	60
Height in Centimeters	88	91	95

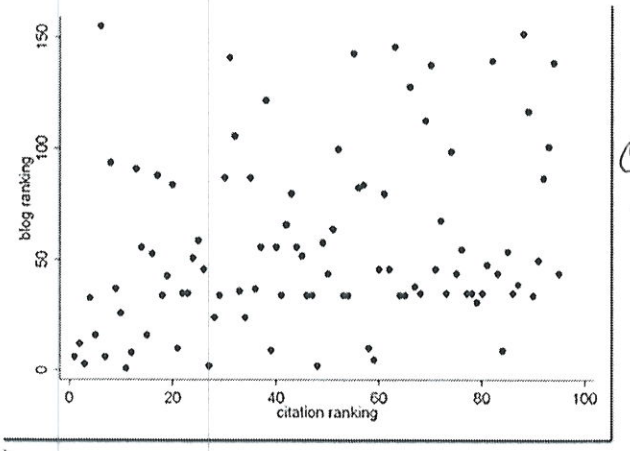
The correlation between the Age and Height variables is close to what value?

1

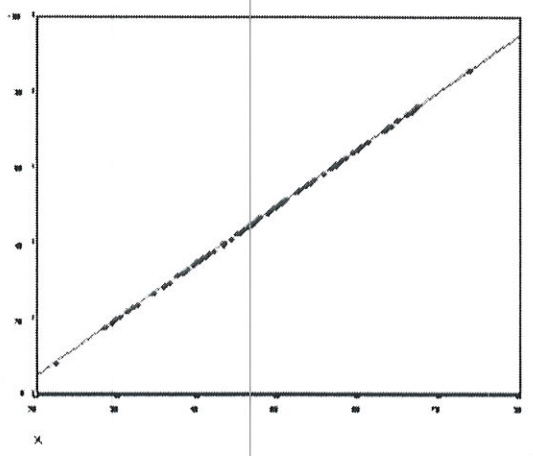
11. What does a correlation of exactly -1 mean for the slope of the graph of the data? What does it mean for the scatterplot? Which of the following scatterplots displays a relationship with a correlation closest to -1?



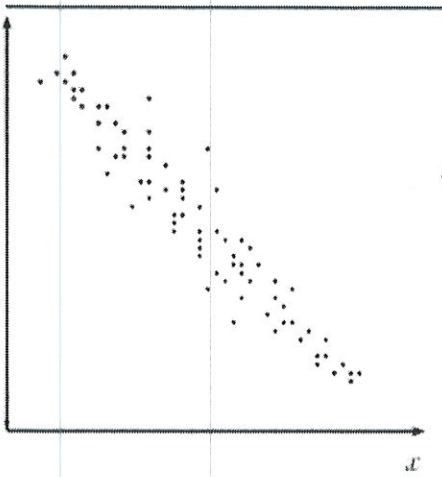
negative moderate



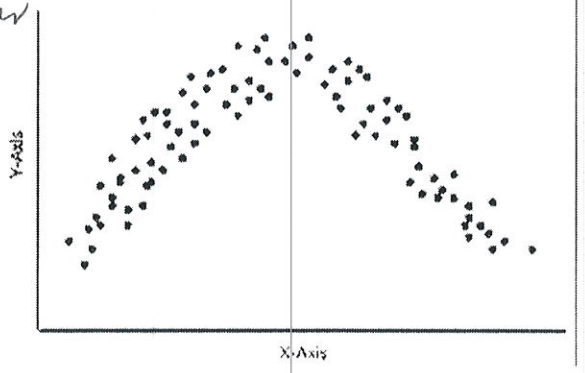
close to zero



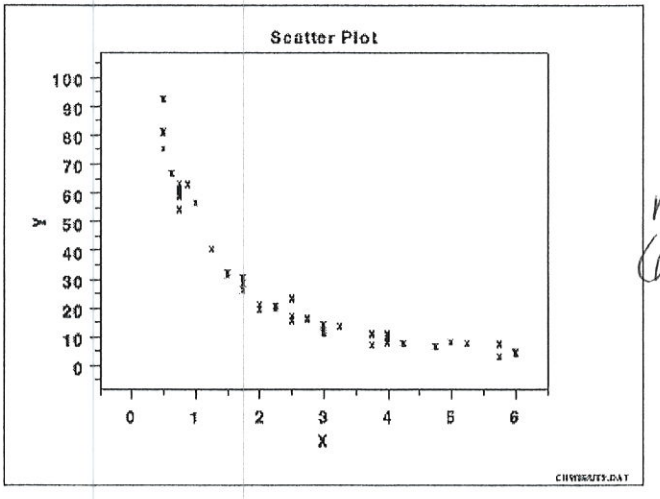
close to one (positive)



← closest to negative one



nonlinear close to zero



nonlinear (negative)

close to -1 means tightly compacted to line w/ negative slope

STAT 1350, 3/12 Discussion Questions

1. Consider a large number of countries around the world. There is a positive correlation between the number of Nintendo games per person x and the average life expectancy y. Does this mean that we could increase the life expectancy in Rwanda by shipping Nintendo games to that country? Why or why not?

No. The lurking variable is wealth. Wealthy countries can afford both games and more medical care.

2. Suppose that the correlation between the scores of students on Exam 1 and Exam 2 in a statistics class is $r = 0.7$. One way to use r is to compute the percent of the variation in Exam 2 scores can be explained by the straight-line relationship between Exam 2 scores and Exam 1 scores. What percent is this, approximately?

$r^2 = .49$ so 49%

3. A least-squares regression line is not just any line drawn through the points of a scatterplot. What is special about a least-squares regression line?

minimizes the total distance between the points and the line. (actually square of distance)

4. A report in a medical journal notes that the risk of developing Alzheimer's disease among subjects who (voluntarily) regularly took the anti-inflammatory drug ibuprofen (the active ingredient in Advil) was about half the risk among those who did not. Is this good evidence that ibuprofen is effective in preventing Alzheimer's disease? Why or why not?

Not really. Since people who were forgetting to take drugs probably had more severe problems to start w/.

5. A high correlation between two variables does not always mean that changes in one cause changes in the other. The best way to get good evidence that cause-and-effect is present is to do what?

Conduct a randomized comparative experiment

6. A study of child development measures the age (in months) at which a child begins to talk and also the child's score on an ability test given several years later. The study asks whether the age at which a child talks helps predict the later test score. The least- squares regression line of test score y on age x is $y = 110 - 1.3x$. According to this regression line, what happens (on the average) when a child starts talking one month later? [Hint: Interpret the slope in the context of the problem.]

IQ goes down by 1.3 points per month. (Score on ability test)

A recent campus bookstore survey sought to determine if there is a relationship between new textbook prices y (dollars) and number of pages in the book x. The resulting least-squares regression line for the study is $y = 38.04 + 0.12x$.

7. Suppose the textbook for your statistics course has 642 pages, how much would you predict it would cost new from the campus bookstore?

$38.04 + .12(642) = \$115.08$

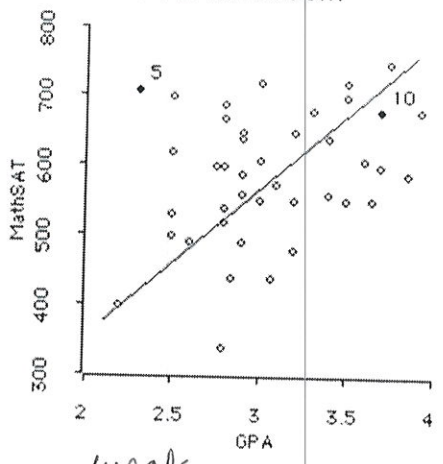
8. We can see from the equation of the line that, as the number of pages x goes up, what happens to y?

the price goes up by an average of 12 cents per page

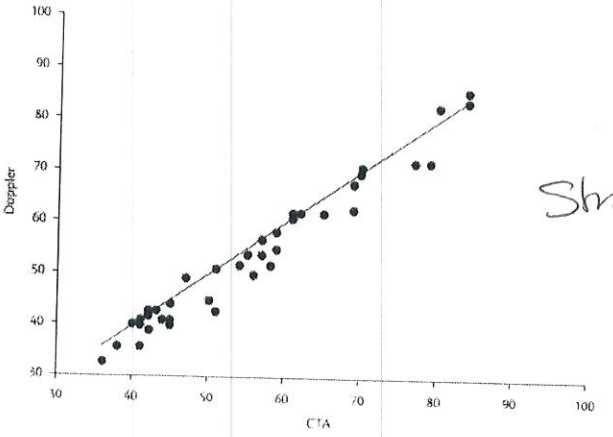
9. For each additional page in a textbook, the average price tends to change by how much?

12 cents

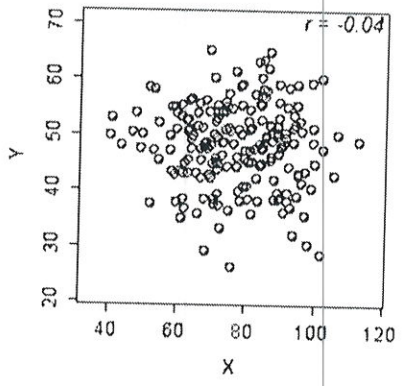
10. Which of the following graphs show a strong correlation? A moderate correlation? A weak correlation? No correlation?



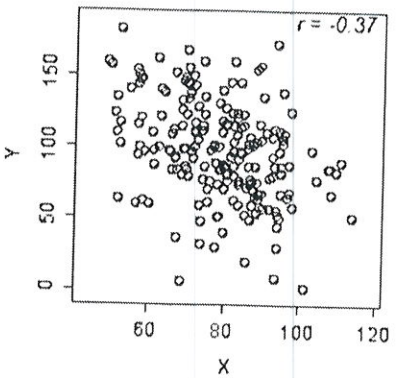
weak



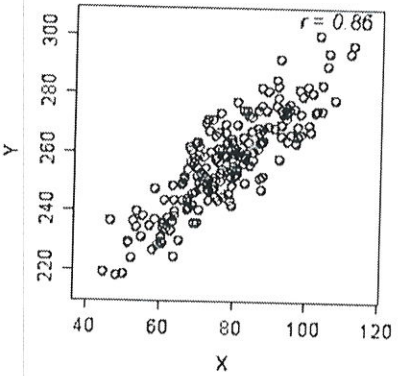
Strong



Near zero

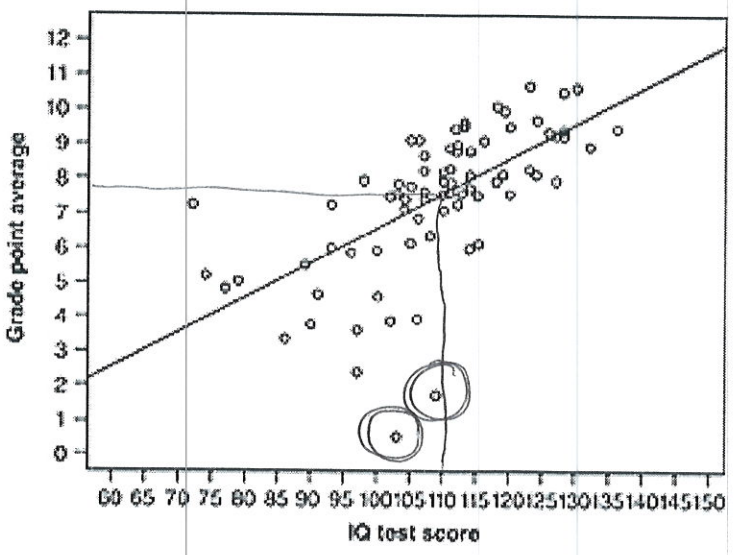


Weak or moderate



Strong

11. An education researcher measured the IQ test scores of 78 seventh-grade students in a rural school, and also their school grade point average (GPA) measured on a 12-point scale. Here is a graph of GPA versus IQ for these students:



The line drawn on the graph is the least-squares regression line of GPA on IQ. Use this line to predict the GPA of a student with IQ 110. Your prediction is GPA about how much based on the graph?

around 7.5

Does the graph appear to have any outliers?

yes, circled

12. The evidence that smoking causes lung cancer is very strong. But it is not the strongest possible statistical evidence because of how it was obtained. How was it obtained and why can't we get stronger data? observational study — we can't get stronger data because it is unethical to ask people to smoke if we think it is harmful