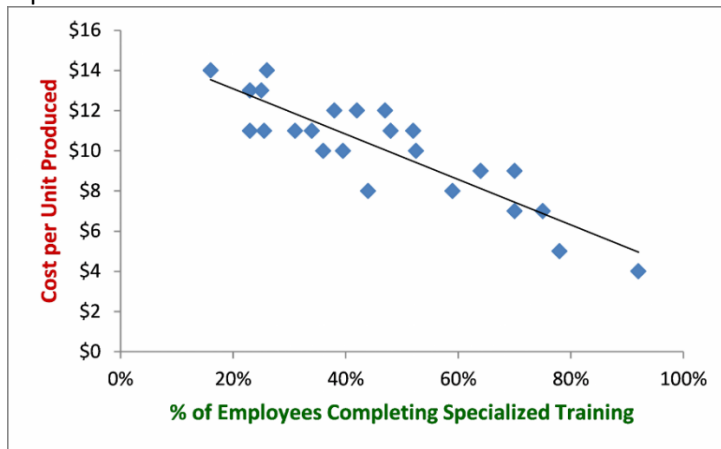


Instructions: Work problems on a separate sheet of paper and attach work to this page. You should show all work to receive full credit for problems. Questions with compact answers can be recorded directly on this page. Graphs and longer answers that won't fit here, indicate which page of the work the answer can be found on and be sure to clearly indicate it on the attached pages. You may use Excel to complete the problems, but then submit Excel files online.

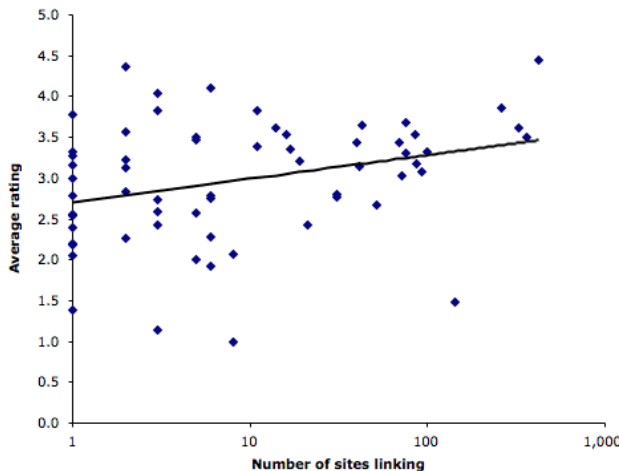
1. What does it mean when we say that: a) the correlation of a data set is exactly 1? b) the correlation of the data set is zero (or near zero)? c) the correlation is near -1?
2. The gas mileage of an automobile first increases as the speed increases and then decreases as the speed increases. Suppose that this relationship is very regular. Make a scatterplot of the mileage vs. speed data shown below and find the correlation (show that $r=0$). Explain why the correlation is zero even though there is a strong relationship?

Speed	25	35	45	55	65
Mileage	20	24	26	24	20

3. For each of the scatterplots below, estimate the correlation (for instance is it strong? Is it positive or negative? Is it close to zero?) The exact value isn't that important here. Are there any datapoints in the graph that might be throwing off the regression line? Is a linear model appropriate for the data?

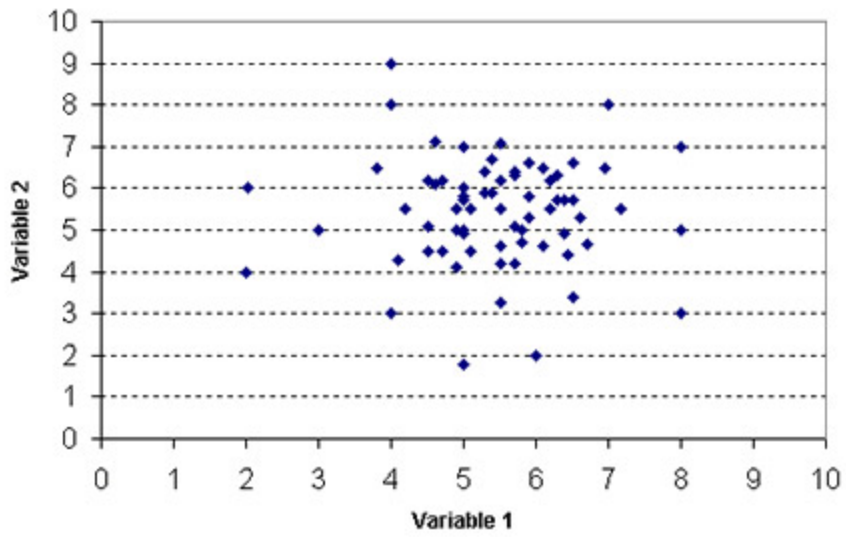


a)

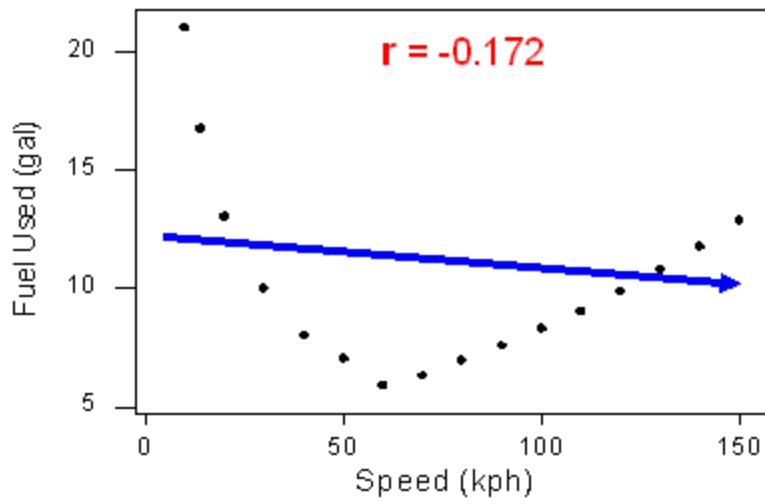


b)

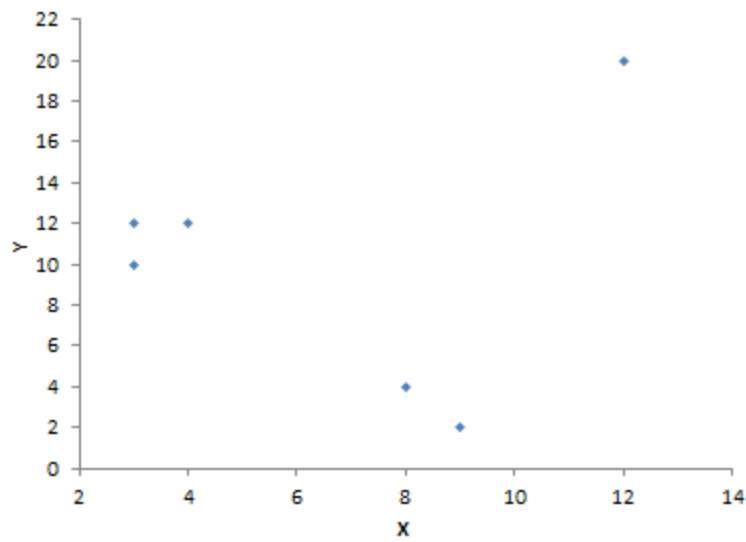
Figure 5. Scattered data points



c)



d)



e)

4. The correlation between the average SAT Mathematics score in the states and the proportion of high school seniors who take the SAT is $r = -0.843$. The correlation is negative. What does that tell us? How well does proportion taking predict average scores? (Hint: Use r^2 .)
5. The least squares line for predicting school GPA from IQ score, based on a sample of 78 students is $GPA = -3.56 + (0.101 \times IQ)$. Explain in words the meaning of the slope $b = 0.101$. Predict the GPA of a student with an IQ of 115.
6. If the correlation between two variables x and y is $r = 0$, there is no straight-line relationship between the variables. It turns out that the correlation is 0 exactly when the slope of the least-squares regression line is zero. Explain why slope 0 means there is no straight-line relationship between x and y . Start by drawing a line with slope 0 and explaining why in this situation x has no value for predicting y .
7. A study of class attendance and grades among first-year students at a state university showed that in general students who attended a higher percentage of their classes earned higher grades. Class attendance explained 25% of the variation in grade index among the students. What is the numerical value of the correlation between percentage of classes attended and grade index?
8. People who use low-calorie salad dressing tend to be heavier than those using regular-calorie salad dressing. Does this mean that using low-calorie dressing causes people to be heavier? Give a more plausible explanation for this association.
9. A February 2, 2008 article in the Columbus Dispatch reported a study on the distance students lived from campus and average GPA. Here is a summary of the results.

<i>Residence</i>	<i>Avg. GPA</i>
<i>Residence Hall</i>	3.33
<i>Walking Distance</i>	3.16
<i>Near campus (long walk or short drive)</i>	3.12
<i>Within the county (not near campus)</i>	2.97
<i>Outside the county</i>	2.94

Based on this data, the association between the distance a student lives from campus and GPA is negative. Many universities require freshman to live on campus, but this data has prompted some to suggest that sophomores should also be required to live on campus in order to improve grades. Does this data imply that living closer to campus actually does improve grades or might there be another factor that should be considered?

10. Go to the website of the Congressional Budget Office (CBO) <http://www.cbo.gov/> and find the prediction for a) the federal budget in five years, b) the federal budget deficit in 5 years, c) the federal debt in 5 years. Do any of these trends seem promising? What assumptions are we required to make when predicting into the future?
11. For each of the data sets below, determine the following:
 - i. Draw a scatterplot of the data
 - ii. Is the graph showing a linear or non-linear relationship

- iii. For linear graphs, find the least-squares regression line and the correlation coefficient.
- iv. What proportion of the data is explained by the linear relationship?
- v. Predict an appropriate measurement.
- vi. Give an example of a measurement you would not want to use with this model.
- vii. How strong is the relationship between the two variables?
- viii. Are there any obvious outliers?

- a. No tortilla chip aficionado likes soggy chips, so it is important to find characteristics of the production process that produce chips with an appealing texture. The following data on x =frying time (sec) and y =moisture content (%) appeared in an article on the subject.

x	5	10	15	20	25	30	45	60
y	16.3	9.7	8.1	4.2	3.4	2.9	1.9	1.3

- b. The following data on y =glucose concentration (g/L) and x =fermentation time (days) for a particular blend of malt liquor was read from a scatterplot in an article.

x	1	2	3	4	5	6	7	8
y	74	54	52	51	52	53	58	71

- c. The data below compared cricket chirps with temperature.

X	Y
20	88.6
16	71.6
19.8	93.3
18.4	84.3
17.1	80.6
15.5	75.2
14.7	69.7
17.1	82
15.4	69.4
16.2	83.3
15	79.6
17.2	82.6
16	80.6
17	83.5
14.4	76.3

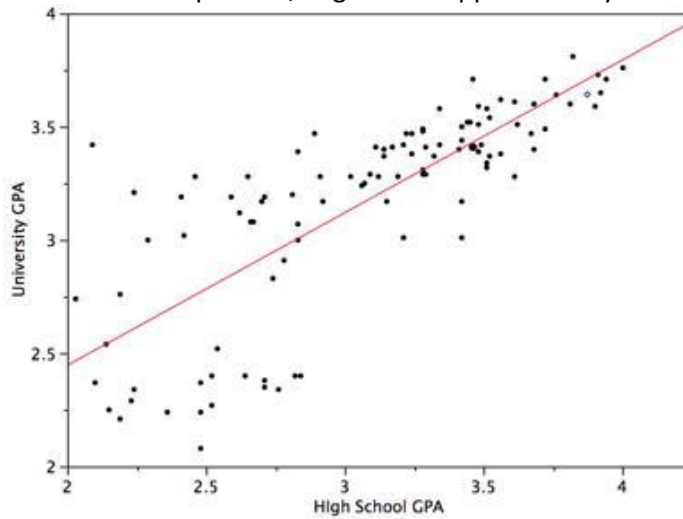
d.

Hours Spent Studying	Math SAT Score
4	390
9	580
10	650
14	730
4	410
7	530
12	600
22	790
1	350
3	400
8	590
11	640
5	450
6	520
10	690
11	690
16	770
13	700
13	730
10	640

e.

House Price in \$1000s (Y)	Square Feet (X)
245	1400
312	1600
279	1700
308	1875
199	1100
219	1550
405	2350
324	2450
319	1425
255	1700

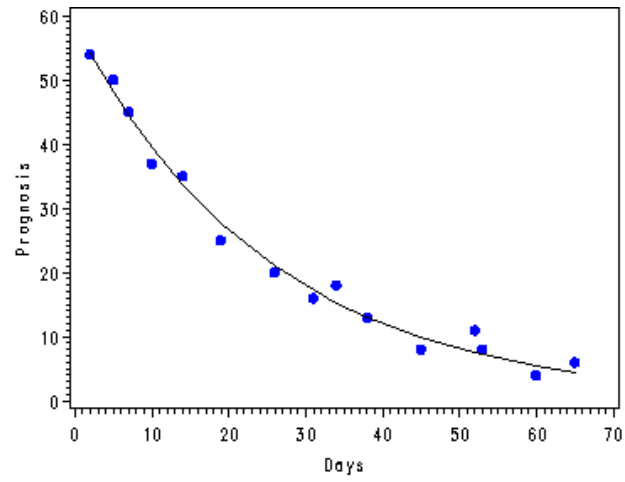
12. For each of the scatterplots below, determine if the best-fit equation is linear or non-linear. If linear, if the correlation positive, negative or approximately zero? Are there any outliers?



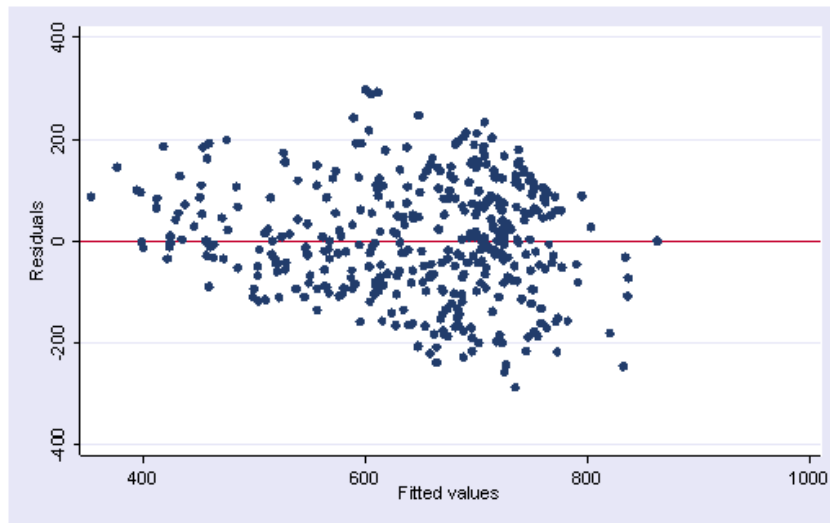
a.



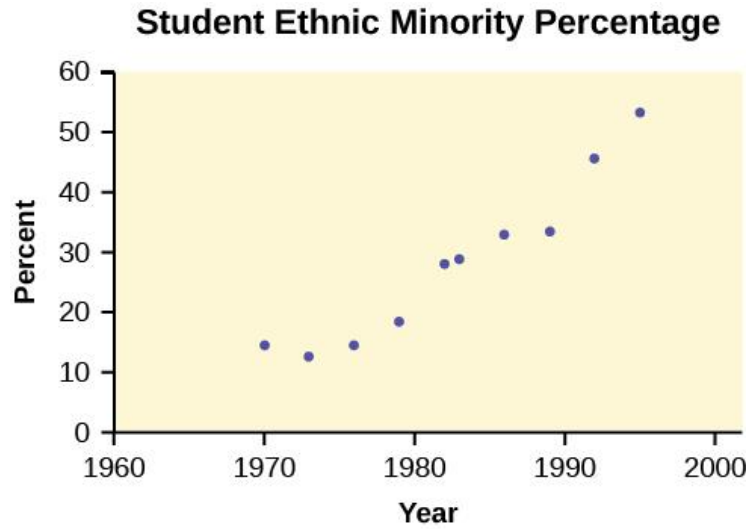
b.



c.



d.



e.

13. Use the data in the Excel file **154data7.xlsx** to answer the following questions.
- Create a scatterplot of the data. Be sure the axes are appropriately labeled and there is a good title.
 - Add a trendline to the graph and display the linear regression equation and R^2 value.
 - What is the correlation of the data?
 - Is the trend positive or negative?
 - Is a linear regression appropriate or is the data nonlinear?
 - Use the regression equation you found to predict the MPG for cars with the following weights: 2500, 3000, 4000 pounds.
 - Interpret the slope in the context of the problem.
 - What is the meaning of the R^2 value in context?