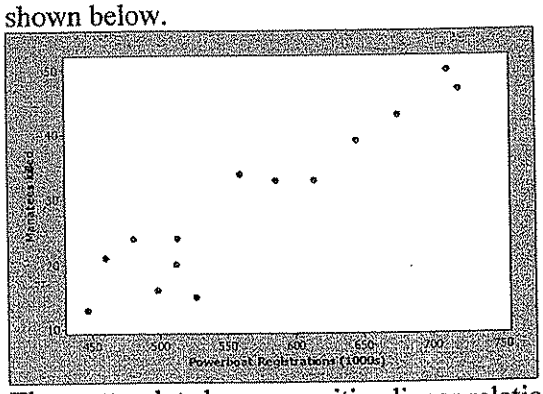
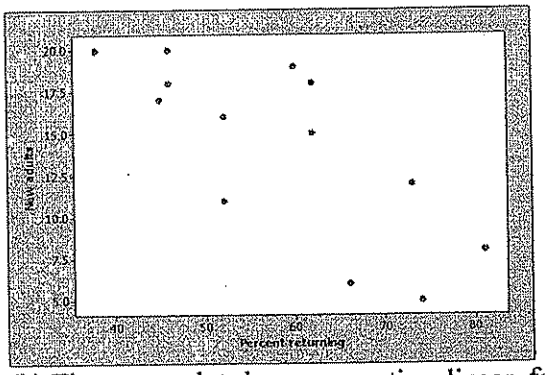
HW pp. 179 – 184 & pp. 193 – 195

3.5 (a) The explanatory variable is the number of powerboat registrations. (b) A scatterplot is shown below.



The scatterplot shows a positive linear relationship between these variables. (c) There is a positive linear association between powerboat registrations and manatees killed. (d) Yes, the relationship between these variables is linear. (e) The relationship is a strong, positive, linear association. Yes, the number of manatees killed can be predicted accurately from powerboat registrations. For 719,000 powerboat registrations, about 48 manatees would be killed by powerboats.

3.6 (a) A scatterplot is shown below.

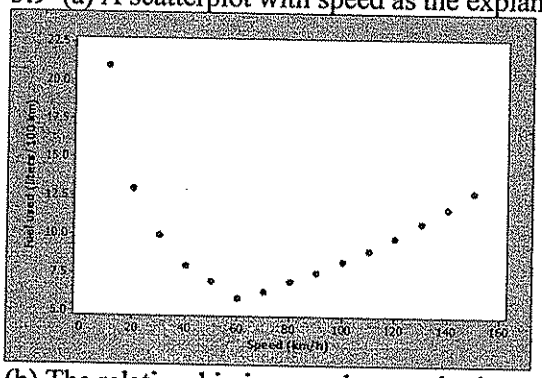


(b) The scatterplot shows a negative, linear, fairly weak relationship. (Note: direction=negative, form=linear, strength=weak.) (c) Because this association is negative, we conclude that the sparrowhawk is a long-lived territorial species.

3.7 (a) A positive association between IQ and GPA means that students with higher IQs tend to have higher GPAs, and those with lower IQs generally have lower GPAs. The plot does show a positive association. (b) The form of the relationship roughly linear, because a line through the scatterplot of points would provide a good summary. The positive association is moderately strong (with a few exceptions) because most of the points would be close to the line. (c) The lowest point on the plot is for a student with an IQ of about 103 and a GPA of about 0.5.

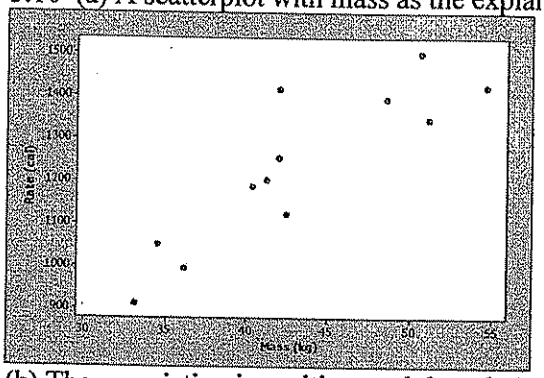
3.8 (a) From Figure 3.5, the returns on stocks were about 50% in 1954 and about −28% in 1974. (b) The return on Treasury bills in 1981 was about 15%. (c) The scatterplot shows no clear pattern. The statement that "high treasury bill returns tend to go with low returns on stocks" implies a negative association; there may be *some* suggestion of such a pattern, but it is extremely weak.

3.9 (a) A scatterplot with speed as the explanatory variable is shown below.

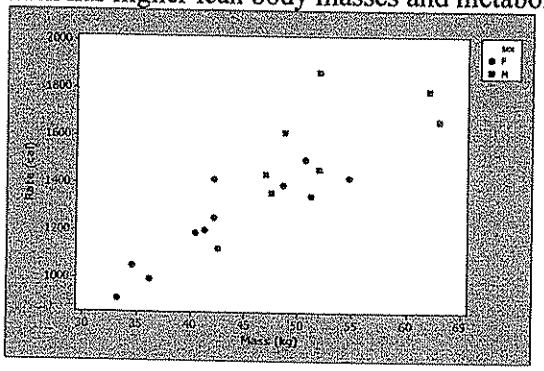


(b) The relationship is curved or quadratic. High amounts of fuel were used for low and high values of speed and low amounts of fuel were used for moderate speeds. This makes sense because the best fuel efficiency is obtained by driving at moderate speeds. (Note: 60 km/hr is about 37 mph) (c) Poor fuel efficiency (above average fuel consumption) is found at both high and low speeds and good fuel efficiency (below average fuel consumption) is found at moderate speeds. (d) The relationship is very strong, with little deviation for a curve that can be drawn through the points.

3.10 (a) A scatterplot with mass as the explanatory variable is shown below.



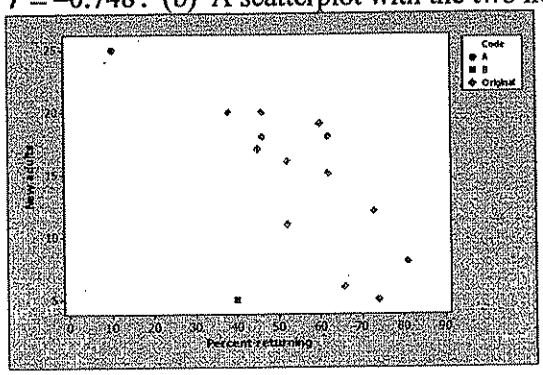
(b) The association is positive, and the relationship is linear and moderately strong. (c) The scatterplot below shows that the pattern of the relationship does hold for men. However, the relationship between mass and rate is not as strong for men as it is for women. The group of men has higher lean body masses and metabolic rates than the group of women.



3.15 (a) The lowest calorie count is about 107 calories and the sodium level for this brand is about 145 mg. The highest calorie count is about 195 calories, and the sodium level for this brand is about 510 mg. (b) The scatterplot shows positive association; high-calorie hot dogs tend to be high in salt, and low-calorie hot dogs tend to have low sodium. (c) The lower left point is an outlier. Ignoring this point, the relationship is linear and moderately strong.

3.16 (a) The correlation r is clearly positive but not near 1. The scatterplot shows that students with high IQs tend to have high grade point averages, but there is more variation in the grade point averages for students with moderate IQs. (b) The correlation r for the data in Figure 3.8 would be closer to one. The overall positive relationship between calories and sodium is stronger than the positive association between IQs and GPAs. (c) The outliers with moderate IQ scores in Figure 3.4 weaken the positive relationship between IQ and GPA, so removing them would increase r. The outlier in the lower left corner of Figure 3.8 strengthens the positive, linear relationship between calories and sodium, so removing this outlier would decrease r.

3.18 (a) The correlation between the percent of returning birds and the number of new adults is $r = -0.748$. (b) A scatterplot with the two new points added is shown below.



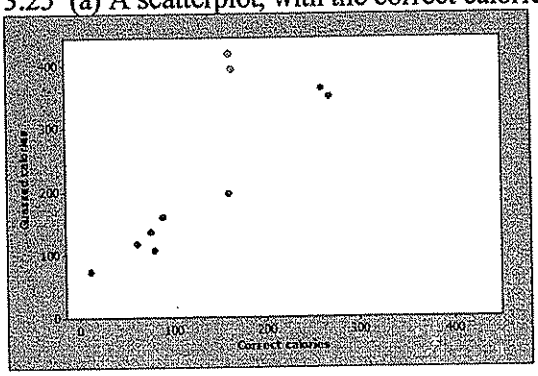
3.19 There is a perfect, positive association between the ages of the women and their spouses, so $r = 1$.

3.21 (a) New York's median household income is about $32,800 and the mean income per person is about $27,500. (b) Both of these variables measure the prosperity of a state, so you would expect an increase on one measure to correspond with an increase in the other measure. Household income will generally be higher than income per person because most households have one primary source of income and at least one other smaller source of income. (c) In the District of Columbia there are a relatively small number of individuals earning a great deal of money. Thus, the income distribution is skewed to the right, which would raise the mean per capita income above the median household income. (d) Alaska's median household income is about $48,000. (e) Ignoring the outliers, the relationship is positive, linear, and moderately strong.

3.23 (a) Gender is a categorical variable and the correlation coefficient r measures the strength of linear association for two quantitative variables. (b) The largest possible value of the correlation coefficient r is 1. (c) The correlation coefficient r has no units.

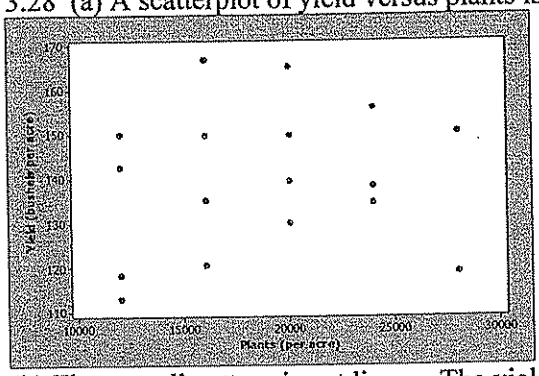
3.24 The paper's report is wrong because the correlation ($r = 0.0$) is interpreted incorrectly. The author incorrectly suggests that a correlation of zero indicates a negative association between research productivity and teaching rating. The psychologist meant that there is no linear association between research productivity and teaching rating. In other words, knowledge of a professor's research productivity will not help you predict her teaching rating.

3.25 (a) A scatterplot, with the correct calories as the explanatory variable, is shown below.



(b) There is a positive, linear relationship between the correct and guessed calories. The guessed calories for 5 oz. of spaghetti with tomato sauce and the cream-filled snack cake are unusually high and do not appear to fit the overall pattern displayed for the other foods. (c) The correlation is $r = 0.825$. This agrees with the positive association observed in the plot; it is not closer to 1 because of the unusual guessed calories for spaghetti and cake. (d) The fact that the guesses are all higher than the true calorie count does not influence the correlation. The correlation r would not change if every guess were 100 calories higher. The correlation r does not change if a constant is added to all values of a variable because the standardized values would be unchanged. (c) The correlation without these two foods is $r = 0.984$. The correlation is closer to 1 because the relationship is much stronger without these two foods.

3.28 (a) A scatterplot of yield versus plants is shown below.



(b) The overall pattern is not linear. The yield tends to be highest for moderate planting rates and smallest for small and large planting rates. There is clearly no positive or negative association between planting rates and yield. (d) The mean yields for the five planting rates are:

Planting rate	Mean
12000	131.025
16000	143.150
20000	146.225
24000	143.067
28000	134.750

A scatterplot with the means added is shown below. We would recommend the planting rate with the highest average yield, 20,000 plants per acre.