Statistics: Chapter 3 Solutions

Lesson 3.1.1

3-1. a. Explanation should relate to the difference between (175 – 110) and (28 + 44).

b. 65 students. Subtract 110 from 175 or add together 21, 7, and 37.

c. Because some students have both and will be counted twice.

3-2. a.

	πPad	no πPad	totals
πPhone	**7**	**21**	28
no πPhone	**37**	110	**147**
totals	44	**131**	175

b. They are similar in that they are both ways to display the information. They are different in that the table lists the totals for each category. In a Venn diagram it is very easy to see and sort the data. In the table you must figure out which row/column to look at, but it displays more information with the margins.

c. *i*. $\frac{7+21+37}{175} = 37.1\%$

ii. $\frac{7}{175} = 4\%$

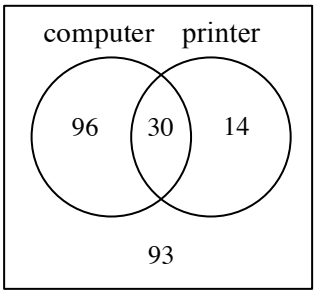
3-3. a.

	Job	No job	
License	99	71	170
No license	9	21	30
	108	92	200

b. P(job and no license) = $\frac{9}{200} = 0.045$

c. P(license given has job) = $\frac{99}{108} = 0.92$

3-4. a. See diagram at at right.



b. P(computer or printer) = $\frac{140}{233} = 0.60$

c. P(printer given computer) = $\frac{30}{126} = 0.238$

3-5. a. P(C or < 20) = $\frac{228}{800} + \frac{200}{800} - \frac{126}{800} = \frac{302}{800} = 0.3775$, or about 38%.

b. Add the number of participants selecting Soda A:
P(A) = $\frac{30+67+88+141}{800} = \frac{326}{800} = 0.4075$, or about 41%.

c. Take only the participants over 60 who selected Soda A out of all those selecting Soda A. P(> 59 given that A) = $\frac{141}{30+67+88+141} \approx 0.43$, or about 43%.

3-6. a. $P(\text{did not check baggage}) = \frac{681}{1000} = 0.681$

b. $P(\text{checked bag or traveling for business}) = \frac{216+103+387}{1000} = 0.706$

c. $P(\text{no checked bag given traveling for business}) = \frac{387}{103+387} = 0.790$

3-7. a. See table at right.
$P(\text{washer or dryer}) = \frac{44+44+20}{177} \approx 61\%$

b. $P(\text{dryer given that washer}) = \frac{44}{88} = 50\%$

	purchase washer	did not purchase washer	
purchase dryer	44	20	64
did not purchase dryer	44	69	113
	88	89	177

3-8. a. $y = x + 2$

b. 19 grams

c. Not confident. It is an extrapolation.
A length of 3 inches is well outside the range of data.

3-9. Moderate negative linear association with no outliers. The data appear to be in two clusters, probably indicating two classes of vehicles.

3-10. a. $b = r\left(\frac{S_y}{S_x}\right) = 1.0563$. Using the means of the explanatory and response variables as a point on the LSRL, $5.04 = a + 1.0563(10.79)$. Solving for $a = -6.36$ cm, making the LSRL equation $\hat{y} = -6.36 + 1.0563x$, where x is root length and $\hat{y}$ is predicted root diameter.

b. For every one foot increase in root length the predicted value of root diameter will increase by 1.0563 cm.

c. $\hat{y} = -6.36 + 1.0563(9.40) = 3.57$ cm

d. The predicted value at $x = 9.9$ feet is $\hat{y} = -6.36 + 1.0563(9.9) = 4.10$ cm. The residual is $y_{\text{observed}} - y_{\text{predicted}} = 4.8 - 4.10 = 0.70$ cm.

e. There is no discernable pattern in the residual plot, confirming that a linear model is most appropriate.

3-11. a. $r = 0.8738$ indicating a strong positive linear association between the tail length versus mass of giant pandas, however, the residual plot shows that a curved model is more appropriate.

b. $r^2 = 0.7635$, meaning that 76% of the variation observed in tail length is explained by a linear relationship with mass. In other words, knowing the giant panda mass associated with each tail length reduces the error in predicting tail length by 76% over using just the mean tail length as a predictor.