Lesson 3.1.2

3-12. The two conditional probabilities are nearly identical ($\frac{7}{44} \approx \frac{21}{131} \approx 16\%$). Therefore there is no association between owning a πPad and owning a πPhone: you are just as likely to own a πPhone whether you own a πPad or not.

3-13. a.

	Boy	Girl	
A	24	7	31
not A	92	22	114
	116	29	145

b. The conditional probability of a boy receiving an A is 21%, while the conditional probability of a girl receiving an A is 24%. So there *is* a weak association between gender and grades.

c.

	Boy	Girl
A	21%	24%
not A	79%	76%

3-14. See table at right. There is no association between blood type and the DapT defect. A person has the same probability of any blood type, regardless of whether they have DapT or not. For example, if you have the DapT defect, there is a 44% likelihood you are Type O and a 4% likelihood you are Type AB; if you do not have the DapT defect, your probabilities for being Type O or Type AB are also 44% and 4% respectively.

	has DapT	no DapT
Type O	$\frac{0.30}{0.68} = 44\%$	$\frac{137}{311} = 44\%$
Type A	41%	42%
Type B	10%	10%
Type AB	4%	4%

3-15. a.

	early curfew	late curfew
does chores	50%	0%
does not do chores	50%	100%

b. Yes, students who had an early curfew did chores 50% of the time, while none of the students who had a late curfew did chores.

c. See table at right. There is still an association: all of students who did chores had an early curfew, while only half of the students who did not do chores had an early curfew.

	does chores	does not do chores
early curfew	9 (100%)	9 (50%)
late curfew	0 (0%)	9 (50%)

d. curfew explanatory:

	early curfew	late curfew
does chores	$33\frac{1}{3}\%$	$33\frac{1}{3}\%$
does not do chores	$66\frac{2}{3}\%$	$66\frac{2}{3}\%$

chores explanatory:

	early curfew	late curfew
does chores	60%	40%
does not do chores	60%	40%

There is no association in either case. If curfew is the explanatory variable, $\frac{1}{3}$ of students with an early curfew do chores, while $\frac{1}{3}$ of students with a late curfew do chores. $\frac{1}{3}$ of students do chores regardless of whether they have an early curfew; there is no association between chores and curfew. If chores are the explanatory variables, $\frac{2}{5}$ of students who do chores have an early curfew, while also $\frac{2}{5}$ of students who do not have chores have an early curfew. So $\frac{2}{5}$ of students have an early curfew whether they do chores or not; there is no association between chores and curfew.

e. All of the rows (or columns), including the totals row, have the same percentages.

3-16. a. It has two variables: Age and living status.

b. $P(\text{live alone}) = \frac{1.45+4.16+3.55+5.57+6.69+4.81+6.50}{312.59} = 10.5\%$

c. $P(\text{over 65}) = \frac{4.81+6.50+16.59+11.26}{312.59} = 12.5\%$

d. Using the complement. 100% – 12.5% = 87.5%

e. $P(\text{under 35 and live alone}) = \frac{1.45+4.16}{312.59} = 1.8\%$

f. Qui counted twice the 1.45 million and 4.16 million people in the top row. He should have computed:
$\frac{\text{first two columns + first row – people counted twice}}{\text{total}} = \frac{111.12+41.57+32.73-1.45-4.16}{312.59} = 57.5\%.$

g. See the conditional relative frequency table below. There is an association. If you are older, you are much more likely to be living alone. For example, only 10% of people 25 to 34 years old live alone, while 37% of people over 75 years live alone.

	Age						
	Under 25	25 to 34	35 to 44	45 to 54	55 to 64	65 to 74	Over 75
Living alone	$\frac{1.45}{111.12} = 1\%$	$\frac{4.16}{41.57} = 10\%$	9%	13%	18%	22%	37%
Living with others	99%	90%	91%	87%	82%	78%	63%

3-17. a. $P(\text{no backpack}) = \frac{3+6+14+16}{100} = 39\%$ or 0.39

b. $P(\text{backpack given junior}) = \frac{18}{32} = 56\%$

c. $P(\text{junior or senior given no backpack}) = \frac{14+16}{3+6+14+16} = 77\%$

d. Conditional relative frequency table:

	Freshmen	Sophomore	Junior	Senior
Backpack	73%	73%	56%	54%
No Backpack	27%	27%	44%	46%

Yes, the juniors and seniors are much less likely to be carrying a backpack.

3-18. a. $\hat{y} = 5.353 - 1.561x$, where $\hat{y}$ is the number of days a cold lasted and x is the number of months a vitamin supplement was taken.

b. Yes, $r = -0.9512$ and there is a linear pattern in the data scatterplot.

c. $R^2 = 90.5\%$. 90.5% of the variability in the length of a cold can be explained by a linear relationship with the amount of time taking supplements.

d. A scatterplot confirms the relationship is linear. It is negative with a slope of –1.561, so an increase in one month of supplements is expected to decrease the length of a cold by 1.561 days. $r = -0.951$. The association is strong, negative, and linear. There are no apparent outliers.

3-19. a. $r = 0.1949$, indicating practically no linear association between the length versus weight of Siberian tigers, which is not contradicted by the residual plot.

b. $r^2 = 0.0380$, meaning that 4% of the variation observed in length is explained by a linear relationship with weight. In other words, knowing the weight associated with each length reduces the error in predicting length by 4% over using just the mean length as a predictor.

c. When making predictions about Siberian tiger lengths by using weights, one can expect to typically be off by 0.337 ft.