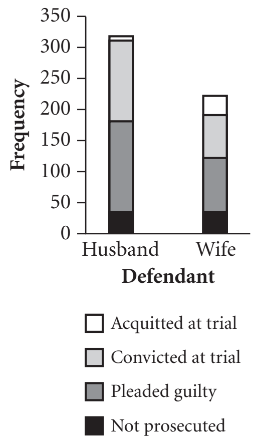
a. There are two populations: husbands who are accused of murdering their wives and wives who are accused of murdering their husbands.

b. A segmented bar graph is suitable.



From the chart it appears that the proportions of husbands and the proportions of wives who fall into each category could be about the same. (The proportions in the acquitted and convicted categories look a bit different, though. More husbands are accused of murdering their wives than wives accused of murdering their husbands.) However, you will see when you finish the chi-square test that the relatively large sample sizes result in a statistically significant difference.

c. *Check conditions:* This situation doesn't quite fit the criteria for a chi-square test of homogeneity. These are not random samples from the populations of husbands accused and wives accused, but all of the cases from the largest counties in one year. Nevertheless, you can reasonably proceed with a chi-square test by thinking of these as two distinct populations and asking only whether the difference can reasonably be attributed to chance. Each case is resolved in exactly one of four ways. The expected number of cases in each cell is 5 or more.

State your hypotheses:

H_0: The difference between the proportion of accused husbands who fall into each of the four categories and the proportion of all accused wives who fall into that category can reasonably be attributed to chance.

H_a: The differences cannot reasonably be attributed to chance. Husbands who fall into that category is not the same as the proportion of wives who fall into that category.

Compute the test statistic. The test statistic is

$$\chi^2 = 0.939 + 1.345 + 0.563 + 0.806 + 1.401 + 2.006 + 0.567 + 15.137 = 32.765$$

Comparing the test statistic to the χ^2 distribution with 3 degrees of freedom, the value of χ^2 from the sample, 32.765, is far out in the tail. The *P*-value is close to 0 (about $3.6 \cdot 10^{-7}$). The test statistic is quite large and would be difficult to see in a sketch.

Write a conclusion in context: You should reject the null hypothesis. You cannot attribute the differences in the ways the cases were resolved to chance alone. A value of χ^2 this large is extremely unlikely to occur if cases are resolved in the same proportions for husbands and wives. You conclude that the proportion in at least one category is different for husbands than for wives. However, you must note that this is not a random sample from a larger population, so all you can be sure about is that in these 75 largest counties in 1988, the distributions of outcomes for husbands and wives do not look like they are random samples from identical populations.

E55. **a.** Description B and Design I
b. Description A and Design III
c. Description C and Design II

Note: Design IV is also a description of a chi-square test of homogeneity, but for a design we don't cover in this textbook that results in a three-way table.

AP Sample Test

AP1. B. For the 136 left-handed children, 64 of the fathers are left-handed. If the variables are independent, we'd expect the same proportion of the 455 right-handed children to have left-handed fathers. Hence $455 \cdot \frac{64}{136}$ is the expected frequency.

AP2. B. There are 8 rows and 2 columns in the table, and thus $(8 - 1)(2 - 1) = 7$ degrees of freedom.

AP3. B. There are 15 days within at most one week from a given date out of 365 days, so the expected frequency is $200 \cdot \frac{15}{356}$, or about 8.219. Expected counts shouldn't be rounded more than necessary. (If students use 14 days, a minor error, they will get an expected frequency of about 7.67 and so will choose A.)

AP4. A. The design of the experiment calls for a matched pairs *t*-test.

AP5. B. We have 50 independent samples, from the 50 state populations.

AP6. A

AP7. C (Specifically, the population proportions must be different for either two or three of the three choices.)

AP8. B

AP9 **a.** From the plot, it looks as if seat belts do make a difference in surviving a crash. The seat belts used part of the *survived* bar is a much greater part of the bar than the same category in the *killed* bar. The conditional proportions are 73.5% of those who survived automobile accidents wore seat belts whereas 34.3% of those who did not survive automobile accidents wore seat belts. But you need to do a test to see if this difference is statistically significant. The observed and expected frequencies are shown.

	Seat Belts Used	Seat Belts Not Used	Total
Killed	Obs: 35 Exp: 74.83	Obs: 67 Exp: 27.17	**102**
Survived	Obs: 16,694 Exp: 16,654.17	Obs: 6,008 Exp: 6,047.83	**22,702**
Total	**16,729**	**6,075**	**22,804**

Check conditions: These data came from a random sample of accidents collected by the National Highway Traffic Safety Administration. There is only one population of all accidents, so a chi-square test of independence would be the most appropriate. Each front seat occupant can be classified into only one cell of the table, and all of the expected frequencies are at least 5 as shown above.

State your hypotheses:

H_0: Seat belt use and surviving an automobile accident are independent.

H_a: Seat belt use and surviving an automobile accident are not independent.

Compute the test statistic, find the P-value, and draw a sketch:

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

$$= \frac{(35-74.83)^2}{74.83} + \frac{(67-27.17)^2}{27.17} + \frac{(16{,}694-16{,}654.17)^2}{16{,}654.17} + \frac{(6{,}008-6{,}047.83)^2}{6{,}047.83}$$

$$\approx 79.930$$

Comparing the test statistic to a χ^2 distribution with 1 degree of freedom, you see that the value of χ^2, 79.93, is far out in the tail. The *P*-value is approximately 0 and would be hard to picture in a sketch.

Write your conclusion in context, linked to your computations: Reject the null hypothesis that seat belt use and surviving an automobile accident are independent. If they were, a random sample of this size would be extremely unlikely to produce a χ^2 value of 79.93 or higher. You have strong evidence that these variables are associated and the difference is highly significant. Seat belts appear to be highly effective in reducing the death rate.

b. From the plot, it looks as if air bags do make a difference in surviving a crash. The *air bags* part of the *survived* bar is a larger part of the bar than the same category in the *killed* bar. The conditional proportions are 54.2% of those who survived automobile accidents had air bags whereas 41.2% of those who did not survive automobile accidents had air bags. But you need to test to see if this difference is statistically significant. The observed and expected frequencies are shown in the table.

	Air Bags	No Air Bags	Total
Killed	Obs: 42 Exp: 55.27	Obs: 60 Exp: 46.73	**102**
Survived	Obs: 12,315 Exp: 12,301.73	Obs: 10,387 Exp: 10,400.27	**22,702**
Totals	**12,357**	**10,447**	**22,804**

Check conditions: Conditions were checked in part a. Each front seat occupant can be classified into only one cell of the table, and all of the expected frequencies are at least 5 as shown above. Conditions are met for a chi-square test of independence.

State your hypotheses:

H_0: Having an air bag and surviving an automobile accident are independent.

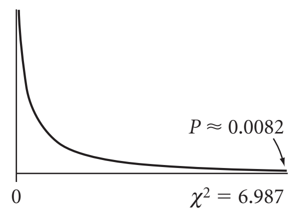
H_a: Having an air bag and surviving an automobile accident are not independent.

Compute the test statistic, find the P-value, and draw a sketch:

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

$$= \frac{(42-55.27)^2}{55.27} + \frac{(60-46.73)^2}{46.73} + \frac{(12{,}315-12{,}301.73)^2}{12{,}301.73} + \frac{(10{,}387-10{,}400.27)^2}{10{,}400.27}$$

$$\approx 6.987$$



Comparing the test statistic to a χ^2 distribution with 1 degree of freedom, you see that the value of χ^2, 6.987, is in the tail. The *P*-value is approximately 0.0082.

Write your conclusion in context, linked to your computations: You should reject the null hypothesis that having an air bag and surviving an automobile accident are independent. If they were, a random sample of this size would be unlikely to produce a χ^2 value of 6.987 or higher. With a *P*-value of only 0.0082, you have strong evidence that these variables are associated and the difference is statistically significant at the 0.01 level. Air bags appear to be effective in reducing the death rate, but less so than seat belts.

c. Both air bags and seat belts are effective in reducing death rates although from this sample, it appears that seat belts are more effective in saving lives than air bags.

AP10. a. From the plot, it looks as if air bags do make a small difference in surviving a crash. The *air bags used* part of the *survived* bar is a larger part of the bar than the category in the *killed* bar. The conditional proportions are 62.7% of those who were wearing seat belts and survived automobile accidents had air bags whereas 54.3% of those who were wearing seat belts and did not survive automobile accidents had air bags. But you need to test to see if this difference is statistically significant. The observed and expected frequencies are shown in the Fathom printout.

From Summary Statistics — Test for Independence

First attribute (categorical): unassigned
Second attribute (categorical): unassigned

		Air Bags: Yes	Air Bags: No	Row Summary
Survived	No	19 (21.9)	16 (13.1)	35
	Yes	10464 (10461.1)	6230 (6232.9)	16694
Column Summary		10483	6246	16729

```
First attribute:          Air Bags
  Number of categories:   2
Second attribute:         Survived
  Number of categories:   2
```
Ho: Air Bags is independent of Survived
```
Chi-square:   1.052
DF:           1
P-value:      0.3
```
The numbers in parentheses in the table are expected counts.

Check conditions: Conditions were checked in part a. Each outcome can be classified into only one cell of the table, and all of the expected frequencies are at least 5 as shown above. Conditions are met for a chi-square test of independence.

State your hypotheses:

H_0: Having an air bag (while using seat belts) and surviving an automobile accident are independent.

H_a: Having an air bag (while using seat belts) and surviving an automobile accident are not independent.

Compute the test statistic, find the P-value, and draw a sketch:

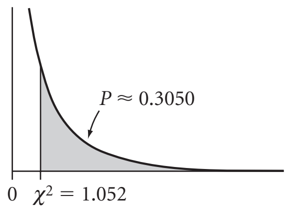
$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

$$= \frac{(19-21.93)^2}{21.93} + \frac{(16-13.07)^2}{13.07}$$

$$+ \frac{(10{,}464-10{,}461.07)^2}{10{,}461.07}$$

$$+ \frac{(6{,}230-6{,}232.93)^2}{6{,}232.93}$$

$$\approx 1.052$$



Comparing the test statistic to a χ^2 distribution with 1 degree of freedom, you see that the value of χ^2, 1.052, is not in the tail. The *P*-value is approximately 0.3050.

Write your conclusion in context, linked to your computations: You should not reject the null hypothesis that having an air bag and surviving an automobile accident are independent when seat belts are also used. These results are typical if there was no association between air bags and surviving an automobile accident when also wearing a seat belt. There is no statistically significant evidence that air bags add any advantage in the presence of seat belts.

b. From the plot, it looks as if air bags do not make a difference in surviving a crash when also not using a seat belt. In fact, it looks as if the opposite is true. The *air bags* part of the *survived* bar is a smaller part of the bar than the same category in the *killed* bar. The conditional proportions are 30.8% of those who weren't wearing seat belts and survived automobile accidents had air bags whereas 34.3% of those who weren't wearing seat belts and did not survive automobile accidents had air bags. You should still test to see if this difference is statistically significant. The observed and expected frequencies are shown in the Fathom table.

From Summary Statistics — Test for Independence

First attribute (categorical): unassigned
Second attribute (categorical): unassigned

		Air Bags: Yes	Air Bags: No	Row Summary
Survived	No	23 (20.7)	44 (46.3)	67
	Yes	1851 (1853.3)	4157 (4154.7)	6008
Column Summary		1874	4201	6075

```
First attribute:          Air Bags
  Number of categories:   2
Second attribute:         Survived
  Number of categories:   2
Ho: Air Bags is independent of Survived
Chi-square:   0.3847
DF:           1
P-value:      0.54
```

The numbers in parentheses in the table are expected counts.

Check conditions: Conditions were checked in part a. Each outcome can be classified into only one cell of the table, and all of the expected frequencies are at least 5 as shown above. Conditions are met for a chi-square test of independence.

State your hypotheses:

H_0: Having an air bag (while not using seat belts) and surviving an automobile accident are independent.

H_a: Having an air bag (while not using seat belts) and surviving an automobile accident are not independent.

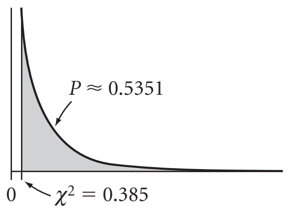
Compute the test statistic, find the P-value, and draw a sketch:

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

$$= \frac{(23-20.67)^2}{20.67} + \frac{(44-46.33)^2}{46.33}$$

$$+ \frac{(1851-1853.33)^2}{1853.33} + \frac{(4157-4154.67)^2}{4154.67}$$

$$\approx 0.385$$



Comparing the test statistic to a χ^2 distribution with 1 degree of freedom, you see that the value of χ^2, 0.385, is not in the tail. The P-value is approximately 0.5351.

Write your conclusion in context, linked to your computations: You should not reject the null hypothesis that having an air bag and surviving an automobile accident are independent when seat belts are not used. These results are typical if there was no association between air bags and surviving an automobile accident when also not wearing a seat belt. There is no statistically significant evidence that air bags add any advantage even when seat belts are not used. In fact, the observed proportion of those killed with air bags is slightly higher than the observed proportion of those killed without air bags. Perhaps this issue should be studied further.