a. **Response: opinion on the best way to fight terrorism (let the terrorists know we will fight back aggressively, to work with other nations to find an international solution) Explanatory: gender (male, female)**

3.6 How to fight terrorism? A survey of 1000 adult Americans (*Rasmussen Reports*, April 15, 2004) asked each whether the best way to fight terrorism is to let the terrorists know we will fight back aggressively or to work with other nations to find an international solution. The first option was picked by 53% of the men but by only 36% of the women in the sample. Assume there were 600 men and 400 women in the sample.

a. Identify the response variable and the explanatory variable, and their categories.

b. Construct a contingency table (similar to Table 3.1) that shows the counts for the different combination of categories.

c. Use a contingency table to display the percentages for the two options, separately for females and for males.

d. Explain why the percentages reported here are *conditional* percentages.

e. Give an example of how results would have to differ from these for you to conclude that there's *no* evidence of association between these variables.

b.

	Opinion on how to fight terrorism		
Gender	Fight back	International	Total
Male	318	282	600
Female	144	256	400

c.

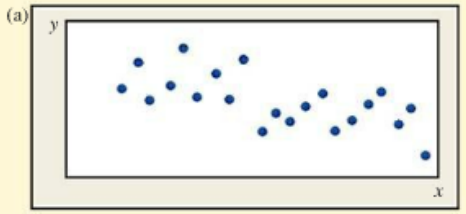
	Opinion on how to fight terrorism		
Gender	Fight back	International	Total
Male	53%	47%	100%
Female	36%	64%	100%

d. **They are the percentages of people have a certain opinion GIVEN their gender.**

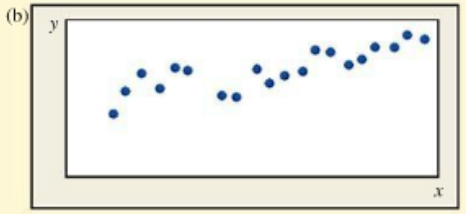
e. **Same percentages for the opinion for both genders. For example 53% of both male and female state to fight back.**

3.16 Match the scatterplot with *r* Match the scatterplots below with the correlation values.

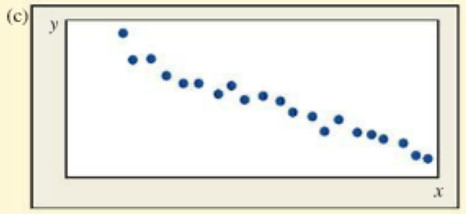
1. $r = -0.9$
2. $r = -0.5$
3. $r = 0$
4. $r = 0.6$



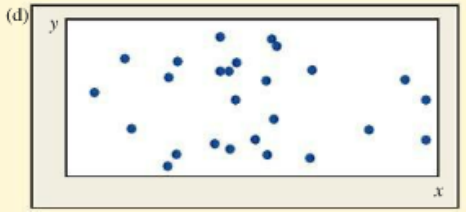
r = - 0.5



r = 0.6



r = - 0.9



r = 0

3.26 Home selling prices The House Selling Prices FL data file on the text CD lists selling prices of homes in Gainesville, Florida, in 2003 and some predictors for the selling price. For the response variable y = selling price in thousands of dollars and the explanatory variable x = size of house in thousands of square feet, $\hat{y} = 9.2 + 77.0x$.

a. How much do you predict a house would sell for if it has (i) 2000 square feet, (ii) 3000 square feet?

b. Using results in part a, explain how to interpret the slope.

c. Is the correlation between these variables positive or negative? Why?

The slope is positive.

d. One home that is 3000 square feet sold for $300,000. Find the residual, and interpret.

a. y-hat = 9.2+77.0X
y-hat (x=2) = 9.2 + 77.0(2) = 163.2
y-hat (x=3) = 9.2 + 77.0(3) = 240.2

b. Slope = 77.0
From part a, one unit increase will have 240.2-163.2 = 77.0 units increase in the home price.
For every increase of one thousand square feet in the size of the house, the home prices are predicted to increase by about $77,000.

d. y-hat (x=3) = 240.2
residual = y – y-hat = 300 – 240.2 = 59.8 or $59,800.
The residual is a prediction error, the selling price is $59,800 higher than the predicted home price using the regression line.

3.40 Oil and GDP An article in the September 16, 2006, issue of *The Economist* showed a scatterplot for many nations relating the response variable y = annual oil consumption per person (in barrels) and the explanatory variable x = gross domestic product (GDP, per person, in thousands of dollars). The values shown on the plot were approximately as shown in the table.

a. Create a data file and use it to construct a scatterplot. Interpret.

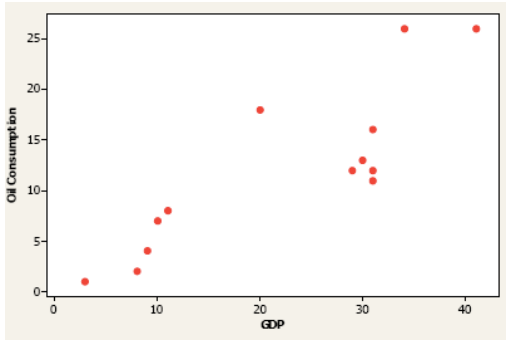
y-hat = -0.105 + 0.5464X

b. Find and interpret the prediction equation.

c. Find and interpret the correlation. r = 0.85

d. Find and interpret the residual for Canada.

Nation	GDP	Oil Consumption
India	3	1
China	8	2
Brazil	9	4
Mexico	10	7
Russia	11	8



d. For Canada, x = 34
y-hat = -0.105 + 0.5464(34) = 18.4726
Residual = y – y-hat = 26 – 18.4726 = 7.5274 or 7.5

c. combined:

Defendant Race	Death penalty	No death penalty
White	(53+0)/(467+16)=0.11 (higher)	(414+16)/(467+16)=0.89
Black	(11+4)/(48+143)=0.08	(37+139)/(48+143)=0.94

(For 3.58 parts a - c, round your answers to two decimal places.)

3.58 Death penalty and race The table shows results of whether the death penalty was imposed in murder trials in Florida between 1976 and 1987. For instance, the death penalty was given in 53 out of 467 cases in which a white defendant had a white victim.

Death Penalty, by Defendant's Race and Victim's Race

Victim's Race	Defendant's Race	Death Penalty: Yes	Death Penalty: No	Total
White	White	53	414	**467**
	Black	11	37	**48**
Black	White	0	16	**16**
	Black	4	139	**143**

Source: Originally published in *Florida Law Review*. Michael Radelet and Glenn L. Pierce, Choosing Those Who Will Die: Race and the Death Penalty in Florida, vol. 43, *Florida Law Review* 1 (1991).

a. First, consider only those cases in which the victim was white. Find the conditional proportions that got the death penalty when the defendant was white and when the defendant was black. Describe the association.

b. Repeat part a for cases in which the victim was black. Interpret.

c. Now add these two tables together to get a summary contingency table that describes the association between the death penalty verdict and defendant's race, ignoring the information about the victim's race.

a. White victim:

Defendant Race	Death penalty	No death penalty
White	0.11 (lower)	0.89
Black	0.23	0.77

b. Black victim:

Defendant Race	Death penalty	No death penalty
White	0 (lower)	1
Black	0.03	0.97

d. The association changes (White have higher rate of death penalty compared to Black but the rate turned out lower) after a third variable is included (victim's race) and data are analyzed at separate levels of that variable.

e. Yes. Both the victim's race and defendant's race predict the death penalty status, and they (victim's race and defendant's race) are associated.
Confounding occurs when two explanatory variables are associated with each other.
In such cases, it is difficult to determine whether either of them truly causes the response.

3.60 Age a confounder? A study observes that the subjects in the study who say they exercise regularly reported only half as many serious illnesses per year, on the average, as those who say they do not exercise regularly. One paragraph in the results section of an article about the study starts out, "We next analyzed whether age was a confounding variable in studying this association."

a. It is possible that those who exercised more were younger, and that rather than exercise, age was responsible for the lower rate of serious illness.

b. If age is not measured in the study, it is a lurking variable.
A lurking variable is one that might have some effects, but is not measured in the study.
If age is included in the study and found to be linked with both the explanatory and response variables, then it is a confounding variable.