**Section 2.4 Graphical Misrepresentations of Data**

**Objective**

1. Describe What Can Make a Graph Misleading or Deceptive

***Objective 1: Describe What Can Make a Graph Misleading or Deceptive***

Objective 1, Page 1

*Answer the following after watching the video.*

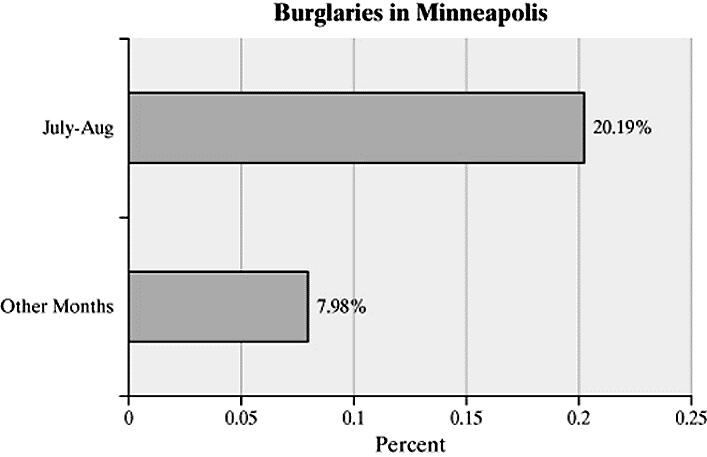
1. Explain the difference between a graph that is misleading and a graph that is deceiving. GRAPH IS MISLEADING IF IT IS UNINTENTIONAL. IT IS DECEIVING IF IT IS INTENTIONAL IMPRESSION
2. List what the most common misrepresentations of data involve.

* Increments between tick marks should be consistent.
* Scales for comparative graphs should be the same.
* The baseline, or zero point, should be at the bottom of the graph.

Objective 1, Page 2

**Example 1 *Misrepresentations of Data***

A home security company located in Minneapolis, Minnesota, develops a summer ad campaign with the slogan "When you leave for vacation, burglars leave for work." According to the city of Minneapolis, roughly 20% of home burglaries occur during the peak vacation months of July and August. The advertisement contains the graphic shown. Explain what is wrong with the graphic.



In this example, we'll take a look

at misrepresentation of data.

A home security company located in Minneapolis, Minnesota

develops a summer ad campaign with the slogan, when

you leave for vacation, burglars leave for work.

According to the city of Minneapolis, roughly 20%

of home burglaries occur during the peak vacation

months of July and August.

The advertisement contains the graphic shown.

Explain what is wrong with the graphic.

So let's consider how the categories of data are defined.

We know for all 12 months, the total has to be 100%.

And if we think about roughly 8% per month times 10 months,

that represents approximately 80% of the burglaries

happen in other months, showing us that this 20.19%

is actually the combined total for July and August,

not per each month.

The unsuspecting reader is misled

into thinking that July and August each have a burglary

rate of 20%.

This graph gives a better picture

of the burglary distribution.

The increased during the month of July

is not as dramatic as the bar graph on the previous screen.

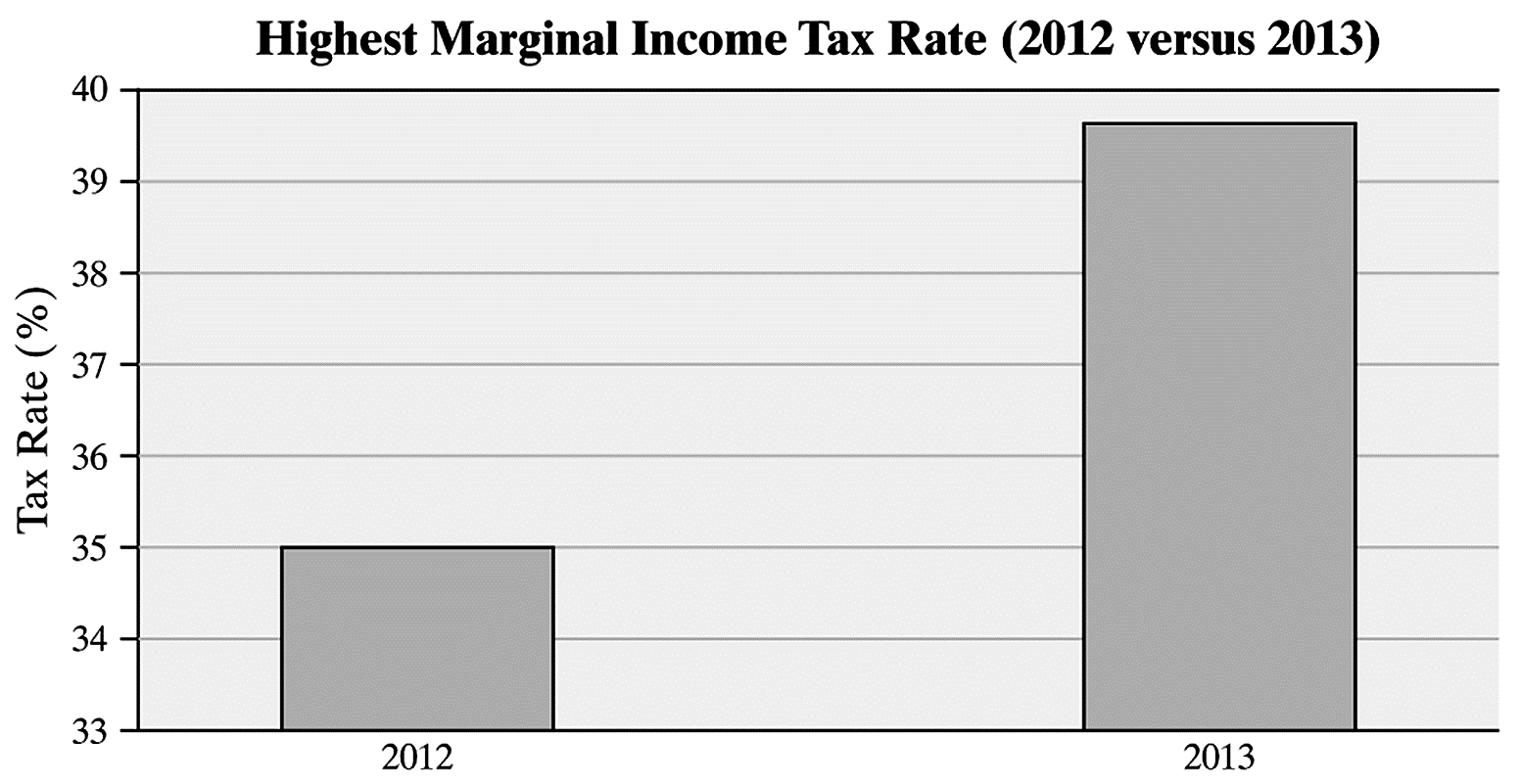
And as it works out, August actually

has fewer burglaries than September or October.

Objective 1, Page 3

**Example 2 *Misrepresentations of Data by Manipulating the Vertical Scale***

A national news organization developed the following graphic to illustrate the change in the highest marginal tax rate effective January 1, 2013. Why might this graph be considered misleading?



In this example, we'll take a look

at misrepresentation of data by manipulating

the vertical scale.

A national news organization developed the following graphic

to illustrate the change in the highest marginal tax

rate effective January 1, 2013.

Why might this graph be considered misleading?

When we take a look at this graph,

that may lead you to believe that marginal tax rates are

more than tripling since the height of the bar for 2013

is more than three times the height of the bar for 2012.

In fact, the difference is only 4.6 percentage points

because the tax rate in 2012 is 35% and in 2013 it's 39.6%.

The reason for this incorrect conclusion

is that the vertical axis does not begin at 0

but, instead, it begins at 33%.

Let's take a look at this graph that does not

distort the difference in tax rates.

The increase in tax rate is still

apparent in the graph-- this bar is definitely

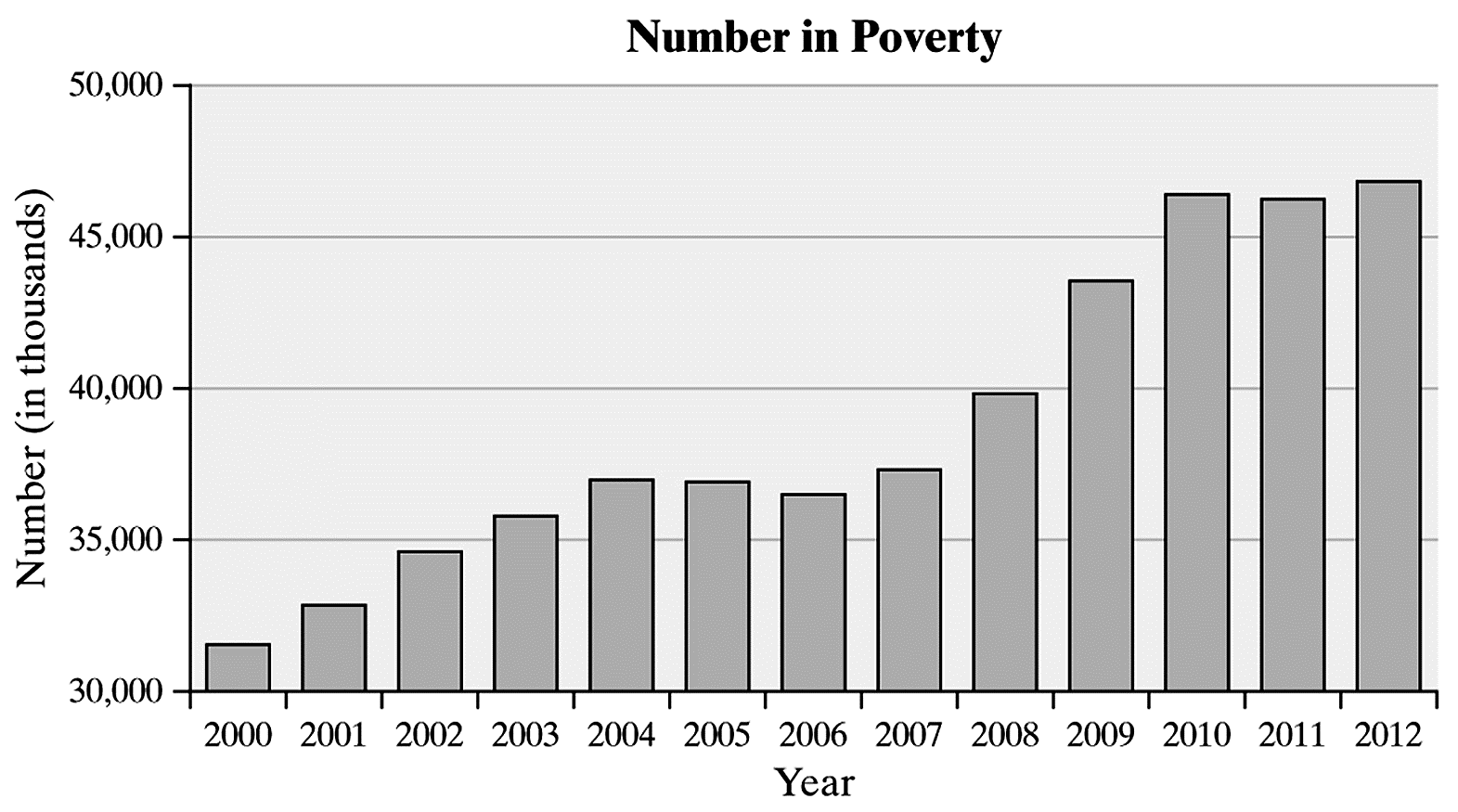
higher than that bar-- and without distorting

the size of the actual increase.

Objective 1, Page 5

**Example 3 *Misrepresentations of Data by Manipulating the Vertical Scale***

The graph depicts the number of residents in the United States living in poverty. Why might this graph be considered misrepresentative?



In this example, we'll talk about misrepresentation of data

by manipulating the vertical scale.

The following graph depicts the number

of residents in the United States living in poverty.

Why might this graph be considered "misrepresentative"?

Here's the graph.

The graph may mislead the reader into believing

that the number in poverty has more than doubled since 2005

because the bar for 2012 is more than twice the height

of the bar in 2005.

However, notice that the vertical axis

begins at 30,000 instead of 0.

This type of scaling is common when the smallest observed

data value is a rather large number.

It's not necessarily done intentionally

to confuse or mislead the reader.

Often, the main purpose in graphs-- particularly time

series graphs-- is to discover a trend rather than

the actual differences in the data.

However, the author of the graph should clearly

indicate that the graph does not begin at 0 by including

the following symbol in the vertical scale.

This symbol indicates that the scale has been truncated

and the graph has a gap.

Although the data does stand out,

it's better to use a time series plot

when displaying time series data rather than a bar plot.

In addition, it's better to use the percent of the population

in poverty rather than the actual number

living in poverty.

This is due to the fact that increases in poverty

may be due to increases in population

as well as a deterioration of the economy.

Here's a time series plot of the percentage of US

residents living in poverty.

The lack of bars allow us to focus on the trend in the data

rather than the relative size or area of the bars.

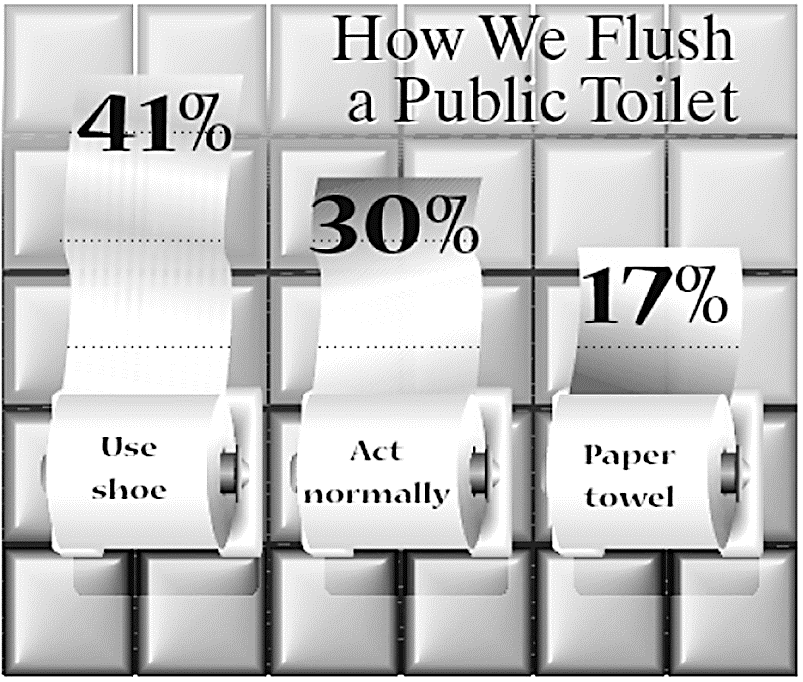
Also, it's clear that the percent in poverty

was highest in 2010, and this is not clear from the bar graphs.

Objective 1, Page 7

**Example 4 *Misrepresentations of Data***

The bar graph shown is a *USA Today*-type graph. A survey was conducted by Impulse Research in which individuals were asked how they would flush a toilet when the facilities are not sanitary. What is wrong with the graphic?



In this example, we'll take a look

at misrepresentation of data.

The following bar graph is a USA Today type graph.

The survey was conducted by Impulse Research

for Quilted Northern Confidential, in which

individuals were asked how they would flush a toilet when

the facilities are not sanitary.

What's wrong with the graph there?

A simple bar chart would have got

the idea across in a clear and concise way.

They used the toilet paper to tie it into the research

question, but it's very misleading.

The first thing that we're unsure of

is whether the roll itself is supposed

to be part of the graph or not.

When we compare 41% to 17%, that's

supposed to be 2.4 times bigger.

If we compare the height including the roll of toilet

paper, then this height is less than twice that height.

Meaning that it's not accurate.

So maybe they intended us to only

use the part up to but not including the roll.

Well, if we did that, it's still about twice that height,

not 2.4 times.

So the vertical scale is off.

Objective 1, Page 9

1. Why is the use of 3-D effects strongly discouraged?
2. Why do we emphasize that the bars or classes should have the same width?

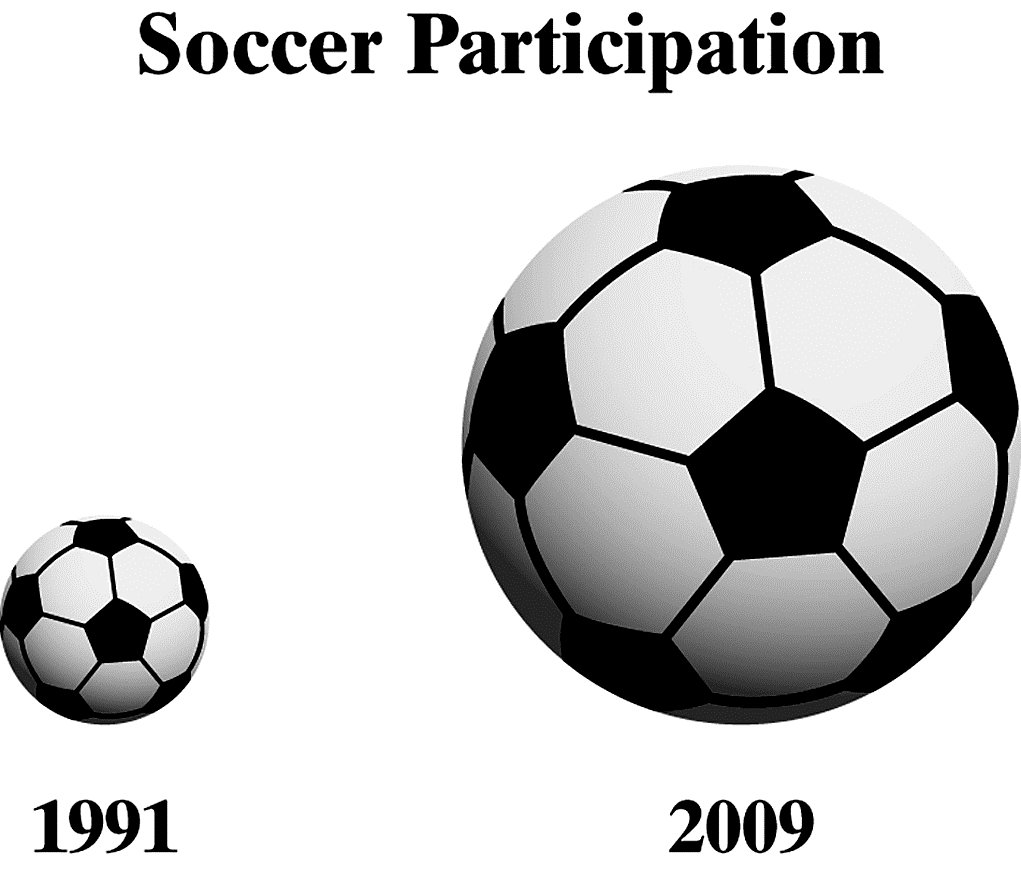
Newspapers, magazines, and Internet sites often go for a "wow" factor when displaying graphs. The graph designer may be more interested in catching the reader's eye than making the data stand out. The two most commonly used tactics are 3-D graphs and *pictograms* (graphs that use pictures to represent the data). The use of 3-D effects is strongly discouraged because such graphs are often difficult to read, add little value to the graph, and distract the reader from the data.

When comparing bars, our eyes are really comparing the *areas* of the bars. That is why we emphasize that the bars or classes should have the same width. Uniform width ensures that the area of the bar is proportional to its height so that we can simply compare the heights of the bars. However, when we use two-dimensional pictures in place of bars, as with pictograms, it is not possible to obtain a uniform width. To avoid distorting the picture when values increase or decrease, both the height and width of the picture must be adjusted. This often leads to misleading graphs.

Objective 1, Page 10

**Example 5 *Misrepresentations of Data by Manipulating Dimension***

Soccer continues to grow in popularity as a sport in the United States. In 1991, there were approximately 10 million participants in the United States aged 7 years and older. By 2009, this number had climbed to 14 million. To illustrate this increase, we could create a graphic like the one shown below. Describe how the graph may be misleading. *Source:* U.S. Census Bureau; National Sporting Goods Association



In this example, we'll take a look at misleading graphs.

Soccer continues to grow in popularity

as a sport in the United States.

In 1991, there were approximately

10 million participants in the United States aged 7 or older.

By 2009, this number had climbed to 14 million.

To illustrate this increase, we could create a graphic

like the following.

Describe how the graph may be misleading.

The actual increase from 10 million to 14 million

participants is an increase of 40%.

Now, I ask you-- does the soccer ball on the right

seem 40% larger than the ball on the left?

Actually, it's four times the size,

which represents an increase of over 300%.

One problem here is that the actual data are not labeled.

If it said "10 million" and "14 million"

next to the two soccer balls, you

could understand the increase just

from the values-- ignoring the size of the pictures.

But since those are missing, there's

not much else you can do, except draw the wrong conclusion as

compared to the right conclusion.

A simple bar graph would have shown the difference

between the two years-- 1991 and 2009-- in a much better way.

We can see the increase-- from 10 million to 14 million--

without distorting the facts.

The bar on the right is definitely

40% larger than the bar on the left.

Now, if you wanted to do this and still

make it look appealing, you could try a graph

like the following where one soccer ball represents

1 million participants.

We can see, from 2009, that there were 4 million

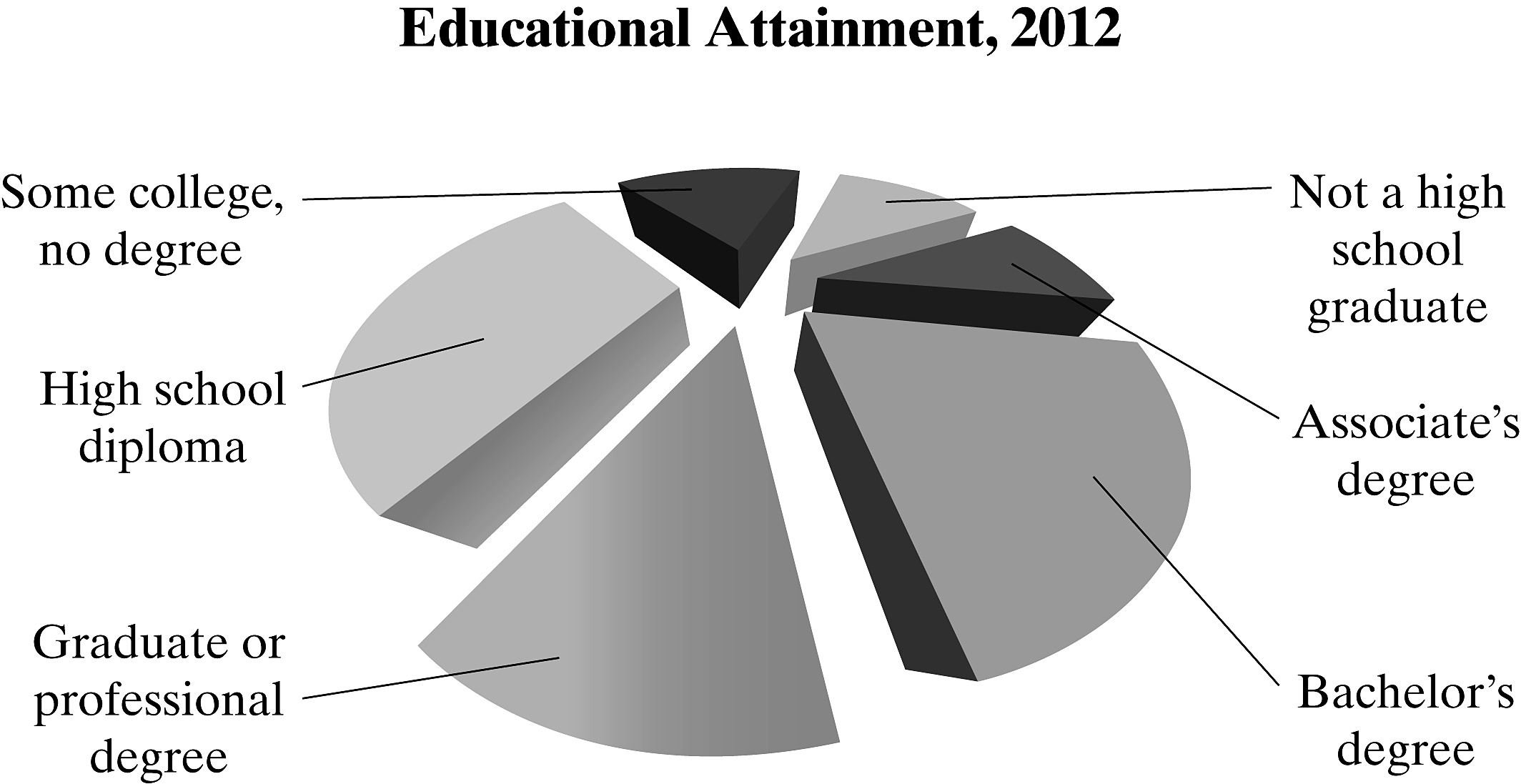
more participants because of the 4 extra soccer balls

than there were in 1991.

Objective 1, Page 11

**Example 6 *Misrepresentations of Data: Three-Dimensional Scale***

The figure represents the educational attainment (level of education) in 2016 of adults 25 years and older who are U.S. residents. Why might this graph be considered misrepresentative?



n this example, we'll discuss misleading graphs.

The figure represents the educational attainment level

of education in 2016 of adults 25 years and older

who are US residents.

Why might this graph be considered misrepresentative?

Here's the graph.

When working with a 3D style graph,

one problem that we run into is that the category

hat's closest to us appears to be overrepresented.

Here, the blue wedge facing us, bachelors degree, is only 21%

of the entire pie.

However, it appears to be much larger than that.

Here are the data that the results are based upon.

And 44,778 out of the total is roughly 21%.

n this case, a standard pie chart

would display the data in a much fairer representation.

Again, if we look at the bachelor's degree here,

we can clearly see that its percentage is less than that

of, say, high school diploma.

But if we go back to the 3D pie chart,

t appears to be close in size to that.

Let's take one quick look at what

happens as we start to rotate these pie slices around

n a three dimensional way.

Here's that pie chart, and we're going

o take a look at what happens when

we rotate our perspective or our view of this pie chart.

Here we go.

Notice that, for instance, the yellow slice

changes as we move it around in terms

of how large that proportion seems to be

or that percentage seems to be.

And that's why these 3D pie charts are not a great way

o represent a set of data.