

# AP Stats

## Chapter 1 1.1 Outline

### VOCABULARY

Categorical variables place individuals into one of several groups or categories.

*individuals are radio stations*

The radio audience rating service Arbitron places U.S. radio stations into categories that describe the kinds of programs they broadcast. Here are two different tables showing the distribution of station formats in a recent year:<sup>3</sup>

*Categorical variable*

Frequency table	
Format	Count of stations
Adult contemporary	1556
Adult standards	1196
Contemporary hit	569
Country	2066
News/Talk/Information	2179
Oldies	1060
Religious	2014
Rock	869
Spanish language	750
Other formats	1579
Total	13,838

Relative frequency table	
Format	Percent of stations
Adult contemporary	11.2
Adult standards	8.6
Contemporary hit	4.1
Country	14.9
News/Talk/Information	15.7
Oldies	7.7
Religious	14.6
Rock	6.3
Spanish language	5.4
Other formats	11.4
Total	99.9

$$\frac{1556}{13838} = .112$$

*roundoff error*

*add up*

*should be 100%*

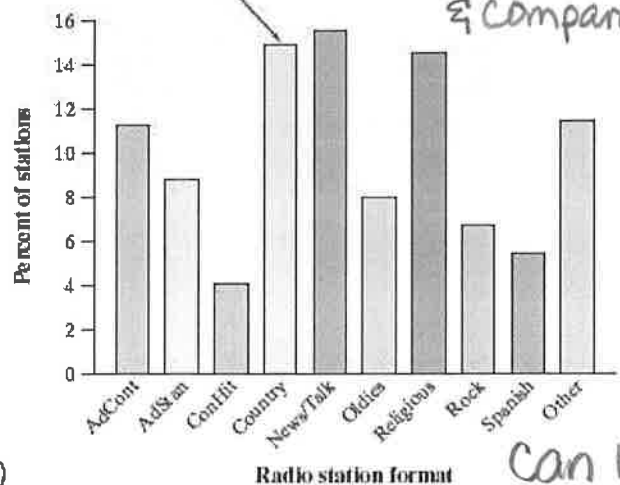
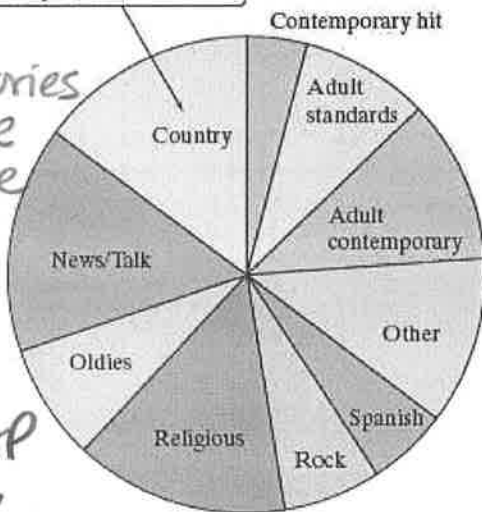
This slice occupies 14.9% of the pie because 14.9% of the radio stations have a "Country" format.

This bar has height 14.9% because 14.9% of the radio stations have a "Country" format.

*all categories that make up a whole  
awkward to make by hand*

*Preferred easier to read & compare*

*NOT an AP Topic!!*



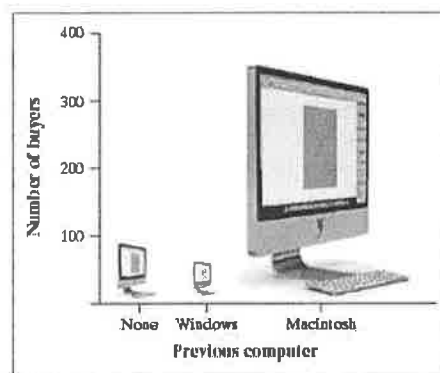
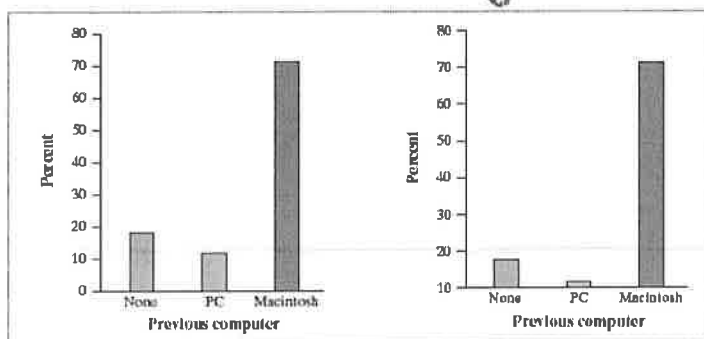
*Can be counts or %*

FIGURE 1.1 (a) Pie chart and (b) bar graph of U.S. radio stations by format.

✓ When you draw a bar graph make the bars EQUALLY WIDE!

Do not replace the bars with pictures for greater eye appeal.

- ✓ Beware of the pictograph →
- ✓ Watch the scales ↓

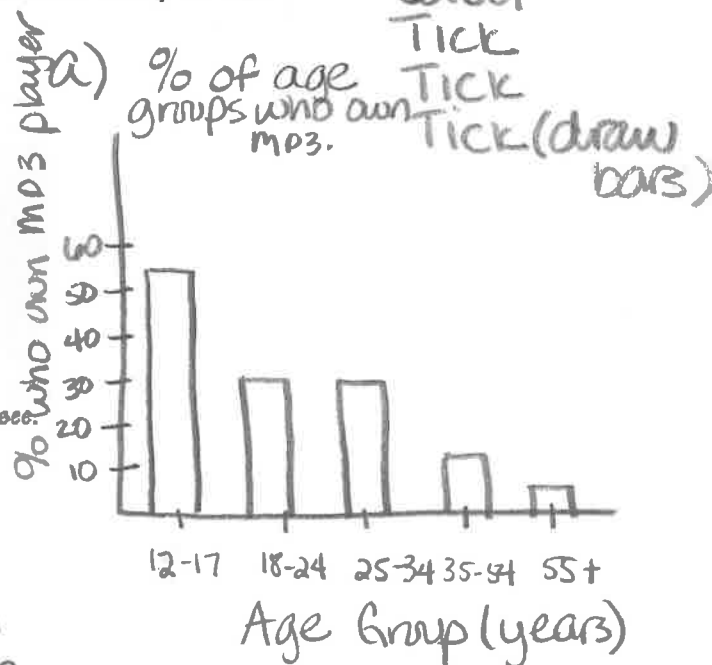


Portable MP3 music players, such as the Apple iPod, are popular—but not equally popular with people of all ages. Here are the percents of people in various age groups who own a portable MP3 player, according to an Arbitron survey of 1112 randomly selected people.<sup>4</sup>

Age group (years)	Percent owning an MP3 player
12 to 17	54
18 to 24	30
25 to 34	30
35 to 54	13
55 and older	5
<hr/>	
132	

**PROBLEM:**

- Make a well-labeled bar graph to display the data. Describe what you see.
- Would it be appropriate to make a pie chart for these data? Explain.



b) making a pie chart to display this data is not appropriate because each percent in the table refers to a different age group, not to parts of a single whole!

When a dataset involves **TWO** categorical variables, we begin by examining the counts or percents in various categories for one of the variables.

- ✓ A **two-way table** describes TWO categorical variables, organizing counts according to a Row variable and a Column variable.

A survey of 4826 randomly selected young adults (aged 19 to 25) asked, "What do you think the chances are you will have much more than a middle-class income at age 30?" The table below shows the responses.<sup>6</sup>

*2 way table* ↓

Young adults by gender and chance of getting rich			
Opinion	Gender		Total
	Female	Male	
Almost no chance	96	98	194
Some chance but probably not	426	286	712
A 50-50 chance	696	720	1416
A good chance	663	758	1421
Almost certain	486	597	1083
Total	2367	2459	4826

*margin*

What are the variables described by this two-way table?

*Gender & opinion about becoming rich.*

How many young adults were surveyed?

*4,826*

- ✓ The **marginal distribution** of one of the categorical variables in a two-way table of counts is the distribution of values of that variable among all individuals described by the table.

- ✓ Percents are more informative than counts. *% of young adults who think they are almost to be rich by 30.* 
$$\frac{\text{Almost certain total}}{\text{Table Total}} = \frac{1083}{4826} = .224 = 22.4\%$$

#### How to examine a marginal distribution:

- Use the data in the table to calculate the marginal distribution (in percents) of the row or column totals
- Make a graph to display the marginal distribution.

Young adults by gender and chance of getting rich			
Opinion	Gender		Total
	Female	Male	
Almost no chance	96	98	194
Some chance but probably not	426	286	712
A 50-50 chance	696	720	1416
A good chance	663	758	1421
Almost certain	486	597	1083
Total	2367	2459	4826

Almost none  $\frac{194}{4826} \approx 4\%$   
 Some chance  $\frac{712}{4826} \approx 14.8\%$   
 50-50  $\frac{1416}{4826} \approx 29.3\%$   
 good  $\frac{1421}{4826} \approx 29.4\%$   
 certain  $\frac{1083}{4826} \approx 22.4\%$

# PROBLEM:

(a) Use the data in the two-way table to calculate the marginal distribution (in percents) of opinions.

(b) Make a graph to display the marginal distribution. Describe what you see.

Title  
 Label  
 Label  
 Tick  
 Tick  
 Tick

