Quantitative variables often take many values. A graph of the distribution is clearer if nearby values are grouped together. The most common graph of the distribution of one quantitative variable is a **histogram.**

Example – Foreign-Born Residents
Making a histogram

What percent of your home state's residents were born outside the United States? The country as a whole has 12.5% foreign-born residents, but the states vary from 1.2% in West Virginia to 27.2% in California. The table below presents the data for all 50 states. The individuals in this data set are the states. The variable is the percent of a state's residents who are foreign-born. It's much easier to see from a graph than from the table how your state compares with other states.

State	Percent	State	Percent	State	Percent
Alabama	2.8	Louisiana	2.9	Ohio	3.6
Alaska	7.0	Maine	3.2	Oklahoma	4.9
Arizona	15.1	Maryland	12.2	Oregon	9.7
Arkansas	3.8	Massachusetts	14.1	Pennsylvania	5.1
California	27.2	Michigan	5.9	Rhode Island	12.6
Colorado	10.3	Minnesota	6.6	South Carolina	4.1
Connecticut	12.9	Mississippi	1.8	South Dakota	2.2
Delaware	8.1	Missouri	3.3	Tennessee	3.9
Florida	18.9	Montana	1.9	Texas	15.9
Georgia	9.2	Nebraska	5.6	Utah	8.3
Hawaii	16.3	Nevada	19.1	Vermont	3.9
Idaho	5.6	New Hampshire	5.4	Virginia	10.1
Illinois	13.8	New Jersey	20.1	Washington	12.4
Indiana	4.2	New Mexico	10.1	West Virginia	1.2
Iowa	3.6	New York	21.6	Wisconsin	4.4
Kansas	6.3	North Carolina	6.9	Wyoming	2.7
Kentucky	2.7	North Dakota	2.1		

Steps to make a histogram:

1. **Divide the range of the data into classes of equal width**. The data in the table vary from 1.2 to 27.2, so we might choose to use classes of width 5, beginning at 0:

0–5 5–10 10–15 15–20 20–25 25–30

But we need to specify the classes so that each individual falls into exactly one class. For instance, what if a state had exactly 5.0% of its residents born outside the United States? Since a value of 0.0% would go in the 0–5 class, we'll agree to place a value of 5.0% in the 5–10 class, a value of 10.0% in the 10–15 class, and so on. In reality, then, our classes for the percent of foreign-born residents in the states are

$$0 \text{ to } < 5 \quad 5 \text{ to } < 10 \quad 10 \text{ to } < 15 \quad 15 \text{ to } < 20 \quad 20 \text{ to } < 25 \quad 25 \text{ to } < 30$$

2. **Find the count (frequency) or percent (relative frequency) of individuals in each class**. Here is a frequency table and a relative frequency table for these data:

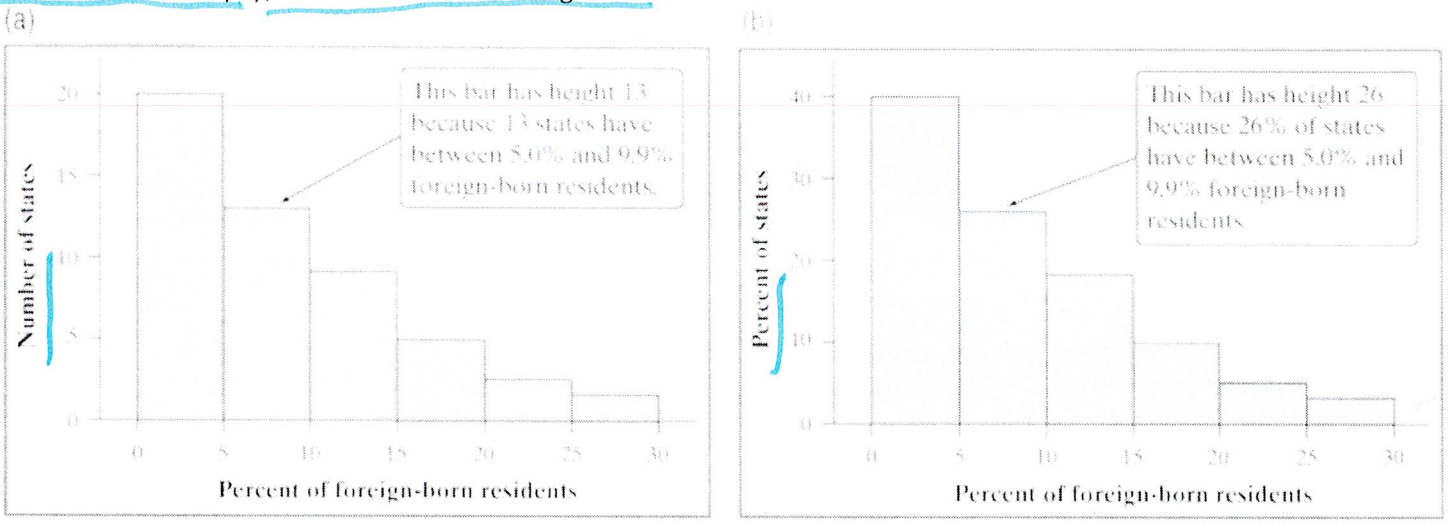
Frequency table

Class	Count
0 to < 5	20
5 to < 10	13
10 to < 15	9
15 to < 20	5
20 to < 25	2
25 to < 30	1
Total	50

Relative frequency table

Class	Percent
0 to < 5	40
5 to < 10	26
10 to < 15	18
15 to < 20	10
20 to < 25	4
25 to < 30	2
Total	100

3. Label and scale your axes and draw the histogram. Label the horizontal axis with the variable whose distribution you are displaying. That's the percent of a state's residents who are foreign-born. The scale on the horizontal axis runs from 0 to 30 because that is the span of the classes we chose. The vertical axis contains the scale of counts or percents. Each bar represents a class. The base of the bar covers the class, and the bar height is the class frequency or relative frequency. Draw the bars with no horizontal space between them unless a class is empty, so that its bar has height zero.



Histogram **(a)** is a frequency histogram. Histogram **(b)** is relative frequency histogram.

Discuss *frequency histogram* versus *relative frequency histogram*. Relative frequency histograms are typically more useful because they make it easier to compare two distributions especially when the number of individuals is very different.

What do the histograms in the Figure above tell us about the percent of foreign-born residents in the states? To find out, we follow our familiar routine: describe the pattern and look for any departures from the pattern.

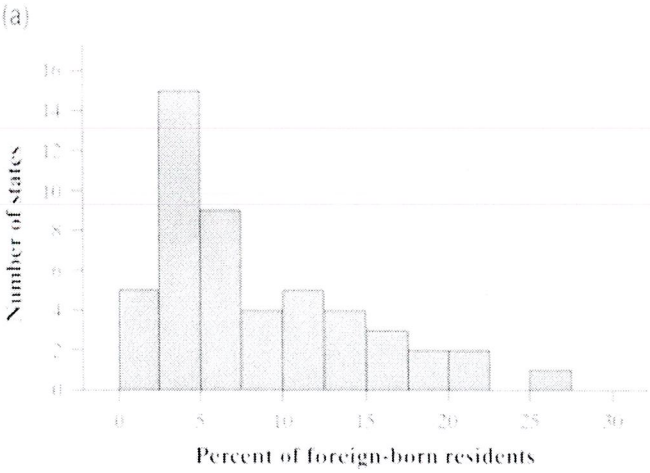
Shape: The distribution is skewed to the right. A majority of states have fewer than 10% foreign-born residents, but several states have much higher percents, so that the graph extends quite far to the right of its peak. The distribution has a single peak at the left, which represents states in which between 0% and 4.9% of residents are foreign-born.

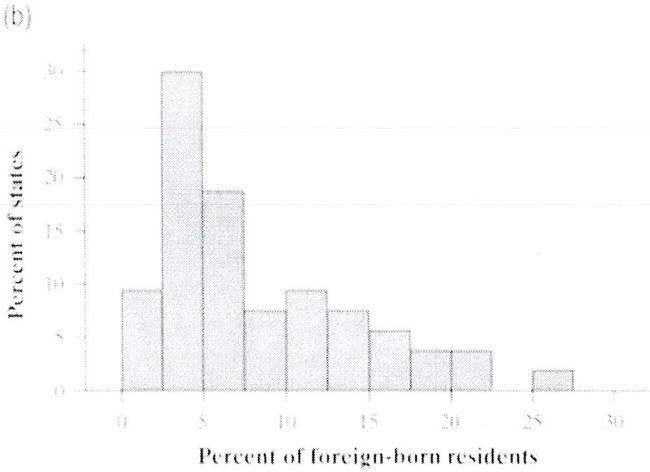
Center: From the graph, we see that the midpoint (median) would fall somewhere in the 5.0% to 9.9% class. Remember that we're looking for the value having 25 states with smaller percents foreign-born and 25 with larger. (Arranging the observations from the table in order of size shows that the median is 6.1%.)

Spread: The histogram shows that the percent of foreign-born residents in the states varies from less than 5% to over 25%. (Using the data in the table, we see that the range is 27.2% − 1.2% = 26.0%.)

Outliers: We don't see any observations outside the overall single-peaked, right-skewed pattern of the distribution. The histograms below shows (a) a frequency histogram and (b) a relat

The histograms below show (a) a frequency histogram and (b) a relative frequency histogram of the same distribution, with classes half as wide. The new classes are 0–2.4, 2.5–4.9, etc. Now California, at 27.2%, stands out as a potential outlier in the right tail. The choice of classes in a histogram can influence the appearance of a distribution. Histograms with more classes show more detail but may have a less clear pattern.





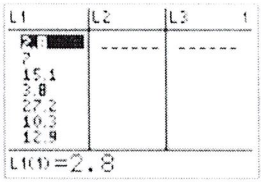
Histogram on the calculator

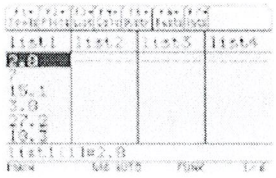
TECHNOLOGY CORNER Histograms on the calculator

TI-83/84 Clear out all equations in y= before starting. TI-89

1. Enter the data for the percent of state residents born outside the United States in your Statistics/List Editor.

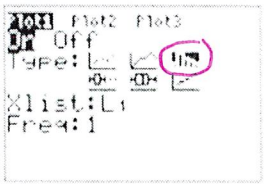
- Press STAT and choose 1:Edit... (TI-83/84)
- Press APPS and select Stats/List Editor. (TI-89)
- Type the values into list L1. (TI-83/84)
- Type the values into list1. (TI-89)

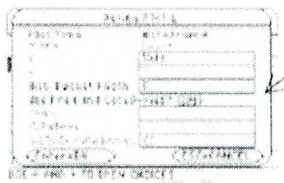




2. Set up a histogram in the Statistics Plots menu.

- Press 2nd Y= (STAT PLOT). (TI-83/84)
- Press F2 and choose 1:Plot Setup... (TI-89)
- Press ENTER or 1 to go into Plot1. (TI-83/84)
- With Plot1 highlighted, press F1 to define. (TI-89)





Set Hist. Bucket Width to 5.

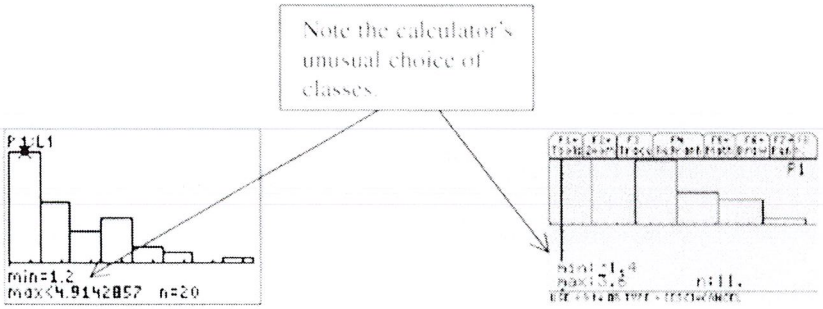
What percent of your home state's residents were born outside the United States? The percentage for each state is listed below:

2.8,7.0,15.1,3.8,27.2,10.3,
12.9,8.1,18.9,9.2,16.3,5.6,
13.8,4.2,3.8,6.3,2.7,2.9,3.2,
12.2,14.1,5.9,6.6,1.8,3.3,
1.9,5.6,19.1,5.4,20.1,10.1,
21.6,6.9,2.1,3.6,4.9,9.7,5.1,
12.6,4.1,2.2,3.9,15.9,8.3,
3.9,10.1,12.4,1.2,4.4,2.7

3. Use ZoomStat (ZoomData on the TI-89) to let the calculator choose classes and make a histogram.

- Press ZOOM and choose 9:ZoomStat.
- Press TRACE and ◄ ► to examine the classes.

- Press F5 (ZoomData).
- Press F3 (Trace) and ◄ ► to examine the classes.



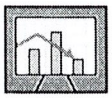
4. Adjust the classes to match those in Figure 1.16, and then graph the histogram.

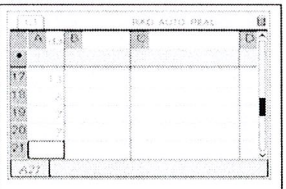
- Press WINDOW and enter the values shown.
- Press GRAPH
- Press TRACE and ◄ ► to examine the classes.

- Press ♦ F2 (WINDOW) and enter the values shown.
- Press ♦ F3 (GRAPH)
- Press F3 (trace and ◄ ► to examine the classes)

TI Inspire Directions

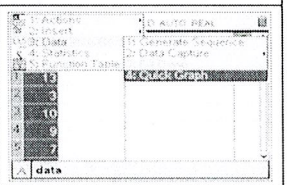
Enter the data into the *Lists & Spreadsheets* area. (See Lists and Spreadsheets for entering data.)





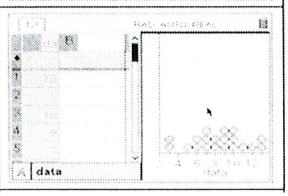
Arrow up to select the entire column.

Press (menu) and choose #3 Data, #4 Quick Graph.



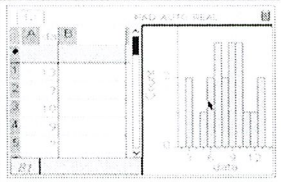
The window will be split and a dot plot will appear to the right.

Remember that CTRL-TAB will move between the split screens.

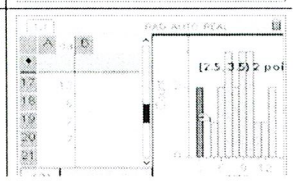


From the right window, press MENU.

Choose #1 Plot Type, #3 Histogram



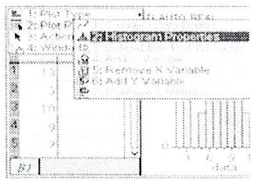
To see the size of the bars and the number of values represented by each bar, click on the bar.



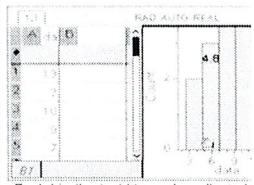
6. To adjust the histogram, press MENU.

Choose #2 Plot Properties, #2 Histogram Properties

#1 Histogram Scale
#1 Count – Displays the data based upon the number of values that occur within each bar (or bin) on the histogram. This is the default data representation when you create a histogram. The Count tool is not available because it is currently in use.
#2 Percent – Displays data in the histogram by each group's percent value of the whole data set.
#3 Density – Displays data in the histogram based upon the density of each value within the data set.
#2 Bin Settings - Displays a dialog for setting the histogram values for bin (bar) width and alignment.
You can adjust the bins' width and number by dragging the side of one bin. A value that occurs on the edge of a bin is counted in the bin to the right. Avoid problems by staying with the default bin widths.



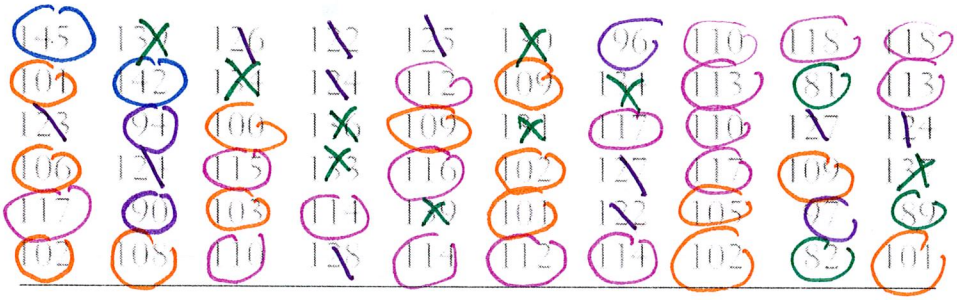
The number of bins displayed depen upon the number of data points and the distribution of these points.



Each bin (bar) width can be adjusted clicking and dragging the side of the

Check Your Understanding

Many people believe that the distribution of IQ scores follows a "bell curve," like the one shown in the margin. But is this really how such scores are distributed? The IQ scores of 60 fifth-grade students chosen at random from one school are shown below.



min = 81

max = 145

X scale: 10

1. Construct a histogram that displays the distribution of IQ scores effectively

n=60

80 - <90 = 3

90 - <100 = 4

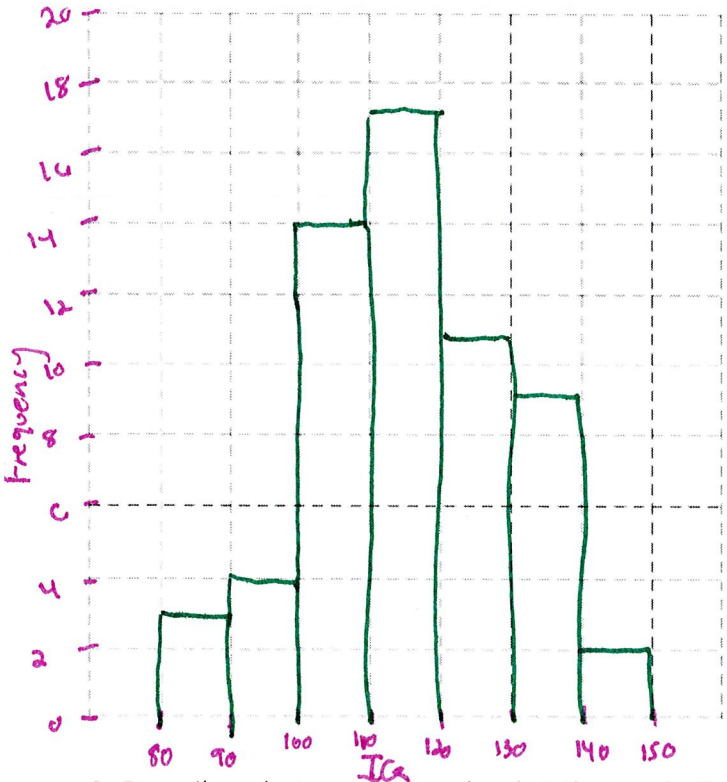
100 - <110 = 14

110 - <120 = 17

120 - <130 = 11

130 - <140 = 9

140 - <150 = 2

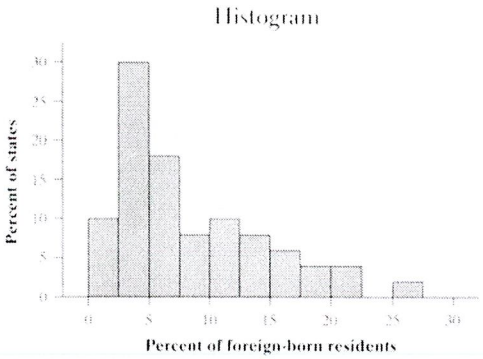


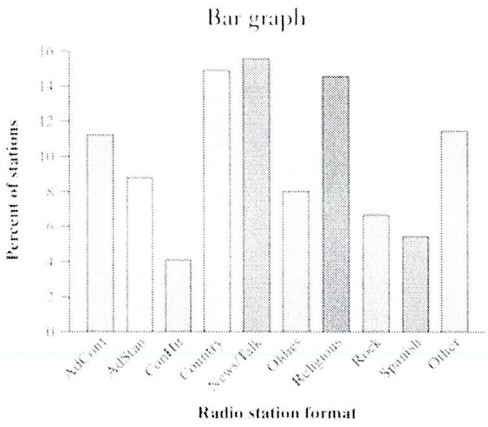
2. Describe what you see. Is the distribution bell-shaped?

The distribution is roughly symmetric and bell shaped. The median IQ appears to be between 110 and 120. The IQ's vary from 80 to 150. There do not appear to be any outliers

Using Histograms Wisely

- Do not confuse histograms and bar graphs
  - Histograms are for quantitative variables
  - Bar graphs are for categorical variables
  - Histograms have no space between bars; bar graphs have a blank space between bars

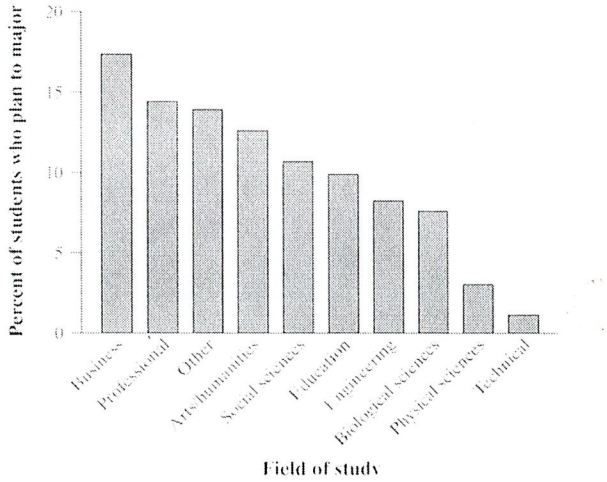




- Do not use counts or percents as data. (on x axis)
- Use percents instead of counts when comparing distributions with different numbers of observations
- Just because a graph looks nice, it is not necessarily a meaningful display of data.

Check Your Understanding

Questions 1 and 2 relate to the following setting. About 1.6 million first-year students enroll in colleges and universities each year. What do they plan to study? The graph displays data on the percents of first-year students who plan to major in several discipline areas.



1. Is this a bar graph or a histogram? Explain.

This is a bar graph b/c the horizontal axis divides the observation up into categories (categorical data)

2. Would it be correct to describe this distribution as right-skewed? Why or why not?

No, the variable on the x axis is categorical, not quantitative.