AP Statistics Answer all exercises on your own paper. Due first class. 1

What have we learned? W's

We've learned that data are information in a context.

- The W's help nail down the context: *Who, What, Why, Where, When,* and *hoW.*
- We must know at least the *Who, What,* and *Why* to be able to say anything useful based on the data. The *Who* are the *cases.* The *What* are the *variables.* A variable gives information about each of the cases. The *Why* helps us decide which way to treat the variables.

We treat variables in two basic ways, as *categorical* or *quantitative.*

- Categorical variables identify a category for each case. Usually we think about the counts of cases that fall in each category. (An exception is an identifier variable that just names each case.)
- Quantitative variables record measurements or amounts of something; they must have *units.*
- Sometimes we treat a variable as categorical or quantitative depending on what we want to learn from it, which means some variables can't be pigeon-holed as one type or the other. That's an early hint that in Statistics we can't always pin things down precisely.

TERMS

Context The context ideally tells *Who* was measured, *What* was measured, *How* the data were collected, *Where* the data were collected, and *When* and *Why* the study was performed.

Data Systematically recorded information, whether numbers or labels, together with its context.

Data table An arrangement of data in which each row represents a case and each column represents a variable.

Case A case is an individual about whom or which we have data.

Variable A variable holds information about the same characteristic for many cases.

Categorical variable A variable that names categories (whether with words or numerals) is called categorical.

Quantitative variable A variable in which the numbers act as numerical values is called quantitative. Quantitative variables always have units.

Units A quantity or amount adopted as a standard of measurement, such as dollars, hours, or grams.

SKILLS

When you complete this lesson you should:

Think

- Be able to identify the *Who, What, When, Where, Why,* and *How* of data, or recognize when some of this information has not been provided.
- Be able to identify the cases and variables in any data set.
- Be able to classify a variable as categorical or quantitative depending on its use.
- For any quantitative variable be able to identify the units in which the variable has been measured (or note that they have not been provided).

Tell

- Be able to describe a variable in terms of its *Who, What, When, Where, Why,* and *How* (and be prepared to remark when that information is not provided).

EXERCISES

For each description of data, identify the W's, name the variables, specify for each variable whether its use indicates it should be treated as categorical or quantitative, and, for any quantitative variable, identify the units in which it was measured (or note that they were not provided).

1. The news. Find a newspaper or magazine article in which some data are reported. For the data discussed in the article, answer the questions above. Include a copy of the article with your report.

2. Investments. According to an article in *Fortune* (Dec. 28, 1992), 401(k) plans permit employees to shift part of their before-tax salaries into investments such as mutual funds. Employers typically match 50% of the employees' contribution up to about 6% of salary. One company, concerned with what it believed was a low employee participation rate in its 401(k) plan, sampled 30 other companies with similar plans and asked for their 401(k) participation rates.

3. Arby's menu. A listing posted by the Arby's restaurant chain gives, for each of the sandwiches it sells, the type of meat in the sandwich, the number of calories, and the serving size in ounces. The data might be used to assess the nutritional value of the different sandwiches.

AP Statistics

What have we learned? Graphs

We've learned that we can summarize categorical data by counting the number of cases in each category, sometimes expressing the resulting distribution as percents. We can display the distribution in a bar chart or a pie chart. When we want to see how two categorical variables are related, we put the counts (and/or percentages) in a two-way table called a contingency table.

- We look at the marginal distribution of each variable (found in the margins of the table).
- We also look at the conditional distribution of a variable within each category of the other variable.
- We can display these conditional and marginal distributions using bar charts or pie charts.
- If the conditional distributions of one variable are (roughly) the same for every category of the other, the variables are independent.

TERMS

Frequency table
Relative frequency table
A frequency table lists the categories in a categorical variable and gives the count or percentage of observations for each category.

Distribution
The distribution of a variable gives
- the possible values of the variable and
- the relative frequency of each value.

Area principle
In a statistical display, each data value should be represented by the same amount of area.

Bar chart
Bar charts show a bar representing the count of each category in a categorical variable.

Pie chart
Pie charts show how a "whole" divides into categories by showing a wedge of a circle whose area corresponds to the proportion in each category.

Contingency table
A contingency table displays counts and, sometimes, percentages of individuals falling into named categories on two or more variables. The table categorizes the individuals on all variables at once, to reveal possible patterns in one variable that may be contingent on the category of the other.

Marginal distribution
In a contingency table, the distribution of either variable alone is called the marginal distribution. The counts or percentages are the totals found in the margins (last row or column) of the table.

Conditional distribution
The distribution of a variable restricting the *Who* to consider only a smaller group of individuals is called a conditional distribution.

Independence
Variables are said to be independent if the conditional distribution of one variable is the same for each category of the other. We'll show how to check for independence in a later chapter.

Simpson's paradox
When averages are taken across different groups, they can appear to contradict the overall averages. This is known as "Simpson's paradox."

SKILLS

When you complete this lesson you should:

Think

- Be able to recognize when a variable is categorical and choose an appropriate display for it.
- Understand how to examine the association between categorical variables by comparing conditional and marginal percentages.

Show

- Be able to summarize the distribution of a categorical variable with a frequency table.
- Be able to display the distribution of a categorical variable with a bar chart or pie chart.
- Know how to make and examine a contingency table.
- Know how to make and examine displays of the conditional distributions of one variable for two or more groups.

Tell

- Be able to describe the distribution of a categorical variable in terms of its possible values and relative frequencies.
- Know how to describe any anomalies or extraordinary features revealed by the display of a variable.
- Be able to describe and discuss patterns found in a contingency table and associated displays of conditional distributions.

EXERCISES Graphs

4. **Graphs in the news.** Find a bar graph of categorical data from a newspaper or magazine.
a) Is the graph clearly labeled?
b) Does it violate the area principle?
c) Does the accompanying article tell the W's of the variable?
d) Do you think the article correctly interprets the data? Explain.

5. **Graphs in the news II.** Find a pie chart of categorical data from a newspaper or magazine.
a) Is the graph clearly labeled?
b) Does it violate the area principle?
c) Does the accompanying article tell the W's of the variable?
d) Do you think the article correctly interprets the data? Explain.

6. **Tables in the news.** Find a frequency table of categorical data from a newspaper or magazine.
a) Is it clearly labeled?
b) Does it display percentages or counts?
c) Does the accompanying article tell the W's of the variable?
d) Do you think the article correctly interprets the data? Explain.

7. **Tables in the news II.** Find a contingency table of categorical data from a newspaper or magazine.
a) Is it clearly labeled?
b) Does it display percentages or counts?
c) Does the accompanying article tell the W's of the variables?
d) Do you think the article correctly interprets the data? Explain.

8. **Magnet schools.** An article in the Winter 2003 issue of *Chance* magazine reported on the Houston Independent School District's magnet schools programs. Of the 1755 qualified applicants, 931 were accepted, 298 were wait-listed, and 526 were turned away for lack of space. Find the relative frequency distribution of the decisions made, and write a sentence describing it.

9. **Magnet schools, again.** The *Chance* article about the Houston magnet schools program described in Exercise 5 also indicated that 517 applicants were Black or Hispanic, 292 Asian, and 946 White. Summarize the relative frequency distribution of ethnicity with a sentence or two (in the proper context, of course).

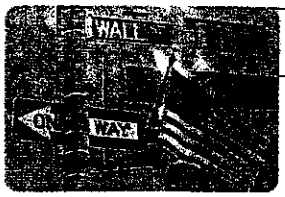
10. **Education.** In a December 2000 report, the U.S. Census Bureau listed the levels of educational attainment for Americans over 65. Create an appropriate display for these data, and write a sentence or two that might appear in a newspaper article about the report.

Education Level	Count (thousands)
No high school diploma	9,945
HS graduate, but no college	11,701
Some college, no degree	4,481
2-year degree	1,390
4-year degree	3,133
Master's degree	1,213
Ph.D. or professional degree	757

11. **Ghosts** A May 2001 Gallup Poll found that many Americans believe in ghosts and other supernatural phenomena. The poll was based on telephone responses from 1012 randomly selected adults. The table shows the percentages of people who expressed belief in various phenomena.

Phenomenon	Percent Expressing Belief
Psychic healing	54
ESP	50
Ghosts	38
Astrology	28
Channeling	15

a) Is it reasonable to conclude that 66% of those polled expressed belief in either ghosts or astrology?
b) Can you tell what percent of people did not believe in any of these phenomena? Explain.
c) Create an appropriate display for these data.



What have we learned? Quantitative Data

We've learned how to make a picture for quantitative data to help us see the story the data have to *Tell*.

- We can display the distribution of quantitative data with a *histogram*, a *stem-and-leaf* display, or a *dotplot*.
- We *Tell* what we see about the distribution by talking about *shape, center, spread,* and any *unusual features*.
- When we want to compare the distributions of a quantitative variable for two different groups, we can look at their histograms, dotplots, or stem-and-leaf displays. Using two displays drawn on the same scale, we compare their shapes, centers, and spreads.
- We can see the trend in a quantitative variable by looking at a *timeplot* of data that have been collected over time.

TERMS

Distribution The distribution of a variable gives
- the possible values of the variable.
- the frequency or relative frequency of each value.

Histogram (relative frequency histogram) A histogram uses adjacent bars to show the distribution of values in a quantitative variable. Each bar represents the frequency (or relative frequency) of values falling in an interval of values.

Stem-and-leaf display A stem-and-leaf display shows quantitative data values in a way that sketches the distribution of the data. It's best described in detail by example.

Dotplot A dotplot graphs a dot for each case against a single axis.

Shape To describe the shape of a distribution, look for
- single vs. multiple modes.
- symmetry vs. skewness.

Center A value that attempts the impossible by summarizing the entire distribution with a single number, a "typical" value.

Spread A numerical summary of how tightly the values are clustered around the "center."

Mode A hump or local high point in the shape of the distribution of a variable is called a "mode." The apparent location of modes can change as the scale of a histogram is changed.

Unimodal Having one mode. This is a useful term for describing the shape of a histogram when it's generally mound-shaped. Distributions with two modes are called **bimodal.** Those with more than two are **multimodal.**

Uniform A distribution that's roughly flat is said to be uniform.

Symmetric A distribution is symmetric if the two halves on either side of the center look approximately like mirror images of each other.

Tails The tails of a distribution are the parts that typically trail off on either side. Distributions can be characterized as having long tails (if they straggle off for some distance) or short tails (if they don't).

Skewed A distribution is skewed if it's not symmetric and one tail stretches out farther than the other. Distributions are said to be **skewed left** when the longer tail stretches to the left, and **skewed right** when it goes to the right.

Outliers Outliers are extreme values that don't appear to belong with the rest of the data. They may be unusual values that deserve further investigation, or just mistakes; there's no obvious way to tell. Don't delete outliers automatically—you have to think about them. Outliers can affect many statistical analyses, so you should always be alert for them.

Timeplot A timeplot displays data that change over time. Often, successive values are connected with lines to show trends more clearly.

SKILLS

When you complete this lesson you should:

Think

- Be able to identify an appropriate display for any quantitative variable.
- Be able to guess the shape of the distribution of a variable by knowing something about the data.

Show

- Know how to display the distribution of a quantitative variable with a stem-and-leaf display (by hand for smaller data sets), a dotplot, or a histogram (made by computer for larger data sets).
- Know how to make a timeplot of data that may vary over time.

Tell

- Be able to describe the distribution of a quantitative variable in terms of its shape, center, and spread.
- Be able to describe any anomalies or extraordinary features revealed by the display of a variable.
- Know how to compare the distributions of two or more groups by comparing their shapes, centers, and spreads.
- Know how to describe patterns over time shown in a timeplot.
- Be able to discuss any outliers in the data, noting how they deviate from the overall pattern of the data.

EXERCISES

12. Statistics in print. Find a histogram that shows the distribution of a variable in a newspaper or magazine article.
a) Does the article identify the W's?
b) Discuss whether the display is appropriate for the data.
c) Discuss what the display reveals about the variable and its distribution.
d) Does the article accurately describe and interpret the data? Explain.

13. Not a histogram. Find a graph other than a histogram that shows the distribution of a quantitative variable in a newspaper or magazine article.
a) Does the article identify the W's?
b) Discuss whether the display is appropriate for the data.
c) Discuss what the display reveals about the variable and its distribution.
d) Does the article accurately describe and interpret the data? Explain.

Example Draw a stem-and-leaf plot for the following distribution of ages:
26, 30, 29, 41, 35, 26, 34, 29, 35, 30, 25, 42, 26, 34, 41, 35

Solution You first obtain the *stem* by using only the tens' digit of each number. These are written to the left of a vertical line. Then, for each age, you record the *leaf*, or units' digit, to the right of the stem. After making an unordered stem-and-leaf plot (as shown at the left below), you should reorder the elements of each row to complete the plot (as shown at the right below).

Stem	Leaf
2	6, 9, 6, 9, 5, 6
3	0, 5, 4, 5, 0, 4, 5
4	1, 2, 1

Stem	Leaf
2	5, 6, 6, 6, 9, 9
3	0, 0, 4, 4, 5, 5, 5
4	1, 1, 2

14. Cholesterol. A study examining the health risks of smoking measured the cholesterol levels of people who had smoked for at least 25 years and people of similar ages who had smoked for no more than 5 years and then stopped. Create histograms for both groups and write a brief report comparing their cholesterol levels.

Smokers				Ex-smokers		
225	211	209	284	250	134	300
258	216	196	288	249	213	310
250	200	209	280	175	174	328
225	256	243	200	160	188	321
213	246	225	237	213	257	292
232	267	232	216	200	271	227
216	243	200	155	238	163	263
216	271	230	309	192	242	249
183	280	217	305	242	267	243
287	217	246	351	217	267	218
200	280	209		217	183	228

15. Nuclear power. For a while in the 20th century, many nuclear-powered electrical generating plants were built, but then growing environmental concerns and construction costs led to increasing reliance on other forms of energy. The table shows the dates of completion (in months after January 1967) and costs (in thousands of dollars per megawatt) of 12 nuclear generators.

Time of Completion (months after Jan 1, 1967)	Construction Cost ($1000/mW)
2	35
3	28
10	32
12	60
17	56
19	63
21	62
26	81
30	84
32	79
41	88
47	80

a) Create a stem-and-leaf display of the costs.
b) Describe the distribution.
c) Create a timeplot of the costs.
d) What information about the construction of nuclear plants can you see from the timeplot that is not obvious in the stem-and-leaf display?

What have we learned? Center, Spread

We've learned how to summarize distributions of quantitative variables numerically.

- The 5-number summary displays the two quartiles, median, min, and max for a variable.
- Measures of center for a distribution include the mean and median.
- Measures of spread include the range, IQR, and standard deviation.
- We'll report the median and IQR when the distribution is skewed. If it's symmetric, we'll summarize with the mean and standard deviation (and possibly the median and IQR, as well).

We've also learned to display distributions with boxplots.

- A boxplot reveals some features of a distribution not easily seen in a histogram—the center, the middle 50%, and outliers. Histograms, though, are better at showing the shape.
- Boxplots are very effective for comparing groups graphically. As always, when we compare groups we'll discuss their shapes, centers, spreads, and unusual features.

TERMS

Center We summarize the center of a distribution with the mean or the median.

Median The median is the middle value with half of the data above and half below it.

Spread We summarize the spread of a distribution with the standard deviation, interquartile range, and range.

Range The difference between the lowest and highest values in a data set. Range = max − min.

Quartile The lower quartile (Q1) is the value with a quarter of the data below it. The upper quartile (Q3) has a quarter of the data above it. The median and quartiles divide data into four equal parts.

Interquartile range (IQR) The IQR is the difference between the first and third quartiles. IQR = Q3 − Q1.

Percentile The *i*th percentile is the number that falls above *i*% of the data.

5-number summary A 5-number summary for a variable consists of:

- The minimum and maximum
- The quartiles: Q1 and Q3
- The median

Boxplot A boxplot displays the 5-number summary as a central box with whiskers that extend to the non-outlying data values. Boxplots are particularly effective for comparing groups.

Mean The mean is found by summing all the data values and dividing by the count.

Variance The variance is the sum of squared deviations from the mean, divided by the count minus one.

Standard deviation The standard deviation is the square root of the variance.

Comparing distributions When comparing the distribution of several groups, consider their

- Shape
- Center
- Spread

Comparing boxplots When comparing groups with boxplots,

- Compare the medians; which group has the higher center?
- Compare the IQRs; which group is more spread out?
- Judged by the size of the IQRs, are the medians very different?
- Check for possible outliers. Identify them if you can.
- Know how to compare the distributions of two or more groups by comparing their shapes, centers, and spreads.
- Be able to compare two or more groups by comparing their boxplots.
- Know how to use the 1.5 IQR rule to identify possible outliers. Interpret outliers found in boxplots made on a computer.

AP Statistics

SKILLS

When you complete this lesson you should: Center, Spread

Think

- Be able to select a suitable measure of center and a suitable measure of spread for a variable based on information about its distribution.
- Know the basic properties of the median: The median divides the data into the half of the data values that are below the median and the half that are above the median.
- Know the basic properties of the mean: The mean is the point at which the histogram balances.
- Know that the standard deviation summarizes how spread out all the data are around the mean.

Show

- Know how to compute the mean and median of a set of data.
- Know how to compute the standard deviation and IQR of a set of data.
- Be able to create a 5-number summary of a variable.
- Be able to construct a boxplot by hand from a 5-number summary.
- Understand that the median and IQR resist the effects of outliers, while the mean and standard deviation do not.
- Understand that in a skewed distribution, the mean is pulled in the direction of the skewness (toward the longer tail) relative to the median.

Tell

- Know how to describe summary measures in a sentence. In particular, know that the common measures of center and spread have the same units as the variable that they summarize, and should be described in those units.
- Be able to describe the distribution of a quantitative variable with a description of the shape of the distribution, a numerical measure of center, and a numerical measure of spread. Be sure to note any unusual features, such as outliers, too.

EXERCISES

16. In the news. Find an article in a newspaper or a magazine that discusses an "average."

a) Does the article discuss the W's for the data?
b) What are the units for the variable?
c) Is the average used the median or the mean? How can you tell?
d) Is the choice of median or mean appropriate for the situation? Explain.

17. In the news II. Find an article in a newspaper or a magazine that discusses a measure of spread.

a) Does the article discuss the W's for the data?
b) What are the units for the variable?
c) Does the article use the range, IQR, or standard deviation?
d) Is the choice of measure of spread appropriate for the situation? Explain.

18. Singers. The frequency table shows the heights (in inches) of 130 members of a choir.

Height	Count	Height	Count
60	2	69	5
61	6	70	11
62	9	71	8
63	7	72	9
64	5	73	4
65	20	74	2
66	18	75	4
67	7	76	1
68	12		

a) Find the 5-number summary for these data.
b) Display these data with a boxplot.
c) Find the mean and standard deviation.
d) Display these data with a histogram.
e) Write a few sentences describing the distribution of heights.

19. Caffeine. A student study of the effects of caffeine asked volunteers to take a memory test 2 hours after drinking soda. Some drank caffeine-free cola, some drank regular cola (with caffeine), and others drank a mixture of the two (getting a half-dose of caffeine). Here are the 5-number summaries for each group's scores (number of items recalled correctly) on the memory test:

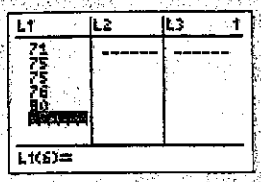
	n	Min	Q1	Median	Q3	Max
No caffeine	15	16	20	21	24	26
Low caffeine	15	16	18	21	24	27
High caffeine	15	12	17	19	22	24

a) Describe the W's for these data.
b) Name the variables and classify each as categorical or quantitative.
c) Create parallel boxplots to display these results as best you can with this information.
d) Write a few sentences comparing the performances of the three groups.

TI Tips

AP Statistics

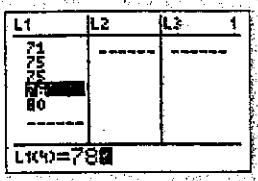
You'll need to be able to enter and edit data in your calculator. Here's how.



To enter data:

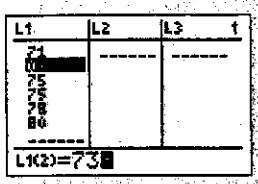
Hit the **STAT** button, and choose **EDIT** from the menu. You'll see a set of columns labeled **L1**, **L2**, and so on. Here is where you can enter, change, or delete a set of data.

Let's enter the heights (in inches) of the five starting players on a basketball team: 71, 75, 75, 76, and 80. Move the cursor to the space under **L1**, type in 71, and hit **ENTER** (or the down arrow). There's the first player. Now enter the data for the rest of the team.



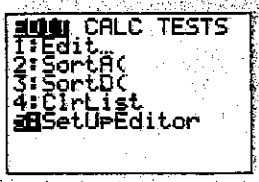
To change a datum:

Suppose the 76" player grew since last season; his height should be listed as 78". Use the arrow keys to move the cursor onto the 76, then change the value and **ENTER** the correction.



To add more data:

We want to include the sixth man, 73" tall. It would be easy to simply add this new datum to the end of the list. However, sometimes the order of the data matters, so let's place this datum in numerical order. Move the cursor to the desired position (atop the first 75). Hit **2nd INS**, then **ENTER** the 73 in the new space.



To delete a datum:

The 78" player just quit the team. Move the cursor there. Hit **DEL**. Bye.

To clear the datalist:

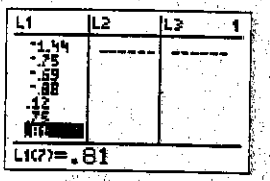
Finished playing basketball? Move the cursor atop the **L1**. Hit **CLEAR**, then **ENTER** (or down arrow). You should now have a blank datalist, ready for you to enter your next set of values.

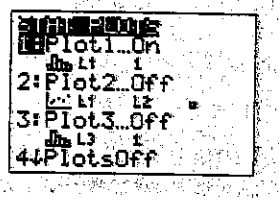
Lost a datalist?

Oops! Is **L1** now missing entirely? Did you delete **L1** by mistake, instead of just *clearing* it? Easy problem to fix: buy a new calculator. No? OK, then simply go to the **STAT EDIT** menu, and run **SetUpEditor** to recreate all the lists.

AP Statistics

TI Tips



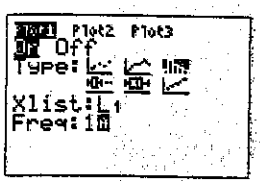


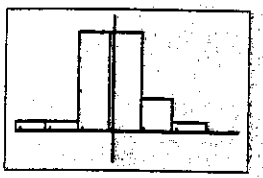
Your calculator can create histograms. First you need some data. Remember the Enron stock price changes at the start of the chapter? Relax—you don't have to type in all of those values to see how it works. We'll just use the Enron data for 1997 and 1998. Enter the 24 monthly values from these two rows of data into L1:

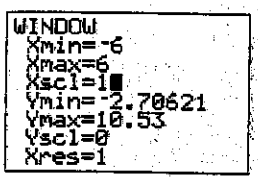
1997	−1.44	−0.75	−0.69	−0.88	0.12	0.75	0.81	−1.75	0.69	−0.22	−0.16	0.34
1998	0.78	0.62	2.44	−0.28	2.22	−0.50	2.06	−0.88	−4.50	4.12	1.16	−0.50

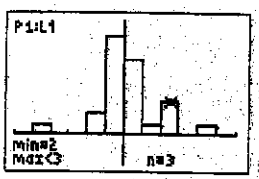
Now set up the calculator's plot:

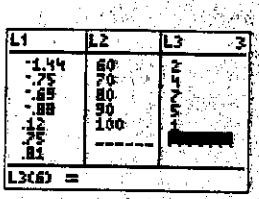
- Go to 2nd STATPLOT, choose Plot1, then ENTER.











- In the Plot1 screen choose On, select the little histogram icon, then specify Xlist:L1 and Freq:1.
- Be sure to turn off any other graphs the calculator may be set up for. Just hit the Y= button, and deactivate any functions seen there.

All set? To create your preliminary plot go to ZOOM, select 9:ZoomStat, and then ENTER.

You now see the calculator's initial attempt to create a histogram of these data. Not bad. We can see that the distribution is roughly symmetric. But it's hard to tell exactly what this histogram shows, right? Let's fix it up a bit.

- Under WINDOW, let's reset the bins to convenient, sensible values. Try Xmin=−6, Xmax=6 and Xscl=1. That specifies the range of values along the *x*-axis and makes each bar span one dollar.
- Hit GRAPH (*not* ZoomStat—this time we want control of the scale!).

Look better? Note that you can now find out exactly what the bars indicate by activating TRACE and then moving across the histogram using the arrow keys. For each bar the calculator will indicate the range of values and the number of data values in that range. We see that the stock price increased between \$2 and \$3 during three of the months.

Play around with the WINDOW settings. A different Ymax will make the bars appear shorter or taller. What happens if you set the bar width (Xscl) smaller? Or larger? You don't want to lump lots of values into just a few bins or make so many bins that the overall shape of the histogram is not clear. Choosing the best bar width takes practice.

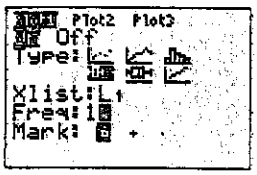
Finally, suppose the data are given as a frequency table. Consider a set of test scores, with two grades in the 60s, four in the 70s, seven in the 80s, five in the 90s, and one 100. Enter the group cutoffs 60, 70, 80, 90, 100 in L2 and the corresponding frequencies 2, 4, 7, 5, 1 in L3. When you set up the histogram STATPLOT, specify Xlist:L2 and Freq:L3. Can you specify the WINDOW settings to make this histogram look the way you want it? (By the way, if you get a DIM MISMATCH error, it means you can't count. Look at L2 and L3; you'll see the two lists don't have the same number of entries. Fix the problem by correcting the data you entered.)

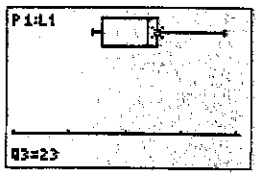
TI Tips

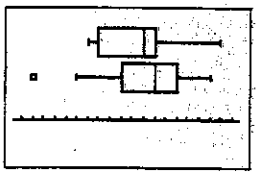
Let's make comparative boxplots for the performances of fourth grade boys and girls on an agility test. The test asks them to jump from side to side across a set of parallel lines, counting the number of lines they can clear in 30 seconds. Here are the results:

Boys: 22, 17, 18, 29, 22, 22, 23, 24, 23, 17, 21

Girls: 25, 20, 12, 19, 28, 24, 22, 21, 25, 26, 25, 16, 27, 22







Enter these data in `L1` (*Boys*) and `L2` (*Girls*).

Set up `STATPLOT`'s `Plot1` to make a boxplot of the boys' data:

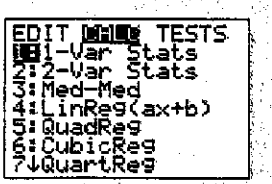
- turn the plot `On`;
- choose the first boxplot icon (you want your plot to indicate outliers);
- specify `Xlist:L1` and `Freq:1`, and select the `Mark` you want the calculator to use for displaying any outliers.

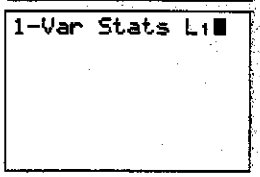
Use `ZoomStat` to display the boxplot for *Boys*. You can now `TRACE` to see the statistics in the five-number summary. Try it!

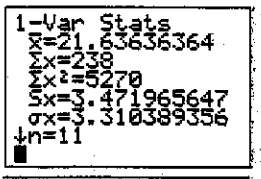
As you did for the boys, set up `Plot2` to display the girls' data. This time when you use `ZoomStat` with both plots turned on, the display shows the parallel boxplots. See the outlier?

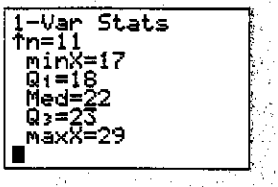
This is a great opportunity to practice your "Tell" skills. How do these fourth graders compare in terms of agility?

TI Tips









Your calculator can easily find all the numerical summaries of data. To try it out, you simply need a set of values in one of your datalists. We'll illustrate using the boys' agility test results from this chapter's earlier TI Tips (still in `L1`), but you can use any data currently stored in your calculator.

- Under the `STAT CALC` menu, select `1-Var Stats` and hit `ENTER`.
- Specify the location of your data, creating a command like `1-Var Stats L1`.
- Hit `ENTER` again.

Voila! Everything you wanted to know, and more. Among all of the information shown, you are primarily interested in these statistics: $\bar{x}$ (the mean), `Sx` (the standard deviation), `n` (the count), and—scrolling down—`minX` (the smallest datum), Q_1 (the first quartile), `Med` (the median), Q_3 (the third quartile), and `maxX` (the largest datum).

Sorry, but the TI doesn't explicitly tell you the range or the IQR. Just subtract: IQR = $Q_3 - Q_1 = 23 - 18 = 5$. What's the range?

By the way, if the data come as a frequency table with the values stored in, say, `L4` and the corresponding frequencies in `L5`, all you have to do is ask for `1-Var Stats L4,L5`.