Introduction

SAY IT WITH CHARTS

It's 9:00 a.m. on the third Tuesday of the month, time for the monthly meeting of the Steering Committee. To set the rest of the day's proceedings in perspective, the committee chairman has asked a bright, fast-rising manager—let's call him Frank—to prepare a brief presentation on the state of the industry in which we compete and our company's performance as a stepping-stone for new investment opportunities.

Intent on doing a good job, Frank has done much research, worked on his story line, and prepared a series of visual aids to help him say it with charts. Like most of us, Frank realizes that charts are an important form of language. They're important because, when well conceived and designed, they help us communicate more quickly and more clearly than we would if we left the data in tabular form.

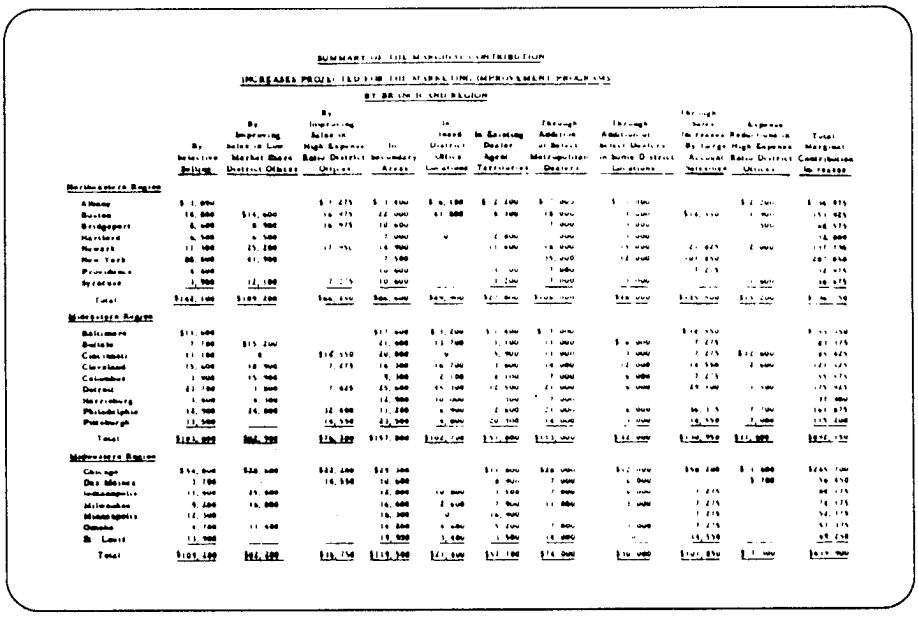
When charts *aren't* well conceived or designed, as we're about to see in Frank's examples, they serve more to confuse than to clarify. Let's sit with the audience and listen to Frank's presentation as we comment, quietly, on the effectiveness of his visuals.

Frank begins: Good morning ladies and gentlemen. My purpose is to present a brief overview of our industry and our company's performance. My objective is to gain your support for expanding into developing countries. I've designed a few visual aids to better place my findings in perspective.

First, let me point out that we compete in a healthy industry. As you can clearly see from this exhibit, for the 11 measures of performance shown across the top and the three types of companies within the industry listed down the side, performance is excellent.

And there you sit in the audience, wondering whether your eyesight is failing, as you try in vain to read the numbers.

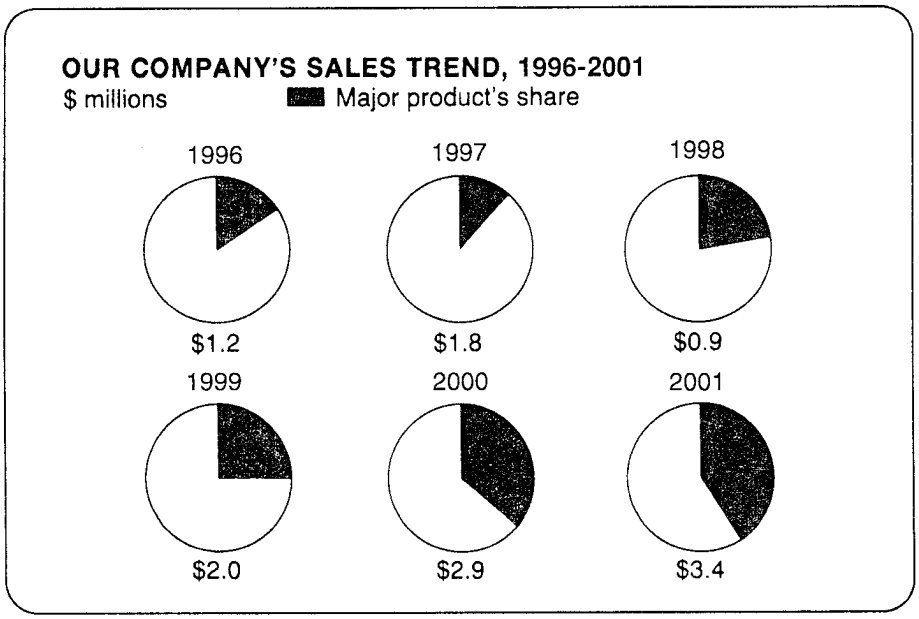
▶ 1



Frank continues: Within the industry, our performance has been outstanding. For instance, our sales have grown, considerably since 1996, in spite of the decline in 1998, which, as you know, was the result of the strike.

"Oops," you whisper," did I just miss something? I could swear I heard Frank say that sales have grown considerably, but what I see is a series of pie charts that show our major product's share increasing. Oh! Wait a minute. I see! He's referring to the figures underneath each pie. . . ."

▶ 2

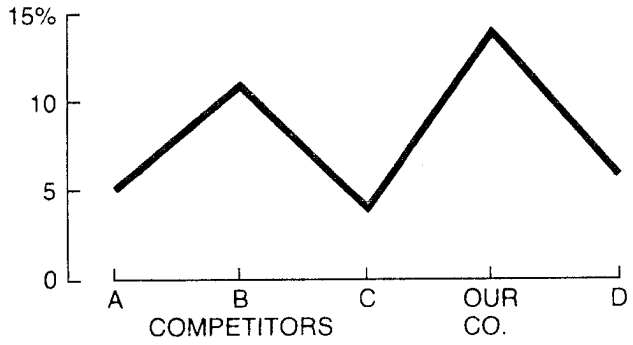


Frank goes on: Compared with our four major competitors, we rank first in return on investment with 14 percent. . . .

"What ranks first? Who ranks first?" you say. "From the visual, I thought the point was that ROI has been fluctuating."

▶ 3

COMPARISON OF 2001 RETURN ON INVESTMENT, BY COMPANY

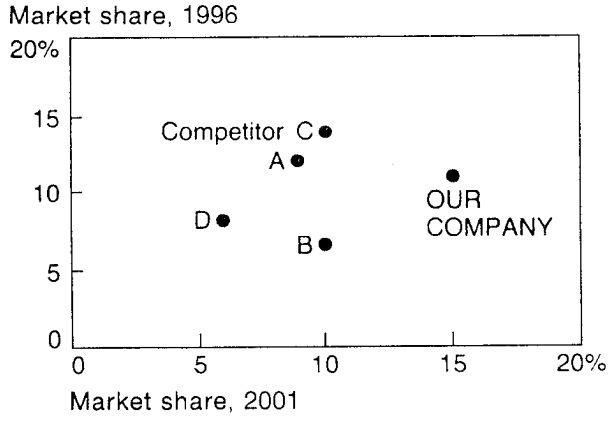


. . . and our market share has increased since 1996 along with that of one competitor, while the other three lost share.

You sigh in frustration "Has someone spiked my orange juice? Why do I feel that my eyes and ears aren't talking to one another, that I'm receiving mixed signals? Is it perhaps that the visuals I see are not supporting the messages I hear?"

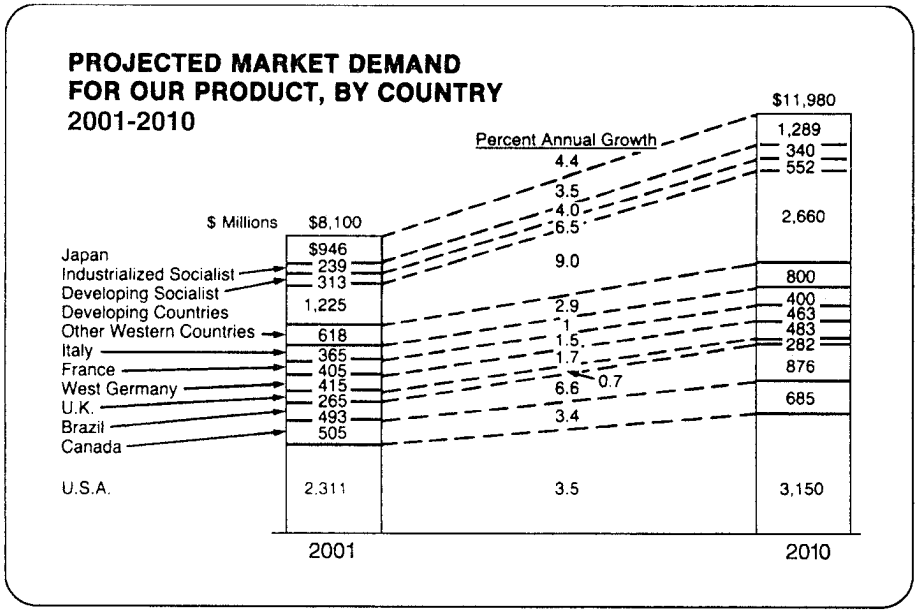
▶ 4

MARKET SHARE TRENDS, BY COMPANY 1996-2001



Frank proceeds: Given these sales, ROI, and market-share trends, we recommend expanding the selling effort for our major product into developing countries. We believe these markets hold considerable potential. Since you may not understand this visual, let me explain it. What I've done is shown the total size of the market worldwide in 2001 and projected it to 2010. On the basis of a lot of research, we forecast that the market will increase from $8 billion to over $11 billion. Then I divided the totals by the 11 countries that make up the market and showed the size of each. Then I calculated the average annual compound

▶ 5



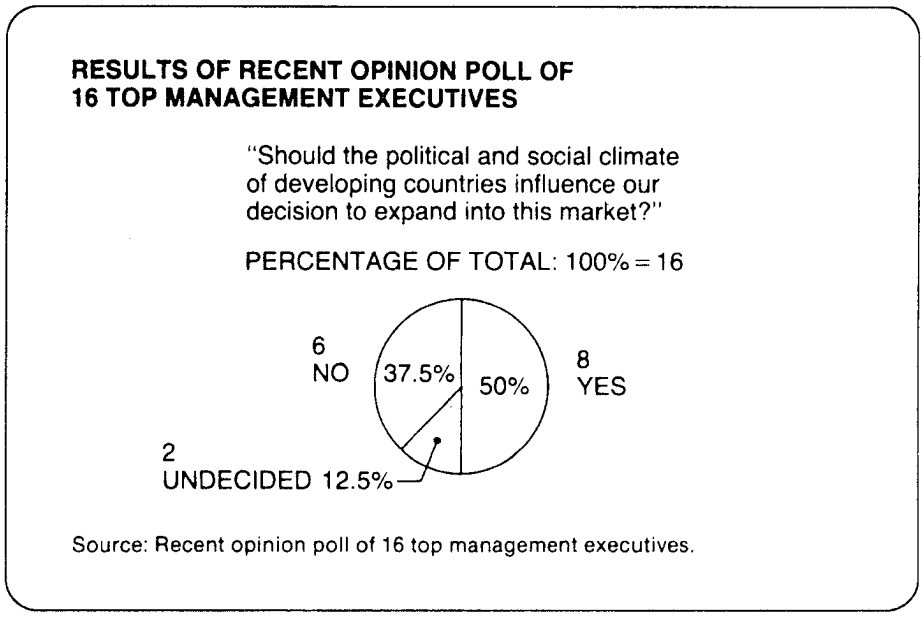
growth rate for each country and listed it in the center. As you can see from these figures, the developing countries are projected to show the fastest growth.

Now I feel your nudge and hear your aside: "Isn't it remarkable how these visual aids designed to aid speakers require so much speaker aid to be understood. I always thought that a picture was worth 1,000 words, not that it required them."

Frank: If we are to move ahead, however, we need first to persuade top management that the political and social climate in these developing countries will not interfere with our plans. A recent poll of 16 top management members reveals they are almost evenly split for and against investing in these countries.

By now, <u>your</u> social unrest is hard to contain, and the pie chart conjures visions of dessert at lunch.

▶ 6



Frank's intention was right: he meant to use charts to support his presentation. His execution was not good: he developed a series of charts that were illegible or unintelligible and that detracted from his presentation. Let's review his charts and see why they don't work.

Visual number ▶ **1** is illegible. Like all illegible visuals, it suffers from the APK—anxious parade of knowledge—syndrome. This is usually the case when the presenter is more concerned with what's been put *into* the chart than in what the audience gets *out* of it.

What Frank didn't realize is that a chart used in a visual presentation must be at least twice as simple and four times as bold as one used in a report. It's the same as the distinction between a billboard that must be read and understood in the time you drive past it and a magazine advertisement that you can study in detail.

At the other extreme is the last chart, number ▶ **6,** which is so simple that it is not needed; the message could have been expressed with words alone. In addition to overly simple charts, there are other situations when you're better off without a chart:

1. Sometimes the chart denotes a sense of accuracy that may be misleading, as is the case with projections or ranges that may be tenuous.

2. Sometimes there are sets of data that the audience or reader has become comfortable with, such as the company's profit and loss statement, and changing the form to a chart could be confusing.

3. Certain individuals may be unaccustomed to, or resistant to, or skeptical about the use of charts.

With charts, a good rule is "fewer is better." Producing charts is time-consuming and expensive. Also, the more charts we use, the less people remember. Use one chart in a report or presentation, and it will receive 100 percent of the audience's attention; use 100, and none will be memorable.

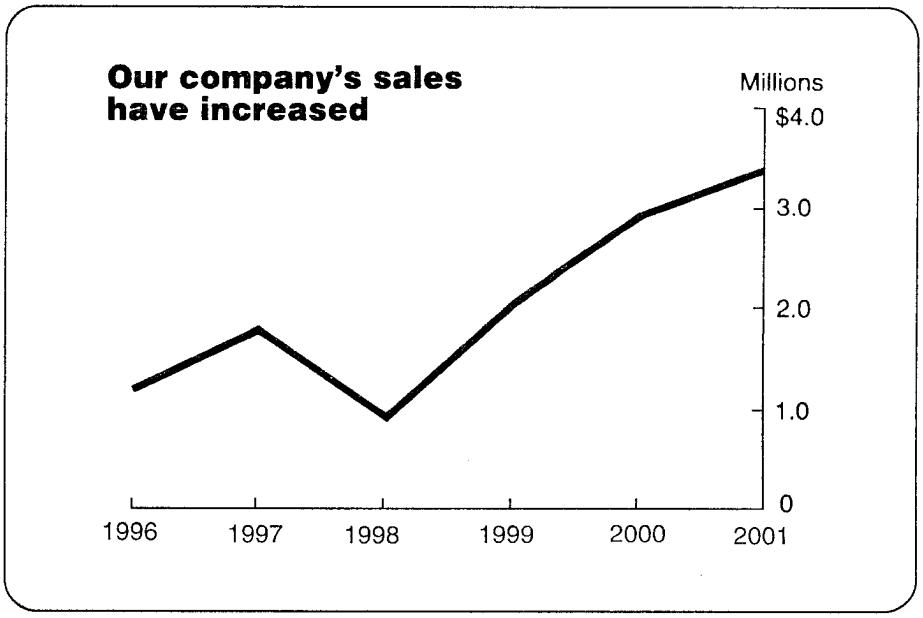
Chart 5, the one describing the world market, is what I call a "charttable"; it can't make up its mind whether to be a chart or a table and decides to be both. The hope is that if the chart doesn't work, the information may. In most cases, neither does. Undoubtedly, this chart helped Frank figure out the important relationships—in this case, the comparison of projected growth rates by country. However, Frank did not translate the data from the form that helped him analyze the problem to a simpler chart that emphasized the results of his analyses.

The remaining three charts, ▶ **2, 3,** and **4,** suffer from what is probably the major problem facing most of us when we translate data into chart form: the wrong form for the message. In Chart 2, pie charts were chosen where a line chart is called for; in Chart 3, a line chart was used where a bar chart would be better; in 4, we see a dot chart instead of column charts.

Here is what these three charts should have looked like to support the spoken message more quickly and clearly.

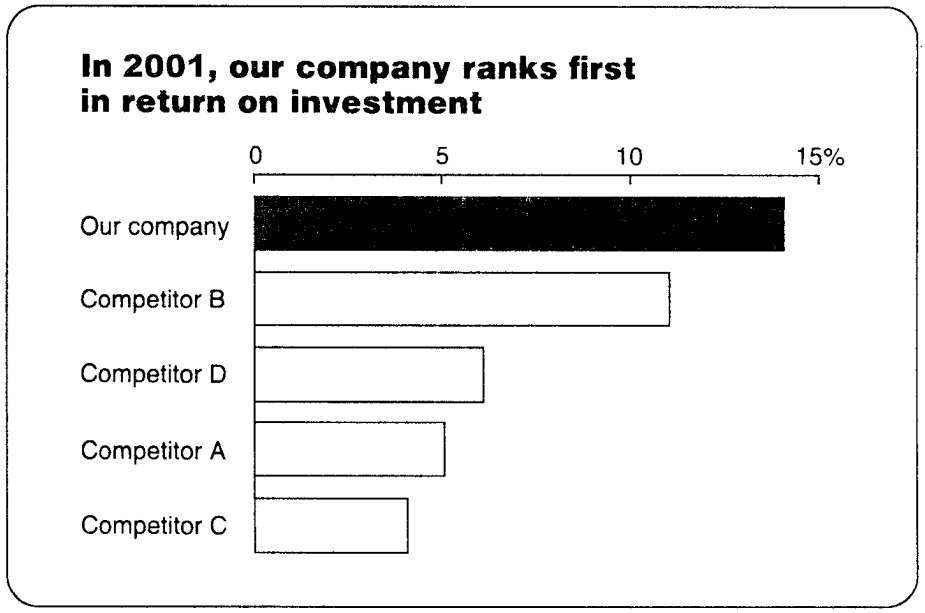
Sales have increased from $1.2 million in 1996 to $3.4 million in 2001 despite the decline in 1998 caused by the strike.

▶ 2



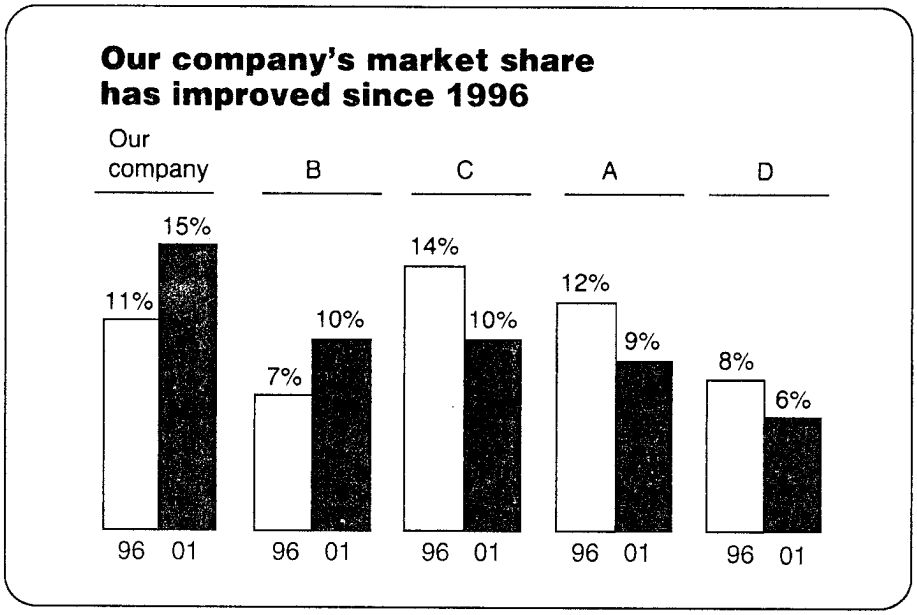
Compared with our four major competitors, we rank first with a 14 percent return on investment in 2001.

▶ 3



Our market share has increased 4 percentage points from 11 percent in 1996 to 15 percent today. Of our four competitors, B also improved, while C, A, and D lost share.

▶ 4



Now these charts work. In each case, the chart form supports the message expressed in the title, and the title reinforces the point the chart demonstrates. In all cases, the message comes across faster and better than it would if the data were left in table form.

And there you have the purpose of this book. Its goal is to help you say it with charts by choosing and using charts that will work for you and your audience no matter where the charts are used—be it in business presentations or reports, in your management information system, in computer graphics software packages, in annual reports, or in magazines or newspaper articles.

In the next portion of this section, we'll work our way through the process that moves us from data to chart. In Section II, we'll examine a portfolio of finished charts that you can refer to for ideas the next time the need arises.

In Section III, we'll show how to convey your message using concept visuals, and visual metaphors.

In Section IV, we'll demonstrate how to design charts for on-screen presentations.

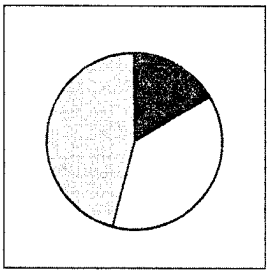
As I said before, charts *are* an important form of language. But as is true of any language we want to become proficient in, it takes time and patience to learn the vocabulary, and practice until the skill becomes second nature. Since no one learns by reading, only by doing, I've incorporated work projects so you can practice as you read. So take pencil in hand, and let's move on to the process of choosing charts.

Section 1

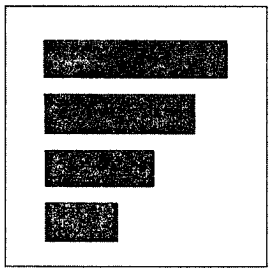
CHOOSING CHARTS

No matter how many business graphics we see in various kinds of communications—including tables, organization charts, flow diagrams, matrixes, maps—when it comes to quantitative charts, there are only five basic chart forms to choose from. As shown here in simple terms, these are:

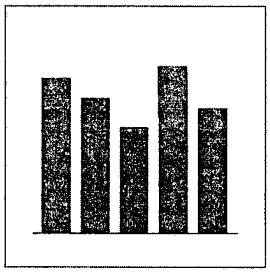
The pie chart



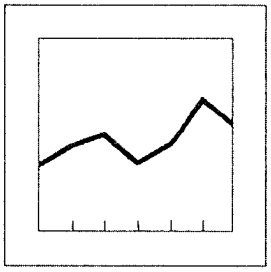
The bar chart



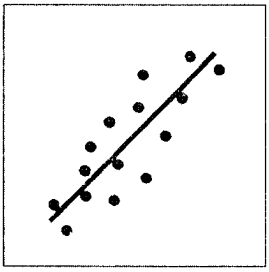
The column chart



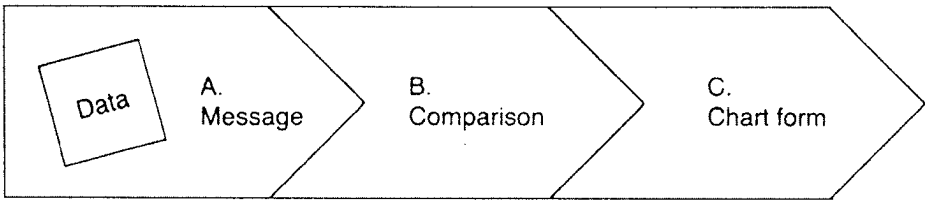
The line chart



The dot chart



Now we know where we're going; the question is, how do we get there? With the following diagram, let me summarize the process that moves us from the data we start with to the specific chart we end with.



Step A:

DETERMINE YOUR MESSAGE

(*from data to message*).

The key to choosing the appropriate chart form is for *you*, as the designer, to be clear, first and foremost, about the specific point you want to make.

Step B:

IDENTIFY THE COMPARISON

(*from message to comparison*).

The message you've determined will always imply one of five basic kinds of comparison: component, item, time series, frequency distribution, or correlation.

Step C:

SELECT THE CHART FORM

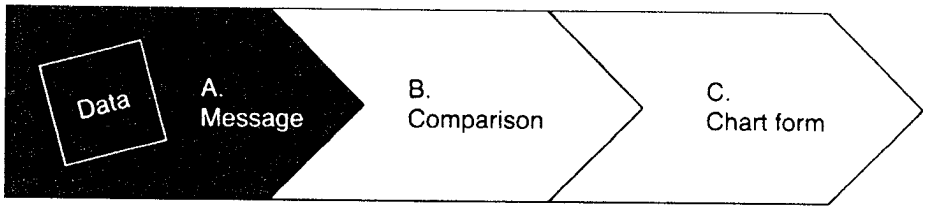
(*from comparison to chart*).

Each comparison will lead, in turn, to one of the five chart forms.

Let's discuss each step in detail.

A. DETERMINE YOUR MESSAGE

(from data to message)



Choosing a chart form without a message in mind is like trying to color coordinate your wardrobe while blindfolded.

Choosing the correct chart form depends completely on *your* being clear about what *your* message is. It is not the data—be they dollars, percentages, liters, yen, etc.—that determine the chart. It is not the measure—be it profits, return on investment, compensation, etc.—that determines the chart. Rather, it is *your* message, what *you* want to show, the specific point *you* want to make.

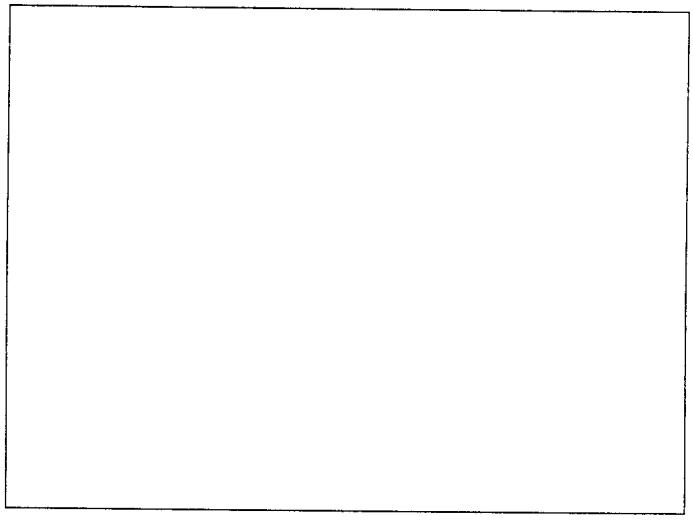
To stress the importance of this first step, sketch as many charts as you can think of in the empty boxes on the next two pages using the data (percentage of sales by region for each company) shown in the upper right-hand box. Don't worry about accuracy—your goal is to draw as many charts as you can before turning to page 14.

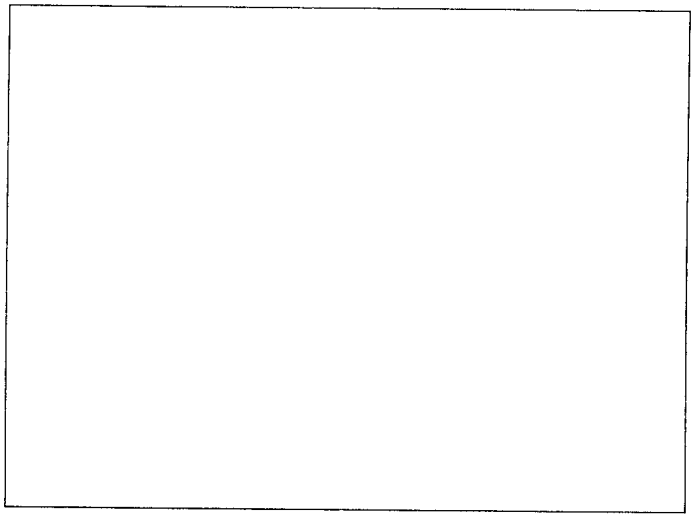
PROJECT

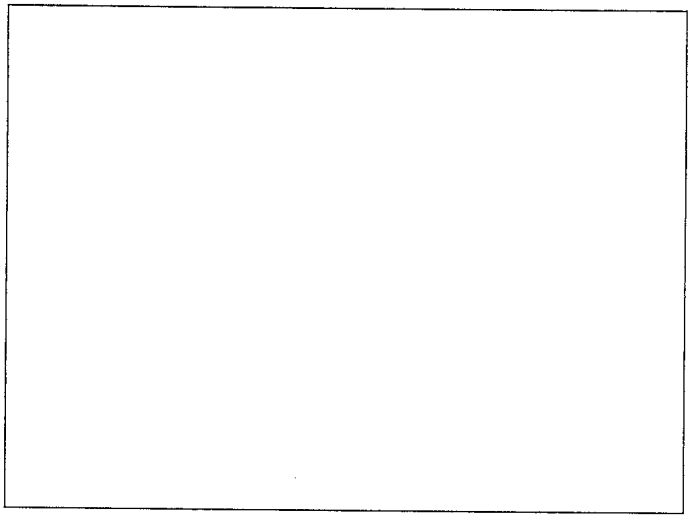
Sketch as many charts as you can think of using these data: the more the better.

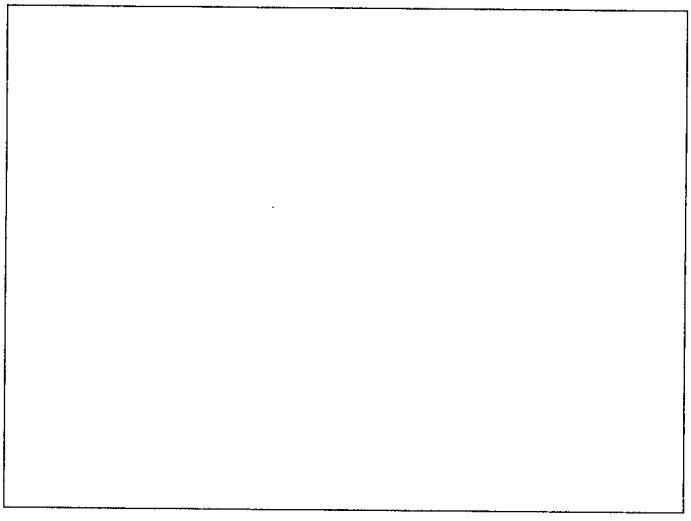
Percentage of January Sales by Region

	Co. A	Co. B
North	13%	39%
South	35	6
East	27	27
West	25	28



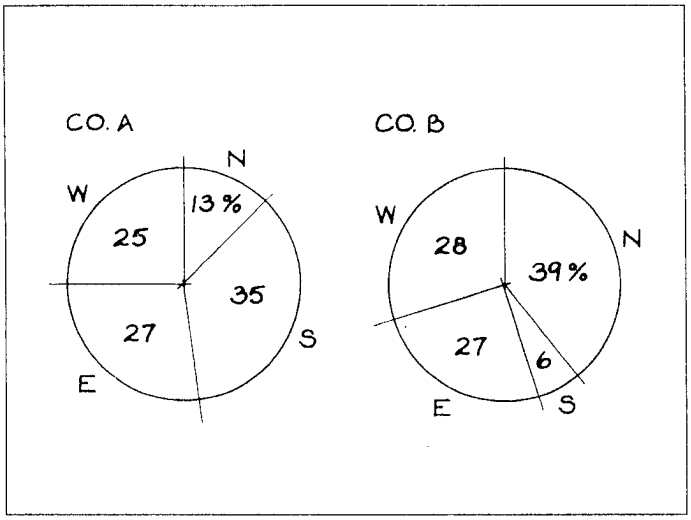




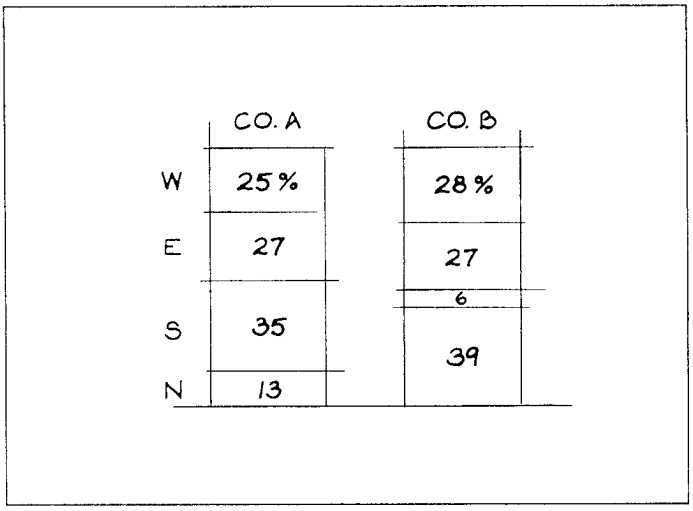


WHICH CHART WOULD YOU CHOOSE?

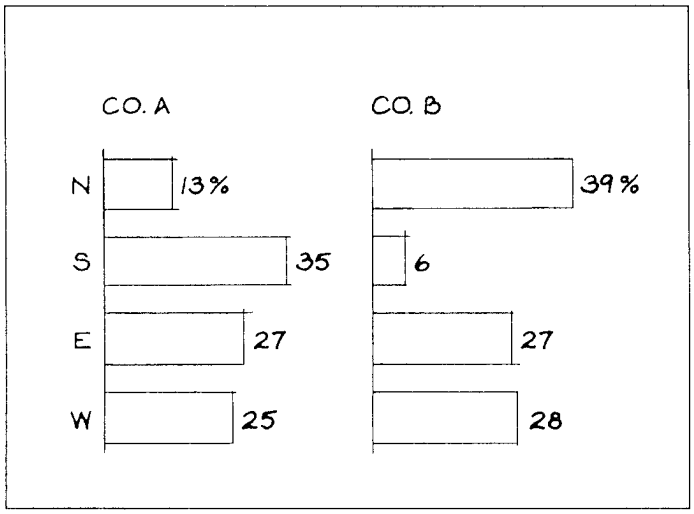
▶ 1



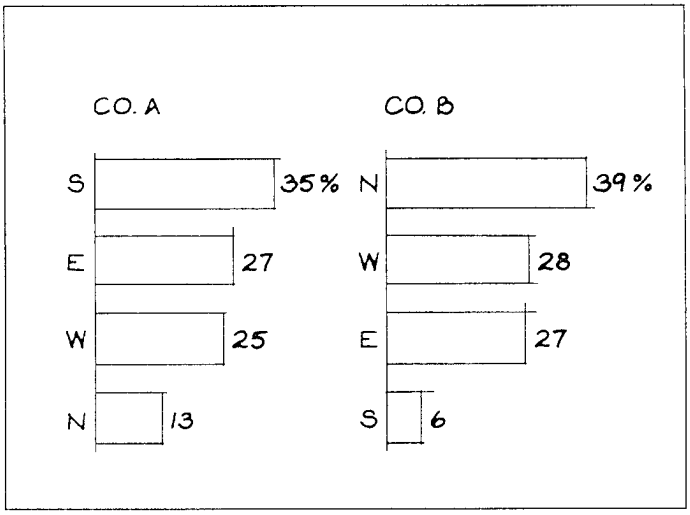
▶ 2



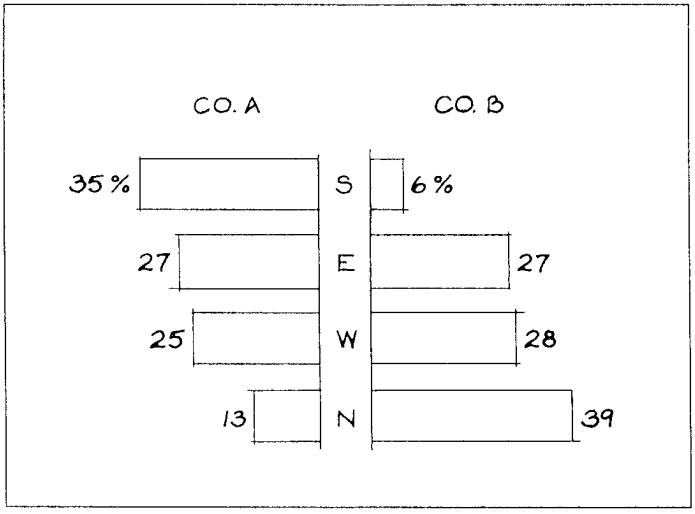
▶ 3



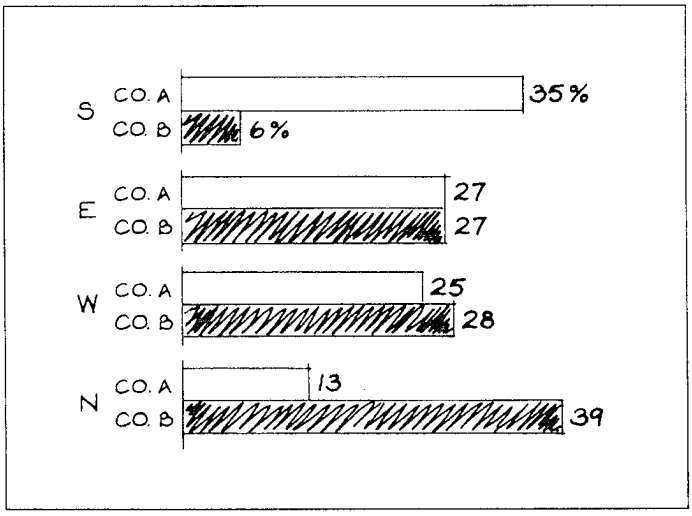
▶ 4



▶ 5



▶ 6



The charts shown on the facing page may be among those you sketched. All the better if you thought of others. But a question remains.

WHICH CHART WOULD YOU CHOOSE?

It all depends! It all depends on the specific point *you* want to make—*your* message. Each chart shown, simply as a function of the way it's organized, is best equipped to emphasize a particular message.

For instance, showing the data as a couple of pie charts or 100 percent columns, you would be emphasizing that:

▶ **1** ▶ **2** The mix of sales is different for Companies A and B.

Or you may have shown the data as two sets of bar charts, sequencing the bars in the order the data were presented in the table. Now the chart is stressing the message that:

▶ **3** The percentage of sales for both Companies A and B varies by region.

On the other hand, you could have ranked the percentage of sales for each company in descending (or ascending) order, now stressing the point that:

▶ **4** Company A is highest in the South; Company B is highest in the North. Or, Company A is lowest in the North; Company B is lowest in the South.

By structuring the bars in a mirror image around the regions, we now demonstrate that:

▶ **5** Company A's share of sales is highest in the South where Company B's is the weakest.

By grouping the bars against a common base, we now compare the gaps by region, showing that:

▶ **6** In the South, Company A leads B by a wide margin; in the East and West, the two are competitive; in the North, A lags B.

Now, it's possible—even probable—that in the early stages of deciding what your message should be, you may need to sketch a number of charts that look at the data from various points of view. A more efficient approach is to highlight the aspect of the data that seems most important and settle on the message that brings out that aspect.

For example, looking at this simplified table, there are three possible aspects of the data that could be highlighted and turned into messages.

Your attention might focus on the overall sales trend from January through May; how the dollar value of sales has changed over time. In that case, your message would be that "Sales have risen steadily since January."

Sales by Product, $000

	Product			
	A	B	C	Total
Jan.	88	26	7	121
Feb.	94	30	8	132
Mar.	103	36	8	147
Apr.	113	39	7	159
May	122	40	13	175

On the other hand, you might want to focus on a single point in time. Reading the figures across for May, for example, you might take note of the ranking of sales for Products A, B, and C. In that case, your message could be: "In May, sales of Product A exceeded those of B and C by a wide margin."

Sales by Product, $000

	Product			
	A	B	C	Total
Jan.	88	26	7	121
Feb.	94	30	8	132
Mar.	103	36	8	147
Apr.	113	39	7	159
May	122	40	13	175

Looking at the same May data from yet another perspective, you might focus on the percentage of total sales accounted for by each product. Then your message might be: "In May, Product A accounted for the largest share of total company sales."

Sales by Product, $000

	Product			
	A	B	C	Total
Jan.	88	26	7	121
Feb.	94	30	8	132
Mar.	103	36	8	147
Apr.	113	39	7	159
May	122	40	13	175
	70%	23%	7%	100%

Note that for these last two examples, we used nearly the same aspect of the data to come up with different messages. The decision to emphasize ranking or share is up to you, and that decision will give you your message.

Suppose you have other data from the same company.

This table shows the distribution of sales by size of sale at one point in time, May. Here your message might be: "In May, most sales were in the $1,000 to $2,000 range."

Number of Sales by Size of Sale, May

Size of sale	Number of sales
<$1,000	15
1,000-1,999	30
2,000-2,999	12
3,000-3,999	8
4,000+	5

This last set of data shows the relationship of the salesperson's experience to the sales he or she generates. Noting that salesperson P, with only two years' experience, generates $23,000 in sales, while salesperson Q, with more than twice the experience, generates only one-quarter the volume would indicate the message that: "There is no relationship between sales and experience."

Relationship of Salesperson's Experience to Sales

Sales-person	Years of experience	Amount of sales
P	2	$23,000
Q	5	6,000
R	7	17,000
S	15	9,000
T	22	12,000

As we've seen, this first step, determining your message, must be completed before you can select the appropriate chart form. Having spent all that time and energy doing so, you may as well make the most of the effort and *let the message become the title of the chart.* Let me elaborate.

For many of the charts we see, the title is little more than a cryptic heading, such as:

COMPANY SALES TREND

PRODUCTIVITY BY REGION

PERCENTAGE OF ASSETS BY DIVISION

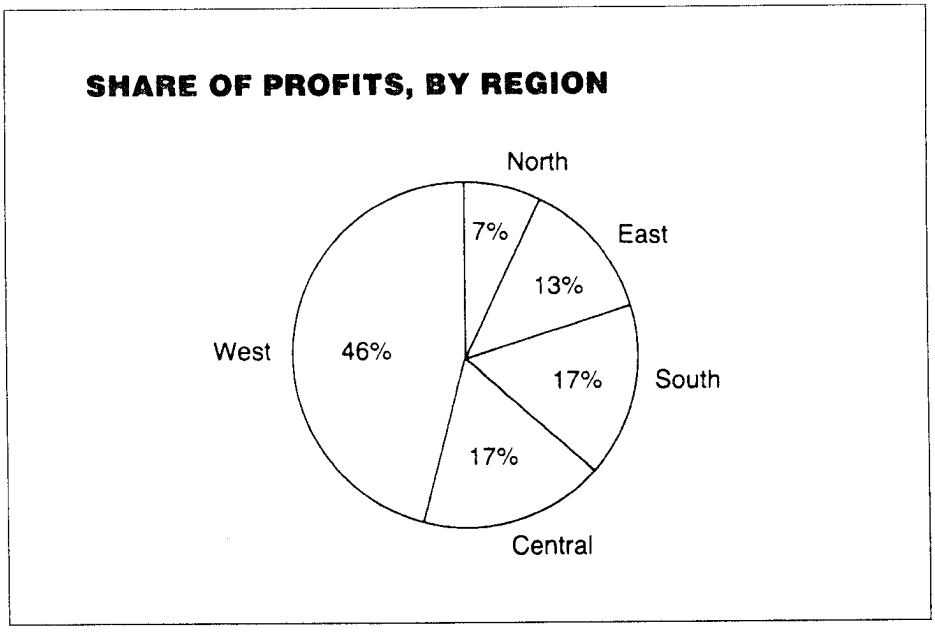
DISTRIBUTION OF EMPLOYEES BY AGE

RELATIONSHIP OF COMPENSATION

TO PROFITABILITY

These titles describe the subject of the chart, but they don't say what's important about it. What about sales performance? What about the distribution of employees? What about the relationship between compensation and profitability? Don't keep it a secret; let your message head the chart. In so doing, you reduce the risk that the reader will misunderstand, and you make sure he or she **focuses on the aspect of the data you want to emphasize.**

Let's look at a couple of examples that demonstrate the difference and establish the advantage of message titles over the topic titles just shown.



Here, the title states the topic of the chart, leaving you to determine the significance of what the chart shows. Studying the chart, most readers would probably focus on the West, believing the message to be emphasized is that the **"West accounts for almost half of profits."**

However, that may not be the point that the designer wants you to focus on. He or she may wish to stress that the *"North* generates the smallest share of profits." In short, with the topic title, you run the risk of being misunderstood. Substituting the message title, **"North generates the smallest share of profits"** reduces that risk by focusing the reader's attention on the aspect of the data we want to stress.

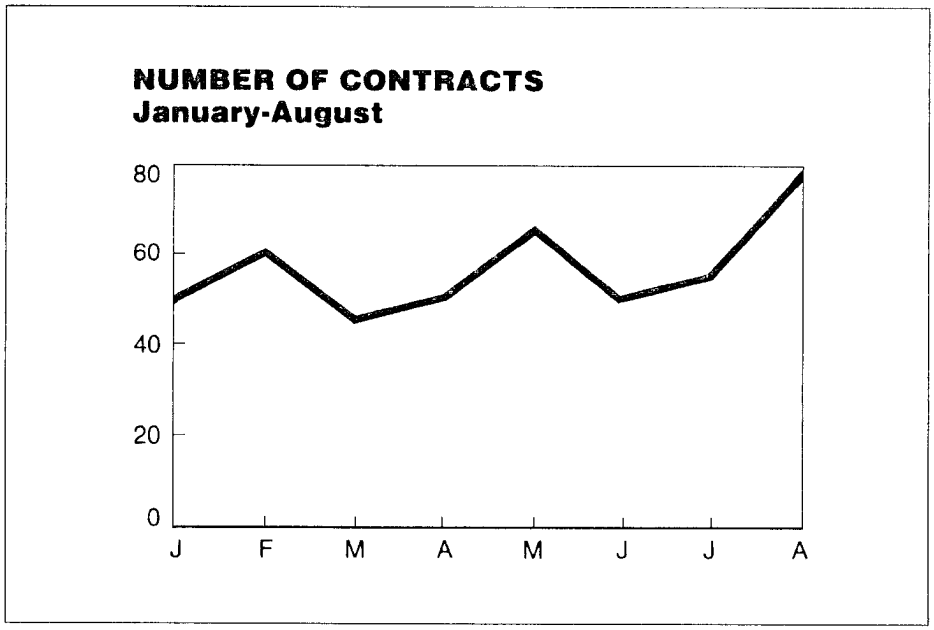
In this second example, the title merely identifies what the trend line stands for—NUMBER OF CONTRACTS—and serves to distinguish the topic of this line chart from that of other line charts we might see in a report or presentation. However, studying the trend, here are four possible aspects we may want to stress.

Message 1
The number of contracts has increased

Message 2
The number of contracts has been fluctuating

Message 3
In August, the number of contracts reached its highest point

Message 4
The number of contracts declined in two of the eight months



To assist readers, let's select **the one message** we wish to emphasize to head the chart.

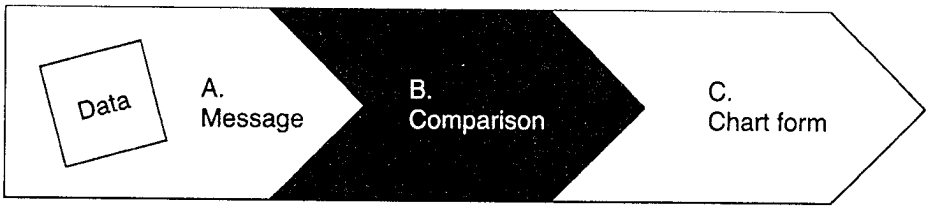
A message title is similar to a headline in your newspaper or magazine; it is brief and to the point and summarizes what you're about to read. For the cryptic headings we showed before, here's what the message titles might look like:

Topic title:	COMPANY SALES TREND
Message title:	**Company sales have doubled**
Topic title:	PRODUCTIVITY BY REGION
Message title:	**Region C ranks fourth in productivity**
Topic title:	PERCENTAGE OF ASSETS BY DIVISION
Message title:	**Division B accounts for 30% of the assets**
Topic title:	DISTRIBUTION OF EMPLOYEES BY AGE
Message title:	**Most employees are between 35 and 45 years old**
Topic title:	RELATIONSHIP OF COMPENSATION TO PROFITABILITY
Message title:	**There is no relationship between compensation and profitability**

Once you've determined *your* message, you'll find the process becomes very specific. So let's move on to the second step, identifying the kind of comparison implied in your message.

B. IDENTIFY THE COMPARISON

(from message to comparison)



If the first step is the message we start with and the third step is the chart we end with, this step is the link between the two.

What's important to recognize here is that any message—that is, any point from the data you wish to emphasize—will always lead to one of five basic kinds of comparison, which I've chosen to call *component, item, time series, frequency distribution,* and *correlation.*

Let's see examples of messages that imply each one of these comparisons. At the same time, let me define the comparisons and give you clues—trigger words—for recognizing them in messages you derive from the data.

1. COMPONENT COMPARISON

In a component comparison, we are interested primarily in showing the *size* of each part *as a percentage* of the total. For example:

- ¶ In May, Product A *accounted* for the largest *share of total* company sales.
- ¶ Client *share of market* in 2001 is less than 10 percent of the industry.
- ¶ Two sources contributed almost *half of total* corporate funds.

Any time your message contains words such as *share, percentage of total, accounted for X percent,* you can be sure you're dealing with a component comparison.

2. ITEM COMPARISON

In an item comparison, we want to compare how things *rank:* are they about *the same,* or is one *more* or *less* than the others? For example:

- ¶ In May, sales of *Product A exceeded* those of *Products B and C.*
- ¶ Client's return on sales *ranks* fourth.
- ¶ Turnover rates in the six departments are *about equal.*

Words indicating *large than, smaller than,* or *equal* are clues to an item comparison.

3. TIME SERIES COMPARISON

This comparison is the one we're most familiar with. We're not interested in the size of each part in a total or how they're ranked, but in how they *change over time,* whether the trend over weeks, months, quarters, years is *increasing, decreasing, fluctuating,* or *remaining constant.* For example:

- ¶ Sales *have risen* steadily since January.
- ¶ Return on investment *has decreased* sharply over the past five years.
- ¶ Interest rates *have fluctuated* over the past seven quarters.

Clues to look for in your message are words like *change, grow, rise, decline, increase, decrease, fluctuate.*

4. FREQUENCY DISTRIBUTION COMPARISON

This kind of comparison shows *how many items fall into a series of progressive numerical ranges.* For instance, we use a frequency distribution to show how many employees earn less than, say, $30,000, how many earn between $30,000 and $60,000, etc.; or how much of the population is under 10 years old, how many people are between 10 and 20, between 20 and 30, etc. Typical messages might be:

- ¶ In May, *most sales* were in the *$1,000* to *$2,000 range.*
- ¶ The *majority of shipments* are delivered in *five to six days.*
- ¶ The *age distribution* of company employees differs sharply from that of our competitor.

Terms to look for that suggest this kind of comparison are *x to y range, concentration,* as well as the words *frequency* and *distribution* themselves.

5. CORRELATION COMPARISON

A correlation comparison shows *whether the relationship between two variables follows*—or fails to follow—*the pattern you would normally expect.* For example, you would normally expect profits to increase as sales increase; you would normally expect sales to increase as the size of the discount offered increases.

Whenever your message includes words like *related to, increases with, decreases with, changes with, varies with,* or the converse such as *doesn't increase with,* it's an instant clue that you're showing a correlation comparison. For example:

- ¶ Sales performance in May shows *no relationship between* sales and the salesperson's experience.
- ¶ Chief executive officer compensation *does not vary* with size of company.
- ¶ Size of policy *increases with* policyholder income.

There you have them, the five kinds of comparison implied in any of the messages you'll be deriving from tabular data. Stated simply:

Component: Percentage of a total.

Item: Ranking of items.

Time Series: Changes over time.

Frequency Distribution: Items within ranges.

Correlation: Relationship between variables.

With this in mind an pencil in hand, study the following 12 typical messages derived from tabular data and identify the kind of comparison implied by each message. Look for the clues—the trigger words—in each, and, if necessary, look back to the definition and the examples we've just discussed. Check your answers with those shown upside down at the bottom of the page.

Typical Messages	Comparison?
1. Sales are forecast to increase over the next 10 years	
2. The largest number of employees earns between $30,000 and $35,000	
3. Higher price of gasoline brands does not indicate better performance	
4. In September, the turnover rates for the six divisions were about the same	
5. The sales manager spends only 15% of his time in the field	
6. Size of merit increases is not related to tenure	
7. Last year, most turnover was in the 30 to 35 age group	
8. Region C ranks last in productivity	
9. Our company's earnings per share is declining	
10. The largest share of total funds is allocated for manufacturing	
11. There is a relationship between profitability and compensation	
12. In August, two plants outproduced the other six by a wide margin	

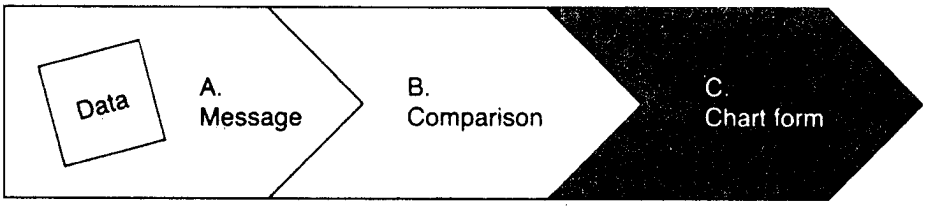
ANSWERS

1. Time Series	5. Component	9. Time Series
2. Frequency Distribution	6. Correlation	10. Component
3. Correlation	7. Frequency Distribution	11. Correlation
4. Item	8. Item	12. Item

Having moved from the data to your message and from your message to a comparison, we're now ready to proceed with the final step—from the comparison to the chart form most appropriate for your message.

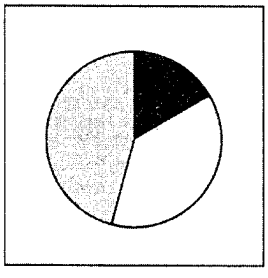
C. SELECT THE CHART FORM

(from comparison to chart)

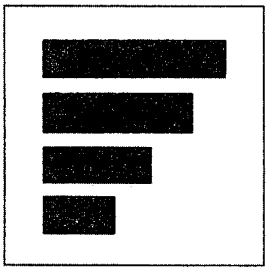


Now you have seen that, no matter what your message is, it will always imply one of the five kinds of comparison. It should come as no surprise that, no matter what the comparison is, it will always lead to one of the five basic chart forms: the pie chart, the bar chart, the column chart, the line chart, and the dot chart.

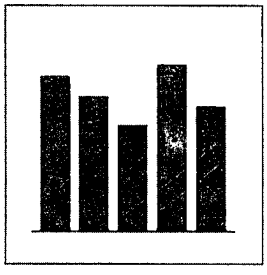
The pie chart



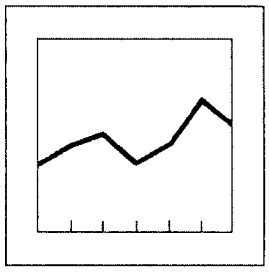
The bar chart



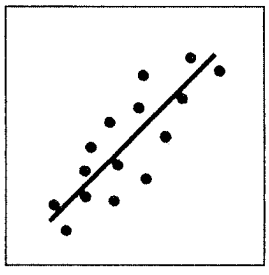
The column chart



The line chart



The dot chart



I've observed that the *pie chart* is the most popular. It shouldn't be; it's the least practical and should account for little more than 5 percent of the charts used in a presentation or report.

On the other hand, the *bar chart* is the least appreciated. It should receive much more attention; it's the most versatile and should account for as much as 25 percent of all charts used.

I consider the *column chart* to be "good old reliable" and the *line chart* to be the workhorse; these two should account for half of all charts used.

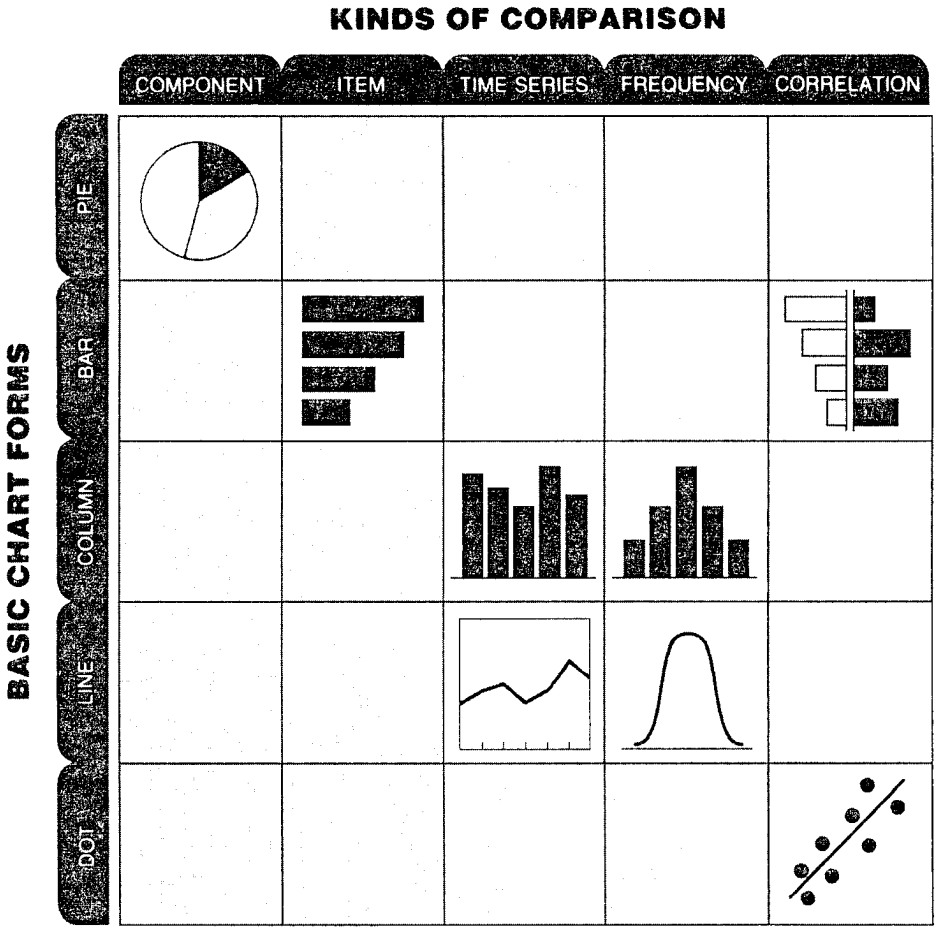
While possibly intimidating at first glance, the *dot chart* has its place 10 percent of the time.

That accounts for 90 percent. The remainder is likely to be these chart forms used in combination—say, a line chart with a column chart or a pie chart with a bar chart.

Let's recognize that each chart form, simply as a function of the way it's designed, is best equipped to illustrate one of the five comparisons.

This matrix illustrates the primary choices. Down the side are the five basic chart forms. Across the top are the five kinds of comparison we've just discussed. For time series, frequency distributions, and correlations, you have two choices of chart forms. Deciding which to use is a function of the amount of data you're plotting. For a time series or frequency distribution, use the column chart when you have few (say, six or seven) data points; use the line chart when you have many. For a correlation comparison, use the bar chart to show few data, the dot chart when you have many.

Let's work our way through the matrix and see why each chart form is recommended to show each comparison. In the process, we'll discuss how to make the most of the chart forms and present variations for each that provide additional information.



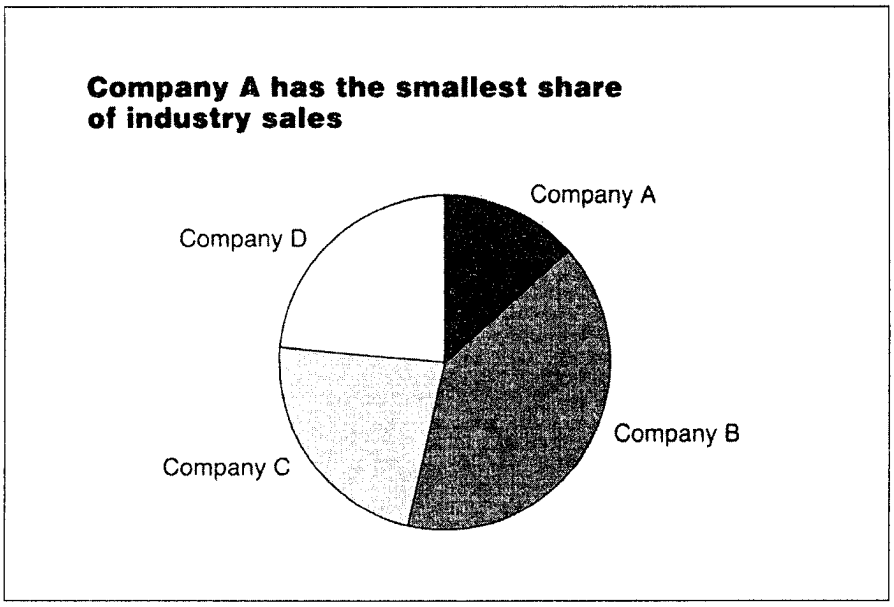
As we proceed, keep in mind that choosing, and especially using, charts is not an exact science. And so you'll note a liberal sprinkling of qualifiers, such as, generally, occasionally, most of the time, some of the time, etc., all of which imply that your judgment must play a role in deciding how best to design the charts. The options presented in the matrix, along with the suggestions for making the most of charts, are guidelines. More often than not, however, you'll find these guidelines will serve you well.

Before turning the page for a specific discussion of each comparison and its recommended chart form, I suggest that you pause for a while and skip to the second section where I present a portfolio of all these charts at work. Browse through this shopping list of charts to gain an appreciation of how effective charts can be when they are well conceived and designed.

1. COMPONENT COMPARISON

A component comparison can best be demonstrated using a pie chart. Because a circle gives such a clear impression of being a total, a pie chart is ideally suited for the one—and only—purpose it serves: showing the size of each part as a percentage of some whole, such as companies that make up an industry.

THE PIE CHART



To make the most of pie charts, you should generally use not more than six components. If you have more than six, select the five components most important to your message and group the remainder into an "others" category.

Because the eye is accustomed to measuring in a clockwise motion, position the most important segment against the 12 o'clock line and, to add emphasis, use the most contrasting color (e.g., yellow against a black

background), or the most intense shading pattern if producing the chart in black and white. If no one segment is more important than the others, consider arranging the components from the largest to the smallest or vice versa and use either the same color or no shading for all segments.

In general, pie charts are the least practical of the five chart forms. They are also the most misused and, worse, the most abused.

For example, on the next page are several pseudo pie charts I've discovered over the years in various presentations, newspapers, magazines, and annual reports. Now I'll grant you each is imaginative and resourceful, even attractive, although example D is a bit macabre. They are also examples of form becoming more important than content and, as a result, they fail to present accurate visual relationships.

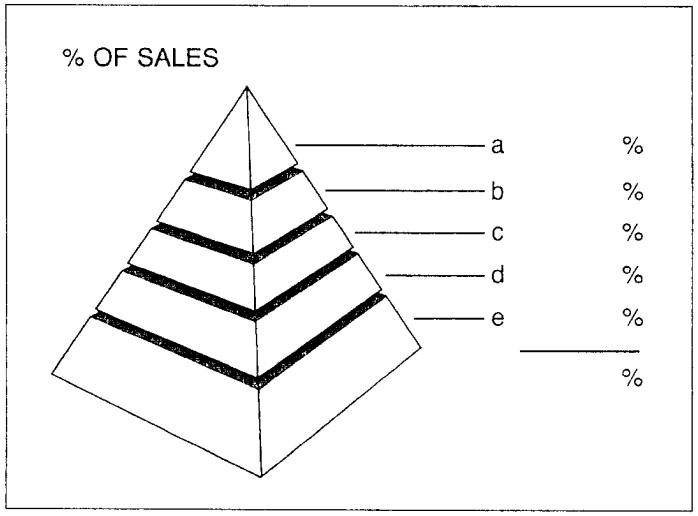
Let me stress that the primary purpose of any chart is to demonstrate relationships more quickly and more clearly than is possible using a tabular form. Whenever the form becomes more important than the content—that is, whenever the design of the chart interferes with a clear grasp of the relationship—it does a disservice to the audience or readers who may be basing decisions on the strength of what they see.

Let's have fun and do an exercise that tests the usefulness of these examples as visual *aids.* To get the most from the following work project, promise that *you will not think;* record your first visual impression. For each example, starting at the top and moving down or around, quickly fill in the percentage of the total corresponding to each component. Then add the totals. Most important, you CAN'T GO BACK, you CAN'T ERASE, and fortunately you CAN'T CHANGE YOUR MIND, since you cannot think.

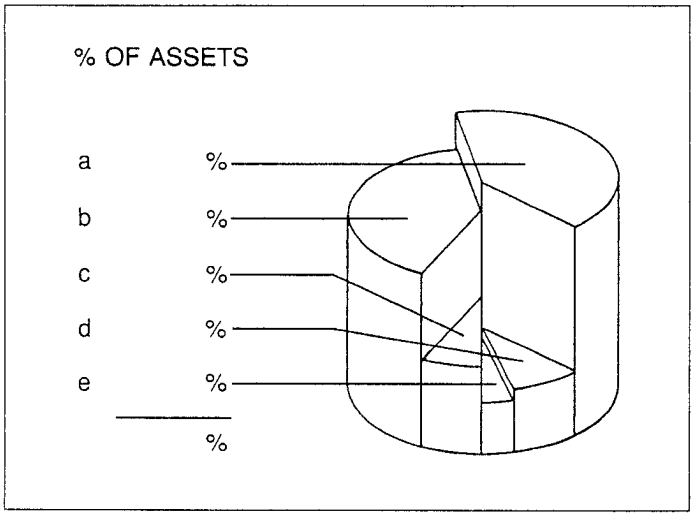
GO!

For each of these six charts, fill in the percentage you feel each segment represents and add them up.

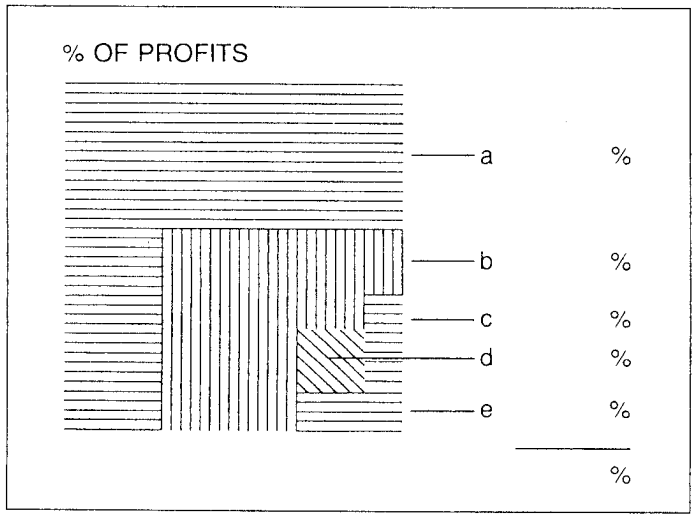
▶ A



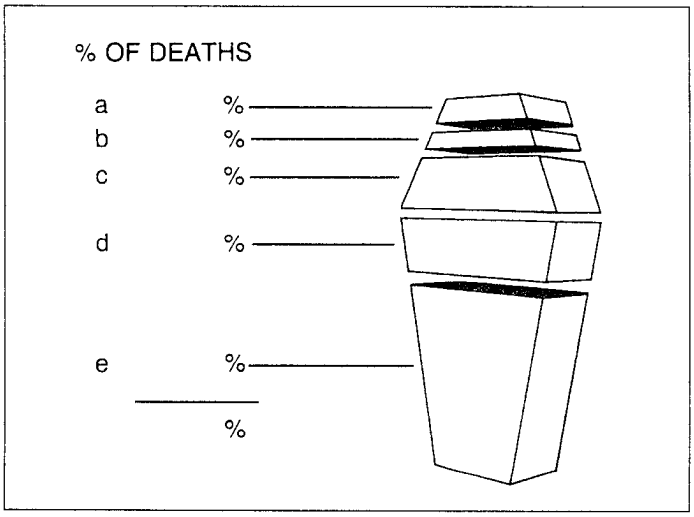
▶ B



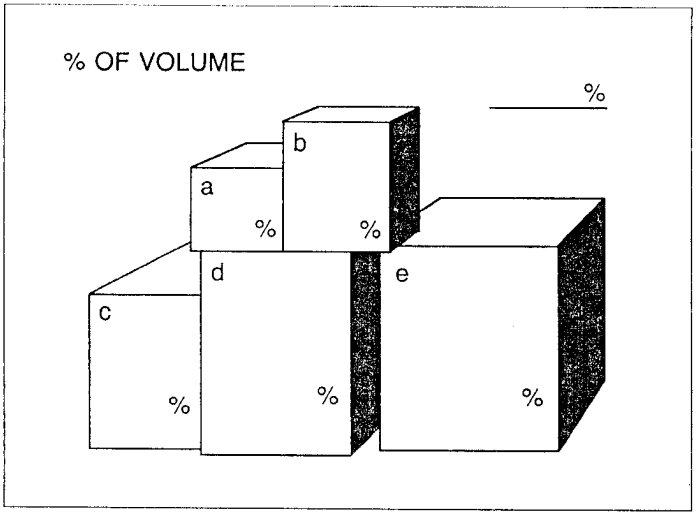
▶ C



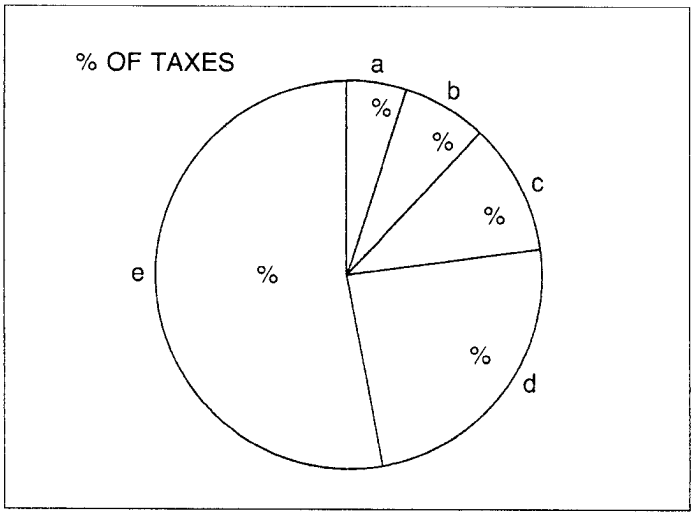
▶ D



▶ E



▶ F



Now compare all your guesses with the actual data that accompanied each example:

	A Percent of sales	**B** Percent of assets	**C** Percent of profits	**D** Percent of deaths	**E** Percent of volume	**F** Percent of taxes
a.	5%	37%	58%	7%	7%	5%
b.	7	31	32	6	15	7
c.	11	10	3	17	18	11
d.	24	14	4	16	25	24
e.	53	8	3	54	35	53
	100%	100%	100%	100%	100%	100%

If your results were radically different from these numbers—at least for Charts A through E—then it's clear that the charts are not doing the job they were intended to do, which is to give you an *accurate* impression of the relationships. I've tested these with many colleagues. Chances are your results were similar to theirs. In few cases did the data add to exactly 100 percent. Instead, the components added to less than 100 percent as often as they did to well over 100 percent. In the most extreme cases, the data added up to only 45 percent at the lower end and to 280 percent at the high end. Even when people arrived at the same total, their proportions were not necessarily similar.

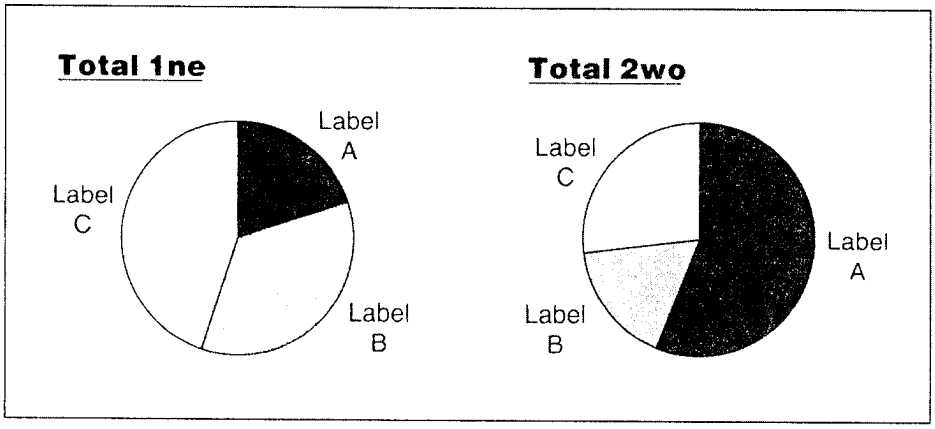
On the other hand, almost everyone was accurate in estimating the percentages in Example F, *Percent of Taxes,* presented as a conventional pie chart. Here, people could more readily *see* that segment *a* is somewhere around 5 percent and that *d* is about 25 percent, while *e* is a bit more than 50 percent. In truth, Example F is based on the same data as Example A. I just changed the titles to see what would happen. Compare the percentage values you filled in for A with those you wrote for F and note how the difference in chart form threw you off.

There's a clear lesson to be learned from this exercise: If your objective is to communicate accurate relationships, overcome the urge to be creative and instead rely on

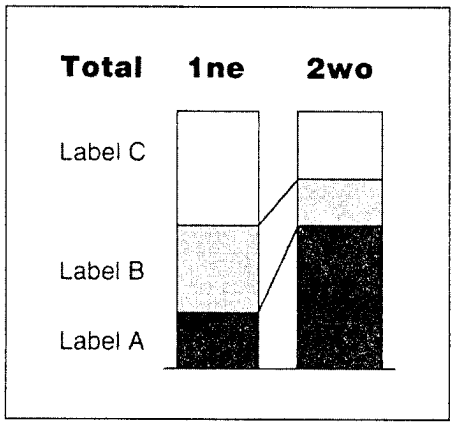
conventional pie charts. Use your creativity to make the charts attractive with handsome layouts, readable type, and constructive use of color or shadings.

A pie chart serves the purpose of showing the components of a single total better than a 100 percent bar or 100 percent column chart. However, as soon as you need to compare the components of more than one total, don't think, don't hesitate. Switch to either 100 percent bars or 100 percent columns. This example shows why.

Poor



Preferred



Notice how the labels must be repeated for each of the two pie charts shown. Of course, we could use a legend. However, this forces the reader to look back and forth between the legend and the components to be clear about which belongs to which. Also, although shadings—or colors—help the viewer to distinguish the three components, the eyes must travel back and forth, from pie to pie, to grasp the relationships.

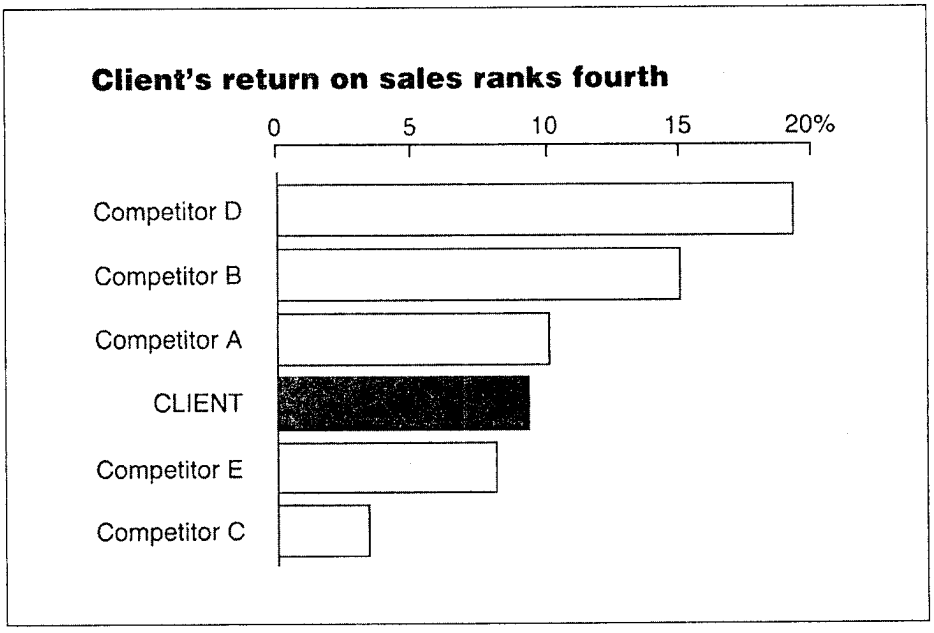
By using two 100 percent columns instead, we reduce the problems. Now the labeling is less redundant, and the relationships between corresponding segments, here reinforced with connecting lines, are more quickly apparent.

2. ITEM COMPARISON

An item comparison can best be demonstrated by a bar chart.

The vertical dimension is not a scale; all it is used for is labeling the measured items—such as countries, industries, companies, salespeople's names. This being the case, you can arrange the bars in any sequence that suits the ranking you want to stress. For example, in a chart that compares the return on sales for a client company with that of its five competitors at one point in time, the bars can be sequenced by company name in alphabetical order, or by date of entry into the industry, or by size of sales, or by the magnitude of the return from either low to high or, as in this example, from high to low (from best to worst).

THE BAR CHART

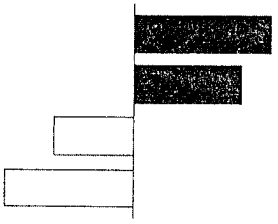


In preparing bar charts, make certain that the space separating the bars is smaller than the width of the bars. Use the most contrasting color or shading to emphasize the important item, thereby reinforcing the message title.

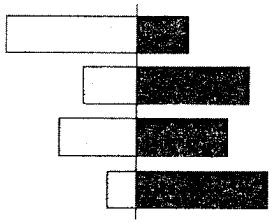
To identify the values, use either a scale at the top (sometimes at the bottom) or numbers at the ends of the bars, not both. Use the scale if all you want is a fast study of the relationships; use the numbers if they are important to your message. At times, it's a good idea to use the scale and *the one number* that needs emphasis. Using both scale and numbers, however, is redundant and adds clutter to the bar chart, as it does, for that matter, to the column chart and the line chart.

When showing numbers, round out the figures and omit decimals whenever they have little effect on your message; a figure such as 12 percent is more easily retained than 12.3 percent or 12.347 percent.

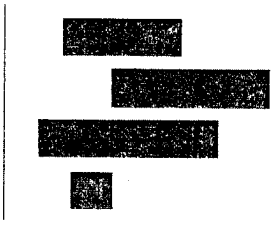
To demonstrate the versatility of the bar chart, here are six variations of the chart form, each providing additional information. Examples of the application of these variations are illustrated in the second portion of this section. You may want to glance at them now. Certainly you will want to incorporate them into your vocabulary of charts at work.



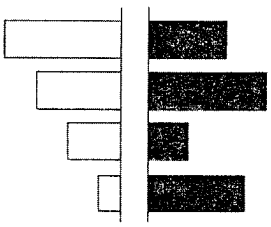
A deviation bar chart
distinguishes the profit winners from the losers



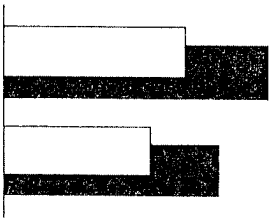
A sliding bar chart
shows the different mix of two components, e.g., percentage of imports versus percentage of exports



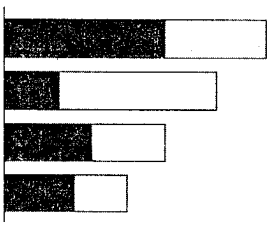
A range bar chart
shows the spread between low and high amounts, e.g., range of discounts



The paired bar chart
shows the correlation between two items, e.g., growth of market versus share of market



A grouped bar chart
compares various aspects of the same item, e.g., with and without discount

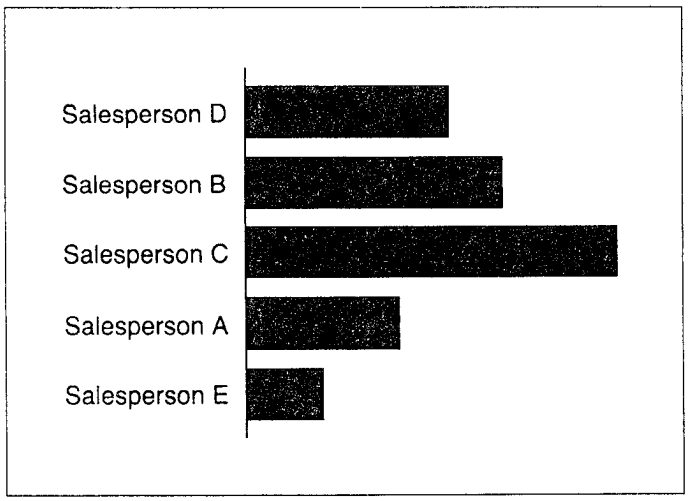


The subdivided bar chart
shows the components that make up the total

At times, you may want to use the column chart instead of the bar chart—vertical bars instead of horizontal bars—to show an item comparison. There's really nothing violently wrong with doing so. However, 9 times out of 10, you're better off with the bar chart for two reasons. First, by reserving bar charts for showing an item comparison, we reduce the possibility of confusion with a time series comparison, for which column charts are more appropriate. To reinforce this distinction, therefore, let's avoid using the bar chart for showing changes over time; in Western culture, we're accustomed to thinking of time moving from left to right, not top to bottom.

The second reason is a practical one. Generally, items have lengthy labels—territories such as Northeast, Southwest; industries such as agriculture, manufacturing; salespeople's names—all requiring space. Notice from the two examples shown that you have all the space you need to the left of the bars to label the various items, whereas, with the column chart, you may have to go through contortions since columns are usually narrow. Here you have to squeeze the label to the point of illegibility, or hyphenate the word, or write it in an awkward manner.

Preferred



Poor

3. TIME SERIES COMPARISON

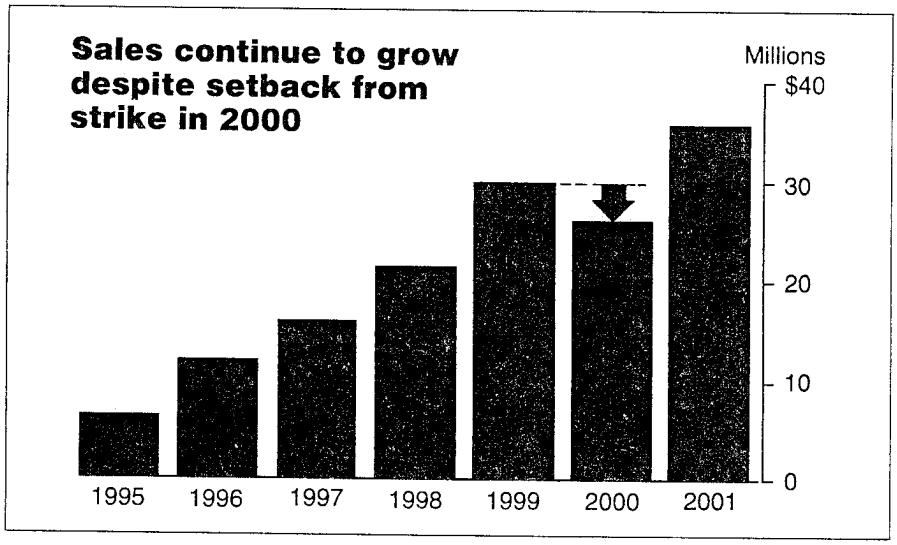
Whereas a component comparison and an item comparison show relationships at one point in time, the time series comparison shows changes over time.

A time series comparison can best be demonstrated with either a column chart or a line chart. Deciding on which to use is simple. If you have only a few points in time to plot (say, up to seven or eight) use the column chart; if, on the other hand, you have to show a trend over 20 years by quarters, you're much better off with the line chart.

In choosing between a column and a line chart, you can also be guided by the nature of the data. A column chart emphasizes levels or magnitudes and is more suitable for data on activities that occur within a set period of time, suggesting a fresh start for each period. Production data fit into this category. A line chart emphasizes movement and angles of change and is therefore the best form for showing data that have a "carry-over" from one time to the next. A good example here is inventory data.

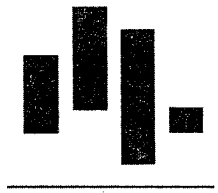
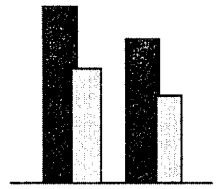
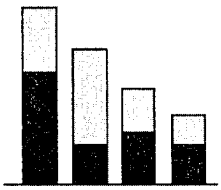
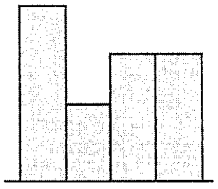
Beyond these distinctions, each chart form has its own characteristics and variations, so let's study them separately.

THE COLUMN CHART

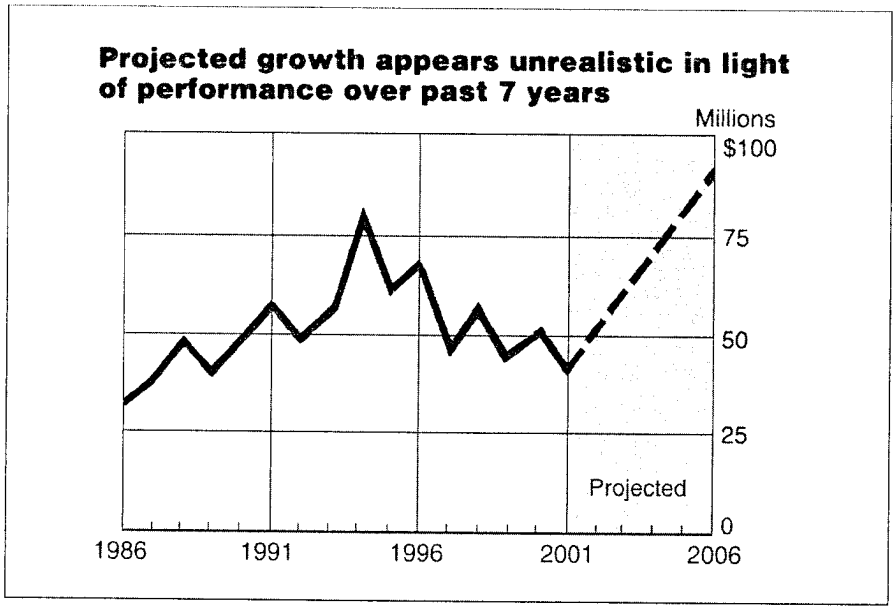


The suggestions for making the most of bar charts also apply to column charts: make the space between the columns smaller than the width of the columns; and use color or shading to emphasize one point in time more than others or to distinguish, say, historical from projected data.

As with the bar chart, there are several variations of the column chart that make it a resourceful and valuable tool; these variations are shown at work in Section 2.

	*The deviation column chart* distinguishes, say, the profit-making years from the losing ones.
	*The range column chart* shows the spread between low and high amounts, like the barometric performance of the stock market.
	The *grouped column chart*, with columns either butting against one another or overlapping, compares two items at each point in time and shows how the relationships change over time, e.g., dollars with inflation and discounted for inflation.
	A *subdivided column chart* shows how the components contributing to the total vary over time, e.g., salary plus fringe benefits adding to total compensation.
	A *step-column chart* leaves no space between the columns and is best used to show data that change abruptly at irregular intervals, such as personnel ceilings or capacity utilization.

THE LINE CHART

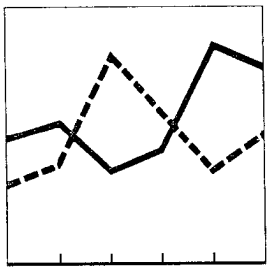


Without doubt, the line chart is the most often used of the five charts, and well it should be since it is the easiest to draw, the most compact, and the clearest for discerning whether the trend is increasing, decreasing, fluctuating, or remaining constant.

When preparing a line chart, make sure the trend line is bolder than the baseline and that the baseline, in turn, is a little bit heavier than the vertical and horizontal scale lines that shape the reference grid.

Think of grid rulings as you would the umpire at a sporting event; they're there for reference purposes, not to dominate the main attraction—in this case, the trend line(s). In other words, you may use vertical grid lines to distinguish the historic from the projected, or to emphasize quarterly periods, or to separate five-year increments. Similarly, a few horizontal scale lines will make it easier for the reader to discern relative values. In short, use your judgment to choose between too many and none at all.

The line chart has only two variations, far fewer than the bar chart or the column chart, but important enough to deserve more discussion.

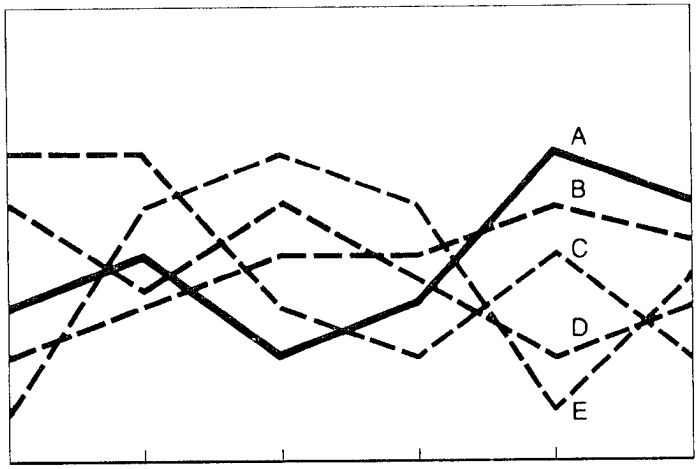


The grouped line chart compares the performance of two or more items. To distinguish, say, your company's trend from those of your competitors, use the most contrasting color or the boldest solid line for your company and less intense colors or thinner or patterned lines (long dashes, short dashes) for the others.

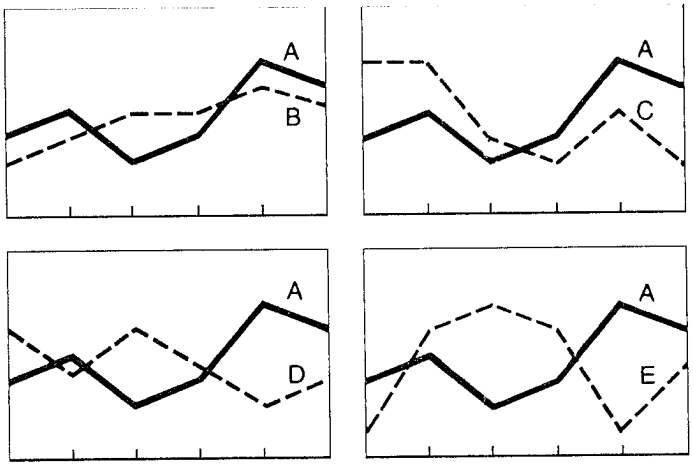
The challenge is deciding how many trend lines we can show simultaneously before the chart looks more like spaghetti than trends. Let's be realistic, a line chart with eight trend lines isn't necessarily twice as useful as one with four lines; twice as confusing, maybe, but not twice as useful.

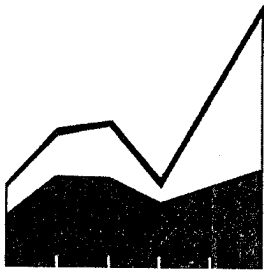
A technique for untangling the mess is to pair your trend with that of each competitor on a set of smaller charts, as you see. Granted, this creates more charts, but simpler comparisons per chart.

The Spaghetti Chart



Untangling the Mess



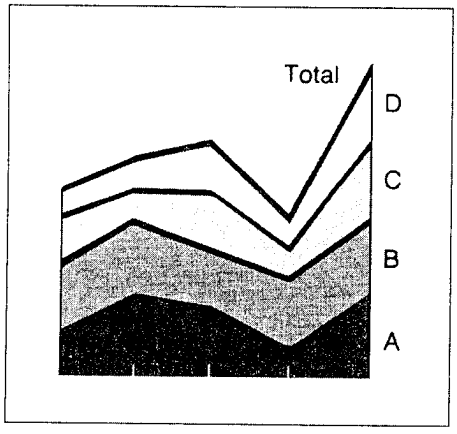


By coloring or shading the surface between the trend line and the baseline, we create the *surface chart*. By subdividing the surface into the components that make up the totals at each point in time, we create the *subdivided surface chart*. Here too, as with the subdivided bar and column charts, limit the number of layers to five or fewer. If there are more than five segments, plot only the four important ones and group the remaining into an "others" category.

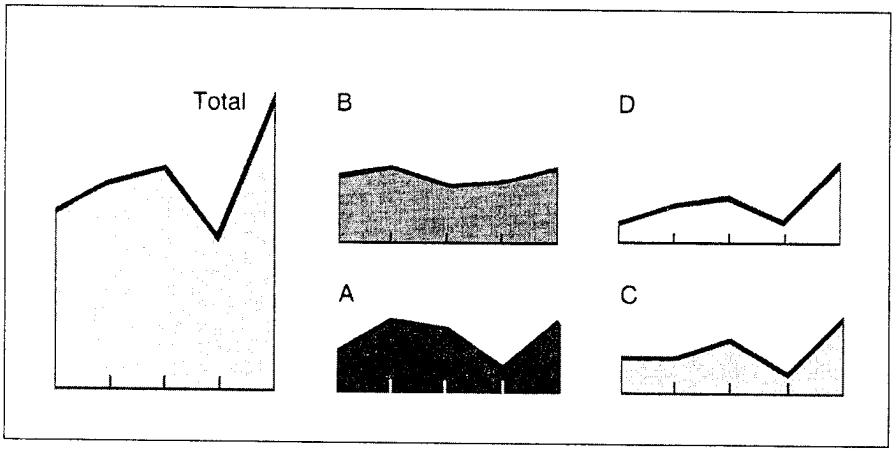
For all subdivided charts, place the most important segment against the baseline, since this is the only segment that is measured against a straight line. All other segments are at the mercy of the ups and downs of that segment.

As with the spaghetti chart we just discussed, the technique for making sense of the sea of layers is to separate the components and show each on its own base, reducing the subdivided chart to simpler surface charts.

From Subdivided Line Chart



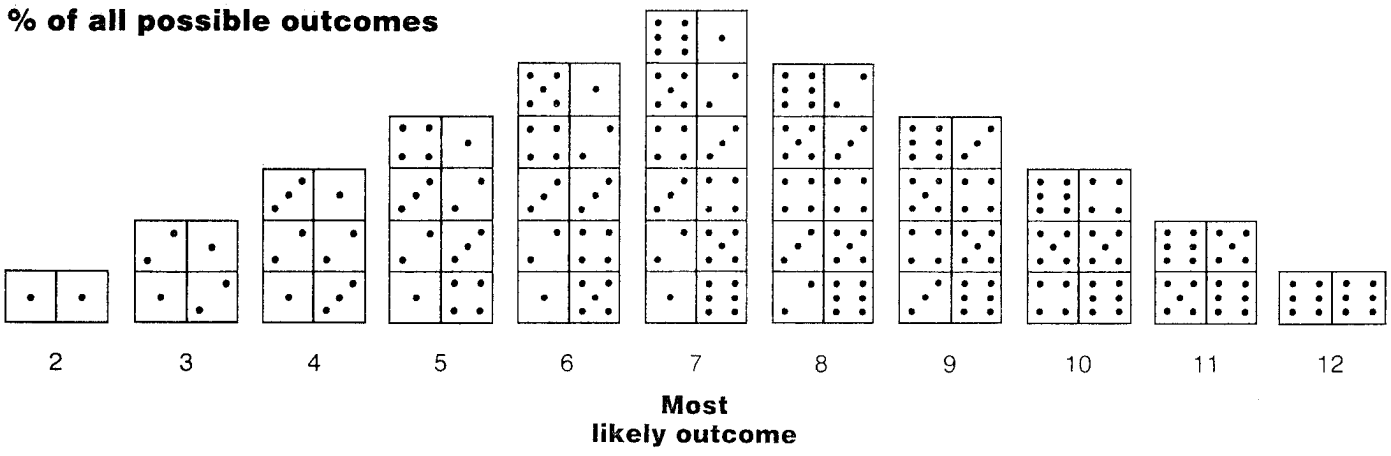
To Simpler Surface Charts



4. FREQUENCY DISTRIBUTION COMPARISON

A frequency distribution comparison shows how many items (frequency) fall into a series of progressive numerical ranges (distribution).

There are two major applications for this kind of comparison. The first is generalizing likely events on the basis of a sample of observations. Here, the frequency distribution is used to predict risk, probability, or chance. One use might be to show that there is a 25 percent chance that shipments will be delivered in five days or less; another might be to describe (un)certainty, such as the odds of rolling a losing seven, as a percentage of all possible outcomes, when shooting craps. (Save your money, the odds are one in six.)

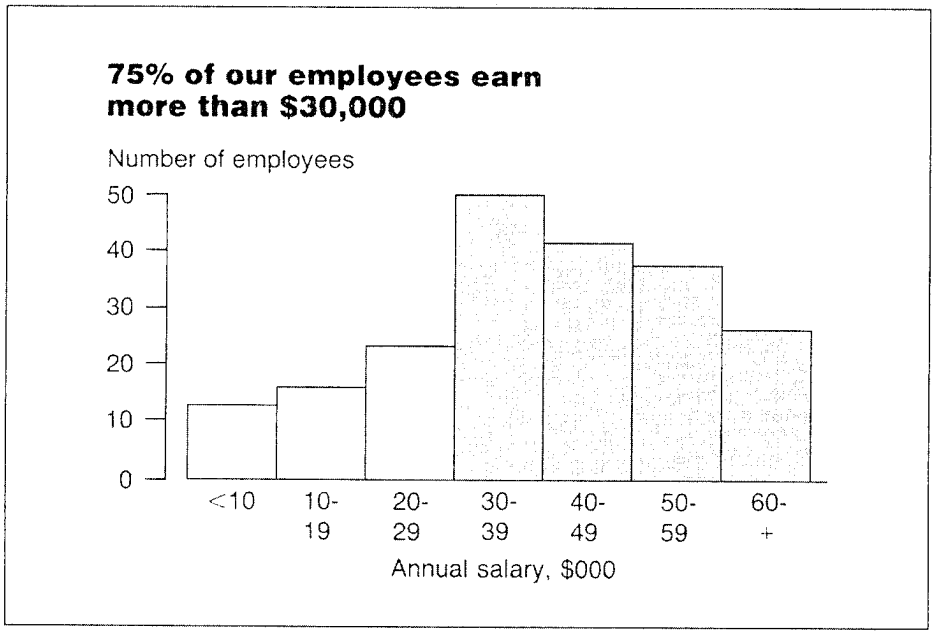


The "bell-shaped" curves and frequency polygons that are associated with this application are governed by mathematical rules. Do yourself a favor and rely on the assistance of a statistician to design them. Since these "curves" are used primarily for analytical purposes, they are not our concern in this book.

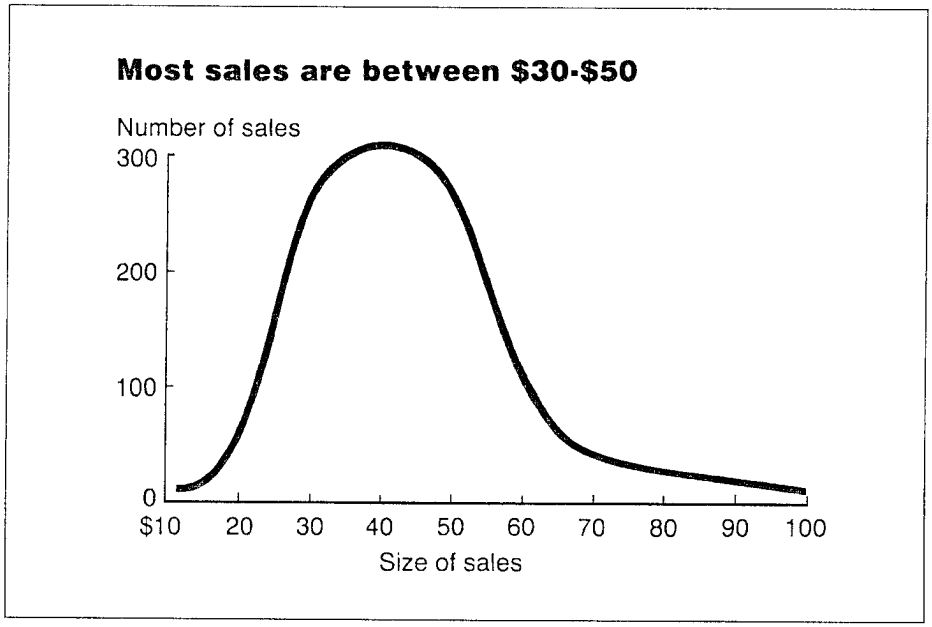
The second application, often seen in business presentations and reports, is summarizing vast amounts of data to demonstrate some meaningful relationship (e.g., 25 percent of the shipments are delivered in five to six days). This application is particularly useful for demographic information such as the number of employees by salary ranges, or the distribution of U.S. families by income levels, or the voting pattern by age group. As you would expect, this use of frequency distribution gains in popularity each time the national census is taken and every four years along with presidential elections.

In this role, the frequency distribution can best be shown by either a *step-column chart (histogram)* or a *line chart (histograph)*. Column charts are better when only a few ranges are used—say, five or seven—and line charts are better when there are many.

THE COLUMN CHART (HISTOGRAM)



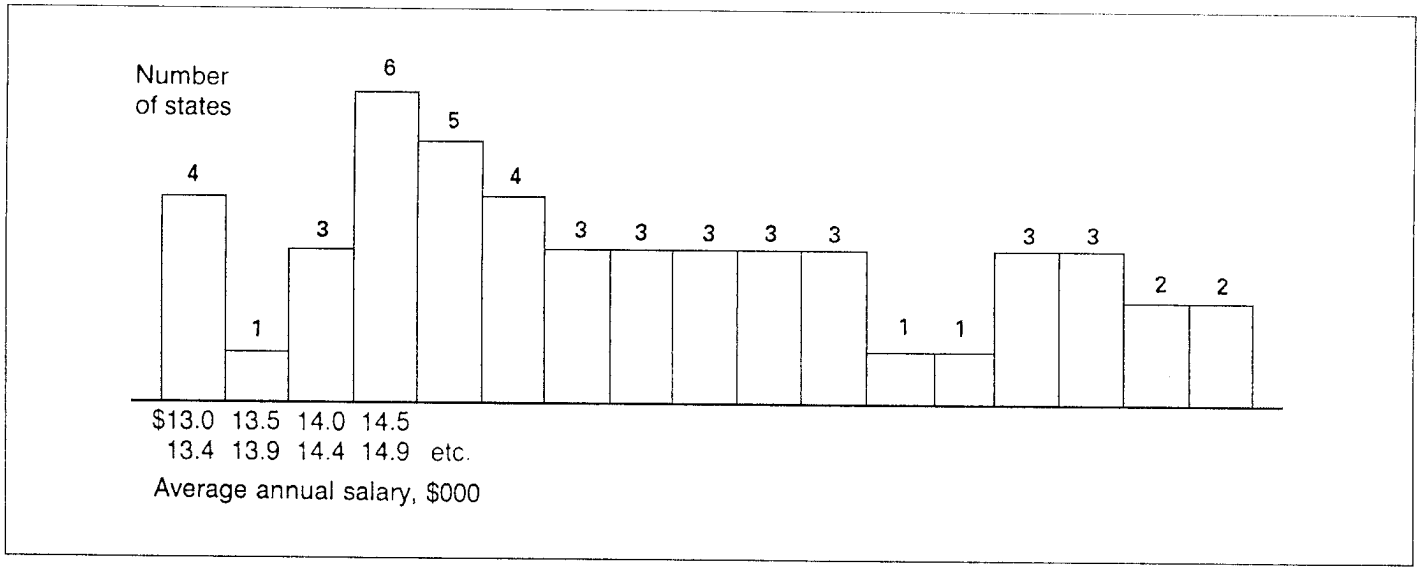
THE LINE CHART (HISTOGRAPH)



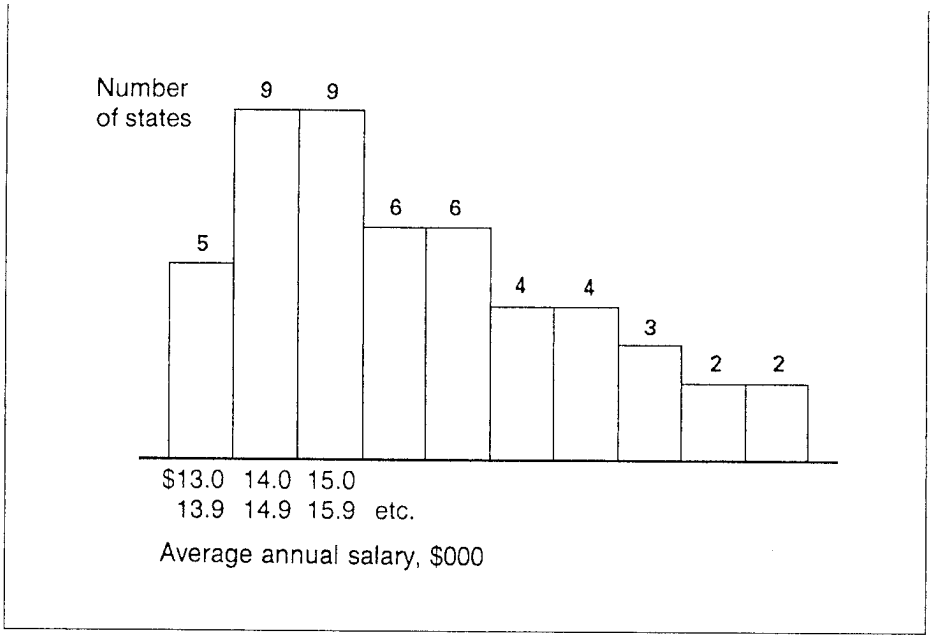
These charts have two scales: the vertical one (frequency) is for the number (sometimes percentage) of items or occurrences; the horizontal one (distribution) is for the ranges. The distribution scale requires special attention.

Size of the Ranges. The size of the ranges—and therefore the number of groups—is important in bringing out the pattern of the distribution. Too few groups hide the pattern; too many groups break it up. As a general rule, use no fewer than 5 groups and no more than 20. Within these extremes, however, you're looking for the number of groups that will demonstrate your intended message. For instance, if we wanted to bring out the pattern of a distribution of average annual salaries paid to public school teachers in the 50 states:

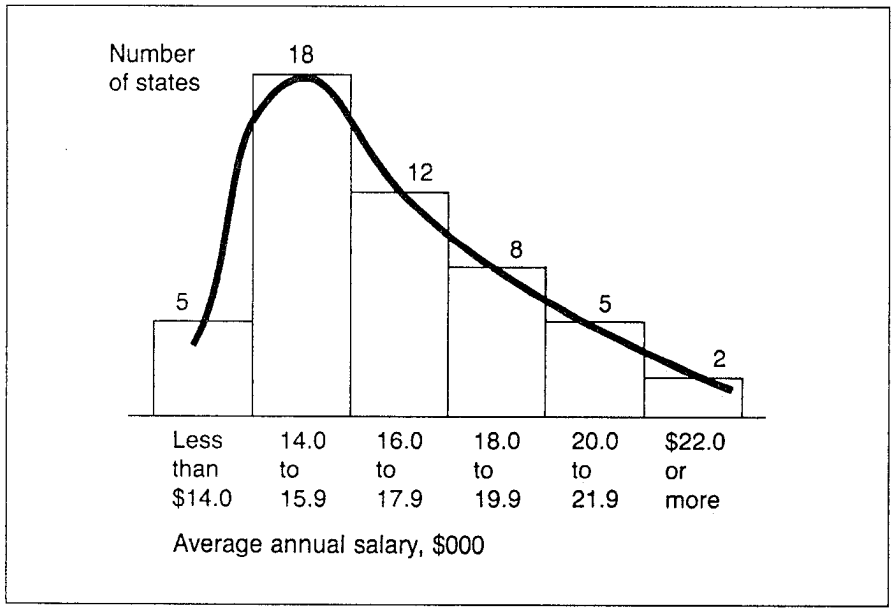
Grouping the ranges by $500 increments reveals no discernible pattern.



Grouping the ranges by $1,000 increments begins to suggest a pattern.



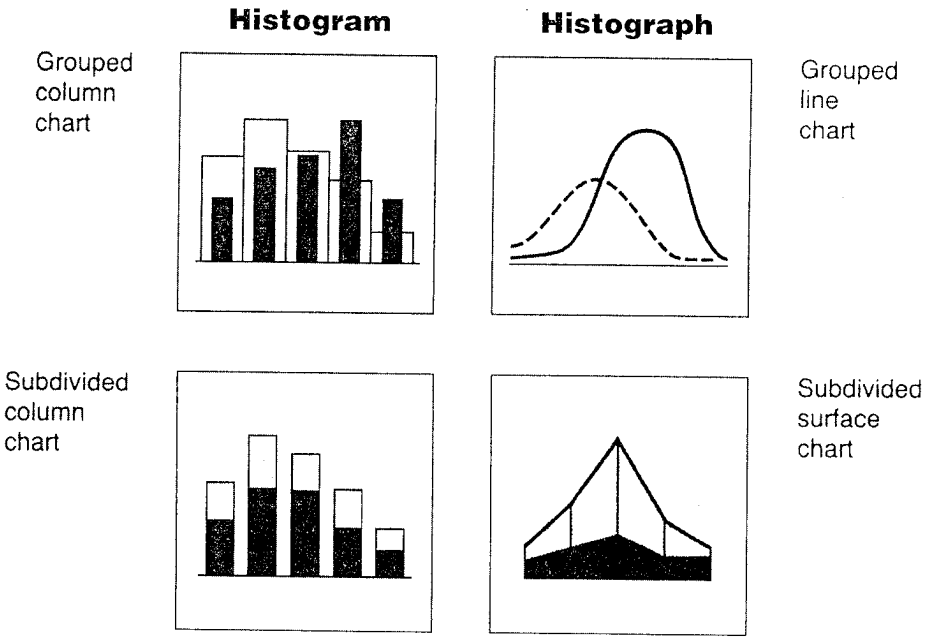
But it isn't until we group the ranges by $2,000 increments that we clearly see the bell-shaped curve usually associated with a frequency distribution. In this example, the curve is skewed to the left—that is, to the lower side of the distribution—indicating a possible message that almost half of the states (23 states out of 50) paid their teachers less than $16,000 in the year.



Size of the Groups. It is best to use groups of equal size. If one grouping represented a $5 range and the next, $20, the shape of the distribution would be distorted. Among the exceptions are cases where the data are not recorded in equal steps (e.g., educational levels) or where unequal steps make better sense, such as personal income tax brackets. Because the range of incomes is so great, and because there are so many people near the lower end and so few near the upper end, equal intervals won't work; $1,000 intervals would result in a chart several yards wide, $40,000 intervals would put virtually everyone in the first interval. The chart would be more informative if smaller intervals were shown at the lower end and larger ranges at the upper end.

Clear Labeling. The size of the groups should be explained clearly. "Overlapping" labels, such as 0–10, 10–20, 20–30, do not tell which groups include the repeated figure. For continuous data, such as dollar sales, the preferred method is less than $10.00, $10.00–19.99, $20.00–29.99, etc. For discrete data, such as number of cars manufactured, the preferred method is less than 10, 10–19, 20–29, etc.

Both the histogram and the histograph can be grouped to show, for instance, the distribution for one year against another, or to compare your employees' age distribution with that of a competitor or perhaps an industry average. Also, when absolute numbers are used, they can be subdivided to show how the components add to their total.[1]

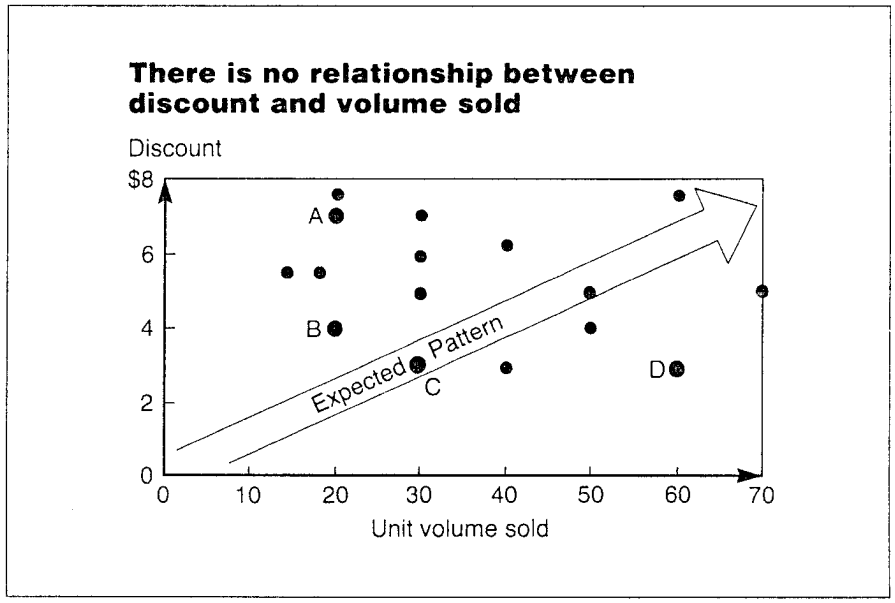


[1] *Caution:* It is misleading to subdivide a frequency distribution when the frequency is expressed in percentage terms. For example, if 60 percent of women earn between, say, $5 and $10 an hour, and 50 percent of men earn the same, it cannot be said that combined, 110 percent of the people earn between $5 and $10 an hour.

5. CORRELATION COMPARISON

A correlation comparison shows whether the relationship between two variables follows—or fails to follow—the pattern you would normally expect. For example, you would normally expect that a salesperson with more experience would generate more sales than one with less experience; you would normally expect that employees with more education would receive higher starting salaries. Such comparisons are best shown by a dot chart, sometimes called a scatter diagram, or by a paired bar chart. Let's look at each in turn.

THE DOT CHART



Here we show 16 transactions in terms of the size of the discount offered and the number of units sold. Normally, you would expect that the larger the discount, the greater the incentive to purchase more units. As the dot chart reveals in this case, there is no such correlation.

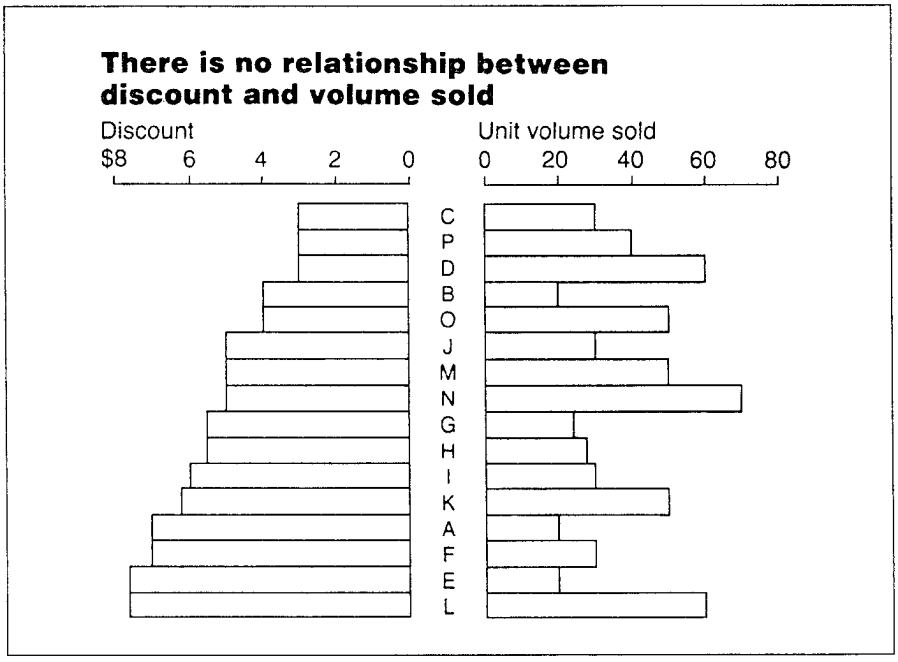
For instance, study the dots labeled *A* and *B*, representing the transactions of two salespeople. Both sold 20 units (horizontal scale). However, *A* offered a $7 discount, while *B* gave only $4 (vertical scale). On the other hand, salespeople *C* and *D* offered the same $3 discount, and yet *C* sold 30 units while *D* sold twice as many. Obviously, the size of the discount offered has little or no effect on the volume bought.

Had there been a correlation, then the dots would have clustered around a diagonal line moving from the lower left of the chart to the upper right, represented here by a faint background arrow. It's often a good idea to include this arrow to reinforce the expected pattern. Of course, at times, the arrow might point down to show, for example, that volume increases as price decreases. Also, let's not confuse this arrow with the mathematically computed "line of best fit" which is a curve fitted through the dots that emphasizes the pattern of the plotted values.

These dot charts are being used increasingly in presentations, reports, and some business magazines. If you plan to use them, be patient with your audience or reader and explain how to read the chart before revealing the message.

Aside from appearing confusing, the problem with these charts is identifying the dots. Including each salesperson's name next to his or her dot not only adds to the confusion but can create a severe case of myopia. An option is to use a legend with each dot identified by a letter or number corresponding to the full name shown somewhere else on the chart. A better option is to use the paired bar chart.

THE BAR CHART



You'll notice that now there is space to label each set of plotted values between the sets of bars. In a paired bar chart, we usually rank the independent variable on the left, in either a low-to-high or high-to-low sequence. When the relationship between the expected pattern and the actual one is consistent, the dependent variable bars on the right will form a mirror image of those on the left. In other words, low discounts will mirror low volume and high discounts will be paired with high volume. When the relationship is not as expected, the two sets of bars will deviate from one another, as they do in this example.

This paired bar chart option will work only when there are relatively few sets of data to plot. Beyond 15 or so, you're better off with the more compact scatter diagram and forgetting about trying to label each dot.