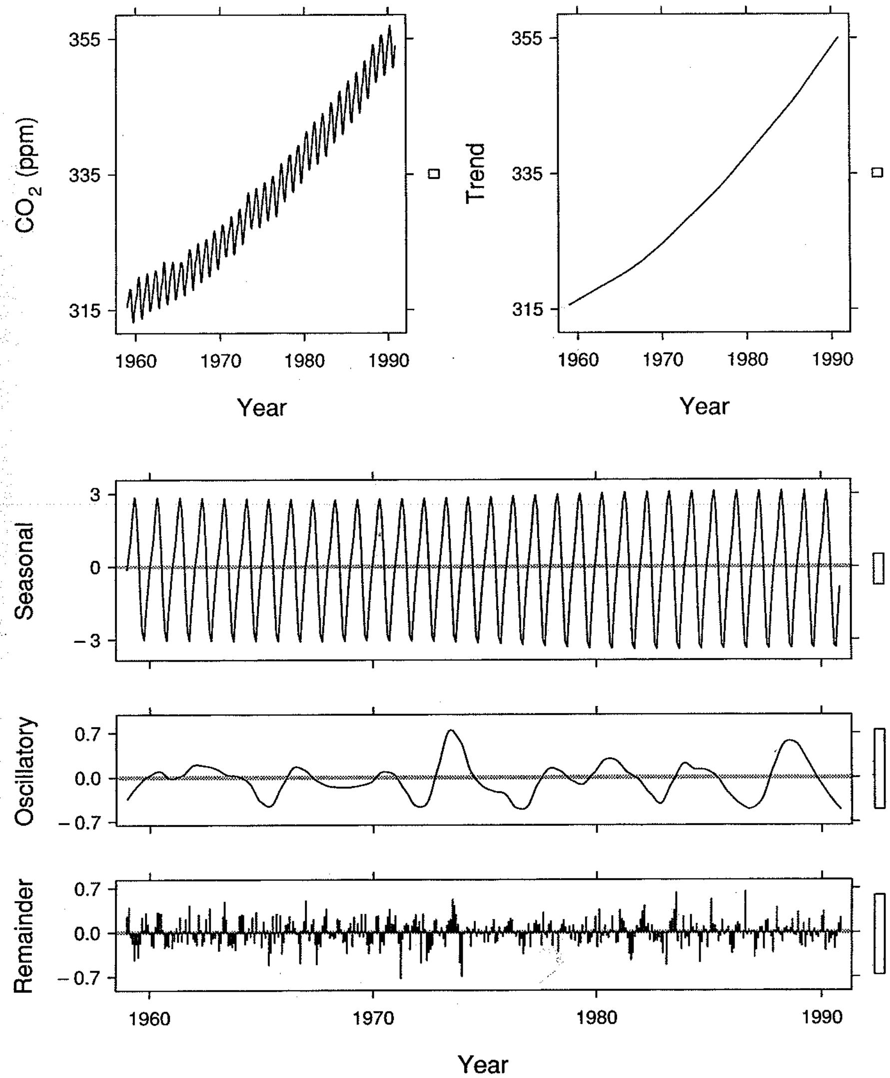


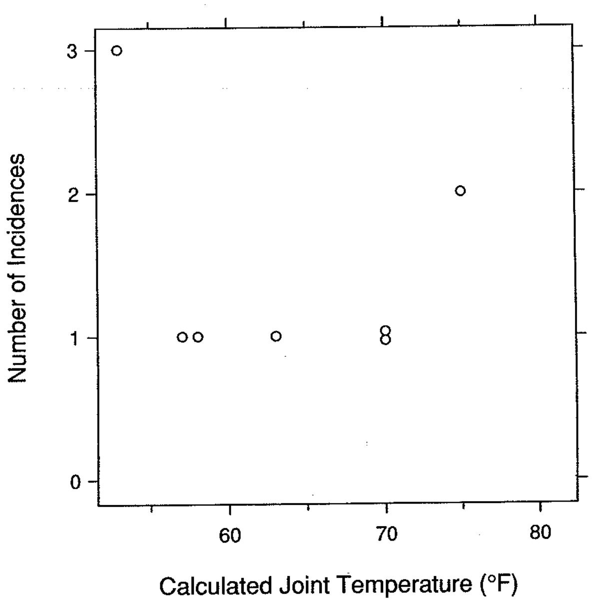
1.2 THE POWER OF GRAPHICAL DATA DISPLAY. Visualization provides insight that cannot be appreciated by any other approach to learning from data. On this graph, the top left panel displays monthly average CO_2 concentrations from Mauna Loa, Hawaii. The remaining panels show frequency components of variation in the data. The heights of the five bars on the right sides of the panels portray the same changes in ppm on the five vertical scales.

ıe
that
s a
6.
rigin;
wn
that
ıy
re
ıis
oints
ukey

ı DISPLAY. This
rations of aerosols.
ms. There is a
tal scale is wrong
scales should be
ı judge the

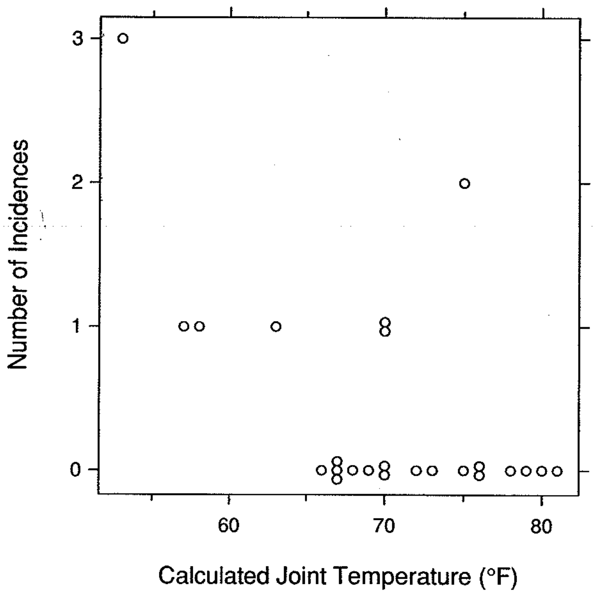
shuttle
een
ıy was
the
ors.

To assess the issue, the engineers studied a graph of the data shown in Figure 1.4. Each data point was from a shuttle flight in which the O-rings had experienced thermal distress. The horizontal scale is O-ring temperature, and the vertical scale is the number of O-rings experiencing distress. The graph revealed no effect of temperature on the number of stress problems, and Morton Thiokel, the rocket manufacturer, communicated to NASA the conclusion that the "temperature data [are] not conclusive on predicting primary O-ring blowby" [43]. The next day Challenger took off, the O-rings failed, and the shuttle exploded, killing the seven people on board.



1.4 STATISTICAL REASONING. These data were graphed by space shuttle engineers the evening before the Challenger accident to determine the dependence of O-ring failure on temperature. Data for no failures was not graphed in the mistaken belief that it was irrelevant to the issue of dependence. The engineers concluded from the graph that there is no dependence.

The conclusion of the January 27 analysis was incorrect, in part, because the analysis of the data by the graph in Figure 1.4 was faulty. It omitted data for flights in which no O-rings experienced thermal distress. Figure 1.5 shows a graph with all data included. Now a pattern emerges. The Rogers Commission, a group that intensively studied the Challenger mission afterward, concluded that the engineers had omitted the no-stress data in the mistaken belief that they would contribute no information to the thermal-stress question [43].



1.5 STATISTICAL REASONING. The complete set of O-ring data is now graphed, including the observations with no failures. A dependence of failure on temperature is revealed.

The graphical analysis of the O-ring data failed, not because of the display method used, as with the aerosol data, but rather because of a poor choice of the statistical information selected for the graph. This arose because of a flaw in the statistical reasoning that underlay the graph. The flaw violated a basic statistical principle: in the analysis of failure data, the values of a causal variable when no failures occur are as relevant to the analysis as the values when failures occur. Statistical thinking is vital to data display. A number of statistical principles are discussed in Chapters 2 and 3.

Brain Masses and Body Masses of Animal Species

Figure 1.6 is a graph from Carl Sagan's intriguing book, *The Dragons of Eden* [107]. The graph shows the brain masses and body masses, both on a log scale, of a collection of animal species. We can see that log brain mass and log body mass are correlated, but this was not the main reason for making the graph.

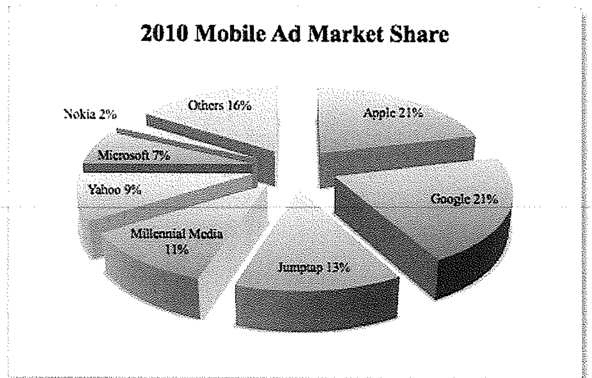
Apple 2.0

Mac news from outside the reality distortion field

Apple is grabbing mobile ad share from Google, Yahoo, Microsoft and Nokia

Posted by Philip Elmer-DeWitt
September 27, 2010 5:45 AM

Smaller rivals like Jumptap and Millennial Media are also gaining, according to IDC



U.S. only. Source: IDC via Bloomberg Businessweek

The pie chart at right is somewhat premature, given that it is IDC's best guess -- via *Bloomberg Businessweek* -- of what the $500 million U.S. mobile advertising market will look three months from now.

But it's an indication of how the winds have shifted. Apple (AAPL), which had 0% share of the market before it bought Quattro Wireless in January and launched iAd in June, will end the year at 21%, according to IDC.

Where did that share come from? Chiefly from Google (GOOG), Yahoo (YHOO), Microsoft (MSFT) and Nokia (NOK), according to the *Businessweek* piece.

IDC's before-and-after numbers for the biggest losers below the fold.

	Dec. 2009	Dec. 2010
Google	27%	21%
Yahoo	12%	9%
Microsoft	10%	7%
Nokia	5%	2%

. . .

[Follow Philip Elmer-DeWitt on Twitter @philiped]

Tags: Apple, Advertising, IDC, Microsoft, Nokia, Yahoo, Mobile, iAd, Google, mobile advertising

0 Comments | Add a Comment

Recommend 7

« Is AAPL over-valued? Not compared with the other ultra-large market caps
Today in Tech: News around the Web »

About This Author

Philip Elmer-DeWitt

Steve Jobs, goes the old joke at Apple, is surrounded by a reality distortion field; get too close and you might believe what he's saying. Apple has made believers out of millions of customers -- and made a lot of investors rich -- but Elmer-DeWitt believes that an ounce of skepticism never hurts when writing about the company. He should know. He's been covering Apple - and watching Steve Jobs operate -- since 1982.

Email This Author
Subscribe to Apple 2.0: RSS feed | email newsletter

BRAINSTORM TECH

- White House defends pace of health-care IT rollout
- Steve Jobs' mark on ABC's Lost
- Is Motorola back?
- The real reason why consumers won't pay online
- EPA to tech companies: build products that break down
- Liberty Global's global retreat

See All Posts

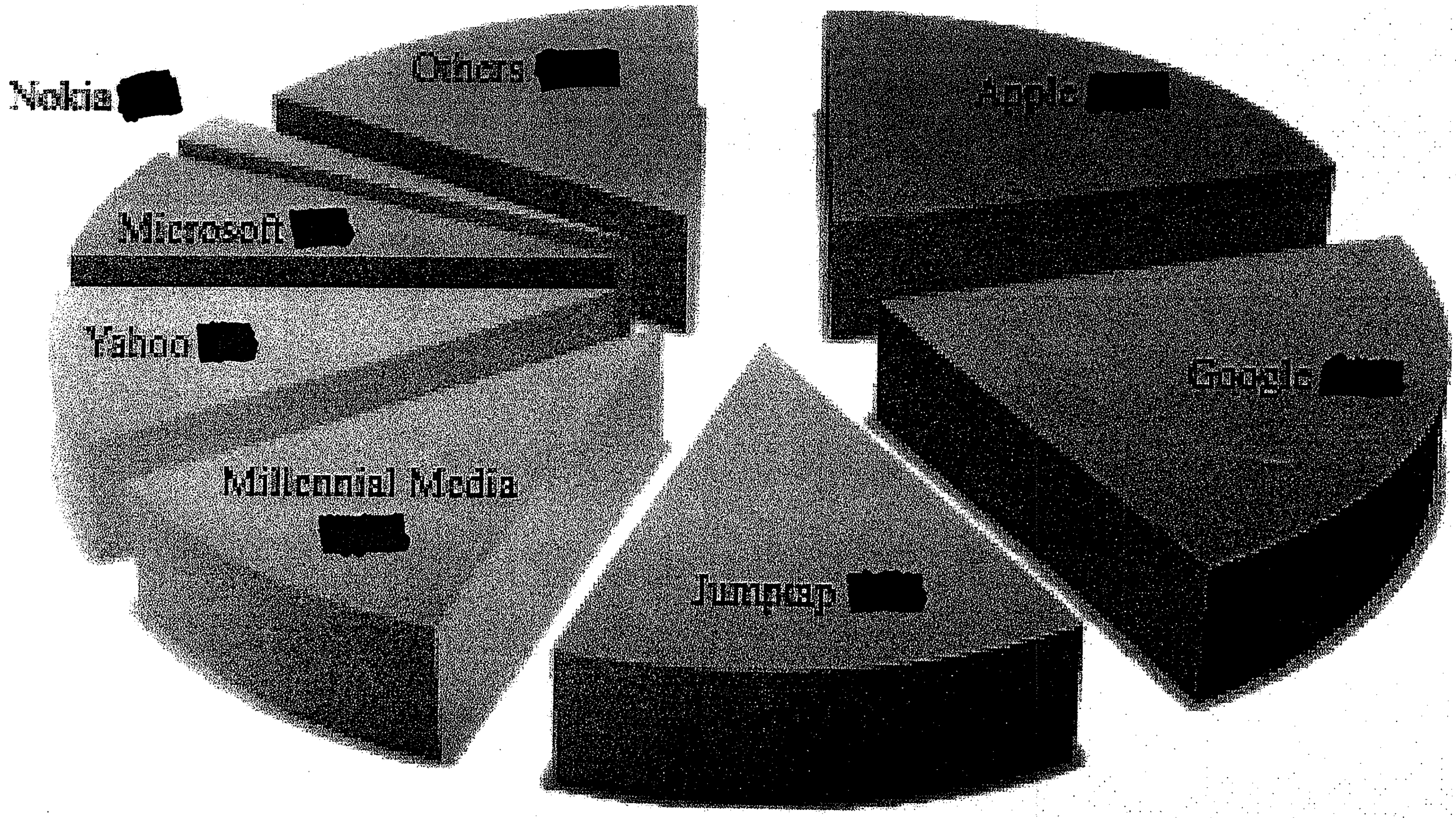
Recent Posts

- Ning CEO: Building a better web site
- Can anybody beat the iPad?
- YouTube celebrates debate on anniversary of Nixon v. Kennedy
- Samsung shows home automation using Android Galaxy Tab
- Today in Tech: News around the Web

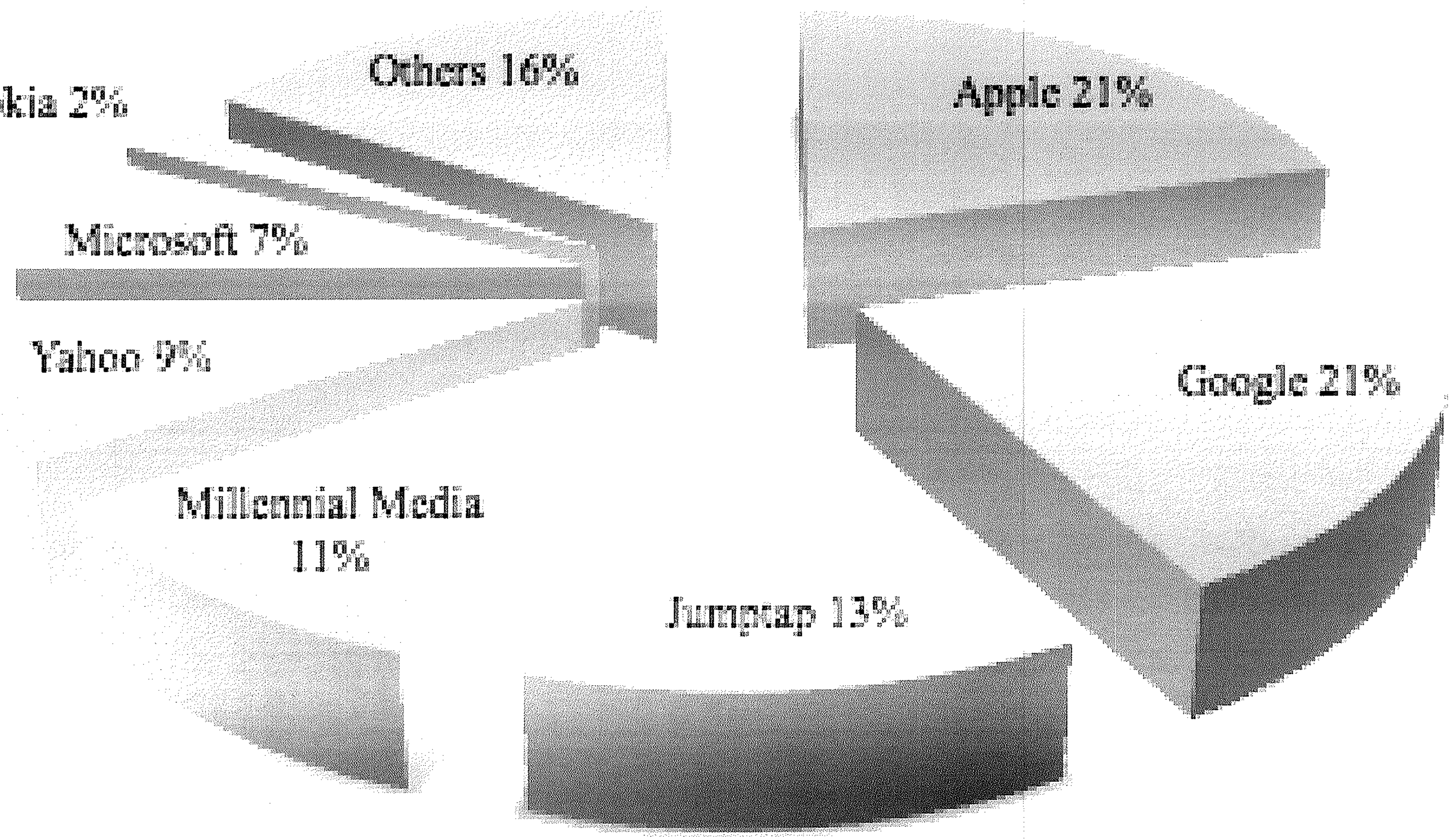
Search

Contributors

2010 Mobile Ad Market Share

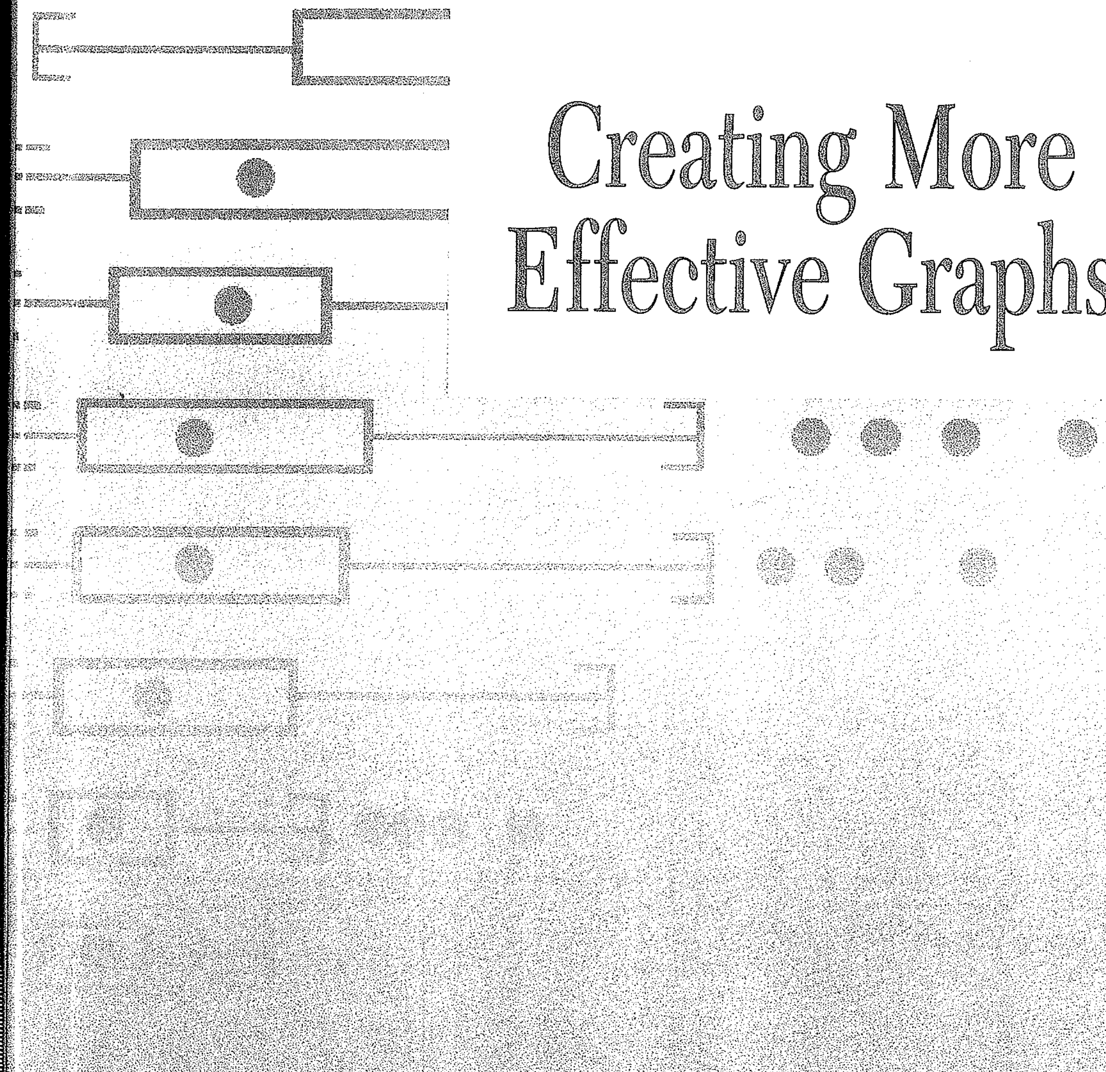


2010 Mobile Ad Market Share



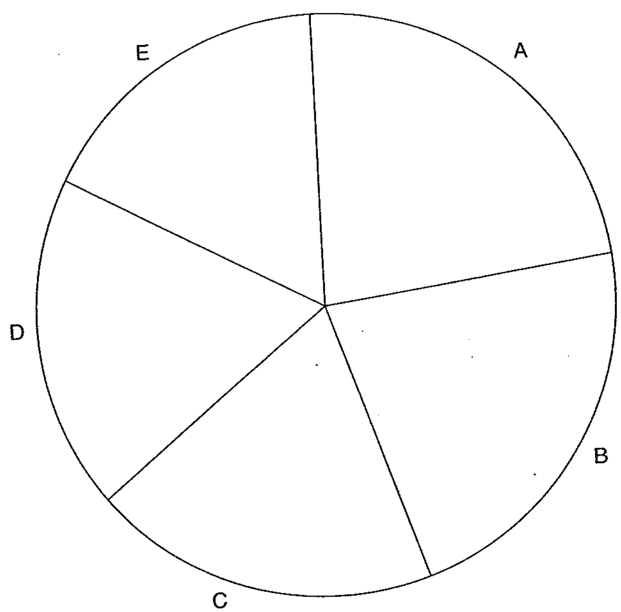
WILEY

Creating More Effective Graphs



NAOMI B. ROBBINS

Fig. 1.1 Similar Pie Wedges

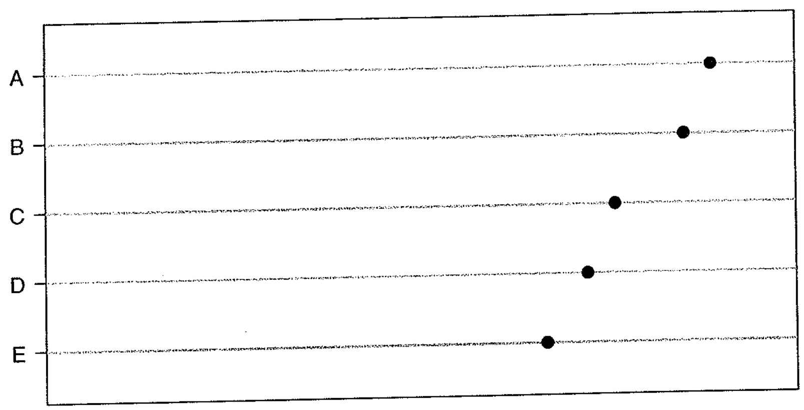


Whether you draw graphs manually or with a computer, and regardless of the software package you use, increasing your knowledge of the principles of effective graphs will greatly improve your work.

1.1 WHAT WE MEAN BY AN EFFECTIVE GRAPH

We begin by discussing what we mean by an *effective graph*. Figure 1.1 is the familiar *pie chart*. This one has five wedges labeled A through E. Study this chart and try to place the wedges in size order from largest to smallest. Use a pencil and paper to write down your results.

Fig. 1.2 Similar Pie Wedges: Dot Plot



The same data are plotted in Figure 1.2, this time using a *dot plot*. Once again take out paper and pencil and order the size of categories A through E.[1] Most of you probably had a lot of trouble placing the wedges of the pie chart in size order. Some of you may have found this to be quite frustrating, even though a few of you probably had no trouble at all. But even those of you who could order the wedges of the pie chart easily must admit that this task is much easier using a dot plot. Cleveland (1984) introduced dot plots to take advantage of the results of experiments on human perception and the decoding of graphical information. In Chapter 3 I discuss this topic briefly.

[1] I have left out the tick marks and labels since a reviewer suggested that it was an unfair comparison to show tick labels on the dot plot and no labels on the pie chart. Tick mark labels help you estimate the values, which is not the task that you were asked to do.

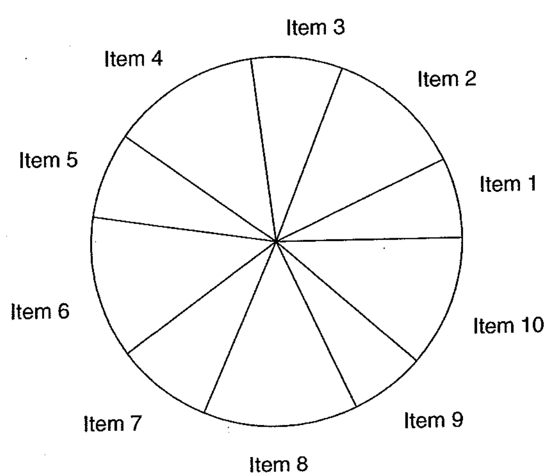
Fig. 1.3 Similar Pie Wedges: Table

A	23.0
B	22.0
C	19.5
D	18.5
E	17.0
Total	100.0

Sometimes graphs do not provide the best solution for presenting data. A table provides another way to show the data in Figures 1.1 and 1.2. The table in Figure 1.3 not only shows the exact values of the categories, but also shows the total. Tables are preferable to graphs for many small data sets.

2 Limitations of Some Common Charts and Graphs

Fig. 2.1 Structured Data Set



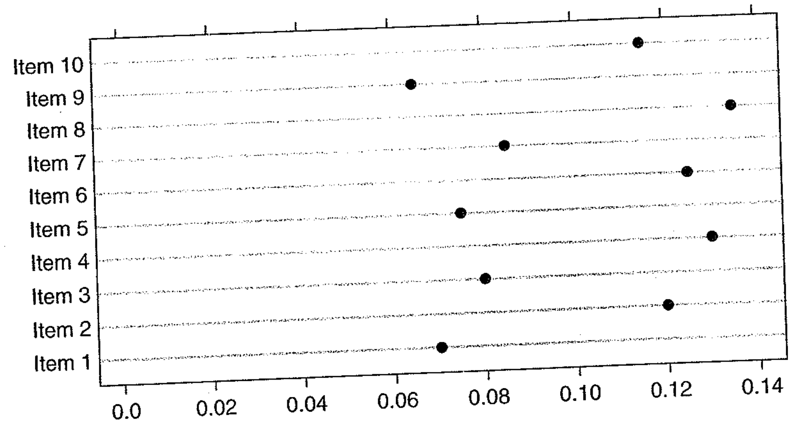
Creating More Effective Graphs, by Naomi B. Robbins
ISBN 0-471-27402-X

2.1 PIE CHARTS

The pie chart in Figure 2.1 shows 10 wedges, which I have labeled items 1 through 10. Many authors suggest ordering pie charts from largest to smallest unless there is a natural ordering[1] of the data, and our labeling here gives a natural ordering. I would like you to study the data and learn from them what you can.

[1] Jan., Feb., and Mar. are ordered categories; so are low, medium, and high and first, second, and third. Apples, oranges, and bananas are categorical data that do not have a natural ordering.

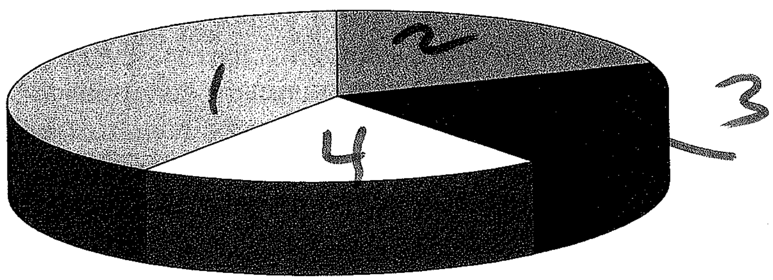
Fig. 2.2 **Structured Data Set: Dot Plot**



Some of you probably noticed in Figure 2.1 that the areas of the wedges labeled with odd numbers were smaller than those with even numbers. Once again we show the same data in a dot plot. Study Figure 2.2 and see what you can learn about the data.

This is a very structured set of data. The values of the items with even-numbered labels are an exact offset of those with odd-numbered labels. Those with odd-numbered labels are centered on 0.075, and each item with an even-numbered label is exactly 0.05 bigger than the preceding one with an odd number. I challenge you to see this in Figure 2.1, but it is clear in Figure 2.2.

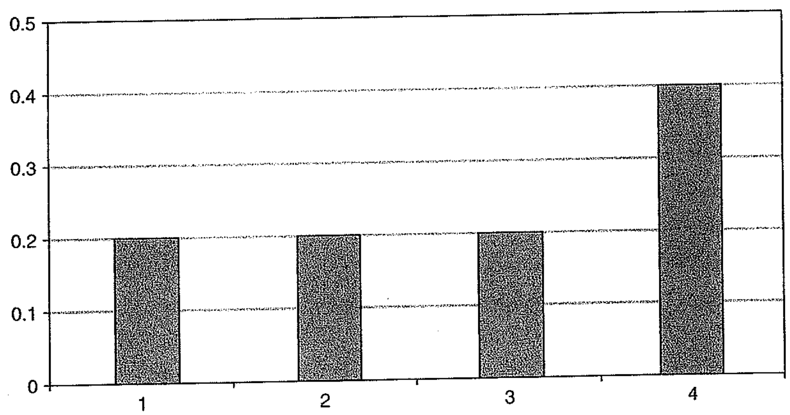
Fig. 2.3 Three-Dimensional Pie Data



2.2 CHARTS WITH A THREE-DIMENSIONAL EFFECT

Tufte said that the only design worse than a pie chart was several of them, but I consider the design shown in Figure 2.3 to be worse than a simple pie chart: It is a *three-dimensional pie chart*. You know that the four wedges add up to 100%. Take your pencil and paper and write down four numbers adding to 100 that represent the percentage of each wedge.

Fig. 2.4 Three-Dimensional Pie Data: Two-Dimensional Bar Chart



In Figure 2.4 I show the same data on a simple *two-dimensional bar chart*. Now you see clearly the percentage of each category. Are the results the same as those you read from Figure 2.3?

Fig. 2.5 Three-Dimensional Pie Data: Two-Dimensional Pie Chart

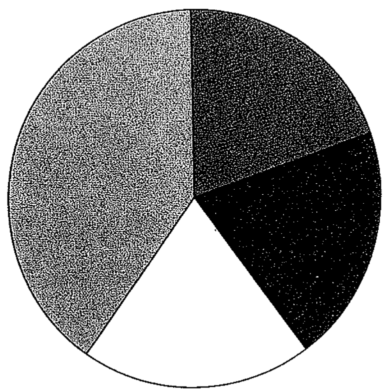
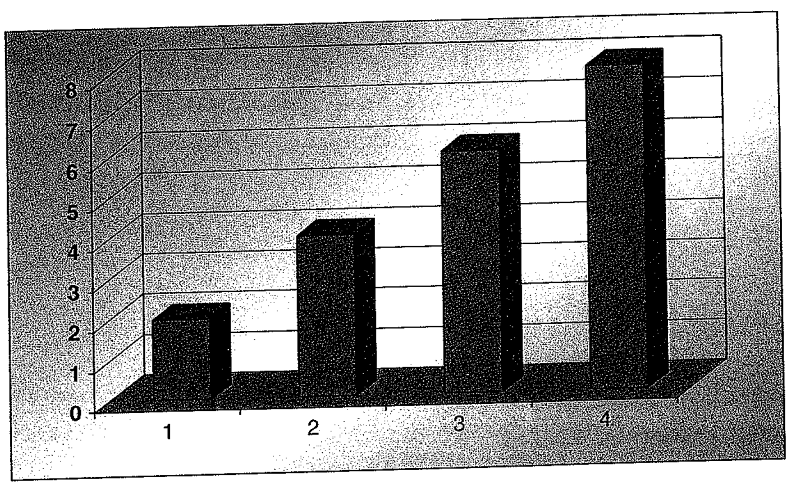


Fig. 2.6 Three-Dimensional Bar Data



We have already seen the limitations of pie charts. But the *two-dimensional pie chart* in Figure 2.5 is certainly an improvement over the three-dimensional version.

Three-dimensional bar charts[2] such as that in Figure 2.6 are ubiquitous. Take your pencil and paper and write down the heights of the bars.

[2] A real three-dimensional bar chart displays three variables; charts with two variables, as in Figure 2.6, are frequently called *pseudo-three-dimensional bar charts*. Common terminology is used here since bar charts with three variables are not discussed.

Fig. 2.7 Three-Dimensional Bar Data: Two-Dimensional Bar Chart

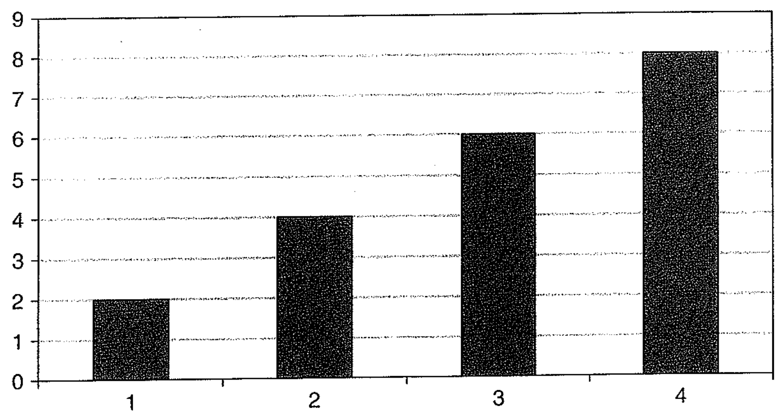
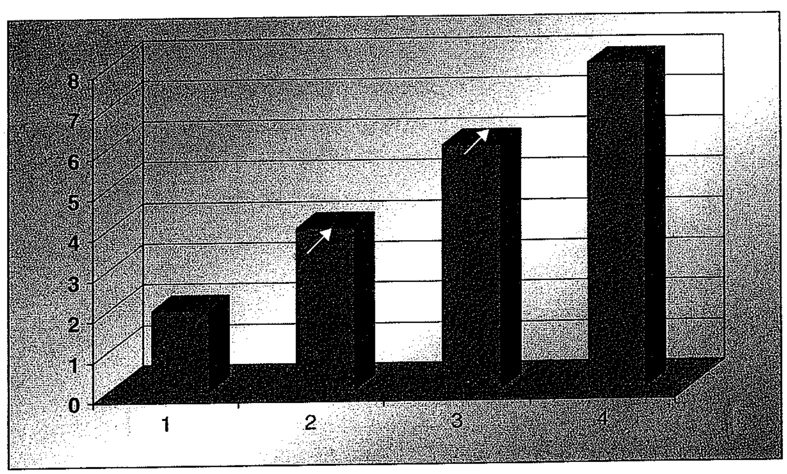


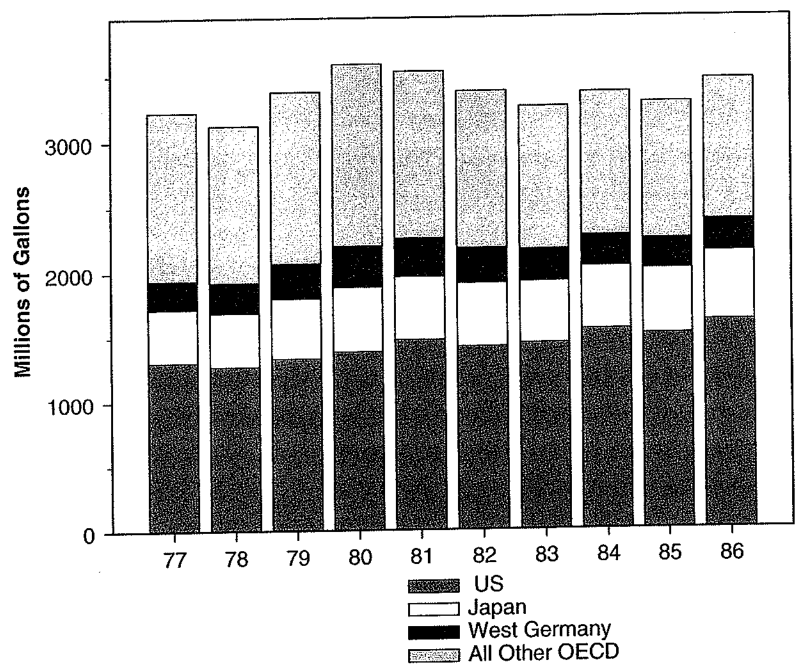
Fig. 2.8 Three-Dimensional Bar Data: Excel



Again, I show the same data in a *two-dimensional bar chart* (Figure 2.7). Do you still agree with the numbers you wrote down?

The problem with three-dimensional bar charts such as Figure 2.8 is that virtually no one knows how to read them. Do you read from the front of the bar, as shown by the arrow on bar 2, or from the back of the bar, as shown by the arrow on bar 3? Figure 2.8 was drawn using Excel 2000. We see in this case that neither supposition is correct. I assume that we are supposed to visualize a plane with a height of 2, and that the top of the first bar is supposed to look tangential to that plane. It doesn't work for me. Does it work for you?

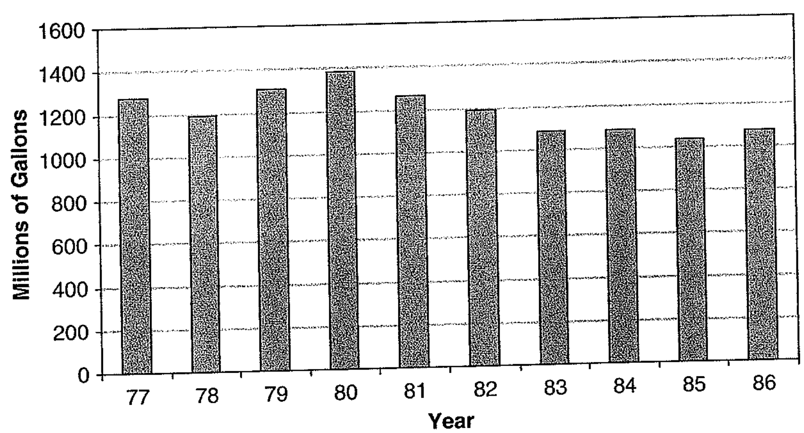
Fig. 2.11 Energy Data



2.3 BAR CHARTS: STACKED AND GROUPED

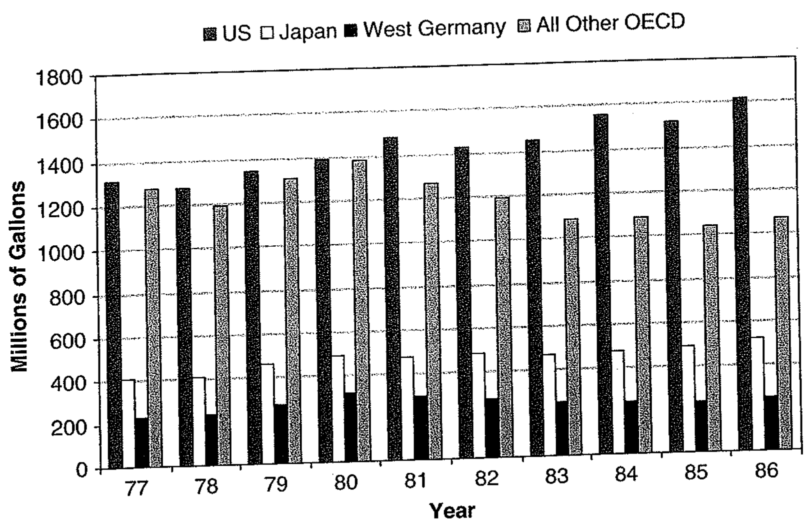
Another common graph form is a *stacked bar chart.* Figure 2.11 shows petroleum stocks from 1977 to 1986 in millions of gallons for the United States, Japan, West Germany, and all other countries of the Organisation for Economic Co-operation and Development (U.S. Dept. Energy, 1986). You probably read the values for the United States and the totals quite accurately. Study the chart and see what you can discover about the other countries.

Fig. 2.12 Energy Data: All Other OECD



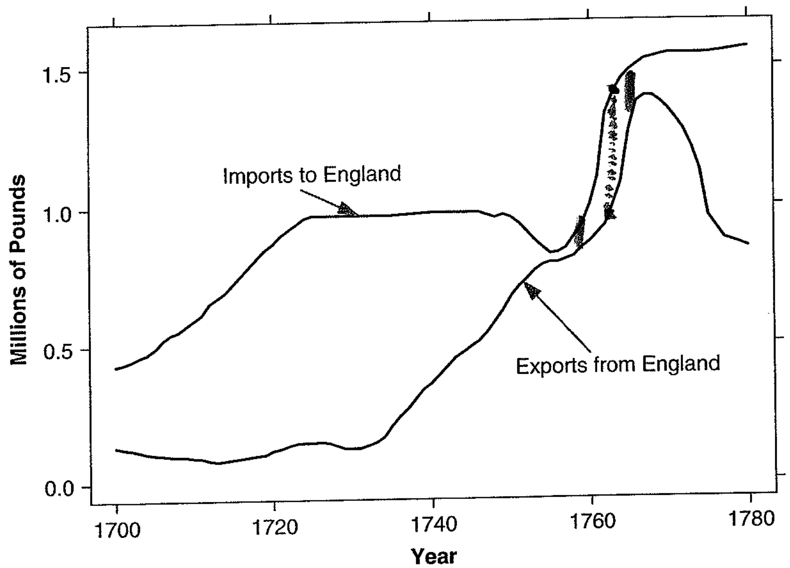
Did you notice in Figure 2.11 that the values for "all other OECD" generally tend to decrease over time? You probably didn't. As we shall see in Chapter 3, it is very difficult to judge lengths that do not have a common baseline.

Fig. 2.13 Energy Data: Grouped Bar Chart



The bars in *grouped bar charts* do have a common baseline. However, a grouped bar chart such as Figure 2.13 becomes difficult to read with even a few groups. It is difficult to follow the trend for a given group such as Japan because the data for the other groups fall between the consecutive values for Japan. Reordering the shadings helps to make the groups distinguishable. The pattern of the "all other OECD" group is certainly clearer than in the stacked bar chart. However, trellis displays, which are discussed in Chapter 5, are far clearer than is a grouped bar chart.

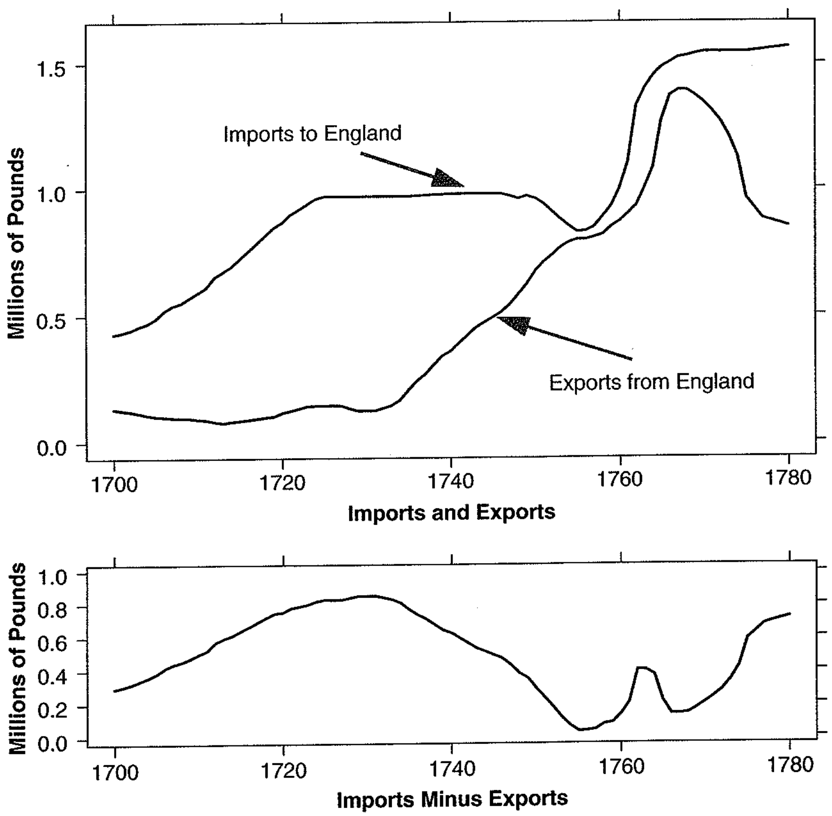
Fig. 2.14 Playfair's Balance-of-Trade Data



2.4 DIFFERENCE BETWEEN CURVES

Most of the graph forms that have been used until recently were introduced by William Playfair in the late eighteenth and early nineteenth centuries. Figure 2.14 uses Playfair's data (Playfair, 1786) to show exports from England and imports to England in trade with the East Indies. We're interested in the balance of trade, which is the difference between exports and imports. We see that the difference is about 0.4 minus 0.2 or 0.2 in 1700, and then it increases for awhile. I'd like you to continue sketching the difference.

Fig. 2.15 Playfair's Balance-of-Trade Data: Imports Minus Exports

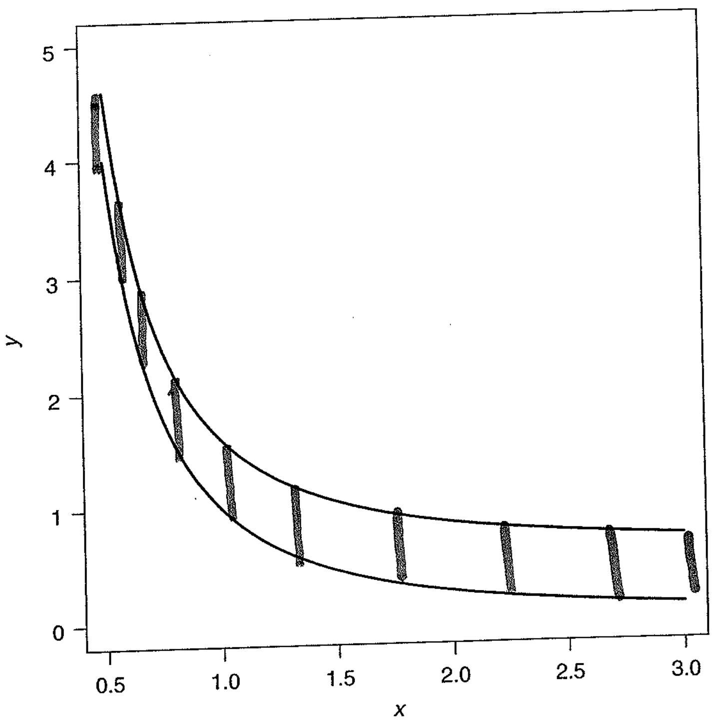


Did you notice the hump after 1760? We miss it because our eyes look at the closest point rather than the vertical distance.

It is important to remember to plot the variable of interest. If interested in the balance of trade, plot the difference rather than just the imports and exports. If we have *before* and *after* data and are interested in improvement, we plot the improvement, not just the *before* and *after* data.

Fig. 2.16 Difference between Curves

NOT RECOMMENDED



Look at the two curves in Figure 2.16. For what values of x are they closest together and also, farthest apart?