

Communicating Data Clearly

A tutorial by Naomi B. Robbins

Strata Conference
London
November 13, 2013

Contact information:

Naomi B. Robbins

11 Christine Court

Wayne, NJ 07470-6523

Phone: (973) 694 - 6009

naomi@nbr-graphs.com

<http://www.nbr-graphs.com>

blogs.forbes.com/naomirobbsins

twitter: @nbrgraphs

Copyright © 2013 Naomi B. Robbins

Communicating Data Clearly

Naomi B. Robbins

Introduction

1. For our purposes, one graph is considered more effective than another if its quantitative information can be decoded more quickly or more easily by most observers.

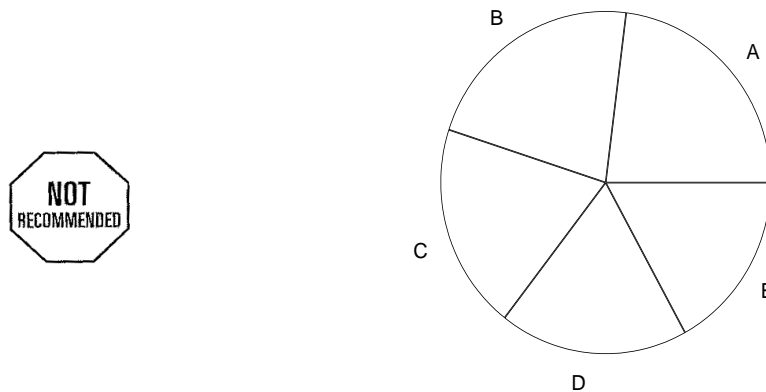


Figure 1. Pie Chart. This pie chart has five wedges. Please order them in size order from largest to smallest.

2. A dot plot is more effective than a pie chart for ordering the sizes of A through E above.
3. Bertin's definition of efficiency: "If, in order to obtain a correct and complete answer to a given question, all other things being equal, one construction requires a shorter observation time than another construction, we can say that it is more efficient for this question." Bertin (1983, p. 139)
4. A table is often very effective for small data sets. Tufte (1983, p. 178)
5. Objectives of Short Course:
 - Understand the distortions and/or limitations of popular displays (e.g., pseudo-three-dimensional bar charts, stacked bar charts, and pie charts), realize that these displays do not communicate as well as alternative forms (dot plots and multipanel plots), and understand why that is so.
 - Know more effective ways to present data and know where to find more information on these graph forms.
 - Be familiar with methods for presenting more than two variables on two-dimensional paper or screens.
 - Recognize common forms of misleading and deceptive graphs so that they will avoid using these forms and also read graphs more critically.
 - Learn general principles for creating clear, accurate graphs.

- Know when to use logarithmic scales.
 - Identify common mistakes and how to avoid these mistakes.
 - Understand that different audiences have varying needs and the presentation should be appropriate for the audience.
6. Summary of Course:
- Limitations of some common graph forms
 - Human perception and our ability to decode graphs
 - Newer and more effective graph forms
 - Trellis graphics and other innovative methods to present more than two variables
 - General principles for creating effective tables and graphs
 - Before and after examples

Limitations of Some Very Common Graph Forms

7. A dot plot shows the structure of the data better than a pie chart does.
8. “A table is nearly always better than a dumb pie chart; the only worse design than a pie chart is several of them, for then the viewer is asked to compare quantities located in spatial disarray both within and between pies.... Given their low data-density and failure to order numbers along a visual dimension, pie charts should never be used.” Tufte (1983, page 178)
9. “Pie charts have severe perceptual problems. Experiments in graphical perception have shown that compared with dot charts, they convey information far less reliably. But if you want to display some data, and perceiving the information is not so important, then a pie chart is fine.” Becker and Cleveland (1996, p. 50)
10. Pseudo-three-dimensional pie charts and exploded pie charts distort the data even more.

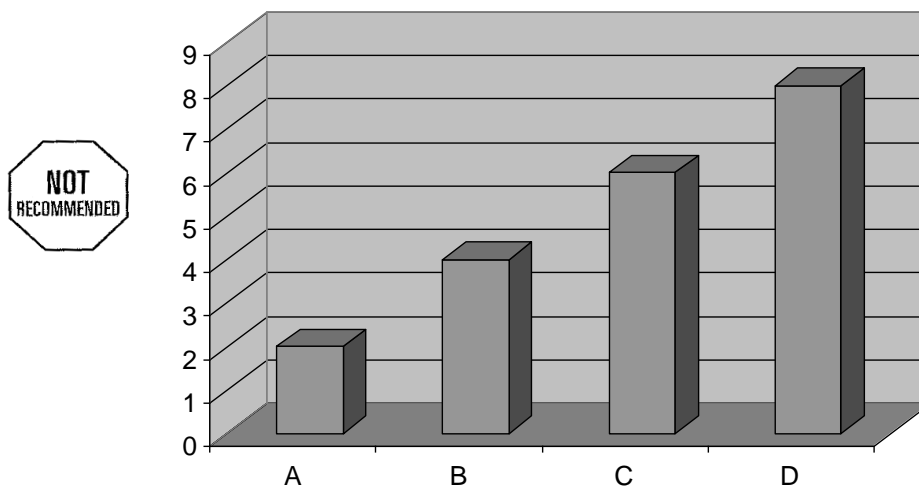


Figure 2. Excel 3-D Bar Chart. Avoid putting extra dimensions in your charts. The pseudo three-dimensional charts are difficult to read. A two-dimensional chart is clearer than a pseudo-three-dimensional one.

11. Avoid putting extra dimensions in your charts. The pseudo-three-dimensional charts are difficult to read. If you know categories and values for each category, a two-dimensional chart is clearer than a pseudo-three-dimensional one.
12. The way to read pseudo-three-dimensional bar charts depends on the software used to create them. However, we're rarely told what software was used.
13. Data labels don't help; they confuse the reader even more.
14. True three-dimensional charts are also confusing to read.

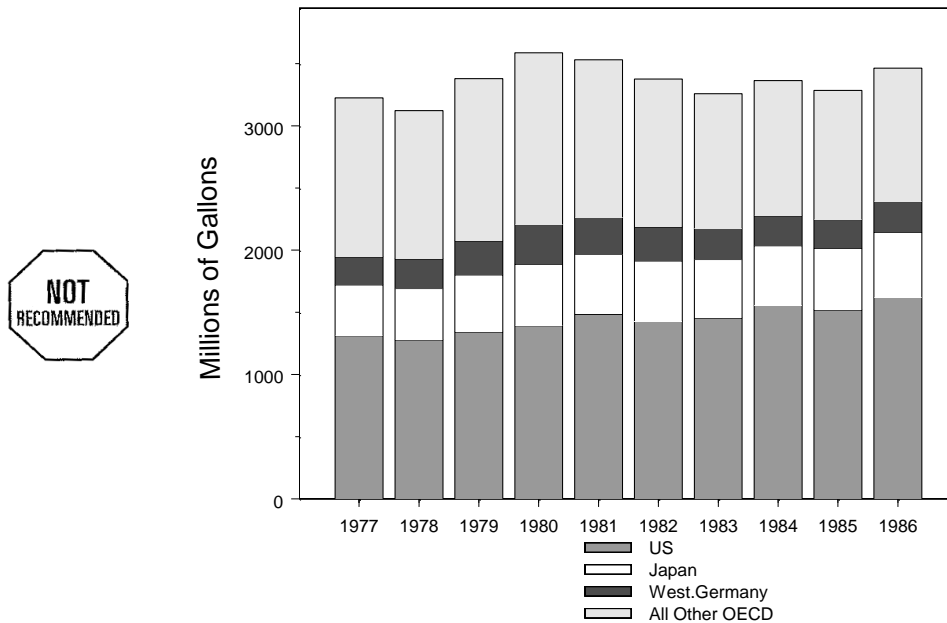


Figure 3 Stacked Bar Charts. The bottom level and the totals are clear but it is difficult to see trends in the other layers.

15. It is difficult to determine trends from stack bar charts unless we are looking at the bottom category or the total since lengths without a common baseline are difficult to compare.
16. Many grouped bar charts are difficult to follow since there is so much extraneous information between the bars we wish to compare.
17. It is difficult to judge the difference between curves. It is usually better to let the computer do the calculations and to plot the difference directly.
18. Areas are difficult to judge. Dot plots are more effective than area or bubble plots.
19. Other charts, such as ribbons, donuts and gauges, don't communicate well.

Human Perception and Our Ability to Decode Graphs

20. There are three stages of memory: iconic, short term, and long term
 - Iconic
 - Very rapid
 - Automatic and unconscious

- Preattentive processing
 - Detect a limited set of visual attributes
 - Short term
 - Temporary
 - Limited storage capacity
 - Long term
21. Some preattentive tasks include size, position, orientation, curvature, gray value, hue, shape, enclosure, and number (up to four).
22. Gestalt rules for perceptually grouping objects include:
- Proximity
 - Similarity
 - Connectedness
 - Continuity
 - Symmetry
 - Closure
 - Size
 - Enclosure

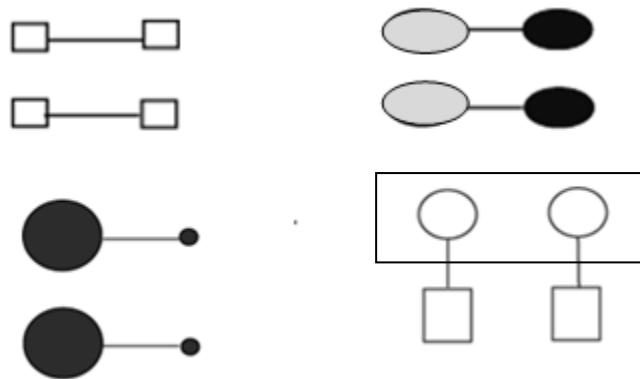


Figure 3a Law of Connectedness. This figure shows that the law of connectedness is stronger than proximity, similarity, size, and shape.

23. Cleveland's list of graphical perception tasks in alphabetical order: Cleveland (1985, p. 254)
- Angle
 - Area
 - Color hue
 - Color saturation
 - Density (amount of black)

- Length (distance)
 - Position along a common scale
 - Position along identical, nonaligned scales
 - Slope
 - Volume
24. Angle judgments are subject to bias.
- Acute angles underestimated
 - Obtuse angles overestimated
 - Angles with horizontal bisectors appear larger than those with vertical bisectors.
25. Area judgments are biased. Area judgments are much less accurate than length and position judgments. Volume judgments are even more biased.
26. Colors become grayer as they become less saturated.
27. Color hue, color saturation, and lightness are very effective for a categorical variable, but not for displaying the values of a quantitative variable.
28. It is difficult to compare lengths without a common baseline. The percentage difference is more important than the absolute difference when comparing lengths.

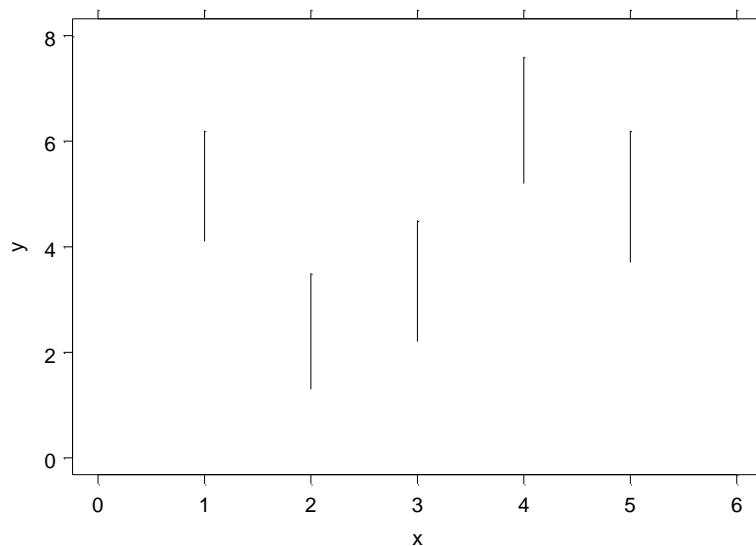


Figure 4. Lengths. It is difficult to compare lengths without a common baseline. The percentage difference is more important than the absolute difference when comparing lengths.

29. Steven's Law: Let x be the magnitude of an attribute of an object such as its length or area. According to Stevens' Law, the perceived scale is proportional to x^β where β usually ranges, as has been determined by experimentation, from 0.9 to 1.1 for length, from 0.6 to 0.9 for area, and from 0.5 to 0.8 for volume.
30. Dot plots allow us to decode the data by making judgments of positions along the common horizontal scale. Experiments have shown that this is the most accurate of the elementary

graphical tasks. We judge position along identical nonaligned scales almost as accurately as position along a common scale.

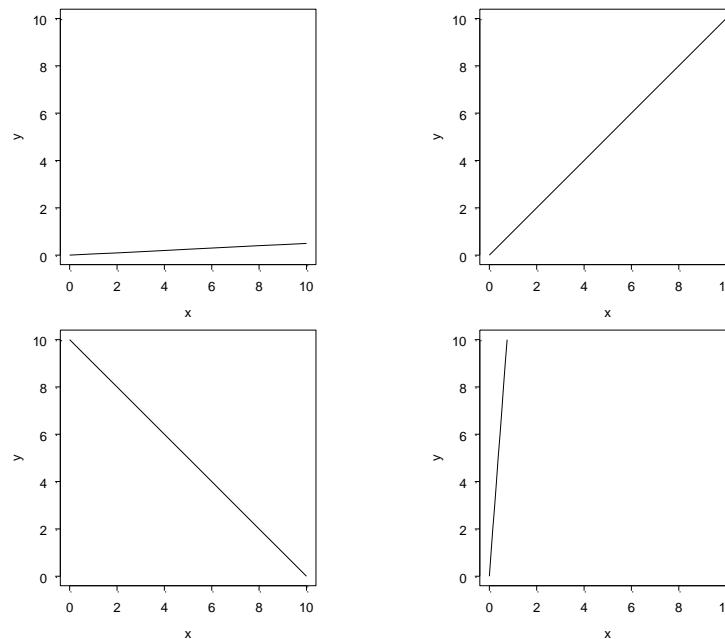


Figure 5. Judging Angles. The accuracy of judgments of slopes of line segments depends on the angle with the horizontal. Poor accuracy will result from angles close to $\pi/2$.

31. People use angle judgments to determine slopes.
32. Cleveland's hierarchy of tasks ordered by our ability to perform accurate judgments:
 - 1. Position along a common scale
 - 2. Position along identical, nonaligned scales
 - 3. Length
 - 4. Angle - Slope
 - 5. Area
 - 6. Volume
 - 7. Color hue - Color saturation – Density
33. Creating a more effective graph involves choosing a graphical construction where the visual decoding uses tasks as high as possible on the ordered list of elementary graphical tasks while balancing this ordering with distance and detection.
 - Distance: The closer together objects are, the easier it is to compare them. As distance between the objects increases, accuracy of judgments decreases.
 - Detection: Before we can perform any of the elementary tasks, we must be able to detect the data. We often cannot if data points overlap or are hidden in the axes or tick marks.

Newer and More Effective Chart and Graph Forms

34. Dot plots use judgments of position along a common scale and are extremely effective.

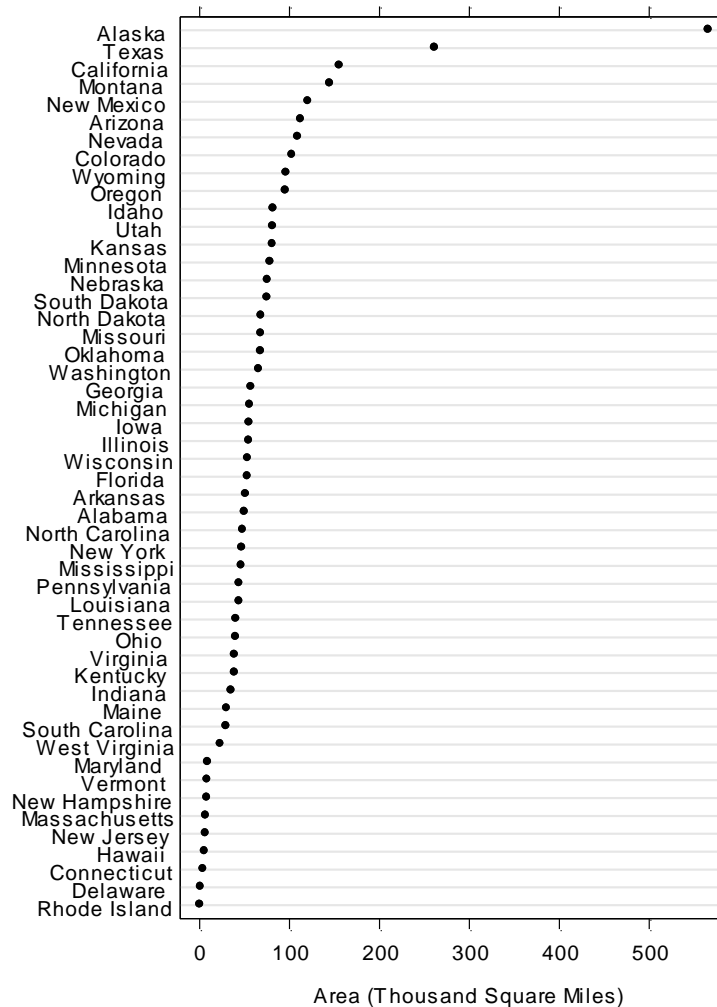


Figure 6. Dot Plots. Dot plots use judgments of position along a common scale. They don't get as cluttered as bar charts do, don't require a zero baseline, and work well with logarithmic scales.

35. Alphabetical order is rarely the most effective. Ordering by size is often better.
36. *Showing data on a logarithmic scale can cure skewness towards large values.*¹
37. The most common base is ten.
38. Base 2 is useful when the data do not range over orders of magnitude.
39. *Use a logarithmic scale when it is important to understand percent change or multiplicative factors.*
40. Know your audience. Not all audiences are comfortable with logarithms.

¹ An italicized statement signifies that the statement is a direct quote from Cleveland's *Elements of Graphing Data*.

41. It is sometimes helpful to use two sets of labels for the pair of scales lines when using logarithmic axes.
42. These useful Web sites describe how to draw dot plots with Excel:
 - <http://www.nbr-graphs.com/resources/other-resources/>
 - <http://www.exceluser.com/dash/dotplot.htm>
 - <http://www.peltiertech.com>²
43. I recommend the utilities at <http://www.peltiertech.com>.
44. Bar plots become cluttered more quickly than do dot plots.
45. Most readers would have little problem understanding either the dot plot or the bar chart. Note that the dot plot is less cluttered, less redundant, and uses less ink.
46. It is possible to add another series to a dot plot without the figure becoming cluttered.
47. Error bars show up better with dot plots than with bar charts.
48. Logarithmic scales may be used with dot plots.
49. Scatterplots are an effective means of showing the relationship between two variables.

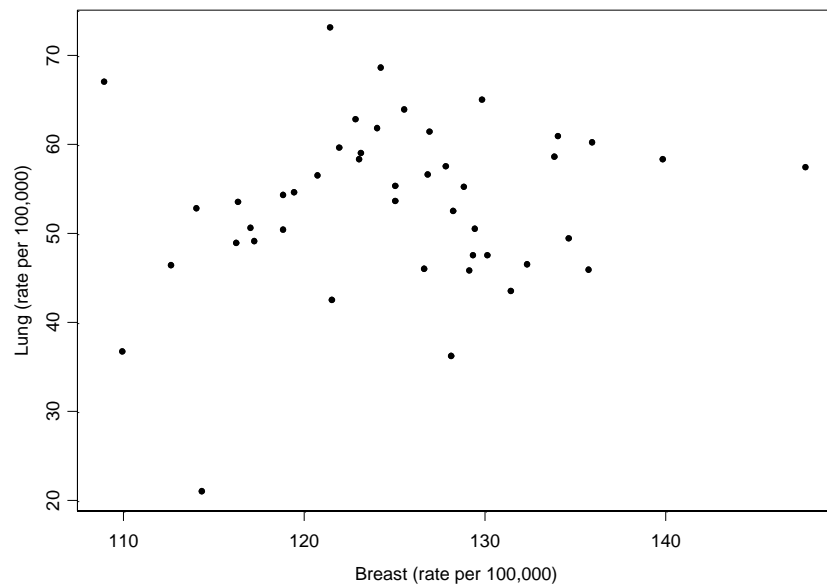


Figure 7. Scatterplot. Scatterplots show the relationship of variables.

50. Using labels as plotting symbols can reduce clutter.
51. Loess (locally weighted regression) is a scatterplot smoother that is useful for fitting nonlinear curves.
52. Histograms are useful for showing the distribution of one variable.
53. Box plots are better for comparing distributions.

² Affiliate

54. Horizontal box plots allow for longer labels.
55. Vertical box plots have values on the vertical axis. Some audiences are confused by values on the horizontal axis.
56. Use consistent style in a presentation or document.

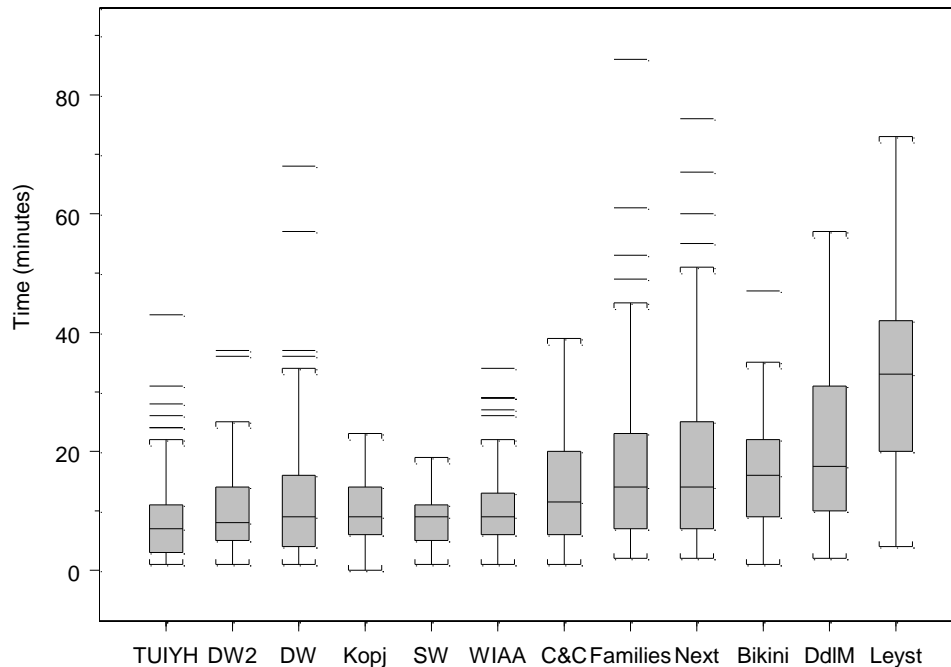


Figure 8. Vertical Box Plots. Vertical box plots have values on the vertical axis. Some audiences are confused by values on the horizontal axis.

57. Tables with microplots help to visualize the information in the table. (Heilberger and Holland, 2005)
58. Time series were one of the earliest graphs drawn.
59. Time series may be drawn with symbols, connected symbols, lines, or vertical lines.
60. "Sparklines are *wordlike* graphics." Tufte (2006, Chapter 2), <http://sparkline.org/>
61. Bullet graphs compare actual performance to a goal or target value.
62. Multiple line plots are often used to show a day of the week or month of the year effect. However, it is difficult to follow the trend over the weeks or months.
63. It is difficult to examine the day of the week or month of the year effect with a time series plot.
64. Month plots, also called cycle plots, are useful for showing a day of week or month of year effect.

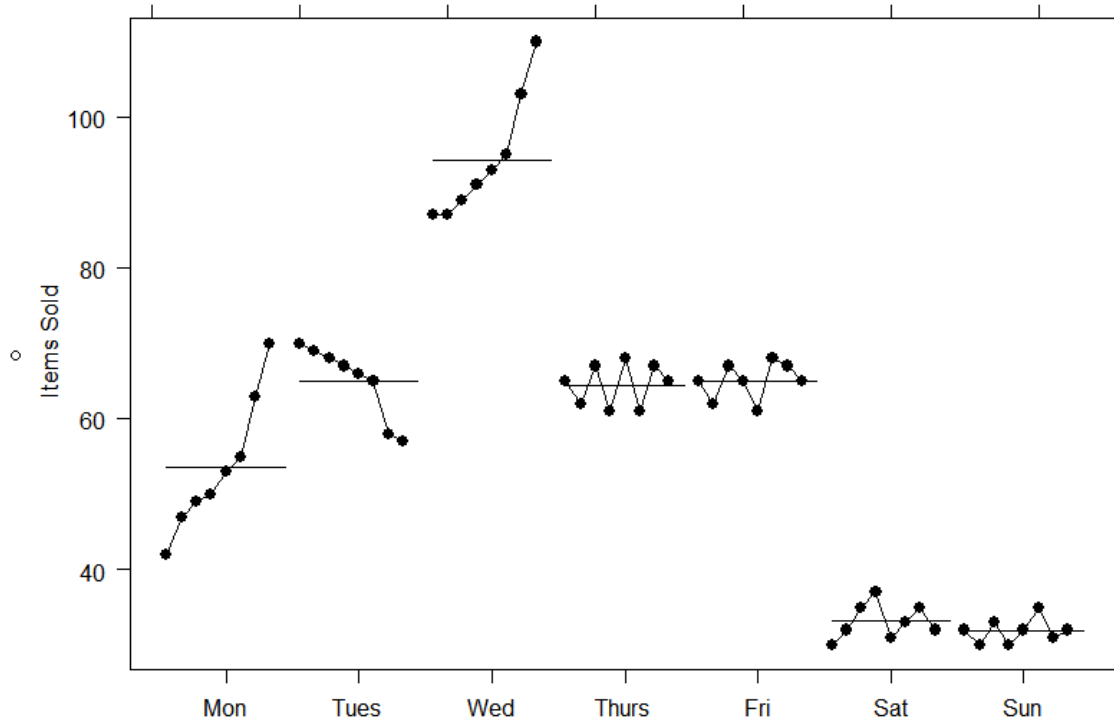


Figure 9. Month plot of items sold. First, all Monday values are plotted, then all Tuesday values, and so forth. Each dot represents a week: from left to right, the eight week period). The trend for each day is shown clearly, yet we still see daily effects such as that sales are highest on Wednesdays. We also see that sales are increasing on Mondays and Wednesdays over the eight week period but decreasing on Tuesdays. The horizontal lines represent the mean for each day.

Trellis Graphics and Other Methods for Showing More than Two Variables

65. Trellis displays provide a framework for multivariate data. They are often extremely useful.
66. A major feature of Trellis displays is multipanel conditioning.
67. Multipanel plots often avoid the need for color.
68. Terminology for multipanel plots and Trellis:
 - Multipanel Plot: any plot with more than one panel.
 - Small Multiples: “a series of graphics, showing the same combination of variables, indexed by changes in another variable.” Tufte (2001, p. 170)
 - Trellis Display: Small multiples with structure imposed; e.g., ordering of panels.
 - Trellis trademarked by Insightful Corporation (now Tibco)
 - Often called lattice displays
 - Term trellis also used for a graphics system in S-Plus

- 69. The Trellis plot of the barley data clearly shows an anomaly of the data that statistical analyses missed.

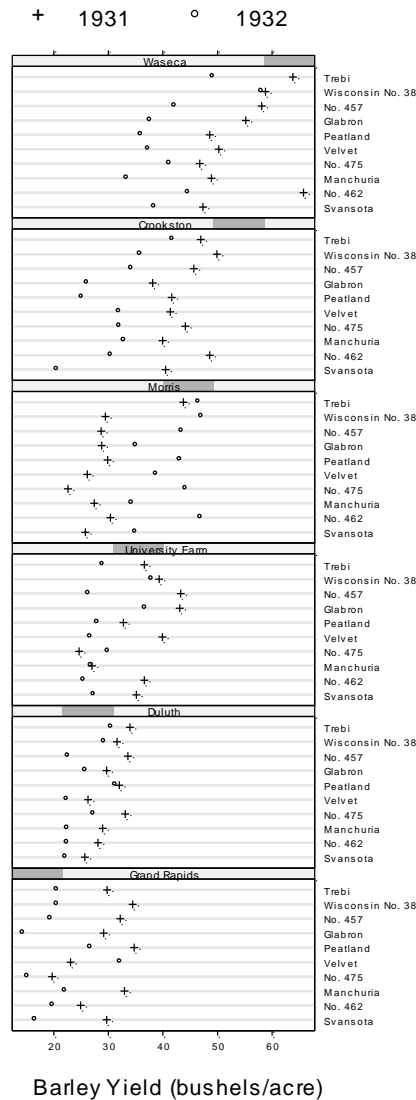


Figure 10. Barley Example. This figure shows the power of visualization and of Trellis displays by showing the anomaly at the Morris site in the barley data set, which was not discovered in 60 years of conventional statistical analyses.

- 70. Bar charts get cluttered more quickly than dot plots.
- 71. A great deal of information can fit on a graph without it being cluttered.
- 72. Diverging stacked bar charts are a preferred way to plot Likert scales.
- 73. Florence Nightingale introduced the rose plot in 1858 in her attempt to improve sanitary conditions for the British forces during the Crimean war.
- 74. Nightingales’s rose plots are also called coxcomb plots, radial area plots, and wedges graphs.

75. The reader does not know whether the values are encoded in the area or the radius of a Nightingale rose. Florence Nightingale encoded the values in the area by making the square root of the radius proportional to the values.
76. Four-fold plots are a variation of the Nightingale rose that are useful for 2x 2 data.

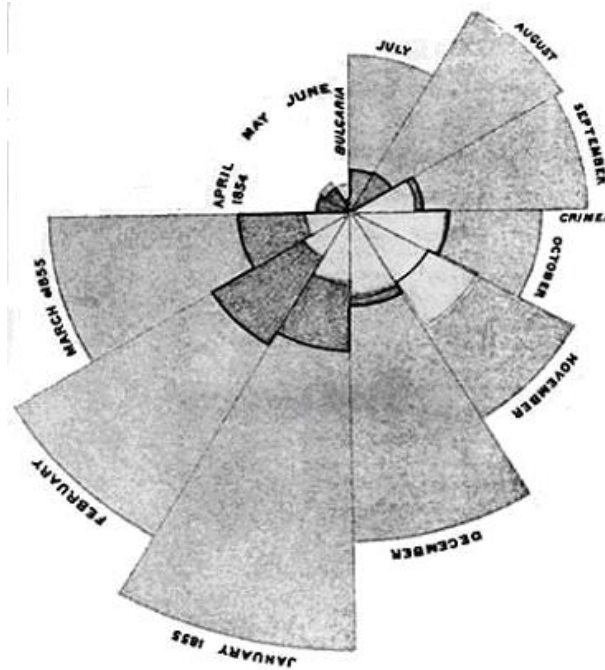


Figure 11. Nightingale's Rose. Florence Nightingale introduced her rose plot in 1858. This figure helped to convince the government to improve sanitary conditions during the Crimean war.

77. Linked micromaps are useful for geographically referenced data.
78. Scatterplot matrices show all pairs of variables. Some audiences find scatterplot matrices difficult to read.
79. The axes are parallel to one another in parallel coordinate plots. One use for parallel coordinate plots is to classify data into subgroups.
80. Mosaic plots are used for multivariate categorical data.
81. Color is useful to distinguish categorical variables. It is sometimes used to show ranges of temperatures as in a weather plot.
82. TableLens replaces the numbers in a table with bars that aid visualization and highlight correlations and exceptions. (Rao, 2006)

General Principles for Creating Effective Charts and Graphs

83. Outline of Section
 - Can we see what is graphed?
 - Can we understand what is graphed?
 - Scales

84. Terminology
 - Scale-line rectangle: the rectangle formed by the horizontal and vertical axes.
 - Data rectangle: the smallest rectangle enclosing the data.
85. Metrics for Graphs
 - Bertin's Efficiency
 - Tufte's Lie Factor -Size of effect in graphic / size of effect in data
 - Tufte's Data Density -Number of entries in data matrix / area of data graphic
 - Tufte's Data-Ink Ratio -Data ink / total ink used to print the graphic
86. *Make the data stand out.* Deemphasize non-data elements.
87. Look at the graph and notice what you see first. The answer should be the data (or model) and not grid lines, long labels, or other graphical elements.
88. Eliminate unnecessary clutter in the graph and the surrounding page.
89. *Use visually prominent graphical elements to show the data.*
90. *Overlapping plotting symbols must be visually distinguishable.*

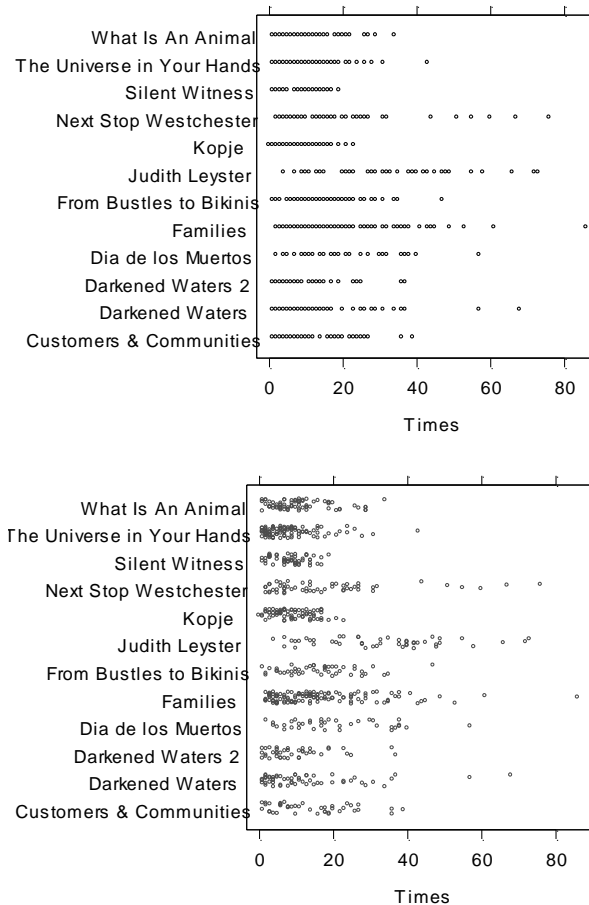


Figure 12. Jittering. The top figure has many overlapping symbols. The bottom figure eliminates the overlapping by adding small random amounts to the data. This is called jittering.

91. Jittering distinguishes overlapping symbols when the sample size is not too large.
92. Hexagonal binning is useful when the sample size is so large that scatterplots would form black blobs.
93. *Superposed data sets must be readily visually assembled.*
 - It is difficult to discriminate groups of data when just shapes are varied.
 - Varying open and closed fill works well when there is no overlap in the group with closed fill and not too many groups.
 - Varying shades of gray works well when there are not too many groups.
 - Cleveland recommends a sequence of symbols including open circles, plus signs, and less than symbols when there is overlap. Cleveland (1994, p. 238)
 - How well letters work depends on the choice of letters. Obviously, combinations like O, C and Q do not work well.
 - Color is the best means for distinguishing groups of data for those with normal color vision. Varying shape as well as color helps for color vision deficits.
94. Trellis displays can avoid the need for superposed symbols.
95. *Do not clutter the interior of the scale line rectangle.*
96. Some ways to reduce clutter:
 - Show axes labels in thousands, millions, or billions instead of including strings of zeros.
 - Avoid the need for data symbols and labels by using labels as symbols.
 - Move axis labels away from zero when there are positive and negative numbers.

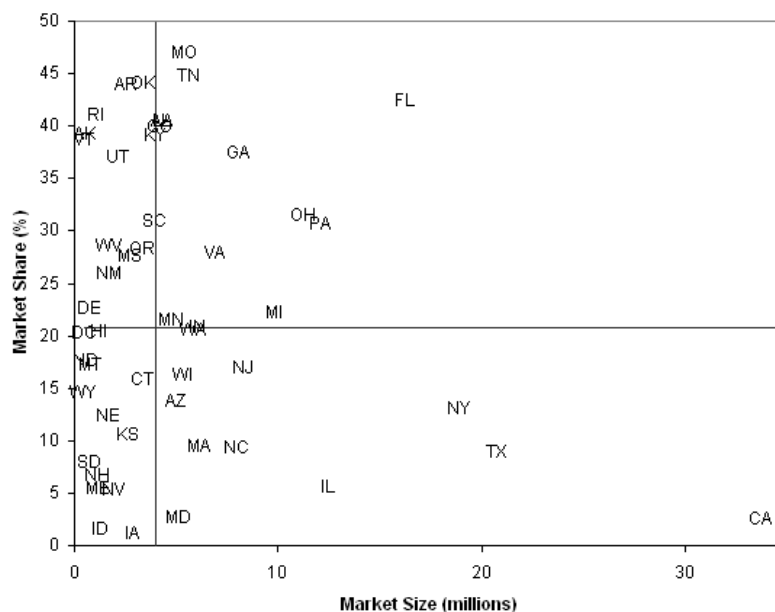


Figure 13. Labels as Plotting Symbols with Excel. Excel can produce graphs that are not on any of its menus and add-ons to Excel are available for downloading.

- Label an axis as percent or dollars rather than including a percent sign or dollar symbol at each tick mark label.
97. Excel can produce graphs that are not on any of its menus and add-ons to Excel are available for downloading. The XY chart labeler to use labels as plotting symbols was downloaded from <http://www.appspro.com/Utilities/ChartLabeler.htm>.
 98. *Do not allow data labels in the interior of the scale-line rectangle to interfere with the quantitative data or to clutter the graph.*
 99. Labels help to spot plotting errors and to see interesting aspects of the data.
 100. Integrate evidence regardless of mode. Tufte (2006, p. 118)
 101. It is often useful to plot data more than one way. Each presentation adds different insights to the data.
 102. Verbal arguments do not resolve design questions. Visual evidence decides visual issues. Tufte (2006, p. 119)
 103. Clutter calls for a design solution. Tufte (2006, p. 120)
 104. *Use a pair of scale lines for each variable.*

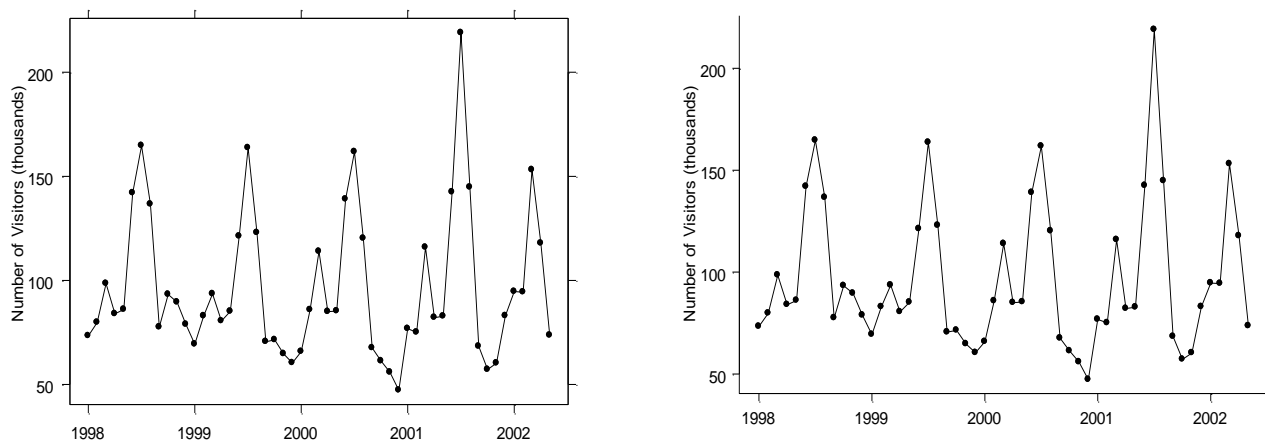


Figure 14. Scale Lines. *Use a pair of scale lines for each variable.* It is more difficult to read values on the right and the top in the figure on the right.

105. Clarity and conciseness are desirable attributes of graphs. Sometimes, however, clarity and conciseness conflict with one another.
106. *Make the data rectangle slightly smaller than the scale line rectangle.*
107. *Tick marks should point outward.*
108. *Do not overdo the number of tick marks* or tick mark labels. Too few labels are also a problem.
109. Tick marks should be at sensible values.
110. Do not overdo the number of decimal places, dollar signs, or percent signs in the labels.

111. Use a reference line when there is an important value that must be seen across the entire graph, but do not let the line interfere with the data.

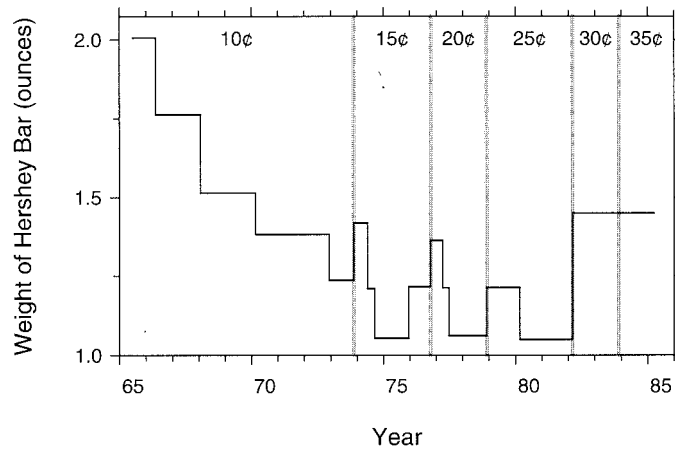


Figure 15. Reference Lines. Reference lines at the times the price of a Hershey Bar changed help to understand the data.

112. Avoid putting notes and keys inside the scale-line rectangle.
 113. Deemphasize grid lines and distinguish grid lines from data.
 114. Visual clarity must be preserved under reduction and reproduction.

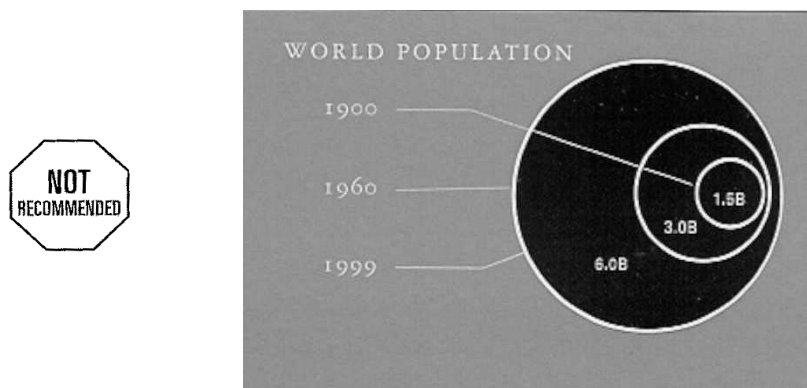


Figure 16. Proofread Graphs. Notice that two lines point to the outer circle and none to the middle one.

115. Proofread graphs.
 116. Put major conclusions into graphical form. Make captions comprehensive and informative.
 117. Captions should:
 - Describe everything that is graphed.
 - Draw attention to the important features of the data.
 - Describe the conclusions that are drawn from the data on the graph.
 118. Error bars should be clearly explained. Error bars can represent:
 - The sample standard deviation of the data.

- An estimate of the standard deviation (also called the standard error) of a statistical quantity.
 - A confidence interval for a statistical quantity.
119. Standard errors form 68% confidence intervals for some distributions, which is not a particularly interesting interval.
 120. Two-tiered error bars show more than one confidence interval.
 121. Draw the data to scale.

NOT
RECOMMENDED

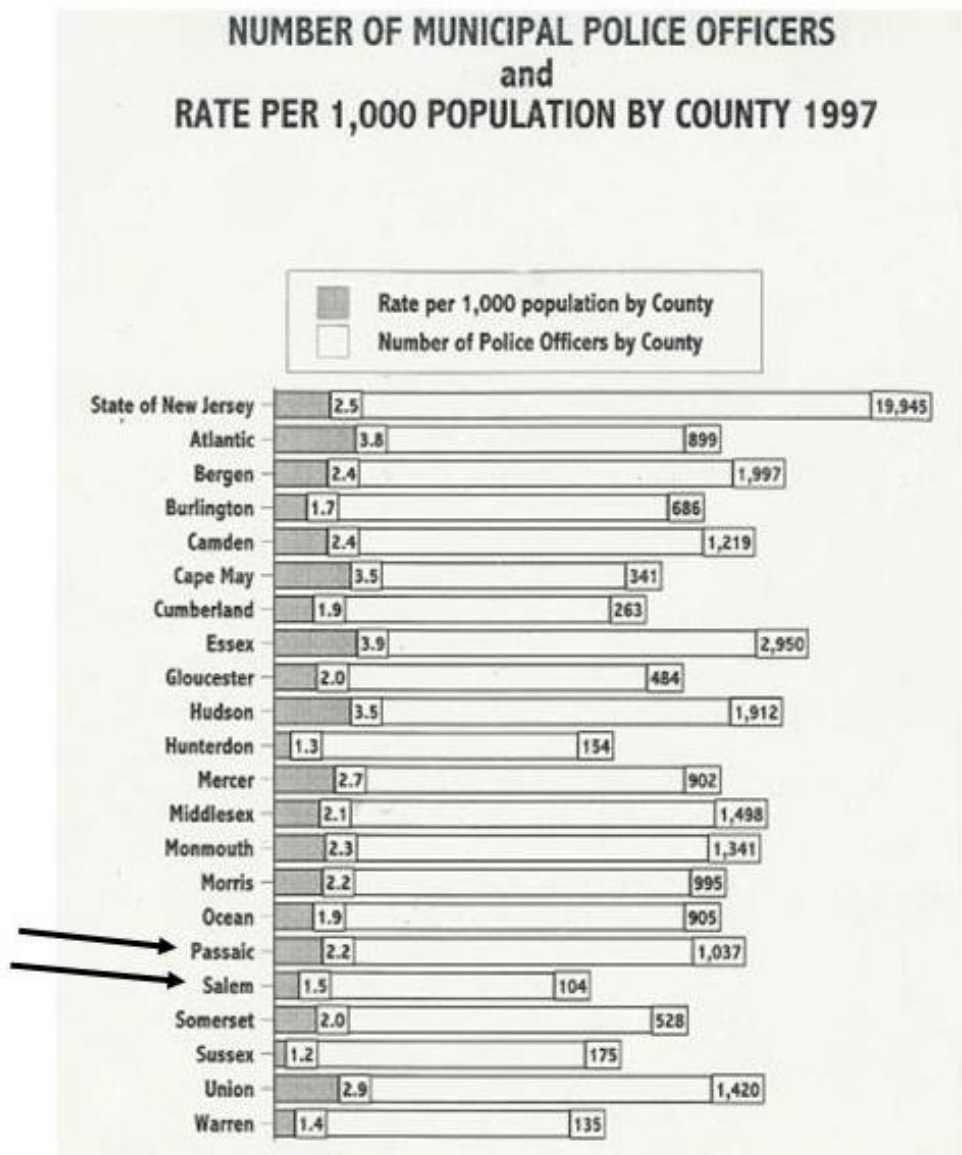


Figure 17. Draw the Data to Scale. Notice that Salem County has 104 police officers and Passaic County has 1037. Does the bar for Salem County appear to you to be about 1/10th the length of the bar for Passaic County?

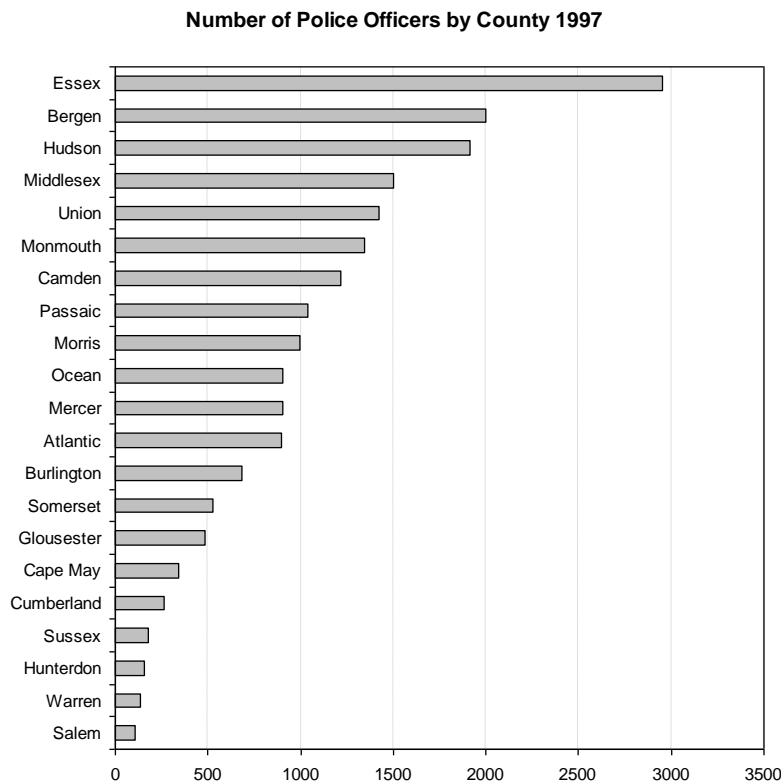


Figure 18. Number of Police Officers Redrawn. Here the number of police officers is drawn to scale.

122. Do not show changes in one dimension by area or volume.

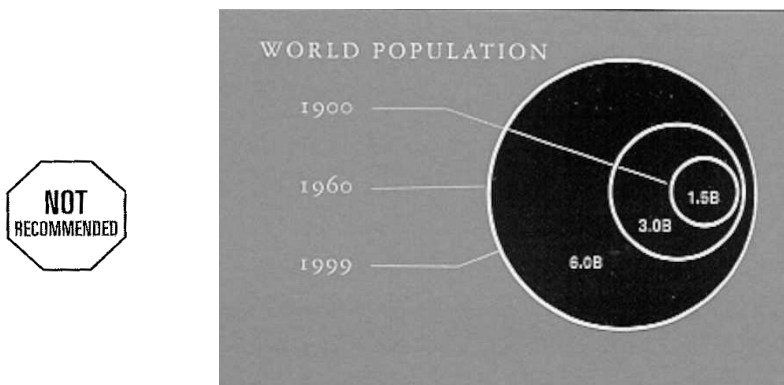


Figure 19. Do not show changes in one dimension by area or volume. The growth in population appears much greater than actually is the case by encoding the data in the diameters while we view the areas.

123. Don't use more dimensions in the figure than variables in the data.
124. Use a common baseline wherever possible.
125. Make your message clear.
126. Think about your data and the message you are trying to communicate before you start a graph. Don't be limited by a small set of graph types that your software's menu offers.
127. Use the title to reinforce your message.

128. Be careful with the placement of labels, legends, etc, so that they don't interfere with the message you are trying to communicate.
129. In presentations, avoid colors that don't project well.
130. A single plot that may be acceptable alone may confuse in a group. Be consistent with colors, scales, and other graphical elements in groups of charts.
131. Choose the principle least likely to mislead if more than one principle applies and they conflict with one another.
132. See <http://www.colorbrewer2.org> for guidelines on color.
133. When choosing colors, vary hue for qualitative differences, vary lightness or saturation for quantitative differences, and use high saturation for highlighting (normal color vision).
134. Consider the requirements of people with color vision deficiencies. See http://www.lighthouse.org/color_contrast.htm
135. Some design choices have a right or wrong: others are style choices.

Scales

136. *The aspect ratio is the ratio of the height of the data rectangle to its width.*
 - The orientation of a line segment is its angle with the horizontal.
 - Centering the orientation of line segments to 45° is called banking to 45° .

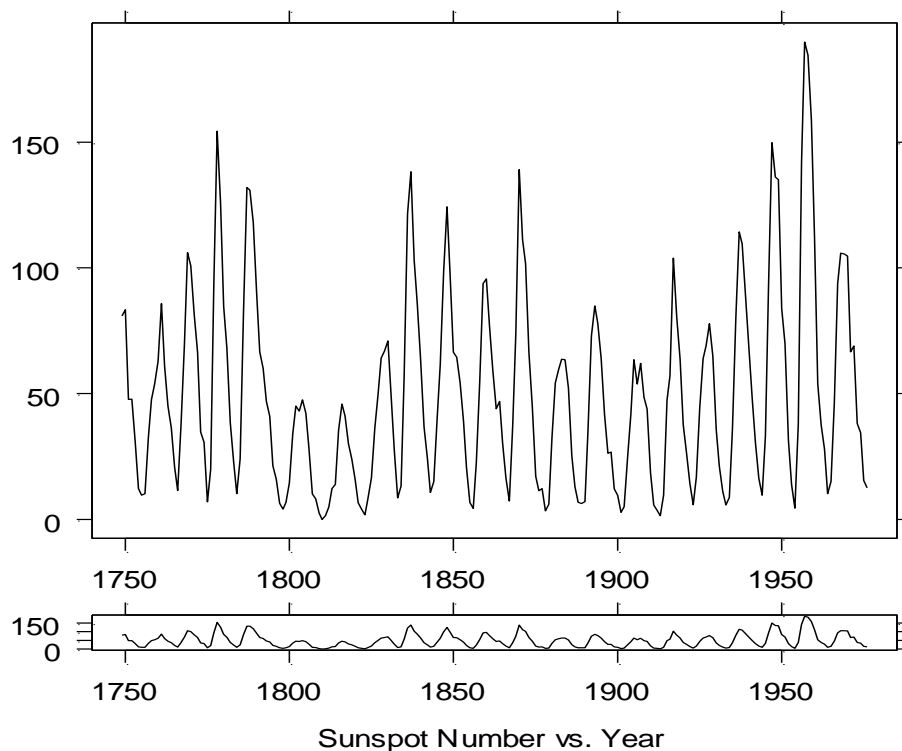


Figure 20. Banking to 45° . The bottom figure is banked to 45° . This helps notice an important property of the data that does not show up clearly in the top figure: that the cycles increase more rapidly than they decrease.

137. Choose the range of tick marks to include or nearly include the range of the data.
138. Do not insist that zero always be included on a scale showing magnitude.
139. Subject to the constraints that scales have, choose the scales so that the data rectangle fills up as much of the scale line rectangle as possible.
140. Whether zero needs to be included on scales is a controversial topic.
141. A bar graph without zero is misleading.

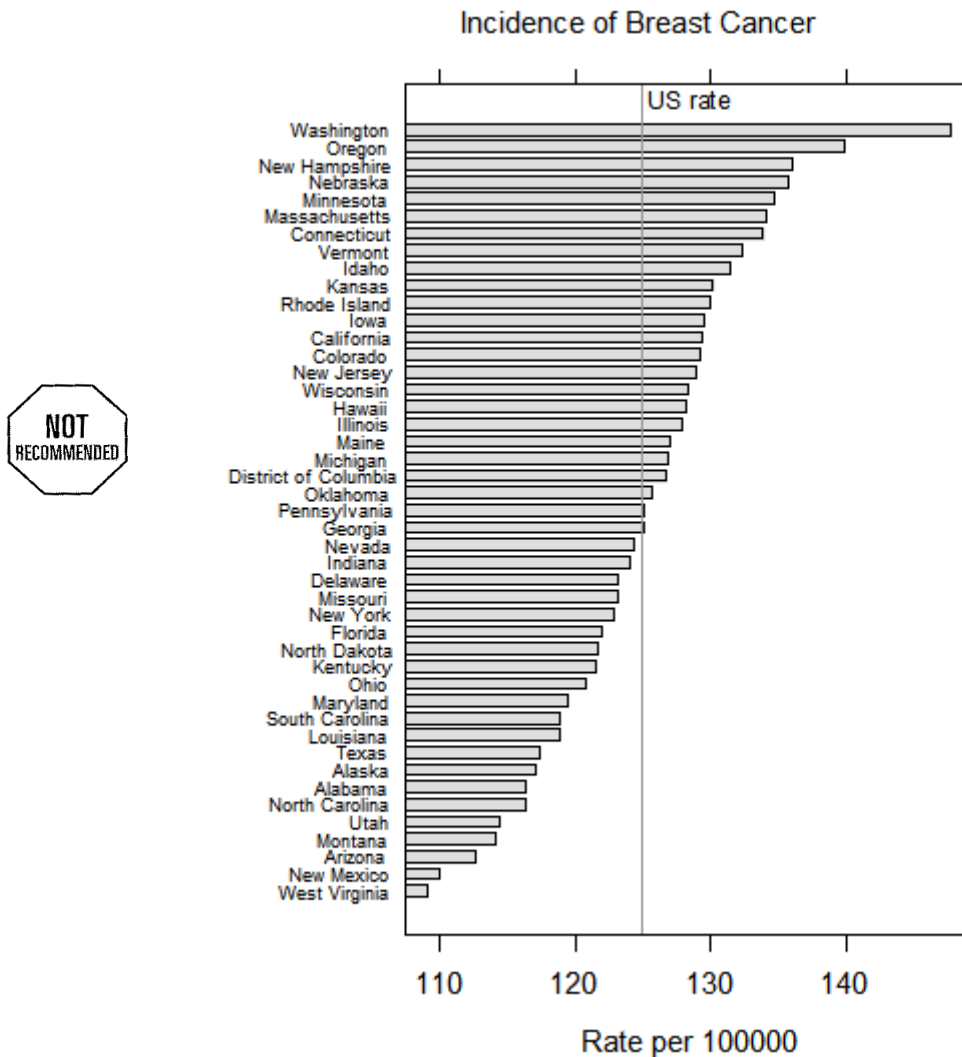


Figure 21. Bar Graph without zero. This figure is a visual lie. It gives the impression of a greater difference in rates than is actually the case. We cannot help seeing and comparing the lengths of the bars.

142. The situation is not as clear with line graphs. My position is that they are not needed for audiences who can be expected to read the scale labels.

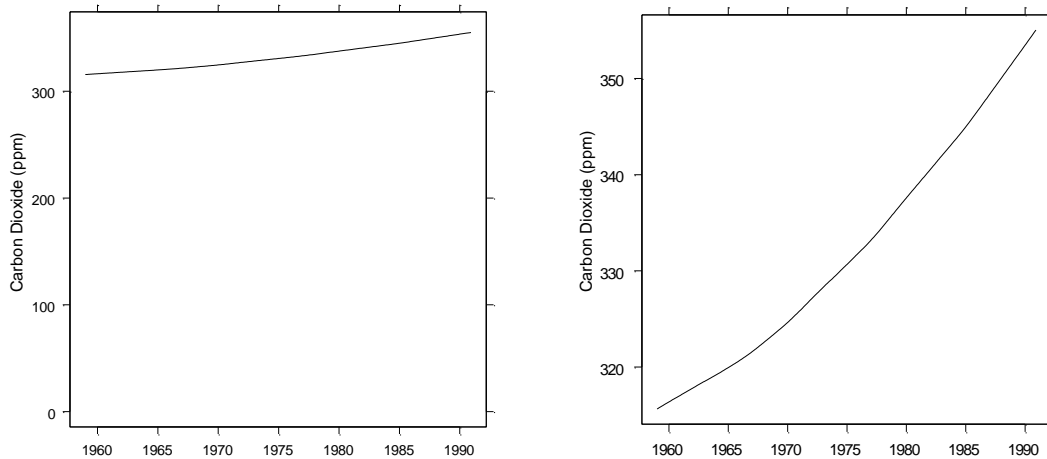


Figure 22. Including zero on a scale. Must zero be included? The left figure show carbon dioxide levels on a scale including zero; the right scale shows only the range of the data. The important fact that the rate of change is increasing is lost in the left figure.

143. *Use a scale break only when necessary. If a break cannot be avoided, use a full scale break. Taking logs can avoid the need for a break.*
144. *Do not connect numerical values on two sides of a break.*
145. *It is sometimes helpful to use the pair of scales lines for a variable to show two different scales.*
146. *Avoid deceptive double-y axes.*
147. *Do not use evenly spaced tick marks for uneven intervals on an arithmetic scale.*

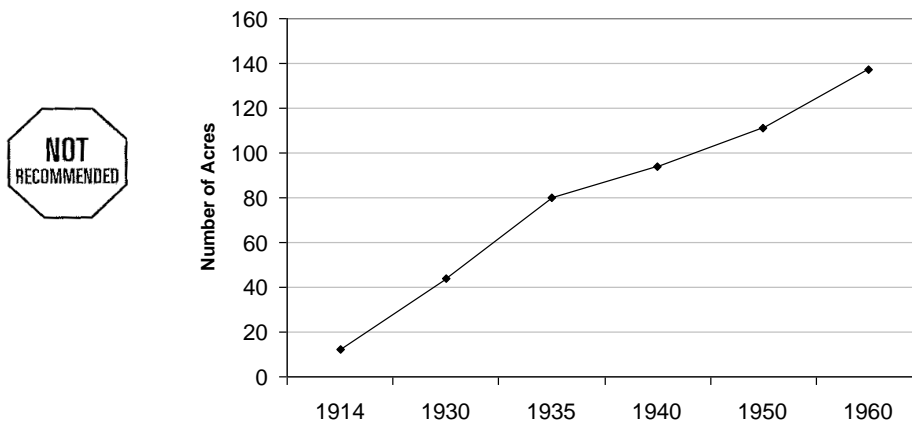


Figure 23. Uneven Intervals. Notice that the years for which data were available are not equally spaced. Excel treats the years as a categorical variable if a line chart is chosen. This is one of the most common mistakes Excel users make.

148. *Choose appropriate scales when data on different panels are compared.*
149. *Choose an aspect ratio that shows variation in the data.*

150. All axes require scales.
151. The horizontal axis should increase from left to right and the vertical axis from bottom to top.

Tables or Graphs

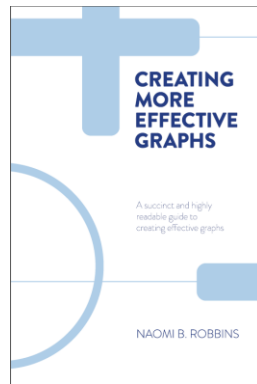
152. Graphs are for the forest; tables are for the trees.
153. Graphs show patterns, shapes and trends. Tables show exact values.
154. Graphs fit much more information in a small space.
155. When costs permit, it is useful to show both.
156. Large tables have no place on projector screens. If they must be shown, use a handout.
157. Semi-graphical tables integrate tables and graphs using small graphs such as sparklines.

Before and After Examples

158. The choice of a graph form has a dramatic affect on a reader's understanding of the data.
159. Trellis displays show characteristics of the data that divided bar charts hide.
160. A sentence or a text box show two numbers clearly; a graph is not needed.
161. Don't ask the reader to perform calculations that the computer can do more easily.
162. Many readers find radar charts to be difficult to read or confusing.
163. Readers expect the data to increase from left to right.
164. It is hard to make comparisons from multiple pie charts.
165. The number of times dot plots and Trellis displays appear in the *after* figures shows how useful they are.
166. Trellis plots often avoid the need for color.

Software

Most graphs in this presentation were drawn using S-Plus or R. R is freely downloadable from www.r-project.org. However, most can be produced in Excel. In my book *Creating More Effective Graphs* I describe dot plots, cycle plots, and other useful but little known graphs. Readers of this book have used their Web pages to provide instructions for Excel users to create these graphs which are not on Excel menus. Links to some of these resources can be found on my website—nbr-graphs.com—if you click on *Resources* from the menu and then *Other* from the drop-down list.



Appendix 1: Sources of Figures

- 1:6 Robbins, Naomi B. 2013. *Creating More Effective Graphs*, Chart House, Ramsey, NJ (reprinted from Wiley 2005). Copyright © 2013 by Chart House. All rights reserved.
- 3a,7 Robbins, Naomi B. 2006. Copyright © 2006 by Naomi B. Robbins. All rights reserved.
- 8 Robbins, Naomi B. 2013. *Creating More Effective Graphs*, Chart House, Ramsey, NJ (reprinted from Wiley 2005). Copyright © 2013 by Chart House. All rights reserved.
- 9 Robbins, Naomi B. 2006. Copyright © 2006 by Naomi B. Robbins. All rights reserved.
- 10 Cleveland, William S. 1994. *The Elements of Graphing Data (Revised Edition)*, Hobart Press, Summit, NJ. Used with permission from William S. Cleveland.
- 11 Nightingale, Florence. 1858. *Notes on Matters Affecting the Health, Efficiency and Hospital Administration of the British Army*, Quoted in <http://www.scottlan.edu/lriddle/women/nightpiechart.htm>.
- 12:14 Robbins, Naomi B. 2013. *Creating More Effective Graphs*, Chart House, Ramsey, NJ (reprinted from Wiley 2005). Copyright © 2013 by Chart House. All rights reserved.
- 15 Cleveland, William S. 1994. *The Elements of Graphing Data (Revised Edition)*, Hobart Press, Summit, NJ. Used with permission from William S. Cleveland.
- 16 Wurman, Saul. 1999. *Understanding USA*. TED Conferences, Inc., Newport, Rhode Island.
- 17 State of New Jersey, Division of State Police, Uniform Crime Reporting Unit. 1997. Uniform Crime Report: State of New Jersey 1997.
- 18 Robbins, Naomi B. 2013. *Creating More Effective Graphs*, Chart House, Ramsey, NJ (reprinted from Wiley 2005). Copyright © 2013 by Chart House. All rights reserved.
- 19 Wurman, Saul. 1999. *Understanding USA*. TED Conferences, Inc., Newport, Rhode Island.
- 20 Cleveland, William S. 1994. *The Elements of Graphing Data (Revised Edition)*, Hobart Press, Summit, NJ. Used with permission from William S. Cleveland.
- 21 Robbins, Naomi B. 2006. Copyright © 2006 by Naomi B. Robbins. All rights reserved.
- 22 Cleveland, William S. 1994. *The Elements of Graphing Data (Revised Edition)*, Hobart Press, Summit, NJ. Used with permission from William S. Cleveland.
- 23 Robbins, Naomi B. 2013. *Creating More Effective Graphs*, Chart House, Ramsey, NJ (reprinted from Wiley 2005). Copyright © 2013 by Chart House. All rights reserved.

Appendix 2: Resources on Color

Choosing colors:

<http://www.colorbrewer2.org>

Zeileis, Achim and Hornik, Kurt and Murrell, Paul. 2007. *Escaping RGBland: Selecting Colors for Statistical Graphics*. Research Report Series / Department of Statistics and Mathematics, 61. Department of Statistics and Mathematics, WU Vienna University of Economics and Business, Vienna. Available from <http://epub.wu.ac.at/1692/>

Color deficiencies:

<http://www.stcsig.org/usability/newsletter/9910-color-blindness.html>

http://www.lighthouse.org/color_contrast.htm

Test for color deficiency

<http://www.mppusers.com/ishihari.htm>

http://www.personal.psu.edu/cab38/ColorBrewer/Steve_Gardner_thesis_PSU.pdf

References

Arditi, Aries. 2005. "Effective Color Contrast: Designing for People with Partial Sight and Color Deficiencies". Lighthouse International. New York, NY.

http://www.lighthouse.org/color_contrast.htm

Becker, Richard and William S. Cleveland. 1996. *S-Plus Trellis Graphics User's Manual*. Mathsoft, Inc., Seattle and Bell Labs, Murray Hill, New Jersey.

Bertin, Jacques. 1973. *Semiologie Graphique*, 2nd Edition, Gauthier-Villars. (English translation: Bertin, J. 1983. *Semiology of Graphics*, University of Wisconsin Press, Madison, WI.

Carr, Daniel B. and Linda W. Pickle. 2010. *Visualizing Data Patterns with Micromaps*. CRC Press, Boca Raton, FL.

Carr, D. B. and S. M. Pierson. 1996. "Emphasizing Statistical Summaries and Showing Spatial Context with Micromaps," *Statistical Computing and Statistical Graphics Newsletter* 7:16-23.

Cleveland, William S. 1994. *The Elements of Graphing Data (Revised Edition)*, Hobart Press, Summit, NJ. (1st Edition, Wadsworth, Inc., Monterey, CA, 1985)

Cleveland, William S. 1993. *Visualizing Data*, Hobart Press, Summit, NJ.

Denby, Lorraine and Colin Mallows. 2009. "Variations on the Histogram." *Journal of Computational and Graphical Statistics*. 18(1): 21-31

- Few, Stephen. 2006. "Beautiful Evidence: A Journey through the Mind of Edward Tufte"
<http://www.b-eye-network.com/view/3226>
- Heiberger, R. M., and B. Holland. 2008. "Structured sets of graphs." In C. Chen, W. Hardle, & A. Unwin (Eds.), *Handbook of Data Visualization*. Springer, New York.
- Koschat, Martin. (2005). "A Case for Simple Tables," *The American Statistician* 59:1, 31-40.
- Playfair, William. 2005. *The Commercial and Political Atlas and Statistical Breviary*, Cambridge University Press. (Original editions published in 1786 and 1801.)
- Rensink, Ronald A. 2006. "Attention, Consciousness and Data Display," *JSM Proceedings*, Statistical Graphics Section. American Statistical Association, Alexandria, VA.
- Robbins, Naomi B. and Joyce Robbins. 2010. "Quantitative Literacy Across the Curriculum: Improving Graphs in College Textbooks",
http://www.perceptualedge.com/articles/visual_business_intelligence/quantitative_literacy_across_curriculum.pdf
- Robbins, Naomi B. 2013. *Creating More Effective Graphs*, Chart House, Ramsey, NJ (reprinted from Wiley 2005).
- Robbins, Naomi B. 2010 "Trellis display," *Wiley Interdisciplinary Reviews: Computational Statistics* 2:600-605. John Wiley and Sons, Hoboken, NJ.
- Robbins, Naomi B. 2008. "Introduction to Cycle Plots,"
http://www.perceptualedge.com/articles/guests/intro_to_cycle_plots.pdf
- Robbins, Naomi B. 2006. "Dot Plots: A Useful Alternative to Bar Charts,"
<http://www.b-eye-network.com/view/2468>
- Tufte, Edward. 2006. *Beautiful Evidence*, Graphics Press, Cheshire, CT.
- Tufte, Edward. 2001. *The Visual Display of Quantitative Information*, 2nd edition, Graphics Press, Cheshire, CT. (First edition 1983)
- Walkenbach, John. 2007. *Excel 2007 Charts*, Wiley, Hoboken, NJ.
- Ware, Colin. 2000. *Information Visualization: Perception for Design*. Morgan Kaufman, San Francisco.

About the Presenter:

Naomi B. Robbins is a consultant and seminar leader who specializes in the graphical display of data. She trains employees of corporations and organizations on the effective presentation of data. She also reviews documents and presentations for clients, suggesting improvements or alternative presentations as appropriate. She is the author of *Creating More Effective Graphs*, published by Chart House in 2013 (first published by John Wiley 2005) and blogged for Forbes at blogs.forbes.com/naomirobbsins. In addition to her one and two day seminars on creating more effective graphs, she offers short programs entitled "Recognizing Misleading and Deceptive Graphs" and "Just because you can doesn't mean you should: Better charts with Excel." Dr. Robbins received her Ph.D. in mathematical statistics from Columbia University, M.A. from Cornell University, and A.B. from Bryn Mawr College. She had a long career at Bell Laboratories before forming NBR, her consulting practice. Naomi was recently voted chair-elect of the Statistical Graphics Section of the American Statistical Association.

Speaker Contact Information

Naomi B. Robbins
NBR
11 Christine Court
Wayne, NJ 07470

Phone: (973) 694-6009
naomi@nbr-graphs.com
<http://www.nbr-graphs.com>
blogs.forbes.com/naomirobbsins
twitter: @nbrgraphs