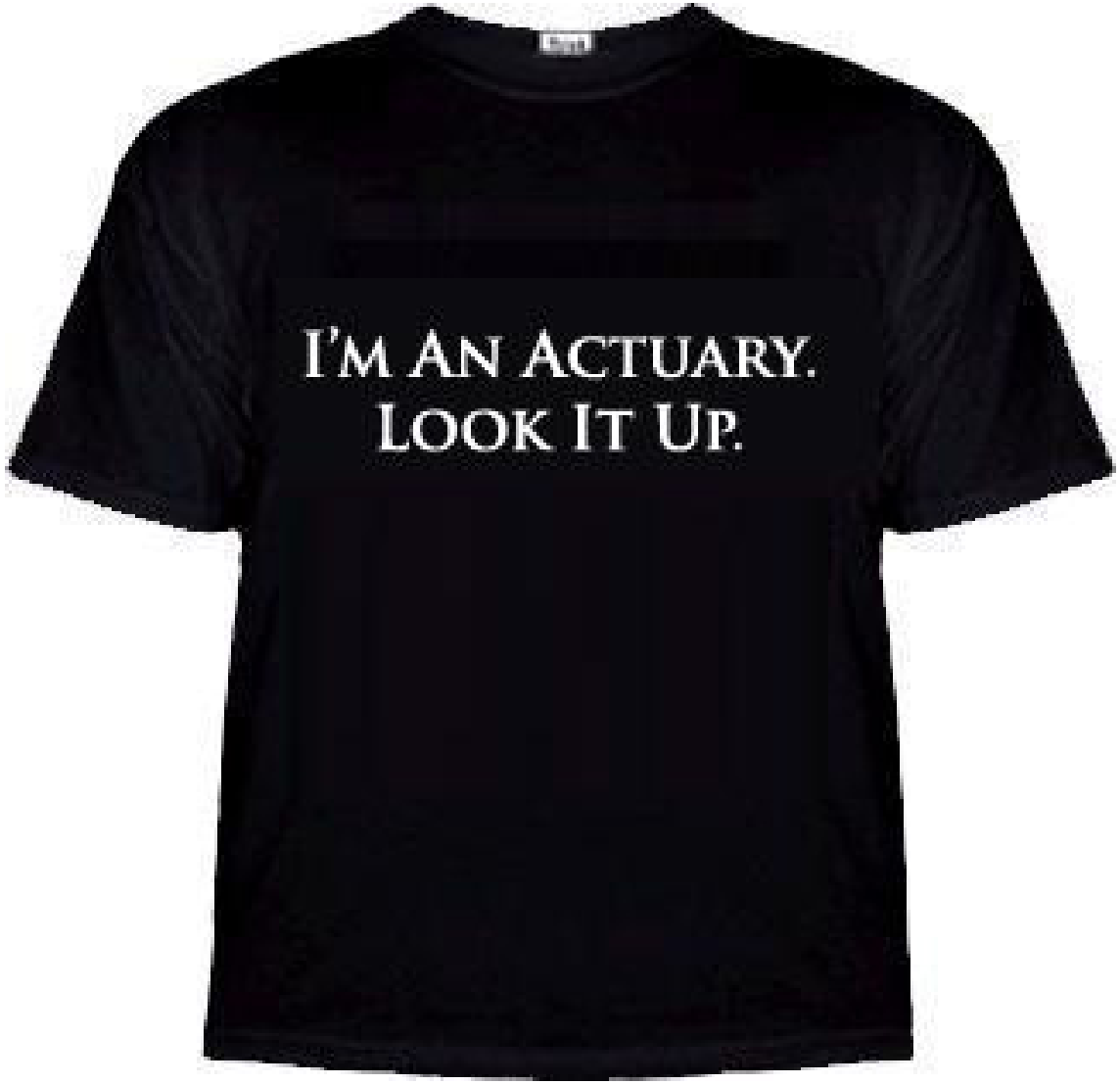
Designing Effective Graphs

Edward W. (Jed) Frees

University of Wisconsin – Madison

An Actuary Is ...



Outline

Resource Files

`http://research3.bus.wisc.edu/RegActuaries`

Goals

- This tutorial seeks to improve graphical practice. I intend to:
  - demonstrate the importance of graphical displays,
  - provide guidelines to improve graphical practice, and
  - introduce some of the scientific underpinnings of good graphical practice.
- Stated differently, the goal is to provide basic tools that viewers can use to become critical consumers and effective producers of graphs
- I also hope that viewers will adopt our enthusiasm and wish to explore the graphical design literature on your own

Basics
What is a “Graph”?

- Think of a “graph” as a pictorial representation of quantitative information
- Writers express quantitative ideas as
  - numbers within paragraphs,
  - numbers within tabular forms,
  - functional relationships such as equations, and
  - data or equations as graphs.
- Graphs should convey information
  - For some purposes, graphs are designed to attract attention or to entertain
  - So, some of the recommendations that described here are not what you will see on television on in the popular press – this is okay, our purposes differ.

Power of Graphs

- Graphs are a simple yet powerful medium for written communication of quantitative ideas. Graphs can
  - present a large amount of data in a small space,
  - express important relationships between quantities,
  - compare different sets of data, and
  - describe data, thus providing a coherent picture of complex systems.
- Graphs do more than merely state an idea; they demonstrate it

Graphs Require Structure

- Graphs are powerful because they are flexible, but flexibility can be a disadvantage because of the potential for abuse
- Well-accepted references dealing with methods of quantitative data presentation mitigate opportunities for abuse
  - The *Chicago Manual of Style* (1993), a standard reference, discusses presentation of in-text data, and
  - Ehrenberg (1977) and Tufte (1983) discuss presentation of tabular data.
- In contrast, we focus on data presentation through *graphical* displays

Structure Through Vigorous Writing Principles

- Principles of vigorous writing can and should be applied to the practice of making effective graphs
- The *Elements of Style* (Strunk and White 1979) summarizes vigorous writing:

Strunk

Vigorous writing is concise. A sentence should contain no unnecessary words, a paragraph no unnecessary sentences, for the same reason that a drawing should have no unnecessary lines and a machine no unnecessary parts. This requires not that the writer make all his sentences short, or that he avoid all detail and treat his subjects only in outline, but that every word tell.

White calls it “a short, valuable essay on the nature and beauty of brevity–sixty-three words that could change the world.”

Structure Through Effective Writing Principles

- Brevity is especially important when making effective graphs
  - This was also understood by Strunk - he said “a drawing should contain no unnecessary lines . . . ”
  - We use the term *chartjunk*, introduced by Tufte (1983), for any unnecessary appendage in a graph
- Just as with writing, effective graphs are the result of repeated revising and editing
- Poorly designed graphs can and do hide information and mislead
- Fancy or pretentious graphs are distracting when simpler graphs suffice

Structure Through Communication Principles

- Writing is processed in a serial manner, word by word, sentence by sentence, with a beginning and an ending
- The process of "reading," or *decoding*, a graph is nonlinear and more complex
- In graphics, the communication process plays a dominant role

Process of Communicating with a Graph

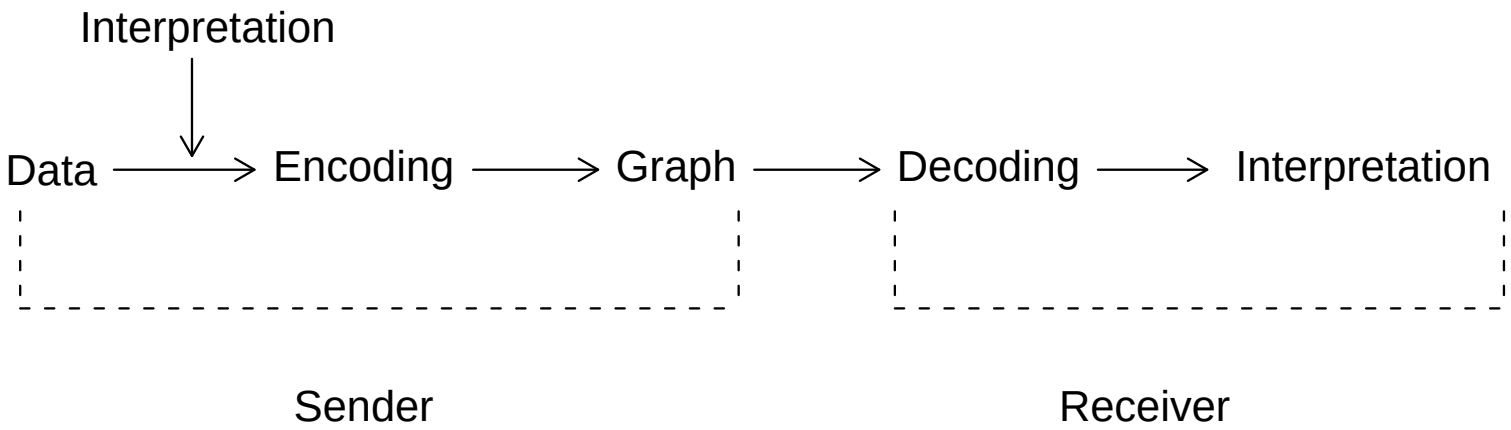


Figure Flow Chart of the Process of Communicating with a Graph. The graph is a crucial intermediary in the process of communicating data interpretation to the receiver.

Communication Process

- The receiver is party to neither the exact interpretation intended by the sender nor the raw data
- The receiver must decode the *graph* and develop an interpretation of its message
- Two issues arise:
  - Whether the interpretation constructed by the receiver is congruent to the interpretation of the sender
  - Whether the receiver's interpretation is consistent with and supported by the data.
- There is substantial room for miscommunication

Summary
Principles of Reading and Writing Graphs

- These principles are important to anyone who deals extensively with quantitative information
- Graphs are powerful – they do more than merely state an idea; they demonstrate it
- Graphs are more complex than text - they are vulnerable to many sources of abuse
- Just as with writing, designing effective graphs requires structure
  - Analogies to writing principles
  - Through graphical perception principles
- In the next video, we show that graphic design choices make a difference

Graphic Design Choices Make a Difference

- Graphs are more complex than text - they are vulnerable to many sources of abuse
- The tutorial presents examples to illustrate how to “lie” with data
- The examples illustrate that graphic design choice makes a difference
  - We have not included gross misrepresentations, such as you might find in Tufte’s *The Visual Display of Quantitative Information* or Wainer’s *Visual Revelations*
  - We seek to show more subtle differences that could result from genuine misinterpretation of data

Graphic Design Choices Make a Difference Examples

- Including Zero to Compress Data
- Perception of Correlation
- Transforming to Logarithmic Scale
- Double Y-Axes
- Aspect Ratio

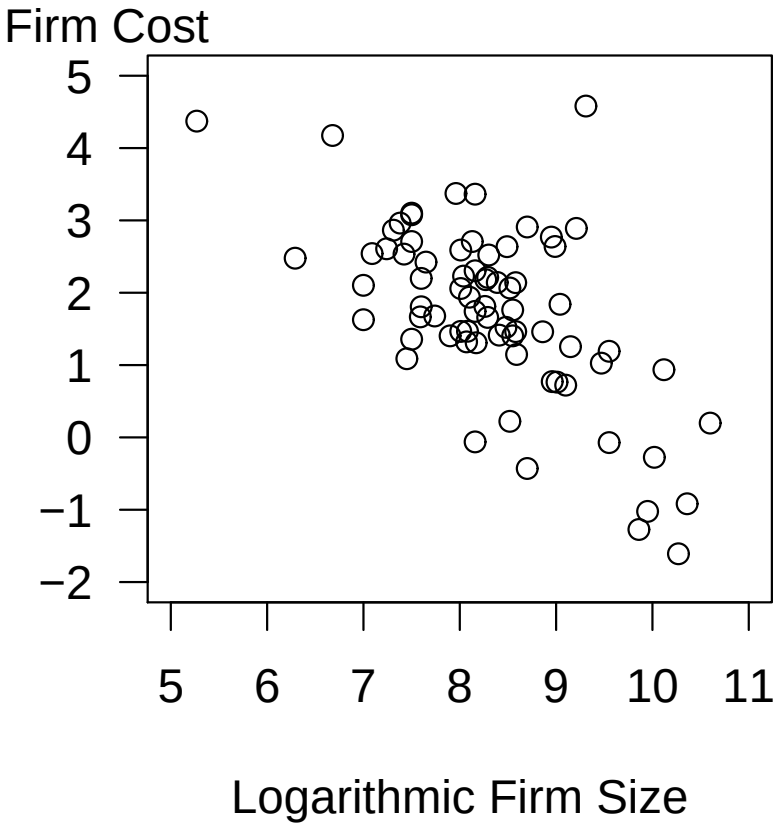
Including Zero to Compress Data

(a) *A stable insurance industry*

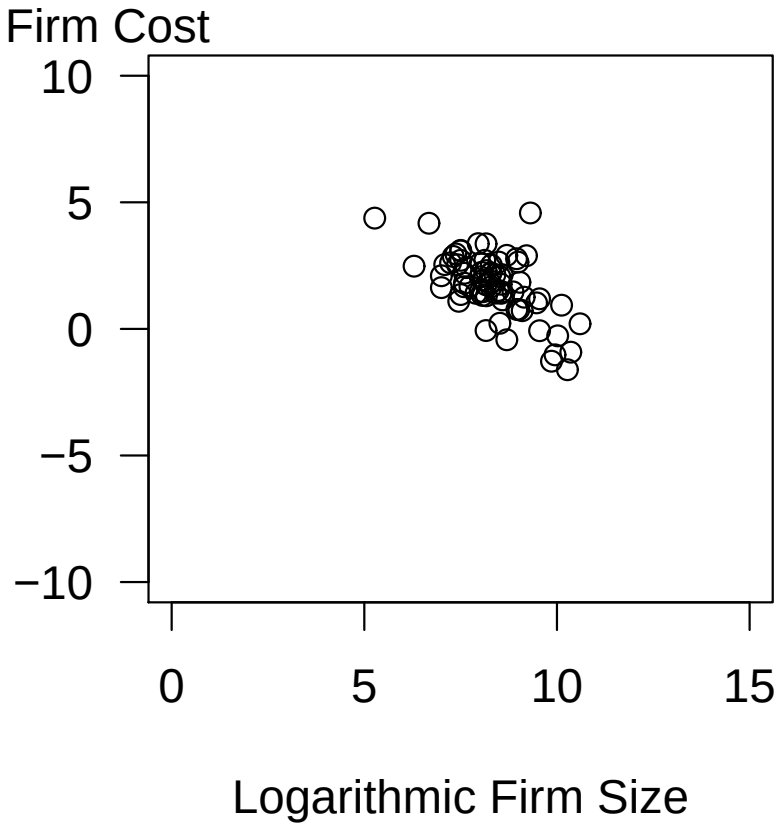
(b) *The insurance industry work force increased dramatically in the 1950s*

Figure Annual Insurance Employees, 1948-1993.

Perception of Correlation



(a) *The data in this figure appear less correlated*



(b) *The data in this figure appear more correlated*

Figure Cost Effectiveness of a Firm's Risk Management Practices Versus Firm Size.

Perception of Correlation 2

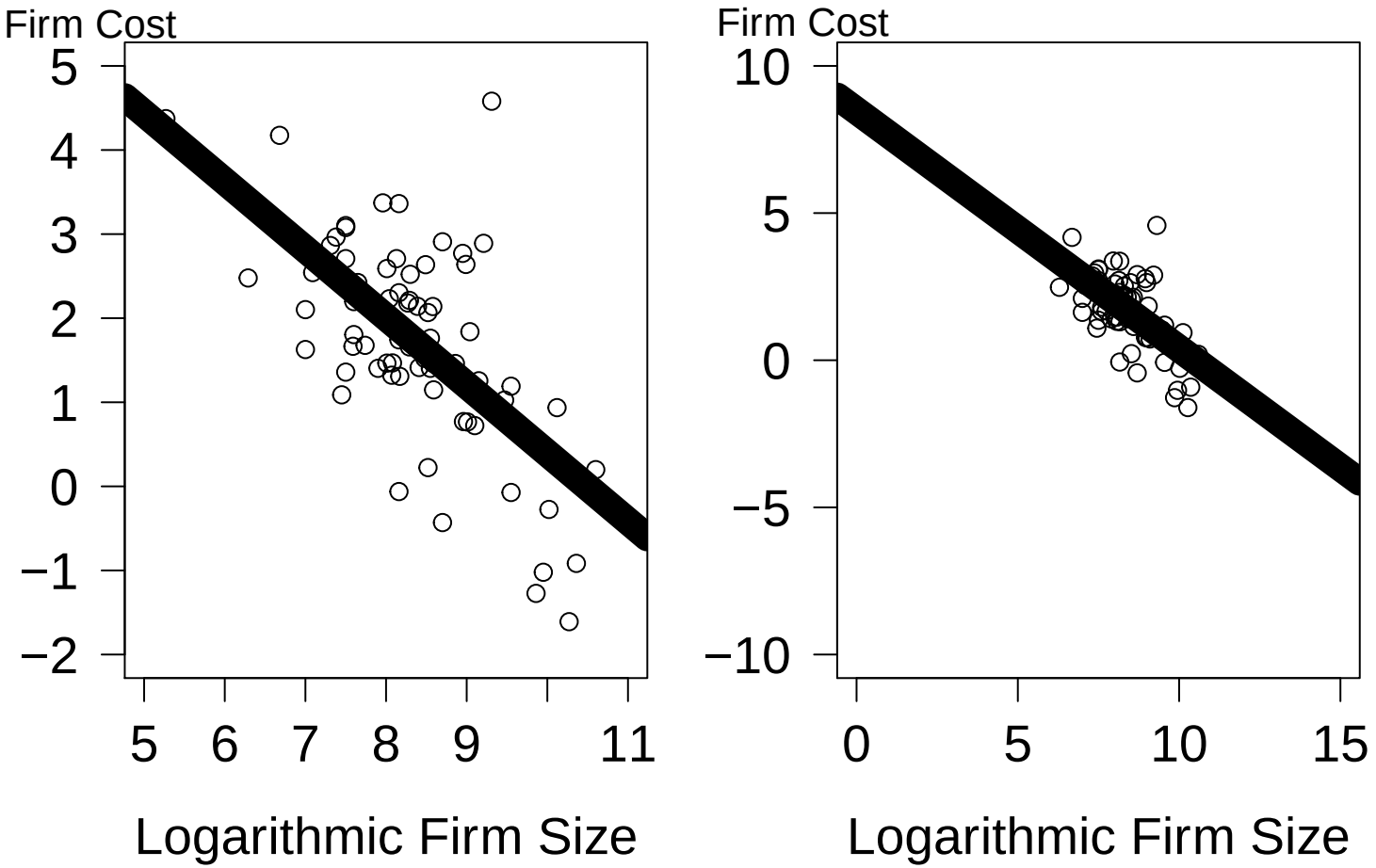
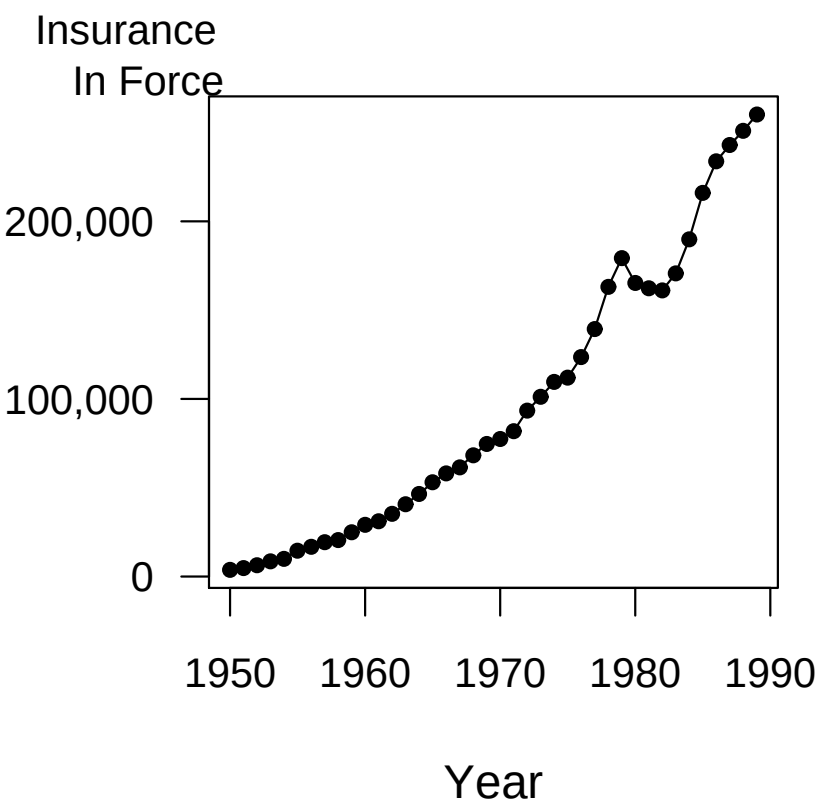
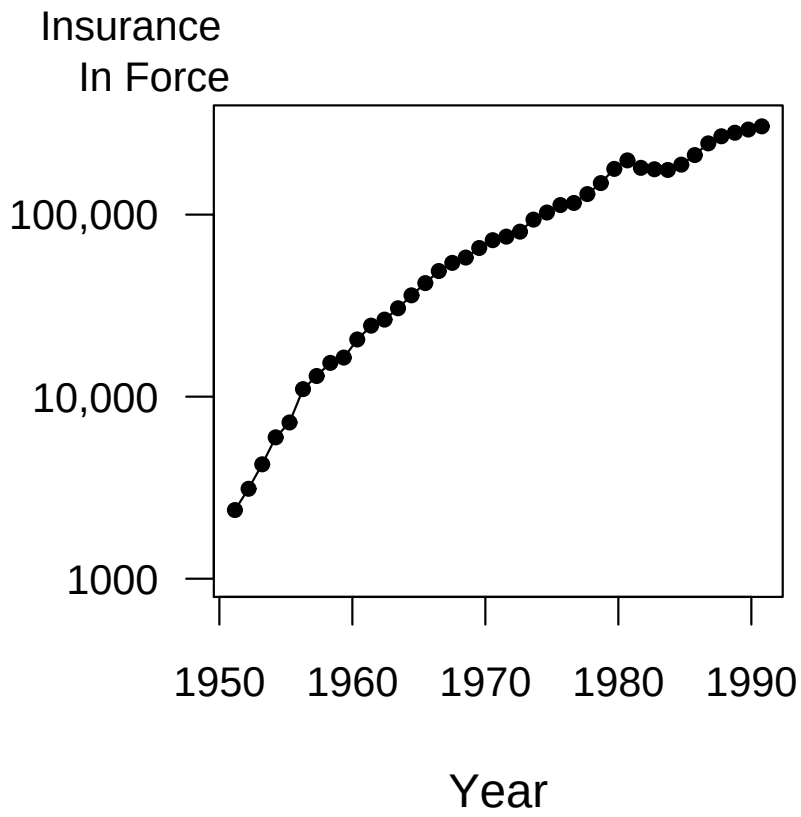


Figure Cost Effectiveness of a Firm's Risk Management Practices Versus Firm Size.

Transforming to Logarithmic Scale



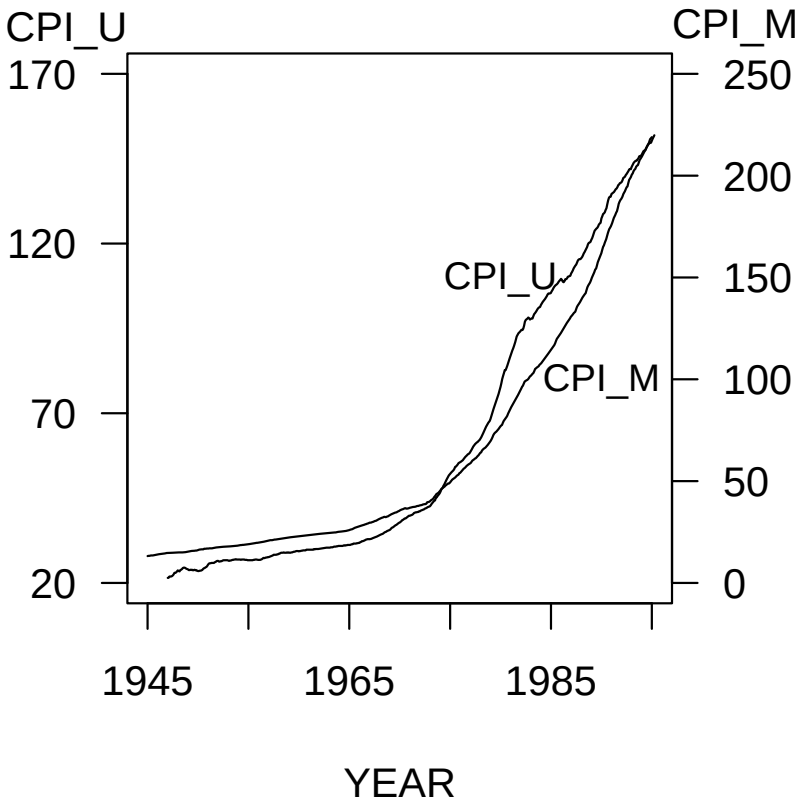
(a) *U.S. credit life insurance market exploding*



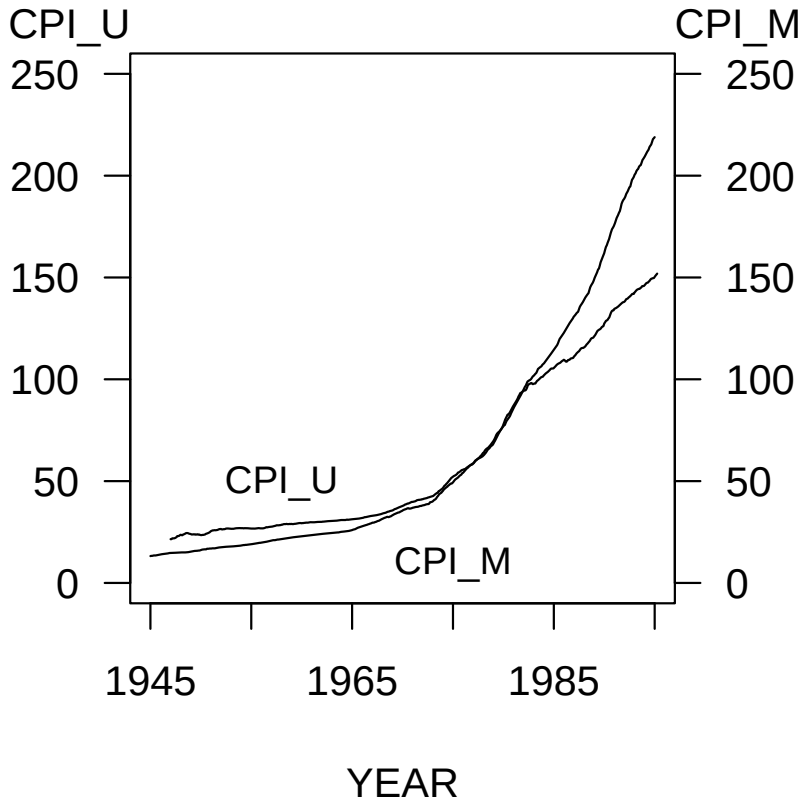
(b) *U.S. credit life insurance market leveling off*

Figure Annual U.S. Credit Life Insurance in Force, 1950-1989.

Double Y-Axes



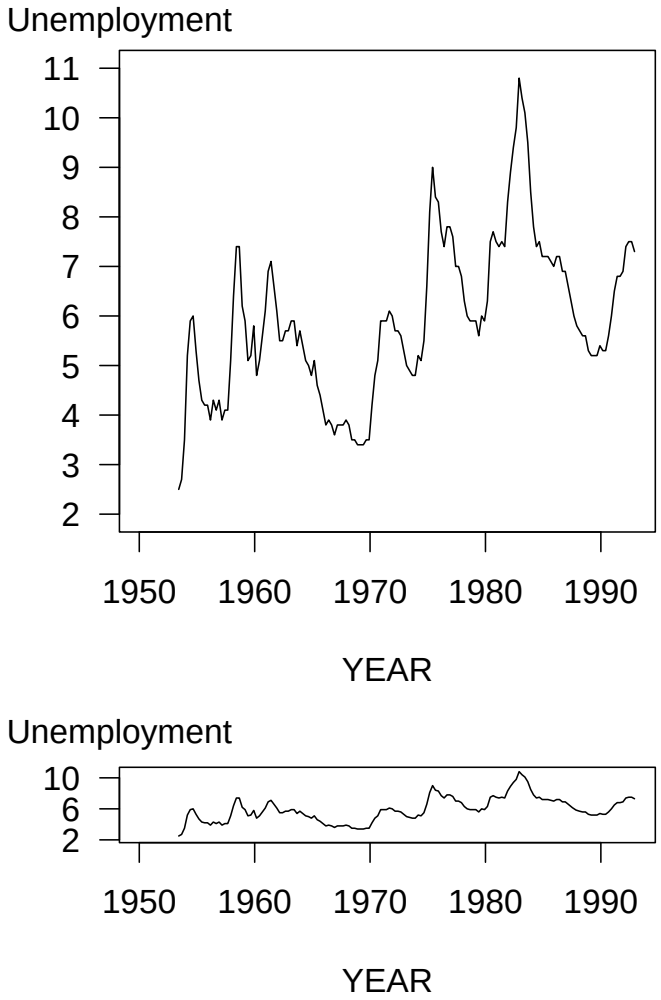
(a) *Overall Consumer Price Index (CPI) is similar to the medical component of the CPI*



(b) *Overall Consumer Price Index (CPI) is increasing more slowly than the medical component of the CPI*

Figure Monthly Values of the Overall Consumer Price Index (CPI) and the Medical Component of the CPI, January 1947 through April 1995.

Aspect Ratio



FigureTime Series Plot of Quarterly Values of the U.S. Unemployment Rate, 1953-1992. The lower panel displays a feature that is not evident in the upper panel; unemployment declines more slowly than it rises.

Summary
Graphic Design Choices Make a Difference

- Graphs are more complex than text - they are vulnerable to many sources of abuse
- The examples illustrate how to “lie” with data – even without gross misrepresentations
  - Including Zero to Compress Data
  - Perception of Correlation
  - Transforming to Logarithmic Scale
  - Double Y-Axes
  - Aspect Ratio
- Graphic design choice makes a difference
- In the next video, we will provide guidelines to enable you to
  - recognize these errors when viewing graphs and
  - avoid these errors when creating graphs.

Design Guidelines
Tips for Making Effective Graphs

Effective Writing and Graphical Perception

Avoid chartjunk
Use small multiples to promote comparisons and assess change

Design Guidelines
Tips for Making Effective Graphs

Effective Writing and Graphical Perception

Avoid chartjunk
Use small multiples to promote comparisons and assess change

Graphical Perception

Use complex graphs to portray complex patterns
Relate graph size to information content
Use graphical forms that promote comparisons

Design Guidelines
Tips for Making Effective Graphs

Effective Writing and Graphical Perception

Avoid chartjunk
Use small multiples to promote comparisons and assess change

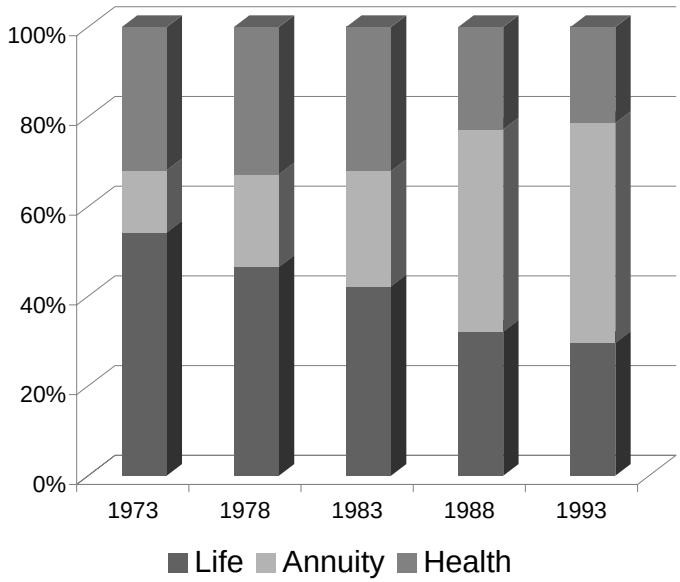
Graphical Perception

Use complex graphs to portray complex patterns
Relate graph size to information content
Use graphical forms that promote comparisons

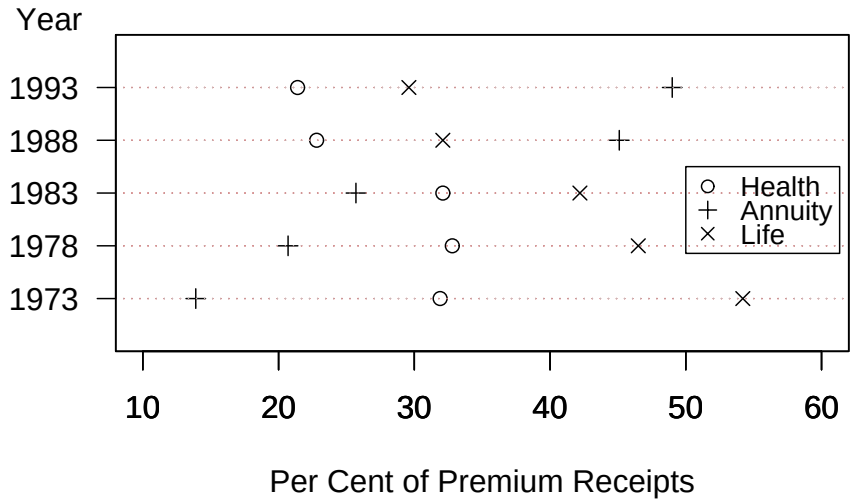
Effective Writing

Integrate graphs and text
Demonstrate an important message
Know your audience

Avoid Chartjunk



(a) *The three-dimensional stacked bar chart is a poor graphical form for making comparisons over time and across lines of business.*



(b) *The dot plot allows for direct comparison over time and across lines of business.*

Figure Distribution of Premium Receipts, 1973-1993. The excessive chartjunk of (a) hides the large change in distribution types between 1983 and 1988.

Small Multiples

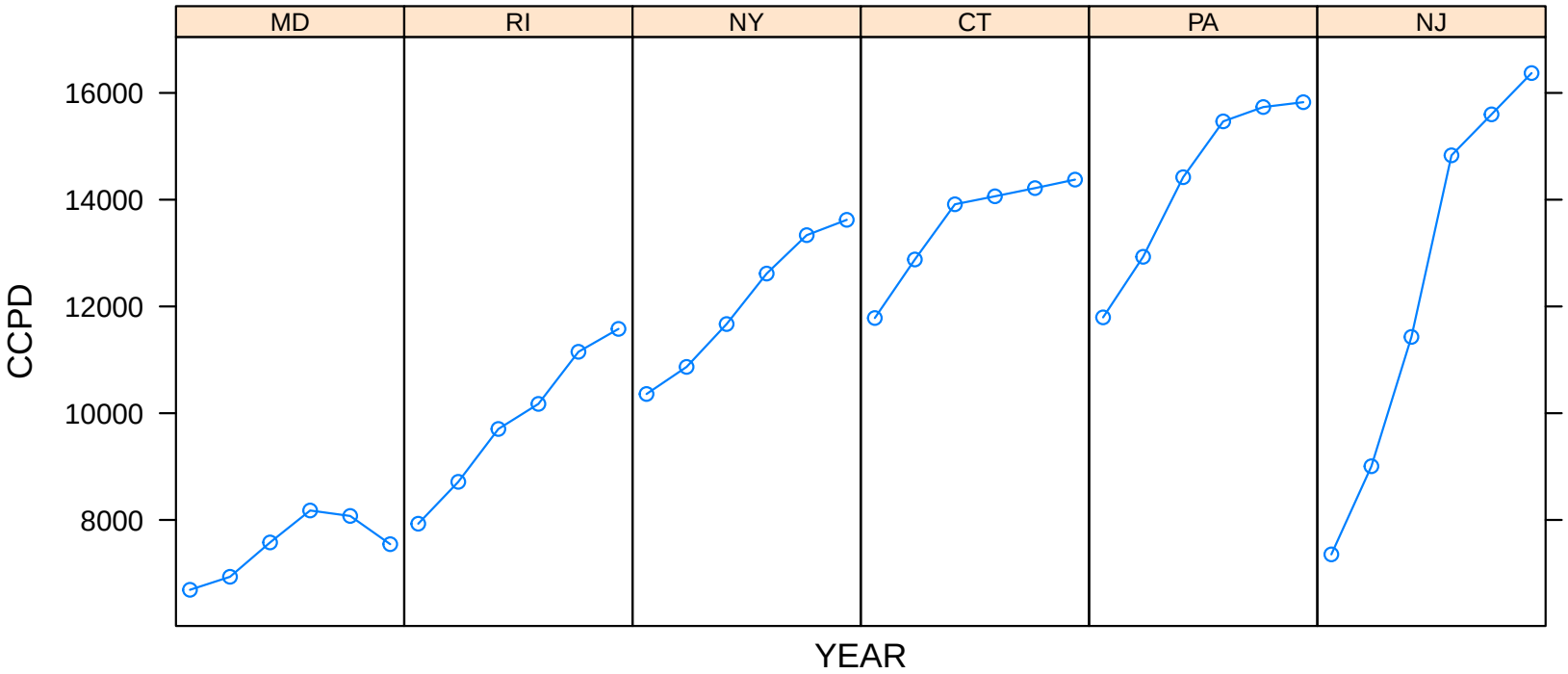
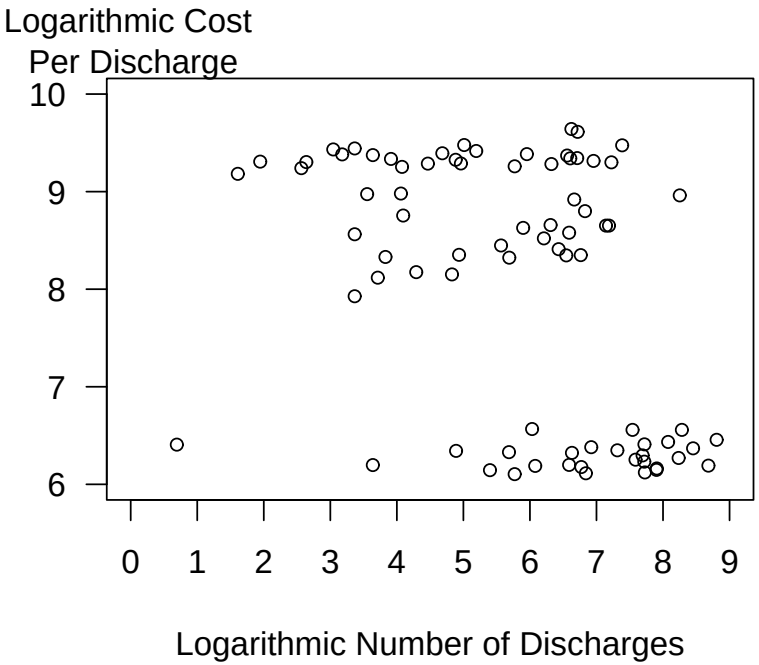
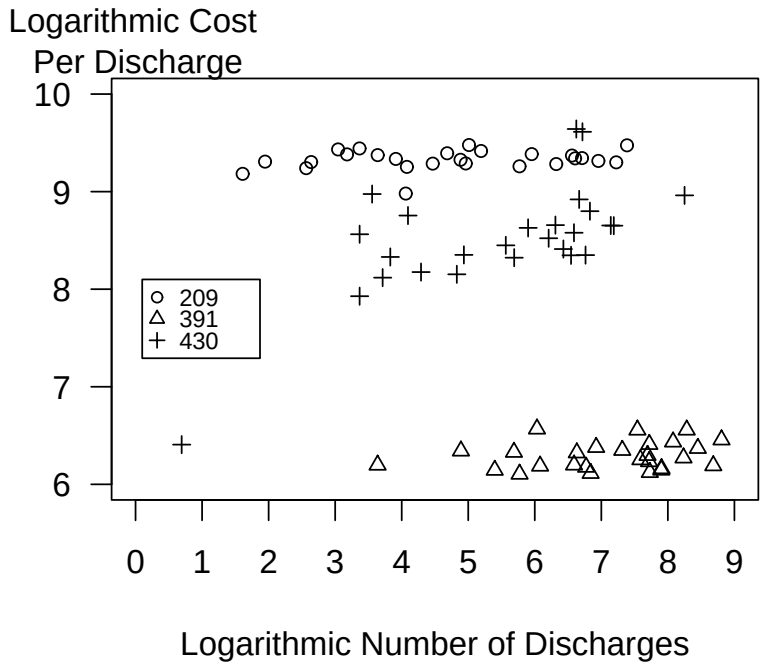


Figure Trellis Plot of Medicare Claims over 1990-1995 for Six States.

Use Complex Graphs to Portray Complex Patterns



(a) *With the exception of one outlying observation in the lower left-hand region, there appears to be a significant negative relationship between cost and number of hospital discharges.*

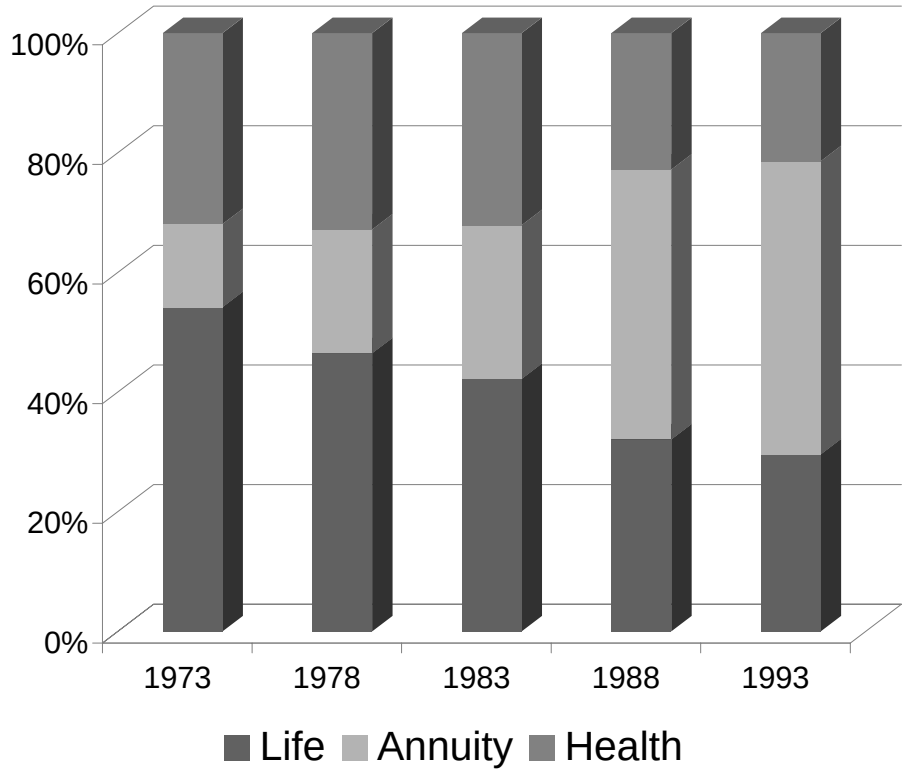


(b) *By introducing the DRG codes, we see a small positive relationship between cost and number of hospital discharges within each group.*

Figure Logarithmic Cost per Discharge Versus the Logarithmic Number of Discharges.

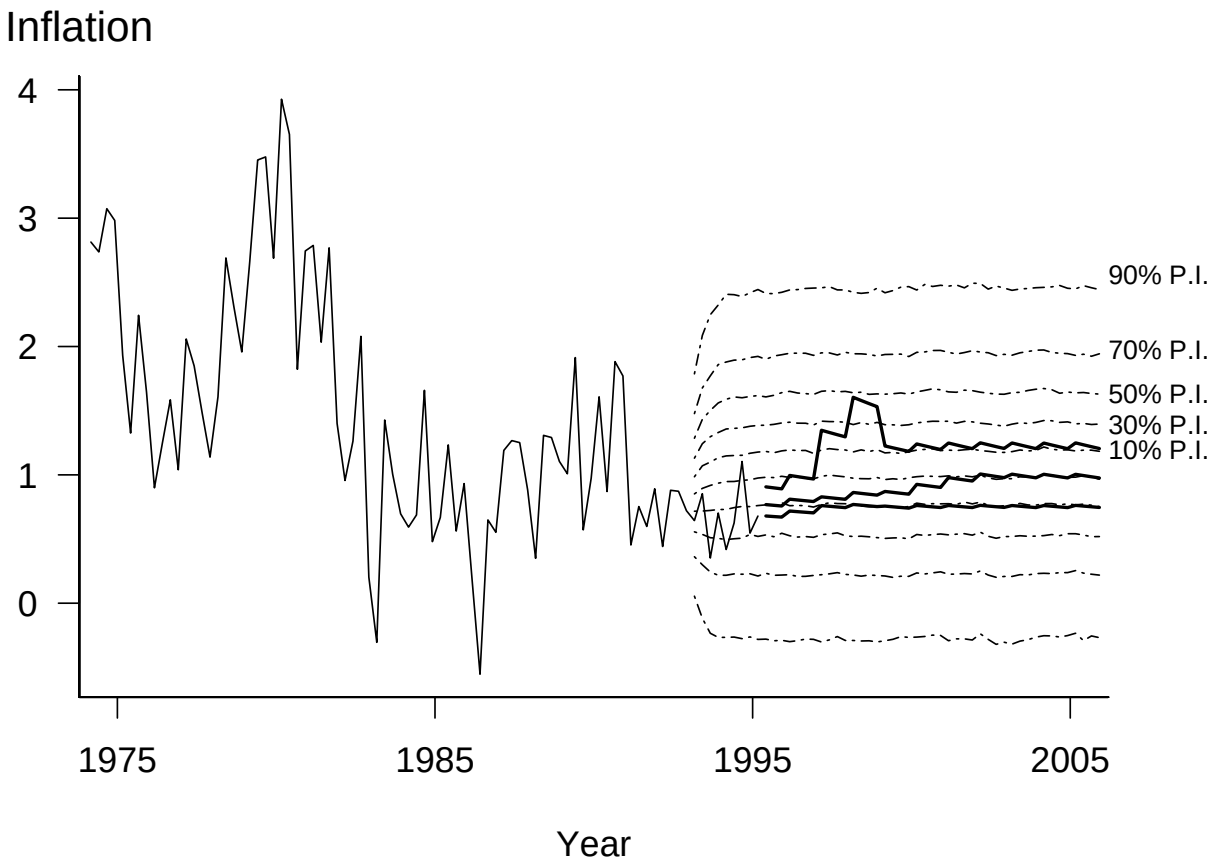
Relate Graph Size to Information Content

- A graph should be proportional to the amount of information that it contains.
- The *data density of a graph* is defined to be the number of data entries per unit area of the graph.



FigureDistribution of Premium Receipts, 1973-1993. Representing only 15 data points, this graph has a low data density.

Relate Graph Size to Information Content



Comparison of Stochastic Prediction Intervals to Held-out Actual Experience and to Social Security's Assumptions. The thin solid lines represent actual inflation rates, and the thick solid lines represent projections by Social Security experts. The dotted lines represent prediction intervals generated by a stochastic time series model. This complex graph allows viewers to make comparisons based on approximately 600 points.

Other Design Guidelines I

- Use graphical forms that promote comparisons - to be discussed as part of the following “empirical foundations” video
- Integrate graphs and text
  - Although “a picture is worth ten thousand words,” a graph needs supporting text
  - Data graphics can be complemented by a tabular presentation of data
    - graphics can highlight relationships among the data, and
    - tables can present precise numerical descriptions of the data.

Other Design Guidelines II

- Demonstrate an important message
  - Detailed legends and graphs should reinforce messages that are developed in the main body of the text
  - Put major conclusions in a graphical form
  - “The greatest value of a picture is when it forces us to notice what we never expected to see.” Tukey (1977)
- Know your audience
  - A basic precept of effective writing, familiarity with one’s audience, is also valid for designing effective graphs
  - You will sometimes prefer a *less* effective form based on the level of audience familiarity
    - We hope that practice will eventually shift from these ineffective modes of communication
    - Creators of graphs should not swim against the current of poor graphic design but rather bend their path towards more effective modes of communication

Summary
Design Guidelines

Effective Writing and Graphical Perception

Avoid chartjunk
Use small multiples to promote comparisons and assess change

Graphical Perception

Use complex graphs to portray complex patterns
Relate graph size to information content
Use graphical forms that promote comparisons

Effective Writing

Integrate graphs and text
Demonstrate an important message
Know your audience

Summary
Design Guidelines

- In the next video, we will discuss foundations upon which these guidelines are based

Empirical Foundations For Guidelines

- This section discusses empirical studies from cognitive psychology and surveys of practice
- We have provided demonstrations on the importance of graphical design and guidelines for improving practice
- Many of the guidelines are based on the science of perception - cognitive psychology.
- “Intuition” is something we learn and cultivate
  - Progress does not always conform to current intuition
  - It was widely believed at one time that the earth was flat and that the sun revolved about the earth
  - The demonstrations of this section may or may not be immediately intuitive but they are logical conclusions from the design guidelines advocated here

Science of Perception

- Return to our communication process figure and think about:
  - How do we process information?
  - How do viewers perceive graphs?

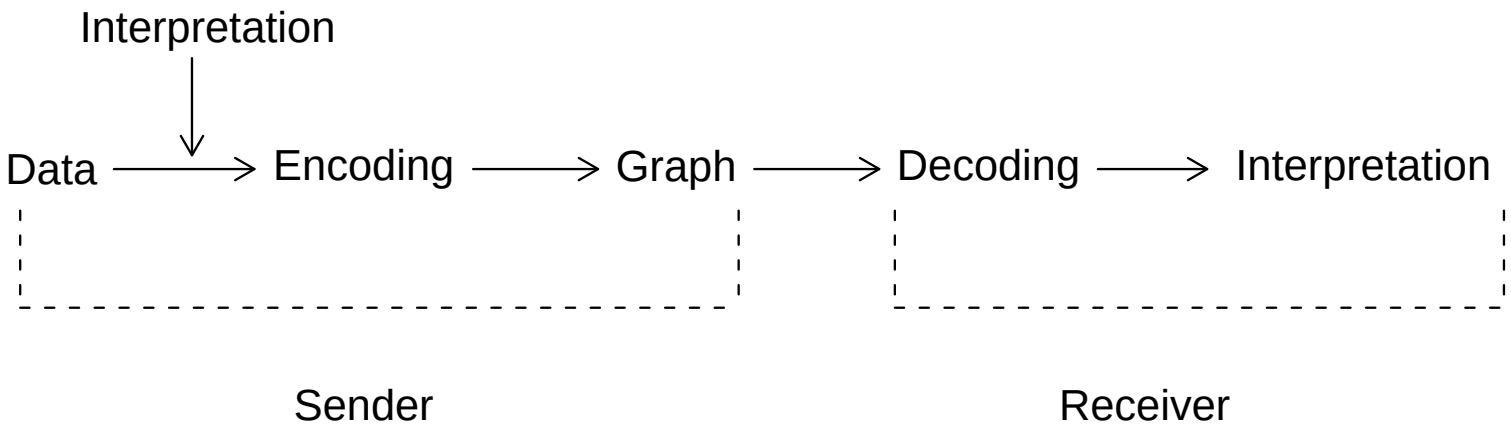
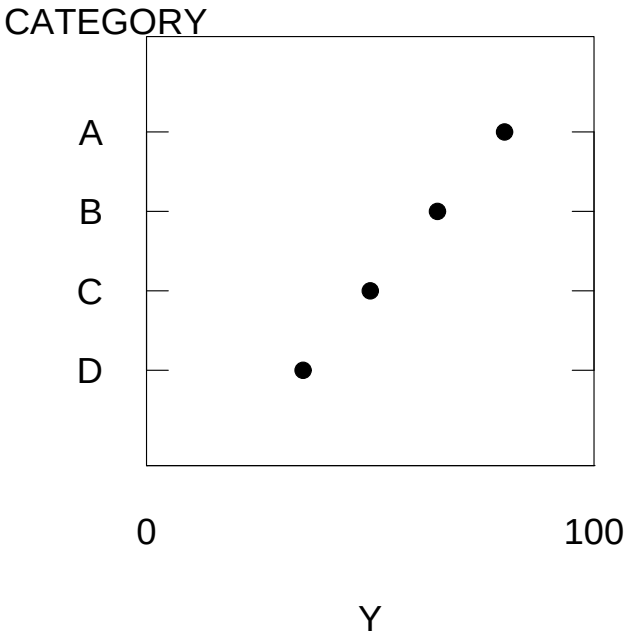


Figure Flow Chart of the Process of Communicating with a Graph

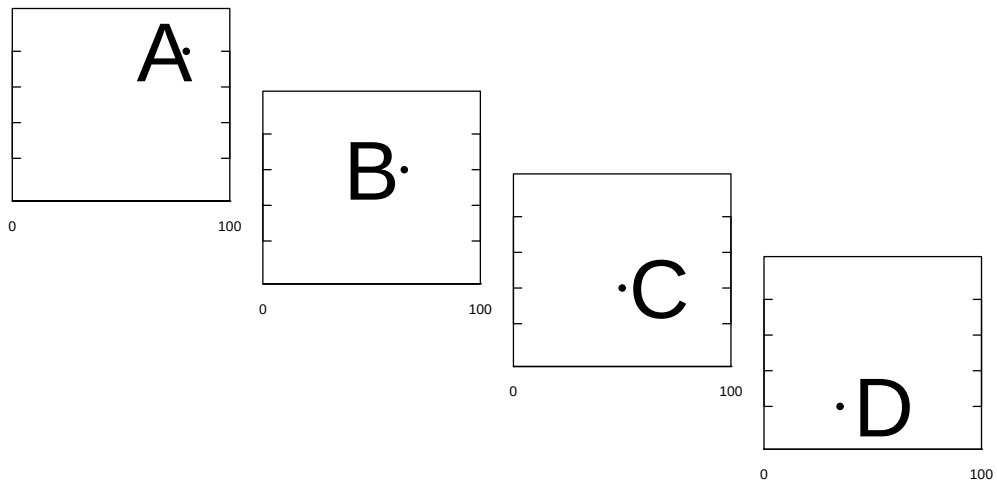
Viewers as Units of Study

- Basic graphical perception tasks (Cleveland, 1994)
  1. Position along a common scale
  2. Position along identical, nonaligned scales
  3. Length
  4. Angles and slopes
  5. Area
  6. Volume
  7. Color and density
- This is an ordered list, so that “position along a common scale” is relatively easy to do compared to “volume” comparisons
- Cleveland and McGill (1984) performed a series of experiments to understand their relative difficulty

Graphical Perception Experiment I



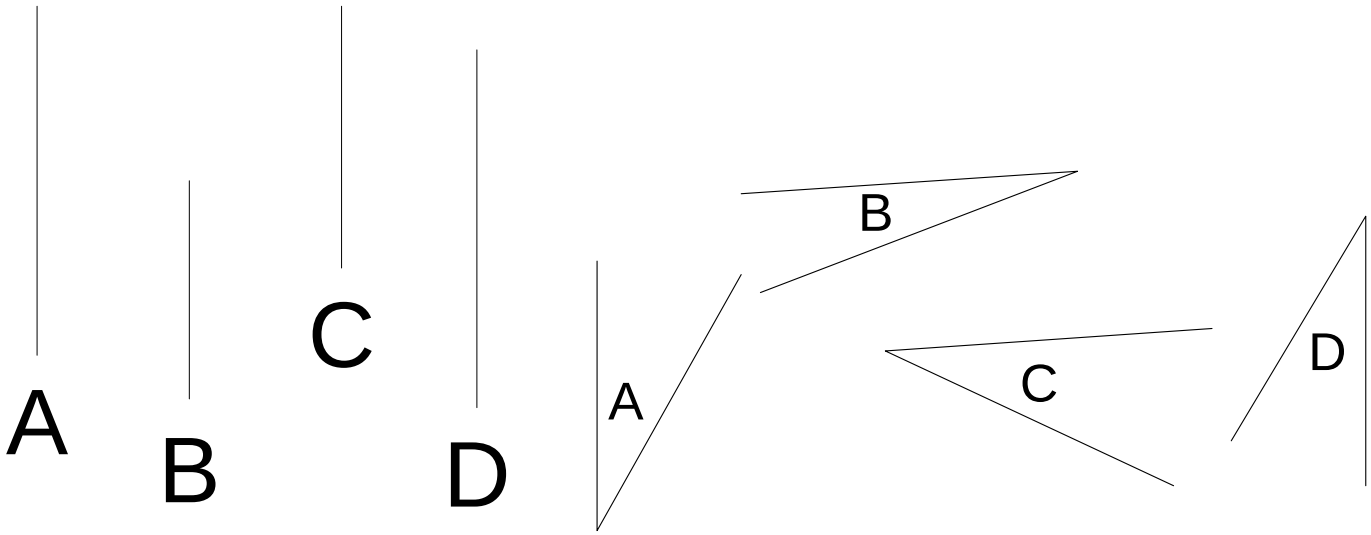
(a) *Experiment to Judge Position along a Common Scale. Assess the relative values of A, B, C and D along this 100-point scale.*



(b) *Experiment to Judge Position along Identical, Nonaligned Scales. Assess the relative values of A, B, C and D on a common 100-point scale.*

Figure Experiments in Judgments about Graphical Perception

Graphical Perception Experiment II

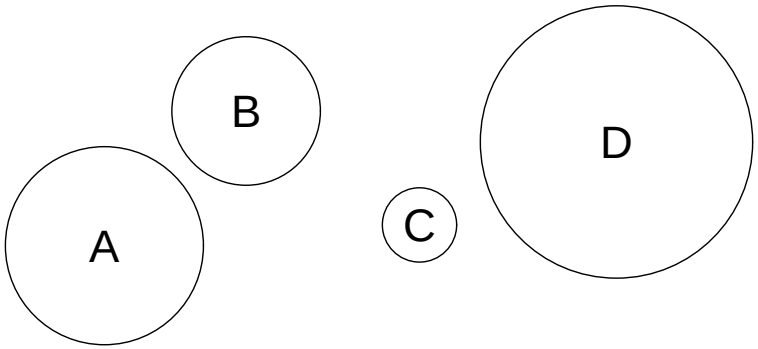


(a) *Experiment to Understand Length Judgments. Suppose line A is 100 units long. Assess the relative lengths of lines B, C and D.*

(b) *Experiment to Understand Angle Judgments. Suppose angle A is 100 units. Assess the relative values of angles B, C and D.*

Figure Experiments in Judgments about Graphical Perception

Graphical Perception Experiment III



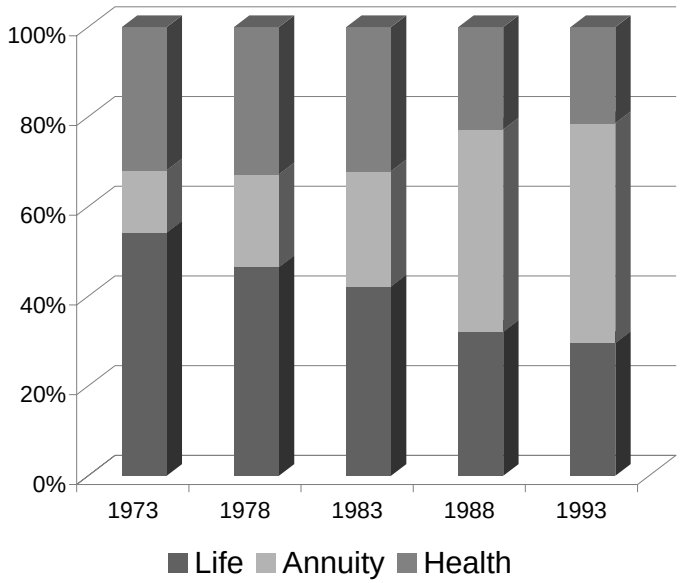
(a) *Experiment to Understand Area Judgments. Suppose circle A has area 100 units. Assess the relative areas of circles B, C and D.*

Figure Experiments in Judgments about Graphical Perception

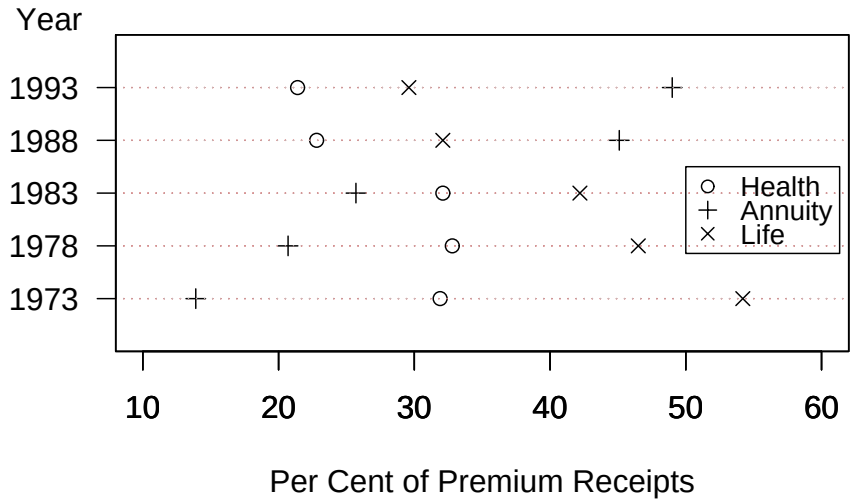
Basic Graphical Perception Tasks

- Be wary of asking viewers of your graphs to make informed judgements about angles, slopes, area, volume, color, and density

Position along Identical, Nonaligned scales



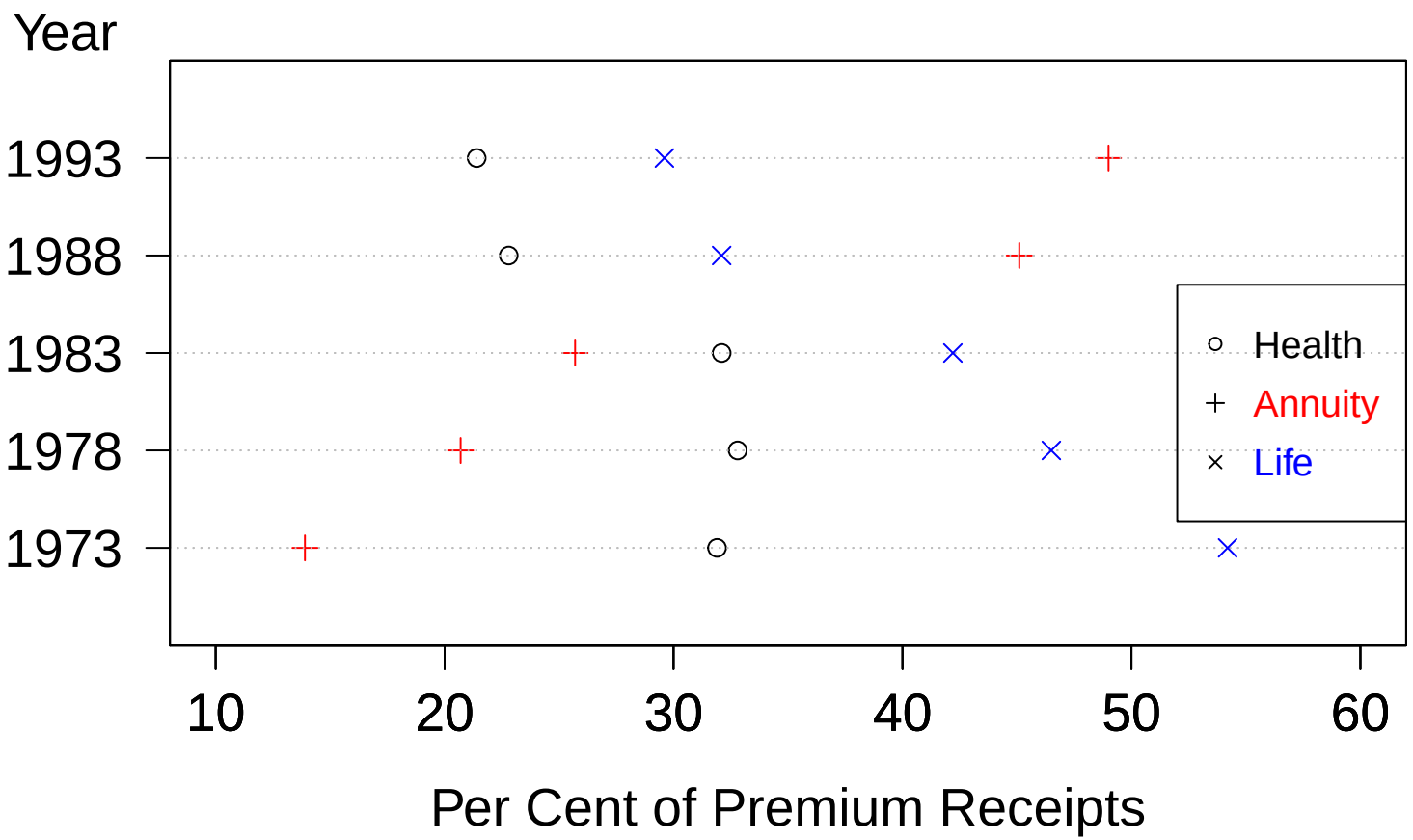
(a) *The stacked bar chart asks you to make comparisons over nonaligned scales.*



(b) *The dot plot allows for direct comparison over a common scale.*

Figure Distribution of Premium Receipts, 1973-1993.

Distribution of Premium Receipts



FigureDistribution of Premium Receipts, 1973-1993. The dot plot allows for direct comparison of positions over a common scale. The use of colors helps to distinguish among types of insurance products.

Basic Graphical Perception Tasks

- Be wary of asking viewers of your graphs to make informed judgements about angles, slopes, area, volume, color, and density
- Creators of scientific (serious) graphs generally avoid pie charts
- The following example shows a poor graphical form - think about how you might improve it using the principles described here

Three-Dimensional Pie Chart
A Poor Graphical Form

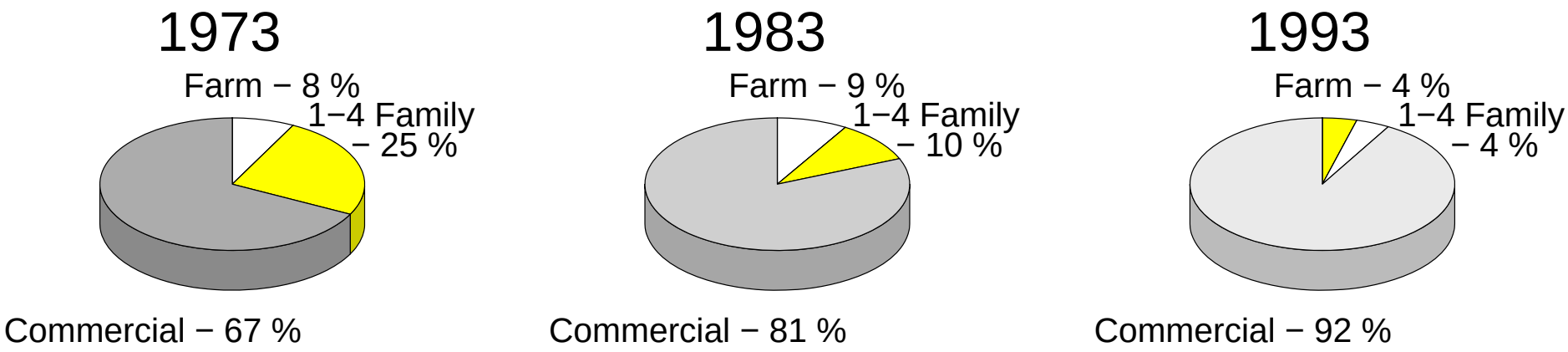


Figure Distribution of Mortgages for the Years 1973, 1983 and 1993. The three-dimensional pie chart is a poor graphical form for making comparisons over time and across types of mortgages.

An Improvement

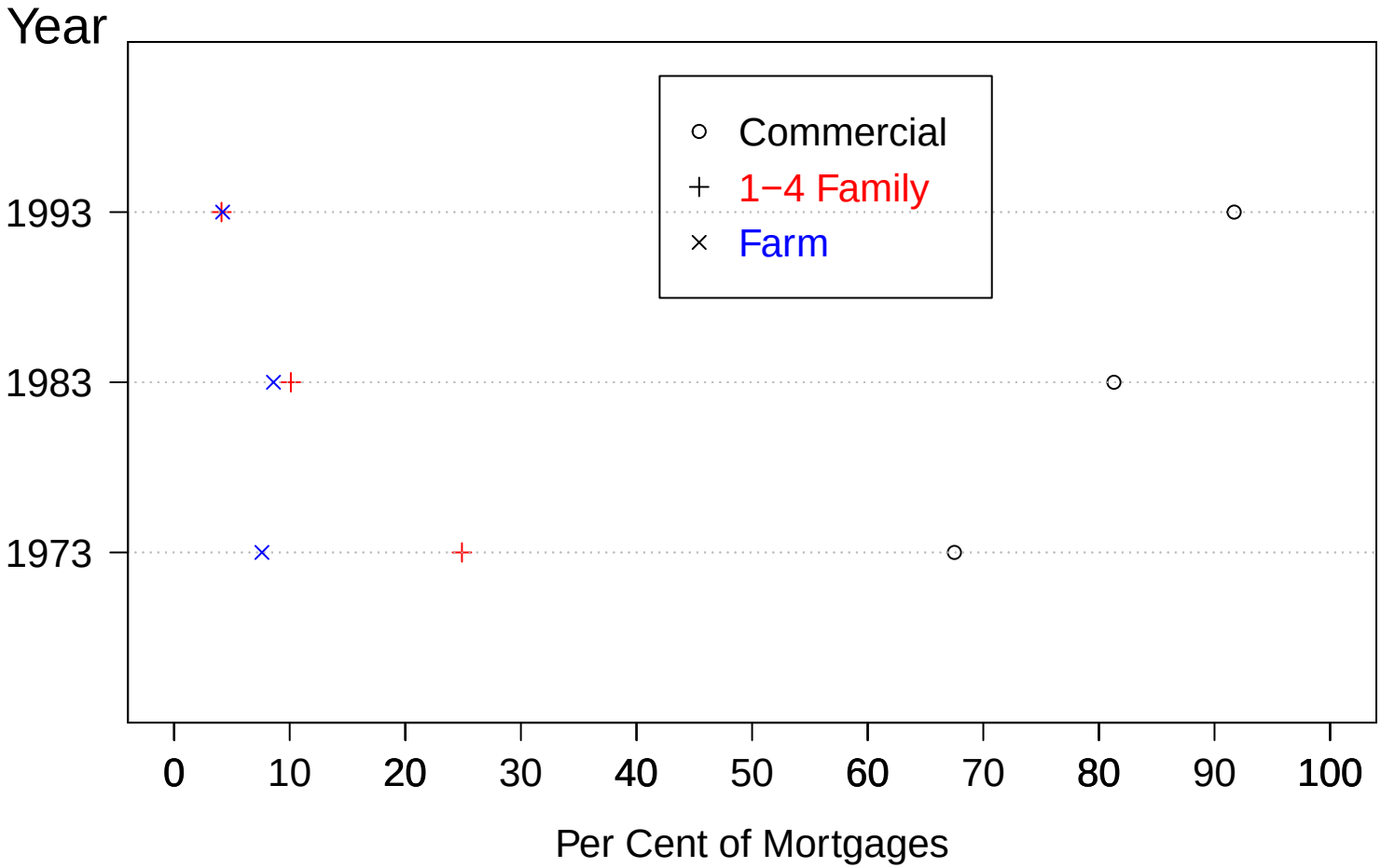


Figure Commercial, 1- to 4-Family, and Farm Mortgages as Percentages of Total Mortgages for 1973, 1983 and 1993. A negative aspect of this graph is the overlap of the 1- to 4-family and farm plotting symbols in 1983 and 1993.

A Table is the Best Choice

Table Commercial, 1- to 4-Family, and Farm Mortgages as Percentages of Total Mortgages for 1973, 1983, 1993

	Year		
Mortgage Type	1973	1983	1993
Commercial	67.5	81.3	91.7
1-4 Family	24.9	10.1	4.1
Farm	7.6	8.6	4.2

Graphs as Units of Study

- Earlier, we noted that “intuition” is something to learn and cultivate
- You can develop your intuition by conducting, or taking part of, experiments where you judge the difficulty of assessing quantitative information graphically
- The Cleveland and McGill is one such empirical study
- You can also develop your intuition by surveying graphical practice in your field of study – see how well creators of graphs are doing at conveying information efficiently!

Surveys of Graphical Practice

- Surveys of graphical practice in professional publications provide an important database with which to assess prevalence of good and bad practice and changes in practice over time
- Tufte (1983) discusses a survey of approximately 4,000 graphs randomly selected from 15 news publications for the years 1974 to 1980
  - The graphs were assessed for "sophistication," defined as presentation of relationship between variables
- Cleveland and McGill (1985) report a similar survey of scientific publications, assessing the prevalence of graphical errors.
- Frees and Miller (1998) report on a survey of factors affecting graphic quality in life insurance publications
- Conduct a survey in your own field of interest!

Summary
Empirical Foundations For Guidelines

- Guidelines require foundations upon which they can be based
- The science of graphical perception helps to identify tasks that are comparatively easy for viewers to decode
  - Less effort decoding means more time and energy for interpretation
  - As a science, theories of graphical perception can be validated by empirical experiments
- You can also gather empirical data to learn about the norms of graphical behavior in your field of study and improve those norms

Summary I

- The first video was on principles of reading and writing graphs
  - Emphasized the power and flexibility of graphs
  - Introduced structure by analogies to writing principles and through graphical perception principles
- The second video emphasized the importance of graphical design through a series of examples
  - “Lies, damned lies, and statistics” is just too easy
  - The examples illustrated how to influence thoughts with data – *without* gross misrepresentations

Summary II

- The third video provided specific guidelines for creating effective graphs
  - Most powerful were guidelines based on both writing principles and graphical perception
  - These were avoiding chartjunk and use of small multiples
  - Guidelines based on graphical perception consisted of:
    Using complex graphs to portray complex patterns, relating graph size to information content, and using graphical forms that promote comparisons
    Guidelines based on effective writing principles consisted of:
    Integrating graphs and text, demonstrating an important message, and knowing your audience.
- In the fourth video, we recommended that you look to empirical studies to develop your sense as to how you wish to convey numerical (empirical) information with graphs

Conclusions

- Just as we need to be taught writing, we need to learn to communicate complex numerical information graphically.
- This starts with our publications and journals, seeps into the classroom and spreads throughout practice.
- With modern-day computing available, it is all too easy to become passive consumers of the defaults programmed by software developers.
- We need to become effective creators and critical consumers of graphical displays!

Further Reading and References
Cleveland and Tufte

- Cleveland, William S. (1994). *The Elements of Graphing Data*. Monterey, Calif.: Wadsworth.
- Cleveland, William S. (1993). *Visualizing Data*. Summit, N.J.: Hobart Press.
- Cleveland, William S., Diaconis, P., and McGill. R. (1982). Variables on scatterplots look more highly correlated when the scales are increased. *Science* 216, 1138-1141.
- Cleveland, William S., and McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association* 79, 531-454.
- Cleveland, William S., and McGill, R. (1985). Graphical perception and graphical methods for analyzing and presenting scientific data. *Science* 229, 828-833.
- Tufte, Edward R. (1983). T*he Visual Display of Quantitative Information*. Cheshire, Conn.: Graphics Press.
- Tufte, Edward R. (1990). *Envisioning Information*. Cheshire, Conn.: Graphics Press.
- Tufte, Edward R. (1997). *Visual Explanations*. Cheshire, Conn.: Graphics Press.

Further Reading and References
Others

- Ehrenberg, A.S.C. (1977). Rudiments of Numeracy. *Journal of the Royal Statistical Society A* 140:277-97.
- Huff, D. (1954). *How To Lie with Statistics*. New York: Norton.
- Schmid, C.F. (1992). *Statistical Graphics: Design Principles and Practices* Malabar, Fla.: Krieger Publishing Co.
- Strunk, W., and White, E.B. (1979). *The Elements of Style*. 3rd ed. New York: Macmillan.
- Wainer, H. (1997). *Visual Revelations*. Springer, New York.
- University of Chicago Press (1993). *The Chicago Manual of Style*. 14th ed. Chicago, Ill.

More References

- Frees, Edward W. and Robert B. Miller (1998). Designing effective graphs. *North American Actuarial Journal* 2(2), 53-76.
- Frees, Edward W. (2010). *Regression Modeling with Actuarial and Financial Applications*. Cambridge University Press, New York.
  - To make the graphs from this presentation, statistical (*R*) code and data are available at `http://research.bus.wisc.edu/RegActuaries`.

The End

Thanks for Watching!