# Assignment 2

Marissa

2/3/2020

## Chapter 1: How to Graph Badly or What NOT to Do

### 1.1 Introduction

- Best learned by studying how to make poor, uninformative graphics
- Skill lies in not making mistakes

### 1.2 Chartjunk

- Chartjunk: extraneous features that add nothing to the information content of a graph
- Psychological studies have shown that people are worse at estimating areas than lengths

#### 1.2.1 Fonts

- Different fonts are good:
  - Distinctive and applied to different elements of the graph
  - Look good together
- Danger is illegibility, not ugliness

#### 1.2.2 Moire Shading

- Shading is very attention-getting
  - Eyes are drawn to the shaded areas

#### 1.2.3 Pseudo 3-D

- Plots into the third dimension can be really useful
  - Only when necessary (not artistic liberty)

#### 1.2.4 Artificial Color

- Color is powerful, but easy to misuse
- Color is expensive in storage, to print
- Use color, but wisely!
  - Use when to make clear

#### 1.2.5 Hype, I: Overinterpretation

- Over-interpreting the data
  - Experimental noise may overwhelm the signal

#### 1.2.6 Hype, II: Graphical Carpet Bombing

- Presenting a lot of complicated information graphically:
  - Varying the graph type(- Mix contour plots, surface plots, line graphs, etc
  - Highlight the key figures in the caption and graphic itself
  - Triage the analysis
    - Spend a lot of words and graphs on the important points( - Few words and possibly no graphs on the unimportant ones
  - Combine many closely-related graphs into a single multi-panel graph

### 1.3 Wainer's Rules for Bad Graphs

- Show as little data as possible [minimize the data density]
- Hide what data you do show [minimize the ratio of data/ink]
- Show the data inaccurately [ignore the visual metaphor and randomize the connection between graphical elements and the numbers]
- Use length as the visual metaphor when the area of 2-D icons is what is actually perceived
- Graph data out of context [sparse captions and vague text]
- Obfuscation #1: Change scales in mid-axis
- Obfuscation #2: Emphasize the trivial [ignore the important]
- Obfuscation #3: Jiggle the baseline [use different axis ranges for two graphs which will be printed side-by-side and need to be compared]
- Obfuscation #4: Order the data by some criterion, such as alphabetical order, which is irrelevant to all of the interesting patterns in the data
- Obfuscation #5: Label (illegibly, incompletely, incorrectly. ambiguously)
- Obfuscation #6: More is murkier (more decimal places, more dimensions)
- Think of a new way to do it

### 1.4 High Data Density

- High density illustrations are good
  - Possible to pack a tremendous amount of information in a single picture if designed carefully
- Low density graph used to emphasize information or concepts that are important

### 1.5 Data-Hiding

- Really bad graph: adding so many extraneous elements that data/information is obscured
- Really bad graph: removing data points and just having a line

#### 1.5.1 Data-Hiding by Graphing Disparate Quantities on the Same Scale

### 1.6 Inconsistent Visual Metaphor

- Visual metaphor: relationship between specific graphical elements and the data

### 1.7 Context-Free Data

- Context of graph: engineering or physics background
- Context remarks:
  - Graph is a failure of the text and caption fail to give enough information to have a comprehensible graph
  - Much of science is comparisons
    - Graph will fail through lack of context if the curves fail to make important comparisons
  - Most important characteristics of good graph:
    - To show enough curves, contain enough information

### 1.8 Area Instead of Length as a Visual Metaphor

- Newspaper artists love dearly to turn simple graphs into works of art

### 1.9 Label Woes

- Illegible labels:
  - Too small type size
    - Reproduction during publishing
  - Poor placement
  - Too few labels
    - Too many labels: more time reading labels than curve
    - Couple labels does not distract from curve

### 1.10 Emphasize the Unimportant

- Emphasize the essential elements

### 1.11 Unnecessary Graphic Novelty  When choosing graphics:

- What is familiar to the audience?
- What will they find comprehensible?

## Chapter 2: The Gospel According to Tufte

### 2.1 Data-Ink

- Definition of Data-Ink
  - The non-erasable core of a graphic
  - Data-Ink Ratio
    - Data-Ink / Total ink use to print the graphic
- Tufte's Five Laws of Data-Ink:
  - Above all else show the data
  - Maximize the data-ink ratio
  - Erase non-data-ink
  - Erase redundant data-ink
  - Revise and Edit

#### 2.1.1 Show the Data

- Context is a mixture of good graphs and explanations that orient the reader like the compass rose on a map

#### 2.1.2 Emphasize the Data

- All graphics require some non-data elements:
  - Axis lines, tic marks, labels, etc.

#### 2.1.3 Erase Non-Data-Ink or Down with the Grid!

- Use of gridlines before the computer age due to hand-drawn graphics Exception to use of grid in graphics:
  - Nomogram: graphical calculator When author expects reader to carefully study curve and determine local maxima and minima
    - Qualitative nomography
  - Three-Dimensional Diagrams
    - Grid carries a message
    - Pierre Welander: the grid was to facilitate the conversation fo the data curve into numerical values
- Key guidelines for a grid:
  - Don't use a grid unless you really have to
  - Make the grid line faint compared to the data-curves
    - Grid as thin as lines or dotted lines
    - Use thick line for the data

#### 2.1.4 Erase Non-Data-Ink: Hurrah for Half-Framing!

- Half-Framing: drawing only the usual horizontal and vertical axes and omitting framing lines on the top and right

#### 2.1.5 Erase Non-Data-Ink: Example of the Simplified Bar Graph

- Bar chart to illustrate that excess can be erased in a chart without compromising the data

#### 2.1.6 Erase Redundant Data-Ink: Symmetry and Wrap-Around and All That

- Redundancy is sometimes needed for complicated visualization
  - Redundancy is good depending on context and readership

#### 2.1.7 Erasing: Eliminating the Graph Entirely

- Sometimes to deal with a flawed graph, eliminate the illustration entirely, and use a table instead

#### 2.1.8 Revise and Edit

Intellectual content is not changed by editing, only the clarity * 1. Experiment: try several versions of the same figure * 2. Redundancy is a matter of subtle judgement( + Redundancy should be there for a reason

## 2.2 High Data Density

- Data Density = Number of entries in data array / Area of data graphic
- High Density Graphs: identify key themes or goals of the figure

#### 2.2.1 The Shrink Principle

- Drink Principle: Graphics can be shrunk way down
- It is much easier to grasp a set of related figures if they are all on one place in a single multi-panel figure rather than scattered over many pages

## 2.3 Multifunctioning Graphical Elements

- Mobilize every graphical element, perhaps several times over, to show the data

## 2.4 Small Multiples or Animations-on-a-Page

- Animations: individual grams must differ by a small amount
- Small Multiple Figure: collection of miniature illustrations, arrayed as a single figure, to be perceived as one
  - Each multi-panel figure must by the same size, same graph species, and all other aspects of the design
- Animations:
  - Shows only six different times
  - Each time is illustrated in two different ways
    - Mesh plot
    - Contour graph
- Interpolate: Morphing if there are too few frames to make it smooth and pleasing to the eye
- Small multiple graph <—> one animation with morphing between frames

## 2.5 One Plus One is Three

- Questions to ask:
  - What to include?
  - How do these part interact?, What should be emphasized?, Can data curves be easily picked out?

## 2.6 Layering, Separation and Rubrication

- Emphasis of separation through color or grayscale

## 2.7 Word-Labels Are Better Than Letter-Labels

- Make labels as clear and explicit on the graph
- Write out labels as whole words or numbers

## 2.8 Collapsing a Dimension or Escaping Flatland

- Ten dimensions, but compacted and have no visible roles
- Experiment: Plot the data in several different forms, and publish an illuminating subset of the graphs

## 2.9 Supplementary Material

### 2.9.1 Small Multiples or Animations-on-a-Page

- All frames are shown in single row with identical format for easy comparison
- Important features are marked with arrows and numerical labels

### 2.9.2 Separation: Inset Graphs

- Use of large graph/inset graph visual format ties two graphs together

## 2.10 Wide is Wonderful: Aesthetics of Aspect Ratio

- External Aspect Ratio: ratio fo width to height as it appears on the printed page
  - Re = width on page / height on page
- Reasons:
  - Human visual system
  - Wider is better

## 2.11 Color or Why the Rainbow Isn't Golden

- Color is most powerful but also easier to misuse
- 5% - 10% red-green color blindness
  - Two-color scheme should avoid red and green
  - Rainbow pallet of colors will create problems for color-blind

## 2.12 Parallelism

- Multiple images are combine in parallel, message is easier to grasp because axes, format are constant

## 2.13 The Friendly Graphic

- Friendly:
  - Words spelled out
  - Words run left to right
  - Little messages help explain data
  - Labels are placed on graphic; avoid legends
  - Graphics attract viewer
  - Colors are strategic and avoid color-blindness
  - Type is clear
  - Type is upper and lower case