Topics in Information Display IMAGES DELETED

CSE 510
Richard Anderson
Ken Fishkin

My Homework

- "When did CHI become a research focus"?
  - SIGCHI – 1983
  - UIST - 1986

Administrivia

- "If you can only attend 1 lecture this quarter".....
  - Wednesday will be the "project gallery"
- **No** extra class 1/17
  - Class as scheduled 3/3

Why Study This?

- How data is organized and presented has a dramatic effect on how the data is conceptualized

Why Study This?

- "When principles of design replicate principles of thought, the act of arranging information becomes an act of insight"

Why Study This?

- *Your* arrangement affects *others* insights

Why Study This?

- "those who discover an explanation are often those who construct its representation"
- When you argue with data, it's better to have the tools to argue well

Why Study This?

- "More stunning is the observation that they *did* have the pertinent data. There were charts they did not image and did not construct that, if created, would have provided the quantitative correlational data required to sustain their position"

Why Study This?

- Most engineers have *no* training in the visual arts.
- But are almost always tasked to make the charts
  - "Collectively, they decided they would need 13 charts. Lund assigned the responsibility for the charts, dividing it among people based on their specializations"

Why study this?

- Like lawyers taking a course in public speaking
- You *will* be tasked to present your results
- It is very easy to make very bad presentations

Tufte to the Rescue!

- Yale Prof in Stats.
- Presenting data is essential
- Has written a series of *very* influential books on designing information displays

Limits of Tufte

- Ignores color-blindness (and visual impairment in general)
- Focus is on static displays – doesn't really address interaction, multimedia, etc.
- We will discuss some of these in later lectures

Art and Science

- Art: "nearly every escape from flatland demands extensive compromise, trading off one virtue against another"
  - Remember periodic table
- Science:"The principles of information design are universal - like mathematics - and are not tied to unique features of a particular language or culture"
  - -- perhaps not….

Digression

- Graphs are (relatively) new
- Cartesian coordinates: 1628
- 1st 1D plot: 1644

Digression

- 1st 2D plot
  - 1769
- Lambert (as in "Lambertian shading")

Digression

- 1st bar chart
  - 1786

Digression

- Playfair: 1st (a) pie chart, (b) multivariate, (c) 2D shape
  - 1801

PART 1: SHOW THE DATA, NOT THE WRAPPER

- Delete unnecessary elements ("Chartjunk")

"Above all else, show the data"

- Rule: "Nothing can be erased without losing information" (or harming understanding)

Occam's razor

- "the smallest effective difference is the Occam's razor of information design"
- "small differences allow *more* differences

Resist Moiré effects

- "Moiré effect is a visual perception that occurs when viewing a set of lines or dots that is superimposed on another set of lines or dots, where the sets differ in relative size, angle, or spacing.
- Generally, "vibration"/"noise"/"bleeding" from nearby graphic elements
- Very easy w/office software
- "noise clouds the flow of information"
- (http://www.mathematik.com/Moire/)

More Moire

- Don't use cross-hatching

More Moire

- Always use dark text on light, solid backgrounds
- WIRED effect
- "it causes 'visual vibration', eye fatigue, and inaccurate reading".
- Well-known in cognitive psych.

Color: hue

- Avoid "visual puzzles", where "the graphic must be interpreted through a verbal rather than a visual process"

Color:hue

- Why? "the mind's eye does not readily give a visual ordering to colors, except possibly for red to reflect higher levels"

Color: hue

- Hue is *cognitively* a circle
- Mapping it to a line means that cognitively *adjacent* colors get "thrown" to opposite ends
- E.g. red and violet

Color:hue

- Where graph A "splits the circle" may not be where graph B does.
- Green is lowest in left picture, middle in right

Color: hue

- Once hue is used for an *ordered* quantity, what's left for unordered ones?

What do you think?

Color:hue

- "rainbow color-coding of quantitative data confounds what happens in a color scheme with what happens in the data"

Color:hue

- Use lightness or saturation for a 1D ordering

Color:hue

- Interpolate through white to interpolate 2 hues

Color: hue

- Use *small spots of intense, saturated color*
- Grey better than white for background

Digression

- "Use small spots of intense, saturated color" also a technique used in cinema to highlight
- Converse (graying out) also used to make things recede
- Might be useful when visualizing uncertainty (later lecture)

Digression

- A common "quasi-hue" ordering is
  - Red: bad
  - Green: good
  - Yellow/Orange: middle
- "quasi-hue" because also appeals to traffic light metaphor
- Issue: doesn't work for 10% males of European descent (red-green color-blind)

Color:hue: summary

- Only use hue for unordered quantities.
- Hue for a linear quantity is "an encoding that now and then reduces perplexed viewers to mumbling color names, and the numbers they represent... Despite our experiences with the spectrum in science textbooks and rainbows, the mind's eye does not readily give an order to ROYGBIV"
- How far is Red from Purple?
- Highest of all is white???

Why Not?

- Quick, which are the highest and lowest rated movies?
  - Eyes going back and forth a sure sign of a bad mapping

How about now?

Don't shout

- All caps, especially if sans-serif, vibrate
- Also "minimize distinctions between letters and words, contributing to the difficulty of reading"

PART 2: BEING TRUTHFUL

Don't distort the data

- Don't change the X (or Y) scale in mid-stream

Don't distort the data

- Use same scales in comparisons

Don't distort the data

- Use same scales in comparisons

Don't distort the data

- Be very careful with aggregation
- "time series are exquisitely sensitive to choices of intervals and end points"

Size in graphic == size in data

- Adding perspective worsens the crime

Size in graphic == size in data

- Don't use 3D for a 1D quantity… ever

Put data in context

- Thiokol only looked at 2 or 8 data points, so could not trend.

Put data in context

- Thiokol only looked at 8% of the data, so could not trend.

Show data in context

- Show scale, orientation, and axes

Charts not always the answer

- Charts have a cognitive "start-up cost", plus need to look over to axes
- If < 20 numbers, perhaps a table is better.

	Participation (%)
College or University	7
Adult	8
Total	15

Homework

- "The owner of the memex, let us say, is interested in the origin and properties of the bow and arrow. Specifically he is studying why the short Turkish bow was apparently superior to the English long bow in the skirmishes of the Crusades"

PART 3: ARGUING BETTER

Cause-and-effect

- Put cause and effect on the axes, with cause on the X axis

What's wrong here?

- Cause/effect mismatch
- Chartjunk
- Extraneous data
- Encodings of damage

Use thin lines for non-data, thick lines for data

Minimize grid lines

- They are the "scaffolding", should be hidden when you "drywall"

Make graphs longer than high

- Analogy to the horizon
- X is typically "cause" axis, emphasizes it
- Try for the golden mean (1.618)
  - 1.0 … 2.2 seems the range
  - But this way you can sound educated!

Two small figures > 1 big one

- "Comparisons must be enforced within the scope of the eyespan"
- Why in this talk I do before and after on same slide

Digression

- Chernoff faces
- Good for showing multiple plots on the same graph
- Exploits our ability to rapidly distinguish faces
- Scales down well

Focus+Context

- Showing multiple levels of data in the same graphic can lend to analysis on both the *macro* (Context) and *micro* (Focus) level

Focus+Context

- long established in art & architecture
- Another more meaningful example later

(Thanks to David Modjeska of U. Toronto)

Focus+Context

- "Such designs can report immense detail, organizing complexity through multiple and (often) hierarchical layers of contextual reading"

Focus+Context

- Don't be afraid of detail, challenge is to maximize it. Humans are very good at focus+context
- "micro/macro designs enforce both local and global comparisons and, at the same time, avoid the disruption of context switching"

Focus+Context

- "Detail cumulates into larger coherent structures. Simplicity of reading derives from the context of detailed and complex information, properly arranged *to clarify and detail*

In sum

- Ask yourself: "what can be taken out?"
- Show data in context
- Use 1D representations of 1D quantities
- Maximize content, minimize annotation

"The Worst Graphic Ever?"

Example #1

- Shows French invasion of Russia
- Shows 6 variables seamlessly
- Shows relations
- Pure data
- "best statistical graphic ever"

Example: Vietnam Memorial

Example: Vietnam Memorial

- Focus + context: each contributes to the other

Example: Vietnam Memorial

- Careful choice of X and Y axes
- Longer than high

Example: Pioneer

Your arrangement affects *others* insights

Homework

- Redo your homework for today. As before, work should fit on an 8" by 4" rectangle, and not suffer from faxing. No time constraints.

Project proposal

- "The principles of information design are universal - like mathematics - and are not tied to unique features of a particular language or culture"
- Conduct user studies to test this hypothesis

Project Proposal

- What do you think Tufte got wrong? Pick a rule of his, argue against it, and then justify your argument by experiment.

Project Proposal

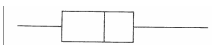
- Write a "Tufte-alyzer", that takes a Powerpoint presentation and detects "hall of shame" situations. Extra credit: suggest corrections.

Project proposal

- CNBC and Yahoo must constantly dynamically create stock graphs. The aspect ratio is (largely) a given. How do they determine the Y axis (scale and offset)? How did they determine colors and fonts? What *should* they use?

Project proposal

- Tukey's box-plot shows uncertainty for a 1D quantity in a plot.



- How would you show uncertainty for an *n*-d quantity in a visualization? (more on this later)