

Info 424

Graphical Excellence

The Visual Display of Quantitative Information (ch 1 & 2)
Show Me the Numbers (ch 1)

Today's lecture

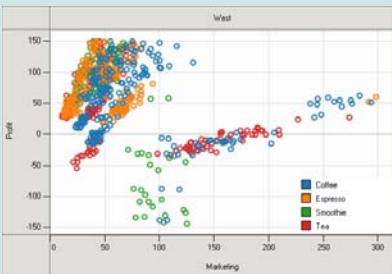
Purposes of Visualization
 Graphical Excellence & Graphical Integrity (Tufte)
 Visualization Critique

Discussion

- Gapminder
- Visualizing train schedules
- Visualizing course schedules

Questions so far?

Data → Pictures → Insight



The goal of visualization is insight (not pictures)

What is the purpose of Vis?

Few

- Analysis
- Monitoring
- Planning
- Communication

Tufte

- Description
- Exploration
- Tabulation
- Decoration

Other

- Education
- Aid to thinking
- Aid to reasoning
- Problem solving
- Enhance cognition
- Insight
- Entertainment?

Thanks to Marti Hearst and Melanie Tory

Do you know the answer?

Yes

- Presentation, communication, education
- Few: data, message, audience
- Tufte: Description, tabulation

No

- Exploration, analysis
- Problem solving, planning
- Aid to thinking, reasoning

Creator or viewer?

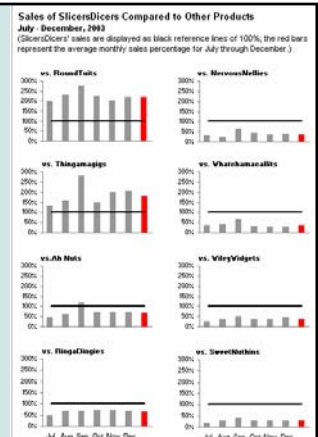
Usually a loop

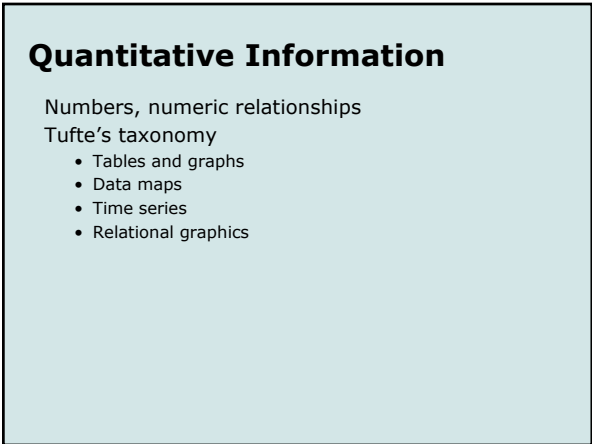
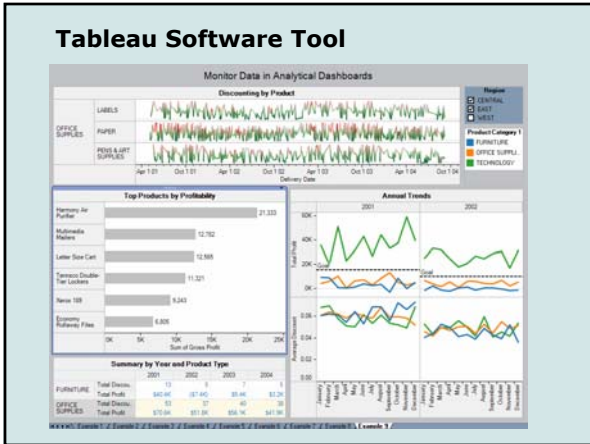
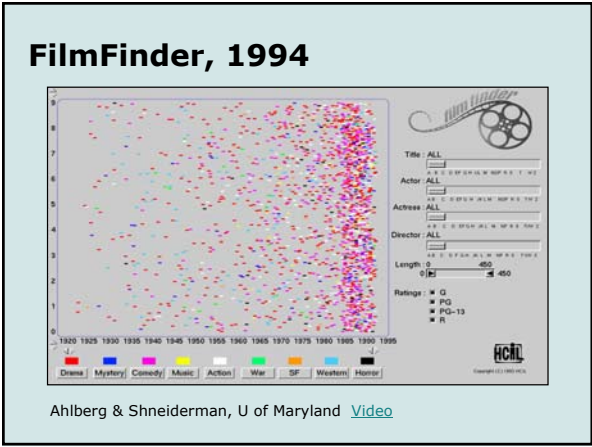
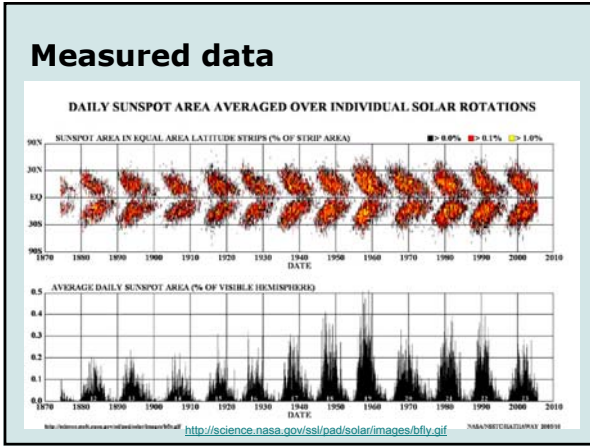
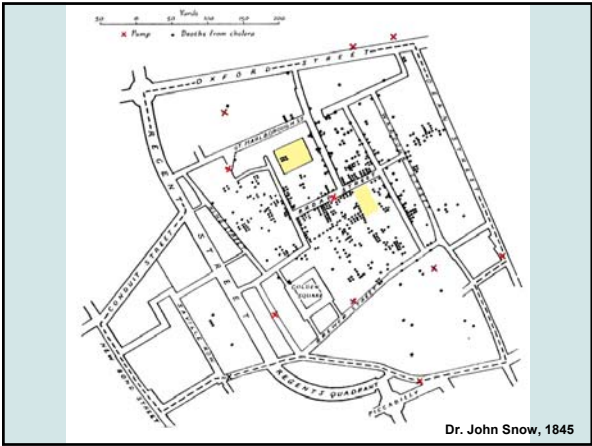
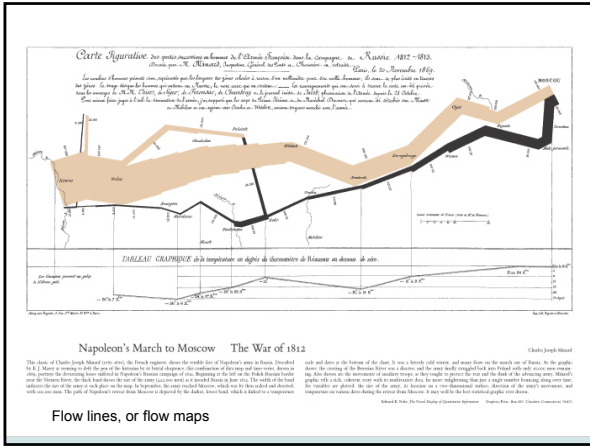
Few Example

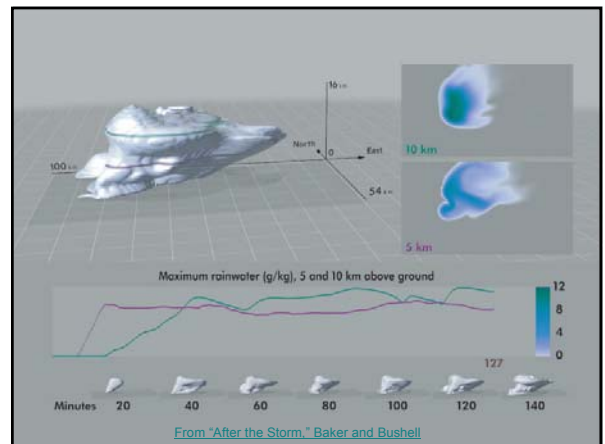
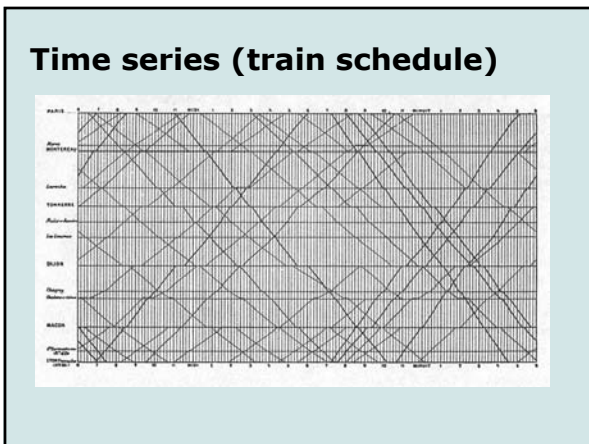
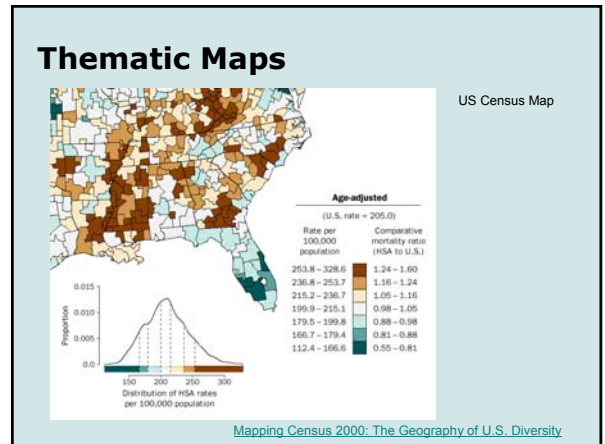
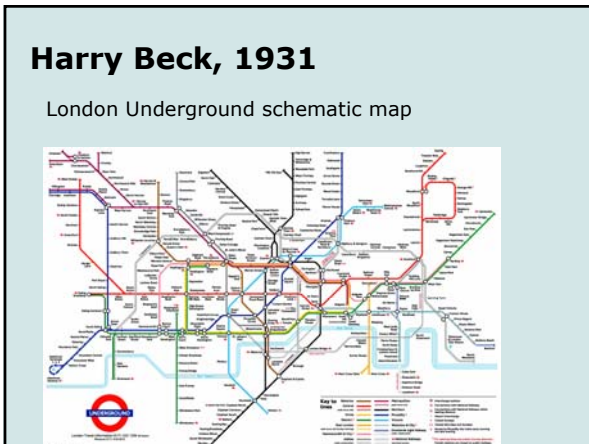
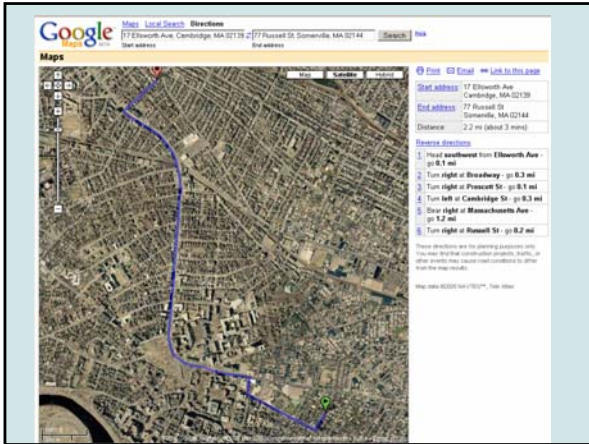
Communicate

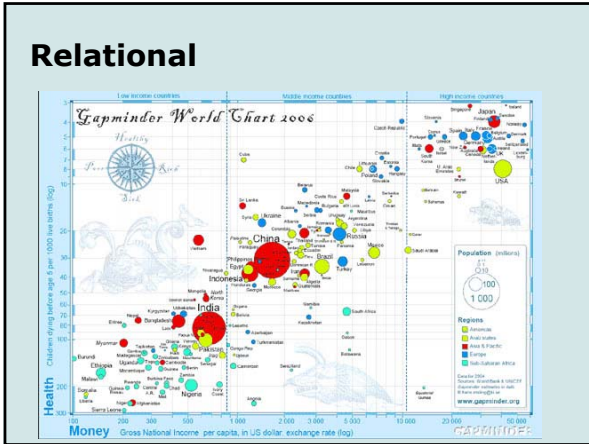
What: Sales relative to other products

To whom: The manager in charge









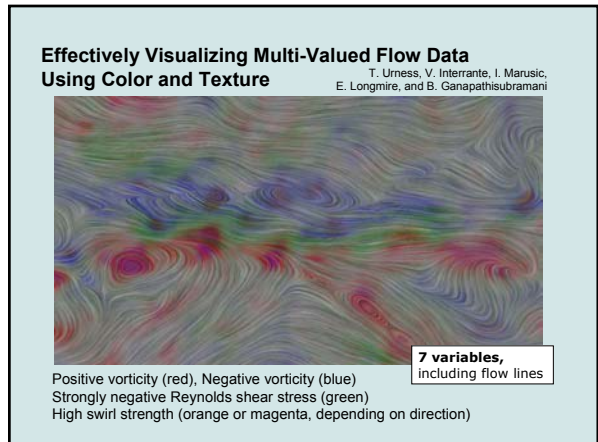
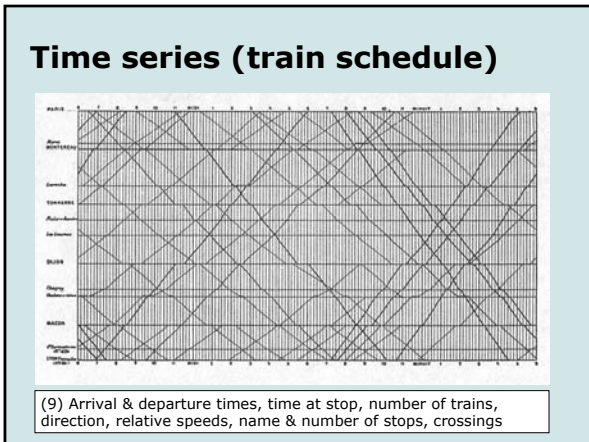
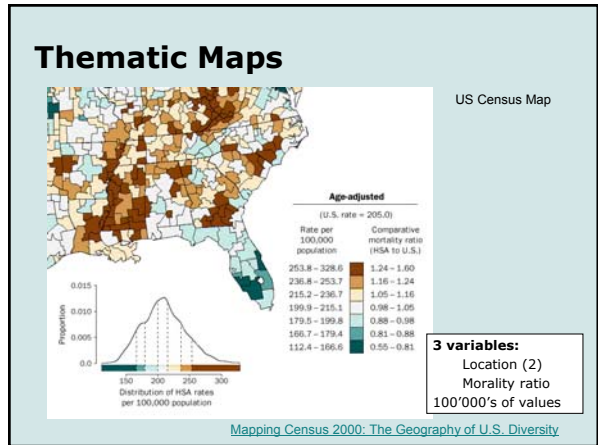
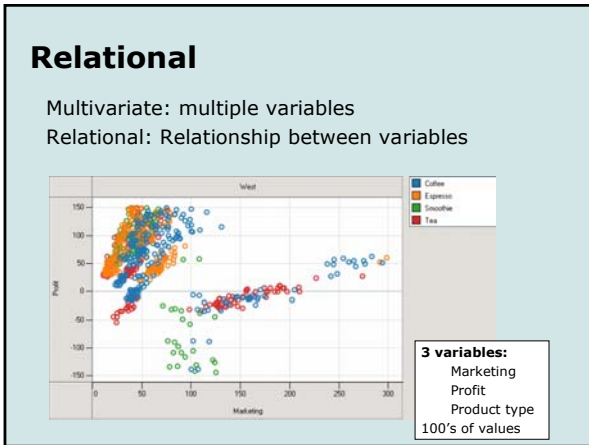
Graphical Excellence

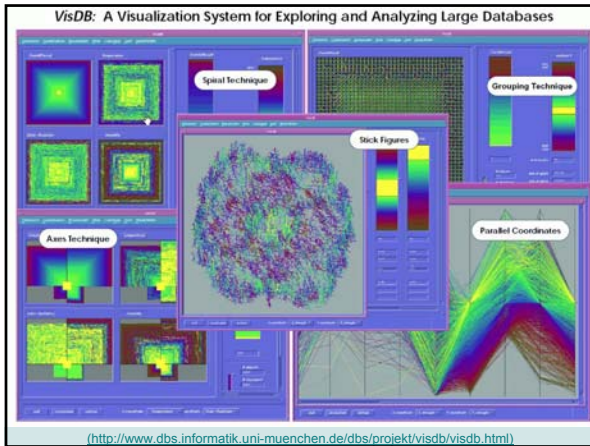
In brief

- Interesting data (complex ideas, multivariate)
- Clear, precise, concise presentation (data-ink ratio)
- Accurate communication (lie factor)

Tufte

- Well-designed presentation of interesting data – a matter of substance, of statistics, and of design
- Complex ideas communicated with clarity, precision and efficiency
- Gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space
- Nearly always multivariate
- Requires telling the truth





Graphical Displays Should

- Show the data
- Induce the viewer to think about the substance, rather than about methodology, graphic design, the technology of graphic productions, or something else.
- Avoid distorting what the data have to say
- Present many numbers in a small space
- Make large data sets coherent
- Encourage the eye to compare different pieces of data
- Reveal the data at several levels of detail
- Serve a reasonably clear purpose: description, exploration, tabulation, or decoration
- Be closely integrated with the statistical and verbal descriptions of a data set.

Tufte VDI Chapter 1, first page

Show the data

Maximize the "Data-Ink Ratio"

- Data-ink = the ink used to show data
- Data-ink ratio = data-ink/total ink used

Process

- Erase non-data ink
- Erase redundant data ink
- Revise and edit

<http://www.tbray.org/ongoing/data-ink/di1>

Graphical Integrity

The representation of numbers as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities represented

Clear, detailed and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data.

Show data variation, not design variation.

In time-series displays of money, deflated and standardized units of monetary measurement are nearly always better than nominal units.

The number of information-carrying (variable) dimensions depicted should not exceed the number of dimensions in the data

Graphics must not quote data out of context

In brief

Present value relationships accurately

- Presentation size precisely matches data
- Avoid area and volume encodings
- Adjust currency values for inflation or other correlated changes (such as population changes)

Label carefully and clearly

Present data in context

Lie Factor

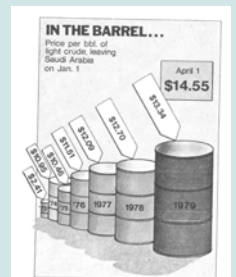
Lie Factor

- (Size of effect in graphic)/(size of effect in data)
- Truth = 0.95-1.05 (5% error)
- 2-5 not uncommon

Lie factor by area = 9.4

Lie factor by volume = 59.4

NB: not adjusted for inflation



Estimating Relationships

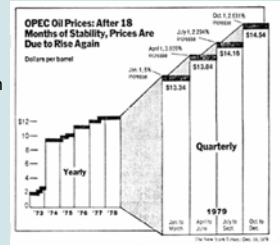
Which is twice the size of the smallest one?

Design Distortions

Gratuitous 3D and perspective
Non-uniform intervals and scales

Design Distortion

- Yearly vs. Quarterly
- Nonlinear scale
- Perspective distortion



Include detailed labels

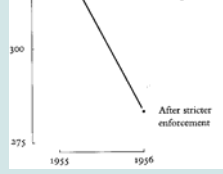
Clear, detailed, complete

- Provides precise values
- Indicates important events
- Defines units, context
- Corrects for presentation distortion

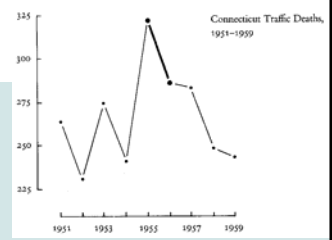
Label directly on the graphic

- Makes graphic self-explanatory
- Maintains focus (compared to legends)

Connection Traffic Deaths, Before (1955) and After (1956) Stricter Enforcement by the Police Against Cars Exceeding Speed limit

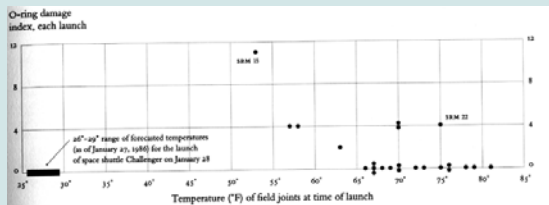
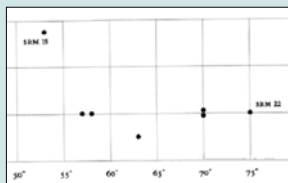


A few more data points add immensely to the account:



Show context

Challenger data
Only complete context shows trend



Summary

Graphical Excellence

- Interesting data (complex ideas, multivariate)
- Clear, precise, concise presentation (data-ink ratio)
- Accurate communication (lie factor)

Graphical Integrity

- Present value relationships accurately
 - Size matches data
 - Avoid area and volume encodings
 - Adjust currency values for inflation, etc.
- Label carefully and clearly
- Present data in context

Visualization Critique

Find an example of a good visualization

Analyze it to explain its quality

Suggested sources

- Google images (charts, graphs, data maps)
- Online news sources
- Online science sources

Due Monday, October 8th at 7am

We'll show in class on Tuesday and Thursday

Analysis Questions

1. Who is the intended audience?
2. What information does this visualization represent?
3. How many data dimensions does it encode?
4. List several tasks, comparisons or evaluations it enables.
5. What principles of excellence best describe why it is good?
6. Can you suggest any improvements?
7. Why do you like this visualization?

VisCritique Example

Baby growth chart ([link](#))

Who is the intended audience?
Doctors, nurses and parents

What information does this visualization communicate?
The distribution of height and weight measurements for infant boys as a function of age. Infants can be plotted on the chart, and their growth evaluated.

How many data dimensions does it encode?
Length, weight, age, percentile
Length and weight in both metric and English units

VisCritique Example

Baby growth chart ([link](#))

List several tasks, comparisons or evaluations it enables.

1. It allows a parent to see where their child's development compares to the population
2. It provides a record that allows a pediatrician to evaluate a child's growth over time
3. It illustrates a normal growth pattern for infant boys

What principles of excellence best describe why it is good?

1. It makes a large data set coherent
2. It encourages comparisons
3. It encodes many numbers

VisCritique Example

Baby growth chart ([link](#))

Can you suggest any improvements?
The grid lines should be less bold

Why do you like it?

1. I found it personally useful when my children were babies for comparing and evaluating their growth.
2. I think it is a very clever layout

Questions?

Discussion

GapMinder

Course Schedules ([html](#), [Flash](#), [XML](#), [2006](#))

Train Schedules (Amtrak vs. Tufte)