NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

BC3406 Business Analytics Consulting

Damien Joseph
Nanyang Business School

Gino K. George
Analytics Executive

A Lee Gilbert
Course Instructor

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Agenda

- Communication of Insights
  - The Art and Science
- Self-Service Visualizatons
- Team Building

Visualization

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

What is Information Visualization?

- Visualize
  - to form a mental image or vision of …<<something>>
  - to imagine or remember as if actually seeing (American Heritage dictionary, Concise Oxford dictionary)
  - "Transformation of the symbolic into the geometric" (McCormick et al., 1987)
  - "... finding the artificial memory that best supports our natural means of perception'' (Bertin, 1983)
- The depiction of information using spatial or graphical representations, to facilitate comparison, pattern recognition, change detection, and other cognitive skills by making use of the visual system.
- Information visualizations are metrics expressed as graphics with labels.

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Information Visualization

- Current problem for visualization:
  - How to understand data and results of analyses?
- Solution
  - External cognitive aids
  - Take better advantage of human perceptual system
  - Convert information into a graphical representation.
- Issues
  - How to convert abstract information into graphical form?
  - Do visualizations do a better job than other methods?

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

External Cognitive Aids

- External cognition
  - Internal and external representation
  - Processing weave together in thought

- External cognitive aids can enhance cognition
  - Examples:
    - Cheatsheets.
    - For multiplication, use of paper reduces the time required by a factor of 5 (for most people) when compared to mental computation. Why?

- An important class of external cognitive aids: Symbolic displays
  - Charts for navigation
  - Diagrams

- A graphic picture may be used to communicate an existing idea or thought

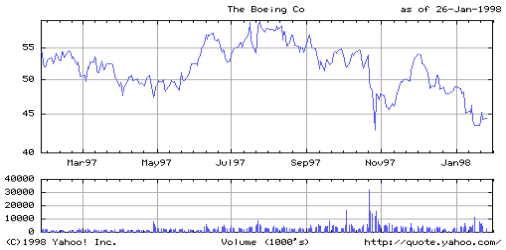
- But graphical aids can also be used in formulating ideas and thoughts
  - Accelerates the identification of hidden patterns in data
  - “A picture is worth a thousand words”

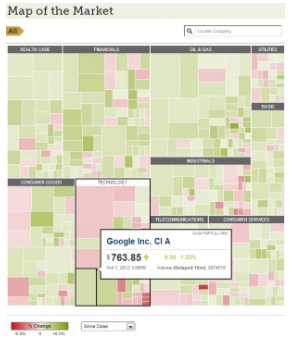
NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

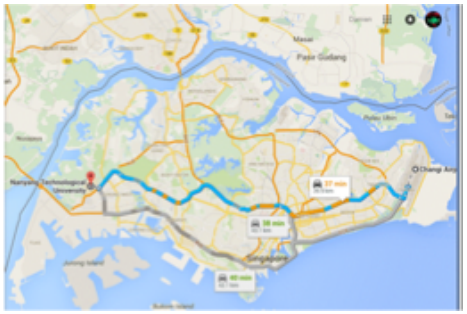
Symbolic Display

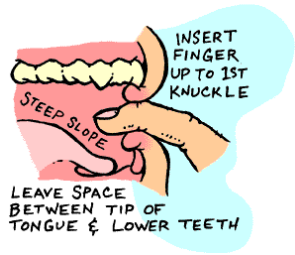
categorical: bar chart
numeric: bar chart, trendline

- A visual display that illustrates one or more relationships among entities
- A shorthand way to present information
- Allows a trend, pattern, or comparison to be easily apprehended
- Types of symbolic displays
  - Graphs
  - Charts
  - Maps
  - Diagrams









NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Image and Image Theory

- An image is the fundamental perceptual unit of a visualization.
  - Ideal visualizations contain only a single image to optimize “processing efficiency" (speed with which observer can extract the information)

- Visual Processing occurs in 3 steps:
  - Formation of retinal image;
  - Decomposition of retinal image into an array of specialized representations; and
  - Reassembly of information into perception.

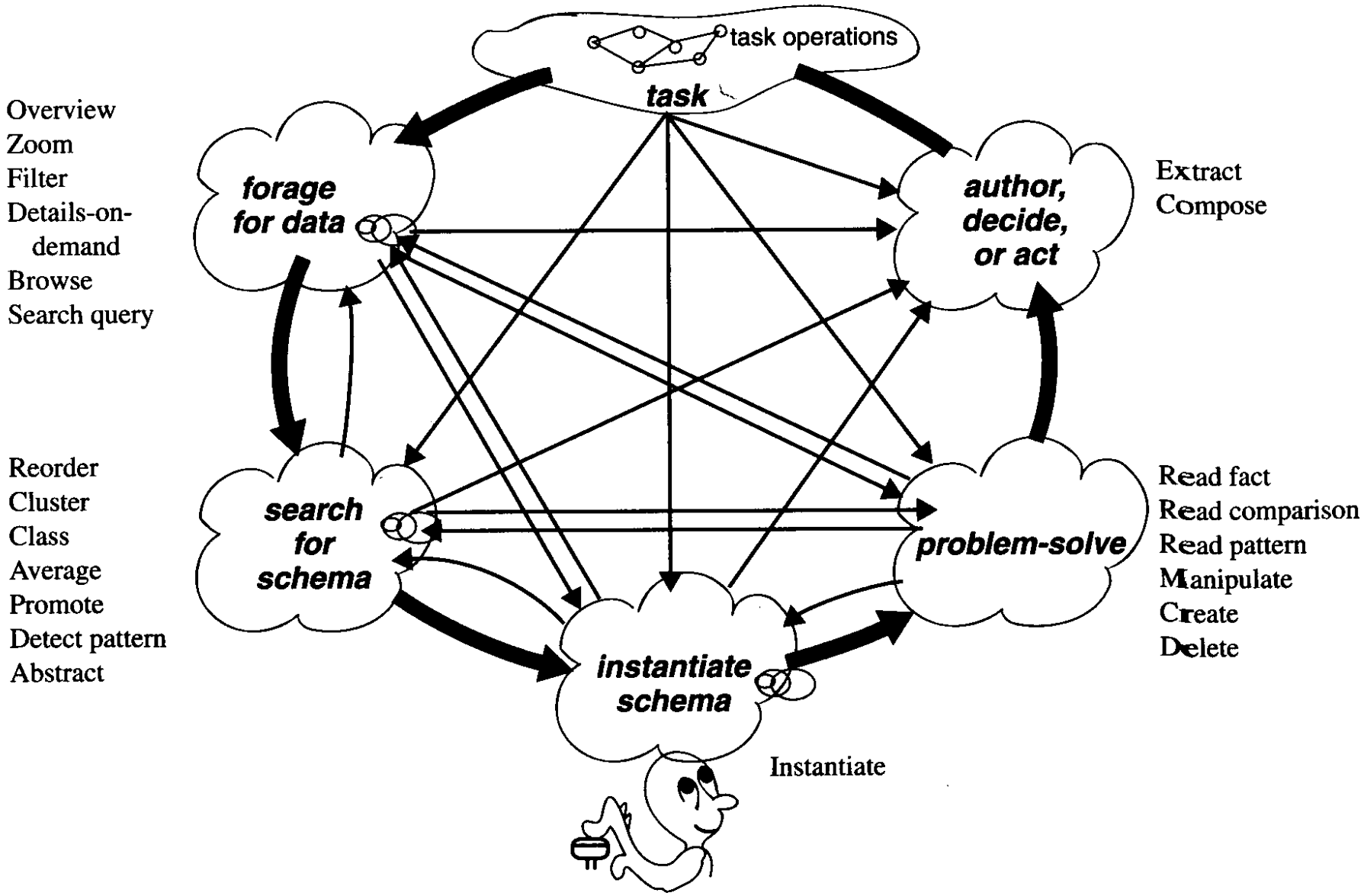
NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Knowledge Representation

- Knowledge crystallization involves getting **insight** about data relative to some **task**
- Approach:
  - **Understand** the problem and **identify** the question
  - Information **gathering (foraging)**
    - from corporate databases; online recourses, human intelligence…
  - Search for/build a **model** (representation)
    - need to know what data/variables to include/omit
    - This would be discussed in Business Analytics I & II
  - **Analyze** model with data
    - Problem solve to **trade-off features of model**
    - May have to develop new/revised models
  - Package results for interpretation in some **output**
- A visualization tool supports and/or automates some of these steps
  - Visualization tool is a cognitive aid during **modeling**

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Knowledge Representation



NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

How Visualization Amplifies Cognition

- Different ways that visualizations could help amplify cognition:
  - By **increasing memory** and processing resources available
  - By **reducing the amount of time to search**
  - **Enhancing the detections of patterns** and enabling perceptual inference operations
  - Aid perceptual **monitoring**
  - By encoding information in a manipulable medium

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Benefits of Visualization?

- Allows users see several different perspectives of the data.
  - Big data makes it difficult to find the nuances that can make a difference.
  - Makes it possible to interpret vast amounts of data
  - Offers the ability to note exceptions in the data.
- Allows the user to analyze trends and patterns in the data.
  - Equips users with the ability to see influences between factors/variables.
  - Translates trends and patterns into insights
  - Makes it a highly effective decision-making tool.
- By simplifying the presentation, Data Visualization can reduce the time and difficulty it takes to move from data to decision making.

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Unresolved Issues

- Does visualization help?
  - Sometimes
  - Provides best support to textual narratives
- Does visualization convince?
  - Sometimes, if done well and meets informational requirements

Challenges:
1. what is purpose of viz
2. why
3. whom

The Art and Science of Visualization

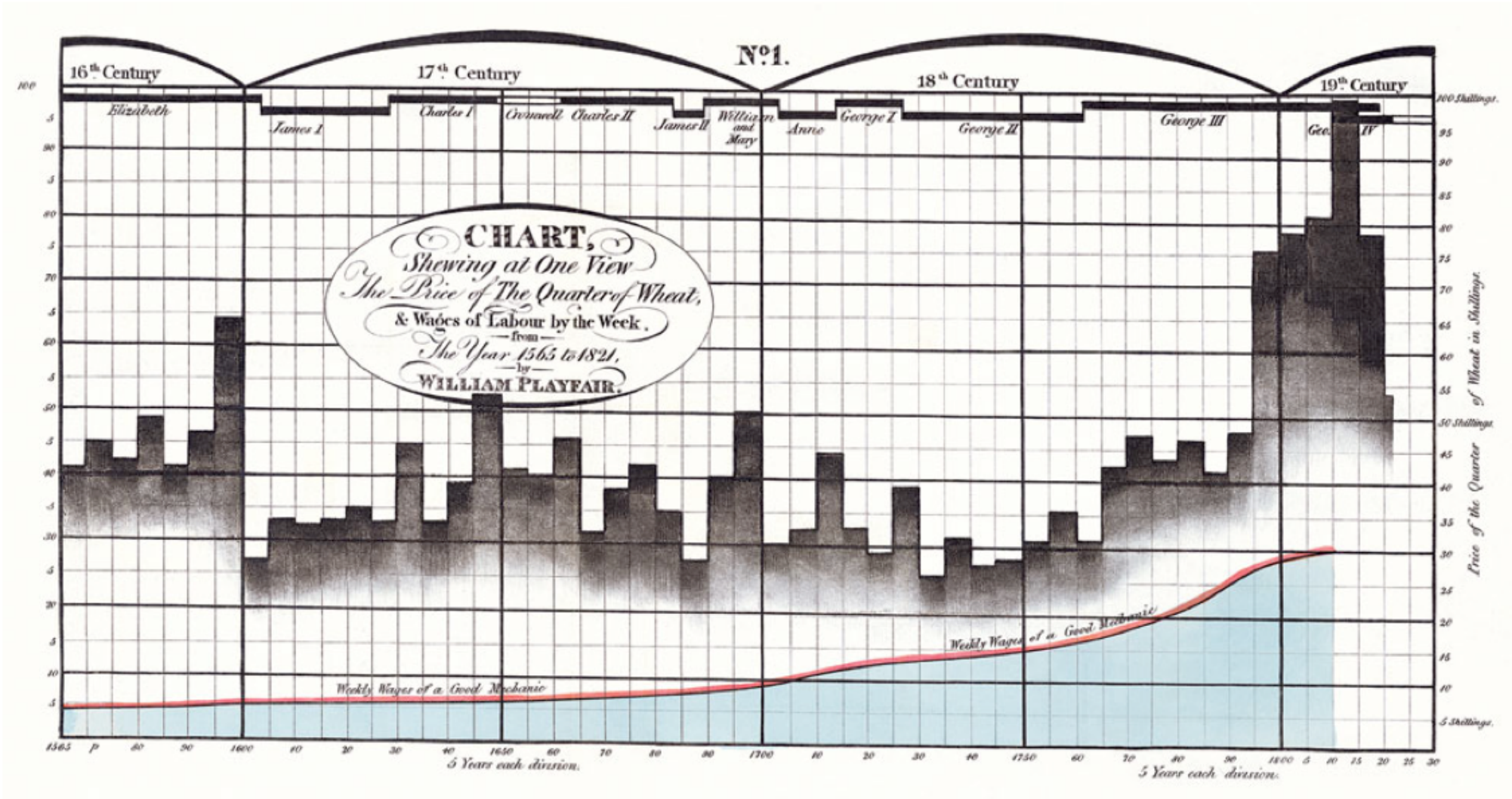
NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Learning from the Giants of Visualization

- Selected Giants
  - William Playfair (1821) – representation of descriptive data.
  - Charles Joseph Minard (1869) – longitudinal analysis.
  - Jacques Bertin (1967) – semiology of graphics
  - John Tukey (1977) – exploratory data analysis
  - Edward Tufte (1983) – statistical graphics standards/practices

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

William Playfair



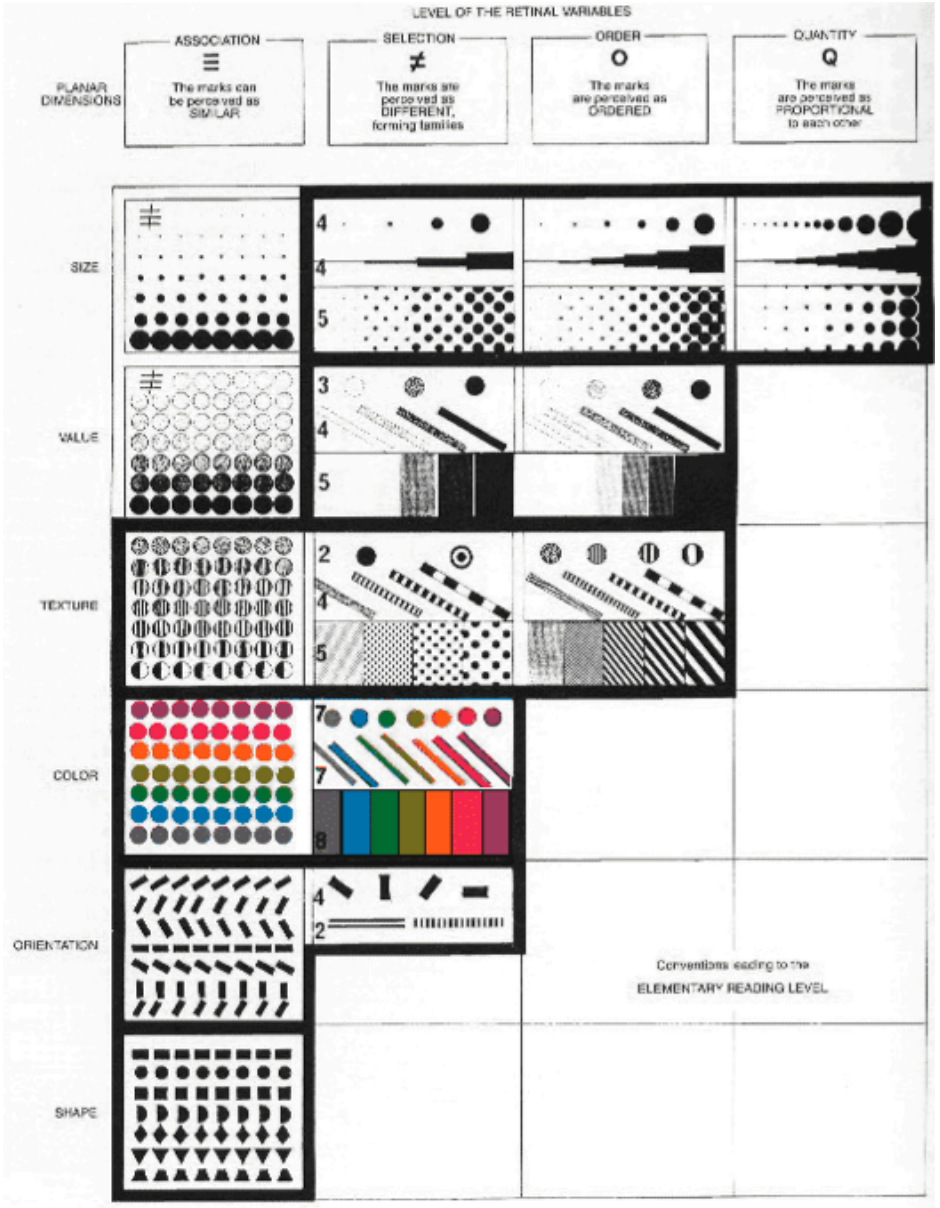
http://apandre.files.wordpress.com/2011/03/oldcombocharwagesofmechanicvspriceofwheat1821.jpg

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Bertin's Visual Variables

- Table of effective mappings between perceptions (questions) and the visual representations

- Bertin recommends
  - Shape be used for associative perceptions;
  - Color can be used for associative and selective perceptions; and
  - Size be used for selective, ordered and quantitative perceptions.

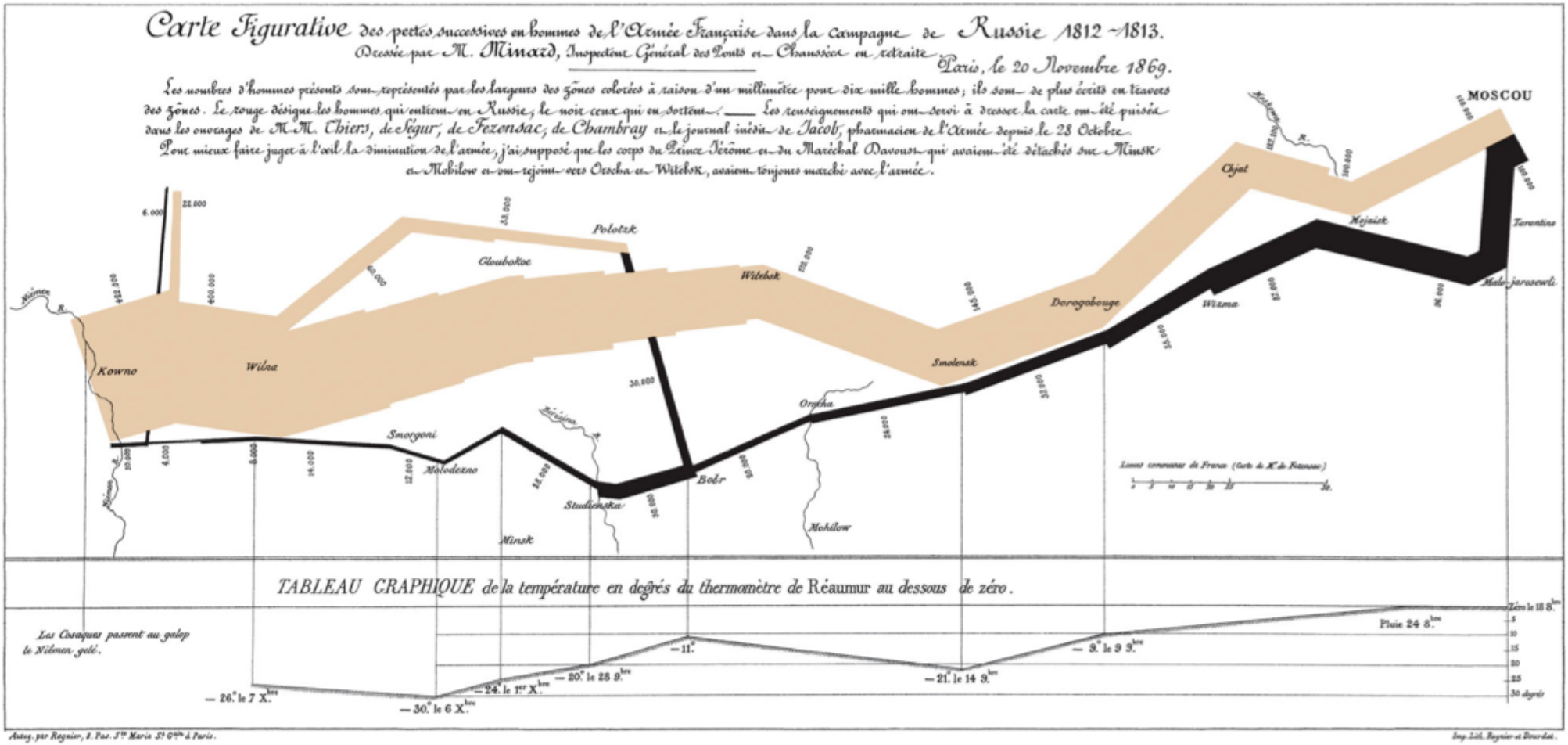


Source: Bertin, J. 1974. *Sémiologie Graphique*. Paris: Editions Gauthier-Villars. Deutsche Übersetzung von Jensch, G.; Schade, D.; Scharfe, W.: Graphische Semiologie.Diagramme – Netze - Karten. Berlin: Walter de Gruyter

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Charles Minard: Napolean's March

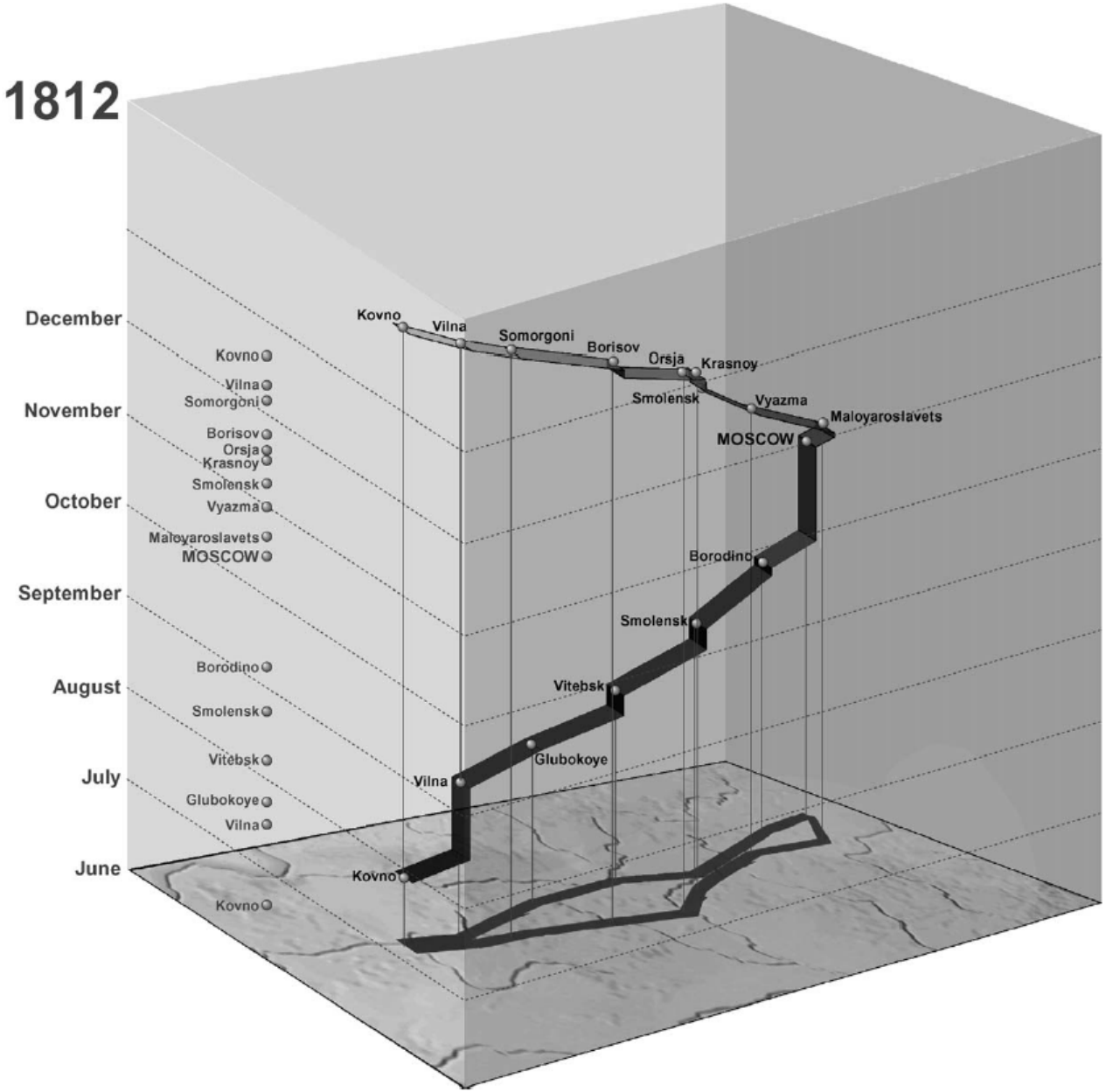
The French engineer, Charles Minard (1781-1870), illustrated the disastrous result of Napoleon's failed Russian campaign of 1812. The graph shows the size of the army by the width of the band across the map of the campaign on its outward and return legs, with temperature on the retreat shown on the line graph at the bottom. Many consider Minard's original the best statistical graphic ever drawn.



NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Menno-Jan Kraak's Reinterpretation

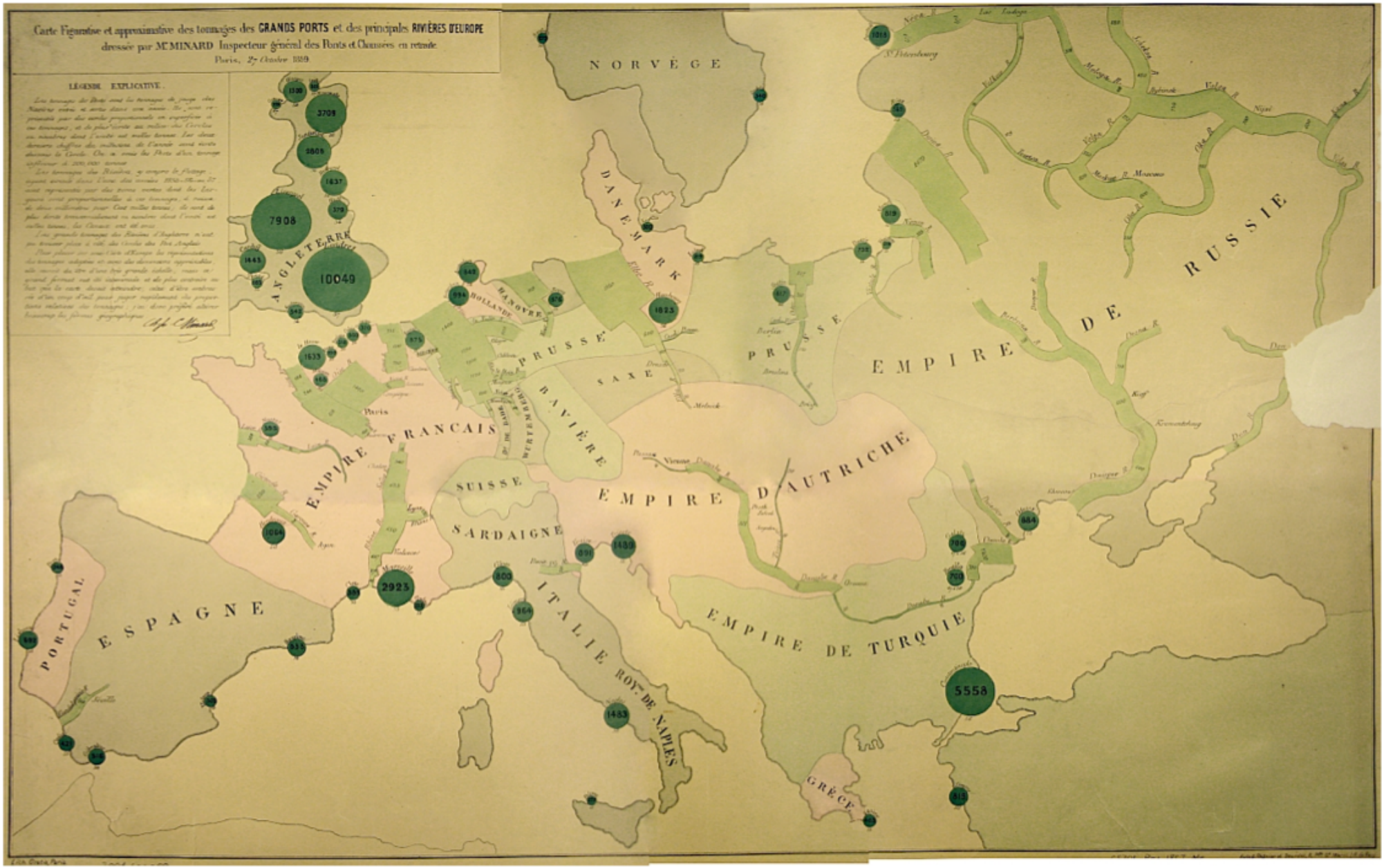
This one unravels time into the 3rd dimension, superposed above the base map.



Source: Kraak, M.J. (2003). Geovisualization illustrated. ISPRS Journal of Photogrammetry and Remote Sensing, 57(5-6), 390–399.

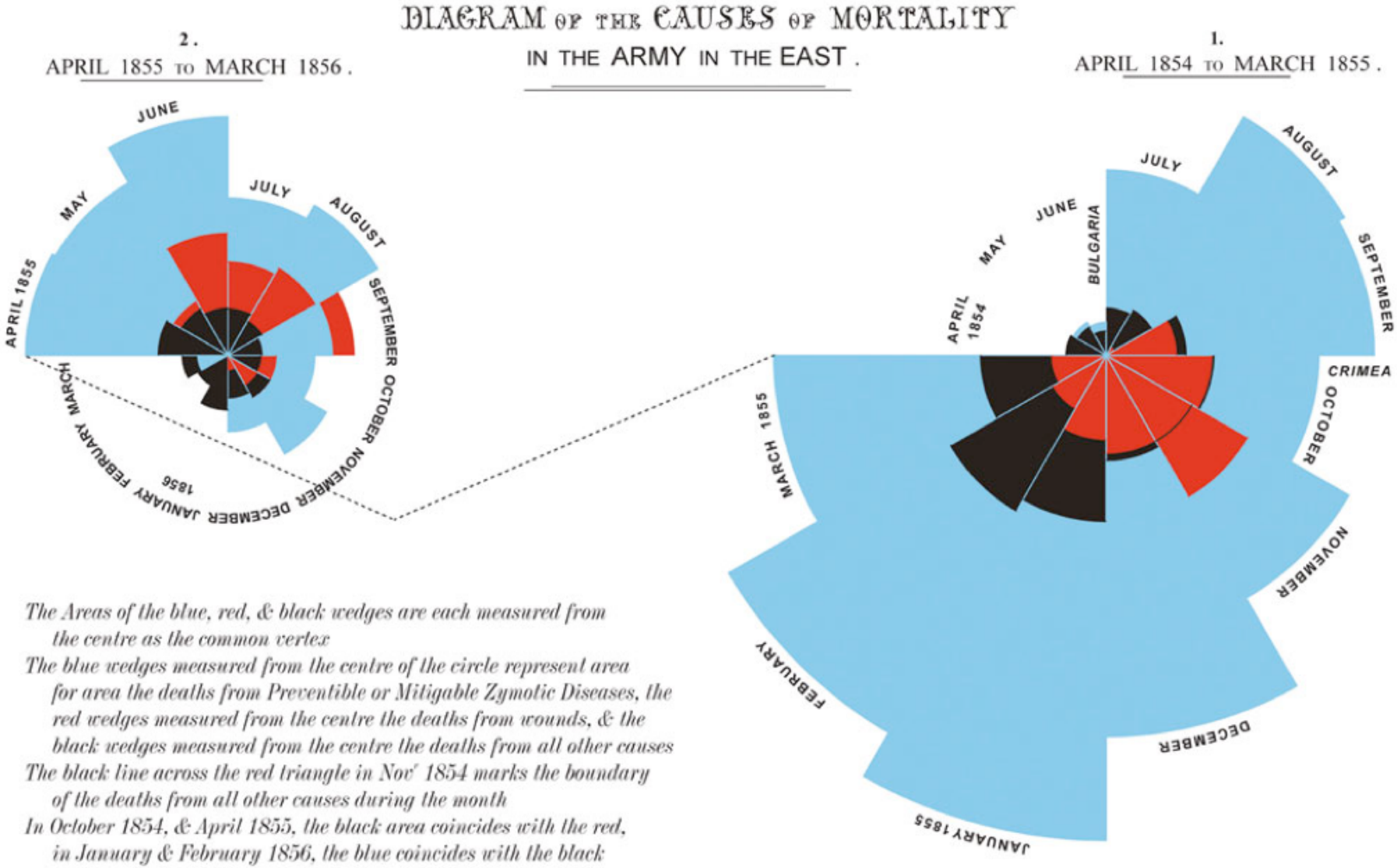
NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Charles Minard: Ports in Europe and Tonnages



NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

William Farr – Coxcomb Diagram

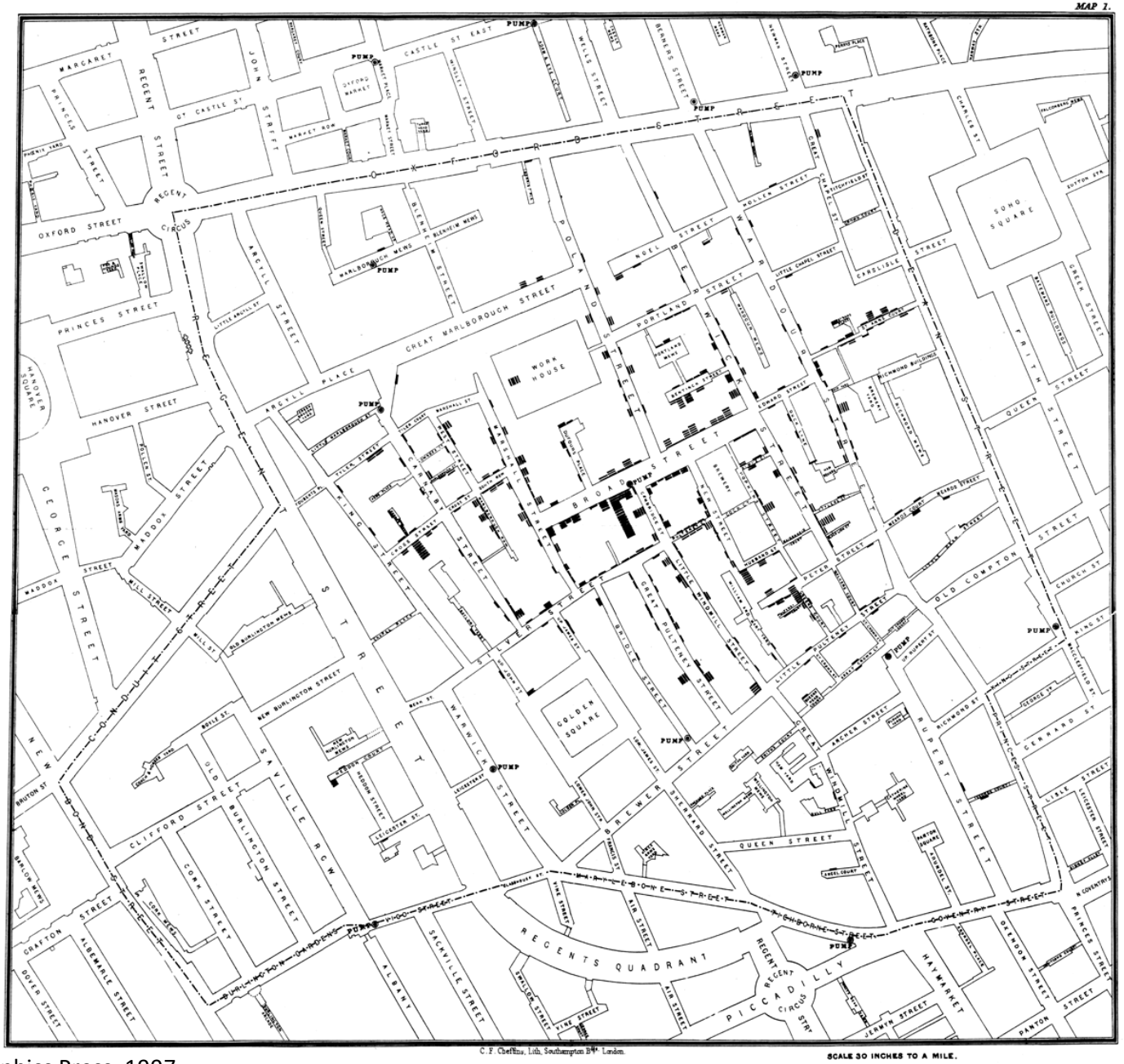


NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Tufte – Analysis of Cholera Epidemic

John Snow's illustration of a cholera epidemic and his deduction that it was caused by a bad water pump, circa 1854.

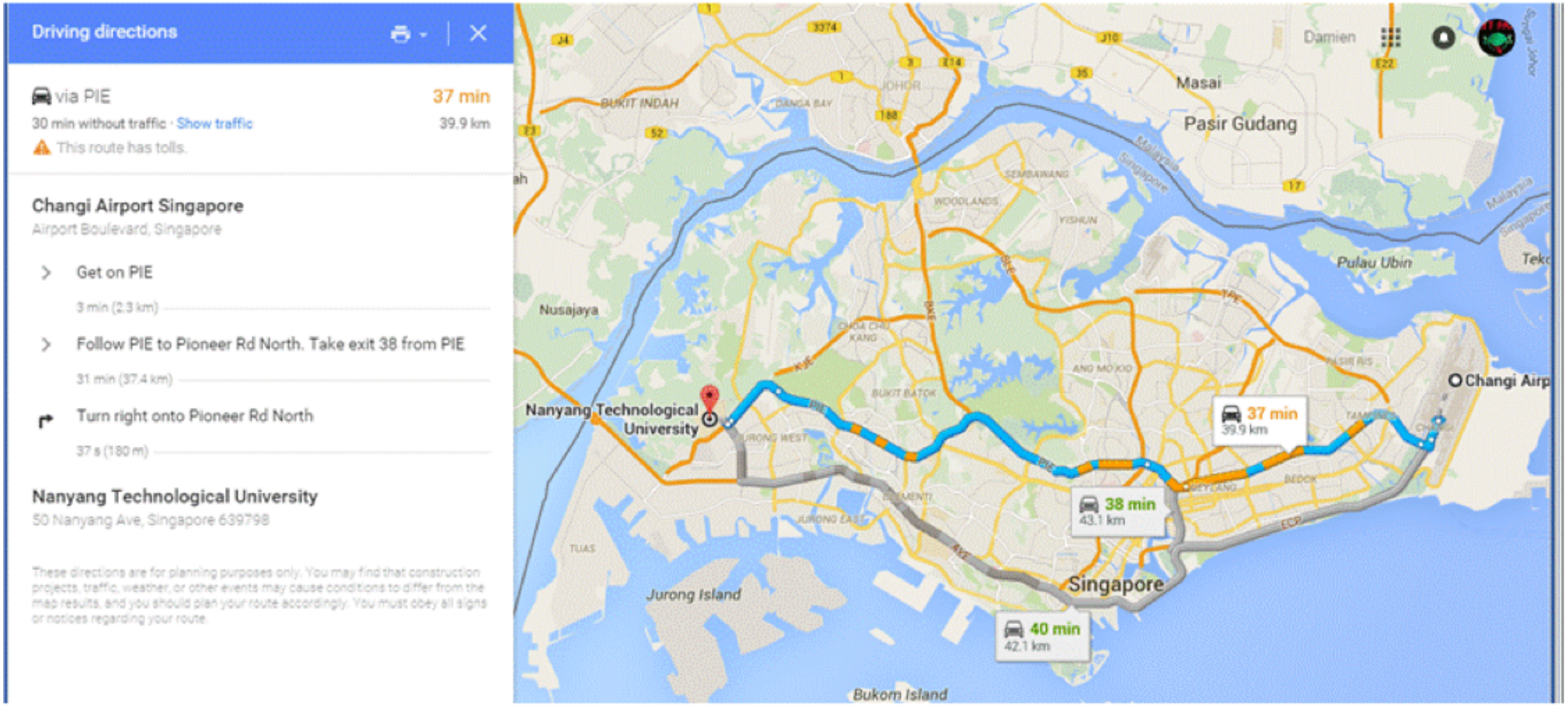
Horizontal lines indicate location of deaths.



Source: Tufte, E. From Visual Explanations, Graphics Press, 1997

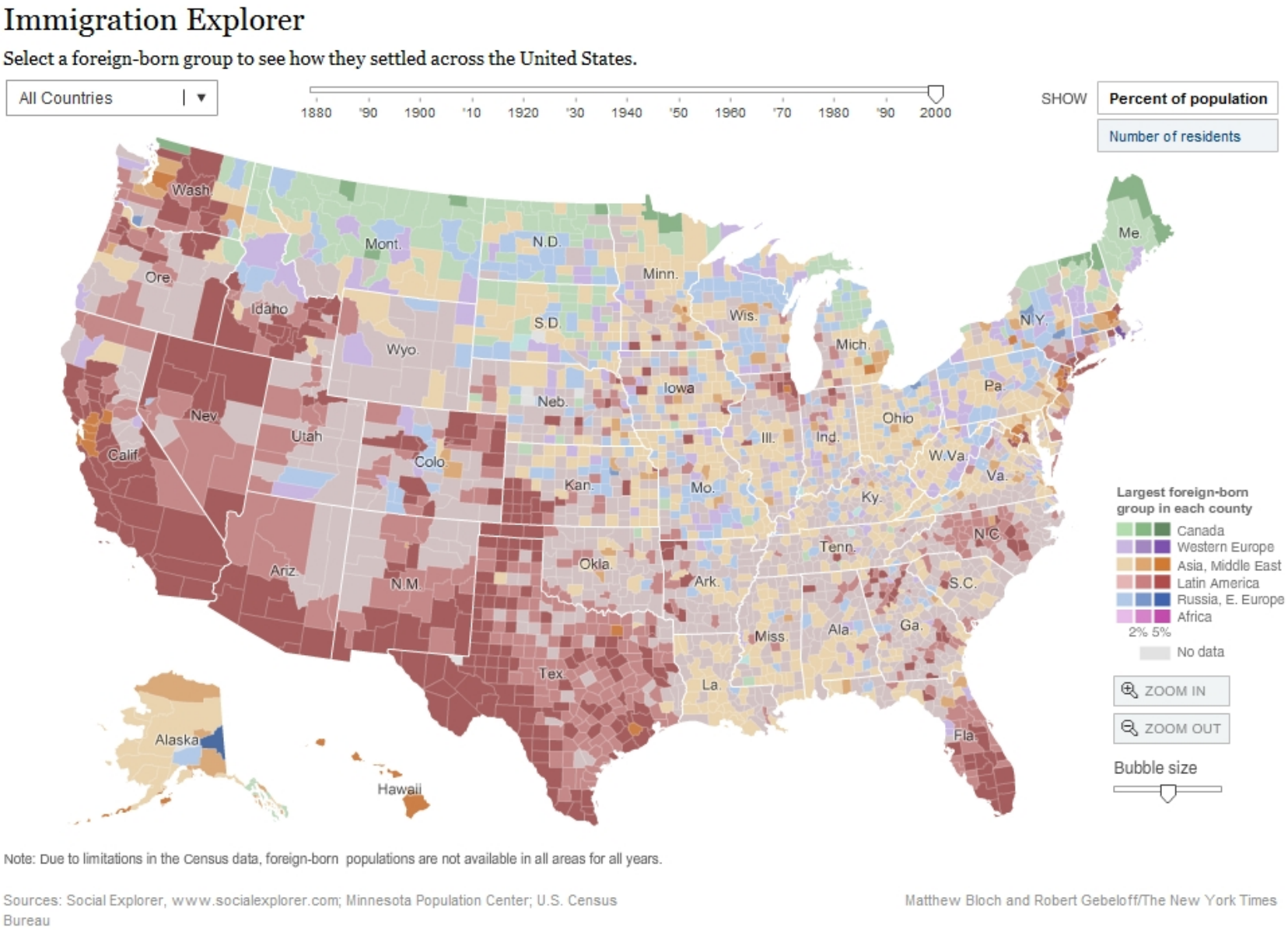
NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Google Maps: Minimalist



NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

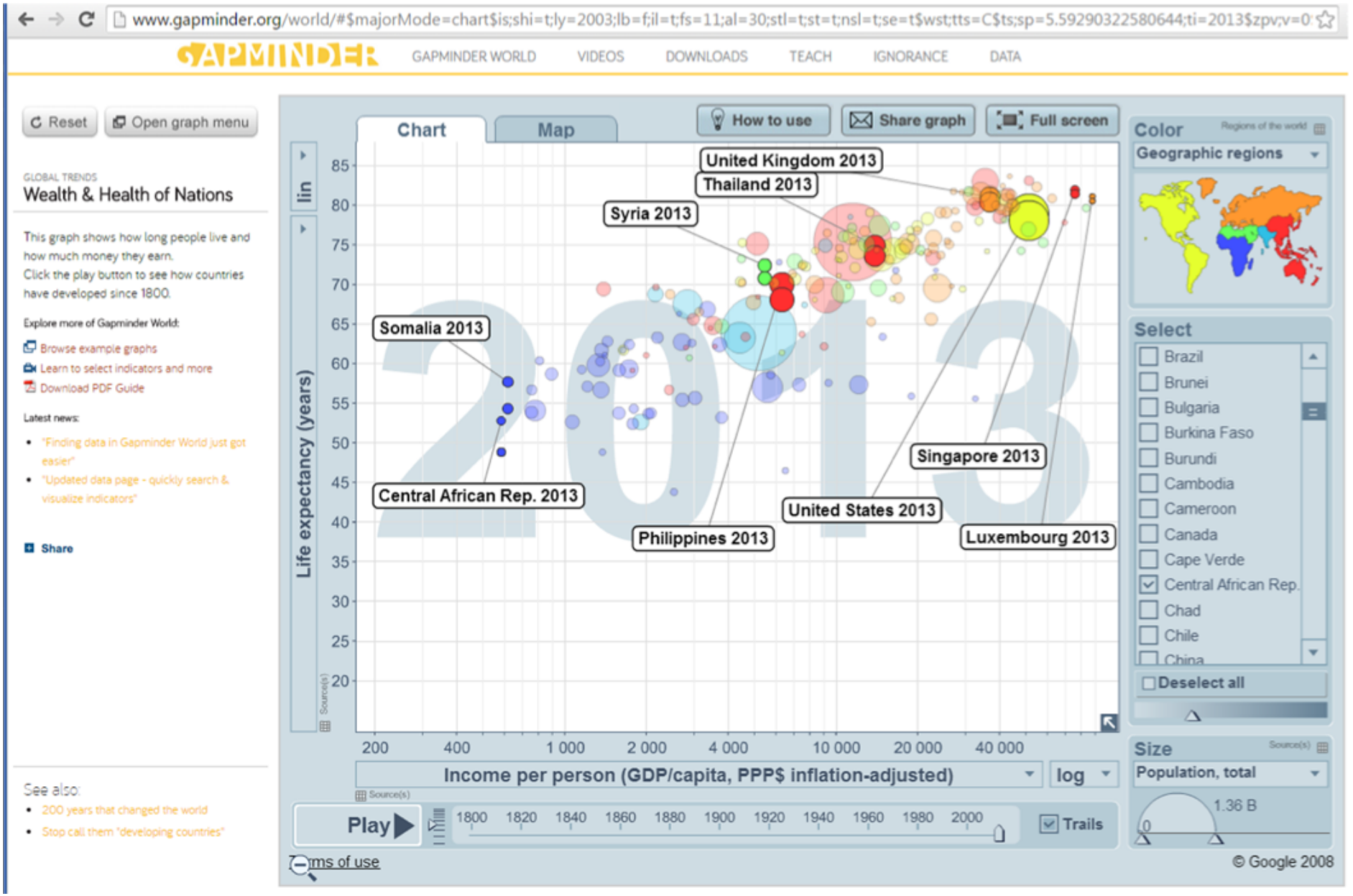
New York Times: Immigration Explorer



Source: www.nytimes.com/interactive/2009/03/10/us/20090310-immigration-explorer.html?hp

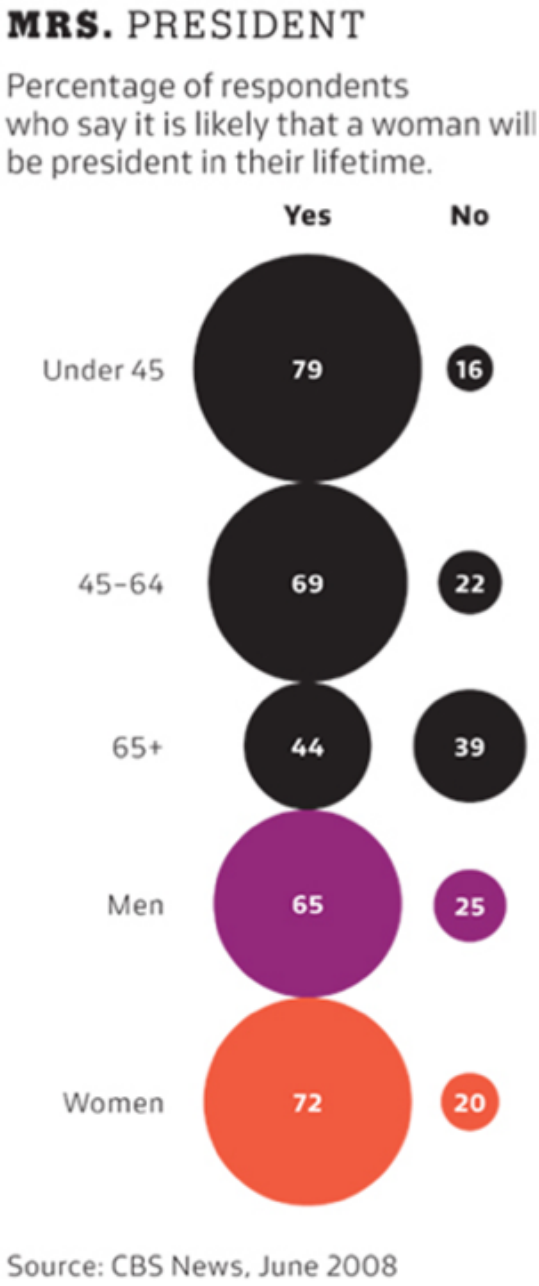
NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Gapminder



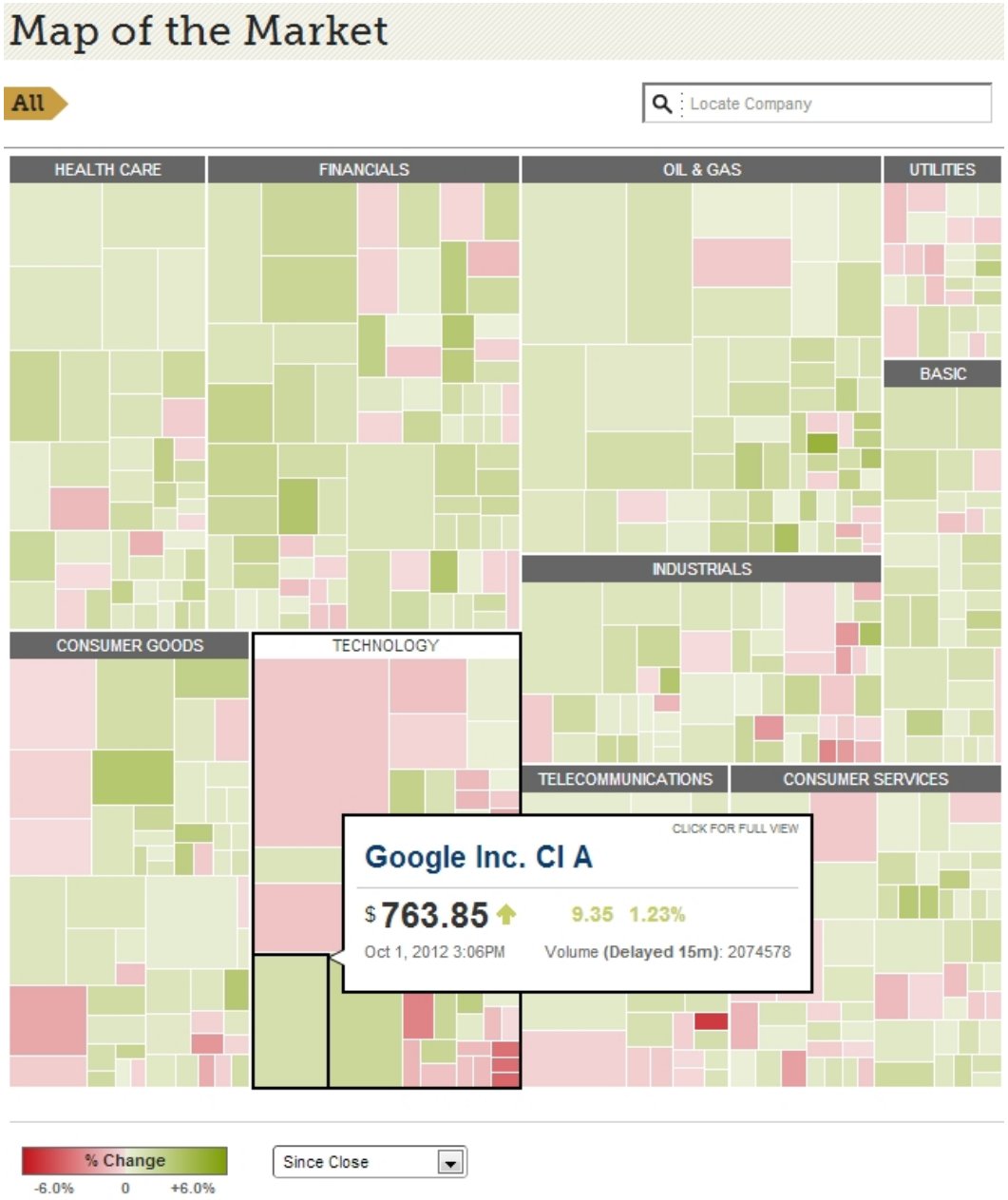
NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

CBS News: Hilary Clinton



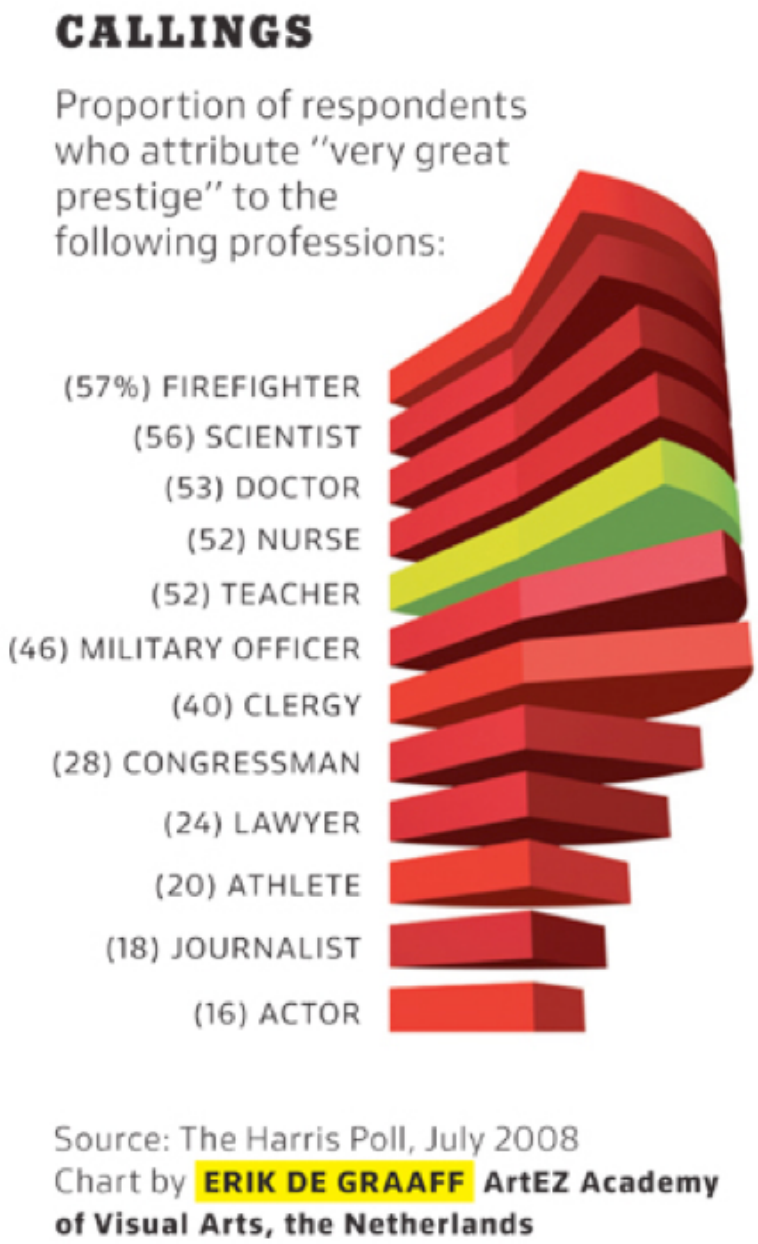
NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

SmartMoney Market Heat Maps



NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

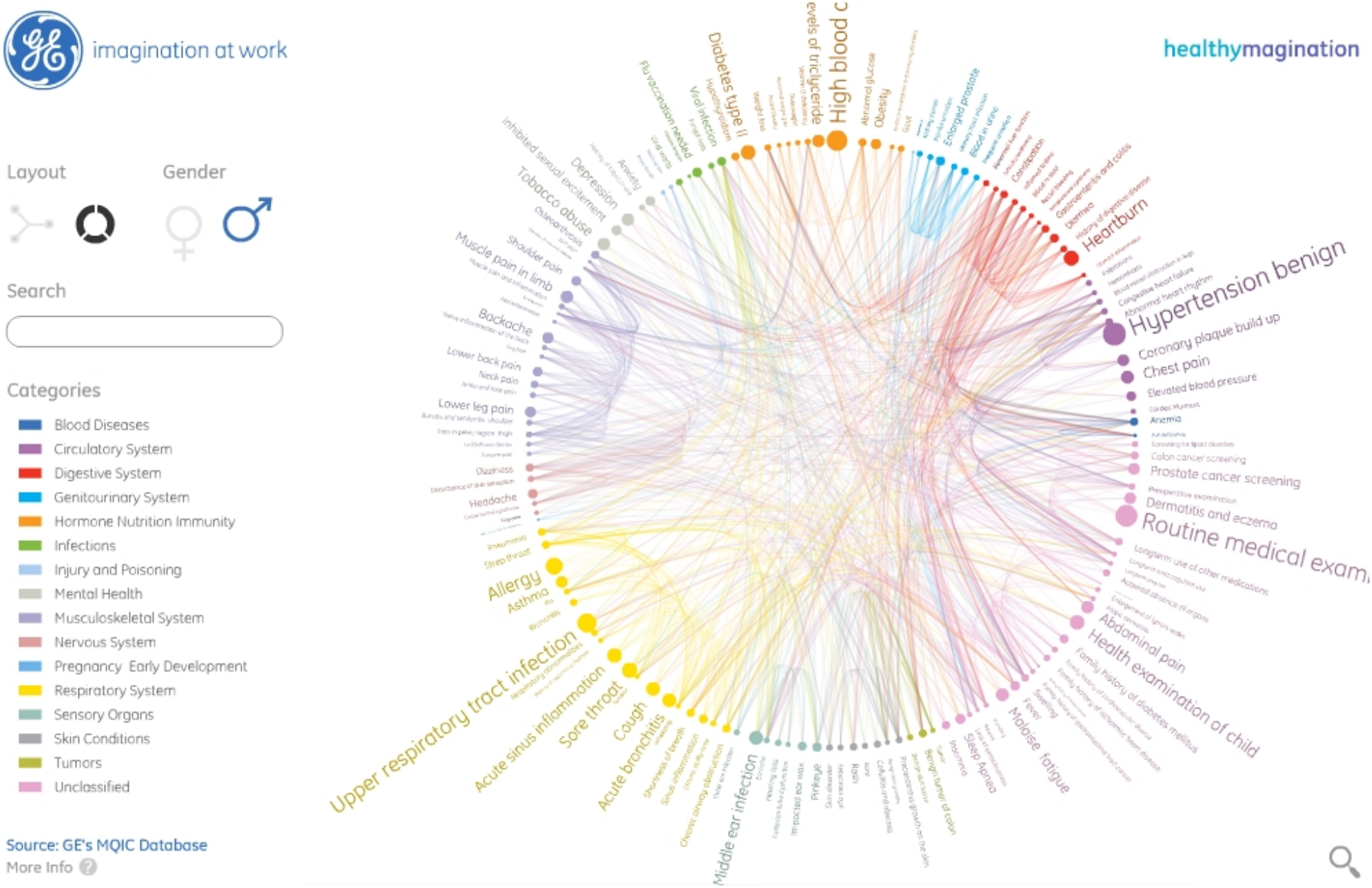
Visualization as Art: Erik De Graaff



Source: eagereyes.org/blog/2008/ny-times-the-best-and-worst-of-data-visualization

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Visualization as Art: Erik De Graaff



NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

If your are interested…..

- Cleveland, W. (1984) The elements of graphing data. Wadsworth: Monterey, Ca.
- Green, Marc (1998) Toward a Perceptual Science of Multidimensional Data Visualization: Bertin and Beyond. Marc Green, PhD
- Tufte, Edward R. The Visual Display of Quantitative Information, Graphics Press, Cheshire, Connecticut, 1989.
- F.J. Anscombe, "Graphs in Statistical Analysis," American Statistician, 27 (February 1973), 17-21.
- http://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen?language=en
- https://www.youtube.com/watch?v=jbkSRLYSojo

Principles of Visualization

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Key Questions when Visualizing

- How do we convert abstract information into a visual representation?
- While still preserving the underlying meaning ?
- And at the same time providing new insight?

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Visual Medium

- Reports
  - Paper (static, with time)
  - Web (dynamic and interactive)

- Presentations (static and dynamic)

Tufte's Principles of Graphical Excellence

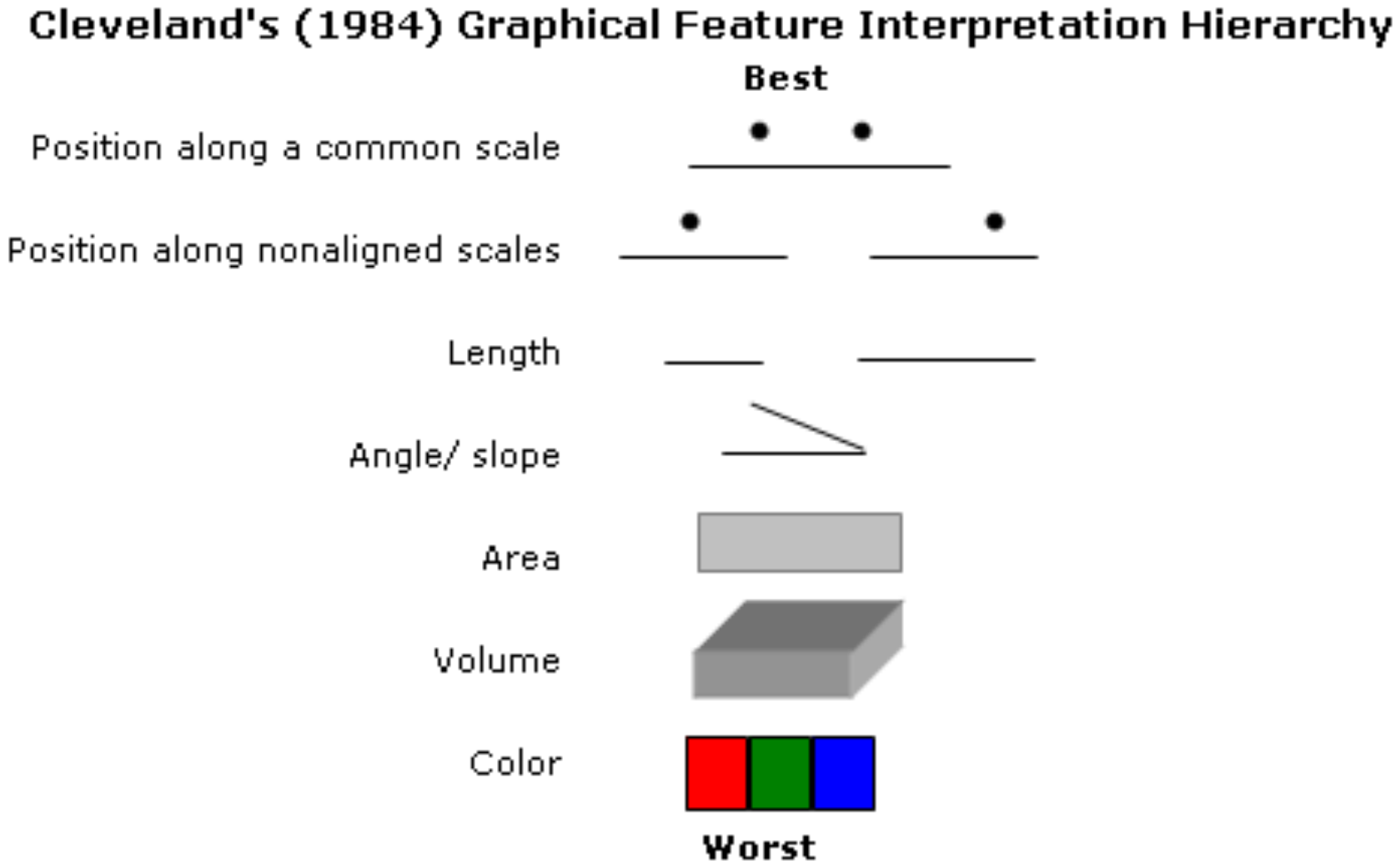
- Graphical excellence is
  - the well-designed presentation of interesting data – a matter of substance, of statistics, and of design
  - consists of complex ideas communicated with clarity, precision and efficiency
  - is that which gives to the viewer the greates number of ideas in the shortest time with the least ink in the smallest space
  - requires telling the truth about the data.

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Visual Properties

- Preattentive Processing
  - A limited set of visual properties are processed preattentively (without need for focusing attention).
  - This is important for design of visualizations
    - what can be perceived immediately
    - what properties are good discriminators
    - what can mislead viewers
  - < 200 - 250ms qualifies as pre-attentive
    - eye movements take at least 200ms
    - yet certain processing can be done very quickly, implying low-level processing in parallel
  - If a decision takes a fixed amount of time regardless of the number of distractors, it is considered to be preattentive.
- Accuracy of interpretation of visual properties increases quality of decision making
- Illusions and the relation to graphical integrity decreases quality of decision making

Graphical Features



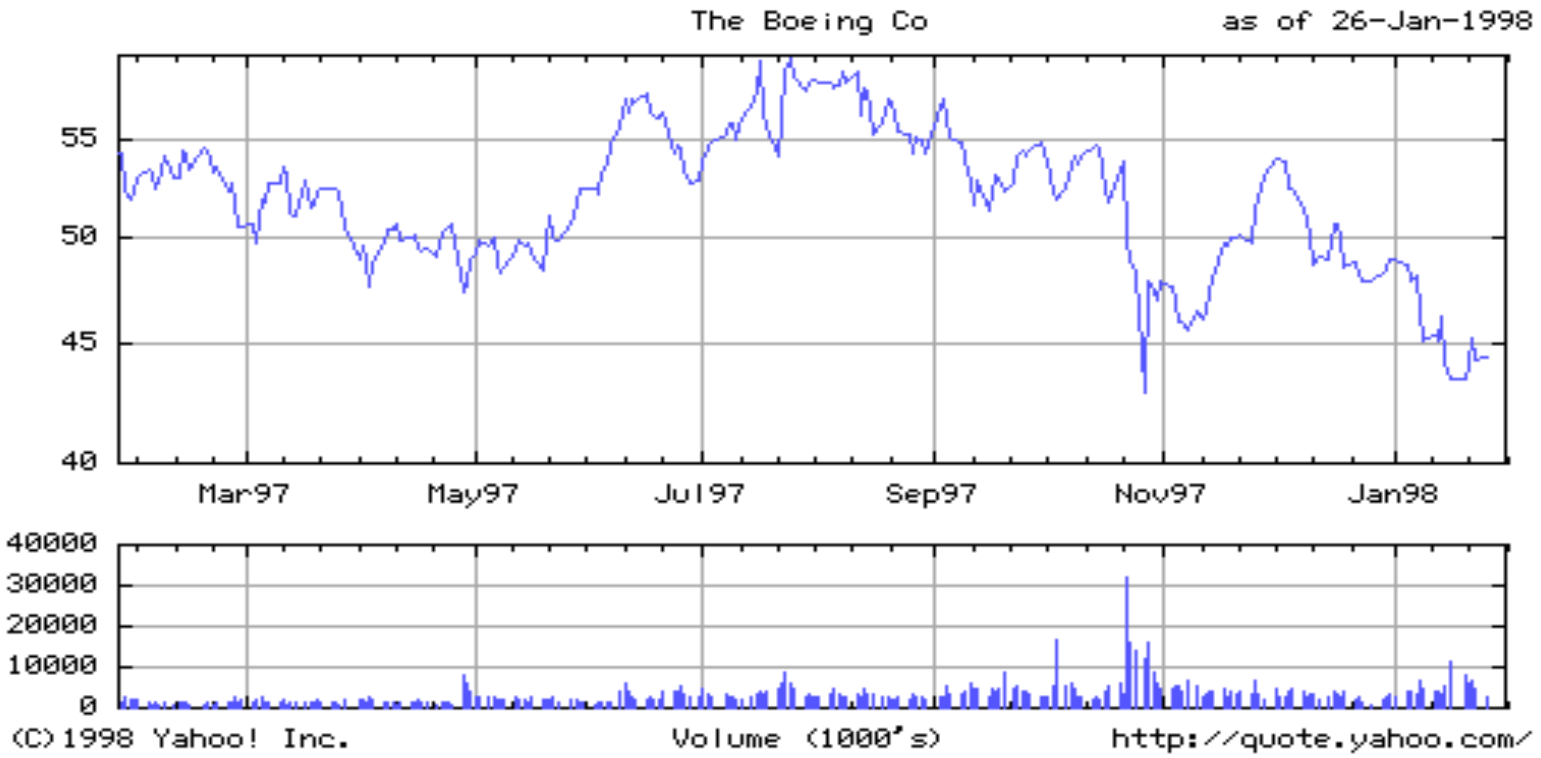
Based on graphic (Figure 2) in Presentation Graphics (white paper) by Leland Wilkinson, SPSS, Inc and Northwestern Uiv.

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Graphs

- At least two dimensions required

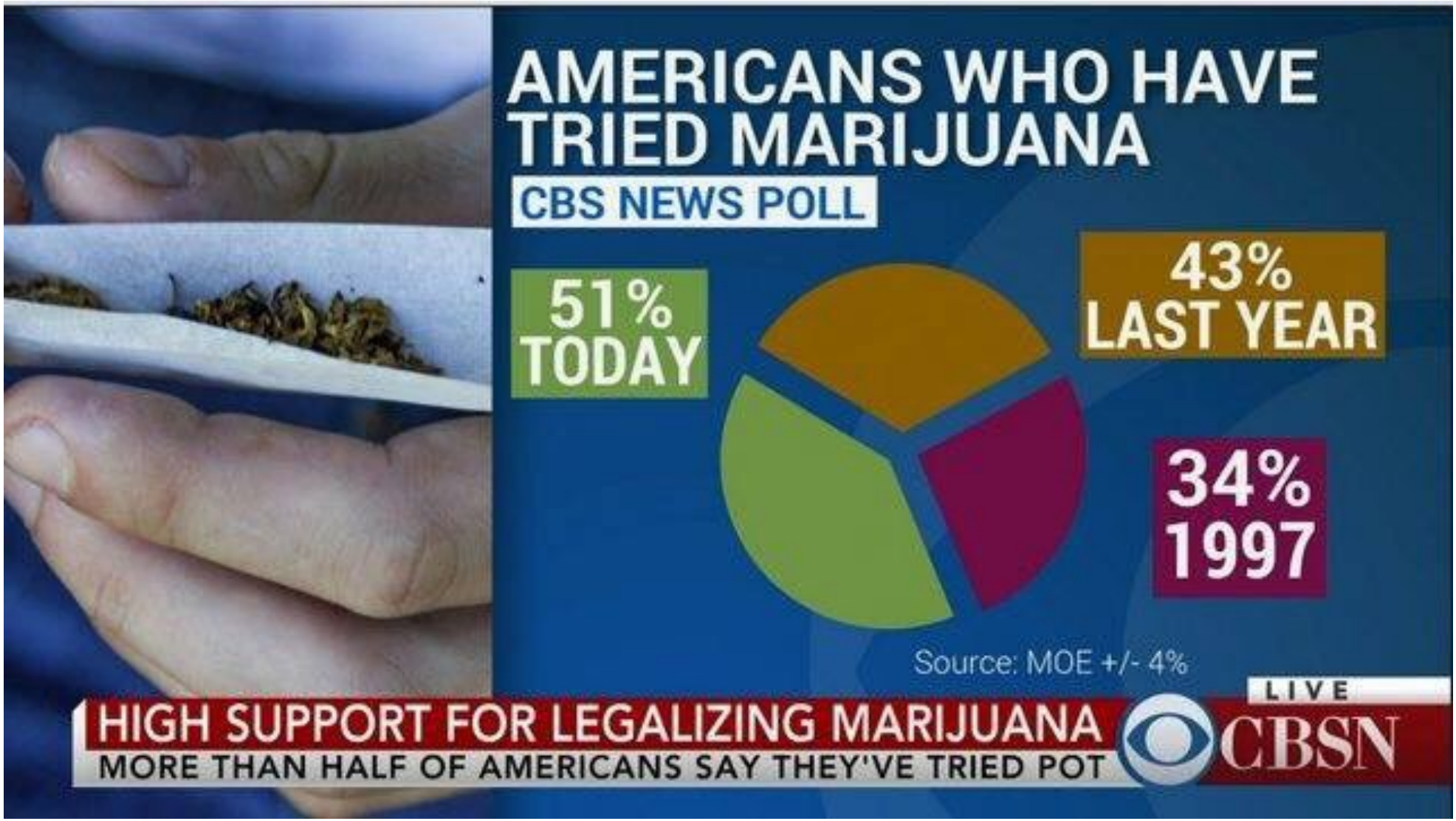
- The dimensions on a graph are related in some way
  - Examples: scatter-plot, bar-chart, layer-graph



NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Graphs: Making Comparisons

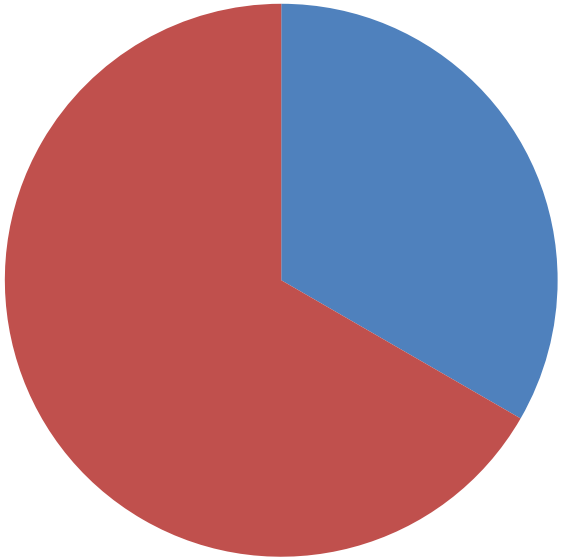
- What is very wrong here?



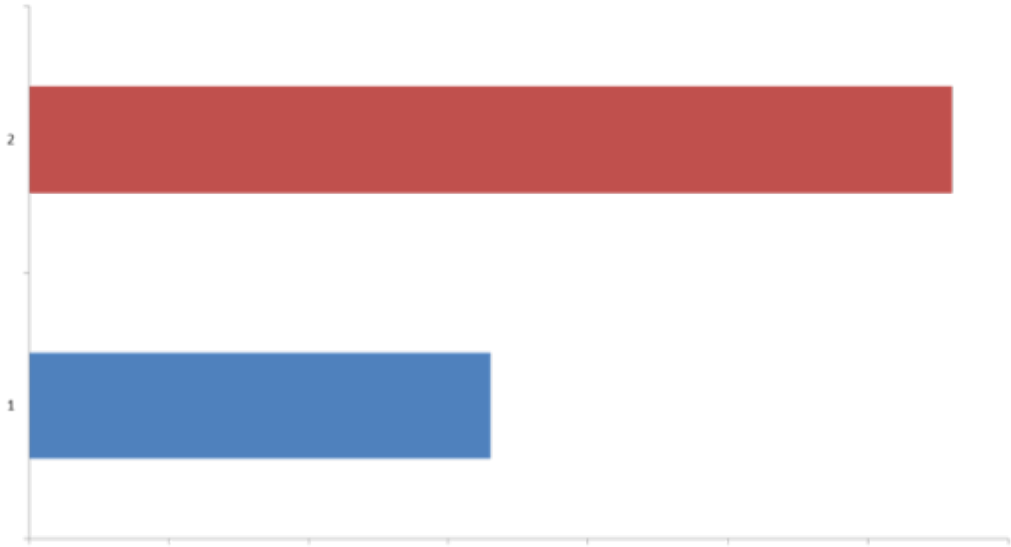
NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Graphs: Making Comparisons

- Percent Blue relative to Red?



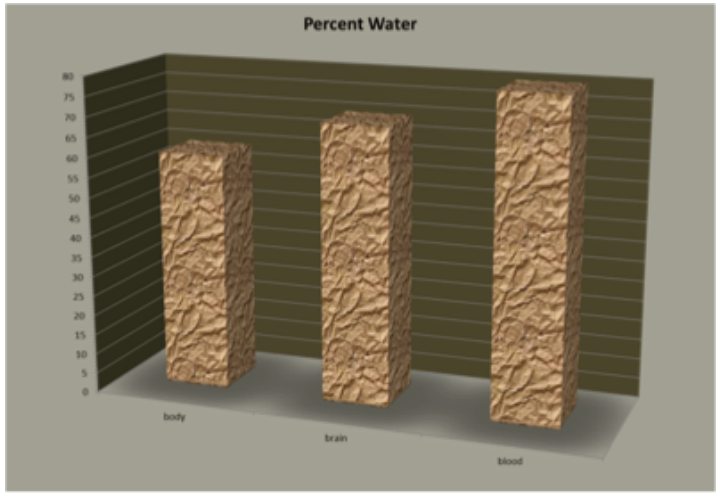
- Which is better?



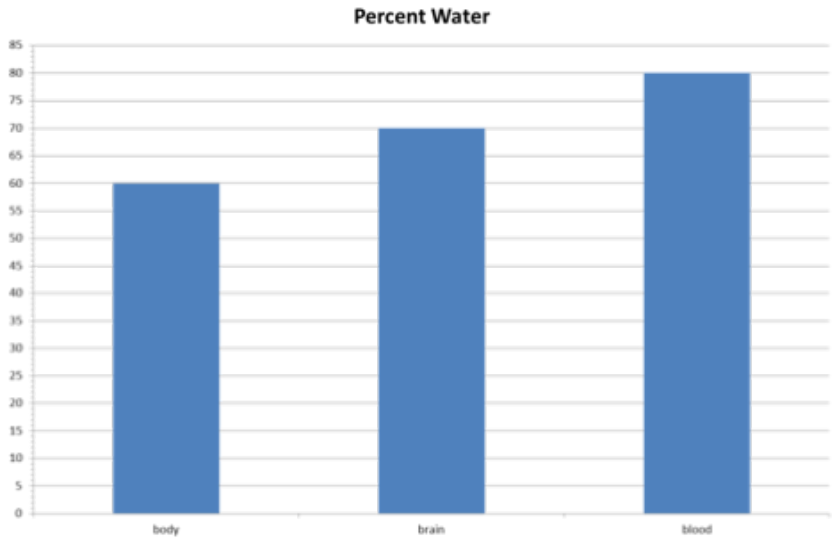
NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

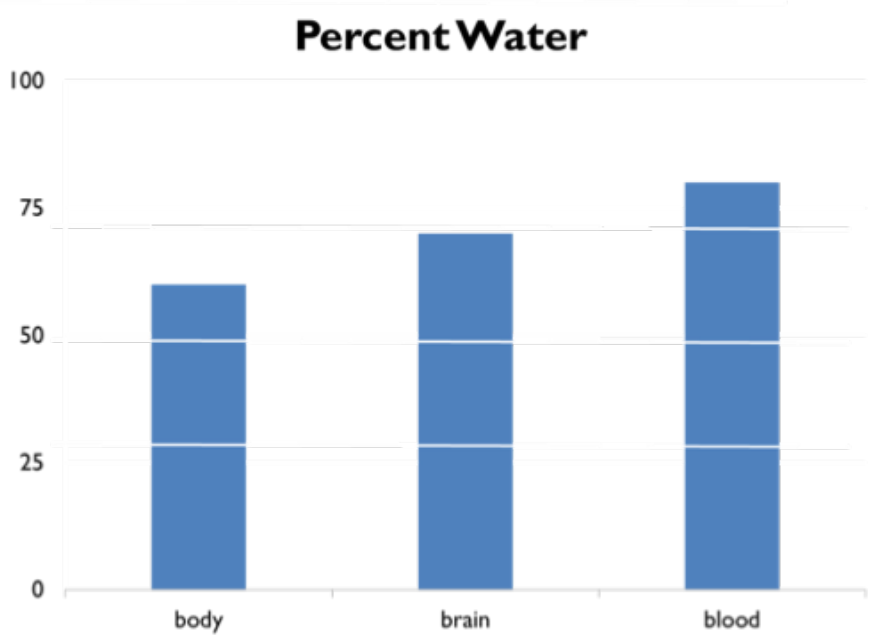
Graphs: Making Comparisons

- Percent water?



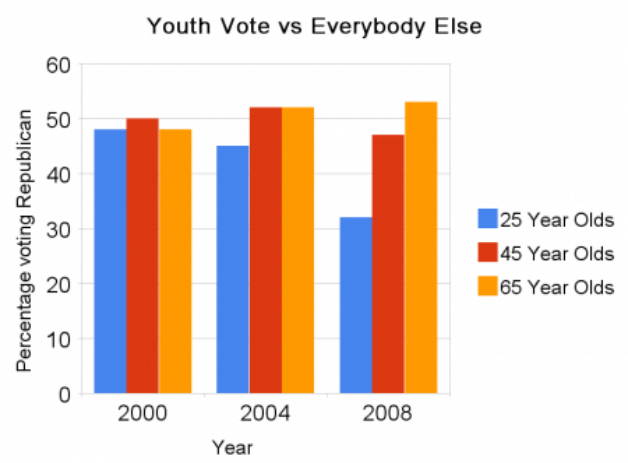
- Which is better?



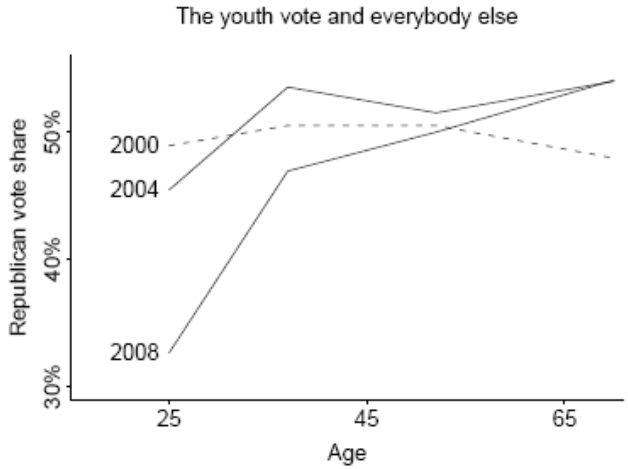


NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

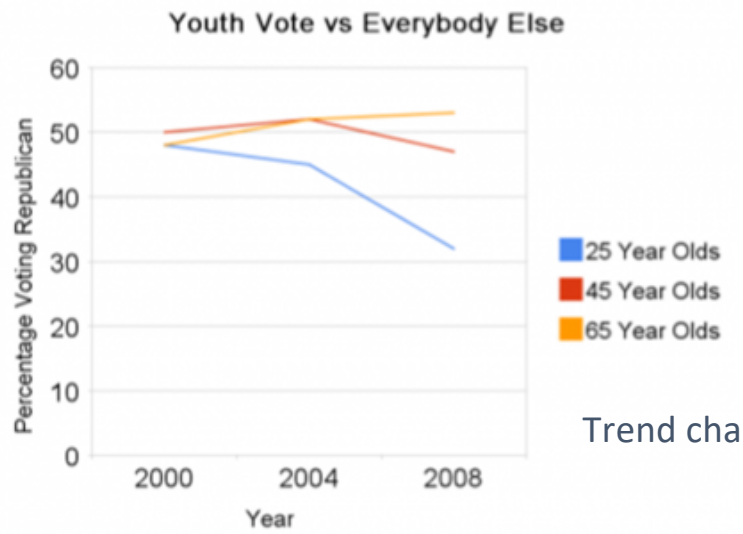
Graphs: Making Comparisons



Bar charts are bad for trends



Trend charts are better; legend is missing

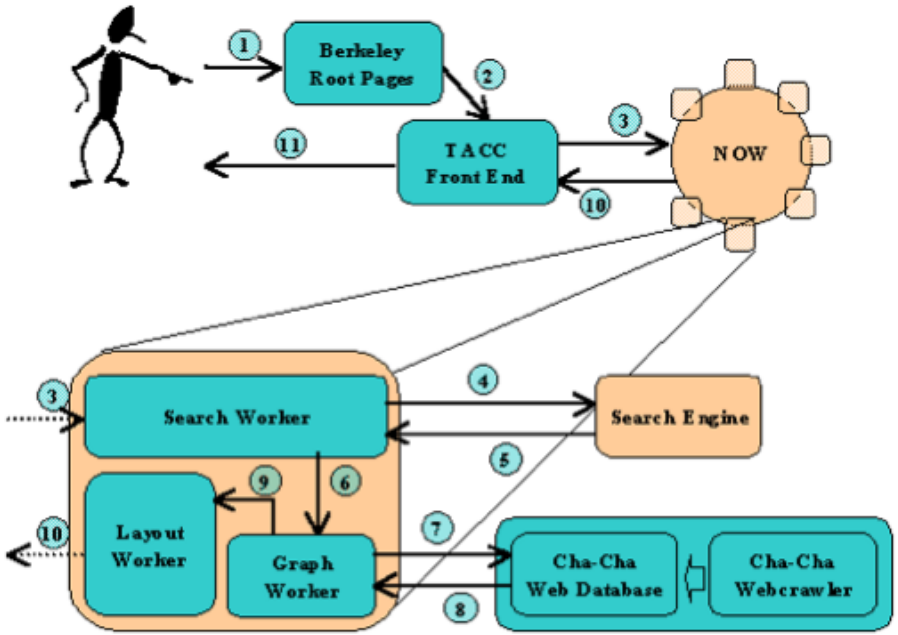


Trend charts are better; with legend

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Charts

- Discrete relations among discrete entities
- Structure relates entities to one another
- Lines and relative position serve as links
  - Examples: Organization structure, Network structures



NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Maps

- Internal relations determined (in part) by the spatial relations

- Labels paired with locations

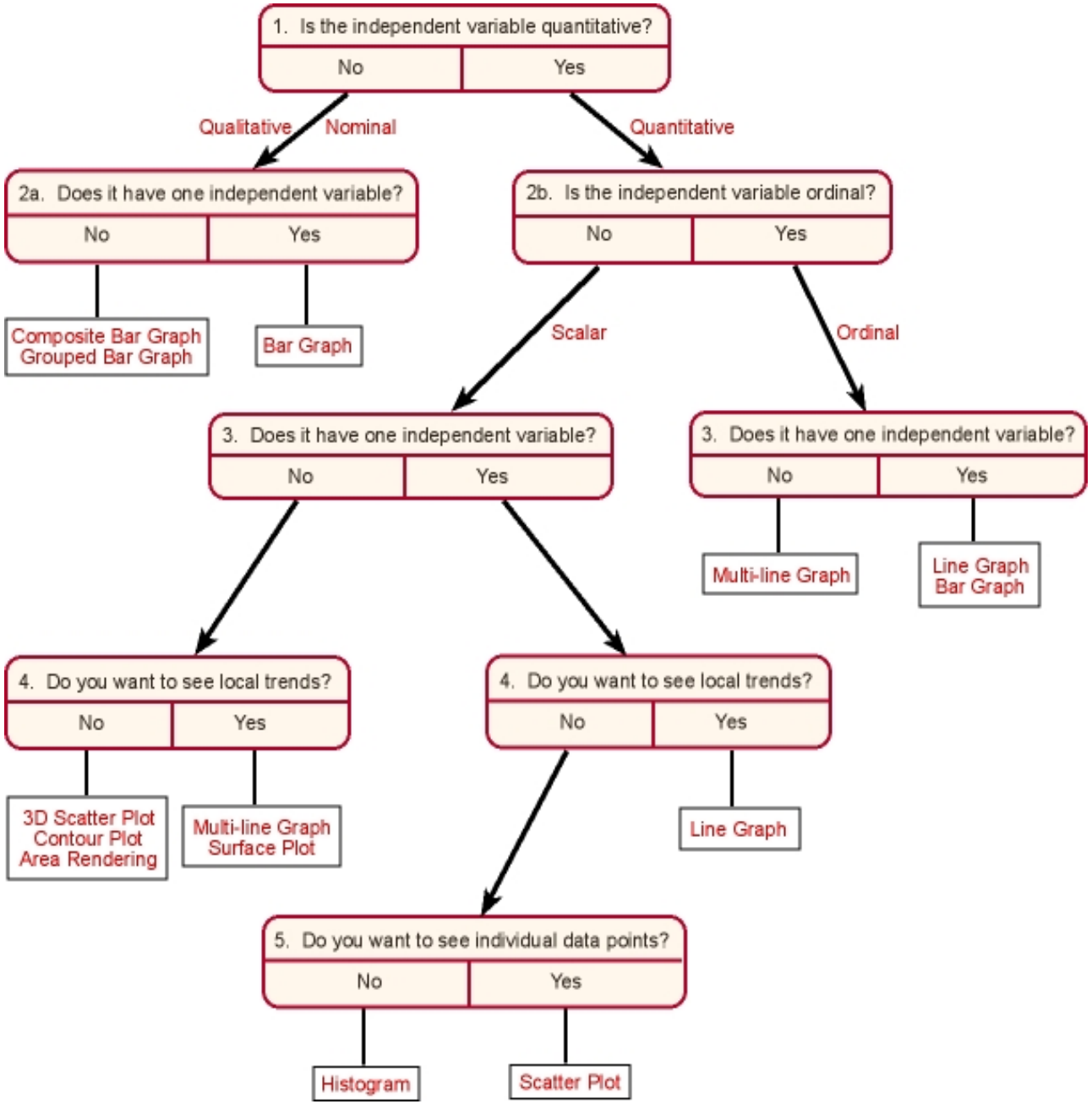
NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Diagrams

- Schematic representation of objects
- Parts may symbolic
  - how-to illustrations
  - figures in a manual
- Examples
  - Circuit diagrams
  - Data flow diagrams
  - Entity-relationship models

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Visuals: A Decision Tree



Visualizations to Avoid

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Chartjunk

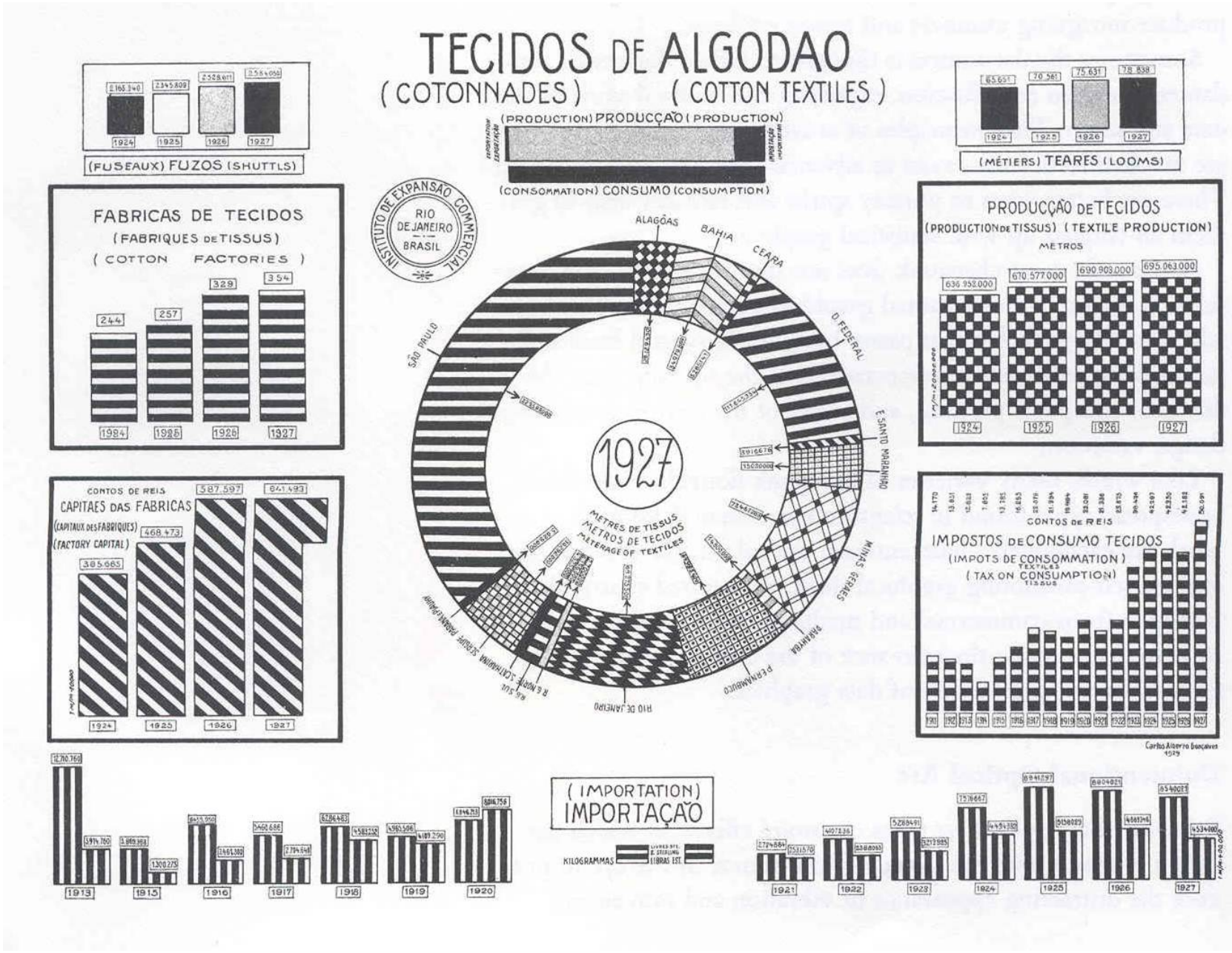
- Chartjunk are non-data-ink or redundant data-ink decoration
  - "In our excitement to produce what we could only make before with great effort, many of us have lost sight of the real purpose of quantitative displays — to provide the reader with important, meaningful, and useful insight.“
    - Stephen Few

- Unintended Optical Art
  - Mainly rely on moiré effects
    - Distracting appearance of vibration and movement
    - The most common form of graphical clutter

- The Grid
  - Dark grid lines are chartjunk
  - The grid should usually be muted or completely suppressed

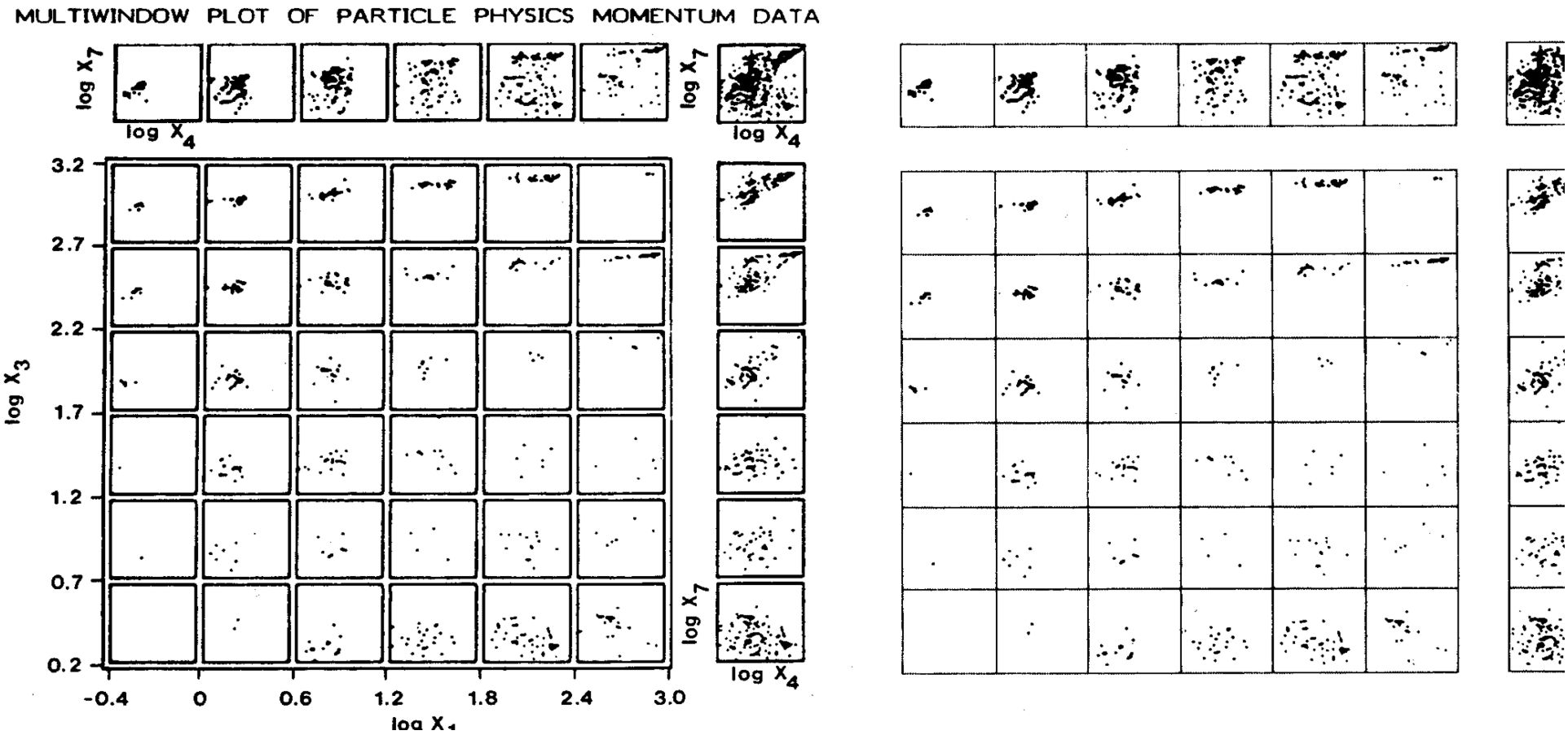
- The Duck
  - Self-promoting Graphics
  - When the data measures become design elements

Moiré Vibrations



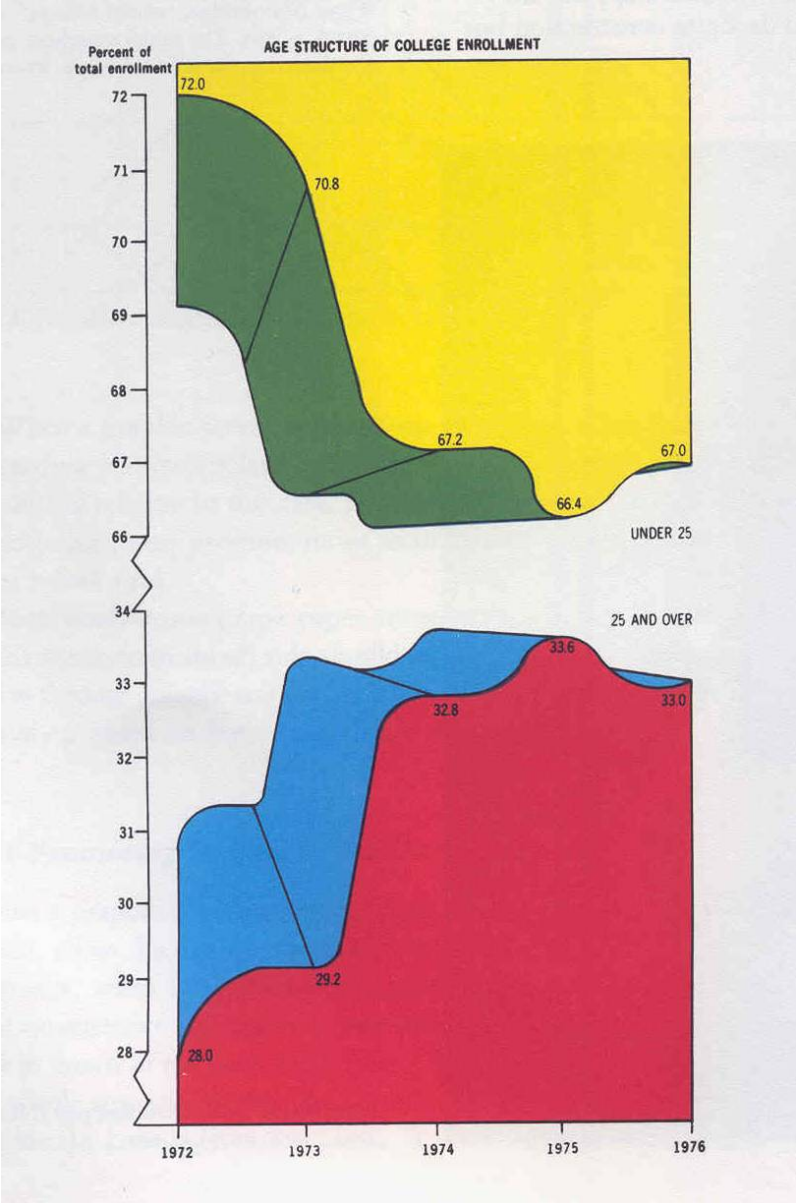
NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

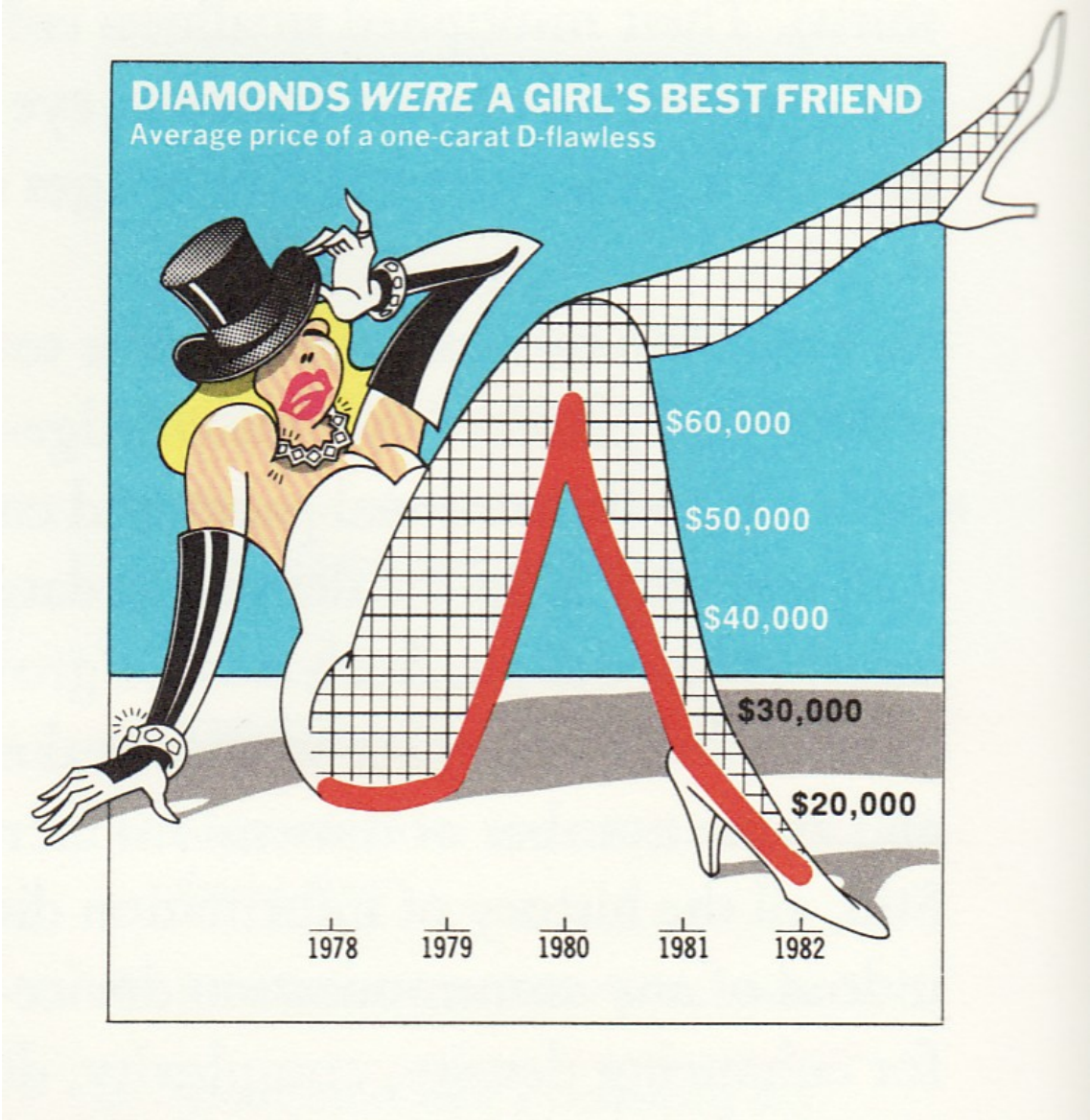
The Grid



NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

The Duck: Examples





Graphical Integrity

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Why do Graphics Lie?

- Lack of quantitative skills of professional artists
- The doctrine that statistical data are boring
- The doctrine that graphics are only for the unsophisticated readers
- *Design is choice. The theory of the visual display of quantitative information consists of principles that generate design options and that guide choices among options. The principles should not be applied rigidly or in a peevish spirit; they are not logically or mathematically certain; and it is better to violate any principle than to place graceless or inelegant marks on paper.*
  - *Edward Tufte, The Visual Display of Quantitative Information*

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Graphical Integrity

- Graphical excellence begins with telling the truth about the data

- Principles
  1. The representation of numbers, as physically measured on the surface of the graphics, should be directly proportional to the numerical quantities represented

  2. Clear, detailed and thorough labeling should be used to defeat distortion

  3. Show data variation, not design variation

  4. The number of information-carrying dimensions depicted should not exceed the number of dimensions in the data

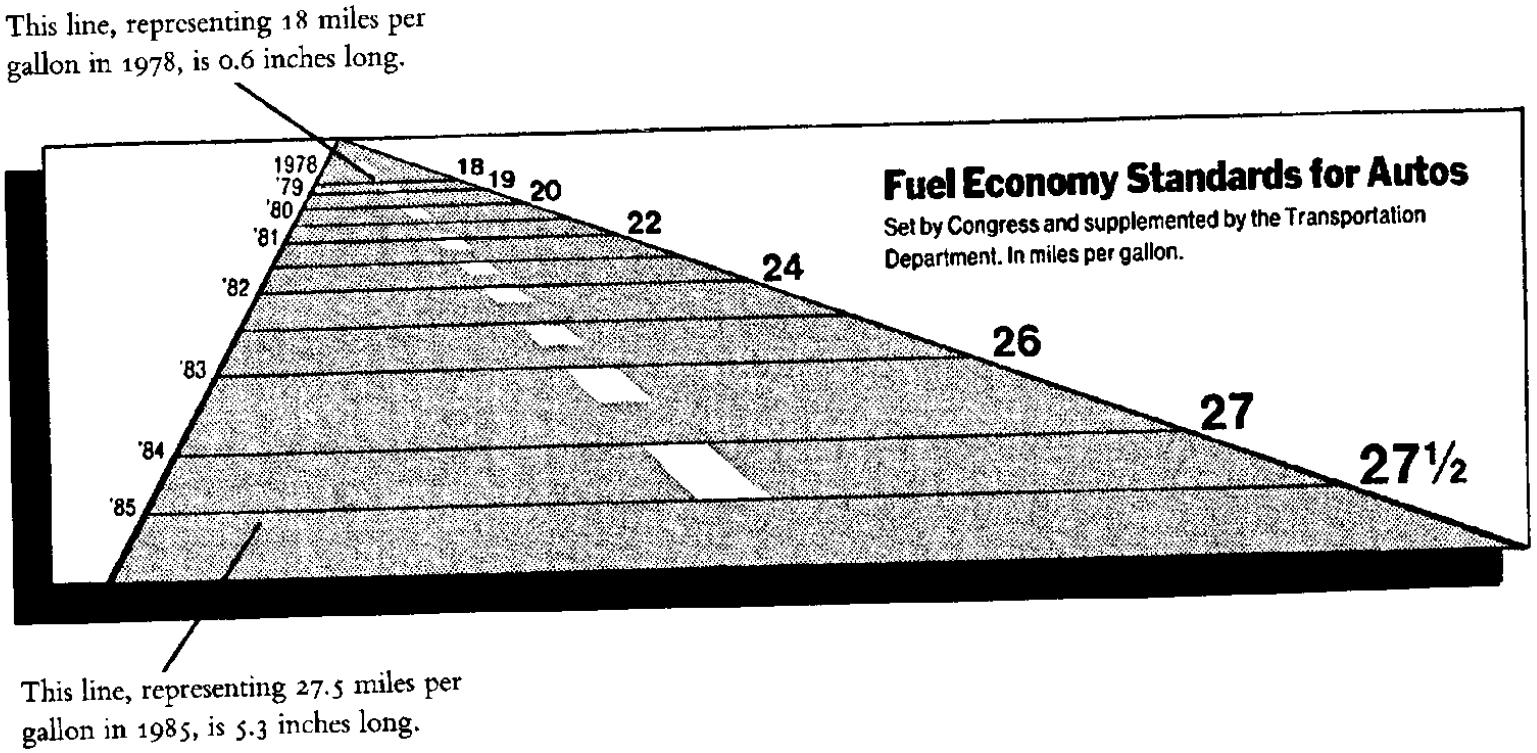
  5. Graphics should not quote data out of context

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Violating Principle 1

- Principle 1 can be measured by Lie factor:
  - Lie Factor = (size of effect shown in graphics) / (size of effect in data)
  - Lie Factor equal to one is ideal
  - Above graphic has a lie factor of 14.8

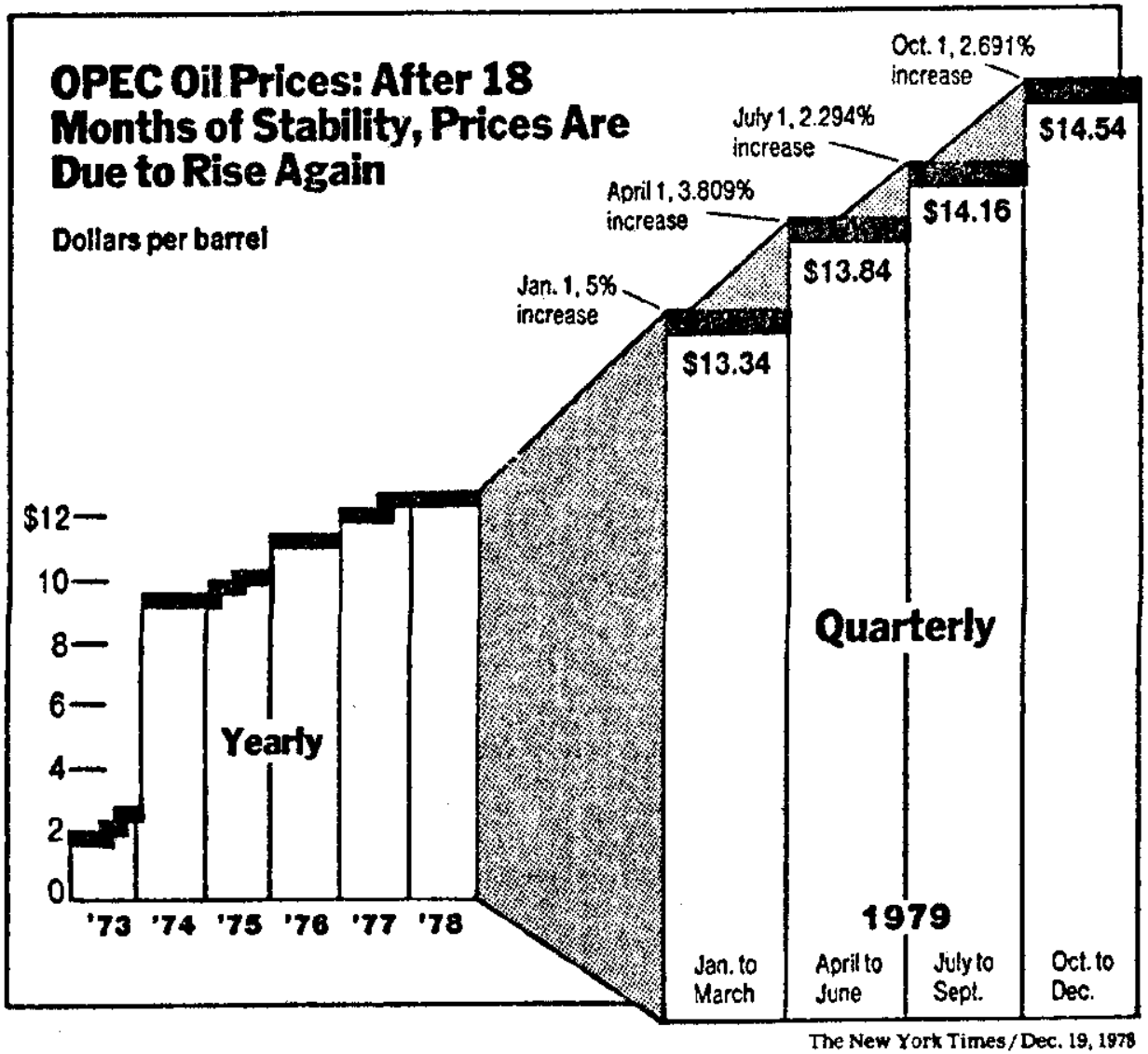
- 18 miles/gallon: 0.6 inches; 27.5miles/gallon: 5.3 inches



NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Design and Data Variation

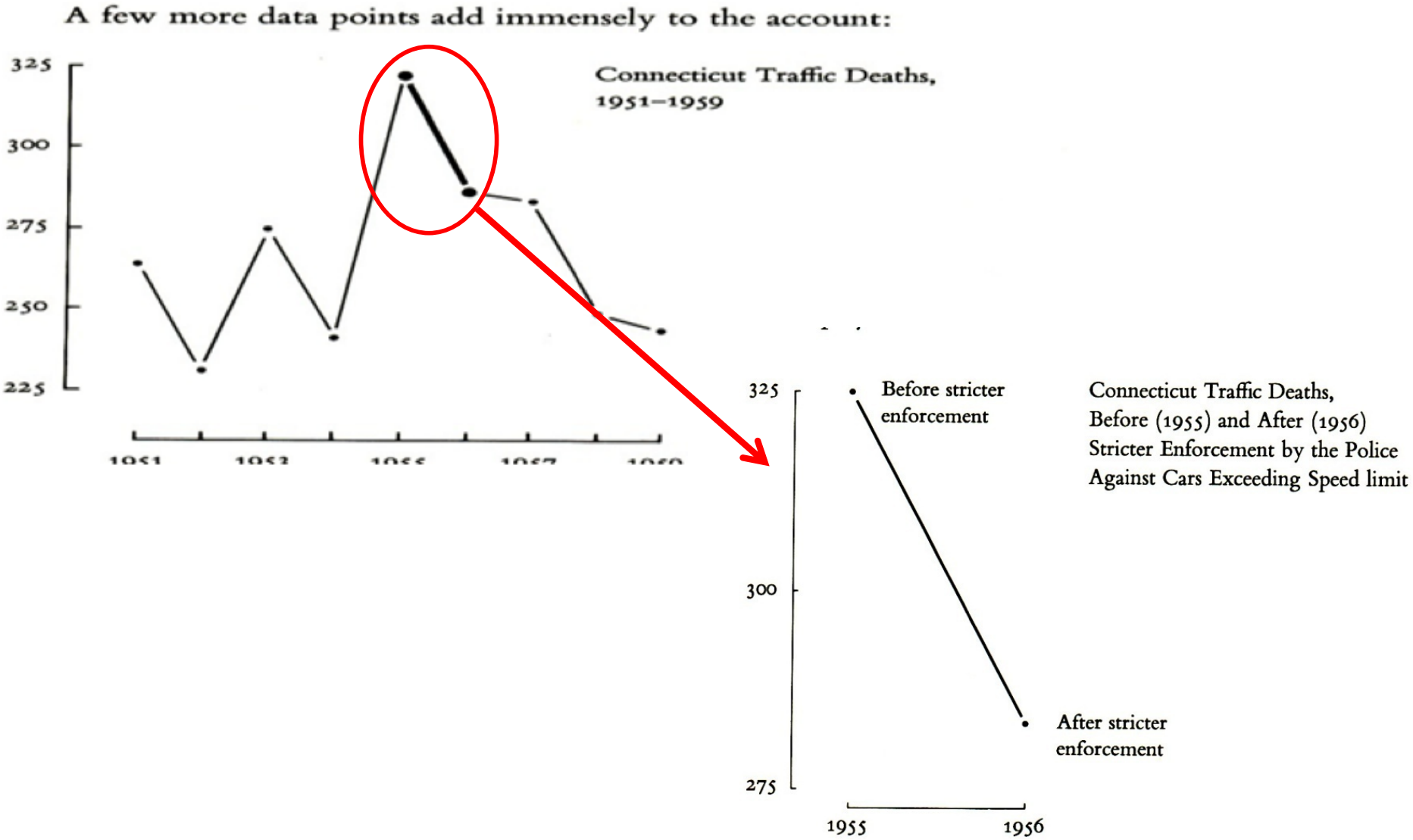
- Show data variation, not design variation
  - 1973-1978: one vertical inch equals to $8.00.
  - 1979: one vertical inch equals $3-4
  - 1973-1978: one horizontal inch equals 3.7 years
  - 1979: one horizontal equals 0.57 year



NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

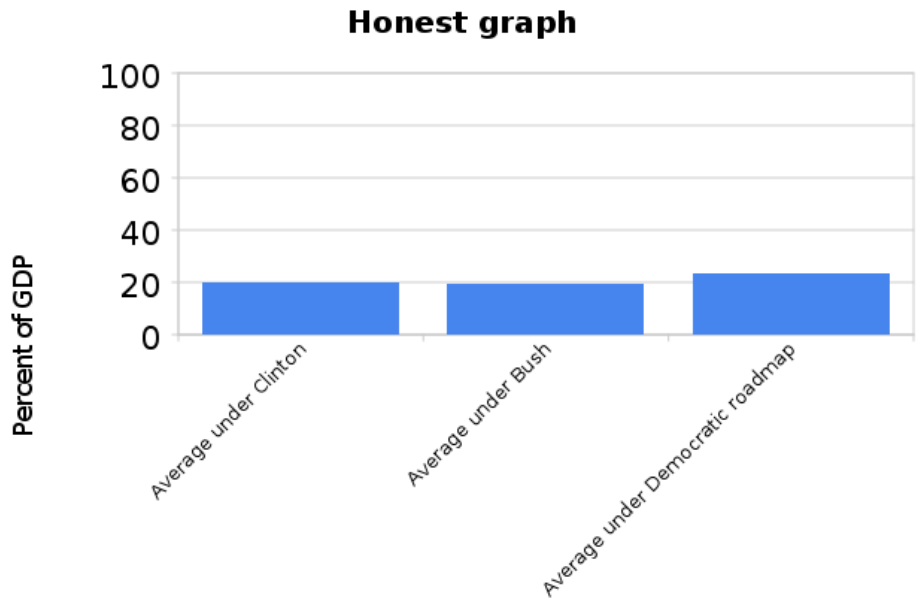
Context is Essential

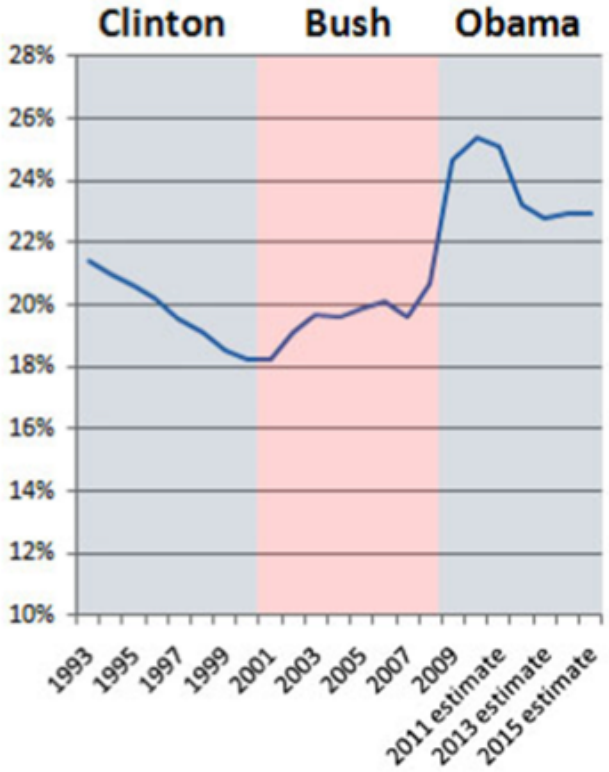
- Graphics must not quote data out of context



NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Accentuating Scales





NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Five Use Cases *(from Gartner)*

Agile Centralized BI	• **Provisioning.** Supports an agile IT-enabled workflow — from data to centrally delivered and managed content — using the self-contained data management capabilities of the platform.
Decentralized Analytics.	• Supports a workflow from data to self-service analytics.
Governed Data Discovery.	• Supports a workflow from data to self-service analytics, to systems-of-record, IT-managed content with governance, reusability and promotability.
Embedded BI.	• Supports a workflow from data to embedded BI content in a process or application.
Extranet Deployment.	• Supports a workflow similar to agile centralized BI provisioning for the external customer or, in the public sector, citizen access to analytic content.

Analytics and Content Creation *(from Gartner)*

Embedded Advanced Analytics.

- Enables users to easily access advanced analytics capabilities that are self-contained within the platform itself or available through the import and integration of externally developed models.

Analytic Dashboards.

- The ability to create highly interactive dashboards and content, with visual exploration and embedded advanced and geospatial analytics, to be consumed by others.

Interactive Visual Exploration.

- Enables the exploration of data via the manipulation of chart images, with the color, brightness, size, shape and motion of visual objects representing aspects of the dataset being analyzed. This includes an array of visualization options that go beyond those of pie, bar and line charts, to include heat and tree maps, geographic maps, scatter plots and other special-purpose visuals. These tools enable users to analyze the data by interacting directly with a visual representation of it.

Mobile Exploration and Authoring.

- Enables organizations to develop and deliver content to mobile devices in a publishing and/or interactive mode, and takes advantage of mobile devices' native capabilities, such as touchscreen, camera, location awareness and natural-language query

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Gartner Magic Quadrant 2016

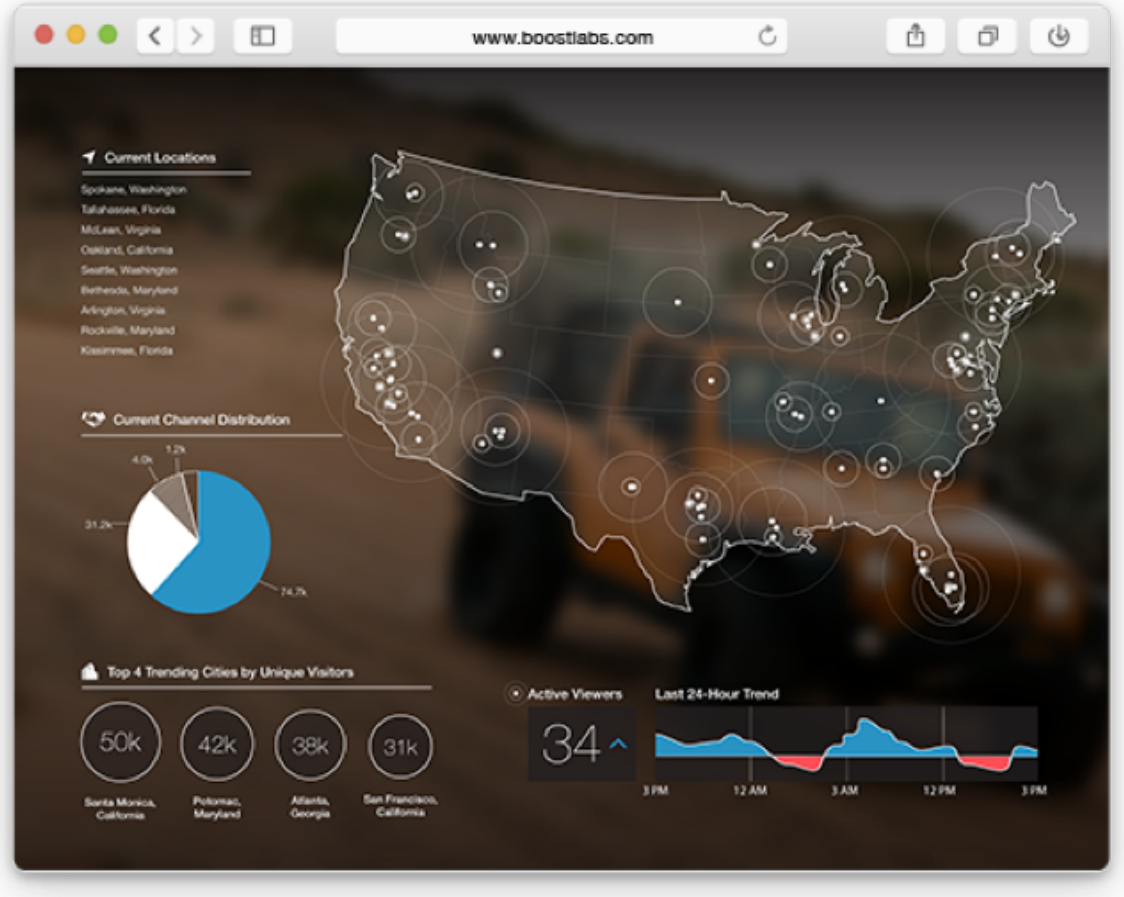
NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

BI Applications - Prototyping

Source: http://www.boostlabs.com/ui-data-visualization-design-services-dc-md-va/

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

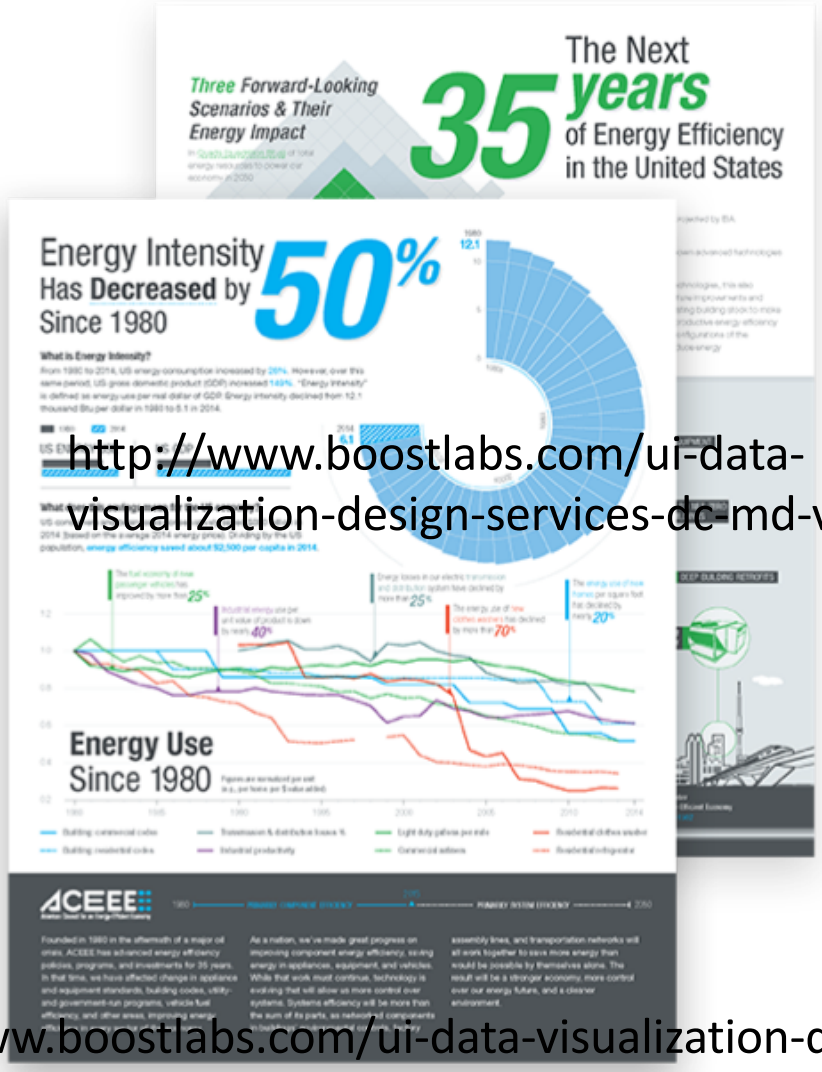
Dashboards



Source: http://www.boostlabs.com/ui-data-visualization-design-services-dc-md-va/

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Infographics



http://www.boostlabs.com/ui-data-visualization-design-services-dc-md-va/

Source: http://www.boostlabs.com/ui-data-visualization-design-services-dc-md-va/

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

Automated Reports

Source: http://www.boostlabs.com/ui-data-visualization-design-services-dc-md-va/

NANYANG TECHNOLOGICAL UNIVERSITY
Nanyang Business School

References for Visual Principles

- Types of Visual Representations:
  - Kosslyn, S. M. (1989). Understanding charts and graphs. *Applied Cognitive Psychology*, 3, 185–226.
- How do people perceive common graphic displays
  - Lohse, G L; Biolsi, K; Walker, N and H H Rueter (1994). A Classification of Visual Representations, *Communications of the ACM*, 37:12, pp 36-49, 1994
- Perceptual properties and visual features
  - Bertin: see reference in an earlier slide
- How to mislead with graphs
  - Tufte: see references (there are at least two) in earlier slides

Questions and Answers