

S. Rinzivillo – rinzivillo@isti.cnr.it

DATA VISUALIZATION AND VISUAL ANALYTICS

Who I Am?

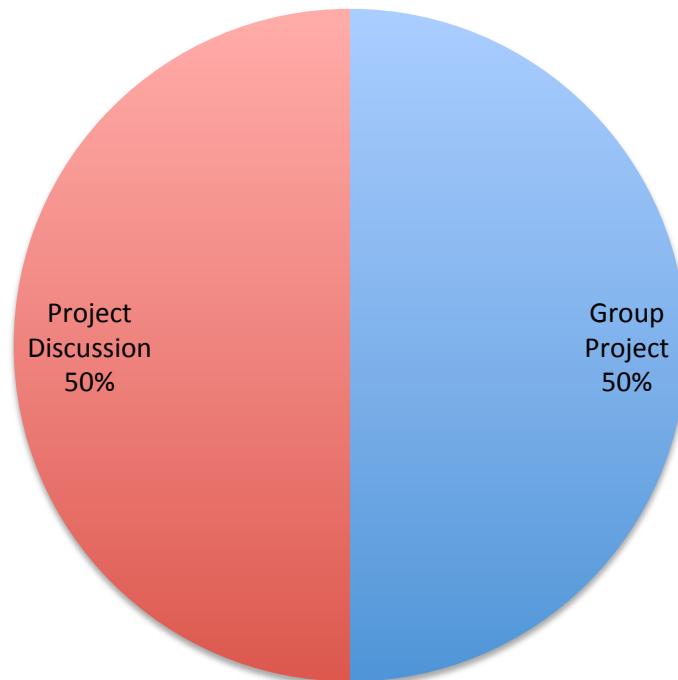
- Salvatore Rinzivillo
 - rinzivillo@isti.cnr.it
- Page course: <http://didawiki.cli.di.unipi.it/>
 - Visual Analytics

Schedule

- On Monday
 - 14:00 to 16:00
 - Room: N1
- On Friday
 - 14:00 to 16:00
 - Room: C1

Grading

- Project (50%)
 - Up to 2 persons per group (!)
- Project discussion (50%)
- Project topic
 - Multidimensional exploration of a dataset
 - One (or two) dataset(s) assigned for all
 - Specific proposal may be discussed

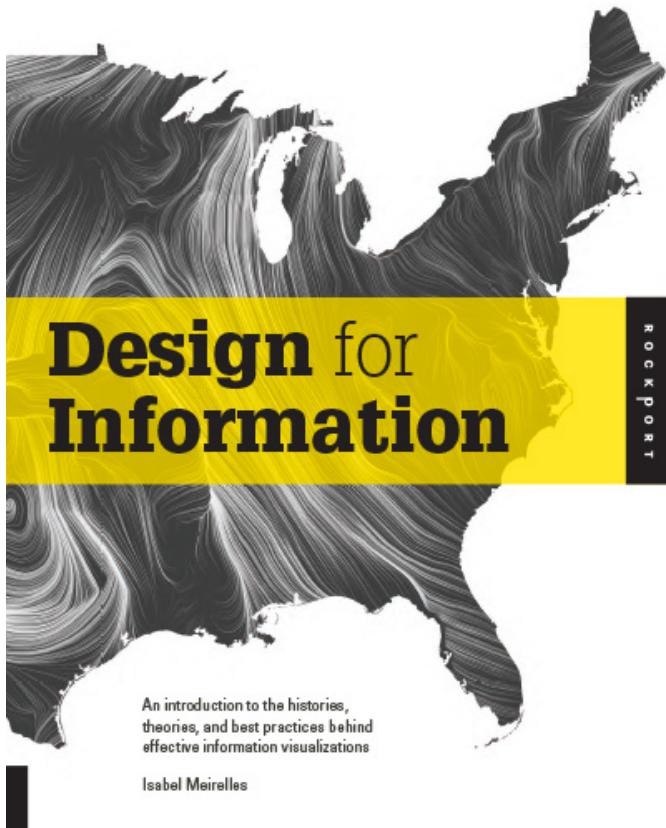


<http://itisaasta.com/nycs/>

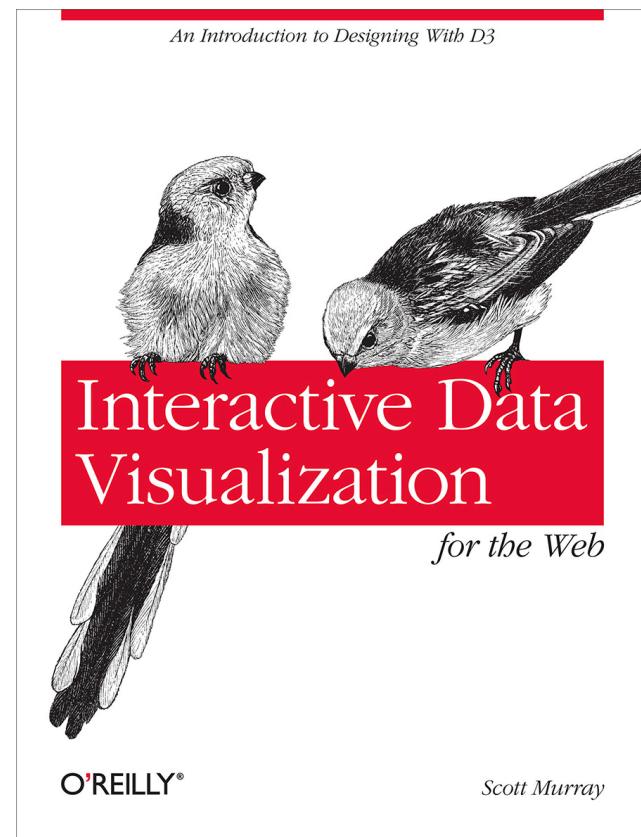
EXAMPLE SCHOOL DISTRICTS

Textbooks

Design for Information Isabel Meirelles



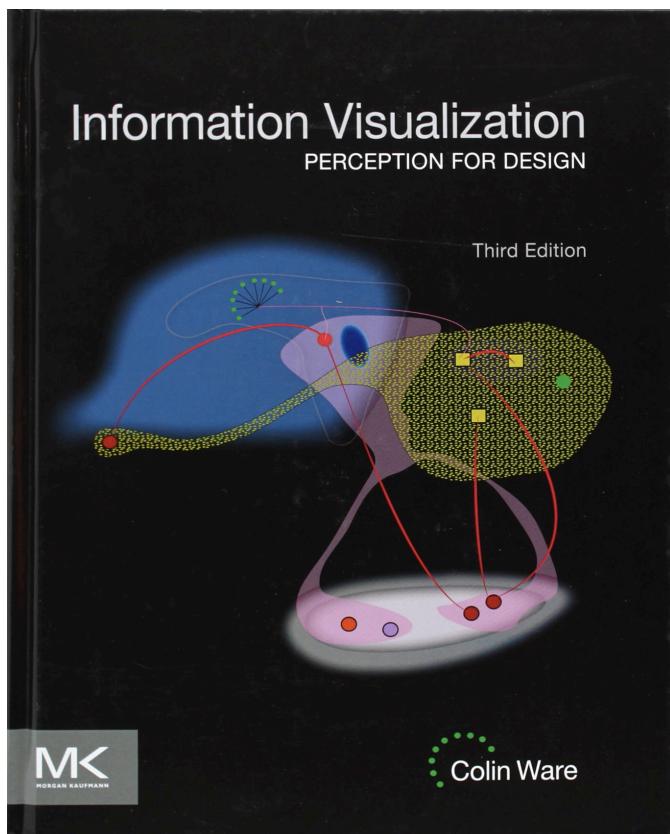
Interactive Data Visualization Scott Murray



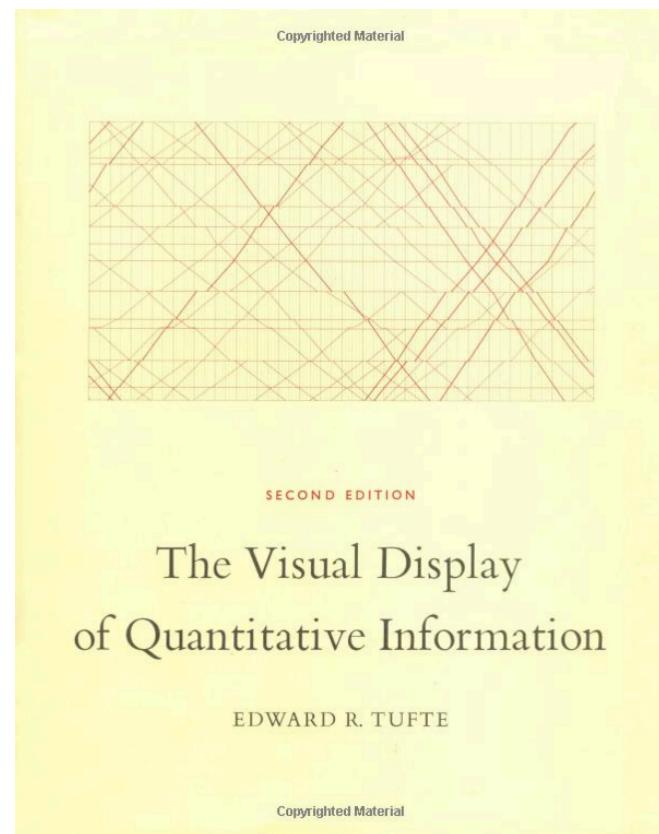
<http://alignedleft.com/tutorials>

Interesting Readings

Information Visualization
Colin Ware



The Visual Display of Visual Information
Edward R. Tufte





DATA VISUALIZATION AND VISUAL ANALYTICS INTRODUCTION

VA - Crash course

- Effective Visual Representation
 - Vision System
 - Visual Variables
- Toolbox – Bootstrap, Node.js, Vue.js, crossfilter.js
- Toolbox – Base visualizations (NVD3, DC.js)
- Toolbox – D3.js
 - Basics
 - Charts
 - Advanced Visualization
- Scientific Visualization
 - Plotting
 - Geography
- Storytelling

Data Visualization

Convey Information through
graphical representation of data

Motivations

- Data everywhere
- No value for raw data
 - Need to extract valuable information
- Information overload:
 - Irrelevant for current task
 - Processed in an inappropriate way
 - Presented in an inappropriate way

Visualization Goal

- Record Information
 - Sketches, photographs, ECG,...
- Analyze data to support decisions (**exploration**)
 - Create and verify hypotheses
 - Identify Patterns
 - Identify Outliers
- Communicate (**explanation**)
 - Share or highlight insights on data
 - Persuade

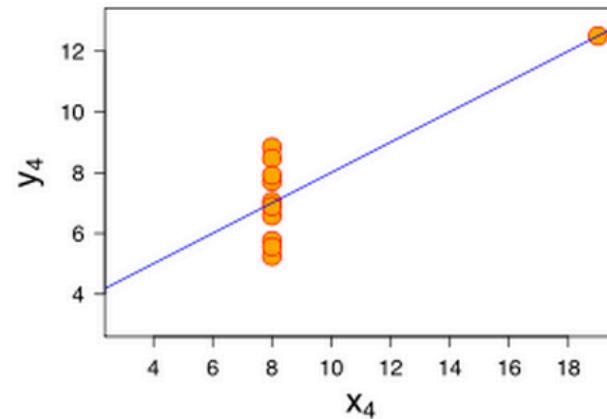
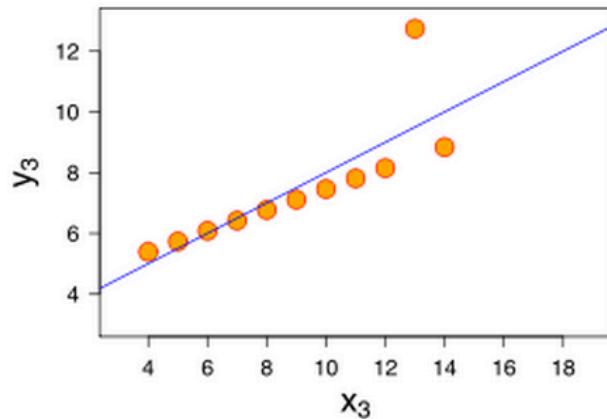
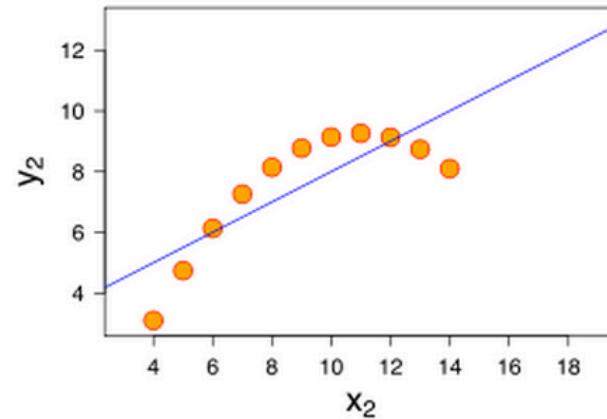
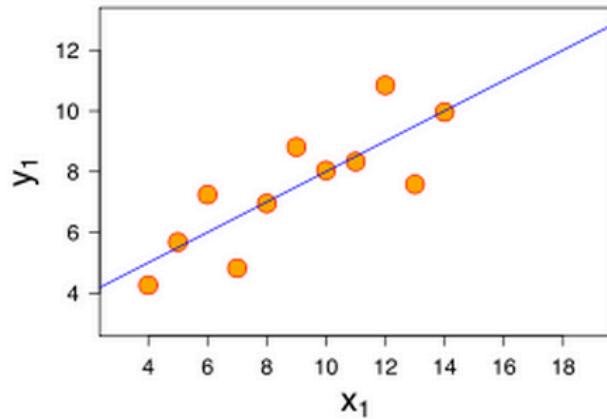
Analyze: Anscombe's quartet - datasets

Data Set A		Data Set B		Data Set C		Data Set D	
X	Y	X	Y	X	Y	X	Y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

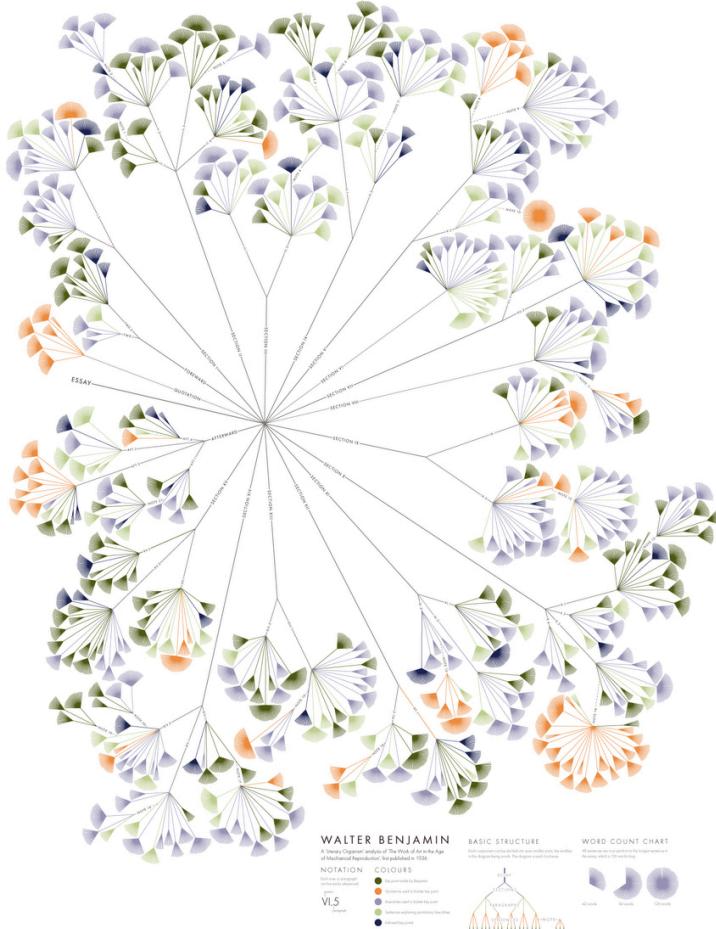
Analyze: Anscombe's quartet - properties

Property	Value
Mean of x in each case	9 (exact)
Sample variance of x in each case	11 (exact)
Mean of y in each case	7.50 (to 2 decimal places)
Sample variance of y in each case	4.122 or 4.127 (to 3 decimal places)
Correlation between x and y in each case	0.816 (to 3 decimal places)
Linear regression line in each case	$y = 3.00 + 0.500x$ (to 2 and 3 decimal places, respectively)

Analyze: Anscombe's quartet – graphics

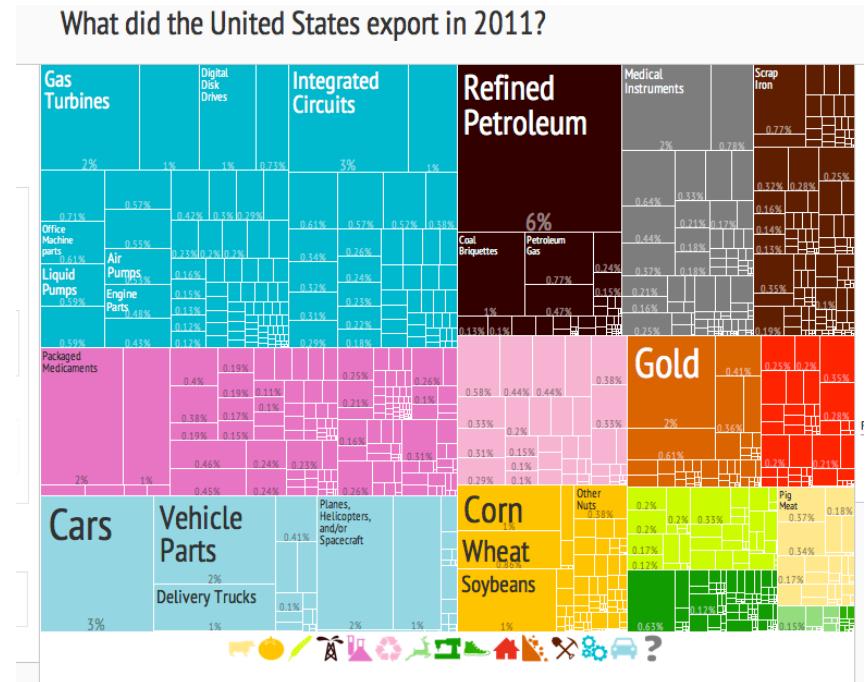


Communicate: Hierarchical Structures



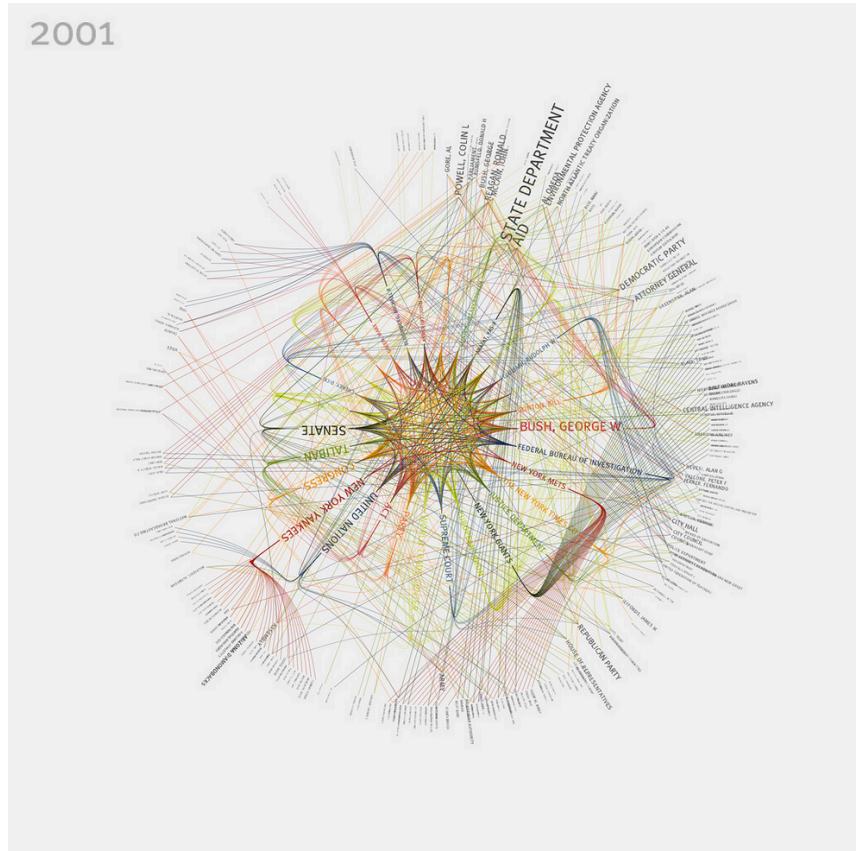
<http://www.stefanieposavec.co.uk/entangled-word-bank/>

<http://atlas.media.mit.edu/>

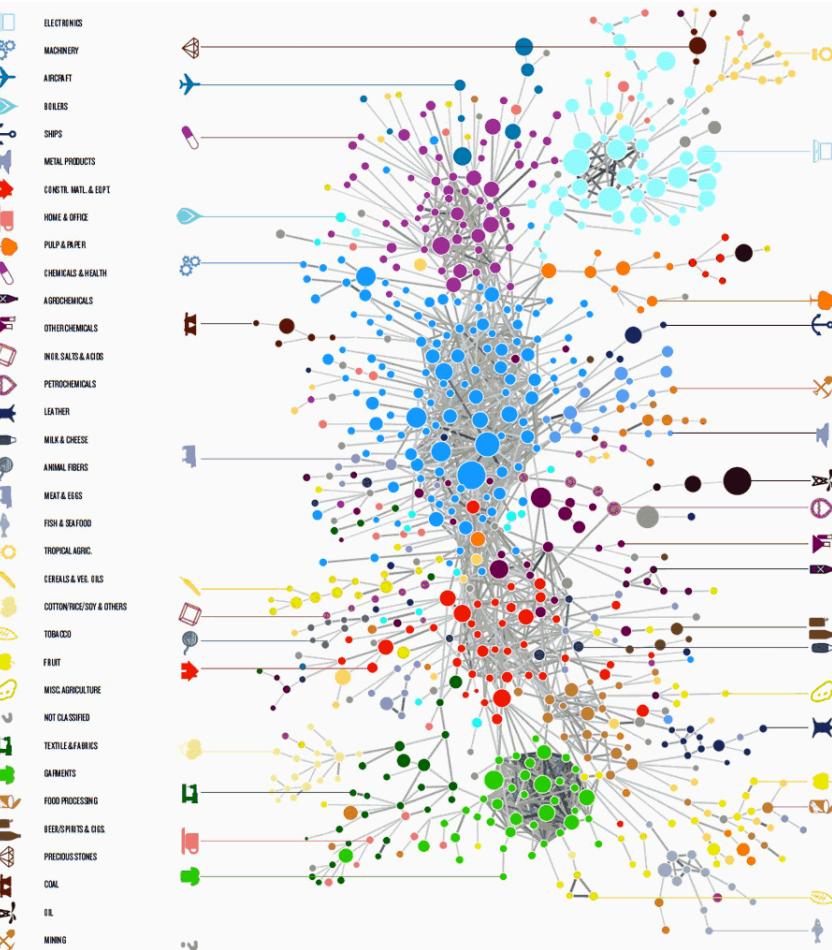


Communicate: Networks

2001

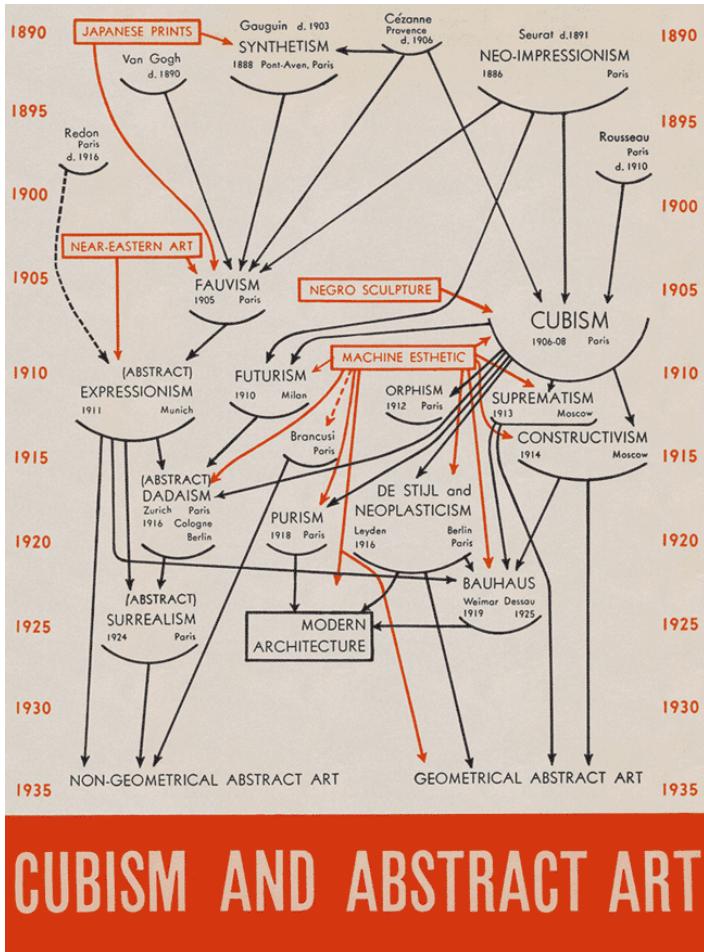


<https://www.flickr.com/photos/blprnt/sets/72157614008027965/>



<http://atlas.media.mit.edu/>

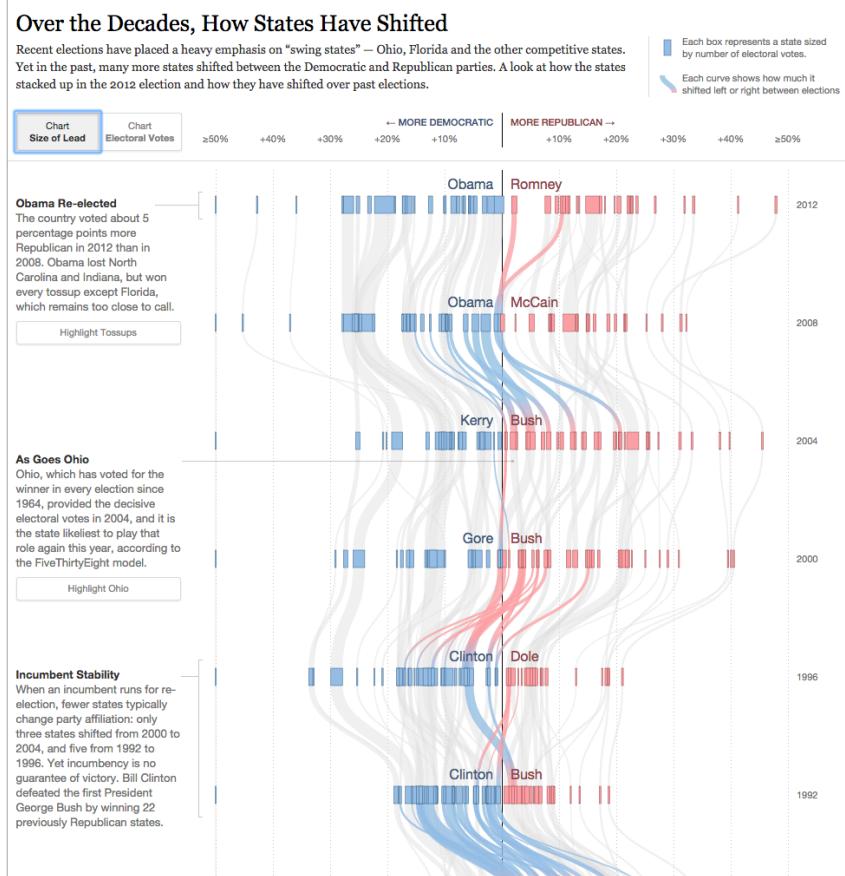
Communicate: Temporal Structures



Cubism And Abstract Art (Alfred H. Barr 1936)

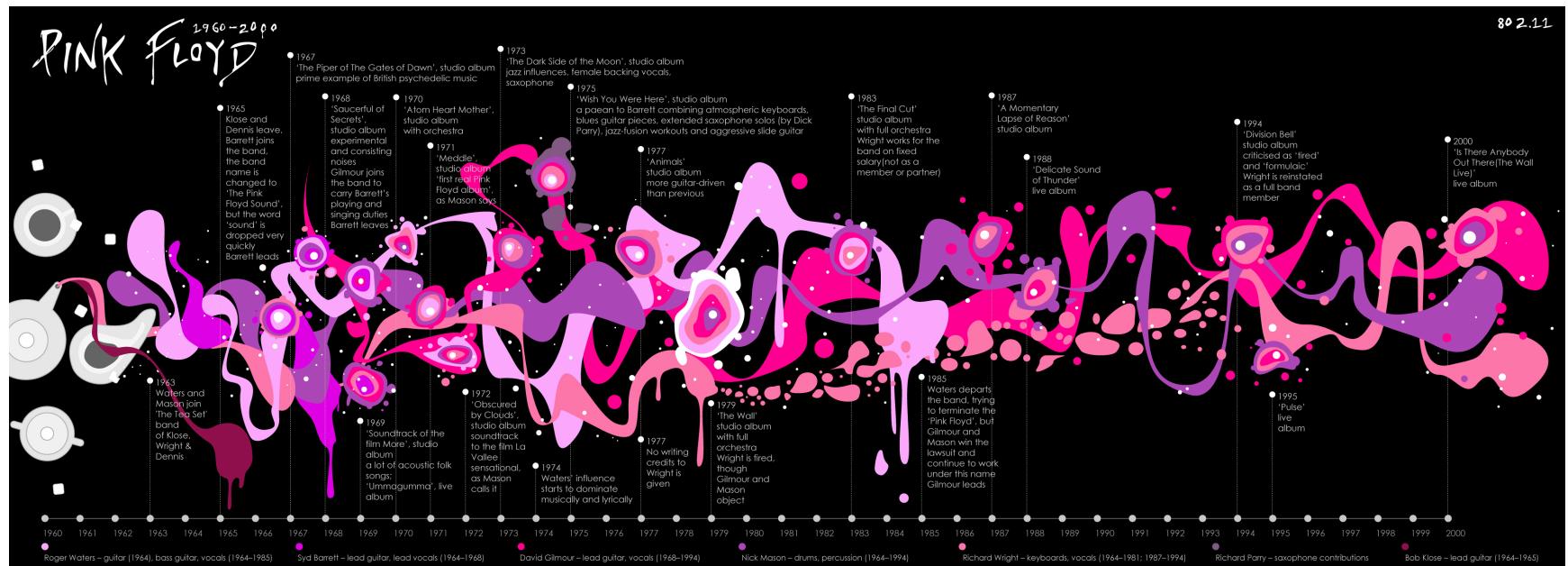
Over the Decades, How States Have Shifted

Recent elections have placed a heavy emphasis on “swing states” — Ohio, Florida and the other competitive states. Yet in the past, many more states shifted between the Democratic and Republican parties. A look at how the states stacked up in the 2012 election and how they have shifted over past elections.



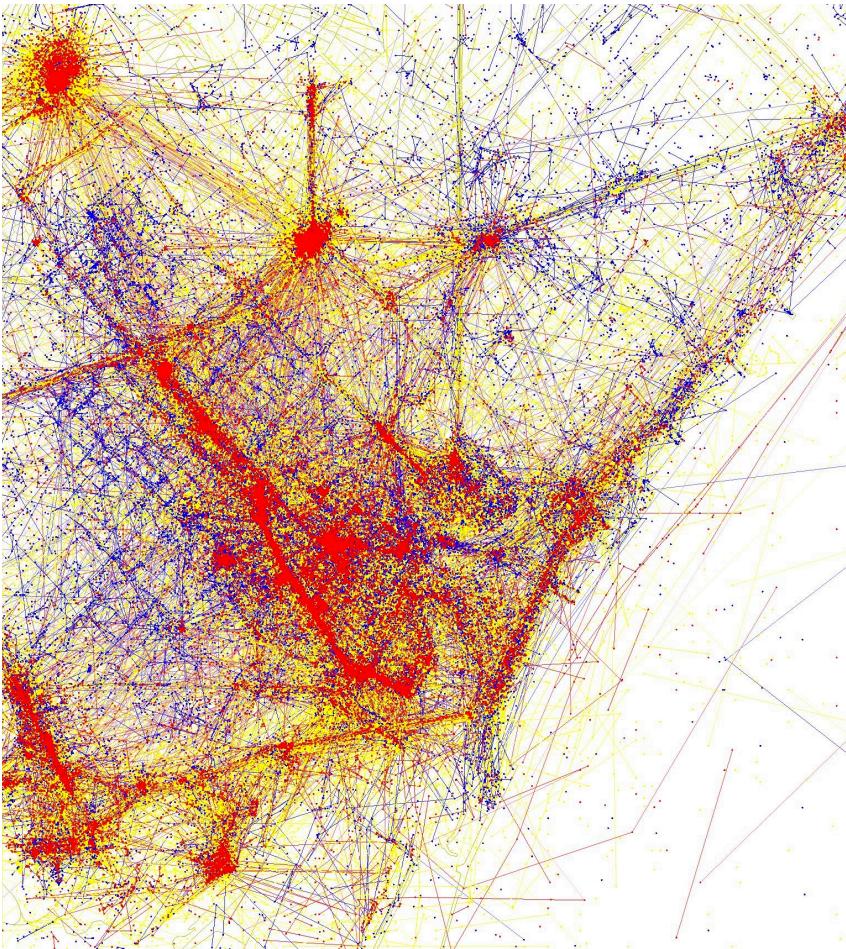
<http://www.nytimes.com/interactive/2012/10/15/us/politics/swing-history.html>

Communicate: Temporal Structures

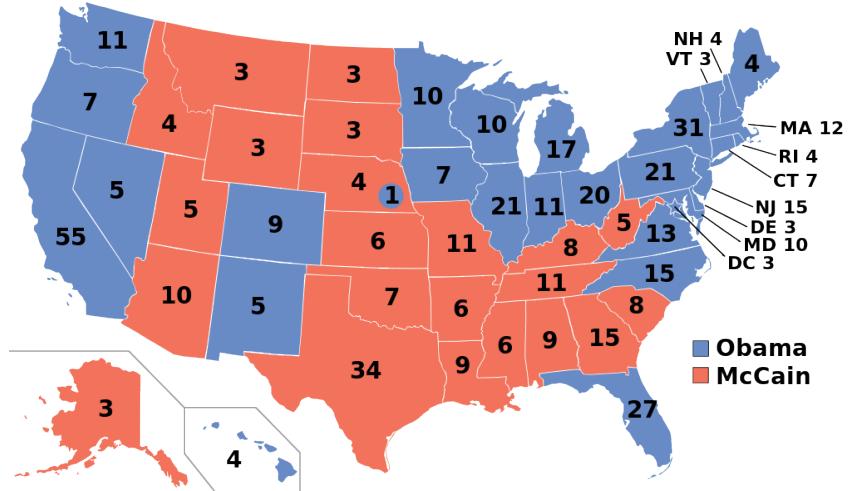


<http://www.80211.cc/>

Communicate: Maps

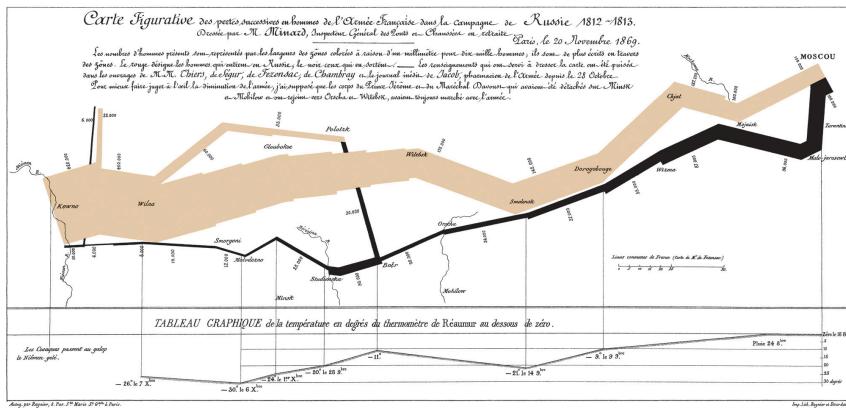


<https://www.flickr.com/photos/walkingsf/sets/72157624209158632/>

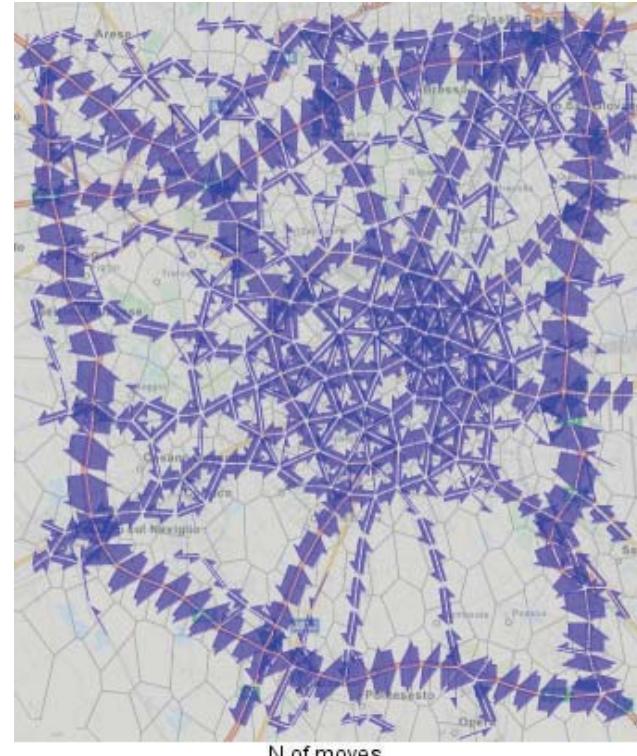


"ElectoralCollege2008" by Gage - Own work. Licensed under Public Domain via Wikimedia Commons - <http://commons.wikimedia.org/wiki/File:ElectoralCollege2008.svg#mediaviewer/File:ElectoralCollege2008.svg>

Communicate: Spatio-Temporal data



"Minard" by Charles Minard (1781-1870) - see upload log. Licensed under Public Domain via Wikimedia Commons - <http://commons.wikimedia.org/wiki/File:Minard.png#mediaviewer/File:Minard.png>

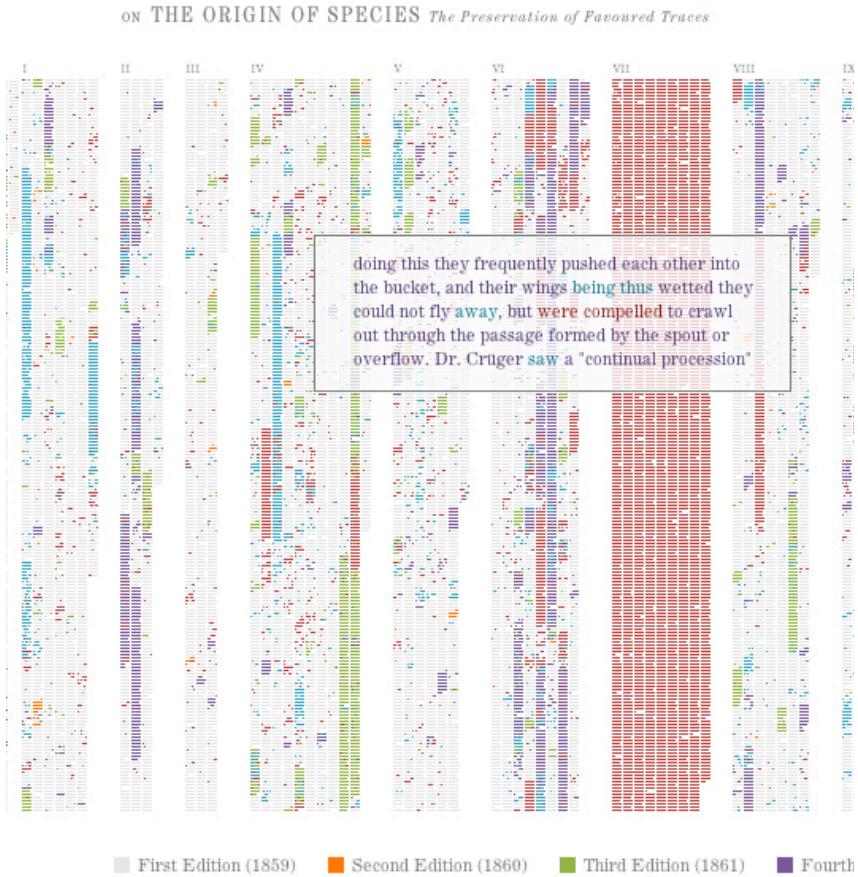


Visual Analytics of Movement.

G. Andrienko, N. Andrienko, P. Bak, D. Keim, S. Wrobel

Springer 2013

Communicate: Text



<http://benfry.com/writing/archives/529>

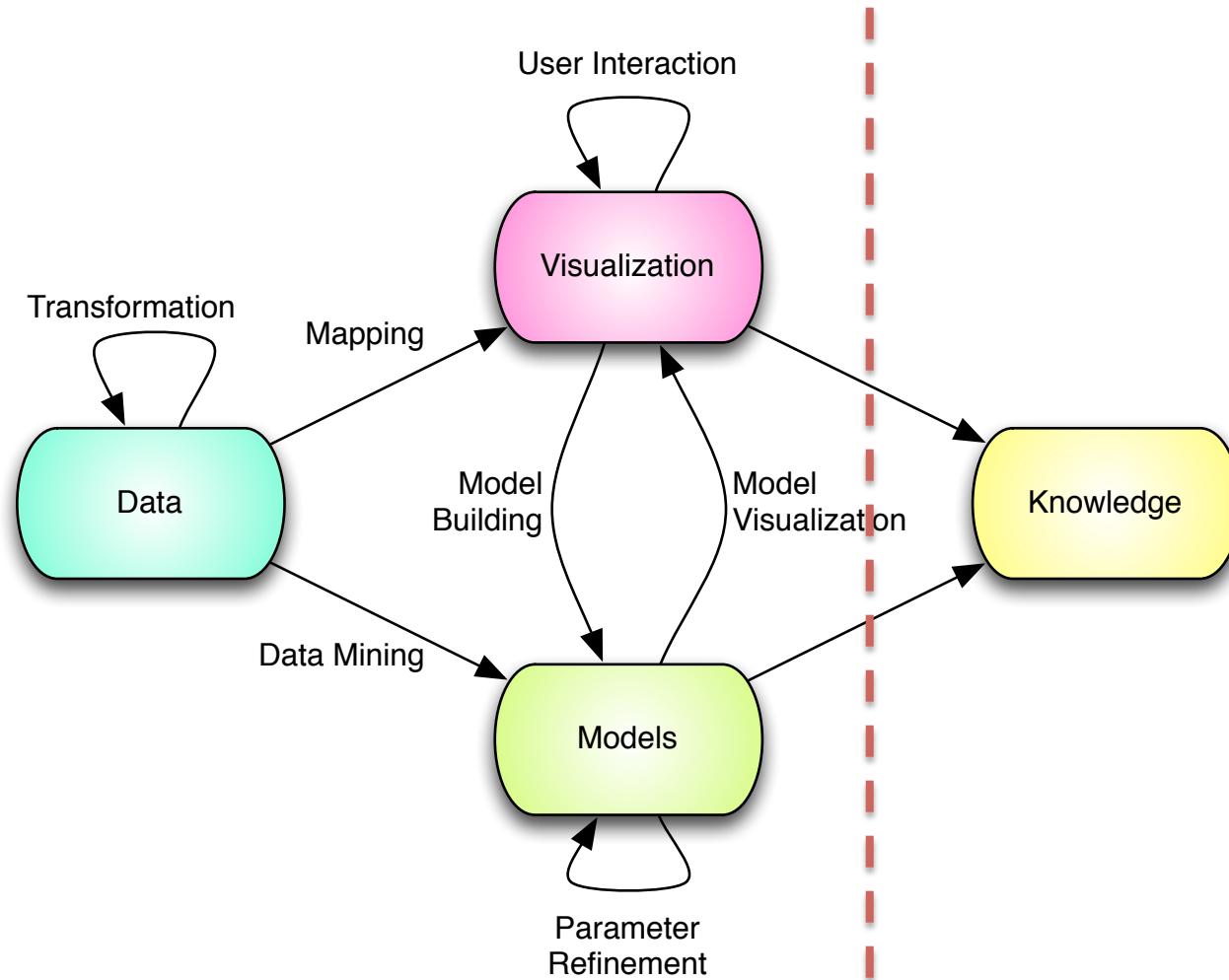
Visualization and Visual Analytics

- Make data and information processing transparent
- Combine strengths of humans and computers

**Computers are
incredibly fast,
accurate,
and stupid;
humans are
incredibly slow,
inaccurate
and brilliant;
together
they are powerful
beyond
imagination.**

Albert Einstein

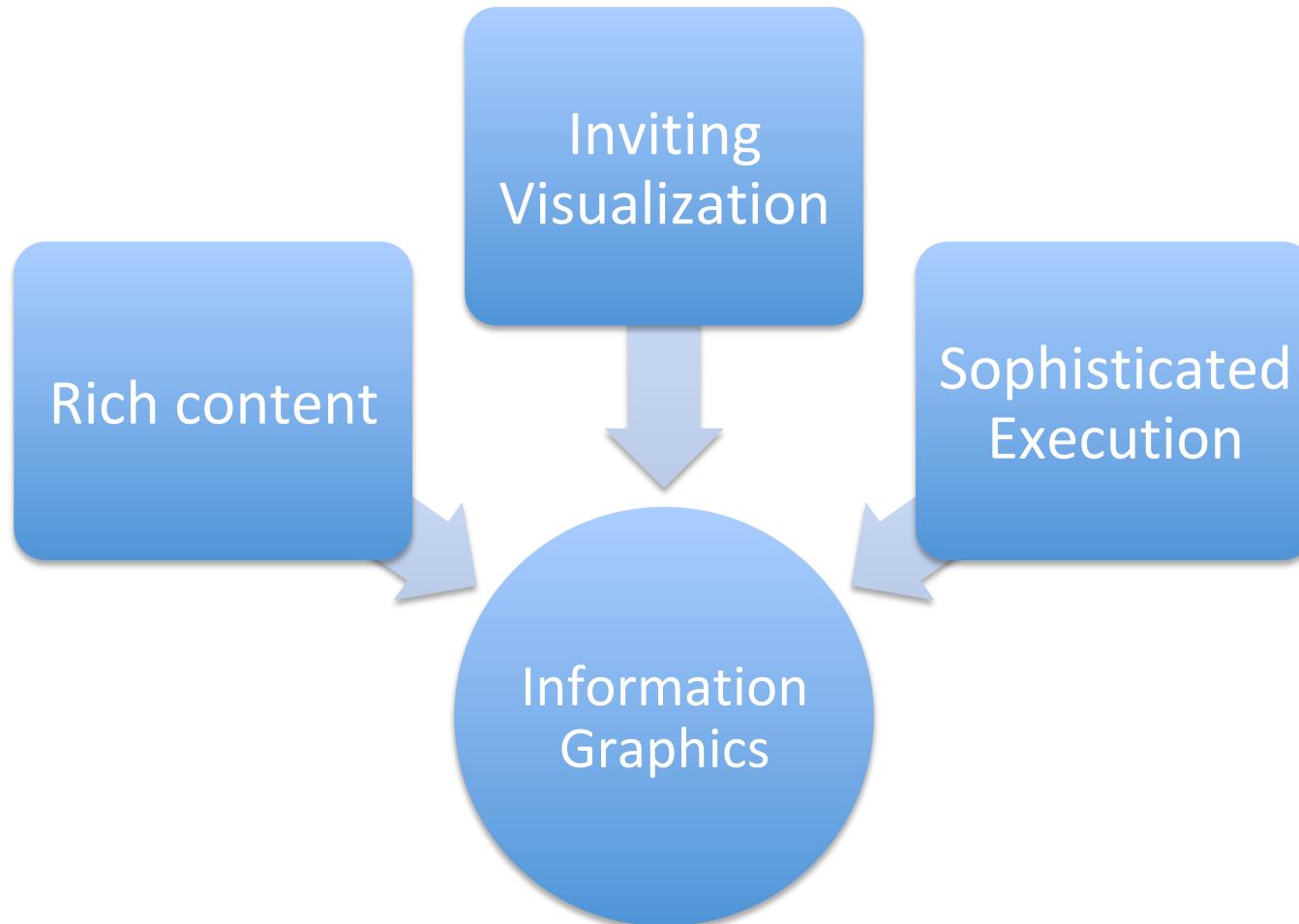
Visual Analytical Process



Adapted from:
Mastering the Information Age
Keim, Kohlhammer, Ellis, Mansmann

! Exploration : Explanation²⁴

Elements of Good Visualization



Importance of valid data



Other Resources

Observe how others resolved design problems

datavisualization.ch

The screenshot shows the homepage of datavisualization.ch. It features a top navigation bar with links for SHOWCASES, EVENTS, TOOLS, OPINIONS, DATASETS, and NOTES. Below this is a "LATEST" section with a colorful bar chart. To the right is a "SPOTLIGHT" section with a link to a curated list of recommended tools. Further down are sections for "SUBSCRIBE" (RSS feed, Twitter, Facebook, YouTube, Tumblr), "POPULAR" (with a red circle icon), and "SPONSOR". The main content area has sections for "RECENT" (Visualizing Substratum), "INTERACTIVELY EXPLORE THE YOLO FLIP", and "ELSEWHERE" (links to Segull Skylights and Data Points). A sidebar on the left contains a "COMMENTERS" section with a "Hitherehi" profile.

informationisbeautiful.net

The screenshot shows the homepage of informationisbeautiful.net. It features a top navigation bar with links for Home, About, Blog, Our Data, Events, Contact, Books, Jobs, and Store. Below this is a "latest most popular" section with a grid of infographics. The grid includes a large pink circle, a world map divided by a diagonal line, a tree with arrows pointing outwards, a sunburst chart, a Venn diagram, a pyramid, a target, a bar chart, a scatter plot, a 3D bar chart, and a grid of small human icons. A search bar is located at the bottom left of the grid.

infosthetics.com

The screenshot shows the homepage of infosthetics.com. It features a top navigation bar with links for SUGGEST, ARCHIVES, ABOUT, and other social media links. Below this is a "A LONG TIME AGO..." section with a message about no images found. The main content area has sections for "CONNECT" (Facebook, Twitter, LinkedIn) and "THE CLASSICS" (links to Brad Falchuk's TED talk, Home Ruling, Interactive Things, Jonathan Harris' We Feel Fine, and others). It also features a "SHOP BOOKS" section with links to books like "Data Flow 2: Visualizing Information in Graphic Design" and "Klanten et al.". The main grid of infographics includes a complex network diagram titled "Visualizing Publicly Available US Government Data On The Web", a bar chart titled "The Disappearing Planet: Comparing The Extinction Rate", a 3D bar chart titled "GitHub: The Universe Of Programming Languages Across", a large 3D bar chart titled "Pi Visualized As A Public Urban Art Mural", and a map titled "The Key Players In The Middle East And Their Relation". A sidebar on the right contains links for "ACKNOWLEDGMENTS", "INFOGRAPHIC PRINTS", and "SPONSORED LINKS".