

ARC²S Group

Applied Research on Computational Complex Systems

Temporal and Textual data visualization

Prof. Giancarlo **Ruffo**

“Analisi e Visualizzazione di Reti Complesse” (9 credits)

Laurea Magistrale in **Informatica**

Università degli Studi di Torino

A.A. 2018/19

@giaruffo



TEMPORAL DATA

References

- Isabel Mairrelles, “Design for Information”, Rockport
Chapter 3: Temporal Structures: Timelines and Flows



- Stephen Few, “Now You See It: Simple Visualization Techniques for Quantitative Analysis,” Analytics Press, California, 2009.

Chapter 7: Time-Series Analysis



Introduction

- Time Series
 - Series of values that change overtime
 - Values **usually** sampled at regular time intervals
- Examples
 - Temperature
 - Stock market
 - Heart beat monitor
 - System logs
 - Version history

Time series patterns

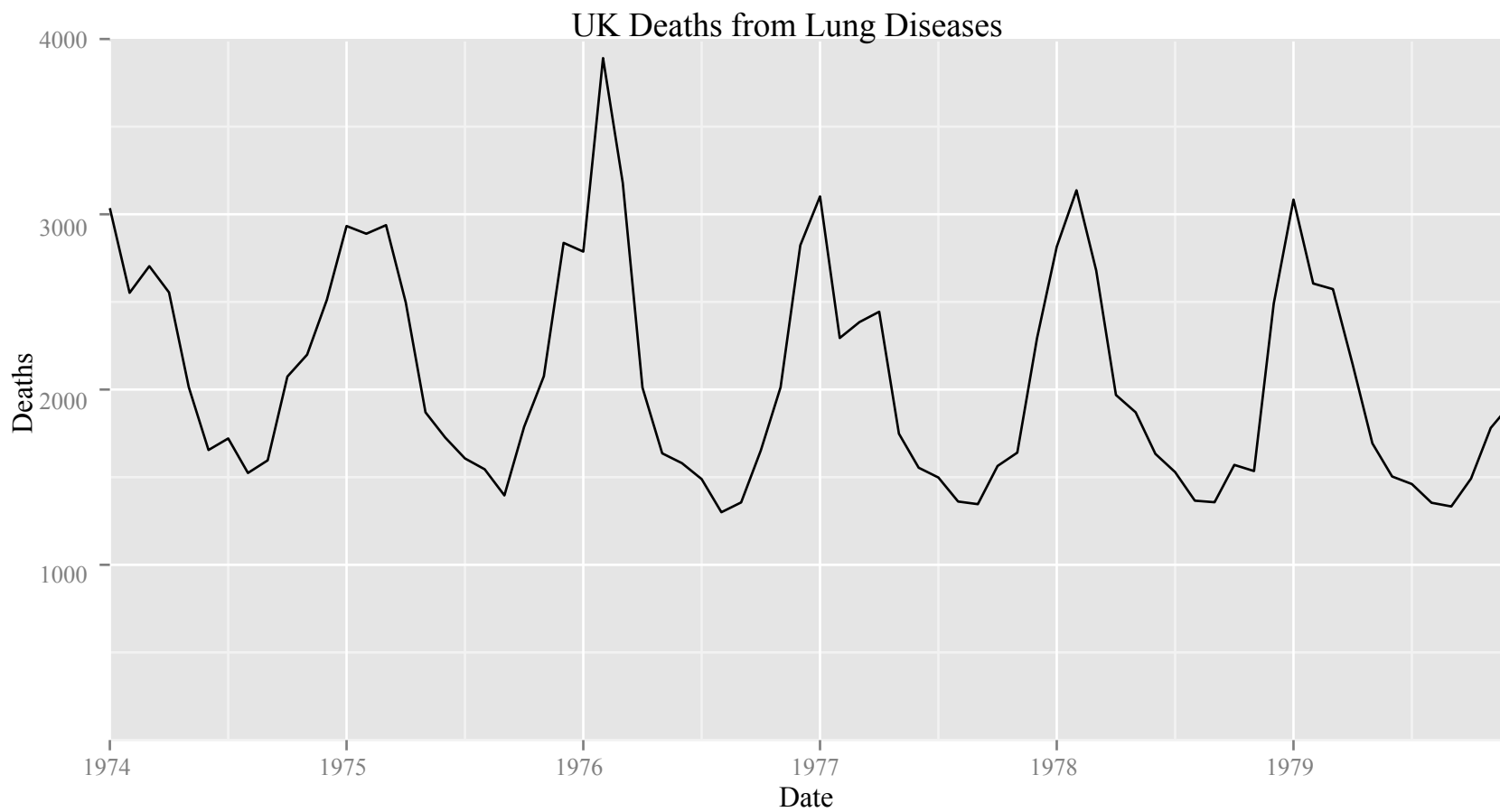
- Trend
- Variability
- Rate of change
- Co-variation
- Cycles
- Exceptions

BASIC TECHNIQUES

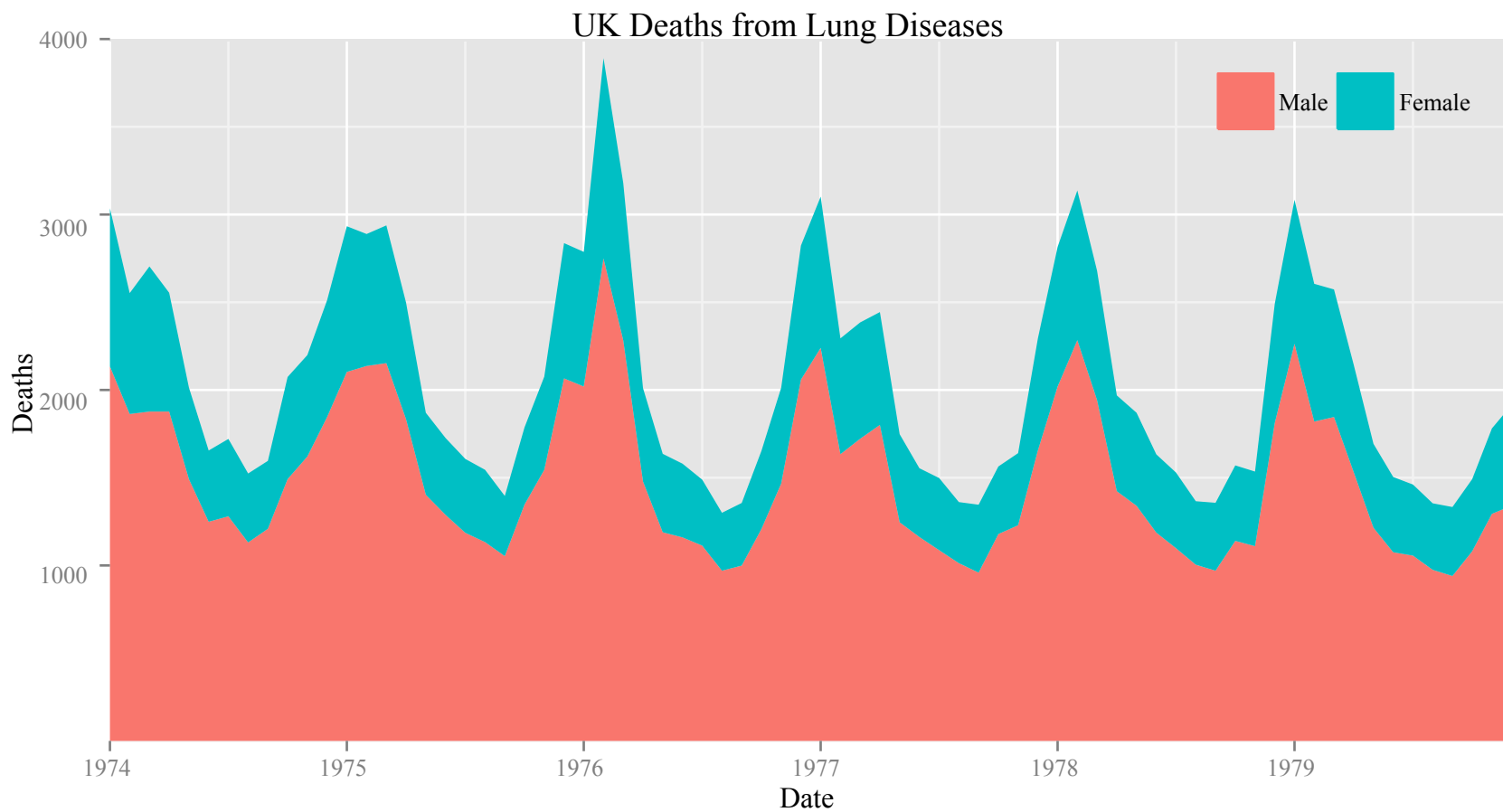
Basic Techniques

- Plot time on x-axis, value on y-axis
 - Line plot
 - Bar plot
 - Area plot
- Plot two variables on x-axis and y-axis, show time via motion by animating changes over time
 - Animated bubble plot
 - Animated geospatial visualizations
 - Animated linked views

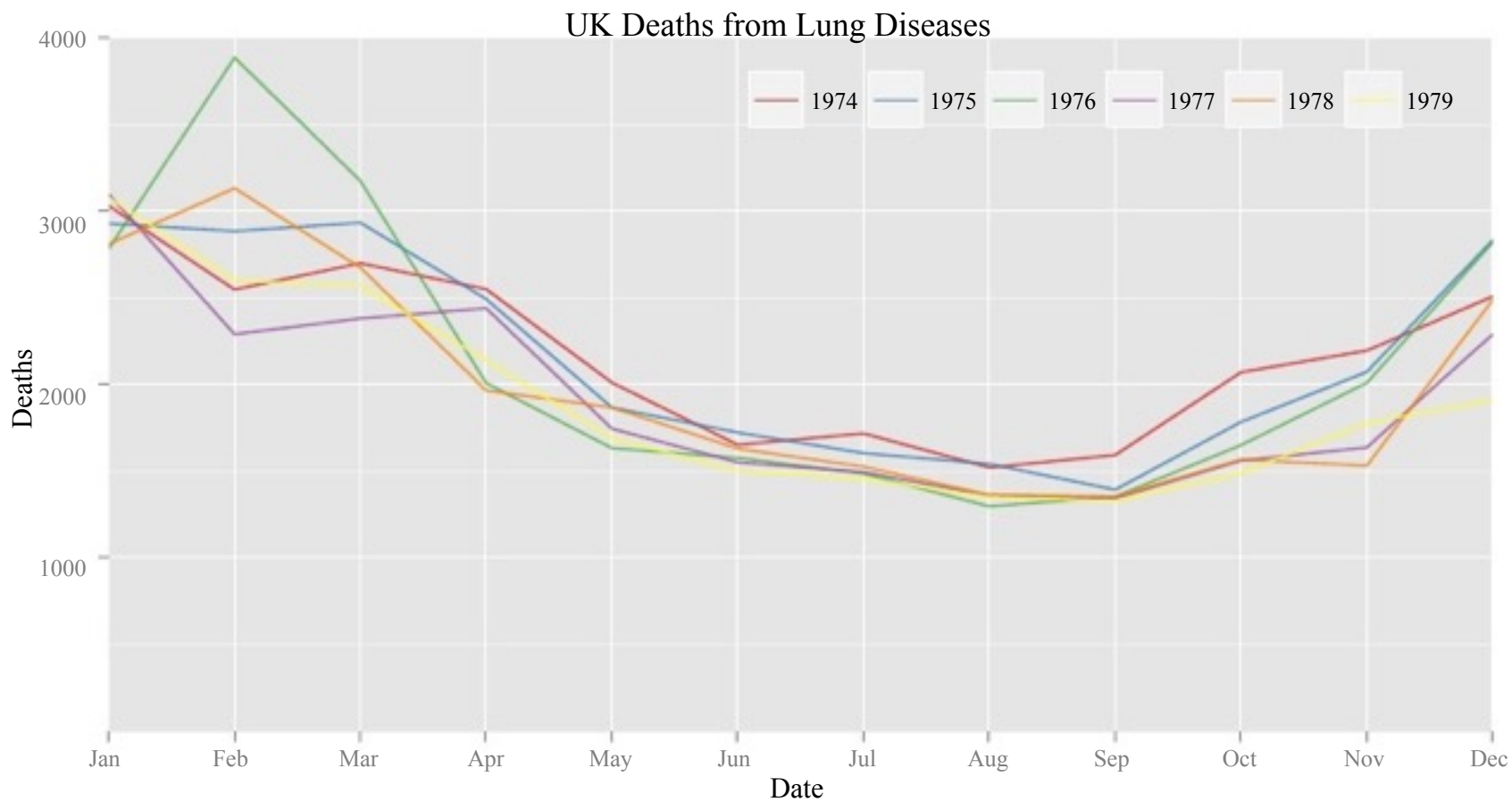
Line Plot



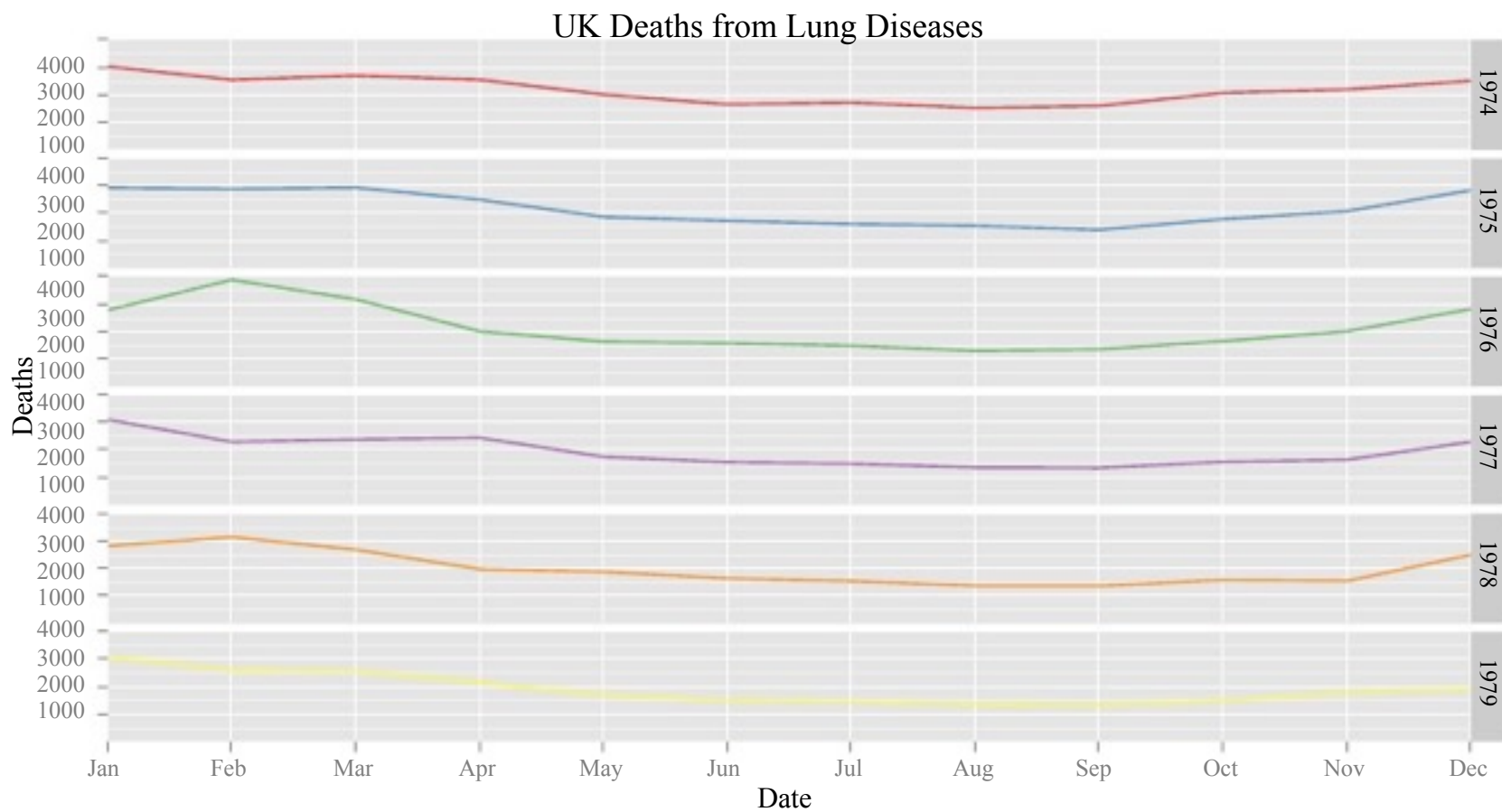
Stacked Area Plot



Multi-Line Plot

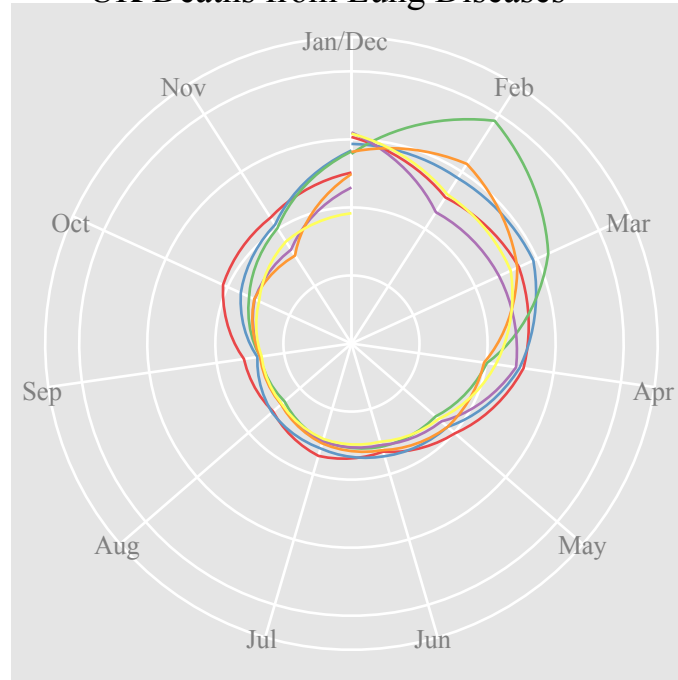


Small Multiples Plot

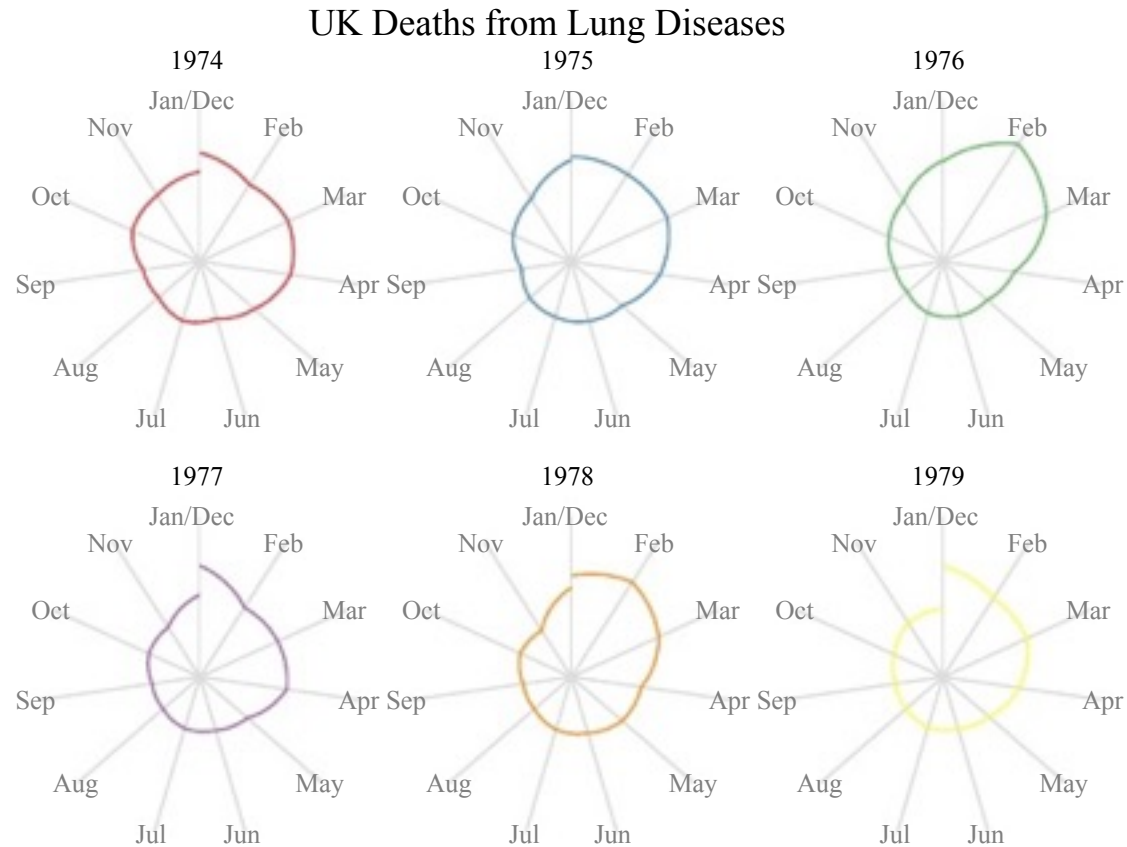


Star Plot*

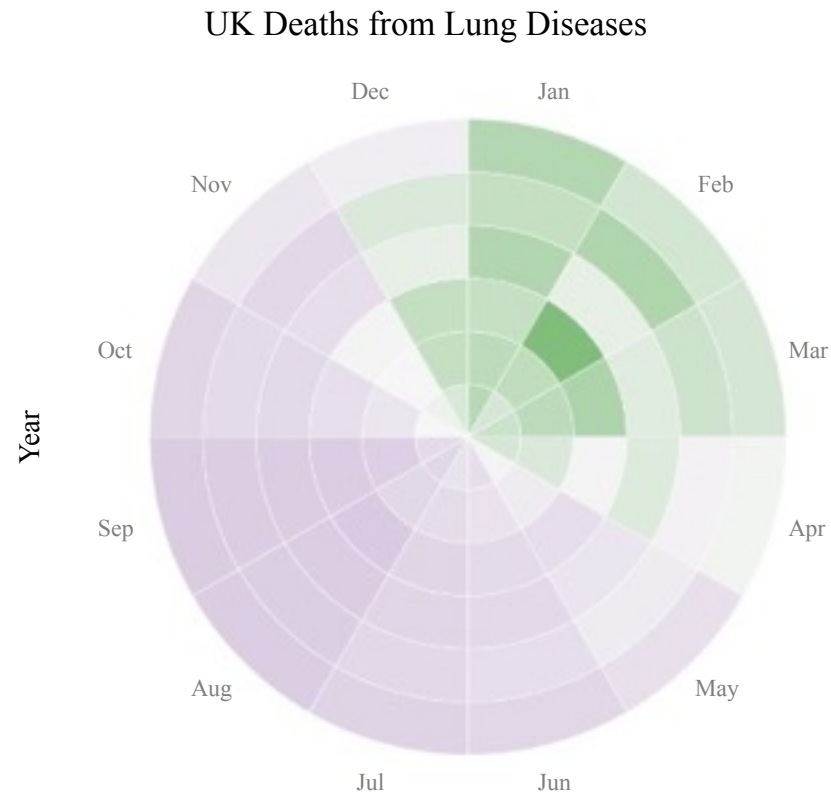
UK Deaths from Lung Diseases



Small Multiple Star Plots*



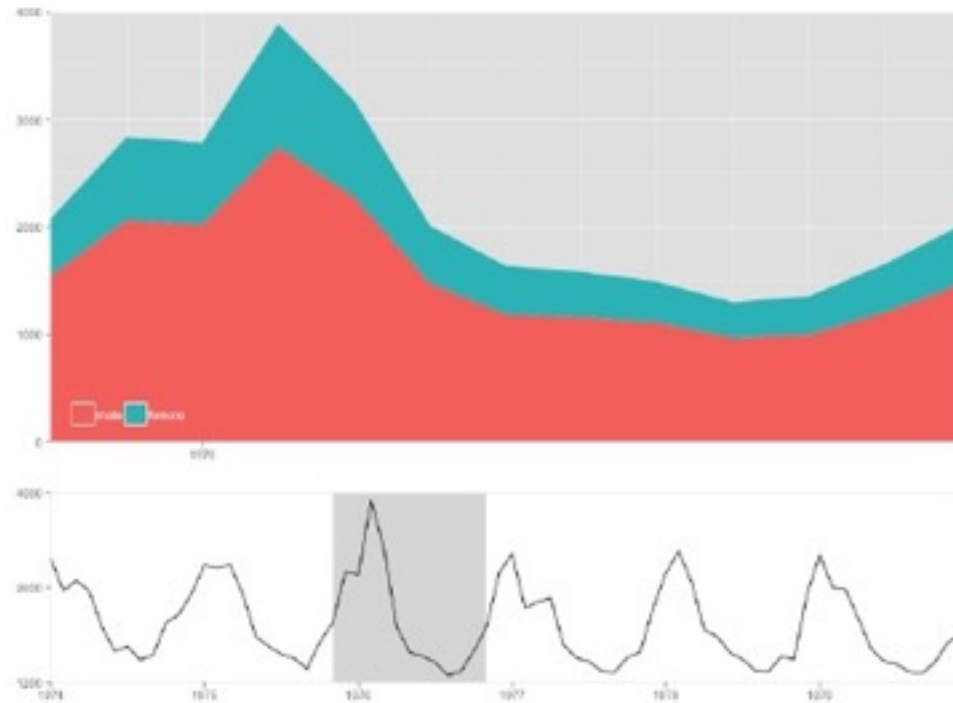
Circle View



<http://dl.acm.org/citation.cfm?doid=989863.989891>

Animated Linked Views

UK Deaths from Lung Diseases



<https://gist.github.com/sjengle/9d5152f0fd4e86d31448>

SHOWCASE

Non-Animated Visualizations

Sunshine Intensity

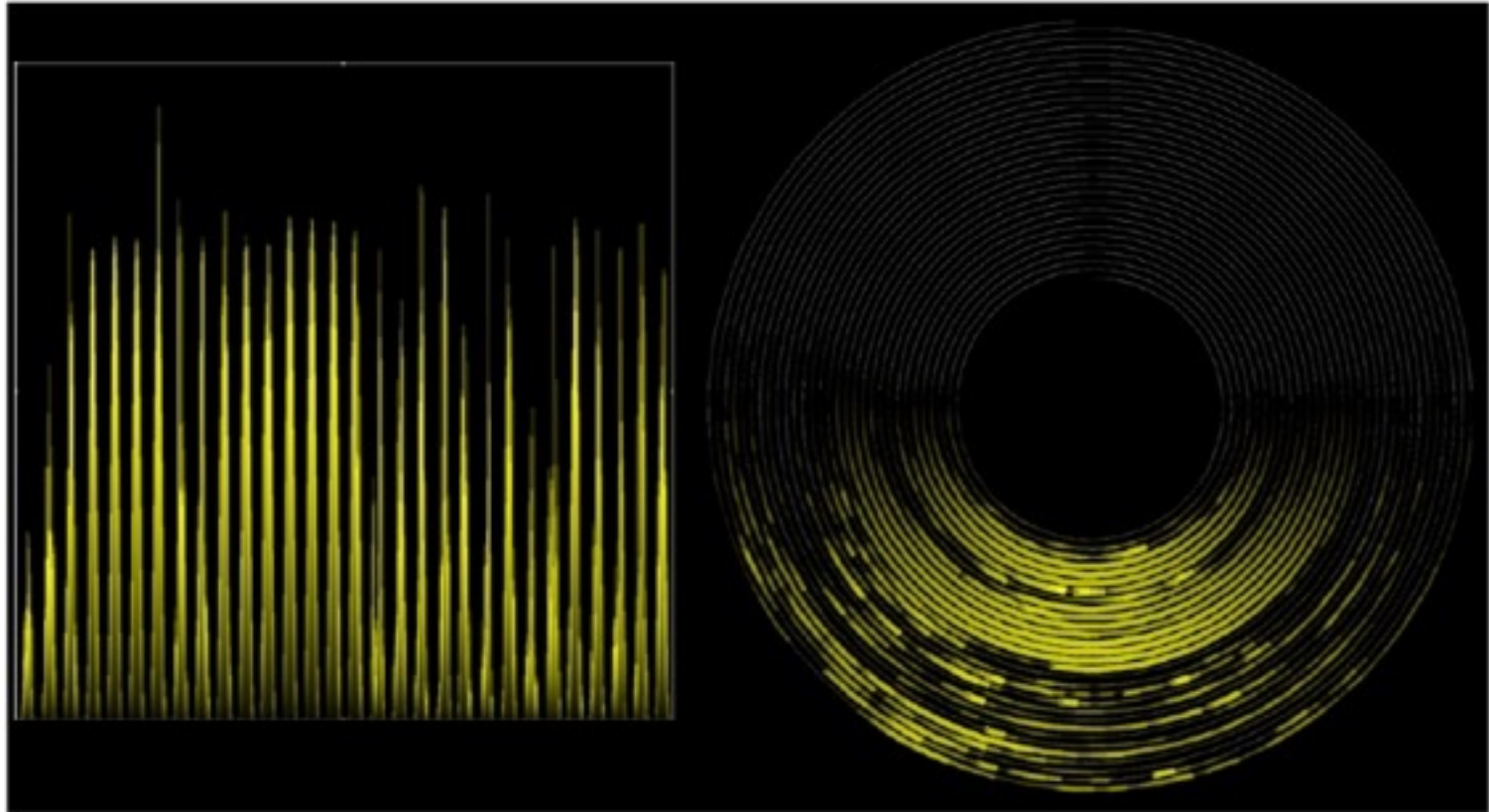
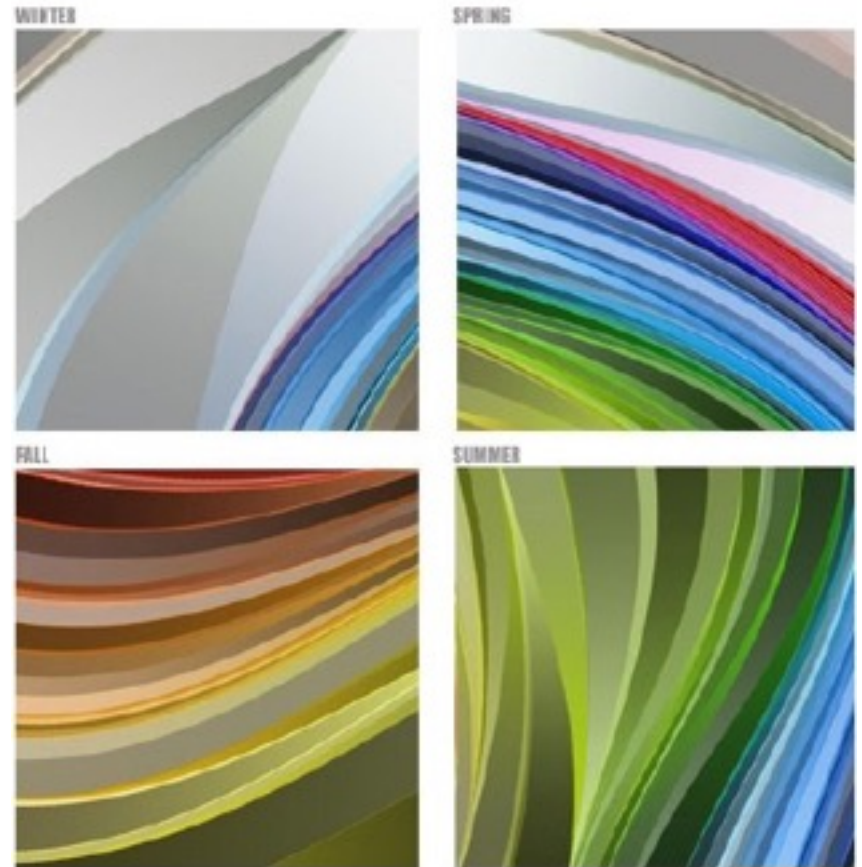


Figure 1: Two visualizations of sunshine intensity using about the same screen real estate and the same color coding scheme. In the spiral visualization it is much easier to compare days, to spot cloudy time periods, or to see events like sunrise and sunset.

<http://dx.doi.org/10.1109/INFVIS.2001.963273>

Flickr Flow



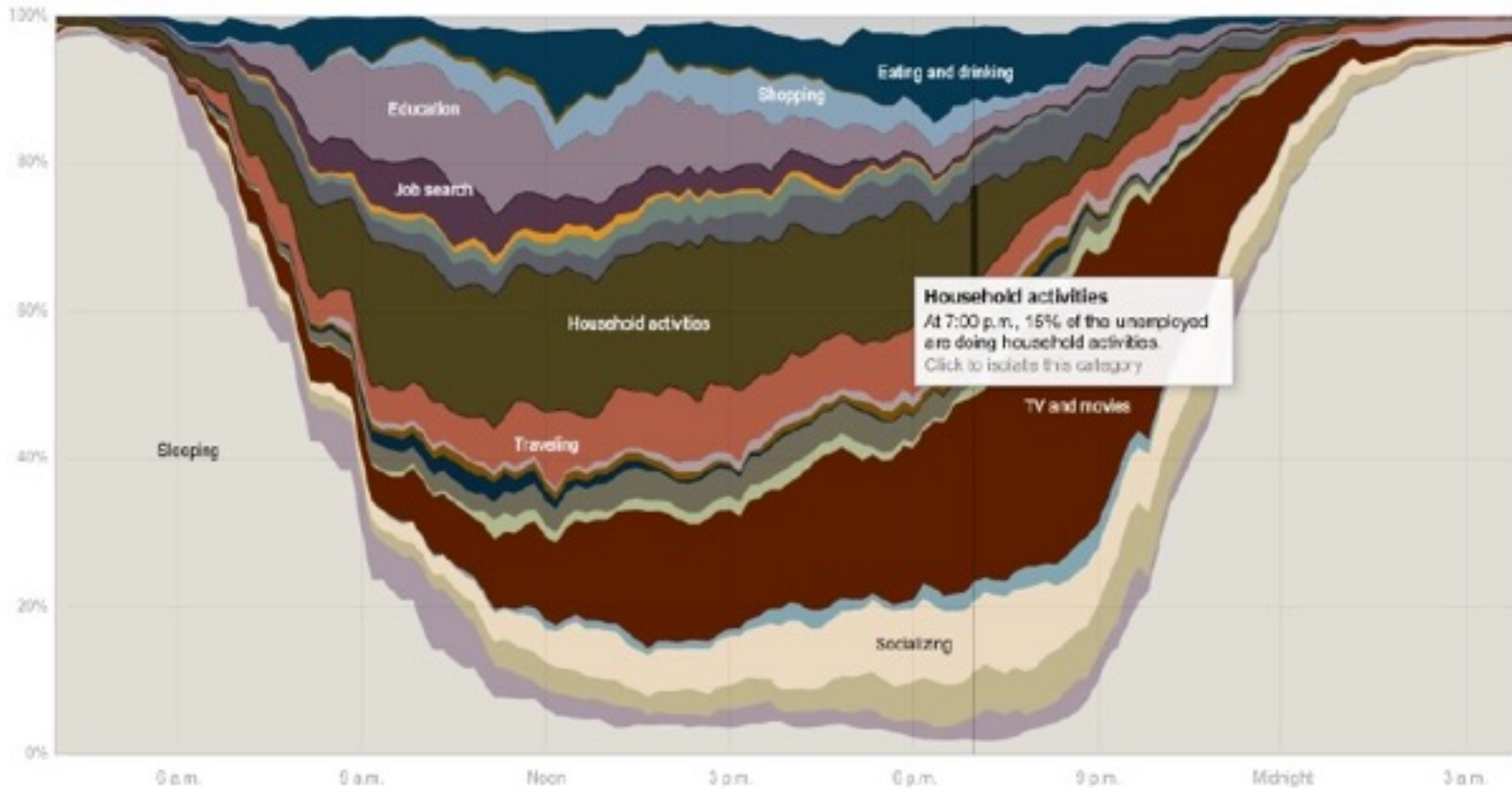
<http://hint.fm/projects/flickr/>

NYT: Stacked Area

The unemployed

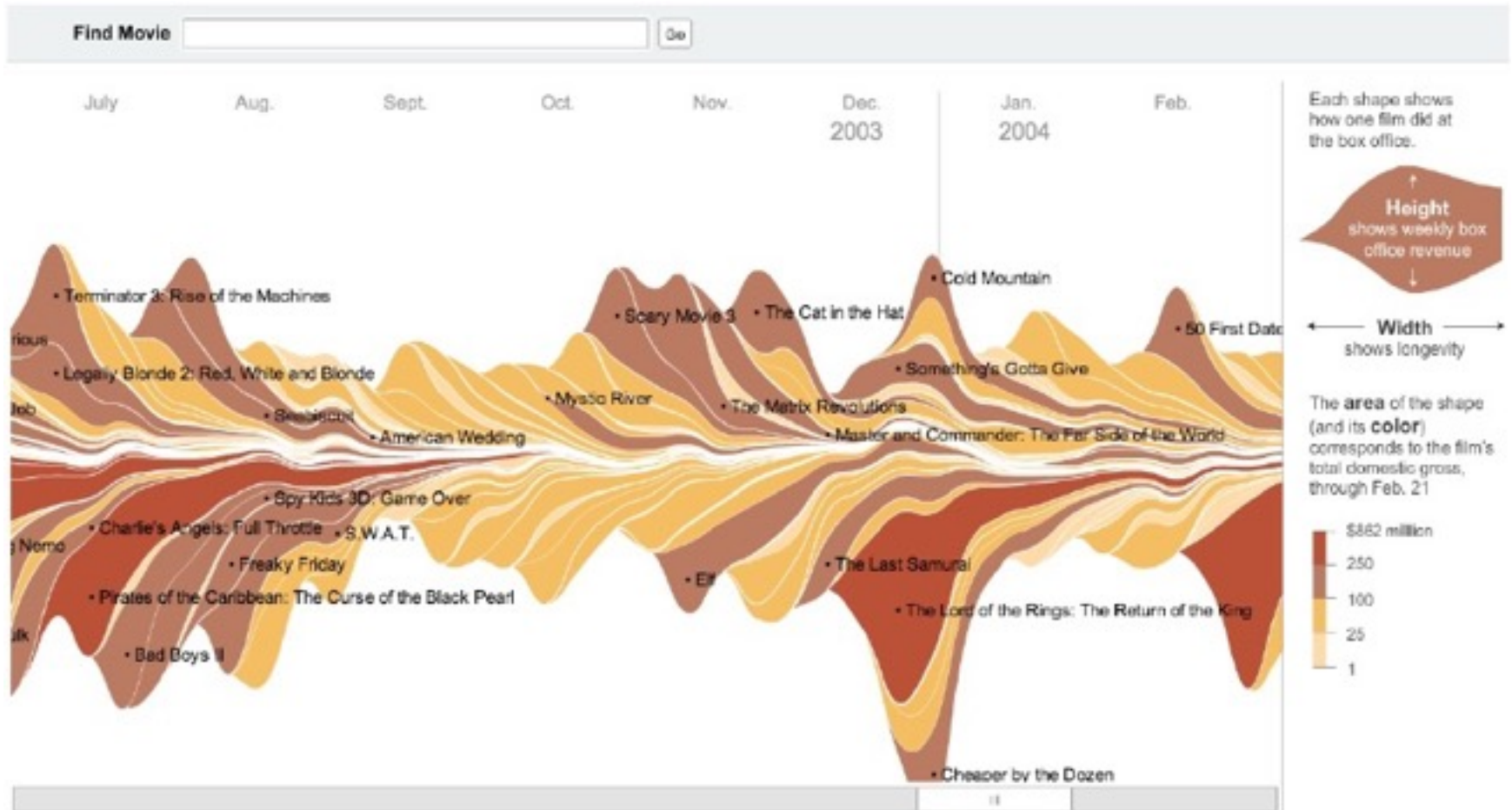
On average, the unemployed spend about a half-hour looking for work. They tidy the house, do laundry and yard work for more than two hours, about an hour more than the employed.

Everyone	Employed	White	Age 15-24	H.S. grads	No children
Men	Unemployed	Black	Age 25-54	Bachelor's	One child
Women	Not in lab.	Hispanic	Age 65+	Advanced	Two+ children



<http://www.nytimes.com/interactive/2009/07/31/business/20080801-metrics-graphic.html>

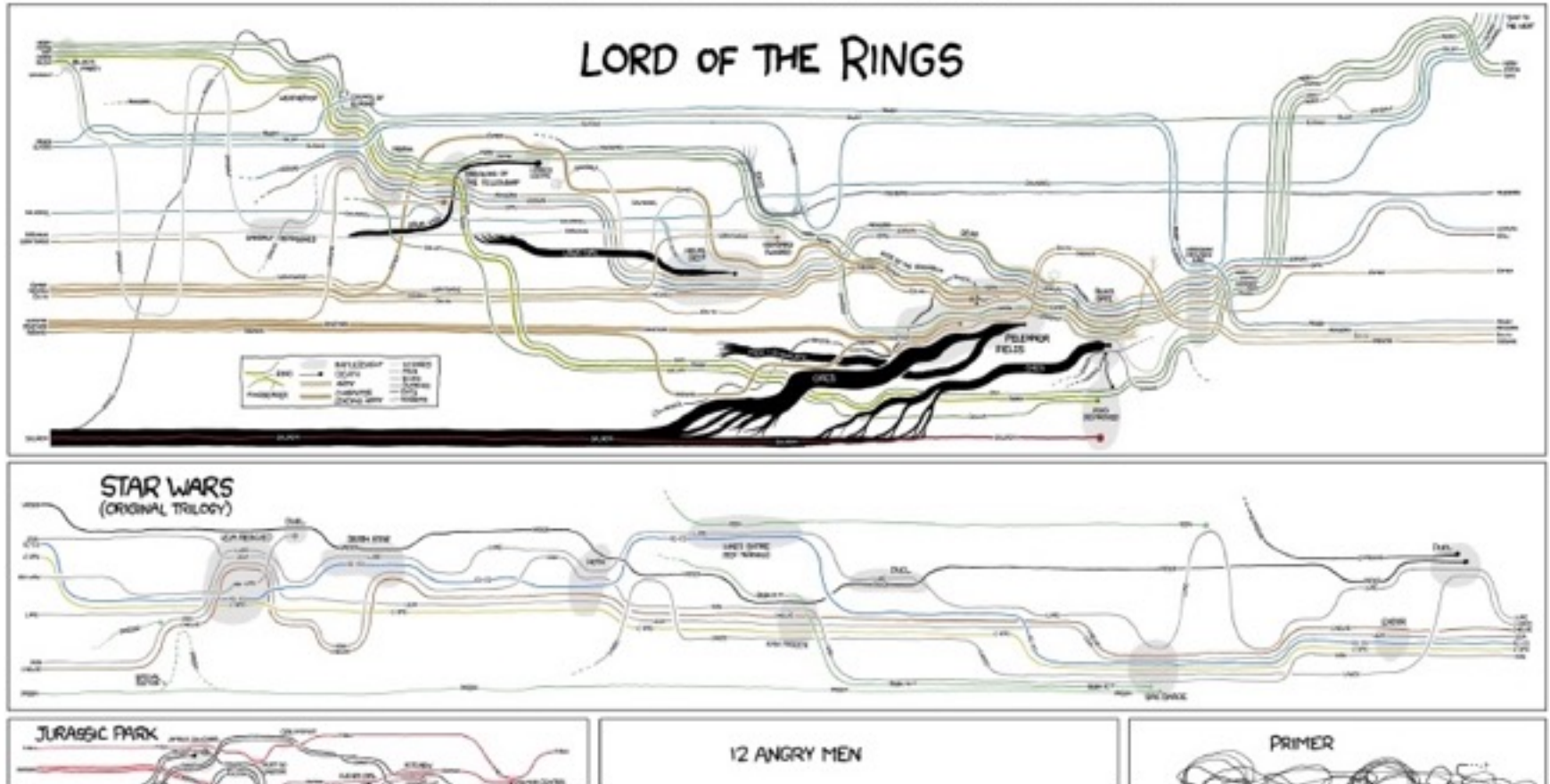
NYT: Streamgraph



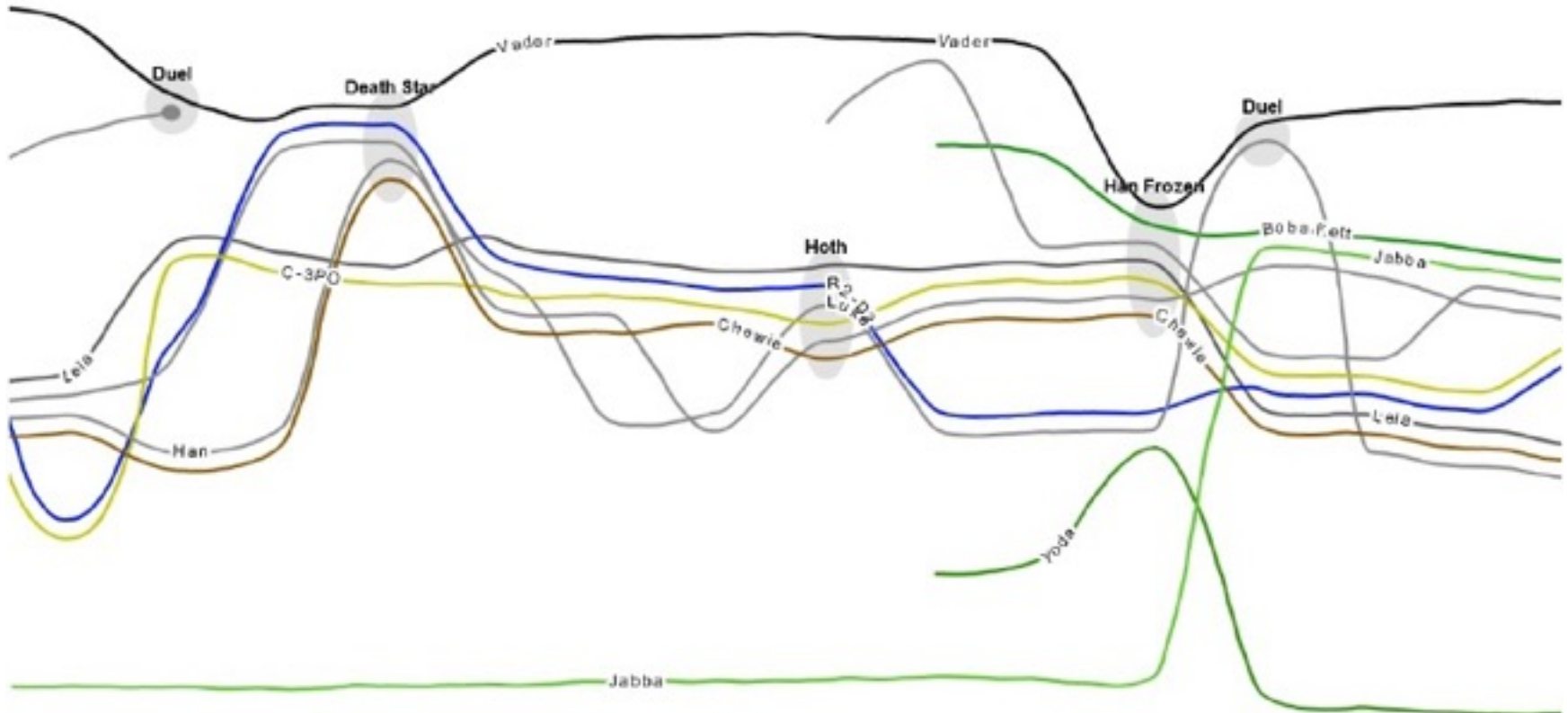
http://www.nytimes.com/interactive/2008/02/23/movies/20080223_REVENUE_GRAPHIC.html

Movie Narrative Charts

THE HORIZONTAL AXIS IS TIME. THE VERTICAL GROUPING OF THE LINES INDICATES WHICH CHARACTERS ARE TOGETHER AT A GIVEN TIME.



Plot Weaver



<http://ogievetsky.com/PlotWeaver/>

SHOW CASE

Animated Visualizations

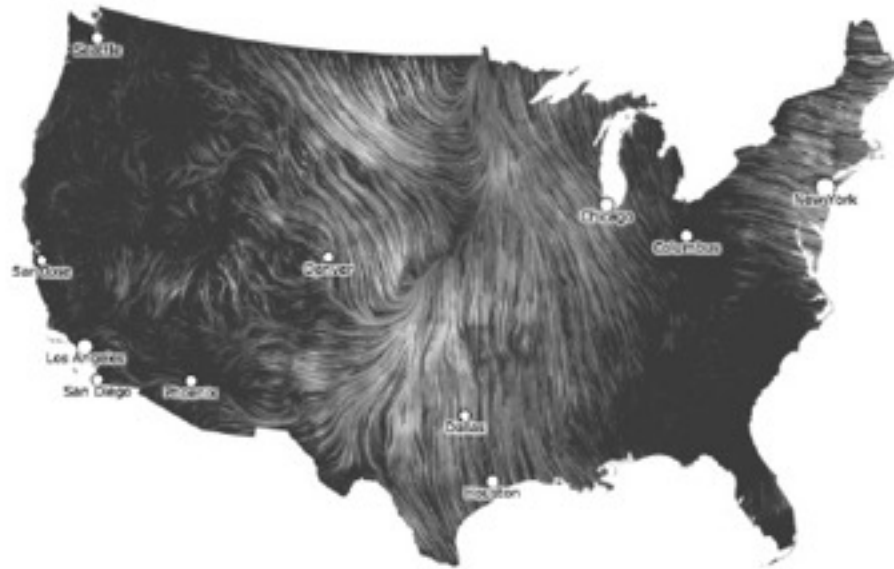
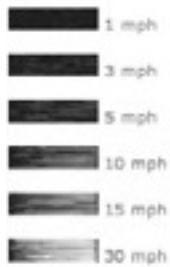
Wind Map

wind map

February 17, 2013

11:59 pm EST
(Time of forecast download)

top speed: 39.4 mph
average: 8.8 mph



Gallery



An invisible, ancient source of energy surrounds us—energy that powered the first explorations of the world, and that may be a key to the future. This map shows you the delicate tracery of wind flowing over the US.

[Donate to Hurricane Sandy relief efforts here.](#)

NEW: wind map prints are available from [Point.B Studio](#).

<http://hint.fm/wind/>

Gapminder

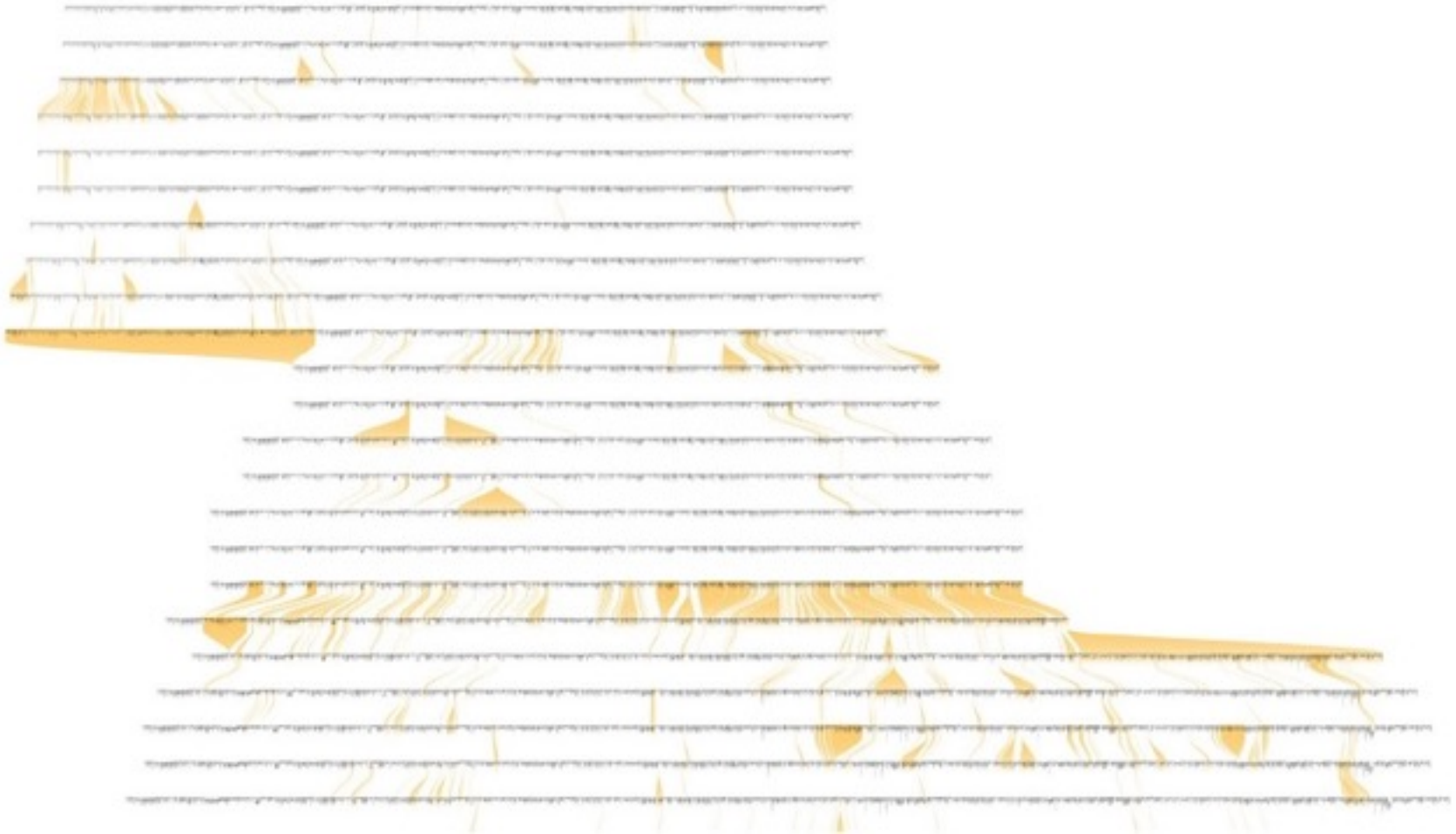


<http://www.gapminder.org/>

SHOWCASE

Revision Visualization

Revisionist



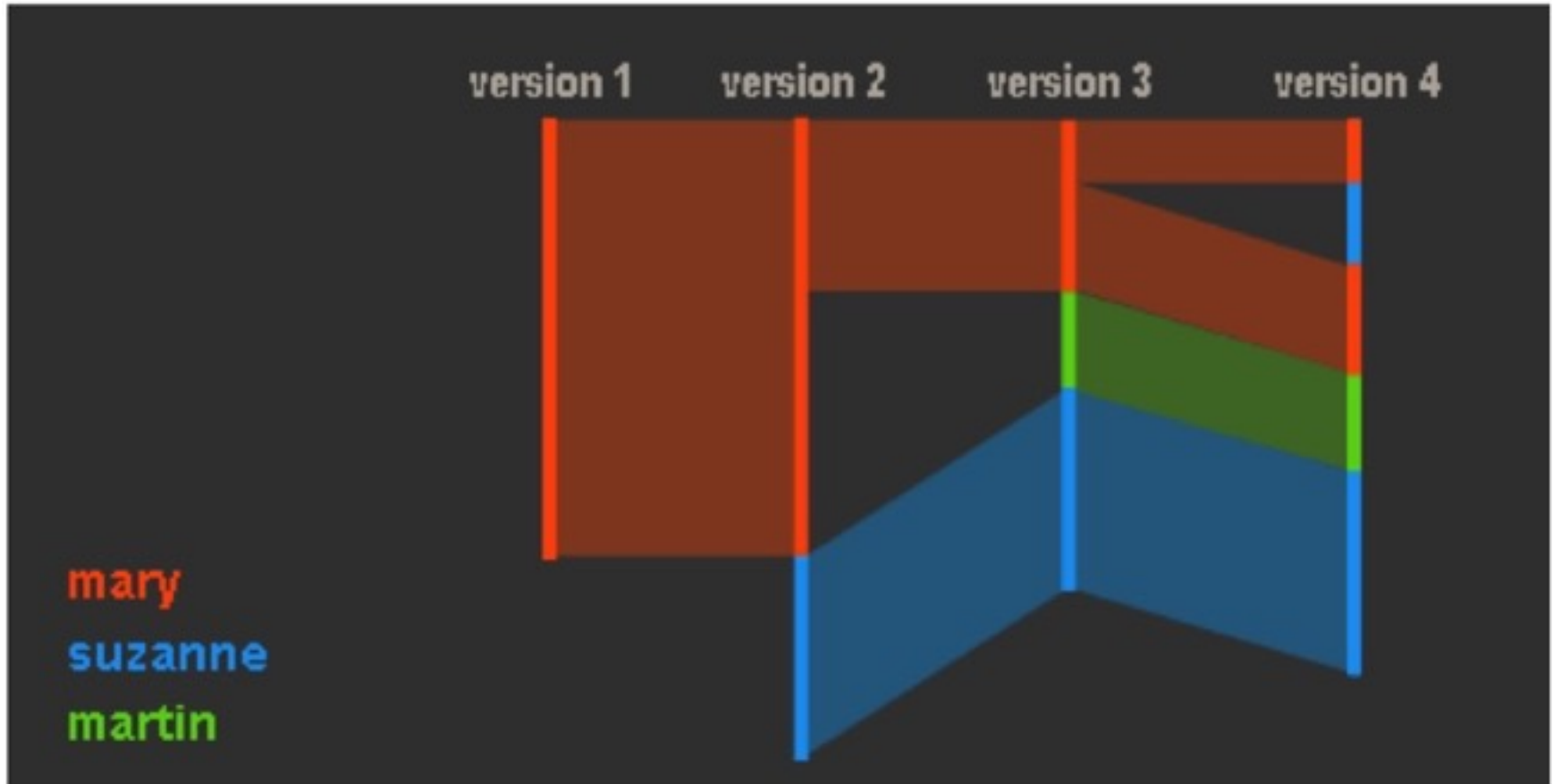
<http://benfry.com/revisionist/>

Revisionist



<http://benfry.com/revisionist/>

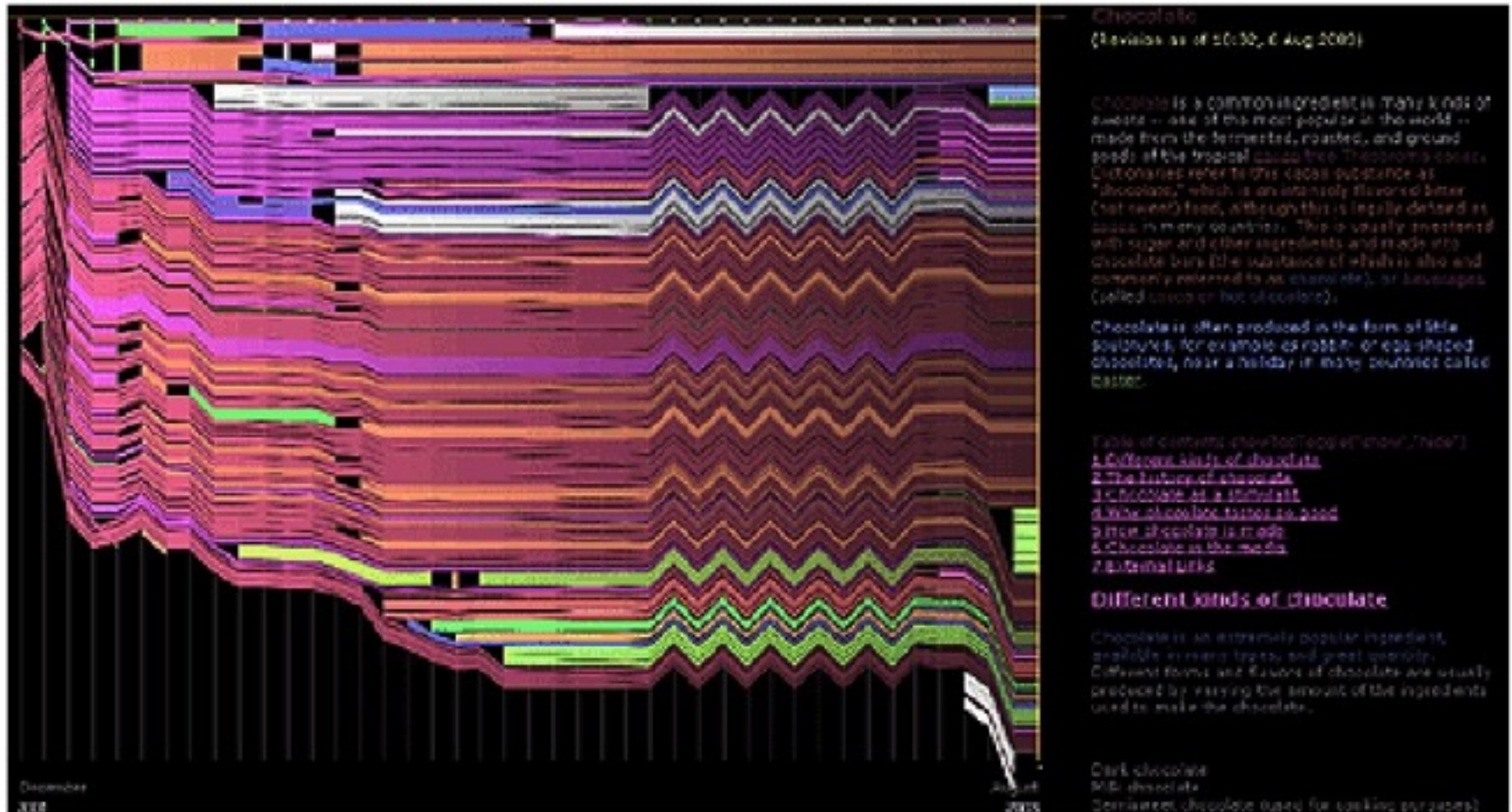
History Flow



<http://hint.fm/projects/historyflow/>

History Flow

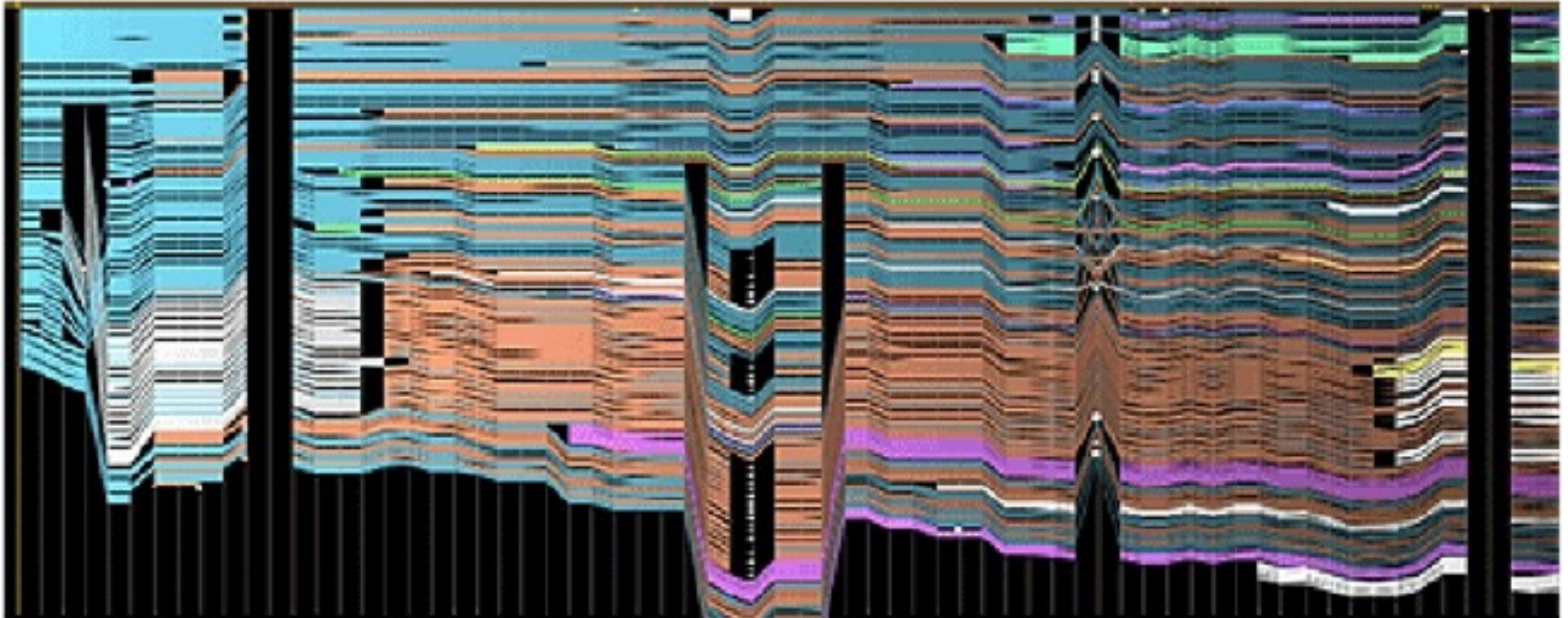
“Chocolate” article on wikipedia



<http://hint.fm/projects/historyflow/>

History Flow

“history” article on wikipedia



<http://hint.fm/projects/historyflow/>

Chromogram

a

Time	Page	Comment
May 21, 9:32 am	Sphere	Add cite
May 22, 10:56 am	Sphere	New Intro
May 22, 1:23 pm	sphere.png	Copyright
May 22, 2:54 pm	Helix	Spell check
May 22, 3:00 pm	Mathematics	Revert
May 24, 11:21 am	Fields Medal	2006 data
May 24, 11:25 am	Talk:Fractal	List of proposed changes
May 24, 11:27 am	Talk:Fractal	List of proposed changes
May 25, 10:13 am	Sphere	Fix tpyo
May 25, 10:23 am	mercator.png	Copyright

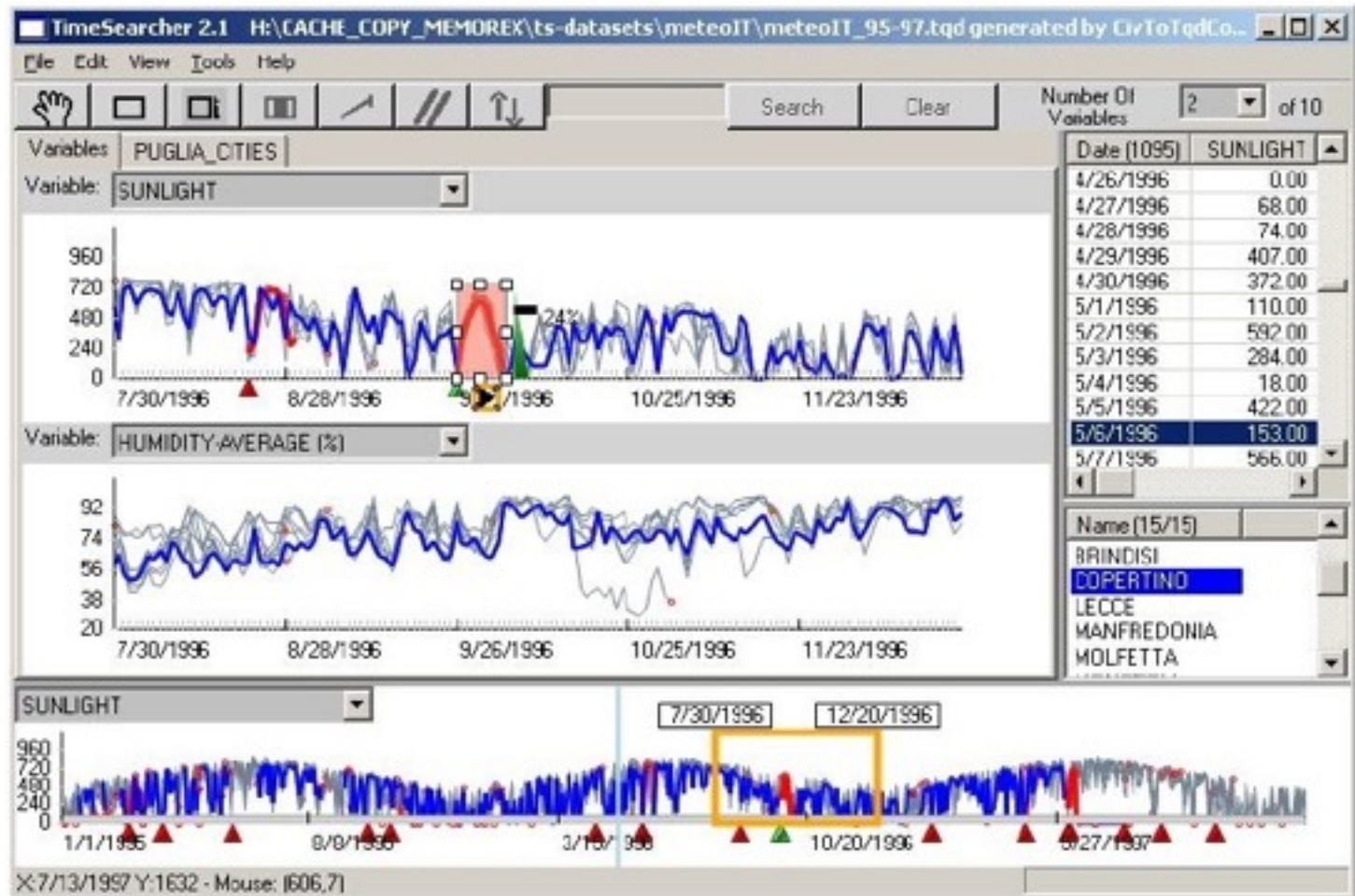


<http://hint.fm/projects/chromogram/>

SHOWCASE

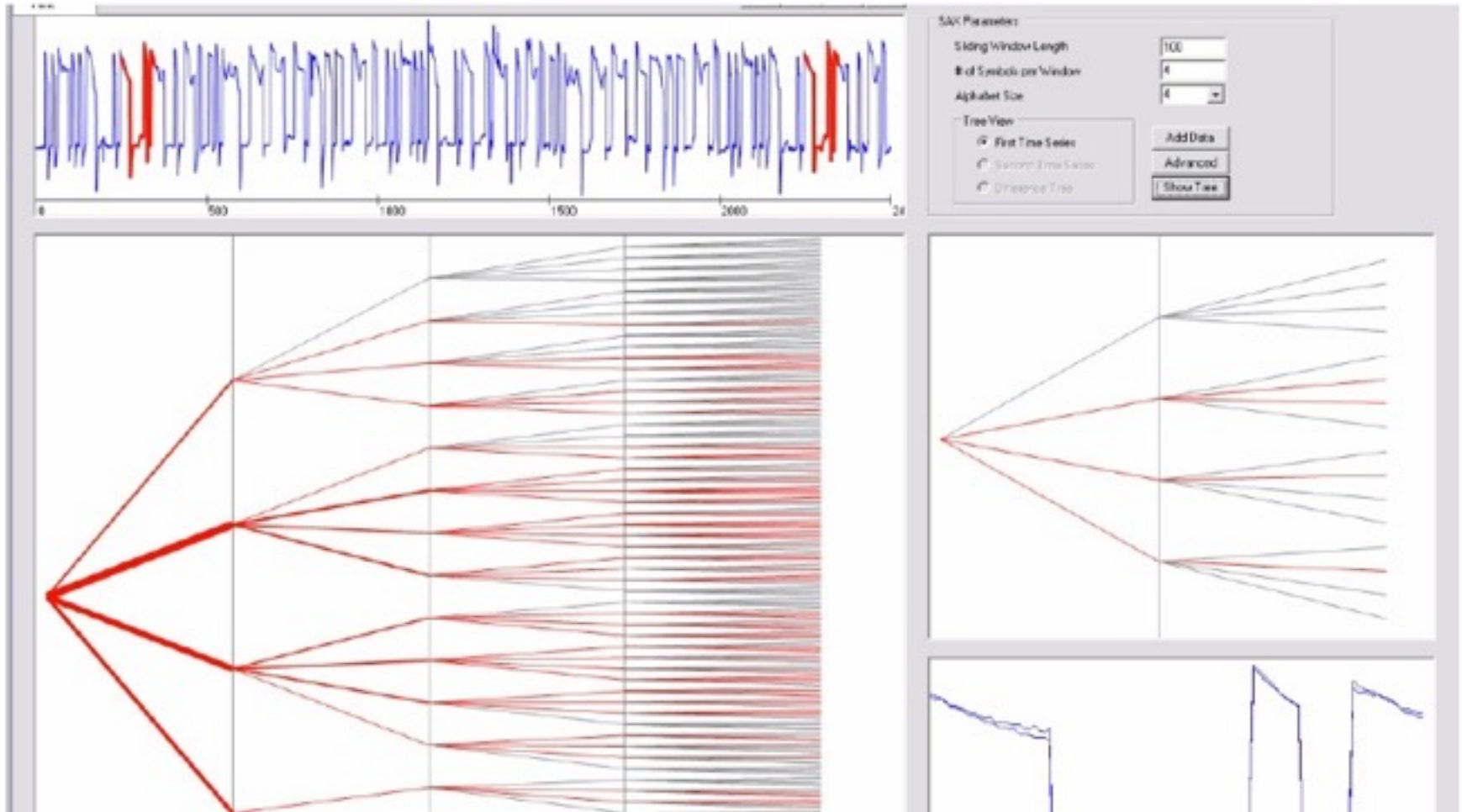
Tools for Time Series

TimeSearcher



<http://www.cs.umd.edu/hcil/timesearcher/>

VizTree



<http://www.cs.gmu.edu/~jessica/viztree.htm>

Similan

Similan - /Volumes/DATADISK/Documents/FlexWorkspace/Similan/bin-debug/data/LL2_Similan_2009Experi... Showing 245 / 246 records

File Edit Settings Help

Target: 00000001
11/7/2008

Sort by: Total Score

Target ID	Target Date	Similarity Score	Similar Patient ID	Similar Patient Date
00000001	11/7/2008	0.99	49988681	11/9/2008
		0.88	45470385	10/4/2008
		0.85	45436079	9/29/2008
		0.84	45551553	10/17/2008

Comparing Target: 00000001 (Fri Nov 7 2008) (above) with 49988681 (Sun Nov 9 2008) (below)

Search Filters Weight

Align by: Admit

Similarity Search

- Select Target (Drag and Drop)
00000001 (Fri Nov 7 2008)
- Select Categories
 - Include in Search
 - Visible
 - Admit (246)
 - Emergency (15)
 - ICU (261)
 - Intermediate (37)
 - Floor (310)
 - Exit (175)
- Choose time difference precision
hour
- Specify Range of Interest (optional)
1) 0Y 0M 0D to 0Y 0M 3D

<http://www.cs.umd.edu/hcil/similan/>

TimeViz Browser

The TimeViz Browser
A Visual Survey of Visualization Techniques for Time-Oriented Data

of Techniques: 110

Search:

Data

Frame of Reference

- Abstract ON
- Spatial ON

Number of Variables

- Univariate ON
- Multivariate ON

Time

Arrangement

- Linear ON
- Cyclic ON

Time Primitives

- Instant ON
- Interval ON

Visualization

Mapping

- Static ON
- Dynamic ON

The image displays a grid of 110 small thumbnail images, each representing a different visualization technique for time-oriented data. The thumbnails are arranged in a 10x11 grid (with the last cell empty). The techniques shown include various types of charts (line, bar, area, pie, radar), maps (2D, 3D, network), diagrams (flow, tree, hierarchical), and other data representations (3D surfaces, network graphs, etc.). Some thumbnails are highlighted with a red border, indicating they are currently selected or featured.

<http://survey.timeviz.net/>

TEXTUAL DATA

References

- Isabel Mairrelles, Design for Information, Rockport
Chapter 6: Textual Structures



Document Mining

- Natural Language Processing
- Digital Humanities
- Quantitative approach vs traditional approach to literary corpus understanding
- Visualization tools to analyze unstructured textual data
- Understanding messages streaming from social media (real time)
- News articles analysis

Nominal data

- Objects, names and concepts: examples of **nominal data**
- Data visualization requires some form of organization that change data type:
 - If we can define an order: from nominal to **ordinal** data
 - Sometimes objects share some characteristics: we can group data, creating **categorical** data
 - Generally speaking: qualitative information becomes **quantitative** with some correlation with other data

“Nominal or categorical variables are difficult to display graphically because they have no inherent ordering. The categorical nature of text, and its very high dimensionality, make it very difficult to display graphically” (Marti Hearts)

Text representation

Lexical: transforms a string of characters into a sequence of atomic entities for further analysis

Syntactic: examines and defines the function of each token.
Decisions on which language model and grammars to use will further define the analytical approach

Semantic: extracts the meaning of the structure derived from the syntactic level toward an analytical interpretation of the full text within a specific context



Codex st. Peter perg 92
XIV century

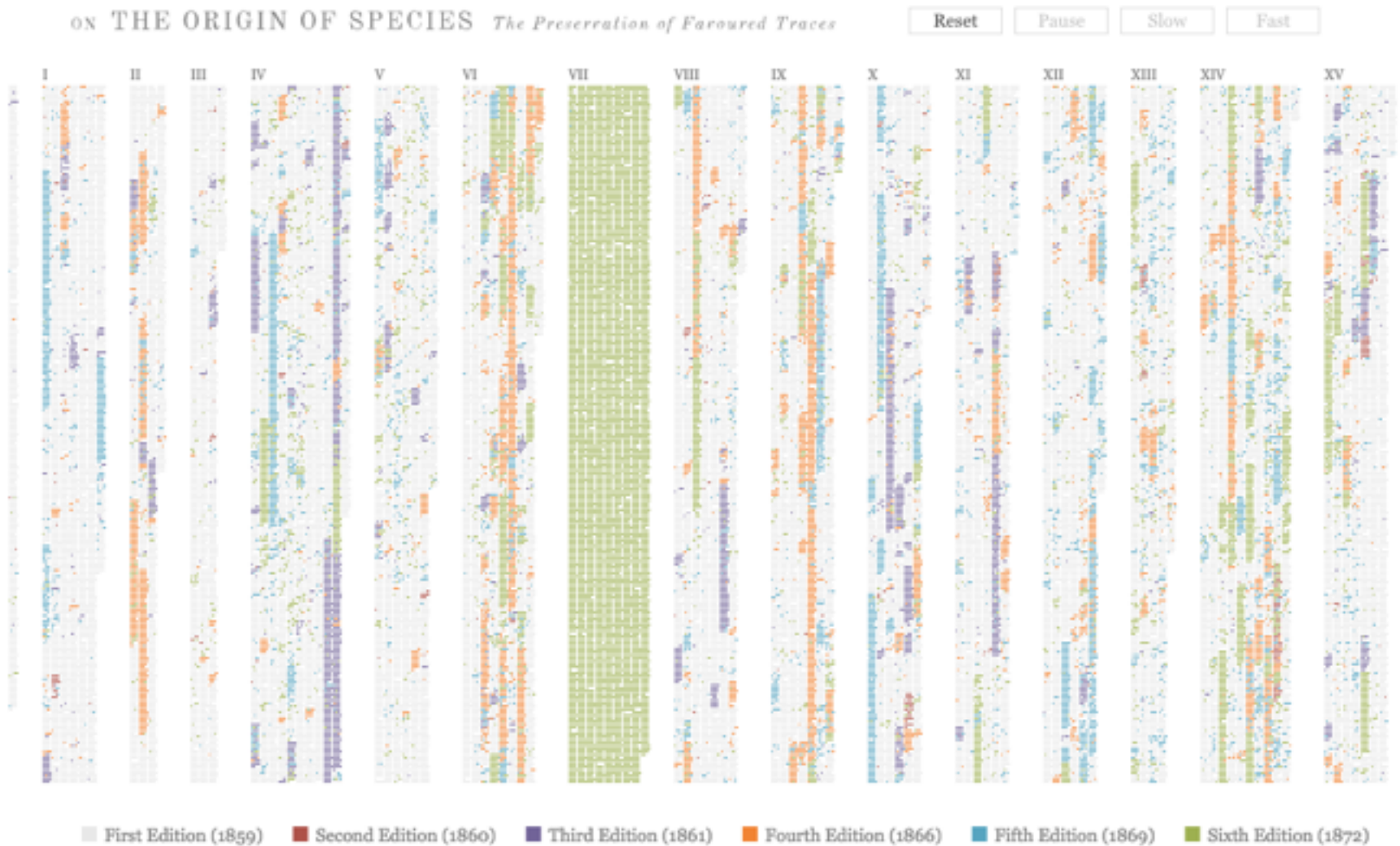
Text Considerations

- Word frequency and distribution
 - Misspelled words?
 - Remove stop words?
 - Stem words?
 - Missing context?
- Phrases
 - Poor grammar?
 - Word order (cat and dog vs dog and cat)?

Three types of textual data visualization

- Viz of **connections** among entities within and across documents
 - text mining, discovery of unknown information, ...
- Viz of document **concordances** and word **frequencies**
 - literature analysis, linguistics, ...
- Viz of **relationships** between words in their usage in language and lexical ontologies
 - literacy analysis and citation analysis

The Origin of the Species (Ben Fry)

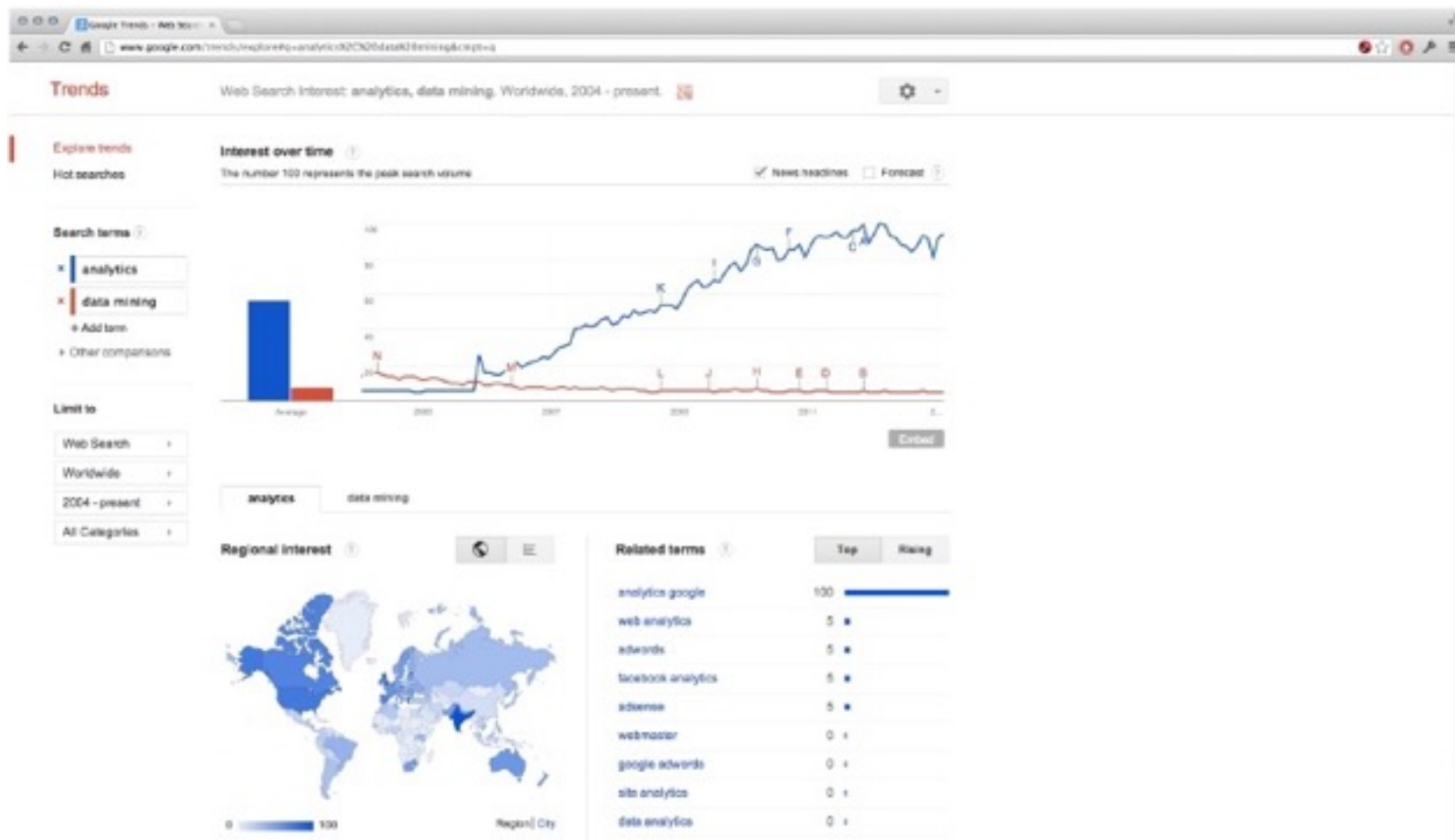


<http://benfry.com/writing/archives/529>

GOOGLE

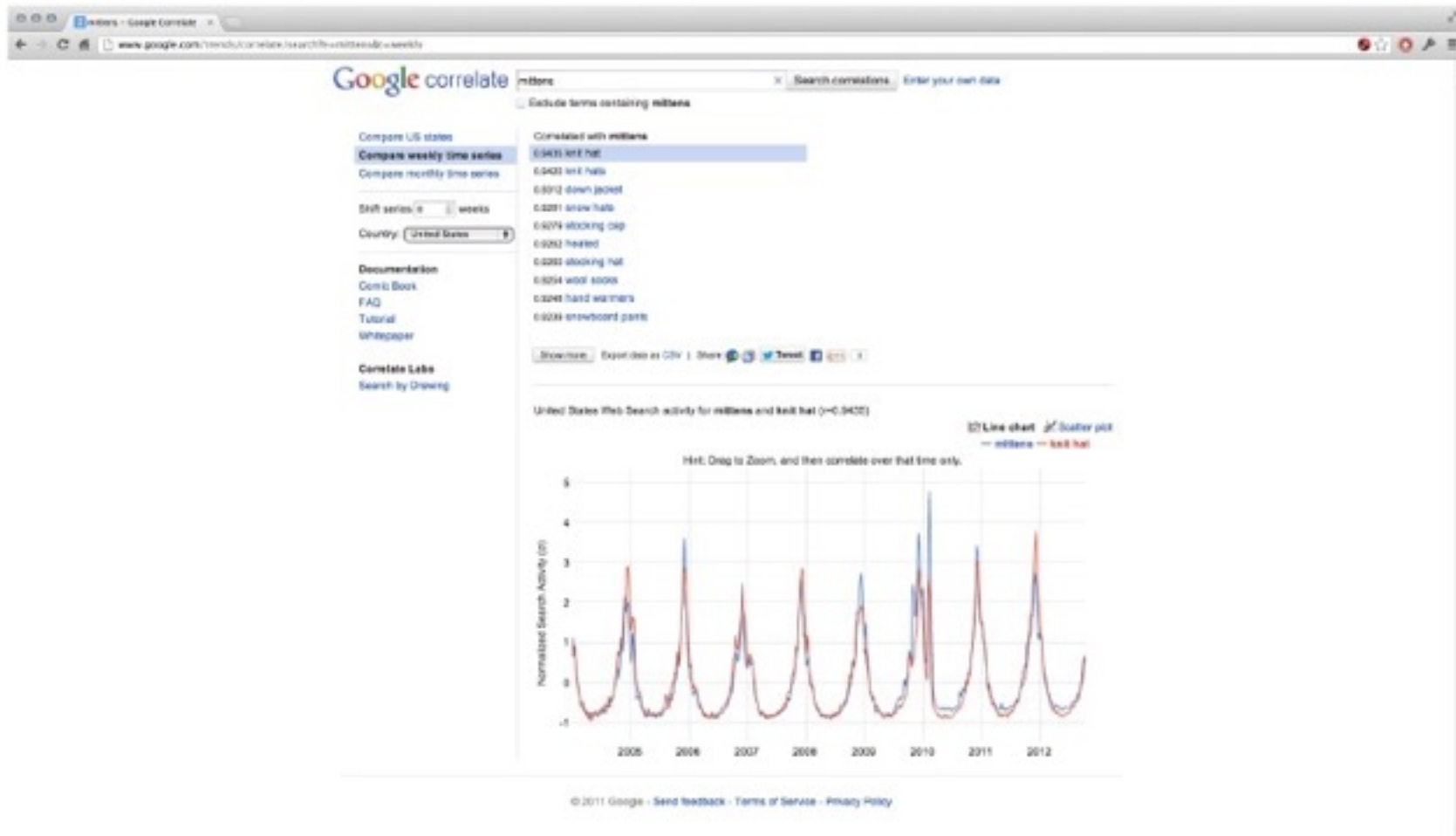
<http://www.google.com>

Google Trends



<http://www.google.com/trends>

Google Correlate



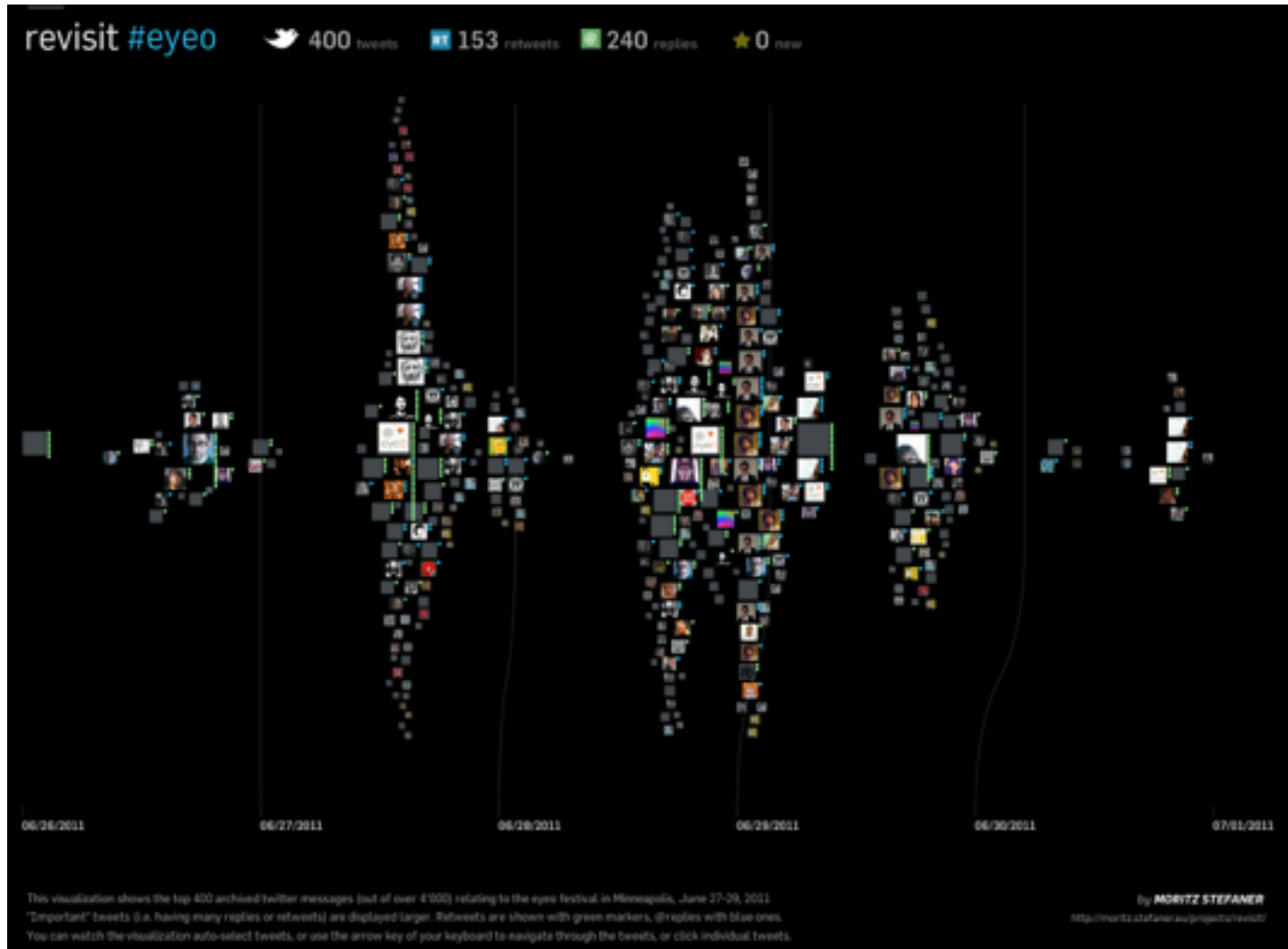
<http://www.google.com/trends/correlate/>

Google Ngram Viewer



<http://books.google.com/ngrams>

Revisit, Moritz Stefaner



<http://truth-and-beauty.net/projects/revisit/>

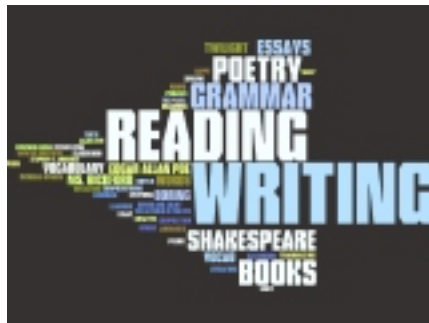
Case studies

Wordle

Phrase Net

Word Tree

Wordle



<http://www.wordle.net>

Pros and Cons of Tag Cloud

Pros

Tag cloud are extremely popular, easy to read, and good to highlight most important topics (two-three greatest “tags”)

Cons

no real ranking; bad for comparisons; no patterns highlighted; no co-occurrences and relationships between words; ...

Bad data density ratio

Bad data ink ratio

**Good (low) lie factor, but very few information/
relationships between data can be displayed**

Tag-Clouds are a no-no in 2019 Data Visualization!

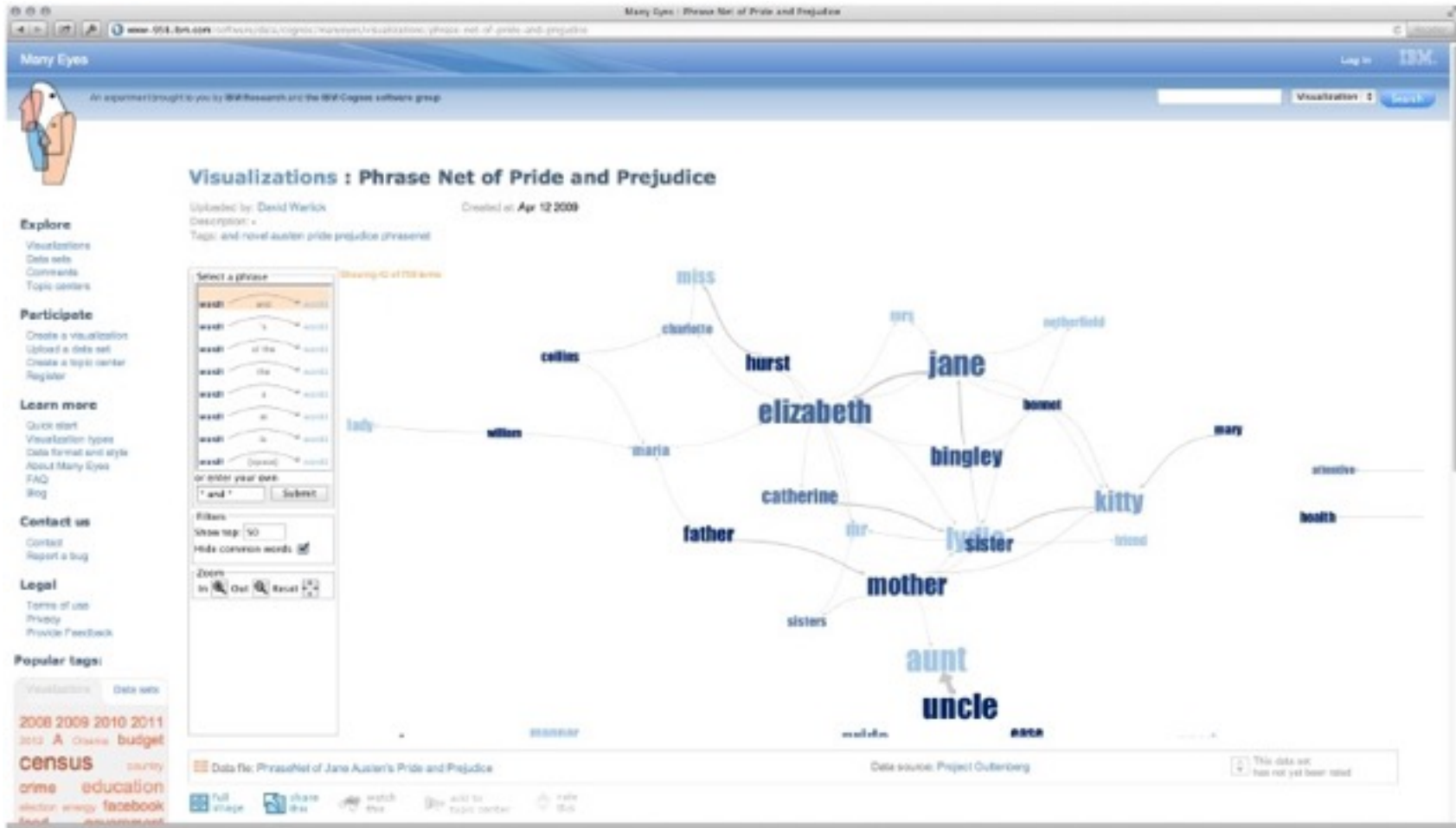
MANY EYES

<http://hint.fm/papers/viegasinfovis07.pdf>

MANY EYES

- Launched in 2007 by Fernanda B. Viégas, Martin Wattenberg, and al. at IBM
- Project closed in June 2015
 - Embedded in IBM Analytics
- Visit Viégas and Wattenberg's site to have some illustration of their contributions: <http://hint.fm>

Phrase Net

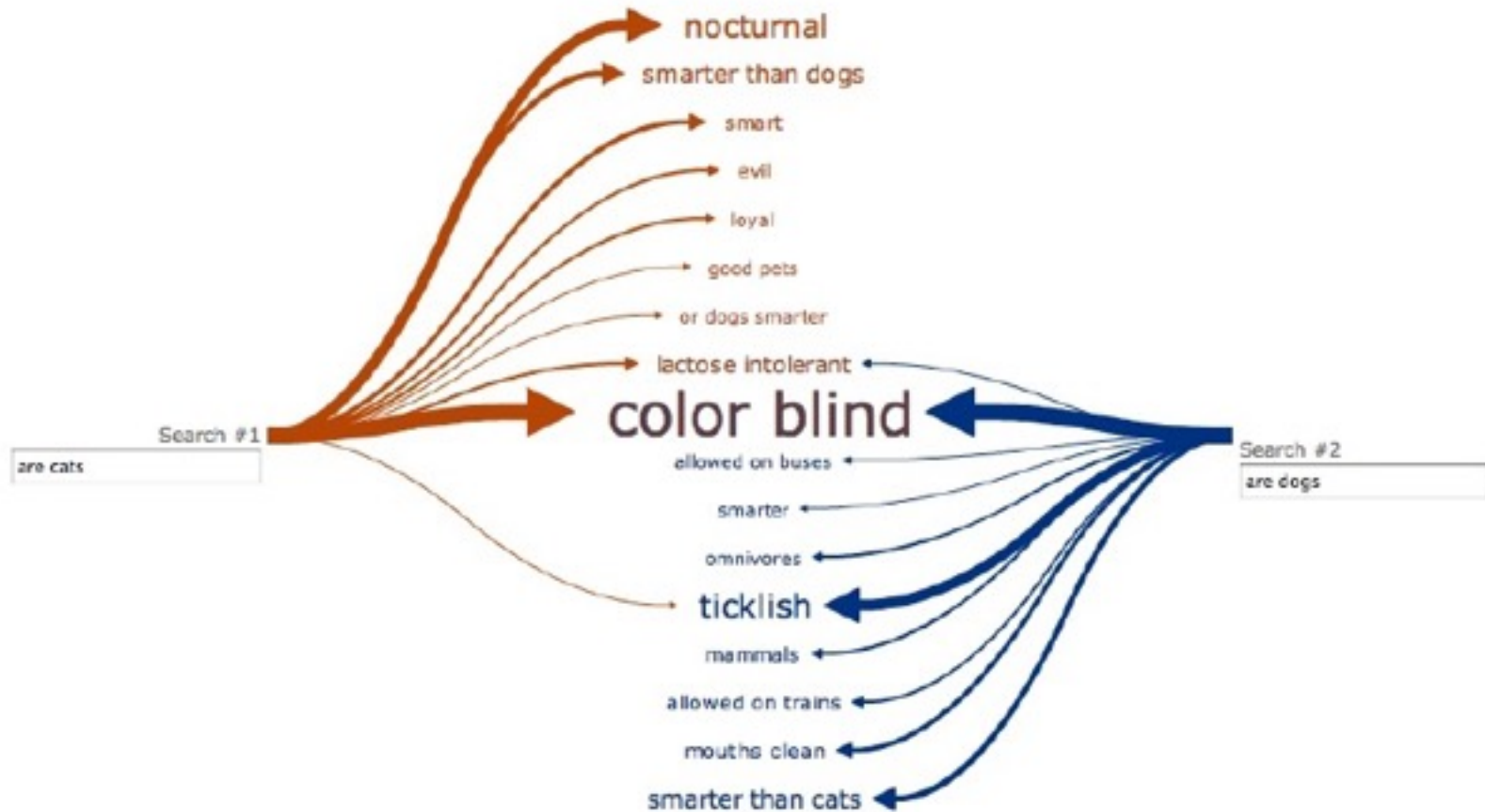


<http://hint.fm/projects/phrasenet/>

FERNANDA VIÉGAS

<http://fernandaviegas.com/>

Web Seer

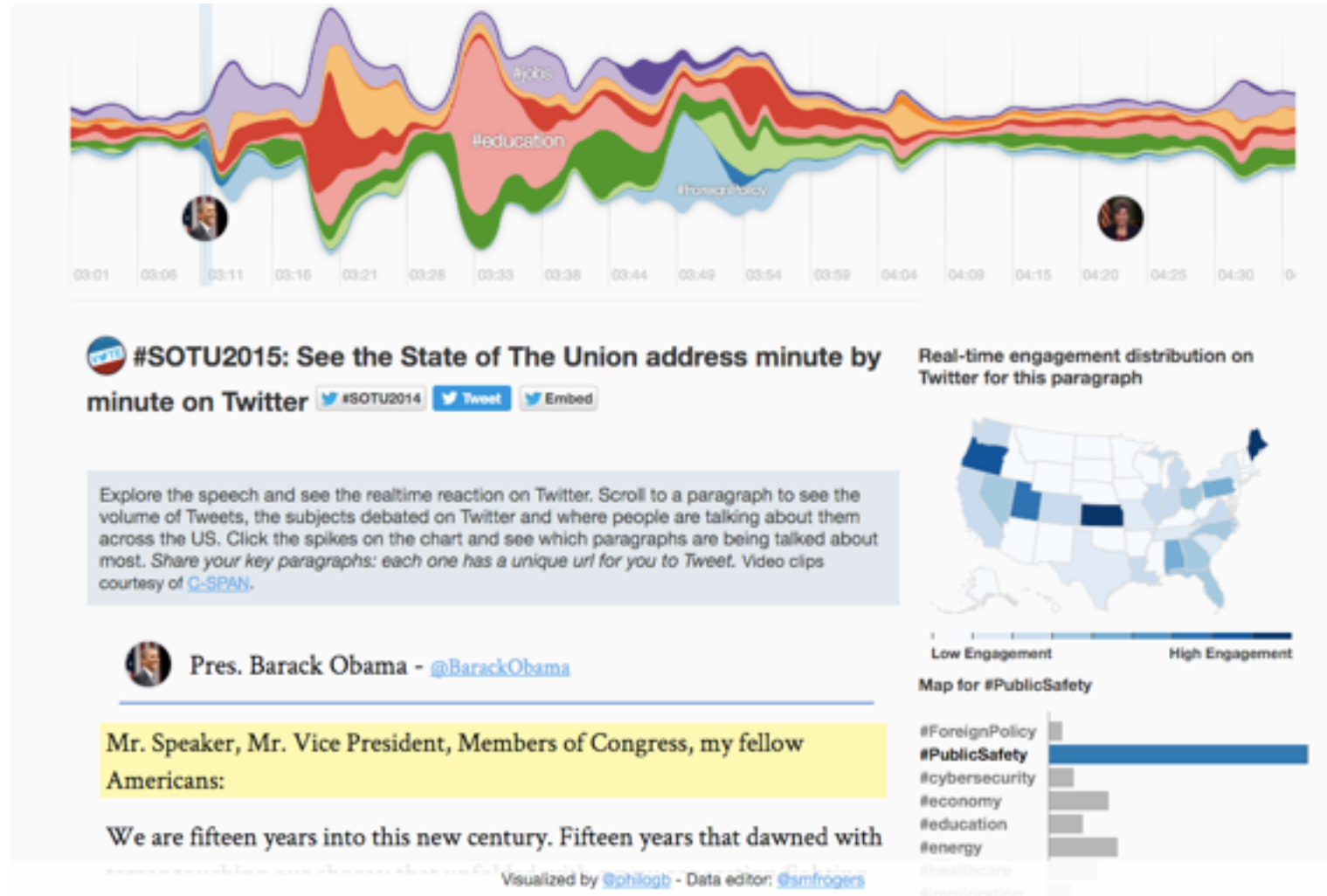


<http://hint.fm/projects/seer/>

NEW YORK TIMES

<http://www.nytimes.com>

State of the Union 2015




<http://twitter.github.io/interactive/sotu2015/#p1>

Inaugural Words


Inaugural Words: 1789 to the Present

Published: July 3, 2011

A look at the language of presidential inaugural addresses. The most-used words in each address appear in the interactive chart below, sized by number of uses. Words highlighted in yellow were used significantly more in this inaugural address than average. (Related Article)



1789 | 1800 | 1820 | 1840 | 1861 | 1880 | 1900 | 1920 | 1940 | 1960 | 1980 | 2000



1861
Abraham
Lincoln

Full text of the address
Article from the Times archive (pdf)

As war loomed, Abraham Lincoln warned that secession was illegal. "In your hands, my disaffected fellow countrymen, and not in mine, is the momentous issue of civil war," he said. "The government will not assail you. You can have no conflict without being yourselves the aggressors. You have no oath registered in heaven to destroy the government, while I shall have the most solemn one to 'preserve, protect, and defend it.'"

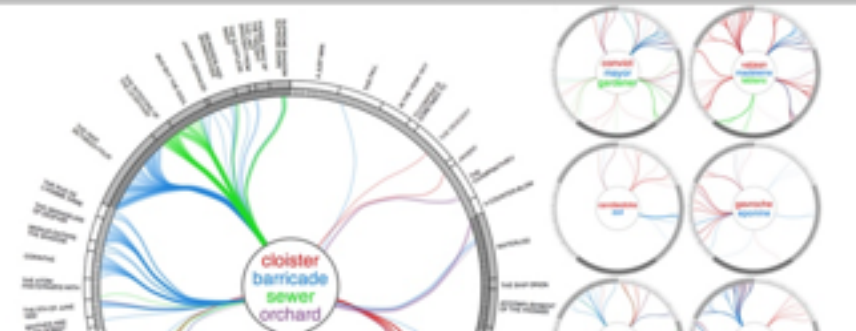
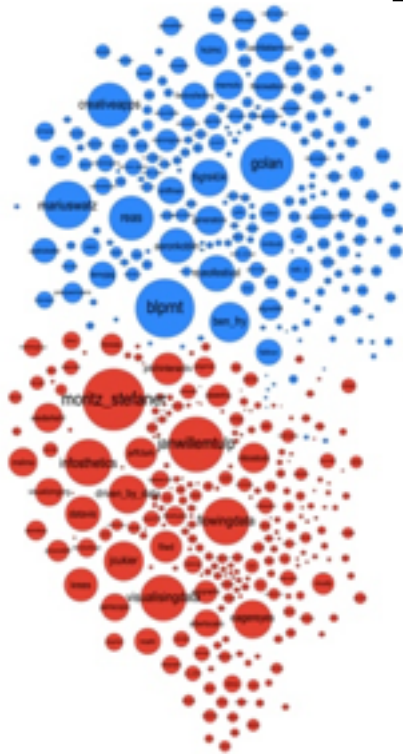
constitution law
union people
government states
right national provision minority state
slave constitutional citizen section authority
party question exist object purpose oath
administration majority unite great time surrender present
congress enforce fugitive labor clause plainly expressly view
service force hold continue amendment frame held secede written
change hand lawfully fly decision controversy vital act side power perfect
slavery office institution term express speak exercise proper duty believe

The New York Times

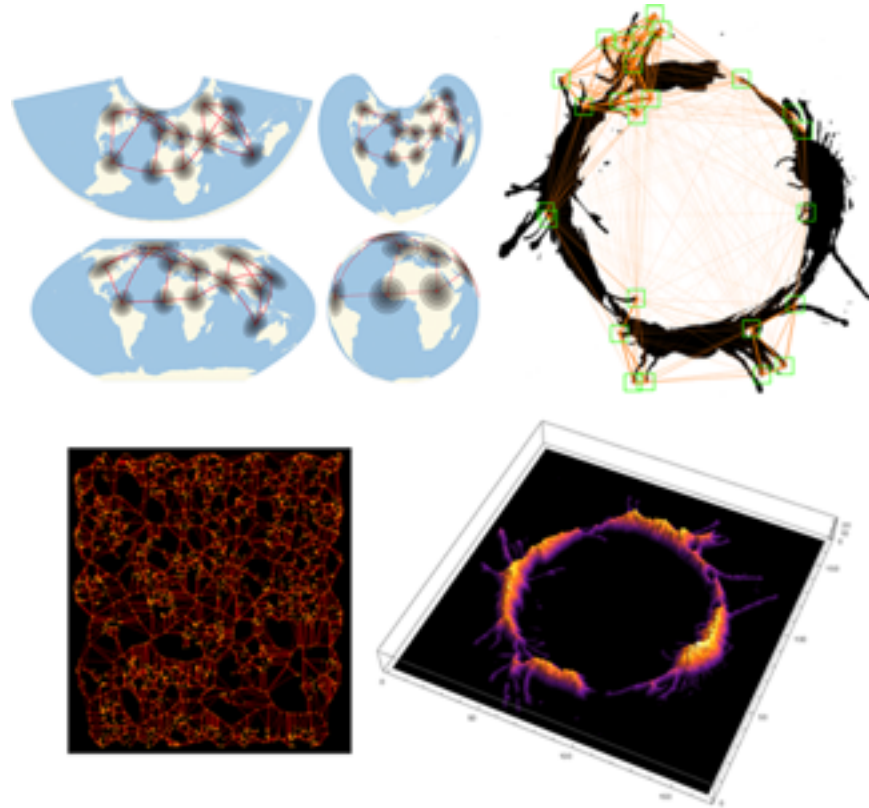
http://www.nytimes.com/interactive/2009/01/17/washington/20090117_ADDRESSES.html

MISCELLANEOUS

http://neoformix.com



What about representing Aliens' unknown languages?



<http://blog.stephenwolfram.com/2016/11/quick-how-might-the-alien-spacecraft-work/>

- Title Bars
<http://people.ischool.berkeley.edu/~hearst/research/tilebars.html>
- Comment Bars
<http://web.media.mit.edu/~dietmar/myspace.html>
- Paper Lens
<http://www.cs.umd.edu/hcil/paperlens/>
- In-Spire
<http://in-spire.pnnl.gov/>
- Jigsaw
<http://www.cc.gatech.edu/gvu/ii/jigsaw/>
- Text Pool
<http://designgraphics.ncsu.edu/pages/projects/textpool/>

QUESTIONS?

*Thanks to
Sophie J. Engle
San Francisco University*

for ideas, suggestions, slides, links, and much other stuff