Time Series Data Visualization

BAIS 6140 – Information Visualization

L. Miguel Encarnação

Time Series Data

- Fundamental chronological component to the data set
- Random sample of 4000 graphics from 15 of world's newspapers and magazines from '74-'80 found that 75% of graphics published were time series
  - Tufte, Vol. 1

Data Sets

- Each data case is likely an event of some kind
- One of the variables can be the date and time of the event
- Example: sunspot activity, baseball games, medicines taken, cities visited, stock prices, etc.

Meta Level

- Consider multiple stocks being examined
- Is each stock a data case, or is a price on a particular day a case, with the stock name as one of the other variables?
- Confusion between data entity and data cases

Data Mining

- Data mining domain has techniques for algorithmically examining time series data, looking for patterns, etc.
- Good when objective is known a-priori
- But what if not?
  - Which questions should I be asking?
  - InfoVis better for that

User Tasks

- What kinds of questions do people ask about time series data?
- Examples
  - When was something greatest/least?
  - Is there a pattern?
  - Are two series similar?
  - Do any of the series match a pattern?
  - Provide simpler, faster access to the series

Other Tasks

- Does data element exist at time t ?
- When does a data element exist?
- How long does a data element exist?
- How often does a data element occur?
- How fast are data elements changing?
- In what order do data elements appear?
- Do data elements exist together?

Muller & Schumann ’03
citing MacEachern ‘95

Taxonomy

- Discrete points vs. interval points
- Linear time vs. cyclic time
- Ordinal time vs. continuous time
- Ordered time vs. branching time vs. time with multiple perspectives

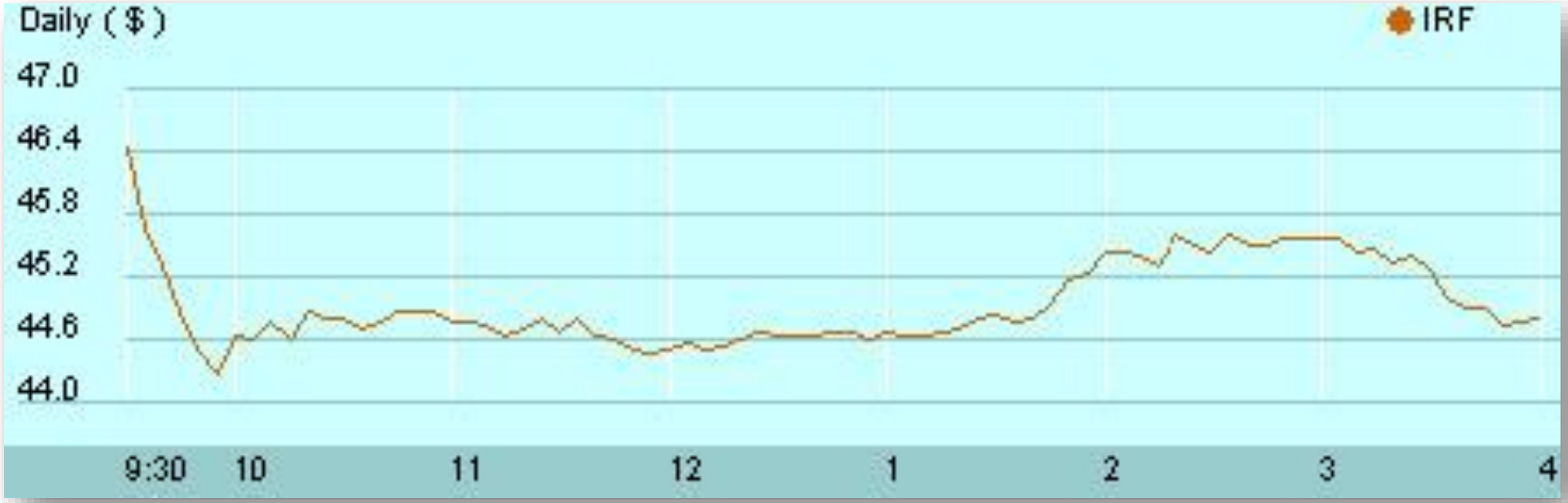
Muller & Schumann ’03
citing Frank ‘98

Fundamental Tradeoff

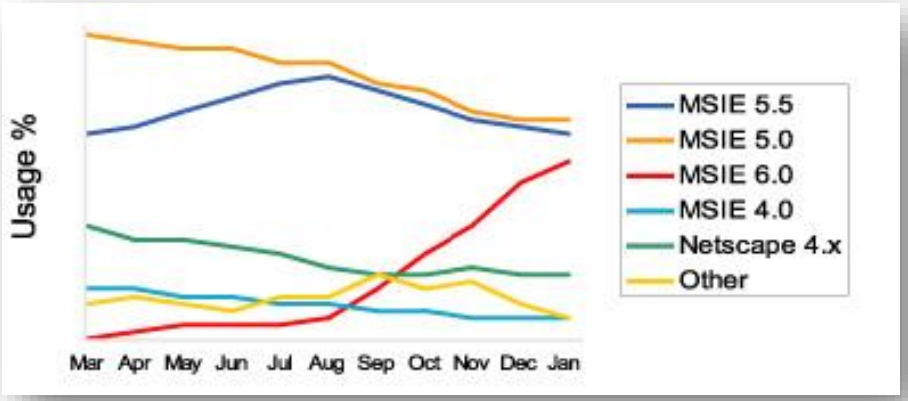
- Is the visualization time-dependent, i.e., changing over time (beyond just being interactive)
  - Static
    - Shows history, multiple perspectives, allows comparison
  - Dynamic (animation)
    - Gives feel for process & changes over time, has more space to work with

Standard Presentation

- Present time data as a 2D line graph with time on x-axis and some other variable on y-axis

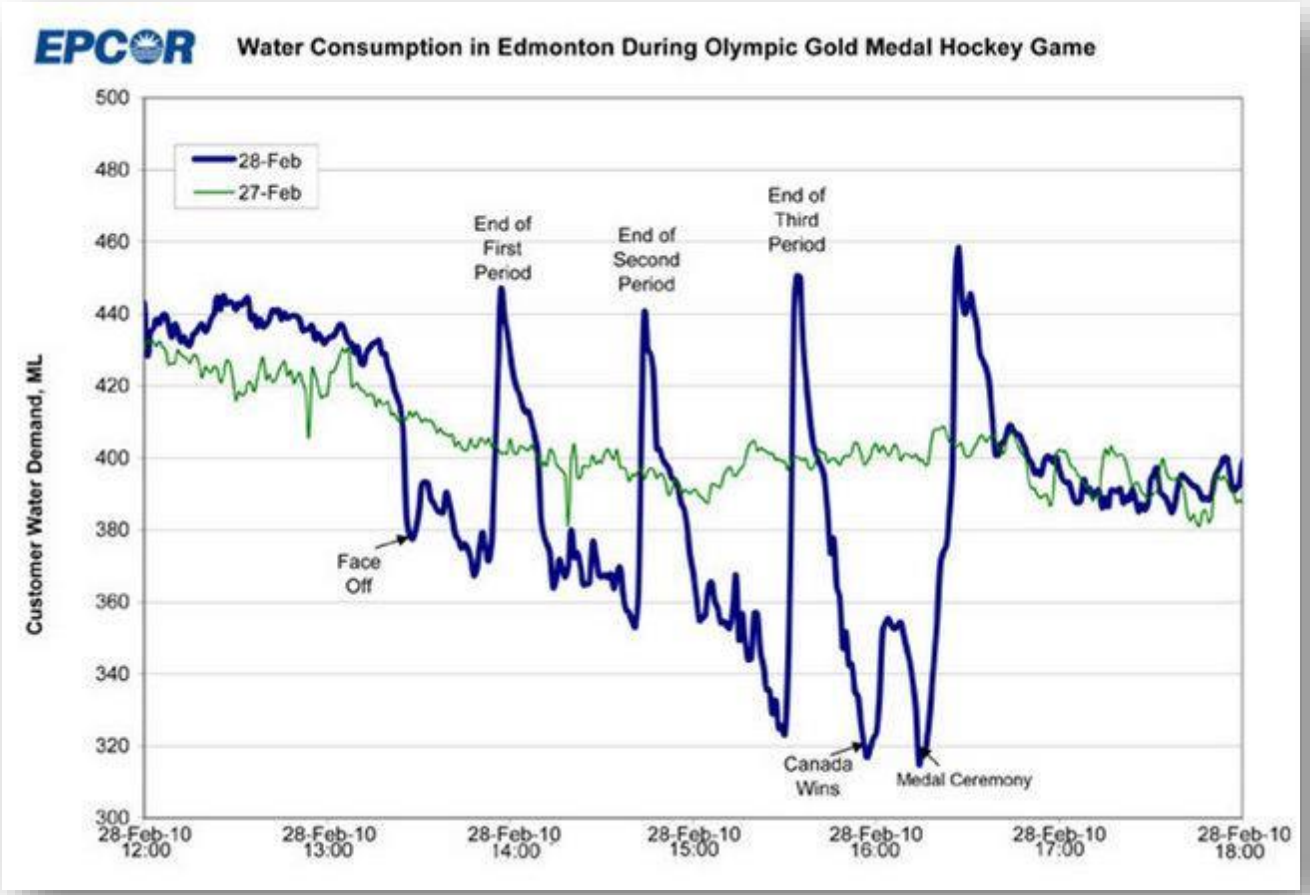


Classic Views



Fun One

- What If Everybody in Canada Flushed At Once?



http://www.patspapers.com/blog/item/what_if_everybody_flushed_at_once_Edmonton_water_gold_medal_hockey_game/

Today's Focus

- Examination of a number of techniques an d case studies
- Learn from some of the different visualization ideas that have been created
- Can you generalize these techniques into classes or categories?

Nice Overview

Visual Methods for Analyzing Time-Oriented Data

Wolfgang Aigner, Silvia Miksch, Wolfgang Müller, Heidrun Schumann, and Christian Tominski

Abstract—Providing appropriate methods to facilitate the analysis of time-oriented data is a key issue in many application domains. In this paper, we focus on the unique role of the parameter time in the context of visually driven data analysis. We will discuss three major aspects—visualization, analysis, and the user. It will be illustrated that it is necessary to consider the characteristics of time when generating visual representations. For that purpose, we take a look at different types of time and present visual examples. Integrating visual and analytical methods has become an increasingly important issue. Therefore, we present our experiences in temporal data abstraction, principal component analysis, and clustering of larger volumes of time-oriented data. The third main aspect we discuss is supporting user-centered visual analysis. We describe event-based visualization as a promising means to adapt the visualization pipeline to needs and tasks of users.

Index Terms—Time-oriented data, visualization, analysis, user.

1 Introduction and Motivation

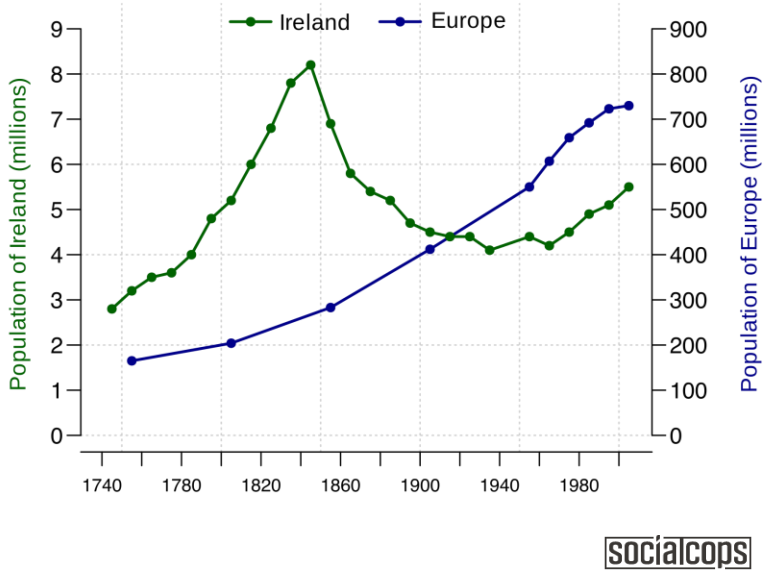
Considering the characteristics of data is vital when designing visual representations. A salient characteristic is whether or not data are related to time. That time is an outstanding dimension is reflected by Shneiderman's Taskby Data Type Taxonomy [1], where temporal data are identified as one of seven basic data types. Nowadays, time-oriented data are ubiquitous in many application domains as, for example, in business, medicine, history, planning, or project management. For a long time, visual methods have been successfully applied to analyze such data. A wide repertoire of interactive techniques for visualizing data sets with temporal dependencies is available. However, many current

- visualization,
- analysis, and
- user.

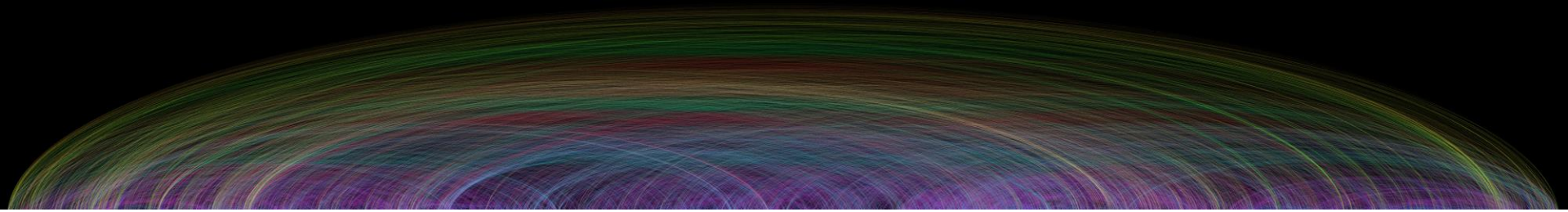
In Section 2, we focus on visualization methods for time-oriented data. We will show that the term *time-oriented data* comprises several types of data with different meanings and applications. Designing or applying visual representations can only be successful if one is aware of these different types. This will be demonstrated with several examples of visualization techniques that stem from our own work or are available in the literature.

Usually, time-oriented data are large—not only in terms

Line graph



- Traps
  - Too many variables
  - Multiple axes
  - Ill-formatted time axis
- Best practice
  - No more than 4 variables
  - Axis consistency in scale and color

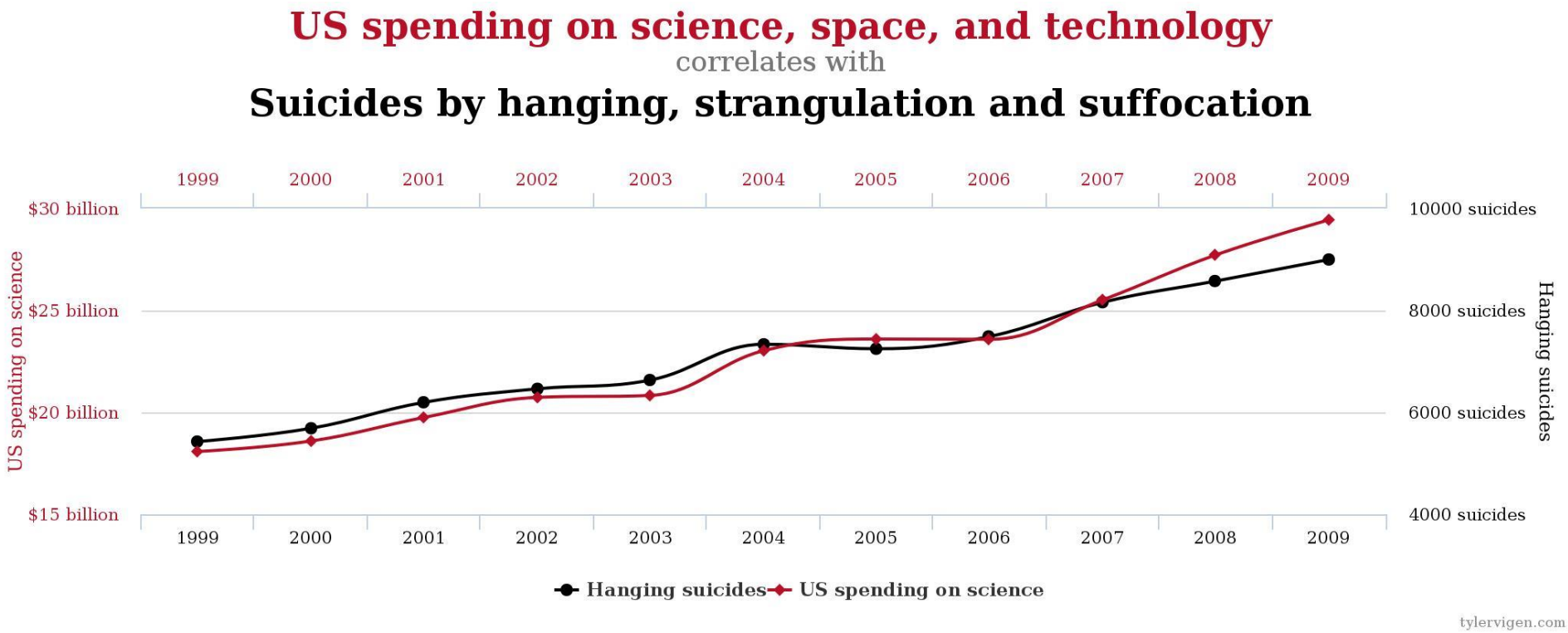


Excursus

CAUSALITY DANGER ZONE

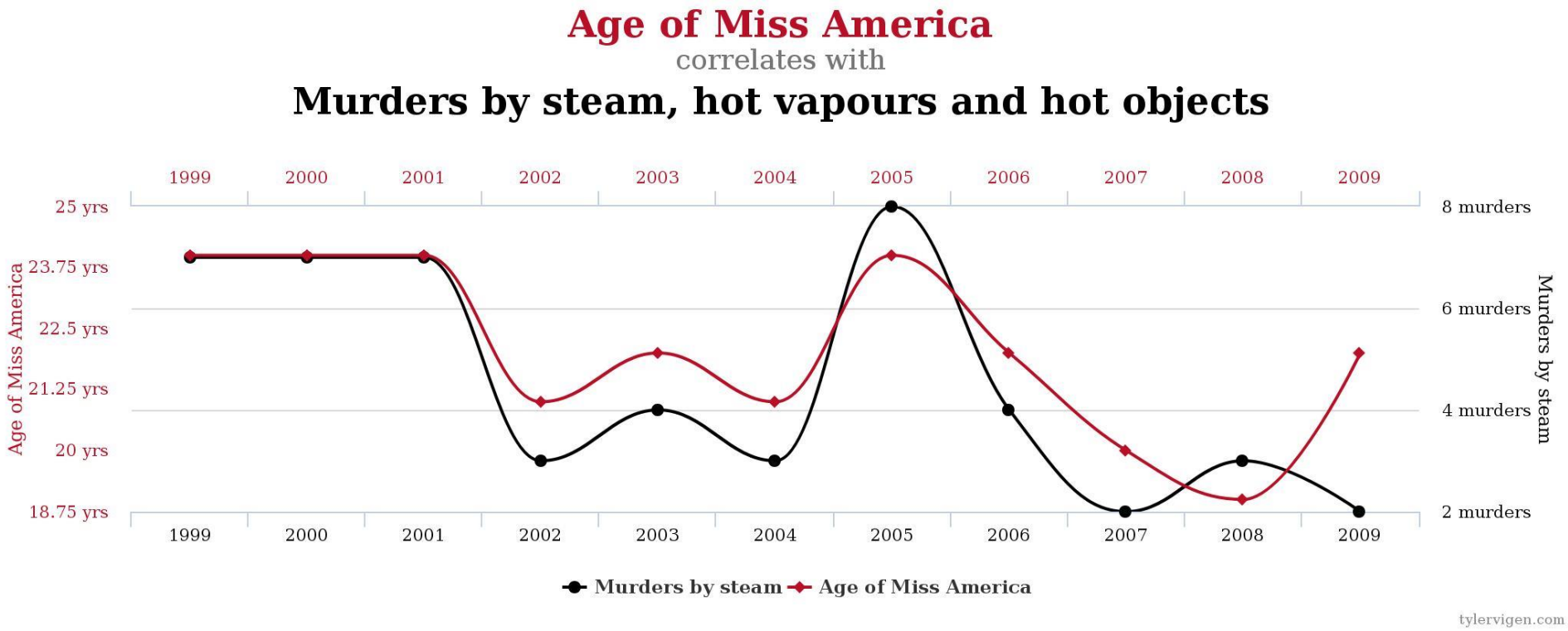
Spurious correlations

http://www.tylervigen.com/spurious-correlations



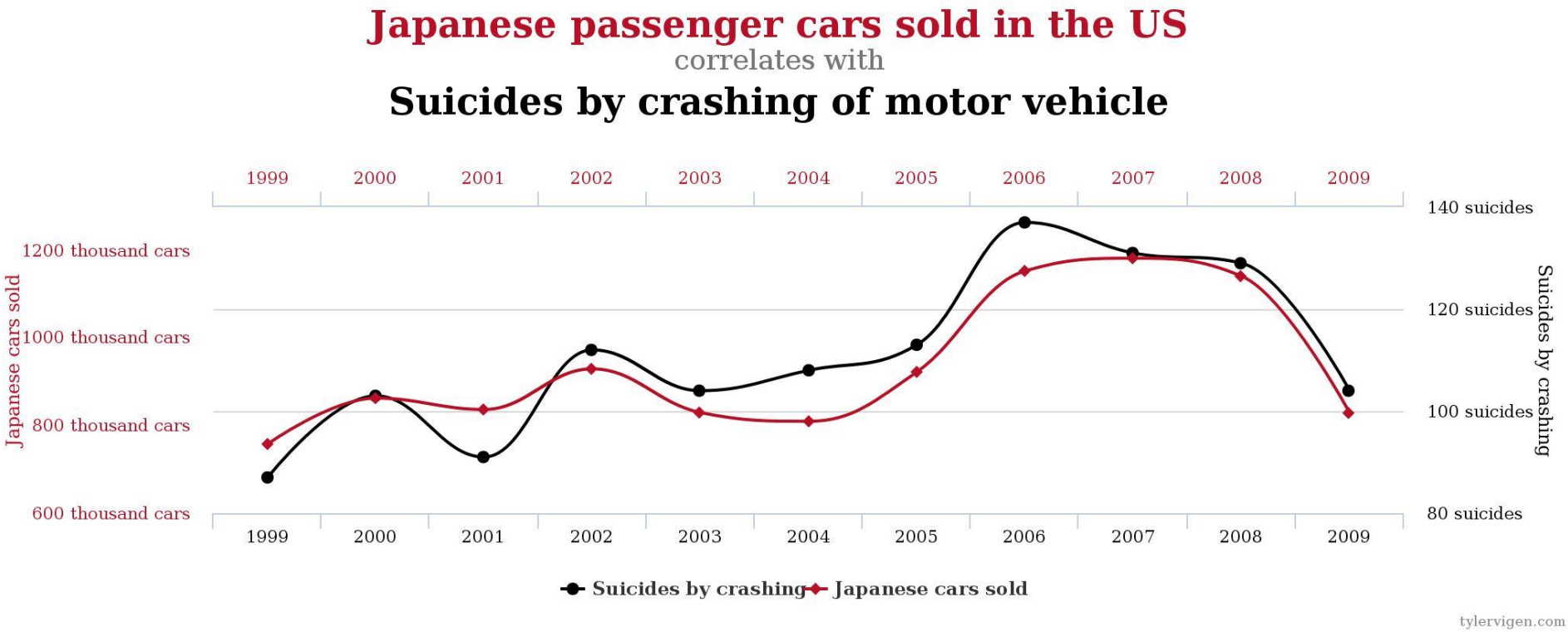
Spurious correlations

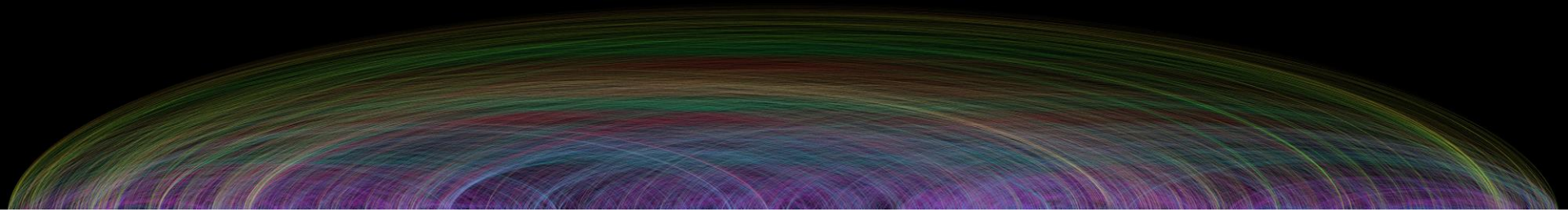
http://www.tylervigen.com/spurious-correlations



Spurious correlations

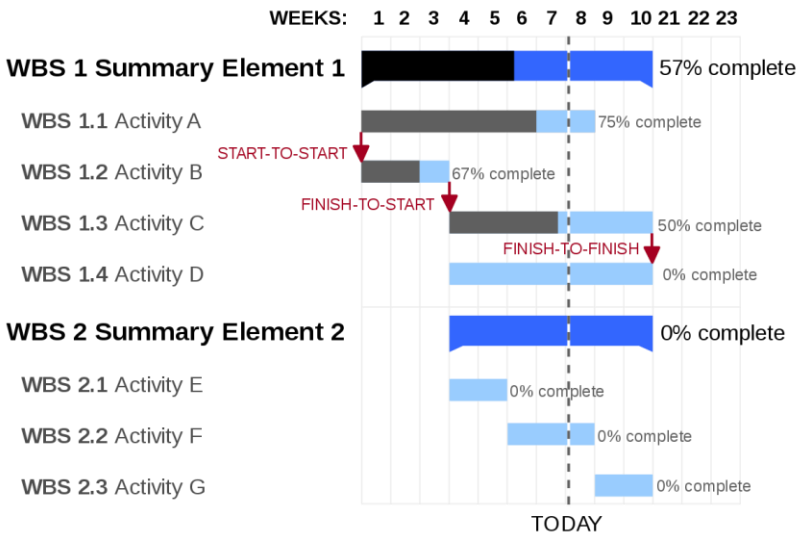
http://www.tylervigen.com/spurious-correlations





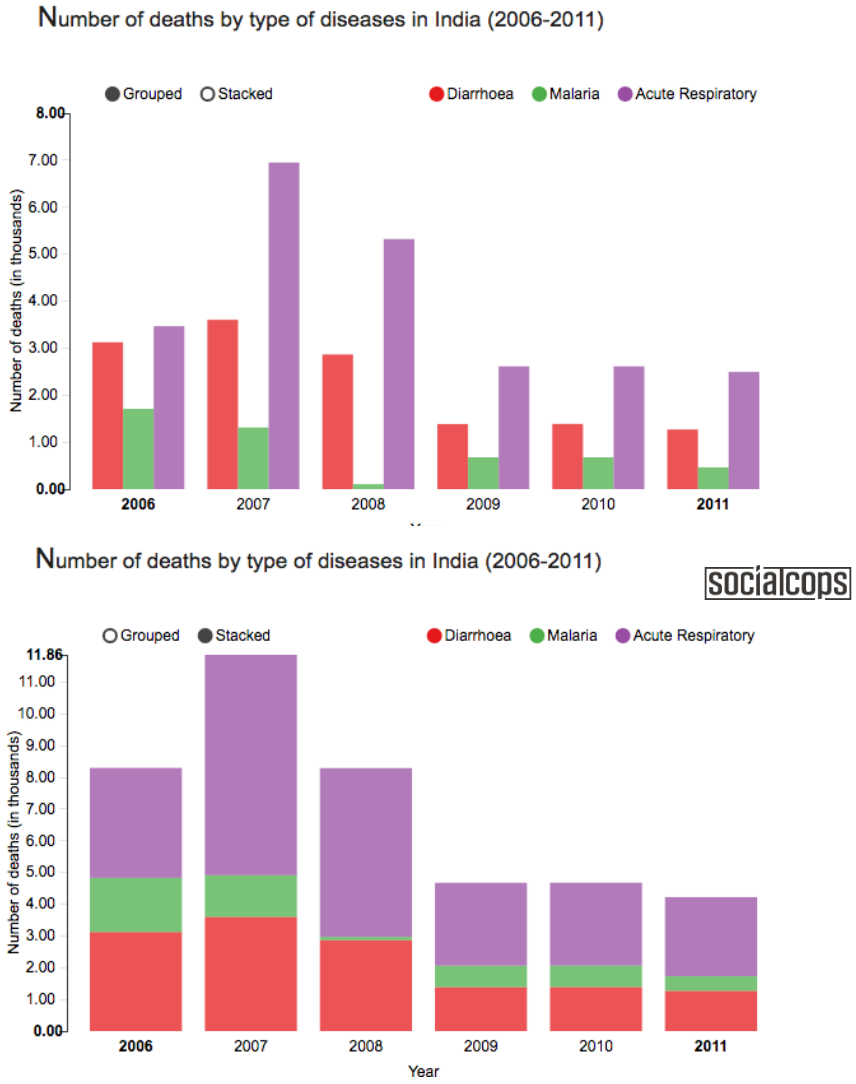
Time to get back to visualization techniques

Gantt chart



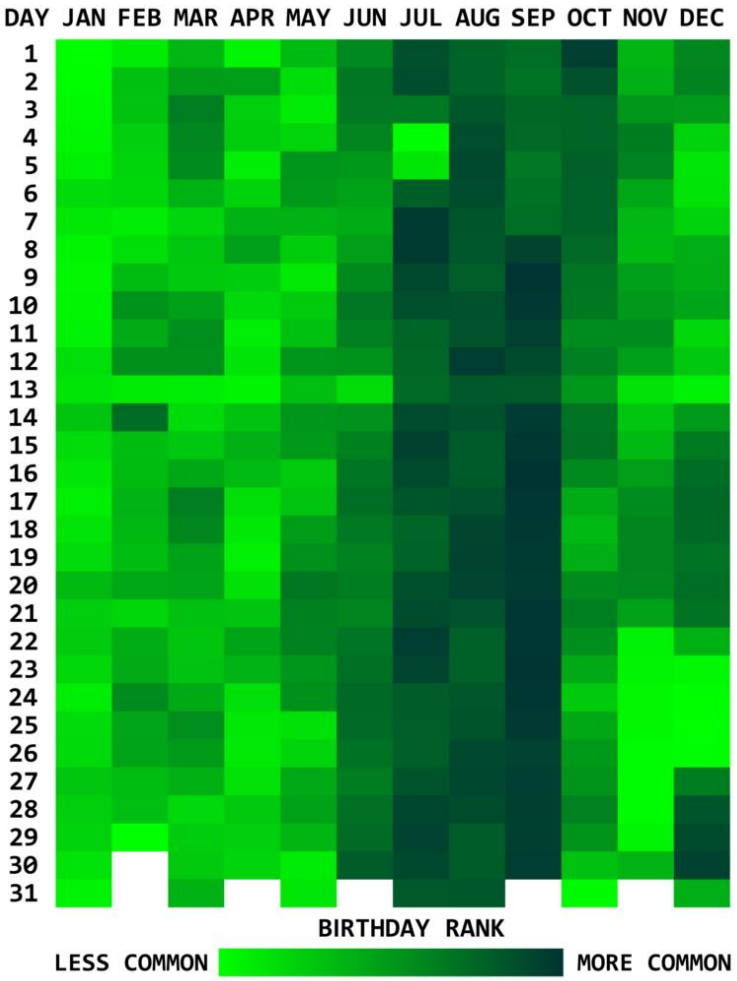
- Traps
  - Sequential perception breaks down with dependencies
  - Complexity increase with multiple projects
- Best practice
  - Display select levels
  - Color states not tasks
  - Diligent update
  - Digital only

Bar graph



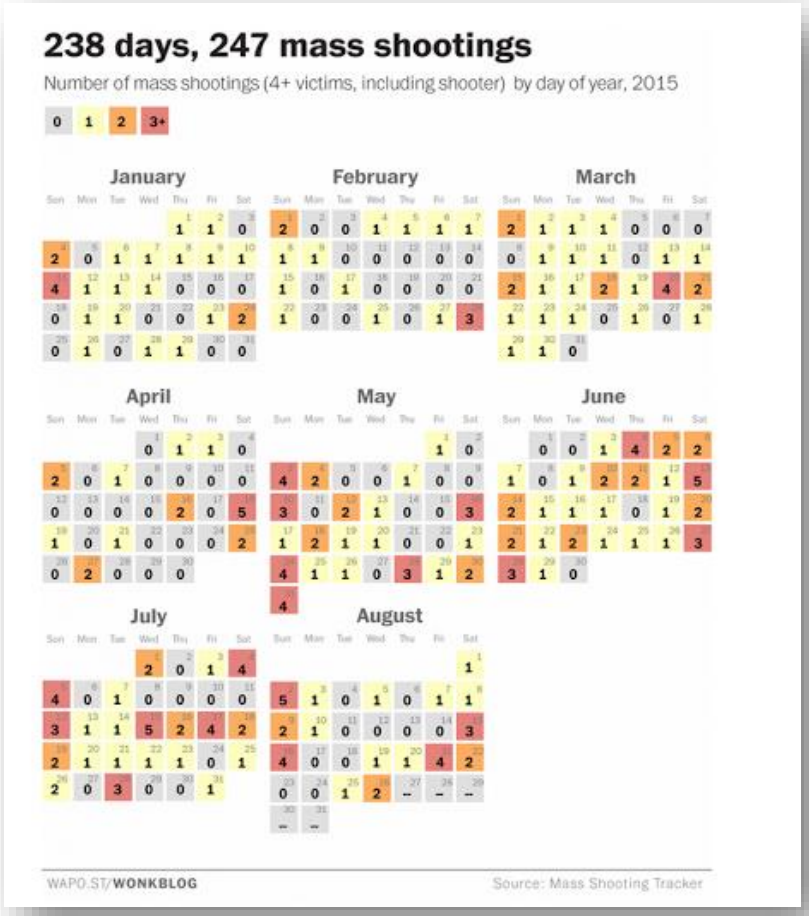
- Traps
  - Base (y) axis truncated
  - Too many variables
  - Clustered bars not well delineated
- Best practices
  - No more than 3 variables
  - Bold discernable colors
  - Sufficient spacing between (clusters of) bars

Heat map



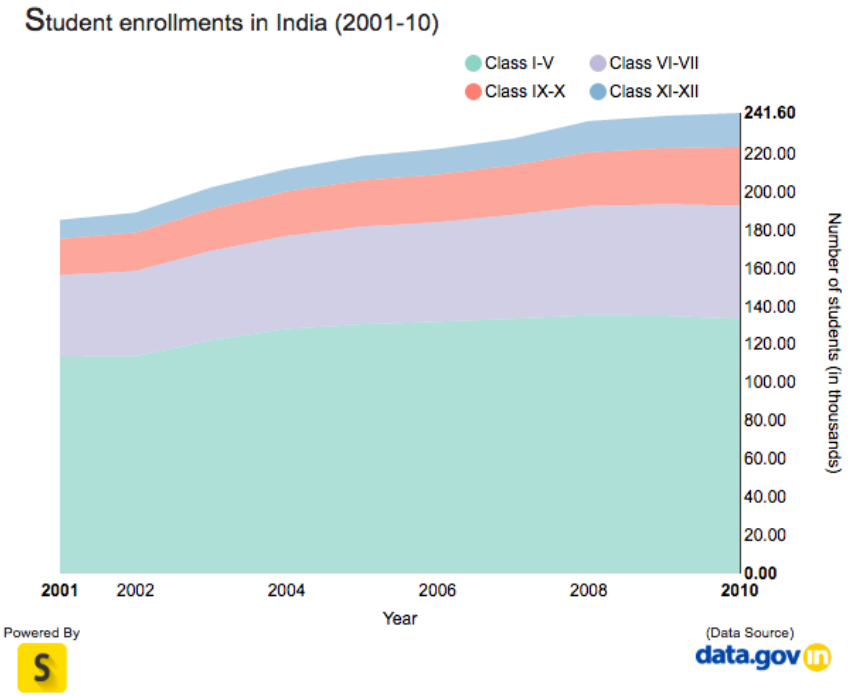
- Traps
  - Color/brightness constancy illusion
  - Apparent motion when animated over time
- Best practices
  - 2D matrix of 2 different time dimensions
  - Only 1 variable displayed

Heat map calendars (Tableau Public)

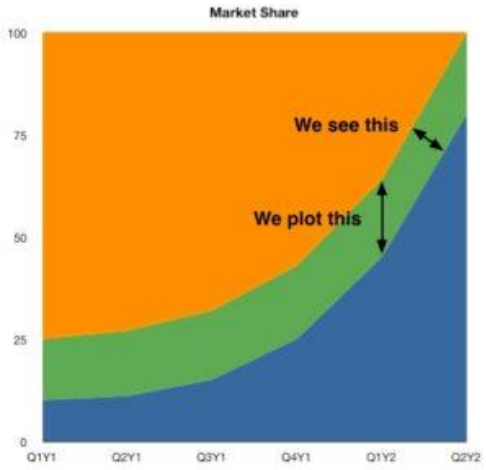


https://public.tableau.com/profile/kevin.taylor#!/vizhome/HeatmapCalendars/Example

Stacked area graph

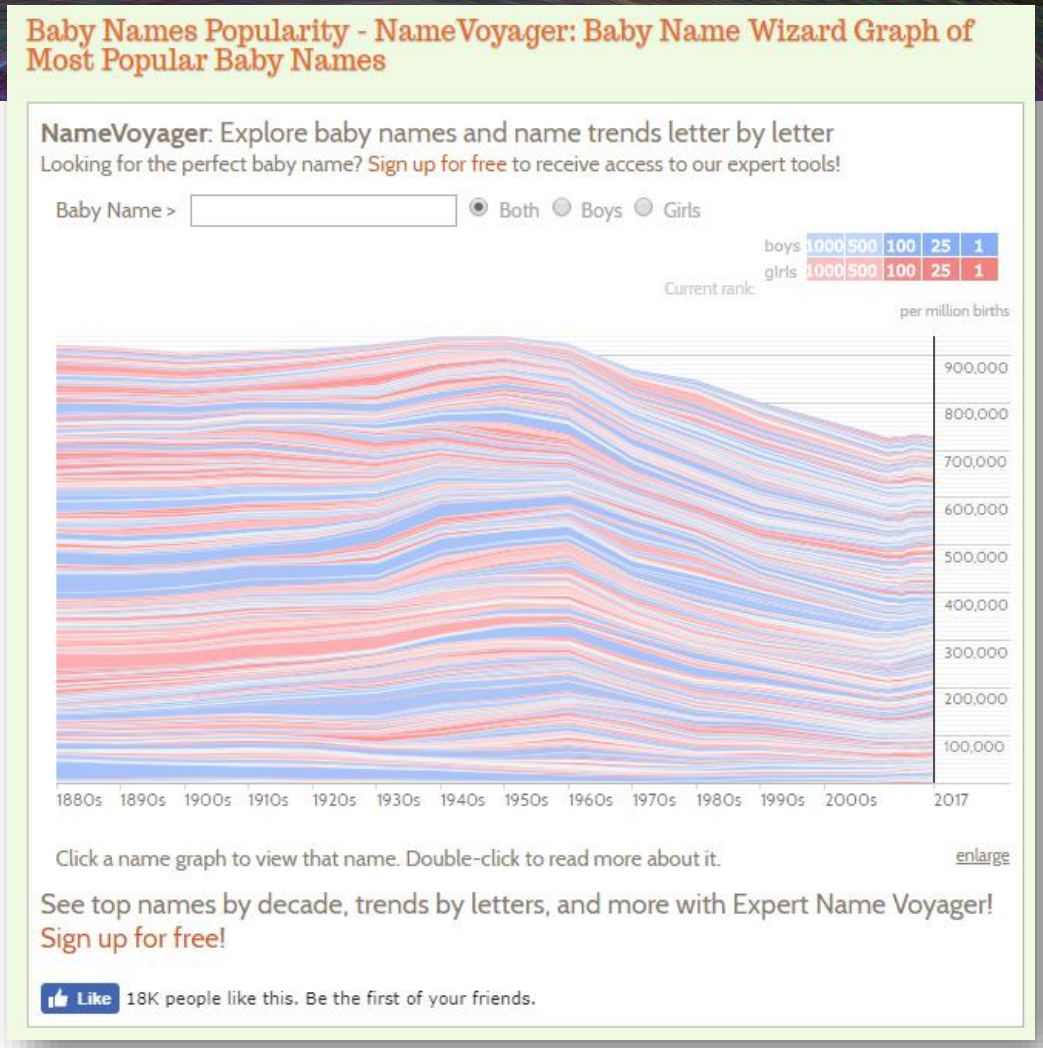


- Traps
  - Perception of differences



- Best practices
  - Useful to show how both a cumulative total and individual components of that total change over time

Example

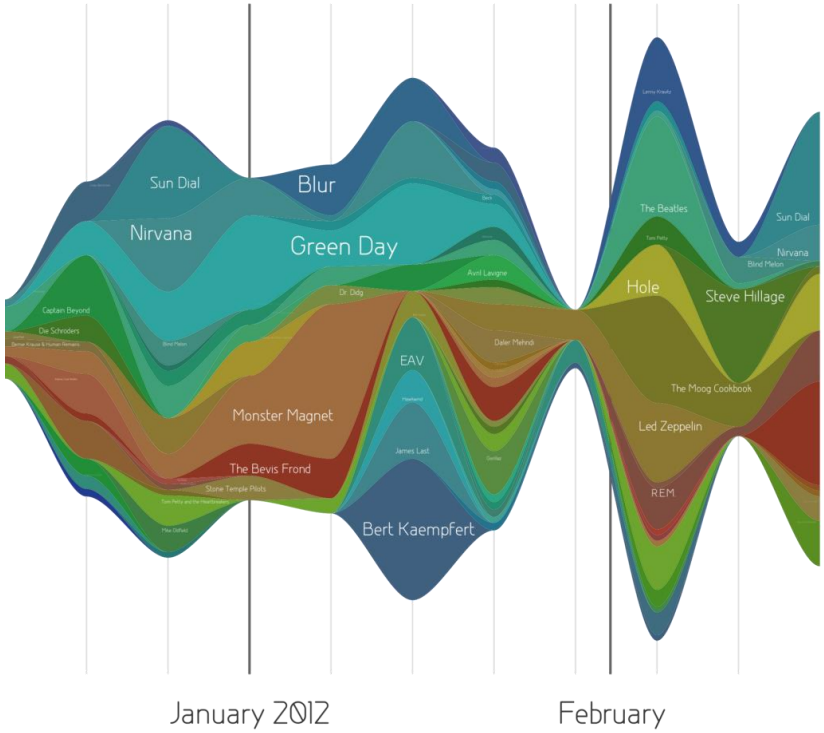


http://www.babynamewizard.com/voyager

Streamgraph (Byron & Wattenberg, *TVCG* '08)

- Similar idea – Stacked graph
- Goals:
  - Show multiple time series
  - Be able to see sum
  - Make labels legible
  - Be able to distinguish different layers
  - Make it aesthetically pleasing

Stream graph



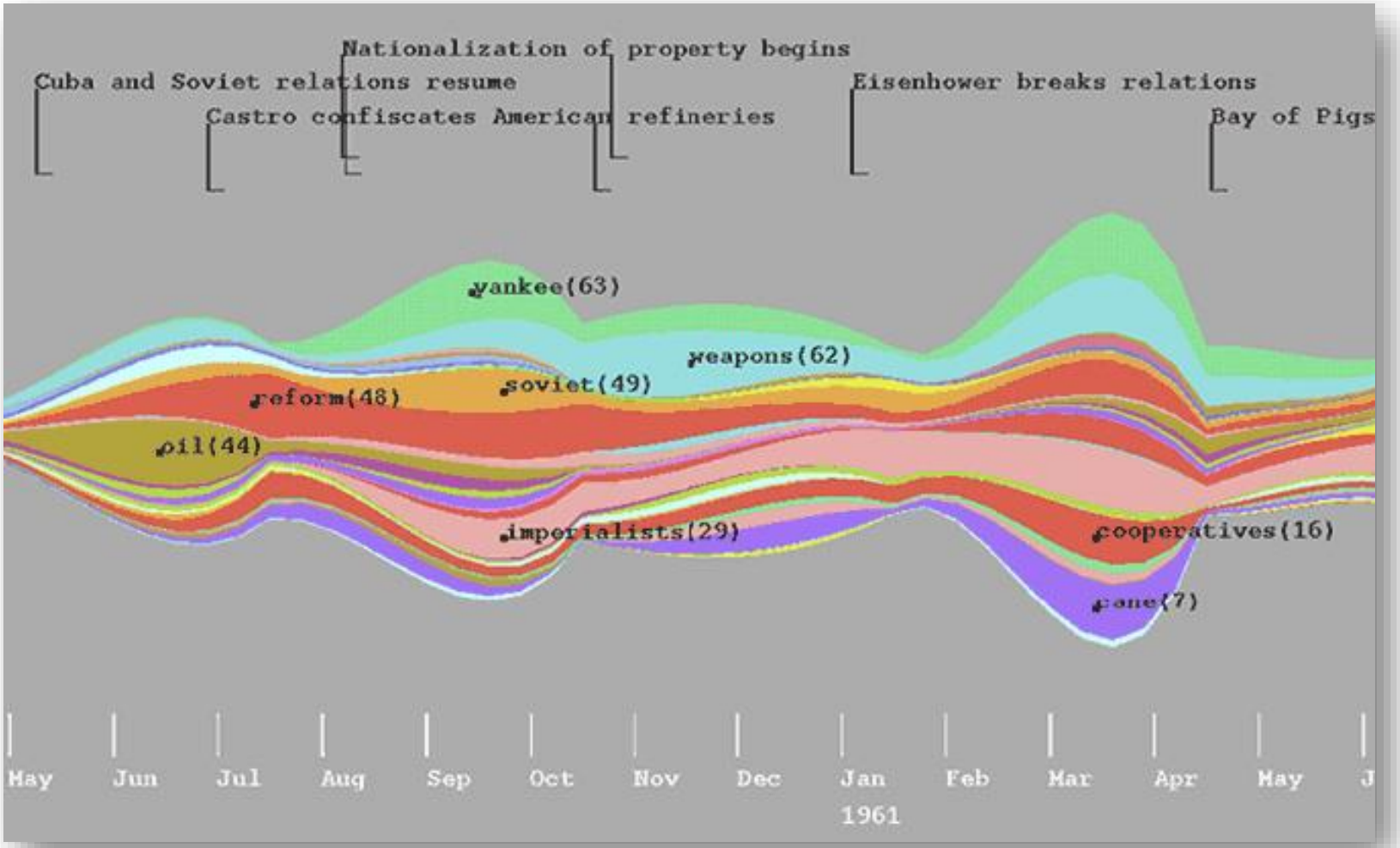
- Traps
  - Variables with small values tend to disappear
- Best practices
  - Multiple variables / large data sets
  - Color choice important

Design Issues

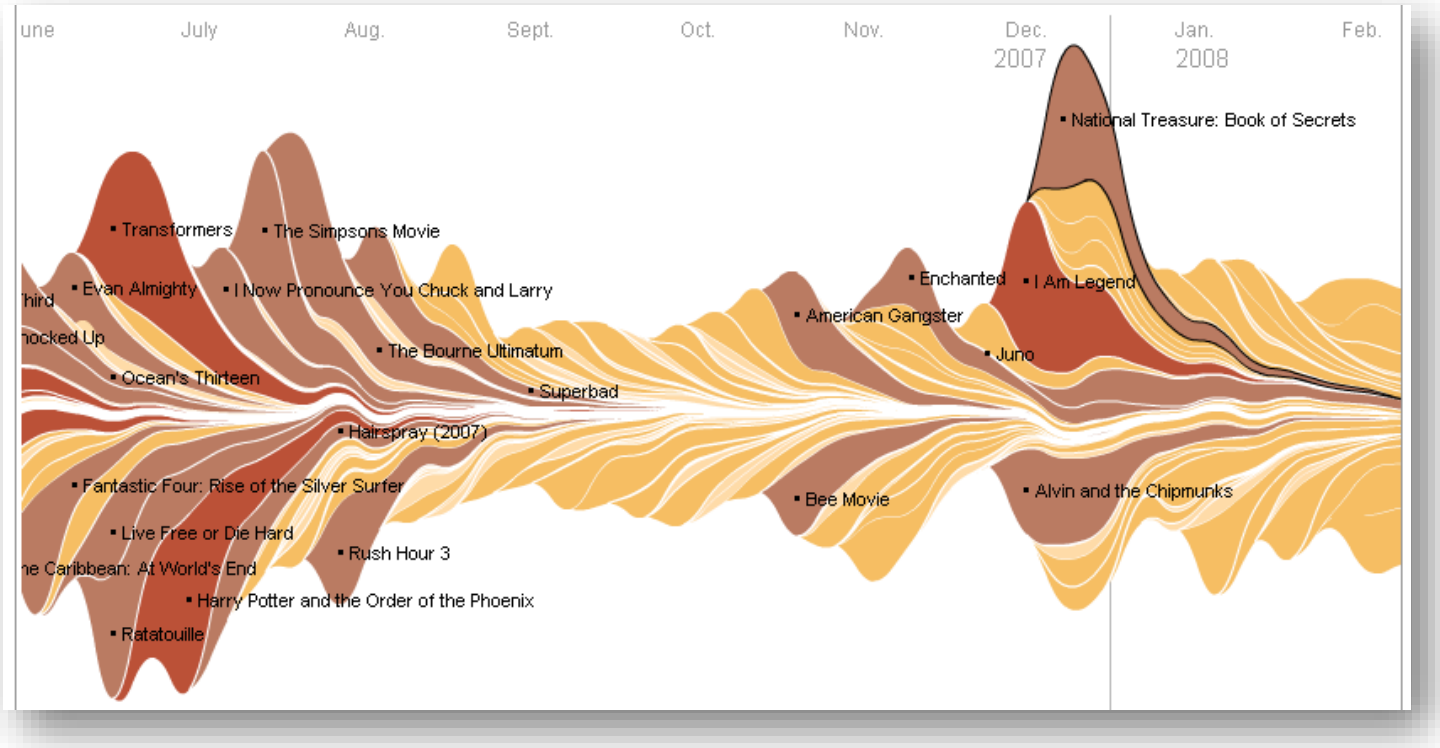
- Curve shape
  - Wiggle, symmetry, balance
  - Definitely some interesting math to do it
- Color choice
- Labeling
- Layer ordering

- *L. Byron & M. Wattenberg* paper provides very nice discussion of this

ThemeRiver (Havre et al., InfoVis '00)

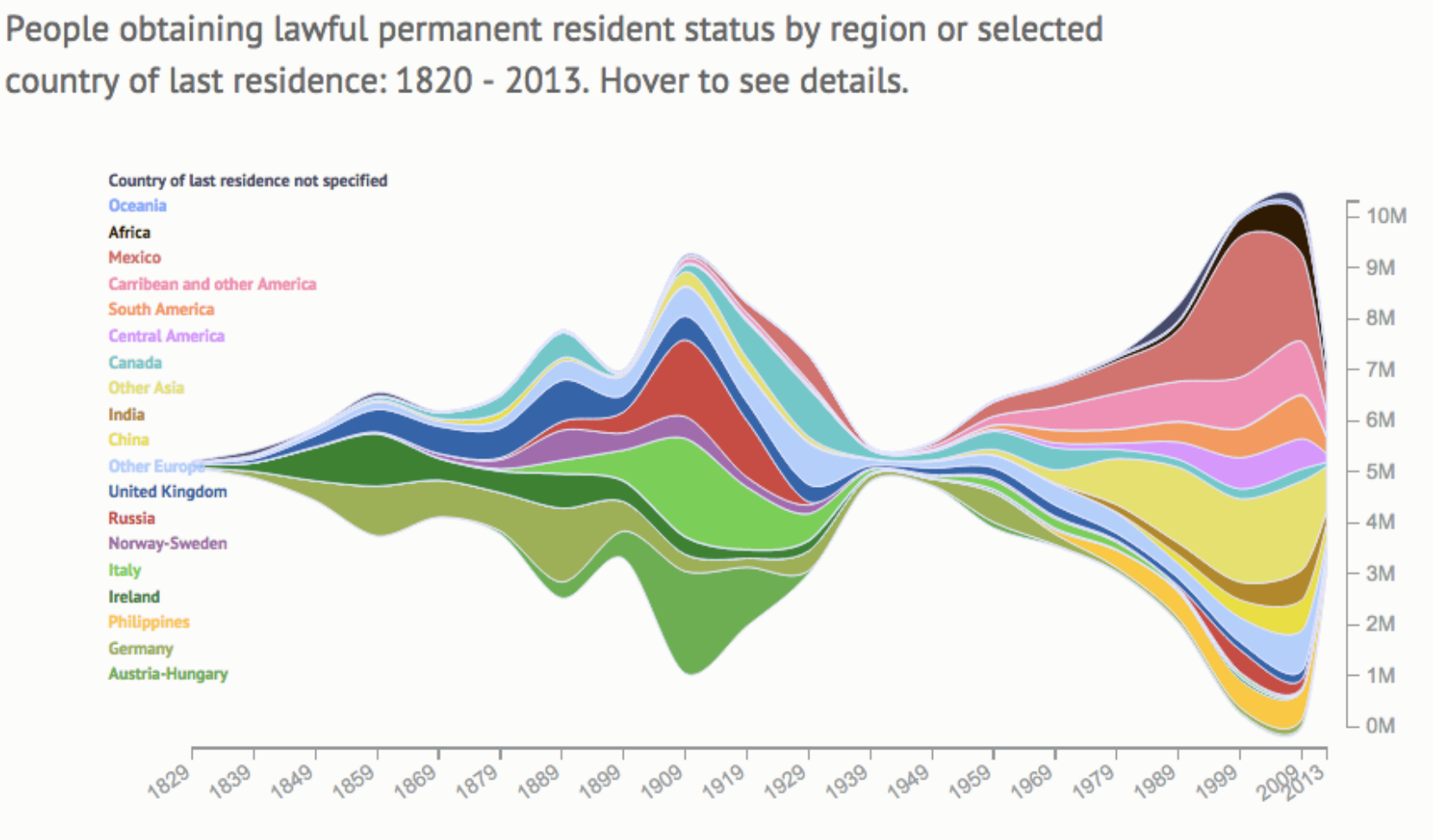


Interactive Application



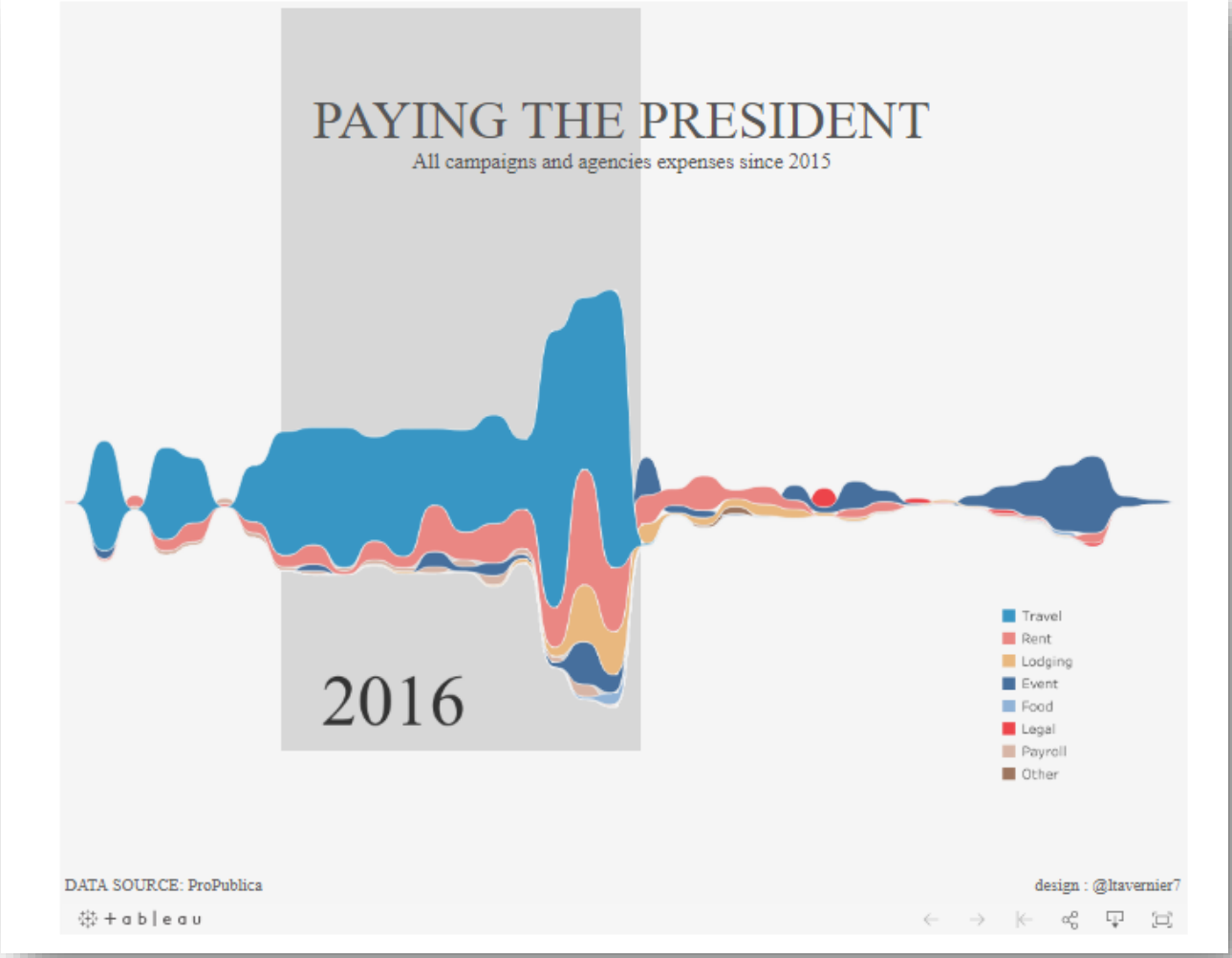
http://www.nytimes.com/interactive/2008/02/23/movies/20080223_REVENUE_GRAPHIC.html

200 Years of US Immigration



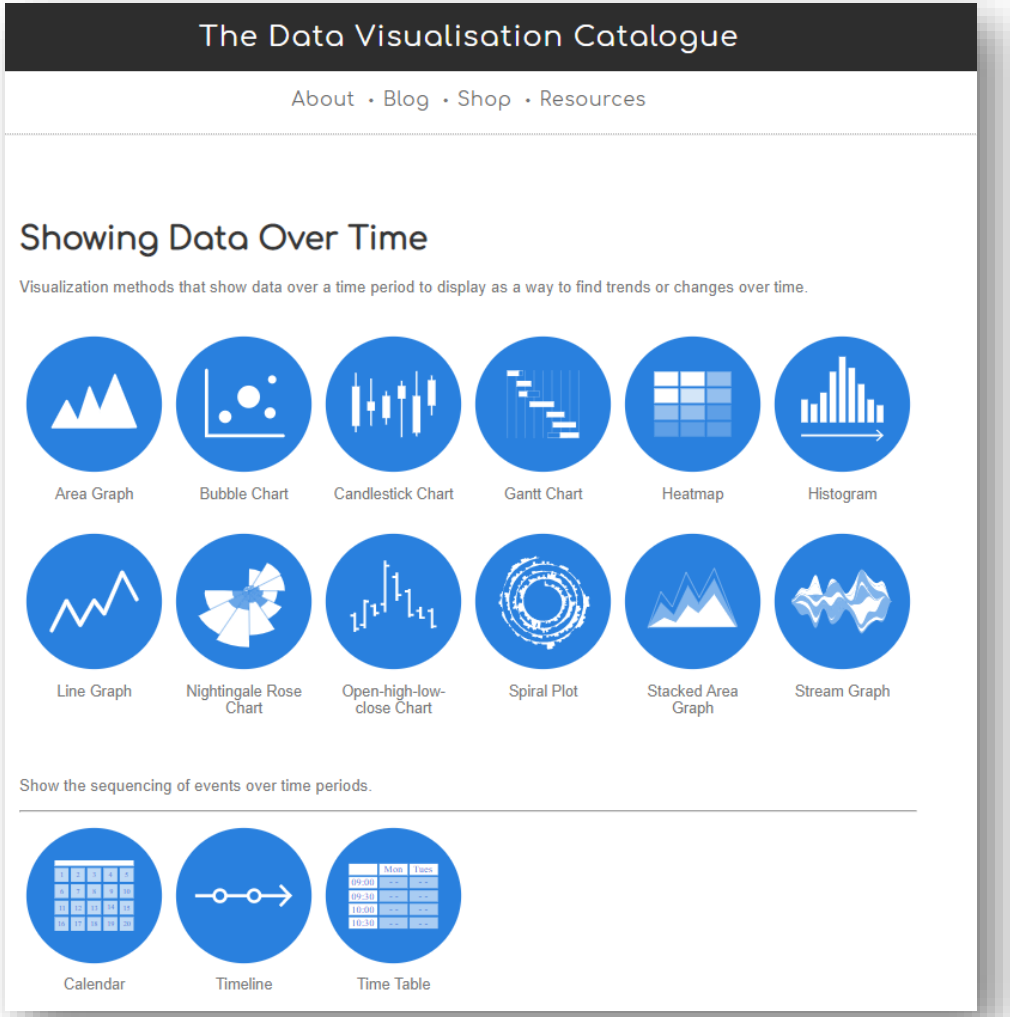
http://insightfulinteraction.comimmigration200years.html

Stream Graph (Tableau Public)

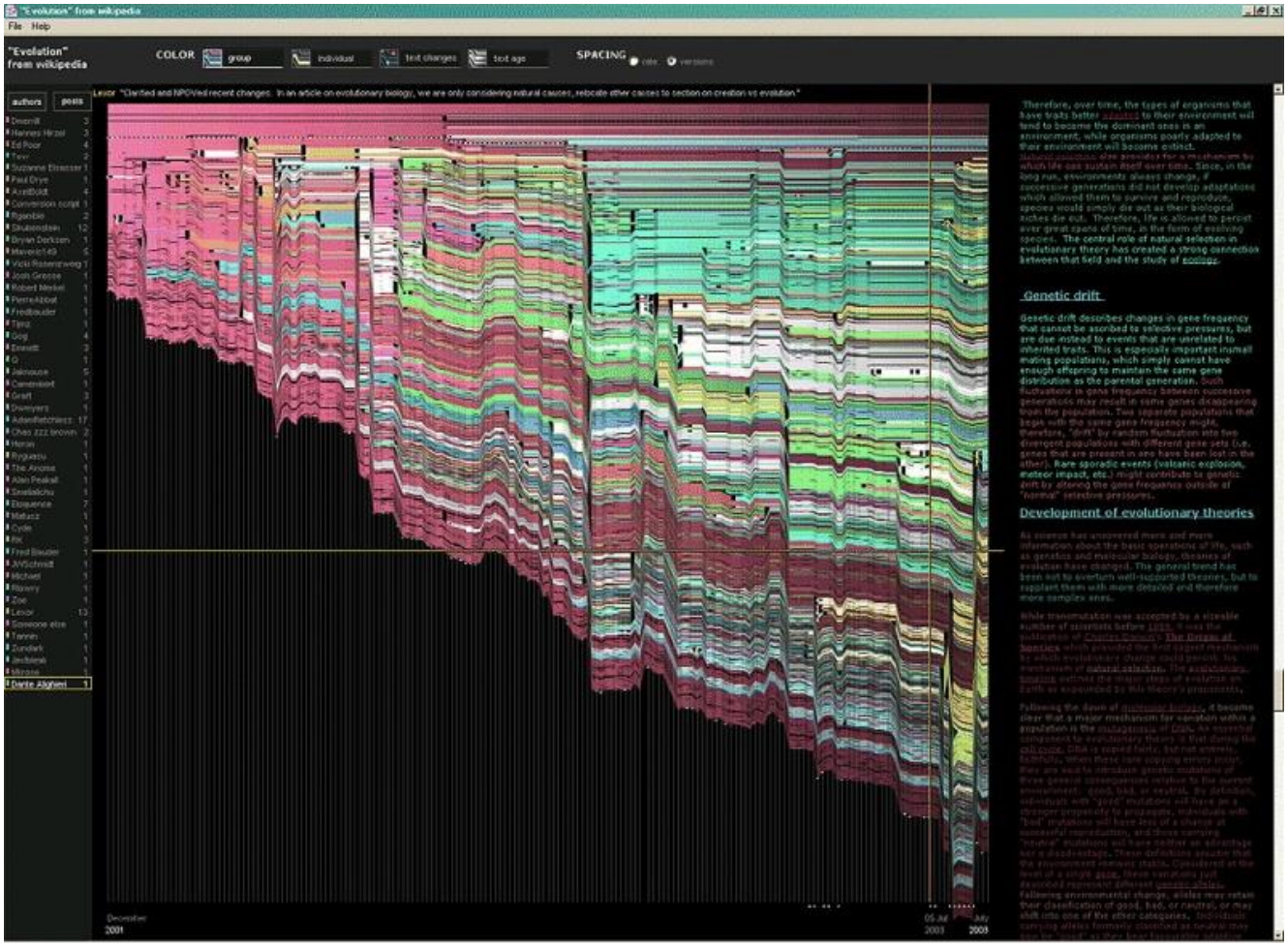


https://public.tableau.com/profile/ludovic.tavernier#!/vizhome/MakeOverMonday-PayingThePresident/PayingThePresident
https://greatified.com/2018/09/17/how-to-build-a-stream-graph-in-tableau-software/

Time Visualization in Business Practice



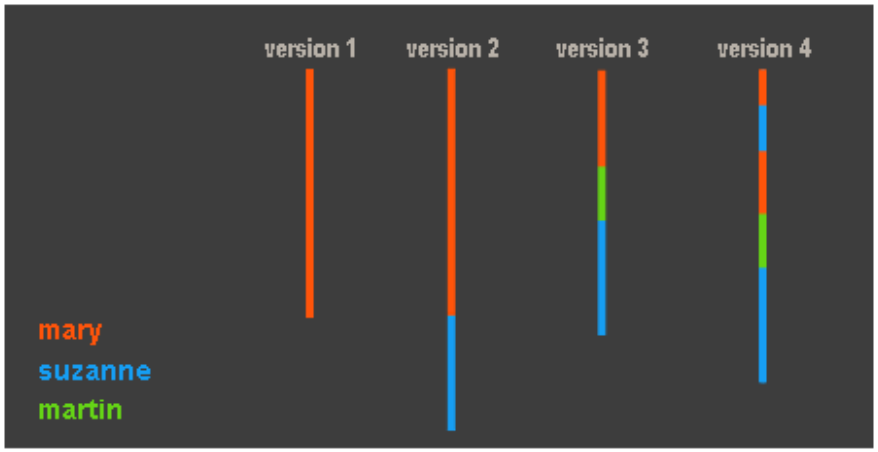
Flow of changes



- Flow of changes across electronic documents

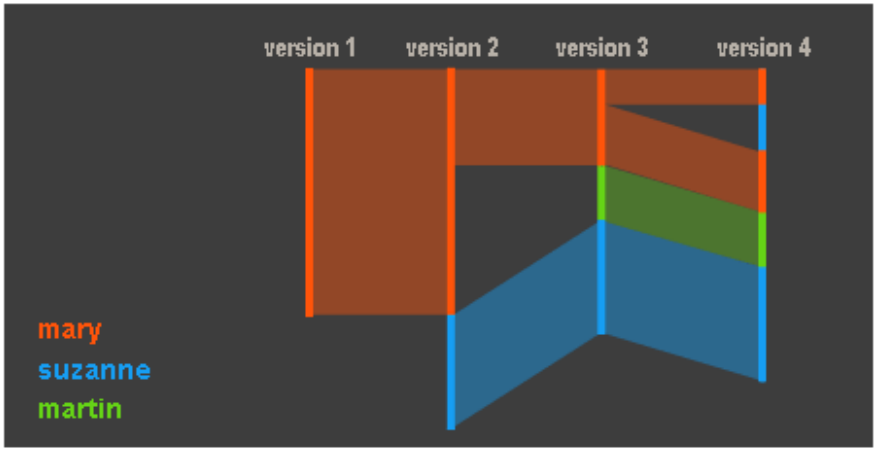
http://researchweb.watson.ibm.com/history/

Technique

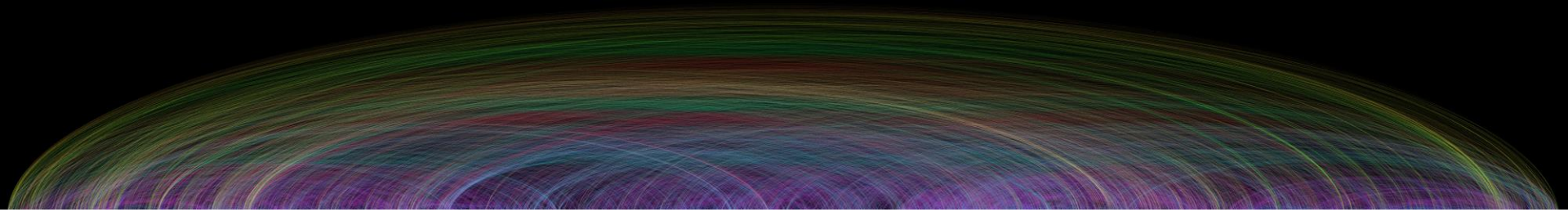


Length – how much text

Time



Make connections



Back to time series …

PERIODIC DATA AND SEASONALITY

Periodic data

- Serial, periodic data
- Data with chronological aspect, but repeats and follows a pattern over time
  - Hinted at in last case study
- How might one visualize that?

Using Spirals (Carlis & Konstan, *UIST* '98)

- Standard x-y timeline or tabular display is problematic for periodic data
  - It has endpoints
- Use spiral to help display data
  - One loop corresponds to one period

Basic Spiral Display

- One year per loop
- Same month on radial bars
- Quantity represented by size of blob

- Is it as easy to see serial data as periodic data?

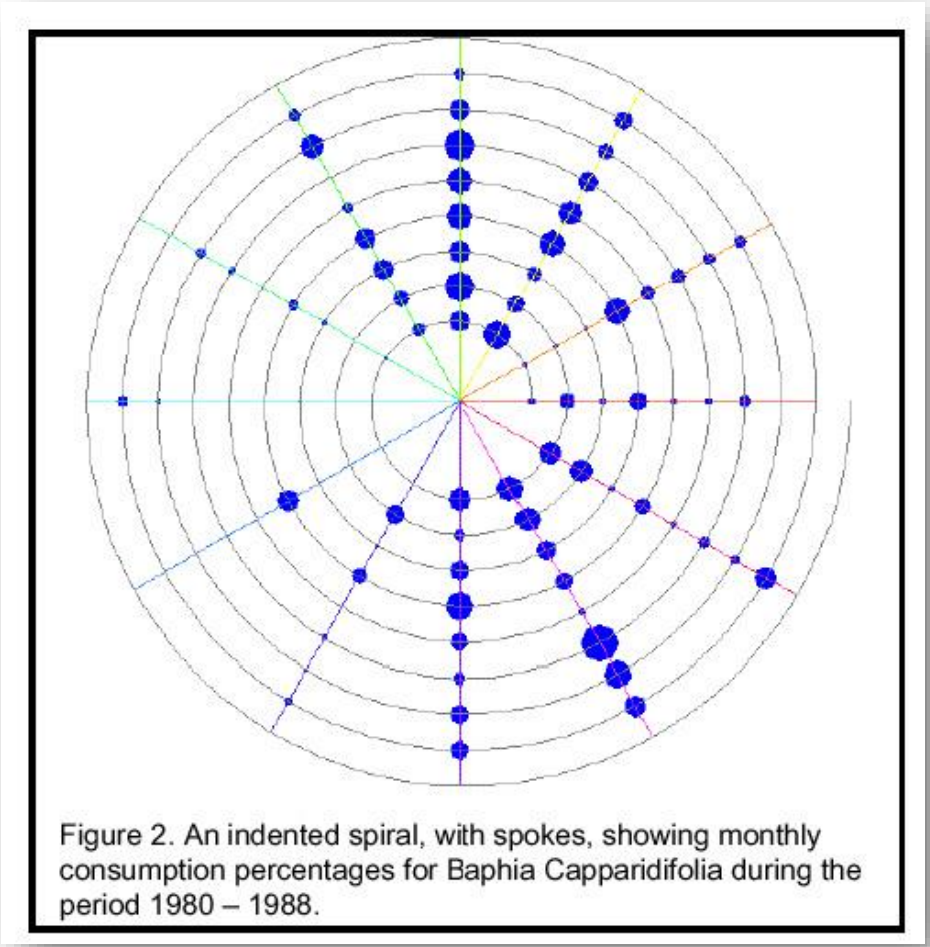


Figure 2. An indented spiral, with spokes, showing monthly consumption percentages for Baphia Capparidifolia during the period 1980 – 1988.

Unknown Periods

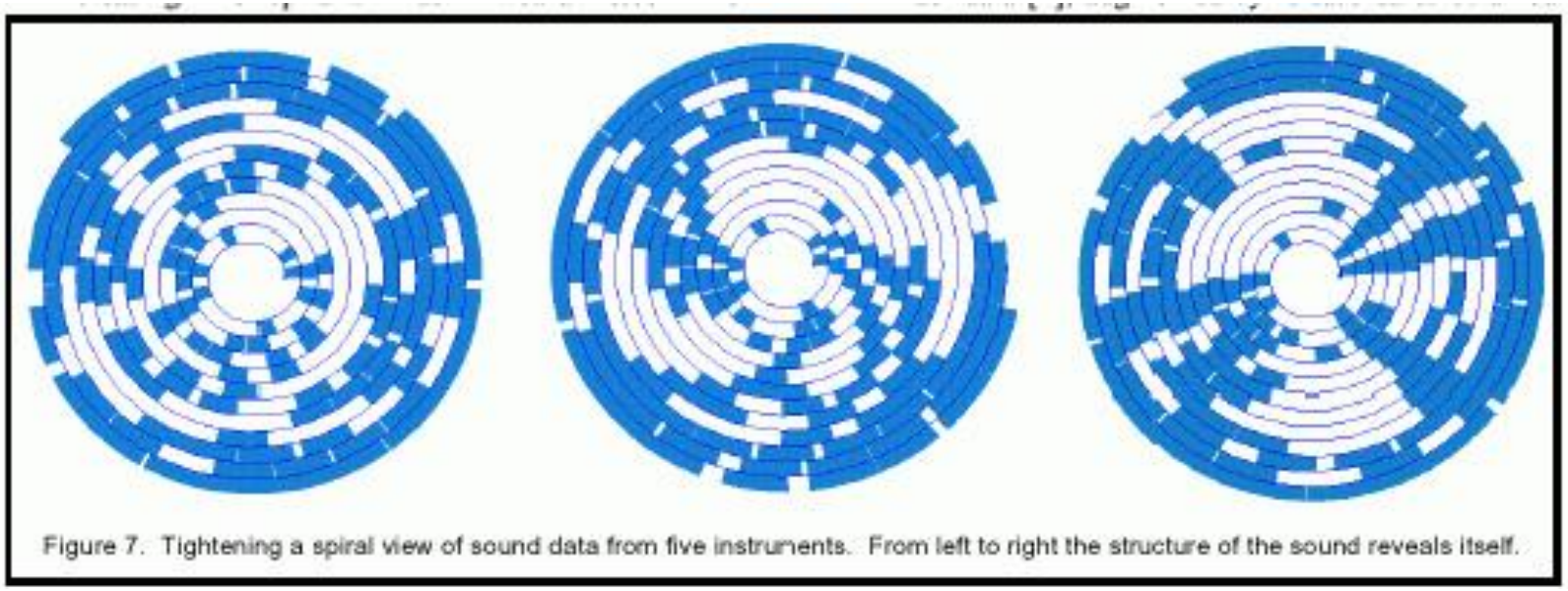
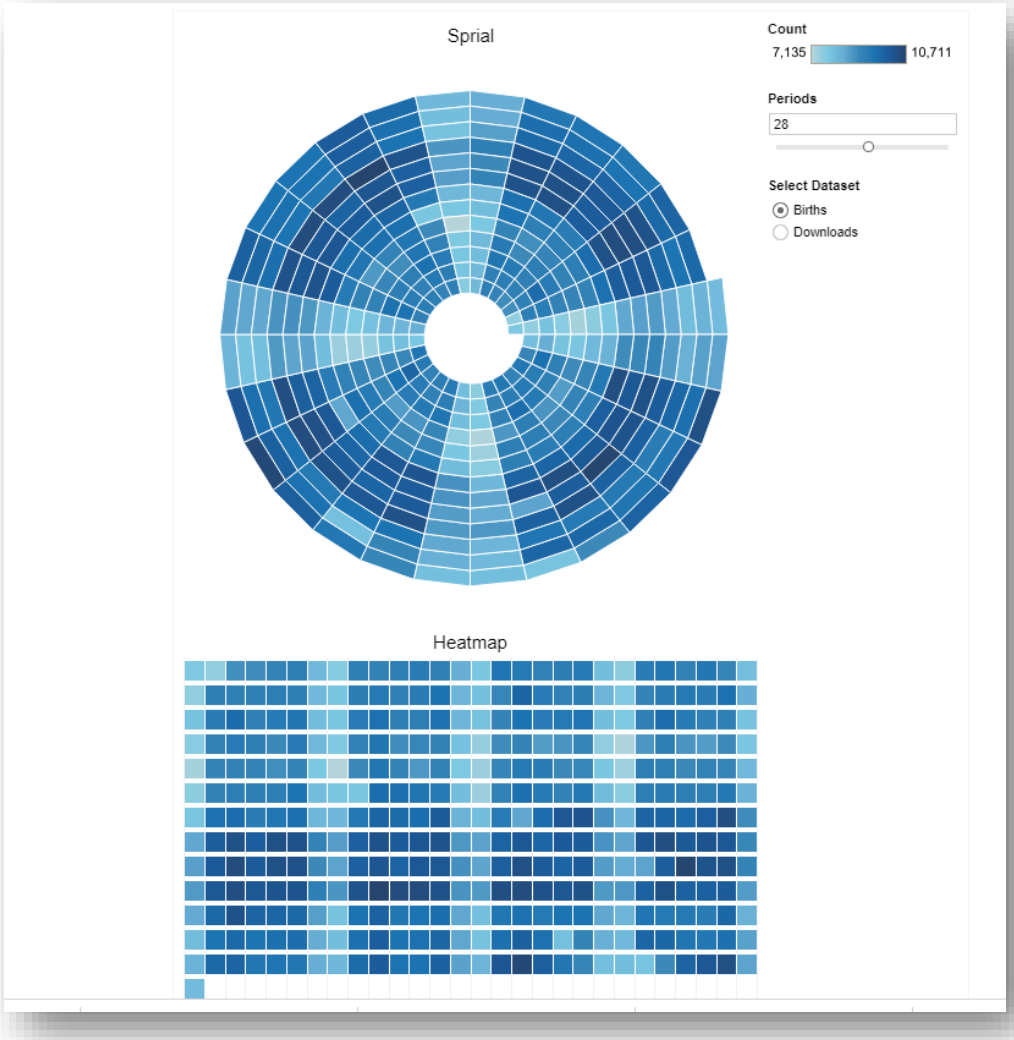


Figure 7. Tightening a spiral view of sound data from five instruments. From left to right the structure of the sound reveals itself.

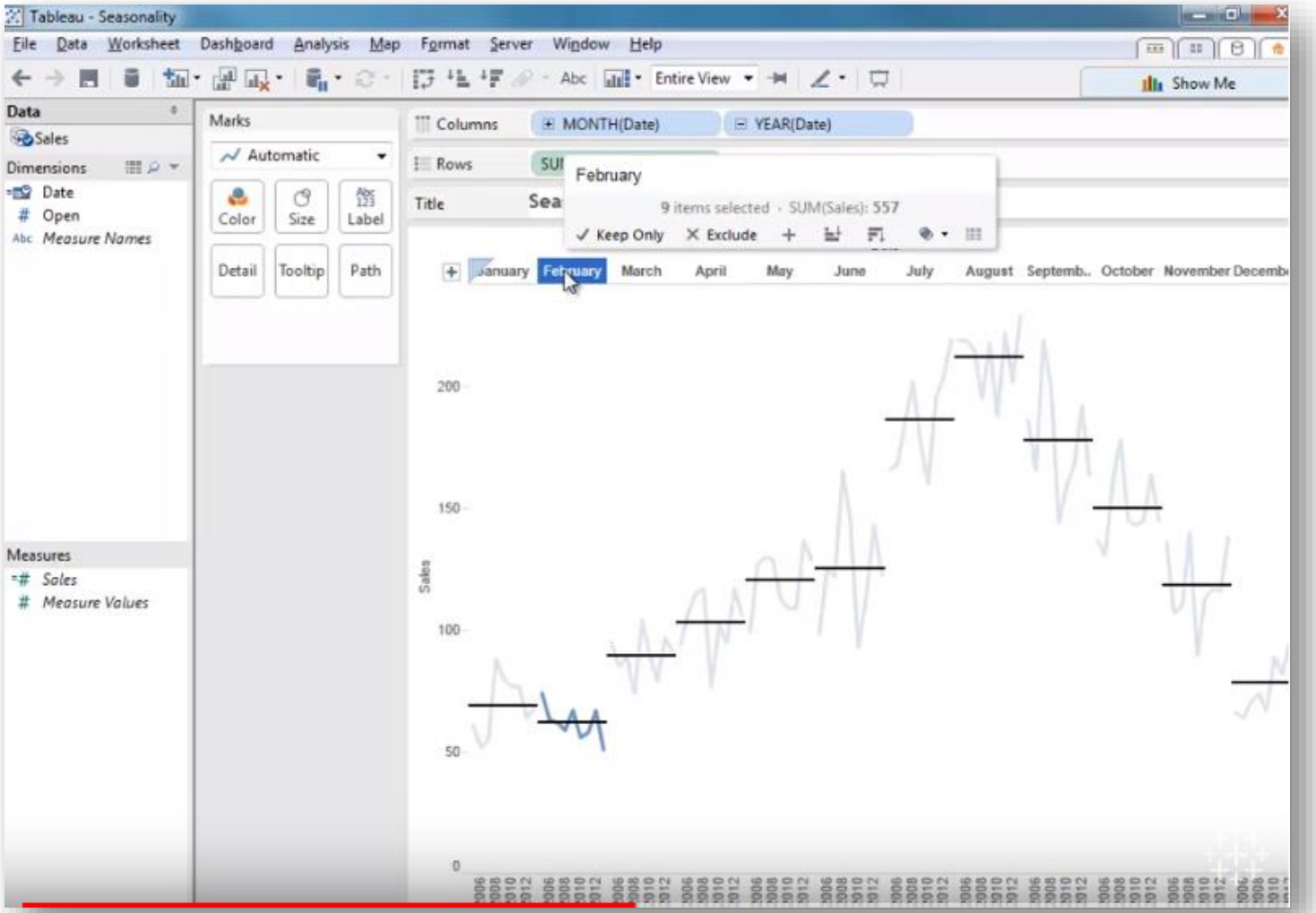
- What if a data set doesn't have a regular temporal period?
- Must do some juggling to align periods

Births in 1978 (Tableau)



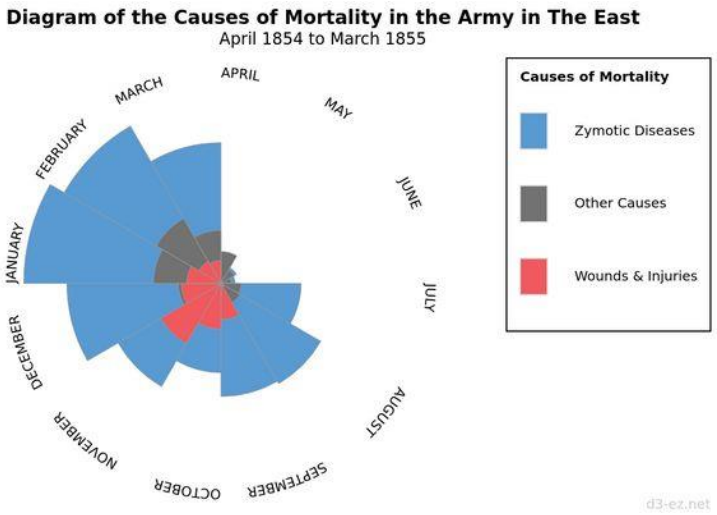
https://public.tableau.com/profile/joe.mako#!/vizhome/SpiralHeatmap/Dashboard

Cycle Plots for Seasonality (Tableau)



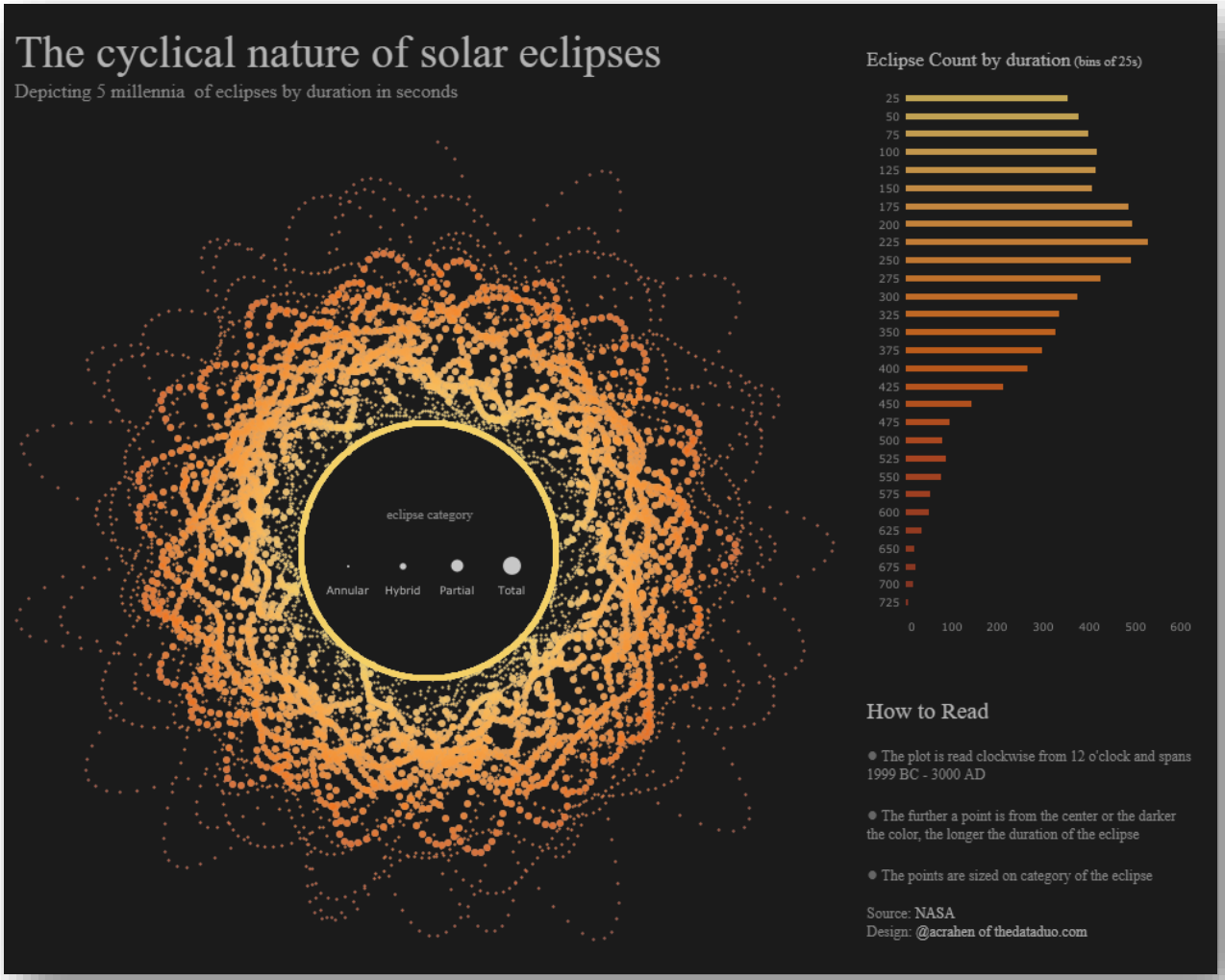
https://www.youtube.com/watch?v=IjeEPBz4puc

Nightingale rose / Polar area diagram

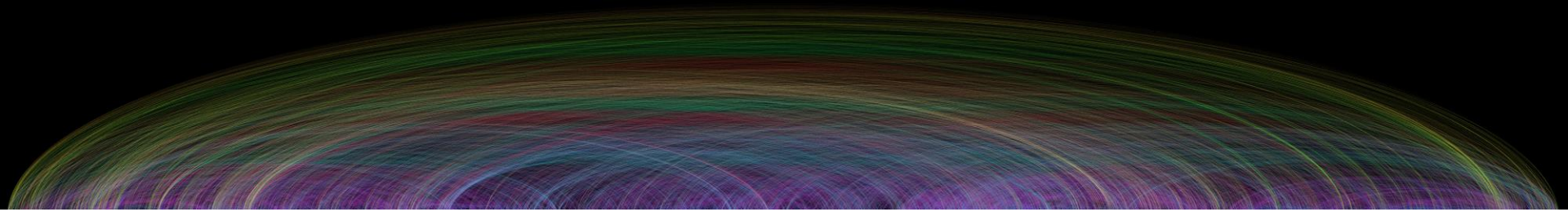


- Traps
  - Positioning of values influences perception of magnitude
  - Size perception of areas (like pie charts!)
  - Small values become lines
- Best practice
  - Choose area of the sectors proportional to the value being represented
  - Limit number of variables to allow for color retention (4)

5 millennia of eclipses



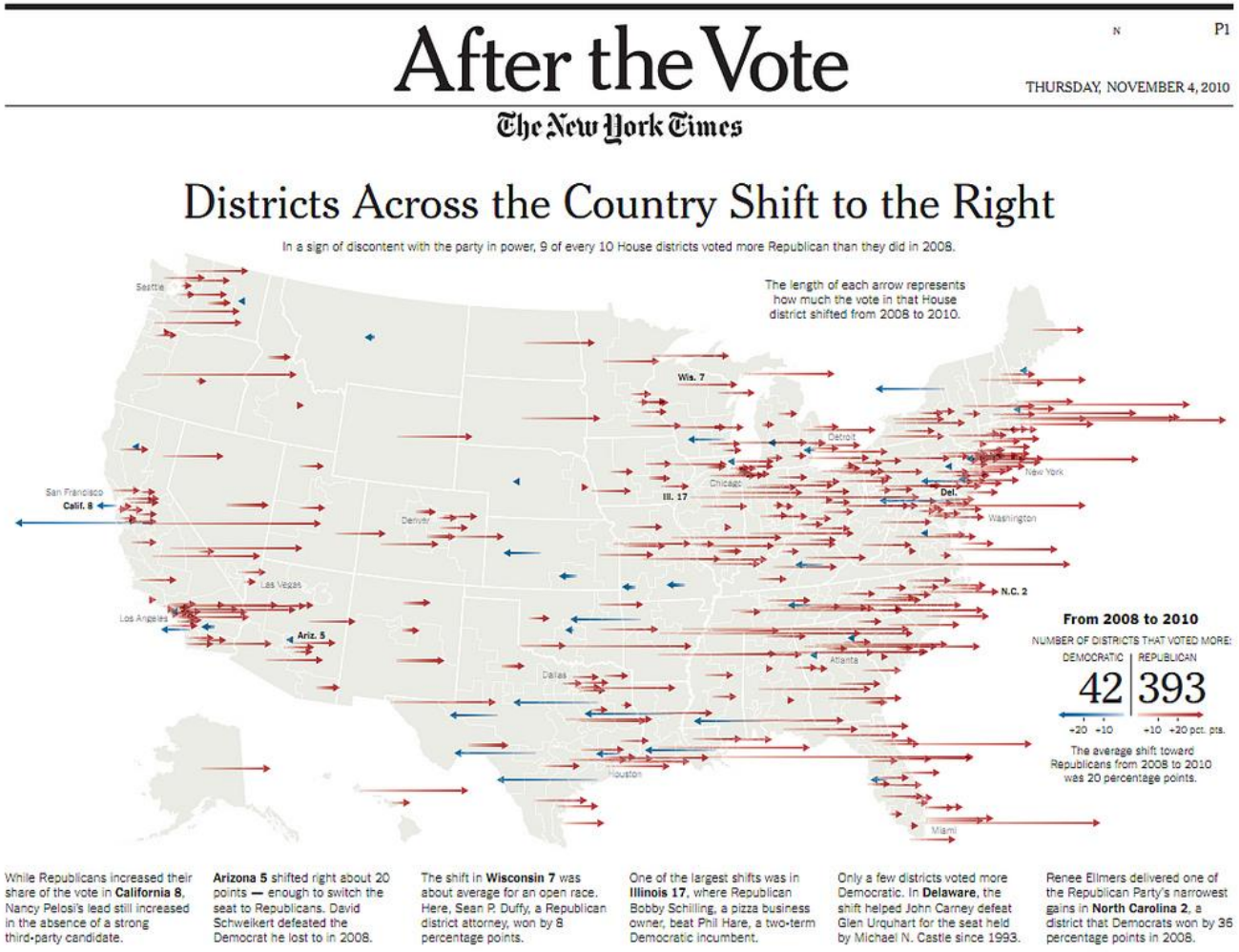
https://public.tableau.com/profile/adam.crahen#!/vizhome/CyclicalNatureofSolarEclipses/Eclipses



Time IS Change!

TIME PATTERNS

Movement / Change of value



After the Vote

N P1

The New York Times

THURSDAY, NOVEMBER 4, 2010

Districts Across the Country Shift to the Right

In a sign of discontent with the party in power, 9 of every 10 House districts voted more Republican than they did in 2008.

The length of each arrow represents how much the vote in that House district shifted from 2008 to 2010.

Seattle
San Francisco
Calif. 8
Las Vegas
Los Angeles
Ariz. 5
Denver
Wis. 7
Chicago
Ill. 17
Detroit
New York
Del.
Washington
N.C. 2
Atlanta
Dallas
Houston
Miami

From 2008 to 2010

NUMBER OF DISTRICTS THAT VOTED MORE:

DEMOCRATIC	REPUBLICAN
42	393
+20 +10	+10 +20 pct. pts.

The average shift toward Republicans from 2008 to 2010 was 20 percentage points.

While Republicans increased their share of the vote in **California 8**, Nancy Pelosi's lead still increased in the absence of a strong third-party candidate.

Arizona 5 shifted right about 20 points — enough to switch the seat to Republicans. David Schweikert defeated the Democrat he lost to in 2008.

The shift in **Wisconsin 7** was about average for an open race. Here, Sean P. Duffy, a Republican district attorney, won by 8 percentage points.

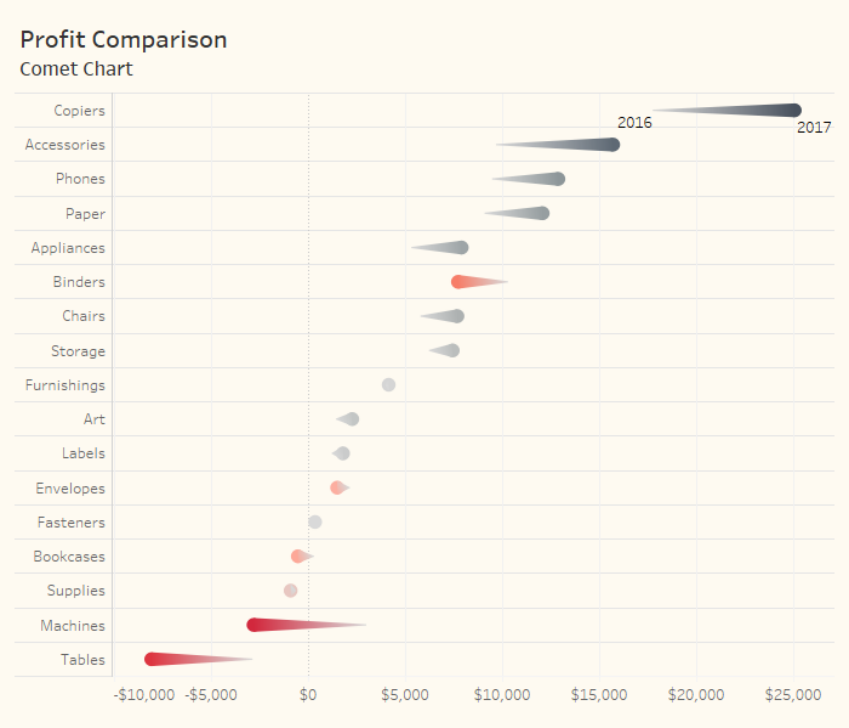
One of the largest shifts was in **Illinois 17**, where Republican Bobby Schilling, a pizza business owner, beat Phil Hare, a two-term Democratic incumbent.

Only a few districts voted more Democratic. In **Delaware**, the shift helped John Carney defeat Glen Urquhart for the seat held by Michael N. Castle since 1993.

Renee Ellmers delivered one of the Republican Party's narrowest gains in **North Carolina 2**, a district that Democrats won by 36 percentage points in 2008.

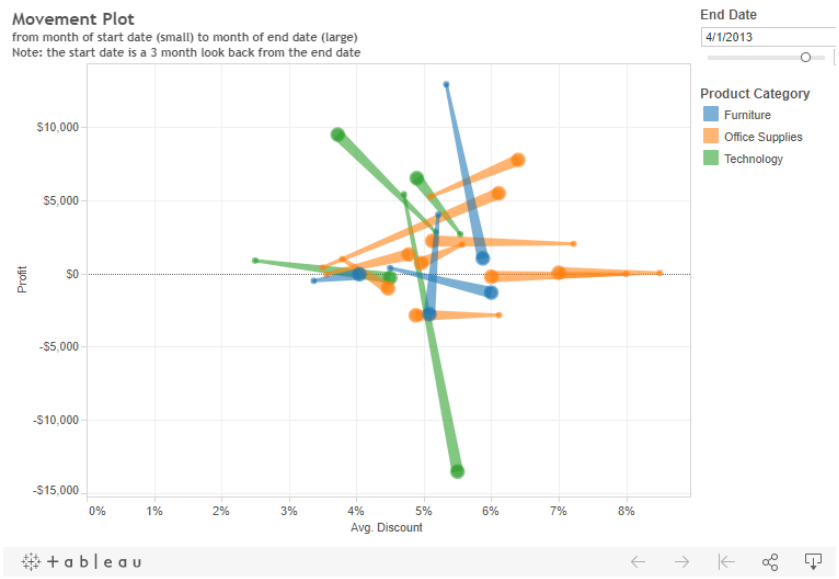
Movement / Change of value

Comet chart



https://www.datarevelations.com/showing-now-versus-then-consider-a-comet-chart.html

Movement plot



https://interworks.com/blog/pmathewson/2014/11/10/tableau-viz-scatter-plot-twist-movement-plot/

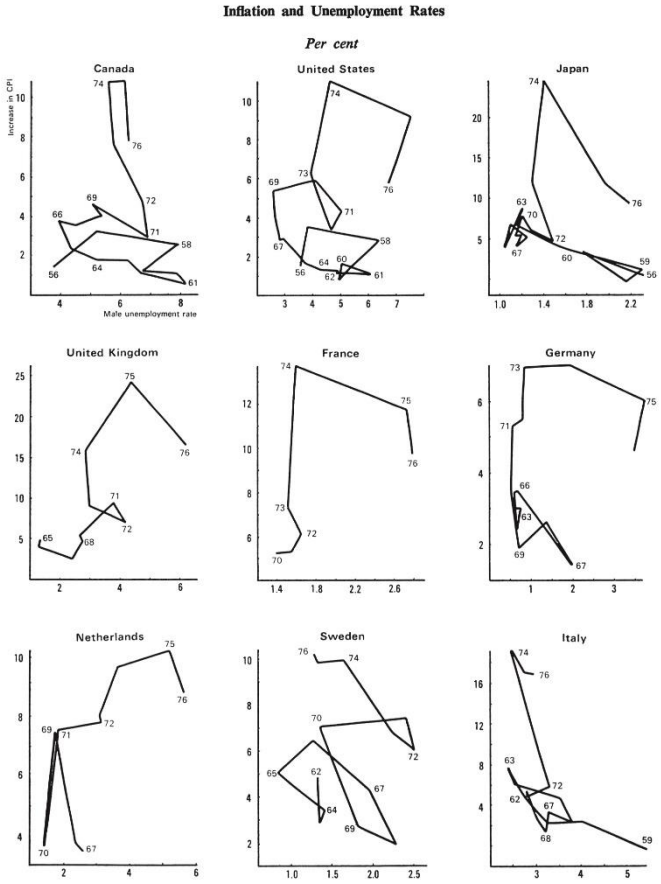
Traces


Sales Forecast
Consumer
Corporate
Home Office
$60,000
$40,000
$20,000
$0
2014
2015
2016
2017
2018
2019
Order Date

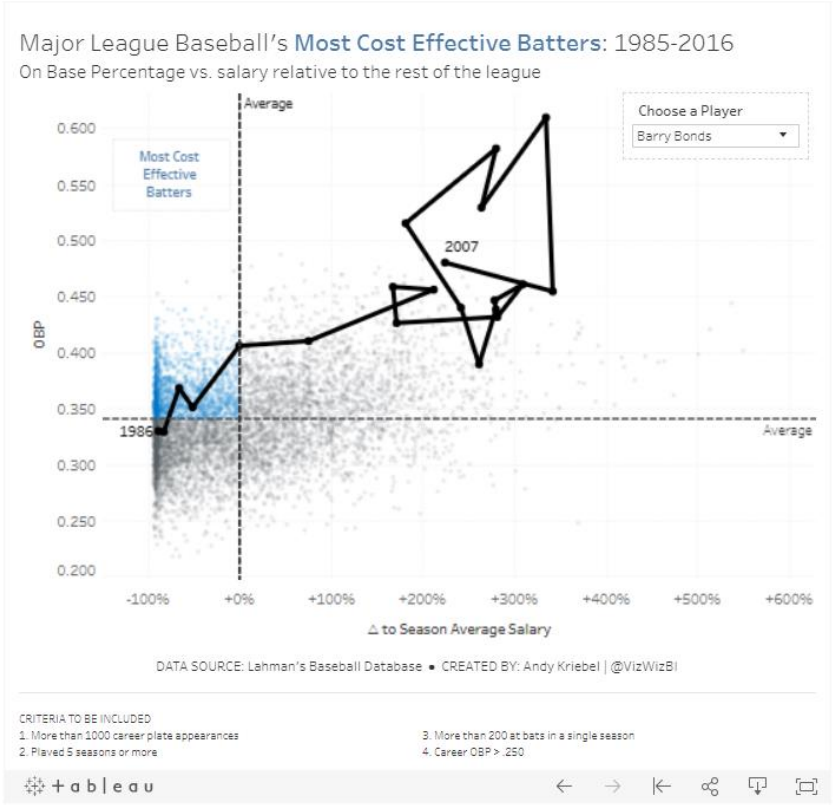
Tufte
Time Paths

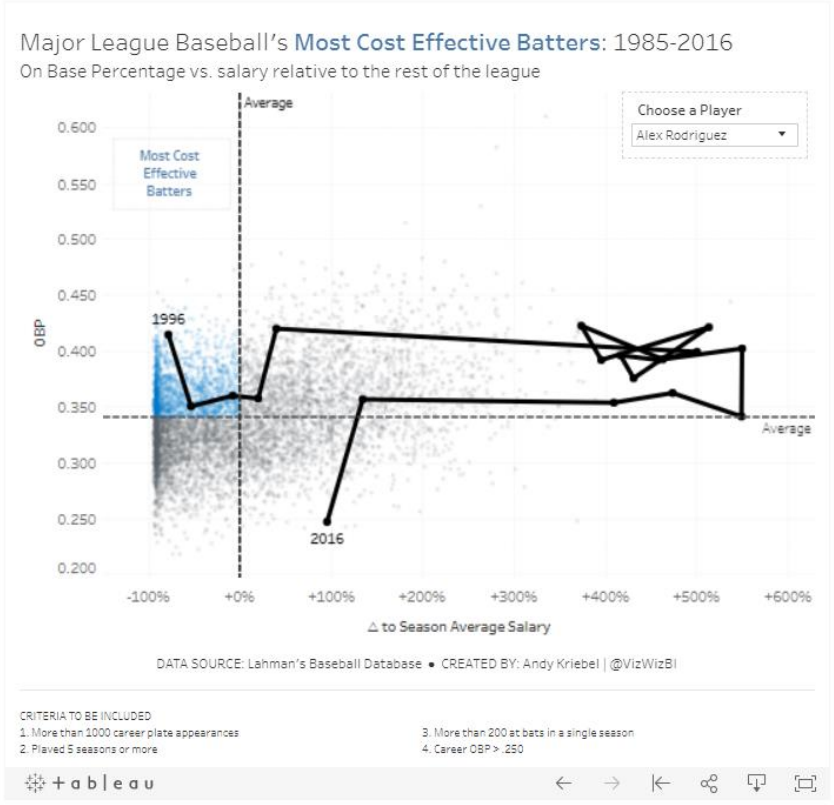
These small-multiple relational graphs show unemployment and inflation over time in "Phillips curve" plots for nine countries, demonstrating the collapse of what was once thought to be an inverse relationship between the variables.

Paul McCracken, et al., *Towards Full Employment and Price Stability* (Paris, 1977), 106.



MLB Batter Effectiveness

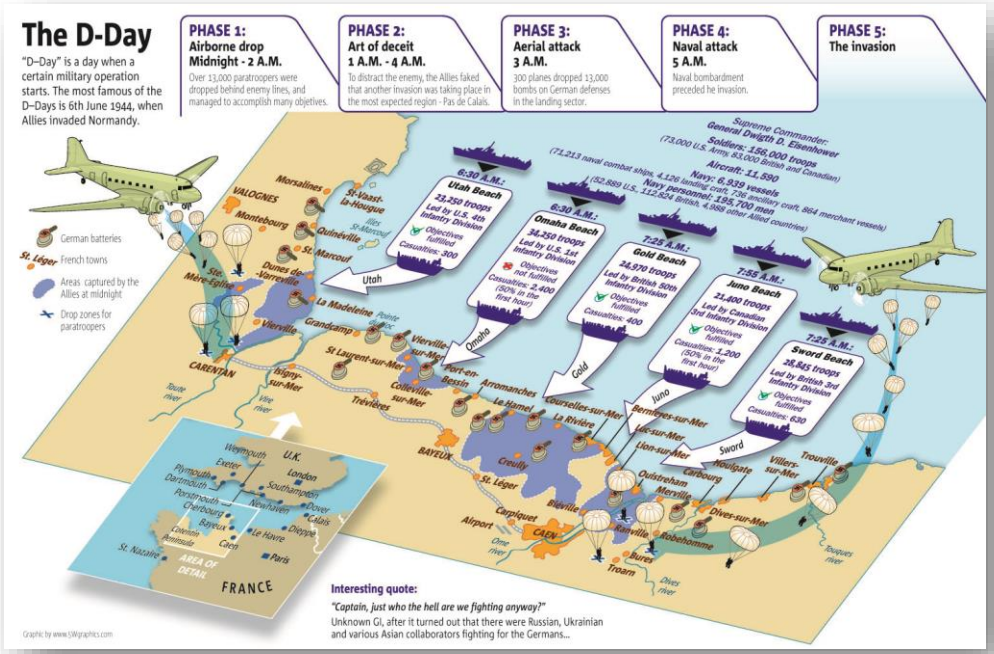


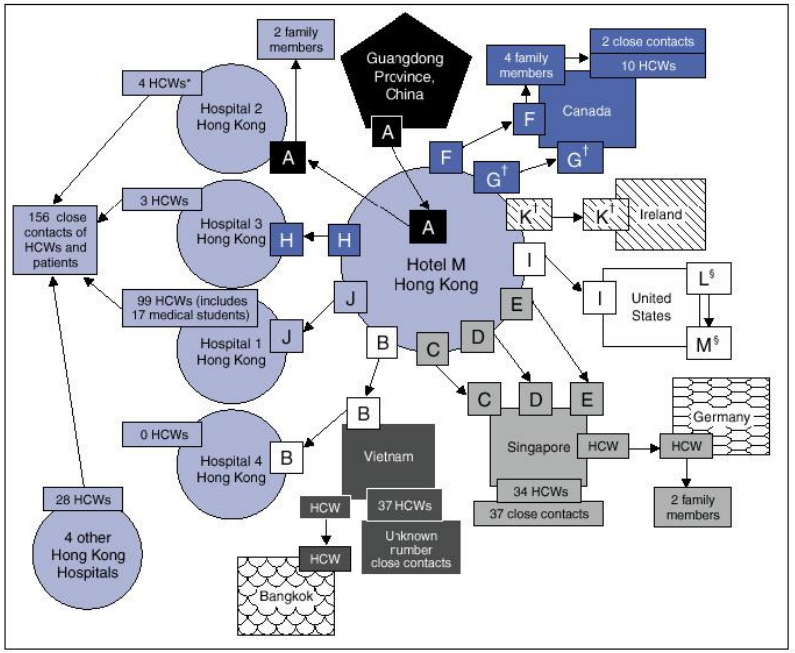


https://public.tableau.com/profile/andy.kriebel#!/vizhome/MLBBattingStats/MLB

Stories in time and place

- How about events in time and place?
  - Many applications of this problem

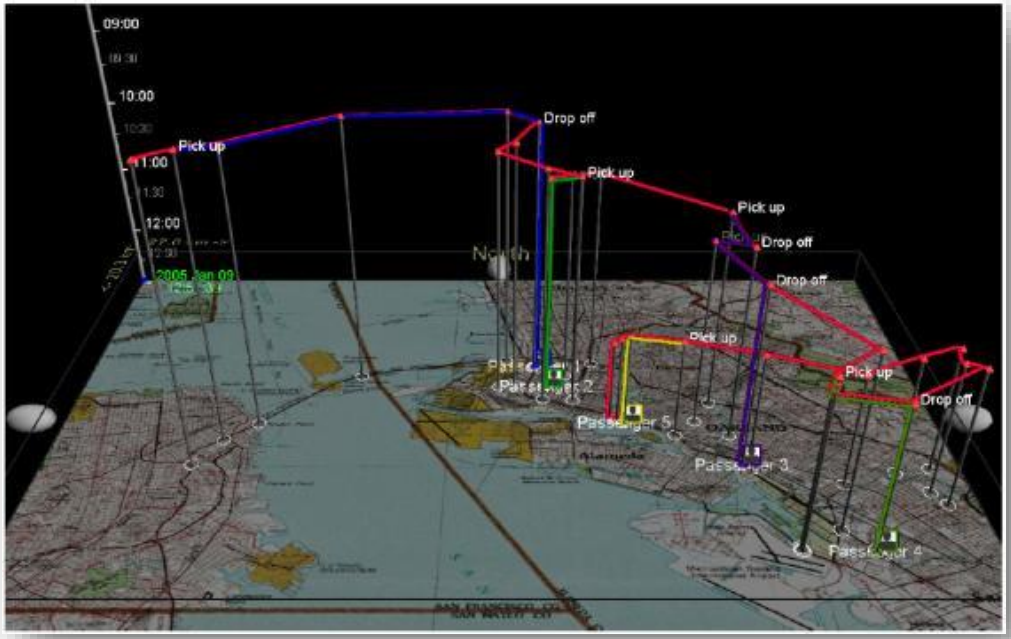


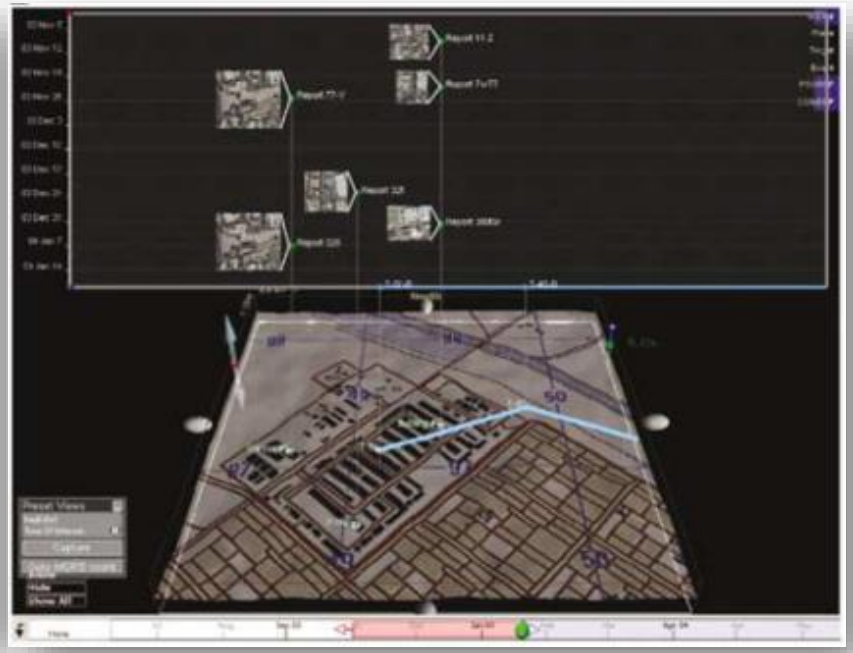


GeoTime (Kapler & Wright, *InfoVis* '04)

- Represent place by 2D plane (or maybe 3D topography)
- Use 3rd dimension to encode time
- Object types:
  - Entities (people or things)
  - Locations (geospatial or conceptual)
  - Events (occurrences or discovered facts)

Overview

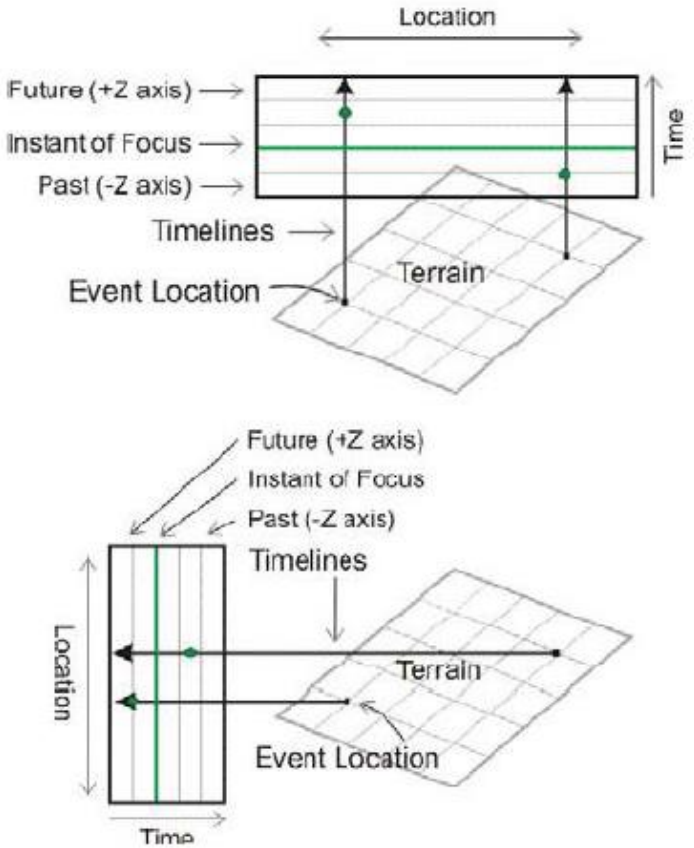




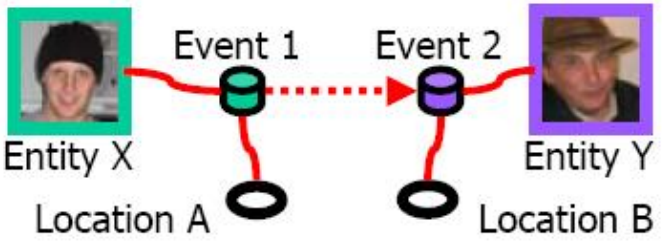
- Objective: visualize spatial interconnectedness of information over time and geography with interactive 3-D view

Source: http://www.oculusinfo.com/

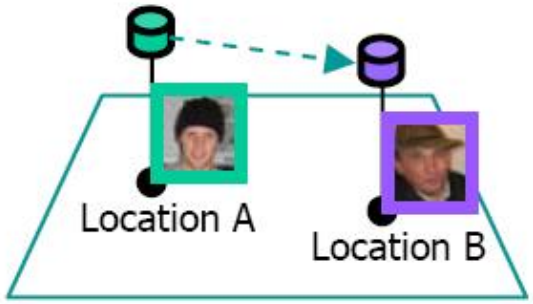
Design Characteristics



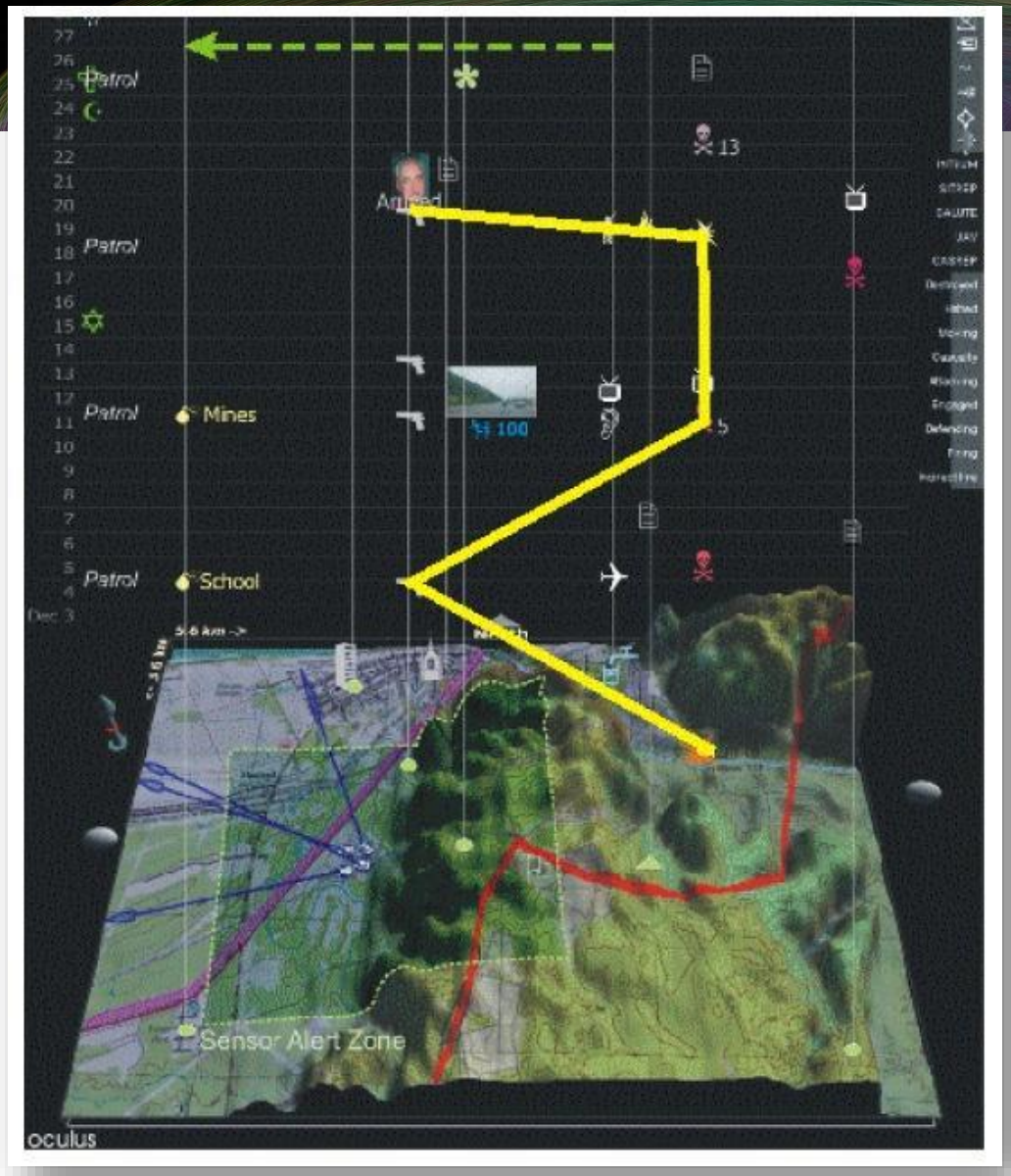
Dimension usage



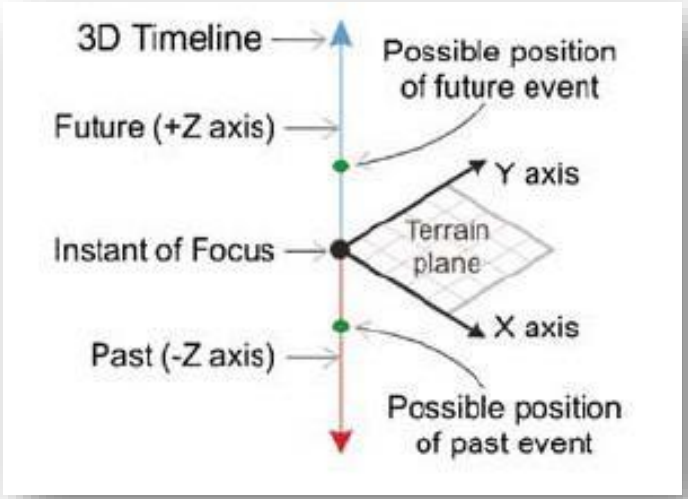
Vector Group with 1+ Actors
(Phone Call, email, money transfer)
5 Associations: Same as above plus.
- *Entity X present at Event 1*
- *Entity Y present at Event 2*

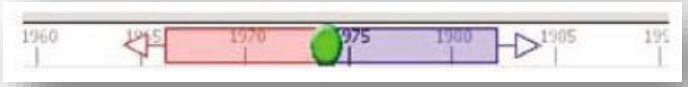


View objects



- Spatial timelines
  - 3-D Z-axis
  - 3-D viewer facing
  - Linked time chart





Sample View

Figure 6: Screenshot of GeoTime with time slider at bottom and moveable time scale at right. The green line traces one entity's movement in time and geography.

Move Time Forward

Video

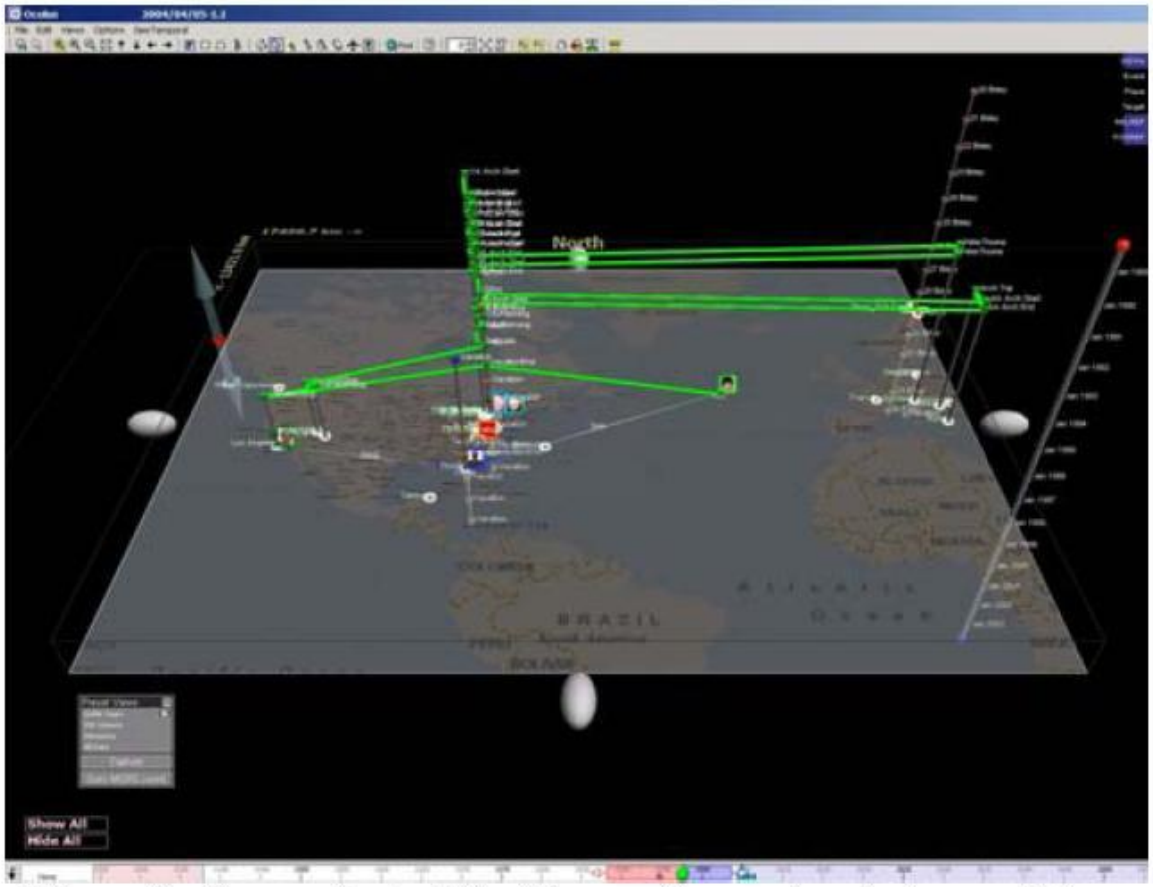


Figure 7: Screenshot of GeoTime with overhead view and time slider advanced forward in time from Figure 6.