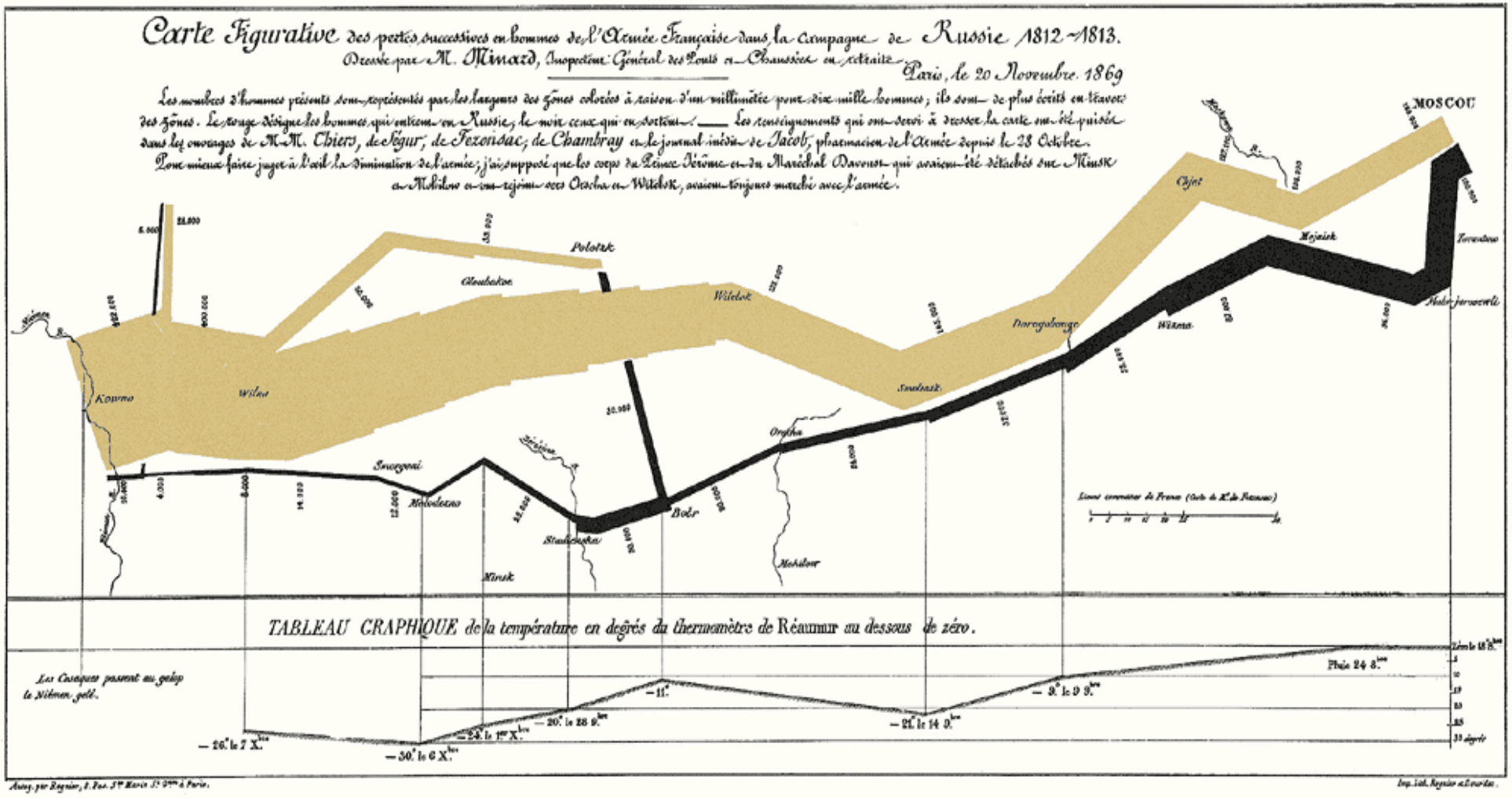
Data Visualisation
Lecture 3 – Trends Over Time

Dr. Cathy Ennis

Learning Outcomes Week 3

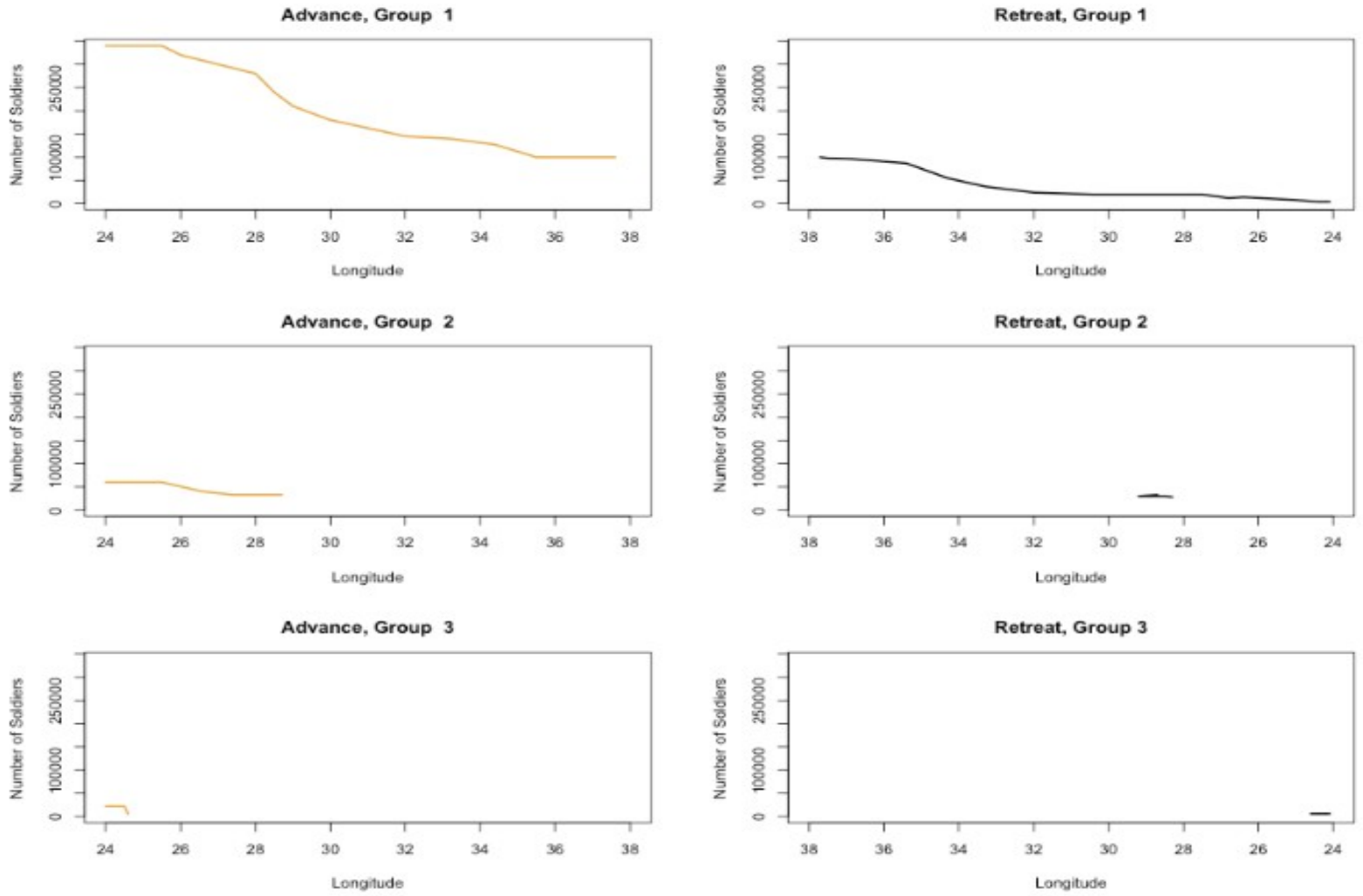
- Design effective visualizations based on principles from perceptual psychology, cognitive science, graphic design and visual art
- Create and deploy successful data visualisations using leading software tools
- Demonstrate an understanding how visualisation is used in date journalism to communicate complex ideas and stories

Visualisation of the week



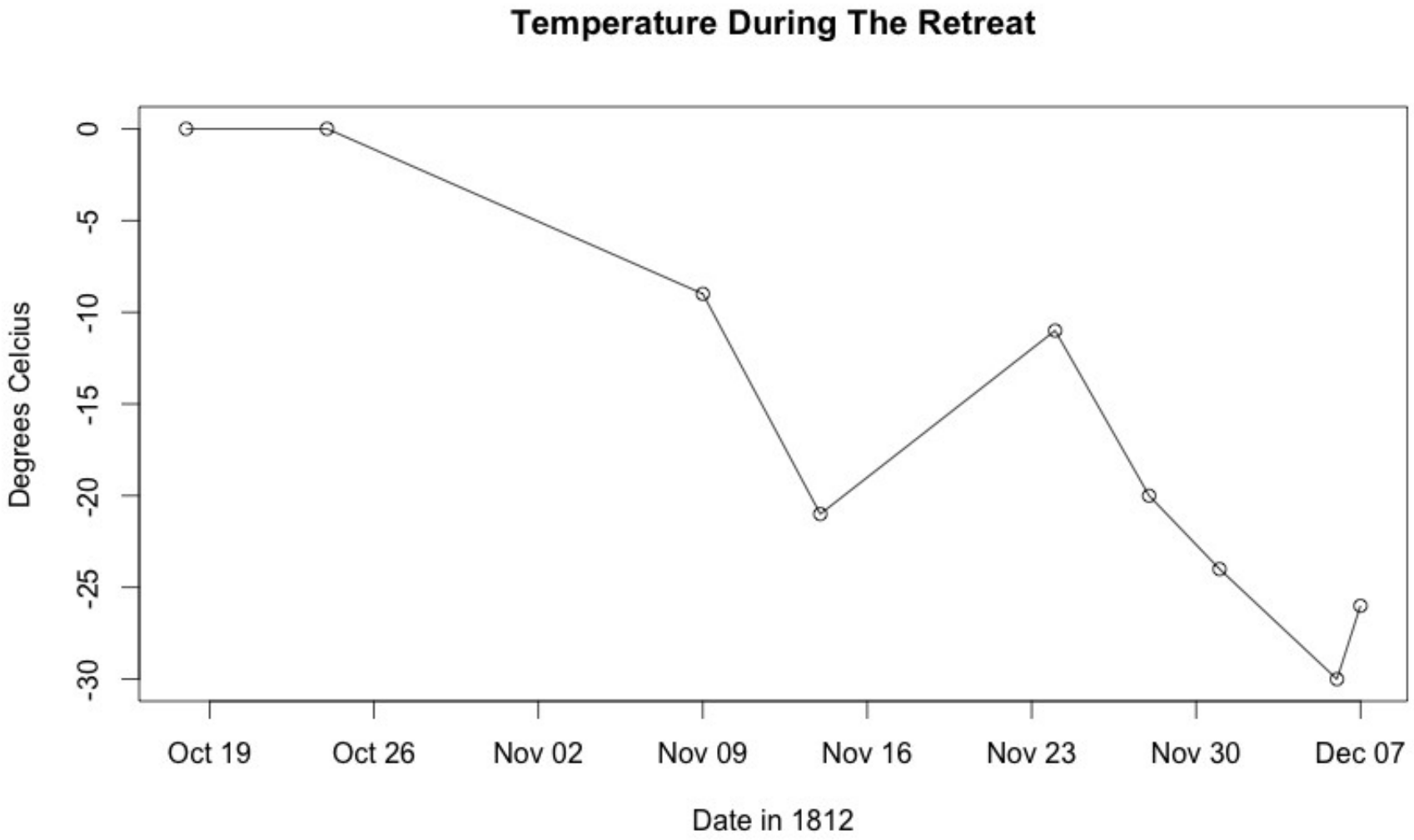
https://www.edwardtufte.com/tufte/minard

Charles Joseph Minard



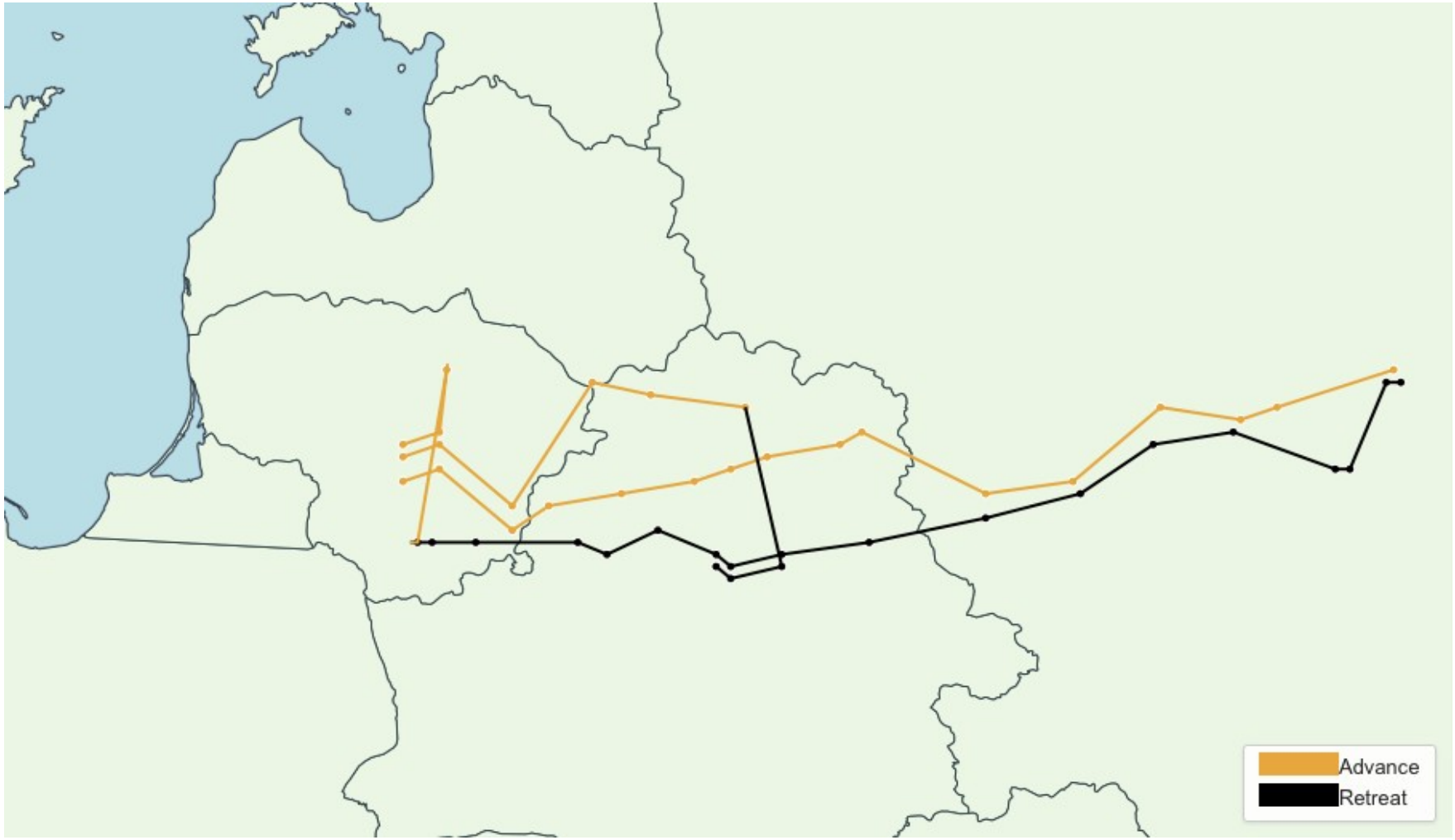
https://robots.thoughtbot.com/analyzing-minards-visualization-of-napoleons-1812-march

Charles Joseph Minard



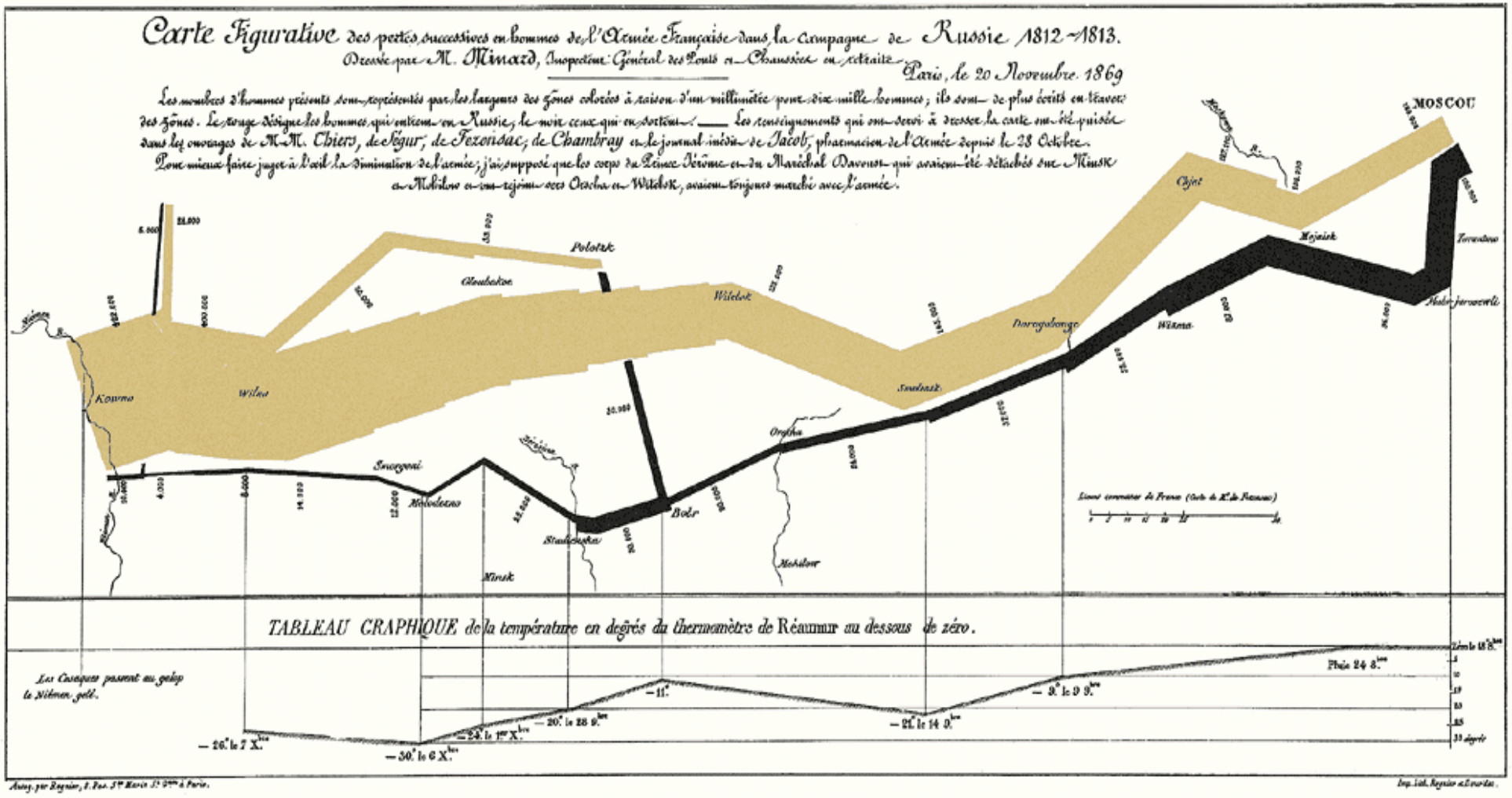
https://robots.thoughtbot.com/analyzing-minards-visualization-of-napoleons-1812-march

Charles Joseph Minard



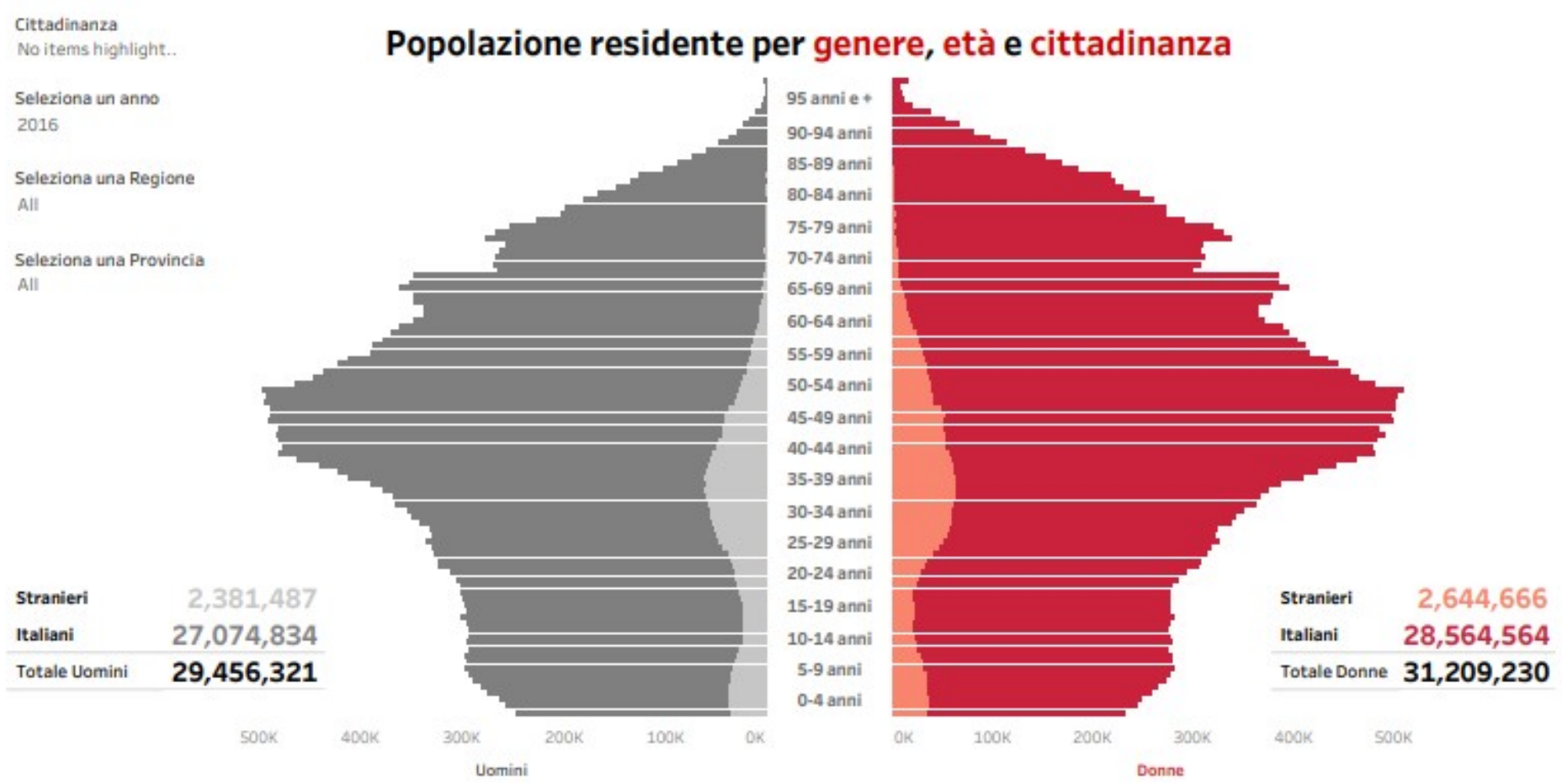
https://robots.thoughtbot.com/analyzing-minards-visualization-of-napoleons-1812-march

Visualisation of the week



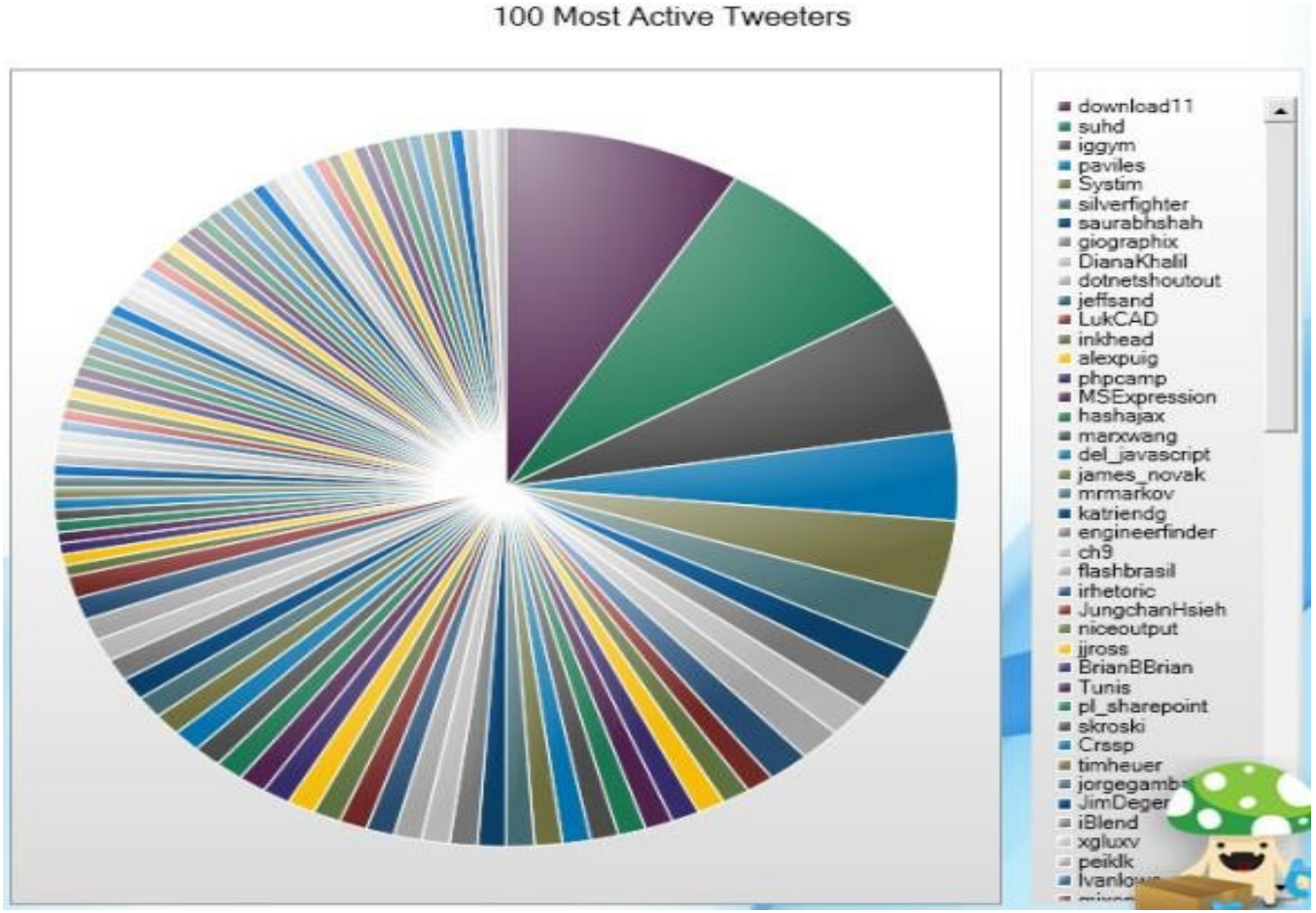
https://public.tableau.com/profile/john.emery#!/vizhome/NapoleonsMarch/FrenchInvasion-French

Visualisation of the week



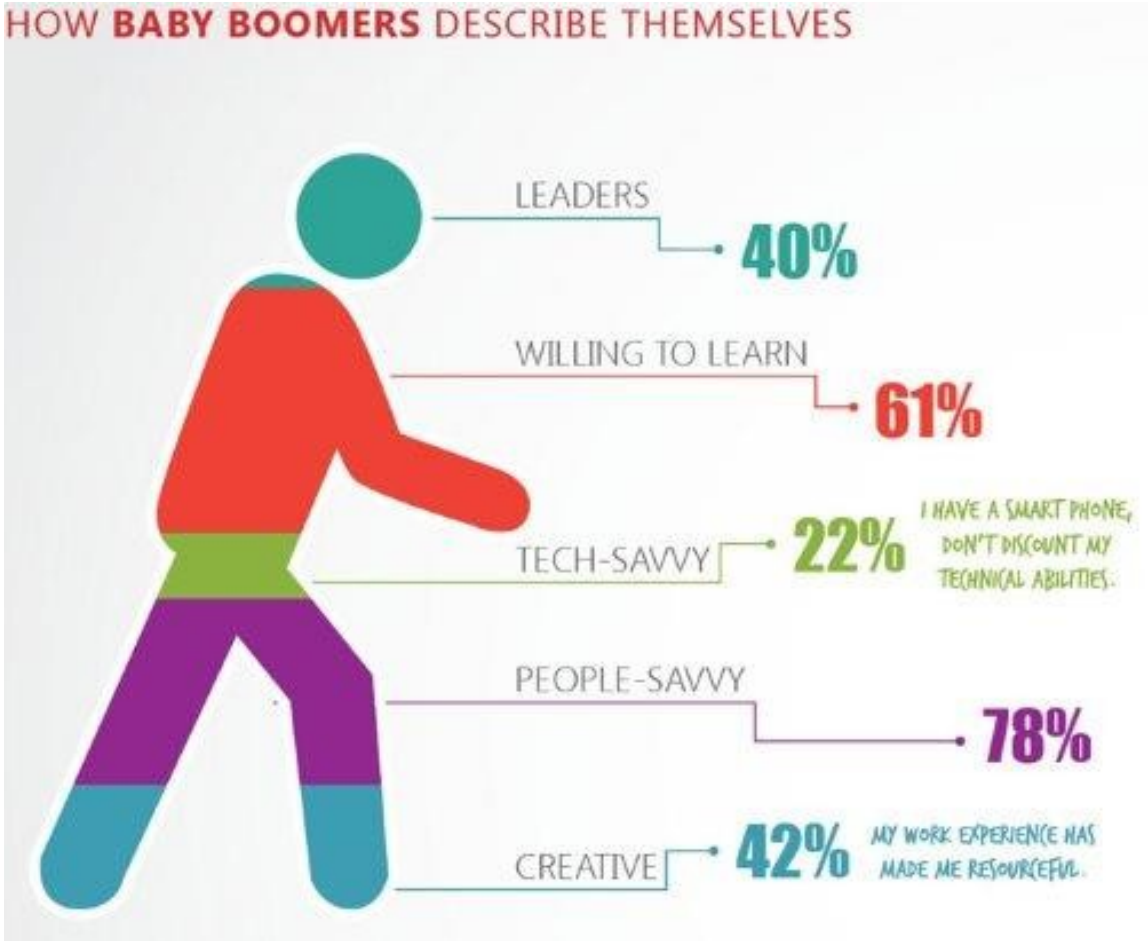
Fonte: elaborazioni Staff EQuIPE 2020 su dati ISTAT (Popolazione residente e Popolazione straniera residente 1 gennaio, period 2003-2016)

!Visualisation of the week



http://livingqlikview.com/the-9-worst-data-visualizations-ever-created/

!Visualisation of the week



http://livingqlikview.com/the-9-worst-data-visualizations-ever-created/

Overview

- Visualising patterns over time
- Discrete points in time
- Continuous points in time
- Multiple dimensions over time

ACCENT for Effective Graphical Displays

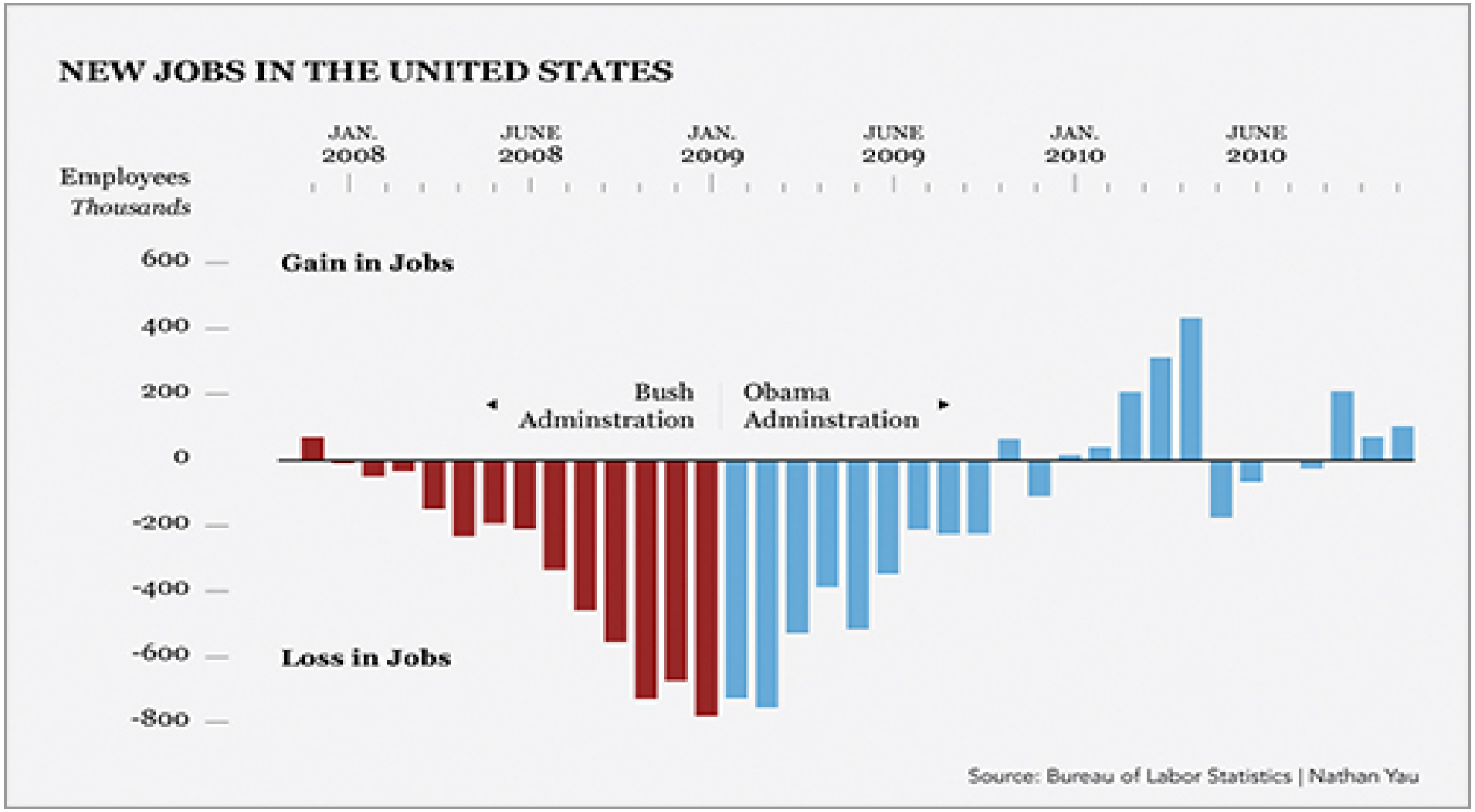
- **A**pprehension: Can you perceive the relationship among the variables?
- **C**larity: Can you visually distinguish all of the elements of the graph?
- **C**onsistency: In a series of graphs, are the meanings of elements, shapes and symbols the same?
- **E**fficiency: Do you portray a complex relationship in the simplest way possible?
- **N**ecessity: Do you need the graph, or will another format (text, table) be more effective?
- **T**ruthfulness: Are the elements accurately positioned and scaled?

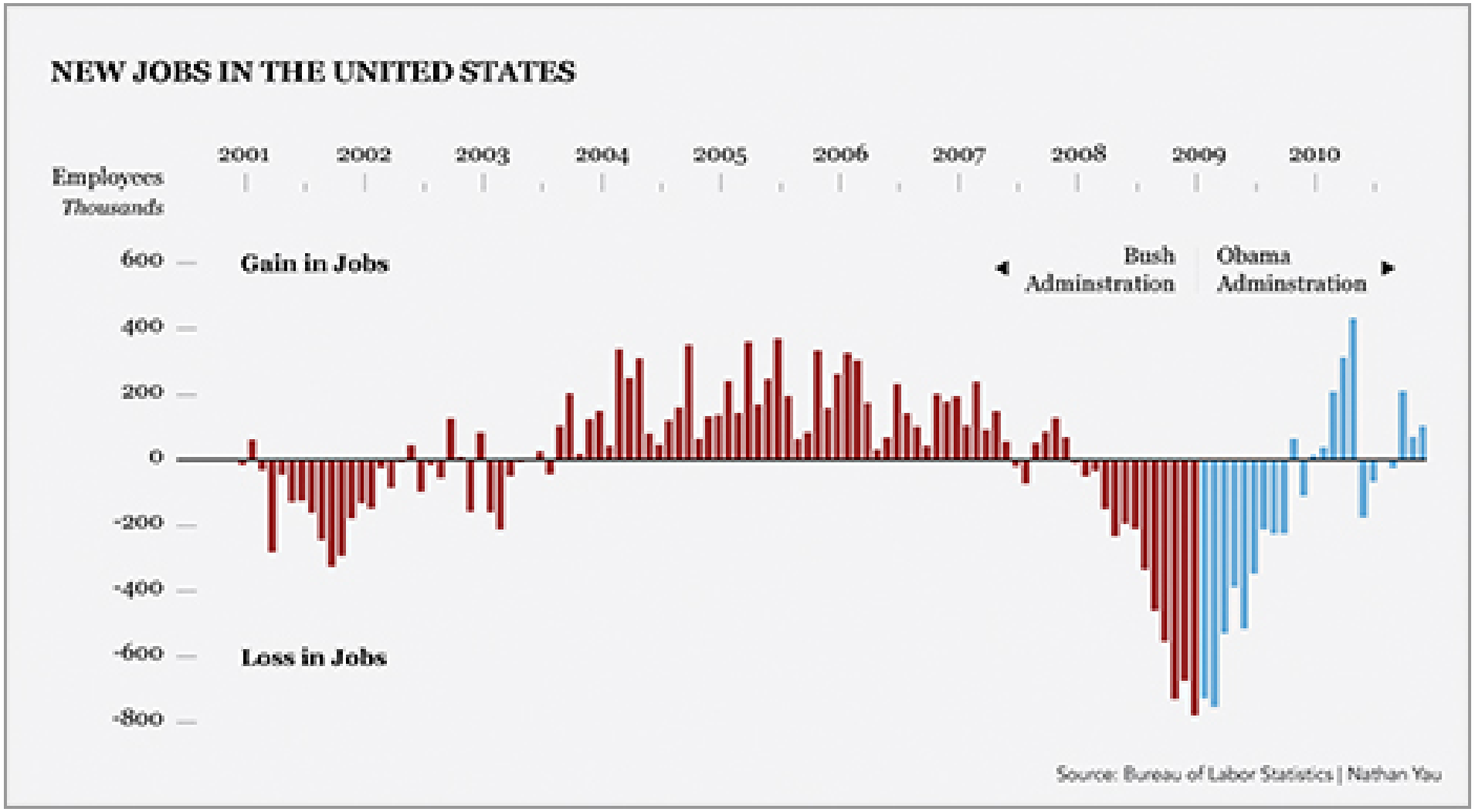
D. A. Burn (1993). Designing Effective Statistical Graphs, in Rao (Ed.) Handbook of Statistics

Trends over Time - What To Look For

- The most common thing you look for in time series, or temporal, data is trends
  - Is something increasing or decreasing?
  - Are there seasonal cycles?

- To find these patterns, you have to look beyond individual data points to get the whole picture


NEW JOBS IN THE UNITED STATES
JAN. 2008
JUNE 2008
JAN. 2009
JUNE 2009
JAN. 2010
JUNE 2010
Employees
Thousands
600
400
200
0
-200
-400
-600
-800
Gain in Jobs
Loss in Jobs
Bush Adminstration
Obama Adminstration
Source: Bureau of Labor Statistics | Nathan Yau


NEW JOBS IN THE UNITED STATES
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
Employees
Thousands
600
400
200
0
-200
-400
-600
-800
Gain in Jobs
Loss in Jobs
Bush Adminstration
Obama Adminstration
Source: Bureau of Labor Statistics | Nathan Yau

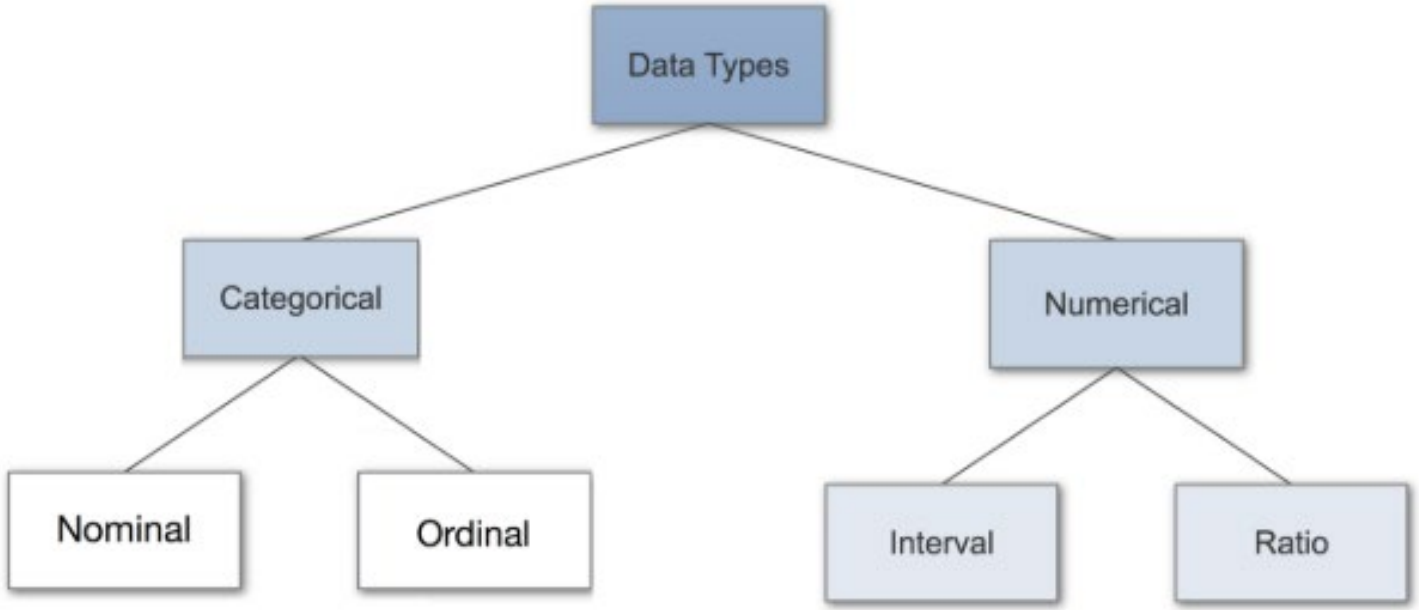
Visualising Different Data Types

- Textual data (qualitative)

- Nominal (qualitative)
- Ordinal (qualitative)

- Interval (quantitative)
- Ratio (quantitative)

- Temporal



Visualising Different Data Types

- Discrete values

- Continuous values

Visualising Patterns Over Time

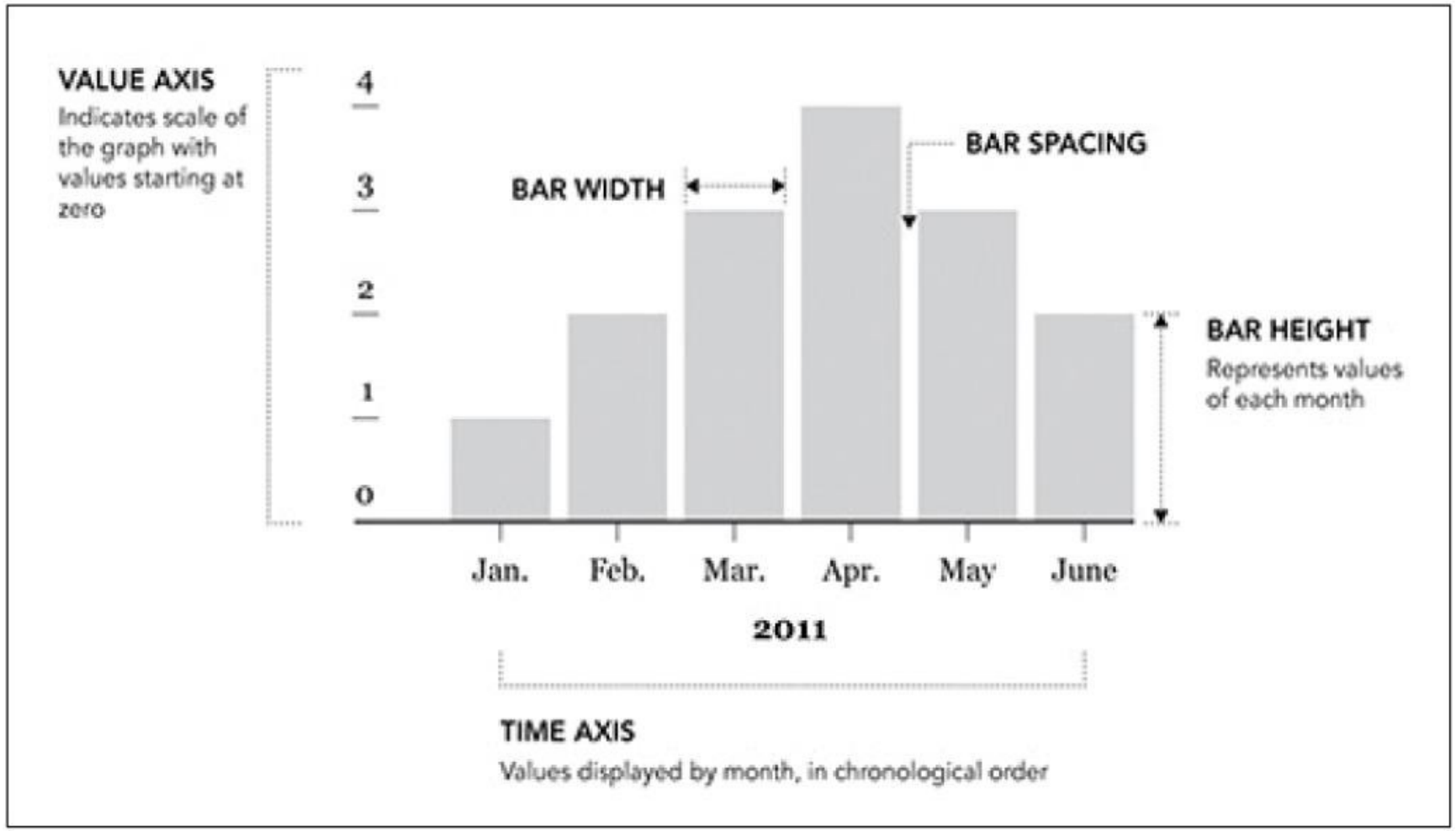
- Discrete points in time
  - Bar Graph /Scatter Plot
- Continuous points in time
  - Line Chart / Step Chart
- Handling multiple dimensions over time
  - Stacked Bar Chart / Area Chart / Stream Graph Small Multiples / Animation

Discrete Points In Time

- Often we have datasets that contain single measurements for a reasonably small number of discrete points in time
  - Profit per year
  - Rainfall per month

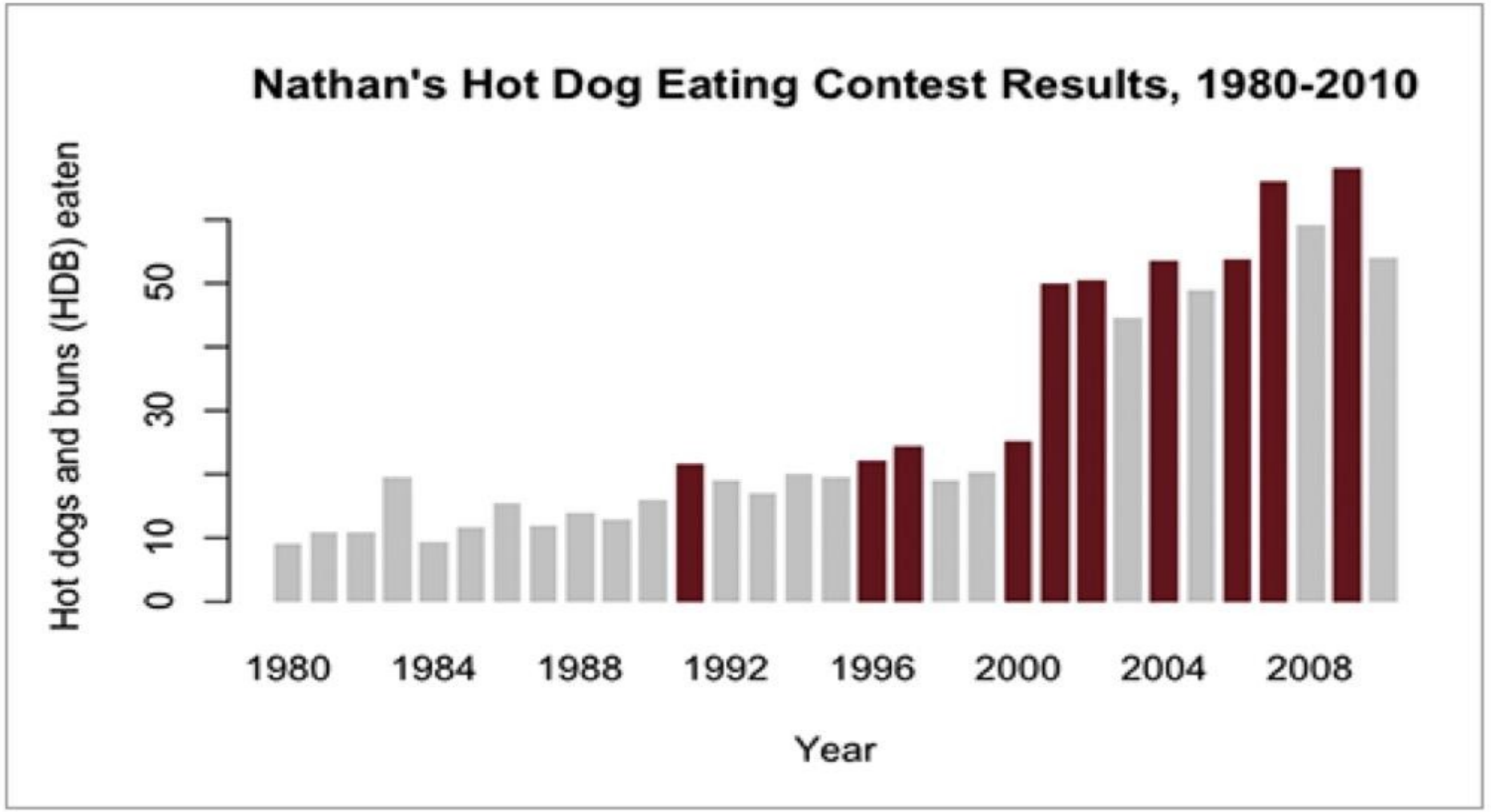
- In these cases a simple bar graph is often appropriate

Simple Bar Graph



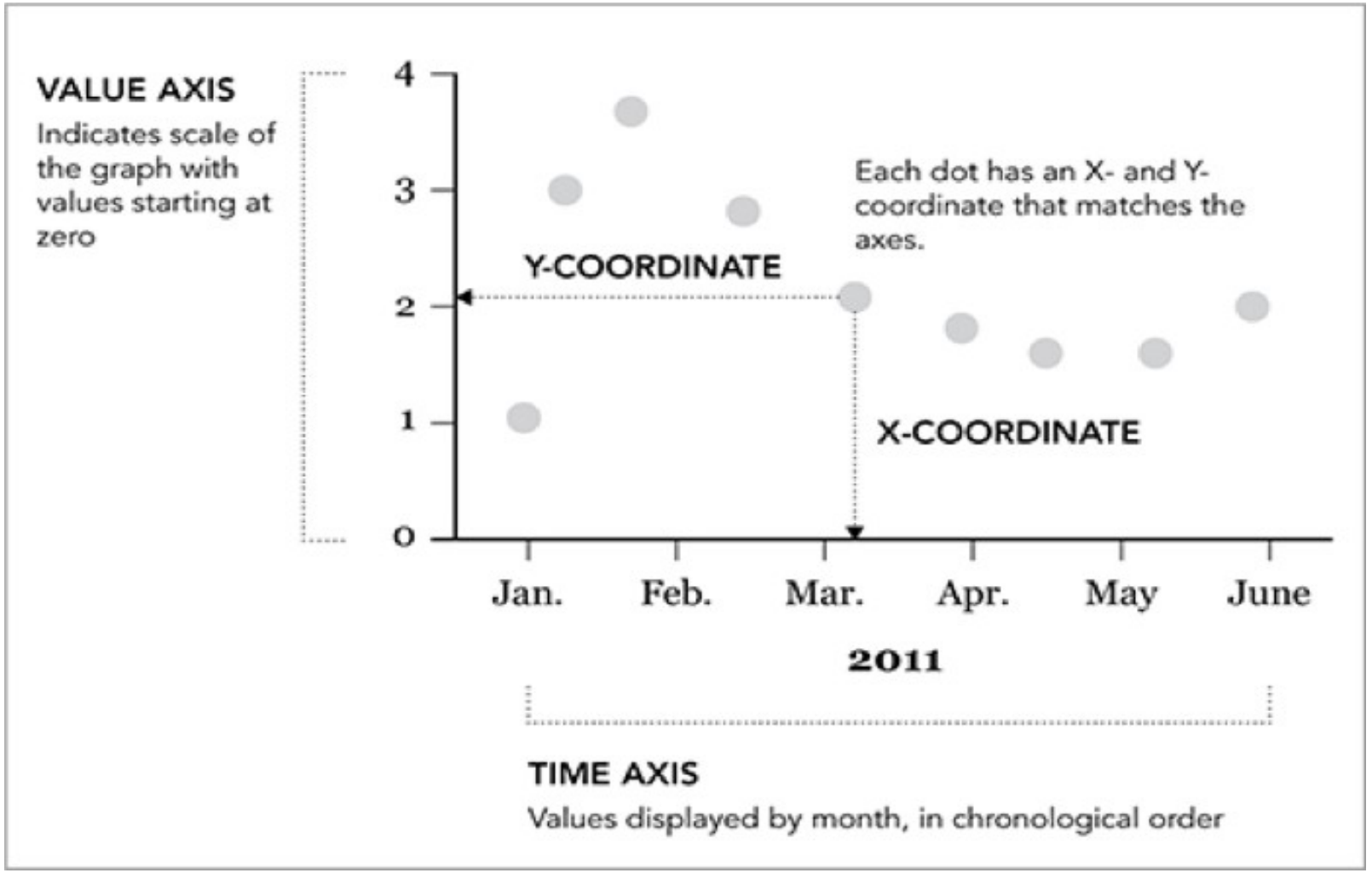
Visualize This: The Flowing Data Guide to Design, Visualization, and Statistics. Nathan Yau. 2011

Simple Bar Graph

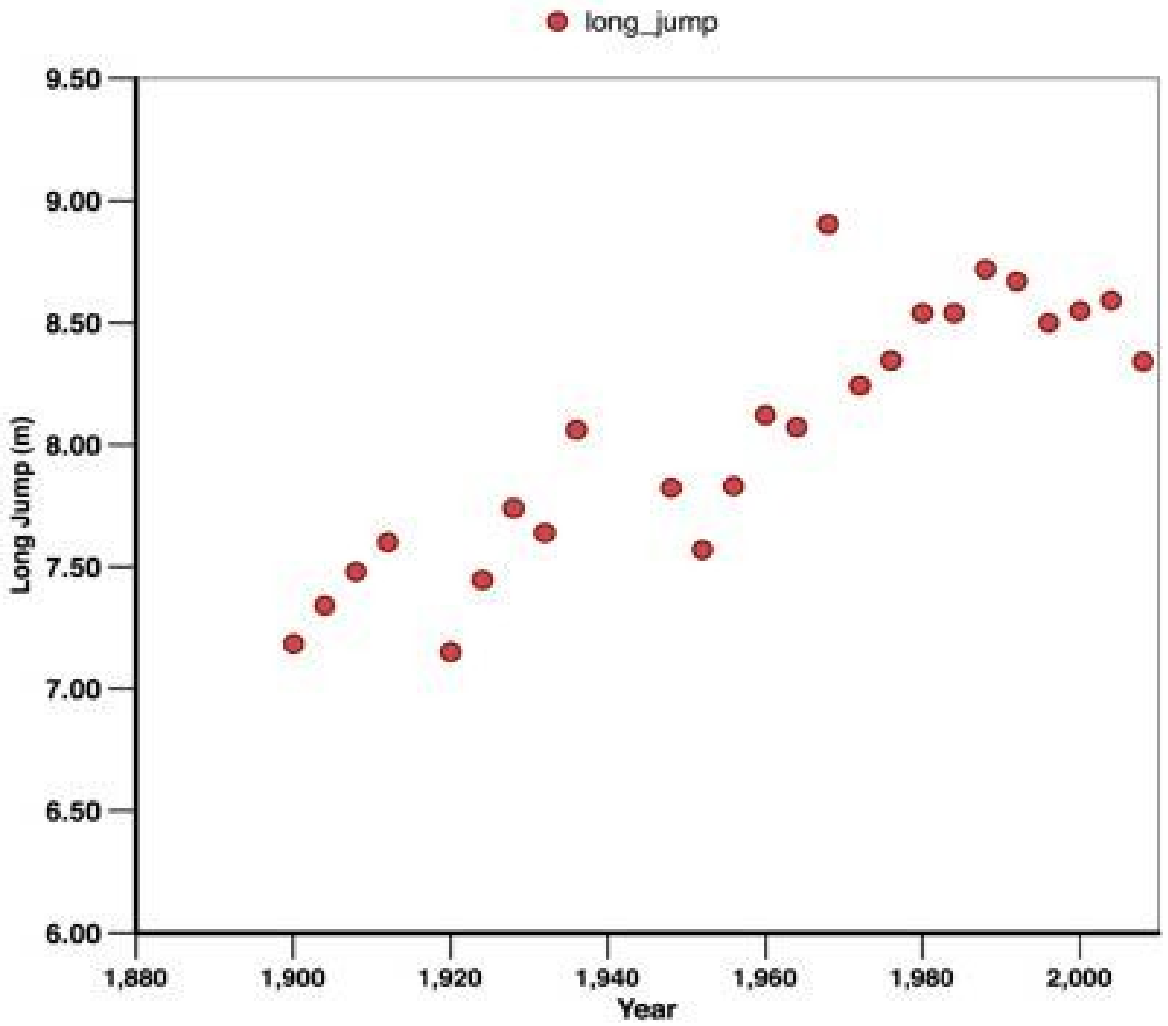


Visualize This: The Flowing Data Guide to Design, Visualization, and Statistics. Nathan Yau. 2011

Simple Scatter



Simple Scatter



https://www.macinchem.or g/reviews/aabel_review_file s/aabel.php

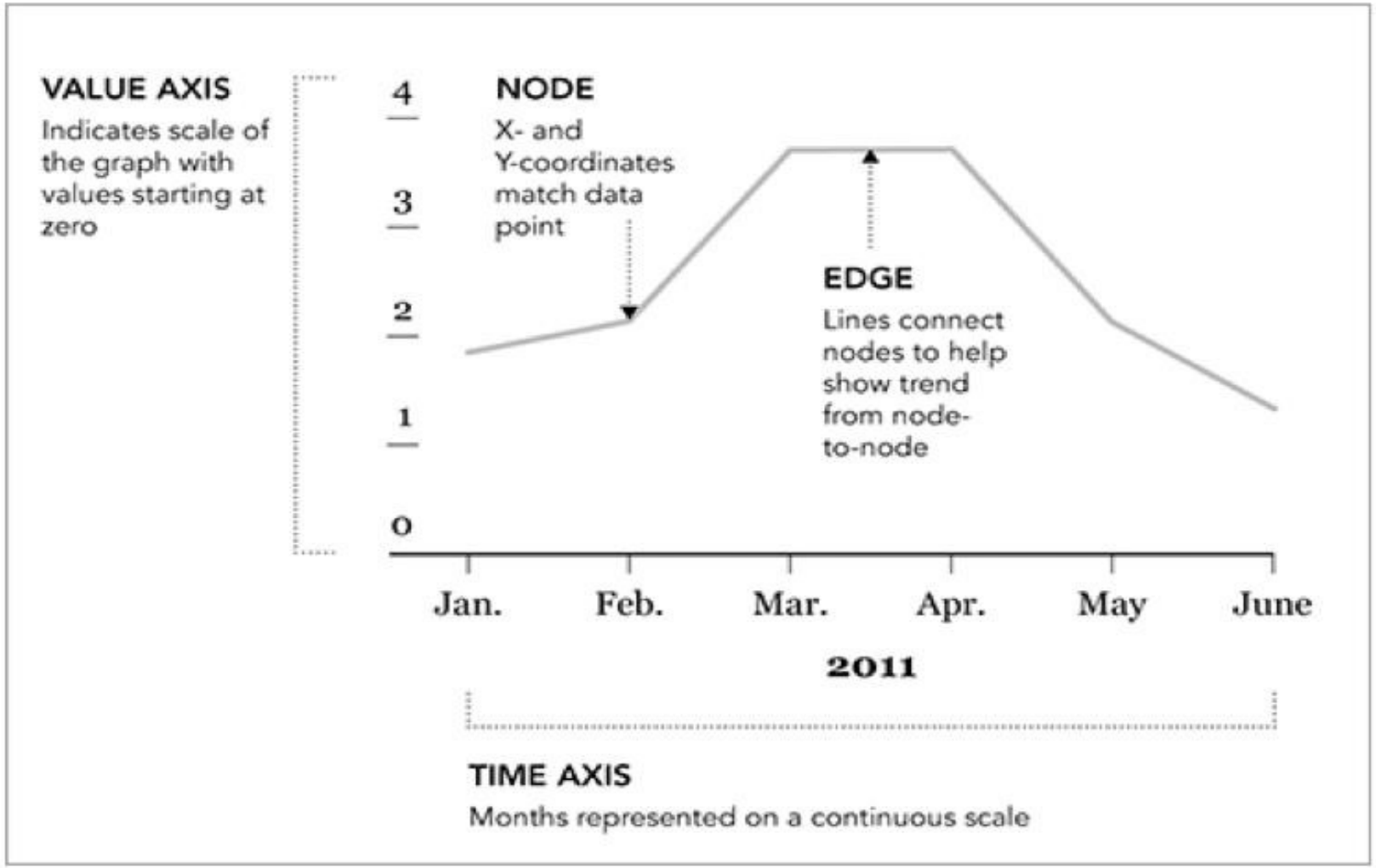
Continuous Points In Time

- Often we have datasets that contain single measurements for a number of continuous points of time
  - Stock prices over time
  - Internet traffic over time

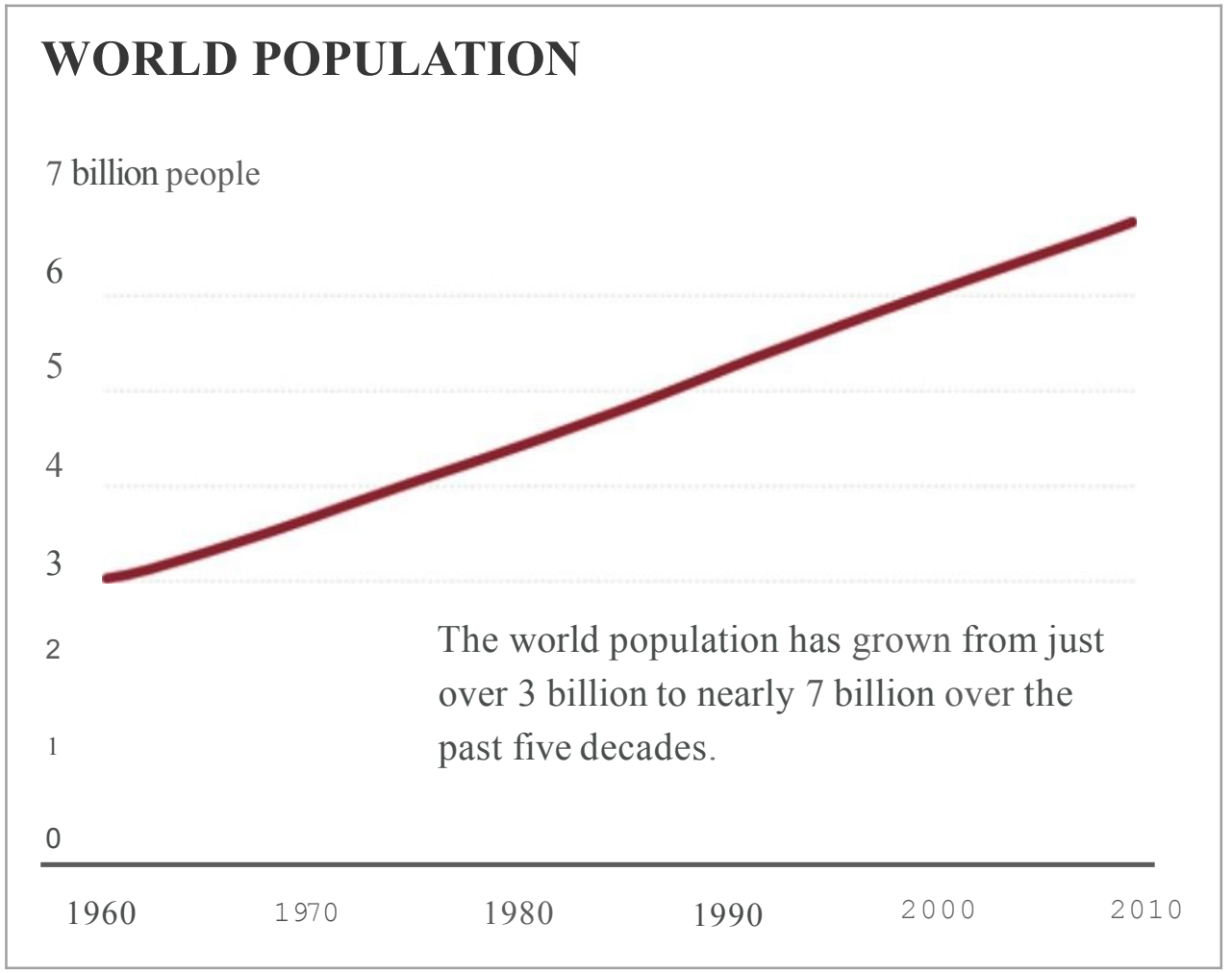
- Visualizing continuous time series data is similar discrete data.
  - Still have a discrete number of data points, even if the dataset is continuous
  - The difference is what they represent in the physical world

- In these cases a line charts are more appropriate

Simple Line Chart



Simple Line Chart



Visualize This: The Flowing Data Guide to Design, Visualization, and Statistics. Nathan Yau. 2011

Simple Line vs.Step Chart

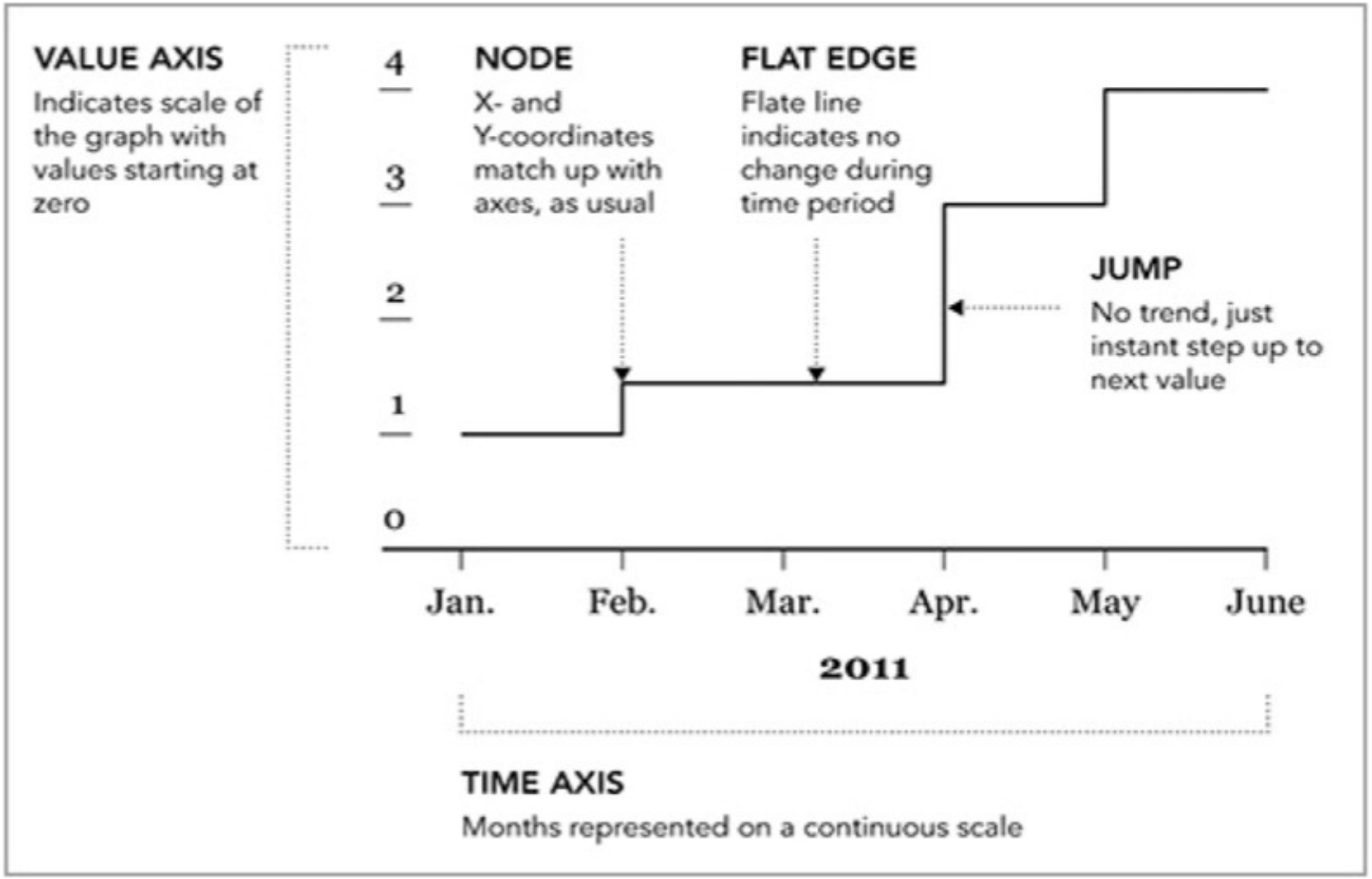
- Basic line charts **interpolate** the change in measurement from one point in time to another.

- If the user zooms in farther than the data resolution supports, our ‘truth’ as approximated by the interpolated lines is questionable.

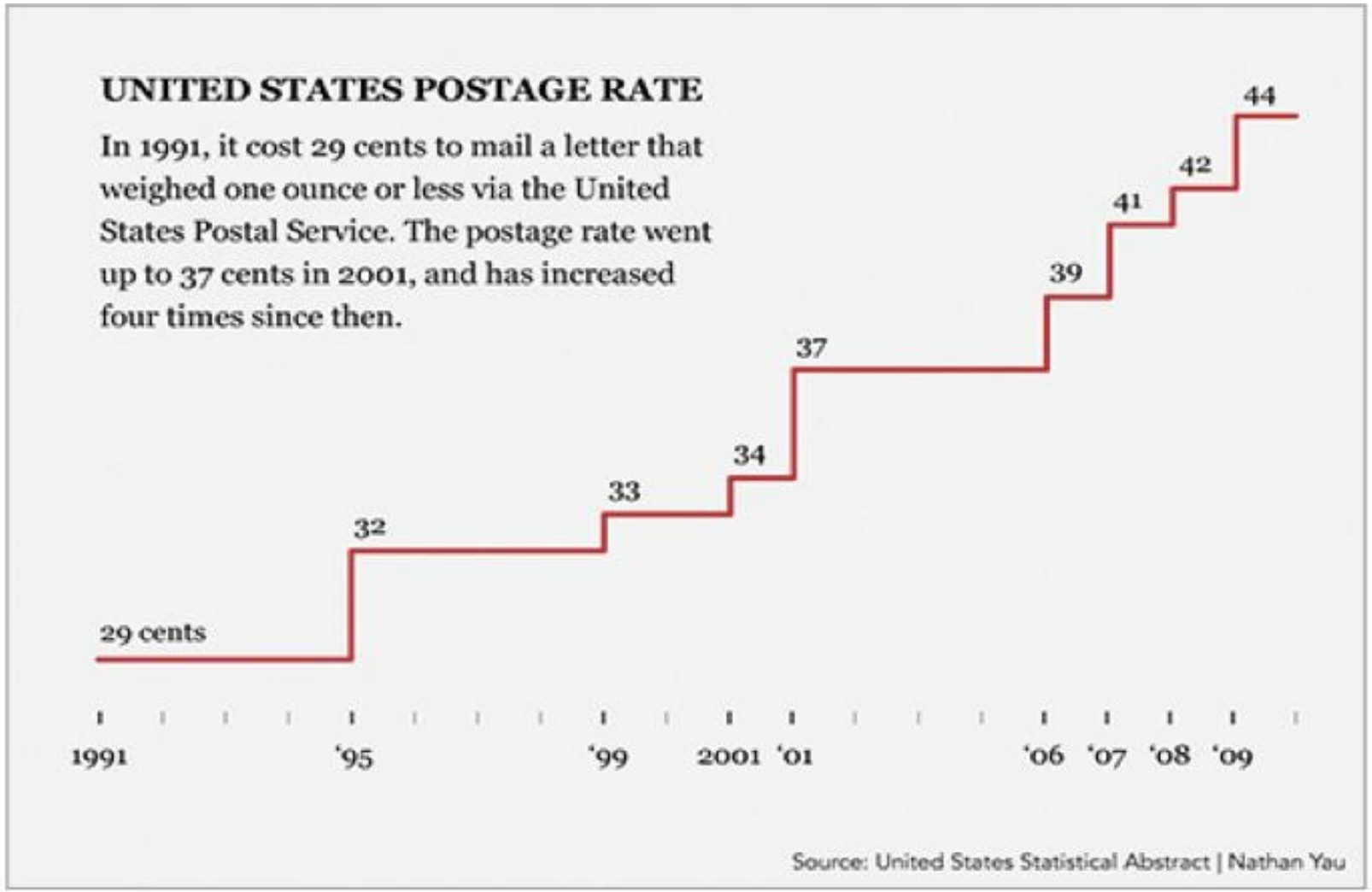
- Often this is not appropriate and in these cases step charts are more appropriate

Visualize This: The Flowing Data Guide to Design, Visualization, and Statistics. Nathan Yau. 2011

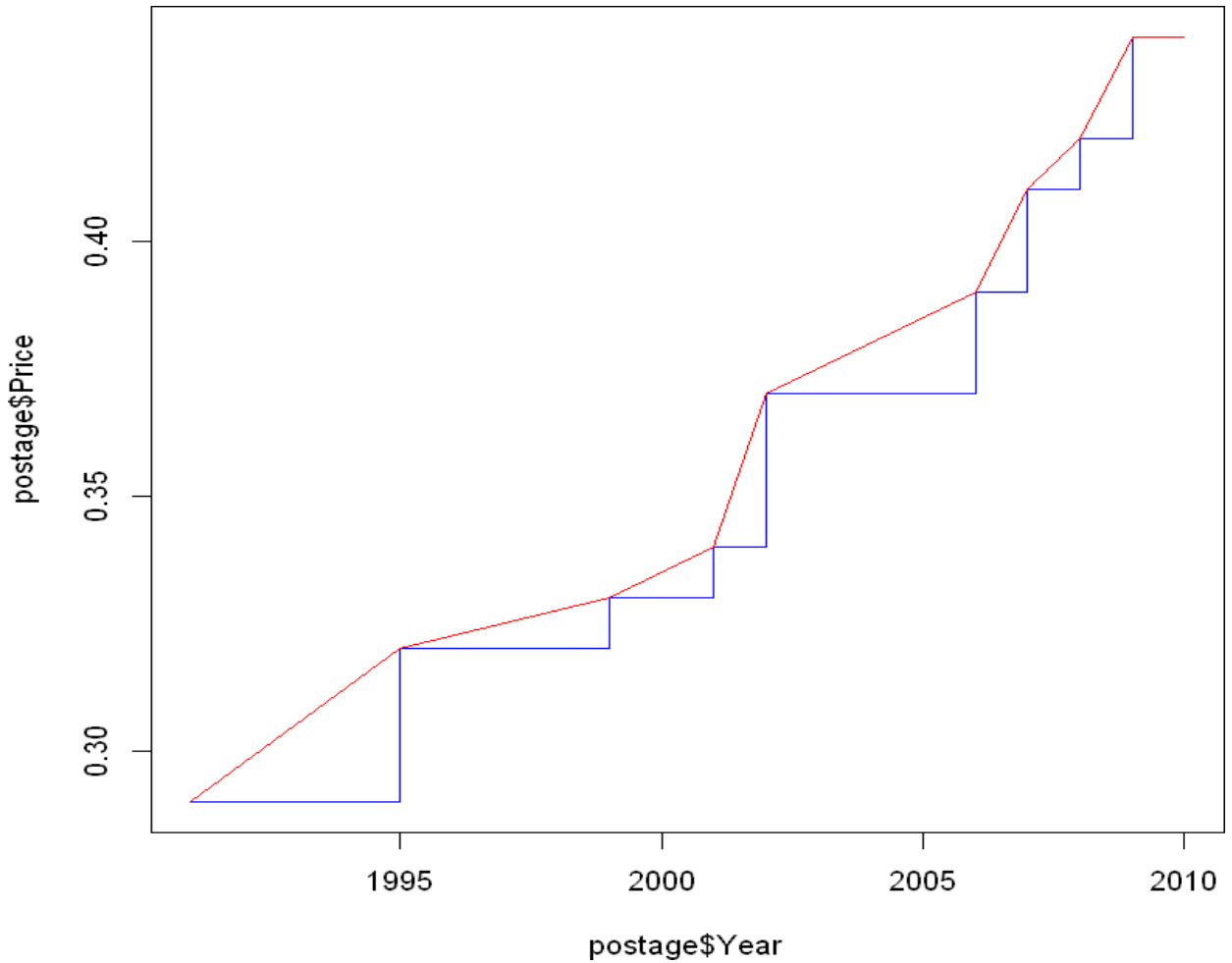
Step Chart


VALUE AXIS
Indicates scale of the graph with values starting at zero
4
3
2
1
0
NODE
X- and Y-coordinates match up with axes, as usual
FLAT EDGE
Flate line indicates no change during time period
JUMP
No trend, just instant step up to next value
Jan.
Feb.
Mar.
Apr.
May
June
2011
TIME AXIS
Months represented on a continuous scale

Step Chart


UNITED STATES POSTAGE RATE
In 1991, it cost 29 cents to mail a letter that weighed one ounce or less via the United States Postal Service. The postage rate went up to 37 cents in 2001, and has increased four times since then.
29 cents
32
33
34
37
39
41
42
44
1991
'95
'99
2001
'01
'06
'07
'08
'09
Source: United States Statistical Abstract | Nathan Yau

Step Chart



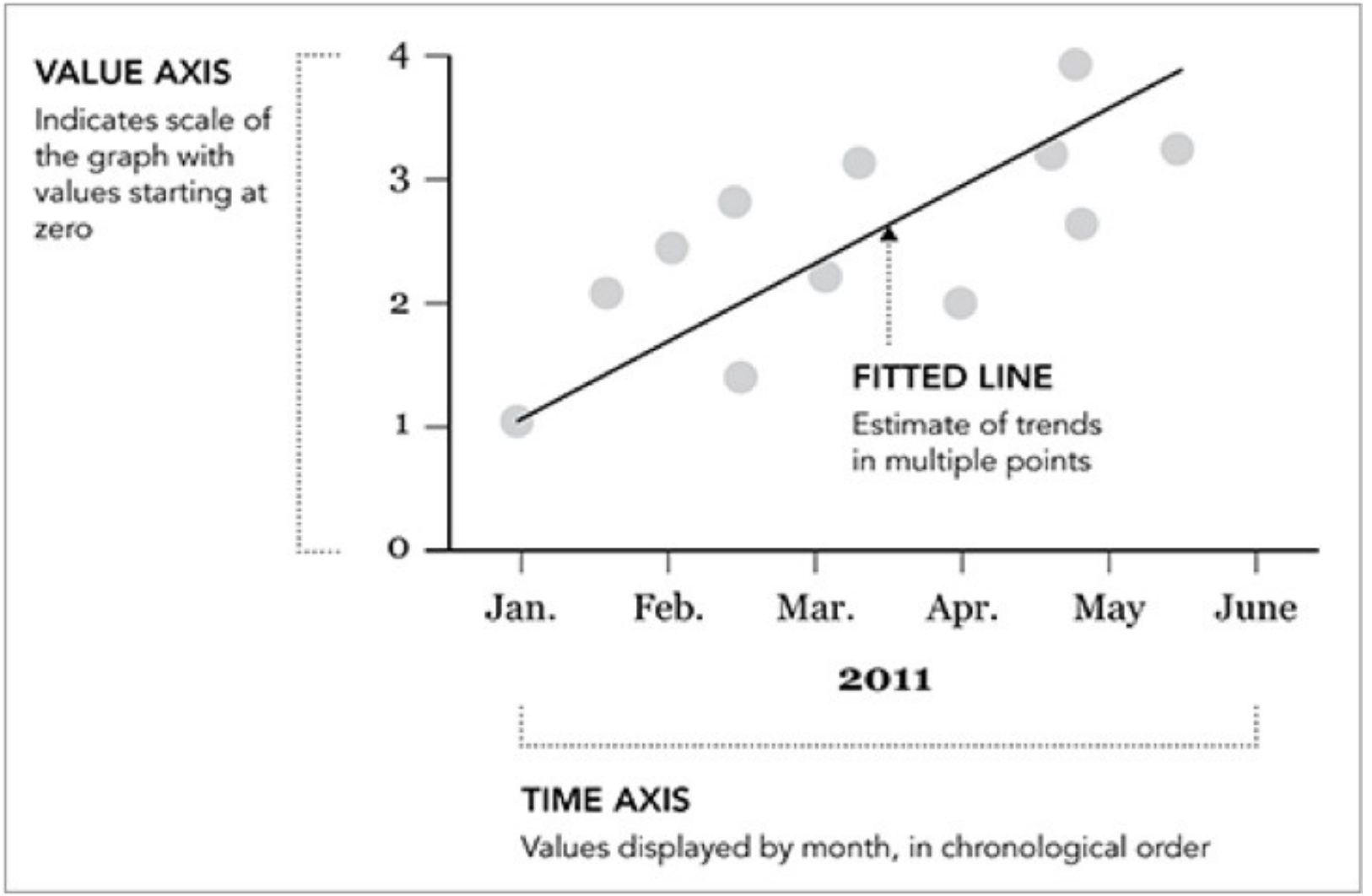
Fitting A Line

- When you have a lot of data, or the data is noisy, it can be hard to spot trends and patterns
  - So to make it easier, you can estimate a trend line, or a curve in some cases

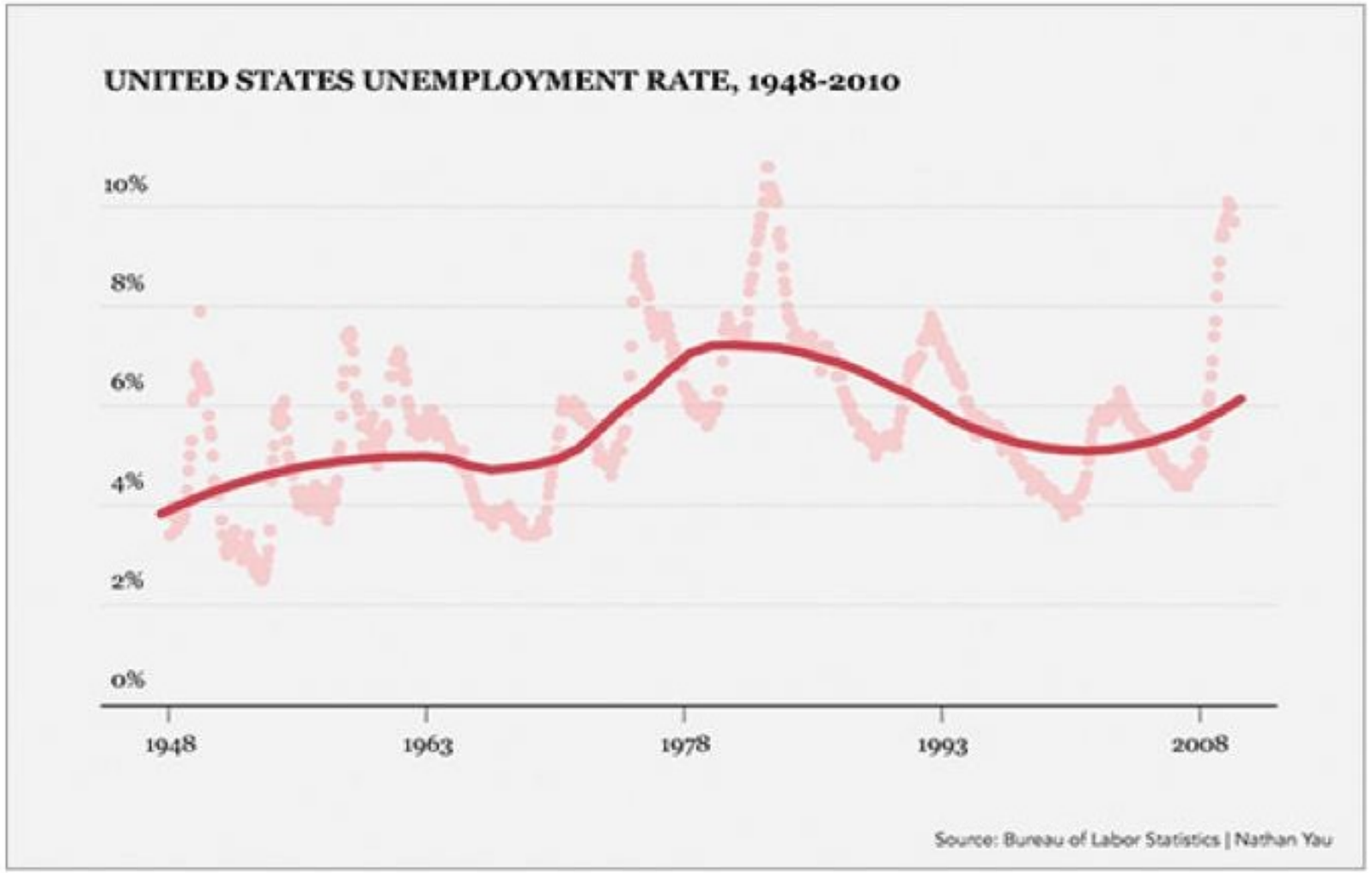
- In some cases adding a line to a scatter plot can mask the trend over time rather than help illustrate it

- Be careful when fitting lines, it is possible to illustrate trends that don’t really exist

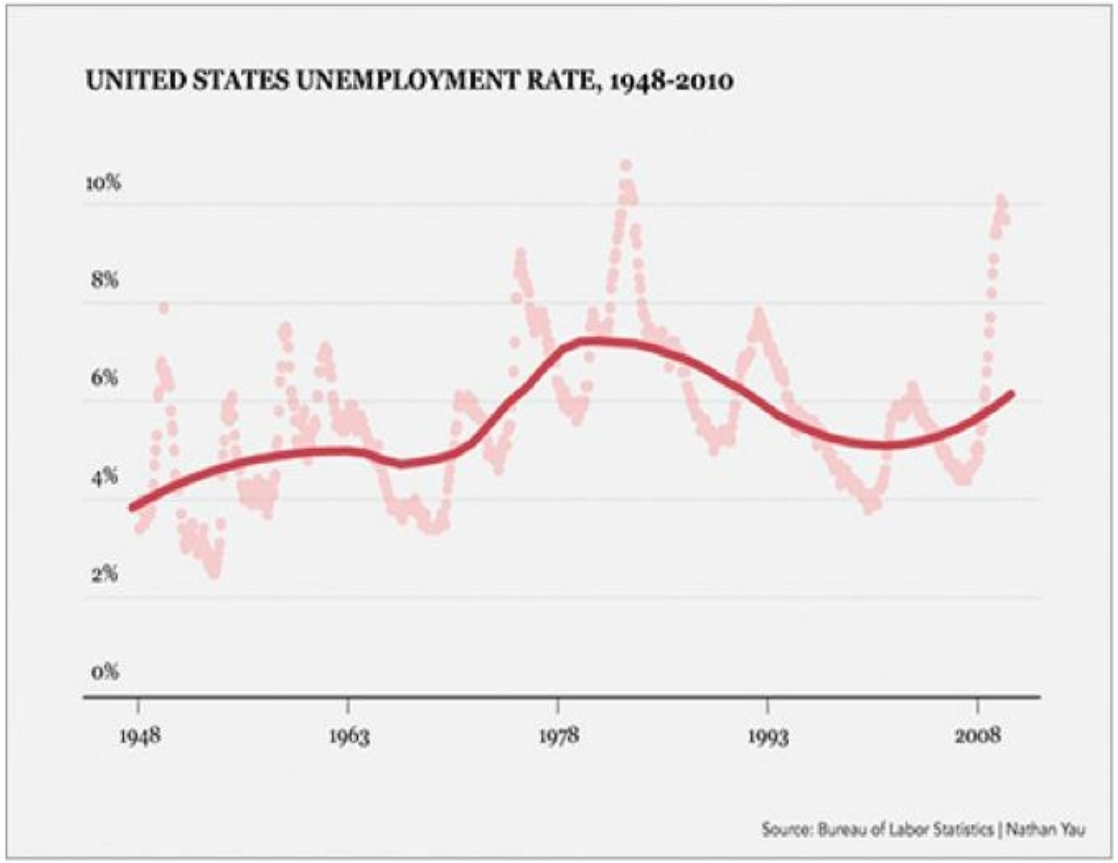
Fitting A Line



Smoothing & Estimation



Smoothing & Estimation



Non-linear trends

Cleveland & Devlin's LOESS Locally weighted scatterplot smoothing

- Starts at the beginning of the data and takes small slices
- At each slice it estimates a low-degree polynomial for the data in the slice
- LOESS moves along fitting a series of tiny curves
- Together they form a single curve

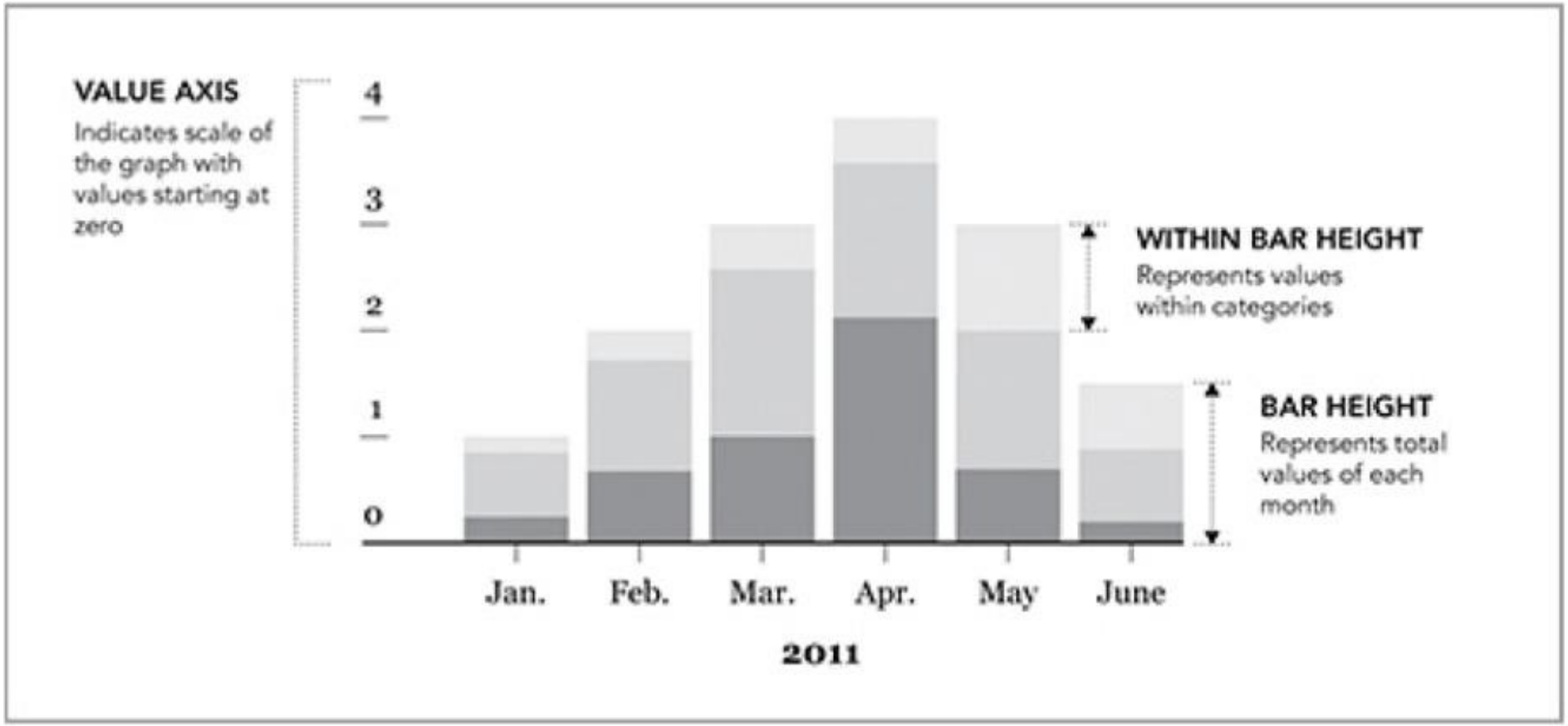
Multiple Dimensions In Time

- Often we have datasets that contain **multiple measurements** for each point in time which gives us **multiple dimensions**

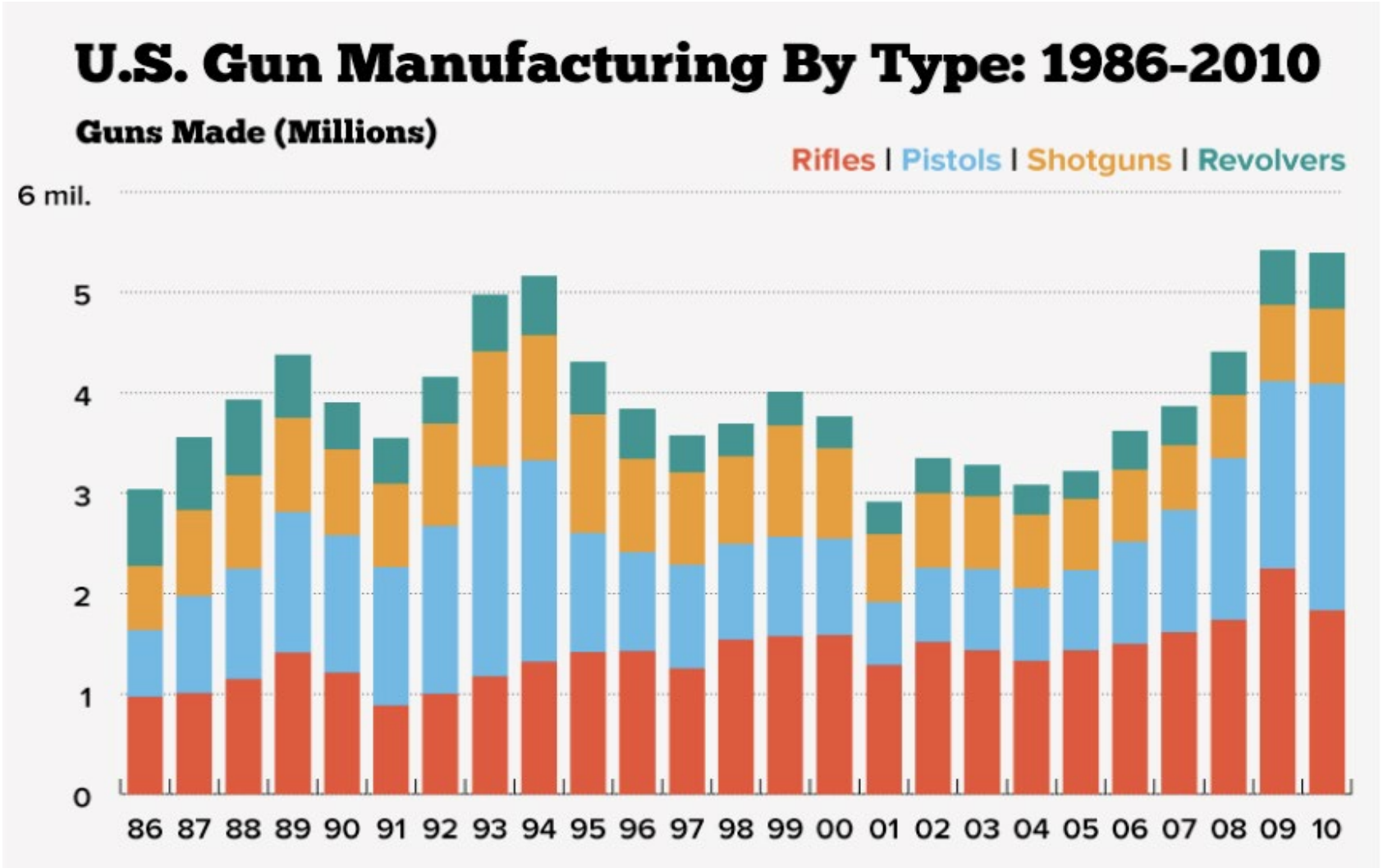
- This makes visualisation more challenging as we need to use newer and different encodings

- There are a number of approaches we can take

Stacked Bar Chart



Stacked Bar Chart

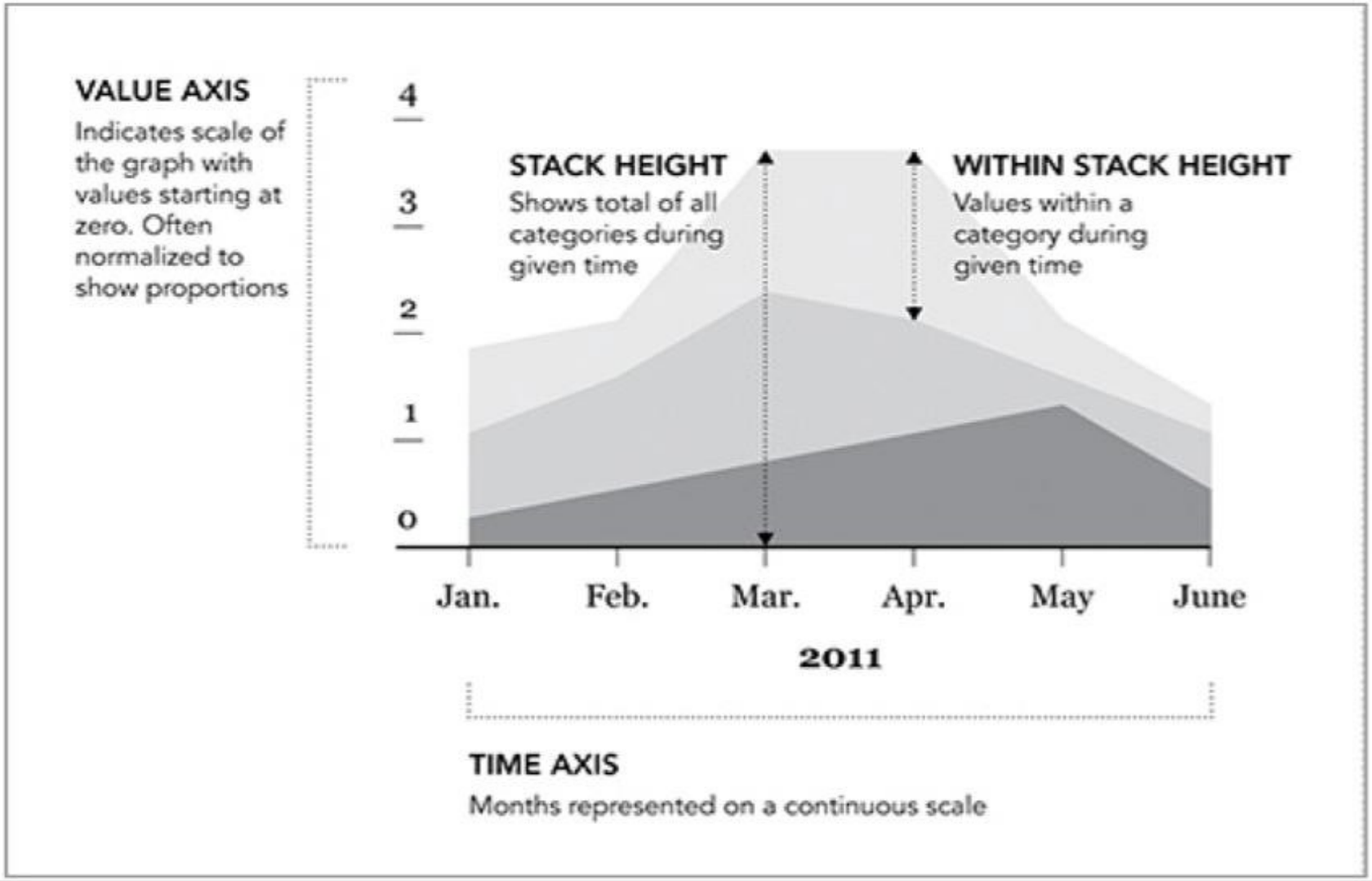


Stacked Bar Chart

- Interpreting multiple time series in a stacked bar chart can be really difficult, especially when we need to interpret the trend within the *higher* categories

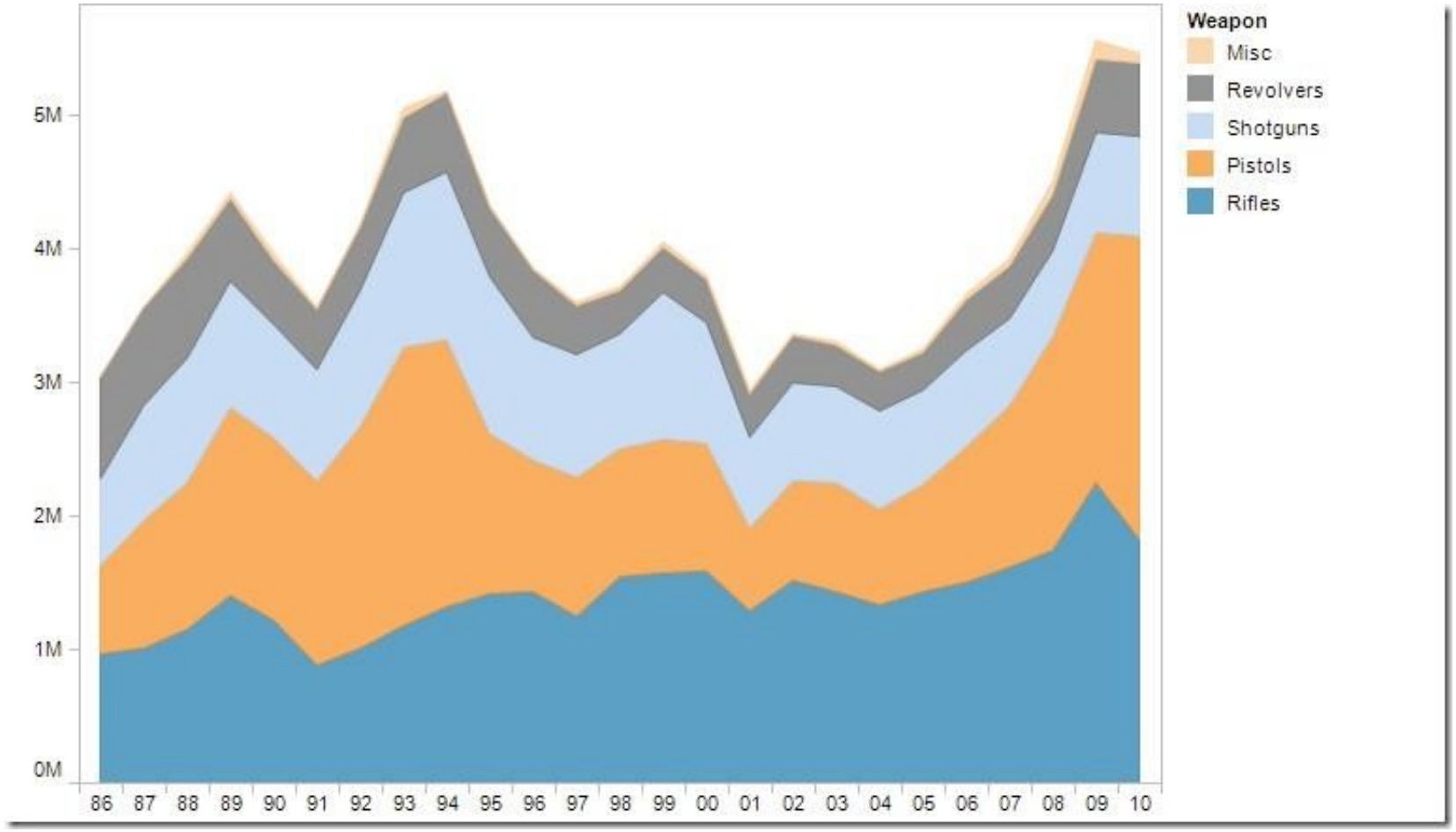
- Only really appropriate when the real trend we want to illustrate is the overall total bar height trend and the individual categories are secondary

Stacked Area Chart



Stacked Area Chart

U.S. Gun Manufacturing By Type



http://vizwiz.blogspot.ie/2012/08/displaying-time-series-data-stacked.html

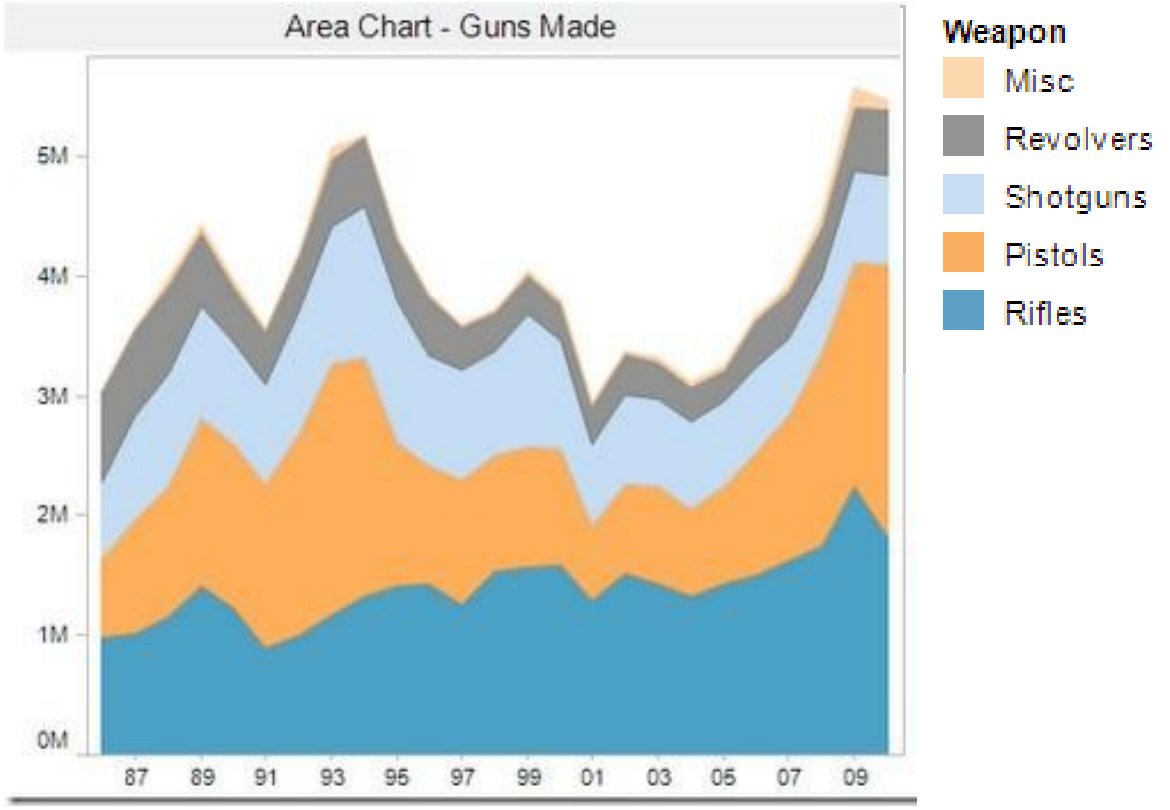
Stacked Area Chart

- Stacked area charts unfortunately suffer from many of the same problems as stacked bar charts

- It can be very hard to interpret the trends of each category as they are stacked on top of each other

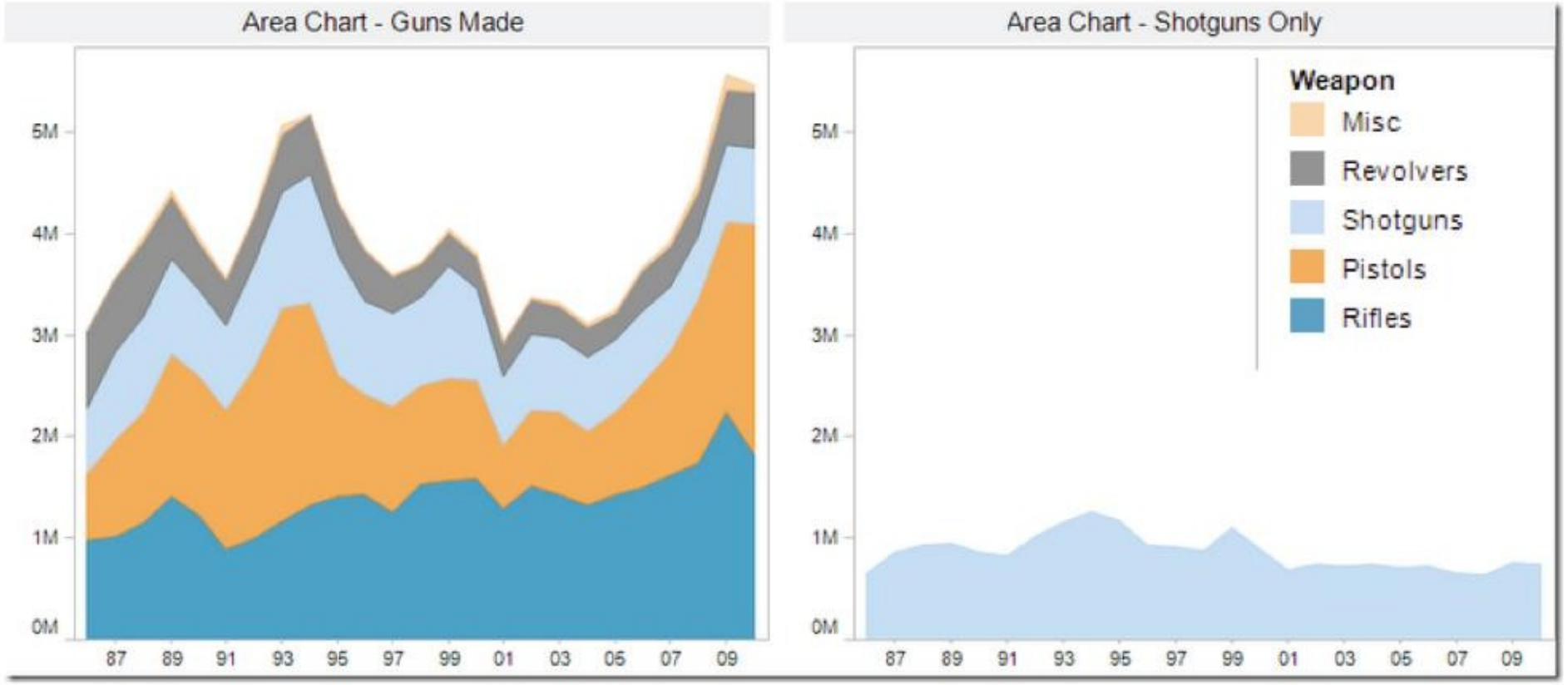
Stacked Area Chart

- What does the shotgun trend look like?



http://vizwiz.blogspot.ie/2012/08/displaying-time-series-data-stacked.html

Stacked Area Chart



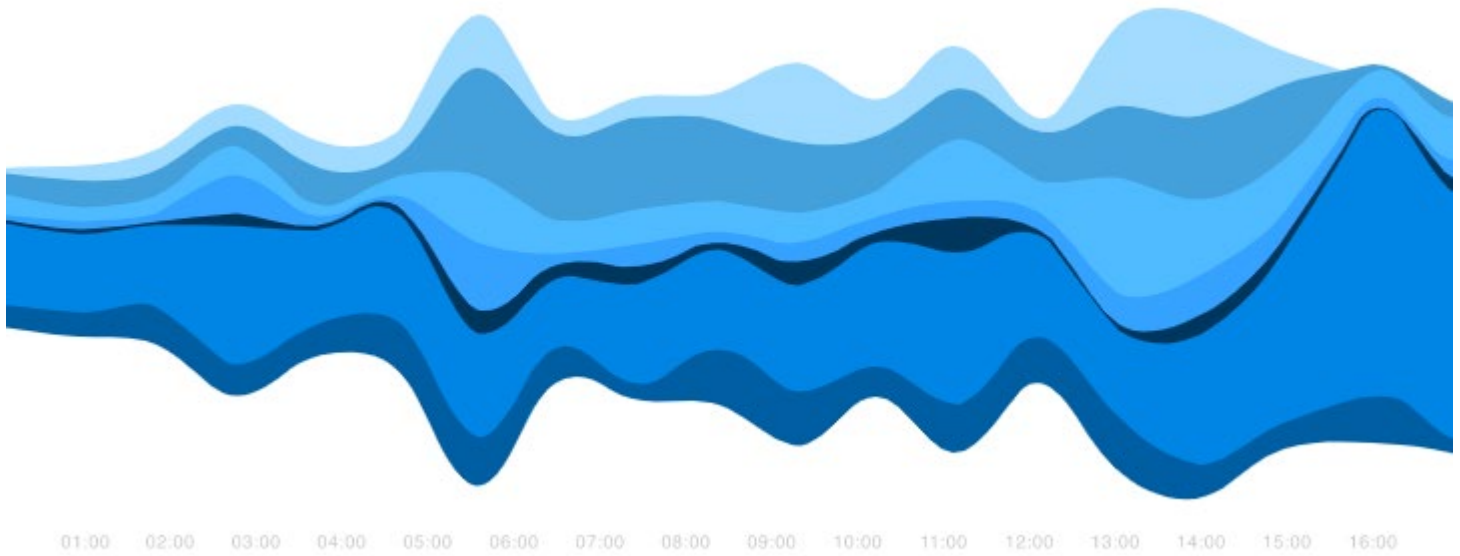
http://vizwiz.blogspot.ie/2012/08/displaying-time-series-data-stacked.html

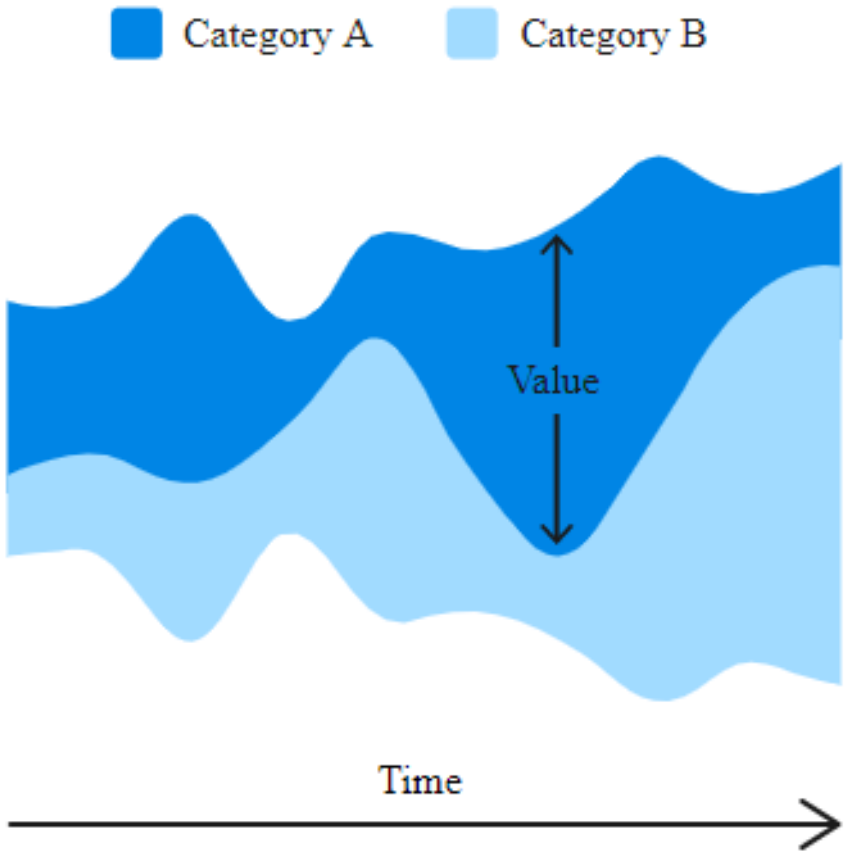
Stream Graph

- The Stream Graph is an attempt to overcome the difficulties associated with stacked bar charts and stacked area charts

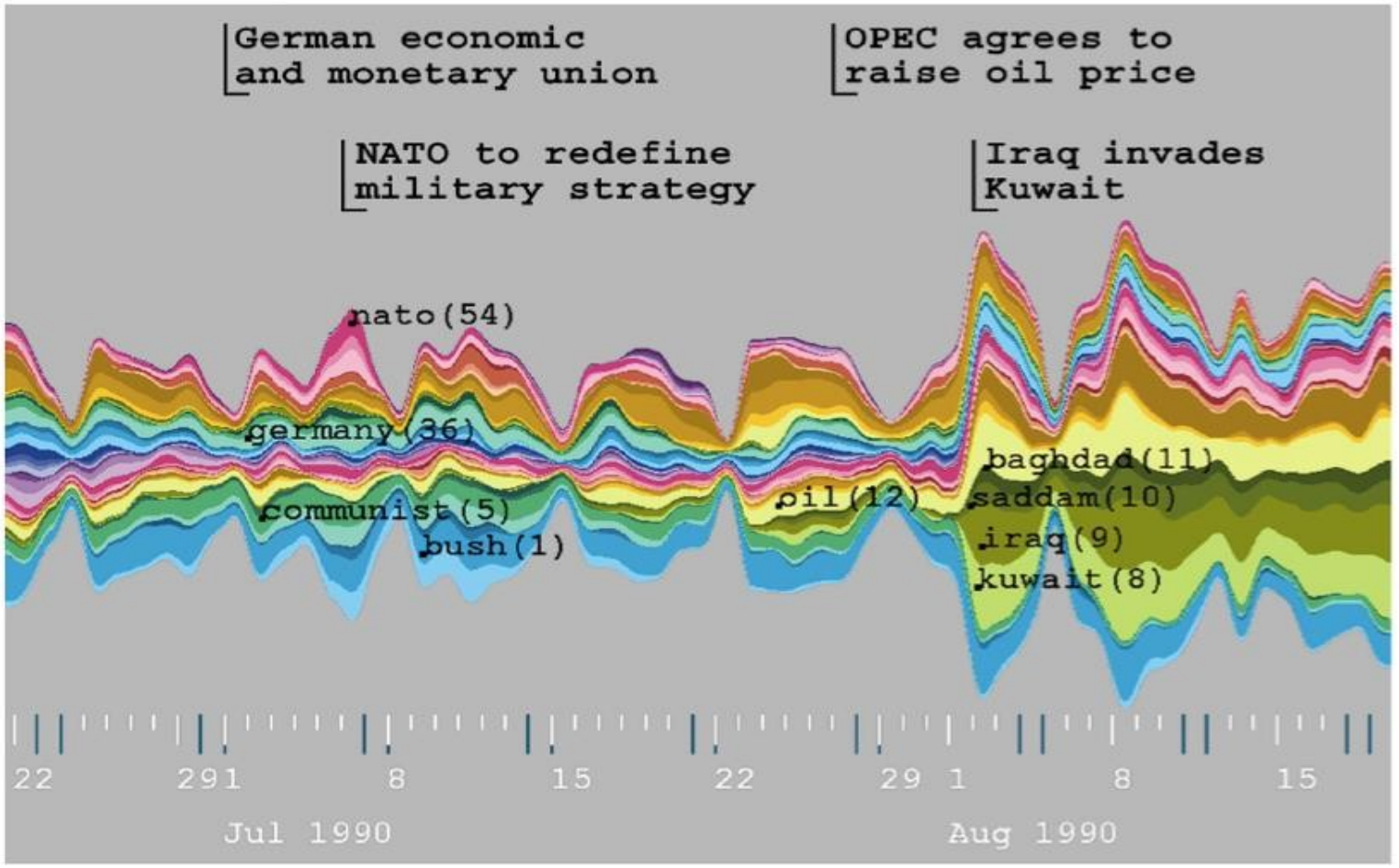
- Instead of plotting values against a fixed, straight axis, a Stream Graph has values displaced around a varying central baseline

Stream Graph





Stream Graph



ThemeRiver: Visualizing Theme Changes over Time, Susan Havre, Beth Hetzler, Lucy Nowell, 2000

Stream Graph

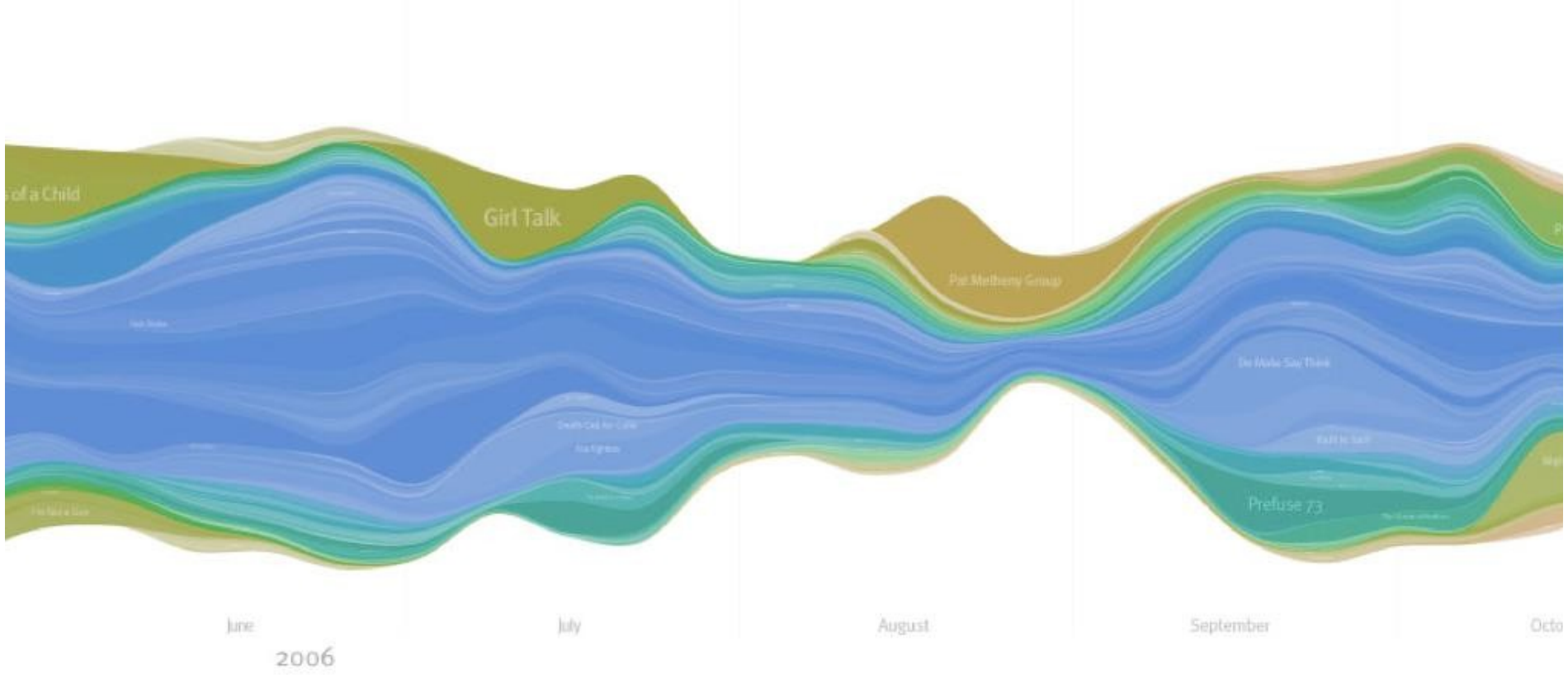
Pros

- Stream Graphs are better for giving a more general view of the data
- They also tend to work significantly better as an interactive piece rather than a static or printed graphics

Cons

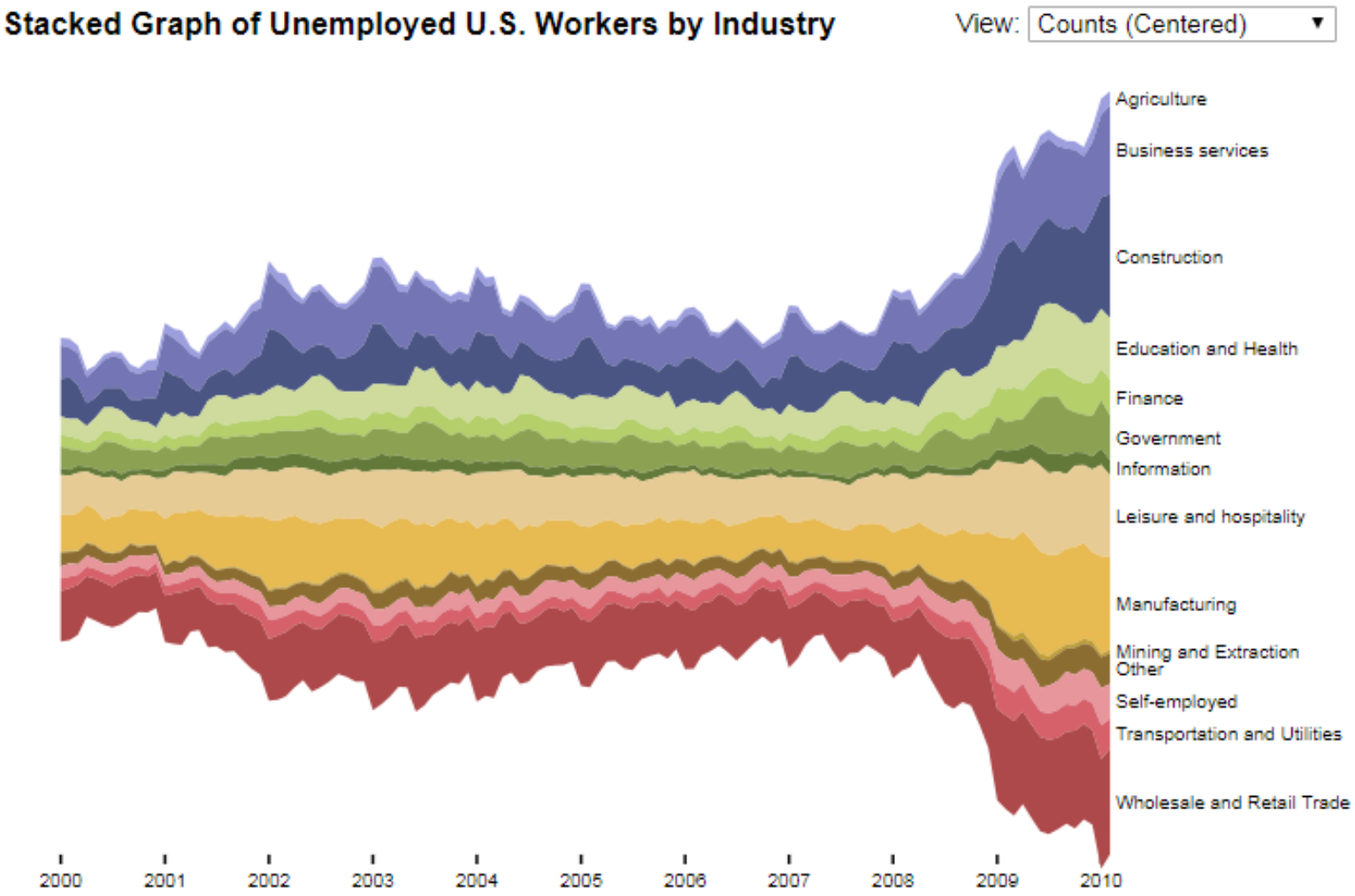
- Suffer from legibility issues, as they are often very cluttered with large datasets
- The categories with smaller values are often drowned out to make way for categories with much larger values, making it impossible to see all the data
- Also, it's impossible to read the exact values visualised in a Stream Graph, as there is no axis to use as a reference

Stream Graph



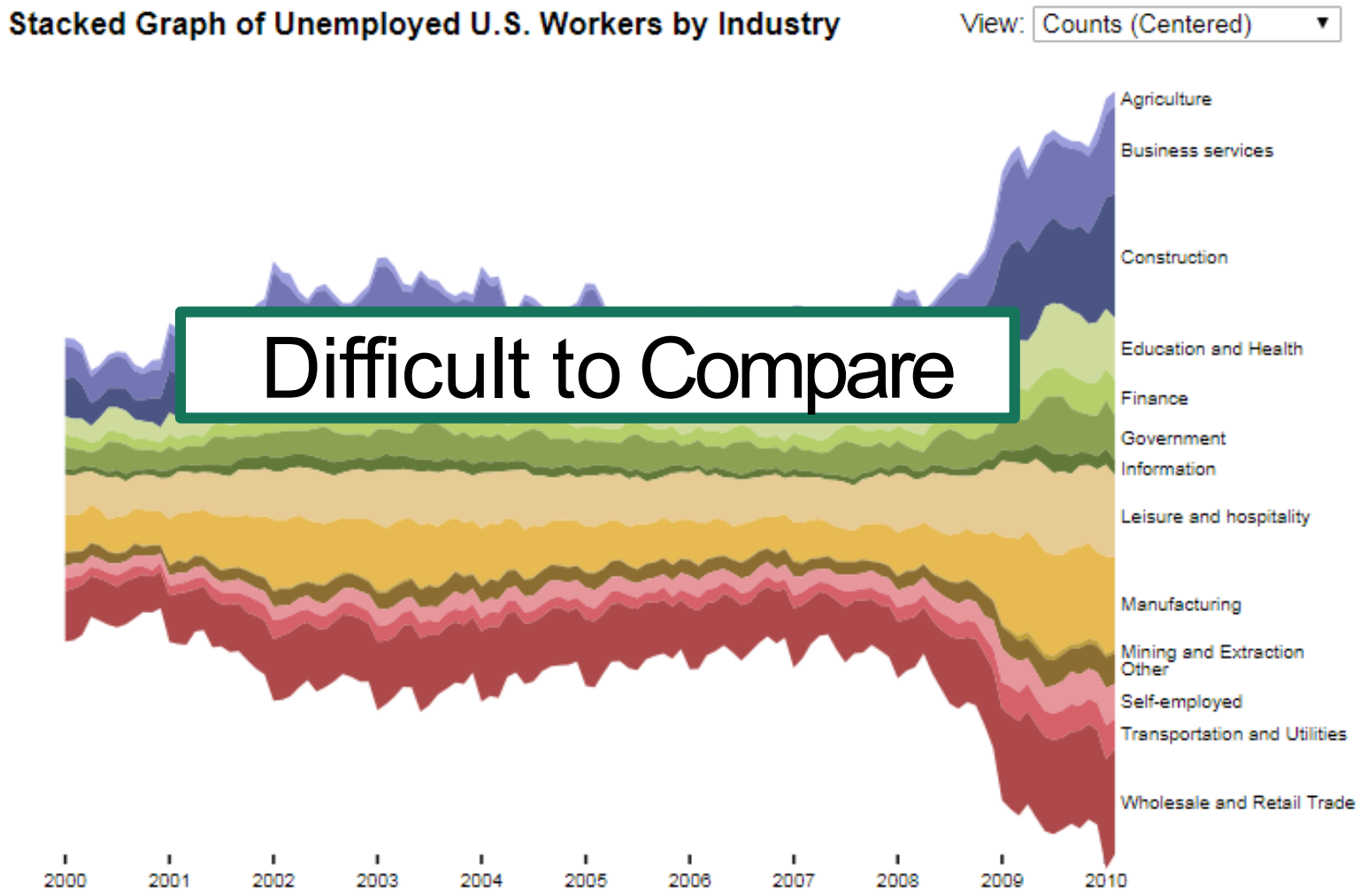
Stacked Graphs - Geometry &Aesthetics, Lee Byron & Martin Wattenberg, 2008.

Stream Graph



https://homes.cs.washington.edu/~jheer//files/zoo/ex/time/stack.html

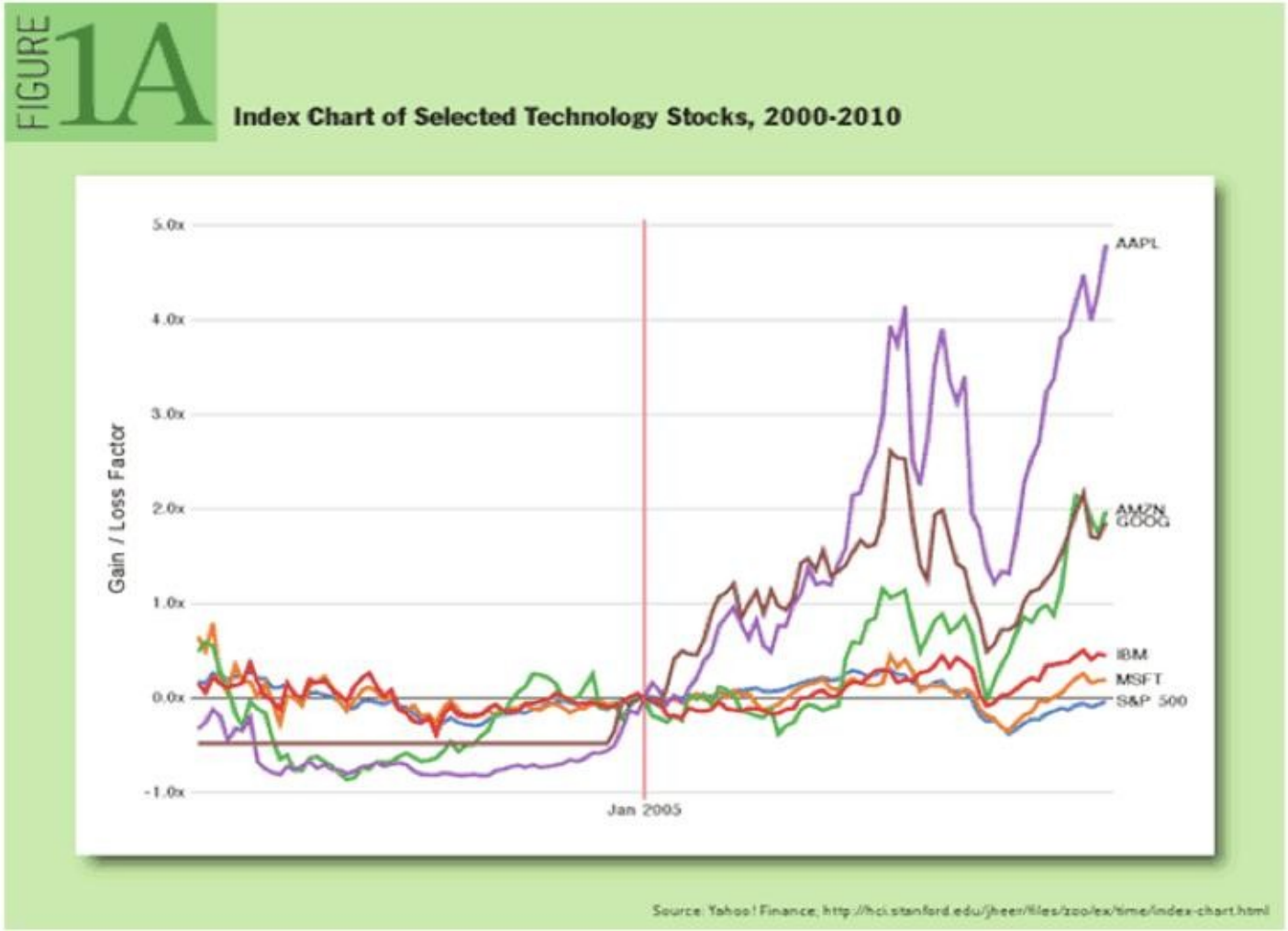
Stream Graph



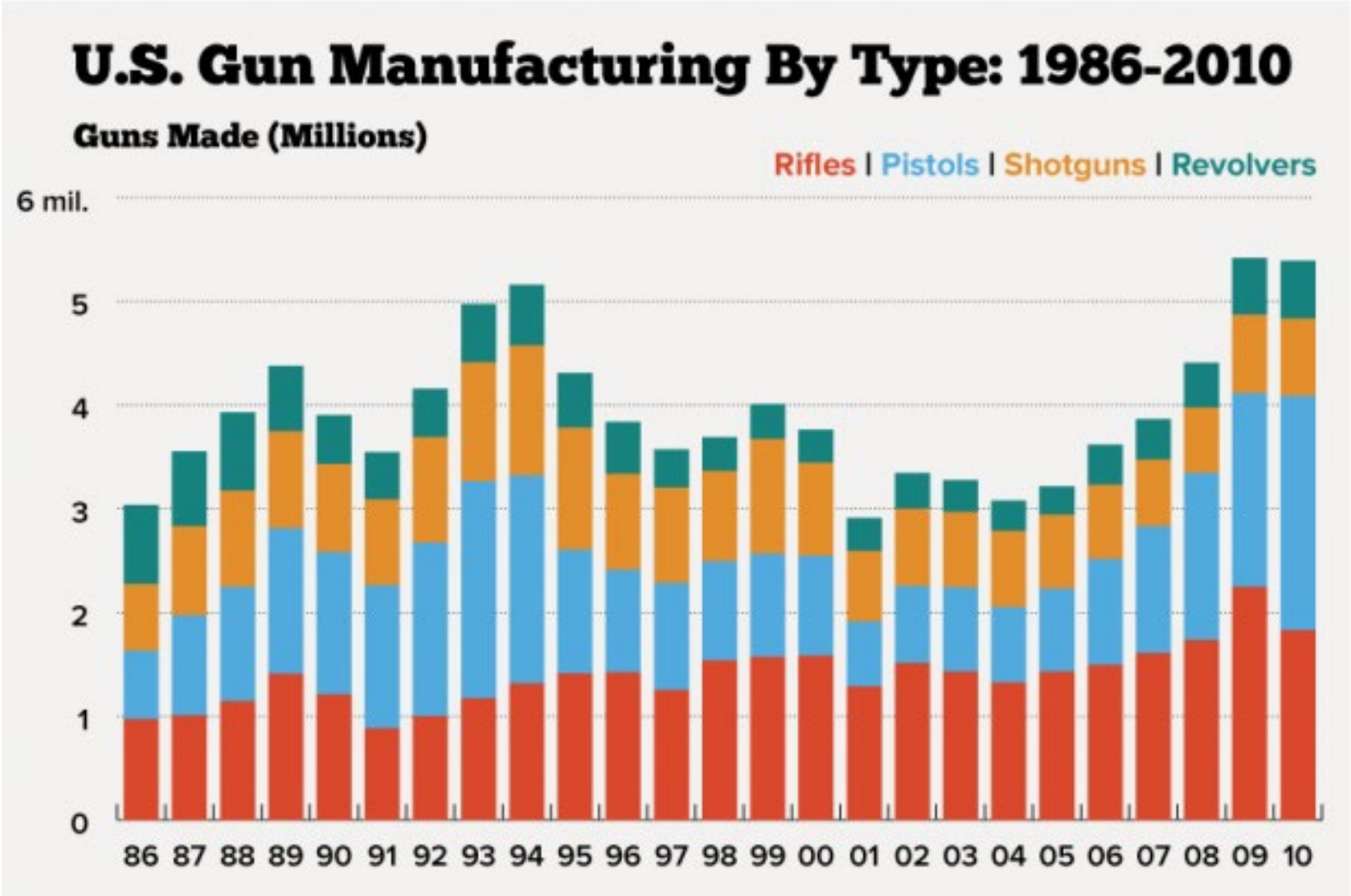
Difficult to Compare

https://homes.cs.washington.edu/~jheer//files/zoo/ex/time/stack.html

Multiple Line Series

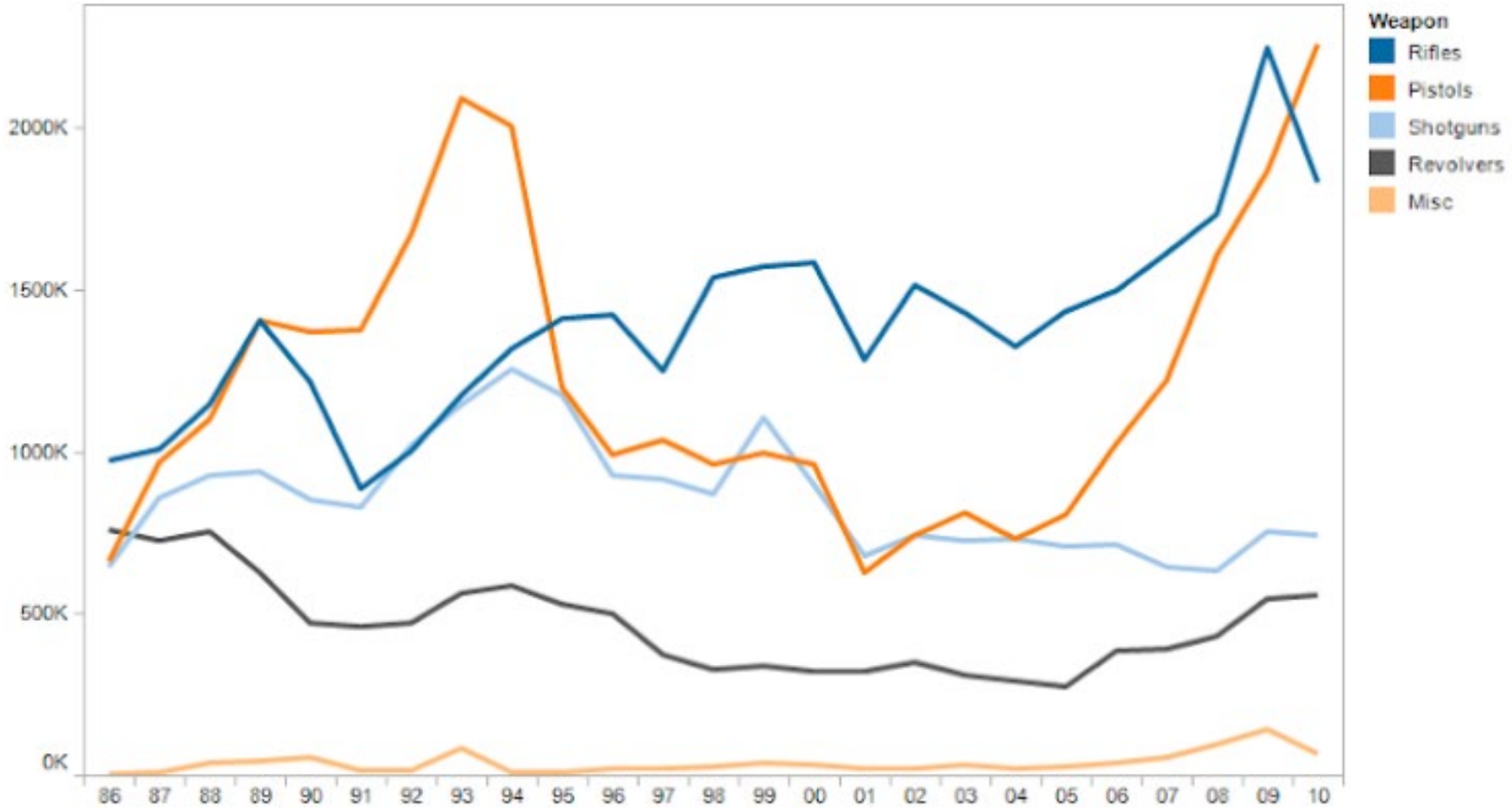


Makeover


U.S. Gun Manufacturing By Type: 1986-2010
Guns Made (Millions)
Rifles | Pistols | Shotguns | Revolvers
6 mil.
5
4
3
2
1
0
86 87 88 89 90 91 92 93 94 95 96 97 98 99 00 01 02 03 04 05 06 07 08 09 10

Multiple Line Series

U.S. Gun Manufacturing By Type

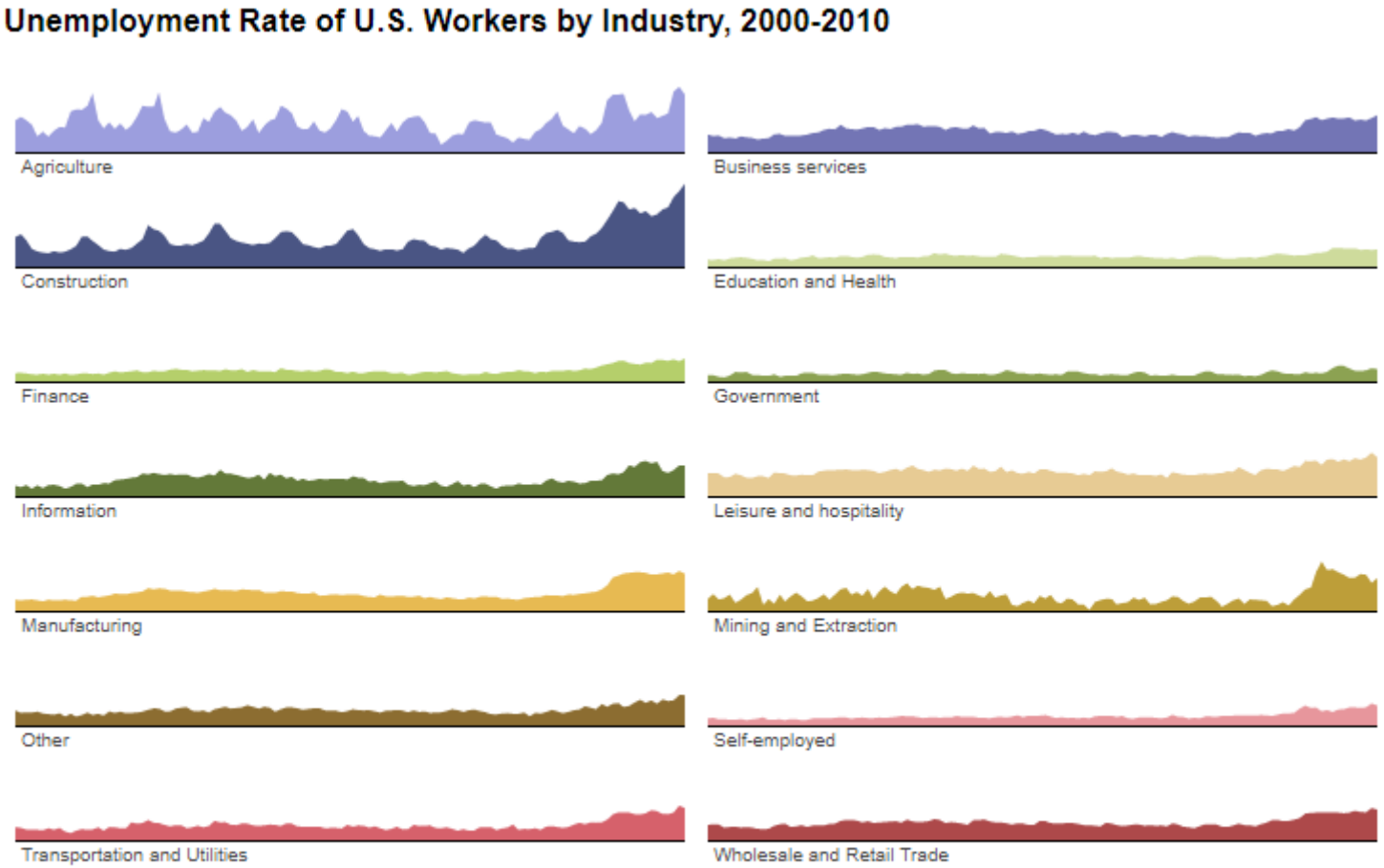


http://vizwiz.blogspot.ie/2012/08/displaying-time-series-data-stacked.html

Small Multiples

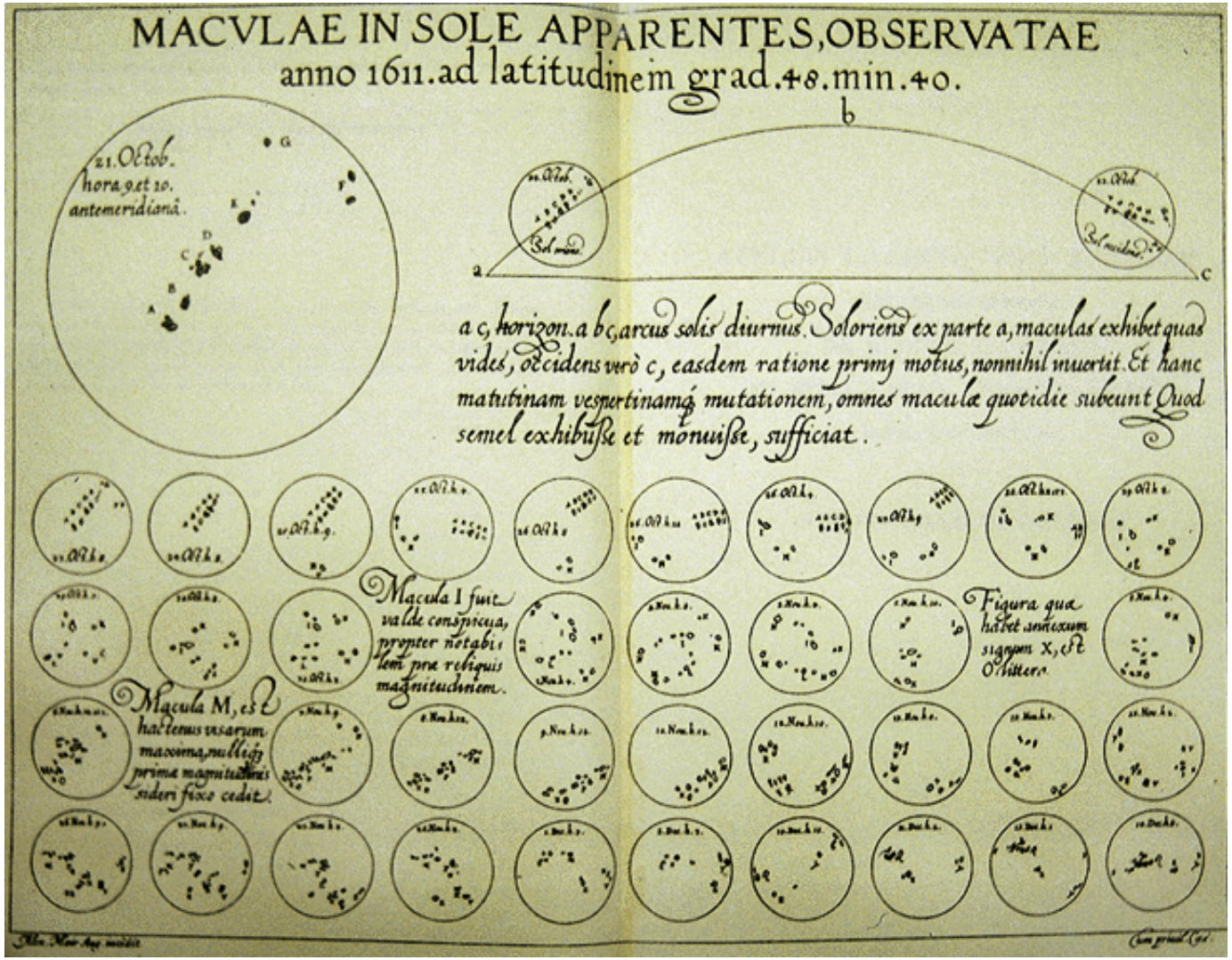
- One solution to illustrating time is to use **small multiples** to show multiple snapshots of the data at different points in time

Small Multiples



https://homes.cs.washington.edu/~jheer//files/zoo/ex/time/multiples.html

Small Multiples



http://galileo.rice.edu/sci/observations/sunspots.html

Animation

- Animation is obviously a great way to illustrate changes over time
- We will look at this more later in the course

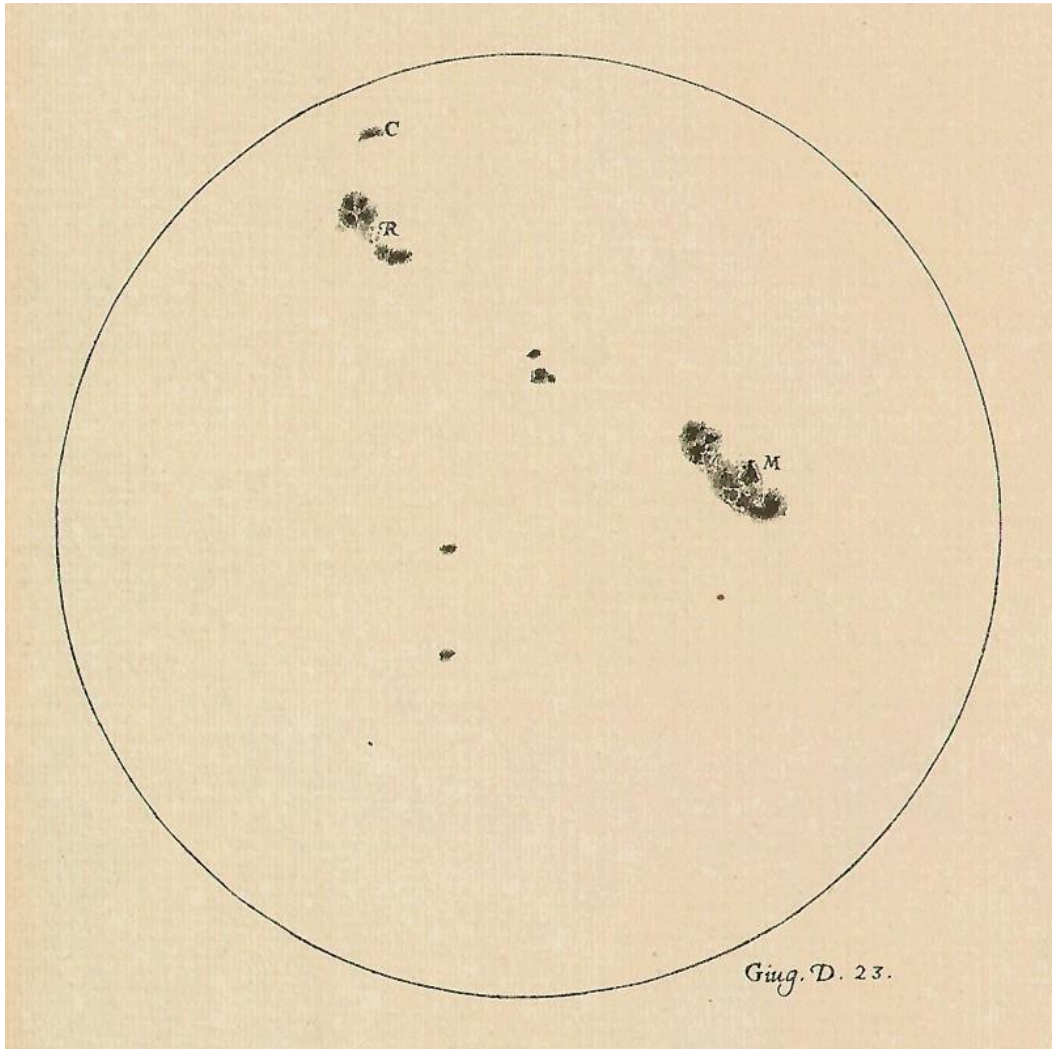
Animation

https://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen?language=en

Animation

- Galileo made flip books of sunspot images to show how they moved over time

- https://www.youtube.com/watch?v=48uECTvqbhE

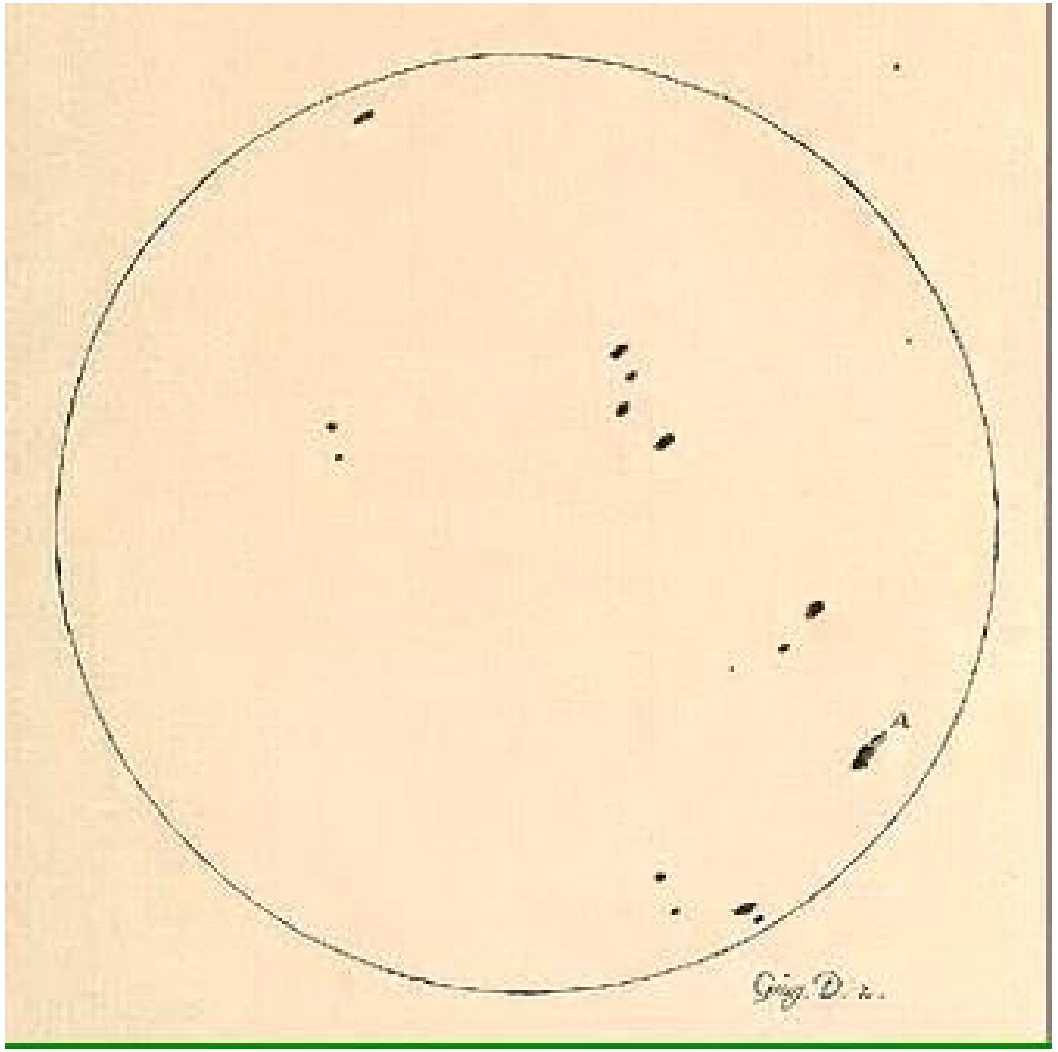


http://galileo.rice.edu/sci/observations/sunspots.html

Animation

- Galileo made flip books of sunspot images to show how they moved over time

- https://www.youtube.com/watch?v=48uECTvqbhE



http://galileo.rice.edu/sci/observations/sunspots.html

Animation

- A day in the life of Americans
- https://flowingdata.com/2015/12/15/a-day-in-the-life-of-americans/

Conclusions

- Visualising across time is a key methodology
- There are a number of approaches that are common and useful

Thanks To

- Marisa Llorens-Salvador, John McAuley, Colman McMahon and Brian Mac Namee for an earlier version of these lecture notes