DS501: Visualization

Prof. Randy Paffenroth
rcpaffenroth@wpi.edu

Worcester Polytechnic Institute

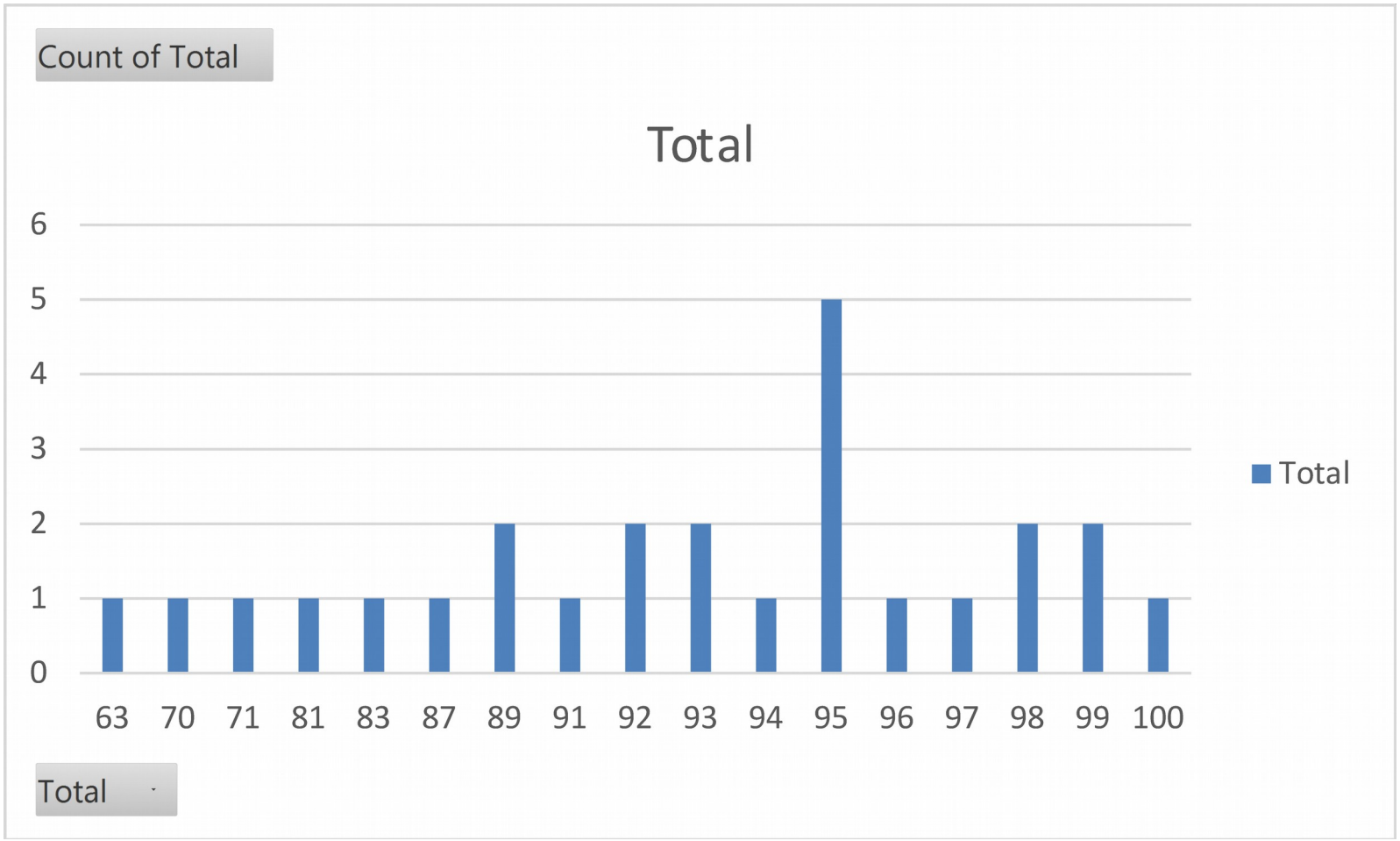
Case study 2

Case study 2

- The comments from the grader:
  - ***I think students have made huge progress from class and most of them did good analysis this time. They are so brilliant and acute to data that I was surprised about their interesting conjectures/analysis when I was reading their reports.***

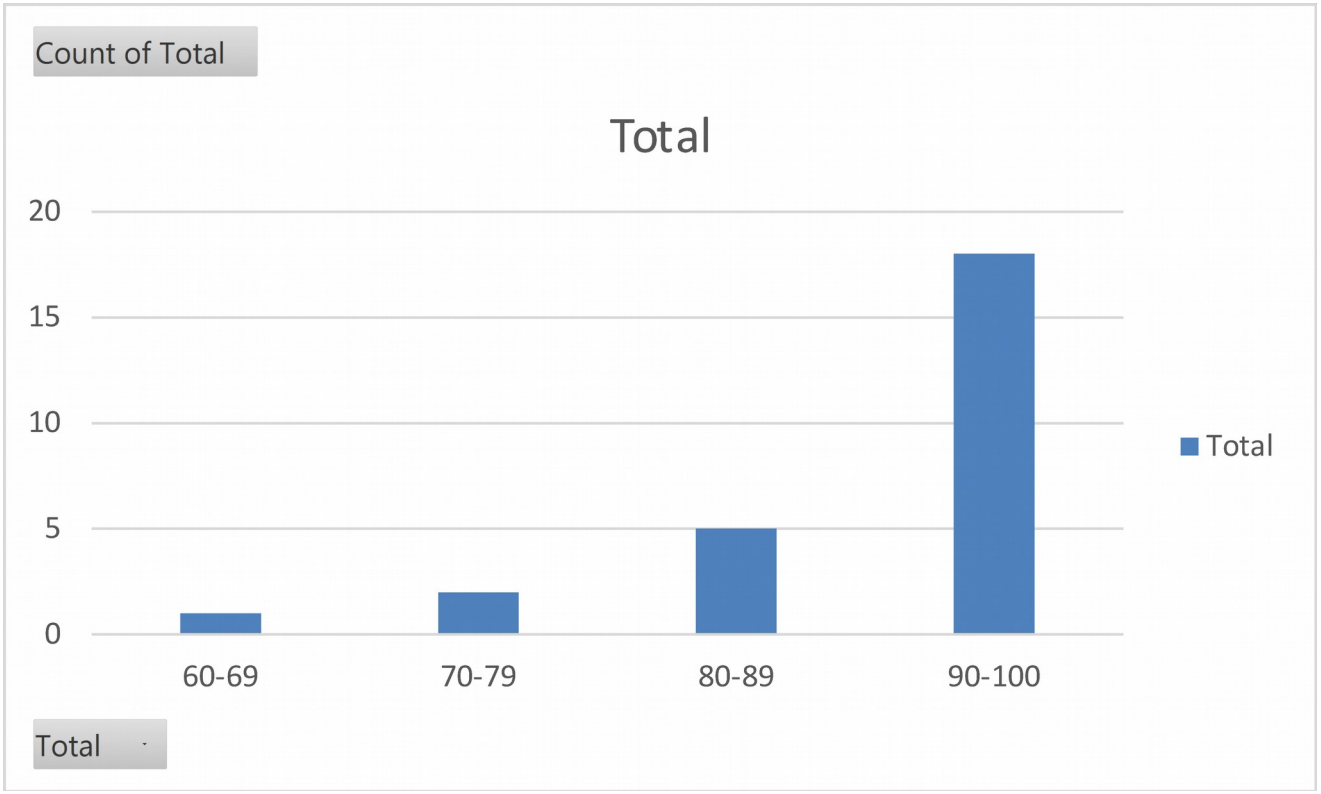
Midterm Grades!

Let's do some Data Science with the midterm...



Let's do some Data Science with the midterm...

Mean	90.38462
Std	9.344175
Median	93.5



Learning **objectives** for this visualization class.

- Visualization
  - Definition
  - Tasks
  - Taxonomy
  - A selection of techniques
  - Interactivity!

We do not have the ability to do justice to the topic in the time we have...

- However, there are many sources to look at for more information.
  - Interactive Data Visualization: Foundations, Techniques, Applications
    - Matthew Ward, Georges Grinstein, Daniel Keim
  - The Visual Display of Quantitative Information
    - Edward Tufte
  - Information Visualization: An Introduction
    - Robert Spence
  - Many more!

Our focus here...

- We will follow the lead of the famous paper
  - Shneiderman, Ben. "The eyes have it: A task by data type taxonomy for information visualizations." In Visual Languages, 1996. Proceedings., IEEE Symposium on, pp. 336-343. IEEE, 1996.
  - Cited almost 3000 times (based upon Google Scholar)

Visualization

- During the class we have actually done a lot of it already!
- However, how would you define it?

Turn data into Graphs

Better understanding

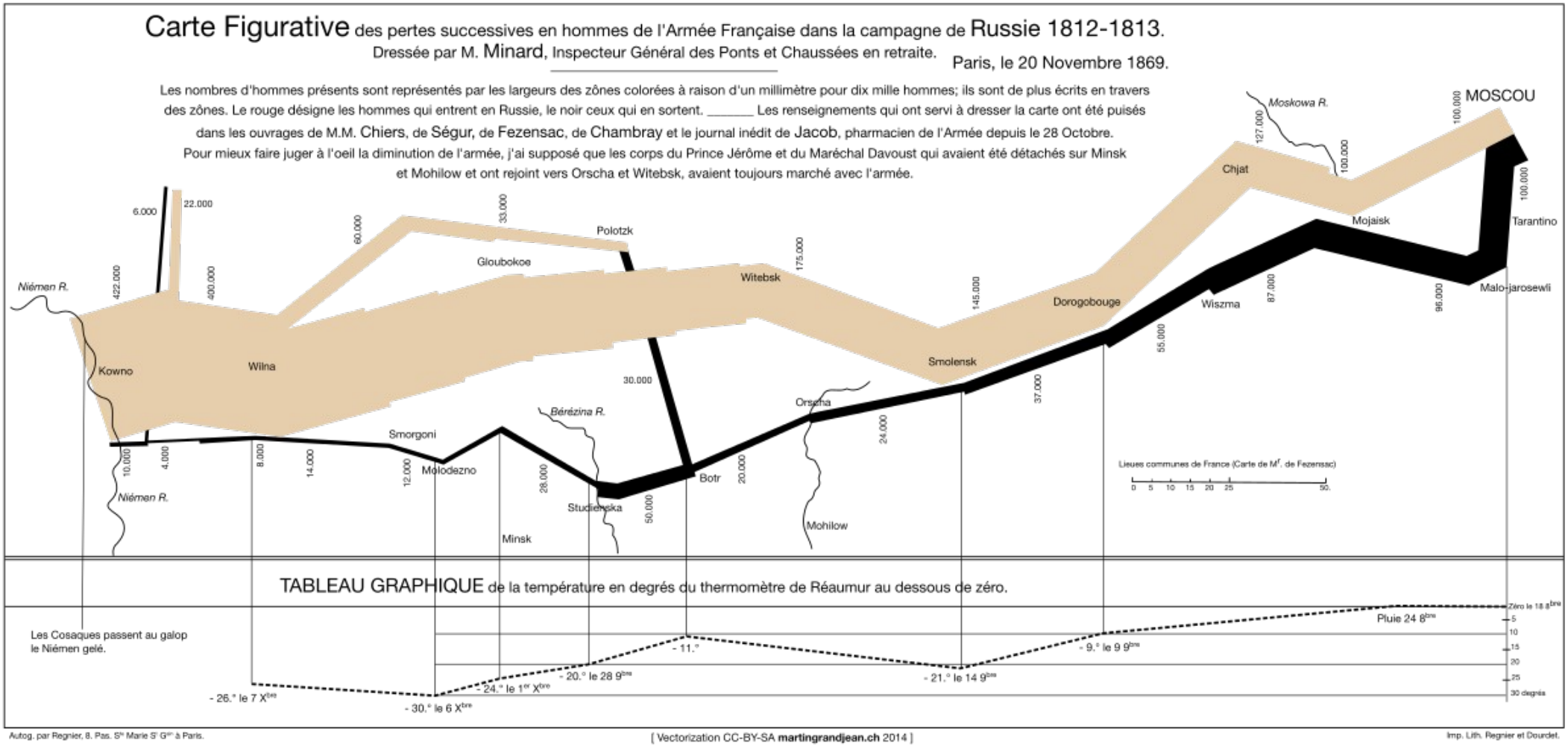
Cool!

Complex ideas => understandable

STORY TELLING

Visualization

- Visualization: The activity of forming a mental model of something
  - R. Spence, Information Visualization, DOI 10.1007/978-3-319-07341-5_1



The numbers of men present are represented by the widths of the colored zones at a rate of one millimeter for every ten-thousand men; they are further written across the zones. The red [now brown] designates the men who enter into Russia, the black those who leave it.

"Minard's Map (vectorized)" by Calvinius - Own work http://www.martingrandjean.ch/historical-data-visualization-minard-map/. Licensed under Creative Commons Attribution-Share Alike 3.0 via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:Minard%27s_Map_(vectorized).svg#mediaviewer/File:Minard%27s_Map_(vectorized).svg

7 Tasks

- Overview: Gain an overview of the entire collection.

Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. Proceedings of IEEE Symposium on Visual Languages - Boulder, CO (pp. 336-343).

7 Tasks

- Overview: Gain an overview of the entire collection.
- Zoom : Zoom in on items of interest.

Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. Proceedings of IEEE Symposium on Visual Languages - Boulder, CO (pp. 336-343).

7 Tasks

- Overview: Gain an overview of the entire collection.
- Zoom : Zoom in on items of interest.
- Filter: filter out uninteresting items.

Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. Proceedings of IEEE Symposium on Visual Languages - Boulder, CO (pp. 336-343).

7 Tasks

- Overview: Gain an overview of the entire collection.
- Zoom : Zoom in on items of interest.
- Filter: filter out uninteresting items.
- Details-on-demand: Select an item or group and get details when needed.

Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. Proceedings of IEEE Symposium on Visual Languages - Boulder, CO (pp. 336-343).

7 Tasks

- Overview: Gain an overview of the entire collection.
- Zoom : Zoom in on items of interest.
- Filter: filter out uninteresting items.
- Details-on-demand: Select an item or group and get details when needed.
- Relate: View relations hips among items.

Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. Proceedings of IEEE Symposium on Visual Languages - Boulder, CO (pp. 336-343).

7 Tasks

- Overview: Gain an overview of the entire collection.
- Zoom : Zoom in on items of interest.
- Filter: filter out uninteresting items.
- Details-on-demand: Select an item or group and get details when needed.
- Relate: View relations hips among items.
- History: Keep a history of actions to support undo, replay, and progressive refinement.

Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. Proceedings of IEEE Symposium on Visual Languages - Boulder, CO (pp. 336-343).

7 Tasks

- Overview: Gain an overview of the entire collection.
- Zoom : Zoom in on items of interest.
- Filter: filter out uninteresting items.
- Details-on-demand: Select an item or group and get details when needed.
- Relate: View relations hips among items.
- History: Keep a history of actions to support undo, replay, and progressive refinement.
- Extract: Allow extraction of sub-collections and of the query parameters.

Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. Proceedings of IEEE Symposium on Visual Languages - Boulder, CO (pp. 336-343).

Visualization Taxonomy

- 1D/Linear
- 2D/Planar (incl. Geospatial)
- 3D/Volumetric
- Temporal
- nD/Multidimensional
- Tree/Hierarchical
- Network

Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. Proceedings of IEEE Symposium on Visual Languages - Boulder, CO (pp. 336-343) and http://guides.library.duke.edu/vis_types.

Tools

- Python
  - Matplotlib – We all know and love this!
  - Mayavi – We have used this too!
  - NetworkX – Graph processing and visualization.
  - Pyglet – Cross domain multimedia library.
  - Vispy – High Performance OpenGL based interactive visualization library.
  - Bokeh – Python interactive visualization in the style of D3.js.
  - plot.ly – Online plotting and data processing.

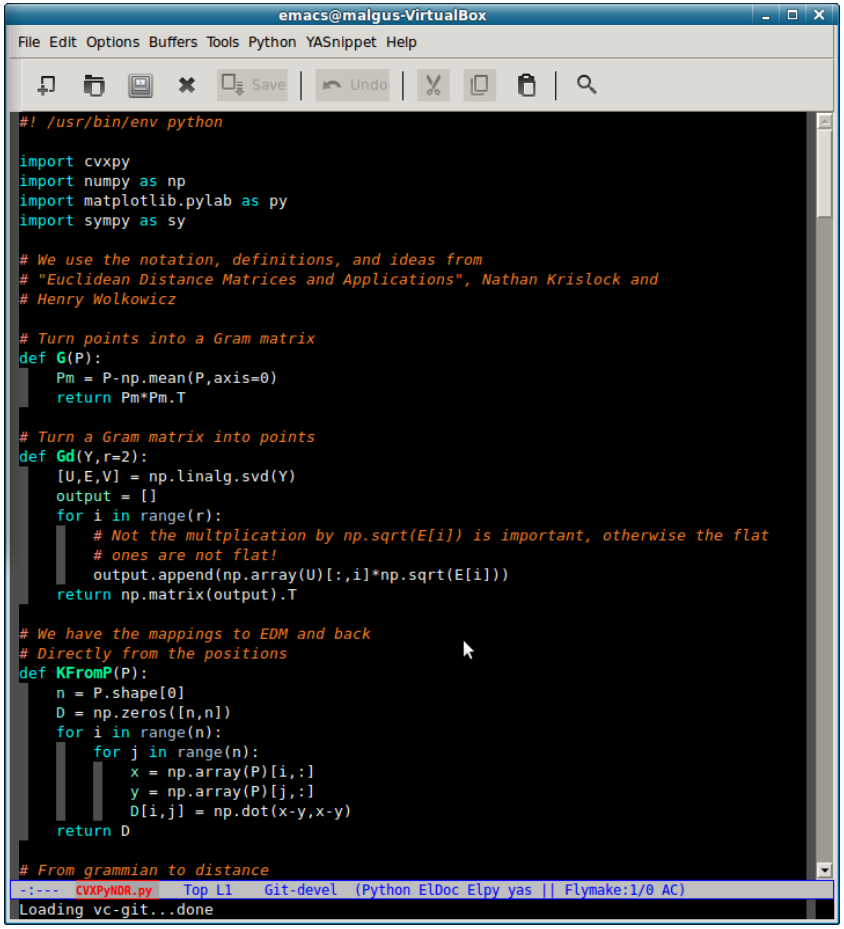
Tools

- Web and Javascript!
  - D3.js – Data Driven Documents
    - NVD3.js – Higher level components
    - Cubism – Focused on time series analysis
    - Rickshaw – Higher level components
    - Many others!
  - Webgl
    - Three.js

1-dimensional

- Purely 1-dimensional data
  - Text documents
  - Program source code
  - Etc.
- Is there anything interesting here?

Yes!


emacs@malgus-VirtualBox
File Edit Options Buffers Tools Python YASnippet Help
Save
Undo
#! /usr/bin/env python
import cvxpy
import numpy as np
import matplotlib.pylab as py
import sympy as sy
# We use the notation, definitions, and ideas from
# "Euclidean Distance Matrices and Applications", Nathan Krislock and
# Henry Wolkowicz
# Turn points into a Gram matrix
def G(P):
Pm = P-np.mean(P,axis=0)
return Pm*Pm.T
# Turn a Gram matrix into points
def Gd(Y,r=2):
[U,E,V] = np.linalg.svd(Y)
output = []
for i in range(r):
# Not the multplication by np.sqrt(E[i]) is important, otherwise the flat
# ones are not flat!
output.append(np.array(U)[:,i]*np.sqrt(E[i]))
return np.matrix(output).T
# We have the mappings to EDM and back
# Directly from the positions
def KFromP(P):
n = P.shape[0]
D = np.zeros([n,n])
for i in range(n):
for j in range(n):
x = np.array(P)[i,:]
y = np.array(P)[j,:]
D[i,j] = np.dot(x-y,x-y)
return D
From grammian to distance
-:--- CVXPyNDR.py Top L1 Git-devel (Python ElDoc Elpy yas || Flymake:1/0 AC)
Loading vc-git...done

2-dimensional

- Maps
- Floorplans
- Etc.

Choropleth: Shaded or patterned map

http://bl.ocks.org/mbostock/4060606

WebGL: World population

- http://data-arts.appspot.com/globe/

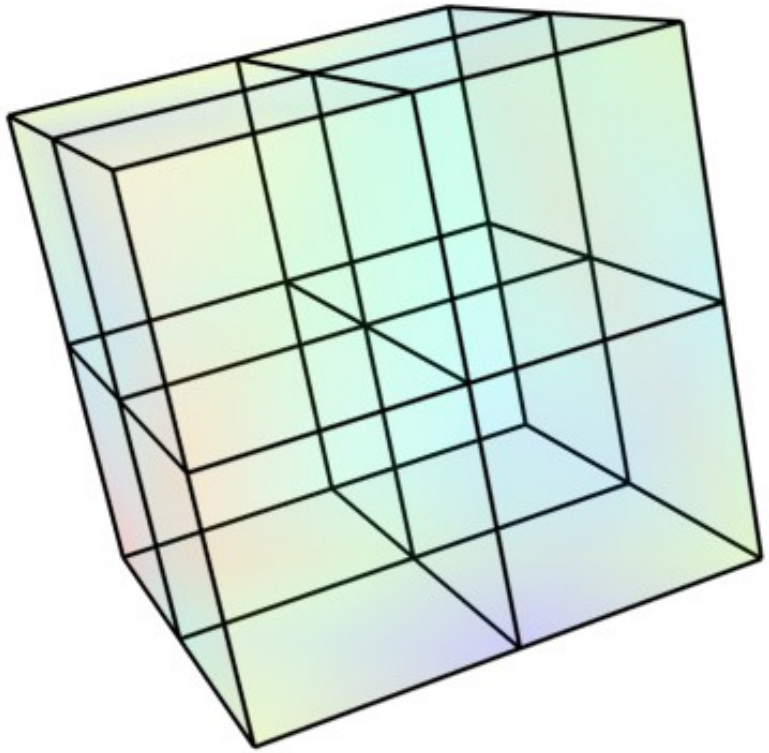
World Bank

http://d3.artzub.com/wbca/

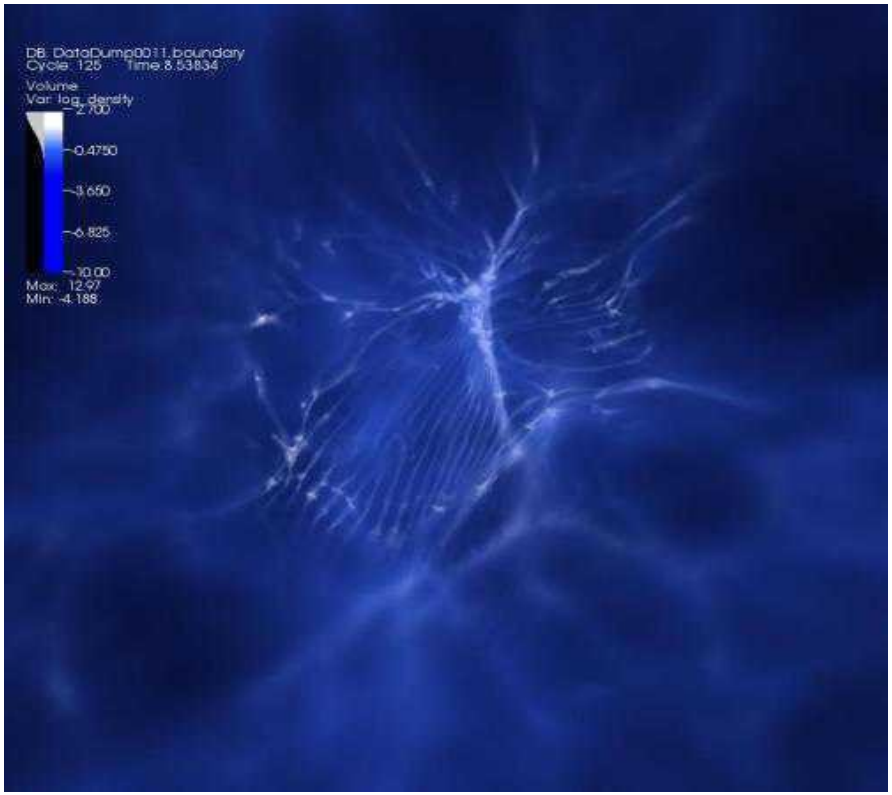
3-dimensional

- Real world objects
  - Buildings
  - Molecules
  - Vehicles
  - Etc.

Data types



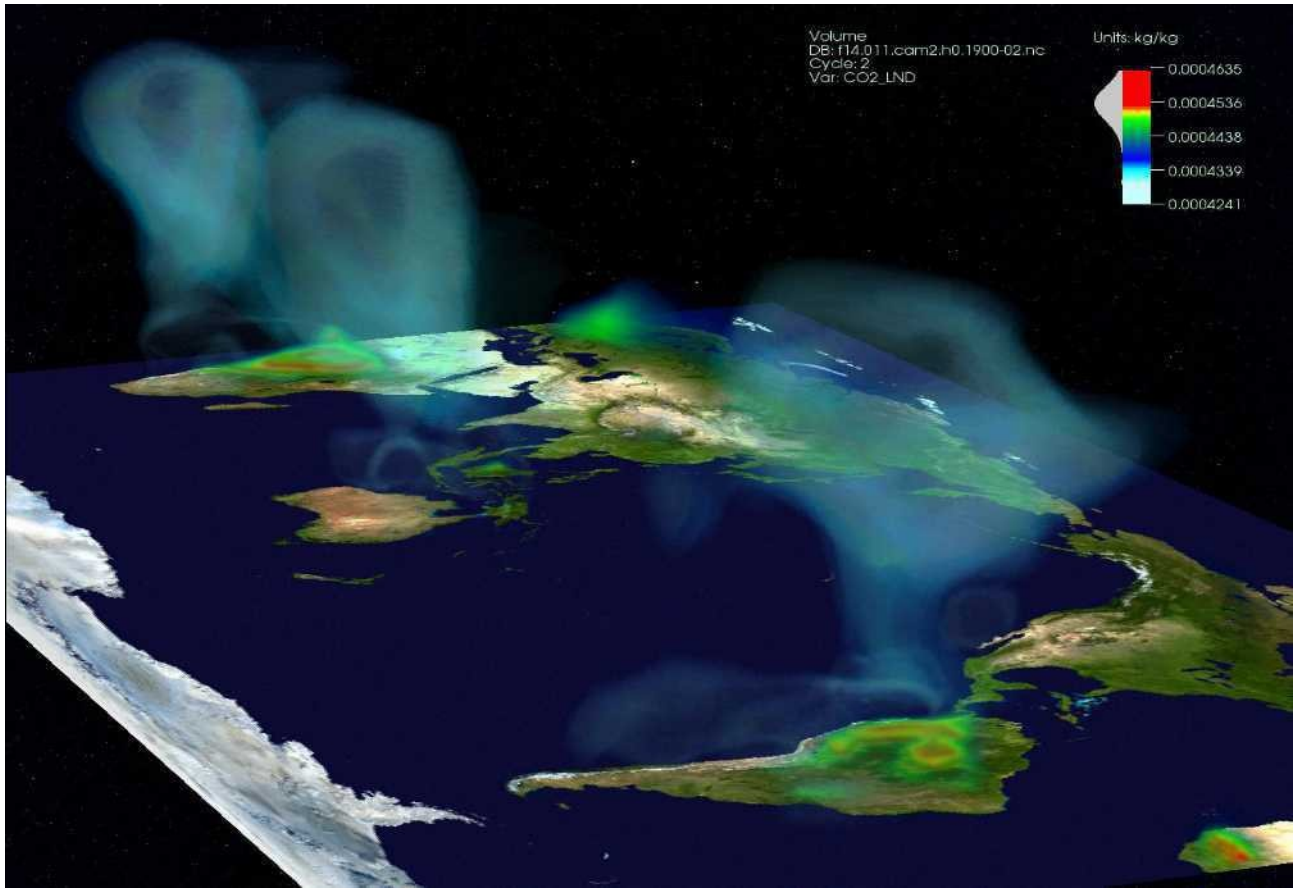
Star formation



Density shown in white-blue and transparency

"Star formation" by UCRL - Visualizations that have been created with VisIt. at wci.llnl.gov. Licensed under Public domain via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:Star_formation.jpg#mediaviewer/File:Star_formation.jpg

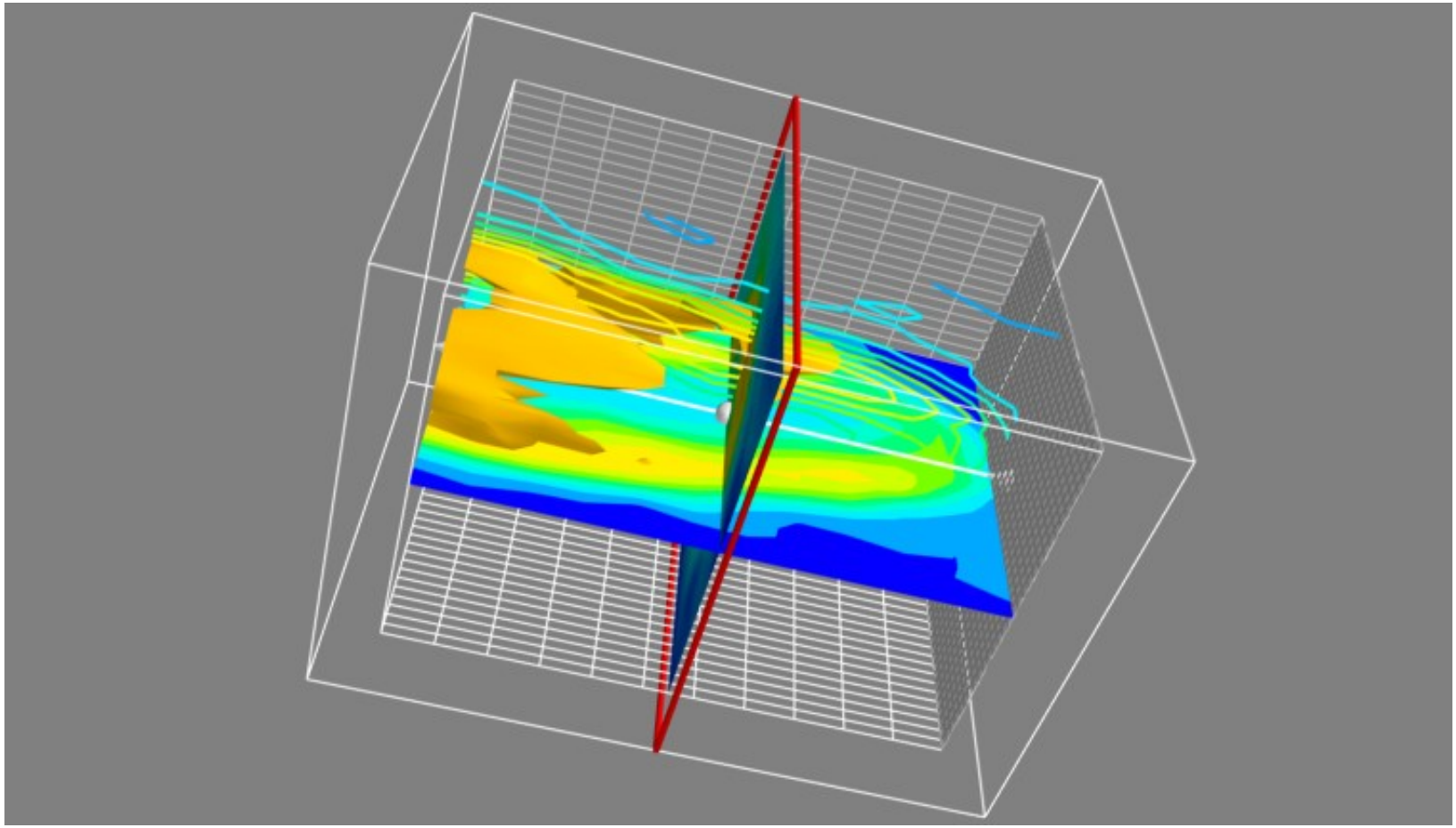
Climate



"Climate visualization" by UCRL and Forrest Hoffman and Jamison Daniel of Oak Ridge National Laboratory - Visualizations that have been created with VisIt. at wci.llnl.gov. Licensed under Public domain via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:Climate_visualization.jpg#mediaviewer/File:Climate_visualization.jpg

CO2 from ocean plumes.

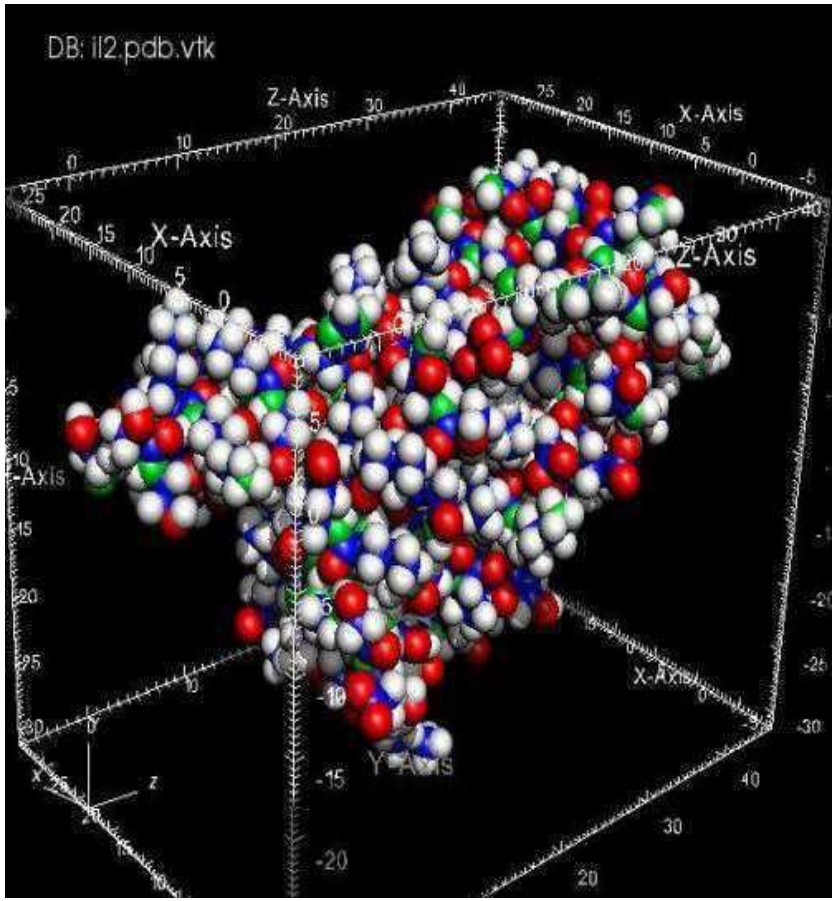
Contours



Python 2

Python 3

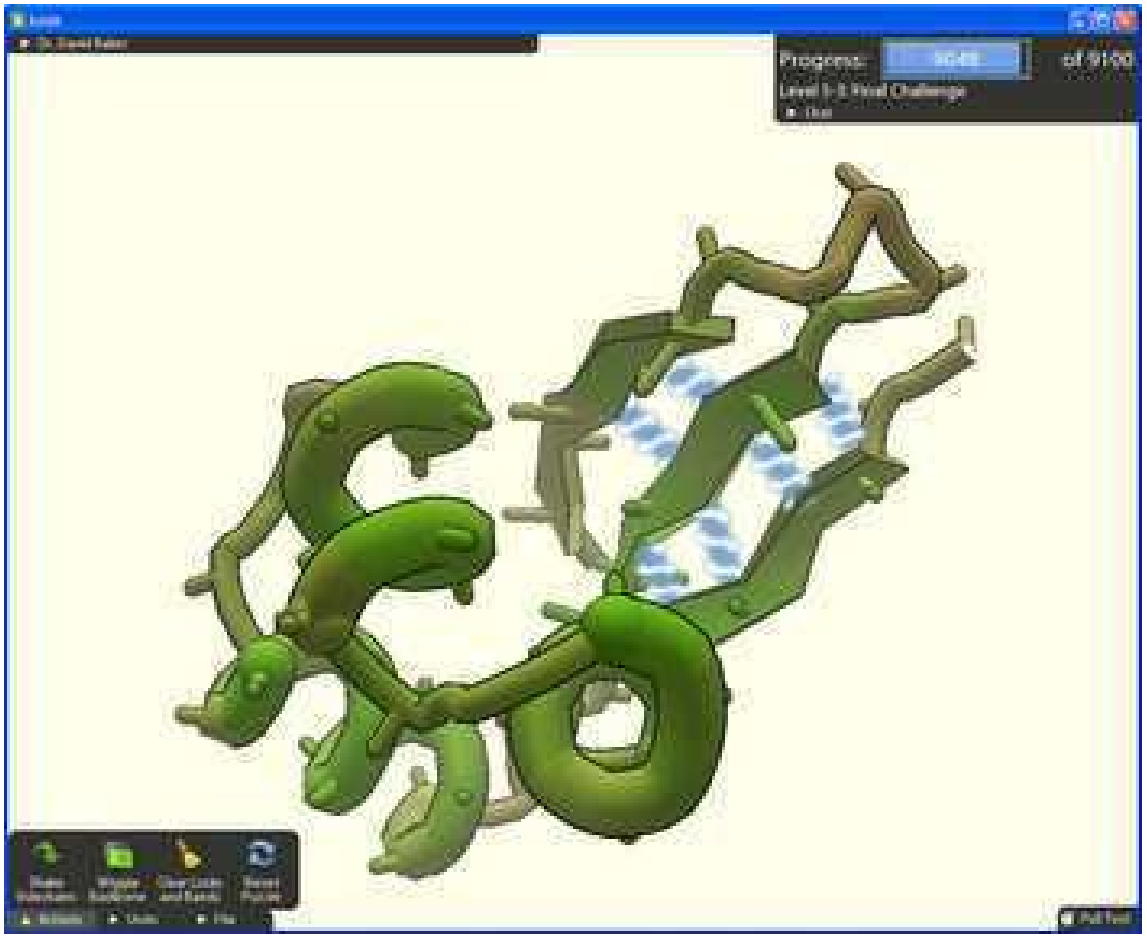
Molecules



Protein

"Molecular rendering" by UCRL-WEB - visualizations that have been created with VisIt. at wci.llnl.gov. Licensed under Public domain via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:Molecular_rendering.jpg#mediaviewer/File:Molecular_rendering.jpg

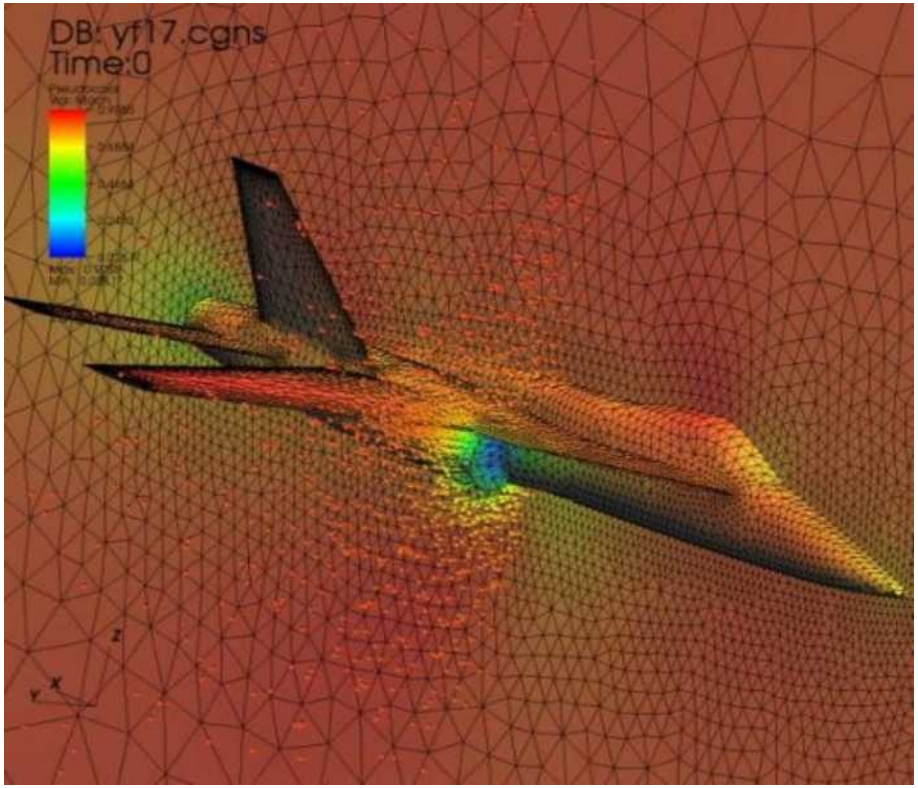
fold.it



A game on protein folding!

"Foldit". Via Wikipedia - http://en.wikipedia.org/wiki/File:Foldit.png#mediaviewer/File:Foldit.png

Finite Element Simulation



Mach number

"YF-17 aircraft Plot" by The dataset was provided by the CGNS user community's - Visualizations that have been created with VisIt. at wci.llnl.gov. Licensed under Public domain via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:YF-17_aircraft_Plot.jpg#mediaviewer/File:YF-17_aircraft_Plot.jpg

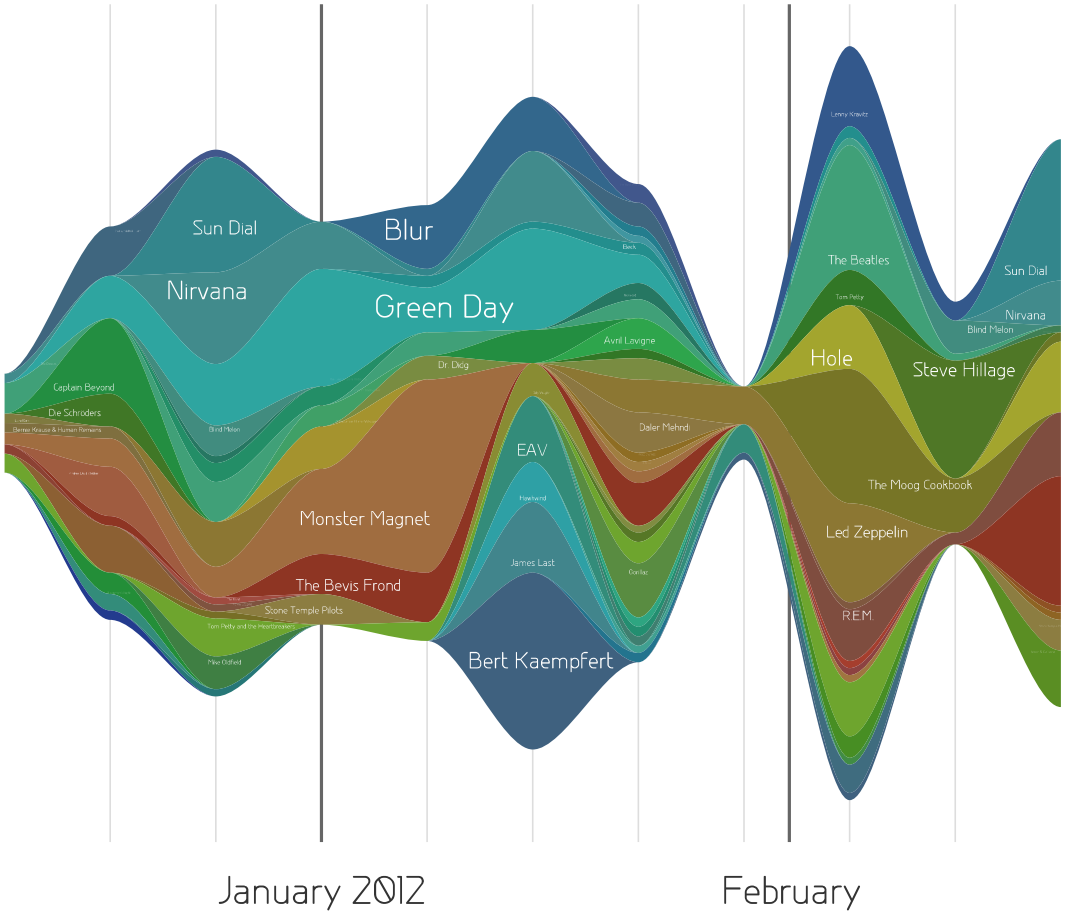
Temporal

- Data with start and finish times
  - Historical events
  - Medical records
  - Marketing
  - Etc.

Cubism

http://square.github.io/cubism/

Streamgraph



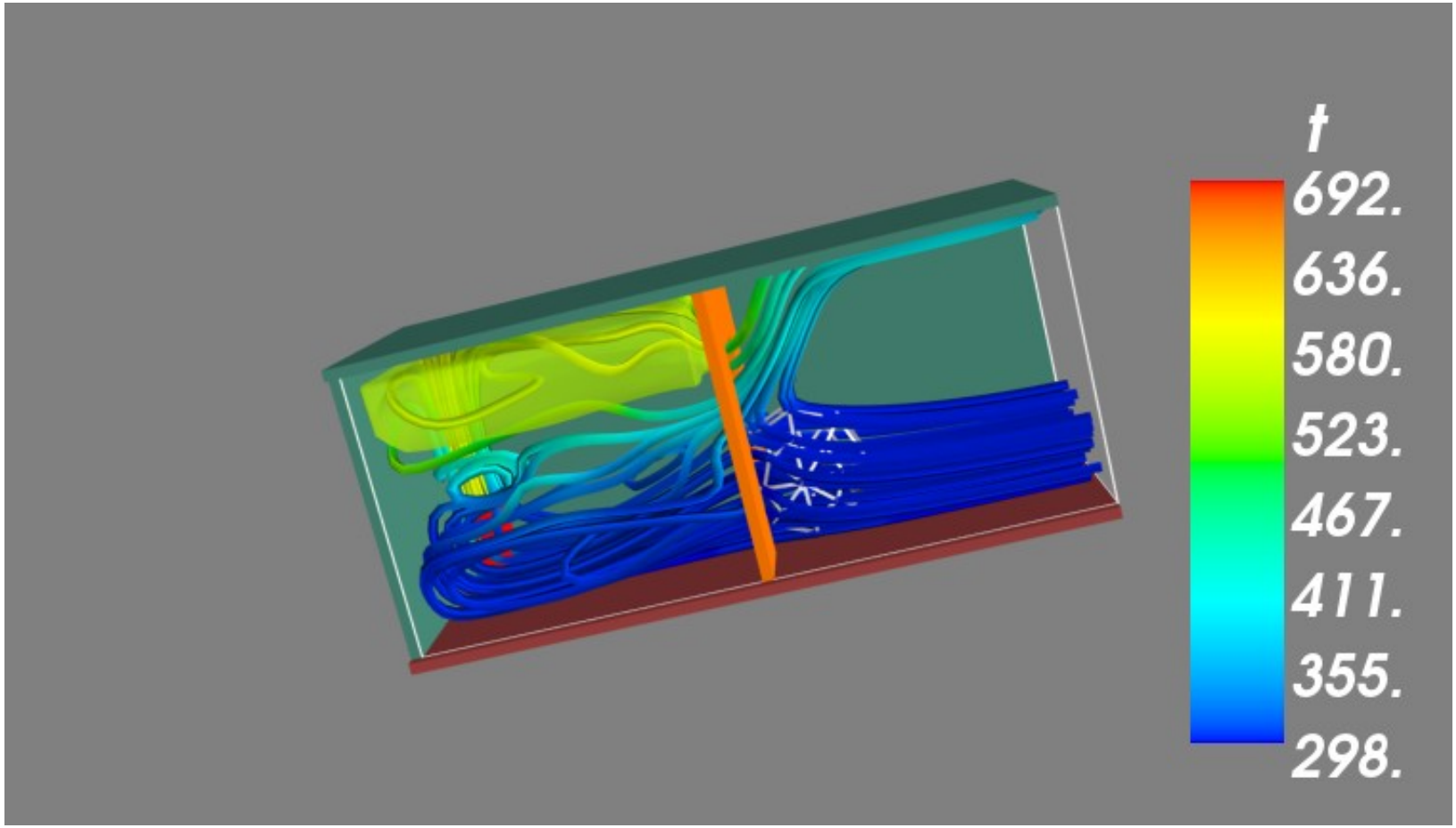
Last.fm person's listening habits.

"LastGraph example" by Psychonaut - Own work. Licensed under Creative Commons Zero, Public Domain Dedication via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:LastGraph_example.svg#mediaviewer/File:LastGraph_example.svg

Streamgraph

- http://bl.ocks.org/WillTurman/4631136

Streams



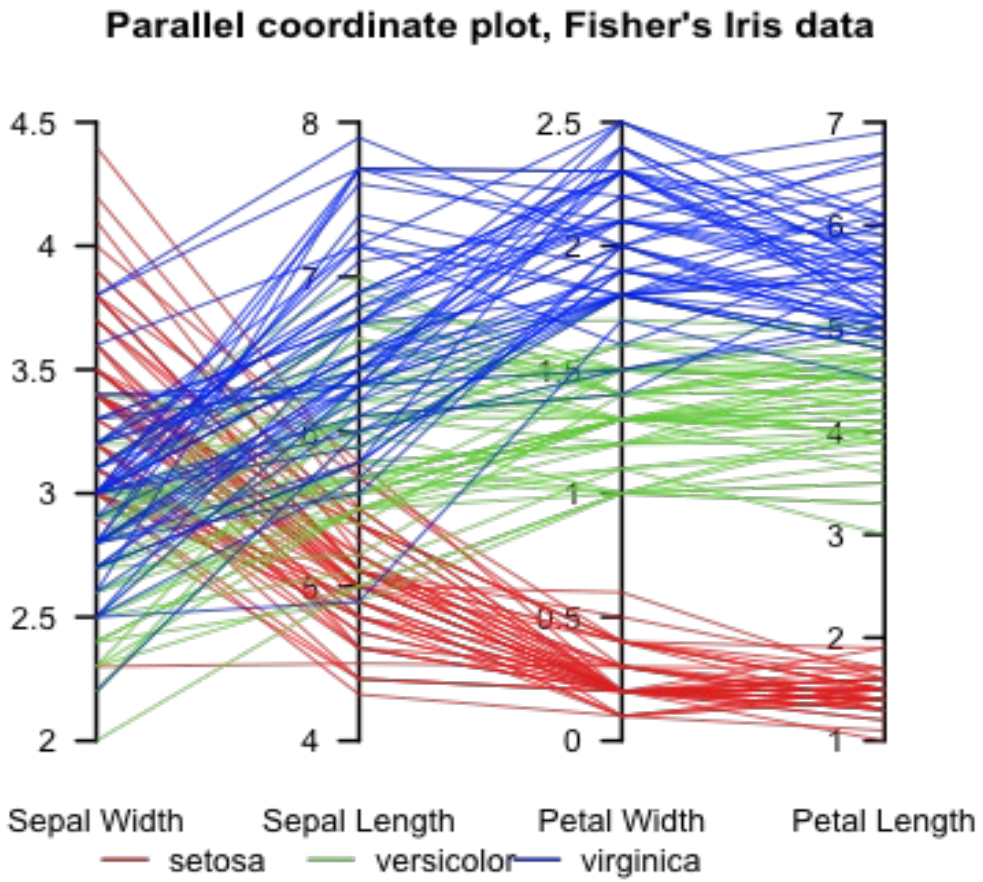
WebGL: Temperature anomalies

- http://www.chromeexperiments.com/detail/temperature-anomalies-200-years/?f=webgl

Multi-dimensional

- As far as Data Science is concerned, this is really where the things get very interesting!
  - Can you give me some examples?

Parallel coordinates



Iris data set!

"ParCorFisherIris". Licensed under Public domain via Wikipedia - http://en.wikipedia.org/wiki/File:ParCorFisherIris.png#mediaviewer/File:ParCorFisherIris.png

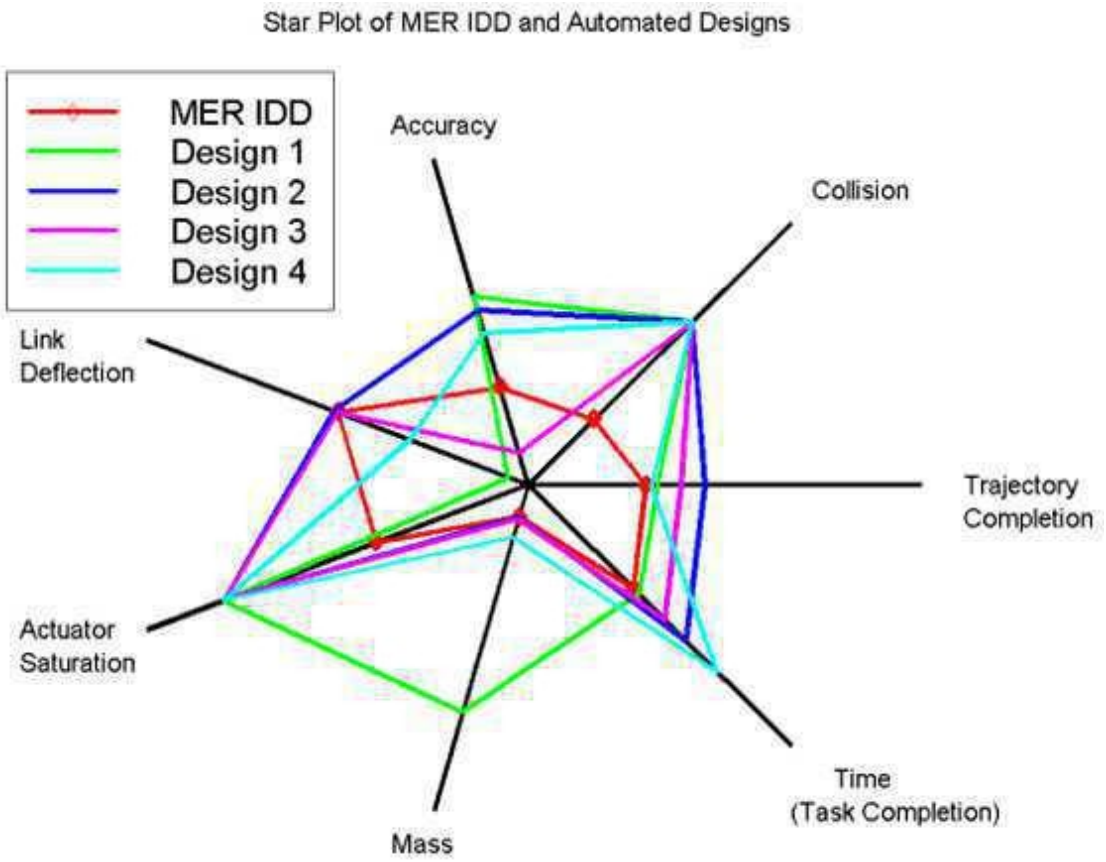
Parallel coordinates:

- http://mbostock.github.io/d3/talk/20111116/iris-parallel.html

Parallel coordinates

- http://exposedata.com/parallel/

Radar Chart



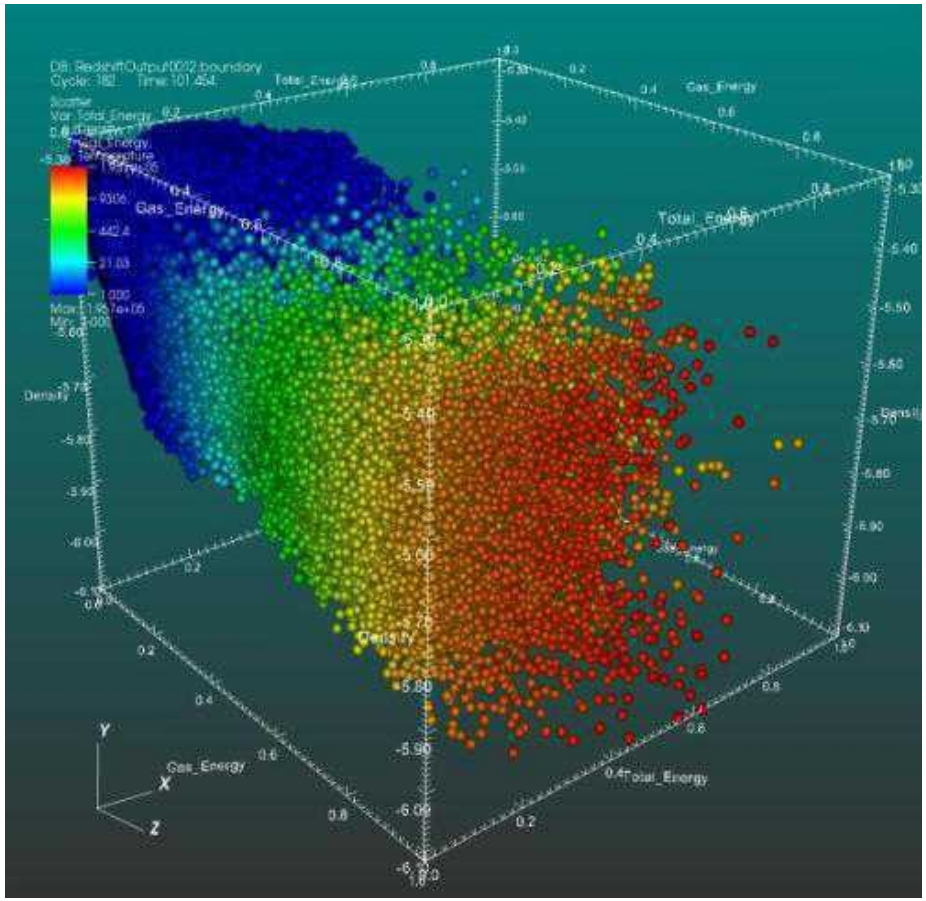
NASA Space system design: Desirable is in the middle.

"MER Star Plot" by NASA Primary START - Automation Tool for Rapid Design of Space Systems. Licensed under Public domain via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:MER_Star_Plot.gif#mediaviewer/File:MER_Star_Plot.gif

Radar Chart

- http://graves.cl/radar-chart-d3/

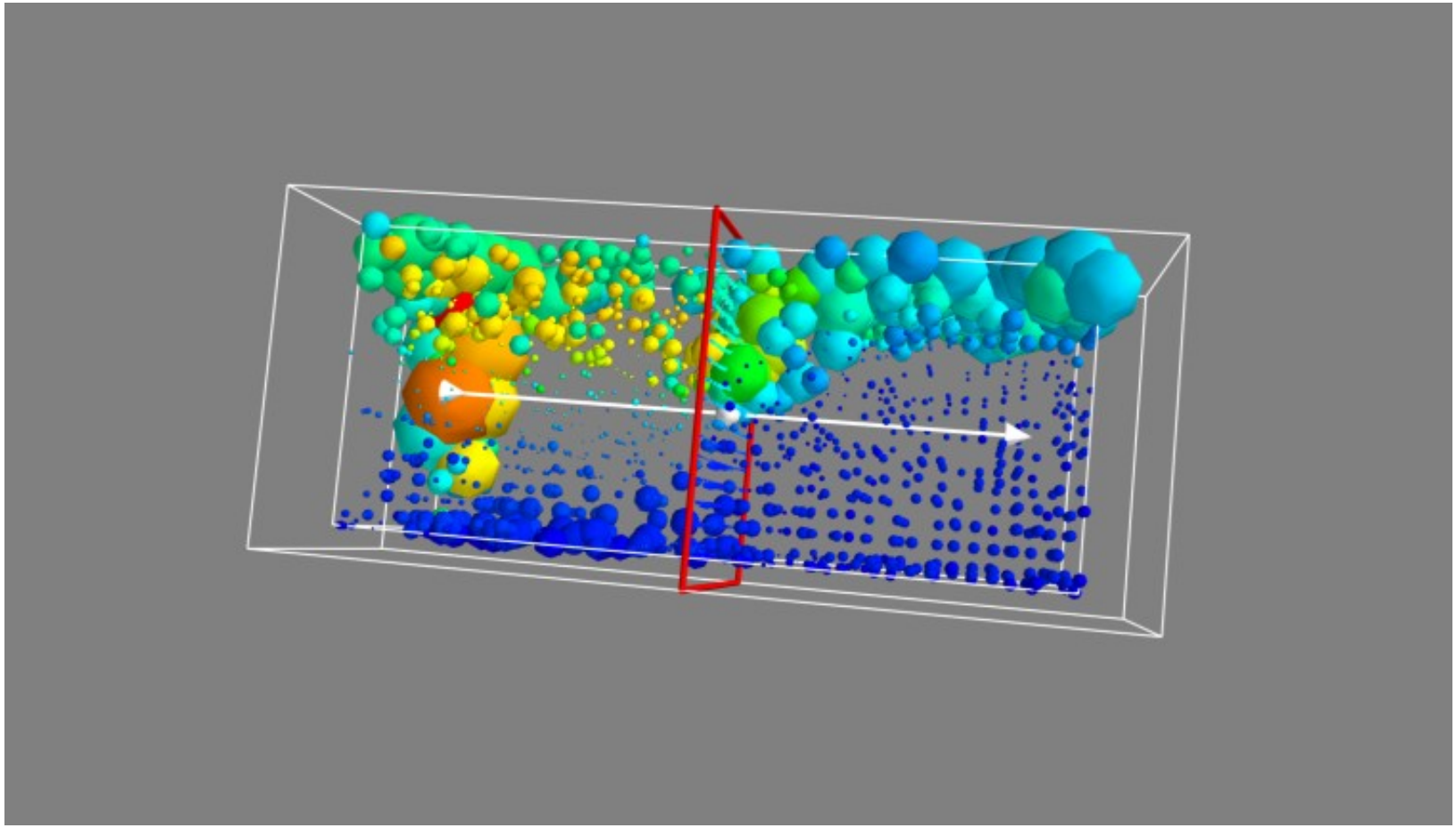
Scatter plot



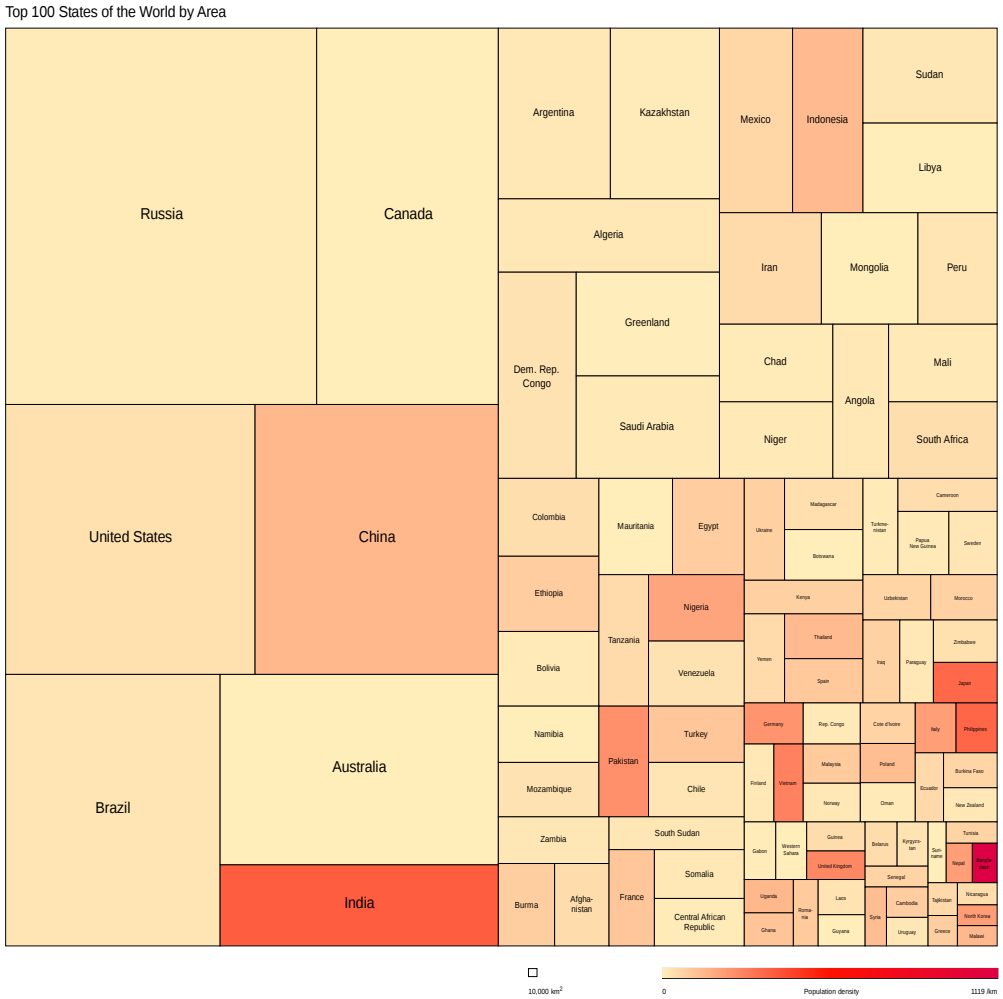
4 dimensions per point

"Scatter plot" by UCRL - Visualizations that have been created with VisIt. at wci.llnl.gov. Licensed under Public domain via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:Scatter_plot.jpg#mediaviewer/File:Scatter_plot.jpg

Glyphs



Tree chart



Size is area, color is density.

"Top100 states area treemap pop-density" by Own work - Data: CIA World Factbook (visited 2012-09-12);Treemapping: Google Chart Tools visualization API. Licensed under Public domain via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:Top100_states_area_treemap_pop-density.svg#mediaviewer/File:Top100_states_area_treemap_pop-density.svg

Tree chart

- http://www.billdwhite.com/wordpress/2012/12/16/d3-treemap-with-title-headers/

Bubble Chart

http://bl.ocks.org/mbostock/4063269

Word cloud

http://www.jasondavies.com/wordcloud/#%2F%2Fwww.jasondavies.com%2Fwordtree%2Fcat-in-the-hat.txt

Tree or hierarchical

- Items that fall into groups that can be further sub-divided.

Federal budget

http://www.brightpointinc.com/interactive/budget/index.html?source=d3js

Radial layout

http://mbostock.github.io/d3/talk/20111018/cluster.html

Network

- Graphs
- Social networks
- Computer networks

Social Network

League of nations personel

"Social Network Analysis Visualization" by Calvinius - Own work : http://www.martingrandjean.ch/wp-content/uploads/2013/10/Graphe3.png. Licensed under Creative Commons Attribution-Share Alike 3.0 via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:Social_Network_Analysis_Visualization.png#mediaviewer/File:Social_Network_Analysis_Visualization.png

Matrix representation

http://bokeh.pydata.org/docs/gallery/les_mis.html

Network

- http://fatiherikli.github.io/programming-language-network/

Network

- http://mbostock.github.io/d3/talk/20111116/force-collapsible.html

Chord Diagram

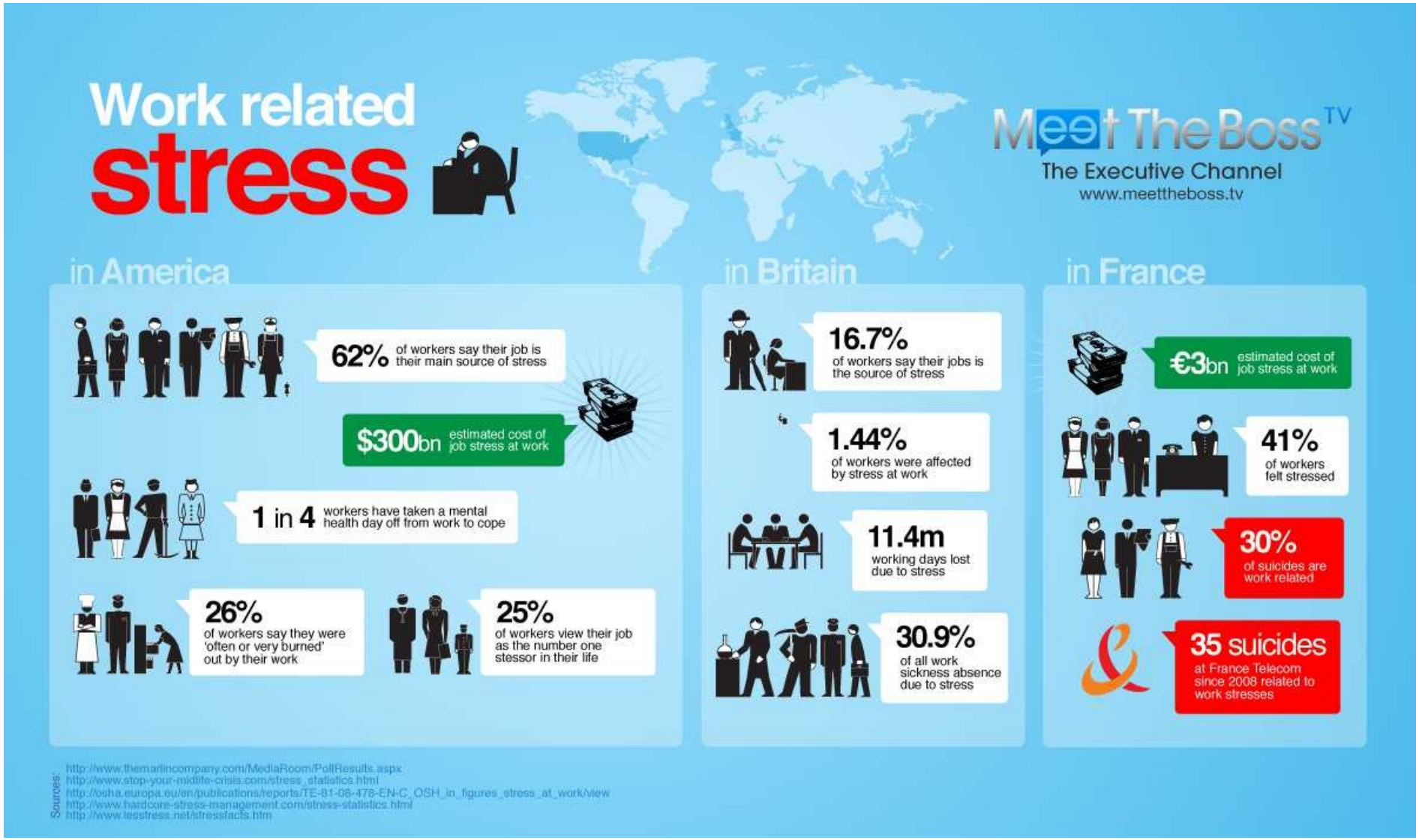
http://bost.ocks.org/mike/uberdata/

Congressional influence

http://www.brightpointinc.com/interactive/political_influence/index.html?source=d3js

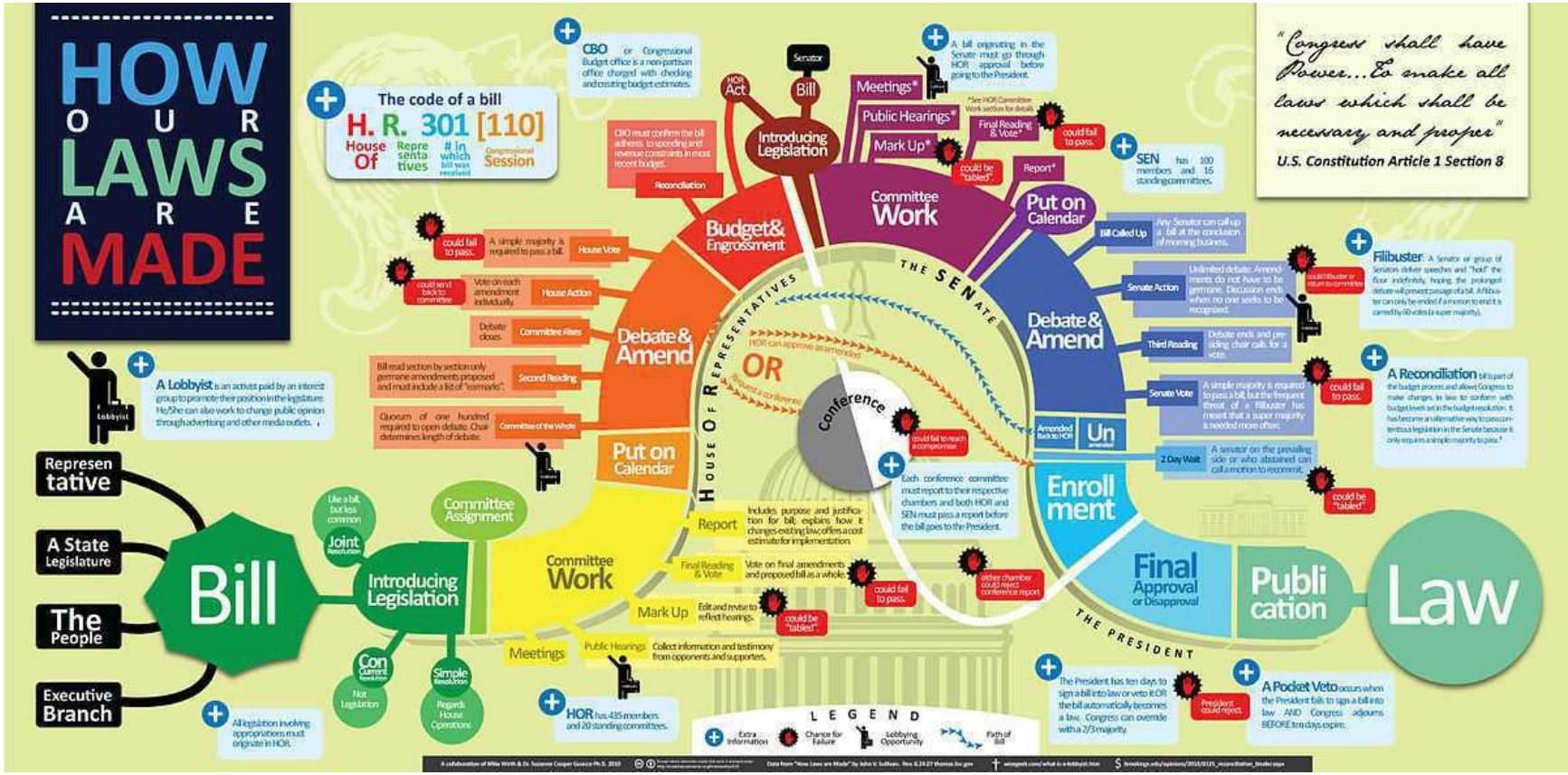
Infographics – Not in Shneiderman

Infographic



https://www.flickr.com/photos/gdsdigital/4514262548/

Infographics



"Visualization-of-How-a-Bill-Becomes-a-Law Mike-WIRTH" by Mike Wirth and Dr. Suzanne Cooper-Guasco - http://www.ritholtz.com/blog/wp-content/uploads/2010/07/howlawsmadeWIRTH2.jpg. Licensed under Creative Commons Attribution 3.0 via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:Visualization-of-How-a-Bill-Becomes-a-Law_Mike-WIRTH.jpg#mediaviewer/File:Visualization-of-How-a-Bill-Becomes-a-Law_Mike-WIRTH.jpg

Backup