

Introduction to Data Visualization

Marci Brandenburg

Justin Joque

Stephanie O'Malley

Outline

- ❖ Visualization Landscape
- ❖ Design Principles
- ❖ Interpretation
- ❖ Campus Resources

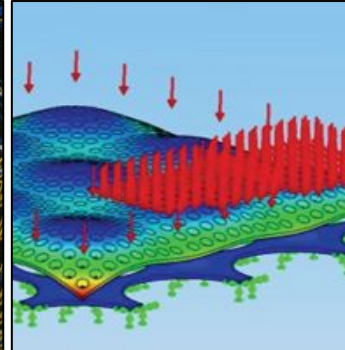
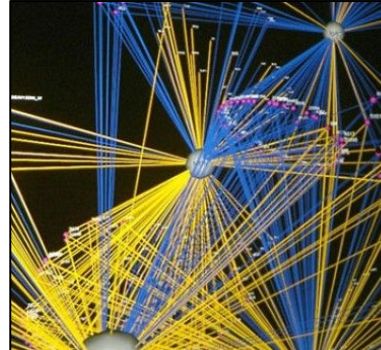
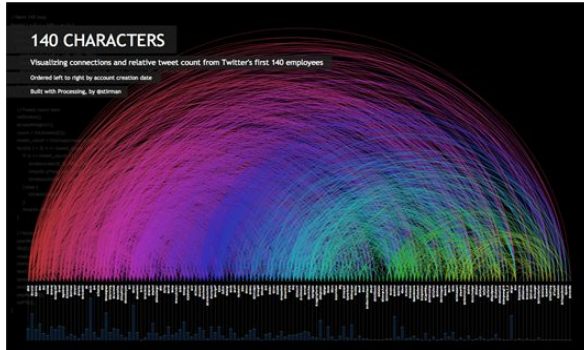
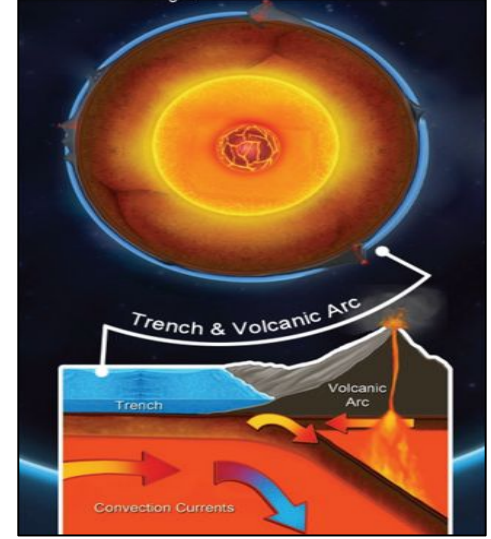
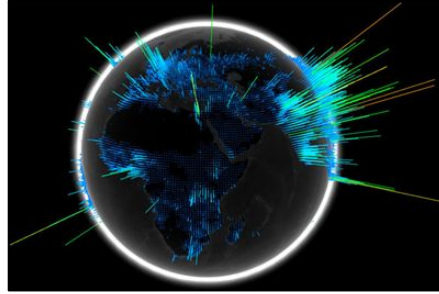
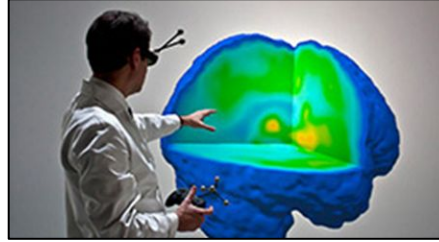
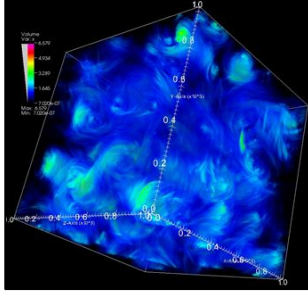
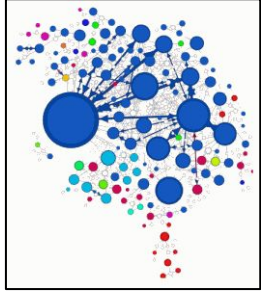
Visualization Landscape

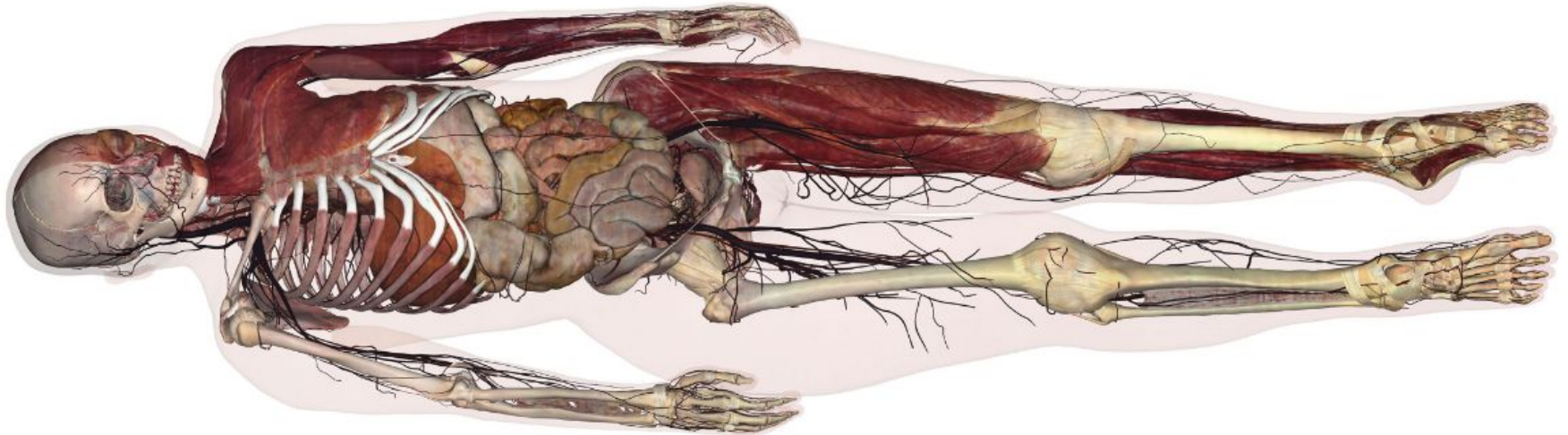
What is Visualization?

“The action or fact of visualizing;
the power or process of forming a mental picture or vision
of something not actually present to the sight;
a picture thus formed.”

-Oxford English Dictionary

Types of Visualization





Map based
example -

Cholera
Outbreak

By John
Snow

(not Game
of
Thrones)

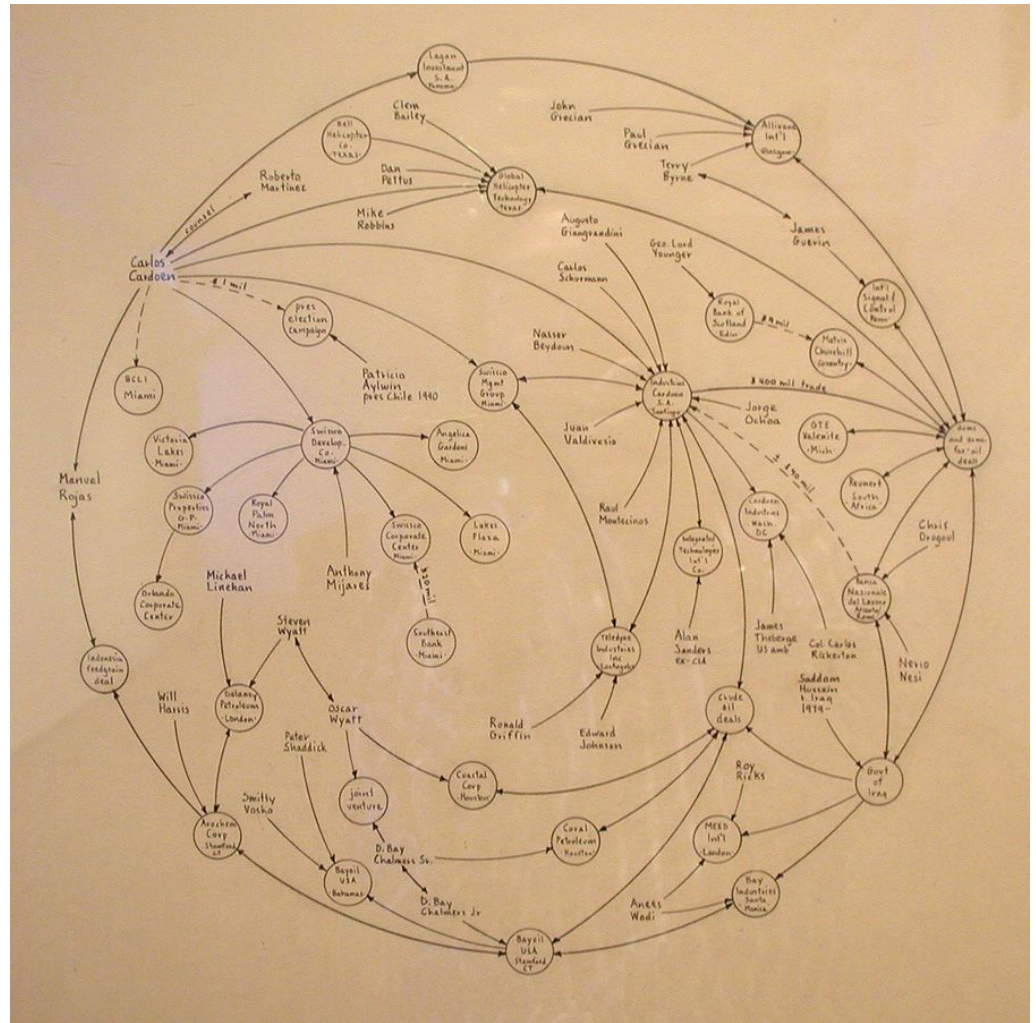


Snow, J. *Snow on Cholera being a Reprint of Two Papers*. London: Oxford University Press, 1936. Print.

(1951-2000) draws on the major political and financial scandals of the day to create large-scale linear diagrams that at first glance look like celestial maps....”

-University Art Museum.
University of Albany, State
University of New York.

<http://www.albany.edu/museum/wwwmuseum/work/lombardi/images/lombardi1.jpg>



wind map

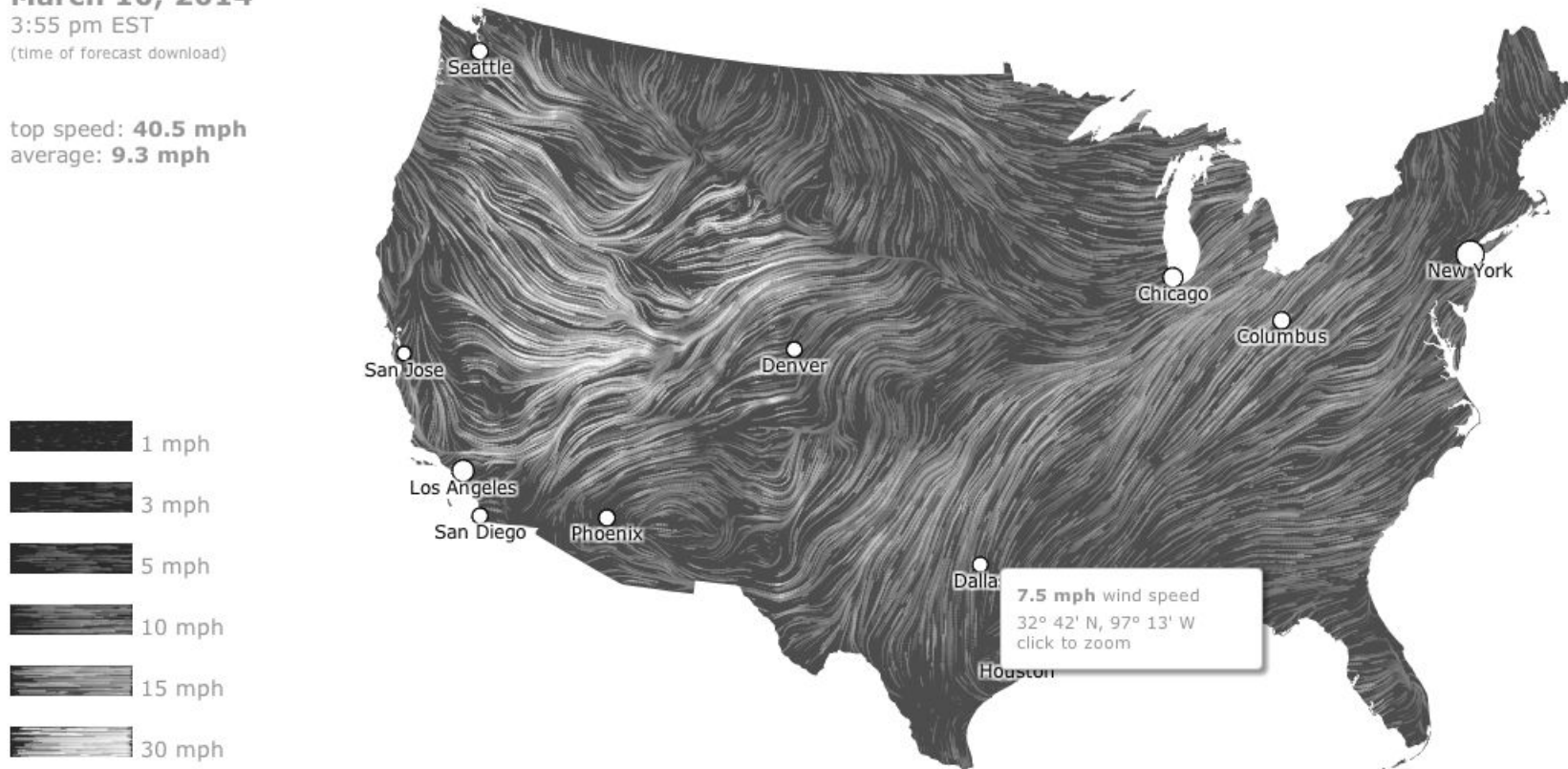
March 10, 2014

3:55 pm EST

(time of forecast download)

top speed: **40.5 mph**

average: **9.3 mph**

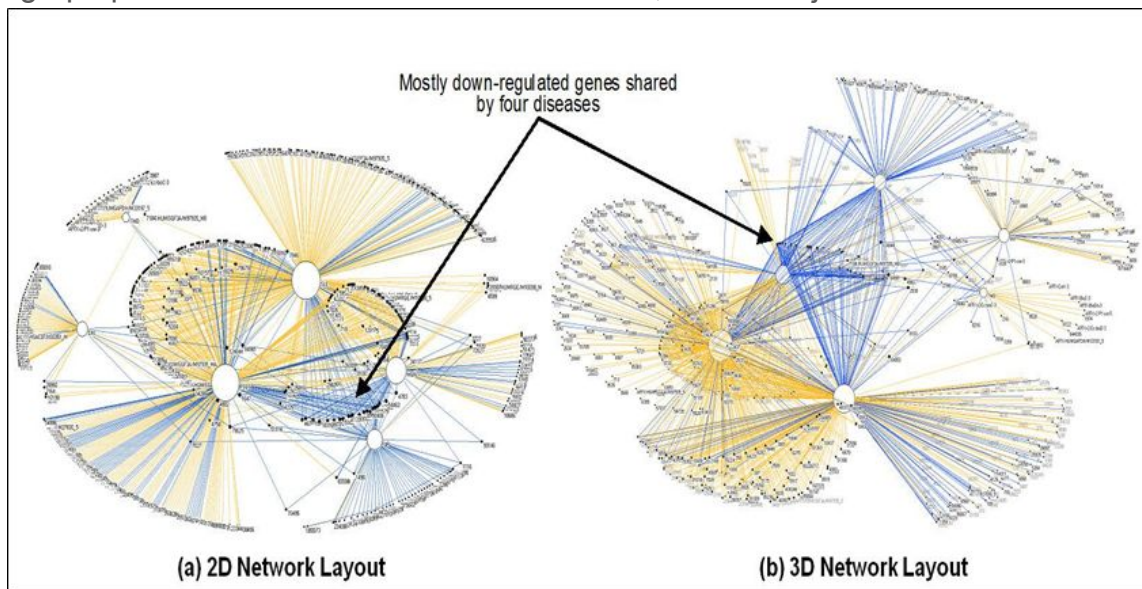


2D vs. 3D Visualization

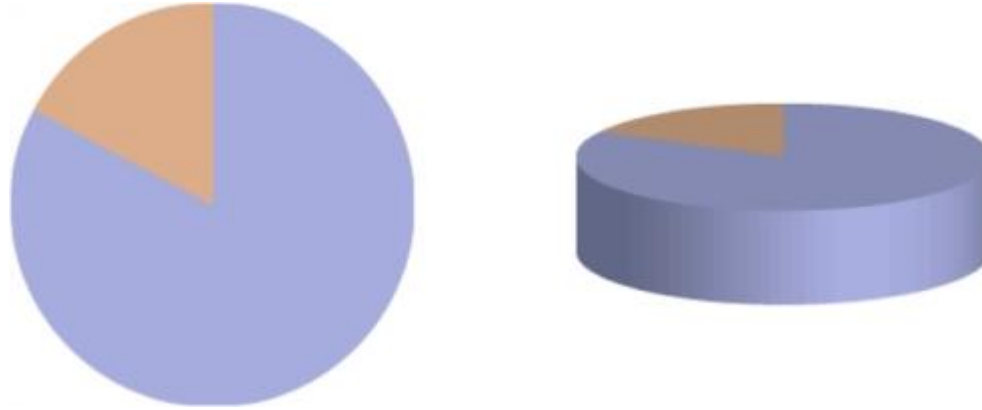
3D can untangle a graph: entire new dimension to place nodes.

Need to think about if our visualization would really benefit from 3D.

- ❖ Does it overcomplicate the visualization?
- ❖ Does the 3D graph provide more information to the viewer, or does it just look different?



Just because we can...



Andy Kirk. *Data visualization: a successful design process*. Birmingham: United Kingdom, 2012. eBook.

Data Visualization

Data visualization deals with communicating information about an existing data set through a visual medium.

Goals

- ❖ Take advantage of the brain's ability to *efficiently* process visuals. A *high volume* of information can often be easily understood through an image.
- ❖ Want the viewer to *learn* something about the data that could not be easily understood by reading the raw data.
- ❖ Visual is presented in a *simple* manner. Don't overload the viewer with many unnecessary details.

Reasons for Visualization

- ❖ Discovery
- ❖ Hypothesis generation
- ❖ Analysis
- ❖ Representation of data for publication or presentation - communication

Common Software/Tools

❖ Scripting Languages

- Python
- R
- Javascript
- PHP

❖ Network Visualization

- Cytoscape
- Gephi
- Pajek

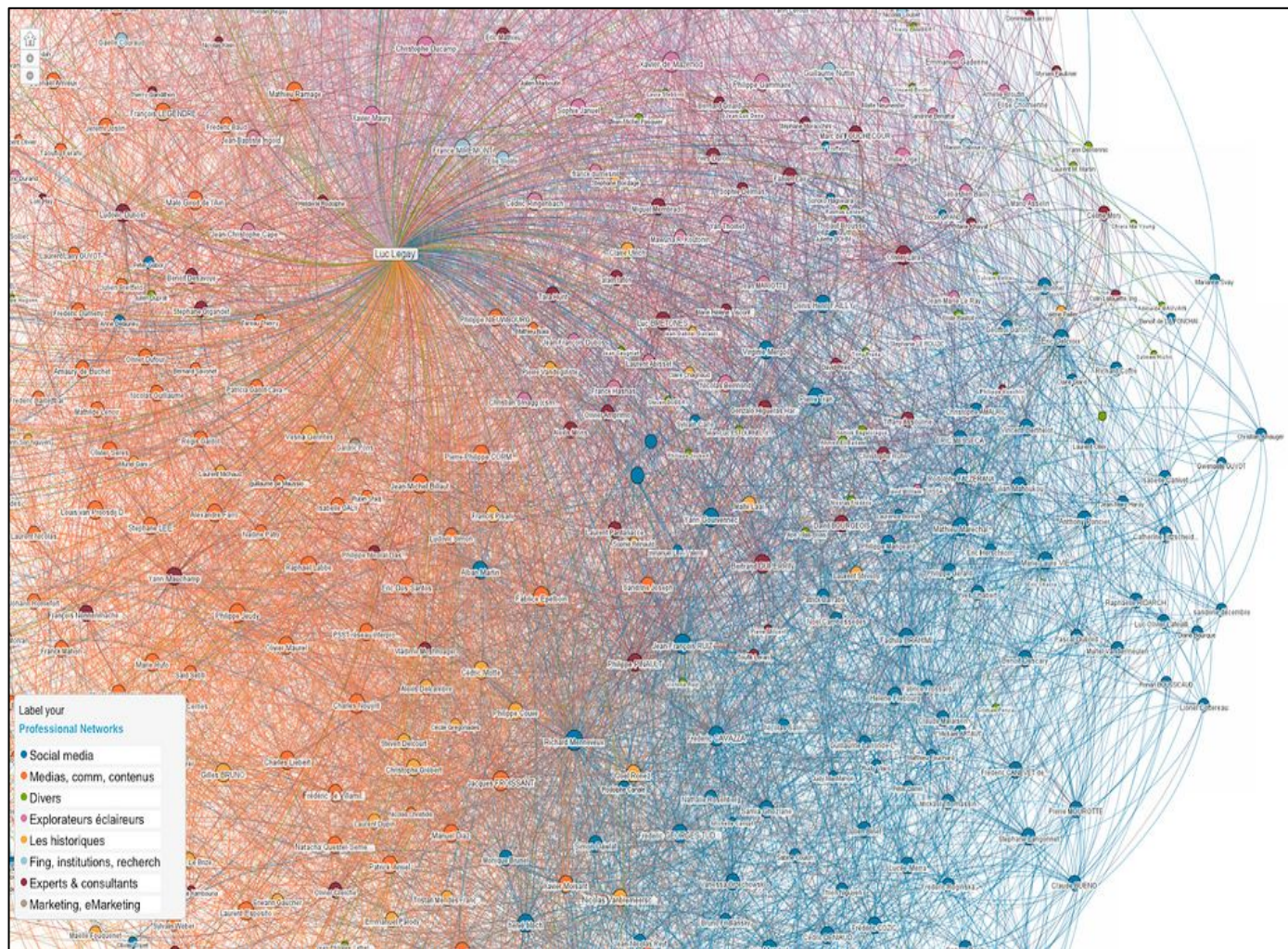
❖ Combination Packages

- VisIt
- Paraview

❖ Other

- GIS
- MatLab
- NVivo
- Omeka
- D3
- Three.js
- Circos
- Tableau Public
- RAW

Design Principles



Visual Weight

Guiding the Eye

Media: Diapers-01.jpg
Time: 00:00:00.003 - 00:00:06.033
Participant filter: All

21.75 secs



If you are not satisfied with the baby leakage protection, you will get your money back. Read more about our leakfree guarantee at www.baby.com

Participant filter: All

15.43 secs



If you are not satisfied with the baby leakage protection, you will get your money back. Read more about our leakfree guarantee at www.baby.com

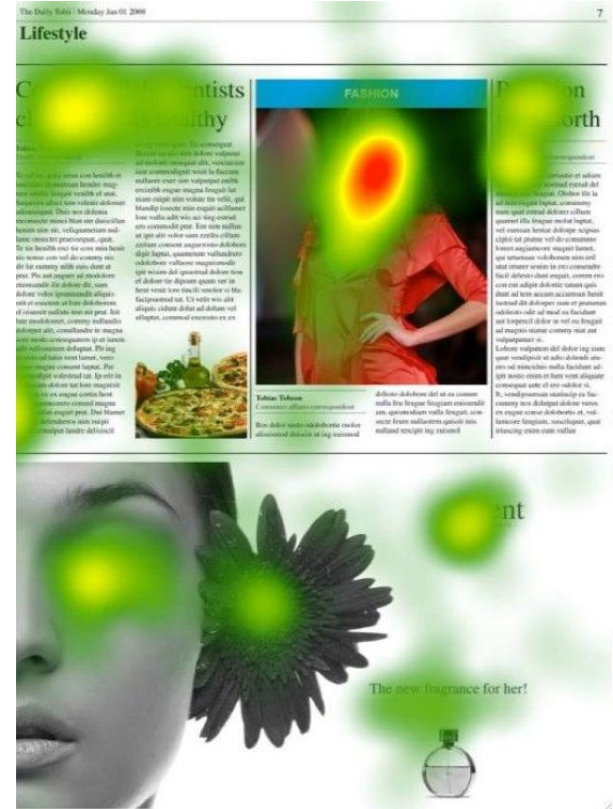
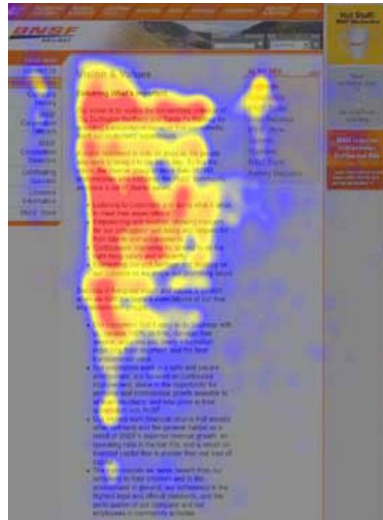
Visual Weight

Guiding the Eye

YOUR EYES HERE

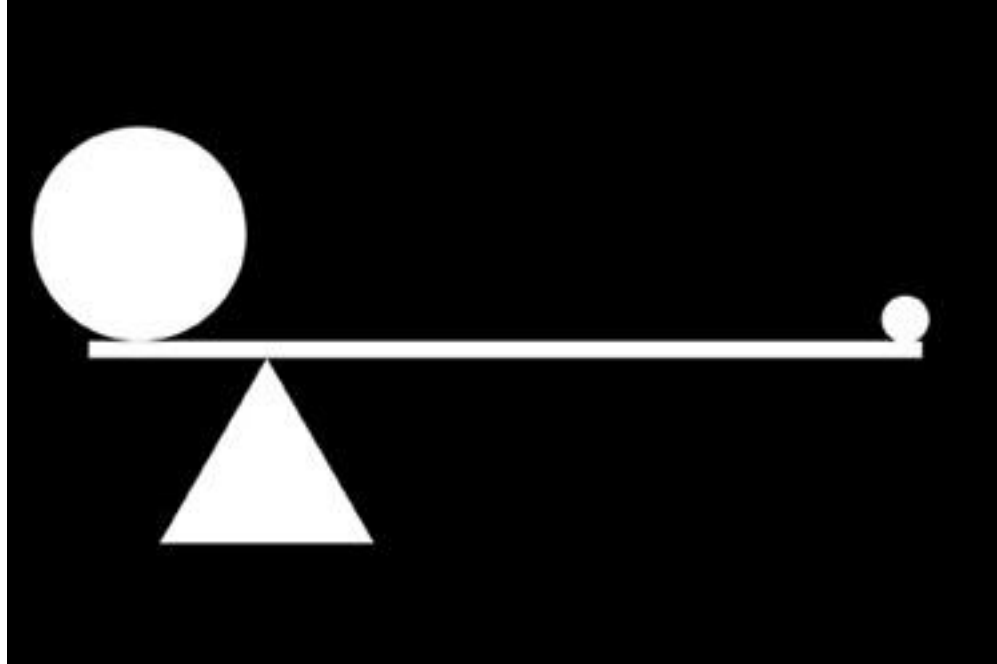
(then here)

Isn't that fascinating?

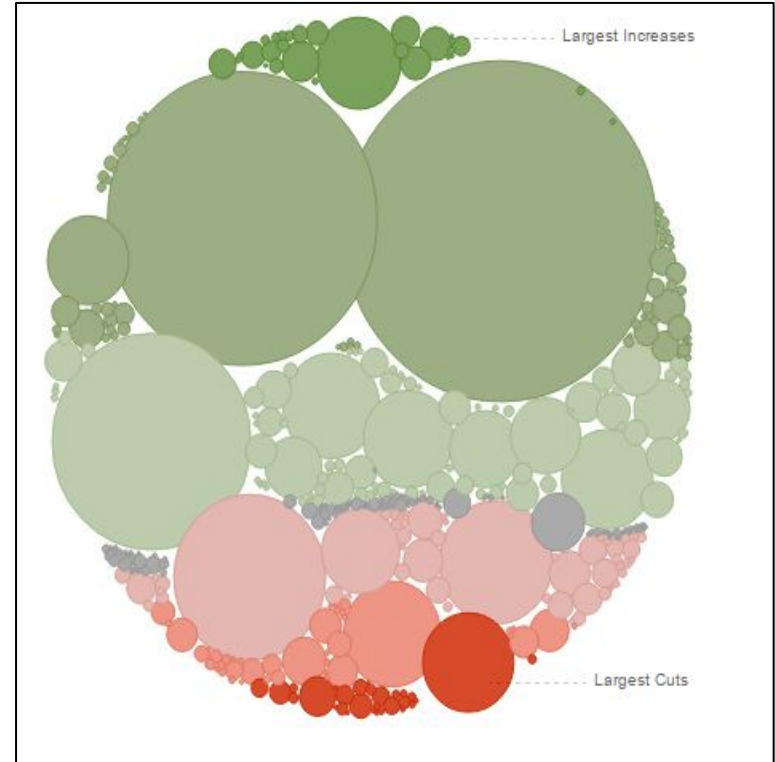
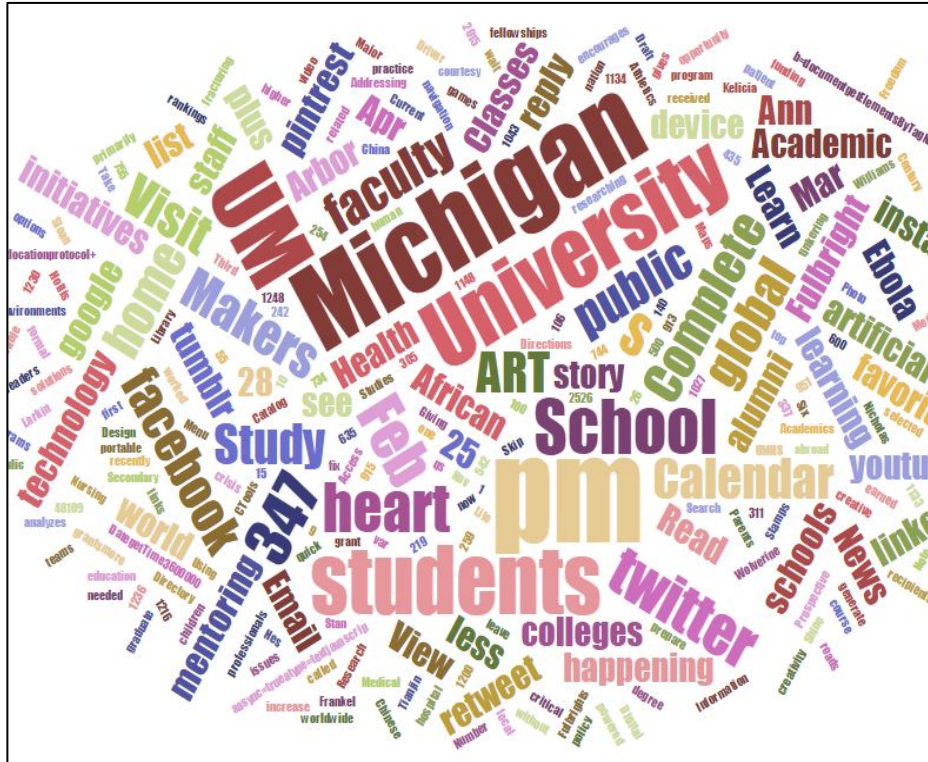


Visual Weight

Guiding the Eye - Size & Shape

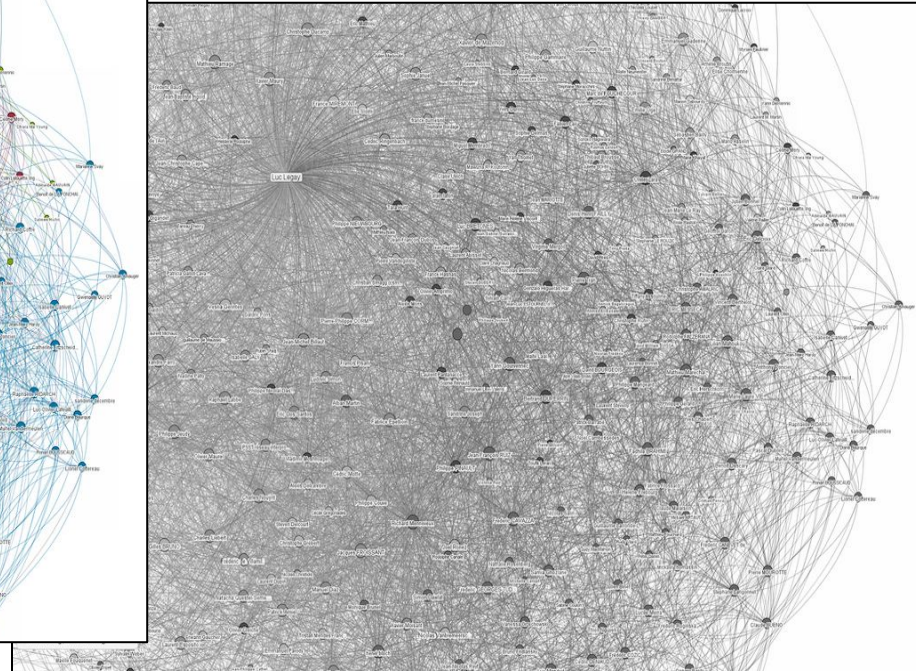
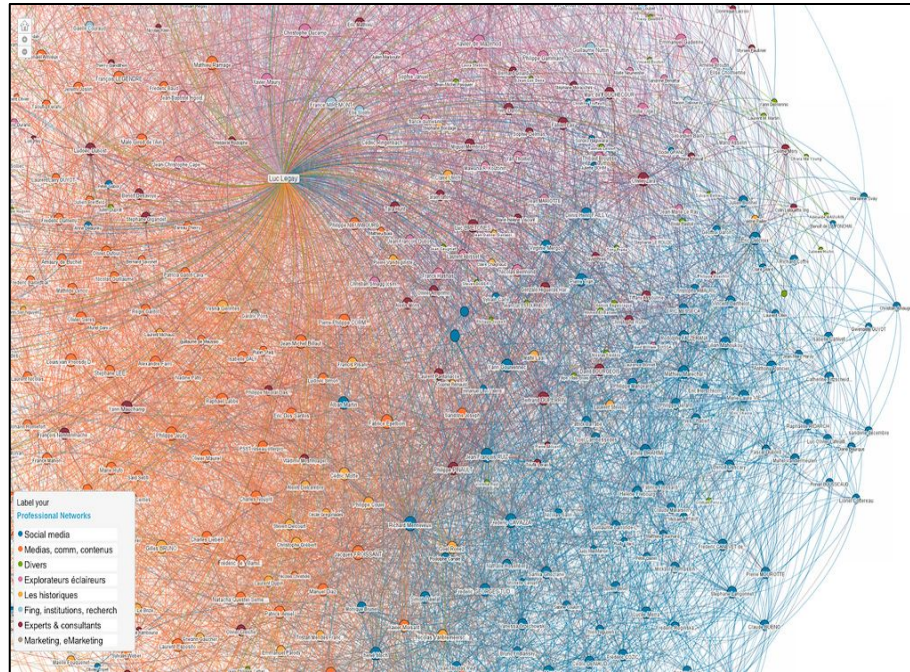


Guiding the Eye - Size & Shape



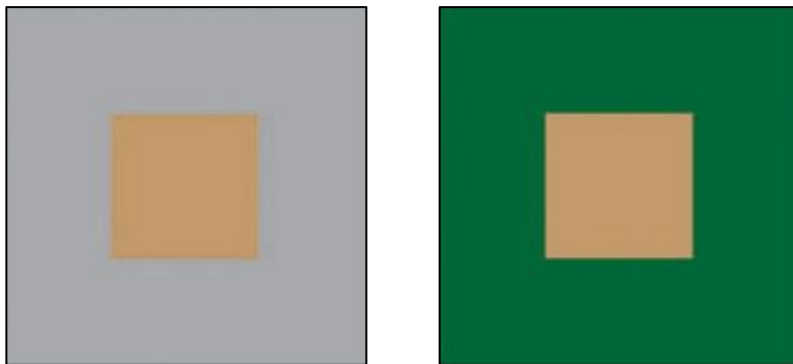
Color

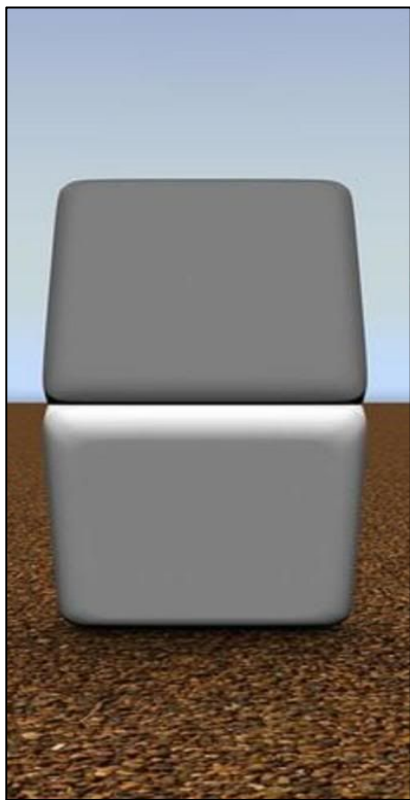
Value & Hue



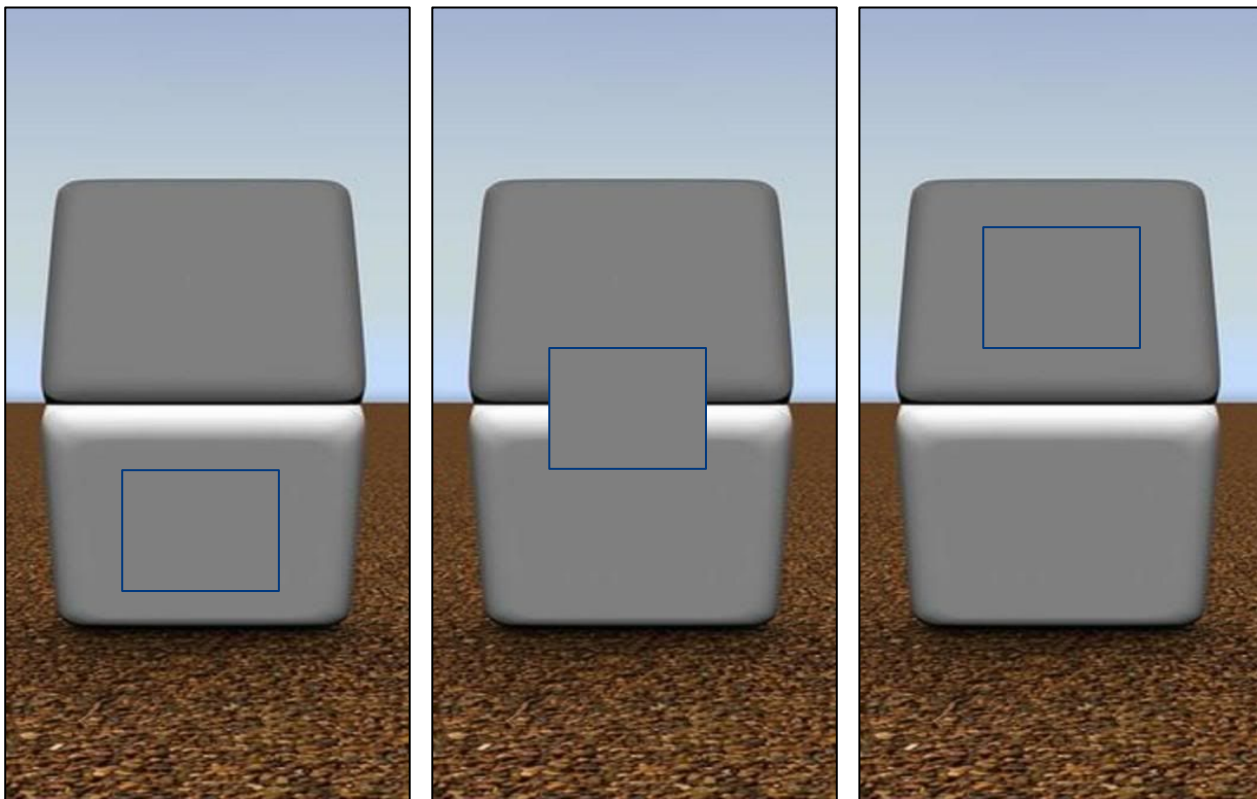
Color

Contrast

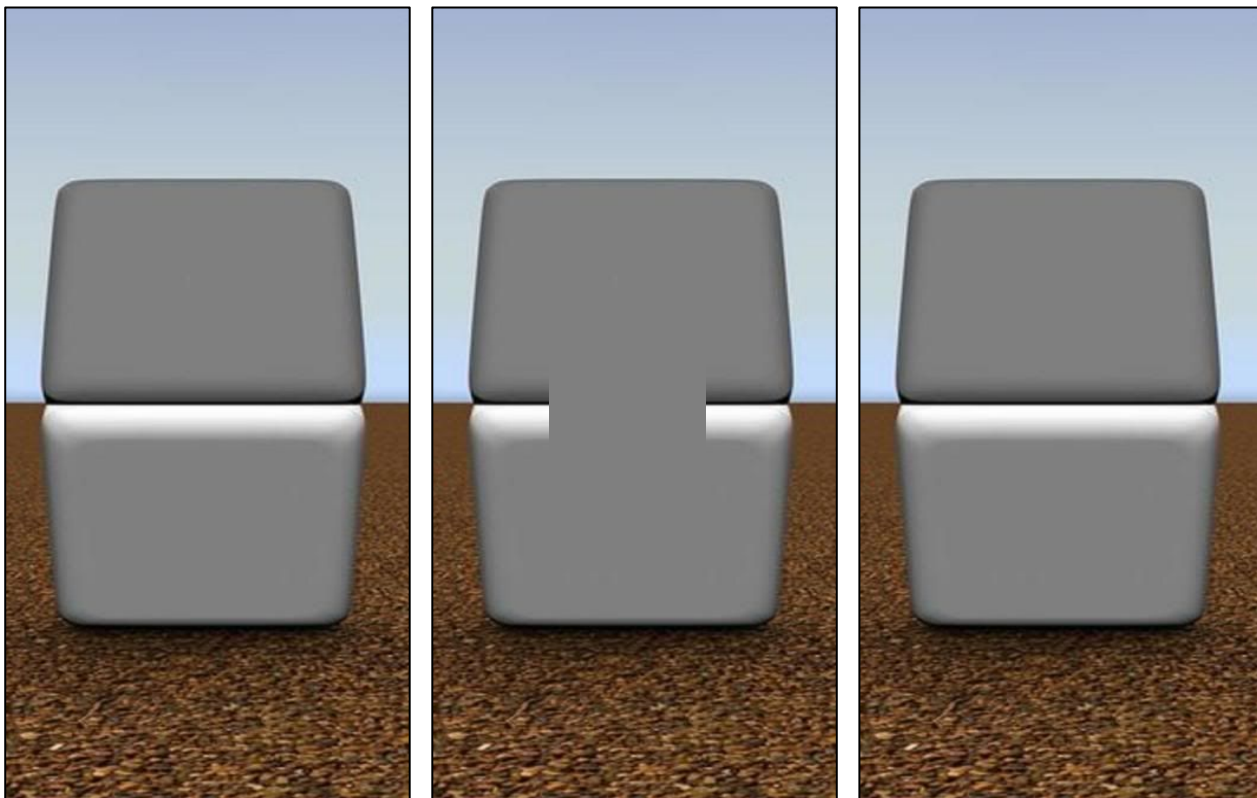




Visual Assumptions



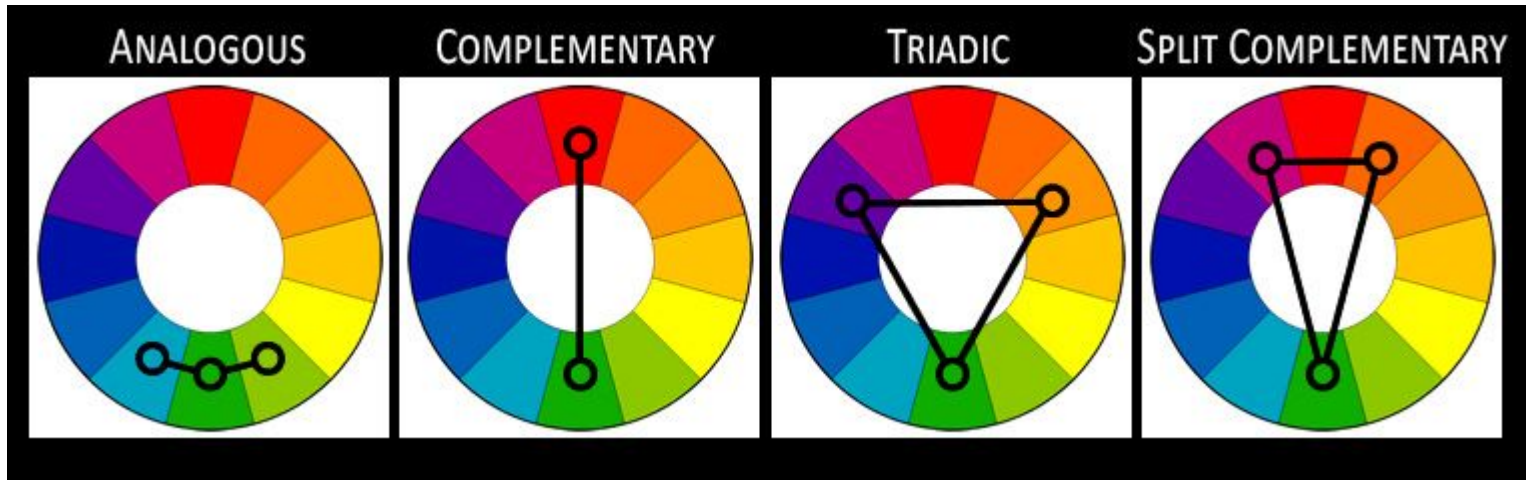
Visual Assumptions



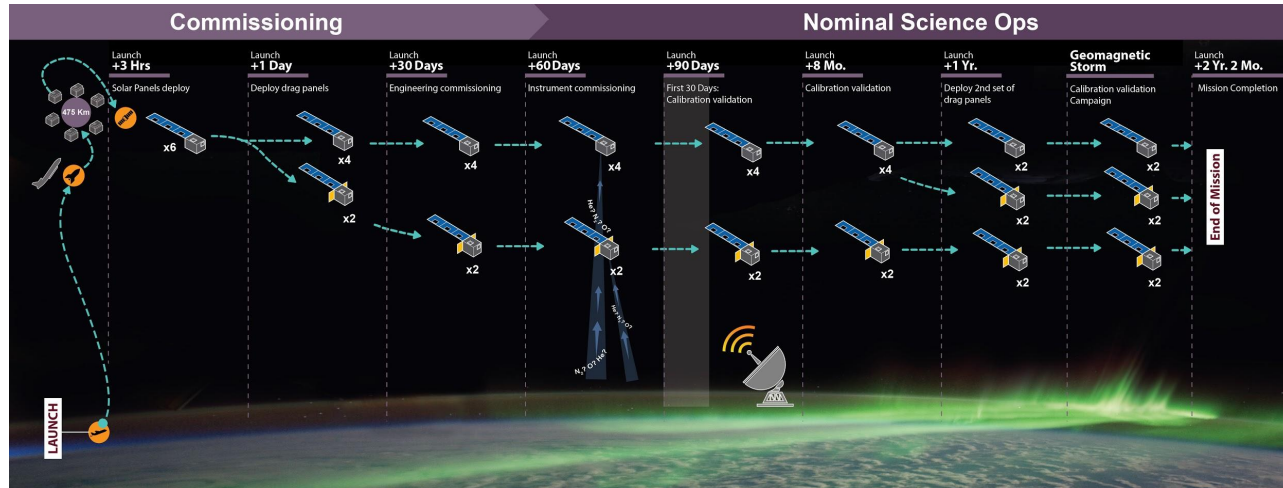
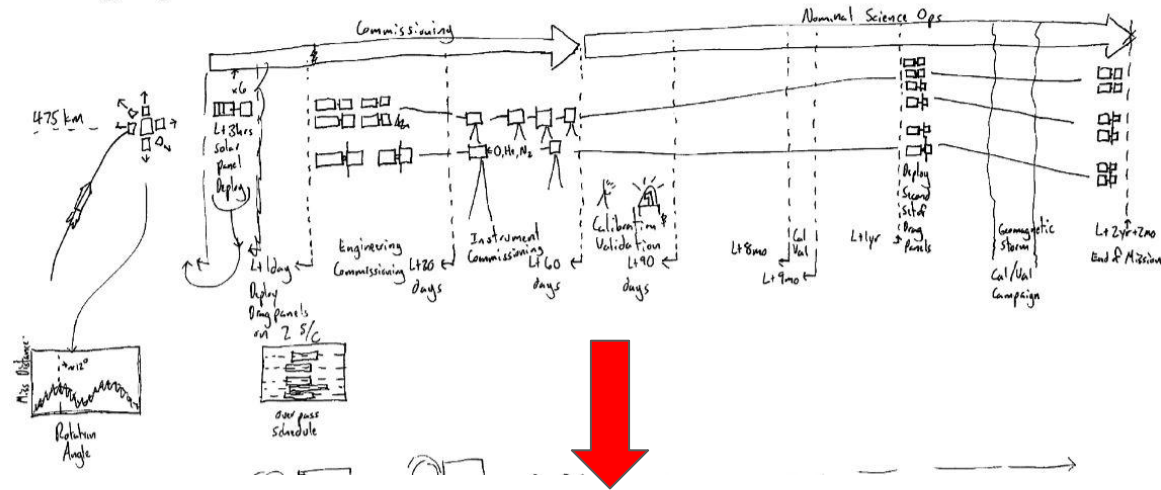
Color

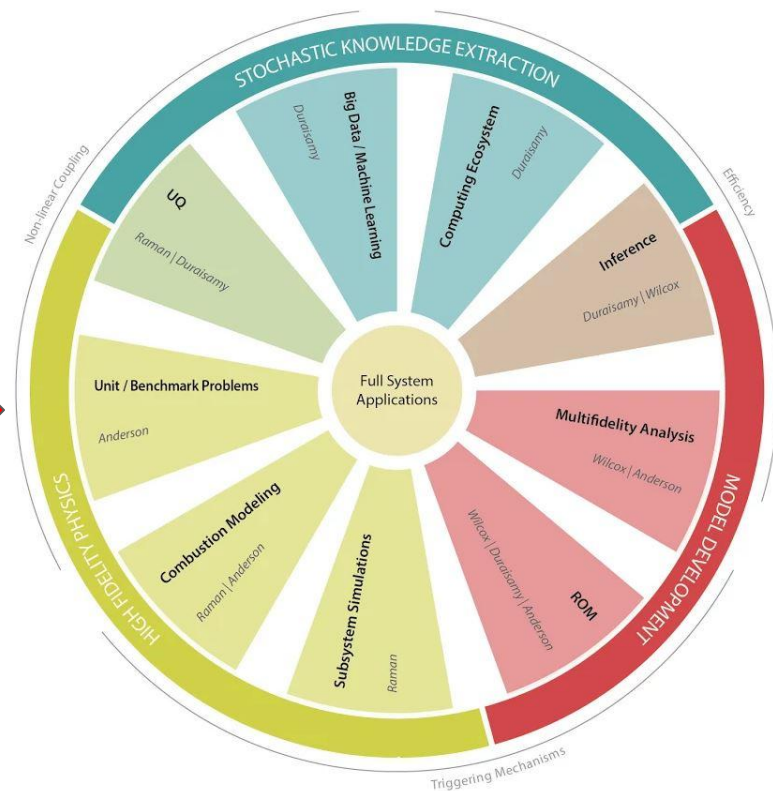
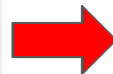
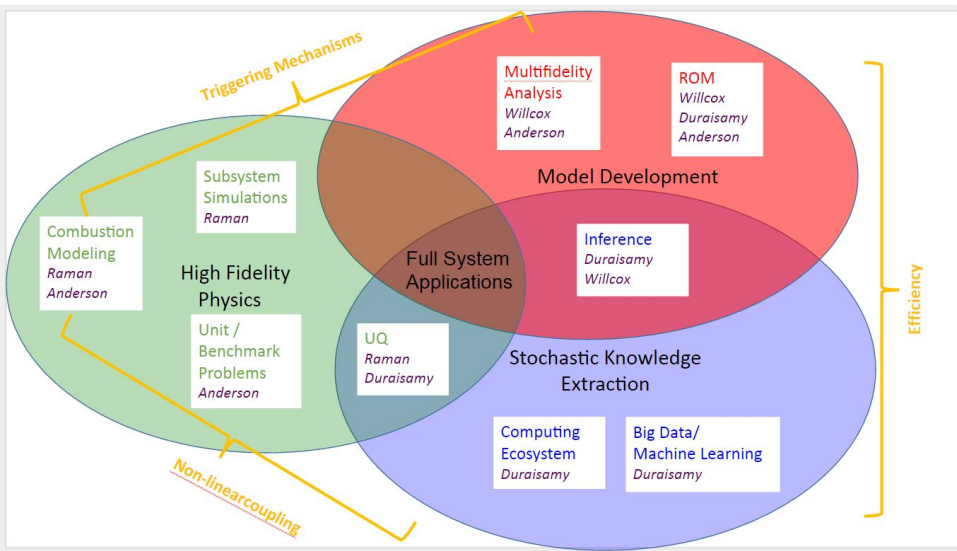
Color Schemes

Are your eyes bleeding yet?
Because this is horrible.



This is going to be F03:





Interpretation

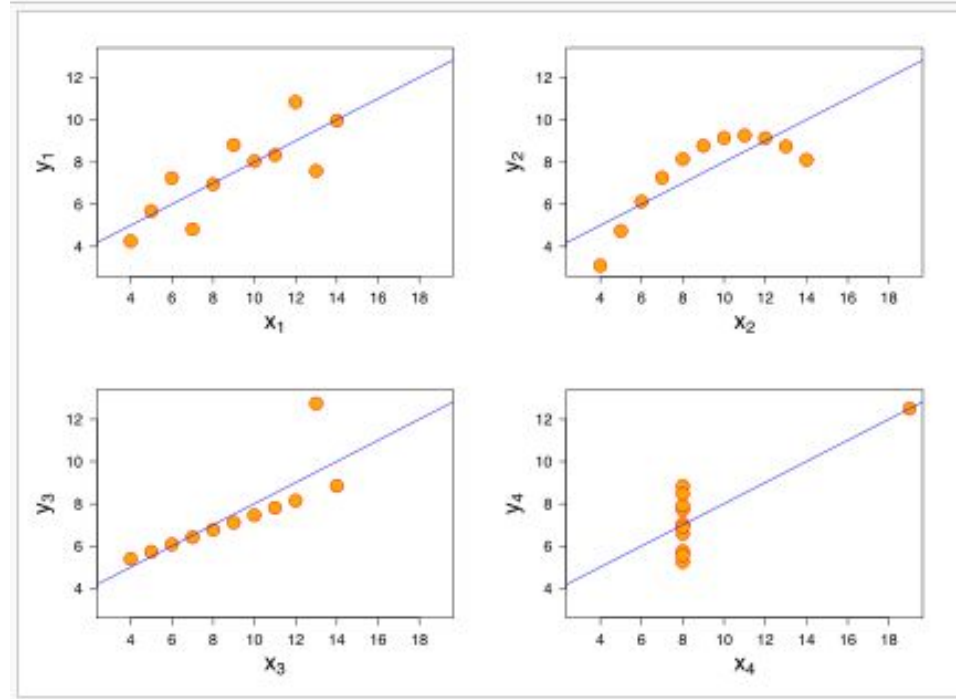
Show me the Data

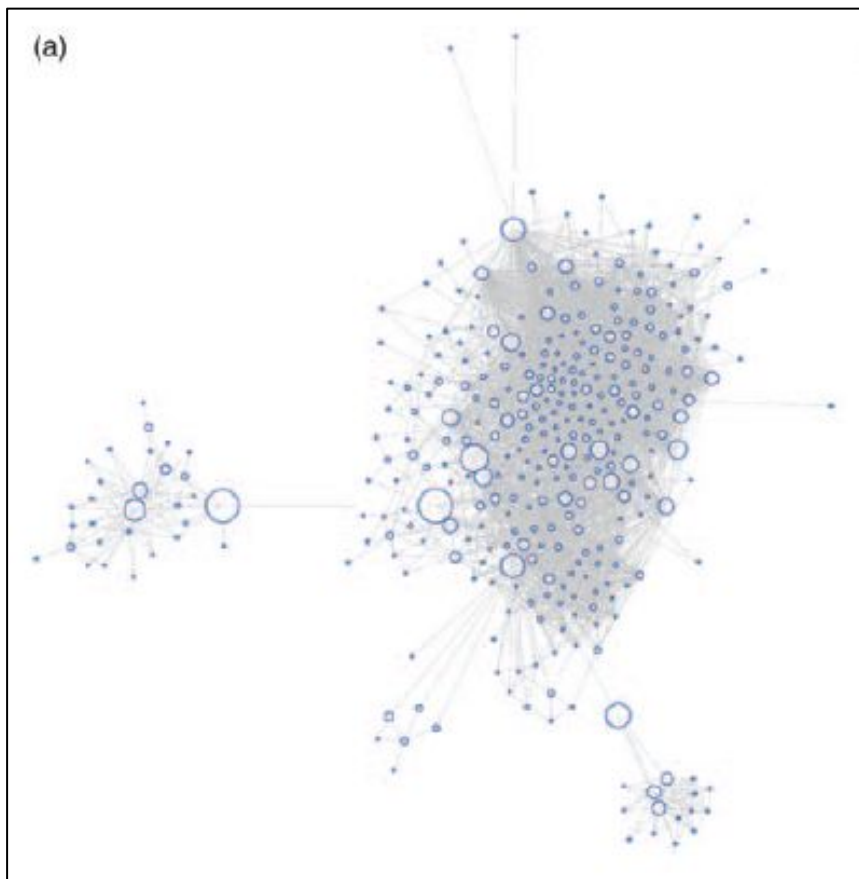
I		II		III		IV	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	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

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)

Anscombe's Quartet

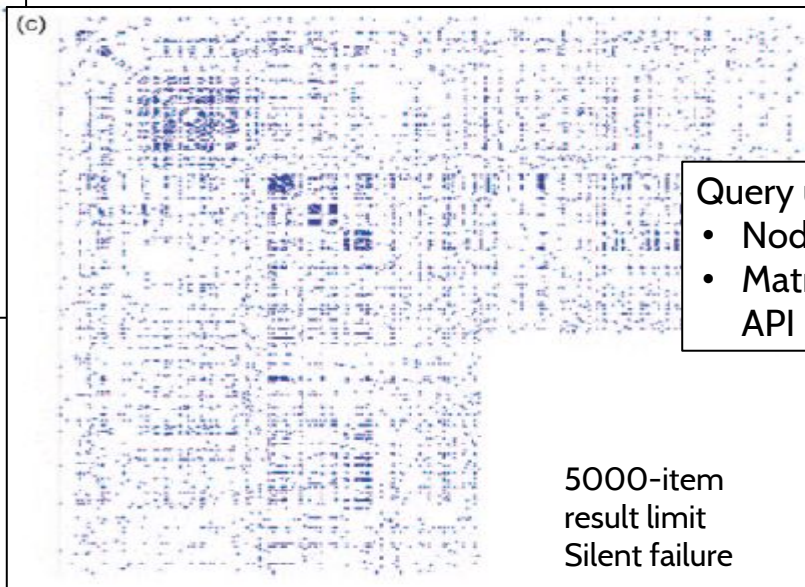
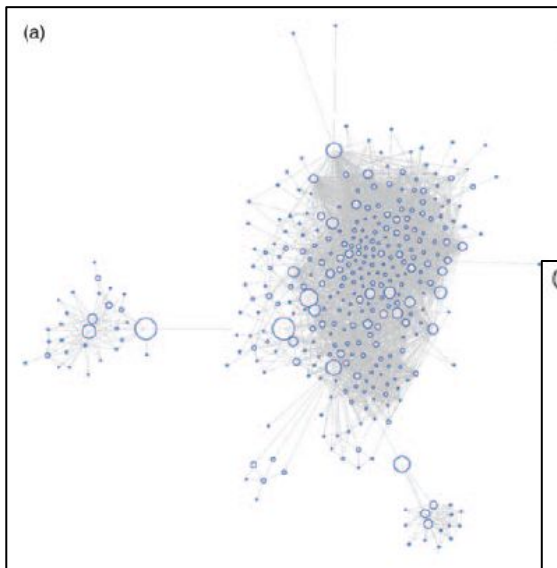
I		II		III		IV	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	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





Query using Facebook API

- Node-link diagram



Query using Facebook API

- Node-link diagram
- Matrix display,
API return order

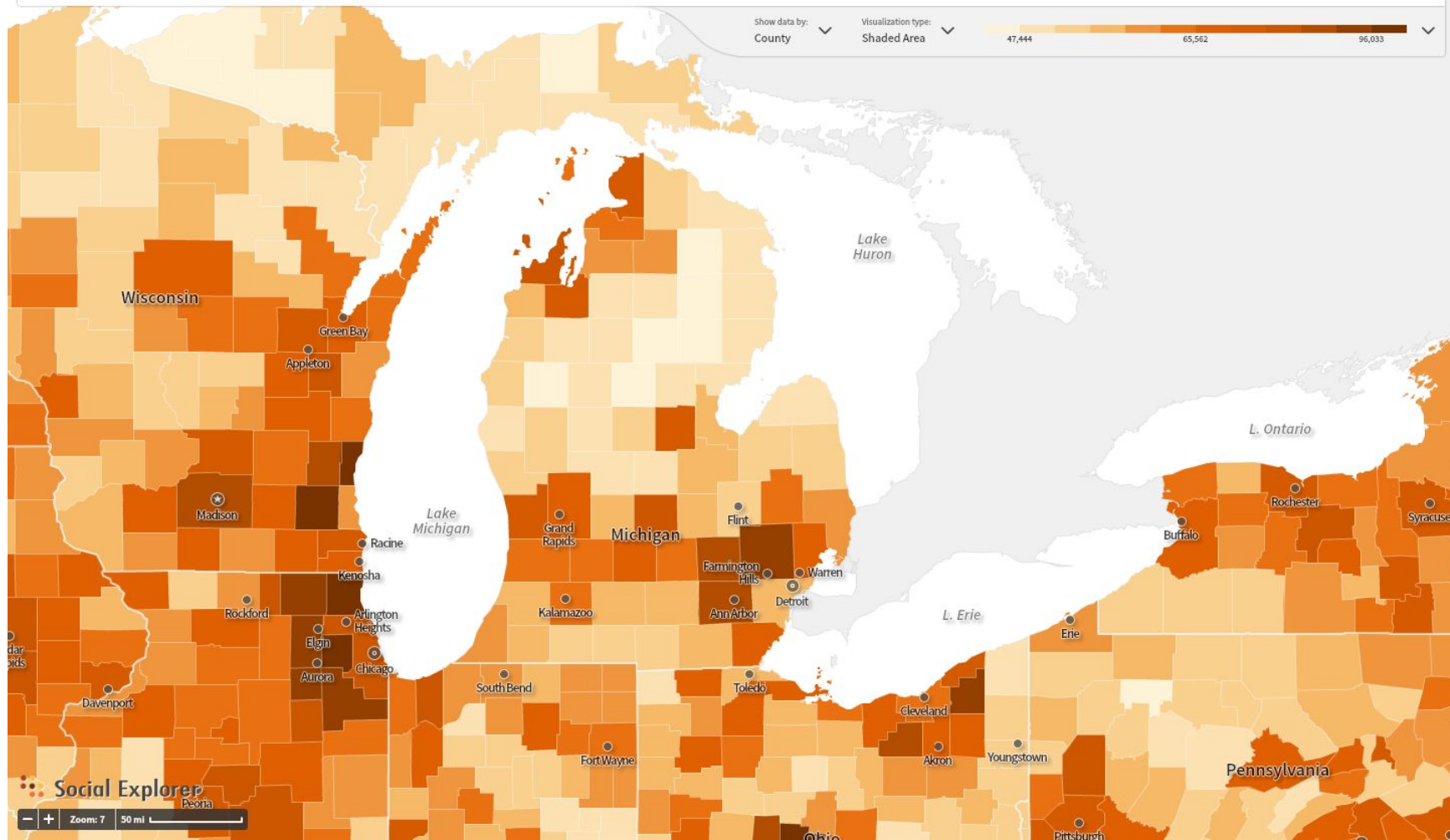
5000-item
result limit
Silent failure

Average household income (In 2014 Inflation Adjusted Dollars):

ACS 2014 (5-Year Estimates)

Show data by:
County

Visualization type:
Shaded Area



Median household income (In 2014 Inflation Adjusted Dollars)

ACS 2014 (5-Year Estimates)

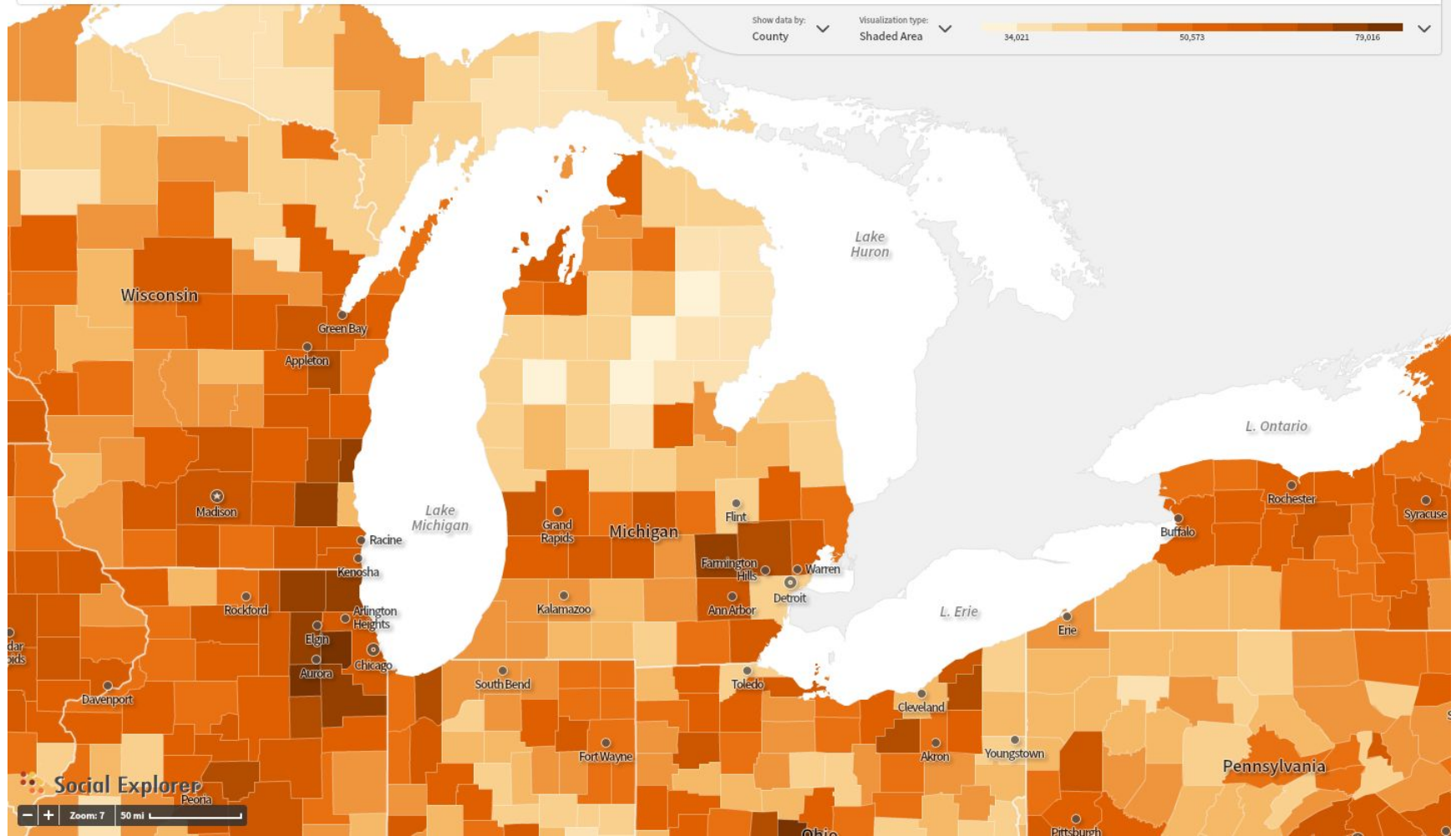
Show data by:
County

Visualization type:
Shaded Area

34,021

50,573

79,016



Median household income (In 2014 Inflation Adjusted Dollars)

ACS 2014 (5-Year Estimates)

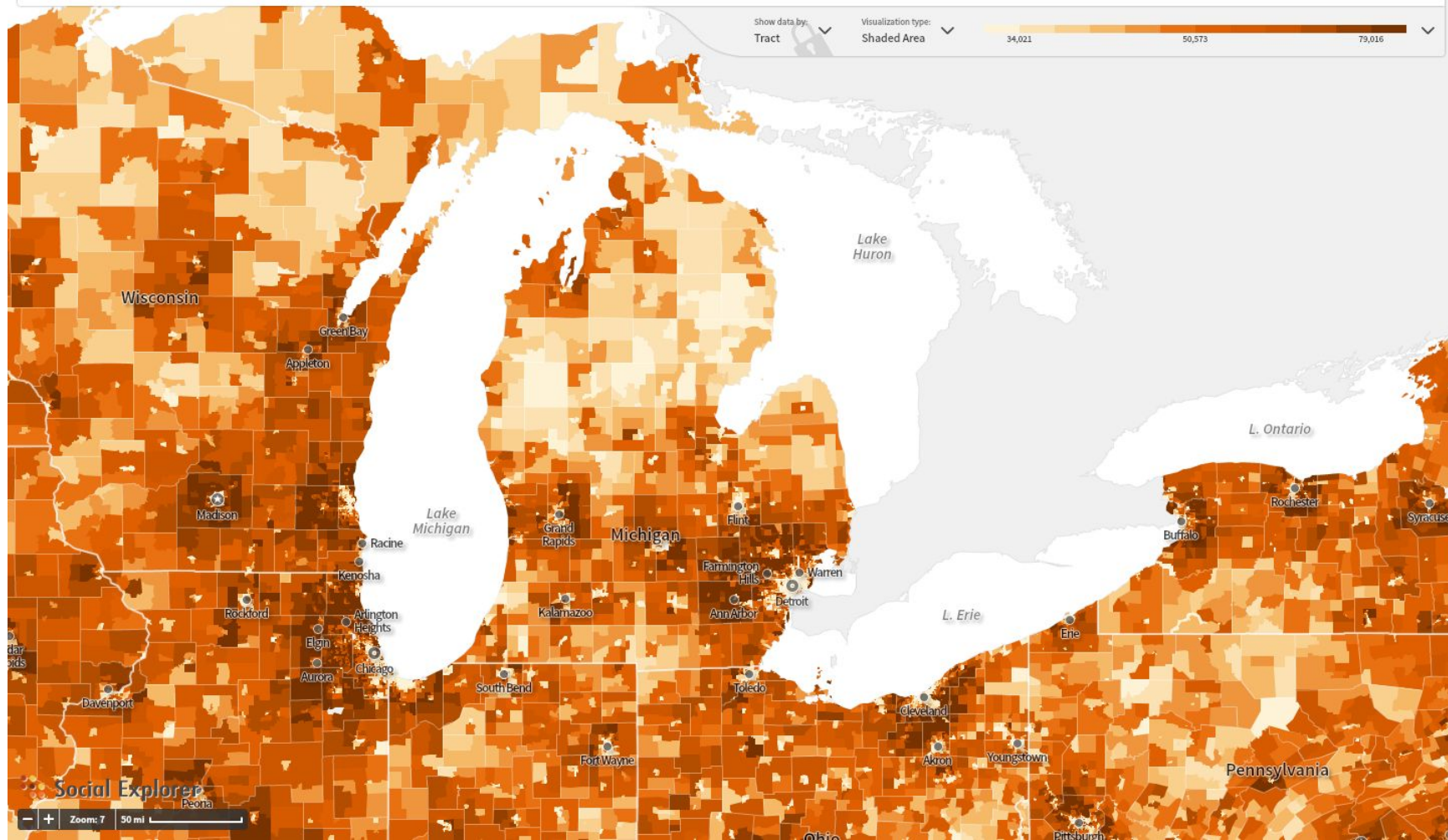
Show data by:
Tract

Visualization type:
Shaded Area

34,021

50,573

79,016

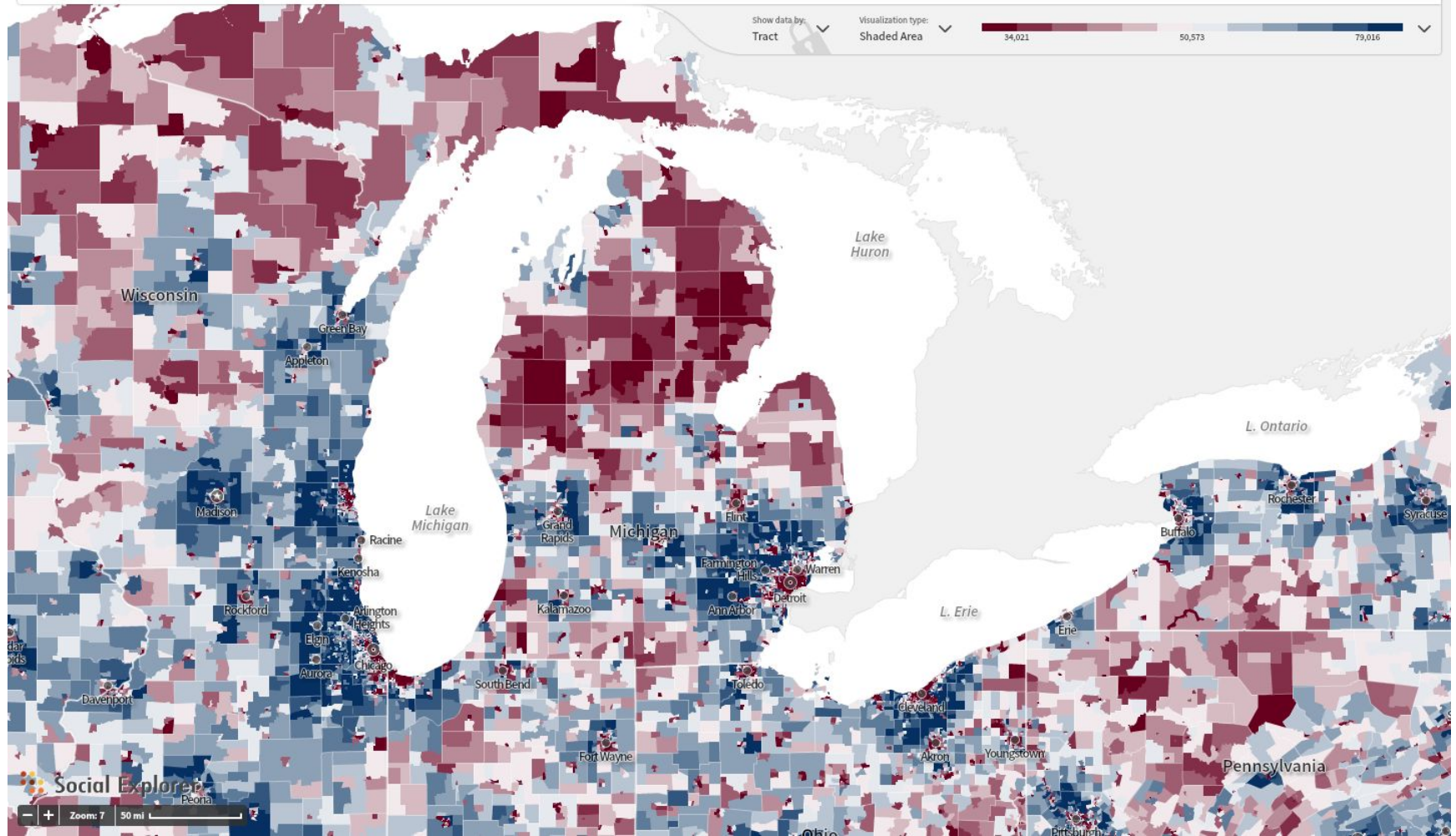
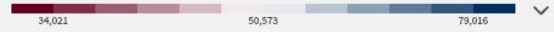


Median household income (in 2014 Inflation Adjusted Dollars)

ACS 2014 (5-Year Estimates)

Show data by:
Tract

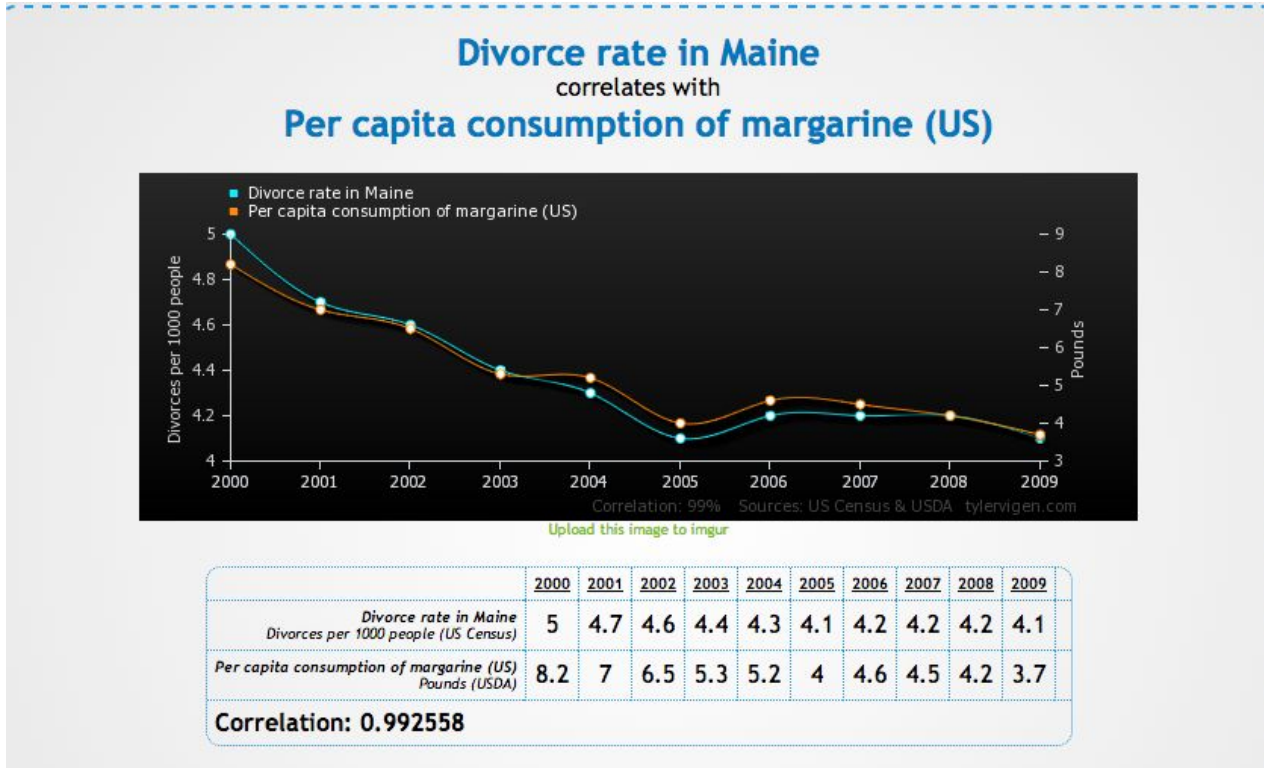
Visualization type:
Shaded Area



Reasons for Visualization (Redux)

- ❖ Discovery
- ❖ Hypothesis generation
- ❖ Analysis
- ❖ Representation of data for publication or presentation - communication

Danger, Will Robinson



Campus Resources

Library

- ❖ Taubman Health Sciences Library
 - Anatomage Table
 - Cytoscape
 - Health-related visualization
- ❖ Clark Library (Hatcher Library)
 - Maps, government information, and spatial and numeric data services
 - Manipulation of diverse types of data
- ❖ UM3D Lab (Duderstadt Center)
 - 3-D visualization via a variety of methods
- ❖ Library Data Visualization Webpage:
<http://www.lib.umich.edu/data-visualization>
- ❖ Library Data Visualization E-Mail Group: libvisualization@umich.edu

Campus

- ❖ [Advanced Research Computing](#) (ARC)
 - Computational science, data science, technology services, statistics consultation
 - High performance computing
- ❖ [Michigan Multimedia](#) (M3)
 - Build websites
 - Specialize in web strategy and graphic design
- ❖ [3D Imaging Laboratory](#) (Department of Radiology)
 - Provide clinically relevant visualization and analysis of medical imaging data
- ❖ [Microscopy & Image Analysis Laboratory](#) (part of BRCF)
 - Equipment for microscopic imaging, including fluorescence microscopy, & scanning and transmission electron microscopy