

The Art *of* Numbers

Alyssa A. Goodman • Harvard University

Relative Strengths



Pattern Recognition
Creativity

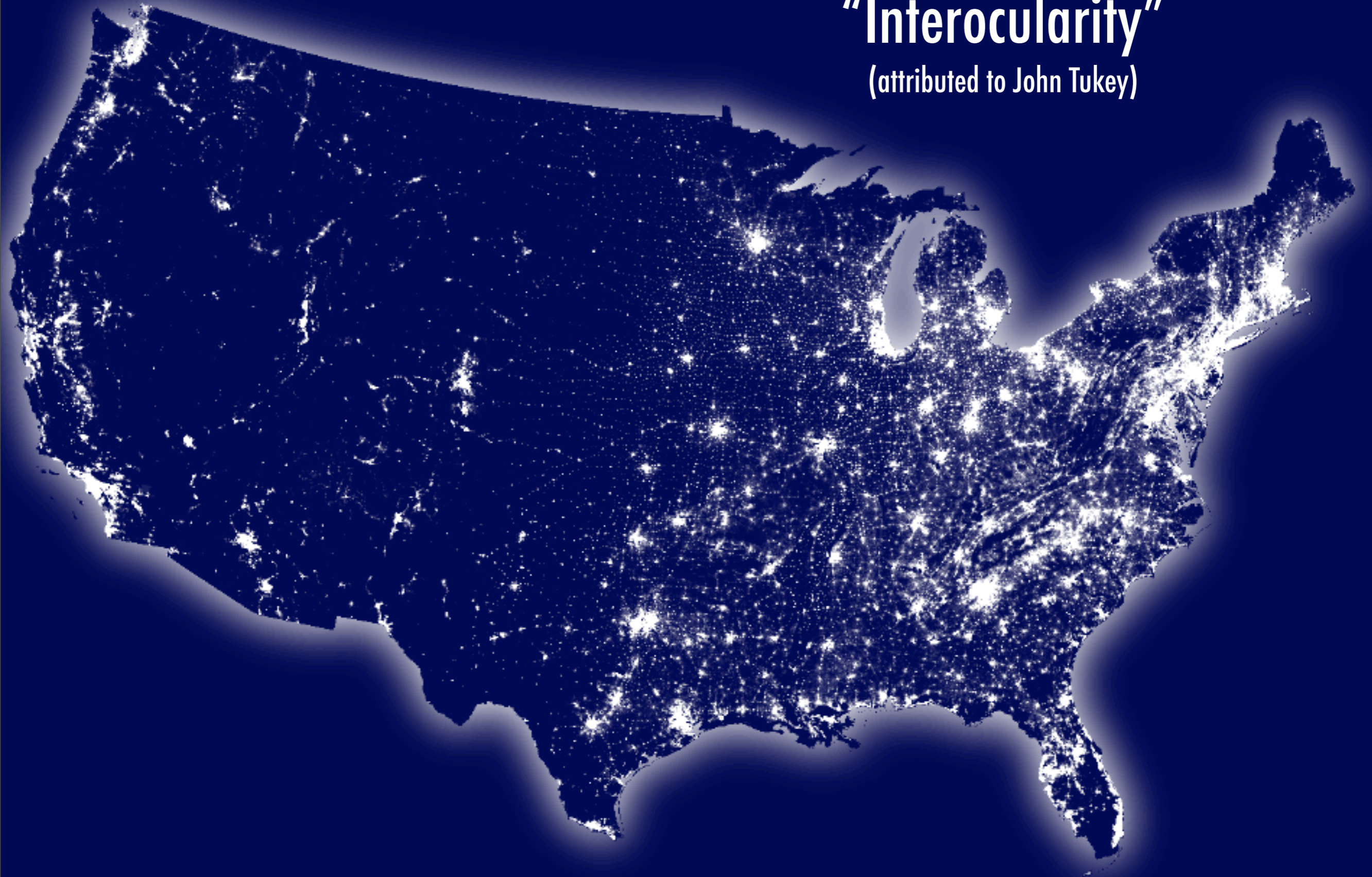


Calculations

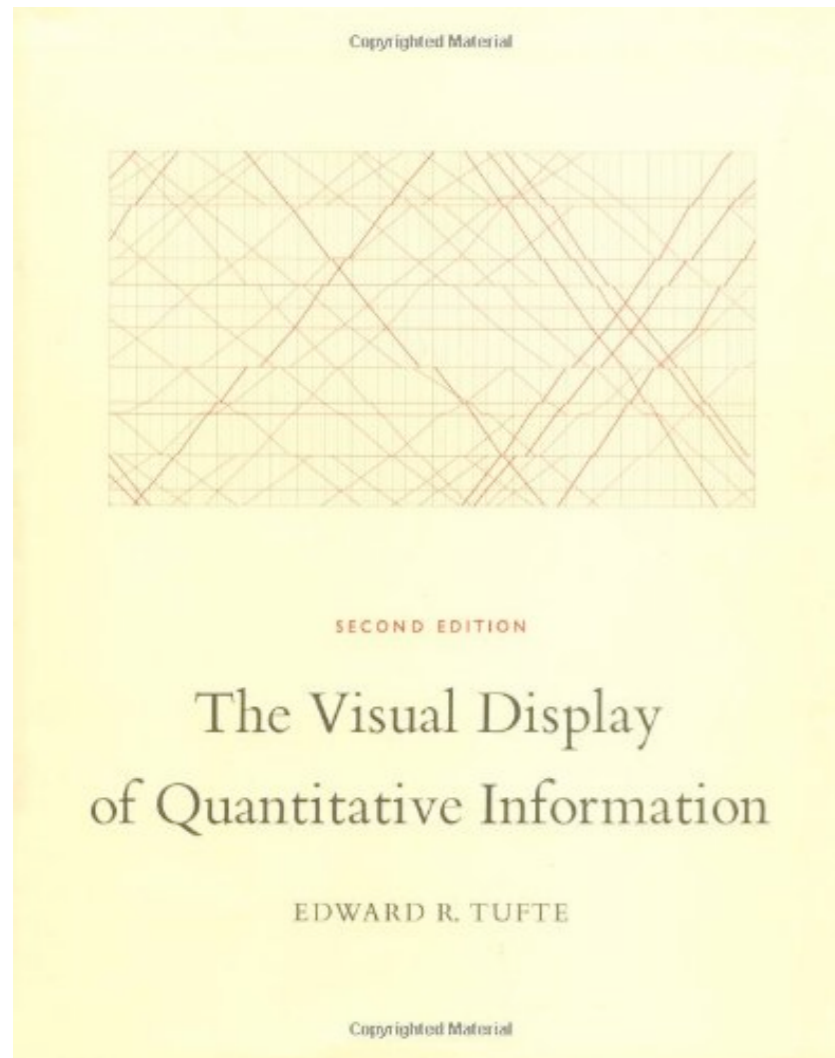


"Interocularity"

(attributed to John Tukey)

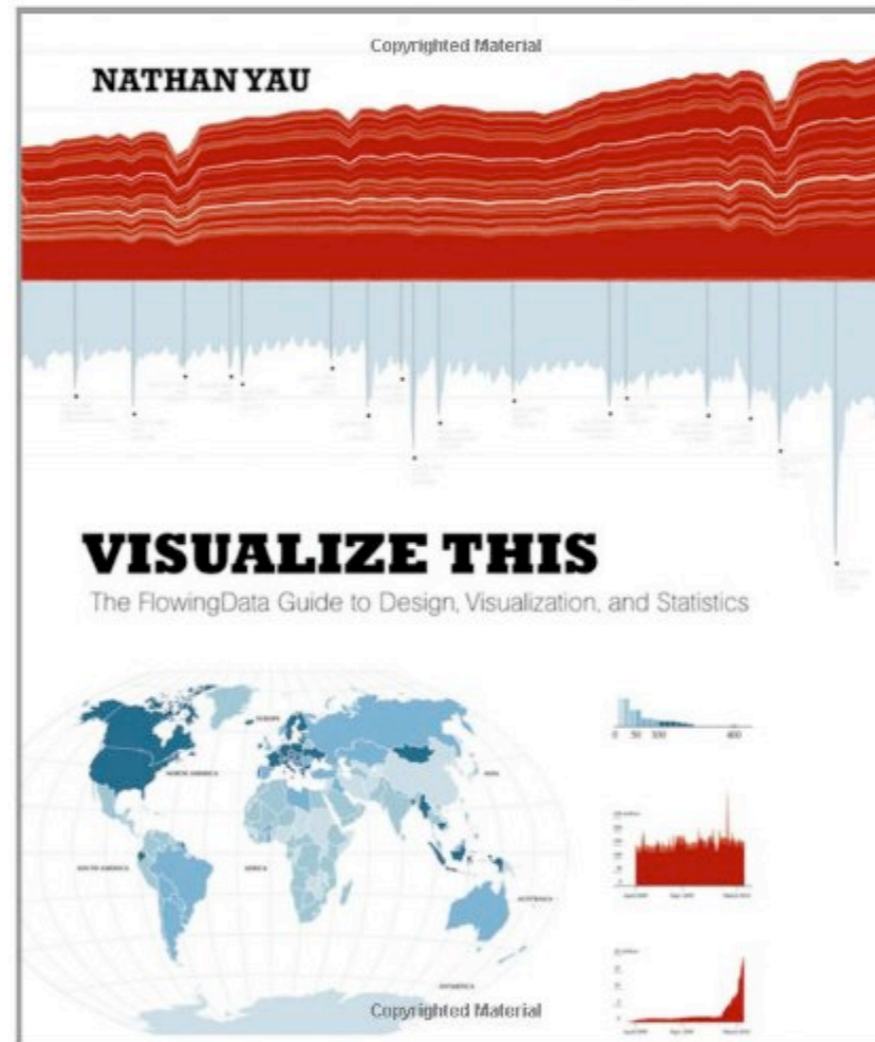


The Classic



1983

Modern "How-to"



2011

Case Studies



2012



Principles of high-dimensional data visualization in astronomy

A.A. Goodman*

Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA

Received 2012 May 3, accepted 2012 May 4

Published online 2012 Jun 15

Key words cosmology: large-scale structure – ISM: clouds – methods: data analysis – techniques: image processing – techniques: radial velocities

Astronomical researchers often think of analysis and visualization as separate tasks. In the case of high-dimensional data sets, though, interactive *exploratory data visualization* can give far more insight than an approach where data processing and statistical analysis are followed, rather than accompanied, by visualization. This paper attempts to chart a course toward “linked view” systems, where multiple views of high-dimensional data sets update live as a researcher selects, highlights, or otherwise manipulates, one of several open views. For example, imagine a researcher looking at a 3D volume visualization of simulated or observed data, and simultaneously viewing statistical displays of the data set’s properties (such as an x - y plot of temperature vs. velocity, or a histogram of vorticities). Then, imagine that when the researcher selects an interesting group of points in any one of these displays, that the same points become a highlighted subset in all other open displays. Selections can be graphical or algorithmic, and they can be combined, and saved. For tabular (ASCII) data, this kind of analysis has long been possible, even though it has been under-used in astronomy. The bigger issue for astronomy and other “high-dimensional” fields, though, is that no extant system allows for full integration of images and data cubes within a linked-view environment. The paper concludes its history and analysis of the present situation with suggestions that look toward cooperatively-developed open-source modular software as a way to create an evolving, flexible, high-dimensional, linked-view visualization environment useful in astrophysical research.

© 2012 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

<http://adsabs.harvard.edu/abs/2012AN...333..505G>

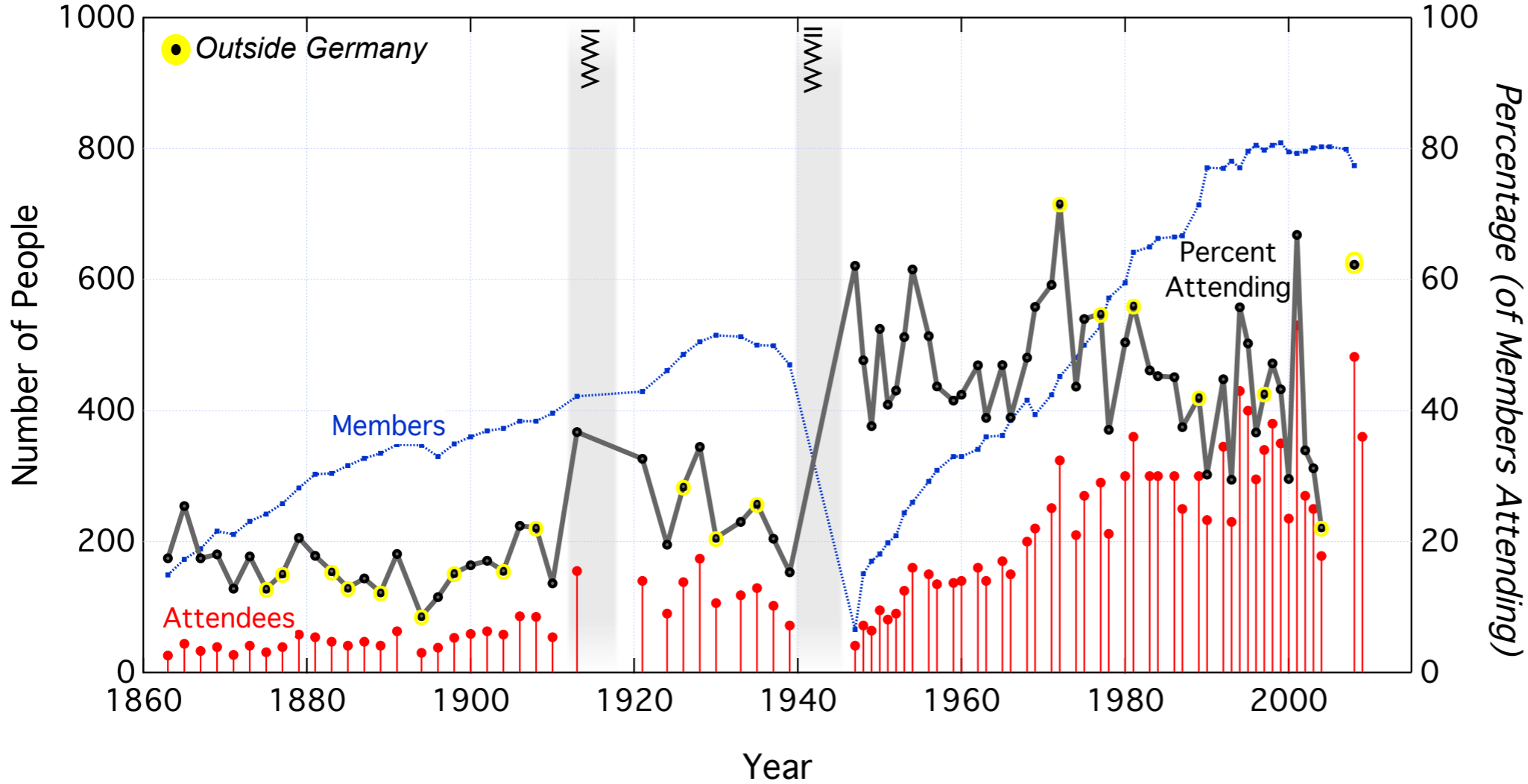
Data • Dimensions • Display

Linked Views

Data • Dimensions • Display

Linked Views

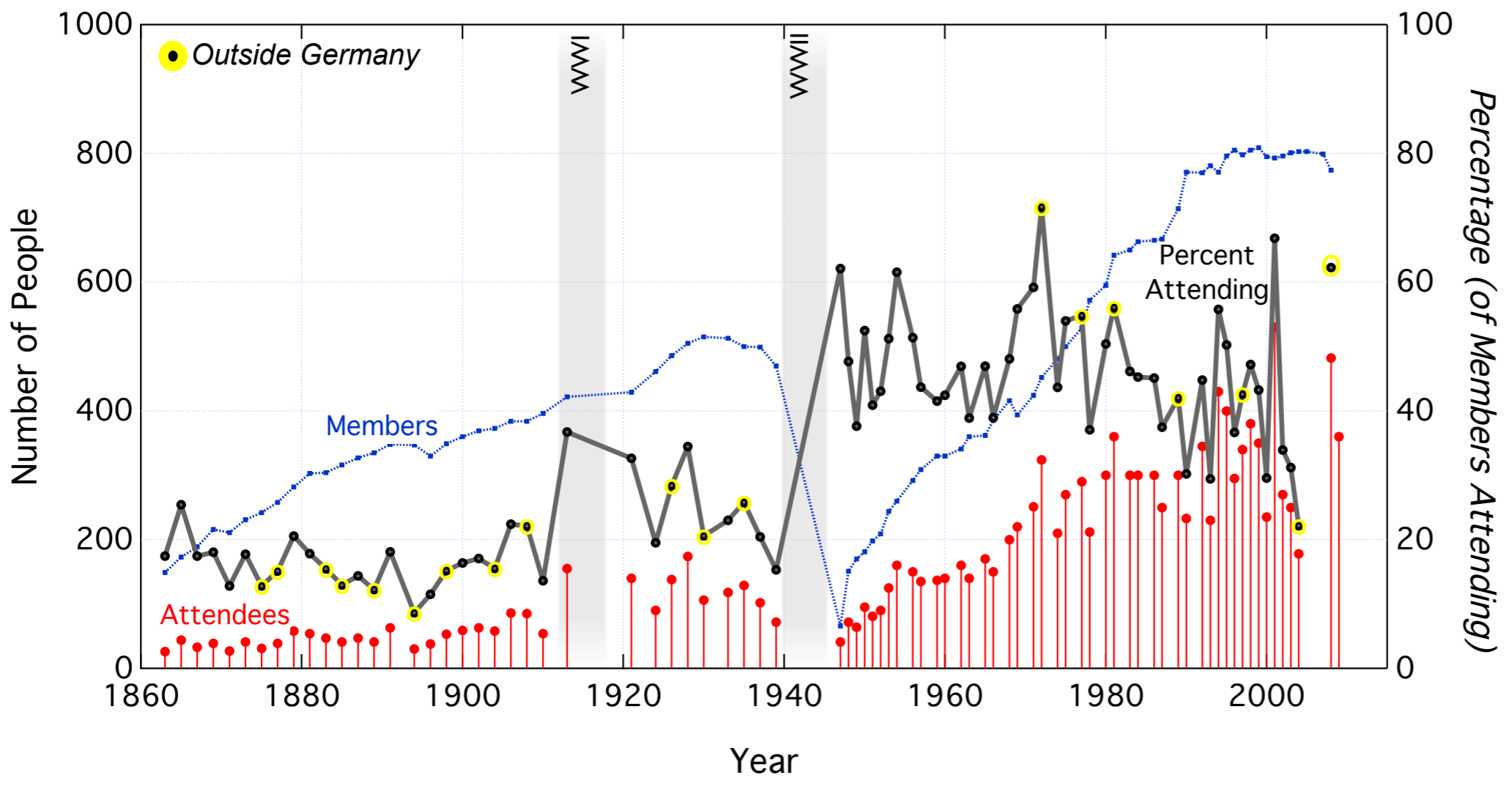
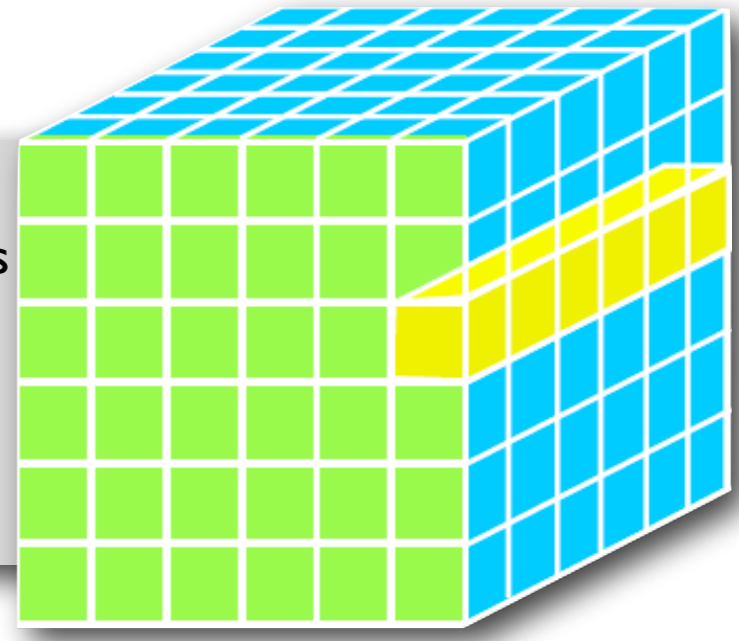
“HIGH-DIMENSIONAL DATA”



figures reproduced from Goodman 2012, "Principles of High-Dimensional Data Visualization in Astronomy"

“HIGH-DIMENSIONAL DATA”

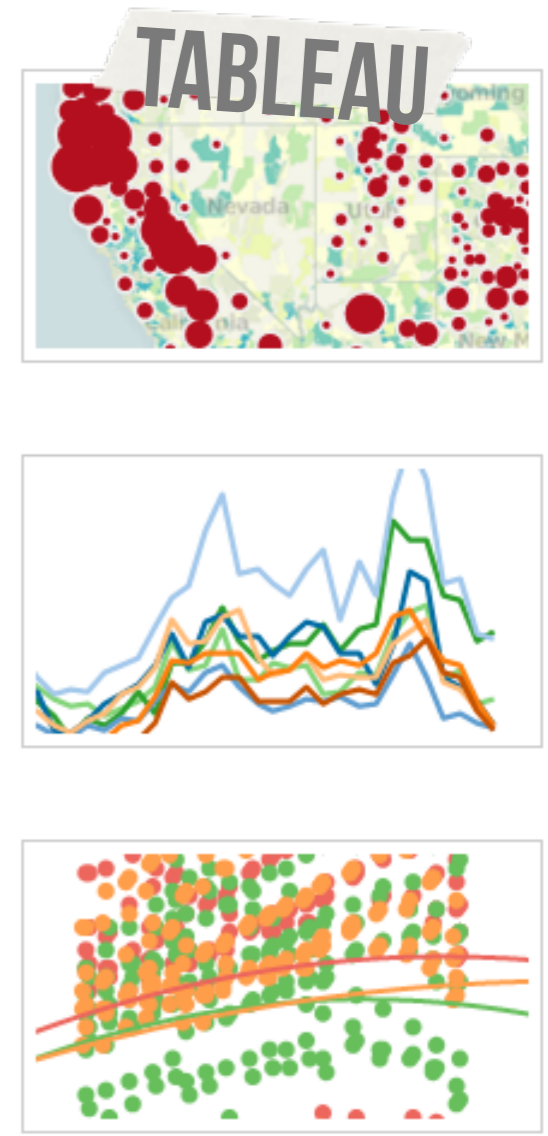
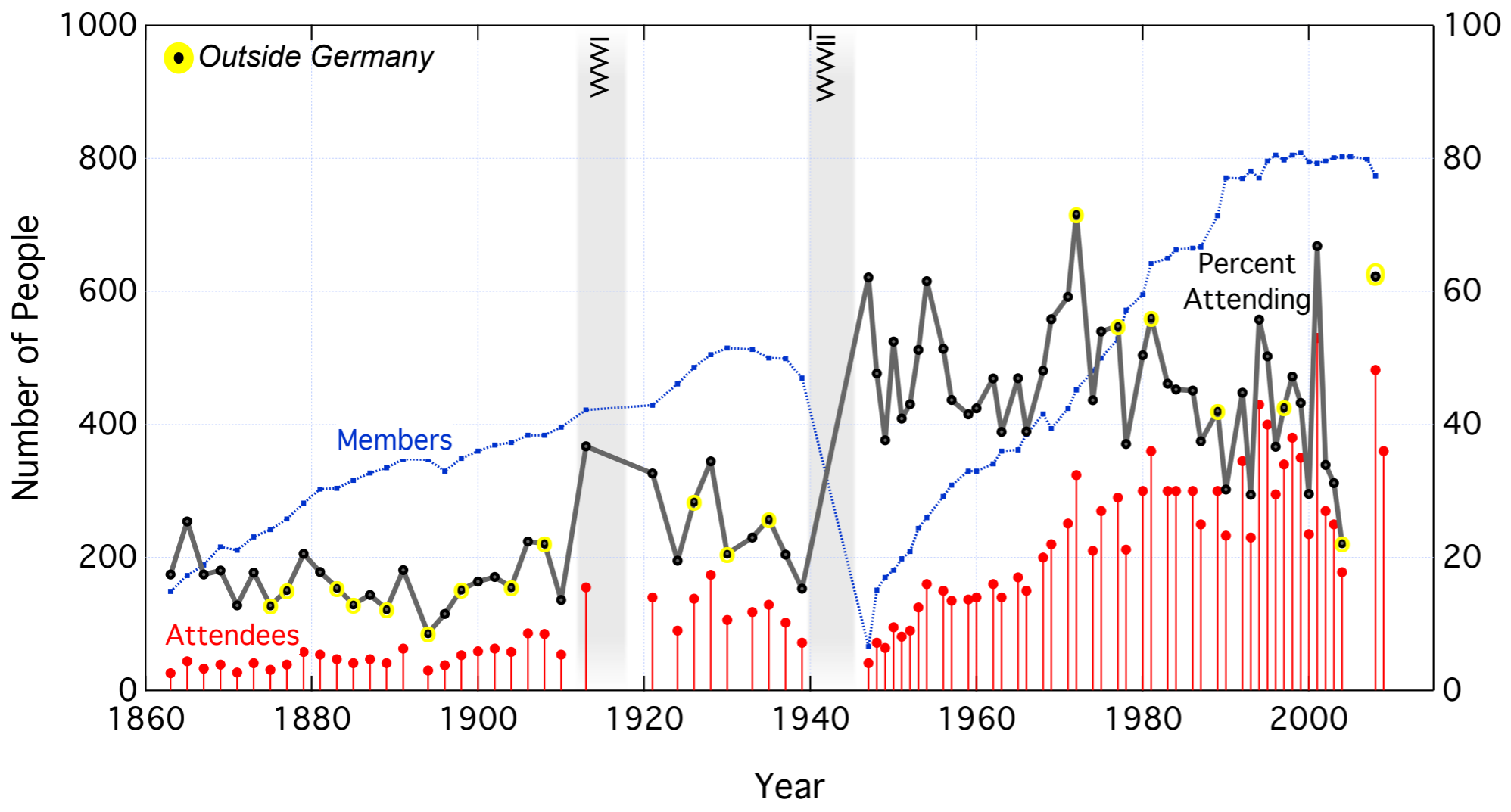
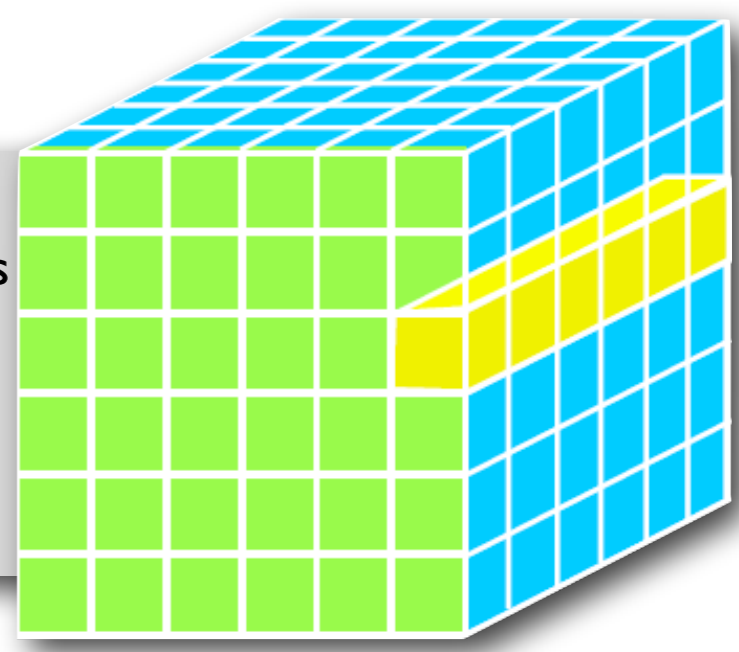
GENERALLY
1D: Columns = Spectra, SEDs, Time Series
2D: Faces or Slices = Images
3D: Volumes = 3D Renderings, 2D Movies
4D: Time Series of Volumes = 3D Movies



figures reproduced from Goodman 2012, “Principles of High-Dimensional Data Visualization in Astronomy”

“HIGH-DIMENSIONAL DATA”

GENERALLY
1D: Columns = Spectra, SEDs, Time Series
2D: Faces or Slices = Images
3D: Volumes = 3D Renderings, 2D Movies
4D: Time Series of Volumes = 3D Movies

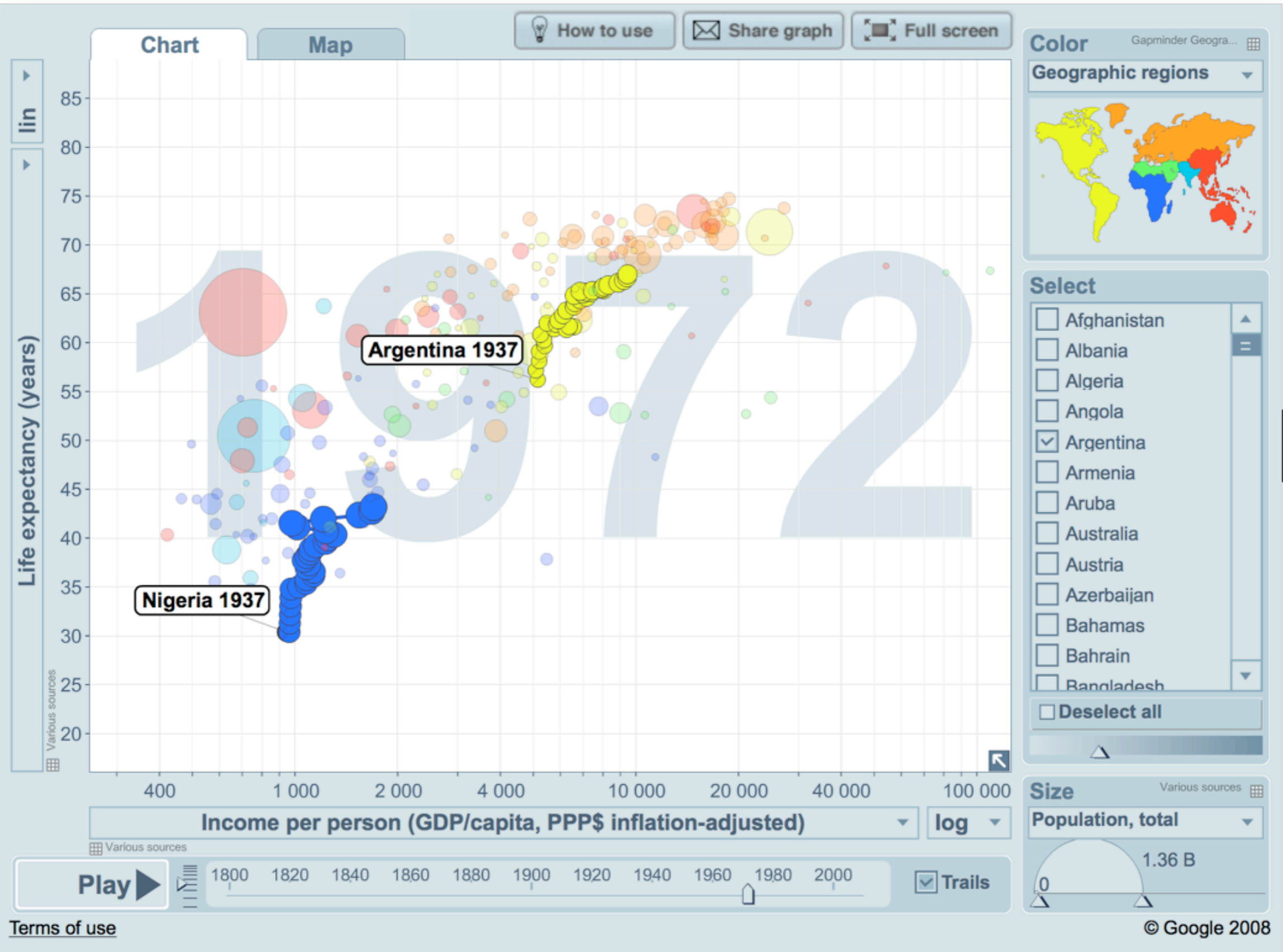


figures reproduced from Goodman 2012, “Principles of High-Dimensional Data Visualization in Astronomy”

Data • Dimensions • Display

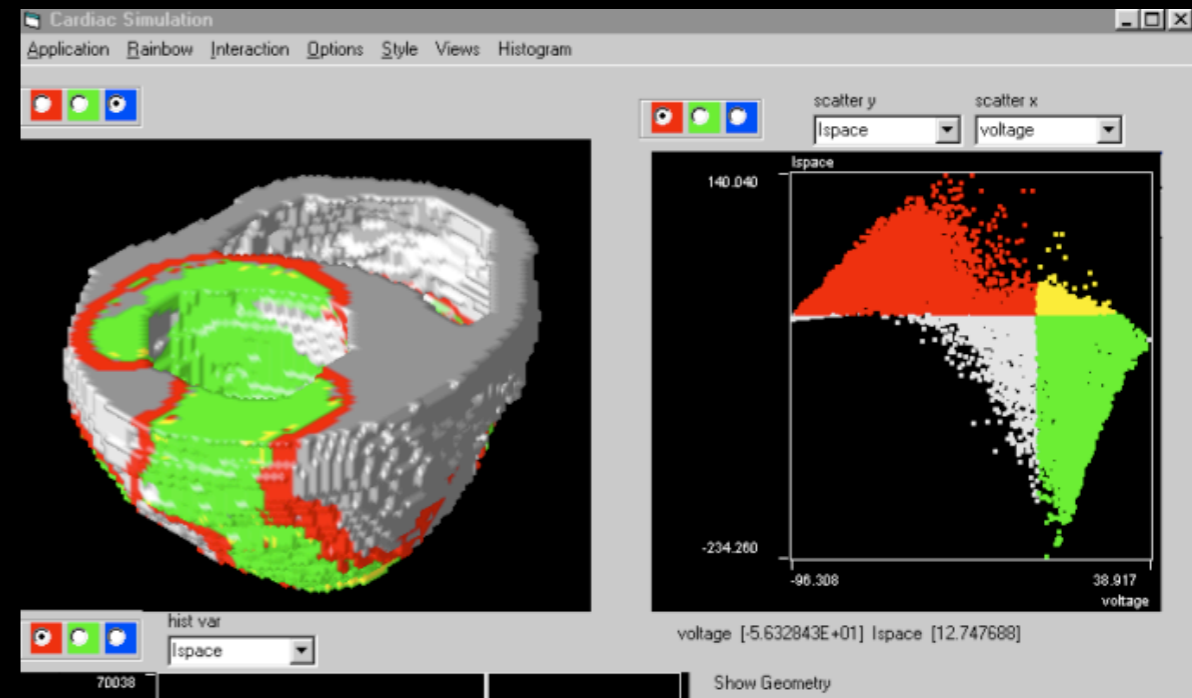
Linked Views

Linked Views

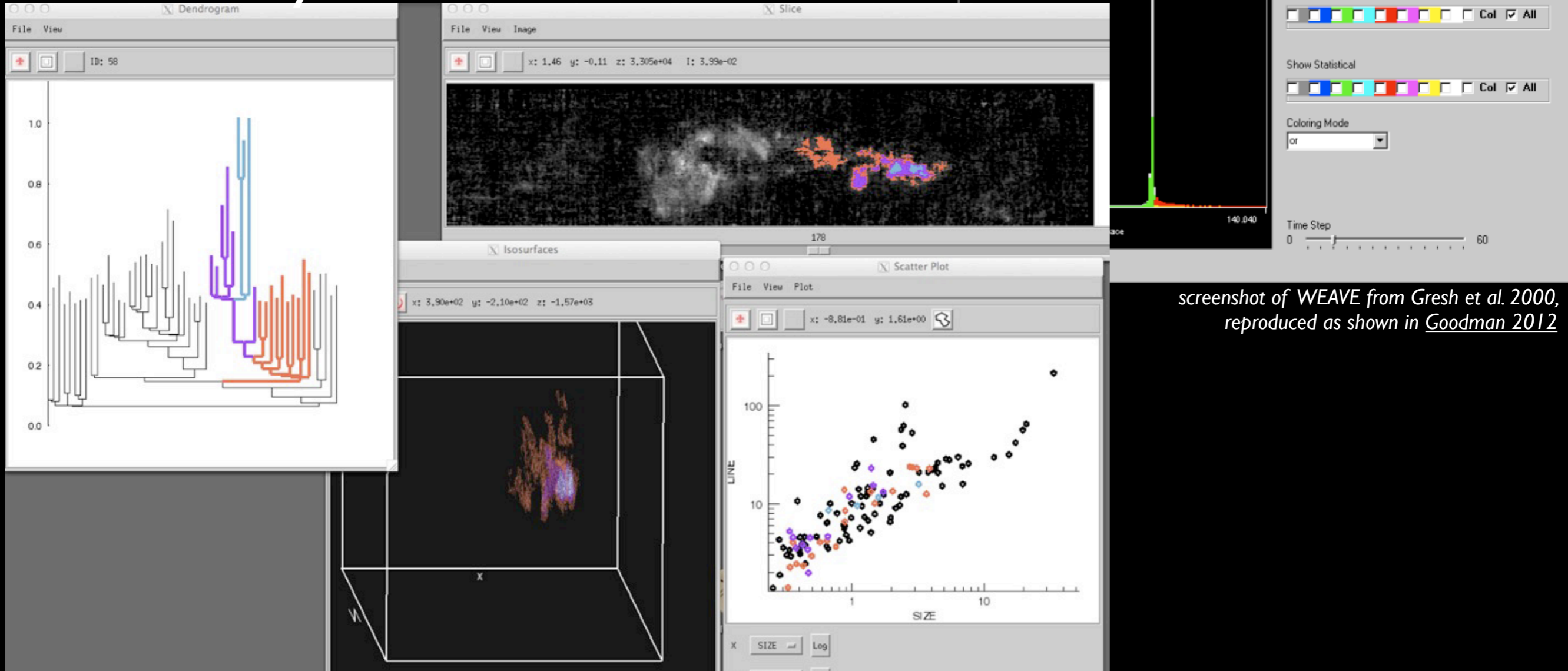


LINKING VIEWS “IN 3D”

Medicine



Astronomy

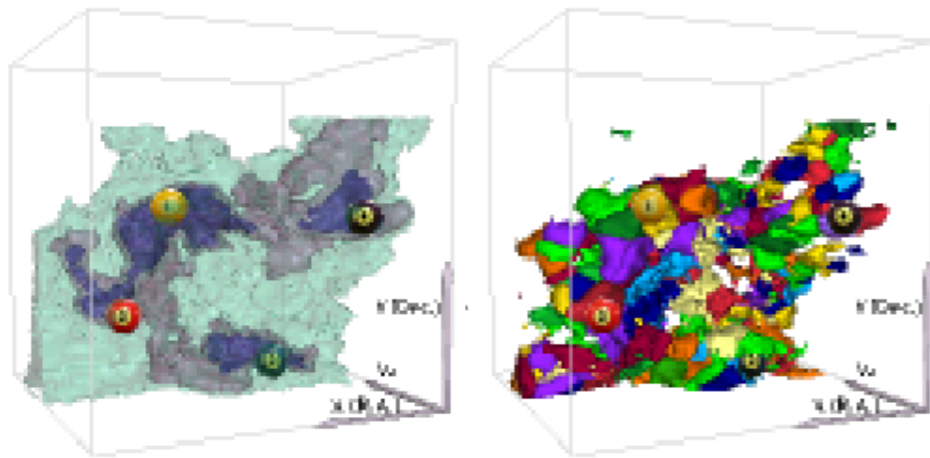


screenshot of WEAVE from Gresh et al. 2000, reproduced as shown in Goodman 2012

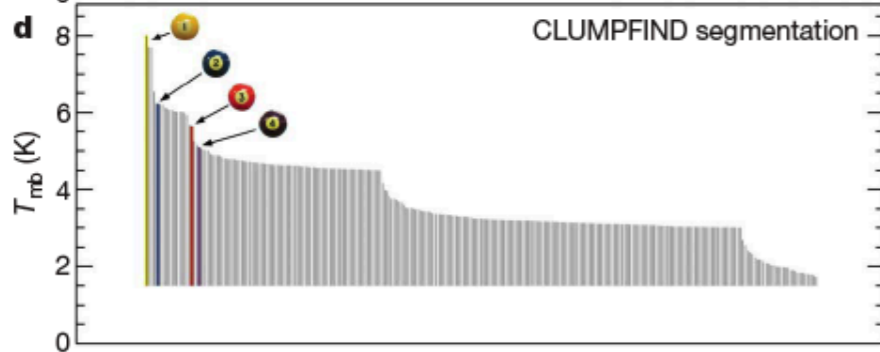
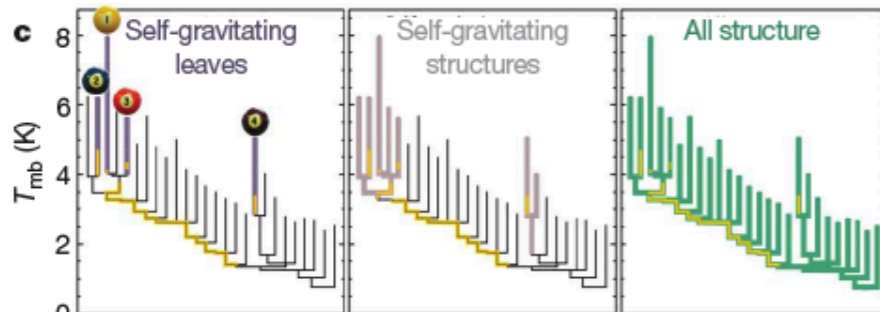
IDL “Dendroviz” screenshot, reproduced from Goodman 2012, “Principles of High-Dimensional Data Visualization in Astronomy” (Dendroviz, created by C. Beaumont, available for [download](#))

"High-dimensional" or "Multivariate" Data (Astronomy=Biology)

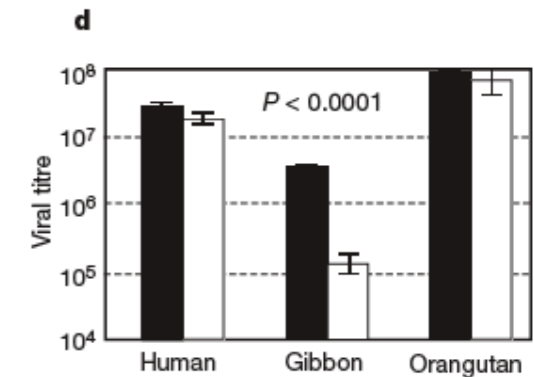
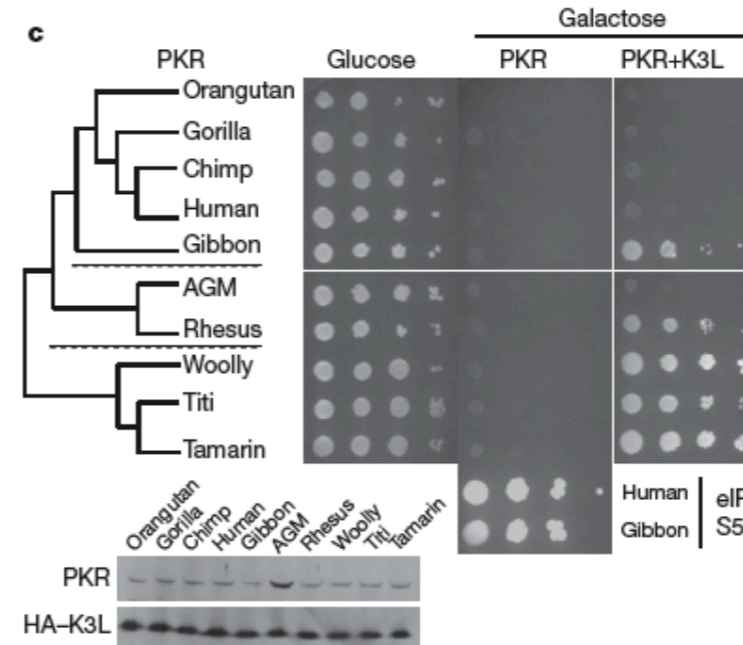
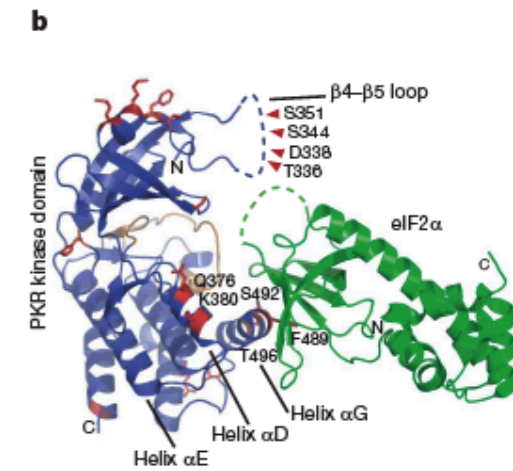
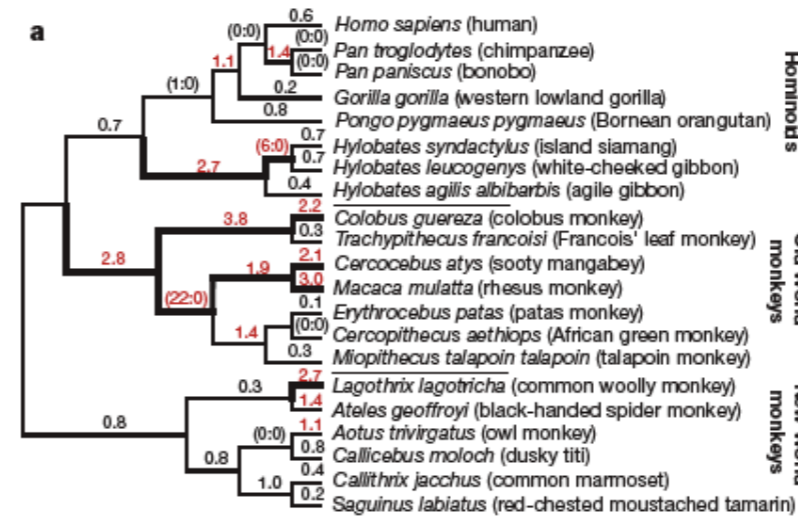
LETTERS



Click to rotate



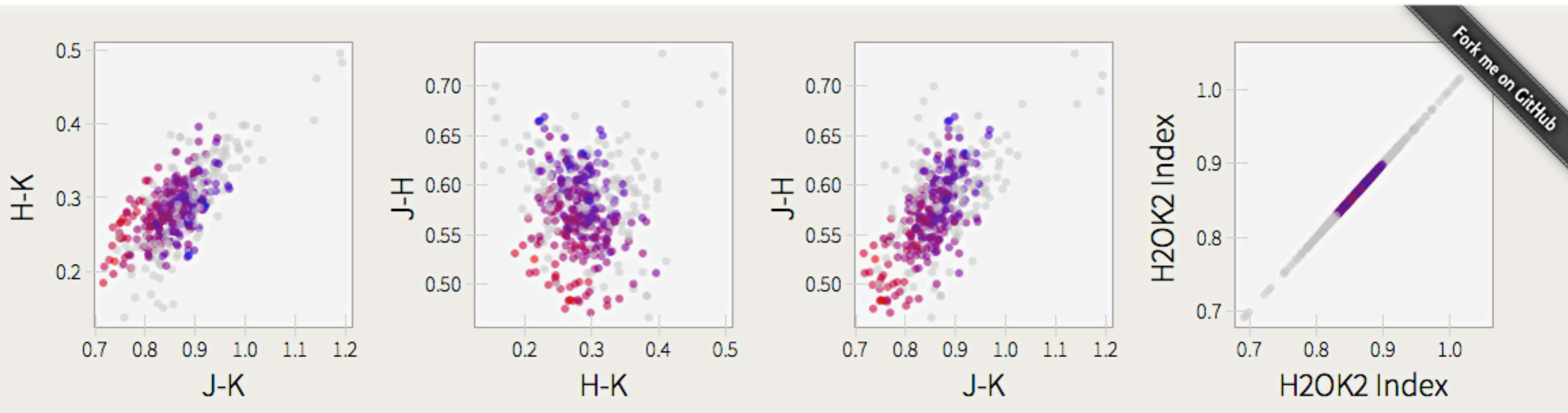
LETTERS



Goodman et al. *Nature*, 2009

Elde et al. *Nature*, 2008

d3po



All the stars

Early M dwarfs

Mid M dwarfs

Late M dwarfs

All the stars and their Na abundance

Early M dwarfs and their Na EW

Mid M dwarfs and their Na EW

...and finally the late M dwarfs and their Na EW.

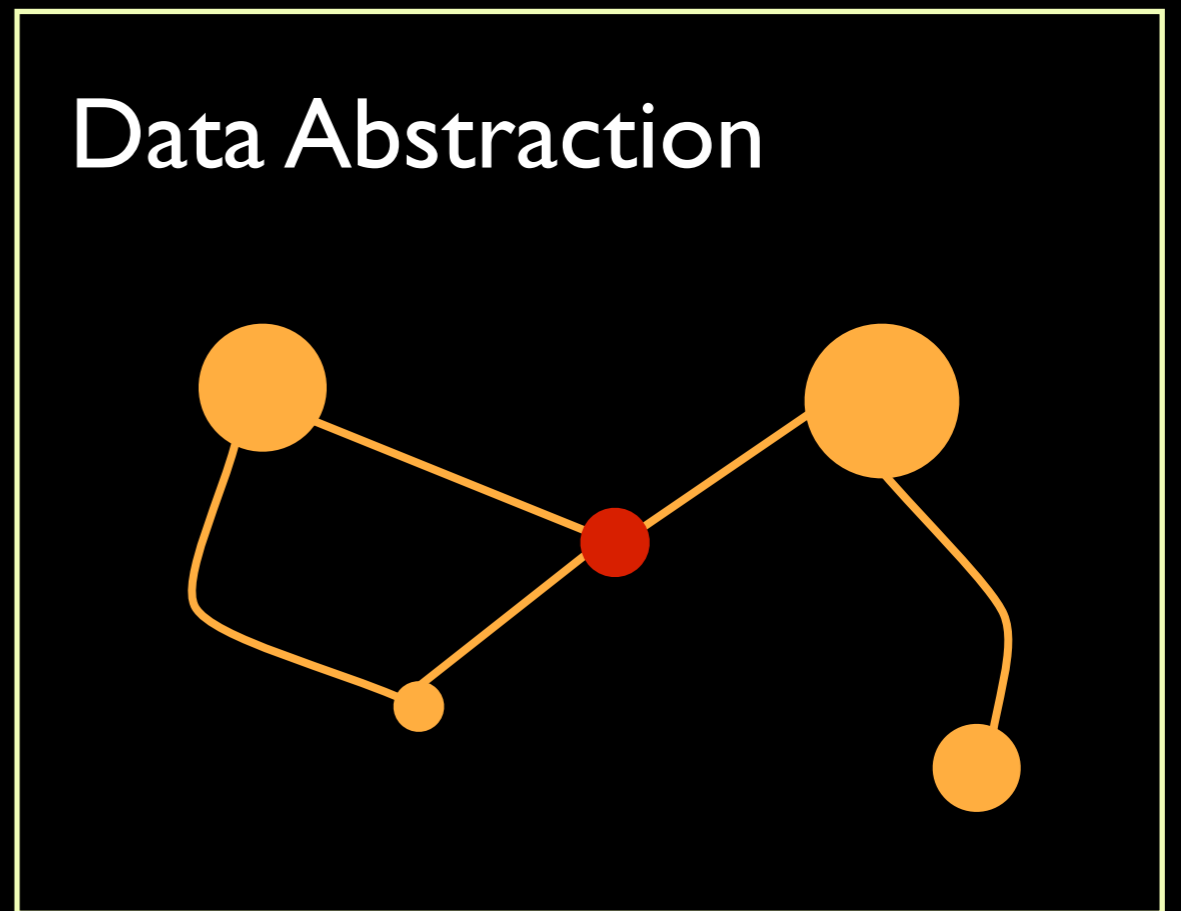
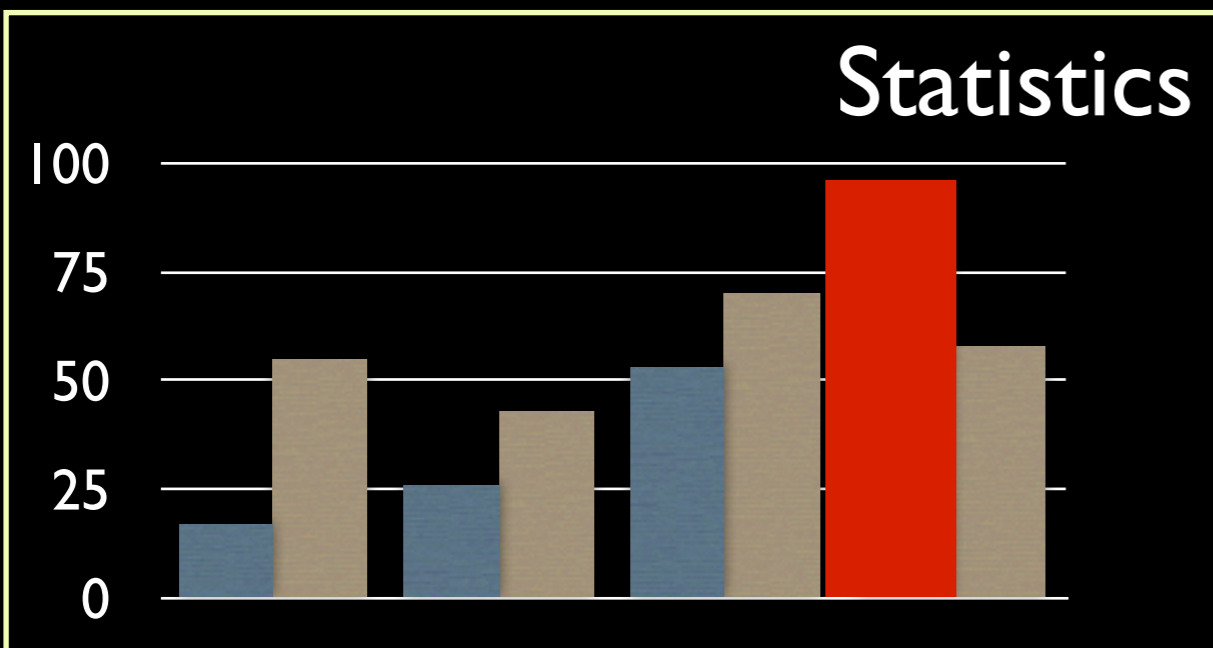
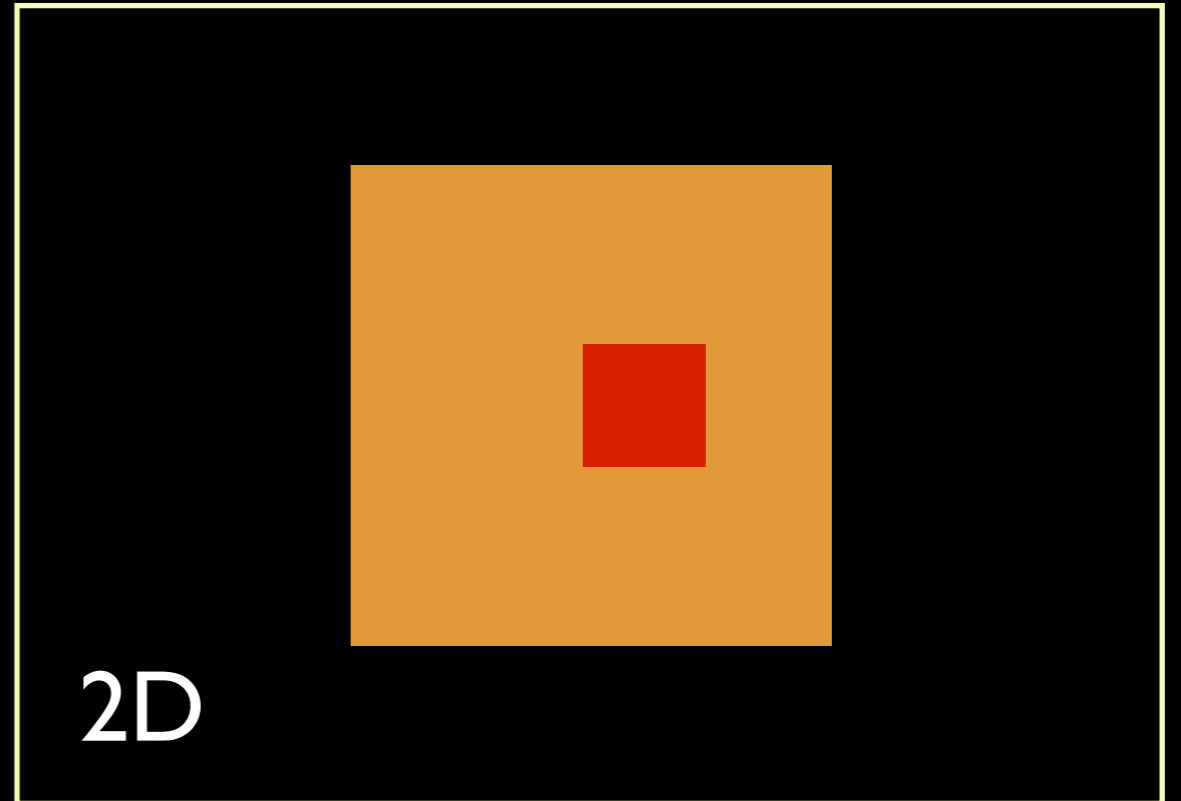
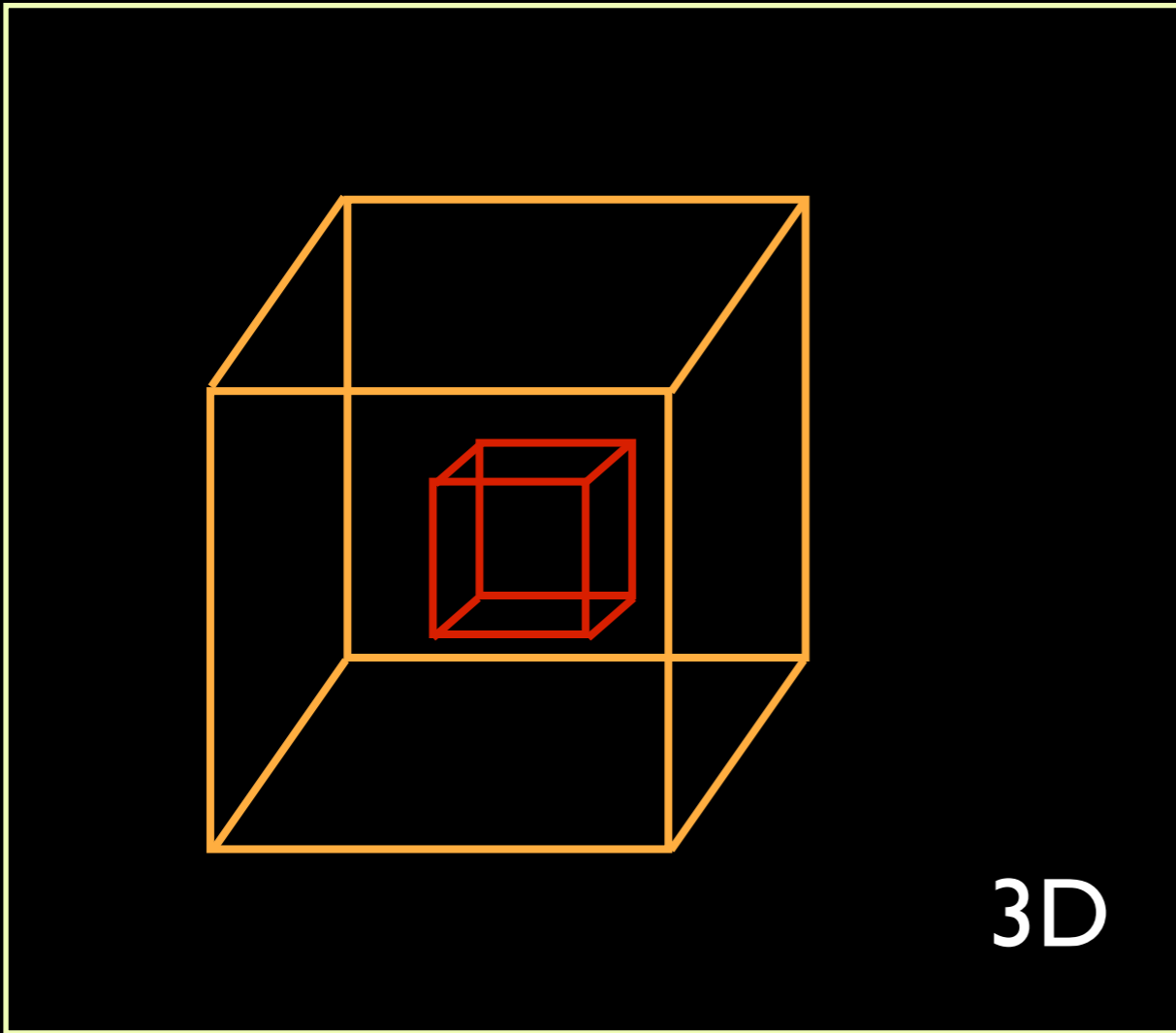
grid

...and the mid M and their EW Na...

A. M. Price-Whelan // J. E. G. Peek // E. R. Newton // M. Borkin // J. Allen // R. Angus // D. Muna // T. Staley //

.astronomy
September 2013

LINKED VIEWS OF HIGH-DIMENSIONAL DATA



figure, by M. Borkin, reproduced from [Goodman 2012](#), "Principles of High-Dimensional Data Visualization in Astronomy"

TUKEY'S "FOUR ESSENTIALS" OF LINKED VIEWS (C. 1972)

Watch the PRIM-9 video at: <http://stat-graphics.org/movies/prim9.html>



TUKEY'S "FOUR ESSENTIALS" OF LINKED VIEWS (C. 1972)

Picturing

Watch the PRIM-9 video at: <http://stat-graphics.org/movies/prim9.html>



TUKEY'S "FOUR ESSENTIALS" OF LINKED VIEWS (C. 1972)

Picturing

Rotation

Watch the PRIM-9 video at: <http://stat-graphics.org/movies/prim9.html>



TUKEY'S "FOUR ESSENTIALS" OF LINKED VIEWS (C. 1972)

Picturing

Rotation

Isolation

Watch the PRIM-9 video at: <http://stat-graphics.org/movies/prim9.html>



TUKEY'S "FOUR ESSENTIALS" OF LINKED VIEWS (C. 1972)

Picturing

Rotation

Isolation

Masking

Watch the PRIM-9 video at: <http://stat-graphics.org/movies/prim9.html>



TUKEY'S "FOUR ESSENTIALS" OF LINKED VIEWS (C. 1972)

Picturing

Rotation

Isolation

Masking

Selection

Watch the PRIM-9 video at: <http://stat-graphics.org/movies/prim9.html>



TUKEY'S "FOUR ESSENTIALS" OF LINKED VIEWS (C. 1972)

Picturing

Rotation

Isolation

Masking

Selection

and these *"need to work together"*
in a *"dynamic display"*

Brushing

Linking

[Watch the PRIM-9 video at: http://stat-graphics.org/movies/prim9.html](http://stat-graphics.org/movies/prim9.html)



TUKEY'S "FOUR ESSENTIALS" OF LINKED VIEWS (C. 1972)

Picturing

Rotation

Isolation

Masking

Selection

and these *"need to work together"*
in a *"dynamic display"*

Brushing

Linking

Results...

1. for immediate **insight**
2. as visual source of **ideas** for statistical algorithms (...relation to SVM)

[Watch the PRIM-9 video at: http://stat-graphics.org/movies/prim9.html](http://stat-graphics.org/movies/prim9.html)



TUKEY'S "FOUR ESSENTIALS" OF LINKED VIEWS (C. 1972)

Picturing

Rotation

Isolation

Masking

Selection

and these *"need to work together"*
in a *"dynamic display"*

Brushing

Linking

Results...

1. for immediate **insight**
2. as visual source of **ideas** for statistical algorithms (...relation to SVM)

Warning

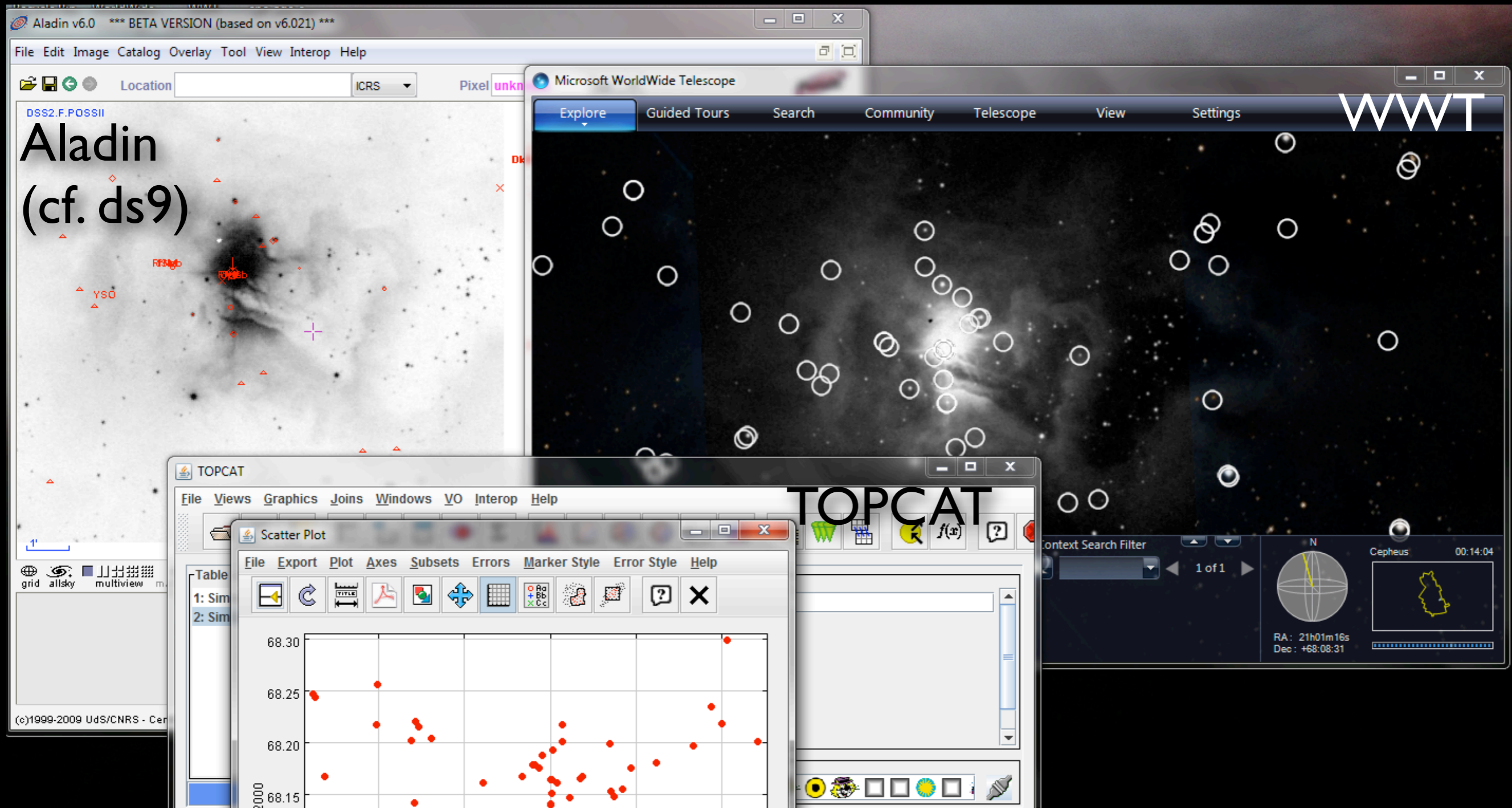
"details of control can make or break such a system"

Watch the PRIM-9 video at: <http://stat-graphics.org/movies/prim9.html>

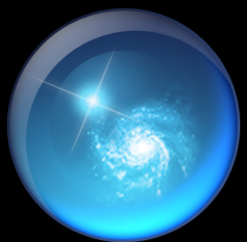


LINKING VIEWS USING SAMP

(SAMP CREATOR IS MARK TAYLOR)



figure, showing SAMP screenshot, reproduced from [Goodman 2012](#), "Principles of High-Dimensional Data Visualization in Astronomy"



Microsoft® Research WorldWide Telescope

worldwidetelescope.org

The screenshot shows the main interface of WorldWide Telescope. At the top, there are navigation tabs: **Explore**, **Guided Tours**, **Search**, **View**, and **Settings**. Below these is a 'Collections' bar with 'All-Sky Surveys' selected, showing a row of image thumbnails: Digitized Sky Survey, VLSS: VLA Low-frequency Sky Survey, WMAP ILC 5-Year Cosmic Microwave Background, SFD Dust Map (Infrared), IRIS: Improved Resolution, 2MASS: Two Micron All Sky Survey, and Hydrogen Alpha Filter. The main view is a 3D sky map with a central circular field of view showing a galaxy. A 'Finder Scope' window is open, displaying details for NGC224, including its classification as a 'Spiral Galaxy in Andromeda' and its coordinates (RA: 00h42m42s, Dec: 41:16:00). Below the main view is a 'Look At' panel with 'Sky' selected, and a 'Context bar' showing thumbnails for NGC221 and M31. A 'Context globe' on the right shows the current field of view on a celestial sphere. At the bottom, there are buttons for 'Research', 'Show Object', and 'Close', along with a '3D' button. A small image of a telescope is at the bottom center.

Seamlessly explore imagery from the best ground and space-based telescopes in the world

Expert led tours of the Universe

Control time to study how the night sky changes

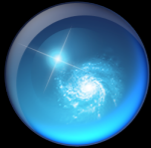
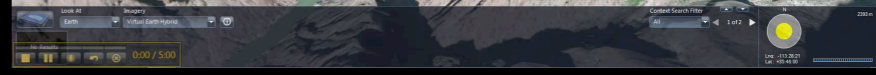
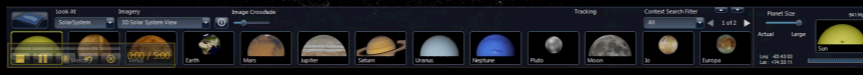
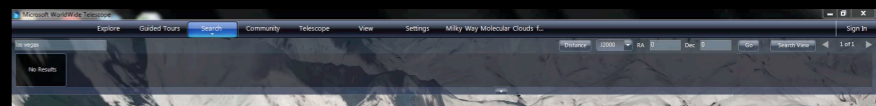
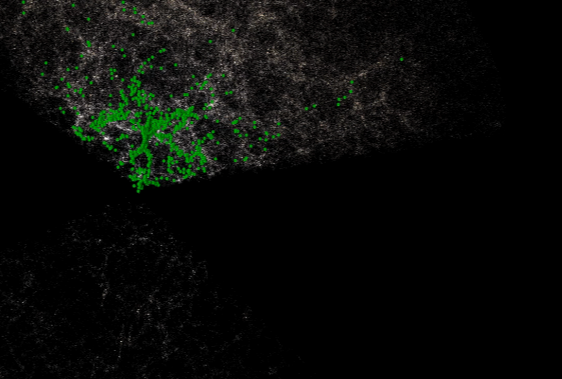
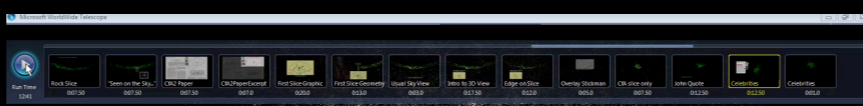
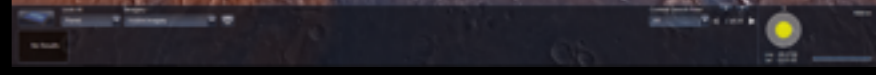
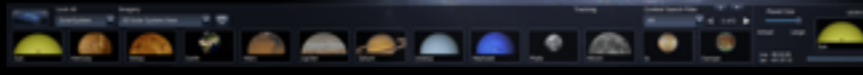
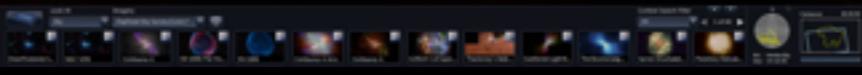
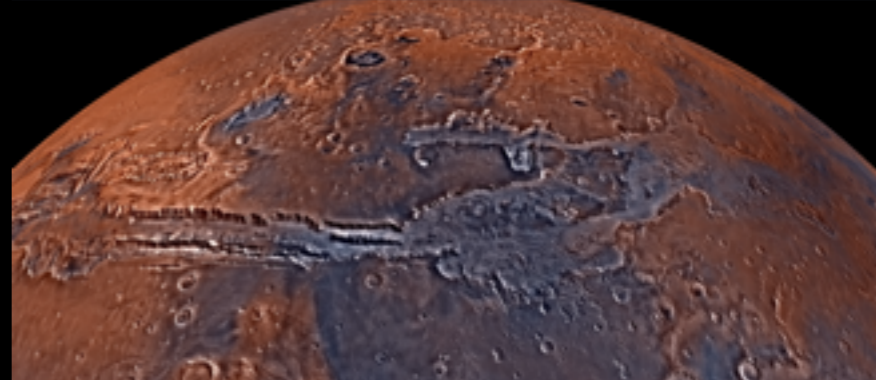
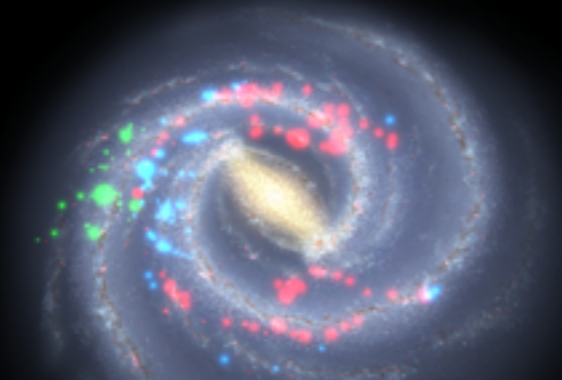
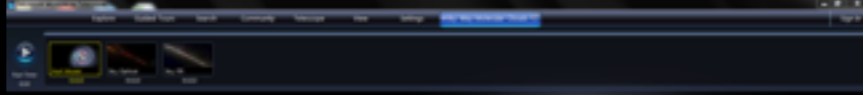
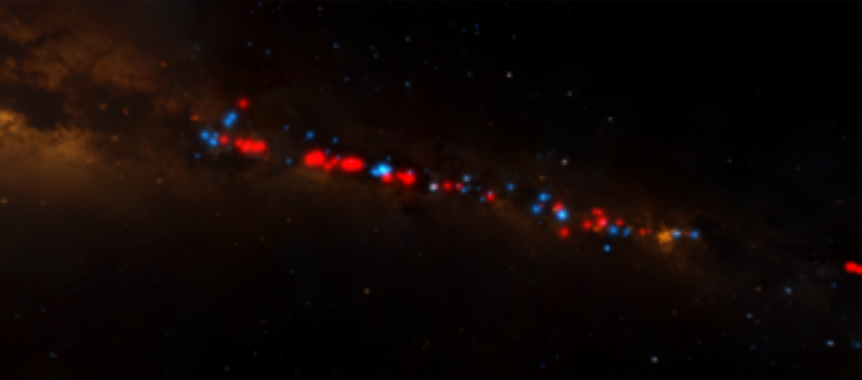
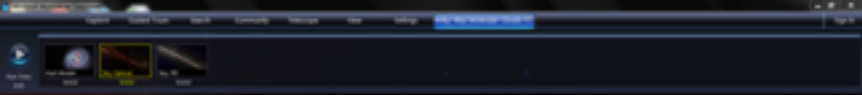
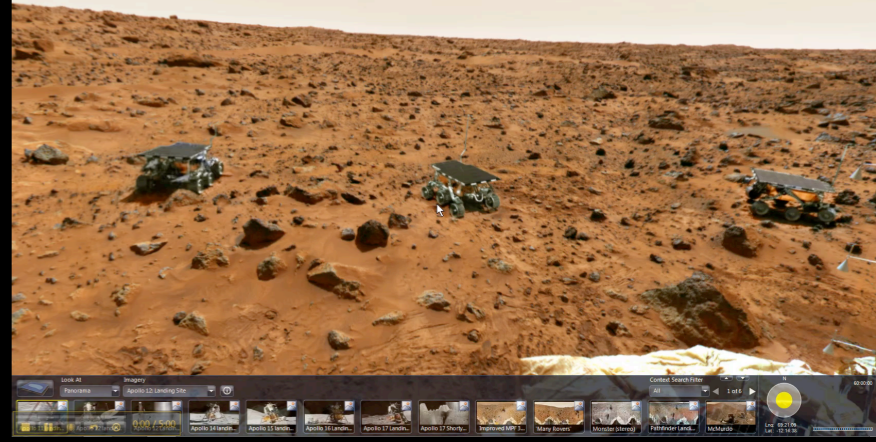
View and compare images from across the electromagnetic spectrum

Much more than "just" the sky at night! 3D features can take you to other planets, stars & galaxies.

Finder Scope links to Wikipedia, publications, and data, so you can learn more

Context bar shows items of interest in current field of view

Context globe shows where you're looking.



Experience WWT at worldwidetelescope.org



glueviz.org

What is glue?

Glue 0.1 documentation > next index

Glue Documentation

glue
multidimensional data exploration

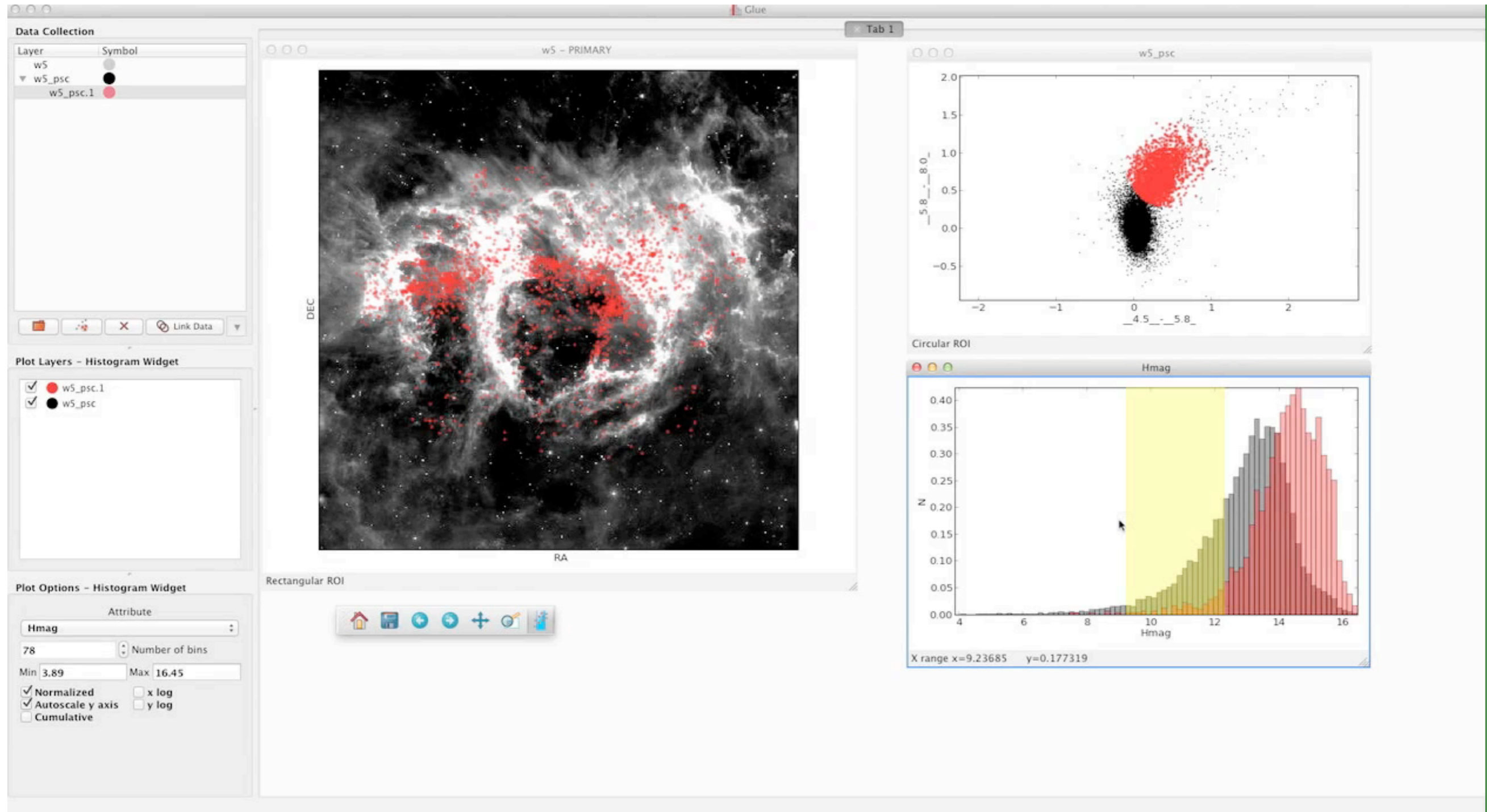
Glue is a Python library to explore relationships within and among related datasets. Its main features include:

- **Linked Statistical Graphics.** With Glue, users can create scatter plots, histograms and images (2D and 3D) of their data. Glue is focused on the brushing and linking paradigm, where selections in any graph propagate to all others.
- **Flexible linking across data.** Glue uses the logical links that exist between different data sets to overlay visualizations of different data, and to propagate selections across data sets. These links are specified by the user, and are arbitrarily flexible.
- **Full scripting capability.** Glue is written in Python, and built on top of its standard scientific libraries (i.e., Numpy, Matplotlib, Scipy). Users can easily integrate their own python code for data input, cleaning, and analysis.

[the film!]

What is glue?

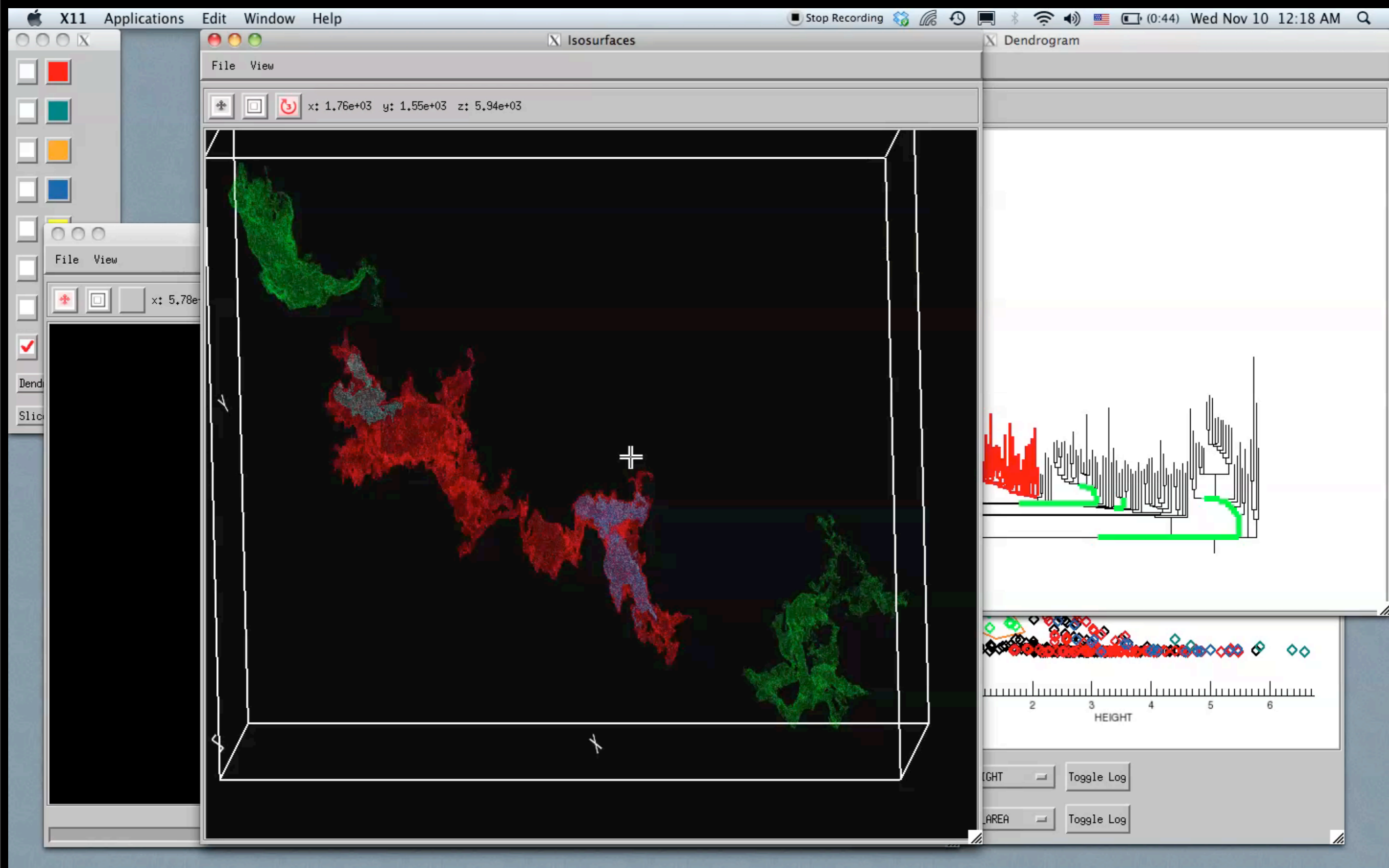
What is glue?



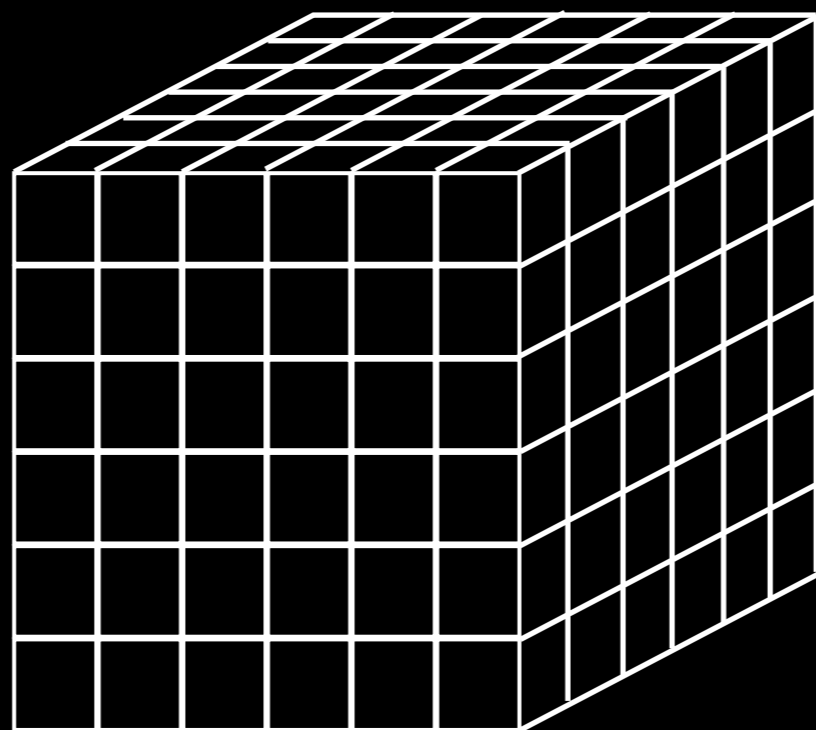
Before Glue: **Linked** Dendrogram **Views** in IDL

*Video & implementation: Christopher Beaumont, CfA/UHawaii;
inspired by AstroMed work of Douglas Alan, Michelle Borkin, AG, Michael Halle, Erik Rosolowsky*

Before Glue: Linked Dendrogram Views in IDL



*Video & implementation: Christopher Beaumont, CfA/UHawaii;
inspired by AstroMed work of Douglas Alan, Michelle Borkin, AG, Michael Halle, Erik Rosolowsky*



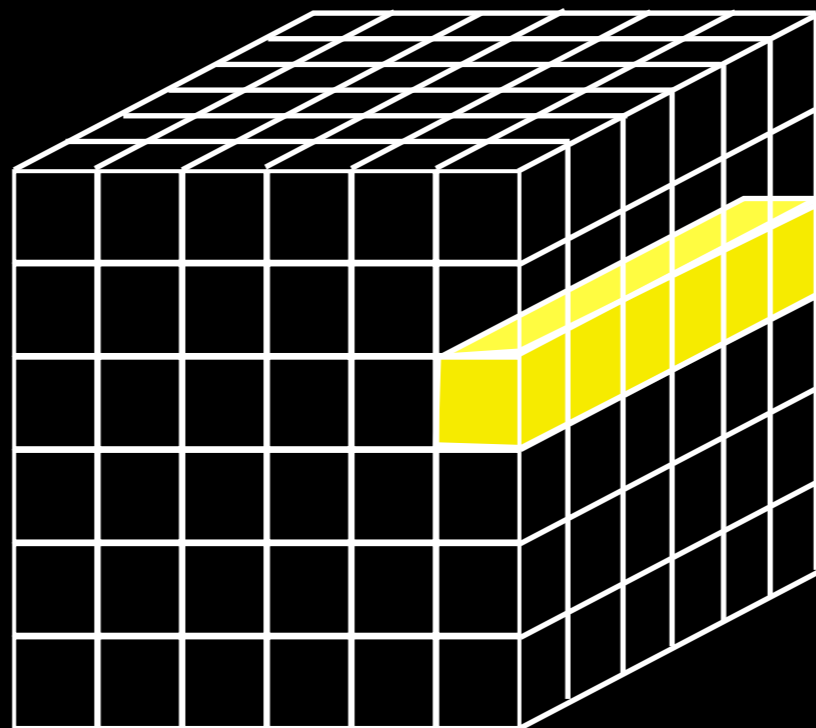
GENERALLY

1D: Columns = “Spectra”, “SEDs” or “Time Series”

2D: Faces or Slices = “Images”

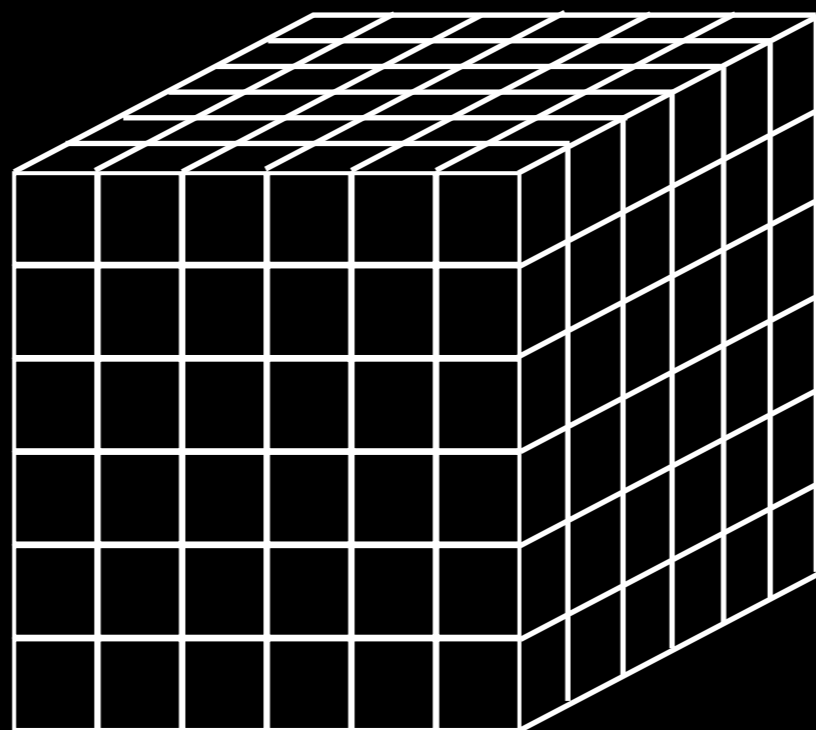
3D: Volumes = “3D Renderings”, “2D Movies”

4D: Time Series of Volumes = “3D Movies”



GENERALLY

- 1D:** Columns = “Spectra”, “SEDs” or “Time Series”
- 2D:** Faces or Slices = “Images”
- 3D:** Volumes = “3D Renderings”, “2D Movies”
- 4D:** Time Series of Volumes = “3D Movies”



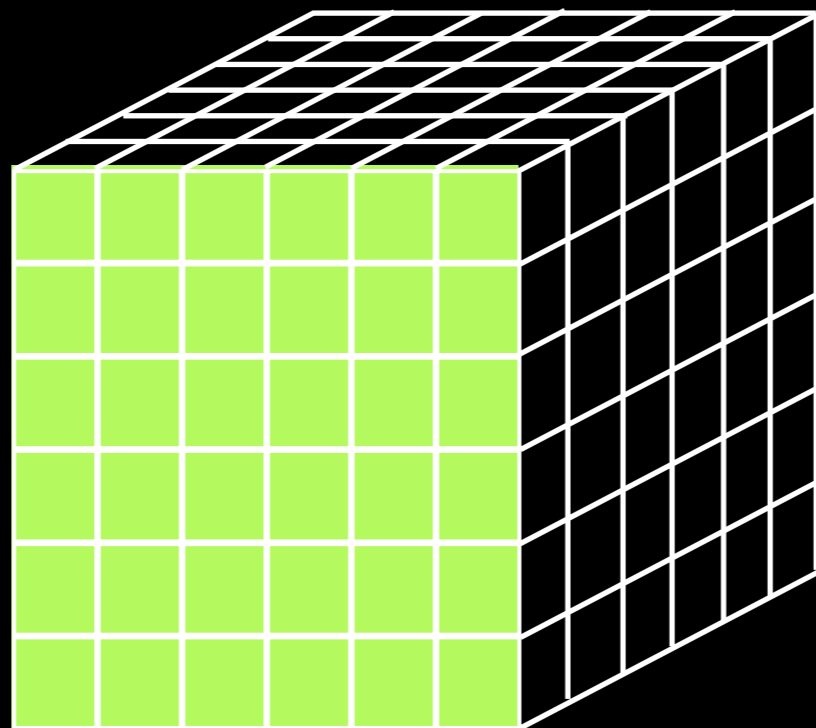
GENERALLY

1D: Columns = “Spectra”, “SEDs” or “Time Series”

2D: Faces or Slices = “Images”

3D: Volumes = “3D Renderings”, “2D Movies”

4D: Time Series of Volumes = “3D Movies”



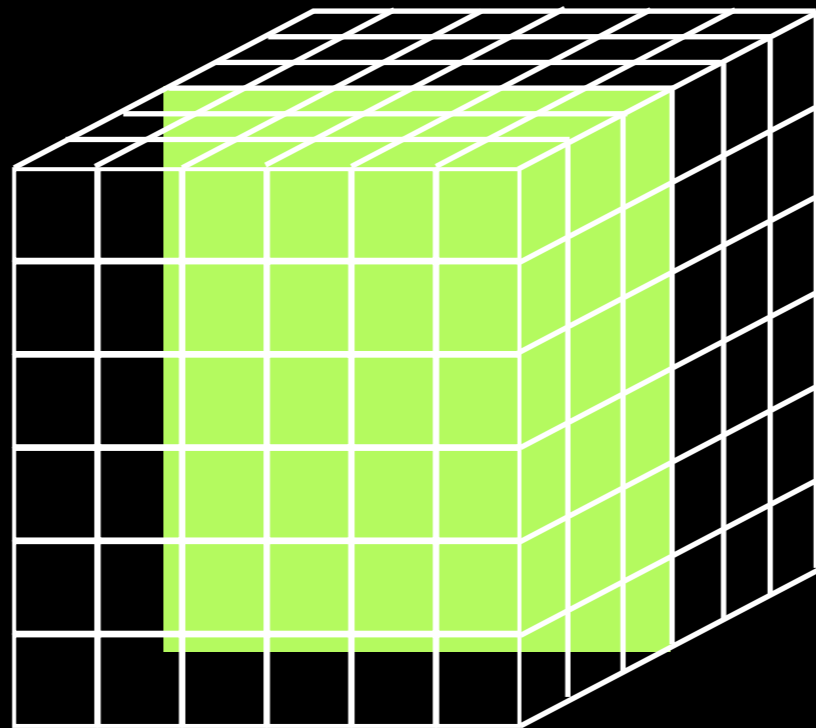
GENERALLY

1D: Columns = “Spectra”, “SEDs” or “Time Series”

2D: Faces or Slices = “Images”

3D: Volumes = “3D Renderings”, “2D Movies”

4D: Time Series of Volumes = “3D Movies”



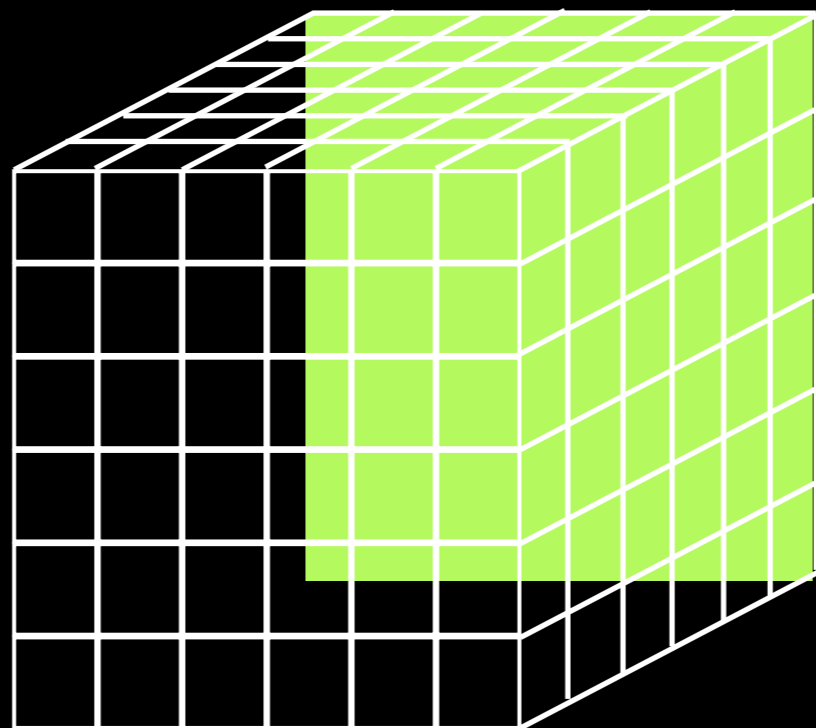
GENERALLY

1D: Columns = “Spectra”, “SEDs” or “Time Series”

2D: Faces or Slices = “Images”

3D: Volumes = “3D Renderings”, “2D Movies”

4D: Time Series of Volumes = “3D Movies”



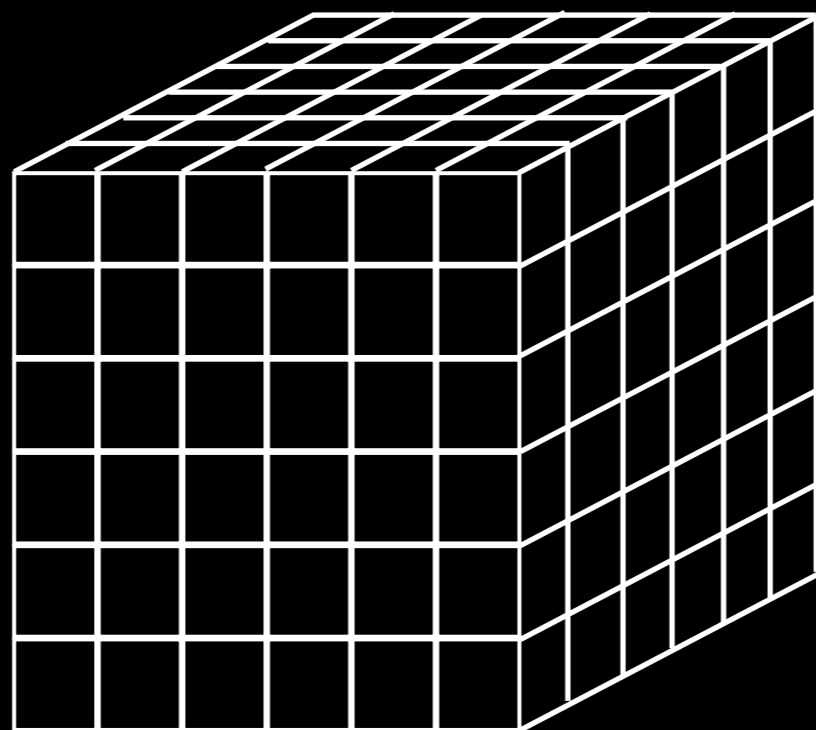
GENERALLY

1D: Columns = “Spectra”, “SEDs” or “Time Series”

2D: Faces or Slices = “Images”

3D: Volumes = “3D Renderings”, “2D Movies”

4D: Time Series of Volumes = “3D Movies”



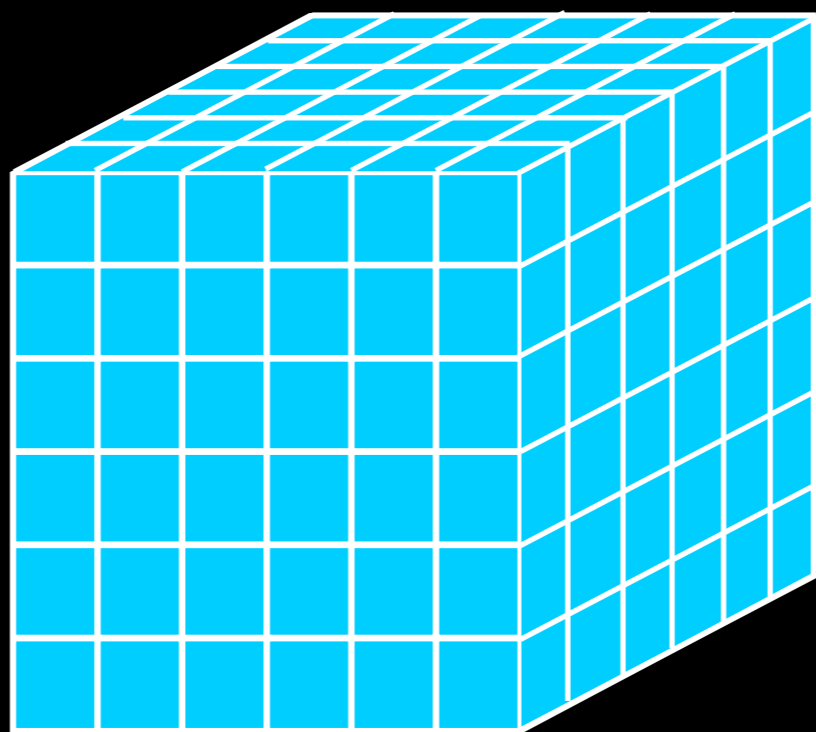
GENERALLY

1D: Columns = “Spectra”, “SEDs” or “Time Series”

2D: Faces or Slices = “Images”

3D: Volumes = “3D Renderings”, “2D Movies”

4D: Time Series of Volumes = “3D Movies”



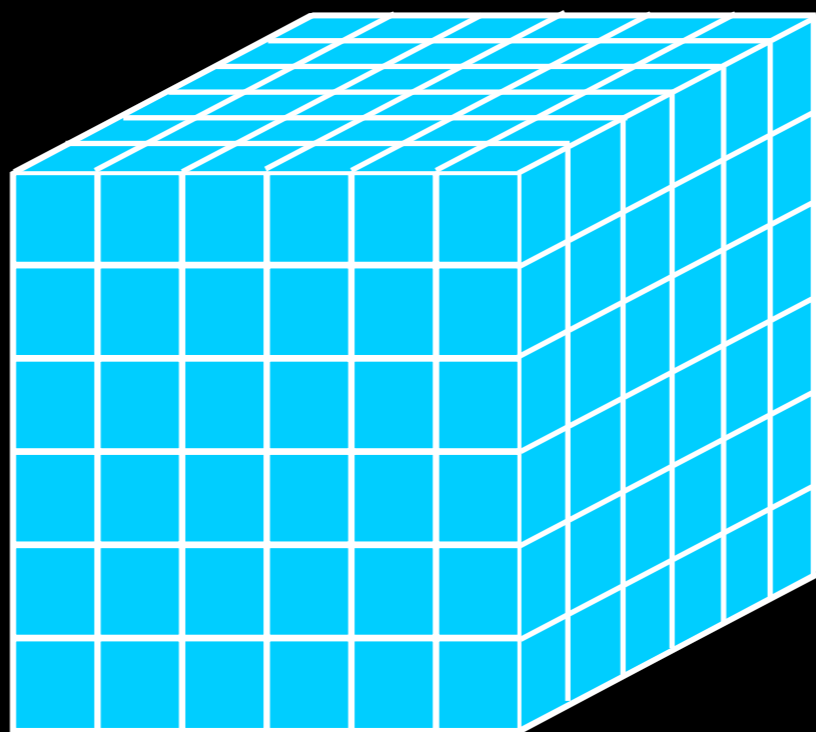
GENERALLY

1D: Columns = “Spectra”, “SEDs” or “Time Series”

2D: Faces or Slices = “Images”

3D: Volumes = “3D Renderings”, “2D Movies”

4D: Time Series of Volumes = “3D Movies”



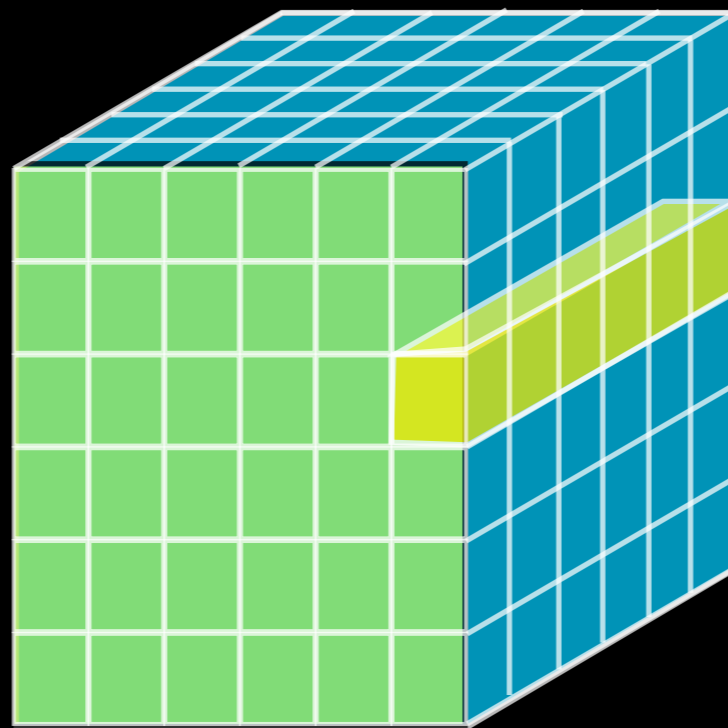
GENERALLY

1D: Columns = “Spectra”, “SEDs” or “Time Series”

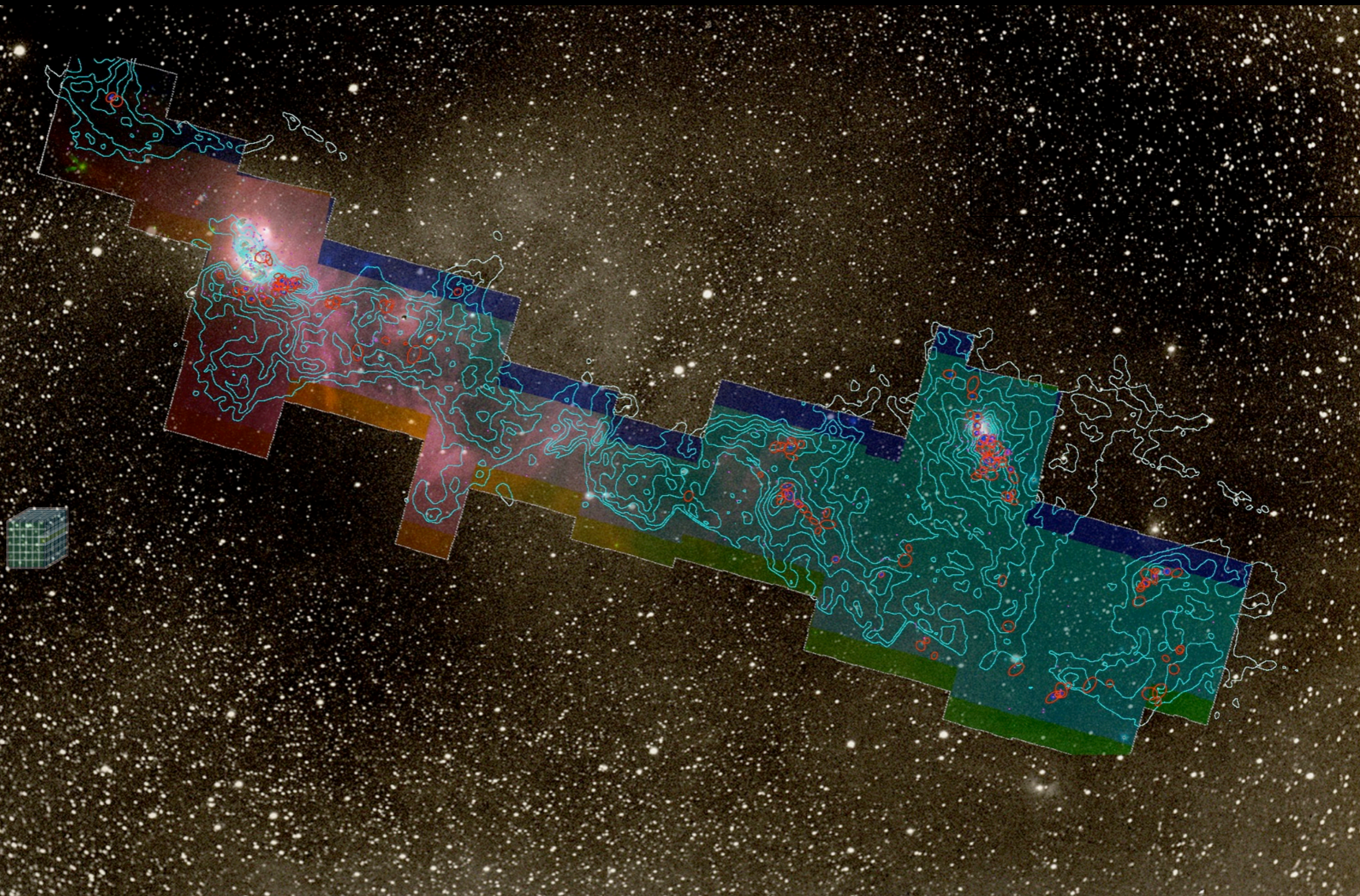
2D: Faces or Slices = “Images”

3D: Volumes = “3D Renderings”, “2D Movies”

4D: Time Series of Volumes = “3D Movies”








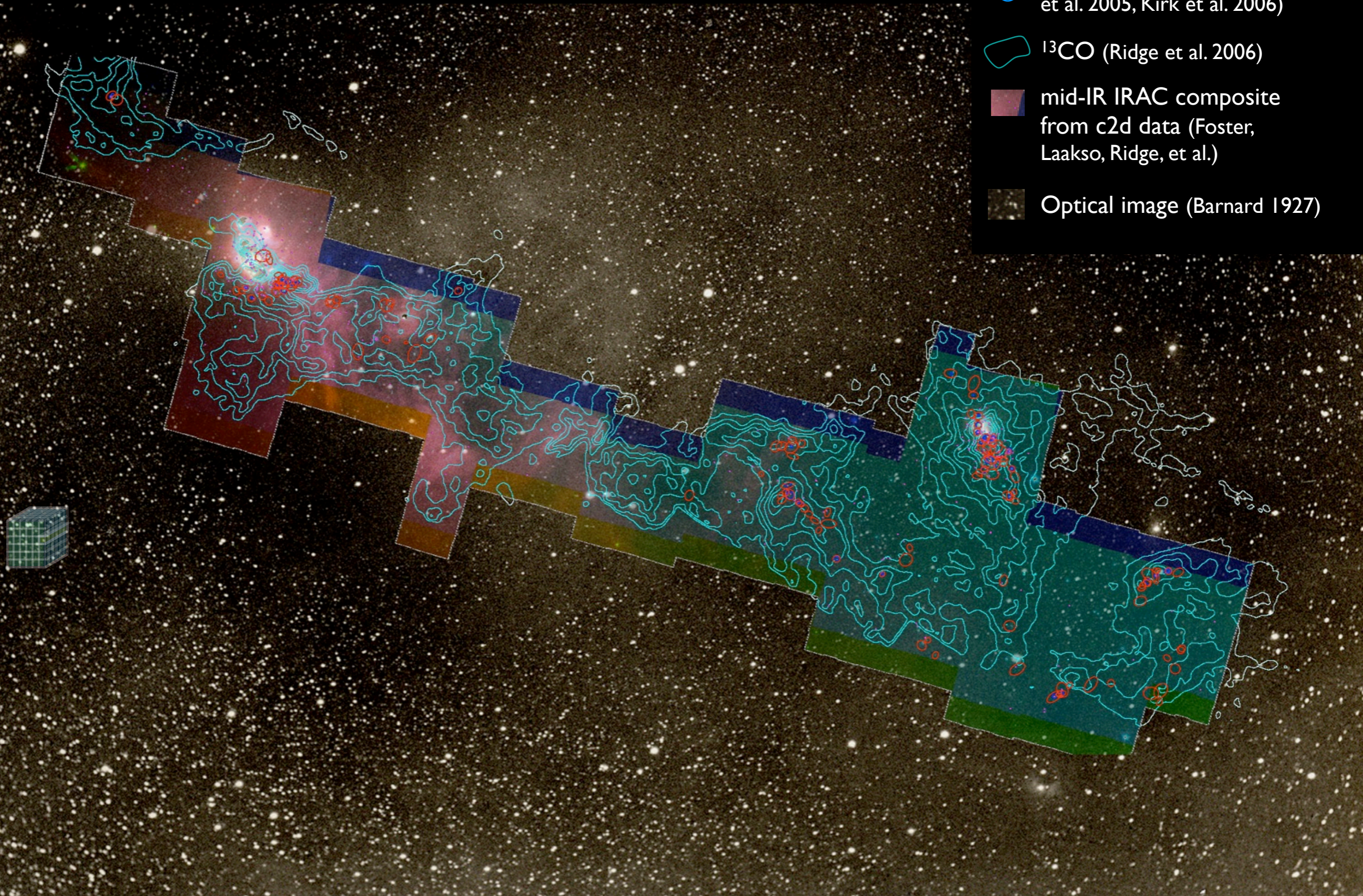
Star Formation in Perseus

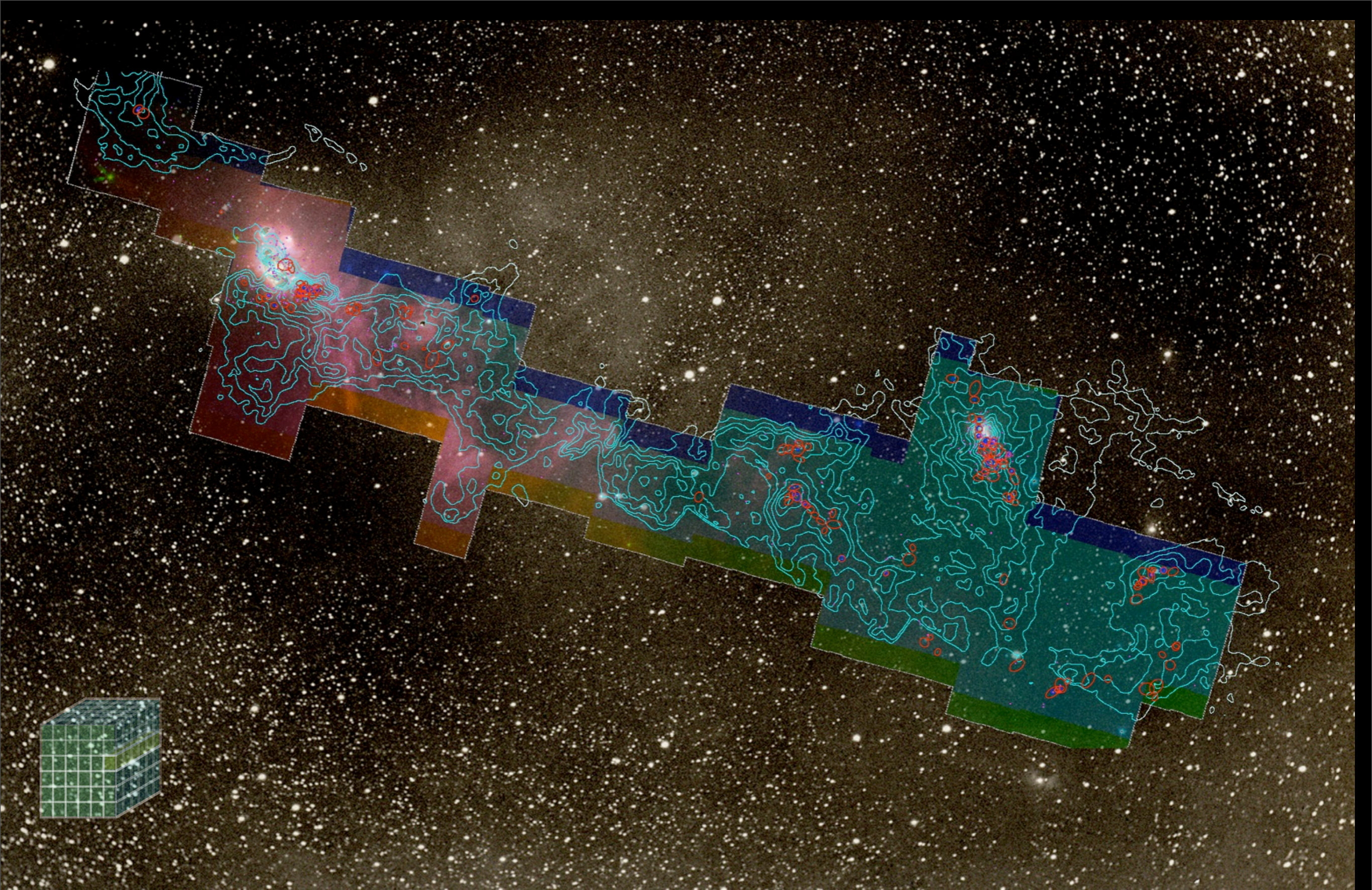


Star Formation in Perseus

COMPLETE

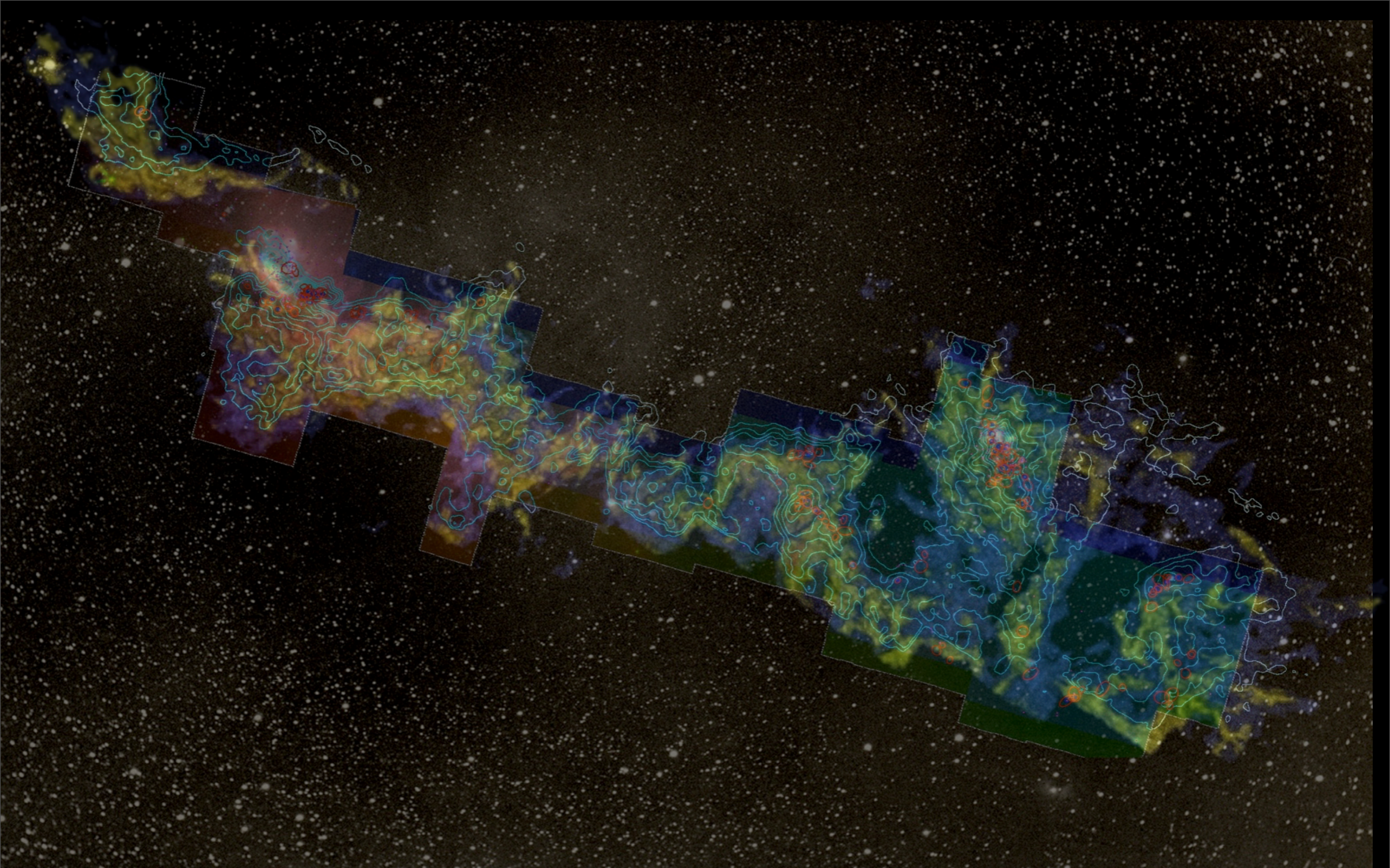
-  mm peak (Enoch et al. 2006)
-  sub-mm peak (Hatchell et al. 2005, Kirk et al. 2006)
-  ^{13}CO (Ridge et al. 2006)
-  mid-IR IRAC composite from c2d data (Foster, Laakso, Ridge, et al.)
-  Optical image (Barnard 1927)





AstronomicalMedicine@iig

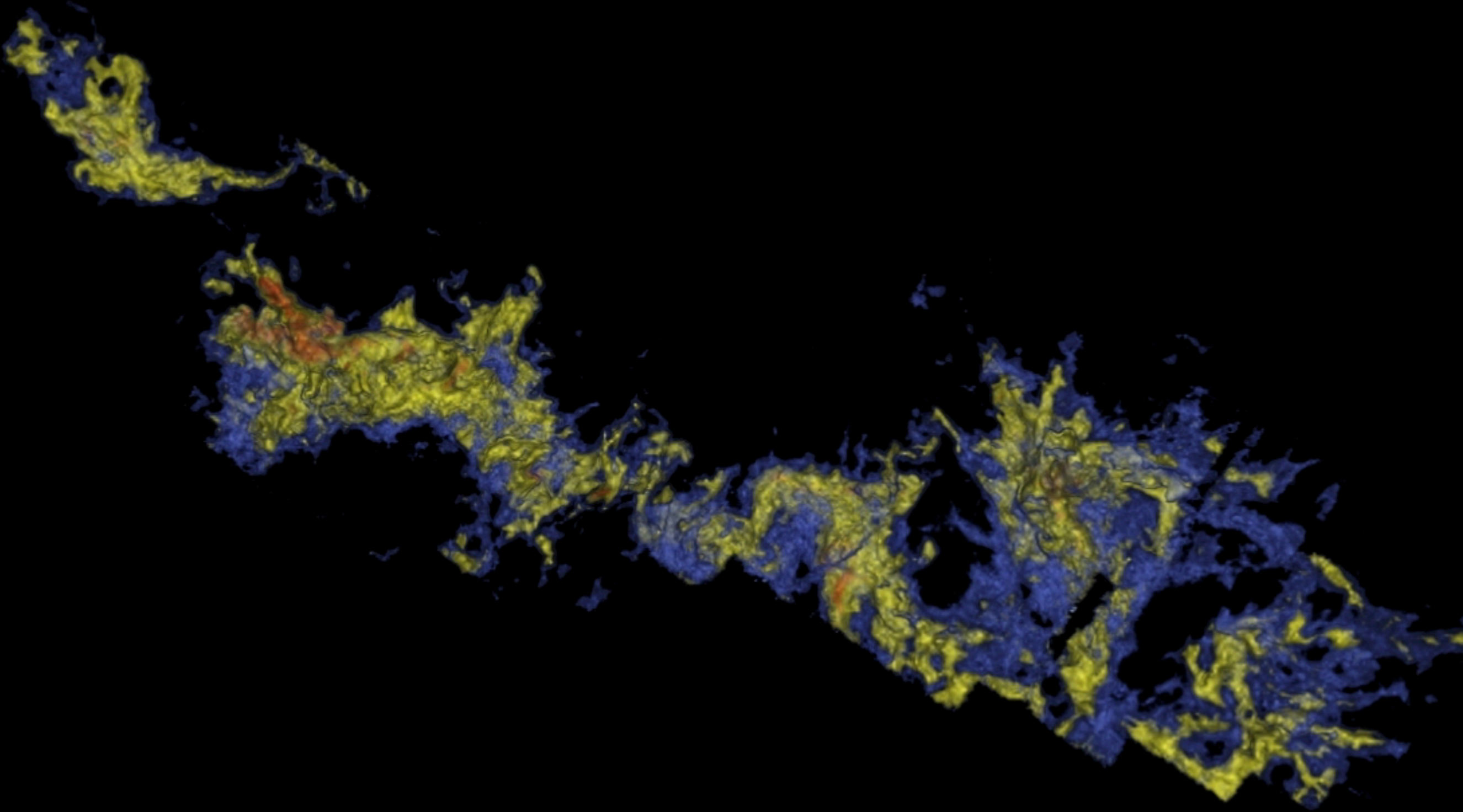
COMPLETE



3D Viz made with VolView

Astronomical**Medicine**@iic

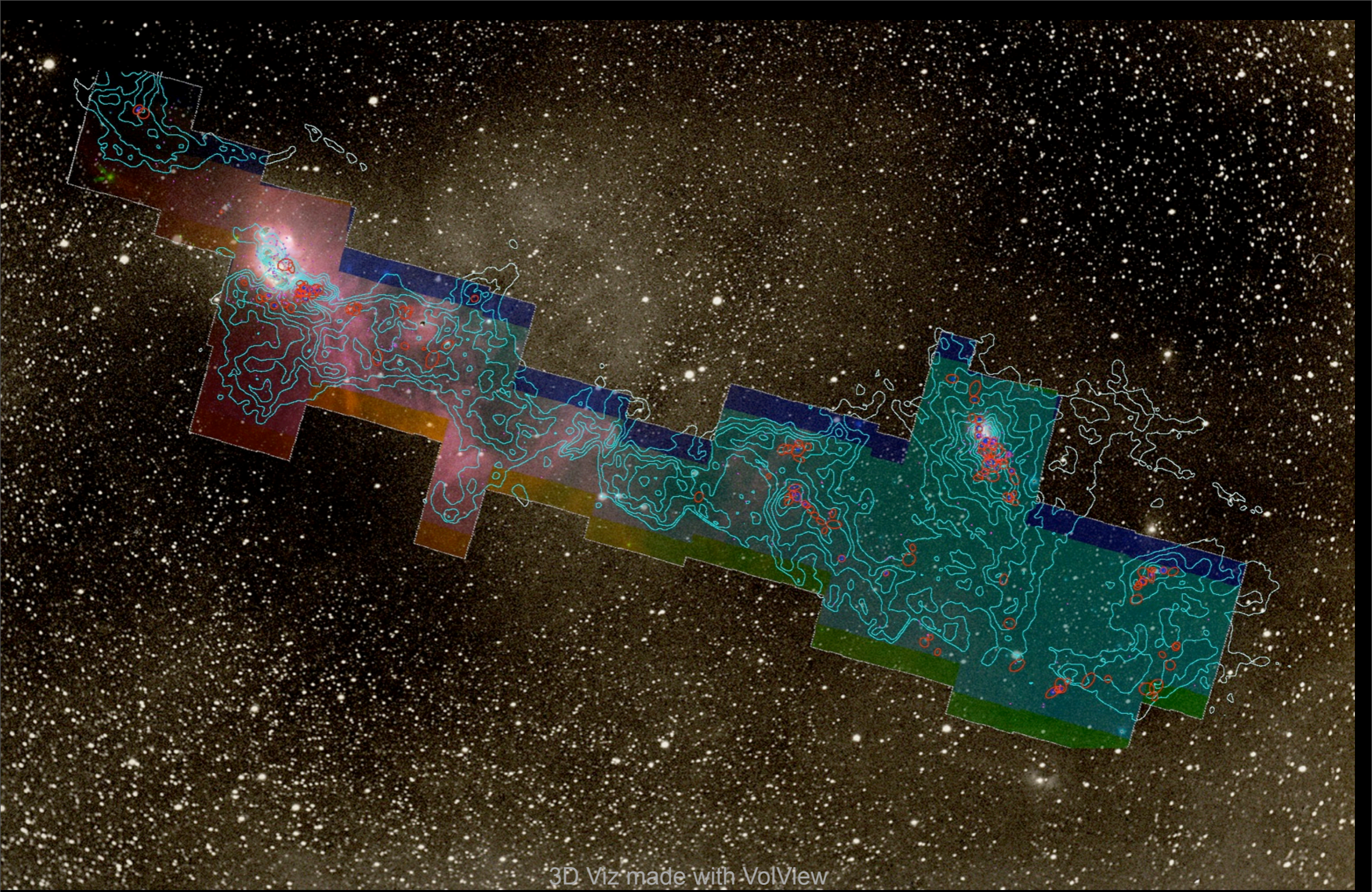
COMPLETE



3D Viz made with VolView

AstronomicalMedicine@iic

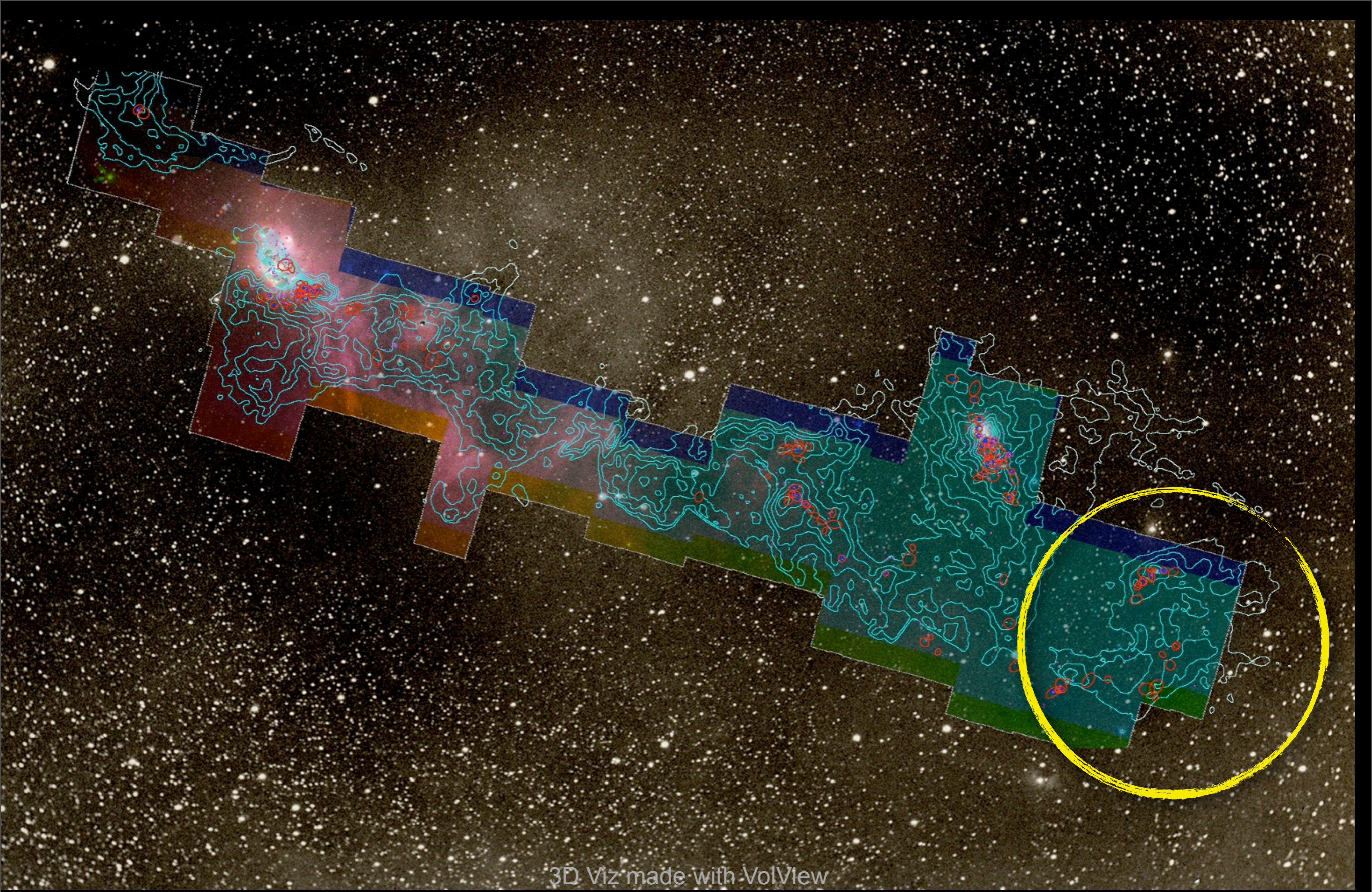
COMPLETE



3D Viz made with VolView

AstronomicalMedicine@iic

COMPLETE



3D Viz made with VolView

AstronomicalMedicine@iic

COMPLETE

Epidemiology

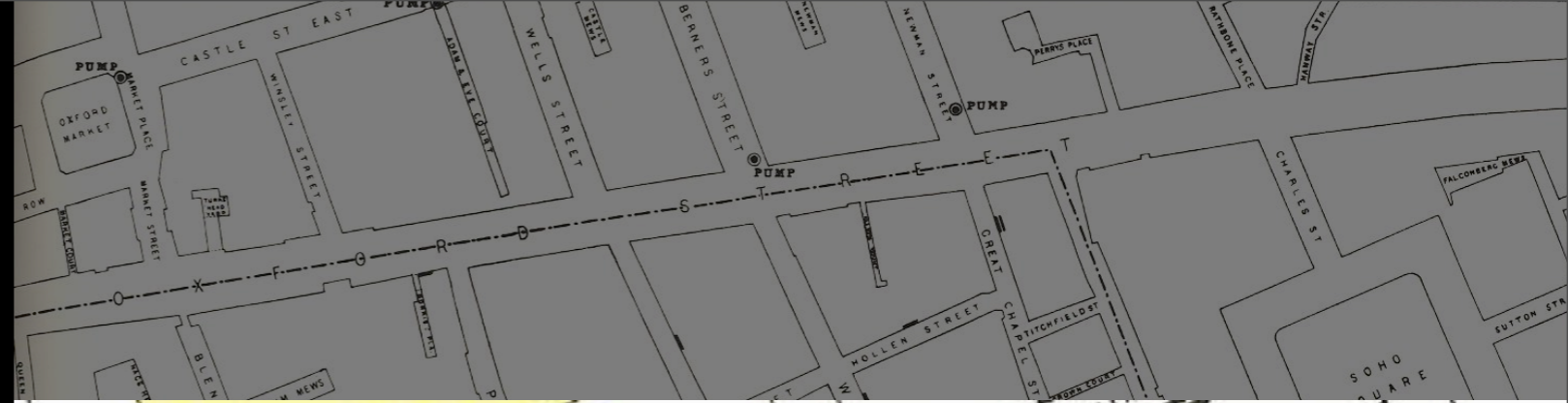
(in 1854)



Reproduced from Visual and Statistical Thinking, © E.R. Tuft 1997, based on Snow's drawing re: 1854 London cholera epidemic.

Epidemiology

(in 1854)



Epidemiology

(in 1854)



Reproduced from Visual and Statistical Thinking, © E.R. Tuft 1997, based on Snow's drawing re: 1854 London cholera epidemic.

Epidemiology (in 1854)

Displaying
“high-dimensional” data

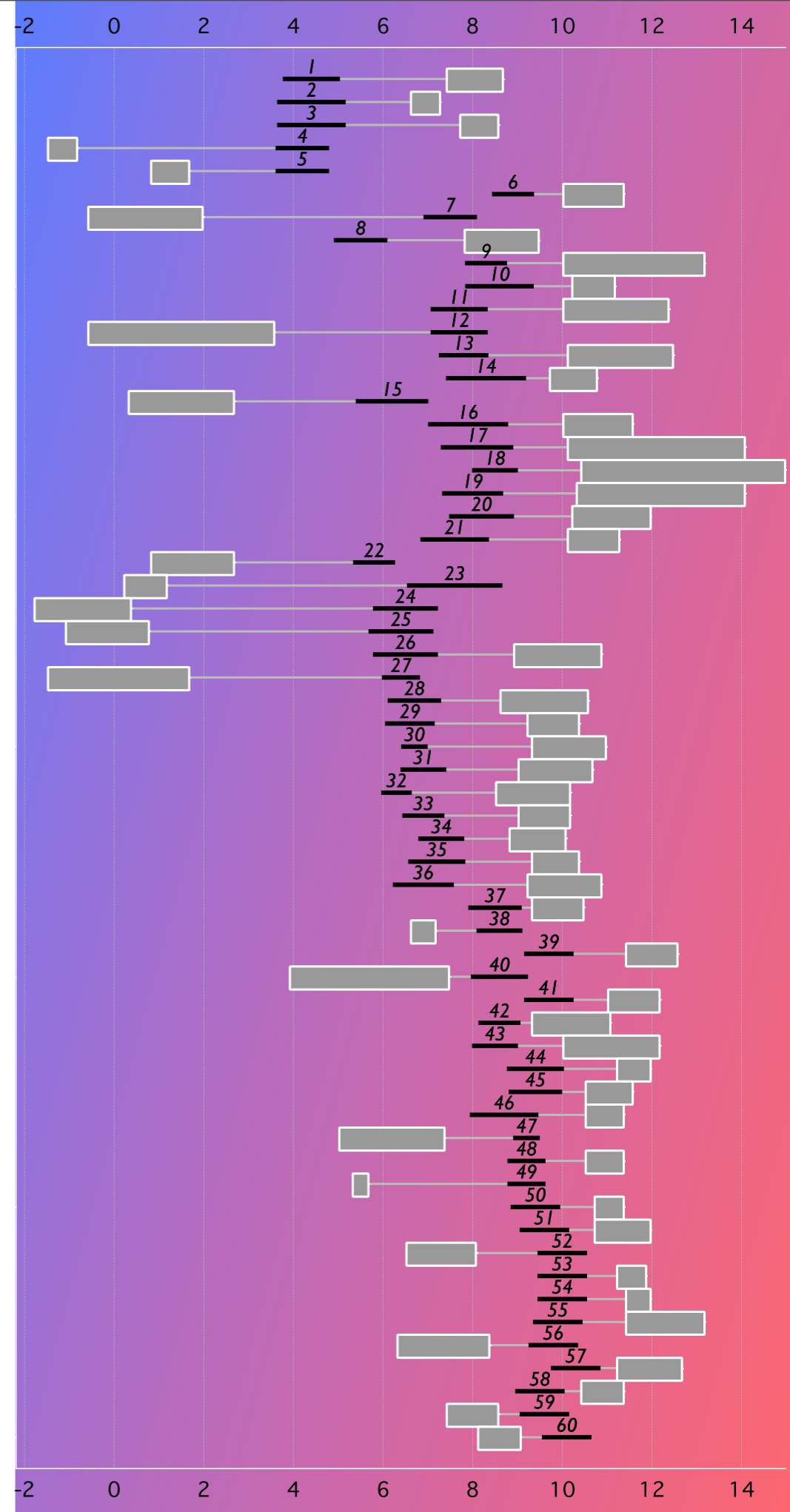
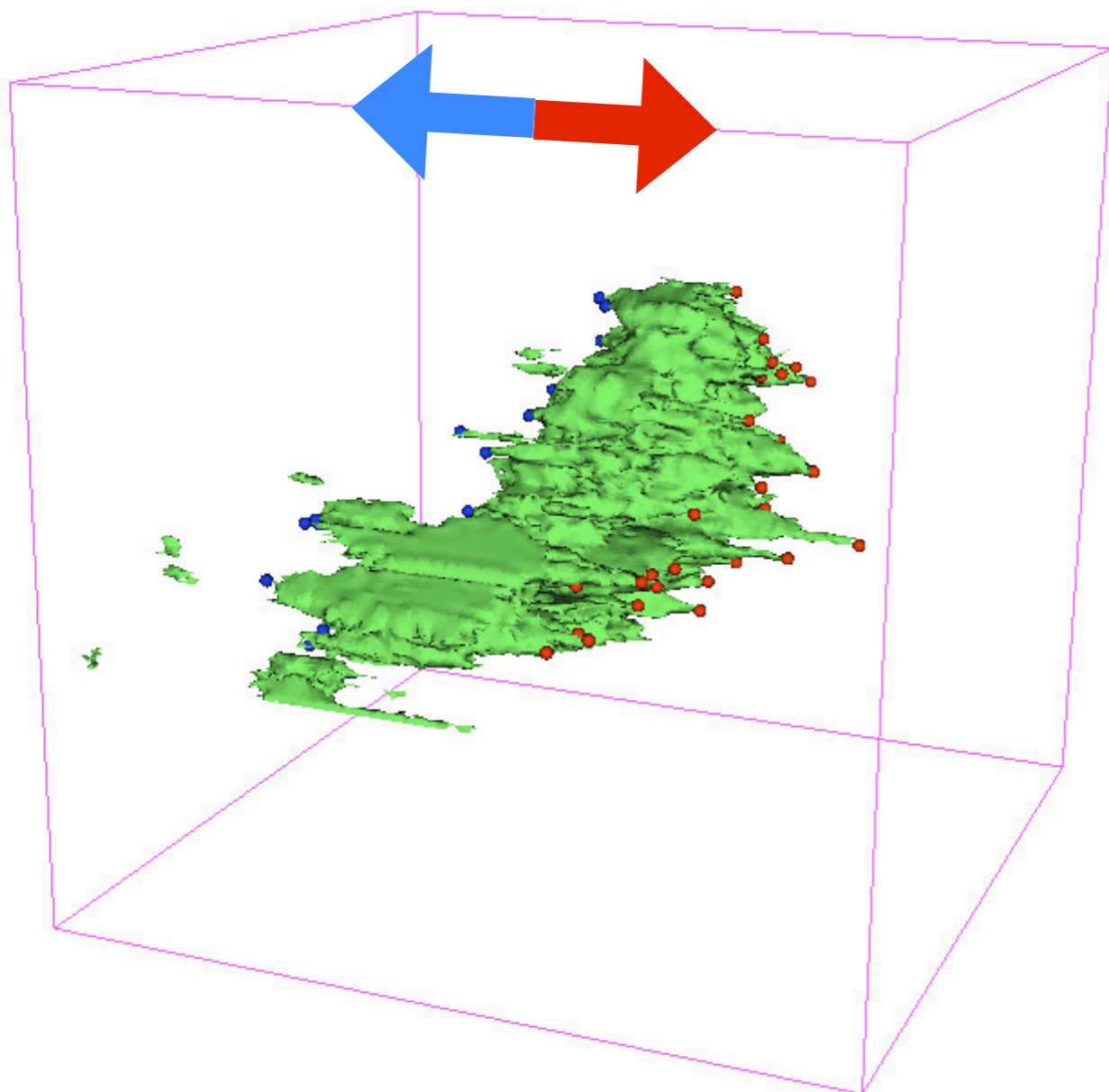
with

“multi-functioning
graphical elements”



Snow couldn't "interact" with the map but we should be able to, with the right **data linkages**, and choice of **dimensions & display**.

COMPLETE Perseus Outflow Candidates

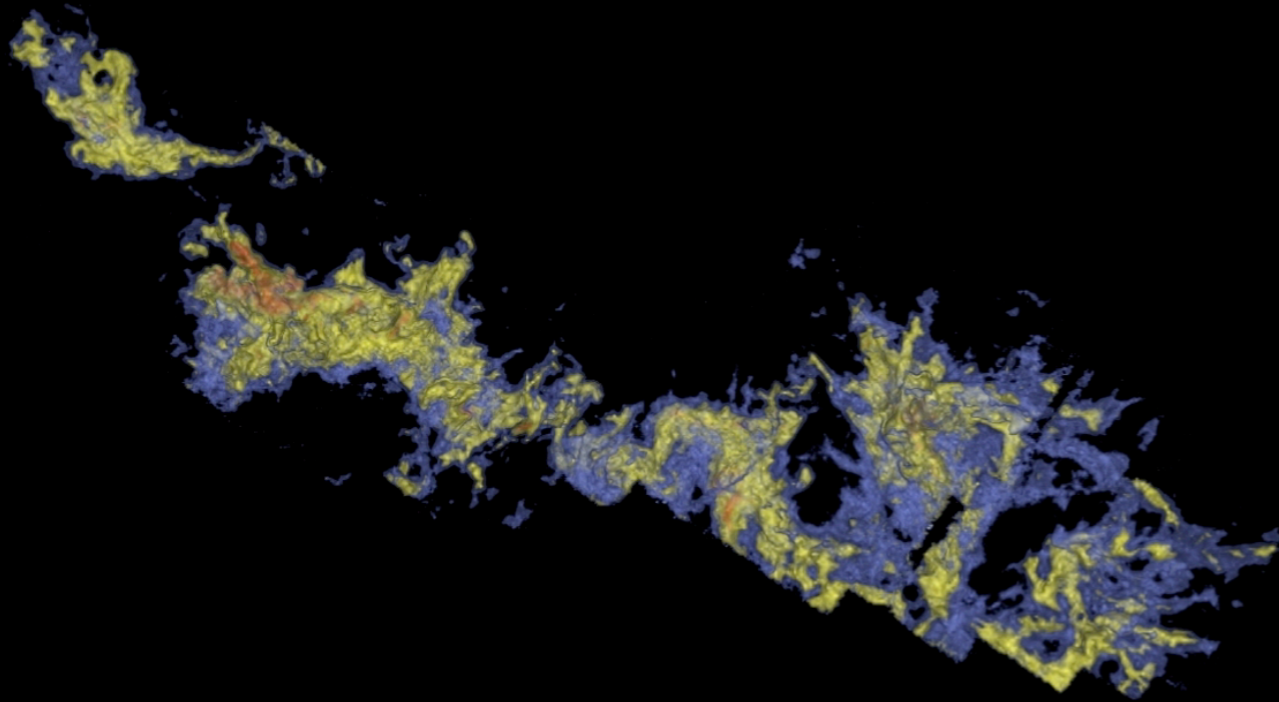


Arce et al. 2010; inset based on Offner et al. 2011

3D Selection

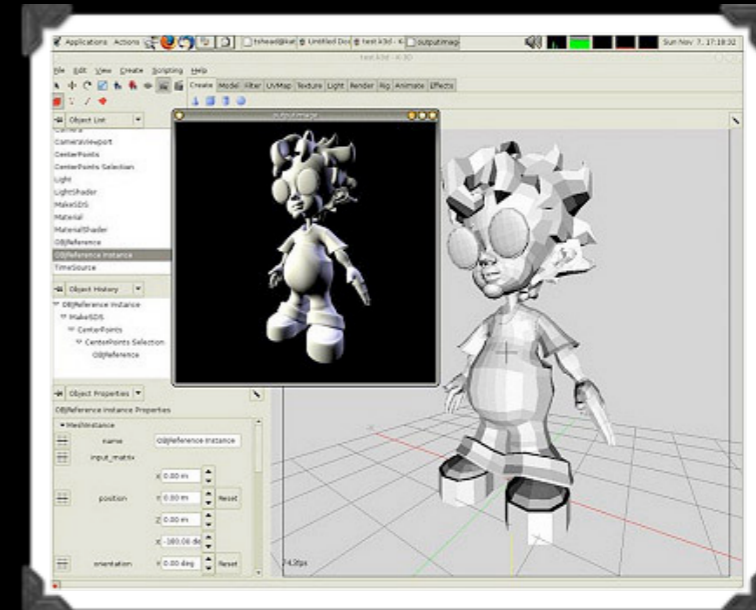
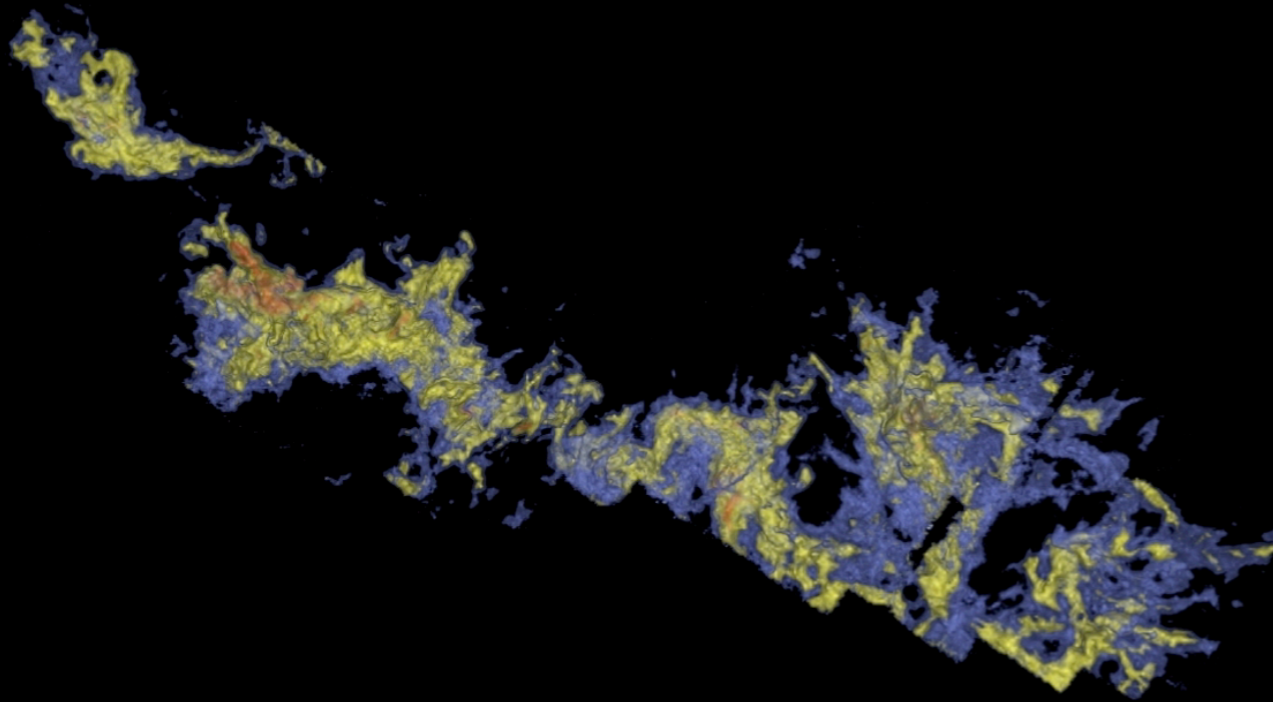
3D Selection

What's the **3D** “magnetic lasso”?



3D Selection

What's the **3D** “magnetic lasso”?
How do you use it with a mouse?

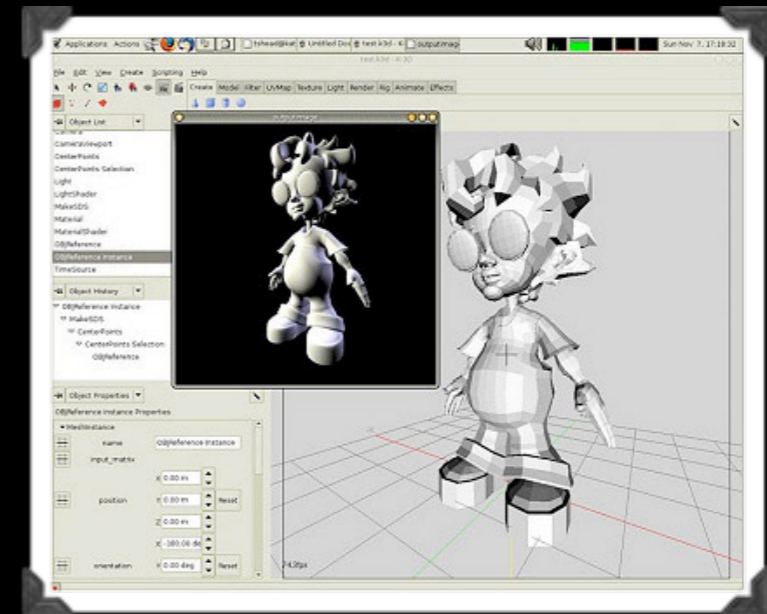
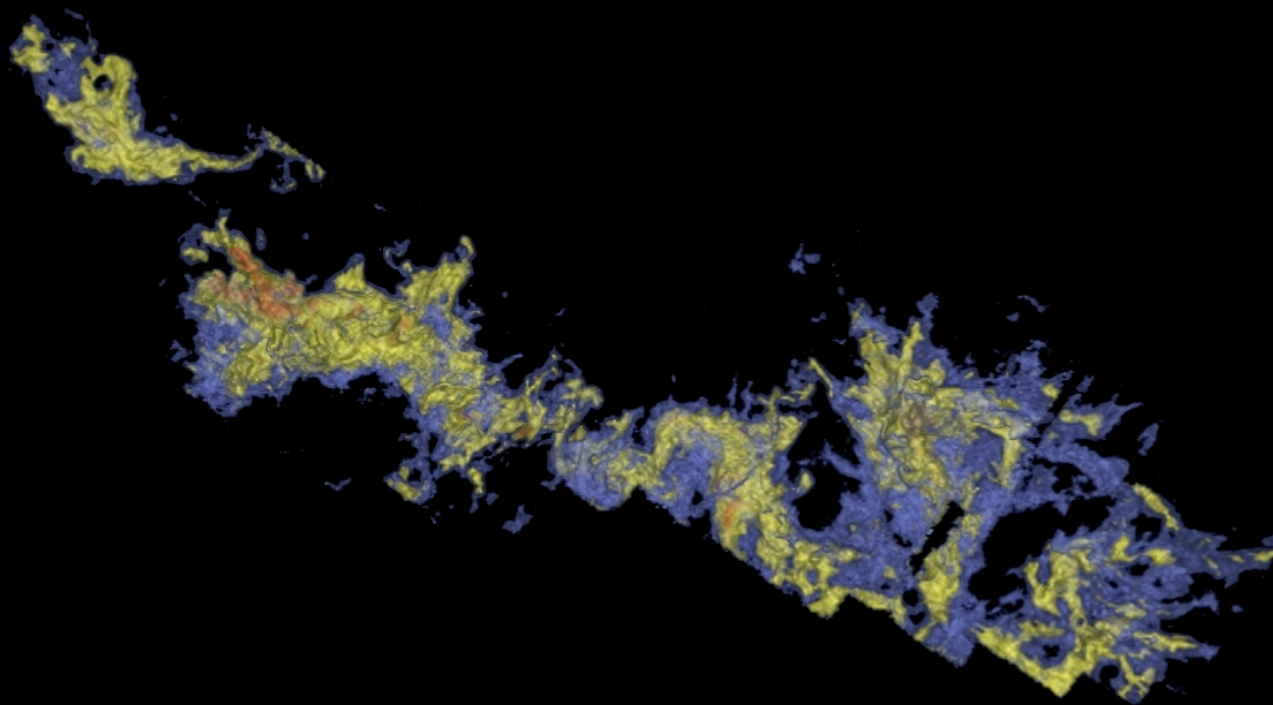


3D Selection

What's the **3D** “magnetic lasso”?

How do you use it with a mouse?

How can a human “steer” computer-aided selection?



3D Selection

Why
How
How



on?