Observing proposals

the scientific justification

The replay

From idea to data

1. What problem do you want to study?

2. What data do you need to tackle this problem?

   Do you need spectra? Images? Radio-data? Do you need archival data?

3. Can you do actually do it?

   It is easy to imagine scientifically interesting data that are unrealistic, or even impossible, to get.

4. What telescope should you use?

   Different telescopes have different instruments, are in different hemispheres, their access rights are different and the application pressure is different.

The TAC - be understandable!

This is composed of researchers that will evaluate the proposals. For big facilities like HST or the VLT there are many sub-committees that consist of experts in a broad area. But in national TACs there is usually only one group.

The reviewer is unlikely to be an expert on your topic!

A proposal must be understandable for a non-expert!

The TAC - be concise

A TAC member for a big facility might have 50-100 proposals to read. How long will they spend on each? Maybe 15 minutes?

A proposal must be concise and clear - get your point across quickly and efficiently!

The proposal

Introduction

Scientific justification

- ✓ Why is this topic important?
- ✓ What are the important questions?
- ✓ Why now?
- ✓ Links to other topics?
- ✓ Why are the proposed observations important?
- ✓ How will you do the analysis?

Technical justification

- ✓ Can you actually carry out your observations?

The proposal

Some hints on writing a good proposal:

- ✓ Clear and concise arguments
- ✓ Make it clear *what* you want to do
- ✓ Make it clear *how* you will do it
- ✓ Write clearly and in correct English - sloppy writing will not help you!
- ✓ Avoid jargon and explain acronyms the first time you use them
- ✓ Carefully follow instructions!

The technical justification

- When is the object visible?
  - Ephemerides, web based tools.
- What filter should I choose - and why?
  - Do you need more than one? Why?
- What exposure time do you need?
  - How bright is your object and what S/N do you need?
- How many observations are needed and at what cadence?

http://home.strw.leidenuniv.nl/~jarle/Teaching/Practicum2015/proposals.html

When

http://catserver.ing.iac.es/staralt/ Check altitude

Moving objects: Ephemerides

http://www.minorplanetcenter.net/iau/MPEph/MPEph.html

Deadlines

11/03 - Scientific justification for observing proposal
1 page

25/03 - Technical justification (1 page) & abstract (max 10 lines)

At this point we need to finalise our filter choices and targets

22/04 - TAC meeting. You grade each other's proposals.

Maximum 2 pages with figures & references
Font: 11 pt Times Roman

Grading

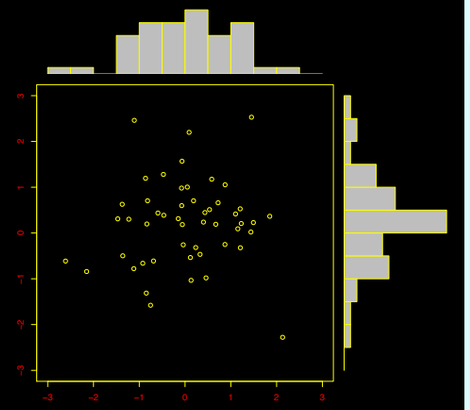
Handing in problem sets is **obligatory**!

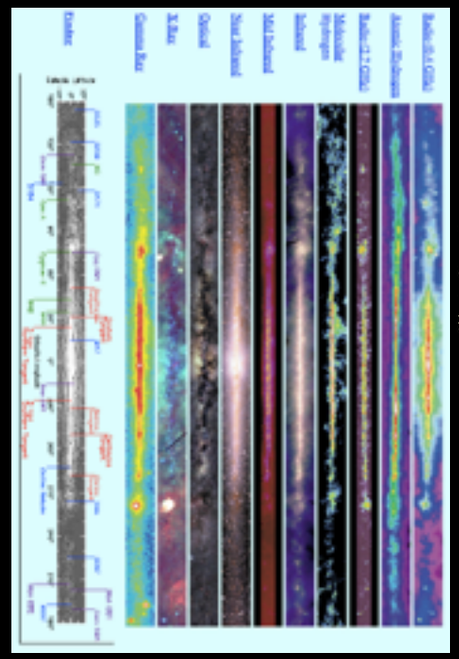
The course grade is the mean of the problem set grades & the observational part of the course.

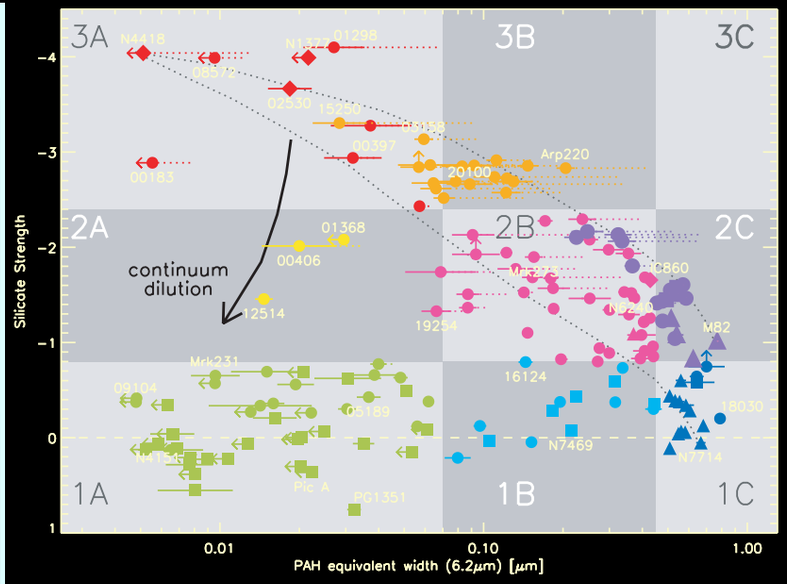
Handing in late next time **will result** in a grade of **0** unless I have given permission to do so (and I will only do for illness).

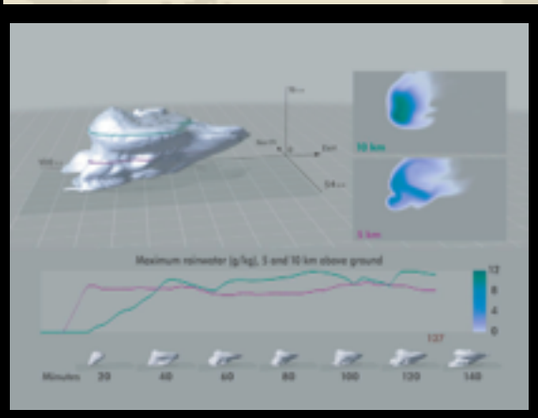
Maximum one 5 from the problem sets unless your average grade is >= 6.

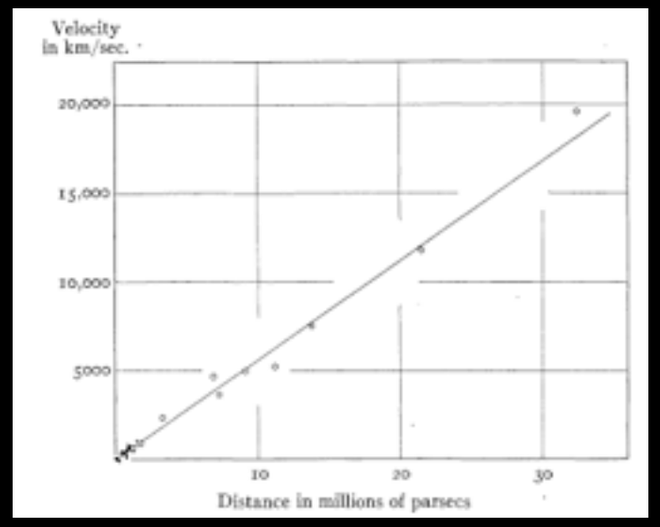
Visual explanations



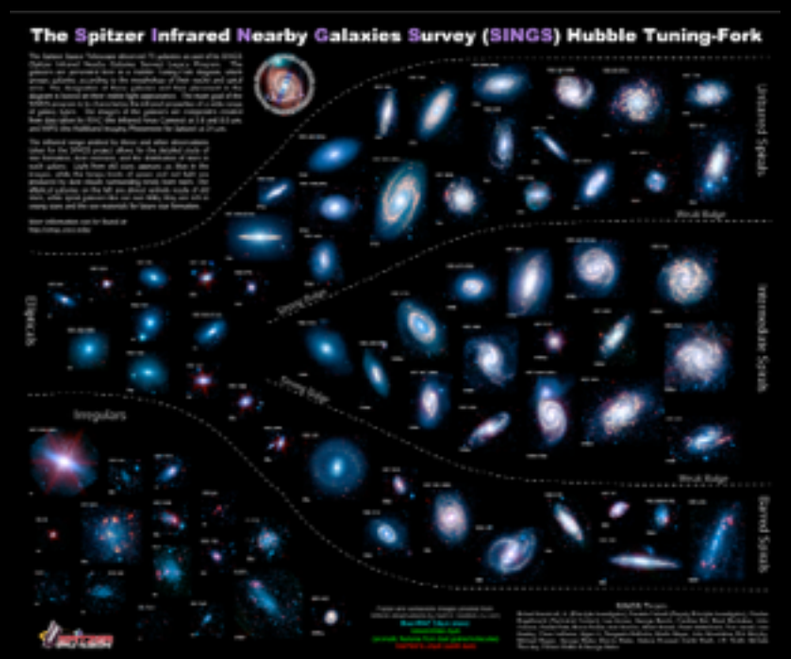


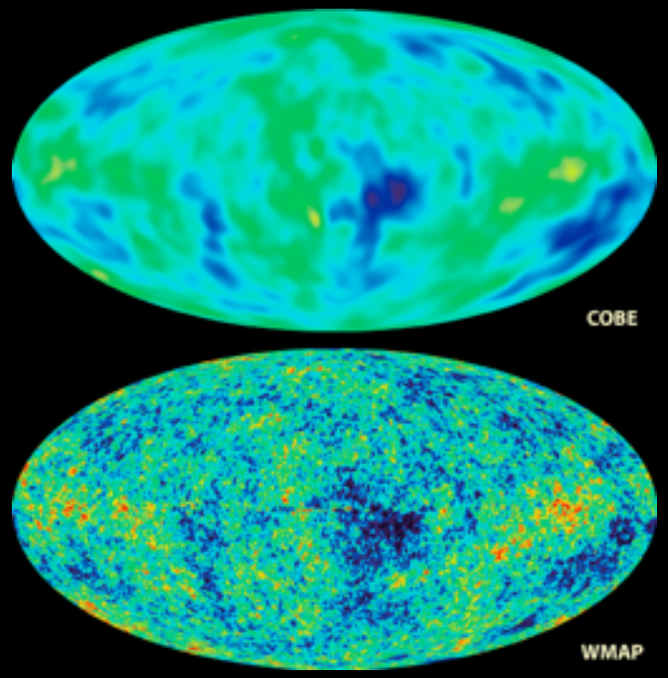






Visual explanations





Making good scientific plots

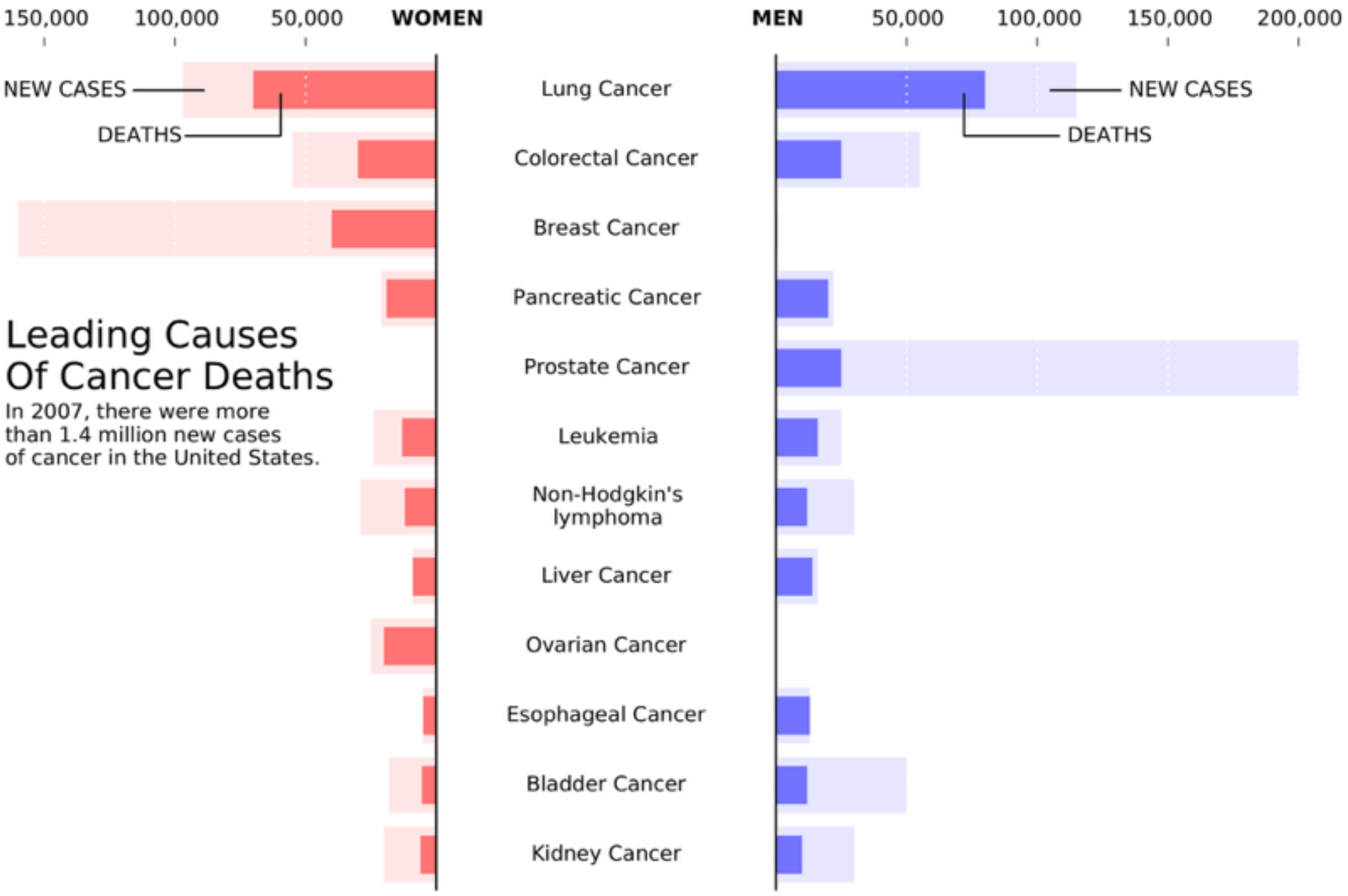
- Know the audience.
- Adapt to the situation.
- Be precise, include all information needed.
- Do not mislead.
- Do not use colour without thinking. Especially bad colour choices.
- Do not overload a figure - or include too little!
- Learn a couple of packages well.

Worth checking out: Rougier NP, Droettboom M, Bourne PE (2014) Ten Simple Rules for Better Figures. PLoS Comput Biol 10(9): e1003833.

http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1003833

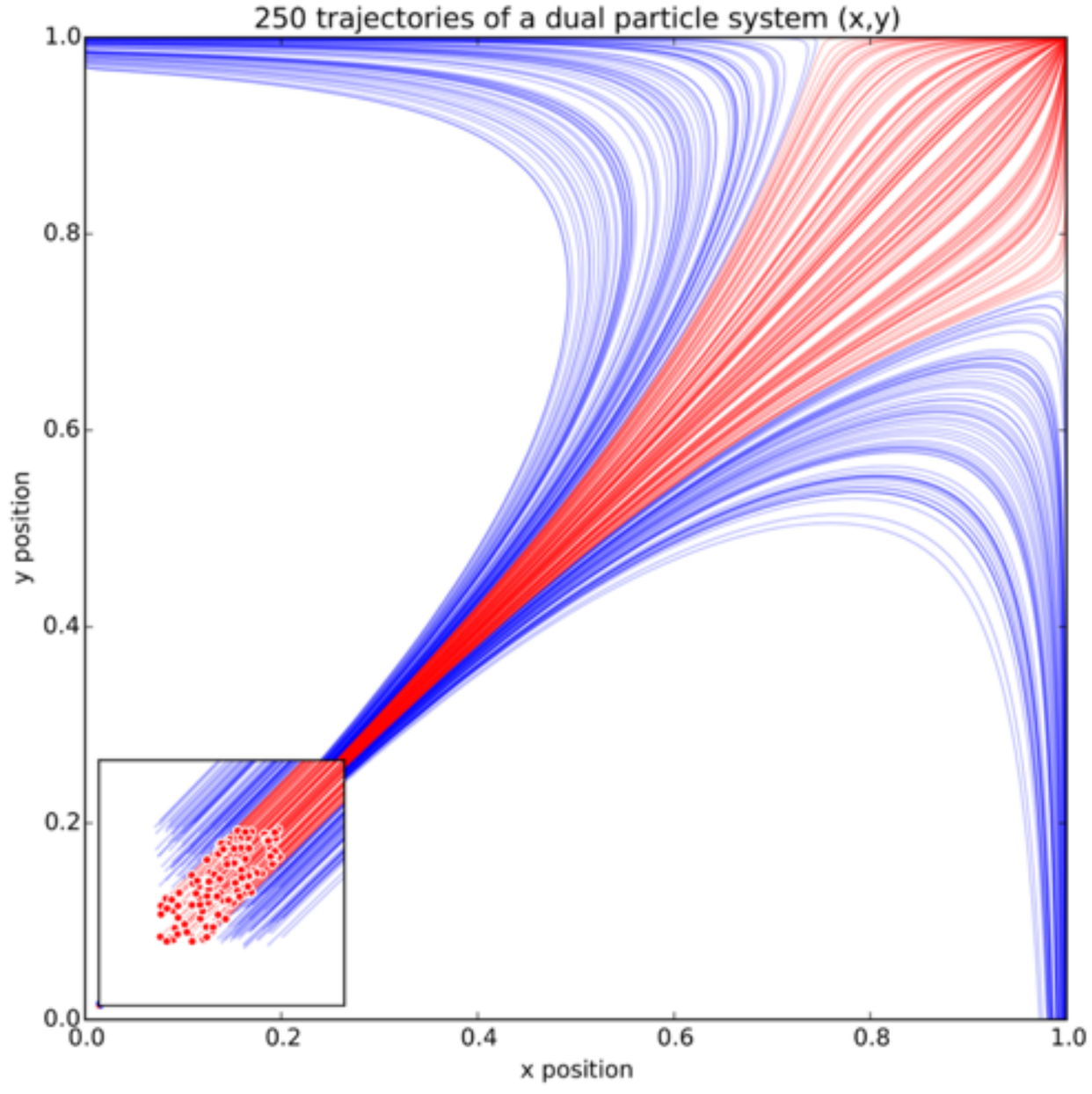
Adapting to the audience

from: Rougier NP, Droettboom M, Bourne PE (2014)

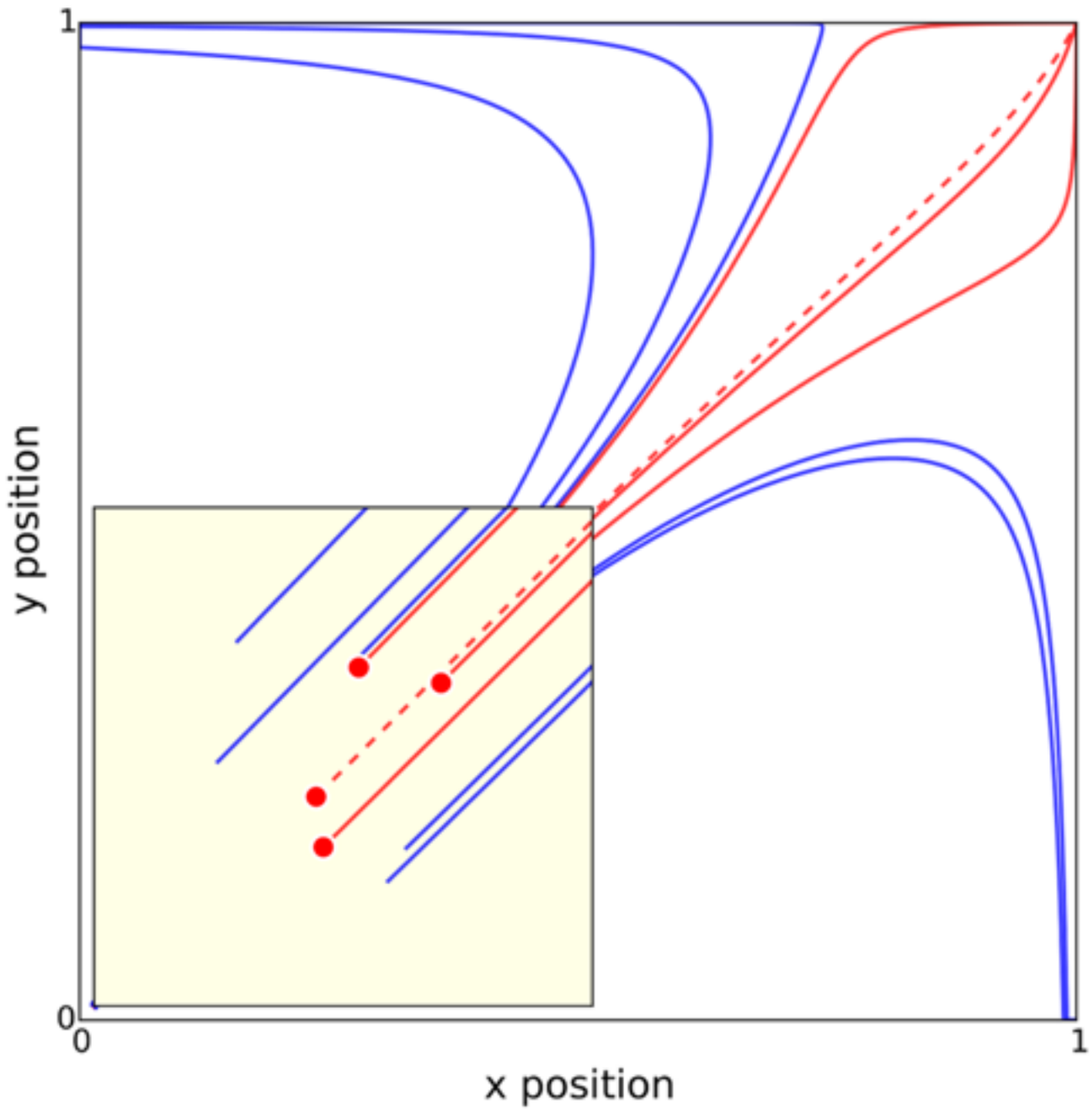


Adapting to the siutation

from: Rougier NP, Droettboom M, Bourne PE (2014)



Scientific paper



Presentation

Why do we create visualisations (plots)?

Why do we create visualisations (plots)?

To help our argumentation.

Why do we create visualisations (plots)?

To help our argumentation.

To summarise a lot of data

Why do we create visualisations (plots)?

To help our argumentation.

To summarise a lot of data

To discover new relationships

Why do we create visualisations (plots)?

To help our argumentation.

To summarise a lot of data

To discover new relationships

To relate different pieces of information

Why do we create visualisations (plots)?

To help our argumentation.

To summarise a lot of data

To discover new relationships

To relate different pieces of information

To develop intuition

An example of *odd* data

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

An example of *odd* data

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

mean(x) = **9**, mean(y) = **7.5**

Var(x) = **11**, Var(y) = **4.12**

The correlation coefficient: **0.816**

The best-fit line: **y = 3 +0.5 x**

An example of *odd* data

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

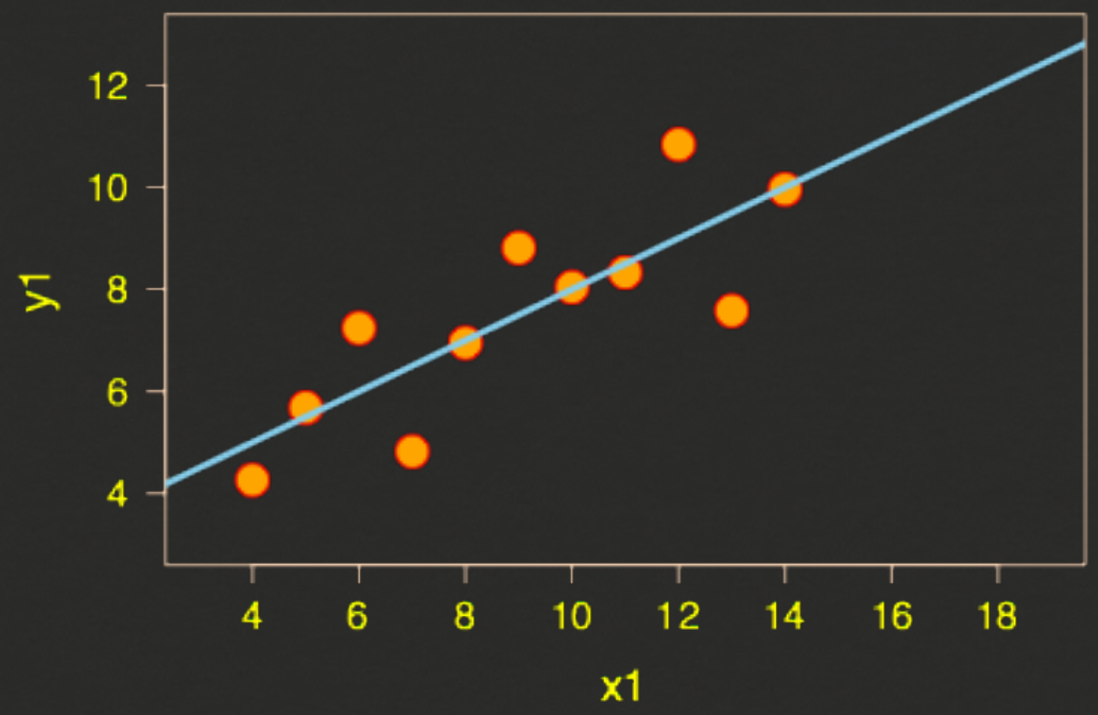
mean(x) = **9**, mean(y) = **7.5**

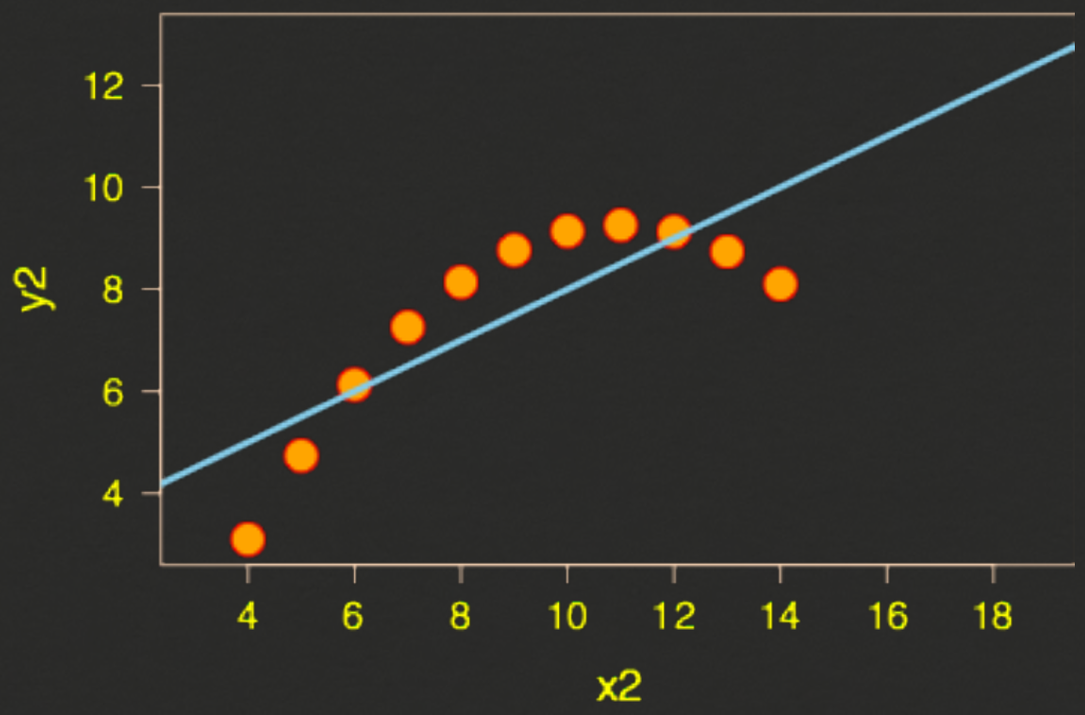
Var(x) = **11**, Var(y) = **4.12**

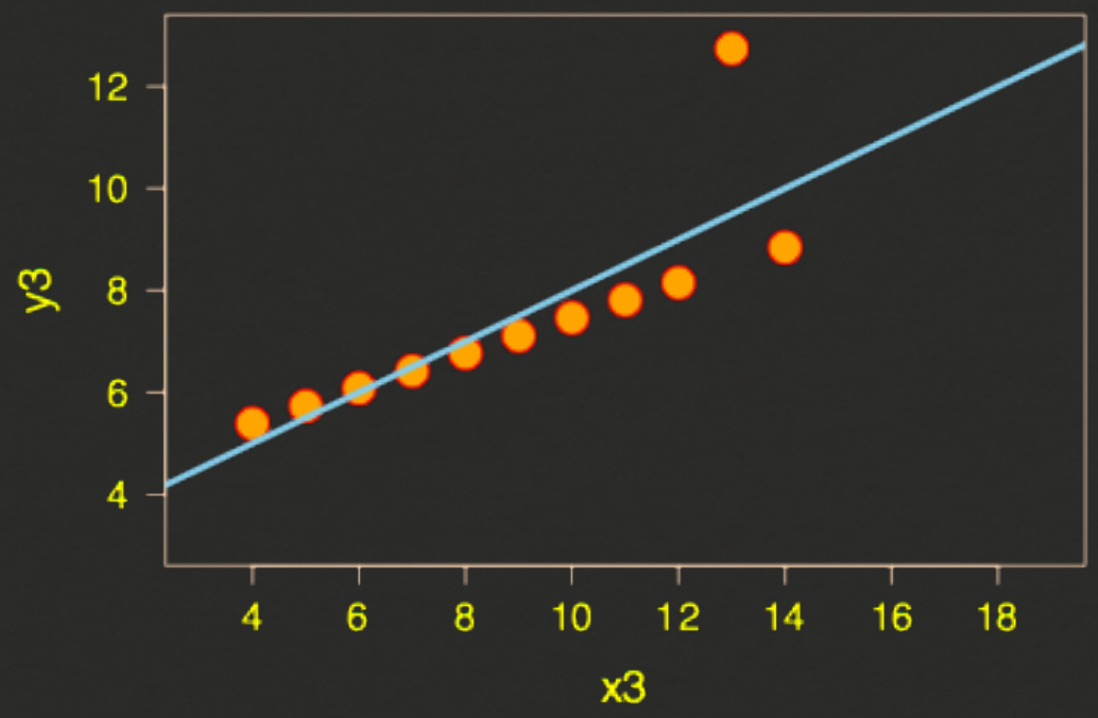
The correlation coefficient: **0.816**

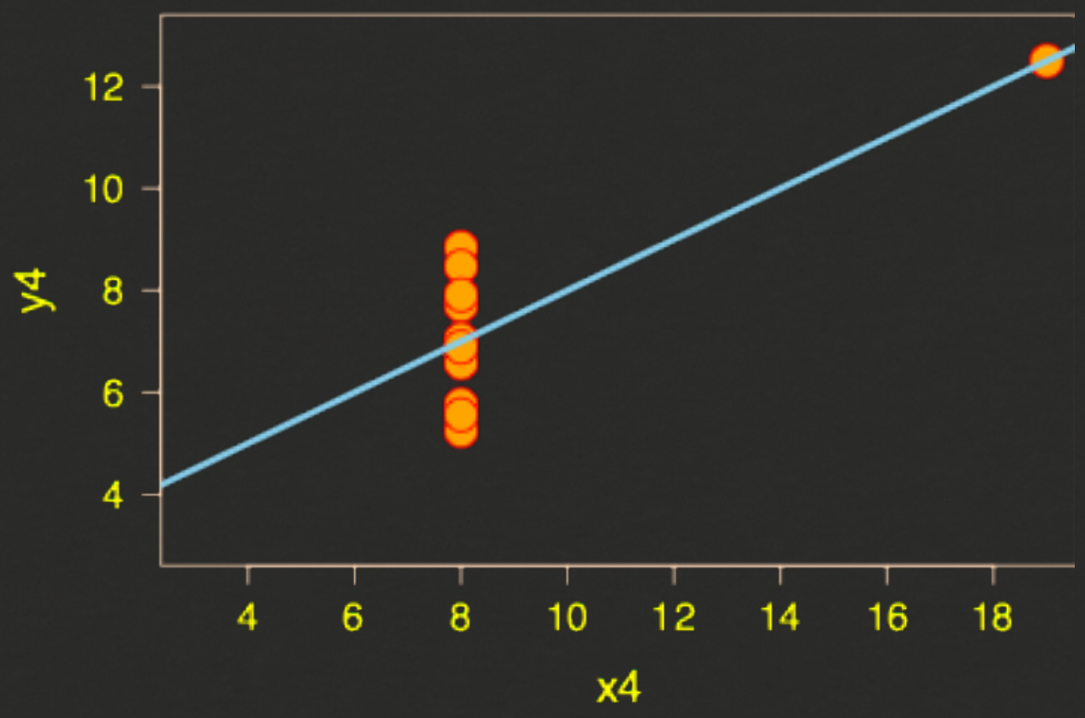
The best-fit line: **y = 3 +0.5 x**

So they seem to be very similar!









The bottom line: Always plot your data! (but you knew that already I hope!)

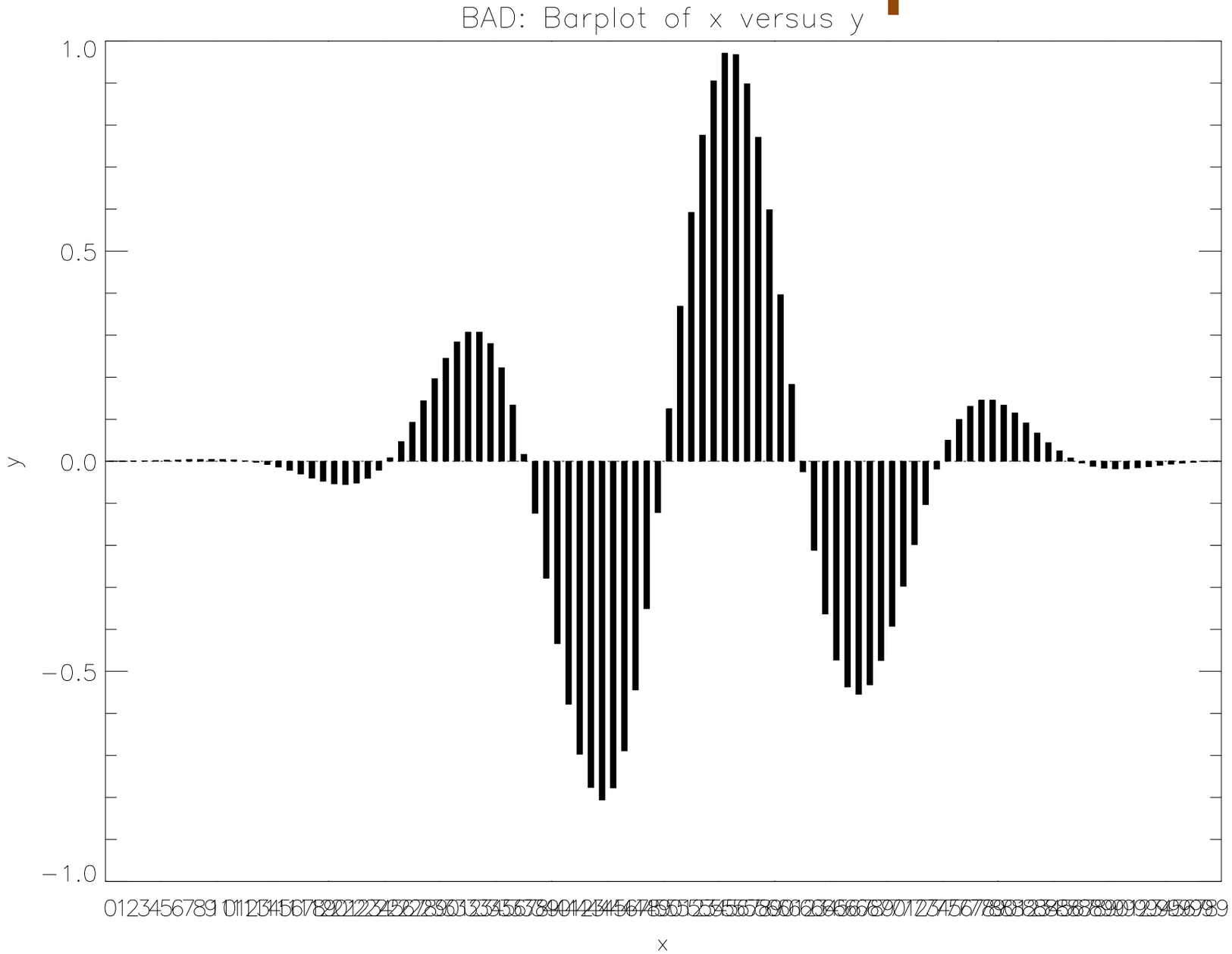
Graphical displays should:

- Show the data.
- Induce the viewer to think about the substance.
- Avoid distorting what the data have to say.
- Present many numbers in a small space.
- Make large data sets coherent.
- Encourage the eye to compare different pieces of data.
- Reveal data at several layers of detail.
- Be relevant.
- **Show units, show scales!**

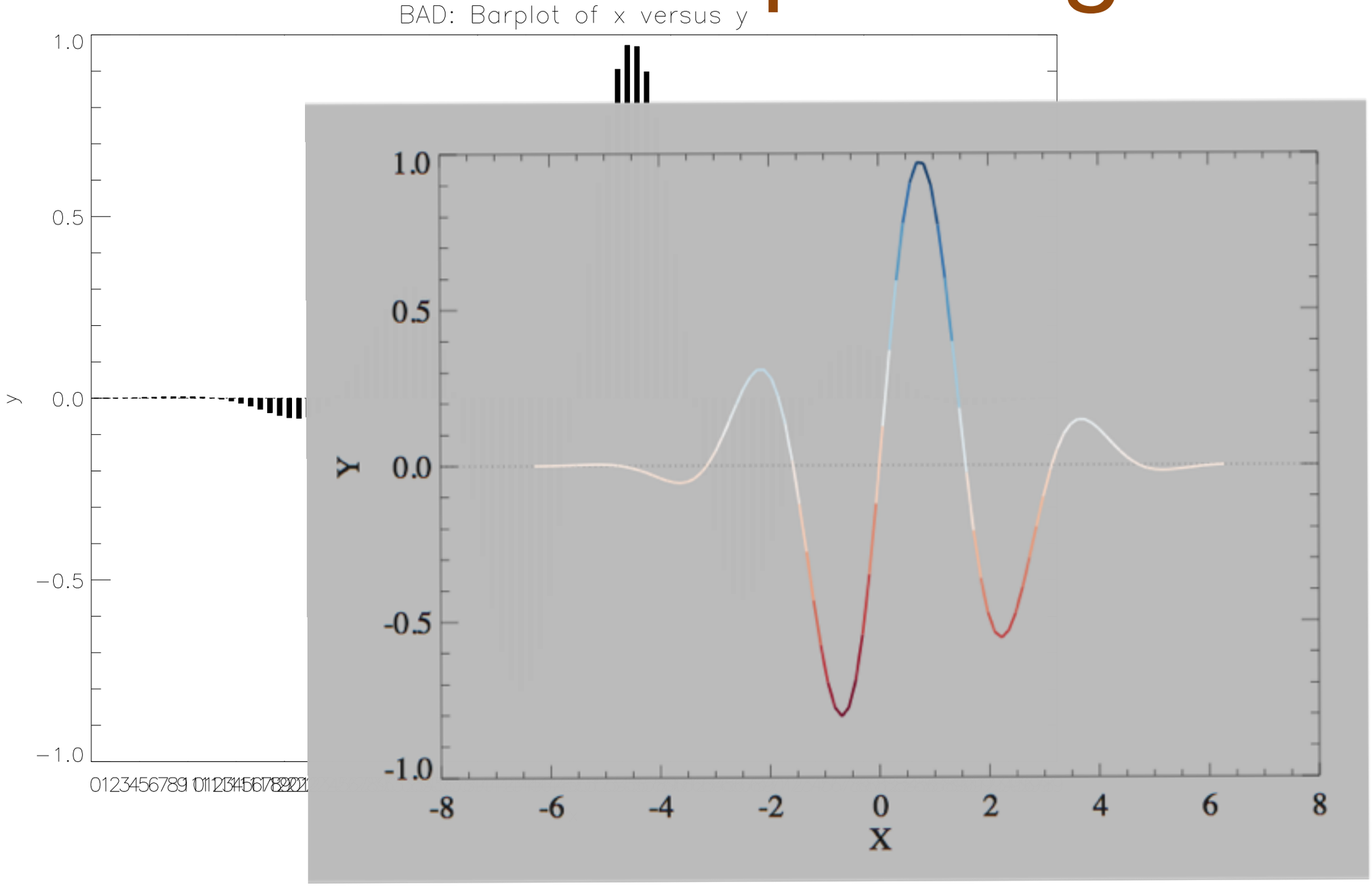
Based on: Tufte: "The Visual Display of Quantitative Information"

The Good, the Bad and the Ugly

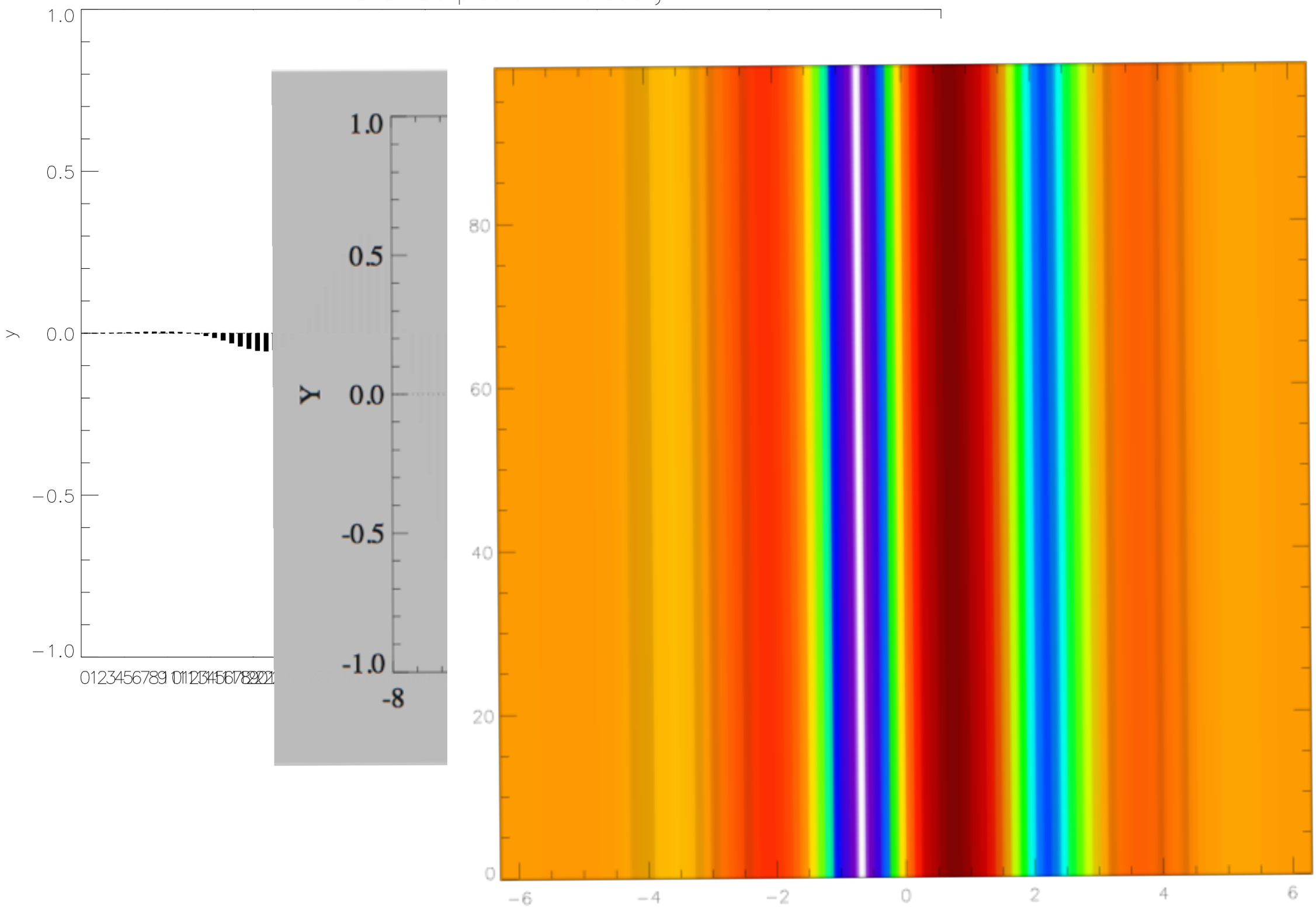
Function plotting.



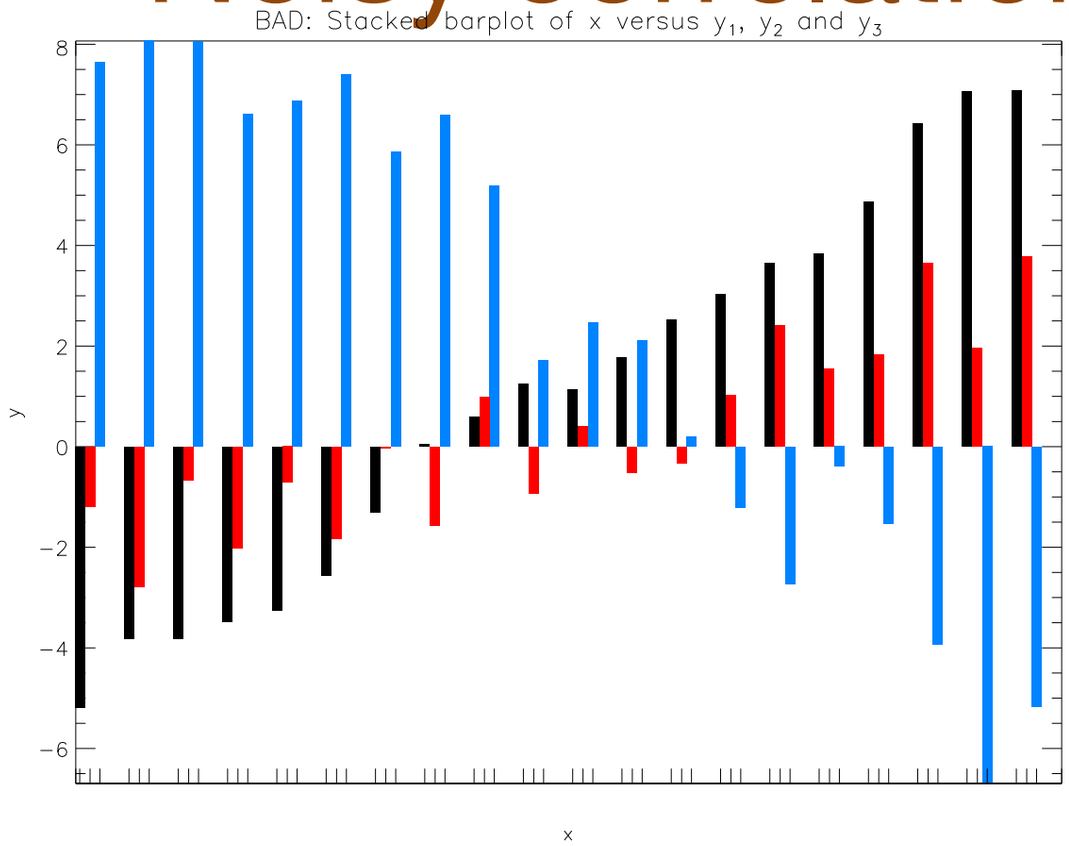
Function plotting.

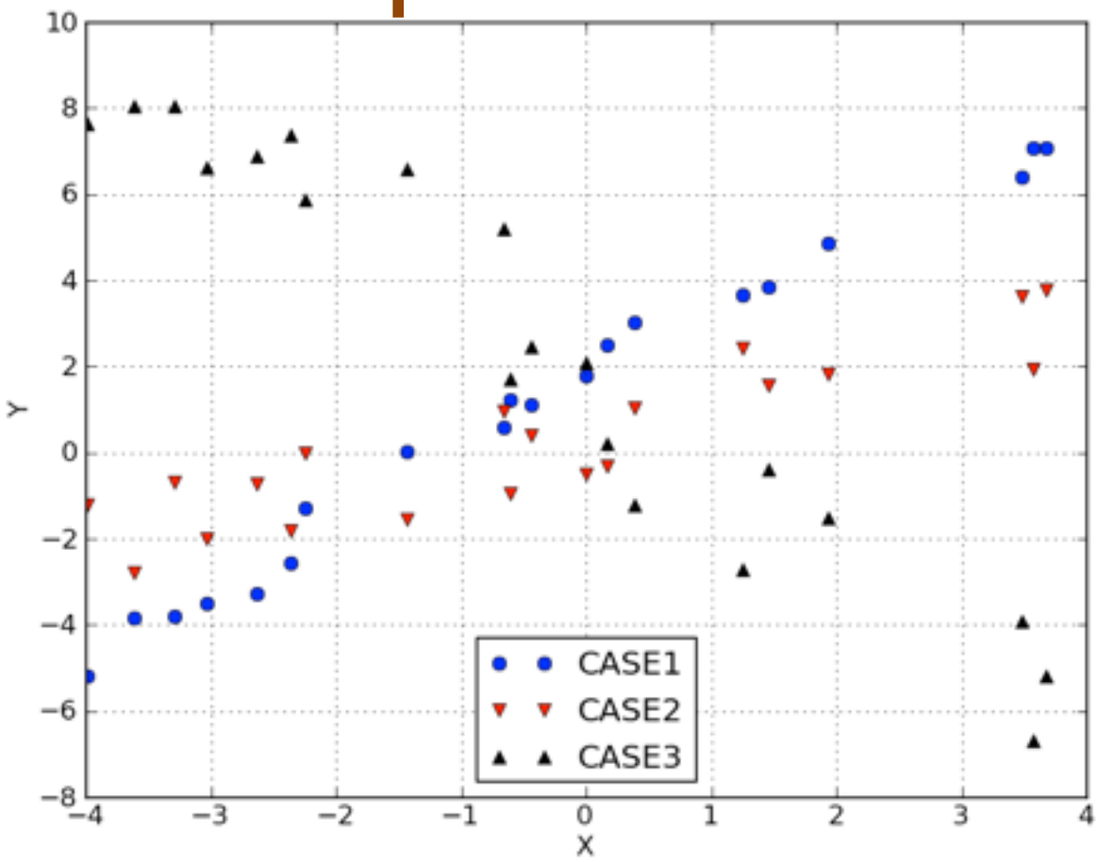


Function plotting.

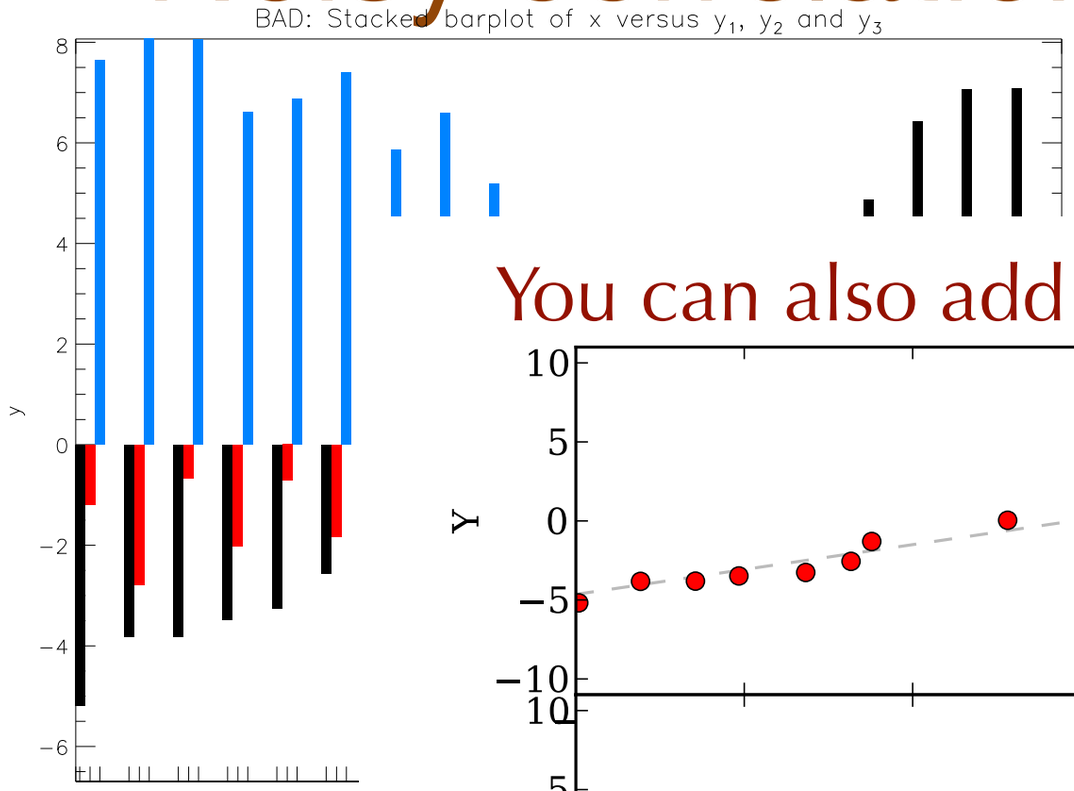


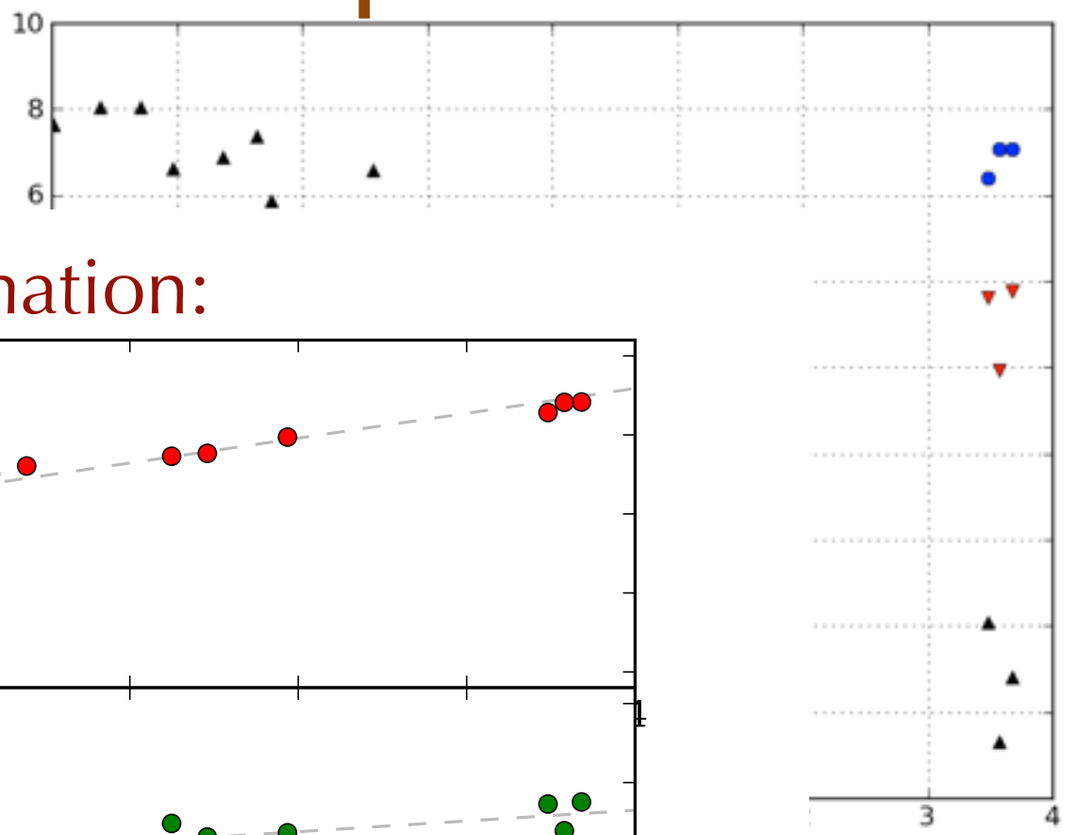
Noisy correlations - other possibilities



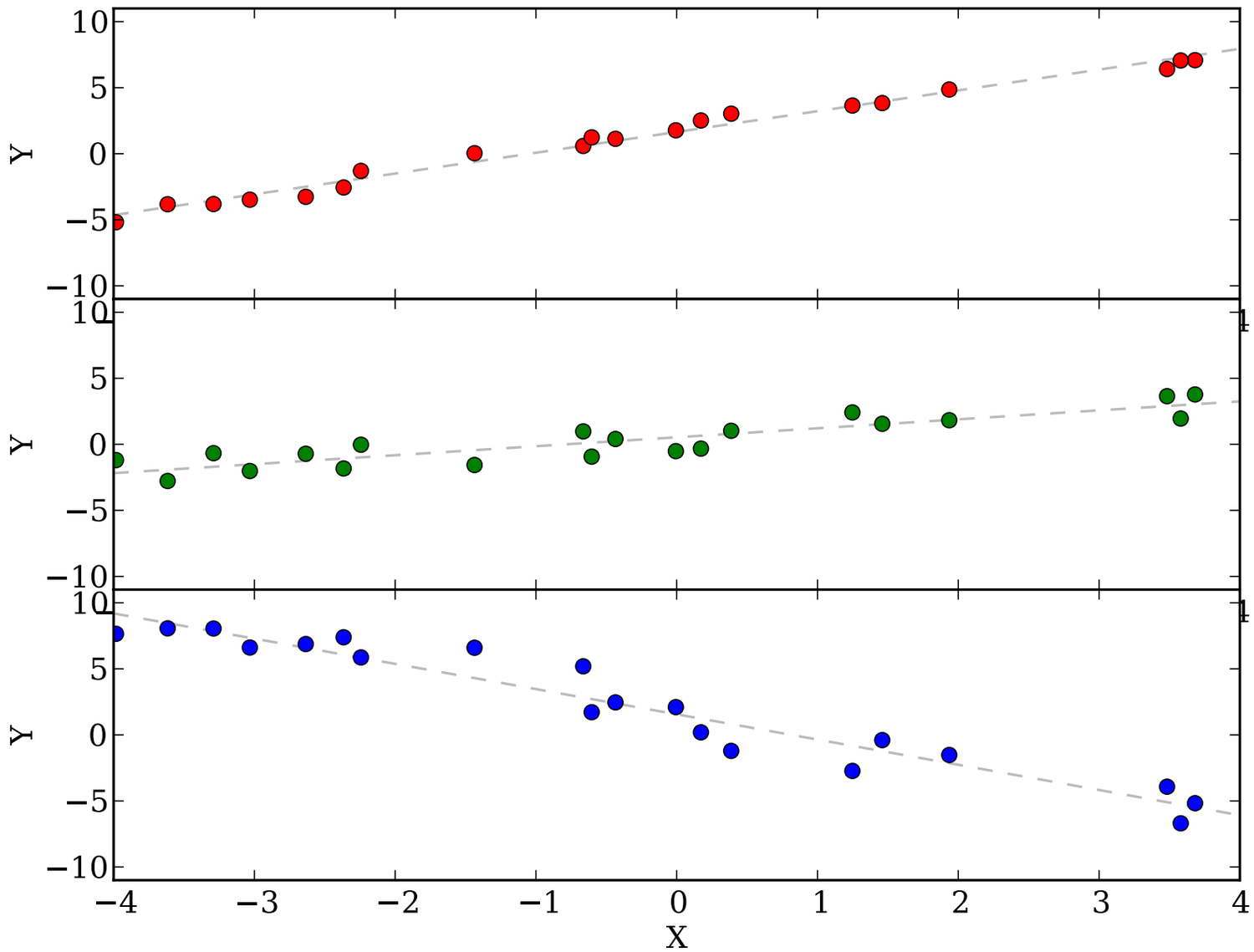


Noisy correlations - other possibilities

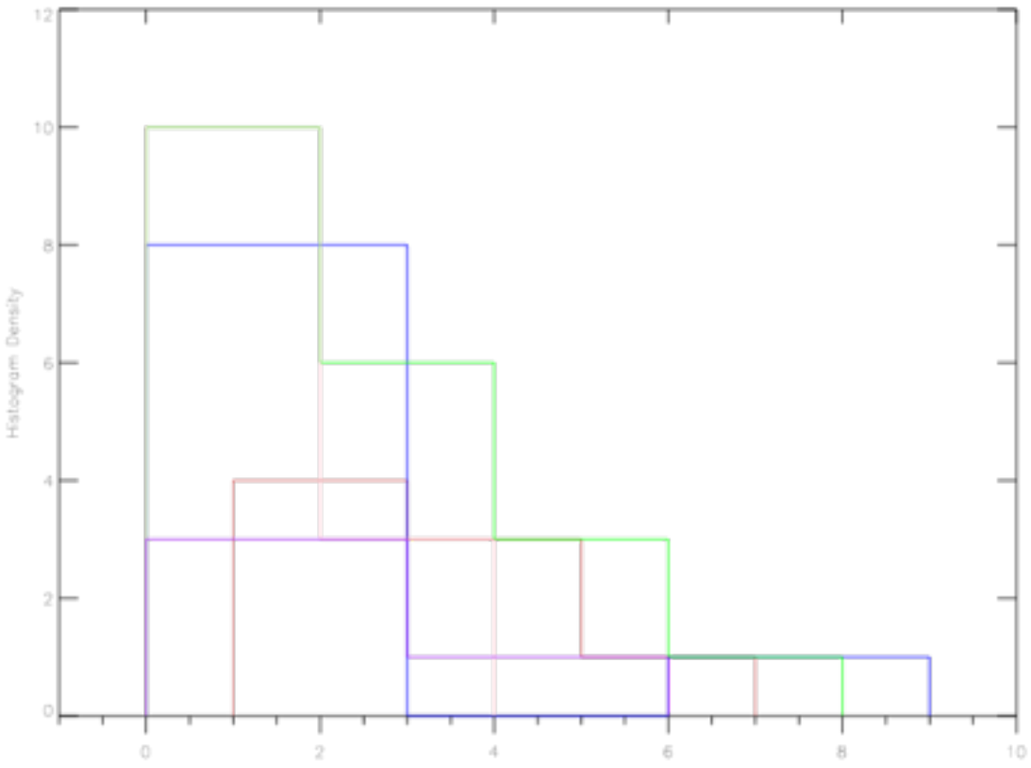


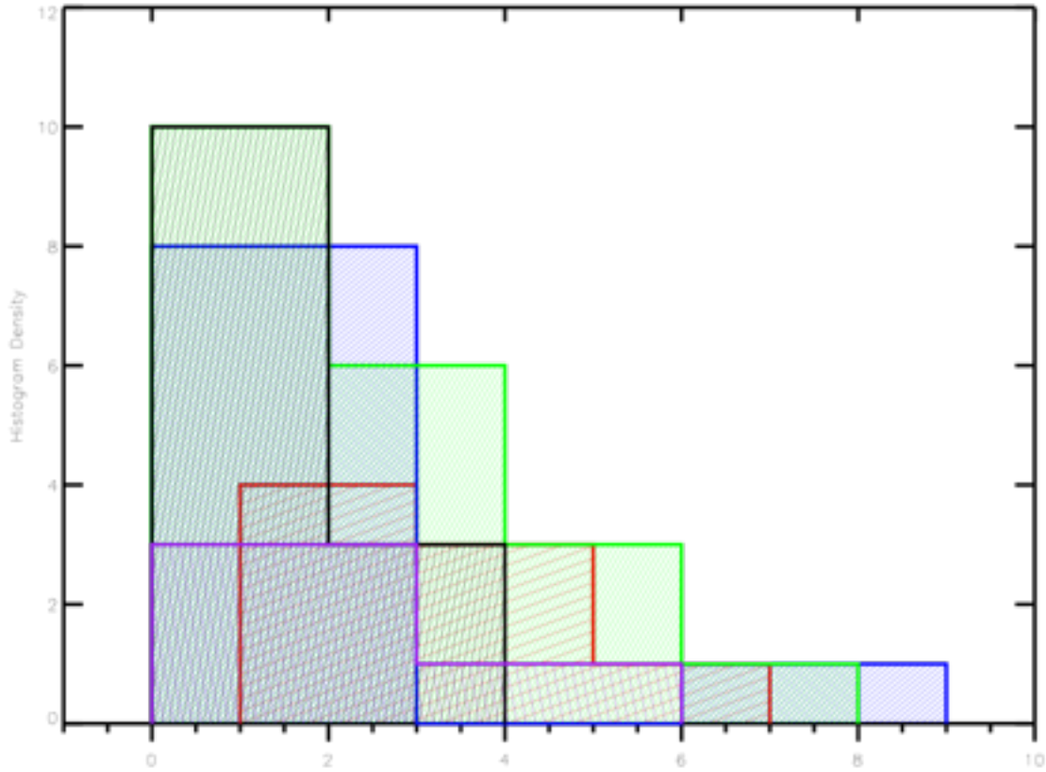


You can also add information:



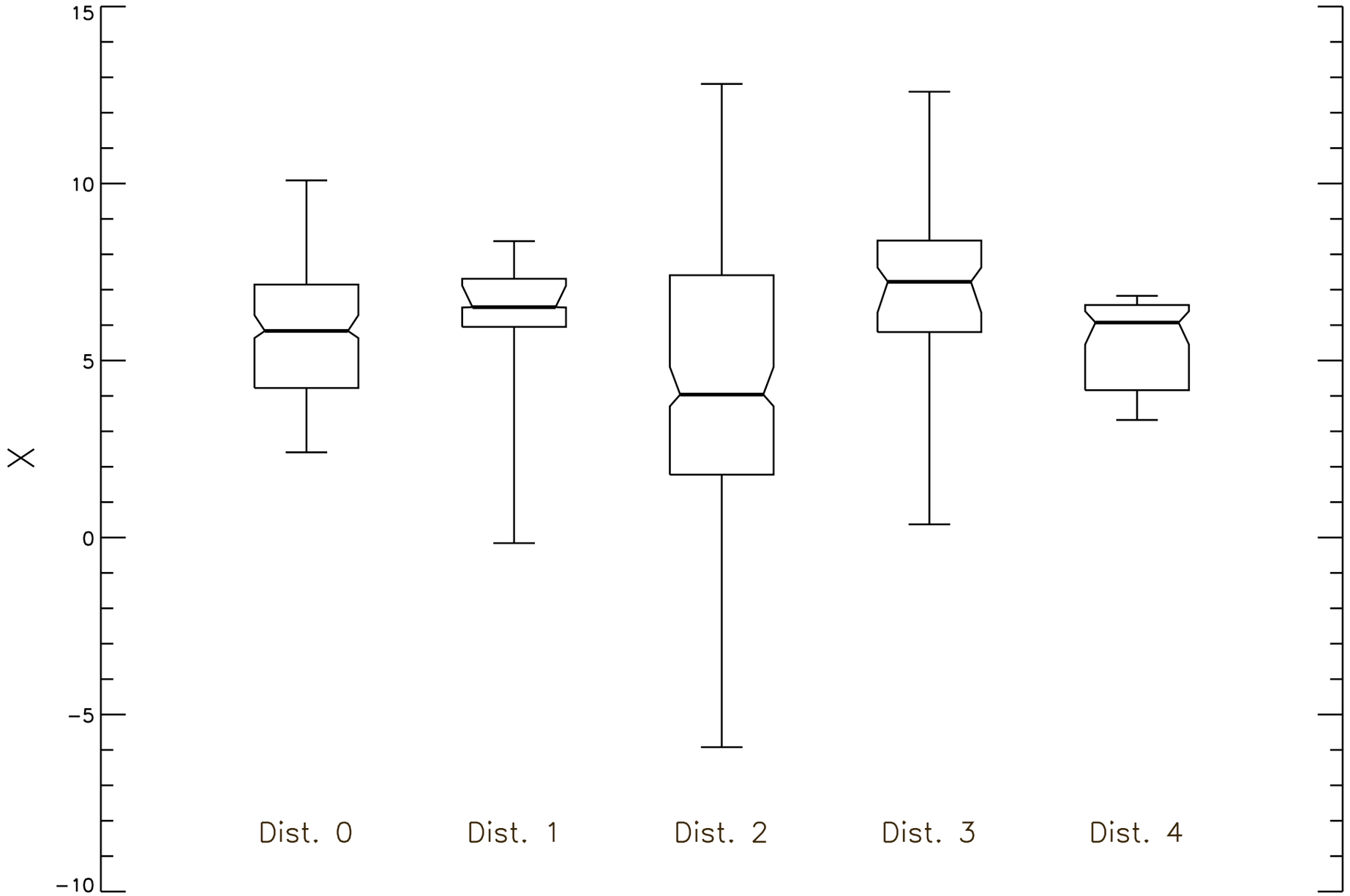
Distributions





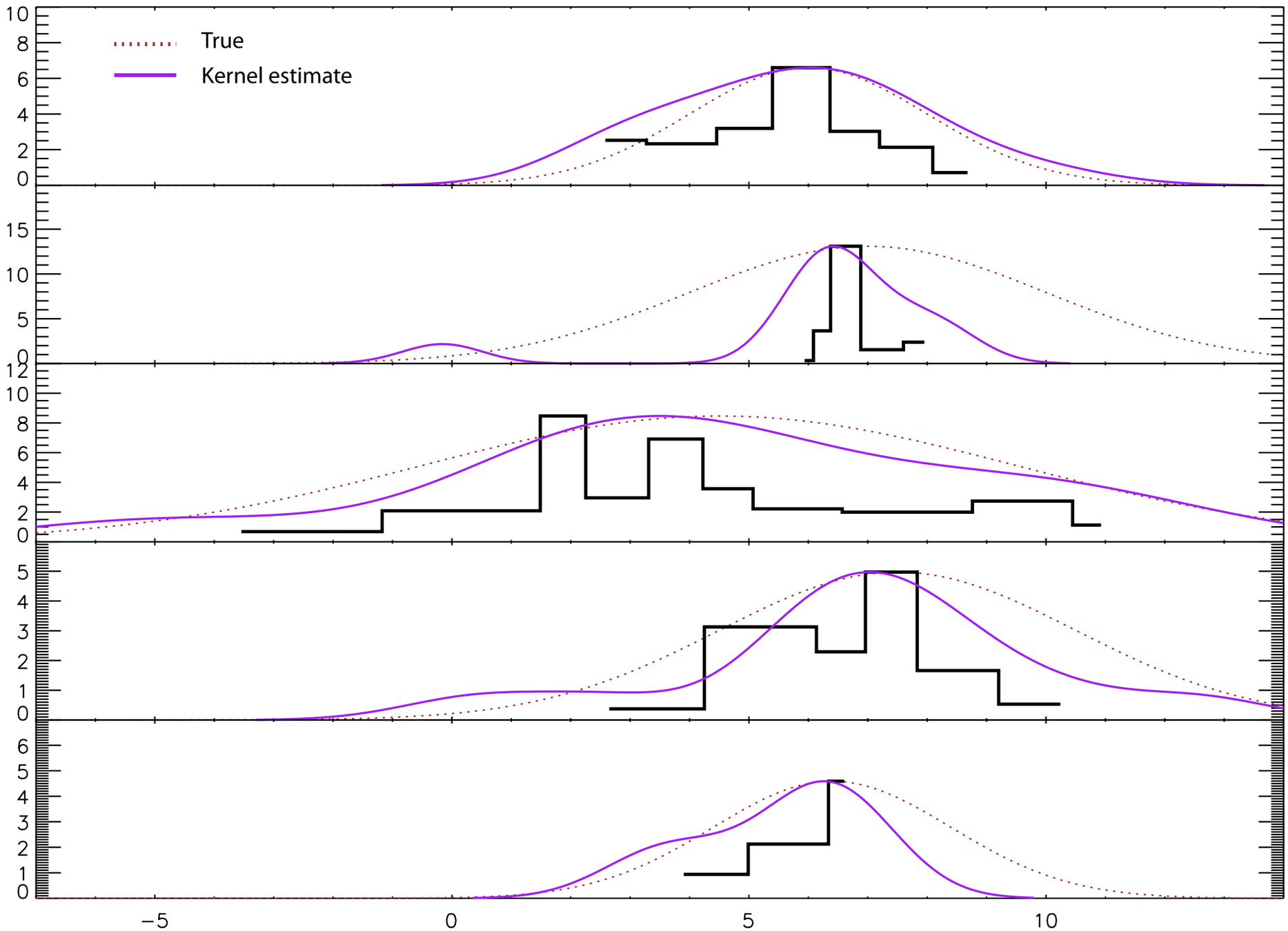
Distributions - some alternatives

Alternatives to histograms (box plot):



Distributions - some alternatives

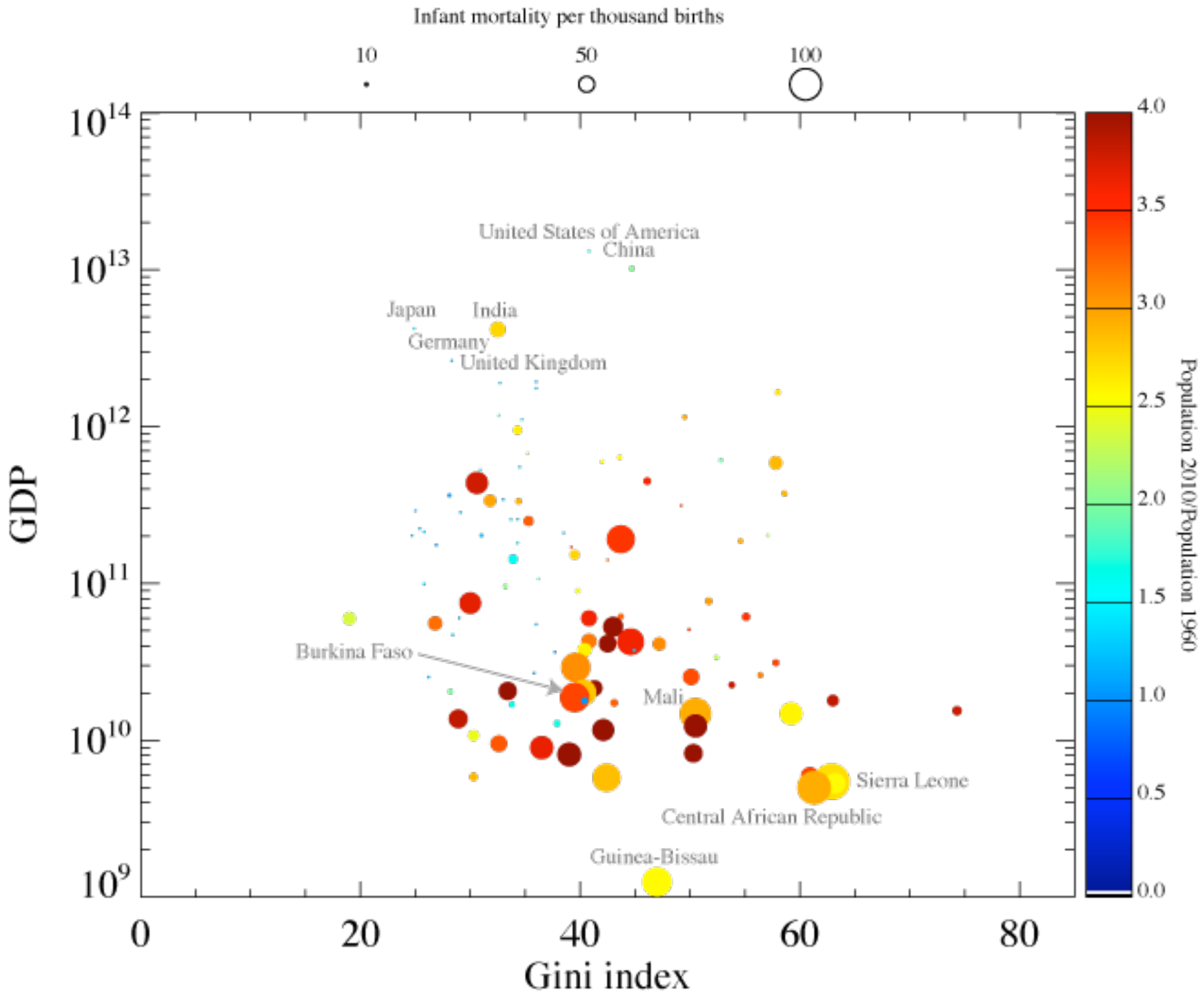
Multiple plot panels + extra information:



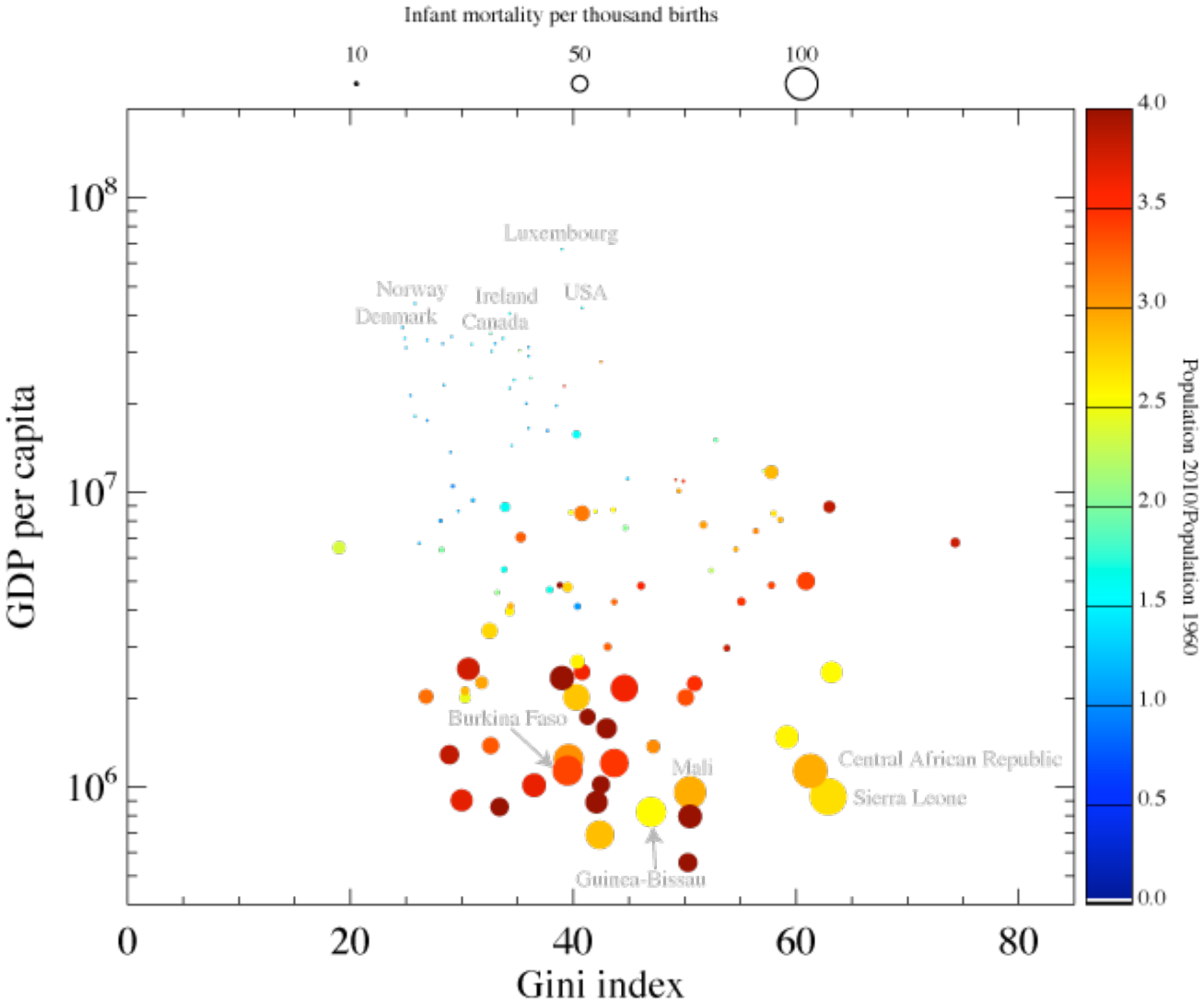
Techniques for visualising high-dimensional data

- 2D histograms/2D scatter plots
  - Colouring with a third variable can illustrate 3D, transparency & size can be used to add another couple.
- Scatter plot matrices
  - Arbitrary high D but not particularly powerful.
- Movies
  - By combining with coloured histograms can represent 4-5 dimensions. Often best when the time-scale of the movie actually is a time-scale and not some abstract quantity.
- Interactive visualisation (e.g. 3D rotation)
  - Very useful for data exploration but in practice hard to use past 3D although colouring can help. Not very useful for publishing.
- Dimensional reduction
  - Use an objective technique to focus on the “important” dimensions of the data, then use the other techniques to look at the results.

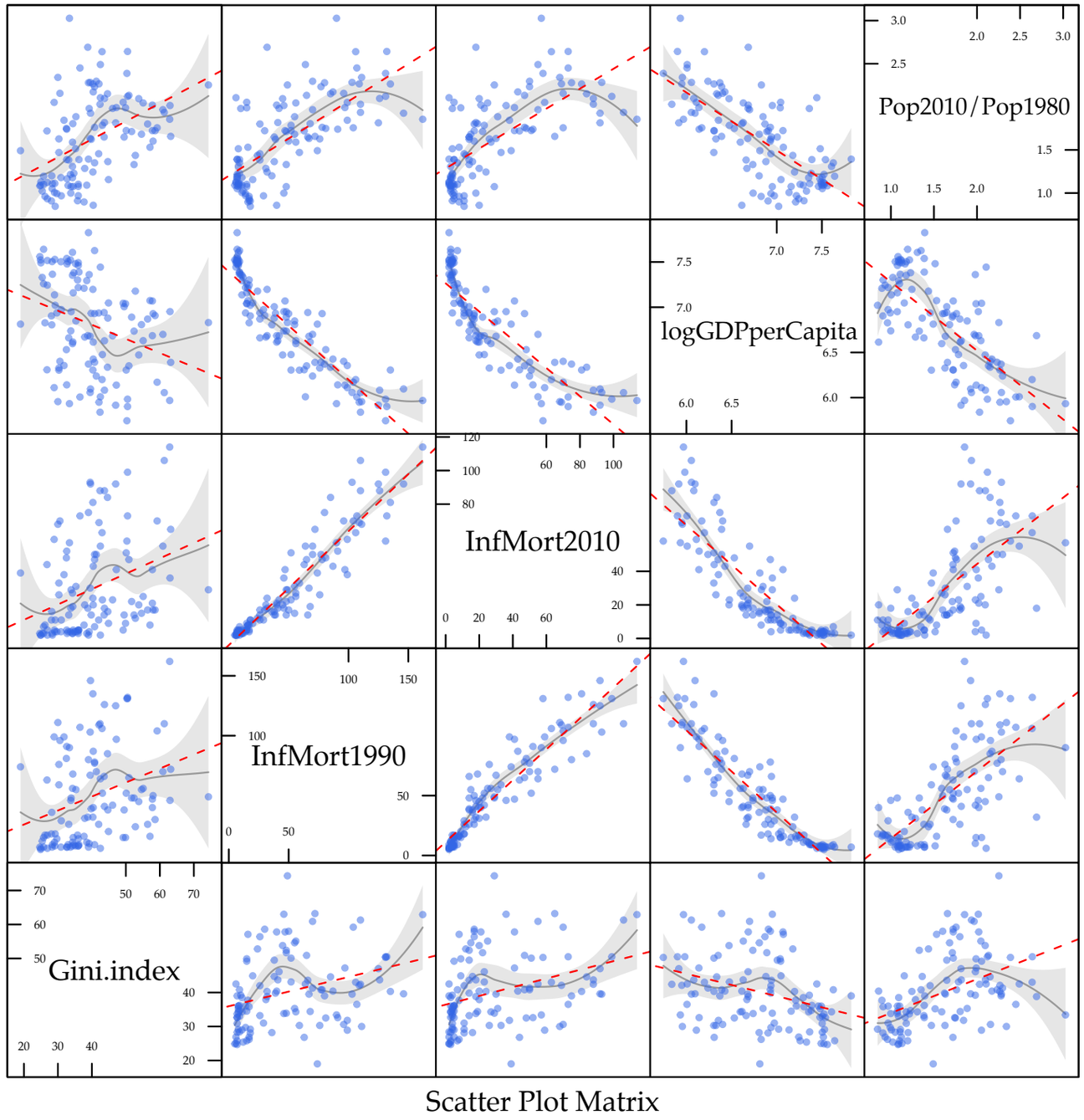
Populations - illustrating some more dimensions



Populations - illustrating some more dimensions



Another look on populations



Scatter Plot Matrix

How to:

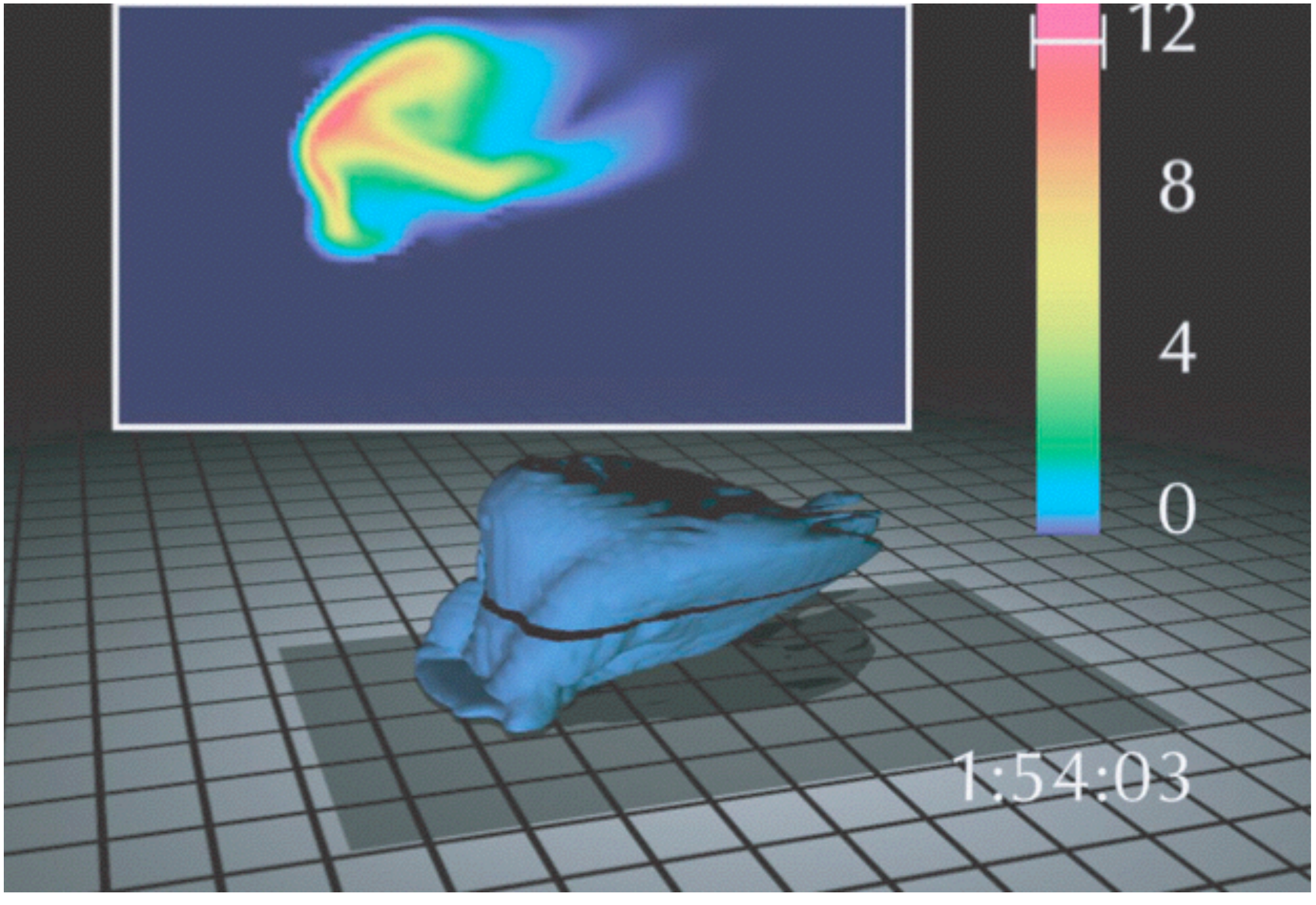
R

plot

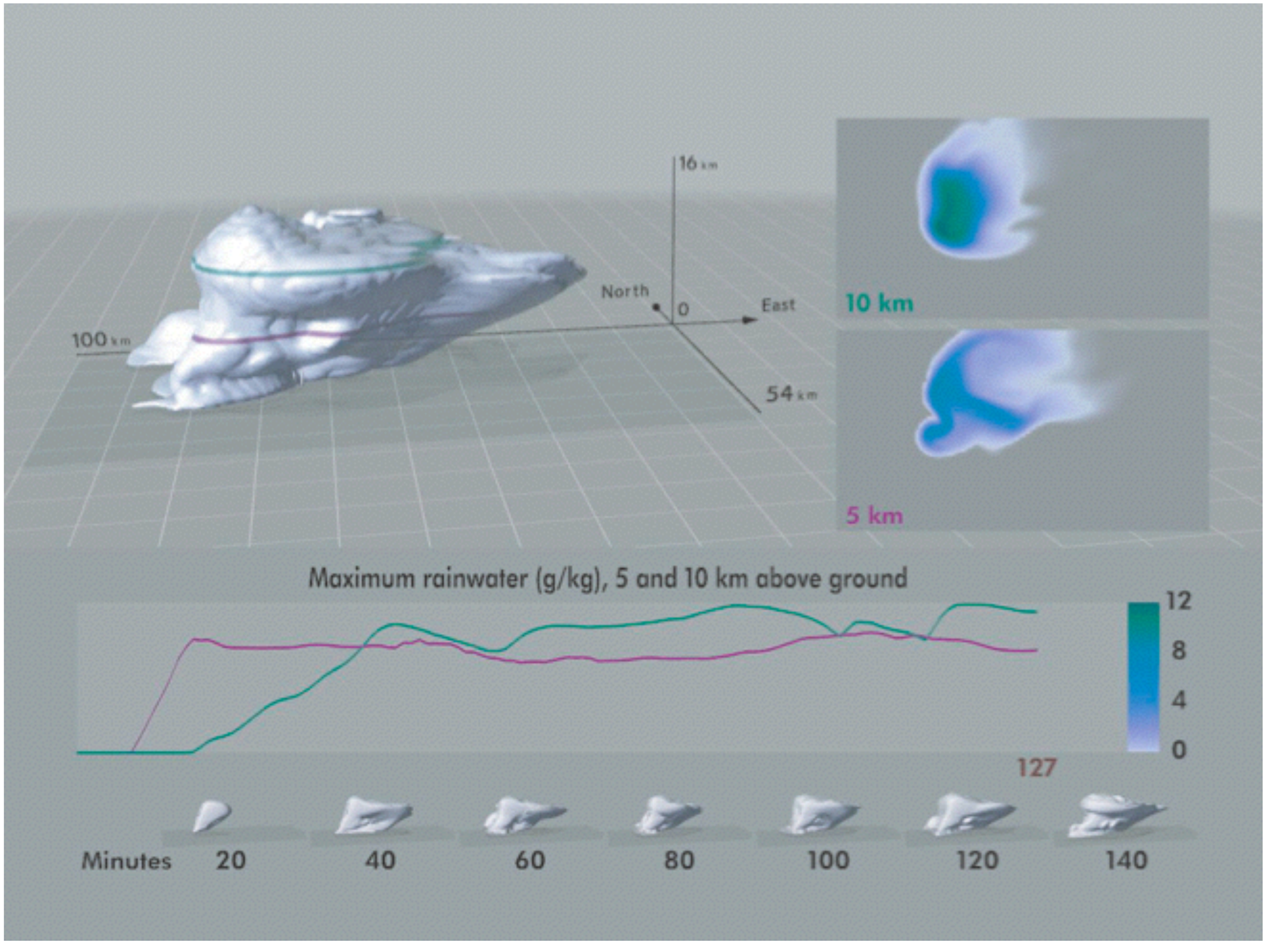
Python

from pandas.tools.plotting import scatter_matrix

Show the data & be quantitative



Show the data & be quantitative

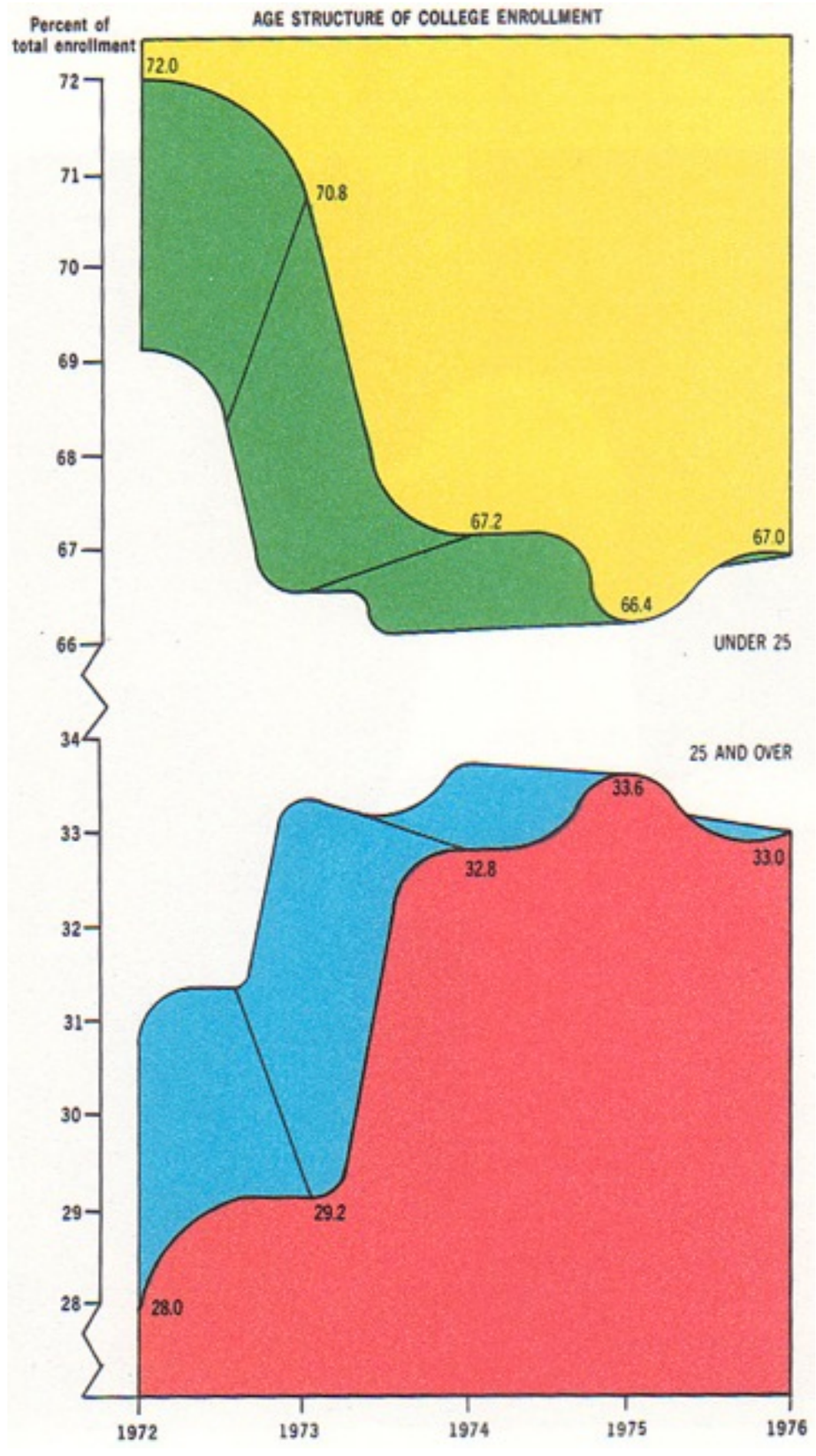


And don't waste!

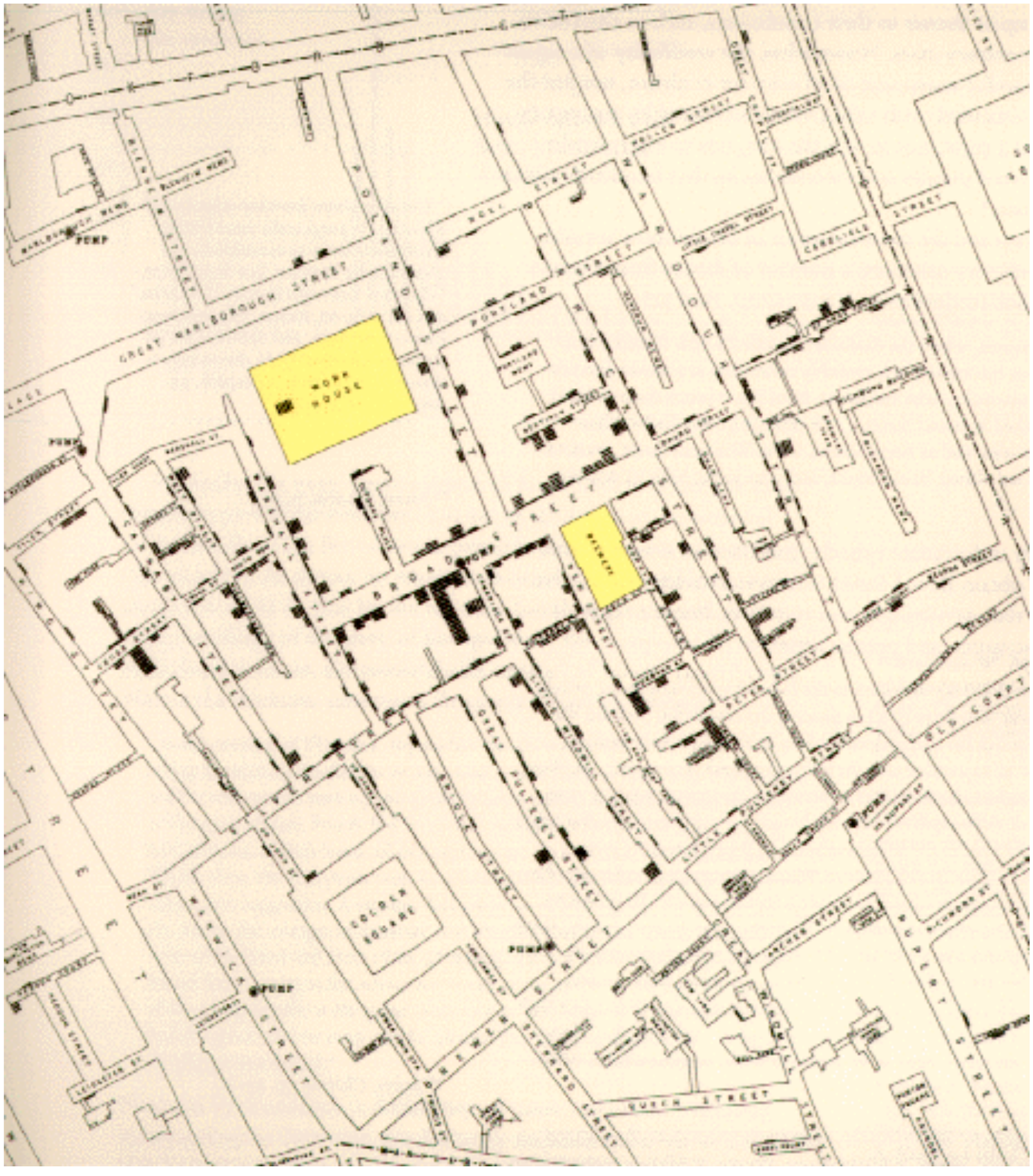
Always think about whether a given shading/symbol etc. is required. Sometimes it is absolutely essential to use different colours & symbols to bring out information.

This is not just about graphics - also be careful that you only use data that are relevant. Thinking before plotting greatly helps data exploration.

Sometimes people do not....

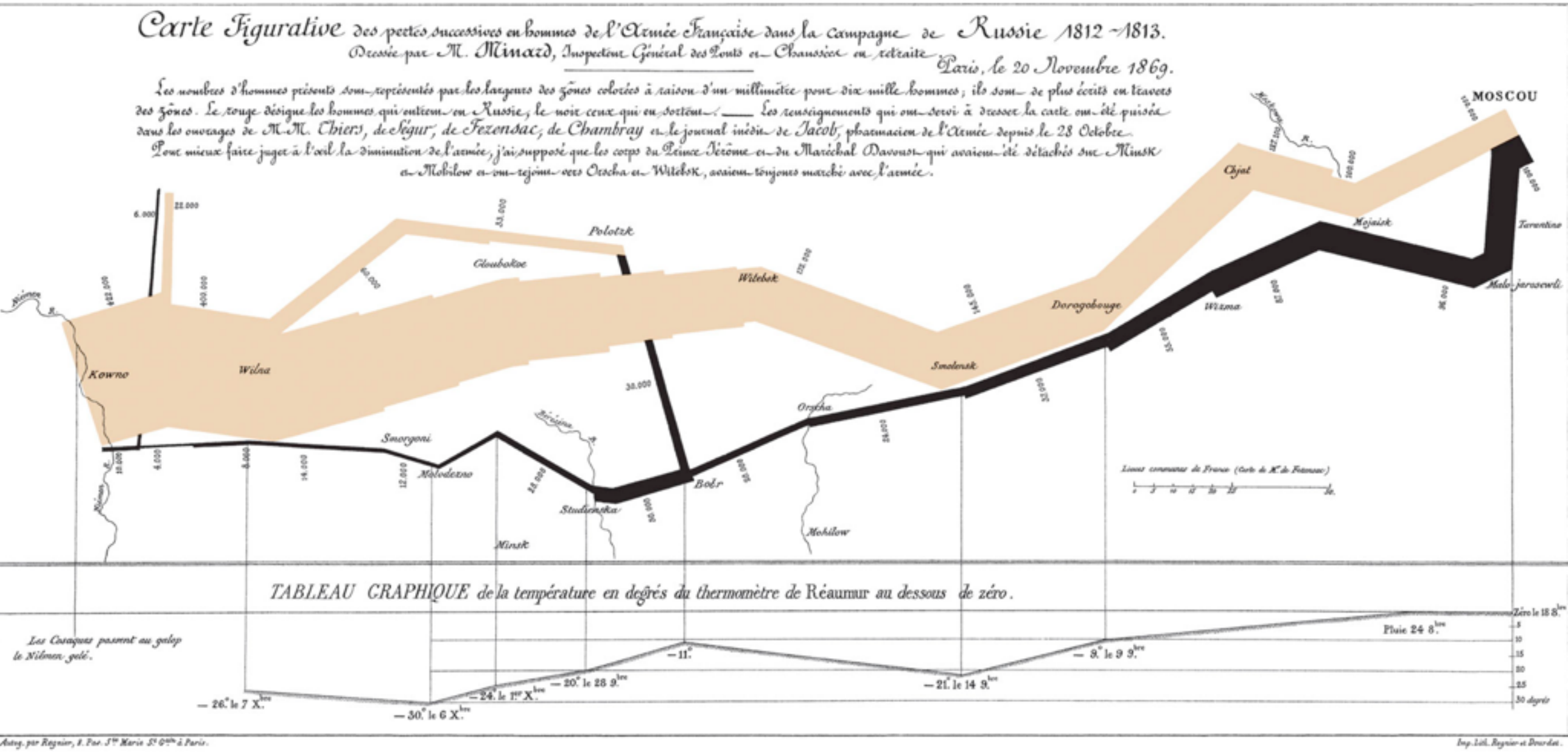


The importance of good displays of data



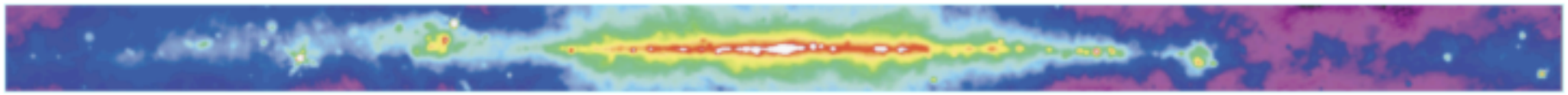
Understanding cholera.

Minard's map of Napoleon's Russian campaign

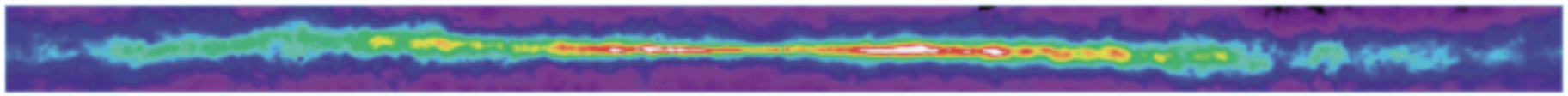


The Milky Way

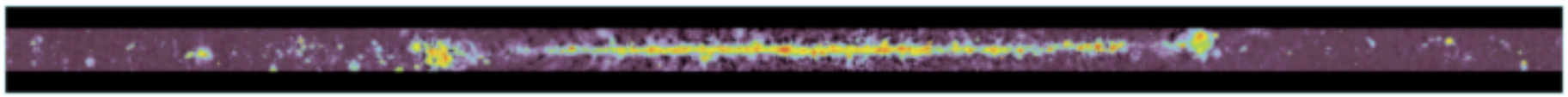
Radio (0.4 GHz)



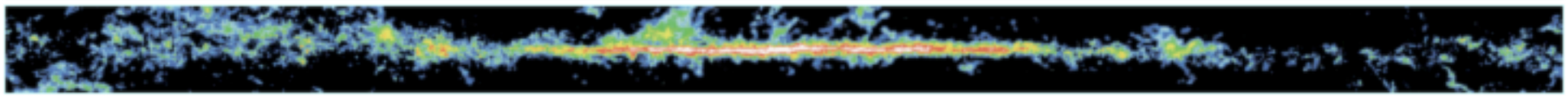
Atomic Hydrogen



Radio (2.7 GHz)

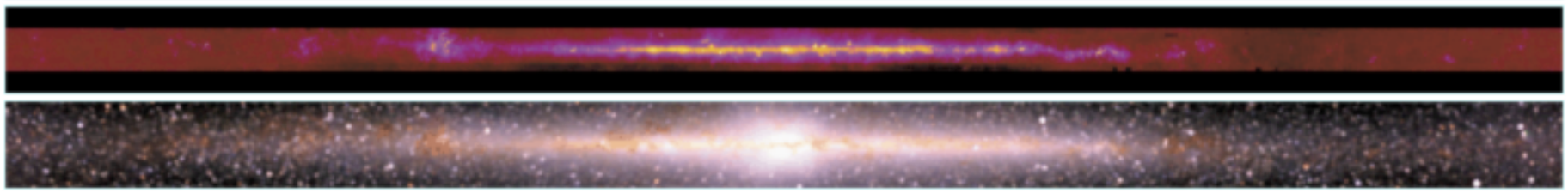


Molecular Hydrogen



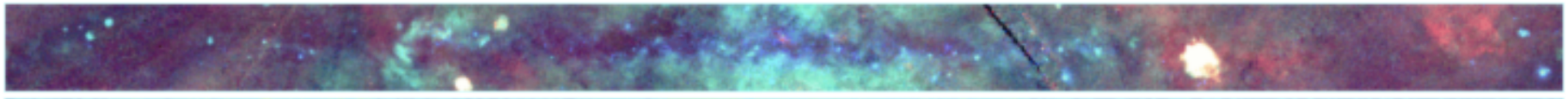
Infrared

Mid Infrared

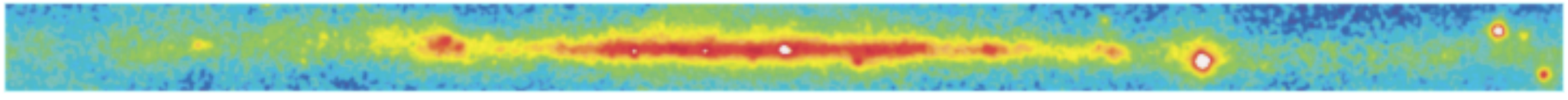


Near Infrared

Optical

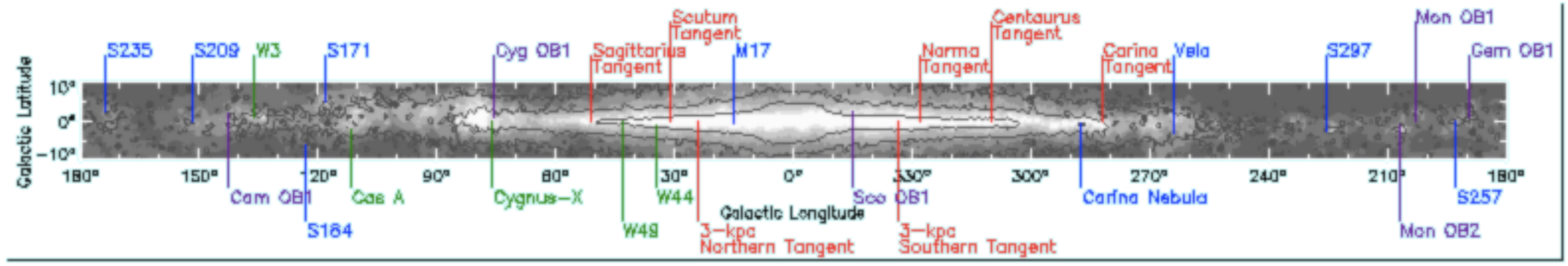


X-Ray

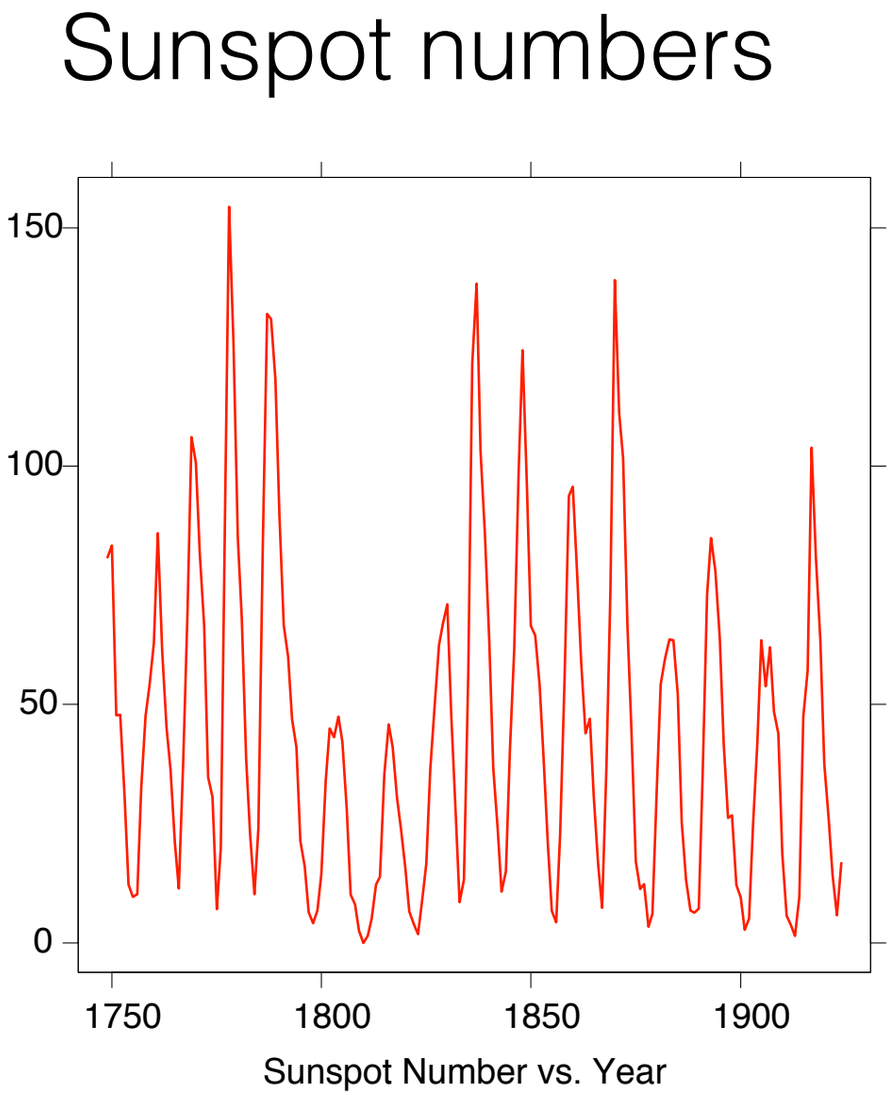


Gamma Ray

Finder

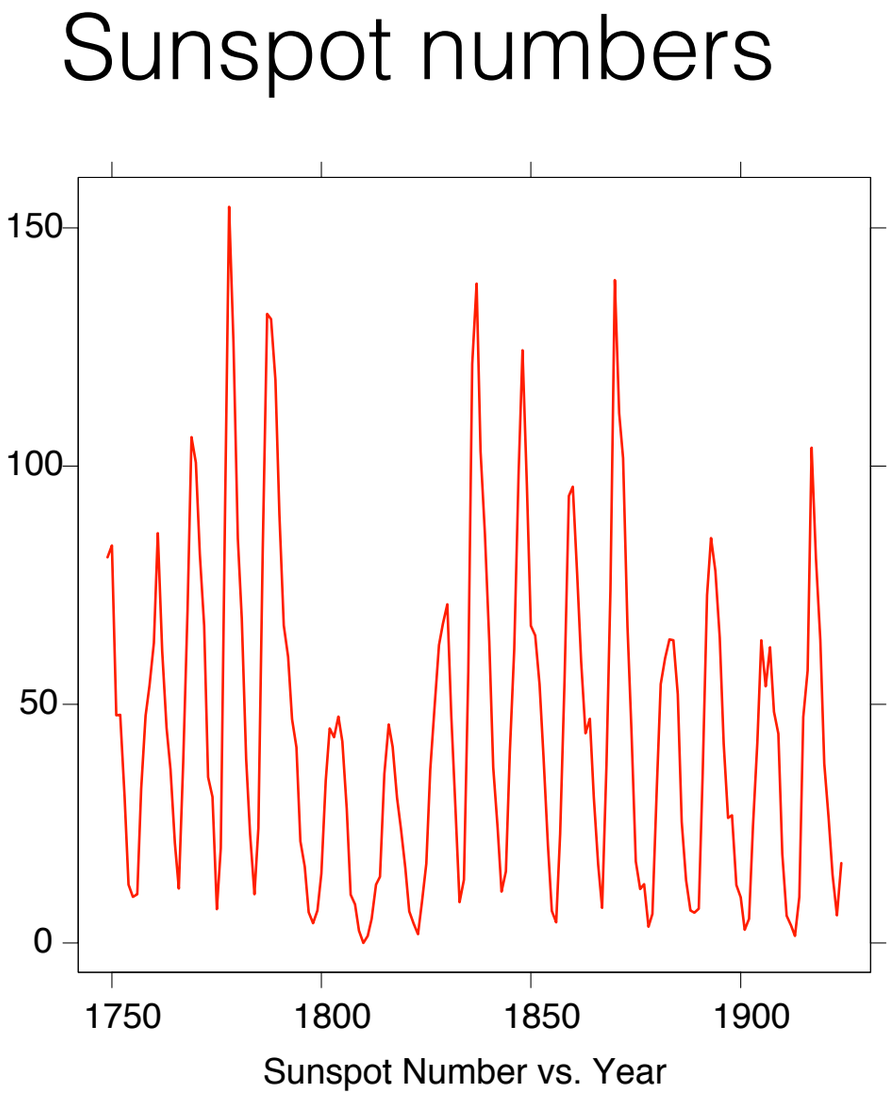


Show the data.



Tufte (1997)
Visual explanations

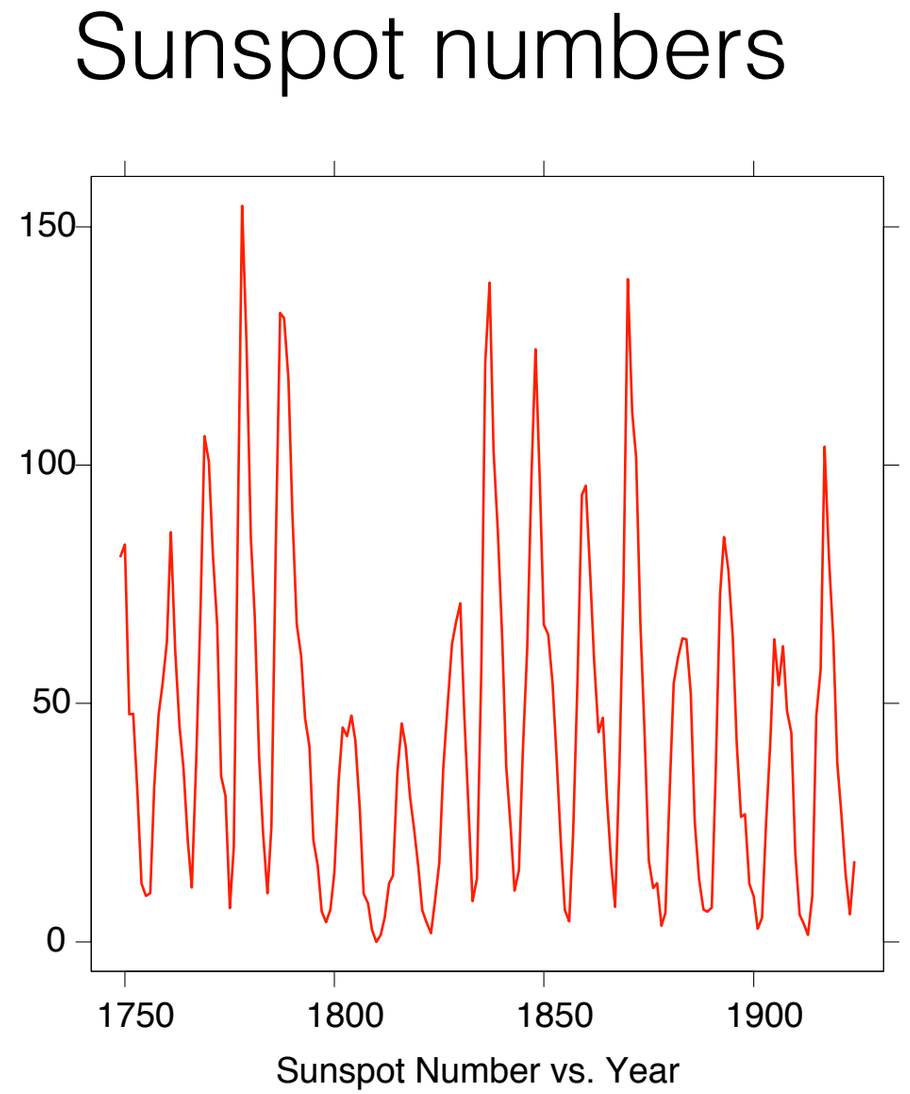
Show the data.

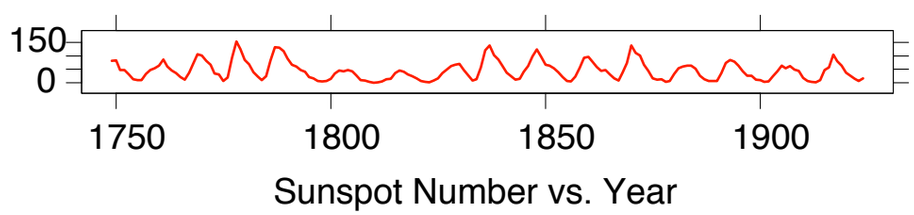


Tufte (1997)
Visual explanations

Show the data.

Tufte (1997)
Visual explanations



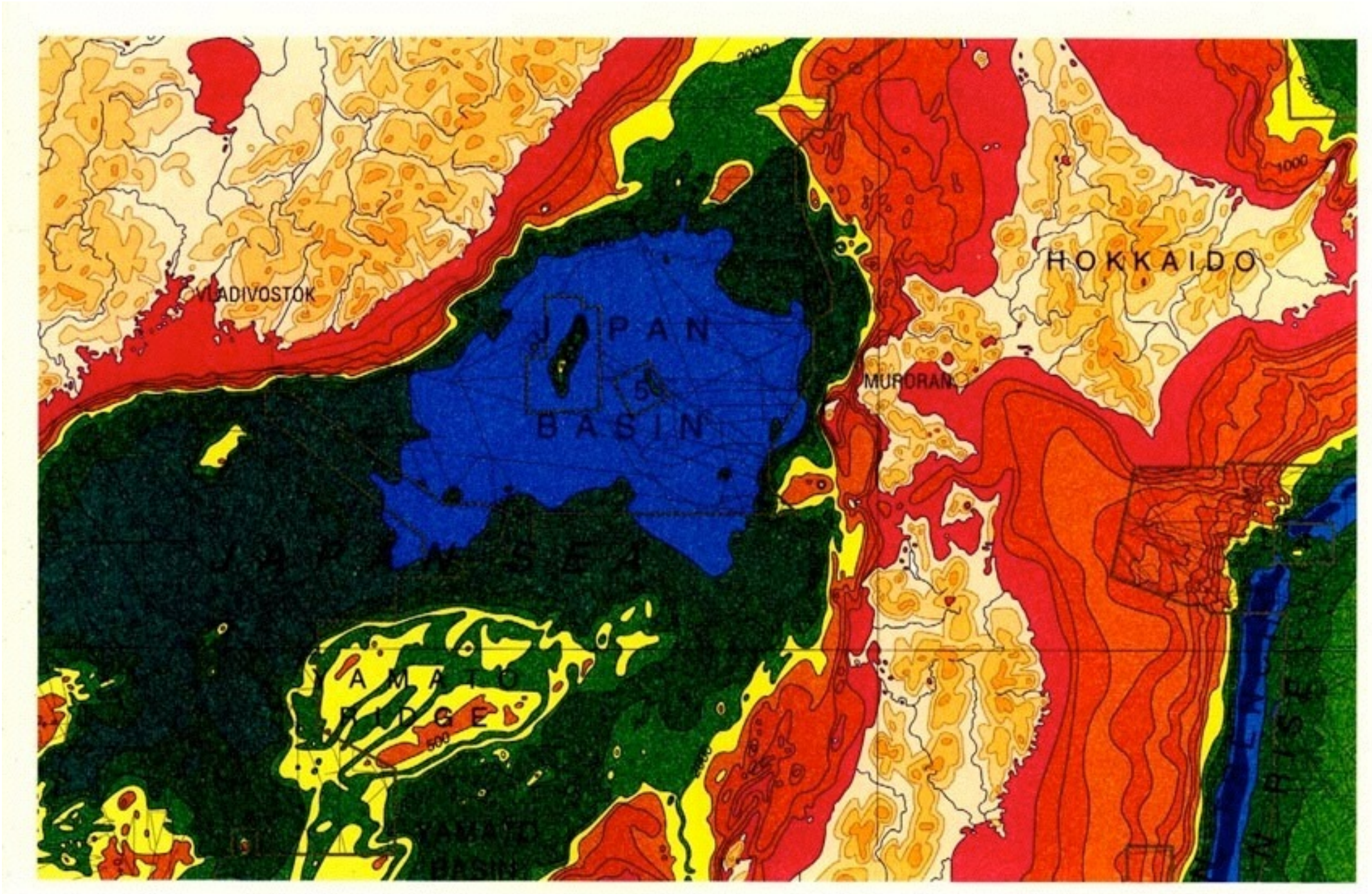


Some things to keep in mind...

The “Smallest effective difference”

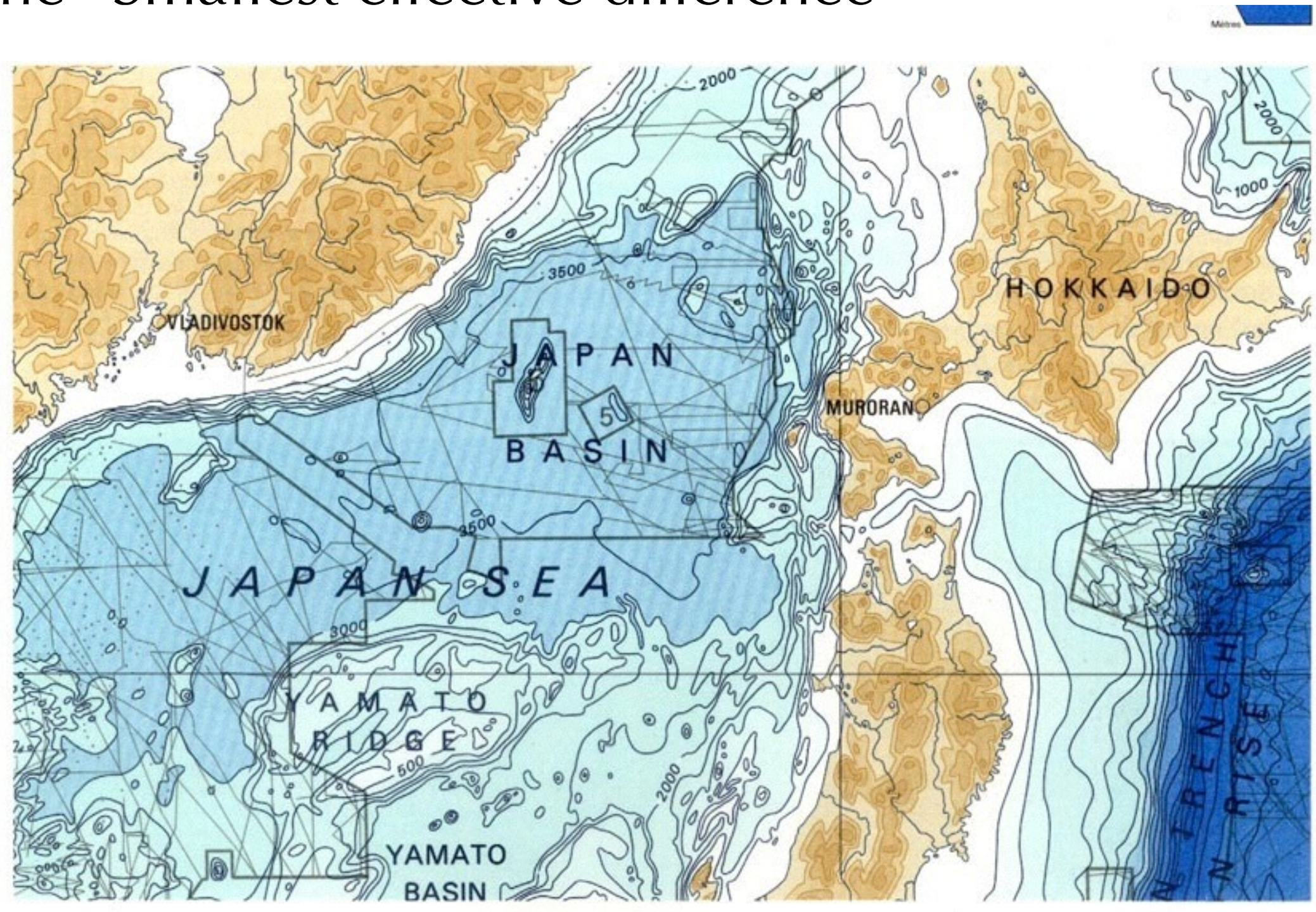
Some things to keep in mind...

The “Smallest effective difference”



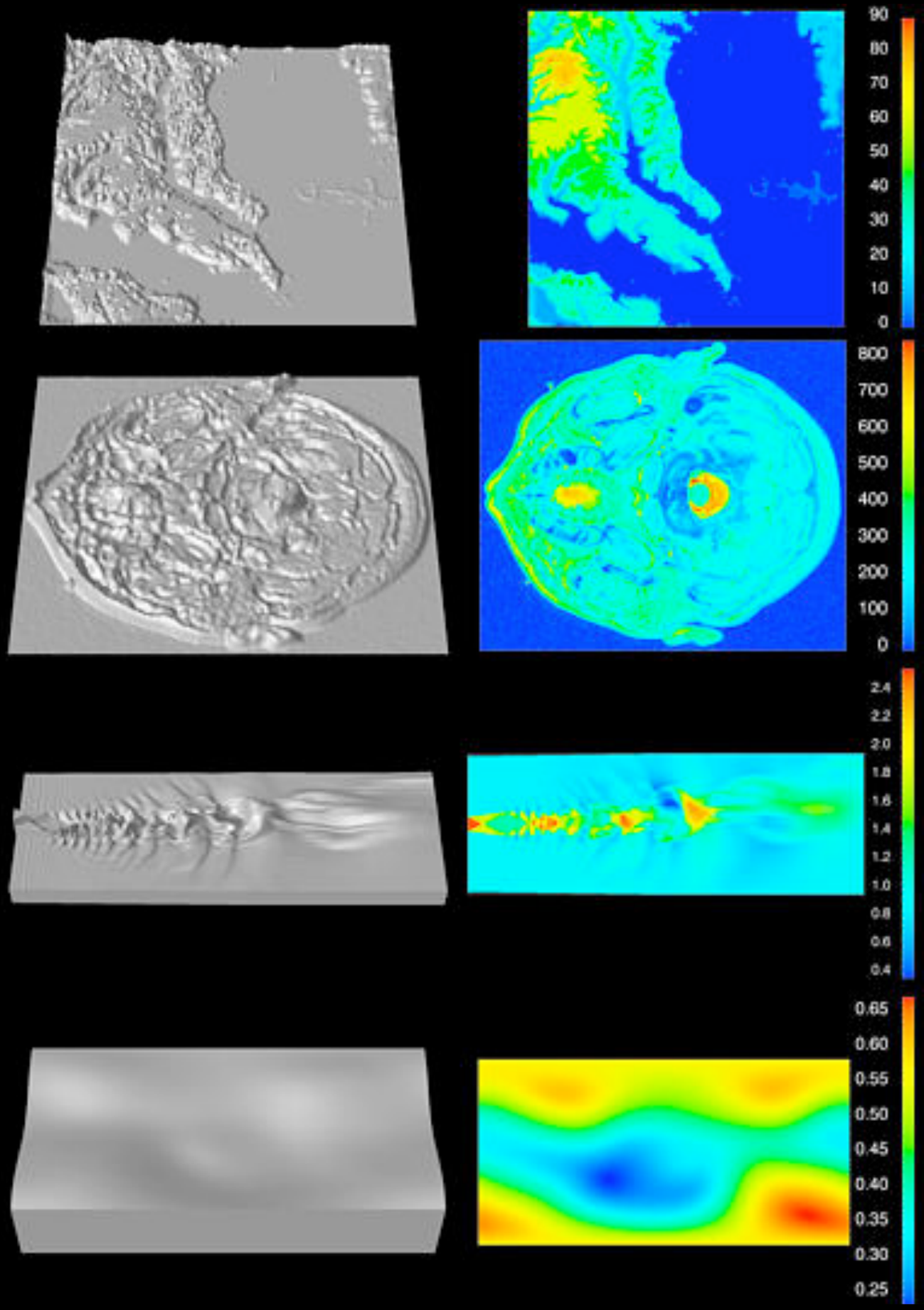
Some things to keep in mind...

The “Smallest effective difference”

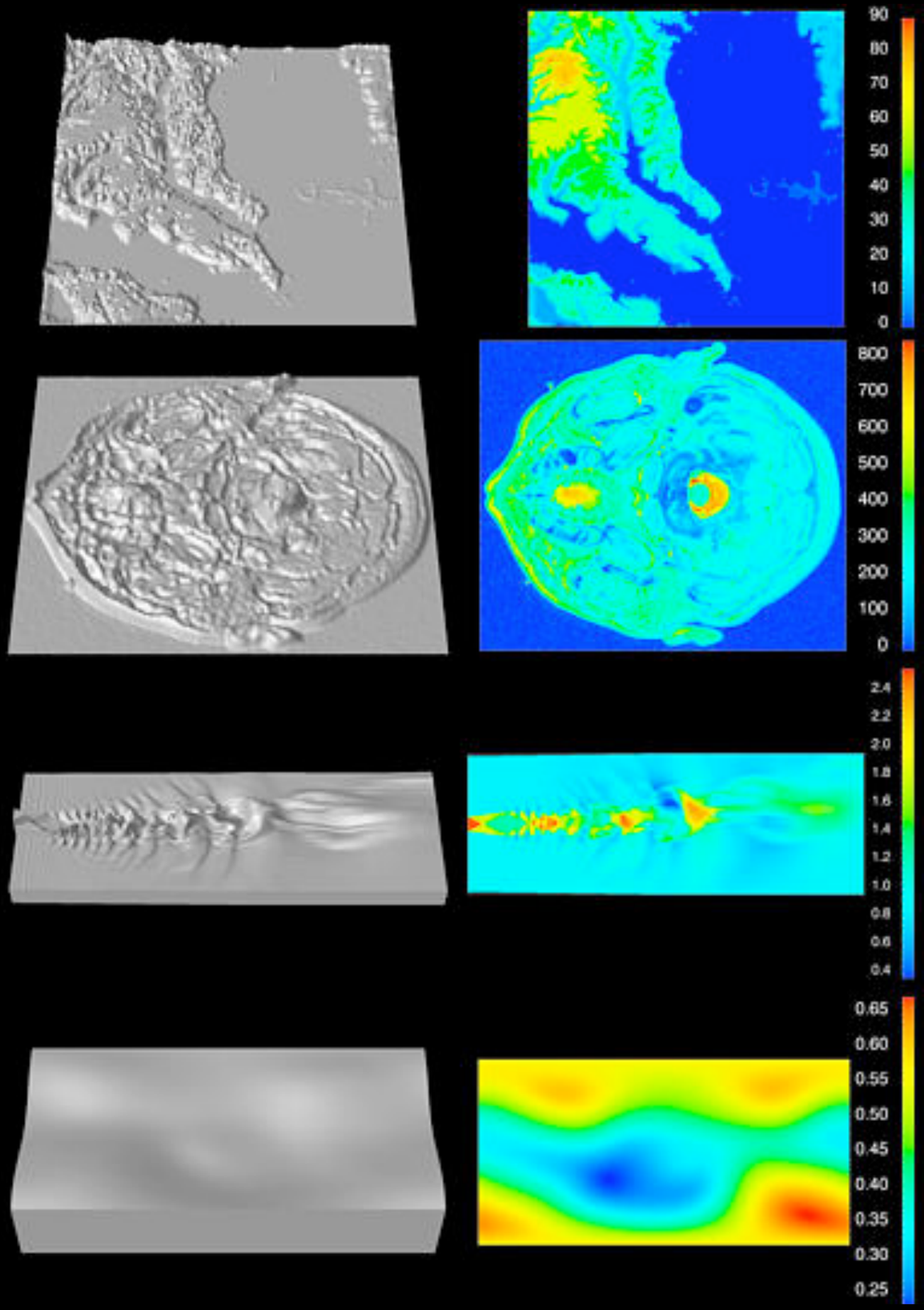


The issue of colour & 2D plots

There is a tendency for astronomers to choose from a small set of colour schemes for their images and plots. Sometimes this leads you to create artificial trends where there are none.

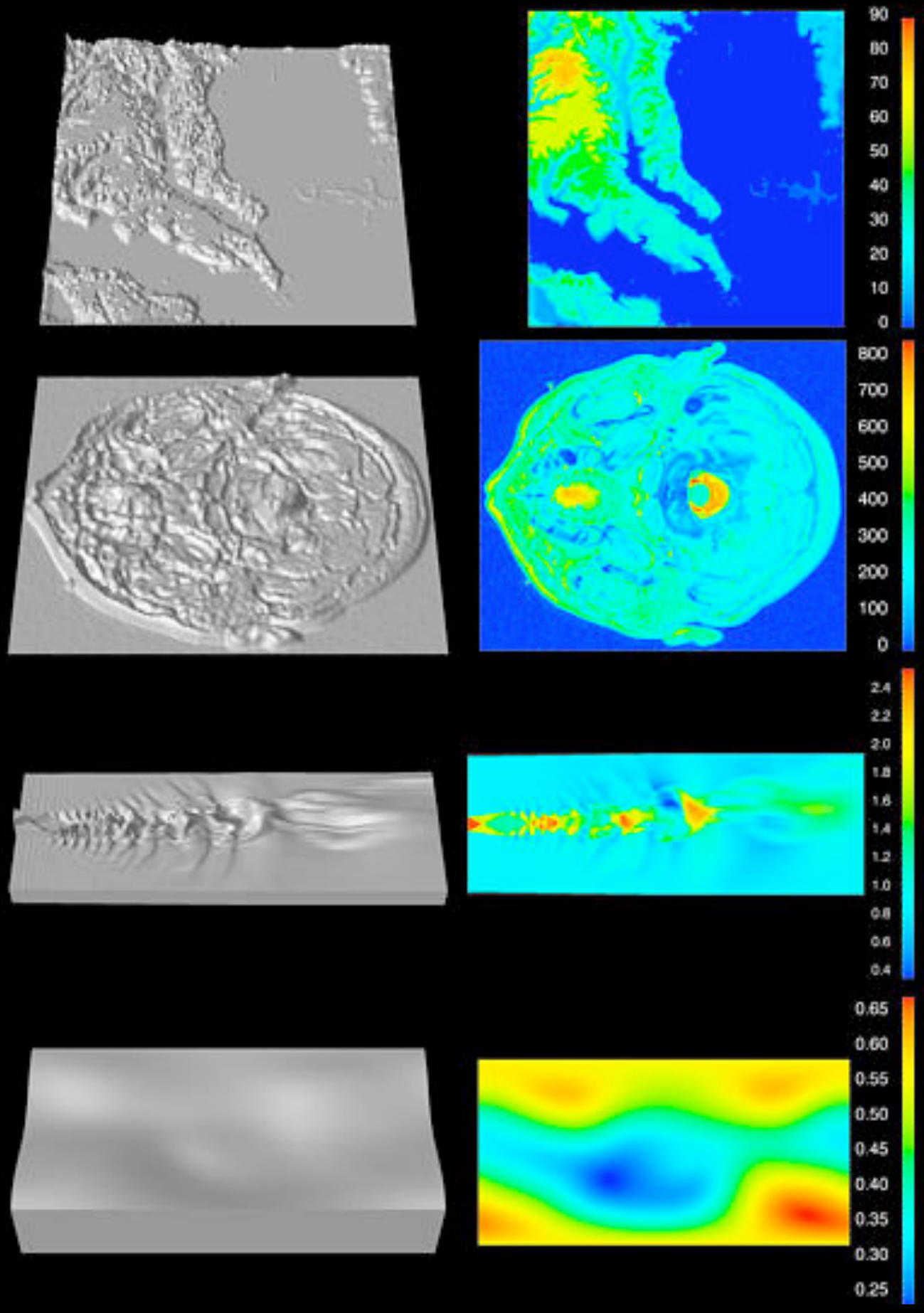

90
80
70
60
50
40
30
20
10
0
800
700
600
500
400
300
200
100
0
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25

Bad Habits



Chesapeake Bay - note the artificial structure at higher altitude when it is actually quite gradual

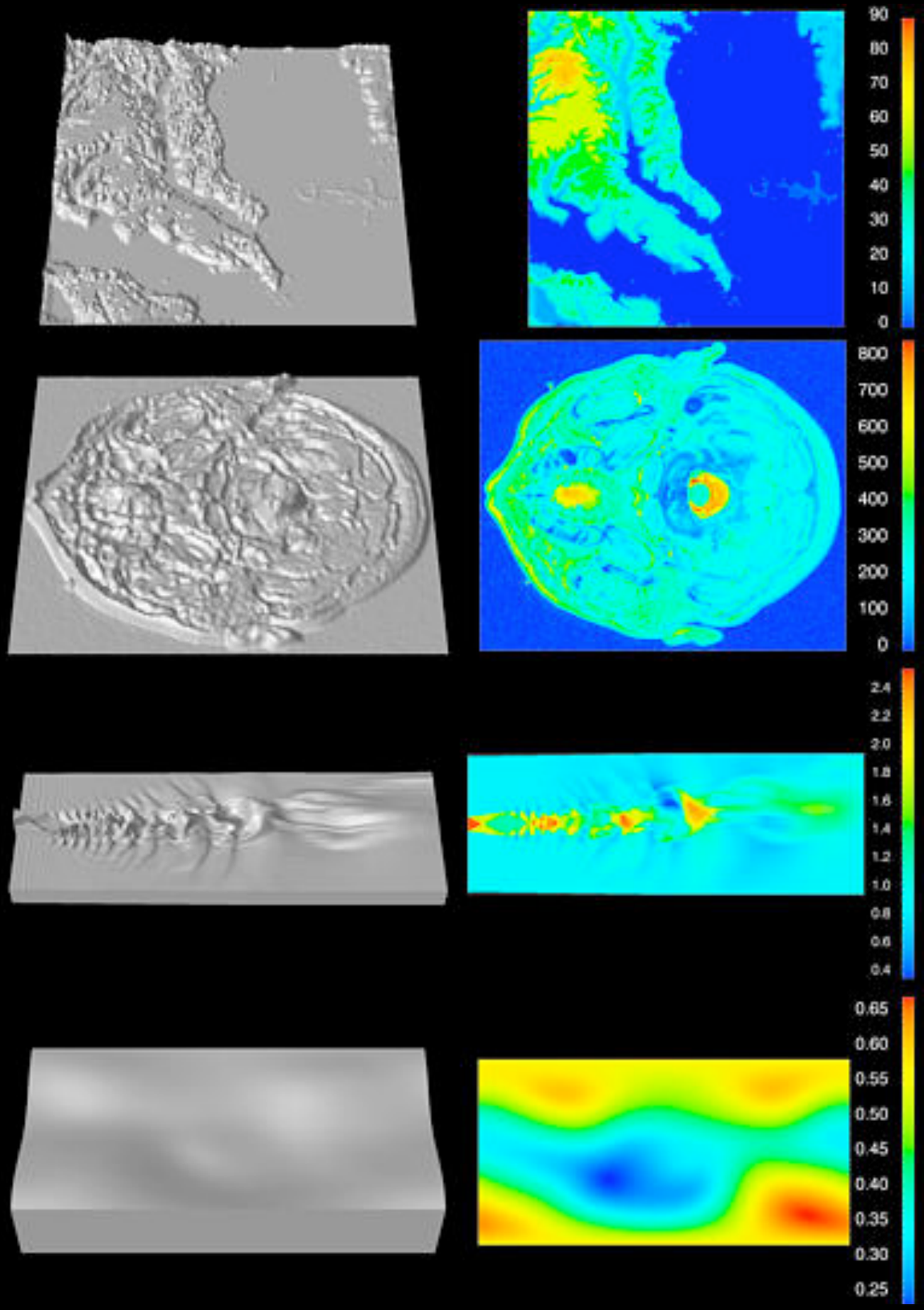
Bad Habits



Chesapeake Bay - note the artificial structure at higher altitude when it is actually quite gradual

An slice of an MRI scan of a human brain. Washing out of detail and artificial structure.

Bad Habits

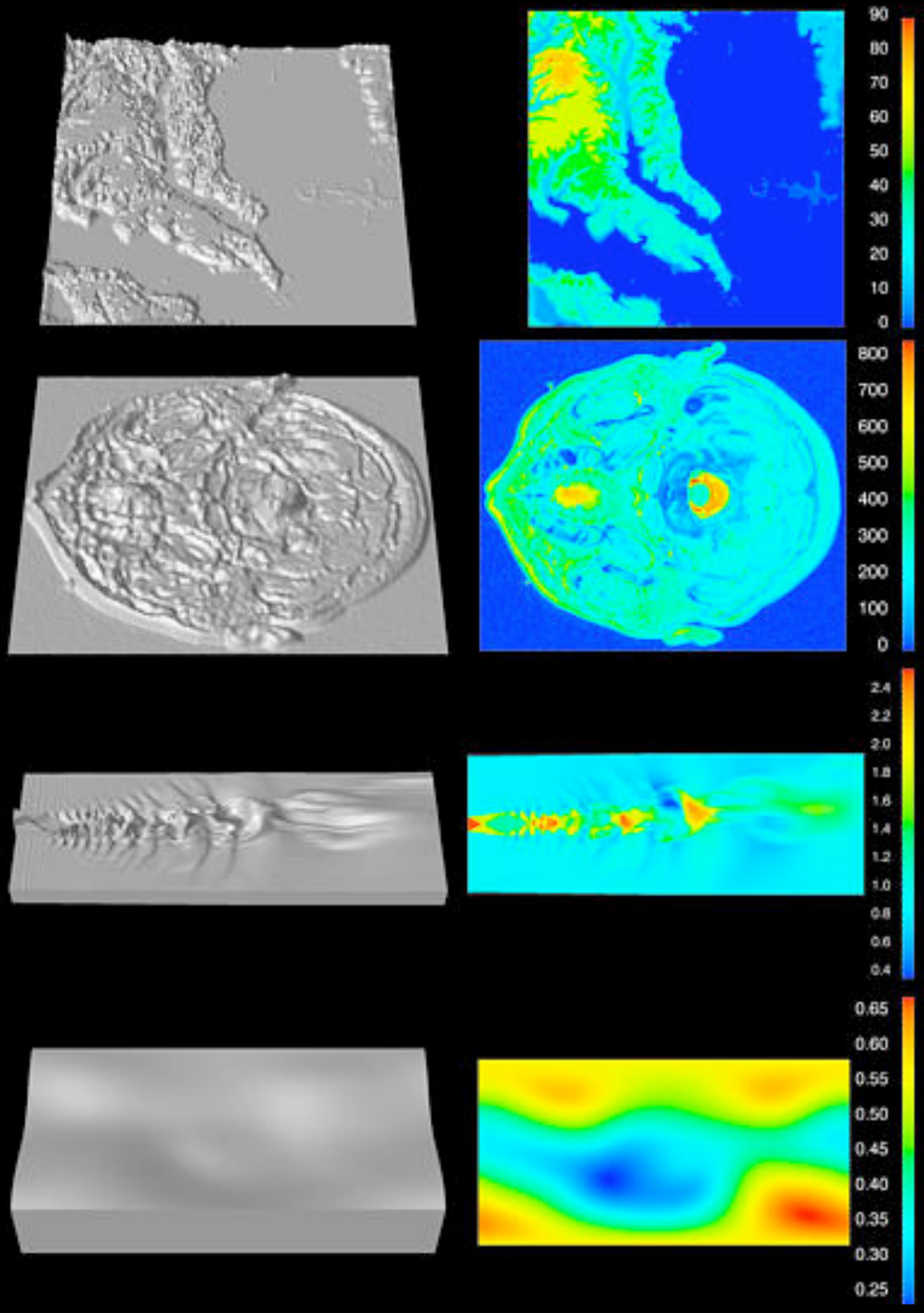


Chesapeake Bay - note the artificial structure at higher altitude when it is actually quite gradual

An slice of an MRI scan of a human brain. Washing out of detail and artificial structure.

Bad Habits

Turbulent flow from a jet engine.



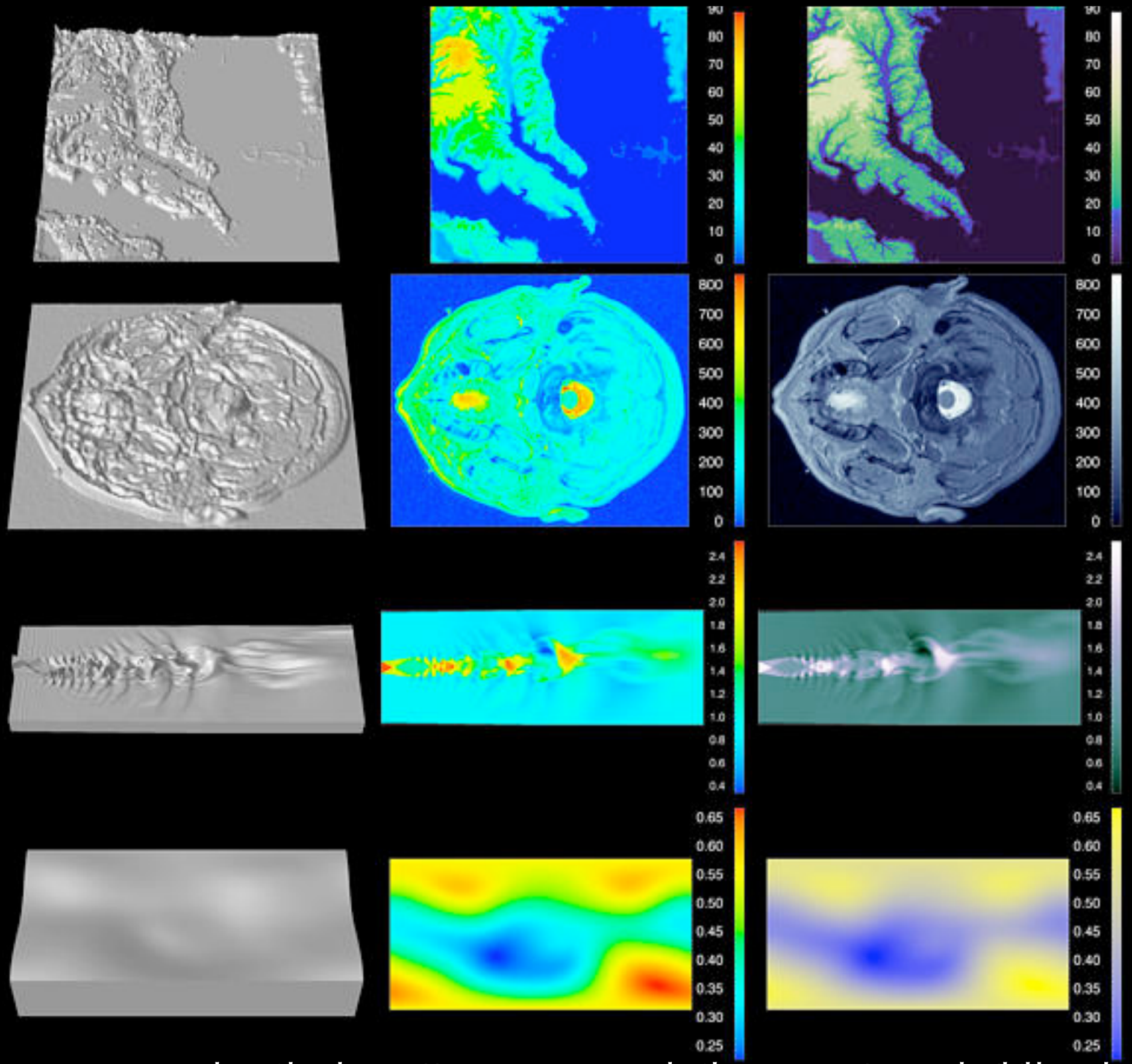
Chesapeake Bay - note the artificial structure at higher altitude when it is actually quite gradual

An slice of an MRI scan of a human brain. Washing out of detail and artificial structure.

Bad Habits

Turbulent flow from a jet engine.

Earth's magnetic field in a Cartesian projection - note the smooth structure.

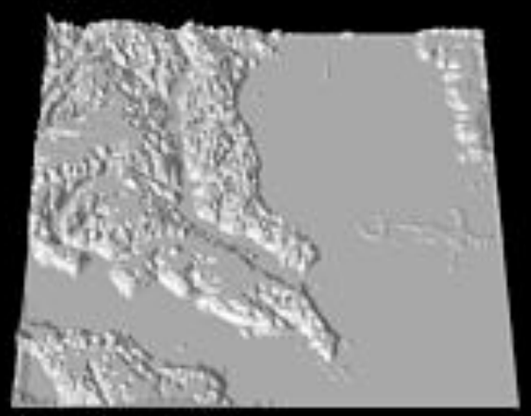


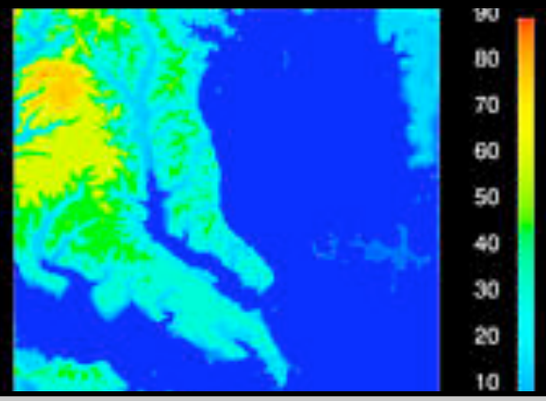
A natural zero

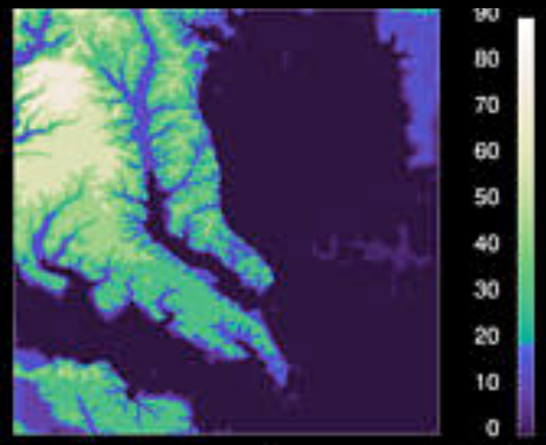
High frequency information - best with variation of luminance

Low frequency information - colour variation (saturation is good)

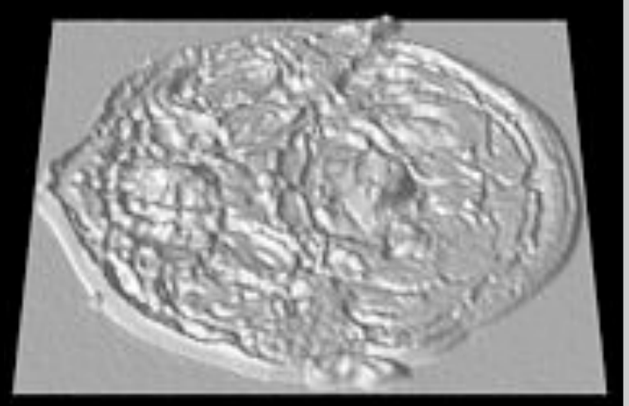
For more details: http://www.research.ibm.com/people/l/lloydt/color/color.HTM





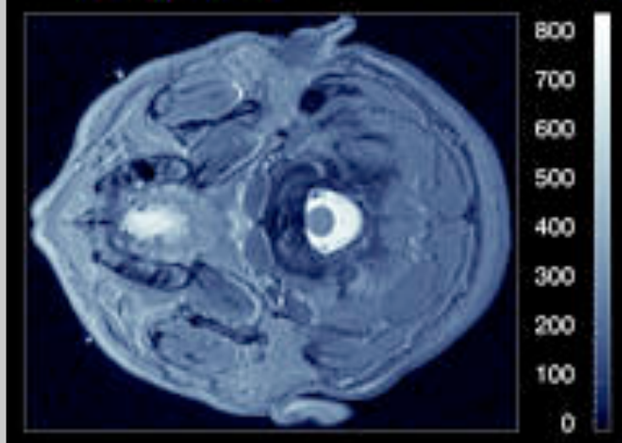


A natural zero

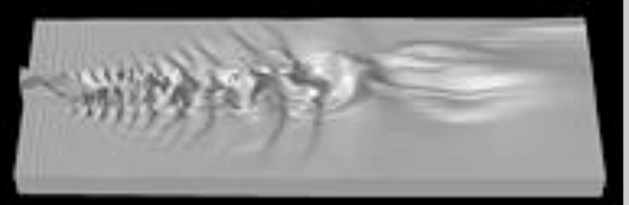


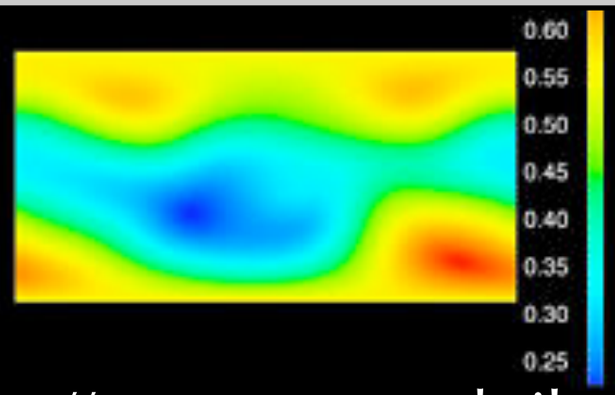
One key lesson:

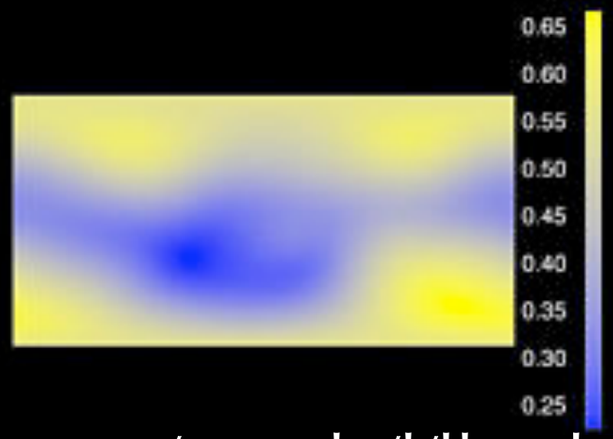
Try different colour maps and do not use a default one mindlessly.



High frequency information - best with variation of luminance



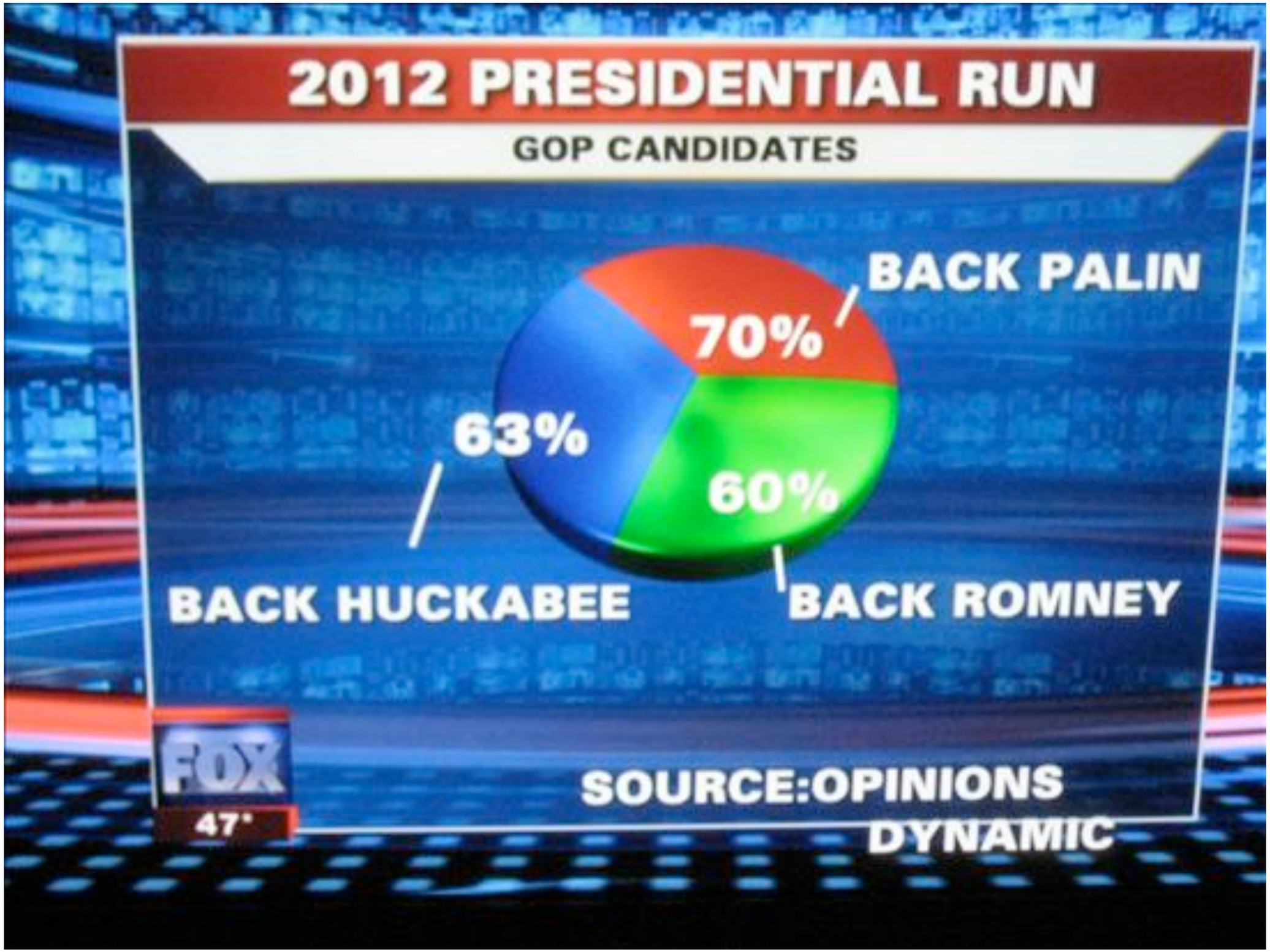




Low frequency information - colour variation (saturation is good)

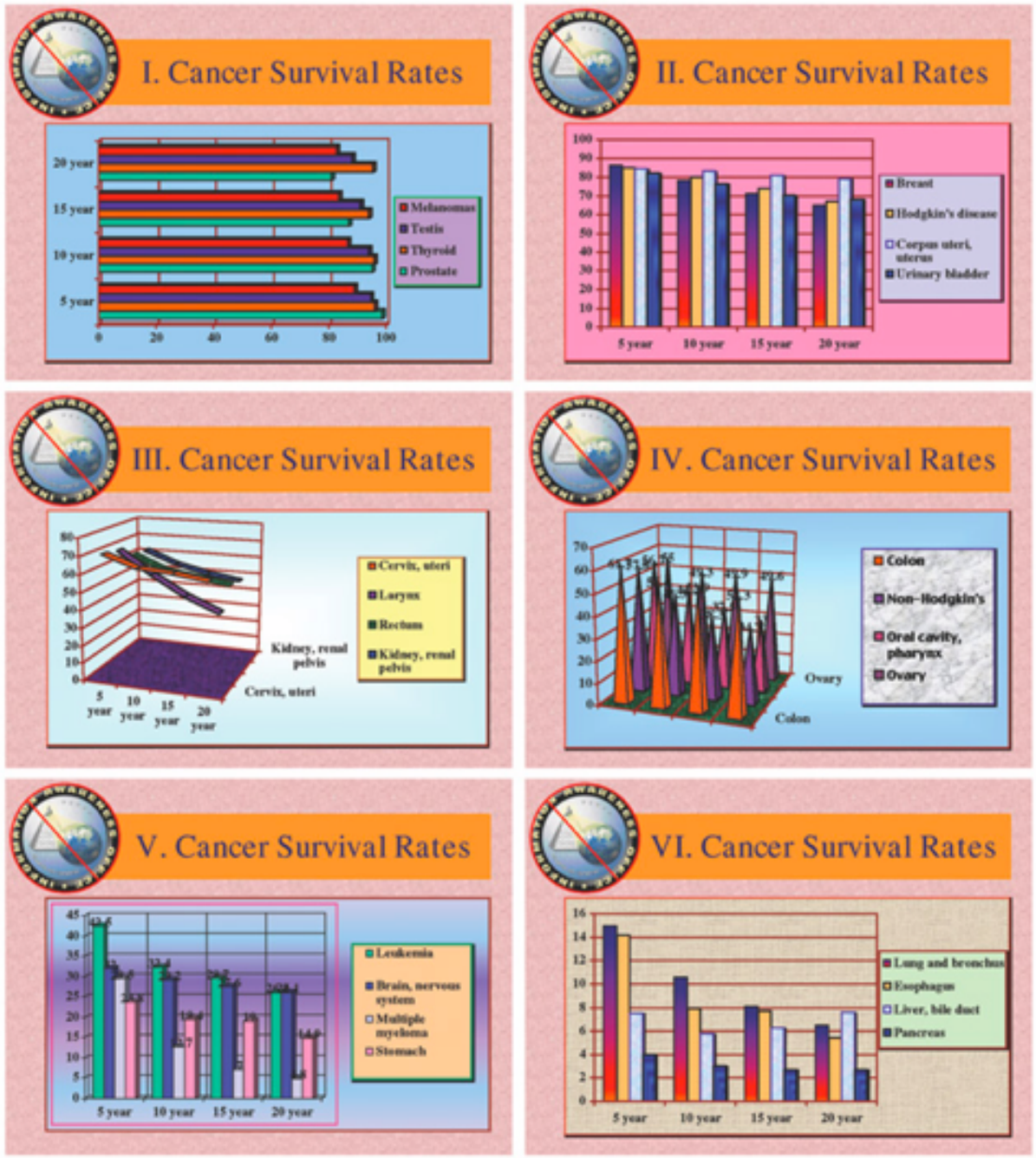
For more details: http://www.research.ibm.com/people/l/lloydt/color/color.HTM

Some bad graphics



An alternative view of a table

But don't do:



http://tinyurl.com/y7uhpp6

The Spitzer Infrared Nearby Galaxies Survey (SINGS) Hubble Tuning-Fork

The Spitzer Space Telescope observed 75 galaxies as part of its SINGS (Spitzer Infrared Nearby Galaxies Survey) Legacy Program. The galaxies are presented here in a Hubble Tuning-Fork diagram, which groups galaxies according to the morphology of their nuclei and spiral arms. The designation of these galaxies and their placement in the diagram is based on their visible-light appearance. The main goal of the SINGS program is to characterize the infrared properties of a wide range of galaxy types. The images of the galaxies are composites created from data taken by IRAC (the Infrared Array Camera) at 3.6 and 8.0 μm, and MIPS (the Multiband Imaging Photometer for Spitzer) at 24 μm.

The infrared range probed by these and other observations taken for the SINGS project allows for the detailed study of star formation, dust emission, and the distribution of stars in each galaxy. Light from old stars appears as blue in the images, while the lumpy knots of green and red light are produced by dust clouds surrounding newly born stars. The elliptical galaxies on the left are almost entirely made of old stars, while spiral galaxies like our own Milky Way are rich in young stars and the raw materials for future star formation.

More information can be found at:
http://sings.stsci.edu/

Ellipticals

Irregulars

Strong Bulge

Strong Bulge

Weak Bulge

Weak Bulge

Unbarred Spirals

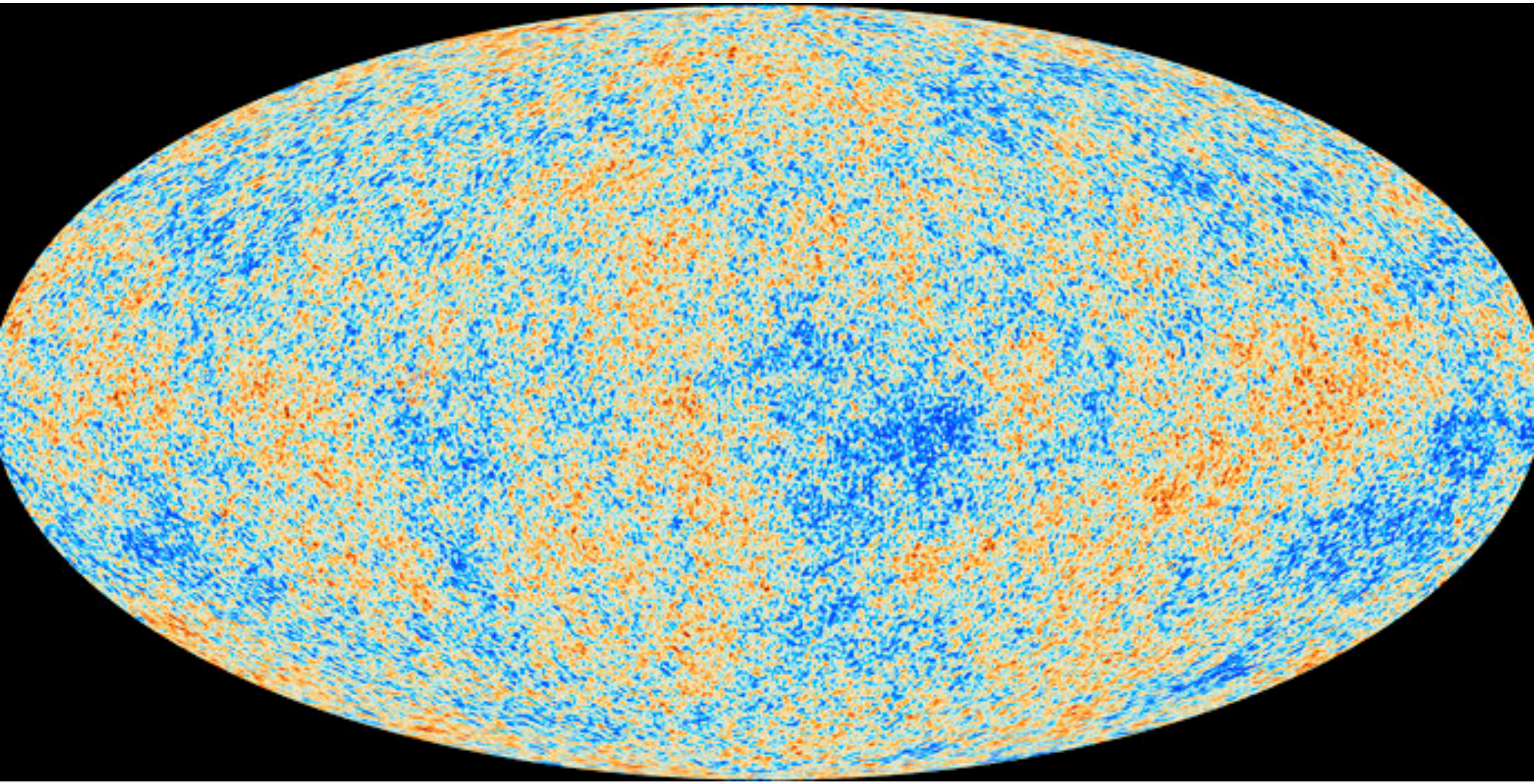
Intermediate Spirals

Barred Spirals

Poster and composite images created from SINGS observations by Karl D. Gordon (Oct 2007)
BlueIRAC 3.6μm (stars)
GreenIRAC 8μm
(aromatic features from dust grains/molecules)

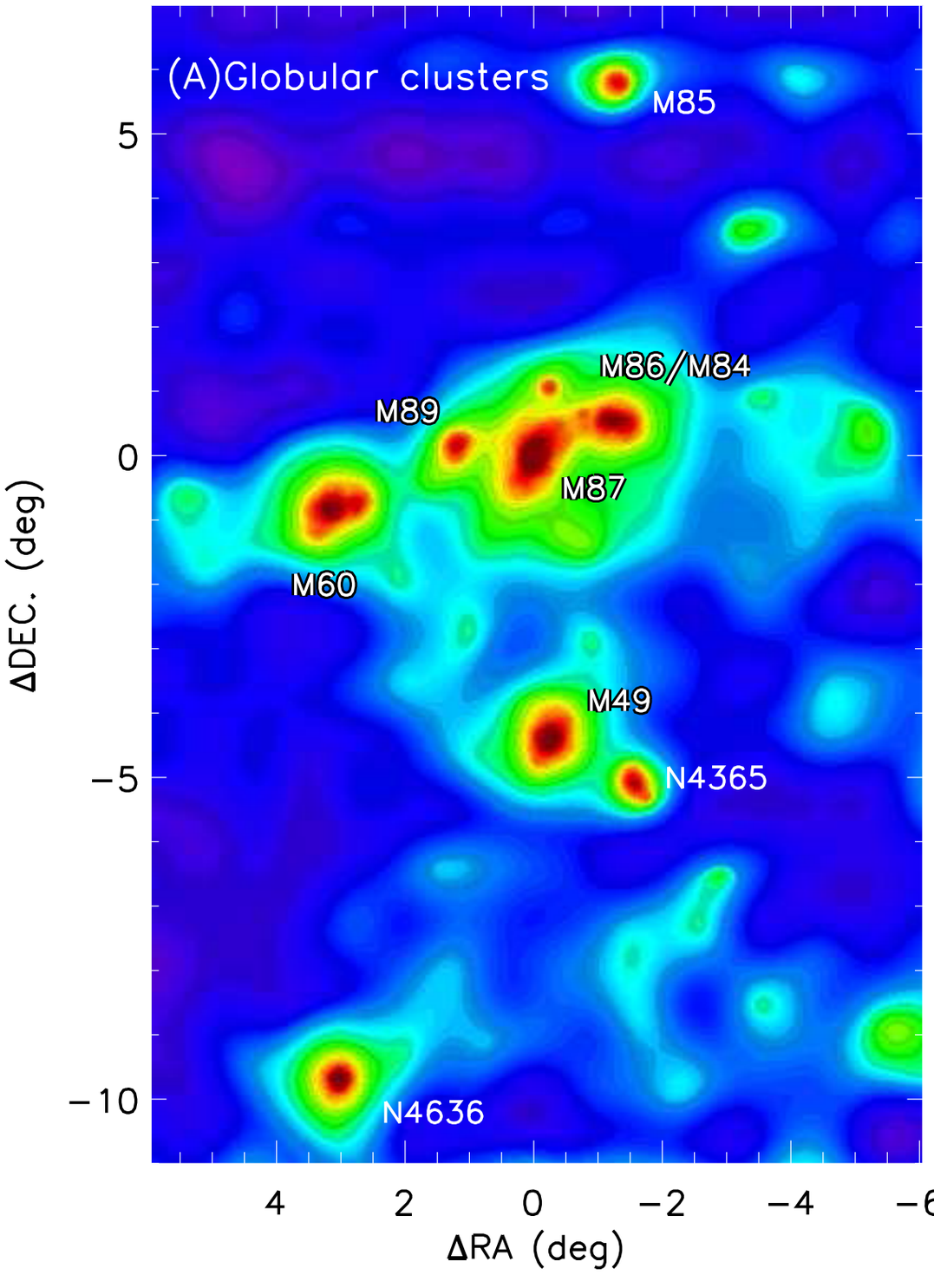
SINGS Team
Robert Kennicutt, Jr. (Principle Investigator), Daniela Calzetti (Deputy Principle Investigator), Charles Engelbracht (Technical Contact), Lee Armus, George Bendo, Caroline Bot, Brent Buckalew, John Cannon, Daniel Dale, Bruce Draine, Karl Gordon, Albert Grauer, David Hollenbach, Tom Jarrett, Lisa Kewley, Claus Leitherer, Aigen Li, Sangeeta Malhotra, Martin Meyer, John Moustakas, Eric Murphy, Michael Regan, George Rieke, Marcia Rieke, Helene Roussel, Kartik Sheth, J.D. Smith, Michele

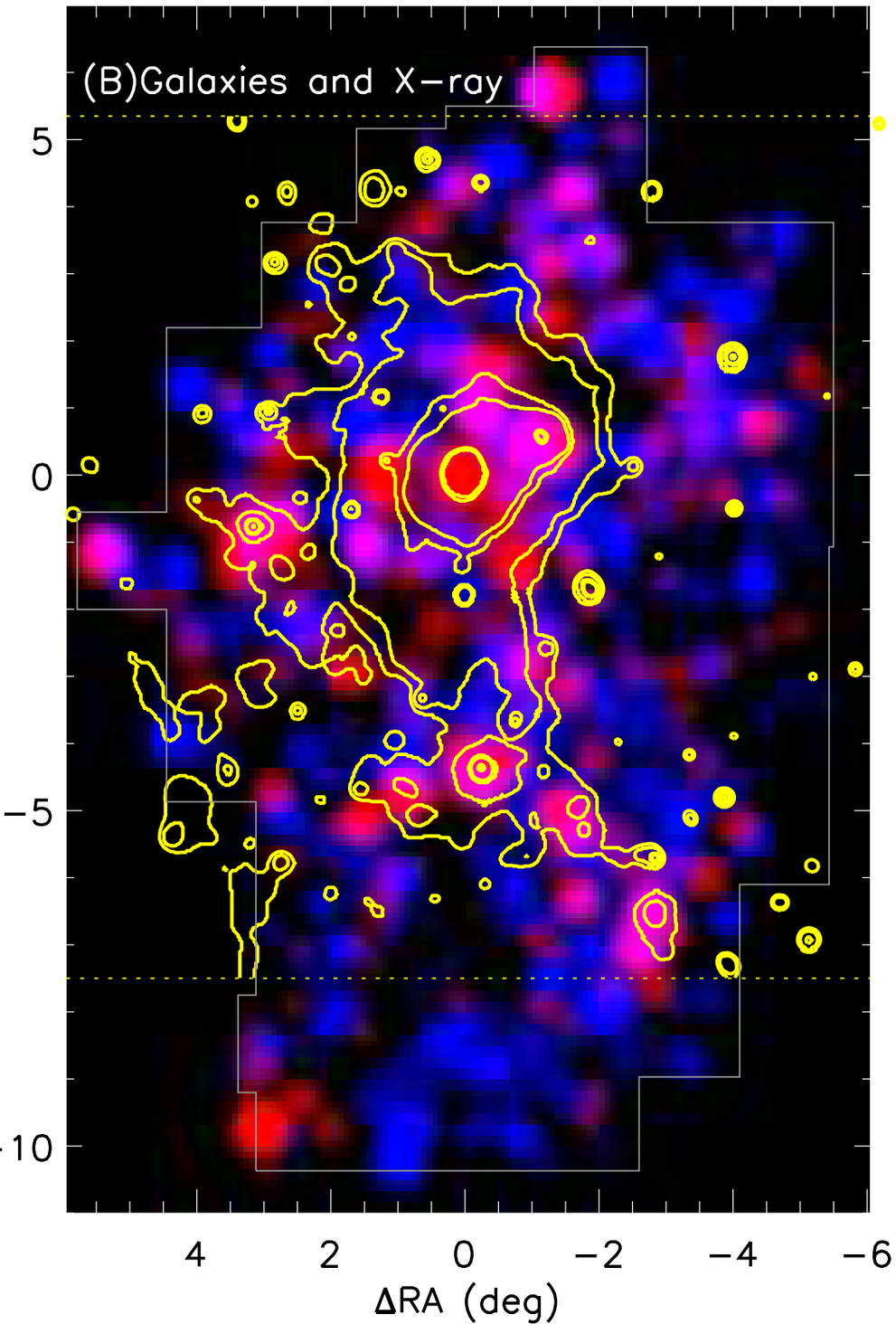
Planck - fluctuations in the temperature of the Cosmic Microwave Background



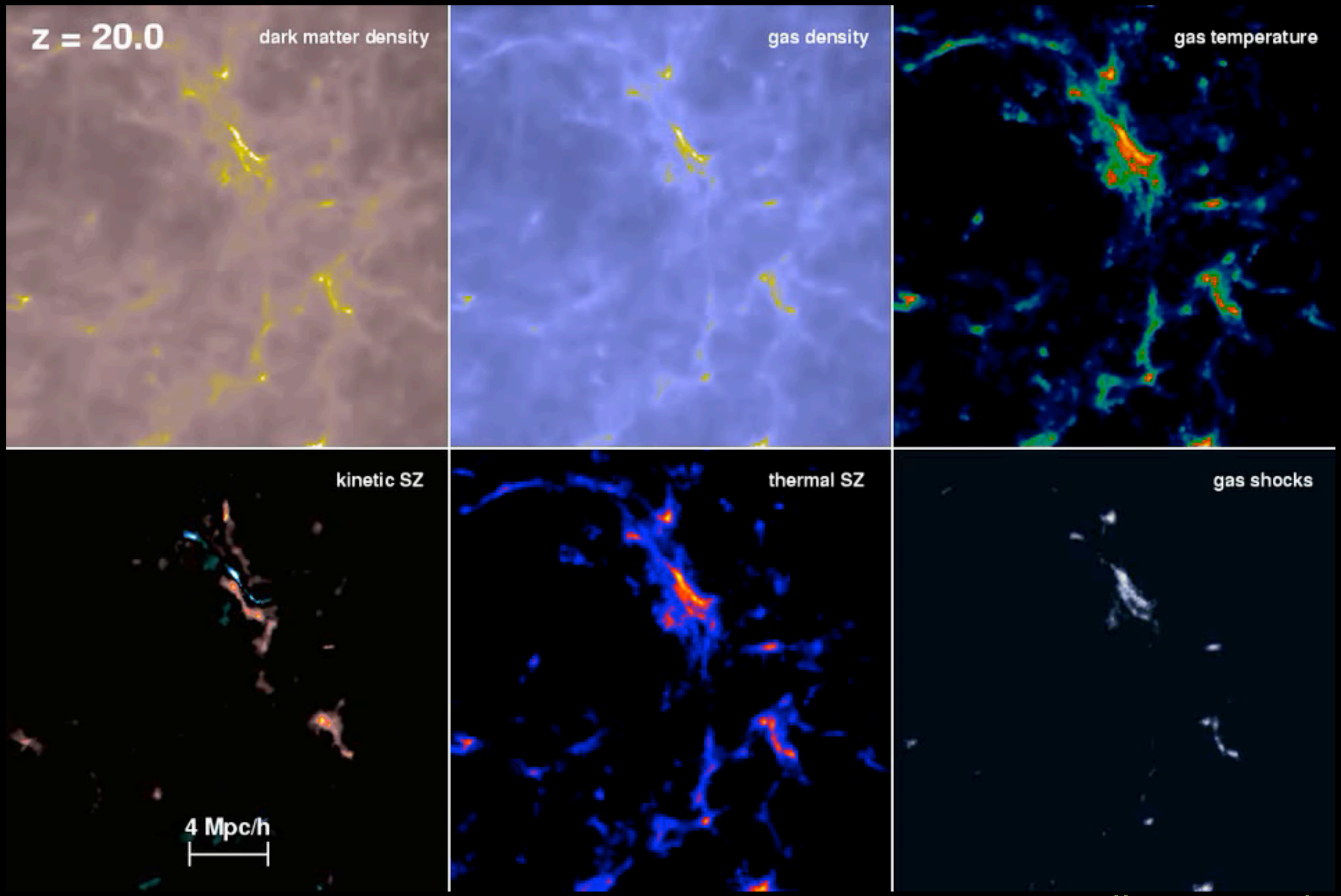
Globular clusters in the Virgo cluster

Lee, Park & Hwang (2010)



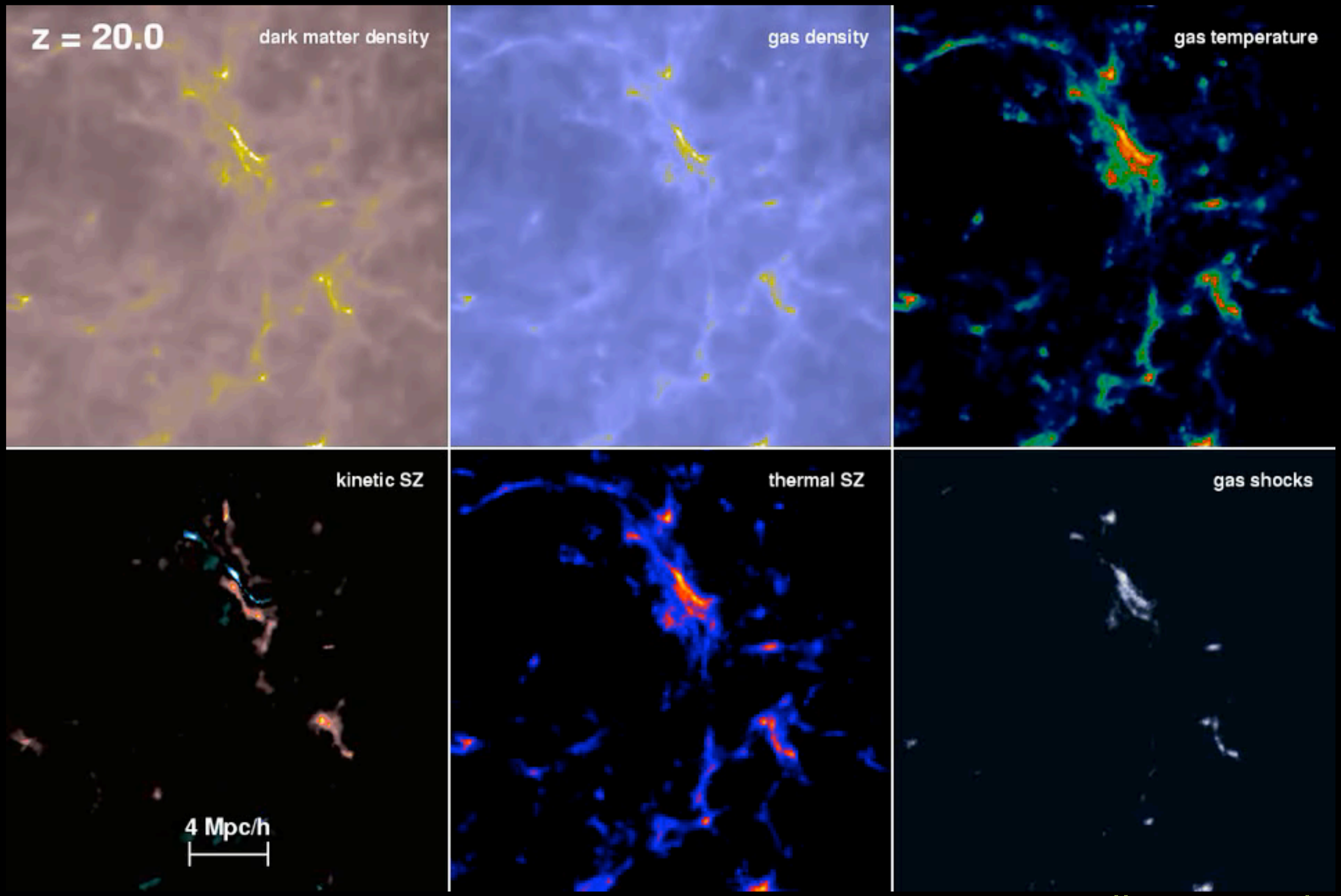


Movies.



Courtesy Volker Springel

Movies.



Courtesy Volker Springel

But how can I do this?

A good place to start is the Matplotlib gallery:

http://matplotlib.org/gallery.html

Another source of ideas are the series of books by Tufte.

http://www.edwardtufte.com/tufte/

More generic visualisations:

http://tylervigen.com/

http://www.informationisbeautiful.net/

http://chartporn.org/

https://eagereyes.org/