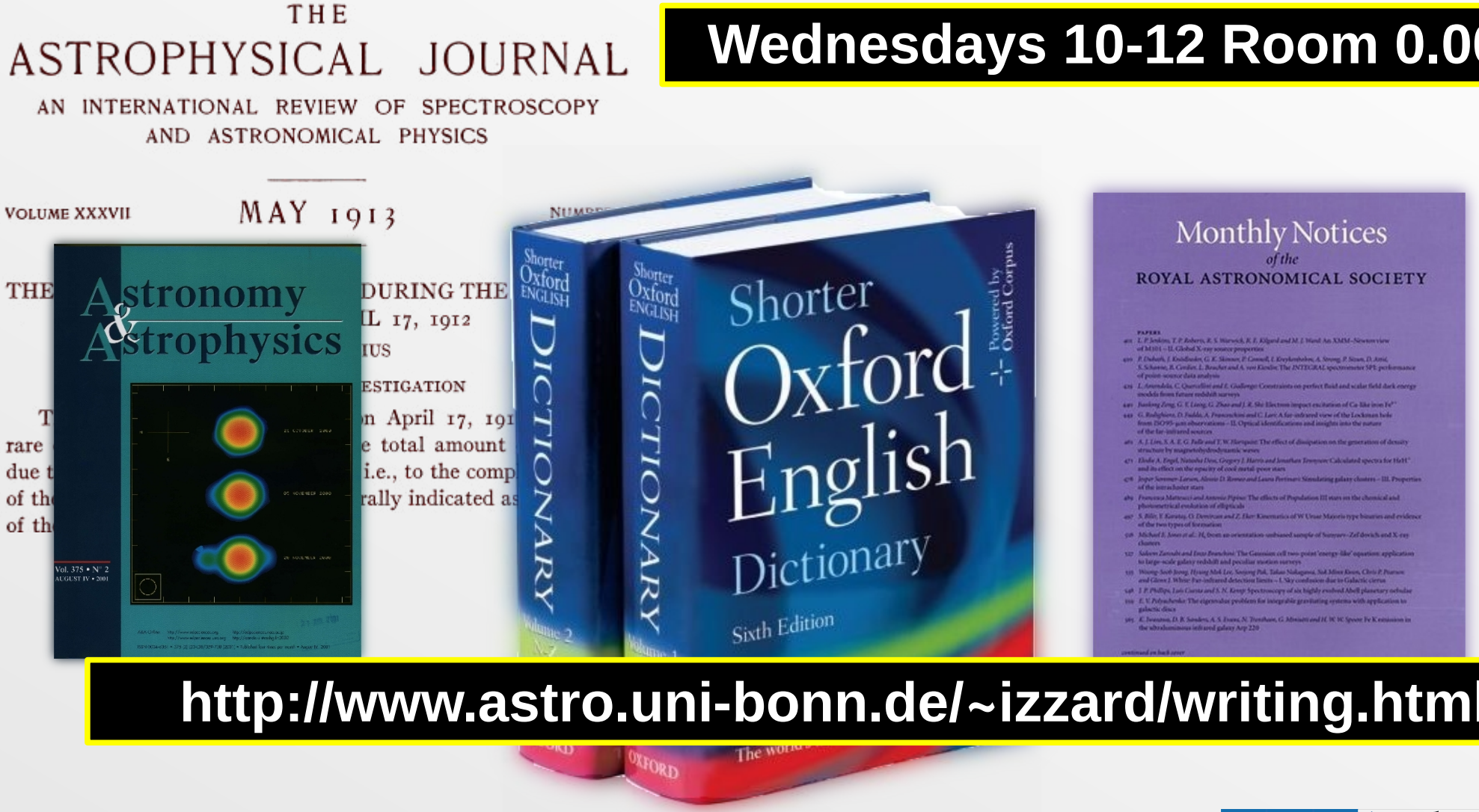
Scientific Writing 6951

Wednesdays 10-12 Room 0.008



http://www.astro.uni-bonn.de/~izzard/writing.html

Argelander Institut für Astronomie

Previously on *Scientific Writing*

- We discussed the different sections in a "traditional" paper:

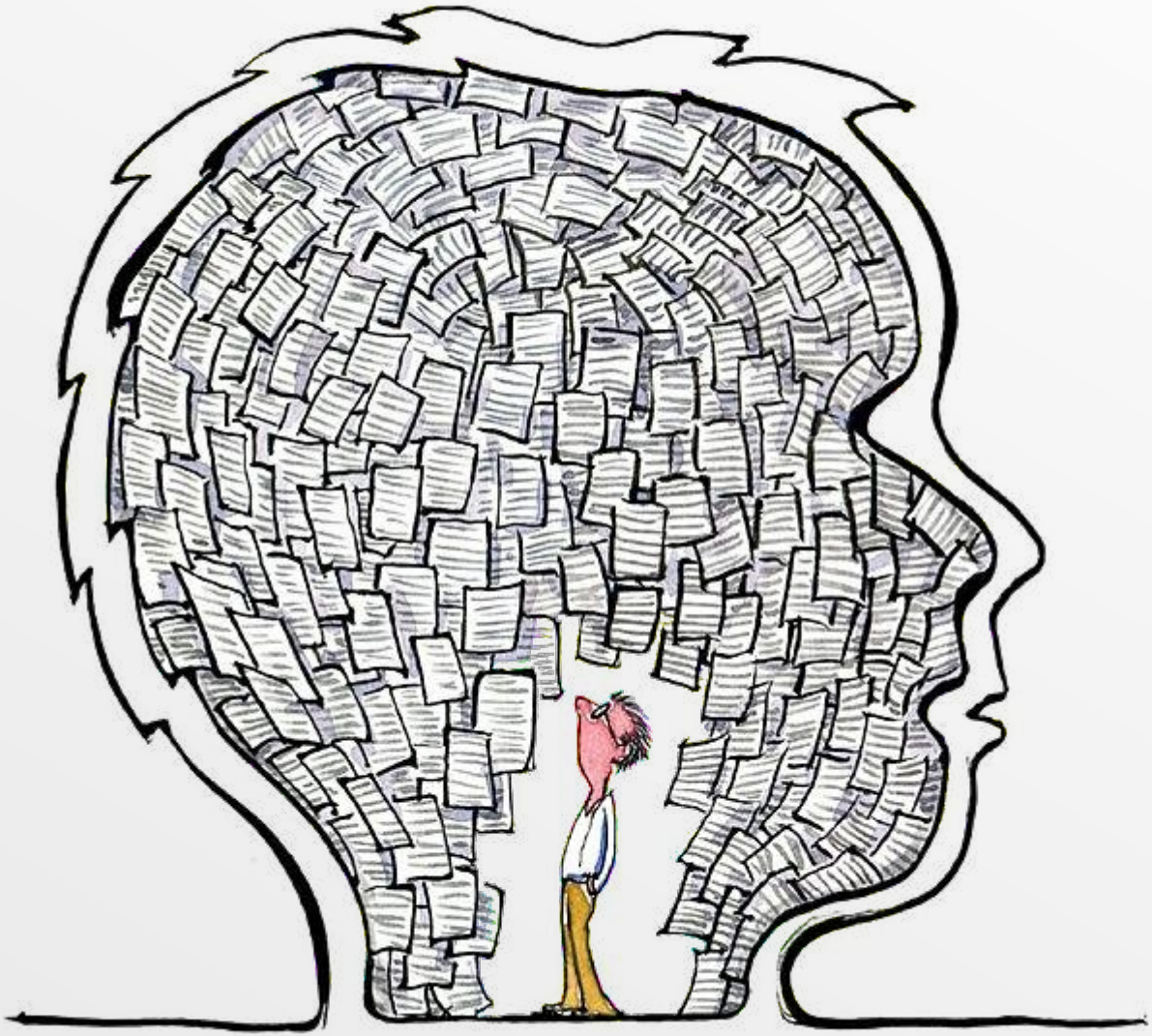
  Intro : Method : Results : Discussion : Conc.

- **Paragraphs**: units of information
- **Sentences**: units of communication
- Logical progression and

  *the **flow** of ideas*

Exercise

- Multiple choice:
  - **Part 2**
- Write flowing sentences
  - **Improve me!**



This week: figures, tables, equations

- ***Why why why?***

  "A picture tells a thousand words"

  Chinese proverb or tram advert?

  or "**Un bon croquis vaut mieux qu'un long discours**"

- Non-linear reading:

  often we look ***only*** **at pictures!**

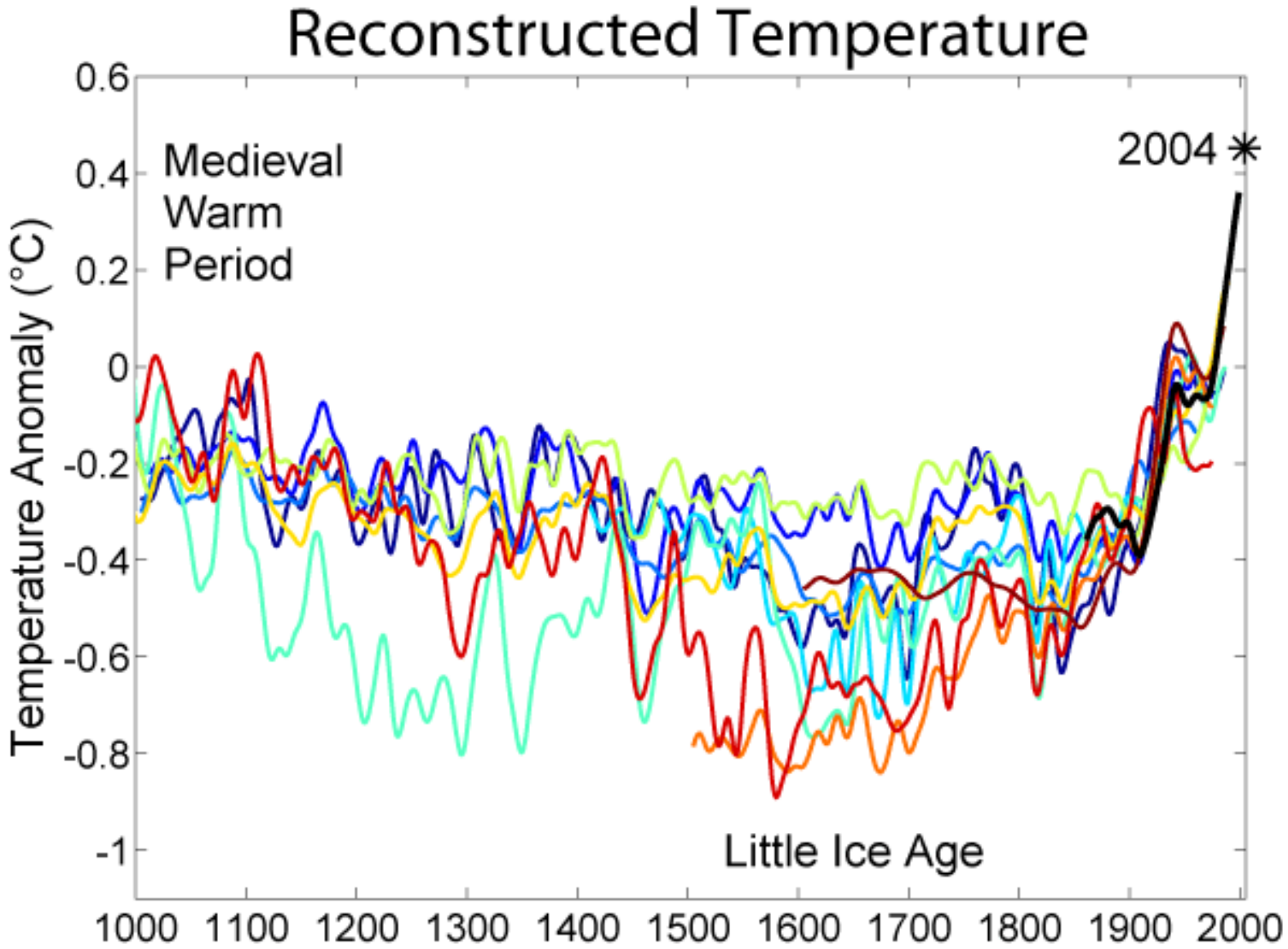
- Talks and posters are (should be)

  75% + pictures!

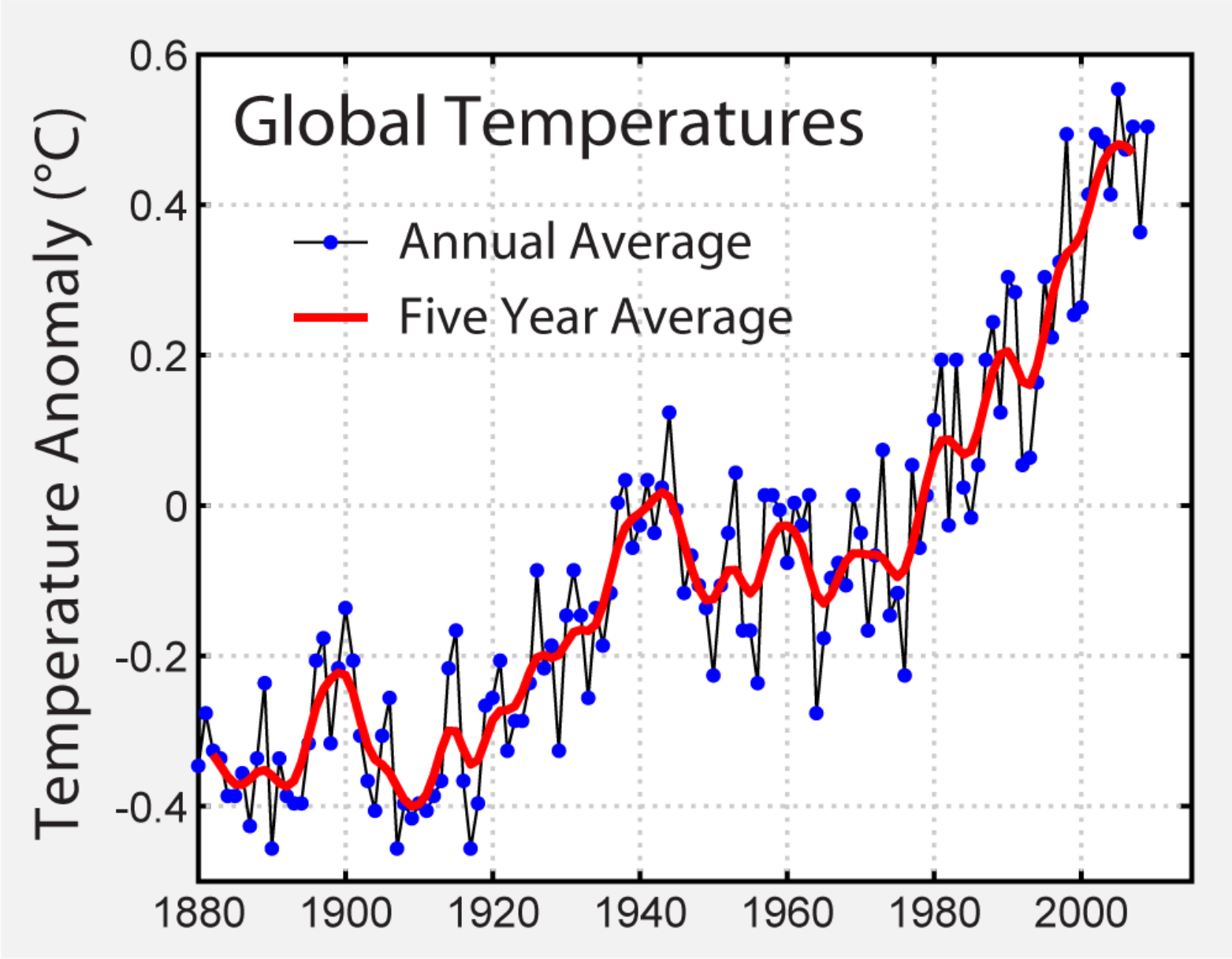
Why use graphics?

- Show **lots of data quickly** and together
- Compare data (e.g. to data, models)
- More "physical" than words
- **~ Language independent**
- Lots of tips in *"Eloquent Science"* and the second *A&A* book on *"Scientific Writing for Young Astronomers"*

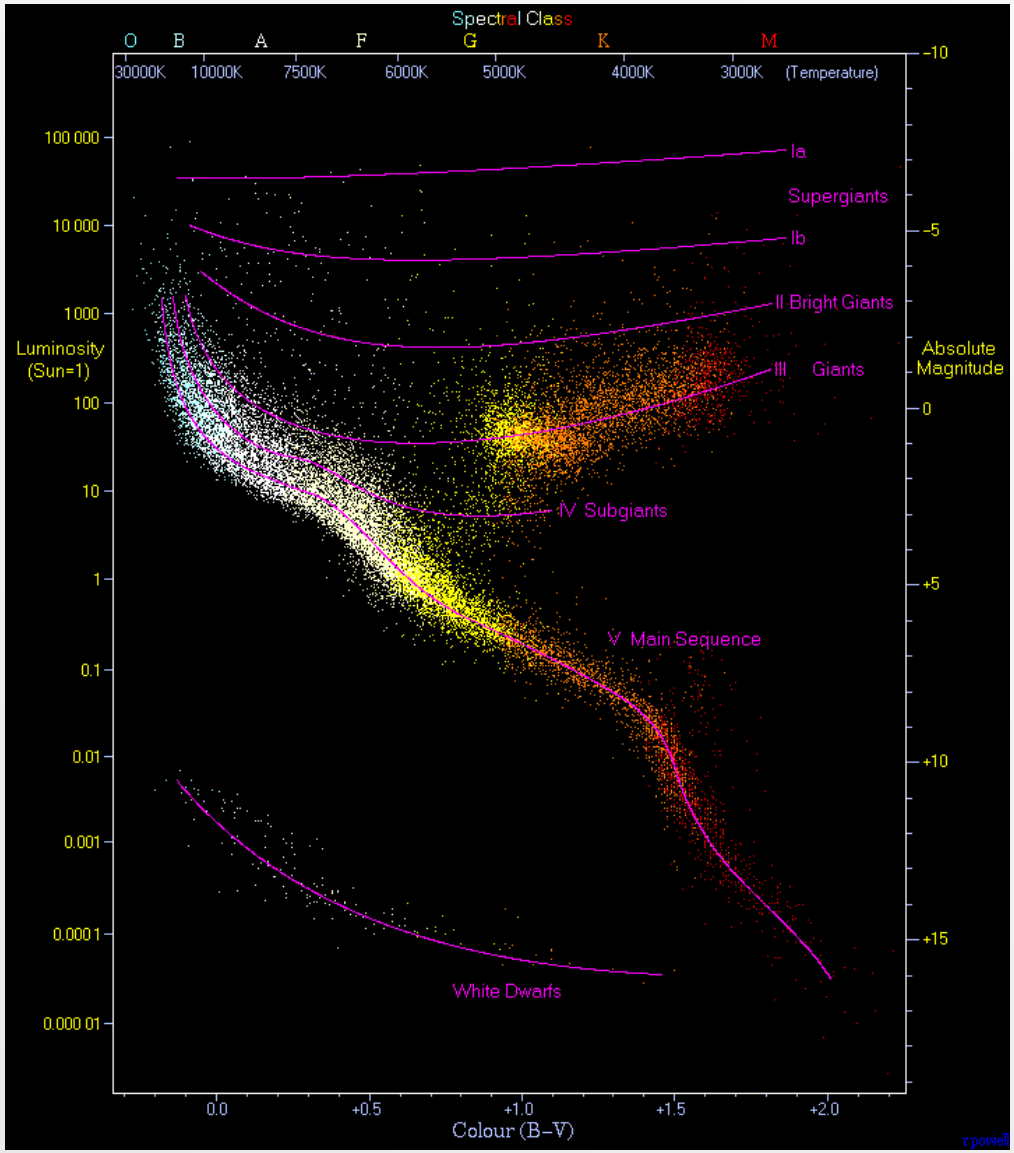
Lots of data

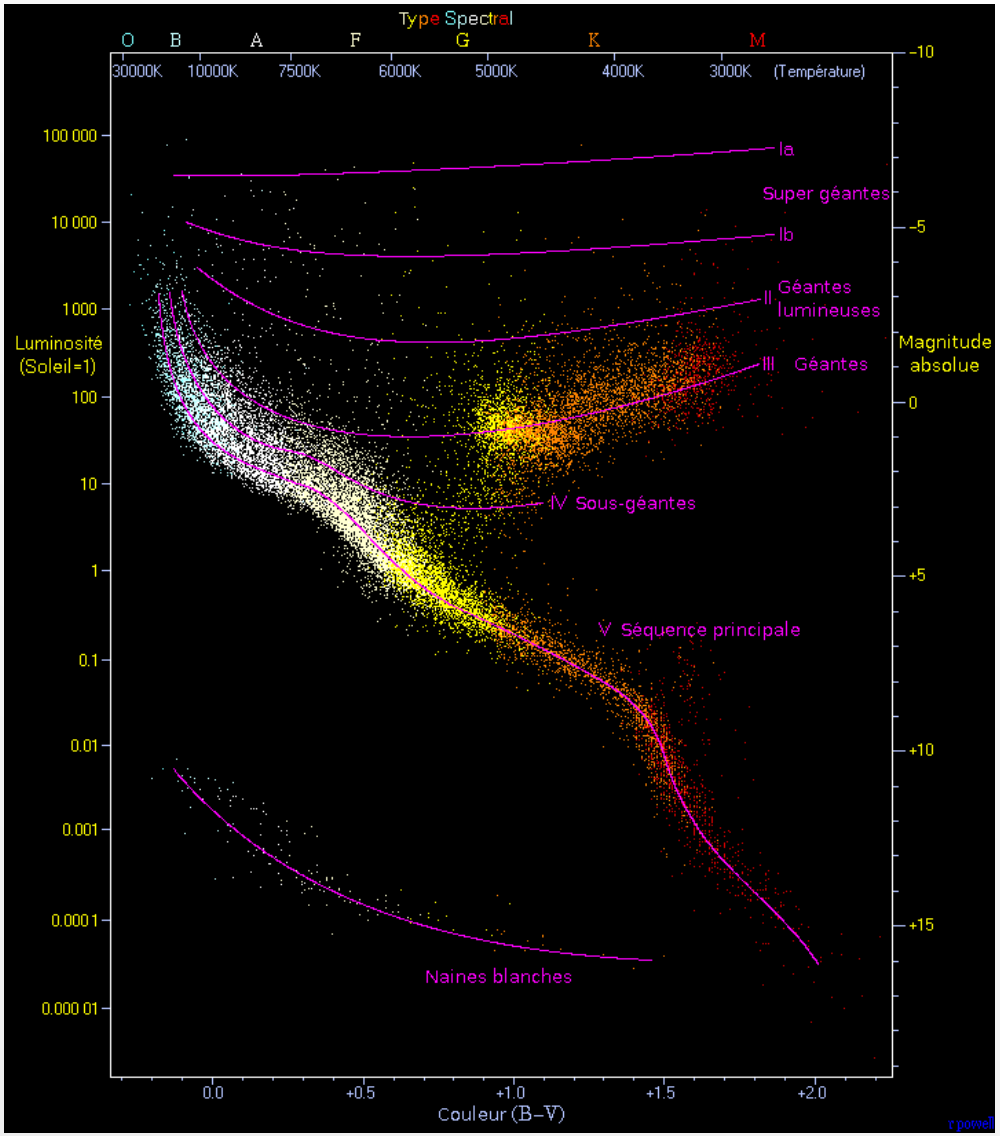

Reconstructed Temperature
Temperature Anomaly (°C)
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
Medieval
Warm
Period
2004
Little Ice Age
1000
1100
1200
1300
1400
1500
1600
1700
1800
1900
2000

Compare data vs data/models



Language independent





http://upload.wikimedia.org/wikipedia/commons/6/6b/HRDiagram.png

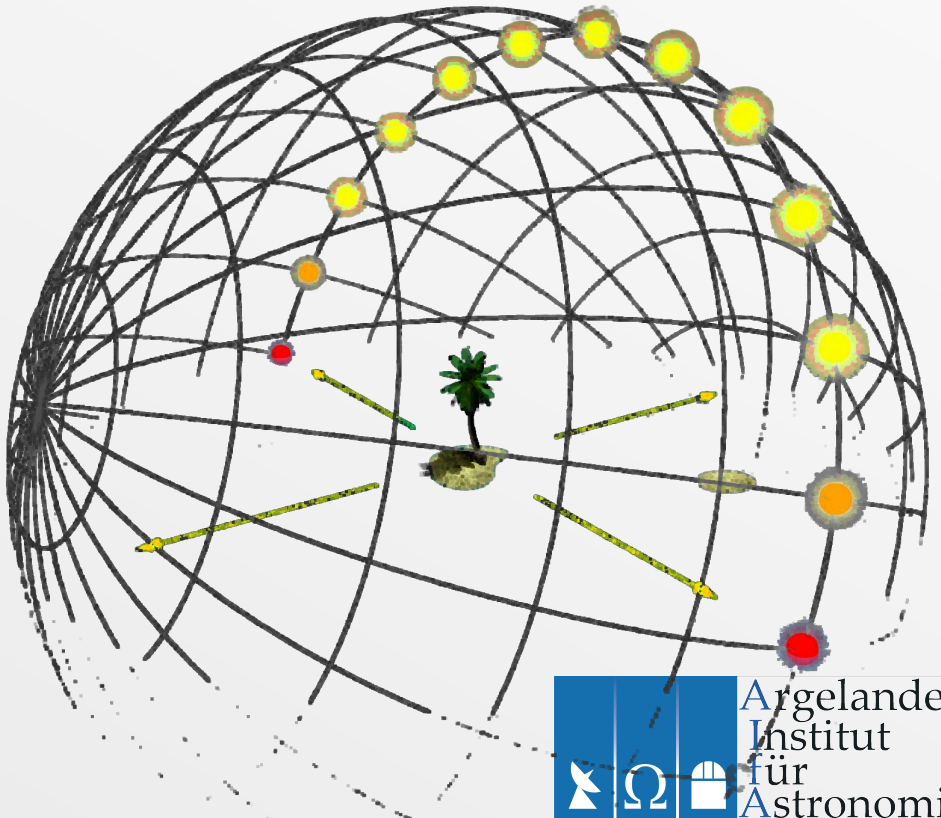
http://upload.wikimedia.org/wikipedia/commons/7/75/HRDiagram-Fr.png

Making figures

- Seems easy ("just put it in Excel")
- ***May take a day to get an image right***

Five key points (see *"Eloquent Science"* p108)

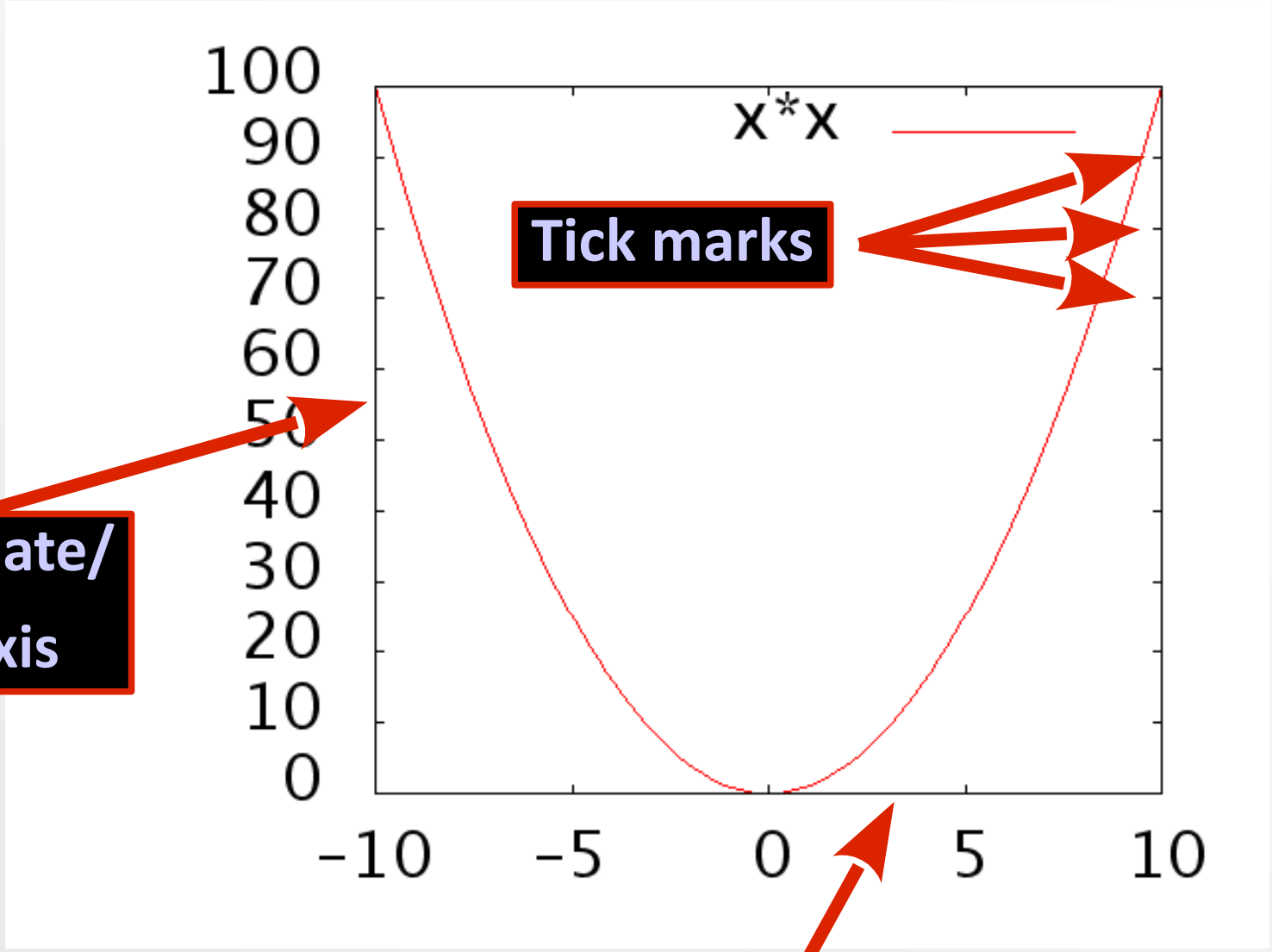
1. **Design**
2. **Size**
3. **Aesthetics**
4. **Consistency**
5. **Annotation**



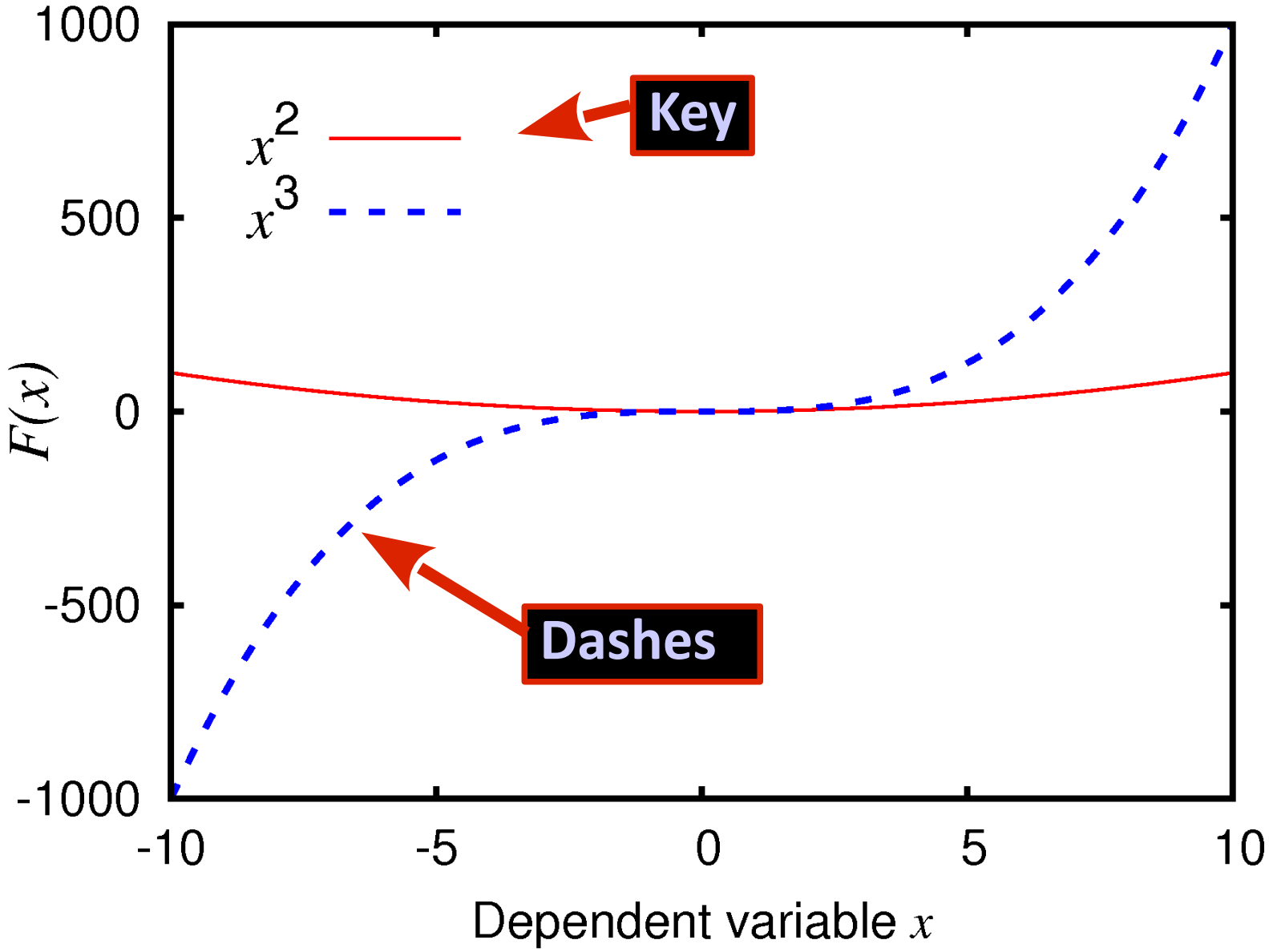
Design

- Many Types of graph:
- **Scatter, Line, Surface, Pie, Bar**
- Choose which you use **carefully**!
- Keep it **simple**, not confusing
- Make the **DATA** stand out, not the borders/axes/grid/labels etc.

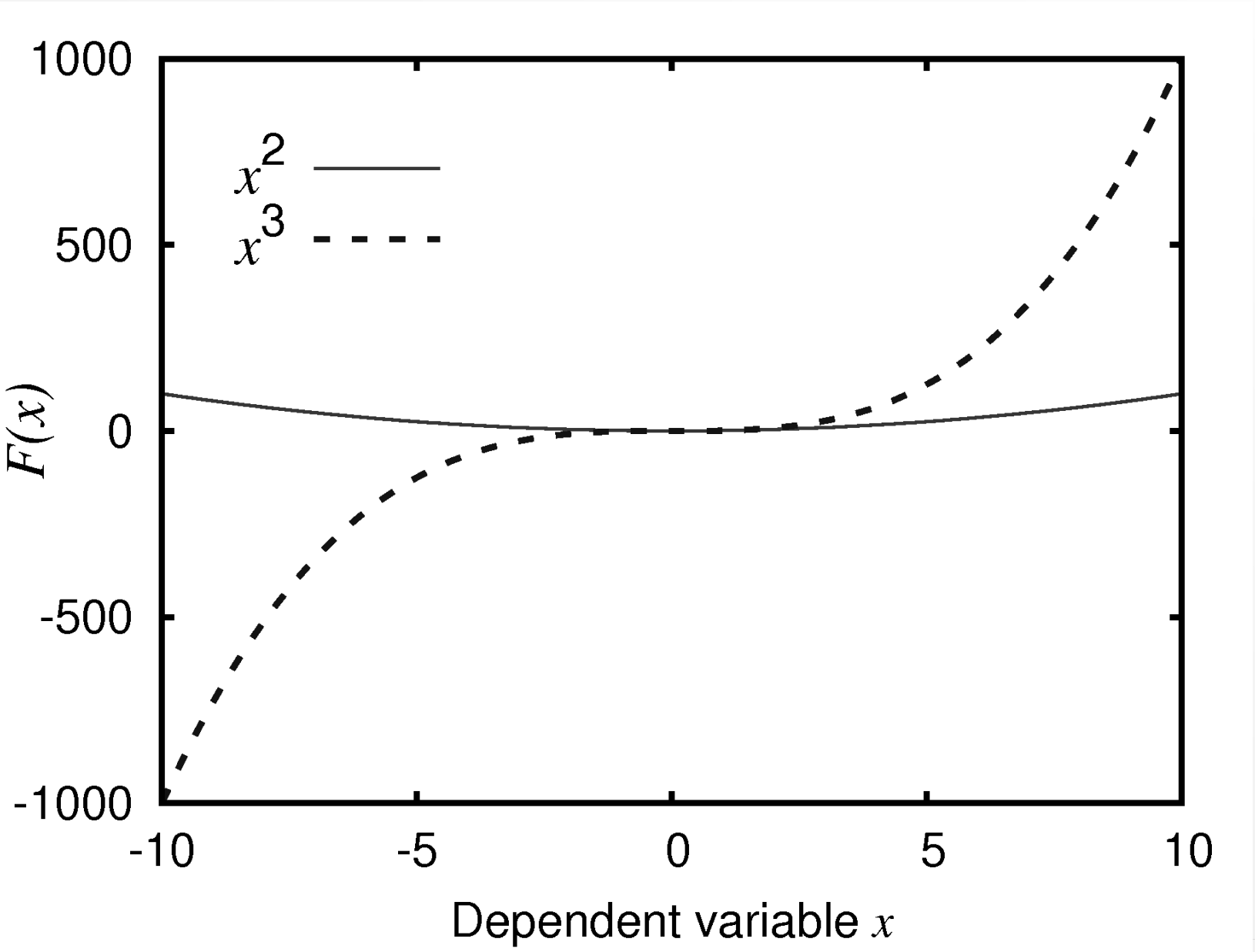
Line graphs


x*x
100
90
80
70
60
40
30
20
10
0
−10
−5
0
5
10
Tick marks
Ordinate/
y-axis
Abscissa / x-axis

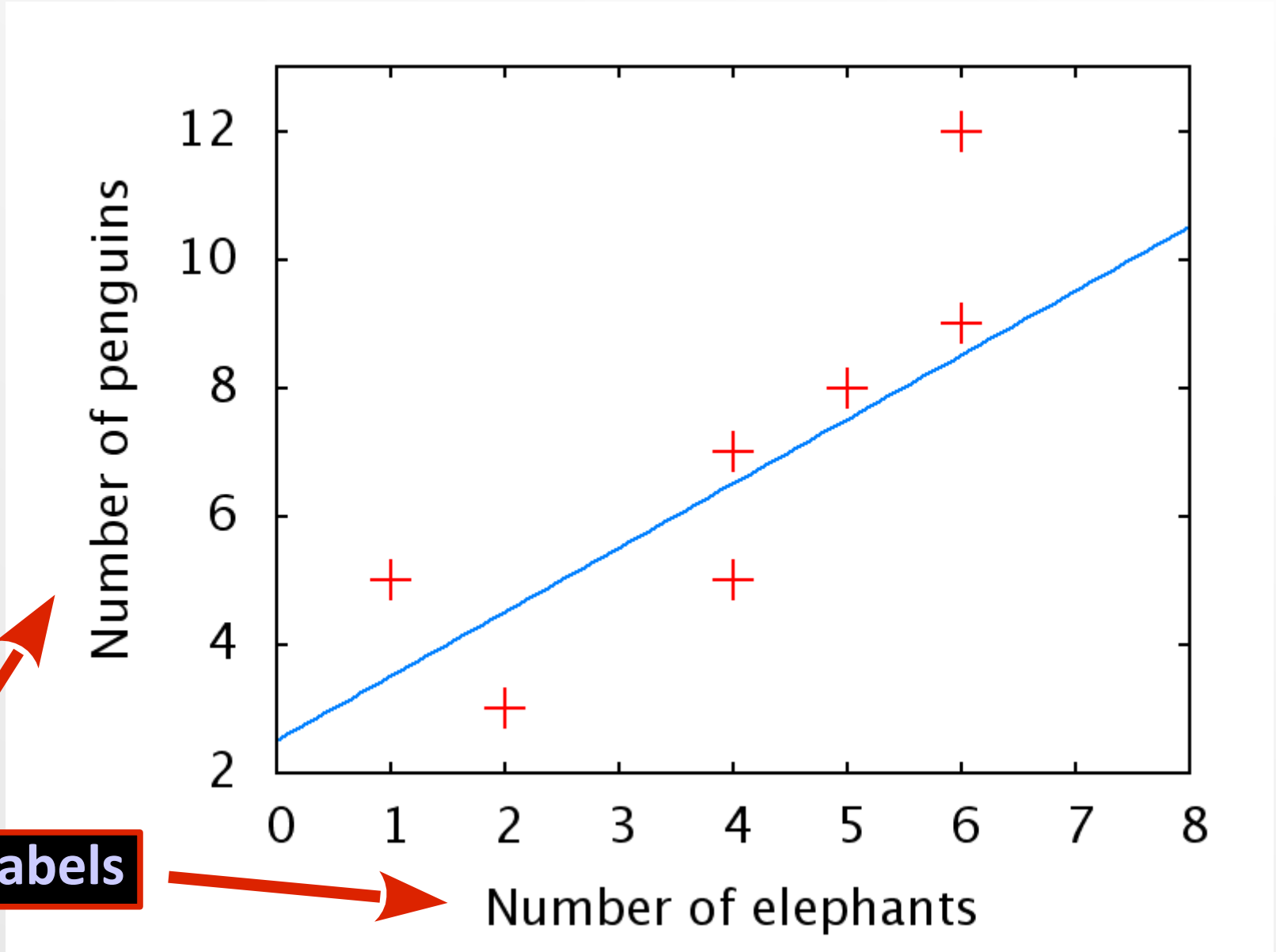
Line graphs


1000
500
0
-500
-1000
-10
-5
0
5
10
F(x)
Dependent variable x
x^2
x^3
Key
Dashes

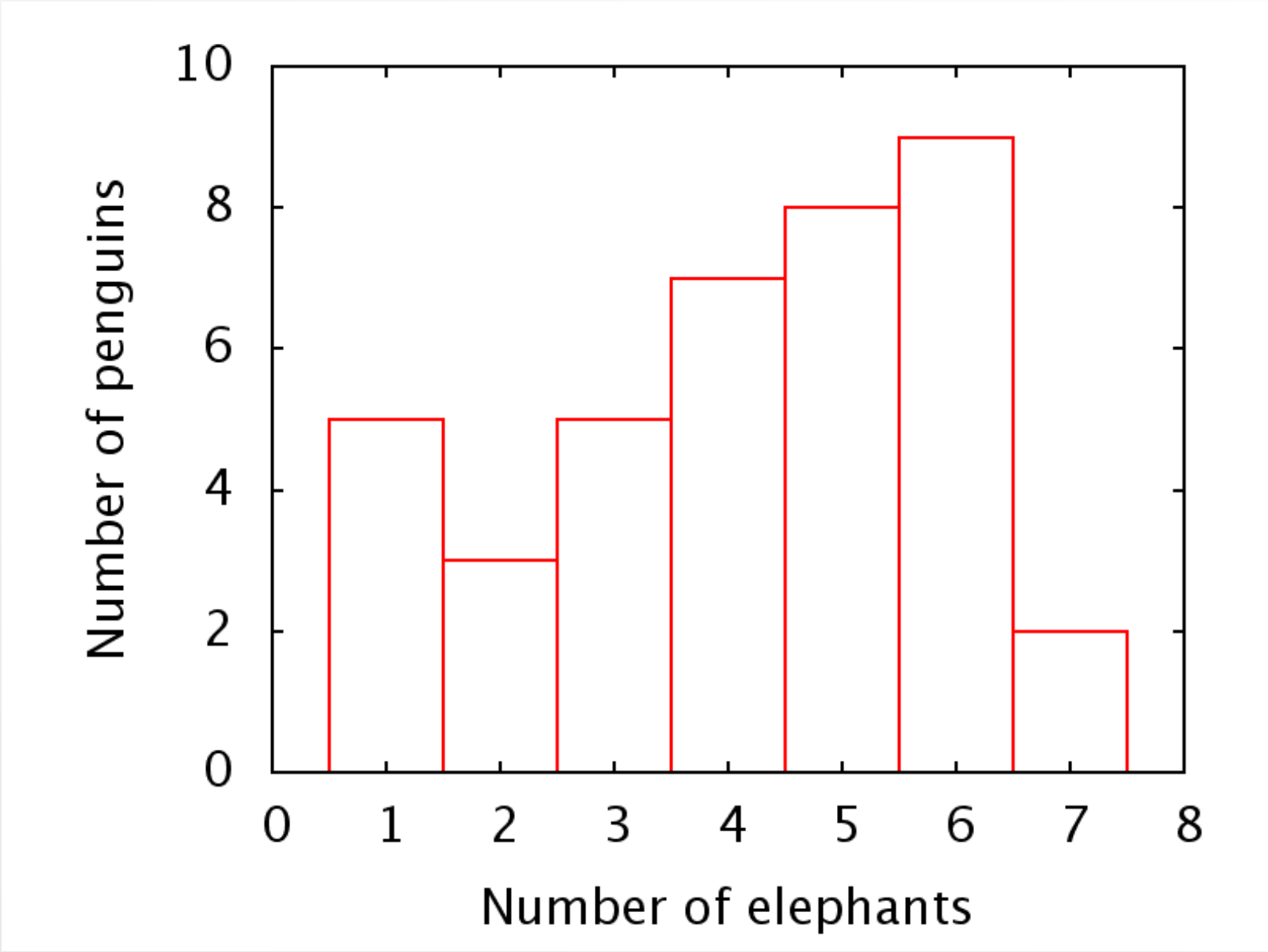
Line graphs



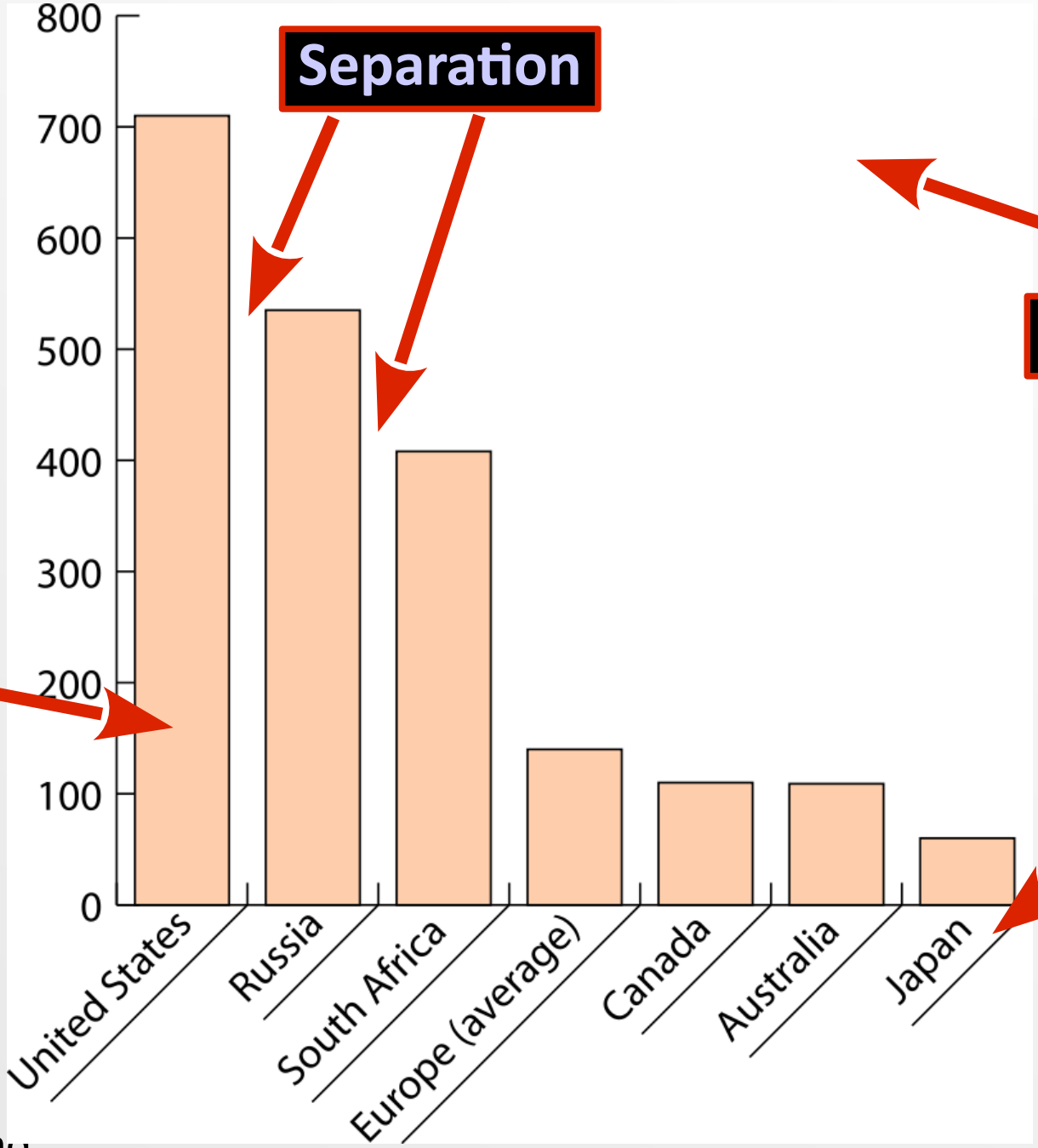
Scatter plot



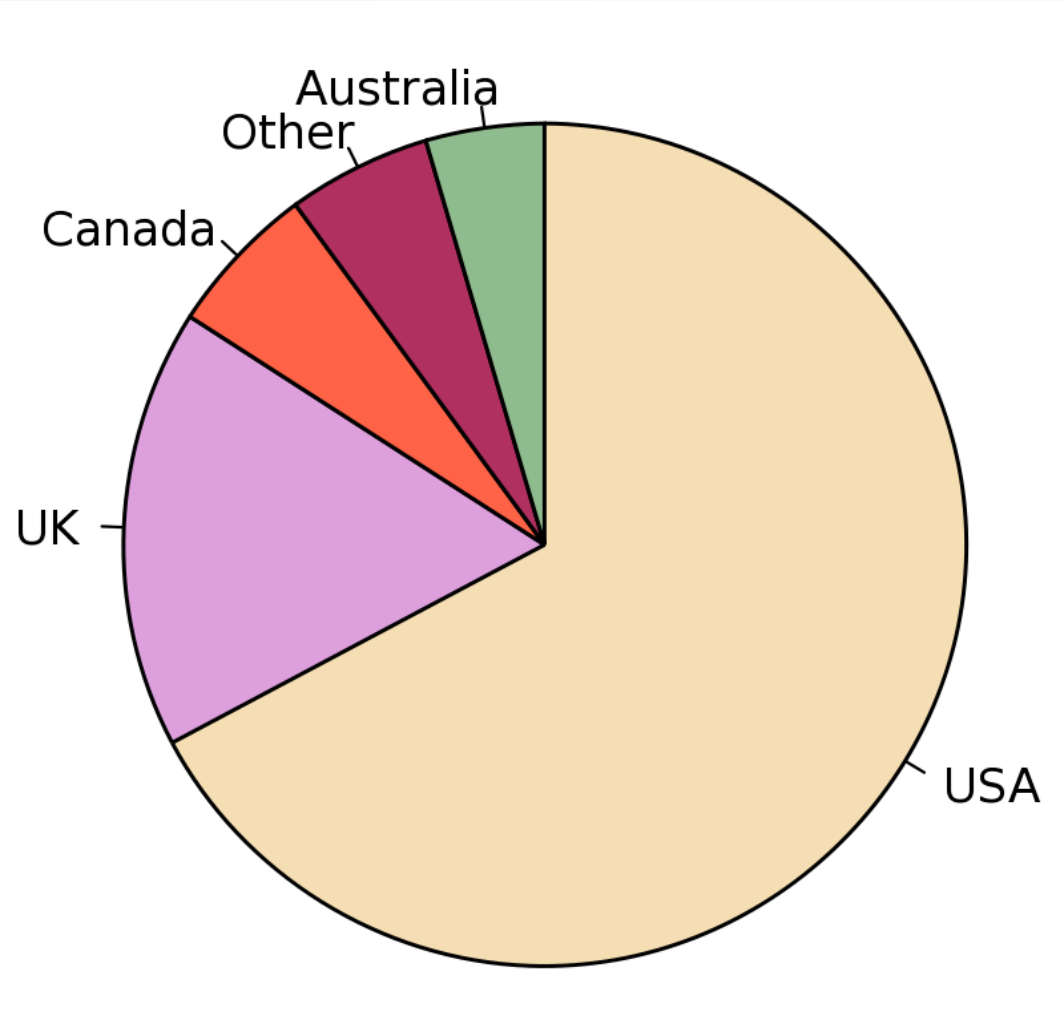
Bar charts



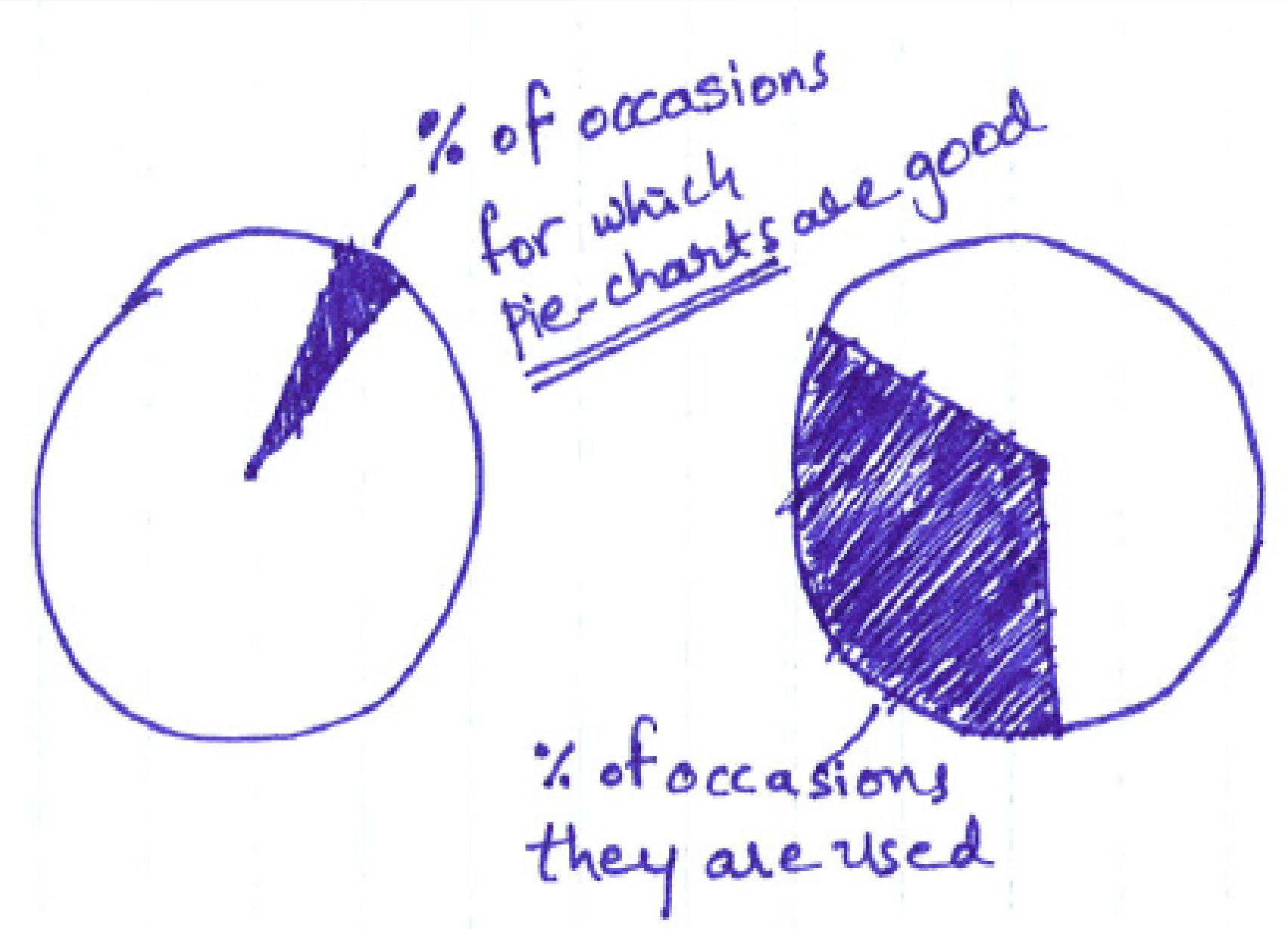
Bar charts


800
700
600
500
400
300
200
100
0
United States
Russia
South Africa
Europe (average)
Canada
Australia
Japan
Separation
Space for key
Fill
Order

Pie charts



Pie charts


% of occasions
for which
pie-charts are good
% of occasions
they are used

Captions

- Explain your figure in a caption
- Not too short, not too long
- A&A *hates* long captions! So...
- Put the information *in the figure:*

  Do not make reading the caption necessary!

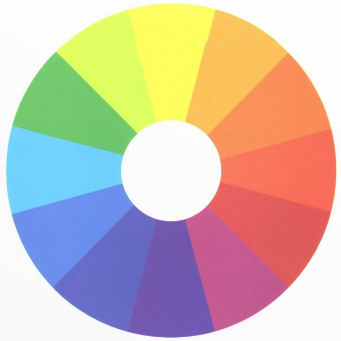
- **NOT**: "This plot shows the velocity of *x*."

  instead "The velocity of *x*."

TALK ABOUT THE (ABSTRACT) DATA

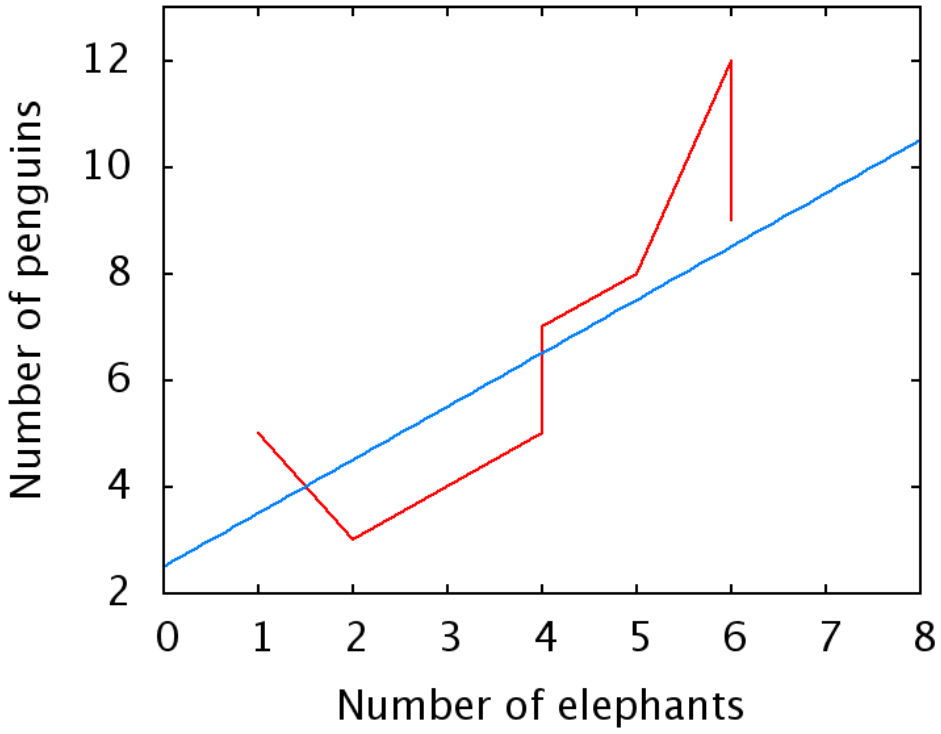
(not the plot itself which is just a load of dots and lines!)

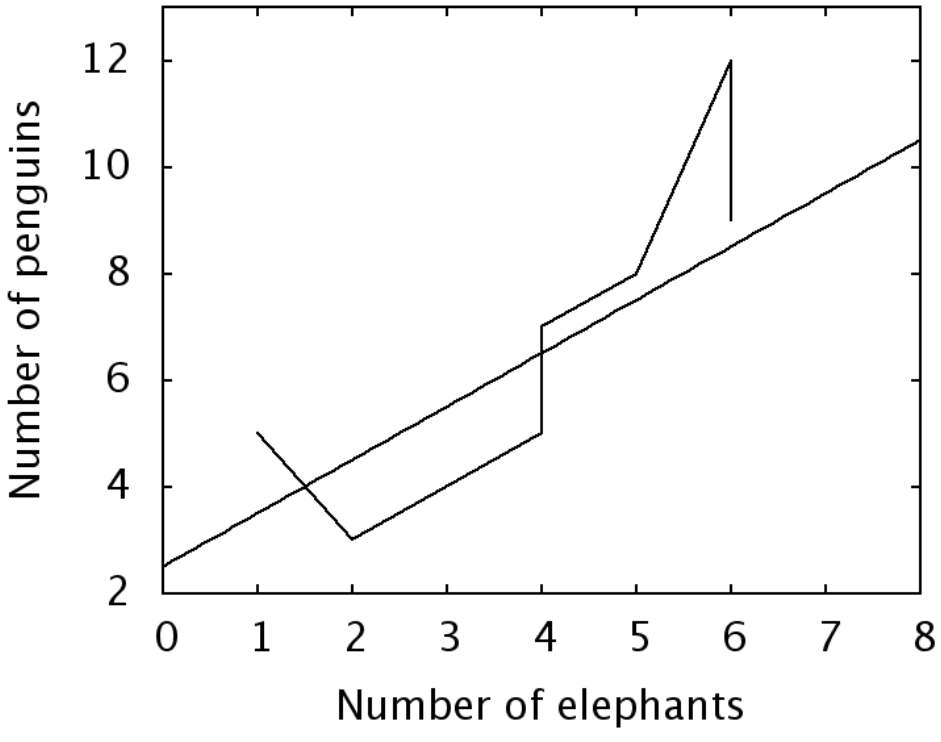
Colours



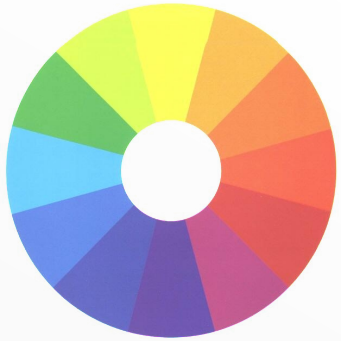
Colours help but **beware**:

your graphics should also look good in **black/white/greyscale**



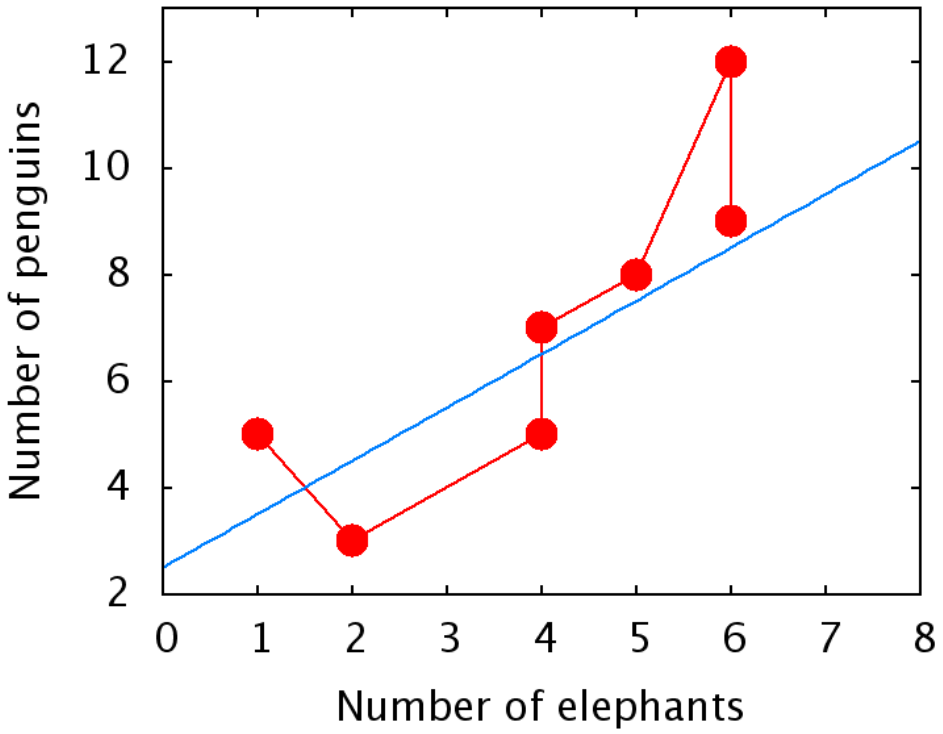


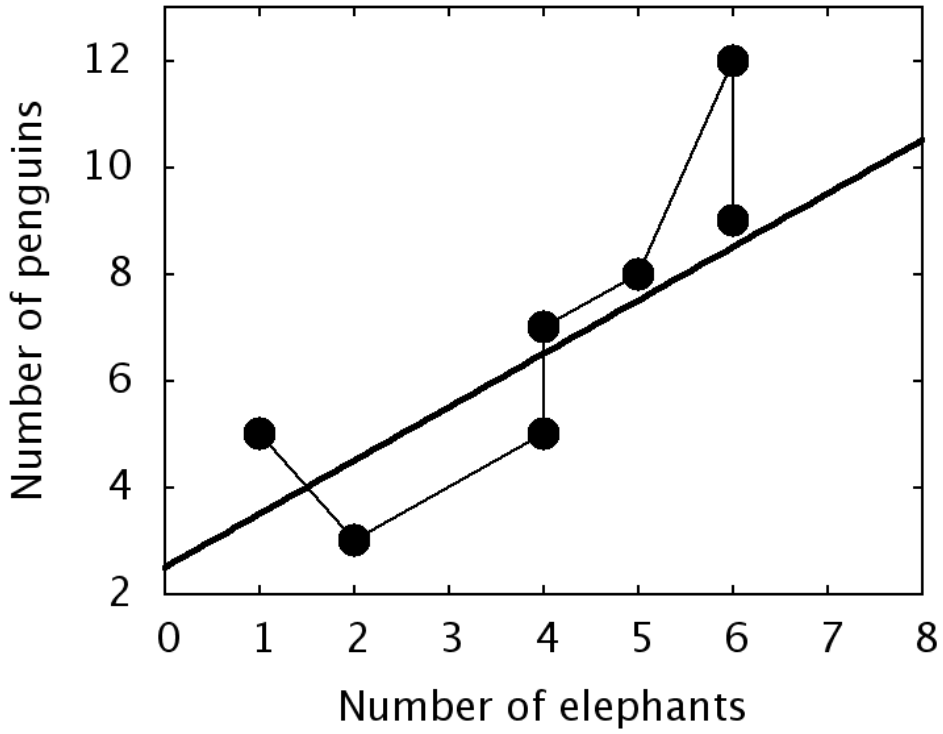
Colours



Colours help but **beware**:

your graphics should also look good in **black/white/greyscale**



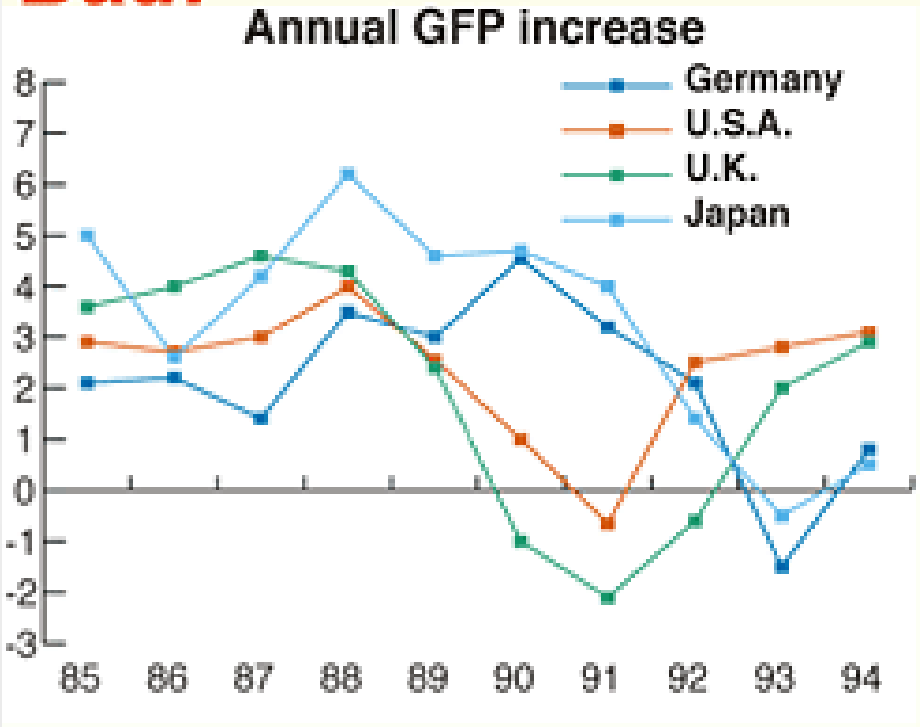


Colours

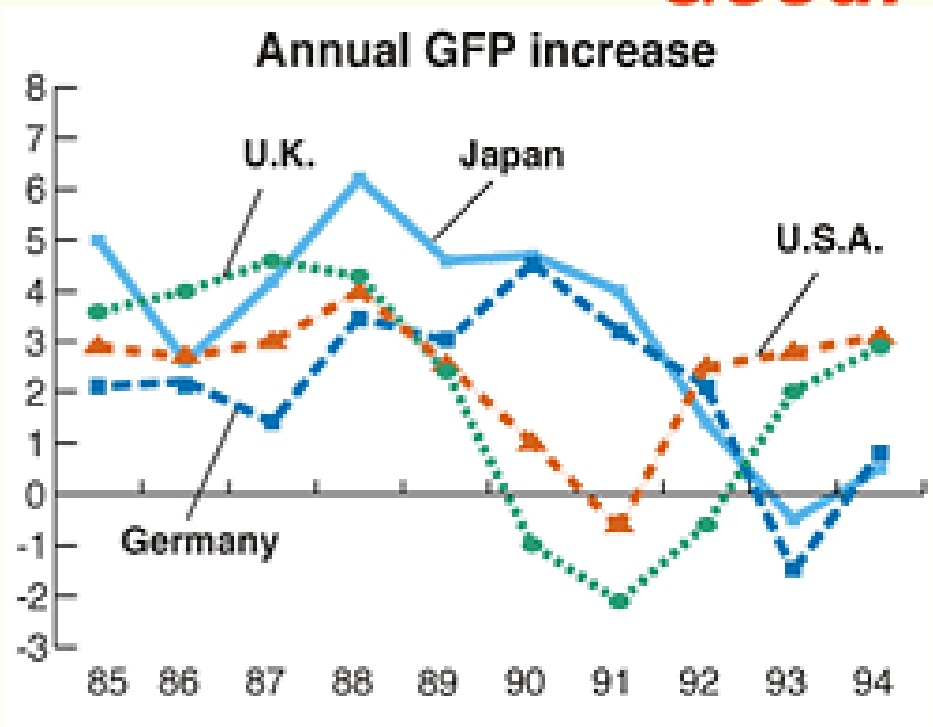
http://jfly.iam.u-tokyo.ac.jp/color/#assign

“Redundant Coding”

Bad!



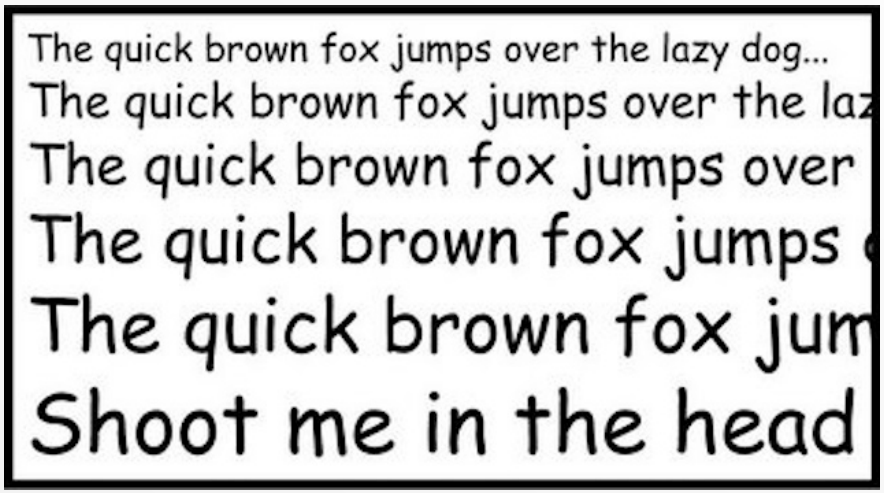
Good!



Do and do not

- Axis labels : **always always always!**
  - Graph title? If necessary.
- Font: Sans serif, size size size size

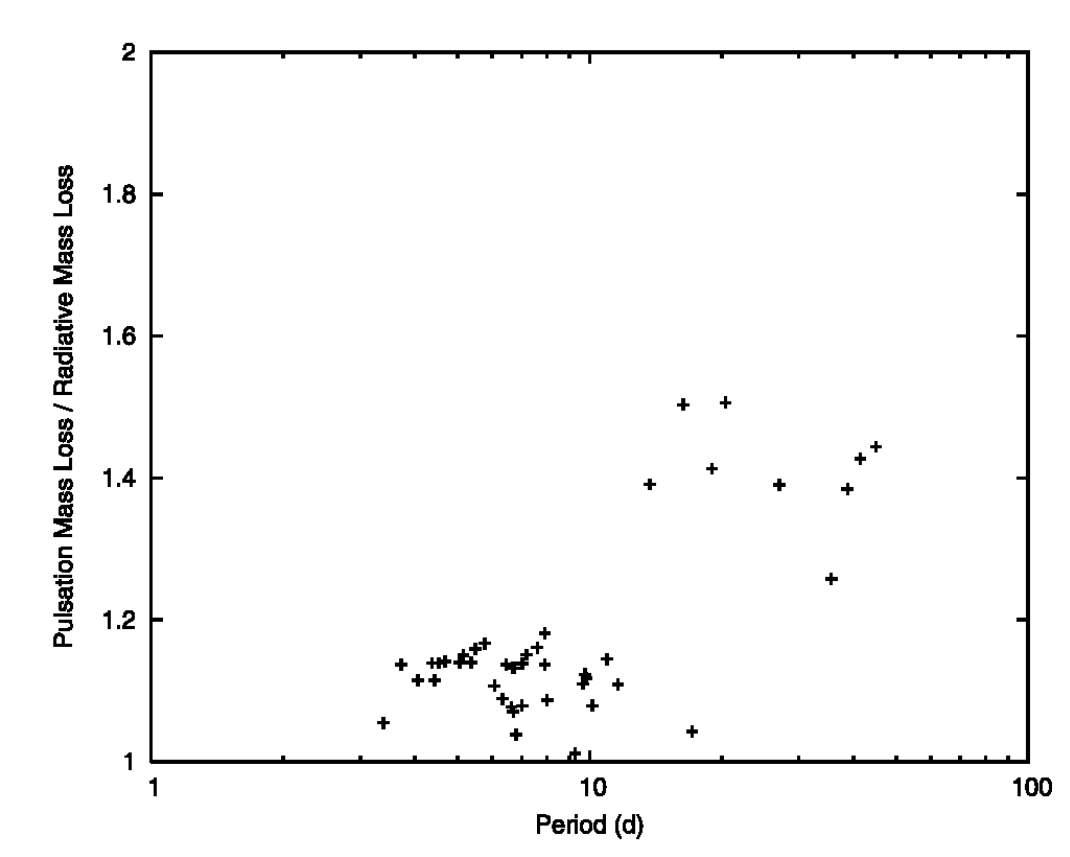
Not comic sans!



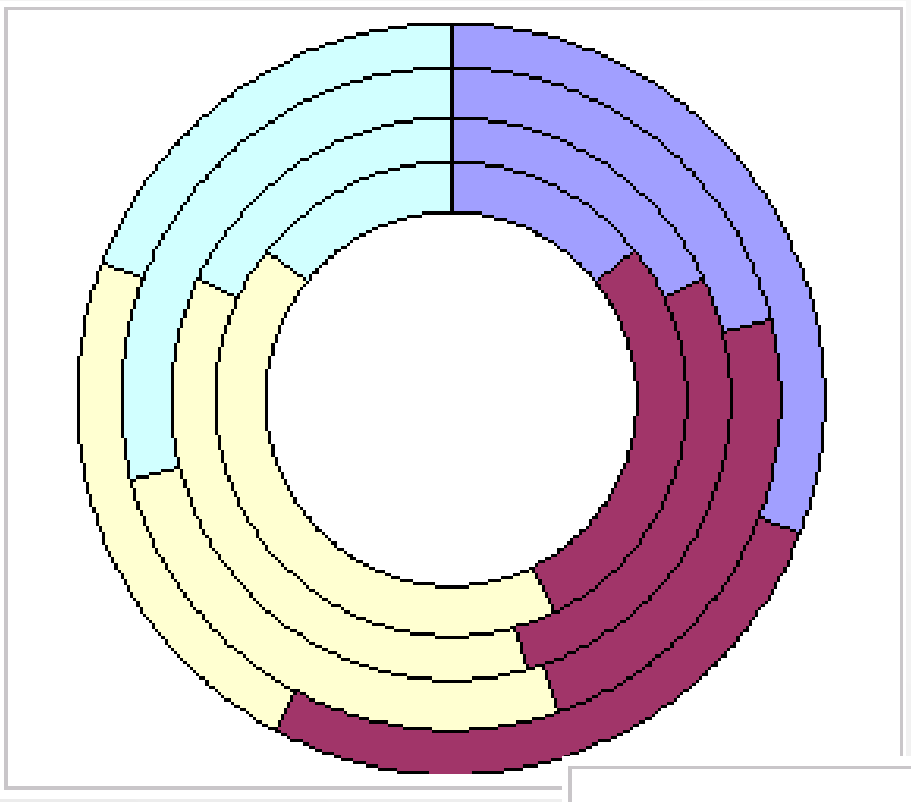
- Line width : **beware rescaling!**
- Colours: aim for **contrast**
- ***The data*** should stand out
- Error bars! Data without errors is *useless*
- Be ***consistent*** throughout the paper!

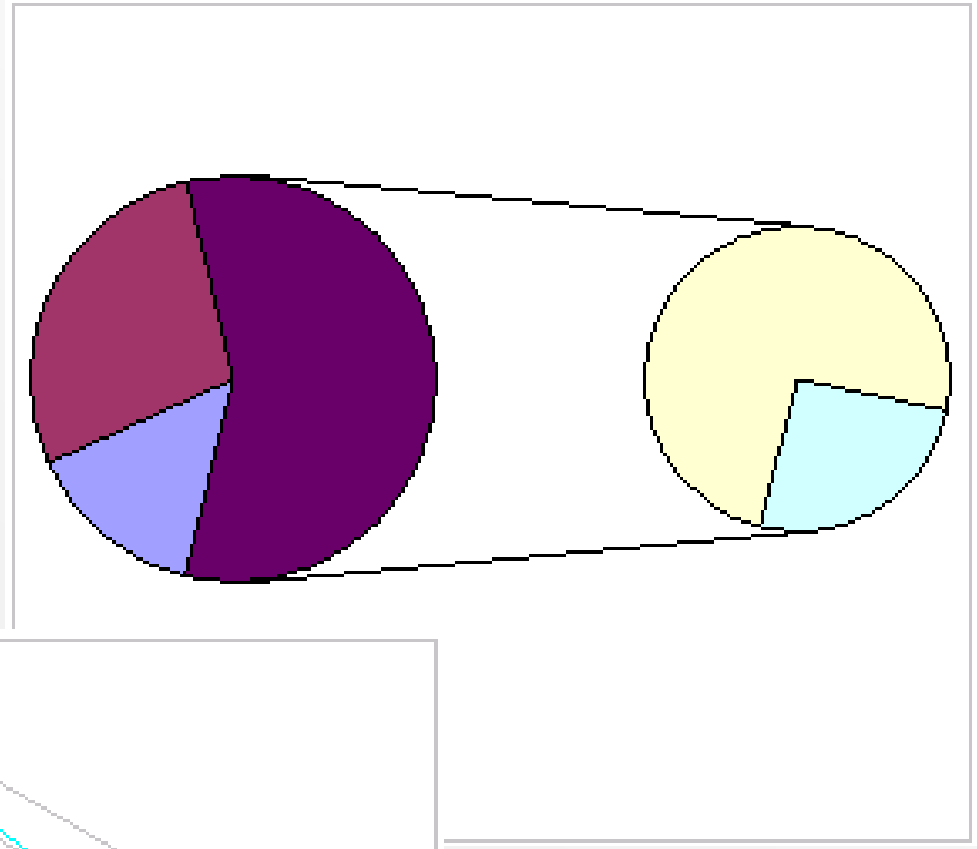
Size of the plot

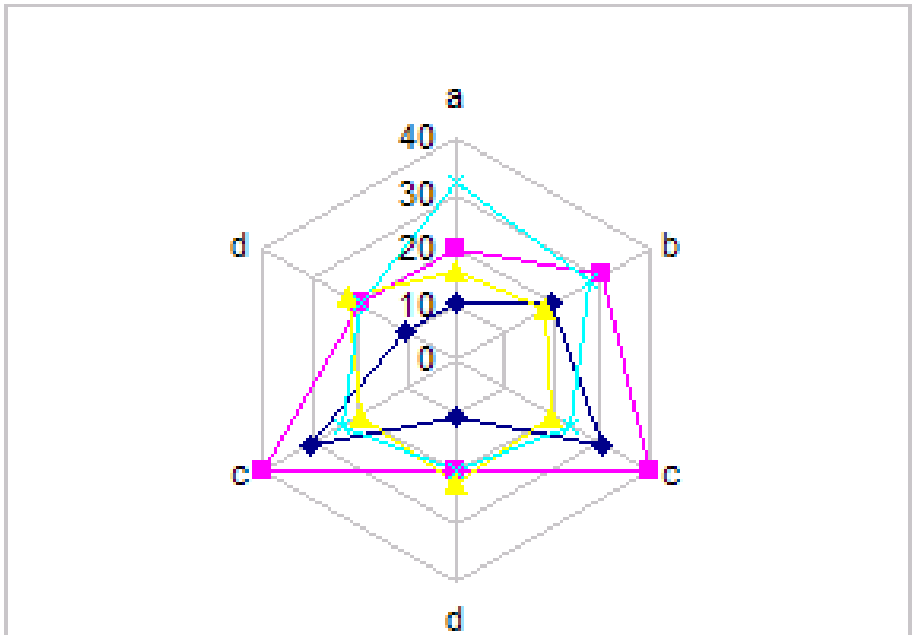
- Single or two-column?
- Do not crowd the plot!
- Panels: label (a), (b) etc. ***inside*** the plot
- Use **all the space**!



Bad examples

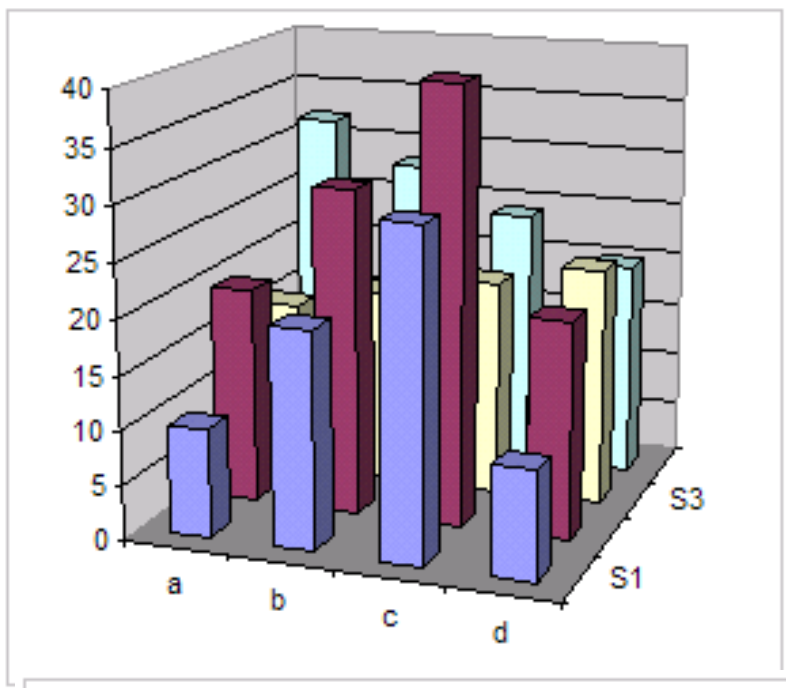


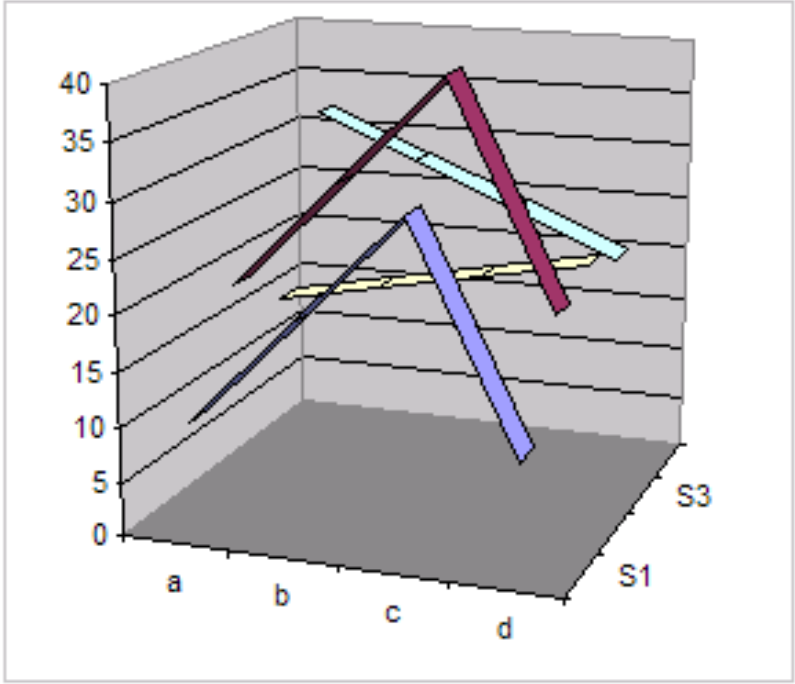


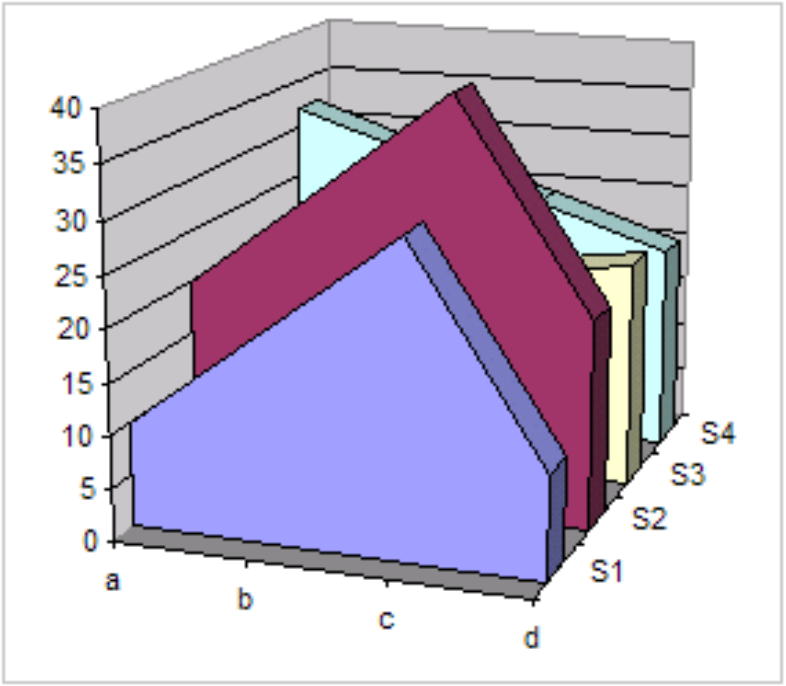


I hate Excel

Bad examples





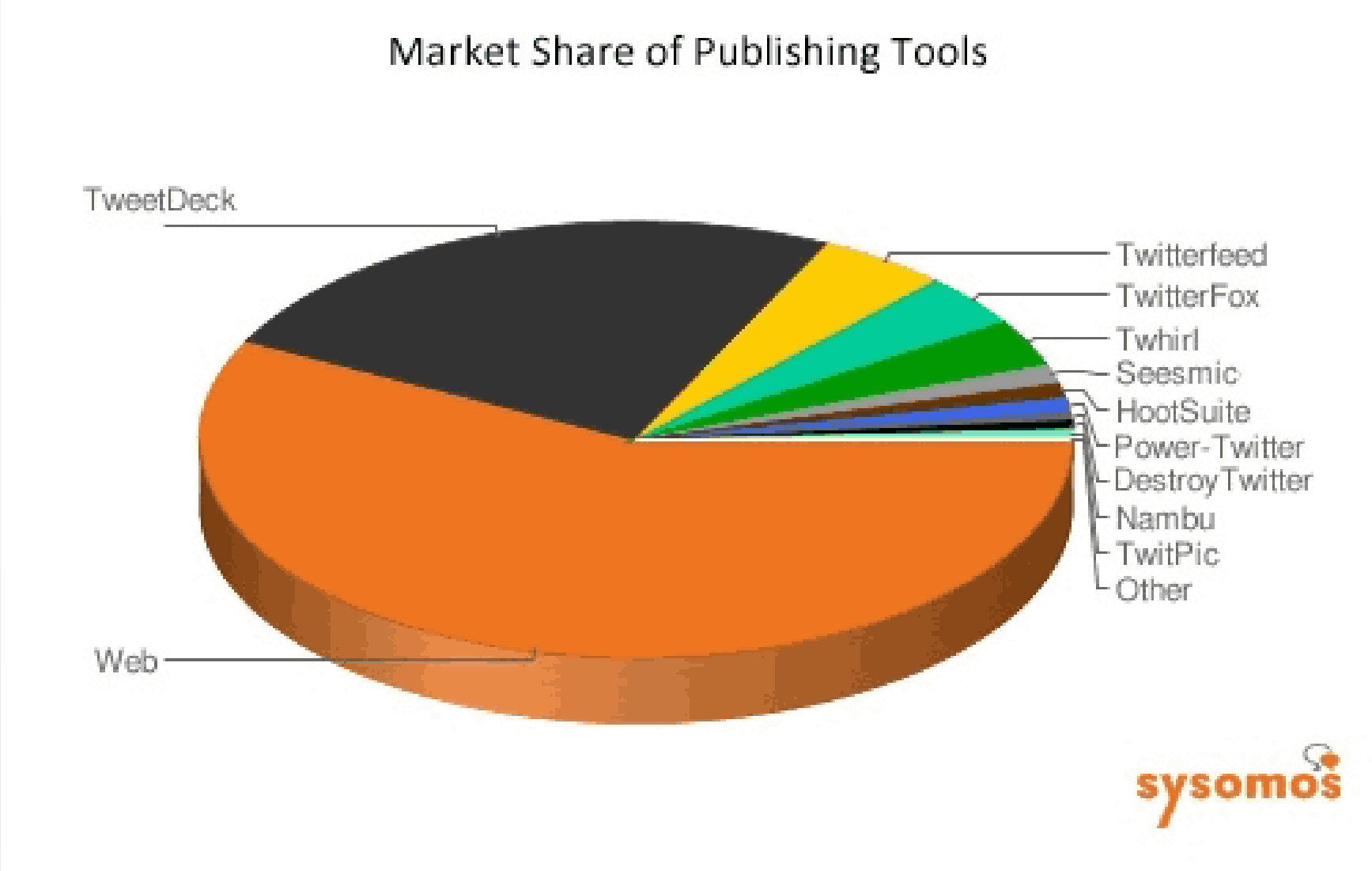


I *really* hate Excel

Bad examples

Which **orange** dot is larger?

Bad examples



So it's mostly on the web then...

Tables

- Shows **precise** information in a list
- More **concise** than textual explanation
- **Emphasise** points from the text
- Is a table ***really*** necessary?
- Order rows/columns
- Consistent (parallel) headings
- Many small tables better than one large
- Beware too many horizontal/vertical lines
- **Appendix ?**

Equations

- Equations are ***part of the flow of text*** !
- Short equations in the text itself
- Use text to **describe** what the eq. is for
- What do the **symbols mean**?
- **Scalars**: Italic $c_s \quad v \quad K \quad \xi$
- **Operators**: Roman $\log(x) \quad \exp(y) \quad \sin(z)$
- **Units**: sometimes Roman $M_\odot \text{ vs } \mathrm{M}_\odot$

$$\mathrm{km\,s^{-1}}$$

Equations

The speed of sound c_s is calculated from,

$$c_s^2 = \frac{\partial p}{\partial \rho}\,, \quad (1)$$

where ρ is the density and p is the pressure.

NOT: where ρ = density and p= pressure.

References:

- The speed of sound is a function of pressure and density (Eq. 1).

NOT:

- Equation (1) gives the sound speed.

Numbers

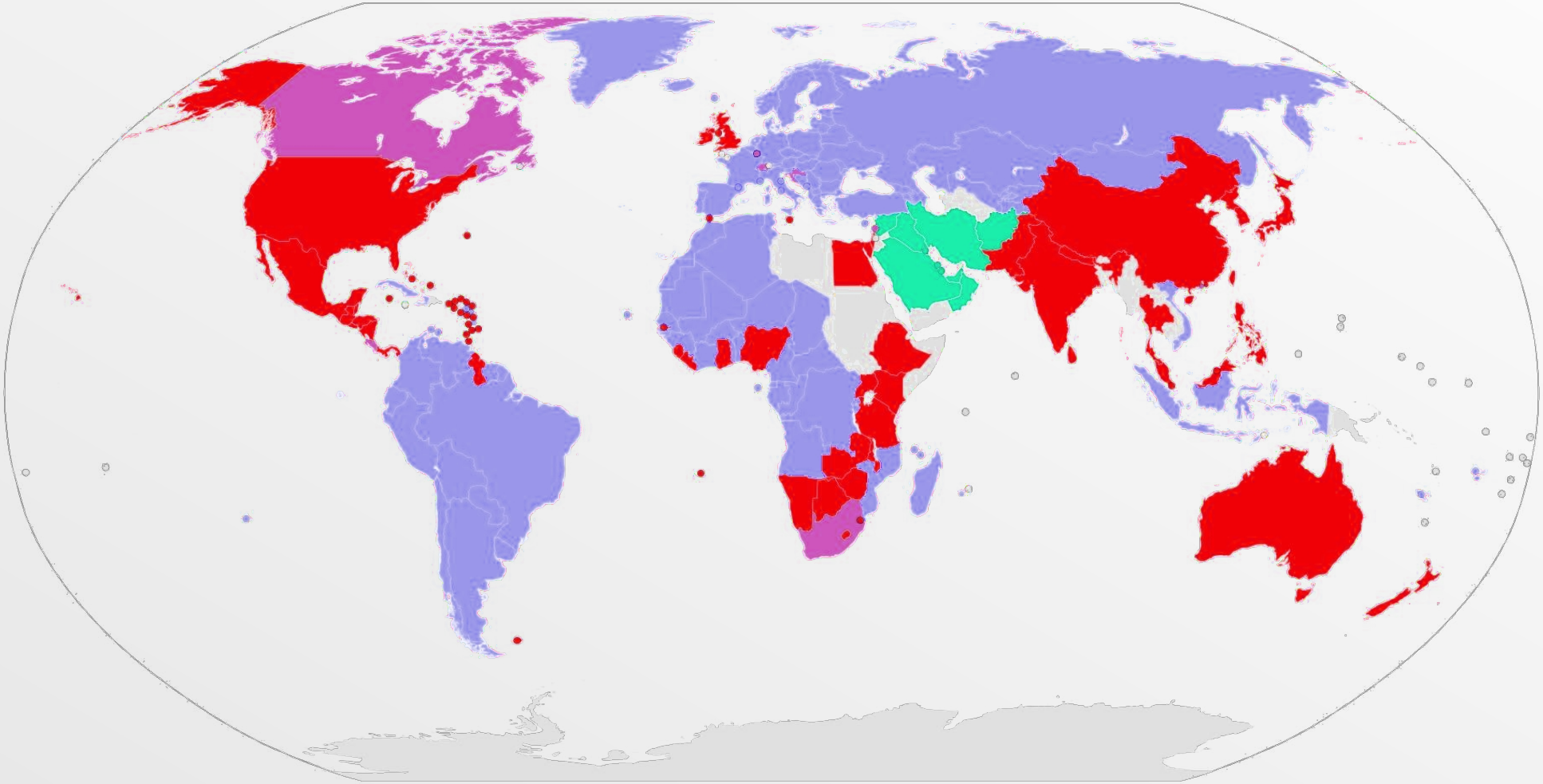
- Small numbers should be written in full

  I found **ten** pints of beer in the fridge.

- Exact ***results*** should be in numbers, with **units**!

$$\text{We find that } c_s = 10\,\mathrm{km\,s^{-1}}.$$

- Beware , (comma) and . (decimal point)
  - In **English**, 10.123 is a bit more than **ten**
  - And 10,123 is ten **thousand**, one hundred and twenty three!



Source: wikipedia

Referencing Fig./Tab./Eq.

- If using **LaTeX** : use **\label** and **\ref**

Do not ever manually number figures / tables / equations!

- Fig. (1) Eq. (2) Table 4

  Brackets () or not? **See your journal's style guide.**

- **All** figs/table/eqs should have numbers!

  (otherwise, why are they there?)

- **Sequential**! i.e. in the **order in which they appear**

- **Appendix** e.g.: See equation (A.2), table (B.3)

Improve me

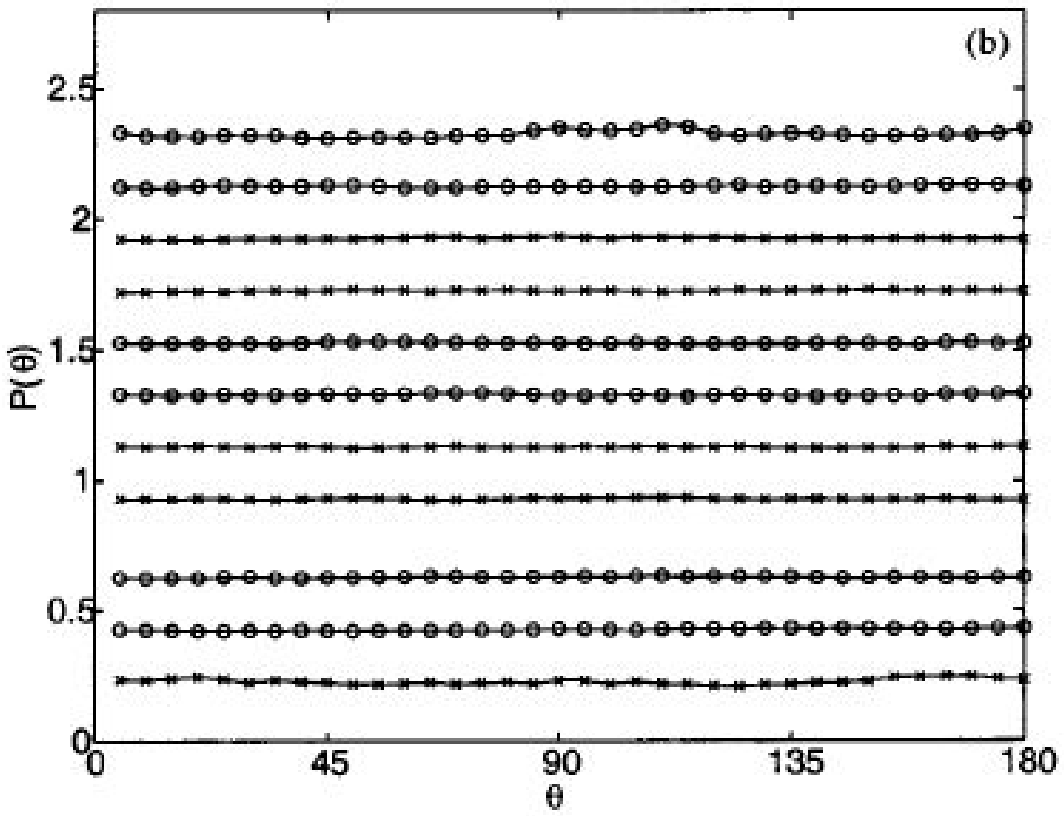
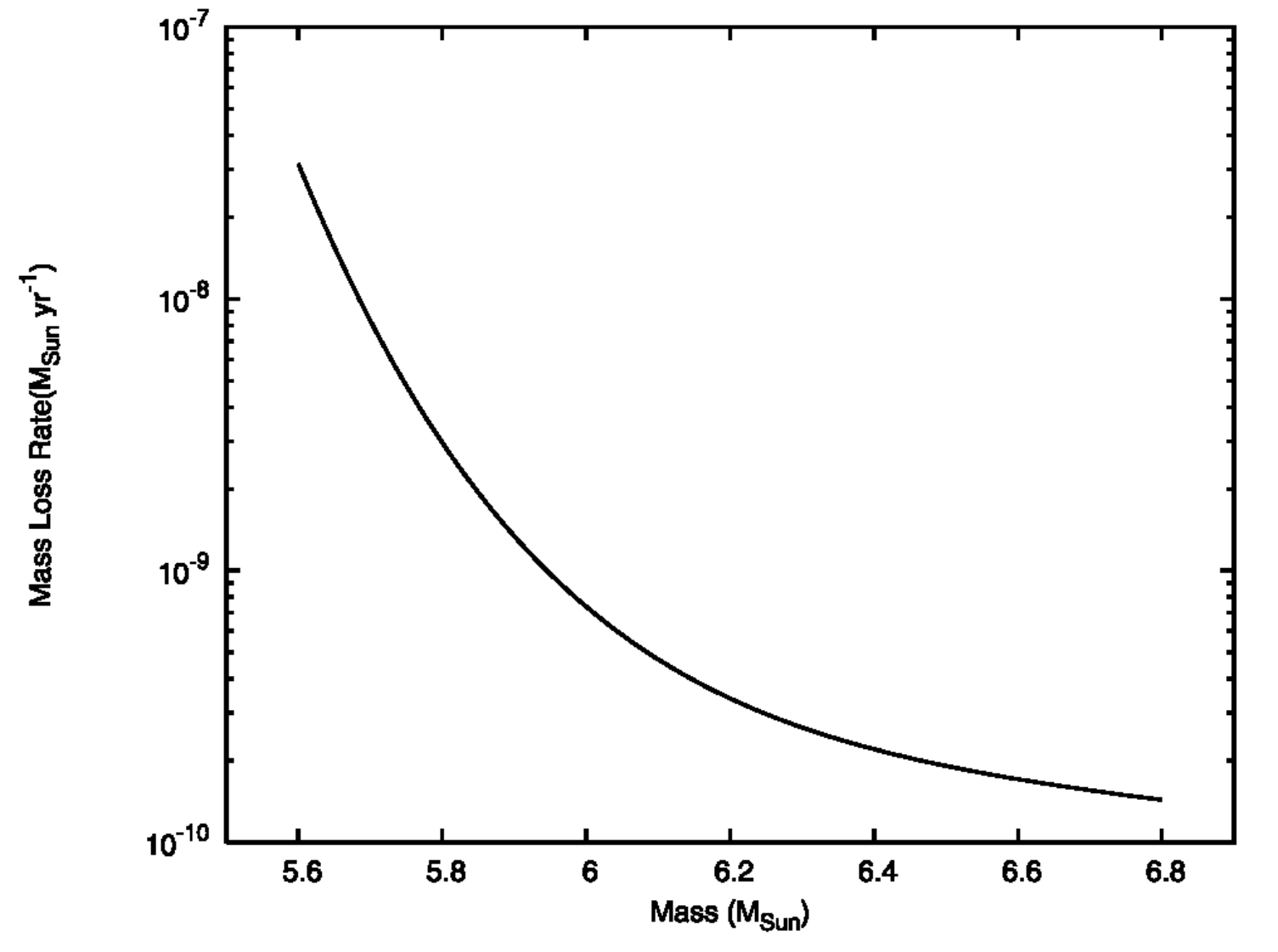


Figure 8. Probability distribution as a function of position in the film for angle made by the largest ellipsoidal axis of the chain with (a) the *x*-axis (ϕ) and (b) the *y*-axis (θ). The successive probability distributions are vertically offset by 0.1 for clarity.

Improve me


Mass Loss Rate(M_{Sun} yr^{-1})
10^{-7}
10^{-8}
10^{-9}
10^{-10}
5.6
5.8
6
6.2
6.4
6.6
6.8
Mass (M_{Sun})

Improve me

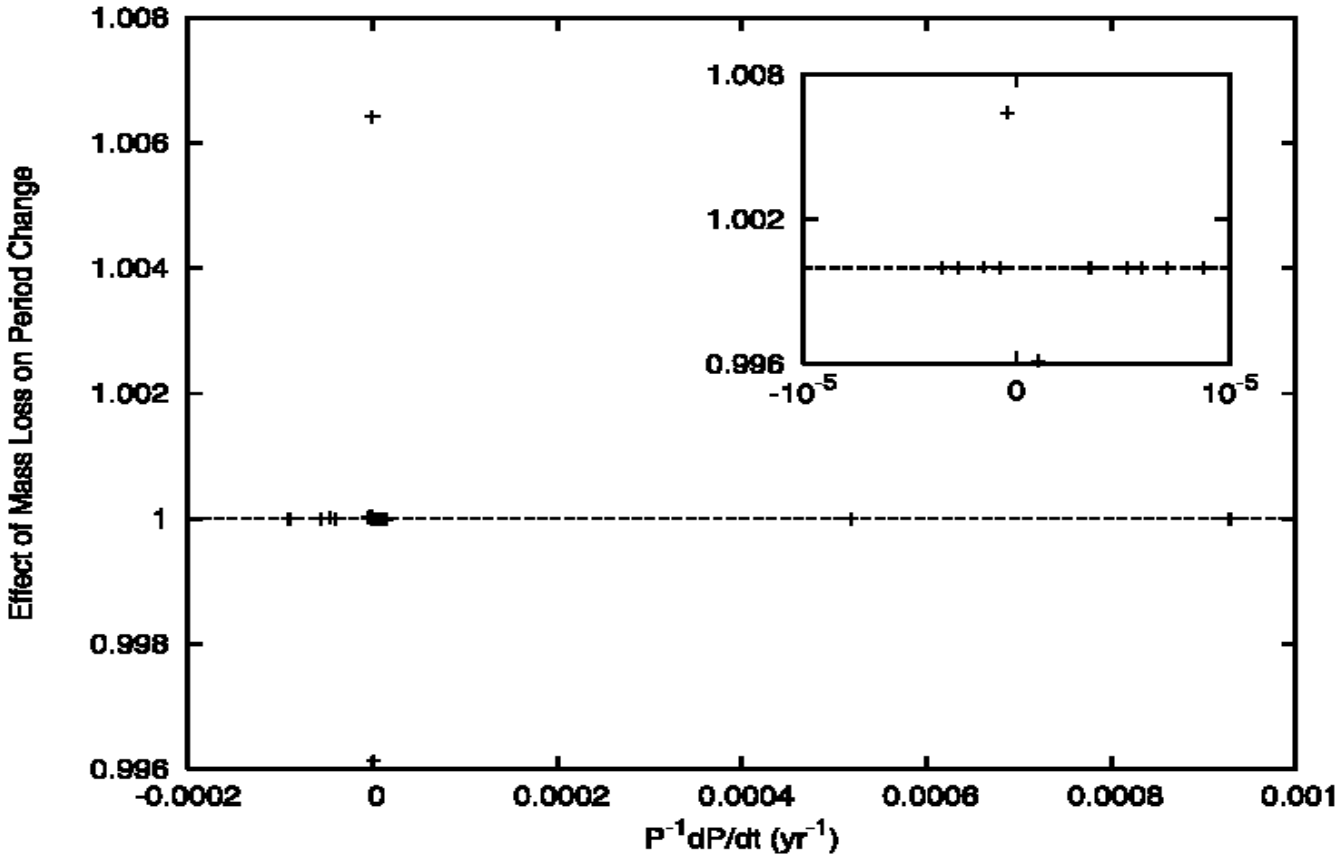
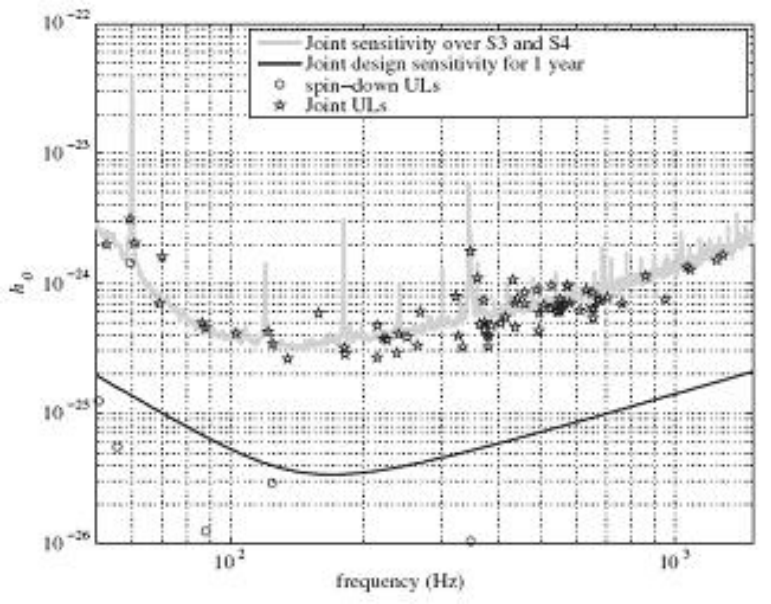
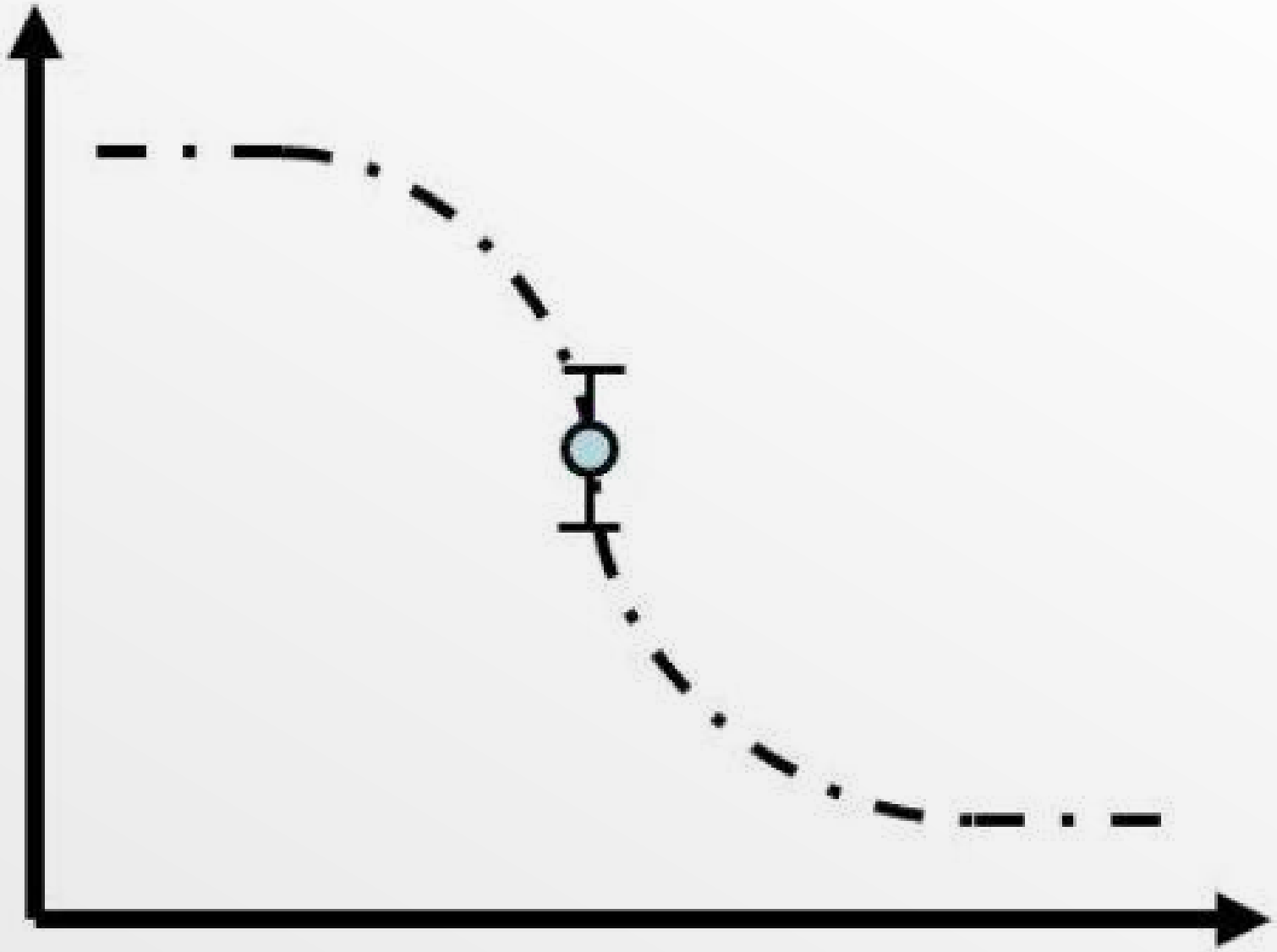
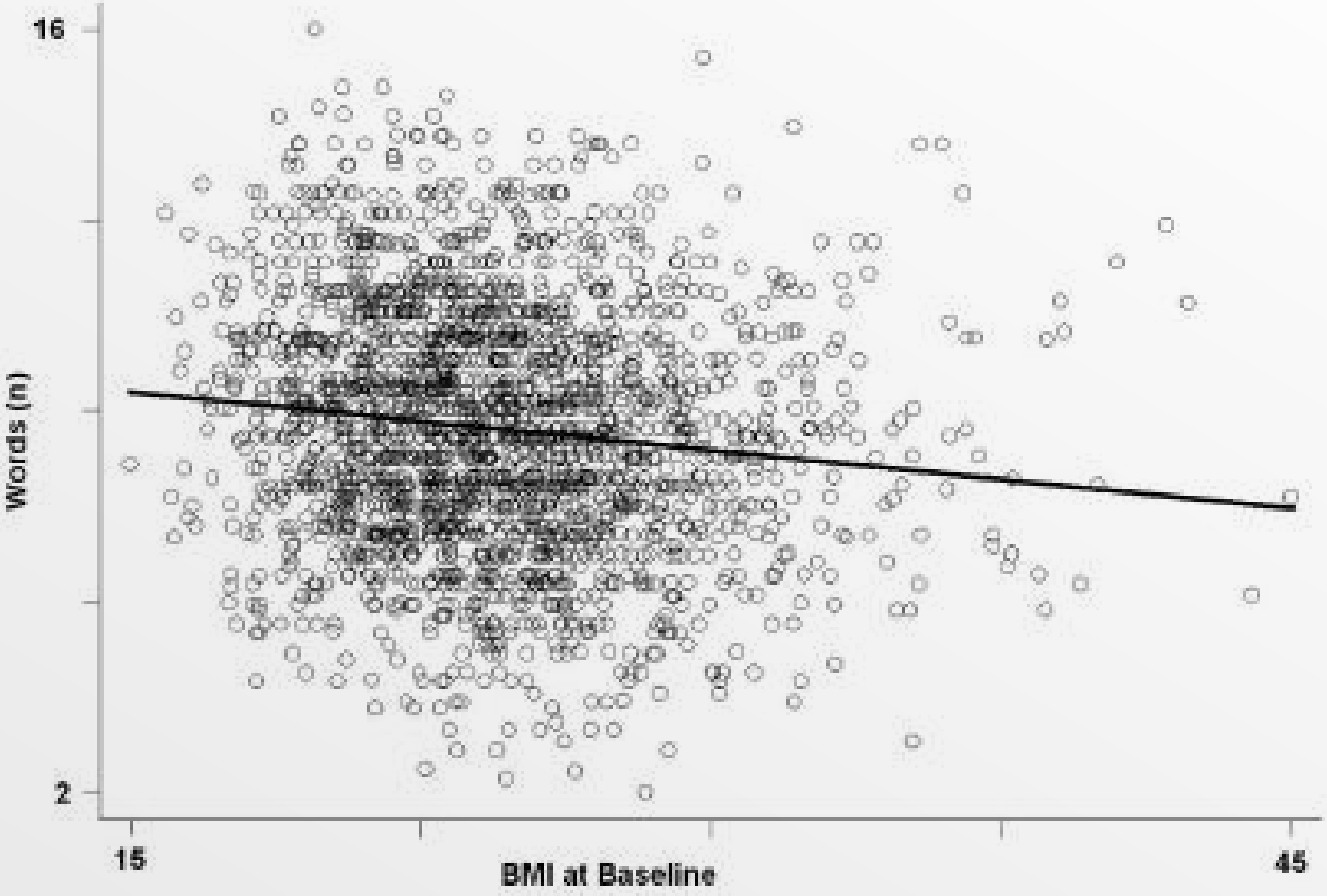


FIG. 11.—Fractional contribution of mass loss toward the period change for Cepheids as a function of period change. Cepheids with a period change not affected by the mass-loss rate would fall on the dashed line. Deviations from the dashed line measure how much mass loss plays a role.

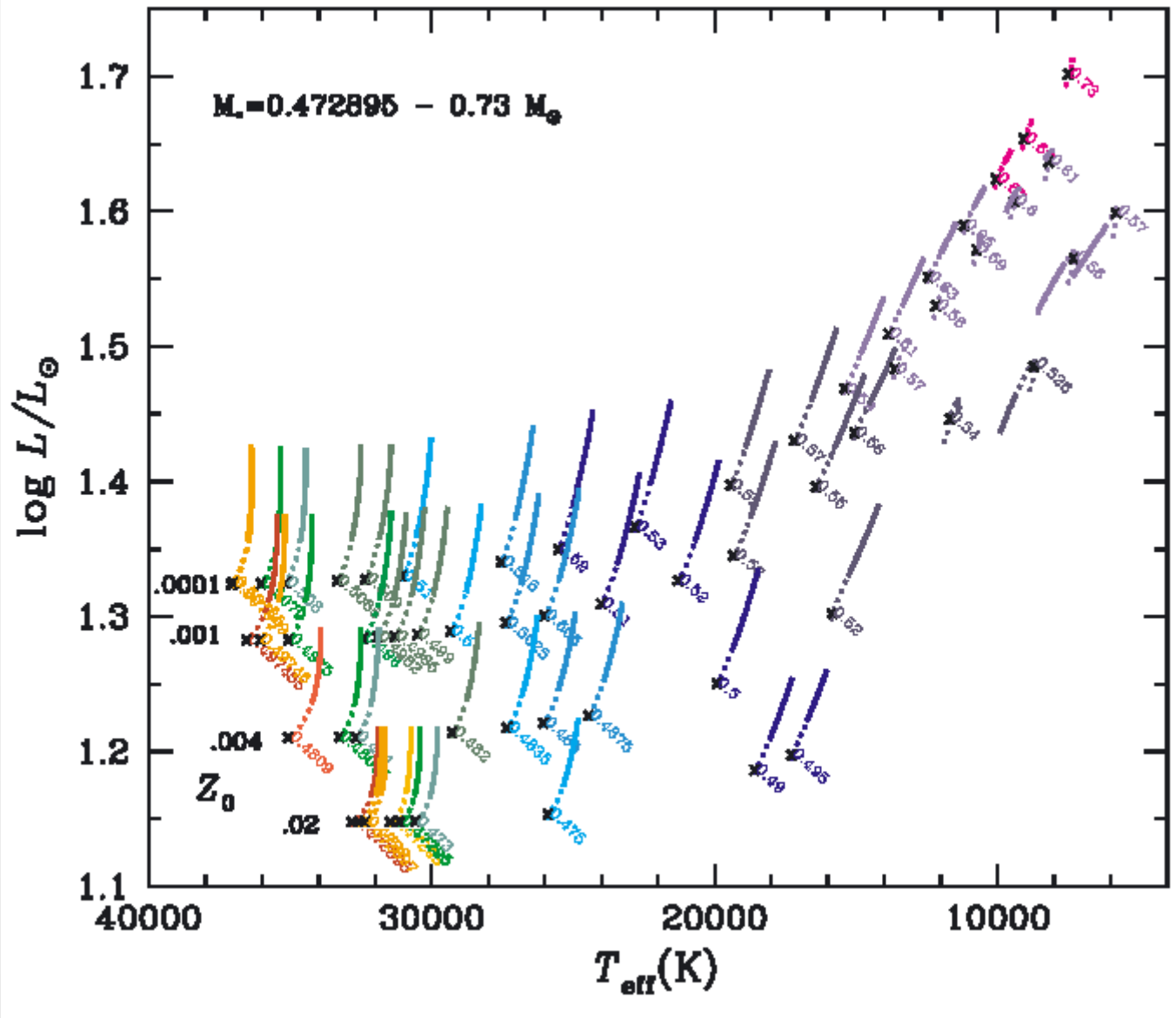
Improve me


Joint sensitivity over S3 and S4
Joint design sensitivity for 1 year
spin–down ULs
Joint ULs
10^{-22}
10^{-23}
10^{-24}
10^{-25}
10^{-26}
h_0
10^{2}
10^{3}
frequency (Hz)




16
2
Words (n)
15
45
BMI at Baseline

Improve me



Michaud et al 2011

Urgh

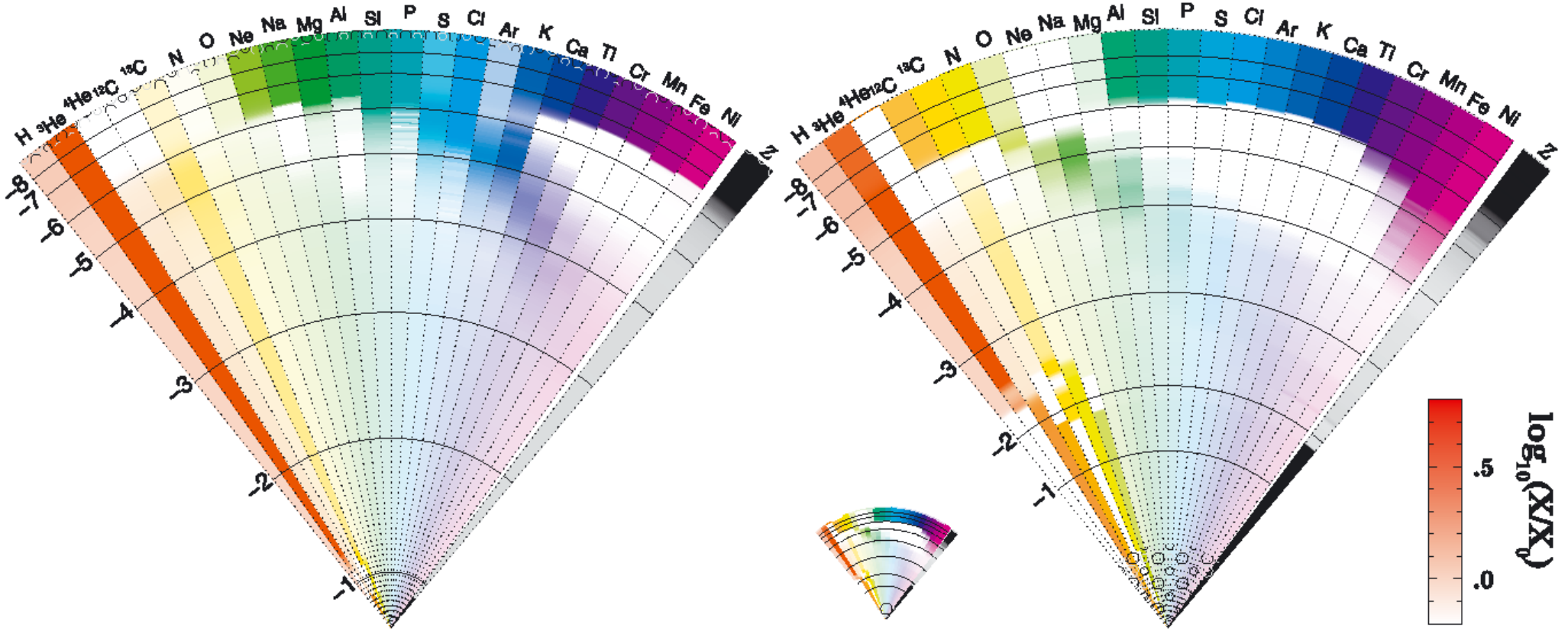
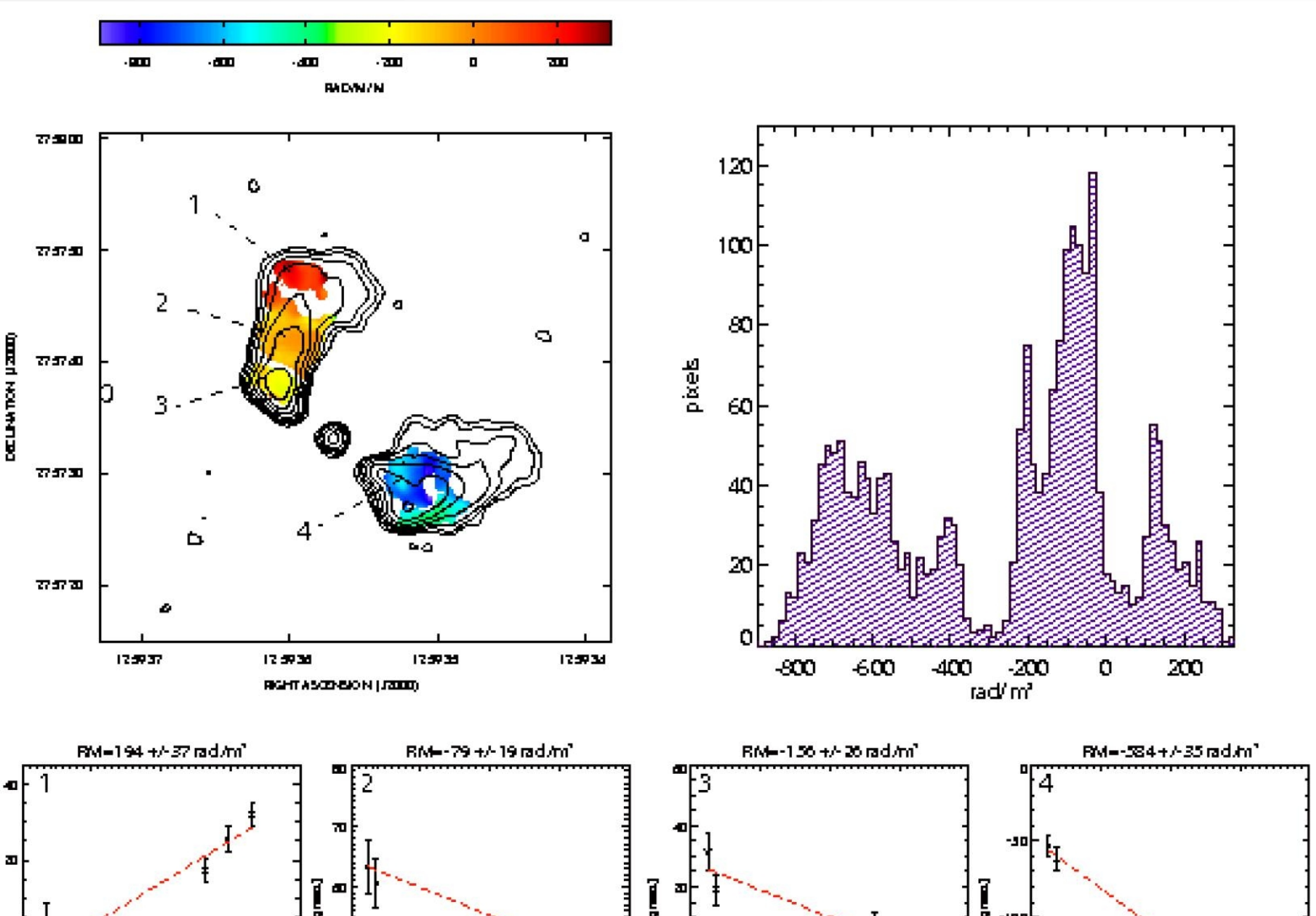


Fig. 4. Color-intensity coded concentrations in two HB stars of the same metallicity after 25 Myr on the HB. *Left panel* with a $T_{\rm eff}$ of 14 000 K (0.59 $M_\odot$) and *right panel* of 30 100 K (0.51 $M_\odot$). The radial coordinate is the radius and its scale is linear, but the logarithmic value of the mass coordinate above a number of points, $\log \Delta M/M_*$, is shown on the left of the horizontal black line. The concentration scale is given in the right insert. Small circles near the top of the left panel mark the extent of the surface convection zone, while similar circles near the center of both models mark the central convection zone. The small inset in between the two panels shows the high $T_{\rm eff}$ star, that is the right panel, on the radius scale of the low $T_{\rm eff}$ star, that is the left panel. For $-7 < \log \Delta M/M_* < -4$ the concentration is quite different for many species. It is surprisingly so for C and O for $\log \Delta M/M_* > -2$. See the text. A black-and-white version of this figure may be found in Fig. A.1.

Michaud et al 2011

Improve me


RADVN/N
DECLINATION (J2000)
RIGHT ASCENSION (J2000)
pixels
rad/m²
RM=194 +/- 37 rad/m²
RM=-79 +/- 19 rad/m²
RM=-156 +/- 26 rad/m²
RM=-384 +/- 35 rad/m²

What's missing?

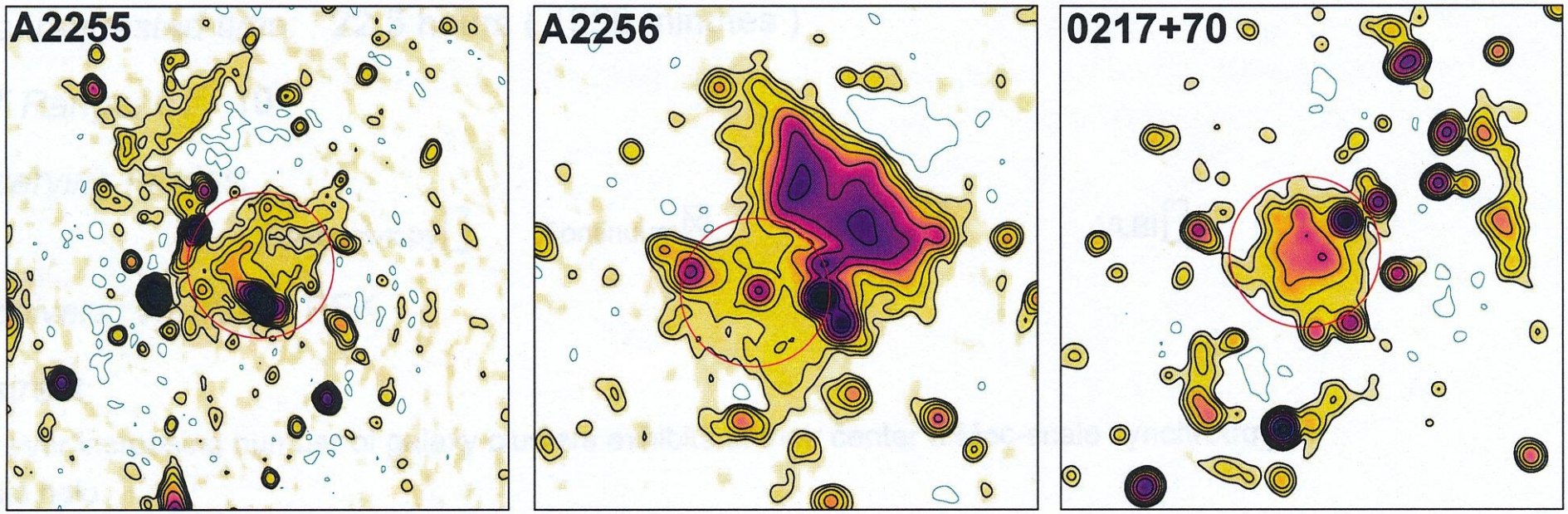
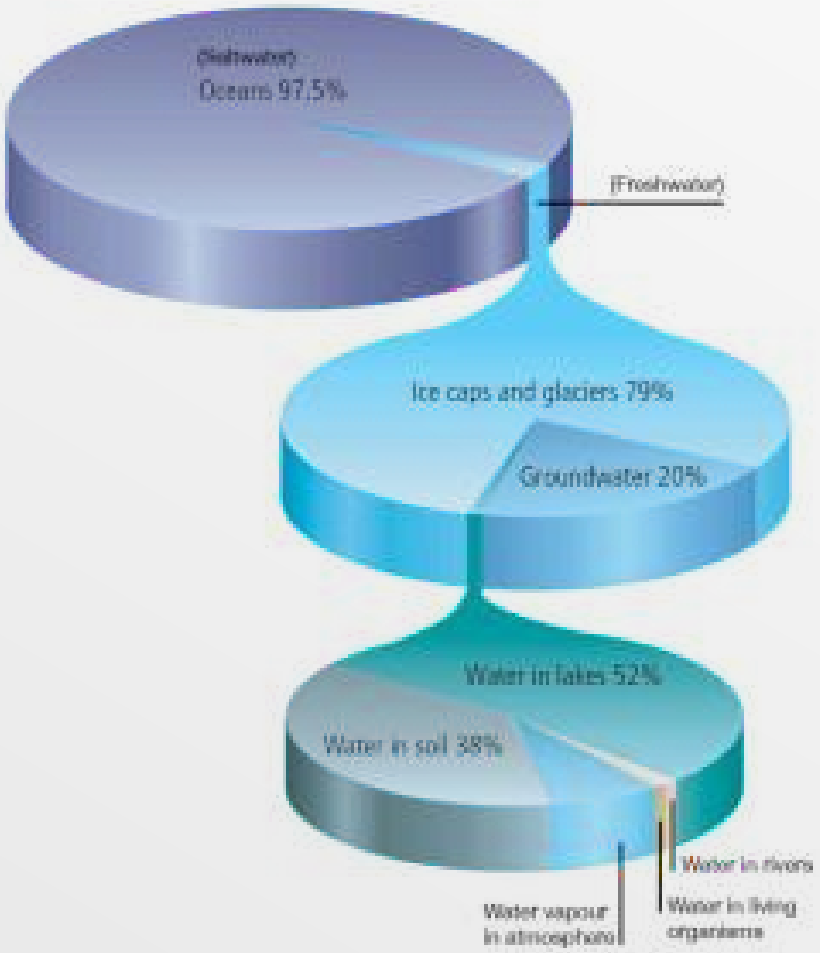
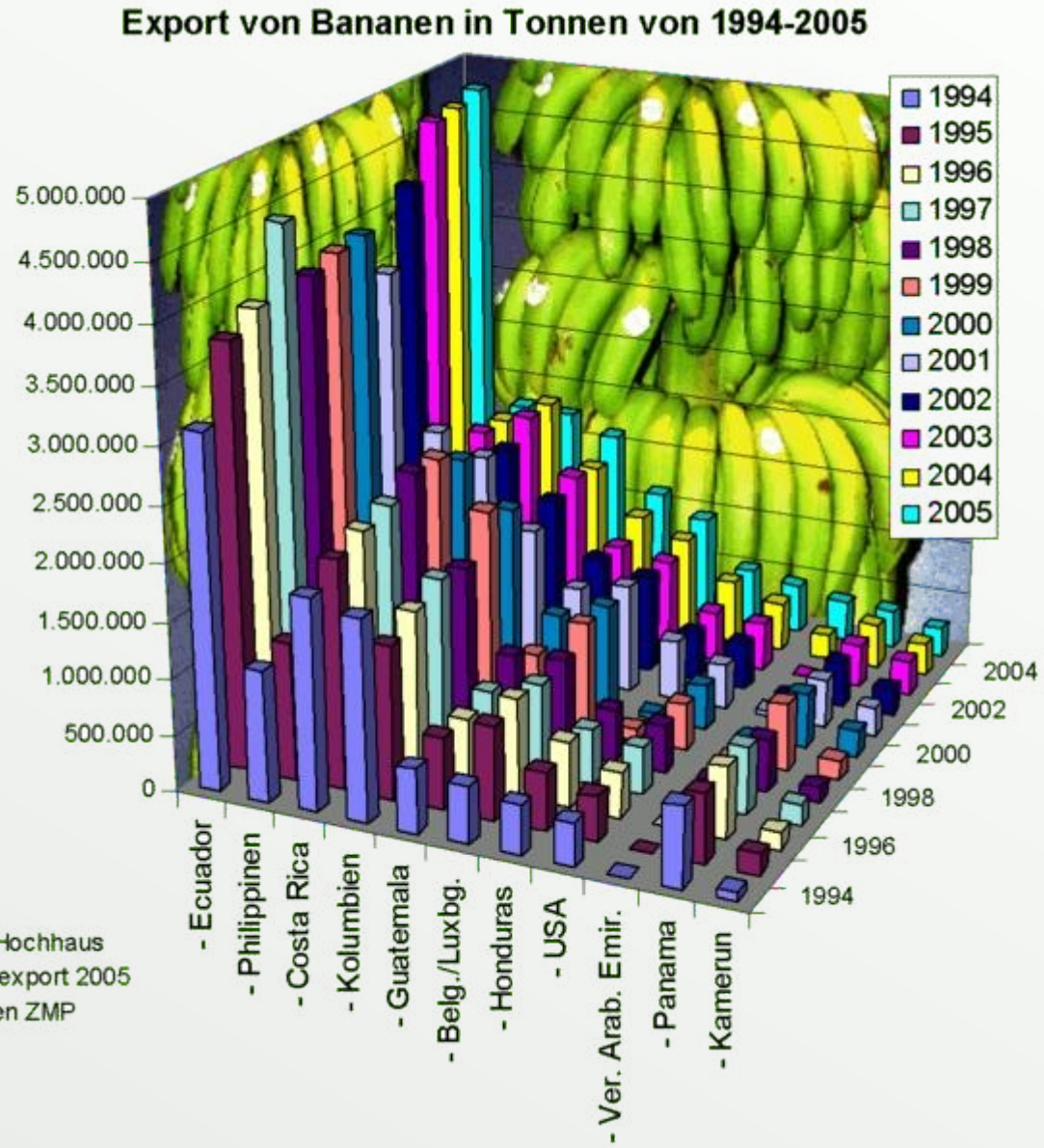
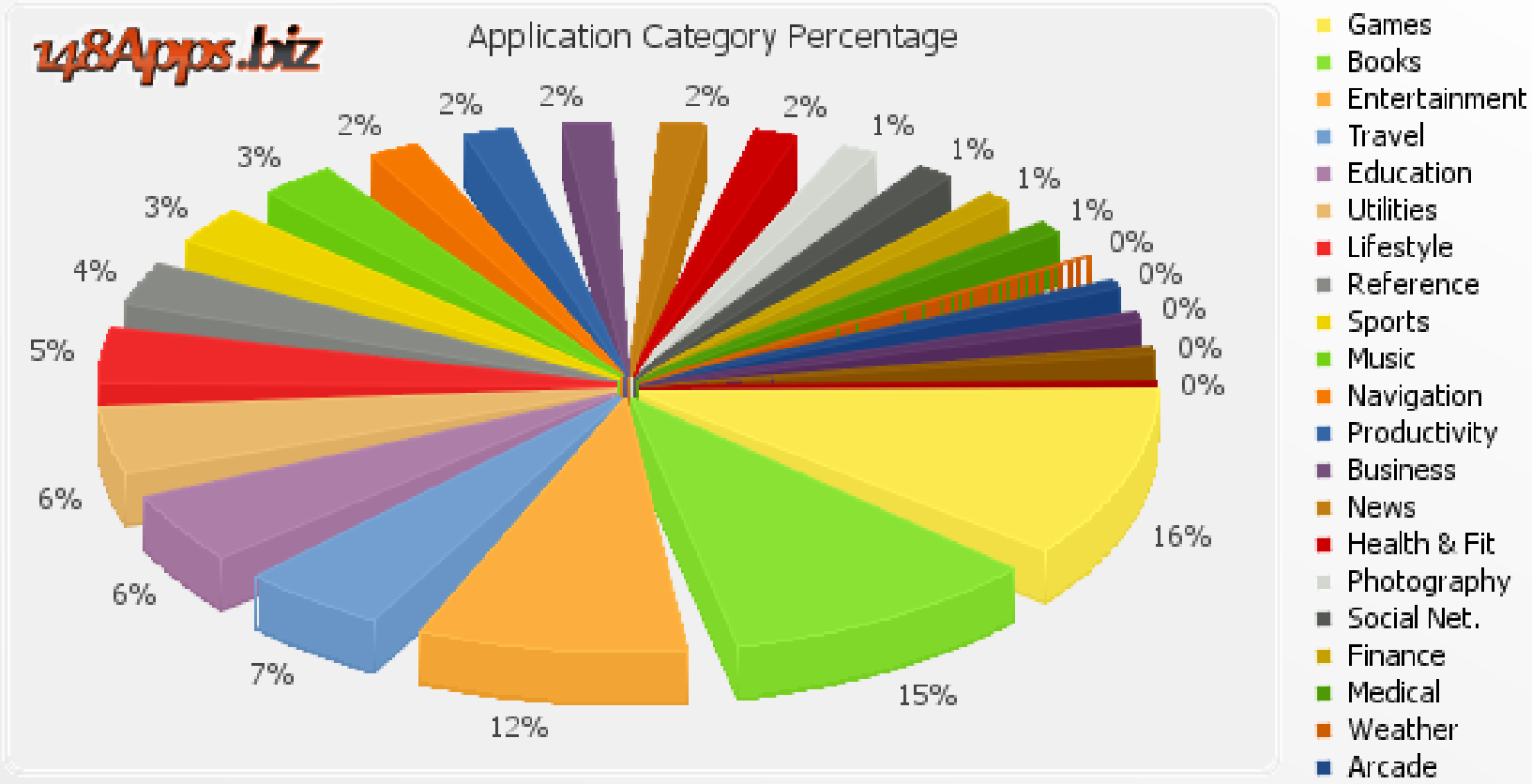


Figure 2 Total intensity VLA images at 1400 MHz of the radio halos sample proposed for Effelsberg 100-m observations: A399 (Murgia et al. 2010), A520 (Govoni et al. 2001), A523 (Giovannini et al. 2011), A665 (Vacca et al. 2010), A1914 (Bacchi et al. 2003), A2219 (Bacchi et al. 2003), A2255 (Govoni et al. 2005), A2256 (Clarke & Ensslin 2006), 0217+70 (Brown et al. 2011). Red circles indicate the Effelsberg beam at 1400 MHz (9.35′). Green lines indicate negative contours.


Oceans 97.5%
(Freshwater)
Ice caps and glaciers 79%
Groundwater 20%
Water in lakes 52%
Water in soil 38%
Water in rivers
Water vapour in atmosphere
Water in living organisms


Export von Bananen in Tonnen von 1994-2005
5.000.000
4.500.000
4.000.000
3.500.000
3.000.000
2.500.000
2.000.000
1.500.000
1.000.000
500.000
0
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
- Ecuador
- Philippinen
- Costa Rica
- Kolumbien
- Guatemala
- Belg./Luxbg.
- Honduras
- USA
- Ver. Arab. Emir.
- Panama
- Kamerun
Dr. Hochhaus
Banexport 2005
Daten ZMP

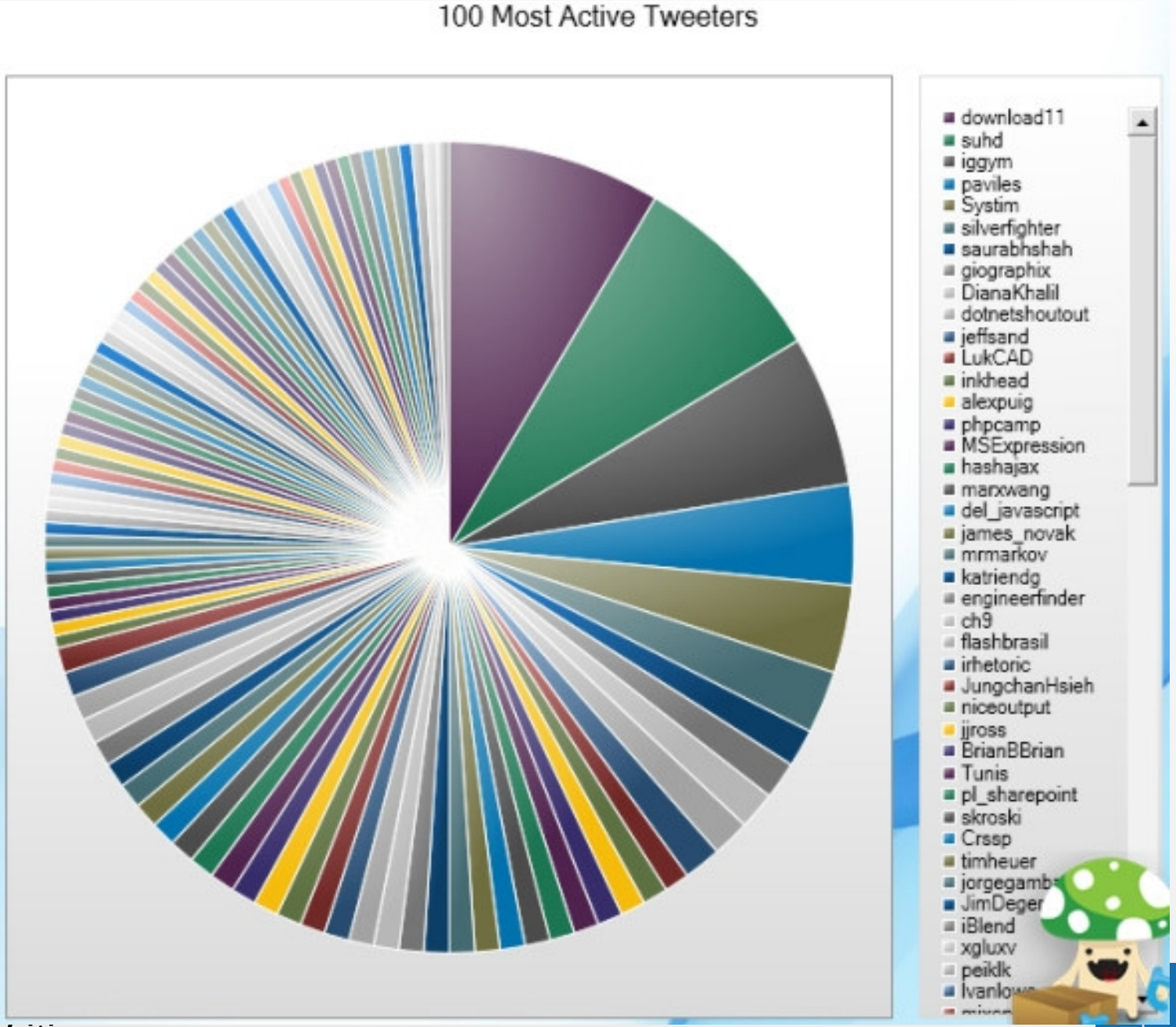


http://peltiertech.com/extra-distortion-in-a-pie-chart/

This clearly illustrates the distortion of the fancy wedge-gapped 3D pie chart. The amount of these pies that are **made up of nothing** is **31%**, leaving **only 69%** for **displaying data**. The largest wedge, which comprises 17% of the data, fills only 11% of the circle.

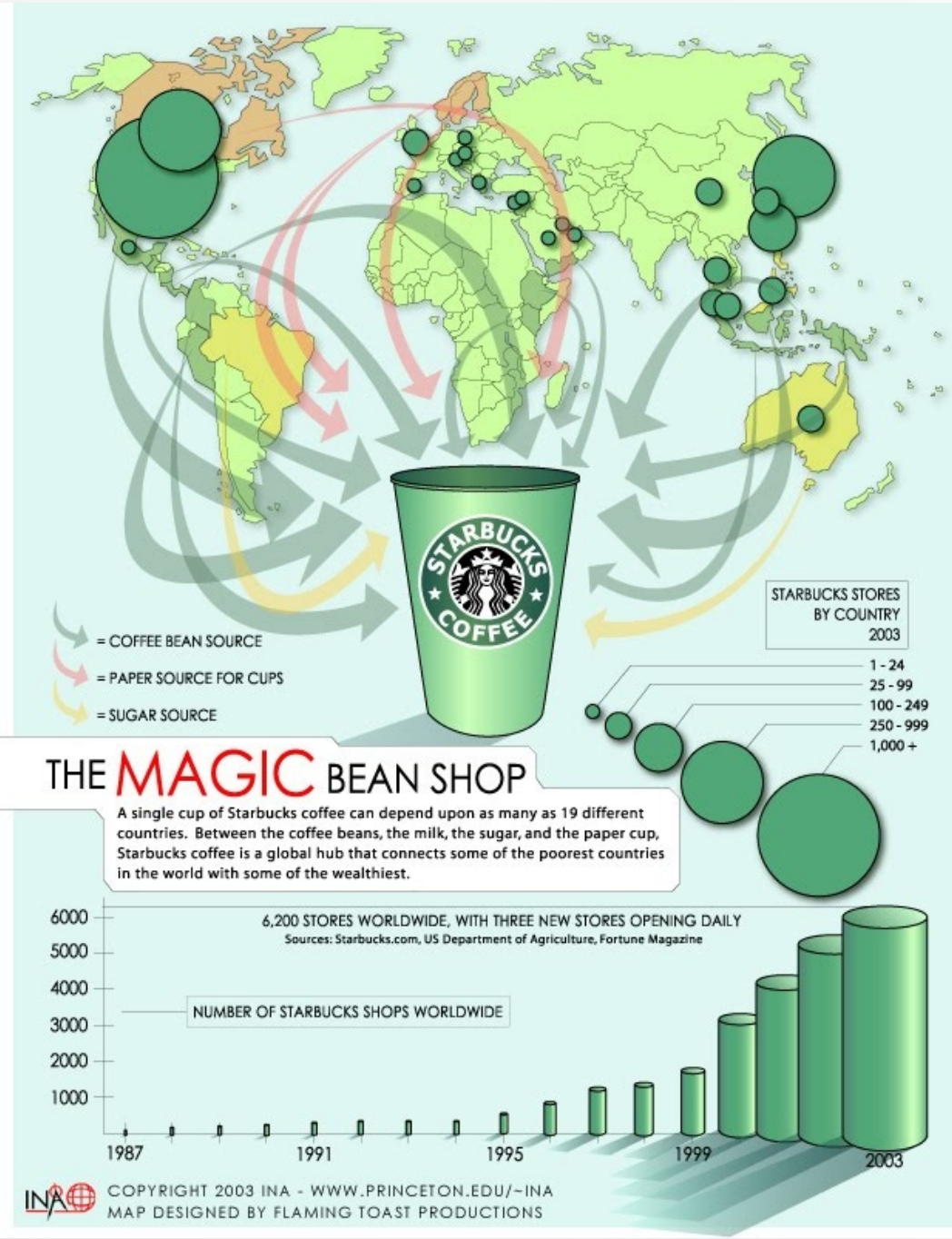
We humans are bad enough at judging areas and angles, so when the angles are distorted in this way, **we're hopeless.**

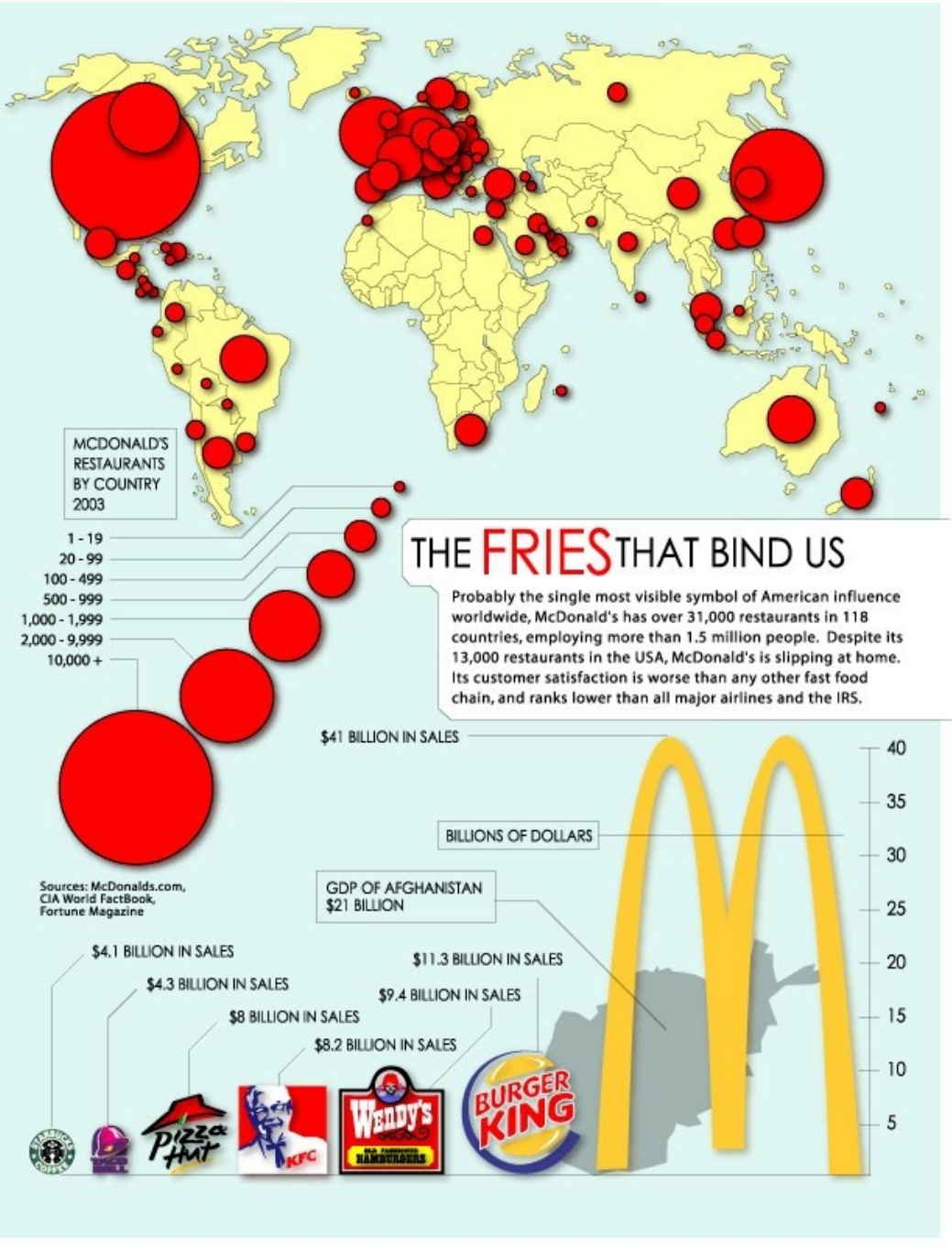
Make it stop!


100 Most Active Tweeters
download11
suhd
iggym
paviles
Systim
silverfighter
saurabhshah
giographix
DianaKhalil
dotnetshoutout
jeffsand
LukCAD
inkhead
alexpuig
phpcamp
MSExpression
hashajax
marxwang
del_javascript
james_novak
mrmarkov
katriendg
engineerfinder
ch9
flashbrasil
irhetoric
JungchanHsieh
niceoutput
jjross
BrianBBrian
Tunis
pl_sharepoint
skroski
Crssp
timheuer
JimDeger
iBlend
xgluxv
peikik

This gauge shows one number. Gauges make very inefficient use of space but are often used on dashboards where space is at a premium.

http://www.statsblogs.com/2012/05/30/winner-of-the-bad-graph-contest-announced/


= COFFEE BEAN SOURCE
= PAPER SOURCE FOR CUPS
= SUGAR SOURCE
STARBUCKS COFFEE
STARBUCKS STORES BY COUNTRY 2003
1 - 24
25 - 99
100 - 249
250 - 999
1,000 +
THE MAGIC BEAN SHOP
A single cup of Starbucks coffee can depend upon as many as 19 different countries. Between the coffee beans, the milk, the sugar, and the paper cup, Starbucks coffee is a global hub that connects some of the poorest countries in the world with some of the wealthiest.
6,200 STORES WORLDWIDE, WITH THREE NEW STORES OPENING DAILY
Sources: Starbucks.com, US Department of Agriculture, Fortune Magazine
NUMBER OF STARBUCKS SHOPS WORLDWIDE
6000
5000
4000
3000
2000
1000
1987
1991
1995
1999
2003
COPYRIGHT 2003 INA - WWW.PRINCETON.EDU/~INA
MAP DESIGNED BY FLAMING TOAST PRODUCTIONS


MCDONALD'S RESTAURANTS BY COUNTRY 2003
1 - 19
20 - 99
100 - 499
500 - 999
1,000 - 1,999
2,000 - 9,999
10,000 +
THE FRIES THAT BIND US
Probably the single most visible symbol of American influence worldwide, McDonald's has over 31,000 restaurants in 118 countries, employing more than 1.5 million people. Despite its 13,000 restaurants in the USA, McDonald's is slipping at home. Its customer satisfaction is worse than any other fast food chain, and ranks lower than all major airlines and the IRS.
Sources: McDonalds.com, CIA World FactBook, Fortune Magazine
$41 BILLION IN SALES
BILLIONS OF DOLLARS
GDP OF AFGHANISTAN $21 BILLION
$4.1 BILLION IN SALES
$4.3 BILLION IN SALES
$8 BILLION IN SALES
$8.2 BILLION IN SALES
$9.4 BILLION IN SALES
$11.3 BILLION IN SALES
Pizza Hut
KFC
Wendy's
BURGER KING
40
35
30
25
20
15
10
5

OMG!

- The Hurley **"52 bug"** (Hurley et al 2002 MNRAS)

$$\frac{1}{\tau_{\rm sync}} = 52^{5/3} \sqrt{\frac{GM}{R^3}} \frac{MR^2}{I} q^2 (1+q)^{5/6} E_2 \left(\frac{R}{a}\right)^{17/2}$$

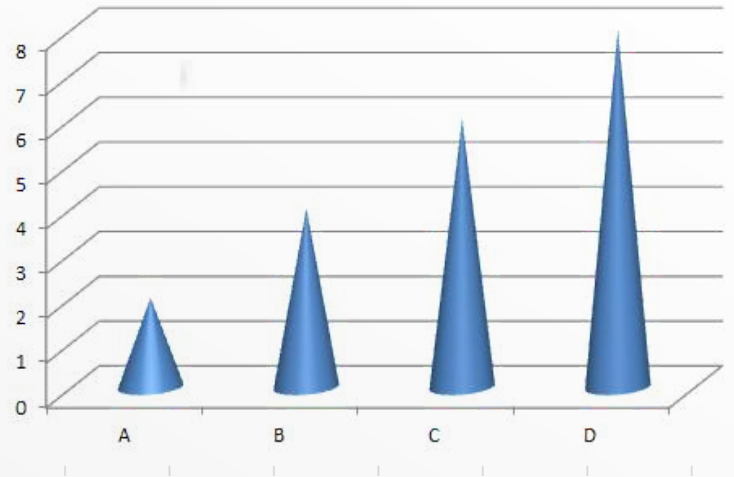
- The original formula from Zahn (1977, A&A)

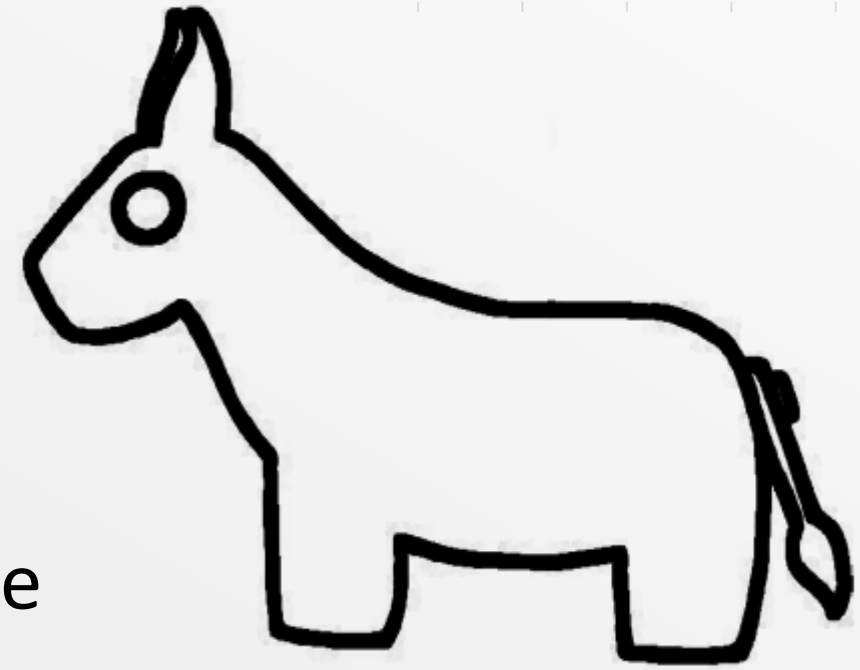
$$\frac{1}{\tau_{\rm sync}} = 5\,2^{5/3} \sqrt{\frac{GM}{R^3}} \frac{MR^2}{I} q^2 (1+q)^{5/6} E_2 \left(\frac{R}{a}\right)^{17/2}$$

Only a factor of *45* different . . .

Better!

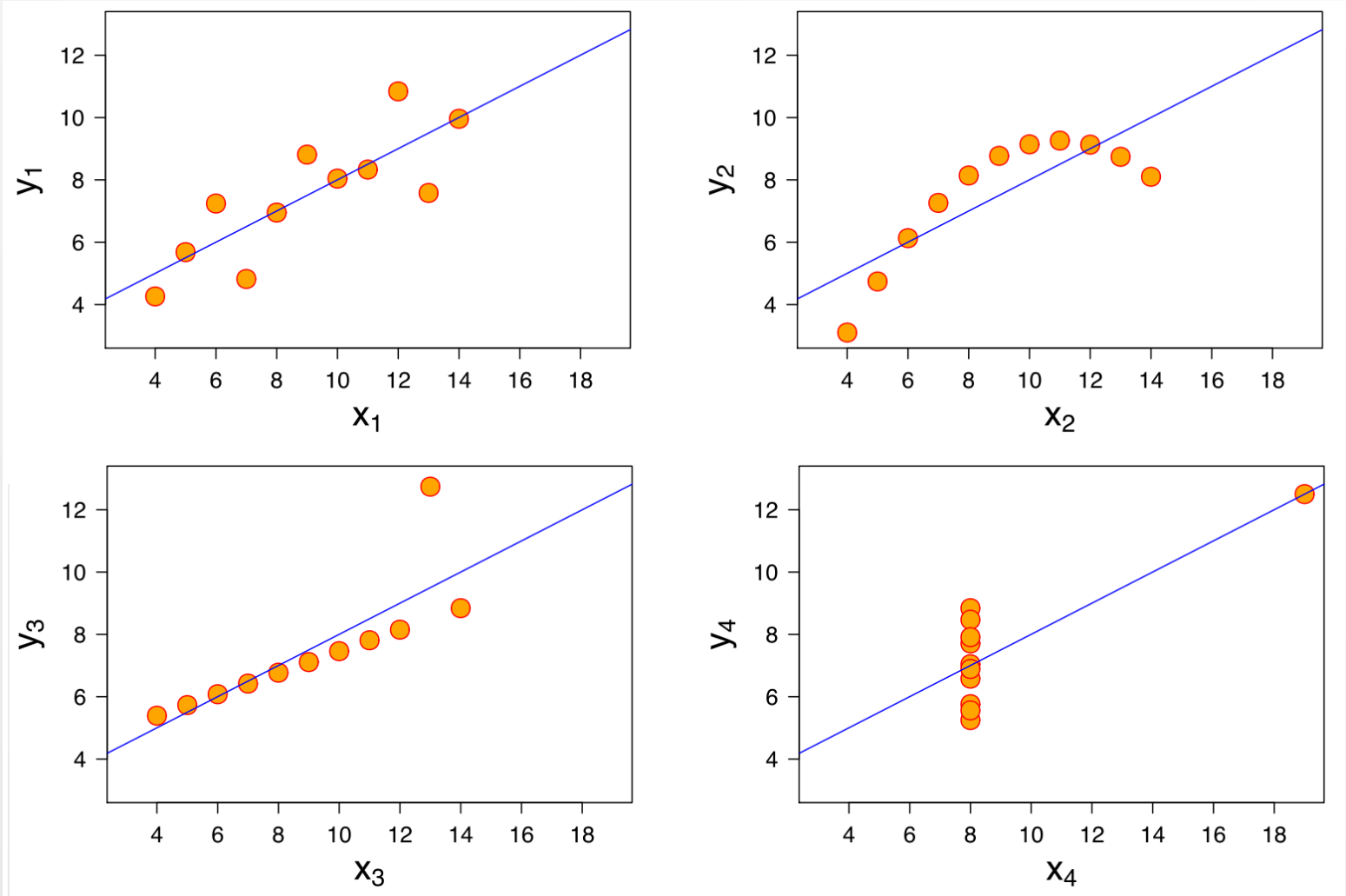
Exercises





- Ithaca times
- US politics
- Anscombe's Quartet
- Discussion on this course
- http://www.informationisbeautiful.net/visualizations/

Anscombe's Quartet



$$\bar{x} = 9$$
$$\sigma_x = 11$$
$$\bar{y} = 7.5$$
$$\sigma_y \simeq 4.12$$
$$R = 0.816$$
$$y = 3.00 + 0.50x$$

http://en.wikipedia.org/wiki/Anscombe%27s_quartet

Anscombe (1973) *The American Statistician vol 27, pp 17*

John McCarthy (famous computer scientist):

"As the Chinese say, 1001 words is worth more than a picture."