Making Better Figures

Theory

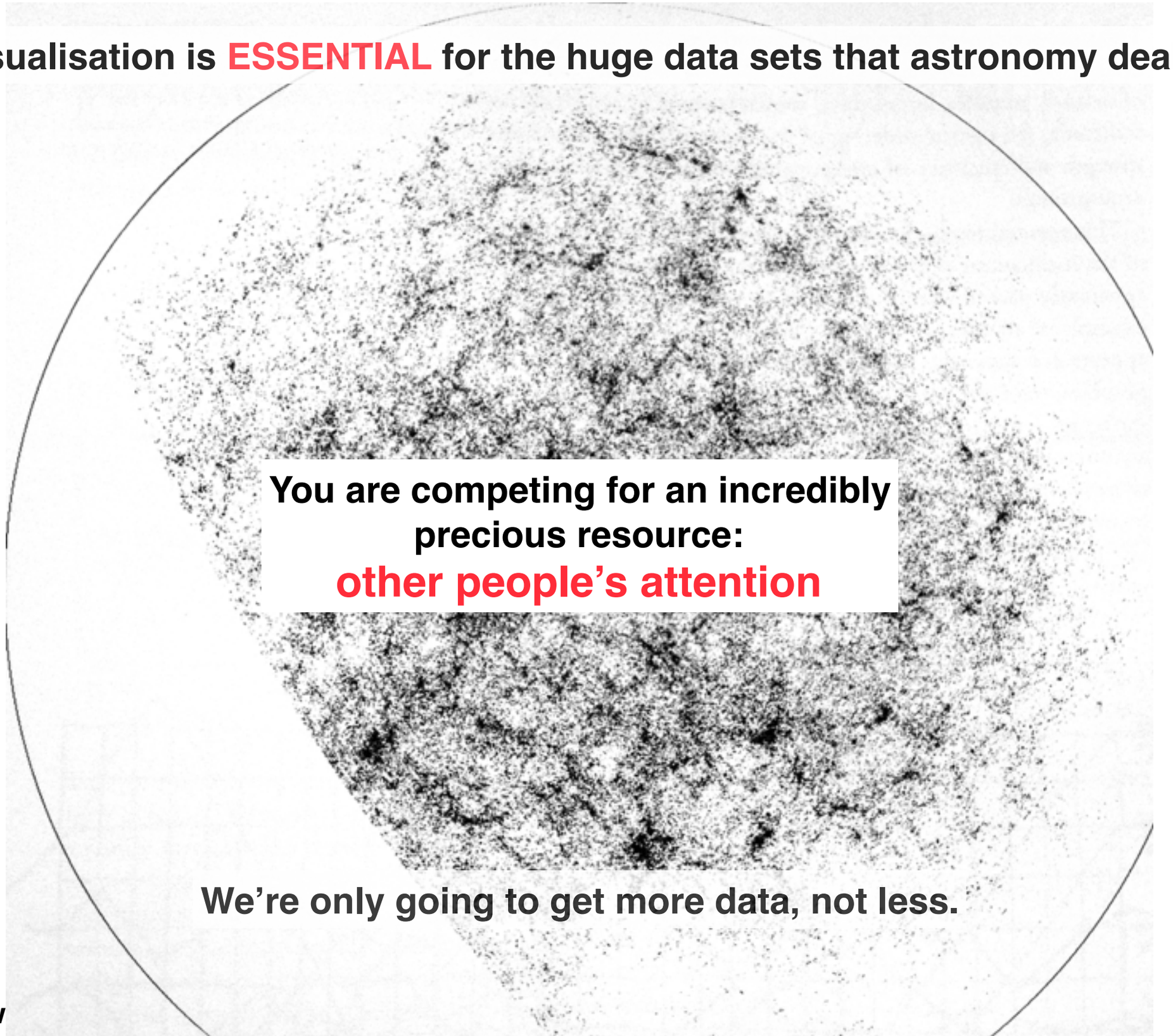
and

How to find that element in matplotlib

Matthew Kenworthy // Leiden Observatory // 2016

Why should you care about this?

Data visualisation is ESSENTIAL for the huge data sets that astronomy deals with.



You are competing for an incredibly precious resource:
other people's attention

We're only going to get more data, not less.

Matthew

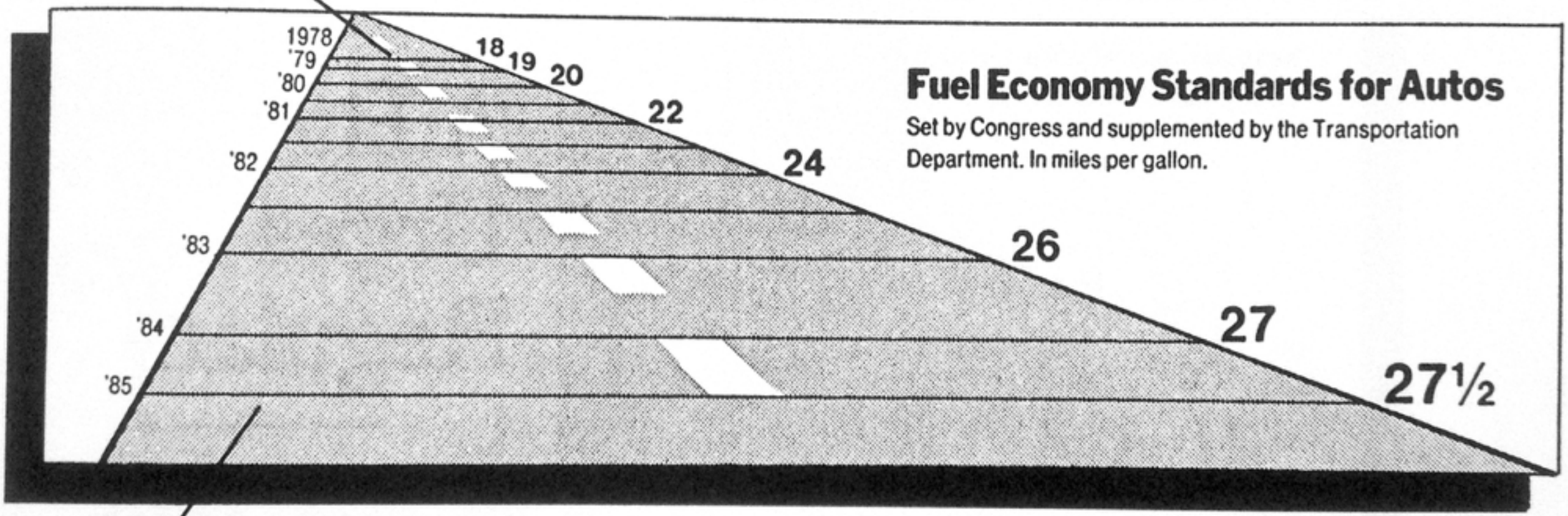
Early Infographics are AWFUL

“I lie to myself, all the time.” —Kenneth on *30 Rock*

Matthew Kenworthy // Leiden Observatory // 2016

Lying with forced perspective

This line, representing 18 miles per gallon in 1978, is 0.6 inches long.

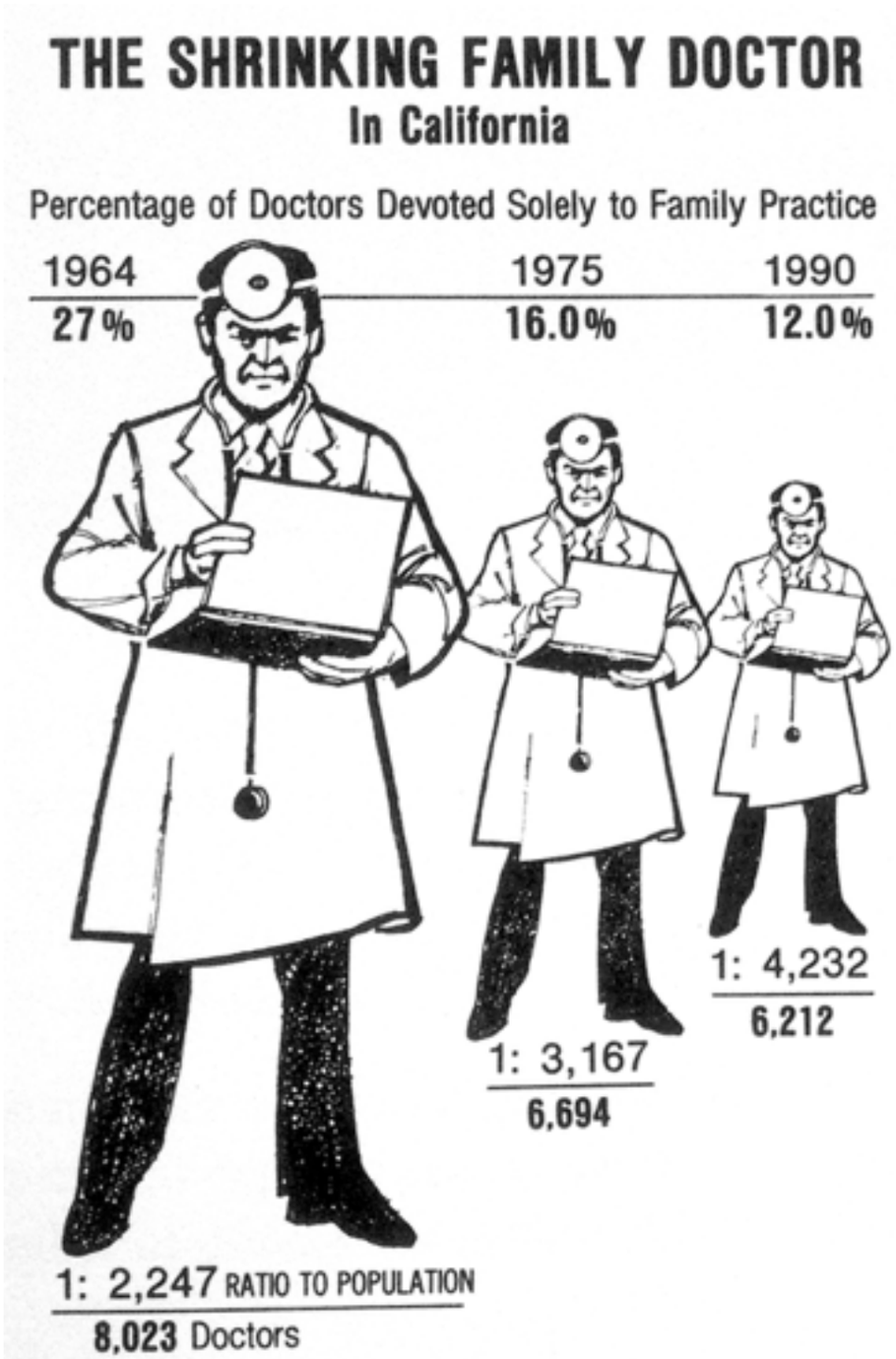


This line, representing 27.5 miles per gallon in 1985, is 5.3 inches long.

New York Times, August 9, 1978, p. D–2.

Matthew Kenworthy // Leiden Observatory // 2016

Is the quantity the area or height?



Answer: it's neither!

But what do people assume?

Answer: Everyone assumes different assumptions

So don't do it!

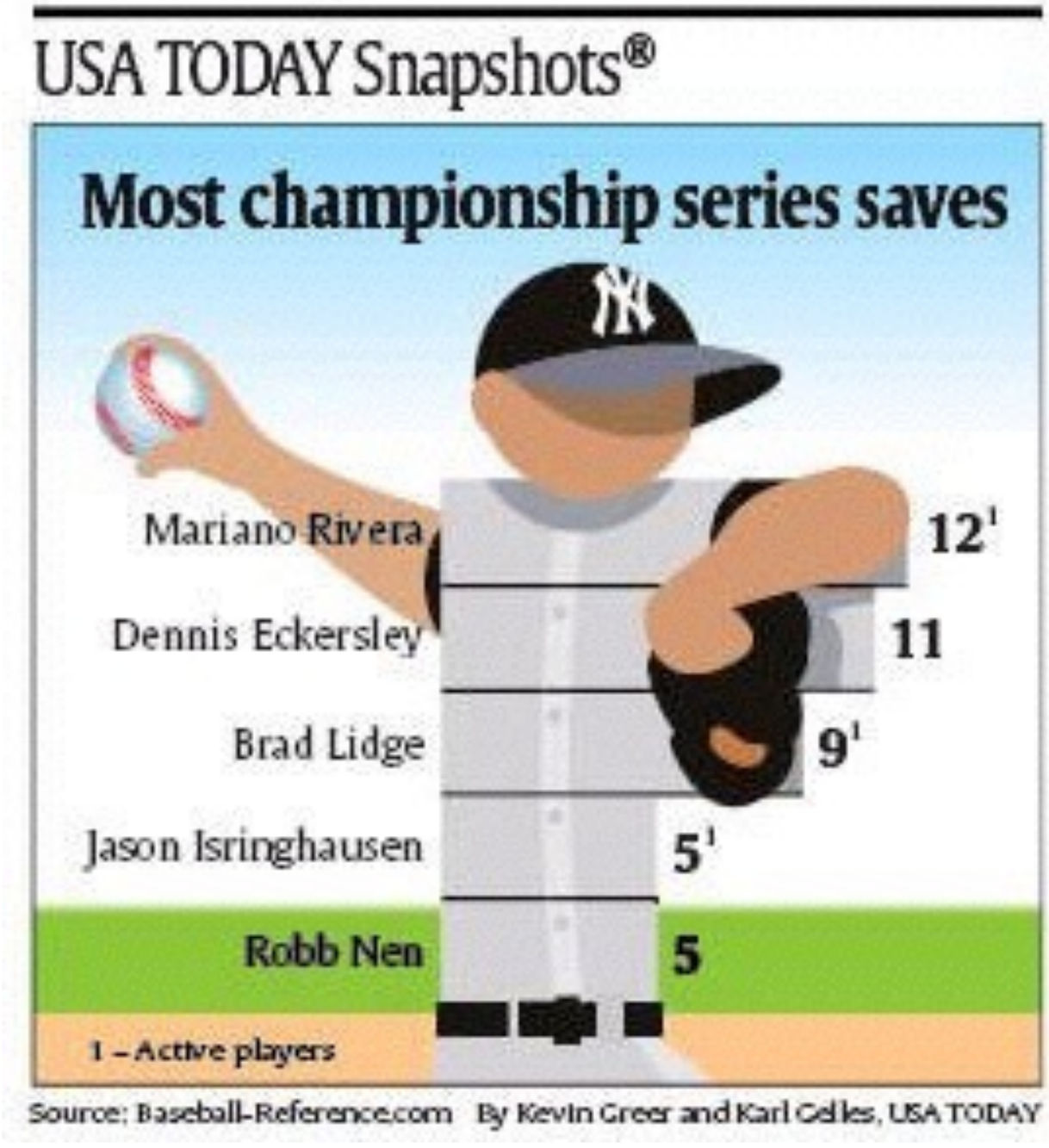
Los Angeles Times, August 5, 1979, p. 3.

USA Today - home of the howler

Black nozzle makes data look larger

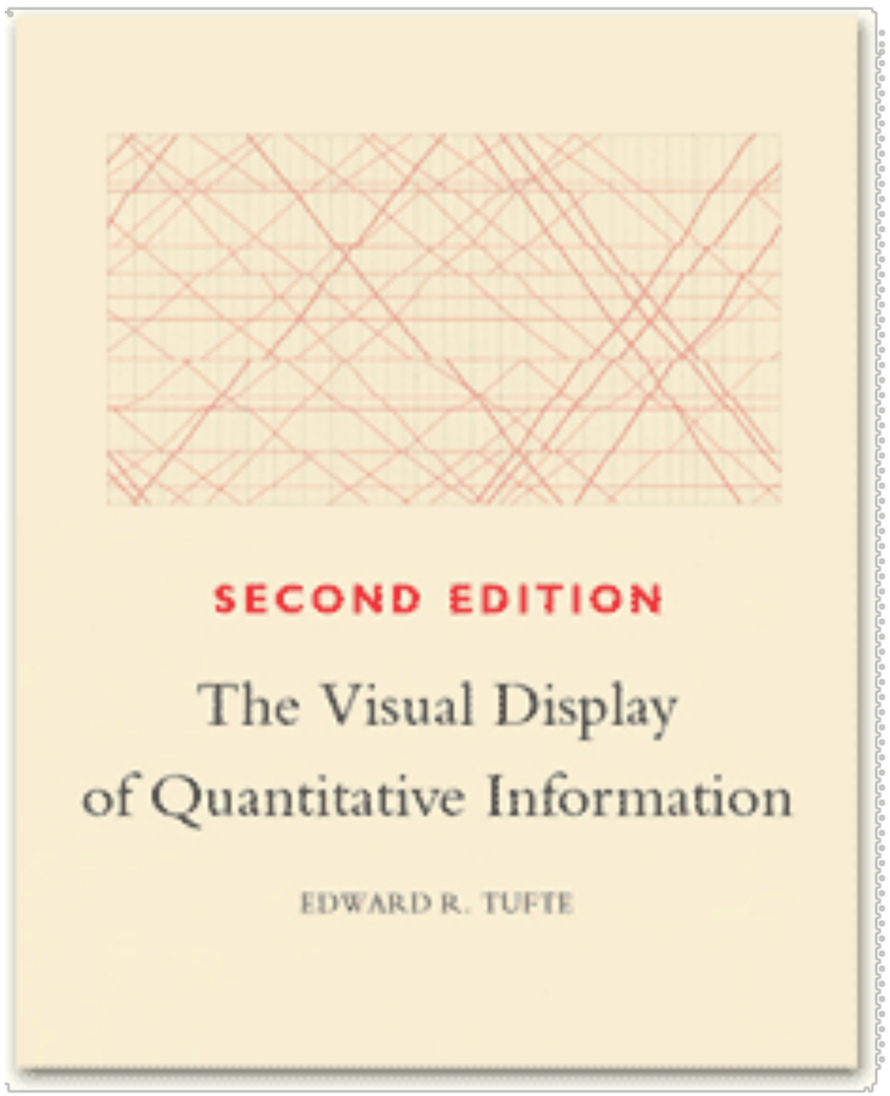
Vertical axis is not scaled properly

This is awful



Matthew Kenworthy // Leiden Observatory // 2016

Edward Tufte's book:
"The Visual Display of Quantitative Information"



Analyses really bad plots and explains why they're bad

Discusses how to present statistical data from many science fields

It was written at a time when computer graphics were only just becoming widespread

It's a beautiful coffee book!

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

Matthew Kenworthy // Leiden Observatory // 2016

Tufte's Suggestions

Illustrated by Merchant for the Brunswick Review
http://www.washingtonmonthly.com/magazine/mayjune_2011/features/the_information_sage029137.php

Above all else show the data.

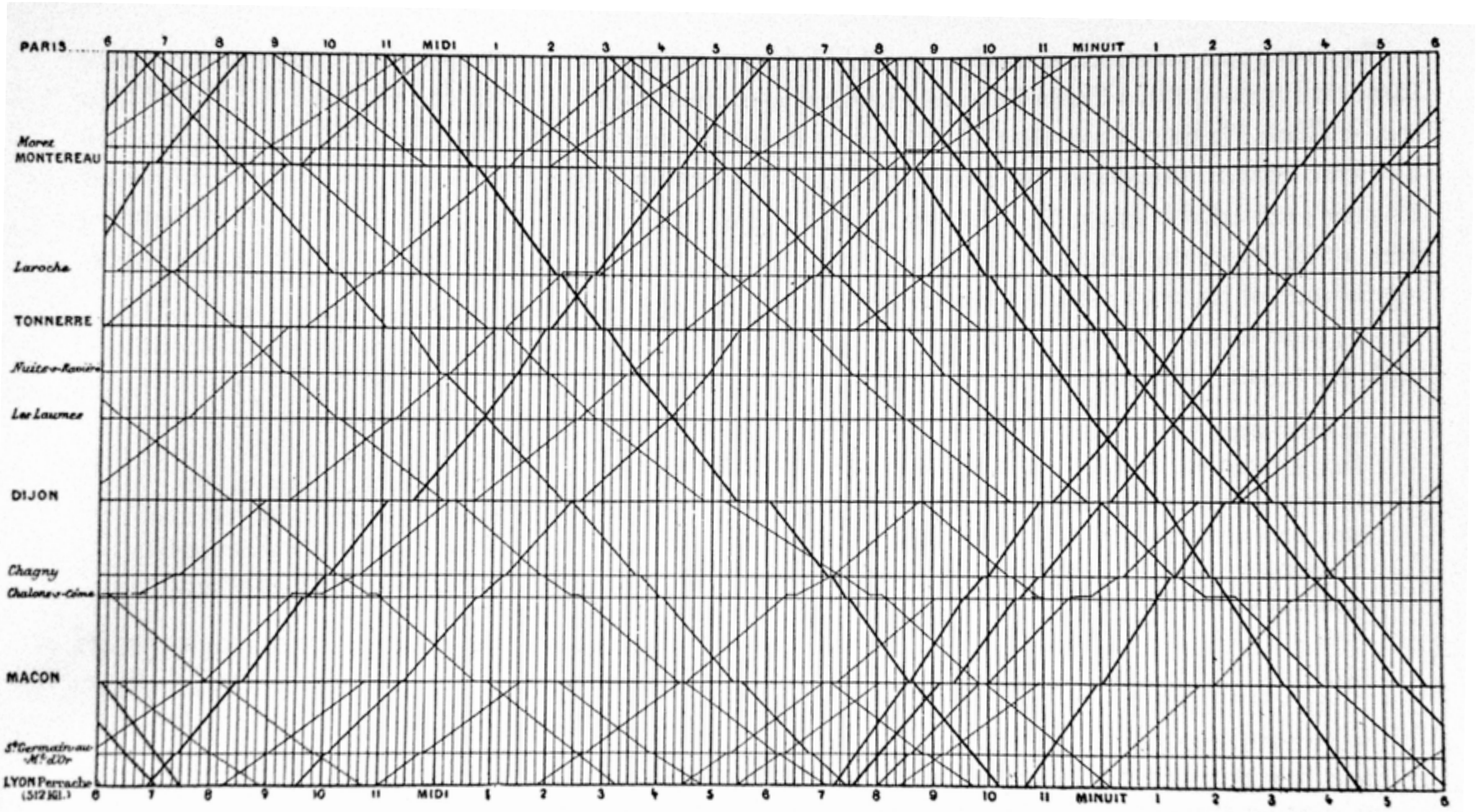
Maximise the data-ink ratio.

Erase non-data-ink.

Erase redundant data-ink.

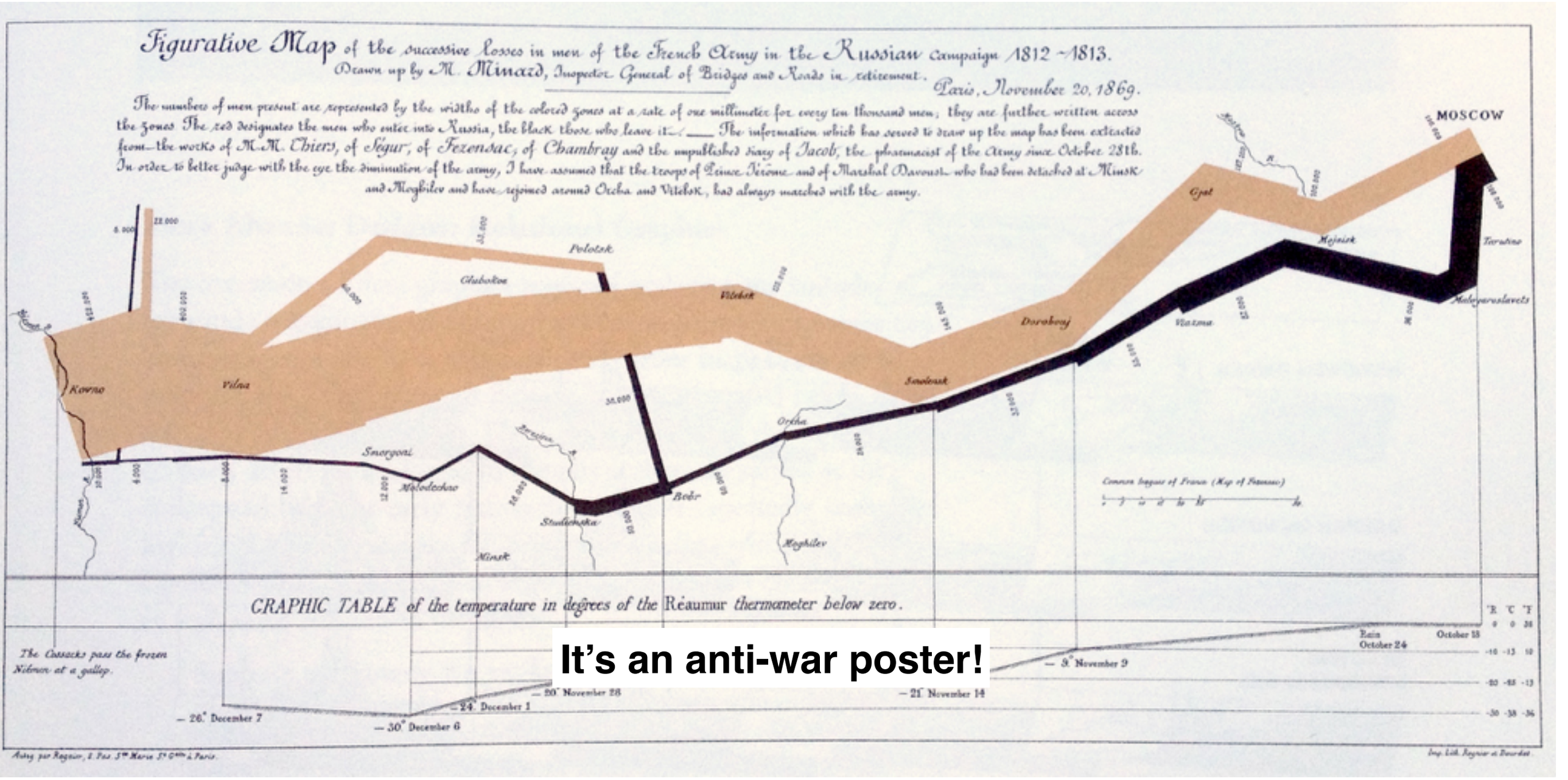
Revise and edit.

Paris to Lyon train timetable



E. J. Marey, *La méthode graphique* (Paris, 1885), p. 20. The method is attributed

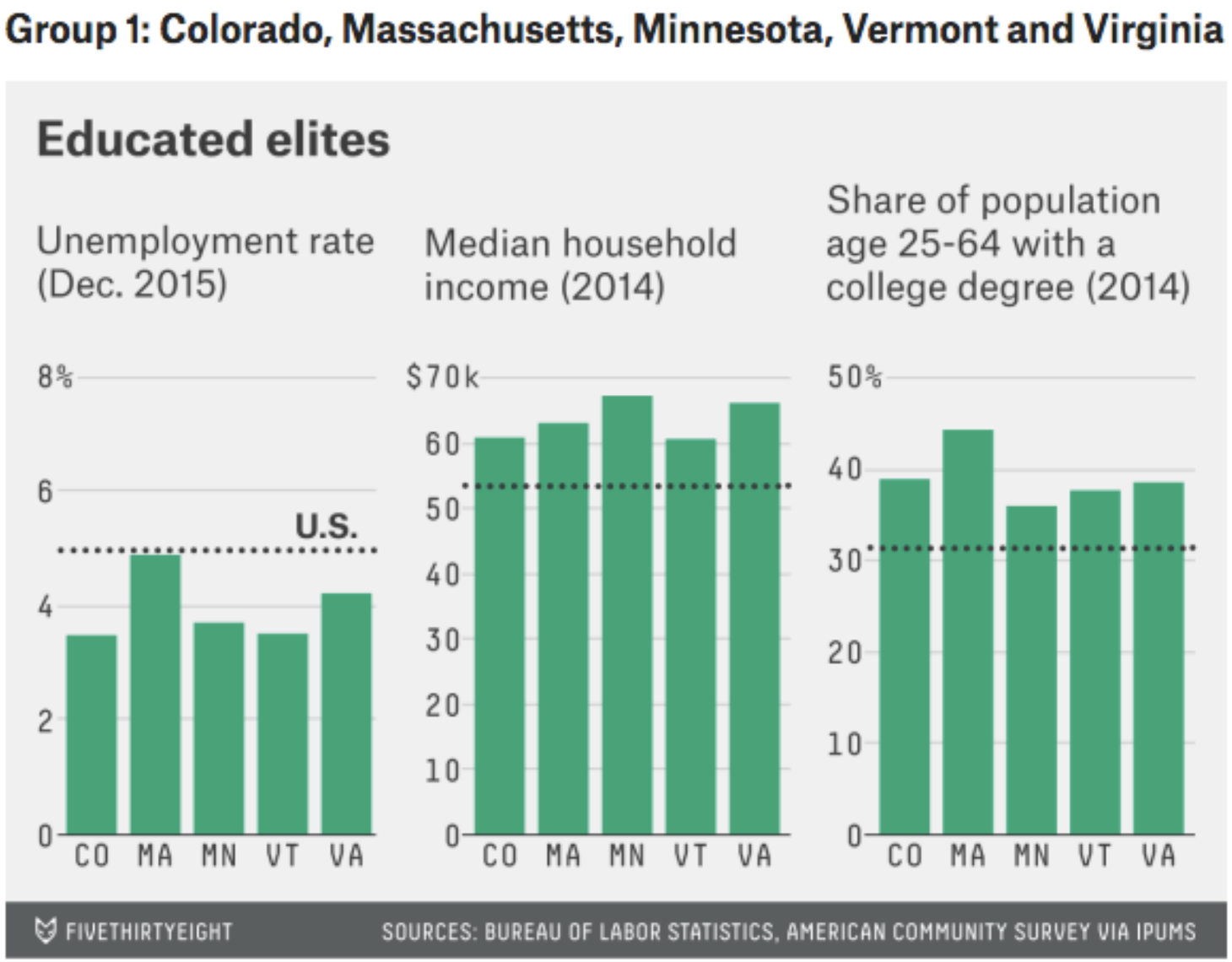
French Army losses in 1812



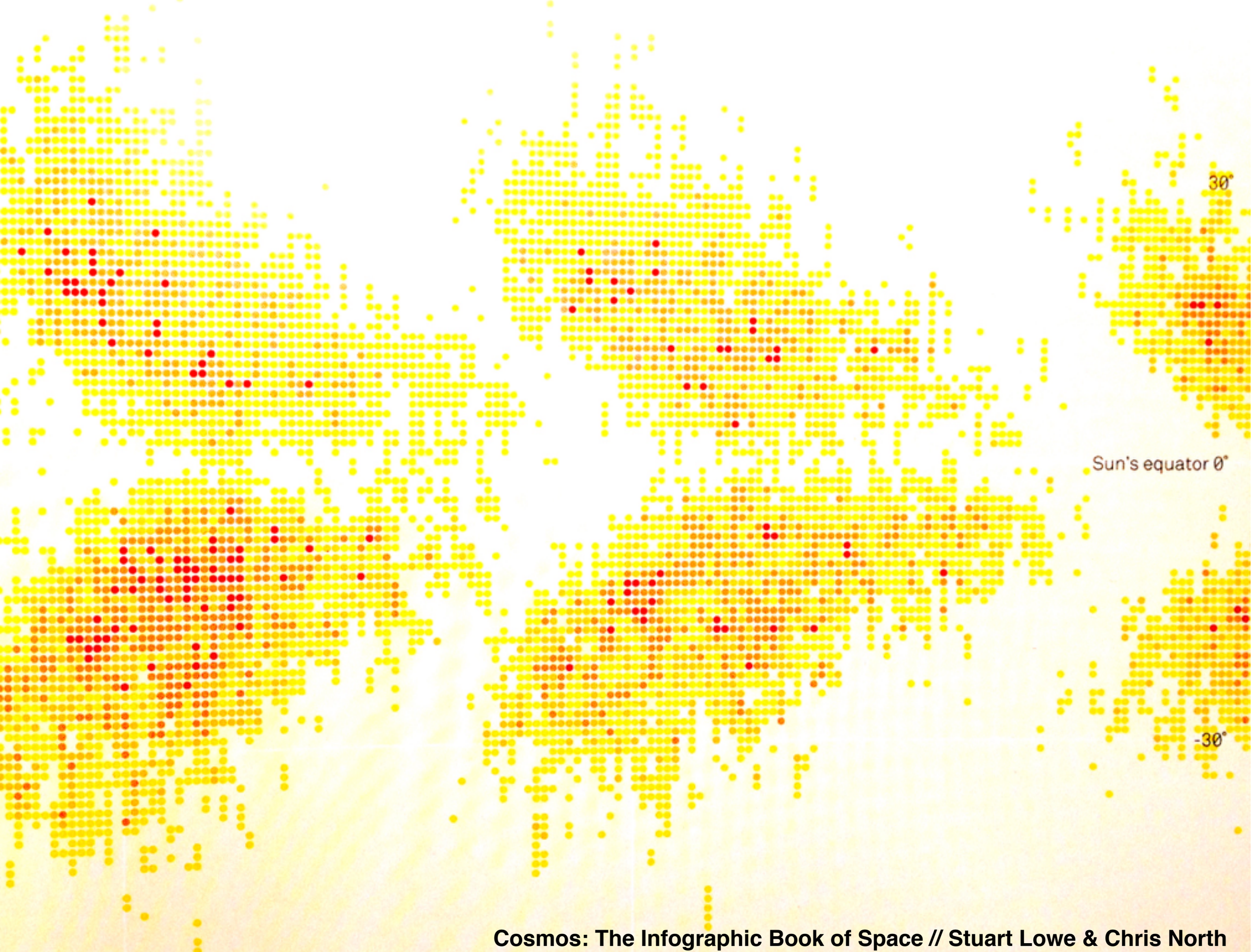
It's an anti-war poster!

Matthew Kenworthy // Leiden Observatory // 2016

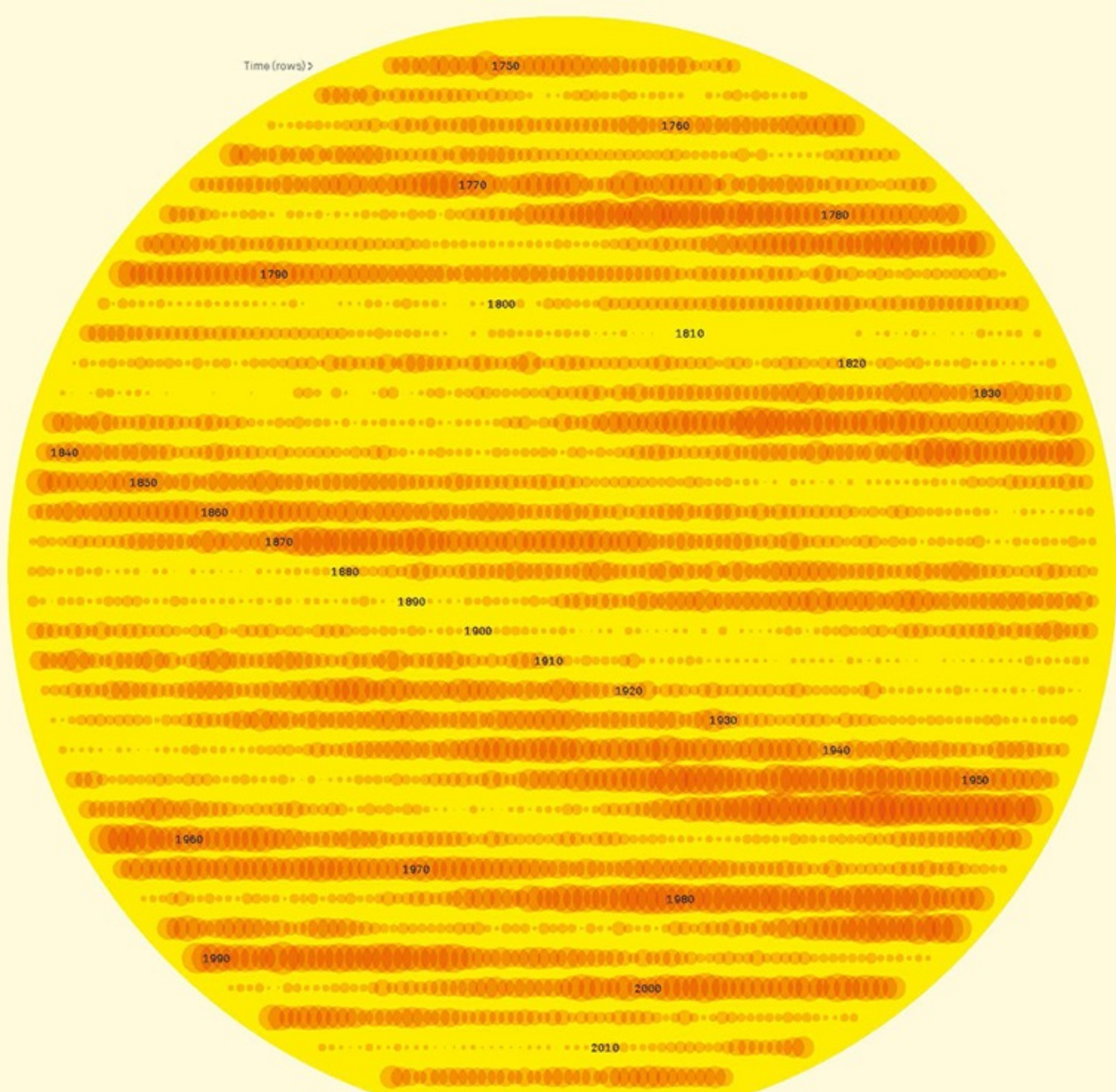
adopters of Tufte's principles



http://fivethirtyeight.com/features/super-tuesday-economy/



Cosmos: The Infographic Book of Space // Stuart Lowe & Chris North

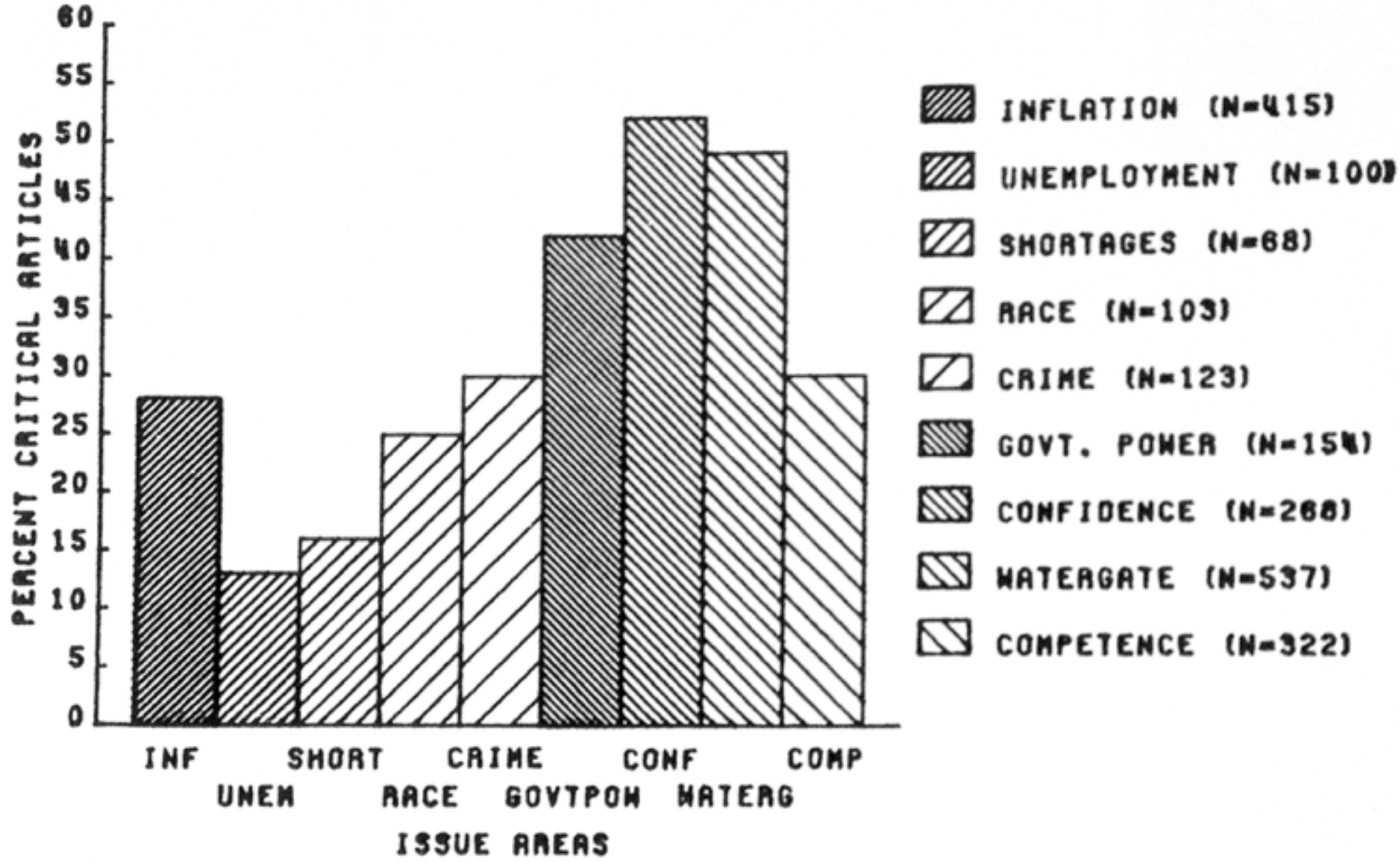


Cosmos: The Infographic Book of Space // Stuart Lowe & Chris North

Plot options are like nuclear weapons…

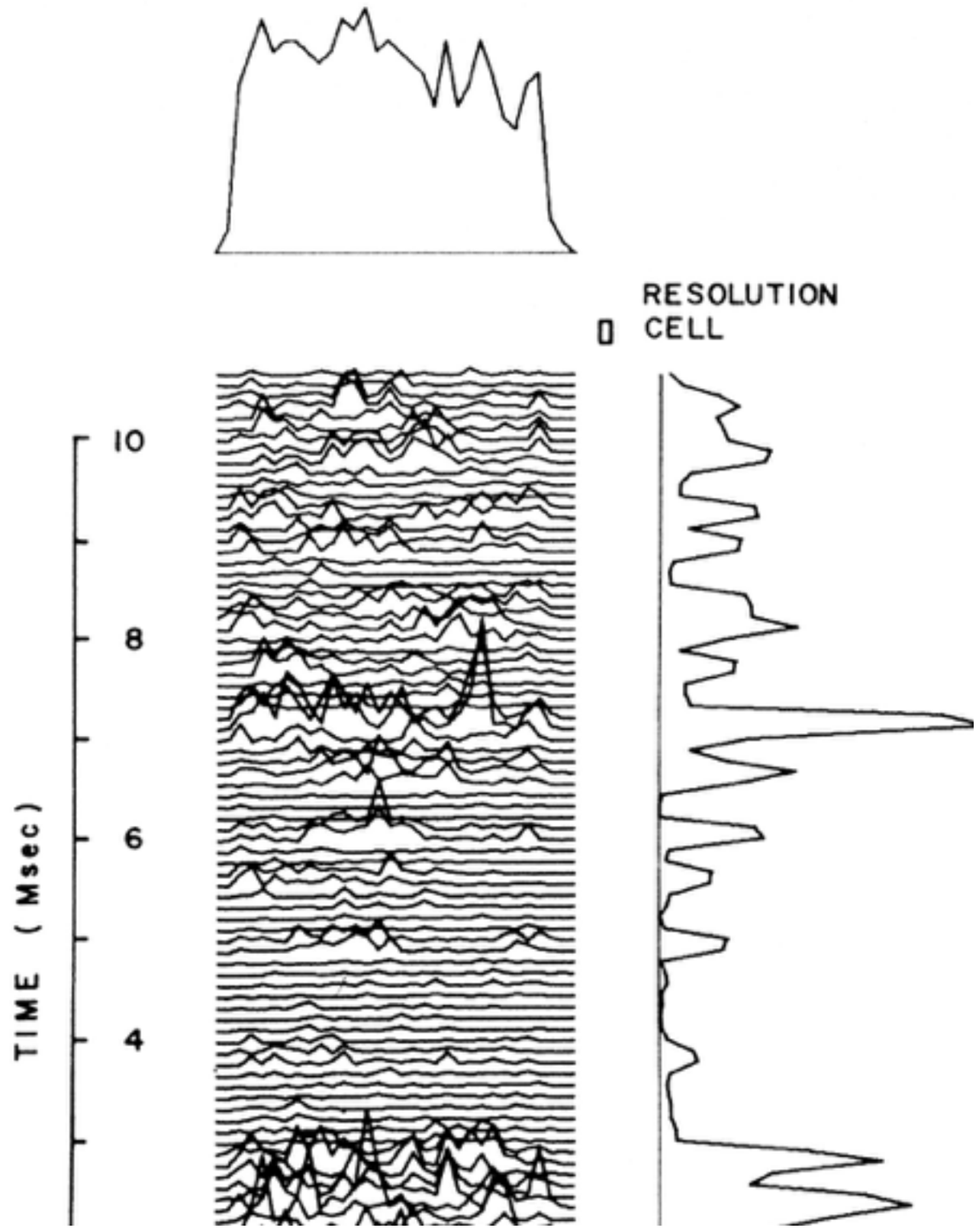
Matthew Kenworthy // Leiden Observatory // 2016

…just because you have them doesn't mean you should use them



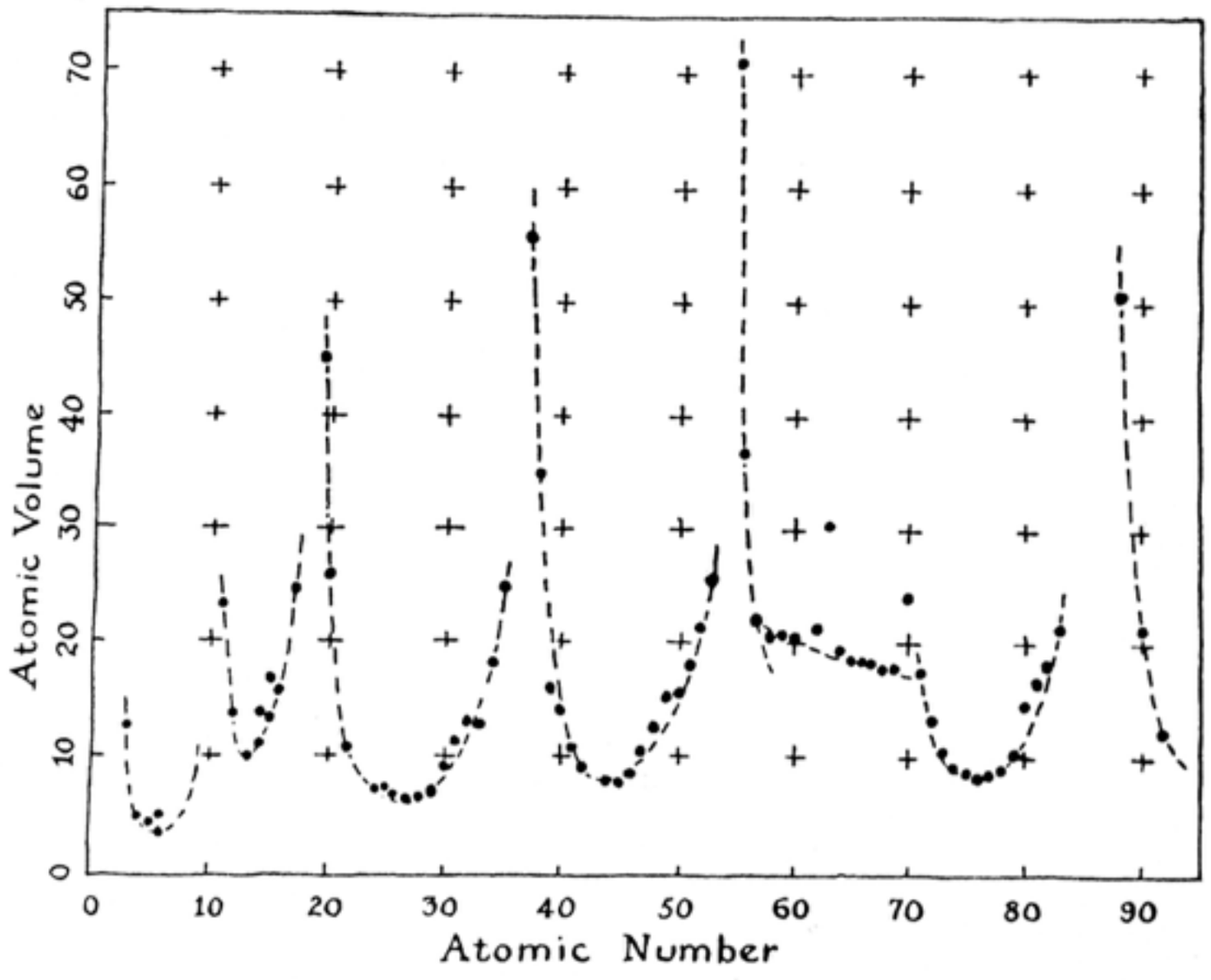
Matthew Kenworthy // Leiden Observatory // 2016

Just the data



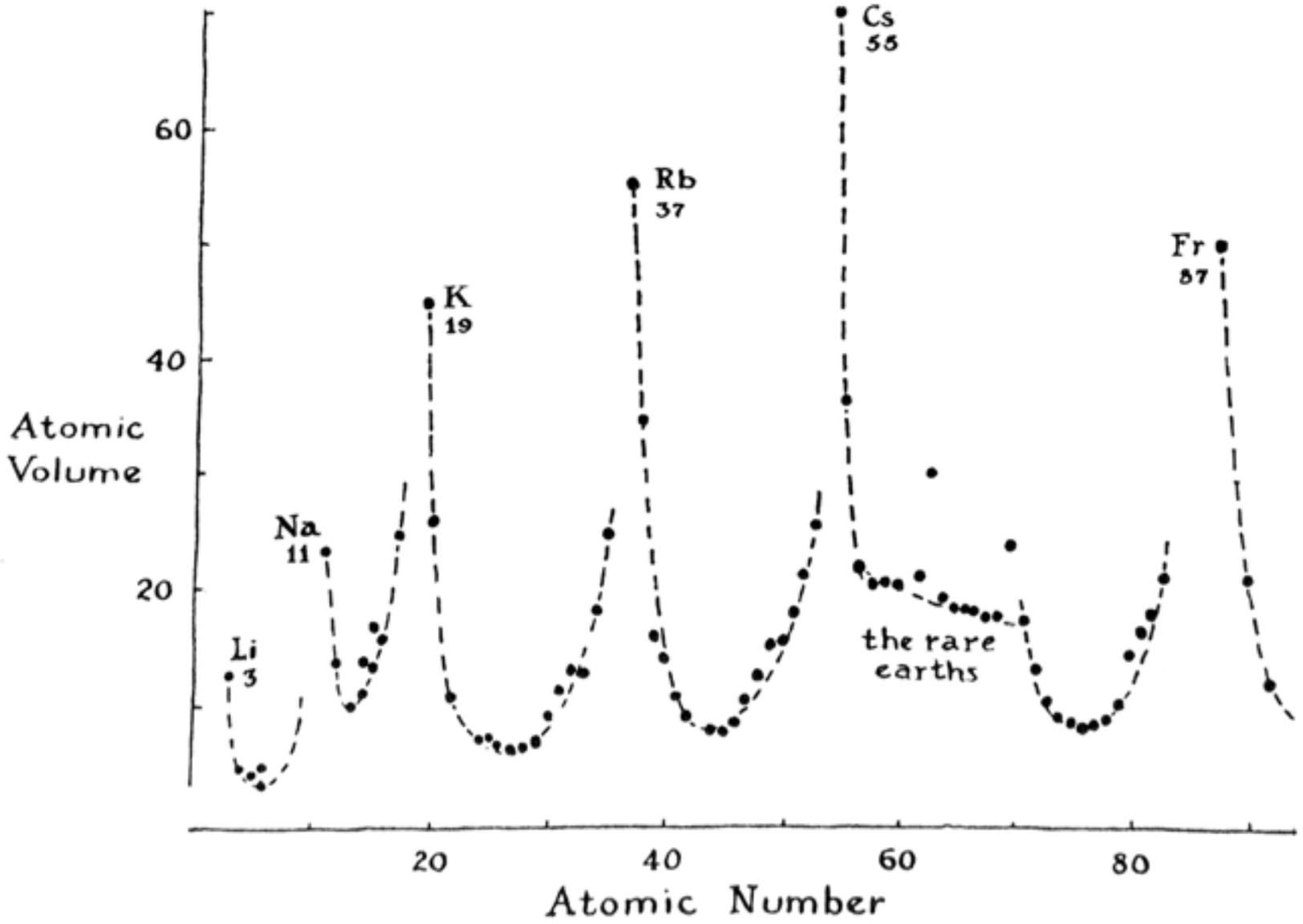
Timothy H. Hankins and Barney J. Rickett, "Pulsar Signal Processing," in Berni Alder, et al., eds., *Methods in Computational Physics, Volume 14: Radio Astronomy* (New York, 1975), p. 108.

Eliminate unnecessary ink



Linus Pauling, *General Chemistry* (San Francisco, 1947), p. 64.

The space opened up by erasing can be effectively used. Labels for the initial elements of each period, an alkali, show the beginning of each cycle in the periodic table of elements—and in the graphic. The unusual rare-earths are indicated. In addition, the label and numbers on the vertical axis are turned to read from left to right rather than bottom to top, making the graphic slightly more accessible, a little more friendly:



Matthew Kenworthy // Leiden Observatory // 2016

the only acceptable pie chart

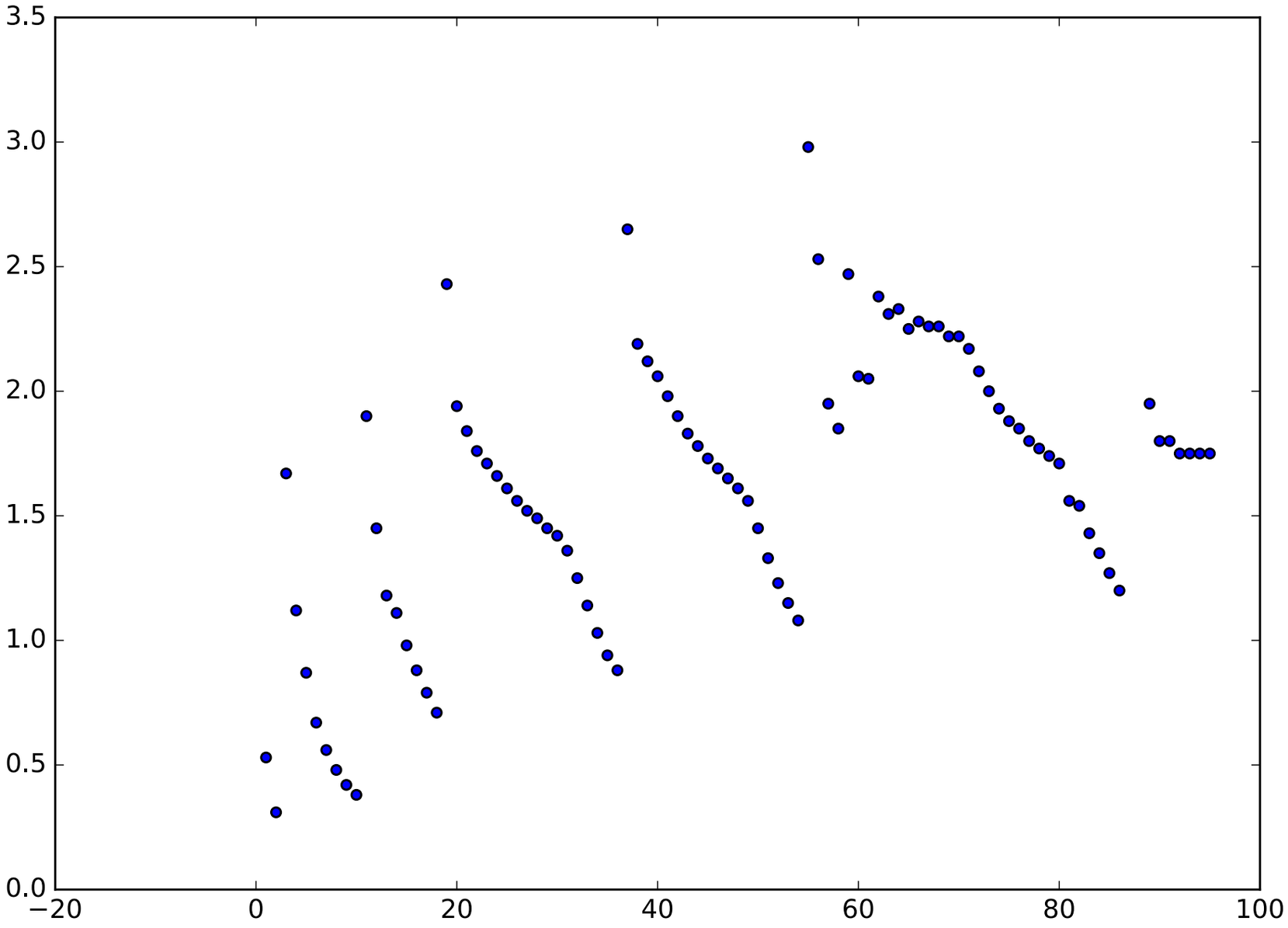
Sky

Sunny side of pyramid

Shady side of pyramid

Matthew Kenworthy // Leiden Observatory // 2016

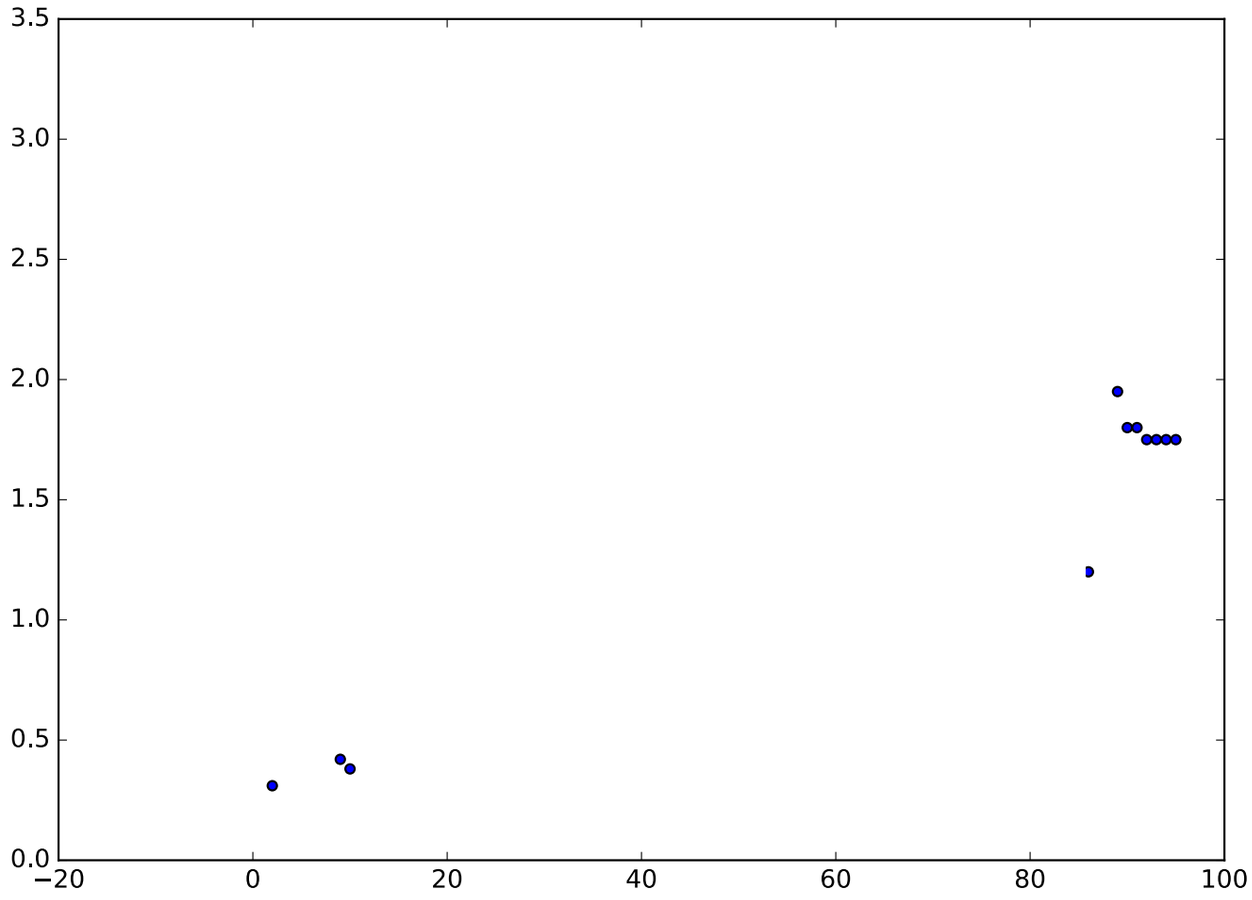
What's that thing on that bit?



Matthew Kenworthy // Leiden Observatory // 2016

The default plot from matplotlib

```
f = plt.figure()
ax = plt.add_subplot(111)
```

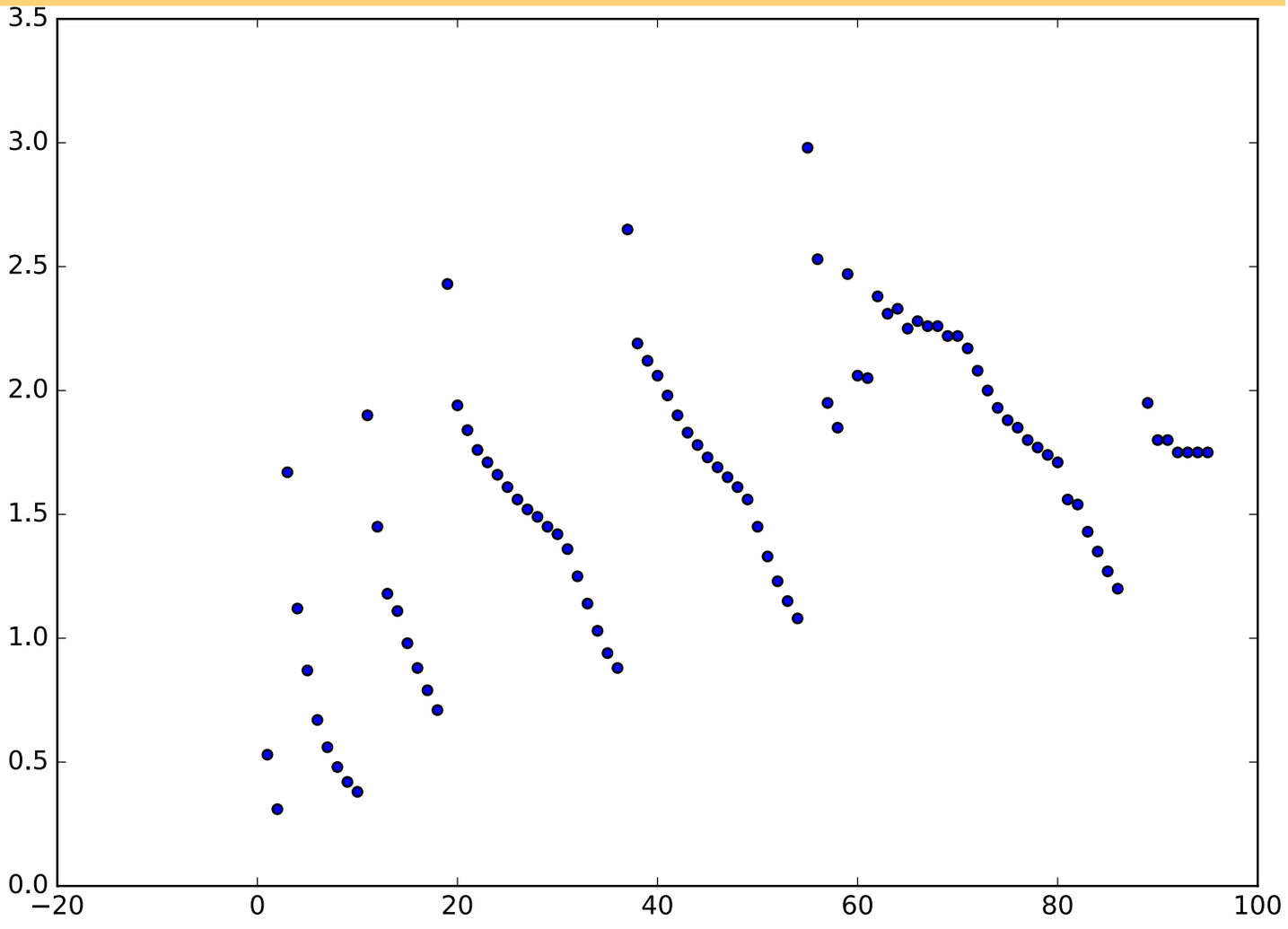


Matthew Kenworthy // Leiden Observatory // 2016

The default plot from matplotlib

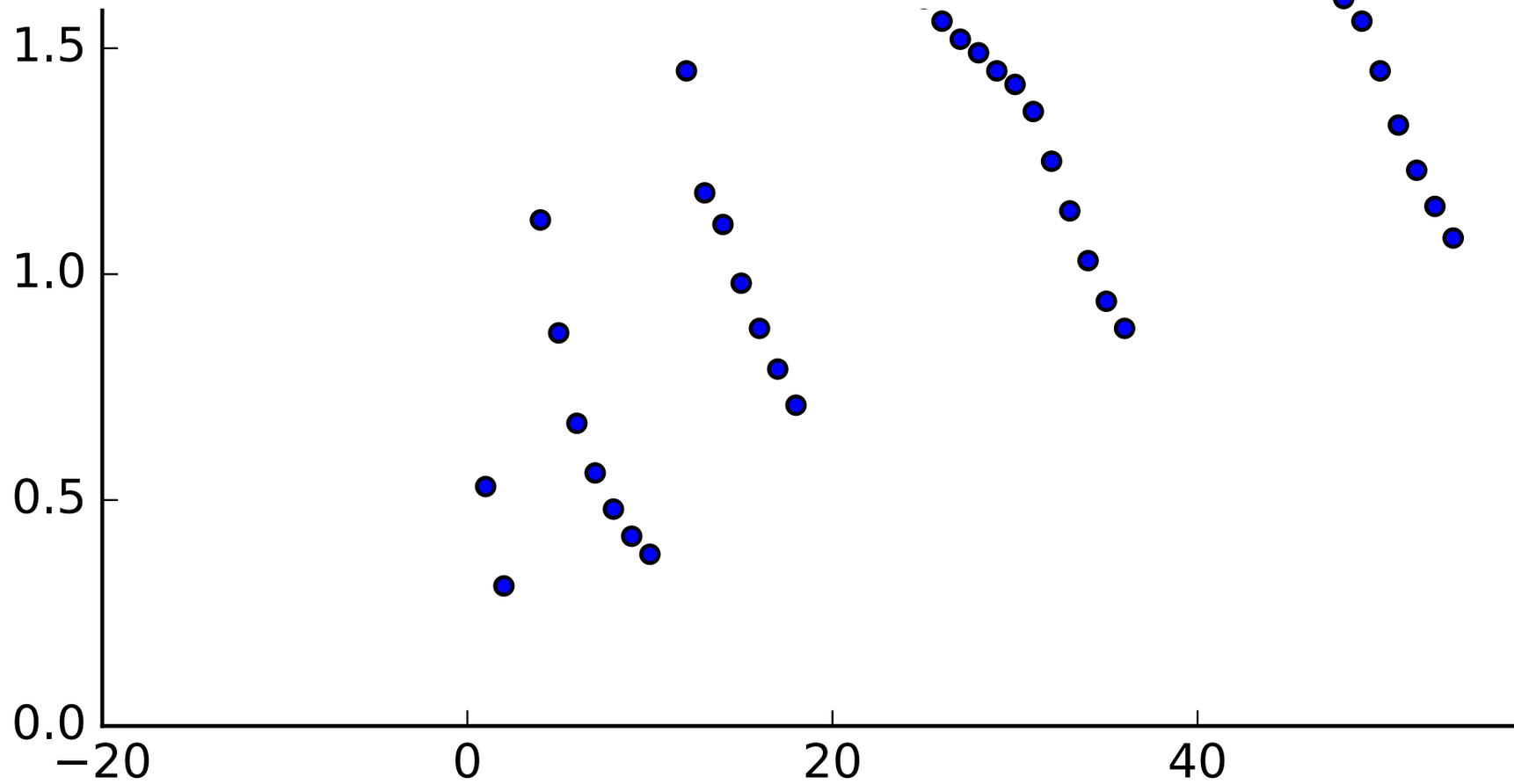
```
f = plt.figure()
ax = plt.add_subplot(111)
lin1 = plt.scatter(x,y)
```

`ax` contains an Axes object which controls the surrounding frame
`lin1` contains the actual points and the lines connecting them



Matthew Kenworthy // Leiden Observatory // 2016

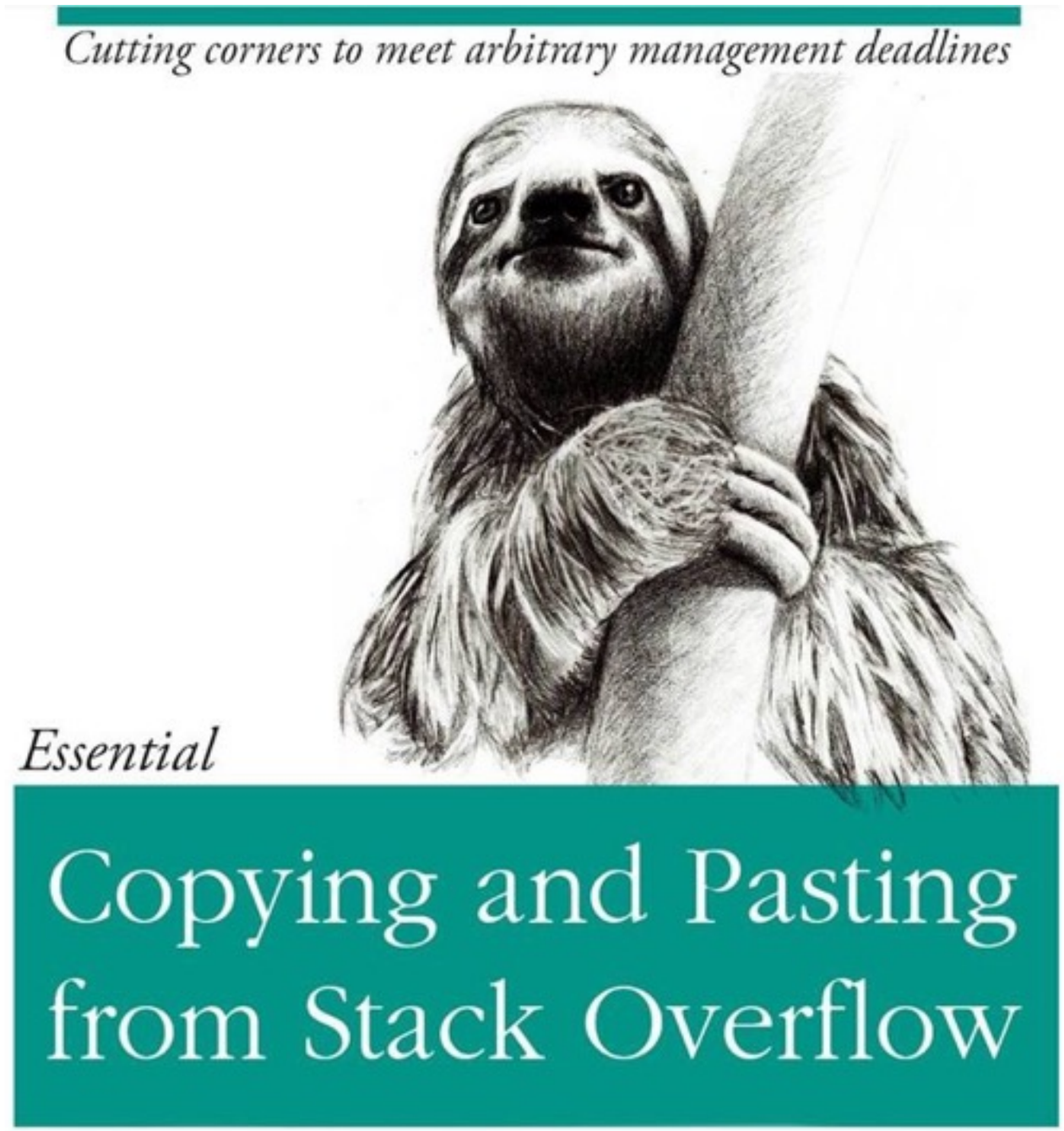
Everything in the plot can be accessed from the `Axes` object



`ax.axes.xaxis`

```
ax.get_xlim() = (-20, 100)
ax.set_xlim(0,65)
```

My very blunt solution to finding out things with matplotlib:



Matthew Kenworthy

O'REILLY®

The Practical Developer
@ThePracticalDev

Recap:

Above all else show the data.

Maximise the data-ink ratio.

Erase non-data-ink.

Erase redundant data-ink.

Revise and edit.

now you have a go!

http://home.strw.leidenuniv.nl/~kenworthy/teaching_better_figures

python notebook Better_Plots_Atomic.ipynb

Making Better Figures

14, 16 and 23 March 2016

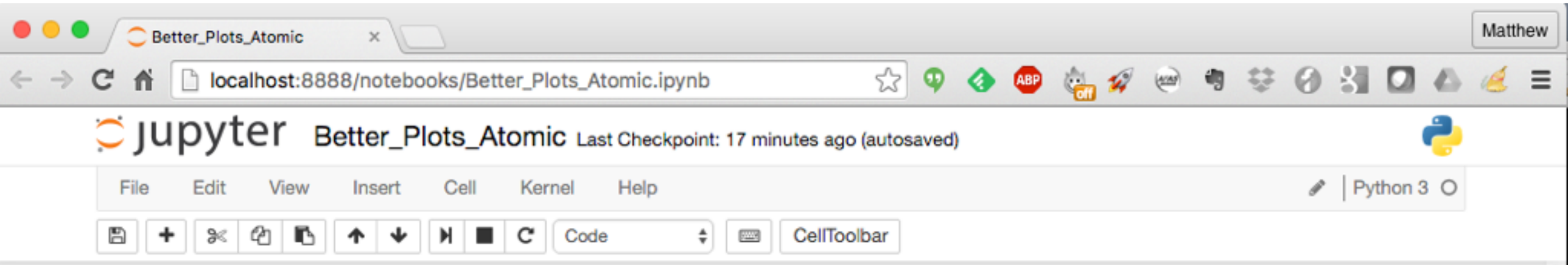
Magnus Persson, Cristobal Sifon, Nadia Murillo, Anna Miotello, Matthew Kenworthy and Jarle Brinchmann

- Download the Python Notebooks for Day 1

now you have a go!

http://home.strw.leidenuniv.nl/~kenworthy/teaching_better_figures

python notebook Better_Plots_Atomic.ipynb



Making Better Plots - Atomic Volume

M Kenworthy 14 March 2016

We'll make a plot of Atomic Volume and then modify it to look similar to the plot shown in the first day slides.

In [10]:
```
# import all the modules we need
import numpy as np
import matplotlib as mpl
import matplotlib.pyplot as plt

%matplotlib inline

from astropy.io import ascii
```

We read in the ascii table of atomic number and atomic radius using astropy:

In [11]:
```
# read in a table of atomic radii
```