Experimental Approaches in Computer Science

Dror Feitelson
Hebrew University

Lecture 2 – Graphs

“Few of us escape being indoctrinated with these notions:

(0) Numerical calculations are exact, but graphs are rough;

(1) For any particular kind of statistical data there is just one set of calculations constituting a correct statistical analysis;

(2) Performing intricate calculations is virtuous, whereas actually looking at the data is cheating."

F. J. Anscombe
The American Statistician **27(1)** Feb 1973

Anscombe's example of 4 datasets:

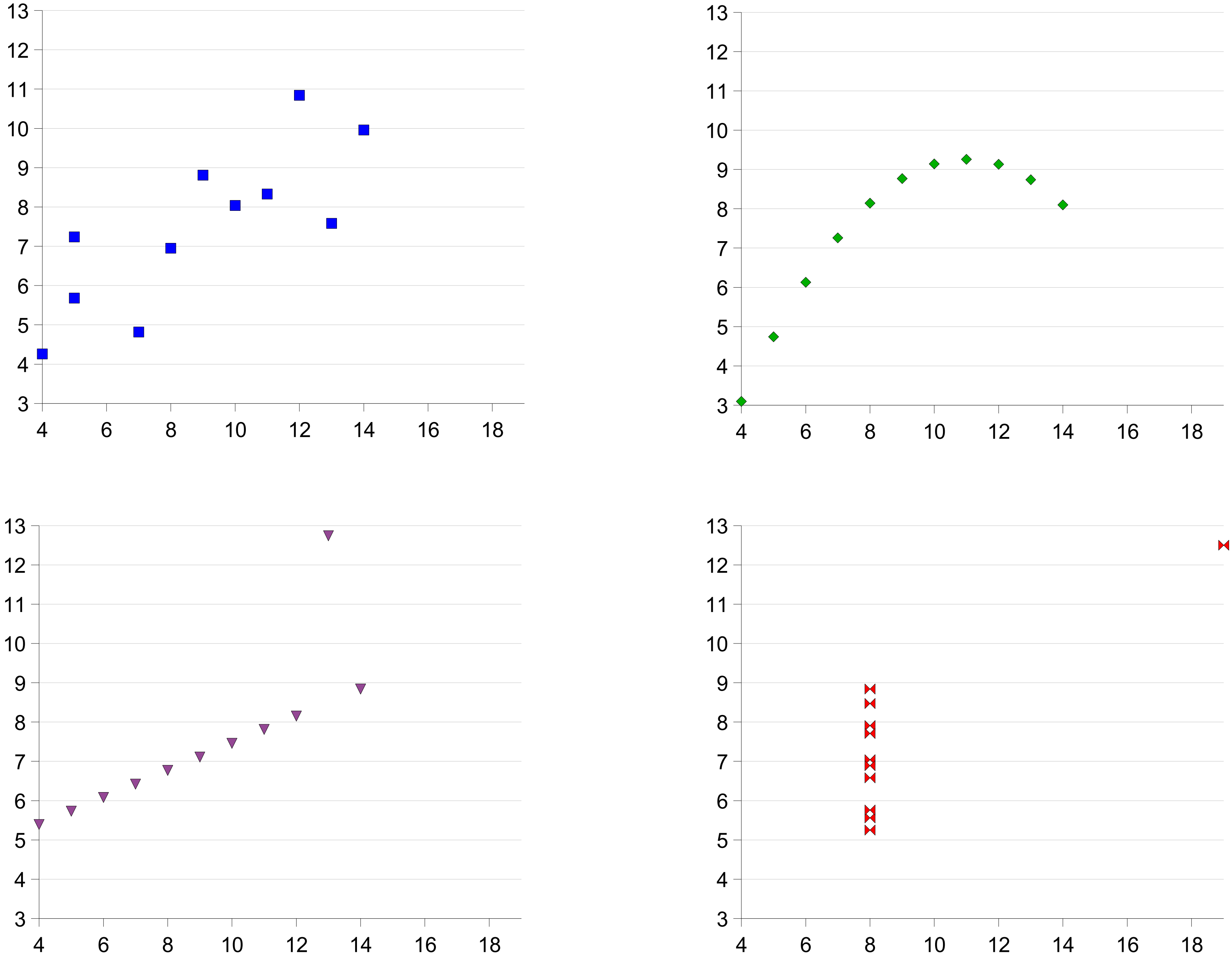
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.1	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	5.39	4	5.39	19	12.5
12	10.84	12	8.15	12	8.15	8	5.56
7	4.82	7	6.42	7	6.42	8	7.91
5	5.68	5	5.73	5	5.73	8	6.89

What can you say about them?

Let's calculate some descriptive statistics for dataset #1:

number of observations:	11
mean of x:	9.0
mean of y:	7.5
linear regression:	$y = 3 + 0.5x$
R^2:	0.667
correlation coefficient:	0.82
sum of squares of *x-avg(x)*:	110.0
regression sum of squares:	27.5
residual sum of squares of y:	13.75
estimated std. error of slope:	0.118

For the other data sets we get the same results!!!
so they are all similar, right?


13
12
11
10
9
8
7
6
5
4
3
4
6
8
10
12
14
16
18

Conclusion:

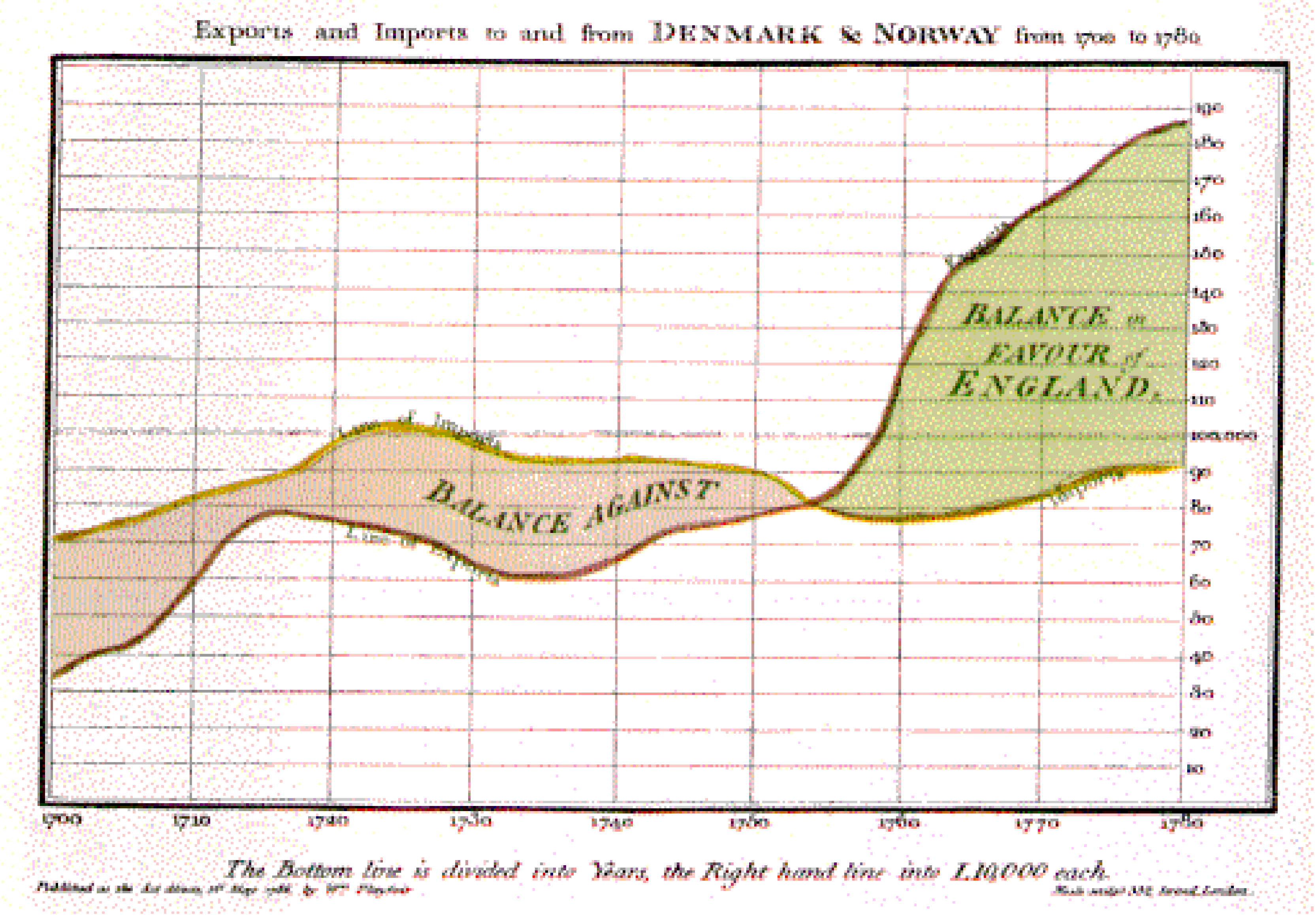
look at the data!

- Discover what the data has to say
  John W. Tukey, *Exploratory Data Analysis*, Addison-Wesley, 1977
- Display your conclusions in the most convincing manner

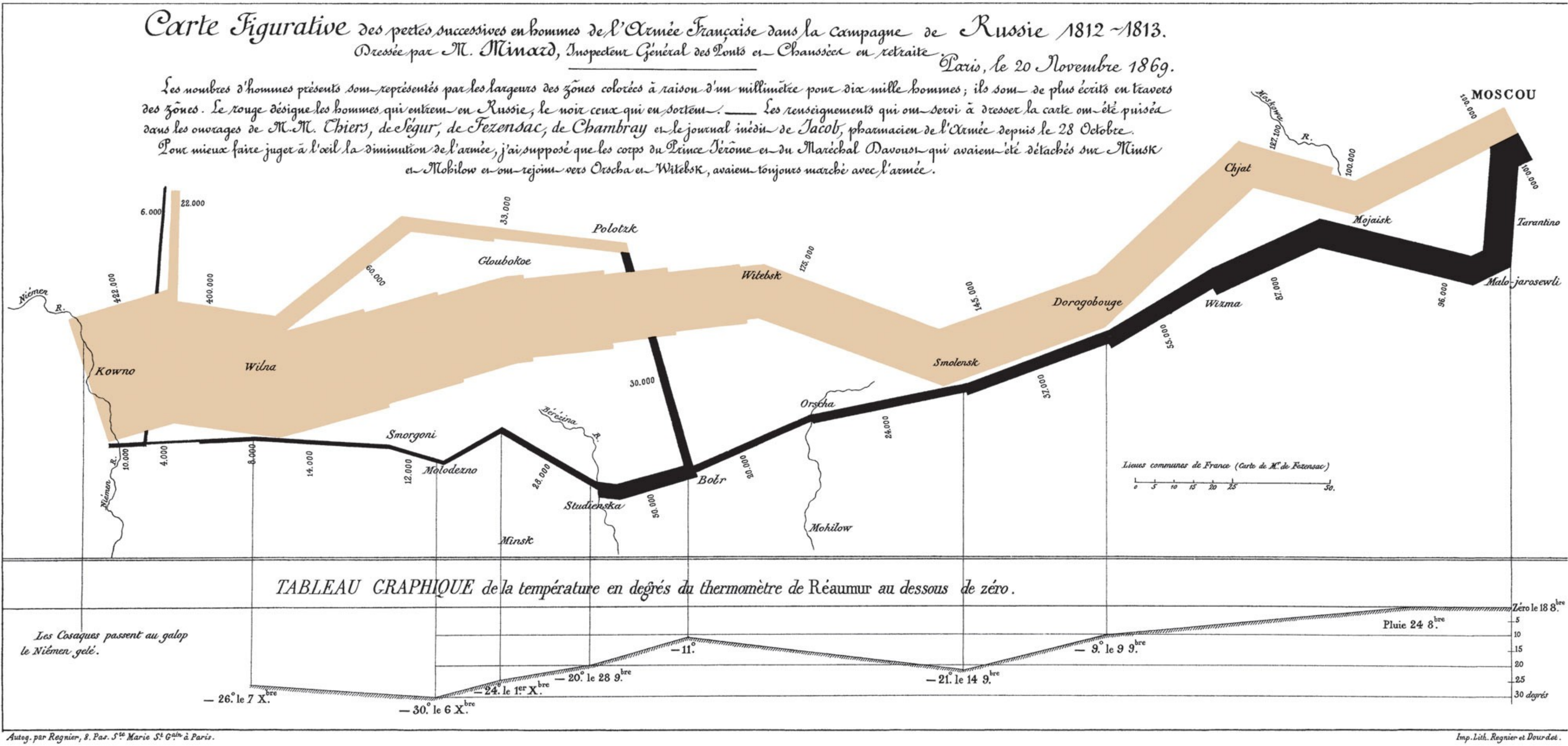
Graphs that made history or illuminate data

Michael Friendly's Gallery of data visulaization

William Playfair, *The Commercial and Political Atlas*, 1786: invented most graphs used today



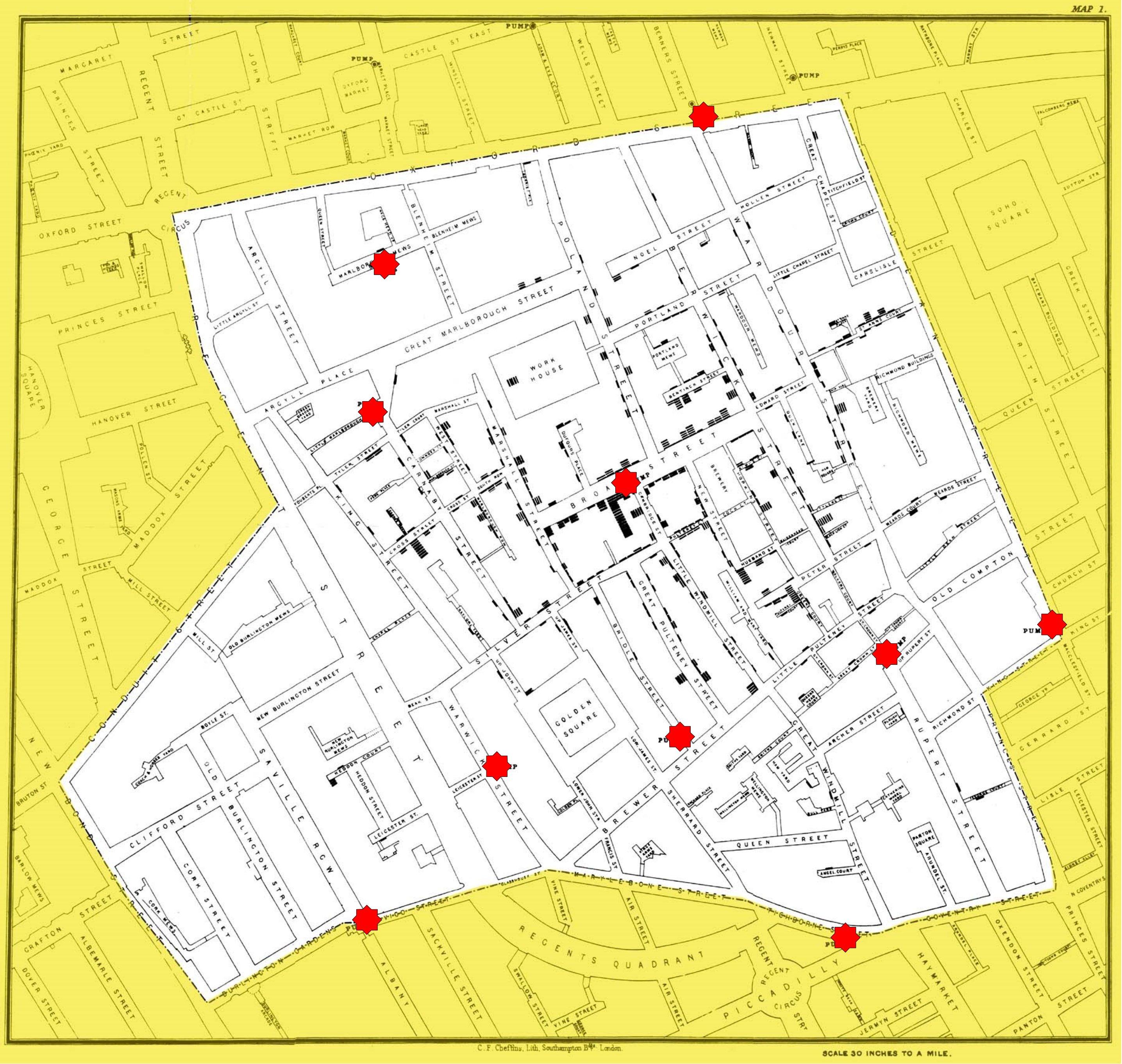
Charles Minard, plot of Napoleon's failed campaign in Russia, 1812



Popularized by Tufte as the best graphic ever

John Snow, deaths in London Cholera epidemic, 1854

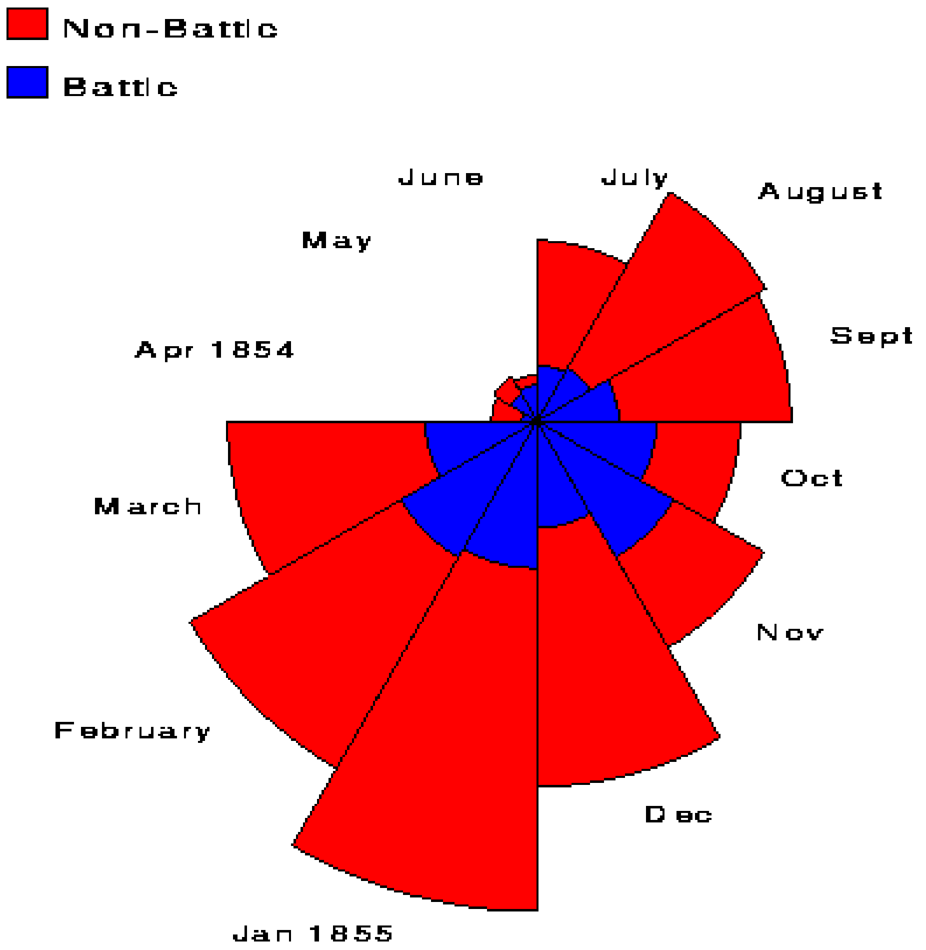
Established link between water quality and health



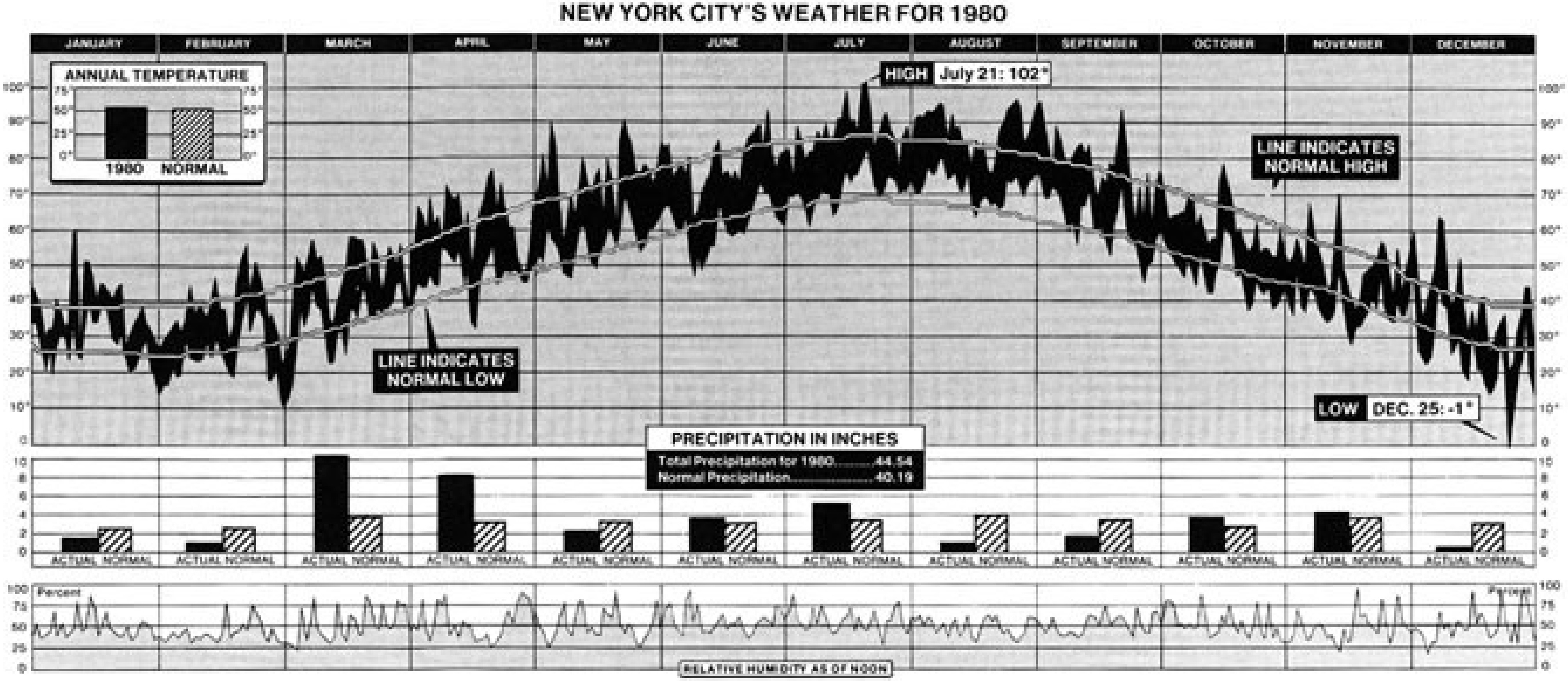
Florence Nightingale, British casualties in Crimean war, 1858

Established sanitation as a decisive factor in hospital operation

Causes of Mortality in the Army in the East
April, 1854 to March 1855

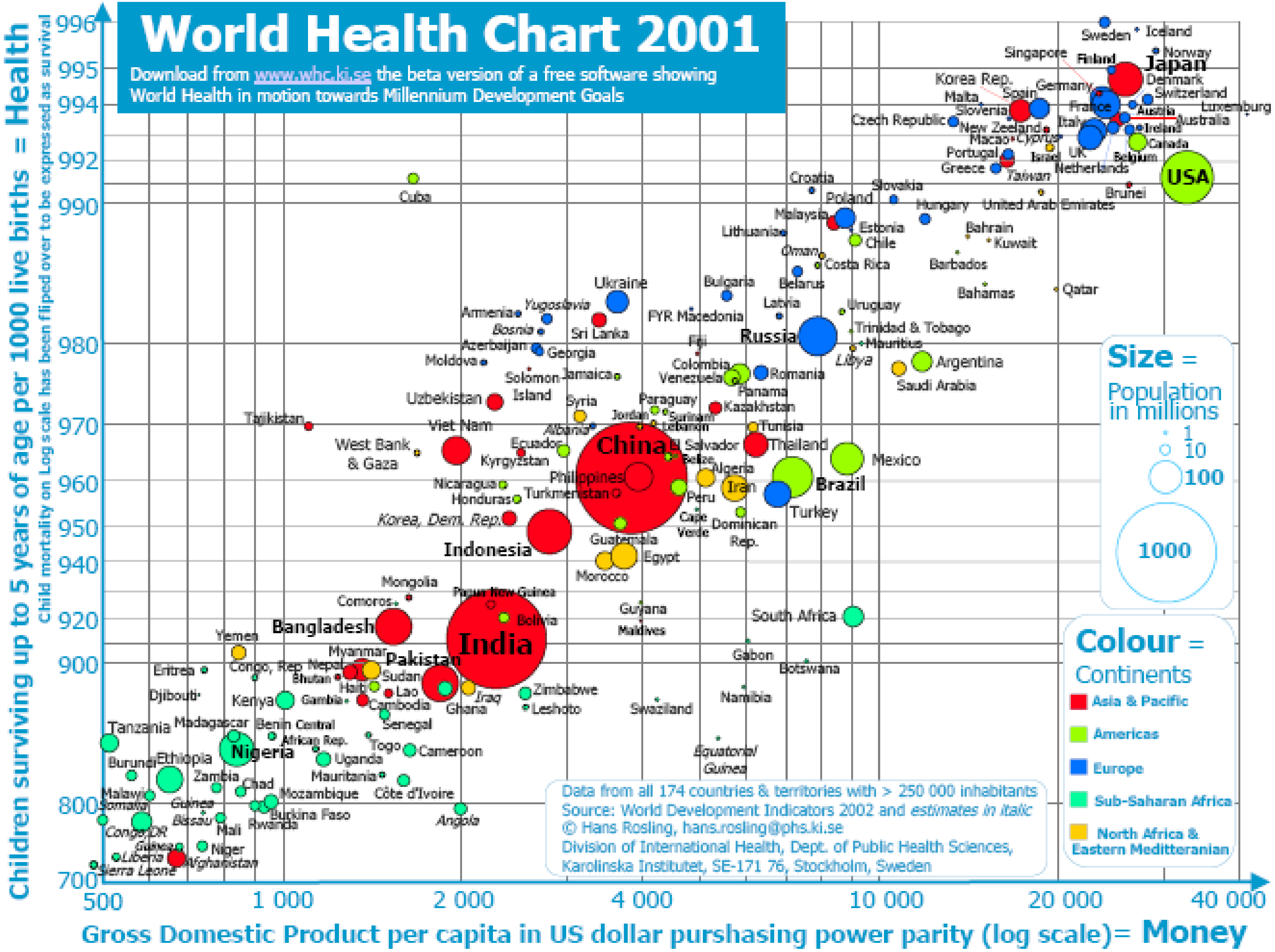


From: F. Nightingale, "Notes on Matters Affecting the Health, Efficiency and Hospital Administration of the British Army", 1858



Summary of a whole year's weather

- Lots of numbers (daily max/min + average max/min + humidity)
- Use of parallel graphs for correlation
- Callouts to emphasize special points



World Health Chart 2001

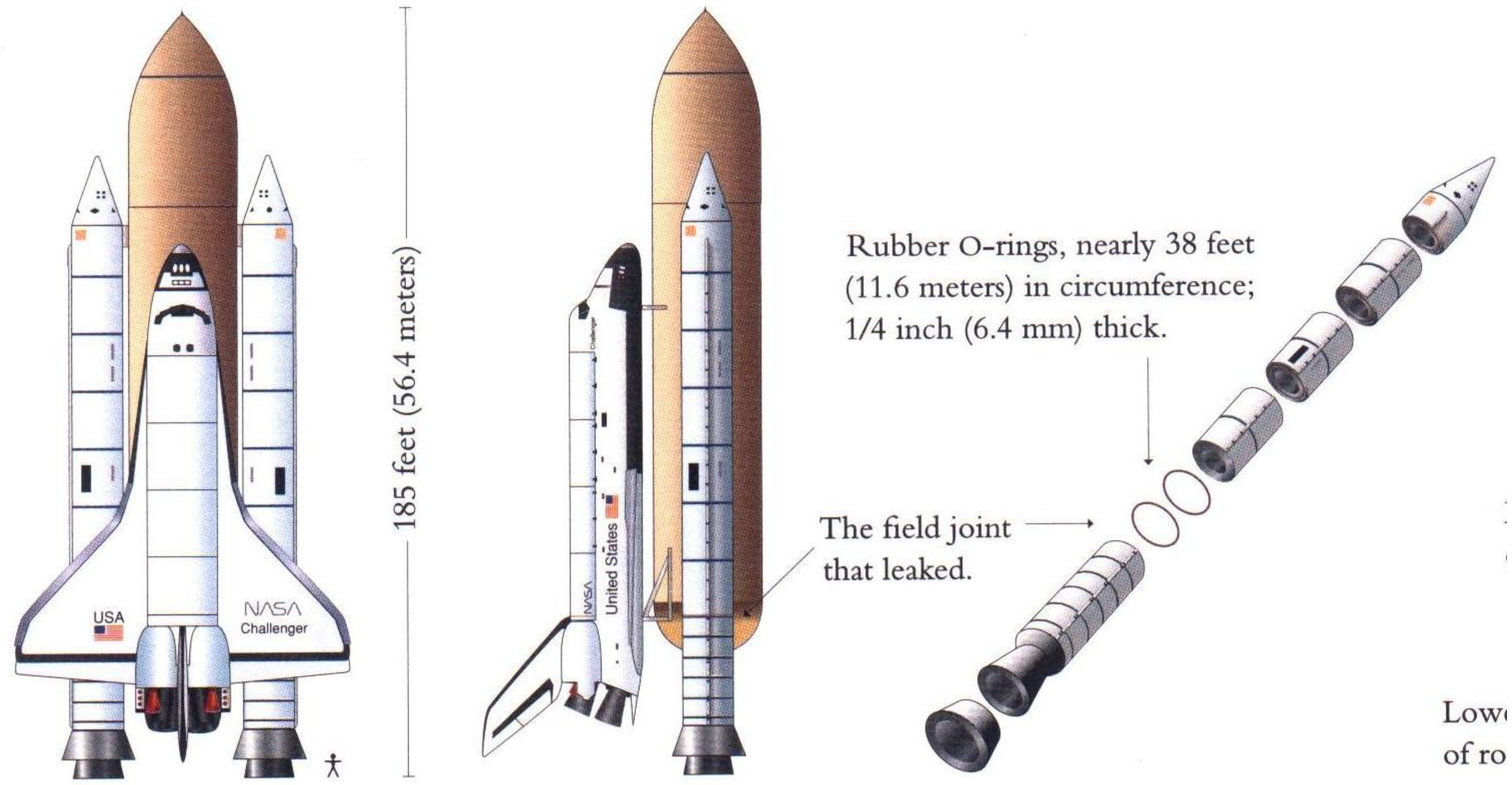
Download from www.whc.ki.se the beta version of a free software showing World Health in motion towards Millennium Development Goals

Data from all 174 countries & territories with > 250 000 inhabitants
Source: World Development Indicators 2002 and *estimates in italic*
© Hans Rosling, hans.rosling@phs.ki.se
Division of International Health, Dept. of Public Health Sciences,
Karolinska Institutet, SE-171 76, Stockholm, Sweden

The harm of bad graphics

Tufte, *Visual Explanations*
using graphs by Morton Thiokol, Inc.

Background: launch of the Challenger space shuttle on 27 January 1986, amid concerns regarding O-ring function in cold weather

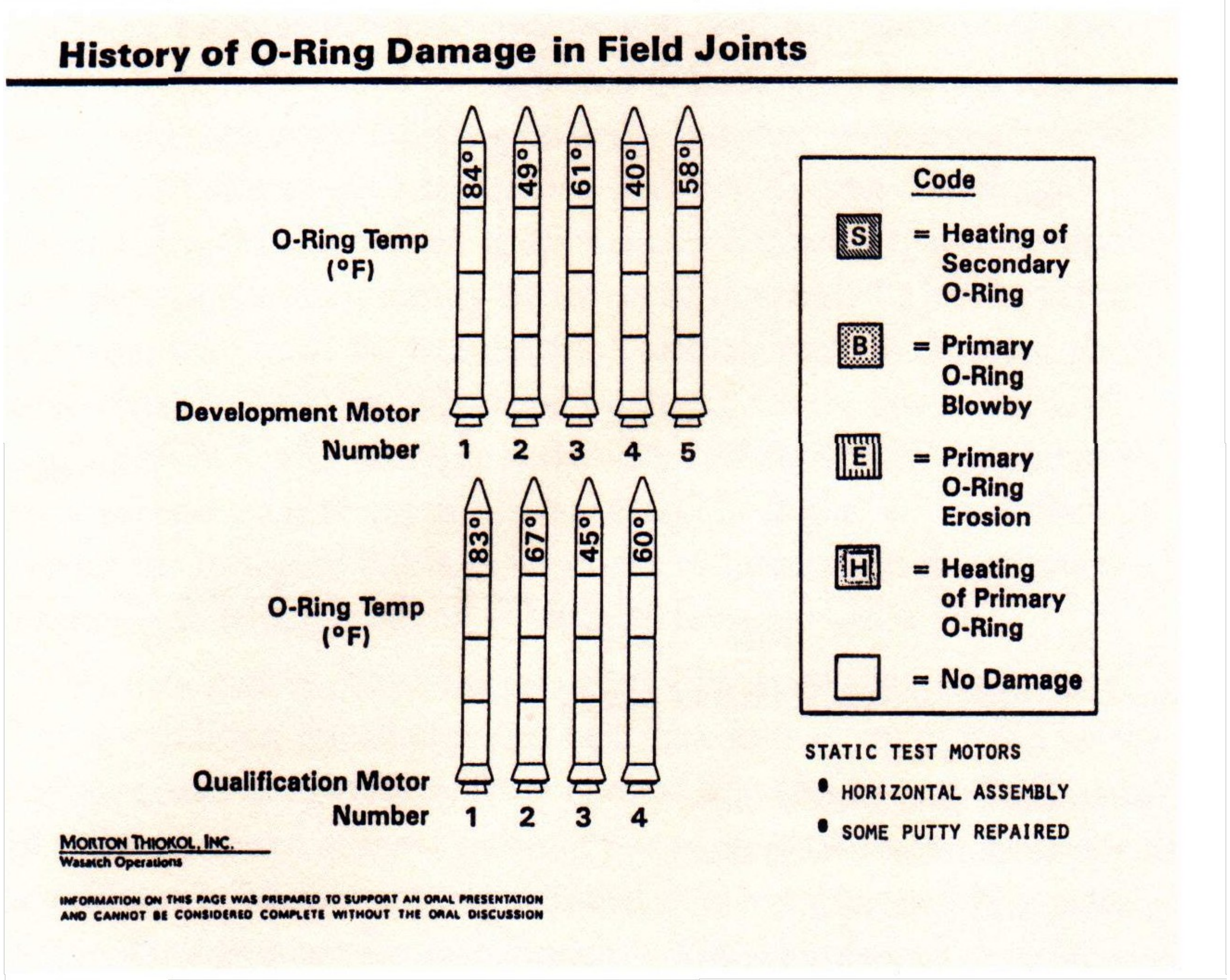


Rubber O-rings, nearly 38 feet (11.6 meters) in circumference; 1/4 inch (6.4 mm) thick.

The field joint that leaked.

Data regarding test rockets from the manufacturer

(chart prepared later; charts used in discussions prior to the launch contained less data)



History of O-Ring Damage in Field Joints

Code

S = Heating of Secondary O-Ring

B = Primary O-Ring Blowby

E = Primary O-Ring Erosion

H = Heating of Primary O-Ring

□ = No Damage

STATIC TEST MOTORS

- HORIZONTAL ASSEMBLY
- SOME PUTTY REPAIRED

MORTON THIOKOL, INC.
Wasatch Operations

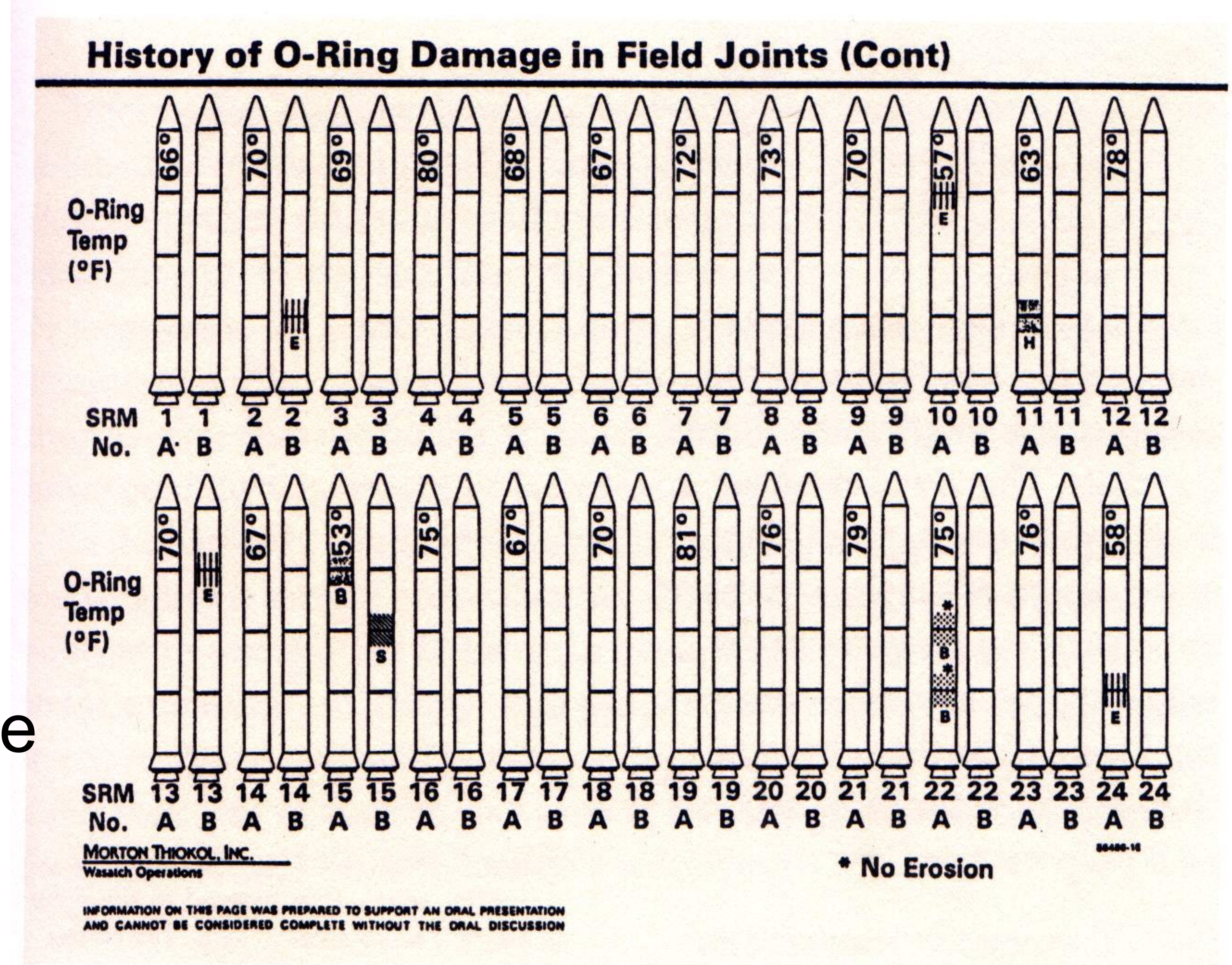
INFORMATION ON THIS PAGE WAS PREPARED TO SUPPORT AN ORAL PRESENTATION AND CANNOT BE CONSIDERED COMPLETE WITHOUT THE ORAL DISCUSSION

Data regarding prior launches

Note that legend is missing (appeared previously)

Contains too much irrelevant data

Does not clarify effect of temperature



History of O-Ring Damage in Field Joints (Cont)

SRM No.	O-Ring Temp (°F)	
1 A	66°	
1 B		
2 A	70°	
2 B		E
3 A	69°	
3 B		
4 A	80°	
4 B		
5 A	68°	
5 B		
6 A	67°	
6 B		
7 A	72°	
7 B		
8 A	73°	
8 B		
9 A	70°	
9 B		
10 A	57°	E
10 B		
11 A	63°	H
11 B		
12 A	78°	
12 B		
13 A	70°	
13 B		E
14 A	67°	
14 B		
15 A	53°	B
15 B		S
16 A	75°	
16 B		
17 A	67°	
17 B		
18 A	70°	
18 B		
19 A	81°	
19 B		
20 A	76°	
20 B		
21 A	79°	
21 B		
22 A	75°	* B * B
22 B		
23 A	76°	
23 B		
24 A	58°	E
24 B		

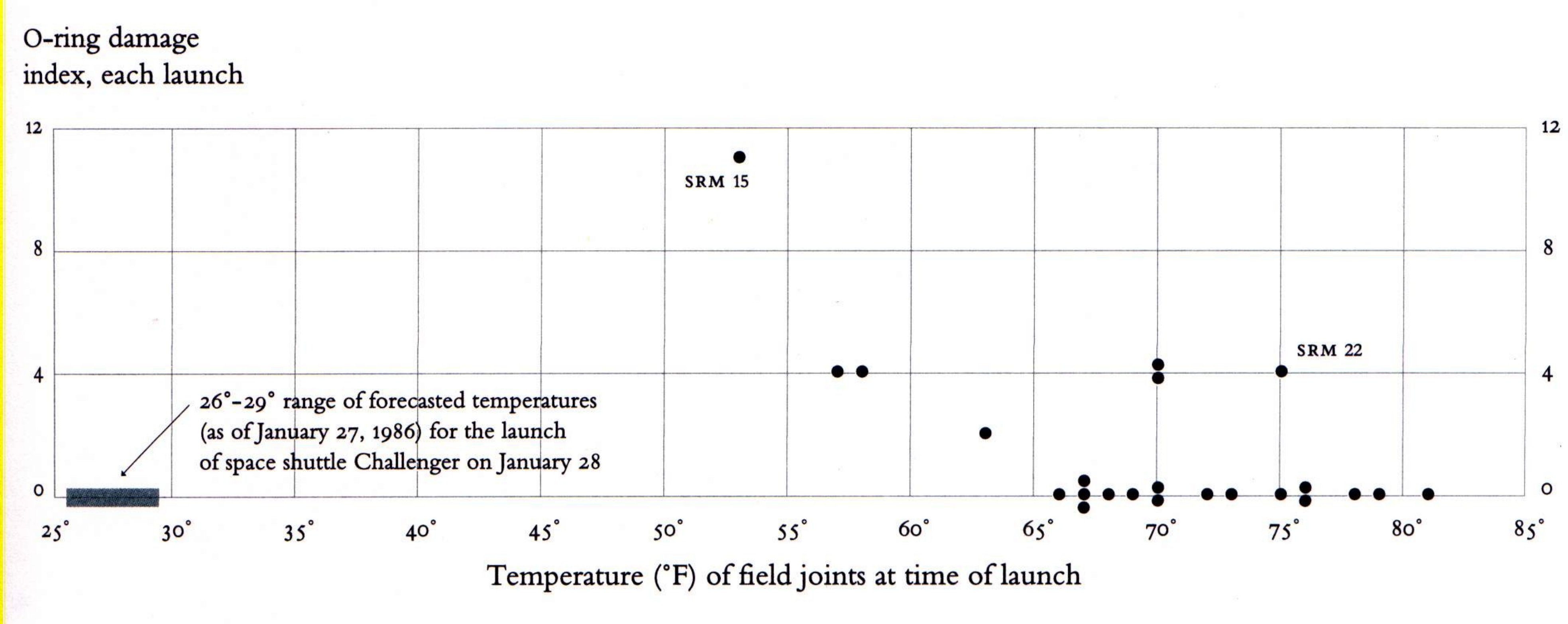
MORTON THIOKOL, INC.
Wasatch Operations

* No Erosion

86486-16

INFORMATION ON THIS PAGE WAS PREPARED TO SUPPORT AN ORAL PRESENTATION AND CANNOT BE CONSIDERED COMPLETE WITHOUT THE ORAL DISCUSSION

Tufte's alternative rendering of the data

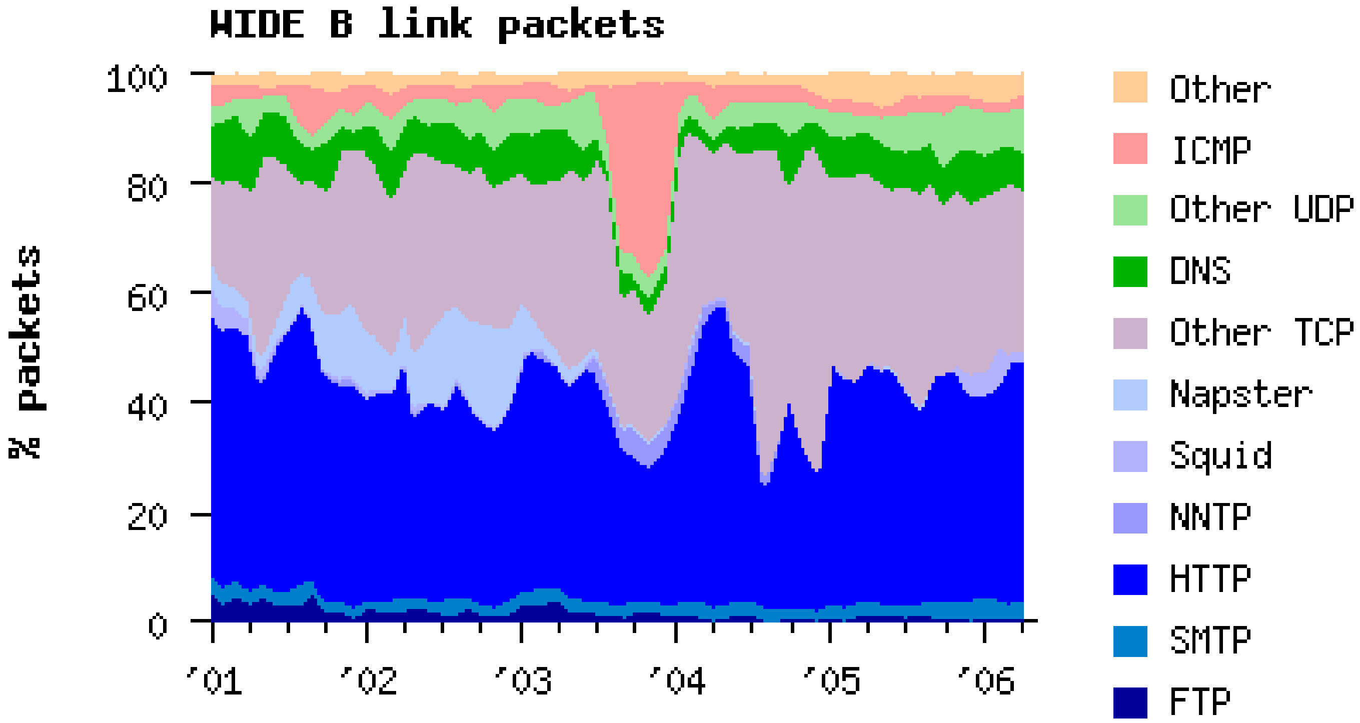


Note exaggerated X scale for emphasis

Examples

A graph should be independent and provide full information

- Title (if relevant)
- Axis labels (including units)
- Tics indicating values
- Legend

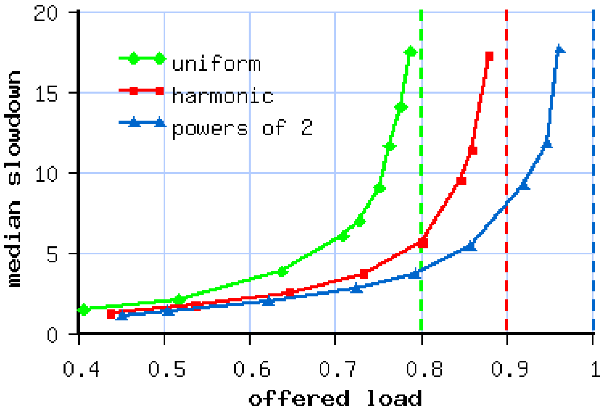


Need to also consider aesthetics

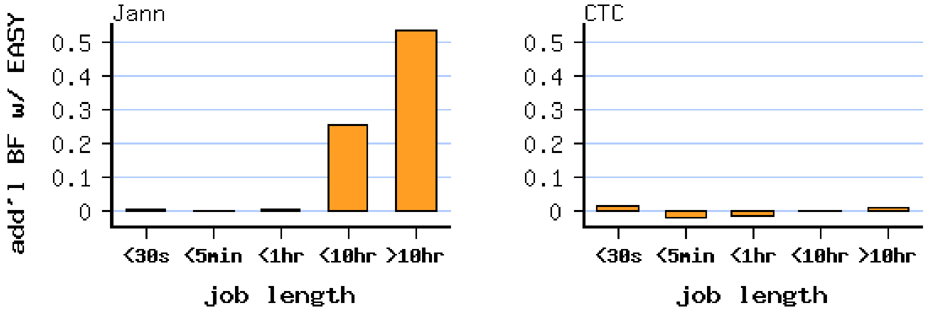
- Proportions
  - Size and placement of labels and legend
  - Size of fonts relative to graphical elements
- Use of color
  - Express gradient with deeper shades
  - Create focus for discussion
  - Should also work in black and white
- Combination of graphical elements
  - Give full picture
  - Connections through consistent use of colors
- Order in legend matches order of graphs

Causal/functional relationship:

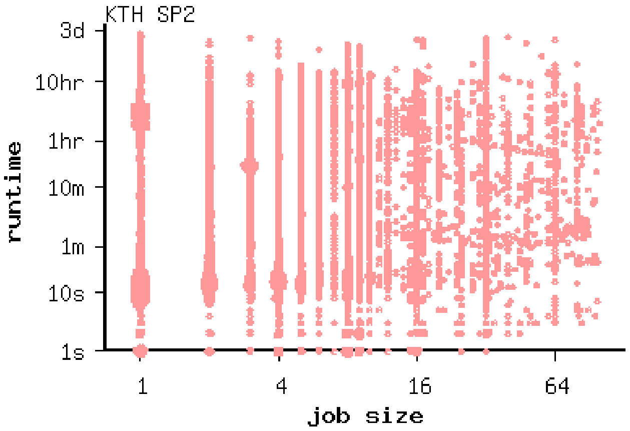
- XY plot (continuous)



- bar chart (categories)

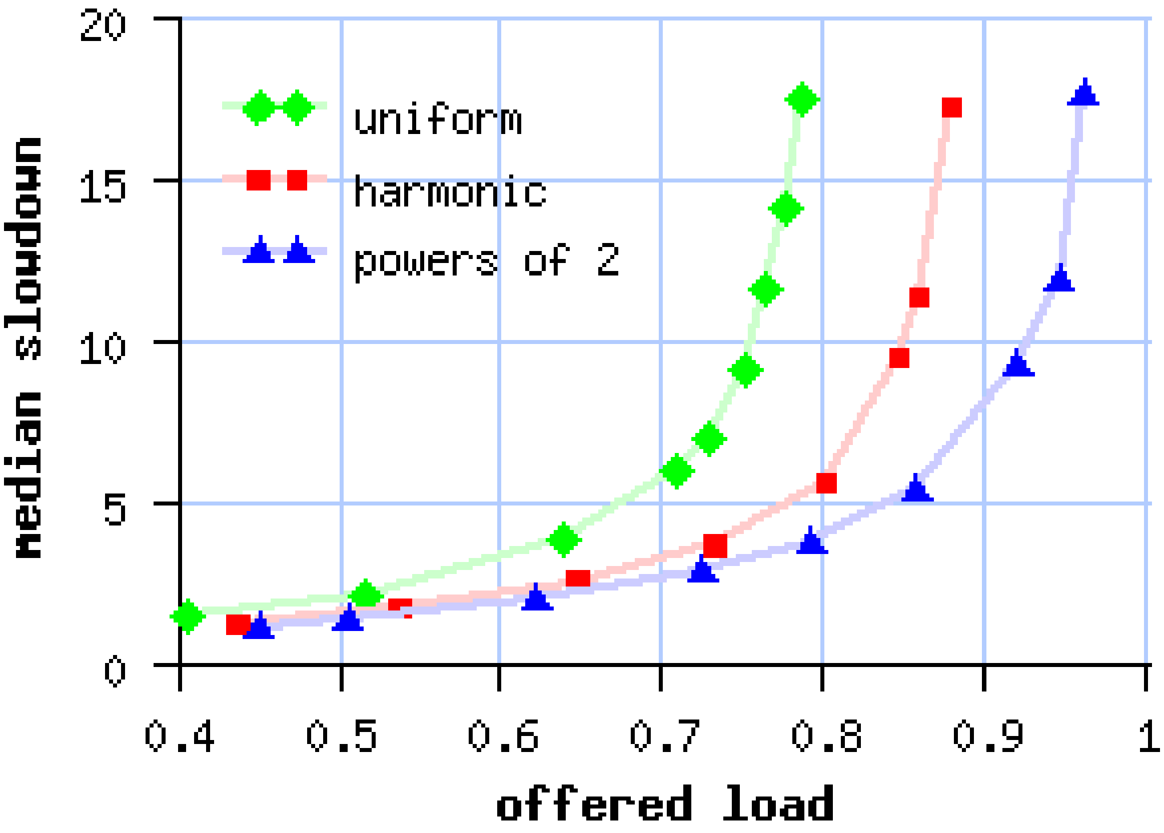


- scatter plot (complicated)



Showing measurements

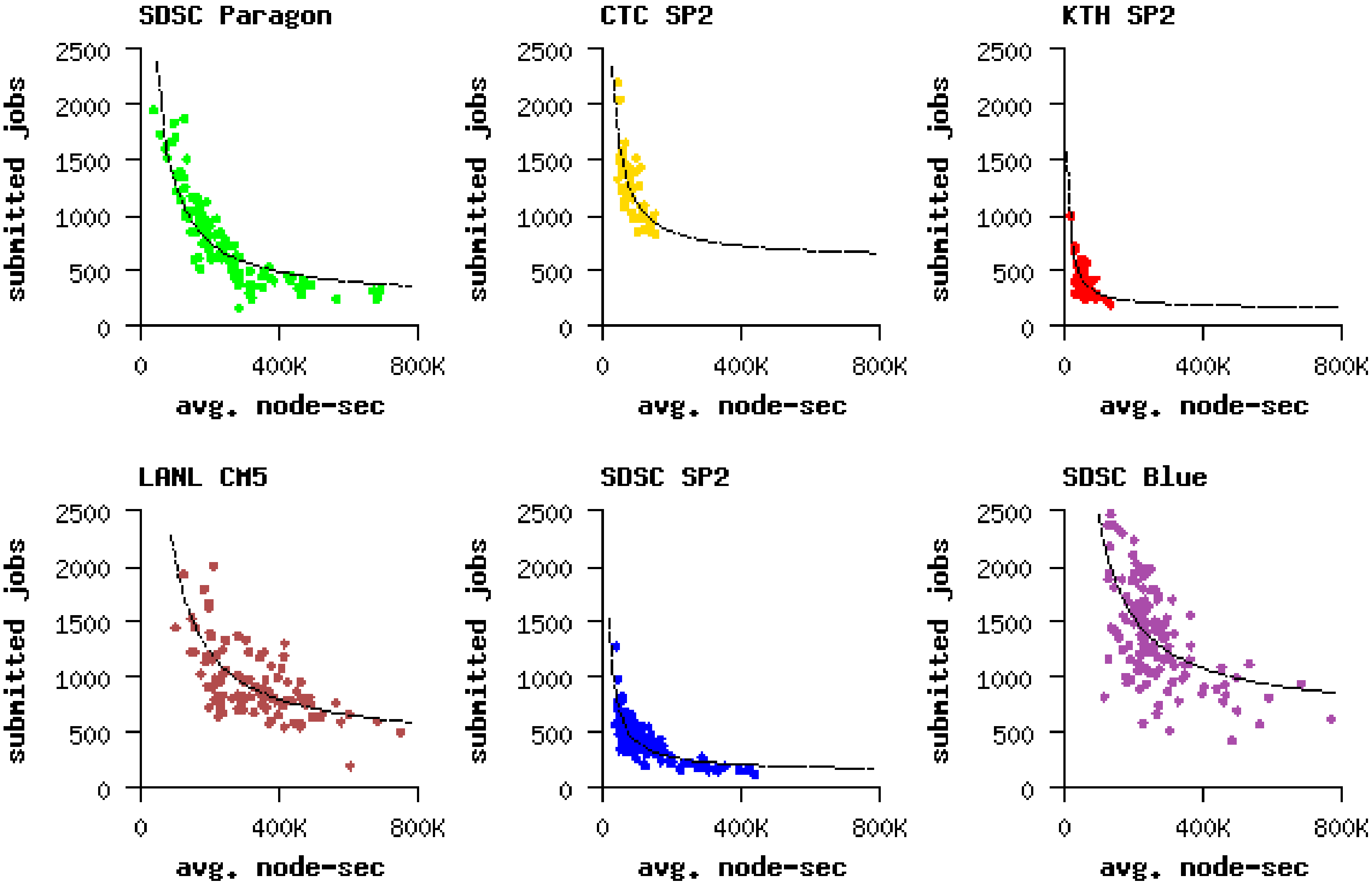
- Emphasize points
- Connect with weaker lines



Or show fitted model line

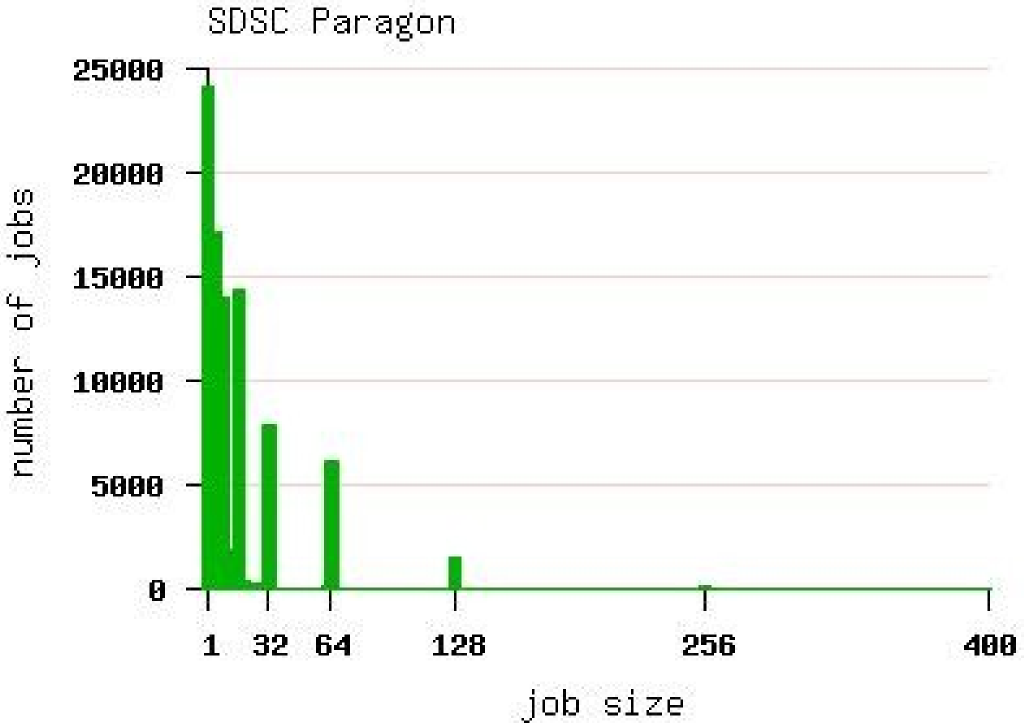
here model is y~1/x

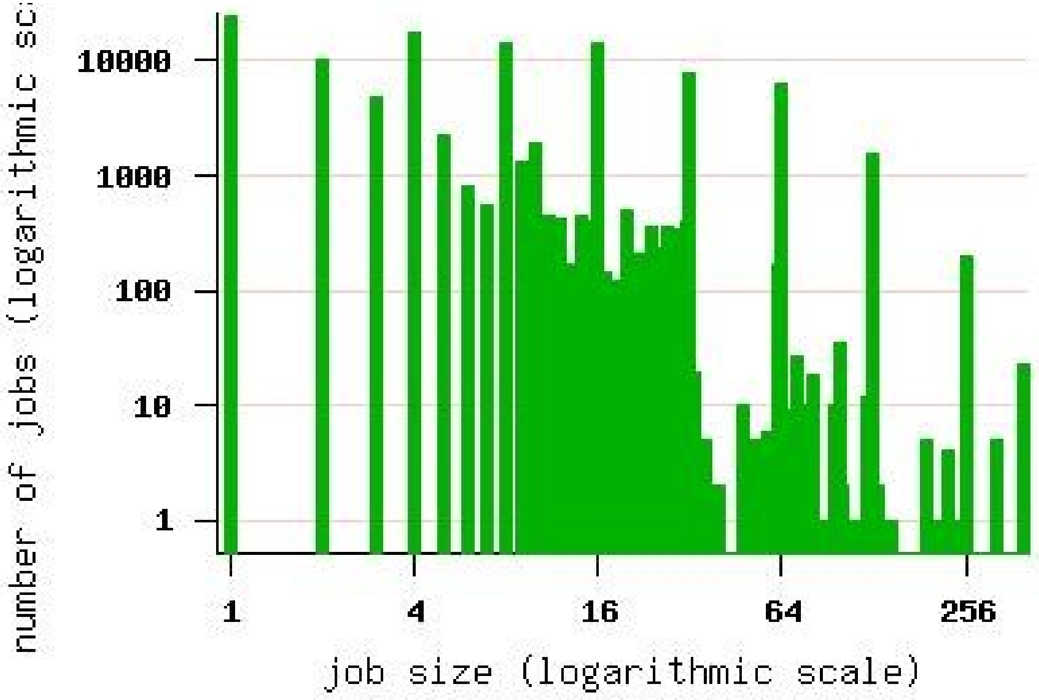
often linear regression



Scales

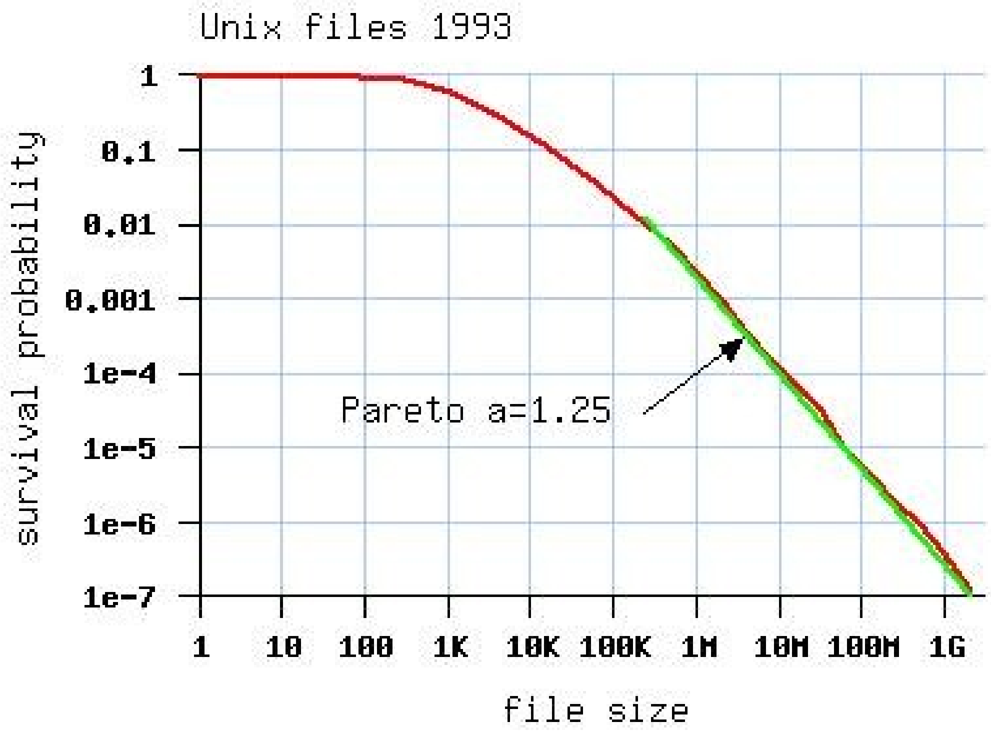
- Linear is best
- Logarithmic if needed

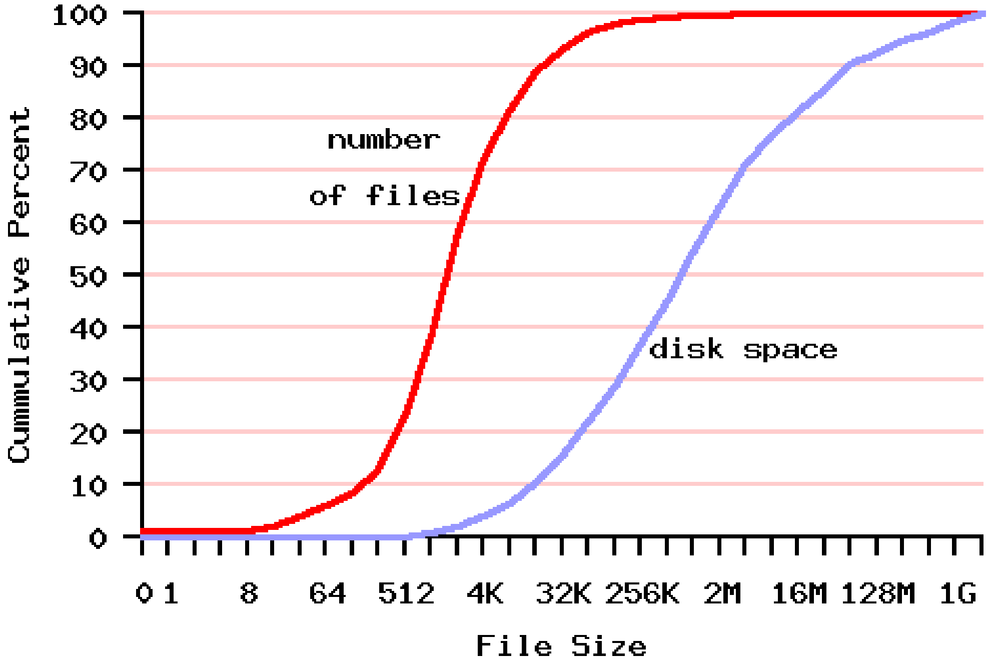




Scales

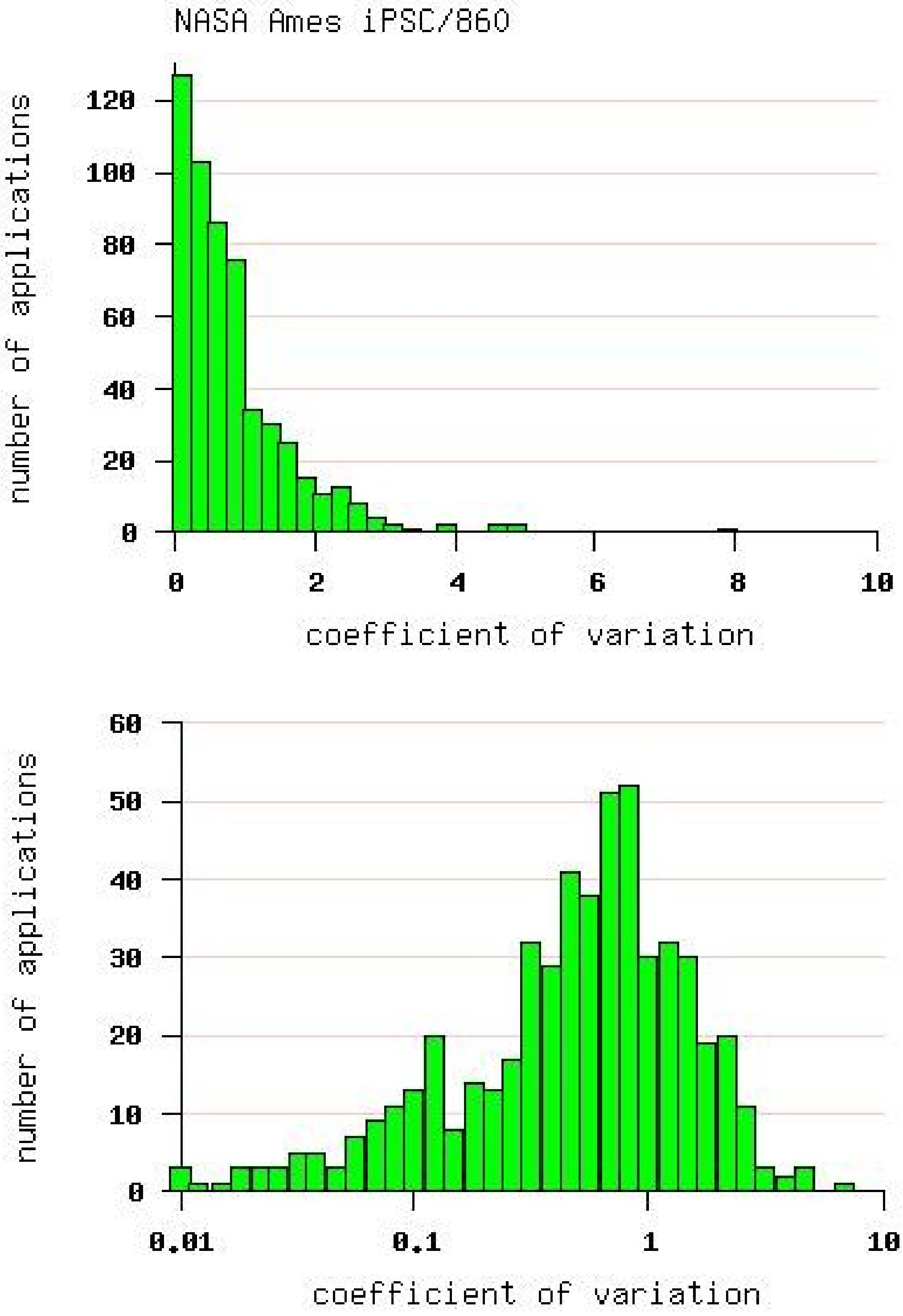
- Logarithmic if needed
- Show values, not their log, in stubs





Log scale

- Beware of expansion in small values



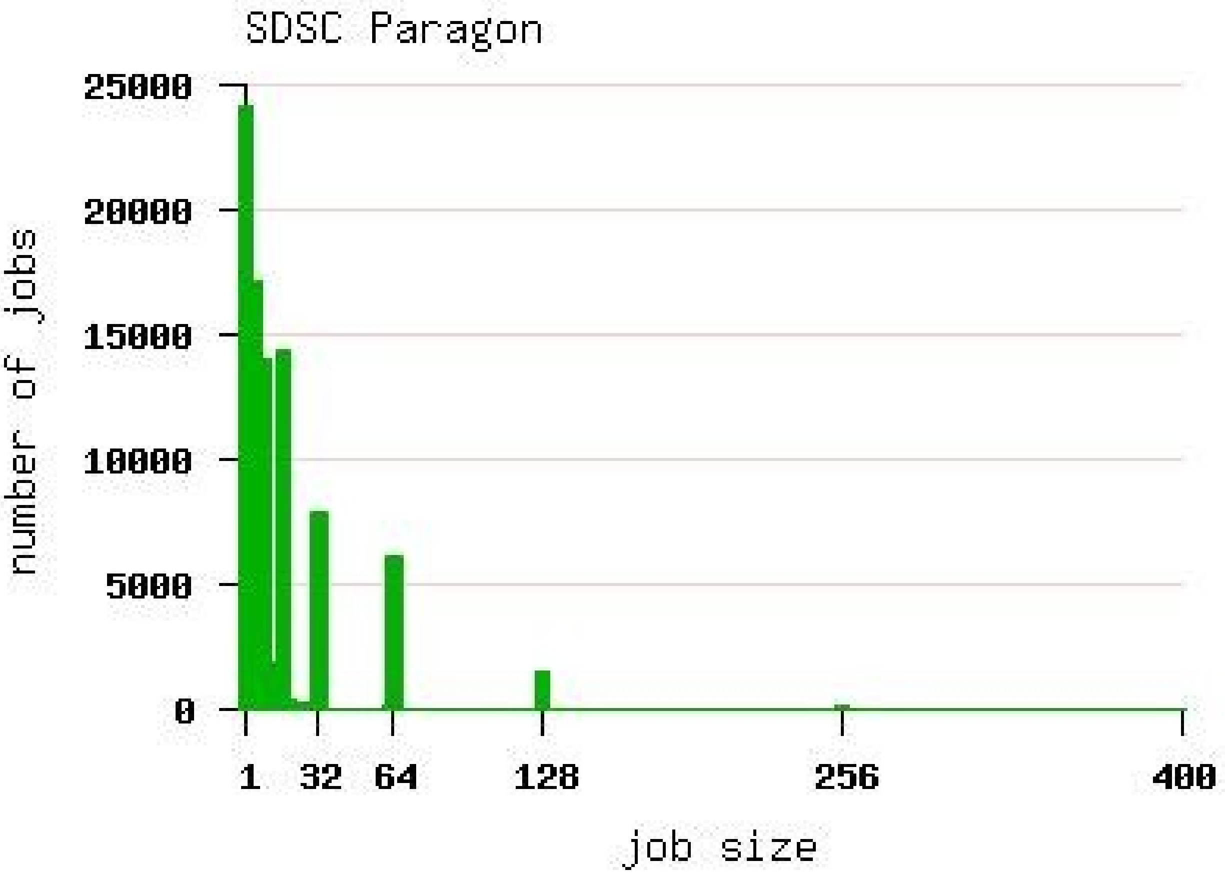
Stubs

- Uniform scale

  (Y axis)
- Match measured values

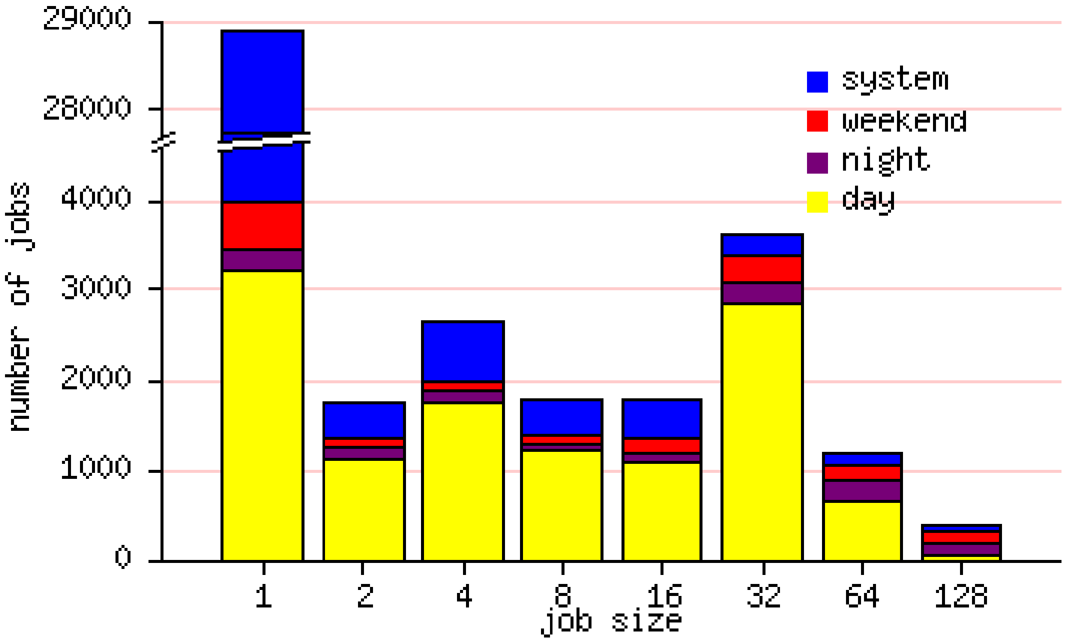
  (powers of 2)
- Show important values

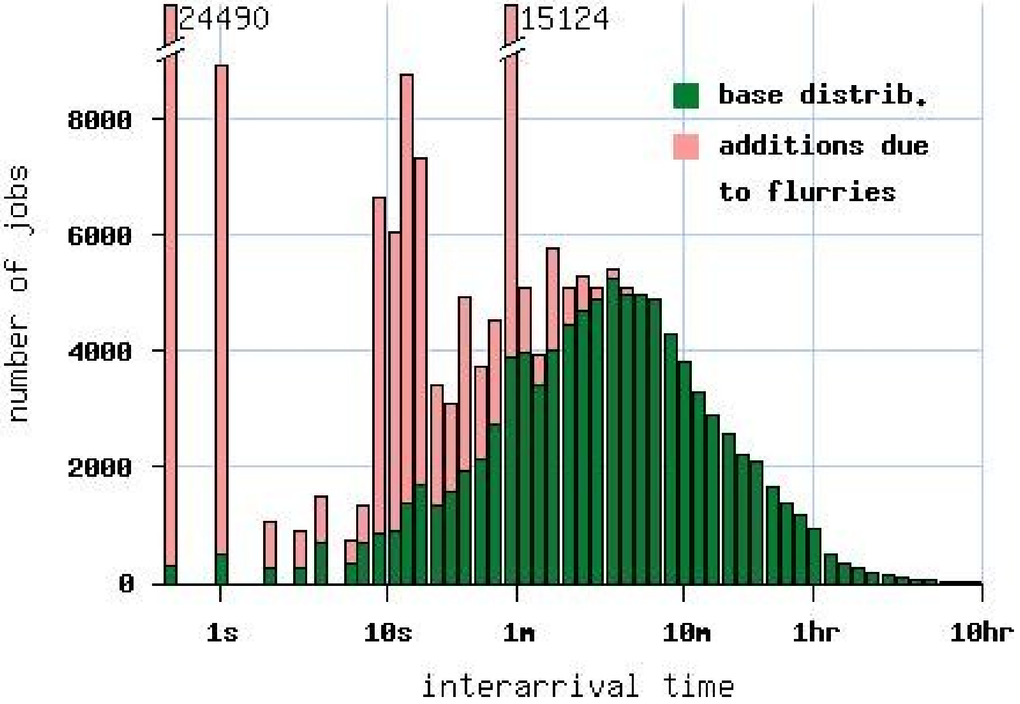
  (maximal size)



Axis break

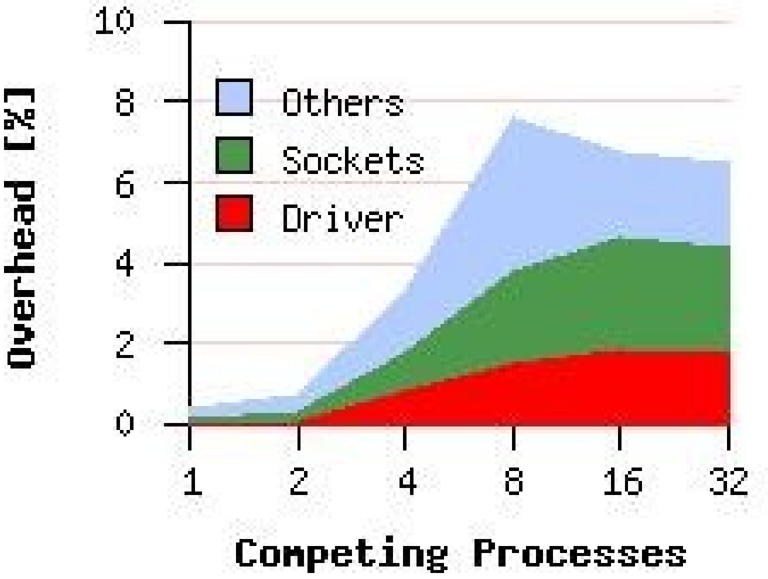
useful for few extreme values

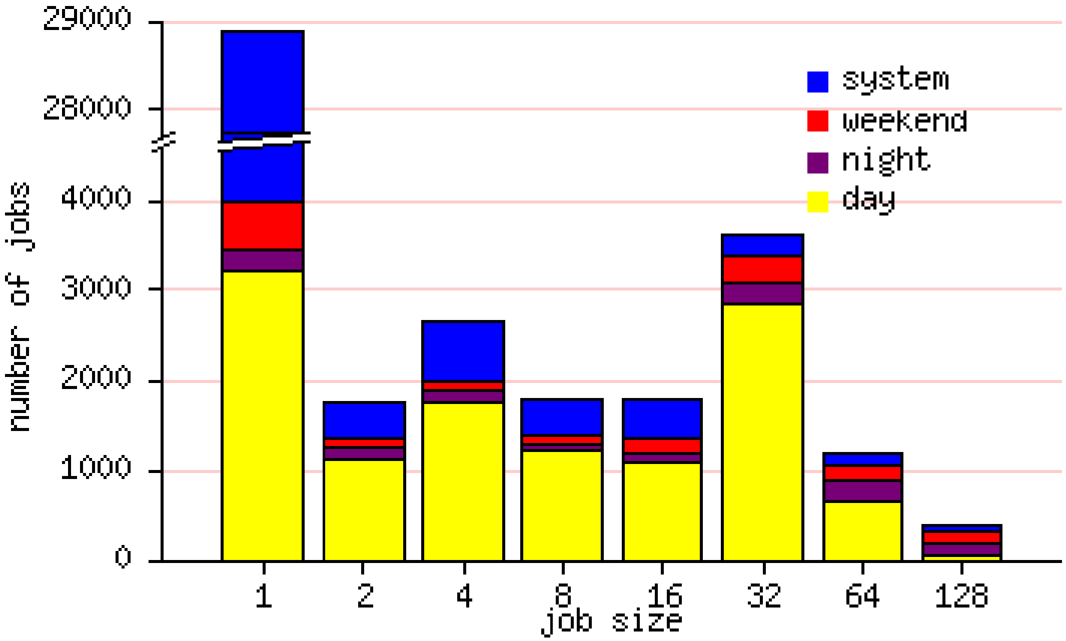




Stacking

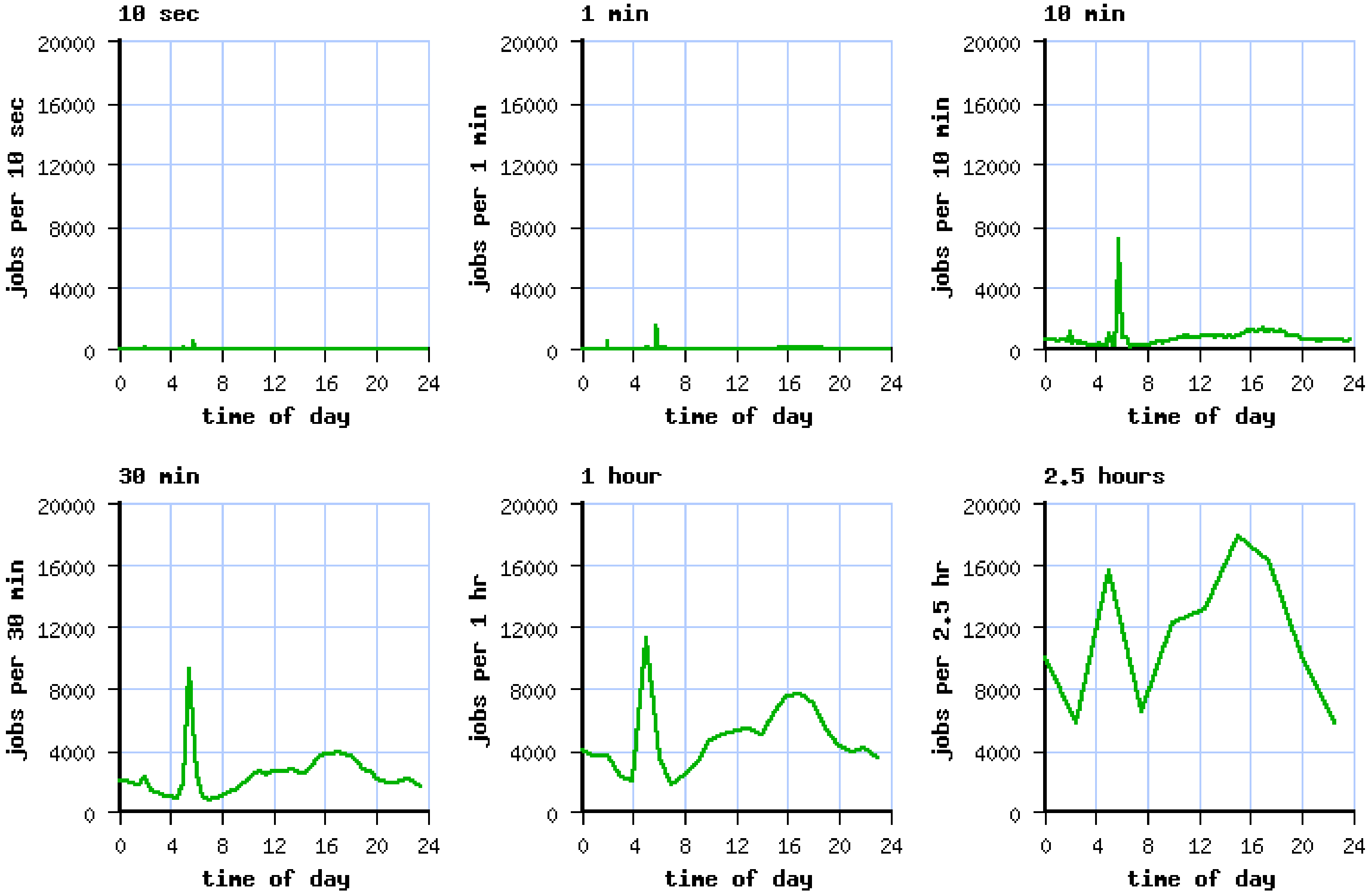
Show individual components and also their sum





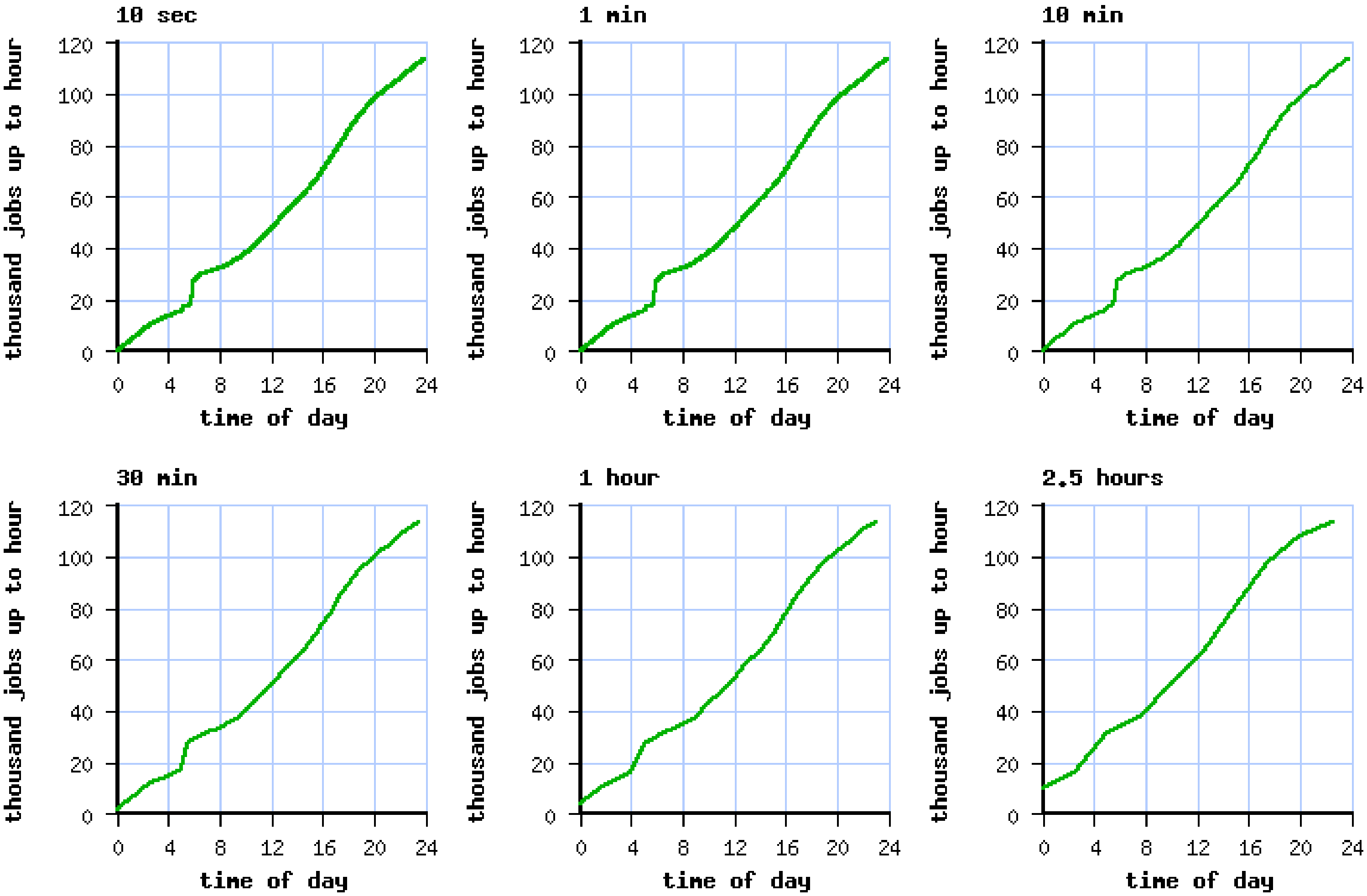
Histograms

- Simplest display of a distribution
- Sensitive to bin size



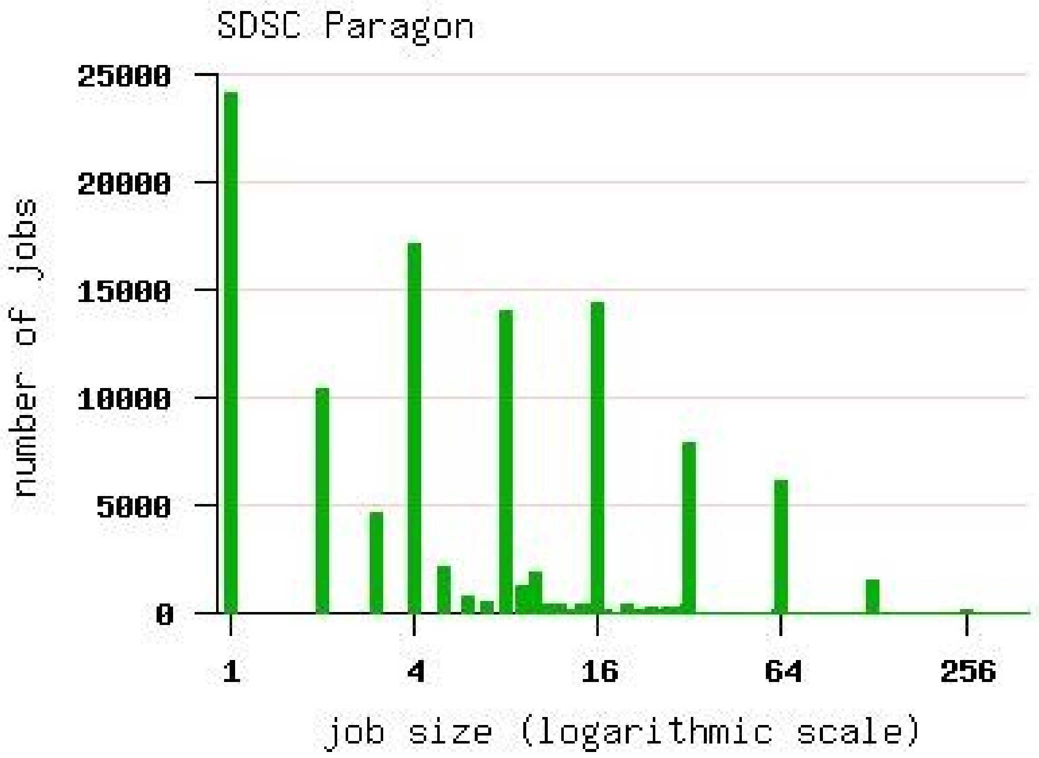
CDF

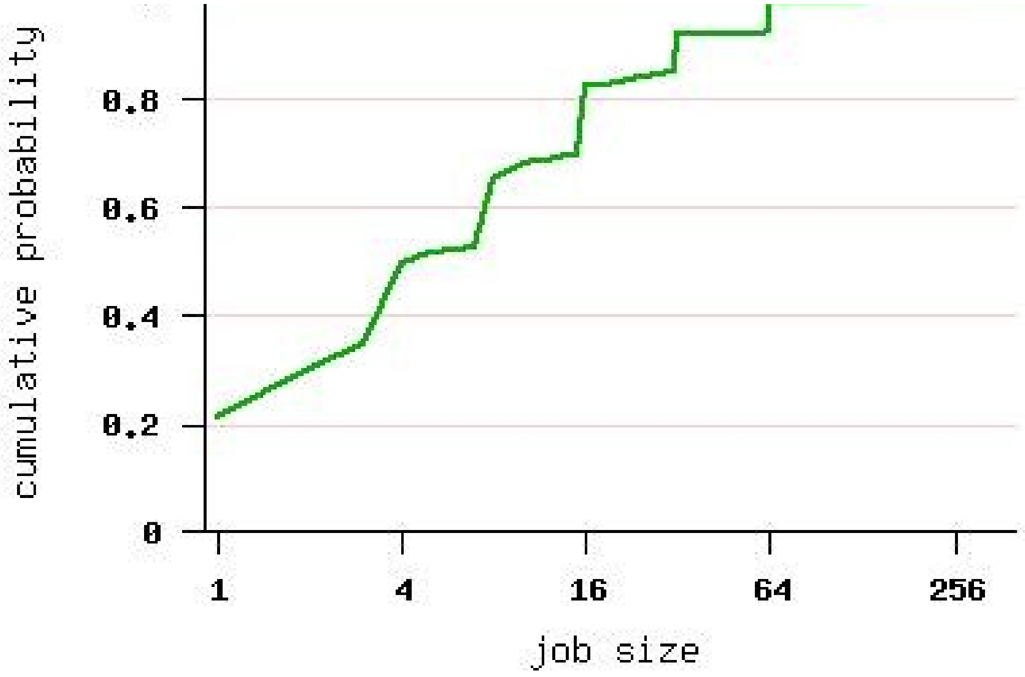
- Robust Alternative to histogram



CDF

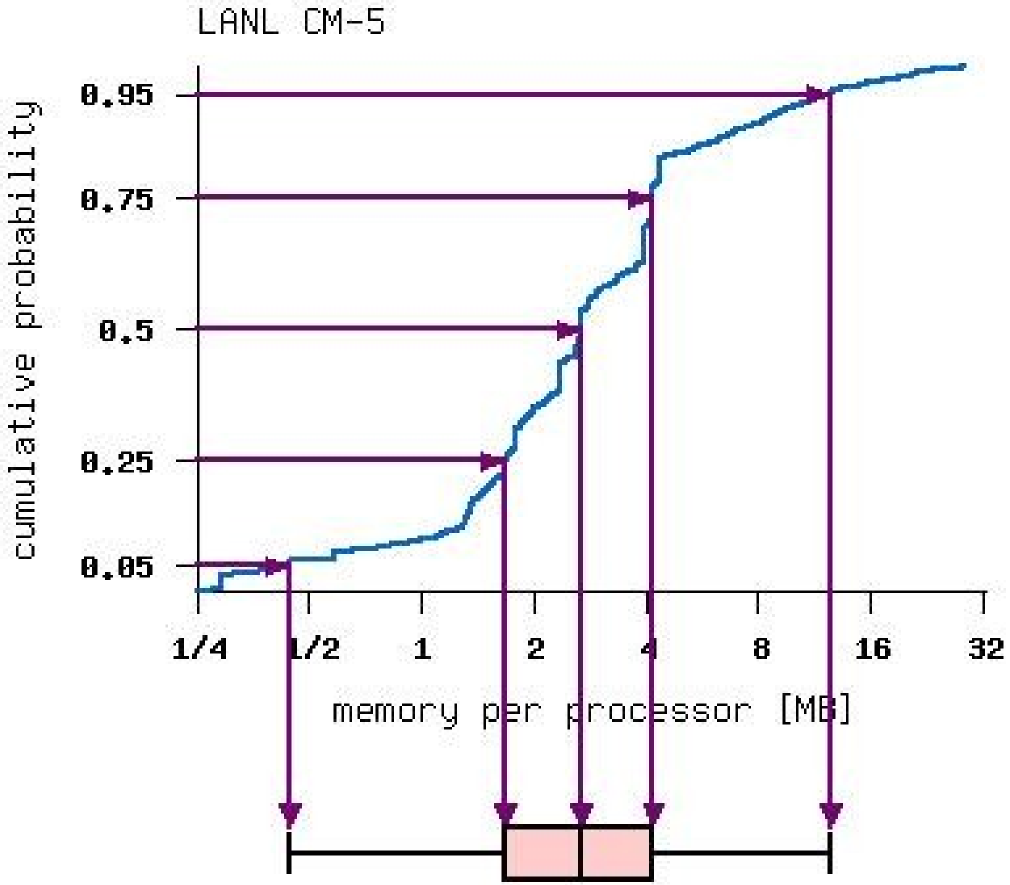
- Modes less prominent



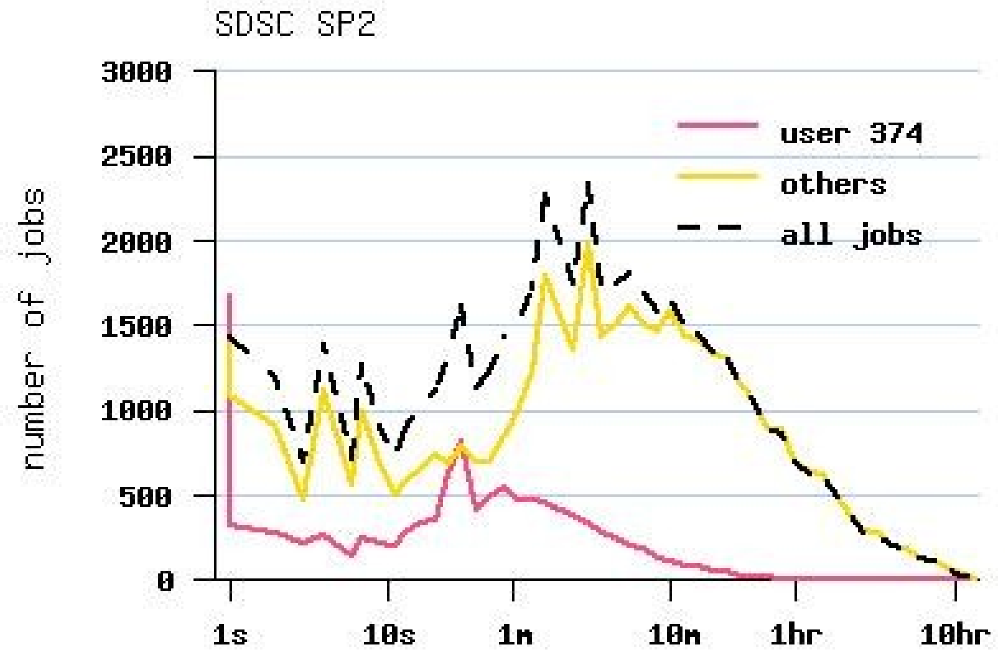


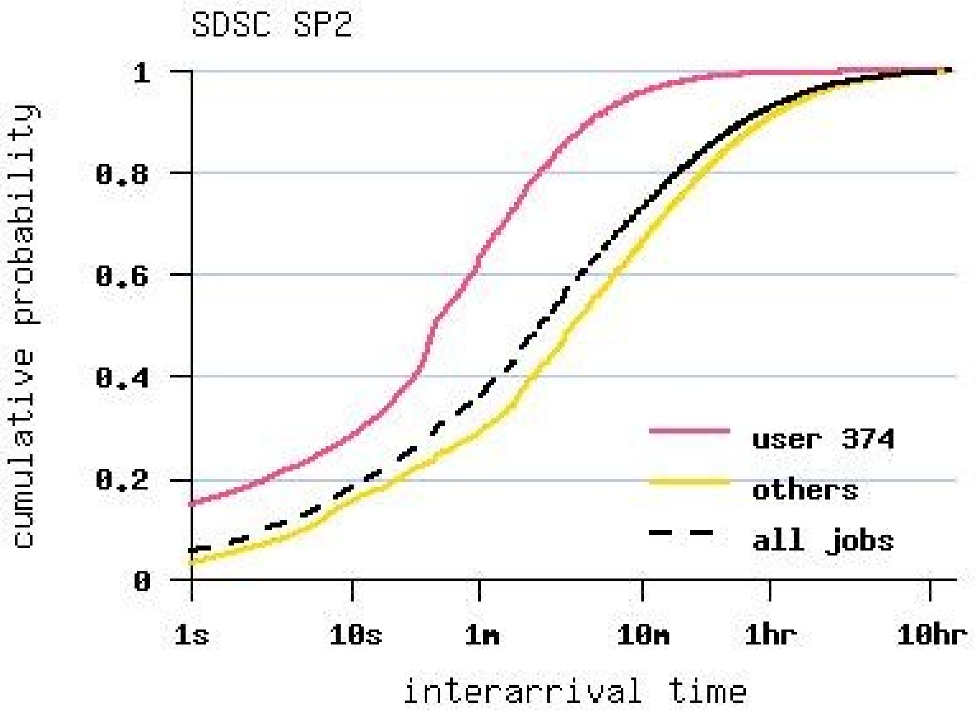
Box plot

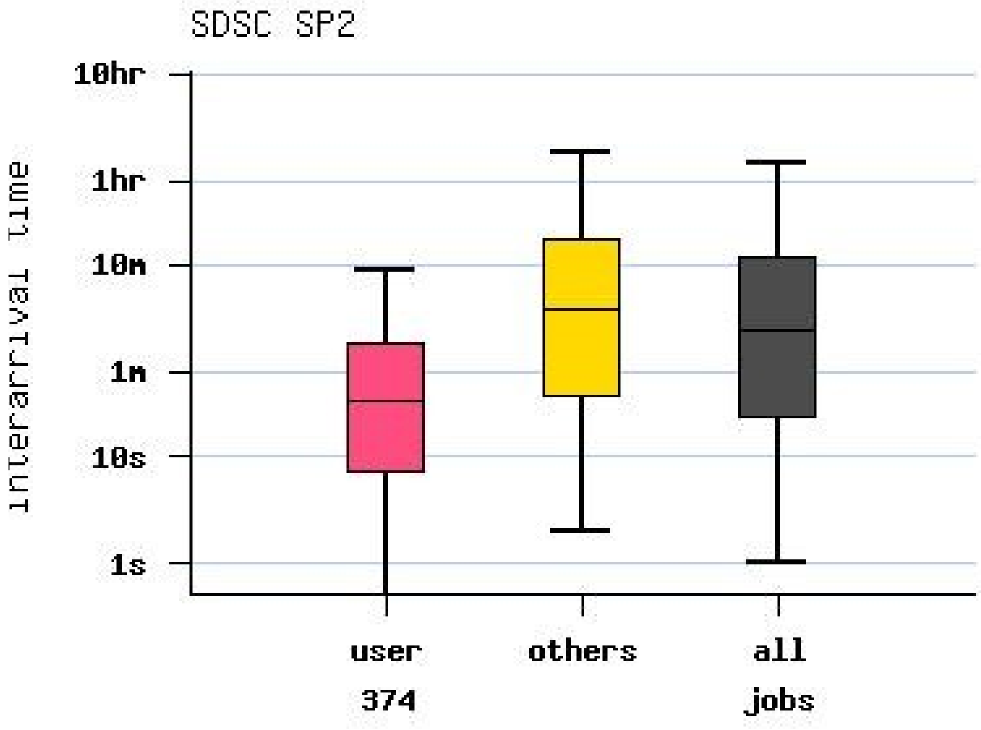
- Summary of a distribution



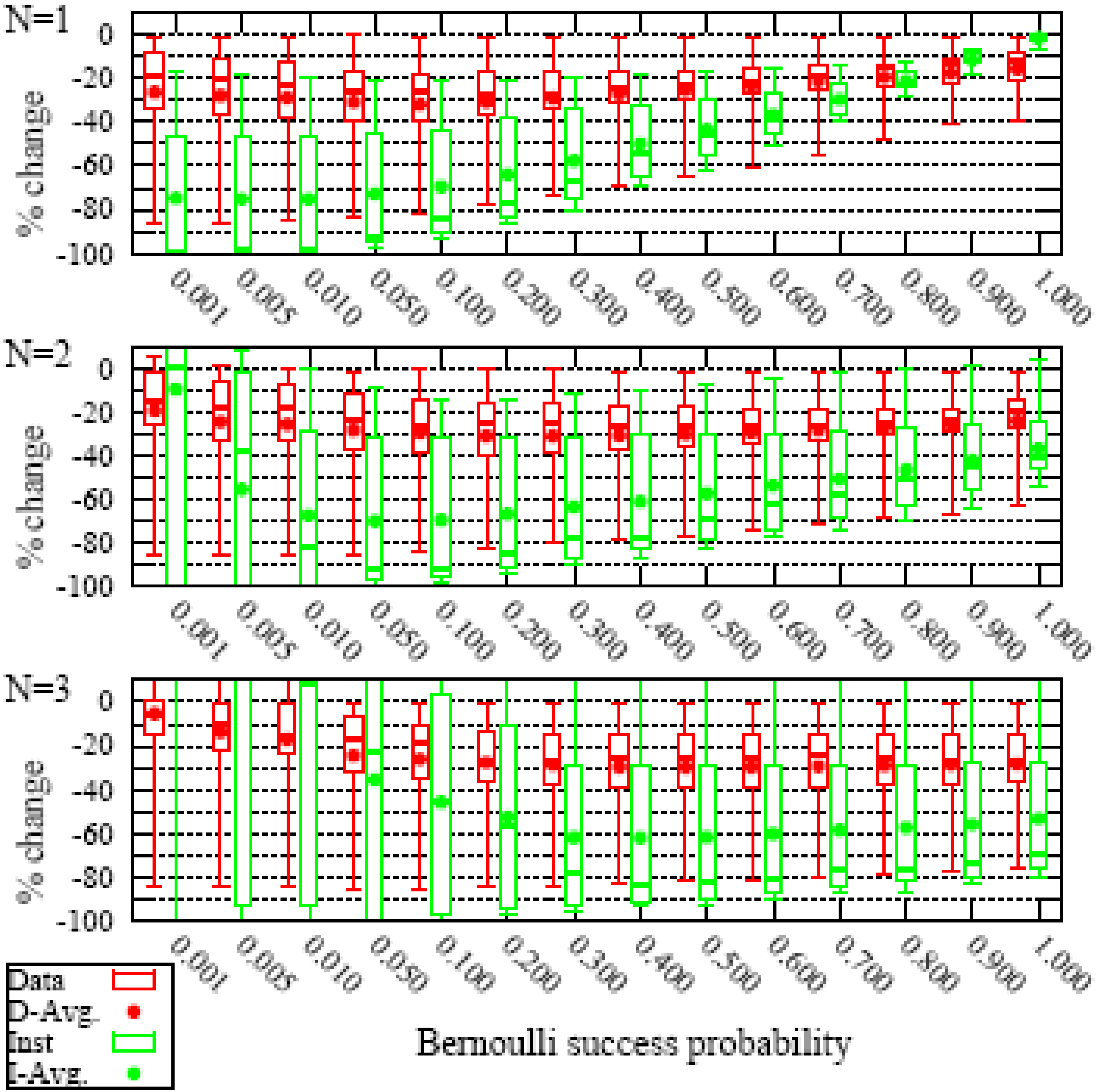
Comparison of distributions





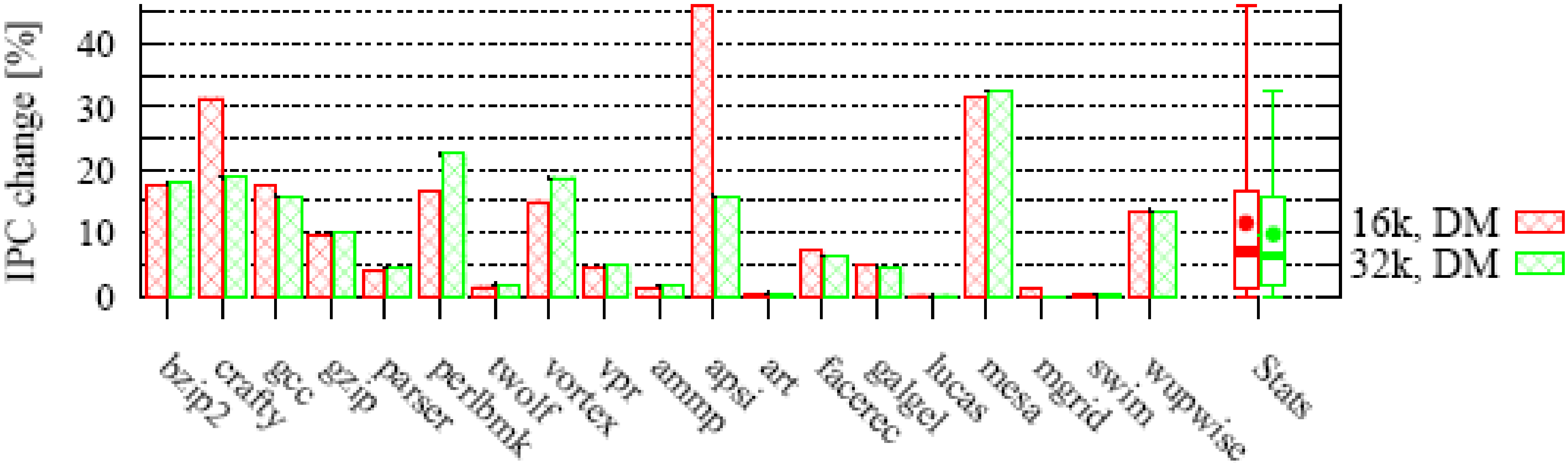


Comparison of distribution of results for different experimental parameter values



Skewed distributions are common

note difference between mean and median


IPC change [%]
bzip2
crafty
gcc
gzip
parser
perlbmk
twolf
vortex
vpr
ammp
apsi
art
facerec
galgel
lucas
mesa
mgrid
swim
wupwise
Stats
16k, DM
32k, DM