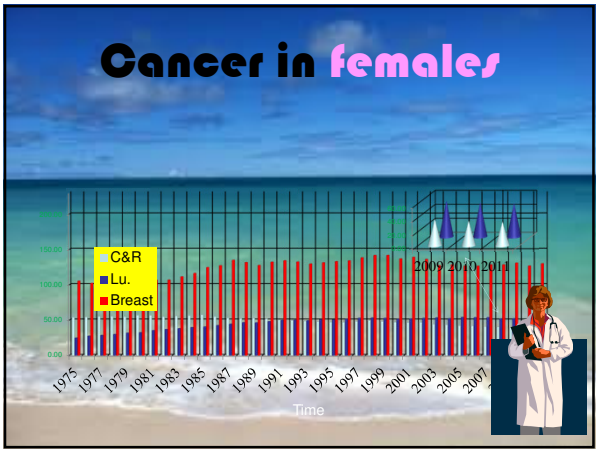
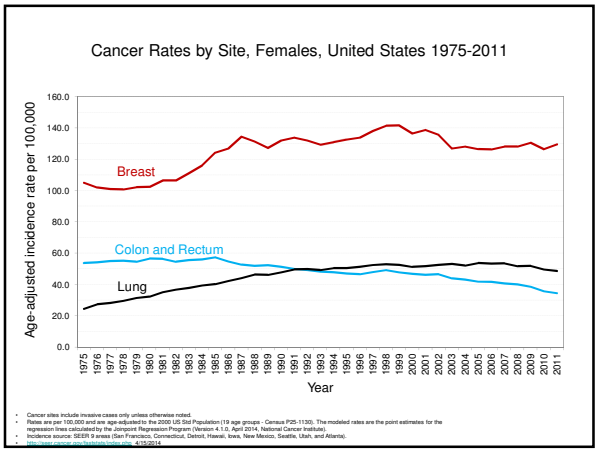
Visual Display of (Public Health) Data - Theory and Practice

Michael C. Samuel, Dr. P.H.
Senior Epidemiologist / Data Scientist





Outline

- Key Issues
  - The Big picture
  - Tufte
- History (les…)
- Software
  - R, PowerPoint, Excel, et. al. (more R…)
- **Type of Displays**
- **Technical Issues**
  - Scale
  - "Nut and Bolts"
    - color, fonts, lines/grids, labels/legends, 3D
  - Production and reproduction (less…)
  - Chart junk, Human touch
- Infographics, query systems
- Interactive Displays and R-Shiny
- "Great Graphs"
- Conclusion

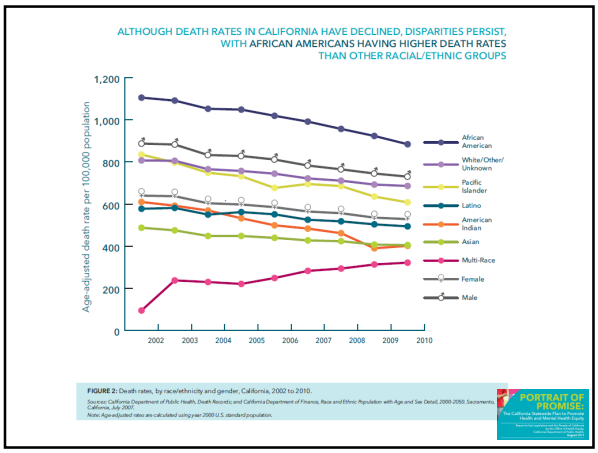
- Note: The example figures in this talk are to discuss form, not the actual substance of these data.

Data → Action

- Program
  - New program
  - Revised program priorities
- New guidelines
- New policy
- New hypothesis (may lead to new action)
- More (or less) money!

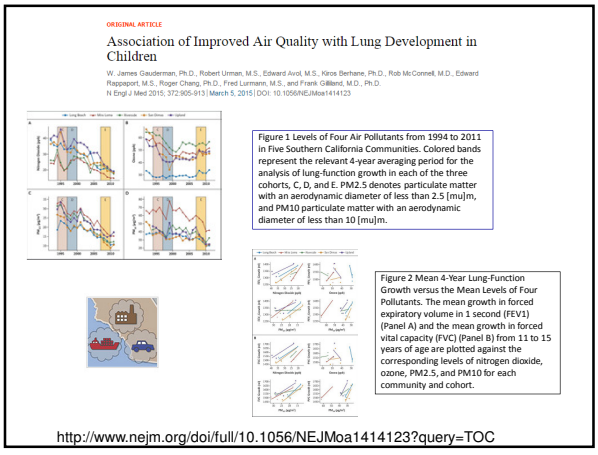
Guidelines for Effective Visual Display

- Communicate important information
- Complexity is good, and…
- Keep it simple, stupid
- Know your audience
- Oral presentation vs. written material
- Data integrity
- Clear labels and annotations
- Use appropriate scale(s)
- Use appropriate type of chart
- Pay attention to details
- Avoid extraneous "Chart Junk"



Emerging Issues

- Investigation/Analysis
- Presentation
  - Data
  - "Stories"
- Delivery Mode
  - E.g. Interactivity
  - E.g. Query Systems
- Nature of the information
  - E.g. Big Data
  - E.g. Open Data



http://www.nejm.org/doi/full/10.1056/NEJMoa1414123?query=TOC

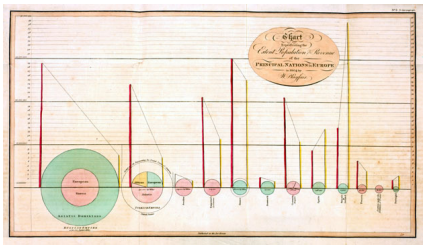
Edward Tufte

- Look at his books!
- Graphical Excellence
- The Lie Factor
- Data Density
- Less is more
- Small Multiples / Parallelism

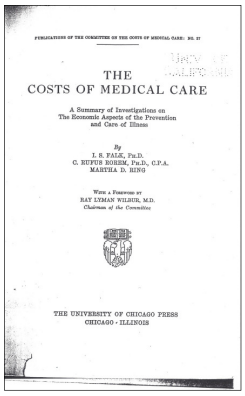
History

http://datavis.ca/milestones/

William Playfair. 1786. The Commercial and Political Atlas: Representing, by Means of Stained Copper-Plate Charts, the Progress of the Commerce, Revenues, Expenditure and Debts of England during the Whole of the Eighteenth Century.



William Playfair, _Inquiry into the Permanent Causes of the Decline and Fall of Wealthy and Powerful Nations_ (1805), figure 2



PUBLICATIONS OF THE COMMITTEE ON THE COSTS OF MEDICAL CARE: NO. 27

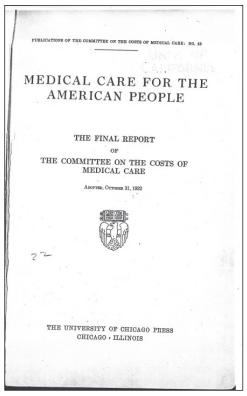
THE
COSTS OF MEDICAL CARE

A Summary of Investigations on
The Economic Aspects of the Prevention
and Care of Illness

By
I. S. FALK, PH.D.
C. RUFUS ROREM, PH.D., C.P.A.
MARTHA D. RING

WITH A FOREWORD BY
RAY LYMAN WILBUR, M.D.
Chairman of the Committee

THE UNIVERSITY OF CHICAGO PRESS
CHICAGO · ILLINOIS



PUBLICATIONS OF THE COMMITTEE ON THE COSTS OF MEDICAL CARE: NO. 10

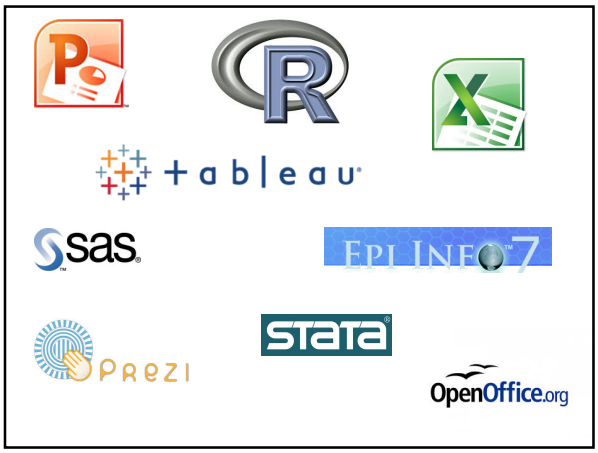
MEDICAL CARE FOR THE
AMERICAN PEOPLE

THE FINAL REPORT
OF
THE COMMITTEE ON THE COSTS OF
MEDICAL CARE

ADOPTED, OCTOBER 31, 1932

THE UNIVERSITY OF CHICAGO PRESS
CHICAGO · ILLINOIS

Software



Software

- Stand alone graphics packages
  - PowerPoint; Open Office Impress
  - Great for presentations; easy to use
- Spreadsheets
  - Excel
  - Easy to use
  - Can be difficult to modify or share
  - Direct integration of data and figures
- Stat packages with graphics
  - SAS, SPSS, Stata; Epi Info
  - Integrate data and graphics
  - Some "point and click", some programming
  - Not as ideal for presentations
- R (S-plus) ←
  - Free
  - Complete integration of data and graphics
  - Completely flexible graphics
  - Harder to learn/use
- Specialized Software
  - Eg. NodeEX, "NetDraw" Network analysis

- Cool Tricks:

- Excel
  - Conditional formatting bars
  - Spark lines / bars

Display Types

- Tables
- Line Charts
- Bar Charts
- Pie Charts
- Scattergrams
- Statistical Charts
  - Box Plots
- Maps
- Others
- Hybrid

Line Graph

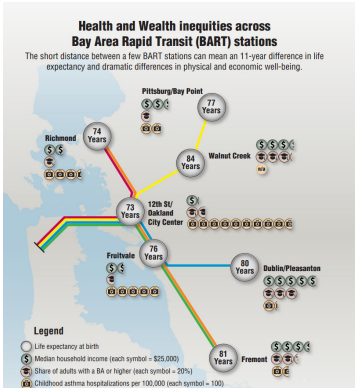
- X-axis truly or close to continuous
- Simple
- Complex: multi-line, 2-axis, logarithmic

Bar Chart

- Very common chart type
- Y-axis: count, rate or percent of something
- X-axis: qualitative variable, or ordered categorical variable
- Vertical bars or horizontal bars
- Simple
- Clustered/Grouped
- Stacked
- 100%
- Histogram=special case

Pie Chart

- Tufte says they should never be used
- But
  - Very familiar to most people
  - Easy to understand
  - Effective if used carefully and sparingly



- http://www.frbsf.org/community-development/files/BART-Health-and-Wealth-Map.pdf

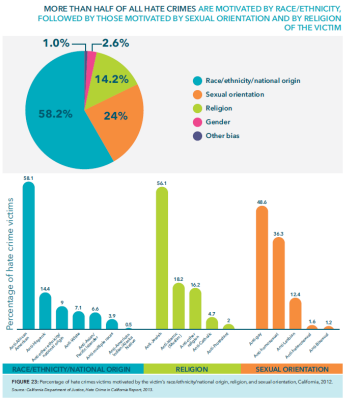
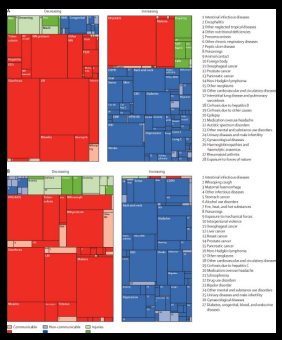


Figure 3



The Lancet 2015 386, 2145-2191DOI: (10.1016/S0140-6736(15)61340-X)

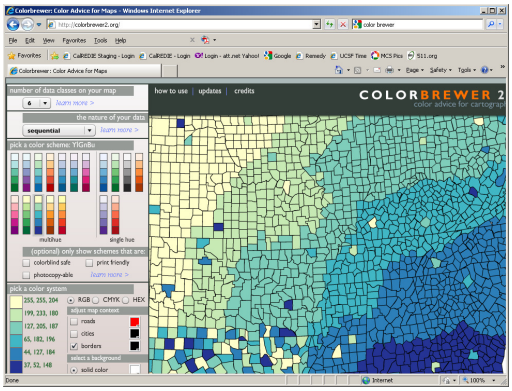
"Nut and Bolts"

- Scale and Proportion
- Labels and Legends
- Grid Lines
- Color
- Animation/"PowerPoint"
- Font
- 3D
- Production/Reproduction
- Chart Junk
- Software

Color

- Use for a reason
- Use nice colors
  - Shades of Blue
  - Shade of Yellow
  - Colors of Nature
- Use color sparingly
- RED can be good for Main Point, if used sparingly
- Red often does not project well with slides and LCDs
- Use consistent colors (and fonts, etc.)

http://colorbrewer2.org/

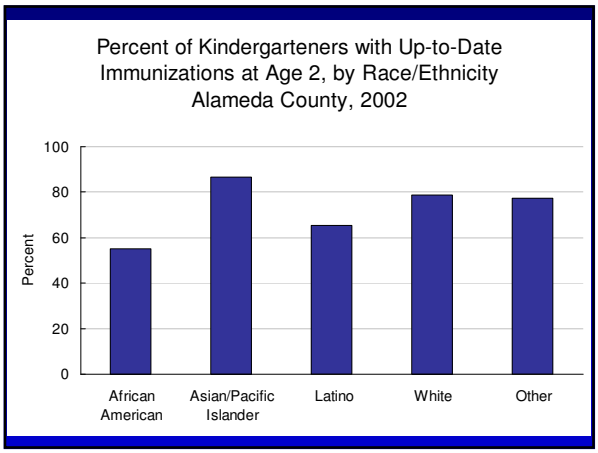


Fonts / Fonts

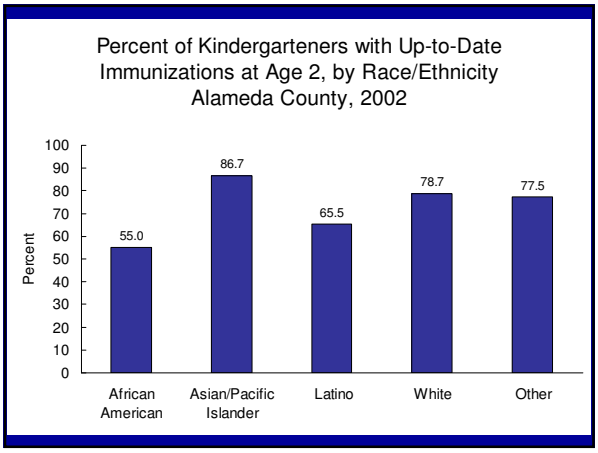
- Use San Serif Fonts, Like Arial
- Not Serif Fonts, Like Times Roman
  - They Are Harder to Read
    - Particularly in Oral Presentations
      - When the Font Is Small
      - See, Isn't This Better
- ALSO, ALMOST NEVER USE ALL CAPS
  - IT'S HARD TO READ TOO
- Big Enough to read

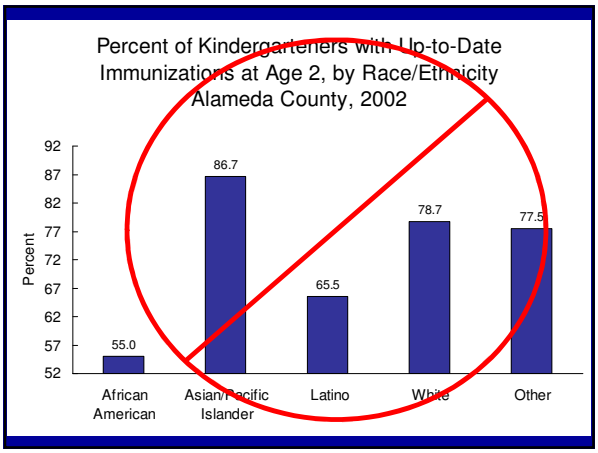
Production / Reproduction

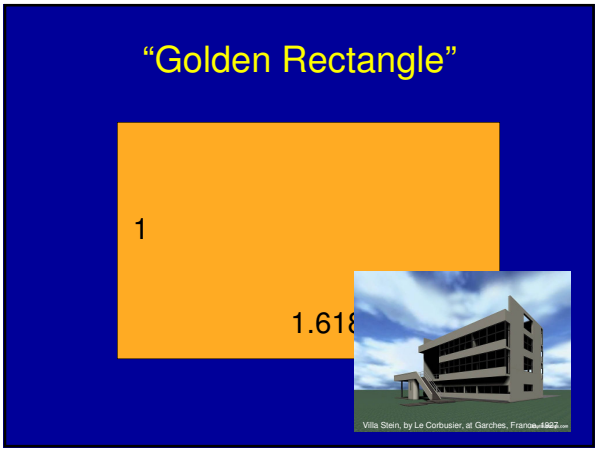
- Test printers, laptops, LCDs before full production is necessary
- Often different colors and styles for:
  - PowerPoint oral presentation
  - Written report or manuscript
- Color
  - May not photocopy (or print) well
  - Can be expensive to reproduce
- Posters made on plotters require special consideration

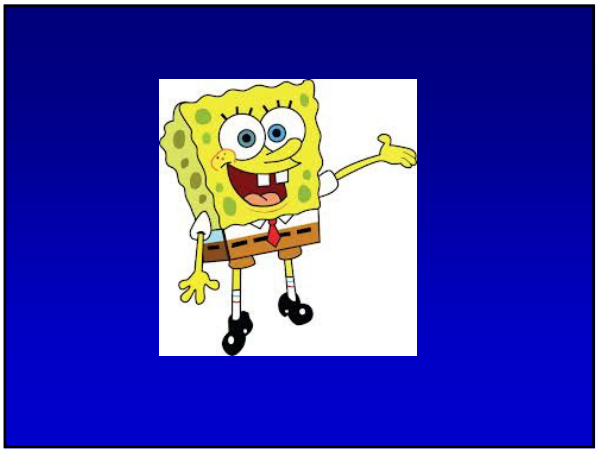

Percent of Kindergarteners with Up-to-Date Immunizations at Age 2, by Race/Ethnicity Alameda County, 2002
Percent
100
80
60
40
20
0
African American
Asian/Pacific Islander
Latino
White
Other

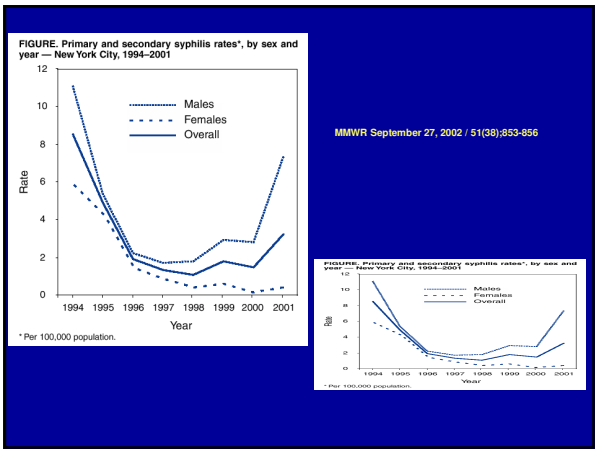
Scale


Percent of Kindergarteners with Up-to-Date Immunizations at Age 2, by Race/Ethnicity Alameda County, 2002
Percent
100
90
80
70
60
50
40
30
20
10
0
55.0
86.7
65.5
78.7
77.5
African American
Asian/Pacific Islander
Latino
White
Other

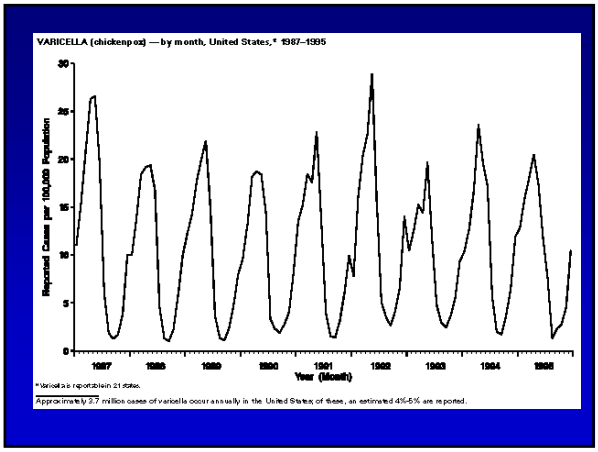

Percent of Kindergarteners with Up-to-Date Immunizations at Age 2, by Race/Ethnicity
Alameda County, 2002
Percent
92
87
82
77
72
67
62
57
52
55.0
86.7
65.5
78.7
77.5
African American
Asian/Pacific Islander
Latino
White
Other

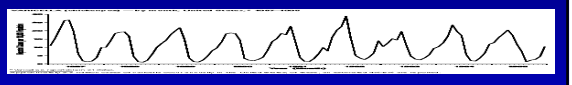

"Golden Rectangle"
1
1.618
Villa Stein, by Le Corbusier, at Garches, France, 1927





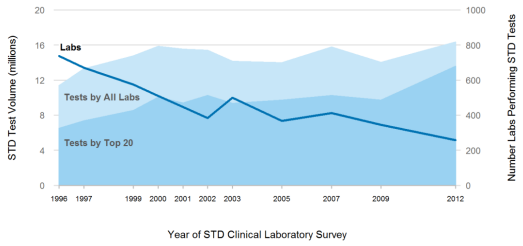
MMWR September 27, 2002 / 51(38);853-856



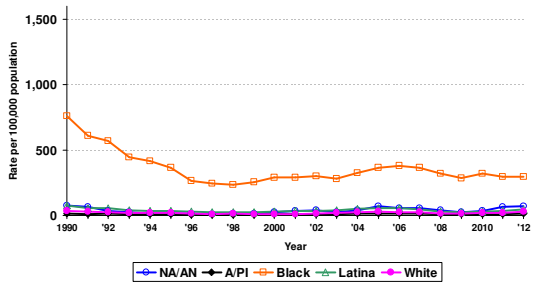


"Bank to 45 degrees"

STD Test Volume Performed by All and Top 20 Laboratories, 1996-2012



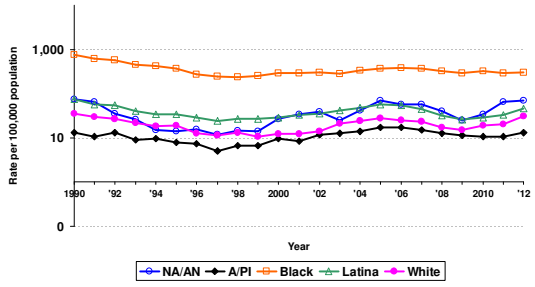
Gonorrhea, Rates for Females by Race/Ethnicity California, 1990–2012



Note: NA/AN = Native American/Alaskan Native, A/PI = Asian/Pacific Islander.
Race/ethnicity "Not Specified" ranged from 29.2% to 43.1% of cases for females in any given year.

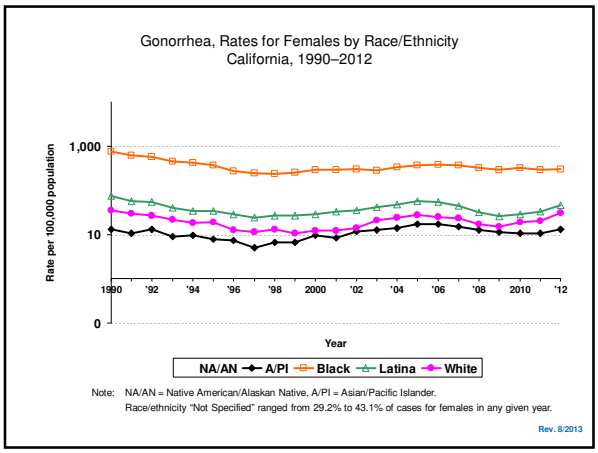
Rev. 8/2013

Gonorrhea, Rates for Females by Race/Ethnicity California, 1990–2012

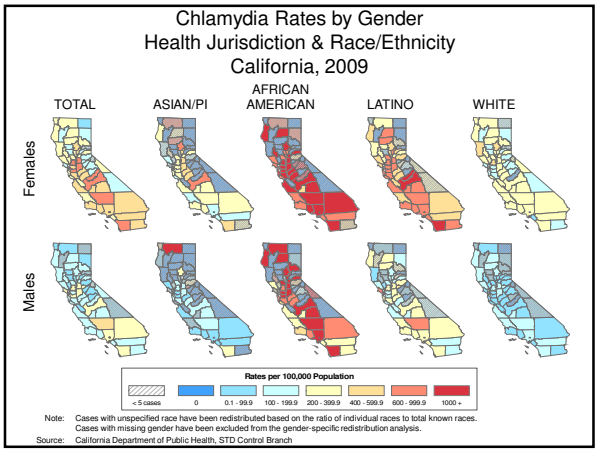


Note: NA/AN = Native American/Alaskan Native, A/PI = Asian/Pacific Islander.
Race/ethnicity "Not Specified" ranged from 29.2% to 43.1% of cases for females in any given year.

Rev. 8/2013

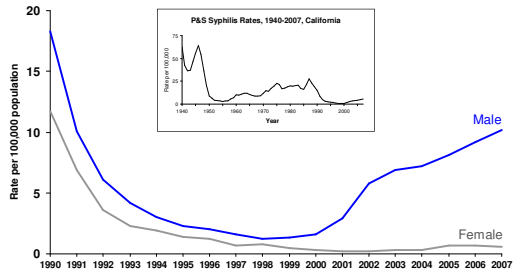

Gonorrhea, Rates for Females by Race/Ethnicity
California, 1990–2012
Rate per 100,000 population
1,000
10
0
1990
'92
'94
'96
'98
2000
'02
'04
'06
'08
2010
'12
Year
NA/AN
A/PI
Black
Latina
White
Note: NA/AN = Native American/Alaskan Native, A/PI = Asian/Pacific Islander.
Race/ethnicity "Not Specified" ranged from 29.2% to 43.1% of cases for females in any given year.
Rev. 8/2013

Cut Points


Chlamydia Rates by Gender
Health Jurisdiction & Race/Ethnicity
California, 2009
TOTAL
ASIAN/PI
AFRICAN AMERICAN
LATINO
WHITE
Females
Males
Rates per 100,000 Population
< 5 cases
0
0.1 - 99.9
100 - 199.9
200 - 399.9
400 - 599.9
600 - 999.9
1000 +
Note: Cases with unspecified race have been redistributed based on the ratio of individual races to total known races.
Cases with missing gender have been excluded from the gender-specific redistribution analysis.
Source: California Department of Public Health, STD Control Branch

Labels and Legends

Primary & Secondary Syphilis, Rates by Gender
California, 1990–2007



CA DPH STD Control Branch (rev 7/2008)

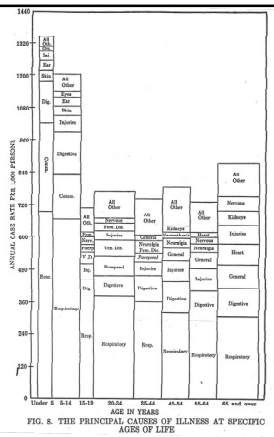
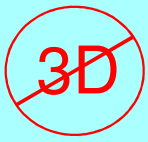
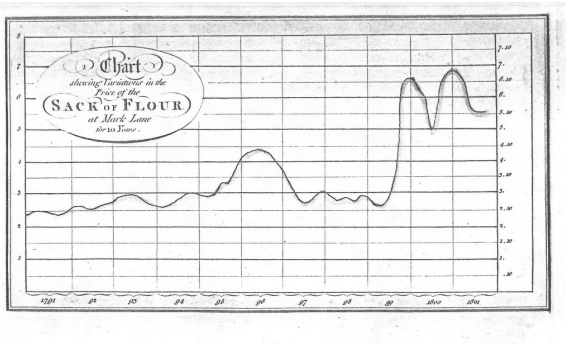


FIG. 8. THE PRINCIPAL CAUSES OF ILLNESS AT SPECIFIC AGES OF LIFE

3D Charts



- Unless there are 3 dimensions and the audience can handle it!



THE COMMERCIAL AND POLITICAL ATLAS
1801, WILLIAM PLAYFAIR

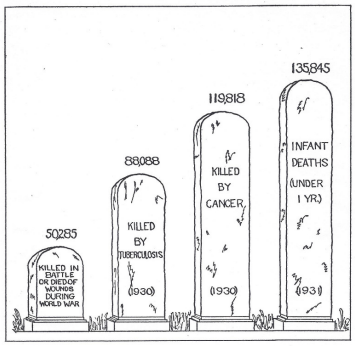
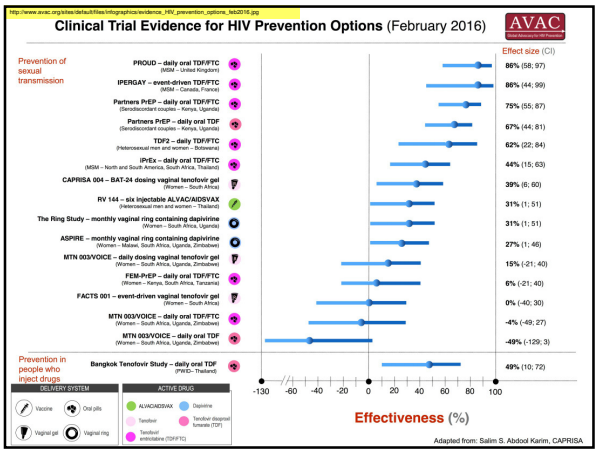
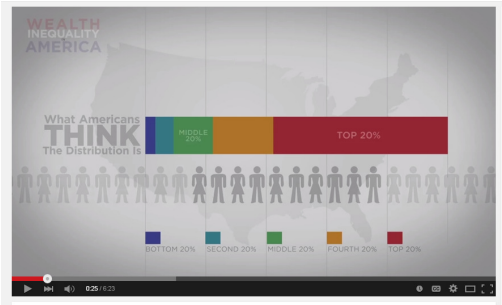


FIG. 18. "HUMAN LIFE IN THE UNITED STATES IS BEING WASTED AS RECKLESSLY, AS SURELY, IN TIMES OF PEACE AS IN TIMES OF WAR"

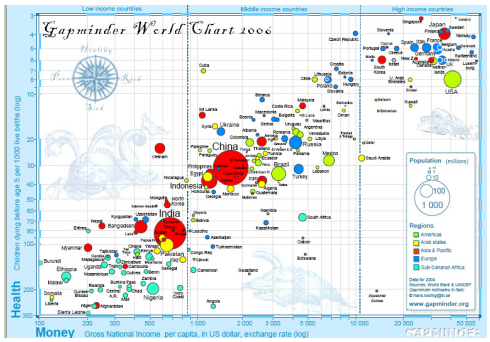




https://www.youtube.com/watch?v=QPKKQnijnsM

http://www.bloomberg.com/dataview/2014-04-17/how-americans-die.html

http://www.gapminder.org/

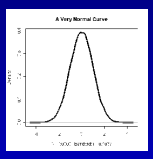


In Conclusion

- Make displays that matter
- Know your audience
- Simple ↔ Complex
- Less is more
- Pay attention to “nuts and bolts” details

For More Information:

- Michael.Samuel@cdph.ca.gov
- 510.620.3198



Part 1 “General Concepts”
Part 1a: http://youtu.be/1c41eMOjt_U
Part 1b: http://youtu.be/XIKA2hgg-rY

Part 2 “Nuts and Bolts”
Part 2a: http://youtu.be/pUDcGIuIfW8
Part 2b: http://youtu.be/YCRyVPpz-yk