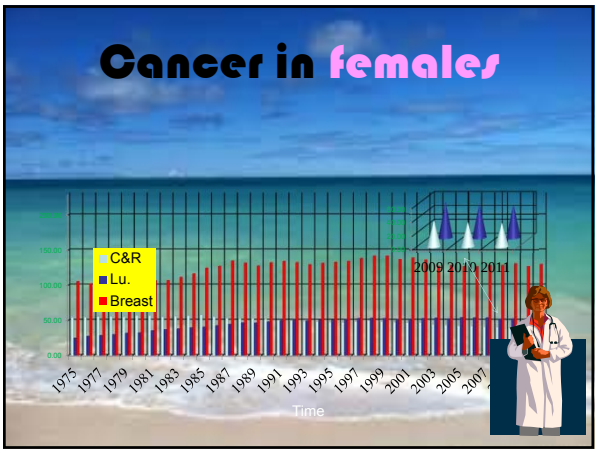
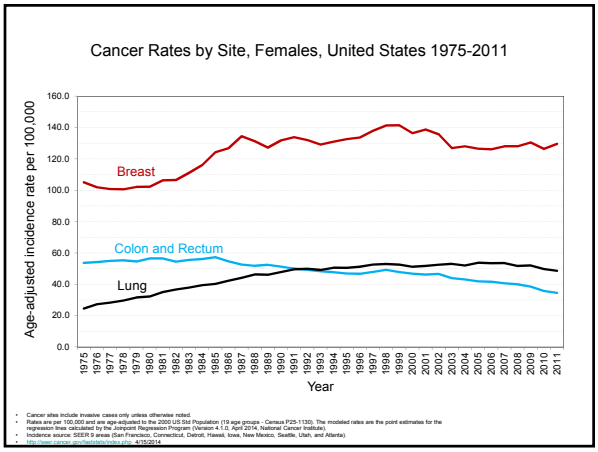
Visual Display of Public Health Data

Michael C. Samuel, Dr. P.H.
CA DHS STD Control Branch





Outline

- Key Issues
  - The Big picture
  - Tufte
  - Sponge Bob
- History (les…)
- Software
  - R, PowerPoint, Excel, et. al. (more R…)
  - Big data
- **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

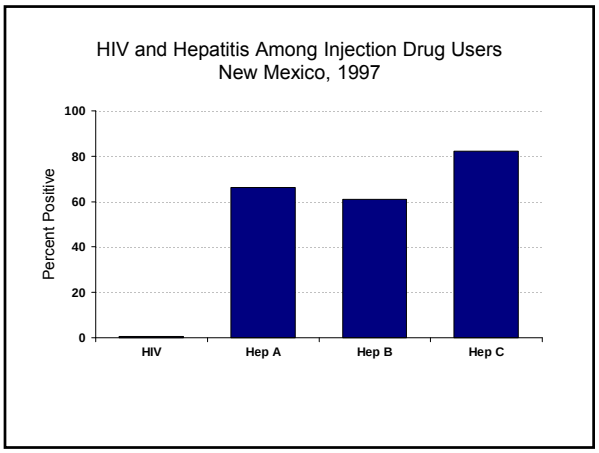
- Note: The example figures in this talk are to discuss form, not the actual substance of these data.

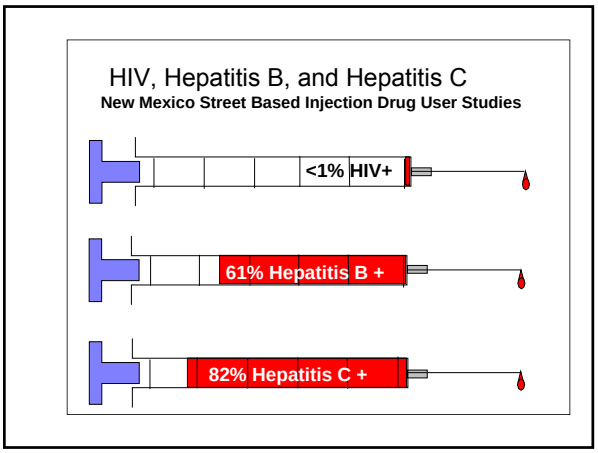
Data → **Action**

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

HIV and Hepatitis among Injection Drug Users
New Mexico, 1997

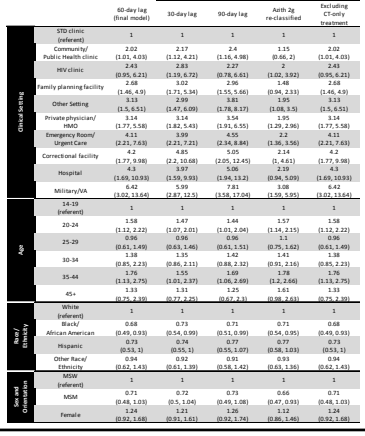
	N tested	% Positive	95% C.I.
HIV	1002	0.5	.16-1.6
Hep A	696	66.1	62.4-69.6
Hep B	950	61.1	57.9-64.2
Hep C	945	82.2	79.6-84.6



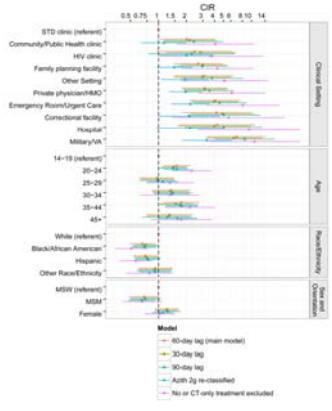


Guidelines for Effective Visual Display

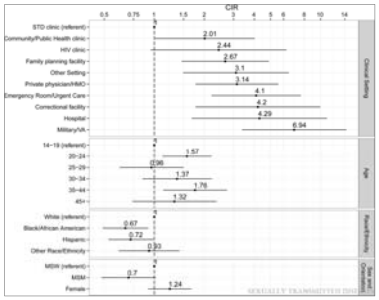
- Communicate important information
- Complexity is good, and…
- Keep it simple, stupid
- Know your audience
- Oral presentation vs. written material
- Data integrity
- Clear labels and annotations
- Use appropriate scale(s)
- Use appropriate type of chart
- Pay attention to details
- Avoid extraneous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



		60-day lag (final model)	Sensitivity Analysis: 30-day lag	90-day lag	Ashth 2g re-classified	Excluding CT-only treatment
Clinical Setting	STD clinic (referent)	1	1	1	1	1
	Community/ Public Health clinic	2.02 (1.01, 4.03)	2.17 (1.12, 4.21)	2.4 (1.16, 4.98)	1.15 (0.66, 2)	2.02 (1.01, 4.03)
	HIV clinic	2.43 (0.95, 6.21)	2.83 (1.19, 6.72)	2.27 (0.78, 6.61)	2 (1.02, 3.92)	2.43 (0.95, 6.21)
	Family planning facility	2.68 (1.46, 4.9)	3.02 (1.71, 5.34)	2.96 (1.55, 5.66)	1.48 (0.94, 2.33)	2.68 (1.46, 4.9)
	Other Setting	3.13 (1.5, 6.51)	2.99 (1.47, 6.09)	3.81 (1.78, 8.17)	1.95 (1.08, 3.5)	3.13 (1.5, 6.51)
	Private physician/ HMO	3.14 (1.77, 5.58)	3.14 (1.82, 5.43)	3.54 (1.91, 6.55)	1.95 (1.29, 2.96)	3.14 (1.77, 5.58)
	Emergency Room/ Urgent Care	4.11 (2.21, 7.63)	3.99 (2.21, 7.21)	4.55 (2.34, 8.84)	2.2 (1.36, 3.56)	4.11 (2.21, 7.63)
	Correctional facility	4.2 (1.77, 9.98)	4.85 (2.2, 10.68)	5.05 (2.05, 12.45)	2.14 (1, 4.61)	4.2 (1.77, 9.98)
	Hospital	4.3 (1.69, 10.93)	3.97 (1.59, 9.93)	5.06 (1.94, 13.2)	2.19 (0.94, 5.09)	4.3 (1.69, 10.93)
	Military/VA	6.42 (3.02, 13.64)	5.99 (2.87, 12.5)	7.81 (3.58, 17.04)	3.08 (1.59, 5.95)	6.42 (3.02, 13.64)
Age	14-19 (referent)	1	1	1	1	1
	20-24	1.58 (1.12, 2.22)	1.47 (1.07, 2.01)	1.44 (1.01, 2.04)	1.57 (1.14, 2.15)	1.58 (1.12, 2.22)
	25-29	0.96 (0.61, 1.49)	0.96 (0.63, 1.46)	0.96 (0.61, 1.51)	1.1 (0.75, 1.62)	0.96 (0.61, 1.49)
	30-34	1.38 (0.85, 2.23)	1.35 (0.86, 2.11)	1.42 (0.88, 2.32)	1.41 (0.91, 2.16)	1.38 (0.85, 2.23)
	35-44	1.76 (1.13, 2.75)	1.55 (1.01, 2.37)	1.69 (1.06, 2.69)	1.78 (1.2, 2.66)	1.76 (1.13, 2.75)
	45+	1.33 (0.75, 2.39)	1.31 (0.77, 2.25)	1.25 (0.67, 2.3)	1.61 (0.98, 2.63)	1.33 (0.75, 2.39)
Race/ Ethnicity	White (referent)	1	1	1	1	1
	Black/ African American	0.68 (0.49, 0.93)	0.73 (0.54, 0.99)	0.71 (0.51, 0.99)	0.71 (0.54, 0.95)	0.68 (0.49, 0.93)
	Hispanic	0.73 (0.53, 1)	0.74 (0.55, 1)	0.77 (0.55, 1.07)	0.77 (0.58, 1.03)	0.73 (0.53, 1)
	Other Race/ Ethnicity	0.94 (0.62, 1.43)	0.92 (0.61, 1.39)	0.91 (0.58, 1.42)	0.93 (0.63, 1.36)	0.94 (0.62, 1.43)
Sex and Orientation	MSW (referent)	1	1	1	1	1
	MSM	0.71 (0.48, 1.03)	0.72 (0.5, 1.04)	0.73 (0.49, 1.08)	0.66 (0.47, 0.93)	0.71 (0.48, 1.03)
	Female	1.24 (0.92, 1.68)	1.21 (0.91, 1.61)	1.26 (0.92, 1.74)	1.12 (0.86, 1.46)	1.24 (0.92, 1.68)



.tiff ~ 6.9 MB
Lectenberg, Samuel, et al.

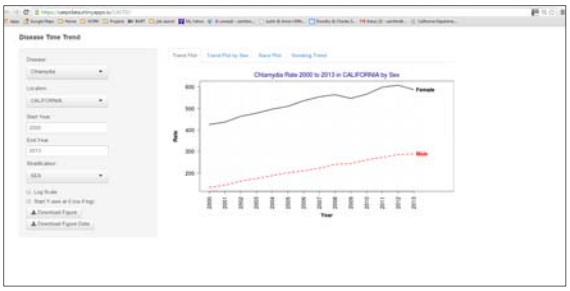


[Variation in Adherence to the Treatment Guidelines for Neisseria gonorrhoeae by Clinical Practice Setting, California, 2009 to 2011]

Lechtenberg, Richard J.; Samuel, Michael C.; Bernstein, Kyle T.; Lahiff, Maureen; Olson, Nicole; Bauer, Heidi M.
Sexually Transmitted Diseases. 41(5):338-344, May 2014.

Figure 2 Results of multivariable statistical modeling. This plot shows cumulative incidence ratios (CIR) and 95% CIs for receiving a non–CDC-recommended treatment of gonorrhea by clinical setting, age, race/ethnicity, and sex and sexual orientation. MSW indicates men who have sex with women only; MSM, men who have sex with men.

Copyright © 2014 Sexually Transmitted Diseases. Published by Lippincott Williams & Wilkins.



https://caepidata.shinyapps.io/CASTD/

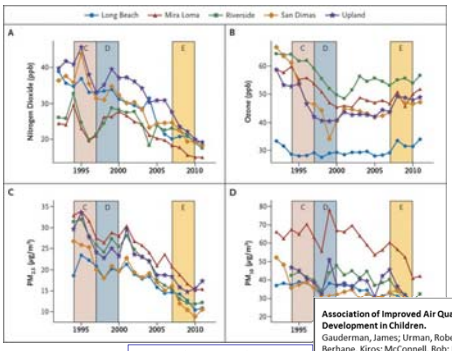


Figure 1 Levels of Four Air Pollutants from 1994 to 2011 in Five Southern California Communities. Colored bands represent the relevant 4-year averaging period for the analysis of lung-function growth in each of the three cohorts, C, D, and E. PM2.5 denotes particulate matter with an aerodynamic diameter of less than 2.5 [mu]m, and PM10 particulate matter with an aerodynamic diameter of less than 10 [mu]m.

Association of Improved Air Quality with Lung Development in Children.
Gauderman, James; Urman, Robert; Avol, Edward; Berhane, Kiros; McConnell, Rob; Rappaport, Edward; Chang, Roger; Lurmann, Fred; Gilliland, Frank

New England Journal of Medicine. 372(10):905-913, March 5, 2015.
DOI: 10.1056/NEJMoa1414123

Edward Tufte

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

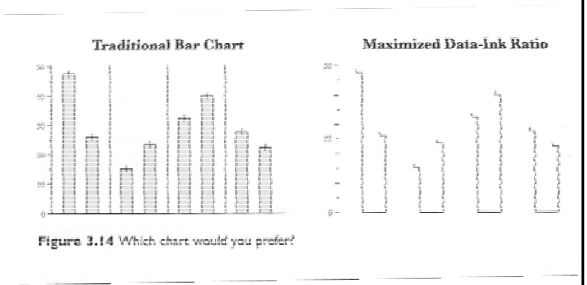
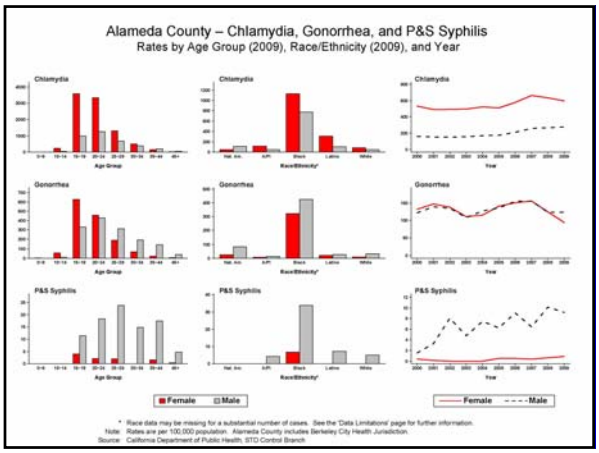
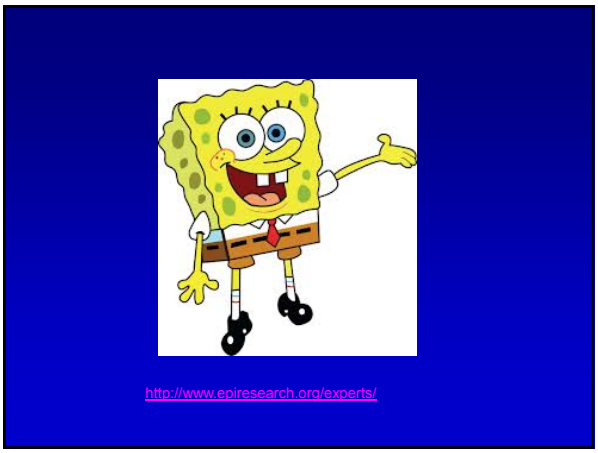


Figure 3.14 Which chart would you prefer?





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

Display Types

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

Tables

Been in a jail or prison in the past 12 months
CA Gonorrhea Cases - 2004

Jurisdiction	Yes %
Alameda	18.8%
Fresno	20.9%
Kern	20.5%
Long Beach	8.5%
Orange	22.2%
San Bernardino	9.0%
Total	100.0%

STD Control Branch

Been in a jail or prison in the past 12 months
CA Gonorrhea Cases - 2004

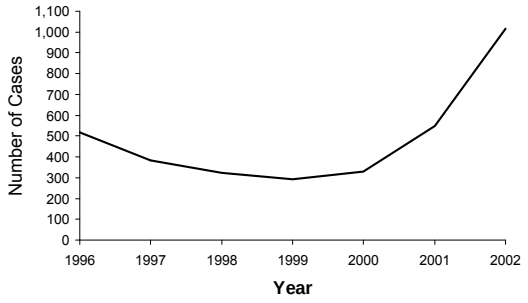
Jurisdiction	N	%
Alameda	207	21.3%
Fresno	203	24.1%
Kern	199	24.1%
Long Beach	201	10.0%
Orange	432	12.0%
San Bernardino	84	25.0%
Total	1326	17.6%

STD Control Branch

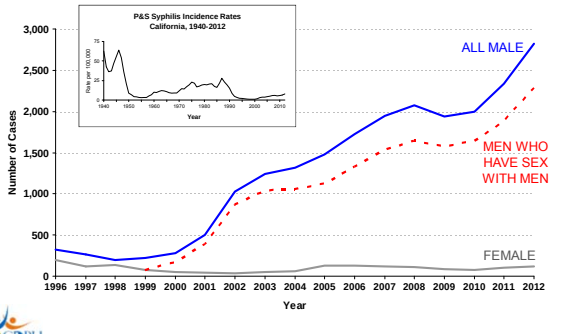
Line Graph

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

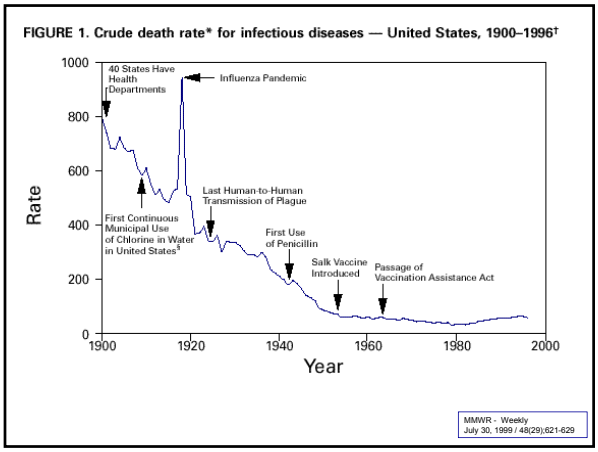
Primary & Secondary Syphilis Cases
California, 1996–2002



Primary & Secondary Syphilis, Cases by Gender
California, 1996–2012

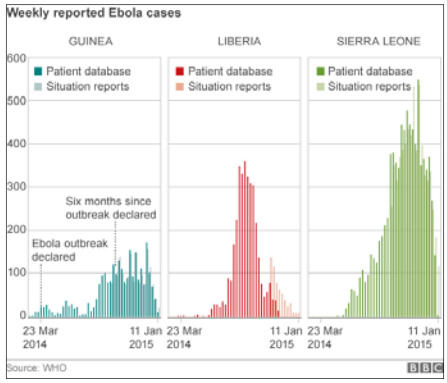


Rev. 8/2013
STD Control Branch

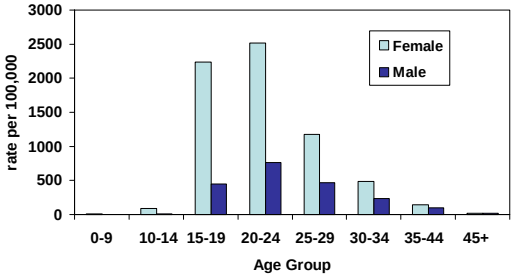


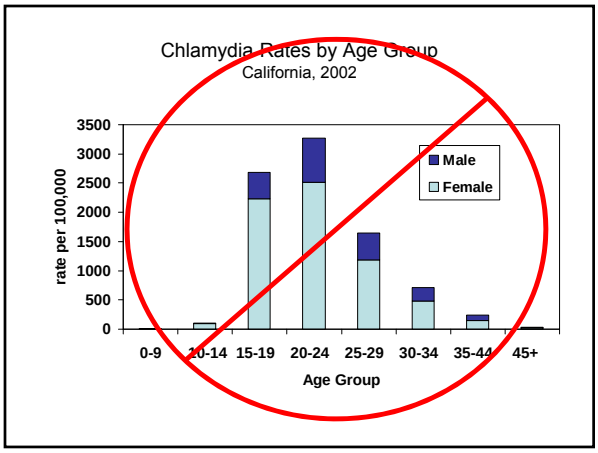
Bar Chart

- 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



http://www.bbc.com/news/health-30932578





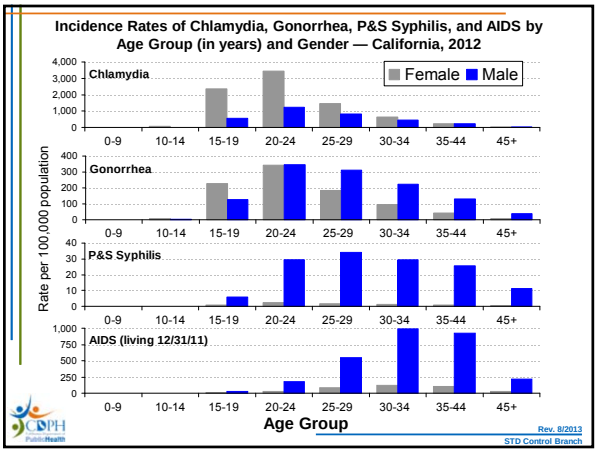
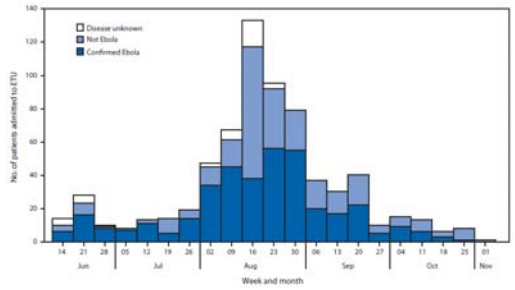


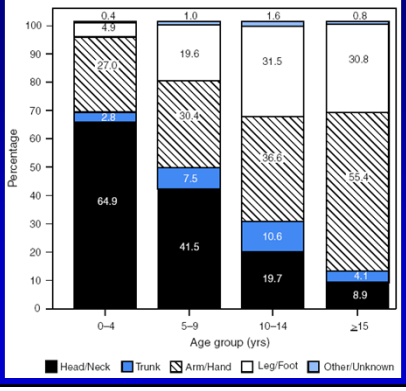
FIGURE 2. Number of patients admitted to an Ebola Treatment Unit (ETU) operated by Médecins Sans Frontières in the town of Foya, by week and final classification — Lofa County, Liberia, June 8–November 1, 2014



MMWR
Evidence for a Decrease in Transmission of Ebola Virus — Lofa County, Liberia, June 8–November 1, 2014
Early Release
November 14, 2014 / 63(Early Release);1-5

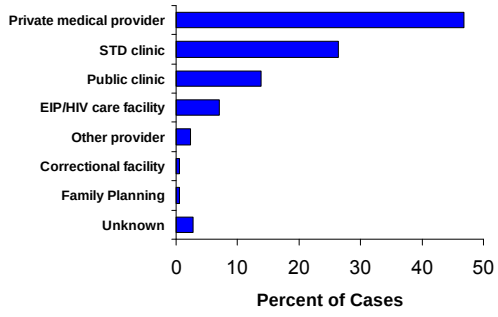
STD Control Branch

FIGURE 2. Percentage of nonfatal dog bite–related injuries treated in U.S. hospital emergency departments, by primary body part affected and age group — National Electronic Injury Surveillance System-All Injury Program, United States, 2001



MMWR
July 4, 2003 / 52(26);605-610

Reporting Source-
MSM P&S Syphilis Cases 2000-2001, California

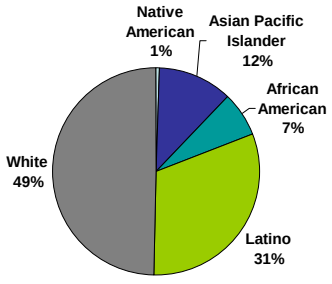


11/2002 **Provisional** Data - CA DHS STD Control Branch

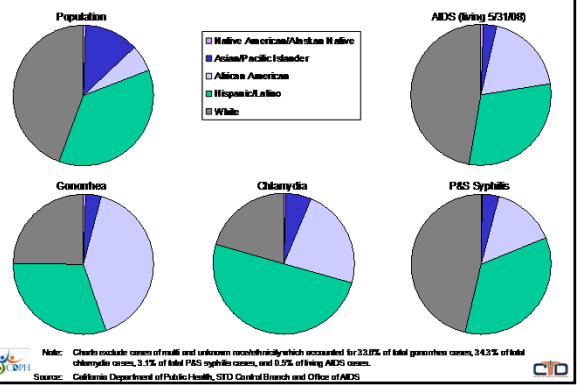
Pie Chart

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

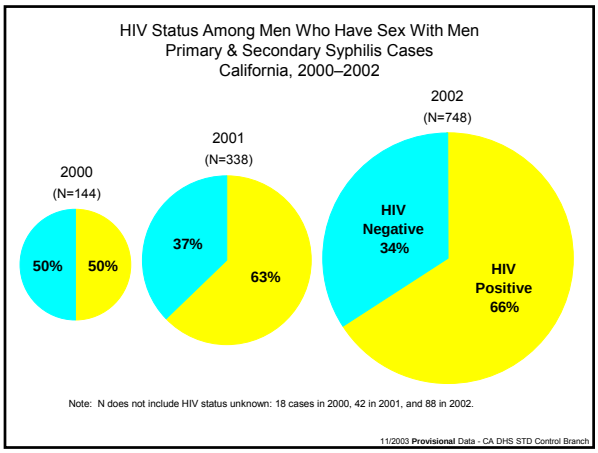
California Population by Race/Ethnicity
2001 CA Dept. of Finance Estimates



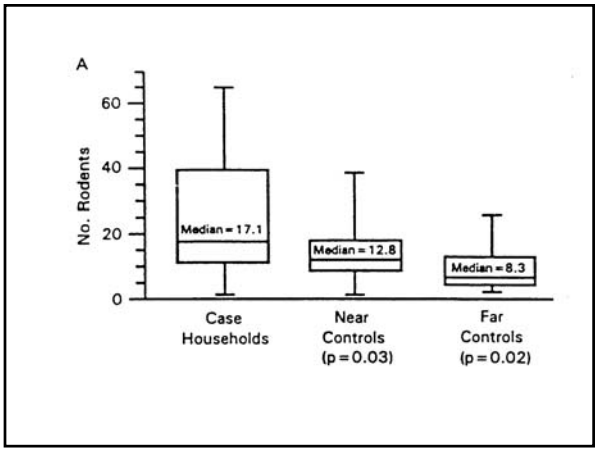
Distribution of Population and Cases by Race/Ethnicity
California, 2007



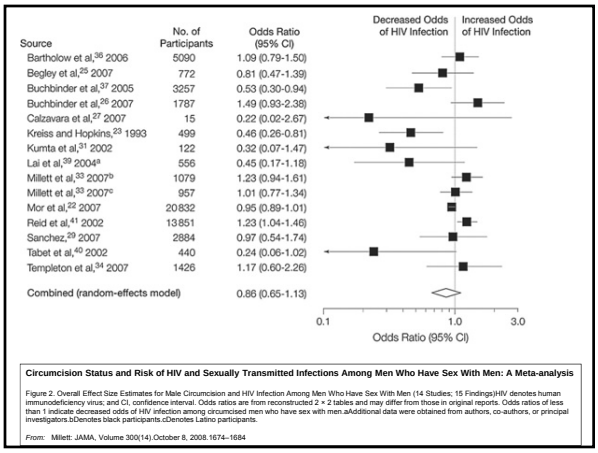
Note: Charts exclude cases of multi and unknown race/ethnicity which accounted for 33.8% of total gonorrhea cases, 34.3% of total chlamydia cases, 3.1% of total P&S syphilis cases, and 0.5% of living AIDS cases.
Source: California Department of Public Health, STD Control Branch and Office of AIDS



Box Plots



Other Statistical Graphics



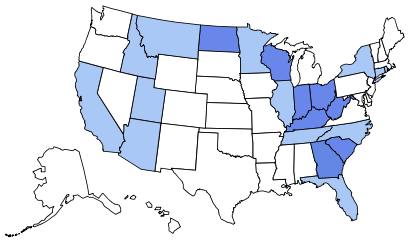
Circumcision Status and Risk of HIV and Sexually Transmitted Infections Among Men Who Have Sex With Men: A Meta-analysis

Figure 2. Overall Effect Size Estimates for Male Circumcision and HIV Infection Among Men Who Have Sex With Men (14 Studies; 15 Findings)HIV denotes human immunodeficiency virus; and CI, confidence interval. Odds ratios are from reconstructed 2 × 2 tables and may differ from those in original reports. Odds ratios of less than 1 indicate decreased odds of HIV infection among circumcised men who have sex with men.aAdditional data were obtained from authors, co-authors, or principal investigators.bDenotes black participants.cDenotes Latino participants.

From: Millett: JAMA, Volume 300(14).October 8, 2008.1674–1684

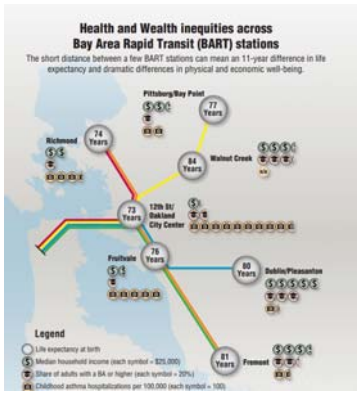
Maps

Obesity Trends* Among U.S. Adults
BRFSS, 1985
(*BMI ≥30, or ~ 30 lbs. overweight for 5' 4" person)



Slides

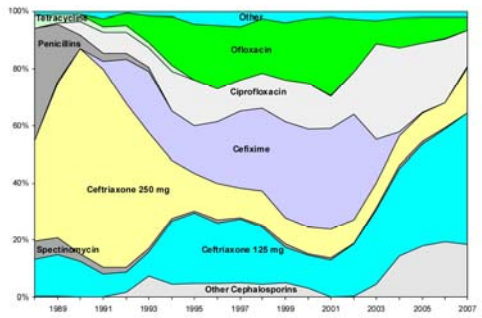
http://www.cdc.gov/nccdphp/dnpa/obesity/trend/maps/



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

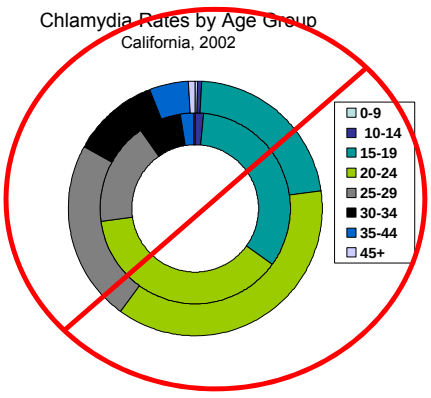
Many Other Types
and
Hybrids

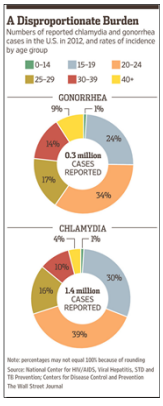
Figure 11. Drugs used to treat gonorrhea in GISP participants, 1988–2007



Note: For 2007, "Other" includes no therapy (0.8%), azithromycin 2 g (0.8%), levofloxacin (0.2%), and other less frequently used drugs.

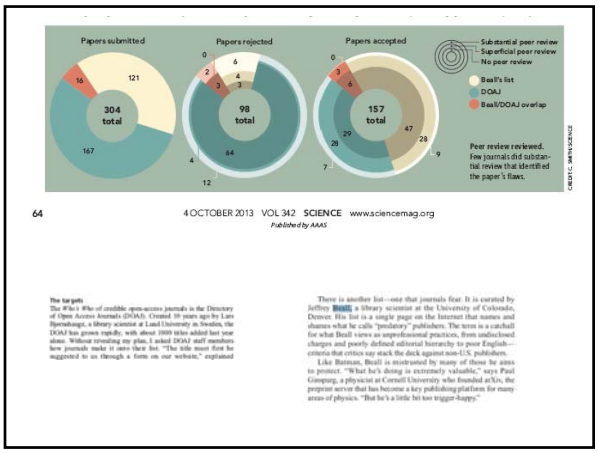
Chlamydia Rates by Age Group
California, 2002

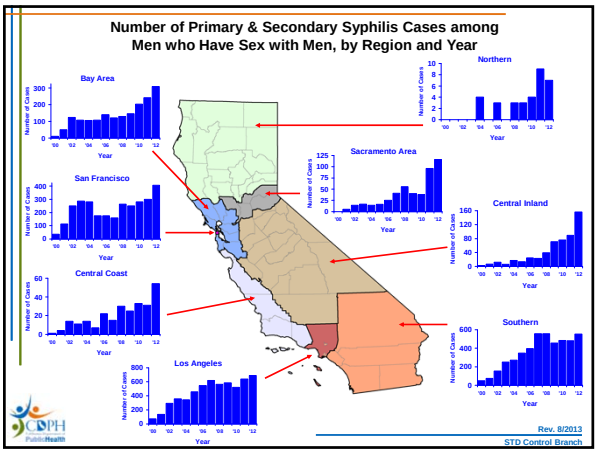




THE WALL STREET JOURNAL.

Many Pediatricians Are Skipping the Sex Talk With Teens
Poor Training, Time Constraints Cause Doctors to Fall Short on Teen Sexual Health
By JEANNE WHALEN
Sept. 14, 2014 4:00 p.m. ET





"Nut and Bolts"

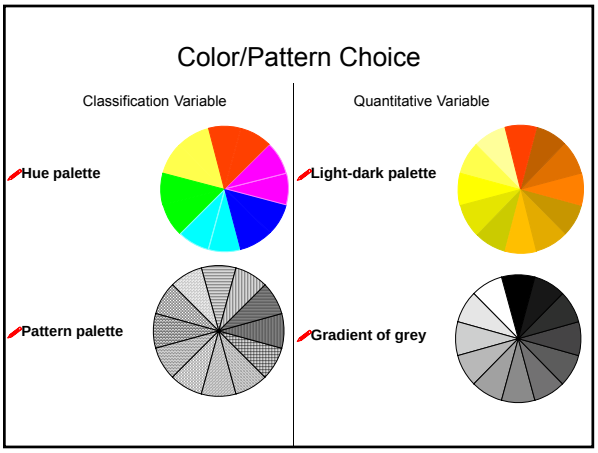
- Scale and Proportion
- Labels and Legends
- Grid Lines
- Color
- Animation/"PowerPoint"
- Font
- 3D
- Production/Reproduction
- Chart Junk
- Software

Color

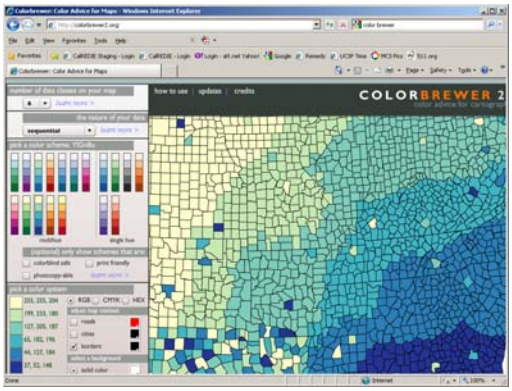
Color

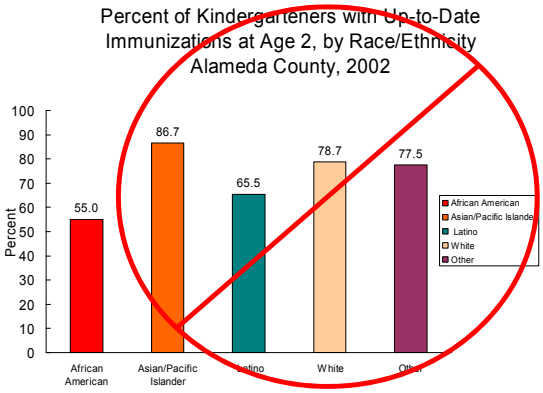
- Use for a reason
- Use nice colors
  - Shades of Blue
  - Shade of Yellow
  - Colors of Nature
- Use color sparingly
- RED can be good for Main Point, if used sparingly
- Red often does not project well with slides and LCDs
- Use consistent colors (and fonts, etc.)

Color/Pattern Choice



http://colorbrewer2.org/





Leading causes of death in Los Angeles County, 2005.[1]

Rank	Cause of death	No. of deaths	Premature death rank
1.	Coronary heart disease	15,154	1.
2.	Stroke	3,775	8.
3.	Lung cancer	3,036	5.
4.	Emphysema/COPD	2,770	13.
5.	Pneumonia/influenza	2,333	19.
6.	Diabetes	2,305	7.
7.	Alzheimer's disease	1,546	50.
8.	Colorectal cancer	1,409	12.
9.	Breast cancer	1,174	10.
10.	Homicide	1,066	2.

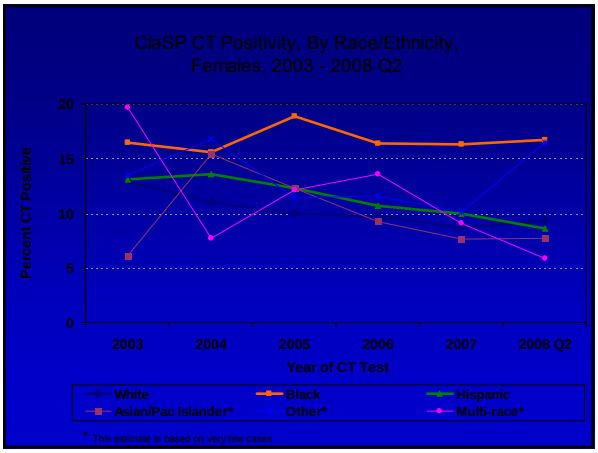
Rank	Cause of death	Years of life lost	Death rank
1.	Coronary heart disease	64,231	1.
2.	Homicide	48,067	10.
3.	Motor vehicle crash	32,394	12.
4.	Suicide	19,375	17.
5.	Lung cancer	18,490	3.
6.	Liver disease	17,897	11.
7.	Diabetes	16,971	6.
8.	Stroke	16,442	2.
9.	Drug overdose	16,138	21.
10.	Breast cancer	13,610	9.

Background Colors

and oral presentations

Background Colors

And Slide "Design Templates"



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

Line Type

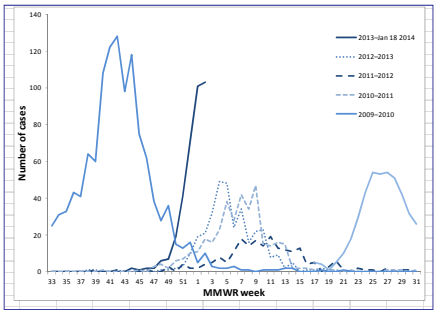
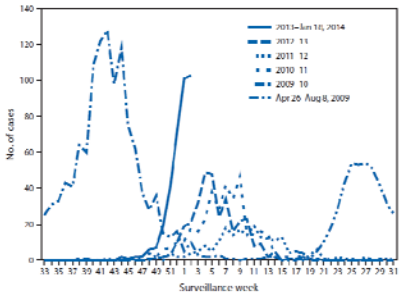


FIGURE 1. Number of cases of severe influenza,* by week of symptom onset — California, April 26, 2009–January 11, 2014†

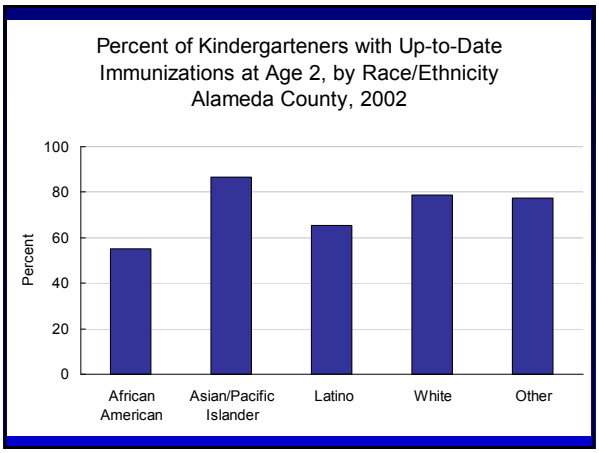


* Severe cases of influenza are defined as influenza infections resulting in intensive care unit (ICU) admission or death.

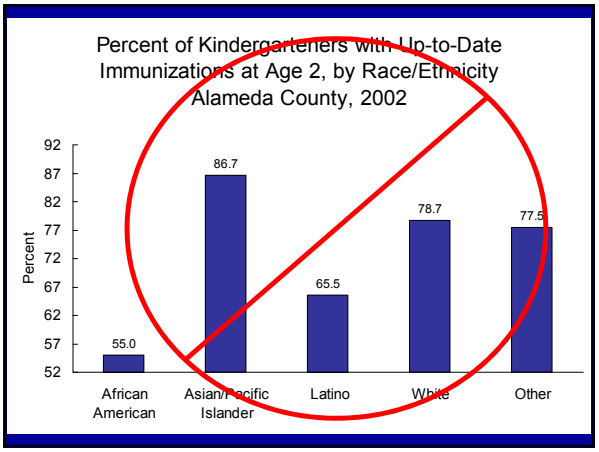
† ICU cases from three large local health jurisdictions have not been fully reported yet for the 2013–14 influenza season; for comparability, their ICU data are excluded from all years in this figure. Only cases occurring through January 11, 2014, are included because reporting for the cases with onset in the week ending January 18, 2014 was incomplete at the time of this report.

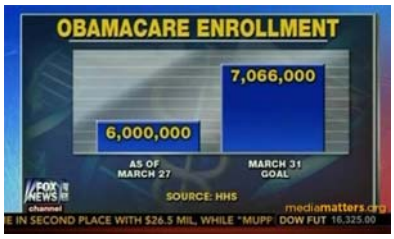
Alternate Text: The figure above shows the number of cases of severe influenza, by week of symptom onset in California during April 26, 2009–January 11, 2014. The 405 severe influenza cases were reported from 41 of 61 local health jurisdictions (67%) in California and had onset dates of October 20, 2013– January 15, 2014. The largest number of severe cases (103) by week of symptom onset occurred during the week ending January 11, 2014. These represent both the highest cumu¬lative number of severe cases at this point in the influenza season and the highest number of new cases in a single week since the 2009 H1N1 pandemic.

Grid Lines

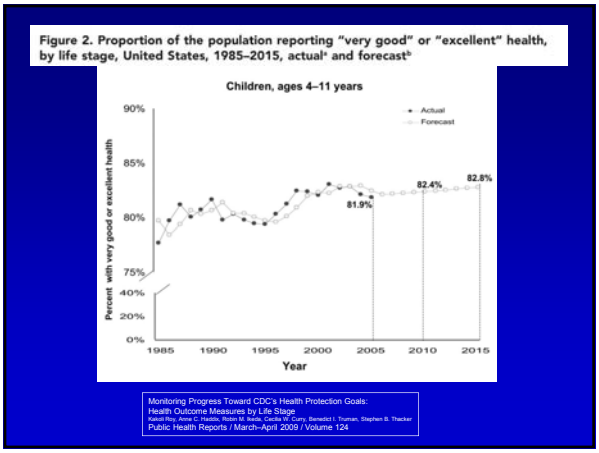


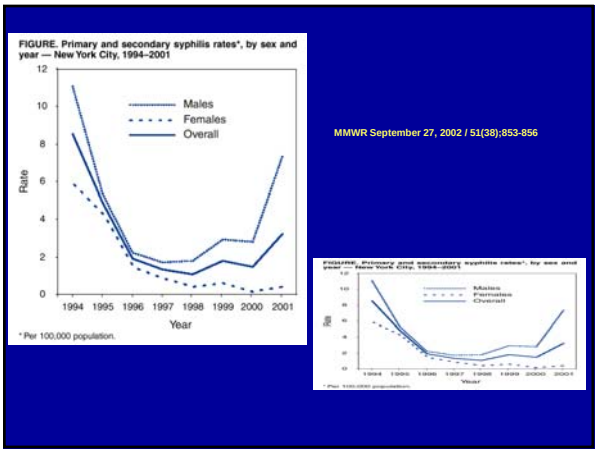
Scale

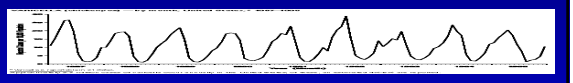




http://flowingdata.com/category/statistics/mistaken-data/







"Bank to 45 degrees"

"Golden Rectangle"

1

1.62

THE SHRINKING FAMILY DOCTOR
In California

Percentage of Doctors Devoted Solely to Family Practice

1964 — 27%
1975 — 16.0%
1990 — 12.0%

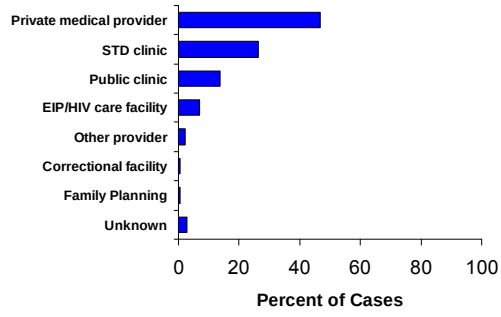
1: 2,247 RATIO TO POPULATION
8,023 Doctors

1: 3,167
6,694

1: 4,232
6,212

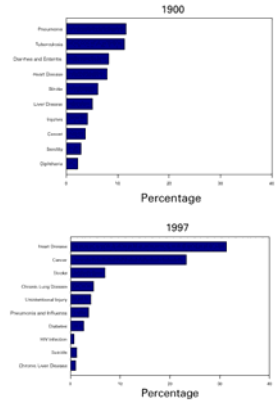
In Tufte, 1983, p. 69.
From LA Times August 5, 1979, p.3.

Reporting Source-
MSM P&S Syphilis Cases 2000-2001, California



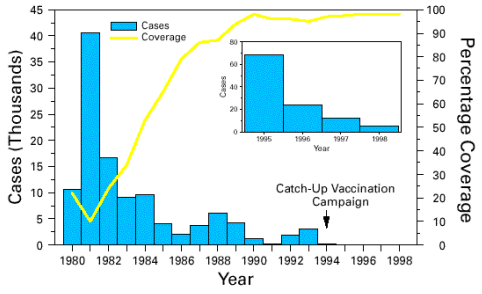
11/2002 **Provisional** Data - CA DHS STD Control Branch

FIGURE 2. The 10 leading causes of death as a percentage of all deaths — United States, 1900 and 1997



MMWR - Weekly
July 30, 1999 / 48(29);621-629

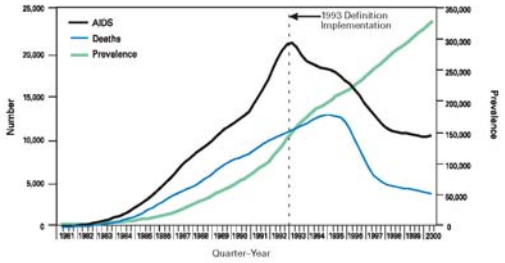
FIGURE 2. Reported number of measles cases and vaccination coverage, by year — Oman, 1980–1998



MMWR

December 03, 1999 / 48(47);1081-6

FIGURE 1. Estimated AIDS incidence*, deaths, and prevalence, by quarter-year of diagnosis/death — United States, 1981–2000

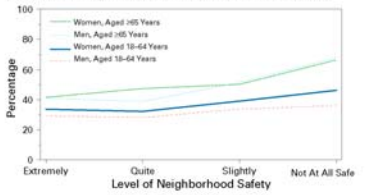


MMWR

HIV and AIDS --- United States, 1981–2000
June 01, 2001 / 50(21);430-4

Physical Inactivity — Continued

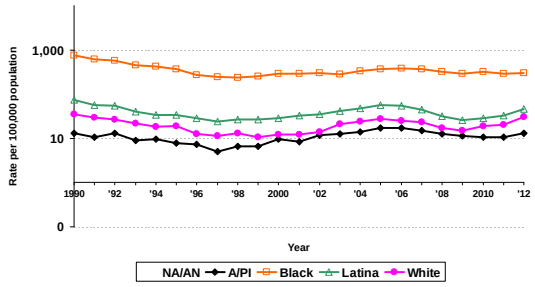
FIGURE 1. Percentage of respondents who reported physical inactivity, by sex, age group, and perceived neighborhood safety level — Maryland, Montana, Ohio, Pennsylvania, and Virginia, Behavioral Risk Factor Surveillance System, 1996



MMWR

Neighborhood Safety and the Prevalence of Physical Inactivity -- Selected States, 1996
February 26, 1999 / 48(07);143-146

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

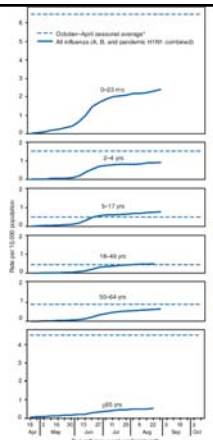


FIGURE 2. Laboratory-confirmed influenza hospitalization rates per 10,000 population, by age group and surveillance week --- Emerging Infections Program, United States, April--August 2009

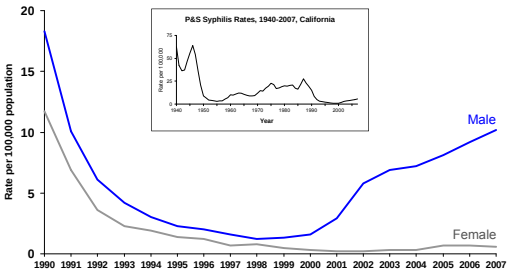
*Age group–specific average influenza rate, October–April, 2005–06, 2006–07, and 2007–08 influenza seasons.

Alternative Text: The above figure consists of six line graphs showing influenza hospitalization rates per 10,000 population by age group and surveillance week in the United States during April–August 2009 and comparing these rates with the October–April seasonal average for the 2005–06, 2006–07, and 2007–08 influenza seasons. The graphs indicate that the April–August 2009 rates, resulting from hospitalizations for 2009 pandemic influenza A (H1N1) were generally lower or similar to the seasonal average. The graphs also indicate that the rates for persons aged 0–23 months and ≥65 years were several times lower for pandemic H1N1 than for seasonal influenza.

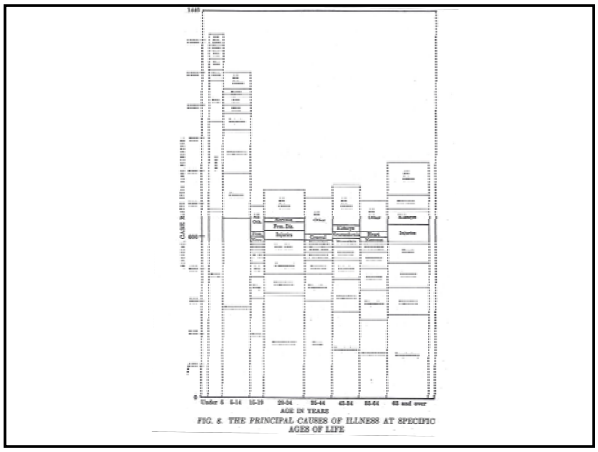
MMWR - ***Early Release*** **September 10, 2009 / 58(Early Release);1-4**

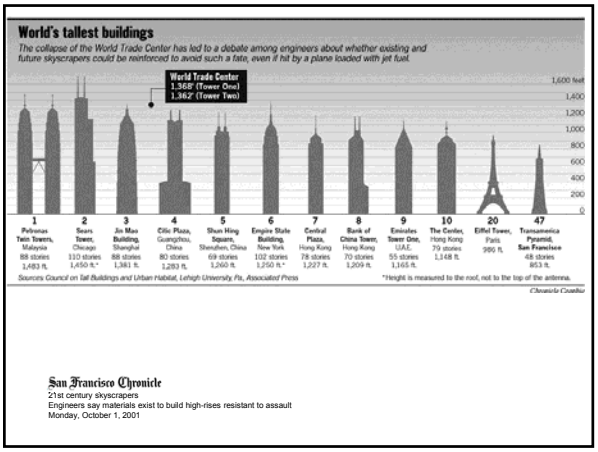
Labels and Legends

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



CA DPH STD Control Branch (rev 7/2008)

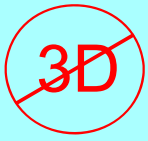




Production and Reproduction

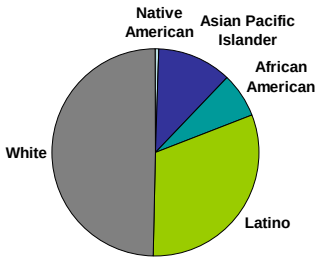
- Paper
- Transparency (from copy, back in the day)
- Transparency (now?)
- LCD (.jpg)
- …better

3D Charts

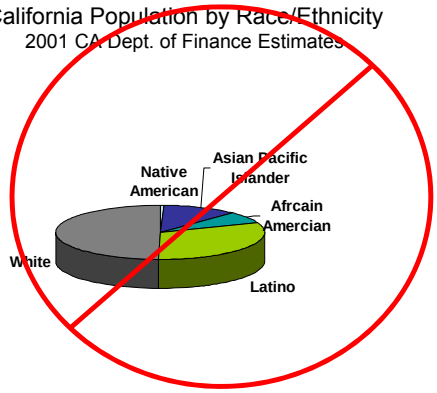


- Unless there are 3 dimensions and the audience can handle it!

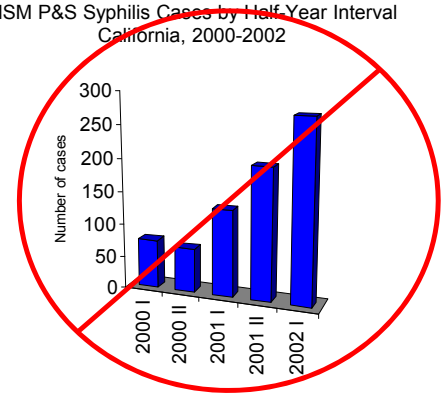
California Population by Race/Ethnicity
2001 CA Dept. of Finance Estimates



California Population by Race/Ethnicity
2001 CA Dept. of Finance Estimates

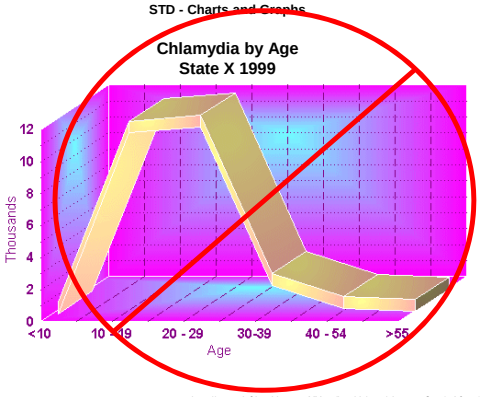


MSM P&S Syphilis Cases by Half-Year Interval
California, 2000-2002



State X Department of Health
STD - Charts and Graphs

Chlamydia by Age
State X 1999



http://www.dsf.health.state.XX.us/health/cwp/view.asp?a=178&q=198607

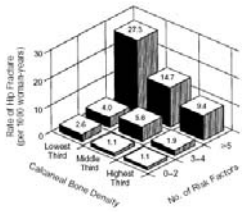
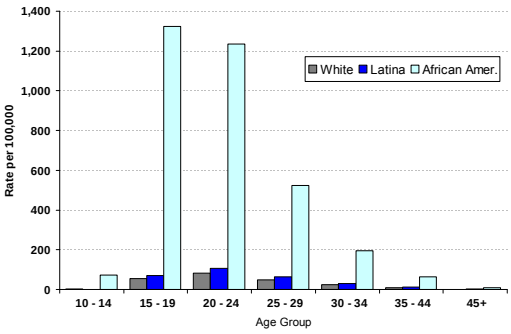


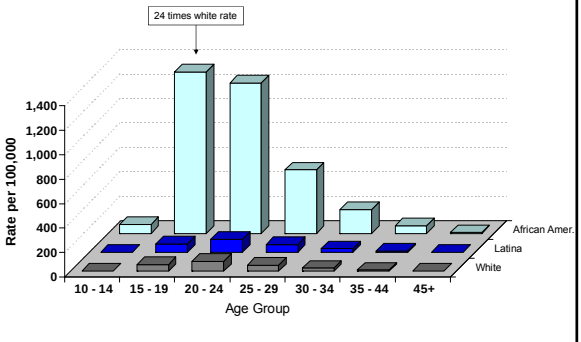
Figure 2. Annual Risk of Hip Fracture According to the Number of Risk Factors and the Age-Specific Calcaneal Bone Density.
The risk factors (from Table 2) are as follows: age >60; maternal history of hip fracture; any fracture (except hip fracture) since the age of 50; fair, poor, or very poor health; previous hyperthyroidism; anticonvulsant therapy; current long-acting benzodiazepine therapy; current weight less than at the age of 25; height at the age of 25 >168 cm; caffeine intake more than the equivalent of two cups of coffee per day; on feet <4 hours a day; no walking for exercise; inability to rise from chair without using arms; lowest quartile (standard deviation >2.44) of depth perception; lowest quartile (<0.70 unit) of contrast sensitivity; and pulse rate >80 per minute.

Gonorrhea Rates, Females, by Race/Ethnicity and Age Group
California, 2009



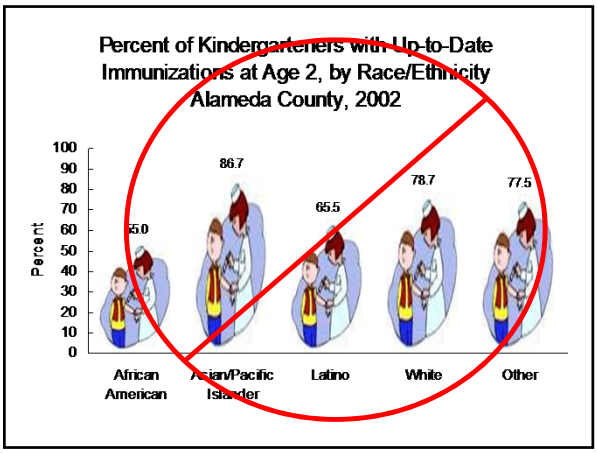
Note: Native American/Alaskan Native and Asian/Pacific Islander rates were excluded due to small case counts.
Source: California Department of Public Health, STD Control Branch

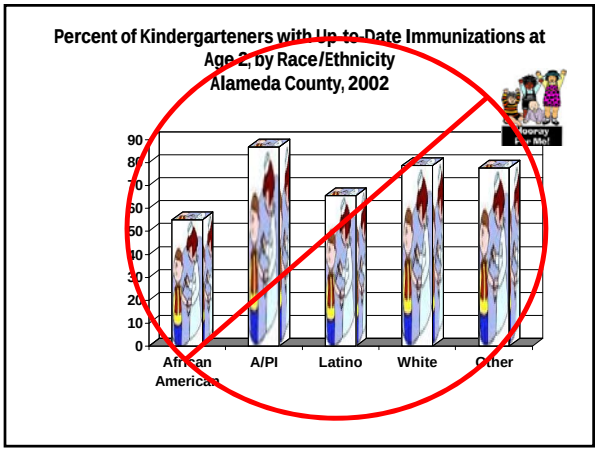
Gonorrhea Rates, Females, by Race/Ethnicity and Age Group
California, 2009



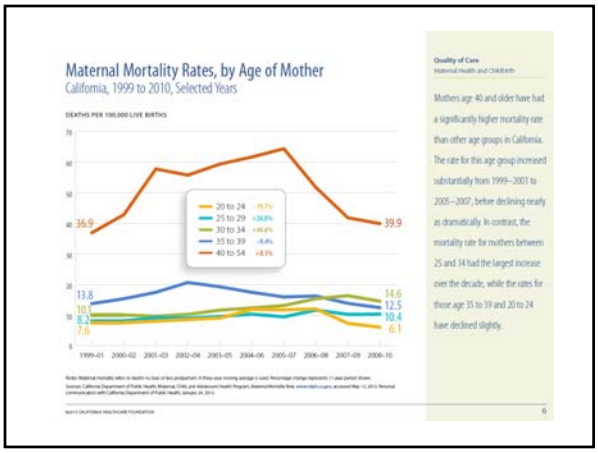
Note: Native American/Alaskan Native and Asian/Pacific Islander rates were excluded due to small case counts.
Source: California Department of Public Health, STD Control Branch

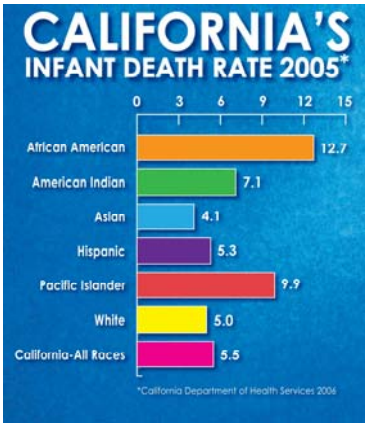
Chart Junk





"Soften the Edges"
'The Human Touch"





“Infographics”

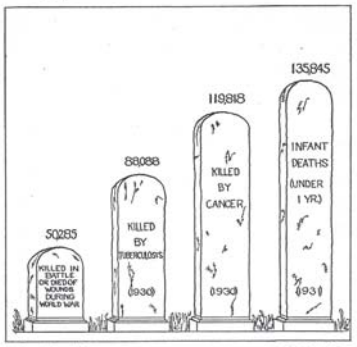
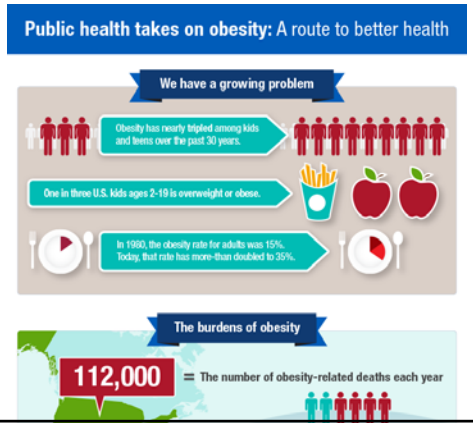
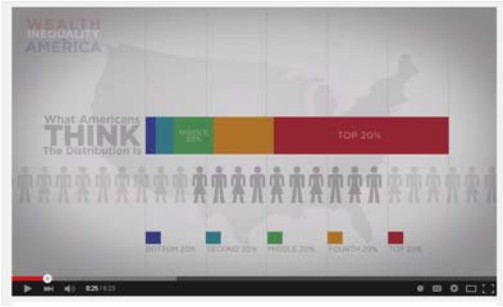


FIG. 18. “HUMAN LIFE IN THE UNITED STATES IS BEING WASTED AS RECKLESSLY, AS SURELY, IN TIMES OF PEACE AS IN TIMES OF WAR”



Stories…

http://insight.livestories.com/countyhealthrankings/Data-With-Weight/

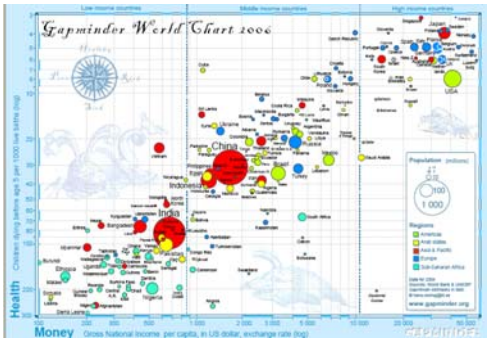


https://www.youtube.com/watch?v=QPKKQnijnsM

Almost done…

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

http://www.gapminder.org/

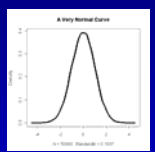


In Conclusion

- 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/pUDcGlulfW8
Part 2b: http://youtu.be/YCRyVPpz-yk