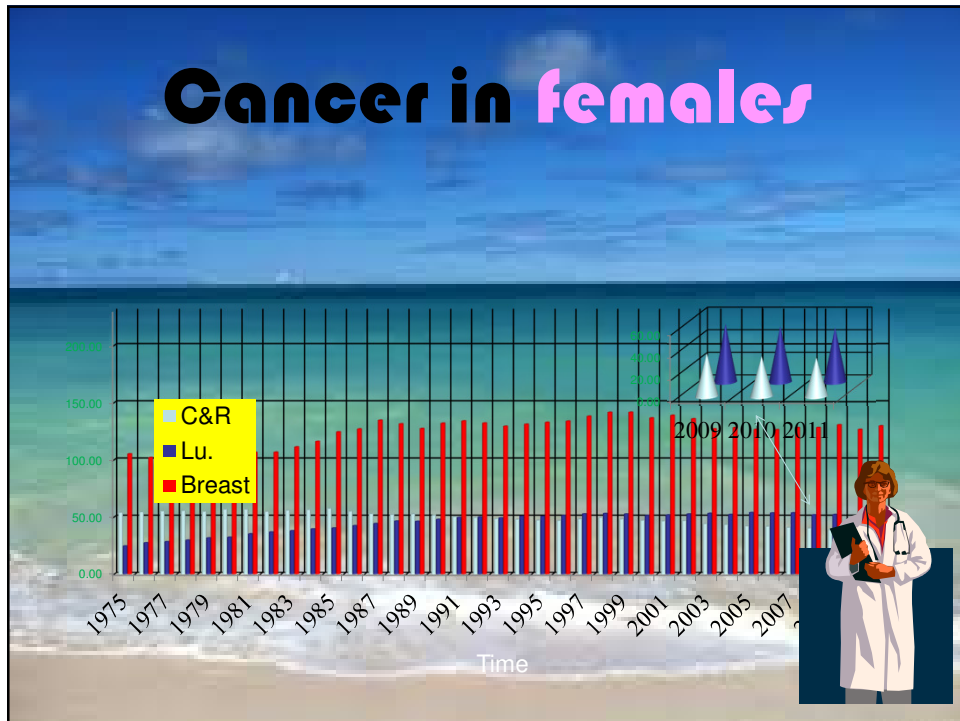
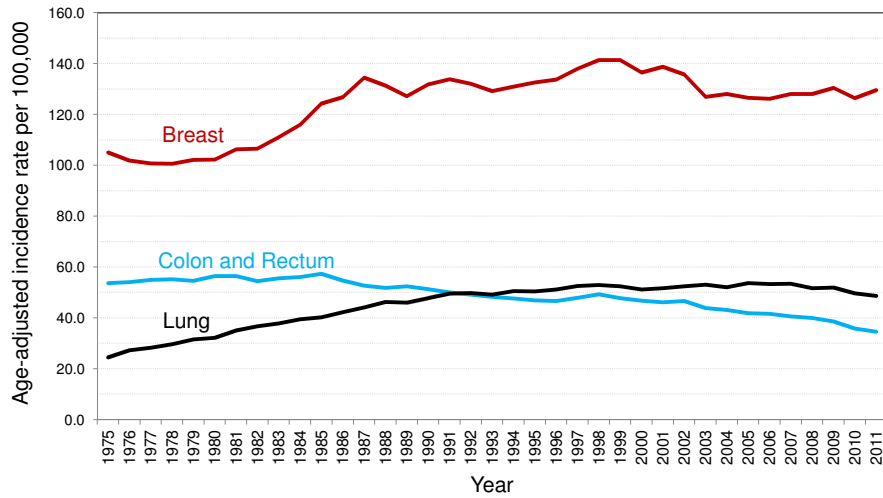


Visual Display of Public Health Data

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



Cancer Rates by Site, Females, United States 1975-2011



* Cancer sites include invasive cases only unless otherwise noted.
 * Rates are per 100,000 and are age-adjusted to the 2000 US Std Population (19 age groups - Census P25-1130). The modeled rates are the point estimates for the regression lines calculated by the Joinpoint Regression Program (Version 4.1.0, April 2014, National Cancer Institute).
 * Incidence source: SEER 9 areas (San Francisco, Connecticut, Detroit, Hawaii, Iowa, New Mexico, Seattle, Utah, and Atlanta).
<http://seer.cancer.gov/faststats/index.php> 4/15/2014

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



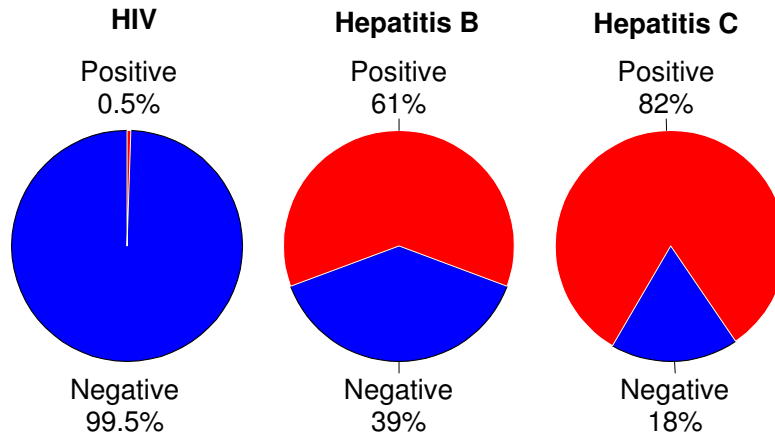
Action

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

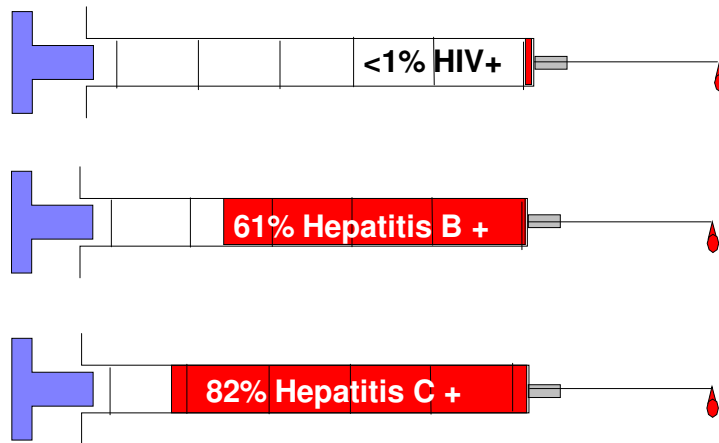
HIV and Hepatitis among Injection Drug Users New Mexico, 1997

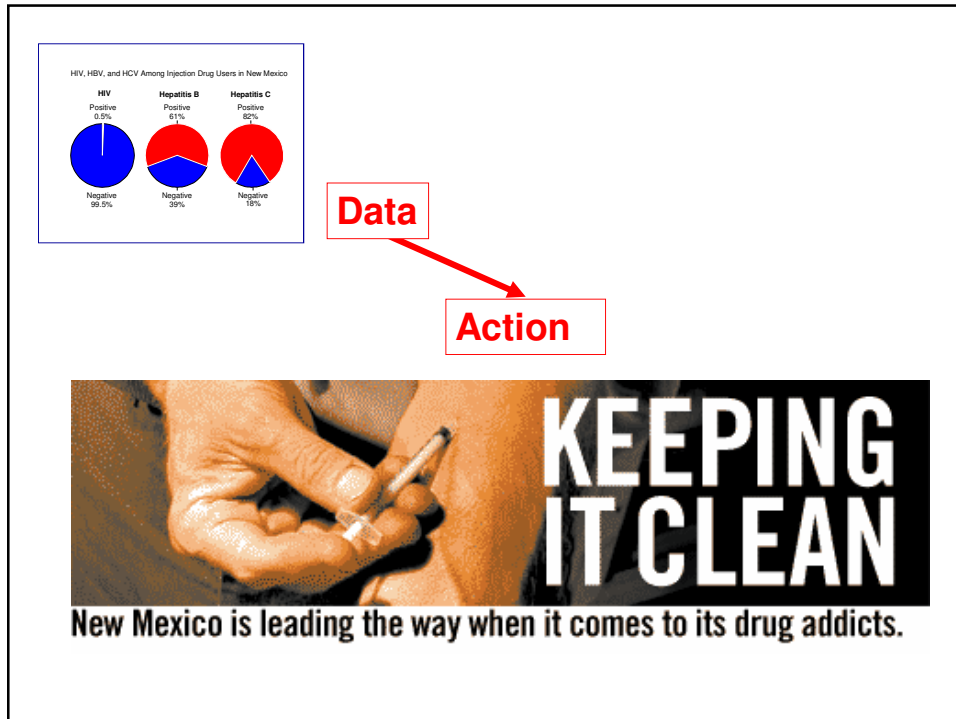
	<u>N</u> <u>tested</u>	<u>%</u> <u>Positive</u>	<u>95% C.I.</u>
HIV	1002	0.5	.16-1.6
Hep A	696	66.1	62.4-69.6
Hep B	950	61.1	57.9-64.2
Hep C	945	82.2	79.6-84.6

HIV, HBV, and HCV Among Injection Drug Users in New Mexico



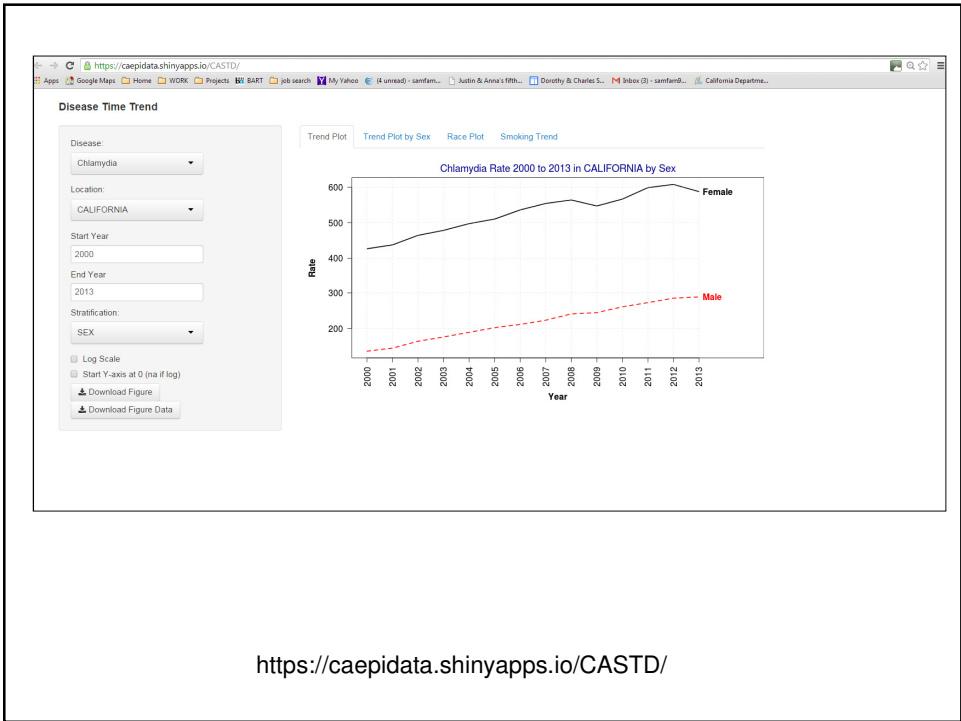
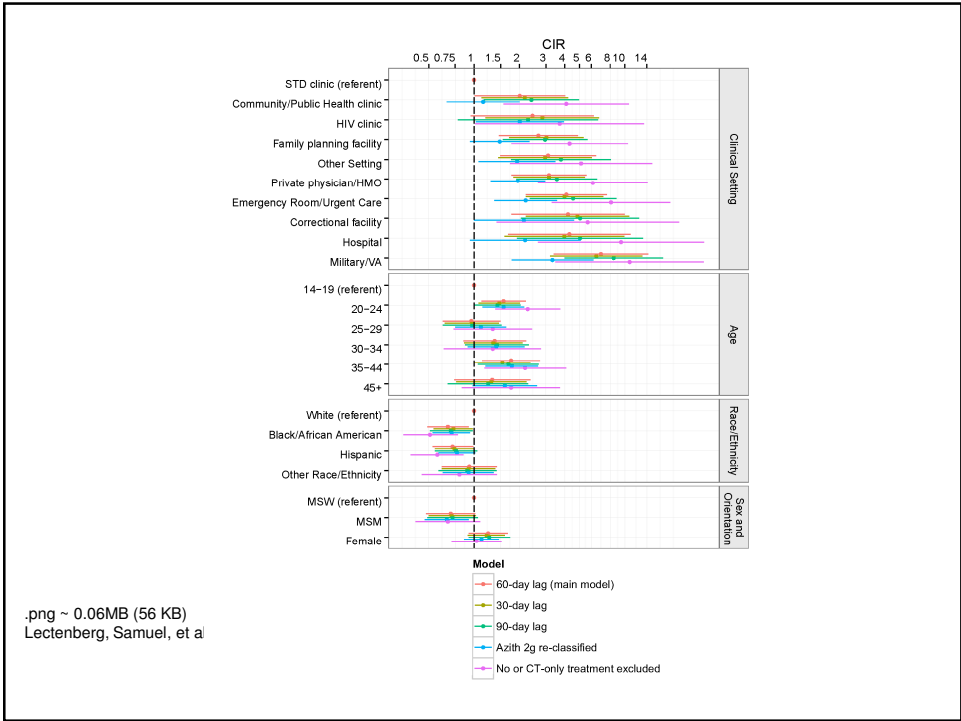
HIV, Hepatitis B, and Hepatitis C
New Mexico Street Based Injection Drug User Studies





Guidelines for Effective Visual Display

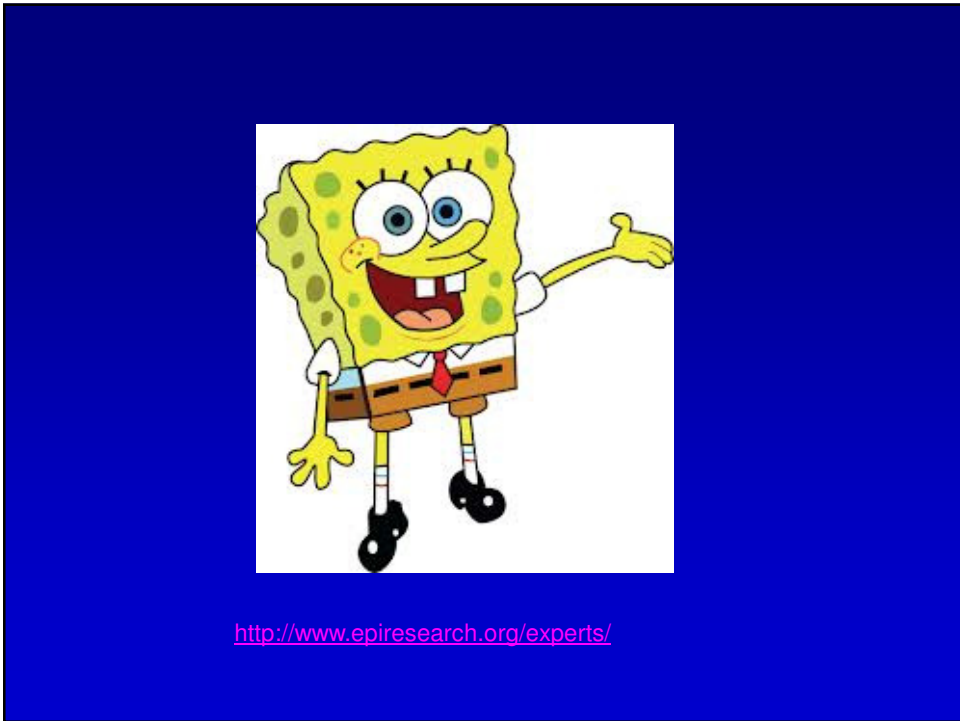
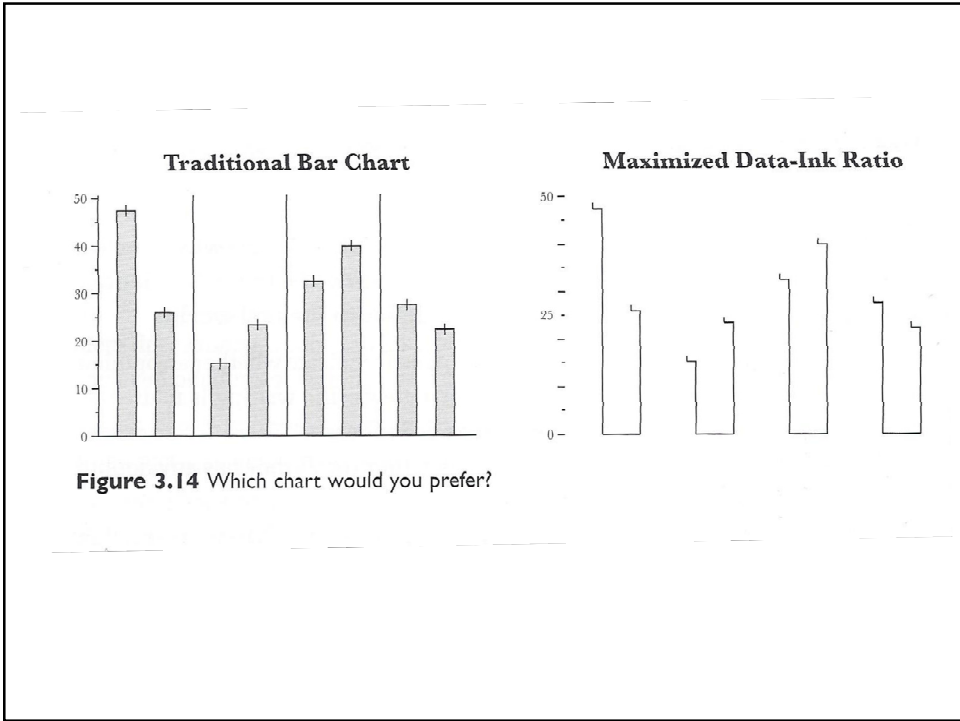
- Communicate important information
- Complexity is good, and...
- Keep it simple, stupid
- Know your audience
- Oral presentation vs. written material
- Data integrity
- Clear labels and annotations
- Use appropriate scale(s)
- Use appropriate type of chart
- Pay attention to details
- Avoid extraneous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



Edward Tufte

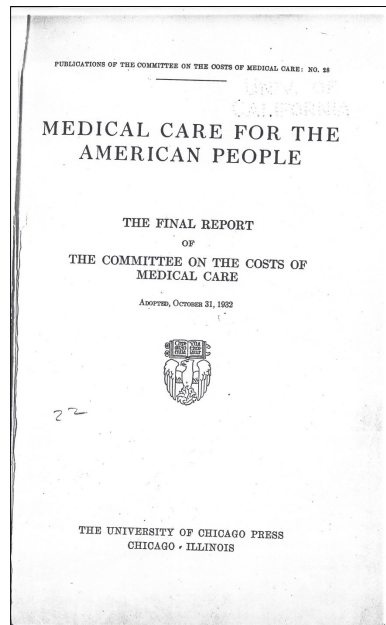
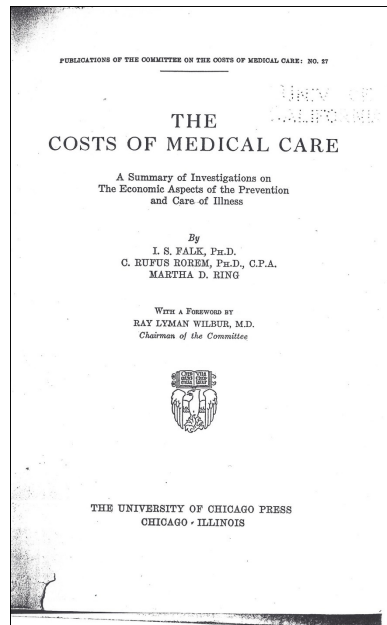
- Look at his books!
- Graphical Excellence
- The Lie Factor
- Data Density
- Less is more
- Small Multiples / Parallelism





History

<http://datavis.ca/milestones/>

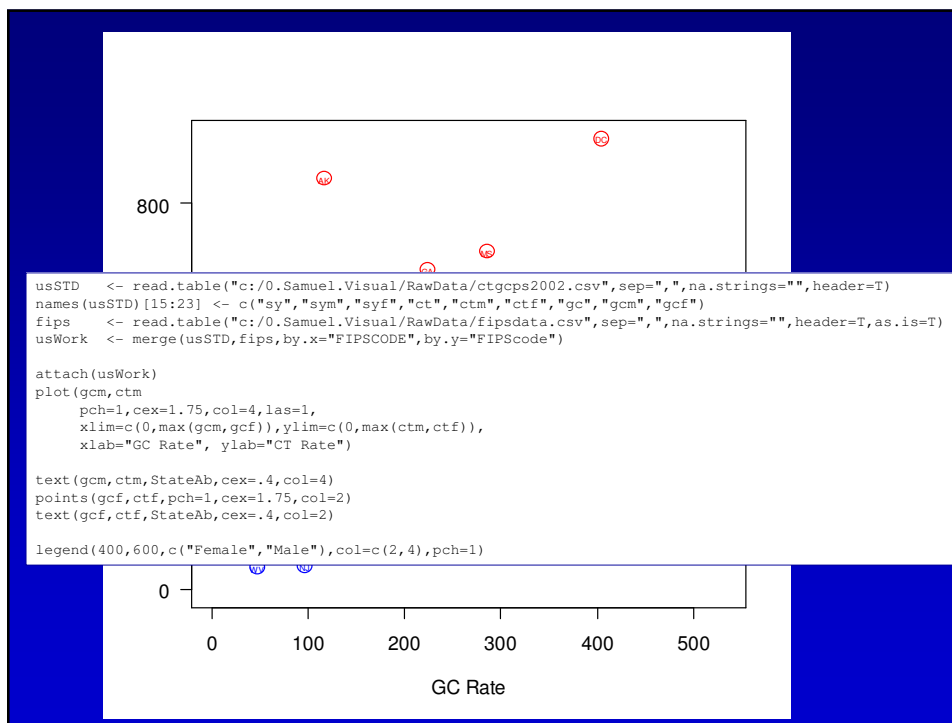


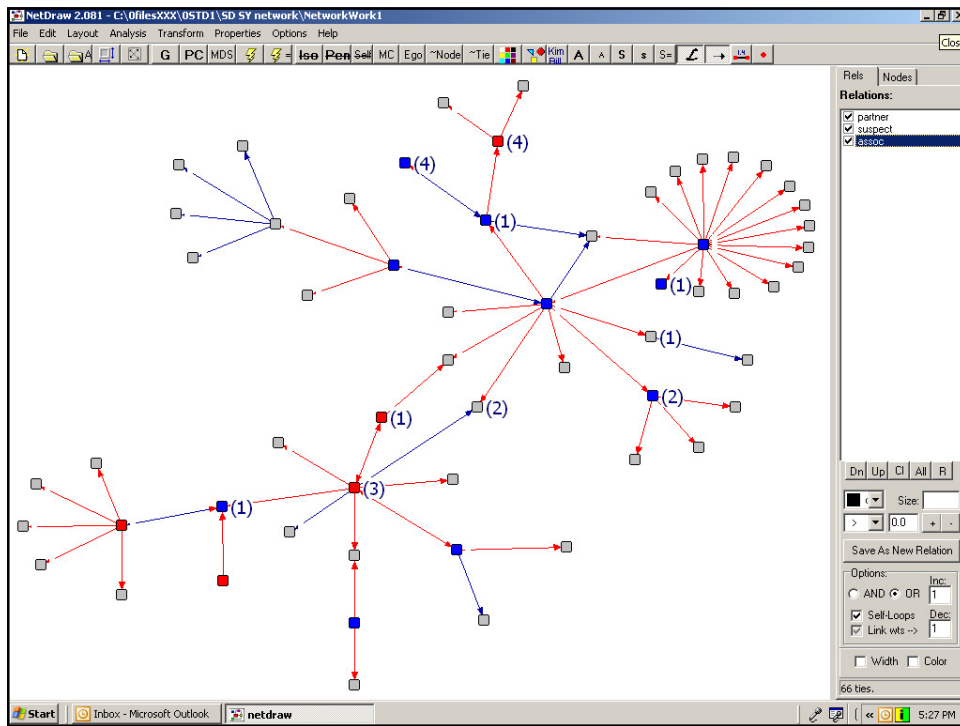
Software



Software

- **Stand alone graphics packages**
 - PowerPoint; Harvard Graphics
 - 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. "NetDraw" Network analysis





Big data

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**

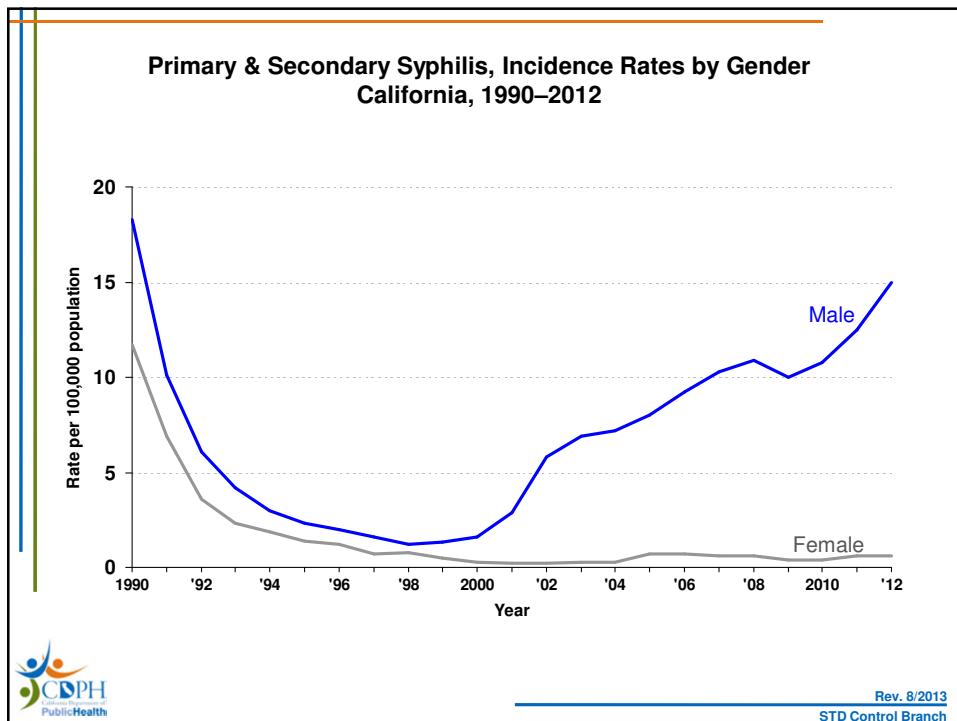
<u>Jurisdiction</u>	<u>N</u>	<u>%</u>
Alameda	207	21.3%
Fresno	203	24.1%
Kern	199	24.1%
Long Beach	201	10.0%
Orange	432	12.0%
San Bernardino	84	25.0%
Total	1326	17.6%

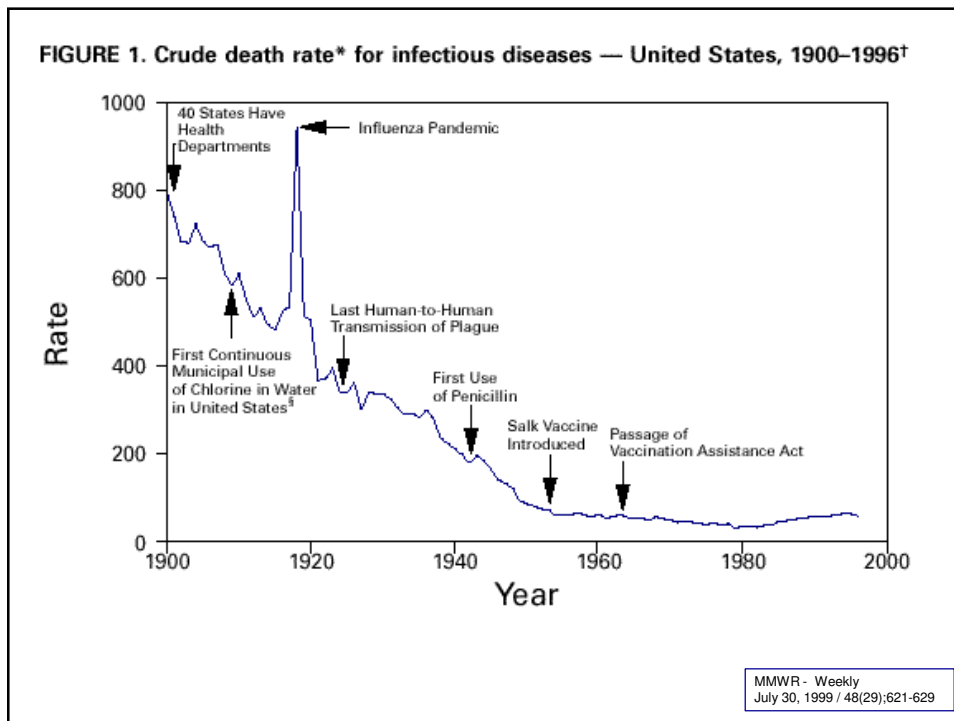
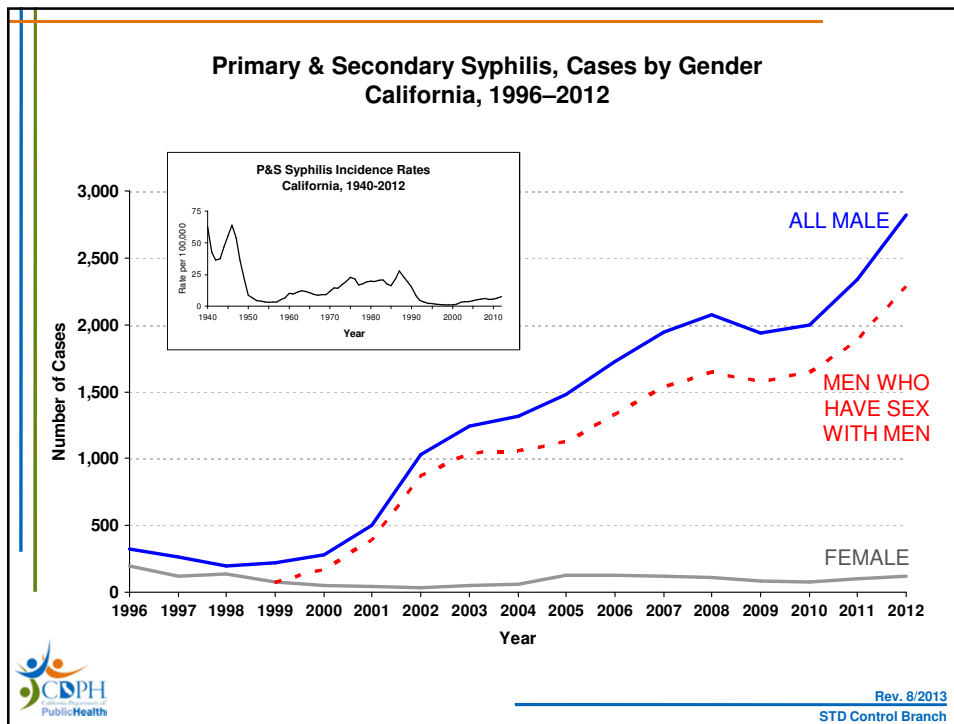


STD Control Branch

Line Graph

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

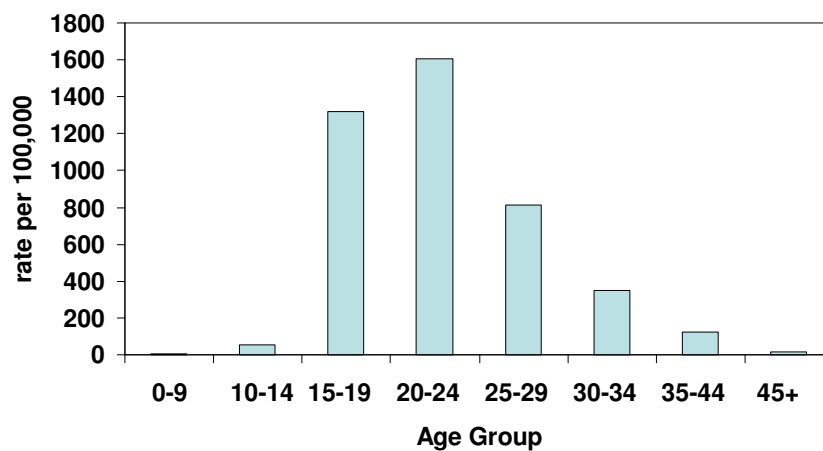




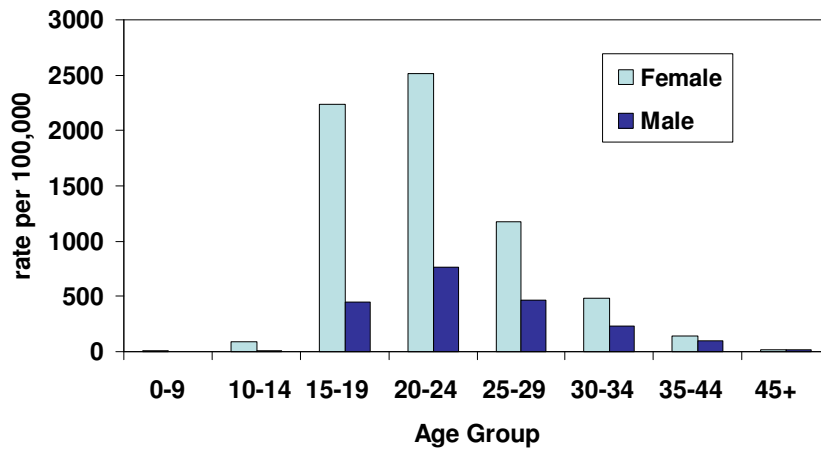
Bar Chart

- 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

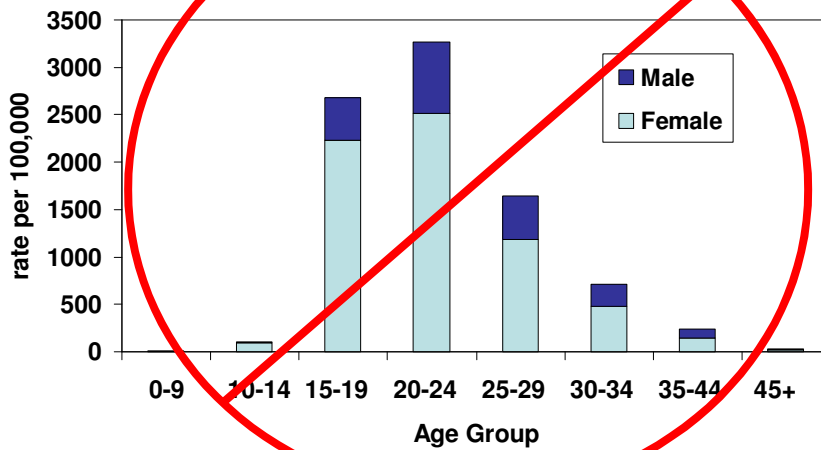
Chlamydia Rates by Age Group
California, 2002



Chlamydia Rates by Age Group
California, 2002



Chlamydia Rates by Age Group
California, 2002



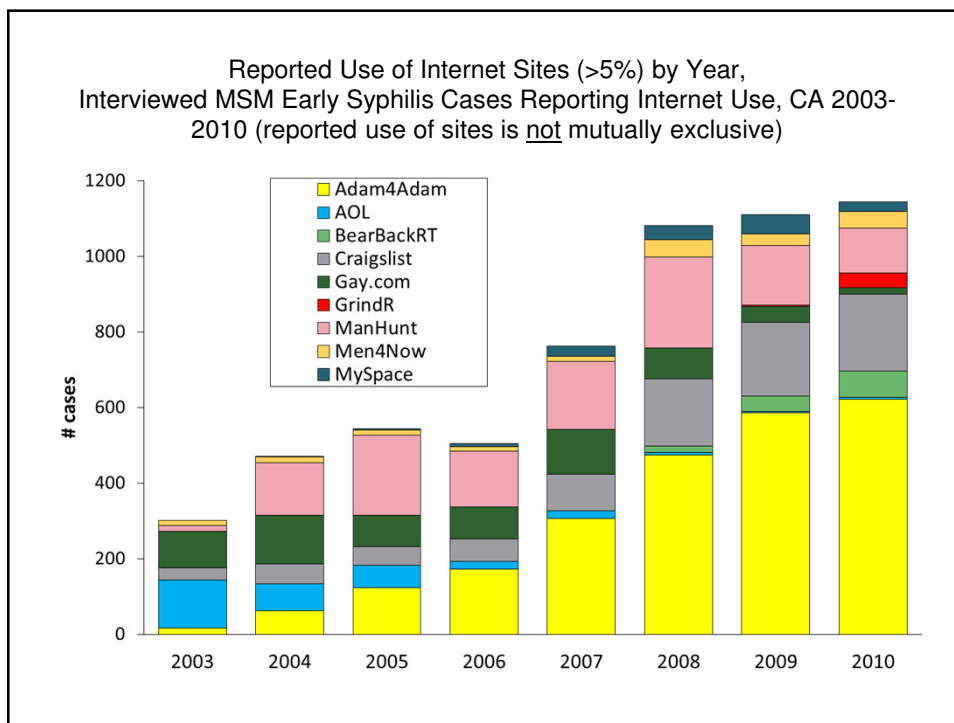
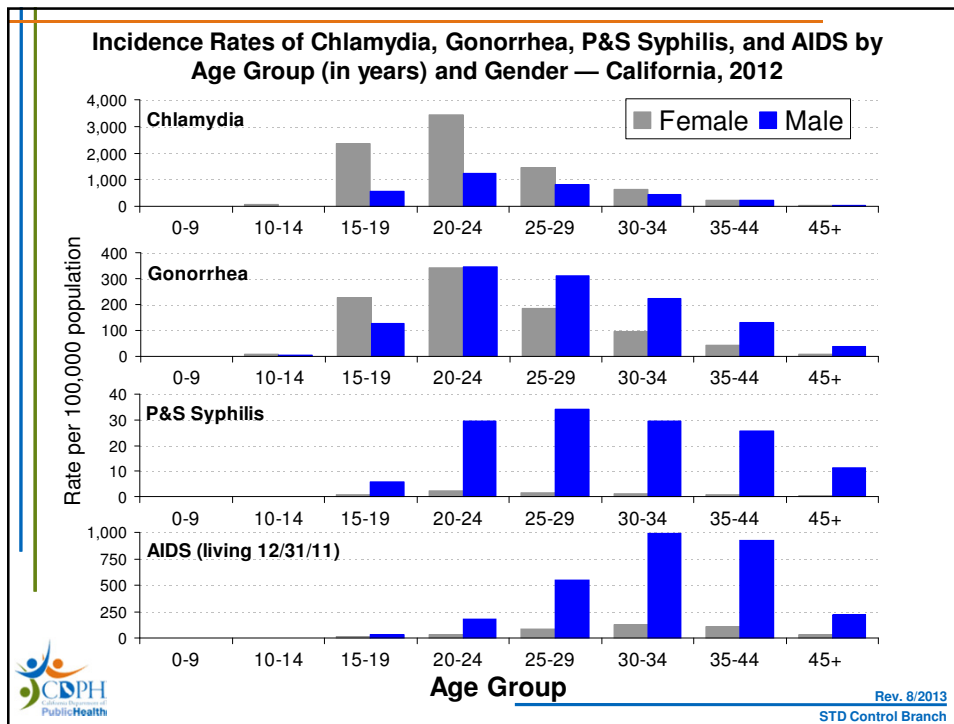
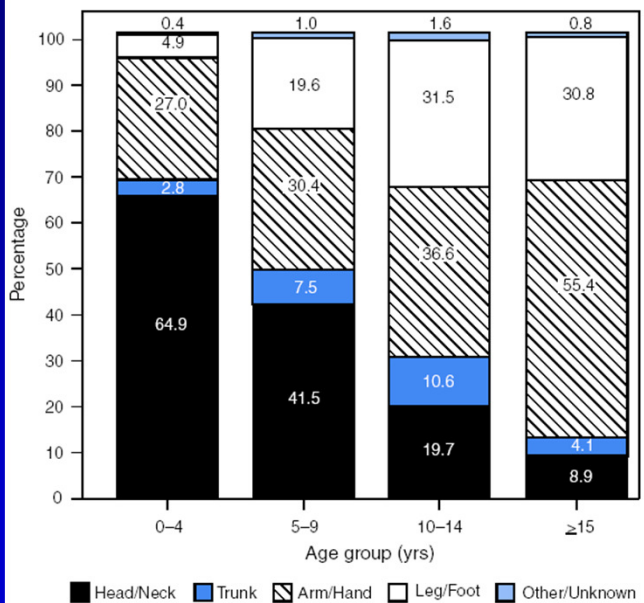
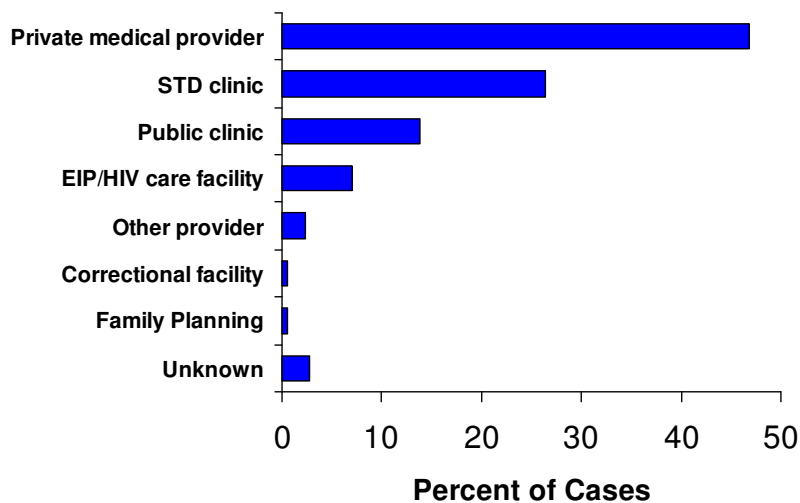


FIGURE 2. Percentage of nonfatal dog bite-related injuries treated in U.S. hospital emergency departments, by primary body part affected and age group — National Electronic Injury Surveillance System-All Injury Program, United States, 2001



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

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

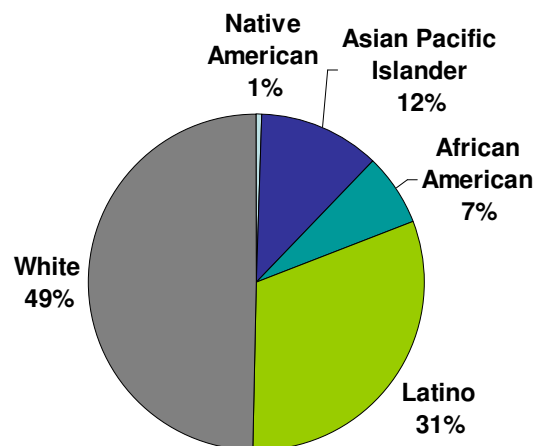


11/2002 Provisional Data - CA DHS STD Control Branch

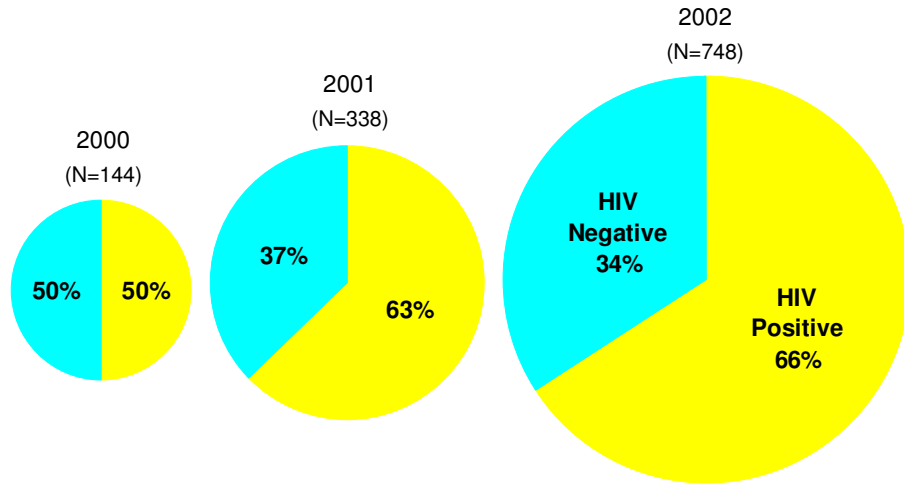
Pie Chart

- 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



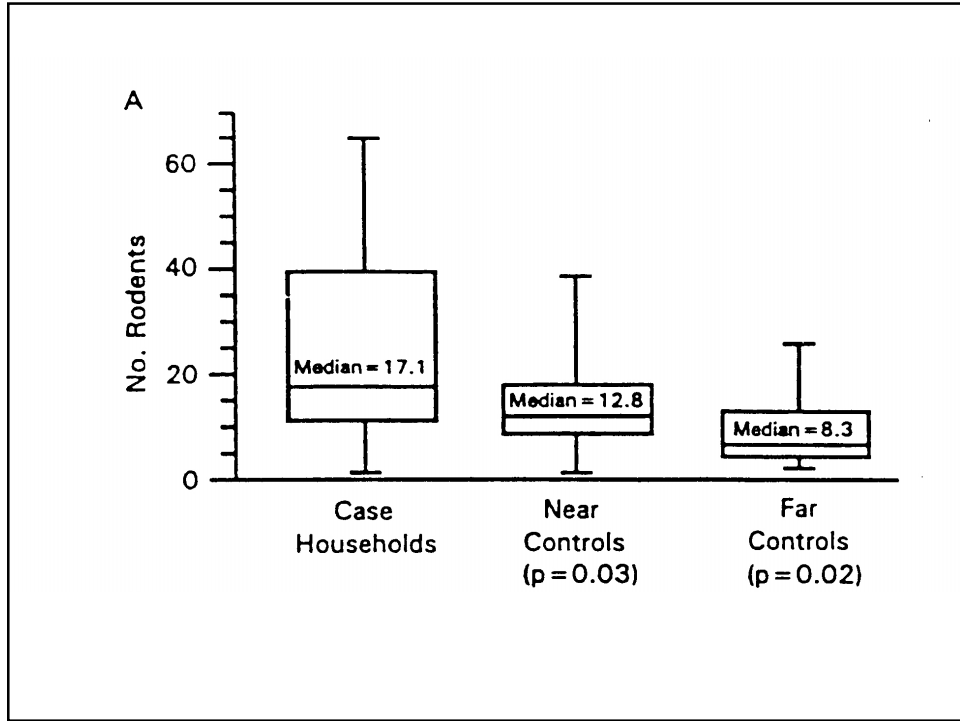
HIV Status Among Men Who Have Sex With Men
Primary & Secondary Syphilis Cases
California, 2000–2002



Note: N does not include HIV status unknown: 18 cases in 2000, 42 in 2001, and 88 in 2002.

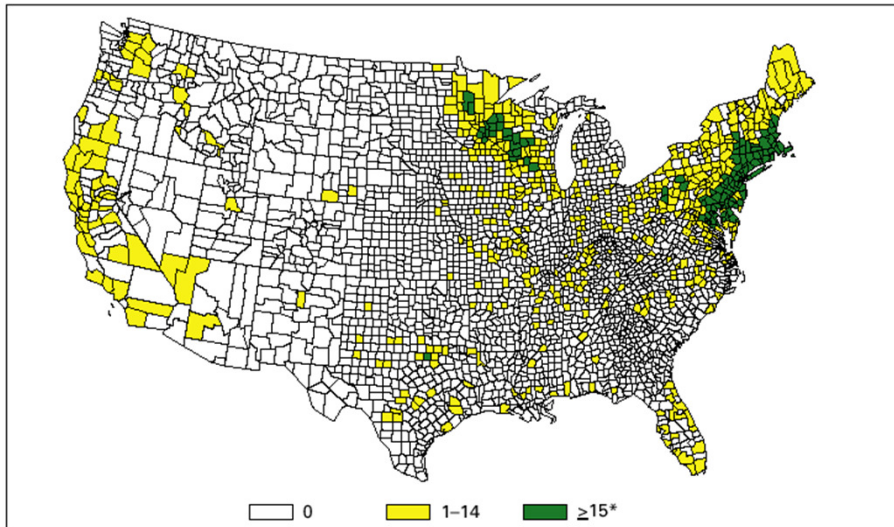
11/2003 Provisional Data - CA DHS STD Control Branch

Box Plots



Maps

Lyme Disease, Reported Cases by County, 2001

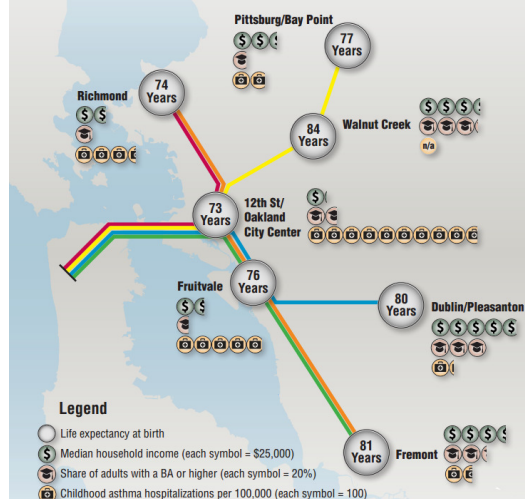


*The total number of cases from these counties represented 90% of all cases reported in 2001.

A total of 17,029 cases of Lyme disease were reported in 2001, 15,998 of these from 10 states with endemic disease (Connecticut, Delaware, Massachusetts, Maryland, Minnesota, New Jersey, New York, Pennsylvania, Rhode Island, Wisconsin). The incidence rate in these states was 23.9/100,000 population. By integrating prevention strategies into community-based programs, CDC and state health departments hope to achieve the *Healthy People 2010* goal of reducing the incidence of Lyme disease to 9.7 cases/100,000 population in states with endemic disease.

Health and Wealth inequities across Bay Area Rapid Transit (BART) stations

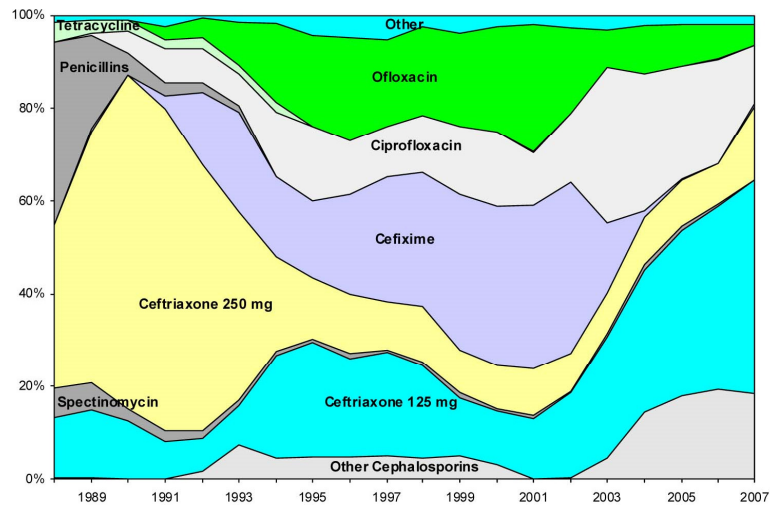
The short distance between a few BART stations can mean an 11-year difference in life expectancy and dramatic differences in physical and economic well-being.



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

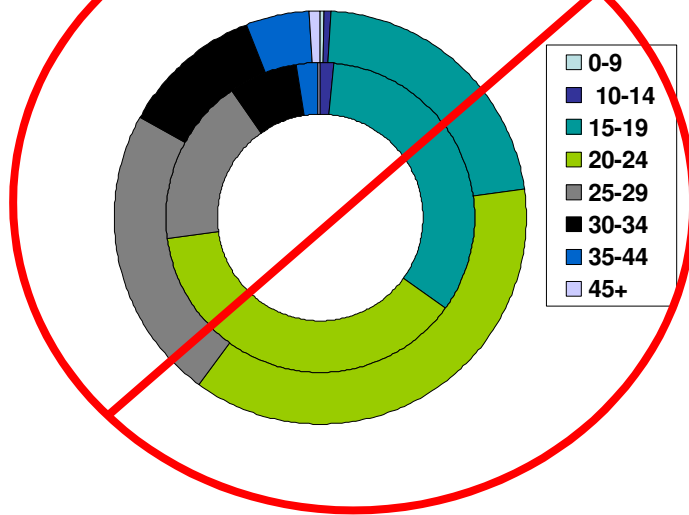
Many Other Types and Hybrids

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

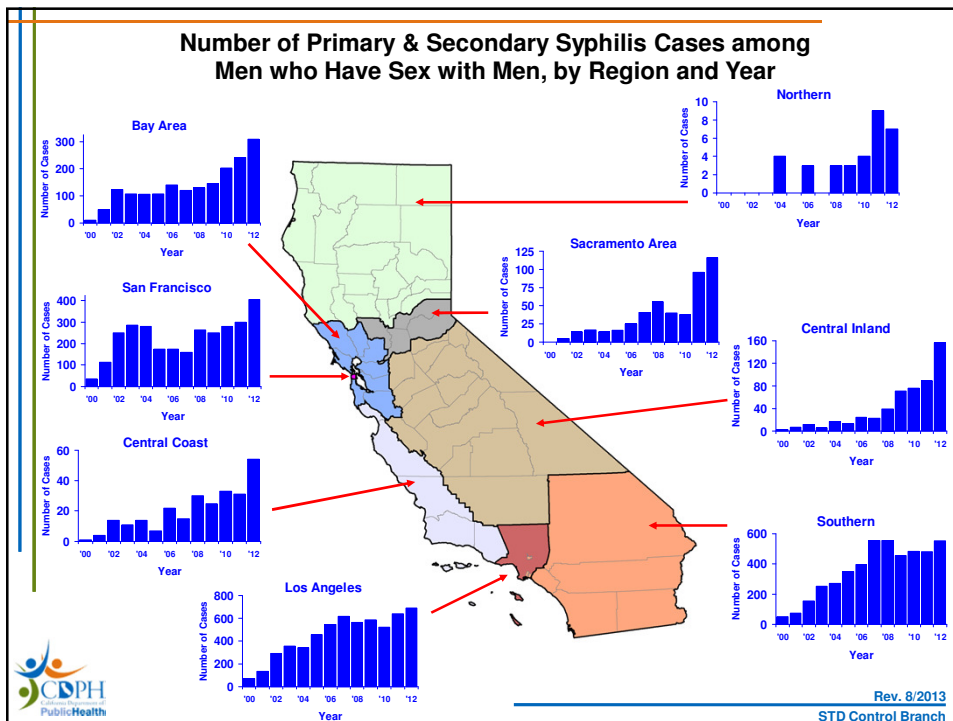


Note: For 2007, "Other" includes no therapy (0.8%), azithromycin 2 g (0.8%), levofloxacin (0.2%), and other less frequently used drugs.

Chlamydia Rates by Age Group California, 2002



Number of Primary & Secondary Syphilis Cases among Men who Have Sex with Men, by Region and Year



“Nut and Bolts”

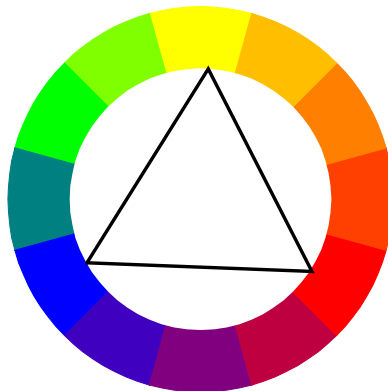
- Scale and Proportion
- Labels and Legends
- Grid Lines
- Color
- Animation/“PowerPoint”
- Font
- 3D
- Production/Reproduction
- Chart Junk
- Software

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.)

How to Select a Color



Color/Pattern Choice

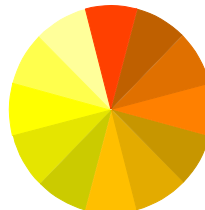
Classification Variable

Quantitative Variable

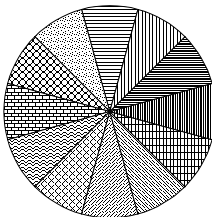
Hue palette



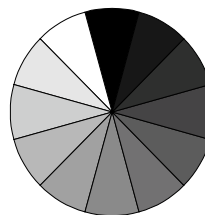
Light-dark palette



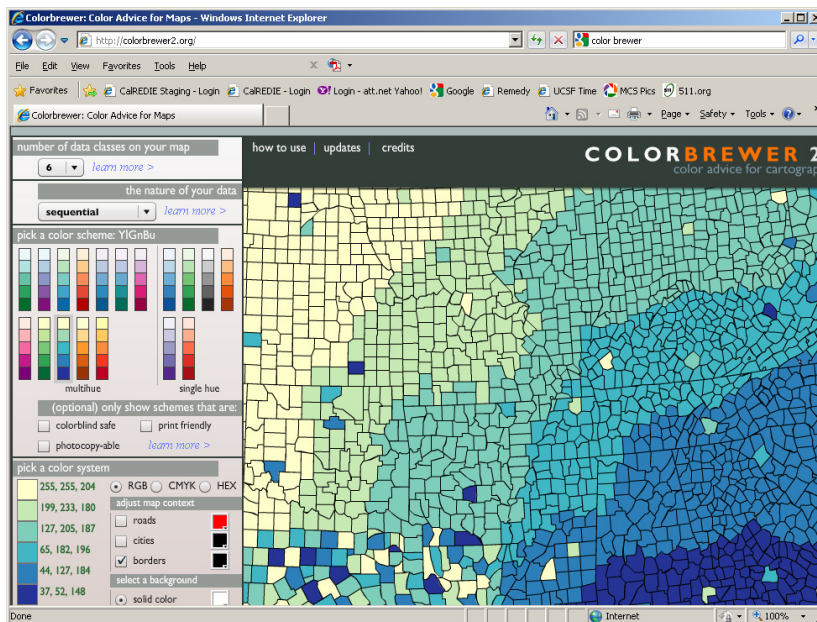
Pattern palette



Gradient of grey



<http://colorbrewer2.org/>



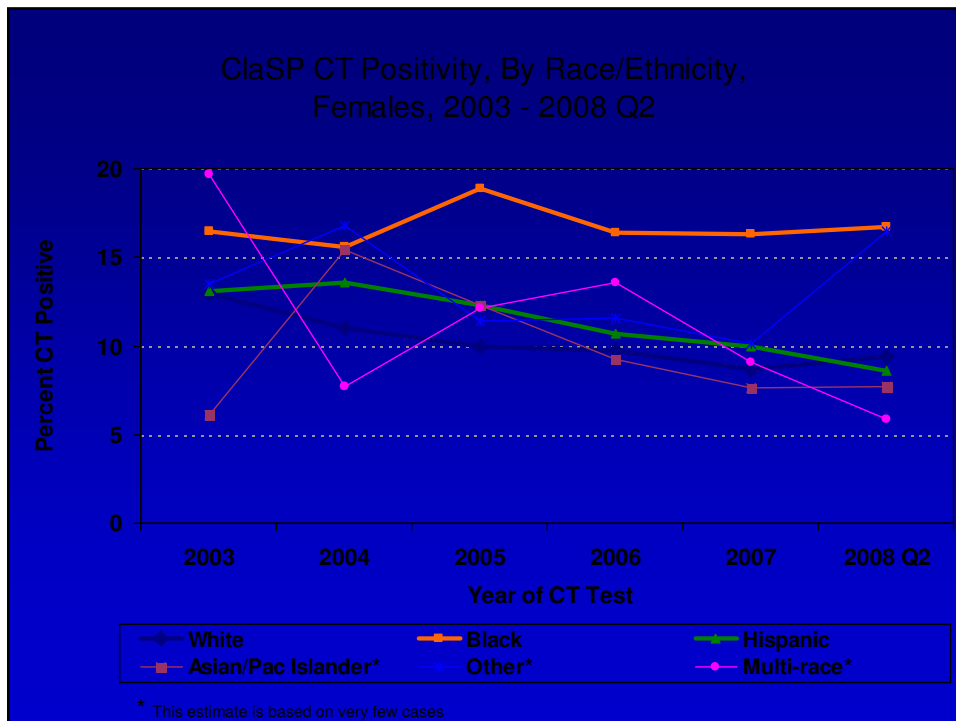
Leading causes of death in Los Angeles County, 2005.¹

Rank	Cause of death	No. of deaths	Premature death rank	Rank	Cause of death	Years of life lost	Death rank
1.	Coronary heart disease	15,154	1.	1.	Coronary heart disease	64,231	1.
2.	Stroke	3,775	8.	2.	Homicide	48,067	10.
3.	Lung cancer	3,036	5.	3.	Motor vehicle crash	32,394	12.
4.	Emphysema/COPD	2,770	13.	4.	Suicide	19,375	17.
5.	Pneumonia/influenza	2,333	19.	5.	Lung cancer	18,490	3.
6.	Diabetes	2,305	7.	6.	Liver disease	17,897	11.
7.	Alzheimer's disease	1,546	50.	7.	Diabetes	16,971	6.
8.	Colorectal cancer	1,409	12.	8.	Stroke	16,442	2.
9.	Breast cancer	1,174	10.	9.	Drug overdose	16,138	21.
10.	Homicide	1,066	2.	10.	Breast cancer	13,610	9.

Background Colors
and oral presentations

Background Colors

And Slide "Design Templates"



Fonts / Fonts

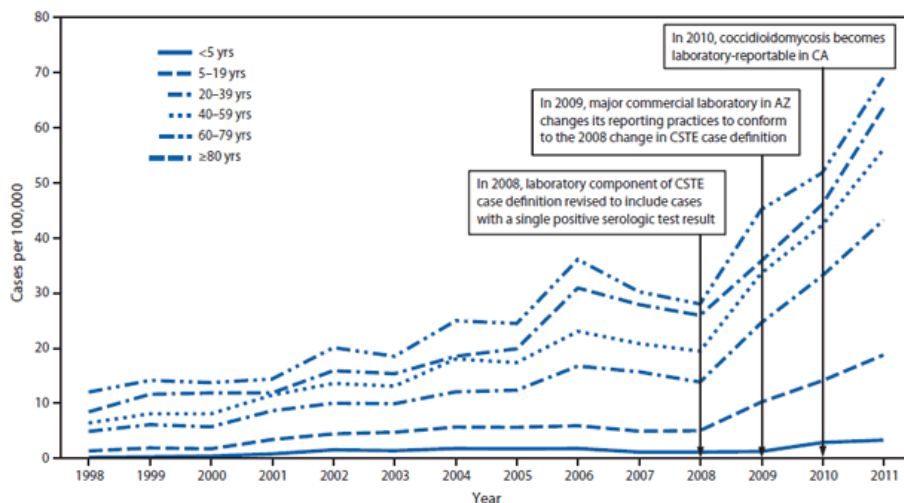
- Use San Serif Fonts, Like Arial
- Not Serif Fonts, Like Times Roman
 - They Are Harder to Read
 - Particularly in Oral Presentations
 - When the Font Is Small
 - See, Isn't This Better
- **ALSO, ALMOST NEVER USE ALL CAPS**
 - IT'S HARD TO READ TOO
- Big Enough to read

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. Coccidioidomycosis incidence per 100,000 population, by age group — Arizona, California, Nevada, New Mexico, and Utah, 1998–2011



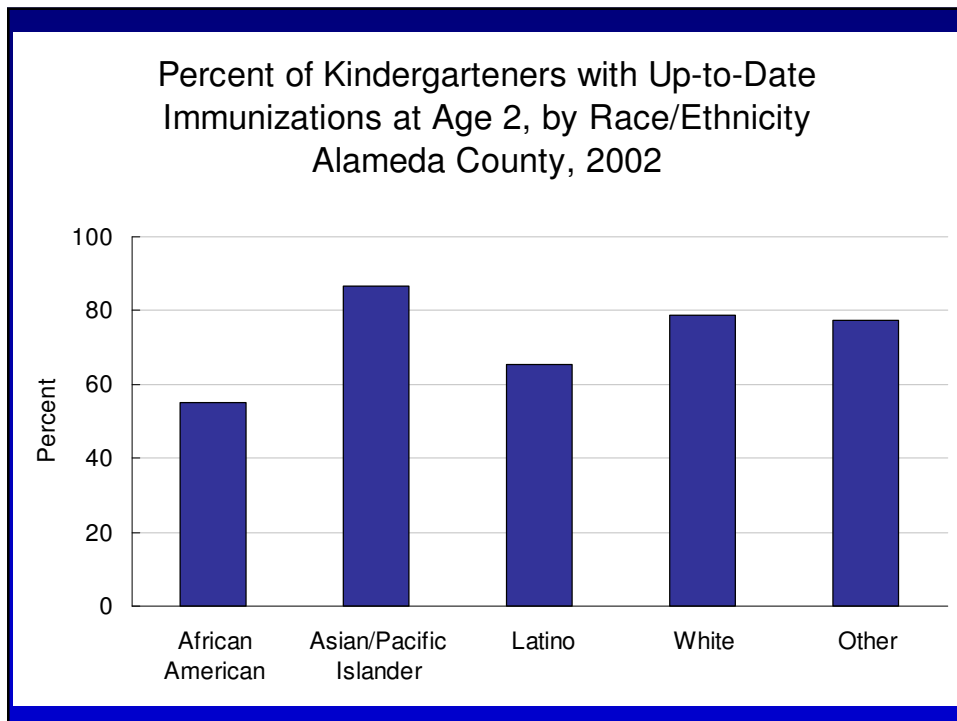
Abbreviations: CSTE = Council of State and Territorial Epidemiologists; AZ = Arizona; CA = California.

Alternate Text: The figure above shows Coccidioidomycosis incidence per 100,000, by age group, in Arizona, California, Nevada, New Mexico, and Utah during 1998-2011. Incidence in coccidioidomycosis-endemic states increased among all age groups during 1998-2011. During this period, incidence typically was highest among the 40-59 year age group in California but was consistently highest among persons aged ≥60 years in Arizona and other coccidioidomycosis-endemic states.

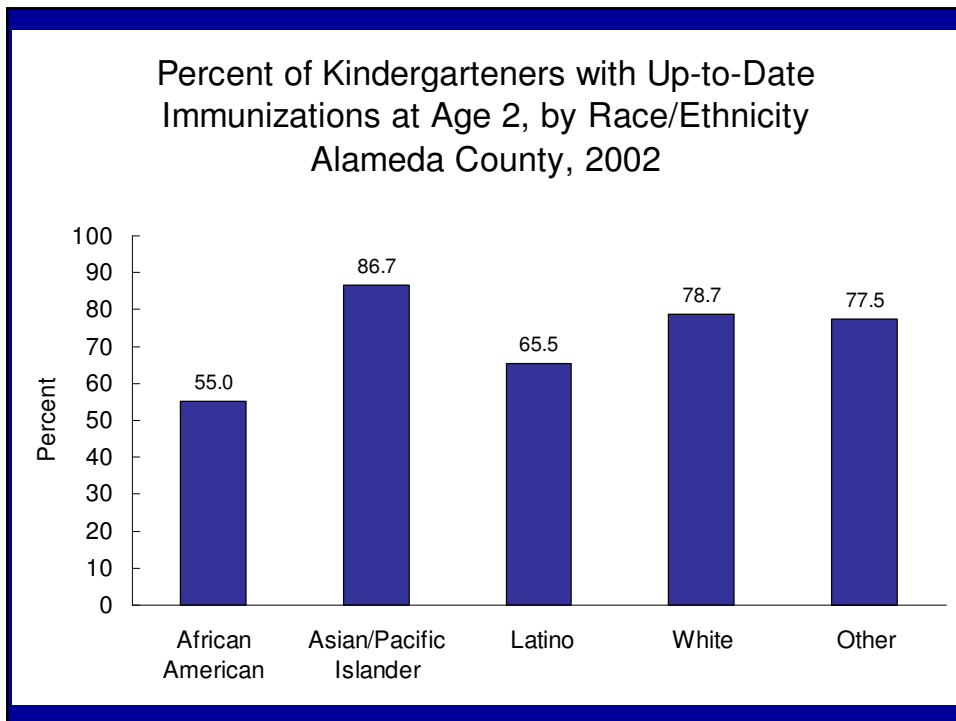
MMWR -- March 29, 2013 / 62(12);217-221

BAD Line types and BAD labeling

Grid Lines



Scale



Percent of Kindergarteners with Up-to-Date Immunizations at Age 2, by Race/Ethnicity
Alameda County, 2002

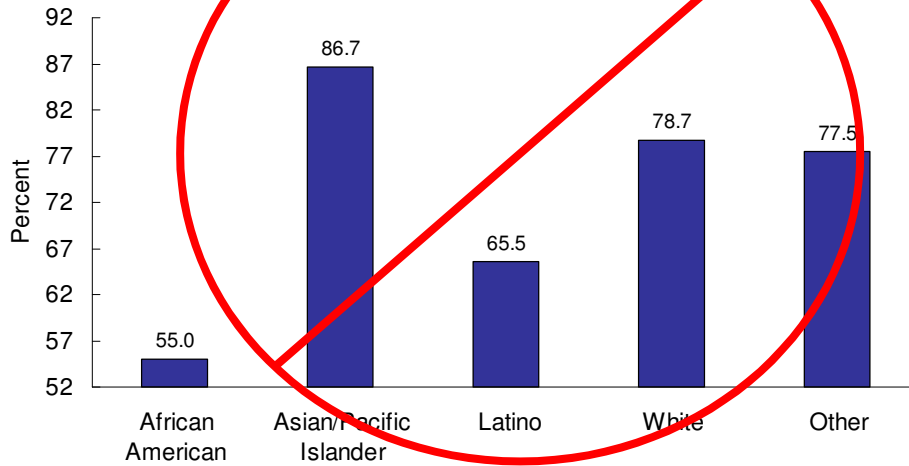
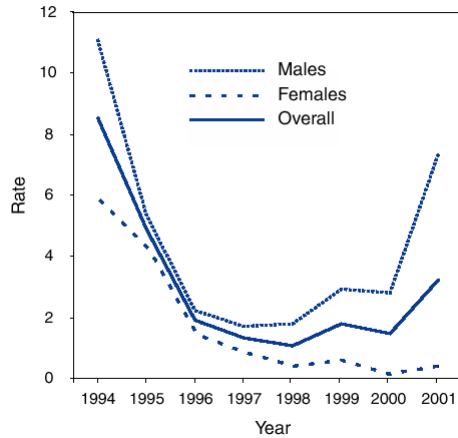


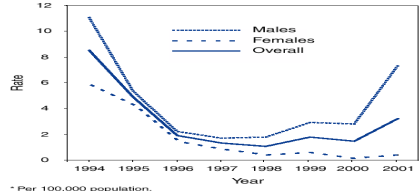
FIGURE. Primary and secondary syphilis rates*, by sex and year — New York City, 1994–2001



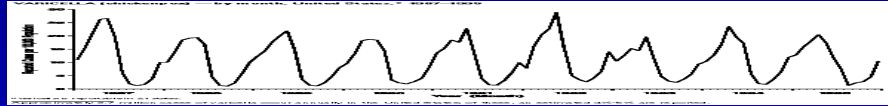
* Per 100,000 population.

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

FIGURE. Primary and secondary syphilis rates*, by sex and year — New York City, 1994–2001

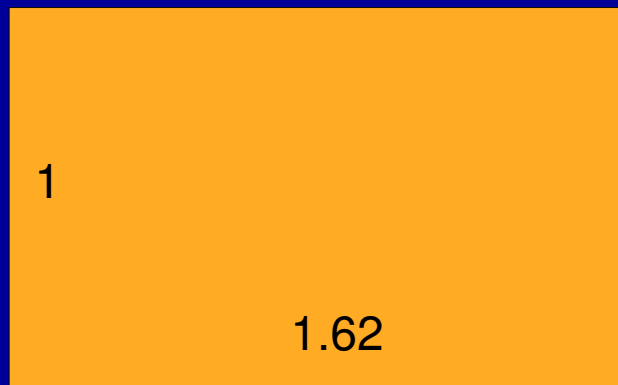


* Per 100,000 population.



“Bank to 45 degrees”

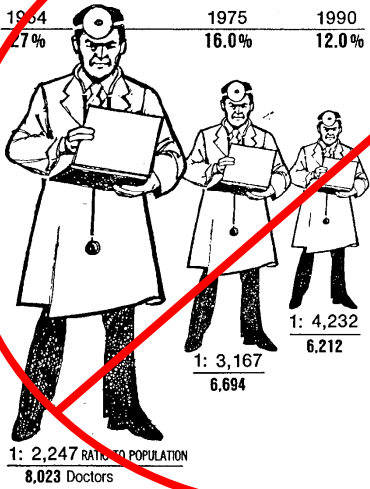
“Golden Rectangle”



THE SHRINKING FAMILY DOCTOR In California

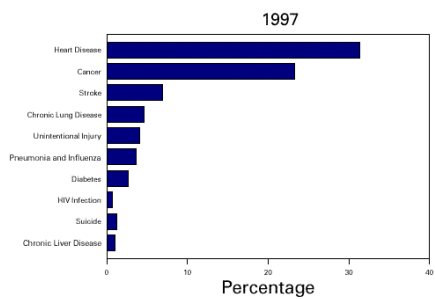
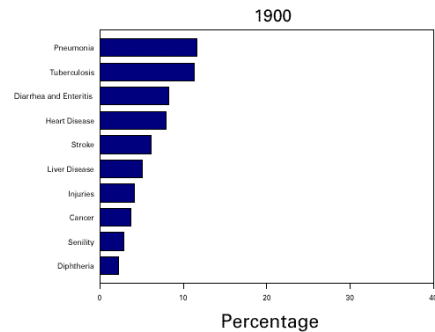
Percentage of Doctors Devoted Solely to Family Practice

1954	1975	1990
27%	16.0%	12.0%



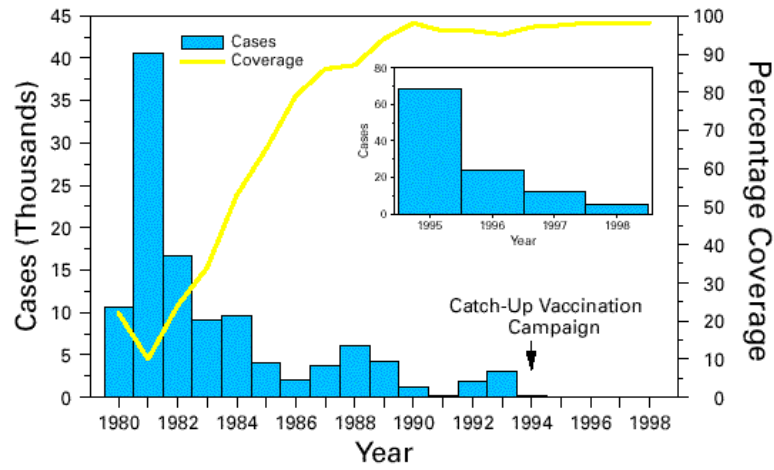
In Tufte, 1983, p. 69.
From LA Times August 5, 1979, p.3.

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



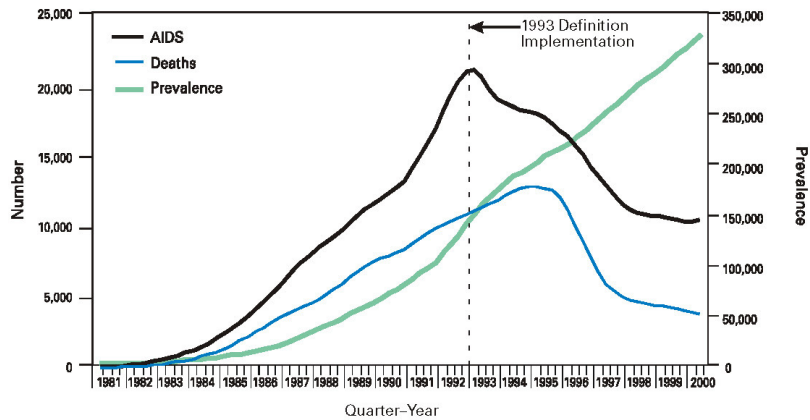
MMWR - Weekly
July 30, 1999 / 48(29):621-629

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



MMWR
Morbidity and Mortality Weekly Report
December 03, 1999 / 48(47);1081-6

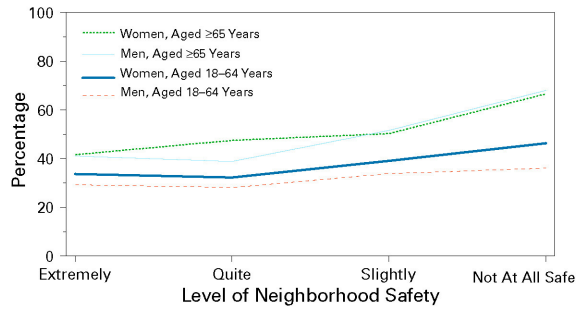
FIGURE 1. Estimated AIDS incidence*, deaths, and prevalence, by quarter-year of diagnosis/death — United States, 1981–2000



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

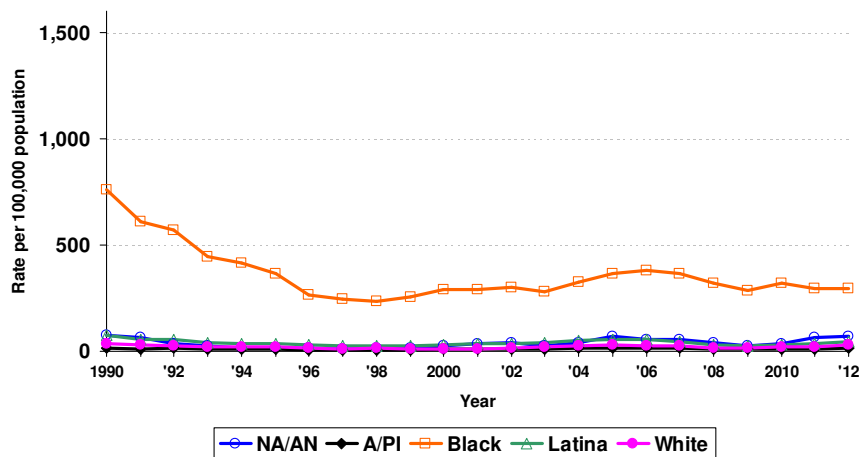
Physical Inactivity — Continued

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



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

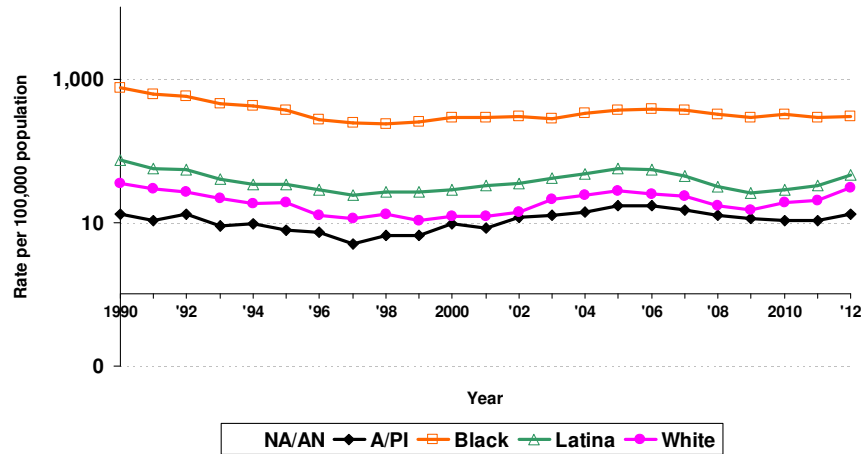
**Gonorrhea, Rates for Females by Race/Ethnicity
California, 1990–2012**



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

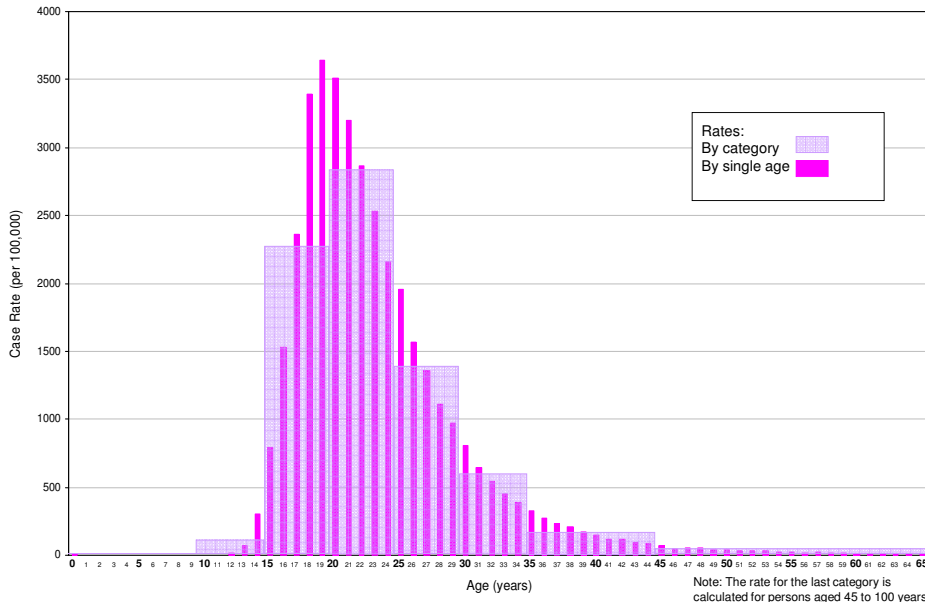
Gonorrhea, Rates for Females by Race/Ethnicity California, 1990–2012



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.

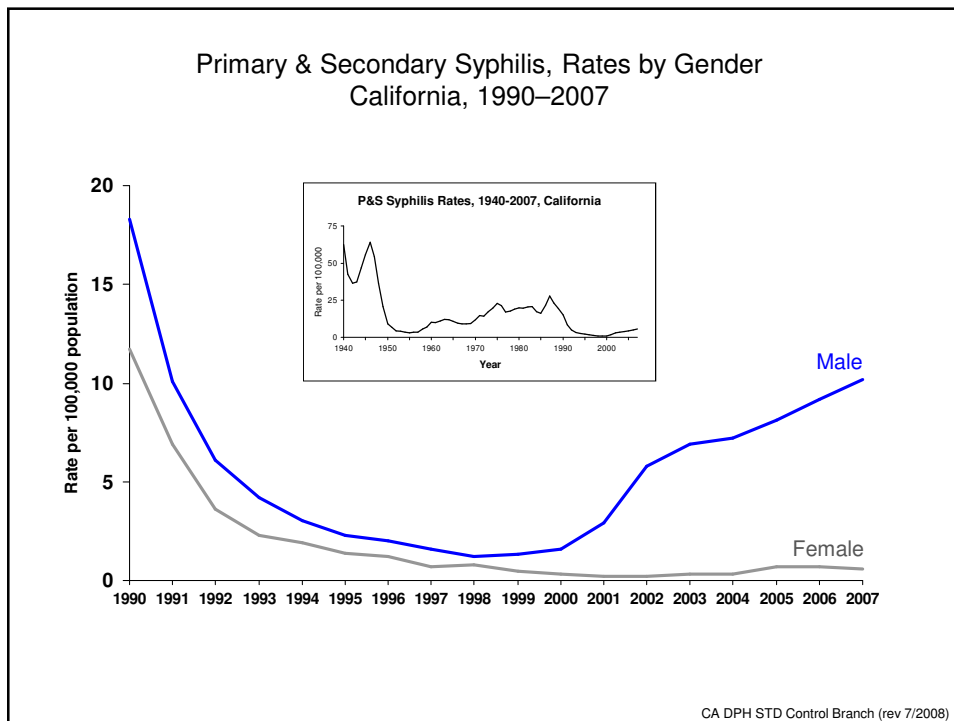
Rev. 8/2013

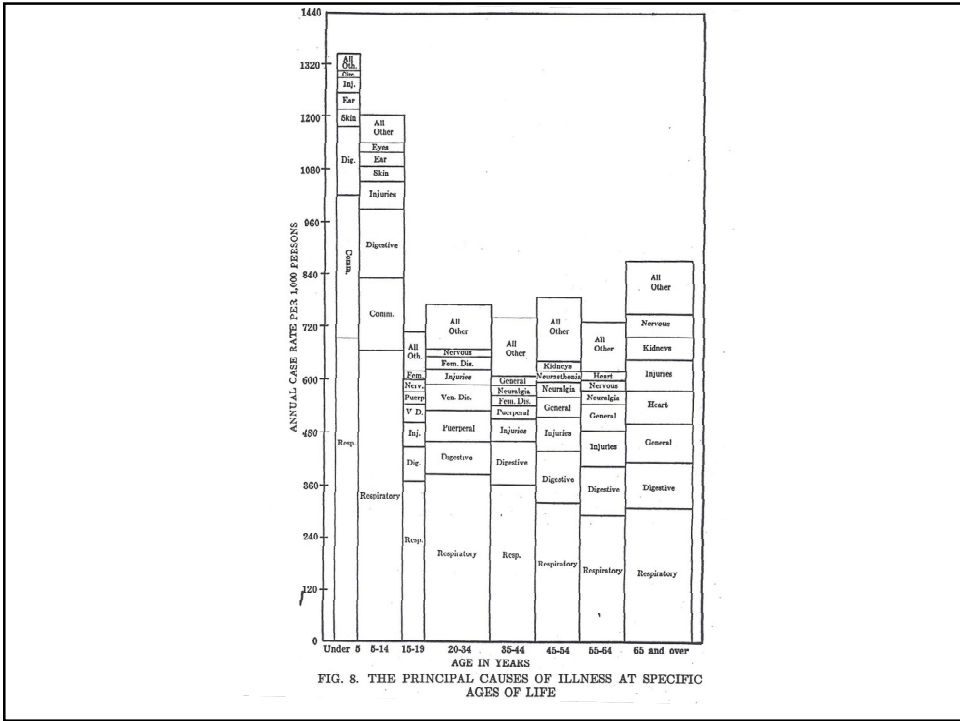
Chlamydia in California Females (2007) Rates by Single Age and Age Category



Note: The rate for the last category is calculated for persons aged 45 to 100 years.

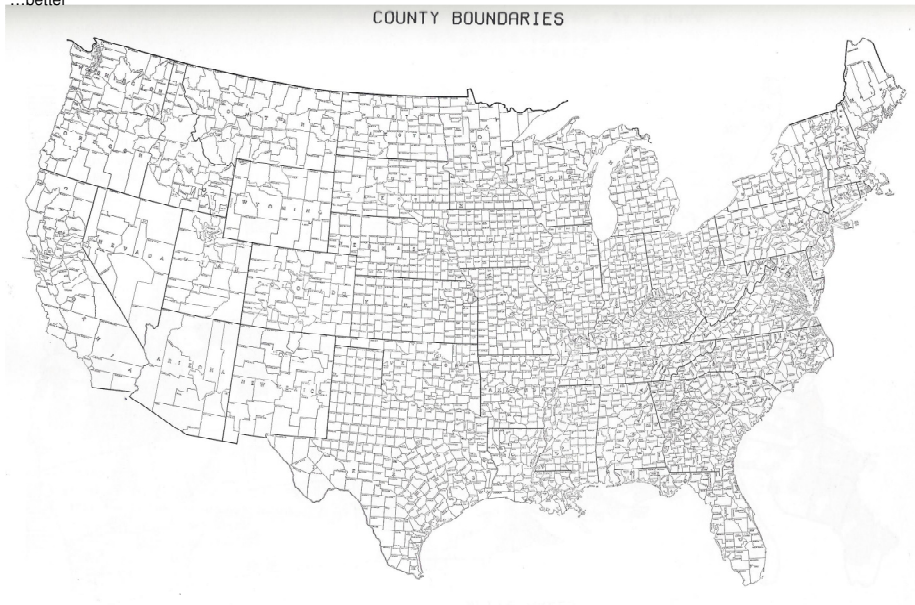
Labels and Legends



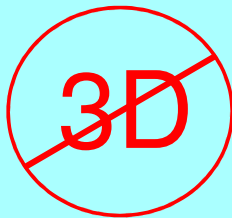


Production and Reproduction

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

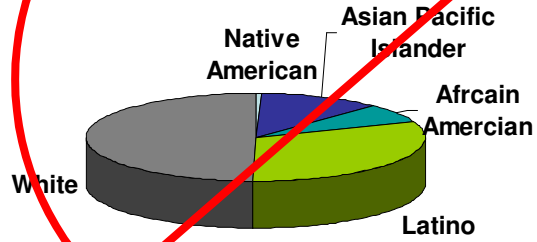


3D Charts

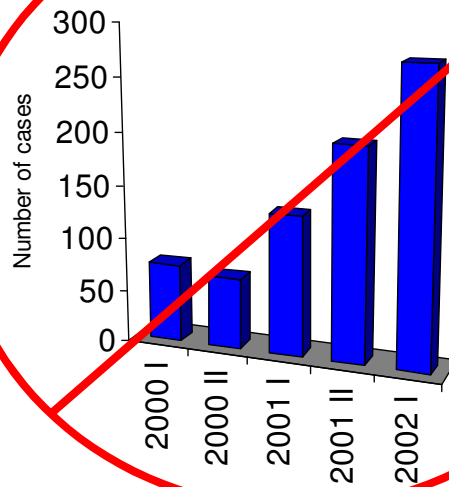


- Unless there are 3 dimensions and the audience can handle it!

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

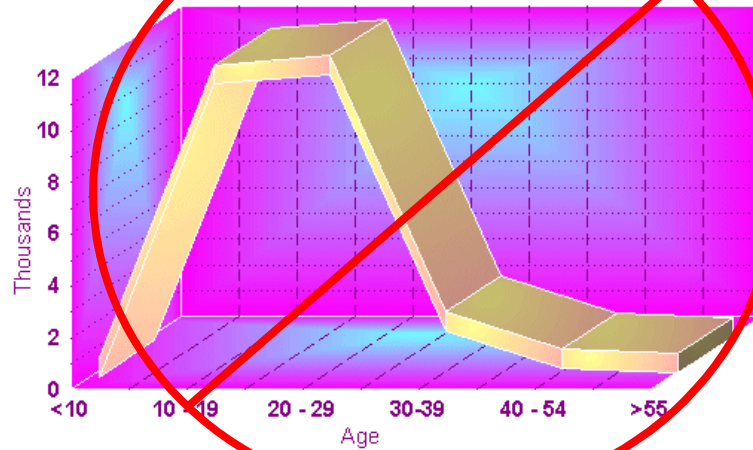


MSM P&S Syphilis Cases by Half-Year Interval
California, 2000-2002



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>

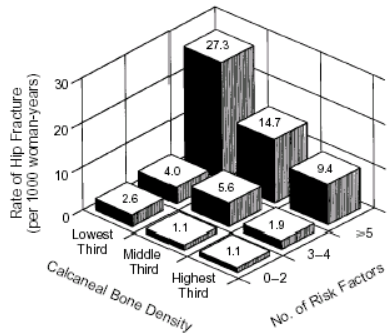
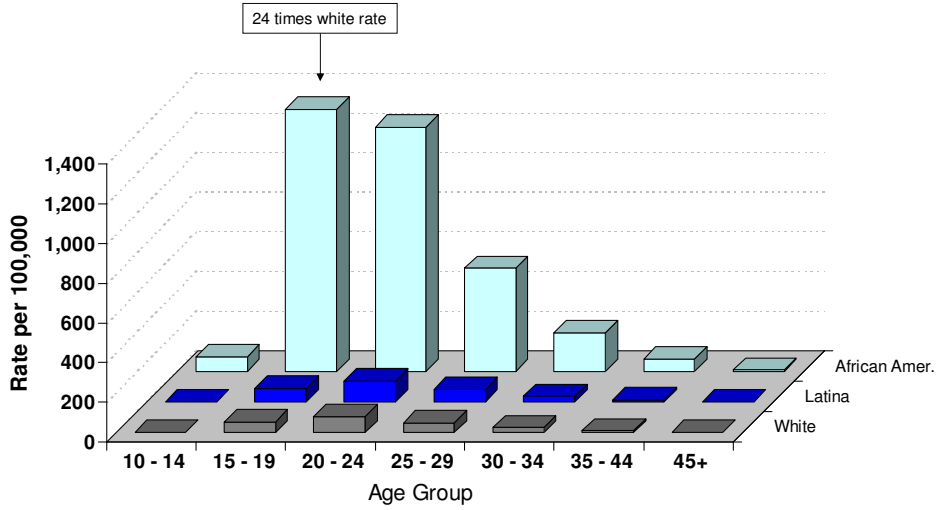


Figure 2. Annual Risk of Hip Fracture According to the Number of Risk Factors and the Age-Specific Calcaneal Bone Density. The risk factors (from Table 2) are as follows: age ≥ 80 ; maternal history of hip fracture; any fracture (except hip fracture) since the age of 50; fair, poor, or very poor health; 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.

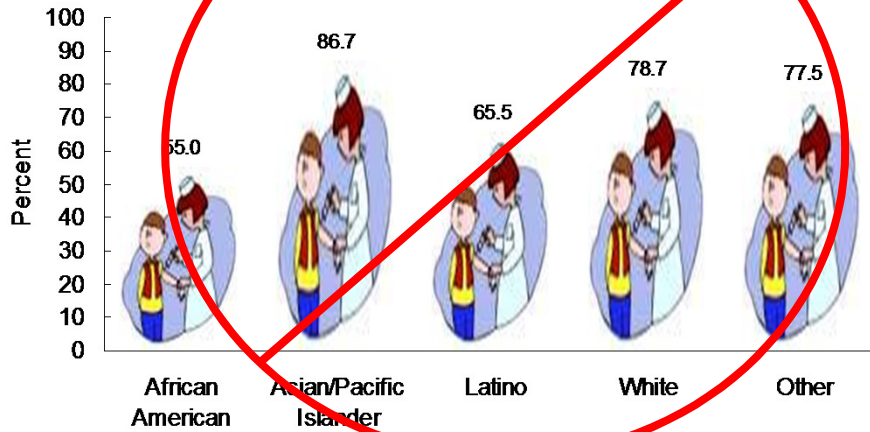
Gonorrhea Rates, Females, by Race/Ethnicity and Age Group California, 2009



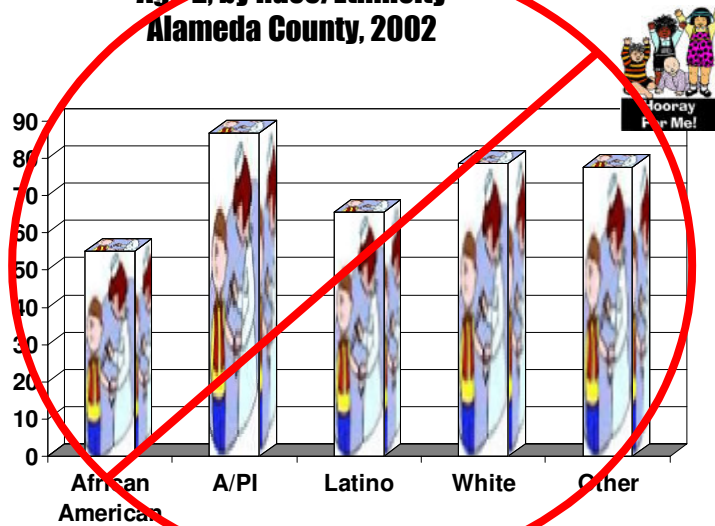
Note: Native American/Alaskan Native and Asian/Pacific Islander rates were excluded due to small case counts.
Source: California Department of Public Health, STD Control Branch

Chart Junk

**Percent of Kindergarteners with Up-to-Date Immunizations at Age 2, by Race/Ethnicity
Alameda County, 2002**



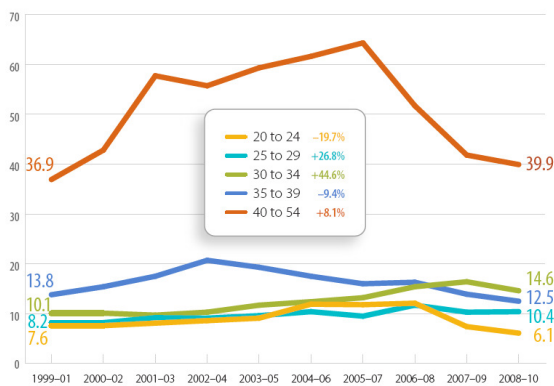
**Percent of Kindergarteners with Up-to-Date Immunizations at Age 2, by Race/Ethnicity
Alameda County, 2002**



“Soften the Edges” “The Human Touch”

Maternal Mortality Rates, by Age of Mother California, 1999 to 2010, Selected Years

DEATHS PER 100,000 LIVE BIRTHS




Notes: Maternal mortality refers to deaths 42 days or less postpartum. A three-year moving average is used. Percentage change represents 11-year period shown.
Sources: California Department of Public Health, Maternal, Child, and Adolescent Health Program, Maternal Mortality Rate, www.cdph.ca.gov, accessed May 13, 2013. Personal communication with California Department of Public Health, January 24, 2013.

©2013 CALIFORNIA HEALTHCARE FOUNDATION

Quality of Care Maternal Health and Childbirth

Mothers age 40 and older have had a significantly higher mortality rate than other age groups in California. The rate for this age group increased substantially from 1999–2001 to 2005–2007, before declining nearly as dramatically. In contrast, the mortality rate for mothers between 25 and 34 had the largest increase over the decade, while the rates for those age 35 to 39 and 20 to 24 have declined slightly.

A Healthy Baby Begins With You!



To Have a Healthy Baby You Should...

Visit your doctor for a complete medical check-up and ask about:

- Taking folic acid to guard against birth defects
- Eating a healthy diet, low in fat and salt
- Being a healthy weight
- Daily exercise
- Any medicine you are taking

To prevent birth defects or a low birth-weight or preterm baby, you should not:




- Drink alcohol
- Use drugs
- Smoke or use tobacco

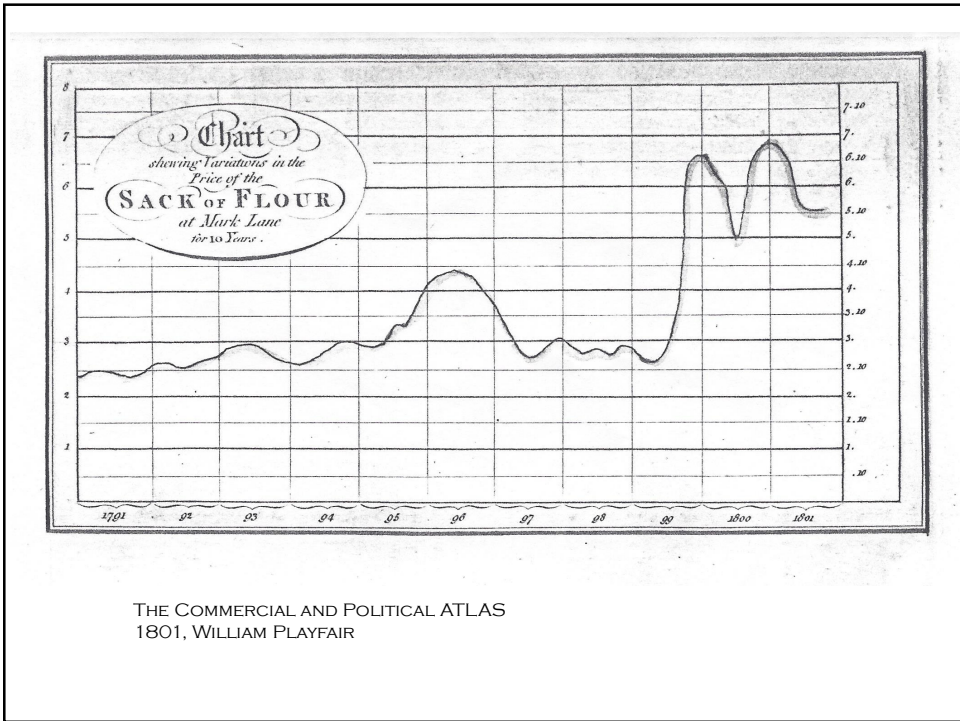
CALIFORNIA'S INFANT DEATH RATE 2005

Group	Rate
African American	12.7
American Indian	7.3
Asian	4.1
Hispanic	3.2
Native American	3.3
White	2.8
California All Race	3.8

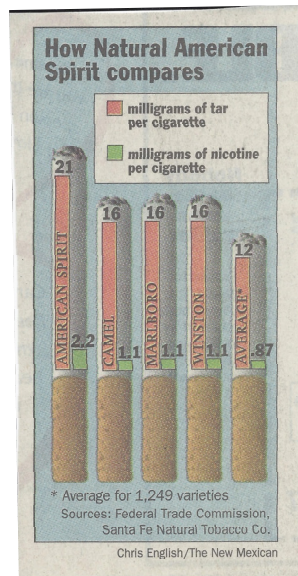
California's infant death rate is highest among African Americans and Pacific Islanders. Proper preconception health is important for the life of every baby. See your doctor if you are pregnant. Risk you are pregnant or planning to become pregnant. This message is brought to you by the Office of Multicultural Health, the California Department of Public Health and California Department of Health Care Services.

MULTICULTURAL HEALTH MONTH APRIL 2009

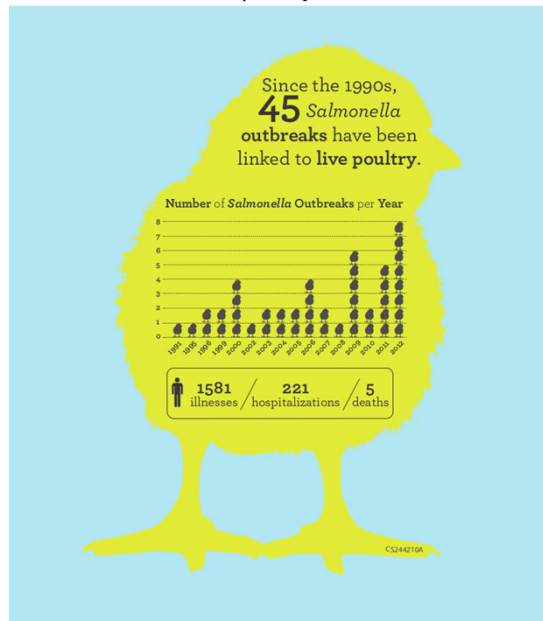






“Infographics”



<http://www.cdc.gov/features/salmonellapoultry/>



SNAP-Ed WORKS

Healthier Lives Through Nutrition Education

THE CHALLENGE

- 1/3 Children in the U.S. are overweight or obese
- 2/3 Adults in the U.S. are overweight or obese
- 1/5 Families with children in the U.S. face hunger/food insecurity
- 1/10 Adults in the U.S. eat the recommended daily amount of both fruits and vegetables

THE SNAP-Ed SOLUTION

- Every \$1 spent on nutrition education saves as much as \$10 in long-term health care costs
- SNAP-Ed works in all 50 states with low-income families and children that receive SNAP (food stamps)
- Teaches SNAP families how to buy and prepare healthy foods
- Helps low-income families stretch tight budgets and buy healthy options
- Introduces school kids to new fruits and vegetables through workshops, classes, and school gardens

THE RESULTS

SNAP-Ed WORKS FOR KIDS

- 3/4 SNAP-Ed participants are children
- 75% of students met the national recommendation for fruit consumption among Georgia SNAP-Ed program participants
- 50% decrease of overweight in elementary school students in a Pennsylvania SNAP-Ed program
- 100% increase of fruit and vegetable consumption in elementary school students in SNAP-Ed's Harvest of the Month program in Ohio, CA

SNAP-Ed WORKS FOR ADULTS & FAMILIES

- Increase in Healthier Foods: 57% increase in the number of participating low-income California adults eating 5+ servings of fruits and vegetables daily
- Increase in Physical Activity: Up to 67% of SNAP-Ed participants were more physically active
- Decrease in Hunger: 40% of SNAP-Ed participants say the program helped them reduce the number of days their families faced food insecurity during the month

Find out how you can protect SNAP-Ed at <http://www.phi.org/resources/?resource=snapedworks>

DESIGNED BY **elefina** designs **PUBLIC HEALTH INSTITUTE**

<http://www.phi.org/resources/?resource=snapedworks>

Query Systems

The screenshot displays the CDC WONDER Multiple Cause of Death, 1999-2010 Request interface. The interface is divided into several sections:

- Request Form:** Includes a search bar, a 'Request Form' section with various filters (Organic Table, Cause of Death, State, Region, Race, Sex, Age, Date), and a 'Crude Rate by State and Race' bar chart.
- Organic Table Section:** Contains a 'Cause of Death' dropdown, 'Organic Table' options, and a 'Crude Rate by State and Race' bar chart.
- Crude Rate by State and Race:** A bar chart showing the crude rate by state and race for the selected query.

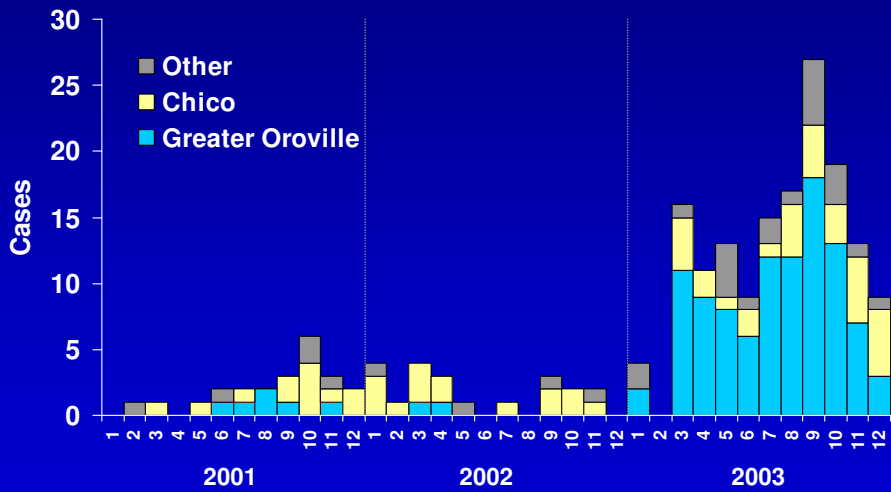
The interface is designed for users to query and analyze multiple cause of death data from 1999 to 2010. It includes a search bar, a 'Request Form' section with various filters (Organic Table, Cause of Death, State, Region, Race, Sex, Age, Date), and a 'Crude Rate by State and Race' bar chart.

Great Graphs

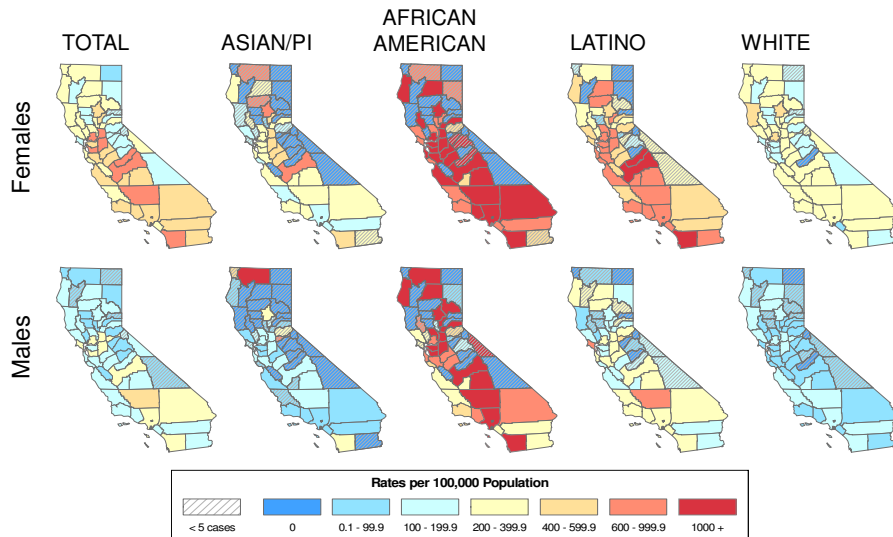
<http://www.bloomberg.com/dataview/2014-04-17/how-americans-die.html>



GC Cases by Month Butte County, 2001-2003

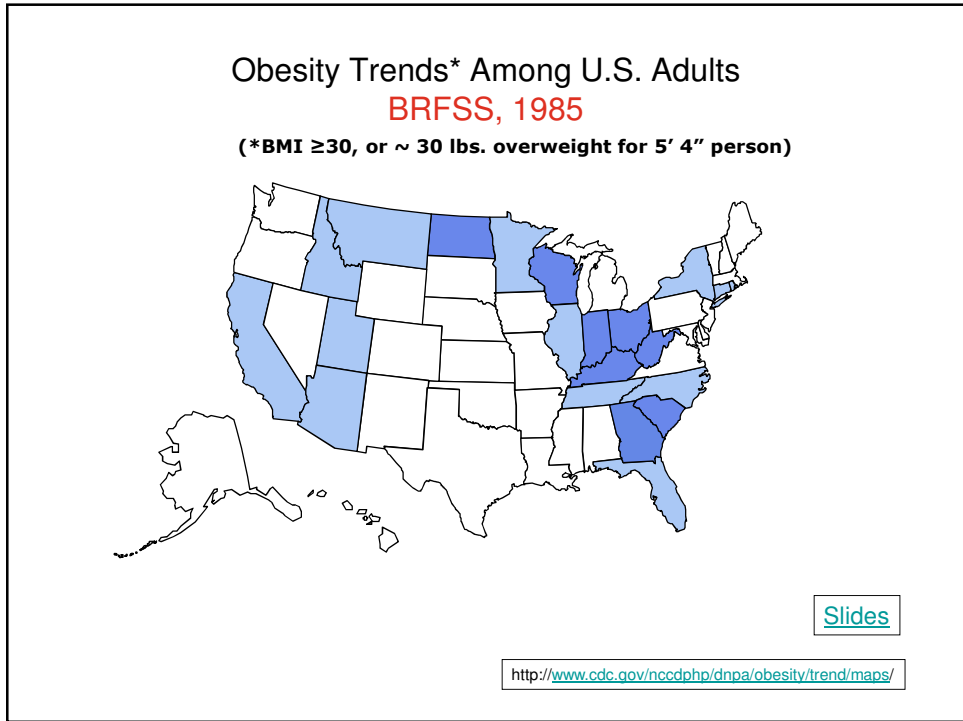
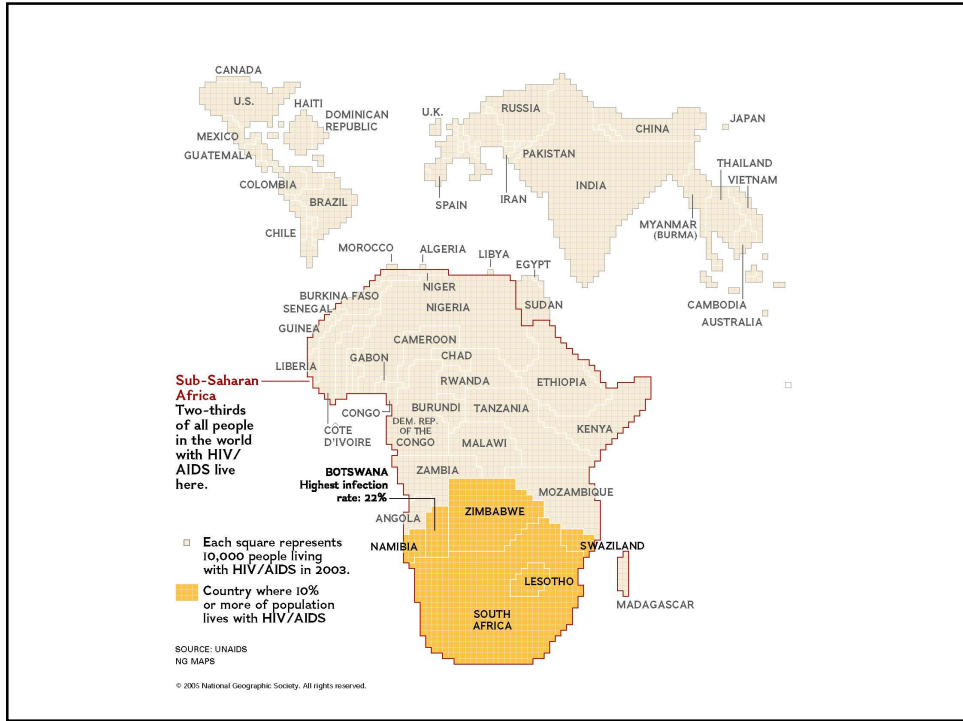


Chlamydia Rates by Gender Health Jurisdiction & Race/Ethnicity California, 2009



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.

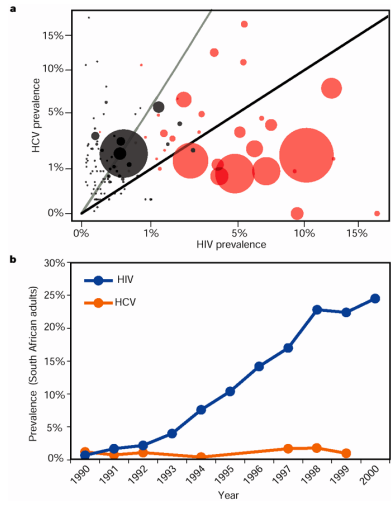
Source: California Department of Public Health, STD Control Branch



Epidemiology: Sexual transmission of HIV in Africa

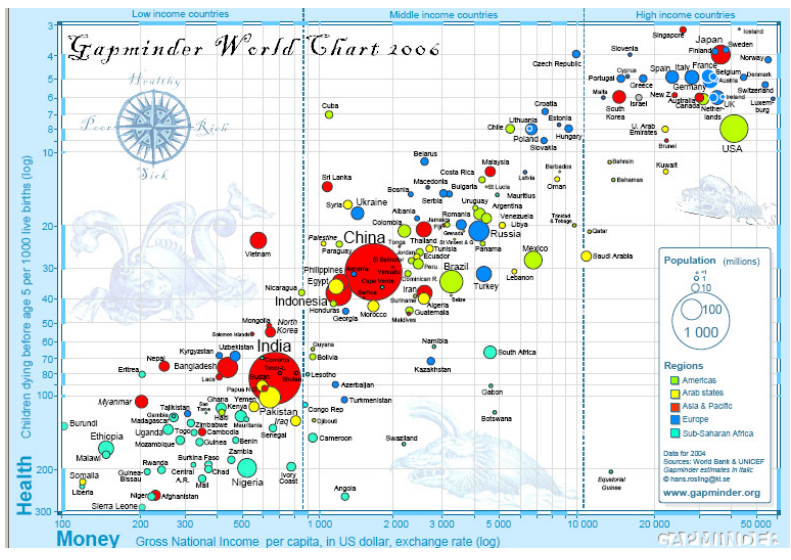
Polly R. Walker, Michael Worobey, Andrew Rambaut, Edward C. Holmes and Oliver G. Pybus

Figure 1 Comparison of epidemic histories of human immunodeficiency virus (HIV) and hepatitis C virus (HCV) in different countries, and in South Africa during the 1990s. **a**, HCV and HIV prevalence in the general population (including children) of every country for which data are available. The arcsine square-root transformation for proportions was used because the original distribution of points was strongly L-shaped. Sub-Saharan African countries are represented by red circles; other countries are represented by black circles; the radius of each circle is proportional to the number of HIV infected individuals in that country. Countries below the black line have a higher prevalence of HIV than HCV; the opposite is true for countries above the black line. The grey line represents an HCV:HIV prevalence ratio of 6, which indicates the relative parenteral transmissibility of the two pathogens **3**. All prevalence figures apart from five HCV values were obtained from World Health Organization sources. **b**, HIV and HCV prevalence in South African adults from 1990 to 2000 (estimates for HIV from ref. **6**; estimates for HCV from sources listed in supplementary information).



NATURE | VOL 422 | 17 APRIL 2003 | www.nature.com

<http://www.gapminder.org/>

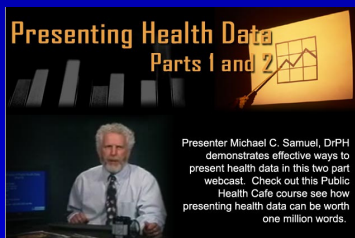
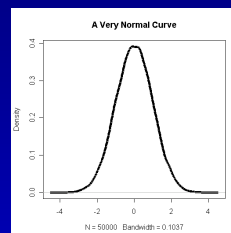


In Conclusion

- 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>