

UCLA CENTER FOR
HEALTH POLICY RESEARCH



Health DATA Program:
Data. Advocacy. Training. Assistance.



Turning Data into *Ta Da!*

California Immunization Coalition Summit
Los Angeles, April 15, 2013

Data: Information for Decision-making

- ▶ Institutions
- ▶ Providers
- ▶ Local Government
- ▶ Community groups and organizations



Why data is important?

- ▶ Basis for planning and priority setting
- ▶ Supports strategizing
- ▶ Strengthens arguments
- ▶ Persuades policymakers and public
- ▶ Motivates action



Data applies to range of information

- ▶ Population
- ▶ Environment
- ▶ Economy
- ▶ Disease and mortality
- ▶ Individual and Institutional behavior
- ▶ Professional practices
- ▶ Policy development



How data is used

- ▶ Assess community trends and context
- ▶ Assess community resources and readiness
- ▶ Establish benchmarks for performance
- ▶ Identify issues
- ▶ Identify gaps in policy
- ▶ Identify audiences and develop messages
- ▶ Support argument, tell your story
- ▶ Refute opponents
- ▶ Polish credibility



Barriers to using data

- ▶ Limited availability of data
- ▶ Limited exposure to data
- ▶ Inability to evaluate the merits of data
- ▶ Inability to generalize existing data
- ▶ Lack of comfort and confidence with data



GOVERNANCE IN ACTION



May 24, 2007

www.ComicStripGenerator.com

-
- ▶ Understanding Data
 - ▶ Determining the Data You Need
 - ▶ Finding Data
 - ▶ Presenting Data
 - ▶ Completing Your Story

Understanding Data

- ▶ Data is factual information used for decision-making. Data comes in many forms.
- ▶ Data can be numerical like statistics or descriptive such as an individual's observations.
- ▶ The most persuasive use of data combines both quantitative data and qualitative data.

Data Types

▶ Quantitative data ←

- ▶ Numerical- # of healthcare facilities in a city

▶ Qualitative data

- ▶ Textual/words, often descriptions- An individual's account of her experience at a healthcare facility

▶ Primary data

- ▶ Collected and analyzed directly by individuals, communities and researchers

▶ Secondary data ←

- ▶ Collected and provided by a third party such as a non-profit, university or government agency (i.e cdc.gov).

Criteria for Evaluating Data

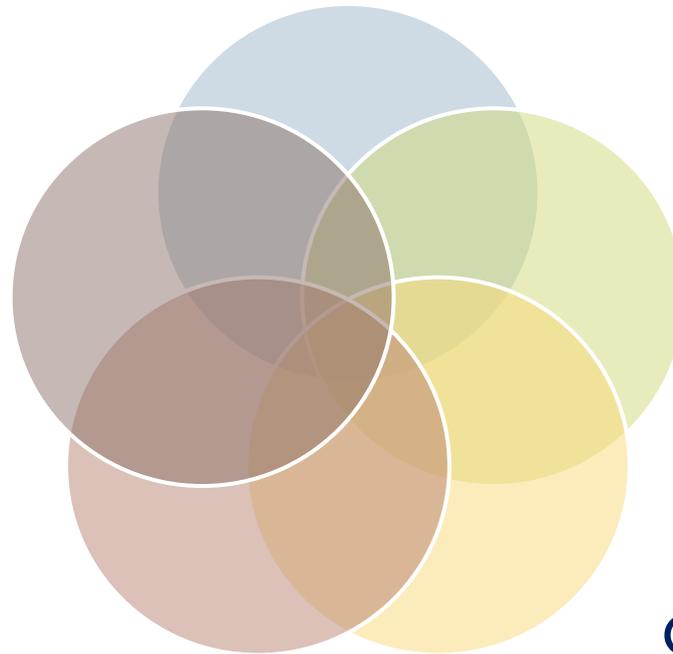
Credibility

Timeliness

Specificity

Reliability

Generalizability



Credibility – who produced the data?

- ▶ **Credibility refers to the source of the data. Can you trust the research entity that produced the data?**
 - ▶ Who paid for, sponsored, or funded the study.
 - ▶ How much of a stake does the data source have in a specific finding? What about their mission or constituents would affect their interpretation of the data?
 - ▶ What is the organization's public image or reputation for their research? Government and academic institutions are considered credible because research is conducted for the public benefit.

Specificity – Is the data precise for a particular population?

- ▶ Specificity refers to data that is limited to particular conditions or factors impacting a specific population.
 - ▶ Research often seeks to answer specific questions.
 - ▶ How close is the relationship between the data provided and the data you need?
 - ▶ For instance, you are looking for information on childcare for Vietnamese children ages 0-5.
 - ▶ Statistics are available on child care for Asians and for children ages 0-5, but nothing specific on Vietnamese children ages 0-5.
 - ▶ If the data is specific to one population you may be able to “generalize” it, that is apply it to another population.

Generalizability – Can you apply data from one population to another population?

- ▶ Generalizability refers to data on a specific population that can be used widely for other populations.
- ▶ If the study was done in Oakland, does it also apply to people in Los Angeles?
- ▶ If it was a national study, can we use the same numbers for Orange County?
- ▶ If the study looked at Latinos, is there anything particular to say about Central Americans as a distinct group?
- ▶ How much can you “generalize” data to your constituents or service population?
 - ▶ It is important to know the “who, what, why, when and where” of the data because it will determine how closely it matches your own need for data.

Reliability- How was the data collected?

- ▶ Reliability refers to the accuracy of the data. Can the data be trusted to be accurate?
- ▶ Has the research that produced the data been reproduced by other researchers?
 - ▶ Did other research studies get the same or similar results? Research studies that result in similar data are said to “validate” or confirm the results of the other similar studies.
- ▶ How was the data collected?
 - ▶ Was it consistent with the mission/goals of the researchers?
 - ▶ Did the researchers adhere to ethical research methods?
- ▶ Is there some kind of bias in who will reply?
 - ▶ Ex: Did they conduct their survey in different languages if they need information about immigrants?
- ▶ You’ll want to scrutinize the methods used to produce the data and collect other studies that validate the data you want to use.

Timeliness – when was the data collected?

- ▶ Timeliness refers to when the research was conducted relative to the changes occurring among a population or conditions impacting the population.
- ▶ When was the study done? And how fast are changes occurring? 1 year ago? 3 years ago? > 10 years ago?
 - ▶ Some data may be accurate over a longer period of time than others.
- ▶ Most comprehensive surveys will be a few years old by the time they are published.
- ▶ Even if the research is “old,” it may be the best possible source if more recent data is not available.
 - ▶ Admit the limitations of this kind of data and try to supplement it with other closely related research.

Challenges of Race/Ethnicity Classifications in Health Data

- ▶ Misclassification of race/ethnicity
- ▶ Great diversity in health needs and actions
- ▶ Impact of length of time in U.S.
- ▶ Lack of disaggregated information



Your Credibility as a Factor

- ▶ You are bringing your own credibility to the data, so it is important to set standards for yourself with regard to the data you plan to use.
- ▶ Researchers judge their data and the studies of others using the above criteria.
- ▶ Use these criteria as well to assess the quality of data that you collected, assess the quality of other sources of data.
- ▶ No data is perfect. Use your own judgment regarding the use of data you think is defensible.
- ▶ Know the assumptions behind the data you use.

Determining the Data You Need

- ▶ Before beginning a search for data, it is important for you to outline and define the questions you need the data to answer.

Defining your question and audience

- ▶ To determine the data you need, answer these questions:
 - 1. What do I need the data to do?**
 - 2. What message does the data have to deliver?**
 - 3. Who are the audiences for the data, and what data does each audience need?**



1. What do I need the data to do?

- ▶ Define the problem
- ▶ Show that your solution alleviates or solves the problem
- ▶ Show the negative consequences of not using your solution
- ▶ Measure program outcomes
- ▶ Monitor trends in a population or community

- ▶ To determine what you need the data to do, get a picture of your social, economic, and political environment:
 1. What are the key public health issues affecting children in your community?
 2. Define the problem central to the issue: i.e. access to health care, a particular chronic disease, lack of places for physical activity, etc. Think about how these affect your community.
 3. Who are the stakeholders in these issues?
 4. Who has the power to advance the policies or programs you have identified?
 5. Barriers/Facilitators

2. What message will the data deliver?

- ▶ What will mobilize your stakeholders and motivate the decision makers into action?
- ▶ Different things will convince different people.
 - ▶ Statistics vs. personal story of overcoming hardships
- ▶ In either case, the statistics and the stories are most effective when they appeal to someone's values.

Your Data's Message

- ▶ Some common values that data can address are cost, quality, access, equity, and rights:
 - ▶ *Cost*—what is the cost of the problem to taxpayers, community, business, individuals, and others.
 - ▶ *Quality*—how is quality of life, environment, services, and programs impacted?
 - ▶ *Access*—who has access to services, programs, insurance, jobs, education, clean air, etc.? Who doesn't?
 - ▶ *Equity*—is there an equitable distribution of resources among segments of a community?
 - ▶ *Rights*—what are the rights of members of a community? What laws, regulations, or constitutional protections confer rights and on whom are the rights conferred?

3. What does your audience need?

- ▶ Elected officials, juries, media, general public need data to understand the scope of the problem (the forest)
- ▶ Committee staffs, judges, special interest groups with legislative analysts need more specific information on who and what is impacted (individual trees)
- ▶ Agencies, courts, academics need details and statistics (roots)

Levels of Information

- ▶ **The Forest -- *Big Picture*:** Politicians, the general public and the media are audiences who tend to need information that is descriptive and easy to understand, often from an overall perspective or big picture point of view.

For all groups, influenza vaccination coverage for the 2010–11 season remained well below the Healthy People 2020 targets.

CDC Seasonal Flu influenza vaccination coverage estimates for the 2010–11

Levels of Information

▶ **Individual Trees -- Some Details:**

- ▶ Committee staff, judges, and special interest groups with legislative analysts tend to need more detail than the big picture.
- ▶ These individuals want to know **what kind of trees are in the forest or how many trees per square acre.**
- ▶ They may want to know if they clear this part of the forest, what does that do to the ecosystem, etc. **This information will have more layers to it;** often the audience understands the general ideas, but does not understand the details.

Healthy People 2020 recommends 80% for persons 6 months and older; and 90% for adults ≥ 65 years and adults 18–64 years with high-risk conditions. Over the past two seasons, there was no improvement in flu vaccine levels for adults ≥ 18 years.

CDC Seasonal Flu influenza vaccination coverage estimates for the 2010–11

Levels of Information

- ▶ **Roots - *Specific Details:***
- ▶ Government agencies, court officials, and academic institutions often need data to be more academically focused or statistically driven to understand and critique.
- ▶ Requires a high degree of accuracy because funding or planning decisions will be made based on those numbers.

Among adults ≥ 18 years, non-Hispanic whites had the highest coverage (43.2%) compared to Hispanics (32.3%), non-Hispanic blacks (34.2%), non-Hispanic others (37.4%), and Asians (38.2%). At Los Angeles County Public Health Outreach clinics, [flu vaccine] recipients 19 to 49 years of age, as a percentage of total doses, remained level at 33%. Blacks and American Indian / Alaskan Native (AI/AN) received the smallest proportions of vaccine administered during clinics. SPA 6: Averaged 16% of flu vaccines for 2012.

CDC Seasonal Flu influenza vaccination coverage estimates for the 2010–11 and Los Angeles County Dept. of Public Health 2011-2012 Influenza Campaign Report

Cutting the Data Question

- ▶ **When looking for data, consider the following questions before you begin your search:**
- ▶ What is the **Problem/Issue** you are trying to resolve?
- ▶ What is the **Cause** of the Problem?
- ▶ What are the **Effects** of the Problem?
- ▶ What are the characteristics of the **Population**?
- ▶ Does **Geography** have an effect on the problem?

Finding Data

- ▶ As you collect, summarize, and interpret information relevant to specific public health problems, you will also encounter some of the real dilemmas that confront those seeking data; a lack of necessary resources, a lack of quality data and a lack of access to the data you need.

Data Search Challenges

- ▶ There are three themes that may arise during your data search.
- ▶ **Limited Resources:** Organizations may lack the internal capacity to search for appropriate data or fully utilize the data that they find. Often the organizations may not have the time, knowledge base, and/or people power to access the data.
- ▶ Plan to spend time looking for data when you do a campaign. Be realistic about how much time this is going to take you and the manpower and expertise that you will need.
- ▶ **Access:** The limited quantity or non-existence of certain data can be very problematic.
 - ▶ May take some time to find the data you need or it may require that you create a local estimate based on national or state data you find.
- ▶ **Quality:** It can be difficult to determine the reliability of information, especially if time constraints are involved in the process.
 - ▶ Refer to the five criteria for evaluating data.

What can you do if the data you need are not available?

- ▶ Use proxy measures
- ▶ Extrapolate state data to make local area estimates
- ▶ Paint a picture with data
- ▶ Ask a researcher

Proxy Measures

- ▶ “Proxy” measures are information that can substitute for the data you need because it is closely related to your issue.
- ▶ For example, Medicaid is a proxy measure for household income.
- ▶ Example: “The extent to which physicians participate in *Medicaid* has been examined frequently as a *proxy measure* of access. .”
- ▶ Not precise, but offers some estimates for when data are not available

Extrapolating Data

Produce local estimates by applying national, state, or county rates to local population

Example: Number of Latinos with diabetes in your community

Nat'l Diabetes

Rates* x Pop. in your community (by the different ages)

$$.02 \times 30,000 \text{ (ages 18-44)} = 600$$

$$.143 \times 11,000 \text{ (ages 45-65)} = 1,573$$

$$.203 \times 2,000 \text{ (ages 65 and over)} = 410$$

Next add up the various populations with diabetes:

$$600 + 1,573 + 410 = 2,583$$

There are an *estimated* 2,500 Latinos with diabetes in your community.

** Centers for Disease Control and Prevention*

Presenting Data

- ▶ Present data visually
 - ▶ Tables
 - ▶ Charts
 - ▶ Graphs
- ▶ Strengthen arguments with data
- ▶ Know data of opponents
- ▶ Counter opponents with data

Painting A Picture

- ▶ Use a combination of different data sources available to “paint a picture” of the problem or solution. Decide how to frame your argument and what information is needed to support any conclusions.

Painting a Picture

- ▶ **Combine Qualitative & Quantitative Data can Strengthen your Case:**
- ▶ *Quantitative data* are usually measured and expressed in the form of numbers or percentages. This data answers the who, what, when and where.
- ▶ *Qualitative data* are usually measured and expressed in the form of words, concepts, themes, or categories rather than numbers. Qualitative data is often used to gain a more in-depth understanding of a particular incident or phenomenon - answering how or why something is occurring.

Ask a Researcher

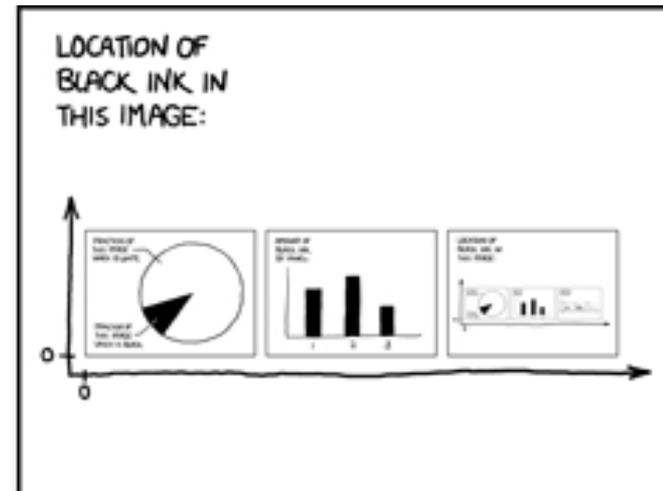
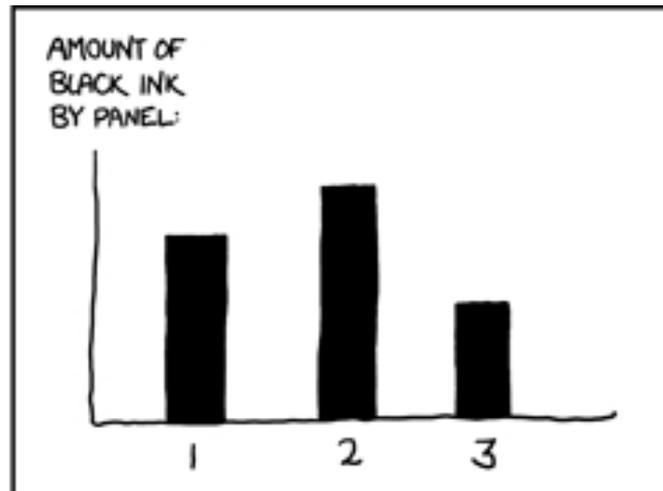
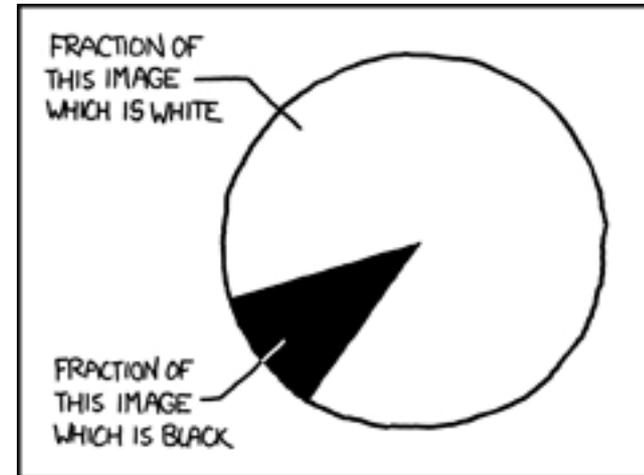
- ▶ Form partnerships with academic institutions, public health departments, individual researchers
- ▶ Form a Community Research Collaborative with other organizations doing similar work to share data and research resources



**Logic: another thing that
penguins aren't very good at.**

Presenting Data

- ▶ Identify what types of graphs, data tables, and narrative analyses are helpful for presentation of different kinds of information to various audiences, ranging from community stakeholders to policy makers.



Tables

Using tables or graphs provides a useful visual presentation of the data.

- ▶ Suitable for providing simple or more complicated numeric or percentage information.
- ▶ Best used for side-by-side comparison of data for various variables or groups.
- ▶ Important to use when you want to show the exact numeric or percentage values.
- ▶ Here is an example of a table showing doctor visits estimates for California residents from the California Health Interview Survey for 2009.

Number of doctor visits in past year

Subset(s): 18 - 100

Entire State of California

	Number of doctor visits in past year		
	Est. N	%	95% C.I.
0 visit	5,420,000	19.7	(18.6 - 20.7)
1 visit	5,359,000	19.5	(18.5 - 20.4)
2 visits	4,346,000	15.8	(15.0 - 16.6)
3 visits	2,968,000	10.8	(10.1 - 11.4)
4 visits	2,469,000	9.0	(8.3 - 9.6)
5-8 visits	3,751,000	13.6	(12.9 - 14.4)
9-12 visits	1,710,000	6.2	(5.7 - 6.7)
13-24 visits	964,000	3.5	(3.1 - 3.9)
25+ visits	559,000	2.0	(1.8 - 2.3)
TOTAL	27,546,000	100.0	n/a

Source: 2009 California Health Interview Survey

Percentage of persons aged ≥ 18 years who reported receiving influenza vaccine by reporting areas, age group and race/ethnicity - Behavioral Risk Factor Surveillance System (BRFSS), 50 states and the District of Columbia, 2001 and 2002 combined

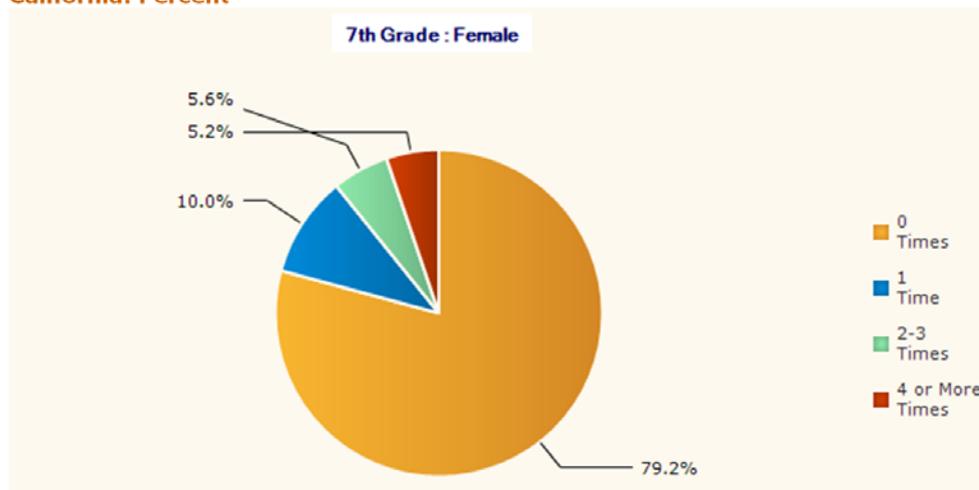
Reporting Area	18-49 years			50-64 years			65 years and older		
	White nH % (95% CI)	Black nH % (95% CI)	Hispanic % (95% CI)	White nH % (95% CI)	Black nH % (95% CI)	Hispanic % (95% CI)	White nH % (95% CI)	Black nH % (95% CI)	Hispanic % (95% CI)
TOTAL	18.7±0.3	19.7±0.8	17.0±0.9	37.3±0.5	29.3±1.8	31.1±2.6	68.1±0.5	49.4±2.5	57.1±4.4
ALABAMA	18.4±1.9	20.8±3.8	28.3±12.4	38.2±3.5	30.6±6.2	*	69.3±3.2	49.6±7.7	*
ALASKA	24.4±2.8	42.6±14.9	31.9±11.1	36.6±4.7	*	*	63.3±7.1	*	*
ARIZONA	17.5±2.4	*	17.3±4.3	34.0±3.9	*	24.8±9.0	66.5±3.5	*	52.9±12.0
ARKANSAS	21.2±1.7	19.0±4.2	26.4±9.8	39.6±2.8	30.6±9.5	30.9±15.4	68.6±2.9	50.6±12.0	*
CALIFORNIA	16.6±1.6	18.8±5.4	14.7±2.2	37.0±3.0	32.9±11.2	29.6±6.6	72.0±2.7	63.7±14.0	64.4±11.6
COLORADO	20.7±1.8	24.9±11.9	18.8±3.5	45.9±3.5	39.1±21.8	28.1±8.9	79.6±3.2	*	54.4±17.6
CONNECTICUT	19.1±1.3	21.8±5.8	22.9±4.3	38.2±2.2	26.3±9.1	56.6±11.8	71.0±2.2	56.0±12.9	60.5±13.6
DELAWARE	19.7±1.9	21.9±4.5	18.2±7.1	39.6±3.3	36.3±9.5	53.8±28.1	71.8±2.9	56.2±11.7	*

Pie Charts

- ▶ Ideal for depicting the size of each part as a percentage of the whole.
- ▶ Best when you have simple percentages and the “slices” of the pie are not too numerous. Avoid dividing the pie graph into too many “slices”. It can lead to confusion when interpreting it.
- ▶ Important to make sure the grayscale or patterns you use to represent the different "slices" are clear and distinguishable from one another. The best option is to display in color, if possible.

Physical Fighting at School, by Gender and Grade Level: 2006-2008

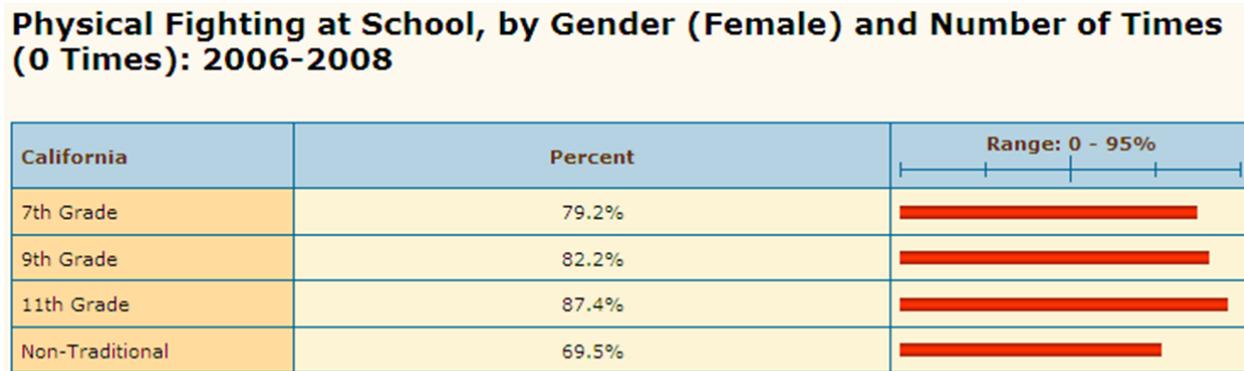
California: Percent



Source: California Department of Education, California Healthy Kids Survey (WestEd). <http://www.wested.org/chks>. As cited on www.kidsdata.org, a program of the Lucile Packard Foundation for Children's Health. Retrieved September 2011.

Bar Graphs

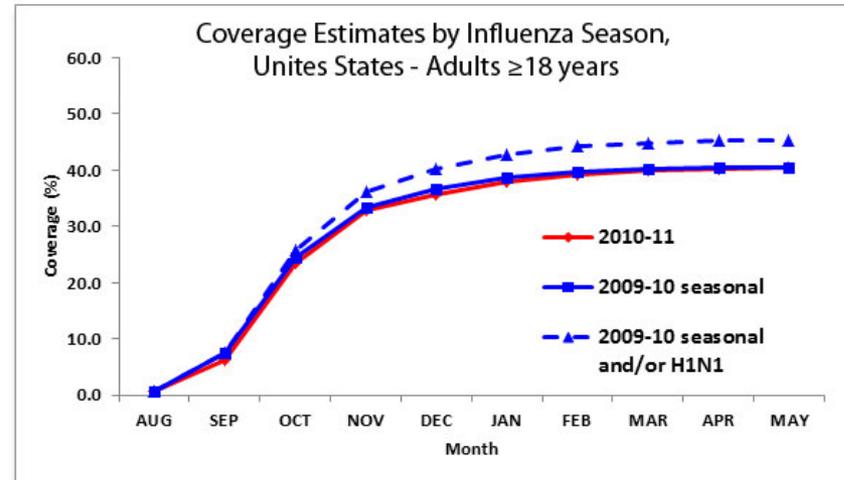
- ▶ Good for comparing quantities and percentages for a single category or timeframe – simple bar lines are easy to read and compare differences between groups.
- ▶ Avoid comparing things that are on different scales - uneven scales can lead to confusion when interpreting the graph
- ▶ Important to make sure the grayscale or patterns you use to represent the different "bars" are clear and distinguishable from each other. The best option is to display in color, if possible.



Source: California Department of Education, California Healthy Kids Survey (WestEd), <http://www.wested.org/chks>. As cited on www.kidsdata.org, a program of the Lucile Packard Foundation for Children's Health. Retrieved September 2011.

Trend Graphs

- ▶ Excellent choice when illustrating trends over time.
- ▶ Line movement, up and down, is easy to understand and interpret.
- ▶ Keeping trend graphs simple by including only a few regions can help avoid creating a graph with too many overlapping lines as seen below.



Maps

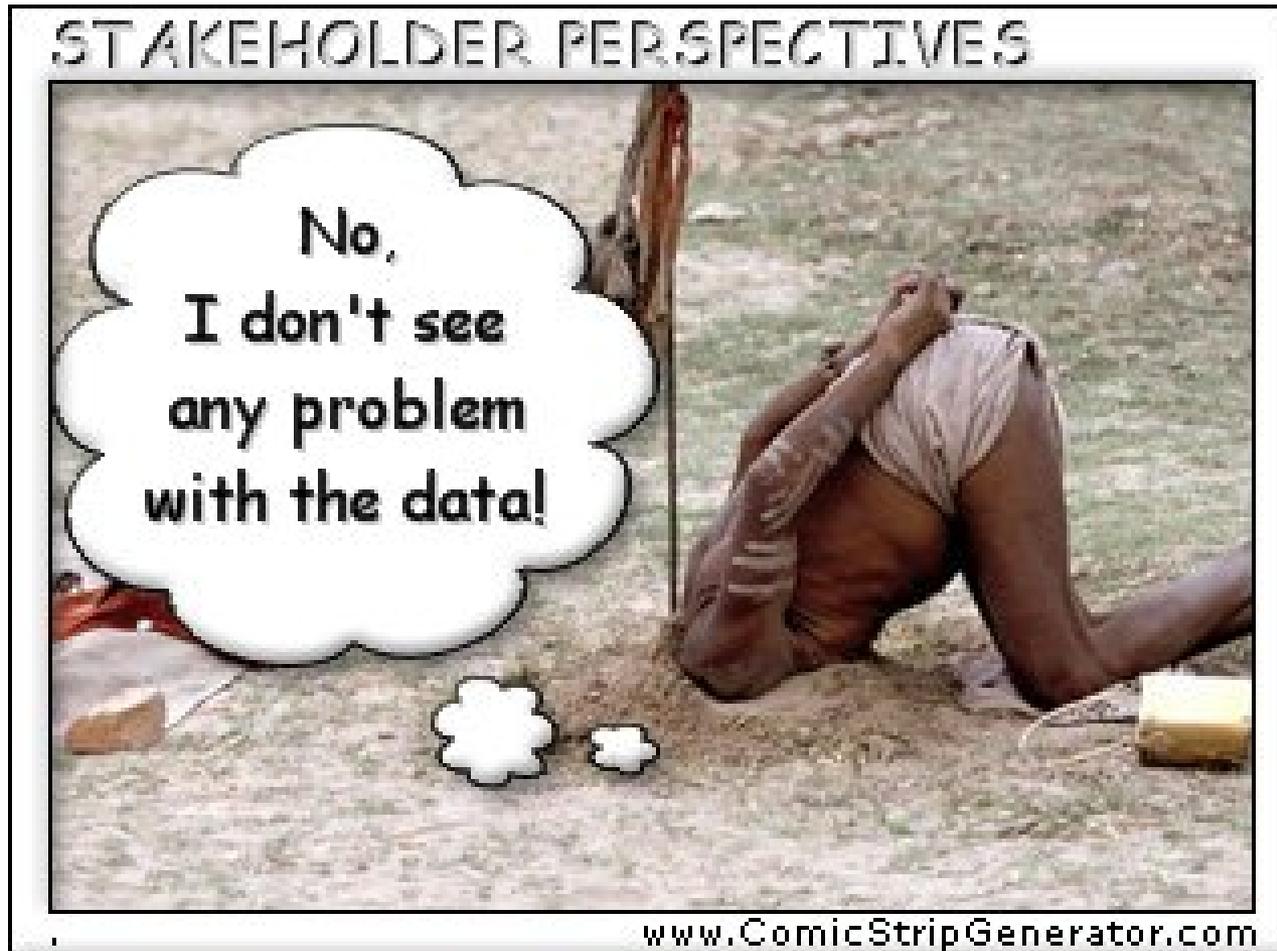
- ▶ Excellent choice when illustrating differences between geographic areas.
- ▶ Demonstrates areas of need or disparity through clear color coding.

Physical Fighting at School, by Gender and Grade Level: 2006-2008
(Grade Level: 7th Grade; Gender: Female; Number of Times: 0 Times)



Source: California Department of Education, California Healthy Kids Survey (WestEd). <http://www.wested.org/chks>. As cited on www.kidsdata.org, a program of the Lucile Packard Foundation for Children's Health. Retrieved September 2011.

Data Relevancy



Facilitators for accessing community data

- ▶ Community expertise
- ▶ Access to the population
- ▶ Experience with issues, solutions
- ▶ Relationships with stakeholders & policymakers
- ▶ History

Collect Your Own Data

- ▶ Start with what you know
- ▶ Research publications and database reviews
- ▶ Government data
- ▶ Legal research on laws and regulation
- ▶ Conduct community research—
Focus groups, interviews, surveys
- ▶ Combine quantitative data (the numbers) with
qualitative data (the words/stories)

Form a Community Research Collaborative

- ▶ Benefits of a Community Research Collaborative:
 - ▶ Pool resources and skills
 - ▶ Access data collection and analytic capability
 - ▶ Capture credible research resources to advance community interests
 - ▶ Expand the reach and acceptability of the data
 - ▶ Enables reflection of diverse community characteristics
 - ▶ Capture local environmental and social influences
 - ▶ Make sure results and reports are widely distributed



Completing Your Story – Incorporating Data

- ▶ Identify target audiences
- ▶ Present your findings
- ▶ Recommend action steps

Questions & Discussion



Health DATA



WWW.HEALTHDATA.UCLA.EDU

hdp@ucla.edu