



Center for
Health
Improvement

Persuasion with numbers: Introduction to data collection, analysis, and presentation

Karen Shore
Center for Health Improvement

Juneau Alaska
August 12, 2009

Why do we care about data?

- Data is used in nearly all aspects of policy-making
 - To describe the problem or problems faced by your community
 - To make your solution to the issue convincing to others
 - To measure progress

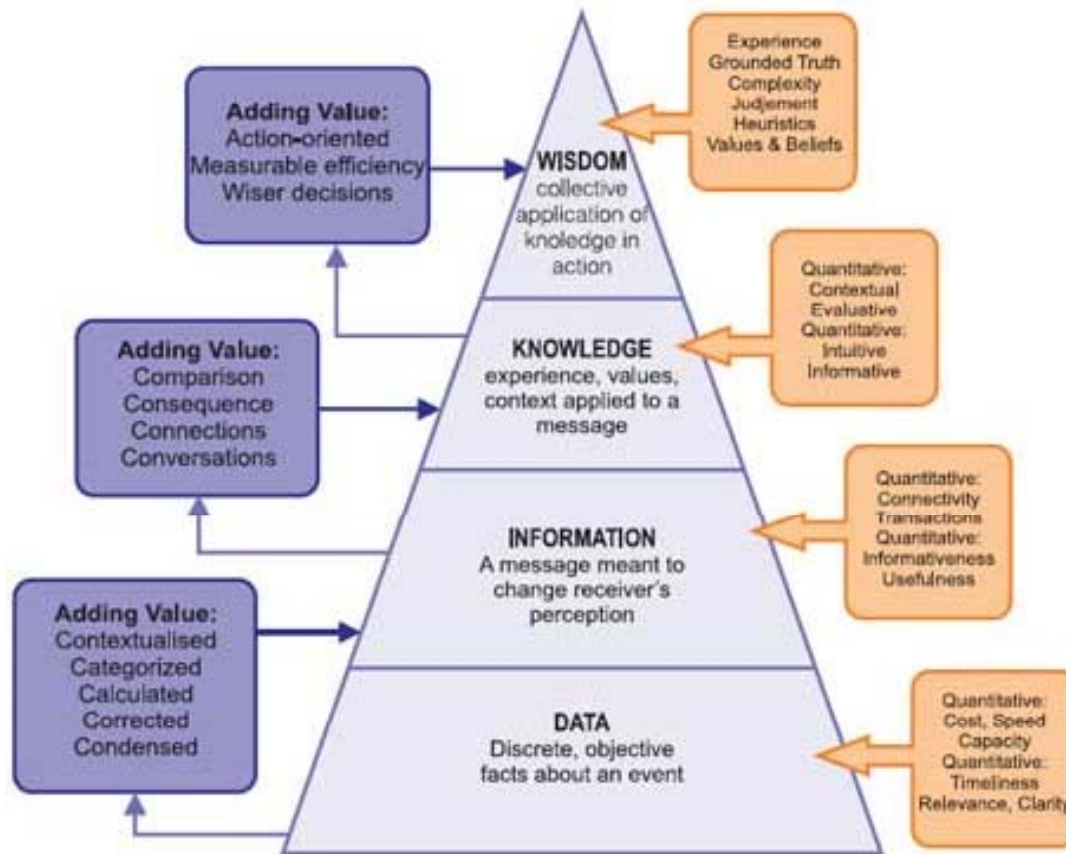


Numbers matter in policy making

“It is not simply the magnitude of numbers that matters for political or policy purpose. It is ***their interpretation*** that most determines how they influence policy making.”

Source: Mark Schlesinger, Health Policy by the Numbers, 2004,
Journal of Health Politics, Policy and Law

Building on data...



Types of data (1)

- Quantitative data – numeric information that is subject to statistical analysis
 - Provide the *what, who, when, and where*
 - Measurable and tangible
 - Involve counts of people, behaviors, conditions, or other discrete events
 - Examples – # of overweight people in your community, # of attendees at a town hall meeting on community health needs

Types of data (2)

- Qualitative data – text-based information that provides descriptive details, often collected from interviews, focus groups, or observations
 - Explain the *why* and the *how*
 - Involves asking or observing people in selected to discover how they feel and why they feel that way.
 - Examples – opinions about a health care provider, experiences with a community clinic

Uses of health data...

- Assess the health of a community or population
- Search for causes of disease, injury, and disability
- Plan programs and policies to meet community needs
- Measure progress in prevention and control efforts

Major types of health data (1)

- Health outcome (health status) data
 - Records (birth, death, marriage records), interviews, questionnaires, registries
- Risk factor data
 - Disease (high blood pressure), personal characteristics (age, gender, race), behavioral factors (exercise, diet, alcohol use), environmental factors (air pollution, contaminated ground water)

Major types of health data (2)

- Resource data
 - Health care facilities, health care professionals, # of smoking cessation classes, # of people completing a smoking cessation class
- Demographic data
 - Population (age, race, sex, personal income, marital status, educational attainment), and households (household size, structure, family type, and household income)

Example of BRFSS data

- Behavioral Risk Factor Surveillance System from Centers for Disease Control & Prevention (CDC)
 - Largest continuously conducted telephone health survey in the world
 - Monitor risk behaviors related to chronic diseases, injuries, and death
 - Data on behaviors, not attitudes or knowledge

BRFSS data on diabetes: SE Alaska

- Six percent of all respondents (6%) have been told by a doctor that they have diabetes
- The incidence of diabetes in Southeast Alaska (7%) is about the same as Alaska overall (8%)
- Nearly six in ten diabetics (59%) check their glucose levels at least once a day. ***However, 16% never check their glucose levels.***

Where do data come from?

- Primary data – data ***you*** collect specifically for the purpose of answering ***your*** questions
 - Surveys, key informant interviews, observations
- Secondary data – data ***someone else*** collected that you can use
 - Disease registries, national surveys like the Behavioral Risk Factor Surveillance System (BRFSS) and the Medical Expenditure Panel Survey (MEPS), cost or utilization data

Data collection methods...

- Surveys and questionnaires
- Focus groups
- Personal interviews
- Observation (e.g., of behaviors or actions)
- Knowledge or performance tests
- Document reviews



Copyright 2002 by Randy Glasbergen.
www.glasbergen.com



What to know before analysis begins

- Identify the purpose of the analysis or project
- Understand the sample(s) under study
- Understand the instruments being used to collect data
- Be aware of data layouts and formats
- Establish a unique identifier if matching or merging of data is necessary

Be sure to think about...

- Who is the audience?
- What is the intended use of the data (e.g., to support decisions and take actions or to monitor performance)?
- What is the basic message you want to communicate?
- What is the presentation format (report, brochure, oral presentation)?

Before we start with the analysis...

Copyright 2005 by Randy Glasbergen.
www.glasbergen.com



“We back up our data on sticky notes because sticky notes never crash.”

Data analysis: Making sense of the numbers

- Process of gathering, modeling, and transforming data with the goal of highlighting useful information, suggesting conclusions, and supporting decision making.
- Make a data analysis plan to include:
 - Descriptive statistics – describe the main features of a collection of data in quantitative terms (e.g., mean, range, frequency, correlations)
 - Hypothesis testing – through statistical tests (e.g., regression analysis, factor analysis)

What analysis program to use

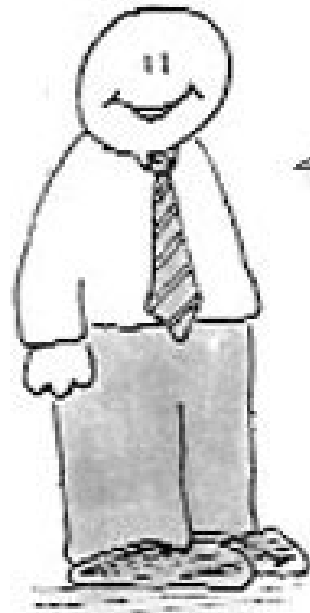
- Microsoft Excel – basic, accessible spreadsheet program
 - Good for general data exploration, simple graphs, histograms, scatter plots
 - Can customize tables
 - Easy to enter into other programs to complete analyses and write reports
 - But, NOT a statistical analysis program
- Statistical programs like SAS and SPSS are designed for specific analytic tasks; others for more specialized analyses

Telling a story with data

- Capture the attention of the public and decision-makers, using data and statistics
- On their own, statistics are just numbers
 - To have meaning, their value must be brought to life
 - Tell a story about the data:
 - What happened
 - Who did it
 - When and where it happened
 - Why and how it happened



GOVERNANCE IN ACTION



**Yes,
I am a data
optimist! Why
do you ask?**

www.ComicStripGenerator.com

What to include in your story

- Use the data to answer the following questions:
 - Is there a trend over time?
 - Should I take any action? What kind of action?
 - What contributes the most to the total (focus on the vital few)?
 - Are we focusing on the highest priority actions?

Communicating your story...

	Strengths	Weaknesses
Prose	<ul style="list-style-type: none">• Easiest way to explain patterns	<ul style="list-style-type: none">• Hard to organize a lot of numbers
Table	<ul style="list-style-type: none">• Holds lots of numbers• Good for detail• Predictable structure	<ul style="list-style-type: none">• Harder to “see” patterns
Graphs	<ul style="list-style-type: none">• Holds lots of numbers• Easy to see general patterns• Predictable structure	<ul style="list-style-type: none">• Difficult to see specific values
Maps	<ul style="list-style-type: none">• Easy to see geographical patterns	<ul style="list-style-type: none">• Difficult to see specific values

Prose: Social math

- It can be challenging to make large numbers meaningful to the audience
- Social math is the process of making data more interesting and understandable – drawing attention to your issue
- The goal of social math – to help people understand numbers and issues in a broad context

Avoid the “blizzard” effect



Social math – Statement A

- It is estimated that 947 million packs of cigarettes worth \$1.26 billion are sold illegally to children under 18 every year. These products generate \$221 million in profits for the tobacco industry. This is only a small part of the profits. In 1997, the four major American tobacco companies, Philip Morris, RJ Reynolds, Brown & Williamson, and Lorillard, reported profits of \$7.8 billion. Overall, the tobacco industry spends about \$5 billion of their profits per year promoting smoking in the U.S.

Social math – Statement B

- The tobacco industry spends more money promoting smoking in a week than the entire federal government spends on preventing smoking in a year.

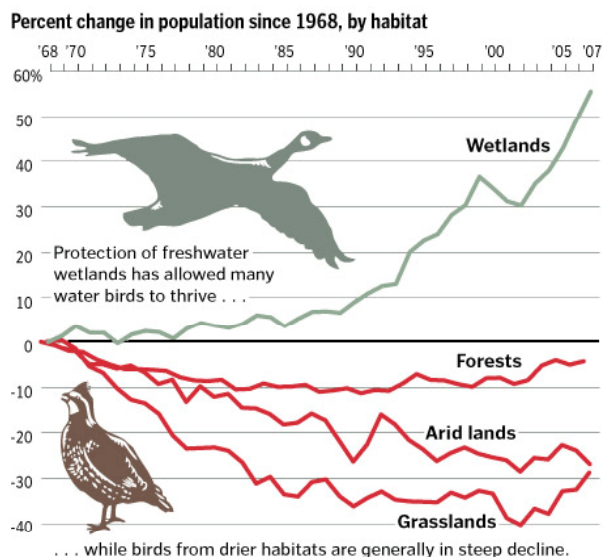
With prose, do...

- Use language that people understand
- Use simple language
- Make bulleted lists for easy scanning
- Use numbers in a consistent fashion (e.g., choose 20 or twenty)
- Use rounded numbers (1-2 decimal points)
- Use active voice

With prose, avoid...

- Table reading (i.e., describing every cell of a complex table in your text)
- Jargon and technical terms
- Acronyms
- “Elevator statistics” (This went up, this went down, this went up)

“A picture is worth a thousand words,
and a graph is worth 1.2 pictures.”



Using graphics

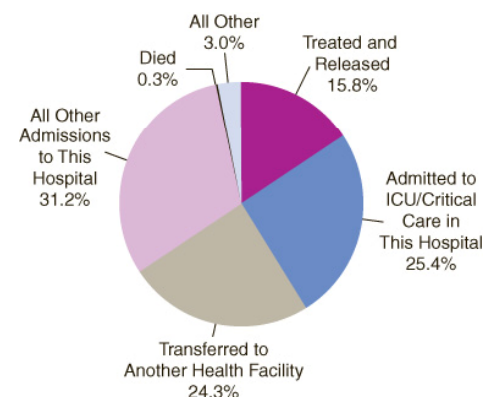
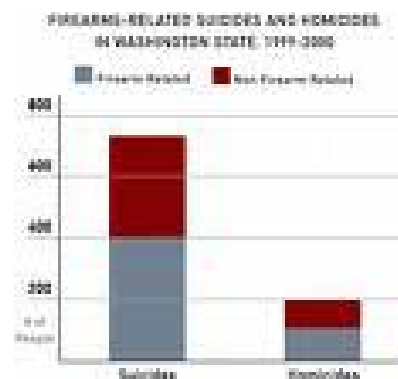
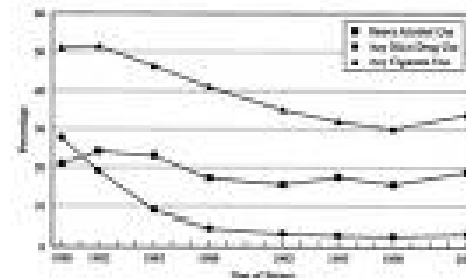
- Can be very effective in showing key results; work best when the data presentation is used to:
 - Communicate a message that is contained in the shape of the data
 - Reveal relationships among many values
- Good graphics:
 - Show the big picture by presenting many data points
 - Convey one finding or a single concept
 - Avoid extra information and distractions
 - Present logical visual patterns

When using graphics, do...

- Show the data
- Avoid distorting the data
- Make it as simple as possible while telling a story
- Induce the viewer to think about the meaning of the graphic and use it as intended, e.g., to make comparisons

Types of graphs

- When creating graphics, let the data determine the type of graph. For example:
 - Use a line graph to show trends over time
 - Use a bar graph for categorical data
 - Use a pie chart when trying to show what contributes most to the total



Don't insist on graphics

- If you can communicate your message clearly, efficiently, and with the desired impact in a simple sentence, do that.
- If your message requires the precision of a table of numbers and text labels to identify what they are, use those.

Good tables...

- Complement the written text and help minimize numbers in the written story
- Present numbers in a concise, well-organized fashion to support the analysis
- Eliminate the need to discuss insignificant variables that are not essential to the story

Tables work best when ...

- The data presentation:
 - Is used to look up individual values
 - Is used to compare individual values
 - Requires precise values
 - Values involve multiple units of measure

Acute or Heavy Alcohol Consumption Among Women of Childbearing Age (age 18-44 and Not Pregnant), by Public Health Region (PHR), Texas BRFSS, 2001-2002 Combined

Region	Sample Size	Percent %	95% CI	
			Lower	Upper
Texas Overall	3,398	13	12	13
Public Health Region				
1 Panhandle	154	8	7	9
2 Abilene/Wichita Falls	98	6	1	10
3 Dallas/Ft. Worth	795	13	12	14
4 Northeast Texas	163	9	6	12
5 Southeast Texas	133	14	2	25
6 Houston/Galveston	702	12	9	15
7 Austin/Temple/Waco	482	15	12	18
8 San Antonio	322	16	12	20
9 Midland/San Angelo/West Texas	101	8	7	9
10 El Paso	114	16	6	27
11 Lower Rio Grande Valley	269	12	8	17
95% CI				
Metropolitan Statistical Area	Sample Size	Percent %	Lower	Upper
Austin	267	17	13	20
Dallas	516	13	10	15
Fort Worth/Arlington	248	14	10	18
Houston	611	12	7	16
San Antonio	244	15	13	16

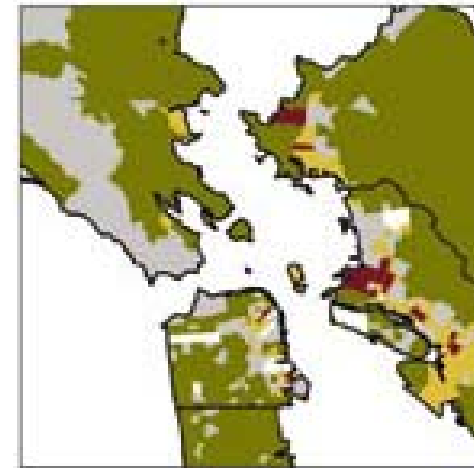
Note: Acute drinking is defined as having five or more drinks on at least one occasion during the past 30 days. Heavy drinking among women is defined as having an average of more than one drink per day.
 Source: Texas Behavioral Risk Factor Surveillance System, Statewide BRFSS Survey, 2001-2002
 Prepared by: Center for Health Statistics Texas Department of Health May 16, 2003

Tables should...

- Rank data by order to make the numbers easy to understand
- Show figures that are the highest and lowest
- Be simple (Save large complex tables for supporting materials such as an appendix)
- Have short titles that describe the table's precise topic or message

Maps

- Can be used to illustrate differences or similarities across geographical areas
- Can reveal local or regional patterns that may be hidden within tables or charts
- Most common type of statistical map –
 - Different shades of color are used to show contrast between regions (e.g., population density)



Summary

- Numbers matter in policy
- Making data useful is a process – from study design to collecting data to analyzing data
- Use data and statistics to bring life to a story; use the right tools for what you want to say and do
- Make your story of numbers resonate with your audience

Questions?

Contact:

kshore@chipolicy.org

916-930-9220