



The following slides are the property of their authors and are provided on this website as a public service. Please do not copy or redistribute these slides without the written permission of all of the listed authors.

Secrets of Data Presentation

April 14, 2011

Dr. Kenneth M. Coleman
Dr. Michael W. Traugott
University of Michigan
(mtrau@umich.edu)

 **SESRI**

معهد البحوث الاجتماعية والاقتصادية المسحية
Social & Economic Survey Research Institute



Secrets of Data Presentation

Dr. Kenneth M. Coleman

Dr. Michael W. Traugott

University of Michigan



THE ORGANIZATION OF DATA PRESENTATION

In any research report (published or supplied to a client), every table or figure should be able to stand alone.

That means the reader should be able to:

1) Understand what relationship was being investigated

2) Interpret what the results were

3) Come to the same conclusion that the researcher did.

Was the hypothesis confirmed or disconfirmed?

THE SECRETS OF DATA PRESENTATION

1. Identify where the data come from, which variable is which, and what the nature of the relationship is.

This enables the reader to reach the same conclusion as the author.



2. As a matter of reference, every presentation of data should be uniquely identified.

Label cross-tabulations as Table 1, Table 2, etc.

Label scatter plots and charts as Figure 1, Figure 2, etc.

This convention tells the reader something about the kinds of data used

KEY ELEMENTS OF A TABLE

1. The title describes the contents appropriately by variable.

Identification of the Context: includes a reference to time or place, usually at the end of the title.



KEY ELEMENTS OF A TABLE

Table 1. The Relationship between Sample Population and Preferred Mode of Interviewing

KEY ELEMENTS OF A TABLE

2. Identification of the **Data Source: Could be in the title or a footnote.**

Table 1. The Relationship between Sample Population and Preferred Mode of Interviewing, 2010 SESRI Omnibus Survey (Qatar).

KEY ELEMENTS OF A TABLE

3. Identify the **independent variable** (the explanatory factor) and the **dependent variable** (the factor to be explained):

Clearly label the rows and columns, by variable and category label

KEY ELEMENTS OF A TABLE

4. **Calculate percentages** across categories of the independent variable.

Define row and column %'s to guide the reader: only put a “%” at the top of the column or at the left of a row, depending on the direction of percentagizing.

Put “100%” at the **bottom of the column or the **end of the row****

Use the “N” notation to help as well

KEY ELEMENTS OF A TABLE

5. Give measures of association and **results of tests of statistical significance** at the bottom of the table or in footnotes associated with the table.

KEY ELEMENTS OF A TABLE

6. Use other **footnote conventions** such as:

Test statistics and levels of significance

Provide the full question wordings

Indicate any special features or qualifiers

Table 1. The Relationship between Sample Population and Preferred Mode of Interviewing

PREFERRED MODE	SAMPLE TYPE*		
	<u>Qataris</u>	<u>Expatriates</u>	<u>Migrant Workers</u>
In person	79%	72%	89%
By phone	5	9	2
Online survey	14	17	5
Don't know, refused	<u>2</u>	<u>2</u>	<u>4</u>
	100%	100%	100%
	(662)	(751)	(682)

¹ The question wording for Preferred Mode of Interview was “In politics TODAY, do you consider yourself a Republican, Democrat, or Independent?”

Sample Type was determined by the stratum from which the respondent was selected.

* Differences in preferred mode of interviewing and sample type are significant at $p < .01$.

KEY ELEMENTS OF A FIGURE

1. Title describes the contents appropriately by variable and contains the unique identifier.

Identification of the Context: includes a reference to time or place, usually at the end of the title.

KEY ELEMENTS OF A FIGURE

Figure 1. The Relationship between Ideal Age for Boys and Girls to Get Married

KEY ELEMENTS OF A FIGURE

2. Identification of the Data Source: Could be in the title or a footnote

**Figure 1. The Relationship between the Perceived Ideal Age for Boys and Girls to Get Married, 2010
SESRI Omnibus Survey**

KEY ELEMENTS OF A FIGURE

3. Use the axis convention:

Independent variable on the X or horizontal axis

Dependent variable on the vertical or Y axis

KEY ELEMENTS OF A FIGURE

4. Label the axes to show the scales used.

Make the units of measurement clear: use a legend if necessary.

Do not "break" either axis and distort the data.

KEY ELEMENTS OF A FIGURE

5. Give measures of association and results of tests of statistical significance at the bottom of the figure or in footnotes there

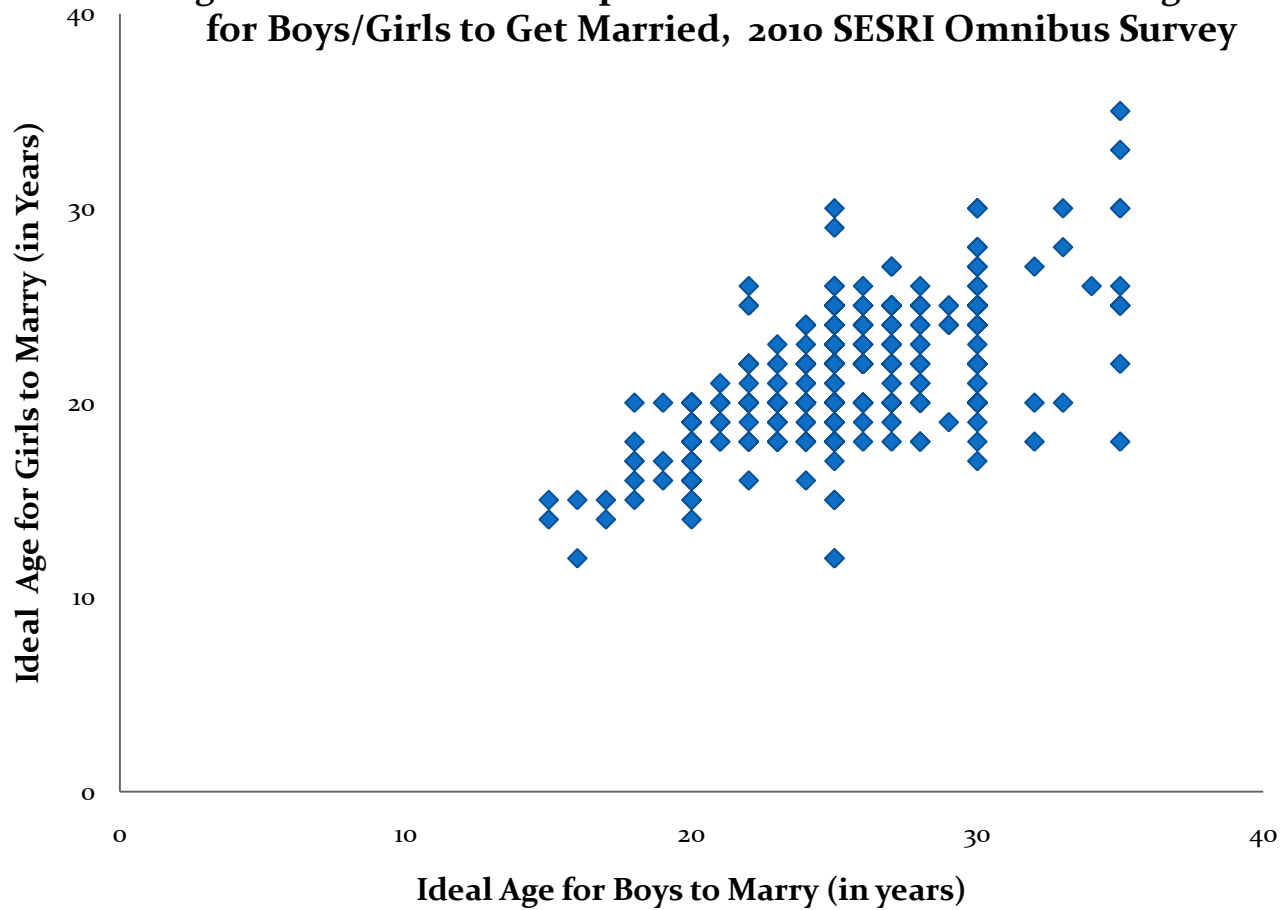
KEY ELEMENTS OF A FIGURE

6. Use other **footnote conventions** such as:

Give the full question wordings

Indicate any special features or qualifiers

Figure 1. The Relationship between the Perceived Ideal Age for Boys/Girls to Get Married, 2010 SESRI Omnibus Survey

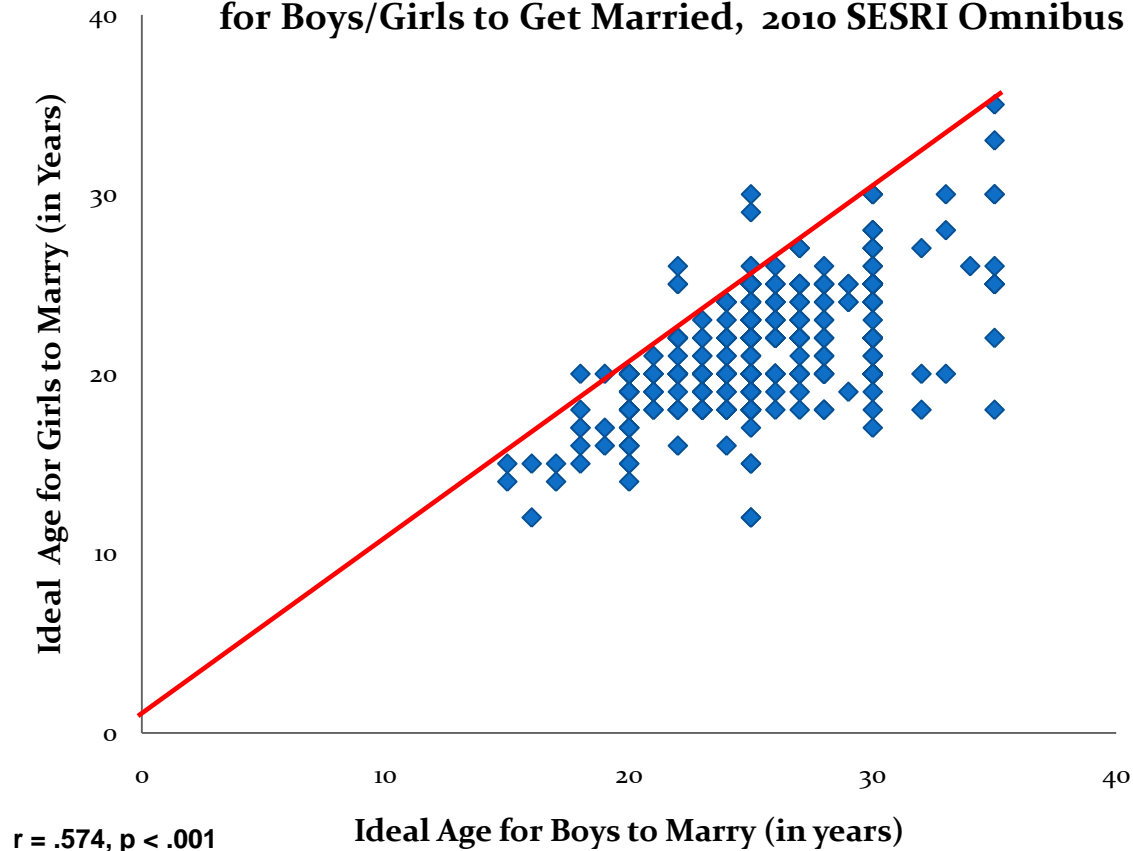


$r = .574, p < .001$

Question wording is “In your opinion, what is the ideal age for boys/girls to get married?”

What Does a Diagonal Line Contribute?

Figure 1. The Relationship between the Perceived Ideal Age for Boys/Girls to Get Married, 2010 SESRI Omnibus Survey



Question wording is “In your opinion, what is the ideal age for boys/girls to get married?”

FUNCTION OF GRAPHS and CHARTS

- **Visual representation of quantities**
- **Summarizing and interpreting quantitative information**
- **Show patterns, trends, anomalies that may not be immediately apparent in data tables**

USING SCALES CORRECTLY

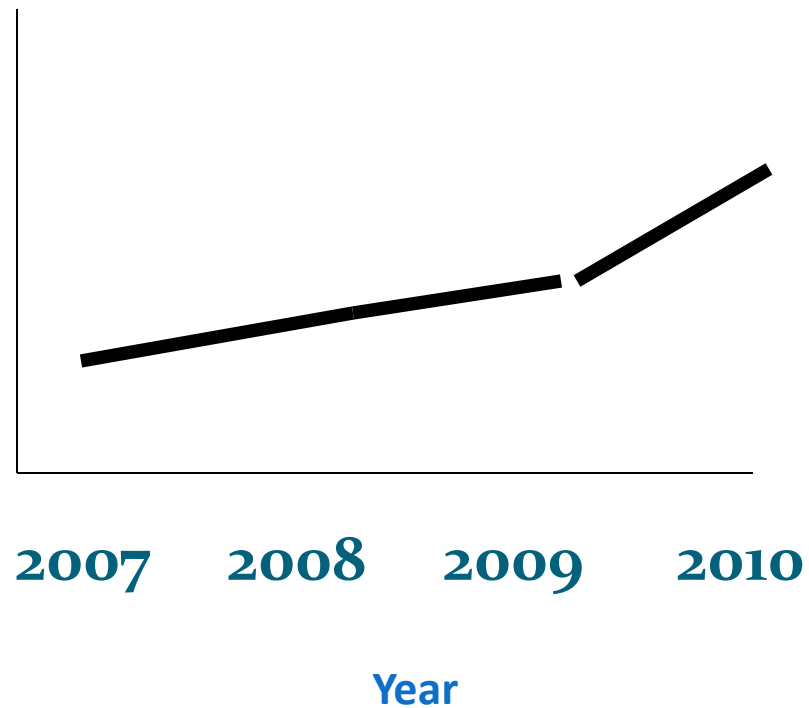
The most common problems with graphs and figures start with the inappropriate use of scales on the axes:

- 1. No scale**
- 2. No origin (suppressed origin)**
- 3. Different scales to emphasize/minimize change.**

USING SCALES CORRECTLY

NO SCALE

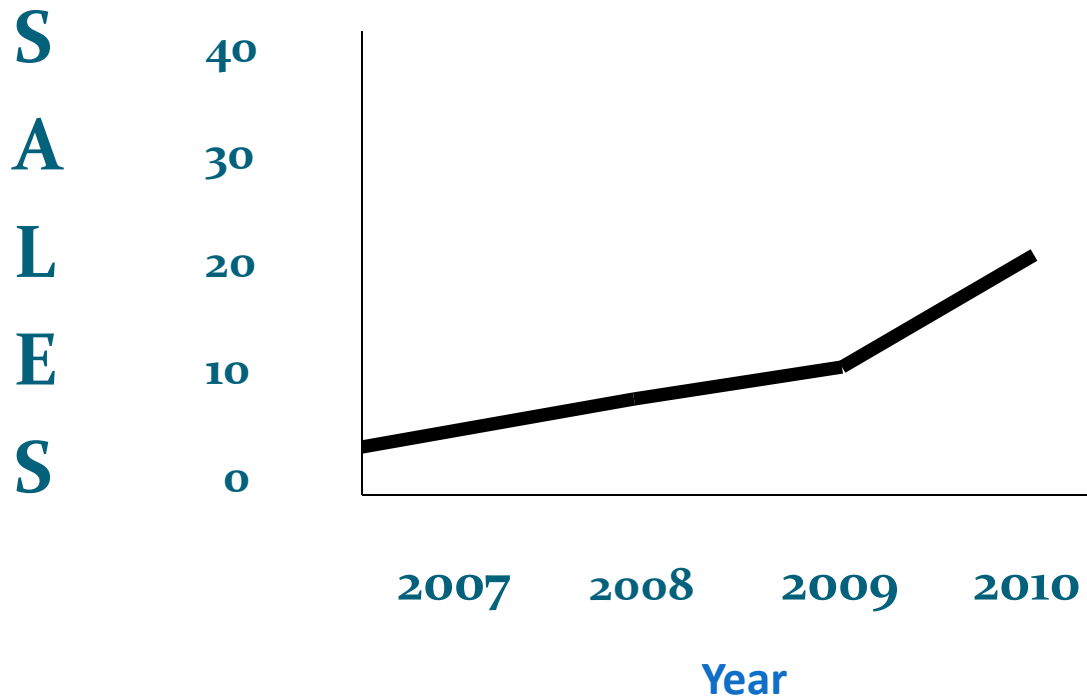
S
A
L
E
S



USING SCALES CORRECTLY

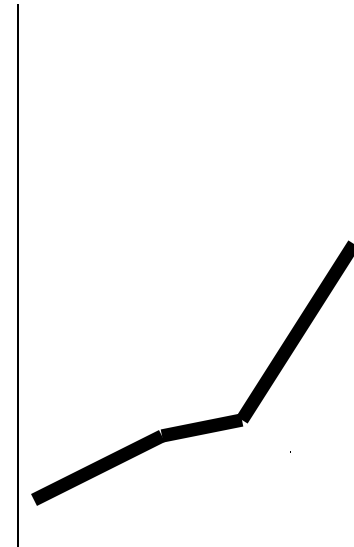
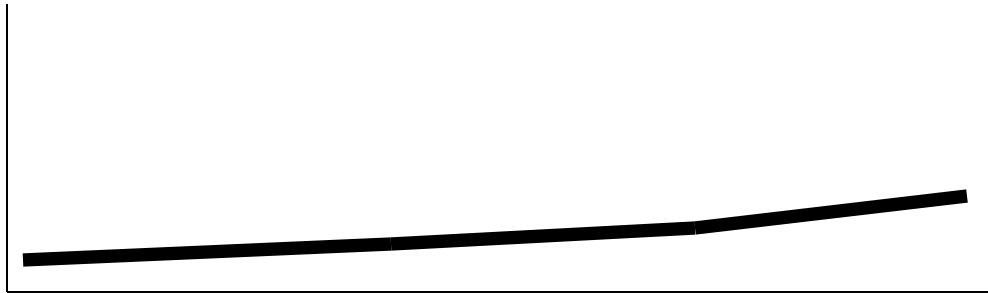
WITH SCALE AND ORIGIN

(\$1,000,000)



USING SCALES CORRECTLY

DIFFERENT SCALES SUGGEST DIFFERENT TRENDS



THE PROBLEM OF OBJECTS DISPLAYED IN 3 DIMENSIONS

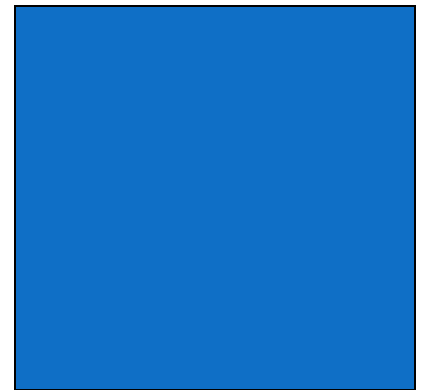
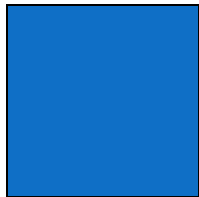
- **The media's frequent use of "objects" to suggest different quantities**
- **They can confuse the relationships between length, area, and volume**

For length, the relationship should be easy:



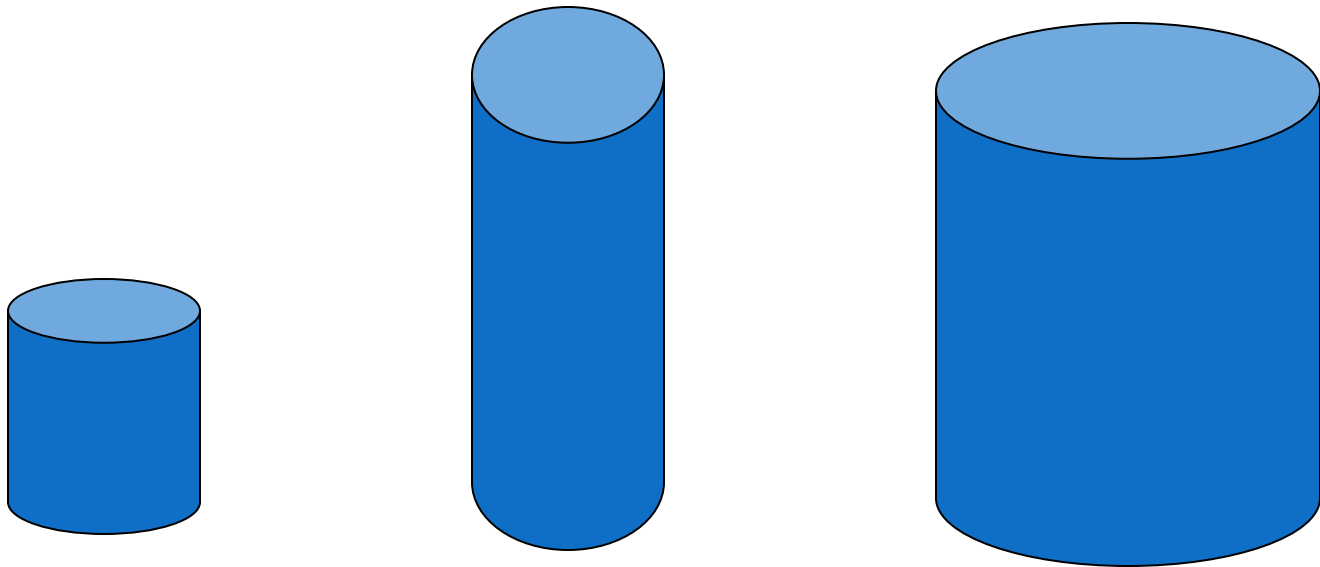
THE PROBLEM OF OBJECTS DISPLAYED IN 3 DIMENSIONS

For area, it gets more complicated because twice as long and twice as high is 4 times the area:



THE PROBLEMS OF OBJECTS DISPLAYED IN 3 DIMENSIONS

For volume, doubling length, width and height makes 8 times the volume



THREE TYPES OF GRAPHS/ CHARTS

- **Maps: Geographic distributions of values**
- **Pie Charts: Show the proportion of “parts” in relation to a “whole”**
- **Bar charts (vertical and horizontal): Used to compare categories along a single measure by relative height/ length**

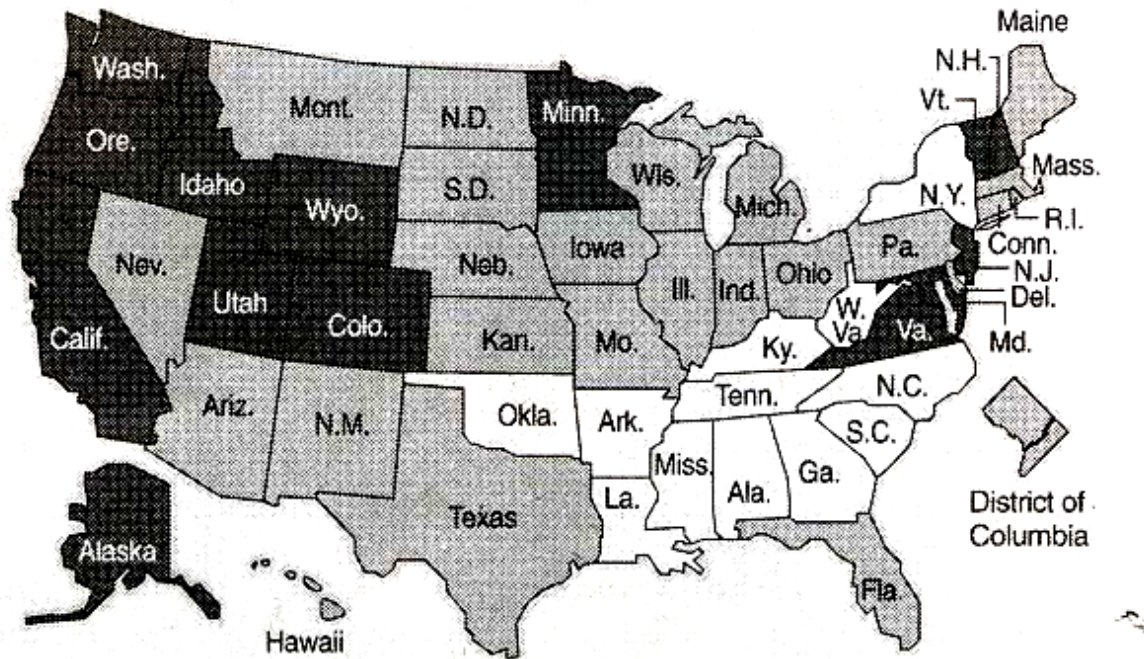
EXAMPLE OF A MAP CHART

Digital divide

In 1998, 42.1 percent of the nation's households had computers.

Percent of households with computers

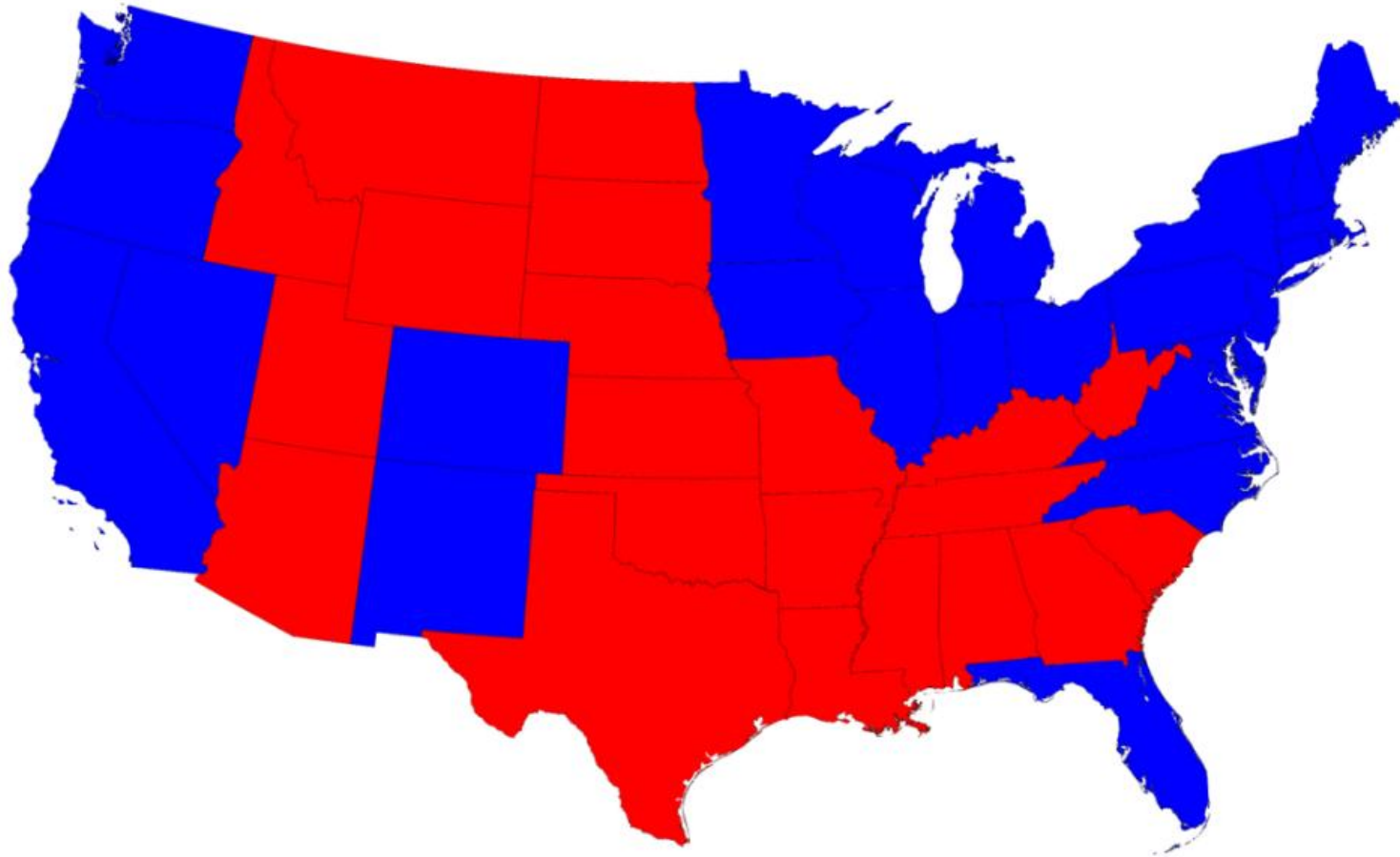
□ 26%-38% ▨ 39% - 44% ■ 46% to 62%



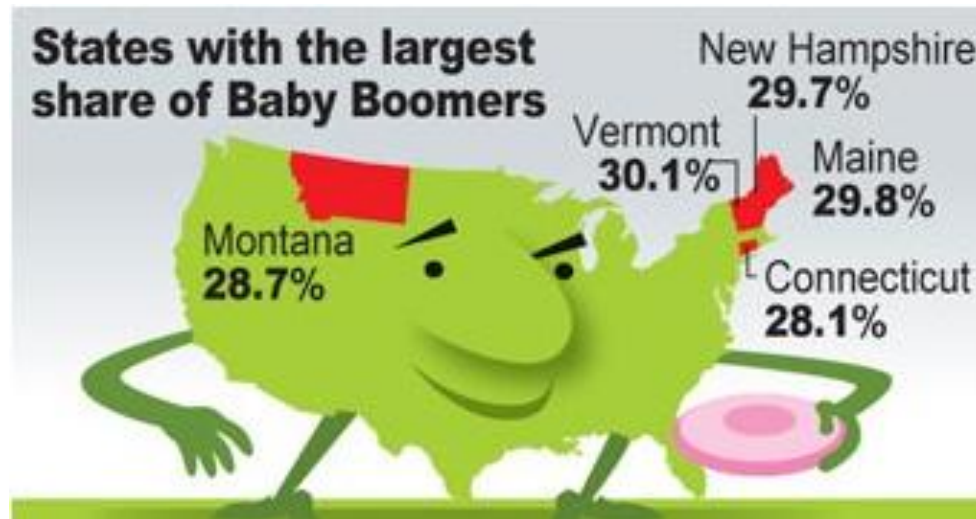
Source: Department of Commerce

Janet Loehrke, Gannett News Service

2008 ELECTORAL VOTE MAP



ALTERING A MAP WITH REFERENCE TO A SPECIFIC TOPIC



REFERRING TO GEOGRAPHY WITHOUT ANY MAP

Homeowners across the USA rate Florida as the most bug-ridden state and insecticide sales validate this belief.



PRINCIPLES OF PIE CHARTS

Best suited to variables with a few categories, distinctly making up 100% of a total

Permit comparison of categories as percent of the total

How to Create Pie Charts in Excel

- 1. Run the frequencies for the selected variable in SPSS**
- 2. Transfer the necessary information (results) into an Excel spreadsheet**
- 3. Select the appropriate Chart function in Excel after highlighting the data series**
- 4. Work through the labeling of “slices” of the chart and preparation of the title**
- 5. Copy the finished chart and paste it into a Word document**
- 6. Save the Excel spreadsheet in case adjustments are needed**

How to Create Pie Charts in Excel

We are interested in the presentation of attitudes about the sponsorship system or “kafala” (ATT10)

The first step is to run the SPSS Frequencies

How to Create Pie Charts in Excel

The screenshot displays the PASW Statistics Data Editor interface. The main window shows a list of variables with columns for Name, Type, Width, Decimals, Label, Values, Missing, Columns, Align, Measure, and Role. A 'Frequencies' dialog box is open, showing a list of variables on the left and 'att10' selected in the 'Variable(s):' field on the right. The 'Charts...' button is highlighted. The 'Display frequency tables' checkbox is checked. The bottom status bar indicates 'PASW Statistics Processor is ready' and the system clock shows 5:47 PM on 4/12/2011.

Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
hr09	Numeric	2	0	hh member's age	None	None	6	Right	Scale	Input
qol05	Numeric	1	0	overall, how wo...	{1, 1. excell...	None	7	Right	Nominal	Input
hr06	Numeric	1	0	gender of hh m...	{1, 1. male}...	None	6	Right	Nominal	Input
qol081	Numeric	1	0	past 12 months...	{0, 0. not s...	None	8	Right	Nominal	Input
qol082	Numeric	1						Right	Nominal	Input
qol083	Numeric	1						Right	Nominal	Input
qol084	Numeric	1						Right	Nominal	Input
qol085	Numeric	1						Right	Nominal	Input
qol086	Numeric	1						Right	Nominal	Input
qol087	Numeric	1						Right	Nominal	Input
qol088	Numeric	1						Right	Nominal	Input
qol089	Numeric	1						Right	Nominal	Input
qol0810	Numeric	1						Right	Nominal	Input
qol0898	Numeric	1						Right	Nominal	Input
qol0899	Numeric	1						Right	Nominal	Input
dem08	Numeric	1						Right	Nominal	Input
gen01a	Numeric	2	0	what is the idea...	{44, 44. god'...	None	8	Right	Nominal	Input
att10	Numeric	1	0	do you think th...	{1, 1. chang...	None	7	Right	Nominal	Input
att01	Numeric	1	0	expatriates and...	{1, 1. very tr...	None	7	Right	Nominal	Input
att07	Numeric	1	0	expatriates and...	{1, 1. very tr...	None	7	Right	Nominal	Input
att06	Numeric	1	0	expatriates and...	{1, 1. very tr...	None	7	Right	Nominal	Input

How to Create Pie Charts in Excel

The screenshot shows the PASW Statistics Viewer interface. The main window displays the following statistics for the variable "do you think the sponsorship system or kafala should be changed?":

Statistics
do you think the sponsorship system or kafala should be changed?

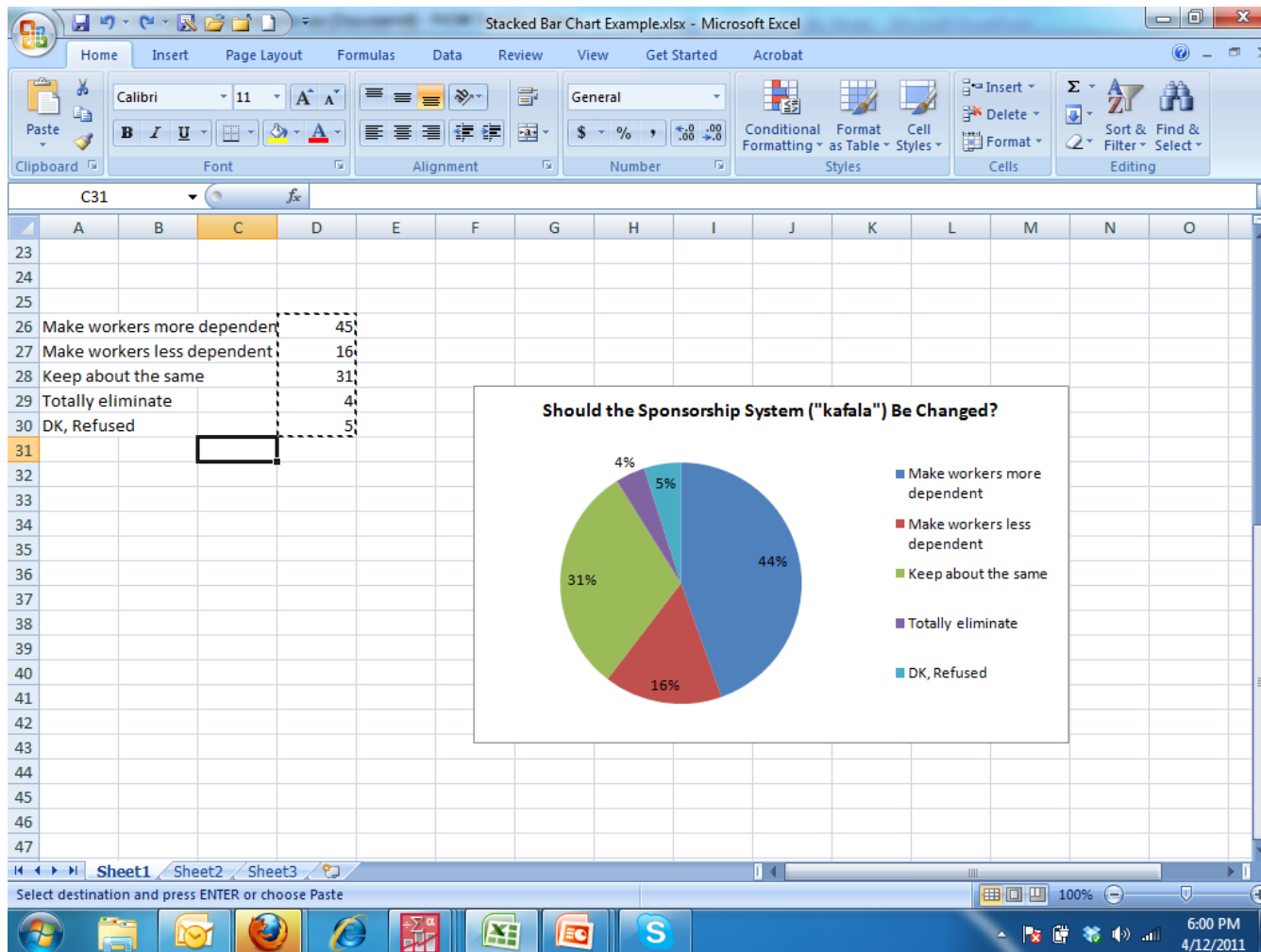
N	Valid	667
	Missing	1472

do you think the sponsorship system or kafala should be changed?

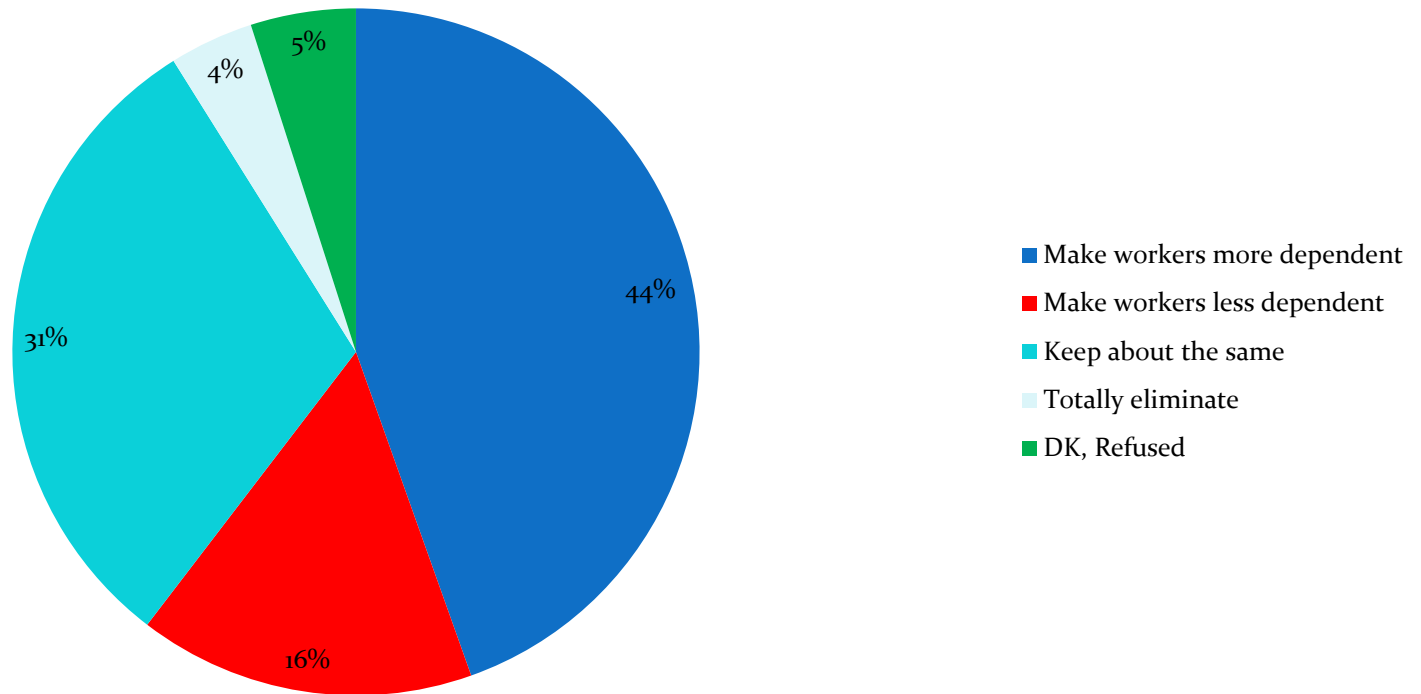
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1. changed/migrant worker more dependent	299	14.0	44.8	44.8
	2. changed/migrant worker less dependent	104	4.9	15.6	60.4
	3. kept about the same	208	9.7	31.2	91.6
	4. totally eliminated	24	1.1	3.6	95.2
	8. don't know	31	1.4	4.6	99.9
	9. refused	1	.0	.1	100.0
	Total	667	31.2	100.0	
Missing	System	1472	68.8		
	Total	2139	100.0		

The interface also shows a file explorer on the left with a tree view of the project structure, including folders for Log, Crosstabs, Title, Notes, Active Dataset, Case Processing, and Frequencies. The bottom status bar indicates "PASW Statistics" and "Solve PC issues: 2 important messages, 6 total messages". The system clock shows 5:50 PM on 4/12/2011.

How to Create Pie Charts in Excel

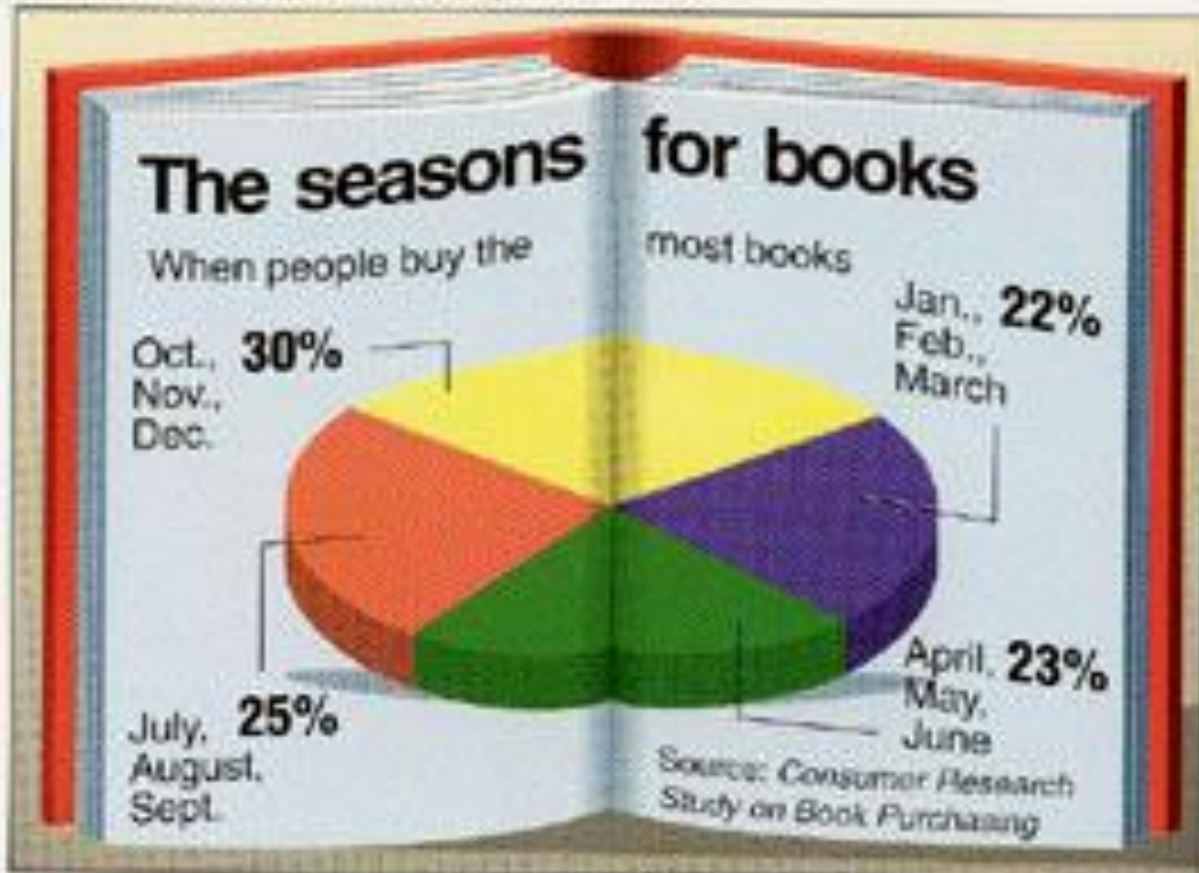


Should the Sponsorship System ("kafala") Be Changed?



USA SNAPSHOTS®

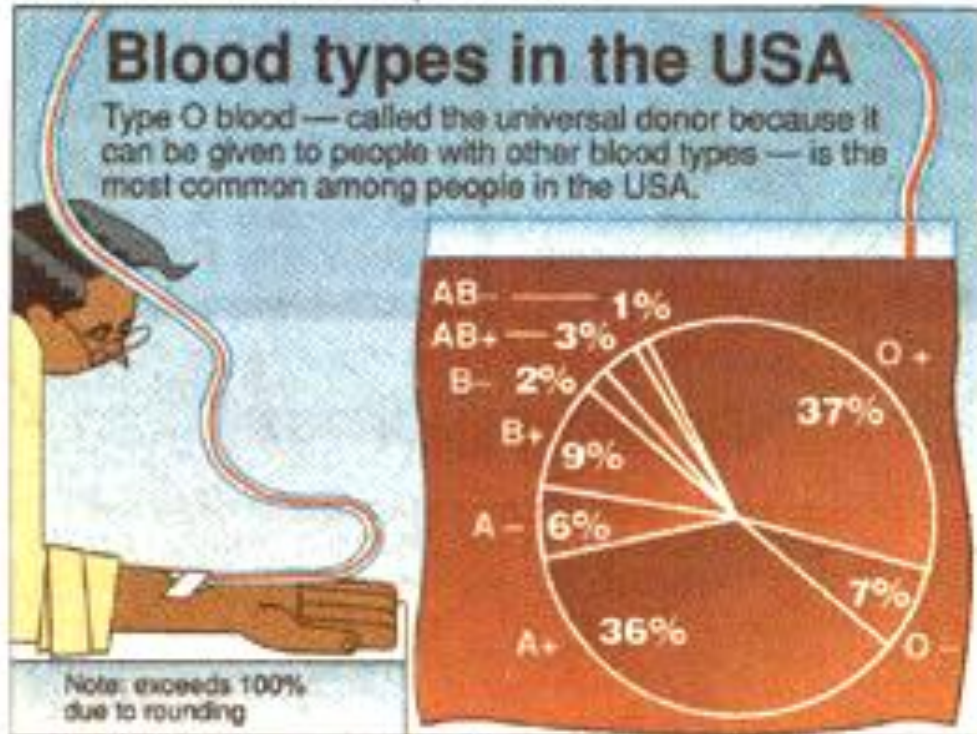
A look at statistics that shape our lives



By Jacqueline Blais and Nick Galifianakis, USA TODAY

USA SNAPSHOTS®

A look at statistics that shape our lives



Source: American Association of Blood Banks

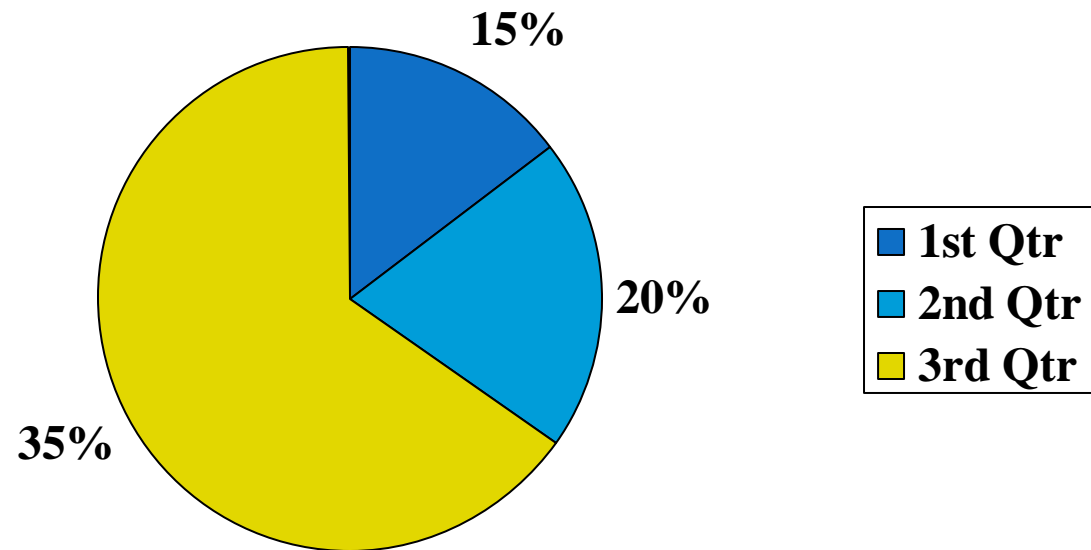
By Sam Ward, USA TODAY

COMMON PIE CHART MISTAKES

Including total as a category itself

Categories listed do not comprise 100% of the total; they are only a few parts of the whole (other parts are not listed, left out of the chart)

Yearly Sales



Problem?



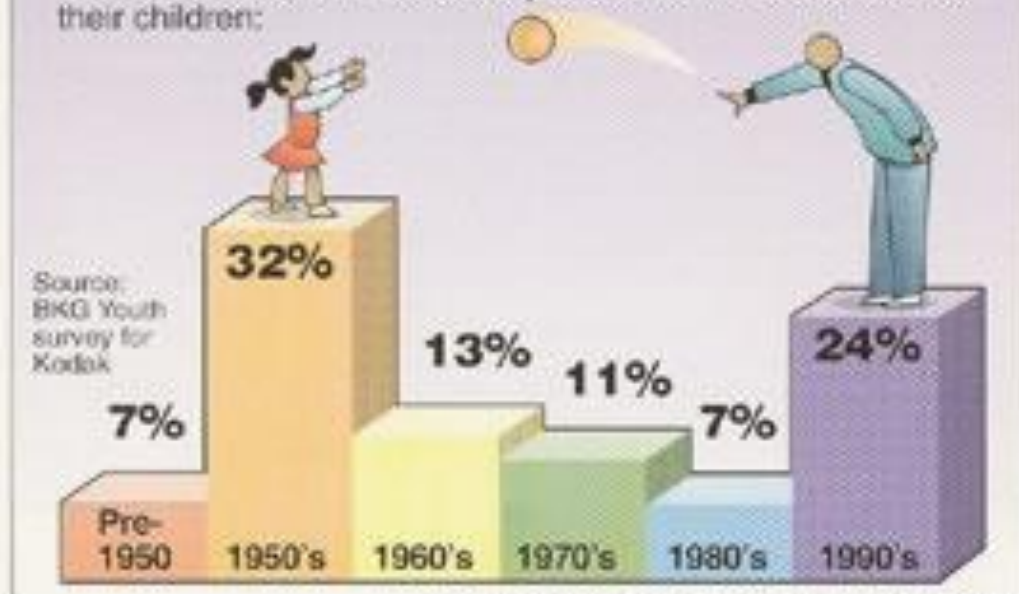
Problem?

USA SNAPSHOTS®

A look at statistics that shape our lives

Wanted: Time machine

Given a choice, which decade parents would select to raise their children:



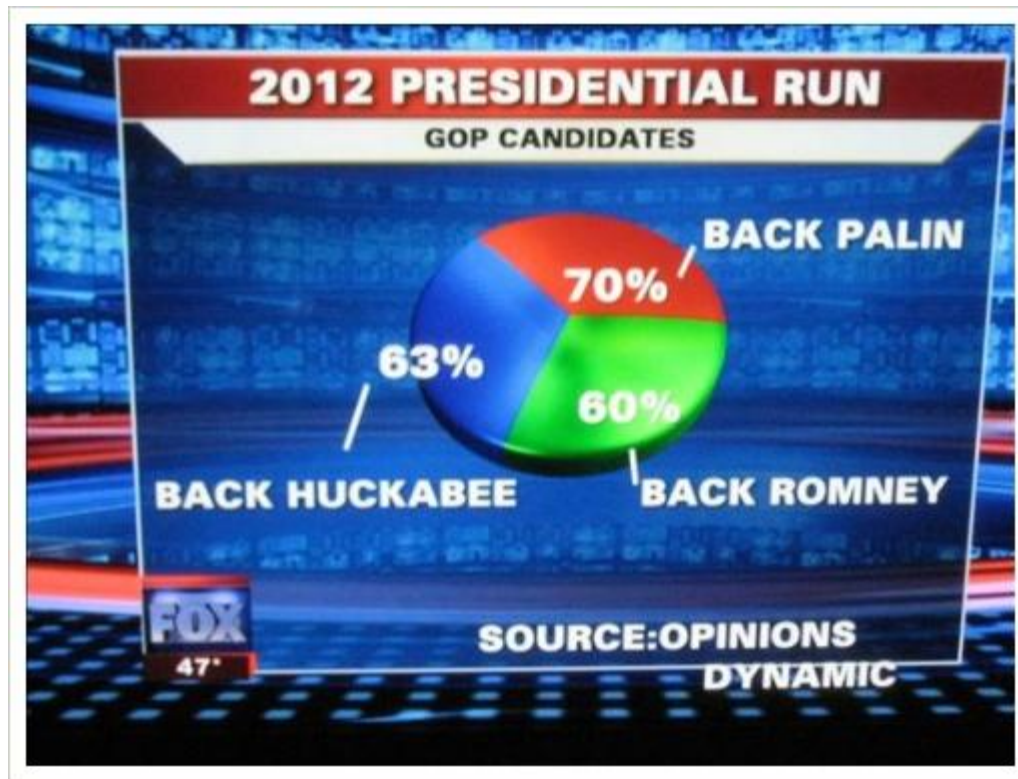
Source:
BKG Youth
survey for
Kodak

By Cindy Hall and Nick Galifianakis, USA TODAY

What's missing?



WHAT'S WRONG WITH THIS CHART FROM FOX NEWS LAST OCTOBER?



PRINCIPLES OF BAR CHARTS

Bar charts are used to compare different categories along a single measure

They can be very effective in displaying differences between groups

They can also be misleading.

How to Create Stacked Bar Charts in Excel

- 1. Run the basic data in SPSS**
- 2. Transfer the necessary information (results) into an Excel spreadsheet**
- 3. Select the appropriate Chart function in Excel after highlighting the data series**
- 4. Work through the labeling of axes, definition of the chart area, and preparation of the title**
- 5. Copy the finished chart and paste it into a Word document**
- 6. Save the Excel spreadsheet in case adjustments are needed**

How to Create Stacked Bar Charts in Excel

We are interested in the presentation of attitudes about the role of higher education of women (GEN06a – GEN06e) and how they differ by gender (HR06)

The first step is to run the SPSS Crosstabs

How to Create Stacked Bar Charts in Excel

The screenshot displays the PASW Statistics Data Editor interface. The main window shows a list of variables with columns for Name, Type, Width, Decimals, Label, Values, Missing, Columns, Align, Measure, and Role. A 'Crosstabs' dialog box is open, showing the following configuration:

- Row(s):** gen06a, gen06b
- Column(s):** hr06
- Layer 1 of 1:** (Empty)
- Display clustered bar charts
- Suppress tables

The dialog box also includes buttons for 'Statistics...', 'Cells...', 'Format...', 'OK', 'Paste', 'Reset', 'Cancel', and 'Help'. The status bar at the bottom indicates 'PASW Statistics Processor is ready' and the system clock shows 5:04 PM on 4/12/2011.

How to Create Stacked Bar Charts in Excel

*Second Runs for Data Display Examples.spv [Document4] - PASW Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Add-ons Window Help

higher education is good for woman because it helps her to have a good husband * gender of hh member
Crosstabulation

			gender of hh member		Total
			1. male	2. female	
higher education is good for woman because it helps her to have a good husband	1. strongly agree	Count	88	113	201
		% within gender of hh member	29.1%	31.8%	30.8%
	2. agree	Count	122	113	235
		% within gender of hh member	40.4%	31.8%	35.8%
	3. disagree	Count	80	124	204
		% within gender of hh member	26.5%	34.9%	31.1%
	4. strongly disagree	Count	12	5	17
		% within gender of hh member	4.0%	1.4%	2.6%
Total	Count		302	355	657
	% within gender of hh member		100.0%	100.0%	100.0%

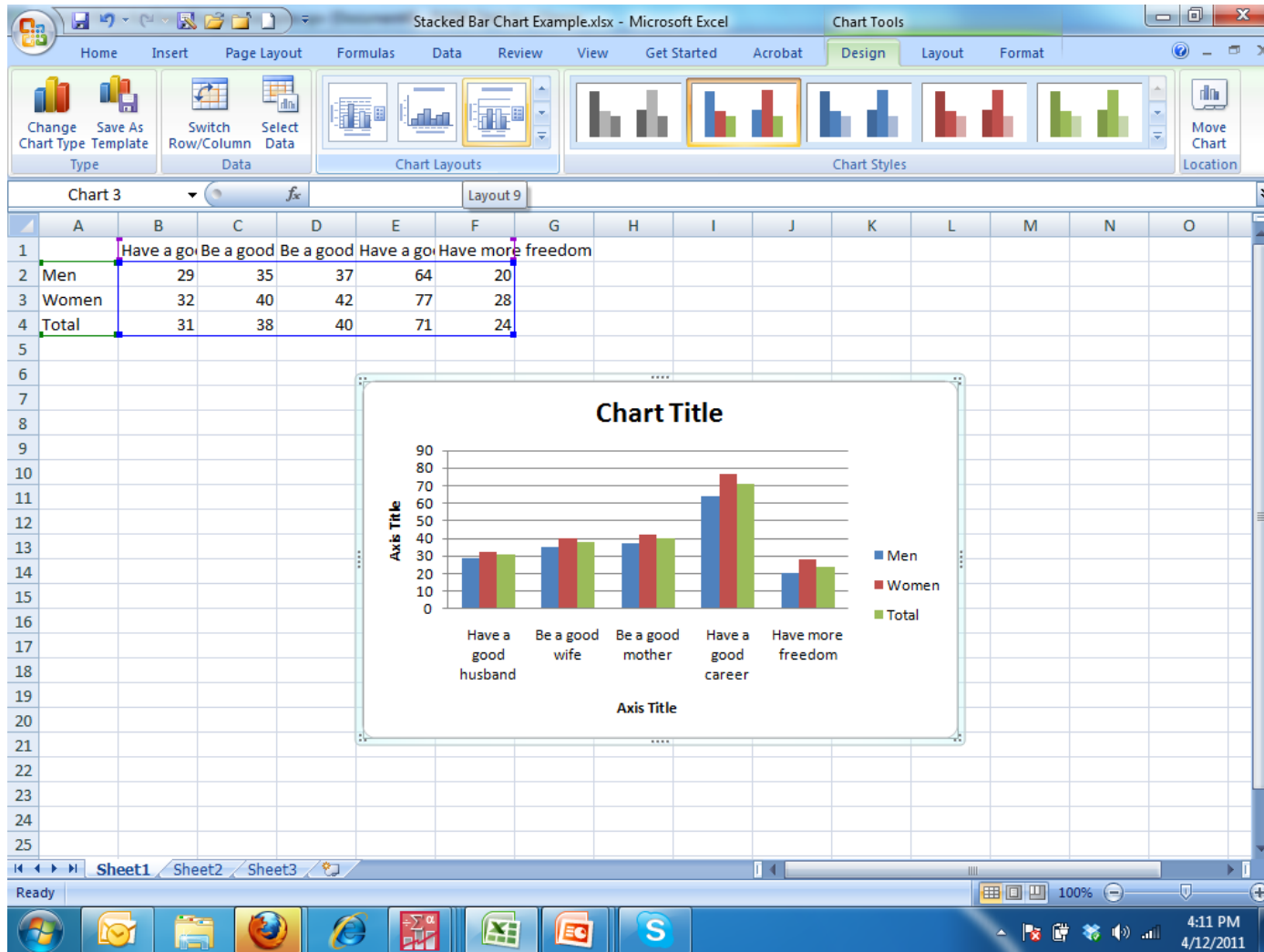
higher education is good for woman because it helps her to be a good wife * gender of hh member
Crosstabulation

			gender of hh member		Total
			1. male	2. female	
higher education is good for woman because it helps her to be a good	1. strongly agree	Count	106	144	250
		% within gender of hh	35.0%	40.3%	37.9%

PASW Statistics Processor is ready

5:10 PM
4/12/2011

How to Create Stacked Bar Charts in Excel



How to Create Stacked Bar Charts in Excel

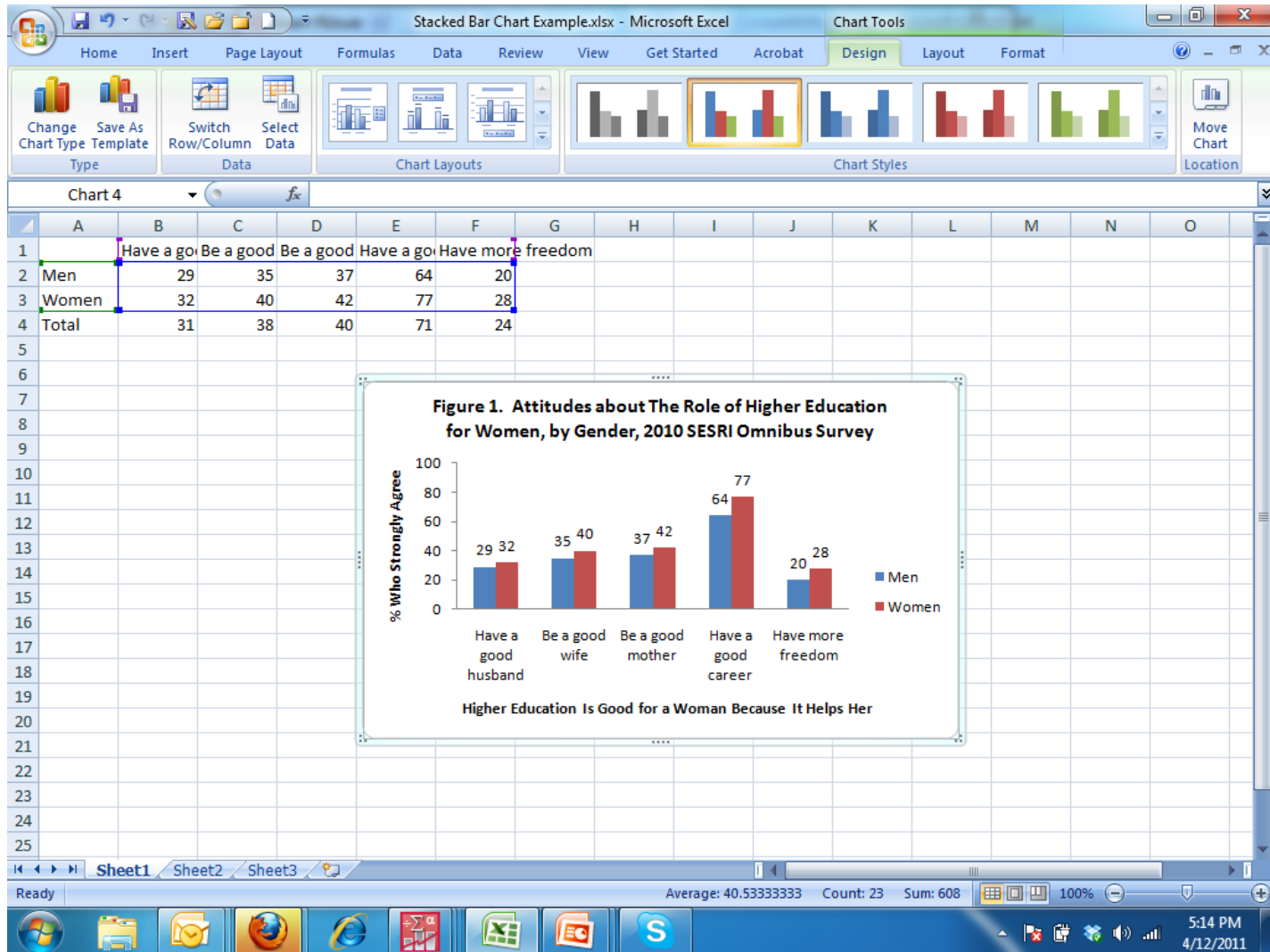
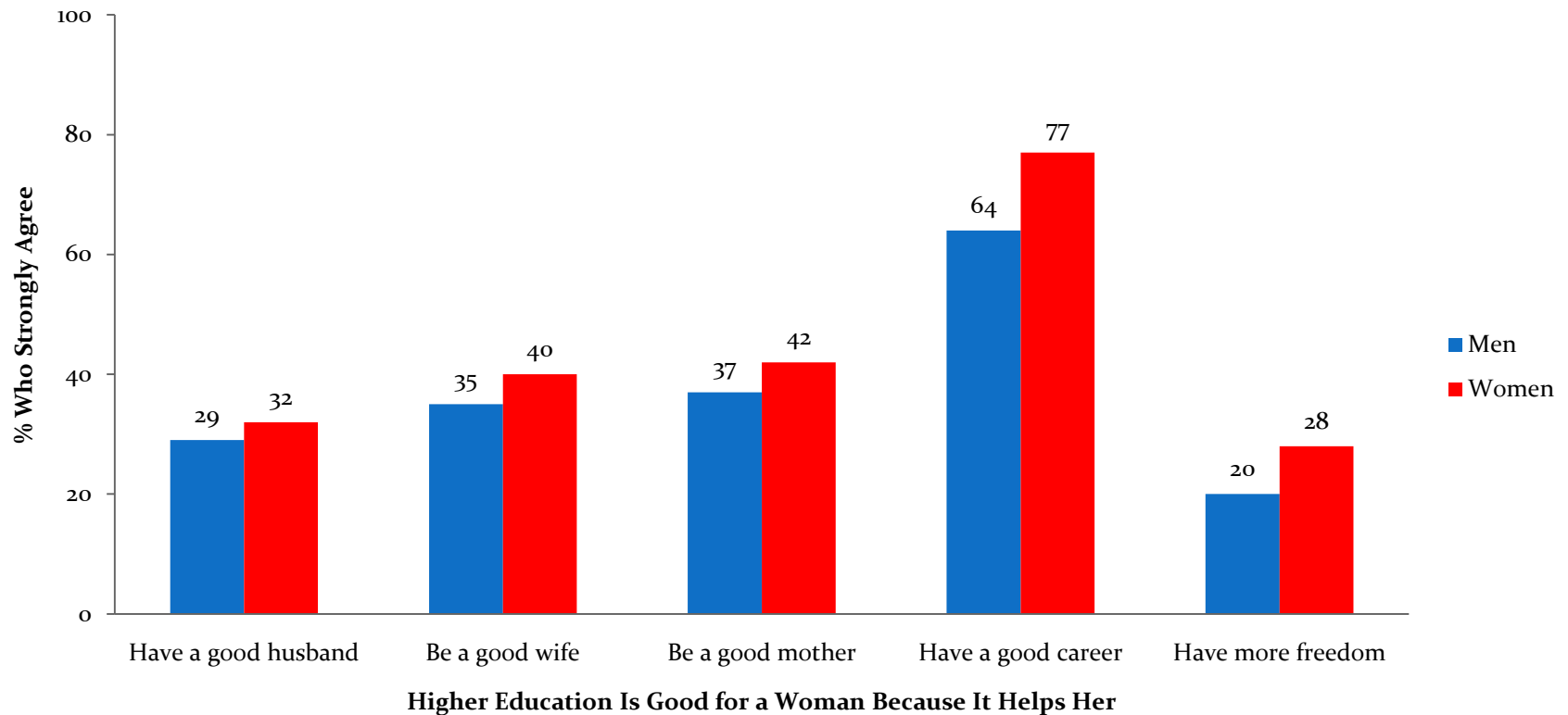


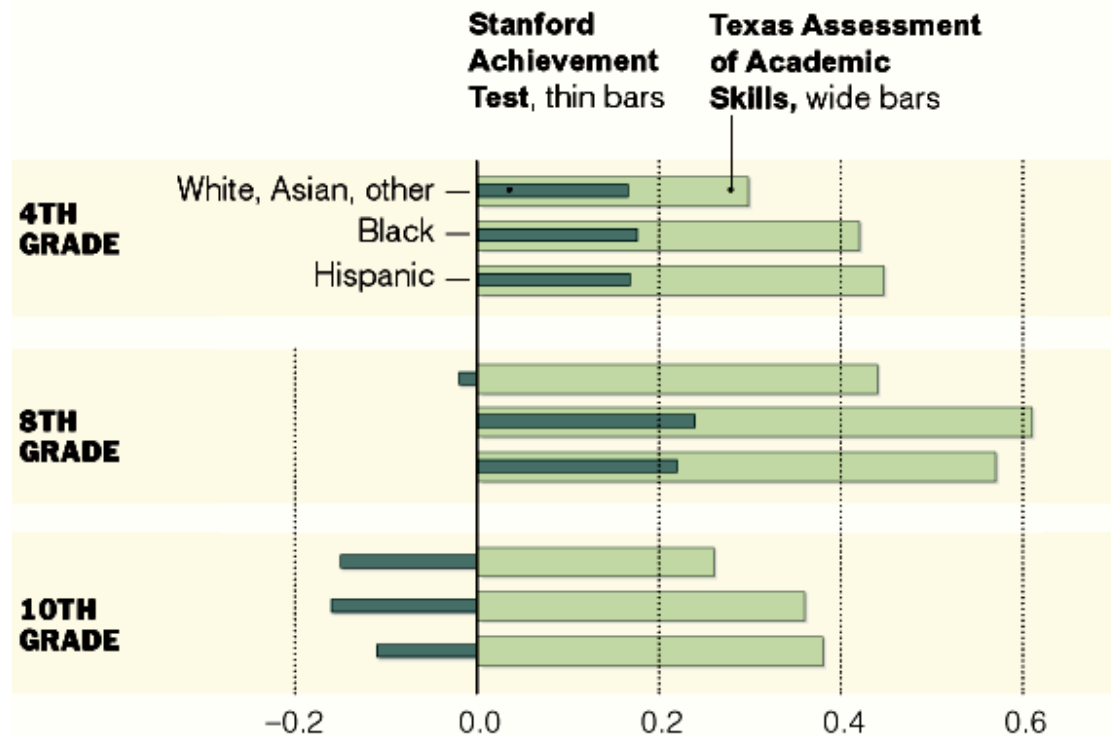
Figure 1. Attitudes about The Role of Higher Education for Women, by Gender, 2010 SESRI Omnibus Survey



EXAMPLES OF GOOD BAR CHARTS

A Tale of Two Tests

Relative change in reading test scores in Houston from 1999 to 2002.*



*Measured by effect size, a statistical measure that allows different kinds of testing to be compared.

COMMON BAR CHART MISTAKES

“Breaking” the axis so that differences are distorted (maximized)

Not paying strict attention to the scale so that “differences” are exaggerated



USA SNAPSHOTS®

A look at statistics that shape your finances

Education pays dividend

Estimated lifetime earnings for those who have:



Source: Census Bureau

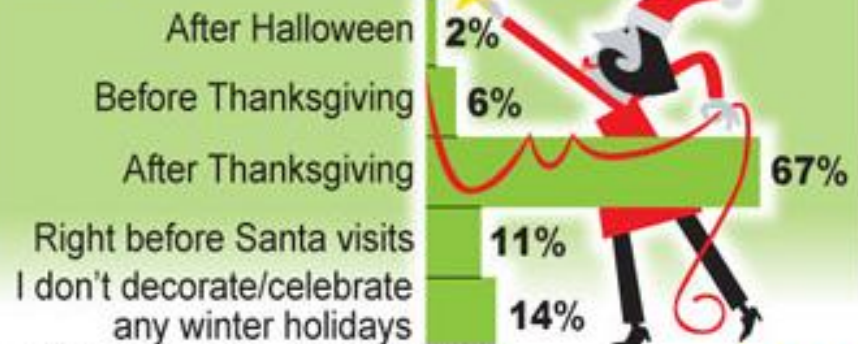
By Cindy Mall and Marcia Staimet, USA TODAY



PROBLEMS WITH TITLES/ LABELS

Among those who plan to decorate, the largest percentage (32%) say they intend to begin 30-37 days before Christmas.

When is it time to 'deck the halls' for the winter holidays?



By Michelle Healy and Sam Ward, USA TODAY
Source: 3M survey of 600 women by Wakefield



PROBLEMS WITH TITLES/ LABELS



PROBLEMS WITH TITLES/ LABELS

Nearly a third of adults ages 65 and older prefer to spend the extra hour reading a good book, much more than all other age groups.

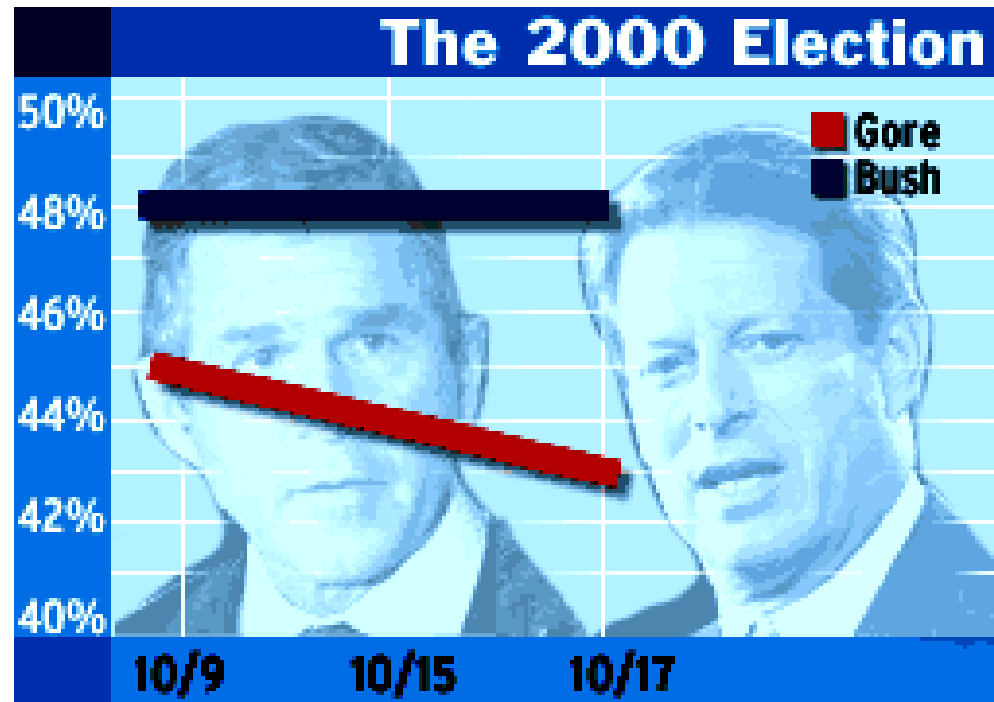


SOMETIMES THE DATA ARE JUST WRONG!



What Is a “Real Difference”?

Was either Bush or Gore really ahead?
(Tracking polls with 500 respondents each day)



What's the Difference Between the Population Size and a Rate?

