NUMBERS FOR REPORTERS

Prof. William Tillinghast
San Jose State University
AJEEP

AJEEP
Numbers for Reporters

Course Description

This course looks at various types of errors in news items that result from in accurate mathematics from basic mistakes in addition and subtraction to failure to calculate proper percent and per capita relationships. Errors often occur because of not comparing two or more numbers in a story (internal inconsistency) to not examining a news story normal to reality (external inconsistency). The course also focuses on the different way numbers are gathered, from simple nominal classifications to ratio level measures of physical reality.

Course Goals and Student Learning Objectives

The first goal is that students lose their fear of numbers through basic familiarization with how they relate to each other and to reality. The second goal is that students can write more accurate and relevant news stories if they are aware of how most errors can occur.

Required Texts/Readings

No text is required
Student may receive handouts and may search the Internet

Assignments and Grading Policy

Any assignments give will primarily focus on gathering addition Internet to support or rebut instructor comments as will finding examples of relevant lecture points. There will probably be two or three 20- or 25-point quizzes to determine whether students are learning techniques and procedures. If the quizzes are given different weights, a total of 100 points may be calculated and the grade based on the standard scale: a minimum grade of 60 to achieve a D; 70 to receive a C; 80 to receive a B; and 90 to receive an A

Numbers for Reporters

Class	Date	Topics, Readings, Assignments, Deadlines
1		Why do mathematical mistakes occur and why precision may less accurate than vagueness
2		Internal and External Consistency: Two or more related numbers in a story must fit with each other and all numbers in a story must match reality
3		Distinction between words and numbers and presentation of numbers in relevant context
4		Nominal to ratio measurement and when to use mean, median, and mode
5		Using percent for comparison of unequals and per capita for overtime distinctions
6		Proper sampling is not only cheaper but better than gathering information from an entire population

The six lectures in this course on numbers will each focus on a different aspect of numbers, the types of errors, and tips or tools to prevent their cropping up in your writing, or at least in remaining in your writing. This lesson focuses on why numerical mistakes occur, ways of preventing their occurrence, as well as when numbers should be used with words. Succeeding lessons focus on (2) jello numbers, and internal and external consistency; (3) sources of numbers; (4) levels of measurement; (5) percent and per capita; and (6) polls and statistics.

You should understand four things about this discussion:

1. The reporting of a number does not automatically make it correct
2. Words can be vague, but numbers are always precise
3. Numerical precision is not always numerical accuracy
3. The five reasons for most math errors

You should always read or listen to news stories in the mass media that contain numbers with a great deal of skepticism. Often – very often – the numbers are wrong!

An accuracy study of U.S. papers, published in 2007 found one of the highest error rates on record, almost 60 per cent of the articles had at least one error, or so the sources said. That percentage is even higher than that found in half a dozen studies since the first one done in 1937 found an error rate of about 50 percent. They weren't all numerical errors, but enough were.

Sometimes numerical inaccuracies occur because of the speed in gathering the news. Sometimes it may be because of the reporter's ability, or lack of it. But sometimes it is because numbers are not like most other words. Numbers are precise. There is no wiggle room. And numbers often are understood only in comparison, or contrast, with other numbers.

For example: It is a long way from Kansas City to Chicago. The victim was shot several times. The budget was raised three percent over last year.

This type of writing is worse than useless. Not only does it not convey sufficient information. It is a waste of time to read. MapQuest says it is 576 miles from Kansas City to Chicago. That is information a reader can use. The victim was shot four times in the stomach. The reader can visualize that. Three percent is useless unless I know either what the old budget was, or what the new budget is.

The long way from Kansas City is descriptive. But although it provides a general sense of description, it is a vague, general term. However, there is nothing vague or general about 576 miles. It is precise. And that is its strength, its value. **Unless of course, you use a wrong number.** A number is much more able to be shown to be wrong. If it is wrong, something that is vague.

The *Columbia Journalism Review* published an article that said there are five reasons that cover why most math mistakes happen. They are:

1. A journalist mishears a correct number given by a source and fails to double-check it.
2. A source unwittingly provides a mistaken piece of information and the journalist fails to verify it.
3. A source deliberately changes the numbers and the journalist fails to verify them.
4. A journalist or editor miscalculates a figure.
5. A journalist re-reports a mistake made by another media outlet.

Mistake #1 can happen for at least two reasons. The reporter is expecting a different number. There are background noises, from other reporters, etc. And some numbers sound alike, such as billion and million. Those two should always be checked in budget stories.

Mistakes #2 and #3 both imply that there are other sources which could validate or rebut if not refute figures provided by the source. It also assumes the reporter has time. Mistake #2 is more likely to happen because of a practice that too many reporters follow when calling for an interview. They often fail to tell the source the various topics they want to cover.

If sources know the issues, they undoubtedly check facts, figures and anecdotes to give to the reporter. A good story for the source is also a good story for the reporter.

Mistake #5 also occurs when reporters are under time pressure and are doing follow-ups on stories that have been printed in the previous issue of the competition.

They assume it is correct and just needs to be updated, as in automobile accident stories, or rewritten with a new lead if all the facts remain the same. So if ages, addresses or other numbers are wrong in the first story, they will be wrong in the second story unless that reporter double-checks, not just numbers but all facts. This, of course, requires time, a luxury in most newsrooms.

But remember, it is the principle that a veteran columnist advised a new reporter: If your mother says she loves you, check it out.

Two journalists who had slaved long hours putting a new state budget into a comprehensive story were called into their managing editor's office the next day. He told them he had received five calls from readers that morning complaining that the full newspaper column of agate (12 lines to the inch) budget department breakdowns totaled more than 100 percent. The editor told the reporters that the readers had to add the full column of two-decimal point numbers at least three times (this was before calculators) in order to ascertain that they were wrong and then the editor asked, if the readers could add them at least three times, why couldn't the reporters?

In other words: Do the math. Always.

Many math errors wouldn't have occurred if the reporter had asked whether such a number was possible, or whether it was meaningful. For example, Martin Gibson in a Writing with Precision column Publisher's Auxiliary in 1987 asked what is wrong with this sentence: In West Texas, the average boy starts shaving at 16.537. Does that make sense, Gibson questioned. It almost specifies the time of the day and what day after age 16. And how would one determine the accuracy to that degree?

An even more ludicrous is one involving just simple doubling. A doctoral student started his dissertations with this statement:

"Every year since 1975, the number of American children gunned down has doubled." His astonished advisor found that the quote was accurately cited from an academic journal. Anyone who takes the time to double each year will see the problem in a few minutes. For the sake of argument, we will start with just one child being gunned down in 1975. Doubling each year after that until 2012, we find that the number of children killed by guns each year has been as follows: (See PowerPoint)

This could continue and would surpass one quadrillion. But you get the point. Absurd? Of course, but only if you think about the number and what is being said.

Another math error committed by far too many people occurred as we entered the 21st century. What day was that? January 1, 2000 or January 1, 2001. Unless you think for a minute it seems logical to say the year 2000.

These are very straightforward scenarios, all of which can be solved with two basic actions. First, journalists need to acquire the basic math skills needed to properly handle numbers and figures. Second, they need to develop the habit of double-checking every number and figure.

The Heritage Foundation reported that 22,000 Americans who were below the poverty line had hot tubs. Conservative publications uncritically spread the word. However, this was an unscientific extrapolation from one case in a survey. It is beyond any logic that you can claim that because one poor family in a sample of 10,000 has a hot tub, 22,000 poor families have hot tubs.

The New York Times once began a story with this paragraph:

"For what experts say is probably the first time, more American women are living without a husband than with one."

If the *New York Times*, of all newspapers says it, then it must be true.

Except the Times used survey data that counted teenagers 15 through 17 as spouseless women. In other words, the *Times* based its lead, and its story, on the percentage of women 15 and older who were not married. Some 90 percent of them were living with their parents. They are still in high school. It is not legal to marry at age 15 in most states.

The value of numbers for journalists can be seen in this quote by William E. Blundell, who wrote The Art and Craft of Feature Writing (The Wall Street Journal Guide). Blundell stated that:

> "We need numbers in almost all our stories, and in some a number may be so important or startling that omitting or generalizing it would weaken the whole piece. I only argue that we be choosy in selecting figures and careful in their treatment.
>
> "In placing numbers in a story, the good writer tries not to stack too many in one paragraph; this builds a wall of abstraction difficult to breach. It becomes impossible to breach when two or more such paragraphs are butted together, a construction that may lead to more unread prose than any other writing fault. Don't do this. Don't ever do this."

A.K. Dewdney, an author of several books on mathematics said:

> "We do not expect reporters to be mathematical geniuses. But we do expect them to sidestep their mind-numbing fear of mathematics long enough to ask, 'Does this make sense?' 'What would I conclude from these numbers?"

Five Reasons Why Math Mistakes Occur

1. A journalist mishears a correct number given by a source and fails to double-check it.

2. A source unwittingly provides a mistaken piece of information and the journalist fails to verify it.

3. A source deliberately fudges the numbers and the journalist fails to verify them.

4. A journalist or editor miscalculates a figure.

5. A journalist re-reports a mistake made by another media outlet.

Skepticism and Logic

The New York Times once began a story with this paragraph:

"For what experts say is probably the first time, more American women are living without a husband than with one."

Inaccurate Numerical Base

The Times used survey data that also counted teenagers 15 though 17 as spouseless women.

The *Times* based its lead, and its story, on the percentage of women 15 and older who were not married.

Some 90 percent of them were living with their parents. They are still in high school. It is not legal to marry at age 15 in most states.

Numbers for Reporters -- 2
Jello Numbers and Consistency

A newspaper can report in a front-page headline that 350 homes have been destroyed in a fire in Southern California.

The next day it will report that the number destroyed was 235.

A television station can report on its evening newscast that about 1,000 people were injured when a meteorite exploded over Russia. It was said to have the force of 500 kilotons of TNT, 20 to 30 times the power of the bombs exploded over Hiroshima and Nagasaki, Japan in World War II.

Days later, that number injured is said to be nearly 1,200 people. The explosive power is now listed as 20 times the power of the World War II bombs.

How can these differences occur? Is someone lying? Can the differences be real over the change from one day to the next?

One of the two numbers on the destruction of the numbers has to be wrong? It cannot be 350 are destroyed the first day, but only 235 on the second. It is impossible for the 115 homes to be miraculously destroyed.

A reporter will add the estimates of several fire captains at different sectors of the fire, for example, 50, 100, 75, 80, and 45 to get 350. But suppose some of the sectors overlap and some of the destroyed homes are being counted twice. Suppose the first two zones both include the same 40 homes, the second and third include 30 in common, while zones 3 and 4 have 25 in common and zones 4 and 5 have 20 in common. That is 40, 30, 25, 20 or 115, the difference between 350 and 235.

On the other hand, if the total was 235 homes the first day, and 350 on the second day, that could be because more homes were destroyed after the first day's listing of fires.

The number of injuries suffered in Russia can increase as authorities get more reports of additional injuries and from medical authorities who had not previously reported. The difference in the explosive power of the explosion is harder to explain. Although one estimate hedges, by saying between 20 and 30 times the explosive power of the World War II bombs, compared to the estimate of just 20 times, the first is misleading. The first estimate provides one calculation and then increases it by 50 percent. This is too great a discrepancy.

These numbers are what I call jello numbers. They are shaky. They are subject to change, increasing or decreasing over time.

It is okay to report these numbers, but only as estimates, and only as estimates that are likely to change as more is learned.

Always attribute these numbers, and never use them in headlines. The source may be blamed for a contextual error, but the headline error is correctly assigned to the news organization.

These types of errors occur in at least two types of stories: breaking news stories and developing news stories.

A breaking news story is one that has just happened, that is happening at the moment it is being reported on. There is confusion among different news sources who are at different places within the news event, be it a natural disaster, a crime in progress or a governmental meeting. Knowledge changes and solidifies over time. But inaccuracies reported as stable fable will be remembered as errors committed by the news organization.

A developing story is similar to a breaking news story but it continues over the course of several days, weeks or longer, such as a manhunt, or a complex governmental action. There are often several sources of information, and some of them may not agree with others. They may also want to use the media to spread their perspective on the news and the issues.

Again, the solution is to always attribute information and to always indicate when the information is established fact, or an estimate.

Both internal and external consistency require that things are in agreement, both numbers and other facts. Internal consistency requires that elements within a news or feature story agree with each other. External consistency requires that the facts or numbers in a story agree with reality.

Numbers can be a common error in both types of stories.

Probably the most common internal inconsistency involving numbers has to do with age. It often occurs in obituaries.

The lead of an obituary might be as follows:

> Former U.S. Senator James Smith died at San Francisco General Hospital Monday following a long battle with pancreatic cancer. He was 73.

Later in the story, there will be references to his accomplishments in life and it might say that he was born February 3, 1950.

His age at death and his birth date are inconsistent. If he was born on February 3, 1950, he can only be 63 on February 18, 2013. Or he was 73, but he was born in 1940. Which is correct will have to be determined by other information but obviously one, or both, are wrong.

Almost any stories dealing with budgets, unemployment figures or other sets of numbers will be inconsistent.

Suppose the lead of a story in Kabul reports that:

> Afghanistan leaders on Tuesday unveiled a three-point plan on how to restructure The country's economic development after next year's withdrawal of American troops.
>
> Government leaders said provincial leaders had agreed to the new plan which focused on increased development of the country's oil and mineral reserves; the sale to the United States of all opium products produced annually; regional export of native vegetables; and development of a tourism industry.

The two paragraphs are numerically inconsistent. The lead lists a three-point plan, but the second paragraph specifies four distinct plans. Either (and probably) the lead should say four, or two of the four in the second paragraph should be combined.

If the lead of a news story lists the city budget for next year as $35.1 million and then later gives the breakdowns, add them up. You might see this:

City administration----------------	$ 4.75 million
Law enforcement -----------------	13.80
Fire department ------------------	11.35
Utilities ----------------------------	3.60
Animal control--------------------	0.75

The above numbers total $36.1 million, exactly $1 million more than the listed total of $35.1 million.

Internal inconsistencies can be spotted by any reader who does the math, who compares a total with the sum of its parts. A trained copyeditor should always examine any number in a story and ask (1) should the number fit with other numbers in the story, (2) does it fit with other numbers, and (3) which numbers are wrong, if the numbers do not fit together.

It is much more difficult to find and correct numerical errors that are inconsistent with reality. These are usually of two types, contemporary and historical.

If a news story reports that Afghanistan has good relations with all six of the countries on its borders: (list them), a copy editor either has to know the countries that border Afghanistan, or have access to a map in order to determine which and how many countries border Afghanistan.

External consistency is a lot harder to achieve. You have to know things about the world, people, governments, culture, history, and so forth. You have to at least even know that you don't, and you have to know where to check to find out that what is written in the news story is accurate, or what the correct information is.

For example, suppose a news story said that no English king or queen had ever lived to be as old as Queen Elizabeth II who is 80.

Is that true? First, is she 80? Second, where do you find a list of English kings and queens to find out how old they were when they died?

What are the rules, or guidelines to remember?

1. Unless I know better, each number referring to the outside world is wrong.
2. Therefore, each number needs to be verified
3. Check with almanacs, encyclopedias, government records, newspaper files, or with people who would know the correct information

These rules also apply to words that are suggesting numbers such as: most, least, highest, lowest, oldest, youngest, richest, etc.

As soon as you report a story using one of those words, someone will inform you that you are wrong. There is something or somewhat who is older, richer, etc.

If necessary, you will have to hedge a bit, such as one of the oldest, or one of the richest, etc.

What you are doing when you examine numbers in a story is essentially the same thing you do with other facts. You challenge the number's right to be in the story.

You turn the spotlight on the number and simply ask: Is this number right? How do I know or find out if it is right or not?

Number (Jello) Errors Over Time

A newspaper headline on Monday reports that fire destroyed 350 homes in Kabul.   The next day it said the number destroyed was 235.

One of the two numbers on the destruction of the numbers has to be wrong? It cannot be 350 are destroyed the first day, but only 235 on the second. It is impossible for the 115 homes to be miraculously destroyed. How can this happen?

A reporter will add the estimates of several fire captains at different sectors of the fire, for example, 50, 100, 75, 80, and 45 to get 350. But suppose, some of the sectors overlap and some of the destroyed homes are being counted twice.

Suppose the first two zones both include the same 40 homes, the second and third include 30 in common, while zones 3 and 4 have 25 in common and zones 4 and 5 have 20 in common. That is 40, 30, 25, 20 or 115, the difference between 350 and 235.

Numerical Errors in Obituaries

A newspaper lead reports that:

> Former U.S. Senator James Smith died at San Francisco General Hospital Monday following a long battle with pancreatic cancer. He was 73.

Later in the story, there will be references to his accomplishments in life and it might say that he was born February 3, 1950.

External Consistency Guidelines

1. Unless I know better, each number referring to the outside world is wrong.

2. Therefore, each number needs to be verified

3. Check with almanacs, encyclopedias, government records, newspaper files, or with people who would know the correct information

Numerical Inconsistencies

A news story begins:

> Afghanistan leaders on Tuesday unveiled a three-point plan on how to restructure the country’s economic development after next year’s withdrawal of American troops.
>
> Government leaders said provincial leaders had agreed to the new plan which focused on increased development of the country’s oil and mineral reserves; the sale to the United States of all opium products produced annually; regional export of native vegetables; and development of a tourism industry.

City’s $35.1 Million Budget

City administration----------------- $ 4.75 million

Law enforcement ------------------   13.80

Fire department -------------------   11.35

Utilities -------------------------------   3.60

Waste and trash -----------------   1.85

Animal control----------------------   0.75

NAME______________________________________

Numbers for Reporters Quiz #1

1. At least half a dozen newspaper accuracy studies have found that the percent of newspaper articles containing one or more errors is about:

   A. 5% B. 10% C. 23.5% D. 50% E. 80%

2. The difference between saying it is a long way from Kansas City to Chicago or saying Mapquest says it is 576 miles from Kansas City to Chicago is:

   A. A "long way" is vague, virtually useless
   B. A motorist cannot calculate driving hours using "long way."
   C. The precision of "576 miles" also makes it easier to prove right or wrong.
   D. A number's precision is its strength, unless the number is wrong.
   E. All of the above are correct

3. At least three of the five reasons for math errors cited by *Columbia Journalism Review* can be attributed to:

   A. deliberate falsification by one or more individuals
   B. reporter's failure to verify, or double-check, a number's accuracy
   C. what a number represents increases or decreases over time.
   D. using numbers for accuracy when description of words would be better
   E. false accuracy, such as young men start shaving at age 17.564 years

TRUE FALSE 4. The first rule of numbers is: "Always do the math." Do the addition, division, etc.

TRUE FALSE 5. The more numbers a reporters includes in a sentence, the more precision and understanding will follow.

TRUE FALSE 6. "Jello" or "shaky numbers" should probably not be used in headlines and not unattributed.

7. The most common numerical error is getting someone's age wrong, especially in obituaries. This is often detected by:

   A. Telephoning the mortuary or reporter who wrote the story.
   B. Comparing the age with those of other family members in the story.
   C. Comparing the age with events the individual participated in, such as subject is 94 in 2013 and was a professional basketball player in 1984.
   D. Looking at subject's face in photo accompanying the story.

8. A knowledge of history, social customs, etc. will help a reporter or editor spot errors that violate:

   A. internal consistency of story elements
   B. common sense
   C. external consistency
   D. improper use of numbers

9. The general guideline for numbers referring to the outside world is:

   A. each number is the article is probably wrong
   B. you need three independent sources for verification.
   C. they are usually typographical errors
   D. if one number is wrong, at least one number will be wrong.

10. You should avoid, or at least qualify or attribute these types of words:

   A. numbers as words, such as twenty-five for 25.
   B. oldest, richest, most, least, etc.
   C. synonyms for numbers
   D. Any quantity over 1,000.

Quiz #1 Answers

1. D
2. E
3. B
4. TRUE
5. FALSE
6. TRUE
7. C
8. C
9. A
10. B

Numbers for Reporters
#3– Sources of Numbers

Where do you get numbers?

You get them every day, from people, from observation, from government and public relations releases, from archival data.

Remember, words describe. They give a picture, a picture that varies for each person who reads or hears the words.

But numbers are precise. Five of something is five for each person who reads or hears the number. So, numbers are accurate. Well, most of the time.

Numbers often need a context, like a series of numbers representing previous years, or other similar items, such as countries. Suppose your news organization received the following drug use report by the United Nations, providing a select list of countries use of opiates in 2010, in alphabetical order.

	Users (Percentage of Population)
Afghanistan	2.65
Argentina	0.13
Australia	0.20
China	0.25
Columbia	0.92
Egypt	6.44
Iran	2.26
Israel	0.61
Myanmar	0.89
Mexico	6.04
Netherlands	0.31
Nigeria	6.70
Russia	1.54
United States	0.57

First, you might think the numbers should be in order of the largest percentage of users, such as this:

Nigeria	6.70
Egypt	6.44
Mexico	6.04
Afghanistan	2.65
Iran	2.26
Russia	1.54
Columbia	0.92
Myanmar	0.89
Israel	0.61
United States	0.57
Netherlands	0.31
China	0.25
Australia	0.20
Argentina	0.13

That would seem to imply that the largest drug-using countries are in the Middle East, Africa, and Mexico rather than the United States. But something is missing. These countries differ greatly in population. We need to include another number, the population of a country, from the CIA 2012 World Book. In that way we can multiply the population by the percentage of users to determine the counties with the most users.

China	0.25	1,343,000,000	33,575,000
United States	0.57	313,847,000	17,889,279
Nigeria	6.70	170,123,000	11,398,241
Russia	1.54	142,517,000	2,194,762
Mexico	6.04	114,975,000	6,944,490
Egypt	6.44	82,688,000	5,325,107
Iran	2.26	78,868,000	1,782,417
Mayanmar	0.89	54,584,000	4,857,976
Columbia	0.92	45,239,000	4,161,988
Argentina	0.13	42,192,000	584,496
Afghanistan	2.65	30,419,000	806,103
Australia	0.20	22,015,000	440,330
Netherlands	0.31	14,730,000	456,630
Israel	0.61	7,590,000	46,299

This raises a number of questions: Are you interested in the total number of opiate users by country for a company-by-company total comparison? Are you interested in the size of the drug problem in a country relative to the proportion of drug users in the country?

Last, but not least, why were these particular countries selected and the other 180 countries omitted?

Thus, the specific context, or lack of context, will give a particular shape to essentially the same pieces of information.

How can you tell when a newspaper is lying?

Here's a hint: watch out for the numbers. Newspapers are filled with many reports from government officials and corporate leaders. They invariably provide numbers to support their statements and opinions.

Journalists write the first draft of history.   They are in a hurry, are often ignorant about the topic they are writing on. They depend on what others tell them. These sources are often biased.

The demand for numbers often leads to in-depth reporting, rich in detail with numbers that aid in understanding.   But it can lead to using precise numbers for guesses. Most politicians use rosy forecasts of next year's budget in planning on spending and if there is a downturn in the economy, they have spent more money than was collected. Ergo a recession can occur.

There are many ways to mislead while allegedly presenting accurate counts or measures to the public.

One can simply make up numbers or just guess. It is expensively to determine all the numbers that reporters ask for.

The National Cancer Institute and the American Cancer Society said that American women have a one-in-eight chance of getting breast cancer. But that only applies to women who have reached the age of 96. The NCI reports that chances are only I in 18,608 for a 25-year-old woman.

Don't trust figures on AIDS in Africa. African health officials increase the number of deaths from AIDS, because AIDS cases attract more foreign aid money than does the usual African diseases and also there is no accurate method of counting.

For years in the 1980s, the media reported that there were three million homeless people in the United States. Why? Because a homeless advocate assumed that 1 percent of Americans were homeless. That projected to 2.2 million in 1980. Since the problem was known to be getting worse, he just reported that it reached three million in 1983.

Unfortunately, erroneous numbers in journalism are not always the result of sincere attempts to quantify the relevant data. If you think might not make the effort to count something or they might have an ulterior motive not to report accurately, be skeptical of the numbers.

The culture of journalism is based on the principle of the citation or quote: if someone else said it, or wrote it, it's okay to repeat it.

Almost any editor or writer would scoff at that brash formulation.

But the reader should not believe that journalists, under the crush of daily deadlines, under the pressure of maintaining long-term relationships with sources, and occasionally under the spell of ideology, always meet that standard. Total skepticism is probably impossible. But greater awareness of the sorts of errors journalists tend to make can only help. Watch out for macroeconomic aggregates; try to figure out where huge counts are coming from and how they are being made; try to check the methodology and phrasing of polls; check on the self-interest of the groups that promulgate scary numbers; and remember that scary stories make great copy and should be mistrusted all the more for that reason.

Hypothetical List of Opiate Users (Percentage of Population)

- Afghanistan 2.65
- Argentina 0.13
- Australia 0.20
- China 0.25
- Columbia 0.92
- Egypt 6.44
- Iran 2.26
- Israel 0.61
- Mayanmar 0.89
- Mexico 6.04
- Netherlands 0.31
- Nigeria 6.70
- Russia 1.54
- United States 0.57

Rank Order List of Users

• Nigeria	6.70
• Egypt	6.44
• Mexico	6.04
• Afghanistan	2.65
• Iran	2.26
• Russia	1.54
• Columbia	0.92
• Mayanmar	0.89
• Israel	0.61
• United States	0.57
• Netherlands	0.31
• China	0.25
• Australia	0.20
• Argentina	0.13

Total skepticism is probably impossible.

1. Watch out for macroeconomic aggregates;
2. Try to figure out where huge counts are coming from and how they are being made;
3. Try to check the methodology and phrasing of polls;
4. Check on the self-interest of the groups that promulgate scary numbers;
5. Remember that scary stories make great copy and should be mistrusted all the more for that reason.

Number of Opiate Users by Country

China	0.25	1,343,000,000	33,575,000
United States	0.57	313,847,000	17,889,279
Nigeria	6.70	170,123,000	11,398,241
Russia	1.54	142,517,000	2,194,762
Mexico	6.04	114,975,000	6,944,490
Egypt	6.44	82,688,000	5,325,107
Iran	2.26	78,868,000	1,782,417
Mayanmar	0.89	54,584,000	4,857,976
Columbia	0.92	45,239,000	4,161,988
Argentina	0.13	42,192,000	584,496
Afghanistan	2.65	30,419,000	806,103
Australia	0.20	22,015,000	440,330
Netherlands	0.31	14,730,000	456,630
Israel	0.61	7,590,000	46,299

Numbers for Reporters -- 4
Levels of Measurement

You can't add some numbers together. Some that you can add are not as good as others that you can add. And, it can be difficult to use numbers to compare.

Welcome to the world of measurement. You intuitively know the difference so understanding them should be fairly easy.

There are four levels of measurement that cover anything in universe. And the four levels vary in their strength and/or ability. From least to most powerful they are: (4) nominal, (3) ordinal, (2) equal interval and (1) ratio. You also have to consider whether you should use the mean, median or mode, and even whether the standard deviation might be appropriate.

But first, the measurement levels.

(4) Nominal. This is the easiest and the least powerful. All that nominal does is classify things as either something or something else.

Nominal is used to group things, usually for computer use. For example, since computers can't add, subtract, multiply, or add men and women, they need to by what the computer can do, which is to calculate, or use numbers.

So, what number is a man? Or a woman? They can be just about any number, but they can't be the same number since the purpose is to distinguish between the categories. Usually, you just make them one and two. Either can be the number one and the other can be two. Or a man could be 2.6 and a woman could be 147.89. It does not matter all the computer does is classify all men under the number designated for men and all women under their number.

You can group countries this way. Give countries with complete freedom the number one and all countries the number two. Obviously that is a crude classification. But it suggests a problem for all measurement.

Note. Measurement categories must me mutually exhaustive and also mutually exclusive.

Gender is easy. Everyone has to be male or female. There are just two sexes so gender defined this why is mutually exhaustive. There are no other categories because gender is an anatomical difference definition. So the numbers one and two, or two other numbers will do.

The categories are also mutually exclusive because you have to be one or the two and you cannot be neither or both. For example, listing apples, oranges, pears, and fruit as categories is a no-no because apples, oranges and pears also fall into the fruit category.

The fruit category does suggest how pollsters can get around the problem of too many groupings under the same overall umbrella. During election years, pollsters focus on the party identification of registered votes. They want to classify them so their views can be compared and contrasted. There are Republicans, Democrats, Independents, Greens, Communists, Socialists, Libertarians and so on. In most elections, polls are only interested in the first three groups, so the might classify Republicans as one, Democrats as 2, Independents as three, and all other as four.

Pollsters use religion this way also. The number of religions has been estimated as high as 750 with 34 major religions. Both are too large for pollsters to handle. They take the four or five main religions in their area of interest and classify all other religions under the “other” umbrella. That is okay as long as it is not large a group. Again, whether Catholics are one or six makes no difference.

Gender defined this way is also mutually exclusive.

It is easy to see that nominal measurement doesn’t tell you much. But many important things can be measured at a higher level, such as ordinal measure.

(3) Ordinal. Ordinal measurement does everything that nominal does, that is classify but it also does one more. It ranks things.

For example, you can rank sports teams. You can say Spain is Number One in soccer, followed by Italy and then others. Or in the United States you can rank college football or basketball teams. You can rank things such as beauty pageants. Who wins is Number One, and the first runner-up is Number Two and so on.

But, while ordinal measure can rank, it cannot tell how much more Spain has of what is being measured, say athletic ability that Italy has, or whether Italy has the same amount more than the team ranked Number Three.

You can often see stories containing ordinal measurement on the Internet. If you go to a farmer’s market, you might find several individuals selling the same thing for about the same price. But either by prior experience, information from friends, or testing the freshness, you might rate one of them as the best and another as second, and so on.

At county fairs in the United States, you often have displays of jams and pies, pigs and cows that have been evaluated with the best in each class being awarded a blue ribbon while second and third get a red and yellow ribbon. But what is not given is how much better Number One was than Number Two. They might have been very close or choosing which was best might have been quite easy because Number One was far superior.

Generally you won't see news stories that try to attach any math to the rankings, except in stories about social science research which will be discussed later.

(2) Equal Interval. Its name tells you what else it does. It classifies and it ranks but it also can tell you that the distance between Number One and Number Two is and that it is the same for all other numbers.

The Fahrenheit temperature scale is a good example. If the temperature at 8:00 a.m. was 70 degrees and it is 80 at noon and then 90 at 3:00 p.m., we know that the temperature went up exactly 10 degrees in each time period. It is the same with tests in school, measuring the level of achievement for students.

Although you can add and subtract and determine some basic indexes, such as the mean, median or mode, there is one major thing you cannot do, and that is describe things in ratio terms. Eighty degrees is not twice as warm as 40 degrees and neither is someone with an I.Q. of 180 twice as smart as a person with an I.Q. of 90. They are in accurate because there is no zero on these scales. In colder climates in the winter, temperature often plummet at night, dropping from a daytime high of 20 degrees to a minus 15 or more. Temperature goes right by the zero point. Zero is simply one point on the temperature scale and does not imply the absence of temperature.

(1) Ratio. Ratio measurement does everything that the others do. It can classify, rank, and determine equal distances. But it can do one more thing. It can express numbers in ratio terms because this measures things that have a real zero

The world is full of ratio measurements and so are news stories. Distance for example. Which city is closer to Paris, Rome or Moscow, and by how much? Rome is 694 miles (1,117 kilometers) from Paris while Moscow is 1,544 miles (2,484 kilometers) away.

A more useful term (unless you are driving) is to say that Moscow is more than twice as far from Paris that is Rome.

You can also say that 60 miles (100 kilometers) per hour is twice as fast as is 30 miles per hour.

By far the most important numbers for journalists to write about are those representing money, dollars, euros, pounds, etc. Stock markets around the world express the estimated value of companies in numbers. Governments try to balance their books, estimating that incoming revenue will equal or surpass governmental spending on social services, security, and so forth.

Much of the time, the importance of numbers is how they can be expressed by one number, by an index.

Suppose you want to do a story about which brand of a given product is the best seller in your community? How do you write about that?

Suppose you want to do a story about which brand of a given product is the best seller in your community? How do you write about that?

Suppose national authorities become concerned about the rise of obesity in the country start releasing statistical data on weight by gender, by geography, by age, by ethnicity, by family income level? How do you write about it?

Suppose you are writing a story about the sales of homes in your community? How do you write about that?

Usually you want numerical information to support quotes, opinions, and other qualitative (authoritative) information. You cannot use hundreds of different numbers. You need an index, one number that can be used for each group you are writing about. You probably need one of these three:

Mode, Mean, Median. Each of these measures indicates a different aspect of any group of information.

Mode. The mode is that number, or value, which occurs most often in any group of data.

Mean. This is another word for average. If 11 men had different amounts of money, between $2 and $50, but the total of their money was $220, the average would be the total, $200 divided by the 11 men or, $20 each.

Median. This is the value at halfway point between the lowest and highest amounts of a group. Fox example, suppose the 11 men had the following amounts of money: $2, $5, $8, $9, $10, $12, $15, $18, $20, $40 and $81. The median is $12 because it is the middle value.

These three indexes have different uses, different strengths for different people.

Suppose you are a street vendor in Kabul and you are selling little dishes of ice cream of a hot summer day. Your refrigerated cart has separate containers for six gallons of ice cream. What flavors do you put in the cart?

The smart vendor will know what flavors sell the best. Suppose half of the residents of Kabul like chocolate the best, and 30 percent like Vanilla. Ten percent like strawberry, five percent prefer lemon, and the other five percent list peppermint.

The smart vendor will sell mostly chocolate and then vanilla. Nothing else.

Suppose you television station wants to increase its primetime entertainment audience. The station surveys people to find out which television shows they like best and this is the result:

Crime dramas -----------	35%
Romance stories --------	25
Situation comedies -----	15
Game shows ------------	10
Hollywood movies ------	7
Documentaries ----------	4
Food and recipies-------	3
Religious programs ----	1

Your station show definitely broadcast crime drams and romance stories, and possibly situation comedies and maybe even game shows, but nothing else.

Remember that while words describe, give a general picture, numbers are precise. Indexes are numbers that sum up a group. Indexes give you a better overall picture of a group.

Levels of Measurement

- Nominal – classifies variables into mutually exclusive and mutually exhaustive categories.
- Ordinal --- does what Nominal does and also rank orders variables or attributes.
- Equal Interval --- does what Ordinal does and has equal space between variables and attributes.
- Ratio --- does what Equal Interval does, but has a true zero and therefore can divide and use proportions.

Nominal

Assignment of number merely places things in categories.

1 = Apples

2 = Oranges

9.3 = Bananas

11.7 = All Others

Nominal

NOTE: "Other" category. It ensures a measurement requirement

Mutual Exhaustiveness: All things must be able to be counted.

Religion: 1= Islam
2= Christianity
3 = Buddhism
4 = All other religions

Nothing is left out.

Nominal

Mutual Exclusivity: All things can only belong to one category.

Which do you like best?	Where do you live?
1 = Apples	1 = Iran
2 = Kiwi	2 = Iraq
3 = Bananas	3 = Afghanistan
4 = fruit	4 = Middle East
5 = oranges	5 = Syria
6 = other	6 = other

Ordinal

Of the places you have visited, which do you like the best?

1. San Francisco
2. Santa Cruz
3. Monterey
4. Other

Ordinal

San Francisco and Santa Cruz may be very close, and both much more liked that Monterey or maybe San Francisco is way ahead of Santa Cruz and Monterey which are very close.

Equal Interval

The distance between two things may be the same but you can not say that 80 degrees temperature is twice as warm as 40 degrees because there is no true zero. Zero is another degree and the temperature can be less than zero.

Ratio

The physical world is ratio level measurement because there is a true zero, where zero equals nothing. And measurements can be proportional.

Diana has ------------ $40

Peter has ------------ $20

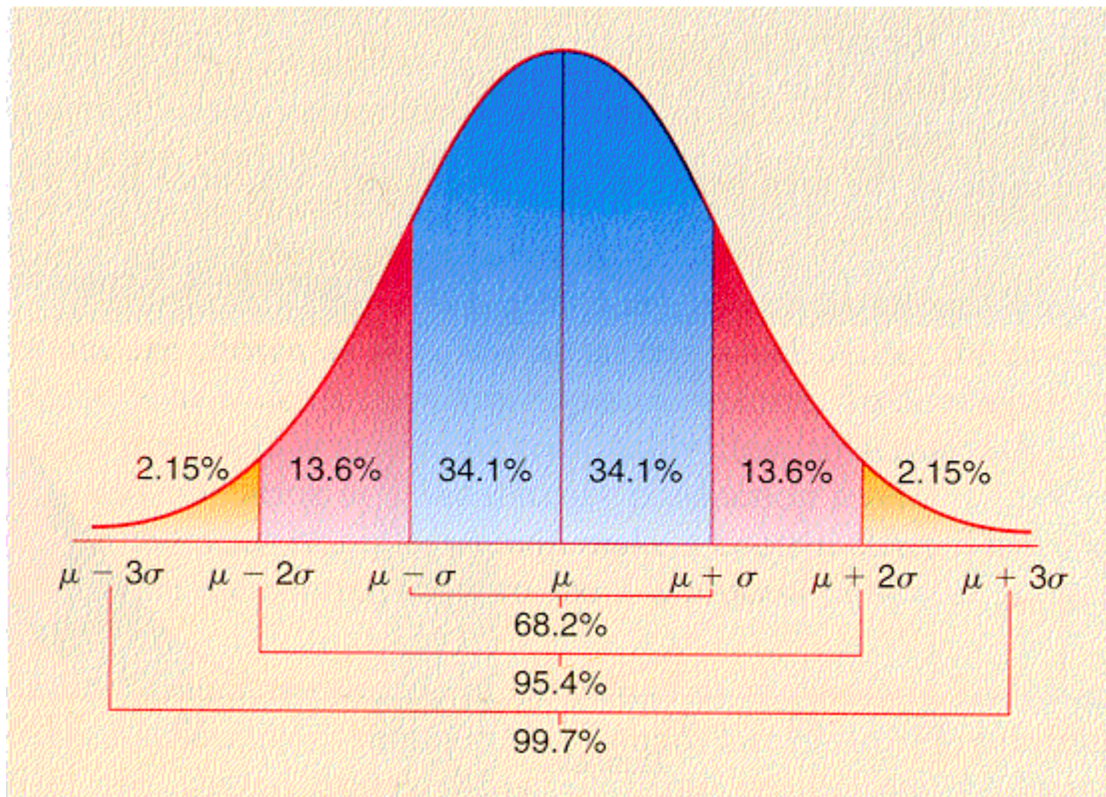
Sandra has ---------- $10

Harry has ------------ $ 5

Bill has ----------------$ 0

Diana has twice as much money as Peter does and four times as much money as Sandra does. Bill has half as much money as Sandra has and on-eighth as much as Diana doe

Normal Curve



Numbers for Reporters
#5 – Percent and Per Capita

Reporters are word people. But a look at any newspaper or many newscasts finds news stories chock full of numbers – and not just simple addition and subtraction not to mention fractions.

Many stories need to compare budgets, rates and other measures with each other and over time, but mostly to examine parts of things. This is where percent and per capita enters the picture.

Percent means 100. Divide any number by 100 and you get one percent. Divide it by 12 and you get 12 percent. And so on. Per cent is anything that can be divided by 100.

Percent is easy – at first. For example, the total amount of money requested by the U.S. military to operate in one year was $793.2 billion. That is a lot of money to be divided up by the various branches of the armed forces. The next question then is how much does each branch get. The Air Force requested 170.6 billion. The others were $244.9 billion by the Army, $29 billion by the Marine Corps, $149.9 billion by the Navy, and there was another $118.7 for defense wide joint activities and $80.1 billion for defense intelligence.

A reporter should first add the individual elements to make sure that they do in fact add up to the total listed. Then present them in table form, such as:

Branch of Service	Amount Requested (in billions)
Army	$ 244.9
Air Force	170.6
Navy	149.9
Joint activities	118.7
Intelligence	80.1
Marine Corps	29.0

That shows in a glance who gets the most but it begs the question: how much more does one branch get than another? This is where percent comes in. For each branch of service, all you have to do is divide the amount of money it receives by the total amount. The Army asked for $244.9 billion. If you divide it by $793.2 you get 0.3150529. To get the percent the Army gets, multiply 0.3150529 by 100. In other words, move the decimal point two numbers to the right and you get 31.50529. Round it off and the Army has asked for 31.5 percent of the total amount. Do the division and multiplication for the other branches and add the percentages to the table. It now looks like this:

Branch of Service	Amount Requested (in billions)	Percent
Army	$ 244.9	31%
Air Force	170.6	21
Navy	149.9	18
Joint activities	118.7	15
Intelligence	80.1	1
Marine Corps	29.0	4

One more thing. Add the per cents. These per cents total 100. But that is because they have been rounded off for better visual clarity. The per cent column would be more accurate if it included one number to the right of the decimal point in order to be more accurate. It would look like this:

Branch of Service	Amount Requested (in billions)	Percent
Army	$ 244.9	30.9%
Air Force	170.6	21.5
Navy	149.9	18.9
Joint activities	118.7	15.0
Intelligence	80.1	10.1
Marine Corps	29.0	3.6

Obviously you would be even more accurate if you include more numbers to the right of the decimal point. The table could look like this:

Branch of Service	Amount Requested (in billions)	Percent
Army	$ 244.9	30.87493%
Air Force	170.6	21.50781
Navy	149.9	18.89813
Joint activities	118.7	14.96469
Intelligence	80.1	10.09833
Marine Corps	29.0	3.65607

When you add those per cents you get 99.99996 per cent. If you multiply the total amount requested, $793.2 by the missing .00006 per cent you get 0.047592.

Remembering that $793.2 divided by 100 equals $7.932 billion which is one percent, the amount of money represented by .047592 per cent can be determined by moving decimal to the right. The zero is the $100 million column, which is one-tenth of one billion in the first column to the left of the decimal point. But there are no numbers in the $100 million column. The 4 is in the $10 million column which means the amount of money represented by .00006 per cent is $4 million plus the $759,200 in the following columns.

Most reporters probably would round off the percentages to even numbers or possibly use one decimal point if they want to be a little more accurate. One might think the table with the percentages carried to five decimal points should be used because they are more accurate. But that is false accuracy because the numbers those percentages could only be accurate if the amount of money requested by each branch of the service was listed down to the last dollar and not rounded off to the nearest billion dollars. Besides which, all of the decimal points detract from the visual clarity of the other tables.

Where most people have problem is with percent change. You must have two of three numbers to calculate percent change. These are (1) the number you start with, (2) the number you end up with, and (3) the difference. If you have two of the numbers, you can determine the third number as well as the percentage.

What is the percent increase from 50 to 75? We can find the difference by subtracting 75 – 50 which equals 25. To determine the per cent increase, divide the difference, 25, by the starting number, 50, and you get .5 which is 50 percent. That is the increase.

Suppose you get a raise at your job. You were making $14 per hour. The new union contact gets you a raise to $15.75. The per cent increase is calculated by taking the difference, $2.75, and dividing it by the starting number, $14, which yields the percent increase, 19.6 percent.

A percentage decrease works the same, except the first number is now the larger number. If we take our first example, a decrease from 75 to 50 we still get 25 as the difference. However, we now divide the difference by 75 which give us 75/25 equals .33 or 33 percent.

Just remember that although the difference is the same in both cases, the per cent is not because the starting number is the smaller number in increases but is the larger number in decreases.

Another way of finding the percentage change is to simply multiply the beginning or first number by the percentage such as 75 times .33 equals 25.

If you multiply a number by one, there is no percent increase. Multiply by 1.3 and you have a 30 per cent increase.

Try some increases that you can find the media every day: (1) how much more money do you get if you receive a 17 per cent in you $650 weekly salary; (2) How many people live in a city that increased 14 percent in 2010 from its 2000 population of 56,000; and (3) how much is this year's state budget if it jumped 30 per cent from last year's $650,000 budget?

Notice again that it takes two numbers to calculate the answer.

Multiply $650 by 1.17, which is your current salary plus the raise, and your new weekly salary is $760.50.

Multiply the 56,000 people by 1.14 and the 2010 population is 63,840.

Multiply the $650,000 budget by 1.30, again 1 for the current budget and .3 for the 30 per cent increase and the new budget is $845,000

We can see where the ability to calculate percent change is helpful for looking at increases such as this hypothetical or the military budget example. Suppose the 2011 budgets for the armed increased to 928.7 billion with the following individual increases:

Branch of Service	Amount Requested (in billions)	
	2010	2011
Army	$ 244.9	$274.5
Air Force	170.6	223.8
Navy	149.9	190.7
Joint activities	118.7	132.3
Intelligence	80.1	91.6
Marine Corps	29.0	30.2

The above table does nothing but list the amounts for the respective years. And although a reasonably intelligent reader could determine the increases, decreases and per cent change to gain more information. Most readers won't do than and will not get the full benefit. Therefore it is up to the reader to include the dollar amount of the change as well as the individual increase or decrease. That table looks like this:

Branch of Service	Amount Requested (in billions)		Difference (in billions)	Per cent Change
	2010	2011		
Army	$ 244.9	$ 274.5	$ 29.6	+12.1%
Air Force	170.6	223.8	53.2	+31.2
Navy	149.9	190.7	40.8	+27.2
Joint activities	118.7	132.3	13.6	+11.5
Intelligence	80.1	77.2	- 2.9	- 3.6
Marine Corps	29.0	30.2	1.2	+ 4.1

It is easy to see why we need per cent changes. They will help tell us not only who are asking for the most money but which branch wants more compared with what they have. In this case, the Air Force is asking for money than are the other services and the Air Force request is also the largest increase. However, although the amount the Air Force is requesting is more than $12 billion above the Navy's request, the per cent difference is only four percent. That difference occurs because of the differences in their original requests. Often a percentage increase for one thing may be larger than a percentage increase of another yet be based on a small actual increase. Consider the GDP of countries.

Russia's GDP skyrocketed 37.8% in 2011 over 2010 compare to the U.S. increase of only 3.17%. Without any other numbers it would appear Russia produced more than ten times as much in economic value. However, a glance at their respective GDPs in the table below tells a different story. Russia's actual growth was less than twice that of the United States

	GDP (in billions)	
	United States	Russia
2010	15.060	2.380
2010	14.582	1.479
Difference	478	901
Per cent Increase	3.1	37.8

Another interesting aspect of freedom is illustrated by the following sentence: Philanthropy in the United States decreased 350 percent last year. Think about it for a moment. Once something has decreased 100 percent, it is all gone. Although an increase can be by percentages greater than 100 percent, decreases cannot.

Percentages also often need to be tied to two other elements in order to report correct information. They are needed when it comes to comparing two items of unequal size and between time periods. To demonstrate their use, the following example was created by Robert Niles, an online publisher. Suppose you read the following two paragraphs, the lead and backup, in a news story:

> Springfield's murder rate increased by 72 percent in the past 10 years compared to only a 19 per cent increase in nearby Capital City.
>
> A total of 50 people were killed in Springfield in 2010, 21 more than the 29 murdered in 2000 while Capital City's 42 killed in 2000 only climbed by eight in 2000.

Suppose Springfield population rose from 450,000 in 2000 to 800,000 in 2010 but Capital City's increase was from 550,000 to 600,000? You probably expect more people to get killed in a larger city. What is missing is per capita, which adjusts for population differences, the number killed in each town as a percentage of the population that ear.

To find that rate, divide the number of murders by the total population of the city. To keep from using such a tiny decimal, statisticians usually multiply the rate by 100,000 and give the result as the number of murders per 100,000 people. Let's do the math for Springfield and Capital City:

Springfield (2010)-------50/800,000 = .0000625 times 100,000 = 6.25%
Capital City (2010)------50/600,000 = .0000833 times 100,000 = 8.33%

That sort of looks like the lead is backward and that Capital City has the higher murder rate. But that was for just the one year. The story is focusing on the increase over time. So we need to first calculate the murder rate for the two cities in 200 which is:

Springfield (2000)-------29/450,000 = .0000644 times 100,000 = 6.44
Capital City (2000)------42/550,000 = .0000763 times 100,000 = 7.63

Obviously Capital City has the higher murder rate for both individual years. To get the comparison overtime we need to follow the same procedure but with the murder rates and with ought 100,000 since it has already been included in the individual calculations. Thus:

Springfield: (6.25 – 6.43)/6.43 = -.18/6.3 = -028% (a decrease over time)
Capital City: (8.33 - 7.64)/7.64 = .69/7.43 = .090 (an increase)

The first paragraph of the story, the lead, needs to be rewritten such as:

> Springfield's murder rate decreased by almost three percent from 2000 to 2010 while nearby Capital City's per capita murder rate increased by more than nine percent in the same time period.

There is a problem in writing the second paragraph in the normal inverted pyramid structural manner, that is in parallel construction supporting both sets of figures in one paragraph, such as:

> Although the number of murders in Springfield increased by 21, from 29 to 50 and Capital City increased only by eight, from 42 to 50, the population of Springfield rose by 200,000 to 800,000 but Capital City only increased by 50,000 to its current 600,000.

There are 10 numbers in that paragraph, counting the years. It is probably better to do the figures for one city at a time although some readers might perceive a bias, favoring Springfield for its decrease it you lead with it or slamming Capital City if you do it first.

Here are some tips gleaned from professional journalists:

1. Remember to take away two decimal places (or multiply by 100) when figuring percents.
2. Try to find an easy way for your readers to picture a rate. Simple fractions give the image of a pie and are easy to digest.
   Correct: Agriculture consumed 35 percent of the budget
   Better: About a third of the budget went to agriculture
   or One of three tax dollars in the state went to agriculture.
3. For small rates, use a larger base than 100 or percent. This is most common in medical stories, like the number of infant deaths per 100,000 instead of per 100.
4. Keep your bases similar; compare similar numbers.
   Correct but meaningless: "Eleven percent of blacks voted for Bush, but 10 percent of Clinton's vote came from blacks."
   This mixes bases: the number of black voters then the number of Clinton voters.
   Better: "Eleven percent of blacks voted for Bush and 60 percent

voted for Clinton." The sentence keeps the base as the number of black voters.

Or: "Two percent of Bush's votes and 10 percent of Clinton's votes came from blacks." It has two bases (the number of votes for each candidate) but keeps the comparison similar.

5. With percent changes greater than 100, revert to ratios, not percent changes. This is expressed as "times as many" or "times as much," not "times more".

   Here are three ways to compare 400 arrests for drunken driving this year compared with 100 last year:

   Wrong: "Four times more" is not only confusing but wrong. Three times more or four times as many. (Compute it as a percent change: (400-100) / 100 = 300/100 = 3.00 or a 300 percent increase.)

   Better: "Four times as many people...

When you subtract numbers expressed as percentages, the result is a percentage point difference not a percent change.

Wrong: Joblessness has grown by 1 percent since the 1960s, from about 4 percent to a little over 5 percent.

Both of these are right; choose your weapon carefully:

"Joblessness grown by one-quarter since the heyday of the 1960s, when unemployment stood at just 4 percent."

or

"Joblessness has grown by only 1 percentage point since the 1960s."

6. Compound, don't multiply, when you have to project changes. Remember to compound, not divide, when you have to figure an annual rate.

   Wrong: If a budget grows 4 percent a year, it will grow 40 percent in 10 years (4 x 10)

   Right: The budget will grow about 48 percent.

   The formula is confusing: 1.04, or 104 percent of each year's value, to the 10th power minus 1 times 100. Or ((1.04^10) - 1) x 100

   (^ stands for "to the power of", and is what you'd use in a spreadsheet)

First Set of Budget Numbers

Branch of Service	Amount Requested (in billions)
Army	$ 244.9
Air Force	170.6
Navy	149.9
Joint activities	118.7
Intelligence	80.1
Marine Corps	29.0

Branch of Service	Amount Requested (in billions)	Percent
Army	$ 244.9	31%
Air Force	170.6	21
Navy	149.9	18
Joint activities	118.7	15
Intelligence	80.1	1
Marine Corps	29.0	4

Branch of Service	Amount Requested (in billions)	Percent
Army	$ 244.9	30.9%
Air Force	170.6	21.5
Navy	149.9	18.9
Joint activities	118.7	15.0
Intelligence	80.1	10.1
Marine Corps	29.0	3.6

Branch of Service	Amount Requested (in billions)	Percent
Army	$ 244.9	30.87493%
Air Force	170.6	21.50781
Navy	149.9	18.89813
Joint activities	118.7	14.96469
Intelligence	80.1	10.09833
Marine Corps	29.0	3.65607

To Calculate Percent Change

You must have two of three numbers to calculate percent change. These are:

(1) the number you start with,

(2) the number you end up with, and

(3) the difference.

If you have two of the numbers, you can determine the third number as well as the percentage.

What is the Percent Increase From 50 to 75?

We can find the difference by subtracting 75 – 50 which equals 25.

To determine the per cent increase, divide the difference, 25, by the starting number, 50, and you get .5 which is 50 percent.

That is the percent increase.

Another Increase Example

Suppose you get a raise at your job. You were making $14 per hour. The new union contact gets you a raise to $15.75.

The percent increase is calculated by taking the difference, $2.75, and dividing it by the starting number, $14, which yields the percent increase, 19.6 percent.

Calculating a Percentage Decrease

A percentage decrease works the same, except the first number is now the larger number.

If we take our first example, a decrease from 75 to 50 we still get 25 as the difference.

However, we now divide the difference by 75 which give us 75/25 equals .33 or 33 percent.

Another Way

Another way of finding the percentage change is to simply multiply the beginning or first number by the percentage such as 75 times .33 equals 25.

If you multiply a number by one, there is no percent increase. Multiply by 1.3 and you have a 30 per cent increase.

Some Examples You Might Find in Media

(1) How much more money do you get if you receive a 17 per cent in you $650 weekly salary?

(2) How many people live in a city that increased 14 percent in 2010 from its 2000 population of 56,000; and

(3) How much is this year's state budget if it jumped 30 per cent from last year's $650,000 budget?

Again, It Takes Two Numbers to Calculate the Answer

(1) Multiply $650 by 1.17, which is your current salary plus the raise, and your new weekly salary is $760.50.

(2) Multiply the 56,000 people by 1.14 and the 2010 population is 63,840.

(3) Multiply the $650,000 budget by 1.30, again 1 for the current budget and .3 for the 30 per cent increase and the new budget is $845,000

Comparing Budget Changes Over Time

Branch of Service	Amount Requested (in billions)	
	2010	2011
Army	$ 244.9	$274.5
Air Force	170.6	223.8
Navy	149.9	190.7
Joint activities	118.7	132.3
Intelligence	80.1	91.6
Marine Corps	29.0	30.2

Dollar and Percent Comparisons

Branch of Service	Amount Requested (in billions) 2010	Amount Requested (in billions) 2011	Difference (in billions)	Per cent Change
Army	$ 244.9	$ 274.5	$ 29.6	+12.1%
Air Force	170.6	223.8	53.2	+31.2
Navy	149.9	190.7	40.8	+27.2
Joint activities	118.7	132.3	13.6	+11.5
Intelligence	80.1	77.2	- 2.9	- 3.6
Marine Corps	29.0	30.2	1.2	+ 4.1

Percent Increases Not Always Enough

GDP
(in billions)

	United States	Russia
2010	$15.060	$ 2.380
2010	14.582	1.479
Difference	$ 478.0	$ 901.0
Per cent Increase	3.1%	37.8%

Deceptive Comparisons

Springfield's murder rate increased by 72% in the past 10 years compared to only a 19% increase in nearby Capital City.

A total of 50 people were killed in Springfield in 2010, 21 more than the 29 murdered in 2000 while Capital City's 42 killed in 2000 only climbed by eight in 2000.

Need for Per Capita Comparisons

Suppose Springfield's population rose from 450,000 in 2000 to 800,000 in 2010 but Capital City's increase was from 550,000 to 600,000?

You probably expect more people to get killed in a larger city. What is missing is per capita, which adjusts for population differences, the number killed in each town as a percentage of the population that ear.

To find that rate, divide the number of murders by the total population of the city. To keep from using such a tiny decimal, statisticians usually multiply the rate by 100,000 and give the result as the number of murders per 100,000 people. Let's do the math for Springfield and Capital City:

Springfield (2010)-----50/800,000 = .0000625 times 100,000 = 6.25%

Capital City (2010)----42/600,000 = .0000833 times 100,000 = 8.33%

Need to Compare Over Time

That looks like the lead is backward and that Capital City has the higher murder rate. But that was for just the one year. The story is focusing on the increase over time. So we need to calculate the murder rate for the two cities in 2000 which is:

Springfield (2000)----29/450,000 = .0000644 times 100,000 = 6.44

Capital City (2000)---42/550,000 = .0000763 times 100,000 = 7.63

Capital City has the higher murder rate for both years. To get the comparison over time we need to follow the same procedure but with the murder rates and without 100,000 since it has already been included in the individual calculations. Thus:

Springfield: (6.25 – 6.43)/6.43 = -.18/6.3 =

- 028% (a decrease over time)

Capital City: (8.33 - 7.64)/7.64 = .69/7.43 =

.090% (an increase)

NAME____

Numbers for Reporters Quiz #2

Please indicate which level of measurement at right is suggested by the example at right.

1.________The number of miles between Paris and Rome.

2.________The main monotheistic religions are Islam, Christianity, and Judaism.

3.________This morning's temperature in either Fahrenheit or Celsius.

4.________A person's weight

A. Ordinal
B. Nominal
C. Ratio
D . Equal Interval

5. Which answer below is the proper power ranking of the four levels of measurement from lowest to highest in measurement power?

   A. ratio, equal interval, ordinal, nominal
   B. nominal ordinal, ratio, equal interval
   C. equal interval, ratio, nomnal, equal interval
   D. ordinal, ratio, nominal, equal interval
   E. nominal, ordinal, equal interval, ratio

TRUE FALSE 6. Percent is anything that can be divided by 100.

TRUE FALSE 7. It is often easier to compare two or more things using percent than using the raw numbers.

TRUE FALSE 8. To determine the increase from one number to another, divide the first number by the difference between the numbers.

TRUE FALSE 9. Per capita can not be used to compare two examples if one of more than twice as large as the other.

10. Which is the correct answer to this statement: If a budget is projected to grow four percent per year, in 10 years it will grow about ___?

A. 50% B. 48% C.42% D. 40% E. 20%

Quiz #2 Answers

1. C
2. B
3. D
4. C
5. E
6. TRUE
7. TRUE
8. TRUE
9. FALSE
10. B

Numbers for Reporters – 6
Polls and Statistics

"There are three kinds of lies: lies, damned lies, and statistics."
Benjamin Disraeli, British politician, 1804-1881

"Statistics: The only science that enables different experts using the same figures to draw different conclusions."
Evan Esar, American humorist, 1899-1999

Some numbers are not precise. They are estimates. The numbers come from surveys or polls, which gather information from, and about, particular population groups.

It is virtually impossible to ask every single person in large, often scattered groups about their lifestyles and opinions. Not everyone can be reached. It would be too time-consuming. It would be far too expensive.

Also, it is unnecessary.

Instead, social scientists, governments, pollsters and political candidates take samples, a small part of that population. How the sample was selected, as well as the numbers, and what they represent, determine the value and the accuracy of the numbers generated.

The first question for the reporter incorporating these numbers into a news story is whether the sample is representative of the particular population. To ensure a qualified representation, individuals conducting surveys use the EPSEM structure, equal probability of selection method. Its foundation is that if every member of the population has the same chance of being randomly chosen, the sample results can be generalized back to the parent population.

By contrast, the ABC-TV news magazine Nightline as well as many local television stations and newspapers reported information, as news, from a volunteer sample, people who called in to give their views. An example of this is having viewers or readers call one phone number if they favored a particular issue or candidate, and call another number if they opposed, such as gun control or abortion. The results, say 8,428 in favor and 9,091 against would then be broadcast.

The numbers are worthless. People who call in or otherwise volunteer are a biased, or non-representative sample and the results cannot be generalized to the over-all population.

In other words, sampling selects a random group of people from a particular population group, whose answers would be the same obtained if every member of the population were interviewed.

How big a sample do you need? Is a sample of 100 enough? 200? 400? Or more? Do you need bigger samples if you have bigger populations, a college, the city of San Francisco, a country?

Sample size depends on a number of things, money, time, purpose . . . and level of accuracy needed, the margin of error. The more people that are sampled, the more confident pollsters can be that the "true" percentage is close to the observed percentage. The margin of error is a measure of how close the results are likely to be.

Since the numbers obtained in a survey are an estimate, we need to know how far off our answer might be, that is how much more or less might the population answer be. Or what is the margin of error?

Suppose that 60 per cent of a sample of 100 put ketchup on their eggs at breakfast and a ketchup company wanted to know the margin of error. Divide the number one by the square root of the sample size and both add and subtract it from 60 percent. One divided by 10 is .10 or ten per cent. Therefore the answer to how many people in the population put ketchup on their eggs ranges 50 per cent to 70 per cent. For a sample size of 400, you divide one by 20 and the answer is .05 or five percent cutting the margin of for the population to between 55 and 65 per cent.

But some things, like elections, the difference of whether issues or candidates are favored or opposed can be much closer, say 53 percent for and 47 per cent against. You can see that if the sample size is 400, adding and subtracting results in saying that between 48 per cent and 58 percent favor or that between 42 per cent and 52 per cent oppose.

That is like saying Candidate A will win or lose? Or that the voters favor or oppose the issue. But you knew that before the survey.

You need a bigger survey. Because of the square root factor in determining the margin of error, or size of the spread, the sample size has to be quadrupled in order to cut the margin of error in half. If the sample size is 1,600, its square is 40 and one divided by 40 is .025 which is 2.5 percent. Now the answer to whether Candidate A with 51 per cent will win is yes, because 2.5 per cent added and subtracted to 53 per cent is within 50.5 per cent and 55.5 per cent.

If you guessed that you need larger samples for larger populations, you guessed wrong. One divided by the square of the size of the sample provides the margin of error, regardless of whether the sample was drawn randomly from Naples, Italy, or from all of Italy.

The major advantage of surveys is that a lot of information can be obtained from a variety of people at a reasonable cost. However, they are getting harder to conduct on a much-surveyed United States, sometimes the wrong respondents are included and bad questions can produce biased results. Surveys can cost between $30 and $50 per respondent.

There are about five types of surveys that you should know the strengths and weaknesses of: (1) personal (face-to-face); (2) telephone; (3) mail; (4) Internet; and (5) mall.

The advantages of the personal interview include: flexibility, interviewer observations and rapport of an identifiable respondent and higher response rate. The drawbacks include that it is the most costly, takes more time than other survey forms, can have interviewer biasing responses and requires interviewer training. It is also the least safe.

The advantages of telephone surveys are that they are the quickest to complete, cheaper than personal interviews and the interviewer can clarify any misunderstanding.

The response rate is higher than all but mail surveys. The disadvantages of telephone surveys include the suspicion you are a telemarketer, the inability to show photos, charts, and graphs plus the difficulty of reaching cell phone users who may differ significantly from users of home or business telephones.

Mail surveys have the advantage of being the cheapest method, eliminating any interviewer bias, are safe, provide the respondent with anonymity, and can cover a large geographic area. However, mail surveys have the lowest response rate, take a longer time, must be completely self-explanatory, the researcher never knows for sure who answered the questions, and if only interested individuals respond, it may be biased concerning the content.

The newest type of survey is the Internet survey. It is generally inexpensive, easy to conduct even with visuals and it can be done quickly. However, Internet surveys are more difficult to generalize to, or be like, the general population, provide no control over the data-gathering process and there is no assurance that the requested respondent is the one who completes the interview.

The mall survey is, as its name implies, conducted generally in shopping malls where researchers roam the corridors or sometimes stake out positions by particular stores in their search for selected individuals. This survey is quick, relatively cheap, but cannot be used to reflect the overall population.

But then, the good mall survey is not interested in reflecting the people in general. In a good mall survey, the researcher is looking for people with the same interest, such as wearing running or jogging shoes and therefore goes where he can find them, going in and out of shoe and sporting goods stores in shopping malls.

The 1936 presidential polls illustrate some of the issues involved in sampling: (1) size does not improve accuracy is sample is biased

1936 Literary Digest & Gallup Polls

The Digest sampled 10 million. 23% said:

Alf Landon -----------------1,293,669 (57.1%)

Franklin Roosevelt-------- 972,897 (42.9%)

Gallup sampled 50,000

and found:	Election Results
Landon---------44.3%	36.5%
Roosevelt------55.7%	60.8%

Reputable outlets, like Gallup, have a track record of results. In its final pre-election polls in the 19 presidential contests since 1936, Gallup has incorrectly predicted just three: 1948, 1976, and 2004, in which the final Gallup survey showed a tie between George W. Bush, the winner, and the Democratic nominee John F. Kerry. It's worth noting that Gallup's pre-election poll in 2000 gave Mr. Bush the advantage over Democrat Al Gore, and while Bush ultimately won the election, Gore won the popular vote.

However, large samples are needed because effects are usually small and there are confounders that change results if not considered. For example, cancer is increasing in the United States but because the increasing number of older people is factored in, the chance of a specific person getting cancer is declining.

You have to watch for spurious clustering

Imagine a chess or checkerboard occupying the bottom of a large cardboard box. Toss in exactly 64 grains of rice. The grains will bounce around and finally come to rest on the 64 squares of the game board. The average incidence is one grain per square. But you're not likely to ever see that in your lifetime. Some squares will have many grains – all by chance. Likewise, some communities will report much larger-than-average incidences of certain diseases, all by chance.

With polls, keep an eye on demographics

When it comes to polling, yes, we can take 850 in an imperfectly-drawn New Hampshire sample and split it 6 ways (young-old, rich-poor, male-female, minoritywhite...) and insinuate the overall statistical power of the overall sample, which isn't that great in the first place, while never once mentioning that New Hampshire's demographics and ground truths have changed a lot since the last hugely contested primary there in 2000, and that younger voters, who have only cell phones, are hard to find – and thus hard to poll. And why? Because bringing up any of this screws up the story!

How Accurate Are Good Samples?

Sample accuracy depends on two numbers: First, we want to be 95% sure, (the percent scientists demand). Second, we need to know the margin of error, that is how much different the sample figures are likely to be from the population.

The margin of error (at 95%) is one divided by the square root of the size of the sample.

"There are three kinds of lies: lies, damned lies, and statistics."

Benjamin Disraeli,

British politician, 1804-1881

"Statistics: The only science that enables different experts using the same figures to draw different conclusions."

Evan Esar,

American humorist, 1899-1999

Margin of Error

The margin of error is a measure of how close the results of a sample are likely to be to the overall population.

Since the numbers obtained in a survey are an estimate, we need to know how far off our answer might be, that is how much more or less might the population answer be.   Or what is the margin of error? How much more of less?

Suppose that 60 per cent of a sample of 100 put ketchup on their eggs at breakfast and a ketchup company wanted to know the margin of error.

Margin of Error

Divide the number one by the square root of the sample size. Then both add and subtract it from 60 percent.

1/square root of 100

= 1/ 10

= .1

Multiply by 100 to get the percent.

.1 times 100 = 10

The answer to how many people in the overall population put ketchup on their eggs is:

60% +/- 10% or between 50% and 70%.

Margin of Error

A 10 percent margin of error is usually too much. To get a smaller margin of error, you need to increase the size of the sample.

Suppose you want only a 5% margin of error. You need to quadruple the sample size to 400.

The square root of 400 is 20.
1 /20 = .05
Multiply .05 by 100 = 5

The margin of error is 60% +/- 5% or between 55% and 65%.

Margin of Error

Many results, such as election polls need an even smaller margin of error.

Suppose your sample size is 1,000.

What is the margin of error?

It is 1 /square root of 1,200.

It is 1/34.641016 = .0288675

Multiply by 100 = 2.88675

The true answer is 60% +/- 2.9% or 57% to 63%.

Sample Size for Different Size Places

To get a margin of error of 2.9%, how big a sample do you need for:

San Jose State University, 30,000 students?

San Jose, 1.1 million people?

California, 37 million people?

Margin of Error

The answer is

1,000 for San Jose State

1,000 for San Jose

1,000 for California

1/square root of 1,000

Famous Survey

1936 Literary Digest & Gallup Polls

The Digest sampled 10 million. 23% said:

Alf Landon ------------------1,293,669 (57.1%)

Franklin Roosevelt-------- 972,897 (42.9%)

Gallup sampled 50,000

and found:	Election Results
Landon---------44.3%	36.5%
Roosevelt------55.7%	60.8%