

2017 Winter Leadership Conference

How Experts Lie with Statistics

Franklind D. Lea, Moderator

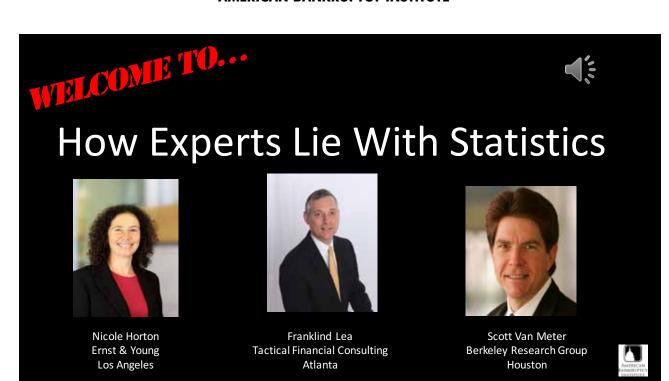
Tactical Financial Consulting, LLC; Alpharetta, Ga.

Nicole R. Horton

EY; Los Angeles

K. Scott Van Meter

Berkeley Research Group; Houston



We live in a world that is built on promises constructed by liars.

AMBRICAN

Lying is done with words and also with silence.

Adrienne Rich



History is a set of lies agreed upon.

Napoleon Bonaparte



I love listening to lies when I know the truth.



The best revenge on a liar is to convince him that you believe what he said.



I'm not upset that you lied to me, I'm upset that from now on I can't believe you.

Friedrich Nietzsche



Being true to yourself is better than being a liar just to impress everyone.



Lying Is an elementary means of self-defense.

Susan Sontag



If you tell the truth, you don't have to remember anything.

Mark Twain





If the words don't add up, its usually because the truth wasn't in the equation.

Gordon Bethune



A single lie discovered is enough to create contagious doubt over every other truth expressed.

Unknown



Sometimes the truth will hurt and a lie will set you free; just as a lie will hurt and the truth will set you free.



Drinking? Why, no, Officer.



The check is in the mail.

It's not the money, it's the principle of the thing.



Half the lies they tell about me aren't true.

Yogi Berra





If you don't want to slip up tomorrow, speak the truth today.

Once a liar always a liar.

Unknown

Bruce Lee



I lie to myself all the time. But I never believe me.

S. E. Hinton



The great masses of the people will more easily fall victim to a big lie than to a small one.

Adolf Hitler



When a man is penalized for honesty he learns to lie.

Criss Jami



The longer the explanation the bigger the lie.

Ziad K. Abdelnour



The most dangerous lies are the ones you tell yourself.

Richard Bach



Never tell a lie – except for practice.

Mark Twain



Any fool can tell the truth, but it requires a man of some sense to know how to lie well.

Samuel Butler



Don't steal, don't lie and don't cheat. The government hates competition.

Hussein Nishah



A lie has speed, but truth has endurance.

Edgar J. Mohn



I've always tried to tell the American People the truth. My mouth just doesn't always cooperate.

Unnamed Political Candidate

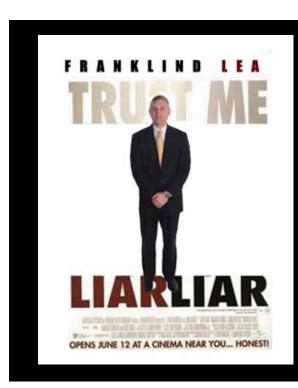


My parents called me a liar...I looked them in the eyes and said "Santa, Easter Bunny, Tooth Fairy" and walked away like a boss.

Unknown

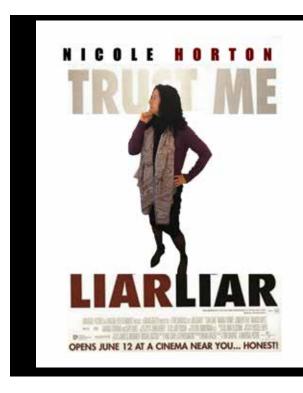


Lies, Damn Lies and Statistics



Franklind Lea

- President, Tactical Financial Consulting
- Financial and Real Estate Expert
- 25+ Years of Experience
- B.S. Management
- M.B.A., Emphasis in Finance
- Master of Real Estate
- Three facts about me:
 - I play(ed) clarinet
 - I compete in equestrian events
 - I had a cat named Boots



Nicole Horton

- Executive Director in EY's Transaction Advisory Services practice
- More than 17 years of experience
- Heavily focused on the financial services sector
- Three facts about me:
 - I play(ed) bassoon
 - I compete in equestrian events
 - I have a cat named Chewbacca



Scott Van Meter

- Managing Director, Berkeley Research Group
- Expertise in valuation and economic damage analysis
- 30+ Years of Experience
- JD, CPA, CIRA
- Three facts about me:
 - I play(ed) the Tuba
 - I compete in equestrian events
 - I have a cat named "Chu Chu"

WHO JUST LIED TO YOU? / CAN YOU IMAGINE A WORLD WHERE YOU COULDN'T LIE







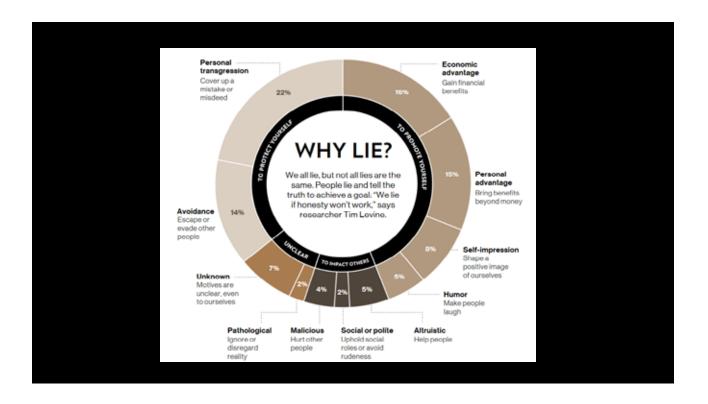






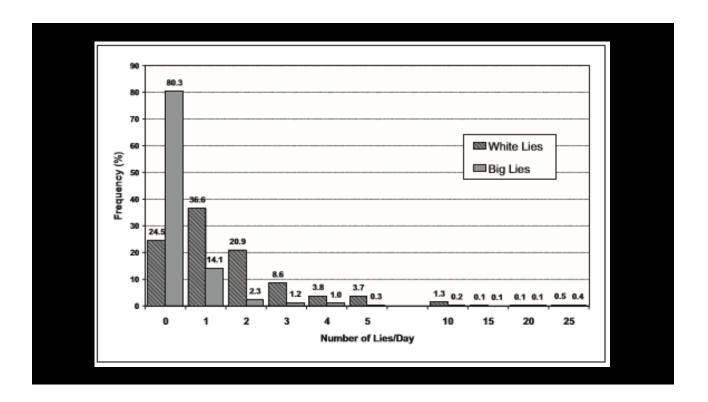


SOCIOLOGISTS TELL US EVERYONE NEEDS TO LIE



What's The Most Common Lie Told?

- A. I'm from the Government and I'm here to help.
- B. It's not you, it's me.
- C. Trust me, I'm a Lawyer.
- D. The check is in the mail.
- E. I'm Fine.

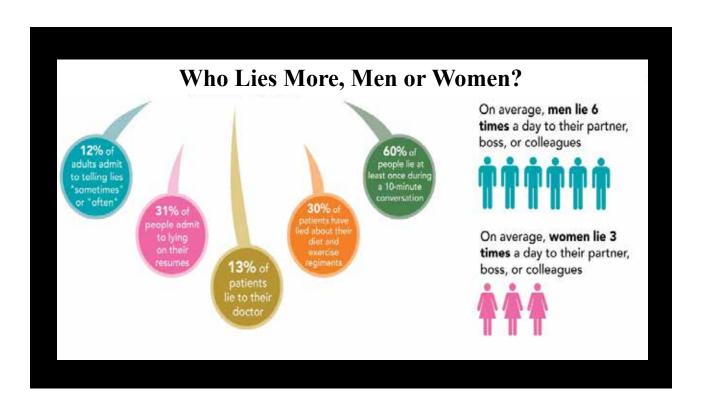


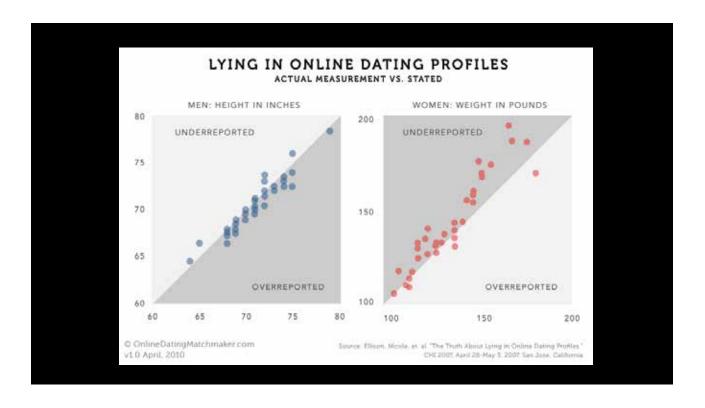


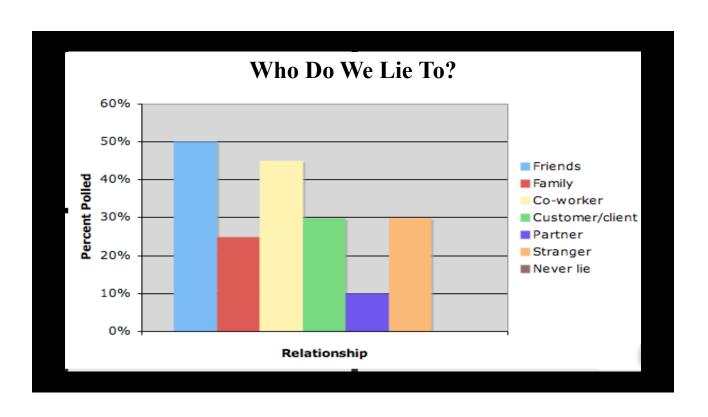
How Many Lies Do You Tell During Your Lifetime?

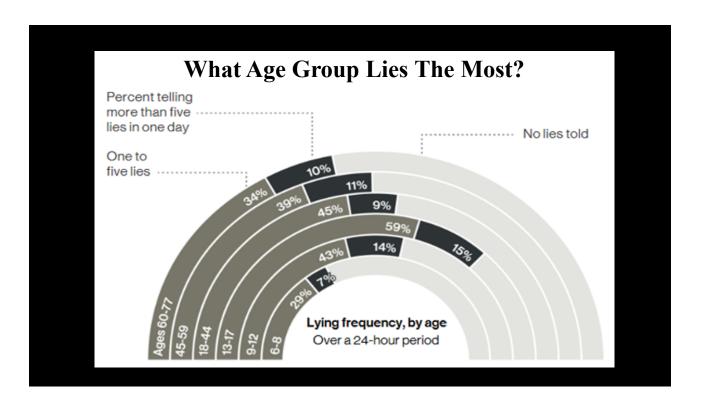
The average person tells 4 lies a day or 1460 a year; a total of 87,600 by the age of 60.

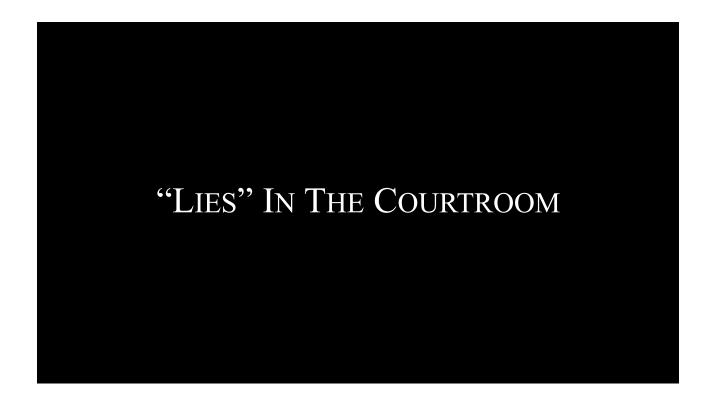
The twelve people now standing will tell over 1,000,000 lies during their lifetimes!

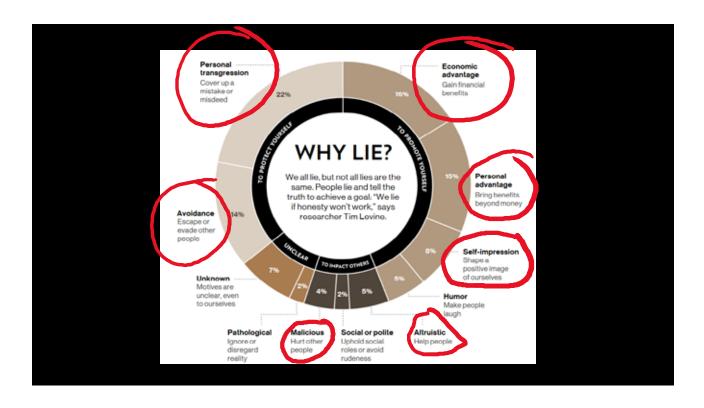


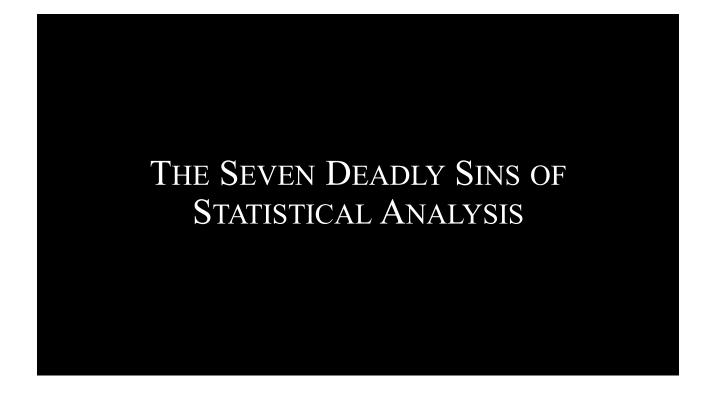












The Seven Deadly Sins of Statistical Analysis (Basis in Academic Literature)

Steven Ross, Ph.D.

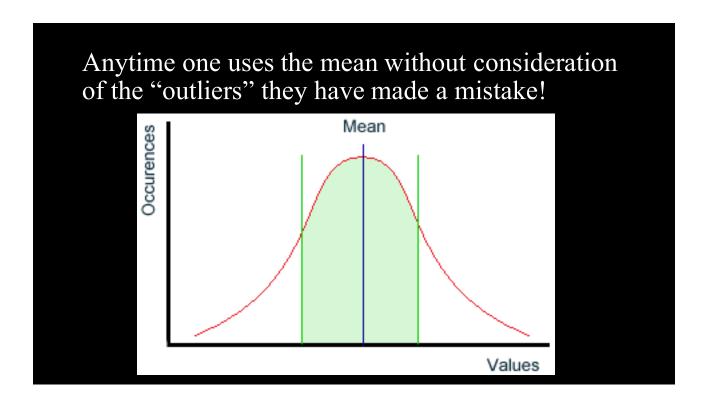
- 1. Non-Representative Sample
- 2. <u>Mistaking Statistical Association</u> <u>for the Cause</u>
- 3. Poisoned Control
- 4. Data Enhancement
- 5. Absoluteness
- 6. Partiality
- 7. A Bad Measuring Stick

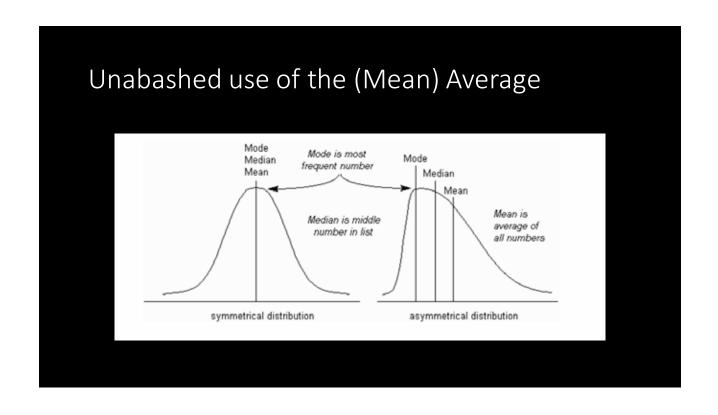
Louis and Chapman

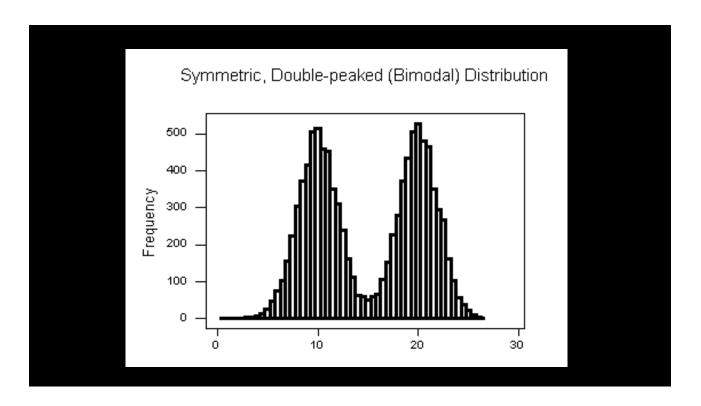
- Assuming Small Differences are Meaningful
- 2. Equating Statistical Significance with Real World Significance
- 3. Neglecting to Look at Extremes
- 4. Trusting Coincidence
- 5. Getting Causation Backwards
- 6. Forgetting to Consider Outside Causes
- 7. <u>Deceptive Graphs</u>

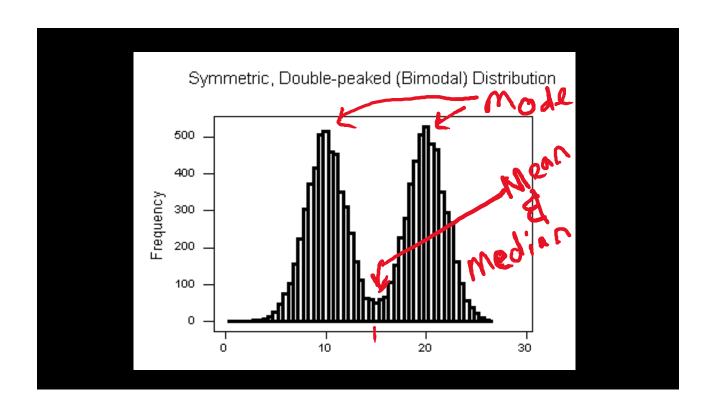
Three Types of Averages

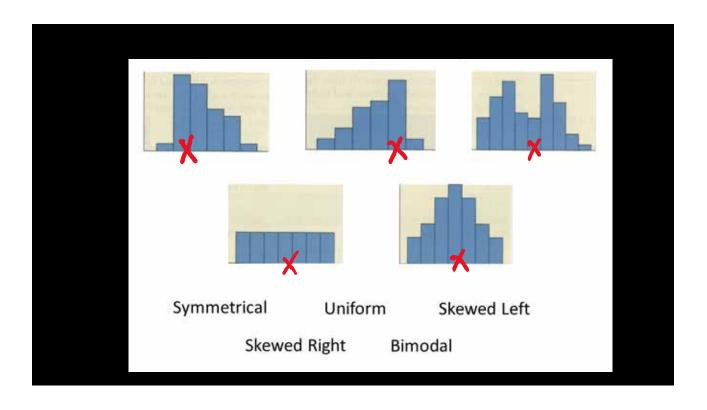
	Data Set #1	Data Set #2	Data Set #3	Data Set #4
	1	0	1	0
	3	2	1	0
	3	3	1	0
	3	5	2	0
	<u>5</u>	<u>5</u>	<u>10</u>	<u>15</u>
Total	15	15	15	15
Mean	15/5=3	15/5=3	15/5=3	15/5=3
Mode	3	5	1	0
Median	3	3	1	0

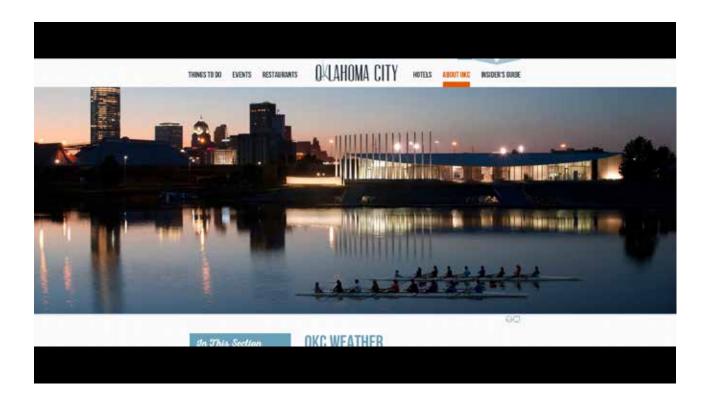


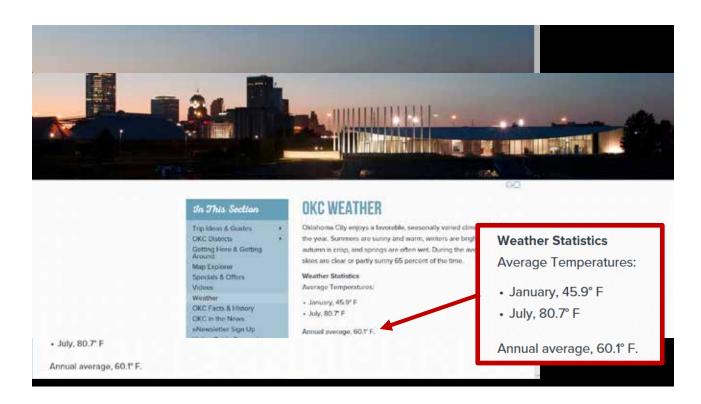












Climate da	ta for Ok	ahoma	City (W	III Roge	rs World	Airport)	, 1981-	2010 ave	erages				[hi
Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Yea
Record high °F (°C)	83 (28)	92 (33)	97 (36)	100 (38)	104 (40)	107 (42)	110 (43)	113 (45)	108 (42)	97 (36)	87 (31)	86 (30)	113 (45
Average high °F (°C)	49.0 (9.4)	53.9 (12.2)	62.7 (17.1)	71.6 (22)	79.4 (26.3)	87.4 (30.8)	99:1 (33:9)	92.8 (33.8)	84.8 (28.9)	72.8 (22.7)	60.8 (16)	49.8 (9.9)	71.4
Daily mean °F (°C)	38.4 (3.6)	42.9 (6.1)	51.3 (10.7)	60.1 (15.6)	69.0 (20.6)	77.1 (25.1)	82 1 (27 B)	81.6 (27.6)	73.1 (22.8)	61.7 (18.5)	49.9 (9.9)	39.7	60.0
Average low "F ("C)	27.7 (-2.4)	31.9	40.0	48.6 (9.2)	58.6 (14.8)	66.8 (19.3)	71.0 (21.7)	70.3 (21.3)	62.2 (16.8)	50.6 (10.3)	39.0 (3.9)	29.6 (-1.3)	49 (9.8
Record low °F (°C)	-11 (-24)	-17 (-27)	(-17)	20 (-7)	32 (0)	46 (8)	53 (12)	49 (9)	35 (2)	16 (-9)	9 (-13)	-8 (-22)	-1 (-2
Precipitation inches (mm)	1,39 (35.3)	1.61 (40.9)	3.06 (77.7)	3.06 (77.7)	4.54 (117.9)	4.95 (125.7)	2.97 (75.4)	3.31 (84.1)	4,07 (103.4)	3.73 (94.7)	2.01 (51.1)	1.88 (47.8)	36.6 (931
Snowfall inches (cm)	2.8 (7.1)	2.0 (5.1)	.9 (2.3)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	.4 (1)	2.3 (5.8)	8.5
Avg. precipitation days (≥ 0.01 in)	5.1	5.9	7.6	7.7	9.8	9.2	5.8	6.7	7.2	8.0	6.0	5.8	84.6
Avg. snowy days (≥ 0.1 in)	1.7	1.3	.6	0	0	0	0	0	0	0	.3	1.7	5.7
Mean monthly sunshine hours	201.5	192.1	244.9	270.0	294.5	327.0	356.5	328.6	264.0	244.9	186.0	179.8	3,089

Oklahoma City Climate Data 1981 -2010

Average Daily Mean Temperature

60.6 Degrees

113 Degrees - Record High

<u>-17 Degrees</u> – Record Low

130 Degrees Swing!

Sampling Errors

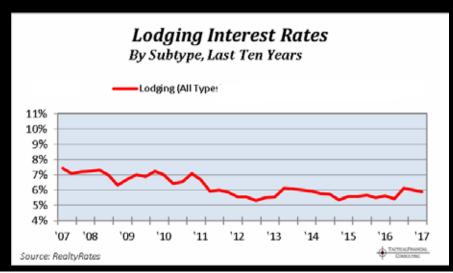
- What is sampling?
 - In instances where it is difficult or impossible to measure the characteristics of an entire population (e.g., every US citizen, all brown dogs), sampling is used to obtain information that is expected to represent the broader population.
 - In order for the sample to provide meaningful data, it is important that the sample be chosen in a way to avoid bias, i.e., using a random selection process.
- How do errors occur?
 - No matter how good the "random" sample is, it can never be fully representative of the entire population – the sample only estimates the characteristics of that population.
- Sampling errors are defined as the differences between the sample observations and the actual characteristics of the population
 - Since the total population's values cannot be determined, the "extent" of the sampling error is estimated using probabilities

Non-Representative Sample

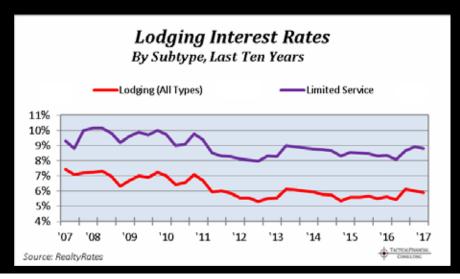
Data-driven decision-making is considered a smart move, but it can be costly or dangerous when something that appears to be true is not actually true.

Even with the best of intentions, some of the world's most famous companies are challenged by skewed results because the data is biased, or the humans collecting and analyzing data are biased, or both.

To Bias or Not To Bias - Should sample represent all available data, the population or should it be purposely skewed toward a certain result?

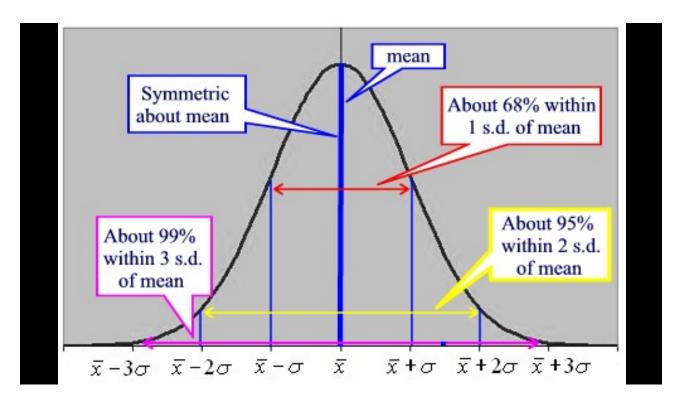


To Bias or Not To Bias - Should sample represent all available data or should it be purposely skewed toward a certain result?



Confidence Intervals

- What is a confidence interval?
 - Confidence intervals consist of a range of values that act as good estimates of the unknown broader population characteristic. Since the data used came from a random sample, the confidence interval is also random.
 - Put more simply, the confidence interval is the range of values we are fairly sure our true value lies in.
- The desired level of confidence is set by the researcher (i.e., it is not determined by the data); typically a 95% confidence level is used

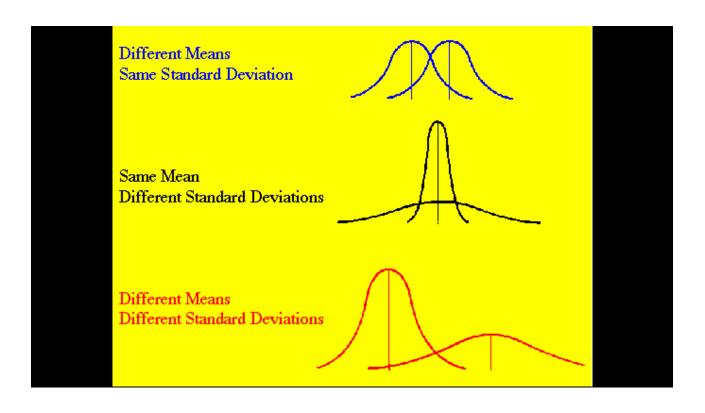


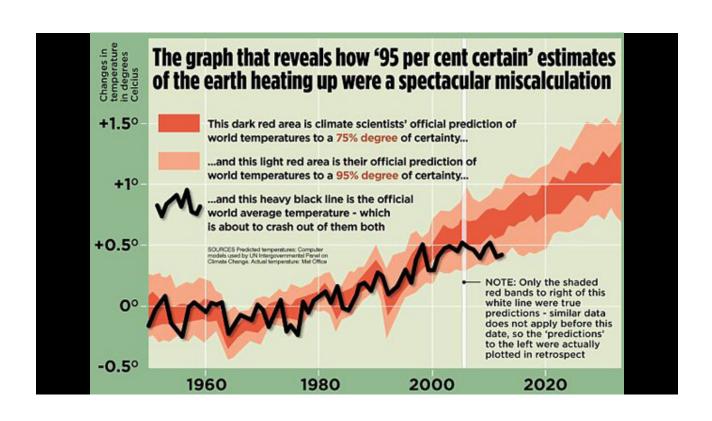
Confidence Intervals - Example

- The height of randomly selected sample of forty men was measured.
 - The mean height was 175 centimeters
 - The standard deviation for the sample was 20 centimeters
- Using a desired level of confidence of 95%, the confidence interval is represented by the following graph:



- Based on the calculations, the conclusion would be that the true mean for all men assuming every man's height could be measured – would fall between 168.8 and 181.2 centimeters – but not necessarily!
- The 95% confidence level means that 95% of our samples would include the true mean, but 5% won't.





False Positives and Negatives

- What is a false positive?
 - A false positive, or "Type 1" error, represents the rejection of a true null hypothesis; in other words, one infers the existence of something that isn't actually there.
- What is a false negative?
 - A false negative, or "Type 2" error, represents the acceptance of a false null hypothesis; in other words, one infers that something doesn't exist, when in fact it does.
- For reference, in statistics, a null hypothesis represents the statement or outcome that one is trying to prove is not true using the evidence/data gathered
 - A researcher will develop the null hypothesis with the intent of rejecting it

False Positives and Negatives - Examples

- Example 1:
 - Hypothesis: "The evidence produced before the court proves that this man is guilty."
 - Null hypothesis (H₀): "This man is innocent."
 - A type I error occurs when convicting an innocent person. A type II error occurs when letting a guilty person go free.
 - A positive correct outcome occurs when convicting a guilty person. A negative correct outcome occurs when letting an innocent person go free.
- Example 2:
 - *Hypothesis:* "A patient's symptoms improve after treatment A more rapidly than after a placebo treatment."
 - Null hypothesis (H₀): "A patient's symptoms after treatment A are indistinguishable from a placebo."
 - A Type I error would falsely indicate that treatment A is more effective than the placebo, whereas a Type II error would be a failure to demonstrate that treatment A is more effective than placebo even though it actually is more effective.

Source: Wikipedia

Regression Analysis

Regression analysis is a mathematical measure of the average relationship between two or more variables in terms of the original units of data.

Types of regression:

- Simple Two variables at a time
- Multiple Three or more variables at a time
- Linear and Non-Linear

Regression Analysis

Linear versus Non-Linear Regression

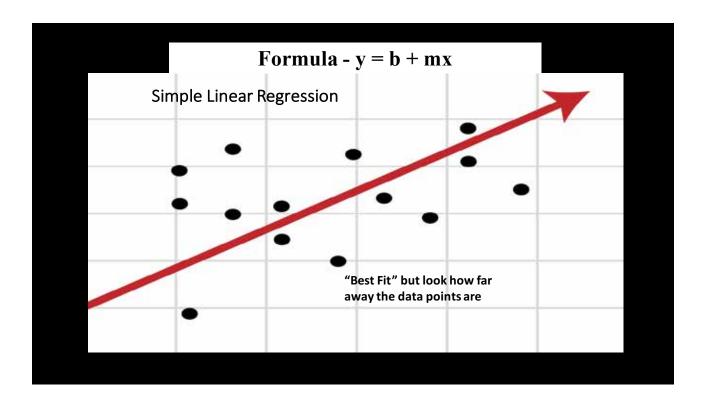
- Linear and nonlinear regression are actually named after the functional form of the models that each analysis accepts.
- A model is linear when each term is either a constant or the product of a parameter and a predictor variable. A linear equation is constructed by adding the results for each term. This constrains the equation to just one basic form:

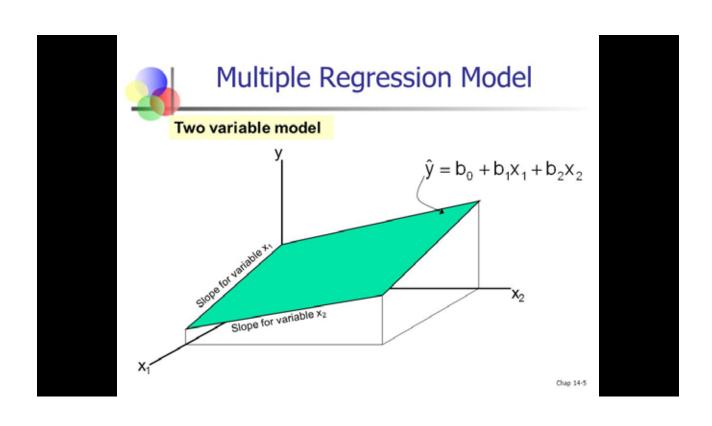
Response = constant + parameter * predictor + ... + parameter * predictor

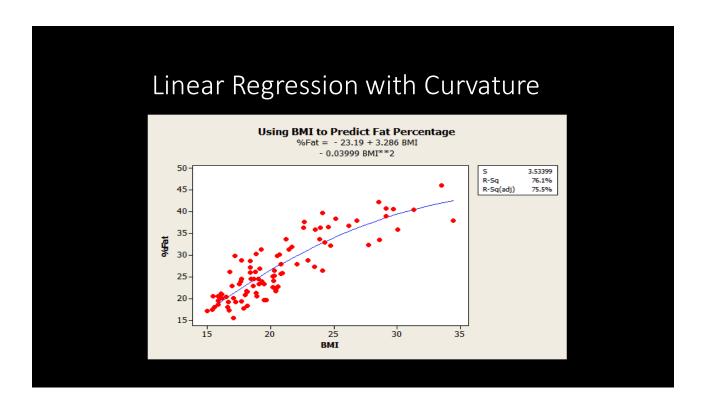
$$Y = b_o + b_1X_1 + b_2X_2 + ... + b_kX_k$$

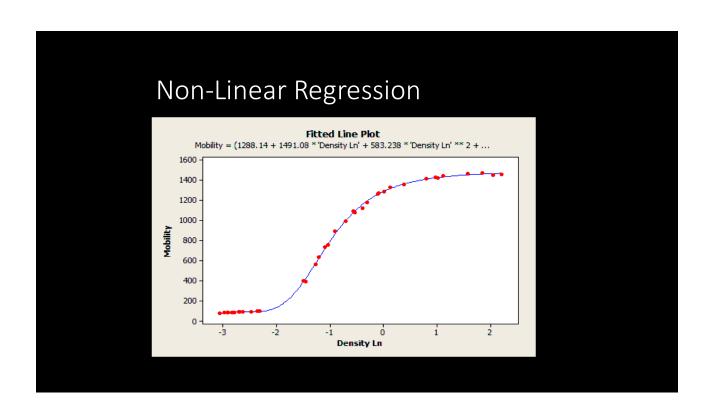
• A linear model can produce a curved line, by squaring one of the variables $Y = b_0 + b_1 X_1 + b_2 X_1^2$

All other models not meeting the parameters for a linear equation are non-linear









Common Mistakes in Regression

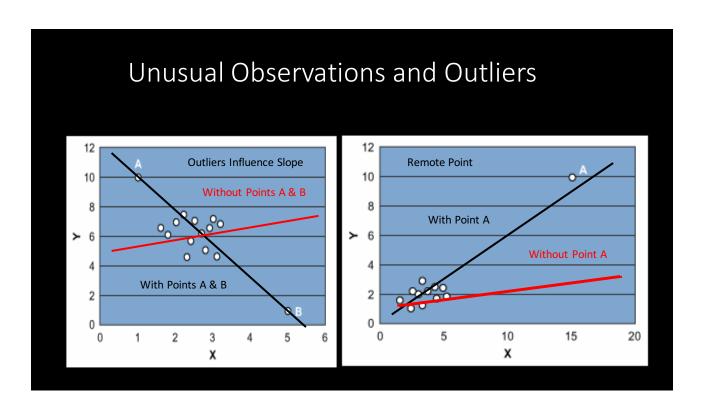
- Model Misspecification
- Unusual Observations and Outliers
- Interpreting the Coefficient of Determination "R-Squared" (R²)

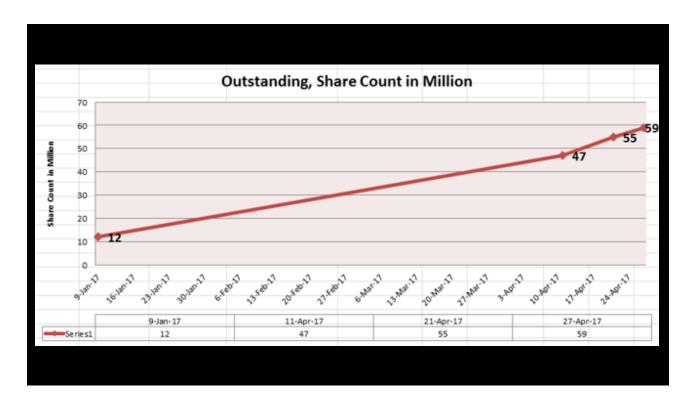
Model Misspecification

- Purpose of regression is to approximate a <u>functional relationship</u> between two or more variables
- Then use that model to predict the variable of interest
- Just because two variables are correlated does not mean they are functionally related. <u>Correlation does not equal Causation!</u>
- For example, a strong statistical relationship between sales of hot chocolate and facial tissues does not mean hot chocolate causes people to need facial tissue.

Unusual Observations and Outliers

- Generally, all data points have an equal weight in estimating the intercept of the regression
- However, the estimate of the <u>slope</u> is more strongly influenced by remote values

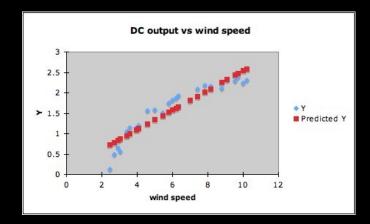




Interpreting Coefficient of Determination

- R² is often misinterpreted:
 - Measures influence of predictor variables on response variable
 - High R² proves correct model has been specified and theory is correct
 - Higher R² in one model means it is better than another model with lower R²
 - R² is simply a measure of the spread of data around a regression line
 - It is not an estimator because there is no relevant population parameter
 - Can be useful in comparing models with same response variable but different predictor variables
 - Use of a test statistic (e.g., t-test, F-test) will show if estimated coefficients are statistically significant

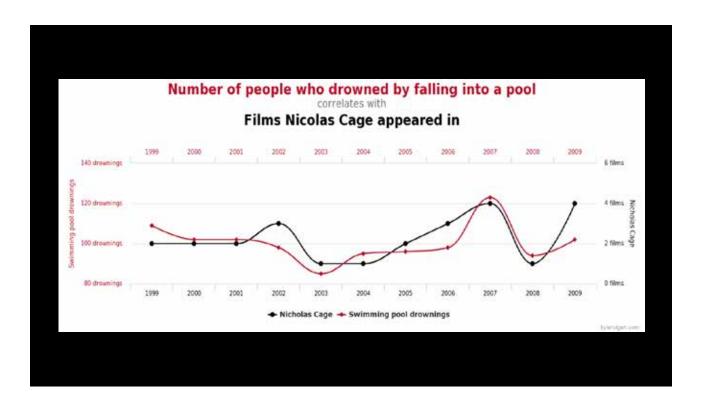
Coefficient of Determination

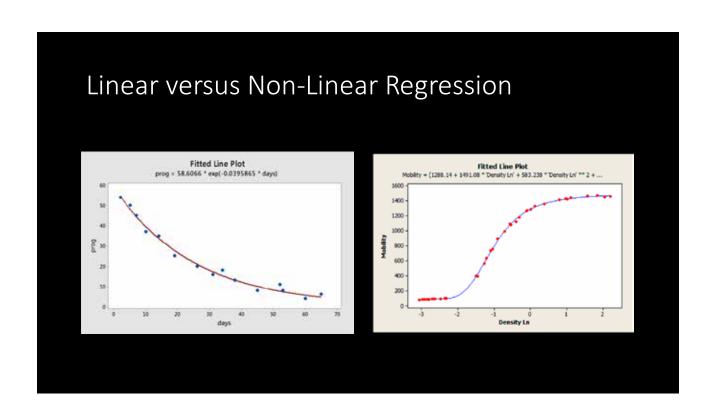


R² equals .87, pretty high
But, data clearly show a curved trend

Mistakes in Thinking About Causation

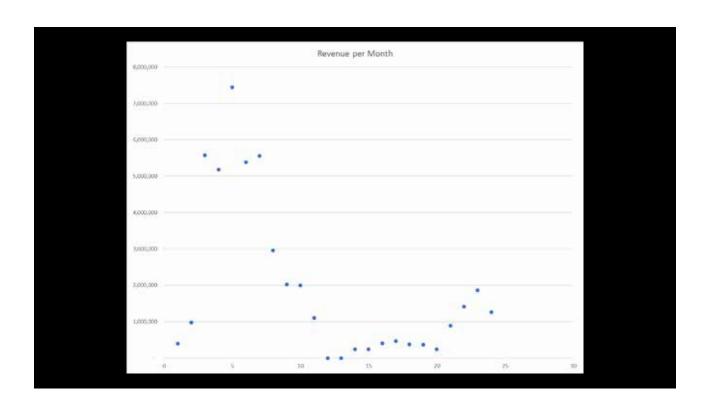
- Correlation is often confused with Causation
 - Example Shoe sizes and scores on a standardized reading exam (given to all students in an elementary school)
 - Data show that students with larger shoe sizes scored higher
 - While <u>correlated</u>, it is absurd to say higher reading scores <u>cause</u> larger shoe sizes (older students have bigger feet AND have read more than younger students, hence higher scores)

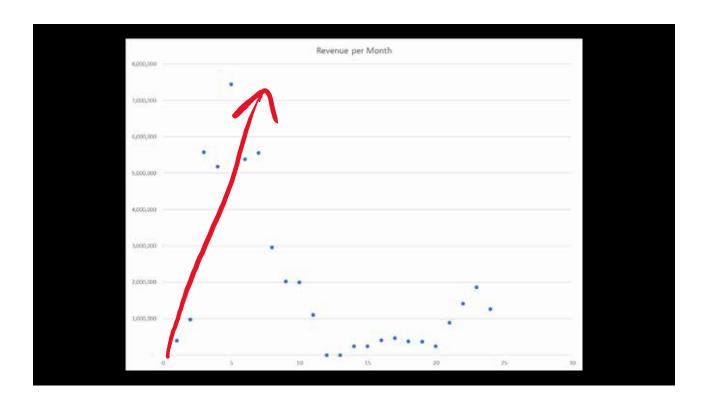


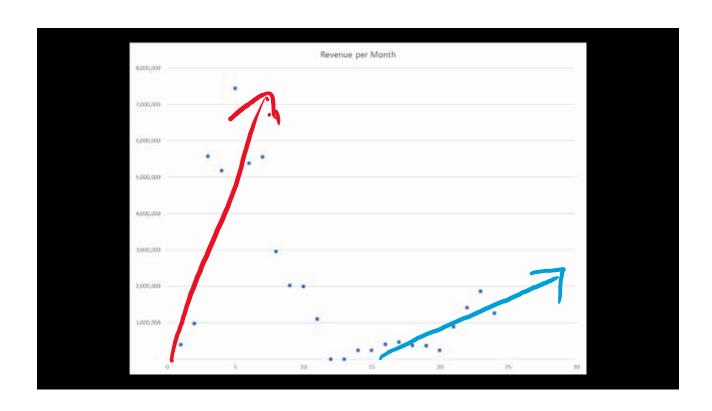


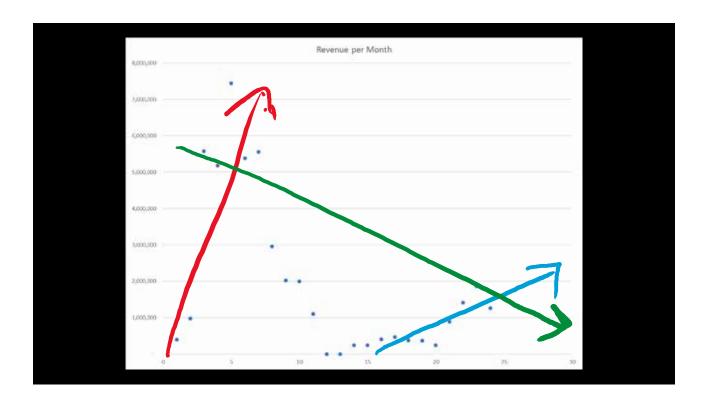
A Simple Case Study

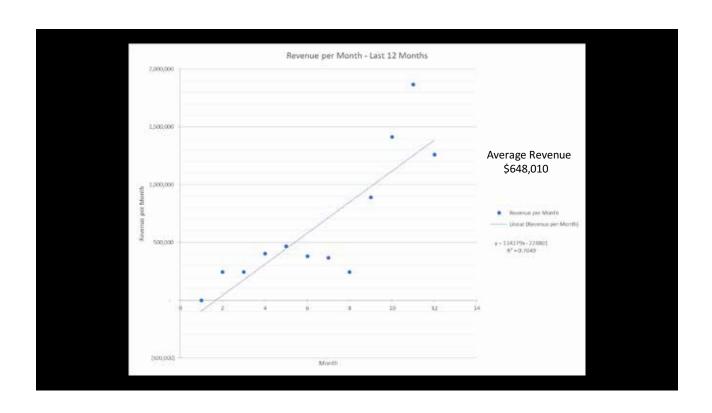
- A partnership was formed to develop a 200 unit condominium project in two phases
- Phase One of the project started out pretty well, but a recession hit toward the end of Year One, slowing sales dramatically and requiring significant discounts to sell the remaining units
- All units from Phase One were sold by the end of Year Two as the economy recovered
- Phase Two is about to kick off, but the parties could not agree on a valuation
- The key issue in the case is how to project revenue per month based on the experience in Phase One

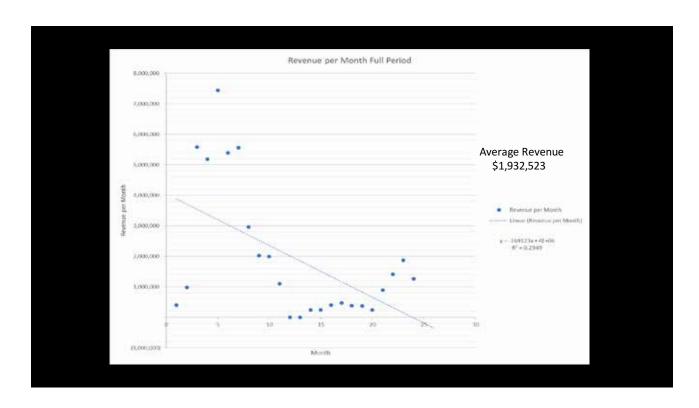


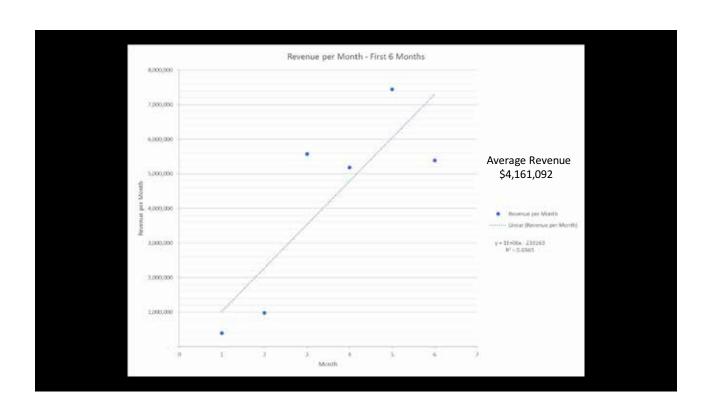








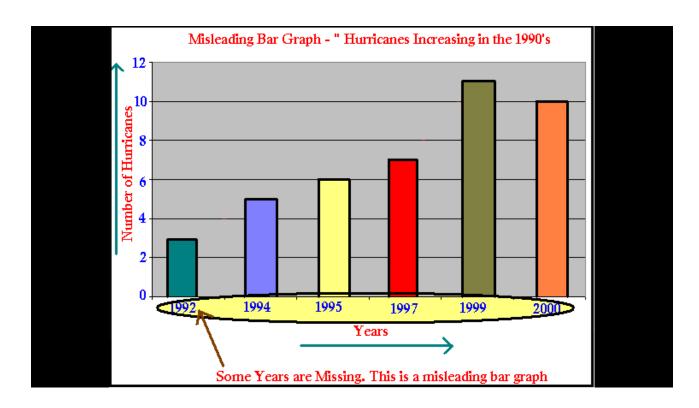


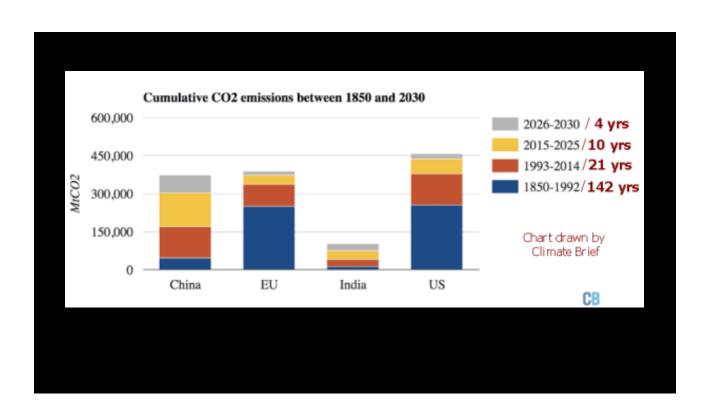


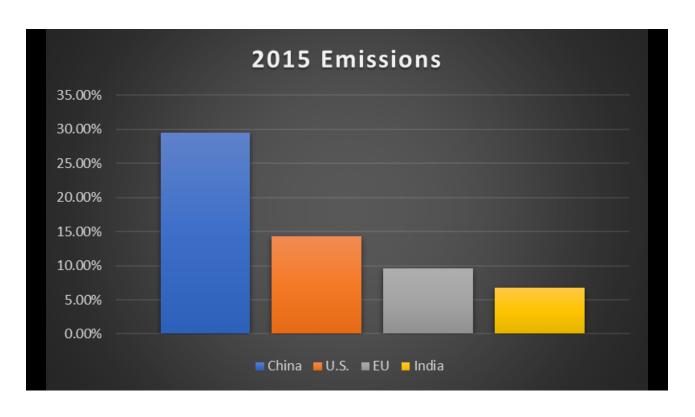
Sample	Revenue		Units Sold	Average Unit Price	
Period	per Month		per Month	per Month	
Full Period	\$	1,932,523	4.17	\$	362,656
First 6 Months	\$	4,161,092	7.83	\$	493,362
First 12 Months	\$	3,217,037	6.50	\$	420,147
First 18 Months	\$	2,241,211	4.67	\$	363,629
Last 12 Months	\$	648,010	1.83	\$	305,165
Last 6 Months	\$	1,006,458	2.67	\$	359,736

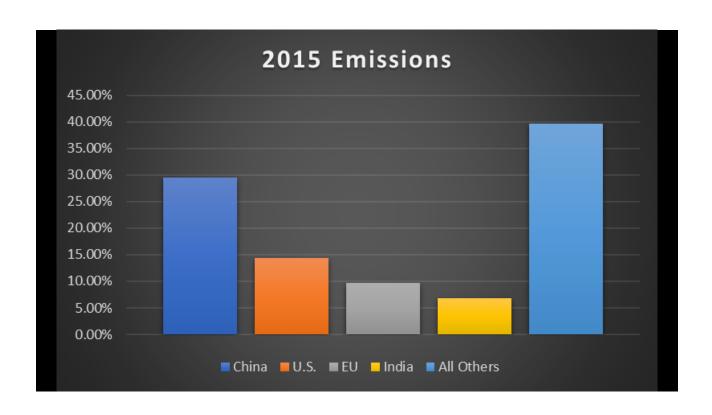
The Gee-Whiz Graph – Deceptive Graphs

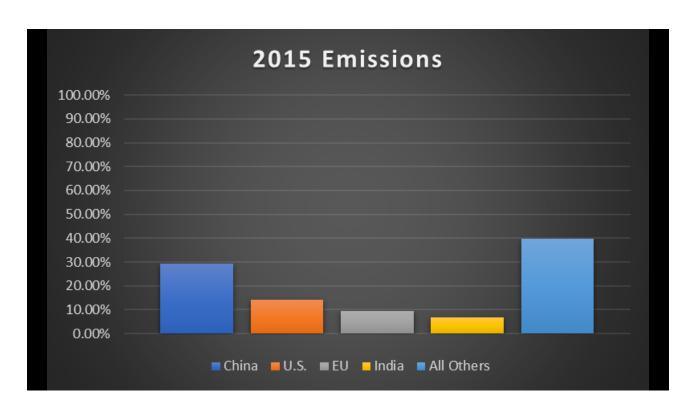
The Seven Deadly Sins of Statistical Analysis



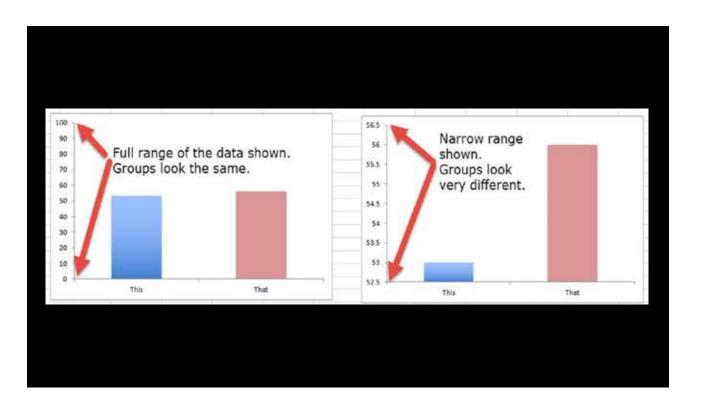


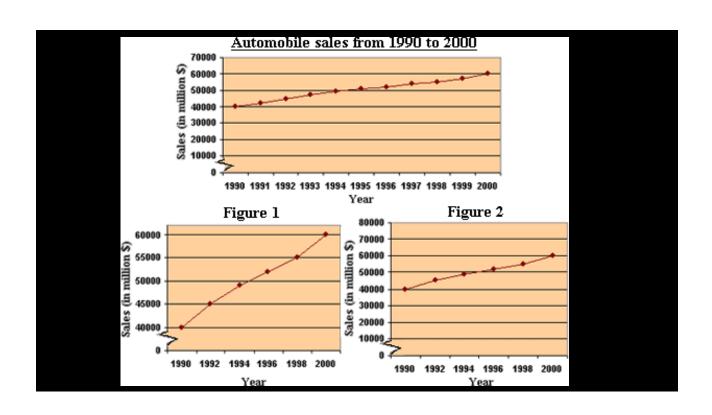


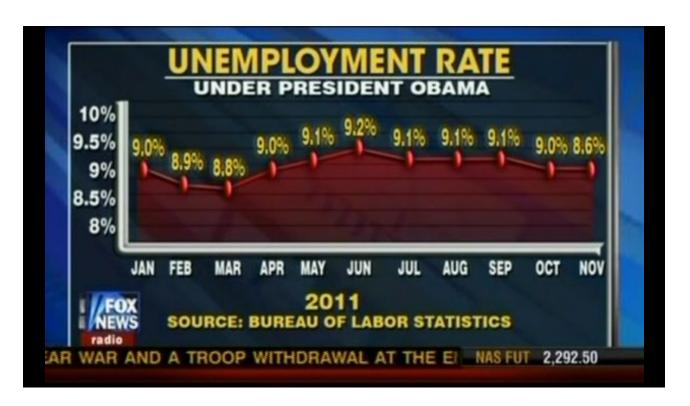


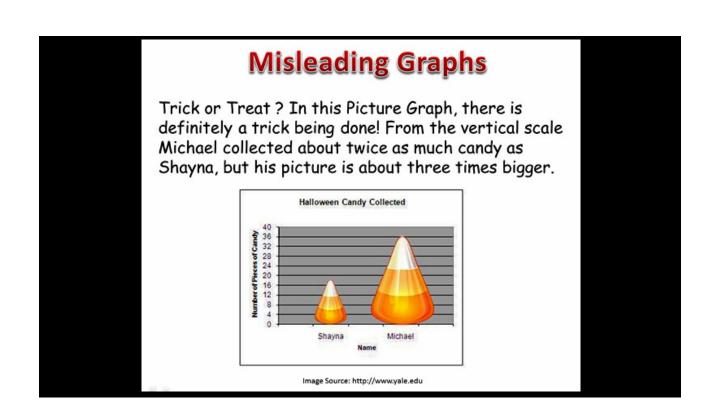


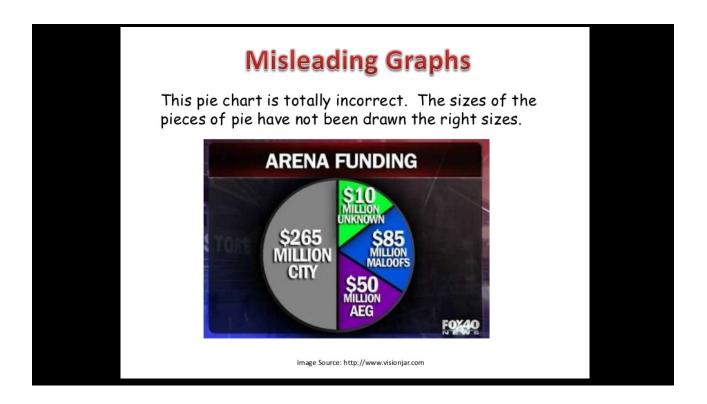


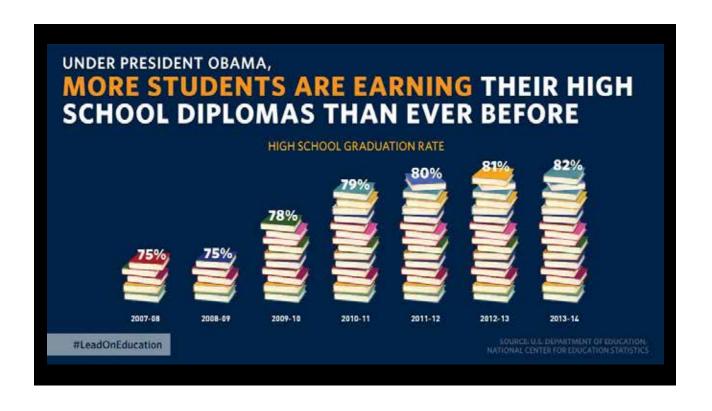


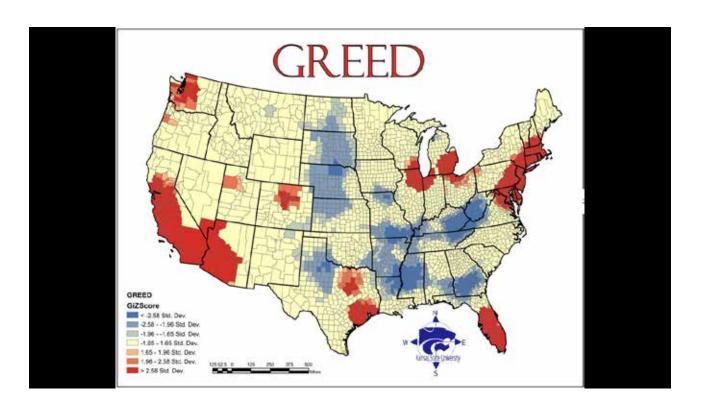


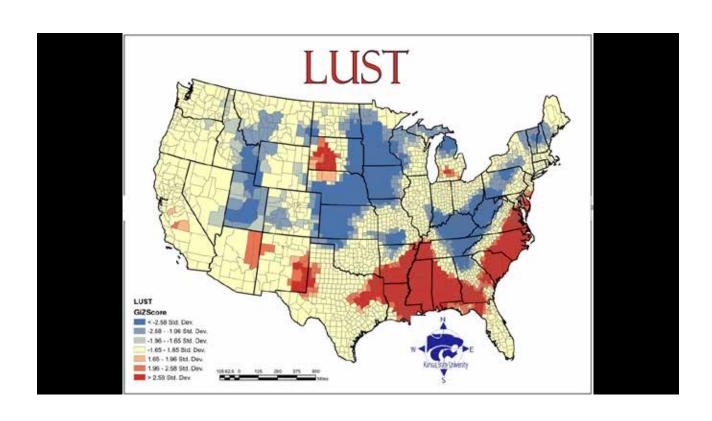














- Problems
- Starts in 1859
- Projects into future (an unknown)
- Ignores Trade Exports to Other Countries using goods made from CO2 emissions
- Emphasizes U.S. however in 2015 the actual releases were:
- China = 29.51%
- U.S. = 14.34%
- EU = 9.62%
- India = 6.81%

One nation, seven sins
Geographers measure propensity for evil in states, counties
By Abigail Goldman
Thursday, March 26, 2009 | 2 a.m.
Seven Deadly Sins in Nevada

Launch slideshow »
Seven Deadly Sins Nationwide

Launch slideshow »
Seven Deadly Sins Nationwide with "Hot Spots"

Launch slideshow »
Seven Deadly Sins Comparison

Launch slideshow »
Beyond the Sun
Kansas State University's Department of Geography
The question of evil and where it lurks has been largely ignored by the scientific community, which is why a

recently released study titled "The Spatial Distribution of the Seven Deadly Sins Within Nevada" is groundbreaking: Never before has a state's fall from grace been so precisely graphed and plotted.

Geographers from Kansas State University have used certain statistical measurements to quantify Nevada's sins and come up with a county-by-county map purporting to show various degrees of lust, gluttony, greed, sloth, wrath, envy and pride in the Silver State. By culling statistics from nationwide databanks of things like sexually transmitted disease infection rates (lust) or killings per capita (wrath), the researchers came up with a sin index. This is a precision party trick — rigorous mapping of ridiculous data.

Their findings were presented Tuesday at the Association of American Geographers' annual meeting at the Riviera, where Kansas State geography research associate Thomas Vought fielded questions while standing next to a poster of his research. Seven maps of Nevada, in seven different colors, for seven different sins. The darker a county, the more evil it is.

Greed was calculated by comparing average incomes with the total number of inhabitants living beneath the poverty line. On this map, done in yellow, Clark County is bile (see map on Page 2).

Envy was calculated using the total number of thefts — robbery, burglary, larceny and stolen cars. Rendered in green, of course, Clark County is emerald.

Wrath was calculated by comparing the total number of violent crimes — murder, assault and rape — reported to the FBI per capita. Vought and his colleagues used the color red to illustrate wrath, so Clark County looks like a fresh welt. Washoe is slightly statistically duller. Everywhere else is a friendly pork pink.

Lust was calculated by compiling the number of sexually transmitted diseases — HIV, AIDS, syphilis, chlamydia and gonorrhea — reported per capita. Here again, Clark and Washoe counties are worst. Carson City County is a close third.

Gluttony was calculated by counting the number of fast food restaurants per capita, and this is one category

where Clark County is bested. First in deep fry goes to Carson City.

Sloth was calculated by comparing expenditures on arts, entertainment and recreation with the rate of employment. Here again Clark County is beat, scoring only average on the scale of sloth.

And pride, lastly, is most important. The root of all sins, in this study, is the aggregate of all data. Vought and his Kansas colleagues combined all data from the six other sins and averaged it into an overview of all evil. So pride, mapped in purple, shows the states two darkest bruises: counties Clark and Carson City.

Yet, in the grand scheme of things, maybe we're not that bad. While Vought and his colleagues spent four weeks on the detailed Nevada study, they also ran the numbers on some 3,000 counties across the country, a nationwide survey of sin.

Turns out Nevada is unremarkable when compared with other states. Sure, we have a little discoloration around Washoe and Clark counties when it comes to wrath, and Southern Nevada as a whole stands out in the nationwide map of greed, but other than that, we're almost colorless, boring even, when compared with Texas, which ranked high for gluttony, or wrath, which was concentrated in Florida and surrounding states. Moreover, the Kansas geographers also compared the level of sin in 10 top casino markets, and while the Las Vegas Strip ranked first for greed, it could muster no better than third place for pride, the aggregate of all sins. It was the southern gambling cities — Lula, Miss.; Biloxi, Miss.; and Shreveport, La., that came out on top of the bottom. Why, exactly, remains to be seen. The Kansas geographers started this project, it seems pretty clear, for the erudite amusement; something to stand out at a 6,000-person convention consumed with the world's heavy questions. But if Tuesday's convention crowd was evidence, the sin study was interesting to other scholars as well. So Vought and colleagues plan to continue their national study of evil.

"It's too much fun," Vought said, smiling in a way that suggested, if not pride, then a good deal of pleasure.