

# GEOG 312

## Data Analysis in Geography

Dr. George H. Allen



Data Analysis in Geography

[www.terragalleria.com](http://www.terragalleria.com)

# Ice Breaker

## 1. Introduce yourself

- Your name
- Your major / degree program
- Where you are from

## 2. What are you excited about for this course?

## 3. What are you worried about for this course?



# About You



Data Analysis in Geography



*“The unity of all science consists  
alone in its method, not its material”*  
(Pearson)

# My Expectations

- To be able to read and understand the empirical literature in your field; to become a critical consumer of information
- To be able to both design and implement a study using quantitative and qualitative methods
- To be able to intelligently participate in research projects, evaluations, and studies that you undertake
- To understand the difference between scientifically acquired knowledge and other kinds of information



- **Hypothesis:** “Social decay in the inner city is caused by a prevalence of dissociative (community destroying) institutions compared to associative institutions”
- **Methods:** Relate number of homicides to the number of liquor stores in different regions of the city
- **Result:** Positive correlation of 0.31, which suggests that more suicides and homicides in areas with more liquor stores
- **Conclusion:** Dissociative institutions cause social decay in inner cities and liquor stores should be removed from inner-city areas
- Should we accept these conclusions?

## Example 1

# Example 2

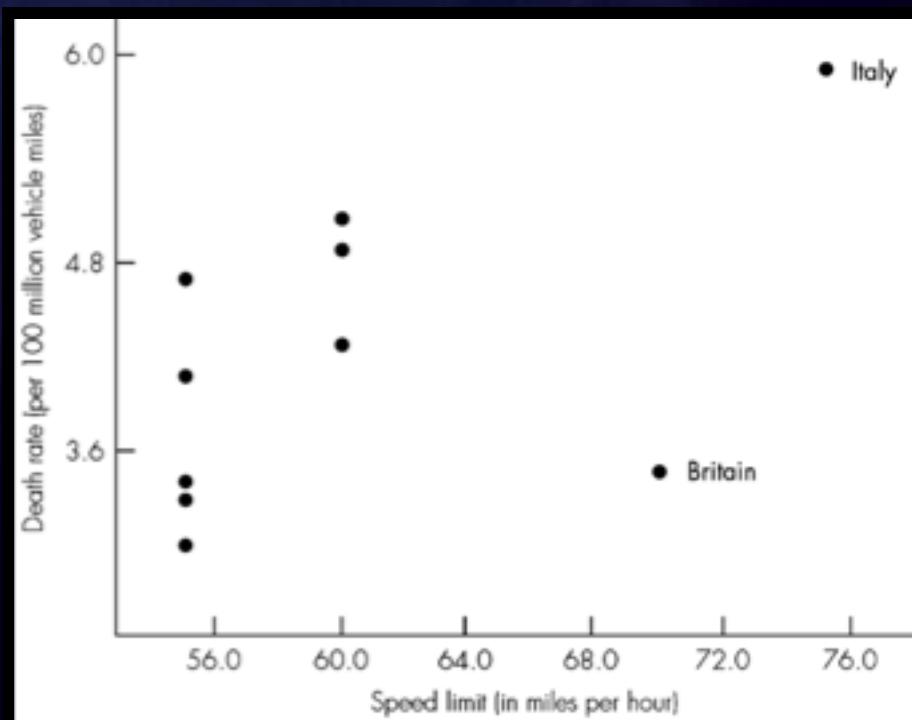
- **Viewer Question:** “Has anyone noticed that life expectancy in Canada under our health system is higher than the USA?”
- **Response:** “Well, that's to be expected Peter, because we have 10 times as many people as you do. That translates to 10 times as many accidents, crimes, down the line.”
- Is this response correct?



COUNTRY	Death Rate (Per 100 Million Vehicle Miles)	Speed Limit (in Miles Per Hour)
Norway	3.0	55
United States	3.3	55
Finland	3.4	55
Britain	3.5	70
Denmark	4.1	55
Canada	4.3	60
Japan	4.7	55
Australia	4.9	60
Netherlands	5.1	60
Italy	6.1	75

- Correlation between death rate and speed limit is 0.55
- If Italy removed, correlation drops to 0.098
- If then Britain removed, correlation jumps to 0.70

## Example 3: Highway Deaths and Speed Limits



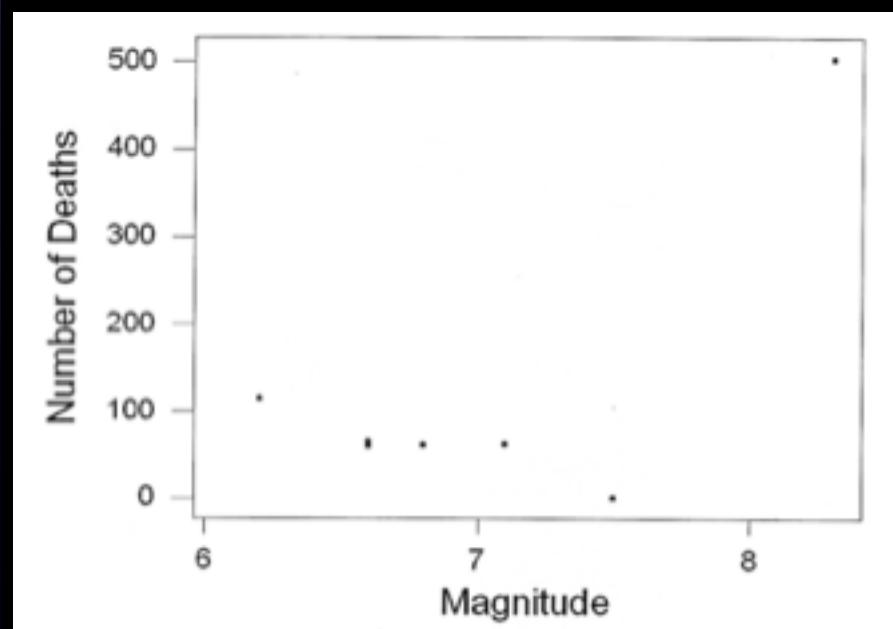
# Example 4: Earthquakes in Continental U.S.

Date	Location	Deaths	Magnitude
August 31, 1886	Charleston, SC	60	6.6
April 18 – 19, 1906	San Francisco	503	8.3
March 10, 1933	Long Beach, CA	115	6.2
February 9, 1971	San Fernando Valley, CA	65	6.6
October 17, 1989	San Francisco area	62	6.9
June 28, 1992	Yucca Valley, CA	1	7.4
January 17, 1994	Northridge, CA	61	6.68

Correlation is **0.689**, relatively strong positive association

If SF earthquake of 1906 removed, correlation is **-0.92**; higher magnitude quakes associated with fewer deaths

The 1906 earthquake was before earthquake building codes enforced. Next largest quake in 1992 occurred in sparsely populated area.



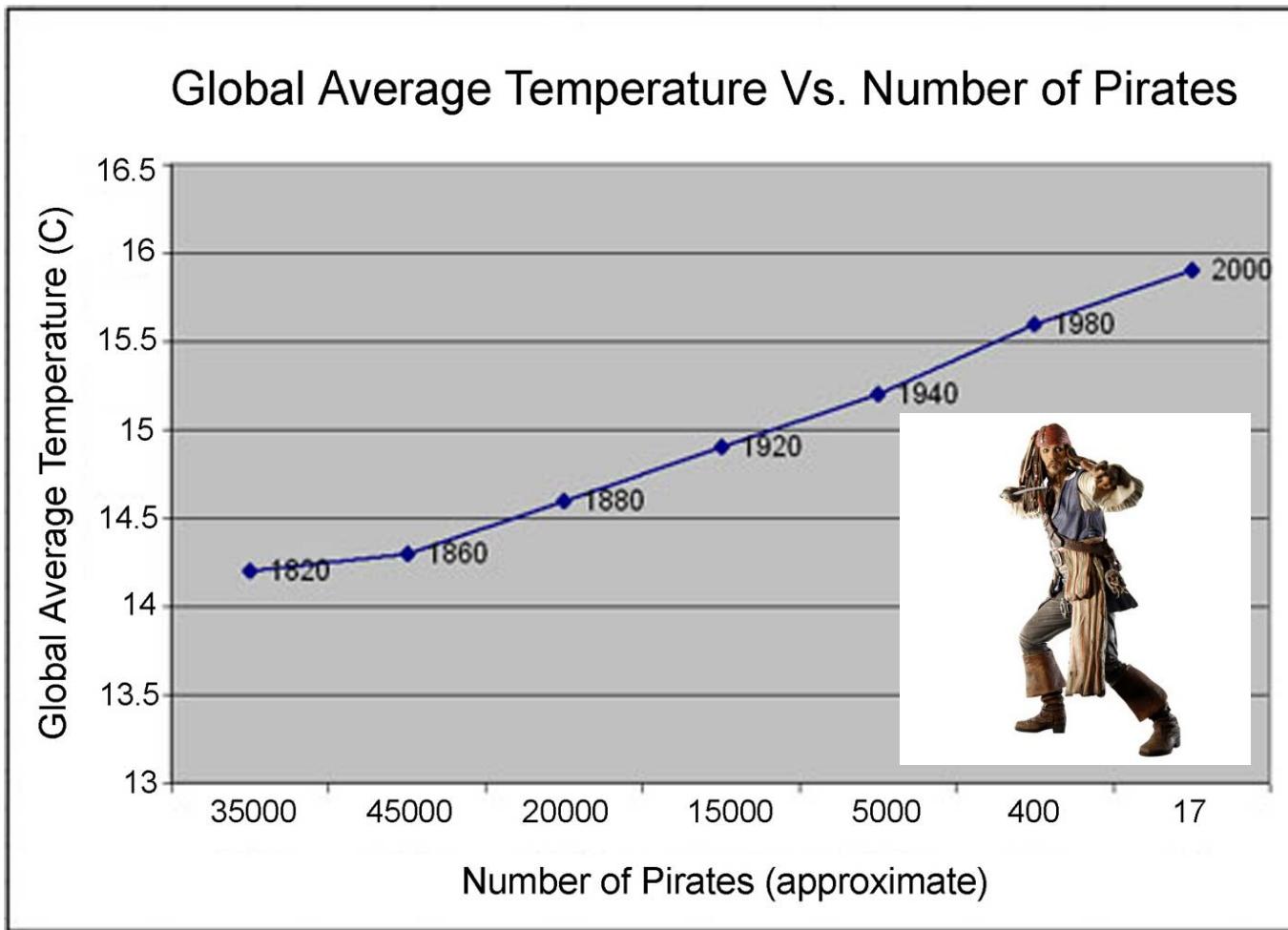
***Positive proof of global warming.***



***18th  
Century      1900      1950      1970      1980      1990***

[http://www.youtube.com/watch?v=emA1eC92a3x\\_s&url=http://alibayes.blogspot.com/2007/09/short-skirts-and-stocks-from-history.html&featureplayer\\_embedded#t=82](http://www.youtube.com/watch?v=emA1eC92a3x_s&url=http://alibayes.blogspot.com/2007/09/short-skirts-and-stocks-from-history.html&featureplayer_embedded#t=82)

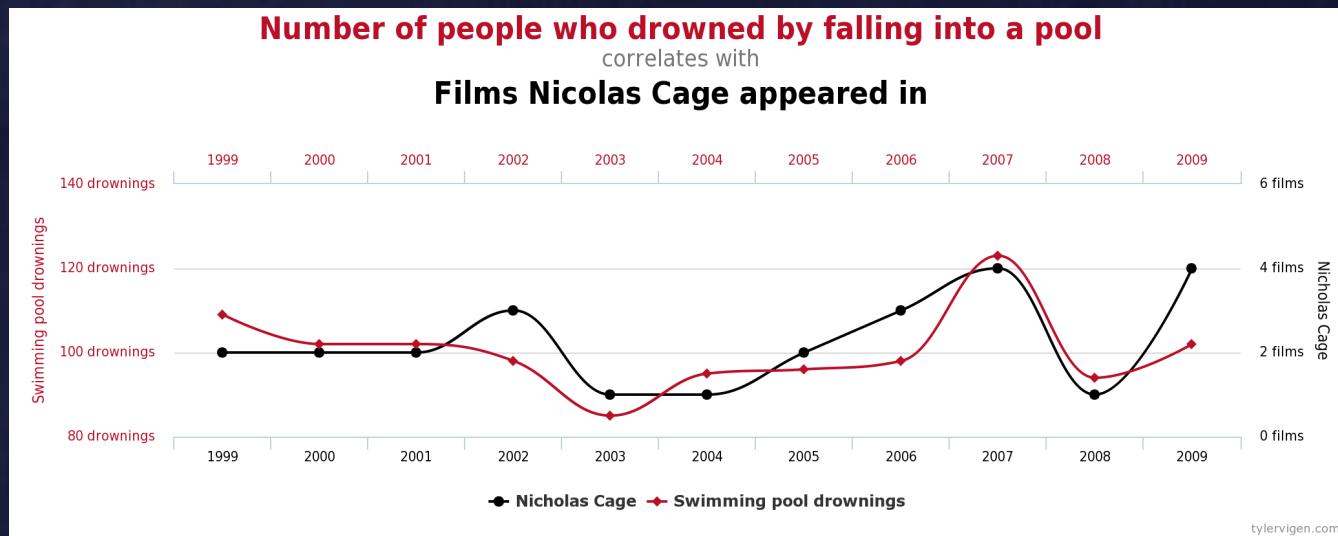
# **STOP GLOBAL WARMING: BECOME A PIRATE**



WWW.VENGANZA.ORG

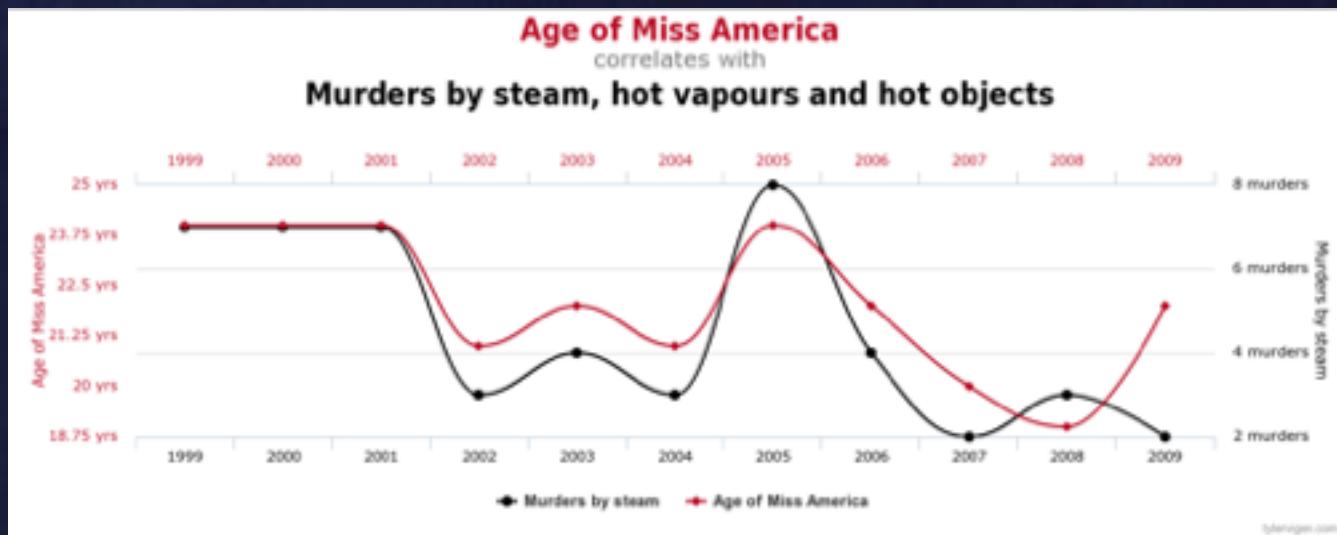


# Spurious Correlations

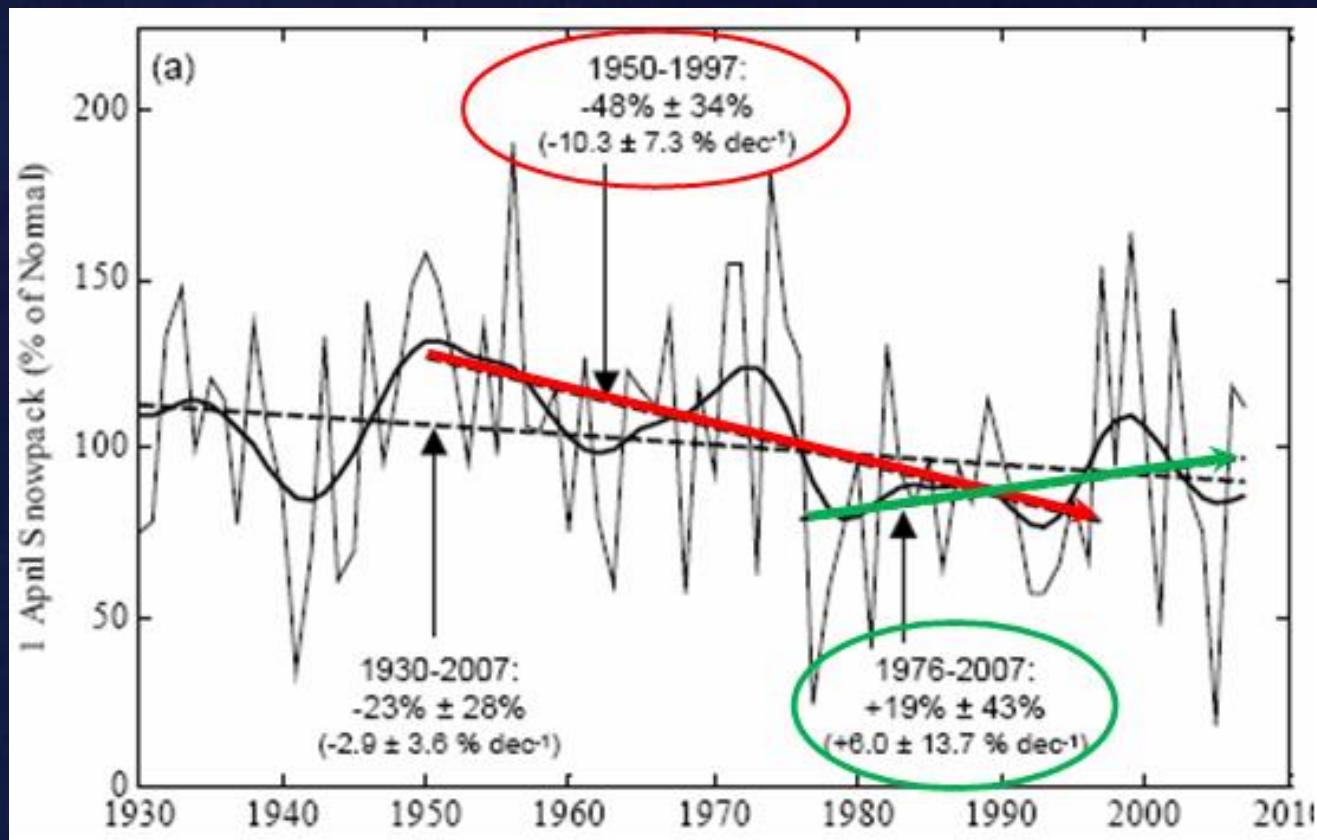


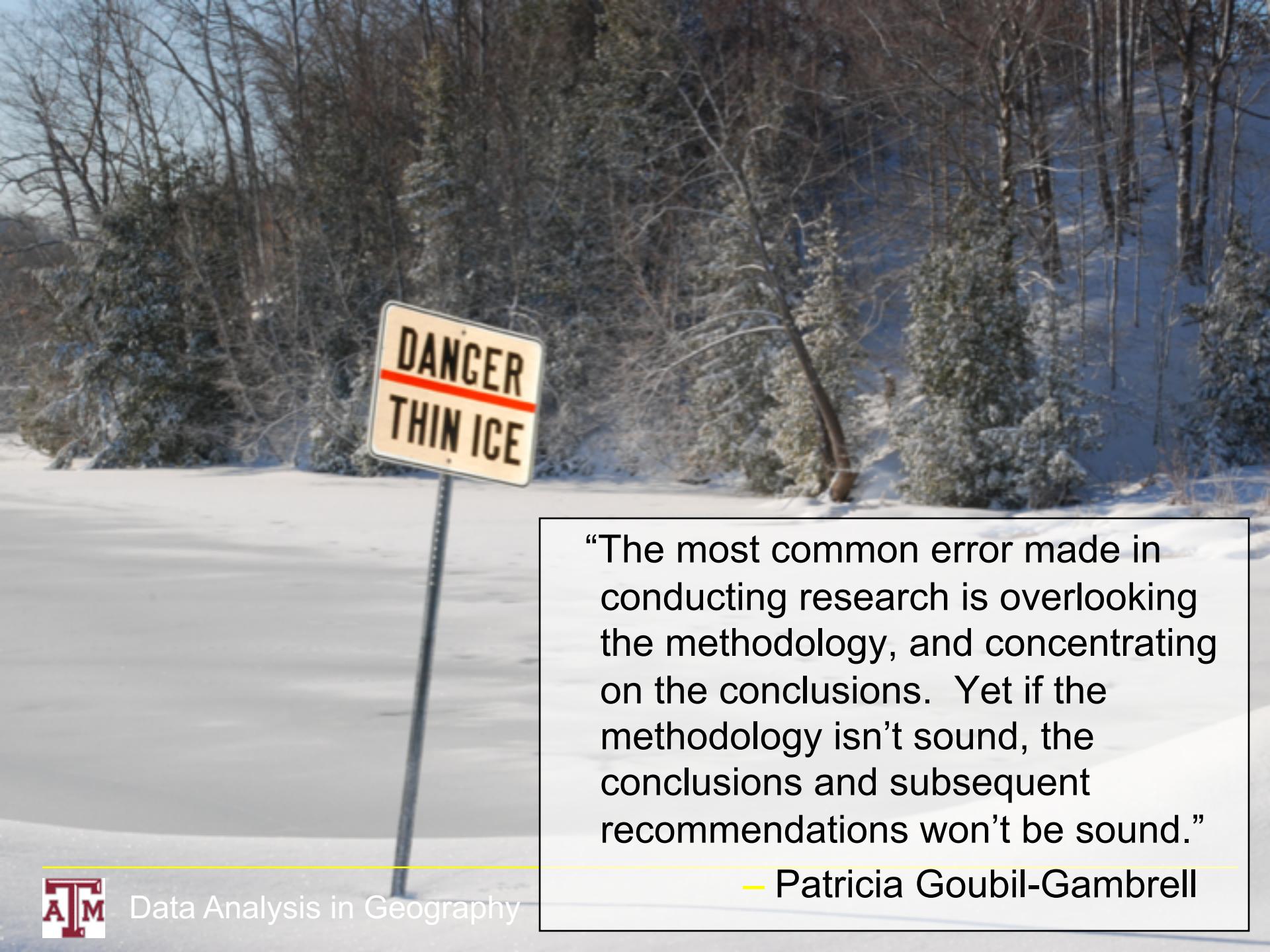
<http://tylervigen.com/>

# Spurious Correlations



<http://tylervigen.com/>





“The most common error made in conducting research is overlooking the methodology, and concentrating on the conclusions. Yet if the methodology isn’t sound, the conclusions and subsequent recommendations won’t be sound.”

— Patricia Goubil-Gambrell

# Course Logistics:

## This will *mainly* be a paperless class

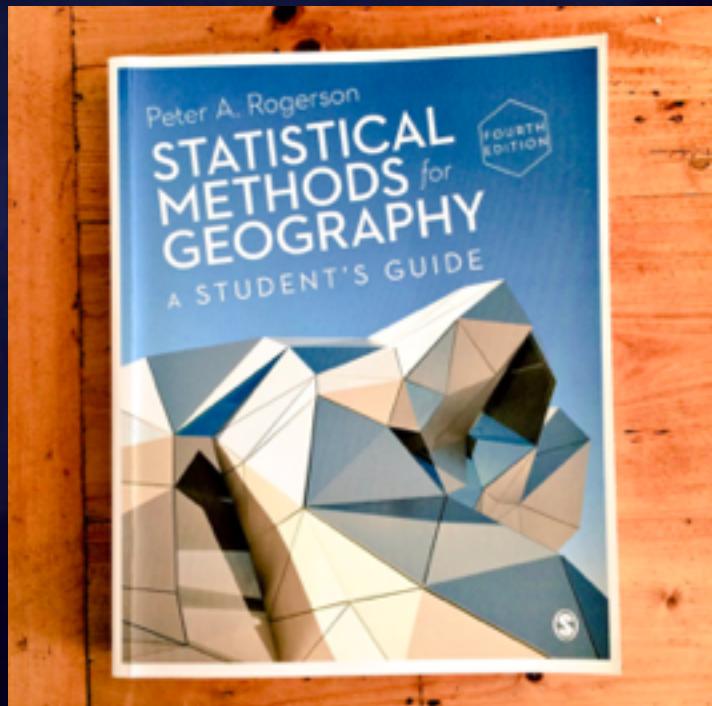
- Syllabus, assignments, and in-class activities will be made available on eCampus... **bring your internet-connected device to class**
- Syllabus → Let's go!



# Textbook

- Statistical Methods for Geography: A Student's Guide. Peter A. Rogerson, SAGE Publications, 2010, 4th edition. ISBN: 9781446295731.
- Additional required readings will be assigned and posted to eCampus

[eCampus.tamu.edu](http://eCampus.tamu.edu)



# Class Setup

- **Tuesday:** Lectures (usually)
- **Thursdays:** R workshops/Data collection (often)



# Are you a useR?



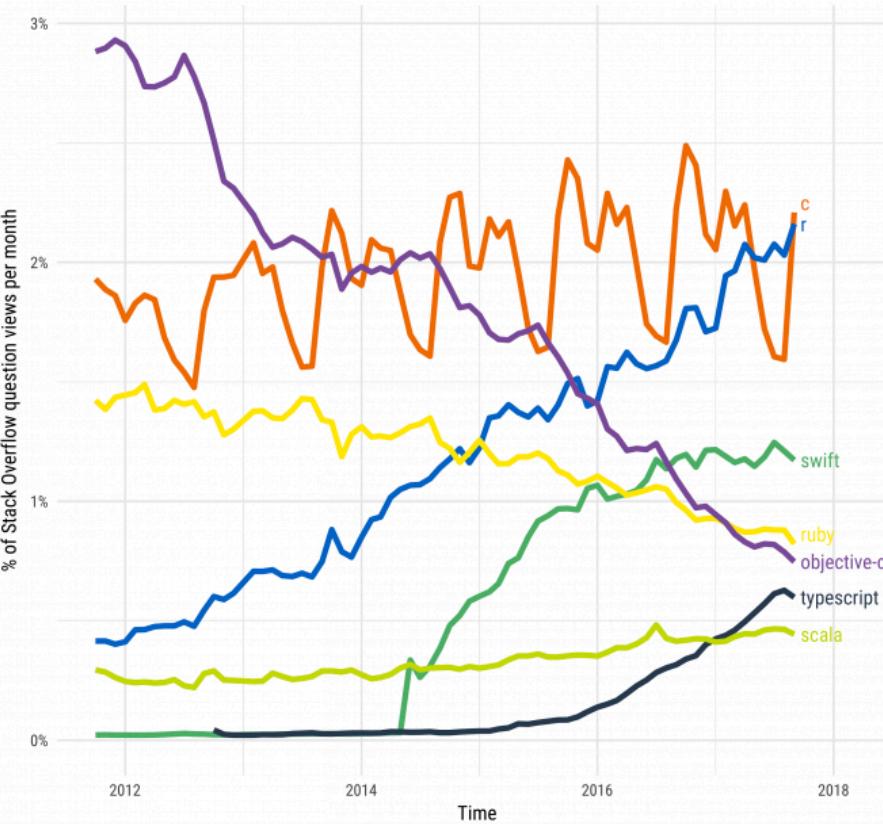
- In this class, we will be using a programming language called R
- R is a language and environment for statistical computing and graphics
- R is open source software (free)
  
- Additionally, we will be using a third-party program called “**RStudio**” which is an “Integrated Development Environment” (IDE), or a program in which we can use R.



# The Impressive Growth of R

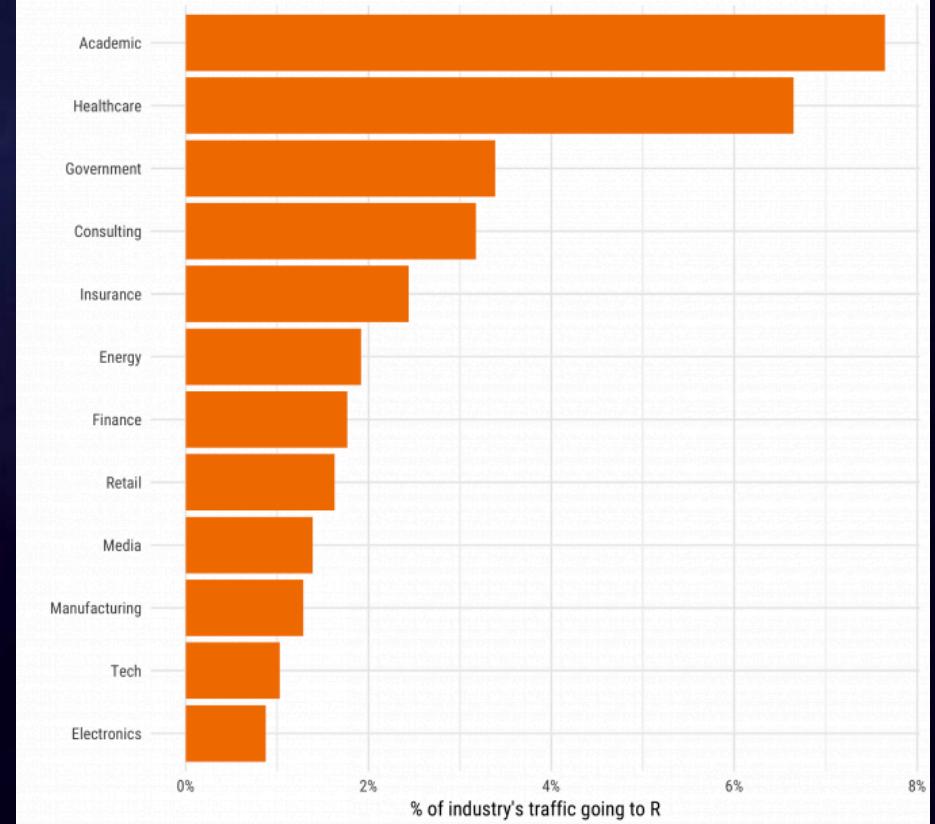
## Stack Overflow Traffic to Programming Languages

Based on visits to Stack Overflow questions from World Bank high-income countries.  
The more-visited languages of Python, JavaScript, Java, C#, and PHP were omitted.



## Visits to R by industry

Based on visits to Stack Overflow questions from the US/UK in January-August 2017.  
The denominator in each is the total traffic from that industry.



# Homework

1. Install R and RStudio on your computer by Thursday's class.
  - Directions can be found on eCampus on this how to document:  
[http://people.tamu.edu/~geoallen/courses/312/install\\_R\\_&\\_RStudio.pdf](http://people.tamu.edu/~geoallen/courses/312/install_R_&_RStudio.pdf)
2. Bring your laptop to class on Thursday, if you have one.
  - Ideally, you will be able to bring a laptop to class everyday. If you forget, or don't have one, you will still be able to follow along.
3. Readings for next time:
  - Rogerson (the course textbook): Chapter 1
  - Clifford et al. (2010) "Key Methods in Geography" Chapter 1:  
[http://people.tamu.edu/~geoallen/courses/312/Ch01\\_Clifford\\_French\\_Valentine.pdf](http://people.tamu.edu/~geoallen/courses/312/Ch01_Clifford_French_Valentine.pdf)



# Questions?

