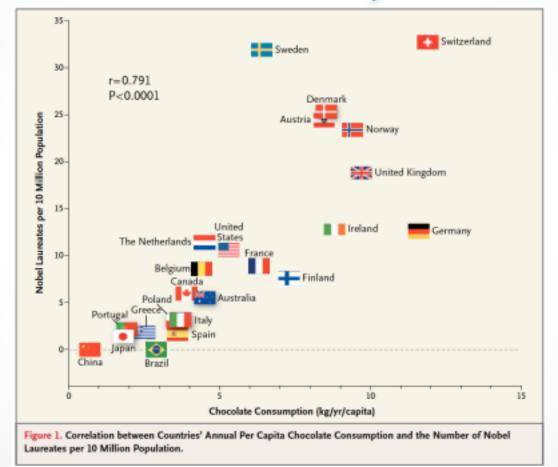
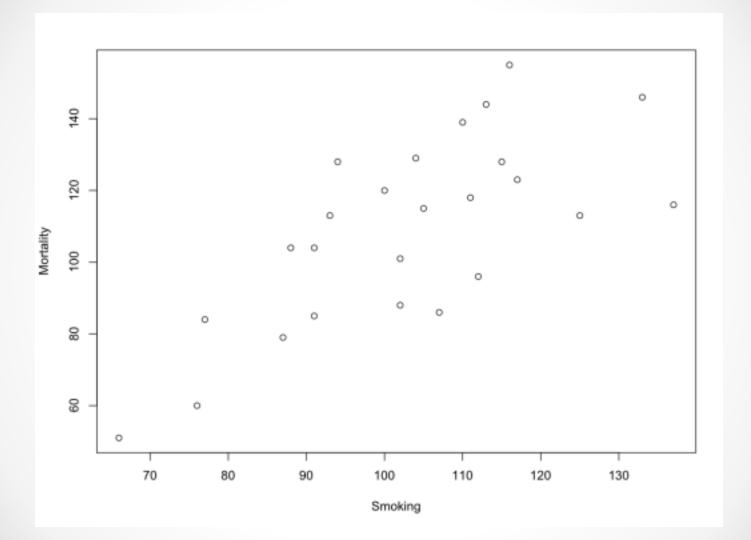
Algorithms week 3-1: Causation

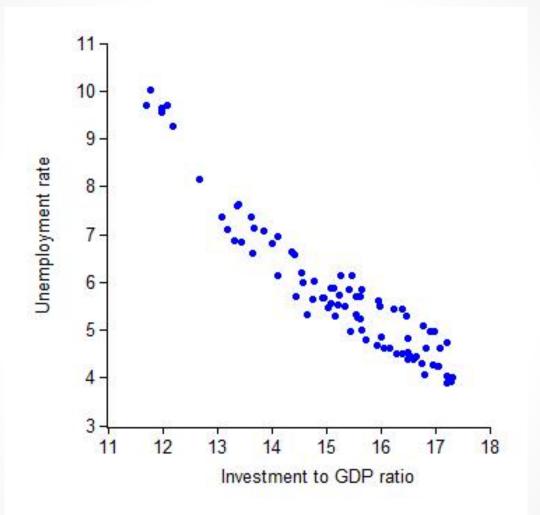
Jonathan Stray
Columbia Lede Program
July 30, 2018

Does chocolate make you smarter?



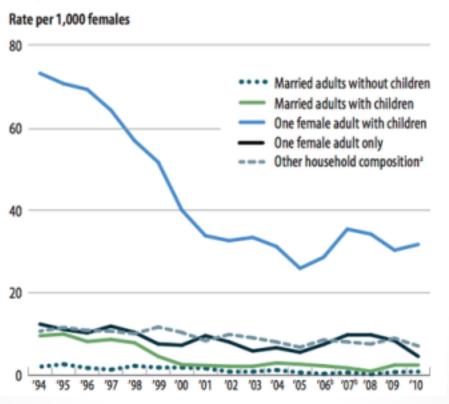
Occupational Group	Smoking	Mortality
Farmers, foresters, and fisherman	77	84
Miners and quarrymen	137	116
Gas, coke and chemical makers	117	123
Glass and ceramics makers	94	128
Furnace, forge, foundry, and rolling mill	116	155
Electrical and electronics workers	102	101
Engineering and allied trades	111	118
Woodworkers	93	113
Leather workers	88	104
Textile workers	102	88
Clothing workers	91	104
Food, drink, and tobacco workers	104	129
Paper and printing workers	107	86
Makers of other products	112	96

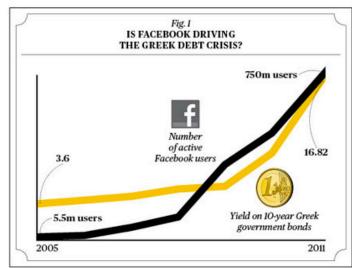


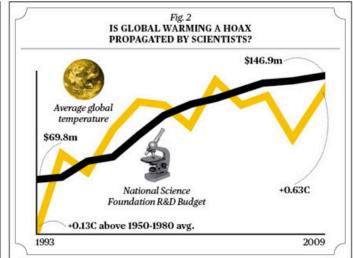


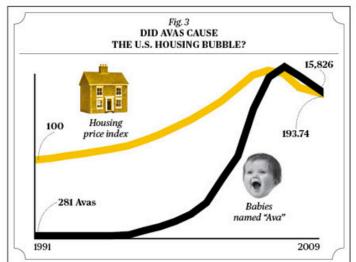
Does marriage make women safer?

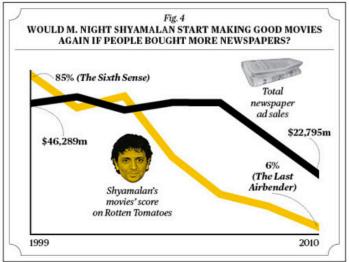
Intimate partner violence against females, by household composition, 1993–2010



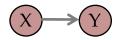




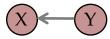




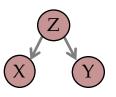
How correlation happens



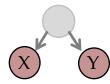
X causes Y



Y causes X



Z causes X and Y



hidden variable causes X and Y





random chance!

Guns and firearm homicides?



if you have a gun, you're going to use it



if it's a dangerous neighborhood, you'll buy a gun





the correlation is due to chance

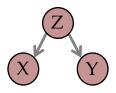


Exactly what to say in a first message, OK Cupid blog, 2009

Messages and responses



Comments on her appearance make her respond less

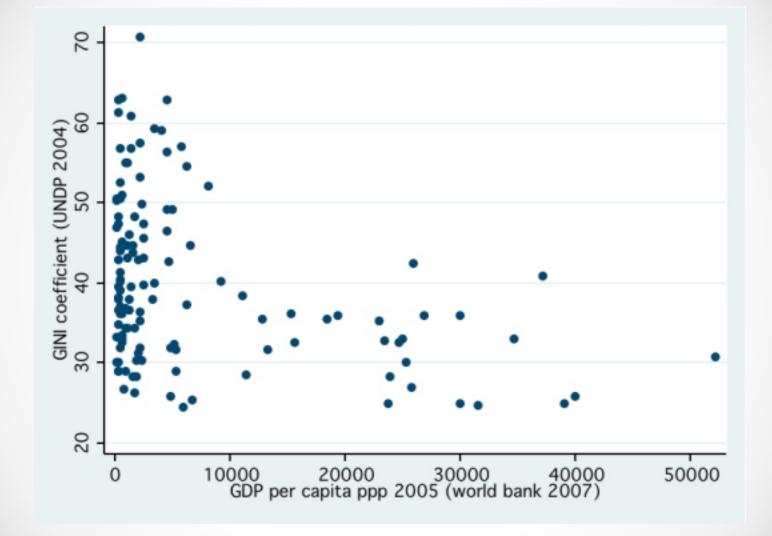


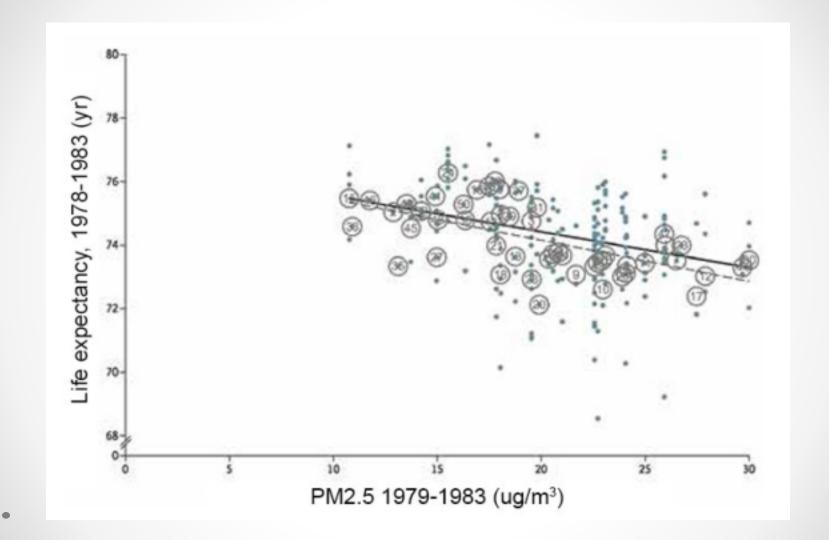
if a woman is attractive,

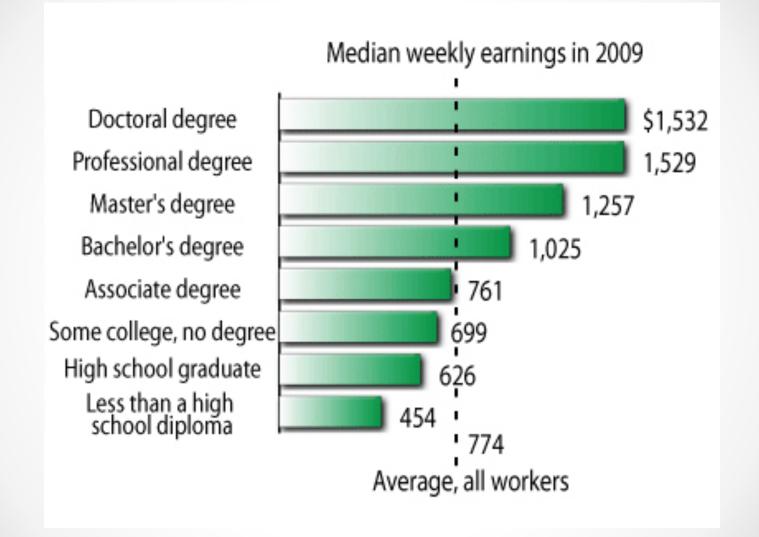
- 1) she'll respond less
- 2) people will tell her that

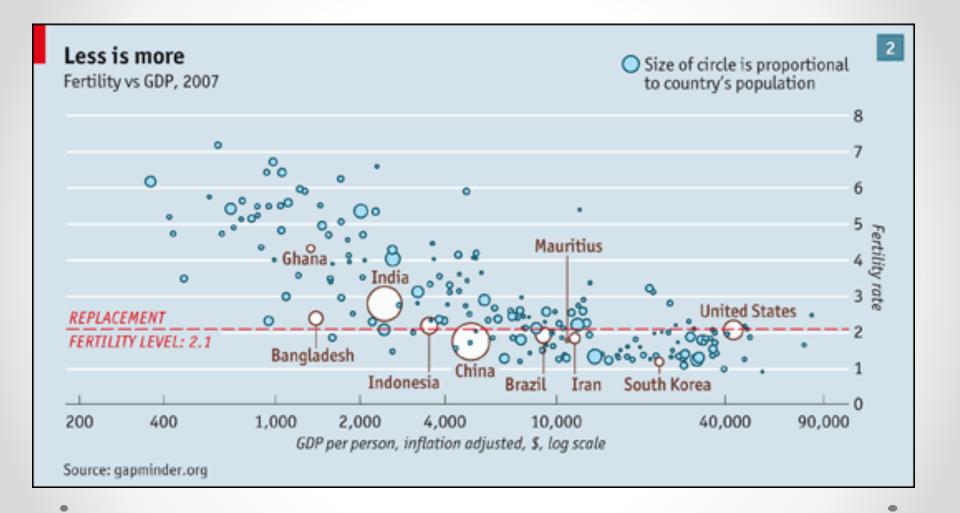
Attractiveness is a "confounding variable." The correlation between words and response rates is not necessarily causal.

Examples

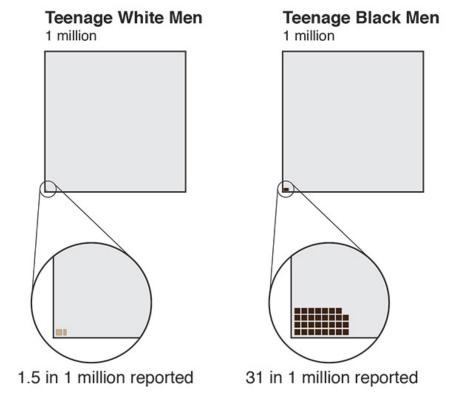








Odds ratio



Reported killings by police during 2010-2012

21 times as likely for black vs. white teenage men

Deadly Force in Black and White, ProPublica 10/10/2014

"Controlling" for a variable

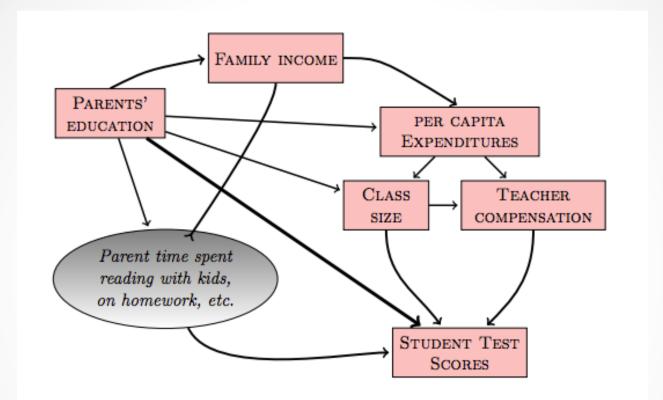
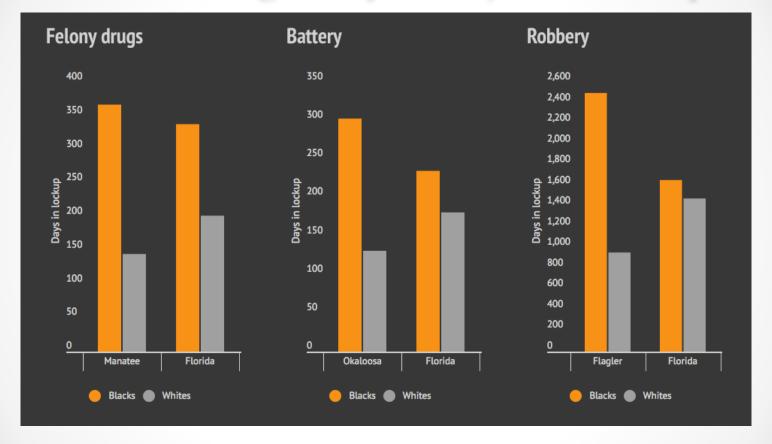


Figure 17.9: A hypothetical causal network relating student test scores to various educational policies and social conditions.

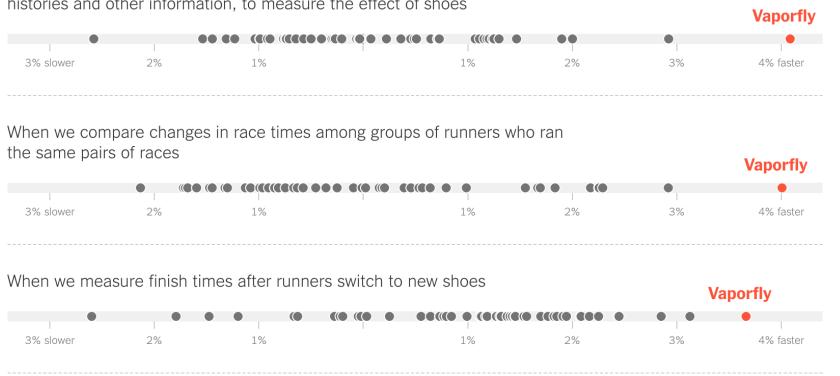
A causal network. From Statistical Modeling: A Fresh Approach

Florida sentencing analysis adjusted for "points"





When we use a statistical model, based on runners' ages, genders, race histories and other information, to measure the effect of shoes



Nike Says Its \$250 Running Shoes Will Make You Run Much Faster, New York Times

Knee Replacement

Replace diseased knee joint with an artificial knee. The most common reason for a knee replacement is osteoarthritis, which is a breakdown of the cartilage in the joint. More information \$\mathbb{G}\$

WILLIAM MCGANN»

PERFORMED PROCEDURE	COMPLICATIONS	RAW COMPLICATION RATE	ADJUSTED COMPLICATION RATE	
			Low Medium High Adjusted Rate of Complications	
166 times	0	0		
			1.6%	

EUGENE WOLF »						
PERFORMED PROCEDURE	COMPLICATIONS	RAW COMPLICATION RATE	ADJUST	ADJUSTED COMPLICATION RATE		
			Low	Medium	High Adjusted Rate of Complications	
83 times	1-10	Redacted		_		
				2.1%		

WILLIAM MONTGOMERY »

PERFORMED PROCEDURE	COMPLICATIONS	RAW COMPLICATION RATE	ADJUSTED COMPLICATION RATE			
			Low	Medium	High Adjusted Rate of Complications	
79 times	1-10	Redacted				
				2.3%		

MICHAEL RIES»

PERFORMED PROCEDURE	COMPLICATIONS	RAW COMPLICATION RATE	ADJUSTED COMPLICATION RATE		
	Low	Medium	High Adjusted Rate of Complications		
91 times	1-10	Redacted			
		2.4%			

Observation vs. Experiments

■1995 ■2010 Female Assistant Professor Salaries as a Proportion of Male Salaries in 1995 and 2010 1.10 1 1.05 1.00 -.95 -.90 -.85 -.80

Math and

Computer Science Physical Sciences Life Science

Social Sciences

Psychology

.75 -

.70 4

Geoscience

Economics

Engineering

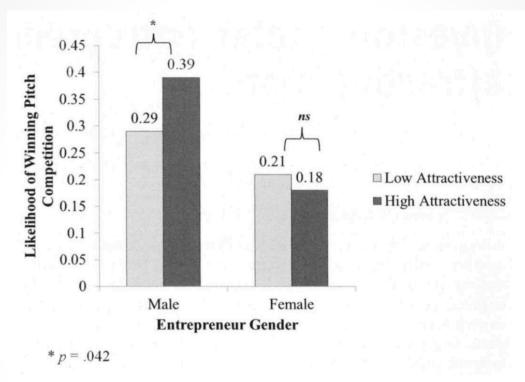


Fig. 1. The effect of entrepreneur gender and physical attractiveness on pitch success rate in a field setting (n = 90). ns, not significant.

Investors prefer entrepreneurial ventures pitched by attractive men,

Brooks et. al. 2014

Does Facebook news feed cause people to share links?

