

Introduction

- 1 The topic
- 2 The course
- 3 Notes

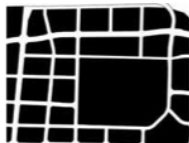
Reality is predictable

Los Angeles Times | ARTICLE COLLECTIONS

Stopping crime before it starts

Sophisticated analysis of data can sometimes tell police where criminals are headed. It's academic now, but the LAPD plans to get involved.

Reality is visualizable



MISSISSAUGA



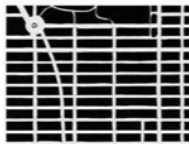
BARCELONA



COPENHAGEN



LONDON



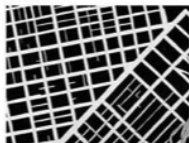
NEW YORK



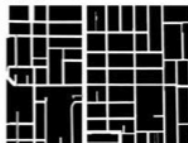
PARIS



ROME



SAN FRANCISCO



TORONTO

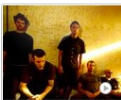
Reality is multidimensional



Bad Religion (317 plays)



David Bowie (220 plays)



Isis (202 plays)



Horace Andy (178 plays)



Army of the Pharaohs
(177 plays)



Biosphere (168 plays)



Bibio (162 plays)



Antonio Vivaldi
(143 plays)



Neil Young (125 plays)



King Crimson
(116 plays)



H.P. Lovecraft
(116 plays)



Virgin Prunes
(115 plays)



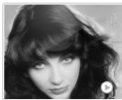
Motorama (109 plays)



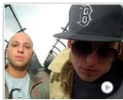
Wax Tailor (104 plays)



Lou Reed (103 plays)



Kate Bush (102 plays)

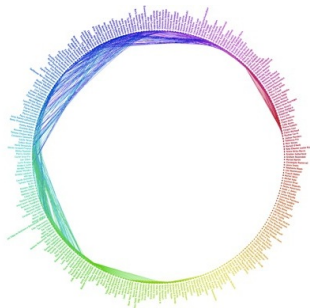


7L & Esoteric
(102 plays)



Gonzales (100 plays)

Reality is relational



Friendship ties on Facebook

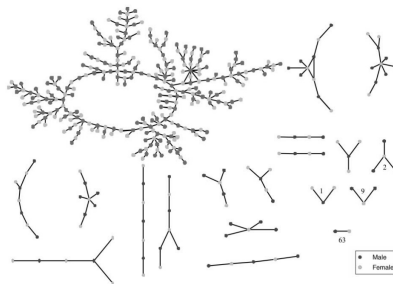


FIG. 2.—The direct relationship structure at Jefferson High

Sexual ties in high school

Data stand as professional assets

OECD Health Data 2010: Statistics and Indicators

AVAILABLE NOW - October 21st - [Internet update for OECD Health Data 2010](#)

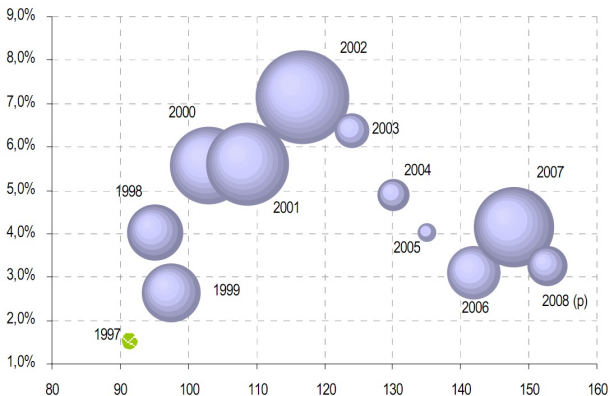
OECD Health Data 2010, released on 29 June 2010, offers the most comprehensive source of comparable statistics on health and health systems across OECD countries. It is an essential tool for health researchers and policy advisors in governments, the private sector and the academic community, to carry out comparative analyses and draw lessons from international comparisons of diverse health care systems.

- [What is OECD Health Data 2010](#)

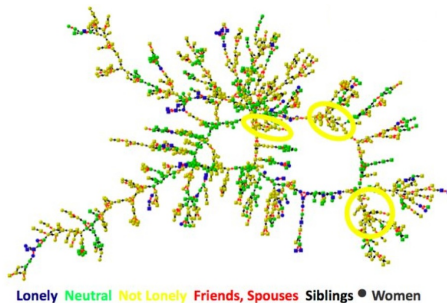


Data stand as policy expertise

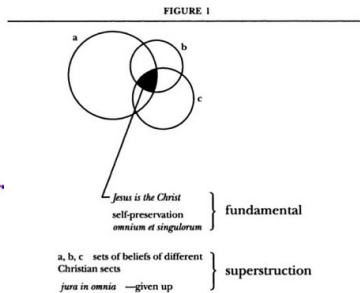
Graphique 1 – Vue d'ensemble de l'ONDAM



Interpretation is key to all analysis



Loneliness in social networks



Sets of Christian beliefs

Interpretation is difficult

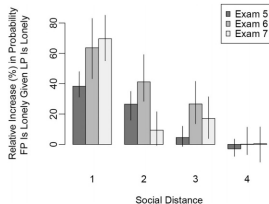
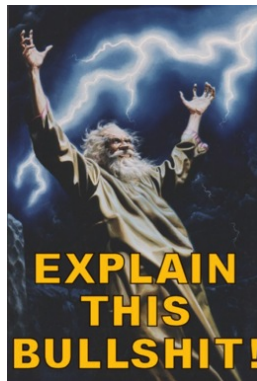


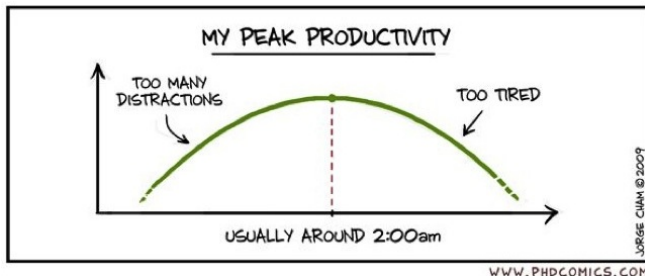
Figure 2. Social distance and loneliness in the Framingham Social Network. This figure shows for each exam the percentage increase in the likelihood a given focal participant (FP) is lonely if a friend or family member at a certain social distance is lonely (where lonely is defined as feeling lonely more than once a week). The relationship is strongest between individuals who are directly connected, but it remains significantly greater than zero at social distances up to three degrees of separation, meaning that a person's loneliness is associated with the loneliness of people up to three degrees removed from them in the network. Values are derived by comparing the conditional probability of being lonely in the observed network with an identical network (with topology and incidence of loneliness preserved) in which the same number of lonely participants are randomly distributed. Linked participant (LP) social distance refers to closest social distance between the LP and FP (LP = Distance 1, LP's LP = Distance 2, etc.). Error bars show 95% confidence intervals.

With explanation



Without explanation

Interpretation is what this course is eventually about



Questions

- What is the **measurement** method for each axis?
- What is the **probability** of 2am being the cutoff point?
- What is the **shape** of the time/productivity relationship?

Bias and measurement

Observational data

- Survey design
- Sampling strategy
- Question wording

Official statistics

- Unreliable aggregates
- Low statistical capacity

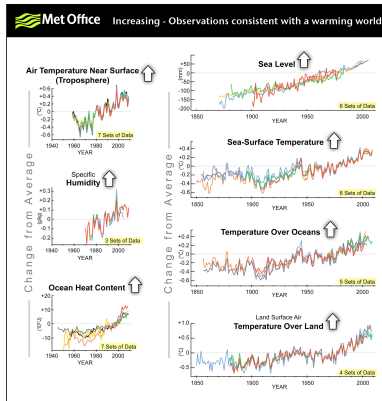
Bias and manipulation

- Media coverage
- Political spin
- Policy implications

Add to that:

- scientific fraud,
- data ethics,
- doubt-mongering,
- publishing bias,

...



The weirdest people in the world?

“The findings suggest that members of Western, Educated, Industrialized, Rich, and Democratic (WEIRD) societies, including young children, are among the least representative populations one could find for generalizing about humans.”

Joseph Henrich

Department of Psychology and Department of Economics, University of British Columbia, Vancouver V6T 1Z4, Canada

joseph.henrich@gmail.com

<http://www.psych.ubc.ca/~henrich/home.html>

Steven J. Heine

Department of Psychology, University of British Columbia, Vancouver V6T 1Z4, Canada

heine@psych.ubc.ca

Ara Norenzayan

Department of Psychology, University of British Columbia, Vancouver V6T 1Z4, Canada

ara@psych.ubc.ca



Course essentials

Core learning objectives

- 1 Data management
- 2 Statistical estimation
- 3 Regression modelling

Core teaching blocks

- | | |
|--------------------------|-----------------|
| ■ Statistical theory | ... readings |
| ■ Statistical computing | ... programming |
| ■ Tons of social science | ... research |

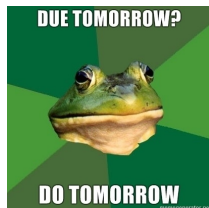
Course essentials

Requirements

- Attendance
- Homework
- No plagiarism

Grading

- Code and paper
- Draft, revised, final
- Project management



This won't work.

Homework

Readings

- | | |
|------------------------|-------------------------------|
| ■ Feinstein and Thomas | ... recommended textbook |
| ■ Urdan | ... shorter, less explanatory |
| ■ Stata Guide | ... practical walkthrough |

Coursework

- Replicate the weekly do-files
- Code your own data analysis
- Write an empirical research paper

Logistics

Elect a student representative!

No estimation without representation. One (wo)man, one vote.

Any questions so far?

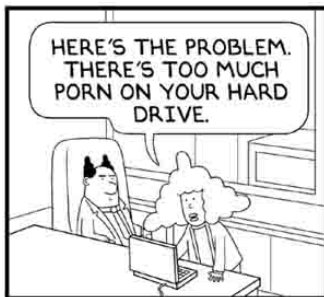
Do not worry about deadlines, these will be discussed in class.



A note on computers

Despite what the general computing industry is telling you, computers are more than

- Music players
- Facebook terminals
- Porn stashes



Important

This course requires that you learn how to work with a computer.

Program or be programmed.

A note on slides

The course slides cover logistics and offer basic guidance. They are absolutely insufficient to complete the course requirements.



Copyright © 2000 United Feature Syndicate, Inc.
Redistribution in whole or in part prohibited

A note on emails

Rules

- 1 Start your email subject line with the “SRQM:” prefix.
- 2 Describe the content of the email in the subject line.
- 3 Attach your code and results to your question(s).

Examples

*“SRQM: Help with regression results”
(include the puzzling output)*

*“SRQM: Problem with plotting ordinal data”
(include the faulty code)*

Welcome on board!

