

Ronald Reagan was a statistician
and other examples
of learning from diverse sources of information

Andrew Gelman
Department of Statistics, Columbia University, New York

16 décembre 2013

“Are you better off than you were four years ago?”



“But the simplest way for each of you to judge recovery lies in the plain facts of your own individual situation. Are you better off than you were last year?”

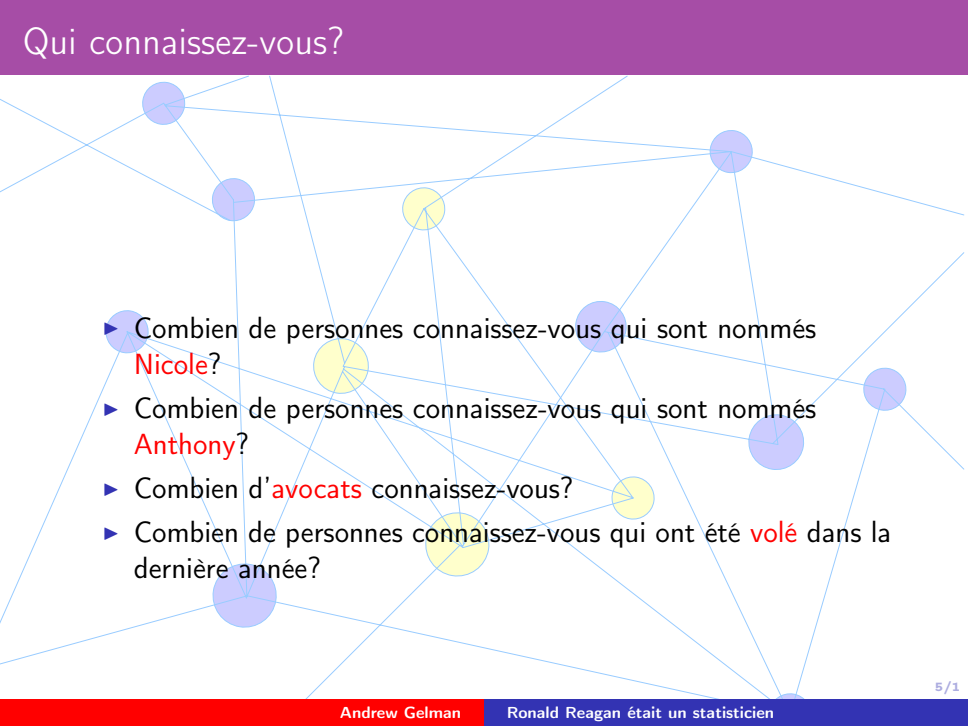




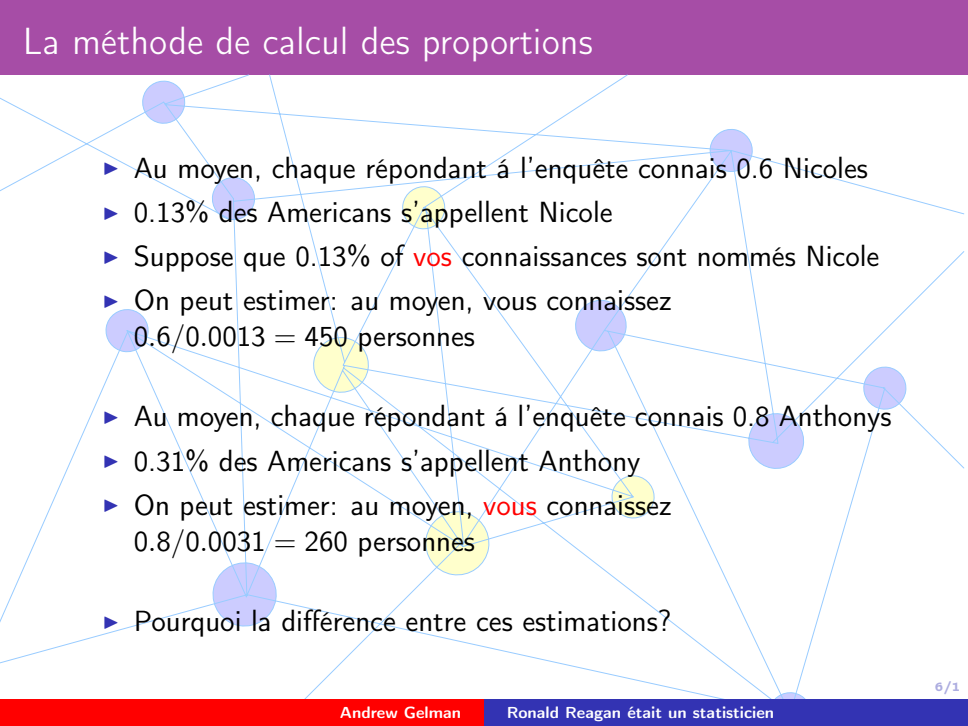
Andrew Gelman

Ronald Reagan était un statisticien

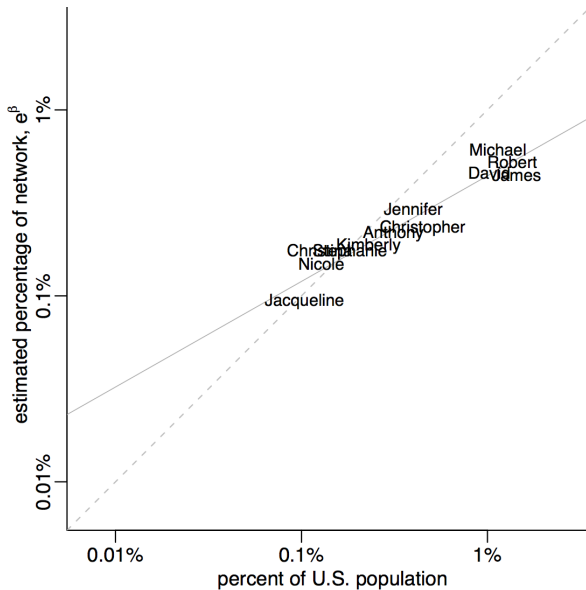
Qui connaissez-vous?

- 
- ▶ Combien de personnes connaissez-vous qui sont nommés **Nicole**?
 - ▶ Combien de personnes connaissez-vous qui sont nommés **Anthony**?
 - ▶ Combien d'**avocats** connaissez-vous?
 - ▶ Combien de personnes connaissez-vous qui ont été **volé** dans la dernière année?

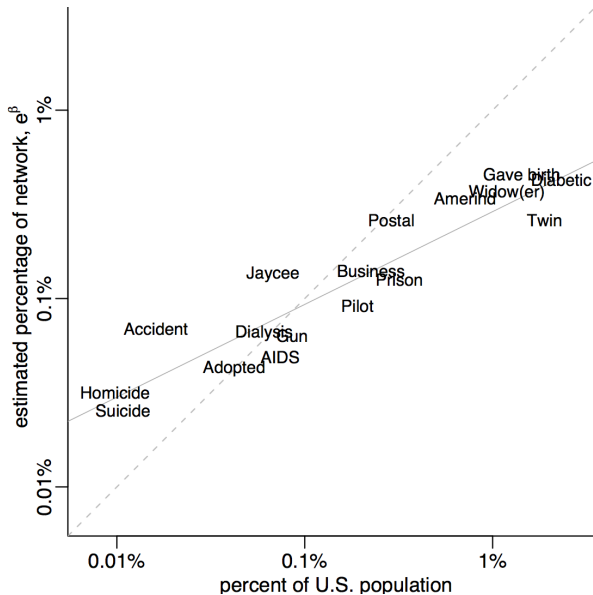
La méthode de calcul des proportions

- 
- ▶ Au moyen, chaque répondant à l'enquête connaît 0.6 Nicoles
 - ▶ 0.13% des Américains s'appellent Nicole
 - ▶ Suppose que 0.13% of **vos** connaissances sont nommés Nicole
 - ▶ On peut estimer: au moyen, vous connaissez $0.6/0.0013 = 450$ personnes
 - ▶ Au moyen, chaque répondant à l'enquête connaît 0.8 Anthony
 - ▶ 0.31% des Américains s'appellent Anthony
 - ▶ On peut estimer: au moyen, **vous** connaissez $0.8/0.0031 = 260$ personnes
 - ▶ Pourquoi la différence entre ces estimations?

La perception et la réalité: Noms

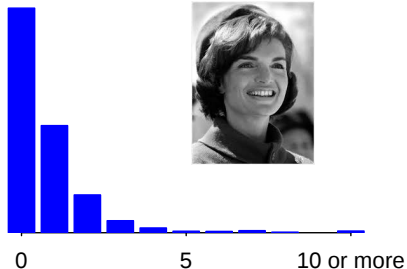


La perception et la réalité: Groupes

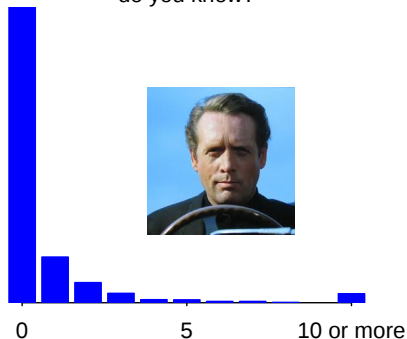


La structure sociale!

How many Jacquelines
do you know?



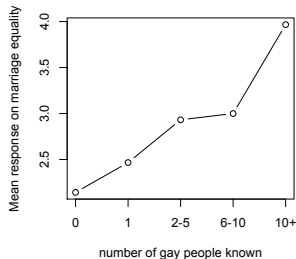
How many men in prison
do you know?



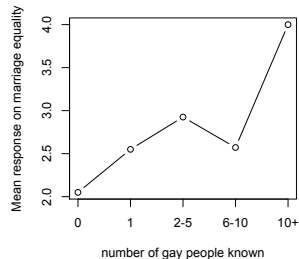
La pénombre des gays



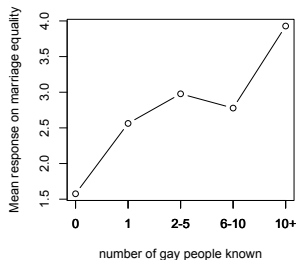
Marriage Support: Everyone



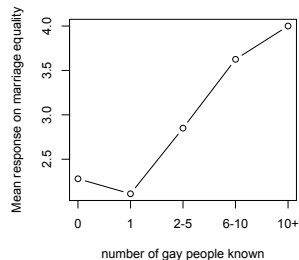
Marriage Support: Democrats



Marriage Support: Republicans



Marriage Support: Independents

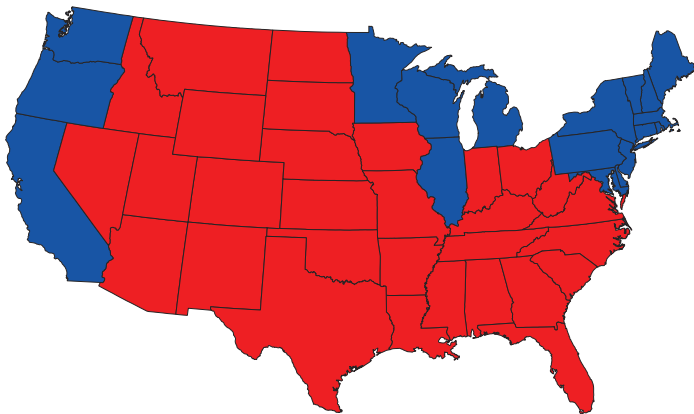


Comment est-ce arrivé?

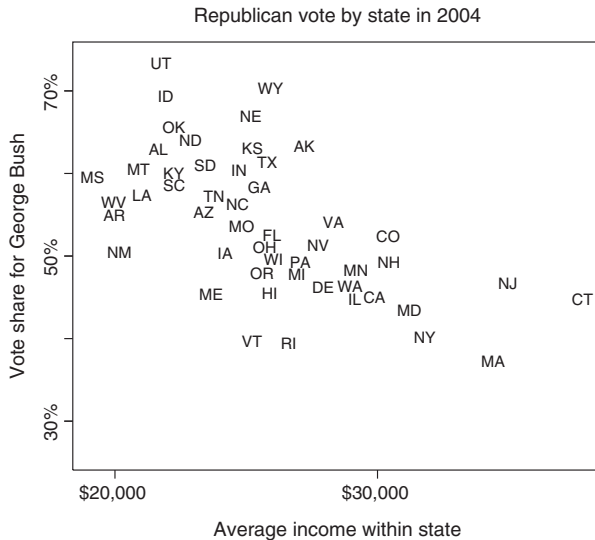


Cela n'a aucun sens!

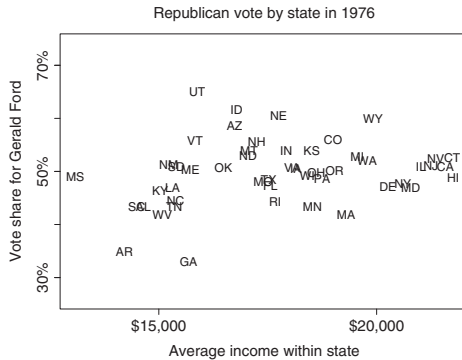
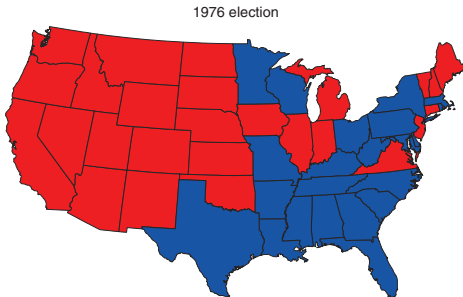
2004 election



C'est pas illusion

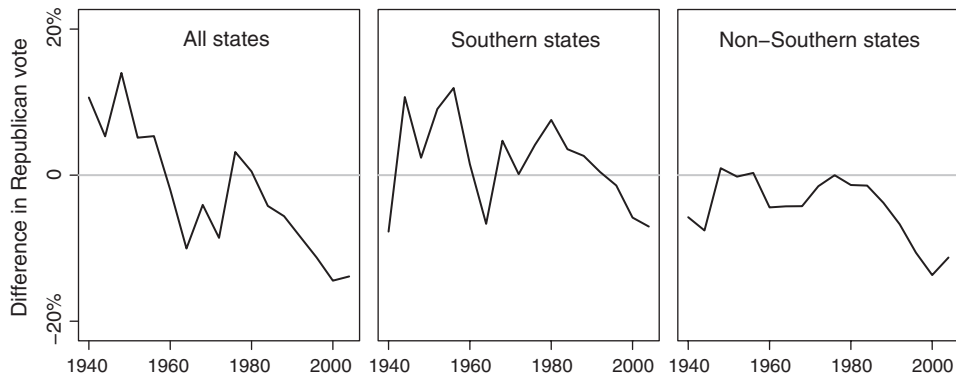


Qu'est-ce qui a changé?



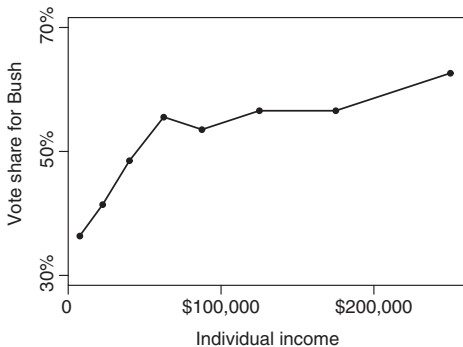
C'était pas seulement les états du sud

Republican vote in rich states minus Republican vote in poor states

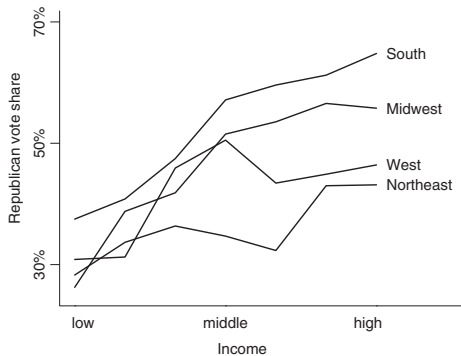


Les riches continue votent pour la droite

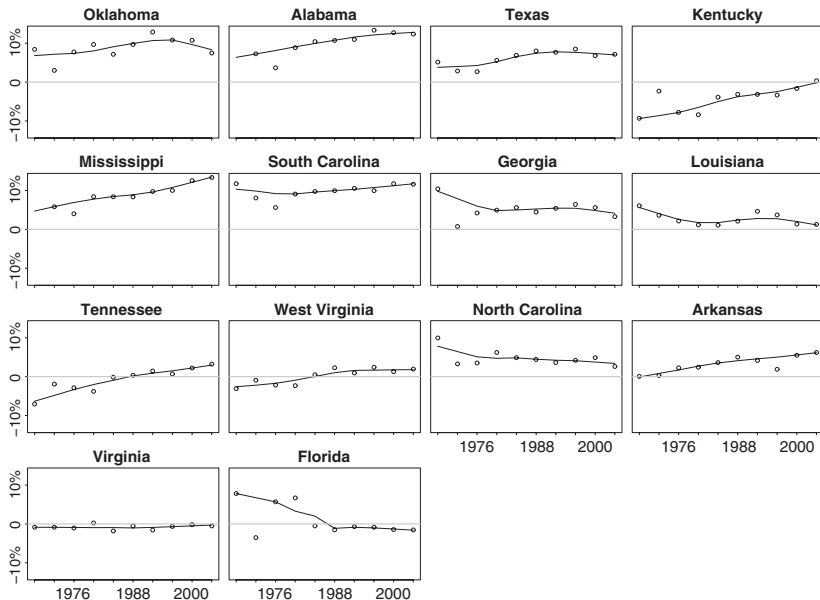
Bush vote in 2004 by income



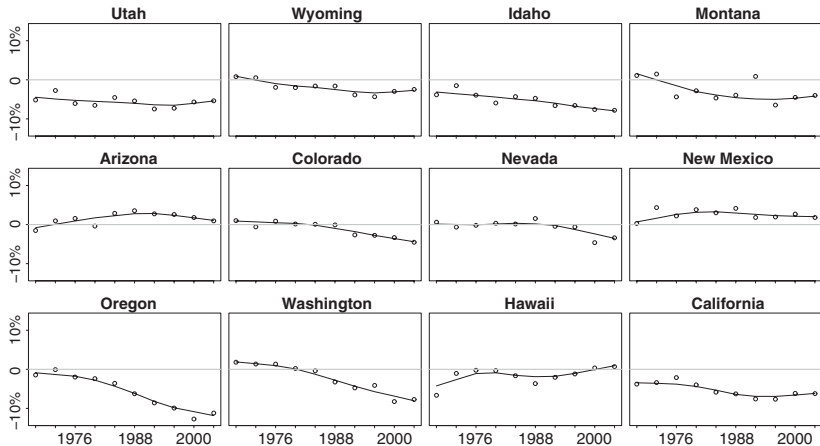
2006 House exit polls



Les comtés riche votent plus pour la droite ...



... Mais pas dans l'ouest!



“Opium du peuple” vs. “Post-matérialisme”

- ▶ “I don’t know that atheists should be considered citizens, nor should they be considered patriots. This is one nation under God.” — George H. W. Bush

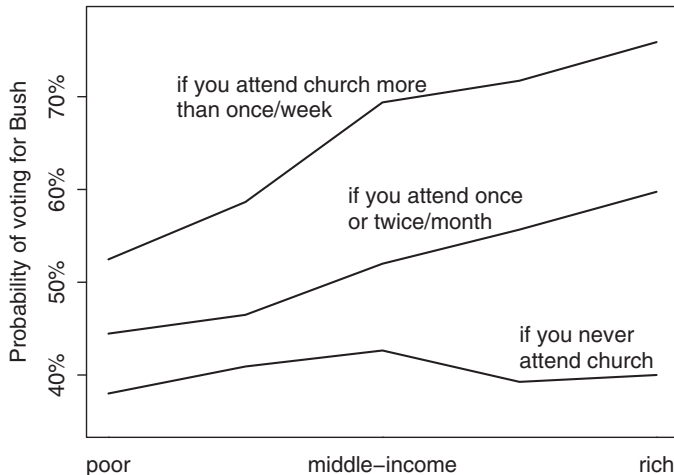


- ▶ “Very few of us . . . could name even five NASCAR drivers, although stock-car races are the best-attended sporting events in the country.” — David Brooks
- ▶ Two theories:
 - ▶ *Opium du peuple*: Les riches votent pour leurs intérêts, les pauvres vote en raison de “Gods, guns, and gays”
 - ▶ *Post-matérialisme*: Les pauvres vote suivant l’économie, les riches ont le luxe de voter en raison des concerns sociales

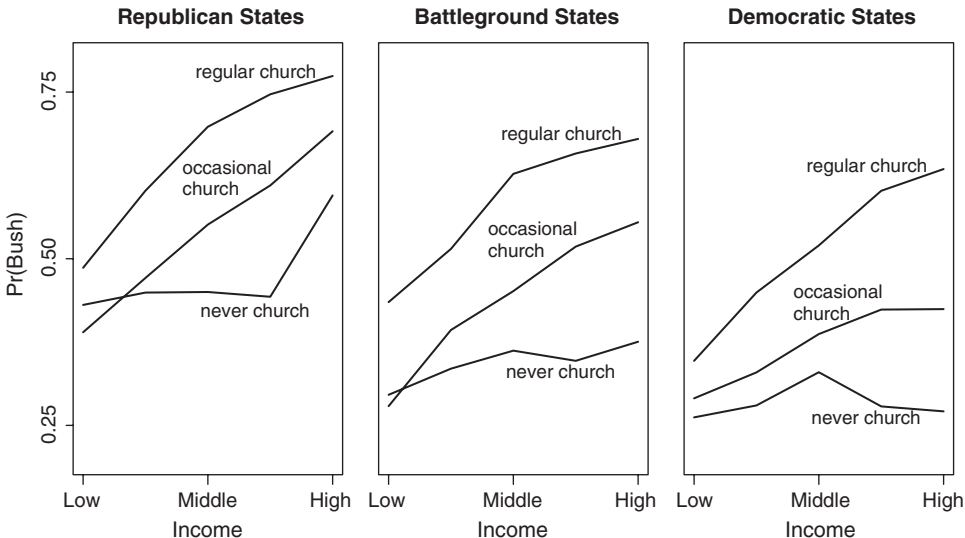


“Opium” non, “Post-matérialisme” oui

Bush vote in 2004 by income and religious attendance



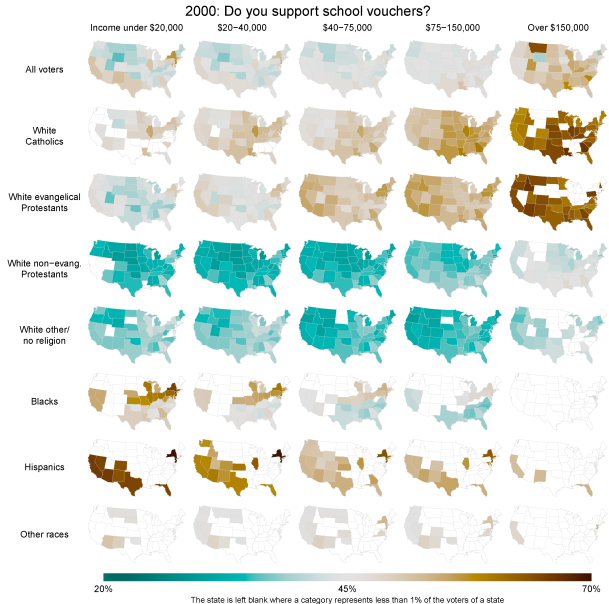
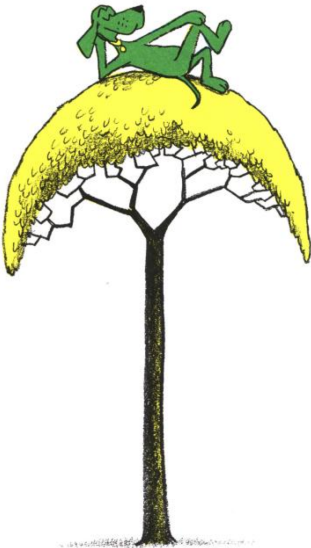
Des tendances similaires dans les régions différents



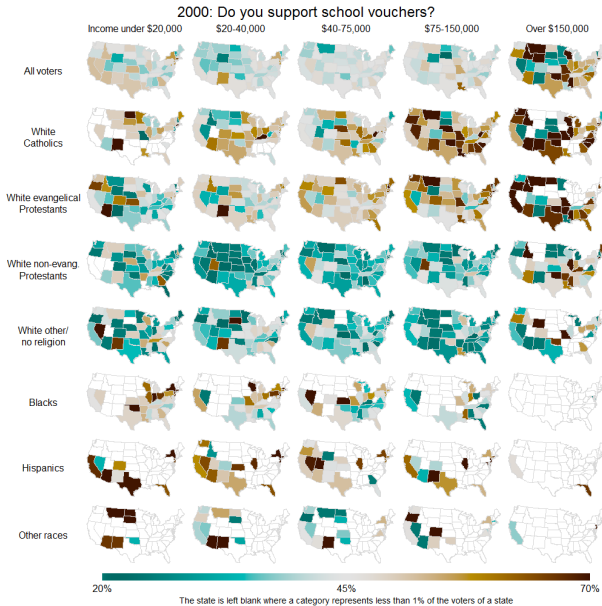
Grand modèle, données clairsemée



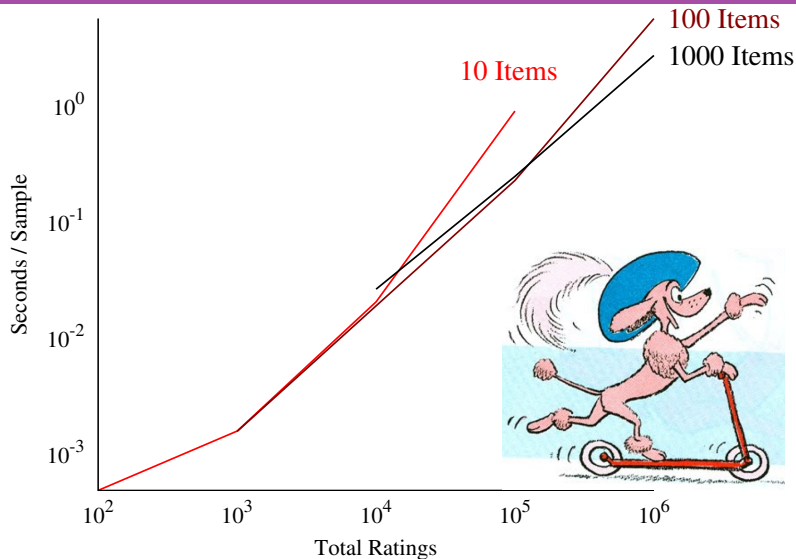
«Grande données» et grand modèle



Les données ne sont pas toujours «parlent d'eux-mêmes»



Grande données, grand modèle, l'informatique pour lesquels on peut changer l'échelle



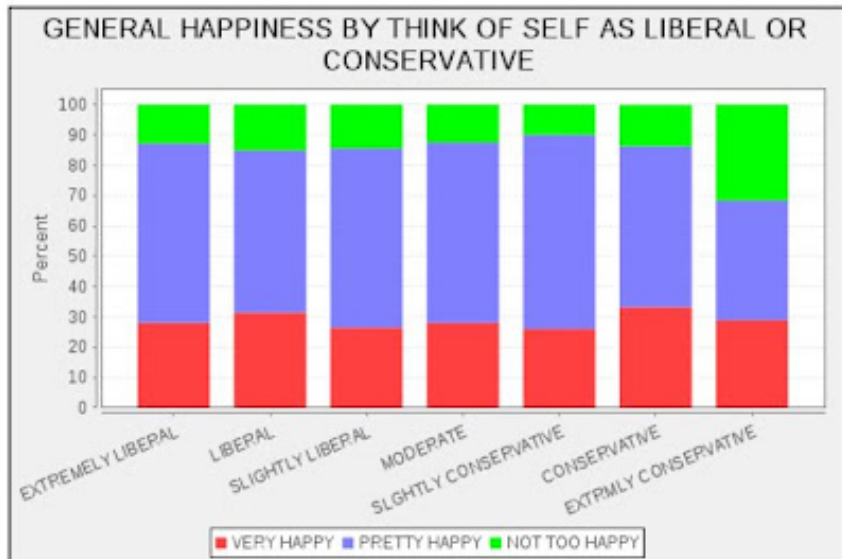
Arthur «pas David» Brooks dans le *New York Times*:

“People at the extremes are happier than political moderates. . . . none, it seems, are happier than the Tea Partiers . . .”

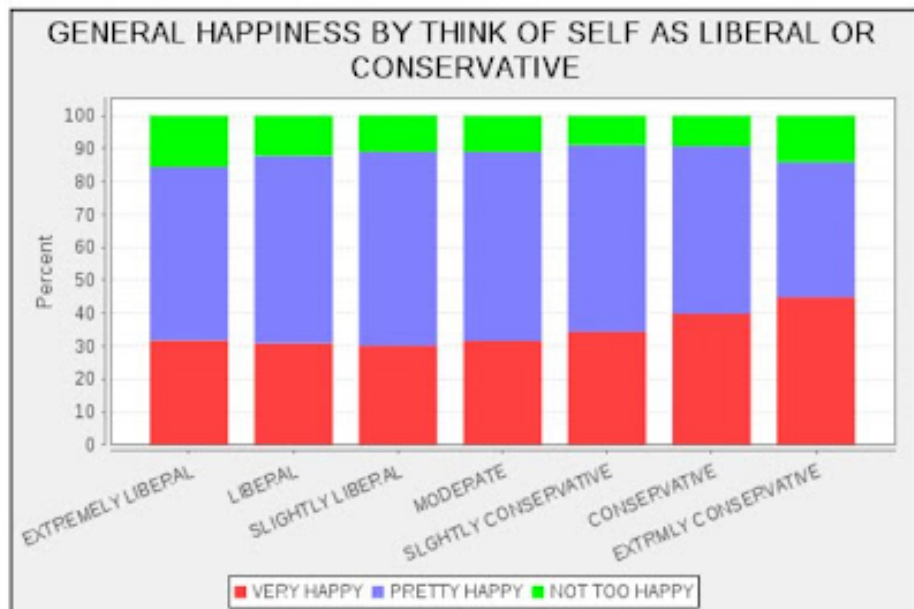
Jay Livingston (sociology, Montclair State University) regarde les données de la «General Social Survey» . . .

“Aucun, paraît-il, sont plus heureux que les Tea Partiers” ??

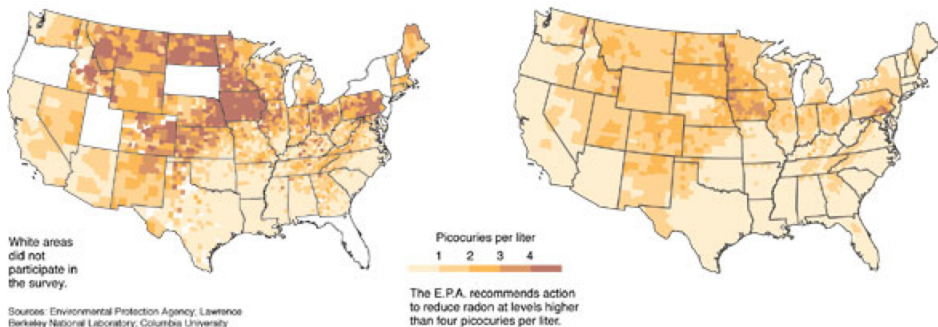
Chart for YEAR = 4(2009-2010)



Mettre les données en commun, 1972–2010



Les données locales, la coordination nationale, les décisions locales



The New York Times

Older Model, left: Observed indoor radon levels, based on short-term measurements collected by the E.P.A. and most states. The measurements, taken in basements in the winter, may overestimate true exposure. **New Model, right:** The predicted average long-term levels of radon that people may actually breathe in their homes.