Classroom activities in applied regression and causal inference, first semester[1]

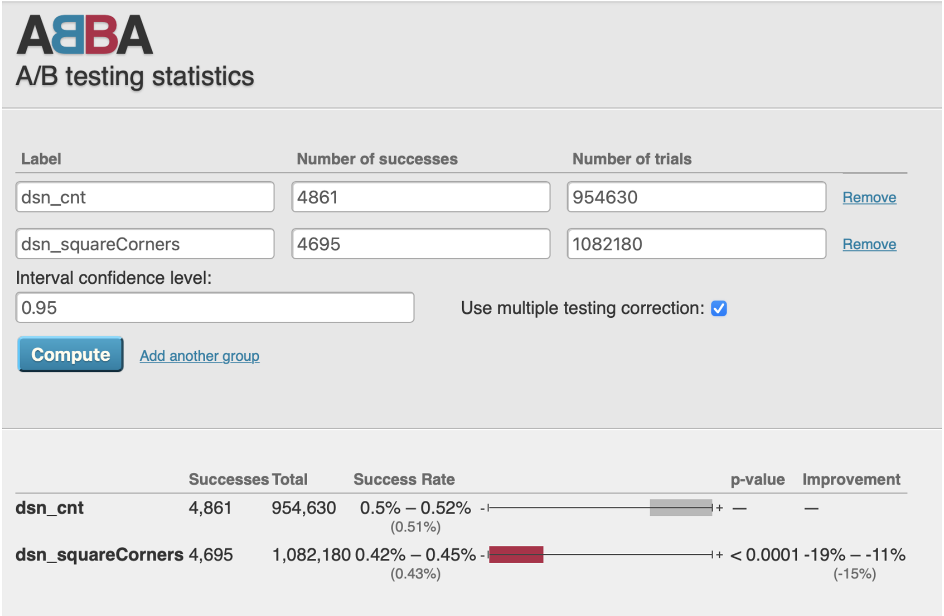
Andrew Gelman and Aki Vehtari

[1] For details, see chapter 3 of *Active Statistics*, by Andrew Gelman and Aki Vehtari, Cambridge University Press (2023), `http://www.stat.columbia.edu/~gelman/active-statistics/`.

Class 1a

Story

Wikipedia experiment



ABBA

A/B testing statistics

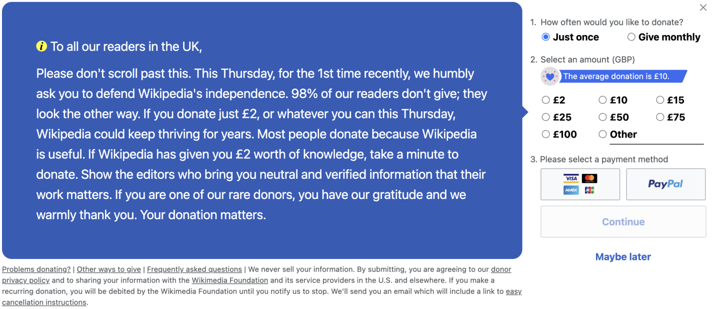
Label	Number of successes	Number of trials	
dsn_cnt	4861	954630	Remove
dsn_squareCorners	4695	1082180	Remove

Interval confidence level: 0.95

Use multiple testing correction: ☑

	Successes	Total	Success Rate	p-value	Improvement
dsn_cnt	4,861	954,630	0.5% – 0.52% (0.51%)	—	—
dsn_squareCorners	4,695	1,082,180	0.42% – 0.45% (0.43%)	< 0.0001	-19% – -11% (-15%)

Wikipedia experiment

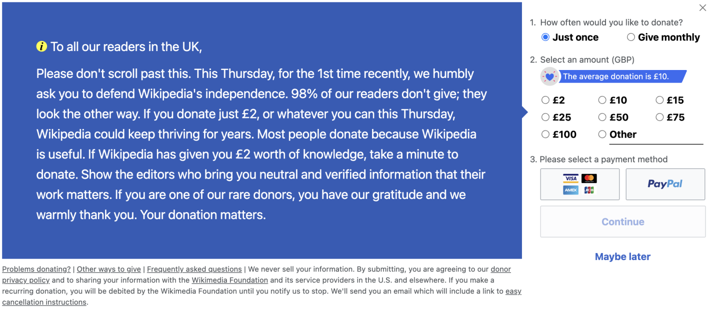


To all our readers in the UK,

Please don't scroll past this. This Thursday, for the 1st time recently, we humbly ask you to defend Wikipedia's independence. 98% of our readers don't give; they look the other way. If you donate just £2, or whatever you can this Thursday, Wikipedia could keep thriving for years. Most people donate because Wikipedia is useful. If Wikipedia has given you £2 worth of knowledge, take a minute to donate. Show the editors who bring you neutral and verified information that their work matters. If you are one of our rare donors, you have our gratitude and we warmly thank you. Your donation matters.

1. How often would you like to donate?
   ● Just once ○ Give monthly
2. Select an amount (GBP)
   The average donation is £10.
   ○ £2 ○ £10 ○ £15
   ○ £25 ○ £50 ○ £75
   ○ £100 ○ Other
3. Please select a payment method

Problems donating? | Other ways to give | Frequently asked questions | We never sell your information. By submitting, you are agreeing to our donor privacy policy and to sharing your information with the Wikimedia Foundation and its service providers in the U.S. and elsewhere. If you make a recurring donation, you will be debited by the Wikimedia Foundation until you notify us to stop. We'll send you an email which will include a link to easy cancellation instructions.



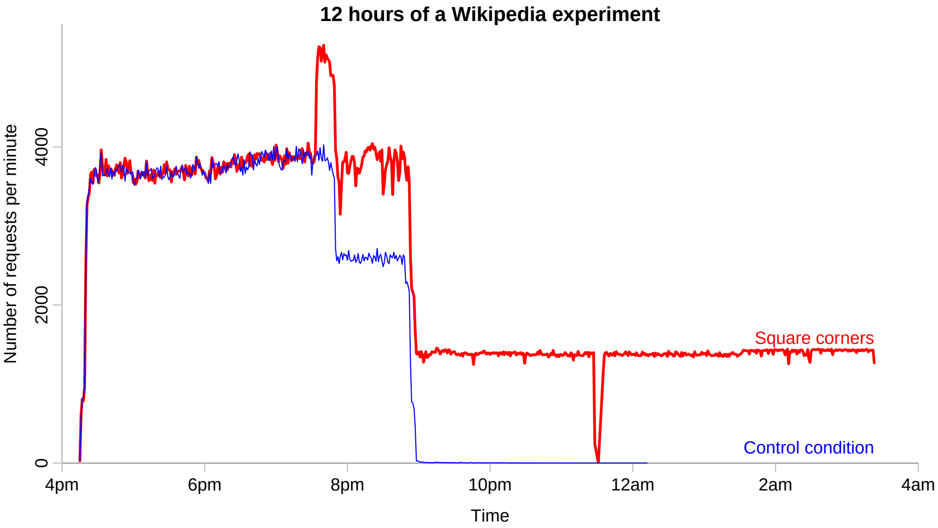
To all our readers in the UK,

Please don't scroll past this. This Thursday, for the 1st time recently, we humbly ask you to defend Wikipedia's independence. 98% of our readers don't give; they look the other way. If you donate just £2, or whatever you can this Thursday, Wikipedia could keep thriving for years. Most people donate because Wikipedia is useful. If Wikipedia has given you £2 worth of knowledge, take a minute to donate. Show the editors who bring you neutral and verified information that their work matters. If you are one of our rare donors, you have our gratitude and we warmly thank you. Your donation matters.

1. How often would you like to donate?
   ● Just once ○ Give monthly
2. Select an amount (GBP)
   The average donation is £10.
   ○ £2 ○ £10 ○ £15
   ○ £25 ○ £50 ○ £75
   ○ £100 ○ Other
3. Please select a payment method

Problems donating? | Other ways to give | Frequently asked questions | We never sell your information. By submitting, you are agreeing to our donor privacy policy and to sharing your information with the Wikimedia Foundation and its service providers in the U.S. and elsewhere. If you make a recurring donation, you will be debited by the Wikimedia Foundation until you notify us to stop. We'll send you an email which will include a link to easy cancellation instructions.

Wikipedia experiment



Activity

Designing a social science study

1. Topic
2. Quantity to measure
3. Definition
4. Data
5. Data collection
6. Estimation
7. Specifics
8. Variation

Introduction to the course

Topics

- Goals of the course
- Components of the course
- Structure of each class period
- Students' responsibilities
- Roles of mathematics, computing, and applications

Computer demonstration

Drill

Designing a social science study

1. Treatments
2. Population
3. Sample
4. Treatment assignment
5. Pre-test measurement
6. Outcome measurement

Discussion problem

Finding the hidden assumption and error

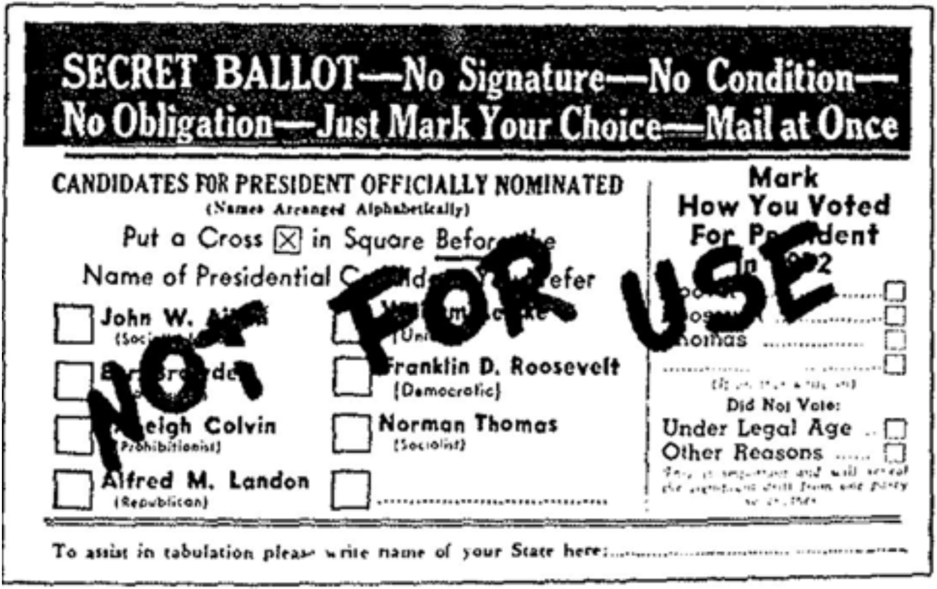
> *"Many theorists claim that domestic instability tends to lead to foreign aggression. Others have made the claim that domestic instability makes it less likely that a country will engage in an aggressive foreign policy. The posited linkages are obvious. Suppose you develop a good measure of both variables, and for each year you compute the total amount of domestic instability in all countries in the international system and correlate this with the total amount of external aggression by all states. You find no correlation at all and conclude that, contrary to both theories, there is no connection between domestic instability and war."*

What's wrong with this argument?

Class 1b

Story

Literary Digest poll



SECRET BALLOT—No Signature—No Condition—No Obligation—Just Mark Your Choice—Mail at Once

CANDIDATES FOR PRESIDENT OFFICIALLY NOMINATED
(Names Arranged Alphabetically)

Put a Cross ☒ in Square Before the Name of Presidential Candidate You Prefer

□ John W. Aiken (Socialist Labor)

□ Earl Browder (Communist)

□ D. Leigh Colvin (Prohibitionist)

□ Alfred M. Landon (Republican)

□ William Lemke (Union)

□ Franklin D. Roosevelt (Democratic)

□ Norman Thomas (Socialist)

□ ....

Mark How You Voted For President In 1932

Hoover .... □

Roosevelt .... □

Thomas .... □

.... □

[illegible]

Did Not Vote:

Under Legal Age .. □

Other Reasons ...... □

This is important and will reveal the significant shift from one party to another [illegible]

To assist in tabulation please write name of your State here:..

Activity

Designing an experiment

1. Topic
2. Two hypotheses
3. Ideal data that would establish hypothesis 1 or 2
4. Scenario of ambiguous data
5. Scenario of data consistent with neither hypothesis
6. Data collection and measurement
7. Inference
8. Specifics

Discuss reading and homework

Computer demonstration

Drill

Generalizing

1. From sample to population
2. From treatment to control group
3. From measurement to underlying construct

Discussion problem

Finding the hidden assumption and error

> *"There is a positive correlation between the per capita GDP of a country and the degree to which it is democratic. Therefore as poor countries get richer, they will also become more democratic."*

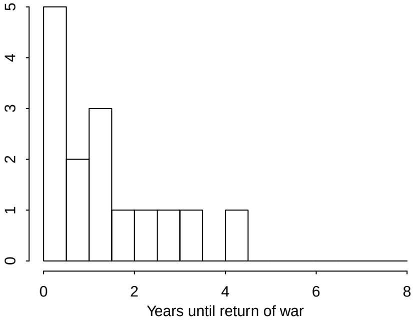
What's wrong with this argument?

Class 2a

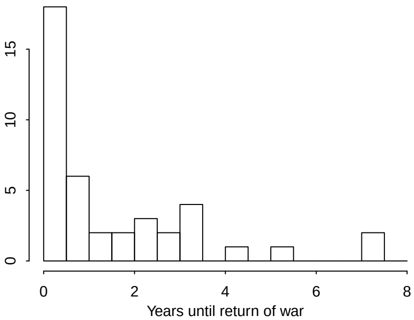
Story

United Nations peacekeeping

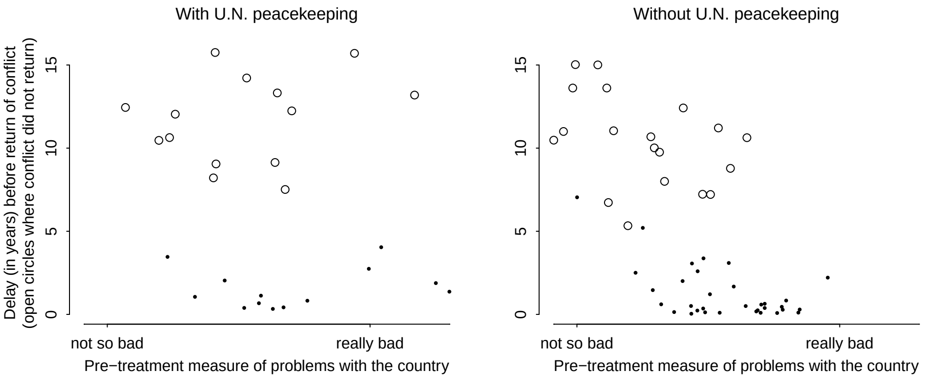
With peacekeeping: 56% of countries stayed at peace.
For others, histogram of time until civil war returned:



Without peacekeeping: 34% stayed at peace.
For others, histogram of time until civil war returned:



United Nations peacekeeping



Activity

Candy weighing

1. Pull 5 candies out of the bag
2. Weigh the candies
3. Write down the weight
4. Put the candies back in the bag!!
5. Pass the scale and bag to your neighbors
6. Silently multiply the weight of the 5 candies by 20

Discuss reading and homework

Computer demonstration

Drill

Describing a fitted regression in words

Explain the meaning of the underlined number, first wrongly and then correctly.

Drill

Simple coding: computing and graphing functions

Give R code.

Discussion problem

Height and earnings

$$\text{earnings} = -26000 + 600 * \text{height} + 10600 * \text{male} + \text{error}$$

Class 2b

Story

Girls and sports

A published claim: "Sports participation [in high school] causes women to be less likely to be religious . . . more likely to have children . . . more likely to be single mothers . . .

A ten percentage-point increase in state-level female sports participation generates a five to six percentage-point rise in the rate of female secularism, a five percentage-point increase in the proportion of women who are mothers, and a six percentage-point rise in the proportion of mothers who, at the time that they are interviewed, are single mothers."

Activity

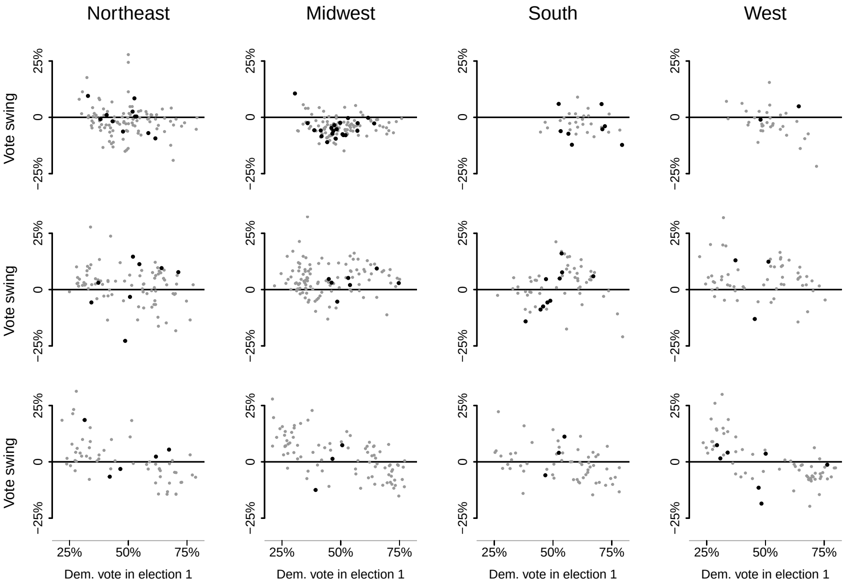
Measurements

What are a few things do you want to learn about each other? We will gather data and make scatterplots.

Discuss reading and homework

Computer demonstration

Graph a function of 4 variables using a grid



Drill

Simple coding: sampling, looping, and vectors

Give R code.

Discussion problem

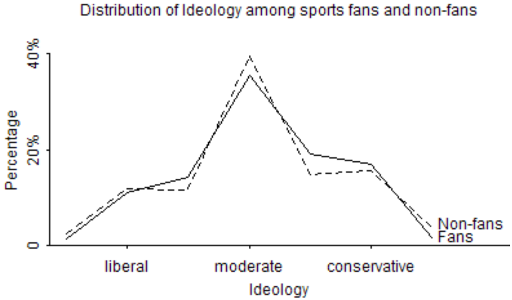
Graphing hypothetical data

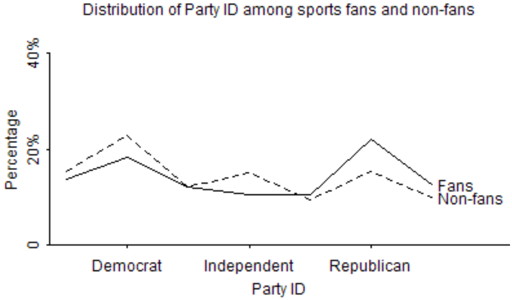
For each scenario, graph hypothetical data, plotting post-test vs. pre-test, using different symbols for treatment and control.

Class 3a

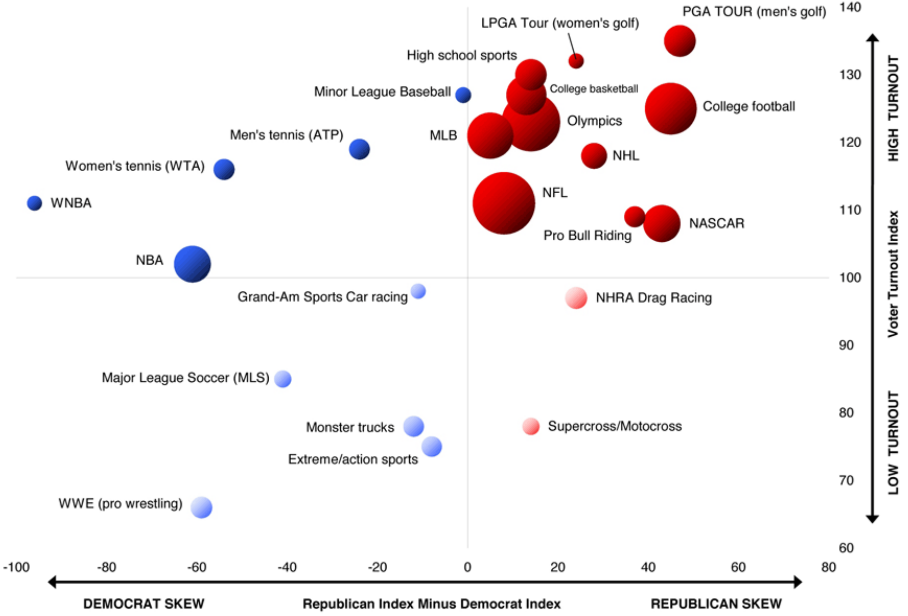
Story

Political leanings of sports fans





Political leanings of sports fans



Source: NMRPP analysis of Scarborough USA+ survey Aug. 2008 - Sept. 2009. N = 218,313 adults.

Activity

Measuring handedness

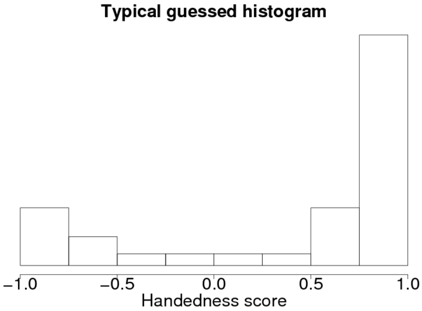
Please indicate which hand you use for each of the following activities by putting a + in the appropriate column, or ++ if you use would never use the other hand for that activity. If in any case you are really indifferent, put + in both columns.
Some of the activities require both hands. In these cases the part of the task, or object, for which hand preference is wanted is indicated in parentheses.

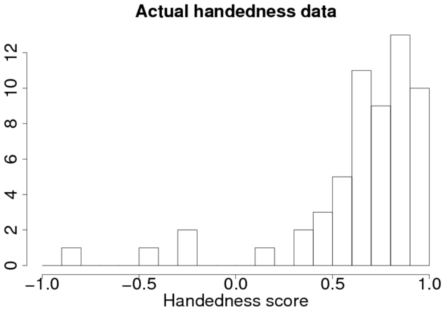
Task	Left	Right
Writing		
Drawing		
Throwing		
Scissors		
Toothbrush		
Spoon		
Total		

Right − Left: Right + Left: $\dfrac{\text{Right} - \text{Left}}{\text{Right} + \text{Left}}$:

Create a Left and a Right score by counting the total number of + signs in each column. Your handedness score is (Right − Left)/(Right + Left): thus, a pure right-hander will have a score of score $(12 - 0)/(12 + 0) = 1$, and a pure left-hander will score $(0 - 12)/(0 + 12) = -1$.

Measuring handedness


Typical guessed histogram


Actual handedness data

Measuring handedness

Please indicate which hand you use for each of the following activities by putting a + in the appropriate column, or ++ if you use would never use the other hand for that activity. If in any case you are really indifferent, put + in both columns.

Some of the activities require both hands. In these cases the part of the task, or object, for which hand preference is wanted is indicated in parentheses.

Task	Left	Right
Writing		
Drawing		
Throwing		
Scissors		
Toothbrush		
Knife (without fork)		
Spoon		
Broom (upper hand)		
Striking match (hand that holds the match)		
Opening box (hand that holds the lid)		
Total		

Right − Left: Right + Left: $\dfrac{\text{Right} - \text{Left}}{\text{Right} + \text{Left}}$:

Create a Left and a Right score by counting the total number of + signs in each column. Your handedness score is (Right − Left)/(Right + Left): thus, a pure right-hander will have a score of score $(20 - 0)/(20 + 0) = 1$, and a pure left-hander will score $(0 - 20)/(0 + 20) = -1$.

Discuss reading and homework

Computer demonstration

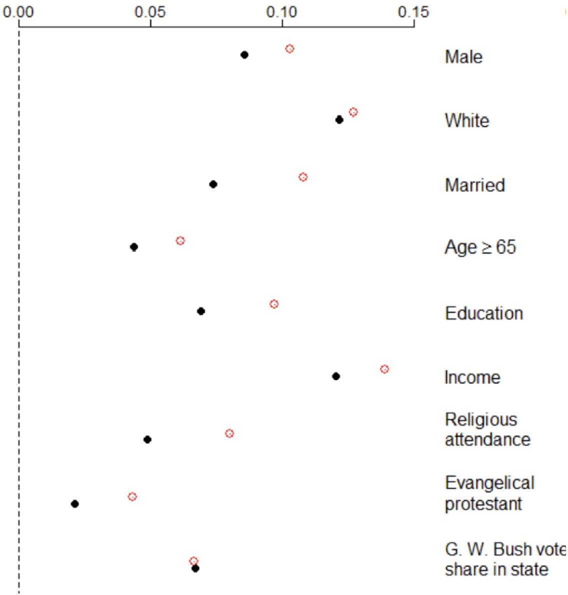
Drill

All graphs are comparisons

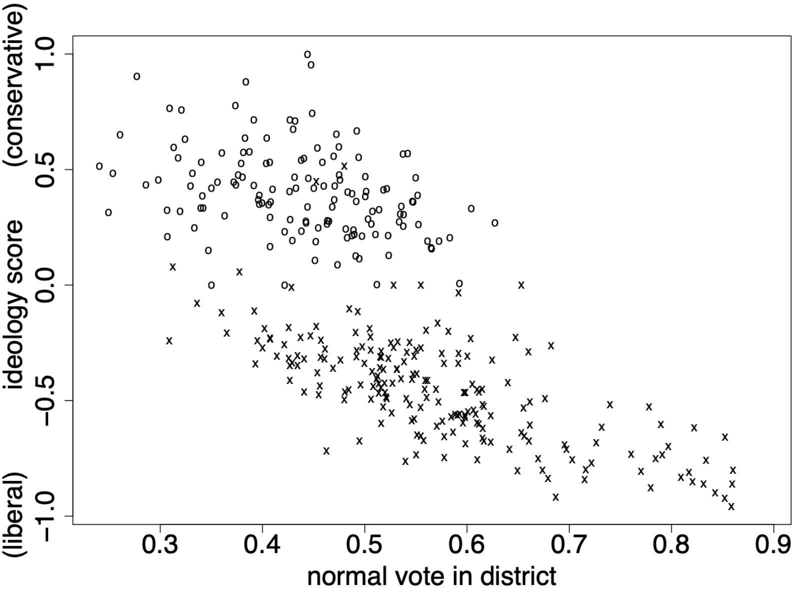
Identify the implicit or explicit comparison that is facilitated by each graph.

All graphs are comparisons

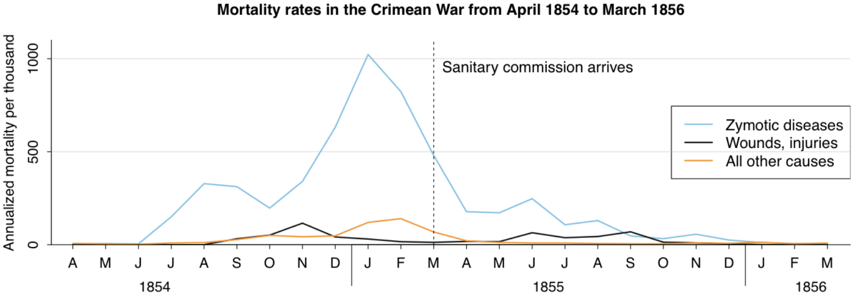
Correlation of opposition to health care reform with...



All graphs are comparisons



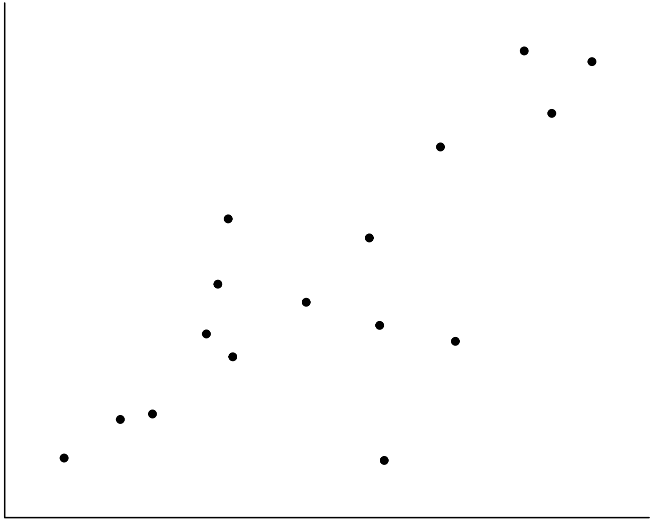
All graphs are comparisons



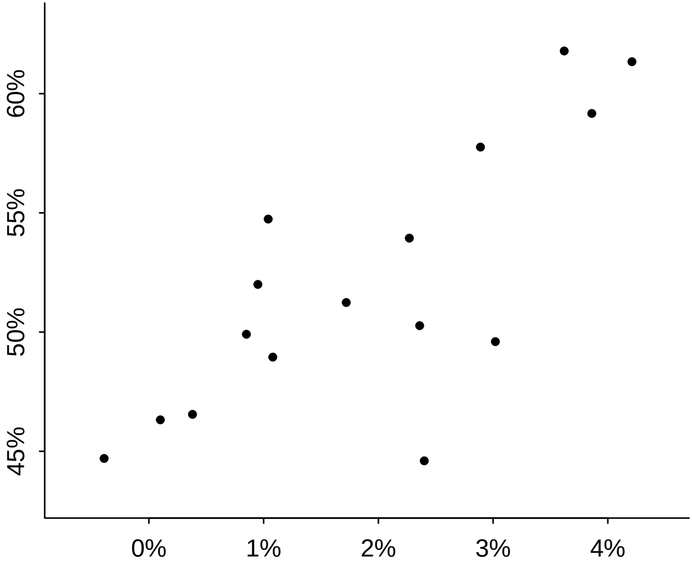
Discussion problem

Activity for next class

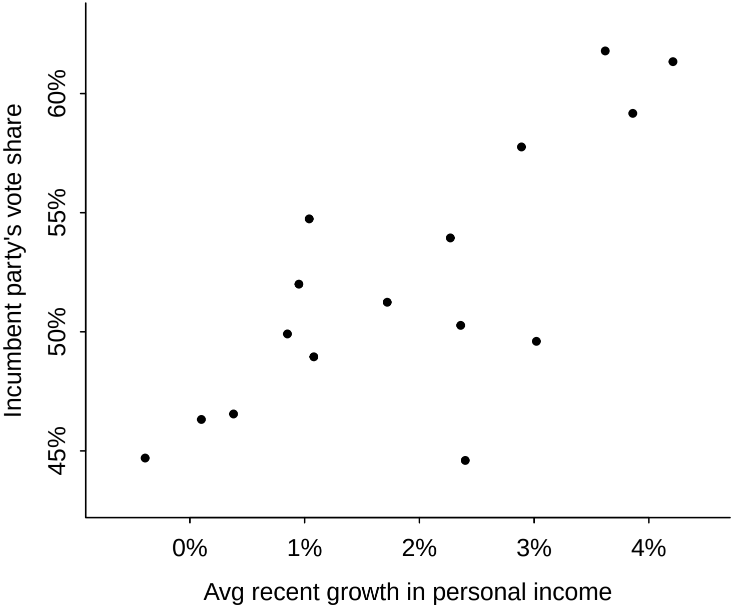
Scatterplot charades



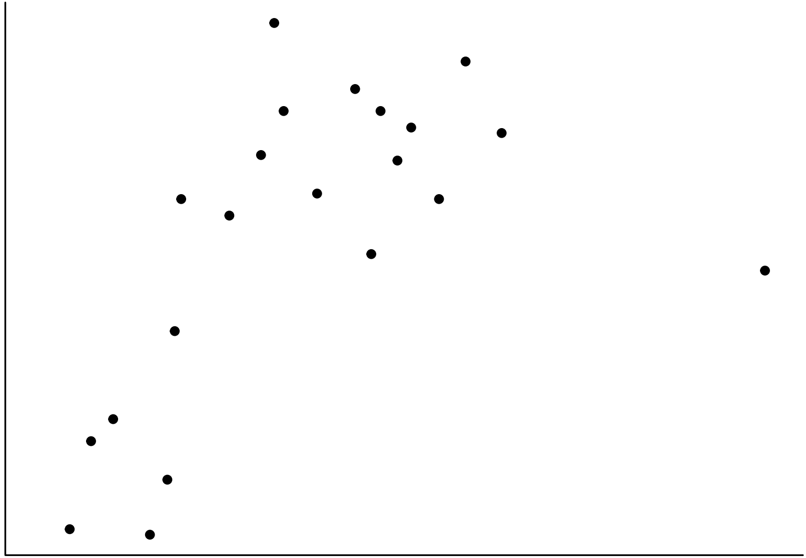
Scatterplot charades



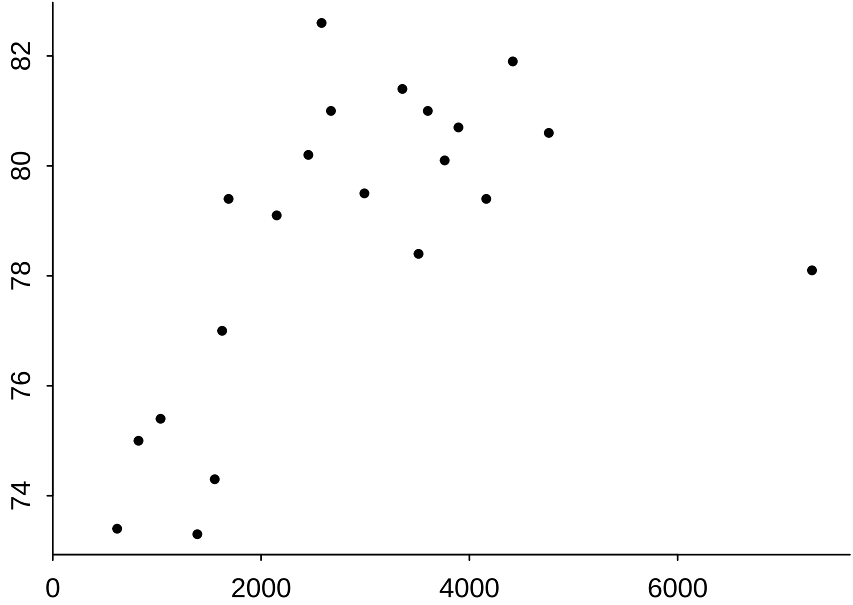
Scatterplot charades



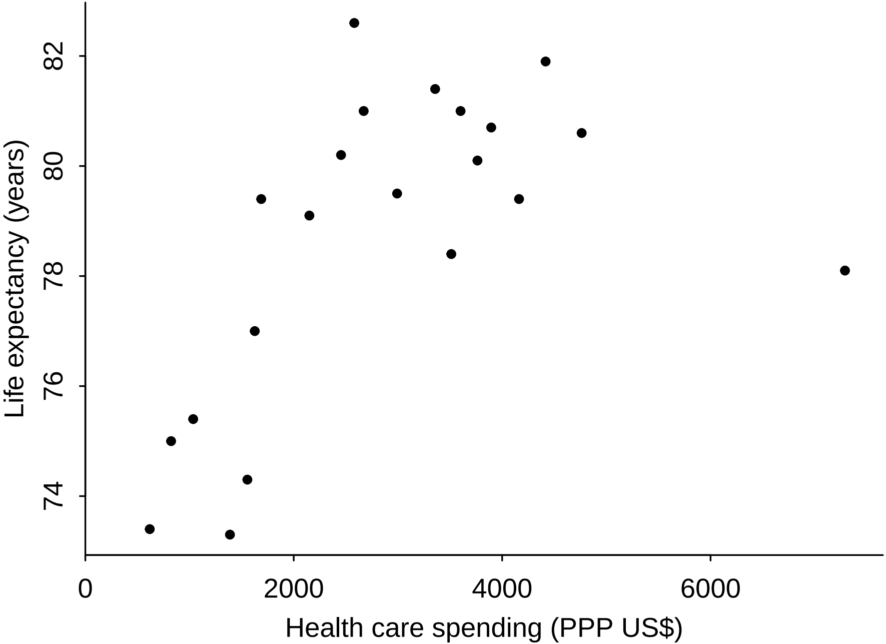
Scatterplot charades



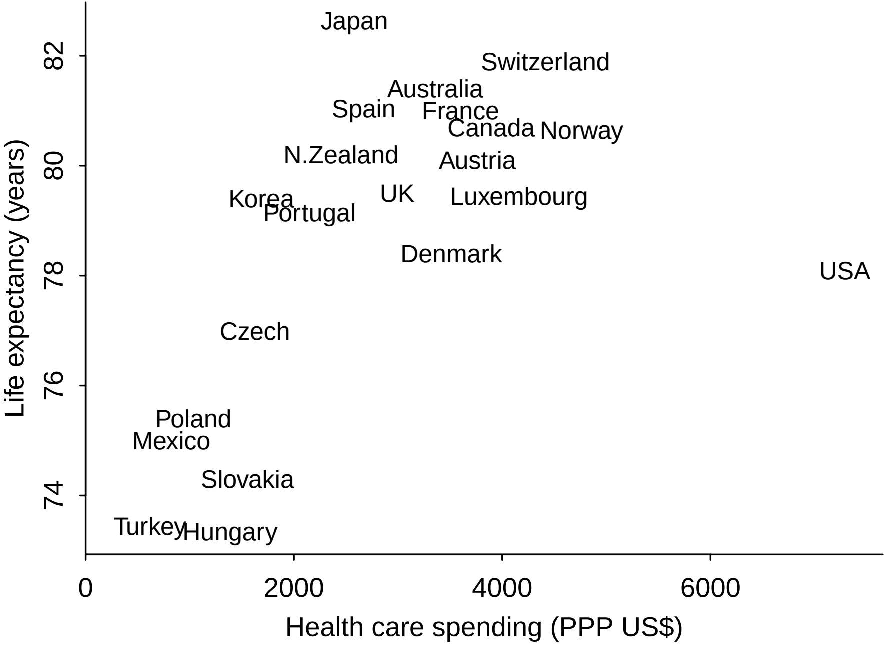
Scatterplot charades



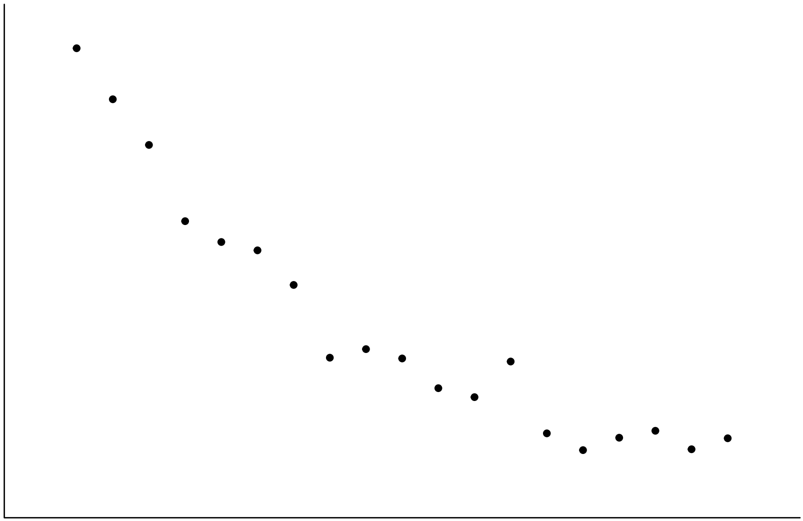
Scatterplot charades



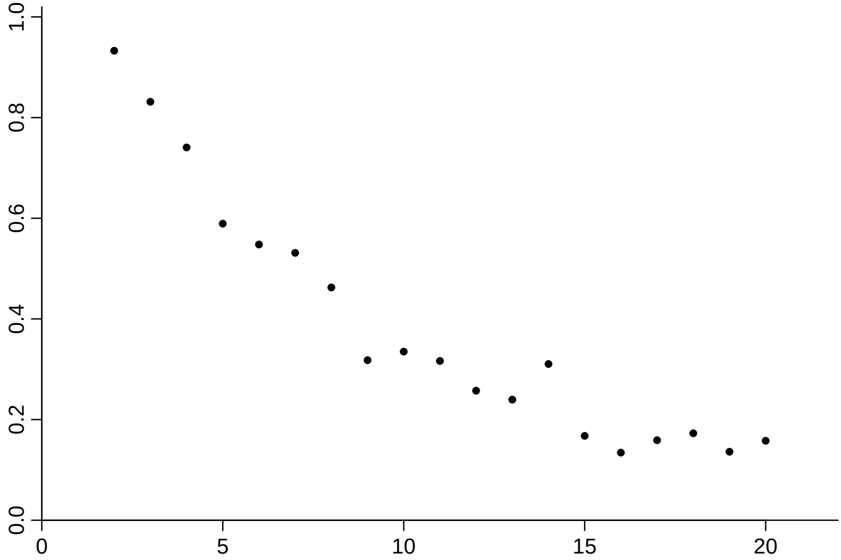
Scatterplot charades



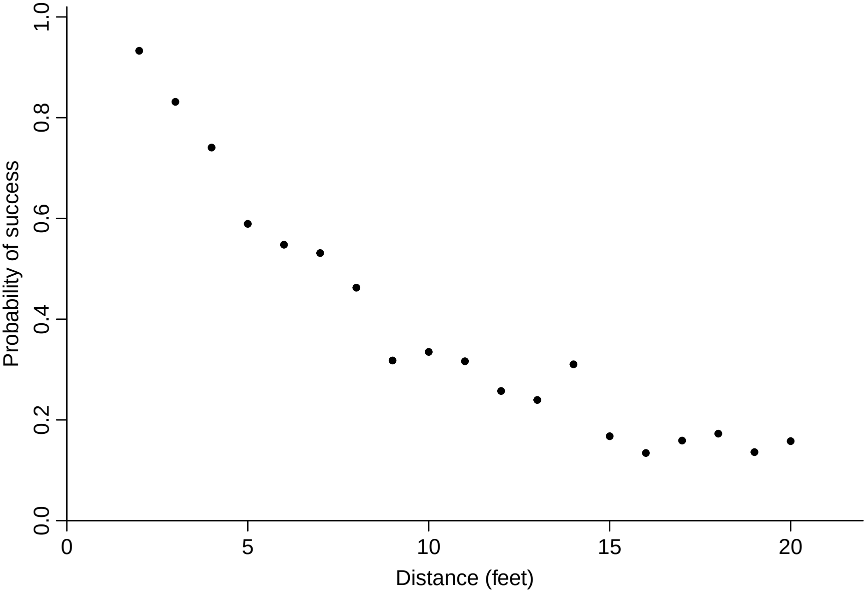
Scatterplot charades



Scatterplot charades



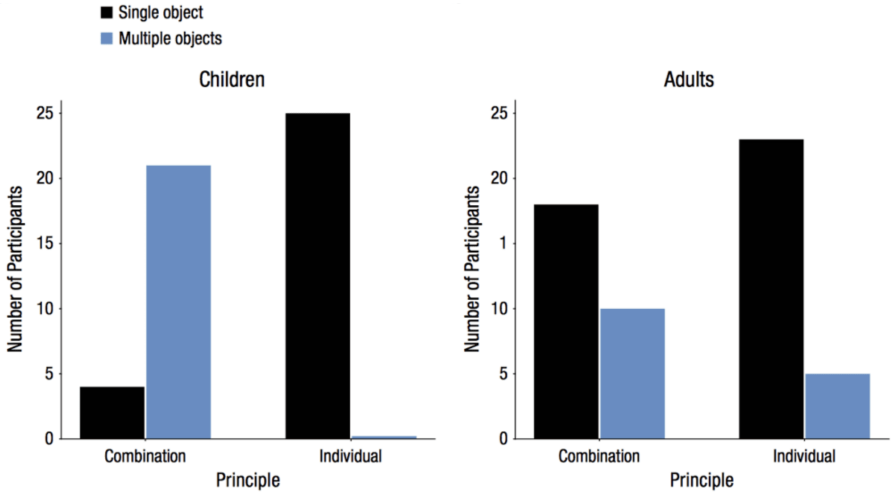
Scatterplot charades



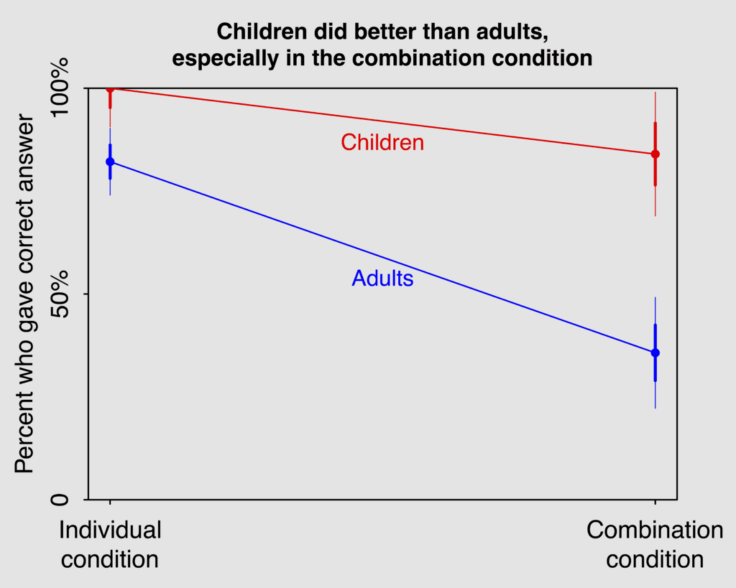
Class 3b

Story

Using the “graphs as comparisons” idea to redraw a graph

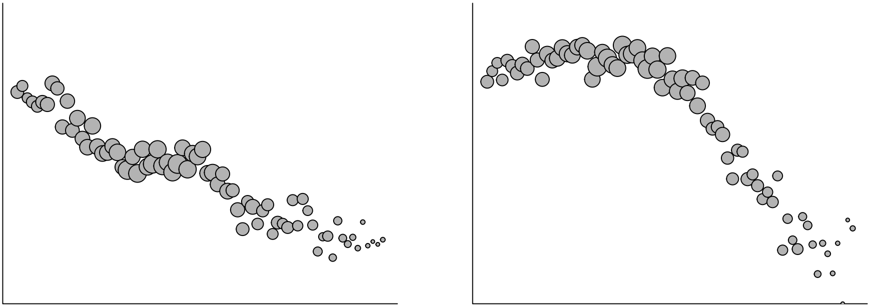


Using the "graphs as comparisons" idea to redraw a graph

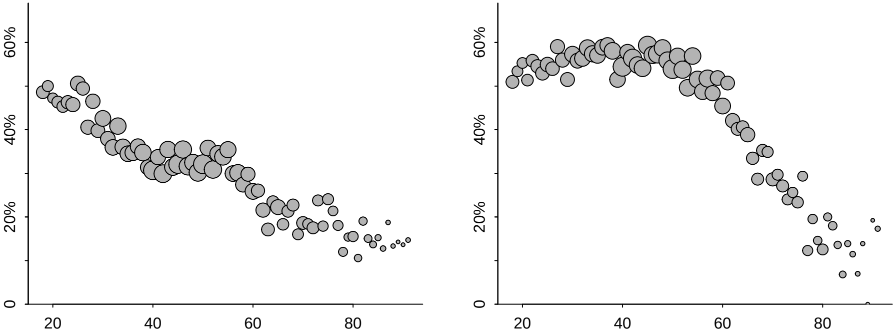


Activity

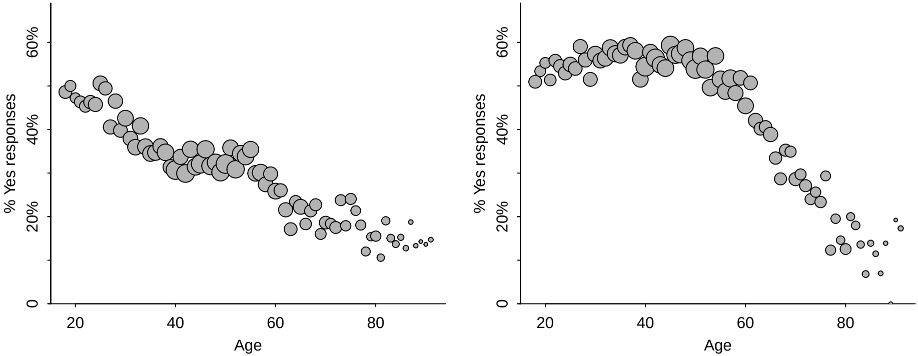
Scatterplot charades



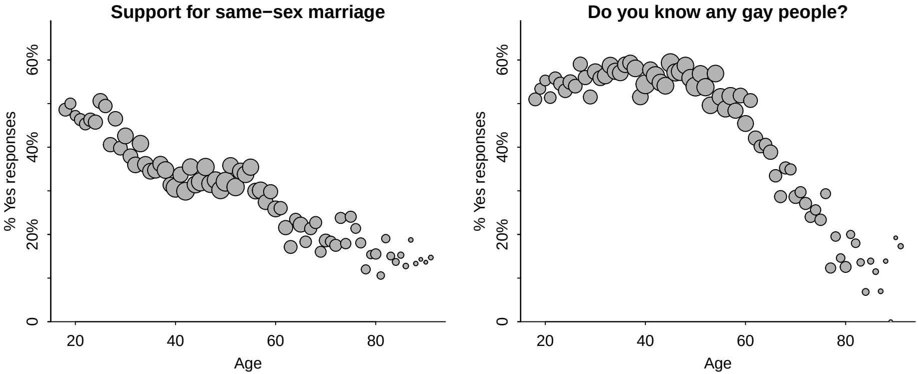
Scatterplot charades



Scatterplot charades



Scatterplot charades



Discuss reading and homework

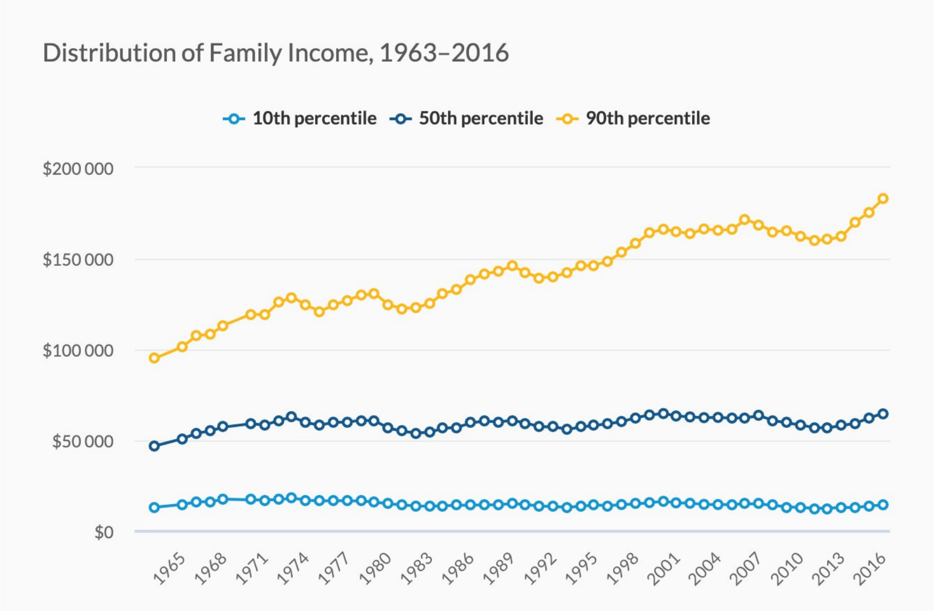
Computer demonstration

Drill

Graph criticism

For each of a series of graphs, identify at least one criticism.

Graph criticism



Sources: Karen Smith, Urban Institute's tabulations from the Current Population Survey 1963–2017.

Graph criticism

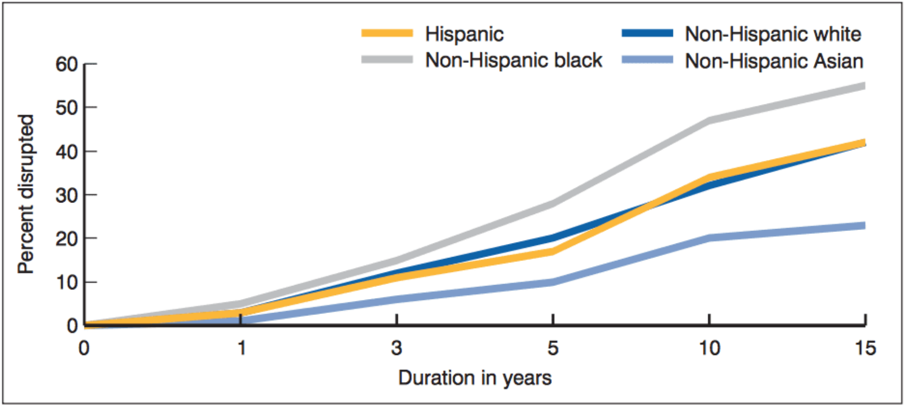
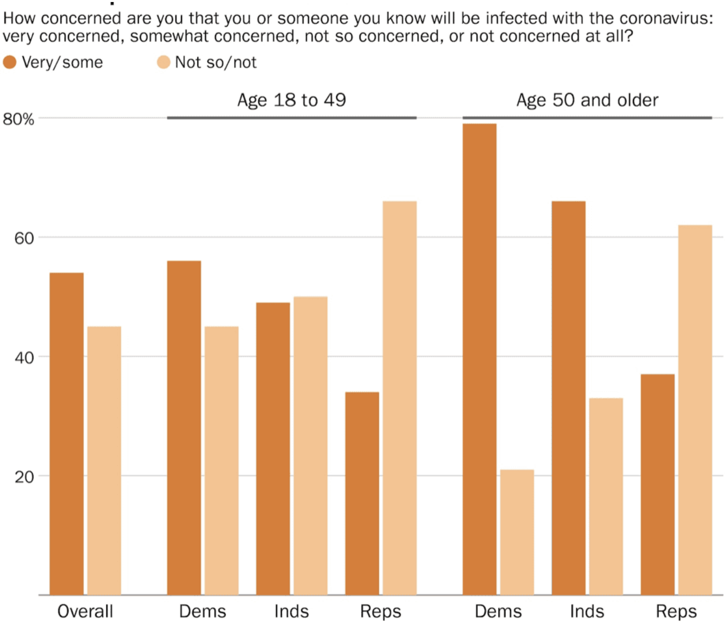


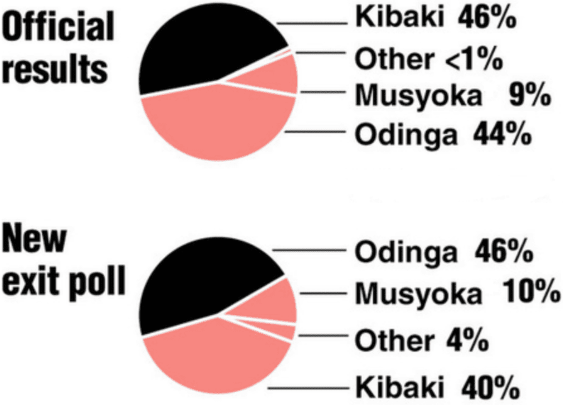
Figure 18. Probability that the first marriage breaks up by duration of marriage and race/ethnicity: United States, 1995

Graph criticism



How concerned are you that you or someone you know will be infected with the coronavirus: very concerned, somewhat concerned, not so concerned, or not concerned at all?

Graph criticism



Discussion problem

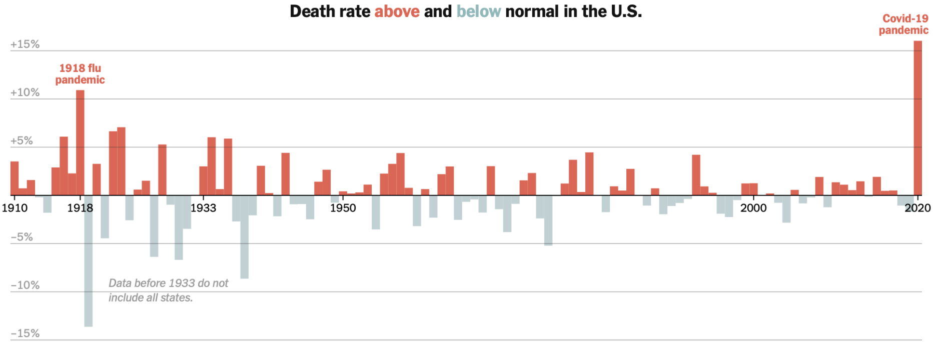
Telling stories with graphs

Mock up a series of graphs.

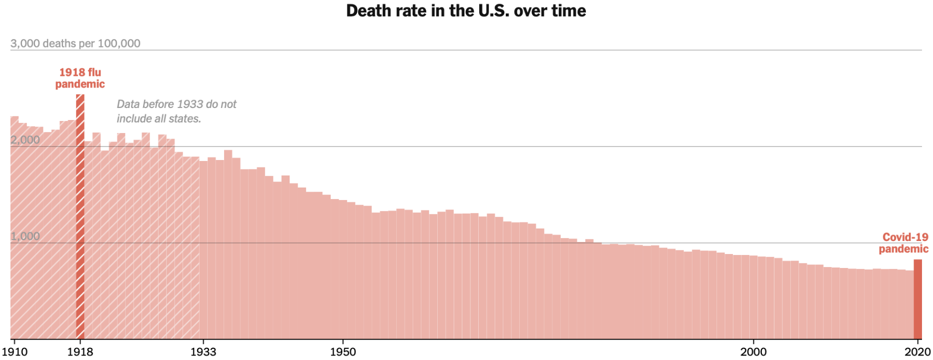
Class 4a

Story

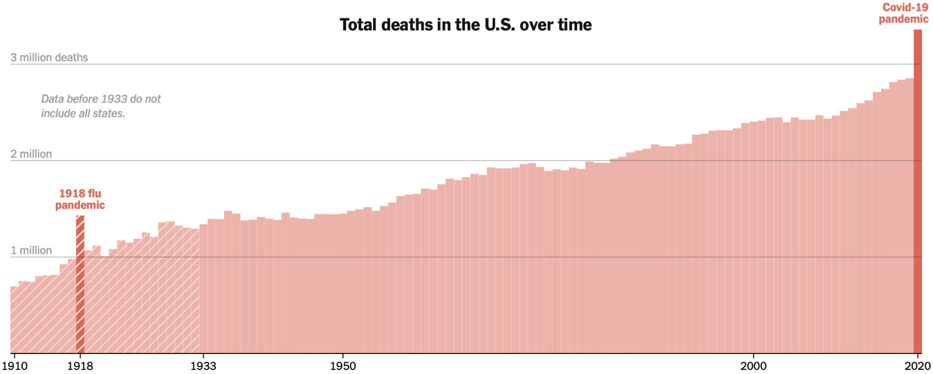
Death rate in the pandemic



Death rate in the pandemic



Death rate in the pandemic



Activity

Amoebas and population growth

- Exponential growth:

$$y = A\exp(bx)$$

$$\log y = a + bx$$

- Exponential decline:

$$y = A\exp(-bx)$$

$$\log y = a - bx$$

Discuss reading and homework

Computer demonstration

Drill

Straight lines

Give R code to graph these lines.

Discussion problem

Squares, cubes, and metabolic rates

- Power-law growth:

$$y = Ax^b$$

$$\log y = a + b \log x$$

- Power-law decline

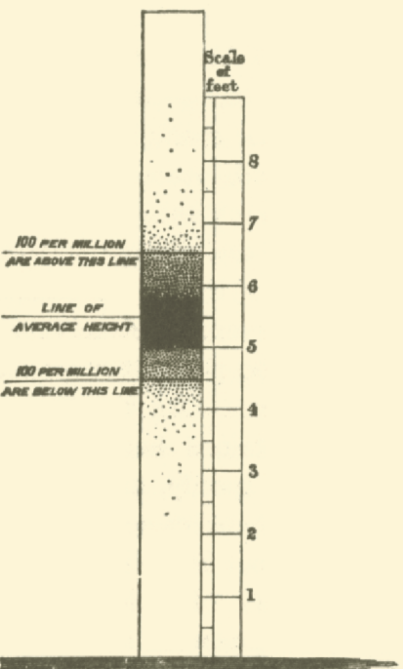
$$y = Ax^{-b}$$

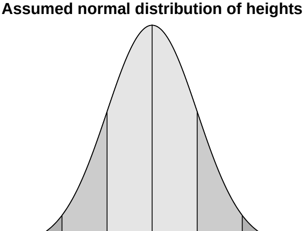
$$\log y = a - b \log x$$

Class 4b

Story

Galton was a hero to most





Activity

Normal and Poisson distributions

- Approximately Poisson distributed
- Approximately normally distributed

Discuss reading and homework

Computer demonstration

Drill

Normal distribution

1. Graph of the distribution
2. Rough estimate of the probability
3. R code to compute the probability

Discussion problem

College admissions and weighted averages

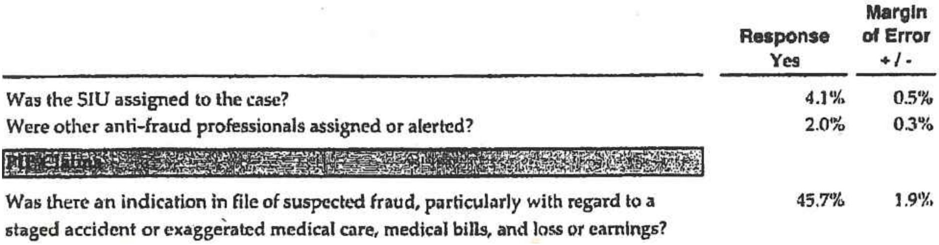
Summary $= a * (\text{Test score}) + b * (\text{Grade point average})$

- Test scores range from 400 to 1600
- Grade point averages range from 0 to 4

Class 5a

Story

They got the wrong standard error



	Response Yes	Margin of Error + / -
Was the SIU assigned to the case?	4.1%	0.5%
Were other anti-fraud professionals assigned or alerted?	2.0%	0.3%
PIP Claims		
Was there an indication in file of suspected fraud, particularly with regard to a staged accident or exaggerated medical care, medical bills, and loss or earnings?	45.7%	1.9%

Activity

Design a bogus social science study, following the Rolf Zwaan model

Rolf Zwaan's steps to produce a clickbait research finding:

1. The idea, based on some popular saying.
2. Theoretical background. Find some remotely relevant connection.
3. The manipulation. Take the expression literally.
4. Outcome measure. Use something fun like candy.
5. Participants in your experiment. Can be anyone.
6. Run experiment 1.
7. Analyze the results. Look for something big in the data.
8. Design experiment 2. Pick a new manipulation.
9. Pick a fun new outcome measure.
10. Repeat steps 5–7.
11. Write your general discussion.
12. Add a quirky celebrity quote.
13. Come up with an amusing title.
14. Hype your findings by overgeneralizing.

Discuss reading and homework

Computer demonstration

Drill

Binomial distribution

A basketball player takes n shots. The shots are independent and she has a 30% chance of making each shot. Let y be the number of shots she makes. What are the mean and standard deviation of y? Sketch the distribution of y.

Discussion problem

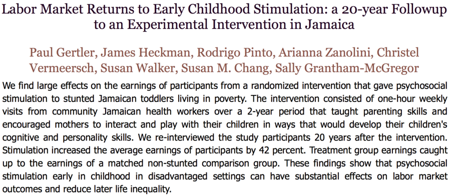
Confidence intervals and true parameter values

Suppose you do 1000 experiments and, from each, you get a 95% interval. You'd expect 950 of these intervals to contain the true parameter values. Assuming your statistical model is correct, would it be a surprise if only 925 of these intervals contained the true parameter values?

Class 5b

Story

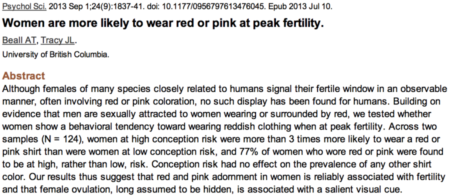
Claims of implausibly large effects



Labor Market Returns to Early Childhood Stimulation: a 20-year Followup to an Experimental Intervention in Jamaica

Paul Gertler, James Heckman, Rodrigo Pinto, Arianna Zanolini, Christel Vermeersch, Susan Walker, Susan M. Chang, Sally Grantham-McGregor

We find large effects on the earnings of participants from a randomized intervention that gave psychosocial stimulation to stunted Jamaican toddlers living in poverty. The intervention consisted of one-hour weekly visits from community Jamaican health workers over a 2-year period that taught parenting skills and encouraged mothers to interact and play with their children in ways that would develop their children's cognitive and personality skills. We re-interviewed the study participants 20 years after the intervention. Stimulation increased the average earnings of participants by 42 percent. Treatment group earnings caught up to the earnings of a matched non-stunted comparison group. These findings show that psychosocial stimulation early in childhood in disadvantaged settings can have substantial effects on labor market outcomes and reduce later life inequality.



Psychol Sci. 2013 Sep 1;24(9):1837-41. doi: 10.1177/0956797613476045. Epub 2013 Jul 10.

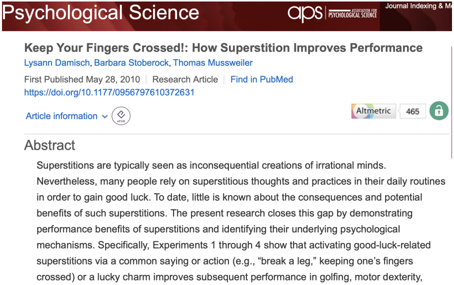
Women are more likely to wear red or pink at peak fertility.

Beall AT, Tracy JL.

University of British Columbia.

Abstract

Although females of many species closely related to humans signal their fertile window in an observable manner, often involving red or pink coloration, no such display has been found for humans. Building on evidence that men are sexually attracted to women wearing or surrounded by red, we tested whether women show a behavioral tendency toward wearing reddish clothing when at peak fertility. Across two samples (N = 124), women at high conception risk were more than 3 times more likely to wear a red or pink shirt than were women at low conception risk, and 77% of women who wore red or pink were found to be at high, rather than low, risk. Conception risk had no effect on the prevalence of any other shirt color. Our results thus suggest that red and pink adornment in women is reliably associated with fertility and that female ovulation, long assumed to be hidden, is associated with a salient visual cue.



Psychological Science aps ASSOCIATION FOR PSYCHOLOGICAL SCIENCE Journal Indexing & M

Keep Your Fingers Crossed!: How Superstition Improves Performance

Lysann Damisch, Barbara Stoberock, Thomas Mussweiler

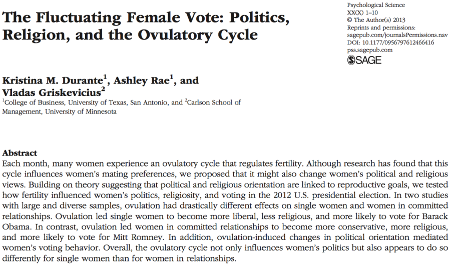
First Published May 28, 2010 | Research Article | Find in PubMed

https://doi.org/10.1177/0956797610372631

Article information ⌄ Altmetric 465

Abstract

Superstitions are typically seen as inconsequential creations of irrational minds. Nevertheless, many people rely on superstitious thoughts and practices in their daily routines in order to gain good luck. To date, little is known about the consequences and potential benefits of such superstitions. The present research closes this gap by demonstrating performance benefits of superstitions and identifying their underlying psychological mechanisms. Specifically, Experiments 1 through 4 show that activating good-luck-related superstitions via a common saying or action (e.g., "break a leg," keeping one's fingers crossed) or a lucky charm improves subsequent performance in golfing, motor dexterity,



The Fluctuating Female Vote: Politics, Religion, and the Ovulatory Cycle

Psychological Science
XX(X) 1–10
© The Author(s) 2013
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797612466416
pss.sagepub.com
SAGE

Kristina M. Durante[1], Ashley Rae[1], and Vladas Griskevicius[2]

[1]College of Business, University of Texas, San Antonio, and [2]Carlson School of Management, University of Minnesota

Abstract

Each month, many women experience an ovulatory cycle that regulates fertility. Although research has found that this cycle influences women's mating preferences, we proposed that it might also change women's political and religious views. Building on theory suggesting that political and religious orientation are linked to reproductive goals, we tested how fertility influenced women's politics, religiosity, and voting in the 2012 U.S. presidential election. In two studies with large and diverse samples, ovulation had drastically different effects on single women and women in committed relationships. Ovulation led single women to become more liberal, less religious, and more likely to vote for Barack Obama. In contrast, ovulation led women in committed relationships to become more conservative, more religious, and more likely to vote for Mitt Romney. In addition, ovulation-induced changes in political orientation mediated women's voting behavior. Overall, the ovulatory cycle not only influences women's politics but also appears to do so differently for single women than for women in relationships.

Activity

Discuss effects in the context of a social science example

1. Consider a topic of interest
2. Consider an outcome measure and hypothesize a treatment effect
3. Construct a hypothetical experiment
4. Specify sample size
5. Hypothesize distribution of outcomes under control and treatment
6. Figure out estimate and standard error
7. Will the experiment give a reliable estimate?

Discuss reading and homework

Computer demonstration

Drill

Sample size and standard errors

How large does n have to be so that your estimate has a standard error of . . . ?

Discussion problem

Approximate standard error for average "feeling thermometer" ratings

From American National Election Study: "I'd like to get your feelings toward some of our political leaders and other people who are in the news these days. I'll read the name of a person and I'd like you to rate that person using something we call the feeling thermometer. Ratings between 50 degrees and 100 degrees mean that you feel favorable and warm toward the person. Ratings between 0 degrees and 50 degrees mean that you don't feel favorable toward the person and that you don't care too much for that person. You would rate the person at the 50 degree mark if you don't feel particularly warm or cold toward the person."

Class 6a

Story

The proportion of identical twins in the population

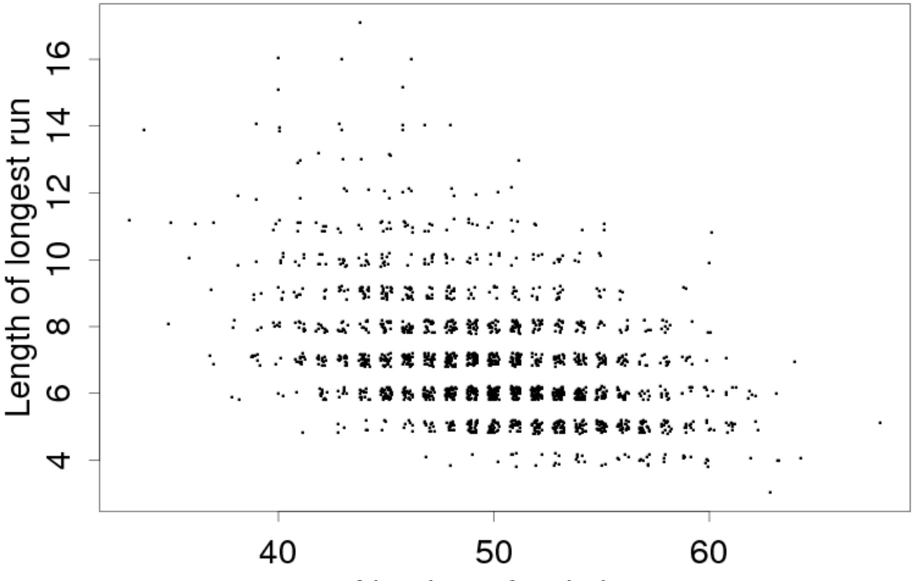
- Probability of fraternal twins: 1/125
- Probability of identical twins: 1/300
- How do we know this?

Activity

Real vs. fake coin flips

1. Instructor and two judges leave the room
2. Group A: Create a sequence of 100 real coin flips
3. Group B; Create a sequence of 100 real coin flips
4. Group C; Create a fake sequence of 0's and 1's that looks like 100 coin flips
5. Group D; Create a fake sequence of 0's and 1's that looks like 100 coin flips
6. Groups A and B write sequences on one board; groups C and D write on the other board
7. We return and figure out which is which!

Real vs. fake coin flips



Discuss reading and homework

Computer demonstration

Drill

Programming in R: distribution of basketball shots (two players)

Write an R function to simulate the outcome of two basketball players shooting n_1 and n_2 baskets, with probability p_1 and p_2 of success. The function should take n_1, n_2, p_1, and p_2, as arguments, simulate the shots, calculate the proportions of shots made for each player, and return the difference in proportions.

Discussion problem

Simulate a mixed discrete/continuous distribution

Simulate the incomes of a hypothetical set of 100 people where there is a probability of zero income and a lognormal distribution otherwise.

Class 6b

Story

Simulating a process of innovation, experimentation, and improvement

How to simulatie a process of innovation, experimentation, and improvement?

Activity

Simulating a probability process

Discuss reading and homework

Computer demonstration

Drill

Propagation of uncertainty

A man applies for n jobs. For each job he has a p_1 chance of getting an interview. If he is interviewed, he has a p_2 chance of getting an offer. Write an R function to simulate this process and compute the number of offers he gets. The function should take n, p_1, and p_2 as inputs and return a single number.

Discussion problem

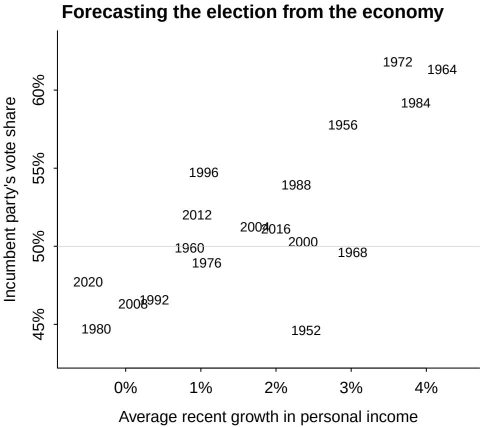
Simulate clustering of buses

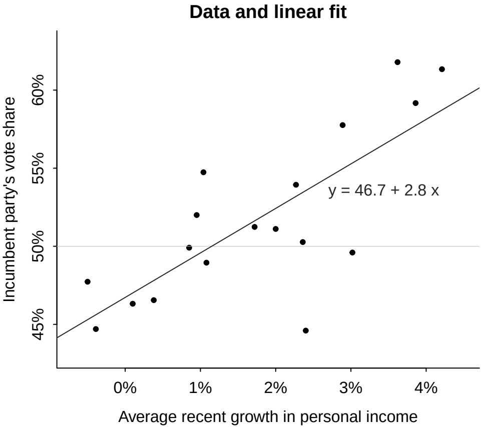
A famous real-world example of a stochastic process is the clustering of buses along a route. Suppose that buses start out equally spaced in time and then have to stop for passengers. The bus in front will pick up the first set of passengers, allowing the next bus to skip some stops if no new passengers arrive. This random process will, on average, lead to the clumping of buses: that annoying phenomenon whereby you have to wait a long time for a bus, and then two or three arrive together.
How can this process be simulated on the computer?

Class 7a

Story

Slope when predicting elections from the economy


Forecasting the election from the economy


Data and linear fit

Slope when predicting elections from the economy

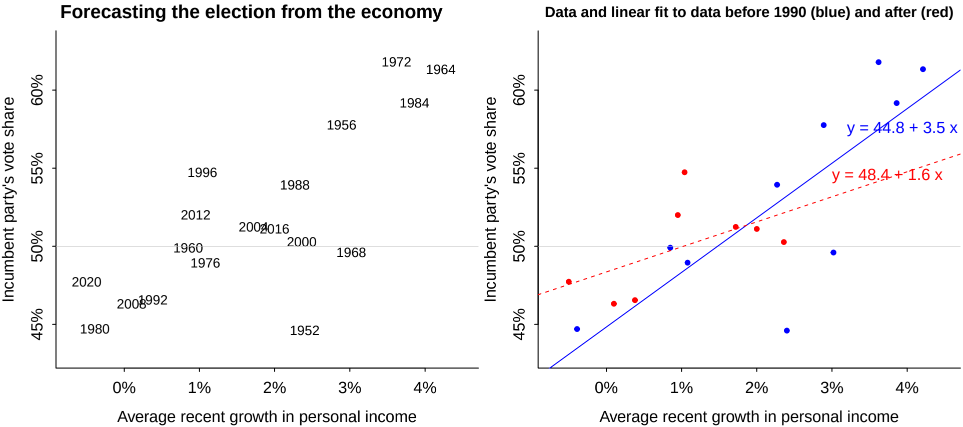
```
            Median MAD_SD
(Intercept) 46.7    1.4
growth       2.8    0.6

Auxiliary parameter(s):
      Median MAD_SD
sigma 3.7    0.7
```

Slide when predicting elections from the economy

```
1948-1988:

             Median MAD_SD
(Intercept) 44.8    2.7
growth       3.5    1.0

Auxiliary parameter(s):
      Median MAD_SD
sigma 4.5    1.2
```

```
1992-2020:

             Median MAD_SD
(Intercept) 48.4    1.5
growth       1.6    1.1

Auxiliary parameter(s):
      Median MAD_SD
sigma 2.8    0.8
```

Slope when predicting elections from the economy



Activity

Fake-data simulation and fitting a regression

1. Come up with a cover story
2. Set up a generative model
3. Choose the parameters
4. Code it up, run it, debug it
5. Play around with the settings

Discuss reading and homework

Computer demonstration

Drill

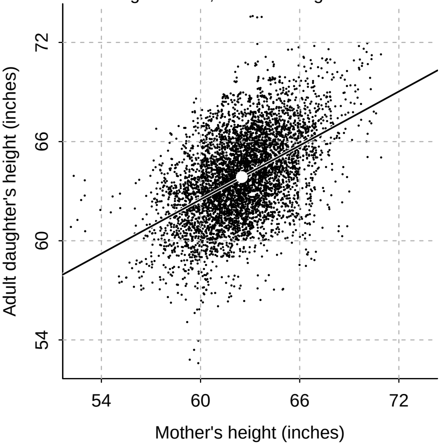
Regression to the mean

What is the student's expected score on the post-test?

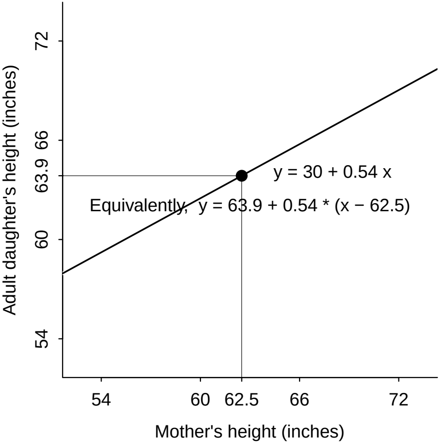
Discussion problem

Other examples of regression to the mean

Mothers' and daughters' heights,
average of data, and fitted regression line



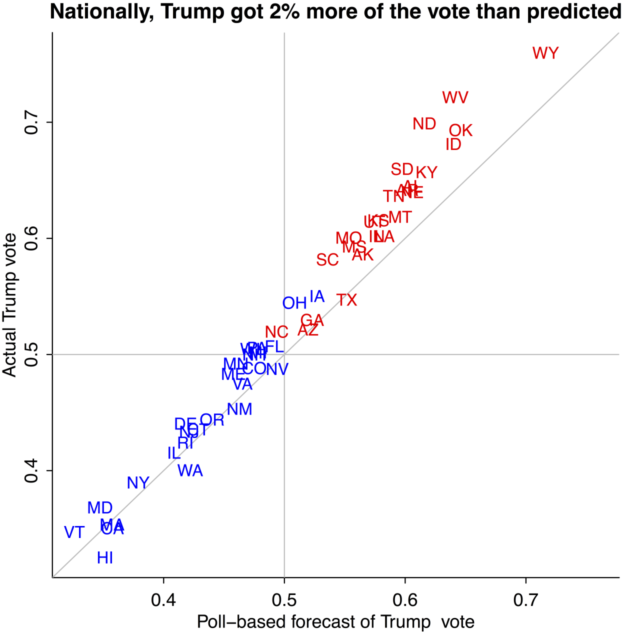
The fitted regression line and the average of the data



Class 7b

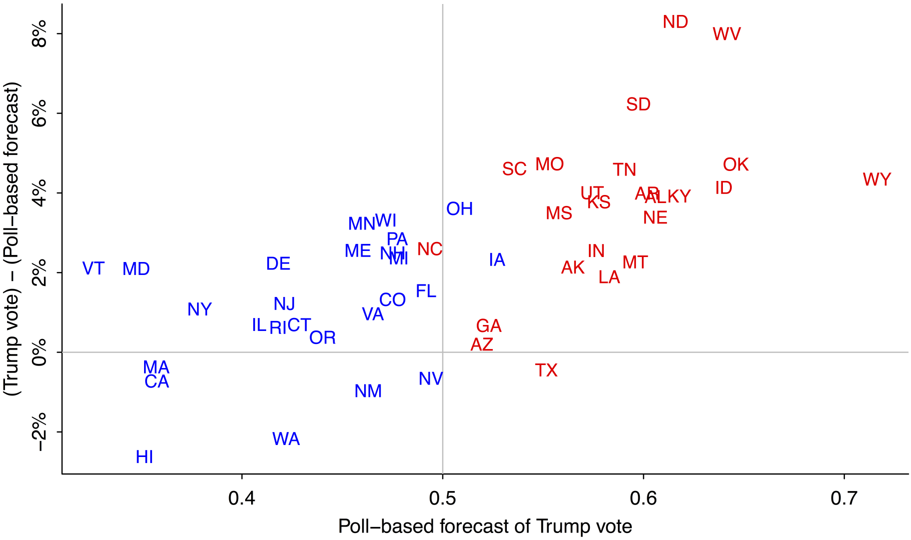
Story

Clinton/Trump vote vs. polls and predictions

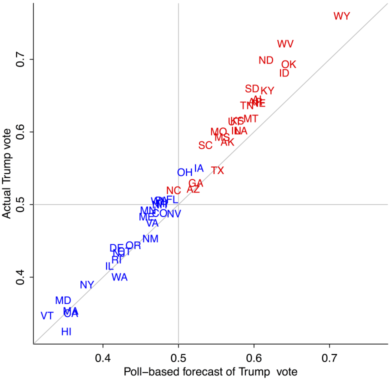


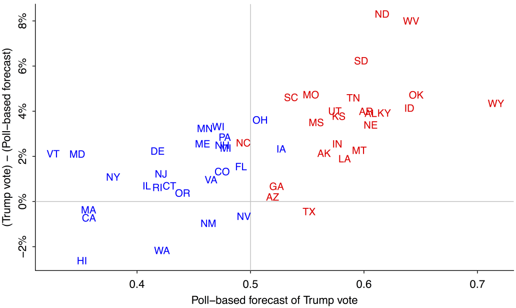
Clinton/Trump vote vs. polls and predictions

Trump did much better than predicted in states that Romney won in 2012



Clinton/Trump vote vs. polls and predictions


Nationally, Trump got 2% more of the vote than predicted


Trump did much better than predicted in states that Romney won in 2012

Activity

Before-after memory tests

brother
bat
beginner
boy
run
experience
end
wheel
jeans
reward
house
theory
train
prose
government
art
song
nation
baseball
flock

Before-after memory tests

Before-after memory tests

brother
bat
beginner
boy
run
experience
end
wheel
jeans
reward

house
theory
train
prose
government
art
song
nation
baseball
flock

Before-after memory tests

cloth
lizard
hook
wheel
school
fight
string
wave
garden
division

boundary
drain
health
wax
car
lace
class
woman
army
fold

Before-after memory tests

Before-after memory tests

cloth
lizard
hook
wheel
school
fight
string
wave
garden
division

boundary
drain
health
wax
car
lace
class
woman
army
fold

Discuss reading and homework

Computer demonstration

Drill

Scatterplots, regression lines, and regression functions

Draw the fitted regression line and a curve of the predicted value of y given x

Discussion problem

Understanding uniform partisan swing (considering regression to the mean)

National elections approximately follow uniform partisan swing at the national and local levels, typically with only small changes from year to year. But over a 20-year period there can be big changes. How can these patterns in the U.S. and elsewhere be understood? Is the concept of regression to the mean relevant here?

Class 8a

Story

$5^2 + 12^2 = 13^2$

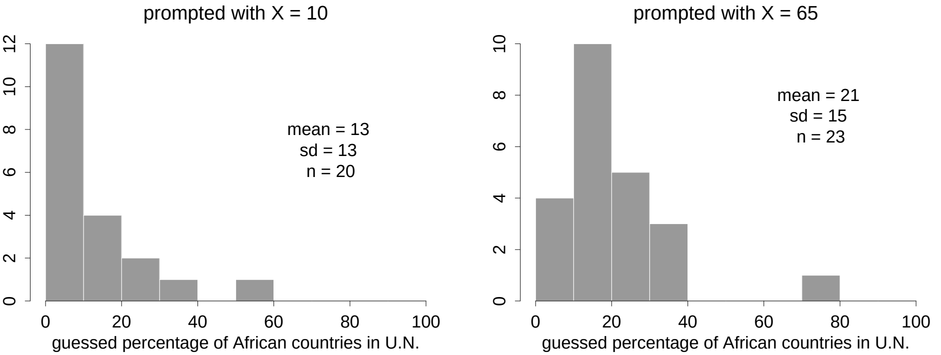
Two sources of uncertainty

Activity

African countries in the United Nations

Answer the two questions on the survey form, then fold it and pass it to the front of the room.

African countries in the United Nations



Discuss reading and homework

Computer demonstration

Drill

Sketch a fitted regression model

Sketch the regression line and data that match both the fitted model and the residual standard deviation.

Discussion problem

Interpreting statistically significant results given huge sample sizes

Suppose you run a regression with a huge sample size, for example a mega-poll with 100 000 respondents or an A/B test at a big company. With a large enough sample size, the standard error will be very small, and so even a very small effect can be statistically significant. How can you interpret such a result?

Class 8b

Story

Interpreting the regression of earnings on height

Activity

Socioeconomic status and political ideology

1. x_1: a socioeconomic measure
2. x_2: a socioeconomic measure
3. y: a measure of political ideology

Discuss reading and homework

Computer demonstration

Drill

Predicting probabilities using regression

Regression predicting final from midterm exam score:

```
            Median  MAD_SD
(Intercept) 24.8    1.4
midterm     0.5     0.1

Auxiliary parameter(s):
      Median  MAD_SD
sigma 11.6    0.3
```

Given a midterm score, calculate the predicted final exam score and the probability that the final exam score will be in the range specified.

Discussion problem

How large was the sample size?

Regression predicting final from midterm exam score:

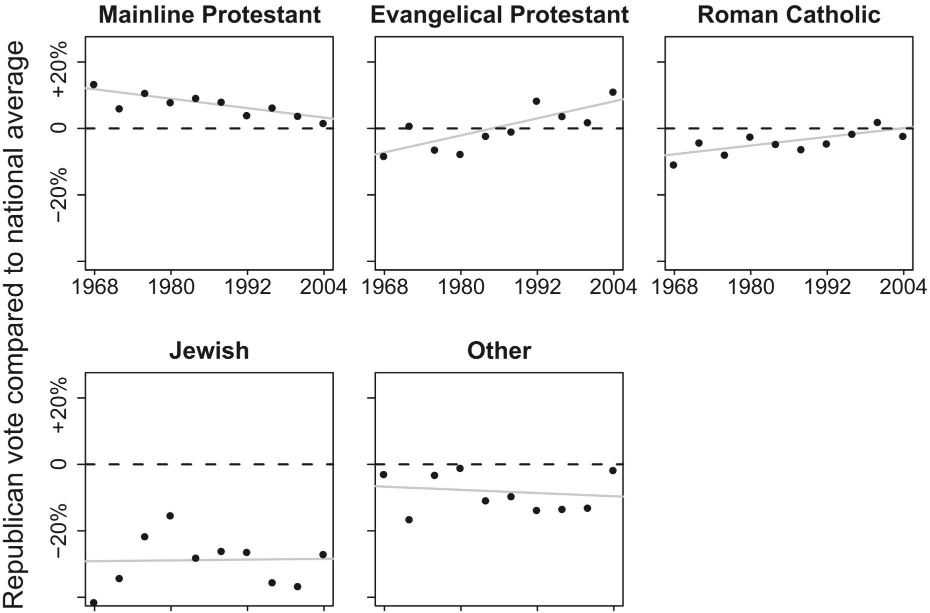
```
            Median  MAD_SD
(Intercept) 24.8    1.4
midterm     0.5     0.1

Auxiliary parameter(s):
        Median  MAD_SD
sigma   11.6    0.3
```

Approximately what was the sample size of this regression?

Class 9a

Story

No, Ronald Reagan did not win "overwhelming support from evangelicals"



Activity

In pairs, simulate and recover regression lines

- Student #1:
  1. Create fake data from the model, $y = a + bx + \text{error}$
  2. Put `x` and `y` in a data frame called `data`
  3. Type `library("rstanarm")`
  4. Type ctrl-L to clear the R console
  5. Type `range(data$x)`
- Student #2:
  1. Take the computer
  2. Fit the regression of `y` on `x` using `stan_glm`
  3. Sketch (not on the computer) your guess of the scatterplot of x and y

Discuss reading and homework

Computer demonstration

Drill

Sample size and standard errors

```
stan_glm
 family:       gaussian [identity]
 formula:      earn ~ height
 observations: 1816
 predictors:   2
------
            Median MAD_SD
(Intercept) -85000   9000
height        1600    100

Auxiliary parameter(s):
      Median MAD_SD
sigma 22000    400
```

Discussion problem

From inference to decision

You run a regression on 100 students studying the effect of an educational intervention on a test score that has a population mean of 500 and standard deviation 100. You get an estimated treatment effect of 20 points with a standard error of 15 points. What does this tell you? Would you do the intervention? What might you do next?

Class 9b

Story

Does having a girl make you more conservative or more liberal?

Study #1: "Using nationally-representative data from the [1994] General Social Survey, we find that female offspring induce more conservative political identification. We hypothesize that this results from the change in reproductive fitness strategy that daughters may evince."

Study #2: "We document evidence that having daughters leads people to be more sympathetic to left-wing parties. Giving birth to sons, by contrast, seems to make people more likely to vote for a right-wing party. Our data, which are primarily from Great Britain, are longitudinal. We also report corroborative results for a German panel."

Does having a girl make you more conservative or more liberal?

Headlines:

- "The Effect of Daughters on Partisanship and Social Attitudes Toward Women"
- "Does Having Daughters Make You More Republican?"
- "Parents With Daughters Are More Likely To Be Republicans, Says New Study"
- "Parents Of Daughters Lean Republican, Study Shows"
- "The Daughter Theory: Does Raising Girls Make Parents Conservative?"

What's missing there?

Activity

How much do you have to move a point to shift the fitted line by a specified amount?

How much do you have to move a point to shift the fitted line by a specified amount?

Discuss reading and homework

Computer demonstration

Drill

Averages and comparisons as regression models

For each statement, express it as a regression in R code and algebra, and give the estimated regression coefficients.

Discussion problem

Sample size and statistical significance

You run an experiment on 200 people and get an estimated treatment effect of 0.20 with standard error 0.15. So, not quite "statistically significant." What might you expect to see if you re-ran with 400 people? Would you expect statistical significance then?

Class 10a

Story

Studying fairness of random exams

- Students randomly assigned to exams
- Average scores:
  - 65 for exam A
  - 71 for exam B
- Should we adjust the students' scores?

Activity

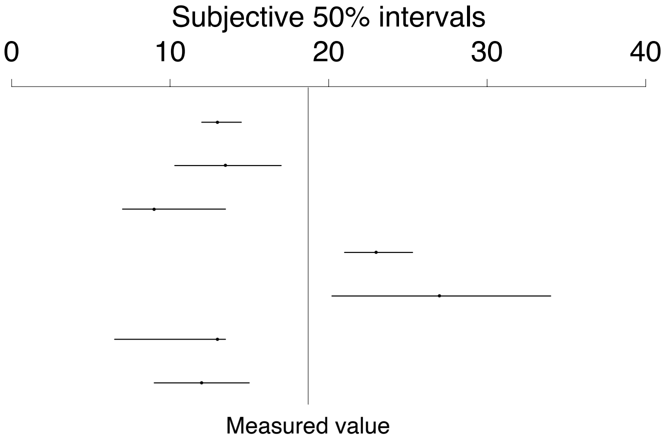
Coverage of prediction intervals

Uncertain quantity	25% lower bound	75% upper bound
% Black		
# eggs		
# airline deaths		
% girl births		
# babies born		
# abortions		
% degrees in CS		
# degrees		
# Super Bowl watchers		
$ median income		

Coverage of prediction intervals

Uncertain quantity	25% bound	75% bound	TRUTH!
% Black			12.4
# eggs			64.6 billion
# airline deaths			299
% girl births			48.8
# babies born			4.06 million
# abortions			857 000
% degrees in CS			4.4
# degrees			2.01 million
# Super Bowl watchers			101.3 million
$ median income			67 500

Coverage of prediction intervals



Discuss reading and homework

Computer demonstration

Drill

Prediction

```
stan_glm
 family:       gaussian [identity]
 formula:      earn ~ height
 observations: 1816
 predictors:   2
------
            Median MAD_SD
(Intercept) -85000   9000
height        1600    100

Auxiliary parameter(s):
      Median MAD_SD
sigma 22000    400
```

Approximately what is the predictive distribution from this regression of the earnings of a person who is . . . ?

Discussion problem

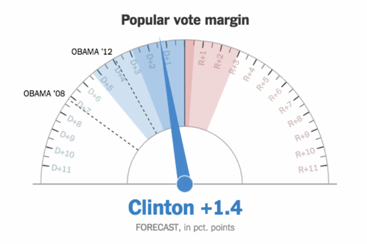
Predictive uncertainties

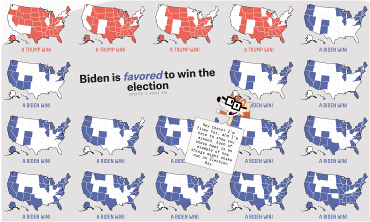
Using a survey with 500 respondents, you fit a regression predicting 0–100 feeling thermometer response on some celebrity (let the students pick someone) given a party identification predictor (on a -3 to 3 scale). Having fit the model, you do `posterior_linpred` and `posterior_predict` for someone who is strongly Republican ($x = 3$). What are the approximate predictive uncertainties?

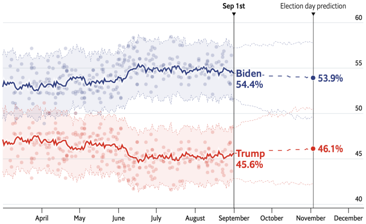
Class 10b

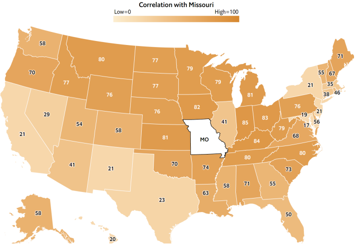
Story

Uncertainties in election forecasts




A TRUMP WIN!
A BIDEN WIN!
Biden is favored to win the election
UPDATED 3 HOURS AGO
Hey there! I'm Fivey Fox, and I'm here to show you around. Each of these maps is an example of how things might shake out on Election Day.





Activity

Prior distributions for real-world quantities

Discuss reading and homework

Computer demonstration

Drill

Elections: calculating Bayesian posterior mean, standard deviation and probability of success

Using a regression model, you forecast that a certain candidate will have 45% support, with forecast standard deviation of 5%. You then do a simple random sample survey of 1000 people, of whom 500 support the candidate.

(i) Give the Bayesian posterior mean and standard deviation of the candidate's support in the population. (ii) What is the posterior probability that this candidate has at least 50% support?

Discussion problem

Real-world examples

1. Give an example of regression prediction.
2. Give an example of Bayesian combination of information.

Class 11a

Story

Incumbency advantage in congressional elections

Predicting Democratic vote share in U.S. House elections in 1988, given incumbency,

$$\begin{cases} +1 & \text{for districts where a Democrat was running for reelection} \\ 0 & \text{for open seats} \\ -1 & \text{for districts where a Republican was running for reelection} \end{cases}$$

```
            Median MAD_SD
(Intercept) 0.50   0.00
inc88       0.17   0.01

Auxiliary parameter(s):
      Median MAD_SD
sigma 0.08   0.00
```

Incumbency advantage in congressional elections

Predicting Democratic vote share in U.S. House elections in 1988, given incumbency,

$$\begin{cases} +1 & \text{for districts where a Democrat was running for reelection} \\ 0 & \text{for open seats} \\ -1 & \text{for districts where a Republican was running for reelection,} \end{cases}$$

Democratic vote share in 1986,
and incumbent *party* (-1, 0, or $+1$) in 1988:

```
            Median MAD_SD
(Intercept)  0.21   0.02  
inc88        0.12   0.01  
v86_adj      0.58   0.04  
incparty88  -0.04   0.02  

Auxiliary parameter(s):
      Median MAD_SD
sigma 0.06   0.00  
```

Incumbency advantage in congressional elections

Compare the two models:

```
            Median MAD_SD
(Intercept) 0.50   0.00
inc88       0.17   0.01

Auxiliary parameter(s):
      Median MAD_SD
 sigma 0.08   0.00
```

```
             Median MAD_SD
(Intercept)  0.21   0.02
inc88        0.12   0.01
v86          0.58   0.04
incparty88  -0.04   0.02

Auxiliary parameter(s):
     Median MAD_SD
sigma 0.06   0.00
```

Incumbency advantage in congressional elections

Predicting Democratic vote share in U.S. House elections in 2020, given incumbency,

$$\begin{cases} +1 & \text{for districts where a Democrat was running for reelection} \\ 0 & \text{for open seats} \\ -1 & \text{for districts where a Republican was running for reelection,} \end{cases}$$

Democratic vote share in 2018,
and incumbent party (-1, 0, or $+1$) in 2020:

```
              Median MAD_SD
(Intercept)   0.01   0.01  
inc2020       0.03   0.00  
v2018_adj     0.92   0.02  
incparty2020 -0.02   0.00  

Auxiliary parameter(s):
      Median MAD_SD
sigma 0.03   0.00  
```

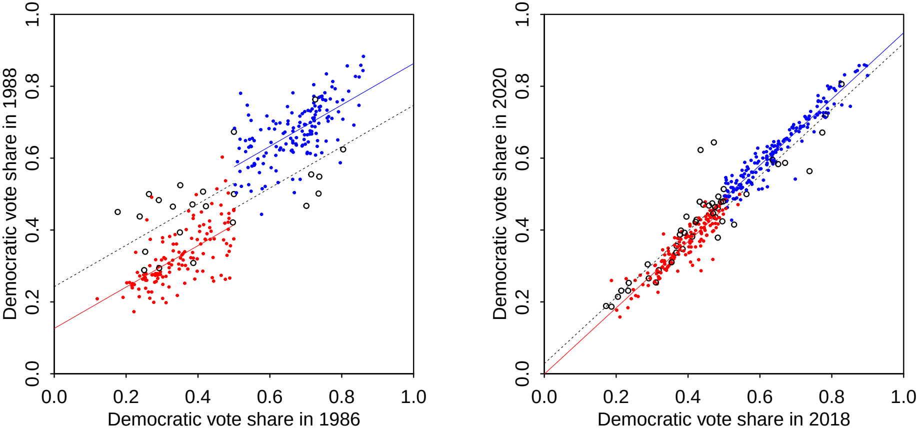
Incumbency advantage in congressional elections

Compare the two time periods:

```
              1986-1988

             Median MAD_SD
(Intercept)  0.21   0.02
inc88        0.12   0.01
v86          0.58   0.04
incparty88  -0.04   0.02

Auxiliary parameter(s):
      Median MAD_SD
sigma 0.06   0.00
```

```
             2018-2020

              Median MAD_SD
(Intercept)   0.01   0.01
inc2020       0.03   0.00
v2018         0.92   0.02
incparty2020 -0.02   0.00

Auxiliary parameter(s):
      Median MAD_SD
sigma 0.03   0.00
```

Incumbency advantage in congressional elections



Activity

Memory quiz with pre-test, treatment, and outcome

friend
cloth
metal
comparison
balloon
tiger
cable
dress
worm
mother
verse
curtain
attempt
size
match
form
maid
expansion
goose
liquid

Memory quiz with pre-test, treatment, and outcome

For second memory quiz:

- If the last digit of your Social Security number is *odd*, you'll get *30 seconds*
- If the last digit of your Social Security number is *even*, you'll get *60 seconds*

Memory quiz with pre-test, treatment, and outcome

friend
cloth
metal
comparison
balloon
tiger
cable
dress
worm
mother

verse
curtain
attempt
size
match
form
maid
expansion
goose
liquid

Memory quiz with pre-test, treatment, and outcome

power
screw
curve
writing
sister
baby
cry
collar
trousers
brick
snail
cake
unit
driving
hair
scarecrow
discussion
channel
sheep
ocean

Memory quiz with pre-test, treatment, and outcome

- x: Score on first memory quiz
- z: Treatment ($z = 1$ if you got 60 seconds, $z = 0$ if you got 30 seconds)
- y: Score on second memory quiz

Consider two regression models predicting y from x and z:

- `y ~ x + z`
- `y ~ x + z + x*z`

Memory quiz with pre-test, treatment, and outcome

power
screw
curve
writing
sister
baby
cry
collar
trousers
brick

snail
cake
unit
driving
hair
scarecrow
discussion
channel
sheep
ocean

Discuss reading and homework

Computer demonstration

Drill

Thinking through predictors

Set up each of these problems as a regression model.

Drill

Causal regressions

Set up each of these problems as a regression model.

Discussion problem

Adjusting for pre-treatment variables

Consider an observational study on a topic of interest, comparing exposed to unexposed groups. Set this up as a regression problem, first as a simple comparison and then adjusting for pre-treatment predictors.

Class 11b

Story

Predicting teaching evaluations from beauty and other variables

```
 formula:      eval ~ beauty + female
 observations: 463
 predictors:   3
------
            Median MAD_SD
(Intercept)  4.09   0.03
beauty       0.15   0.03
female      -0.20   0.05

Auxiliary parameter(s):
      Median MAD_SD
      sigma 0.54   0.02
```

Predicting teaching evaluations from beauty and other variables

```
              Median MAD_SD
(Intercept)    4.11   0.03
beauty         0.20   0.04
female        -0.21   0.05
beauty:female -0.11   0.06

Auxiliary parameter(s):
      Median MAD_SD
sigma 0.54   0.02
```

Predicting teaching evaluations from beauty and other variables

```
              Median MAD_SD
(Intercept)    4.21   0.15 
beauty         0.19   0.04 
female        -0.22   0.05 
age            0.00   0.00 
beauty:female -0.11   0.06 

Auxiliary parameter(s):
      Median MAD_SD
sigma 0.54   0.02  
```

Predicting teaching evaluations from beauty and other variables

```
              Median MAD_SD
(Intercept)    4.21   0.14
beauty         0.19   0.05
female        -0.22   0.05
age10         -0.02   0.03
beauty:female -0.11   0.06

Auxiliary parameter(s):
      Median MAD_SD
sigma 0.54   0.02
```

Predicting teaching evaluations from beauty and other variables

```
              Median MAD_SD
(Intercept)    4.23   0.14 
beauty         0.19   0.04 
female        -0.21   0.05 
age10         -0.02   0.03 
nonenglish    -0.34   0.10 
beauty:female -0.11   0.06 

Auxiliary parameter(s):
      Median MAD_SD
sigma 0.53   0.02  
```

Activity

Designing a study with regression in mind

- x: Score on first memory quiz
- z: Treatment
- y: Score on second memory quiz

Consider two regression models predicting y from x and z:

- `y ~ x + z`
- `y ~ x + z + x*z`

Discuss reading and homework

Computer demonstration

Drill

Interpret interaction coefficients

For each model, describe each coefficient in words.

Discussion problem

What is gained by including a pre-test?

Consider a randomized experiment:

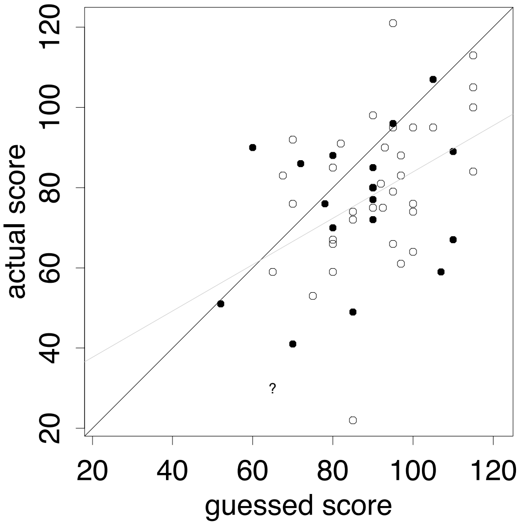
- Regression of post-test on treatment: `y ~ z`
- Regression of post-test on treatment and pre-test: `y ~ z + x`

What is gained by adjusting for pre-test?

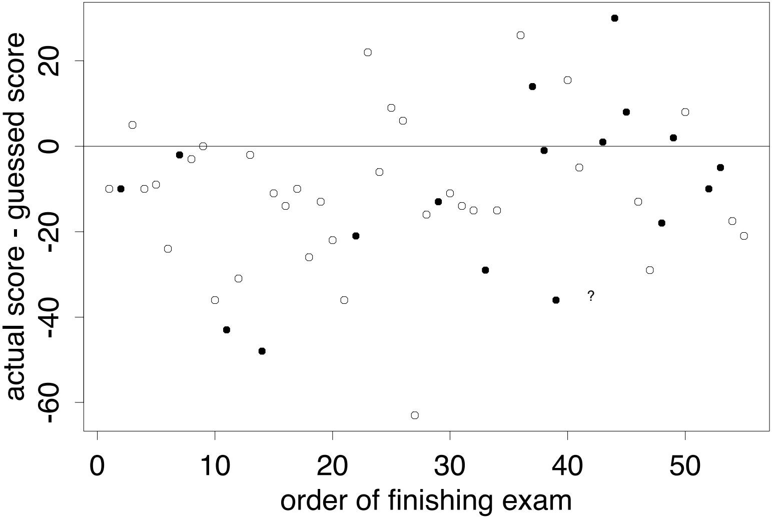
Class 12a

Story

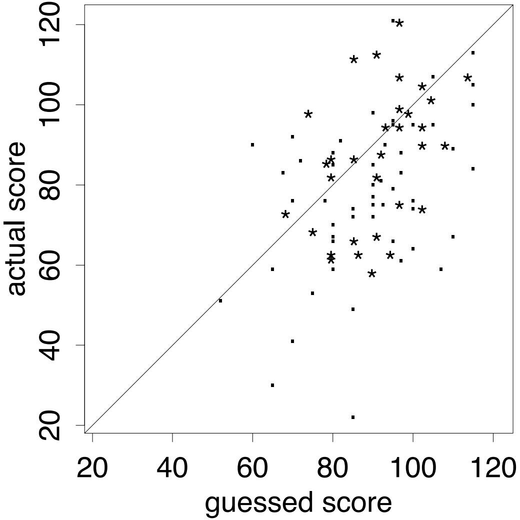
Actual vs. guessed exam scores



Actual vs. guessed exam scores



Actual vs. guessed exam scores



Activity

Sample size and statistical significance

Fit a regression to predict some outcome of interest from the General Social Survey, for example happiness or trust or behavior or attitude on some social or political issue.

Discuss reading and homework

Computer demonstration

Drill

Assumptions of regression and and how they can fail

Consider a regression fit to a set of different countries, predicting the rate of some illegal behavior (for example, tax evasion or speeding) given country-level predictors (per-capita income, average education level, etc.). For each assumption, give an example of how it can fail.

Discussion problem

Consider the implications of regression assumptions for a real-world study

1. Validity
2. Representativeness
3. Additivity
4. Linearity
5. Independence of errors
6. Equal variance of errors
7. Normality of errors

Class 12b

Story

Bill James does model checking

From Bill James: "Total Baseball has Glenn Hubbard rated as a better player than Pete Rose, Brooks Robinson, Dale Murphy, Ken Boyer, or Sandy Koufax, a conclusion which is every bit as preposterous as it seems to be at first blush.
To a large extent, this rating is caused by the failure to adjust Hubbard's fielding statistics for the ground-ball tendency of his pitching staff. Hubbard played second base for teams which had very high numbers of ground balls, as is reflected in their team assists totals. The Braves led the National League in team assists in 1985, 1986, and 1987, and were near the league lead in the other years that Hubbard was a regular. Total Baseball makes no adjustment for this, and thus concludes that Hubbard is reaching scores of baseballs every year that an average second baseman would not reach, hence that he has enormous value."

Activity

Assumptions of regression

1. Validity
2. Representativeness
3. Additivity
4. Linearity
5. Independence of errors
6. Equal variance of errors
7. Normality of errors

Discuss reading and homework

Computer demonstration

Drill

Explain how regression assumptions can be tested, using real-world examples

For each of the assumptions of regression, explain how it can be tested, using a real-world example.

Discussion problem

Patterns of residuals

Anna takes continuous data x_1 and binary data x_2 and creates fake data y from the model, $y = a + b_1x_1 + b_2x_2 + b_3x_1x_2 + \text{error}$, and gives these data to Barb, who, not knowing how the data were constructed, fits a linear regression predicting y from x_1 and x_2 and makes a plot of y vs. x_1, using dots and circles to display points with $x_2 = 0$ and $x_2 = 1$, respectively. The residual plot indicates to Barb that she should fit the interaction model. Sketch the residual plot that Barb could have seen when she fit the regression without the interaction.

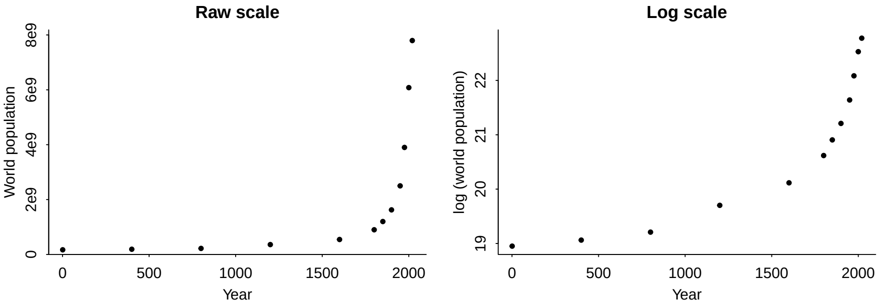
Class 13a

Story

Logarithm of world population

Year	Population	log (population)
1	170 million	18.95
400	190	19.06
800	220	19.21
1200	360	19.70
1600	545	20.12
1800	900	20.62
1850	1200	20.91
1900	1625	21.21
1950	2500	21.64
1975	3900	22.08
2000	6080	22.53
2020	7795	22.78

Logarithm of world population



Logarithm of world population

```
> fit <- stan_glm(log_pop ~ year, data=population)
> print(fit)

            Median  MAD_SD
(Intercept) 18.3    0.5
year        0.0     0.0

Auxiliary parameter(s):
        Median  MAD_SD
sigma   0.7     0.2
```

Logarithm of world population

```
> fit <- stan_glm(log_pop ~ year, data=population)
> print(fit)

            Median  MAD_SD
(Intercept) 18.3    0.5
year         0.0    0.0

Auxiliary parameter(s):
        Median  MAD_SD
sigma   0.7     0.2

> print(fit, digits=3)

            Median  MAD_SD
(Intercept) 18.276  0.489
year         0.002  0.000

Auxiliary parameter(s):
        Median  MAD_SD
sigma   0.728   0.163
```

Logarithm of world population

```
> population$year_1000 <- population$year/1000
> fit_2 <- stan_glm(log_pop ~ year_1000, data=population)
> print(fit_2)

            Median  MAD_SD
(Intercept) 18.3    0.5
year_1000   1.7     0.3

Auxiliary parameter(s):
        Median  MAD_SD
sigma   0.7     0.2
```

Activity

Predictive uncertainties

Predicting a “feeling thermometer” survey response

Discuss reading and homework

Computer demonstration

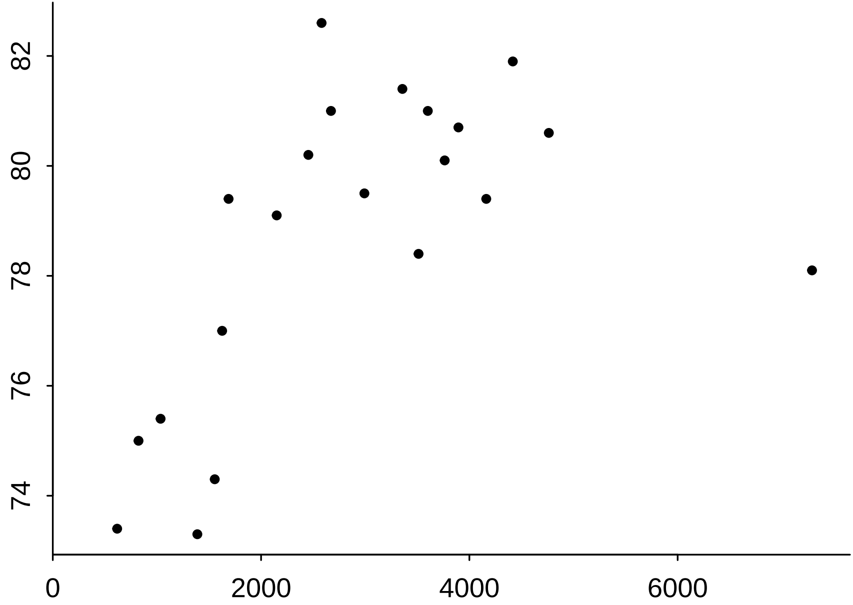
Drill

Log and antilog

Follow the instructions to either log or exponentiate the following expressions.

Discussion problem

General rules of when to use the log scale



Class 13b

Story

Price elasticity of demand

$$\log(\text{demand}) = a + b \log(\text{price})$$

Activity

Combining predictors to create a total score

1. Construct several questions to measure an underlying construct of interest
2. Use these to create a combined score

Discuss reading and homework

Computer demonstration

Drill

Examples of exponential growth and decline

Examples of exponential growth and decline

Discussion problem

Straight line fit to an exponential or power-law pattern

Fit a linear model to data that roughly follow exponential or power-law growth or decline. What happens?