

API-205
POLITICS AND POLICIES: WHAT CAN DATA TELL US?**Deborah Hughes Hallett**Office: B-122; phone: 617-495-9640
Deborah_Hughes_Hallett@Harvard.edu**Meetings: Tuesday and Thursday 2:45 pm- 4 pm in L-140****Review Session: Friday 2:45 pm-4 pm in L-140****Shopping: Tuesday August 29 from 2:45-3:15 pm and from 3:30-4 pm in L-140****Teaching Fellow**

Jack Sennett (Jack_Sennett@hks18.harvard.edu)

Course Assistants

Joshua Albert (joshua_albert@hks18.harvard.edu)

Jonathan Mazumdar (jonathan_mazumdar@hks18.harvard.edu)

Jimena Romero (jimena_romero@hks18.harvard.edu)

Niriksha Shetty (niriksha_shetty@hks18.harvard.edu)

Overview

Intended for decision leaders, this course provides a broad overview of the major concepts of statistics focusing on the critical interpretation of data and reports. The course aims to develop confidence in reporting on data and program evaluations. Topics includes experimental design, sampling, inference, multiple regression, and the contribution of “big data” Using case studies, the course asks what insight data can provide—and what it cannot—and compares the perspectives of statistics and ethics

Course Meetings

Lectures center on the development of concepts and their application to advocacy and policy making. Your perspective is warmly welcomed! Friday Review sessions and office hours provide the opportunity to ask questions and get started on assignments.

Do you need a text-book?

Lecture notes are available for each class and assignments are posted on the course page, along with the assignments. Some people find it helpful to have a book to see a second approach to topics; other do not. Before spending the money, make sure you need a book!

Textbooks: Two Possibilities

Statistical Reasoning for Everyday Life, 4th edition, by Jeffrey Bennett, William Briggs, Mario Triola. (Pearson, 2014, ISBN #0-321-81762-1.)

Student Solutions Manual for *Statistical Reasoning 4th edition*, by Bennett, Briggs, Triola (Pearson, 2014, ISBN #0-32181763-X.) The text and solutions manual are in the Coop and the HKS Library.

Or

The Practice of Statistics for Business and Economics, 4th edition, by David. S. Moore et al (Freeman, 2016, ISBN# 978-1-4641-3227-8).

What is the difference between the two main texts? You may not need a text; certainly *not* both.

Statistical Reasoning has fewer formulas and is more talkative. The second part of API-205 is not in *Statistical Reasoning*. *The Practice of Statistics for Business and Economics* goes further—but will be harder going at the start. Before choosing a text, look at both books at the Coop. Second hand versions of the books or older editions are available on line.

Source Readings: Optional

The source reading posted are all optional, but provide background for those who are interested. They will not be assessed.

Prerequisites

The mathematical background required is kept to a minimum, though there is some algebra and a calculator is needed. Data is given in Excel. The most important prerequisite is the willingness to ask questions and grapple with ideas in a policy context.

Case Studies and Examples

The course website contains data sets and additional readings that provide background for the case discussions and examples. These may include:

- Election fraud: Allegations and evidence in Albania, Iran, US, Venezuela
- Exit polls and polling regulation around the world
- Redistricting and Voter IDs in US: Voting Rights Act of 1965
- Regulating the fashion industry
- Value added measures of teaching and educational attainment
- Corruption: Identification, impact, relation to gender
- Education of Afro-Caribbean boys in UK
- Set-asides of university places in India: Education of underprivileged castes
- Legal cases in which statistics played a decisive role: Howland, Collins, Clark
- War deaths in Vietnam, Iraq, and D. R. Congo
- Human Development Index, World Bank development indicators and governance indicators
- Labor market discrimination in the US: New Haven firefighters; NCAA football coaches
- Workers' protection: Denim and sandblasting
- Roadside stops and racial profiling: New York and Los Angeles
- Madrassa enrollment in Pakistan; impact of the Hajj
- Cross-country relationships: Growth, taxes, debt, wealth, and happiness
- Taxes and debts: Cross country comparisons
- Medical testing policy: HIV, pap smears, and mammograms
- The death penalty: Deterrence and justice
- Measles, mumps, rubella (MMR) vaccine controversy
- Malaria: Impact and prevention
- Voting patterns of women judges: Sonia Sotomayor's Supreme Court nomination
- IQ and immigration
- Randomized Control Trials (RCTs): Dr. Frances Kelsey and thalidomide; development policy
- Discovery of the cause of cholera
- Ethical issues in Statistical Experiments: HIV in Tanzania, Syphilis, Tearoom Trade
- Press Freedom: Reporters without Borders

Assignments

There will be eight written **assignments** due Sept 15, 21, 28, Oct 5, 12, 19, Nov 2, 9. You are welcome to discuss assignments with the course staff or other students. Each student must write his or her own solutions individually. Solutions may not be shared; sources and collaborators should be listed.

Quizzes and Assessments

There will be eight online **quizzes**.

There will be one **progress check** in class Thursday October 26 and an *optional final assessment* on Wednesday December 6 from 9-12 noon. In preparation for each assessment, sample materials will be available.

Project and Presentation

There will be an *optional final project* showing an application of the course material to a professional or policy context. This is an opportunity you to research a topic that you care about, find and analyze data, and make a presentation. The project can be done in groups. The presentation will be made to the class on Nov 30 or filmed and posted on YouTube on Nov 28-29.

You must complete *at least one* of the **final assessment** or the **final project** (you are also welcome to do both)

Grading

The lowest two assignment scores and the lowest two quiz scores will be dropped. If the final assessment score is higher than the progress check score, it will replace the progress check.

Final grades will be based on the highest score generated by the following three methods:

Assessment	Method 1	Method 2	Method 3
Assignments (in total)	20%	20%	20%
Quizzes (in total)	20%	20%	20%
Progress check	20%	20%	20%
Final presentation	40%	20%	
Final assessment		20%	40%

Letter grades are determined in accordance with the HKS recommended grade distribution:

<http://www.hks.harvard.edu/degrees/registrar/faculty/exams-and-grading/grades>

Course Schedule:

<u>Date</u>	<u>Lecture Topic</u> <u>Reading:</u> SR is <i>Statistical Reasoning</i> SBE is <i>Statistics for Business and Economics</i>	<u>Assignments</u>
Week 1 Tue Aug 29	<u>Shopping:</u> “Let the Numbers Speak”: Goals of API-205 (Source unclear: http://stats.stackexchange.com/questions/91789/who-said-let-the-data-speak-for-themselves)	<u>Shopping</u> <u>Location:</u> L-140 2:45-3:15 pm and 3:30 –pm
Thu Aug 31	<u>Lecture #1: Assessing a Statistical Study</u> Components of a Statistical Study: Framing, Collecting, Concluding. Pitfalls and Pushback New directions: Visualization and Big Data <u>Cases:</u> <i>Chlorine pills and diarrhea; Bribes for government services; Madrassa enrollments; Cell phones and driving; Racial profiling at airports; Mathematical performance and gender; Laptops in the classroom; Exit polls</i> <u>Reading:</u> SR Text: Section 1.1, 1.4 SBE Text: Section 1.1 <u>Sources:</u> Syria: http://www.syria-visualized.com/ “The Promises and Limitations of Big Data” HBS May 5, 2017	
Week 2 Mon Sept 4 Tue Sept 5	<u>Labor Day: Holiday</u> <u>Lecture #2: How Should Data be Collected?</u> Sampling and Data Collection: Random, Stratified, and Cluster Samples. Sampling frame; Internal and External Validity. <u>Examples:</u> <i>Calling an Election (1948), Current Population Survey, Household surveys</i> <u>Case Study:</u> <i>Madrassa enrollment in Pakistan; War deaths in Vietnam, Iraq, D.R. Congo</i> <u>Reading:</u> SR Text: Section 1.2. SBE Text: Section 3.1 <u>Sources:</u> “The Human Cost of War” by Les Roberts. <i>The Lancet Medical Journal</i> “Fifty years of violent war deaths from Vietnam to Bosnia: analysis of data from the world health survey programme” by Ziad Obermeyer, <i>British Medical Journal</i>	

“Religious School Enrollment in Pakistan: A Look at the Data” December 2005, by T. Andrabi, J. Das, A. Khwaja, T. Zajonc

“Mortality in the Democratic Republic of Congo: a nationwide survey”, by B. Coghlan, et al, *The Lancet*, 2006.

Thu Sept 7

Lecture #3: Experimental Design

Observational Studies vs. Experiments.

Placebos, Controls, Hawthorne Effect, Confounding. Randomization and Natural Experiments.

Ethical Issues in Experimentation.

Cases and Examples: Vaccination Policy; Hormone Replacement Therapy, Physicians' Health Study, Video-games and Disease, Identification of Cholera Source, Circumcision, MMR vaccine, Tuskegee Syphilis Study, Tearoom Trade, Tanzania HIV Study

Reading: SR Text: Sections 1.3, 1.4.
SBE Text: Section 3.1

Sources:

“America's Toe-Tapping Menace” by Laura MacDonald, *New York Times*, Sept 2, 2007

“Statistical Models and Shoe Leather”, by David A. Freedman, *Sociological Methodology*, 1991

“No Link Between MMR Vaccine and Autism” Chris Guillot, N.C. Journal of Law and Technology, Feb 18, 2009

“Italy makes 12 vaccinations compulsory for children” BBC, May 19, 2017

“Estimating the Impact of the Hajj: Religion and Tolerance in Islam’s Global Gathering” D. Clingingsmith, A. Khwaja, M. Kremer, 2008

“Male Circumcision”, American Academy of Pediatrics, Aug 27, 2012

“Cultural Bias in the AAP’s 2012 Technical Report and Policy Statement on Male Circumcision” *Pediatrics*, April 2013

“Harvard Researchers Accused of Breaching Students’ Privacy”, *The Chronicle of Higher Education*, July 10, 2011

“ “But the data is already public”: on the ethics of research in Facebook”, M. Zimmer, *Ethics Information Technology*, June 4, 2010

Week 3

Lecture #4: Central Limit Theorem for Proportions

Tue Sept 12

Sampling Distribution of Proportions, Central Limit Theorem. Normal distributions

Assignments

#1 available

Case Study: Albania’s mayoral election, May 2011; Venezuela’s 2004 referendum on the recall of Hugo Chavez. New Haven firefighters; coaches for US college football

Reading: SR Text: Section 8.1.
SBE Text: Section 3.3 and 8.1 (page 458)

Sources:

“Who You Know & Who Knows You: The Hiring Process & Practice of NCAA FBS & FCS Head Coaching Positions” Keith Harrison, *Black Coaches and Administrators*, 2008

Thu Sept 14 Lecture #5: Central Limit Theorem for Means
Sampling Distribution of Means, Central Limit Theorem. Data types. Normal distributions

Example: *Repeated sampling; Baby weights: Russia, Texas, Indonesia, Spain*

Reading: SR Text: Section 5.3, 8.1.
SBE Text: Section 4.4

Review on Normal Distribution:

SR text: Sections 5.1, 5.2.
SBE text: Sections 1.4

Week 4 Lecture #6: Probability
Tue Sept 19 Probability, Conditional Probability, and Equity

Assignment

#1 due
#2 available

Policy Questions: *Vitamin A Study; Should Afro-Caribbean boys in UK be educated separately? Should Set-Asides in Indian Colleges be Expanded, US Voter ID Laws*

Reading:

SR Text: Sections 6.1, 6.2, (to mid p.203), 6.3 (p.208-mid p. 209, p.211-212)
SBE Text: Section 4.1, 4.2

Sources:

“Bad attitudes; Race and education” *Economist*, Mar 10, 2005

“Tackling low educational achievement” by Robert Cassen and Geeta Kingdon, Joseph Rowntree Foundation, 2007

“Under-Representation of Disadvantaged Classes in Colleges: What Do the Data Tell Us?” by Rana Hasan and Aashish Mehta. *Economic and Political Weekly*, Sept 2, 2006, 3791–3796

Thu Sept 21 Lecture #7: Independence and Statistical Significance
Independence, *P*-values, and Significance Levels

Examples: *Age and Gender, Sally Clark, Vitamin A Study, Polls,*

Reading: SR Text: Section 4.4 and 6.1, 6.5 (p. 224).
SBE Text: Section 5.1, 5.2

Sources:

“Like Mother, Like Daughter? An Economic Comparison of Immigrant Mothers and Their Daughters” Gary Smith, Margaret Hwang Smith, *International Migration*, 2013

“US Supreme Court to hear Ohio voter registration case” May 30, 2017. *Yahoo News*

Week 5

Tue Sept 26

Lecture #8: Estimation of Means

Confidence Intervals: Means. Margin of Error and Sample Size.

Assignments

#2 due

#3 available

Example: *Radar detectors and speeding*

Case Discussion: *Civilian Deaths in Iraq: Lancet studies. War Deaths in Vietnam*

Reading: SR Text: Section 8.2

SBE Text: Section 6.1

Sources:

“The Human Cost of War” by Les Roberts. *The Lancet Medical Journal*

“Fifty years of violent war deaths from Vietnam to Bosnia: analysis of data from the world health survey programme” by Ziad Obermeyer, *British Medical Journal*

Thu Sept 28

Lecture #9: Estimation of Proportions

Confidence Intervals: Proportions. Margin of Error and Sample Size.

Examples: *2011 Venezuelan polls, Presidential approval ratings*

Reading: SR Text: Section 8.3

SBE Text: Section 8.1

Week 6

Tue Oct 3

Lecture #10: Testing One Sample Mean

Hypothesis Testing: Means: One Sample. Practical and Statistical Significance

Assignments

#3 due

#4 available

Example: *Malaria and drug testing; slide in SAT scores*

Reading: SR Text: Sections 9.1–9.2

SBE Text: Sections 6.2 and 7.1

“Frances Oldham Kelsey, FDA scientist who kept thalidomide off U.S. market, dies at 101”, Aug 7, 2015, *Washington Post*

“J&J forms ‘compassionate use’ panel for experimental drugs”, May 8, 2015, *Aljazeera America*

Thu Oct 5

Lecture #11: Testing Two Sample Means

Hypothesis Testing: Means: Two samples

Examples: *Do poets die young? Do women talk more than men? Did the Washington Consensus work? Comparison of growth rates of developing countries; Rule of Law and development*

Reading: SBE Text: Section 7.2

Sources:

“Goodbye Washington Consensus, Hello Washington Confusion?”, by Dani Rodrik (2006) p. 1-2

“Are Women Really More Talkative Than Men?” Matthias Mehl et al *Science*, July 6, 2007

Week 7
Mon Oct 9

Columbus Day: Holiday

Tue Oct 10

Lecture #12: Redistricting and Gerrymandering with Colleen Mathis, Chair of Arizona Independent Redistricting Committee

Assignments

#4 due

#5 available

Case Discussion: *Redistricting in Arizona and the cases before the US Supreme Court*

Sources:

“Gerrymandering is the biggest obstacle to genuine democracy in the United States. So why is no one protesting?” *Washington Post*, Feb 10, 2017

“Analysis indicates partisan gerrymandering has benefited GOP” David A. Lieb, *AP*, June 25, 2017

“This is actually what America would look like without gerrymandering” *Washington Post*, Jan 13, 2016. Includes video.

“This is the best explanation of gerrymandering you will ever see” *Washington Post*, March 1, 2015

“The research that convinced SCOTUS to take the Wisconsin gerrymandering case, explained” Nicholas O. Stephanopoulos, *Vox* July 11, 2017

“Partisan Gerrymandering and the Efficiency Gap” by Nicholas O. Stephanopoulos & Eric M. McGhee *University Chicago Law Review*, Oct 1 2014

“How the Efficiency Gap Works” Eric Petry, *Brennan Center for Justice*

Thu Oct 12

Lecture #13: Testing One Sample Proportions
Hypothesis Testing: Proportions: One Sample

Examples: *Hispanics and Jews on Juries; Army suicide rates; testing a HIV vaccine*

Case Study: *Madrassa Enrollment in Pakistan; Kissing study*

Reading: SR Text: Section 9.3

SBE Text: Section 8.1

“Pakistan’s Madrassas, Extremism and the Military,” Crisis Group, July 29, 2002, July 15, 2005

“Religious School Enrollment in Pakistan: A Look at the Data” December 2005, by T.Andrabi, J. Das, A. Khwaja, T.Zajonc

Week 8
Tue Oct 17

Lecture #14: Testing Two Sample Proportions
Hypothesis Testing: Proportions: Two samples

Assignments

#5 due
#6 available

Case Studies: *NYPD stop and frisk; LAPD roadside search; Labor market discrimination; Miller-El vs. Cockrell*

Reading: ES Text: Chapter 21, p. 420
SBE Text: Section 8.1, 8.2
Summary sheet for formulas

“Judge Rejects New York’s Stop-and-Frisk Policy” *New York Times*, Aug 12, 2013

“Driving While Black in the City of Angels” by Lily Khadjavi, in *Chance*, 2006

“Surprising New Evidence Shows Bias in Police Use of Force but Not in Shootings”
Q. Bui and A. Cox, *New York Times*, July 11, 2016

“An Empirical Analysis of Racial Differences in Police Use of Force”
Roland G. Fryer, Jr, July 2016, NBER 22399

“Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination” by Marianne Bertrand and Sendil Mullainathan, *American Economic Review*, 2004

“Discrimination against Female Migrants wearing Headscarves” Doris Weichselbaumer, *Institute for the Study of Labor (IZA)*, September 2016

Thu Oct 19

Lecture #15: Bivariate Regression
Least Squares Regression with One Explanatory Variable, Prediction, and Hypothesis Testing

Examples: *Cigarettes and lung cancer death*

Case Study: *Social cohesion between Jews and Arabs in Haifa, Israel. The 2002 Alabama gubernatorial election*

Reading: SR Text: Section 7.1, 7.2, 7.3
SBE Text: Section 2.3, 10.1

“Positive prejudice: Really loving your neighbour” *Economist*, 2007

“Allophilia” by Todd Pittinsky, in *Compass*, 2005

“Strange Case of an Imprisoned Alabama Governor” *New York Times*, Sept 10, 2007

“A Statistical Analysis of Possible Electronic Ballot Box Stuffing; The Case of Baldwin County Alabama Governor’s Race in 2002”, by James Gundlach, April 11, 2003.

Week 9 Tue Oct 24	<p><u>Lecture #16: with Christopher Robichaud, Lecturer in Ethics, HKS</u> The Search for Causation: Correlation and Causation; Confounding.</p> <p><i>Policy Question: Should We Have the Death Penalty?</i></p> <p><i>Reading:</i> SR Text: Section 7.1, 7.2, 7.4 SBE Text: Sections 2.1, 2.1, 2.4</p> <p>“Does Death Penalty Save Lives?” <i>NY Times</i>, Nov 18, 2007 (<i>plus links</i>)</p> <p>“The State of the Death Penalty: Evaluating Fairness and Accuracy in State Death Penalty Systems” American Bar Association, Oct 2007</p>	<u>Assignments</u> #6 due
Thu Oct 26	<i>Progress Check</i>	
Week 10 Tue Oct 31	<p><u>Lecture #17: Multiple Regression</u> Regression with Several Explanatory Variables and Hypothesis Testing</p> <p><i>Examples: Cross country comparison of Central Bank independence and inflation; Voice pitch and reproductive success</i></p> <p><i>Reading:</i> SBE Text: Section 11.1, 11.2</p>	<u>Assignments</u> #7 available
Thu Nov 2	<p><u>Lecture #18: Multiple Regression with Dummy Variables</u> Regression and Hypothesis Testing: Dummy (Binary) Variables. Reading a Regression Table; Interpreting Coefficients</p> <p><i>Reading:</i> SBE Text: Section 11.3</p>	
Week 11 Tue Nov 7	<p><u>Lecture #19: Program Evaluation</u> Regression Coefficients and Significance</p> <p><i>Case Discussion: Deworming and Education</i></p> <p><i>Reading:</i> “How to Read a Regression Table” G. Meirer, J. Rauch, in <i>Leading Issues in Economic Development</i>, Oxford, 2005</p>	<u>Assignments</u> #7 due #8 available
Thu Nov 9	<p><u>Lecture #20: Reading a Regression Table with Dummy Variables</u> Reading a Regression Table: Interpretation of Coefficients; Log & Dummy Variables</p> <p><u>Discussion:</u> What is the impact of International Criminal Court (ICC) cases?</p> <p>“Unintended Positive Complementarity: Why International Criminal Court Prosecutions May Increase Domestic Human Rights Prosecutions” G. Dancy and F. Montal, Oct 30, 2015 http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2736519</p> <p><u>Discussion:</u> What has been the long-term impact of the slave trade?</p> <p>“The Long-Term Effects Of Africa’s Slave Trades”, Nathan Nunn, <i>The Quarterly Journal of Economics</i>, February 2008</p>	

Fri Nov 10	<i>Veteran's Day Holiday</i>	
Week 12	<u>Lecture #21: Regression Table with Log and Dummy Variables</u>	<u>Assignments</u>
Tue Nov 14	Reading a Regression Table: Interpretation of Coefficients: Log & Dummy Variables	#8 due
	<u>Discussion:</u> What is the impact of employment on women's fertility?	
	"Jobs and Kids: Female Employment and Fertility in China" H. Fang, K. Eggleston, J. Rizzo, R. Zeckhauser, Faculty Research Working Paper, HKS, Nov 2012.	
Thu Nov 16	<u>Lecture #22: Regression Table with Log Variables</u>	
	Reading a Regression Table: Interpretation of Coefficients; Logarithmic Variables	
	"The Impact of Public Officials' Corruption on the Size and Allocation of U.S. State spending" by C. Liu, J. Mikesell, <i>Public Administration Review</i> , May/June 2014	
	"Focus on Corruption", Aaron Ausland and Alfonso Tolmos, Second Year Policy Analysis Paper, HKS-MPAID, 2005	
Week 13	<u>Lecture #23: Regression Table with Lagged Variables</u>	
Tue Nov 21	Reading a Regression Table: Interpretation of Coefficients; Dummy Variables	
	<u>Case Discussion:</u> <i>What is the impact of value added metrics for teachers?</i>	
	<u>Source:</u> "The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood" Raj Chetty, John N. Friedman, Jonah E. Rocko, Dec 2011	
	"The Case for \$320,000 Kindergarten Teachers" <i>New York Times</i> , July 27, 2010	
Thu Nov 23	<i>Thanksgiving Holiday</i>	
Week 14	<u>Lecture #24: Regression Table with Log and Dummy Variables</u>	
Tue Nov 28	Reading a Regression Table: Interpretation of Coefficients; Logarithmic Variables	
	<u>Case Discussion:</u> <i>Are women less corrupt than men?</i>	
	"Gender and Corruption" by Anand Swamy, Stephen Knack, Young Lee, and Omar Azfar, <i>Journal of Development Economics</i> , Volume 64, Issue 1, Feb 2001, pp. 25-55	
Thu Nov 30	Final Student Presentations	
Wed Dec 6 9-12 noon	Final Assessment	