HARVARD UNIVERSITY
SPRING 2019

Sociology 90EI

Research Lab: Ethnicity and Immigration

Thursday 9:00–11:00
William James Hall 601

Instructor: Dr. Hilary J. Holbrow
Email: hholbrow@fas.harvard.edu
Office: 616 William James Hall
Office Hours: Wednesday 2pm-4pm (please book at http://holbrow.youcanbook.me/)

COURSE DESCRIPTION

Students will build skills in conducting original quantitative data analysis, using the Children of Immigrants Longitudinal Study (CILS) and the Immigrant Second Generation New York (ISGMNY) survey. These datasets cover children of immigrants from more than 77 countries, with particular concentrations of those from Central America, the Caribbean, East Asia, and Southeast Asia. Students may use these data to study questions about the relationships between national, racial, ethnic, family background on one hand, and attitudes and outlook, school performance, and socio-economic attainment on the other. While exploring how ethnicity and national background shape young people's opportunities and experiences, students will also gain experience with data analysis, data visualization, and the interpretation of statistical analyses.

This is a small, hands-on class. I may modify the syllabus to better fit students' needs and interests.

COURSE REQUIREMENTS

Class Attendance and Participation 30%

Weekly Assignments 30%

Final Project Outline 5%

Presentation 15%

Results Write-Up 20%

COURSE READINGS

Students should purchase *Data Visualization: A Practical Introduction* by Kieran Healy. (https://www.amazon.com/gp/offer-listing/0691181624). We will be reading it in March, so please make sure you have it by the end of February. All other readings will be made available on the course website.

COURSE POLICIES

Grading: Final Project Outline, Presentation, and Results Write-Up will be graded on a letter-grade basis. Weekly assignments will be graded out of 6 points. Attendance and Participation in weekly class meetings will also be graded out of 6 points.

Late Assignments: Late weekly assignments will receive half-credit if submitted within 48 hours of the deadline. Weekly assignments received more than 48 hours after the deadline will not receive credit. Letter-graded assignments (Project Outline, Presentation, and Results Write-up) will be marked down by one grade step for each 24 hour period past the deadline the paper is submitted (i.e. an A paper submitted within 24 hours past the deadline will receive an A-; an A paper submitted between 24 and 48 hours after the deadline would receive a B+).

Collaboration Policy: You are encouraged to discuss assignments and projects with other students, but your assignments must reflect your own thinking and writing. Students may share Stata and R code. However, if you share code, must write the names of students with whom you shared code (or from whom you received code) at the top of any code-based assignment. When you submit Stata or R code, you should comment on any section of the code on which you collaborated with other students, shared with other students, or received from other students. Comments should include the names of collaborators and the nature of the collaboration (e.g. shared code with Charles; received code from Maria; collaborated on code with Yuqi).

COURSE SCHEDULE

Week 1 (January 31) Introducing the Datasets & the Course

Week 2 (February 7) Major Debates about the Immigrant Second Generation

To do before class:

- Course readings (see below)
- Weekly assignment (see below)

Readings:

- Haller, William, Alejandro Portes, and Scott M. Lynch. 2011. "Dreams Fulfilled, Dreams Shattered: Determinants of Segmented Assimilation in the Second Generation." *Social Forces* 89 (3): 733-762.
- Alba, Richard, Philip Kasinitz, and Mary C. Waters. 2011. "Comments on Haller, Portes, and Lynch: The Kids are (Mostly) Alright: Second-generation Assimilation." *Social Forces* 89 (3): 763-773.
- Haller, William, Alejandro Portes, and Scott M. Lynch. 2011. "On the Dangers of Rosy Lenses." *Social Forces* 89(3): 775-782.

Submit at the start of class:

- Weekly Assignment: A 2-3 paragraph summary of the debate between HPL and AKW. What questions are they interested in? How are their answers similar or different?

In class:

- Discuss readings. Interpreting an OLS regression table.

Week 3 (February 14) Forming Research Questions and Hypotheses

To do before class:

- Course readings (see below)
- Weekly assignment (see below)
- Create ICPSR account; remember username and password!

Readings:

- Luthra, Renee, Thomas Soehl, and Roger Waldinger. 2017. "Reconceptualizing Context: A Multilevel Model of the Context of Reception and Second-Generation Educational Attainment." *International Migration Review.* https://doi.org/10.1111/imre.12315.
- Feliciano, Cynthia. 2017. "How Family, Immigrant Group, and School Contexts Shape Ethnic Educational Disparities." *Ethnic and Racial Studies* 41 (2): 189-209.

Submit at the start of class:

- Weekly Assignment: Both LSW and Feliciano are *primarily* interested in how characteristics of immigrant groups and their social context shape educational attainment of the second generation (as opposed to how family or individual-level factors shape attainment). Please describe which characteristics of immigrant groups and their social context LWS and Feliciano think are important, and how they empirically measure the concepts of interest. (This can take the form of a list or table if you prefer). Briefly compare and contrast the conclusions they draw.

In class: Download datasets; discuss readings, with focus on the relationships between concepts and measurements; discuss how scholars develop research questions; brief overview of how to hunt for academic articles.

Week 4 (February 21) Developing our own Research Questions

To do before class:

- Browse CILS & ISGMNY Codebooks.
- Identify 5 research topics you might use the CILS and/or the ISGMNY to pursue; find and read 3 academic articles on one of your topics; come prepared to discuss your topics and summarize the articles to your classmates. (This will be graded as participation.) Your weekly assignment will be the hard copy of the list of the research topics and list of three citations related to one of these questions/topics.

Readings: (see above)

Submit at the start of class:

- Weekly assignment: List of topics (at least 5) you might use the CILS/ISGMNY to pursue, along with the bibliographic references of the sources you read (bring copies for everyone).

In class: Each student will share their list of research questions and summarize the articles they read on the topic, followed by discussion of how to use CILS/ISGMNY to purse that topic.

Week 5 (February 28) Basic Data Manipulation in Stata

To do before class:

- Download and install StataMP on your computer. (Available at: https://downloads.fas.harvard.edu/download)
- Weekly assignment (see below)

Readings:

- Acock, Alan C. 2018. *A Gentle Introduction to Stata*. College Station, TX: Statacorp. pp. 51-121.

Submit at the start of class:

- Weekly assignment: Draft of the research questions you will pursue for the final project, with 1-2 paragraphs explaining why these questions are interesting and important; should include at least 3 citations; a list of variables you will be working with and why you plan to include these variables, and what concepts you are trying to capture with particular variables.

In class: Working with Stata; summarizing data, generating new variables, etc.

Week 6 (March 7) Basic Data Visualization with R

To do before class:

- Download and install R on your computer. (Available at: https://cran.r-project.org/bin/windows/base/)
- Download and install RStudio.
- Course Readings (see below)
- Weekly assignment (see below)

Readings:

- Wickham, Hadley and Garrett Grolemund. 2017. *R for Data Science*. Sebastabol, CA: O'Reilly Media. pp3-43.
- Healy, Kieran. 2018. *Data Visualization: A Practical Introduction*. Princeton, NJ: Princeton University Press. Chapters 1 and 3.

Submit at the start of class:

- Weekly assignment: Create a table of descriptive statistics related to your project, along with a short write-up of patterns you find interesting, and your ideas about how those patterns emerge. Include any Stata code you used to get the variables into a workable format.

In class: Working with R.

Week 7 (March 14) - Advanced Data Visualization with R (I)

To do before class:

- Course Readings (see below)
- Weekly Assignment (see below)

Readings:

- Healy, Kieran. 2018. *Data Visualization: A Practical Introduction*. Princeton, NJ: Princeton University Press. Chapters 4 and 5.

Submit at the start of class:

- Weekly assignment: Improved descriptive stats tables.

In class: Working with R, again.

SPRING BREAK (March 21)

Week 8 (March 28) – Advanced Visualization with R (II)

To do before class:
- Weekly Assignment (see below)

Readings:
- No new readings. Review Healy chapters from last week.

Submit at the start of class:
- Weekly assignment: Prepare 5 data visualizations that show important or interesting features of your data (related to your research question). Try to include different types of visualizations (e.g. scatterplots, distributions, bar charts.) Use the Healy book and/or online resources to add appropriate axis labels. Write a short paragraph about each graph. What is the most striking pattern it reveals? What narrative does the graph suggest to you? Are there any issues (such as non-linear relationships in scatterplots) or highly uneven distributions across groups (from bar charts) that you notice?

In class: Refining our data visualizations.

Week 9 (April 4) Regression 1 - Conducting and Interpreting Regressions

To do before class:
- Readings (see below)
- Weekly assignment (see below)

Readings:
- Vickers, Andrew. 2010. *What is a P-value Anyway*? Boston, MA: Pearson Education. pp. 54-89.
- Optional reading (a more technical discussion of regression): Angrist, Joshua D. and Jörn-Steffen Pischke. 2015. *Mastering 'Metrics*. Princeton, NJ: Princeton University Press. pp47-97.

Regression Resources (These are useful references; refer to them as necessary for Week 10's weekly assignment):
- OLS regression: "Multiple Linear Regression Analysis." (https://statistics.laerd.com/stata-tutorials/multiple-regression-using-stata.php)
- Interactions & categorical variables: "Working with Categorical Data and Factor Variables." (Course site)
- Logistic regression: "Logistic Regression | Stata Data Analysis Examples" https://stats.idre.ucla.edu/stata/dae/logistic-regression/.

Submit at the start of class:
- Weekly assignment: 3 updated data visualizations, made more beautiful and clear.

In class: Interpreting regressions; using Stata to run regressions; using interactions & categorical variables.

Week 10 (April 11) Regression 2 - Building Models

To do before class:

- Readings (see below)
- Weekly Assignment (see below)

Readings:

- Elwert, Felix and Christopher Winship. 2014. "Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable." *Annual Review of Sociology*: 31-53.

Submit at the start of class:

- Weekly assignment: A short summary, in your own words, of the rules of conditioning in regressions, as described in the reading, along with questions about things you found confusing. Also, please prepare a regression table based on your project.

In class: Drawing causal graphs; improving regression models.

Week 11 (April 18) - Visualizing Regression Results

To do before class:

- Weekly Assignment (see below)
- Course readings

Readings: None.

Submit at the start of class:

- Weekly Assignment: Draw a causal graph or graphs associated with your project. List which variables from the graph you will include as conditioning variables and which you will exclude. If any are excluded, explain why in 2-3 sentences.
- Submit your final project outline.

In class: From regression output to effective visual presentation of results.

Week 12 (April 25) Student Presentations

Final Project Write-Up Due May 6 @ 5pm