Syllabus
Research Design for Social Data Analytics
SoDA 308
Fall 2020
T/R 9:05-10:20am, ZOOM

Instructor: Omer F. Yalcin

Office:	Zoom link in Canvas
Email:	oxy4@psu.edu
Web Page:	omerfyalcin.com
Office Hours:	W 9–10, R 3–4 & by appt.

Course Description: This course engages students in the study and use of research design tools for the analysis of social systems and "big data". Topics to be addressed include: how the scientific method relates to a practice of establishing the validity of propositions and the role that analytics can play in that process when the observations are vast and varied; how the validity of systematic patterns in data are assessed as well as how spurious or biased patterns in the data are ruled out; and how the scientific method can guide the use of exploratory techniques. Through the course, students will learn to develop innovative research designs in an effort to improve the statistical analyses used with social data and how to present these analyses to nontechnical audiences, such as non-profits, employers, and the general public.

Prerequisites: PLSC 309 or equivalent training in statistical analysis

Learning Objectives: There are four related learning goals that underpin the design and content of this course. Students who successfully complete the course will develop:

- A mastery of the process of designing research to quantitatively evaluate causal claims about the social world.

- A familiarity with a diverse set of examples of social research designs, especially those investigating the digital social world.

I would like to thank Bruce Desmarais for permitting me to adopt portions of his syllabus.

- A competence in using several critical computational tools for carrying out social data analytics projects.

- A capacity to convey the motivations, designs, and results of projects to non-specialists.

Required Book:

- Title: Social Scientific Research

- Author: Dawn Brancati

- Publisher: SAGE, 2018

- ISBN: 1526452790, 9781526452795

Penn State honors and values the socioeconomic diversity of our students. If you require assistance with the costs of textbooks for this course, contact the Office of Student Care and Advocacy (129 Boucke Building, 814-863-2020, http://studentaffairs.psu.edu/studentcare). For additional need related to socioeconomic status please visit http://sites.psu.edu/projectcahir.

Other readings: The "Bit by Bit" book Chapter 2 is available online at https://www.bitbybitbook.com/en/1st-ed/observing-behavior/. All the other readings are provided in Canvas.

Software: The software of the course is the free and open source R statistical programming software. It is required that all problem sets and analyses reported in the research paper be completed using R .

Article component presentation: Each Thursday beginning on 9/10, the first part of class will consist of up to six 5–10min student presentations. Each student is responsible for delivering three presentations over the course of the semester. Each of the three presentations must be delivered on a different day, and cover a different article component (i.e., students may not give more than one presentation on a given day or on a given article component). Each presentation will provide a brief overview of a particular component of the research article that the class reads for that day. The article components are drawn from the Huck reading (assigned for 8/27), and are listed below. A PDF version of the slides to be presented must be uploaded to a Canvas discussion thread by the student

by 11pm on the Tuesday before the presentation date. Students should sign up for their presentations on SignUpGenius, at the following link https://www.signupgenius.com/go/10C0D4FAEA92FA4F9CF8-thursday, by class time on Thursday, 9/3.

- Background (see Huck)

- Purpose & Hypotheses (see Huck)

- Participants & Materials (see Huck)

- Statistical Plans & Results (see Huck)

- Applying Brancati (or Salganik): Commentary on the ways in which the reading from Tuesday applies to the article.

- Critique: Discussion of limitations, shortcomings, or possible errors in the research.

Research Project: Students are required to complete a group research paper by the end of the semester. Groups will be assigned on the basis of the article interest survey/quiz. Please complete this on Canvas no later than Tuesday, 9/1. The paper is expected to be original research—applying the research design principles covered in the course. The final paper is due 12/16

- **Process**: Groups should make progress on the project each week. The files associated with the paper should be shared among the students in the group and the instructor using a folder on box.psu.edu. It is possible to assign tasks related to a file to individual collaborators in Box. Each week during the group project time, the group should assign a task, to be completed over the next week, to each student in the group. Each student is responsible for making sure they have a task to work on, and complete that task, each week.

- **Topic**: The project can address virtually any topic in which students are interested. Considering Brancati's discussion of research topics, this document should include a brief discussion of why this is a good topic. There are two constraints in terms of a hypothesis. First, students must define at least one hypothesis that constitutes an "effect of a cause" statement. Second, the hypothesis should be amenable to a design that can be implemented before the end of the semester. The document should include a brief justification for your hypothesis. Please provide proper citation/reference for any literature discussed in the topic document. The topic of the research project,

including the hypotheses, should be written up in a one-page document. This is due on 9/25.

- **Design**: Students should write up a research design. A research design will consist of complete introduction, literature review, theory and research design sections, with the results yet to come. This is due on 10/23.

Research tool tutorial: Each student is responsible for writing a 3–5 page appendix to the group research project in which a data analysis tool/method that is used in the group project is explained. The tutorial should include a definition and overview of the tool/method, example code, a data analysis example that is separate from the group project (i.e., doesn't use any of the same data), a link to an online archive (e.g., Box, Google Drive, Dropbox, Github) at which the example data can be downloaded, and a discussion of the sorts of research projects and objectives for which the tool could be used. Students will submit a one-page outline of the tutorial by 11/6, and the final product is due on 11/20.

Pop quizzes: There will be 4 or 5 pop quizzes administered throughout the semester. Each quiz will include a combination of multiple choice and true/false questions covering material that has previously been covered in lectures. Each student's lowest quiz grade will be dropped. If you miss a quiz day with an excused absence, you will be given the opportunity to make up the quiz. The objective of these quizzes is to assure that students are keeping up with the lecture material.

Grading:

- Article component presentations–30%

- Research tool tutorial–20%

- Group project–30%

- Pop quizzes–20%

Grading Scale. Grade values will not be rounded - any grade that is greater than or equal to 'Lower' and less than 'Upper' will receive the respective grade.

Grade	Lower	Upper
A	92	101
A-	90	92
B+	88	90
B	82	88
B-	80	82
C+	78	80
C	72	78
C-	70	72
D+	68	70
D	62	68
D-	60	62
F	0	60

Course Schedule: The schedule below gives the required reading. The date corresponds to the week beginning on the respective Monday. The readings listed for a particular day should be read before class time that day.

1. 8/24, Reading a scientific study

   - Tuesday: Syllabus/course overview
   - Thursday: Chapter 1, "Reading Statistics and Research," Schuyler W. Huck.

2. 8/31, Social Research, Definition and Ethics

   - Tuesday: Brancati, Chapter 1 & 2
   - Thursday: Group Project Kickoff

3. 9/7, Background and Research Objectives

   - Tuesday: Brancati, Chapter 3 & 4
   - Thursday: McClendon (2016)

4. 9/14, Theory and Causality

   - Tuesday: Brancati, Chapter 5 & 6
   - Thursday: Beaman, Duflo, Pande and Topalova (2012)

5. 9/21, Methods Selection

   - Tuesday: Brancati, Chapter 7 & 8
   - Thursday: Burke and Kraut (2014)

6. 9/28, Case selection

   - Tuesday: Brancati, Chapter 9
   - Thursday: Howard, Savage, Saviaga, Toxtli and Monroy-Hernández (2016)

7. 10/5 Characteristics of Big Data

   - Tuesday: Salganik, Chapter 2 (2.1, 2.2, 2.3)
   - Thursday: Michel, Shen, Aiden, Veres, Gray, Pickett, Hoiberg, Clancy, Norvig, Orwant et al. (2011)

8. 10/12 Research Strategies with Big Data

   - Tuesday: Salganik, Chapter 2 cont'd (2.4, 2.5)
   - Thursday: Coviello, Sohn, Kramer, Marlow, Franceschetti, Christakis and Fowler (2014)

9. 10/19, Quantitative measurement

   - Tuesday: Brancati, Chapter 15
   - Thursday: Karaman and Pamuk (2013)

10. 10/26, Quantitative Data

    - Tuesday: Brancati, Chapter 16
    - Thursday: Lutz and Lutz (2015)

11. 11/2, Content Analysis

    - Tuesday: Brancati, Chapter 17
    - Thursday: Windsor, Dowell, Windsor and Kaltner (2018)

12. 11/9, Experiments

    - Tuesday: Brancati, Chapter 19

- Thursday: Rhodes, Leslie and Tworek (2012)

13. 11/16, Surveys

   - Tuesday: Brancati, Chapter 18
   - Thursday: Samuels and Zucco Jr (2014)

   11/23, **Thanksgiving Break, No Classes**

14. 11/30, Observational data analysis

   - Tuesday: Brancati, Chapter 20
   - Thursday: Cavaille and Marshall (2019)

15. 12/7, Project Week

   - Tuesday: Project presentations
   - Thursday: Project presentations

Instruction Mode: The instruction mode is remote synchronous. Classes will be held on Zoom at the time scheduled in LionPATH. The Zoom link for each class session is available in the Zoom tab in Canvas. During class meetings, to avoid background noise, please keep yourself muted except when you are talking. Having your cameras on is optional but highly encouraged.

Office Hours: Office hours will be held at the designated time via Zoom using the Office Hour Zoom link in Canvas. To meet the instructor for office hours, students should sign up for an office hour time slot at TBA

Attendance: Regular attendance is critical for building on the skills and knowledge developed throughout the class. Students who participate have a more complete understanding of the material presented and are more likely to succeed in the class. Attendance will not be taken during each class meeting, but if there appears to be an attendance problem attendance will be taken. In-excused absences beyond the first one may be penalized with a one-point reduction in the final grade.

Late Submission Policy: A penalty of 20% will accrue for each (rounded up) day that an assignment is late.

Extended Absences: During your enrollment at Penn State, unforeseen challenges may arise. If you ever need to miss an extended amount of class in such a circumstance, please notify your professor so you can determine the best course of action to make up missed work. If your situation rises to a level of difficulty you cannot manage on your own with faculty support, reach out to the Student Care & Advocacy office by phone at (814-863-2020) or email them at StudentCare@psu.edu. Office hours are Monday-Friday, 8 a.m. to 5 p.m.

Disability Accommodation Statement Penn State welcomes students with disabilities into the University's educational programs. Every Penn State campus has an office for students with disabilities. Student Disability Resources (SDR) website provides contact information for every Penn State campus (http://equity.psu.edu/sdr/disability-coordinator). For further information, please visit Student Disability Resources website (http://equity.psu.edu/sdr/).

In order to receive consideration for reasonable accommodations, you must contact the appropriate disability services office at the campus where you are officially enrolled, participate in an intake interview, and provide documentation: See documentation guidelines (http://equity.psu.edu/sdr/guidelines). If the documentation supports your request for reasonable accommodations, your campus disability services office will provide you with an accommodation letter. Please share this letter with your instructors and discuss the accommodations with them as early as possible. You must follow this process for every semester that you request accommodations.

Academic Integrity Statement Academic integrity is the pursuit of scholarly activity in an open, honest and responsible manner. Academic integrity is a basic guiding principle for all academic activity at The Pennsylvania State University, and all members of the University community are expected to act in accordance with this principle. Consistent with this expectation, the University's Code of Conduct states that all students should act with personal integrity, respect other students' dignity, rights and property, and help create and maintain an environment in which all can succeed through the fruits of their efforts.

Academic integrity includes a commitment by all members of the University community not to engage in or tolerate acts of falsification, misrepresentation or deception. Such acts of dishonesty violate the fundamental ethical principles of the University

community and compromise the worth of work completed by others.

Students with questions about academic integrity should visit http://www.la.psu.edu/ and then click on "Academic Integrity."

Counseling and Psychological Services Statement Many students at Penn State face personal challenges or have psychological needs that may interfere with their academic progress, social development, or emotional wellbeing. The university offers a variety of confidential services to help you through difficult times, including individual and group counseling, crisis intervention, consultations, online chats, and mental health screenings. These services are provided by staff who welcome all students and embrace a philosophy respectful of clients, cultural and religious backgrounds, and sensitive to differences in race, ability, gender identity and sexual orientation.

Counseling and Psychological Services at University Park (CAPS) (http://studentaffairs.psu.edu/counseling/): 814-863-0395

Counseling and Psychological Services at Commonwealth Campuses (http://senate.psu.edu/faculty/counseling-services-at-commonwealth-campuses/)

Penn State Crisis Line (24 hours/7 days/week): 877-229-6400
Crisis Text Line (24 hours/7 days/week): Text LIONS to 741741

Educational Equity/Report Bias Statements Penn State takes great pride to foster a diverse and inclusive environment for students, faculty, and staff. Consistent with University Policy AD29, students who believe they have experienced or observed a hate crime, an act of intolerance, discrimination, or harassment that occurs at Penn State are urged to report these incidents as outlined on the University's Report Bias webpage (http://equity.psu.edu/reportbias/)

Syllabus Change Policy: This syllabus is a guide and every attempt is made to provide an accurate overview of the course. However, circumstances and events may make it necessary for the instructor to modify the syllabus during the semester and may depend, in part, on the progress, needs, and experiences of the students. Changes to the syllabus will be made with advance notice.

References

Beaman, Lori, Esther Duflo, Rohini Pande and Petia Topalova. 2012. "Female leadership raises aspirations and educational attainment for girls: A policy experiment in India." *Science* p. 1212382.

Burke, Moira and Robert E Kraut. 2014. Growing closer on facebook: changes in tie strength through social network site use. In *Proceedings of the SIGCHI conference on human factors in computing systems*. pp. 4187–4196.

Cavaille, Charlotte and John Marshall. 2019. "Education and anti-immigration attitudes: Evidence from compulsory schooling reforms across Western Europe." *American Political Science Review* 113(1):254–263.

Coviello, Lorenzo, Yunkyu Sohn, Adam DI Kramer, Cameron Marlow, Massimo Franceschetti, Nicholas A Christakis and James H Fowler. 2014. "Detecting emotional contagion in massive social networks." *PloS One* 9(3):e90315.

Howard, Philip N., Saiph Savage, Claudia Flores Saviaga, Carlos Toxtli and Andrés Monroy-Hernández. 2016. "Social Media, Civic Engagement, and the Slacktivism Hypothesis: Lessons from Mexico's "El Bronco"." *Journal of International Affairs* 70(1):55–73.
URL: *https://www.jstor.org/stable/90012597*

Karaman, K Kivanc and Şevket Pamuk. 2013. "Different paths to the modern state in Europe: the interaction between warfare, economic structure, and political regime." *American Political Science Review* pp. 603–626.

Lutz, Brenda J and James M Lutz. 2015. "Globalisation and terrorism in the middle east." *Perspectives on terrorism* 9(5):27–46.

McClendon, Gwyneth H. 2016. "Race and responsiveness: An experiment with South African politicians." *Journal of Experimental Political Science* 3(1):60–74.

Michel, Jean-Baptiste, Yuan Kui Shen, Aviva Presser Aiden, Adrian Veres, Matthew K Gray, Joseph P Pickett, Dale Hoiberg, Dan Clancy, Peter Norvig, Jon Orwant et al. 2011. "Quantitative analysis of culture using millions of digitized books." *Science* 331(6014):176–182.

Rhodes, Marjorie, Sarah-Jane Leslie and Christina M Tworek. 2012. "Cultural transmission of social essentialism." *Proceedings of the National Academy of Sciences* 109(34):13526–13531.

Samuels, David and Cesar Zucco Jr. 2014. "The power of partisanship in Brazil: Evidence from survey experiments." *American Journal of Political Science* 58(1):212–225.

Windsor, Leah, Nia Dowell, Alistair Windsor and John Kaltner. 2018. "Leader language and political survival strategies." *International Interactions* 44(2):321–336.