

Research Design for Social Data Analytics
SoDA 308
Fall 2023
M/W/F 1:25-2:15PM, Borland Bldg 110

Instructor: Cassandra Tai

Office: B001 Sparks, Pennsylvania State University
Zoom: <https://psu.zoom.us/my/yhcasstai>
Email: yhcasstai@psu.edu
Office Hours: F 11:00AM–13:00PM & By appointment

Course Description: This course engages students in the study and use of research design tools for the analysis of social phenomenon, including the use of “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 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
- 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

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 Wednesday & Friday beginning on 9/8, the class will consist of three or four student presentations (each presentation up to 15 minutes). 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 8/23 & 8/25), and are listed below. A PDF version of the slides to be presented must be uploaded to Canvas by the student by 11:59AM on the presentation date. Students should sign up for their presentations on this **Google Sheet**, by Tuesday 11:59PM, 9/5.

- Background & Purpose & Hypotheses (see Huck)
- Methods & Results (see Huck)
- Applying Brancati (or Salganik): Commentary on the ways in which the reading from Monday applies to article.

Team Research Project: Students are required to complete a team research paper by the end of the semester. Each student will be assigned to one team. Each team consists of three or four students. The paper is expected to be original research—applying the research design principles covered in the course. The final paper is due on 12/15.

- **Process:** Teams should make progress on the project each week. The files associated with the paper should be shared among the students in the team and the instructor using an online folder. Each week during the team project time, the team should assign a task, to be completed over the next week, to each student in the team. 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 team members 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 9/24. *(5% of final grade)*
- **Design:** Teams 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 10/27 11:59 PM. *(5% of final grade)*
- **Presentation:** Teams will present their research design in Week 11 (11/1 & 11/3) and their final project in Week 15 (12/4 & 12/8). For both, each team will present for 15–20 minutes and have 10 minutes for Q/A. The expectation for the presentation at Week 11 is as follows: *Introduction (Topic and Puzzle), Literature Review, Theory and Hypotheses, Research Design (Model and Data)*. *(15% of final grade)*
- **Final Paper:** The paper is expected to be original research—applying the research design principles covered in the course. The final paper is due on 12/15. *(15% of final grade)*

Method proposal: Each student is responsible for writing a 2-page method proposal for the group research project. This proposal should explain a data analysis tool/method which

could be potentially used in the group project. The proposal should encompass the following elements:

- **Definition and Overview:** Provide a clear definition and comprehensive overview of the chosen data analysis tool/method.
- **Example Code:** Include sample code demonstrating the application of the tool/method.
- **Discussion:** Discuss the types of research projects and objectives for which the chosen tool/method could be effectively employed.

Students are required to submit their proposals by 10/15 and then discuss their proposals with their group members to decide which one will be chosen in their final projects.

Quizzes: There will be four 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%
- Team project: 40%
- Method proposal: 15%
- In-class quizzes: 15%

Critical Dates:

- 9/24: Topic document
- 10/15: Method proposal
- 10/27: Research design

- 10/31: Research design slides
- 12/5: Project slides
- 12/15: Project paper

Grading Scale: Grade values will not be rounded. That is, any grade value 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 readings listed for a particular day should be read before class time that day.

- 8/21, 8/23, 8/25, Reading a scientific study
 - Monday: Syllabus/course overview, online survey
 - Wednesday: Chapter 1, “Reading Statistics and Research,” Schuyler W. Huck.
 - Friday: Chapter 1, “Reading Statistics and Research,” Schuyler W. Huck
- 8/28, 8/30, 9/1, Social Research, Definition and Ethics
 - Monday: Brancati, Chapter 1 & 2
 - Wednesday: Brancati, Chapter 1 & 2 (Recorded)

- Friday: Online team project kickoff
3. 9/4, 9/6, 9/8, Background and Research Objectives
 - Monday: No Class (Labor Day)
 - Wednesday: Brancati, Chapter 3 & 4
 - Friday: Munger (2017)
 4. 9/11, 9/13, 9/15, Theory and Causality
 - Monday: Brancati, Chapter 5 & 6
 - Wednesday: Brancati, Chapter 5 & 6, Beaman, Duflo, Pande and Topalova (2012)
 - Friday: Quiz 1, Beaman et al. (2012)
 5. 9/18, 9/20, 9/22, Methods Selection
 - Monday: Brancati, Chapter 7 & 8
 - Wednesday: Brancati, Chapter 7 & 8, Burke and Kraut (2014)
 - Friday: Burke and Kraut (2014)
 6. 9/25, 9/27, 9/29, Characteristics of Big Data
 - Monday: Salganik, Chapter 2 (2.1, 2.2, 2.3)
 - Wednesday: Salganik, Chapter 2 (2.1, 2.2, 2.3), Michel, Shen, Aiden, Veres, Gray, Pickett, Hoiberg, Clancy, Norvig, Orwant et al. (2011)
 - Friday: Michel et al. (2011)
 7. 10/2, 10/4, 10/6, Research Strategies with Big Data
 - Monday: Salganik, Chapter 2 cont'd (2.4, 2.5)
 - Wednesday: Salganik, Chapter 2 cont'd (2.4, 2.5), Coviello, Sohn, Kramer, Marlow, Franceschetti, Christakis and Fowler (2014)
 - Friday: Quiz 2, Coviello et al. (2014)
 8. 10/9, 10/11, 10/13, Quantitative measurement
 - Monday: Brancati, Chapter 15
 - Wednesday: Brancati, Chapter 15, Karaman and Pamuk (2013)

- Friday: Karaman and Pamuk (2013)
9. 10/16, 10/18, 10/20, Quantitative Data
- Monday: Brancati, Chapter 16
 - Wednesday: Brancati, Chapter 16, Blumenstock, Cadamuro and On (2015)
 - Friday: Blumenstock, Cadamuro and On (2015)
10. 10/23, 10/25, 10/27, Content Analysis
- Monday: Brancati, Chapter 17
 - Wednesday: Brancati, Chapter 17, Siegel, Nikitin, Barberá, Sterling, Pullen, Bonneau, Nagler, Tucker et al. (2021)
 - Friday: Quiz 3, Siegel et al. (2021)
11. 10/30, 11/1, 11/3, Project Week I
- Monday: Preparation Day (No Class)
 - Wednesday: Project presentations
 - Friday: Project presentations
12. 11/6, 11/8, 11/10, Experiments
- Monday: Brancati, Chapter 19
 - Wednesday: Brancati, Chapter 19, Bail, Argyle, Brown, Bumpus, Chen, Hunzaker, Lee, Mann, Merhout and Volfovsky (2018)
 - Friday: Bail et al. (2018)
13. 11/13, 11/15, 11/17, Surveys
- Monday: Brancati, Chapter 18
 - Wednesday: Brancati, Chapter 18, Druckman and Levendusky (2019)
 - Friday: Druckman and Levendusky (2019)

11/20 - 11/24, **Thanksgiving Break, No Classes**

14. 11/27, 11/29, 12/1, Observational data analysis

- Monday: Brancati, Chapter 20
- Wednesday: Brancati, Chapter 20, Cavaille and Marshall (2019)
- Friday: Quiz 4, Cavaille and Marshall (2019)

15. 12/4, 12/6, 12/8, Project Week II

- Monday: Preparation Day (No Class)
- Wednesday: Project presentations
- Friday: Project presentations

Instruction Mode: The instruction mode is in-person. However, some classes might be offered remotely or asynchronously under certain circumstances. Any change to the mode of instruction will be announced in advance.

Attendance: 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. Regular attendance is critical for building on the skills and knowledge developed throughout the class. Students who participate more actively have a more complete understanding of the material presented and are more likely to succeed in the class.

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.

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

Office Hours: Office hours will be held at the designated time. To meet the instructor for office hours, students can walk in.

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 more information, 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: Penn State defines academic integrity as “the pursuit of scholarly activity in an open, honest and responsible manner” (Senate Policy 49-20). Dishonesty of any kind will not be tolerated in this course. Dishonesty includes, but is not limited to, cheating, plagiarizing, fabricating information or citations, facilitating acts of academic dishonesty by others, having unauthorized possession of examinations, submitting work of another person or work previously used without permission from the instructor, or tampering with the academic work of other students. Students facing allegations of academic misconduct should not drop the course; those who do will be added to the course again and will be expected to complete course work and meet course deadlines. If the allegations are dismissed, then the drop will be permitted. Students found responsible for academic misconduct often receive academic sanctions, which can be severe, and put themselves at risk for disciplinary sanctions assigned by the University’s Office of Student Conduct (see Senate Policy G-9). 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 Statement: 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 will be 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.

References

- Bail, Christopher A, Lisa P Argyle, Taylor W Brown, John P Bumpus, Haohan Chen, MB Fallin Hunzaker, Jaemin Lee, Marcus Mann, Friedolin Merhout and Alexander Volfovsky. 2018. “Exposure to opposing views on social media can increase political polarization.” *Proceedings of the National Academy of Sciences* 115(37):9216–9221.
- 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.
- Blumenstock, Joshua, Gabriel Cadamuro and Robert On. 2015. “Predicting poverty and wealth from mobile phone metadata.” *Science* 350(6264):1073–1076.
- 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.
- Druckman, James N and Matthew S Levendusky. 2019. “What do we measure when we measure affective polarization?” *Public Opinion Quarterly* 83(1):114–122.
- 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.
- 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.
- Munger, Kevin. 2017. “Tweetment effects on the tweeted: Experimentally reducing racist harassment.” *Political Behavior* 39(3):629–649.

Siegel, Alexandra A, Evgenii Nikitin, Pablo Barberá, Joanna Sterling, Bethany Pullen, Richard Bonneau, Jonathan Nagler, Joshua A Tucker et al. 2021. “Trumping Hate on Twitter? Online Hate Speech in the 2016 US Election Campaign and its Aftermath.” *Quarterly Journal of Political Science* 16(1):71–104.