Designing Social Inquiry: An Introduction to Causal Inference
Political Science 3105
Professor Vanessa Baird
Office Ketchum 114E; Email Vanessa.Baird@colorado.edu
https://vbaird.com/
Office Hours: Tuesday 2-5; Wednesdays 10-12
(also by appointment)
https://vbaird.com/

Course Objectives

This is not an ordinary statistics class. This course focuses less on the mathematics of statistics (though there is some) and programming (though there is some) and focuses more on the logic of causal inference, the common mistakes people (even scientists) make, and the tools to uncover those mistakes. There is no memorization (exams and quizzes are open note, open book, open internet), and exams test your ability to identify mistaken inferences.

Most statistics courses will tell you that correlation is not necessarily causation, but this course enumerates the logical and statistical tools to uncover the most common threats to the validity of a causal inference, using examples from a wide variety of fields: economics, psychology, medicine, and life sciences.

My story: I studied languages and literature in college and avoided math as much as I could. I started my PhD with the goal of teaching in political philosophy, but the program required some statistics. Six weeks into that first course, I caught the bug. I suddenly realized that these were the tools I needed to answer the questions I had always wondered about. I wrote my first paper on what makes people socialized in communist contexts to support democratic institutions. This was so much fun! After that, I took every statistics course the department offered. I want students who might be as reluctant as I was to learn "math" to get a sense of how much fun data analysis can be!

Who should take this course? I assume no prior knowledge. For anyone interested in a career in any science that uses statistics: biology, any social science, data science, policy, or the health sciences, this course is foundational. It is also intended for students who can keep up with the reading and assignments from week to week, as each week builds on knowledge from previous weeks. **Students who fall behind find it very difficult to catch up.** But the good news is that you may produce work that gets recognized for an **undergraduate award** from ICPSR or our own political science department. I have also had quite a number of students present their work at national conferences.

The paper: Students will produce original research explaining why people do the things they do (e.g. use drugs, vote, recycle) or why people believe what they do (e.g. abortion, gun control, helping the poor).
In class assignments: There will be short writing assignments due in class). There are four paper milestones; each is a stage in the final paper and include revisions of the previous milestone.

Exams and quizzes will be open note, open book, and open internet. You will be asked to identify illogical inferences. There are two exams and four quizzes (one can be dropped). Quizzes are online and the two exams are in class. Every exam and quiz is comprehensive and will test (at least in part) aspects learned in the class so far.

Grades

There are three tracks in this class, depending on how much you want to engage.

Points	Grade
910	A
880	A-
850	B+
810	B
780	B-
750	C+
700	C
600	D

Track I Estimated grade B-/C: 950 points possible

Track I: canned paper option: You can choose a relatively easy canned (by me) paper, where I guide you on the variables you choose. This is an option for those who just want to pass the class with a B- or C. A "canned" paper is worth 350 of the total points for the paper, so the total possible points out of 1000 are 850 (though there are options for extra credit).

Canned paper option (300 points)
Three milestones (for those doing a canned paper)   50, 50, 100
Final paper (for those doing a canned paper)   100

Track II Estimated Grade A-/B: 1100 points possible

Track II: You do the class without a study group. The total possible points are 1000, but there are some extra credit options available.

Paper 450 points
Three of four paper milestones (in order of due date)   50, 100, 150
Final paper   150

Track III Estimated grade A: 1250 points possible

Track III: You do the same individual paper project, but also work in small study groups of 4-5
There is **extra credit for working in groups (150)**. This 15% is graded partially as a function of the assessment of other members (and you) about your relative contribution over the semester. For those working in groups: part of the credit of paper milestones is composed of the quality of your peer reviews of the papers of your group members.

Note: I will put students in groups based on which statistical package you choose (See R versus Stata)

Paper (working in weekly study groups) 600 points
Three of four paper milestones (in order of due date)   30, 60, 100
Quality of peer reviews of three of four milestones   20, 40, 50

Final paper	150
Semester study group extra credit	150

All tracks (550 points)

"Prepare for class" assignments, pop quizzes, participation	150
Two midterms	200
Three quizzes (out of four total)	150
Final paper presentation during Final Exam session	50

Extra credit available to all students (100 points)

Turning in three milestones on time	10 points each (30)
Non-group peer reviews (must review five papers each milestone)	10 points each (30)
Dec 2 preliminary presentation	30 points
Quality of Dec 2 critique of others' papers	10 points

Note: Grades are not calculated using Canvas. There are some grades there, but it is not a way to see your grade so far. I have created three Excel spreadsheets with the target due dates and grades so that you can keep track of your grades there.

Assignments and TARGET due dates

Assignment	Due date	Where to turn in
Online quiz 1	September 3 at 5pm	Canvas
Paper milestone 1	Friday, September 17 at 5pm	Kritik or Canvas
Peer review 1	Friday, September 24 at 5pm	Kritik
Paper milestone 2	Friday, October 1 at 5pm	Kritik or Canvas
Peer review 2	Friday, October 8 at 5pm	Kritik
Online quiz 2	Opens Tuesday 5pm DUE Wednesday October 13 at 5pm	Canvas
Midterm 1	Opens NO CLASS: Tuesday, October 19 DUE Wed Oct 20th at 5pm	~~In person~~ Canvas
Online quiz 3	Friday October 29 at 5pm	Canvas
Paper milestone 3	Friday, November 12pm	Kritik or Canvas
Peer review 3	Friday November 17 at 5pm	Kritik
Online quiz 4	Monday November 15 at 5pm	Canvas
Midterm 2	Thursday, November 18	In person
Rough draft (optional milestone 4)	Friday, December 3	Canvas
Final paper due	Monday December 13	Canvas
Final presentations	Final exam period	In person

Note: If you are in Track 3, OR you want extra credit for peer reviews, you must use Kritik. for Paper milestones. Otherwise, Canvas.

Statistical package: The pros and cons of Stata versus R

You are welcome to use any statistical package that you would like. It is my experience that you can easily use Google for answers to most questions about any statistical code, so it is not the focus of this class. In my experience, it is relatively easy to learn any package once you understand the underlying logic of making inferences from data.

Pros and cons of R: You may have experience with R. You can also download R or RStudio free. R is also becoming the most used statistical package in data science familiarity with it is a very marketable skill. The downside is the steep learning curve.

Pros and cons of Stata: In my experience, Stata is easier to learn and use, but it also **costs $50 for a student six month lease**. You can also use Stata on the various computer labs across campus. It has been my experience in this class that most students, even when they have experience with R, choose Stata (though many stick with R).

I will provide code for what you need in both languages, but when I use a program in class, it will be in Stata. The department has a TA who knows the answers to most questions about both R and Stata. Google know the answers to most questions in both R and Stata. I will also provide written and video guides to both languages. You can also purchase the R or Stata companion to Political Analysis.

Costs

There is no text; all readings are posted on Canvas.
The peer review app (Kritik) costs $24, but Kritik gives scholarships to students who say this will be a hardship. Stata costs $50 if you want to use Stata and want the convenience of a personal copy.

Book chapters are used from the following sources
Applied Multiple Regression-Correlation Analysis for the Behavioral Sciences, 3rd Edition by Jacob Cohen, Patricia Cohen, Stephen G. West, Leona S. Aiken (CCWA)
Statistics Concepts And Controversies by David S. Moore, William I. Notz (MN)
Understanding Political Science Research Methods The Challenge of Inference by Maryann Barakso, Daniel M. Sabet, Brian Schaffner (BSS)
Open Source Textbook (OST)

Course Outline

Guide to color codes:
Required reading and "prepare for class" assignments
What we will do in class
Graded assignments, exams, and quizzes

Part I. The logic of inference

Tuesday, August 24: Introduction

In class
Outline of the final paper: each section should have dates in class that help write that portion

Thursday, August 26: Logic of inference

Prepare for class
Play PowerPoint slideshow with audio: Evaluating the kinds of statements
Read OST Chapters 1 -3
BSS: Logic of inference

In class
1. History of philosophy
2. Philosophy of science
3. 3. Hans Rosling Ted talk, 0-4:50
4. Logic of normative, interpretation, causal, and factual

Tuesday, August 31: Causality: the importance of concepts and units

Prepare for class
Read OST Chapters 5-8
BSS: Descriptive conceptualization

In class
PowerPoint lecture: Introduction to scientific inquiry
Learn the nature of a research question
Learn what works as a concept (and which concepts can be about people)
Make a causal argument; draw the resulting scatterplot

Example research questions	
Why do people believe in science?	Why do people support civil liberties for people they dislike?
Why do people support war?	Why do people support harsh punishments?
Why do people vote (or protest)?	Why do people recycle?
Why do people vote the way they do?	Why do people trust government?

Thursday, September 2: The ethical, causal, and descriptive nature of why does Y matter?

Prepare for class
Read BSS Research Question
Purview questionnaires or codebooks for potential data sources) that can be used for this class
**   (you can use other surveys but it may add a weekend of headache)**
https://electionstudies.org/
https://gss.norc.org/
https://cces.gov.harvard.edu/ (I wrote questions for the CCES 2016 and CCES 2020)
Pro-tip: Pick a dependent variable TODAY. The least creative aspect of your question is your dependent variable; just choose something you care about. The creative element is your independent variable.
Pro-tip: Choose one of the data sets TODAY. The GSS has the best independent variables, so you have more creativity with that one.

Recommended: "Gerring: Mere description"

In class
1. The ethics and causal aspect of: why does Y matter?
2. Practice quiz over logic of inference
3. Go over practice quiz

<p style="text-align:center">DUE September 3 at 5pm: Online quiz 1</p>

Topic: Logic of inference (categorizing statements), causality, concepts and units

Tuesday, September 7 Three kinds of literature reviews

Prepare for class
Read: Rule of Law Literature Review example
"Enigmas of Tolerance"
Data Gordon Intro Literature Excerpts tables: how to read a table
Bring your research question to class, in the form of why does Y vary? (Units must be people)

 In class
Go over quiz
PowerPoint lecture: The three kinds of literature reviews

Thursday, September 9: How to write a literature review efficiently

Prepare for class
Bring a (short) summary of (at least) four articles that can be used in each of the four kinds of literature review
Epidemiology EMF: using puzzles from observed correlations to infer which hypotheses to test
How to Read a Journal Article

Recommended: Watch video: Project advice: How to write a literature review efficiently

In class
How to not read everything: focus on the abstract, data presentation, conclusion
How to take notes: by concept
How to not plagiarize (even by accident)
Pro-tip: do not miss class; it will save you twenty (or more) hours of work

<p style="text-align:center">Part II. Probabilistic linear causal relationships</p>

Tuesday, September 14 Causality: an introduction to exogeneity

Prepare for class
Read: BSS Theory
Theory causality: application in education research
Shipley Causal Inference Logic: application in biostats (botany research)
Example of literature review - political participation

Recommended: Elaborate Theory
Recommended: logic quality-meets-quantity-case-studies-conditional-probability-and-counterfactuals

In-class
Understanding exogeneity: lecture on Death Camp Eldorado
Causality in cross-sectional observational study

Thursday, September 16: Introducing the GSS, ANES, and the CCES: how to understand the utility of a data source

Prepare for class

if you are using Stata

Easygoing Reference Stata Intro

Easygoing Reference Stata descriptive statistics

If you are using R

Intro to R
Intro R
R descriptive stats

In-class
Understanding life development as a way to exogeneity

DUE Friday, September 17 at 5pm: Paper milestone 1

Write your research question in the correct format.
List about ten independent variables from your data, copy and paste the survey questions for each, name the concepts and make all causal arguments with a single sentence.
With regard to your most nonobvious cause, write a causal argument using the relevant literature.

Tuesday, September 21 Statistical descriptive inferences: some statistics basics

Prepare for class
Reading: BSS: Descriptive inference
Reading: CCWA Visualization Histogram
MN statistics descriptions (long but big print and lots of pictures)

Recommended: Read KKV: Descriptive inference

In class
PowerPoint lecture: Statistical descriptive inference
1. Mean, mode, median, probability distributions
2. Variance, standard deviation

Thursday, September 23 From variation to co-variation and correlation

Prepare for class
Run and interpret the following analyses: a histogram, frequency, and standard deviation of your DV
CCWA Visualization Scatter
MN correlation scatterplots (long but big print and lots of pictures)

In class
PowerPoint lecture: Correlations: Covariance Correlation

DUE Friday, September 24 at 5pm: Peer reviews of paper milestone 1

Tuesday, September 28: Bivariate linear regression: Ordinary Least Squares

Prepare for class
Read BSS Observation
Read CCWA Linear regression
MN OLS bivariate (long but big print and lots of pictures)

Recommended: Gordon Intro OLS

For those using Stata

Easygoing Reference Stata correlation regression

For those using R

R Correlation Regression

Recommended, for fun:
What is life without an 80s song about the Central Limit Theorem?

Part III. Types of statistical relationships

Thursday, September 30: Confounding versus additive relationships

Prepare for class
Bring a crosstab table, a correlation, and bivariate regression with interpretation to class your most important IV and DV
Read chapter 4 from Open Source Text
CCWA Intro MRC Confounding Mediating
Gordon Indirect and Omitted Variable Bias

In class
The logic of a statistical control: DAGs, Venn diagrams, scatterplots, crosstabs
PowerPoint lecture: 2021 spurious intervening suppression

DUE Friday, October 1 at 5pm: Paper milestone 2

Using the literature, one page on why the question is important
Describe the DV: variations, measures
Using the literature, write a paragraph on the causal effects of each IV on the DV
Include measures of all independent variables, in their respective section
Subheadings should be the name of each concept

Tuesday, October 5: Mediation and suppression (Simpson's paradox)

Prepare for class
CCWA Intro to Confounding and Mediating

In class
PowerPoint lecture: 2021 spurious intervening suppression, continued

Thursday, October 7: Moderation: an introduction to interactions

Prepare for class
CCWA Ch7 Interactions

In class
PowerPoint lecture: 2021 conditional

Topic: descriptive inference, correlation, linear regression, causality, crosstabs, Venn diagrams, DAGs, and interpreting statistical controls: confounding, mediation, suppression, and moderation

Tuesday, October 12: Review for midterm 1: mediating or confounding, moderating?

In class Go over quiz

Thursday, October 14: Midterm 1

Topic: descriptive inference, correlation, linear regression, causality, crosstabs, Venn diagrams, DAGs, and interpreting statistical controls: confounding, mediation, suppression, and moderation

Part IV: Measurement

Tuesday, October 19: Scaling I: Dimensionality of Likert scales

Prepare for class
Chapter 9, from Open Source Text
MN measuring

In class
PowerPoint Lecture: Conceptualization validity

Thursday, October 21: Complex scales and validity

Prepare for class
Chapter 10, from Open Source Text
MC Unidimensional scaling

In class
PowerPoint Lecture: Scaling
Application: Support for abortion

Tuesday, October 26: Reliability

Prepare for class
CZ Reliability validity

In class PowerPoint Lecture: Reliability
Application: Measuring life circumstances

Part V: Multiple regression

Thursday, October 28: Multiple regression

Prepare for class
CCWA Intro MRC Ballantine
BSS Linear multiple regression

Recommended DSI_CCWA_MRC_Partial_Semipartial

In class

PowerPoint lecture: Trivariate

DUE: Friday October 29 at 5pm: Online quiz 3

Topic: measurement: scaling, reliability, and validity
All other material so far

Tuesday, November 2 Using multiple regression to diagnose the type of relationship among the following: alternative causes, confounding, mediating, suppression

Prepare for class
Bring to class Run and interpret the following analyses: reliability analyses for 3 multiple item scales, defend the validity and reliability
CCWA Ch11 MRC Causal

In class
PowerPoint lecture: Etiology

DUE Friday, November 5pm: Paper milestone 3

Why the question is important
The literature review for all independent variables
Measure every variable
Evaluate validity and reliability
Preliminary caveats

Thursday, November 4: Dummy variables, interactions with dummies, curvilinearities

Prepare for class
CCWA Ch8 Nominal Dummy Variables
CCWA Ch9 Dummy Interaction
Dummy interaction models notes
Data CCWA Ch6 Polynomials Curvilinear

In class
Application: Stealth democracy and concern for success

Tuesday, November 9: Continuous interactions

Prepare for class
Nelson Gibson histogram interactions
Gibson political freedom

Part VI: Caveats and implications

Thursday, November 11: Effect of systematic measurement error, systematic missing data: DV, IV, control

Prepare for class
CCWA Ch11 Missing Data

In class
PowerPoint lecture: Measurement and inference

Topic: Interpreting multiple regression

Tuesday, November 16: Review for midterm 2

Bring to class Run and interpret the following analyses: dummy variable regression, interaction effects with a dummy variable, continuous interactions
Go over quiz

Thursday, November 18: Midterm 2

Topic: Interpreting multiple regression
All other material so far

Thanksgiving break: no class November 23, and 25

Tuesday, November 30: Go over exam and discuss the logic of caveats
Thursday, December 2: Understanding implications

DUE Rough draft due December 3, if you want feedback by Dec. 8 (optional)

Tuesday, December 7: Present initial findings from paper (optional)
Thursday, December 9: Paper questions

Final paper due Monday December 13

Final exam period: Final presentations of final paper

Appendix: Some helpful videos

General helpful videos

[Simple linear regression](#)
[Trivariate regression](#)
[Confounding variables](#)
[More about confounding](#)
[Mediation](#)
[Effect modification](#)
[Understanding collinearity (multicollinearity)](#)
[Summary of regression, mediation, confounding and modifier (interaction)](#)
[Building a regression model in R, and estimating the effect size](#)

Learning or reviewing R

[Why use R?](#)
[How to install R and Rstudio](#)
[Importing data in R and RStudio](#)
[Working with variables in R and RStudio](#)
[Setting up a working directory in R](#)
[How to install packages in R and RStudio](#)
[Histograms in R](#)
[Mean, standard deviations in R](#)
[Scatterplots in R](#)
[Modifying plots in R](#)
[Add text to a plot in R](#)
[Add legend to plot in R](#)
[Simple linear regression in R](#)
[Trivariate regression in R](#)
[Confounding variables in R](#)
[More confounding variables in R](#)
[Collinearity in R](#)
[Interpreting multiple regression with modifying variable (aka interaction) in R](#)
[Interaction with two categorical (dummy) variables in R](#)

Learning Stata

[There are a ton of very good introductory videos here](#)

Rules and Regulations

Classroom Behavior

Both students and faculty are responsible for maintaining an appropriate learning environment in all instructional settings, whether in person, remote or online. Those who fail to adhere to such behavioral standards may be subject to discipline. Professional courtesy and sensitivity are especially important with respect to individuals and topics dealing with race, color, national origin, sex, pregnancy, age, disability, creed, religion, sexual orientation, gender identity, gender expression, veteran status, political affiliation or political philosophy. For more information, see the policies on classroom behaviorLinks to an external site. and the Student Conduct & Conflict Resolution policiesLinks to an external site..

Requirements for COVID-19

As a matter of public health and safety due to the pandemic, all members of the CU Boulder community and all visitors to campus must follow university, department and building requirements and all public health orders in place to reduce the risk of spreading infectious disease. Students who fail to adhere to these requirements will be asked to leave class, and students who do not leave class when asked or who refuse to comply with these requirements will be referred to Student Conduct and Conflict ResolutionLinks to an external site.. For more information, see the policy on classroom behaviorLinks to an external site. and the Student Code of ConductLinks to an external site.. If you require accommodation because a disability prevents you from fulfilling these safety measures, please follow the steps in the "Accommodation for Disabilities" statement on this syllabus.

As of Aug. 13, 2021, CU Boulder has returned to requiring masks in classrooms and laboratories regardless of vaccination status. This requirement is a temporary precaution during the delta surge to supplement CU Boulder's COVID-19 vaccine requirement. Exemptions include individuals who cannot medically tolerate a face covering, as well as those who are hearing-impaired or otherwise disabled or who are communicating with someone who is hearing-impaired or otherwise disabled and where the ability to see the mouth is essential to communication. If you qualify for a mask-related accommodation, please follow the steps in the "Accommodation for Disabilities" statement on this syllabus. In addition, vaccinated instructional faculty who are engaged in an indoor instructional activity and are separated by at least 6 feet from the nearest person are exempt from wearing masks if they so choose.

Students who have tested positive for COVID-19, have symptoms of COVID-19, or have had close contact with someone who has tested positive for or had symptoms of COVID-19 must stay home. In this class, if you are sick or quarantined, let me know so by email. I will make old Zoom lectures available that you can use as an alternative to coming to class, if you are sick. **I do not require "doctor's notes" or quarantine verifications for classes missed due to illness or quarantine; campus health services do not provide "doctor's notes," appointment verifications or quarantine verifications.}**

Accommodation for Disabilities

If you qualify for accommodations because of a disability, please submit your accommodation letter from Disability Services to your faculty member in a timely manner so that your needs can be addressed. Disability Services determines accommodations based on documented disabilities in the academic environment. Information on requesting accommodations is located on the Disability Services websiteLinks to an external site.. Contact Disability Services at 303-492-8671 or dsinfo@colorado.edu for further assistance. If you have a temporary medical condition, see Temporary Medical ConditionsLinks to an external site. on the Disability Services website.

Preferred Student Names and Pronouns

CU Boulder recognizes that students' legal information doesn't always align with how they identify. Students may update their preferred names and pronouns via the student portal; those preferred names and pronouns are listed on instructors' class rosters. In the absence of such updates, the name that appears on the class roster is the student's legal name.

Honor Code

All students enrolled in a University of Colorado Boulder course are responsible for knowing and adhering to the Honor Code academic integrity policy. Violations of the Honor Code may include, but are not limited to: plagiarism, cheating, fabrication, lying, bribery, threat, unauthorized access to academic materials, clicker fraud, submitting the same or similar work in more than one course without permission from all course instructors involved, and aiding academic dishonesty. All incidents of academic misconduct will be reported to the Honor Code (honor@colorado.edu); 303-492-5550). Students found responsible for violating the academic integrity policy will be subject to nonacademic sanctions from the Honor Code as well as academic sanctions from the faculty member. Additional information regarding the Honor Code academic integrity policy can be found on the Honor Code websiteLinks to an external site..

Sexual Misconduct, Discrimination, Harassment and/or Related Retaliation

The University of Colorado Boulder (CU Boulder) is committed to fostering an inclusive and welcoming learning, working, and living environment. CU Boulder will not tolerate acts of sexual misconduct (harassment, exploitation, and assault), intimate partner violence (dating or domestic violence), stalking, or protected-class discrimination or harassment by or against members of our community. Individuals who believe they have been subject to misconduct or retaliatory actions for reporting a concern should contact the Office of Institutional Equity and Compliance (OIEC) at 303-492-2127 or email cureport@colorado.edu. Information about OIEC, university policies, reporting optionsLinks to an external site., and the campus resources can be found on the OIEC websiteLinks to an external site..

Please know that faculty and graduate instructors have a responsibility to inform OIEC when made aware of incidents of sexual misconduct, dating and domestic violence, stalking, discrimination, harassment and/or related retaliation, to ensure that individuals impacted receive information about their rights, support resources, and reporting options.

Religious Holidays

Campus policy regarding religious observances requires that faculty make every effort to deal reasonably and fairly with all students who, because of religious obligations, have conflicts with scheduled exams, assignments or required attendance. In this class, **{Faculty: *insert your procedures here*}.**

See the campus policy regarding religious observancesLinks to an external site. for full details.