

Data and Policy Summer Scholar (DPSS) Program

2026 Syllabus

Professors: Austin Wright, PhD (Academic Director) & Sheng-Hao Lo, PhD (Data Analytics Instructor)

Graduate Instructor: Peizan Sheng (Programming in R)

Teaching Assistants: Rushikesh Jadhav, MPP'26 & Alex Zhao, MPP'25

Course Description:

This course is an introduction to research design and data analytics. We will focus on establishing a foundation for studying public policy problems through data. This course consists of two complementary sections – Data Analytics and Programming in R. These two sections are equivalent to one full course at the University of Chicago Harris School of Public Policy.

Data Analytics [DA] introduces the statistical foundations, tools, and methods employed by public policy researchers. In this section we will explore the fundamental problem of causal inference, and learn how to use data, research design, and statistical modeling to navigate around this problem. The suite of tools covered in Data Analytics will vary by learning track.

Programming in R [R] introduces students to the powerful statistical modeling software known as R. This is a free and open-source software that is constantly being expanded and upgraded. Knowledge of programming in this language is in high demand in policy job markets. In this section of the course, you will be introduced to many of the key uses of the software — including data wrangling, applications of causal inference methods, data visualizations, and more. The suite of tools covered in R Programming will vary by learning track.

The Capstone is an opportunity to apply the coursework of Data Analytics and R Programming in a research project with open access data. Datasets have been selected in three areas and students may choose one based on interest. One or two of these datasets shall serve as the basis for a couple of assignments in the Programming in R assignments. You will work with a small group of peers ("capstone group") to collaborate and share ideas, but you will each submit an individual project. Each student must submit a project that reflects independent work. The final product is something you will be able to showcase in your resume, LinkedIn, job, or graduate school application. Professional track students will provide a Policy Memo while academic track students will deliver a Research Note. Capstone topics may include:

1. American Electoral Policy: Evaluating the determinants of voting patterns in the U.S.
2. Labor Economics and Public Policy: Analyzing data from nationally representative microdata from the U.S. through the IPUMS-CPS dataset
3. Economic Development Policy: Analyzing data from Ghana

Virtual Format:

Academic lectures for Data Analytics and R Programming are delivered via weekly video modules (pre-recorded lectures). These modules will provide students with the knowledge and materials to understand the basics of data analytics as applied to public policy. During the Capstone Weeks, faculty will host synchronous lectures for each research project. You are strongly encouraged to attend the synchronous lecture. Live office hours are held multiple times throughout each week, enabling students to connect with faculty and graduate teaching assistants. The schedule and how to access the virtual office hours will be available on the Canvas calendar.

Course Materials:

All necessary course materials will be made accessible via Canvas. There are no required textbooks. If you would like a supplementary text that covers many of the same topics, then we recommend by category the following:

Data Analytics:

- *Thinking Clearly With Data: A Guide to Quantitative Reasoning and Analysis* by Bueno De Mesquita and Fowler.
- *Mastering Metrics: The Path from Cause to Effect* by Angrist and Pischke is recommended for Quantitative Analysis.
- *Mostly Harmless Econometrics* by Angrist and Pischke.

R Programming:

- *R for Data Science (R4DS)* by Hadley Wickham and Garret Grolemond
 - This online [textbook](#) is free, and a new version has been posted as well.
- In HW1, you will **download R and RStudio**. Please use the following resources to do so:
 - Download R: [here](#)
 - Download R Studio: [here](#)
 - Guide for Downloading R: [here](#)

Assignments and Grading:

The graded assignments for this course will consist of homework assignments. Each homework assignment will be released via Canvas, and you will submit your homework in Canvas by the required deadline. The assignments will be due on Mondays by 11:59 pm Central Daylight Time and include separate evaluations for both R Programming and Data Analytics.

The concept quizzes will be available after each video lecture, and the questions will relate directly to content from the preceding lecture. These concept quizzes will be an excellent way to gauge your knowledge of the material covered in lecture, and they will prepare you for questions on your homework assignments. Concept quizzes are optional and not graded.

Communication:

Communication from instructors and TAs to students will happen through posting of materials on Canvas, including postings to Announcements and on Ed Discussion. The most effective way for the teaching team to support students is for students to submit questions in the following forms:

- Questions regarding course material should be asked in the live Office Hours or posted on Ed Discussion, a forum that is monitored by the teaching assistants and instructors.
- Questions regarding scheduling should be directed to Peizan Sheng.

Ed Discussion Board: This course will also have a dedicated Ed Discussion board. Ed Discussion is a free online gathering place where students can engage with each other regarding course material, under the guidance of their instructors. Setting up an Ed Discussion account will be part of your first assignment, but participation is optional. On Ed Discussion, you can both ask and answer questions related to homework assignments and course materials. If you provide a strong answer, one of the teaching assistants might endorse your answer.

University of Chicago Academic Integrity Policies:

All students are expected to uphold the highest standards of academic integrity and honesty. Among other things, this means that students shall not represent another's work as their own.

The University's policies regarding academic integrity and dishonesty are described [here](#) and below:

"It is contrary to justice, academic integrity, and to the spirit of intellectual inquiry to submit another's statements or ideas as one's own work. To do so is plagiarism or cheating, offenses punishable under The University's disciplinary system. Because these offenses undercut the distinctive moral and intellectual character of the University, we take them very seriously. Proper acknowledgment of another's ideas, whether by direct quotation or paraphrase, is expected. In particular, if any written or electronic source is consulted and material is used from that source, directly or indirectly, the source should be identified by author, title, and page number, or by website and date accessed. Any doubts about what constitutes "use" should be addressed to the teaching fellows."

Much of this class will be spent coding in R. We encourage using online resources and discussion boards such as Stack Overflow. Code you find online should be referenced via a link to where you found the solution to your problem. Cite all code you copy, even if just a single line snippet. You may screen share your code with your peers, but do not copy each other's code. All work you turn in must be your own and copying code will be considered a violation of the terms of academic honesty.

The use of Artificial Intelligence (AI) tools (e.g., ChatGPT, Gemini, Claude, or similar products) to generate code or written responses for assignments is restricted. Assignments may be reviewed using plagiarism and AI-detection tools; however, these tools are imperfect and may be used as one factor among several when evaluating potential academic integrity concerns. Students may use AI tools as a learning aid only (e.g., to clarify concepts or troubleshoot errors) and must clearly cite the tool used and the prompt provided. Failure to disclose AI assistance, or submitting AI-generated work as one's own, may result in an academic integrity review and assignment penalties.

University Policies and Resources: Participants are expected to familiarize themselves with and adhere to all [University of Chicago policies and standards of conduct](#). These policies promote a safe, inclusive, respectful, and academically rigorous learning environment for all members of the University community.

Late Policy:

Requests for an extension to submit a homework assignment must be submitted through [this form](#), once a student submits, they will be notified in 24 hours if your request has been accepted. In the instance of an extenuating circumstance (i.e., health issue, family emergency) and at their discretion, DPSS faculty and TAs may grant a homework extension (maximum of 2 days) with no penalty. Barring any extenuating circumstances, any late submissions will result in a 5% penalty each day past the original deadline (maximum total penalty of 50% or 10 days). Exceptions apply for the capstone project.

Schedule:

Week	ACADEMIC TRACK		PROFESSIONAL TRACK		Assignments
	Data Analytics	R Programming	Data Analytics	R Programming	
Week 1	1.1 - Foundations of Causal Inference for Public Policy	2.1 - Intro to R and RStudio	1.1 - Foundations of Causal Inference for Public Policy	2.1 - Intro to R and RStudio	HW1 - 10 pts
	1.2 - Fundamentals of Working with Data	2.2 - Intro to tidyverse	1.2 - Fundamentals of Working with Data		
Week 2	1.3 - Randomized Control Trials	2.3 - Data Transformation and Cleaning	1.2 - Fundamentals of Working with Data Cont... 1.3 - Randomized Control Trials (Optional)	2.2 - Intro to tidyverse	HW2 - 30 pts
	1.4 - Bivariate regression	2.4 - Data visualizations	1.4 - Bivariate regression	2.3 - Data Transformation and Cleaning	
Week 3	1.5 - Multivariate regression	2.5 - APIs	1.5 - Multivariate regression	2.3 - Data Transformation and Cleaning Cont...	HW3 - 30 pts
		2.6 - Programming concepts (for loops, functions, control flow)		2.4 - Data visualizations 2.5 - APIs	

Week 4	1.6 - Functional Forms, Interactions, and Indicator Variables	2.7 - Casual Inference Stats in R	1.6 - Functional Forms, Interactions, and Indicator Variables 1.7 - Binary Dependent Variables 1.8 - Differences in Differences (Optional)	2.6 - Programming concepts (for loops, functions, control flow)	HW4 - 30 pts
	1.7 - Binary Dependent Variables	2.8 - Spatial Data			
	1.8 - Differences in Differences				
Week 5	1.9 - Panel data and fixed effects models	2.9 - R Markdown	1.9 - Panel data and fixed effects models (Optional) 1.10 - Regression discontinuity designs (Optional) 1.11- Instrumental variables (Optional)	2.7 - Casual Inference Stats in R (Optional) 2.8 - Spatial Data 2.9 - R Markdown	
	1.10 - Regression discontinuity designs				
	1.11- Instrumental variables				
Week 6	Capstone Project				
Week 7					

Passing the Course & Certification of Credential Completion:

In order to pass this course, you must successfully pass at least two of the three academic components: Data Analytics, R Programming, and the Capstone Research Project. The passing score in each of the three components is 60%.

Upon successful completion of the coursework, you will be issued a Certification of Credential Completion by the University of Chicago Harris School of Public Policy. The Certification of Credential Completion contains a point grade for each academic component: Data Analytics, R Programming, and Capstone Research Project. You will receive an email from our office (harriscredential@uchicago.edu) with the electronic certification approximately four weeks after the program.

Community Engagement Award: The Community Engagement Award recognizes participants who make the most of their DPSS experience through meaningful engagement with co-curricular programming, peer learning, and professional development opportunities. Recipients demonstrate a strong commitment to professional growth, community building, and active participation throughout the program.

1. Attend at least 7 synchronous speaker series events or professional development workshops
2. Complete at least 10 community engagement contributions through any combination of:
 - Participating in WhatsApp or WeChat community check-ins
 - Posting thoughtful questions, insights, or constructive responses on Ed Discussion

Participants who satisfy the Community Engagement Award requirements will receive formal recognition for their engagement, leadership, and contributions to the DPSS learning community.

For reference, sample copies of both the Community Engagement Award and the Certification of Credential Completion are included with your program materials. The Certification of Credential Completion contains a point grade for each academic component: Data Analytics, R Programming, and the Capstone Research Project.



The University of Chicago Harris School of Public Policy

certifies

Full Name

has successfully completed the non-degree, non-credit

Data and Policy Summer Scholar Program

an immersive seven-week experience with training in Data Analytics in Public Policy, Introduction to R Programming,
a capstone research project with faculty, and rich connections with the UChicago community

June 30 – August 15, 2025

Course Dates

Austin Wright

Associate Professor and Faculty Lead



This is the official report for the academic performance of

Full Name

with the completion of the following non-credit coursework for the

2025 Data and Policy Summer Scholar Program

Data Analytics in Public Policy

Hours of Instruction: 20

Grade: «D_point»

Introduction to Programming in R

Hours of Instruction: 20

Grade: «R_point»

Capstone Research Project

Hours of Instruction: 10

Grade: «Capstone_point»



The University of Chicago Harris School of Public Policy

awards

Full Name

the

DPSS Community Engagement Award

for exemplary engagement in synchronous DPSS events and activities



June 30 – August 15, 2025

Course Dates

Shilin Liu, MPP'17

Associate Dean of Executive Education and Partnerships