

## Data and Policy Summer Scholar (DPSS) Program Sample Syllabus

### Instructor

Austin Wright, Assistant Professor and DPSS Faculty Director  
[Professor Wright's Biography](#)

### 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.

**Data Analytics [DA]** provides an introduction to 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.

**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.

These two sections compare to one full course at UChicago Harris.

### Capstone Research Project

In the capstone research project, you will collaborate with Professor Wright and a group of peers on a real-world problem and design a policy recommendation. The project hours are approximately 15-20.

You will harness the skills of research design, policy analysis, and team collaboration to conduct a research project using open-source or faculty-provided datasets. There are elements of data collection, analysis, and visualization, and result in a policy memo.

Project topics are determined during the program. Professor Wright presents three-four project topics, and students vote to select their preferred topic. You will collaborate with peers in small groups (your “capstone peer group”) on the same topic. The skills gain in the project are transferable for further research in your area of interest.

### 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.

The Capstone Project is combination of live lecture and interactive working sessions with Professor Wright, graduate teaching assistants, and capstone peer group.

Office hours are held multiple times throughout each week, enabling students the connect with faculty and graduate teaching assistants. Office hours are scheduled based on the current cohort of students as we accommodate the global time zones and work schedules.

Students are encouraged to connect regularly with their optional study groups and assigned capstone peer group.

### **Course Materials**

All necessary course materials will be made accessible via the course canvas page. There are no required textbooks. If you would like a supplementary text that covers many of the same topics, then R for Data Science (R4DS) by Hadley Wickham and Garret Golemund's is recommended for Programming in R. This online textbook is free. Mastering `Metrics: The Path from Cause to Effect by Angrist and Pischke is recommended for Data Analytics.

A recommended, not required, read is [Thinking Clearly with Data: A Guide to Quantitative Reasoning and Analysis](#) by Ethan Bueno de Mesquita and Anthony Fowler

### **Assignments and Grading**

The assignments will consist of homework assignments, and optional concept quizzes. Each homework assignment will be released via canvas. Details of points per assignment and assignment due dates are available in the Canvas Syllabus.

The concept quizzes will be available after each video lecture, and the questions will relate to the content from the preceding lecture. These concept quizzes are an excellent way to gauge your knowledge of the material covered in lecture, and prepare you for questions on homework assignments. Concept quizzes are not graded, and you can take them as many times as you would like.

The Capstone Research Project is graded separately. Please refer to Capstone Syllabus.

### **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 subsequent participation is recommended not required. This discussion board will be accessible via canvas.

On Ed Discussion, you can ask and answer questions related to homework assignments and course materials. If you provide a strong answer, the teaching team might endorse your answer.

## Schedule

	Data Analytics	R Programming
Week 1	1.0 - Course Preview 1.1 - Foundations of Causal Inference for Public Policy	2.0 - Course Preview 2.1 - Intro to R and RStudio (working dirs, projects, panes, R basics, etc)
	1.2 - Fundamentals of Working with Data	2.2 - Intro to tidyverse, fundamentals of data, basic visualization
Week 2	1.3 - Difference in means: RCTs (experimental ideal)	2.3 - Tidy data, data wrangling, and simple data cleaning
	1.4 - Bivariate regression: properties, testing, interpretation	2.4 - Recoding, data transformation, and joins (plus more wrangling)
Week 3	1.5 - Multivariate regression: testing, interpretation, omitted variable bias	2.5 - Data visualization and exploration (ggplot2, summarization)
	1.6 - Binary outcomes and functional form	2.6 - APIs and policy applications (working with Census data)
Week 4	1.7 - Panel data designs: fixed effects, first differences	2.7 - Programming concepts (for loops, functions, control flow)
	1.8 - Difference in Differences Design	2.8 - Causal inference stats in R (lm, sample, distributions, stargazer)
Week 5	1.9 - Regression discontinuity designs	2.9 - Introduction to spatial data (sf, tmap, ggmap)
	1.10 - Instrumental variables	2.10 - Literate programming (RMarkdown, code syntax), GitHub
	<b>Capstone Research Project</b>	
Weeks 6 & 7	3.1 - Capstone Project Kick-off Meeting 3.2 - Policy Memo Writing Workshop 3.3 - Capstone Working Group 3.4 - Capstone Mid-cycle Check Meeting 3.5 - Capstone Presentation Summit	

## Credential Award

Upon successful completion of the coursework, participants will receive two documents, issued electronically, upon successful completion of the credential program:

- **Official UChicago Transcript** with pass/fail marks (non-credit) for each academic component: Data Analytics, R Programming, and Capstone research project. Students will be able to access this transcript at the end of summer (approximately early October), and will have lifetime access to request a transcript from the University Registrar. This is a non-credit program.
- **Certificate of Credential Completion**, issued by the Harris School of Public Policy, contains a grade for each academic component: Data Analytics, R Programming, and Capstone research project. You will receive an email from our office ([harriscredential@uchicago.edu](mailto:harriscredential@uchicago.edu)) with the electronic certificate approximately four weeks after the program.

## 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, use un-allowed materials during exams, or otherwise gain unfair academic advantage.

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 Exchange. 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.

Please see: The University of Chicago Student Manual: <https://studentmanual.uchicago.edu/Policies>