

Syllabus

SMPA 2152: Data Analysis for Journalism and Political Communication (Spring 2025)

Meeting Times: Mondays and Wednesdays, 8:00-9:15am

Classroom: MPA 307

Professor: Nicholas Bell, Ph.D. (he/him/his)
nicholasbell@gwu.edu

Office Hours (on Zoom): Wednesdays 7:30-8:30pm
(Zoom link for office hours:
<https://gwu-edu.zoom.us/j/94237095955>)

I prefer to meet during office hours or by appointment.
However, I am available by email, and I try to respond to
emails by the end of the next business day (M-F).

Course Description

Data has been democratized. More data is available to the ordinary person than ever before, and leaders in every industry – including journalism and political communication – want to join the big data revolution. However, most of us lack the data literacy skills to make good use of these resources, and this can lead to the misapplication and misuse of data. To fully leverage the promise of big data, we must become familiar with the basic challenges inherent in data analysis and how to overcome them. This course is an introduction to the principles and practices of data analysis. The goal is for students to become responsible consumers and producers of data. Students will learn how to critically evaluate claims derived from data. Students will also learn how to ethically present data in compelling and persuasive ways to non-expert audiences. This class includes a special discussion of political polling, which is widely used in journalism and political communication but has come under increasing scrutiny in recent years. Students require only a basic aptitude in numeracy (e.g. percentages and averages) for this course.

In addition to developing data literacy, students will be introduced to the R programming language. There are many advantages to learning R: it is free and open-source, meaning that developers are continually releasing new tools to make coding easier; it is widely used by news organizations and researchers around the world; and R is one of the most powerful programming languages for statistical analysis. Students will

learn data literacy by applying the same tools and techniques used by professional data scientists.

Learning Objectives

1. You will be able to assess the pragmatic and ethical issues in collecting, manipulating, and analyzing data, known as “data literacy.”
2. You will be able to obtain publicly-available data and perform basic manipulations on that data using the programming language R.
3. You will be able to visualize and present data in accurate and persuasive ways.
4. You will be familiar with the statistical concepts of sampling, uncertainty, hypothesis testing, and linear regression and how to conduct basic statistical analyses in R.

Course Materials

Students are required to purchase a [Posit Cloud](#) student access plan for the duration of the course. Posit Cloud is a web-based version of the popular R development environment RStudio. Even if you are familiar with the desktop version of RStudio, you will need to purchase a Posit Cloud membership. All of the code we work on during this class will be available on Posit Cloud. At the time of this writing, a membership costs \$5 per month. If the course materials pose a financial burden, please contact me and we will work something out.

Time Required

This course requires 2.5 hours of direct instruction and a minimum of 5 hours of independent learning per week for a combined minimum total of 7.5 hours per week or 112.5 hours per semester.

Assessment

Your course grade is calculated as your grade on each of the following course components weighted by:

Assignments	35%
Mid-term exam	20%
Final project	35%
Attendance	10%

Course grades are converted into letter grades according to the following rubric:

93-100 = A (4.0 GPA points)
90-92 = A- (3.7 GPA points)
87-89 = B+ (3.3 GPA points)
83-86 = B (3.0 GPA points)
80-82 = B- (2.7 GPA points)
77-79 = C+ (2.3 GPA points)
73-76 = C (2.0 GPA points)
70-72 = C- (1.7 GPA points)
67-69 = D+ (1.3 GPA points)
63-66 = D (1.0 GPA points)
60-62 = D- (0.7 GPA points)

Assignments

There will be eight assignments in this course that will ask you to apply the R skills we learn in class. You may complete these assignments on your own or in collaboration with other students. This means that you may work together to write code and/or solve problems. Do not split up the questions or combine independent work. **It is a violation of the academic integrity policy to submit any code to which you did not contribute as your own.** If you work with other students, please indicate their names at the top of your submission. Each student must submit an assignment on Blackboard.

Your lowest assignment grade will be dropped. This means that your highest seven assignment grades will each contribute 5% to your overall course grade, for a total of 35%.

Assignments are typically due by 11:59pm Eastern of the following Sunday. This means that you have 6 days to complete each assignment. Late assignments will not be accepted after the start of class on Monday, unless you communicate with the professor before the deadline. We will review the solutions to the assignment at the start of our class meeting on Monday.

Mid-term Exam

The mid-term exam is scheduled for March 3 from 8:00-9:15am on Blackboard and may be completed from the location of your choosing. The mid-term exam will cover material from the data analysis track through February 12. A review session is scheduled for February 26.

Final Project

Your final project is a data storytelling exercise that is due at the conclusion of the final exam block assigned to this course (there is no in-person final exam). More in-

formation will be provided during the course.

There will be no class on Wednesday, April 30 (the designated Monday). This time is reserved for one-on-one meetings with the professor to discuss your final project.

Attendance

Attendance is mandatory. If you miss a class due to illness, family emergencies, University-scheduled events, and other unusual circumstances, you must email me and let me know. You do not need to provide proof of your reason for missing class, but misrepresentation of your reason for excusal is a violation of the [Code of Student Conduct](#).

Your attendance grade is the percentage of class meetings with an unexcused absence deducted from 100 (rounded up). For example, if you have two unexcused absences, your attendance grade is $100 - ((2/27) * 100) = 93$.

University Policy on Observance of Religious Holidays

Students must notify faculty during the first week of the semester in which they are enrolled in the course, or as early as possible, but no later than three weeks prior to the absence, of their intention to be absent from class on their day(s) of religious observance. If the holiday falls within the first three weeks of class, the student must inform faculty in the first week of the semester. For details and policy, see "Religious Observance Policy" at provost.gwu.edu/policies-procedures-and-guidelines.

Support for Students with Disabilities

Any student who may need an accommodation based on the impact of a disability should contact the Office of Disability Support Services (DSS) to inquire about the documentation necessary to establish eligibility and to coordinate a plan of reasonable and appropriate accommodations. DSS is located in Rome Hall, Suite 102. For additional information, please call DSS at 202-994-8250, or consult disabilitysupport.gwu.edu/.

Academic Integrity

Academic integrity is an essential part of the educational process, and all members of the GW community take these matters very seriously. As the instructor for this course, my role is to provide clear expectations and uphold them in all assessments. Violations of academic integrity occur when students fail to cite research sources properly, engage in unauthorized collaboration, falsify data, and otherwise violate the [Code of Academic Integrity](#). If you have any questions about whether or not particular academic practices or resources are permitted, you should ask me for clarification. If you are reported for an academic integrity violation, you should contact Conflict Education and Student Accountability (CESA), formerly known as Student Rights and Responsibilities (SRR), to learn more about your rights and options in the process.

Consequences can range from failure of assignment to expulsion from the university and may include a transcript notation. For more information, please refer to the [CESA website](#), email rights@gwu.edu, or call 202-994-6757.

Course Policy on Generative AI

Generative Artificial Intelligence (GAI) tools are increasingly used in academic and professional settings to efficiently complete certain specific tasks. Therefore, this course permits limited use of GAI tools on work submitted for evaluation, such as homework assignments and exams.

What is permissible:

- Using GAI tools to answer questions, similar to using a search engine. For example, "What is the function to generate a mean in R?" or "What does 'Error: object not found' mean in R?" would be permissible.
- Using GAI tools to generate code that is not specific to a data set. For example, if you are asked to produce a line graph of presidential approval, it would be permissible to ask a GAI tool to provide you with an example of R code to generate a line graph.

What is not permissible:

- Using GAI tools to generate code that is specific to a data set. For example, if you are asked to produce a line graph of presidential approval, it would be impermissible to provide the data to a GAI tool and make only minor modifications to the resulting code.
- Using GAI tools to debug (fix) code that you have written. This is a valuable use of GAI in the professional world, but relying on GAI tools to debug your code in this class will hinder your learning.
- Any other use not specifically permitted under the "What is permissible" section above.

By submitting work for evaluation in this course, you represent it as your own intellectual product. Impermissible use of GAI tools for work submitted for evaluation constitutes cheating under the GW's [Code of Academic Integrity](#). If you have any questions about the application of this policy (or any other questions about academic integrity in this course), please email the instructor.

Class Recordings and Use of Electronic Course Materials

Class meetings will be audio/video recorded and made available to other students in this course. As part of your participation in this course, you may be recorded. If you do not wish to be recorded, please contact me during the first week of class to discuss alternative arrangements.

Students are encouraged to use electronic course materials, including recorded class sessions, for private personal use in connection with their academic program of study. Electronic course materials and recorded class sessions should not be shared or used for non-course related purposes unless express permission has been granted by the instructor. Students who impermissibly share any electronic course materials are subject to discipline under the [Code of Student Conduct](#). Please contact the instructor if you have questions regarding what constitutes permissible or impermissible use of electronic course materials and/or recorded class sessions. Please contact [Disability Support Services](#) if you have questions or need assistance in accessing electronic course materials.

Additional Resources for Students

Academic Support

- **Academic Commons**

Academic Commons is the central location for academic support resources for GW students. To schedule a peer tutoring session for a variety of courses visit go.gwu.edu/tutoring. Visit academiccommons.gwu.edu for study skills tips, finding help with research, and connecting with other campus resources. For questions email academiccommons@gwu.edu.

- **Writing Center**

GW's Writing Center cultivates confident writers in the University community by facilitating collaborative, critical, and inclusive conversations at all stages of the writing process. Working alongside peer mentors, writers develop strategies to write independently in academic and public settings. Appointments can be booked online at gwu.mywconline.com.

Health and Wellness Support

- **Disability Support Services**

202-994-8250

Any student who may need an accommodation based on the potential impact of a disability should contact Disability Support Services at disabilitysupport.gwu.edu to establish eligibility and to coordinate reasonable accommodations.

- **Student Health Center**

202-994-5300

The Student Health Center (SHC) offers medical, counseling/psychological, and psychiatric services to GW students. More information about the SHC is available at healthcenter.gwu.edu. Students experiencing a medical or mental health emergency on campus should contact GW Emergency Services at 202-994-6111, or off campus at 911.

GW Campus Emergency Information

GW Emergency Services: 202-994-6111 For situation-specific instructions, refer to [GW's Emergency Procedures guide](#).

GW Alert

GW Alert is an emergency notification system that sends alerts to the GW community. GW requests students, faculty, and staff maintain current contact information by logging on to alert.gwu.edu. Alerts are sent via email, text, social media, and other means, including the Guardian app. The Guardian app is a safety app that allows you to communicate quickly with GW Emergency Services, 911, and other resources. Learn more at safety.gwu.edu.

Protective Actions

GW prescribes four protective actions that can be issued by university officials depending on the type of emergency. All GW community members are expected to follow directions according to the specified protective action. The protective actions are Shelter, Evacuate, Secure, and Lockdown (details below). Learn more at safety.gwu.edu/gw-standard-emergency-statuses.

- Shelter
 - Protection from a specific hazard.
 - The hazard could be a tornado, earthquake, hazardous material spill, or other environmental emergency.
 - Specific safety guidance will be shared on a case-by-case basis.
 - **Action:**
 - * Follow safety guidance for the hazard.
- Evacuate
 - Need to move people from one location to another.
 - Students and staff should be prepared to follow specific instructions given by first responders and University officials.

- **Action:**
 - * Evacuate to a designated location.
 - * Leave belongings behind.
 - * Follow additional instructions from first responders.
- Secure
 - Threat or hazard outside of buildings or around campus.
 - Increased security, secured building perimeter, increased situational awareness, and restricted access to entry doors.
 - **Action:**
 - * Go inside and stay inside.
 - * Activities inside may continue.
- Lockdown
 - Threat or hazard with the potential to impact individuals inside buildings.
 - Room-based protocol that requires locking interior doors, turning off lights, and staying out of sight of corridor window.
 - **Action:**
 - * Locks, lights, out of sight.
 - * Consider Run, Hide, Fight.

Classroom emergency lockdown buttons

Some classrooms have been equipped with classroom emergency lockdown buttons. If the button is pushed, GWorld Card access to the room will be disabled, and GW Dispatch will be alerted. The door must be manually closed if it is not closed when the button is pushed. Anyone in the classroom will be able to exit, but no one will be able to get in.

Course Outline

This course consists of two parallel tracks, a Coding track that will be taught on Mondays and a Data Analysis track that will be taught on Wednesdays. The Coding track will teach you how to conduct data analyses using the statistical programming language R. The Data Analysis track will introduce you to key principles of data analysis and core concepts in statistics.

The Data Analysis track homework includes assigned readings. The readings assigned for each class should be completed before the following meeting of that track, e.g., the readings assigned in the Data Analysis track on January 29 are to be completed by February 5. Readings with an embedded link can be accessed online. All other readings are available on Blackboard.

For the Coding track, I provide links to reference texts that provide additional information on the material covered during that day's class. The reference texts are not assigned readings. References to Ismay & Kim refer to Statistical Inference via Data Science: A ModernDive into R and the Tidyverse by Chester Ismay and Albert Y. Kim, an open-source online textbook that is available at <https://www.moderndive.com>.

Week	Coding Track (Mondays)	Data Analysis Track (Wednesdays)
1	January 13: Introduction Homework (for January 22) • Fry (2021), "When Graphs Are a Matter of Life and Death" (The New Yorker) (on Blackboard)	January 15: Introduction to R References • Ismay & Kim, Preface: Introduction for Students • Ismay & Kim, Chapters 1.1 to 1.3
2	January 20: No Class (MLK Day/Inauguration Day)	January 22: Data Visualization Homework • Retro Report (2021), "What's in a Number?" (video)
3	January 27: Data Visualization I References • Ismay & Kim, Chapters 2.1 to 2.5 Homework • Problem Set #1 due February 2	January 29: Researcher Choices and Bias Homework • Aviv (2024), "Conviction" (New Yorker) (on Blackboard)
4	February 3: Data Visualization II References • Ismay & Kim, Chapters 2.6 to 2.9 Homework • Problem Set #2 due February 9	February 5: Correlation vs. Causation Homework • Bouie (2021), "Quantifying the Pain of Slavery" (New York Times)
5	February 10: Quarto References • Wickham, Çetinkaya-Rundel, and Grolemund, <u>R for Data Science</u> , 2nd ed., Ch. 29: Quarto	February 12: Data Ethics and Responsibility Homework • Complete CITI Research Ethics training (due February 18)
6	February 17: No Class (President's Day)	February 19: Sampling Homework (for March 19) • Excerpt of Morris (2022), Strength in Numbers (on Blackboard)

7	<p>February 24: Data Wrangling I</p> <p><u>References</u></p> <ul style="list-style-type: none"> • Ismay & Kim, Chapter 4.1 • Ismay & Kim, Chapters 3.1 to 3.6 <p><u>Homework</u></p> <ul style="list-style-type: none"> • Problem Set #3 due February 28 	<p>February 26: Review for Mid-term Exam</p> <p><u>Homework</u></p> <ul style="list-style-type: none"> • None
8	<p>March 3: Mid-term exam</p>	<p>March 5: No class</p> <p><u>Homework</u></p> <ul style="list-style-type: none"> • None (enjoy Spring Break!)
9	<p>March 17: Data Wrangling II (video lecture)</p> <p><u>References</u></p> <ul style="list-style-type: none"> • Ismay & Kim, Chapters 3.7 to 3.9 • Ismay & Kim, Appendix C.1 • Ismay & Kim, Chapter 4.2 <p><u>Homework</u></p> <ul style="list-style-type: none"> • Problem Set #4 due March 23 	<p>March 19: Political Polling</p> <p><u>Homework</u></p> <ul style="list-style-type: none"> • Excerpt of Hickey (2023), You Are What You Watch: How Movies and TV Affect Everything (on Blackboard)
10	<p>March 24: Political Polling II</p> <p><u>Homework</u></p> <ul style="list-style-type: none"> • None 	<p>March 26: Guest Speaker: Walt Hickey, Pulitzer Prize-winning data journalist, author of "You Are What You Watch" (Workman, 2023), and Executive Editor of Sherwood News</p> <p><u>Homework</u></p> <ul style="list-style-type: none"> • Excerpt of Tetlock and Gardner (2015), Superforecasting: The Art and Science of Prediction (on Blackboard)
11	<p>March 31: Analyzing Polls in R</p> <p><u>Homework</u></p> <ul style="list-style-type: none"> • Problem Set #5 due April 6 	<p>April 2: Election Prediction Models</p> <p><u>Homework</u></p> <ul style="list-style-type: none"> • None

12	<p>April 7: Guest Speaker: Matthew Herdman, Democratic communications and campaign strategist</p> <p><u>Homework</u></p> <ul style="list-style-type: none"> • None 	<p>April 9: Guest Speaker: Jennifer Agiesta, Director of Polling and Election Analytics at CNN</p> <p><u>Homework</u></p> <ul style="list-style-type: none"> • Final Project Data Set due April 13
13	<p>April 14: Hypothesis Testing (lecture)</p> <p><u>Homework</u></p> <ul style="list-style-type: none"> • None 	<p>April 16: Hypothesis Testing (coding)</p> <p><u>Homework</u></p> <ul style="list-style-type: none"> • Problem Set #6 due April 22 • Final Project Wrangling Code due April 23
14	<p>April 21: Regression (lecture)</p> <p><u>Homework</u></p> <ul style="list-style-type: none"> • None 	<p>April 23: Regression (coding)</p> <p><u>Homework</u></p> <ul style="list-style-type: none"> • None
15	<p>April 28: Putting it all together (lecture)</p> <p><u>Homework</u></p> <ul style="list-style-type: none"> • None 	<p>April 30: No class (one-on-one meetings with the professor on April 24-25)</p>

Version: 4
Last Updated: April 16, 2025
Subject to change.