

Syllabus

SMPA 2152: Data Analysis for Journalism and Political Communication (Fall 2024)

Meeting Times: Mondays and Wednesdays, 5:45-7:00pm

Classroom: MPA B01

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

Office Hours (on Zoom): Thursdays 7:30-8:30pm and by appointment
(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.

You must complete all eight assignments, however, 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 due by 11:59pm Eastern of the following Sunday. This means that you have 6 days to complete each assignment. Late assignments will be deducted 50%, and no assignments will be accepted after the start of class on Monday. 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 October 23 and it is an in-class exam. The mid-term exam will cover material from both the data analysis and coding tracks through October 16. A review session is scheduled for October 21.

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 information will be provided during the course.

There will be no class on December 9. 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 readings assigned for each class should be completed before the following meeting of that track, e.g., the readings assigned in the Coding track on September 9 are to be completed by September 16, while the readings assigned in the Data Analysis track on September 11 are to be completed by September 18. Readings with an embedded link can be accessed online. All other readings are available on Blackboard. 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	<p>August 26: Introduction</p> <p>Homework (for Wednesday)</p> <ul style="list-style-type: none"> • Ismay & Kim, Preface: Introduction for Students • Ismay & Kim, Chapters 1.1 to 1.3 <p>Homework (for Sept. 4)</p> <ul style="list-style-type: none"> • Fry (2021), "When Graphs Are a Matter of Life and Death" (The New Yorker) (on Blackboard) 	<p>August 28: Introduction to R</p> <p>Homework</p> <ul style="list-style-type: none"> • None
2	<p>September 2: No Class (Labor Day)</p>	<p>September 4: Data Visualization</p> <p>Homework (for Monday)</p> <ul style="list-style-type: none"> • Ismay & Kim, Chapters 2.1 to 2.5 <p>Homework (for Wednesday)</p> <ul style="list-style-type: none"> • Retro Report (2021), "What's in a Number?" (video)
3	<p>September 9: Data Visualization I</p> <p>Homework</p> <ul style="list-style-type: none"> • Ismay & Kim, Chapters 2.6 to 2.9 • Assignment #1 due September 15 	<p>September 11: Researcher Choices and Bias</p> <p>Homework</p> <ul style="list-style-type: none"> • Aviv (2024), "Conviction" (New Yorker) (on Blackboard)
4	<p>September 16: Data Visualization II</p> <p>Homework</p> <ul style="list-style-type: none"> • Wickham, Çetinkaya-Rundel, and Grolemund, R for Data Science, 2nd ed., Ch. 29: Quarto • Assignment #2 due September 22 	<p>September 18: Correlation vs. Causation</p> <p>Homework</p> <ul style="list-style-type: none"> • Bouie (2021), "Quantifying the Pain of Slavery" (New York Times)
5	<p>September 23: Quarto</p> <p>Homework</p> <ul style="list-style-type: none"> • Ismay & Kim, Chapter 4.1 • Ismay & Kim, Chapters 3.1 to 3.6 	<p>September 25: Data Ethics and Responsibility</p> <p>Homework</p> <ul style="list-style-type: none"> • TBA

6	September 30: Data Wrangling I Homework <ul style="list-style-type: none"> • Ismay & Kim, Chapters 3.7 to 3.9 • Ismay & Kim, Appendix C.1 • Ismay & Kim, Chapter 4.2 • Assignment #3 due October 6 	October 2: Hold for Guest Speaker Homework <ul style="list-style-type: none"> • Rosenbaum (1970), "Statisticians Charge Draft Lottery Was Not Random" (New York Times) (on Blackboard)
7	October 7: Data Wrangling II Homework <ul style="list-style-type: none"> • Wickham, Çetinkaya-Rundel, and Golemund, R for Data Science, 2nd ed., Chs. 15, 17, & 18: Strings, Factors, and Dates and times • Assignment #4 due October 13 	October 9: Sampling Homework <ul style="list-style-type: none"> • Excerpt of Morris (2022), Strength in Numbers (on Blackboard)
8	October 14: Data Wrangling III Homework <ul style="list-style-type: none"> • Assignment #5 due October 20 	October 16: Political Polling Homework (for October 30) <ul style="list-style-type: none"> • Excerpt of Tetlock and Gardner (2015), Superforecasting: The Art and Science of Prediction (on Blackboard)
9	October 21: Review for Mid-term Exam Homework <ul style="list-style-type: none"> • None 	October 23: Mid-term exam Homework <ul style="list-style-type: none"> • None
10	October 28: Analyzing Polls in R Homework <ul style="list-style-type: none"> • Assignment #6 due November 3 	October 30: Election Prediction Models Homework <ul style="list-style-type: none"> • Beeler (2014), "Science by the numbers: Researchers ask, 'How true are our findings?'" (WHYY) (audio; on Blackboard)
11	November 4: Making Maps in R Homework <ul style="list-style-type: none"> • Assignment #7 due November 10 	November 6: Hypothesis Testing Homework <ul style="list-style-type: none"> • TBA

12	November 11: Hypothesis Testing <u>Homework</u> • Excerpt of Gelman, Hill, & Vehtari (2020), <u>Regression and Other Stories</u> (on Blackboard) • Assignment #8 due November 17	November 13: Guest Speaker: Matthew Herdman, Senior Communications Strategist, 50+1 Strategies <u>Homework</u> • TBA
13	November 18: Regression (lecture) <u>Homework</u> No homework	November 20: Guest Speaker: Emily Guskin, Deputy Polling Director, The Washington Post <u>Homework</u> No homework
14	December 2: Regression (coding) <u>Homework</u> No homework	December 4: Putting it all together <u>Homework</u> No homework
15	December 9: One-on-one meetings	
Final project due date TBA		

Version: 2
Last Updated: August 26, 2024
Subject to change.