# PADM-GP 4505: R Coding for Public Policy

R is among the most popular coding and development packages from the new generation of powerful and versatile softwares used in public policy and other research settings.

Contemporary data engineering and analysis in the service of public policy depends on a mastery of interlocking tools, including statistical techniques, visualization, and reproducibility; therefore the goal of the course is to lead students into the R world, help them master the basic tools, approaches, and thinking, and establish a platform for future independence and expertise in these tools. This course offers students basic programming, data engineering, and data analysis skills in R in the context of public policy-making and policy evaluation, with a particular emphasis on balancing theory with implementation.

# Course and Learning Objectives

Students who successfully complete this course will install R and RStudio and become familiar with the IDE, understand and utilize core R concepts such as objects and commands from a number of key libraries, and utilize best-practices for project reproducibility and management. The course will also emphasize use cases for R, focusing on cleaning, exploring, and analyzing data.

## Takeaways

Upon completion of the course, you will be able to:

1. Install and set up R and RStudio
2. Find, install, and use R packages
3. Understand basic programming concepts and how they apply to the R language
4. Read, manipulate, and clean data
5. Plot simple, clear graphics for effective policy analysis
6. Conduct regression using R packages
7. Apply data management best practices using R
8. Demonstrate additional insight into software ecosystem elements like LaTeX and GitHub

## Instructor

Emil Hafeez, MS, MSPH

Assistant Adjunct Professor of Public Service emil.hafeez@nyu.edu

## Class Sessions

194 Mercer Street, Room 204 9/7/2023 - 10/19/2023

6:45 PM - 8:25 PM

# Learning Resources

## Hardware

* Access to a computer and an internet connection is necessary for completing the course goals, and as such, bringing a laptop to each class is important enough for me to make it necessary. If you are not able to bring a laptop to class, contact Emil at emil.hafeez@nyu.edu. Please note that there are available resources from NYU, including here.

## Software

* R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://[www.R-project.org/.](http://www.R-project.org/)
* RStudio Team (2020). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL [http://www.rstudio.com/.](http://www.rstudio.com/)

## Recommended References

* The Internet (StackOverflow, Google, blog posts)
* R’s `?` operator
* R for Data Science 2e by H Wickham, M Çetinkaya-Rundel, and G Grolemund:
	+ H. Wickham, M. Çetinkaya-Rundel, G. Grolemund (2023). R for Data Science. O'Reilly Media. R for Data Science.
* Exploratory Data Analysis with R by R Peng:
	+ Peng, R. D. (2020). Exploratory Data Analysis with R. Springer. Exploratory Data Analysis with R.
* R Programming for Data Science by R Peng
	+ Peng, R. D. (2022). R Programming for Data Science. Leanpub. R Programming for Data Science.
* Advanced R by H Wickham:
	+ Wickham, H. (2023). Advanced R. Chapman and Hall/CRC. Advanced R.
* R Graphics Cookbook by W Chang:
	+ Chang, W. (2023). R Graphics Cookbook. O'Reilly Media. R Graphics Cookbook.
* R Cookbook by JD Long and P Teetor:
	+ Long, J. D., & Teetor, P. (2019). R Cookbook. O'Reilly Media. R Cookbook.
* Pro Git by S. Chacon and B Straub:
	+ Chacon, S., & Straub, B. (2023). Pro Git. Apress. Pro Git.
* R Packages by H Wickham
	+ Wickham, H. (2020). R Packages. O'Reilly Media. R Packages.

## Brightspace

All announcements, resources, and assignments will be delivered through the Brightspace site. I may modify assignments, due dates, and other aspects of the course as we go through the term, with advance notice provided as soon as possible through the course website. You are expected to check the Brightspace and your student email account regularly.

# Assignments and Evaluation

The Course Grade is based on the following:

## Individual Assessment

* 5 Assignments: 90%
* Participation: 10%

The instructions for each assignment will be released after the lecture and the assignment is due before class the following week, or in some choice circumstances upon completion of each course topic when specified directly. Participation refers to your presence and engagement in class, as well as responding to any class surveys and/or Brightspace queries posted by the teaching team.

## Late Work Policy

Points will be deducted from late submissions, 15% of full credit for each late day. Extensions will be granted only on a case by case basis and in case of emergency.

## Collaboration Policy

The course policy on collaboration applies to all homework assignments and is as follows.

The course goal is to scaffold students' future independence and mastery of R by supplying challenging but fundamental skills – and accordingly, it's recognized that learning rarely occurs in isolation, and many data-oriented professional workflows occur in teams. Students are encouraged to collaborate in a way that improves their own individual understandings of the material without derogating from anyone else's. This may include reviewing lecture material and recommended references, reviewing examples, and even coarse-grained outlines to assignment problems. However, note that fair assessment of each individual is critical to any course, and as such any work submitted in completion of an assignment must reflect the individual's own skill, effort, and development. You are not allowed to copy code, literally or otherwise. This includes from classmates, code found online or in references, or generated by LLMs like ChatGPT or Bard.

If you’re unsure about what this means or if you do not know if something is permitted, you’re welcome to email Emil Hafeez, emil.hafeez@nyu.edu. My guess is that if you have to ask, it’s probably not allowed. If you’re having trouble keeping up with the course and want to stay

consistent with the collaboration policy, it’s recommended you attend an office hour or reach out to Emil.

## Grading Scale and Rubric

Students will receive grades according to the following scale:

* A = 4.0 points
* A- = 3.7 points
* B+ = 3.3 points
* B = 3.0 points
* B- = 2.7 points
* C+ = 2.3 points
* C = 2.0 points
* C- = 1.7 points
* There are no D+/D/D-
* F (fail) = 0.0 points

Student grades will be assigned according to the following criteria:

1. Excellent: Exceptional work for a graduate student. Work at this level is unusually thorough, well-reasoned, creative, methodologically sophisticated, and well written. Work is of exceptional, professional quality.

(A-) Very good: Very strong work for a graduate student. Work at this level shows signs of creativity, is thorough and well-reasoned, indicates strong understanding of appropriate methodological or analytical approaches, and meets professional standards.

(B+) Good: Sound work for a graduate student; well-reasoned and thorough, methodologically sound. This is the graduate student grade that indicates the student has fully accomplished the basic objectives of the course.

1. Adequate: Competent work for a graduate student even though some weaknesses are evident. Demonstrates competency in the key course objectives but shows some indication that understanding of some important issues is less than complete. Methodological or analytical approaches used are adequate but student has not been thorough or has shown other weaknesses or limitations.

(B-) Borderline: Weak work for a graduate student; meets the minimal expectations for a graduate student in the course. Understanding of salient issues is somewhat incomplete. Methodological or analytical work performed in the course is minimally adequate. Overall performance, if consistent in graduate courses, would not suffice to sustain graduate status in “good standing.”

(C/-/+) Deficient: Inadequate work for a graduate student; does not meet the minimal expectations for a graduate student in the course. Work is inadequately developed or flawed by numerous errors

and misunderstanding of important issues. Methodological or analytical work performed is weak and fails to demonstrate knowledge or technical competence expected of graduate students.

1. Fail: Work fails to meet even minimal expectations for course credit for a graduate student. Performance has been consistently weak in methodology and understanding, with serious limits in many areas. Weaknesses or limits are pervasive.

## Style Note

The use of proper style is important in this course and when writing code; for that reason, in addition to the points allocated for particular homework assignments, points can be deducted for coding and style issues including:

* + Incomplete sentences and particularly sparse explanation
	+ Poor text formatting and legibility
	+ Inappropriate spacing and indentation in code
	+ Unclear variable naming conventions
	+ Disorganized code and file structure

Whitespace and organization is covered by most R style guides including the Advanced R textbook. Occasional typos won’t be penalized but consistent or severe errors may be.

# Overview of the Course

Expect to come to class with a laptop, and participate in some combination of lecture, demonstration, practical implementation, troubleshooting, and collaborative work. Each week with a deliverable listed, you can expect to turn in this assignment before class.

* Week 1
	+ Topic: Introduction to R and Rstudio
	+ Deliverable: Install R and RStudio, Create a GitHub account
* Week 2
	+ Topic: Data objects, functions, and the tidyverse
	+ Deliverable: Assignment 1
* Week 3
	+ Topic: Application I: Data quality & cleaning data
	+ Deliverable: Assignment 2
* Week 4
	+ Topic: Graphics in R
	+ Deliverable: Assignment 3
* Week 5
	+ Topic: Application II: Exploratory Data Analysis
	+ Deliverable: Assignment 4
* Week 6
	+ Topic: R for analyses
	+ Deliverable: Assignment 5
* Week 7
	+ Topic: Other real-world tools & applications

# Additional Policies and Resources

## Academic Integrity

Academic integrity is a vital component of Wagner and NYU. All students enrolled in this class are required to read and abide by Wagner’s Academic Code. All Wagner students have already read and signed the Wagner Academic Oath. Plagiarism of any form will not be tolerated and students in this class are expected to report violations to me. If any student in this class is unsure about what is expected of you and how to abide by the academic code, you should consult with me.

## Henry and Lucy Moses Center for Student Accessibility

Academic accommodations are available for students with disabilities. Please visit the Moses Center for Students with Disabilities (CSD) website and click the “Get Started” button to apply for accomodation. You can also call or email CSD (212-998-4980 or mosescsd@nyu.edu) for information. Students who are requesting academic accommodations are strongly advised to reach out to the Moses Center as early as possible in the semester for assistance.

## NYU’s Calendar Policy on Religious Holidays

NYU’s Calendar Policy on Religious Holidays states that members of any religious group may, without penalty, absent themselves from classes when required in compliance with their religious obligations. You must notify Emil in advance of religious holidays or observances that might coincide with exams, assignments, or class times to schedule mutually acceptable alternatives. Students may also contact religiousaccommodations@nyu.edu for assistance.

## NYU’s Wellness Exchange

NYU’s Wellness Exchange has extensive student health and mental health resources. A private hotline (212-443-9999) is available 24/7 that connects students with a professional who can help them address day-to-day challenges as well as other health-related concerns.