

SPP 608 Statistics for Public Policy in R

Lecture and Lab (4 credits)

*A *Statistic* Class – Spring 2024*

Lecture: Friday 1:30 - 4:00pm | Lab: Tuesday 2:00 - 3:00pm

Instructor: Viviana Chiu Sik Wu, LLB, MS, PhD

(Dr. Wu / Professor Wu)

Email: vivianachius@umass.edu

Office: Thompson 628

Office hours: Friday 12-1pm or by appointment

Website: <https://connectivecommons.cc/>

TA: Kiam Jamrog-McQuaid

Email: kjamrogmcqua@umass.edu

Office hours: Monday 11-1

(by appointment)

Lab: [Zoom Link](#)

Course Objectives

This is an introductory statistical course designed for public policy students to analyze and solve statistical problems in the real world. The goal of this course is to familiarize you with a number of statistical techniques commonly used for analyzing different types of policy data. Throughout the course, it emphasizes two crucial aspects of students' learning: 1) statistical analysis for policy data, and 2) presentation and interpretation of statistical results. While the course will cover some basic theories behind most statistical methods, it will focus mostly on their applications. You will learn when each statistical test should be used and the assumptions behind each test. In addition, you will learn how to describe and interpret the output of the analyses they run. Methods covered in this course include exploratory data analysis, correlation and bivariate analysis, linear regression, panel data regression, and probit/logistic regression. This course is designed around the concepts of **flipped classroom, collaborative and experimental learning**, with plenty of hands-on exercises and problem-solving on a project basis similar to policy consultants. The free R open-source software will be used for conducting statistical analysis and data visualization.

Overview of Class Design

This class will have a combination of in-person lectures, discussions and R demonstration. We will meet every Friday for lectures. The lab session on Tuesday will go through the R tutorial in detail and review lab assignments in real time. There will be weekly lab assignments. Students will partner up and work on a data analysis project throughout the semester. Collective and reflective learning will be the key cornerstone for excelling in this class!

Course Values

- Empathy: Care and respect for each other and be empathetic and compassionate.
- Openness: Be proactive in asking questions. Be willing to learn by trial and error.
- Integrity: Take responsibility for your own and teamwork. Be fair and trustworthy.
- Professionalism: Have high standards and solid preparation for the class and the team.
- Perseverance and passion: Actively engage and participate to strive for your own learning goals.

Class Participation Agreement **(Resources)**

- Come from curiosity and generosity not judgment.
- Embrace multiple perspectives and a diversity of voices.
- Share talking time. Step Forward/Step Back.
- One person speaks at a time. Avoid interrupting and talking over others.
- Acknowledge what another person has said. Ask clarifying questions.
- Challenge ideas and arguments not people.
- Agree to take a problem-solving approach when conflicts arise.
- Write down thoughts or questions if there is no time or you don't feel safe to voice them during the discussion.
- Acknowledge that stereotypes, bias, discrimination, and oppression based on race/ethnicity, social class, gender, sexuality, etc. exist and that we will actively try to combat them

Required Software

Installing R & R Studio

Bring your own personal computer to each class. You will need to **first install R** (the free and open source statistical computing language used in this course), **then R Studio** (the graphical user interface for R) *in advance before the first class.*

1. [Install R](#) [Base version: Download the latest version R 4.3.2]
2. [Install R Studio Desktop](#) [Free version]

The R statistical software is used. There are infinitely many new packages available; a pretty interface to explore the publicly available R packages is available via Microsoft. Using the free RStudio, an interface for writing R documents and working with data is highly recommended.

You will use R Studio Desktop (on your computer) to work on your assignment. You can also use R Markdown (.Rmd file) to "knit" your assignment and quiz submissions in the .pdf format. Alternatively, you can use R script (.R file) for the assignments and quizzes and transfer your codes and results screenshots to a Word file. More instructions will follow in the first few labs.

Recommended Readings

All lab materials and slides are accessible on RPubS: <https://rpubs.com/Vivianacswu>

We will adopt a **reference textbook (ITER)** for this class; readings for each week will be posted on Canvas. Other textbooks and resources are listed below.

[Reference textbook] Hanck, C., Arnold, M., Gerber, A., & Schmelzer, M. (2019). *Introduction to Econometrics with R*. [ITER] Essen: University of Duisburg-Essen. Available here: <https://www.econometrics-with-r.org/ITER.pdf>

Walker, K. (2023). *Analyzing US Census Data: Methods, Maps, and Models in R*. Chapman & Hall. Retrieved from <https://walker-data.com/census-r/mapping-census-data-with-r.html>

Gillespie, B. J., Hibbert, K. C., & Wagner III, W. E. (2020). *A guide to R for social and behavioral science statistics*. SAGE Publications, Incorporated. (\$45)

https://www.amazon.com/dp/1544344023/ref=cm_sw_em_r_mt_dp_8qU2FbKQT9S75?encoding=UTF8&psc=1

Assignments and Gradings

Your grade for this course will be determined by the following:

- Class Participation 10%
- Assignments 30%
- Quizzes 30%
- Final project 30%

Class Participation (10% of grade)

This year's course format will be interactive and engaging, combining in-person lectures, discussions, small group exercises, and collaborative research projects. We will optimize learning by holding labs in both in-person and online formats through Zoom, providing flexibility and ensuring access for everyone. Active participation and engagement are essential for achieving each week's learning milestones. This means attending all lectures and labs is mandatory. *Exception* to this principle applies if you face an unforeseen situation or other extraordinary circumstances that prevent you from continued participation. We are here for you and want to support you in ways we can, please do reach out to me if your circumstances require exceptions and additional support along the way. You are permitted *two absences* without hurting your participation grade, provided you email the Professor in advance to explain and request a leave of absence. Similarly, your lab grade will reflect both your attendance and the quality of your participation during lab sessions.

By coming prepared, actively participating, and taking ownership of the material in lectures and labs, you will better grasp the conceptual and practical aspects of statistical methods and data analysis. Your active participation and preparation also enhance the collective learning atmosphere, benefiting your own learning and your peers. The use of digital devices is restricted to class-related activities.

Assignments (30% of grade)

There will be five assignments in this course to refresh and challenge your understanding of each week's materials. Each is due the following **Thursday at 11:59pm EST** after the corresponding Tuesday lab session that week. You are encouraged to help each other figure out the answers to the problem sets, but it is expected that you write up your answers independently. Once the assignments are submitted on [GradeScope](#) (Entry Code: **YDXEDB**), students will be given a detailed answer key to **grade their assignments and make corrections and notes to explain each error made by the following Thursday**. The TA and I will re-check all self-graded assignments afterward. No matter the score it receives, each self-graded assignment (*with thorough corrections and reflection*) can earn full credit. An assignment not self-graded then will get zero points. In other words, careful work with self-reflection will be counted. The whole point of self-grading is to foster reflective and intentional learning—to give you a safe space to reflect on your learning, identify mistakes shortly after making them, and revise as needed. The process empowers you to learn effectively and take initiative in your learning. *Please note:* To qualify for self-grading credit, you must first submit a genuine attempt at the assignment by the original deadline.

Late Work Policy

It is imperative that you manage your workload properly for this course. *We will allow late assignments and self-grading up to 3 days late, with a 10% penalty per day.* Note that lateness will be determined by timestamp on Gradescope submissions, i.e. after 11:59 PM is considered late. If you foresee that you cannot finish the class assignments on time, a negotiated change of due date may be possible if you notify me at least 48 hours before the deadline. If there are last-minute emergencies (loss in family, hospitalizations), please let me know as soon as possible. I expect you will make every effort to promptly complete work for assignments and communicate consistently with group members.

Quizzes (30% of grade)

Additionally, there will be two quizzes – covering the first half and second halves of the course, respectively. **You are expected to work on quizzes without the assistance of any other person.** Students will receive a detailed answer key to self-grade the quizzes (Quiz 2), and the instructor will re-check and adjust the grades if needed. While the quizzes will not be adjusted for self-reflection and correction, they will count toward bonus points.

Research Article Review

~~This assignment aims to deepen your understanding of research methodologies, statistical analysis, and data interpretation for empirical research articles. After carefully reviewing the assigned article, your task is to provide a comprehensive summary, thorough analysis, and critical evaluation of the article's approach and findings.~~

Policy Research Report (Team project) (30% of grade)

The final project is designed so that each of you will bring a problem of personal interest to the class. You will need to identify a problem to tackle with a data set that either you collect/extract or find. This project team involves a group of up to 4 members. A complete write up is required. This would be a good project to put in your CV if desired. A monthly check-in with the professor is highly recommended to meet the monthly goal.

- A well-motivated, relevant topic is most desirable.
- Originality, complexity, and challenge will be another plus.
- A complete, professional write-up is a must.
- Maximum of 15 double-spaced pages in APA 7th format.
- Potential working datasets will be provided.
- [UMass Library](#), [IPUMS](#), [Google Dataset](#), [Kaggle](#) are some good places to find a data set.

The final project will be accompanied by a rubric describing the allocation of points and criteria for evaluation. Each group will have a chance to review and provide feedback on the presentation and report from another group. The instructor will grade the final project. The final grade of the course will be adjusted for peer evaluation from your teammates.

How to be Successful in the Course?

In general, any submitted work written work (quizzes and final project) is assessed on these evaluative criteria:

- **Thoroughness** – each question is addressed and well-explained
- **Rigorous analysis** – analytically rigorous in tackling the assignment
- **Professional presentation** – explain codes with notes, accurate interpretation of results
- **Problem-solving acumen** – show accumulated learning to approach statistical problems

Calculation of Grades

Student performance in the course will be graded based on the following scale:

A 93-100%	B- 80-81%
A- 90-92%	C+ 78-79%
B+ 87-89%	C 70-77%
B 83-86%	F 69% or below

Assignments and Quizzes Timeline

Schedule	Submission Due at 23:59	Self-grading Due at 23:59
Lab 2 HW	Thu 2/15	Thu 2/22
Lab 3 HW	Thu 2/22	Thu 2/29
Quiz 1 (MC)	Fri 3/1 (During Lecture)	---
Lab 6 HW	Thu 3/14	Thu 3/28
Lab 7 HW	Thu 4/4	Thu 4/11
Lab 8 HW	Thu 4/11	Thu 4/18
Quiz 2 (MC + R)		

Policy Research Project Timeline

February 2024	March 2024	April 2024
<ul style="list-style-type: none"> ▪ Research Question 	<ul style="list-style-type: none"> ▪ Data cleaning & transformation 	<ul style="list-style-type: none"> ▪ Interpretation of results
<ul style="list-style-type: none"> ▪ Variable measures 	<ul style="list-style-type: none"> ▪ Statistical models 	<ul style="list-style-type: none"> ▪ Report write-up
<ul style="list-style-type: none"> ▪ Select data set(s) 	<ul style="list-style-type: none"> ▪ Data analysis & visualization 	<ul style="list-style-type: none"> ▪ Class presentation

Class Schedule

Tentative and subject to change.

Week	Topics	Weekly Learning Goals
Week 1, Feb 2	Course Overview: Introduction to Statistics & R	<ul style="list-style-type: none"> ✓ Course expectation and plan for success! ✓ Recognize the different types of data and research questions
Lab 1, Feb 6	Create your first R file	
Week 2, Feb 9	Assumptions of statistical inference Random variables, probability distribution, central limit theorem	<ul style="list-style-type: none"> ✓ Assumptions of statistical inference (larger sample size!) ✓ Use R to graph and explain descriptive statistics of variables
Lab 2, Feb 13	Exploratory data analysis	
Week 3, Feb 16	Hypothesis testing, significance level, t-statistics	<ul style="list-style-type: none"> ✓ Write hypotheses and interpret statistical results ✓ Estimate the causal effects of experiments using differences of means
Lab 3, Feb 20	Compute t-test, ANOVA test, construct confidence interval for treatment effect	
Week 4, Feb 23	Review Lecture + Simple Linear Regression or Ordinary Least Square (OLS) regression	<ul style="list-style-type: none"> ✓ Understand Best Linear Unbiased Estimator (BLUE) assumptions ✓ Interpret and write regression results
Lab 4, Feb 27	Review Week 1 to Week 3	
Week 5, Mar 1	Quiz 1	
	No Lab: make up exam	
Week 6, Mar 8	Simple Linear Regression + Regression Diagnosis	<ul style="list-style-type: none"> ✓ Diagnose regression results ✓ Recognize violation of assumptions
Lab 6, Mar 12	Compute simple regression models Run diagnosis plots	
Week 7, Mar 15	Multiple Linear Regression, Omitted Variable Bias, Measures of fit Compute multiple regression models	<ul style="list-style-type: none"> ✓ Interpret regression results ✓ Diagnose model fitness and violation of assumptions
No Lab	----- Spring Break (3/17-24) -----	

Week 8, Mar 29	Logarithm and Variable Transformation, Binary variables	<ul style="list-style-type: none"> ✓ Improve model fitness ✓ Interpret regression results with transformed variable
Lab 7, Apr 2	Create and plot nonlinear terms	
Week 9, Apr 5	Review for Quiz 2	<ul style="list-style-type: none"> ✓ Study Week 4 to Week 8
Lab 8, Apr 9	Review Wks 4 - 8	
Week 10, Apr 12	Monday schedule – no class Take home Quiz 2	
Lab 9, Apr 16	Go over Quiz 2 Answer Key	
Week 11, Apr 19	Check in with Professor	
Lab 11, Apr 23	Lab Optional: Project Support	
Week 12, Apr 26	Project Preparation	
Lab 12, Apr 30	Lab Optional: Project Support	
Week 13, May 3	Project Presentation	
	Final project Report due on <u>May 13th Monday</u> → <u>May 17th Friday by 23:59</u>	

No Use of Generative AI Permitted except when Instructor said Otherwise

Intellectual integrity is vital to an academic community and for my fair evaluation of your work. All work completed and/or submitted in this course must be your own, working individually or in groups, completed in accordance with the University's Guidelines on Academic Integrity as stated above. **Except when the instructor said otherwise, students should not have another person/entity do the writing of any substantive portion of an assignment for them, which includes hiring a person or a company to write assignments and using artificial intelligence tools like ChatGPT and Google Bard for any assignments.**

No Social Loafing/Free Riding

Social loafing or free riding occurs when a group member relies on the remaining members to complete group tasks and does not contribute to group work. To discourage free-riding, individual group member's contributions to the group project will be assessed by confidential peer evaluations at the end of the semester. If the peer evaluations you receive clearly show that you are a free rider, your individual grade on the group project will be adjusted.

Writing Support

The Writing Center provides free support on any writing you do while at UMass, whether for a course or not. Trained tutors are available to work with you ([currently online only](#)) as you plan, draft, and revise your writing. Please visit www.umass.edu/writingcenter for hours and information about making appointments. If you have questions about the UMass Writing Center, please email writingcenter@acad.umass.edu or call 413-577-1293.

Academic Honesty Statement

Since the integrity of the academic enterprise of any institution of higher education requires honesty in scholarship and research, academic honesty is required of all students at the University of Massachusetts Amherst. Academic dishonesty is prohibited in all programs of the University. Academic dishonesty includes but is not limited to: cheating, fabrication, plagiarism, and facilitating dishonesty. Appropriate sanctions may be imposed on any student who has committed an act of academic dishonesty. Instructors should take reasonable steps to address academic misconduct. Any person who has reason to believe that a student has committed academic dishonesty should bring such information to the attention of the appropriate course instructor as soon as possible. Instances of academic dishonesty not related to a specific course should be brought to the attention of the appropriate department Head or Chair. Since students are expected to be familiar with this policy and the commonly accepted standards of academic integrity, ignorance of such standards is not normally sufficient evidence of lack of intent (http://www.umass.edu/dean_students/codeofconduct/acadhonesty/)

Accommodation Statement

The University of Massachusetts Amherst is committed to providing an equal educational opportunity for all students. If you have a documented physical, psychological, or learning disability on file with Disability Services (DS), you may be eligible for reasonable academic accommodations to help you succeed in this course. If you have a documented disability that requires an accommodation, please notify me within the first two weeks of the semester so that we may make appropriate arrangements. For further information, please visit Disability Services (<https://www.umass.edu/disability/>)

Title IX Statement

In accordance with Title IX of the Education Amendments of 1972 that prohibits gender-based discrimination in educational settings that receive federal funds, the University of Massachusetts Amherst is committed to providing a safe learning environment for all students, free from all forms of discrimination, including sexual assault, sexual harassment, domestic violence, dating violence, stalking, and retaliation. This includes interactions in person or online through digital platforms and social media. Title IX also protects against discrimination on the basis of pregnancy, childbirth, false pregnancy, miscarriage, abortion, or related conditions, including recovery. There are resources here on campus to support you. A summary of the available Title IX resources (confidential and non-confidential) can be found at the following link: <https://www.umass.edu/titleix/resources>. You do not need to make a formal report to access them. If you need immediate support, you are not alone. Free and confidential support is available 24 hours a day / 7 days a week / 365 days a year at the SASA Hotline 413-545-0800.