Syllabus
PUBM 5910: Advanced Data Analysis
&
Management

Instructor: Professor Chengxin Xu

Seattle University

Summer 2025

Course Information:
Class Hours: Tuesday 6:00-9:00pm
Classroom: Pigott Building 205

Contact Me:
E-mail: cxu1@seattleu.edu
Website: https://cxuchengxin.github.io/website/
Office: Casey 210-09
Office Hours: By appointment via Bookings

Contents

1 Course Description

Public agencies are increasingly interested in unlocking the potential of large-scale data to improve service delivery and inform policy efforts. Computational tools capable of making productive use of big data have proliferated in recent years, drastically decreasing the barriers to entry for interested parties. This 8-week course will be divided into two following sections:

1.1 Section One: Data Analysis with Regressions

The first part of this course will explore the practice of using data to improve organizational performance, including techniques for data collection, analysis, and behavior change. Regression serves as the foundation for modern quantitative program evaluation techniques, so the first part of this course establishes a baseline set of skills that can be extended in your future self-studies.

Regression analysis is not a single thing – it is a powerful set of tools used to examine relationships in data and test hypotheses concerning the significance of these relationships. Regression can be used to analyze observational data, in which case it can be used to identify important correlations that help us predict when events will occur together.

In the program evaluation context we are specifically interested in causal analysis, which allows

us to determine whether a management practice, a nonprofit or government program, or a specific public policy has a positive impact. When certain conditions are met we can use regression analysis to estimate the impact of a policy or program. This first part of the course will build a foundational understanding of the mechanics of multiple regression, the proper interpretation of regression models, and the circumstances in which we can use regression results for causal reasoning.

1.2 Section Two: Data Management

The second part of the course focuses on effective data management for public agencies. Students will operate as their own laboratory through a data journaling exercise, and devise strategies for incorporating data into management practices of public and nonprofit organizations.

One of the key course take-aways is that data can make your organization more effective, but data itself is not sufficient without strong management frameworks. The course is built around a quantified-self experiment where you will use life-logging tools and a journaling system to learn about managing information overload and building goal-oriented, evidence-based routines.

Students will also explore the social, political, and ethical considerations associated with building and managing data analytics programs in the public sphere. Students will engage critical dilemmas of data privacy, data protection, predictive analytics, personalized service delivery and resource provision, algorithmic regulation, and large-scale data analytics for administrative efficiencies and resource management optimization, among others.

Students will read several case studies that explore the use of data in organizations, will engage in discussions about socio-economic policy considerations, and write policy guidance frameworks on best practices in evidence-based management, open data, and privacy.

2 Student Learning Outcomes

On successful completion of this course (i.e. by passing this course), you will be able to:

1. Run multiple regression model and interpret the results, which include:
    - Understanding how adding or removing control variables from a model will change the accuracy (bias) and precision (efficiency) of the model.
    - Understanding the causes and consequences of omitted variable bias, it's impact on causal analysis, and the circumstances in which we can interpret regression results as correlational or causal.
    - The ability to translate regression results to an articulation of program impact using effect sizes and confidence intervals to offer responsible evidence-based recommendations to managers and policy-makers.
    - Diagnosis of common specification problems in regression analysis such as outliers, nonlinearity, multi-collinearity, and measurement error.

2. Describe how public agencies harness large-scale data to inform policy design, increase stakeholder engagement, and improve service delivery.
3. Recognize situations where it's possible to collect data to inform organizational processes.
4. Intelligently consider the social, political, and ethical considerations of using data analytics.

3 Course Prerequisites:

It is recommended to pass **PUBM 5440 Applied Data Analysis** prior to taking this course. To be successful in this course, you need some basic background in statistics including familiarity with the following topics:

- Variable types – numeric, categorical, binary
- Descriptive statistics such as the mean, median, mode, quantiles, and standard deviation
- Scatterplots, bar charts, and histograms
- Basic probability and sampling
- Confidence intervals
- Hypothesis testing

4 Books and Materials

4.1 Required Reading and Materials

Lewis-Beck, Michael S., (2016) *Applied Regression: An Introduction*. SAGE. (Access to E-book via Lemieux Library)

Llaudet, Elena and Kosuke Imai, (2023) *Data Analysis for Social Science: A Friendly and Practical Introduction*. Princeton University Press.

O'Neil, Cathy. (2016) *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown.

Eagel, Nathan and Kate Greene, (2014) *Reality Mining: Using Big Data to Engineer a Better World*. The MIT Press.

4.2 Required Electronic Devices

Your learning and practice in this class will heavily rely on the use of PC/Mac. Please bring your own device to the class so that you can practice during the lab session.

Any laptop with internet browsers (e.g., Chrome, Safari, Edge, etc.) will work. No software is required to be installed in your device.

You may also borrow a laptop from the Lemieux Library at Seattle University. You may find the instruction here: https://library.seattleu.edu/collections/technology.

5 Course Evaluation

Students are expected to keep up with assigned readings and to participate actively in class discussions. There will be four projects and a final exam. Grades will be based on the following:

Assignment	Relevant Learning Outcomes	Relative Value	Due Date
Problem Set 1	Outcome 1.1	10%	See Canvas
Problem Set 2	Outcome 1.2	10%	See Canvas
Problem Set 3	Outcome 1.3	10%	See Canvas
Problem Set 4	Outcome 1.4	10%	See Canvas
Discussion Lead 1	Outcome 2	10%	See Canvas
Discussion Lead 2	Outcome 3	10%	See Canvas
Discussion Lead 3	Outcome 4	10%	See Canvas
Take-home Final Exam	Outcomes 1–4	30%	See Canvas

Table 1: Overview of Graded Assignments

5.1 Formatting and Submitting Your Assignments

When necessary, APA (see a `guidance` here) is required in terms of writing style of your assignments. In-text citations are preferred to footnotes or endnotes. Writing clarity and quality are considered in grading. Please use a standard font such as Arial, Calibri, or Times New Roman. Pages should use 1-inch margins, double-spaced, with 12 point fonts. If you think you need help with your writing, talk to the professor. There are resources on campus to help you. Also feel free to seek for services of the `Writing Center at Seattle University`.

Due dates of all assignments are listed on the Course Schedule and your Canvas Calendar.

5.2 Grading Scale

Students' final grades will be calculated as follows:

A	94-100	A−	90-93	B+	87-89	B	83-86
B−	80-82	C+	77-79	C	73-76	C−	70-72
D+	67-69	D	63-66	D−	60-62	F	59 or less

Any students who want to challenge their course grade shall follow the Academic Grading Grievance Policy at Seattle University. The detail of the policy can be found here. An agreement to receive an Incomplete (I) grade may be negotiated if your circumstances do not allow you to finish the course on time. The Incomplete Grades Policy of the university is available on the Office of the Registrar website.

5.3 Assignment Deadlines and Extensions

The assignments in this course build on one another, so the feedback process is an important part of your learning. My plan is therefore to return your assignments as soon as possible after you have submitted them. (My normal policy is to return papers within five working days; I hope to be able to keep to that this quarter.) Faster feedback means you're more likely to remember what you wrote and can more easily take the feedback on board for the next assignment.

It would greatly help the learning process if you are able to submit your assignment by the stated deadlines. If you find that you are unable to complete an assignment on time, please communicate with me in advance. I do understand that life happens and can sometimes get in the way. I'll do my best to be flexible; for that to happen, we need to keep in dialogue so that I know where you are in your learning and how I can best support you.

To help keep yourself on track, I strongly encourage you to plan ahead: Check the session-by-session schedule on Canvas to see when assignments are due. This enables you to block out time in your calendar now so that you know when you will be working on assignments for this course. If you're wondering how much time you need to allow for completing assignments well, then I suggest you use the approach I take for my own work: make a rough estimate of time and then multiply it by 2.5. This sounds like a lot, but I think you'll find it more realistic for managing your expectations and planning.

An agreement to receive an Incomplete (I) grade may be negotiated if your circumstances do not allow you to finish the course on time. The Incomplete Grades Policy of the university is available on the Office of the Registrar web site: https://www.seattleu.edu/redhawk-axis/academic-policies/.

6 Academic Integrity Policy

Students must follow the school's integrity policy. Please be aware that plagiarism and any forms of cheating will never be excused under any circumstances. Violation of these policies leads to immediate failure of the course. If you are not familiar or clear about the school's academic integrity policy, please refer to the tutorial here.

6.1 The Use of Generative Artificial Intelligence (AI)

The use of the work or intellectual property of other persons or the outputs of Generative Artificial Intelligence (AI) programs (e.g., ChatGPT, Gemini, Grok, DeepSeek, etc.) presented as one's own work without appropriate citation or acknowledgment is now considered as plagiarism by the university policy.

It is not recommended to use Generative AI for coding. You should code by yourself in the first place. Other than that, the use of AI is encouraged in this class, but please notice at the end of the assignment how AI has helped you to finish your work, including exact prompts you use to get the results. Appropriate use of AI includes brainstorming, literature search and review, debugging, and language polishing.

7 Disability Services

Seattle University is committed to providing students with disabilities an equal opportunity to participate in its classes and activities. This goal is fundamental to its belief in social justice. To achieve this goal, faculty and staff work together to remove barriers in physical access, communication, attitudes, and policies that can prevent students from taking part in programs. In addition, the University provides supplemental support services through the Disability Services Office.

Students who are requesting academic, housing, and meal plan accommodations for the first time are required to complete the new student application via the myDS New Student Application link and meet with `Disability Services` **before receiving services**. During the intake meeting, the student will discuss their goals at Seattle University, the assistance they may need in reaching them, and the creation of a personalized accommodations plan. Students must also provide documentation from a qualified diagnostician that:

- Establishes the nature of their disability.
- Provides evaluative data with test names and scores supporting their current need for specific accommodations.
- May list services recommended for the student.

Documentation is confidential and will not be released or discussed with the student's permission. For more information, please see the Documentation page and Requirements for Documentation of Disabilities.

If the student needs, or expects to need, an accommodation not supported by their documentation, the student should discuss options for additional evaluation with DS Staff.

For more information regarding disability services, please see `here`.

8 Statement of Diversity

Seattle University celebrates educational excellence achieved through diversity. In the meanwhile, diversity is a central topic in the field of public and nonprofit administration. In this class, we will frequently face opportunities and challenges of diversity, equity, and inclusion. We will have discussion of the normative and pragmatic value of diversity and learn how to embrace it.

Students are encouraged to honor the uniqueness of themselves and peers, and to appreciate the opportunities we have to learn from each other. In particular, students are welcomed to share their personal experience regarding diversity policies and management. Students have to demonstrate respect to others' opinions and refrain from personal attacks, hatred language, and personal judgments.

9 Policy on Religious Accommodations

Consistent with Washington State law (RCW 28.10.039), it is the policy of Seattle University to reasonably accommodate students who, due to the observance of religious holidays, expect to be absent or endure a significant hardship during certain days of their academic course or program.

A student seeking accommodations under this policy must provide a written request to the student's faculty member: (i) within fourteen (14) calendar days from the first day of classes of the academic quarter in which the student is requesting accommodations; and (ii) identifying the specific dates the student requests accommodations.

For more details about the policy on religious accommodations for students, please see here.

Please see the course schedule at the next page.

10 Course Schedule

10.1 Regression for Data Analysis

Week 01, 06/23 - 06/27: Fundamentals of Regressions
Regression Mechanics: Variance, Slopes and Residuals;
Hypothesis Testing: Standard Errors, Confidence Intervals, and p-values

Week 02, 06/30 - 07/04: Multivariate Regressions
Control Variables & Partitioned Regressions; Omitted Variable Bias

Week 03, 07/07 - 07/11: Dummy Variables
Dummy variables and nonlinear regressions

Week 04, 07/14 - 07/18: Regressions for Causal Identification
Randomized Experiment; Two-way Fixed Effect; Differences-in-Differences

10.2 Data Management for Public Services

Week 05, 07/21 - 07/25: The Value of Data
Readings:

- Eagle, N., & Greene, K. (2014). Reality mining: Using big data to engineer a better world. MIT Press. CH7 mobile and internet data
- Lohr, S., (2012). The Age of Big Data

Week 06, 07/28 - 08/01: Challenges of Big Data

Week 07, 08/04 - 08/08: Managing with Data

Week 08, 08/11 - 08/15: Data-Driven Human Resources