Quantitative Research Methods

Sociology/Criminology 3040-001

John McLevey, Memorial University

Class	Tuesday and Thursdays, 1:30-2:50 pm, in **CP 2003**
Instructor	Professor John McLevey (he/him)
Email	`mclevey@mun.ca`
Office	AA 4055
Office hours	3:00–4:00 PM Tuesdays and Thursdays or by appointment

Note on expectations for email communication: I will reply to course related e-mail within 24 hours between Monday and Friday. I do not check or respond to e-mails on weekends or holidays. Please use your `@mun.ca` email account for all course-related correspondence.

COURSE DESCRIPTION

SOCI 3040, Quantitative Research Methods, introduces students to foundational concepts, principles, and practices in contemporary quantitative social science. Core topics include research computing, data collection and processing, descriptive and inferential statistics, measurement, visualization, and statistical modelling. Research ethics, transparency and reproducibility, and effective communication are emphasized throughout. This course follows the Faculty of Humanities and Social Sciences' **Quantitative Reasoning Course Guidelines** (see www.mun.ca/hss/qr).

Extended Course Description

SOCI 3040, Quantitative Research Methods, introduces foundational concepts, principles, and practices in contemporary quantitative social science through a carefully scaffolded sequence of three thematic modules that are designed to develop your research capabilities progressively over the course of the semester. It follows the Faculty of Humanities and Social Sciences' Quantitative Reasoning Course Guidelines (see www.mun.ca/hss/qr).

Module 1: Collecting, Processing, and Describing Data establishes foundations in research computing using Python, data collection (web-based and surveys), and descriptive statistics. Students learn to gather, clean, and analyze real social science data while developing programming skills that are essential for contemporary quantitative social science.

Module 2: Measurement and Latent Constructs explores measurement theory, reliability and validity assessment, and techniques for identifying underlying constructs in complex social phenomena. Students work with survey data, text analysis, and formal measurement models to understand how abstract concepts are quantified in social research.

Module 3: Uncertainty, Inference, and Statistical Modelling introduces statistical inference from a "generative" perspective, emphasizing uncertainty quantification and principled model interpretation. Students learn to build and validate statistical models while connecting results to substantive social science questions.

The course culminates in a **quantitative research portfolio** that integrates your work across all three core modules, demonstrating how contemporary quantitative methods address substantive questions in sociology, criminology, and related social sciences. Throughout all modules, research ethics, transparency and reproducibility, and effective communication of quantitative findings are emphasized.

LEARNING OBJECTIVES

By the end of this course, students will be able to:

- Compute, clean, and manage social data reproducibly using Python and Colab, producing documented notebooks that others can run.
- Collect data ethically (e.g., web, APIs, surveys), process it using reasonable, transparent, and reproducible methods.
- Describe and visualize data using appropriate summaries and graphics, interpreting patterns clearly and honestly.
- Apply measurement theory to evaluate reliability and validity, and to distinguish scales and indicators for key social constructs.
- Analyze association and structure (e.g., correlations) to motivate simple measurement models and understand where more complex methods are necessary.
- Reason probabilistically about uncertainty, applying a generative lens to interpret estimates and credible intervals.
- Specify, fit, and interpret core models (correlation, multiple linear regression, logistic regression) and connect results to substantive questions.
- Diagnose and validate models with appropriate checks and transparent reporting of limitations.
- Communicate quantitative findings, in writing and presentations, to technical and non-technical audiences with clarity, transparency, and intellectual humility.
- Practice transparency, reproducibility, and responsible GenAI use: document decisions, attribute tools, and verify outputs.

REQUIRED COURSE MATERIALS

I am developing an open access textbook (see below) specifically for this course. It integrates seamlessly with course activities and provides all necessary background reading, programming tutorials, and conceptual frameworks for successful completion of the course. The book is freely available online and will be updated throughout the semester based on student feedback and learning needs. A link to the text is provided on the course Brightspace page.

1. **Software**: This course is taught using **Python** and **Google Colab** ("Colab"), both of which are free to use. You do not need to install any software on your computer, but you will need a free Google account to use Colab. If you do not already have a Google account, you can create one for free at `https://accounts.google.com/signup`.
2. **Textbook (Open Access)**: McLevey, John. *Quantitative Social Science: Core Concepts, Skills, and Stories*. Open access, available online. A link to the text is provided on the course Brightspace page.

COURSE DESIGN AND ORGANIZATION

Sociology/Criminology 3040-001 is a **project-based course** organized into sequential modules designed to build quantitative research skills progressively throughout the semester. Each module focuses on a specific aspect of quantitative research methods, with carefully scaffolded assignments that deepen understanding and practical capabilities.

Module Overview

Module	Topic	Weeks	Weight
1	Collecting, Processing, and Describing Data	2-4	25%
2	Measurement and Latent Constructs	5-7	25%
3	Uncertainty, Inference, and Statistical Modelling	8-10	25%
4	Final Research Portfolio	11-12	25%

Each module spans three weeks and includes three interactive notebooks that combine reading, reflection, and hands-on data analysis. Each module culminates in an assignment that synthesizes and reinforces the skills and knowledge that you are actively acquiring.

The module descriptions below provide an overview of the key topics, skills, and assignments for each module. Note that I may make some minor adjustments to the schedule or content based on class progress and feedback, but any such changes will be communicated clearly in advance and will not affect the overall structure or assessment of the course. Note that details on grading/assessment are available in the Assignments and Assessment section below.

Module 1: Collecting, Processing, and Describing Data

This foundational module introduces essential skills for gathering, cleaning, and analyzing data in social science research by integrating Python programming fundamentals with core data science concepts through real research applications. Rather than learning programming in isolation, students develop technical capabilities while addressing authentic research questions, beginning with Python basics (variables, data types, functions) through practical data collection tasks and progressing to web scraping Canadian election polling data from Wikipedia. Students learn string processing, data structures (lists, dictionaries), and HTML analysis while simultaneously exploring survey methodology concepts through polling analysis.

The module culminates with descriptive statistics and data visualization, where students master measures of central tendency and variability, exploratory data analysis techniques, and data visualization for pattern recognition. Throughout this progression, students develop both computing skills (loops, conditionals, pandas Series and DataFrames) and research skills (literature search and synthesis, web APIs, web scraping, survey methodology foundations), while learning advanced data cleaning techniques for handling missing data, outliers, and transformations alongside survey methodology foundations and polling aggregation methods.

Module 2: Measurement and Latent Constructs

This module explores fundamental measurement principles and introduces techniques for identifying and analyzing underlying constructs that cannot be directly observed, emphasizing the

theoretical and practical challenges of measuring abstract social phenomena. Students begin with measurement theory foundations, including measurement scales (nominal, ordinal, interval, ratio) and reliability and validity assessment principles, applying these concepts through a case study of measuring sentiment and emotion in text data. The module progresses to examining relationships between variables through correlation analysis and introduces cultural schemas and belief systems in survey data, helping students understand population heterogeneity in political attitudes.

The module concludes with formal measurement models, specifically Classical Test Theory (CTT) and Item Response Theory (IRT), where students learn to specify, fit, and validate measurement models using V-DEM expert survey data to measure democracy and autocracy. Throughout this sequence, students develop advanced pandas operations, correlation analysis, and basic statistical modeling while building research skills in measurement theory application, reliability and validity assessment, latent construct identification, and comparative measurement approaches across text, survey, and expert data sources.

Module 3: Uncertainty, Inference, and Statistical Modelling

This module introduces statistical inference from a contemporary generative perspective, emphasizing probabilistic thinking and uncertainty quantification as students move beyond descriptive analysis to develop inferential capabilities. Students begin with a gentle introduction to Bayesian reasoning and probabilistic modeling, learning about prior and posterior distributions and uncertainty quantification while fitting and interpreting univariate statistical models. The module progresses to bivariate analysis, covering correlation and simple linear regression from a probabilistic perspective, revisiting cultural schemas and belief systems with statistical models that account for population heterogeneity in political attitudes.

The module culminates with multiple regression and applied inference, where students learn multiple linear regression and logistic regression, including model specification, fitting, interpretation, and diagnostic checking. Through applications modeling public opinion dynamics using polling data, students confront key challenges in applied statistics including regression interpretation, latent constructs, and generalization while developing computing skills in statistical modeling, regression analysis, and model diagnostics alongside research skills in statistical inference principles, uncertainty quantification, model validation, and connecting statistical results to substantive social science conclusions.

Final Research Portfolio

The final research portfolio synthesizes learning from all previous modules into a coherent research demonstration that pulls together work completed throughout the semester, emphasizing integration, reflection, and communication of developing quantitative research skills rather than requiring extensive new analysis. Students connect their work from Modules 1-3 to address a substantive research question of their choice, demonstrating how different methodological approaches contribute to understanding their chosen topic while analyzing their learning progression and explaining how quantitative methods enhance their research capabilities.

The portfolio includes both a 10-minute professional presentation covering the student's substantive research question, data sources and methods used, key findings, and personal reflection on skill development, as well as a brief written report summarizing their work with emphasis on methodological choices and implications. This culminating experience allows students to demonstrate their ability to integrate technical skills with research applications while reflecting

on the strengths, limitations, and future applications of quantitative methods in their disciplinary context.

ASSIGNMENTS AND ASSESSMENT

This course uses a portfolio-based assessment system designed to reward consistent effort and skill development throughout the semester. All assignments build cumulatively, allowing you to develop and demonstrate progressively sophisticated research capabilities.

Component	Weight	Due Date
Module 1 Assignment	25%	October 2
Module 2 Assignment	25%	October 28
Module 3 Assignment	25%	November 20
Final Research Portfolio	25%	December 2-4

Grading Philosophy

This course rewards consistent effort and persistent engagement with challenging material. The assignments are designed to recognize that learning quantitative research methods effectively requires steady work over time. For better or worse, you cannot compress this learning into last-minute efforts.

A detailed rubric for all assignments will be provided in the second week of the semester, but in short, my grading philosophy in this class is based on the following principles:

- The content is challenging but accessible with dedicated time and effort.
- Confusion may be unpleasant and frustrating, but it is a normal byproduct of engaging with challenging content. Mistakes are part of the learning process.
- Steady effort and consistent work throughout the semester should be fairly rewarded.
- There should be no assessment-related surprises. If you complete the work thoughtfully, you will earn the points. The grading criteria are simple, transparent, and apply consistently across all assignments.

Assignment Details

Module Assignments (75% total)

Each module concludes with a mini-portfolio assignment that synthesizes the skills and concepts from that three-week period. These mini-portfolio assignments include:

- Completed notebook exercises with code, analysis, and interpretation
- Written reflection connecting technical skills to research applications
- Documentation of your problem-solving process and learning progression
- Application of module concepts to data relevant to your research interests

They are developed *incrementally* throughout each module, with clear guidelines and support provided in class. All assignments will require additional work outside of class time, but the bulk of the work will be done during class sessions.

Final Research Portfolio (25%)

The culminating assignment integrates learning from all three modules:

- Connect work from Modules 1-3 to address a substantive research question. Present your research question, methods, findings, and reflection to the class.
- Produce a brief written report summarizing your methodological journey and key findings, as well as a critical reflection and analysis of your learning progression and future research applications.

I will provide assessment rubrics for each assignment.

Late Assignment Policy

Module assignments must be submitted by the specified deadlines to maintain the cumulative learning structure. Unless prior arrangements are made for documented emergencies or accommodations, late submissions will receive reduced credit of 5% a day up to one week after the deadline, after which assignments will not be accepted. For the final portfolio, extensions may be available with advance notice, but presentation dates cannot be rescheduled due to the collaborative nature of the final presentations.

Exception to the late assignment policy: I will grant everyone **2 grace days** for late assignments that can be used at any time during the semester without penalty. You do not need to explain or justify using these grace days, just submit your assignment within 2 days of the deadline and you will receive full credit. Use them wisely!

Submission Guidelines

- Download your completed assignment from Colab in **both** `.ipynb` and `.py` format. Do not alter them after downloading, they are already in the format I need to grade them.
- Submit both files on the course Brightspace page.

COURSE SCHEDULE

As noted above, the assigned readings for this course are in the open-access textbook. New content will be released at least one week before it is needed in class, beginning in Week 2.

Note on Week 1: We'll use our first meetings to review the syllabus, expectations, and overall course design. On Thursday we'll focus on getting set up and familiarizing ourselves with Google Colab. There is no textbook reading for Week 1. Instead, please read Chapter 1 from Rohan Alexander's *Telling Stories with Data* (available at `https://tellingstorieswithdata.com/01-introduction.html`). We will draw on this data-storytelling framework throughout the course.

Week	Date	Topic	Module
1	**Tue Sep 9**	Course Introduction and Setup	Intro
1	**Thu Sep 11**	Research Computing Foundations	Intro
2	**Tue Sep 16**	Python Fundamentals & Data Collection	1
2	**Thu Sep 18**	Working with Data in Python	1
3	**Tue Sep 23**	Web Scraping: Canadian Election Polls	1
3	**Thu Sep 25**	Data Structures and Survey Methods	1
4	**Tue Sep 30**[1]	National Day for Truth and Reconciliation	—
4	**Thu Oct 2**	Descriptive Statistics + **Module 1 Assignment Due**	1
5	**Tue Oct 7**	Measurement Theory Foundations	2
5	**Thu Oct 9**	Measuring Sentiment and Emotion	2
6	**Tue Oct 14**[2]	Fall semester break	—
6	**Thu Oct 16**	Correlation and Cultural Schemas (meets as Tuesday)	2
7	**Tue Oct 21**	Association and Belief Networks	2
7	**Thu Oct 23**	Formal Measurement Models (CTT/IRT)	2
8	**Tue Oct 28**	Measurement + **Module 2 Assignment Due**	2
8	**Thu Oct 30**	Probability and Statistical Thinking	3
9	**Tue Nov 4**	Uncertainty and Statistical Models	3
9	**Thu Nov 6**	Bivariate Analysis and Inference	3
10	**Tue Nov 11**[3]	Remembrance Day (university closed)	—
10	**Wed Nov 12**[4]	Class meets on Wednesday at your usual Tuesday time	3
10	**Thu Nov 13**	Cultural Schemas Revisited	3
11	**Tue Nov 18**	Multiple Regression and Applications	3
11	**Thu Nov 20**	Regression + **Module 3 Assignment Due**	3
12	**Tue Nov 25**	Portfolio Development Workshop	4
12	**Thu Nov 27**	Research Integration and Reflection	4
13	**Tue Dec 2**	**Portfolio Presentations**	4
13	**Thu Dec 4**	**Portfolio Presentations**	4

1. There will be no class on Tuesday September 30th as it is National Day for Truth and Reconciliation.
2. There will be no class on Tuesday October 14th as it is the Fall Semester Break / Thanksgiving.
3. There will be no class on Tuesday November 11th as it is Remembrance Day.
4. Class will meet on Wednesday November 12th as if it were Tuesday to make up for the Remembrance Day holiday.

COURSE POLICIES

Attendance and Participation

Regular attendance is essential for success in this hands-on course. Each class session builds directly on previous work, and collaborative learning activities cannot be replicated outside class. If you must miss a session:

- Contact a classmate to understand what was covered
- Review session materials and complete any missed exercises
- Attend office hours if you need clarification on missed content

More than two unexcused absences may result in difficulty completing assignments successfully.

Technology and Device Policy

Computers are required in class for hands-on programming sessions. Please use devices responsibly and avoid non-course related activities during class.

Academic Integrity

Students are expected to adhere to the principles of academic integrity as outlined in the University's Academic Integrity Policy. Academic misconduct includes, but is not limited to, plagiarism, cheating, misrepresentation, and unauthorized collaboration. All violations will be reported to the appropriate authorities and may result in failure of the assignment, course, or more severe sanctions. Please consult www.mun.ca/student/supports-and-services/student-conduct-and-academic-integrity/

All work submitted must be your own, though collaboration and discussion are encouraged during class sessions and office hours. When working with code and data analysis:

- You may discuss approaches and problem-solving strategies with classmates
- You must write and submit your own code and analysis
- Properly cite any external resources, tutorials, or assistance you receive
- Never copy code directly from classmates or online sources without attribution

Course Policy on Generative Artificial Intelligence (GenAI)

The use of generative artificial intelligence (GenAI) tools in this course is **permitted with the following guidelines**:

- **Attribution Required:** Clearly document any use of GenAI tools. Example: "I used ChatGPT to help debug this function. The original error was X, and the AI suggested Y modification, which I then tested and verified." I will provide templates for documenting and disclosing your use of GenAI tools in Week 2.
- **Understanding Required:** Use AI tools to help you understand concepts and debug problems, not to generate complete solutions you don't understand. If AI provides code, you must be able to explain how it works and modify it if needed.

- **Learning Progression:** In early modules focused on learning fundamentals, prioritize writing code yourself with minimal AI assistance. In later modules, AI tools can be more helpful for complex tasks once you've demonstrated basic competency.
- **Responsibility:** AI-generated code may contain errors or be inefficient. You are responsible for testing, understanding, and validating any AI-generated content before submission.

Note that violations of academic integrity could result in automatic failures of the assignment or the course, depending on severity. All cases will be reported.

Class Cancellations

Changes to the schedule due to weather, illness, or other circumstances will be communicated via email to your `@mun.ca` accounts. Check email before traveling to campus during severe weather. Makeup sessions or alternative arrangements will be provided as needed to ensure course learning objectives are met.

Accessibility and Accommodations

Memorial University is committed to supporting students with disabilities. Students who may need accommodations are encouraged to contact the Glenn Roy Blundon Centre for Students with Disabilities for a confidential consultation as early as possible in the semester.

Contact Information

- **Location:** University Centre, UC-5000
- **Phone:** 709-864-8000
- **Email:** blundon@mun.ca
- **Website:** www.mun.ca/blundon/

I am committed to ensuring all students can succeed in this course and will work with you to implement approved accommodations effectively.

Respectful Learning Environment

Memorial University is committed to creating a respectful, inclusive learning environment free from harassment and discrimination. All students, faculty, and staff have the right to learn and work in an environment that promotes dignity and respect. This is a shared responsibility.

RESOURCES FOR STUDENTS

Student Well-being and Support

Student Wellness and Counselling Centre

- Location: University Centre, UC-5400
- Phone: 709-864-8500
- Services: counselling, crisis support, mental health resources
- Website: www.mun.ca/student-wellness-and-counselling-centre/

Student Success Centre

- Academic coaching and study skills support
- Time management and stress reduction workshops
- Location: University Centre, UC-5000
- Website: www.mun.ca/student/supports-and-services/student-success-centre/

Sexual Harassment and Misconduct Reporting

- Confidential reporting and support services available
- Anonymous reporting options for sexual harassment and assault
- Website: www.mun.ca/sexualharassment/

Financial Support

- Student Financial Aid: www.mun.ca/student/money-matters/
- Emergency bursaries and assistance programs available

Research and Library Support

QEII Library

- Research consultation services
- Data and statistical support
- Access to academic databases and resources
- Website: www.library.mun.ca

Student Life (ASK)

- Location: University Centre, UC-3005
- Information about courses, housing, books, financial matters, and health services
- General student support and guidance

Additional Academic Support

If you are struggling with course material, personal issues, or need additional support, please reach out during office hours or via email. Early communication allows us to address challenges before they become overwhelming. The university provides extensive support services designed to help you succeed academically and personally.