# POLI 312: Intermediate Quantitative Methods

Department of Political Science
McGill University
Winter 2026

| | |
|---|---|
| **Instructor:** Mark Williamson | **TA:** Teresa Pian |
| **Email:** mark.williamson@mcgill.ca | **Email:** teresa.pian@mail.mcgill.ca |
| **Lecture**: MW 1:05 to 2:25pm | **Conferences**: W 3:35 to 4:25 pm |
| **Place**: Trottier 1100 | **Place**: Burnside 934 |

**Course Description:** POLI 312 is designed to continue the introduction to data driven quantitative political science students begin in POLI 311. Our main goal will be establishing a strong foundation in statistical theory and computational methods, which is critical if you are considering graduate school in a social science and/or to plan on a career in applied data science.

There are three components to the course. We start by reviewing the fundamentals of statistics, presenting a more mathematically rigorous view of the basic statistical concepts first introduced in POLI 311. Next, we will revisit regression modeling, adding important techniques like multiple regression and interaction effects to the basic regression toolkit. Finally, we apply what we've learned to newer methods used at the cutting edge of quantitative political science, such as models for causal inference and the analysis of textual and spatial data.

**Prerequisites:** POLI 311 is a prerequisite for this course. If you did NOT take POLI 311, you may only enroll if you provide evidence of equivalent past coursework. This course assumes a working knowledge of R and familiarity with basic statistics.

**Land Acknowledgment:** McGill University is on land which has long served as a site of meeting and exchange amongst Indigenous peoples, including the Haudenosaunee and Anishinabeg nations. We acknowledge and thank the diverse Indigenous people whose footsteps have marked this territory on which peoples of the world now gather.

Acknowledging the history of this land is a sign of respect toward its original peoples, but also a call to action for all those who benefit from the land today. I encourage you to learn more about the Indigenous history of Tiohtià:ke/Montreal and reflect on how you can use what you learn in this course to promote improved relations between Indigenous and non-Indigenous peoples.

## Course Organization

**Lectures**: We will meet every Monday and Wednesday from 1:05 to 2:35pm in Trottier 1100. Some of these sessions will be in-class coding workshops. Please bring laptop computers to class (fully charged!) for use in classroom activities.

**Conferences**: Throughout the term, there will be conference sections on Wednesdays from 3:35 to 4:25pm in Burnside 934. Conferences for this course will usually serve as

open office hours with the TA. Students are encouraged to get answers to any remaining questions about course material and discuss solutions to practice problems, quizzes, and assignments from coding workshops (see below). In some cases, conferences will be specifically set up to review certain material from workshops or in preparation for an assignment; see the schedule below for specific dates of these sessions.

**Class Discussion Board**: You will be invited to a discussion board on Slack that you can use to ask questions and discuss the course material. I encourage you to use this as first recourse for any questions you may have, rather than email. Please never post your code or an answer to specific homework questions on the discussion board. Keep questions general!

**Office Hours**: The instructor holds office hours to discuss course material and assignments on Mondays 2:40-5:00pm, either virtually via Zoom (link) or in-person at 3610 McTavish rm. 26-4. Please sign up for a meeting time within that time block using the link here. Office hours will be subject to change, but the times listed at that link will be kept updated.

## Materials

**Textbooks**: The main textbook for this course ("ROS") is:

> Gelman, Andrew, Jennifer Hill, and Aki Vehtari (2020). *Regression and Other Stories*. Cambridge University Press.

In general, each week of the course is based on a chapter from ROS. You will find it easier to follow the lectures if you complete the recommended reading before coming to class.

We will also use the textbook from POLI 311 ("QSS"), both as a reminder of what you learned last year and an introduction to some new concepts in later weeks:

> Imai, Kosuke and Lori D. Bougher (2021). *Quantitative Social Science*. Princeton University Press.

Both textbooks are available at the bookstore or in the library. Regression and Other Stories is also available as a free pdf online: https://users.aalto.fi/ ave/ROS.pdf

An additional resource for coding in R that is also available free online is:

> Grolemund, Garrett and Hadley Wickham (2016). R for Data Science. http://r4ds.had.co.nz/.

**Software**: Like POLI 311, POLI 312 is taught in the R statistical programming language. I also encourage the use of the RStudio integrated development environment (IDE), which makes learning R and compiling your assignments easier.

**Course Website**: There is a course website on MyCourses. All important documents (syllabus, supplementary readings, etc.) will be posted there. It is your responsibility to regularly check the page for updates. Course assignments will also be submitted there.

## Evaluation

There are four components to your grade:

- Workshops: 15%
- Quizzes: 25%
- Problem sets: 25%
- Research project:

    Descriptive analysis: 10%

    Regression analysis: 10%

    Final report: 15%

For all assignments, use a standard 12-point font, double-spacing, and 1 inch margins. To complete your work, you must use the math-friendly typesetting program R Markdown (.rmd) and compile your document into a .pdf, which is easily implemented through RStudio and can then be uploaded to MyCourses. Do not submit handwritten assignments or type equations directly into a document editor such as Microsoft Word; these submissions will be not be graded. Submissions should be a single document that includes all code used to generate results in-line, as well as a written explanation of the results, where appropriate. All assignments are due at 11:59pm ET on the dates listed in the schedule below.

**Workshops**: Over the course of the semester, there will be five coding workshops, which will take place during class time. Each workshop will occupy a full 80 minute class session. During these workshops, students will tackle data analysis assignments related to the workshop's theme in small groups. Students can use lecture notes and R package help files, or directly ask for help from the instructor or TA, but no other coding assistants or guides are allowed.

Grades will be based on active participation:

- Students who clearly engage in workshop activities will receive full points. Clear engagement includes attendance throughout the activity and active involvement in the assigned activity.
- Students who clearly do not engage in workshop activities or who are absent will receive zero points. Clear lack of engagement includes not participating in the assigned activity, doing something other than the assigned activity, or leaving the class instead of participating in the activity. Students with justified absences may come to the following office hours to answer questions about the material and potentially receive the missing points. A maximum of two justified absences will be considered.

Students are encouraged to take home any activities they do not have time to finish during the workshop and complete them on their own. Solutions will be discussed during conference the week after each workshop.

**Quizzes**: There will be three in-person, closed-book quizzes over the course of the term.

- Each quiz will contain short answer and multiple choice questions on the concepts covered in the corresponding unit.
- Quizzes will be based purely on concepts; no coding or extensive calculations required.
- Students will have 40-50 minutes to complete each quiz, after which we will go over answers together.

**Problem sets:** There will be 3 problem sets. Problem sets will consist of R coding exercises designed to help you practice applying methods and skills covered in workshops.

After you submit your practice problems, you will be given a solution set. You will have one week to resubmit a revised version of your practice problems, which includes the following:

- You must mark each of your answers as correct or incorrect. If you mark an answer as incorrect, you may provide a corrected answer. If your answer was correct, but you arrived at it in a different way, you should explain why and which approach you believe is preferable.
- Corrections/explanations should be submitted as annotations on the original submission. The annotations should be typed up below the original submission in a different font color.
- Each annotation must explain why the original answer was incorrect and how you corrected it. Reflections that use generic explanations (e.g., 'I misunderstood the question') or simply restate the solution without explanation will receive limited credit.
- To get credit on corrections, you must have made a good-faith attempt at completing the question on the original submission. If you do not meaningfully attempt a question on the original submission, it will not be eligible for corrections.

Practice problems will be graded out of ten according to the rubric below.

**Research project:** The final research project in this course will be based on independent quantitative data analysis. Students should select a dataset based on their own interests. Datasets can either come from publicly available statistical data, such as the surveys run by the Canadian Elections Study, World Values Survey, or European Social Survey, or use replication data from a published political science paper (for instance, from the Harvard Dataverse); but note that if using replication data you MUST go beyond the original paper, you can't just reproduce the same analysis. When selecting a dataset, the goal should be to use the dataset to test a theory proposing a relationship between explanatory variable(s) and an outcome variable. A formal rubric will be posted on MyCourses.

The grade for the research project will be based on a collection of data analyses conducted over the course of the semester.

- **Descriptive analysis** (10%): A memo of 2-4 pages that introduces the dataset(s) that you are going to use and the question you will ask. The memo should include 3-4 figures and tables that describe the distributions of key variables and/or the relationships between them.

Table 1: Problem set rubric

| **Task** | **Descriptions** | **Tips / What We Are Looking For** |
|---|---|---|
| Good-faith initial submission (3 points) | • **3 points:** All questions attempted; code runs or is clearly intended to run; evidence of independent reasoning, even if incorrect.<br>• **2 points:** Partial attempt; several missing or incomplete answers; limited engagement with the problems.<br>• **1 point:** Minimal attempt; many answers missing or nonsensical.<br>• **0 points:** No submission or essentially blank. | Correctness is *not* the focus here. We want to see genuine, independent effort and engagement with the problems. |
| Quality of corrections and reflection (4 points) | • **4 points:** Errors correctly identified; clear explanation of why the original answer was incorrect or could be improved, and how it was fixed; corrections are written in the student's own words.<br>• **3 points:** Corrections mostly correct; explanations present but somewhat vague or incomplete.<br>• **2 points:** Corrections made, but explanations are superficial or largely restate the solution.<br>• **1 point:** Minimal reflection; corrections mostly copy the solution without explanation.<br>• **0 points:** No meaningful corrections or reflection. | Corrections must explain reasoning, not just present the right code. Generic statements receive limited credit. |
| Correct solutions in initial submission (1 point) | • **1 point:** At least some nontrivial answers are correct.<br>• **0 points:** Most answers incorrect or missing. | This component rewards early understanding but is intentionally low-stakes. |
| Correctness and coherence of final submission (2 points) | • **2 points:** Final answers are largely correct; code, output, and explanations are consistent.<br>• **1 point:** Some remaining errors, but clear improvement over the initial submission.<br>• **0 points:** Little improvement or answers remain largely incorrect. | We are looking for improvement and internal consistency between code, results, and interpretation. |

- **Regression analysis** (10%): A memo of 2-4 pages that applies a regression model to the dataset in order to test a theory linking explanatory variable(s) to an outcome variable. The memo should report and interpret results from a regression model with at least two regressors.
- **Final report** (15%): A memo of 4-6 pages that adds advanced techniques to the initial regression analysis, which may include interaction effects, variable transformations, or any other technique covered in POLI 312 but not POLI 311. The final report should discuss the limitations of the analysis. Reports should focus on the research methods and do not need to have extensive citations or literature reviews.

## Course Policies

**Language of Submission**: In accord with McGill University's Charter of Students' Rights, students in this course have the right to submit in English or in French any written work that is to be graded. Conformément à la Charte des droits de l'étudiant de l'Université McGill, chaque étudiant a le droit de soumettre en français ou en anglais tout travail écrit devant être noté.

**Academic Integrity**: McGill University values academic integrity. Therefore, all students must understand the meaning and consequences of cheating, plagiarism and other academic offences under the Code of Student Conduct and Disciplinary Procedures" (see McGill's guide to academic honesty for more information).

**Late policy**: All assignments are due at 11:59pm on MyCourses on the dates specified in the schedule below. 5 percentage points per day will be deducted for all late assignments, including on weekends, for up to a total of 25 points (5 days). For the research project components, submissions will not be accepted after five days and will instead be assigned a grade of zero. In the case of problem sets, students may only submit up to one day late (with the usual 5 point penalty), because the solution sets must be posted the next day; submissions more than one day late for a problem set will receive a zero. If you face extraordinary circumstances and require an extension, please contact the instructor (appropriate documentation is required). Please note, however, that "K" grades (i.e. extensions beyond the term) will not be granted for this class.

**Collaboration**: Collaboration is part of learning how to code. I encourage you to collaborate! But you will not learn how to do statistical programming if you do not write your own code. Please feel free to work with colleagues, but **do not** copy each others' code verbatim. You must also write your **own** interpretations of the results.

**Generative AI**: AI can be a useful tool for learning statistical concepts and programming. Students may use AI to reinforce course material by, for example, asking for explanations of concepts or code seen in class. It can also help with debugging code after students have made a good faith effort to solve the problem on their own first. In this class, AI should not be used to write code or full responses to problem set questions or research project components. If students do use AI assistance, they must indicate where and how it was used in their work. See here for details on how to do that with ChatGPT. The instructor or TA may ask students to walk through submitted work in person if

there is reason to believe this policy is being violated. Violations of the above principles will be treated as an infringement of the university's standards for academic integrity and subject to applicable disciplinary procedures. Finally, students are encouraged to consider the ethical, social, and environmental consequences of AI use before deciding on whether and how to use this technology.

**Communication**: Please check your McGill email and MyCourses regularly for course updates. When emailing the instructor, please use your McGill email account and include POLI 312 in your subject line. I will do my best to respond within 48 hours – please do not expect instant replies to emails, especially in the evening or over the weekend. Please email for logistical issues or clarification. For more substantive issues, please see me during drop-in hours.

**Decorum**: Students are expected to arrive to class on time and behave in a manner that is respectful to the instructors and to fellow students. Opinions held by other students should be respected; harassment, derogatory comments, personal attacks on others, or interrupting the class will not be tolerated. Please avoid the use of cell phones and electronics for non-class related purposes.

**Accommodation**: Students experiencing an extraordinary personal situation, or a temporary illness may request additional assistance and support in order to meet certain academic obligations. The Student Accessibility and Achievement Office is available to meet with students to discuss ways to provide some flexibility in the program and to accommodate particular circumstances.

Students who, because of religious commitments, cannot undertake or submit an assessment task in a course have the right to request reasonable accommodation in fulfilling the assessment in accordance with the Policy for the Accommodation of Religious Holy Days. Pregnant students and students caring for dependents have the right to request reasonable accommodation in fulfilling an assessment in a course in accordance with the Guidelines for the Academic Accommodation of Pregnant Students and Students Caring for Dependants.

**Extraordinary Circumstances**: In the event of extraordinary circumstances beyond the University's or instructor's control, the content and/or evaluation scheme in this course is subject to change.

## Course Outline & Readings

Note: weekly readings and topics are subject to change. Refer to MyCourses for the most up-to-date version of the syllabus. For each week in the term, please read the assigned materials before the indicated date in the schedule below.

## Section I: Reviewing Fundamental Statistics

**Week 1** **Jan. 5-7**

- Jan 5: Introduction
  - Course goals, structure, and requirements
  - Description, causation, and prediction
  - ROS 1
- Jan 7: Data refresher
  - The tidyverse, ggplot2, data visualization, reshaping data
  - ROS 2; QSS 1, 3
  - **Conference**: Refresher on R and RMarkdown

**Week 2** **Jan. 12-14**

- Jan 12: Workshop (Data)
- Jan 14: Random Variables I
  - Probability and cumulative distributions, location and dispersion
  - ROS 3; QSS 6
  - **Conference**: Workshop review

**Week 3** **Jan. 19-21**

- Jan 19: Random Variables II
  - Marginal, joint, conditional distributions, association
  - ROS 3; QSS 6
  - **Problem Set 1 due**
- Jan 21: Inference I
  - Estimation, standard errors, sampling distributions, confidence intervals,
  - ROS 4; QSS 7.1-7.2

**Week 4** **Jan. 26-28**

- Jan 26: Inference II
  - Hypothesis testing
  - ROS 4; QSS 7.1-7.2
  - **Problem Set 1 revision due**
- Jan 28: Workshop (Simulation + Inference)
  - ROS 5

**Week 5:** **Feb. 2-4**

- **Feb 2: Quiz 1**

## Section II: Revisiting Regression

- Feb 4: Regression I
  - Non-parametric regression, local linear regression
  - ROS 6
  - **Conference**: Workshop review

**Week 6:** **Feb. 9-11**

- Feb 9: Regression II
  - Linear regression model
  - ROS 7, 8.1, 11.1, 11.6
  - QSS 4.1 - 4.2
- Feb 11: Regression III
  - Inference for regression
  - ROS 7, 8.1, 11.1, 11.6
  - QSS 7.3
  - **Research Project Descriptive Analysis due**

**Week 7:** **Feb. 16-18**

- Feb 16: Multiple regression
  - Multiple regression, omitted variable bias
  - ROS 10.1-10.2
- Feb 18: Interactions and interpretation
  - Indicator variables, interaction effects, polynomials
  - ROS 10.3-10.4, 11, 12.1-12.4

**Week 8:** **Feb. 23-25**

- Feb 23: Workshop (Regression)
- **Feb 25: Quiz 2**
  - **Conference**: Workshop review

**Week 9: Reading Week** **Mar. 2-4**

## Section III: Applications

**Week 10:** **Mar. 9-11**

- Mar 9: Randomized experiments I

  - Potential outcomes, average treatment effects, randomization, identification
  - ROS 18.1-18.3
  - QSS 2.3 - 2.5, 4.3
  - **Problem Set 2 due**
- Mar 11: Randomized experiments II
  - Balance, covariate adjustment, heterogeneous treatment effects
  - ROS 18.1-18.3
  - QSS 2.3 - 2.5, 4.3

**Week 11:** **Mar. 16-18**

- Mar 16: Ethics and transparency in experimentation
  - Teele, Dawn Langan. "Reflections on the ethics of field experiments." In Teele, Dawn Langan (ed.). *Field experiments and their critics: Essays on the uses and abuses of experimentation in the social sciences* (2014): 115-40.
  - Nosek, Brian A., et al. 2018. "The Preregistration Revolution." *Proceedings of the National Academy of Sciences 115* (11): 2600–2606.
  - **Problem Set 2 revision due**
- Mar 18: Regression for causal inference I
  - Difference-in-differences
  - QSS 2.5

**Week 12:** **Mar. 23-25**

- Mar 23: Regression for causal inference II
  - Regression discontinuities
  - QSS 4.3.4
  - ROS 21.3
  - **Research Project Regression Analysis due**
- Mar 25: Workshop (Causal inference)
  - **Conference**: Workshop continuation

**Week 13:** **Mar.30-Apr. 1**

- **Mar 30: Quiz 3**
- Apr 1: Descriptive inference I
  - Text-as-data

  - QSS 5.1
  - **Problem Set 3 due**

**Week 14:** **Apr. 6-8**

- Apr 6: Easter Monday (no class)
- Apr 8: Descriptive inference II
  - Spatial data
  - QSS 5.3
  - **Problem Set 3 revision due**

**Week 15:** **Apr. 13**

- Apr 13: Workshop (Descriptive)

**Research Project Final Report Due: Apr. 20**