

# SOCI 350: Statistics in Social Research

Instructor: Luca Maria Pesando

Fall 2019

E-mail: [lucamaria.pesando@mcgill.ca](mailto:lucamaria.pesando@mcgill.ca)

Phone: 438-525-7848

Office Hours: T 4-6 pm (drop-in), TH 2-3 pm (by appt)

Class Hours: T,TH 10:05-11:25

Office: Stephen Leacock Building, Room 821

**Class Room: CAMPUS1 109**

---

TA: **Ann-Marie Helou**

E-mail: [ann-marie.helou@mcgill.ca](mailto:ann-marie.helou@mcgill.ca)

Office Hours: M 10-12 am

Conference/Lab Hours: TBD

Office: TBD

Conference/Lab Room: TBD

---

TA: **William Zachary Marshall**

E-mail: [william.marshall2@mcgill.ca](mailto:william.marshall2@mcgill.ca)

Office Hours: T 12.30-2.30 pm

Conference/Lab Hours: TBD

Office: 318 Peterson Hall

Conference/Lab Room: TBD

See footnote for **important instructions** about how to access the classroom. <sup>1</sup>

## 1 Course Description

This course introduces students to the basics of statistics and data-driven analysis, with a focus on how they are used in social science research. At the end of this course, students should be able to use various numerical and graphical tools to describe a single variable, and to summarize the distribution of a variable using measures of central tendency and spread. In addition, they should be able to use correlation and — time permitting — basic regression to describe the relationship between a pair of variables.

---

<sup>1</sup>Process for students requiring elevator access to the McGill classrooms located at CAMPUS1 MTL

1. The student will check in at the front desk with security at the main Campus1 entrance. They will be instructed to sign into the guest log.
2. Security will issue a key fob on a purple lanyard to the student.
3. The student will be accompanied to the elevator by security and up to the 1st floor in the elevator.
4. The student will exit the elevator vestibule by pressing the automatic door opener on the wall.
5. Once the student is ready to exit the premise, they will use their key fob to open the elevator lobby door.
6. The student will press the down button to call the elevator, and choose RC level once in the elevator.
7. The student will then be able to exit the speed gates on the RC level by using their FOB.
8. The student needs to sign out of the guest binder and return their fob to the security agent.

The course also covers the basics of statistical inference, i.e., a set of methods that allow to draw conclusions about some broader population on the basis of sample data. It discusses the concepts underlying probability sampling, what a sampling distribution is, and the role of a sampling distribution in statistical inference. Students apply these concepts by learning how to test hypotheses about means, proportions, and pairs of means and proportions. They also learn how to calculate the confidence intervals associated with these tests. Finally, the course provides students with real-world skills that they can put on their resumés when applying for jobs.

### **1.1 Who is this course for?**

- This is your first semester-long course on quantitative methods for data analysis. If you have already taken a full semester quantitative (or statistics) course, you are recommended not to enroll.
- You want to be able to better read and interpret quantitative research in the social sciences (even if you only intend to do qualitative work yourself).
- You are willing to spend time outside of the classroom to learn the course materials, as data analysis is a skill learned through actual practice.
- You want to be able to apply quantitative methods to your papers and future career.

### **1.2 Prerequisites**

Some familiarity with elementary algebraic notation is assumed. This is an introductory course, so no previous experience in statistics is required. The course serves as a background for further statistics and data-analysis courses, helping to provide the intuition that can sometimes be lost amid the formulas.

## **2 Course Objectives**

- Learn the basic tools of empirical research in the social sciences.
- Gain competency in the critical evaluation of empirical analyses of the social world you are confronted with in civic life.
- Obtain the tools both technical and intellectual to perform statistical analyses and critiques of your own.
- Gain real-world skills (such as coding or project-management skills) that will help you obtain jobs in careers of the future.

### 3 Instructional Method

The course will be organized as a mix of lectures in which I review the conceptual tools behind specific sets of statistical analyses, and of applied demonstration/exercise/lab sessions aimed at helping you learn how to perform these analyses using one popular statistical software named Stata. Ideally, I will hold an applied/lab session after every 3 lectures. You will find the demonstration/exercise sessions most useful if you can follow me on your computers and be active participants in the exercises I provide. Therefore, I highly encourage you to have your computers or tablets with you in class, **but only when the applied/lab sessions are held** — see Sections 7 and 10 at the end of the syllabus.

Required readings are assigned. These readings will come from two main sources. First, I will regularly assign chapters from an introduction to statistics textbook. At the end of each chapter from this textbook, a series of simple exercises are proposed. I strongly recommend working through as many of these exercises as you can every week. They will help you think through the concepts we learn and make sure you have understood them. Second, I will assign — yet occasionally — journal articles that will be posted to MyCourses ahead of class. When this happens, I will ask you to come prepared to discuss the readings.

#### 3.1 Textbook

The textbook for this course will be:

Moore, David, William I. Notz, Michael A. Fligner. *The Basic Practice of Statistics*, **Seventh Edition**. New York: W.H. Freeman.

If you find a cheaper version which is an older edition (Sixth), that is also fine. I will list the discrepancies and provide reading references for both the Sixth and Seventh editions. Note that I myself have the Seventh edition. The Eighth edition — likely more expensive — is also fine, and there are almost no discrepancies with respect to the Seventh. I am aware it can take time to obtain the book. It will be fine to obtain it by the beginning of the third class. I will provide an online PDF of the first and second chapters to help students get started while they wait for their book.

The textbook includes a formula card that might be helpful for this course. When you print the formula card, review it carefully to make sure that all symbols print correctly. If you have problems obtaining the card, please email either myself or the TAs.

Any additional readings will be provided on the MyCourses page and advertised before the respective class. You won't have to look for additional readings yourselves, I will make them available for download.

### 3.1.1 Other Resources (optional)

- David Pevalin and Karen Robson. 2009. *The Stata Survival Manual*. McGraw–Hill. This book provides guidance on how to use STATA, a statistical software widely used in the social sciences. **Note:** This book is optional, as both myself and the TAs will guide you towards learning Stata both in class and during office hours. Also, there are multiple online resources that are good substitutes of this book — feel free to ask us if interested.
- Moore’s text comes with useful online supplements that you are encouraged to explore (<https://www.macmillanlearning.com/studentresources/college/collegebridgepage/bps8e.html>). These include self-quizzes, additional exercises, statistical applets, and data sets.

### 3.2 Course Website

I will use **MyCourses** to post resources used in this class. These include the syllabus, the required readings, the databases we will use to perform statistical analyses, the demos run in class or labs, and lecture overheads. It is the students’ responsibility to familiarize themselves, if they have not done so already, with MyCourses. The MyCourses website is located at <https://mycourses2.mcgill.ca/d2l/home>.

I will attempt to post the lecture overheads the day before or the morning (early) of each class so you can take notes on them. However, I remind students that simply looking at the overheads is not a substitute for attending lectures. I will further post revised overheads after the class if any change/correction is to be made.

### 3.3 Preparing for Lectures

One way to study: to prepare for class, read the chapter, just skimming the problems, before it is covered in class. After class, read the chapter again and do the homework problems; do additional problems if you are having trouble. Then read the text of the chapter again to solidify what you learned. At the beginning of every class, I will make an effort to review briefly what we have covered in the previous class.

### 3.4 Computer Software

The **software** I will use in this class for demonstration/exercise sessions and that I also expect you to utilize to complete assignments is called Stata. This software is installed on many computers at McGill Library: <https://www.mcgill.ca/library/services/computers>. For instance, at the Data Lab in McLennan Library Building (M2-37A), in rooms 313 and 413 at Schulich Library of Physical Sciences, Life Sciences, and Engineering, and at Macdonald Campus Library (ezone).

You can also purchase a temporary student license to use Stata on your own computers. If you choose to do so, you will be able to follow and replicate in real time and on your own computers the demos I run in class. However, be sure that all the syntax I produce during these demos will also be posted to MyCourses the day after class. For this class, Stata/IC will work. The price for a temporary Stata/IC license is \$89 for 1-year or \$45 for six months (<https://www.stata.com/news/student-pricing/>). If you think you might use Stata again in the future for your own research, however, I recommend opting for Stata/SE instead of Stata/IC.

### 3.5 Calculators

You will need a scientific calculator for some assignments and the exams (not the phone). You may want to choose a calculator that also has some statistical functions (mean, standard deviation, and correlation) built in. You are responsible for learning how to use the statistical functions on your calculator. We do not provide specific instruction on how to use calculators.

## 4 Evaluation Policy

Below is a description of the means of evaluation to be used in the course. Note that there will be **no final exam**. Your final grade will be based on the following components:

- **Two midterms/quizzes, each worth 25% of your final grade (50% overall).** These will be held in class and will last approximately 1 hour. They will be non-cumulative. Tentative exam dates are on the schedule below, but dates may change. You must attend the exam. Exam questions will include true/false questions, multiple choice questions, and open-ended questions that require discussion, data analysis and calculation, or the selection of appropriate statistical methods. A copy of the quick reference card from the Moore book will be provided for your use during the exams; you may use this reference card and a calculator, but no other material. The quick-reference card has copies of almost all of the formulas covered in the course.
- **Three assignments, worth 15%, 15%, and 20% of your final grade (50% overall).** I recommend starting working on the assignments as soon as they become available, which will give you the opportunity to follow the course as topics are being discussed, and also give you ample time to troubleshoot any possible issues with me and/or the Teaching Assistants. I will post assignments on MyCourses the day announced, and discuss them in class subsequently. Unless you meet an exception criterion, no late assignments are accepted. More details on the assignments below:
  1. **Assignment 1, individual assignment, 15%:** In this assignment, you will be asked to perform sets of statistical analyses using Stata. These analyses will be similar to, and inspired by, the demonstrations I present in class. You can

discuss this assignment with other students in the class to further your understanding of the material, but you must write it up independently. To simply copy another student's assignment and turn it in is cheating. You will be expected to turn in both your answers to the assignment's questions and the Stata syntax you have employed to reach them (you will be taught on how to do this).

2. **Assignment 2, individual assignment, 15%:** Same type of assignment and same rules as for Assignment 1.
3. **Assignment 3, group assignment/project, 20%:** This is a different type of assignment that you are supposed to complete in groups of 3/4 students, where each team member will obtain the same grade — and only one write-up and code per group will be submitted. One of the goals of this course is to help you develop your skills as a producer of social research. As such, this assignment consists of a data-analysis project focused on descriptive statistics and graphical display of data, accompanied by a 5-page write-up — basically, a short academic paper on a research question of your own interest. We will provide you with a clean and simplified dataset that you can use for your project. You are also welcome to use your own dataset, but please share it with us beforehand so we can approve it and provide proper guidance. Details on this group project will be specified in class during the course. I will also create a Group discussion on MyCourses to facilitate you finding group members. If you have troubles finding group members, reach out to me or the TAs. This group project will mark the completion of this course, as there will be no final exam.

**Optional assignment** - see also Section 6: To make it easier for those who miss either Assignment 1 or 2 for emergency reasons (properly documented), I will arrange an additional individual assignment which they can do by another set date. Students who already completed Assignments 1 and 2 are not supposed to do this assignment. However, if they feel like doing it, and they are not satisfied with the grades they obtained in the first two, they can still do the optional assignment. In that case, I will compare the three individual assignments and keep only the two with the highest grade. This rule does not apply to Assignment 3 (the group one), which is compulsory for every student, and the grade of which cannot be replaced in any way. To summarize, the chronological order of exams and assignments is below:

Assignment 1 (Individual)	15%
Midterm/Quiz 1	25%
Assignment 2 (Individual)	15%
Midterm/Quiz 2	25%
Assignment 3/Final Project (Group)	20%

#### 4.1 Re-Grading

Students who wish to contest a grade for an assignment or exam must do so in writing (by email, sent to me) providing the reasoning behind their challenge to the grade received within two weeks of the day on which the assignments are returned. The TA who graded the assignment will re-grade your assignment, and may **raise or lower the grade**. If you are still unsatisfied after the re-assessment, you can re-submit the assignment to me (original copy with TA comments), along with your justification. I will then re-evaluate the paper, but also reserve the right to **raise or lower the grade**.

#### 4.2 Attendance

It is your responsibility to go to class and attend the lectures. As such, attendance won't be recorded class by class and no clickers will be used. Attendance will not count towards your final grade. However, during two or three lectures throughout the term I will hand in short surprise quizzes (15-mins approx., **not graded**) to see who is in class. This information will be used to raise or lower grades in case of borderline situations (e.g., if someone is very close to reaching a higher grade threshold and is found to be in class all three times, I will consider raising her/his grade.).

### 5 Communication Policy and Class Discussion List

My preference is to answer questions in class or in office hours — this format is best for avoiding misunderstandings (which are common via email or when conversations are rushed). My office hours are listed above. For questions regarding using Stata, the TAs will also have office hours to help you and answer your questions. Please, stick to their office hours (or in class). Also, if at any time you feel that what I am doing is not advancing your learning, please let me know (in a respectful manner) – I want each and every one of you to feel safe and to learn, so please let me know if that is not happening. Outside of office hours, I may be contacted by email at the address specified above.

I will set up a class discussion list on *MyCourses*. I encourage you to use this mailing list to ask questions you may have. Do not post your code or specific homework questions on the course list server. Please post general questions. If you post homework codes on the website, it will be taken down and your grade may be lowered.

### 6 Make-Up Work Policy

If you are absent for an emergency, medical or family reason, I will try my best to arrange an alternative exam date provided the absence is properly documented. If you cannot provide a valid reason for your absence, you will receive no (zero) points for the missed exam. As for the assignments, I will arrange an optional one that students who missed either Assignment 1 or 2 could do (see Optional Assignment in Section 4). Also, keep in mind that you cannot turn in anyone else's homework for them.

If you must miss a class, an examination, a work assignment, or a project in order to observe a religious holiday, you will be given an opportunity to complete the missed work within a reasonable time after the absence.

## 7 Technology Policy

### 7.1 Screens in the Classroom Policy

Mobile computers, telephones and tablets (MC) are strictly limited in this class. The current [literature](#) recognizes that MCs inhibit learning. I have found that computers distract the student using the computer as well as her peers. So no computers, tablets or phones are allowed to be used in the classroom unless specified by the instructor. **Computers are only allowed when we hold a lab/demonstration session in class** — check the detailed schedule to see when these sessions are arranged — typically after every 3 lectures. Indeed, bring your computer to office hours if you have doubts. The TAs and myself will help.

## 8 Academic Integrity

### 8.1 Course Policy on Computer Code

Just like writing a paper, copying other people's computer code constitutes plagiarism. Moreover, data programming is learned through trial and error. **Please do not under any circumstance copy another student's code.** If you are found to have done so, you may be referred to the appropriate Dean. The instructors reserve the right to use software to compare the code that has been written by different students.

### 8.2 McGill Policy

"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 <https://www.mcgill.ca/students/srr/honest/> for more information).

## 9 Other Policies

### 9.1 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.

### 9.2 Disabilities Policy

As the instructor of this course I endeavor to provide an inclusive learning environment. However, if you experience barriers to learning in this course, do not hesitate to discuss them with me and the Office for Students with Disabilities, 514-398-6009.



### 9.3 Other McGill Resources and Support

If you need counseling and mental health support, do not hesitate to contact Counseling Services, where Psychologists, Social Workers, Counseling, Psychotherapies and Psychiatrists support the wellbeing and mental health of McGill students. You can reach the Student Wellness Hub at <https://www.mcgill.ca/wellness-hub/>.

If you have been impacted by sexual violence — sexual harassment or assault, gender-based or intimate partner violence, cyberviolence —, do not hesitate to contact the Office for Sexual Violence Response, Support and Education. You can reach the Office by mail at [osvrse@mcgill.ca](mailto:osvrse@mcgill.ca), in person at 550 Sherbrooke O. Suite 585 (West Tower elevators 1-11) and visit <https://www.mcgill.ca/osvrse/>.

### 9.4 End-of-Course Evaluations

End-of-course evaluations are one of the ways that McGill works towards maintaining and improving the quality of courses and the student's learning experience. You will be notified by e-mail when the evaluations are available. Please note that a minimum number of responses must be received for results to be available to students.

## 10 Class Schedule

Scheduling of topics for the course listed below and associated readings are subject to change at my discretion. Such changes, should they take place, will be announced in class and through MyCourses. Students are responsible for keeping abreast of any changes made.

### Part a. Design of Data, Data Analysis, and Descriptive Statistics

#### Week 1

- 03/09: Introduction, class syllabus, and course overview (Moore chapter 0, **optional**)
- 05/09: Picturing distributions with graphs (Moore chapter 1)

#### Week 2

- 10/09: Describing distributions with numbers (Moore chapter 2)
- 12/09: The Normal distributions (Moore chapter 3)

#### Week 3

- 17/09: Review and Stata Lab - Introduction to coding and stata syntax (very important session - **bring laptop**)
- 19/09: Scatterplots and correlation (Moore chapter 4) — **Assignment 1 out**

#### Week 4

- 24/09: Regression, part 1 (Moore chapter 5)
- 26/09: Regression, part 2 (Moore chapter 5)

#### Week 5

- 01/10: Review and Stata Lab (**bring laptop**)
- 03/10: Two-way tables (Moore chapter 6) — **Assignment 1 due**, bring hard copy to class

### Part b. Data Production

#### Week 6

- 08/10: Producing data: Sampling (Moore chapter 8)
- 10/10: **Midterm/quiz 1** (covering chapters 1-2-3-4-5-6 - chapter 7 provides a summary of all these chapters and might be helpful for a review)

### Week 7

- 15/10: Producing data: Experiments (Moore chapter 9)
- 17/10: Introducing probability (Moore chapter 12; corresponding to chapter 10 in Edition 6th)

### Part c. Probability and Statistical Inference

### Week 8

- 22/10: Sampling distributions (Moore chapter 15, corresponding to chapter 11 in Edition 6th) — **Assignment 2 out**
- 24/10: Review and Stata Lab (**bring laptop**)

### Week 9

- 29/10: Confidence Intervals - The basics (Moore chapter 16, corresponding to chapter 14 in Edition 6th)
- 31/10: Tests of significance - The basics (Moore chapter 17, corresponding to chapter 15 in Edition 6th)

### Week 10

- 05/11: Recap. on confidence intervals and significance tests (Moore chapters 16 and 17) — **Assignment 2 due, bring hard copy to class**
- 07/11: Review and Stata Lab (**bring laptop**)

### Week 11

- 12/11: **Midterm/quiz 2** (covering chapters 8-9-12-15-16-17)
- 14/11: Inference in practice (Moore chapter 18, corresponding to chapter 16 in edition 6th)

### Week 12

- 19/11: Inference about a population mean (Moore chapter 20, corresponding to chapter 18 in edition 6th)
- 21/11: Two-sample problems (Moore chapter 21, corresponding to chapter 19 in edition 6th)

### Week 13

- 26/11: Inference about a population proportion (Moore chapter 22, corresponding to chapter 20 in edition 6th)

- 28/11: Recap. on inference about a population mean and proportion (Moore chapters 20-21-22)

**Week 14**

- 03/12: Review and Stata Lab (**bring laptop**) - class dedicated to discussion of group projects

**Assignment 3** (Group project) is due on **December 15 by 6pm**