POL 232H1: Introduction to Quantitative Reasoning II

Course Outline

Winter 2017
(Section L0201)

Class Time: **Thursdays, 10AM–12 PM**
Class Location: **SS 561 (Sidney Smith Hall 561)**

Instructor: Prof. Ludovic Rheault
Email: ludovic.rheault@utoronto.ca

Office Hours: Thursdays, 1–3 PM
Office Location: Sidney Smith 3005

Teaching Assistant: Eve Bourgeois
Email: eve.bourgeois@mail.utoronto.ca

Course Description

We live in an information society where the ability to analyze data has become an essential skill for college-educated students integrating the job market or pursuing graduate studies. Throughout their professional career—whether in government, business or non-governmental organizations—students will have to deal with empirical data on a regular basis. For instance, they may be asked to interpret statistics comparing economic outcomes across countries, measuring social mobility for various groups of population, or assessing the environmental impact of public policies.

To prepare them adequately for the modern work environment, this course provides students with an introduction to applied statistical analysis. We will cover not only the theory behind statistical methods, but also the practical aspects of data analysis using computers. Students will be introduced to fundamental concepts in statistics. In particular, the course covers descriptive statistics, visualization techniques, measures of association between variables, regression analysis, and deals with the key concept of *statistical inference*. The course also includes concrete training with one of the most widely used programming languages for statistical computing: R. By the end of this course, students will be able to conduct basic statistical analysis independently.

Course Format

The pedagogical approach emphasizes examples based on real-world political data, so that students can more easily relate the methods they learn to concrete applications. The course consists of lectures given in a computer lab on Thursdays, combined with tutorials directed by a teaching assistant (TA) every other week. The tutorials are taking place in the same classroom, before or after the lectures, with the class divided into two smaller groups. Tutorials provide students with opportunities to practice exercises with the help of the teaching assistant.

Collaborative work is strongly encouraged in this course. Participation in class is invited, and students are encouraged to expand their learning experience through interactions with their classmates. In particular, students can work in teams to complete the exercises and term paper.

Contact Policy

Questions about the course material can be asked on the Blackboard portal for the course or in person during office hours. Emails should be restricted to administrative and formal questions, since they are sub-optimal for providing feedback and explanations about course content.

Marking Scheme

Exercise #1	20%	**Due:** February 9, 2017.
Exercise #2	20%	**Due:** March 2, 2017.
Final Test	25%	March 23, 2017.
Term Paper	25%	**Due:** April 5, 2017.
Participation	10%	Throughout the semester.

Readings

REQUIRED
- **Alan Agresti and Barbara Finlay. 2009.**
  ***Statistical Methods for the Social Sciences.* 4th Ed. Prentice Hall.**
  - Available at the UofT Bookstore. For those students with a tight budget, know that the textbook has been used in POL 232 in previous years, which means that there may be used copies in resale stores. Used copies may also be available online on websites such as amazon.com.
  - Abbreviated as **AF** below.
- **James E. Monogan III. 2015.**
  ***Political Analysis Using R.* Springer.**
  - Available in electronic format for registered students via the UofT library, or at the Bookstore.
  - Abbreviated as **M** below.

OTHER USEFUL GENERAL REFERENCES
- David S. Moore, William I. Notz and Michael A. Fligner. 2013. *The Basic Practice of Statistics.* 6th Ed. W. H. Freeman and Company.

  - An exhaustive substitute to the main textbook **AF**. Students may opt for this one as long as they are willing to find the corresponding chapters to follow along.

- Phil H. Pollock III. 2014. *An R Companion to Political Analysis.* CQ Press.

  - Could be used as a substitute to **M**.

- Andy Field, Jeremy Miles and Zoé Field. 2012. *Discovering Statistics Using R.* Sage.

  - Another introductory book on R available at the Robarts library.

Software

For practical exercises, the course relies upon the R language for statistical computing, one of the most popular choices for data analysis in both the academia and private organizations. In-class examples will be provided from RStudio, a free graphical interface to use the R language. Both R and RStudio (non-commercial) are free to use and students may choose to install them on their personal computers to practice exercises if they wish to.

The required software will be available on all stations in the SS 561 computer lab. For students who do not have access to a personal computer, know that R is available in most computer labs across the campus and in the Department of Political Science.

Teaching Assistant and Tutorial Groups

Eve Bourgeois, PhD student in the Department of Political science, is the teaching assistant for this course. She will be responsible for tutorial sessions and for grading evaluations. Students are required to attend the tutorial sessions to which they have registered via ACORN.

Tutorials take place immediately before or after the class, depending on the group students have chosen during registration. The room is the same as for the lectures (SS 561).

Evaluations

Exercises

The exercises are problem sets for which students are asked to analyze political data using the methods learned during class. These problem sets can be done in teams of two students and are to be handed in at the due date.

Term Paper

The term paper is an original research report (2,000–2,500 words in length, or 7–8 pages) in which students are asked to put into practice at least some of the methods learned during the course. Students may use one of the datasets at their disposal or collect original data. Topics require approval from the TA. Collaborative work is once again encouraged, and the term work can be done in teams of up to two students.

Late Penalty: The **late penalty** for the exercises and term paper is **2 percentage points per day**, including weekends.

Participation

Developing skills with statistical analysis requires practice. Participation points are used to ensure that students are actually completing in-class exercises. The marks will be attributed proportionally to students' attendance and participation during tutorials. One participation point will be granted for filling out the course evaluation.

Final Test

The final test takes place at the end of the semester during regular class hours. This is a closed-book test. Relevant formulas and memory help will be made available to students.

Turnitin.com Statement

> Normally, students will be required to submit their course essays to Turnitin.com for a review of textual similarity and detection of possible plagiarism. In doing so, students will allow their essays to be included as source documents in the Turnitin.com reference database, where they will be used solely for the purpose of detecting plagiarism. The terms that apply to the University's use of the Turnitin.com service are described on the Turnitin.com web site.

Academic Integrity Statement

Please read the following statement regarding academic integrity. For more information, consult the link provided below to consult UofT's Code of Behaviour.

> Academic integrity is fundamental to learning and scholarship at the University of Toronto. Participating honestly, respectfully, responsibly, and fairly in this academic community ensures that the U of T degree that you earn will be valued as a true indication of your individual academic achievement, and will continue to receive the respect and recognition it deserves.
>
> Familiarize yourself with the University of Toronto's Code of Behaviour on Academic Matters (http://www.governingcouncil.utoronto.ca/policies/behaveac.htm). [...]

Class Schedule: Summary of Readings, Tutorials and Due Dates

The following table summarizes the schedule for lectures and tutorials. Readings for each week are indicated using acronyms: **AF** refers to the main textbook (Agresti and Findlay 2009) and **M** refers to the R textbook (Monagan 2015). Numbers refer to chapters.

Date	Description	Readings	Tutorials	Due Dates
January 5	Course introduction and R Language	**AF1 M1**		
January 12	R Language & Univariate Statistics	**AF2 M1–2**		
January 19	Measures of Central Tendency	**AF3 M4**	Tutorial 1	
January 26	Distributions and Visualization	**AF4 M3**		
February 2	Statistical Inference	**AF5 M5**	Tutorial 2	
February 9	Measures of Association I	**AF6-7 M5**		**Exercise #1 Due**
February 16	Measures of Association II	**AF8 M5**	Tutorial 3	
February 23	[*Reading Week–No Class*]			
March 2	Linear Regression I	**AF9 M6**	Tutorial 4	**Exercise #2 Due**
March 9	Linear Regression II	**AF10–11 M6**		
March 16	Linear Regression III	**AF12,14 M6**	Tutorial 5	
March 23	**Final Test**			**Final Test**
March 30	Logistic Regression	**AF15 M7.1**	Tutorial 6	
April 5	[*End of Term*]			**Term Paper Due**

Note: Dates for each topic and the tutorial schedule may be adjusted depending on the progress made during class and to maximize the pedagogical benefits to students. Note

Detailed Schedule

---

1. Introduction to Data Analysis and R
(January 5-12)

- **Required Readings:**
  - **AF**1–2; M1–2.
- **Topics:**
  1. Computers and the (big) data revolution.
  2. Overview of fundamental concepts in data analysis.
  3. Causality and inference.
  4. Sources of data in political science.
  5. Introduction to R: data types, input/output, arrays, functions, plots.

---

2. Measures of Central Tendency and Dispersion
(January 19-26)

- **Required Readings:**
  - **AF**3–4; M4.
- **Topics:**
  1. Variable types.
  2. Measures of central tendency.
  3. Measures of dispersion.

Tutorial 1: Introduction to R (January 19)

---

3. Distributions and Visualization Techniques
(January 26)

- **Required Readings:**
  - **AF**4; M3.
- **Topics:**
  1. Probability and distributions.
  2. The normal distribution.
  3. Summarizing data and distributions by variable type.
  4. Graphs and visualization techniques in R.

---

4. Statistical Inference
   (February 2)

- **Required Readings:**
    - **AF**5; M5.
- **Topics:**
    1. Statistical inference.
    2. Central limit theorem.

Tutorial 2: Univariate Analysis and Visualization (February 2)

5. Measures of Association I
   (February 9)

- **Required Readings:**
    - **AF**6–7; M5.
- **Topics:**
    1. Measures of association for continuous variables.
    2. Significance testing: z-test and t-test.
    3. Mean-comparison tests.
    4. Correlation and covariance.
- **Exercise #1 Due on February 9**.

6. Measures of Association II
   (February 16)

- **Required Readings:**
    - **AF**8; M5.
- **Topics:**
    1. Measures of association for discrete variables.
    2. Significance testing: Chi-square test.

Tutorial 3: Measures of Association (February 16)

February 23: Reading Week—No Class

7. Linear Regression I
(March 2)

- **Required Readings:**
  - **AF**9; M6.
- **Topics:**
  1. Linear regression framework.
  2. Coefficient of determination.
- **Exercise #2 Due on March 2**.

Tutorial 4: Measures of Association and Linear Regression (March 2)

8. Linear Regression II & III
(March 9)

- **Required Readings:**
  - **AF**10–11; M6.
- **Topics:**
  1. Linear regression: estimation, inference and interpretation.
  2. Multiple (multivariate) linear regression.

8. Linear Regression III
(March 16)

- **Required Readings:**
  - **AF**12,14; M6.
- **Topics:**
  1. Multiple linear regression (continued).
  2. Analysis of variance (ANOVA).
  3. Regression diagnostics.
  4. Violations of assumptions of linear regression.

Tutorial 5: Linear Regression II (March 16)

March 23: Final Test

(Location and Time: Same as regular class.)

9. Logistic Regression
(March 30)

- **Required Readings:**
  - **AF**15; M7.1.
- **Topics:**
  1. Logistic regression.
  2. Maximum likelihood.
  3. Overview of advanced topics in statistics.

Tutorial 6: Logistic Regression (March 30)

April 5: Term Paper Due