

Autumn 2018 Syllabus for STA302/1001

Methods of Data Analysis I

Instructor: Professor Mark Ebden

Email: mark (dot) ebden (at) utoronto (dot) ca

Instructor's office hours: In EP 102 by appointment only, from 6-7 pm on Mondays (click [here](#) to book) and 2-3 pm on Thursdays (click [here](#) to book).

Background: My experience has mainly been in the analysis of *scientific data* using probabilistic approaches. This is my second time teaching STA302.

Teaching Assistants' office hours: To be finalized once TA hiring is complete.

Course webpage: Accessible through Quercus at g.utoronto.ca

Classroom sessions:

- **Section 1:** Twenty-three sessions (excluding midterm) from Thursday 6 Sept to Tuesday 4 Dec; in room OI G162, 10-11 am (Thurs) & 10 am – 12 pm (Tuesdays)
- **Section 2:** Eleven sessions (excluding midterm) from Thursday 6 September to Thursday 29 November; in room ES 1050, 5-8 pm

Prerequisites

- STA238H1/STA248H1/STA255H1/STA261H1/ECO227Y1
- CSC108H1/CSC120H1/CSC121H1/CSC148H1
- MAT221H1(70%)/MAT223H1/MAT240H1

You will need to recall basic matrix operations which you learned in the linear algebra course. An example of a review of matrix algebra is available in the first ten pages [here](#).

Course Content

This course covers the theory and application of linear regression analysis. In particular this course will cover: initial examination of data, correlation, simple- and multiple regression models using least squares, geometry of least squares, inference for regression parameters for normally distributed errors, confidence and prediction intervals, model diagnostics and remedial measures when the model assumptions are violated, interactions & dummy variables, ANOVA, model selection, and ridge regression.

The learning objectives of this course are:

- To gain a solid understanding of linear regression analysis
- To learn practical skills for developing linear regression models for inference and prediction, and to interpret the results

Textbook

A Modern Approach to Regression with R by Simon Sheather (2009)

- ISBN: 978-0387096070

- Most of chapters 2-7

- The [textbook website](#) has datasets and other resources

Additional reference material (optional)

Applied Linear Regression Models, 4th edition, by Michael Kutner, Christopher Nachtsheim, and John Neter (2004)

- ISBN: 978-0073014661

- Much of Chapters 1-8, and some content from chapters 9-11, overlaps with our course

Applied linear regression, 4th edition, by Sanford Weisberg (2014)

- ISBN: 978-1118386088

Evaluation

Undergraduate- and graduate students will be evaluated according to the following marking scheme.

Assessment	Weight	Notes
Quizzes	10%	Occurring during class time
Assignments	10%	Not to be discussed in office hours
Term test	35%	Occurring during class time
Final exam	45%	Expected 8-21 December

Graduate students will be evaluated at the graduate level according to the [University Assessment and Grading Practices Policy](#).

You are not evaluated on attendance or similar forms of participation. But, generally the more you put into a course the more you'll get out of it.

Assessments

You must bring your **student identification** to the quizzes, term test, and final exam. These are closed book and closed notes. The location of the midterm will be announced later. The midterm papers may differ between the two course sections but the final exams are identical. Practice problems will be posted on Portal to help you prepare for the midterm and exam and are not to be handed in. You need to know basic R syntax for the midterm and exam, and you must also know how to interpret output from R.

Marking concerns

Any requests to have your quiz/test/assignment mark re-evaluated must be made in writing within *one week* of the date the work was returned. The request must contain a justification for consideration. Send your request to sta302sec1@gmail.com or sta302sec2@gmail.com, depending on your section.

Submitting your assignments

- Assignments are submitted electronically on Quercus, never via email messages.
- If you leave your submission to the last minute and an overburdened Quercus records a late timestamp while processing multiple requests, this is your responsibility. If you're not *feeling lucky*, submit at least 30-60 minutes early.
- Assignments submitted late by less than 24 hours will be penalized 10%.
- Those submitted late by less than 48 hours will be penalized 20% total.
- Those submitted more than 48 hours late will receive a grade of zero except in the case of medical reasons documented properly (see below section). The weight for the assignment will be transferred to the weight for the *final exam*.
- After you have submitted your work, staff may use *turnitin* software to help consider the class's assignment submissions.

What if I miss a quiz or test?

- If the quiz/test is missed for a valid medical reason, you can submit the proper documentation (see below section), and the *final exam* will be worth more.
- Other reasons for missing a quiz/test will require prior approval by your instructor. If prior approval is not received for non-medical reasons then you will receive a grade of zero for that quiz/test.

Submitting medical documentation:

- Submit the University of Toronto Verification of [Student Illness or Injury form](#) to your instructor within one week of the quiz/test/assignment deadline. (Come to any lecture or office hours.)
- The form will only be accepted as valid if the original paper form is submitted, filled out according to the instructions on the form.
- The form must indicate that the degree of incapacitation on academic functioning is moderate, serious, or severe in order to be considered a valid medical reason for missing the term test. If the form indicates that the degree of incapacitation on academic functioning is negligible or mild then this will *not* be considered a valid medical reason.

Computing and Calculators

For the quizzes, test and exam, you will need a basic scientific hand-calculator, with statistical functions, logarithmic functions etc. Calculators on phones or other devices equipped to communicate with the outside world (for example, through the internet or cellular or satellite phone networks) will not be permitted during the term test and the final exam. Programmable calculators are not allowed on the quizzes, midterm or exam.

We will make extensive use of the R language (and RStudio) to analyse data. The main advantages of R are that it is freeware and that there is a lot of help available online. R is freely available for download at cran.r-project.org for Windows, Mac, and Linux operating systems. *RStudio* is a good integrated development environment to R. It is freely available at www.rstudio.com/products/rstudio You may also like to sign up for a CQUEST account. To get an account and find out more information about using CQUEST go to www.cquest.utoronto.ca

If you need help installing R and RStudio, and learning the basic syntax of R, a helpful document is [here](#). Another good online R reference is [here](#), and there is a downloadable book called *Introduction to R* by William Venables and David Smith. The alternative reference textbooks above, and websites such as datacamp.com, are also good to help you pick up R.

I'll give you a quick introduction to R and provide you with R source code for the examples in lecture. Note that there are many graphics options available to produce plots, but we'll focus on the basics, sufficient for the assignments. You may like to use *Rmarkdown* to type up your work.

Online Discussion Board

You will have the option to use Piazza for class discussion. If you decide not to use Piazza it will not disadvantage you in any way, and will not affect official University outcomes (e.g., grades and learning opportunities). If you choose not to opt-into Piazza then you can ask questions or discuss course material with the instructor or TAs during office hours.

Please read Piazza's [Privacy Policy](#) and [Terms of Use](#), taking time to understand and be comfortable with them. They provide for substantial sharing and disclosure of your personal information held by Piazza, which affects your privacy. If you decide to participate in Piazza, only provide content that you are comfortable sharing under the terms of the Privacy Policy and Terms of Use.

The Piazza system is highly catered to getting you help quickly and efficiently from classmates, the TA, and the lecturers. Rather than emailing questions to the teaching staff, we encourage you to post your questions on Piazza. To sign up for the discussion forum, click on the link: piazza.com/utoronto.ca/fall2018/sta3021001

TAs are assigned to answer questions you have on Piazza. If you post your questions there and don't get a response in three days, please inform me as soon as possible.

Additional help

Practice problems from the textbook for your home preparation will be posted on the web-site. They are not to be handed in.

Need extra help with the coursework? Here are some options:

- For continued class discussion and questions outside of class, try posting on the discussion forums. The instructor and TAs will be monitoring the
- You may choose to join (or create) an STA302 study group
- You can visit the instructor or teaching assistants during their office hours
- E-mail should only be used for emergencies or personal matters

Questions concerning the **assignments** won't generally be addressed in office hours.

How to communicate with your instructor

Questions about course material such as:

- How do I do this question in the textbook?
- What is standard deviation?
- When is the midterm?

can be posted on the Piazza discussion forum. If you are shy, questions can be posted anonymously (so that the author is anonymous to other students but not to the instructors).

For private communication, such as "I missed the test because I was ill," e-mail me. Use your utoronto.ca e-mail account and include your full name and student number.

You may post entirely anonymous feedback [here](#). Nobody will know who you are. In past courses this has helped sort out small problems and big ones.

Academic integrity

You are responsible for knowing the content of the University of Toronto's Code of Behaviour on Academic Matters at

www.governingcouncil.utoronto.ca/policies/behaveac.htm. If you have any questions about what is or is not permitted in this course, please do not hesitate to contact your instructor.

Accessibility needs

The University of Toronto is committed to accessibility. If you require accommodations for a disability, or have any accessibility concerns about the course, the classroom, or course materials, please contact Accessibility Services as soon as possible: accessibility.services@utoronto.ca or <http://accessibility.utoronto.ca>.

Your responsibilities

The classroom sessions for this class are designed to actively engage you in the course material. We hope you'll find them interesting, challenging, fun, and an excellent opportunity to truly learn the material.