Instructor: Karen H. Wong - sta248@utoronto.ca

Office Hours: Will be posted on Quercus, held on Bb Collaborate/MS Teams

Course Description: A survey of statistical methodology with emphasis on data analysis and applications. The topics covered include descriptive statistics, data collection and the design of experiments, univariate and multivariate design, tests of significance and confidence intervals, power, multiple regression and the analysis of variance, and count data. Students learn to use a statistical computer package as part of the course (Note: STA248H1 does not count as a distribution requirement course).

Prerequisites: STA237H1/STA247H1/STA257H1/STA256H5/STAB52H3; CSC148H1/ CSCA48H3/CSC148H5
Prerequisites are strictly enforced by the department and prerequisite waivers are not accepted.

Textbooks:

1. *A Modern Introduction to Probability* by *Dekking et. al.*, available through the library here

2. *Modern Mathematical Statistics with Applications*, 2nd ed. by *Devore and Berk* available through the library here

3. *R for Data Science* by *Grolemund and Wickham* is a recommended reference text for students new to using R. Chapters 3-9 is recommended reading to reinforce skills in R. Available here

Course Structure: During these unprecedented and challenging times, this course will operate differently from the regular sitting. Our course will be a combination of synchronous and asynchronous lectures with irregular tutorial activities. Synchronous lectures will be recorded and made available on our Quercus page, but tutorials are self-paced individual and group work, and are **mandatory**. Note that all course times listed are in Toronto time!

Suggested Weekly Routine

Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	Monday
Post-Sync: Attempt 2-3 practice questions while the memory is fresh! Video lectures posted. Watch at your own pace, take your own notes. Post your questions and answers on discussion board. Work on tutorial, if applicable.		*Attempt 5-10 suggested problems each day. Post your questions to the discussion board. Finish up tutorial work.*		*REST DAY or work on group tutorial, if applicable.*	*Review notes from the week's videos to prepare for sync session. Rework any R material/activities done in class. Start on tutorial, if applicable.*	

Everyone learns at a different pace. In addition to our course times ($\sim$ 3-4 hours per week), students should expect to commit *at least an additional 3 hours* as self-study, or 6-7 hours per week. Self-study includes: creating course notes for yourself, working on suggested problems, reviewing class material (including R analyses), posting on the discussion board, and completing tutorial work.

Quercus Page: All lecture slides, video recordings, quizzes/tests/assignments, and other course material will be posted on Quercus under Modules. Course materials provided on Quercus are for the use of students *currently enrolled in this course only*. **Distributing course materials to anyone outside of the course is considered unauthorized use.**

Grading:

Weekly Discussion	12%	Throughout Winter Term
Syllabus Quiz	2%	Jan. 28
Tutorial Activities	7% each	(Best 3 of 4): Weeks of Jan. 25, Feb. 8, Mar. 15, Apr. 5
Quizzes	9% each	(3): Feb. 4, Mar. 11, Apr. 1
Midterm	15%	March 4: 10 AM - 12 PM EST
Final Assessment	23%	TBD

Discussion Board: There will be weekly topic threads on the Quercus Discussion Board as a space for students to discuss textbook problems, course topics, and exchange peer support. The instructor and TAs will be monitoring and contributing regularly to keep the discussions going, guide students back on track, or answer any lingering questions. The expectation is that students contribute by working collaboratively with each other and build on your understanding of concepts by answering each others' questions. The grading scheme will be posted and discussed during our first class.

Tutorials: Tutorials will be a combination of small progress 'check-in' problems and activities explore concepts through using R. The tutorials should take no longer than 3 hours, depending on individual progress and fluency with course material. Tutorials will either be individual or group learning activities. In the latter case, you will complete an *individual* pre-tutorial in the week leading up to the tutorial, before you collaborate over the weekend with at least two other classmates to complete and submit the group component of the tutorial.

   Tutorial Weeks: Jan. 25 (individual), Feb. 8 (group), Mar. 15 (TBD), Apr. 5 (TBD)

Quizzes: Take place during our Thursday scheduled synchronous sessions on Quercus from 10 AM - 11:05 AM. Quizzes will be timed and question-locked, but are open course-notes and course textbooks. The format will be a mix of multiple choice/true-false and short written responses with an emphasis on conceptual understanding of topics. If file uploads are required, an extra 15 minutes will be allotted (i.e. until 11:20 AM) for uploading. You will always be notified ahead of the quiz if you should expect to upload any work.

Midterm & Final Assessment: Assessment information will be available on Quercus approximately 1 week prior to the date of the assessment. The midterm will be a <u>timed</u> assessment on **Thursday Mar. 4, 10 AM - 12 PM EST**. A *mirrored time slot* on the same day for 10 PM - 12 AM will also be offered if required. A survey will be available a week before for you to register for your time slot.

Any time conflicts with the midterm must be communicated as soon as you are aware, and <u>**no later than 2 weeks prior to the midterm**</u>. No accommodations will be offered for late notice of time conflicts.

Minimum Passing Requirement: Students must complete and submit at minimum:

- 2 of the 4 mandatory tutorials

- 2 of the 3 quizzes

- Both the midterm and the final assessment.

If you miss more than these minimum passing requirements, even with accommodation, you will not be able to pass this course. There are no make-up assessments except for the midterm. If there are extenuating circumstances that will affect your performance in the course in the long term, it is your responsibility to contact your college registrar and the instructor as early as possible.

Accommodations: Accommodation requests for mini-homework/individual tutorials must be made by email to the course instructor (sta248@utoronto.ca) **at least 24 hours before** the assignment due date. Extensions will be arranged where possible. There are no accommodations for tutorials due to the self-paced time frame and the group component.

Missed Work: Missed term work (tutorials and quizzes up to a maximum of 2, and midterm) due to legitimate reasons (e.g., injury, illness, or other exceptional circumstances) should be communicated to the instructor <u>within 1 week</u> following the assessment **and** have the absence declaration form on ACORN completed and attached to the email. For documented missed work:

- Missed tutorial work will be redistributed among the tutorials and midterm equally.

- Missed quizzes will be redistributed proportionally among the quizzes

- Missed midterm: A make-up midterm will be scheduled after March 11 for students who submitted documentation.

Grading: Marking schemes/solutions will be provided for all assessments except the final assessment. These should be reviewed before requesting a reread. If you still have concerns about your grading, send me an email with a brief explanation **no later than 1 week** after the assessments have been released back to you. Late requests will not be accepted.

Course Conduct:

- **Email:** Any administrative or personal concerns regarding the course should be addressed to the course instructor (sta248@utoronto.ca). Questions regarding course material and concepts should be left for office hours or the discussion board. I will usually respond within 48 hours but may take up to 5 business days.

- **Programming Languages:** We will be using R Statistical Software regularly, R statistical software which is available for free download. An installation guide will be provided in Week 1.

  - If you have problems with setting up R, you may find it easier to use through R Studio Cloud which offers free 15 project hours per month.
  - You can also access R Studio through JupyterHub (Select 'R Studio', or if that's not readily available, click 'New', and 'R Studio' on drop down menu).
  - Instructions on using R as well as demos will be done regularly during our synchronous sessions, and starter code will be provided where appropriate.

  By the end of the course, you are expected to be able to run simple simulations on your own, write your own functions, and more generally, understand how various code works, apply and adapt them to perform statistical analyses, and predict the output that will be generated. Check the weekly agenda to see if you will need to have R ready to use during the synchronous lecture times.

Accessibility Services: The University of Toronto provides accommodations through accessibility services to students with diverse learning styles and needs. If you have a disability or health consideration that may require accommodations, please feel free to reach out to Accessibility Services at 416-978-8060 or through accessibility.utoronto.ca.

Academic Integrity: 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 Matter. Potential offences include, but are not limited to:

- Obtaining or providing unauthorized assistance on any independent assessment including:

  - working in groups on individual assessments, including giving hints to the answer!
  - having someone rewrite, edit, or add material to your independent work
  - researching for inspiration, hints, or answers to any graded problem
  - posting active assessment questions on discussion boards/private tutoring companies for hints/solutions
  - sharing your code for graded assessments, or letting another classmate look over your code for 'ideas/inspiration'

- Lending your work to a classmate who submits it as their own with or without your permission

The University of Toronto treats cases of academic misconduct very seriously. All suspected cases of academic dishonesty will be investigated following the procedures outlined in the Code. The consequences for academic misconduct can be severe, including a failure in the course and a notation on your transcript. **If you have any questions about what is or is not permitted in this course, please do not hesitate to contact the instructor.** If you are experiencing personal challenges that are having an impact on your academic work, please speak to the instructor or seek the advice of your college registrar.

Tentative Schedule:

Week Of	Topics
1: Jan. 12	Introduction to the Course, EDA: Summary Statistics
2: Jan. 19	EDA: Graphical Summaries, Sampling Methods
3: Jan. 26	Approximations and Limiting Theorems, Statistical Model, Estimators and their Properties
4: Feb. 2	Estimators and their Properties, MOM Estimation
5: Feb. 9	MLE and MLE Estimator Properties, Intro to Bayesian Model
6: Feb. 16	WINTER READING BREAK
7: Feb. 23	Bayesian Statistics and Estimation
8: Mar. 2	MIDTERM, Bootstrap Principle (Why Bootstrapping?)
9: Mar. 9	Parametric Bootstrap, Common Sampling Distributions
10: Mar. 16	Confidence Intervals & Bootstrapped CIs
11: Mar. 23	Introduction fo Hypothesis Testing, Testing Errors
12: Mar. 30	Simulated P-Values, Count Data & Goodness of Fit Tests
13: Apr. 6	Simple Linear Regression