<div align="center">

**University of Toronto**
**Department of Statistical Sciences**
**STA220H1: Practice of Statistics I**
**Tuesdays and Thursdays, 7:10p.m.–10:00p.m., ES1050**
**Summer 2019**

</div>

**Instructor:** Asal Aslemand, PhD
**Office Hours and Location:** Fridays, 11:00a.m.–1:00p.m., Stewart Building in Room 103A
**E-mail:** Click here to Email Asal or use asal.aslemand@utoronto.ca

> **Note:** Please use your U of T email address when sending me an e-mail.
> In the subject line of your email, please write STA220.

**Course Website:** Quercus

**Teaching Assistant (T.A.) and Tutorial Information:**

- Please note that tutorials start on Thursday, May 9th.

- There is a tutorial session before each lecture (except for the first lecture and on the day of the test).

- There are six tutorial sections in this course and all are scheduled from 6:10p.m. to 7:00p.m.

- Students are assigned to a tutorial section by their last names alphabetically.

| Student Lastname Begining with Letter(s) | Tutorial Room | Teaching Assistant TA | TA Office Hour in Room SS623B |
|---|---|---|---|
| A to C | SS2125 | Jessica Chau | Tuesdays, 5:00-6:00p.m. |
| D to J | SS1074 | Hana Dampf | Thursdays, 5:00-6:00p.m. |
| K to Lin | SS1070 | Irwin Khuu | Tuesdays, 5:00-6:00p.m. |
| Liu to Q | SS1072 | Michal Malyska | Thursdays, 5:00-6:00p.m. |
| R to Wang | WI523 | Henry Lu | Thursdays, 5:00-6:00p.m. |
| Wei to Z | MP134 | Lu Yu | Tuesdays, 5:00-6:00p.m. |

<div align="center">

**Welcome to STA220**

</div>

Data are everywhere. Statistic is science of data. Statistical science involves collecting, classifying, analysing, presenting, interpreting and communicating information - often messy information that requires cleaning and organizing. In everyday life, we need to use quantitative data to help us understand and make sense of the world and choose how to act. Statistical methods help us separate sense from nonsense. That is, statistics helps us be objective.

## Course Description:

This course will provide an intuitive introduction to fundamental statistical concepts and reasoning. The course will cover effective graphical and numerical displays, estimating and describing the natural variability in data, methods of data collection, and the key ideas in how statistical tests can be used to separate significant differences from those that are only a reflection of the natural variability in data. In this course, we will draw from real-life examples to explicate statistical ideas.

## Course Topics:

The first half of the course introduces descriptive statistics, contingency tables, Normal model, correlation and linear regression, methods of data collection (e.g., sample surveys, experiments and observational studies), probability rules, and random variables. The second half of the course introduces inferential statistical methods. These topics include sampling distribution models, point and interval estimations, test of hypotheses for proportions ($z$-test), and inferences about means ($t$-test).

## Course Textbook:
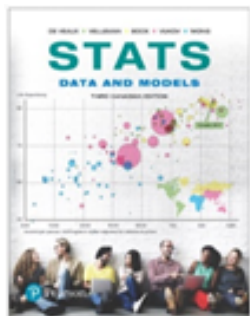
De Veaux, R. D., Velleman, P. F., Bock, D. E., Vukov, A. M. , & Wong, A. (2018). *Stats: Data and Models.* Third Canadian Edition. Pearson Canada Publication.

## Statistical Software:

We use R & RStudio for performing statistical analysis.

- You need to first download and install R (a free statistical package) from: https://www.r-project.org/.

- After installing R, download and install RStudio (its free version) from: https://www.rstudio.com/

- I will provide text and video instructions regarding how to download and install R and RStudio on your computer and any course activity that you need to use R for completing it. The instructions will be posted on our Quercus page for the course.

## Calculator:

You need to use a basic (non-programmable) calculator for tutorials, midterm test, and final exam. Your simple calculator should allow you to add, subtract, multiply, divide, and take a square root.

**Please Note:** Phone calculators or any other communicating devices are **NOT** allowed during the midterm test, and final exam.

## Course Components:

- ### Class Lectures:

  Each lecture covers important course materials by guiding you through examples to develop your statistical literacy, reasoning and thinking. Every lecture builds on the materials that we discussed in the previous one. This means that statistical ideas are connected and have a logical flow. I highly recommended you attend the lectures regularly for deeper understanding of and making connections among statistical ideas.

- ### Weekly Tutorials:

  Tutorials start on Thursday May 9<sup>th</sup>. There are 10 tutorials scheduled in this course. Prior to each tutorial, I will post on our Quercus page for the course, in "Tutorial Preparation" module, the information regarding your upcoming tutorial. These include a set of practice problems, which are related to the previous lecture's contents. Some of the assigned problems require that you use RStudio to explore and analyze data. You need to bring the printouts of your statistical analyses, which we refer to them as R outputs, to your tutorial and will be required to hand them in at the end of your tutorial to your TA. I will provide you with the necessary instructions – both in text and video formats – that you need for preparing for each of the tutorial's statistical output activity.

  In each tutorial you will receive a worksheet, which is based on your tutorial preparation. You need to refer to your R outputs to answer questions on your worksheets. You may work in small-groups with your peers and your TA will facilitate group discussion.

  **Please note that, you are expected to complete <u>only your own worksheet</u> during each tutorial session**. You may, however, discuss statistical ideas with your peers in class and may also refer to your course notes, and the course textbook during the tutorials.

  The completion of each tutorial worksheet will account for your tutorial mark. Your lowest tutorial mark (one out of the ten worksheets) will be dropped. Each worksheet is marked out of 10 points. Your TA will mark your tutorial worksheet and will return it to you at the next tutorial. Please keep all your marked worksheets for your record in case of any discrepancy of mark-input on your Quercus's grade page.

- ### Data Analysis Project:

  In this course, there is a data analysis project based on a realistic data set. The objective of this course project is to give you the opportunity to use R software and explore some of the statistical techniques that you learn in making sense of the given data set. Furthermore, you can work (it is optional) with two other students in this course.

  The data set is retrieved from the Organization for Economic Co-operation and Development (OECD) website. This OECD data set is cleaned and is posted on our Quercus page for the course in a Module named " Data Analysis Project". In that module, the instruction on how to analyze this OECD data set is available. The instruction-document includes a set of questions that requires you to answer. A template for inserting your answers is created and is made available in the same module. Please use the template to answer the related questions.

This course project enables you to make informal statistical inferences. You will obtain, present and interpret descriptive statistics for univariate and bivariate data, and investigate relationships between variables in the data. On the due date of this course project, you will submit two PDF files on Assignments' Quercus page. One file is your answers to the related data analysis questions using the provided template on Quercus. The second file is your R outputs. If you worked in a group, only one student from your group will submit your group's data analysis with its R outputs with all group member's first and last names included on the top of the completed template's page.

   ○ **Late Submission Policy**:

   No late submission will be accepted after one week from the due date of this course project. Please note that you (and all your group members) will be deducted 5 points out of the 50 points allocated to this course project (i.e., 0.5% of your final course grade) for each day (not including the weekends) that your submission is late. If you experience an unseen circumstance preventing you complete the course component in time, please email Asal no later than one week from the due date to seek an alternative arrangement that is within the University of Toronto's Academic Policy and Code of Conduct.

● **Midterm Test and Final Exam:**

A one-hour and thirty minutes midterm test is scheduled on Thursday May 30th and starts at 7:15p.m. The midterm test covers topics from Week 1 to 6 of the course. It will consist of True or False, multiple-choice, and short-answer questions. A three-hour final exam will be given during June exam period, which covers the entire course. The format of final exam is the same as the midterm test.

You are allowed to bring one sheet of handwritten letter-sized $(8.5'' \times 11'')$ using only one side of the sheet for the midterm test but both sides for the final exam. You need to have a calculator for both the midterm test and final exam. Statistical tables that you may need will be provided.

**Course Assessments:**

| Type | Number of Occurrence | Due Dates | Weight |
|---|---|---|---|
| Tutorial Worksheets | 10 (Counting best 9 of 10) | Before Lectures 6:10p.m.–7:00p.m. Starting Thursday, May 9th No tutorials on the day of the test | 15% |
| Data Analysis Project | 1 | June 9th (11:59p.m. on Quercus) | 5% |
| Midterm Test[1] | 1 | Thursday, May 30th Midterm test starts at 7:15p.m. | 30% |
| Final Exam[2] | 1 | To Be Announced | 50% |
| | | Total | 100% |

[1] The location for our midterm test will be announced later but before the test on our Quercus page for the course.

[2] The Final Exam is scheduled by Faculty of Arts and Science during June 2019 examination period.

## Missed Tutorials, and/or Midterm Test:

- There are no makeup tutorials or makeup test.

- If you miss a tutorial, it will become your lowest tutorial worksheet mark (mark of zero), but it will be the one you drop.

- If you miss more than one tutorial or if you miss the midterm test with a valid reason in order to be excused, you must submit to me (Asal) an appropriate documentation (e.g. U of T medical certificate) within one calendar week of the missed tutorial and/or missed midterm test. U of T medical certificate is available at:
  http://www.illnessverification.utoronto.ca/index.php.
  Please, print on it your name, student number, course number, and the date.

- If a tutorial is missed for a valid reason, I will still drop your lowest tutorial worksheet mark among the remaining ones and will calculate your average mark accordingly.

- If the midterm test is missed for a valid reason, its weight will be shifted to the final exam. In that case, your final exam will account for 80% of your total course grade.

- Note that I will put the **code: "-1"** for an excused missed assessment (e.g., worksheet, midterm test) in your grade's page on Quercus. This code means that at the end of the term, I will omit this excused (missed) assessment from your course grade calculation.

## Course Incentives:

- **Arts on Statistics – Thinking Beyond Numbers:**

  Based on the Data Analysis Project, you (individually or with your group-members) are welcome to make a visual or vocal presentation (e.g., 3-D model, video, painting, drawing, music, song, poem, infographics, Prezi, pamphlet, concept map, mind map) that can convey the statistical information of your data analysis. Ssubmit your creative work by Tuesday, June 11$^{\text{th}}$ to receive 1% bonus mark added to your course grade.

  I invite you to think of this opportunity as "thinking beyond numbers". Some creative works of former students are posted in the "Course Incentive" module. A guideline with tips regarding how to think creatively is posted in the same module. If your creative work is:

  ○ like any of these forms: 3-D model, painting, drawing, concept map, or mind map, please bring it with you to our class on Tuesday, June 11$^{\text{th}}$. Also, please take a picture of your creative work. Print your picture in COLOUR, and state on it your first and last name and student number and submit it to me (Asal) on the same day.

  ○ in a form of an audio or video format, email your link to Asal on Tuesday, June 11$^{\text{th}}$.

  ○ a pamphlet or digital poster, print it in colour and submit it in class on Tuesday, June 11$^{\text{th}}$.

  **Note: All emails should have a subject line of "STA220 Creative Work".**

## Tentative Weekly Lecture Coverage:

This is our tentative weekly lecture coverage. Please note that modifications to this plan may be made based on the rate at which we are able to cover the material.

| Lecture | Chapter |
|---|---|
| 1 | Ch.1. Statistics Starts Here |
| | Ch.2. Displaying and Describing Categorical Data |
| 2 | Ch.3. Displaying and Summarizing Quantitative Data |
| | Ch.4. Understanding and Comparing Distributions |
| 3 | Ch.5. The Standard Deviation as a Ruler and the Normal Model |
| 4 | Ch.6. Scatterplot, Association, and Correlation |
| | Ch.7. Linear Regression |
| 5 | Ch.7. Linear Regression |
| | Ch.8. Regression Wisdom |
| 6 | Ch.9. Sample Surveys |
| | Ch.10. Experiments and Observational Studies |
| 7 | Ch.11. From Randomness to Probability |
| | Ch.12. Probability Rules! |
| 8 | Ch.13. Random Variables |
| 9 | Ch.14. Sampling Distribution Models |
| | Ch.15. Confidence Intervals for Proportions |
| 10 | Ch.16. Testing Hypotheses About Proportions |
| | Ch.17. More About Tests |
| 11 | Ch.18. Inference About Means |
| 12 | Ch.19. Comparing Means |
| | Ch.20. Paired Samples and Blocks |
| | Ch.21. Comparing Two Proportions |

## Course Learning Outcomes:

Upon completing all course components, you will be able to:

- Understand the nature of statistical data;
- Identify type of variables (e.g., categorical) and investigate relationships between them;
- Construct, present, and interpret graphical displays, summary statistics, confidence intervals, and significance tests (e.g., $p$-values);
- Know for what objectives and under what conditions a statistical method is applicable;
- Verify that data satisfy the conditions required for conducting a statistical analysis and the research objectives are compatible with the objectives of the statistical methods being used;
- Identify and investigate the impacts of outliers in a data;
- Realize common misinterpretations and know how to avoid such interpretations.

## Accessibility Centre:

Students with diverse learning styles and needs are welcome in this course. In particular, if you have a disability/health consideration that may require accommodations, please feel free to approach me and/or contact the Accessibility Services Office as soon as possible at accessibility.services@utoronto.ca or visit http://www.accessibility.utoronto.ca.
I will work with you and Accessibility Services to ensure you can achieve your learning goals in this course. Enquiries are confidential.

## Academic Integrity Statement:

Academic integrity is essential to the pursuit of learning and scholarship in a university, and to ensuring that a degree from the University of Toronto is a strong signal of each student's individual academic achievement. As a result, University of Toronto treats cases of cheating and plagiarism very seriously. The University of Toronto's Code of Behaviour on Academic Matters outlines the behaviours that constitute academic dishonesty and the processes for addressing academic offences. Please refer to the following related links regrading:

- Code of Behaviour on Academic Matters:
  http://www.governingcouncil.utoronto.ca/policies/behaveac.htm

- Academic Integrity:
  http://academicintegrity.utoronto.ca/

## Final Note:

In this course, no previous knowledge in statistics or using statistical software are required. Also, no advanced mathematical training is necessary; however, I appreciate your *effort* and *willingness to work hard* on statistical concepts that we will cover in this course. Try to stay on tasks and attend all lectures, tutorials, and visit office hours. Research in education has shown that one's motivation to achieve is determined by their extent in which they *value* the reward of *succeeding at a task* (e.g., learning statistics will better enable me understand research in my field of study), and their *growth mind set* [believing that by working hard on doing a task (e.g., statistical problems) abilities (e.g., ability in doing statistics) can be developed]. Thus, the best learning strategy in this or any course is to work hard regularly throughout the term and realize the relevance of learning the subject matter.

Keep in mind that learning quantitative methods is not just about doing maths and calculations, however, it is about understanding concepts. Make sure that you know why you are doing a statistical calculation and what the results will tell you; when in doubt seek help from me (Asal) and/or your Teaching Assistants. I wish you all the best in this course.