Statistical Sciences
UNIVERSITY OF TORONTO

STA 220: THE PRACTICE OF STATISTICS I

Fode Tounkara

Summer 2021

Meeting Information

- **Instructor:**   Dr. Fode Tounkara
- **E-mail:**   f.tounkara@utoronto.ca.
- **Lecture:**   Tuesday, Thursday; 6 PM-9 PM.
- **Location:**   online-Blackboard Collaborate (Bb Collaborate) in Quercus.
- **Office Hours:**   Wednesday 10:00-11:00 AM, Online-Blackboard Collaborate.
- **TA Office Hours:**   TBD

How and when will the course operate?

The course will be offered entirely online this summer, with weekly synchronous meetings held in Bb Collaborate. Save the monthly course schedules found on Quercus (q.utoronto.ca) to stay on top of the weekly tasks! All listed times are in local Toronto time. If you are in a different time zone, you are responsible for any time conversions. It is recommended that you schedule important dates and times into your calendar at the beginning of term with reminders to avoid missing anything!

Course Overview

Course Description

This course will provide an intuitive introduction to fundamental statistical concepts and reasoning. The course will cover: methods of data collection; constructing effective graphical and numerical displays; estimating and describing the natural variability in data; 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.

Course Objectives

The learning objectives of the course are:

- Understand the ideas, principles, and considerations that are common to all statistical methods,

- Develop a statistical toolbox of some methods for the collection, analysis, and display of data,

- Identify appropriate uses of the statistical methods, including their strengths and limitations, and

- Develop statistical literacy, including the ability to recognize the importance of data in decision-making and understand the social and scholarly applications of statistics.

Topics to be covered

| Topic	Description
A first look at data	Summary statistics and graphical displays for a single categorical or quantitative variable and for relationships between two variables.
Collecting Data	Sampling, Observational studies and experiments. The effect of confounding and concluding causation.
Probability	Probability models, Bayes' theorem, the normal distribution, the Law of Large Numbers, the Central Limit Theorem, sampling distributions.
Confidence Intervals	Confidence intervals for proportions and means.
Statistical Tests	Tests of significance for proportions and means.
Two Samples	Tests of significance and confidence intervals for proportions and means in the two-sample case.
Linear Regression	Method of least squares, evaluating model fit, the effects of outliers and influential observations.

Textbook

There is no required course textbook. All course material is contained in the video lectures and their corresponding notes. Two recommended textbooks are:

1. OpenIntro Statistics 4th Ed. Diez, D. Barr, C. D., and Cetinkaya- Rundel Mine.

https://www.openintro.org/book/os/

OpenIntro Statistics (4th edition) is free and available to download. You can also choose to pay what you can for the textbook, purchase a physical copy on Amazon for $20 or $60. This is an excellent textbook that is less conversational compared to the Velleman et al. textbook. However, the concepts are clearly explained. A nice feature of the text and website is that many of the examples and vignettes used to illustrate the concepts are based on real applications of statistics.

2. Stats: Data and Models, third Canadian edition, by Richard D. De Veaux, Paul F. Velleman, David E. Bock, Augustin M. Vukov, and Augustine C.M. Wong. 3rd ed.
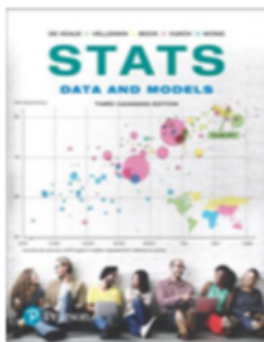
The digital textbook is available for purchases **HERE**. It is extremely easy to read and is written in a conversational style. Most of the concepts are clearly explained and there are lots of fun and interesting vignettes that illustrate statistical concepts.

Your responsibilities

The course is designed to actively engage you in the course material. We hope you'll find the subject of statistics interesting, challenging, and fun, and an excellent opportunity to truly learn the material. Each week, you will be introduced to the course material through video lectures, with fine-tuning and learning support in your synchronous meetings in the following week. In order for the in-class sessions to be effective, preparing by learning about the week's concepts through the video lectures and notes is essential. Make use of the posted weekly schedules and the suggested study routine on Quercus, see the course outline.

Computing

Calculators

You will need a calculator. A scientific calculator is recommended for ease of use, but any calculator that can perform basic operations, and take powers and logarithms is sufficient.

Software

The course includes a lot of numerical calculation. We will use **Statistical programming language and environment R** software to support the class. **Using R and understanding of R outputs are required**, on the level explained in the class.

R is freely available for download at

- **https://www.r-project.org**

Rstudio is a good integrated development environment to **R**. It is strongly recommended that you also download R Studio to accompany R for a nicer user interface. It is also freely available at:

- **https://www.rstudio.com**

If you'd prefer to use a browser version of RStudio, please use the U of T JupyterHub (**https://jupyter.utoronto.ca/hub/login** with RStudio selected). You can access this platform with your U of T credentials.

Please find below a good online reference to get started with R.

- **Get started with R**

For your assignments and final project, it is recommended to use **Rmarkdown** to write your solution. R Markdown is a file format for making dynamic documents with R. Please find below a cheat sheet that may help to get started with R Markdown.

- **R Markdown**

Discussion Forum

We will be using Piazza for discussions surrounding course content. This will be a space for you to post your questions (general, textbook, practice), your answers (what you tried, where you are getting stuck), or suggestions to your peers. Piazza will be monitored regularly by instructors and teaching assistants. You can join Piazza for our course <u>**HERE**</u>. Be sure to familiarize yourself with their **privacy policy** before opting in to use!

Additional Help

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

- Post your question on the class discussion forum on Piazza. The instructor and TAs will monitor this discussion forum regularly.

- The instructor and TAs will have weekly office hours on Bb Collaborate.

- E-mail should only be used for emergencies or personal matters. If you email course questions to the instructor or TA, you will be asked to post your question on the discussion forum. Individual questions about the course content will not be answered via email.

Evaluation

Mark Breakdown

- Quizzes (best 5 of 6× 5 %) worth 25%
- Assignments (2 × 12.5 %) worth 25%
- Midterm worth 20 %
- Final Project worth 30 %

Assessment	Weight	Due data	Time	Location
Weekly Quiz (best 5 of 6)	5× 5 % each	Saturday	11:59 pm	Submit online (Quercus)
Assignment 1	12.5 %	May 15	11:59 pm	Submit online (Crowdmark)
Term Test	20 %	May 27	7:30-8:40 PM	Crowdmark
Assignment 2	12.5 %	June 5	11:59 pm	Submit online (Crowdmark)
Final Projet	30 %	June 19	11:59 pm	Submit online (Crowdmark)

Weekly Quizzes

You will be given an online quiz on the topics covered in the weekly modules. These quizzes should be completed after watching the videos and reading the notes posted each week. You have four attempts for each quiz, and the highest grade will be recorded.

- The quizzes will be held on Quercus and are always due on Saturday at 11:59 PM (23:59) Toronto time. Please see the course schedule for specific due dates.

- The weekly quiz will cover material in that week's module (video lectures and accompanying notes)

- The quiz will consist of multiple choice and true/false questions, randomly chosen from a pool of questions.

- The number of questions will vary from week to week but the quizzes will be equally weighted.

- You will find out your score immediately and you can take the quiz up to four times until the Saturday 11:59 PM deadline.

- Your final quiz score will be the highest score from all of your attempts. Note that you will get a different randomly generated quiz each time. This means that you will not be penalized for taking the quiz again even if you obtain 100% on a previous attempt.

Assignments

- There will be two equally-weighted assignments, each assignment will consist of an individual component to be completed and submitted on your own.

  - A1 : Saturday, May 15th, due 23:59

  - A2 : Saturday, June 5th, due 23:59

- The assignments will each be of a practical nature, for which the use of R might be required

- Must be written as PDF files using R Markdown (word or Latex) and submitted online into **Crowdmark** by 11:59 PM on the due dates

- Late submissions will not be allowed beyond the due date

- Email submission will not be accepted.

- **Due to the nature of these assignments, there will be no extensions granted under any circumstances**

Term Test

- **80-minutes multiple-choice and short-questions** test

- Thursday, May 27, from 7:30-8:50 PM

- The term test will be written online (Crowdmark).

- 10 minutes to submit your answers on Crowdmark

- The information on coverage will be posted on Quercus in advance.

Final Project

- There will be a final project, which consists of an individual component to be completed and submitted on your own.

- It will be of a practical nature, for which the use of R might be required

- Must be written as PDF files using R Markdown (word or Latex) and submitted online into **Crowdmark** by 11:59 PM on June 19

- Late submissions will not be allowed beyond the due date

- Email submission will not be accepted.

Missed assessments

- **There will be no accommodations made for the weekly quizzes beyond a best 5 of 6.**

- If a midterm or an assignment is missed for a valid medical reason, you must fill out the absence declaration form on ACORN and inform your instructor within **three days following the test**. Requests after this deadline will not be accepted.

- If you miss a test without valid documentation, then a grade of zero will be assigned for the test.

- **An alternative assessment will be arranged at the instructor's discretion** if accommodation is granted for a missed term test or missid assignment. Note that this alternative assessment may have a different format (e.g., oral assessment) and may be scheduled in June after classes end.

- Other reasons for missing a 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 the missed midterm test.

Minimum Passing Requirement:

Students must complete and submit at minimum:

- Half (3) of the weekly quizzes
- At least one Assignment
- The term test
- The final project

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 quizzes. If there are extenuating circumstances that will affect your performance in the course in the long term, you will be responsible for contacting and informing your instructor as early as possible.

Course outline

This lecture outline is tentative and subject to change as the semester progresses.

	DATES	MODULES	IMPORTANT REMINDERS
Week 1	May 3-7	Introduction to the course, Summarizing: one variable, Relationship between two variables.	May 3, class starts **Quiz 1 due on May 8 at 11:59 PM**
Week 2	May 10-14	Probability: Events, Probability: Random Variables.	**Quiz 2 and A1 due on May 15 at 11:59 PM**
Week 3	May 17-21	Part-1: Sampling Distributions, Part-2: Data Collection.	**Quiz 3 due on May 22 at 11:59 PM**
Week 4	May 24-28	Part-1: Confidence Intervals Part 1, Part-2: Confidence Intervals Part 2.	**Midterm on May 27 at 7:30 PM** **Quiz 4 due on May 29 at 11:59 PM**
Week 5	May 31- June 4	Part-1: The Process of Statistical Tests, Part-2: The Effective Use of Statistical Tests.	June 1-Last day to drop **Quiz 5 and A2 due on June 5 at 11:59 PM**
Week 6	June 7-11	Part-1: Comparing Two Groups, Part-2: Simple Linear Regression.	**Quiz 6 due on June 12 at 11:59 PM**
Week 7	June 14-18	Jun 17-28 – Final assessment period	Monday, June 14-Classes end Tuesday, June 16-Study day **Final Project due on June 19 at 11:59 PM**

Course Policies

- **Online board participation:** Questions about course material should be posted on the discussion forums on Piazza. Review discussion postings before posting your own to avoid redundancy.

- **To communication with your instructor:** Questions on online board discussion can be posted anonymously (so that the author is anonymous to other students but not to the instructors), if desired. If your communication is private, such as, "I missed the test because I was ill", then e-mail your instructor. Use your utoronto.ca e-mail account to ensure that your message doesn't automatically go to a Junk folder and include your full name, student number, and course code (STA220).

- **Deadlines for assignment and final project submissions**
  - Late submissions will not be allowed beyond the due date
  - Email submission will not be accepted.

Institutional Policies and Support

Academic integrity

You are responsible for knowing the content of the University of Toronto's Code of Behaviour on Academic Matters **Here**. 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, or course materials, please contact Accessibility Services as soon as possible: accessibility.services@utoronto.ca or http://accessibility. utoronto.ca.