

**UNIVERSITY OF TORONTO SCARBOROUGH**  
**Department of Computer and Mathematical Sciences**  
**STAB22H3 STATISTICS I Winter 2017**

**Course Description:** Statistics is the science of collecting, organizing and interpreting data. In science, society and everyday life, people use data to help them understand the world and choose how to act, and statistical methods help to separate sense from nonsense.

In this course, we learn about some of the most important techniques used in statistical work. The emphasis of this course is on concepts and techniques and will be useful to students who seek to gain an understanding of the use of statistics in their own field. Our ultimate goal is to gain understanding from data, going from data collection to analysis to conclusions.

Content, emphasis, etc. of the course is defined by means of the lecture material - not only the textbook. It is important to attend all lectures, as there is normally no simple way to make up for missed lectures (perhaps obtain another student's notes). There will also be many lecture examples using statistical software, which students will be using.

**Important announcements, problem sets, additional examples, and other course info will be posted on blackboard. Check it regularly.**

**Instructors**

LEC01: Mahinda Samarakoon

LEC02: Caren Hasler

**E-mail**

caren.hasler@utoronto.ca

mahinda@utsc.utoronto.ca

Note 1: When sending e-mail to an instructor, please use your U of T e-mail address.

Note 2: Depending on e-mail volume, we might not be able to reply to every email received. Dr Ken Butler has prepared an FAQ page which gives answers to some of the most frequent questions that we have received from students. On our course web page I have added a link to this FAQ page. (or just click <http://fisher.utstat.utoronto.ca/~mahinda/stab22/b22faq.html>) Before sending us emails, please check if the answer to your question is on this FAQ page.

**Office:** IC344 (Caren), IC442 (Mahinda)

**Office hours:**

- Caren: Mon and Wed 12:00-13:00 (after class) in the classroom or in IC344
- Mahinda: Wed, Fri 11:00-12:00 in room IC404 (Note: This is right after the lecture on both days. I have noticed that most students like to ask questions in classroom right after the lecture. If our classroom does not have an event right after the lecture, we will be able to have a part (or full) of this office hour in the classroom.)

**Lectures:**

**LEC01:** Wed 10:00-11:00 AC223, Fri 10:00-11:00 in IC130

**LEC02:** Mon, Wed 11:00-12:00 in AA112

**Webpage:** Course information will be posted on blackboard course page.

**Textbook and Software**

- **The text**
  - Stats: Data and Models, DeVeaux, Velleman, Bock, Vukov, Wong  
Second Canadian edition, publ. Pearson Canada
  - The bookstore will have a “custom edition” with only the chapters we need. This should be cheaper than buying the full 2nd edition.
  - Textbook comes with (custom) student solutions manual and the software MyStatLab and StatCrunch
- **Software**
  - We use StatCrunch software (on web).
  - Requires access code (comes with text or can buy separately).
  - Easy to learn
  - I show you what to do.
  - Learning StatCrunch enables you to analyze realistic data.
  - You will need to interpret output from StatCrunch.

## Assessment

Item	Percentage of grade
Quizzes and MyStatLab/webwork assigns	20%
Midterm Test	30%
Final Exam	50%

How do we calculate the grade for Quizzes and MyStatLab/webwork assigns (i.e. for item 1 above)?

For each student, I assign two grades for Quizzes and MyStatLab/webwork assigns as described below and choose the maximum of the two as the grade for Quizzes and MyStatLab/webwork assigns.

- Grade 1: 15% for quizzes + 5% for MyStatLab or webwork assigns
- Grade 2: 20% for quizzes

Usually 20%-25% of the students in this course get A's. Less than 5% of all students who complete the course work fail.

Notes: The above calculation implies that MyStatLab or webwork assigns are not mandatory. You can get the full 20% for component 'Quizzes and MyStatLab/webwork assigns' from only quizzes. These online tools MyStatLab/webwork do have problems and I may not be able to provide solutions for some problems. If you are not comfortable with these tools, please don't use them. Most past students have told me that these tools do have problems, but still they like them and so I'm keeping them for anyone who wants to use them. This of course can help boost your grades for the component 'Quizzes and MyStatLab/webwork assigns' if your quiz grades are low.

## Quizzes

- There will be a quiz in each tutorial, starting from the second week of lectures (i.e the week of January 9th).
- Intended to be straightforward if you are keeping up with material.
- During quizzes, you can refer to your notes and/or the textbook.
- I drop your worst quiz when calculating your quiz grade.
- You **must** write quizzes in the tutorial for which you are registered. If it is impossible to write in your tutorial, you must seek permission from the course coordinator to write in another tutorial for one week only. If you write a quiz in a tutorial other than the one you are registered for, without permission, you get *zero*.

- Your quiz grades should appear on blackboard within one of writing the quiz (i.e. before the next quiz). TAs will distribute the marked quizzes during the tutorial, one week after writing the quiz. Please make sure that you collect all your quizzes and keep the marked quizzes until the term is over. Please check blackboard every week and if your quiz grade doesn't appear, you should check with your TA. If your quiz grade, still doesn't appear on blackboard, you may contact your course instructor. The course instructor will ask for the marked quiz. If you do not have the marked quiz with you, the instructor may not be able to proceed any further.

### **webwork**

- Free web-based problem sets with instant feedback and unlimited trials until the deadline.
- Access with UTorID and password.

### **MyStatLab**

- Similar to webwork. This is bundled with the textbook (only from UTSC bookstore)
- This needs registration. Details on how to register for MyStatLab are posted on blackboard.

### **Test and Exam**

- The midterm test and the final exam are based on multiple choice questions.
- Allowed "cheat sheets": 1 for midterm, 2 for final exam, but no other books/notes. These sheets must be handwritten. You may use both sides of the sheet(s).
- You need a calculator for test, exam and quizzes.

### **Calculators**

Hand calculators are cheap and useful. Any cheap one with a square root and one memory button will do. Mean, standard deviation, sum, and sum of squares keys may save you a bit of time on occasion, but we do not recommend the purchase of expensive calculators to get keys with special statistical calculations. Tests and exams will be designed so that those calculators give no advantage. We emphasize the use of StatCrunch software for doing any tedious or complex calculations. However, **it is important to have a calculator during tutorials/test/exam.**

### Missed Tests and quizzes

There are **no make up tests or quizzes** in this course. If the test is missed for a valid reason, you must submit appropriate documentation (e.g. UTSC medical certificate) to the course instructor **within one calendar week of the missed test or quiz**. UTSC medical certificate is available at [http://www.utsc.utoronto.ca/~registrar/resources/pdf\\_general/UTSCmedicalcertificate.pdf](http://www.utsc.utoronto.ca/~registrar/resources/pdf_general/UTSCmedicalcertificate.pdf)

Print on it your name, student number, course number, and date. **If documentation is not received in time, your test mark will be zero. If a test is missed for a valid reason, its weight will be shifted to the final exam.**

If a quiz is missed due to a valid reason and if the supporting are provided on time, you will be excused from that quiz and your tutorial grade will be calculated from the remaining quizzes only. If you are excused from a quiz, we will indicate that by the code “-1” in place of your grade for that quiz on Blackboard.

### Computing

Students will be using StatCrunch for computing. No previous computing experience is assumed. With this software, you will analyze the data sets in the textbook exercises. The data sets can be found on the publisher’s web site.

### Facilitated Study Groups

- These weekly study sessions are open to everyone in the class.
- Attendance is voluntary, but students who attend regularly often earn higher grades.
- Please be sure to fill out the survey in the first week of class to help ensure the study groups are scheduled at optimal times.
- If you have any questions, please ask your facilitator, or visit the FSG website at <http://ctl.utsc.utoronto.ca/home/fsg>.

### Frequently asked questions

Currently living at:

<http://www.utsc.utoronto.ca/~butler/b22/faq.html>

Check before you ask your instructor. **Blackboard**

- Course announcements, quiz marks etc. will be on Blackboard.
- You are responsible for keeping up with announcements from instructors, course coordinator etc.

- Any problems with recording of marks should be brought to the attention of your TA (quizzes) or course coordinator (exams). Contact the course coordinator if you cannot resolve problems with your TA.

## **ACCESSABILITY STATEMENT**

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 the AccessAbility Services Office as soon as possible. I will work with you and AccessAbility Services to ensure you can achieve your learning goals in this course. Enquiries are confidential. The UTSC AccessAbility Services staff (located in S302) are available by appointment to assess specific needs, provide referrals and arrange appropriate accommodations (416) 287-7560 or [ability@utsc.utoronto.ca](mailto:ability@utsc.utoronto.ca).

## **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, the University treats cases of cheating and plagiarism very seriously. The University of Toronto's Code of Behaviour on Academic Matters (<http://www.governingcouncil.utoronto.ca/policies/behaveac.htm>) outlines the behaviours that constitute academic dishonesty and the processes for addressing academic offences. Potential offences include, but are not limited to:

**IN PAPERS AND ASSIGNMENTS:** Using someone else's ideas or words without appropriate acknowledgement. Submitting your own work in more than one course without the permission of the instructor. Making up sources or facts. Obtaining or providing unauthorized assistance on any assignment.

**ON TESTS AND EXAMS:** Using or possessing unauthorized aids. Looking at someone else's answers during an exam or test. Misrepresenting your identity.

**IN ACADEMIC WORK:** Falsifying institutional documents or grades. Falsifying or altering any documentation required by the University, including (but not limited to) doctor's notes. All suspected cases of academic dishonesty will be investigated following procedures outlined in the Code of Behaviour on Academic Matters. If you have questions or concerns about what constitutes appropriate academic behaviour or appropriate research and citation methods, you are expected to seek out additional information on academic integrity from your instructor or from other institutional resources (see

<http://academicintegrity.utoronto.ca/>).

## STAB22 - TENTATIVE LECTURE GUIDE

We cover chapters 1-21 of the customized text (chapters 1-18 and 20-22 of full 2nd edition). A tentative schedule is given below:

**Week 1:** Statistics starts here (Ch1), Variables (p5), Categorical, nominal, ordinal quantitative variables. Displaying and describing categorical data (Ch2, p18), frequency tables (p19), bar charts, pie charts (p21), contingency tables (p22), conditional distributions, (p24)

**Week 2:** Displaying and summarizing quantitative data (Ch 3, p52). Histograms (p53), The shape of a distribution (p57), The centre of a distribution, means (p60) and median (p62), skewness. Spread of a distribution (p64), Range, interquartile range (p68), and standard deviation (p64). The five number summary (p69)  $1.5 \times IQR$  rule for outliers (p70), Boxplot (p70), Understanding and comparing distributions (Ch 4), comparing groups with histograms (p100), comparing groups with boxplots (p101)

**Week 3:** The standard deviation as a ruler and the Normal distribution (ch 5, p131). Standardized values, z-scores (p132), Shifting data (p124), rescaling data (p126), linear and non linear transformations of data (p136). density curves and the normal model (p137), 68-95-99.7 rule for Normal models (p141), finding normal percentiles (p143), Normal probability (quantile) plot, form percentiles to scores (p148).

**Week 4:** Scatterplots, association and correlation (Ch 6, p165), describing scatterplots (p166), role of variables (p170), correlation (p170), Correlation conditions (p173), correlation properties (p175), Linear regression (Ch 7, p200), the least squares line (p200), predicted values and residuals (p201), residual plots (p208), regression assumptions and conditions (p214), R-square, the variation accounted for (p211)

**Week 5:** Regression wisdom (ch 8, p237), residuals, nonlinear relations (p237), subsets of data (p239), outliers and influence (p240), lurking variables and causation (p245), extrapolation (p243), , working with summary values (p248), restricted range (249)

**Week 6:** Sample surveys (ch 10, p290), population, sample, bias (p291), randomization (p291), sample size (p292), census, Populations and parameters, samples and statistics (p293), simple random samples, (p294), stratified samples (p296), cluster and multistage sampling (p297), systematic samples (p300), Common sampling mistakes (p304).

**Week 7:** Experiments and observational studies (ch 11, p318), observational studies (p318), experiments (p319), principals of experimental design (p321), does the difference make a difference (p325), experiments and samples (p326), control groups, blinding (p326), placebos (p328), blocking (p329), more factors (p331), confounding (p332), probability (ch 12, p350), the law of large numbers, empirical probability (p352), theoretical probability (p353), personal (subjective) probability (p354).

**Week 8:** Probability rules (p369), Conditional probability (p369), independence (p371), general multiplication rule (p372), multiplication rule for independent events (p373), independence and disjointness (p376)

**Week 9:** Random variables (ch 14, p394), discrete random variables, probability distributions (models), expectation (means) of a random variable (p394), standard deviation of a random variable (p397), linear transformations (p400), two or more random variables (p401), combining random variables (p399), binomial model (p405), binomial tables, Normal approximation



to the binomial distribution (p415).

**Week 10:** Sampling distribution models (ch 15, p433), sampling distribution for sample proportions, CLT for sample proportions (p438), sampling distribution for sample mean (p441), CLT (p443), confidence intervals for proportions (ch 16, p467), margin of error (p471), sample size (p476), testing hypotheses about proportions (ch 17, p496), null, alternative hypotheses (p496), p-values (p498), one-sided and two-sided tests (p502)

**Weeks 11 and 12:** More about tests (ch 18, p522), alpha levels (p528), confidence intervals and hypothesis tests (p532), inference about means (ch 20, p555), one sample t-interval for the mean (p559), tests for the mean (p566), Comparing two means (Ch 21, p596), two sample t-interval for the difference between two means (p597), a test for the difference between two means (p603). Paired sample (Ch 22, p631), the paired t-test (p632).