UNIVERSITY OF CALGARY
FACULTY OF SCIENCE

Department of Mathematics & Statistics
MS 476
403 220-5210
mathinfo@ucalgary.ca

COURSE OUTLINE

1. **Course:** DATA 603, Statistical Modelling with Data - Fall 2019

    Lecture 01: TR 09:00 - 12:00 in ICT 517

Instructor	Email	Phone	Office	Hours
Dr Thierry Chekouo Tekougang	thierry.chekouotekou@ucalgary.ca	403 220-7196	MS 554	TuTh 1-2:30 PM
Dr. Thuntida Ngamkham	thuntida.ngamkham@ucalgary.ca	403 220-6305	MS 550	TR: 15:00-16:00 , W: 13:00-15:00

    Lecture 02: TR 17:00 - 20:00 in ICT 517

Instructor	Email	Phone	Office	Hours
Dr. Thuntida Ngamkham	thuntida.ngamkham@ucalgary.ca	403 220-6305	MS 550	TR: 15:00-16:00 , W: 13:00-15:00

    Course Site:

    D2L: DATA 603 L01-(Fall 2019)-Statistical Modelling with Data

    Note: Students must use their U of C account for all course correspondence.

2. **Requisites:**

    See section 3.5.C in the Faculty of Science section of the online Calendar.

    Prerequisite(s):
    Data Science 602 and admission to the Graduate Certificate in Fundamental Data Science and Analytics or the Graduate Diploma in Data Science and Analytics.

3. **Grading:**

    The University policy on grading and related matters is described in F.1 and F.2 of the online University Calendar. In determining the overall grade in the course the following weights will be used:

Component(s)	Weighting %	Date
Assignment (4)	40%	Nov 1,Nov 22, Nov 29, Dec 6
Quizzes (2)	30%	Nov 7, Dec 5
Project (1)	30%	Dec 3

    Each piece of work (reports, assignments, quizzes, midterm exam(s) or final examination) submitted by the student will be assigned a grade. The student's grade for each component listed above will be combined with the indicated weights to produce an overall percentage for the course, which will be used to determine the course letter grade.

    The conversion between a percentage grade and letter grade is as follows.

	A+	A	A-	B+	B	B-	C+	C	C-	D+	D
Minimum % Required	95 %	90 %	85 %	80%	75%	70 %	65 %	60%	55%	50 %	45 %

    This course has a registrar scheduled final exam.

4. **Missed Components Of Term Work:**

    In the event that a student misses the midterm or any course work due to illness, supporting documentation, such as a medical note or a statutory declaration will be required (see Section M.1; for more information regarding the use of statuary declaration/medical notes, see FAQ). Absences must be reported within 48 hrs.

    The regulations of the Faculty of Science pertaining to this matter are found in the Faculty of Science area of the

Calendar in <u>Section 3.6</u>. It is the student's responsibility to familiarize themselves with these regulations. See also <u>Section E.3</u> of the University Calendar.

5. **Scheduled Out-of-Class Activities:**

   There are no scheduled out of class activities for this course.

6. **Course Materials:**

   Recommended Textbook(s):

   > Gareth James • Daniela Witten • Trevor Hastie Robert Tibshirani, *An Introduction to Statistical Learning with Applications in R* : Springer New York Heidelberg Dordrecht London.
   > Wickham and Grolemund, *R for Data Science*: O'Reilly Media.

7. **Examination Policy:**

   Using your personal laptop is permitted for this course.

   Students should also read the Calendar, <u>Section G</u>, on Examinations.

8. **Approved Mandatory And Optional Course Supplemental Fees:**

   There are no mandatory or optional course supplemental fees for this course.

9. **Writing Across The Curriculum Statement:**

   For all components of the course, in any written work, the quality of the student's writing (language, spelling, grammar, presentation etc.) can be a factor in the evaluation of the work. See also Section <u>E.2</u> of the University Calendar.

10. **Human Studies Statement:**

    Students will not participate as subjects or researchers in human studies.

    See also <u>Section E.5</u> of the University Calendar.

11. **Reappraisal Of Grades:**
    A student wishing a reappraisal, should first attempt to review the graded work with the Course coordinator/instructor or department offering the course. Students with sufficient academic grounds may request a reappraisal. <u>Non-academic grounds are not relevant for grade reappraisals</u>. Students should be aware that the grade being reappraised may be raised, lowered or remain the same. See <u>Section I.3</u> of the University Calendar.

    a. **Term Work:** The student should present their rationale as effectively and as fully as possible to the Course coordinator/instructor within **10 business days** of either being notified about the mark, or of the item's return to the class. If the student is not satisfied with the outcome, the student shall immediately submit the Reappraisal of Graded Term work form to the department in which the course is offered. The department will arrange for a re-assessment of the work if, and only if, the student has sufficient academic grounds. See sections <u>I.1</u> and <u>I.2</u> of the University Calendar

    b. **Final Exam:** The student shall submit the request to Enrolment Services. See <u>Section I.3</u> of the University Calendar.

12. **Other Important Information For Students:**

    a. **Mental Health** The University of Calgary recognizes the pivotal role that student mental health plays in physical health, social connectedness and academic success, and aspires to create a caring and supportive campus community where individuals can freely talk about mental health and receive supports when needed. We encourage you to explore the mental health resources available throughout the university community, such as counselling, self-help resources, peer support or skills-building available through the SU Wellness Centre (Room 370, MacEwan Student Centre, <u>Mental Health Services Website</u>) and the Campus Mental Health Strategy website (<u>Mental Health</u>).

    b. **SU Wellness Center:** The Students Union Wellness Centre provides health and wellness support for students including information and counselling on physical health, mental health and nutrition. For more information, see <u>www.ucalgary.ca/wellnesscentre</u> or call <u>403-210-9355</u>.

    c. **Sexual Violence:** The University of Calgary is committed to fostering a safe, productive learning environment. The Sexual Violence Policy <u>(https://www.ucalgary.ca/policies/files/policies/sexual-violence-policy.pdf)</u> is a fundamental element in creating and sustaining a safer campus environment for all

community members. We understand that sexual violence can undermine students' academic success and we encourage students who have experienced some form of sexual misconduct to talk to someone about their experience, so they can get the support they need. The Sexual Violence Support Advocate, Carla Bertsch, can provide confidential support and information regarding sexual violence to all members of the university community. Carla can be reached by email (svsa@ucalgary.ca) or phone at 403-220-2208 .

d. **Misconduct:** Academic misconduct (cheating, plagiarism, or any other form) is a very serious offence that will be dealt with rigorously in all cases. A single offence may lead to disciplinary probation or suspension or expulsion. The Faculty of Science follows a zero tolerance policy regarding dishonesty. Please read the sections of the University Calendar under Section K. Student Misconduct to inform yourself of definitions, processes and penalties. Examples of academic misconduct may include: submitting or presenting work as if it were the student's own work when it is not; submitting or presenting work in one course which has also been submitted in another course without the instructor's permission; collaborating in whole or in part without prior agreement of the instructor; borrowing experimental values from others without the instructor's approval; falsification/ fabrication of experimental values in a report. **These are only examples**.

e. **Assembly Points:** In case of emergency during class time, be sure to FAMILIARIZE YOURSELF with the information on assembly points.

f. **Academic Accommodation Policy:** Students needing an accommodation because of a disability or medical condition should contact Student Accessibility Services in accordance with the procedure for accommodations for students with disabilities available at procedure-for-accommodations-for-students-with-disabilities.pdf.

Students needing an accommodation in relation to their coursework or to fulfill requirements for a graduate degree, based on a protected ground other than disability, should communicate this need, preferably in writing, to the Associate Head of the Department of Mathematics & Statistics, Jim Stallard by email jbstall@ucalgary.ca or phone 403-220-3953. Religious accommodation requests relating to class, test or exam scheduling or absences must be submitted no later than **14 days** prior to the date in question. See Section E.4 of the University Calendar.

g. **Safewalk:** Campus Security will escort individuals day or night (See the Campus Safewalk website). Call 403-220-5333 for assistance. Use any campus phone, emergency phone or the yellow phones located at most parking lot pay booths.

h. **Freedom of Information and Privacy:** This course is conducted in accordance with the Freedom of Information and Protection of Privacy Act (FOIPP). Students should identify themselves on all written work by placing their name on the front page and their ID number on each subsequent page. For more information, see Legal Services website.

i. **Student Union Information:** VP Academic, Phone: 403-220-3911 Email: suvpaca@ucalgary.ca. SU Faculty Rep., Phone: 403-220-3913 Email: sciencerep@su.ucalgary.ca. Student Ombudsman, Email: ombuds@ucalgary.ca.

j. **Internet and Electronic Device Information:** Unless instructed otherwise, cell phones should be turned off during class. All communication with other individuals via laptop, tablet, smart phone or other device is prohibited during class unless specifically permitted by the instructor. Students that violate this policy may be asked to leave the classroom. Repeated violations may result in a charge of misconduct.

k. **Surveys:** At the University of Calgary, feedback through the Universal Student Ratings of Instruction (USRI) survey and the Faculty of Science Teaching Feedback form provides valuable information to help with evaluating instruction, enhancing learning and teaching, and selecting courses. Your responses make a difference - please participate in these surveys.

l. **Copyright of Course Materials:** All course materials (including those posted on the course D2L site, a course website, or used in any teaching activity such as (but not limited to) examinations, quizzes, assignments, laboratory manuals, lecture slides or lecture materials and other course notes) are protected by law. These materials are for the sole use of students registered in this course and must not be redistributed. Sharing these materials with anyone else would be a breach of the terms and conditions governing student access to D2L, as well as a violation of the copyright in these materials, and may be pursued as a case of student academic or non-academic misconduct, in addition to any other remedies available at law.

Course Outcomes

Multiple Linear Regression

1. Model the multiple linear relationship between a response variable (Y) and all explanatory variables (both categorical and numerical variable) with interaction terms. Interpret model parameter estimates, construct confidence intervals for regression coefficients, evaluate model fits by (R^2adj and RMSE).

2. Visualize correlations between a response variable (Y) and all explanatory variables (X) by graphs (scatter plot, residual plot) to assess model validity. Check the multi -linearity problem of a MLR model.

3. Use Stepwise/Forward/Backward variable selection procedures to identify the most significant subset of predictor variables that contribute to the explanation of the variation Y.

4. Check all assumptions by residual plots and statistical tests (testing for non constant variance, testing for normality) and outliers.

5. Apply transformations (logarithmic transformations, power functions) to some/all of the response variable and the predictor variables to create a better model.

6. Once the fitted model exists, predict the response variable at a certain level of the explanatory variables.

7. Use open data set to build the best fitted model, write reports for a group research project. Deliver professional presentations and discuss group research project results with the entire class.

Logistic Regression

8. Model the logistic regression for a binary response variable with a set of independent variables (both categorical and numerical variable), interpret model parameter estimates in terms of odd ratio, construct confidence intervals for logistic regression coefficients

9. Visualize correlations between a response variable (Y) and all explanatory variables (X) by graphs.

10. Check all necessary assumptions by residual plots, multicollinearity and outliers.

11. Predict the probability of response variable $P(Y=1|X)$ at a certain level of the explanatory variables

Experimental Design

12. Understand methods of sampling from population, definitions of Design of Experiment, and some concepts of DOE.

13. Test for difference in averages between levels for one factor design (complete randomized design), block design, two factor factorial design and nested design, use pairwise comparison procedures for comparing in pairs to determine which means are different from others, discuss the output in class and make conclusions.

14. Check all assumptions by residual plots and statistical tests (testing for non constant variance, testing for normality) for a model.

15. Apply non-parametric test when some assumptions are not satisfied.

Software

16. Learn to use statistical software R to analyse a model, write some R codes, and apply R functions to interpret results from outputs.

Electronically Approved - Sep 16 2019 17:05

Department Approval