University of Manitoba – Department of Statistics

STAT 7360 T11: Advanced Topics in Statistics 4

Winter Term 2018

Course Details

Course Number & Title:	STAT 7360, Advanced Topics in Statistics 4 Computational Statistics and Sports Analytics
Section & CRN:	Section T11, CRN: 56555
Course Schedule:	From January 3 to February 16 Tuesday/Thursday, 1:00 to 2:30 p.m. (Slot 7) 419 Machray Hall
Prerequisites:	There are no formal prerequisites for this course, but experience with the R statistical software will be an important asset.

Instructor Contact Information

Instructor:	Alexandre Leblanc
Office Location:	367 Machray Hall
Phone:	(204) 474-6273
Email:	Alex.Leblanc@umanitoba.ca
Office Hours:	Monday, from 1:30 to 3:00 p.m. and Wednesday, from 10:30 to 11:30 a.m., or by appointment.

General Goals for this Course

This course will be focusing on learning

- some modeling/computational/numerical aspects of classical Statistics,
- some modern computer-intensive methods used in Statistics,
- the application of these methods in the world of Sports Analytics, with a major focus on soccer.

Obviously, you will need to make extensive use of the computer (and the R statistical software) throughout the whole course.

In addition, I find that students in graduate programmes in Statistics often have a difficult time reading scientific papers, even though reading papers is a crucial aspect of graduate level training. Interestingly, many sports related papers are easily understood as they provide a good balance between well grounded data analysis using standard techniques and the more technical aspects of statistical modelling and inference. Hence, students will be asked to read many papers (an initial reading list is provided below) throughout the course, but these readings are expected not to be a big burden on the students, and in fact, to be rather enjoyable. Phenomenon known or believed to be important in sports analytics will be discussed depending on the students' interests: strength of schedule, home field advantage, streakiness and/or the 'hot hand' phenomenon, etc.

Textbook and Other Materials

Textbook: There is no formal textbook for this course.

Lecture notes: Lecture notes and other materials (e.g. assignments, sample codes and papers to read) will be posted on the UM Learn system.

Statistical Software: As mentioned above, you will need to make extensive use of the computer throughout the whole course for your homework/assignments.

Specifically, we will use the statistical software R which can be downloaded for free (for Windows, Mac, UNIX or LINUX) from the *Comprehensive R Archive Network* at: `http://cran.r-project.org/`

Sample codes seen in class will be provided to students through the UM Learn system.

Useful online resources on the R software can be found at:

- `http://cran.r-project.org` (R software, documentation and contributed packages)
- `http://www.rstudio.com` (for downloading R Studio)
- `http://rmarkdown.rstudio.com` (to learn about R Markdown)

Other references: Although the course material cannot be found in textbooks, many other relevant and useful references are available for free (downloadable in PDF) through the online catalogue of the University of Manitoba Libraries. Such interesting general references are:

Computational Statistics:

- Gentle, J.E. (2009), *Computational Statistics*, Springer.
- Lange, K. (2010), *Numerical Analysis for Statisticians*, Springer.

Statistics using *R*:

- Marin, J.-M. and Robert, C.P. (2014), *Bayesian Essentials with R*, Springer.
- Robert, C.P. and Casella, G. (2011), *Introducing Monte Carlo Methods with R*, Springer.

Programming in *R*:

- Lafaye de Micheaux, P., Drouilhet, R. and Liquet, B. (2013), *The R Software: Fundamentals of Programming and Statistical Analysis*, Springer.
- Chambers, J.M. (2008), *Software for Data Analysis: Programming with R*, Springer.

Reading List

A list of papers will be available (and regularly updated) on UMLearn. The list is expected to be modified throughout the term depending on the students' interests/needs.

Assignments, Test and Exam

Readings/Homework: You will be required to do some reading – research – homework (not to be turned in for marking) to prepare for each class. As mentioned above, you should find those readings to be enjoyable and reasonably simple – the papers you will be assigned all have a very applied nature, with a few of them also having some theoretical aspects that are directly related to the problems at hand. Some papers were published in statistical journals, but others come from the sports analytics literature.

Assignments: Assignments (to be graded) will be handed out at a rate of about one assignment per two weeks of class, for of total of 3 assignments that shouldn't be too long. See below for details on how to turn in your work.

Tests and Exams: There will be no test or exam in this course.

Project/Presentation: The class will end with each student selecting one of the two options for a final project.

The first option is for the student to present (using R Markdown or LaTeX) a mutually agreed upon paper related to sports that they find interesting. An important aspect to be discussed/highlighted should be the computational components of the work presented. Ideally, the student can report on their own simulation study related to the paper.

The second option is for the student to participate in the team that will prepare forecasting and visualization tools for the Statistics group at Info Days (to be held during the reading week). In this case, each student would be responsible for a different set of R functions to be used to display some of the statistical tools/techniques that can be used for Sports Analytics during Info Days. Each student would be expected to turn in a short report (prepared in R Markdown or LaTeX) on their contributions, including good clean R code, and to help the Info Days team in getting familiar with the tools that have been created through giving a short presentation.

Grading timeline: Work should normally be graded and returned promptly, within two weeks of submission.

Course Evaluation and Grading Scheme

Final Mark: The final mark for the course will be obtained from the following rule.

Assignments (3)	50%
Presentation	50%

Outline of Covered Topics and Approximate Timeline

The course is expected to proceed as follows.

Topics	Approximate Duration (in weeks)
1. Basic modelling techniques	1
2. Some useful computational techniques (Bootstrap, EM, Data Augmentation, etc.)	1
3. More advanced models	1
4. Rating team abilities	1
5. Measuring home advantage	1
6. Final projects – Work & Discussions	0.5
7. Other topics	0.5

Important Dates

The following dates are important to how the course will progress throughout the term.

Date	Information
January 4	First lecture
January 16	End of the registration revision period
Feb 13 & 15	Last week of lectures – Final Project
Feb 20 - Feb 23	Reading week – Info Days

Changes to the above are subject to Section 2.8 of the ROASS Procedure.

Work Submission Guidelines and Policies

Unless otherwise noted, all assignments must be submitted in the R Markdown format, which allows for incorporating R code within regular text, and should be written in a style that smoothly integrates prose, code, tables and graphics. All your code should be *human-readable*.

Make sure your submission properly compiles with "Knit HTML" before turning it in. Every file you submit should have your name and be properly labeled (Assignment 1, etc.)

Work submitted without following the above guidelines will not be marked and given a grade of 0.

Work submitted late will normally not be marked and be given a grade of 0.

Collaboration, Copying, and Plagiarism

Students are encouraged to discuss course material, including assignments and the final project. However, each student must hand in his or her own copy of each assignment/project with personalized solutions, including comments, discussions, explanations and interpretations, and R code. Copying from anywhere, including other students, books and the web,

1. does not help you to learn the course material,
2. constitutes a case of academic dishonesty and could have serious consequences.

If you are unclear on what is acceptable, please ask me or visit the Faculty of Science page on cheating and plagiarism at: http://umanitoba.ca/faculties/science/undergrad/resources/webdisciplinedocuments.html

Technology in the Classroom

It is the general University of Manitoba policy that all technology resources are to be used in a responsible, efficient, ethical and legal manner. Students should restrict their use of technology to those approved by the instructor *for educational purposes only.* Electronic messaging, email, social networking, gaming, etc. should be avoided during class time. Cell phones should be turned off. If a student is on call for emergencies, their cell phone should be on vibrate mode and the student should leave the classroom before using it.

Class Communications

The University requires all students to activate an official U of M email account, which should be used for all communications between yourself and the university (including all your instructors). All these email communications should comply with the University's policy on electronic communication with students, which can be found at: http://umanitoba.ca/admin/governance/governing_documents/community/electronic_communication_with_students_policy.html

Copyrights

Copyrighted Materials: We will use copyrighted content in this course. I have ensured that the content I use is appropriately acknowledged and is copied in accordance with copyright laws and University guidelines. Copyrighted works, including those created by me, are made available for private study and research and must not be distributed in any format without permission.

Lectures: No audio or video recording of lectures or presentations is allowed in any format, openly or surreptitiously, in whole or in part without my permission.

More details are available online at http://umanitoba.ca/copyright/.

ROASS Schedule A

Schedule "A" of the *Responsibilities of Academic Staff with regards to Students* (ROASS) policies of the University of Manitoba lists resources and policies for students. It is important that you familiarize yourself with these resources and policies. This document is available from the Department of Statistics web page at: http://umanitoba.ca/science/statistics/.

This document is dated January 4, 2018.