STA5075Z: Data Science Masters Statistical Computing

Course Outline - 2023

Welcome! This is an intensive two-week short course that introduces you to the R programming language and the RStudio IDE with a strong focus on solving statistical problems. While the course is very applied and practical, there are several key statistical concepts that you will need to learn.

Course objective

The overall objective of this short course is to introduce you to knowledge and tools you can use to tackle data science problems with computer programming in R.

Learning philosophy

This is a hands-on course where you as the student are primarily responsible for learning the necessary theory and programming skills to solve the variety of problems presented to you. There are no lectures. You are required to have your own laptop on which you will install R & RStudio. You will be provided with ample material in the form of textbooks, weblinks, examples & exercises with which you are to engage in order to master the content and assessments in this course. In summary, please adopt a proactive approach to this course. Your success in this course is firmly in your hands!
#takeresponsibilityforyourlearning

Collaboration via Face2face practical sessions

A key component to this course is collaboration. Should you take up a data scientist position in the near future, there is a high probability that you will end up working in a team. A crucial element of teamwork is the ability to clearly communicate with one another to solve the problems you are facing. This course aims to simulate that environment by randomly placing you into a team of 6-8 people with whom you will collaborate for the duration of this course.

You will be meeting in these groups every weekday (Monday to Friday) from 8am to 12pm to work on the practical problems. Each group will have a tutor who will be present each day to offer assistance when needed. Each day you will be presented with some problems ('Practicals') to solve together as a team. You are encouraged to communicate with one another to solve these Practicals.

If you are a student who is familiar with R and finds the problems easy then use this as an opportunity to sharpen up your communication skills in helping your classmates to solve the problems. You only understand something as well as you can explain it to someone else!
Note that while you are encouraged to collaborate in solving problems, you will need to submit your own work at all times.

You group allocation and meeting venue will be communicated on the Vula site closer to the start of the course. If you are not able to attend these face2face Practicals, please email Sulaiman Salau (sulaiman.salau@uct.ac.za) ASAP.

Daily Practicals

Each day you will be given some exercises that you will need to complete and submit before midnight of the next weekday. You are encouraged to collaborate with your group members when working on these exercises. Each student will then submit their own attempt at the daily exercise by midnight of the next day on the Assignments tab on Vula. You will get feedback on your submission from your tutor during the following days consultation session.

Tutors

Our tutors are available to assist your learning. Note that they have been specifically instructed NOT to give you 'the answer' but rather help guide you along as you attempt to solve the problems presented to you. They are there to help you after you have made a thorough attempt to understand underlying statistical concepts or solving the problems presented to you i.e., after engaging with the resources available, you are encouraged to pose as many questions to your tutor as you deem necessary.

Please note that in all cases when you ask for help with a Practical, you need to have evidence that demonstrates that have made a proper attempt before getting help from your tutor.

Each tutor will be allocated to a student group for the duration of the course i.e., they will be present at each of the daily practical sessions and a member of your MS Teams group.

Assessment

In addition to the daily Practicals, there will be two assignments and an exam during the first semester. The dates of these assessments, as well are indicated in the table below.

Assessment	Release date	Due Date	% of overall mark
Practicals	Each day	Midnight of the next weekday	20 (average)
Assignment 1	Tue 14 February	Tue 21 February	10
Assignment 2	Wed 22 February	Wed 01 March	10
Exam	TBA (April or June)	TBA (April or June)	40

All assessments require that you submit an Rmarkdown document that contains all your code and the compiled pdf or html that displays the relevant output for the assessment.

There are a total of 17 Practicals that you need to complete. These will be assessed according to the following rubric:

Description	Mark (%)
Not submitted on time or little attempt made	0
Submitted on time, incomplete code	25-50
Submitted on time, code runs but incorrect solution	51-75
Submitted on time, code runs without errors and correct solution	76-100

DP requirements

Note that it is compulsory to submit your own work for each Practical and both assignments in order to be granted DP and be allowed to take the exam.