Statistical Methods in R: Assessment Information

Assessment Title: End-of-Class Test
Module Name: Statistical Methods in R

Module Code:	**7011DATASCI**
Level:	**7**
Credit Rating:	**20**
Weighting:	**30%**
Maximum mark available:	**60**

Lecturer:	Daniel Perley
Contact:	Email: d.a.perley@ljmu.ac.uk Office: Liverpool Science Park IC2, Room 322

Hand-out Date:	4 January 2022, 1:30 PM
Hand-in Date:	4 January 2022, 4:50 PM
Hand-in Method:	Canvas
Feedback Date:	before 21 January 2022
Feedback Method:	Canvas

Learning outcomes to be assessed

- Critically apply a variety of statistical techniques to problems in data science, using the R computer language
- Critically evaluate the appropriateness of various techniques for a particular problem
- Critically analyze the outcomes of hypothesis testing in an informed fashion
- Synthesize a combination of statistical methods to investigate a complex problem

Detail of the task

The end-of-class test will consist of **four** problems, of which students are expected to answer **any three** of their choice. Each problem will contain a series of related questions, usually in association with an accompanying data set. Questions may ask the student to (for example) calculate a value, test a hypothesis, produce a plot, or create a data table. Most questions are designed to be solvable with a few lines of R code. Later questions may build on the results of previous ones from the same problem. Students will have three hours and twenty minutes to complete all chosen problems. Each problem will be worth 20 marks in total.

Students are expected to carry out the test at the reserved IT lab on campus (Henry Cotton Room 209). In the event that it is impossible to come to campus on the day of the test due to (e.g.) pre-arranged travel or documented illness, students may request in writing to complete the test at another location. Tests completed at home without prior permission will not be marked and will receive a zero.

Only R/Rstudio should be used: neither the data files nor the report should be edited in Excel, Word, or other software. (Students may compile to Word and export to PDF, however.) Students may use standard R plotting/analysis libraries (those used in this module and in the other Data Science modules).

Submission format

An Rmarkdown template will be provided containing all questions and empty code chunks, in a format similar to what is used in the Computer Lab tutorials. Students will be expected to fill in this template with their responses, and then submit a compiled PDF (using Rmarkdown's Knit function) and the associated .rmd code file, as two separate file uploads to Canvas.

Students should make sure that any responses to the question they elect not to answer are completely empty. If all four problems are attempted, only the first three will be marked, regardless of the degree of completion.

Output provided by code chunks should generally be restricted to the value, values, or graphics that the relevant question requests, or larger outputs (e.g. from the summary() command) from which those values are derived. (Lines producing other outputs should be deleted or commented out.) If the question requests a single numerical answer and the code chunk produces only a single numerical output, no further annotation is necessary. If more than one number is output by a code chunk (e.g. from a summary() command), the answer(s) to the question should be clearly indicated by writing it explicitly below the code chunk. If the question asks for multiple numbers, each should be written and clearly labeled below the code chunk. Marks will not be awarded if multiple potential "answers" could be inferred from the output. For questions requesting a written response, this should be written in the RMarkdown document below the question (and below the code chunk, if calculations and written responses are both required by the same question).

A page limit of **16 pages (including the cover page)** will apply to the submitted PDF document. A deduction of **two marks per page** (up to a maximum deduction of 10 marks) will apply to submissions in excess of this. The questions themselves and the cover-page instructions should **not** be deleted.

Students can resubmit up to the deadline. Work submitted after the deadline will be subject to a penalty of **two marks** deducted if submitted no more than 12 minutes late; submissions more than 12 minutes late will be penalized at an additional **two marks per additional minute late**. For students with an exemption to take the test at home, no accommodation can be made in the event of computer or internet disruption or for any other reason. In the event of computer/connectivity issues at home leading to inability to submit, the only alternative option will be to retake the test as a referral over the summer.

Policy on co-operation, assistance, and misconduct

Students may consult the textbook, the R online help documentation, completed coursework and lab activities, notes, the module Canvas website, and any other internet resource. Code taken or adapted from an online source should be attributed using a URL.

Students may not communicate with each other or with anyone else in any way (e.g. by e-mail, text, messaging client, discussion forum, in person). An invigilator will be available to answer procedural questions about the test but will not be able to provide advice on questions about the material.

The Faculty takes academic misconduct seriously and suspected cases will be investigated according to the University's standard policy: https://www.ljmu.ac.uk/about-us/public-information/student-regulations/academic-misconduct

Extenuating Circumstances

Students who are unable to complete the assessment at the scheduled time for a documented reason will be eligible for a deferral. The instructor must be notified well in advance of the test to make use of this option. More guidance is available at https://www.ljmu.ac.uk/about-us/public-information/student-regulations/guidance-policy-and-process