Prifysgol Cymru
Y Drindod Dewi Sant
University of Wales
Trinity Saint David

APPENDIX GA36d

LEVEL 7 ASSESSMENT SPECIFICATION

Student name:		**Student ID number:**	
Programme:	**MSc Data Science**		
Module:	**Advanced Statistical Programming**		
Module code:	**ACCA7022**	**Contribution to Overall Module Assessment (%):**	**50%**
Lecturer:	**Dr Nik Whitehead**	**Internal Verifier:**	**Dr Tim Bashford**
Assignment Title:	**Data Analysis**	**Word count (or equivalent):**	**3000**
Submission deadline:	**16:00 May 14th 2021**	**Return date of provisional marks & written feedback:**	**14/6/2021**
Submission method:	All written assessments, where practical and possible, must be submitted via Turnitin unless otherwise instructed by the Lecturer. (Please DO NOT put this assessment specification into Turnitin or it will match many similarities with other students' submissions.) **Alternative submission method (if applicable):** **Late submission of the assessment will result in a late penalty mark.** Penalties for late submission: Up to one week late, maximum mark of 50%. Over one week late, 0%. Only the Extenuating Circumstances Panel may approve a change to submission dates.		
Academic honesty / referencing:	Academic honesty is required. In the main body of your submission you must give credit to authors on whose research and ideas your work is based. Append to your submission a reference list that indicates the books, articles, etc. that you have used, cited or quoted in order to complete this assessment.		

Module Learning Outcomes (from module syllabus)
Upon the successful completion of this module, the student should be able to: 1. Select and apply appropriate statistical methods for the analysis of experimental data 2. Manipulate data and implement statistical programming techniques on large datasets 3. Select and apply appropriate visualisation methods for the presentation of analytical data 4. Perform data exploration using appropriate visualisation tools.

TASK DESCRIPTION

Task 1 (50%):
You should write an R script called `wineanalysis.r` to perform the following:

1. The file *WineSamples.xlsx* contains data from the test results on the composition of a number of different samples of three different types of wine. Further details of the dataset are available from the UCI Dataset Repository at https://archive.ics.uci.edu/ml/datasets/wine. Load this data into a data frame *wineData*, ensuring that the headers are imported and that the type of wine is imported as a factor. Your script should assume that the data file is stored in the same folder as the script file. Display *wineData*.
   (*5 marks*)
2. Add a column to *wineData* that shows proline as a factor with five equal-sized levels (v. low, low, medium, high, v. high) and give this an appropriate heading. Display *wineData*.
   (*10 marks*)
3. Add a column to *wineData* that shows the ratio of hue to colour intensity and give this an appropriate heading. Display *wineData*.
   (*5 marks*)
4. Using a loop to count the number of instances of each type of wine, record the type of wine and number of samples in a new data frame *wineSummary*. Display *wineSummary*.
   (*5 marks*)
5. For each type of wine, calculate the means and standard deviations of alcohol and phenols and add suitable columns to *wineSummary* to record these. Display *wineSummary*.
   (*10 marks*)
6. For each type of wine, calculate how the percentage of samples with a high or v. high level of proline and add a suitable column to *wineSummary* to record this. Display *wineSummary*.
   (*10 marks*)

A further 5 marks will be awarded for commenting, white space and code presentation.

Task 2 on next page.

Task 2 (50%):

You should write an R script called `covidanalysis.r` to perform the following:

1. The file *Covid.xlsx* contains data from the European Centre for Disease Prevention and Control showing testing rates across Europe during the current Covid-19 pandemic. Load this data into a data frame *countryData*, ensuring that the headers are imported. Your script should assume that the data file is stored in the same folder as the script file.

   (*5 marks*)
2. Create a data frame *CovidSummary* that shows descriptive statistics for the total number of new cases across Europe for weeks 21-25 and weeks 36-41. This should contain the following measures: mean, standard deviation, median, max, min, and range.

   (*15 marks*)
3. Using the Shapiro test for normality, and the boxplot plot to identify any outliers, clean your data to remove any outliers and show that the resulting data is normal. You should explain what you are doing and why you are doing this in the comments.

   (*10 marks*)
4. Perform a t-test on your normal data to identify whether there is a significant difference in differences new cases during the two study periods. You should state your experimental hypotheses, the result of the test, and whether you accept or reject the null hypothesis in your comments.

   (*5 marks*)
5. Perform an ANOVA test on your original data to identify if there are statistically significant differences between the data for the UK, France, Italy, and Spain. If such a difference is found, perform the appropriate ad-hoc tests to identify where the differences lie. You should state your experimental hypotheses, the result of the test, and whether you accept or reject the null hypothesis in your comments, and should explain what you are doing and why you are doing this in the comments.

   (*10 marks*)

A further 5 marks will be awarded for commenting, white space and code presentation.

Submitting your code

You should upload your two script files via the assignment links on Moodle. These will not go via Turnitin as Turnitin does not handle code well. Make sure that both scripts have your name and student number in a header comment.

You may be asked to talk through your code in a short viva after the submission date.

GUIDANCE FOR STUDENTS IN THE COMPLETION OF TASKS

NOTE: The guidance offered below is linked to the five generic assessment criteria overleaf.

1. Engagement with Literature Skills

Your work must be informed and supported by scholarly material that is *relevant* to and *focused* on the task(s) set; you should make use of scholarly reviews and primary sources, as appropriate (for example, refereed research articles and/or original materials appropriate to the discipline). You should provide evidence that you have accessed a wide *range* of sources, which may be academic, governmental and industrial; these sources may include academic journal articles, textbooks, current news articles, organisational documents, and websites. You should consider the *credibility* of your sources; academic journals are normally highly credible sources while websites require careful consideration/selection and should be used sparingly. Any sources you use should be current and up-to-date, mostly published within the last five years or so, though seminal/important works in the field may be older. You must provide *evidence* of your research/own reading throughout your work, using correctly a suitable referencing system, including in-text citations in the main body of your work and a reference list at the end of your work.

Guidance specific to this assessment: *Not applicable*

2. Knowledge and Understanding Skills

At level 7, you should be able to demonstrate a systematic understanding of knowledge, and a critical awareness of current problems and/or new insights, much of which is at, or informed by, the forefront of your academic discipline, field of study or area of professional practice, with a comprehensive understanding of techniques applicable to your own research or advanced scholarship. Your work must demonstrate your growing mastery of these concepts, principles, current challenges, innovation and insights associated with the subject area. *Knowledge* relates to the facts, information and skills you have acquired through your learning. You demonstrate your *understanding* by interpreting the meaning of the facts and information (knowledge). This means that you need to select and include in your work the contemporary concepts, techniques, models, theories, etc. appropriate to the task(s) set. You should be able to explain the theories, concepts, etc. meaningfully to show your understanding. Your mark/grade will also depend upon the *extent* to which you demonstrate your knowledge and understanding; ideally each should be complete and detailed, with comprehensive coverage.

Guidance specific to this assessment: *Not applicable*

3. Cognitive and Intellectual Skills

You should be able to: evaluate critically current research and advanced scholarship in the discipline; evaluate methodologies and develop critiques of them and, where appropriate, to propose new hypotheses; deal with complex issues both systematically and creatively to make sound judgements in the absence of complete data. Your work must contain evidence of logical, analytical thinking, evaluation and synthesis. For example, to examine and break information down into parts, make inferences, compile, compare and contrast information. This means not just describing what! But also justifying: Why? How? When? Who? Where? At what cost? At all times, you must provide justification for your arguments and judgements. Evidence that you have reflected upon the ideas of experts within the subject area is crucial to you providing a reasoned and informed debate within your work. Your choice of methodologies to gather data and information must be rigorously defended. Furthermore, you should provide evidence that you are able to make sound judgements and convincing arguments using data and concepts. Sound, valid, persuasive conclusions are necessary and must be derived from the content of your work. Where relevant, alternative solutions and recommendations may be proposed.

Guidance specific to this assessment: *Not applicable*

4. Practical Skills

At level 7, you should be able to demonstrate originality in the application of knowledge, together with a practical understanding of how established techniques of research and enquiry are used to create and interpret knowledge in the discipline. This includes acting autonomously in planning and implementing tasks at a professional or equivalent level, originality in tackling and solving problems, and decision-making in complex and unpredictable contexts or situations.

You should be able to demonstrate mastery of the leading edge subject-related concepts and ideas as they relate to real world situations and/or particular contexts. How do they work in practice? You will deploy models, methods, techniques, and/or theories, in those contexts or circumstances, to assess current situations, perhaps to formulate plans or plausible, justifiable recommendations to solve problems, or to propose new models, or to create artefacts, which may be innovative and creative, thereby demonstrating your understanding of how the boundaries of knowledge are advanced through research and/or application. This is likely to involve, for instance, the use of real world artefacts, examples and cases, the application of a model within an organisation and/or benchmarking one theory or organisation against others.

Guidance specific to this assessment: *The two tasks need different skillsets. Task 1 requires you to be able to pre-process data, adding and converting data between different data modes.. Task 2 requires you to perform statistical analysis in code. In both cases, properly commented code that gives the right answer will be enough to pass this assignment. Higher marks will require well-structured code that shows an understanding of the deeper concepts of R as a programming language.*

5. Transferable Skills for Life and Professional Practice

Your work must provide evidence of the qualities and transferable skills necessary for postgraduate-level employment in circumstances requiring sound judgement, personal responsibility and initiative in complex and unpredictable professional environments. This includes demonstrating: the independent learning ability for continuing professional development to advance existing skills and acquire new competences of a professional nature that will enable you to assume significant responsibility within organisations; that you can initiate and complete tasks, projects and procedures, whether individually and/or collaboratively, to a professional level; that you can use appropriate media to effectively communicate information, arguments and analysis in a variety of forms for a variety of audiences; fluency of expression; clarity and effectiveness in presentation and organisation. Work should be coherent and well-structured in presentation and organisation.

Guidance specific to this assessment: *Remember that all code should be commented so that another developer can easily understand it. The comments should include a header with the programmer's name, the title and purpose of the file, and the month it was written.*

STUDENT FEEDBACK FORM

This section details the *extent* to which the assessment criteria are demonstrated by you, which in turn determines your mark. The marks available for each category of skill are shown. Lecturers will use the space provided to comment on the achievement of the task(s), including those areas in which you have performed well and areas that would benefit from development/improvement.

Generic Assessment Criteria	Marks available	Marks awarded
1. Engagement with Literature Skills		
	0	
2. Knowledge and Understanding Skills		
	0	
3. Cognitive and Intellectual Skills		
	0	
4. Practical Application Skills		
Good working code that answers the questions set and demonstrates an understanding of R data structures and their manipulation.	90	
5. Transferable Skills for Life and Professional Practice		
Professional structuring and commenting of code.	10	

Assessment Mark (Assessment marks are subject to ratification at the Exam Board. These comments and marks are to give feedback on module work and are for guidance only until they are confirmed.)	**Late Submission Penalties (tick if appropriate)**		**%**
	Up to 1 week late (50% Max)		
	Over 1 week late (0%)		

GENERIC ASSESSMENT CRITERIA

Level 7
In accordance with the Framework for Higher Education Qualifications, at the end of Level 7 students should be able to demonstrate: a systematic understanding of knowledge, and a critical awareness of current problems and/or new insights, much of which is at, or informed by, the forefront of their academic discipline or area of professional practice; a comprehensive understanding of techniques applicable to their own research or advanced scholarship; originality in the application of knowledge, together with a practical understanding of how established techniques of research and enquiry are used to create and interpret knowledge in the discipline; conceptual understanding that enables the student to evaluate critically current research and advanced scholarship in the discipline to evaluate methodologies and develop critiques of them and, where appropriate, to propose new hypotheses. They will be able to deal with complex issues both systematically and creatively, make sound judgements in the absence of complete data, and communicate their conclusions clearly to specialist and non-specialist audiences. They will demonstrate self-direction and originality in tackling and solving problems, and act autonomously in planning and implementing tasks at a professional or equivalent level to continue to advance their knowledge and understanding, and to develop new skills to a high level. They the qualities and transferable skills necessary for employment requiring the exercise of initiative and personal responsibility; decision-making in complex and unpredictable situations/professional environments; and the independent learning ability required for continuing professional development.

Level 7	FAIL	FAIL	MARGINAL FAIL	SATISFACTORY (Pass)	GOOD to VERY GOOD (Merit)	EXCELLENT (Distinction)	EXCEPTIONAL (Distinction)
Category	**0-29%**	**30-44%**	**45-49%**	**50-59%**	**60-69%**	**70-84%**	**85-100%**
Engagement with literature (including current research, referencing, academic conventions and academic honesty)	Little or no evidence of reading and/or reliance on inappropriate sources. Views and findings mostly unsupported and non-authoritative. Referencing conventions used incoherently or largely absent.	Poor engagement with essential reading. No evidence of wider reading. Reliance on inappropriate sources, and/or indiscriminate use of sources. Heavily reliant on information gained through class contact. Inconsistent and weak use of referencing.	Engagement with a very limited range of relevant and credible sources. Some omissions and minor errors. Referencing conventions evident though not always applied accurately or consistently.	Engagement with an appropriate range of research-informed literature, including sources retrieved independently. Some over-reliance on texts. Referencing may show minor inaccuracies or inconsistencies.	Engagement with a wide range of research-informed literature, including sources retrieved independently. Selection of relevant and credible sources. Very good use of referencing, with no/very few inaccuracies or inconsistencies.	Engagement with an extensive range of relevant and credible literature, informed by the latest research. Consistently accurate application of referencing.	Exceptional engagement with an extensive range of relevant and credible literature, informed by the latest research. High-level referencing skills consistently and professionally applied.
Knowledge and	Major gaps in knowledge and	Gaps in knowledge, with	Limited knowledge and	Knowledge is accurate and	Knowledge has a well-defined	Excellent mastery of a	Exceptional mastery of a

Level 7	FAIL	FAIL	MARGINAL FAIL	SATISFACTORY (Pass)	GOOD to VERY GOOD (Merit)	EXCELLENT (Distinction)	EXCEPTIONAL (Distinction)
Category	**0-29%**	**30-44%**	**45-49%**	**50-59%**	**60-69%**	**70-84%**	**85-100%**
understanding (A systematic, conceptual understanding of knowledge, and a critical awareness of current problems and/or new insights, much of which is at, or informed by, the forefront of the discipline; a comprehensive understanding of techniques applicable to their own research)	understanding of the subject matter is not systematic. Substantial inaccuracies. No awareness of current problems, insights or the latest research and/or advanced scholarship within the discipline.	only superficial systematic understanding. Some significant inaccuracies and/or irrelevant material. No critical awareness of current problems, insights, or latest research within the discipline.	systematic understanding of the concepts and principles within the subject area, which to some marginal extent, is informed by current research and scholarship. Some critical awareness of current problems and/or new insights, but often under-developed.	reasonably detailed. A systematic understanding of the field of study informed by, to some extent, current research and scholarship, A critical awareness of current problems and/or new insights though this may be under-developed occasionally.	focus, which is reasonably extensive, coherent and detailed, with a critical awareness of current problems and/or new insights. Exhibits good understanding of the breadth and depth of contemporary and established views, and the work is, at least in part, well-informed by current research and scholarship.	complex and specialised area of knowledge. A systematic, excellent understanding of the concepts of the subject informed by current research and scholarship. Highly critical awareness of current problems and/or new insights. A critical, awareness of how the boundaries of knowledge are advanced through research.	complex and specialised area of knowledge. An exceptionally critical awareness of current problems and/or new insights. An outstanding understanding of the concepts of the subject, well-informed by current research and scholarship. A critical, sophisticated and nuanced awareness of how the boundaries of knowledge are advanced through research.
Cognitive and intellectual skills (Critical evaluation of current research and methodologies and develop critiques of them and, where appropriate, to propose new hypotheses; make sound	Wholly or almost wholly descriptive work. Little or no evaluation or critique or attempt at a systematic approach. Failure to develop arguments, leading to illogical or invalid	Largely descriptive work, with superficial use of critical evaluation of research and methodologies. Absent or weak development of hypotheses and judgements. Information accepted uncritically, uses	Limited attempt at critical evaluation/ critique of current research and methodologies, tending towards description. Limited attempt to propose new hypotheses. Can deal with complex issues	Some critical evaluation/ critique of current research and methodologies, though slightly underdeveloped in places, Proposes adequate but limited new hypotheses, where relevant.	Sound critical evaluation/ critique of current research and methodologies, Proposes new hypotheses, where appropriate. Can deal with complex issues systematically and with some	Excellent critical evaluation/ critique of current research and methodologies, Proposes innovative hypotheses, where appropriate. Can synthesise complex issues systematically	Exceptional critical evaluation/ critique of current research and methodologies, Proposes innovative hypotheses, where appropriate. Exemplary systematic and

Level 7	FAIL	FAIL	MARGINAL FAIL	SATISFACTORY (Pass)	GOOD to VERY GOOD (Merit)	EXCELLENT (Distinction)	EXCEPTIONAL (Distinction)
Category	**0-29%**	**30-44%**	**45-49%**	**50-59%**	**60-69%**	**70-84%**	**85-100%**
judgements in the absence of complete data.)	judgements. Unsubstantiated generalisations, made without use of credible evidence.	generalised statements made with scant evidence and unsubstantiated opinions. Ideas sometimes illogical and contradictory.	but not systematically or creatively. Some evidence to support emerging judgements but these may be underdeveloped or with a little inconsistency / mis-interpretation. May assert rather than argue a case.	Can deal with complex issues but not fully systematically or creatively. Ability to make judgements based on data (that may be incomplete) but with some tendency to assert/state opinion rather than argue on the basis of reason and evidence.	creativity. Ability to make sound judgements based on data (that may be incomplete)	and creatively. Ability to investigate contradictory or incomplete information and make strong, persuasive, arguments and sophisticated judgements.	creative synthesis of complex issues. Ability to investigate contradictory or incomplete information and make strong, persuasive, arguments and sophisticated, nuanced, judgements. Potential for journal publication or doctoral research.
Practical skills (Originality / creativity in the application of knowledge, tools and techniques and in tackling and solving problems in complex and unpredictable professional situations; practical use of established techniques of research and enquiry to create and interpret knowledge in the discipline.)	Limited or no use of methods, materials, tools and/or techniques. Little or no appreciation of the context of the application. Limited understanding of the application of theory to practice or making appropriate links between the two. Very weak problem-solving skills in complex and	Rudimentary application of methods, materials, tools and/or techniques but without consideration and competence. Flawed appreciation of the context of the application. Weak understanding of the application of theory to practice, with only occasional evidence of	An awareness and mostly appropriate application of well-established methods, materials, tools and/or techniques, with occasional errors. Basic appreciation of the context of the application. Theoretical knowledge and understanding applied in practice, but not always making	An appropriate application of standard methods, materials, tools and/or techniques. Clear appreciation of the context of the application. Mainly consistent, accurate and logical application of theory to practice, making appropriate links between the two.	A very good application of a range of methods, materials, tools and/or techniques. Very good consideration of the context of the application, with perceptive insights. Can identify problems and propose appropriate solutions in complex and unpredictable contexts.	An advanced application of knowledge, methods, materials, tools and/or techniques. The context of the application is well considered, and insightful. Can identify complex problems and propose excellent solutions. An excellent grasp of techniques applicable to own research or	Exceptional application skills in complex, unpredictable, contexts, drawing skilfully on the latest research within the discipline. Can identify complex problems and propose sophisticated, original solutions. An outstanding application of techniques applicable to own research or

Level 7	FAIL	FAIL	MARGINAL FAIL	SATISFACTORY (Pass)	GOOD to VERY GOOD (Merit)	EXCELLENT (Distinction)	EXCEPTIONAL (Distinction)
Category	**0-29%**	**30-44%**	**45-49%**	**50-59%**	**60-69%**	**70-84%**	**85-100%**
	unpredictable contexts.	making appropriate links between the two. Weak problem-solving skills in complex and unpredictable contexts.	logical links between the two. Can identify problems and propose basic solutions without fully appreciating the complexity of unpredictable contexts.	Can identify problems and propose mostly appropriate solutions in complex and unpredictable contexts, with elements of originality.	Evidence of originality and creativity.	advanced scholarship. Shows originality in application of knowledge and techniques, and of how established techniques of enquiry create and interpret knowledge in the discipline.	advanced scholarship. Shows originality in application of knowledge and techniques, and of how established techniques of enquiry create and interpret knowledge in the discipline with assimilation and development of cutting edge processes and techniques.
Transferable skills for life and professional practice (Exercise of self-direction, autonomy and personal responsibility; plan and implement tasks at a professional level; independent learning; use appropriate media to communicate effectively and professionally to a variety of audiences; fluency of expression; systematic	Communication medium is inappropriate or misapplied. Work is poorly structured, disorganised and/or confusingly expressed. Very weak use of language and/or very inappropriate style. Little or no evidence of autonomy (or collaboration, where relevant) in the completion of tasks. Little or no evidence of	Communication medium is poorly designed and/or not suitable for the audience. Work is poorly presented in a disjointed manner. It is loosely, and at times incoherently, structured, with information and ideas often poorly expressed. Weak use of language and/or inappropriate style. Weak independent	Can communicate in a suitable medium but with some room for improvement. Mostly ordered presentation and structure in which relevant ideas / concepts are reasonably expressed. Work may lack coherence in places. Can work as part of a team, but with limited involvement in group activities.	Can communicate effectively in a suitable format, but may have minor errors. Mostly coherent, organised work, in a suitable structure and is for the most part clearly expressed. Can work effectively independently and/or as part of a team, with clear contribution to group activities. Demonstrates the skills	Can communicate well, confidently and consistently in a suitable format. Work is coherent, fluent, well-structured and organised. Can work very well autonomously and/or as part of a team, with a good contribution to group activities. Demonstrates comprehensive professional, postgraduate	Can communicate professionally confidently and consistently in a suitable format. Work is coherent, very fluent and is presented professionally. Can work autonomously with initiative. Where relevant can work professionally within a team, showing leadership skills as appropriate, managing	Can communicate with an exceptionally high level of professionalism. Work is remarkably coherent, very fluent and is presented professionally. Can work outstandingly well and professionally within a team, showing advanced leadership skills. Demonstrates exemplary

Level 7	FAIL	FAIL	MARGINAL FAIL	SATISFACTORY (Pass)	GOOD to VERY GOOD (Merit)	EXCELLENT (Distinction)	EXCEPTIONAL (Distinction)
Category	**0-29%**	**30-44%**	**45-49%**	**50-59%**	**60-69%**	**70-84%**	**85-100%**
approach; clarity and effectiveness in presentation and organisation.)	the skills required in professional, postgraduate employment.	initiative (or collaboration, if relevant). Limited evidence of the skills required in professional, postgraduate employment.	Demonstrates some but not all of the basic skills required in professional, postgraduate employment, with some areas of minor weakness.	required in professional, postgraduate employment, with some areas of strength and some of minor weakness.	employment skills.	conflict and meeting obligations. Demonstrates excellent professional, postgraduate employment skills and a strong appetite for further development.	professional, postgraduate employment skills and a strong appetite for further development.

THIS DOCUMENT IS ALSO AVAILABLE IN WELSH