Further Statistics and Probability: Assessment Briefing

MODULE TITLE:	Data science				
PROGRAMME:	BASc in Interdisciplinary Problems and Methods	**CREDITS:**	15	**LEVEL:**	5
ASSESSMENT OVERVIEW:	This module has THREE (3) assessments.				
SUBMISSION DEADLINES:	Summative Assessment #1: Every week Summative Assessment #1: Week 6 Summative Assessment #2: Week 12				
SUBMISSION METHOD:	ELECTRONIC VIA CORTEX. Please see submission details below.				
MODULE STAFF:	Matthew Brett				

Assessment #1

Formative / Summative	**Assessment method and limits**	**Weighting (% of whole module)**	**Module Learning Outcome(s) assessed by this task**	**Programme Learning Outcome(s) assessed by this task**
Summative	Weekly notebook exercises in class. Notebook exercises are questions asking for answers in code, multiple choice answers and free-text answers, embedded in a narrative framework of text and code.	10%	Develop fluency in using notebook interface in narrative and executable data analysis. Extend ability to clean, select, tabulate and visualize data. Develop ability to understand the nature of real data and the world from which they are derived. Show understanding of probability, variability, and data modelling using code.	Adopt a positive attitude to continual learning, showing some evidence of metacognition. (A1) Knowledge of several discipline-centred concepts from across the sciences, social sciences, arts and humanities. (K1) Use both quantitative and qualitative methods, with some awareness of their respective limitations, to illuminate understanding of real-world issues. (S2)

				Knowledge of some concepts and methods associated with the tackling of problems using an interdisciplinary approach. (K2)

Deadline and submission details:

You should submit your notebooks by **Midnight on Saturday of each week.**

- Submission: Your assessment must be submitted as one notebook (.ipynb) file Cortex has guidance about Submissions, including how to save and upload the notebook file.

Filename: You must put your student ID number (NOT your name) on your work and use the following naming convention when saving your work for submission: [QT2001_weekly1_student number], where QT2001 names the course, and "weekly1" means the first of the weekly submissions.

We may not be able to find and mark your submission, if you don't follow these instructions carefully.

Assessment Task description:

Each week you will be working on questions in a Jupyter notebook during a supervised lab session, in which you can ask questions and discuss answers. Generally we expect you will be able to fill in most or all of the notebook during the lab session, and you may well want to submit them at the end of the session, but you have until the end of the week to finish and submit them.

The notebooks to submit consist of explanations and questions, sometimes in the form of code for you to fill in, sometimes in the form of text with free-text or multiple-choice answers. The code answers will often have some skeleton for you to start from and will usually have some tests for you to run to check you are on the right track, or even, whether you have the right answer. You can submit the notebooks that you work on during the class as files, via Cortex.

Further guidance about the assignment and opportunities to ask questions will be available as part of the teaching on the module. Please do also feel free to contact Matthew Brett as the module lead on matthew.brett@lis.ac.uk.

Please ensure that you are familiar with the additional support available to you via the Student Support team. Visit Cortex - Student Support for more information.

Marking criteria:

We will mark the lab notebook for each week, for a total of 10 notebooks over the course. Each notebook therefore counts for 1% of the total course mark.

We will give you 40% of the marks per notebook, for submitting any notebook. The remainder (60%) of the marks per notebook are from the correctness of your code answers, plots and text answers.

Pass: 40+, Merit: 60+, Distinction: 70+.

Assessment #2

Formative / Summative	Assessment method and limits	Weighting (% of whole module)	Module Learning Outcome(s) assessed by this task	Programme Learning Outcome(s) assessed by this task
Summative	Coding work sheets for Python	40%	Understand the concepts and data structures of python (or another coding language) Write and debug code in a way that reflects best practice Automate routine computational tasks	Comprehensive understanding of the techniques to analyse linguistic, numerical, and image data using state of the art methods (i.e. use methods developed within relevant disciplines for analysing data associated with the three main systems of cultural representation).

Deadline and submission details:

You should submit your work by **Monday 12 noon, Week 6.**

- Submission: Your assessment must be submitted as one notebook (.ipynb) file Cortex has guidance about Submissions, including how to save and upload the notebook file.

Filename: You must put your student ID number (NOT your name) on your work and use the following naming convention when saving your work for submission: [QT2001_mid_student number], where QT2001 names the course, and "mid" means the mid-term assessment.

We may not be able to find and mark your submission, if you don't follow these instructions carefully.

Assessment Task description:

The assessment consists of a Jupyter notebook, like the ones we will be using in the classes, and that you submitted for your weekly labs. The notebooks consist of explanations and questions, sometimes in the form of code for you to fill in, sometimes in the form of text with free-text or multiple-choice answers.

We will give you a week to work on the notebook before submitting. Please remember that you are not allowed to share answers or work on your notebooks together. We will check for evidence that you have shared answers when we mark the notebooks.

The notebook will be highly structured. The code answers will often have some skeleton for you to start from and will usually have some tests for you to run to check you are on the right track. Each question in the notebook generally counts for the same marks, except where we indicate otherwise.

Further guidance about the assignment and opportunities to ask questions will be available as part of the teaching on the module. Please do also feel free to contact Matthew Brett as the module lead on matthew.brett@lis.ac.uk.

Please make sure that you are familiar with the additional support available to you via the Student Support team. Visit Cortex - Student Support for more information.

Marking criteria:

We mark the code questions with automated tests to see if you got the right result, although we give partial marks for answers that are close, or where you get the wrong answer because of a previous mistake. We mark your graphs and plots, and your free-text answers by hand according to pre-defined criteria.

The relationship of numerical marks to quality descriptors is from the LIS generic marking criteria:

Quality	**Mark**	**Descriptor**
Bare Pass	45 (40-49)	The piece of work meets enough of the criteria to demonstrate learning from the course. There are substantial ways in which it could be improved.
Secure Pass	55 (50-59)	The piece of work meets the criteria in a way that demonstrates secure learning from the course. There are some important ways in which it could be improved.
Merit-worthy	65 (60-69)	The piece of work meets all basic criteria and some of the criteria for excelling. There are some ways in which it could be improved.
Distinguished	75 (70-79)	The piece of work meets several of the criteria for excelling. There are minor ways in which it could be improved.

Professional/publishable	85 (80-100)	The piece of work is exemplary. It is unimprovable and would be ready for sharing with an external audience.

Assessment #3

Formative / Summative	Assessment method and limits	Weighting (% of whole module)	Module Learning Outcome(s) assessed by this task	Programme Learning Outcome(s) assessed by this task
Summative	Project: 15 minute group oral presentation. Reproducible data analysis, code and data. Self-assessment of contribution with evidence.	60%	Show initiative in choosing research question and matching data sets for data research project. Develop teamwork for delegating tasks, evaluating the results, and taking joint ownership of the chosen question, data and analysis methods. Learn to collaborate with others on a code and data analysis task in order to draw real-world conclusions. Choose informative data analysis methods, apply them correctly, and draw sound conclusions. Take joint responsibility for management of task and team. Show joint understanding with	Knowledge of several discipline-centred concepts from across the sciences, social sciences, arts and humanities. (K1) Use both quantitative and qualitative methods, with some awareness of their respective limitations, to illuminate understanding of real-world issues. (S2) Carry out small-scale interdisciplinary research activities, and/or design an extended research project, with appropriate guidance and supervision. (S4) Communicate across disciplinary boundaries and to a variety of audiences, using written, oral and visual forms. (S5) Be able to exercise some initiative and

			final presentation, followed by questions and answers.	independence of mind. (A2) Be able to work collaboratively as part of a team to deliver collective goals, while also taking individual responsibility and demonstrating decision-making ability. (A3)

Deadline and submission details:

There are three parts to the final assessment. The final assessment is a group project analyzing data of your own choice, as agreed with the module leader. The three parts are:

- Group presentation: deadline is a session with your group and the module leader, by arrangement, during week 10 of the module.
- Your group project: deadline to finish work on the project is **Wednesday 12 noon, Week 12 of the module.** You will be working on your group project on our online code system, JupyterHub. We will tell you more about how to do this during the course. Your group work will be notebooks and code to describe and implement your group data analysis. The submission process is automatic; we take a snapshot of your files on the online system at the deadline and mark those files.
- Personal contribution: submit your personal contribution as a PDF document to Cortex, by Friday 12 noon, week 12 of the module. Please see Cortex for guidance about submissions.

Filename for personal contribution document: You must put your student ID number (NOT your name) on your work and use the following naming convention when saving your work for submission: [QT2001_personalcontribution_studentnumber].

Your assessment may go missing if you do not use the naming scheme above.

Assessment Task description:

The final assessment is a substantial group project, working on data of your choice, but as agreed with the module leader. The assessment is to give you experience of a real-world data analysis, in the real-world setting of collaborating closely with a group. We do expect you to understand and be able to use the techniques we have we have shown you in the course – if they are the right ones for your data. You should not use any techniques that you do not understand. We would far prefer that you do simple, clear analyses using basic techniques than complex analyses that you do not fully understand. Your job as a data scientist is to draw clear conclusions from data. Often this will just involve selecting and

plotting relevant data and making an argument from the results. We will be looking for good judgement in choosing the right methods to analyse and explain your data

A useful data analysis must be reproducible – someone else should be able to pick up your analysis and run it, to get the key results you use in your conclusions. For that reason, your final report should be in the form of a Notebook, or similar executable document that interweaves code for the analysis and explanation of the analysis and results.

Further guidance about the assignment and opportunities to ask questions will be available as part of the teaching on the module. Please do also feel free to contact Matthew Brett as the module lead on matthew.brett@lis.ac.uk.

Please make sure that you are familiar with the additional support available to you via the Student Support team. Visit Cortex - Student Support for more information.

Marking criteria:

Group project:

Each of the following 7 sections is worth 10 marks, for a total of 70, representing 70^ of the project mark and therefore 70% * 60% = 42% of the overall module mark.

For each row, the table gives a description of work that will earn: between 0 and 50% in this category (fail to bare pass); 50-75% (bare pass to good pass); and 75-100% (good pass to excellent).

Notice that your project presentation is one category; see above.

	0-50%	50-75%	75-100
Questions	Questions overly simplistic, unrelated, or unmotivated	Questions appropriate, coherent, and motivated	Questions well motivated, interesting, insightful, and novel
Analysis	Choice of analysis is overly simplistic or incomplete	Analysis appropriate	Analysis appropriate, complete, advanced, and informative
Results	Conclusions are missing, incorrect, or not based on analysis Inappropriate choice of plots; poorly labelled plots; plots missing	Conclusions relevant, but partially correct or partially complete Plots convey information but lack context for interpretation	Relevant conclusions explicitly tied to analysis and to context Plots convey information correctly with adequate and appropriate reference information
Readability	Code is messy and poorly organized; unused or irrelevant code distracts when reading code. Variables and	Code is reasonably well organized. There is little unused or irrelevant code, or this code has been moved out of the main project files. Variable and	Code very well organized. No irrelevant or distracting code. Variable and function names have clear relationship to their purpose in the code. Code

	functions names do not helpful to understand code.	function names generally meaningful and helpful for understanding.	is easy to read and understand.
Presentation	Verbal presentation is illogical, incorrect, or incoherent. Visual presentation is cluttered, disjoint, or illegible Verbal and visual presentation unrelated	Verbal presentation partially correct but incomplete or unconvincing Visual presentation is readable and clear Verbal and visual presentation related	Verbal presentation is correct, complete, and convincing Visual presentation is appealing, informative, and crisp Verbal and visual presentation clearly related
Writing	Explanation is illogical, incorrect, or incoherent	Explanation is correct, complete, and convincing	Explanation is correct, complete, convincing, and elegant
Reproducibility	Code didn't run	Recipes in project directory correctly load data and generate all results and figures in report	Recipes additionally validate data against its source (such as URL or other download). The recipes generate all exploratory work and supplementary analysis

Personal contribution

The personal contribution contributes 30% to the project mark, and therefore 30% * 60% = 18% of the overall project mark.

Each team member should submit a document of up to 1500 words describing their contribution to the project, under any of the following headings:

- Development of question / hypothesis;
- Data research: search for relevant data to contribute to question;
- Literature review;
- Analysis strategy;
- Analysis code;
- Code review;
- Work planning and organization;
- Improving teamwork and collaboration;
- Testing code and procedures;
- Writing the report.

You should describe your own contribution, and any work you did to help other people contribute to the same area.

You can add other headings if you think we should consider them.

For each heading, give any evidence for your contribution from project files or other data that is accessible to us, your graders.

The mark guidelines for this part are:

- 0-50%: little evidence of contribution or collaboration. Contributions under few headings. Little effort to help other team members contribute.
- 50-75%: moderate evidence of contribution and collaboration. Contributions under many headings; substantial contributions to more than one heading. Evidence that you helped other team members contribute across some headings.
- 75-100%: strong evidence of contribution and collaboration. Some contribution to nearly all headings; substantial contributions across the majority of headings. Strong evidence that you helped other team members contribute across several headings.

IMPORTANT NOTES AND FURTHER INFORMATION

Students are reminded that it is their responsibility to organise themselves so that they can submit on time and they must familiarise themselves with LIS policy on late submission.

Extension/Late Penalty Information: Any work handed in after the deadline will be subject to penalties. Work handed in within 6 days of the deadline will be subject to a deduction of 10 percentage points per day and work handed in over 7 days after the deadline will be given a mark of zero.

Both these penalties apply unless there has been a prior agreement by the relevant staff member for an extension of the deadline or the submission is permitted late through an accepted extenuating circumstances claim. Your attention is drawn to the Extensions and Extenuating Circumstances policy available on the website and Cortex.

Feedback. The School's policy on feedback is that feedback on summative work will be provided no later than 4 weeks after the submission deadline. All marks are provisional until the Examination Board. Summative feedback Is normally provided via e-mail to your School e-mail account.

Self-Plagiarism. Students cannot submit any work for this assignment that they have submitted for any previous assignment unless it is appropriately referenced. Exceptions apply If this Is a resubmission of an earlier (failed) piece of work.

Authenticity and Academic Integrity: Plagiarism, cheating and collusion and attempting to obtain an unfair academic advantage are entirely unacceptable and not allowed. As such, these and any other forms of academic misconduct will be subject to disciplinary regulations.

In submitting assessments students accept responsibility for any copyright infringement, collusion, plagiarism or other form of academic misconduct. The School's Academic Misconduct Policy is available on the School website and on Cortex.

Your work will be analysed by plagiarism detection software.

Referencing. Your work must be referenced appropriately using a valid method of academic referencing highlighted by the School. Information on acceptable forms of referencing (such as the Harvard system of referencing) is available on Cortex.

Type of Assessment

- Individual: All work submitted for each element of this assessment must be your own individual work. Except for the group tasks you are not permitted to work with anyone else on these assessment tasks.
- Group work: If one or more tasks involves group work then please check the guidance on group work available on Cortex.