

 University of Central Lancashire UCLan School of Psychology and Computer Science	UCLan Coursework Assessment Brief		Academic Year 2022/23
	Module Title: Data Science	Module Code: CO3722	Level 6
	Practical Assessment: Emerging Technologies		This assessment is worth 50% of the overall module mark

THE BRIEF/INSTRUCTIONS

Context

Businesses and organisations need to continually develop their systems and technologies in order to be able to meet the challenges of the modern globalised world. Their challenge, however, is to be able to make rational decisions about which technologies they should incorporate in their businesses. Academia and other places of research continually invest in new Artificial Intelligence (AI) concepts, including data science and analytics.

One of the key sources for identifying emerging IT products is the annual Gartner Hype curves for a range of application areas, see <https://www.gartner.com/technology/research/methodologies/hype-cycle.jsp>

The Task

Using your Research Report from assignment 1, you will develop a small program in the Python programming language that generates business insights from an appropriate dataset for your chosen business and Use Case. In addition to your programming outcomes, you will summarise your findings discussing the implementation details of your Python model, dataset and insights gained.

Key points to reflect on and consider before implementation of your model, which will form part of your submission, include discussion on:

- Business/Use Case and thus, Stakeholders selected and the types of questions that would provide insight and competitive advantage. Consider your original problem-base in assignment 1.
- Data pre-processing/cleaning, preparation, analysis and visualisations of the chosen dataset.
- Machine learning technologies to automate processes.

The contents of the assignment should focus on extracting all the knowledge within your selected dataset for the purpose of “telling a story” about the data. This should include a demonstration of your analytics skills, which should reflect on importing data into an analytics model, data cleansing, data preparation such as merging, appending, etc., data insights and creating various reports such as tables, charts, graphs. The purpose of ‘telling a story’ here implies you to explain the nature of the data, what inferences you are looking for, and how they are achieved. You must include your results of the analysis in your submission. Coding used should be included in the appendices. Your Practical Assessment submission will include samples of your program coding and written reflection, which will have the following sections with respective weighing:

Question Design (20%) – this introduces the dataset and the objectives of your analytics, focusing on the stakeholders, questions, and how your analytics can benefit your business and Use Case.

Program Design (40%) – this section details how you are constructing the phases of analytics, namely pre-processing such as importing, cleansing, data preparation, analysis and post-processing such as presenting results in a variety of tables, charts and graphs. What methods, machine learning algorithms and experiments have you carried out? What challenges have you faced and decisions made to gain appropriate insights for Use Case? Describe your experience and the process that you have gone through whilst evaluating your chosen dataset. Show evidence of outputs using screen shots and/or sample coding.

The Story/Insights (40%) – this section discusses the story of your analytics, and the results of your experiments, including statistics computed, trends identified, etc. with visuals when required for discussion. What are your concluding insights and how would this be appropriately disseminated to your Use Case?

Formatting Requirements

You are recommended to use appropriate headings and sub-headings to provide structure.

Please submit the assignment as a single Word document in the Blackboard Assignment Submission link for the module. Sample code can be included in your discussions and any other remaining code can be included in your appendices. Please note that all assignments will be checked for plagiarism using computer software (Turnitin) as part of the submission process.

Learning Outcomes for Assessment: 1, 2, 3 & 4

- 1. Perform and evaluate basic descriptive, exploratory and confirmatory data analysis**
- 2. Write programs in an appropriate language for analysing data**
- 3. Visualise data in multiple ways**
- 4. Summarise, organise and communicate knowledge from data analysis**

To achieve 40% +

- you will have a basic program written in Python.
- You will be able to carry out some analysis of your dataset with little or no justification of the results and insights produced.
- Limited or no evidence of data cleaning and your approach used to prepare the dataset for analysis.
- There is limited discussion of your algorithm and use of visualisations. Visualisations are used to inform decisions, which is unclear in your reflective account.
- There is limited use of machine learning and its appropriateness for your model. There may be some use of statistics to demonstrate comparison between testing and predictive outcomes.
- A fair report with descriptions of the findings, visualisations and the final model.

To achieve 50% + you will have a satisfactory program written in Python. Your implementation will demonstrate some consideration of the requirement to review and pre-process data with reference to 'null' values and outliers and an approach for managing such data. Visualisations are described with some consideration of the possible choice of machine learning algorithm for your model. There is some description of the output and insights drawn from the experiments carried out using test sets and statistical measures such as variance, absolute and mean squared error. Overall, a satisfactory model is produced that could be enhanced through further examination of the features of the dataset. A satisfactory report with descriptions of the insights identified and final model.

To achieve 60% + you will have a very good program written in Python. There is evidence of research into the different types of algorithms for data science and use within your model. This will be reflected in your discussions and use of visualisations. Your dataset is appropriately pre-processed assessing any 'null' values and outliers with some descriptions of the decisions made to amend the data before data analysis. Further, statistical measures are used such as mean, standard deviation and variance to support such decisions. Evidence of more than one experiment is used to evaluate the impact that one or more features has on the target data. Several visualisations are used to communicate findings appropriately. The model is tested using at least the 'hold out' approach and supported through statistical measures such as variance, absolute and mean square error. Discussion is used to evaluate the model and its effectiveness for future predictions. A very good reflective report is presented with mostly solid justifications and evaluation of the model.

To achieve 70% + you will have an excellent/outstanding program written in Python. This will be shown through experimentation with different machine learning algorithms and use of visualisations to inform critical decisions of the approach to take for the model. These decisions are justified in your reflective account. You have carefully reviewed your dataset demonstrating your approach to pre-processing the data and critical discussion of the

decisions made such as managing 'null' values and outliers. Use of basic statistics is shown to inform these justifications such as mean, standard deviation and variance. Evidence of several experiments are applied to evaluate the impact that different features have on the target data. Your findings include discussion of the insights produced and a model that is suitable for predicting future trends. You use visualisations to present the effectiveness of the model and use statistical measures such as absolute and mean square error for reference. The model is thus, tested using the 'hold out' approach and evaluated further using cross validation. The reflective report is excellently presented with justifications of decisions, insights from experiments and use of the final model for future predictions.

PREPARATION FOR THE ASSESSMENT

Formative assessment will take place each week. This will not, however contribute to your final mark and is meant as a teaching and learning exercise.

Preparation and development for this assignment will take place during semester two within seminars, tutorials and weekly formative assessment.

RELEASE DATE AND HAND IN DEADLINE

Assessment Release date: **23rd January 2023**

Assessment Deadline Date and time: **20th April 2023, 23:55.**

Please note that this is the final time you can submit – not the time to submit!

Your feedback/feed forward and mark for this assessment will be provided on **15th May 2023.**

SUBMISSION DETAILS

The Practical Assessment assignment including program code and a written reflection of 1000 words (+/-10% expected). All coding examples must be included in your appendices section at the end of the document but must be referenced throughout when discussed. Your appendices do not contribute to your written reflection word count.

Submission of one Word processed document is required through Blackboard using Turnitin. University referencing style should be used throughout, if required.

HELP AND SUPPORT

- You will find information links to all our Library resources in the Library area of the Student Hub. For support with using these resources, please contact your subject librarian at SubjectLibrarians@uclan.ac.uk.
- You can get support with your academic skills (academic writing, critical thinking and referencing) through WISER. For details of the WISER support services go to the [Study Skills section of the Student Hub](#).
- If you have not yet made the university aware of any disability, specific learning difficulty, long-term health or mental health condition, please complete a [Disclosure Form](#). The [Inclusive Support team](#) will then contact to discuss reasonable adjustments and support relating to any disability. For more information, visit the [Inclusive Support page of the Student Hub](#).
- To access mental health and wellbeing support, please complete our [online referral form](#). Alternatively, you can email wellbeing@uclan.ac.uk, call 01772 893020, attend a drop-in, or visit our [UCLan Wellbeing Service](#) pages for more information.
- If you have any other query or require further support you can contact Student Support via studentsupport@uclan.ac.uk. Speak with us for advice on accessing all the University services as well as the Library services. Whatever your query, our expert staff will be able to help and support you. For more information please visit the Student Hub.
- If you have any valid mitigating circumstances that mean you cannot meet an assessment submission deadline and you wish to request an extension, you will need to apply online prior to the deadline.

Disclaimer: The information provided in this assessment brief is correct at time of publication. In the unlikely event that any changes are deemed necessary, they will be communicated clearly via e-mail and a new version of this assessment brief will be circulated.

Version: 1

