Summer Semester Quarter 2

Assessment Brief -

AI and Applications

Summer 2022

Part I: General Information

Module	M504 AI and Applications
Term	Semester 2, Quarter 2
Primary Assessment Task	Individual Final Project
Weighting	70% Primary Assessment Task 15% Online Assessments 15% Class Participation
Distributed on:	W/C July 18, 2022
To be submitted on:	September 22, 2022. 18:00 Berlin Time.
Submission Method	This assignment must be submitted as a Jupyter Notebook (both the *.ipynb and *.pdf files) on the corresponding submission folder to be found on Canvas. You must submit your work accompanied by an Assessed Submission Form, which must be completed in full. The assignment will not be accepted by the Registry unless the form is completed correctly.
Length	The submitted Jupyter Notebook should contain all the required textual and code cells. No specific length or word count criteria need to be met.

Part II: Assessment Details

Primary Assessment Task Topic	You are a data science consultant. Your client company has a dataset and a bunch of business questions. Therefore, you are required to build an exploratory data analysis pipeline in a Jupyter Notebook to answer these business questions. Your designed and implemented pipeline will be submitted to your client company.
Assessment Guidelines	The submitted Jupyter Notebook should contain both textual sections and runnable codes in a rational structure. The texts and codes should be written in a clear and easy-to-follow manner. All the design decisions should be made in a principled manner. In fact, all the choices should be justified in the notebook, either by explaining the intuition or by conducting empirical experiments. Be creative. Get inspired from any public documents (e.g., blogs, documentation, open-source projects) but design and implement your notebook yourself. Reproducing another source will lead to plagiarism issues.

	Please clearly mention the URL of your dataset. It allows us later to rerun your notebook, if necessary, as the dataset is accessible via its URL. Using data visualization is not required but appreciated. The notebook should contain the following information: • The business context. For example, who is your client company? Which kind of data do they have? Why do they need you as a data scientist? • A data exploration discussion on the characteristics of the given dataset. For example, how do distributions of the attributes look like? What are the most interesting metadata? • The data preprocessing step. For example, what are the common data quality issues of the data? What kind of data curation techniques are required? How do you clean their data? • The data analysis part. Define 5-10 interesting business questions and answer them. The questions should be diverse and require different Python/Pandas skills/features. So, please make sure that, overall, the submitted Jupyter notebook showcases your in-depth Python programming skills. Define genuine business questions that you are personally interested in, not some trivial questions that are easy to answer. For example, these business questions could be "What time could be the best time to display our advertising to customers?", "What product combinations were most often sold?", etc. For each business question do the following: o Add a new subsection in the Jupyter Notebook; o Explain the importance of this query for the business; o Write a code cell including Python codes to find the answer; o Add a text cell to explain your Python code, elaborate the result, and how it answers the original business question. • A final discussion and conclusion. For example, what are the overall strengths and limitations of your exploratory data analysis? What are the insights and implications of the results for the business? What are your data-driven recommendations for your client company?
Purpose	Designing and implementing such exploratory data analysis pipelines with Python is one of the required skills for data scientists. This assignment is designed to assess your ability to build such pipelines. We are especially interested to see that you can apply various

	techniques that you have learned in the module in a systematic and principled way.
Links to Module Intended Learning Outcomes	The assignment relates to the following intended learning outcomes for the module: • Understand and evaluate the business contexts that can benefit from advanced Python-based technologies. • Critically design and develop data science projects for various problems using Python and its technologies. • Integrate a range of Python development technologies. • Understand and evaluate Jupyter Notebooks, NumPy, and Pandas as they compare with other technology options.
Special Instructions	
Additional Assessment Components	GISMA University rewards in-class attendance, and engagement with asynchronous content, at a rate of 30% per module. Students attending ≥ 80% (factoring on possible extenuating circumstances) of their synchronous classes as per their due mode of delivery, will gain 15% towards their final module mark. Students successfully engaging with asynchronous material on the gamification/microlearning path and completing all summative assessments in the asynchronous environment, will equally gain 15% towards their final module mark. The above also entails that, students falling below 80% of attendance, although they will be still allowed to submit, will have their final mark capped at 85/100. Equally, if they fail to engage with the asynchronous material and complete the short summative assessments included in specific checkpoints during each term (usually 4), their module mark, irrespective of their engagement and participation in synchronous delivery, will drop by a maximum rate of 15%.

Part III: Marking Criteria / Assessment Criteria

Mark Weight 100%	Fail (0 - 49%) 5,0	Sufficient (50 – 59%) 4,0 - 3,7	Satisfactory (60 – 74%) 2,7-3,3	Good (75-89%) 1,7-2,3	Very Good (90-100%) 1,0-1,3
Marking Criteria	Does not fulfil the requirements of the assessment.	Demonstrates acceptable knowledge and understanding of the subject-matter and	Demonstrates substantial knowledge and understanding of the subject-matter and achievement of learning	Demonstrates a comprehensive knowledge and understanding of the subject-matter and achievement of	Demonstrates a comprehensive knowledge and understanding of the subject-matter

		achievement of learning outcomes at low to average level of performance.	outcomes at average to above average performance levels.	learning outcomes at well above average levels of performance.	and achievement of learning outcomes at high (highest) levels of performance.

Primary Assessment Task Criteria	Your primary assessment task will be assessed based on the following criteria: • The correctness, completeness, and conciseness of runnable codes. (40%) • The structure of the report, quality of writing, and critical evaluation of codes in the text. (30%)
Notes about Marking	

Part IV: Tips for Successfully Engaging with this Assessment

Answer the Question	It may seem obvious, but make sure you are answering the question you have been set, not the question you would prefer to answer. If the brief has a number of tasks or parts, answer all of them. Parts that involve evaluation or analysis are usually longer and worth more marks than parts that ask for description or explanation. Keep the brief in front of you and check it regularly.
How to Use Assessment Criteria	The assessment criteria document is not usually a guide to the structure of your assignment. Each section of the criteria is not a separate paragraph in your assignment, but qualities that you need to demonstrate throughout. Treat the assessment criteria as a checklist at the end not as a plan at the beginning. Also, the criteria document often tells you what to demonstrate (e.g., critical analysis) but not necessarily how to do it. For how to do it, look back at the skills and activities you have covered in the rest of the module. Above all, remember this is not a test of how much you know or how much you have read about the topic. It is a test of how well you can use your knowledge to answer the specific question set.
Planning and Preparation	Make sure you attend the lectures, especially the first and the last one, where we will be 'unpacking' this assignment in greater detail.
Referencing	GISMA Business School requires that students use Harvard Referencing.
Plagiarism and Cheating	Your attention is drawn to the University's stated position on plagiarism. THE WORK OF OTHERS THAT IS INCLUDED IN THE ASSIGNMENT MUST BE ATTRIBUTED TO ITS SOURCE (a list of references and bibliography must be submitted).

	Please note that this is intended to be an individual piece of work. Ensure that you read through your work prior to submission. Action will be taken where a student is suspected of having cheated or engaged in any dishonest practice. Students are referred to the University regulations on plagiarism and other forms of academic misconduct. Students must not copy or collude with one another or present any information that they themselves have not generated. For more information on Plagiarism, please see the relevant section in your Programme Handbook.