MONASH University

Department of Econometrics and Business Statistics

ETW3482 DATA MINING AND PREDICTIVE MODELLING

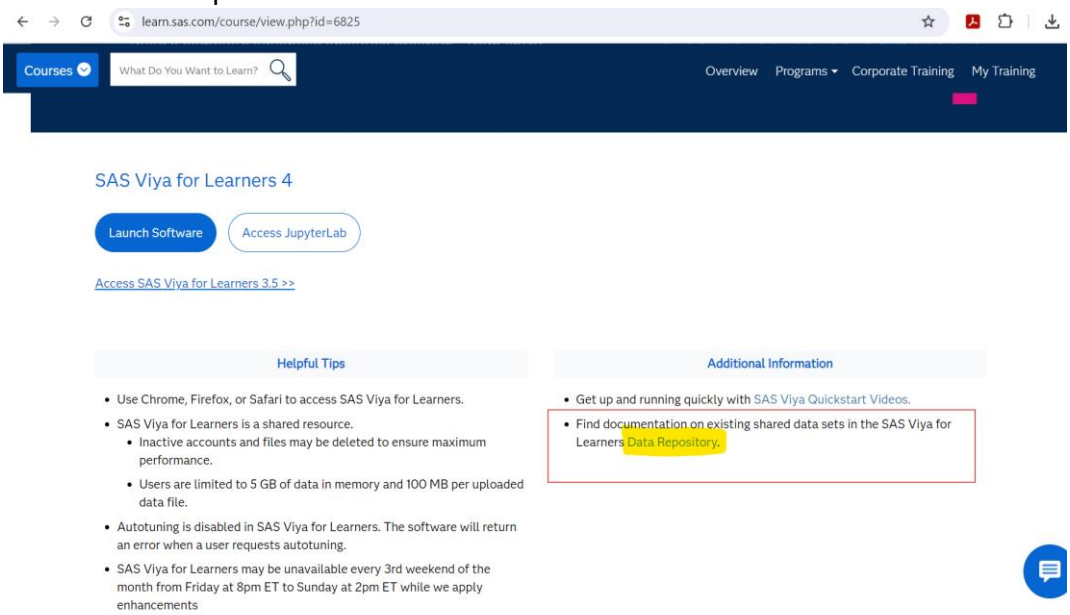
Semester 2, 2024
Assignment: Predictive Modelling Project Proposal
Due date: Monday, 19 August 2024, 11.55 p.m.

This assignment is worth 20% of your overall grade for this unit. The total number of marks for this project is 100. The objectives of this assignment are:

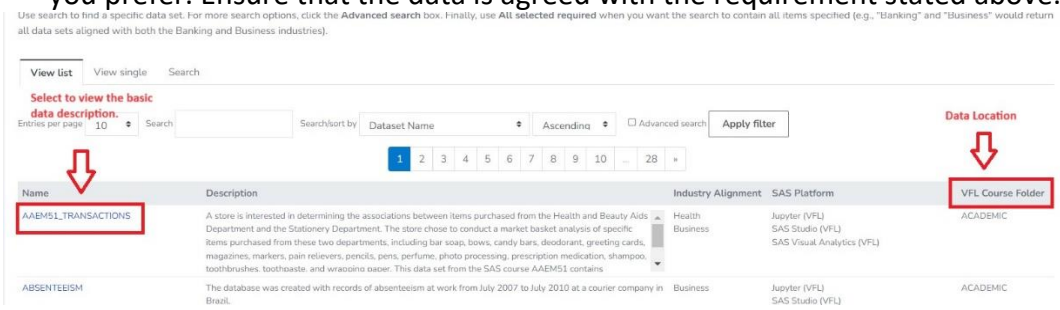
1. Identify data for predictive modelling.
2. Compose a rational analytics question for a business case.
3. Produce a business analytics proposal.

INSTRUCTIONS

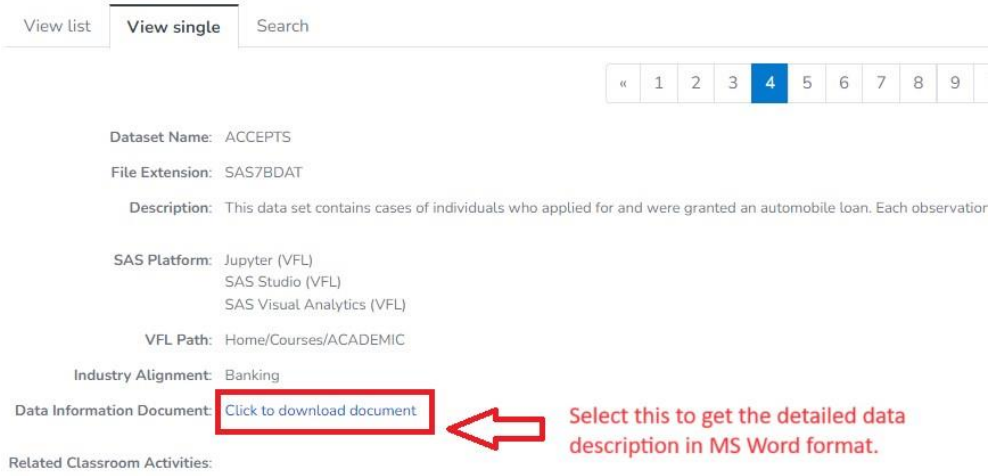
1. Follow the following steps to choose a suitable data set for the predictive modelling project. The selected data must have a **binary target and at least 15 inputs**.
   (i) Go to SAS Viya for Learners 4 Data Repository to identify suitable data and obtain the description of the data.



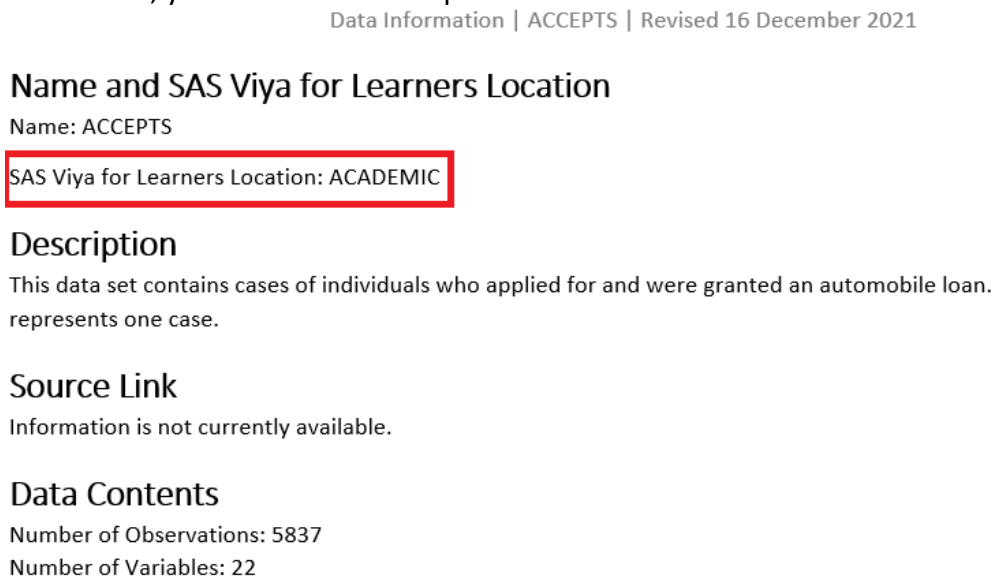
   (ii) Data name, brief description and location can be found in the View List tab in SAS Viya for Learners 4 Data Repository. Use search to get details on the data you prefer. Ensure that the data is agreed with the requirement stated above.



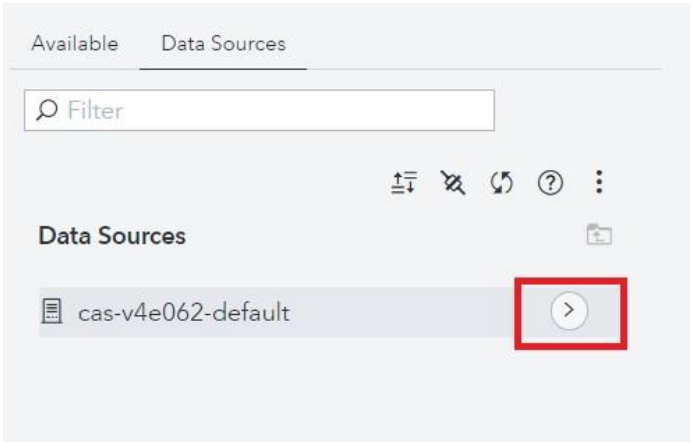
(iii) Select the Variable Name to view the basic description of the data in the View single tab. Select the "Data Information Document" to obtain a detailed data description in Microsoft Word format.

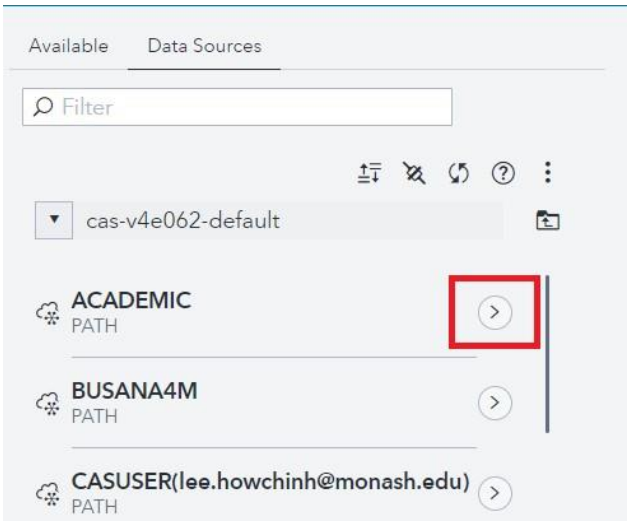


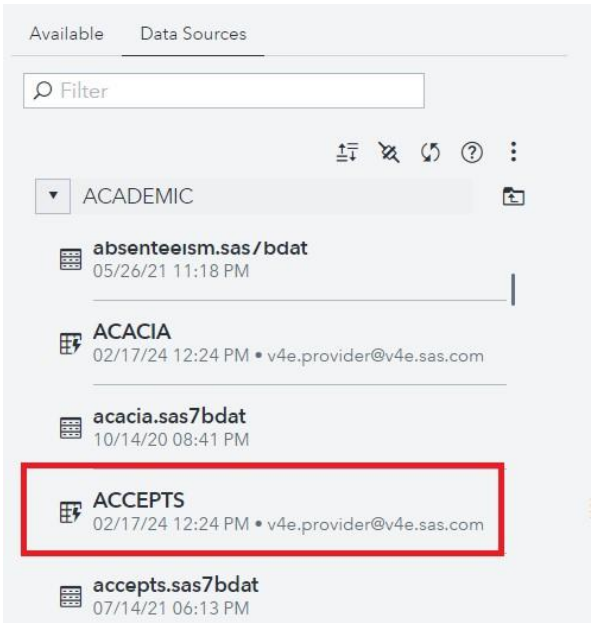
(iv) After identifying suitable data, you can access it in SAS Viya for Learners according to the location in the data description. For instance, the location of the data "ACCEPTS" can be identified as follows in the Microsoft Word document obtained from the above step. The file location is ACADEMIC. Besides the location, you can find other important data information in this document.



(v) The data can be accessed from SAS Viya following the diagrams below.







Avoid spending too much time searching for the data. Focus on the data that meets the requirements stated above, and ensure that the data description meets your expectations.

2) Prepare a predictive modelling project proposal based on the selected data. The proposal must consist of the following components:
    1. **Title Page:**
        - ✓ Include the ETW3482 Assignment 1: Predictive Modelling Project Proposal as the title.
        - ✓ Include your name, student ID and submission date.
    2. **Executive Summary:**
        - ✓ Provide a concise project overview, highlighting key objectives and potential impact.
    3. **Introduction:**
        - ✓ Introduce the context of your predictive modelling project.
        - ✓ Clearly state the purpose and relevance of the project to the identified problem or question.
    4. **Objectives and Goals:**
        - ✓ Clearly define a specific objective of your predictive modelling project.
        - ✓ Make sure your objective is measurable and aligns with the purpose/objective of the project.
    5. **Data Sources:**
        - ✓ Identify and describe the data sources for your project.
    6. **Scope of Work:**
        - ✓ Detail the tasks and activities involved in your project based on the data mining process.
        - ✓ Specify the scope and limitations of your predictive modelling initiative.
    7. **Tools and Technologies:**
        - ✓ List the SAS applications in SAS Viya that you intend to use for your predictive modelling. Justify your choices.
    8. **Timeline:**
        - ✓ Develop a timeline with milestones and deadlines for your project within this semester.
        - ✓ Consider creating a Gantt chart or a timeline diagram for visual representation.
    9. **Risk Management:**
        - ✓ Identify potential risks and challenges in your project.
        - ✓ Propose strategies for mitigating each identified risk.
    10. **Conclusion:**
        - ✓ Summarise your predictive modelling proposal.
        - ✓ Include any specific next steps or actions needed.

3) Writing a proposal for a predictive modelling project requires a combination of technical clarity, persuasive communication, and a structured approach. Here are some key writing skills and style considerations for crafting an effective proposal:
    a. **Clarity and Precision:**
        - Use clear and concise language to convey your ideas.
        - Avoid unnecessary jargon or technical terms that may confuse the readers.
        - Clearly define terms and concepts to ensure a shared understanding.
    b. **Structured Organisation:**
        - Follow a logical structure with clearly defined sections and headings.
        - To enhance readability, use a consistent formatting style for headers, subheadings, and bullet points.
        - Ensure a smooth flow of information from one section to the next.
    c. **Audience Awareness:**
        - Tailor your language and level of detail to your audience's expertise.
        - Assume they have a general understanding but may not be experts in predictive modelling.
        - Clearly articulate the value of the project and its relevance to stakeholders.
    d. **Purposeful Introduction:**
        - Begin with a compelling introduction that grabs the readers' attention.
        - Clearly state the purpose and significance of the predictive modelling project.
        - Set the stage for the reader to understand the context and importance of your proposal.
    e. **Objectivity and Professional Tone:**
        - Maintain an objective and professional tone throughout the proposal.
        - Avoid subjective language and focus on facts, evidence, and a rational argument.
    f. **Detail Orientation:**
        - Provide sufficient detail to demonstrate a thorough understanding of the project.
        - Clearly explain the methodology, data sources, and tools you plan to use.
        - Anticipate potential questions and address them pre-emptively.
    g. **Consistent Terminology:**
        - Use consistent terminology and definitions throughout the proposal.
        - Ensure that your chosen terms align with industry standards and commonly accepted practices.
    h. **Visual Aids and Examples:**
        - Include visual aids, such as charts or graphs, to support your points.
        - Use examples to illustrate complex concepts and enhance understanding.
    i. **Engaging Conclusion:**
        - Summarise the key points in a compelling conclusion.
        - End with a call to action or a clear indication of the next steps.
    j. **Proofreading and Editing:**
        - Thoroughly proofread your proposal to identify grammatical errors and typos.
        - Ensure consistency in formatting, style, and tone throughout the proposal.

k. **Coherence and Cohesion:**
   - Ensure that each section of the proposal flows logically from the previous one.
   - Use transitions to connect ideas and maintain a cohesive narrative.

l. **Conciseness:**
   - Be concise in your writing, avoiding unnecessary repetition or verbosity.
   - Clearly convey your ideas without overloading the reader with information.

By incorporating these writing skills and style considerations, you can create a compelling and well-structured proposal that effectively communicates your predictive modelling project to your intended audience.

4) Proper citation is an essential aspect of academic writing. As you prepare your project, please ensure you cite all sources used in your work, including direct quotes, paraphrases, and ideas borrowed from other authors or generative AI. Failure to cite sources properly could result in charges of plagiarism, which could negatively affect your marks for this assessment. Use APA citation style. You can refer to this webpage for the APA citation style and generate the APA citation using this APA referencing generator. You can refer to this Monash page and this webpage for citing resources obtained from generative AI and nonrecoverable sources.

5) Any students caught plagiarising or permitting others to plagiarise their work will receive a zero mark on this project. Students should be aware of what constitutes plagiarism and collusion, and the procedure should one be suspected of committing such acts.

6) Students should emphasise the narration and how the results are presented and interpreted. Students should endeavour to ensure that the report is complete and well-composed. **Poor presentation, poor command of English writing and/or failure to comply with instructions may result in a mark penalty.** You are encouraged to access Studiosity to improve your report writing.

*Please note that the services available for you in Studiosity (accessible via Moodle site) are supplementary to this unit. Studiosity is a third-party provider contracted by Monash University to assist you with generic skills such as essay writing, grammar, referencing, etc. They do not provide specific comments on unit content or the appropriateness of your answer regarding assessment tasks and learning outcomes. Rather, they address your key skills of argument, structure, expression, and referencing.
Evaluation of your work for assessment purposes is conducted solely by your Monash teachers (chief examiners or tutors). You should use consultation hours provided by Monash teachers if you have concerns or questions about unit content and your understanding of that content or if you have questions specifically about assessment tasks.

7) Your report should not exceed **1000 words** (excluding diagram and table titles, Reference List and Appendix).
    (i) Use default format, paragraph and margin settings (These settings are in default mode whenever you open a new Word document)
    (ii) Font type: Times New Roman. Font size: 12.
    (iii) 1.5 line spacing between lines.
    (iv) **All diagrams should be numbered and labelled and in line with the text** for ease of reading and not placed in an Appendix at the end of the report.

8) All submissions will be via Moodle. Submit your proposal as a **PDF document**.

Additional Notes:

- You are encouraged to seek feedback from the Chief Examiner, your tutor or peers during the proposal development process.

- The proposal should reflect a clear understanding of the data mining process and predictive modelling concepts covered in weeks one and two.