|  |  |
| --- | --- |
|  | School of Applied Business **Bachelor of Accounting** |

**ACCY6105 Data Analytics**

**Individual Assignment Semester 2, 2022**

|  |  |  |
| --- | --- | --- |
| **Date Issued:** | Week 7 | |
| **Due date and time:** | Week 13**, 5/11/2023, 11:59 pm** | |
| **Delivery:** | Turnitin via Moodle | |
| **Total Marks:** | 100 marks | |
| **Weighting:** | 50% | |
| **Expected number of hours to spend on this assessment:** |  | |
| **Word limit:** | NA | |
| **Instructions:** | Complete this **cover sheet** and attach itto your assignment.  Collusion, copying or plagiarism may result in disciplinary action. | |
| **Lecturer:** |  | |
| **Individual declaration:** | I confirm that:   * My submission is original and entirely my own work. * Where I have used ideas, tables, diagrams etc. of other writers, I have acknowledged the source in every case using APA7 style. * My submission has not been submitted as assessed work for any other academic course. | |
| **Student name:** | **Student ID:** | **Student Signature** |
|  |  |  |

|  |  |
| --- | --- |
| **This assignment relates to the following Learning Outcomes:** | |
| 2. | Analyse data requirements and implement queries in an organisational context. |
| 3. | Apply statistical tools and techniques to support decision-making in an organisational context. |
| 4. | Discuss the ethical implications associated with data utilisation in an organisational context. |

|  |
| --- |
| **Aim of the Assignment** |
| The aim of the assignment is to develop the knowledge, skills, and competencies relevant to business data and decision-making. |

|  |
| --- |
| **Assignment Instructions** |
| For this assessment, you are allowed to use Generative AI tools such as ChatGPT, Bard, Grammarly Go etc to:Clarify concepts, theories, ideas, etc, during your assessment preparation.Spelling/grammar checkers.You are NOT allowed to use Generative AI to:Generate definitions or writing used in your final report.Produce counterarguments or refine thinking on your final submission.Write any code that is needed for the assessment. **Assessment Components**   |  |  |  |  | | --- | --- | --- | --- | | **Assignment Description** | Due Date | Marks | Percentage (overall) | | Part A: Data preparation | Week 12 | 20 | 20% | | Part B: Data analysis and report | Week 12 | 80 | 80% |  You will be provided an Access database containing two tables, a financial data table and a company table. The financial data table includes 190+ financial indicators, such as revenue, cost of revenue, gross profit, R&D expenses, free cash flow margin, EBITDA, etc., for 3500+ stocks within the US stock market from 2014 to 2018. The financial indicators have been scraped from Financial Modelling Prep API and are those found in the 10-K filings that publicly traded companies release yearly. The class column (distinct from financial indicator) in financial data table is your target variable which represents,if the value of a stock increases during that year, then class = 1.if the value of a stock decreases during that year, then class = 0.For example, stocks that belong to class 1 are stocks that one should buy at the start of year 2015 and sell at the end of year 2015.Part A: Data preparation (20 marks)Before commencing the analysis, kindly complete the following tasks:Randomly Select a set of 100 company stocks along with 15 financial indicators according to your preferences from the provided database. Upon completion of the data selection, store the chosen data within a new table in Access. Then, demonstrate your final data selection with your steps in the Word (including the list of 15 chosen financial indicator, the list of selected company and overall row count in your joint dataset). (8 marks)Examine the ethical implications of data utilisation concerning data accuracy and bias and provide two comments in Word. (6 marks)Use Power Query or Orange to clean the selected data and provide a comprehensive statement about your data cleaning procedures in the Word, include step-by-step explanations. (6 marks)Part B: Data analysis and report (80 marks)In this part, kindly complete following tasks:Employ Excel or Orange to perform an in-depth exploratory data analysis (EDA) on the chosen data in part A and generate visualizations to present the findings and insights from your analysis. The outcomes should integrate into your report.Ascertain the top 10 most significant financial indicators from 15 financial indicators in part A with respect to the target variable. And then, contract a predictive machine learning model (decision tree) in Orange to understand whether it is possible to classify the future performance of a stock by looking at the chosen financial indicators. Ensure that the model performance outputs and comments regarding both the model and its outputs are integrated into your report.The report should be no more than 1,000 words and no less than 900 words (not include cover page, table of contents, reference, and Appendix). There will be penalty for exceeding the maximum word limit specified.You are free to generally follow the report structure below,  1. Cover page 2. Introduction 3. Visualization & Data Analysis 4. Investigations and Findings 5. Conclusion 6. References 7. Appendix (if applicable)  Please refer to the Grading Rubric on the mark allocation and as a guide to ensure you satisfy the criteria for the Assignment. **Note:**   * Download the database files by clicking on Assessment 2 folder on Moodle. * Submit your **Microsoft Word file** withthe format "**Word\_YourName\_Student ID**" to the IndivAssignment\_Word link on Moodle. * Submit your **Microsoft Excel file** with the format “**Excel\_YourName\_Student ID”** to IndivAssignment\_Excel link on Moodle. * Submit your **Orange file** with the format “**Orange\_YourName\_Student ID**” to IndivAssignment\_Orange link on Moodle. * A word limit is recommended to guide you on the length required. Marks may be deducted for exceeding the word limit. * Use supporting evidence, referenced, to support your comments. Use the APA7 referencing style to acknowledge all sources. For guidelines on APA7 referencing go to <http://libguides.unitec.ac.nz/apareferencing>. * Formatting should be Calibri 12 pt font / 1.5 line spacing, all pages numbered. * Please note Unitec’s late assessment policy detailed in the Course Information Sheet on Moodle. |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Marking Rubric for Individual Assignment (report)** | | | | | | | | |
| Course code and title | | ACCY6105: Data Analytics | | Assessment | | Individual Assignment | | |
| Ākonga name | |  | | Ākonga ID number | |  | | |
| **Performance Criteria** | **Level of Achievement** | | | | | | |  |
| **Achieved with excellence**  **(≥80%)** | | **Achieved with merit**  **(60-79%)** | | **Achieved**  **(40-59%)** | | **Not achieved**  **(≤39%)** | **Mark** |
| Purpose and objective | The purpose and objective of the report is made clear, and the report addresses the objective in a focused and logical manner. | | The purpose and objective of the report is made clear, and the report addresses the objective | | The purpose and objective of the report is vaguely discussed, and the report partially address the objective | | The report does not clearly address the objectives. | /10 |
| 8 ≤ mark ≤ 10 | | 6 ≤ mark < 8 | | 4 ≤ mark < 6 | | 0 ≤ mark < 4 |  |
| Data analysis, | Data analysis methods are fully and correctly applied. | | Most data analysis methods are fully and correctly applied, with only 1 – 2 instances of misapplications. | | Some data analysis methods are applied but with 3 – 4 significant errors or omissions. | | Nearly all data analytics methods are either misapplied or absent. | /20 |
| 16 ≤ mark ≤ 20 | | 12 ≤ mark < 16 | | 8 ≤ mark < 12 | | 0 ≤ mark < 8 |  |
| Findings, and discussion | Findings and discussion are clearly oriented to the purpose. It is organised in a considered, relevant manner, and leads the reader logically from the findings to the recommendations and conclusions. | | Discussion shows alignment with purpose but needs improvement for effectiveness. Justifications and explanations for conclusions and recommendations are present, but clarity of connections could be enhanced for the reader. | | Discussion somewhat aligned with purpose but lacks effectiveness. Provides justifications and explanations for conclusions and recommendations, yet consistently vague to the reader. | | Discussion is poorly organised and leaves the reader wondering how the conclusions and recommendations were made. | /20 |
| 16 ≤ mark ≤ 20 | | 12 ≤ mark < 16 | | 8 ≤ mark < 12 | | 0 ≤ mark < 8 |  |
| Visualization | * Visualizations effectively address topics with conciseness and clarity. * Legend describes every graphic variable type present in the visualization. | | * Most visualizations effectively address topics with conciseness and clarity, with only 1 - 2 visuals that appear unclear and lack conciseness. * Legend describes graphic variable type present in the most visualization. | | * Most visualizations somewhat address topics, with only 3 - 4 visuals that appear unclear and lack conciseness. * Legend describes only a few of graphic variable type present in the visualizations. | | * Visualizations poorly address topics, appear unclear and lack conciseness. * Either there is no legend, or it does not describe any of the graphic variable types present in the visualization. | /20 |
| 16 ≤ mark ≤ 20 | | 12 ≤ mark < 16 | | 8 ≤ mark < 12 | | 0 ≤ mark < 8 |  |
| Structure and presentation | * The report demonstrates a well-organized structure that is effortlessly comprehensible. The flow of paragraphs is smooth and coherent, making it easy for the reader to follow the content. * Very few spelling errors (less than 2), no grammatical errors, grammatically correct and complete sentences. | | * The report is structured, and written expression is generally good and clear. * few spelling errors (less than 5), limited grammatical errors (less than 5) that don’t interfere with communication. | | * The report is adequately structured, and the written expression is generally satisfactory. * Some spelling errors (less than 10), limited grammatical errors (less than 10) that don’t interfere with communication. | | * The report lacks proper structure, and the paragraphs suffer from unclear and vague content. * Many errors in spelling and grammar that interfere with communication. | /10 |
| 8 ≤ mark ≤ 10 | | 6 ≤ mark < 8 | | 4 ≤ mark < 6 | | 0 ≤ mark < 4 |  |
| **Total** | | | | | | | | **/80** |

**Help with your Assignment**

To do the best that you can do on this assignment you may:

* + Talk it over with your lecturer, or Academic Programme Manager.
  + Visit the “Ask Me” desk in Te Puna (B180, Level 1) to book learning advice and support or Ph 0800 10 75 10, email [learningadvisors@untiec.ac.nz](mailto:learningadvisors@untiec.ac.nz)
  + Seek counselling or medical advice at the Medical Centre (B180 Te Puna, level 0). Ph 0800 10 85 10, email [studentwellbeing@unitec.ac.nz](mailto:studentwellbeing@unitec.ac.nz)
  + Visit the Maia Marae team (B180, Level 1, Room 1001) for academic support for Māori.
  + Visit the Pacific Centre (B180, Level 1) for Pacific Development and Support, o r email [askpacific@unitec.ac.nz](mailto:askpacific@unitec.ac.nz) (the Pacific Centre is in Building 180, Level 1)
  + Contact Student Support for general advice and support 0800 10 75 10, email [studentsupport@unitec.ac.nz](mailto:studentsupport@unitec.ac.nz)
  + Students with learning disabilities can receive support from Access4Succes by emailing [disabilityservices@unitec.ac.nz](mailto:disabilityservices@unitec.ac.nz) or phone 021 0221 1686.