TUNKU ABDUL RAHMAN UNIVERSITY OF MANAGEMENT AND TECHNOLOGY

FACULTY OF COMPUTING AND INFORMATION TECHNOLOGY

ACADEMIC YEAR 2025/2026

BMDS2003 DATA SCIENCE ASSIGNMENT

COURSE NAME	**:**	**Data Science**
COURSE CODE	**:**	**BMDS2003**
SESSION	**:**	**202505**
TOTAL MARK	**:**	**100%**
WEIGHTAGE TO FINAL MARK	**:**	**42%**
Group	**:**	**3 to 4 members**
SUBMISSION DEADLINE	**:**	**Week 12 (Monday, 8:30 am)**
PRESENTATION	**:**	**Weeks 12, 13, and 14**

Learning Outcome Being Assessed:

CLO1:	Apply appropriate data science concepts, statistics, and machine learning to make predictions on data (C3, PLO2).
CLO2:	Practice exploratory data analysis and visualisation methods with the aid of appropriate data science tools (P3, PLO3).
CLO3:	Analyse relevant techniques, statistical methods, and Machine Learning algorithms for data analytics (C4, PLO7).

The following is a list of datasets:

1. Obesity Levels
2. Drug Consumption
3. Communities and Crime
4. Traffic Flow
5. Aids Clinical Trials
6. Productivity of Garment Employees
7. Diamonds
8. Heart Disease
9. Online Gaming
10. Taiwanese Bankruptcy
11. Tezpur University Android Malware
12. Seoul Bike Sharing Demand

Group Assignment Instructions

Choose **ONE (1)** dataset from the list provided for your group assignment. Click the link to download the dataset.

Important:

- Each dataset is only allowed to be selected by one group per practical session. **FIRST COME, FIRST SERVED.**
- You may refer to the Google Spreadsheet provided by your tutor to check dataset availability and register your selection.
- This assignment should cover a wide range of content discussed in the lectures and practical sessions.

You will need to:

- Present your assignment to the class.
- Submit a report that outlines a strategic plan to utilise the available data to help make better business decisions.

Assignment Guidelines

1. **Project and Business Understanding**
   - Define the business problem and clarify customer goals from a business perspective.
   - Balance competing objectives and constraints to uncover key factors influencing project outcomes.
   - Develop a strategic plan to address the identified business problem using data analytics.

2. **Data Understanding**
   - Acquire relevant data resources and describe the data collection process.
   - Load and integrate data from multiple sources where necessary.
   - Analyse and summarise the characteristics of the dataset, including variables and data types.

3. **Data Preparation**
   - Clean and transform raw data to prepare it for analysis.
   - Address data quality issues, including handling missing values and outliers.
   - Standardise data formats and enrich source data where necessary.
   - Document the preprocessing steps in the report.

4. **Modelling**
   - Select and apply suitable modeling techniques based on the project objectives.
   - If multiple techniques are applied, perform and document each task separately.
   - Justify the choice of models and describe the configuration of parameters used.

5. **Evaluation**
   - Assess the performance of the models and determine whether they meet project objectives.
   - Conduct testing on real-world scenarios to evaluate the robustness and accuracy of the models.
   - Provide a critical assessment of the model's strengths and limitations.

6. **Deployment Strategy**
   - Develop and propose a strategy for deploying the model results in a business context.
   - Outline potential challenges and recommend mitigation measures for effective deployment.

7. **Dashboard and Visualisation**
   - Create an interactive and informative dashboard to present key findings.
   - Use appropriate visualisation techniques to communicate insights clearly.
   - Justify the choice of visualisations and describe their relevance to stakeholders.

8. **Conclusion**
   - Summarise the key findings of the project, emphasising the advantages and limitations of the models used.
   - Highlight the overall contribution of the project to business decision-making.
   - Reflect on the lessons learned and potential improvements for future projects.

9. **References**
   - Include proper in-text citations and a comprehensive list of references using the APA referencing system.

Submission Requirements:
Written Report: A well-structured and detailed report covering all the sections outlined above.
Presentation: A 10-15 minute group presentation summarising the key aspects of the project.

Late Submission:
No late assignments will be accepted (get zero). Please **DO NOT** argue with your tutor if you really failed to submit your assignment on time as the consequence of late submission has been given in advance. However, in certain circumstances, the students may be allowed to turn in the assignment late. With the exception of Extenuating Mitigation Circumstance (EMC) reasons, penalty for the late submission shall be imposed after submission deadline/extended submission deadline.
The penalty is as follows:

Late submission within 1 - 3 days: Total marks to be deducted is 10 marks.
Late submission within 4 - 7 days: Total marks to be deducted is 20 marks.
Late submission after 7 days: Reject coursework and zero marks shall be awarded.

Note: For EMC cases, students need to apply for the Leave of Application from the FOCS office as a record of proof of MCs, Death Certificate. SPM Examination, MUET Examination, etc. is given.

No-Cheating Policy:
A reminder on the no-cheating policy: You are **NOT ALLOWED** to share your work with your peers, but please feel free to discuss with your peers. If cheating is discovered, both parties will take the equal blame (get zero). Please note that the assignment should be your own work, although you may incorporate ideas or techniques from books, online resources, etc. Copying materials directly from any sources of materials will lead to zero. You have been warned. Whenever you face any problems, please seek advice from your tutor.