University of Greater Manchester

Greater Manchester Business School

Module Code	MBA7064
Module Title	Digital Business and Data Analytics
Module Instructor	Dr. Adnan Shoaib
Word Count	5000 words
Assessment Type	Summative Assessment
Date of Announcement	25th November, 2025
Submission Deadline	31st December, 2025
KIS Weightage	100%
Marks	100 pts

Notes for Students:

Assessment Instructions

The submission deadline for this assessment is **23:00 PST on Wednesday, December 31, 2025**. Kindly be advised that you may submit this assessment prior to the deadline.

A deduction of five marks will be imposed for tardy submissions. The repercussions of non-submission are significant and may entail severe penalties. If you cannot finish your open assessment by the specified submission date due to Exceptional Circumstances, you may request an extension. In the event of unforeseen and extraordinary circumstances, you must promptly seek assistance and furnish evidence at the time of the incident. Applications must be submitted prior to the deadline for consideration.

Instructions for completion and submission of written assignment:

- ➢ This assessment constitutes **100%** of the module's total marks.
- ➢ You are required to submit your assessment electronically via Moodle.
- ➢ The word count is **5000 words**, with a permissible variation of +/- 10%. Exceeding the word limits is not advisable.
- ➢ The words listed in the bibliography are not included in the word count.
- ➢ Utilise exclusively Harvard referencing.

➢ All cited websites must be properly referenced; references from any "pedia" sources (such as Wikipedia or Investopedia) are prohibited.
➢ The assessment text must be presented in a 12-point Arial font with a minimum line spacing of 1.5.
➢ Assignment must be **Microsoft word Processed**

Academic Integrity

When submitting work for evaluation, it is imperative to comprehend the University's standards for upholding academic integrity and to ensure that the work adheres to university regulations. Examples of unacceptable behaviour include plagiarism, reusing previously assessed work, collusion with others, and purchasing assignments from third parties. For further details regarding academic offences, poor academic conduct, and associated penalties, please consult chapter four of our academic regulations.

Please Note that all Assessment of this module allows partial plagiarism and AI content with upper **limit for partial plagiarism < 10% and for AI < 20%.**

Return of Marked Work

You can expect to have feedback returned to you within two weeks after the submission of assessments.

Assessment Details
(Learning Outcome 1, 2, 3, 4)

This assessment involves a practical task. You are asked to source a set of data, clean, manipulate, and use it to produce insights that would be useful to a specific audience or for a defined purpose. You need to produce a report of the process undertaken and the tools you have used for collecting, processing, and analysing data. For this assignment, each student is required to:

1. *Select an Appropriate Dataset. (LO1)* Begin by carefully selecting a dataset to apply for this assignment. Your selection can be from any sources so long as there are no copyright restrictions that limit the use of the data. The dataset(s) you select should be those that you think are interesting to a particular group of people. You can find your own sources and take some guidance from lecture material. Other sources of useful datasets might be:
   - Google Dataset Search - https://datasetsearch.research.google.com
   - www.Data.gov.uk
   - https://data.gov
   - UK Data Service - https://ukdataservice.ac.uk
   - World Bank DataBank - https://databank.worldbank.org
   - OECD Data Explorer - https://data-explorer.oecd.org/
   - Eurostat Database - https://ec.europa.eu/eurostat/web/main/data/database
   - Etc.

   You can scrape data from public websites where this is appropriate - but we are looking for rather large datasets (more than 500 observations/records with more than 6-7 variables/features) as a key element of your Data Story, not just a few numbers.

   Avoid redundancy by ensuring that the chosen dataset differs from the one utilized in your class activity. In other words, refrain from using the same dataset for Assignments.

2. *Identify the Target Audience (LO4).* Clearly define the target audience for your visualizations. Explain why this audience would be interested in the data and how they are expected to use it once provided.

3. *Prepare the Dataset.* The dataset you find may need restructuring, cleaning and editing to improve its quality and suitability for your purpose. You may use any tool to clean the data, including (but not necessarily limited to) Python, Excel etc.

4. _Perform Exploratory Data Analysis (LO3)_ (EDA) to understand and summarize the key characteristics of a dataset, identify patterns or anomalies, and prepare the data for modelling. Use Excel or Tableau (or both) to explore the data. Your EDA should include a range of visualizations, such as:

   - Histograms to examine distributions of numeric variables.

   - Boxplots to detect outliers and understand the spread of data.

   - Scatterplots to explore relationships between pairs of variables.

   - Bar charts, line charts, or Pivot Charts to analyse trends or compare categories.

   - Summary tables to present counts or percentages of each categorical variables, and descriptive statistics like mean, median, mode, and standard deviation for numerical variables.

   - Other relevant visual tools based on the characteristics of your dataset. The primary objectives of your EDA are to:

   - Gain a good understanding of the variables, including the main characteristics of each variable (e.g., distributions, central tendencies, and variability).

   - Identify patterns, trends, and relationships between variables.

   - Detect missing values, outliers, or anomalies in the data and propose strategies to handle them.

5. _Select and build appropriate data modelling (LO2)._ The choice of model depends on the nature of the business problem, the goals of the analysis, and the structure of your dataset. This may involve:

   Regression model:

   - Purpose: Analyse and predict numerical dependent/outcome variable based on independent variables/predictors.
   - Examples: Predicting sales revenue, customer lifetime value, or housing prices.
   - Approach:
     - Select appropriate regression techniques such as multiple linear regression or non-linear regression models
     - Evaluate the model's performance using metrics like R-squared and interpret coefficients.

Times-series model:

- Purpose: Understand and model patterns, trends, and temporal dependencies in time-ordered data.
- Examples: Forecasting sales, stock prices, demands, or website traffic trends.
- Approach:
  - Decompose the time series into trend, seasonality, and residual components.
  - Apply models like Moving Average, Exponential Smoothing, and Regression-based forecasting.
  - Evaluate predictions using metrics like Mean squared error (MSE) or Mean Absolute Percentage Error (MAPE).

Classification model:

- Purpose: Analyse and predict categorical dependent/outcome variable based on independent variables/predictors.
- Examples: Fraud detection, customer risk classification, or churn behaviour.
- Approach:
  - Use classification algorithms such as k-Nearest Neighbours, logistic regression or decision trees.
  - Evaluate model performance with metrics like accuracy.

Unsupervised model:

- Purpose: Identify natural groupings or patterns within the data without predefined labels.
- Examples: Market/Customer segmentation, grouping similar products, market-basket analysis or text mining.
- Approach:
  - Use unsupervised algorithms like K-Means, hierarchical clustering, sentiment analysis.
  - Evaluate cluster validity using metrics such as silhouette score or Elbow method
  - Visualize clusters using scatter plots, dendrograms, or silhouette plots.

Structure of your data analytics report (5,000 words maximum):

Chapter 1- Business Understanding:

- Express the primary goal of the data analytics.
- Explain social or business (or organizational/industrial) problem that is being addressed,

Chapter 2- Data Understanding:

- Explain why you chose the dataset(s) you did. You must provide a link to the dataset(s) used.
- Describe the data: its size in terms of number of records (observations) and variables (features). Provide data dictionary, including the variable names, formats (e.g., numeric, categorical), descriptions, and examples of data values.

Note: Ensure the dataset is different from the one used in your assignments to avoid redundancy.

Chapter 3- Data Preparation:

- Explain the process you used to clean, edit or constructing the data. If you discarded any data say why this was done.
- If you merged or integrate datasets, explain how and why you did this.
- Describe what problems you encountered and how you overcame them.

Chapter 4- Exploratory Data Analysis:

- Use the methods of descriptive analytics and visualisation (such as crosstabulation, histogram, bar charts, line charts, scatterplots, heat maps, Pivot Charts, etc) to explore the data.
- Explain and interpret the results of the visualizations, highlighting trends, patterns, relationships, or anomalies. Discuss the implications of these findings in the context of the business or problem at hand.

Chapter 5- Modelling:

- Depends on the business problem and dataset, apply regression, times series analysis, classification or clustering techniques to build business analytics model(s).
- Explain and interpret the results of the model(s) in relation to the business objectives.
- If applicable, include comparisons of multiple models to identify the best performing approach.

Submission guidance:

1. Ensure each chapter flows logically and builds upon the previous one and keep the report concise, focusing on key insights and actionable findings.
2. Include the screengrabs of data visualisations in your word submission to help give the word document context.

3. Provide links to the source dataset(s) you used - otherwise we cannot audit the validity of your data, and you will drop marks.
4. Your report should be submitted in the form of a Word document, use minimum 12pt font, and at least 1.5 line spacing.

Module Learning Outcomes:

This assignment tests the following module learning outcomes:

- Collect, analyse and interpret data analytics to make informed business decisions.
- Appraise how digital business and data analytics can be used to generate actionable insights for managers and decision-makers.
- Communicating, presenting and disseminating analysis of the data.

Grading Criteria

Your work will be assessed on these Criteria

S. No.	Section			Marks
1	Data Selection	-	Have relevant dataset(s) been selected and are useful for the audience identified?	10
		-	Has this selection been justified?	
		-	Is the dataset interesting and substantial (more than 500 observations/records with more than 6-7 variables/features)?	
2	Business Understanding	-	Has the primary goal of the data analytics been clearly expressed and relevant?	5
		-	Has the social/business/industrial problem been clearly explained?	
		-	Has the target audience been clearly identified? and has the purpose for which they might use the data analytics results been well-articulated?	
3	Data Understanding	-	Has the data been described in terms of size (number of records and variables)?	5
		-	Has the data dictionary been provided, including variable names, formats, descriptions, and examples of data values??	
		-	Is the dataset different from the one used in the group assignment?	
4	Data Preparation	-	Has the data cleaning and editing process been thoroughly explained?	20
		-	Have appropriate techniques/tools been used to edit and clean the data? Is it clear why these techniques/tools were used?	

		-	If any data was discarded or merged, has the reasoning been provided?	
		-	Have challenges encountered and solutions implemented been described?	
5	Exploratory Data Analysis	-	Have appropriate methods of descriptive analytics and visualization been used to explore the data?	30
		-	Have the results of the visualizations been explained and interpreted effectively?	
		-	Have the implications of the findings been discussed in the context of the business or problem at hand?	
6	Modelling	-	Have the appropriate data modelling techniques been selected based on the business problem and dataset?	30
		-	Have the results of the model(s) been explained and interpreted in relation to the business objectives?	
Total		-		**100**