UTHM
Universiti Tun Hussein Onn Malaysia

UNIVERSITI TUN HUSSEIN ONN MALAYSIA

FINAL EXAMINATION
SEMESTER II
SESSION 2023/2024

COURSE NAME	:	DATA QUALITY AND GOVERNANCE
COURSE CODE	:	BWB 11403
PROGRAMME CODE	:	BWB
EXAMINATION DATE	:	JULY 2024
DURATION	:	2 HOURS 30 MINUTES
INSTRUCTIONS	:	1. ANSWER ALL QUESTIONS

2. THIS FINAL EXAMINATION IS CONDUCTED VIA
   ☐ Open book
   ☒ Closed book

3. STUDENTS ARE **PROHIBITED** TO CONSULT THEIR OWN MATERIAL OR ANY EXTERNAL RESOURCES DURING THE EXAMINATION CONDUCTED VIA CLOSED BOOK

THIS QUESTION PAPER CONSISTS OF **FOUR (4)** PAGES

TERBUKA

Q1 Ensuring good data is crucial for businesses to operate smoothly. The organization needs to make sure their data is accurate and reliable. That is where a solid data quality strategy comes in.

(a) Explain **FOUR (4)** strategies that businesses can develop to create a comprehensive data quality strategy and optimize their operations.

(8 marks)

(b) Discuss **FOUR (4)** essential elements of a robust data quality plan and how they impact improving organizational performance.

(8 marks)

(c) Provide **ONE (1)** example to illustrate how companies can address typical hurdles in ensuring data integrity across different phases of data management.

(4 marks)

Q2 Data quality management is the cornerstone of any successful business operation in today's data-driven world. Ensuring that data is accurate, reliable, and consistent is not just a matter of good practice, it is essential for making informed decisions, maintaining customer trust, and achieving organizational goals.

(a) Briefly explain data quality management.

(2 marks)

(b) What makes a proactive approach different from a reactive one in data quality management?

(2 marks)

(c) Clarify the potential drawbacks of relying solely on a reactive approach to data quality management.

(2 marks)

(d) Explain technology's roles in supporting proactive and reactive data quality management strategies.

(2 marks)

(e) Describe the **FIVE (5)** phases of the Six Sigma data quality management DMAIC process and how each phase improves data quality within an organization.

(10 marks)

Q3   Data cleansing is a critical aspect of data quality management. Data cleansing, also known as data cleaning or scrubbing, is identifying and correcting errors, inconsistencies, and inaccuracies in a dataset to improve its quality and reliability.

(a)   Explain **FOUR (4)** data quality techniques involved in data cleansing.

(8 marks)

(b)   Identify scenarios in which data cleansing becomes the only solution for resolving data quality issues, outlining the necessity for its application in each case. Provide examples to illustrate these scenarios.

(9 marks)

(c)   Describe how data parsing contributes to improving data quality, particularly in handling unstructured or semi-structured data.

(3 marks)

(d)   **Table Q3.1** displays a list of addresses stored in a single data field. Refer to this table to answer the questions below.

Table Q3.1

Address
20, Jalan Murni 1, Tmn Murni, 83000, Batu Pahat, Johor Darul Takzim, Malaysia
24, Jln Kemboja 5, Taman Kemboja, BP 83000, Johor, Malaysia
80, Jalan U1/5, Taman Perkasa, KL 53000 Wilayah Persekutuan Kuala Lumpur, Malaysia
56, Jalan Pudu, Taman Maluri, 55100 KL, Malaysia
15, Jalan Parameswara, Taman Bukit Katil, 75450 Melaka, Malaysia
12, Persiaran Perdana, Precinct 1, 62000 Putrajaya, Malaysia

(i)   Perform data parsing by decomposing and transforming address data in **Table Q3.1**.

(6 marks)

(ii)   Construct data standardisation based on data parsing in **Q3(d)(i)**.

(6 marks)

Q4   Data quality myths are misconceptions or misunderstandings about data quality that may lead to ineffective data management practices or decision-making.

(a)   One of the myths is that data cleansing is a one-stop solution for all data quality issues.

(i)   How does this misconception hinder the development of comprehensive data quality management practices? Provide a detailed explanation.

(5 marks)

(ii)   Give a real-world example to support your explanation.

(2 marks)

(b)   The myth that data quality is solely the responsibility of the IT department creates challenges for fostering a culture of data ownership and accountability across different departments within an organisation. How can organisations overcome this misconception to promote collaborative data management practices?

(3 marks)

- END OF QUESTIONS -

4