



---

**UNIVERSITI TEKNOLOGI MARA**  
**TEST**

---

<b>COURSE</b>	:	<b>PRINCIPLES OF DATA MANAGEMENT</b>
<b>COURSE CODE</b>	:	<b>ICT550</b>
<b>DATE</b>	:	<b>JUNE 2025</b>
<b>TIME</b>	:	<b>08.30 AM – 10:30 AM</b>
<b>STUDENT NAME</b>	:	<u>MUHAMMAD FAIZ BIN ABD RAZAK</u>
<b>UiTM ID NO.</b>	:	<u>2023911655</u> <b>GROUP:</b> <u>NBCS2405A</u>
<b>LECTURER</b>	:	<u>AHMAD ZAMBRI B SHAHUDDIN</u>

**INSTRUCTIONS TO CANDIDATES**

1. This question paper consists of **PART A** and **PART B**
2. Answer **ALL** questions in English.
3. Please send only **ONE (1)** attachment as your answer file.
4. Rename the answer file with STUDENTNAME\_STUDENTID\_COURSECODE  
For example:  
**SOFIA JANE ABDULLAH\_2022111222 ICT550**
5. No discussions are allowed during the test.

Question	Full Marks	Your Marks
A	45	
B	70	
<b>Total marks</b>	115	

---

**DO NOT TURN THIS PAGE UNTIL YOU ARE TOLD TO DO SO**

---

*This question paper consists of 4 printed pages*

## ANSWER

### QUESTION 1

- a) Data management is the process of acquiring, storing, organizing, and using data efficiently and securely.

b)

Structured (S)	Unstructured (U)
A database table storing employee IDs, names, and department codes.	A voice recording of a customer service call.
A JSON file with weather data fields like "Temperature," "Humidity," and "Wind Speed".	A printed doctor's handwritten notes about a patient's symptoms.
	A collection of social media comments on a viral post.

- c) 1. **Format and Structure:** Unstructured data has no set schema and can be in the form of text, audio, video, or social media posts, whereas structured data has a predetermined format, such as rows and columns (such as tables in a database).

2. **Simplicity of Analysis and Storage:** Conventional tools like SQL make it simple to store, search, and analyze structured data. On the other hand, meaningful analysis of unstructured data necessitates the use of specialized technologies such as text mining or natural language processing (NLP).

- d) Master Data Management (MDM) is the process of creating and maintaining a single, consistent, and trusted source of key business data such as customer, product, or employee information across an organization.

- e) **Architecture of the Registry** a simple method that uses indexing and matching algorithms to connect master data across systems without requiring the data to be physically moved. It leaves real data in source systems but preserves references.

**Architecture of Consolidation** Source systems continue to be the systems of record even while master data from several systems is gathered and stored in a central repository for analysis and reporting. Consistency is enhanced without complete data synchronization.

## QUESTION 2

- a) The degree of accuracy, completeness, dependability, and relevance of data for its intended use is referred to as data quality. Efficient decision-making and corporate operations are supported by high-quality data
- b) **Accuracy:** Information accurately depicts the thing or event it represents in the real world. The name and address of a consumer, for instance, must correspond with their official records.

**Completeness:** Every necessary data field is there. Data dependability is decreased by missing information, such as an empty phone number column.

**Consistency:** The same data value shows up in several records or systems in the same way. A product pricing that appears differently in two systems, for example, is inconsistent.

- c) Inaccurate, partial, redundant, out-of-date, or improperly formatted data are all considered dirty data and can result in deceptive analysis or operational mistakes.
- d) 1. A customer's date of birth recorded as “31/02/1999” (an invalid date).  
2. The same customer appearing twice with slightly different spellings (e.g., “John Smith” and “Jon Smith”).
- e) **Errors in manual data entering** The dataset may contain errors and inconsistencies due to human error, such as typos, improper formatting, or missing values during data entry.

**Absence of data standards** Mismatches and integration challenges arise when disparate data sources adhere to inconsistent standards (such as date styles or naming conventions).

**Systems that are redundant or outdated** The overall dependability of the data originating from legacy systems may be impacted by the presence of redundant, out-of-date, or badly managed data.

### QUESTION 3

- a) **Manual Data Integration:** Information is gathered and assembled by hand from various sources. It's easy, but it takes a long time and is prone to mistakes.

**Middleware-based Integration:** This method connects several systems via software, enabling automated data interchange without changing databases that are already in place.

**Application-based integration:** is the dynamic collection, processing, and unification of data from several sources through the use of applications or APIs.

- b) **Single Source of Truth:** Unifying data to remove redundancy and guarantee consistency across systems is known as the "single source of truth."

**Better Data Accessibility:** Ensuring that users have centralized access to comprehensive data.

**Improved Decision-Making:** Better analysis and better-informed choices are supported by integrated data.

**Operational Efficiency:** Enhances corporate workflows by reducing human labor and data silos.

**Data Quality and Consistency:** Accuracy and uniformity of data across departments are encouraged by data quality and consistency.

- c) **Schema heterogeneity:** Mismatches or compatibility problems may arise from variations in data formats or structures among systems.

**Semantic Inconsistencies:** In some systems, the same data field may have distinct meanings (for example, "client" versus "customer").