“Your team is expected to prepare and deliver a design report of your intended development work for the client organisation, and you may refer to the Lecturecast in this unit for guidance. Your design report should capture the following:

- ***Logical design*** *- data items/entities, attributes of the data items chosen, relationships and associations. Identify and explain the data types used and data formats selected.*
- *Produce a proposal of the database build, creating an intended* ***database model design****. You should propose a database management system that you will be using for the build, considering the client requirements of storage, user access, and the manipulation and retrieval of data within the proposed database.*
- ***Critically evaluate the data management pipeline*** *process with regards to discussing the capturing of the data used and detailing its source,* ***documenting how you implemented data cleaning techniques and the stages*** *that have been carried out during the cleaning process.”*

Step 1: Select the Application and Client Profile

- ✔ **Brainstorm potential industries**:
  - List areas of interest or familiarity within the team (e.g., retail, transport, healthcare).
  - Choose a specific client type (e.g., an e-commerce store needing inventory tracking or a postal service optimizing delivery routes).
- ✔ **Decide together**: Select the most feasible and interesting idea for the team.

Step 3: Logical Database Design

- ✔ **Identify Data Items/Entities**:
  - List key entities (e.g., for retail: Products, Customers, Orders).
  - Define attributes for each entity (e.g., Product: ProductID, Name, Price).
- ✔ **Determine Relationships**:
  - Map how entities relate (e.g., "A Customer places an Order").
  - Create an Entity-Relationship Diagram (ERD) to visualize relationships.
- ✔ **Select Data Types and Formats**:
  - For each attribute, define its data type (e.g., integer, string, date).
  - Standardize formats (e.g., YYYY-MM-DD for dates).

Responsibility:

- One team member creates the ERD.
- Others review and provide feedback.

Step 4: Database Build Proposal

- ✔ **Select a Database Management System (DBMS)**:
  - Research systems (e.g., MySQL, PostgreSQL, MongoDB).
  - Choose a DBMS based on client needs (e.g., scalability, ease of access).
- ✔ **Design a Database Model**:
  - Develop the schema in line with the logical design.
  - Ensure normalization to avoid redundancy.

- ✔ **Address Storage, Access, and Data Manipulation**:
    - Propose solutions for storage capacity, user access roles, and query handling.
    - Justify these choices in the report.

Responsibility:

- Divide research tasks for DBMS options and justification.
- One member drafts the model.

Step 5: Data Management Pipeline

- ✔ **Capture and Source Data**:
    - Decide where data originates (e.g., client logs, external APIs).
    - Document sources clearly in the report.
- ✔ **Implement Data Cleaning Techniques**:
    - Identify potential issues (e.g., duplicates, missing values).
    - Describe how these are handled (e.g., deduplication, null value replacement).
- ✔ **Document the Pipeline**:
    - Explain the steps from raw data capture to cleaned data ready for the database.

Responsibility:

- One member leads the pipeline discussion and writes the related section.
- Others review and enhance.

Step 6: Write the Report

- ✔ **Draft Sections**:
    - Assign sections to team members (e.g., Introduction, Logical Design, Pipeline).
    - Write concisely and meet the word count requirement (1,000 words max).
- ✔ **Include Visuals**:
    - Add diagrams (ERD, database schema).
    - Include charts or flow diagrams for the pipeline, if helpful.
- ✔ **Proofread and Integrate**:
    - Merge all sections into one cohesive document.
    - Ensure consistent tone and formatting.
    - One person handles final edits and submission.

Responsibility:

- Assign one member to coordinate and finalize the report.

An **Entity-Relationship Diagram (ERD)** is a visual representation of the data model for a system. It shows the entities (tables), their attributes (columns), and the relationships between them. ERDs are a critical step in database design, as they help to organize and plan how data will be stored and related.

Core Elements

- ✔ **Entities**: Represented as rectangles, these are the main objects in the database (e.g., *Customer*, *Order*, *Product*).
- ✔ **Attributes**: Represented as ovals or listed within entities, these are the properties or details of an entity (e.g., *CustomerName*, *OrderDate*).
- ✔ **Relationships**: Represented as diamonds or lines connecting entities, they show how entities interact (e.g., *Customer places Order*).

- **Lucidchart / https://www.lucidchart.com/pages/er-diagrams**
- **Draw.io (diagrams.net)**

Steps to Create an ERD Using a Tool (e.g., Lucidchart or Draw.io)

- ✔ **Define Entities**:
  - Create rectangles for each entity (e.g., *Customer*, *Order*).
  - List attributes inside the rectangles.
- ✔ **Establish Relationships**:
  - Connect entities with lines to show relationships.
  - Label the lines with relationship types (e.g., "places", "contains").
- ✔ **Add Cardinality**:
  - Use notations like 1:1, 1:N, or N:M to show relationships (e.g., "One Customer places many Orders").
- ✔ **Review and Export**:
  - Review for completeness and accuracy.
  - Export the diagram as a PDF or PNG for the report.