Assignment 1: Data Modeling with ERDs

The purpose of this assignment is to practice converting a business problem into a data model, while thinking about the types of data anomalies that can appear throughout the process.

Consider the following description of a **college:**

- The course WMDD 4820 is taught by Mort (Certified College Instructor).
    - Below are the students and their grades. For privacy reasons, they are referred to only by their integer ID numbers.
        - 1 (B-)
        - 2 (C)
        - 3 (A-)
    - Mort is the father of student #1
- WMDD 4835 is taught by Bob (No Certification).
    - The students:
        - 1 (B+)
        - 4 (C-)
        - 3 (A+)
- WMDD 4936 is taught by Mort (Certified College Instructor).
    - Students:
        - 1 (A-)
- WMDD 4980 is taught by Linda (Certified College Instructor).
    - The students:
        - 5 (B)

The college has asked us to create an application (and its accompanying database), using the above data, to answer the following questions:

A) Do courses taught by certified instructors have higher or lower average grades?
B) Do students taught by family relatives have higher or lower average grades?

Note that even though the above two problems involve average grades, your database design doesn't have to include the averages, only the individual grades. (The application will calculate the averages AFTER pulling the grades from the database.)

Create a new text document called **answers** (**.txt** or **.md**) to write your answers. Each written answer should be clearly labeled with the task number and letter. For making tables, use **Excel** or **Google Sheets** and save the file in Excel **.xlsx** format, naming the file as directed. For ERDs, either draw them on paper or use **Lucidchart;** take a photo or export as an image and name the file as directed.

Design the Data Model

To design the data model for this application, follow these steps:

1. **[2 marks]** Create a **single table** that contains all the data above. (Name the file **1-single-table.xlsx)**
    a. [**2 marks**] Find at least **two** anomalies that this single-table data model is susceptible to and explain them with examples.
2. We obviously need to split this data into more than one entity! Create an **ERD** with the following entities: **Student, Instructor,** and **Course.** It should have **attributes** for each entity and **relationships** between entities. Each entity should have an attribute that acts as a unique identifier; you can make up a new attribute for this if one doesn't exist already. (Name the image **2-erd-three-entities**)
    a. **[1 mark]** Explain and justify any assumptions you must make about aspects of the data that are not clearly stated in the problem description.
    b. **[3 marks]** For the relationships, include multiplicities with **ordinality** and **cardinality.**
    c. **[1 mark]** Create tables for each of these entities (with their attributes as **columns**) and fill them with all the data from the first table. (Name the file **2-c-three-tables.xlxs)**
    d. **[2 marks]** Explain which anomalies were resolved from the previous step, using examples.
3. **[2 marks]** Your ERD from the previous step ought to have M:M relationships, which we generally do not want. Create *another* ERD that has *no* **M:M** relationships. (name the image **3-erd)**
    a. **[2 marks]** For each of the two tasks that our application must accomplish (A and B), explain the sequence of steps that would be taken to get the necessary data from the database using the model you designed. You may find it helpful to transfer the data to tables.
4. Suppose were to add the following rule to the problem description: *courses can be taught by more than one instructor*. (Maybe the instructors take turns, or maybe they split the topics of the course.)
    a. **[1 mark]** How does this complicate the problem? How must your data model (and therefore ERD) be updated to accommodate this change? Feel free to make an ERD or table to supplement your explanation if you want (name it **4-b-table/erd**)**.**

Total: **16**

Hand In

Place all files in a folder called **a1-firstname-lastname** with your own first and last name. Zip the folder AFTER renaming it. Don't use some other archive format like 7z or rar, as we may not be able to open it. Hand in the zip archive to the appropriate assignment folder; copy the text from your **answers** document to the **Comments** area of the submission page.