Ollscoil na Gaillimhe
University of Galway
Scoil Ghnó agus Eacnamaíochta J.E. Cairnes
J.E. Cairnes School of Business and Economics

MS805 Group Assignment Specification

"Supporting the Data Needs for SDG Applications"
For MS805 Database Systems 2024–2025

Lecturer: Dr. Blair Wang, blair.wang@universityofgalway.ie

---

1. Introduction

In this group assignment, you and your team will be **designing a relational database system based on a case study** that is relevant to one of the United Nations Sustainable Development Goals (SDGs). In order to facilitate diversity in the SDGs that various student teams are engaging with, there will be a process of allocating SDGs to teams based on an allocation process outlined in section 2 of this document.

In order for your relational database system to be useful, relevant, and supportive of the organisation specified in your case study, you will need to produce THREE (3) design options, write some reports explaining your designs, and write some examples of SQL queries that demonstrate the functionality of your database design. These are the **deliverables** outlined in section 3 of this document.

This group assignment is worth **30% of your total marks** for MS805.

2. Next Steps

1. ***Await case study notice from the lecturer:*** Your lecturer will email your team with details of a case study relevant to one of the United Nations SDGs.
2. ***Complete Deliverables A to E:*** as outlined in section 3 of this document, below.

3. Deliverables

3.1. Deliverable A — Assignment Cover Page and Peer Evaluation Form

What to upload to Canvas:

- ***Required:*** *1 × PDF file.*
- *No other files are expected for this deliverable.*
- *There is no word limit for this deliverable. Furthermore, this deliverable does NOT contribute to the page limit of this assignment.*

Each team will be required to submit a completed **MS805 Assignment Cover Page and Peer Evaluation Form**. More information will be made available on Canvas closer to the assignment due date.

3.2. Deliverable B — ERD and Relational Model (3 × Design Options)

What to upload to Canvas:

- ***Required:*** *1 × PDF file, with multiple pages.*
- ***Optional:*** *1 × ZIP file, providing additional supplementary materials (e.g., ".mermaid" files, PNG files, SVG files) as you see fit.*
- *There is no word limit for this deliverable. There is, however, a page limit of the entire assignment overall (please refer to section 4 of this document).*

This deliverable requires you produce at least THREE (3) possible design options for the relational database system that you are designing for the organisation specified in your case study. These designs are to be depicted as Entity-Relationship Diagrams (ERDs). Here are some tips to consider:

1. You may use whatever tool you like for the purpose of drawing the ERD. Whilst Mermaid is recommended (https://mermaid.js.org/), **you are very welcome to use any appropriate software or tools that you prefer** based on your team's personal preferences (provided that the software is legally sourced, free of malware, does not compromise your academic integrity or the academic integrity of others, and in general complies with the ICT Regulations and Policies of the University of Galway). You may even use pen and paper if you'd like, provided that you can upload a sufficiently high-resolution photograph or scan of any pen-and-paper diagrams that you draw.
2. Please ensure that each of your ERDs is clear and legible. If you choose to use Mermaid, you may find it help to submit your Mermaid source files (".mermaid" files) to ensure that this additional information is available for the marking process. If you choose to use other diagram-drawing software, you may find it helpful to submit original image files (e.g., PNG, SVG). If you choose to use pen-and-paper, you may find it helpful to submit the full-size photograph (e.g., JPEG), but be sure to not include your private information (e.g., geolocation coordinates).[1] All such sources files, images, and/or photograph files should be packaged as a ZIP file for upload to Canvas.

[1] For example, please refer to: https://www.macworld.com/article/351888/how-to-remove-location-data-in-shared-photos.html

3. As a general guideline, you should have somewhere in the range of 4~12 entities in your ERD, each with approximately 4~8 attributes. **This is a general guideline only and you are NOT required to comply with it**, but you may find this guideline helpful for determining whether or not your team's work is "on the right track".
4. If you find that it is not possible to fit all entities *and their attributes* on the same page, you may consider drawing an ERD with only entities, their relationships and the connectivities of those relationships, i.e., without attributes. You could then provide the list of entities and their attributes separately, in the form of a *data dictionary*.

If you would like, your team may take the option of using generative artificial intelligence (e.g., OpenAI's ChatGPT, Microsoft's Copilot) to generate, in part or in whole, ONE (1) of the three (3) possible designs that you are producing. **Marks will be awarded for skilful use of generative artificial intelligence in this manner**, provided that you:

1. Clearly identify which of your three designs was based on work produced using generative artificial intelligence; *and*
2. Clearly acknowledge, by name, which generative artificial intelligence product you used (e.g., OpenAI's ChatGPT, Microsoft's Copilot, etc.); *and*
3. Provide the prompts that you used to generate your design using generative artificial intelligence; *and*
4. Critically compare the design produced using generative artificial intelligence with the other designs. This is to be done as part of Deliverable C (described below).

3.3. Deliverable C — Design Flexibility Report

What to upload to Canvas:

- ***Required:*** *1 × PDF file.*
- *No other files are expected for this deliverable.*
- *There is no word limit for this deliverable. There is, however, a page limit of the entire assignment overall (please refer to section 4 of this document).*

This deliverable requires you to write a brief report about your design options (ERDs), commenting on:

1. **Assumptions:** What assumptions have you made in order to produce your ERDs? How have these assumptions impacted on the designs of your ERDs and lead to various design decisions, in which areas and in what ways?
2. **Comparative advantages and limitations:** When comparing the three ERDs that you have designed, what are the relative (comparative) advantages of each ERD? For example, you may say that *Design #1 is particularly good for so-and-so reason whereas Design #2 is not as advantageous for that aspect but has other advantages according to such-and-such*. If you have chosen to design one of your ERDs based on generative artificial intelligence, please draw particular attention to that ERD and comment on the particulars of what the generative artificial intelligence has done well and likewise what the generative has done not so well.
3. **Flexibility:** In your professional opinion, how much work would be involved in changing the design to work with other possible designs both (a) prior to database rollout and (b) when the database is already live and in use?

4. **Sociotechnical considerations:** which may include some or all of the following (depending on your best judgement), and any others you may consider relevant:
    - Stakeholders and sensemaking
    - Organisational silos and boundary spanning
    - Time and temporality
    - Consensus and the Abilene paradox
    - Upgrading with care
    - Corroboration and entity resolution

There is no particular word limit for this deliverable — apart from that of the entire assignment overall (please refer to section 4 of this document).

3.4. Deliverable D — Normalisation Report

What to upload to Canvas:

- ***Required:*** *1 × PDF file.*
- *No other files are expected for this deliverable.*
- *There is no word limit for this deliverable. There is, however, a page limit of the entire assignment overall (please refer to section 4 of this document).*

This deliverable requires you to write a very brief report about your three design options (ERDs). For each design option (ERD), please comment on:

1. Which normal form (1NF, 2NF, 3NF, other?) have you reached?
2. How do you know that you have reached that normal form?
3. Why is that normal form appropriate for your design? (You may wish to refer to points that you made in Deliverable B.)
4. If you have reached only 1NF or 2NF — would further normalisation into higher normal forms have a positive, negative, or neutral impact on your work?

Again, there is no particular word limit for this deliverable. However, it would be anticipated that, for most teams, Deliverable D would be shorter than Deliverable C.

3.5. Deliverable E — SQL Scripts

What to upload to Canvas:

- ***Required:*** *1 × ZIP file, containing your SQL scripts.*
- ***Required:*** *1 × PDF file, with multiple pages.*
- *There is no word limit for this deliverable. There is, however, a page limit of the entire assignment overall (please refer to section 4 of this document).*

This deliverable requires you to submit SQL scripts (i.e., plaintext files containing SQL code, with the file extension ".sql"). Your SQL scripts should do the following:

1. Using data definition language (DDL), establish the database with its tables, columns, primary key / foreign key / unique column constraints, etc.
2. Using data manipulation language (DML), establish at least 10 rows of sample data ("dummy data") for each table.
3. Using SELECT statements, provide at least FIVE (5) queries that could be useful for various stakeholders.

Please note that your SQL scripts should be packaged as a ZIP file for upload to Canvas.

Additionally, in relation to the *"FIVE (5) queries that could be useful for various stakeholders"*, please provide a brief report (PDF file) describing these queries, what they are doing, which stakeholders they would be useful for and why, and showing screenshots of the expected output.

3.5. Interim Progress Report

What to upload to Canvas:

- ***Optional:*** *1 × PDF file, with multiple pages.*
- ***Optional:*** *1 × ZIP file, containing any work in progress that you wish to showcase or seek feedback on.*
- *There is no word limit; furthermore, the Interim Progress Report does NOT contribute to the page limit of this assignment.*

Prior to the final due date for Deliverables A to E, you are invited to submit an **Interim Progress Report** about how you are progressing. Whilst there are no marks allocated to the Interim Progress Report, you are strongly advised to participate in the Interim Progress Report process in order to receive feedback and facilitate the early detection of any issues that may affect your understanding of, and performance in, this group assignment.

There is no strict structure to follow for the Interim Progress Report. However, you are recommended to write a page or two on:

1. What you have completed and what you have not yet completed
2. What you are finding difficult or confusing about the assignment
3. Which aspects of your work would like feedback about

In relation to point 3: you may wish to provide, packaged as a ZIP file, your working drafts for your deliverables A to E. Your lecturer will provide some feedback, based on which aspects of your work you have indicated that you would like feedback about.

4. Word Limit, Page Limit, Formatting Requirements

Given that this assignment involves the submission of SQL scripts, **no word limit is imposed on your work** for this assignment.

However, **there is a page limit of 100 pages** in the sense that, if the total number of A4 pages in all the PDF files uploaded for Deliverables B~E somehow exceeds 100 pages, then the lecturer reserves the right to only assess the first 100 pages. This constraint is in place to avoid situations where an unreasonably large number of pages are sent in for assessment. In practice, it is very unlikely that student work is exceeding 100 pages.

Please **ensure that your work is appropriate formatted** so that it is legible and carries a professional appearance. There are no specific prescriptions for typesetting, font size, margin, or line spacing, but please use your best judgement. PDF documents should be created with appropriate settings, including ensuring that images are not overly compressed. SQL scripts should be written according to appropriate naming conventions, indentation, and line breaks.

5. Timeline and Due Dates

Key Date	What's due?	Notes
Friday 18 October 2024 (Week 6)	**Interim Progress Report** to be uploaded to Canvas.	*(See section 5.)*
Friday 01 November 2024 (Week 8)	**(Main Assignment Due Date) Deliverables A to E** to be uploaded to Canvas.	*(See section 3.)*

Late submission of the group assignment is not advisable. A **late penalty of 10% of the total marks available** will apply for each business day that a deliverable is submitted late.

Exceptions will be made in line with the University's QA209 Extenuating Circumstances Policy and QA208 Alternative Assessment Policy.

6. Academic Integrity and Referencing

You are required to operate in alignment with the University's QA220 Academic Integrity Policy. Whilst all aspects of that policy apply, please particularly note that:

1. The policy prohibits *collusion* (p. 1), which is when part of your work for this assignment was produced by students outside of your team.
2. The policy prohibits the use of generative artificial intelligence *and claiming it as your own work* (p. 2). Therefore, the use of generative artificial intelligence in Deliverable B is permitted if and only if it is acknowledged appropriately (including the name of the generative artificial intelligence product and the details of the prompts used) as outlined in section 3 of this document, for only for ONE (1) of three (3) designs for Deliverable B.

The policy also requires that you must use appropriate referencing when building on the work of others. Please familiarise with the University's resource on appropriate referencing, provided by the University Library and available at the following web address:

https://libguides.library.universityofgalway.ie/friendly.php?s=research-skills/citingandreferencing

You are advised to use the **APA** referencing format. Referencing software like EndNote, Zotero or BibTeX may be helpful for this purpose.

7. Mark Allocations and Learning Outcomes

Deliverable	**Marks Allocated**	**Learning Outcomes**
A. Assignment Cover Page and Peer Evaluation Form	10%	—
B. ERD and Relational Model (3 × Design Options)	25%	1, 2, 4
C. Design Flexibility Report	20%	1, 2, 4
D. Normalisation Report	20%	3
E. SQL Scripts	25%	4, 5, 6, 8

The MS805 Learning Outcomes referenced above are as follows:

1. *Explain basic principles of relational database design.*
2. *Model data requirements using techniques such as Entity-Relationship and UML Class Diagrams.*
3. *Apply the technique of normalisation to optimise data structures in relational database design.*
4. *Understand how a relational database system enforces data integrity constraints and transaction management rules.*
5. *Use the Structured Query Language (SQL) to create a relational database with tables and views.*
6. *Use SQL to extract, summarise and aggregate data in a relational database.*
7. *Use advanced features of SQL such as fuzzy matching, spatial data queries, and triggers.*
8. *Understand principles of SQL query optimisation and indexes.*