

# Modelling Database Requirements with E-R Diagrams

IS5102 – Database Management Systems

Due date: Thursday 6th October 2022 (week 4), 21:00  
33.0% of coursework for the module  
(MMS is the definitive source for deadline and credit details)

In this assignment, you are asked to model a scenario for its database aspects, using entity-relationship models to express models and requirements.

You are expected to have read and understood all the information in this specification at least a week before the deadline. You must contact the lecturer regarding any queries well in advance of the deadline.

## Purpose

This practical will help practice and develop your skills in:

- analysing and defining database requirements;
- modelling data using E-R model techniques;
- understanding how to utilise functional requirements;
- understanding and reporting requirements.

## Summary

This is an **individual assignment**.

In lectures we have talked about modelling database scenarios, in particular at the conceptual level using E-R models. This assignment presents a short scenario and asks you to model it using the techniques presented.

## Submission

The primary output will be a **report**, which should be in **PDF format**, with sections per tasks 1-3 as below. Include a picture (machine-generated or scanned) of your model in the report. Submit this report via MMS by the deadline, checking to make sure that the version you submit is the one you mean to submit.

## Scenario

A university lab needs to organise their experimental data with the help of a database management system. The following initial specification has been given:

We operate several machines to analyse chemical samples. Each machine processes samples in batches. Our machines have varied characteristics, including maximal number of samples in a batch, batch runtime, maximal weight of a single sample, as well as the model and year of manufacturing. Each machine also has a unique name under which it is known to its operators and users.

Researchers bring their samples to the lab for analysis, and collect them back afterwards. Upon receiving by the lab, each sample is given a unique ID and weighted. We aim at processing each sample within 10 days, and send a notification to a researcher when it's completed. We also have limited storage facilities in the lab, so we need to send a reminder for samples not collected within another 10 days after they were analysed. We will however keep indefinitely long the results of the analysis (a CSV-file associated with a particular processing of a sample as a part of a batch) which may be retrieved by the researchers whenever they need them. For this purpose we record the date, time and unique ID number of the batch run in which this file has been produced for a given sample. It may also happen that a sample needs to be analysed several times, possibly on different machines, in case of any problems with initial analysis.

Our users may be either staff members or students. In both cases, we record their name, email, and department. For staff members, we also record their job title, and for students - their start date. Each student must be supervised by a staff member.

Finally, for billing purposes any analysis of any sample must be charged to a specific research project. This information is provided by a researcher when a sample is received by the lab. Each research project has a name, and a unique combination of a sort code and account number. We also record which researchers are working on this project, so when a sample is brought to the lab we can check whether a staff member or a student has the credentials to charge its analysis to this project. At the end of each month, we generate reports with the number of samples analysed for each project, and pass them to Finances for processing.

Besides using the database for tracking samples, storing analysis results, and charging research projects for services, we would also like to generate some service statistics to analyse the demand and efficiency of our work.

## Tasks

### Task 1

From the given scenario, document the intended use of the data and any other factors that you consider relevant. You should think particularly of the questions that will be asked of the data. As part of this, come up with (at least) **three queries** that you feel are likely to be made by stakeholders (use plain English to formulate questions and queries).

Write a **specification** of the data from the scenario that will be stored in the relational database. It should be written in plain English (instead of using a formal notation). You should think of what constraints may exist on the data, and how to express any such constraints. This specification will be the basis for your database definition, so it should be as precise as possible. Document clearly any assumptions you make.

## Task 2

Design a representation of the data in terms of entities, attributes and relationships between entities. Construct an **E-R diagram** to depict this representation. Your E-R diagram should be submitted electronically as part of your report. It is recommended that you use some drawing software in the construction of your model (<https://www.diagrams.net/>, with its online version also known as <https://draw.io/>, or drawing tools in Microsoft 365 components will suffice). However, if you wish, you can use scanned versions of hand-drawn diagrams. These are acceptable provided they are neat and legible.

In your report, indicate where the data to answer the queries from Task 1 “lives” in your data model, showing that your data model does satisfy the requirements specification.

## Task 3

Reflect on the process of doing database modelling, linking back to the work done in Tasks 1 and 2. What did you feel you did well? What did you find challenging? What problems did you encounter, if any, and how did you try to resolve those? Did you find gaps between the theory taught in class and/or presented in textbooks, and the practical application? What would you do differently, if anything? This should be a **short reflective section** of your report. You can say as much or as little as you want, but as guidance, between half a page to a page should be enough.

## Marking

A specification capturing the requirements, a correct E-R model which captures the scenario, and a reflective section of the report make up the basic requirements. A good report covering all of these will be sufficient to get marks up to 16.

To get the higher marks (17 and above), advanced work is required. This can include (for example)

- insightful and relevant questions for the requirements;
- exceptional clarity and understanding in the requirement specification;
- good style in E-R modelling including (for example, where applicable) weak entity sets; specialisation/generalisation; a range of simple, compound and derived attributes etc; or
- report demonstrating excellent understanding of the modelling process.

## Policies and Guidelines

### Marking

See also the standard mark descriptors in the School Student Handbook:

[https://info.cs.st-andrews.ac.uk/student-handbook/learning-teaching/feedback.html#Mark\\_Descriptors](https://info.cs.st-andrews.ac.uk/student-handbook/learning-teaching/feedback.html#Mark_Descriptors)

### Lateness

The standard penalty for late submission applies (Scheme B: 1 mark per 8 hour period, or part thereof):

[https://info.cs.st-andrews.ac.uk/student-handbook/learning-teaching/assessment.html#Lateness\\_Penalties](https://info.cs.st-andrews.ac.uk/student-handbook/learning-teaching/assessment.html#Lateness_Penalties)

### Good Academic Practice

The University policy on Good Academic Practice applies:

<https://www.st-andrews.ac.uk/students/rules/academicpractice/>