

ISYS1055/3412 (Practical) Database Concepts

Assessment 1: Database Design



Assessment type: PDF

Word limit: N/A



Draft Due Date: 31 March 2024 at 23:59 (Melbourne time) – Week 4 (otherwise 2 mark deduction)

Final Due Date: 21 April 2024 at 23:59 (Melbourne time) – Week 6



Weighting: 20%, 20 Marks

Overview

The objective of this assignment is to measure your understanding of the basic concepts in the relational database model and using entity-relationship model for database design. The assessment is in two parts, split into four tasks which cover Basic ER Modelling and Basic Relational Modelling. The tasks are as follows.

Part A: Entity-Relationship Modelling (12 Marks)

1. Design and plan for the implementation of a database system, diagramming the design to a high standard using UML notation through the diagramming tool Lucidchart.
2. Model the activities of an organisation and present the model as an Entity-Relationship (ER) diagram. Analyse this ER diagram, and possibly modify it, based on additional client requirements.
3. Map an ER diagram into a relational database schema, showing every step of the mapping.

Part B: Relational Database Model (8 Marks)

4. Answer a series of short questions about a Relational Database model.

To complete this assessment, you must be familiar with Lucidchart, which is covered during the Week 1-4 activities.

Assessment criteria

This assessment will measure your ability to:

- Describe various data modelling and database system technologies.
- Explain the main concepts for data modelling and characteristics of database systems.

Course learning outcomes

This assessment is relevant to the following course learning outcomes:

CLO1	Describe the underlying theoretical basis of the relational database model and apply the theories into practice.
CLO2	Explain the main concepts for data modelling and characteristics of database systems.
CLO3	Develop a sound database design using conceptual modelling mechanisms such as entity-relationship diagrams.
CLO4	Develop a database based on a sound database design.

Task 1: Designing an Entity-Relationship Model**Metro Trains Study**

You are asked to design the train service management database for Metro Trains Corporation. Description of the miniworld is given below.

- There are several train lines from the city centre to suburbs. Each line has a unique name and is also described by the distance between the city centre and the terminus station. For example, the Lake Villa line runs from the city centre to Lake Villa and has a distance of 30 kilometres.
- Each station has a unique name, a contact phone number. A station is on at least one train line and some stations are on several train lines.
- Standard facilities of stations include vending machines, newspaper machines, and ticket machines. Each machine has a unique facility number. Some premium stations have such facilities.
- Trains are identified by unique train Nos. Other information about trains includes passenger capacity, model, manufacturer and year-made. A train line has several trains and a train is assigned to one line if not under maintenance.
- Employees of Metro Trains are described by unique employee Nos, and the first name, last name, and birth date. Drivers are assigned to work on several train lines, and a line has several drivers.
- During peak hours (7:30am—9:30am on workdays) premium stations have supervisors to provide customer service. A premium station has a supervisor, and a supervisor can supervise at most one premium station.

Based on the given description, model the given business rules, and present your model as an Entity-Relationship (ER) diagram. Carefully state any assumptions that you make. In your ER diagram, you must properly denote all applicable concepts, including weak or strong entities, keys, composite or multi-valued attributes, relationships and their cardinality and participation constraints.

If you cannot represent any of this information in the ER model, clearly explain what limitations in the ER model restrict you from representing your model.

You **must** use UML notation and the diagramming tool [Lucidchart](#) to draw your diagram. Your diagram must be drawn to a high standard with minimal clutter. You are **not** required to map the ER model to relational model.

A special note: This is an open-ended question with many different models that can be derived. Your model is assessed based on how accurately it represents business rules described above.

Task 2: Optimize an Entity-Relationship Model

Advanced Metro Trains Study

Initial design of Metro Trains management database is now presented to users. After some detailed discussions, complete requirements for the database are finalised as follows (changes are in bold):

- There are several train lines from the city centre to suburbs. Each line has a unique name and is also described by the distance between the city centre and the terminus station. For example, the Lake Villa line runs from the city centre to Lake Villa and has a distance of 30 kilometres.
- Each station has a unique name, a contact phone number. A station is on at least one train line and some stations are on several train lines. **A station also has a number counting outwards from the city centre. A station may be on several train lines with the same number. For example, Happy Valley is the 2nd station on both the Lake Villa line and Eastwood line. Smallville is the 10th station on both the Eastwood line and Eastern Cross line**
- Standard station facilities include vending machines, newspaper machines, and ticket machines. Each machine has a unique facility number. Some premium stations have such facilities.
- Trains are identified by unique train Nos. Other information about trains includes passenger capacity, model, manufacturer and year-made. A train line has several trains and a train is assigned to one line if not under maintenance.
- Employees of Metro Trains are described by unique employee Nos, and the first name, last name, and birth date. Drivers are assigned to work on several train lines, and a line has several drivers.
- During peak hours (7:30am—9:30am on workdays) premium stations have supervisors to provide customer service. A station has a supervisor, and a supervisor can manage at most one station.
- **An employee may have a line manager, who is also an employee at Metro Trains. Start date of supervision should also be recorded.**
- **A station may have several platforms numbered from one. For example, Happy Valley has 8 platforms from No. 1 to No. 8. Smallville has 4 platforms from No. 1 to No. 4.**
- **Each line has inbound and outbound train runs that are numbered from one. For example, Eastern line has inbound run No. 1, outbound run No. 1, inbound No. 2, outbound No. 2 etc. Each train run is assigned a primary driver and an assistant driver. Each train run is also assigned a train. On each line, a train run is timetabled to depart a platform of a station at certain time. Express train runs do not stop at all stations on a line.**

Please modify / optimize your initial ER diagram from Task 1 to enforce the additional requirements specified above. A complete ER diagram (initial ER with amendments) should be presented. Carefully state any assumptions that you make. In your ER diagram, you must properly denote all applicable concepts, including weak or strong entities, keys, composite or multi-valued attributes, relationships and their cardinality and participation constraints.

You **must** use UML notation and the diagramming tool [Lucidchart](https://www.lucidchart.com) to draw your diagram. Your diagram must be drawn to a high standard with minimal clutter. You are **not** required to map the ER model to relational model.

A special note: This is an open-ended question with many different models that can be derived. Your model is assessed based on how accurately it represents business rules described above.

Task 3: Mapping an ER Model to a Relational Database Schema

Consider the following ER diagram, which shows aspects of a studio management system.

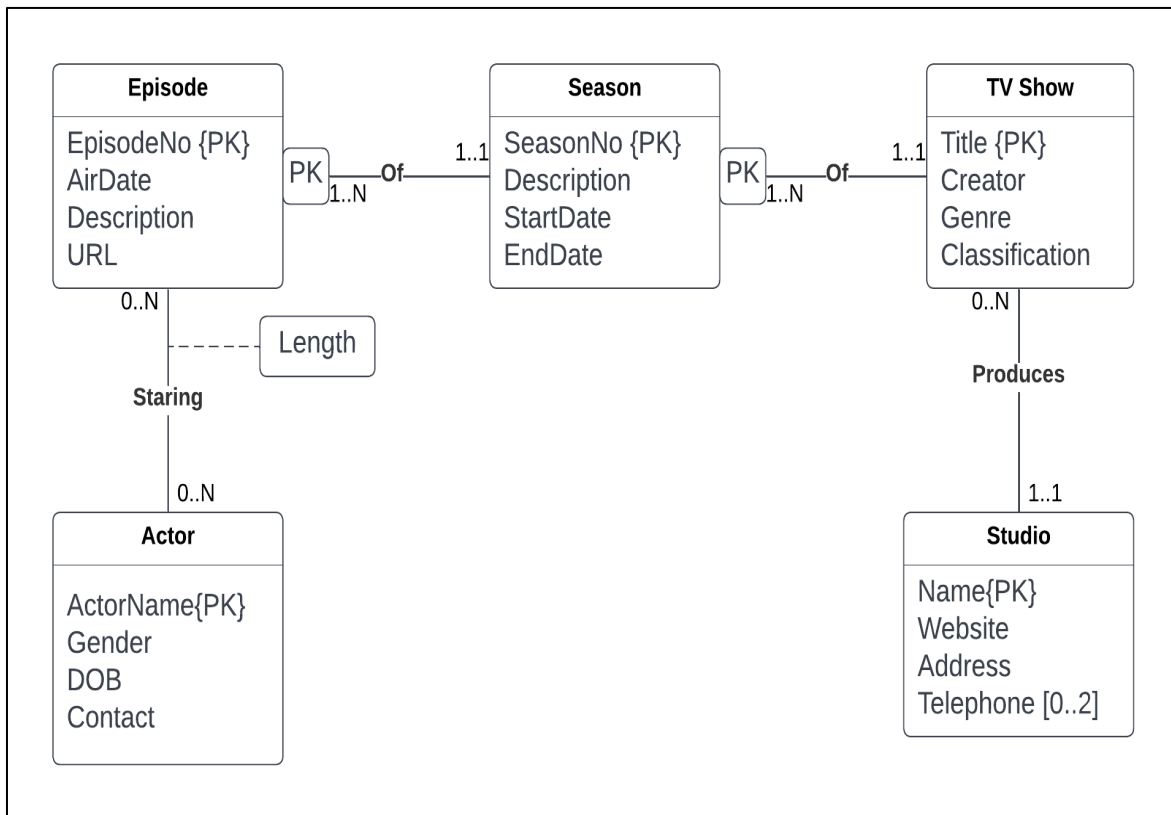


Figure 1: Employee Database Instance

You are requested to map the above ER diagram into a relational database schema. Show every step of the mapping.

No marks are awarded to the final schema if you do not show the partially built schema at the end of each step.

Clearly indicate the primary key (underlined) and foreign keys (with an asterisk) in each relation.

Part B: Relational Database Model

Task 4: Relational Database Model

This section contains the schema and a database instance for the Employee database that stores employee data for an organisation. The data includes items such as personal info (e.g., name, phone, salary), departments of the organisation (e.g., name and location of each department, who the manager is), jobs (e.g., titles, salary range), and a history for past contracts with each employee. A database instance is shown in Figure 2 followed by the database schema.

Employees							
employee_id	first_name	last_name	phone_number	hire_date	empjob_id	salary	department_id
50	Adam	Smith	1234	26/10/2009	22	\$66,000	2
66	Tom	Moosa	1235	10/12/2016	10	\$140,000	2
10	Jonny	Deans	1236	21/08/2002	33	\$70,000	1
12	Adam	Jones	1247	8/08/2009	10	\$138,000	1
18	Joseph	Ryan	1277	5/05/2020	10	\$150,000	3

Departments			
department_id	department_name	manager_id	location_id
1	IT Services	12	10
2	Accounting	66	20
3	Human Resource	18	30

Jobs			
job_id	job_title	min_salary	max_salary
10	Dep Manager	\$120,000	\$150,000
22	Accountant	\$60,000	\$80,000
33	Programmer	\$60,000	\$80,000
45	Senior Programmer	\$70,000	\$120,000

Locations					
location_id	street_address	postal_code	city	state_province	country_id
10	123 Collins St	3000	Melbourne	VIC	1
20	222 Bourke St	3000	Melbourne	VIC	1
30	555 Swanston St	3000	Melbourne	VIC	1

Countries	
country_id	country_name
1	Australia
2	Vietnam
3	Spain

Job History				
employee_id	start_date	end_date	job_id	department_id
10	1/01/2001	10/04/2002	33	1
10	11/04/2002	20/08/2002	33	1
12	1/01/1998	5/10/2003	33	1
12	6/10/2003	6/10/2004	33	1
12	7/10/2004	7/08/2009	33	1

Figure 2: Employee Database Instance

The database schema is shown below, and the meaning of most attributes is self-explanatory. “Job History” is simplified as “JobHistory”. Primary keys are underlined. However, only some foreign keys are annotated with a *

Employees(employee_id, first_name, last_name, phone_number, hire_date, empjob_id, salary, department_id)
Departments(department_id, department_name, manager_id, location_id*)
Jobs(job_id, job_title, min_salary, max_salary)
Locations(location_id, street_address, postal_code, city, state_province, country_id*)
Countries(country_id, country_name)
JobHistory(employee_id, start_date, end_date, job_id, department_id)

The following table further clarifies the connection between the keys across multiple tables.

Locations.country_id -----> Countries.country_id
Departments.location_id -----> Locations.location_id

Figure 3: Keys

The following questions must be answered based on the given database schema and instance. Where explanation is required, each answer should be a SHORT passage of at most several lines.

QUESTION 4.1: What foreign key constraints are missing from Figure 3? Please identify all foreign key constraints in the database schema and write down the missing constraints in the format shown in figure 3. Present the updated database schema reflecting **all** the constraints.

QUESTION 4.2: Does the database schema allow an employee to work on different jobs (jobID) in the same period? Explain your answer.

Questions 4.3: Does the database schema allow an employee to work for different departments (in different period) in the organization? Explain your answer.

QUESTION 4.4: Does the current database schema allow multiple different locations for the same department? If yes, explain your answer. If no, what changes would you need to make it possible?

QUESTION 4.5: The employee named Adam Smith changed his job to become a Programmer on 2/10/2012. The following SQL statement intends to make the required changes to reflect this change.

```
UPDATE Employees SET empjob_id=33, hire_date='2012-10-02' WHERE salary=66000;
```

Explain if there any issues with the outcome of the update, and how it should be fixed? After running the above query, consider the request “find all the past contracts that Adam Smith used to have”. Can this request be completed using the given database schema and after the above statement is run? If yes, explain how the request can be answered. If no, explain what is missing and how it should be fixed.

QUESTION 4.6: The organization plans to relocate all offices from 123 Collins Street to other locations due to lease expiration. The following SQL statement intends to delete this location from database.

```
DELETE FROM Locations WHERE location_id=10;
```

Will this query run properly? Explain your answer.

QUESTION 4.7: Write an SQL statement to create the JobHistory table including all the constraints, assuming all the tables that 'JobHistory' table depends on already exist in the database. Make reasonable assumptions for the data type associated with each field. Your SQL statement must be valid for SQLite Studio environment and free of any errors. Please use the updated database schema from Question 4.1.

QUESTION 4.8: On 01/01/2020, a new employee Scott Wallace was hired and was appointed as manager of the new Art department. The following SQL statements intend to record all the changes required in the database instance.

```
INSERT INTO Employees VALUES (88, 'Scott', 'Wallace', '1111', '2020-01-01', 10, 140000, 4);  
INSERT INTO Departments VALUES (4, 'Art', 88, 20);
```

Will they work properly? If they are sufficient to achieve the requirements specified above, explicitly mention they are sufficient. If no, explain the problem and provide correct SQL statements. Your SQL statement must be valid for SQLite Studio environment, free of any errors, and compatible/consistent with existing data in the instance in Figure 2. Please use the updated database schema from Question 4.1.

Submission format

You should submit one PDF document with all answers together. Do not submit Word files.

You must use Lucidchart to work on Part 1 of your assignment. You may use Word or any other word processor to compile your submission. Use section titles to indicate which question you are answering. At the end, convert your answer sheet into PDF format. Microsoft Word has the option of saving your document in PDF format. If the conversion option is not available on your system or word processor, there are free PDF converters online you can utilise (e.g., <http://convertonlinefree.com/>).

Academic integrity and plagiarism

Academic integrity is about honest presentation of your academic work. It means acknowledging the work of others while developing your own insights, knowledge, and ideas.

You should take extreme care that you have:

- Acknowledged words, data, diagrams, models, frameworks and/or ideas of others you have quoted (i.e., directly copied), summarised, paraphrased, discussed, or mentioned in your assessment through the appropriate referencing methods.
- Provided a reference list of the publication details so your reader can locate the source if necessary. This includes material taken from Internet sites.

If you do not acknowledge the sources of your material, you may be accused of plagiarism because you have passed off the work and ideas of another person without appropriate referencing, as if they were your own.

RMIT University treats plagiarism as a very serious offence constituting misconduct.

Plagiarism covers a variety of inappropriate behaviours, including:

- Failure to properly document a source
- Copyright material from the internet or databases
- Collusion between students

For further information on our policies and procedures, please refer to the [University website](#).

Referencing guidelines

Where referencing is required, use RMIT Harvard referencing style for this assessment.

Refer to the [RMIT Easy Cite](#) referencing tool to see examples and tips on how to reference in the appropriated style. You can also refer to the library referencing page for more tools such as EndNote, referencing tutorials and referencing guides for printing.

Penalties for late submissions

Assignments received late and without prior extension approval or special consideration will be penalised by a deduction of 10% of the total score possible per calendar day late for that assessment.

Assessment declaration

When you submit work electronically, you agree to the [assessment declaration](#).