> ## FIT5137 S2 2023 Assignment 1:Take home test (5%)
> ## Due date: Week 4 - Wednesday, 16 August 2023, 11:55pm (AEDT)

A. Learning Outcomes

LO1. Design multi-dimensional databases and data warehouses.
LO2. Use fact and dimensional modelling.

B. Case Study: G9 Education

G9 Education is a private Australian provider of high quality early childhood education and care. They have recently implemented a simple database system. The system contains information about their childcare centres, teachers and children attending the centre. The simple database system consists of the following tables and attributes:
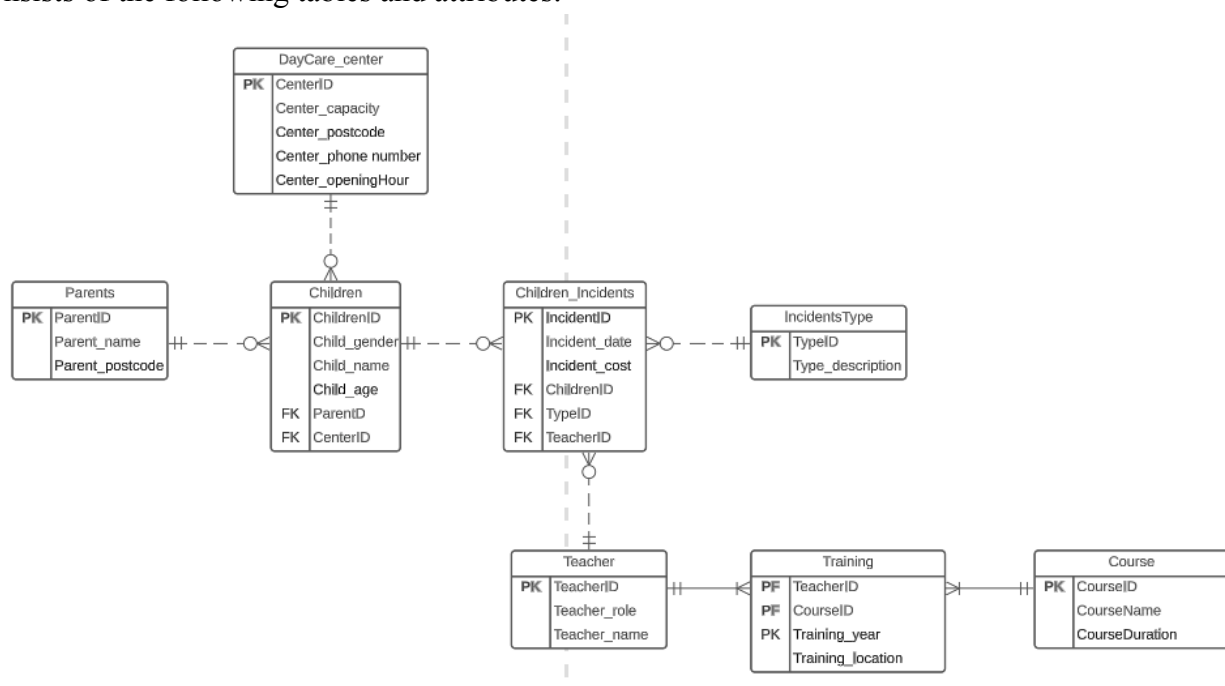
Due to an increasing number of incidents in the day care centres, the manager of G9 Education needs a data warehouse for analysis purposes. Analysis is needed to identify the most common type of incident, the total cost per type of incident, etc.

You are required to design a small data warehouse to keep track of the statistics. The manager is particularly interested in analysing the *number of incidents* and *total incident cost* by incident type, month (e.g., Jan, Feb), daycare center, teacher, and age group points of view.

For the age group, the manager wants to be able to drill down based on two different *age groups*: pre-kinder (between 1 and 2 years old), and kinder (between 3 and 5 years old).

The operational database tables from G9 Education can be accessed through the **MonChild** account. You can, for example, execute the following query:

```
select * from MonChild.<table_name>;
```

C. Tasks

1. Develop a G9 Education star schema. Identify the fact table, dimensions and attributes required to support the schema. The result of this task is a star schema diagram. You can use any drawing tool, such as Lucidchart, to draw the star schema.

   > **Note:**
   > **Please provide a concise and well-reasoned explanation for your design choices. Clearly articulate the reasons for choosing this particular scheme for the G9 education data, <u>ensuring that your design aligns with the assignment scenario.</u>**

2. Validate your star schema using the two-column table methodology. You are required to illustrate the two-column tables for all fact measures and dimensions in your star schema design.

3. Write the SQL commands to create the fact and dimension tables. You need to create a script file containing the appropriate SQL commands to create the fact and dimension tables. Each dimension must have an ID or a surrogate key. The operational database tables are accessible through the **MonChild** account.
   Note: The result of this task is the SQL commands. You will also need to show SQL commands and the contents of the tables that you have created.

4. Write the SQL commands to answer the following queries using the star schema you created above, and display the results:
   A. Show the total number of incidents and total incident costs by age group.
   B. Show the total number of incidents and total incident costs for the teachers whose roles are Early childhood teacher.
   C. Show the total number of incidents and total incident costs by incident type in May.
   D. Show the total number of incidents and total incident costs by daycare center.
   E. Show all information about the teacher who has the lowest number of incidents, including:teacherID,teacher_role,teacher_name,total_num_incident,total_incidentCost

D. Submission checklist:

1. One **pdf file** containing:
    A. The star schema diagram – ***Task 1***
    B. The Two-Column Table Methodology illustration – ***Task 2***
    C. The SQL commands to create all dimension and fact tables, <u>as well as</u> the contents of these tables (the contents can be in the form of screenshots) – ***Task 3***
    D. The SQL commands to answer the queries in ***Task 4*** <u>and</u> the query results. You can take screenshots of the results.
    E. Name your pdf file in this format: THT_ <mark>Yourname</mark>_StudentID.

2. Two **.sql file** containing:
    A. The SQL commands to create the data warehouse.
    B. The SQL commands to answer the queries in *Task 4*.
    C. Save all the .sql files from above to a folder. Name the folder THT_SQL_StudentID.
    D. ZIP the folder to **THT_ SQL_StudentID.zip**. This **must** be a ZIP file and **not** other types of compressed folder. The zip file should contain the prescribed files as listed in the Submission Checklist.

(<u>**Note: The marker will run the .sql file. You need to ensure that the SQL commands are written correctly and clearly with proper comments in the .sql file and runnable without errors**</u>).

E. Submission Method

- The assignment submission must be made through Moodle by the due date: Wednesday, 16 August 2023, 11:55 PM (AEDT).
- The submission of this assignment must be in the form of **a single PDF file AND a single ZIP file**. <u>No other forms will be accepted.</u>
- Penalty for late submission: 10% deduction for each day, including the weekends.
- Submission Cut-off time: **Wednesday, 23 August 2023, 11:55 PM** (Submission link will be unavailable after the cut-off date).

F. Marking Criteria:

FIT5137	Take Home Test Marking Criteria	
TASK	**MARKING CRITERIA**	**MARK**
1. Star Schema	Identifying and creating dimensions. Adding correct attributes to dimensions. Identifying and creating fact measures. Adding correct dimension IDs into fact. Adding correct relationship between fact and dimension table Using correct design notations and relevance towards the case study.	28
2. Two-Column Table Methodology	Creating two-column tables for related dimensions. Using appropriate discrete value selection and relevance towards the star schema design.	8
3. Data Warehouse Implementation	Correct schema implementation following the star schema design in Task Code runs without error.	44
4.Queries	Relevant query implementation. Correct fact and dimensions join to answer the queries. Code runs without error.	15
Additional Criteria	PDF document readability with required screenshots SQL Code readibility with proper comments Timely submission of correct and relevent files	5

G. Authorship

This assignment is an **individual assignment** and the final submission must be identifiably your own work. Breaches of this requirement will result in an assignment not being accepted for assessment and may result in disciplinary action.

As per the University's policy on the guidelines and practice pertaining to the usage of Generative AI, this assignment restricts all use of generative AI. In this assessment, **you must not use generative artificial intelligence (AI) to generate any materials or content in relation to the assessment task.** In addition, generative AI tools may produce inaccurate content or design in the area of star schema design, which could have a negative impact on students' understanding of data warehouse topics. Therefore, the use of generative AI is prohibited for this assignment.

H. Late Penalty:

Late assignments submitted without an approved extension may be accepted up to a maximum of **seven days** with the approval of the Chief Examiner and/or Lecturer but will be **penalised at the rate of 10% per day (including weekends and public holidays)**. Assignments submitted more than seven days after the due date will receive a zero mark for that assignment and may **not receive any feedback**.

Please note:

- An inability to manage your time or computing resources will not be accepted as a valid excuse. (Several assignments being due at the same time are a fact of university life.)
- Hardware failures, whether of personal or university equipment, are not normally recognised as valid excuses. Failure to back up assignment files is also not recognised as a valid excuse.

I. Special Consideration:

From this semester onwards, students will no longer seek extensions from the Chief Examiner/Unit Teaching Team. All extensions / special considerations will now be handled by the central Spec Con team. **Please do not email teaching staff to request an extension or special consideration.**

Extensions and other individual alterations to the assessment regime will only be considered using the University Special Consideration Policy. Students should carefully read the [Special Consideration website](), especially the details about what formal documentation is required.

All special consideration requests should be made using the [Special Consideration Application]().

Please do not assume that submission of a Special Consideration application guarantees that it will be granted – you must receive an official confirmation that it has been granted.

G. Getting help and support:

What can you get help for?

- ***Consultations with the Teaching Team***
Talk to the Teaching Team: https://lms.monash.edu/course/view.php?id=162086&section=2
- ***English language skills***
Talk to English Connect: https://www.monash.edu/english-connect
- ***Study skills***
Talk to a learning skills advisor: https://www.monash.edu/library/skills/contacts
- ***Counselling***
Talk to a counsellor: https://www.monash.edu/health/counselling/appointments

K. Plagiarism and Collusion:

Monash University is committed to upholding standards and academic integrity and honesty. Please take the time to view these links.

[Academic Integrity Module]()

Student Academic Integrity Policy

Test your knowledge, collusion (FIT No Collusion Module)

All the best with your Take Home Test!