FIT3003 S2 2022 Take Home Test (10%)
## Due date: Week 5 - Monday, 22 August 2022, 11:55pm (AEST)

Learning Outcomes:

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

Case Study: Pet9

Pet9 is a pet services salon that offers professional, cage-free services to domestic pets, particularly dogs and cats. The staff at Pet9 are well-trained in different specialised services, including washing, pet grooming, tidy-up, and massages. Pet9 is well known for using environmentally friendly products in their services. Each visit is charged based on the service, pet type and pet size. Customers can choose either a junior, senior, or professional staff member to provide the services. An invoice is issued for each visit.

The following entity-relationship diagram (ERD) shows the data stored in Pet9's operational database:
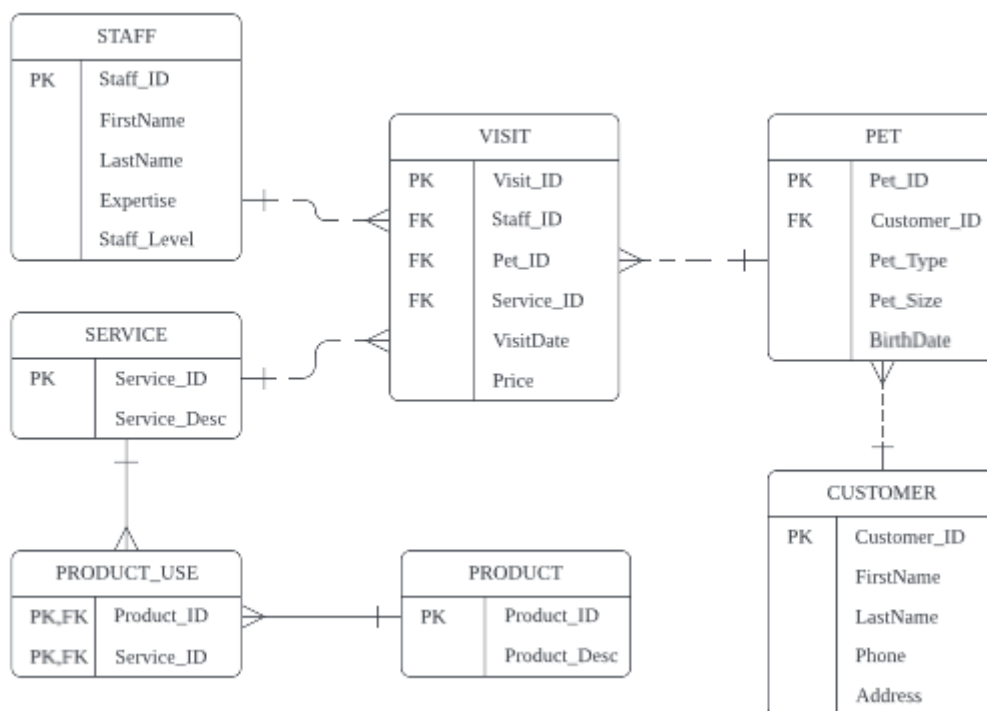
(Note: Customer_ID is the FK for the Pet table. The original typo error(PK) has been removed)

After being in operation for more than 5 years, the management at Pet9 would like to build a data warehouse to analyse their business performance. The data warehouse should provide the following information:

1. The number of visits by seasons and by pet types,
2. The number of visits handled by different levels of staff expertise,
   (Note: level of staff expertise is based on the Staff level, e.g.: Junior,Senior, etc)

3. The number of visits by different products,
4. The total income by services, for different pet types, and
5. The total income by different products.

The seasons are defined as: summer (December to February), autumn (March to May), winter (June to August) and spring (September to November).

You are required to design a data warehouse for Pet9. The operational database can be accessed through the **Pet9** account. You can, for example, execute the following query:

select * from Pet9.<table_name>;

Example: select * from Pet9.Product;

Tasks:
[1]. Develop a star schema named ***Pet9_DW***. Identify the fact table, dimensions, and attributes required to support the schema. The result of this task is a star schema diagram. If you are using a bridge table, make sure to include the <u>weight factor</u> and <u>list aggregate attributes</u>.
You can use any drawing tool, such as Lucidchart or draw.io, to draw the star schema.

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

[3]. Write the SQL commands to create the fact and dimension tables. You must create a runnable 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 **Pet9** account. The results of this task are the SQL commands. You will also need to show the contents of the tables you have created.

[4]. Write the SQL commands to answer the following queries. For each of the following queries, write the SQL and show the results:
a) The number of visits by season and by pet types,
b) The number of visits handled by different levels of staff expertise,
   (Note: level of staff expertise is based on the Staff level, e.g.: Junior,Senior, etc)
c) The number of visits by different products,
d) The total income by services, for different pet types, and
e) The total income by different products.

Submission checklist:

A. One **pdf file** containing:
    i.   The star schema diagram – ***Task 1***
    ii.   The Two-Column Table Methodology illustration – ***Task 2***
    iii. The SQL commands to create all dimension and fact tables, <u>as well as the contents of these tables</u> (the contents can be in the form of screenshots) – ***Task 3***
    iv. The SQL commands to answer the queries in ***Task 4*** <u>and</u> the query results. You can take screenshots of the results.

B. Two **.sql file** containing:
    i.   The SQL commands to create the data warehouse.
    ii.   The SQL commands to answer the queries in *Task 4*.

    (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).

    Zero marks will be given for the implementation if

    ● **SQL files are not runnable, or**
    ● **SQL files are missing, or**
    ● **Wrong SQL files are submitted, including showing inconsistent SQL statements between pdf and SQL files.**

C. Name your **pdf file** containing all tasks listed in [A] in this format: **THT_ Report_StudentID.**

D. One **ZIP folder**:
    i.   Save all .sql files from [B] and in one folder.
    ii.   Name the folder **THT_SQL_StudentID**.
    iii. ZIP the folder to **THT_ SQL_StudentID.zip**. This **must** be a ZIP file and **not** other types of compressed folders. The zip file should contain the prescribed files as listed in the Submission Checklist.

Submission Method:

1. The assignment submission must be made through Moodle by the due date: **Monday, 22 August 2022, 11:55 PM (AEST).**

2. 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>

3. Penalty for late submission: 10% deduction for each day, including weekends.

    Submission cut-off time: **Monday, 29 August 2022, 11:55 PM.** The submission link will be unavailable after the cut-off date.

Marking Criteria:

TASK	MARKING CRITERIA	MARK
FIT3003 S2 2022 Take Home Test Marking Criteria		
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

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.

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=140811&section=2 (Clayton)
  https://lms.monash.edu/course/view.php?id=139836&section=18 (Malaysia)
- ***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

Extensions:

If you are experiencing difficulties that you think will impact your ability to meet this deadline, you may apply for an assignment extension. You must apply **no later than two university working days after the due date** of this assignment.

The extension application can be found under *Moodle > Assessments > How to Apply for an Extension*. Please allow two business days for your application to be processed. Please ensure your application is supported by appropriate documentation.

Special Consideration:

Students should carefully read the [Special Consideration website](#), especially the details about the formal documentation 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.

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**.

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)](#)

<u>END OF TAKE HOME TEST</u>