IFB130 Database Management
Project - part A

Project overview
The IFB130 project gives you an opportunity to apply all the concepts and skills you acquire in the unit to a realistic database design scenario and reflect on the data requirements of an organisation.

The submission is divided into 2 parts due at different times during the semester. These parts will cover:

A. Design of a database
B. Creation and use of a database

The task for part A
You are to design a solution for a database management system that addresses the needs of an interactive visualization tool.

In this part you will design a database by:

1. Creating an ER model for the database based on a use case;
2. Creating a relational schema for the database; and,
3. Assessing an alternative relational schema for the database using normalisation.

Weighting
Part A is worth 15 marks, for 15% of the unit.

Groups
This assignment can be done individually or in pairs.

If you choose to work in a pair, only one student should submit the assignment. Please provide the name and student number of the person you worked with on the front page. No consideration will be given to students who claim they did more work in their pair than the other student because this assignment can be done individually.

Doing the assignment in undeclared groups, or groups larger than two students, will be treated as plagiarism. Pairs that work together and then split due to difficulties must not submit any of the same work, or it will be treated as plagiarism.

Due date
Monday 2 April 2018 at 11:59 pm (end of week 6).

Submission
For this assessment you should submit a professionally presented Word document containing your solution to each task. ER models should be created using a computer drawing program, such as draw.io or Lucid Chart. Partial or fully handwritten submissions will not be accepted. This needs to be uploaded to the submission link on Blackboard.

Late submission
Assessment work submitted after the due date will be marked only with an approved extension (MOPP E/6.8.2). Assessment work submitted after the due date without an approved extension or, where an extension has been granted, after the extended due date, will not be marked and a grade of 1 or 0% will be awarded against the assessment item.

Please note: extensions will not be granted for group submissions, regardless of the reason.

Background

You have been hired by the Brisbane based company Pulse Learning. Pulse Learning specialises in data visualisation to help educators better understand their student's health and improve the workload within their units. The product they would like you to design will display correlations between sleep duration and assignments. The idea is that students will be provided with smart watches that will record their daily sleeping patterns throughout the semester.

After hours of meetings and a lot of emails the company have agreed on the screenshots of what educators should be able to see. It is now your task to **design the database** to allow the data to be stored so that it can ultimately be displayed.

The scenario – *data visualisation*

Pulse Learning will build a database for each of their clients (Universities). Therefore, this scenario will refer to "the University" as one such client.

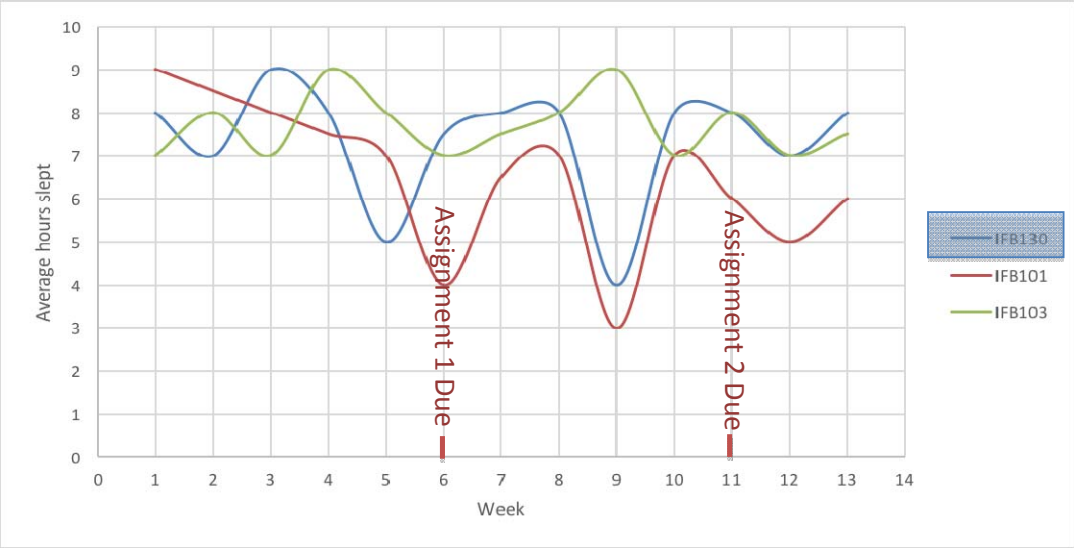
Each **student** in the university is provided with one smart watch, which sends his or her **sleep patterns** (daily totals for time spent asleep and time spent awake) to the server. Students can choose which smart watch they prefer. The available choices are FitBit, Suunto, Apple Watch and the Samsung Gear. Every smart watch also has a unique serial number.

The University provides Pulse Learning with information on the units and student enrolments. This includes:

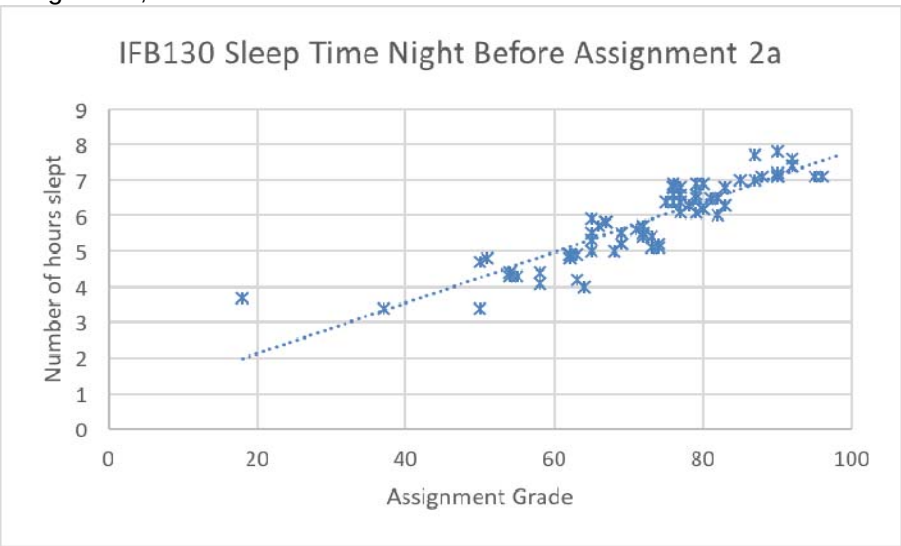
- basic information about students (their student id, first and last name, email address, full mailing address and phone number. Each student could have up to three phone numbers recorded.)
- smart watch information about students (which type of watch they were given, and what is its serial number).
- basic information about the available **units** (the unit code, the year and semester the unit is available, the name of the unit)
- which unit(s) each student is enrolled in
- when the due dates for **assignments** are in each unit and the name of the assignment
- the grades of each student for each assignment
- list of student pairs in the buddying system
- staff id, full names and email addresses of **tutors** assigned to each student in each unit.

A dynamic, overall grade point average (GPA) should feature in each student's file and is calculated from all grades in their assignments.

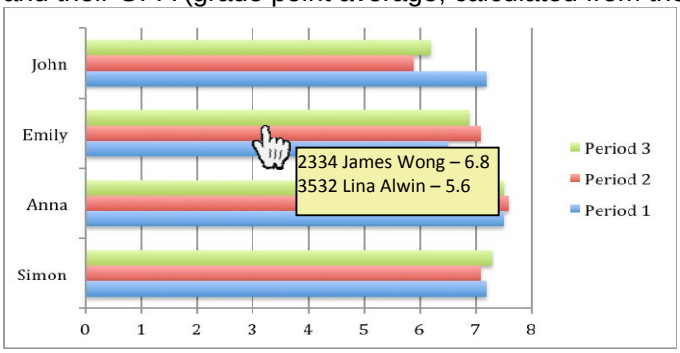
The first interactive plot they would like to display would show how students sleep throughout the semester, averaged by week, in a number of selected units. When clicking on one unit, the teacher should be able to see on the graph when are the assignments due for that particular unit, as shown below.



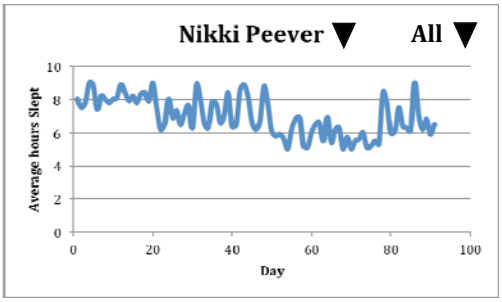
The second interactive plot they would like to display is, within a selected unit and for a selected assignment, of each student per the number of hours they slept and their grade for this particular assignment, as shown below.



Some students can have a buddy if they wish: a buddy is another student who has done the course before. A student can only have one buddy throughout their studies. However, a student can serve as buddy to as many students they wish to, or none. The third interactive plot should show for each buddy, how well the students they are partnering with are sleeping within 3 periods of the semester, on average. By clicking on a buddy's line, the name of the students they are supporting should appear, as well as their student number and their GPA (grade point average, calculated from their grades in all assignments across all their units) .



Finally, for each unit students enrol in, they are assigned one tutor. Tutors are staff members of the University, and they can support students across several units. The fourth interactive plot should show the sleeping paterns of students supported by a given tutor. There should be an option to show this averaged across all students in all units the tutor is involved in, or averaged across students in each unit the tutor is involved in.



Part A tasks

Part A requires you to complete a number of sequential tasks to fulfill the requirements of the scenario. In this part you will design a database by:

1. Creating an ER model for the database;
2. Deriving the relational schema for the new database;
3. Demonstrating the constraints of the relational schema through examples; and
4. Verifying a previous (different) relational schema for the database using normalization;

Task 1 [5 marks]

Considering the use case provided, create an **Entity Relationship Diagram** that correctly models the data requirements of *Pulse Learning.* You must ensure that everything that is required for the visualisations and provided by the University is recorded. Your ER diagram (using UML notation) needs to show:

- the **entity types** (0.5 marks),
- the **binary relationship types** (1 mark),
- the **complex relationship type** (1 mark),
- the **multiplicity** of relationship types (1 mark),
- the **attributes** (and if applicable their **domains**) of entity types (0.5 marks),
- the **attributes** (and if applicable their **domains**) of relationship types (0.5 marks), and
- the **primary key** of each entity type (0.5 marks).

State any assumptions you make.

Task 2 [5 marks]

Derive a **Relational Model** from the conceptual schema you provided in Task 1 (it should match exactly).

1. List all the relations derived from the entity types of the ERD, and their initial attributes. (1 mark)
2. For each relationship type in the ERD justify whether it leads to a new relation, a new attribute in a relation derived from an entity type, or to a merger of two relations derived from two entity types. (1 marks)
3. List all the relations in the final relational model, and all their attributes (1 mark)
4. List all the constraints included in the conceptual schema in addition to the relevant primary keys and foreign keys. (2 marks)

Task 3 [2 marks]

Create four rows of sample data for each of the relations you identified in Task 2.

Based on one of the relations of your relational model and its sample data, specify an example of each of the following, and explain why you selected it:

- A delete operation that would run successfully
- An update operation that would run successfully
- An update operation that would not run successfully
- An insert operation that would not run successfully

Task 4 [3 marks]

Pulse Learning had an intern who prototyped a former version of the database, but they could never implement it, and they have already been told this is really **poor design**. This is what it looks like, with some sample data:

Student

ID	Units
3345	IA33 FR12 QA25
3279	FR54 QA25
8754	IA33
4579	QA45

Tutor

Student ID	Student Name	Unit Code	Tutor ID	Tutor First Name
3467	Adele X	IA33	3345	Nikki
4666	John Snow	FR54	3345	Nikki
2357	Malcom Turnbull	FR54	5798	Jack
8097	Mickey Mouse	FR54	5798	Jack

Coordinator

Unit code	Unit Name	Coordinator ID	Coordinator Name
FR54	French Studies	2556	Emma Macron
IA33	Data Mining	7643	Mark Zuck
QA25	Natural Language Processing	2234	Laurianne Sitbon
QA54	Search Engines	2234	Laurianne Sitbon

For each of the 3 relations proposed (Student, Tutor, Coordinator), explain which normal form (0NF, 1NF or 2NF) it is in, and why.

BONUS MARKS (2 marks, but your overall grade for this assignment will be capped to 15)

Write a scenario and the corresponding ERD that would display at least 5 entity types, at least one relationship type of each multiplicity (one-to-one, one-to-many, many-to-many), at least one complex relationship type (degree 3 or more), and at least one recursive relationship type. The model should also feature at least one derived attribute and one composite attribute.

Write in a statement if you agree for your scenario to be used in the current and future offerings of the unit as additional practice material (this decision will not affect the marks for this bonus question). If you wish, you can also share it on Slack with other students as soon as it is ready. If you do so, please make sure to email a copy to the teaching team beforehand in order to avoid risks related to plagiarism.