ACS-4904   Assignments 3, 4, 5   due on or before Tuesday April 5, 2016

For assignments 3, 4, 5 choose 3 different questions from below.

For each of these questions you are either working with an MS Access database or PostgreSQL database.

For MS Access you must submit the database with data and any required queries (and where pertinent supply instructions for running queries)

For PostgreSQL you can demonstrate your solutions to the instructor, or you may submit the necessary SQL to recreate your database along with any required queries (and where pertinent supply instructions for running queries)

Your marker may need to verify your work by re-creating a solution using the queries you provide.

1. Create the databases shown in figures 6-4 and 6-5 (pages 124-5). Populate your databases with data shown in those figures and more as is needed. Note that Policy is a type 2 dimension and must have a current indicator field, an effective date and an expiry date.

    Develop pairs of queries, for figure 6-4 and for figure 6-5, to demonstrate your designs are equivalent.

2. Consider the paper:

    http://www.cis.drexel.edu/faculty/song/publications/p_Song_M_N_DMDW_final.pdf

    where multi-valued dimensions are discussed relating to health care diagnoses. You should only need to focus on sections 3.0 and 3.1.

    Create the schema shown in Figure 2 and populate with data from Tables 1, 2, 3.

    Create the queries shown in Situations 1 and 2, and create a query to list patients who have diagnoses of Lung and Cancer.

3. Given a recursive hierarchy (e.g. 'reports to' in Northwind) use SQL to create a bridge table as discussed in our text. Your code would access the standard OLTP table and create the hierarchy bridge table.

4. Like the above but this time for the nested sets implementation of a hierarchy.

5. Figure 9-10 (page 210) shows a bridge table used for a case where there is a multivalued attribute *industry*. Create the database shown but with more than the data given in Figure 9-10; include order facts and more customers that have some industries in common with *CustomerC.*

    Create the queries on pages 211-2 and verify you get the results they show.

6. Create the Base Star <u>and two</u> of the Aggregate stars in Figure 15-2. Populate your database and create two queries that answer the same question (one uses the aggregate and the other does not). The query based on the aggregates must be obtained (manually) using Kimball's algorithm for rewriting a query.

7. Implement the schemas in Figure 11-1 and 11-2. Create queries to determine the account balance at the end of the day (Feb 9) for account 7922-3002.

8. Implement the schema in figure 12-2 and write two sets of queries (see page 294) : one set using Count(…) and the other set using SUM(…).

9. Implement the schemas in figure 14-2. This exercise requires that you first implement the schema in the top half of the figure, and then populate the schema in the lower half using the necessary DDL/SQL. Create two queries to generate the types of reports shown on page 336.

10. Create a source student table and a corresponding Type 2 student dimension. Create ETL to update the student dimension. To simplify you are concerned with updates to one attribute of the student dimension.

11. Consider the Declarations Schema shown later in this document. Modify the schema to permit a student to declare multiple majors. For example, at the UW we allow a student to declare any number of majors and so a student can graduate with a major, with a double major, with a triple major, etc.   As discussed in Chapter 9, a bridge table can be used to

accommodate the requirement that a student be allowed to declare multiple majors. Use the approach presented in Chapter 9 to modify the schema.

Provide rows in the Declarations fact table, the bridge table and the Major table to accommodate 3 cases

- A student declaring a single major in IS (Information Systems)

- A student declaring a single major in HI (Health Informatics)

- A student declaring a double major in HI and BPHE (Bachelor of Physical and Health Education)

Design a query to list the first names of students who have declared a major in HI

open to suggestions …

E.g. For those who know the Evangelium Project, design a star schema for Evangelium. Give examples of queries

More questions are likely to be added to the above.
Where necessary clarifications to the above can be added.