Assignment for Data Warehouse course
Fall 2022

Purpose: Design and manipulation of a data warehouse.

Problem: The University of Savannah at Richmond Hill (USRH) has collected data about her graduates for several years (Data are attached at the end of the assignment). The authorities of the university want you to design a *graduates* data warehouse. This data warehouse facilitates the study of the number of graduates for different majors, degrees, colleges, GPAs, semesters, and/or any combination of them. To satisfy the needs of the authorities do:

1. Design a data warehouse using one of the schemas presented in class. Justify your choice of schema. Your justification and explanation must be included in a word document. (Name this document Project Report). For your schema, identify and justify dimensions and their attributes, fact table and its attributes, concept and schema hierarchy (ies) for each dimension. All of your dimensions, their table's definitions, and justifications along with all of your diagrams (computer generated diagrams are the only type of diagrams that are acceptable) must be included in the document Project Report under proper headings of your choice.

2. Create your data warehouse (as a database) using MYSQL (a free relational database management system, RDBMS, that can be downloaded from internet). Populate your data warehouse using the attached file. Your mechanism for finding "facts" out of the attached dataset for using in the subject table of your data warehouse is of a great interest. Clearly describe how your mechanism of choice does the work.

   Your explanations and justifications about the data warehouse creation and the mechanism for the "fact" findings must be included in the Project Report document under proper headings of your choice.

3. Write an interface for the data warehouse (Computer language of your choice) that mimics an OLAP system. That is, for the cuboid of your data warehouse user can issue a sequence of OLAP operations to express a query within the interface.

   Your interface must guide the user upon choice of an operation to complete the rest of codes needed by the operation. Therefore, as one solution, a template may be created by the interface to guide the user to fill in the template. (An example was provided in the recorded session of (8-DW-OLAP Operations-II) available in Folio)

   The interface:

   a. Accepts a query expressed by the Olap operations as input
   b. Converts the query into SQL statements
   c. Submits the SQL statement to MYSQL for implementation and displays the result produced by MYSQL.

   (The details of these steps may be found in the recorded session of (8-DW-OLAP Operations-II) in Folio)

4. Test your interface by implementing the following four queries using OLAP operations: (Express each query using OLAP operations, convert the OLAP operations into SQL statements, and then implement the SQL statements.)
    a. *Get the number of graduates for the Cyber College*
    b. *Get the number of graduates with a BS in Computer science in summer of 1988.*
    c. *Get the number of graduates with a high GPA in the college of business.*
    d. *Get the number of international students graduated in years 89, 90, and 91*

    For each one of the above queries, include the followings in the Project Report under proper headings :
    i. Query itself,
    ii. The query solutions using OLAP operations (these OLAP operations will be used as input to the interface),
    iii. List of SQL codes produced by your interface from translation of the output of step ii (above) into SQL, and
    iv. The outcome produced by MYSQL implementing the output of step iii. (You are not permitted to type this outcome. I must see the outcome as it is generated by MYSQL. You may take a screen shot and include the image in your Project report).

5. Add a cover page, table of contents, table of figures, introduction, and a conclusion to your Project Report. Also, add your program codes as an appendix to the Project Report and submit your Project Report by email to me not later than **Dec. 7, 2022**

6. You have almost **six weeks** to finish this assignment. However, you must submit your assignment on an **incremental basis** as follows:
    A. The deadline for submitting the completed **item 1** (above) along with the Project Report of the assignment for the completed item 1 to me is **Nov. 6, 2022**. (You can consider this deliverable as a take home test for your **Exam 2** and it will be graded accordingly.)
    B. The deadline for submitting the completed **items 2** (above) along with the Project Report of the assignment for the completed items 1and 2 to me is **Nov. 16, 2022**. (You can consider this deliverable as your first assignment, and it will be graded accordingly.)
    C. The deadline for submitting the completed **item 3** (above) along with the Project Report of the assignment for the completed items 1, 2, and 3 to me is **Nov. 28, 2022**. (You can consider this deliverable as your **second assignment**, and it will be graded accordingly.)
    D. The deadline for submitting the completed **items 4 and 5** (above) along with the Project Report of the assignment for the completed items 1, 2, 3, 4, and 5 to me is **Dec. 7, 2022**. (You can consider this deliverable as a take home test as your **Final Exam**, and it will be graded accordingly.)

Collected Data by the University authority

COLLEGES

NAME	MAJORS
Cyber College	Computer Sc Information Sc Applied Sc
College of Business	Accounting Business Admin Economics
College of Education	Elementary Ed Secondary Ed
College of Art and Science	Biology Chemistry

If you find majores other than the above ones in your dataset, then they need to be fixed because they have been missed during the integration process.

STUDENTS

ID	Name	Address	Major	Degree	GPA	MonthDay	Y	Status
356	C. Brown	Denver	Applied Sc	PhD	4.0	Dec.15	89	I
667	K. Chen	Conway	Elementary	EdD	3.2	Aug. 15	92	O
529	M. Yu	Conway	Secondary	MS	3.2	May 15	90	I
552	K. Ahmad	Atlanta	Secondary	EdD	3.9	July 48	87	O
687	Y. Morty	Atlanta	Computer Sc	BS	3.9	July 4	84	N
688	Y. Crema	Little Rock	Biology	AS	2.2	June 1	87	N
653	C. Alshukri	Denver	Applied Sc	PhD	4.0	Dec.15	91	O
100	A. Jones	Little Rock	Computer Sc	BS	3.1	May 15	84	I
986	Q. Brady	Pine Bluff	Computer Sc	MS	3.8	May 15	87	N
579	U. Jones	Little Rock	Computer Sc	BS	3.1	May 15	84	N
200	B. Brady	Little Rock	Applied Sc	PhD	3.2	May 15	92	I
300	C. Cook	Little Rock	Biology	BS	2.2	June 1	87	I
400	D. Morty	Atlanta	Biology	BS	3.9	July 4	84	N
422	U. Yu	Conway	Secondary	MS	3.2	May 15	90	O
423	U. Mooshe	Atlanta	Secondary	EdD	3.9	July 4	90	O
401	D. Morgan	Conway	Biology	BS	3.9	July 4	87	I
702	Y. Brady	Conway	Applied Sc	PhD	3.2	May 15	87	I
201	W. Brady	Conway	Applied Sc	PhD	3.2	May 15	87	I
202	A. Bank	Conway	Applied Sc	PhD	3.2	July 4	87	N
415	T. King	Conway	Elementary	EdD	3.2	Aug. 15	90	N
901	C. Crema	Little Rock	Biology	AS	2.2	June 1	87	O
222	D. Abul	Savannah	Economics	BA	3.9	Aug. 15	90	N
333	G. Halk	Atlanta	B Admin	MBA	2.9	Aug. 15	90	N
431	W.Goos	Savannah	Accounting	BA	4.1	Aug. 15	90	O
421	T. Chen	Conway	Elementary	EdD	3.2	Aug. 15	91	I
524	U. Alramzy	Denver	Applied Sc	PhD	4.0	Dec.15	92	I
689	Y. Morgan	Conway	Biology	BS	3.9	July 4	87	N

195	D. Red	Seattle	Info Sc	BS	2.1	Dec. 15	90	O
703	Y. Bank	Conway	Applied Sc	PhD	3.2	July 15	87	O
665	Z. Silver	Seattle	Elementary	EdD	3.3	Aug. 15	91	I
739	A. Moon	Little Rock	Computer Sc	BS	3.1	June 1	84	N
839	Q. Robe	Pine Bluff	Computer Sc	MS	3.8	June 18	87	N
939	U. Wang	Little Rock	Computer Sc	AS	3.1	July 15	84	O
356	C. Brown	Denver	Applied Sc	PhD	4.0	Dec.15	89	O
500	K.Bronz	Dallas	Chemistry	MS	3.2	July 4	88	O
132	A. White	Orlando	Computer Sc	AS	3.9	Aug 15	92	N
100	A. Jones	Little Rock	Computer Sc	MS	3.2	May 15	87	N
232	K. Blue	Little Rock	Computer Sc	AS	3.1	Aug 15	91	I
352	C. Black	Pine Bluff	Applied	PhD	3.8	Aug 15	90	O
600	C. Jones	Dallas	Info Sc	BS	4.0	Dec. 15	88	O
680	Y. Brady	Pine Bluff	Computer Sc	MS	3.8	May 15	90	O
682	Y. Cook	Little Rock	Biology	BS	2.2	June 1	87	O
400	D. Morty	Seattle	Info Sc	BS	2.1	Dec. 15	88	I
452	N. Brady	Little Rock	Applied Sc	PhD	3.2	Aug 15	88	I
590	X. Cook	Little Rock	Biology	AS	2.2	June 1	92	N
690	D. Brown	Atlanta	Biology	BS	3.9	July 4	88	N
175	K. Glob	Dallas	Applied Sc	PhD	3.2	July 4	89	N
681	Y. Brady	Little Rock	Applied Sc	PhD	3.2	May 15	91	N
185	X. Jones	New York	Info Sc	BS	4.0	Dec. 15	90	N
700	A. Boss	Seattle	Elementary	MS	3.3	Aug. 15	84	I
800	K. Queen	Conway	Elementary	EdD	3.2	Aug. 15	89	O
416	T. Gola	Conway	Secondary	MS	3.2	May 15	90	N
420	T. Silver	Seattle	Elementary	EdD	3.3	Aug. 15	91	N
900	B. Gola	Conway	Secondary	MS	3.2	May 15	87	I
955	K. Booth	Atlanta	Secondary	MS	3.9	May 15	87	O
956	C. Jessy	Orlando	Applied Sc	MS	4.0	Dec.15	89	O
150	A. Cool	Orlando	Computer Sc	MS	3.9	Aug 15	86	I
250	K. Jones	Little Rock	Computer Sc	BS	3.1	Aug 15	86	N
350	C. Cool	Pine Bluff	Applied	MS	4.0	Aug 15	86	N
229	M. Katchal	Savannah	Economics	BA	3.0	Aug. 15	91	N
339	H.Poory	Atlanta	B Admin	BA	3.9	Aug. 15	87	N
439	H. Mishu	Savannah	Accounting	BA	2.8	Aug. 15	87	O
329	G. Kooper	Savannah	Accounting	BA	3.0	Aug. 15	91	N
339	M. Bayraq	Savannah	MBA	MBA	3.9	Aug. 15	87	N
337	N. Mikhy	Savannah	B Admin	BA	2.8	Aug. 15	87	O
650	D. Sorty	Atlanta	Biology	BS	3.9	July 4	84	N
750	K.Gold	Dallas	Applied Sc	PhD	3.2	July 4	84	N
850	F. Jones	New York	Info Sc	BS	4.0	Dec. 15	85	I
950	D. Golden	Seattle	Info Sc	BS	2.1	Dec. 15	85	I
165	A. Silver	Seattle	Elementary	MS	3.3	Aug. 15	85	O
417	T. Dool	Atlanta	Secondary	MS	3.9	July 4-9	92	O
418	T. Brown	Denver	Applied Sc	PhD	4.0	Dec.15	92	O
419	T. Red	Seattle	Info Sc	BS	2.1	Dec. 15	92	O
215	K. King	Conway	Elementary	EdD	3.2	Aug. 15	84	N
315	M. Gola	Conway	Secondary	MS	3.2	May 15	84	N
255	K. Dool	Atlanta	Secondary	MS	3.9	July 4	86	N
299	M. Golab	Seatel	Chemistry	BS	3.0	May 15	92	N
199	K. Duell	Atlanta	Biology	AS	4.0	Dec. 15	88	I
473	M. Browny	Seattle	CS	PhD	4.0	May.15	92	I
407	T. Doolak	Seattle	BusinessAD	MBA	3.1	June 1	92	O

205	K. Minges	Conway	Elementary	EdD	3.2	Aug. 15	84	N
305	H. Molar	Conway	Biology	MS	3.2	July 4	84	N
207	K. Goolar	Atlanta	Chemistry	MS	3.9	July 4	86	N
107	M. Kochaal	Savannah	Chemistry	BS	3.9	May 15	92	N
305	M. Searsy	Conway	Biol	MS	3.2	July 4	84	N
207	P. Little	Atlanta	Chemistry	MS	3.9	July 4	86	N
107	M. Big	Little Rock	Chemistry	BS	3.9	May 15	92	O
109	P. Board	Atlanta	Chem	AS	4.0	Dec. 15	88	I
305	M. Korp	Conway	Biology	MS	3.2	July 4	84	N
109	G. Bear	Little Rock	Biology	AS	4.0	Dec. 15	88	O
591	Z. Kochaal	Savannah	Chemistry	BS	3.9	May 15	92	N
391	X. Searsy	Conway	Biology	MS	3.2	July 4	84	N
201	Y. Little	Atlanta	Chemistry	MS	3.9	July 4	86	N
117	W. Big	Little Rock	Chemistry	BS	3.9	May 15	92	O
612	V. Board	Atlanta	Chemistry	AS	4.0	Dec. 15	88	I
305	M. Kartsoo	Conway	Biology	MS	3.2	July 4	84	N
609	N. Gearsal	Little Rock	Biology	BS	4.0	Dec. 15	88	O
615	K. Jiring	Conway	Element.	BA	3.2	Aug. 15	84	N
715	M. Chendo	Conway	Secondary	MS	3.2	May 15	84	N
855	Z. Megart	Atlanta	Secondary	MS	3.9	July 4	86	N
699	A. Woo	Seattle	Elementary	BA	3.0	May 15	92	N

Notes:

Fall semester ends in Dec. 15
Spring semester ends in May. 15
Summer 1 ends in June 1
Summer 2 ends in July 4
Summer 3 ends in July 15
Summer 4 ends in Aug 15

$4 \geq GPA \geq 3$ is high
$3 > GPA > 2$ is medium
$2 \geq GPA$ is low.

Status:	I	International Student
	O	Out of state
	N	Non-of-the above

Degrees	AS	Associate degree
	BS	Bachelor of Science
	BA	Bachelor of Art
	MS	Master of Science
	MA	Master of Art
	MBA	Master of Business Administration
	EdD	Doctor of Education
	PhD	Doctor of Philosophy

The assumption is that students have a unique ID. If you find two students with the same ID, please change one of them to a different 3-digit number.