REAL WORLD DESIGN AND IMPLEMENTATION IN THE STUDENT’S FIRST DATABASE COURSE

Jeanne M Baugh, Robert Morris University, baugh@rmu.edu

ABSTRACT

All Computer Information Systems programs require a database course for their majors. This paper describes an approach to such a course in which real world examples, both design projects and actual database application projects are incorporated throughout the semester. Working on real projects for real users gives the students valuable experience. Database consulting scenarios are also discussed, showing both good and bad examples of real issues a real consultant encountered. Students are expected to apply the traditional database concepts to actual database storage problems. The design of the database is emphasized and a number of “tricks” are offered to aid the student in this process. User interaction is required in connection with the student’s actual individual database project. The details of a database written by the Author to organize all information surrounding the student’s project are presented. Course content is discussed and recommendations are made.

Keywords: Database first course, Real world projects, Database project, and IS education

INTRODUCTION

All students in an Information Systems degree are usually required to take one database course as part of the core for the major [13]. Such a course would emphasize database design and implementation based on a thorough analysis of requirements and information modeling. An introduction to relational database technology would be provided, highlighting the use of Structured Query Language (SQL) and report generation. An advanced course could go into more detail in the areas of advanced SQL considerations, PL/SQL, database performance and security issues, multimedia, parallel, and distributed database management systems, data warehousing, and object-oriented databases. At the Author’s institution, the undergraduate student takes only the first, beginning course.

In the database course described in this paper, students are given as much real world experience as possible. Therefore, a great deal of time is spent in design work of simulated real world problems as well as a detailed design of the student’s individual semester project. If the course material can be made more interesting to the student, then he will be more inclined to want to learn it. Real world projects allow the students to “learn better through a particular domain of their interest” and “see the practical value of what they learned” [18]. Others have had success with a “real world” approach, such as the partnership of a University with a large Insurance company to implement database assignments for introductory and advanced database courses [20]. To give students such experiences in the “real world”, a bridge must be established between business and academia [5]. In teaching beginning database for many years, it has been the Author’s experience that students have a relatively naive view of the complexity of the real-world business.

Students are expected to also communicate with actual users. Database programming skills along with communication skills go hand in hand for successful employment [20]. In fact, employers do demand this of their entry level employees [10]. Thus, large semester projects along with documentation should be key elements of any database course [8]. And, gaining practical experience is invaluable when the student is on that first job search.

DATABASE COURSE EMphasIZING DESIGN WORK

Two different approaches to teaching a database course could be used. One focuses on the actual application software, Oracle, SQL Server etc… The other focuses on the theory of database design and implementation, with the vehicle (the database package) being secondary. This paper describes the later approach. A database can not be implemented until the design is clean [7]. Or perhaps better stated, it can not be implemented successfully if there are normalization problems.

This Author has worked as a database consultant for many years, with the most recent projects being complicated databases in the Orthopedic department of a large city hospital. Many of the experiences are brought into the
Database design is a major element of the course work described in this paper. The students were told on the first day of the course that design would be the hardest thing they would do during the course [4]. After a few weeks, the students did believe this! It is the Author’s opinion that many of the database text books do not spend enough time teaching the actual design of a database [16]. In this course, the students spend at least a month on database design only, before they actually work with any physical database on the computer. Throughout the semester the students are periodically given design problems and one exam is entirely comprised of such problems.

The students are provided with many examples of interactions with users [11]. Communication skills of the database designer are extremely important [1]. Students are taught to “listen” to the user. Students not only need to hone their technical skills, but they need to develop a communication expertise that will allow them the best possible chance at drawing out of the user what their data needs may be. Often the users will not know what to tell the database designer about their data needs. Sometimes they will even try to tell the designer what actual fields should be in a table. They will speak in terms of the outcomes of the database, for example the desired reports. Students are taught to take those required reports and “back design” the database from them. They are also taught to ask questions (as many as possible) to help understand how all of the data fits together. For example, if they are designing a surgical database, and they are told that the Doctor wants to store the type of anesthesia used in the surgery, the student should then ask if there can be more than one anesthesia administered during the surgery. The students are provided with many examples of user issues and some “mistakes” the Author has had to deal with in designing and implementing databases at the hospital. The “anesthesia” is one mistake this Author made on an actual database. By not accounting for multiple anesthesia, the students saw that a mistake can cause a major database table change. They were able to appreciate how difficult it was to deal with this mistake after 6 months of data had been entered. Seeing how actual design errors cause very big problems is a valuable tool. By knowing what the user expects from the database, the design can be adjusted accordingly before data has ever been entered.

Although database text books contain design assignments and projects, this Author has written many the students work through. The kind of text that is written in these design assignments has come from this Author’s many years of dealing with users and helping them define what their needs really are. The Author acts as the user in these scenarios and students must ask as many questions as possible of the user to understand what the design should be. A few examples of these design problems are listed below. (The students were to create the Entity Relationship diagram for each problem). An entity is a distinguishable item and a relationship is the connection between the sets of entities.

1. You and your family have an internet business selling items on EBay. You are having some problems keeping track of which items belong to which individual. You have been asked by your family to organize the EBay records. (some of your items are also consignment items from friends and others) Associated with the sale of the item you will have shipping information, including the customer who purchased the item. The reports you will need include:
   • All information on items sold within a specific time frame
   • Which family member (or friend) is selling the most items
   • All open auctions
   • List of all shipping costs
   • Newsletter to all customers
   • History of the bids on each item

2. You are the coach for a youth soccer team. Your players are children between the ages of 6 and 8. You have both male and female members on your team. You not only want to keep track of goals scored and who the opponents were, you must also know who played in the games, because the players will get a trophy if they attend all 10 games. You must also keep track of who is allowed to pick up each child after practice (the guardian). (Every individual guardian must be listed for each player) The reports you will need include:
   • List of all players and their guardians
   • Record of the games won and lost along with the opponent
   • List of all goals each player scored and against which opponent
• List of all players attending each game
• Mailing list for the end of year banquet

3. You have been hired to work for the Universal Travel Agency as a database administrator. This is a business that books tour packages to the Caribbean exclusively. Their tour packages include both air and land arrangements. Each tour package may have multiple airline reservations associated with it, and may include stops to several islands. The company represents several resorts at each island. The owner also wishes to reward those employees who sell a large number of packages with a monthly bonus. The following reports are among those requested by the owner on a weekly basis:
• All tour package details
• Sales history per employee
• Resorts most frequently chosen by customers
• Airlines most frequently used

The students used these and many other simulated problem statements to work through Entity-Relationship (E-R) diagrams. Often the students worked on these problems in groups. The solution E-R diagram for each was then provided to the students. The youth soccer solution is provided in Figure 1.

Figure 1. Solution to Youth Soccer Design Problem

A number of tricks/or shortcuts have been developed by the Author to help the student to create the E-R diagrams. For example, the students are taught to listen for any plural words in the problem statements. (a patient has surgeries, a student is enrolled in courses, an insurance agent has clients) When a plural exists, that is an indication that the entities should not be stored together. The E-R diagram for the patient and surgeries is in Figure 2.

Figure 2. E-R Design (Patient has multiple surgery)
Another trick the Author has presented concerning the E-R design is a trick to define the cardinality. In the example in Figure 2, cardinality is the relationship between the patient entity set and the surgery entity set. The relationship is a diamond in the diagram and in this diagram it is called “Has”. Is the relationship one-to-one, one-to-many, or many-to-many? Defining the type of relationship is often very difficult for the first time database student [20]. The students are taught to start with one of the entities in the relationship and ask if each of those entities can have multiple connections on the other side. If they say Yes, then an “M” is placed on the other side. Then they go to the other side of the relationship and ask the same question in the opposite direction. Using the E-R diagram in Figure 2, the method is applied as such:

Start with the Patient entity set and ask “Can each patient have multiple surgeries”? If the answer is YES, then an “M” is placed next to the surgery entity set. Then go to the surgery set side and ask, “Can a surgery be related to more than one patient”? In this case the answer is NO, so a “1” is placed next to the Patient entity set. This method will work, no matter which entity set you may start with. This is not the notation database books use for cardinalities, but in learning to create the designs, students do well with this method. Later in the semester the standard notations are discussed [9]. (for example, (1:N) or (M:N) or (1:1)). Students are also given examples of multiple relationships. One example of this is the special case when one of the related entities participates as 1 and the others as N.

After the students have had about a month of design work, they then work with translating the design into an actual database. This process can also be somewhat difficult for the beginning database student. The normalization process is studied later in the semester, but a few rules are defined for the students to help insure that the resulting database will have as few normalization problems as possible.

The rules for defining the physical database from the E-R diagram are:

- Entity Sets (rectangles) become tables
- Relationships (diamonds) that are one to many, do not become tables, but the key from the one side has to be a field in the many side table
- Relationships (diamonds) that are many to many, become tables and the data in that table are at least the keys of the entity sets they relate

STUDENT INDIVIDUAL PROJECT

The students then move on to define their own databases project. They are asked to find a real project to implement. The Author approved the project based on the following guidelines:

- Was the idea a valid one? (a subjective judgment on the part of the Author)
- Was the problem something the student could reasonably finish in the time allotted?
- Would the design provide approximately 10 tables once implemented?
- Were there sufficient relationships among the data? (one to many, many to many, etc…)
- Were there sufficient attributes to be defined?

Some of the project created were:

- Database of Fundraisers and their Products
- Prospective student athlete information for the men's lacrosse team
- Well device and communication information for an oil & gas company
- Track services provided by a company with multiple types of audio visual meeting rooms for rent
- Sales incentive payment information for a financial services company
- Track information relating to children who attend a Church Sunday School
- Patient lab test Appointments and results

After approval of the initial database idea, the detailed design was done. Because they were doing work for a real user, it was important that their designs were correct [2]. They provided a great deal of information concerning their database project and several passes were made to try to create a correct E-R design. It must be noted that a significant amount of this Author’s time was spent with each student to insure that their database design was correct. At this point in the semester, the students really did not have enough experience in design work to know if their design was a good one or not. Since the student was creating a “real” project, this Author felt a responsibility to the
users who have agreed to participate in the student’s learning experience to ensure that the resulting database was correct. Below are some of the specifications for the initial database design assignment:

You are to explain your idea for a database to be created this semester.

✓ Provide a written summary of the database you will model. You must give as much detail as possible as to the information you are modeling. This must be more than one sentence!
✓ List each entity set and provide a short description of the data that will be contained in it. (You may list the fields that you will be storing)
✓ List each relationship and describe it’s functionality
✓ You are to create an initial Entity Relationship diagram for your database project. On this diagram, you are to include the following:
  o All entities sets
  o All relationships (give a name for each)
  o All cardinalities of the relationships
  o Define a key for each entity set (place it on the ER diagram)

It became apparent to the Author that because each of the students were working on a different project with different issues, a way to organize this information was essential. Therefore, the Author wrote an Access database to store and track all information. The Author also created an Excel spreadsheet with various column headings related to information about the student’s individual database project. After the student’s database project had been approved, they each filled in a spreadsheet and sent it to the Author. The information was then imported into the Access database. All students were able to “view” the data in the database. Again, a “real world” database was used in the course. And, the “real world” issue of how to integrate spreadsheet data into an existing database was a teaching moment as well. Often data to be stored in a database may come from some external electronic source [15]. Figure 3. shows a sample screen of imported data from a student who created a database for a friend who owned a bar. The data on the left of this figure was imported from the student’s spreadsheet. And, on the right of this figure you can see several reports that were assigned to this student by the Author.

![Database Project Information](image)

**Figure 3.** Sample screen of student database project

Students were given access to this database so they could not only track requirements assigned to them but also view a summary of the work of their fellow students. This database also kept track of individual problems or issues that must be addressed by each student. This database was a solution for the Author to the problems of:

- Keeping track of what each student was working on
- Keeping track of problems/ issues that each student must address for his individual database
- Keeping track of custom reports assigned to each student
• Keeping track of individual student E-R diagrams (these were imported into a field that was an OLE object. Please see Figure 4. For a student E-R Example)

![Figure 4. Student stored E-R diagram](image)

Students in the course were also able to see the progress of all other student databases that were real world projects. Thus they experienced many issues beyond their own for their own database. Figure 5 shows a sample screen of some issues a student had with his database implementation. The Author entered these issues/problems throughout the semester.

![Figure 5. Sample screen of Student database project issues/problems](image)
REAL MEDICAL DATABASE EXAMPLE

As was stated earlier, this Author is a database consultant for a surgical practice in a large hospital. Many times during the semester issues arose on various databases the Author had written for the hospital. The students were often presented with “real” scenarios that the Author had dealt with. The students were given the database with simulated data and were to figure out how the Author addressed the specific issues. The Author then presented the “real” solutions to the issues or requirements of the “real” database.

Figure 6 shows a screen shot of the “real” surgical database the students worked with. Figure 7 is an example of a student assignment dealing with some of the “real” database issues the Author had dealt with.

1. The entire database is linked by Social Security Number (a string). But many patients are no longer providing this. The primary key will now be Athena number (also a string). This number is assigned to the patient and all will have one. No patient will be entered without one. The user will also still want to store the social security number if the patient provides it.

PROBLEM: how to convert to Athena number and also keep the Social Security Number

2. The patient’s employer is stored as a field in the claim table. This field is a text entry. The user now wants reports based on the entry in this field. PROBLEM: The data in the employer field is text, so the user has spelled the entries in various ways. What should be done to fix this so the employer name is entered in the database only once?

3. There are contacts associated with each claim. These contacts get information on the Worker’s Comp claim sent to them. These contacts can be Doctors, companies, individual nurses etc… Each contact belongs to a Company. The database also stores the patient’s employer as well as individuals who work for those employers. The User wants to add an employer to the contact company table and add various individuals from the employer employee list to the contacts list. PROBLEM: The user wants to just click on a button from the employer screen to add that employer to the contact company and add specific employees to the contacts list. But no duplicated must be added to either table

4. There are many reports listed in the database. All are run from the Report Section form. However, Report 28 is one that the user wishes to run from the patient claim screen 

PROBLEM: Report 28 (The Patient Cover Sheet) prompts the user for the Athena number and the claim number. But the user wants to just click on a button and print the patient’s information that is showing on the screen

5. There are three major contacts listed for each patient. (Contact 1, 2 and 3) There are three fields in the claim table that store the id for each of the three contacts. The contact id comes from the contact table. The user wants various reports listing the contact information for each patient. One of the reports counts the number of patients for each employer during a specific range of dates PROBLEM: It is relatively easy to do a report and group the patient by the contact number. But there are three contact numbers and the patient needs to be grouped by all three on one report
In the database course, the students also learned SQL (Structured Query Language). They used this with the Author’s sample databases as well as their own databases. For the most part, the students did use Access for their projects. However, some used Oracle or Sequel Server. But the students were required to use standard SQL at all times. Even though it is very easy to create tables in Access from the wizard screen they were required to write all data definition language (DDL) statements that would have been necessary for the creation of the tables in their database [19]. These included statements to create the tables, primary keys, foreign keys and various constraints. The students were assigned a number of reports specific to their data by the Author. Again, the Author’s Access database was essential to track what was assigned to each student. The students used SQL to create the underlying queries for the reports and then used the software features, such as the access report wizard, to create the actual reports from the SQL query statements. The students were also given a number of assignments using DDL statements and DML (data manipulation language statements). Again, this Author is of the mind that the actual database product used is not as important as the database topic to be mastered…in this case that being standard SQL.

NORMALIZATION

Students in a first database course must also be introduced to the process of normalization… (proving that the database design is a good one) Because of the “The rules for defining the physical database from the E-R diagram” referred to earlier in this paper, in most cases, the databases designs did not have too many problems. Students were to prove that all tables were in at least third normal form. A brief definition of first, second and third normal form is provided in Figure 8 [12].

<table>
<thead>
<tr>
<th>First Normal Form</th>
<th>Second Normal Form</th>
<th>Third normal Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>- The table has a primary key.</td>
<td>- The tables meet the criteria for first normal form.</td>
<td>- The tables meet the criteria for second normal form.</td>
</tr>
<tr>
<td>- No single attribute (column) has multiple values.</td>
<td>- If the primary key is a composite of attributes (contains multiple columns), the non key attributes (columns) must depend on the whole key.</td>
<td>- Each non-key attribute in a row does not depend on the entry in another key column.</td>
</tr>
<tr>
<td>- The non-key attributes (columns) depend on the primary key.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 8. Brief summary of Normal Forms

STUDENT RESULTS AND ATTITUDES

The Author has been teaching database similar to this for a number of years. In the Fall 2013 semester there were 35 students enrolled in the database course. Fifty percent of the students received an A for the course. An A was given if the student completed all of the assigned work on their database with only minor problems. This included user documentation in the form of a User’s Guide. They also had to successfully complete a number of smaller assignments as well as passing two exams. The students reported that they had a high interest in completing the individual database project because it was a “real” assignment for a “real” user. Further, they said that they could not “put down” the database and spent many additional hours getting it to look and run “just right”. Many of them also said that they added additional functionality beyond what was required and that they would take an advanced database course if given the opportunity. There is existing research that supports the theory that one course is not enough for the computer student to adequately understand database topics [21] [20]. It was the Author’s observation that the students are always more interested in a real project. By the end of the semester when the students presented their databases to the class, it was clear to see the pride on the face of those presenting. It was also just as rewarding for the Author to see students giving other fellow students accolades for their work. The Author will teach this course again in the upcoming Fall semester and an in-depth survey will be distributed to the students to assess their views of the course in a more complete manner.
CONCLUSIONS

This Author was able to bring many personal experiences as a database consultant into the classroom on a daily basis. The students benefitted from the discussions of real encounters with real users and real issues/problems. “Real” projects have always been a great way to get the student involved in what he is suppose to be learning, especially a project that the he has selected for himself [9].

Another issue that is of importance and not mentioned in this paper is the area of dishonesty. When using a format such as the one discussed in this paper, there is no possible chance of students copying from each other. In fact, often one student would help a fellow student with a problem he might be having in his database project. This gave the student experiences with issues he might not have encountered in his own database.

If teaching a course in this manner, the Author has a few recommendations:

- Be careful to limit the scope of each student’s project. The student may not realize that the project he wishes to do may be too much for someone new to database work. Also, the project must be completed in the Semester time limit. Perhaps deleting some of the functionality he may wish to include will be necessary. By the end of the semester, he should be able to incorporate additional functionality on his own.
- Make sure students interact with the user often. Any issues concerning the user must be communicated to the Instructor as soon as possible. It also might be a good idea for the instructor to briefly interact with the student’s user a few times during the semester.
- A detailed evaluation of the design is necessary in order for the student to complete the database successfully. This will require a significant amount of time on the instructor’s part. Because the students are new to database work, and the design is done first in the semester, they will need help to ensure that the design is a solid one.
- Create some way to organize what each student is working on. This Author wrote a database to store detailed information surrounding the student database projects. This database also allowed the students to keep track of what others in the class were working on. Thus, it gave them another outlet to examine “real world” projects.
- Find as many ways as possible of bringing real database experiences into the classroom. If the instructor is not doing actual database consulting work, bringing in a database programmer to talk to the class would be advised. Any way that the students can understand what really happens in the “real world” is advantageous.
- Make sure the students create sample data for database development. Also, do not include sensitive user information in the student’s database prototype. Make sure the user understands that the student’s work may be shown in class and therefore user’s private information must not be included.
- Make sure the students understand the grading rubric for the project. This could include an evaluation from the actual user. But is essential that students know how their grade is being calculated. This protects both the student and the faculty member.

This course will help the students become better database designers and implementers. They will come away with a wide range of practical knowledge that is not easily taught in a typical lecture setting. They will succeed in two arenas, technology and communication. It is becoming more and more important to bring “real” experiences into the classroom. The course described here is a step in the right direction [14].

REFERENCES