

INST 327

Section: 0101

Team Project Final Submission

December 14, 2020

Team 3

Shima Abdulla, Ilyas Dasti, Marwa Mansour, Parker Robinson, Antonio Saco

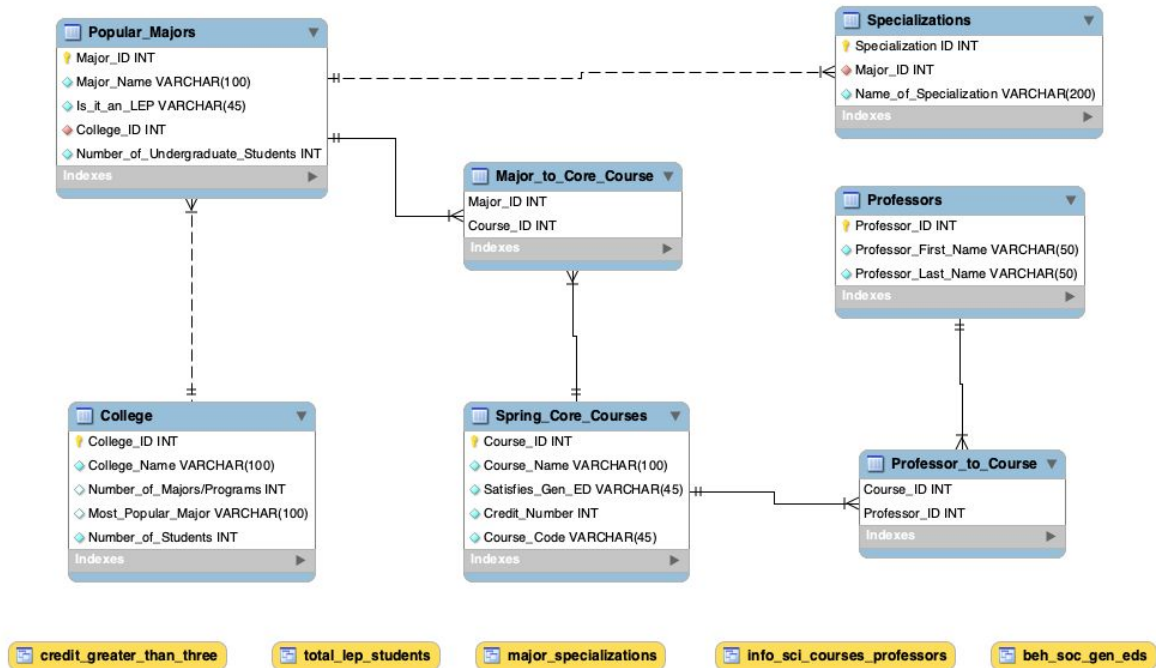
Introduction

As college students, it is often difficult to find the right resources to fulfill information needs, especially academic ones. Scrambling to find the right class is a struggle that we all know too well. The database should serve as a guiding force for new or incoming students as they learn to navigate the many programs and courses offered at UMCP. Our database provides a straightforward, simple solution for UMD college students who may need a resource to look up their most important courses. Many factors go into choosing a course including GEP requirements, number of credits, whether it satisfies Gen Ed requirements, and more. To simplify this decision process we created a database in which students are able to specify their course requirements and determine the appropriate course to register for. As students ourselves, creating this database is important as it solves a common issue that many new students like our former selves may have had in terms of deciding what class to take. Our final database has the following tables: (Popular_Majors, Specializations, College, Core_Courses, Major_to_Core_Courses, Professors, and Course_to_Professors).

Database Ethics

Database ethics played a small but notable role in the design of our project. One problem is that students often pick classes based on what professor is teaching, but bias is avoided by including all the professors rather than the most popular ones. The overall goal for the project was to be accessible and usable by all; The project actively avoids large problems, such as discrimination. Any student that uses the database should feel welcome and not have any trouble at all interacting with the provided data.

Logical Design



Physical Database

Our database consists of seven tables. Our main table is the College table which represents the different colleges or schools housed under the University of Maryland College Park. Although the university technically has 12 different colleges in terms of major related ones, we also included the University of Maryland Shady Grove as a separate entry as students can receive permission to enroll alongside their UMCP courses. Additionally the Graduate school was included to account for undergraduates who enroll in graduate level courses, Lastly the office of undergraduate studies was also included as it is listed by the university as a separate school on their site.

The College table has a one-to-many relationship with our Popular_Majors table. The Popular_Majors table depicts the fifteen largest undergraduate majors in terms of size according to enrollment data provided by the university of students who have declared the major as either a secondary or primary major for Fall 2019.

The spring core courses table represents projected core courses of the fifteen popular majors that as of December 11th, 2020 are expected to be taught during the Spring 2020

semester. To fill this dataset we first determined the core courses for each major as listed by their department then physically removed the courses that are listed on testudo prior to importing the data. We determined that one core course can be required by multiple majors, such as a CMSC course being required by both compsci and computer engineering majors, and a major may have multiple core courses thus there was a many to many relationship between Spring_Core_Courses and Popular_Courses leading to the linking table Major_to_Core_Course.

Likewise the Professors table represented data on current planned instructors for Spring 2020. To connect this table to the Spring Core Courses table a linking table, namely the Professor_to_Course table, was created as there is a many to many relationship between courses and professors as a course can be taught by multiple professors and a professor can teach multiple courses.

Lastly our specializations table represented the different specializations or concentrations offered by the 15 most popular majors. It is important to note that although two majors may have an identically named specialization, such as data science under information science and computer science, the two specializations are different. Overall since one major can have multiple specializations, a one to many relationship was created between the Popular_Majors table and the Specializations table.

Sample Data

Below is the data in our Popular_Majors table.

Major_ID	Major_Name	Is_it_an_LEP	Number_of_Undergraduate_Students	School_ID
1	Computer Science	Yes	2727	5
2	Engineering: Mechanical	Yes	1244	7
3	Information Science	No	1216	9
4	Public Health Science	No	1119	10
5	Finance	Yes	1117	15
6	Criminology and Criminal Justice	Yes	1022	4
7	Communication	Yes	840	3
8	Biological Sciences	Yes	1777	5
9	Kinesiology	No	732	10
10	Psychology	Yes	710	4
11	Economics	No	705	4
12	Government and Politics	Yes	622	4
13	Marketing	Yes	617	15
14	Engineering: Aerospace	Yes	617	7
15	Engineering: Electrical & Computer Engineering	Yes	509	7

When creating our datasets we attempted to utilize as much real data as possible through sources such as testudo, department websites, and other university reports. This stated, although the data for our dataset could be found online, an already created dataset with all the information we wanted was not available, thus we utilized google sheets to compile and insert data found online.

Views / Queries

View Name	Req. A	Req. B	Req. C	Req. D	Req. E
Total_LEP_Courses		X	X		
Beh_Soc_Gen_Eds	X	X		X	
Info_Sci_Courses_Professors	X	X		X	X
Credits_Greater_Than_Three	X	X	X	X	X
Major_Specializations	X				

Query 1: Creates view of the total number of undergraduate students in a major that has Limited enrollment

Query 2: Creates view of the core courses that satisfy a general education requirement and belong to a major that falls under the College of Behavioral and Social Sciences

Query 3: Creates view of the names of professors who teach an Information Science core course, along with the course name and the course code

Query 4: Creates view with the name of the major if it has courses offered at more than three credits and the total number of these courses

Query 5: Creates view of the different specializations per Major

Changes from original design

We have made a few changes since our initial Project Proposal. In our Proposal, our plan included listing courses for multiple semesters. However, upon diving deeper into our project we realized that the amount of data would be too large for the scope of our project, so we limited our database to the semester of Spring 2021. Furthermore, deciding upon representing core courses offered only offered for the Spring 2020 term made sense, as this would give students a more

accurate representation of classes they should enroll in during future/upcoming terms (as opposed to say representing Fall 2020 courses which wouldn't be as useful since not all those classes will be offered in the next/later semester). We also eliminated a couple of tables we planned on having (LEPs, Pre-Professional courses, and Minors), because we realized that having them in our database would either be too complex for the scope of the project or create redundant data. Instead we included a professors table. This table seemed necessary as currently many students pick and choose their courses based off of the professor who teaches them that semester. Furthermore initially we neglected the need for linking tables but after further examination of our data set there were many-to many relationships in which a linking table was needed resulting in the creation of two linking tables.

Lessons Learned

During our group's time working on this project, we have learned some important lessons both with regards to SQL and team work. In terms of SQL most of the lesson we learned came as a result of making mistakes. One of the most recent lessons learned comes from the titles in SQL. Usually when writing titles proper grammar is utilized including having spaces between words and punctuation marks including question marks where needed. This stated initially our columns and table titles included spaces in between each word and question marks where necessary. When we first created the tables it didn't seem like a problem. However when we started writing our queries the spaces and question mark made selecting the titles tedious as we could no longer just drag over the name but rather had to use quotes and be more specific in a way for SQL to recognize the columns. This stated, we tried to work with the spaces, but as we continued it was getting more and more complicated to have them stay. Our group decided it would be much easier and more time efficient to just edit the table titles. Ultimately, we altered our tables so that between each word had an underscore rather than a space and removed the question mark. This simplified the query writing process. Looking back at it this method of labeling is utilized in many of our other tech classes and is rather efficient. By simplifying the labeling process the debugging process is also made easier. It's the little mistakes like these that our group faced but we all came together and thought of a solution.

Perhaps the largest problem we faced whilst completing this project appeared in the beginning when creating our ERD and more specifically deciding upon our tables and how to

link all of our tables together. Initially we wanted to add in tables about pre-professional programs and LEP majors separately. That stated we later determined through examining the potential data that would be included in such a table that the data would be redundant. After some guidance from our team advisor and our section TA we were able to decide upon the final tables. Overall we learned that sometimes having more tables isn't better. If the data is redundant, perhaps adding another column, such as a boolean for if a major is an LEP, would include the information needed and simplify the process. This also made understanding the linking better as having unique data is easier to link than redundant.

Lastly, one of our most important lessons that we learned, how to cooperate as a team. Whenever our group had a problem we either communicated over chat or if it was more of a major problem we would set up a meeting. We found this form of communication and teamwork was the most efficient way to deal with problems.

Potential Future Work

When creating a database on courses the potential is endless. Currently the data imported into our database is based on the information that was available on testudo up until the 11th of December. This stated, not all the professors are listed now. In fact, instructors are still listed as TBA on testudo thus in those areas we had to neglect the professor. In the future once that is updated we could add that in. Furthermore we could add a fall core courses table to represent courses that were taught this Fall, so that students could start to get a general image of what may be taught. This method could also be used for spring and winter semesters. Furthermore currently we only offer data on the fifteen largest undergraduate programs but we could expand the data to include that of all other majors. This could benefit all students at UMD, not just a subset of them like our current database does. If we wanted to expand this database even more we could use data from other schools and have a universities table. This could be helpful for all students applying for college and trying to decide what school is best for them. This would benefit them because they can search for class information and cross check them against other schools. For example, they can filter out schools that have less than 20 classes in the computer science major. Overall the potential is endless for this database.

Citations

“The Colleges and Schools.” *2020-2021 Catalog Undergraduate*, University of Maryland, academiccatalog.umd.edu/undergraduate/colleges-schools/.

- Utilized for the dataset as information on different majors, major core courses, major specializations, and the different colleges was extracted through navigating this cite

“Limited Enrollment Programs.” *University of Maryland*, University of Maryland, admissions.umd.edu/explore/majors/limited-enrollment-programs.

- Utilized in dataset to determine if a major is an LEP

“Number of Registered Majors.” *Reports.umd.edu*, Office of Institutional Research, Planning & Assessment, 2020, reports.umd.edu/reportHolder.html.

- Used in dataset to determine the number of students per major, number of students per program, most popular majors, and most most popular majors

“Schedule of Classes - Spring 2021.” *Schedule of Classes*, University of Maryland , 2020, app.testudo.umd.edu/soc/.

- Used in dataset to determine courses offered in Spring 2020 and the professors teaching the courses