INST 327: Database Design and Modeling
Section: 0201
Final Report
May 14, 2021
Team 7
Sophia Chen, Mingqian Huang,
Kevin Hung, Noble Nwokoma,
Maham Sohail

Team 7 Final Report

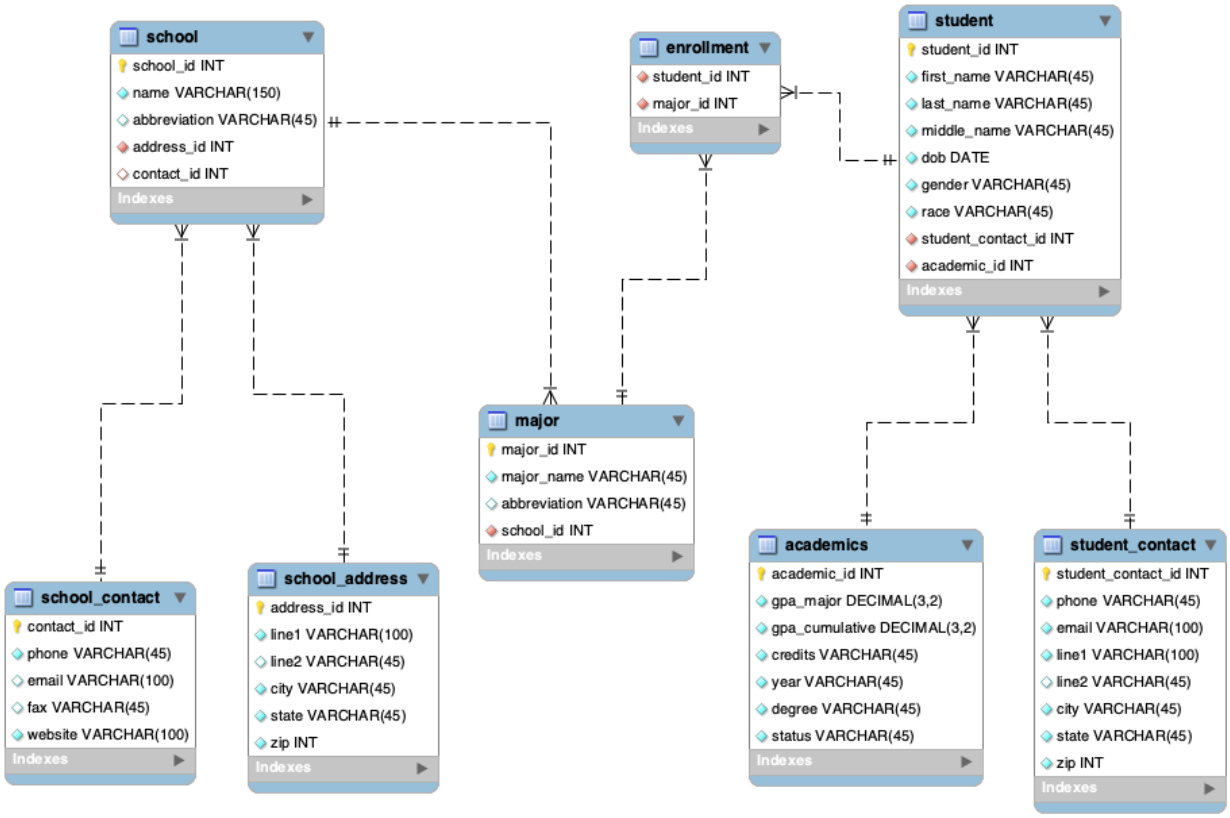
Introduction

Our database is a collection of all the departments, schools, and majors at the University of Maryland along with a collection of student names, addresses, phone numbers, GPA, and enrollments. The goal of this database is to create a system where a user can find information about different departments at the University of Maryland along with student information such as cumulative and major GPA, the major the student is currently in, and how many credits the student currently has. This database is useful for information retrieval and access to information regarding student progress, and majors and departments offered at the university.

Database Description

Logical Design

The ERD model contains eight different tables; school table contains the different departments at the University of Maryland, the school contact table contains the different forms of contact information for each department, school address lists the addresses for each department, the major table lists all the available majors at the University of Maryland along with a corresponding id (major_id). The student table contains the names of the students, the date of birth, gender, and race. The academics table lists the students gpa, credits, year, degree, and status. student_ contact lists the contact information for each student. The enrollment table contains two foreign keys and is used as a linking table for the student table and major table.



Physical Database

The goal of this database is to create an all-inclusive guide for University of Maryland students. In regards to assisting students in learning more about different majors and departments at the University of Maryland, the database also contains student information which could be potentially helpful for faculty and staff. After the ERD model was completed, forward engineering which is a process of converting ERD models to schemas was implemented to create our database. We attempted to make a well-rounded database that provided students with all the information that they might need about the college. There were many issues that we encountered when trying to import data into our schemas but those issues were minor and we fixed them quickly. Overall, the physical Database (backup database) is successful when uploaded by all members.

Sample Data

Data used to collect information for the different departments are found on the official University of Maryland website: https://www.umd.edu/, other websites were also used to collect the University of Maryland school departments information:

- https://www.umd.edu/colleges-and-schools This website provides the user with a list of colleges and schools at the University of Maryland. Each one of the Colleges and Schools is linked to its official website, which provides more detailed information about each College/School.
- https://www.admissions.umd.edu/explore/majors This website contains a list of majors of each college/school at the University of Maryland. Our team used this information to collect data for the major table.

The students used in the database are hypothetical or made up. We have not collected any information from real students. This data was primarily used for our students, academics, and student_contacts tables. Creating a database with hypothetical students helped the team save the team a lot of time. Additionally, we randomly generated values for much of our numerical data, including GPA, credits, and id numbers. This is an example of our academics table database for students.

academic_id	gpa_major	gpa_cumulative	credits	year	degree	status
114526	3.30	1.27	69	Year 3	B.S	full
186548	3.82	2.76	86	Year 3	B.S	part
212561	3.74	2.19	37	Year 1	B.A	part
248403	2.67	3.64	99	Year 3	B.A	full
257507	2.82	2.56	102	Year 4	B.A	part
262771	2.20	3.12	114	Year 4	B.S	full
422212	3.76	2.15	109	Year 4	B.A	full
531446	2.78	1.57	54	Year 2	B.A	part
561335	3.97	3.01	73	Year 2	B.S	full
578169	3.10	2.77	39	Year 1	B.S	full
585848	3.22	2.67	32	Year 1	B.A	full
613609	3.25	2.73	91	Year 3	B.A	part
623878	2.03	2.49	113	Year 4	B.S	full
636338	2.12	2.03	74	Year 2	B.S	part
647689	2.90	3.83	58	Year 2	B.S	part
729491	3.03	1.43	38	Year 1	B.S	part
736422	2.73	3.17	101	Year 4	B.S	full
770332	2.86	2.55	28	Year 1	B.A	full

Views / Queries

Our views and queries focused on five primary questions. These queries are helpful in better understanding the functionality and use for these types of databases as well as helping the user understand the importance for these types of databases. The five questions that we answered are (question 1) How many students are majoring in history? For this question we used a VIEW, JOIN and WHERE clause to capture all the student names who majored in history, the second question is how many students have a GPA of 3.0 and higher? This question used VIEW, JOIN and WHERE clauses where we captured all the students who have a gpa of 3.0. Question 3 asks for a list of students who have a B.S degree. For this query we used a VIEW, JOIN, SUBQUERY and WHERE clauses. Finally the fourth question asks what is the average cumulative GPA of each school in UMD? This question implemented the VIEW, JOIN, GROUPBY and aggregation. Finally question five asked for the number of students Male/Female in each school for which we used the JOIN clause and an aggregation method.

Query Name	Req. A	Req. B	Req. C	Req. D	Req. E
history_major_student	X	X		X	
high_gpa_student	X	X			
bs_degrees	X	X		X	X
avg_school_gpa	X		X	X	
gender_count	X		X	X	

Discuss whether the presentation on Database Ethics impacted your project

Data ethics was not much of a factor when creating our database. The database did not require any real, or factual information. All the information is either false and created or the information such as the school departments is already information used for public consumption. There was no issue regarding privacy of any persons and we did not collect any personal information from anyone. When creating the table for students, we modeled our data based on the demographic reports provided by the school. Since we used those categories, the "gender" column only includes "female" and "male" and not people who are non-binary. Furthermore, our "race" column was also limited; we didn't include "mixed race" in our data. Though we used existing categories for our data, they don't represent people who don't fit into those categories.

Changes Made From Original Design

At the beginning of the semester there was a different concept planned for our database and queries, the plan was that we would collect and create a database for all schools in Maryland. As the semester progressed the team decided to create a database that would only look at the departments and students at the University of Maryland. There were also many modifications made in perfecting the ERD model and schemas. For example, a table named pathways was implemented to show the different pathways each student was in within their individual major but problems arose when we couldn't find a table to link so we decided to delete the pathways table and create a new table named enrollments which were used to link our tables as required for the subqueries.

Lessons Learned

When creating the database the team had to make many edits and modifications to excel files and schemas. Also when creating our Views and Queries the team was not sure about how to implement all the requirements but, we were successful. Also, the team communicated effectively and learned how to work together to complete assignments.

Potential Future Work

Potential changes and modifications which can further perfect our project is that the database could use more tables to create a more cohesive dataset.The dataset could potentially be changed

in a way to help solve multiple problems. Another concept we forgot to include is to include students who are mixed-race individuals and including people who are non-binary.