Data 331

Fall 2023

Team Project

**Project Overview**

The purpose of the team project is to give students the opportunity to collaborate in a team setting to deliver a solution to a data problem that is representative of problems commonly found in a professional environment. Working in teams, students will be given a set of requirements that define what their solution should accomplish. Teams will be expected to follow the system analysis and design process covered in the course to come up with a solution.

Students will be placed into teams of 5, and all teams will work on the same problem. There is no single correct solution, and teams will have the autonomy to reason and come up with their own solution approach. The primary criteria for success on the project will be following the system analysis and design process, submitting the required project documentation, collaborating effectively within the team, and effectively presenting the solution to the class.

There will be two significant milestones for the project: first, submitting documentation of the requirements analysis and solution design, and second, submitting the completed solution and giving a presentation on the work that was done. Details on the required documentation and submission dates are described below in the Project Deliverables section.

**Project Problem and Requirements**

It’s the summer of 2023, and a small liberal arts college has recently made the decision to migrate from their old Student Information Management System to a new, more modern one that will be custom built. All data for the 2023-24 academic year will be entered into the new system, but data from the old system will need to be migrated into the new system database that will be designed to meet the requirements of the administration team. As part of this migration project, your team will be responsible for the following:

1. Analyzing the data that is extracted from the old system.
2. Designing a new database that will support the given requirements of the administration team.
3. Transforming and loading the data from the old system into the newly designed database
4. Validating that all data from the legacy system was migrated successfully.
5. Producing analysis that meets the business intelligence requirements of the administration team.
6. Producing required documentation for your analysis and design.

Your team has the following technology at your disposal:

1. Microsoft Excel – Data from the old system will be provided in Excel, which can also be used to analyze the old data. Excel is also a good choice for some forms of project documentation.
2. Microsoft Azure SQL – The new system will be built using Microsoft Azure SQL. Your team will have full administrative access to an Azure SQL database that you can use to design, build, and store data.
3. Microsoft SQL Server Management Studio (SSMS) – SSMS can be used to connect to databases in Azure SQL Server, as well as data in many other sources, including Excel. Once connected, SSMS is used to query, manage, and analyze data. It can also be used to execute basic data migration tasks.
4. R and R Studio – R and R studio can be used to connect to data and perform more advanced analysis, as well as programmatic data transformation and migration,
5. Microsoft Word and PowerPoint – Word is used for delivery of project documentation, and PowerPoint will be used for the project presentation.
6. Draw.io - Draw.io can be accessed as part of your Google account and is a useful tool for creating diagrams, specifically database design diagrams.

The following requirements have been defined:

**Student Data Requirements**

* Student data should include:
  + An ID (required, unique)
  + First Name (required)
  + Middle Name
  + Last Name (required)
  + Email Address (required, unique)
  + Major (required)
  + Class Level (Freshman, Sophomore, Junior, Senior) (required)
  + Status (Active, Inactive, Alumni)(required)
  + First Enrolled Date (required)
  + The student’s faculty Advisor (required)
  + Address (Country, Region/State, City, Street Address, ZIP/Postal Code)
  + Birth Date (required)
* Student records include all records for all students that have been enrolled at any point. Active students are currently enrolled in good standing, inactive students are not currently enrolled but were at one point, and graduates have completed a degree.
* Majors must be a currently active major.
* If a student has not declared a major, an undeclared major should be assigned.
* Faculty advisor must be a currently active advisor.
* For the sake of simplicity, Class Level corresponds to the number of years a student has been enrolled in the college (first year = freshman, second year = sophomore, etc.)

**Course Data Requirements**

* Every course has a single course record.
* Every course record name is unique.
* The combination of Course Number and Active status is Unique (i.e. there can only be one active course with a specific course number at a time).
* Course data should include:
  + An ID (required)
  + The Course Number (required)
  + The Course Name (required)
  + Course Description (required)
  + The Department the course belongs to (required)
  + Status (Active, Inactive) (required)
* A course offering is created every time a course is offered. Course offering data should include:
  + An ID (required)
  + The course being offered (required)
  + The instructor teaching the course (required)
  + The academic year of the course offering (required)
  + The term of the course offering (required) (Fall, Winter, Spring)
  + The section of the course (required)
  + Credit hours
* The combination of course, academic year, term and section is unique.

**Enrollment Data Requirements**

* Students enroll in specific course offerings.
* A student can only be enrolled in a course offering one time.
* There should be a single, unique record for each enrollment.
* Enrollment data should include:
  + Student (required)
  + Course Offering (required)
  + Status (Enrolled, Withdrawn, Complete) (required)
* Enrollment changes from Enrolled to Complete when an academic term ends for all students that completed the course (regardless of final grade)

**Grade Data Requirements**

* For students that complete an enrolled course, the final grade is recorded.
* There should be a single, unique record for each student grade.
* Grade data includes:
  + The student (required)
  + The course offering enrollment(required)
  + The final letter grade (required)(A,B,C,D,F,I). Plus and minus values are allowed for all grades except I.
  + Total grade points (required)
  + Status (Final, In Progress) (required)
* A student can have a grade of I while the grade is in progress. Once a grade status is Final, the grade must be A,B,C,D, or F, with Plus and Minus variants.

**Faculty Data Requirements**

* Faculty data should include historical faculty data.
* Faculty data should include:
  + ID (required, unique)
  + First Name (required)
  + Middle Name
  + Last Name (required)
  + Email Address (required, unique)
  + Is an Advisor (required)
  + Is a Department Chair (required)
  + Is an Instructor (required)
  + Is Tenured (required)
  + Is Full Time (required)
  + Status (required)(Active, Terminated, Emeritus)
  + Hire Date (required)
  + Termination Date
  + Salary (required)
* Only Full Time faculty can be a Chair, Advisor or Tenured.
* Only active Faculty that are instructors can be assigned to a course offering in the current term. After the current term is over, the faculty status can change to a status other than active.
* Only faculty with a status of Inactive should have a termination date.
* Faculty with an inactive status cannot be an advisor, chair, instructor, full time or tenured.

**Department Data Requirements**

* A department represents an area of academic focus in the college, and each department contains one or more Major areas of study.
* Department data includes:
  + ID (required, unique)
  + Name (required)
  + Chair ID (required)
  + College (required) (Business, Liberal Arts and Sciences, Engineering)
* The chair of a department must be an active Faculty member that is designated as a chair.

**Major Data Requirements**

* Major data includes:
  + An ID (required, unique)
  + The Department the major is associated with (required)
  + The Major Name (required)
  + Whether or not the major is active (required)
* Currently enrolled/active students can only be associated with an Active major.
* A general “undeclared” major should exist for students that have not yet declared a major.

**Project Deliverables**

**Mid-Term Deliverables (Due Monday, October 30, 2023)**

Data Analysis Documentation

* Excel Data Dictionary (Template Provided)
* Data Requirements Analysis (Template Provided)

Database Design Diagram

* E-R Diagram of Proposed SQL Database (Use Draw.io or a tool of your choice)

Database Data Dictionary

* SQL Data Dictionary (Template Provided)

**Final Exam Deliverables (Due Wednesday, December 13, 2023)**

Data Flow/Mapping Documentation

* Template Provided

Azure SQL Database

* Each team will create an Azure SQL database and populate it with data from the provided Excel spreadsheet.

Team Programming Assignment

* Complete the SQL Scripts described in the Team Project Programming assignment.