E-Learning Platform



**Group - 4 - EduTech Innovators**

Derleen Saini

Kaveri Kesarwani

Ritika Verma

Tanya Jain

# 

# Table of Contents

# Business Description …………………………………………………………………..2-3 1.1 Overview: Nature and Scope …………………………………………………………2 1.2.Purpose and Benefits: Optimizing Educational Delivery …………………………….2

1.3. Anticipated Functionalities and Features …………………………………………….3

# Conceptual Design (Enhanced Entity Relationship Diagram) ........................................4-7

# 2.1. Data Entities and Attributes ………………………………………………………….5

2.2. Entity-Relationships ………………………………………………………………….7

# Logical Design (Relational Model) ....................................................................................8

# 3.1. Normalization (Third Normal Form)............................................................................8

# Implementation - SQL Database ……………………………………………………...9-12

# 4.1. Without Using Joins ………………………………………………………………...10

4.2. Using Joins ………………………………………………………………………….10

4.3. Using Sub-Queries ………………………………………………………………….11

4.4. Using Views ………………………………………………………………………...11

4.5. Using Stored Procedure …………………………………………………………….12

4.6. Using Trigger ……………………………………………………………………….12

# Administrative Functions ……………………………………………………………13-15

# 2.1. Potential Issues ………………………………………………………………….13-14

2.2. Measures and Features Incorporated ……………………………………………….14

# Conclusion ………………………………………………………………………………15

# 

# 

# 

# 

# 1. Business Description

Founded in 2012 by Stanford University's Andrew Ng and Daphne Koller, Coursera Inc has become a leading provider of Massive Open Online Courses (MOOCs) in the U.S., aiming to democratize education by partnering with over 300 universities and corporations worldwide. By 2024, Coursera expanded its offerings to include around 7,000 courses, playing a crucial role in connecting learners with top-tier instructors and educational content. This effort supports Coursera's mission to improve global access to education, promoting lifelong learning and professional growth, and highlights the role of digital platforms in breaking down traditional educational barriers, offering scalable, high-quality educational solutions globally.

# 1.1 Overview: Nature and Scope

Coursera is an e-learning platform that aims to offer a diverse range of courses across various disciplines to a global audience, providing accessible and flexible learning opportunities. The e-learning platform serves as a virtual classroom, connecting instructors with learners worldwide. It caters to individuals seeking to expand their knowledge, acquire new skills, or enhance existing ones. The scope encompasses a wide array of subjects, including but not limited to technology, business, arts, sciences, and humanities. The database will enable course enrollment, progress tracking, instructor management, forum interaction, analytics etc for informed decision-making.

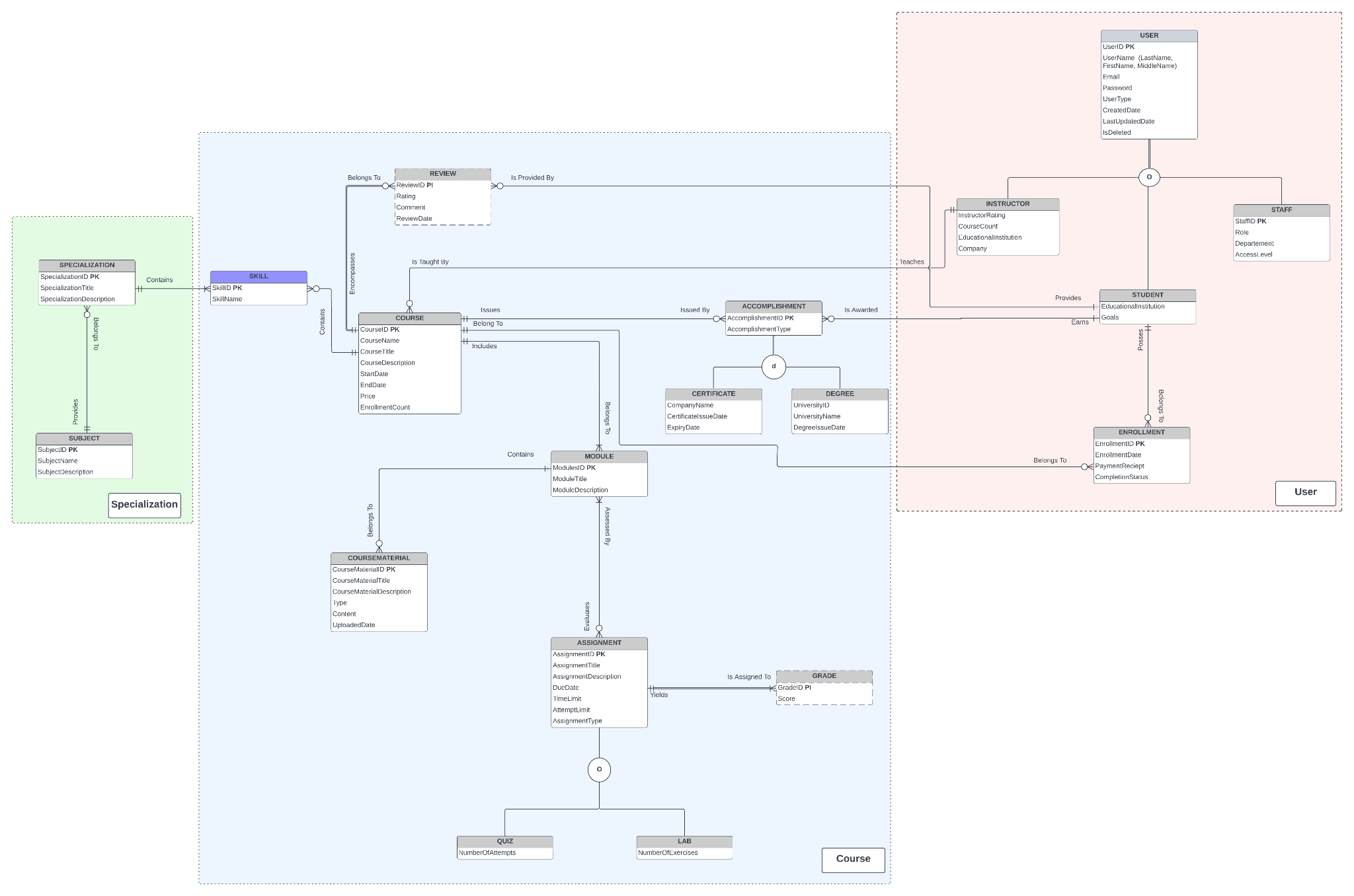
# 1.2. Purpose and Benefits: Optimizing Educational Delivery

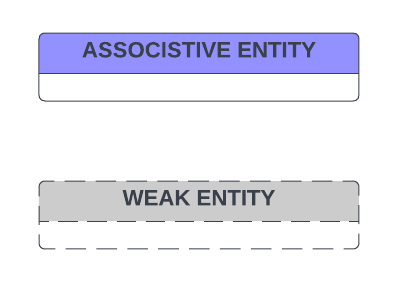
1. **Operational Efficiency**: The database's streamlined structure ensures efficient data management, enabling quick delivery of educational content and maintaining high service quality.
2. **Scalability**: Designed to handle growth, the database supports an expanding user base and course catalog, ensuring Coursera remains responsive and accessible without compromising performance.
3. **Data-Driven Insights**: Aggregating and analyzing data, the database informs course improvements and strategic decisions, tailoring offerings to meet learner needs and enhancing platform effectiveness.
4. **Personalized User Experience**: Utilizing stored data, Coursera offers personalized course recommendations, enhancing engagement by aligning learning paths with individual preferences.
5. **Professional and Academic Advancement**: Managing certifications and degrees, the database provides learners with credentials that support their career and academics.
6. **Robust Security Framework**: Incorporating encryption, access controls, and authentication, alongside continuous security audits and anomaly detection, the database safeguards user information, ensuring a secure learning environment.
7. **Resilience and Recovery**: With comprehensive backup and disaster recovery plans, Coursera protects users' educational investments and guarantees swift service restoration in case of disruptions.

# 1.3. Anticipated Functionalities and Features

1. **User Data Management**: Securely stores and manages millions of user profiles, safeguarding personal information and preferences while ensuring easy access.
2. **Course Catalog Management**: Dynamically accommodates the growing range of courses, facilitating updates and categorization, which aids user navigation of courses.
3. **Enrollment and Progress Tracking**: Essential for recording enrollments and tracking learner progress, including quizzes and assignments, to offer personalized learning paths.
4. **Certification and Degree Management**: Supports the issuance and verification of credentials, crucial for learners to leverage their achievements professionally.
5. **Advanced Analytics and Personalization**: Utilizes data to tailor the learning experience, enhancing user engagement through personalized course recommendations.
6. **Content Management and Delivery**: Ensures course materials are up-to-date and adaptable to individual learning styles, supports real-time updates and adaptive learning.
7. **Integration and Accessibility**: Facilitates integrations with third-party tools and ensures all content is accessible, meeting diverse learner needs.

# 2. Conceptual Design (Enhanced Entity Relationship Diagram)





# 

# 

# 2.1. Data Entities and Attributes

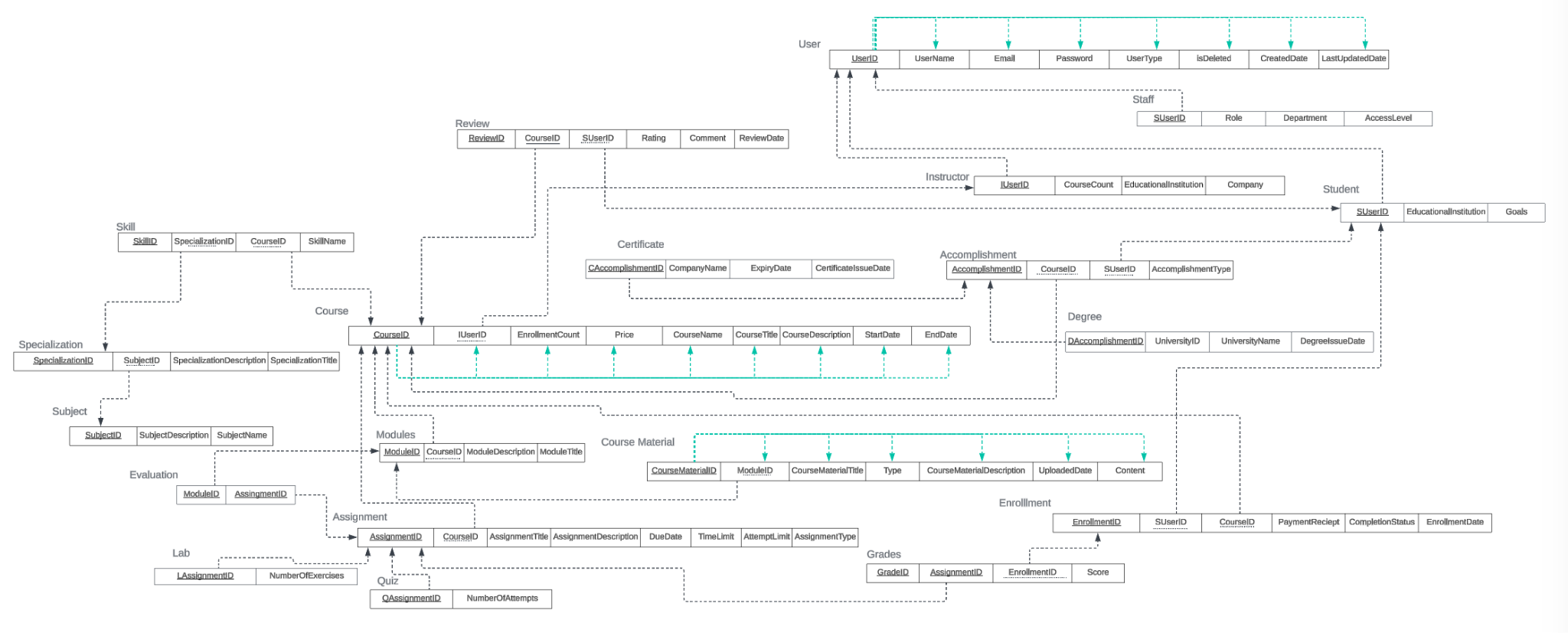
| Entity | Entity Definition | Attributes / Attributes Definition |
| --- | --- | --- |
| **USER** | Anyone who has registered on the platform with a unique username and email. | UserID, Username, Password, Email, UserType, CreatedDate, LastUpdatedDate, IsDeleted |
| **STUDENT** | Registered user who enrolls in courses to gain knowledge and skills. | Inherits basic USER attributes.  EducationalInstitution, Goals |
| **STAFF** | Users who are employees of the platform, involved in administrative and operational duties. | Inherits basic USER attributes.  Role, Department, AccessLevel |
| **INSTRUCTOR** | User with expertise who creates and facilitates courses. | Inherits basic USER attributes.  InstructorRating, CourseCount, EducationalInstitution, Company |
| **COURSE** | Structured educational program available on the platform, not to be confused with individual classes or materials. | CourseID, CourseName, CourseDescription, StartDate, EndDate, Price,EnrollmentCount |
| **MODULE** | Component of a course, representing a subset of the course's content. | ModuleID, ModuleTitle, ModuleDescription |
| **ASSIGNMENT** | Tasks set by the instructor within a course can be quizzes, exams, or both. | AssignmentID, AssignmentTitle, AssignmentDescription, DueDate, TimeLimit, AssignmentType |
| **QUIZ** | Set of questions aimed at assessing student understanding within a course | Inherits basic ASSIGNMENT attributes, NumberOfAttempts |
| **LAB** | Practical assignment aimed at assessing student understanding within a course distinct from assignments. | Inherits basic ASSIGNMENT attributes, NumberOfExcercises |
| **GRADE** | Reflects the evaluation of a student's performance on an assignment or quiz. | GradeID, Score |
| **ENROLLMENT** | An enrollment record is created when a student registers for a course and can be deleted if the student withdraws or the course ends. | EnrollmentID, EnrollmentDate, PaymentReceipt, CompletionStatus |
| **REVIEW** | Feedback provided by students on a course, separate from internal course assessments or discussions. | ReviewID, Rating, Comment |
| **COURSE MATERIAL** | Encompasses the educational content provided in a course, not including external resources. | MaterialID, CourseMaterialTitle, Description, Type, Content, UploadedDate |
| **SPECIALIZATION** | Represents a series of interconnected courses designed to develop expertise in a specific area. | SpecializationID, SpecializationTitle, SpecializationDescription |
| **SUBJECT** | Area of study within or across specializations, not to be confused with specific courses or skills. | SubjectID, SubjectName, SubjectDescription |
| **SKILL** | Competencies that subjects or courses aim to impart. | SkillID, SkillName |
| **ACCOMPLISHMENT** | User's achievements on the platform, such as course completion. | AccomplishmentID, AccomplishmentType |
| **CERTIFICATE** | Formal certificate of completion and expertise awarded. | Inherits basic ACCOMPLISHMENT attributes., CompanyName, ExpiryDate |
| **DEGREE** | Formal degree of completion and expertise awarded. | Inherits basic ACCOMPLISHMENT attributes, UniversityName |

2.2. Entity-Relationships

The relationships between entities are clearly defined to ensure the integrity and usefulness of the stored data. Few relationships between entities are defined below.

1. STUDENT "Posses" zero or many ENROLLMENT. But ENROLLMENT “Belongs To” only 1 STUDENT.
2. STUDENT “Earns” zero or many ACCOMPLISHMENT, and an ACCOMPLISHMENT is “Awarded To” to only one STUDENT.
3. COURSE “Belong To” at least one or more MODULE, but MODULE “Belongs To” only one COURSE.
4. MODULE can be “Assessed By” zero or more ASSIGNMENT, and one ASSIGNMENT “Evaluates” at least one MODULE.
5. Each SPECIALIZATION “contains” multiple SKILL', and each SKILL can be part of multiple COURSE.
6. GRADE is a weak entity dependent on ASSIGNMENT.
7. REVIEW is a weak entity dependent on COURSE.
8. STAFF, STUDENT and INSTRUCTOR are subtypes of USER that have total specialization and overlapping constraints
9. LAB and QUIZ are subtypes of ASSIGNMENT relation that have partial specialization and overlapping constraints.
10. CERTIFICATE and DEGREE are subtypes of ACCOMPLISHMENT relation that have partial specialization and disjoint constraints.

3. Logical Design (Relational Model)



* Apart from the entities listed above, an Evaluation Relation is created to justify a many to many relationship between Module and Assignment
* All the relations are in the third normal form as depicted with the Course, User and Course Material relations through green lines.
* Solid underlines for primary keys
* Dotted underlines for Foreign Keys

3.1 Normalization ( Third Normal Form )

* **Checked for Multi-Valued Attributes:** Ensured atomicity,where each attribute held indivisible values.
* **Removed of Partial Dependencies:** The "user" relation was decomposed into "student," "staff," and "instructor" to eliminate partial dependencies and ensure each non-key attribute is fully functionally dependent on the primary key to bring to the 2NF.
* **Elimination of Transitive Dependencies:**Reviewed relations for transitive dependencies.Ensured attributes such as "skill" and "specialization" were normalized,resolving any transitive dependencies by restructuring them into separate relations instead of a single one to bring the relation to 3rd Normal Form.
* **Verification of Full Functional Dependencies:**Confirmed that every non-key attribute is fully functionally dependent on the entire primary key.This step ensures robust data integrity, with each attribute uniquely determined by the primary key.
* **Redundancy Minimization:**Checked for and eliminated redundant data across relations to enhance data consistency and storage efficiency within the relational model.

4. Implementation - SQL Database

All transactions will be handled by the application and correspond to actions taken by users within the application. We do not expect or anticipate that the database admin will need to intervene with manual transactions. However, they will be the only other user able to transact with the database. The application will insert the attribute values as submitted by the user via the application.Through leveraging the database, Coursera can deploy a range of business functions aimed at extracting actionable insights from the available data. These functionalities are poised to enhance user experience and drive engagement on the platform. Key among these functions is:

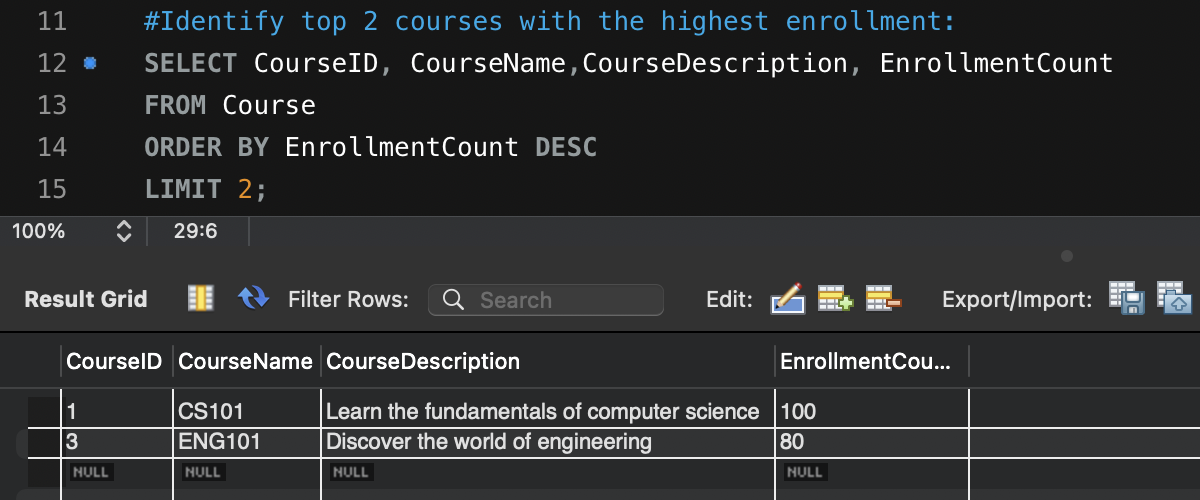
* Utilize data on user preferences, past enrollments, and completed courses to develop a personalized course recommendation engine.
* Analyze user behavior patterns to suggest relevant courses based on their interests, skill level, and career goals.

We have used the INSERT statement to feed sample values into the table for our purposes. Please see the example transactions, as they would be entered by the application during normal operations, below:

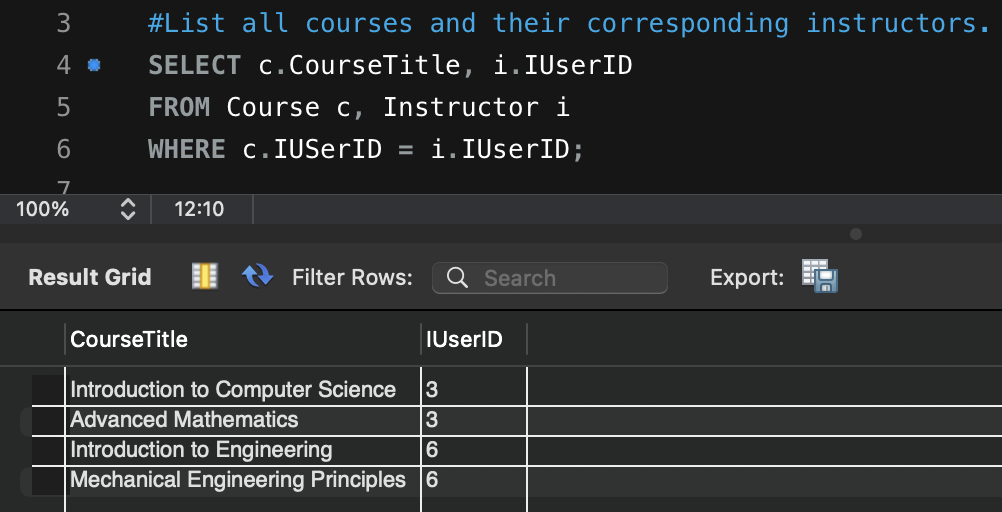
**Note**-The sql files of DDL, DML & Business Functionalities are attached with the submission for reference.

4.1. Without Using Joins

1. To identify courses with highest enrollments

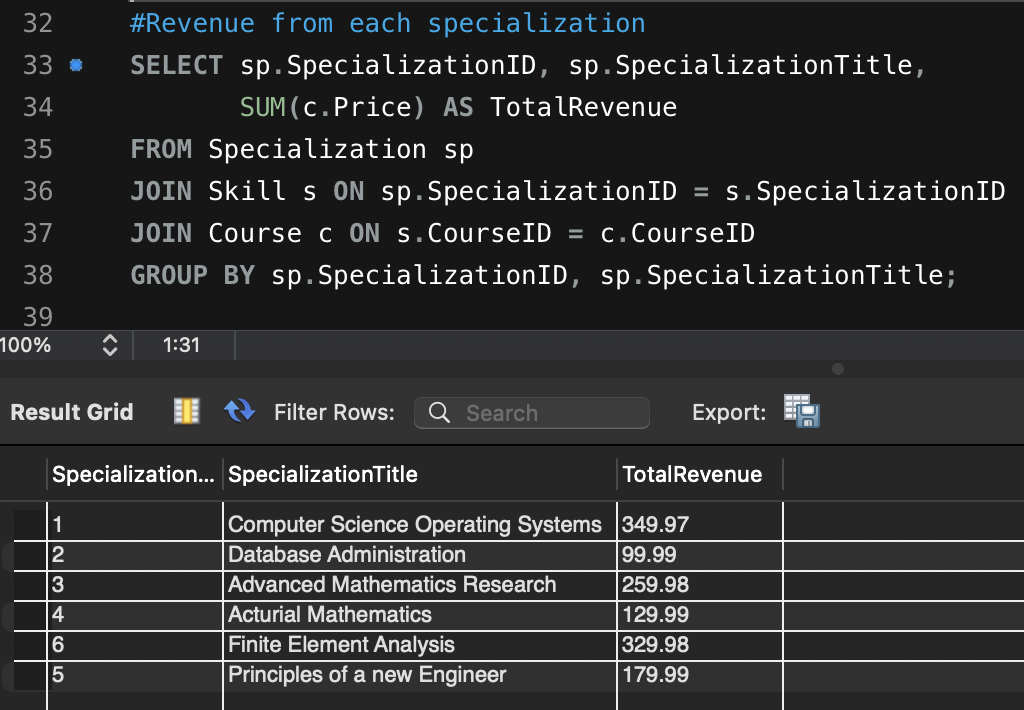


1. Find Courses and their corresponding Instructors

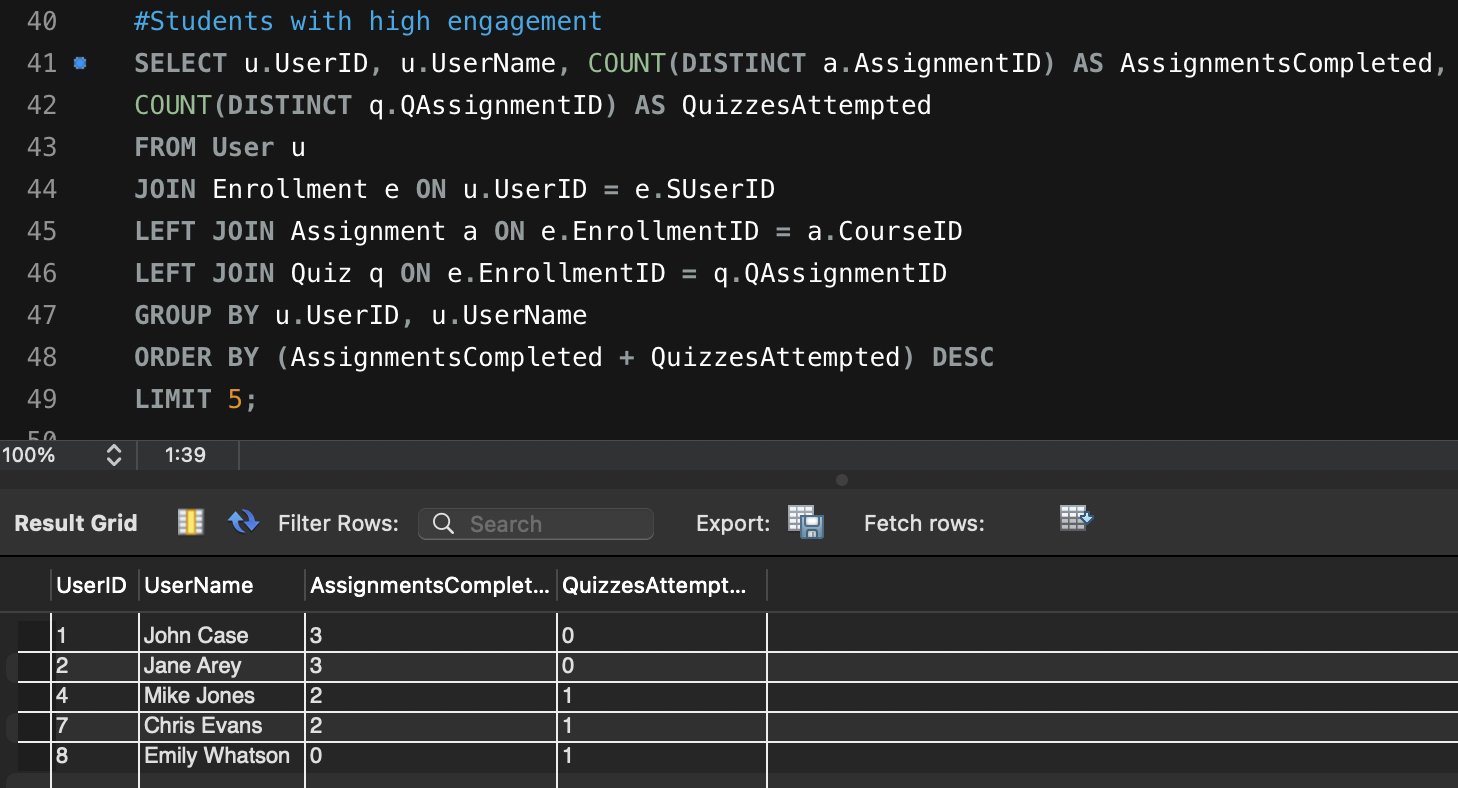


4.2. Using Joins

1. Revenue from each specialization

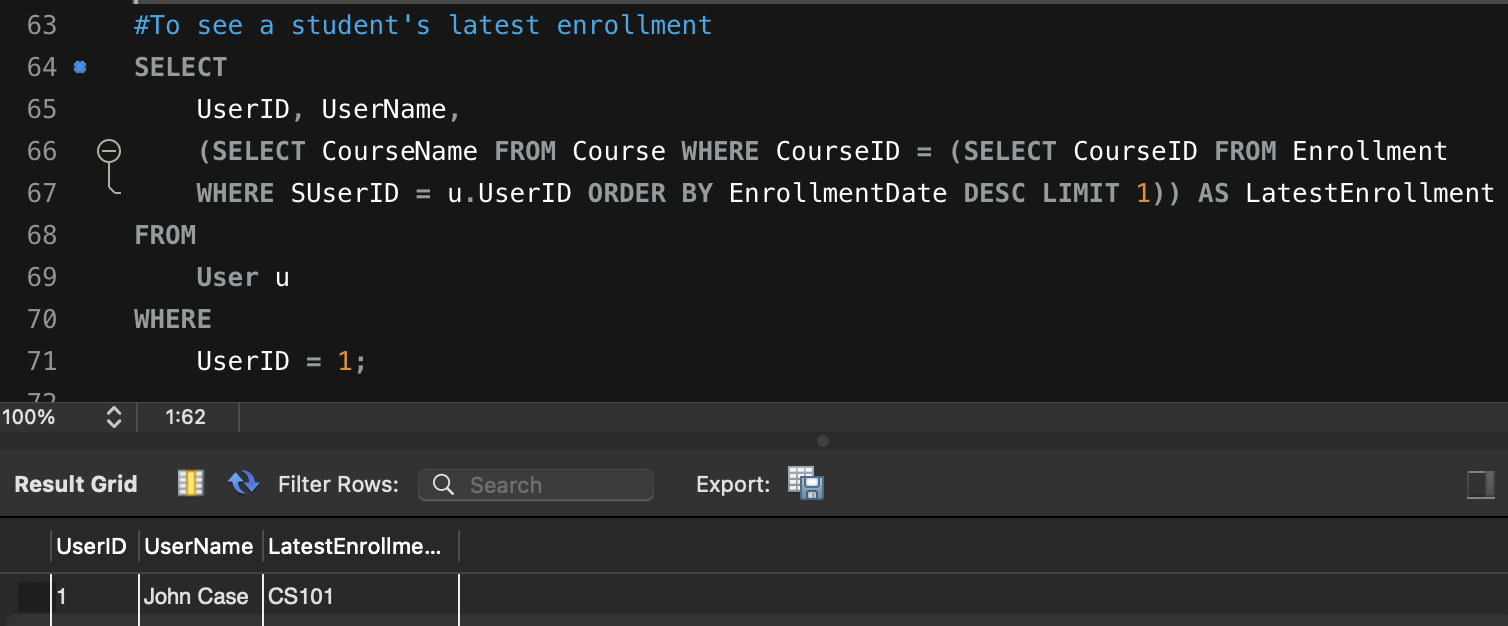


1. Students with high engagement



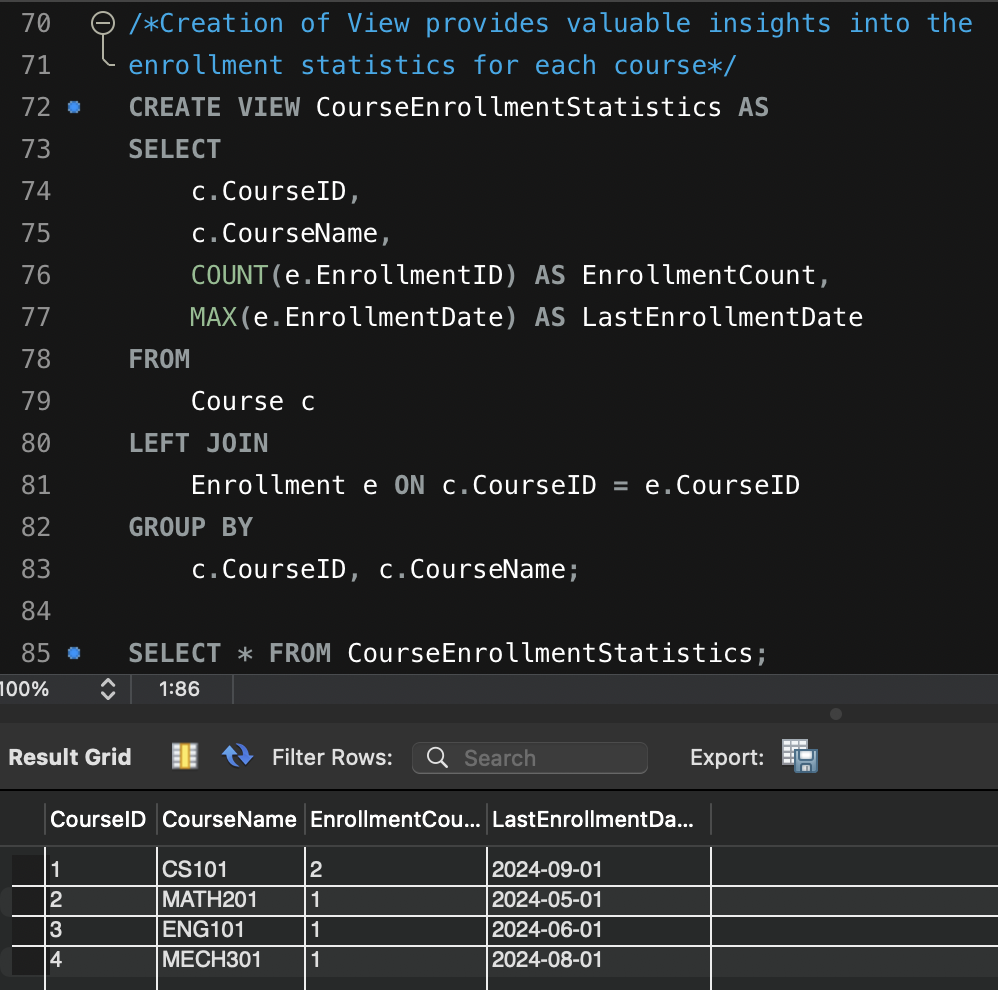
4.3. Using Sub-Queries

1. To see a student’s latest enrollment



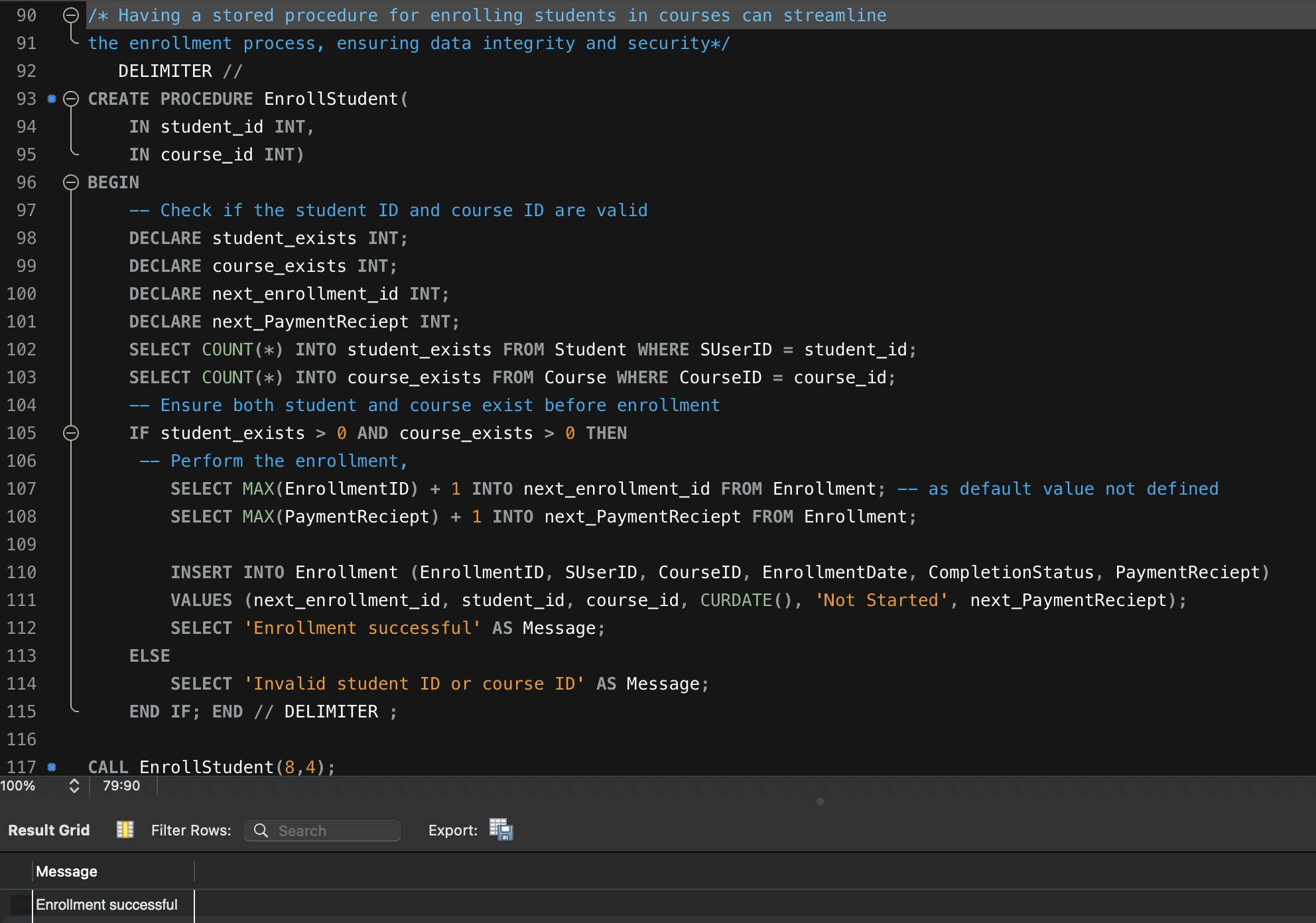
4.4. Using Views

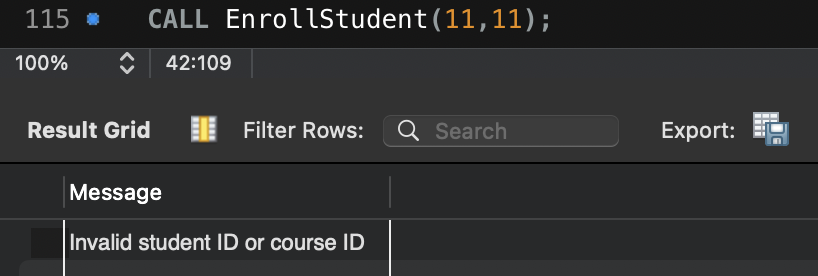
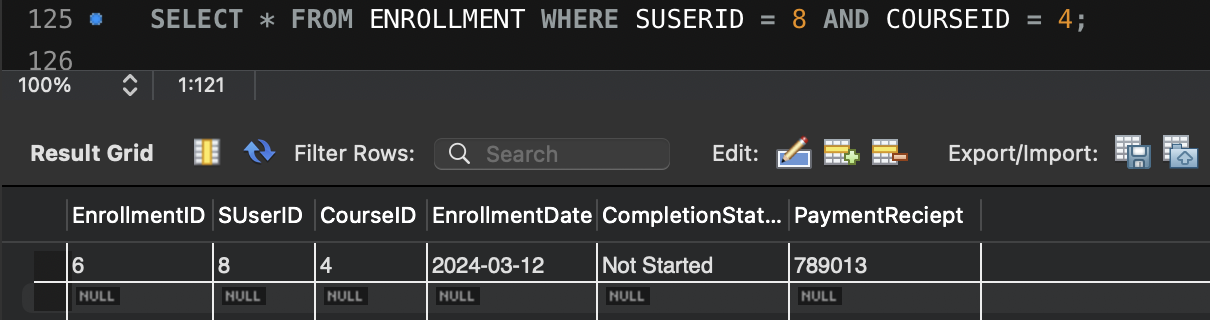
1. View to provide enrollment statistics for each course



4.5. Using Stored Procedure

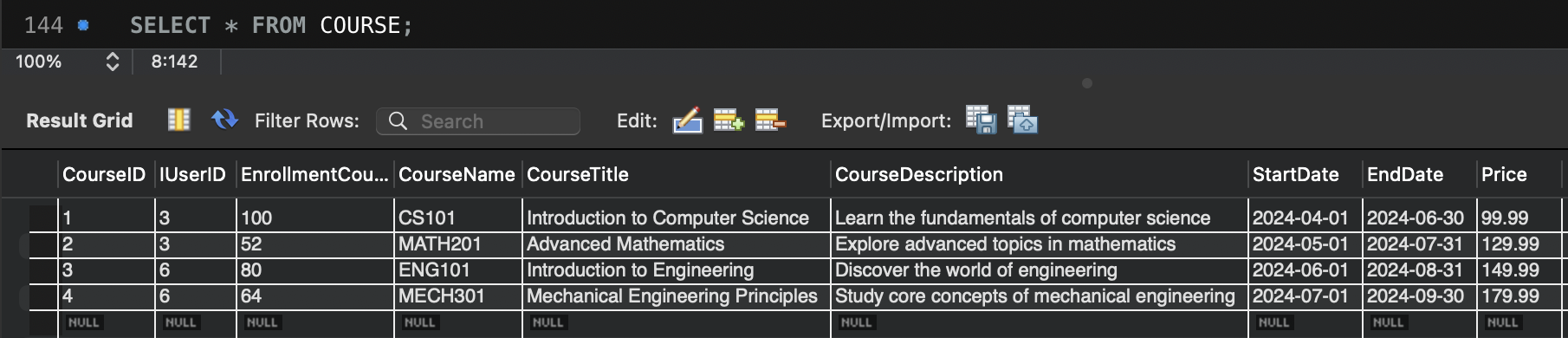
1. To auto populate the enrollment table when there is a new user with a course

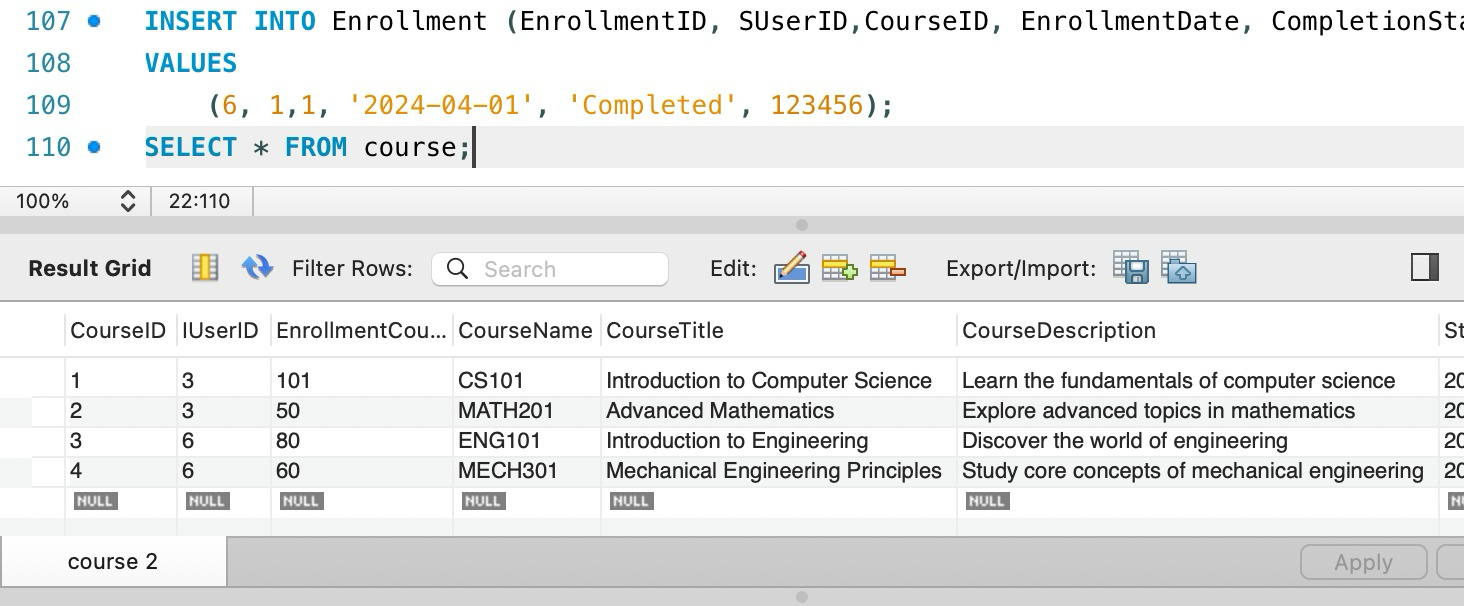
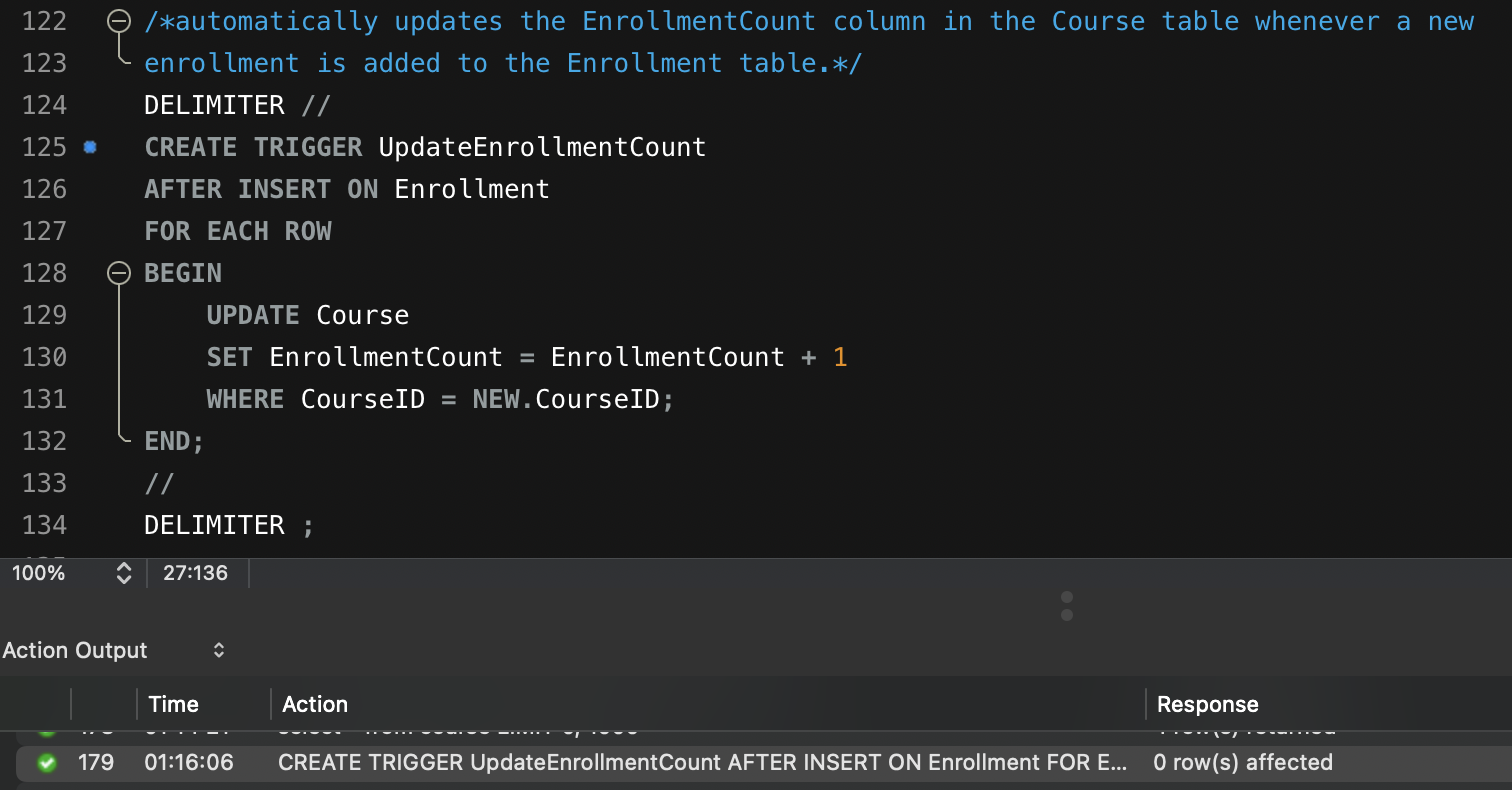




4.6. Using Trigger

1. Updating enrollment count in Course table for new enrollments





5. Administrative Functions

* Some administrative functions implemented within the organization to support the database system’s management and maintenance is to continuously monitor the database system to detect and resolve any issues that may pop up.
* Establish a backup and recovery procedure to protect against any data loss and system failures. We took in account many security controls for the data administration.
* First is access control, the database must be on a separate partition and remote access must be limited by having secured protocols. The database must be installed on the separate partition as OS, so that the DB cannot impact the base OS or any other server installed. We are ensuring Surface hardening, by disabling any unnecessary services installed.
* Then we have Network security to have IP whitelisting, so the database is only accessible by authorized personnel only. Implementing encryption for both data at rest and data in transit. Next, all passwords must adhere to the company's password policies.
* Making sure to enable logging and monitoring of important activities which might be needed at the time of audit or incidents.
* Keeping the database up-to-date is also crucial to prevent any zero-day vulnerability. Lastly, backups must be taken on a separate drive to ensure availability incase of any failure.

5.1. Potential Issues

Some potential issues that may arise during the implementation and application of the proposed database are:

* Data quality issues such as there being errors in the data entry, having duplications, inconsistencies and the naming convention which means having the same uniform format for the attributes and entities.
* Issues such as data format mismatches, data transformation errors or synchronization issues can occur during data exchange processes. If there is large data volume then that can lead to the query running slowly and might show that the data is not able to scale effectively.
* We also had to install many DBMS tools such as MySQL Workbench which required a MySQL server, and connectors. We had to manage these dependencies and had to ensure that they are correctly installed and configured which can be challenging.
* Next, as we started implementing the database we discovered that the data types for a few of the attributes needed to be modified since some of the data types did not match with what we had initially. Changing data types after inserting values into the table challenges the data integrity of the particular attribute.
* We had to verify the accuracy and completeness of the database implementation through rigorous testing and validation procedures which is essential.

5. 2. Measures and Features Incorporated

* In order to address the issues mentioned above we ensure data quality management.
* We used data validation rules, constraints and data cleansing to ensure accuracy and consistency.
* We also used standardized naming conventions and enforce data integrity constraints to prevent errors and duplications. For example, before our attributes all were shown in different formats such as Start\_date. We changed all the attribute formats so it matched this format accordingly: StartDate.
* To address the data type management we would conduct analysis to accurately define data types for each attribute.
* We would convert data types as needed and make sure there was minimal impact on data integrity.
* We would check the data types during implementation to make sure it was consistent throughout.
* We created dummy variables and ran test scripts to verify the accuracy, completeness and functionality of the database implementation.

6. Conclusion

The database management system, as we know, plays a massive part in the designing and functioning of applications. However, the business side plays a crucial part in designing a system. A robust system will not only save many resources for the company but also has the potential to be efficient and create a solid user base. Furthermore, a database is an intrinsic part of an organization that maintains the workflow. Hence the database team needs to work in tandem with business analysts, database security, and the application/developer team to create a seamless architecture that can benefit the business in the long run.

Our Coursera database project has been instrumental in illustrating the practical application of database management systems within the context of an e-learning platform. By designing and implementing a robust database system, we have effectively addressed the complex data management requirements inherent to such platforms.

One of the key takeaways was that the database design process is an agile process rather than a waterfall process. Our team had to go through multiple iterations to trace the business requirement and assess the scalability of the attributes and tables. This introduced us to the process that the Database Management team goes through during the development cycle.