

# University Student Record – Part II

## 1. What extraction and schema changes will be required to add new data such as course enrolments?

**Ans:** To add new data such as course enrolments, we would need to modify the existing schema to include a new model for courses and a new model to store the enrolments for each student. We would also need to modify the CSV files to include the new data and update the code to read and transform the data appropriately.

## 2. How would you modify just the output json if the student also had multiple course enrolments, each with an assigned letter grade?

**Ans:** If the student had multiple course enrolments, each with an assigned letter grade, we would need to modify the 'plan' dictionary in the output JSON to include a new key for the list of courses, each with their respective letter grades.

## 3. Scalability: If the input had 5000 students instead of just 2:

### 1. What factors should be considered when extracting data from database. How would you optimally handle “retrieving” queries?

**Ans:** When dealing with a larger number of students (5000 in this case) instead of just 2, there are several factors that should be considered:

**Data volume:** The volume of data being extracted is larger when dealing with 5000 students, it means that the extraction process will take longer and put a greater strain on the database.

**Query optimization:** To optimize the query, you can use indexing, which can speed up retrieval queries by allowing the database to quickly locate the relevant data.

**Resource allocation:** The database server needs to have sufficient resources to handle the increased workload, including CPU, memory, and storage.

**Data modeling:** The database schema needs to be optimized for the increased volume of data.

**Concurrency:** With a larger number of students, there may be more concurrent queries being executed at the same time.

It is important to optimize the query to retrieve only the necessary data and avoid unnecessary joins, subqueries, or other operations that can slow down the query. DE normalizing tables, partitioning data across multiple tables, or using other techniques to reduce the amount of data being retrieved and processed.

### 2. How would you update your “storing” queries?

**Ans:** When updating "storing" queries for a larger number of records. It's important to optimize the queries, ensure data consistency, test them thoroughly, and monitor performance.

### 3. What about 1 million usernames? What factors need to be considered if you have records of 1 million students?

**Ans:** Overall, scalability is considered when dealing with larger volumes of data like 1 Million. By optimizing queries, allocating resources appropriately, optimizing data modeling, ensuring concurrency control, you can ensure that your database system can handle the increased data volume without any hindrance.