

PART-2 Report: Data Migration

The migration process involved several transformations:

- 1) **Denormalization:** The many-to-many relationship between students and courses was handled by embedding the enrollment data within each student's document (courses field) and each course's document (students field).
- 2) **Data Cleaning:** The data from PostgreSQL was cleaned to ensure it was in the correct format for MongoDB. This included:
 - Converting PostgreSQL types (e.g., SERIAL, DECIMAL, CHAR) to appropriate MongoDB types (e.g., NumberInt, NumberDecimal, String).
 - Checking for consistency in foreign key references between tables before embedding them in MongoDB.
- 3) **Data Extraction:** Using psycopg2, data was fetched from each PostgreSQL

fetch_data: Fetches records from PostgreSQL for the specified table.
- 4) **Data Transformation:** After fetching the data, it was transformed into MongoDB's document structure:
 - transform_students: Transforms student data, embedding an empty list for courses that the student will later be enrolled in.
 - transform_courses: Transforms course data, embedding an empty list for students who will be enrolled.
 - transform_enrollments: Maps the enrollment relationships between students and courses.
 - transform_instructors and transform_departments: Convert instructor and department data into a format compatible with MongoDB.