Unit 7 - Normalisation Task

Normalisation steps performed

1. **Analyse raw data** - checked the un-normalised data sheet to understand layout and detect repeating student blocks and empty cells. The data table is not atomic and has repeating course rows per student.
2. **Convert to 1NF** - duplicated student-level fields so each row contains complete student details plus a single course record. Purpose is to ensure atomic rows.
    - Result: 15 student course enrolment rows
3. **Identify dependencies**:
    - StudentNumber has student attributes (name, date of birth, support)
    - CourseName has course attributes (exam board, teacher)
    - StudentNumber + CourseName has exam scores
4. **Create 2NF tables** - split the 1NF table into three tables to remove partial dependencies:
    - **Student** (StudentNumber, StudentName, DateOfBirth, Support) - 5 rows
    - **Course** (CourseName, ExamBoards, TeacherName) - 6 rows
    - **Enrolment** (StudentNumber, CourseName, ExamScore) - 15 rows
5. **Confirm 3NF (remove transitive deps)** - checked that non-key attributes were stored with the entities they depend on (teacher and exam board kept in Courses) and that Enrolments contains only attributes dependent on the composite key. Final schema meets 3NF.
6. **Handle student names** - split StudentName into Forenames and Surname for clearer storage and matching.
7. **Data quality check** - observed the same ExamScore repeated across multiple course rows for individual students. Need to confirm whether ExamScore is a student-level overall score or a per-course enrolment score. I duplicated the ExamScore pending confirmation.

Dean Carey – September 2025