

Unit 7: Constructing Normalised Tables and Database Build

In this unit we:

- Learned how to take a flat file and turn it into a clean, normalised database structure.
- Built a relational database and tested how well it works in practice.
- Figured out which data items depend on each other and how they're connected.
- Looked at any limitations or restrictions that come with the dataset and how it can be used.
- Identified which data attributes are actually important and necessary for working with that dataset.

Normalisation Task

We were given a table with data in un-normalised form and were requested to normalise the data to 3rd form (3NF), showing each process, from 1NF to 3NF.

Table

Student Number	Student Name	Exam Score	Support	Date of Birth	Course Name	Exam Boards	Teacher Name
1001	Bob Baker	78	No	25/08/2001	Computer Science Maths Physics	BCS EdExcel OCR	Mr Jones Ms Parker Mr Peters
1002	Sally Davies	55	Yes	02/10/1999	Maths Biology Music	AQA WJEC AQA	Ms Parker Mrs Patel Ms Daniels
1003	Mark Hanmill	90	No	05/06/1995	Computer Science Maths Physics	BCS EdExcel OCR	Mr Jones Ms Parker Mr Peters
1004	Anas Ali	70	No	03/08/1980	Maths Physics Biology	AQA OCR WJEC	Ms Parker Mr Peters Mrs Patel
1005	Cheuk Yin	45	Yes	01/05/2002	Computer Science Maths Music	BCS EdExcel AQA	Mr Jones Ms Parker Ms Daniels

Normalisation Process

1. First Normal Form(1NF)

- This is achieved when every column is atomic and each row is unique (Rob & Coronel, 2007), as seen below.

1	Student Number	Student Name	Exam Score	Support	Date of Birth	Course Name	Exam Boards	Teacher Name
2	1001	Bob Baker	78	No	2001-08-25	Computer Science	BCS	Mr Jones
3	1001	Bob Baker	78	No	2001-08-25	Maths	EdExcel	Ms Parker
4	1001	Bob Baker	78	No	2001-08-25	Physics	OCR	Mr Peters
5	1002	Sally Davies	55	Yes	1999-10-02	Maths	AQA	Ms Parker
6	1002	Sally Davies	55	Yes	1999-10-02	Biology	WJEC	Mrs Patel
7	1002	Sally Davies	55	Yes	1999-10-02	Music	AQA	Ms Daniels
8	1003	Mark Hanmill	90	No	1995-06-05	Computer Science	BCS	Mr Jones
9	1003	Mark Hanmill	90	No	1995-06-05	Maths	EdExcel	Ms Parker
10	1003	Mark Hanmill	90	No	1995-06-05	Physics	OCR	Mr Peters
11	1004	Anas Ali	70	No	1980-08-03	Maths	AQA	Ms Parker
12	1004	Anas Ali	70	No	1980-08-03	Physics	OCR	Mr Peters
13	1004	Anas Ali	70	No	1980-08-03	Biology	WJEC	Mrs Patel
14	1005	Cheuk Yin	45	Yes	2002-05-01	Computer Science	BCS	Mr Jones
15	1005	Cheuk Yin	45	Yes	2002-05-01	Maths	EdExcel	Ms Parker
16	1005	Cheuk Yin	45	Yes	2002-05-01	Music	AQA	Ms Daniels

2. Second Normal Form(2NF)

- I separated student-specific data, course-specific data and exam data. This step removed partial dependencies (Elmasri & Navathe, 2015)

Student Table

	A	B	C	D
1	Student Number	Student Name	Date of Birth	Support
2	1001	Bob Baker	2001-08-25	No
3	1002	Sally Davies	1999-10-02	Yes
4	1003	Mark Hanmill	1995-06-05	No
5	1004	Anas Ali	1980-08-03	No
6	1005	Cheuk Yin	2002-05-01	Yes

Course Table

	A	B	C
1	Course Name	Exam Board	Teacher Name
2	Computer Science	BCS	Mr Jones
3	Maths	EdExcel	Ms Parker
4	Physics	OCR	Mr Peters
5	Maths	AQA	Ms Parker
6	Biology	WJEC	Mrs Patel
7	Music	AQA	Ms Daniels

Exam Table

	A	B	C
1	Student Number	Course Name	Exam Score
2	1001	Computer Science	78
3	1001	Maths	78
4	1001	Physics	78
5	1002	Maths	55
6	1002	Biology	55
7	1002	Music	55
8	1003	Computer Science	90
9	1003	Maths	90
10	1003	Physics	90
11	1004	Maths	70
12	1004	Physics	70
13	1004	Biology	70
14	1005	Computer Science	45
15	1005	Maths	45
16	1005	Music	45

3. Third Normal Form(3NF)

- I retained the Students Table, Courses Table, and Exams Table from 2NF
- To eliminate transitive dependencies,
 - I introduced a unique Course ID in the Courses Table, making it the primary key
 - Replaced the composite identifier in the Exam table with the Course ID as the foreign key

Student Table

	A	B	C	D
1	Student Number	Student Name	Date of Birth	Support
2	1001	Bob Baker	2001-08-25	No
3	1002	Sally Davies	1999-10-02	Yes
4	1003	Mark Hanmill	1995-06-05	No
5	1004	Anas Ali	1980-08-03	No
6	1005	Cheuk Yin	2002-05-01	Yes

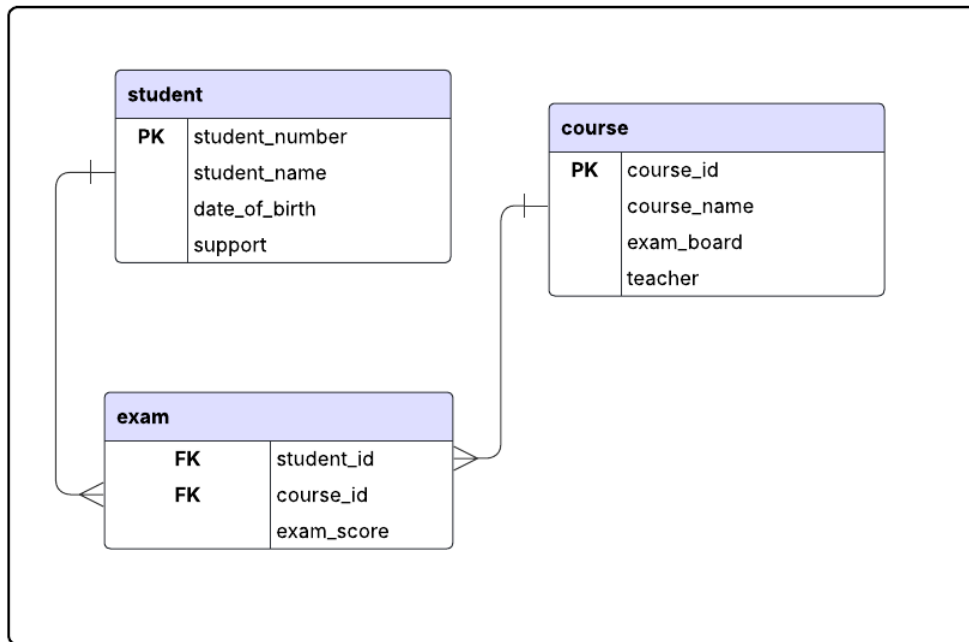
Course Table

	A	B	C	D
1	Course ID	Course Name	Exam Boards	Teacher Name
2	1	Computer Science	BCS	Mr Jones
3	2	Maths	EdExcel	Ms Parker
4	3	Physics	OCR	Mr Peters
5	4	Maths	AQA	Ms Parker
6	5	Biology	WJEC	Mrs Patel
7	6	Music	AQA	Ms Daniels

Exam Table

	A	B	C
1	Student Number	Course ID	Exam Score
2	1001	1	78
3	1001	2	78
4	1001	3	78
5	1002	4	55
6	1002	5	55
7	1002	6	55
8	1003	1	90
9	1003	2	90
10	1003	3	90
11	1004	4	70
12	1004	3	70
13	1004	5	70
14	1005	1	45
15	1005	2	45
16	1005	6	45

Data Build Task



References

Elmasri, R. and Navathe, S.B. (2015) *Fundamentals of Database Systems*. 7th edn. Boston: Pearson Education.

Rob, P. and Coronel, C. (2007) *Database Systems: Design, Implementation, and Management*. 7th edn. Boston, MA: Thomson Course Technology.