Task: normalise the below table to 3rd Normal Form (3NF) showing each step of the process

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

1st normal form (1NF) - each row is unique and each value is atomic

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

2NF - must be in 1NF and each column that is not primary key must be dependent on the primary key

- primary key would be student number.
- Course name, exam boards and teachers are not dependent on student number;
  - therefore need new tables

1. Student Details

Student Number	Student Name	Exam Score	Support	Date of Birth
1001	Bob Baker	78	No	2001-08-25
1002	Sally Davies	55	Yes	1999-10-02
1003	Mark Hanmill	90	No	1995-06-05
1004	Anas Ali	70	No	1980-08-03
1005	Cheuk Yin	45	Yes	2002-05-01

2. Course Details

Course Name	Exam Boards	Teacher Name
Computer Science	BCS	Mr Jones
Maths	EdExcel	Ms Parker
Physics	OCR	Mr Peters
Biology	WJEC	Mrs Patel
Music	AQA	Ms Daniels

3. Courses Taken

Student Number	Course Name
1001	Computer Science
1001	Maths
1001	Physics
1002	Maths
1002	Biology
1002	Music
1003	Computer Science
1003	Maths
1003	Physics
1004	Maths
1004	Physics
1004	Biology
1005	Computer Science
1005	Maths
1005	Music

3NF - must be in 2NF and every column not part of primary key is only dependent on the primary key

Student details table remains the same
Table 1; student details

Student Number	Student Name	Exam Score	Support	Date of Birth
1001	Bob Baker	78	No	2001-08-25
1002	Sally Davies	55	Yes	1999-10-02
1003	Mark Hanmill	90	No	1995-06-05
1004	Anas Ali	70	No	1980-08-03
1005	Cheuk Yin	45	Yes	2002-05-01

Course details table should be divided as exam boards may cover different subjects and teachers may teach more than one subject

Table 2; exam boards

EBID	Exam Boards
EB1	BCS
EB2	EdExcel
EB3	OCR
EB4	WJEC
EB5	AQA

Table 3; teachers

TID	Teacher
T1	Mr Jones
T2	Ms Parker
T3	Mr Peters
T4	Mrs Patel
T5	Ms Daniels

Table 4; courses taken

Student Number	Course Name
1001	Computer Science
1001	Maths
1001	Physics
1002	Maths
1002	Biology
1002	Music
1003	Computer Science
1003	Maths
1003	Physics
1004	Maths
1004	Physics
1004	Biology
1005	Computer Science
1005	Maths
1005	Music

Table 5. Link teacher to course

Course Name	TID
Computer Science	T1
Maths	T2
Physics	T3
Biology	T4
Music	T5
Music	T5

Table 6. Link EBID to course

EBID	Course Name
EB1	Computer Science
EB2	Maths
EB3	Physics
EB4	Biology
EB5	Music

Observations; learnings

It is not always straightforward to determine what the primary key(s) are and this will determine how the tables are separated. I found that it becomes harder to read as a human as we progress through the normal forms, although it is clear why this helps reduce redundancies and limit the potential for data conflicts or inaccuracies. Entering the data into the tables needed to be closely checked as one mistake early on can be propagated across subsequent tables as moving from1st to 2nd to 3rd Normal forms.