

**Analysis, Design and Implementation of a Relational  
Database System PRAC1**

**Kyle Cuthbert – st20249341**

## Table of Contents

Normalization.....	3
Data Specified .....	3
Separate Normalization .....	4
Revised Normalization .....	8
Schema.....	11
Entity Relation Diagram .....	13
Data dictionary.....	14
Creation Script .....	16
Insert Data Script .....	17
Various Query Examples .....	18

## Normalization

### Data Specified

For the assessment brief we have been given the below information which represents the requirements of the users:

- For students
  - Name
  - Place of birth
  - Date of birth
  - Home School
  - City
  - Country
  - Contact number
  - Courses studied and grades (at home school not from course list)
  - Recommendation letter (*Yes/No Field, to indicate if it is included*)
  - Invitation from the host (*Yes/No Field, to indicate if it is included*)
- For researchers:
  - Name
  - Place of birth
  - Date of birth
  - Home university
  - City
  - Country
  - Contact number
  - Required fund (*Amount required*)
  - Proposal (*One proposal only*):
    - Intro
    - Background
    - Risks
    - Methods

Course title	Course code	Host university	Duration	Type
Computer Security	CS1022	Cardiff Met University	36 months	BSc
Data Science	DS2221	Swansea University	12 months	MSc
Bioinformatics	BI0011	MIT	36 months	PhD

Grant title	Available fund	Duration
Short term	20K	6 months
ESRC	250K	24 months
Welsh Government	150K	12 months

## Separate Normalization

The first step that we must take in the normalization process is to put the related data together to compose a ‘bad’ table and populate it with example test data. This ‘bad’ table will then be decomposed into several smaller ‘good’ tables that will be the process of normalization.

The stages of normalization are called normal forms with the higher the number of normal forms representing a better design.

Having interpreted the information provided I initially decided that two separate normalization processes will be carried out, one for the student data and one for the researcher data. However, after completing the steps of normalization I noticed that I could link the two separate sets of data with a table for common personal data.

Below I have explained the steps taken for both normalizations.

Test data has been added to represent the un-normalized data:

### Student Data

Name	Place of birth	Date of birth	Home School	City	Country	Contact number	Courses and grade	Recommendations letter	Invitation from host	Course title	Course code	Host University	Duration	Type
Wyatt Page	Cardiff	26/05/1995	John Frost High	London	England	0711111111	History A, Maths B	Y	Y	Computer Security	CS1022	Cardiff Met University	36	BSc
Karlie Decker	Newport	03/08/2000	Duffryn Comp	Manchester	England	0711111112	IT A, Maths C	Y	Y	Data Science	DS2221	Swansea University	12	MSc
Ishaan Harvey	Swansea	26/08/1971	St Michael's	Birmingham	England	0711111113	Geography D	N	Y	Bioinformatics	BI0011	MIT	36	PhD
Laila Murray	Wrexham	08/10/1986	Gaer	Bristol	England	0711111114	English C	N	Y	Computer Security	CS1022	Cardiff Met University	36	BSc
Leonidas Jefferson	Monmouth	11/04/1974	Pill	Sheffield	England	0711111115	IT A, Science D	Y	Y	Data Science	DS2221	Swansea University	12	MSc
Joy Rush	Bridgend	24/01/1993	John Frost High	Glasgow	Scotland	0711111116	Maths C, IT D, English C	Y	Y	Bioinformatics	BI0011	MIT	36	PhD

### Researcher

Name	Place of birth	Date of birth	Home School	City	Country	Contact number	Required fund	Proposal	Grant title	Available fund	Duration
Amaya Hebert	Cardiff	29/04/1978	John Frost High	London	England	0722222222	1000	Intro	Short term	20K	6 months
Gavin Stephenson	Newport	20/11/1978	Duffryn Comp	Manchester	England	0722222223	200	Background	ESRC	250K	24 months
Chase Lewis	Swansea	30/01/2000	St Michael's	Birmingham	England	0722222224	600	Risks	Welsh Government	150K	12 months
Cael Boyer	Wrexham	07/08/1996	John Frost High	London	England	0722222225	800	Methods	Short term	20K	6 months
Glenn Mccann	Monmouth	06/05/1981	Duffryn Comp	Manchester	England	0722222226	1000	Intro	ESRC	250K	24 months
Emilia Peters	Bridgend	15/10/1992	St Michael's	Birmingham	England	0722222227	200	Background	Welsh Government	150K	12 months
Jeremiah Blackburn	Carmarthen	14/09/1966	St Brides	Madrid	Spain	0722222228	600	Risks	Short term	20K	6 months

The next stage that we must take is normalizing the data to 1NF. For data to meet 1NF a primary key must be included.

As part of 1NF columns need to be unique, i.e., each cell can only contain one value and cannot have mixed types of data (there are no examples of mixed data types here). As a result, the ‘Courses and grades’ column will need to be separated into two separate columns.

The next stage of making the data conform to 1NF will be to remove the repeating attributes. The Course Code will remain in the student table as the foreign key.

## Student

Student Id	Name	Place of birth	Date of birth	Home School	City	Country	Contact number	Courses	Grade	Recommendations letter	Invitation from host	Course code
ST1	Wyatt Page	Cardiff	26/05/1995	John Frost High	London	England	0711111111	History	A	Y	Y	CS1022
ST1	Wyatt Page	Cardiff	26/05/1995	John Frost High	London	England	0711111111	Maths	B	Y	Y	CS1022
ST5	Leonidas Jefferson	Monmouth	11/04/1974	Pill	Sheffield	England	0711111115	IT	A	Y	Y	DS2221
ST5	Leonidas Jefferson	Monmouth	11/04/1974	Pill	Sheffield	England	0711111115	Science	D	Y	Y	DS2221
ST4	Laila Murray	Wrexham	08/10/1986	Gaer	Bristol	England	0711111114	English	C	N	Y	CS1022
ST2	Karlie Decker	Newport	03/08/2000	Duffryn Comp	Manchester	England	0711111112	IT	A	Y	Y	DS2221
ST2	Karlie Decker	Newport	03/08/2000	Duffryn Comp	Manchester	England	0711111112	Maths	C	Y	Y	DS2221
ST6	Joy Rush	Bridgend	24/01/1993	John Frost High	Glasgow	Scotland	0711111116	Maths	C	Y	Y	BI0011
ST6	Joy Rush	Bridgend	24/01/1993	John Frost High	Glasgow	Scotland	0711111116	IT	D	Y	Y	BI0011
ST6	Joy Rush	Bridgend	24/01/1993	John Frost High	Glasgow	Scotland	0711111116	English	C	Y	Y	BI0011
ST3	Ishaan Harvey	Swansea	26/08/1971	St Michael's	Birmingham	England	0711111113	Geography	D	N	Y	BI0011

Course code	Course title	Host University	Duration	Type
CS1022	Computer Security	Cardiff Met University	36	BSc
DS2221	Data Science	Swansea University	12	MSc
BI0011	Bioinformatics	MIT	36	PhD

## Researcher

Researcher id	Name	Place of birth	Date of birth	Home School	City	Country	Contact number	Required fund	Proposal	Grant title	Available fund	Duration
R1	Amaya Hebert	Cardiff	29/04/1978	John Frost High	London	England	0722222222	1000	Intro	Short term	20K	6 months
R2	Gavin Stephenson	Newport	20/11/1978	Duffryn Comp	Manchester	England	0722222223	200	Background	ESRC	250K	24 months
R3	Chase Lewis	Swansea	30/01/2000	St Michael's	Birmingham	England	0722222224	600	Risks	Welsh Government	150K	12 months
R4	Cael Boyer	Wrexham	07/08/1996	John Frost High	London	England	0722222225	800	Methods	Short term	20K	6 months
R5	Glenn McCann	Monmouth	06/05/1981	Duffryn Comp	Manchester	England	0722222226	1000	Intro	ESRC	250K	24 months
R6	Emilia Peters	Bridgend	15/10/1992	St Michael's	Birmingham	England	0722222227	200	Background	Welsh Government	150K	12 months
R7	Jeremiah Blackburn	Carmarthen	14/09/1966	St Brides	Madrid	Spain	0722222228	600	Risks	Short term	20K	6 months

The next step to be taken as part of the normalization process is to take the data from 1NF to 2NF. Separate tables should be created for sets of values that apply to multiple records. These must be related with a foreign key. Please note that the student grades have a composite key.

### Student

Student Id	Name	Place of birth	Date of birth	Home School	City	Country	Contact number	Recommend	Invitatio	Course code
ST1	Wyatt Page	Cardiff	26/05/1995	John Frost High	London	England	0711111111	Y	Y	CS1022
ST5	Leonidas Jefferson	Monmouth	11/04/1974	Pill	Sheffield	England	0711111115	Y	Y	DS2221
ST4	Laila Murray	Wrexham	08/10/1986	Gaer	Bristol	England	0711111114	N	Y	CS1022
ST2	Karlie Decker	Newport	03/08/2000	Duffryn Comp	Manchester	England	0711111112	Y	Y	DS2221
ST6	Joy Rush	Bridgend	24/01/1993	John Frost High	Glasgow	Scotland	0711111116	Y	Y	BI0011
ST3	Ishaan Harvey	Swansea	26/08/1971	St Michael's	Birmingham	England	0711111113	N	Y	BI0011

Course code	Course title	Host University	Duration	Type
CS1022	Computer Security	Cardiff Met University	36	BSc
DS2221	Data Science	Swansea University	12	MSc
BI0011	Bioinformatics	MIT	36	PhD

Student Id	Courses	Grade
ST1	History	A
ST1	Maths	B
ST5	IT	A
ST5	Science	D
ST4	English	C
ST2	IT	A
ST2	Maths	C
ST6	Maths	C
ST6	IT	D
ST6	English	C
ST3	Geography	D

## Researcher

<u>Researcher_id</u>	Name	Place of birth	Date of birth	Home School	City	Country	Contact number	Required fund	Proposal	Grant_id
R1	Amaya Hebert	Cardiff	29/04/1978	John Frost High	London	England	0722222222	1000	Intro	G1
R2	Gavin Stephenson	Newport	20/11/1978	Duffryn Comp	Manchester	England	0722222223	200	Background	G2
R3	Chase Lewis	Swansea	30/01/2000	St Michael's	Birmingham	England	0722222224	600	Risks	G3
R4	Cael Boyer	Wrexham	07/08/1996	John Frost High	London	England	0722222225	800	Methods	G1
R5	Glenn Mccann	Monmouth	06/05/1981	Duffryn Comp	Manchester	England	0722222226	1000	Intro	G2
R6	Emilia Peters	Bridgend	15/10/1992	St Michael's	Birmingham	England	0722222227	200	Background	G3
R7	Jeremiah Blackburn	Carmarthen	14/09/1966	St Brides	Madrid	Spain	0722222228	600	Risks	G1

<u>Grant_id</u>	Grant title	Available fund	Duration
G1	Short term	20K	6 months
G2	ESRC	250K	24 months
G3	Welsh Government	150K	12 months

The next stage of normalization is to take the form of 3NF which is to eliminate any transitive dependencies where a non-key attribute does not depend on another non-key attribute. In the student data, we do not have any transitive dependencies, but we do in the researchers.

<u>Researcher_id</u>	Name	Place of birth	Date of birth	Home School	City	Country	Contact number	Grant_id	Proposal_id
R1	Amaya Hebert	Cardiff	29/04/1978	John Frost High	London	England	0722222222	G1	1
R2	Gavin Stephenson	Newport	20/11/1978	Duffryn Comp	Manchester	England	0722222223	G2	2
R3	Chase Lewis	Swansea	30/01/2000	St Michael's	Birmingham	England	0722222224	G3	3
R4	Cael Boyer	Wrexham	07/08/1996	John Frost High	London	England	0722222225	G1	4
R5	Glenn Mccann	Monmouth	06/05/1981	Duffryn Comp	Manchester	England	0722222226	G2	1
R6	Emilia Peters	Bridgend	15/10/1992	St Michael's	Birmingham	England	0722222227	G3	2
R7	Jeremiah Blackburn	Carmarthen	14/09/1966	St Brides	Madrid	Spain	0722222228	G1	3

<u>Grant_id</u>	Grant title	Available fund	Duration
G1	Short term	20K	6 months
G2	ESRC	250K	24 months
G3	Welsh Government	150K	12 months

<u>Proposal_id</u>	Required fund	Proposal
1	1000	Intro
2	200	Background
3	600	Risks
4	800	Methods

## Revised Normalization

As mentioned above I noticed that there was common data between the two separate sets of data and yet there was no link between the two. This meant that there could never be any cross over e.g., a researcher that was also a student or the ability to add another 'role'. I then decided to perform the steps of normalization again which can be seen below and in the '**Joint**' tab of the '**Normalisation.xlsx**' spreadsheet:

### Un-normalized

Name	Place of birth	Date of birth	Home School	City	Country	Contact number	Required fund	Proposal	Grant title	Available fund	Duration	Courses and grade	Recommendations	Invitation from host	Course title	Course code	Host University	Duration	Type
Amaya Hebert	Cardiff	23/04/1978	John Frost High	London	England	0722222222	1000	Intro	Short term	20K	6 months	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
Gavin Stephenson	Newport	20/11/1978	Duffryn Comp	Manchester	England	0722222223	200	Background	ESRC	250K	24 months	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
Chase Lewis	Swansea	30/09/2000	St Michael's	Birmingham	England	0722222224	600	Risks	Welsh Government	150K	12 months	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
Cael Boyer	Wrexham	07/08/1998	John Frost High	London	England	0722222225	800	Methods	Short term	20K	6 months	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
Glenn McCann	Monmouth	06/05/1981	Duffryn Comp	Manchester	England	0722222226	1000	Intro	ESRC	250K	24 months	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
Emilia Peters	Bridgend	15/10/1982	St Michael's	Birmingham	England	0722222227	200	Background	Welsh Government	150K	12 months	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
Jeremiah Blackburn	Carmarthen	14/09/1966	St Brigid's	Madrid	Spain	0722222228	600	Risks	Short term	20K	6 months	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
Wyatt Page	Cardiff	26/05/1995	John Frost High	London	England	0711111111	NULL	NULL	NULL	NULL	NULL	History A, Maths B	Y	Y	Computer Security	CS1022	Cardiff Met University	36	BSc
Karlie Decker	Newport	03/08/2000	Duffryn Comp	Manchester	England	0711111112	NULL	NULL	NULL	NULL	NULL	IT A, Maths C	Y	Y	Data Science	DS2221	Swansea University	12	MSc
Ishaan Harveg	Swansea	26/08/1971	St Michael's	Birmingham	England	0711111113	NULL	NULL	NULL	NULL	NULL	Geography D	N	Y	Bioinformatics	BI0011	MIT	36	PhD
Laila Murray	Wrexham	08/10/1986	Gaer	Bristol	England	0711111114	NULL	NULL	NULL	NULL	NULL	English C	N	Y	Computer Security	CS1022	Cardiff Met University	36	BSc
Leondas Jefferson	Monmouth	19/04/1974	Pill	Sheffield	England	0711111115	NULL	NULL	NULL	NULL	NULL	IT A, Science D	Y	Y	Data Science	DS2221	Swansea University	12	MSc
Joy Rush	Bridgend	24/09/1993	John Frost High	Glasgow	Scotland	0711111116	NULL	NULL	NULL	NULL	NULL	Maths C, IT D, English Y	Y	Y	Bioinformatics	BI0011	MIT	36	PhD

### 1NF

ApplicationGrant_Id	Name	Place of birth	Date of birth	Home School	City	Country	Contact number	Required fund	Proposal	Grant title	Available fund	Duration	Completed Courses	Grades	Recommendations	Invitation from host	Course title	Course code	Host University	Duration	Type
1	Amaya Hebert	Cardiff	23/04/1978	John Frost High	London	England	0722222222	1000	Intro	Short term	20K	6 months	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
2	Gavin Stephenson	Newport	20/11/1978	Duffryn Comp	Manchester	England	0722222223	200	Background	ESRC	250K	24 months	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
3	Chase Lewis	Swansea	30/09/2000	St Michael's	Birmingham	England	0722222224	600	Risks	Welsh Government	150K	12 months	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
4	Cael Boyer	Wrexham	07/08/1998	John Frost High	London	England	0722222225	800	Methods	Short term	20K	6 months	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
5	Glenn McCann	Monmouth	06/05/1981	Duffryn Comp	Manchester	England	0722222226	1000	Intro	ESRC	250K	24 months	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
6	Emilia Peters	Bridgend	15/10/1982	St Michael's	Birmingham	England	0722222227	200	Background	Welsh Government	150K	12 months	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
7	Jeremiah Blackburn	Carmarthen	14/09/1966	St Brigid's	Madrid	Spain	0722222228	600	Risks	Short term	20K	6 months	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
8	Wyatt Page	Cardiff	26/05/1995	John Frost High	London	England	0711111111	NULL	NULL	NULL	NULL	NULL	History A	Y	Y	Computer Security	CS1022	Cardiff Met University	36	BSc	
9	Wyatt Page	Cardiff	26/05/1995	John Frost High	London	England	0711111112	NULL	NULL	NULL	NULL	NULL	Maths B	Y	Y	Computer Security	CS1022	Cardiff Met University	36	BSc	
10	Karlie Decker	Newport	03/08/2000	Duffryn Comp	Manchester	England	0711111113	NULL	NULL	NULL	NULL	NULL	IT A	Y	Y	Data Science	DS2221	Swansea University	12	MSc	
11	Karlie Decker	Newport	03/08/2000	Duffryn Comp	Manchester	England	0711111114	NULL	NULL	NULL	NULL	NULL	Maths C	Y	Y	Data Science	DS2221	Swansea University	12	MSc	
12	Ishaan Harveg	Swansea	26/08/1971	St Michael's	Birmingham	England	0711111115	NULL	NULL	NULL	NULL	NULL	Geography D	N	Y	Bioinformatics	BI0011	MIT	36	PhD	
13	Laila Murray	Wrexham	08/10/1986	Gaer	Bristol	England	0711111116	NULL	NULL	NULL	NULL	NULL	English C	N	Y	Computer Security	CS1022	Cardiff Met University	36	BSc	
14	Leondas Jefferson	Monmouth	19/04/1974	Pill	Sheffield	England	0711111117	NULL	NULL	NULL	NULL	NULL	IT A	Y	Y	Data Science	DS2221	Swansea University	12	MSc	
15	Leondas Jefferson	Monmouth	19/04/1974	Pill	Sheffield	England	0711111118	NULL	NULL	NULL	NULL	NULL	Science D	Y	Y	Data Science	DS2221	Swansea University	12	MSc	
16	Joy Rush	Bridgend	24/09/1993	John Frost High	Glasgow	Scotland	0711111119	NULL	NULL	NULL	NULL	NULL	Maths C	Y	Y	Bioinformatics	BI0011	MIT	36	PhD	
17	Joy Rush	Bridgend	24/09/1993	John Frost High	Glasgow	Scotland	0711111120	NULL	NULL	NULL	NULL	NULL	IT D	Y	Y	Bioinformatics	BI0011	MIT	36	PhD	
18	Joy Rush	Bridgend	24/09/1993	John Frost High	Glasgow	Scotland	0711111121	NULL	NULL	NULL	NULL	NULL	English C	Y	Y	Bioinformatics	BI0011	MIT	36	PhD	

## 2NF

ApplicationGrant_Id	Person_Id	Required fund	Proposal	Grant_id	Recommendations lot	Invitation from host	Course title	Course code	Host University	Duration	Type	
1	1	1000	Intro	1	NULL	NULL	NULL	NULL	NULL	NULL	NULL	
2	2	200	Background	2	NULL	NULL	NULL	NULL	NULL	NULL	NULL	
3	3	600	Risks	3	NULL	NULL	NULL	NULL	NULL	NULL	NULL	
4	4	800	Methods	1	NULL	NULL	NULL	NULL	NULL	NULL	NULL	
5	5	1000	Intro	2	NULL	NULL	NULL	NULL	NULL	NULL	NULL	
6	6	200	Background	3	NULL	NULL	NULL	NULL	NULL	NULL	NULL	
7	7	600	Risks	1	NULL	NULL	NULL	NULL	NULL	NULL	NULL	
8	8	NULL	NULL	Y	Y	Computer Security	CS1022	Cardiff Met Universit	36	BSc		
9	9	NULL	NULL	NULL	Y	Computer Security	CS1022	Cardiff Met Universit	36	BSc		
10	9	NULL	NULL	NULL	Y	Data Science	DS2221	Swansea University	12	MSc		
11	9	NULL	NULL	NULL	Y	Data Science	DS2221	Swansea University	12	MSc		
12	10	NULL	NULL	N	Y	Bioinformatics	BI0011	MIT	36	PhD		
13	11	NULL	NULL	N	Y	Computer Security	CS1022	Cardiff Met Universit	36	BSc		
14	12	NULL	NULL	N	Y	Data Science	DS2221	Swansea University	12	MSc		
15	12	NULL	NULL	N	Y	Data Science	DS2221	Swansea University	12	MSc		
16	13	NULL	NULL	NULL	Y	Y	Bioinformatics	BI0011	MIT	36	PhD	
17	13	NULL	NULL	NULL	Y	Y	Bioinformatics	BI0011	MIT	36	PhD	
18	13	NULL	NULL	NULL	Y	Y	Bioinformatics	BI0011	MIT	36	PhD	
Person_Id	Name	Place of birth	Date of birth	Home School	City	Country	Contact number	PositionCode				
1	Amaya Hebert	Cardiff	23/04/1978	John Frost High	London	England	0722222222	1				
2	Gavin Stephenson	Newport	20/11/1978	Duffryn Comp	Manchester	England	072222223	1				
3	Chase Lewis	Swansea	30/01/2000	St Michael's	Birmingham	England	072222224	1				
4	Cael Boyer	Wrexham	07/08/1936	John Frost High	London	England	072222225	1				
5	Glenn McCann	Monmouth	06/05/1981	Duffryn Comp	Manchester	England	072222226	1				
6	Emilia Peters	Bridgend	15/10/1992	St Michael's	Birmingham	England	072222227	1				
7	Jeremiah Blackburn	Carmarthen	14/03/1966	St Brides	Madrid	Spain	072222228	1				
8	Wyatt Page	Cardiff	26/05/1935	John Frost High	London	England	071111111	2				
9	Karlie Decker	Newport	03/08/2000	Duffryn Comp	Manchester	England	071111112	2				
10	Ishaan Harvey	Swansea	26/08/1971	St Michael's	Birmingham	England	071111113	2				
11	Laila Murray	Wrexham	08/10/1986	Gars	Bristol	England	071111114	2				
12	Leonidas Jefferson	Monmouth	11/04/1974	Pill	Sheffield	England	071111115	2				
13	Jay Rush	Bridgend	24/01/1993	John Frost High	Glasgow	Scotland	071111116	2				
PositionCode	Position											
1	Researcher											
2	Student											
Person_Id	Completed_Courses	Grade										
8	History	A										
8	Maths	B										
9	IT	A										
3	Maths	C										
10	Geography	D										
11	English	C										
12	IT	A										
12	Science	D										
13	Maths	C										
13	IT	D										
13	English	C										
Grant_id	Grant title	Available fund	Duration									
1	Short term	20K	6 months									
2	ESRC	250K	24 months									
3	Welch Government	150K	12 months									

## 3NF

GrantApplication_Id	Person_Id	Proposal_Id	Grant_id
1	1	1	1
2	2	2	2
3	3	3	3
4	4	4	1
5	5	1	2
6	6	2	3
7	7	3	1

Application_Id	Person_Id	Recommendations let	Invitation from host	Course code
1	8	Y	Y	CS1022
2	8	Y	Y	CS1022
3	9	Y	Y	DS2221
4	9	Y	Y	DS2221
5	10	N	Y	BI0011
6	11	N	Y	CS1022
7	12	Y	Y	DS2221
8	12	Y	Y	DS2221
9	13	Y	Y	BI0011
10	13	Y	Y	BI0011
11	13	Y	Y	BI0011

Person_Id	Name	Place of birth	Date of birth	Home School	City	Country	Contact number	PositionCode
1	Amaya Hebert	Cardiff	1978-04-29	John Frost High	London	England	0722222222	1
2	Gavin Stephenson	Newport	1978-11-20	Duffryn Comp	Manchester	England	0722222223	1
3	Chase Lewis	Swansea	2000-01-30	St Michael's	Birmingham	England	0722222224	1
4	Cael Boyer	Wrexham	1996-08-07	John Frost High	London	England	0722222225	1
5	Glen McCann	Monmouth	1981-05-06	Duffryn Comp	Manchester	England	0722222226	1
6	Emilia Peters	Bridgend	1992-10-15	St Michael's	Birmingham	England	0722222227	1
7	Jeremiah Blackburn	Carmarthen	1966-03-14	St Brides	Madrid	Spain	0722222228	1
8	Wyatt Page	Cardiff	1935-05-26	John Frost High	London	England	0711111111	2
9	Karlie Decker	Newport	2000-08-03	Duffryn Comp	Manchester	England	0711111112	2
10	Ishaan Harvey	Swansea	1971-08-26	St Michael's	Birmingham	England	0711111113	2
11	Laila Murray	Wrexham	1986-10-08	Gaer	Bristol	England	0711111114	2
12	Leonidas Jefferson	Monmouth	1974-04-11	Pill	Sheffield	England	0711111115	2
13	Joy Rush	Bridgend	1993-01-24	John Frost High	Glasgow	Scotland	0711111116	2

PositionCode	Position
1	Researcher
2	Student

Person_Id	Completed_Course	Grade
8	History	A
8	Maths	B
3	IT	A
3	Maths	C
10	Geography	D
11	English	C
12	IT	A
12	Science	D
13	Maths	C
13	IT	D
13	English	C

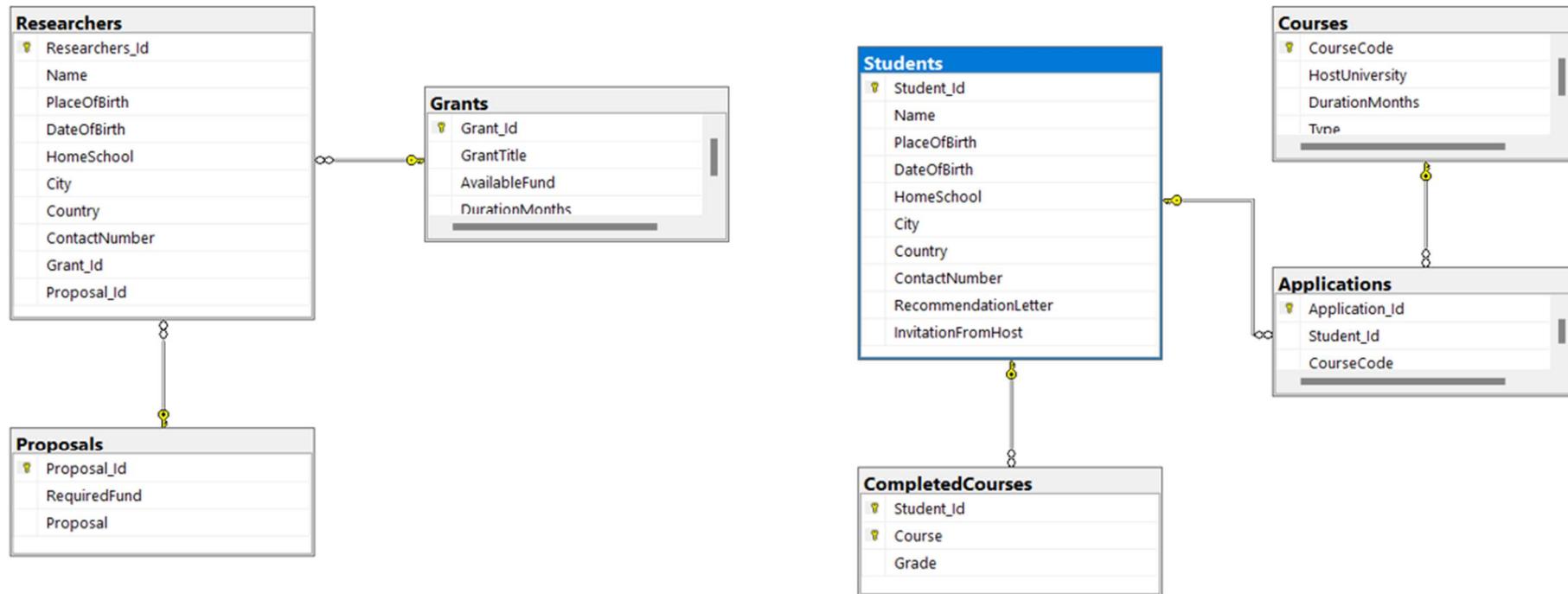
Grant_id	Grant title	Available fund	Duration
1	Short term	20K	6 months
2	ESRC	250K	24 months
3	Welch Government	150K	12 months

Course code	Course title	Host University	Duration	Type
CS1022	Computer Security	Cardiff Met University	36	BSc
DS2221	Data Science	Swansea University	12	MSc
BI0011	Bioinformatics	MIT	36	PhD

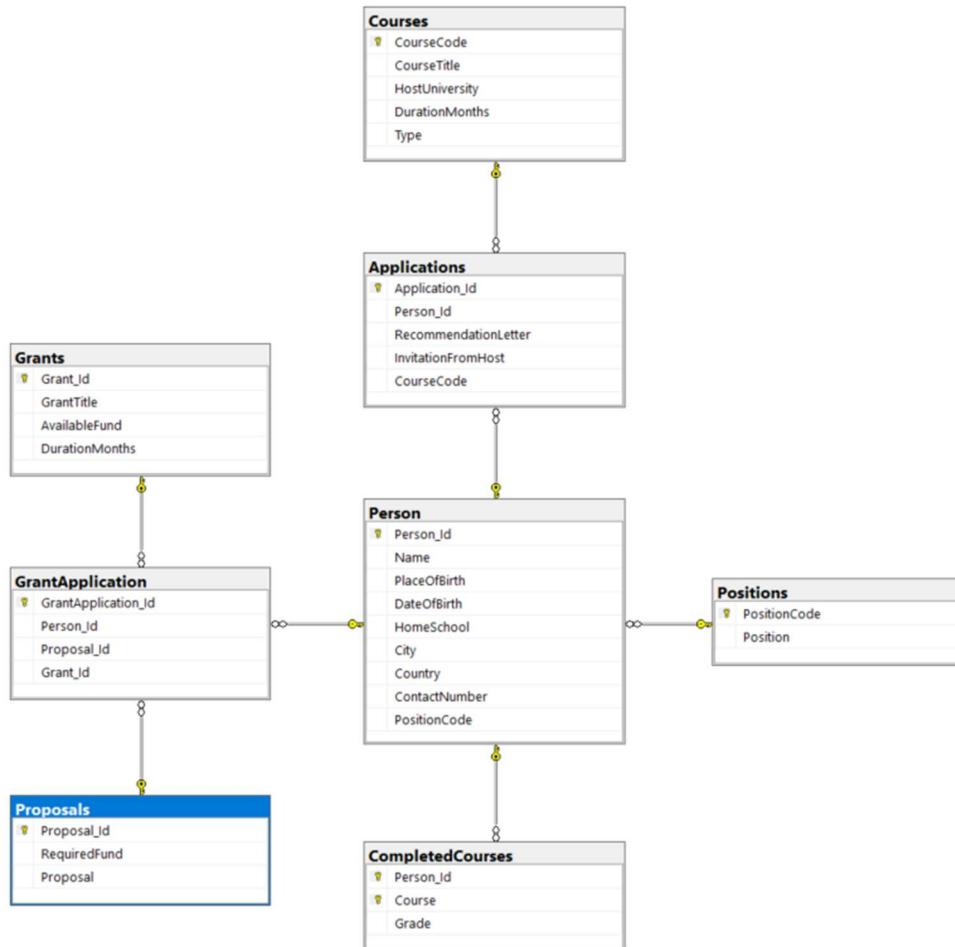
Proposal_Id	Required fund	Proposal
1	1000	Intro
2	200	Background
3	600	Risks
4	800	Methods

## Schema

After the first attempt at normalization, I created a schema which highlighted the disconnect between the two sets of data:



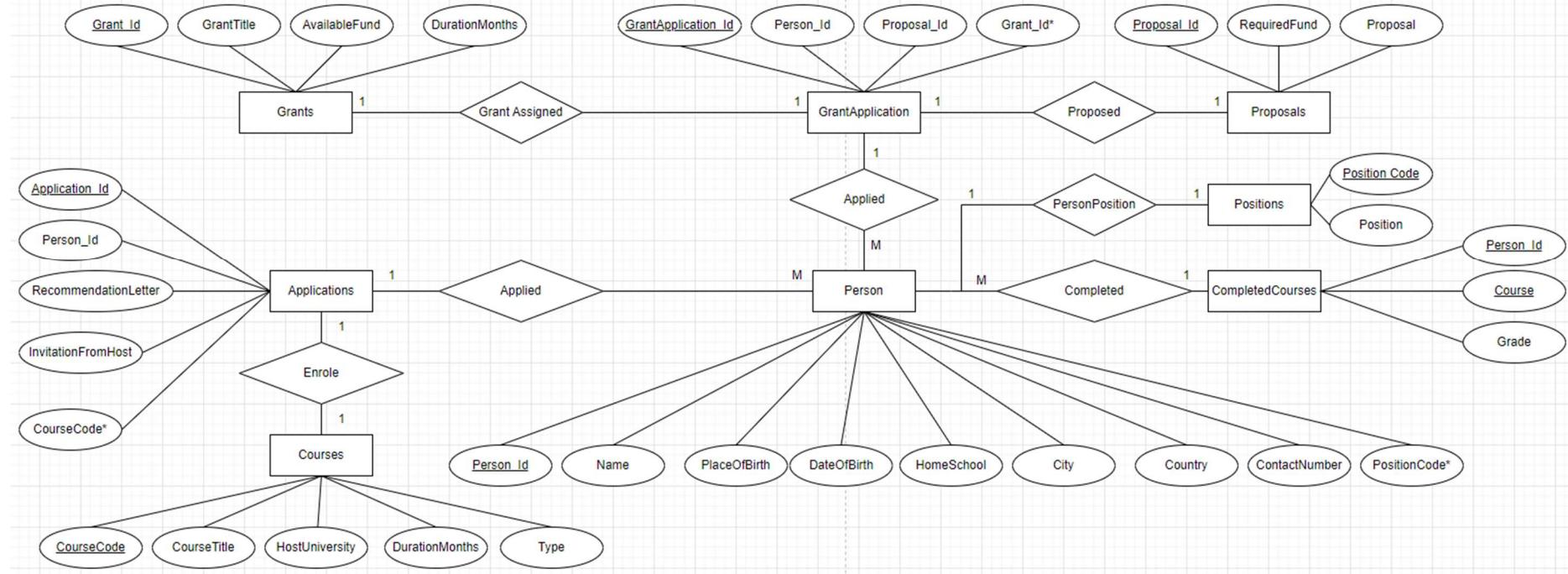
Once the normalization had been improved, I created a new schema which can be found in the file named ‘UniversitySchema.png’.



## Entity Relation Diagram

The entity relation diagram can be opened using diagrams.net (<https://www.diagrams.net/>) and can be found in the file named ‘ERD-EntityRelationshipDiagram.drawio’.

**Entity Relationship Diagram - Chen Notation**



## Data dictionary

To assist with the database design a data dictionary has been created and is named '**Data Dictionary.xlsx**'.

Entity Name	Entity Description	Column Name	Column Description	Data Type	Length	Primary Key	Foreign Key	Nullable	Unique	Example
Person	A person that can be a student or a researcher depending on the PositionCode	Person_Id	Primary key for this table	INT	10	TRUE	FALSE	FALSE	TRUE	1
		Name	The person's name	NVARCHAR	255	FALSE	FALSE	FALSE	FALSE	Joe Bloggs
		PlaceOfBirth	The person's place of birth	NVARCHAR	255	FALSE	FALSE	FALSE	FALSE	Cardiff
		DateOfBirth	The person's date of birth	DATE		FALSE	FALSE	FALSE	FALSE	2000-12-08
		HomeSchool	The person's home school	NVARCHAR	255	FALSE	FALSE	FALSE	FALSE	Llandaff High School
		City	The person's city	NVARCHAR	255	FALSE	FALSE	FALSE	FALSE	Cardiff
		Country	The person's country	NVARCHAR	255	FALSE	FALSE	TRUE	FALSE	Wales
		ContactNumber	The person's contact number	NVARCHAR	255	FALSE	FALSE	TRUE	TRUE	07111111112
		PositionCode	The person's position code	INT	10	FALSE	TRUE	FALSE	FALSE	1
Entity Name	Entity Description	Column Name	Column Description	Data Type	Length	Primary Key	Foreign Key	Nullable	Unique	Example
Positions	The valid positions that a person can be	PositionCode	Primary key for this table	INT	10	TRUE	FALSE	FALSE	TRUE	1
		Position		NVARCHAR	255	FALSE	FALSE	FALSE	TRUE	Student
Entity Name	Entity Description	Column Name	Column Description	Data Type	Length	Primary Key	Foreign Key	Nullable	Unique	Example
CompletedCourses	Courses that a student person has completed	Person_Id	Composite Primary key for this table	INT	10	TRUE	FALSE	FALSE	TRUE	1
		Course	for this table and the completed Course name	NVARCHAR	255	TRUE	FALSE	FALSE	TRUE	History
		Grade	the person in the course	NCHAR	1		FALSE	FALSE	FALSE	A
Entity Name	Entity Description	Column Name	Column Description	Data Type	Length	Primary Key	Foreign Key	Nullable	Unique	Example
Courses	Courses that a student person can apply for	CourseCode	Primary key for this table and the course code	NCHAR	6	TRUE	FALSE	FALSE	TRUE	CS1022
		CourseTitle	The course title	NVARCHAR	255	FALSE	FALSE	FALSE	TRUE	Computer Security
		HostUniversity	The host university of the course	NVARCHAR	255	FALSE	FALSE	FALSE	FALSE	Cardiff Met University
		DurationMonths	The duration of the course	INT	10	FALSE	FALSE	FALSE	FALSE	36
		Type	The type of course qualification	NCHAR	3	FALSE	FALSE	FALSE	FALSE	BSc

Entity Name	Entity Description	Column Name	Column Description	Data Type	Length	Primary Key	Foreign Key	Nullable	Unique	Example
Applications	student persons that contains recommendations and initiations and the course code	Application_Id	Primary key for this table	INT	10	TRUE	FALSE	FALSE	TRUE	1
		Person_Id	The primary key from the person table	INT	10	FALSE	FALSE	FALSE	FALSE	8
		RecommendationLetter	Y/N whether a recommendation was received	NCHAR	1	FALSE	FALSE	FALSE	FALSE	Y
		InvitationFromHost	Y/N whether a invitation was received	NCHAR	1	FALSE	FALSE	FALSE	FALSE	Y
		CourseCode	The primary key from the Course table and the course code	NCHAR	6	FALSE	TRUE	TRUE	FALSE	CS1022

Entity Name	Entity Description	Column Name	Column Description	Data Type	Length	Primary Key	Foreign Key	Nullable	Unique	Example
Grants	Grants that are available to researcher	Grant_Id	The primary key of this table	INT	10	TRUE	FALSE	FALSE	TRUE	2
		GrantTitle	The title of the grant	NVARCHAR	255		FALSE	FALSE	TRUE	ESRC
		AvailableFund	The amount of funds available	INT	10		FALSE	FALSE	FALSE	250K
		DurationMonths	The duration of the grant	INT	10		FALSE	FALSE	FALSE	24 months

Entity Name	Entity Description	Column Name	Column Description	Data Type	Length	Primary Key	Foreign Key	Nullable	Unique	Example
Proposals	Proposals available to researchers	Proposal_Id	This table's primary Key	INT	10	TRUE	FALSE	FALSE	TRUE	2
		RequiredFund	The required fund	INT	10	FALSE	FALSE	FALSE	FALSE	200
		Proposal	The proposal type	NVARCHAR	255	FALSE	FALSE	FALSE	TRUE	Background

Entity Name	Entity Description	Column Name	Column Description	Data Type	Length	Primary Key	Foreign Key	Nullable	Unique	Example
GrantApplication	Grant applications that have been made by researcher persons included the id of the proposal	GrantApplication_Id	Table primary key	INT	10	TRUE	FALSE	FALSE	TRUE	2
		Person_Id	Primary id from the persons table	INT	10	FALSE	TRUE	FALSE	FALSE	2
		Proposal_Id	Primary id from the proposal table	INT	10	FALSE	TRUE	FALSE	FALSE	2
		Grant_Id	Primary id from the grant table	INT	10	FALSE	TRUE	FALSE	FALSE	2

## Creation Script

The database creation script is named '**CreateTables.sql**'.

Run the 'CREATE DATABASE' in isolation and then comment the command out.

Remove the comment characters for the 'USE' command and execute the query.

```
--CREATE DATABASE University;
--USE University;

CREATE TABLE Positions(
    PositionCode int PRIMARY KEY,
    Position NVARCHAR(255)
);

CREATE TABLE Person(
    Person_Id int PRIMARY KEY,
    "Name" NVARCHAR(255),
    PlaceOfBirth NVARCHAR(255),
    DateOfBirth Date,
    HomeSchool NVARCHAR(255),
    City NVARCHAR(255),
    Country NVARCHAR(255),
    ContactNumber NVARCHAR(255),
    PositionCode int NOT NULL FOREIGN KEY REFERENCES Positions(PositionCode)
);

CREATE TABLE CompletedCourses(
    Person_Id int,
    Course NVARCHAR(255),
    Grade NCHAR(1),
    CONSTRAINT PK_CompletedCourses PRIMARY KEY (Person_Id, Course),
    CONSTRAINT FK_Person_Id FOREIGN KEY (Person_Id) REFERENCES Person(Person_Id)
);

CREATE TABLE Courses(
    CourseCode nchar(6) NOT NULL PRIMARY KEY,
    CourseTitle NVARCHAR(255),
    HostUniversity NVARCHAR(255),
    DurationMonths int,
    "Type" NCHAR(3)
);

CREATE TABLE Applications (
    Application_Id int NOT NULL PRIMARY KEY,
    Person_Id int NOT NULL FOREIGN KEY REFERENCES Person(Person_Id),
    RecommendationLetter NCHAR(1),
    InvitationFromHost NCHAR(1),
    CourseCode nchar(6) NOT NULL FOREIGN KEY REFERENCES Courses(CourseCode)
);

CREATE TABLE Grants(
    Grant_Id int PRIMARY KEY,
    GrantTitle NVARCHAR(255),
    AvailableFund int,
    DurationMonths int
);

CREATE TABLE Proposals(
    Proposal_Id int PRIMARY KEY,
    RequiredFund int,
    Proposal NVARCHAR(255)
);

CREATE TABLE GrantApplication(
    GrantApplication_Id int PRIMARY KEY,
    Person_Id int,
    Proposal_Id int,
    Grant_Id int,
    CONSTRAINT FK_Persons_Id FOREIGN KEY (Person_Id) REFERENCES Person(Person_Id),
    CONSTRAINT FK_Grant_Id FOREIGN KEY (Grant_Id) REFERENCES Grants(Grant_Id),
    CONSTRAINT FK_Proposal_Id FOREIGN KEY (Proposal_Id) REFERENCES Proposals(Proposal_Id)
);
```

## Insert Data Script

The script to insert test data is named '**InsertDataIntoTables.sql**'.

```
USE [University]
GO

INSERT INTO [dbo].[Positions]
VALUES
(1,'Student'),
(2,'Researcher')
GO

INSERT INTO [dbo].[Courses]
VALUES
('CS1022','Computer Security','Cardiff Met University',36,'BSc'),
('DS2221','Data Science','Swansea University',12,'MSc'),
('BI0011','Bioinformatics','MIT',36,'PhD')
GO

INSERT INTO [dbo].[Grants]
VALUES
(1,'Short term',20000,6),
(2,'ESRC',250000,24),
(3,'Welsh Government',150000,12);
GO

INSERT INTO [dbo].[Proposals]
VALUES
(1,1000,'Intro'),
(2,200,'Background'),
(3,600,'Risks'),
(4,800,'Methods');
GO

INSERT INTO [dbo].[Person]
VALUES
(1,'Amaya Hebert','Cardiff','1978-04-29','John Frost High','London','England','0722222222',1),
(2,'Gavin Stephenson','Newport','1978-11-20','Duffryn Comp','Manchester','England','0722222223',1),
(3,'Chase Lewis','Swansea','2000-01-30','St Michaels','Birmingham','England','0722222224',1),
(4,'Cael Boyer','Wrexham','1996-08-07','John Frost High','London','England','0722222225',1),
(5,'Glenn McCann','Monmouth','1981-05-06','Duffryn Comp','Manchester','England','0722222226',1),
(6,'Emilia Peters','Bridgend','1992-10-15','St Michaels','Birmingham','England','0722222227',1),
(7,'Jeremiah Blackburn','Carmarthen','1966-09-14','St Brides','Madrid','Spain','0722222228',1),
(8,'Wyatt Page','Cardiff','1995-05-26','John Frost High','London','England','0711111111',2),
(9,'Karlie Decker','Newport','2000-08-03','Duffryn Comp','Manchester','England','0711111112',2),
(10,'Ishaan Harvey','Swansea','1971-08-26','St Michaels','Birmingham','England','0711111113',2),
(11,'Laila Murray','Wrexham','1986-10-08','Gaer','Bristol','England','0711111114',2),
(12,'Leonidas Jefferson','Monmouth','1974-04-11','Pill','Sheffield','England','0711111115',2),
(13,'Joy Rush','Bridgend','1993-01-24','John Frost High','Glasgow','Scotland','0711111116',2)
GO

INSERT INTO [dbo].[CompletedCourses]
VALUES
(8,'History','A'),
(8,'Maths','B'),
(9,'IT','A'),
(9,'Maths','C'),
(10,'Geography','D'),
(11,'English','C'),
(12,'IT','A'),
(12,'Science','D'),
(13,'Maths','C'),
(13,'IT','D'),
(13,'English','C')
GO

INSERT INTO [dbo].[Applications]
VALUES
(1,8,'Y','Y','CS1022'),
(2,8,'Y','Y','CS1022'),
(3,9,'Y','Y','DS2221'),
(4,9,'Y','Y','DS2221'),
(5,10,'N','Y','BI0011'),
(6,11,'N','Y','CS1022'),
(7,12,'Y','Y','DS2221'),
(8,12,'Y','Y','DS2221'),
(9,13,'Y','Y','BI0011'),
(10,13,'Y','Y','BI0011'),
(11,13,'Y','Y','BI0011')
GO

INSERT INTO [dbo].[GrantApplication]
VALUES
(1,1,1,1),
(2,2,2,2),
(3,3,3,3),
(4,4,4,1),
(5,5,1,2),
(6,6,2,3),
(7,7,3,1)
GO
```

## Various Query Examples

To test various queries and how the tables are connected I created a file containing various queries. This file is called '**VariousQueries.sql**'.

```
--delete queries
DELETE FROM [University].[dbo].[GrantApplication] WHERE Person_Id = 1;
DELETE FROM [University].[dbo].[Person] WHERE Person_Id = 1;

DELETE FROM [University].[dbo].[Applications] WHERE Person_Id = 10;
DELETE FROM [University].[dbo].[Person] WHERE Person_Id = 10;

--Aggregate queries
SELECT Country
FROM [University].[dbo].[Person]
GROUP BY Country;
GO

SELECT COUNT(DISTINCT City) AS 'Number of unique cities'
FROM [University].[dbo].[Person]
GO

SELECT AVG(AvailableFund) AS 'Average available fund'
FROM [University].[dbo].[Grants]
GO

--Wildcard Selections
SELECT *
FROM [University].[dbo].[Person]
WHERE [Name] LIKE '%ecker' or City LIKE '%ming%'
GO

--Inner Joins
SELECT [Name], HomeSchool,Course,Grade
FROM [University].[dbo].[Person]
INNER JOIN [University].[dbo].[CompletedCourses]
ON [University].[dbo].[Person].Person_Id = [University].[dbo].[CompletedCourses].Person_Id;

SELECT [Name] As 'Researcher',GrantTitle As 'Grant Applied For',AvailableFund As 'Fund Available'
FROM [University].[dbo].[Person]
INNER JOIN [University].[dbo].[GrantApplication]
ON [University].[dbo].[Person].Person_Id = [University].[dbo].[GrantApplication].Person_Id
INNER JOIN [University].[dbo].[Grants]
ON [University].[dbo].[GrantApplication].Grant_Id = [University].[dbo].[Grants].Grant_Id;

--Left out join
SELECT [Name],Position
FROM [University].[dbo].[Person]
LEFT OUTER JOIN [University].[dbo].[Positions]
ON [University].[dbo].[Person].PositionCode = [University].[dbo].[Positions].PositionCode

--returns all even if there is no course title
SELECT Distinct([Name]),CourseTitle
FROM [University].[dbo].[Person]
LEFT OUTER JOIN [University].[dbo].[Applications]
ON [University].[dbo].[Person].Person_Id = [University].[dbo].[Applications].Person_Id
LEFT OUTER JOIN [University].[dbo].[Courses]
ON [University].[dbo].[Applications].CourseCode = [University].[dbo].[Courses].CourseCode
```

```
--Right joins (does not return any without course title
SELECT Distinct([Name]),CourseTitle
FROM [University].[dbo].[Person]
RIGHT JOIN [University].[dbo].[Applications]
ON [University].[dbo].[Person].Person_Id = [University].[dbo].[Applications].Person_Id
RIGHT OUTER JOIN [University].[dbo].[Courses]
ON [University].[dbo].[Applications].CourseCode = [University].[dbo].[Courses].CourseCode

SELECT Distinct([Name]),RecommendationLetter,InvitationFromHost,CourseTitle,Course,Grade
FROM [University].[dbo].[Person]
RIGHT JOIN [University].[dbo].[Applications]
ON [University].[dbo].[Person].Person_Id = [University].[dbo].[Applications].Person_Id
RIGHT JOIN [University].[dbo].[Courses]
ON [University].[dbo].[Applications].CourseCode = [University].[dbo].[Courses].CourseCode
RIGHT JOIN [University].[dbo].[CompletedCourses]
ON [University].[dbo].[Person].Person_Id = [University].[dbo].[CompletedCourses].Person_Id

SELECT [Name],GrantTitle
FROM [University].[dbo].[Person]
RIGHT JOIN [University].[dbo].[GrantApplication]
ON [University].[dbo].[Person].Person_Id = [dbo].[GrantApplication].Person_Id
RIGHT JOIN [University].[dbo].[Grants]
ON [University].[dbo].[GrantApplication].Grant_Id = [dbo].[Grants].Grant_Id
```