

Data Management & Database Design

ASSIGNMENT 4: Normalization

Team Name: MStudents (Team: 13)

Team Members: Renuka Huddar (002750433)

Darshana Rokade (002705752)

Table 1 : Student

Schema:

	Field	Type	Null	Key	Default	Extra
►	StudentID	int	NO	PRI	NULL	
	FirstName	varchar(50)	YES		NULL	
	LastName	varchar(50)	YES		NULL	
	Gender	varchar(5)	YES		NULL	
	DeptName	varchar(100)	YES		NULL	
	College	varchar(70)	YES		NULL	
	DeptID	int	NO	MUL	NULL	
	FeePayment	varchar(5)	YES		NULL	
	CreditsEarned	int	YES		NULL	
	IntakeSemester	varchar(10)	YES		NULL	
	IntakeYear	int	YES		NULL	

Sample Data:

	StudentID	FirstName	LastName	Gender	DeptName	College	DeptID	FeePayment	CreditsEarned	IntakeSemester	IntakeYear
►	101	Antwan	Davies	M	Information Systems	College of Engineering	207	No	16	Fall	2021
	102	Keshawn	Frey	F	Health Informatics	Bouvé College of Health Sciences	203	No	16	Fall	2021
	103	Isla	Andrade	M	JD	School of Law	218	Yes	24	Spring	2021
	104	Lellani	Wilkerson	F	MS in Information Design and Data Visualization	College of Arts, Media and Design	205	Yes	17	Spring	2022
	105	Ivan	Ochoa	F	FlexJD	School of Law	219	Yes	16	Spring	2022
	106	Emilia	Guerrero	F	Exercise Science Certificate for Clinicians	Bouvé College of Health Sciences	202	No	16	Fall	2021
	107	John	zhho	M	Online LLM	School of Law	220	No	16	Spring	2022
	108	Mark	Orozco	F	Human Movement and Rehabilitation Sciences	Bouvé College of Health Sciences	204	Yes	16	Spring	2022
	109	Misael	Cowan	F	Medicinal Chemistry and Drug Discovery	Bouvé College of Health Sciences	216	No	17	Spring	2022
	110	Jeremiah	Hobbs	F	Exercise Science Certificate for Clinicians	Bouvé College of Health Sciences	202	No	16	Spring	2022
	111	Landyn	Kelly	F	Health Informatics	Bouvé College of Health Sciences	203	Yes	16	Fall	2021
	112	Wendy	Rojas	M	Human Movement and Rehabilitation Sciences	Bouvé College of Health Sciences	204	Yes	24	Spring	2021
	113	Emanuel	Powers	M	Medicinal Chemistry and Drug Discovery	Bouvé College of Health Sciences	216	Yes	16	Fall	2021

1NF:

The table is already in 1NF as it has StudentID as Primary Key. There are no multivalued attributes, each attribute has unique values with no repeating groups.

2NF:

1NF condition is satisfied.

This table had partial dependencies(IntakeYear), which is not accepted at 2NF. As a result, we created 2 separate tables: Fall & Spring, which have StudentID and IntakeYear fields. So the table is now in 2NF.

SQL Query:

```
ALTER TABLE Student DROP COLUMN IntakeSemester, DROP COLUMN IntakeYear;
```

StudentID	FirstName	LastName	Gender	DeptID	FeePayment	CreditsEarned
101	Antwan	Davies	M	207	No	16
102	Keshawn	Frey	F	203	No	16
103	Isla	Andrade	M	218	Yes	24
104	Leilani	Wilkerson	F	205	Yes	17
105	Ivan	Ochoa	F	219	Yes	16
106	Emilia	Guerrero	F	202	No	16
107	John	zhho	M	220	No	16
108	Mark	Orozco	F	204	Yes	16
109	Misael	Cowan	F	216	No	17
110	Jeremiah	Hobbs	F	202	No	16

3NF:

2NF conditions are satisfied.

Since *collegeName* was dependent on *deptName*, and *deptName* was dependent on *deptID* and *deptID* is dependent on *StudentID(PK)*, therefore we removed the *collegename* column from the *student* table. Since *collegename* is already present in the *college* table, we can fetch details of students by joining *student* and *college* table.

SQL Query:

```
ALTER TABLE Student DROP College;
```

```
ALTER TABLE Student DROP DeptName;
```

```
ALTER TABLE StudentJob ADD CONSTRAINT FOREIGN KEY (StudentID) REFERENCES Student (StudentID);
```

StudentID	FirstName	LastName	Gender	DeptID	FeePayment	CreditsEarned
101	Antwan	Davies	M	207	No	16
102	Keshawn	Frey	F	203	No	16
103	Isla	Andrade	M	218	Yes	24
104	Leilani	Wilkerson	F	205	Yes	17
105	Ivan	Ochoa	F	219	Yes	16
106	Emilia	Guerrero	F	202	No	16
107	John	zhho	M	220	No	16
108	Mark	Orozco	F	204	Yes	16
109	Misael	Cowan	F	216	No	17
110	Jeremiah	Hobbs	F	202	No	16

Table 2 : Fall

Schema:

	Field	Type	Null	Key	Default	Extra
►	StudentID	int	NO	MUL	NULL	
	IntakeYear	int	NO		NULL	
	ID	int	NO	PRI	NULL	

Sample Data:

	StudentID	IntakeYear	ID
►	101	2021	1
	102	2021	2
	106	2021	3
	111	2021	4
	113	2021	5
	115	2021	6
	116	2021	7
	124	2021	8
	125	2021	9
	126	2021	10

SQL Query:

ALTER TABLE Spring ADD CONSTRAINT FOREIGN KEY (StudentID) REFERENCES Student (StudentID);

1NF:

The table is already in 1NF as it has ID as Primary Key. There are no multivalued attributes, each attribute has unique values with no repeating groups.

2NF:

1NF condition is satisfied.

The table has no partial dependencies and calculated data. Therefore, the table is in 2NF.

3NF:

2NF condition is satisfied.

The table has no transitive dependencies. Therefore, the table is in 3NF.

Table 3 : Spring

Schema:

	Field	Type	Null	Key	Default	Extra
►	StudentID	int	NO	MUL	NULL	
	IntakeYear	int	NO		NULL	
	ID	int	NO	PRI	NULL	

Sample Data:

	StudentID	IntakeYear	ID
▶	103	2021	1
	104	2022	2
	105	2022	3
	107	2022	4
	108	2022	5
	109	2022	6
	110	2022	7
	112	2021	8
	114	2022	9
	117	2022	10

Spring 66 ×

SQL Query:

```
ALTER TABLE Fall ADD CONSTRAINT FOREIGN KEY (StudentID) REFERENCES Student (StudentID);
```

1NF:

The table is already in 1NF as it has ID as Primary Key. There are no multivalued attributes, each attribute has unique values with no repeating groups.

2NF:

1NF condition is satisfied.

The table has no partial dependencies and calculated data. Therefore, the table is in 2NF.

3NF:

2NF condition is satisfied.

The table has no transitive dependencies. Therefore, the table is in 3NF.

Table 4 : StudentJob

Schema:

	Field	Type	Null	Key	Default	Extra
▶	StudentID	int	NO	UNI	NULL	
	JobID	int	NO		NULL	
	ID	int	NO	PRI	NULL	

Sample Data:

	StudentID	JobID	ID
▶	101	2084	1
	102	2591	2
	103	1370	3
	104	690	4
	105	1775	5
	106	1283	6
	107	2642	7
	108	1141	8
	109	1270	9

1NF:

The table is already in 1NF as it has ID as Primary Key. There are no multivalued attributes, each attribute has unique values with no repeating groups.

SQL Query:

```
CREATE TABLE StudentJob (StudentID INT NOT NULL UNIQUE, JobID int NOT NULL, ID INT NOT NULL PRIMARY KEY);
```

2NF:

1NF condition is satisfied.

The table has no partial dependencies and calculated data. Therefore, the table is in 2NF.

3NF:

2NF condition is satisfied.

The table has no transitive dependencies. Therefore, the table is in 3NF.

Table 5 : StudentOnCampus

Schema:

	Field	Type	Null	Key	Default	Extra
▶	StudentID	int	NO	MUL	NULL	
	JobID	int	YES		NULL	
	ID	int	NO	PRI	NULL	

Sample Data:

	StudentID	JobID	ID
▶	108	324	1
	112	319	2
	113	314	3
	115	321	4
	119	318	5
	126	314	6
	128	309	7
	131	303	8
	133	312	9

StudentOnCampus 63 x

1NF:

The table is already in 1NF as it has ID as Primary Key. There are no multivalued attributes, each attribute has unique values with no repeating groups.

2NF:

1NF condition is satisfied.

The table has no partial dependencies and calculated data. Therefore, the table is in 2NF.

3NF:

2NF condition is satisfied.

The table has no transitive dependencies. Therefore, the table is in 3NF.

Table 6 : OnCampusEmployment

Schema:

	Field	Type	Null	Key	Default	Extra
▶	JobID	int	NO	PRI	NULL	
	Employer	varchar(100)	NO		NULL	
	Location	varchar(100)	NO		NULL	

Sample Data:

	JobID	Employer	Location
▶	301	Dunkin Donuts (Shillman Hall)	Shillman Hall
	302	Dunkin Donuts (Richards/Hayden Hall)	Richards/Hayden Hall
	303	Caffe Strega (ISEC)	ISEC
	304	Chartwells	Northeastern
	305	Reprographics (Ell Hall)	Ell Hall
	306	Wollastons	Northeastern
	307	Vanguard (Columbus Parking Garage)	Columbus Parking Garage
	308	B.Good (Marino Center)	Marino Center
	309	Cafe 716	Northeastern

OnCampusEmployment 66 x

1NF:

The table is already in 1NF as it has ID as Primary Key. There are no multivalued attributes, each attribute has unique values with no repeating groups.

2NF:

1NF condition is satisfied.

To remove partial dependencies, we created a new table: *OnCampusEmploymentStats*, which has *JobID*, *skillsRequired*, *SalaryInUSD*, *DurationInMonths*, *PostingDate*, *Deadline* and *Vacancy* as fields.

Now there are no partial dependencies and calculated data so it is in 2NF.

SQL Query:

```
CREATE TABLE OnCampusEmploymentStats (JobID INT NOT NULL, SkillsRequired  
VARCHAR(100) NOT NULL, SalaryInUSD INT NOT NULL,  
DurationInMonths VARCHAR(100), PostingDate DATETIME, Deadline DATETIME, Vacancy  
INT);
```

3NF:

2NF condition is satisfied.

To remove transitive dependencies, we removed columns: *SkillsRequired*, *SalaryInUSD*, *DurationInMonths*, *PostingDate*, *Deadline* and *Vacancy*.

SQL Query:

```
ALTER TABLE OnCampusEmployment DROP COLUMN SkillsRequired, DROP COLUMN  
SalaryInUSD, DROP COLUMN DurationInMonths, DROP COLUMN PostingDate, DROP  
COLUMN Deadline, DROP COLUMN Vacancy;
```

Now, the table is in 3NF.

Table 7 : OnCampusEmploymentStats

Schema:

	Field	Type	Null	Key	Default	Extra
►	ID	int	NO	PRI	NULL	
	JobID	int	NO	MUL	NULL	
	SkillsRequired	varchar(100)	NO		NULL	
	SalaryInUSD	int	NO		NULL	
	DurationInMonths	varchar(100)	YES		NULL	
	PostingDate	datetime	YES		NULL	
	Deadline	datetime	YES		NULL	
	Vacancy	int	YES		NULL	

Sample Data:

	ID	JobID	SkillsRequired	SalaryInUSD	DurationInMonths	PostingDate	Deadline	Vacancy
►	1	301	Fluency In English	10	4	2022-11-01 20:30:00	2022-12-01 20:30:00	2
	2	302	Fluency In English	10	4	2022-11-02 20:30:00	2022-12-02 20:30:00	4
	3	303	Fluency In English	12	6	2022-11-03 20:30:00	2022-12-03 20:30:00	8
	4	304	Fluency In English	11	4	2022-11-04 20:30:00	2022-12-04 20:30:00	12
	5	305	Fluency In English	13	4	2022-11-05 20:30:00	2022-12-05 20:30:00	15
	6	306	Fluency In English	12	4	2022-11-06 20:30:00	2022-12-06 20:30:00	10
	7	307	Fluency In English	13	6	2022-11-07 20:30:00	2022-12-07 20:30:00	4
	8	308	Fluency In English	12	4	2022-11-08 20:30:00	2022-12-08 20:30:00	5
	9	309	Fluency In English	10	4	2022-11-09 20:30:00	2022-12-09 20:30:00	6

OnCampusEmploymentStats 73 x

1NF:

The table is already in 1NF as it has ID as Primary Key. There are no multivalued attributes, each attribute has unique values with no repeating groups.

2NF:

1NF condition is satisfied.

The table has no partial dependencies and calculated data. Therefore, the table is in 2NF.

3NF:

2NF condition is satisfied.

The table has no transitive dependencies. Therefore, the table is in 3NF.

Table 8 : College

Schema:

	Field	Type	Null	Key	Default	Extra
►	CollegeName	varchar(100)	NO		NULL	
	DeptName	varchar(100)	NO		NULL	
	DeptID	int	NO	PRI	NULL	
	DeanName	varchar(100)	NO		NULL	

Sample Data:

	CollegeName	DeptName	DeptID	DeanName
►	Bouvé College of Health Sciences	Applied Behavior Analysis (CAGS, MS, Grad Cer...	201	Carmen Sceppa
	Bouvé College of Health Sciences	Exercise Science Certificate for Clinicians	202	Carmen Sceppa
	Bouvé College of Health Sciences	Health Informatics	203	Carmen Sceppa
	Bouvé College of Health Sciences	Human Movement and Rehabilitation Sciences	204	Carmen Sceppa
	College of Arts, Media and Design	MS in Information Design and Data Visualization	205	Elizabeth Hudson
	College of Arts, Media and Design	MS in Game Science and Design	206	Elizabeth Hudson
	College of Engineering	Information Systems	207	Gregory D. Abowd
	College of Engineering	Pharmaceutical Engineering	208	Gregory D. Abowd
	College of Engineering	Telecommunication Networks	209	Gregory D. Abowd

1NF:

The table is already in 1NF as it has ID as Primary Key. There are no multivalued attributes, each attribute has unique values with no repeating groups.

2NF:

1NF condition is satisfied.

The table has no partial dependencies and calculated data. Therefore, the table is in 2NF.

3NF:

2NF condition is satisfied.

The table has no transitive dependencies. Therefore, the table is in 3NF.

Table 9 : InternshipsAndCoop

Schema:

	Field	Type	Null	Key	Default	Extra
►	JobID	int	NO	PRI	NULL	
	JobTitle	varchar(500)	YES		NULL	
	SalaryEstimateLower	int	YES		NULL	
	SalaryEstimateUpper	int	YES		NULL	
	JobDescription	varchar(1000)	YES		NULL	
	Rating	float	YES		NULL	
	CompanyName	varchar(500)	YES		NULL	
	Location	varchar(500)	YES		NULL	
	HeadQuaters	varchar(500)	YES		NULL	

Sample Data:

	JobID	JobTitle	SalaryEstimateLower	SalaryEstimateUpper	JobDescription	Rating	C
►	401	Data Analyst, Center on Immigration and Justic...	37000	66000	Are you eager to roll up your sleeves and harne...	3.2	Vi
	402	Quality Data Analyst	37000	66000	Overview Provides analytical and technical sup...	3.8	Vi
	403	Senior Data Analyst, Insights & Analytics Team ...	37000	66000	We are looking for a Senior Data Analyst w...	3.4	Si
	404	Data Analyst	37000	66000	Requisition NumberRR-0001939 Remote:Yes W...	4.1	C
	405	Reporting Data Analyst	37000	66000	ABOUT FANDUEL GROUP FanDuel Group is a w...	3.9	Fi
	406	Data Analyst	37000	66000	About Cubist Cubist Systematic Strategies is on...	3.9	Ph
	407	Business/Data Analyst (FP&A)	37000	66000	Two Sigma is a different kind of investment man...	4.4	Ti
	408	Data Science Analyst	37000	66000	Data Science Analyst Job Details Level Experie...	3.7	G

1NF:

The table is already in 1NF as it has ID as Primary Key. There are no multivalued attributes, each attribute has unique values with no repeating groups.

2NF:

1NF condition is satisfied.

The table has no partial dependencies and calculated data. Therefore, the table is in 2NF.

3NF:

2NF condition is satisfied.

The table has no transitive dependencies. Therefore, the table is in 3NF.

Table 10 : EventDetails

Schema:

	Field	Type	Null	Key	Default	Extra
►	EID	int	NO	PRI	NULL	
	Ename	varchar(100)	YES		NULL	
	EventHeadID	int	NO	MUL	NULL	
	Location	varchar(50)	YES		NULL	
	Purpose	varchar(1000)	YES		NULL	
	Dateandtime	datetime	YES		NULL	
	FareinUSD	int	YES		NULL	

Sample Data:

	EID	Ename	EventHeadID	Location	Purpose	Dateandtime	FareinUSD
►	501	Northeastern University Professor on Marketing...	756	Hayden Hall	Northeastern University Distinguished Professor...	2022-11-09 20:30:00	20
	502	Carry the Love: Northeastern University	456	Snell Engg center	Uniting students on campus for two days of wor...	2022-02-10 19:00:00	80
	503	Northeastern State University- College Jazz Night	1000	Snell Library	NSU College Jazz Combos night at LowDown!	2022-11-19 14:00:00	0
	504	Chinese Language and Culture Seminar	235	Online	Learn more about Chinese culture, Mandarin, a...	2022-11-11 23:00:00	0
	505	Demo Day Fall 2022	456	Ruggles	Demo Day is Northeastern University's largest c...	2022-11-09 20:30:00	0
	506	College Admission 101	214	Tremont Street	This presentation will help families separate coll...	2022-12-07 20:00:00	0
	507	International Education Week Kick-Off Event	1023	West Village	Drop by for some snacks and to find out more a...	2022-11-17 12:45:00	0
	508	Next Steps for Critical Infrastructure & Cyber S...	1257	Church Hill	Next Steps for Critical Infrastructure & Cyber S...	2022-11-26 10:00:00	60
	509	The Promise and Progress Towards A Green Ne...	689	Ell Hall	Join us for an informative discussion on the pro...	2022-11-18 11:00:00	23

1NF:

The table is already in 1NF as it has ID as Primary Key. There are no multivalued attributes, each attribute has unique values with no repeating groups.

2NF:

1NF condition is satisfied.

The table has no partial dependencies and calculated data. Therefore, the table is in 2NF.

3NF:

2NF condition is satisfied.

The table has no transitive dependencies. Therefore, the table is in 3NF.

Table 11 : StudentClubs

Schema:

	Field	Type	Null	Key	Default	Extra
►	ClubID	int	NO	PRI	HULL	
	ClubName	varchar(100)	YES		HULL	
	StudentID	int	NO	MUL	HULL	
	Contacts	varchar(100)	YES		HULL	
	Motive	varchar(1000)	YES		HULL	

Sample Data:

	ClubID	ClubName	StudentID	Contacts	Motive
►	601	Association for Computing Machinery	107	cbw@ccs.neu.edu	engage informative speakers, public forums, an...
	602	CS Student Association	128	khoury-prospect@northeastern.edu	connect students at Northeastern through the ...
	603	Code 4 Culture	130	c4cneu@gmail.com	social and networking club for Black and Latinx s...
	604	Code4Community	134	c4cneu@gmail.com	student organization that partners with non-pro...
	605	CS Mentoring Organisation	145	b.hescott@northeastern.edu	connects students across the discipline of comp...
	606	DATA Club	154	khoury@northeastern.edu	The DATA Club provides speaker events, works...
	607	FirstBYTE	241	e.strange@northeastern.edu	provides affordable and user-friendly resource...
	608	American Institute of Chemical Engineers	273	nuaiche@gmail.com	focuses around enriching the education of chem...
	609	Alliance for Diversity in Science and Engineering	284	northeasternadse@gmail.com	organizations that reach out to students and sci...

1NF:

The table is already in 1NF as it has ID as Primary Key. There are no multivalued attributes, each attribute has unique values with no repeating groups.

2NF:

1NF condition is satisfied.

The table has no partial dependencies and calculated data. Therefore, the table is in 2NF.

3NF:

2NF condition is satisfied.

The table has no transitive dependencies. Therefore, the table is in 3NF.