CSCI UA.0060 Spring 2025

Assignment 3 – Normalization

Deliverables

Upload a pdf with the answers to the questions below to the Github repository for this assignment.

Filenames should be FirstInitialLastName-Ass3.pdf. For example, John Smith's file would be JSmith-Ass3.pdf

Overall Requirements

This assignment will give you more practice on Normalization. There are three exercises. In each one, you must first show all the data in one table. This table might be in 1NF, 2NF or 3NF (unlikely!!) or might not be in a Normal Form at all.

For each exercise, make sure you show each step of the process, including both a graphical view of the data and the specific reasons why each piece of data changes from one step to the next. Don't just show what the table(s) look like at the end of the process.

In each iteration, make sure that fields that are part of a PK are shown in a bold underlined font. For the 3NF version, make sure that FKs are italicized (including FKs that are also part of a PK).

Specific Requirements

1. The following is a handwritten Parts Order Form for a local Supply company. You can assume that employees help any and all different customers, i.e. there is no special relationship between specific employees and customers. A cage code is the identifier of the cages (or shelves) that the inventory is stored in.

   The plan is to replace the manual processes at the company with an application and database.

   Answer the questions shown below:

Happy Supplies
Parts Warehouse

Customer Name:	Jeff Peterson	**Date:**	7/1/2024
Customer Number:	HG54587	**Time:**	10:30am
		Employee:	D. Harrison
Customer Type:	Consumer		

Part Number	Name	Type	Cage Code	Quantity Ordered	Unit Price
10654	Float Control	Plumbing	G413	4	12
10456	Modulator	Electrical	H433	3	7
10776	Hose Assembly	Plumbing	G413	7	9
10657	Float Assembly	Plumbing	G413	5	10

   a. **State any assumptions that you make about the data and the attributes shown.**

      customers are able to purchase unlimited amounts of each part, only one employee helps a customer

   b. **Describe and illustrate the process of normalizing the attributes shown in this form.**

      The table is already in 1NF with PKs customerNumber, date, time, and cageCode

1NF

customerNumber(PK)	customerName	customerType	date (PK)	time (PK)	employee	partNumber(PK)	name	type	cage code	quantity ordered	unit price
HG54587	Jeff Peterson	consumer	7/1/2024	10:30 AM	D. Harrison	10654	float control	plumbing	G413	4	12
HG54587	Jeff Peterson	consumer	7/1/2024	10:30 AM	D. Harrison	10456	modulator	electrical	H433	3	7
HG54587	Jeff Peterson	consumer	7/1/2024	10:30 AM	D. Harrison	10776	hose assembly	plumbing	G413	7	9
HG54587	Jeff Peterson	consumer	7/1/2024	10:30 AM	D. Harrison	10657	float assembly	plumbing	G413	5	10

Moving onto 2NF, the table is not already in 2NF because customerName and customerType only depend on CustomerNumber, employee does not depend on anything, name, type, unitPrice and cageCode only depend on partNumber

2NF

customer		
customerNumber(PK)	customerName	customerType
HG54587	Jeff Peterson	consumer

part					
partNumber(PK)	name	type	cage code	quantity ordered	unit price
10654	float control	plumbing	G413	4	12
10456	modulator	electrical	H433	3	7
10776	hose assembly	plumbing	G413	7	9
10657	float assembly	plumbing	G413	5	10

customerNumber(PK)	date (PK)	time (PK)	partNumber(PK)	employee	quantity ordered
HG54587	7/1/2024	10:30 AM	10654	D. Harrison	4
HG54587	7/1/2024	10:30 AM	10456	D. Harrison	3
HG54587	7/1/2024	10:30 AM	10776	D. Harrison	7
HG54587	7/1/2024	10:30 AM	10657	D. Harrison	5

This table is also at 3NF since there are no transitive dependencies present

3NF

customer		
customerNumber(PK)	customerName	customerType
HG54587	Jeff Peterson	consumer

part					
partNumber(PK)	name	type	cage code	quantity ordered	unit price
10654	float control	plumbing	G413	4	12
10456	modulator	electrical	H433	3	7
10776	hose assembly	plumbing	G413	7	9
10657	float assembly	plumbing	G413	5	10

order					
customerNumber(PK, FK)	date (PK)	time (PK)	partNumber(PK, FK)	employee	quantity ordered
HG54587	7/1/2024	10:30 AM	10654	D. Harrison	4
HG54587	7/1/2024	10:30 AM	10456	D. Harrison	3
HG54587	7/1/2024	10:30 AM	10776	D. Harrison	7
HG54587	7/1/2024	10:30 AM	10657	D. Harrison	5

c. **Identify the primary keys at each step of the normalization process and both the primary and foreign keys in your 3NF relations.**

The primary keys at the initial form and 1NF are customerNumber, date, time, partNumber

The primary keys at 2NF and 3NF customerNumber for customer table, partNumber for part table, and customerNumber, date, time, and partNumber for order table

The foreign keys for 3NF are customerNumber and partNumber

d. **Ensure that the 3NF relations are named.**

customer

customerNumber(PK)	customerName	customerType
HG54587	Jeff Peterson	consumer

part

partNumber(PK)	name	type	cage code	quantity ordered	unit price
10654	float control	plumbing	G413	4	12
10456	modulator	electrical	H433	3	7
10776	hose assembly	plumbing	G413	7	9
10657	float assembly	plumbing	G413	5	10

1:M

1:M

order

customerNumber(PK, FK)	date (PK)	time (PK)	partNumber(PK, FK)	employee	quantity ordered
HG54587	7/1/2024	10:30 AM	10654	D. Harrison	4
HG54587	7/1/2024	10:30 AM	10456	D. Harrison	3
HG54587	7/1/2024	10:30 AM	10776	D. Harrison	7
HG54587	7/1/2024	10:30 AM	10657	D. Harrison	5

2. The data shown below is used by the Panacea Mental Health Corporation to track its therapists. Therapists may work at a number of different branches, but they only see patients at one specific branch on any given day. A patient is given an appointment at a specific time and date at a particular branch with one therapist. Patients may have multiple appointments in any given day and with multiple different therapists.

staffNo	therapistName	patNo	patName	appointment		branchNo
				date	time	
S1011	Fred Smith	P100	Lily White	9/12/2022	10:00	M15
S1011	Fred Smith	P105	Jill Baker	9/12/2022	12:00	M15
S1024	Heidi Pierce	P108	Andy McKee	9/12/2022	10:00	Q10
S1024	Heidi Pierce	P108	Andy McKee	9/14/2022	14:00	Q10
S1032	Richard Levin	P105	Jill Baker	9/14/2022	16:30	M15
S1032	Richard Levin	P110	Jimmy Winter	9/15/2022	18:00	B13

a. State any assumptions you make about the data and the attributes shown in this table.

Patients only are only registered under one branch

b. Describe and illustrate the process of normalizing the table to 3NF relations.

In order to set the table to 1NF, the appointment column must be split into two columns since two values are held within the same column.

1NF

staffNo(PK)	therapistName	patNo(PK)	patName	branchNo	date (PK)	time(PK)
S1011	Fred Smith	P100	Lily White	M15	9/12/2022	10:00:00
S1011	Fred Smith	P105	Jill Baker	M15	9/12/2022	12:00:00
S1024	Heidi Pierce	P108	Andy Mckee	Q10	9/12/2022	10:00:00
S1024	Heidi Pierce	P108	Andy Mckee	Q10	9/14/2022	14:00:00
S1032	Richard Levin	P105	Jill Baker	M15	9/14/2022	16:30:00
S1032	Richard Levin	P110	Jimmy Winter	B13	9/15/2022	18:00:00

To set the table to 2NF, we must ensure that all non PK fields are dependent on all PK fields. patientName is only dependent on patNo and therapistName is only dependent on staffNo, so they must be separated into their own tables.

2NF

therapists	
StaffNo (PK)	therapistName
S1011	Fred Smith
S1024	Heidi Pierce
S1032	Richard Levin

patients	
patNo (PK)	patName
P100	Lily White
P105	Jill Baker
P108	Andy Mckee
P110	Jimmy Winter

appointment				
staffNo (PK, FK)	patNo (PK, FK)	date (PK)	time (PK)	branchNo
S1011	P100	9/12/2022	10:00	M15
S1011	P105	9/12/2022	12:00	M15
S1024	P108	9/12/2022	10:00	Q10
S1024	P108	9/14/2022	14:00	Q10
S1032	P105	9/14/2022	16:30	M15
S1032	P110	9/15/2022	18:00	B13

When we accounted for non-PK dependencies, we were able to also eliminate transitive dependencies, so the table is also in 3NF

c. Identify the primary keys at each step of the normalization process and both the primary and foreign keys in your 3NF relations.

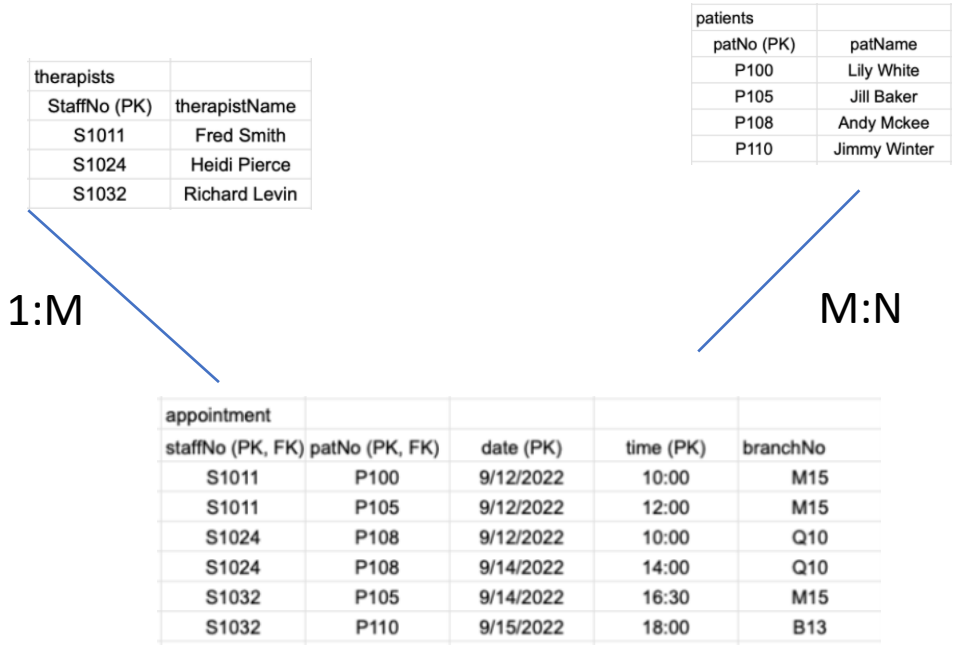
The primary keys are the initial table is staffNo, patNo, and appointment

The primary keys in 1NF are staffNo, patNo, date and time

The primary keys in 2NF and 3NF are staffNo in the therapist table, patNo in the patient table, and staffNo, patNo, date, and time in the appointment table

In 3NF, staffNo and patNo are also foreign keys since they reference the therapist and patient tables respectively

d. Ensure that the 3NF relations are named.



3. The Maid Better temp agency supplies help to Event Management companies within the New York area . Below is the data that the company uses to track employee hours against different contracts. The Employee Number (eNo) is unique for each member of staff. Each contract only applies to one event. There may be different contracts for an event depending upon different service needs.

eNo	contractNo	hours	eName	eventNo	eventLoc
1135	C1024	16	Smith J	H25	Queens
1057	C1024	24	Hocine D	H25	Queens
1068	C1025	28	White T	H4	Yonkers
1135	C1025	15	Smith J	H4	Yonkers
1135	C1026	10	Smith J	H25	Queens

a. State any assumptions you make about the data and the attributes shown in this table.

each contract applies to one event but each contract within an event is unique to each employee (Ex: Smith and Hocine are under C1024 but their hours differ based on need)

b. **Describe and illustrate the process of normalizing the table to 3NF relations.**

the table is already in 1NF, there are no repeating groups and has the PKs eNo, contractNo, and eventNo

1NF

eNo (PK)	contractNo (PK)	hours	eName	eventNo (PK)	eventLoc
1135	C1024	16	Smith J	H25	Queens
1057	C1024	24	Hocine D	H25	Queens
1068	C1025	28	White T	H4	Yonkers
1135	C1025	15	Smith J	H4	Yonkers
1135	C1026	10	Smith J	H25	Queens

In order to move onto 2NF, we must ensure that all non PK fields are dependent on all PK fields. Hours depends on the entire primary key. eName only depends on eNo and eventLoc only depends on eNo.

2NF

employee	
eNo (PK)	eName
1135	Smith J
1057	Hocine D
1068	White T

event	
eventNo (PK)	eventLoc
H25	Queens
H4	Yonkers

contract-event			
eNo (PK)	contractNo (PK)	eventNo (PK)	hours
1135	C1024	H25	16
1057	C1024	H25	24
1068	C1025	H4	28
1135	C1025	H4	15
1135	C1026	H25	10

after completing 2NF, we can see that the tables are already in 3NF as there are no transitive depencies

employee	
eNo (PK)	eName
1135	Smith J
1057	Hocine D
1068	White T

event	
eventNo (PK)	eventLoc
H25	Queens
H4	Yonkers

contract-event			
eNo (PK, FK)	contractNo (PK)	eventNo (PK, FK)	hours
1135	C1024	H25	16
1057	C1024	H25	24
1068	C1025	H4	28
1135	C1025	H4	15
1135	C1026	H25	10

c. **Identify the primary keys at each step of the normalization process and both the primary and foreign keys in your 3NF relations.**

The primary keys in the initial form and 1NF are eNo, contractNo, and eventNo

The primary keys in 2NF are eNo for employee table, eventNo for event table, and eNo, contractNo and eventNo for contract-event assignment table

The primary keys in 3NF are eNo for employee table, eventNo for event table, and eNo, contractNo, and eventNo for the contract event table. Contract event table also has foreign keys eNo and eventNo

d. Ensure that the 3NF relations are named.

employee	
eNo (PK)	eName
1135	Smith J
1057	Hocine D
1068	White T

event	
eventNo (PK)	eventLoc
H25	Queens
H4	Yonkers

1:M

contract-event			
eNo (PK, FK)	contractNo (PK)	eventNo (PK, FK)	hours
1135	C1024	H25	16
1057	C1024	H25	24
1068	C1025	H4	28
1135	C1025	H4	15
1135	C1026	H25	10

1:M

Grading Rubric

See Brightspace for Grading Rubric