

## Database Normalisation 1-8-2 (PI)

### What is Normalization?

Normalization is the process of efficiently organizing data in a database. There are two goals of the normalization process: eliminating redundant data (for example, storing the same data in more than one **table**) and ensuring **data dependencies** make sense (only storing related data in a table).

Both of these are worthy goals as they reduce the amount of space a database consumes and ensure that data is logically stored.

### The Normal Forms

The database community has developed a series of guidelines for ensuring that databases are normalized. These are referred to as normal forms and are numbered from one (the lowest form of normalization, referred to as **first normal form** or 1NF) through three (third normal form or 3NF). That said, let's explore the normal forms from next slide.

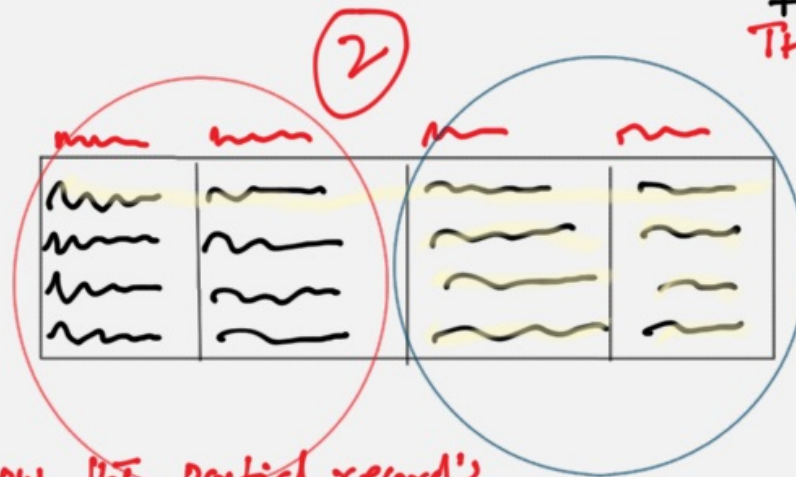
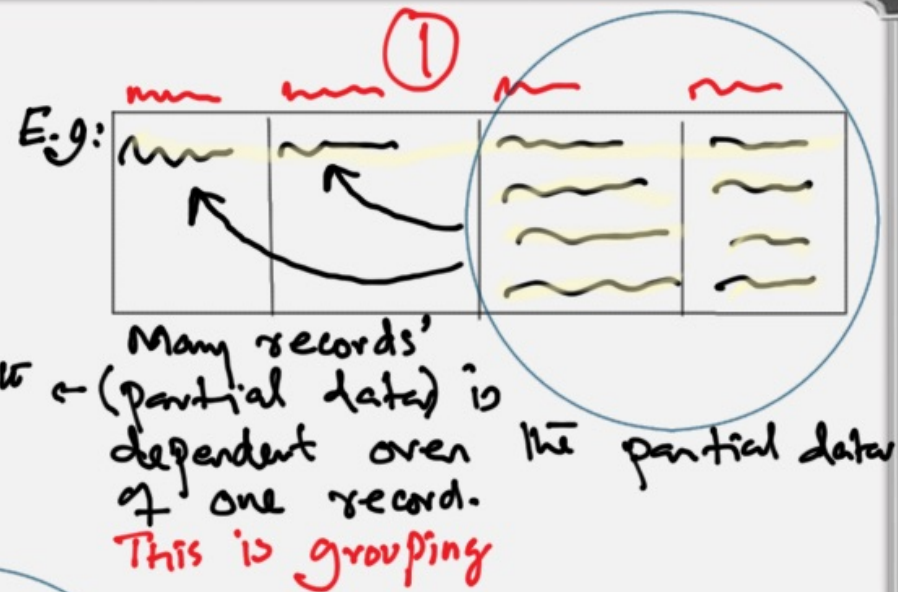


# NORMALISATION:

## 1<sup>st</sup> Normal Form (1NF):

- \* NO GROUPING
- \* NO REPETITION
- \* USE OF PRIMARY KEY

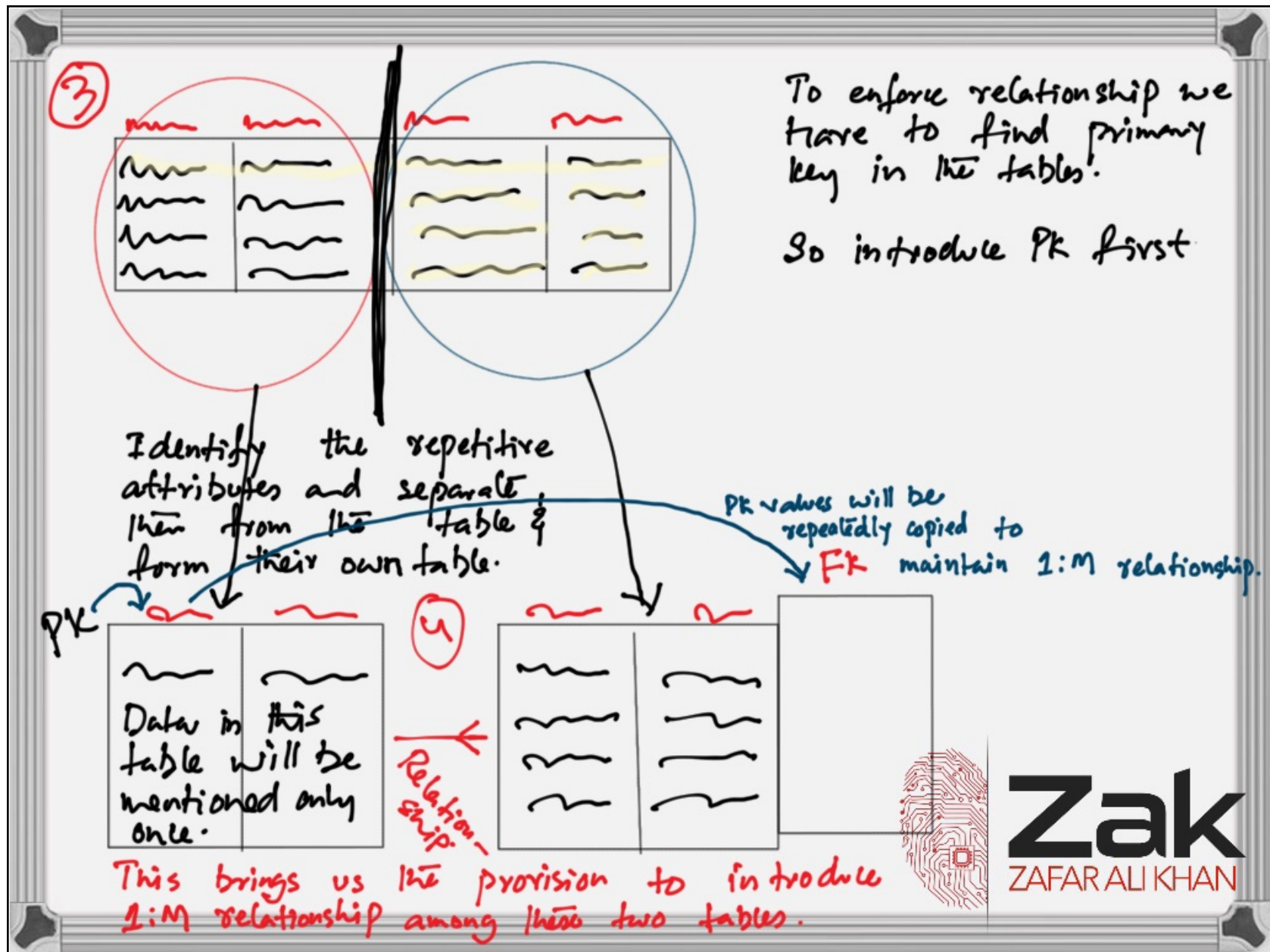
incomplete record



Now the partial record's data is copied to all other record's to get rid of grouping. BUT this will introduce repetition.



**Zak**  
ZAFAR ALI KHAN





StudentSubjects **NON NORMAL FORM.**

StudentName	TutorGroup	Tutor	Subject	Level	SubjectTeacher
Tom	6	SAN	Physics	A	SAN
Tom	6	SAN	Chemistry	A	MEB
Tom	6	SAN	Gen. Studies	AS	DIL
Joe	7	MEB	Geography	AS	ROG
			French	AS	HEN
Samir	6	SAN	Computing	A	VAR
			Chemistry	A	MEB
			Maths	A	COR
			Gen. Studies	A	DIL

**Grouping**  
Gaps show it!

**Repetition**  
Artificial gap.  
Filling creates it!

Normalisation Class - Microsoft Excel (Product Activation Failed)

DL: 0.16 mbps UL: 4.78 mbps

A16

Student (1NF)			
StudentName	TutorGroup	Tutor	
Tom	6	SAN	
Joe	7	MEB	
Samir	6	SAN	
PK			
KEY	NON-KEY		

StudentSubjects (1NF)			
StudentName	Subject	Level	SubjectTeacher
Tom	Physics	A	SAN
Tom	Chemistry	A	MEB
Tom	Gen. Studies	AS	DIL
Joe	Geography	AS	ROG
Joe	French	AS	HEN
Samir	Computing	A	VAR
Samir	Chemistry	A	MEB
Samir	Maths	A	COR
Samir	Gen. Studies	A	DIL
FK			
PK		NON-KEY	
COMPOUND KEY			

STUDENT — STUDSUB

Sheet1 Sheet1 (2) Sheet1 (3) Sheet1 (4) Sheet2 Sheet3

Ready 130% 6:08 PM 10/26/2016

## 2<sup>nd</sup> Normal Form:

NB: It is only applied to tables with compound keys. Rest of the tables are automatically 2NF clean.

\* Every non-key should be fully dependent over all parts of the compound key.

Those attributes that are partially dependent will be detached from the mother table along with the copy of the dependant compound key part.



Normalisation Class - Microsoft Excel (Product Activation Failed)

DL: 0.04 mbps UL: 2.11 mbps

File Home Insert Page Layout Formulas Data Review View Acrobat

Clipboard Font Alignment Number Styles Cells Editing

G13 fx NON-KEY

Student (1NF, 2NF)			
StudentName	TutorGroup	Tutor	
Tom	6	SAN	
Joe	7	MEB	
Samir	6	SAN	
PK			
KEY		NON-KEY	

StudentSubjects (1NF)		
StudentName	Subject	Level
Tom	Physics	A
Tom	Chemistry	A
Tom	Gen. Studies	AS
Joe	Geography	AS
Joe	French	AS
Samir	Computing	A
Samir	Chemistry	A
Samir	Maths	A
Samir	Gen. Studies	A
FK	FK	
PK		NON-KEY
COMPOUND KEY		

Subject (1NF, 2NF)	
Subject	SubjectTeacher
Physics	SAN
Chemistry	MEB
Gen. Studies	DIL
Geography	ROG
French	HEN
Computing	VAR
Maths	COR

STUDENT ↔ STUSUB ↔ SUBJECT

Sheet1 Sheet1 (2) Sheet1 (3) Sheet1 (4) Sheet2 Sheet3

Ready 100%

6:27 PM 10/26/2016



### 3<sup>RD</sup> Normal Form:

NB: Only applies to Table with more than one attributes in non-key.

\* No non-key should be dependent over another non-key.

Those attributes, which are non-key and dependent over another non-key are detached and put in a separate table along with the copy of non-key they are dependent over.





Normalisation Class - Microsoft Excel (Product Activation Failed)

DL: 0.16 mbps UL: 5.91 mbps

H22

File Home Insert Page Layout Formulas Data Review View Acrobat

Student (1NF, 2NF, 3NF)

StudentName	TutorGroup
Tom	6
Joe	7
Samir	6

PK NON-KEY

TutorGroup (1NF, 2NF, 3NF)

TutorGroup	Tutor
6	SAN
7	MEB

PK NON-KEY

StudentSubjects (1NF, 2NF, 3NF)

StudentName	Subject	Level
Tom	Physics	A
Tom	Chemistry	A
Tom	Gen. Studies	AS
Joe	Geography	AS
Joe	French	AS
Samir	Computing	A
Samir	Chemistry	A
Samir	Maths	A
Samir	Gen. Studies	A

FK FK

PK NON-KEY

COMPOUND KEY

Subject (1NF, 2NF, 3NF)

Subject	SubjectTeacher
Physics	SAN
Chemistry	MEB
Gen. Studies	DIL
Geography	ROG
French	HEN
Computing	VAR
Maths	COR

STUDENT

STUSUB

SUBJECT

TUTOR GROUP

Sheet1 Sheet1 (2) Sheet1 (3) Sheet1 (4) Sheet2 Sheet3

Ready

100%

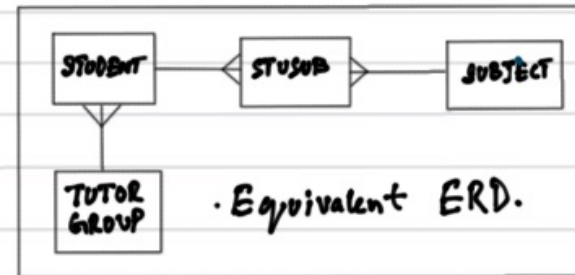
1:22 PM 10/29/2016

## Database Relationships in Domain Form:

A normalised database apart from ERD can also be shown using domain form.

E.g. the database that we have just normalised can also be shown in domain form, as:

**TutorGroup** (Tutor Group, Tutor)  
**Student** (StudentName, TutorGroup)  
**StudentSubject** (StudentName, Subject, level)  
**Subject** (Subject, SubjectTeacher)



Where keys are:

- underlined attributes are "Primary keys"; also called "keys"
- a primary key mentioned in another table and may have duplicate but related items to the primary key is "Foreign key"; it maintains relationship.
- a primary key composed of two or more attributes is called "Compound key".
- attributes other than PK are "non keys"
- more than one PK options are candidate keys, chosen one is PK & rest are secondary/alternate keys.



**Zak**  
ZAFAR ALI KHAN