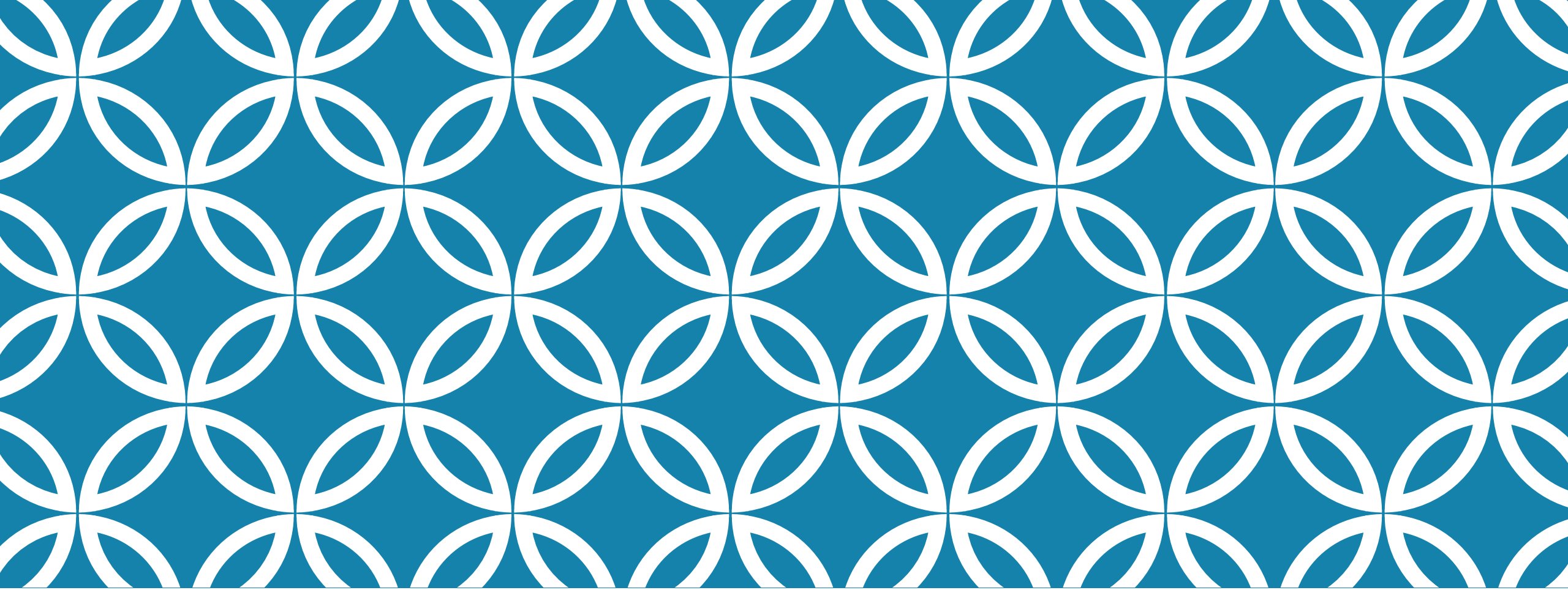


DATABASE NORMALIZATION

Manish joshi
Parul Institute of Computer Applications

WHAT IS NORMALIZATION?

This is the process which allows you to remove redundant data within your database.

This involves restructuring the tables to successively meet higher forms of Normalization.

A properly normalized database should have the following characteristics

- Atomic values in each fields
- Absence of redundancy.
- Minimal use of null values.
- Minimal loss of information.

FIRST NORMAL FORM (1NF)

- Each column must have only atomic values

Roll_no	Name	Subjects
101	Rahul	OS,JAVA
102	Priya	CN,DS
103	Kiran	OS

Roll_no	Name	Subjects
101	Rahul	OS
101	Rahul	JAVA
102	Priya	CN
102	Priya	DS
103	Kiran	OS

SECOND NORMAL FORM (2NF)

- Table must be in 1 NF
- It should not have partial dependency
- Partial Dependency – Proper subset of candidate key determines non key attribute.

Roll_no	Course_no	Course_fee
101	1001	1000
102	1002	3000
101	1004	5000
103	1005	3000

{roll_no,course_no} is the candidate key

- Multiple courses are having same fees
- Course_fee can not decide value of course_no or roll_no
- COURSE_FEE together with roll_no cannot decide the value of COURSE_NO
- COURSE_FEE together with COURSE_NO cannot decide the value of roll_no

2NF

Roll_no	Course_no
101	1001
102	1002
101	1004
103	1005

Course_no	Course_fees
1001	1000
1002	3000
1004	5000
1005	3000

THIRD NORMAL FORM (3NF)

Table must be in 2NF

There is no transitive dependency for non key attributes.

- A->B & B->C then A->C

Stud_no is the candidate

STUD_NO	STUD_NAME	STUD_STATE	STUD_COUNTRY	STUD_AGE
1	RAM	HARYANA	INDIA	20
2	RAM	PUNJAB	INDIA	19
3	SURESH	PUNJAB	INDIA	21

STUD_NO -> STUD_STATE and STUD_STATE -> STUD_COUNTRY are true.
So STUD_COUNTRY is transitively dependent on STUD_NO.

3NF

STUD_NO	STUD_NAME	STUD_STATE	STUD_COUNTRY	STUD_AGE
1	RAM	HARYANA	INDIA	20
2	RAM	PUNJAB	INDIA	19
3	SURESH	PUNJAB	INDIA	21

Stud_no	Stud_name	Stud_state	Stud_age
…			

State	country
…	

BOYCE CODD NORMAL FORM (BCNF)

Table must be in 3NF

For every X->Y , X is superkey

Roll_no + subject is the primary key

Roll_no	Subject	Professor
101	Java	Mr. J1
101	OS	Mr. O
102	Java	Mr. J2

{roll_no,subject} ->professor
Professor->subject (professor is not super key)

BCNF

Roll_no	Subject	Professor
101	Java	Mr. J1
101	OS	Mr. O
102	Java	Mr. J2

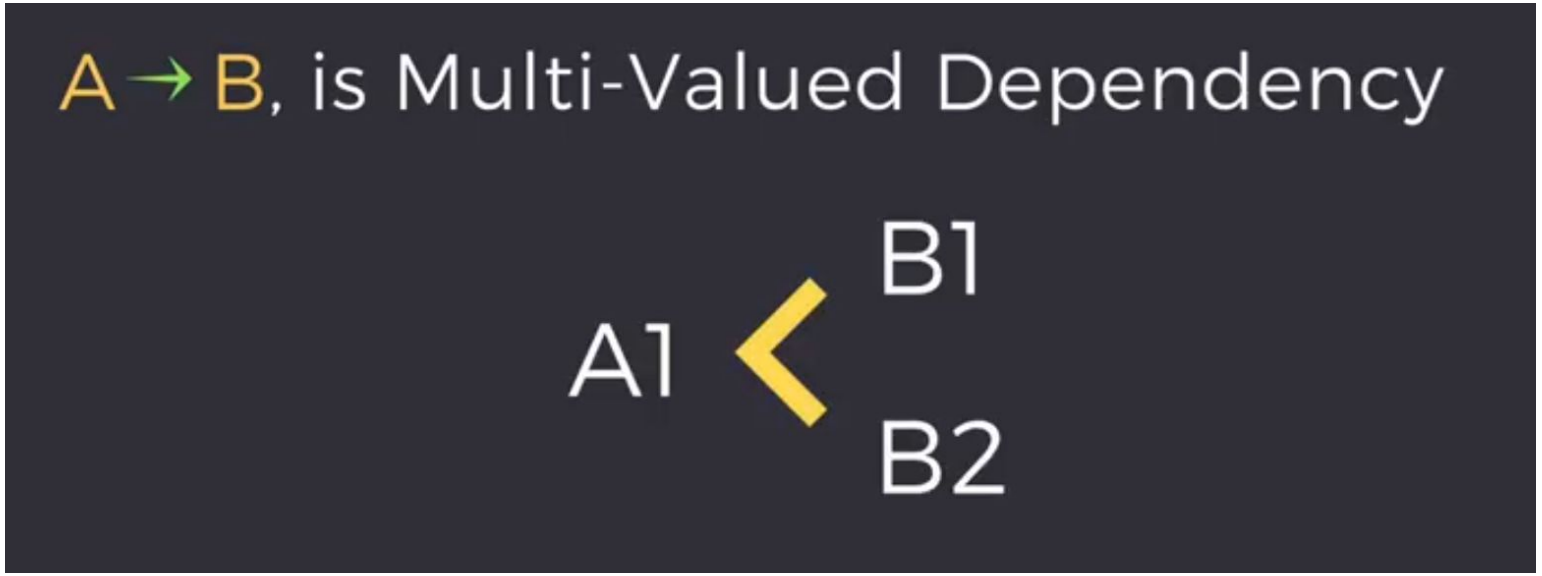
Roll_no	Prof_id
101	P1
101	P2
102	P3

Prof-id	Professor	Subject
P1	Mr. J1	Java
P2	Mr. O	OS
P3	Mr. J2	Java

FOURTH NORMAL FORM (4NF)

Table must be in BCNF

It should not have Multivalued Dependency.



4NF

- A ↠ B, for a single value of A, more than one value of B exist.
- Table should have at-least 3 columns.
- For this table with A, B, C columns, B and C should be independent.

4NF

s_id	course	hobby
1	Science	Cricket
1	Maths	Hockey
2	C#	Cricket
2	Php	Hockey

s_id	course
1	Science
1	Maths
2	C#
2	Php

s_id	hobby
1	Cricket
1	Hockey
2	Cricket
2	Hockey

FIFTH NORMAL FORM (5NF)

. Fifth normal form is satisfied when all tables are divided into as many tables as possible in order to avoid redundancy. Once it is in fifth normal form it cannot be broken into smaller relations without changing the facts or the meaning.

Student-ID	Mobile Number	Hobby
123	9999900000	Dancing
124	9999900000	Singing
124	9999900000	Dancing
123	8975622122	Singing
123	9999900000	Singing

=

Student-ID	Mobile Number
123	9999900000
123	8975622122
124	9999900000

+

Student-ID	Hobby
123	Dancing
123	Singing
124	Singing
124	Dancing

+

Mobile Number	Hobby
9999900000	Dancing
9999900000	Singing
8975622122	Singing

THANK YOU