



Prifysgol Cymru
Y Drindod Dewi Sant
University of Wales
Trinity Saint David

200
1822 - 2022

Normalization

Man Geni Addysg Uwch yng Nghymru
Where Higher Education Began in Wales

www.uwtsd.ac.uk

Anomalies

- 3 problems that can occur when we work with the data:
- **INSERT ANOMALY:** This refers to the situation when it is impossible to insert certain types of data into the database.
- **DELETE ANOMALY:** The deletion of data leads to unintended loss of additional data, data that we had wished to preserve.
- **UPDATE ANOMALY:** This refers to the situation where updating the value of a column leads to database inconsistencies (i.e., different rows on the table have different values).

Anomalies

Rollno	Game	Fee structure
1	Basketball	500
2	Basketball	500
3	Basketball	500
4	Cricket	600
5	Cricket	600
6	Cricket	600
7	Tennis	400

student table has from all three anomalies

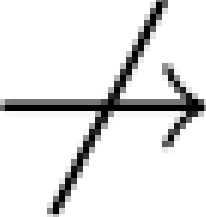
- Possibility of :
 - Inserting new game without student enrolment?
- Losing information
Eg: deletion of Roll no 7- lost tennis.
- Changes made in more than one place.

What is Normalization?

- Normalization is the process of organizing data in a database.
- creating tables and establishing relationships between those tables
- Adhering to data protection rules
- Flexible database by eliminating redundancy and inconsistent dependency.

Inconsistent Dependency?

- Inconsistent dependencies can make data difficult to access because the path to find the data may be missing or broken.

Example : customer table  Employee salary

Levels of database normalization

- 1st Normal Form (1NF)
- 2nd Normal Form (2NF)
- 3rd Normal Form (3NF)
- Boyce-Codd Normal Form (BCNF)
- 4th Normal Form (4NF)
- 5th Normal Form (5NF)

First Normal Form

- A database is in first normal form if it satisfies the following conditions:
- Contains only atomic values
- There are no repeating groups

First Normal Form-Example

TABLE_PRODUCT

Product ID	Color	Price
1	red, green	15.99
2	yellow	23.99
3	green	17.50
4	yellow, blue	9.99
5	red	29.99

This table is not in first normal form

To bring this table to first normal form, we split the table into two tables

First Normal Form-Example

Product ID	Color	Price
1	Red	15.99
1	Green	15.99
2	Yellow	23.99
3	Green	17.50
4	Yellow	9.99
4	Blue	9.99
5	red	29.99

First Normal Form-Activity 1

Student_No	Student_Name	Phone_Number	City	Country
07895	James Moller	079856 56743	Swansea	UK
07986	Alan Peter	078955 76846 079899 67546	Cardiff	UK
18789	Robert Murray	079653 87432	Swansea	UK
17865	Enola Holmes	077657 54367 076854 34587	London	UK
18965	Enola Holmes	079846 54635	Edinburgh	UK

First Normal Form-Activity 2

<i>branchNo</i>	<i>branchAddress</i>	<i>telNos</i>
B001	8 Jefferson Way, Portland, OR 97201	503-555-3618, 503-555-2727, 503-555-6534
B002	City Center Plaza, Seattle, WA 98122	206-555-6756, 206-555-8836
B003	14 – 8th Avenue, New York, NY 10012	212-371-3000
B004	16 – 14th Avenue, Seattle, WA 98128	206-555-3131, 206-555-4112

(a) Why is this table not in 1NF?

(b) Describe and illustrate the process of normalizing the data shown in this table to first normal form (1NF).

Second normal form

- The 2nd Normal Form (2NF) eliminates the Partial Dependency.
- A database is in second normal form if it satisfies the following conditions:
- It is in first normal form
- All non-key attributes are fully functional dependent on the primary key

Second normal form

The prime key attributes are **StudentID** and **CourseNo**

Student ID	Student Name	Course No	Course Name
S01	George	P001	Course1
S02	Alexander	P002	Course2

non-prime attributes i.e. **StudentName** and **CourseName** should be functionally dependent on candidate key.

- The **StudentName** can be determined by **StudentID**, which makes the relation Partial Dependent.
- The **CourseName** can be determined by **CourseNo**, which makes the relation Partial Dependent.

Second normal form

Student id	Student name	Course No
S01	George	P001
S02	Alexander	P002

Course No	Course Name
P001	Course1
P002	Course2

Create separate tables for sets of values and relate these tables with a foreign key.

Second normal form - Activity

Artist ID*	Artist Name	Venue ID*	Concert Venue	Location of Venue	Concert Date*	Agent First Name	Agent Last Name	Agent Email
RO01	Rock On	O21	O ₂ Academy	Birmingham	01/06/2016	Alice	Jones	ajones@welovemusic.com
SF01	Stage Fright	HA1	Hammersmith Apollo	London	01/06/2016	Bill	Richards	billyrich@amazingartists.com
JS01	John Smith	O22	O ₂ Academy	Leeds	10/06/2016	Henry	Cooper	henry.cooper@me.com
RO01	Rock On	BC1	Barbican Centre	London	06/06/2016	Alice	Jones	ajones@welovemusic.com
RO01	Rock On	O23	O ₂ Academy	Liverpool	10/06/2016	Alice	Jones	ajones@welovemusic.com
SF01	Stage Fright	RA1	The Royal Albert Hall	London	20/06/2016	Bill	Richards	billyrich@amazingartists.com
JY01	Jimmy	RA1	The Royal Albert Hall	London	28/06/2016	Henry	Cooper	henry.cooper@me.com
TD01	The Delights	O21	O ₂ Academy	Birmingham	13/06/2016	Alice	Jones	ajones@welovemusic.com
RO01	Rock On	HA1	Hammersmith Apollo	London	13/06/2016	Alice	Jones	ajones@welovemusic.com
JY01	Jimmy	BC1	Barbican Centre	London	30/06/2016	Henry	Cooper	henry.cooper@me.com

Third normal form

- A database is in third normal form if it satisfies the following conditions:
- It is in second normal form
- There is no transitive functional dependency

Third normal form

Consider the following example:

Emp_ID	Emp_Name	Emp_Postalcode	Emp_state	Emp_city
111	Harry	20101	UP	Boston
222	Steve	0228	UK	Noida
333	Lan	06347	US	Chicago

- Superkey :{(Emp_ID),(Emp_ID,Emp_name),(Emp_id,Emp_name,Postalcode)}.....
- we have transitive functional dependency
- Emp_State and Emp_city dependent on Emp_postalcode
- Emp_postalcode dependent on Emp_ID
- To bring this table to third normal form, we split the table into two as follows:

Third normal form

- Employee table

Emp_ID	Emp_name	Emp_Postalcode
111	harry	20101
222	steve	0228
333	Lan	06347

Employee_postal table

Emp_Postalcode	Emp_state	Emp_city
20101	UP	Boston
0228	UK	Noida
06347	US	Chicago

Third normal form – Activity

ArtistDetails

Artist ID*	Artist Name	Agent First Name	Agent Last Name	Agent Email
RO01	Rock On	Alice	Jones	ajones@welovemusic.com
SF01	Stage Fright	Bill	Richards	billyrich@amazingartists.com
JS01	John Smith	Henry	Cooper	henry.cooper@me.com
JY01	Jimmy	Henry	Cooper	henry.cooper@me.com
TD01	The Delights	Alice	Jones	ajones@welovemusic.com

Third normal form – Activity

VenueDetails

Venue ID*	Concert Venue	Location of Venue
O21	O ₂ Academy	Birmingham
HA1	Hammersmith Apollo	London
O22	O ₂ Academy	Leeds
BC1	Barbican Centre	London
O23	O ₂ Academy	Liverpool
RA1	The Royal Albert Hall	London

Third normal form – Activity

ArtistsBookings

Concert ID*	Artist ID	Venue ID	Concert Date
CON001	RO01	O21	01/06/2016
CON002	SF01	HA1	01/06/2016
CON003	JS01	O22	10/06/2016
CON004	RO01	BC1	06/06/2016
CON005	RO01	O23	10/06/2016
CON006	SF01	RA1	20/06/2016
CON007	JY01	RA1	28/06/2016
CON008	TD01	O21	13/06/2016
CON009	RO01	HA1	13/06/2016
CON010	JY01	BC1	30/06/2016

Exercise 1

Consider the below table for three unique ORDER_IDs,

- Convert to the First Normal Form
- Identify the functional dependencies
- Convert the table created in (i) to the Second Normal Form. Provide the Table name, column names, and appropriate key types
- Convert the table created in (iii) to the Third Normal Form. Provide the Table name, column names, and appropriate key type

ORDERS

ORDER_ ID	DATE	CUST_ ID	CUST_ NAME	STATE	ITEM_ _NUM	ITEM_ DESCRIP	QUANTITY	PRICE
1001	7/7	A004	George	VA	6531	Tape	5	100.00
					7890	Table	1	400.00
					7956	Pens	1	2.00
1002	7/8	A999	Ben	PA	3456	Stove	1	300.00
1003	7/9	A201	Tom	PA	4234	Paper	15	234.00
					7956	Pens	10	20.00

Exercise 2

- Hotels in Scotland receive part-time or temporary pstaff through a company named Enterprise. The table below lists the time spent by company staff working at various hotels. The national insurance number (NIN) is unique for every member of staff

NIN	ContractNo	Hours	eName	hNo	hLoc
1135	C1024	16	Smith J.	H25	East Killbride
1057	C1024	24	Hocine D.	H25	East Killbride
1068	C1025	28	White T.	H4	Glasgow
1135	C1025	15	Smith J.	H4	Glasgow

1. The above table is susceptible to update anomalies. Provide examples of insertion, deletion and update anomalies.
 2. Normalize this table to third normal form.
- State any assumptions.