



Team –1

JavaBots

NORMALIZATION



What is Normalization?

- Normalization is the process of organizing the data in the database.
- Normalization divides the larger table into smaller and links them using relationships.
- The normal form is used to reduce redundancy (**having several copies of the same data**) from the database table.

Sales Records:

Cust Name	Item	Shipping Address	Newsletter	Supplier	Supplier Phone	Price
Alan Smith	Xbox One	35 Palm St, Miami	Xbox News	Microsoft	(800) BUY-XBOX	250
Roger Banks	PlayStation 4	47 Campus Rd, Boston	PlayStation News	Sony	(800) BUY-SONY	300
Evan Wilson	Xbox One, PS Vita	28 Rock Av, Denver	Xbox News, PlayStation News	Wholesale	Toll Free	450
Alan Smith	PlayStation 4	47 Campus Rd, Boston	PlayStation News	Sony	(800) BUY-SONY	300

1st Normal Form(1NF)

- Each cell to be single valued.
- Entries in a column should be of same type.
- Rows should be uniquely identified.

Primary Key

Cust ID	Cust Name	Item	Shipping Address	Newsletter	Supplier	Supplier Phone	Price
at_smith	Alan Smith	Xbox One	35 Palm St, Miami	Xbox News	Microsoft	(800) BUY-XBOX	250
roger25	Roger Banks	PlayStation 4	47 Campus Rd, Boston	PlayStation News	Sony	(800) BUY-SONY	300
wilson44	Evan Wilson	Xbox One	28 Rock Av, Denver	Xbox News	Microsoft	(800) BUY-XBOX	250
wilson44	Evan Wilson	PS Vita	28 Rock Av, Denver	PlayStation News	Sony	(800) BUY-SONY	200
am_smith	Alan Smith	PlayStation 4	47 Campus Rd, Boston	PlayStation News	Sony	(800) BUY-SONY	300

2-NF

- It should be in 1NF.
- It should not hold partially Dependency.

Primary Key

Cust ID	Cust Name	Shipping Address	Newsletter
at_smith	Alan Smith	35 Palm St, Miami	Xbox News
roger25	Roger Banks	47 Campus Rd, Boston	PlayStation News
wilson44	Evan Wilson	28 Rock Av, Denver	Xbox News
wilson44	Evan Wilson	28 Rock Av, Denver	PlayStation News
am_smith	Alan Smith	47 Campus Rd, Boston	PlayStation News

Primary Key

Item	Supplier	Supplier Phone	Price
Xbox One	Microsoft	(800) BUY-XBOX	250
PlayStation 4	Sony	(800) BUY-SONY	300
PS Vita	Sony	(800) BUY-SONY	200

Primary Key

Primary Key

Cust ID	Item
at_smith	Xbox One
roger25	PlayStation 4
wilson44	Xbox One
wilson44	PS Vita
am_smith	PlayStation 4

3-NF

- It should be in 2NF
- There should not be transitive functional dependency.
- If $A \rightarrow B$ then A should be super key or B is prime Attribute.

Primary Key

Cust ID	Cust Name	Shipping Address	Newsletter
at_smith	Alan Smith	35 Palm St, Miami	Xbox News
roger25	Roger Banks	47 Campus Rd, Boston	PlayStation News
wilson44	Evan Wilson	28 Rock Av, Denver	Xbox News
wilson44	Evan Wilson	28 Rock Av, Denver	PlayStation News
am_smith	Alan Smith	47 Campus Rd, Boston	PlayStation News

Primary Key	Foreign Key	
Item	Supplier	Price
Xbox One	Microsoft	250
PlayStation 4	Sony	300
PS Vita	Sony	200

Primary Key	Primary Key
Cust ID	Item
at_smith	Xbox One
roger25	PlayStation 4
wilson44	Xbox One
wilson44	PS Vita
am_smith	PlayStation 4

Primary Key

Supplier	Supplier Phone
Microsoft	(800) BUY-XBOX
Sony	(800) BUY-SONY

BCNF

- BCNF is a special case of 3NF, where we have multiple overlapping candidate keys.

Popularity	Item	Date of Release	Year of Release
2	Xbox One	15 June 2012	2012
1	PlayStation 4	9 July 2013	2013
3	PS Vita	20 July 2013	2013
1	PlayStation 5	10 May 2014	2014

Candidate key:

- {Item}.
- {Popularity, Release Year}.
- {DOR, Popularity}

Here we can see that Year Of release is dependent upon Date of Release. This may give rise to inconsistent data.

For example we can see that year of release and date of release contradict themselves. TO avoid such cases BCNF is used.

Here we can see that now Date Of Purchase is not dependent on year of purchase.

2	PS Vita 2.0	30 May 2014	2016
---	-------------	-------------	------

Popularity	Item	DOR	YOR
at_smith	Xbox One	15 June	2012
wilson44	Playstation 4	9 July	2014

4-NF

- It should be in BCNF.
- It should not have multivalued dependency.
- For a dependency $A \twoheadrightarrow B$, if for a single value of A there exists multiple values in B then this relation is called multivalued dependency.

Primary Key

Cust ID	Cust Name	Shipping Address
at_smith	Alan Smith	35 Palm St, Miami
roger25	Roger Banks	47 Campus Rd, Boston
wilson44	Evan Wilson	28 Rock Av, Denver
am_smith	Alan Smith	47 Campus Rd, Boston

Primary Key

Item	Supplier	Price
Xbox One	Microsoft	250
PlayStation 4	Sony	300
PS Vita	Sony	200

Primary Key

Cust ID	Newsletter
at_smith	Xbox News
roger25	PlayStation News
wilson44	Xbox News
wilson44	PlayStation News
am_smith	PlayStation News

Primary Key

Cust ID	Item
at_smith	Xbox One
roger25	PlayStation 4
wilson44	Xbox One
wilson44	PS Vita
am_smith	PlayStation 4

Primary Key

Primary Key

Supplier	Supplier Phone
Microsoft	(800) BUY-XBOX
Sony	(800) BUY-SONY

5-NF

- It should be in 4NF.
- Does not have join dependency and joining should be lossless.
- 5NF is satisfied when all the tables are broken into as many tables as possible in order to avoid redundancy.
- 5NF is also known as Project-join normal form.

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

Table P1

Student-ID	Mobile Number
123	9999900000
123	8975622122
124	9999900000

Table P2

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

5-NF

If natural join is performed on all the three relations then there will be no extra tuples.

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

Table P3

Mobile Number	Hobby
9999900000	Dancing
9999900000	Singing
8975622122	Singing

So now we can say that P1, P2 and P3 are in fifth normal form.