



Normal Forms

Prepared By,
K.Revathi, AP/IT



NORMALIZATION



Definition:

- Normalization is the process of organizing the data in the database.
- **Normal forms** are used to eliminate or reduce redundancy in database tables.



TYPES



	1NF	2NF	3NF	4NF	5NF
Decomposition of Relation	R	R ₁₁ R ₁₂	R ₂₁ R ₂₂ R ₂₃	R ₃₁ R ₃₂ R ₃₃ R ₃₄	R ₄₁ R ₄₂ R ₄₃ R ₄₄ R ₄₅
Conditions	Eliminate Repeating Groups	Eliminate Partial Functional Dependency	Eliminate Transitive Dependency	Eliminate Multi-values Dependency	Eliminate Join Dependency



Normalization: Advantages and Disadvantages

Advantages and Disadvantages of Normalization

- Advantages

Eliminate modification anomalies

Reduce duplicated data

- Eliminate data integrity problems

- Save file space

Single table queries will run faster

- Disadvantages

More complicated SQL required for multitable subqueries and joins

Extra work for DBMS can mean slower applications



First Normal Form



Definition:

A relation is said to be in **1NF** (**first normal form**), if it doesn't contain any multi-valued attribute.



Example



Emp_Id	Emp_Name	Emp_Address	Emp_Mobile
101	Herschel	New Delhi	8912312390
102	Jon	Kanpur	8812121212 , 9900012222



Solution: 1NF



Emp_Id	Emp_Name	Emp_Address	Emp_Mobile
101	Herschel	New Delhi	8912312390
102	Jon	Kanpur	8812121212
102	Jon	Kanpur	9900012222



Second normal form (2NF)



A table is said to be in 2NF if both the following conditions hold:

- Table is in 1NF (First normal form)
- No non-prime attribute is dependent on the proper subset of any candidate key of table.



EXAMPLE

Teacher Table

Teacher_Id	Subject	Teacher_Age
111	Maths	38
111	Physics	38
222	Biology	38
333	Physics	40
333	Chemistry	40



Candidate Keys:

{Teacher_Id, Subject}

Non prime attribute:

Teacher_Age



This violates the rule for 2NF as the rule says “no non-prime attribute is dependent on the proper subset of any candidate key of the table”.



SOLUTION IN 2NF



Teacher_Details table

Teacher_Id	Teacher_Age
111	38
222	38
333	40



Teacher_Subject table:

Teacher_Id	Subject
111	Maths
111	Physics
222	Biology
333	Physics
333	Chemistry



Third Normal form (3NF)



A table design is said to be in 3NF if both the following conditions hold:

- Table must be in 2NF
- Transitive functional dependency of non-prime attribute on any super key should be removed.



EMPLOYEE_DETAIL table:

EMP_ID	EMP_NAME	EMP_ZIP	EMP_STATE	EMP_CITY
222	Harry	201010	UP	Noida
333	Stephan	02228	US	Boston
444	Lan	60007	US	Chicago
555	Katharine	06389	UK	Norwich
666	John	462007	MP	Bhopal



- **Super key in the table above:**
- {EMP_ID}, {EMP_ID, EMP_NAME}, {EMP_ID, EMP_NAME, EMP_ZIP}....so on
- **Candidate key:** {EMP_ID}
- **Non-prime attributes:** In the given table, all attributes except EMP_ID are non-prime.



Solution in 3NF

EMPLOYEE table:

EMP_ID	EMP_NAME	EMP_ZIP
222	Harry	201010
333	Stephan	02228
444	Lan	60007
555	Katharine	06389
666	John	462007



EMPLOYEE_ZIP table:

EMP_ZIP	EMP_STATE	EMP_CITY
201010	UP	Noida
02228	US	Boston
60007	US	Chicago
06389	UK	Norwich
462007	MP	Bhopal



Boyce Codd normal form (BCNF)

- BCNF is the advance version of 3NF.
- A table is in BCNF if every functional dependency $X \rightarrow Y$, X is the super key of the table.



- **EMPLOYEE table:**

EMP_ID	EMP_COUNTRY	EMP_DEPT	DEPT_TYPE	EMP_DEPT_NO
264	India	Designing	D394	283
264	India	Testing	D394	300
364	UK	Stores	D283	232
364	UK	Developing	D283	549



Solution: BCNF



we decompose it into three tables

EMP_COUNTRY table:

EMP_ID	EMP_COUNTRY
264	India
264	India



EMP_DEPT table:

EMP_DEPT	DEPT_TYPE	EMP_DEPT_NO
Designing	D394	283
Testing	D394	300
Stores	D283	232
Developing	D283	549



EMP_DEPT_MAPPING table:

EMP_ID	EMP_DEPT
D394	283
D394	300
D283	232
D283	549



Fourth normal form (4NF)

A relation will be in 4NF if it is in Boyce Codd normal form and has no multi-valued dependency.



Fourth normal form (4NF)

A relation will be in 4NF if it is in Boyce Codd normal form and has no multi-valued dependency.



Example

STUDENT Table

STU_ID	COURSE	HOBBY
21	Computer	Dancing
21	Math	Singing
34	Chemistry	Dancing
74	Biology	Cricket
59	Physics	Hockey



- The given STUDENT table is in 3NF, but the COURSE and HOBBY are two independent entity. Hence, there is no relationship between COURSE and HOBBY.
- So to make the above table into 4NF, we can decompose it into two tables:



STUDENT_COURSE Table

STU_ID	COURSE
21	Computer
21	Math
34	Chemistry
74	Biology
59	Physics



STUDENT_HOBBY Table

STU_ID	HOBBY
21	Dancing
21	Singing
34	Dancing
74	Cricket
59	Hockey



Fifth normal form (5NF)

- A relation is in 5NF if it is in 4NF and not contains any join dependency and joining should be lossless.
- 5NF is also known as Project-join normal form (PJ/NF).



Example Teacher Table

SUBJECT	LECTURER	SEMESTER
Computer	Anshika	Semester 1
Computer	John	Semester 1
Math	John	Semester 1
Math	Akash	Semester 2
Chemistry	Praveen	Semester 1



So to make the above table into 5NF, we can decompose it into three relations

P1, P2 & P3



P1

SEMESTER	SUBJECT
Semester 1	Computer
Semester 1	Math
Semester 1	Chemistry
Semester 2	Math



P2

SUBJECT	LECTURER
Computer	Anshika
Computer	John
Math	John
Math	Akash
Chemistry	Praveen



Sixth normal form (6NF)

- A relation is in 6NF, only if, It is in 5NF, and every join dependency on the relation is trivial.



Example Student Table

Student_ID	Student_FirstName	Student_LastName	Marks
S01	Tom	Alter	90
S02	Jacob	Watson	80
S03	Harry	Smith	85



Solution: 6NF

StudentFirstName

Student_ID	Student_FirstName
S01	Tom
S02	Jacob
S03	Harry



StudentLastName

Student_ID	Student_LastName
S01	Alter
S02	Watson
S03	Smith



StudentResult

Student_ID	Marks
S01	90
S02	80
S03	85