

Database Management Systems (CSN-351)

Relational Database Design (contd.)

BTech 3rd Year (CS) + Minor + Audit

Instructor: **Ranita Biswas**
Department of Computer Science and Engineering
Indian Institute of Technology Roorkee
Roorkee, Uttarakhand - 247 667, India



1NF

EMP_PROJ		Projs	
Ssn	Ename	Pnumber	Hours

1NF

EMP_PROJ

		Projs	
Ssn	Ename	Pnumber	Hours

Ssn	Ename	Pnumber	Hours
123456789	Smith, John B.	1	32.5
		2	7.5
666884444	Narayan, Ramesh K.	3	40.0
453453453	English, Joyce A.	1	20.0
		2	20.0
333445555	Wong, Franklin T.	2	10.0
		3	10.0
		10	10.0
		20	10.0
999887777	Zelaya, Alicia J.	30	30.0
		10	10.0
987987987	Jabbar, Ahmad V.	10	35.0
		30	5.0
987654321	Wallace, Jennifer S.	30	20.0

1NF

EMP_PROJ		Projs	
Ssn	Ename	Pnumber	Hours

1NF

EMP_PROJ

Projs			
Ssn	Ename	Pnumber	Hours

EMP_PROJ1

Ssn	Ename
-----	-------

EMP_PROJ2

Ssn	Pnumber	Hours
-----	---------	-------

1NF

DEPARTMENT

Dname	<u>Dnumber</u>	Dmgr_ssn	Dlocations

The diagram illustrates the normalization process for the DEPARTMENT table. It shows three vertical arrows pointing upwards from the table rows to the column headers Dnumber, Dmgr_ssn, and Dlocations respectively, indicating that these columns are candidate keys. A horizontal dashed arrow points from the right side of the Dmgr_ssn column towards the right edge of the table, representing a foreign key relationship.

1NF

DEPARTMENT

Dname	<u>Dnumber</u>	Dmgr_ssn	Dlocations

DEPARTMENT

Dname	<u>Dnumber</u>	Dmgr_ssn	Dlocations
Research	5	333445555	{Bellaire, Sugarland, Houston}
Administration	4	987654321	{Stafford}
Headquarters	1	888665555	{Houston}

1NF

DEPARTMENT

Dname	<u>Dnumber</u>	Dmgr_ssn	Dlocations


DEPARTMENT

Dname	<u>Dnumber</u>	Dmgr_ssn	Dlocations
Research	5	333445555	{Bellaire, Sugarland, Houston}
Administration	4	987654321	{Stafford}
Headquarters	1	888665555	{Houston}

DEPARTMENT

Dname	<u>Dnumber</u>	Dmgr_ssn	<u>Dlocation</u>
Research	5	333445555	Bellaire
Research	5	333445555	Sugarland
Research	5	333445555	Houston
Administration	4	987654321	Stafford
Headquarters	1	888665555	Houston

2NF

Definition (2NF)

A relation schema R is in 2NF if every *nonprime* attribute A in R is *fully functionally dependent* on the primary key of R .

2NF

Definition (2NF)

A relation schema R is in 2NF if every *nonprime* attribute A in R is *fully functionally dependent* on the primary key of R .

Definition (Prime and Nonprime Attributes)

An attribute of relation schema R is called a prime attribute of R if it is a member of some candidate key of R . An attribute is called nonprime if it is not a prime attribute — that is, if it is not a member of any candidate key.

2NF

Definition (Full Functional Dependency)

A functional dependency $X \rightarrow Y$ is a full functional dependency if removal of any attribute A from X means that the dependency does not hold any more; that is, for any attribute $A \in X$, $(X - \{A\})$ does not functionally determine Y .

2NF

Definition (Full Functional Dependency)

A functional dependency $X \rightarrow Y$ is a full functional dependency if removal of any attribute A from X means that the dependency does not hold any more; that is, for any attribute $A \in X$, $(X - \{A\})$ does not functionally determine Y .

Definition (Partial Functional Dependency)

A functional dependency $X \rightarrow Y$ is a partial dependency if some attribute $A \in X$ can be removed from X and the dependency still holds; that is, for some $A \in X$, $(X - \{A\}) \rightarrow Y$.

2NF

EMP_PROJ

Ssn	Pnumber	Hours	Ename	Pname	Plocation
FD1					
FD2					
FD3					

The diagram illustrates three functional dependencies (FD1, FD2, and FD3) for the attributes of the EMP_PROJ relation. FD1 is shown as covering the attributes Ssn, Pnumber, and Hours. FD2 is shown as covering the attributes Ssn, Pnumber, and Ename. FD3 is shown as covering the attributes Ssn, Pnumber, and Pname.

2NF

EMP_PROJ

Ssn	Pnumber	Hours	Ename	Pname	Plocation
-----	---------	-------	-------	-------	-----------

FD1



FD2



FD3

**2NF Normalization****EP1**

Ssn	Pnumber	Hours
-----	---------	-------

FD1

**EP2**

Ssn	Ename
-----	-------

FD2

**EP3**

Pnumber	Pname	Plocation
---------	-------	-----------

FD3



3NF

Definition (3NF)

A relation schema R is in 3NF if it satisfies 2NF and no nonprime attribute of R is *transitively dependent* on the primary key.

3NF

Definition (3NF)

A relation schema R is in 3NF if it satisfies 2NF and no nonprime attribute of R is *transitively dependent* on the primary key.

Definition (Transitive Dependency)

A functional dependency $X \rightarrow Y$ in a relation schema R is a transitive dependency if there exists a set of attributes Z in R that is neither a candidate key nor a subset of any key of R and both $X \rightarrow Z$ and $Z \rightarrow Y$ hold.

3NF

EMP_DEPT

Ename	Ssn	Bdate	Address	Dnumber	Dname	Dmgr_ssn



3NF

EMP_DEPT

Ename	Ssn	Bdate	Address	Dnumber	Dname	Dmgr_ssn

Diagram illustrating dependencies in the EMP_DEPT table:

- Ename → Dmgr_ssn
- Ssn → Dmgr_ssn
- Bdate → Dmgr_ssn
- Address → Dmgr_ssn
- Dnumber → Dmgr_ssn

3NF Normalization

**ED1**

Ename	Ssn	Bdate	Address	Dnumber

Diagram illustrating dependencies in ED1:

- Ename → Dmgr_ssn
- Ssn → Dmgr_ssn
- Bdate → Dmgr_ssn
- Address → Dmgr_ssn
- Dnumber → Dmgr_ssn

ED2

Dnumber	Dname	Dmgr_ssn

Diagram illustrating dependencies in ED2:

- Dnumber → Dmgr_ssn
- Dname → Dmgr_ssn
- Dmgr_ssn → Dmgr_ssn

Summary of Normal Forms

Normal Form	Test	Remedy (Normalization)
First (1NF)	Relation should have no multivalued attributes or nested relations.	Form new relations for each multivalued attribute or nested relation.
Second (2NF)	For relations where primary key contains multiple attributes, no nonkey attribute should be functionally dependent on a part of the primary key.	Decompose and set up a new relation for each partial key with its dependent attribute(s). Make sure to keep a relation with the original primary key and any attributes that are fully functionally dependent on it.
Third (3NF)	Relation should not have a nonkey attribute functionally determined by another nonkey attribute (or by a set of nonkey attributes). That is, there should be no transitive dependency of a nonkey attribute on the primary key.	Decompose and set up a relation that includes the nonkey attribute(s) that functionally determine(s) other nonkey attribute(s).

General Definitions: 2NF

Definition (2NF — Old)

A relation schema R is in 2NF if every nonprime attribute A in R is fully functionally dependent on the primary key of R .

General Definitions: 2NF

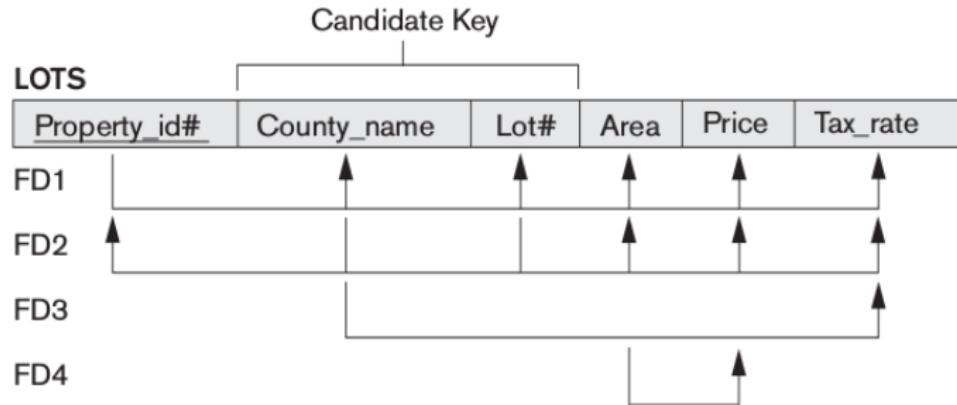
Definition (2NF — Old)

A relation schema R is in 2NF if every nonprime attribute A in R is fully functionally dependent on the primary key of R .

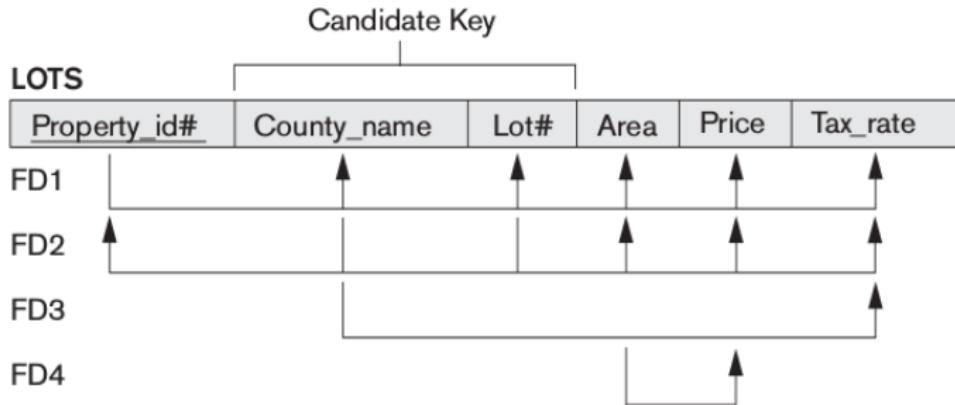
Definition (2NF — General)

A relation schema R is in second normal form (2NF) if every nonprime attribute A in R is not partially dependent on any key of R .

General Definitions: 2NF



General Definitions: 2NF

**LOTS1**

	<u>Property_id#</u>	County_name	Lot#	Area	Price
FD1					
FD2					
FD4					

LOTS2

	<u>County_name</u>	Tax_rate
FD3		

General Definitions: 3NF

Definition (3NF — Old)

A relation schema R is in 3NF if it satisfies 2NF and no nonprime attribute of R is transitively dependent on the primary key.

General Definitions: 3NF

Definition (3NF — Old)

A relation schema R is in 3NF if it satisfies 2NF and no nonprime attribute of R is transitively dependent on the primary key.

Definition (3NF — General)

A relation schema R is in third normal form (3NF) if, whenever a nontrivial functional dependency $X \rightarrow A$ holds in R , either (a) X is a superkey of R , or (b) A is a prime attribute of R .

General Definitions: 3NF

Definition (3NF — Old)

A relation schema R is in 3NF if it satisfies 2NF and no nonprime attribute of R is transitively dependent on the primary key.

Definition (3NF — General)

A relation schema R is in third normal form (3NF) if, whenever a nontrivial functional dependency $X \rightarrow A$ holds in R , either (a) X is a superkey of R , or (b) A is a prime attribute of R .

Definition (3NF — General Alternative)

A relation schema R is in 3NF if every nonprime attribute of R meets both of the following conditions:

- It is fully functionally dependent on every key of R .
- It is nontransitively dependent on every key of R .

General Definitions: 3NF

LOTS1

	Property_id#	County_name	Lot#	Area	Price
FD1					
FD2					
FD4					

FD1: Property_id# → County_name, Lot#, Area, Price
FD2: Property_id# → County_name, Lot#, Area
FD4: Property_id# → Price

LOTS2

	County_name	Tax_rate
FD3		

FD3: County_name → Tax_rate

General Definitions: 3NF

LOTS1

<u>Property_id#</u>	County_name	Lot#	Area	Price
FD1				
FD2				
FD4				

LOTS2

County_name	Tax_rate
FD3	

LOTS1A

<u>Property_id#</u>	County_name	Lot#	Area
FD1			
FD2			

LOTS1B

Area	Price
FD4	

BCNF

Definition (BCNF)

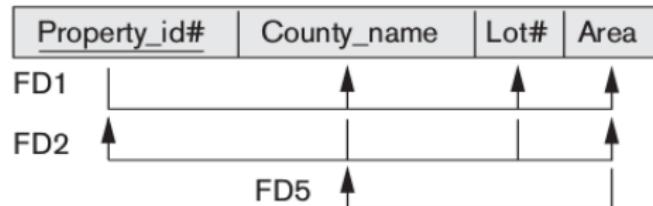
A relation schema R is in Boyce-Codd normal form (BCNF) if whenever a nontrivial functional dependency $X \rightarrow A$ holds in R , then X is a superkey of R .

BCNF

Definition (BCNF)

A relation schema R is in Boyce-Codd normal form (BCNF) if whenever a nontrivial functional dependency $X \rightarrow A$ holds in R , then X is a superkey of R .

LOTS1A



BCNF

Definition (BCNF)

A relation schema R is in Boyce-Codd normal form (BCNF) if whenever a nontrivial functional dependency $X \rightarrow A$ holds in R , then X is a superkey of R .

LOTS1A

Property_id#	County_name	Lot#	Area
FD1			
FD2			
			FD5

BCNF Normalization

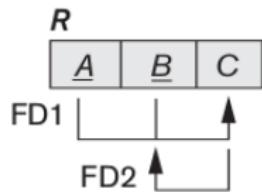
LOTS1AX

Property_id#	Area	Lot#

LOTS1AY

Area	County_name

3NF, but not BCNF



Example

TEACH

Student	Course	Instructor
Narayan	Database	Mark
Smith	Database	Navathe
Smith	Operating Systems	Ammar
Smith	Theory	Schulman
Wallace	Database	Mark
Wallace	Operating Systems	Ahamad
Wong	Database	Omiecinski
Zelaya	Database	Navathe
Narayan	Operating Systems	Ammar