

Redundancy Problem:

Problem ==> Redundancy!! ==> Anomalies

1. Insert=> Here address is repeating so, its redundant data. Now, if i want to insert the same data twice or thrice depending on emp's, Its a tedious task to insert every time and we may make some mistakes while typing the same data again again.
2. Update=> For same redundant data. Let say, i want to update address by adding postal code in the address column. Here i need to update that same postal code in every address column that has same address with id #1546. So specially here, Its a tedious task to update all columns and we may make some mistakes while entering the same data repeatedly.
3. Delete=> I want to delete the address which is holding the store id #1506. Here, i have to start deleting the same data repeatedly as long as data is repeating which is a tedious task again.

EmpID	FirstName	LastName	DoB	Position	Department	StoreID	Address
#20399	John	Ford	1996/2/12	Manager	HR	#1506	1200 W Dillon Rd, Louisville
#30123	Anne	Brand	2001/3/12	Intern	Marketing	#1546	1600 29th Street, Boulder
#12524	David	Biden	2000/2/20	Assistant	Sales	#1524	1271 Sheridan Blvd, Broomfield
#14517	William	Potter	2001/9/12	Senior Manager	HR	#1506	1200 W Dillon Rd, Louisville
#15214	Mary	Alexander	2001/9/12	Assistant	IT	#1524	1271 Sheridan Blvd, Broomfield
#11032	Rose	Smith	1999/1/21	Intern	IT	#1503	10003 Grant Street, Thornton
#20212	Julie	Smith	1977/12/1	Senior Manager	IT	#1503	10003 Grant Street, Thornton
#78123	Angela	White	1967/4/4	Senior Manager	HR	#1546	1600 29th Street, Boulder
#21342	John	Ford	1963/1/11	Manager	IT	#1546	1600 29th Street, Boulder

EmpID	FirstName	LastName	DoB	Position	Department	StoreID
#20399	John	Ford	1996/2/12	Manager	HR	#1506
#30123	Anne	Brand	2001/3/12	Intern	Marketing	#1546
#12524	David	Biden	2000/2/20	Assistant	Sales	#1524
#14517	William	Potter	2001/9/12	Senior Manager	HR	#1506
#15214	Mary	Alexander	2001/9/12	Assistant	IT	#1524
#11032	Rose	Smith	1999/1/21	Intern	IT	#1503
#20212	Julie	Smith	1977/12/1	Senior Manager	IT	#1503
#78123	Angela	White	1967/4/4	Senior Manager	HR	#1546
#21342	John	Ford	1963/1/11	Manager	IT	#1546

StoreID	Street	City	Zip
#1506	1200 W Dillon Rd	Louisville	80027
#1546	1600 29th Street	Boulder	80301
#1524	1271 Sheridan Blvd	Broomfield	80020
#1517	7125 W 88th Ave	Westminster	80021
#1548	16420 Washington Street	Thornton	80023
#1503	10003 Grant Street	Thornton	80029
#1502	5215 Wadsworth Blvd	Arvada	80002

No redundancy!!

First Normal Form:

NOT a relation – NF²

RegularUser

Email	Interest	Since Age	Birth Year	Current City	Salary
user1@gt.edu	Music	10	1985	Seattle	27,000
	Reading	5			
	Tennis	14			
user2@gt.edu	Blogging	13	1993	Austin	43,000
	Meditation	21			
	Surfing	19			
user3@gt.edu	Music	11	1967	San Diego	45,000
	Reading	6			
user4@gt.edu	DIY	18	1988	San Francisco	24,000
user9@gt.edu			1988	Las Vegas	24,000
user10@gt.edu			1986	Dallas	26,000
user12@gt.edu			1974	San Diego	38,000

Multi Values = not Supported.

Second Normal Form:

Basics of Second Normal Forms:

1. Identify Candidate key.
2. Find prime and Non prime attributes.
3. Find partial Dependencies.

==> Remove partial deps.

1.) What is Partial Dependencies?

Sol==> if any non-prime attribute depends upon subset of any prime attribute then, that particular non-prime attribute will leads to partial dependency.

Example: R: "A, B, C, D" Note: A, B==> Prime Attributes and C, D==> Non-Prime Attributes.

Ans: Here $\textcircled{C} \Rightarrow \textcircled{D}$: C= depends on b, here c= non prime attribute which depends upon ...subset of... prime att=b==> partial Dependency.

2.) Problem with Partial Dependency?

Sol==>

A	B
1	5
	3
2	

A, B==> Prime Attributes. Meaning ==> they are Candidate keys.

Sol==> 1. Pick partial dependency ==> $\textcircled{B} \Rightarrow \textcircled{C}$

2. See side table on 3rd row, the b cell is empty so, on empty b cell the functional dependence cannot give us \textcircled{C} because its empty, which leads to invalid query which is a data inconsistency problem.

Prime Attributes:

Finding Candidate Keys:

① R(A B C D E) Step 1:- Map Dependencies:

$\Rightarrow C \rightarrow D$
 $\Rightarrow D \rightarrow B$
 $\Rightarrow C \rightarrow A$

R(A B C D E)

Step 2:- Find incoming Attributes that doesn't have any "incoming edges" -

Ex:- $\Rightarrow C, E \rightarrow$ doesn't have any incoming edges.

Step 3:- Our Aim is to find single attribute (or "Minimal" attributes) which gives us "all Attributes" upon deriving closure set.

Ex:-

- i) $(C)^+ = \{C, A\} \rightarrow$ Not giving all attributes (X)
- ii) $(E)^+ = \{E, B\} \rightarrow$ (X)
- iii) $(CE)^+ = CE$
 - = ACE
 - = ACDE
 - = ABCDE

$\therefore (CE)^+ =$ giving all attributes
 $\therefore (CE)^+ =$ Candidate Key

Partial Dependencies:

2/17

Partial Dependence

Subset of PA

Prime Attributes

Non-Prime Attributes

iff NPA depends on Subset of PA
↳ Leads to (PD).

Problem

FD's

* $A, B = \text{PA's}$ (4) (= NPA)

$A, B \rightarrow C$

$B \rightarrow C$

Table

A	B	C
0	K ₅	x ₁
K ₁	0	x ₂
K ₂		x ₃
K ₃	K ₄	x ₄

$B \rightarrow C$

K₅ → x₁

0 → x₂

K₃ → x₄

(X) → (x₃x)

↳ This is Data
-inconsistency
Problem?

Third Normal Form:

3NF

Transitive Dependence

FD's

i) $A, B \rightarrow C$
 ii) $C \rightarrow D$ } $\Rightarrow A, B \rightarrow C \rightarrow D$

note:- 'C' will lead us into Problem.

Ex:-

	A	B	C	D
1	x ₁	x ₂	y ₁	z ₁
2	x ₃	x ₄	y ₂	z ₂
3	x ₅	x ₆	(x)	z ₃
4	x ₇	x ₈	y ₄	z ₄

A B C D
 x₁ x₂ → y₁ → z₁
 x₃ x₄ → y₂ → z₂
 x₅ x₆ → (x) → z₃

→ what if a user Request 4th Row on, knowingly.
 ∴ This will lead to Data Inconsistency Problem.

$R(\underline{A}, \underline{B}, C, D)$

$R_1(\underline{A}, \underline{B}, C)$
 $R_2(\underline{C}, \underline{D})$ } 3NF

Note: Here A, B ==> prime att's and C, D ==> non prime att's.
 The table dependencies says that, A, B ==> gives C and C ==> gives D which is called as----
Transitive Dependency.

But Here There's a big Problem associated with it, The problem is:
 ==> Indirectly A, B together can give D by taking the help with C.
 But Note:==> Here A, B ==> together prime keys.

So, Since C is a non-prime att, There may be a chance of being C as NULL.
 In case, C is NULL, A, B together cannot give D as per given Functional Dependencies.

∴ It Will Lead to DATA INCONSISTENCY PROBLEM!!!

↓

Solution for Above Problem.
 Removing Transitive Dependence