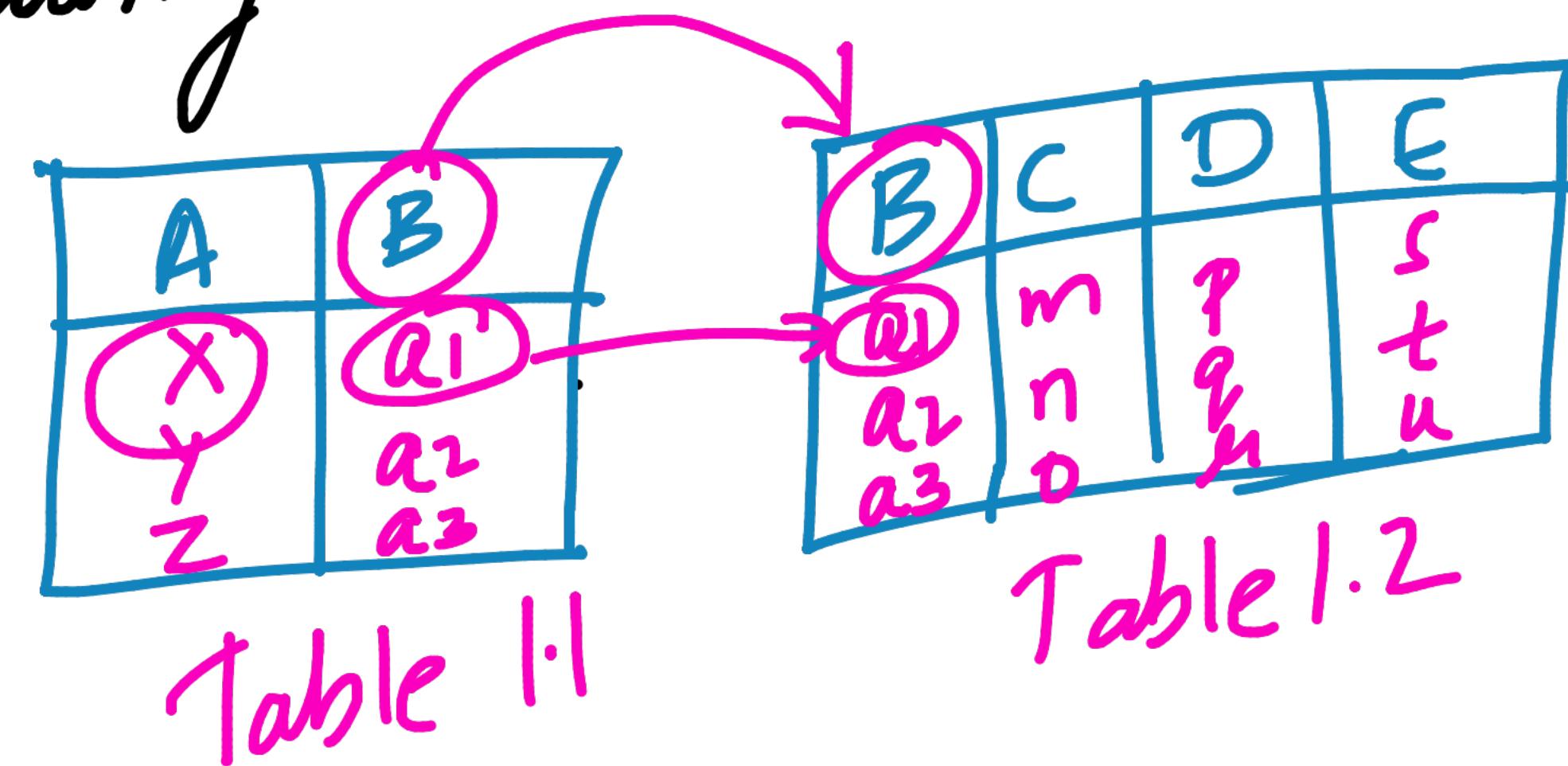


Normalization

Breaking down of table along the columns by following some rules.

A	B	C	D	E
X	a ₁₁	m	p	s
Y	a ₂₁	n	q	t
Z	a ₃₁	o	r	u

Table 1



Why do we need Normalization

~~1 million~~

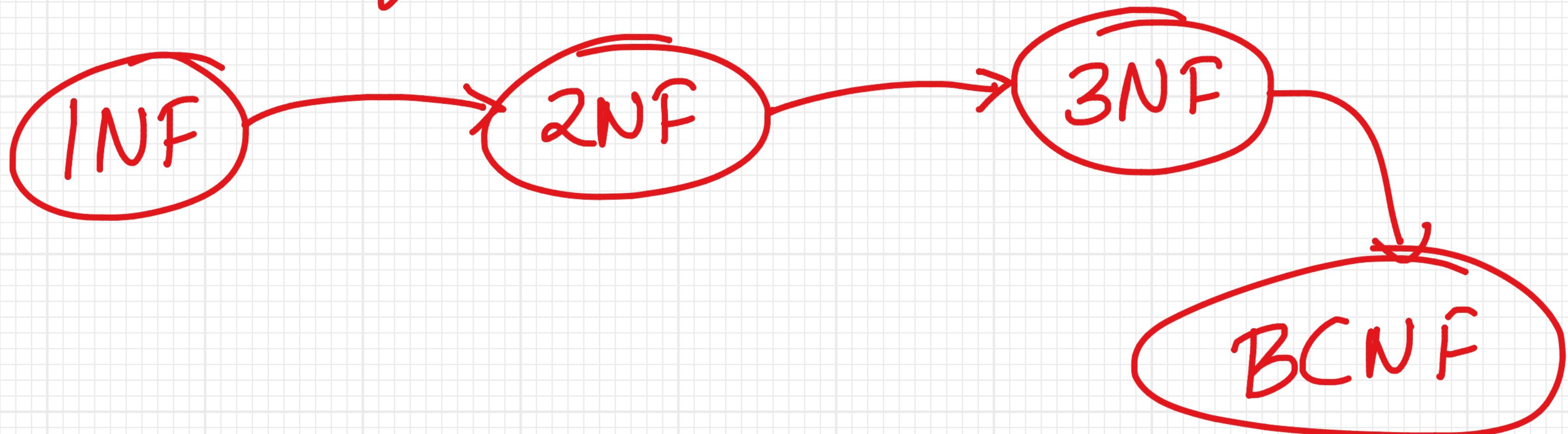
EmpID	EName	DepID	DepName	SupID	Group
101	Ashish	A	CS	Null	x
102	Khal	B	IT	101	y
103	Akslit	A	CS	101	z
104	Siddandi	B	IT	102	a
105	Pooja	A	CS	102	b

Dep ID	Dep Name
A	CS
B	IT

↑ ✗

- ① Redundancy \Rightarrow Repeated Data
- ② Anomalies \Rightarrow
 - ① Insertion \rightarrow Insert a New Dep.
 - ② Deletion \rightarrow what if I delete all the employees of Dep A
 - ③ Updation \rightarrow Change Dep Name.
CS \rightarrow Computer Science

Normalization Forms



Boyce - Codd
Normal Form

Functional Dependency

X	A	B	C
m	1	a	b
n	2	a	c
o	3	d	e
p	4	f	g

$$A \rightarrow BC$$

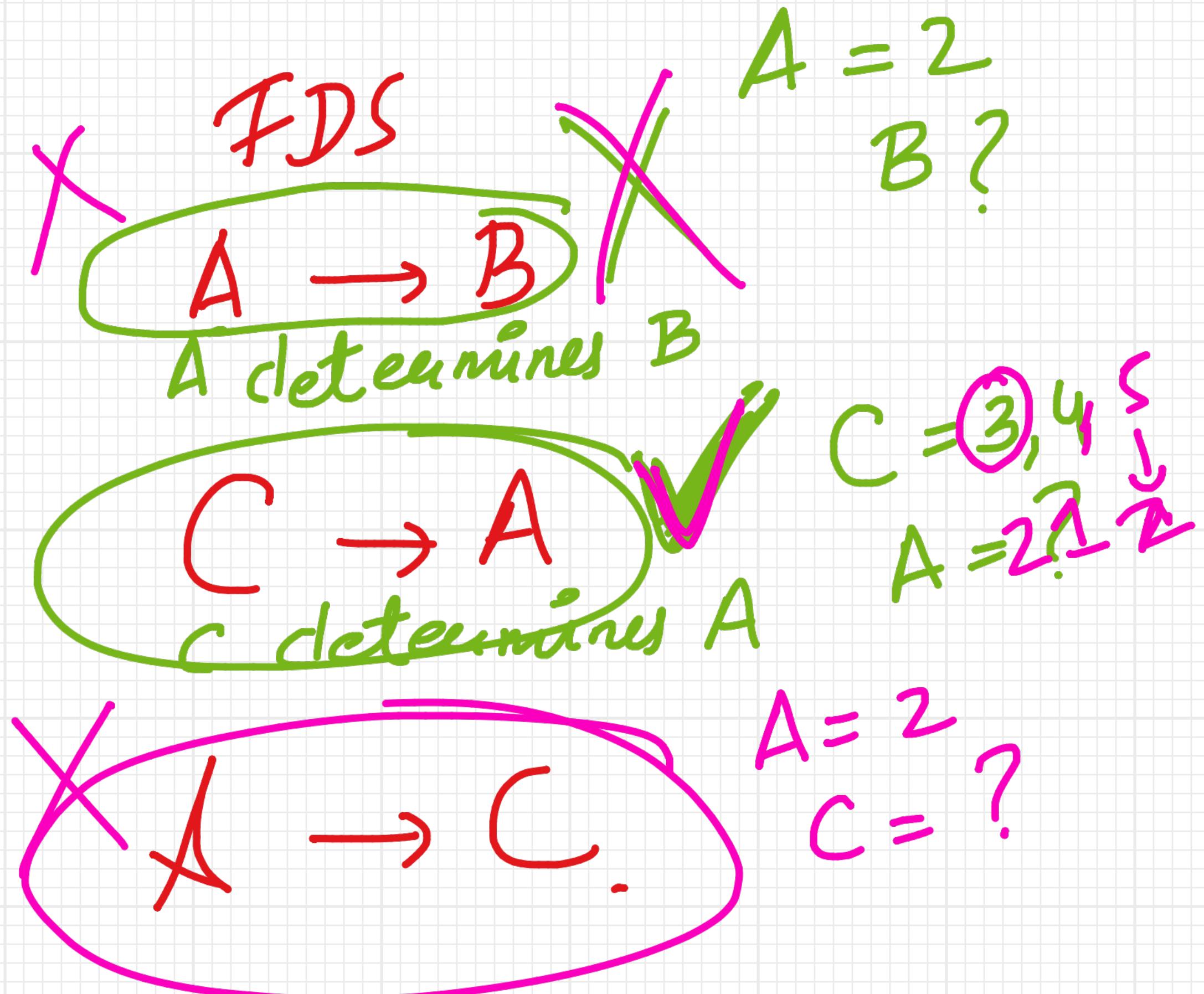
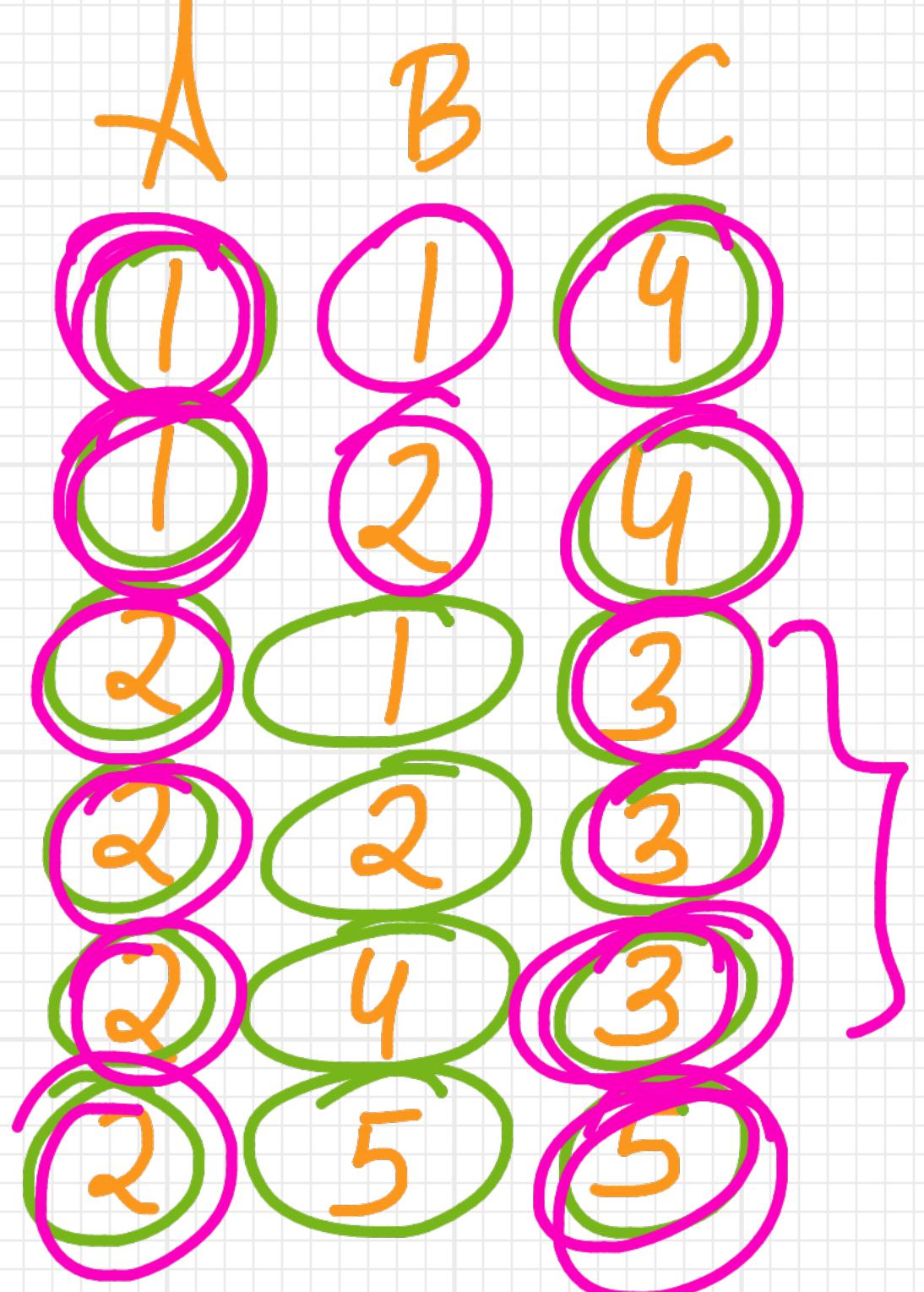
A determines BC

Given the value of A

B & C

you will be able to identify
without any confusion.

$$A = 1$$



Closure

A	B	C	D	E

$A^+ = \{B, D, A\}$

Closure of A

$B^+ = \{B, D\}$

~~A*~~

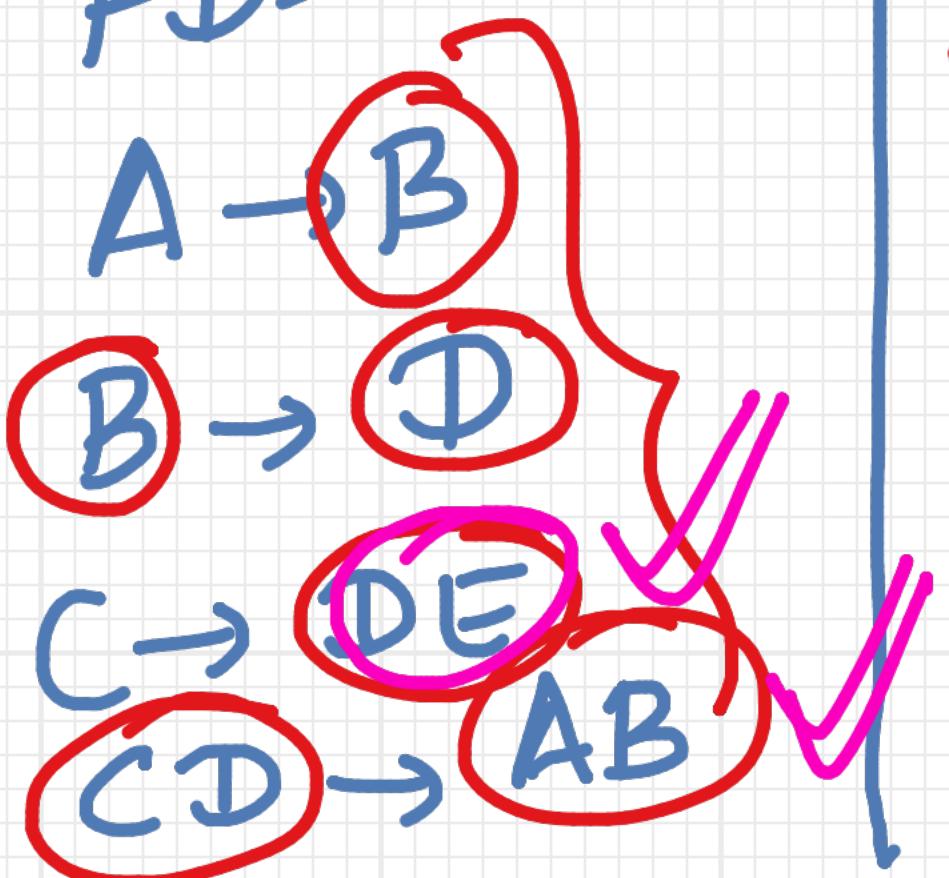
$C^+ = \{C, D, E, A, B\}$

$D^+ = \{D\}$

$E^+ = \{E\}$

$AB^+ = \{A, B, D\}$

FDS



$AC^+ = \{C, D, E, A, B\}$

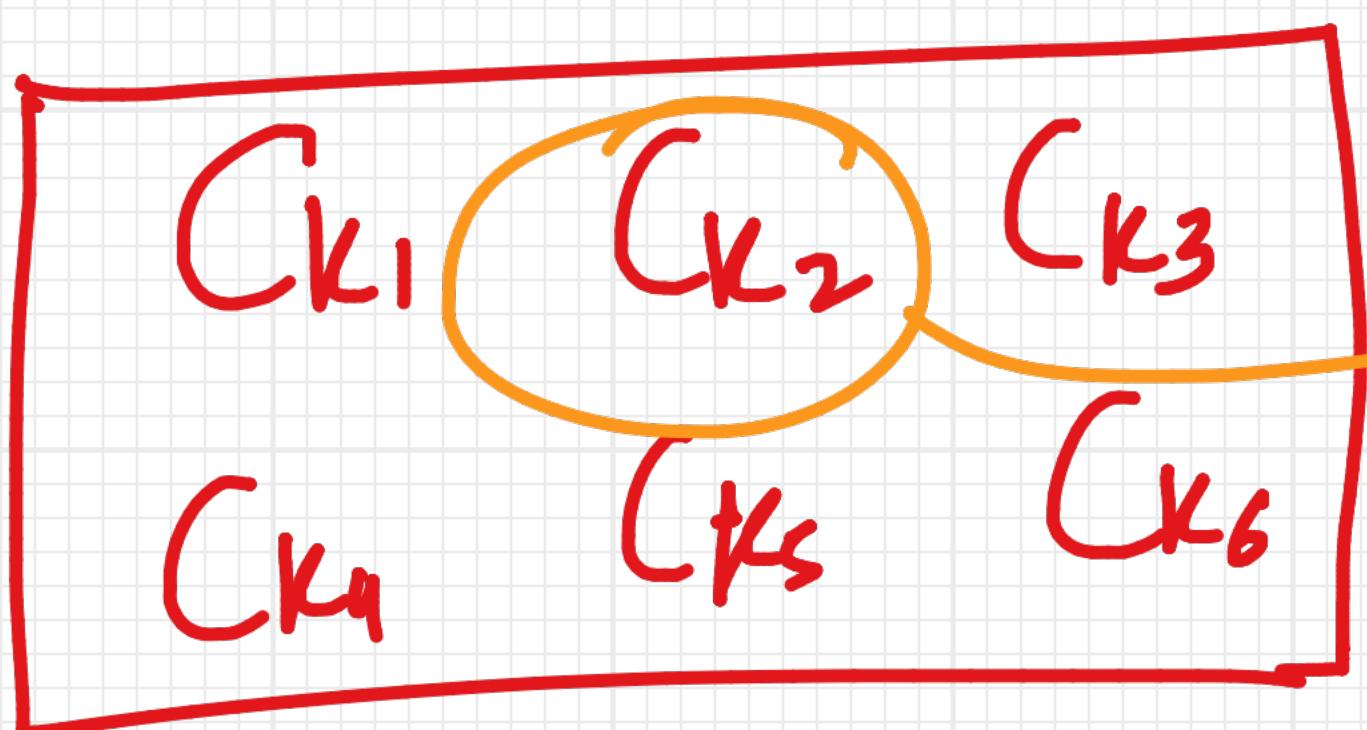
$AD^+ = \{A, D, B\}$

Stopped -

Note: The minimal closure which finds everything like C, is a candidate key.

* Candidate Key

Given an set of minimal columns we will be able to uniquely identify a row



One which we chose becomes a Primary Key

Aadhar No	DL No	- - -
--------------	----------	-------

Employee

$Ck_1 = \text{Aadhar No}$ ^{Pk}

$Ck_2 = \text{DL No}$

Note: There can
be only 1
Primary Key.

** Super Key can be so many
= Candidate Key + Waste Col.

Roll No	Name	Marks

Student

- $C_K = \text{Roll No}$ PK
- $S_{K_1} = \text{Roll No + Name}$
- $S_{K_2} = \text{Roll No + Marks}$
- $S_{K_3} = \text{Roll No + Name + Marks}$

~~1ST Normal Form~~

Name	Gender	Phone Nos
A	M	900- - - - 81 - - - -
B	F	76 - - - - 91 - - - -

~~Normalization~~

Name	Gender
A	M
B	F

~~1ST Normal F.~~

~~Multiple values~~

~~X. Atomicity~~

Name	P1	P2
A	900- - - -	81 - - - -
B	76 - - - -	91 - - - -

Reg* \Rightarrow Atomicity = 1 Col Contains only 1 data value

2nd Normal form \rightarrow ① 1NF ✓ → ② No Partial Dependency

* Given FDs

	A	B	C
1	a	C_1	$AB \rightarrow C$
2	a	C_1	$B \rightarrow C$
1	b	C_2	
2	b	C_2	
1	c	C_3	
2	c	C_3	
3	c	C_4	
4	c	C_5	

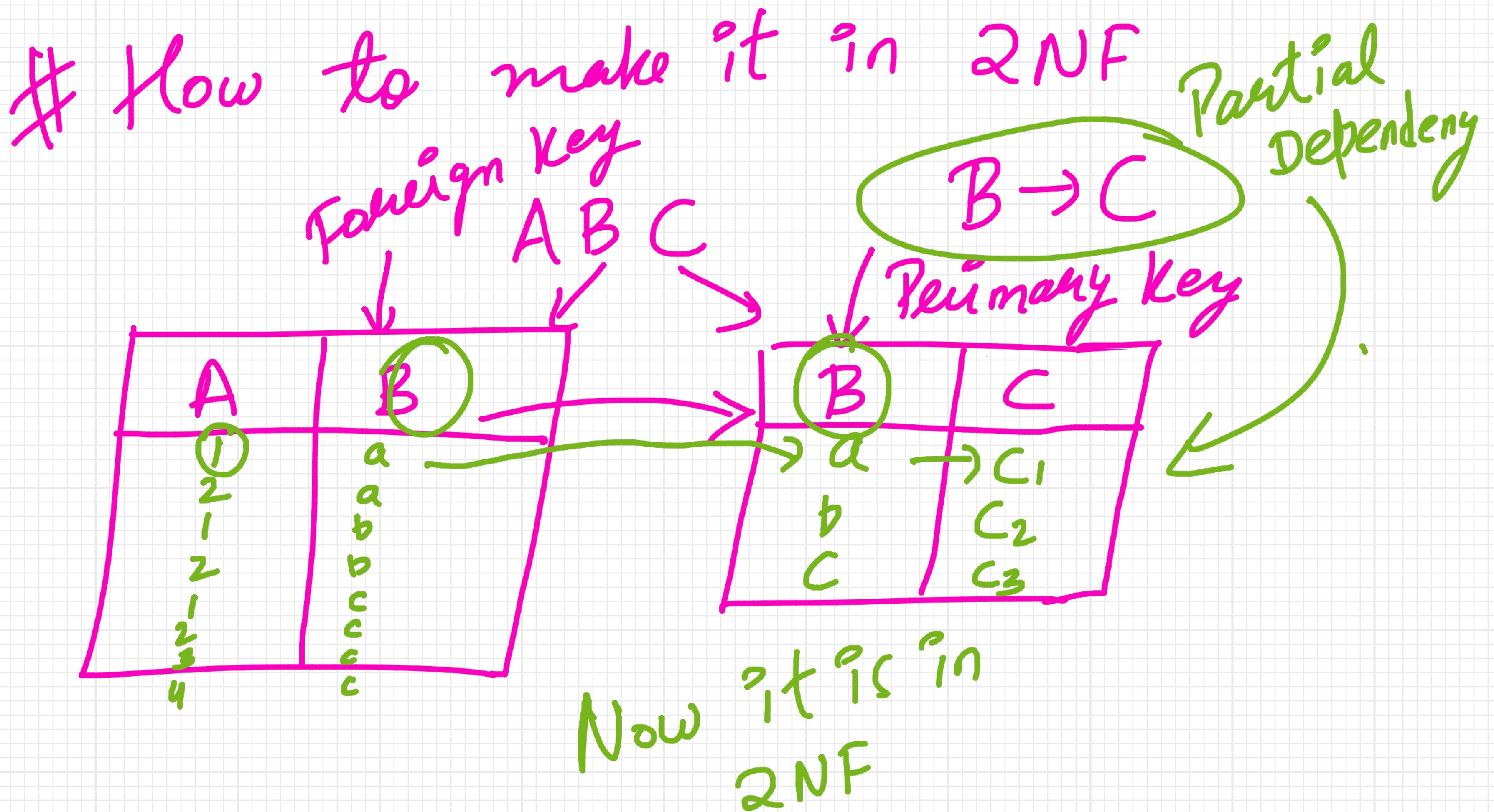
Part of C_k is determining something else.

$$A^+ = \{A\} \quad B^+ = \{B, C\}$$

$$C^+ = \{C\}, \quad AB^+ = \{A, B, C\}, \quad AC^+ = \{A, C\}$$

$$BC^+ = \{B, C\}, \quad ABC^+ = \{A, B, C\}$$

NOT in 2NF Super Key

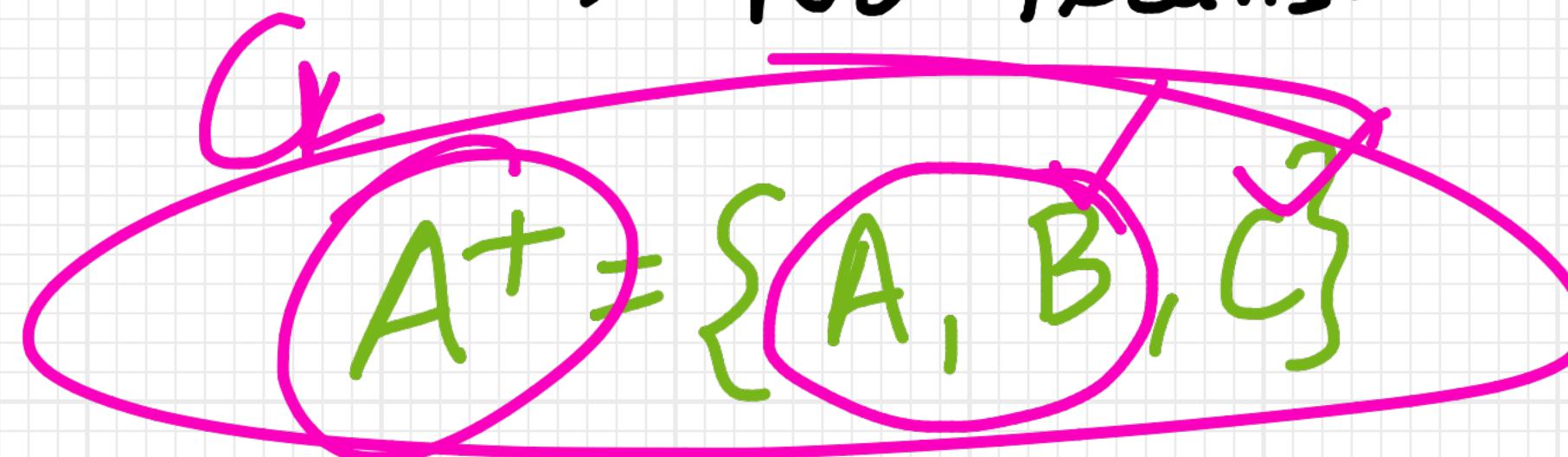


3rd Normal Form

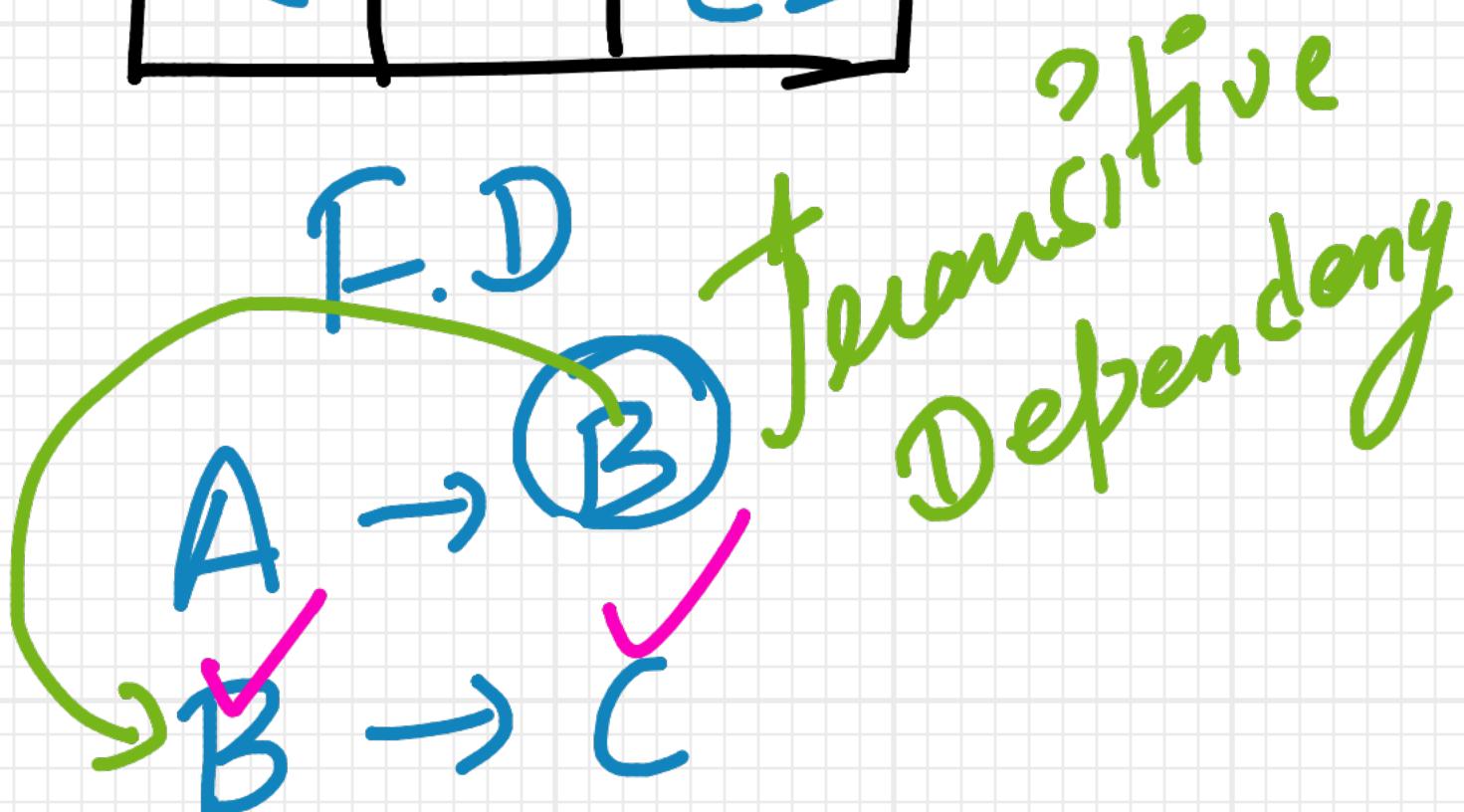
→ 2NF

→ No Transitive Dependency

A	B	C
1	a	c_1
2	b	c_2



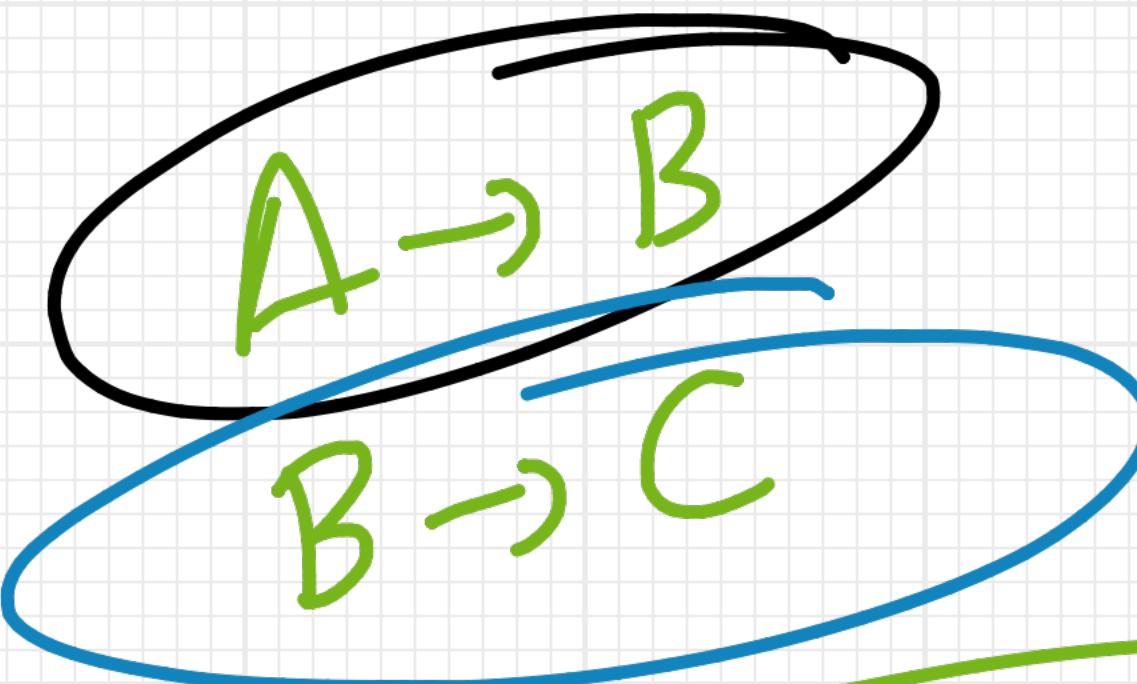
$$B \rightarrow C$$



A	B
1	a
2	b

B	C
a	c_1
b	c_2

	A	B	C
1	a	a	x
2	a	a	x
3	a	b	y
4	b	b	y
5	b	b	y
6	b	b	y



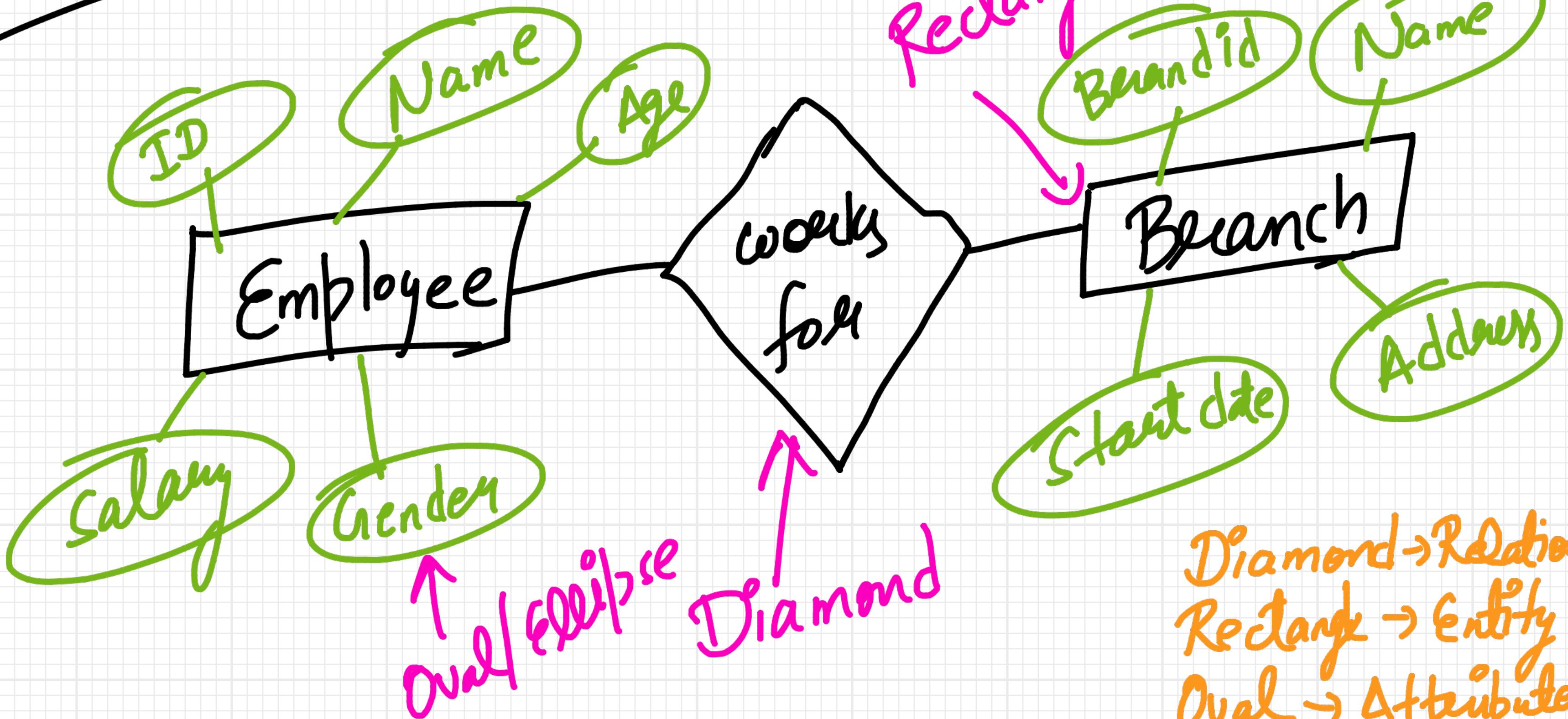
	A	B
1	a	a
2	a	a
3	a	b
4	b	b
5	b	b

3NF

	B	C
1	a	x
2	a	y
3	b	y

3NF

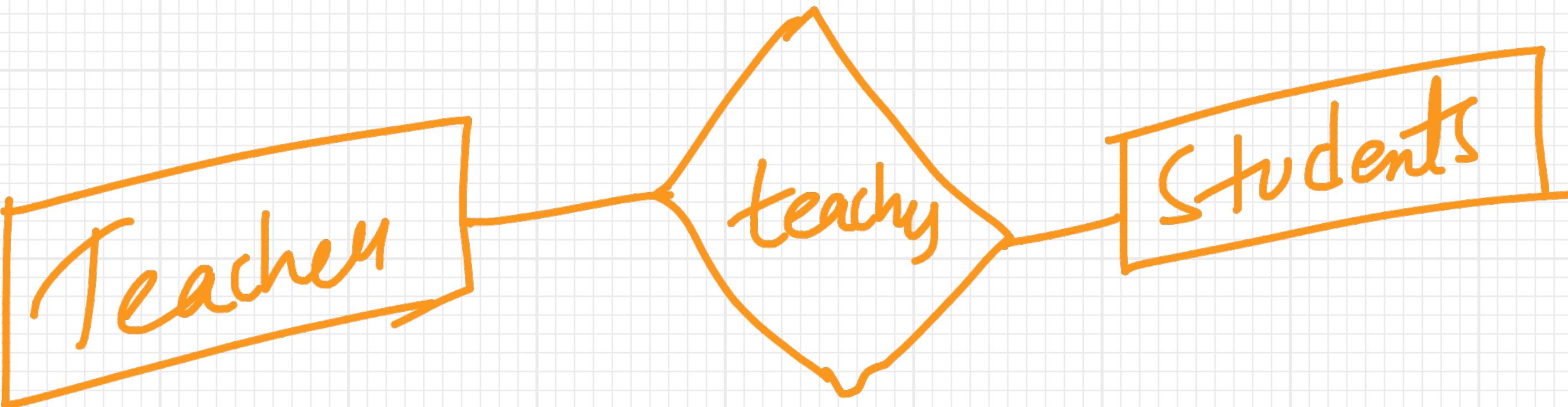
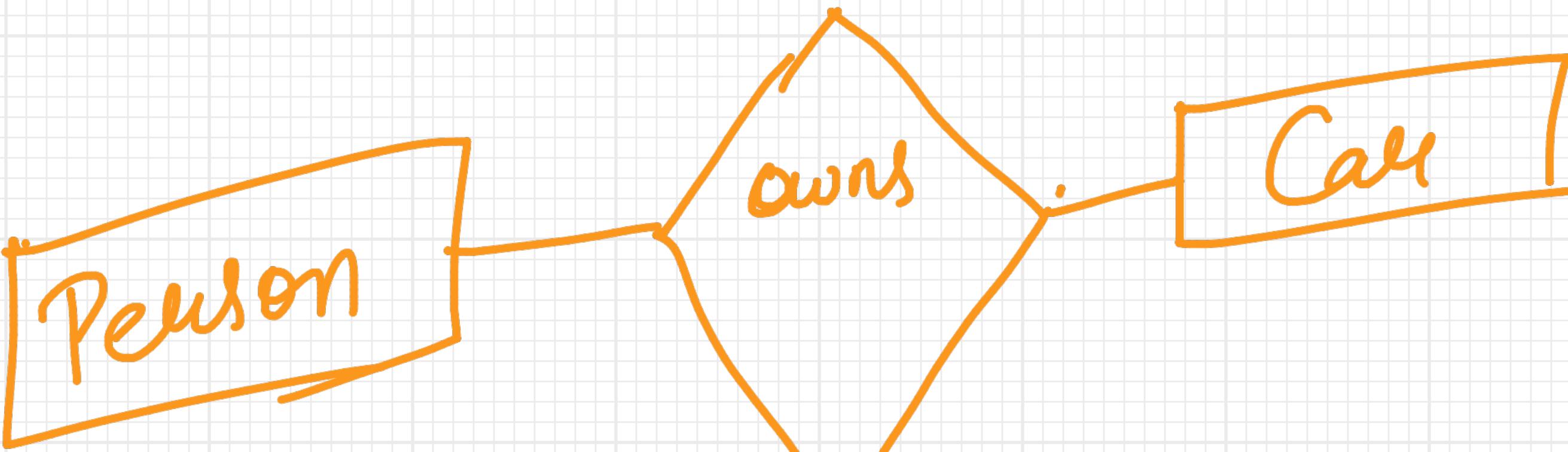
ER Diagrams



Diamond → Relations

Rectang → Entity

Oval → Attributes



Employee

emp_id	first_name	last_name	birth_date	sex	salary	super_id	branch_id
100	David	Wallace	1967-11-17	M	250,000	NULL	1
101	Jan	Levinson	1961-05-11	F	110,000	100	1
102	Michael	Scott	1964-03-15	M	75,000	100	2
103	Angela	Martin	1971-06-25	F	63,000	102	2
104	Kelly	Kapoor	1980-02-05	F	55,000	102	2
105	Stanley	Hudson	1958-02-19	M	69,000	102	2
106	Josh	Porter	1969-09-05	M	78,000	100	3
107	Andy	Bernard	1973-07-22	M	65,000	106	3
108	Jim	Halpert	1978-10-01	M	71,000	106	3

Branch

branch_id	branch_name	mgr_id	mgr_start_date
1	Corporate	100	2006-02-09
2	Scranton	102	1992-04-06
3	Stamford	106	1998-02-13

Works_With

emp_id	client_id	total_sales
105	400	55,000
102	401	267,000
108	402	22,500
107	403	5,000
108	403	12,000
105	404	33,000
107	405	26,000
102	406	15,000
105	406	130,000

Client

client_id	client_name	branch_id
400	Dunmore Highschool	2
401	Lackawana Country	2
402	FedEx	3
403	John Daly Law, LLC	3
404	Scranton Whitepages	2
405	Times Newspaper	3
406	FedEx	2

Branch Supplier

branch_id	supplier_name	supply_type
2	Hammer Mill	Paper
2	Uni-ball	Writing Utensils
3	Patriot Paper	Paper
2	J.T. Forms & Labels	Custom Forms
3	Uni-ball	Writing Utensils
3	Hammer Mill	Paper
3	Stamford Lables	Custom Forms

