Account

a/c no	Custno	Cname	Mobile	Email	Type	balance
100	1	Kishori	34445	dff@ffnj	demat	50000
101	1	Kishori	444444	dff@ffnj	saving	60000
102	1	Kishori	444444	dff@ffnj	current	60000
103	2	Rajan	3453	Asda@asd	saving	60000
	3	Revati	65867	asdas@lsdj		-----Not possible

If the data is not normalized, then in the data lot of redundancy is there. Hence it causes following problem

1. Insertion anomaly--- in the given table, I cannot keep the information of customer, who do not open the a/c, is called as insertion anamaly.
2. Updation anomaly- if the redundant data changes at one place we may not guarantee that changes will be up to date at all places, is called as updation anamoly.
3. Deletion anomaly---if any customer closes the a/c, and if that is the only a/c then along with account information, we may lose customer information also, is clled as deleteion anamaly.

account

a/c no	Custno	Type	balance	Mobile
100	1	demat	50000	44444
101	1	saving	60000	44444
102	1	current	60000	44444
103	2	saving	60000	3453

Customer

Custno	Cname	Mobile	Email
1	Kishori	5555	dff@ffnj
2	Rajan	3453	Asda@asd
3	Revati	65867	asdas@lsdj

Types of normalization ----- 1NF, 2NF, 3Nf, BCNF(3.5NF)

First normal form (1NF)

As per the rule of first normal form, an attribute (column) of a table cannot hold multiple values. It should hold only atomic values.

Stud no	Sname	marks

1	Rajas	56,78,67
2	Rajat	45,56
3	rohit	45

Since every row does not contain atomic value so it is not in 1NF

Stud no	Sname	marks
1	Rajas	56
1	Rajas	78
1	Rajas	67
2	Rajat	45
2	Rajat	56
3	rohit	45

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.

  Step1

  Check whether table is in 1NF

  Step 2:

  Find out primary key, and if its composite key then only checks for 2NF

All the attributes which are part of primary key are called as prime attributes, and remain attributes are called as non-prime attribute

The fees is dependent on teacher and the subject

Duration is dependent on subject

teacher_id	subject	teacher_age	fees	duration
111	Maths	38	25000	20 days
111	Physics	38	30000	30 days
222	Biology	38	50000	40 days
333	Physics	40	20000	30 days
333	Chemistry	40	35000	50 days

There should not be any partial functional dependency

Primary key ---- teacherid+subject

Prime attribute

Teacherid---→ teacher age

Subject--→duration

Teacherid+subject---→ fees

teacher_id	subject	fees
111	Maths	25000
111	Physics	30000
222	Biology	50000
333	Physics	20000
333	Chemistry	35000

teacher_id	teacher_age
111	38
222	38

333	40

subject	duration
Maths	20 days
Biology	40 days
Physics	30 days
Chemistry	50 days

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.

a-→b b-→c so a-→c

the following table are in 3NF

a/c no	Custno	Type	balance
100	1	demat	50000
101	1	saving	60000
102	1	current	60000
103	2	saving	60000

Transitive dependency is there so the table is not in 3NF

a/cno---→ custno---→ cname, mobile, email

Custno	Cname	Mobile	Email
1	Kishori	444444	dff@ffnj
2	Rajan	3453	Asda@asd

BCNF (3.5) NF

Usually BCNF is required if the table has more than one candidate key

It is an advance version of 3NF that's why it is also referred as 3.5NF.

BCNF is stricter than 3NF. A table complies with BCNF if it is in 3NF and for every functional dependency X->Y, X should be the super key of the table.

Rollno	name	voterid	age	Ward no
1	Rajan	3456	20	10
2	Revati	2345	25	23
3	Swati	4567	20	2
4	soham	5678	25	10

Ck={rollno, voterid}

Dependendecies

Rollno→ name, age

Voterid-> name, age,wardno

If rollno is primary key then we will not be able to get ward no

And since in our system we want to identify students by rollno we cannot keep voter id as a primary key

voterid	Ward no
3456	10
2345	23
4567	2
5678	10

Rollno	name	voterid	age
1	Rajan	3456	20
2	Revati	2345	25
3	Swati	4567	20
4	soham	5678	25

Normalize the given table:

One employee can work on many projects.

Proj Code	Proj Type	Proj Desc	Empno	Ename	Grade	Sal scale	Proj Join Date Time	Time allocated	
001	APP	LNG	46	JONES	A1	5	12/1/1998	24	
001	APP	LNG	92	SMITH	A2	4	2/1/1999	24	
001	APP	LNG	96	BLACK	B1	9	2/1/1999	18	
004	MAI	SHO	72	JACK	A2	4	2/4/1999	6	
004	MAI	SHO	92	SMITH	A2	4	5/5/1999	6	
005	MAI111	SHOrty	92	SMITH	A2	4	5/5/1999	6	

1.Check whether table is in 1 NF

Yes, it is in 1NF

2. check whether the table is in 2NF

Is in in 1NF ----yes

Is there any partial functional dependency

Candidate key {proj_code+Empno}

Prime attributes -→ project code, empno

Non prime attributes--→ project type, project desc, ename,grade,sal scale, proj join date, time allocated

Projcode---→ project type, proj desc

Empno---→ ename,grade, sal scale

Projcode + empno--→ proj join date,Time allocated

project

Proj Code	Proj Type	Proj Desc

001	APP	LNG
004	MAI	SHO
005	MAI111	SHOrty

Employee

Empno	Ename	Grade	Sal scale
46	JONES	A1	5
92	SMITH	A2	4
96	BLACK	B1	9
72	JACK	A2	4

Emp_proj

Proj Code	Empno	Proj Join Date Time	Time allocated	
001	46	12/1/1998	24	
001	92	2/1/1999	24	
001	96	2/1/1999	18	
004	72	2/4/1999	6	
004	92	5/5/1999	6	
005	92	5/5/1999	6	

4. Check whether the table is in 3NF
   Is it in 2NF -----yes
   Check whether transitive dependency is there

   Project table is in 3NF

Proj Code	Proj Type	Proj Desc
001	APP	LNG
004	MAI	SHO
005	MAI111	SHOrty

Employee table is not in 3NF currently so break it into 2 tables
Empno-→grade--→sal scale

Empno	Ename	Grade	Sal scale
46	JONES	A1	5
92	SMITH	A2	4
96	BLACK	B1	9
72	JACK	A2	4

Grade

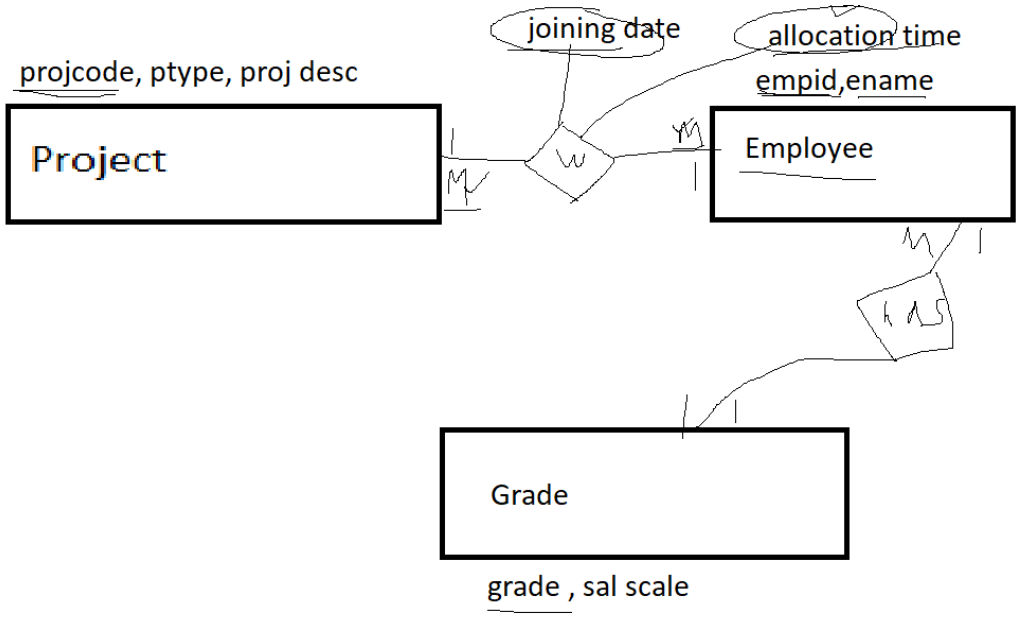
Grade	Sal scale
A1	5
A2	4
B1	9

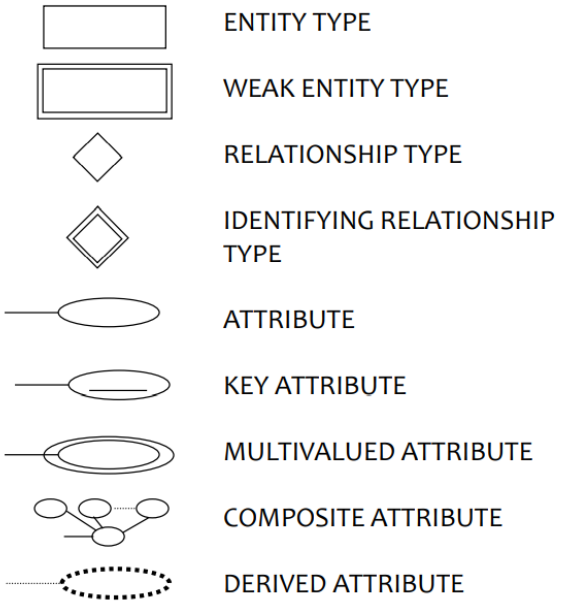
Employee

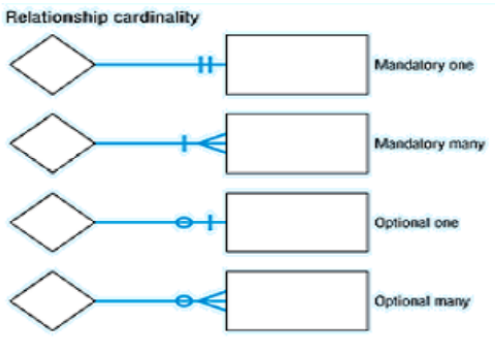
Empno	Ename	Grade
46	JONES	A1
92	SMITH	A2
96	BLACK	B1
72	JACK	A2

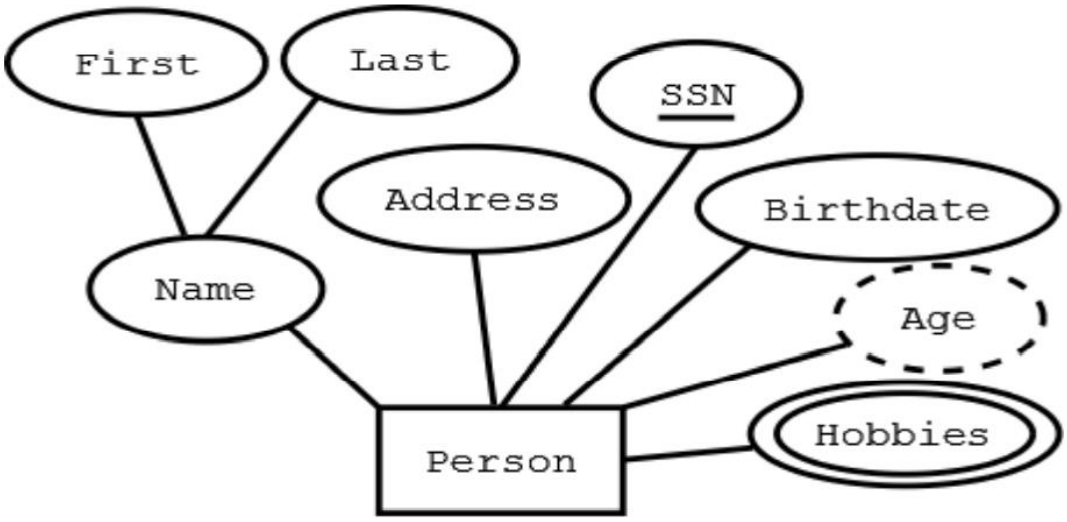
Emp_proj is already in 3 NF

Proj Code	Empno	Proj Join Date Time	Time allocated	
001	46	12/1/1998	24	
001	92	2/1/1999	24	
001	96	2/1/1999	18	
004	72	2/4/1999	6	
004	92	5/5/1999	6	
005	92	5/5/1999	6	









Example of E-R diagram

Drawbacks of Normalization

- Typically, in a normalized database, more joins are required to pull together information from multiple tables.
- Joins require additional I/O to process, and are therefore more expensive from a performance standpoint than single-table lookups.
- Additionally, a normalized database often incurs additional CPU processing. CPU resources are required to perform join logic and to maintain data and referential integrity.

Example 1:

- **roomno**
- **customerno**
- **booking date**
- **roomname**
- **roomlocation**
- **cname**
- **adress**
- **city**
- **state**
- **Email list**
- **phonenumber**
- **rateofroom**
- **mgrno**
- **managername**

Rules
1.One customer can book many rooms
2.One room can be booked by many customers on different date
3. Booking of a customer is managed by one manager
4. One manager can book many room for many customer

Rom no	custno	Bkdt	rname	rloc	cname	address	Phnum	Rate of room	Bkrt	Std rate	Mgr no	mname

1	1	10 nov								1	
1	1	20 nov								1	
2		10 nov									

Email table

custno	Email
1	aaa@sdfjsh
1	xxx@ldjkdj
2	xxdd@ashh
3	asjhdj@sdj

To check whether table is in 2NF

Primary key is Room no + bk dt

Rom no	custno	Bk dt	rname	rloc	cname	address	Ph num	Rate of room	Bk rt	Mgr no	mname

Room no----→rname,rloc,rate of room,mgr no,mname

Bk dt ---→

Room no + bk dt--→cust no,cname,address,ph number,bk rt room

Rom no	rname	rloc	Rate of room	Mgr no	mname

Room-booking

Rom no	custno	Bk dt	cname	address	Ph num	Bk rate

The room table is not in 3NF

Room no--→ mgr no---→ mname

Rom no	rname	rloc	Rate of room	Mgr no

Manager

Mgr no	mname

Room no+bk dt---→ cust no---→ cname,address, phone,state,city,

Room booking

Rom no	custno	Bk dt	Bk rate

customer

custno	cname	address	Ph num	City	state

Example 2

- Orderno
- Orderdate
- Itemno
- Qty
- Price
- Cname
- Custno
- Email
- Orderamt
- Salespersonno
- Salespersonname
- Locationid ----------location from where item dispatched
- Location name

One customer can place many order
One order contains many items
One order will be managed by one salesperson
One order belong to one customer
One order can be dispatched from different location

orderno	Orderdt	itemno	Ordered qty	buying price	Cust no	St item no	St item qty	St price	cname	email	amt	Sale per no	Sales name	Loc id	Loc Name
1	1 nov	1	2												
1	1 nov	2													
2		1	5												

Orderno---→order dt,cname,email,amt,sale per no,sales per name,
Item no---→

Orderno+item no--→ orde qty,buying price, loc id, loc name

Stock item no--→st item name,st item qty,st price

orderno	itemno	Ordered qty	buyingprice	Loc id	Loc name
1	1	2			
1	2				
2	1	5			

St item no	St itemqty	St price

orderno	Orderdt	cname	email	Cust no	amt	Sale per no	Sales name
1	1 nov						
1	1 nov						
2							

For coverting tables into 3NF

orderno	itemno	Ordered qty	buyingprice	Loc id
1	1	2		
1	2			
2	1	5		

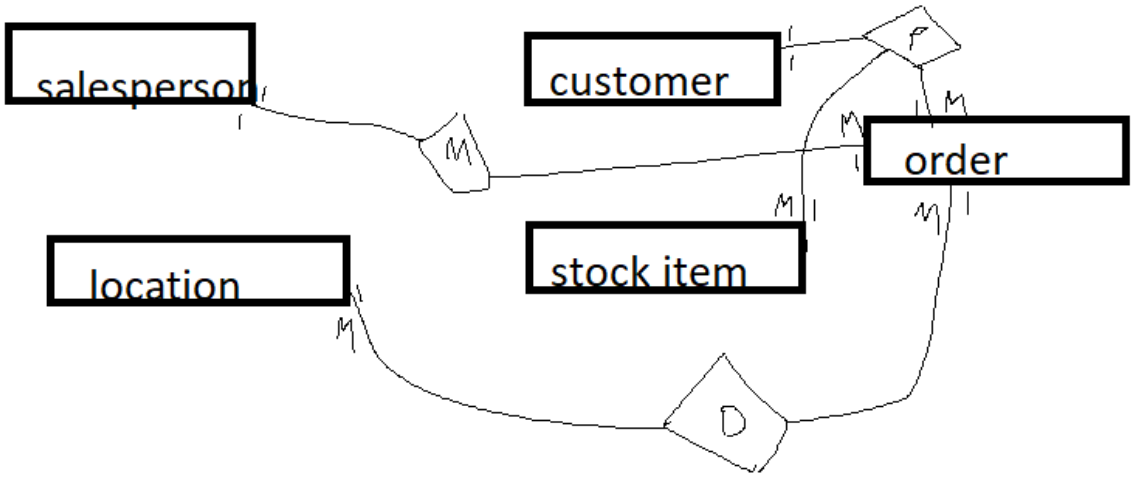
Loc id	Loc name

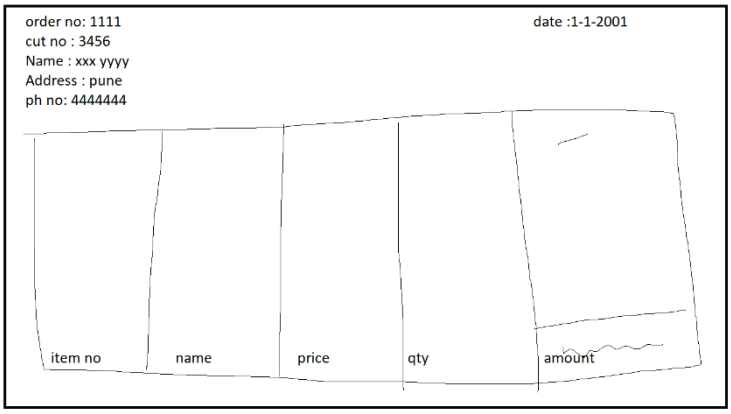
St item no	St itemqty	St price

orderno	Orderdt	Cust no	amt	Sale per no

cname	email	Cust no

Sale per no	Sales name





user name, password, name,address, email,ph num, itemno i stock, item name stock, qty , rate