

# Database Normalisation

# What is Normalization

**Database normalization** is the process of organizing data to minimize data redundancy (data duplication), which in turn ensures data consistency.

EmployeeName	Gender	Salary	DeptName	DeptHead	DeptLocation
Sam	Male	4500	IT	John	London
Pam	Female	2300	HR	Mike	Sydney
Simon	Male	1345	IT	John	London
Mary	Female	2567	HR	Mike	Sydney
Todd	Male	6890	IT	John	London

## Problems of Data Redundancy

1. Disk Space Wastage
2. Data Inconsistency
3. DML queries can become slow

## Normalized Table Design

DeptId	DeptName	DeptHead	DeptLocation
1	IT	John	London
2	HR	Mike	Sydney

EmployeeId	EmployeeName	Gender	Salary	DeptId
1	Sam	Male	4500	1
2	Pam	Female	2300	2
3	Simon	Male	1345	1
4	Mary	Female	2567	2
5	Todd	Male	6890	1

**Database normalization is a step by step process.** There are 6 normal forms, First Normal form (1NF) thru Sixth Normal Form (6NF). Most databases are in **third normal form (3NF)**. There are certain rules, that each normal form should follow.

# First Normal Form (1NF)

A table is said to be in 1NF, if

1. The data in each column **should be atomic**. No **multiple values**, separated by comma.
2. The table does not contain any **repeating column groups**
3. Identify each record uniquely using **primary key**.

## Non Atomic Employee Column

DeptName	Employee
IT	Sam, Mike, Shan
HR	Pam

## Problems of Non Atomic Columns

It is not possible to SELECT, INSERT, UPDATE and DELETE just one employee

## No Repeating Column Groups

DeptName	Employee1	Employee2	Employee3
IT	Sam	Mike	Shan
HR	Pam		

## Problems of Repeating Column Groups

More than 3 employees - Table structure change required  
Less than 3 employees - Wasted disk space

First Normal Form  
1 NF

Primary Key		Foreign Key	
DeptId	DeptName	DeptId	Employee
1	IT	1	Sam
2	HR	1	Mike
		1	Shan
		2	Pam

First Normal Form  
1 NF



# Second Normal Form (2NF)

A table is said to be in 2NF, if

1. The table meets **all the conditions of 1NF**
2. Move **redundant data** to a separate table
3. Create **relationship** between these tables using **foreign keys**.

EmpId	EmployeeName	Gender	Salary	DeptName	DeptHead	DeptLocation
1	Sam	Male	4500	IT	John	London
2	Pam	Female	2300	HR	Mike	Sydney
3	Simon	Male	1345	IT	John	London
4	Mary	Female	2567	HR	Mike	Sydney
5	Todd	Male	6890	IT	John	London

**Problems of Data Redundancy**

1. Disk Space Wastage
2. Data Inconsistency
3. DML queries can become slow

## Table Design in Second Normal Form

DeptId	DeptName	DeptHead	DeptLocation
1	IT	John	London
2	HR	Mike	Sydney

EmpId	EmployeeName	Gender	Salary	DeptId
1	Sam	Male	4500	1
2	Pam	Female	2300	2
3	Simon	Male	1345	1
4	Mary	Female	2567	2
5	Todd	Male	6890	1

# Third Normal Form (3NF)

A table is said to be in 3NF, if the table

1. Meets all the conditions of 1NF and 2NF

2. Does not contain columns (attributes) that are **not fully dependent** upon the **primary key**

EmpId	EmployeeName	Gender	Salary	AnnualSalary	DeptId
1	Sam	Male	4500	<del>54000</del>	1
2	Pam	Female	2300	<del>27600</del>	2
3	Simon	Male	1345	<del>16140</del>	1
4	Mary	Female	2567	<del>30804</del>	2
5	Todd	Male	6890	<del>82680</del>	1

EmpId	EmployeeName	Gender	Salary	DeptName	DeptHead
1	Sam	Male	4500	IT	John
2	Pam	Female	2300	HR	Mike
3	Simon	Male	1345	IT	John
4	Mary	Female	2567	HR	Mike
5	Todd	Male	6890	IT	John

EmpId	EmployeeName	Gender	Salary	DeptId
1	Sam	Male	4500	1
2	Pam	Female	2300	2
3	Simon	Male	1345	1
4	Mary	Female	2567	2
5	Todd	Male	6890	1

DeptId	DeptName	DeptHead
1	IT	John
2	HR	Mike

---

## Product

---

Name	Color	Supplier	Phone
Widget	Blue	Contoso	555-12345
Thingybob	Blue, Red	Northwind	555-54321
Knicknack	Red, Black	Contoso	555-12345

## Product

Name	Color	Supplier	Phone
Widget	Blue	Contoso	555-12345
Thingybob	Blue, Red	Northwind	555-54321
Knicknack	Red, Black	Contoso	555-12345



---

## Product

---

Name	Color	Supplier	Phone
Widget	Blue	Contoso	555-12345
Thingybob	Blue	Northwind	555-54321
Knicknack	Red	Contoso	555-12345
Thingybob	Red	Northwind	555-54321
Knicknack	Black	Contoso	555-12345



## Product

Name	Color	Supplier	Phone
Widget	Blue	Contoso	555-12345
Thingybob	Blue	Northwind	555-54321
Knicknack	Red	Contoso	555-12345
Thingybob	Red	Northwind	555-54321
Knicknack	Black	Contoso	555-12345

## Product

ProductID	Name	Color	Phone	Supplier
1	Widget	Blue	555-12345	1
2	Thingybob	Blue	555-54321	2
3	Knicknack	Red	555-12345	1
4	Thingybob	Red	555-54321	2
5	Knicknack	Black	555-12345	1

## Supplier

SupplierID	Supplier
1	Contoso
2	Northwind

## Product

ProductID	Name	Phone	Supplier
1	Widget	555-12345	1
2	Thingybob	555-54321	2
3	Knicknack	555-12345	1

## Color

ColorID	Color
1	Blue
2	Red
3	Black

## Supplier

SupplierID	Supplier
1	Contoso
2	Northwind

## Product

ProductID	Name	Phone	Supplier
-----------	------	-------	----------

1	Widget	555-12345	1
2	Thingybob	555-54321	2
3	Knicknack	555-12345	1

## Color

ColorID	Color
---------	-------

1	Blue
2	Red
3	Black

## Supplier

SupplierID	Supplier
------------	----------

1	Contoso
2	Northwind



## Product

ProductID	Name	Phone	Supplier
1	Widget	555-12345	1
2	Thingybob	555-54321	2
3	Knicknack	555-12345	1

## Color

ColorID	Color
1	Blue
2	Red
3	Black

## Supplier

SupplierID	Supplier
1	Contoso
2	Northwind

## ProductColor

Product	Color
1	1
2	1
2	2
3	2
3	3

## Product

ProductID	Name	Phone	Supplier
1	Widget	555-12345	1
2	Thingybob	555-54321	2
3	Knicknack	555-12345	1

## Color

ColorID	Color
1	Blue
2	Red
3	Black

## Supplier

SupplierID	Supplier
1	Contoso
2	Northwind

## ProductColor

Product	Color
1	1
2	1
2	2
3	2
3	3

## Product

ProductID	Name	Supplier
-----------	------	----------

1	Widget	1
2	Thingybob	2
3	Knicknack	1

## Color

ColorID	Color
---------	-------

1	Blue
2	Red
3	Black

## Supplier

SupplierID	Supplier	Phone
------------	----------	-------

1	Contoso	555-12345
2	Northwind	555-54321

## ProductColor

Product	Color
---------	-------

1	1
2	1
2	2
3	2
3	3