**NORMALIZATION**

**NORMALIZATION:**

1. **Normalization** is the process of splitting the bigger table into many small tables without changing its functionality.
2. It is generally carried out during the design phase of SDLC.

**Advantages:**

1) it reduces the redundancy (unnecessary repeatation of data)

2) avoids problem due to delete anamoly (inconsistency)

Normalization is a step-by-step process and in each step, we have to perform some activities.

**STEPS IN NORMALIZATION:**

1) 1NF – 1st Normal form

2) 2NF – 2nd Normal form

3) 3NF – 3rd Normal form

**1NF**

1. We should collect all the required attributes into 1 or more bigger entities.
2. We have to assume no 2 records are same (i.e, records should not be duplicated)
3. Identify the probable primary key

At the end of 1NF, our data looks like this,

|  |  |
| --- | --- |
| **COLLEGE** | |
|
| RegNo **- PK** | |
| Sname |  |
| Semester |  |
| DOB |  |
| MailID |  |
| Phone |  |
| BookNo - **PK** | |
| Bname |  |
| Author |  |
| DOI |  |
| DOR |  |
| Fine |  |

**2NF:**

To perform 2NF,

1. The tables have to be in 1NF
2. Here, we identify all the complete dependencies and move them separately into different tables.

At the end of 2NF, our data looks like this,

|  |  |  |  |
| --- | --- | --- | --- |
| **STUDENTS** | |  | **BOOKS** |
|  |
| RegNo - PK | |  | BookNo - PK |
| Sname | |  | RegNo - FK |
| Semester | |  | Bname |
| DOB | |  | Author |
| MailID | |  | DOI |
| Phone | |  | DOR |
|  |  |  | Fine |

**3NF:**

1. The table will have to be in 2NF
2. Here, we identify all the partial dependencies and move such columns to a separate table.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **STUDENTS** |  | **BOOKS** | |  | **LIBRARY** | |
|  |  |
| RegNo - PK |  | BookNo - PK | |  | BookNo - FK | |
| Sname |  | Bname | |  | RegNo - FK | |
| Semester |  | Author | |  | DOI | |
| DOB |  |  |  |  | DOR | |
| MailID |  |  |  |  | Fine | |
| Phone |  |  |  |  |  |  |

**Disadvantage of Normalization:**

1. The only minor disadvantage is we may have to write complex queries as we have more number of tables to be accessed.

**Denormalization** is the process of combining more than 1 smaller table to form 1 bigger table is called as denormalization.

**CODD rules** ( **Differentiates between DBMS & RDBMS** ):

1) should support NULL values

2) should support creation of relationship between tables

3) should support DDL, DML, TCL

4) should support constraints like PK, Unique, CHK

5) should support query techniques like sub – queries, joins, grouping etc.

**Oracle 9i Features(i means internet):**

1. TIMESTAMP datatype
2. SYSTIMESTAMP function
3. ANSI style joins
4. Renaming a column

**Oracle 10g features (g means grid)**

1. Recycle Bin

**ERD - Entity Relationship Diagram**

1. It is the pictorial representation of all the entities and their relationships (tables).

|  |
| --- |
| **STUDENTS** |
| RegNo - PK  Sname  Semester  DOB  MailID  Phone |

|  |
| --- |
| **STUDENTS \_ INTERNALS** |
| RegNo - FK  Sid  Marks |

|  |
| --- |
| **BOOKS** |
| BookNo - PK  BName  Author |

|  |
| --- |
| **LIBRARY** |
| BookNo - FK  RegNo - FK  DOI  DOR  Fine |