**Normalization**

* Normalization is the process of organizing the data in database.
* It is performed to reduce data redundancy in database. Data redundancy means having the same data at multiple places.
* In simpler terms, normalization means breaking down a large, complex table into smaller and simpler tables while maintaining data relationships.
* Normalisation Anomaly**:** Update, Delete, Insert.

**Types of Normal Forms:**

* **1NF: It is first normal from.** The data must be unique, single entry and no duplicate values

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Roll\_no | Branch | Languages |
| Samiksha | 21 | IT | Hindi, English |
| Pratiksha | 22 | CSE | English, Marathi |
| Maya | 23 | EE | Hindi |

The decomposition of the table into 1NF has shown below:

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Roll\_no | Branch | Languages |
| Samiksha | 21 | IT | Hindi |
| Samiksha | 21 | IT | English |
| Pratiksha | 22 | CSE | English |
| Pratiksha | 22 | CSE | Marathi |
| Maya | 23 | EE | Hindi |

* **2NF: It is second normal from** It should be in 1NF and not contain any partial dependency (primary key should not contain duplicate values).

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Roll\_no | Branch | Languages |
| Samiksha | 21 | IT | Hindi |
| Samiksha | 21 | IT | English |
| Pratiksha | 22 | CSE | English |
| Pratiksha | 22 | CSE | Marathi |
| Maya | 23 | EE | Hindi |

Here, roll\_no is the primary key, but it contains duplicate values. So, we will perform 2NF here,

**Separate table for no partial dependency:**

|  |  |  |
| --- | --- | --- |
| Name | Roll\_no | Branch |
| Samiksha | 21 | IT |
| Pratiksha | 22 | CSE |
| Maya | 23 | EE |

**Separate table for partial dependency:**

|  |  |
| --- | --- |
| Roll\_no | Languages Known |
| 21 | Hindi |
| 21 | English |
| 22 | English |
| 22 | Marathi |
| 23 | Hindi |

* **3NF: It is third normal from**. It must be in 2NF, and there should be no transition dependency.

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Roll\_no | Branch | Fee |
| Samiksha | 21 | IT | 15,000 |
| Pratiksha | 22 | CSE | 10,000 |
| Maya | 23 | EE | 20,000 |

**The decomposition of the table into 3NF has shown below:**

|  |  |  |
| --- | --- | --- |
| Name | Roll\_no | Branch |
| Samiksha | 21 | IT |
| Pratiksha | 22 | CSE |
| Maya | 23 | EE |

|  |  |
| --- | --- |
| Branch | Fee |
| IT | 15,000 |
| CSE | 10,000 |
| EE | 20,000 |

* **BCNF(Boyce Codd normal form):** It is advance version of 3NF. In every functional dependency A should depend on B, and A should be ‘SUPER KEY’ of a table.
* **4NF: It is forth normal from.** A relation will be in 4NF if it is in Boyce Codd's normal form and has no multi-valued dependency.
* **5N**F: **It is fifth normal from.** A relation is in 5NF. If it is in 4NF and does not contain any join dependency, joining should be lossless.