Question 1: Normalization

Apply Normalization on the below table and return the list of tables after 1NF, 2NF, 3NF, 4NF, 5NF( if applicable)

Normalize the table to 1NF, 2NF, and 3NF, explaining the steps you took at each normalization level.

Table - Employee Information

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Employee ID | Employee Name | Department | Project ID | Project Name | Start Date | End Date | Salary |
| 101 | John Doe | HR | 001 | Project A | 2023-01-15 | 2023-06-30 | 5000 |
| 101 | John Doe | HR | 002 | Project B | 2023-04-01 | 2023-08-31 | 5200 |
| 102 | Jane Smith | Marketing | 001 | Project A | 2023-02-01 | 2023-05-31 | 5500 |
| 103 | Mike Johnson | IT | 002 | Project B | 2023-03-10 | 2023-08-15 | 6000 |
| 103 | Mike Johnson | IT | 003 | Project C | 2023-06-15 | 2023-11-30 | 6200 |
| 104 | Sarah Brown | HR | 002 | Project B | 2023-06-15 | 2023-11-30 | 4800 |
| 105 | Robert Lee | Finance | 001 | Project A | 2023-06-15 | 2023-11-30 | 5200 |
| 106 | Lisa Wang | IT | 001 | Project A | 2023-06-15 | 2023-11-30 | 5800 |

Solution -

Table is already in 1st Normal form. Because -

* All values are atomic
* Each column has an unique name
* Columns have the values of same type

2nd Normal form -

* The table should be in 1NF
* All non-prime attributes should be fully functionally dependent on the primary key. In this case, the primary key is {Employee ID, Project ID}.
* Create a new table for Employee data: (Employee ID, Employee Name, Department).
* Create a new table for Project data: {Project ID, Project Name, Start Date, End Date}.
* The original table becomes: {Employee ID, Project ID, Start Date, End date, Salary}.

Table - Employee Data

|  |  |  |
| --- | --- | --- |
| Employee ID | Employee Name | Department |
| 101 | John Doe | HR |
| 102 | Jane Smith | Marketing |
| 103 | Mike Johnson | IT |
| 104 | Sarah Brown | HR |
| 105 | Robert Lee | Finance |
| 106 | Lisa Wang | IT |

Table - Projects

|  |  |
| --- | --- |
| Project ID | Project Name |
| 001 | Project A |
| 002 | Project B |
| 003 | Project C |

Table - Employee Information

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Employee ID | Project ID | Start Date | End date | Salary |
| 101 | 001 | 2023-01-15 | 2023-06-30 | 5000 |
| 101 | 002 | 2023-04-01 | 2023-08-31 | 5200 |
| 102 | 001 | 2023-02-01 | 2023-05-31 | 5500 |
| 103 | 002 | 2023-03-10 | 2023-08-15 | 6000 |
| 103 | 003 | 2023-06-15 | 2023-11-30 | 6200 |
| 104 | 002 | 2023-04-20 | 2023-07-31 | 4800 |
| 105 | 001 | 2023-05-05 | 2023-09-30 | 5200 |
| 106 | 001 | 2023-06-01 | 2023-12-31 | 5800 |

3rd Normal form -

Tables are in 3NF because -

* In 3NF, the table should be in 2NF.
* No transitive dependencies should exist.

This is the final normalized structure. Further normalization (4NF, 5NF) is not necessary in this case, as the table does not exhibit any multivalued dependencies or other complex relationships.