

## CP363 - Assignment 8

**Normalization** is a process of organizing data in a database to reduce data redundancy and improve data integrity. Normalization involves dividing a table into two or more tables and defining relationships between the tables. The goal of normalization is to create tables that are in a specific normal form.

**BCNF (Boyce-Codd Normal Form)** is a higher level of normalization than 3NF (Third Normal Form). A table is in BCNF if and only if every determinant in the table is a candidate key. In other words, a table is in BCNF if there are no non-trivial functional dependencies between attributes in the table.

To verify whether a table is in BCNF, you need to identify all the functional dependencies (FDs) that exist between the attributes in the table. A functional dependency is a relationship between two attributes in which the value of one attribute determines the value of another attribute. Once you have identified all the FDs in the table, you need to check if any of the determinants (the attributes on the left side of the FD) are not candidate keys. If there is at least one determinant that is not a candidate key, the table is not in BCNF.

If a table is not in BCNF, you need to decompose the table into two or more tables to eliminate the non-trivial FDs. Each new table should have a primary key, and you need to define relationships between the tables using foreign keys.

To modify a table to make it BCNF, you may need to add new tables and define relationships between the tables. If you add new data to the modified tables, you need to make sure that the data satisfies the BCNF requirements.

Based on the given functional dependencies(FDs), we can identify the primary keys and non-key attributes for each table.

- For the **Organizations table**, the primary key is organization\_id. The non-key attributes are org\_date\_created, org\_name, org\_address, org\_desc, org\_netWorth, and payment\_cycle.

|   | organization_id | org_date_created    | org_name          | org_address     | org_desc                        | org_netWorth | payment_cycle |
|---|-----------------|---------------------|-------------------|-----------------|---------------------------------|--------------|---------------|
| ▶ | 1               | 2023-02-27 00:34:03 | AmazonPrime       | 123 sre st      | New Random Desc                 | 1021.65      | 2023-06-22    |
|   | 2               | 2023-02-26 08:04:18 | Microsoft         | 45 erb st       | software development company    | 2554.50      | 2023-02-20    |
|   | 3               | 2023-02-26 08:04:18 | Netflix           | 256 bvl st      | internet entertainment services | 1200.45      | 2023-06-18    |
|   | 4               | 2023-02-26 08:04:18 | Apple             | 556 brown st    | software development company    | 1054.50      | 2023-01-06    |
|   | 5               | 2023-02-26 08:30:11 | DaBouf            | Trap            | Narcotics                       | 1000000.00   | 2069-04-20    |
|   | 6               | 2023-02-27 04:38:26 | Test Organization | 423 qwerqoinqwr | New Orgaization                 | 10000.00     | 2023-06-24    |

- For the **Employee table**, the primary key is employee\_id. The non-key attributes are emp\_date\_created, emp\_firstName, emp\_lastName, emp\_address, emp\_phone, emp\_username, emp\_email, emp\_password, emp\_type, emp\_hourlyWage, and emp\_salary.

|     | employee_id         | emp_date_created | emp_firstName | emp_lastName        | emp_address | emp_phone | emp_username     | emp_email  | emp_password | emp_type | emp_hourlyWage | emp_salary | organization | department | bank_id |
|-----|---------------------|------------------|---------------|---------------------|-------------|-----------|------------------|------------|--------------|----------|----------------|------------|--------------|------------|---------|
| ▶ 1 | 2023-02-26 08:04:18 | John             | Doe           | 123 Main St         | 13212412    | jdoe      | jdoe@email       | 123asdF1   | full_time    | NOKEY    | 10000.11       | 1          | 1            | 1          | 1       |
| 2   | 2023-02-26 08:04:18 | Mary             | Ann           | 178 Kelp St         | 22765839    | mann      | mann@email       | 568aerg0   | part_time    | 16.00    | NOKEY          | 1          | 1            | 2          | 2       |
| 3   | 2023-02-26 08:04:18 | Elon             | Musk          | 178 Texas St        | 22765983    | bigtwit   | musk@email       | 568aerg0   | full_time    | NOKEY    | 100000.11      | 1          | 2            | 3          | 3       |
| 5   | 2023-02-26 08:17:56 | John             | Doe           | 123 Main St         | 13212412    | jdoe23    | jdoe@emf34fail   | 123asdF1   | full_time    | NOKEY    | 10000.11       | 1          | 1            | NOKEY      | NOKEY   |
| 6   | 2023-02-26 08:20:09 | Harri            | Siva          | 23e23e              | 3223212     | oin23f    | noin@fovef       | 122ff      | full_time    | NOKEY    | 110000.00      | 1          | 1            | NOKEY      | NOKEY   |
| 7   | 2023-02-26 08:32:57 | El               | Chapo         | US Prison           | 911         | chapito   | fuck@dea.com     | nonoino3d2 | full_time    | NOKEY    | 100000.00      | 5          | 1            | NOKEY      | NOKEY   |
| 8   | 2023-02-26 20:52:04 | Test             | Employee      | Somewhere in a cave | 0           | nope      | no@pe.com        | ad12d      | part_time    | 20.00    | NOKEY          | 1          | 2            | NOKEY      | NOKEY   |
| 9   | 2023-02-27 04:35:00 | Test2            | Employee      | pomapomfrfq         | 123123124   | qpm1r12r  | apodna@gafqw.com | qwfqwfpm12 | full_time    | NOKEY    | 100000.00      | 1          | 1            | NOKEY      | NOKEY   |

Both tables are in at least 2NF (Second Normal Form) as there are no partial dependencies in the functional dependencies listed. However, we need to check for transitive dependencies to determine if they are in 3NF (Third Normal Form) or BCNF (Boyce-Codd Normal Form).

Assuming that emp\_type refers to the type of employment (e.g., full-time, part-time, etc.), we can see that there is a transitive dependency in the Employee table.

emp\_type → emp\_hourlyWage, emp\_salary, since the type of employment determines the wage/salary. To eliminate this transitive dependency, we can create a new table called EmploymentType with emp\_type as the primary key and emp\_hourlyWage and emp\_salary as non-key attributes.

After normalization, the Organizations table would remain unchanged. The Employee table would be split into two tables:

1. Employee (employee\_id, emp\_date\_created, emp\_firstName, emp\_lastName, emp\_address, emp\_phone, emp\_username, emp\_email, emp\_password)
2. EmploymentType (emp\_type, emp\_hourlyWage, emp\_salary).

Both new tables would be in 3NF as there are no transitive dependencies or partial dependencies. The Employee table would also be in BCNF since the determinant employee\_id is a candidate key and there are no non-trivial functional dependencies.

- For the **Performance table**, the primary key is performance\_id. The non-key attributes are perf\_date\_created, perf\_notes, date\_Achieved, and KPI\_Achieved.

|   | performance_id | perf_date_created   | perf_notes | date_Achieved | KPI_Achieved | appraiser_id | appraised_id | project_id |
|---|----------------|---------------------|------------|---------------|--------------|--------------|--------------|------------|
| ▶ | 27             | 2023-02-27 03:45:04 | completed  | 2022-04-20    | 50           | 1            | 2            | 3          |
| * | NULL           | NULL                | NULL       | NULL          | NULL         | NULL         | NULL         | NULL       |

- For the **Project table**, the primary key is project\_id. The non-key attributes are prj\_date\_created, prj\_name, prj\_desc, and KPI\_goal.

|   | project_id | prj_date_created    | prj_name | prj_desc       | KPI_goal |
|---|------------|---------------------|----------|----------------|----------|
| ▶ | 3          | 2023-02-26 00:00:00 | Test     | Random Test    | 20       |
|   | 6          | 2023-02-27 00:46:24 | NewProj  | Will Be Edited | 89       |
|   | 8          | 2023-03-23 00:00:00 | NewProj2 | Newfwef        | 0        |

Both tables are in at least 2NF (Second Normal Form) as there are no partial dependencies in the functional dependencies listed. However, we need to check for transitive dependencies to determine if they are in 3NF (Third Normal Form) or BCNF (Boyce-Codd Normal Form). There are no apparent transitive dependencies in either table. Thus, both tables are already in 3NF and BCNF.

**Therefore, no further normalization is required.**

- For the **Department table**, the primary key is Department\_id. The non-key attributes are dep\_date\_created, dept\_name, dept\_desc, and dept\_budget.

|   | department_id | dep_date_created    | dept_name | dept_desc                                   | dept_budget | organization_id | manager_id |
|---|---------------|---------------------|-----------|---|-------------|-----------------|------------|
| ▶ | 1             | 2023-02-26 08:04:18 | Marketing | Responsible in identifying customer         | 2000.00     | 1               | 1          |
|   | 2             | 2023-02-27 00:44:47 | Finance   | Acquiring and utilizing money for financing | 2030.00     | 1               | 3          |
|   | 4             | 2023-02-26 08:30:30 | Supply    | Sinolans                                    | 10000000.00 | 5               | NULL       |

- For the **Bank table**, the primary key is bank id. The non-key attributes are bnk\_date\_created, institute\_number, transit\_number, and account\_number.

|   | bank_id | bnk_date_created    | institute_number | transit_number | account_number |
|---|---------|---------------------|------------------|----------------|----------------|
| ▶ | 1       | 2023-02-26 08:04:18 | 001              | 11242          | 123123123123   |
|   | 2       | 2023-02-26 08:04:18 | 002              | 11363          | 11231442342256 |
|   | 3       | 2023-02-26 08:04:18 | 003              | 12363          | 1123144231256  |

- For the **Transactions table**, the primary key is transaction id. The non-key attributes are transaction\_date, wage, EI\_pay, vacation\_pay, bonus\_pay, overtime\_pay, and net\_pay.

|   | transaction_id | transaction_date | wage     | EI_pay | vacation_pay | bonus_pay | overtime_pay | net_pay | employee_id | bank_id |
|---|----------------|------------------|----------|--------|--------------|-----------|--------------|---------|-------------|---------|
| ▶ | 1              | 2023-02-16       | 100.00   | 0.00   | 0.00         | 0.00      | NULL         | 100.00  | 1           | 1       |
|   | 2              | 2022-01-09       | 100.00   | 0.00   | 0.00         | 50.00     | NULL         | 150.00  | 2           | 2       |
|   | 8              | 2023-02-26       | 12312.00 | 124.00 | 421.00       | 12.00     | NULL         | 412.00  | 1           | 1       |

Both the **Department and Bank tables** are in at least 2NF as there are no partial dependencies in the functional dependencies listed. However, we need to check for transitive dependencies to determine if they are in 3NF or BCNF.

For the **Department table**, there are no transitive dependencies. Thus, it is already in 3NF and BCNF.

For the **Bank table**, there is a transitive dependency between the bank id and the non-key attributes institute\_number, transit\_number, and account\_number. This means that the table is not in 3NF or BCNF.

To normalize the Bank table, we can split it into two tables:

Bank\_1 {bank\_id} -> {bnk\_date\_created}

Bank\_2 {bank\_id} -> {institute\_number, transit\_number, account\_number}

Now, both Bank\_1 and Bank\_2 tables are in 3NF and BCNF.

**Therefore, the Transactions table is already in 3NF and BCNF, and the Bank table is now in 3NF and BCNF after normalization.**

Here is the summary for all the given tables:

1. **Organizations:** The table has a primary key **organization\_id** and contains information about organizations including their creation date, name, address, description, net worth, and payment cycle. The table is in BCNF since there are no non-trivial functional dependencies where a determinant is not a superkey.
2. **Employee:** The table has a primary key **employee\_id** and contains information about employees including their creation date, first name, last name, address, phone, username, email, password, type, hourly wage, and salary. The table is in BCNF since there are no non-trivial functional dependencies where a determinant is not a superkey.
3. **Performance:** The table has a primary key **performance\_id** and contains information about performance including the creation date, notes, date achieved, and KPI achieved. The table is in BCNF since there are no non-trivial functional dependencies where a determinant is not a superkey.
4. **Project:** The table has a primary key **project\_id** and contains information about projects including the creation date, name, description, and KPI goal. The table is in BCNF since there are no non-trivial functional dependencies where a determinant is not a superkey.
5. **Department:** The table has a primary key **department\_id** and contains information about departments including the creation date, name, description, and budget. The table is in BCNF since there are no non-trivial functional dependencies where a determinant is not a superkey.
6. **Bank:** The table has a primary key **bank\_id** and contains information about banks including the creation date, institute number, transit number, and account number. The table is in BCNF since there are no non-trivial functional dependencies where a determinant is not a superkey.

7. **Transactions:** The table has a primary key **transaction\_id** and contains information about transactions including the transaction date, wage, EL pay, vacation pay, bonus pay, overtime pay, and net pay. The table is in BCNF since there are no non-trivial functional dependencies where a determinant is not a superkey.