**CS614 Assignment #1 Spring 2024**

Name: Abdul Rehman

VU-ID: BC220424444

**Question #1:**

Table 1: StaffBranch given below reflects a specific form of normalization.

**Table 1: StaffBranch**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| StaffNo. | Name | position | Salary | BranchNo. | BranchAddress | TelNo. |
| S17005 | Rizwan Ahmed | Manager | 300000 | B001 | 7th – st. Al Anayat Co. G-11 Islamabad. | 302-3332-2 |
| S15002 | Arifa Javed | Assistant | 80000 | B001 | 7th – st. Al Anayat Co. G-11 Islamabad. | 302-3332-2 |
| S23361 | Ahmed Ali | Manager | 300000 | B002 | City Center Plaza, Karachi | 420-5847-5 |
| S40021 | Salman Ahmed | Assistant | 750000 | B002 | City Center Plaza, Karachi | 420-5847-5 |
| S00129 | Hafza Khan | Supervisor | 90000 | B003 | Naz Digital Plaza, Lahore. | 205-2584-3 |
| S01352 | Ameer Ali | Manager | 200000 | B004 | Yusif plaza Rawalpindi | 305-4561-9 |

You are required to identify the normalization form(NF) of the given table and also transform the given table into the 3rd normalization form (3NF) and provide the resultant tables.

**Note:** To convert in 3NF, first you need to convert it in 2NF, and resultant tables of 3NF are only required.

**Solution:**

According to the given scenario, we need these steps to figure out the normalization form (NF) of the table and change it to the 3rd normalization form (3NF).

**Normal Forms.**

**First Normal Form (1NF):** Check if each row has a unique identifier and if each column holds single values. If yes, it's in 1NF.

The table has a unique ID for each staff member (StaffNo.) and each column contains single values, so it meets 1NF requirements.

**Second Normal Form (2NF):** Ensure all non-key attributes are fully dependent on the entire primary key.

Attributes like Position, Salary, BranchAddress and TelNo. depend only on StaffNo., the primary key. So, it's in 2NF.

**Third Normal Form (3NF):** Make sure there are no transitive dependencies; every non-key attribute depends only on the primary key.

But we have BranchAddress and TelNo. depending on BranchNo., which isn't the primary key. To fix this, we need to break down the table further.

**Convert to 3rd Normal Form (3NF):**

Now, split the given table into smaller tables to remove transitive dependencies.

**Table 1: Staff**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **StaffNo.** | **Name** | **position** | **Salary** | **BranchNo.** |
| S17005 | Rizwan Ahmed | Manager | 300000 | B001 |
| S15002 | Arifa Javed | Assistant | 80000 | B001 |
| S23361 | Ahmed Ali | Manager | 300000 | B002 |
| S40021 | Salman Ahmed | Assistant | 750000 | B002 |
| S00129 | Hafza Khan | Supervisor | 90000 | B003 |
| S01352 | Ameer Ali | Manager | 200000 | B004 |

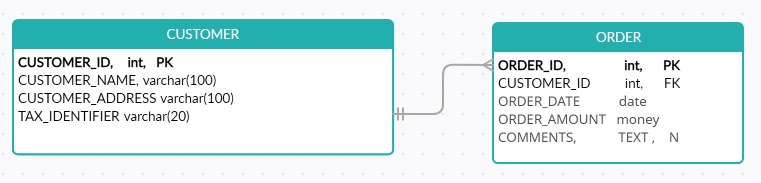
**Table 2: Branch**

|  |  |  |
| --- | --- | --- |
| **BranchNo.** | **BranchAddress** | **TelNo.** |
| B001 | 7th – st. Al Anayat Co. G-11 Islamabad. | 302-3332-2 |
| B001 | 7th – st. Al Anayat Co. G-11 Islamabad. | 302-3332-2 |
| B002 | City Center Plaza, Karachi | 420-5847-5 |
| B002 | City Center Plaza, Karachi | 420-5847-5 |
| B003 | Naz Digital Plaza, Lahore. | 205-2584-3 |
| B004 | Yusif plaza Rawalpindi | 305-4561-9 |

Now, our tables are in the 3rd Normal Form (3NF). Each table has its own theme, and there are no unnecessary dependencies.

**Question #2:**

Consider the following information about CUSTOMER and ORDER tables.



*Figure 1: CUSTOMER\_ORDER Relationship*

**Table1 (Master):**

Header Size of the CUSTOMER table: = 60 Bytes

Number of records stored in CUSTOMER table: = 70000

**Table2 (Detail):**

Header size ORDER table: = 70 Bytes

Number of records stored in ORDER table: = 5000000

Suppose you have applied the pre-joining De-Normalization technique on the ‘CUSTOMER’ and ‘ORDER’ tables.

Calculate the size of the resultant table in Megabytes (MB), Gigabytes (GB), and Terabytes (TB). The reference column in these two tables is 8 bytes.

Hint: There would be 5000000 records in the De-normalized table.

**Solution:**

According to the given question we have the following given data:

CUSTOMER Table includes:

* Header Size = 60 Bytes.
* Number of Records = 70000.

ORDER Table includes:

* Header Size = 70 Bytes.
* Number of Records = 5000000.

Reference column in these 2 tables = 8 bytes.

Now, to find the size of the resulting table after pre-joining de-normalization, combine data from CUSTOMER and ORDER tables in each record, considering header sizes and the reference column. Then, we will multiply by the total number of records.

**Calculate the size of a single record in the de-normalized table.**

= Header Size of CUSTOMER table + Header Size of ORDER table - Size of Reference Column = (60 + 70 - 8) Bytes = 122 Bytes.

**Total size of all records in the de-normalized table.**

= Size of a single record \* Number of records = 122 Bytes/record \* 5,000,000 records.

**Convert the total size into MB, GB, and TB.**

Size of Resultant Table(in bytes) = Header Size of Resultant Table x Number of records in Resultant Table.

Size(in bytes) = 122 bytes x 5000000 records = 610000000 bytes

**Convert the Size into Megabytes(MBs), Gigabytes(GBs), and Terabytes(TBs).**

1 Megabyte(MB) = 1000000 bytes.

Size in Megabytes(MB) = 610000000 bytes / 1000000 = 610 MB.

1 Gigabyte(GB) = 1000000000 bytes.

Size in Gigabytes(GB) = 610000000 bytes / 1000000000 = 0.61 GB.

1 Terabyte(TB) = 1000000000000 bytes.

Size in Terabytes(TB) = 610000000 bytes / 1000000000000 = 0.00061 TB.

Now, the size of the resultant de-normalized table after pre-joining the ‘CUSTOMER’ and ‘ORDER’ table is:

610 Megabytes(MBs), 0.61 Gigabytes(GBs), and 0.00061 Terabytes(TBs).