CSCI 5408

DATA MANAGEMENT AND WAREHOUSING

LAB-2: Data Modelling Process

Banner ID: B00982253

Table of Contents

1. INVOICE Table to 1NF……... 3
2. INVOICE Table to 2NF……... 5
3. INVOICE Table to 3NF……... 8
4. De-Normalization……. 12

1: Normalize the INVOICE table to 1NF

Current Invoice Table:-

Table 1 : Invoice Table as provided

InvoiceID	Date	Branch	Branch Location	City	Customer Type	Customer Name	Gender	Product	Unit Price	Quantity
1101	2/5/2024	1,004, 008	Quinpool, Windsor, Dartmouth	Halifax	Member, Normal	Alex, Rohan, Mark	Female , Male	Phone, Battery, Cover	200, 50, 25	1, 2, 5
2104	3/3/2024	3, 007, 008	Dowrey Street, Abbey Lane, GlenView	Toront o	Member, Normal	Suzan, Carla, Andreq	Female , Male	Code, Power bank	10, 35	20, 45

Note: I take one assumption from the above table as there is one ambiguity present in the given table.

- The branch code is only unique within one particular city, so it may be possible that the same branch code is assigned to different stores present in different cities. For example, branch code 008 is the same for Dartmouth and GlenView which are present in Halifax and Toronto respectively.

Explanation:-

- For a table to be in 1NF, there must be only one value present per column.
- Here we see that Branch, BranchLocation, CustomerType, CustomerName, Gender Product, UnitPrice, and Quantity have multiple values in for one entry.
- After normalizing to 1NF, we can see that InvoiceID and Product combining uniquely identify each row.

After 1st Normal Form:-

Table 2 : Invoice Table In 1NF

InvoiceID	Date	Branch	Branch Location	City	Customer Type	Customer Name	Gender	Product	Unit Price	Quantity
1101	2/5/2024	1	Quinpool	Halifax	Member	Alex	Female	Phone	200	1
1101	2/5/204	004	Windsor	Halifax	Member	Rohan	Male	Battery	50	2
1101	2/5/204	008	Dartmouth	Halifax	Normal	Mark	Female	Cover	25	5
2104	3/3/2024	3	Dowrey Street	Toronto	Member	Suzan	Female	Code	10	20
2104	3/3/2024	007	Abbey Lane	Toronto	Member	Carla	Female	Code	10	20
2104	3/3/2024	008	GlenView	Toronto	Normal	Andreq	Male	Power bank	35	45

Dependency Diagram:-

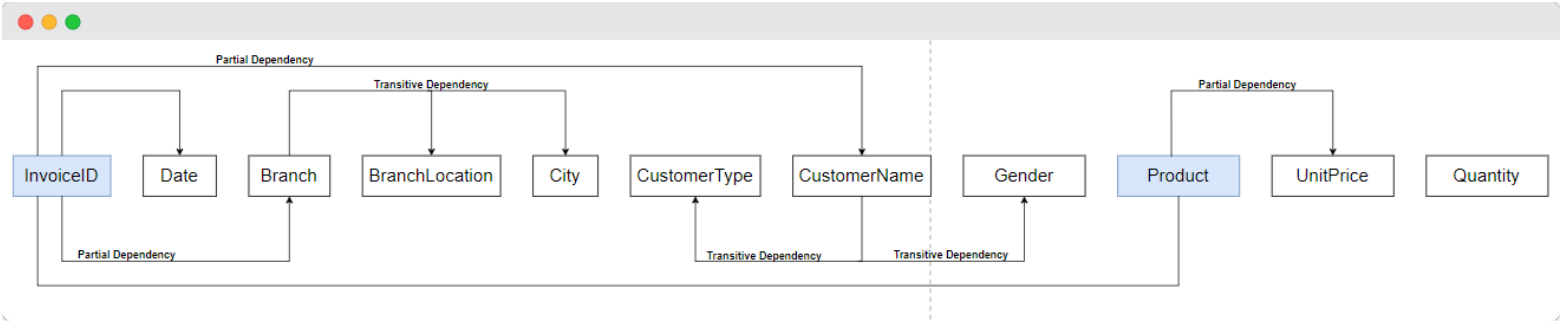


Figure 1 : Dependency Diagram after 1NF

2: Normalize the INVOICE table to 2NF

Explanation:-

- For a table to be in 2NF, there is not a partial dependency present in a table.
- Here we see in *Figure 1*, that Branch which is a non-prime attribute is dependent on part of the primary key which is InvoiceID, in a similar manner CustomerName is dependent on InvoiceID, and UnitPrice is dependent on Product.
- So, removing the partial dependency on the invoice table results in 4 tables.

After 2nd Normal Form:-

Table 3 : Invoice Table In 2NF

InvoiceID	Date	Product	Quantity
1101	2/5/2024	Phone	1
1101	2/5/2024	Battery	2
1101	2/5/2024	Cover	5
2104	3/3/2024	Code	20
2104	3/3/2024	PowerBank	45

Table 4 : Branch Table

InvoiceID	Branch	BranchLocation	City
1101	1	Quinpool	Halifax
1101	004	Windsor	Halifax
1101	008	Dartmouth	Halifax
2104	3	Dowrey Street	Toronto
2104	007	Abbey Lane	Toronto
2104	008	GlenView	Toronto

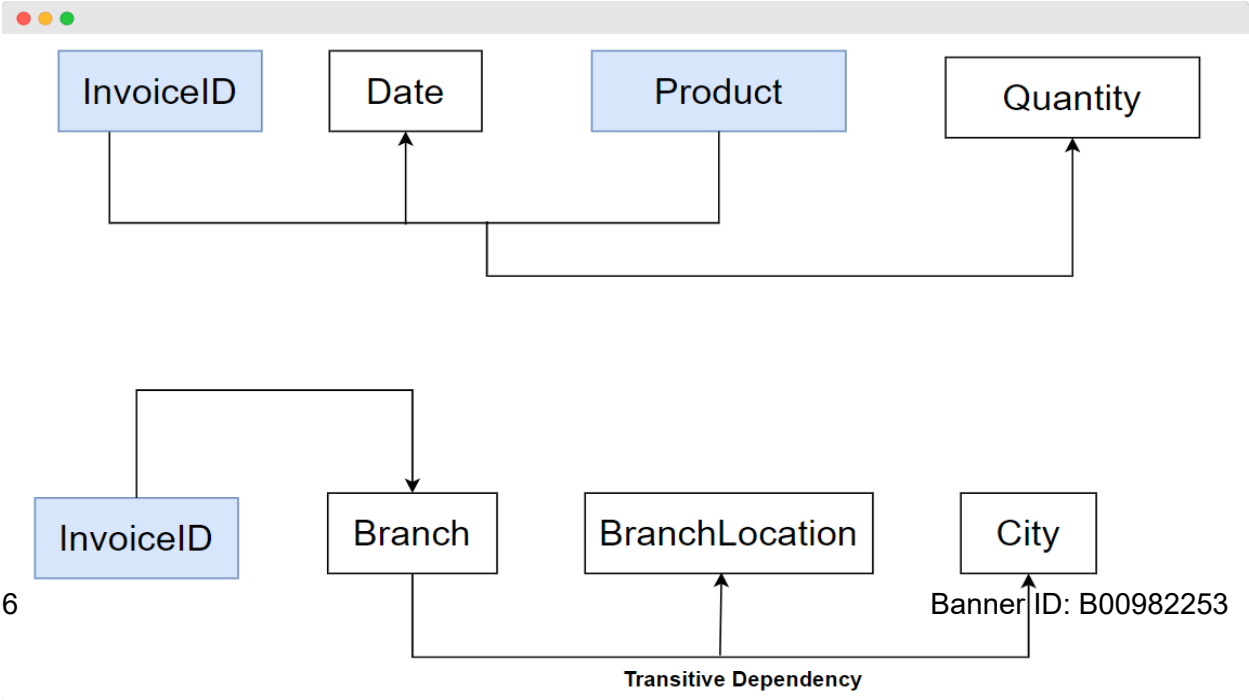
Table 5 : Customer Table

InvoiceID	CustomerName	CustomerType	Gender
1101	Alex	Member	Female
1101	Rohan	Member	Male
1101	Mark	Normal	Female
2104	Suzan	Member	Female
2104	Carla	Member	Female
2104	Andreq	Normal	Male

Table 6 : Product Table

Product	Unit Price
Phone	200
Battery	50
Cover	25
Code	10
Power Bank	35

Dependency Diagram:-



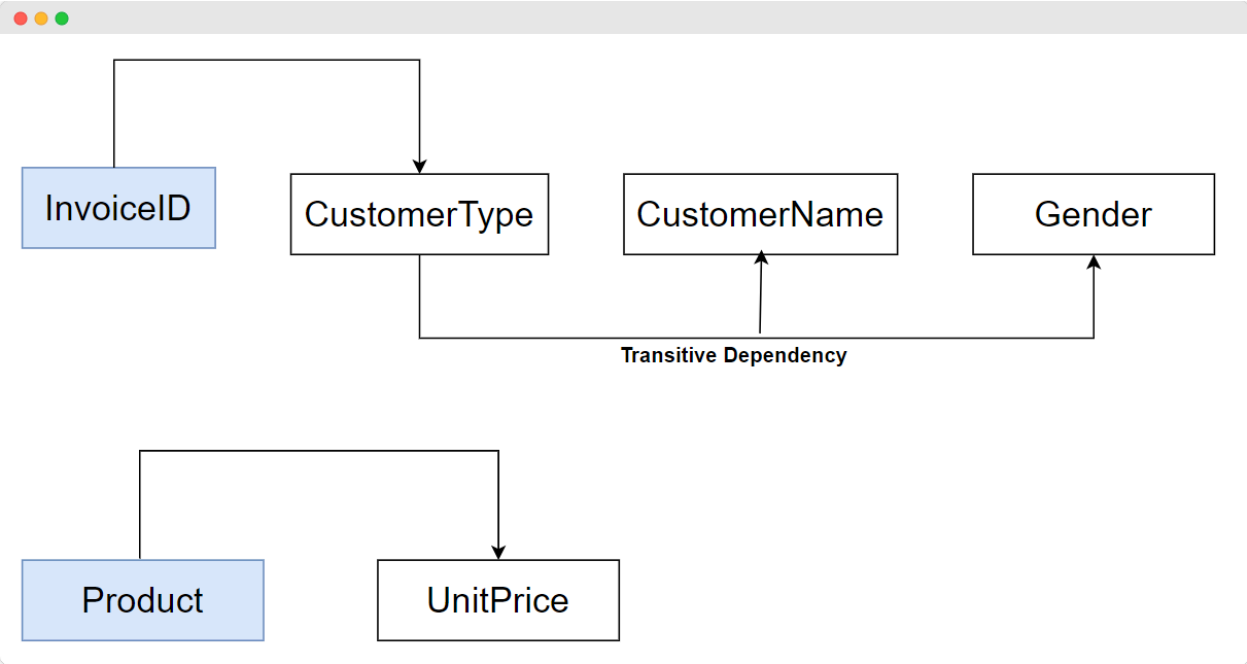


Figure 2 : Dependency Diagram after 2NF

3: Normalize the INVOICE table to 3NF

Explanation:-

- For a table to be in 3NF, there must not be a transitive dependency.
- Here we see in *Figure 2*, BranchLoaction is dependent on a combination of Branch and City, in a similar way CustomerType and Gender are dependent on CustomerName.
- We resolve this dependency by creating a separate table called Invoice_Branch, and Invoice_Customer.

After 3rd Normal Form:-

Table 7 : Invoice Table In 3NF

InvoiceID	Date	Product	Quantity
1101	2/5/2024	Phone	1
1101	2/5/2024	Battery	2
1101	2/5/2024	Cover	5
2104	3/3/2024	Code	20
2104	3/3/2024	PowerBank	45

Table 8 : Branch Table

InvoiceID	Branch	City
1101	1	Halifax
1101	004	Halifax
1101	008	Halifax
2104	3	Toronto
2104	007	Toronto
2104	008	Toronto

Table 9 : Invoice_Branch Table

Branch	BranchLocation	City
1	Quinpool	Halifax
004	Windsor	Halifax
008	Dartmouth	Halifax
3	Dowrey Street	Toronto
007	Abbey Lane	Toronto
008	GlenView	Toronto

Table 10 : Customer Table

InvoiceID	CustomerName
1101	Alex
1101	Rohan
1101	Mark
2104	Suzan
2104	Carla
2104	Andreq

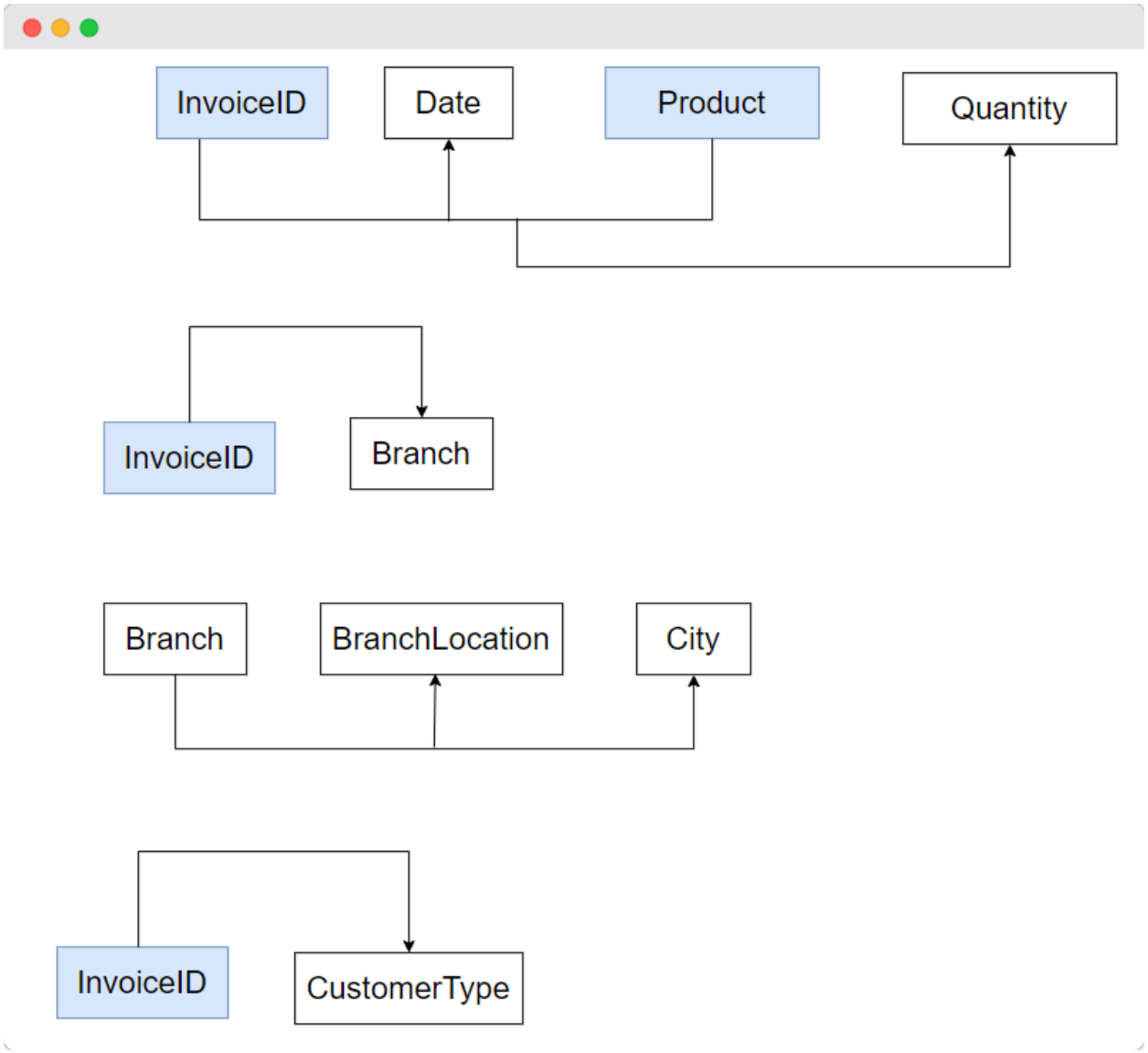
Table 11 : Invoice_Customer Table

1101	Alex	Member	Female
1101	Rohan	Member	Male
1101	Mark	Normal	Female
2104	Suzan	Member	Female
2104	Carla	Member	Female
2104	Andreq	Normal	Male

Table 12 : Product Table

Product	Unit Price
Phone	200
Battery	50
Cover	25
Code	10
Power Bank	35

Dependency Diagram:-



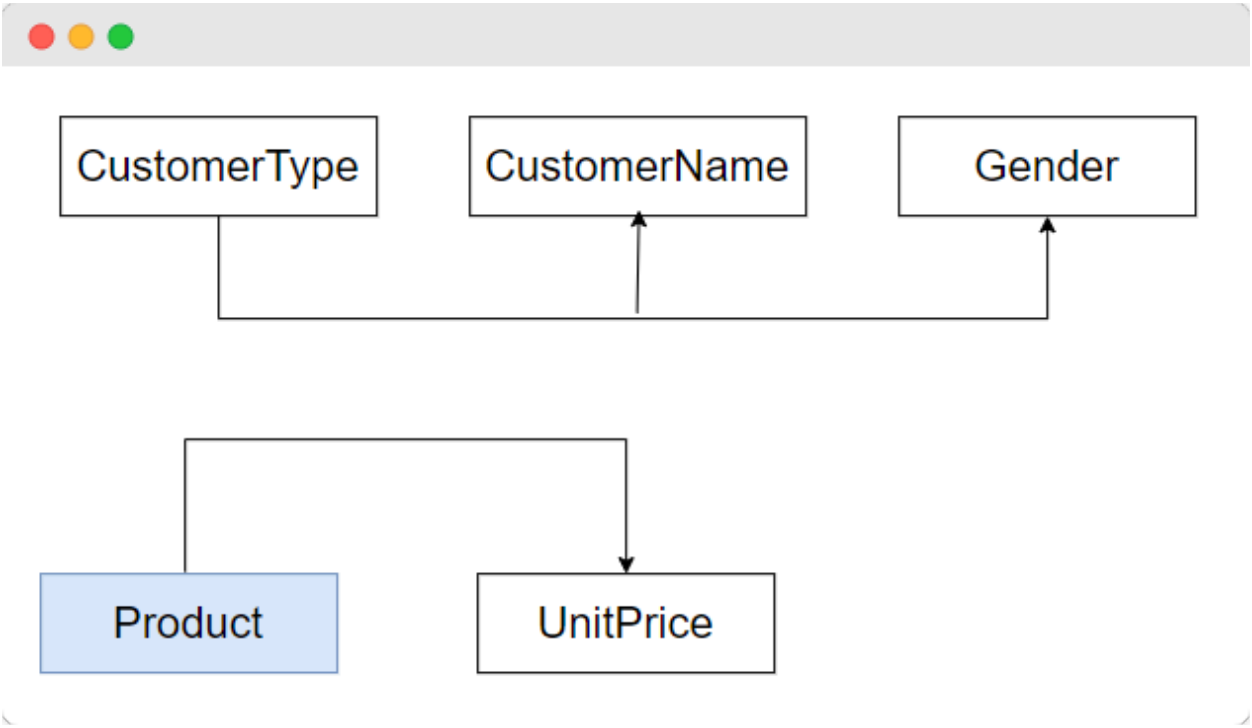


Figure 3 : Dependency Diagram after 3NF

4: De-Normalization

Question:-

Do you think it will be a good design decision to de-normalize the normal forms generated for the Product table in the above steps? Why/Why not? Provide a brief explanation.

Answer:-

- According to me, it is not feasible to store all the data in one table, because the main reason is that the usage of this database is more for write operations as compared to read operations.
- This database has information on the customer and their purchase order, so whenever any customer buys something there are multiple write operations performed on this database, while to get a daily report they have to iterate just once a day or for weekly or monthly reports almost 30-50 read operations in the entire month.
- In conclusion, in my opinion, the de-normalization is not a good idea for this kind of database.