

# CSCI-5408

## DATA MANAGEMENT, WAREHOUSING, & ANALYTICS

### LAB ASSIGNMENT - 2

Banner ID: B00977669

GitLab Assignment Link:

[https://git.cs.dal.ca/saji/csci5408\\_w24\\_b00977669\\_christin\\_saji](https://git.cs.dal.ca/saji/csci5408_w24_b00977669_christin_saji)

## Table of Contents

Problem Statement 1 .....	3
Problem Statement 2 .....	4
Problem Statement 3 .....	5
Problem Statement 4 .....	6

## Question: Normalize INVOICE Table

InvoiceID	SaleDate	ProductDescription	SupplierCode	SupplierName	QuantitySold	TotalBillAmount
2211102	1/22/2024	Samsung Galaxy S23, Apple Iphone 14, Apple Air Pods Pro	339	Technology Ent.	3	1500

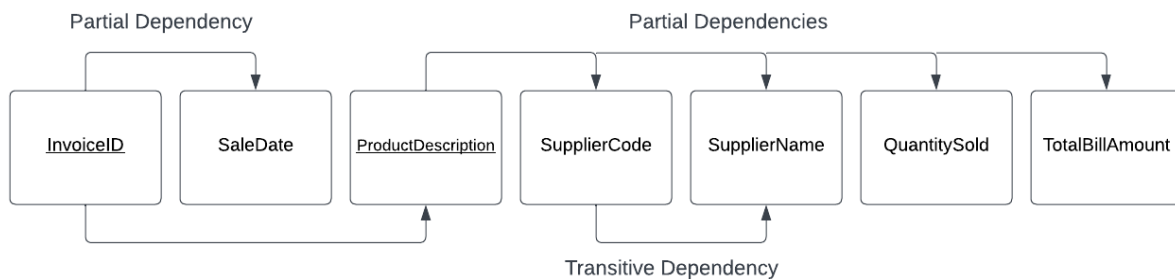


Figure 1 Invoice Table: Partial and Transitive Dependencies

### Problem Statement 1:

Normalize INVOICE table to 1NF.

InvoiceID	SaleDate	ProductDescription	SupplierCode	SupplierName	QuantitySold	TotalBillAmount
2211102	1/22/2024	Samsung Galaxy S23	339	Technology Ent.	3	500
2211102	1/22/2024	Apple Iphone 14	339	Technology Ent.	3	500
2211102	1/22/2024	Apple Air Pods Pro	339	Technology Ent.	3	500

Figure 2 Table Name: Invoice

To ensure that a table is in the first normal form (1NF), it should not have any multi-valued attributes, and each field should contain only atomic values. Therefore, I have separated all the multi-valued attributes from the 'ProductDescription' and created different rows, each containing only one 'ProductDescription' in a field.

## Problem Statement 2:

Normalize the 1NF tables to 2NF.

InvoiceID	ProductDescription
2211102	Samsung Galaxy S23
2211102	Apple Iphone 14
2211102	Apple Air Pods Pro

Figure 3 Table Name: Product Invoice

InvoiceID	SaleDate
2211102	1/22/2024

Figure 4 Table Name: Sale Date

ProductDescription	SupplierCode	SupplierName	QuantitySold	TotalBillAmount
Samsung Galaxy S23	339	Technology Ent.	3	500
Apple Iphone 14	339	Technology Ent.	3	500
Apple Air Pods Pro	339	Technology Ent.	3	500

Figure 5 Table Name: Product Details

To be in the second normal form (2NF), a table must already be in the first normal form (1NF), and is should not include any partial dependencies. In this context, SaleDate depends on InvoiceID, which is part of the primary key. Similarly, SupplierCode, SupplierName, QuantitySold, and TotalBillAmount depend only in ProductDescription, which is part of the composite primary key where the keys are InvoiceID and ProductDescription.

### Problem Statement 3:

Normalize 2NF tables to 3NF.

InvoiceID	ProductDescription
2211102	Samsung Galaxy S23
2211102	Apple Iphone 14
2211102	Apple Air Pods Pro

Figure 6 Table Name: Product Invoice

InvoiceID	SaleDate
2211102	1/22/2024

Figure 7 Table Name: Sale Date

ProductDescription	QuantitySold	TotalBillAmount	SupplierCode
Samsung Galaxy S23	3	500	339
Apple Iphone 14	3	500	339
Apple Air Pods Pro	3	500	339

Figure 8 Table Name: Product Sold

SupplierCode	SupplierName
339	Technology Ent.

Figure 9 Table Name: Supplier

For a table to be in third normal form (3NF), it must already be in second normal form (2NF), and it should not have any transitive dependencies. In this table, only SupplierName was dependent on the non-primary attribute SupplierCode, so it needs to have another table.

### Problem Statement 4:

In the database design process for a large retail organization like Walmart, invoices play a crucial role in recording items brought by customers. Do you think it will be a good design decision to de-normalize the normal forms generated for the INVOICE table in the above steps? Why/Why not? Provide a brief explanation.

Answer: For a large retail organization like Walmart, it would be a good design decision to denormalize the normal forms generated for the INVOICE table. This can improve the query performance by reducing the number of joins needed, especially for frequently read operations like invoices.