



Team Member's Name: Mir Ebtaysaam Nasir

Roll no: BSE-F23-M13

Dep: Software Engineering

1. Original Invoice (Un-normalized Data)

This is a sample invoice from a store.
All information is written together in one place.

Invoice #: INV-1001

Invoice Date: 2025-11-20

Customer ID: C101

Customer Name: Alice Johnson

Customer Phone: (555) 123-4567

Salesperson: John Doe

Line	Product ID	Product Name	Category	Quantity	Unit Price	Line Total
1	P50	Wireless Mouse	Accessory	2	25.00	50.00
2	P25	Mechanical Keyboard	Accessory	1	80.00	80.00
3	P10	Laptop Sleeve	Accessory	1	20.00	20.00

Total: \$150.00

Problems (Anomalies)

1. Repeated Data (Redundancy)

- Customer name and phone are repeated in every line of the invoice.
- Product name and category are also repeated for every product line.

2. Update Problem

If a product's category changes, you must update it in every invoice where it appears.
This can cause mistakes.

3. Insert Problem

You cannot add a new product into the database unless it appears in an invoice.

4. Delete Problem

If you delete the only invoice that has a product, you also lose all data about that product.

2. Normalization Steps

Step 1: First Normal Form (1NF)

Rule:

- No repeating groups
- Each cell should have only one value
- Each record should be simple and flat

We separate each line item into its own row.

1NF Table (All data in one big table)

Primary Key: (**InvoiceID, ProductID**)

Invoi ceID	Invoic eDate	Custo merID	Custome rName	Custome rPhone	Salesp erson	Produ ctID	Product Name	Cate gory	Qua ntity	Unit Price	Line Total
INV- 1001	2025- 11-20	C101	Alice Johnson	(555)123- 4567	John Doe	P50	Wireless Mouse	Acces sory	2	25.00	50.00
INV- 1001	2025- 11-20	C101	Alice Johnson	(555)123- 4567	John Doe	P25	Mechani cal Keyboar d	Acces sory	1	80.00	80.00
INV- 1001	2025- 11-20	C101	Alice Johnson	(555)123- 4567	John Doe	P10	Laptop Sleeve	Acces sory	1	20.00	20.00

Step 2: Second Normal Form (2NF)

Rule:

- Must be in 1NF
- No partial dependency
(A column should not depend on only part of the key)

Key = (InvoiceID, ProductID)

What depends on InvoiceID only?

- InvoiceDate
- CustomerID
- CustomerName
- CustomerPhone
- Salesperson

What depends on ProductID only?

- ProductName
- Category
- UnitPrice

So we split into 3 tables:

Invoice Header (depends only on InvoiceID)

Invoice (InvoiceID, InvoiceDate, CustomerID, CustomerName, CustomerPhone, Salesperson)

Product Table (depends only on ProductID)

Product (ProductID, ProductName, Category, UnitPrice)

Invoice Line Table (depends on InvoiceID + ProductID)

InvoiceLine (InvoiceID, ProductID, Quantity, LineTotal)

Step 3: Third Normal Form (3NF)

Rule:

- Must be in 2NF
- No transitive dependency
(No column should depend on another non-key column)

In Invoice Table:

$\text{InvoiceID} \rightarrow \text{CustomerID} \rightarrow \text{CustomerName, CustomerPhone}$

This means CustomerName and CustomerPhone should be in a separate table.

In Product Table:

`ProductID → ProductName → Category`
Category should also be in its own table.

So we create new tables to fix this.

Final Tables in 3NF (Clean and Normalized)

1. Customer

- `CustomerID (PK)`
- `CustomerName`
- `CustomerPhone`

2. Invoice

- `InvoiceID (PK)`
- `InvoiceDate`
- `Salesperson`
- `CustomerID (FK → Customer)`

3. Product

- `ProductID (PK)`
- `ProductName`
- `UnitPrice`

4. Category

- `CategoryName (PK)`

5. ProductCategory (Many-to-Many link, safer design)

- `ProductID (FK → Product)`
- `CategoryName (FK → Category)`
- Primary Key: (`ProductID, CategoryName`)

6. InvoiceLine

- `InvoiceID (PK + FK → Invoice)`
- `ProductID (PK + FK → Product)`
- `Quantity`
- `LineTotal`

