

Final Project

- Normalization report -

PROG2111

Name: Zumhliansang Lung Ler | Sungmin Leem | Nick Turco)

Course: Software Engineering Technology

Date of Submission: 30 November 2025

TABLE OF CONTENTS

Phase 3: Normalization.....	3
Original Data (Unnormalized Form)	3
First Normal Form(1NF).....	4
Second Normal Form(2NF)	5
Third Normal Form(3NF).....	6
Phase 4: DDL statement.....	9

Phase 3: Normalization

Original Data (Unnormalized Form)

OrderID	CustomerInfo	BookInfo	PublishInfo	OrderDate	TotalPrice
1001	D Beckham, C001, david@gmail.com, 7 ManU st, Madrid, ON, M7U R7D, 777-777-2323	FreeKickMaster: B001: 2 : ISBN 978-1-4028-9462-6 : Sports Last Dance: B004: 1 : ISBN 978-0-545-01022-1 : Autobiography	FIFABook: P001, MyLifeBook: P002	11/30/2025	54.97
1002	C Ronaldo, C002, CR7@gmail.com, 7 Real st, Porto, ON, R7D P7G, 777-777-7777	How To Be Winner: B002: 1 : ISBN 978-1-59327-584-6 : Autobiography	MyLifeBook: P002	11/29/2025	29.99
1003	D Beckham, C001, david@gmail.com, 7 ManU st, Madrid, ON, M7U R7D, 777-777-2323	Become a Rich: B005: 1 : ISBN 978-0-306-40615-7 : Investment	FIFABook: P001	11/28/2025	34.99
1004	L Messi, C003, Goat@gmail.com, 9 Nukamp, Barcia, ON, B9N A9H, 999-999-9999	G.O.A.T!: B003: 1 : ISBN 978-1-250-03095-7 : Autobiography Last Dance: B004: 2 : ISBN 978-1-118-02212-8 : Autobiography	AnimalBook: P003 MyLifeBook: P002	11/27/2025	54.48

Issues and Redundancy

- Multiple Values in One Cell:** The CustomerInfo column has name, ID, email, and address all together. We cannot search for just a name or email.
- Repeating Groups:** The BookInfo column has multiple books separated by commas. It is hard to know how many books were sold.
- Redundancy:** The same customer information (D Beckham) is repeated in rows 1001 and 1003.

First Normal Form(1NF)

Do:

- **Split Information:** We separated CustomerInfo into CustomerID, CustomerName, Email, Address, and Phone. We did the same for BookInfo
- **Removed Repeating Groups:** For Order 1001 (which has 2 books), we created two separate rows. Now, each row has only one book.

-Order

Order ID	Customer ID	Customer Name	Customer Email	CustomerAddress	Customer Phone	OrderDate	Book ID	BookTitle	BookISBN	Category	Qty	UnitPrice	Publisher ID	Publisher Name	Total Price
1001	C001	D Beckham	david@gmai.com	7 ManU st, Madrid, ON, M7U R7D	777-777-2323	11/30/2025	B001	FreeKickMaster	978-1-4028-9462-6	Sports	2	20	P001	FIFABook	54.97
1001	C001	D Beckham	david@gmai.com	7 ManU st, Madrid, ON, M7U R7D	777-777-2323	11/30/2025	B004	Last Dance	978-0-545-01022-1	Autobiography	1	14.97	P002	MyLifeBook	54.97
1002	C002	C Ronaldo	CR7@gmai.com	7 Real st, Porto, ON, R7D P7G	777-777-777	11/29/2025	B002	How To Be Winner	978-1-59327-584-6	Autobiography	1	29.99	P002	MyLifeBook	29.99
1003	C001	D Beckham	david@gmai.com	7 ManU st, Madrid, ON, M7U R7D	777-777-2323	11/28/2025	B005	Become a Rich	978-0-306-40615-7	Investment	1	34.99	P001	FIFABook	34.99
1004	C003	L Messi	Goat@gmai.com	9 Nukamp, Barcia, ON, B9N A9H	999-999-9999	11/27/2025	B003	G.O.A.T!	978-1-250-03095-7	Autobiography	1	24.5	P003	AnimalBook	54.48
1004	C003	L Messi	Goat@gmai.com	9 Nukamp, Barcia, ON, B9N A9H	999-999-9999	11/27/2025	B004	Last Dance	978-1-118-02212-8	Autobiography	2	14.99	P002	MyLifeBook	54.48

Second Normal Form(2NF)

Problem: The Primary Key is (OrderID, BookID).

- CustomerName depends only on OrderID. It has nothing to do with BookID. This is a partial dependency.
- BookTitle depends only on BookID. It has nothing to do with OrderID.

Solution: We split the table into three parts:

- Orders Table: Keeps information about the order (Customer, Date) → depends on OrderID.
- Books Table: Keeps information about the book (Title, Category, Publisher) → depends on BookID.
- OrderDetails Table: Keeps Qty because quantity depends on both Order and Book.

- Orders

OrderID	Customer ID	Customer Name	Customer Email	CustomerAddress	Customer Phone	OrderDate	Total Price
1001	C001	D Beckham	david@gmaiil.com	7 ManU st, Madrid, ON, M7U R7D	777-777-2323	11/30/2025	54.97
1002	C002	C Ronaldo	CR7@gmaiLcom	7 Real st, Porto, ON, R7D P7G	777-777-7777	11/29/2025	29.99
1003	C001	D Beckham	david@gmaiil.com	7 ManU st, Madrid, ON, M7U R7D	777-777-2323	11/28/2025	34.99
1004	C003	L Messi	Goat@gmaiil.com	9 Nukamp, Barcia, ON, B9N A9H	999-999-9999	11/27/2025	54.48

- Books

BookID	BookTitle	BookISBN	UnitPrice	Category	PublisherID	PublisherName
B001	FreeKickMaster	978-1-4028-9462-6	20	Sports	P001	FIFABook
B002	How To Be Winner	978-1-59327-584-6	29.99	Autobiography	P002	MyLifeBook
B003	G.O.A.T!	978-1-250-03095-7	24.5	Autobiography	P003	AnimalBook
B004	Last Dance	978-0-545-01022-1	14.97	Autobiography	P002	MyLifeBook
B005	Become a Rich	978-0-306-40615-7	34.99	Investment	P001	FIFABook

- OrderDetails

OrderID	BookID	Qty
1001	B001	2
1001	B004	1
1002	B002	1
1003	B005	1
1004	B003	1
1004	B004	2

Third Normal Form(3NF)

Problem:

- In the Orders table, CustomerName depends on CustomerID, not OrderID. ($\text{OrderID} \rightarrow \text{CustomerID} \rightarrow \text{CustomerName}$). This is a transitive dependency.
- In the Books table, PublisherName depends on PublisherID, and Category (name) is repeated text.

Solution: We moved these dependent columns to their own tables.

- Created a Customer Table. Now Orders only keeps CustomerID.
- Created a Publisher Table. Now Books only keeps PublisherID.
- Created a Category Table. Now Books only keeps CategoryID.

- Customer

CustomerID (PK)	CustomerName	CustomerEmail	CustomerAddress	CustomerPhone
C001	D Beckham	david@gmail.com	7 ManU st, Madrid, ON, M7U R7D	777-777-2323
C002	C Ronaldo	CR7@gmail.com	7 Real st, Porto, ON, R7D P7G	777-777-7777
C003	L Messi	Goat@gmail.com	9 Nukamp, Barcia, ON, B9N A9H	999-999-9999

- Book

BookID (PK)	BookTitle	BookISBN	BookPrice	BookStock	PublisherID (FK)	CategoryID (FK)
B001	FreeKickMaster	978-1-4028-9462-6	20	15	P001	CAT001
B002	How To Be Winner	978-1-5932-7-584-6	29.99	8	P002	CAT002
B003	G.O.A.T!	978-1-250-03095-7	24.5	12	P003	CAT002
B004	Last Dance	978-0-545-01022-1	14.97	20	P002	CAT002
B005	Become a Rich	978-0-306-40615-7	34.99	10	P001	CAT003

- Order

OrderID (PK)	OrderDate	TotalAmount	CustomerID (FK)
1001	11/30/2025	54.97	C001
1002	11/29/2025	29.99	C002
1003	11/28/2025	34.99	C001
1004	11/27/2025	54.48	C003

- OrderDetail

OrderDetailID (PK)	OrderID (FK)	BookID (FK)	Qty	UnitPrice
1	1001	B001	2	20
2	1001	B004	1	14.97
3	1002	B002	1	29.99
4	1003	B005	1	34.99
5	1004	B003	1	24.5
6	1004	B004	2	14.99

- Publisher

PublisherID (PK)	PublisherName
P001	FIFABook
P002	MyLifeBook
P003	AnimalBook

- Category

CategoryID (PK)	CategoryName
CAT001	Sports
CAT002	Autobiography
CAT003	Investment

Phase 4: DDL statement

