

|  |  |  |
| --- | --- | --- |
| **Module Code** | **:** | CT042-3-1-IDB – Introduction to Database |
| **Intake Code** | **:** | APDIF2206 |
| **Lecturer Name** | **:** | Lai Chew Ping |
| **Hand in Date** | **:** | 06 December 2022 |
| **Group No.** | **:** | Group 16 |
| **Group Leader** | **:** | Ng Shao Hwa |

|  |  |
| --- | --- |
| **Student ID** | **Student Name** |
| TP065019 | Ng Shao Hwa |
| TP068977 | Soong Yau Joe |
| TP068687 | Soo Jiun Guan |
| TP065322 | Adleena Binti Ahmad Nizam |

Table of Contents

[Introduction 3](#_Toc126505098)

[Part 1 4](#_Toc126505099)

[a. Database and Database Management System 4](#_Toc126505100)

[a.1 Disadvantages of file-based system 4](#_Toc126505101)

[a.2 Advantages of Database and DBMS 5](#_Toc126505102)

[a.3 Functions of DBMS 6](#_Toc126505103)

[b. Business Rules & Normalization 8](#_Toc126505104)

[b.1 Business Rules 8](#_Toc126505105)

[b.2 Normalization 9](#_Toc126505106)

[b.2.1 UNF 9](#_Toc126505107)

[b.2.2 INF 10](#_Toc126505108)

[b.2.3 2NF 12](#_Toc126505109)

[b.2.4 3NF 14](#_Toc126505110)

[b.2.5 Tables implemented 15](#_Toc126505111)

[c. Entity Relationship Diagram 16](#_Toc126505112)

[Workload Matrix – Part 1 17](#_Toc126505113)

[References 18](#_Toc126505114)

# Introduction

The existence of books has been surrounding us for centuries until now and it is unquestionably one of the essential and relatively prevalent pathways that people used to gain information and knowledges back in the days where advance computers haven’t received that many improvements. Although there is a big leap in our technologies and lead to the invention of e-books, there are still thousands of bookstores out there supporting the needs of physical books for bibliophiles and advocation of reading physical books are still exist ubiquitously in every country. There is an apparent dilemma that bookstores struggled in the past decades which is the inefficiency in storing those data that rush into their store in a daily manner which includes the records of payment from the customer and their physical order list from the publisher for books replenishment.

Fortunately, as the development of digital era begins to thrive, bookstores are allowed to save up significant amount of spaces for other functions or even enhance their bookstore in terms of aesthetic by implementing database system into their company. Furthermore, database system successfully makes the evaluation and analyzing of the bookstore’s data become easier and efficient compared with those old-fashion ways of maintaining the data. In this case study, the team will be going to design a delicate database system for the bookstore which located in Kuala Lumpur Malaysia and looking forward for developing an online E-Bookstore Database System. The constructions of a practical database system will be determined based on the requirements of the bookstores after analyzing the scenario of the bookstore.

The understanding of creating a list of business rules, normalization of the entities, and the capability of designing an Entity Relationship Diagram (ERD) will act as an important base for the team to develop a concrete framework to carry the project until the very last part of it. SQL Data Definition Language (DDL) will be the core essential software the team use to create the tables for the DBMS as it is more user-friendly and simpler to implement. When the developing process has ended, the group is going to use Data Manipulation Language (DML) to write SQL statement and run certain amount of queries to ensure there is no flaws exist throughout all the tables that had created by the team.

# Part 1

## a. Database and Database Management System

### a.1 Disadvantages of file-based system

File-based system which also known as a file system is one of the types of software that makes organizing and accessing limited groups of data by the users possible and it was considered somehow very useful for quite a long period of time since it’s being introduced. Nonetheless, people will be surely notice those flaws within this system as it starts to embed into their daily workflows and this system has gradually become obsolete as new type of system was being introduced to replace the incompleteness of the previous system.

Speaking of disadvantages of using a file-based system, the most concern issues will probably be the difficulties in administrating the system (Watt, 2014) and data as the number of files starts to grow rapidly. If the bookstore is trying to implement this system to manage their data and files, it is no surprise an imaginable calamity will be occur in the future. For instance, the fact that the bookstore itself who has becomes a place to contains thousands of titles of books already indicate the plethora of data they need to cope with in the management level. Additionally, the restless amount of data that keep on rushing into the file-based system due to the roll out of new titles in every month will eventually causing the process of retrieving certain data based on given requirements become more suffering and time-consuming. (Roomi, 2021)

The second drawback of the file-based system will be the occurrence of data redundancy and data inconsistency as number of files starts to rise. (Singh, 2022) There will be a comparably high chances for the system to store the same but unmatching data in different places. The severity of this problem will not seem to be serious until it comes to the part where data relates to personal information of the customers or members of the bookstores. Delivery, as being an essential part of an online purchasing service is also a part of the process that is driven by a ton of data in behind it and one of the categories of data will be the address of every customer. Try to imagine what if there is a part being forgotten on updating customers’ addresses in the system. This will be definitely brings up a lot of problems which it will be including the delivery of books to an incorrect address.

### a.2 Advantages of Database and DBMS

Despite there are several detrimental flaws available in the file-based system, human will always search for solutions to resolve those dilemmas they are facing. The existence of the file-based system has successfully become a steppingstone and foundation to foster the development of the Database and Database Management System which is also known as DBMS. Nowadays, most of the companies and industries are highly rely on the DBMS to help them store and manage those data they had collected in the past few years or decades due to its nearly impeccable performance and great recognition among users across the industries.

One of the primary advantages that DBMS offer will be its ability to relate and store data in a single data repository (TutorialsPoint, 2020) and this enable the bookstore to generate higher quality information to make sure they have a clearer picture of the inclination of their company and plan out a constructive tactic for them to run their business more smoothly and help them to understand the demands of their customers at the same time. (Thiru, n.d.) For example, the bookstore analyzes the data they collected from the customers wisely to find out which books has an outstanding sale throughout a specific period of time to give better recommendation to their buyers and keep them on track with the latest trends in the world of bibliography. Additionally, popularity of books enables the bookstore manager to decide the quantity of each title they have to order to prevent surplus situation.

Furthermore, one of the benefits which is increment of the productivity of the users (*Benefits of Database Management Systems (DBMS) | ZoomInfo*, n.d.) promoted by the DBMS has become one of the quintessential and yet important features that most of the users required. By implementing DBMS into the bookstore company, it will save up a lot of time for the manager to focus his or her effort on the high-values activities and lower the time to spend on cleaning data or debugging the program. Moreover, DBMS allows bookstore manager to back up and store their data in the backup server to ensure duplicate data is always available for them when data recovery procedures are needed. (TutorialsPoint, 2021)

In terms of privacy and security of using a DBMS, the system can be said to be much safer and secure compared to file-based system as the security framework of the DBMS is more accessible and more user-friendly. (Singhal, 2022) As a result of the impressive security features of the DBMS, the bookstore manager will be able to ensure all the data are being encrypted safely in their server and greatly reduce the risk of invasion of their data from malignant hackers by managing the accessibility of the workers based on their obligation.

### a.3 Functions of DBMS

The main function of the database management system is to arrange the given data in the table. Moreover, it also manages and controls the database in a unified way to ensure the security and integrity of the database. There are some extra functions of DBMS have been listed below:

* **Multiuser Access Control**

DBMS uses sophisticated algorithms to ensure multiple users access the data simultaneously without destroying file integrity and causing inconsistencies. (Thiru, n.d.) In contrast, spreadsheets do not provide different users with the ability to view and edit the different data in the same file. In simple terms, once the first user opens the file, another user cannot access the file. Alternatively, other users can read the file, but cannot edit the data at the same time. This will help to provide efficient access control for the role separation of the E-Bookstore Database System.

* **Backup and Recovery Management**

DBMS allows administrators to execute backup and recovery procedures in order to ensure data integrity and safety. (“Study Notes for DB Design and Management Exam 1 (Chapters 1-2-3)”, n.d.) Database backup means database will create a duplication of the data regularly and then store it on a safe location while administrator can use recovery to retrieve the lost data of the database after a failure. With this, the E-Bookstore could maintain its business continuity from any disruptive incident.

* **Security Management**

It ensures data is not accessed by unauthorized users that utilize it for criminal activities, as well as avoids user privacy and data security from malicious threats. Many security rules have been used in database to establish a reliable and safe environment for users. In recent years, there are many companies that neglected to manage the security issues of the data have caused various problems to occur. Therefore, data security issues cannot be ignored.

* **Data Dictionary Management**

Data dictionary refers to define and describe details and relationship about the data elements that are used in a database project. Whenever there is any change happens in the data dictionary, the structure is automatically modified, so the users will access the changed structure of the database. A data dictionary can make data easier to analyze due to its useful information on relationship between different databases tables.

## b. Business Rules & Normalization

### b.1 Business Rules

1. A Publisher\_ID must associate with one or many Book\_ID
2. A Book\_ID must be associate with only one Publisher\_ID
3. A Bookstore\_Order must be associated at least one Publisher
4. A Member\_ID must be associated with one and only one Cart\_ID
5. A Rating must be associated with one and only one Member\_ID
6. A Payment\_ID must be associated with one and only one Member\_ID
7. A Member\_ID must associate with at least one Payment\_ID
8. A Member\_Order can associate with zero or one Rating
9. A Rating must associate with one and only one Member\_Order
10. Title can be associated with one or more Publisher\_ID
11. Book\_ID must be associated with at least one Genre
12. Author must be associated with at least one Book\_ID
13. A Book\_ID must be associated with at least one Author
14. A Book\_ID must only be associated with only one Unit\_Price\_(RM)
15. A Member\_ID must associate with only one Delivery\_Add
16. A Member\_ID can have zero or many MBOrder\_Num
17. A Shopping\_Cart must only be associated with only one Member\_ID

### b.2 Normalization

Normalization is a database design technique for organizing data in an orderly and consistent manner. A database system without normalization may slow down work progress and provide inaccurate information for the users. Therefore, it is necessary for a database management system because it can reduce data redundancy, remove undesired characteristics, as well as ensure only relevant data is stored in each table. In short, the application of normalization can simplify the design of the database and achieve optimal structure in order to allow users easier to retrieve and find data in the database. There are three different normal forms are required for normalization, which are 1NF, 2NF and 3NF.

### b.2.1 UNF

Table

Description automatically generatedUnnormalized form, or UNF, is a table that contains a jumble of data. The data in UNF are redundant and unstructured, therefore this type of database model is thought to be the simplest. Despite it being disorganised, UNF is still an important step in normalisation process since it allows the reconstruction of data to create 1NF table. Table 1 shows an example of an UNF data table of the E-bookstore sales order.

*Table 1*

### b.2.2 INF

The rules applied in first normalized form states that:

1. Present data in a tabular format.
2. Each row and column intersection should only contain one value or a null.
3. Each attribute should contain an appropriate data value.
4. Each attribute should have a unique key called Primary Key.
5. Columns and rows should not be duplicated.

To create the first normalization form, all the data from table 1 must be presented in a proper table. Next, separate multiple or duplicated data to separate rows and columns in order to make each attribute in each cell has one relationship with the other attributes. Then, identify the primary key to make sure the dataset is unique. In the ERD, the primary key is identified with an underline. The table in 1NF is not separated yet, all data is compiled in one table, therefore data redundancy is still there. Table 2 shows the example of 1NF that is created from the attributes of the UNF table in table 1. Moreover, there are partial dependencies that can be identified from the table that must be solved in the following NF.

First normal form

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Member\_ID** | **Name** | **Delivery\_Add** | **MBOrder\_Num** | **Date** | **Status** | **Book\_ID** | **Title** | **Quantity** | **Unit\_Price\_(RM)** | **Line\_Total** | **Order\_Total** |
| MEM0000152 | Jimmy Burgess | No. 42C, Batu 10 1/2, Jalan Cheras Perdana | C003871 | 29 January 2023 | Processing | B0008 | Heart Bones | 2 | 50.80 | 101.60 | 362.70 |
| MEM0000152 | Jimmy Burgess | No. 42C, Batu 10 1/2, Jalan Cheras Perdana | C003871 | 29 January 2023 | Processing | B0879 | The Hating Game | 3 | 46.90 | 140.70 | 362.70 |
| MEM0000152 | Jimmy Burgess | No. 42C, Batu 10 1/2, Jalan Cheras Perdana | C003871 | 29 January 2023 | Processing | B0945 | The Fine Print | 2 | 60.20 | 120.40 | 362.70 |

Identify PK: Member\_ID, MBOrder\_Num, Book\_ID

*Table 2*

### b.2.3 2NF

The second step of the normalization process is the second normalized form (2NF). In this step, 2NF must identify all dependencies between attributes and eliminate any partial dependency. When an attribute depends only on **SOME** of the PK and not **ALL** of the PK, it is known as partial dependency. If the attribute is dependent on **ALL** the PK, it is called as full dependency. 2NF must make sure that each non-key attributes, or non-determinant, or non-PK is fully dependent on its determinant (PK).

Reconstruction of the data to 2NF contributes many different advantages such as:

1. Data redundancy is reduced greatly.
2. Data in the database is more consistent.
3. The overall structure of the data is improved.
4. Security of the database is improved.

Table 3 shows an example of 2NF. ‘Member\_ID’, ‘MBOrder\_Num’, and ‘Book\_ID’ are utilized as the primary key. ‘Name’, ‘Delivery\_Add’, ‘Date’, ‘Status’, ‘Title’, ‘Quantity’, ‘Unit\_Price\_(RM)’, ‘Line\_Total’, and ‘Order\_Total’ are all identified as partial dependency.

Second normal form

|  |  |  |  |
| --- | --- | --- | --- |
| **Composite PK** | **Non-key attributes** | **Check List** | **Action** |
| Member\_ID, MBOrder\_Num, Book\_ID | Name | Is Name fully dependent on all PK?  Answer: No, Name depend only on Member\_ID | Partial dependency exists. Remove Name together with its determinant (Member\_ID) to form a new table. |
| Member\_ID, MBOrder\_Num, Book\_ID | Delivery\_Add | Is Delivery\_Add fully dependent on all PK?  Answer: No, Delivery\_Add depend only on Member\_ID | Partial dependency exists. Remove Delivery\_Add together with its determinant (Member\_ID) to form a new table. |
| Member\_ID, MBOrder\_Num, Book\_ID | Date | Is Date fully dependent on all PK?  Answer: No, Date depends only on MBOrder\_Num | Partial dependency exists. Remove Date together with its determinant (MBOrder\_Num) to form a new table. |
| Member\_ID, MBOrder\_Num, Book\_ID | Status | Is Status fully dependent on all PK?  Answer: No, Status depends only on MBOrder\_Num | Partial dependency exists. Remove Status together with its determinant (MBOrder\_Num) to form a new table. |
| Member\_ID, MBOrder\_Num, Book\_ID | Title | Is Title fully dependent on all PK?  Answer: No, Title depends only on BookID | Partial dependency exists. Remove Title together with its determinant (Book\_ID) to form a new table. |
| Member\_ID, MBOrder\_Num, Book\_ID | Quantity | Is Quantity fully dependent on all PK?  Answer: No, Quantity depends only on both Book\_ID & MBOrder\_Num | Partial dependency exists. Remove Quantity together with its determinant (Book\_ID & MBOrder\_Num) to form a new table. |
| Member\_ID, MBOrder\_Num, Book\_ID | Unit\_Price\_(RM) | Is Unit\_Price\_(RM) fully dependent on all PK?  Answer: No, Unit\_Price\_(RM) depend only on Book\_ID | Partial dependency exists, Remove Unit\_Price\_(RM) together with its determinant (Book\_ID) to form a new table. |
| Member\_ID, MBOrder\_Num, Book\_ID | LineTotal | Is LineTotal fully dependent on all PK?  Answer: No, Line\_Total depend only on both Book\_ID & MBOrder\_Num | Partial dependency exists, Remove Line\_Total together with its determinant (Book\_ID & MBOrder\_Num) to form a new table. |
| Member\_ID, MBOrder\_Num, Book\_ID | OrderTotal | Is OrderTotal fully dependent on all PK?  Answer: No, Order\_Total depend only on MBOrder\_Num | Partial dependency exists, Remove OrderTotal together with its determinant (MBOrder\_Num) to form a new table. |

*Table 3*

### b.2.4 3NF

The third normalized form (3NF) is the third step of the normalization process. In 3NF, if transitive dependency is identified, it must be eliminated. Transitive dependency happens when a non-key attribute depends on another non-key attribute. When the tables are in 3NF, the normalization process is completed.

Tables in third normal form

|  |  |  |
| --- | --- | --- |
| Member\_ID | Name | Delivery\_Add |
| MEM0000152 | Jimmy Burgess | No. 42C, Batu 10 1/2, Jalan Cheras Perdana |

|  |  |  |
| --- | --- | --- |
| Book\_ID | Title | Unit\_Price\_(RM) |
| B0008 | Heart Bones | 50.80 |
| B0879 | The Hating Game | 46.90 |
| B0945 | The Fine Print | 60.20 |

|  |  |  |  |
| --- | --- | --- | --- |
| MBOrder\_Num | Date | Status | Order\_Total |
| C003871 | 29 January 2023 | Processing | 362.70 |

|  |  |  |  |
| --- | --- | --- | --- |
| Book\_ID | MBOrder\_Num | Quantity | Line\_Total |
| B0008 | C003871 | 2 | 101.60 |
| B0879 | C003871 | 3 | 140.70 |
| B0945 | C003871 | 2 | 120.40 |

### b.2.5 Tables implemented

Member

|  |  |  |
| --- | --- | --- |
| Member\_ID | Name | Delivery\_Add |
| MEM0000152 | Jimmy Burgess | No. 42C, Batu 10 1/2, Jalan Cheras Perdana |

Book

|  |  |  |
| --- | --- | --- |
| Book\_ID | Title | Unit\_Price\_(RM) |
| B0008 | Heart Bones | 50.80 |
| B0879 | The Hating Game | 46.90 |
| B0945 | The Fine Print | 60.20 |

Member\_Order

|  |  |  |  |
| --- | --- | --- | --- |
| MBOrder\_Num | Date | Status | Order\_Total |
| C003871 | 29 January 2023 | Processing | 362.70 |

Member\_Order\_Details

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Detail\_ID | Book\_ID (FK) | MBOrder\_Num (FK) | Quantity | Line\_Total |
| D0021 | B0008 | C003871 | 2 | 101.60 |
| D0017 | B0879 | C003871 | 3 | 140.70 |
| D0092 | B0945 | C003871 | 2 | 120.40 |

## Diagram, engineering drawing Description automatically generatedc. Entity Relationship Diagram

*Figure 1*

# Workload Matrix – Part 1

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Part** | **Component** | **Student 1**  **Ng Shao Hwa (TP065019)** | **Student 2 Soong Yau Joe (TP068977)** | **Student 3**  **Soo Jiun Guan (TP068687)** | **Student 4 Adleena Binti Ahmad Nizam (TP065322)** | **Total** |
|  | a) Database and Database  Management System | **25%** | **25%** | **25%** | **25%** | **100%** |
|  | b) Business Rules & Normalization | **25%** | **25%** | **25%** | **25%** | **100%** |
|  | c) Entity Relationship Diagram | **25%** | **25%** | **25%** | **25%** | **100%** |

# References

*Benefits of Database Management Systems (DBMS) | ZoomInfo*. (n.d.). Pipeline Blog. <https://pipeline.zoominfo.com/operations/6-benefits-of-using-database-management-systems-dbms>

Roomi, M. (2021, February 18). *7 Advantages and Disadvantages of File Management System | Limitations & Benefits of File Management System*. HitechWhizz - the Ultimate Tech Experience. <https://www.hitechwhizz.com/2021/02/7-advantages-and-disadvantages-limitations-benefits-of-file-management-system.html>

Singh, C. (2022, July 24). *Advantages and Disadvantages of DBMS: DBMS vs file System.* BeginnersBook. <https://beginnersbook.com/2015/04/dbms-vs-file-system/>

Singhal, P. (2022, December 13). *Database Security in DBMS*. Scaler Topics. <https://www.scaler.com/topics/database-security-in-dbms/>

*Study Notes for DB Design and Management Exam 1 (Chapters 1-2-3).* (n.d.). <https://www.ccri.edu/faculty_staff/comp/kelly/COMP1200/Study%20Notes%20for%20DB%20Design%20and%20Management%20Exam%201.pdf>

Thiru, T. (n.d.). *Functions of Database Management System (DBMS)*. <http://www.myreadingroom.co.in/notes-and-studymaterial/65-dbms/465-functions-of-dbms.html>

Thiru, T. (n.d.). *What is DBMS? Advantages and Disadvantages of Database Management System(DBMS)*. <http://www.myreadingroom.co.in/notes-and-studymaterial/65-dbms/462-advantages-and-disadvantages-of-dbms.html>

TutorialsPoint. (2020, June 19). *Advantages of Database Management System*. <https://www.tutorialspoint.com/Advantages-of-Database-Management-System>

TutorialsPoint. (2021, July 28). *DBMS - Data Backup*. <https://www.tutorialspoint.com/dbms/dbms_data_backup.htm>

Watt, A. (2014, October 24). *Chapter 1 Before the Advent of Database Systems – Database Design – 2nd Edition*. Pressbooks. <https://opentextbc.ca/dbdesign01/chapter/chapter-1-before-the-advent-of-database-systems/>