Assesment Project: Publisher-Implementation

Module: Database Systems

24/04/2023

Rafal Krzempek

BSc in Information Technology

Student No: 20096679

# Table of Contents

[Table of Contents 1](#_Toc129896275)

[Introduction 2](#_Toc129896276)

[Business Description 3](#_Toc129896277)

[Conceptual Data Model (Enhanced ER Diagram) 4](#_Toc129896278)

[Logical Data Model 5](#_Toc129896279)

[Printing House 5](#_Toc129896280)

[Bookshop 5](#_Toc129896281)

[Salary 5](#_Toc129896282)

[Author 5](#_Toc129896283)

[Manuscript (3 options) 5](#_Toc129896284)

[Book 6](#_Toc129896285)

[Sale 6](#_Toc129896286)

[Storage Representation 7](#_Toc129896287)

[Table Design 13](#_Toc129896288)

[Printing House 13](#_Toc129896289)

[Bookshop 13](#_Toc129896290)

[Salary 14](#_Toc129896291)

[Author 14](#_Toc129896292)

[Manuscript 15](#_Toc129896293)

[Non-Fiction 15](#_Toc129896294)

[Fiction 16](#_Toc129896295)

[Book 16](#_Toc129896296)

[Sale 17](#_Toc129896297)

[Security & View 18](#_Toc129896298)

[Queries 19](#_Toc129896299)

[Conclusion 20](#_Toc129896300)

# Introduction

As part of my Database Systems module, I am required to revisit the previous semester's project and design and develop its extended database, which will be used as documentation for implementing the database into the MySQL environment. Its purpose is to show the process of improving database models and introducing new elements, from the conceptual phase, through the logical phase, to the creation of a Physical Data Model for implementation in the database. Additional utilities like storage representation and security will also be introduced. Each level of design is logically interconnected. The developed project concerns a publishing house that prints and sells books.

# Business Description

A publishing house called 'Big A' works with fifteenth authors, sixteenth printing houses and forty bookshops. Once a manuscript submitted to the publishing house by one author has been accepted, its printing is commissioned to one of the printing houses working with the publishing house. Only one printing house is commissioned to print each title. Each title has one author. For each title, the publisher pays the author in the form of a salary and may also pay in advance (dates of these transactions must be recorded). Each payment is made to one author. Once the printed titles (books) have been delivered from the printers, they will be sold to bookshops, and a single sale may involve one or more titles. Information about each transaction, such as the date of sale, the date of due payment, the number of items, the amount of the sale and the terms and conditions of the sale should be recorded.

**The following information will be collected for each author: PPSN, first name, last name, address with street, city, country, postal code, e-mail address, telephone number and whether the author has sent his/her biography to the publisher.**

**The manuscript, on the other hand, in addition to its content and date of receipt, must include information on the date of the decision to put it into print, the number of pages, whether additional attachments from the author have been provided, and may specify whether the book is classified as literary fiction or non-fiction. The genre or discipline of the manuscript should then be described.**

**For the printing house is needed data including its name, address with street, city, country, postal code, e-mail address, telephone number and name of the person who is responsible for cooperation with the publishing house.**

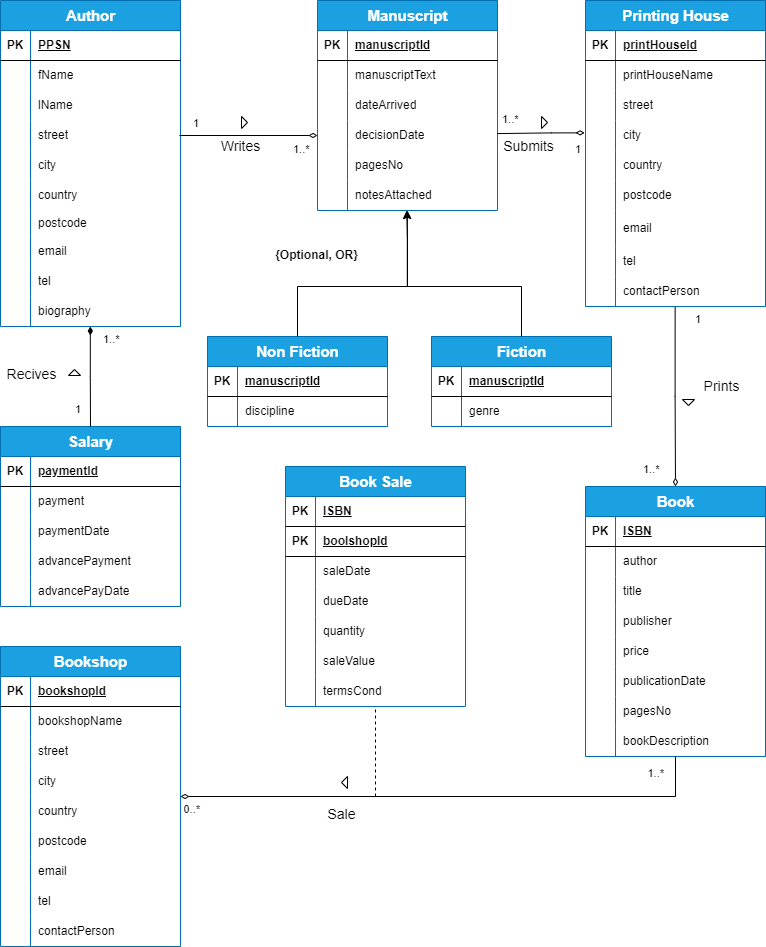
**The table named book should contain records about the book's ISBN number, title, publisher's name, price, publication date, number of pages and an indication if there is a description of the book.**

**As books are sold to bookshops, it is also necessary to collect data about the name of the bookshop, the street where it is located, city, country, postal code, e-mail address, telephone number and name of the person who is responsible for cooperation with the publishing house.**

**Annually, the publishing house receives 60 manuscripts, of which 40 may be classified as Fiction and the rest as Non-Fiction.**

**Since the publishing house's inception, 58 titles have been printed, and 58 have hit bookshops.**

# Conceptual Data Model (Enhanced ER Diagram)



The table has been renamed from 'Sale' to 'Book Sale' and two PKs ('ISBN' from Book and 'bookshopId' from Bookshop) have been added there.

# Logical Data Model

*CamelCase notation was used in the following examples to meet the naming requirements of MySql.*

## Printing House

**PrintingHouse** (printHouseId, printHouseName, street, city, country, postcode, email, tel, contactPerson)

Primary key printHouseId

## Bookshop

**Bookshop** (bookshopId, bookshopName, street, city, country, postcode, email, tel, contactPerson)

Primary key bookshopId

## Salary

**Salary** (paymentId, payment, paymentDate, advancePayment, advancePayDate)

Primary key paymentId

## Author

**Author** (PPSN, fName, lName, street, city, country, postcode, email, tel, biography)

Primary key PPSN

Foreign key PaymentId references **Salary**(paymentId)

## Manuscript (3 options)

Three Tables:

**Manuscript** (manuscriptId, manuscriptText, dateArrived, decisionDate, pagesNo, notesAttached)

Primary key manuscriptId

Foreign key PPSN references **Author** (PPSN)

Foreign key printHouseId references **PrintingHouse** (printHouseId)

**NonFiction** (manuscriptId, discipline)

Primary Key manuscriptId

**Fiction**(manuscriptId, genre)

Primary Key manuscriptId

Two Tables:

**NonFiction** (manuscriptId, manuscriptText, dateArrived, decisionDate, pagesNo, notesAttached, discipline)

Primary Key manuscriptId

Foreign key PPSN references **Author** (PPSN)

Foreign key printHouseId references **PrintingHouse** (printHouseId)

**Fiction**(manuscriptId, manuscriptText, dateArrived, decisionDate, pagesNo, notesAttached, genre)

Primary Key manuscriptId

Foreign key PPSN references **Author** (PPSN)

Foreign key printHouseId references **PrintingHouse** (printHouseId)

One Table:

**Manuscript** (manuscriptId, manuscriptText, dateArrived, decisionDate, pagesNo, notesAttached, discipline, genre)

Primary key manuscriptId

Foreign key PPSN references **Author** (PPSN)

Foreign key printHouseId references **PrintingHouse** (printHouseId)

## Book

**Book** (ISBN, author, title, publisher, price, publicationDate, pagesNo, bookDescription)

Primary key ISBN

Foreign key printHouseId references **PrintingHouse** (printHouseId)

## BookSale

**BookSale** (saleDate, dueDate, quantity, saleValue, termsCond)

Primary key (ISBN, bookchopId) composite Key

Foreign key ISBN references **Book** (ISBN)

Foreign key bookshopId references **Bookshop** (bookshopId)

*The following changes were made to the above table:*

* *renamed the table from 'Book' to 'BookSale'*
* *two PKs were added (ISBN and bookshopId)*

# Storage Representation

|  |  |
| --- | --- |
| Table: PrintingHouse | |
| Attributes | **Size - Bytes** |
| printHouseId | 2 |
| printHouseName | 60 |
| street | 70 |
| city | 20 |
| country | 50 |
| postcode | 8 |
| email | 30 |
| tel | 12 |
| contactPerson | 50 |
| Total: | **302** |

The publishing house cooperates with sixteen printing houses.

16 \* 302 = **4,832** bytes are needed for the Printing House table.

|  |  |
| --- | --- |
| Table: Bookshop | |
| Attributes | **Size - Bytes** |
| bookshopId | 3 |
| bookshopName | 60 |
| street | 70 |
| city | 20 |
| country | 50 |
| postcode | 8 |
| email | 30 |
| tel | 12 |
| contactPerson | 50 |
| Total: | **303** |

The publishing house cooperates with forty bookshops.

40 \* 303 = **12,120** bytes are needed for the Bookshops table.

|  |  |
| --- | --- |
| Table: Salary | |
| Attributes | **Size - Bytes** |
| paymentId | 4 |
| payment | 12 |
| paymentDate | 10 |
| advancePayment | 12 |
| advancePayDate | 10 |
| Total: | **48** |

The publisher has paid salaries for 60 manuscripts.

60 \* 48 = **2,880** bytes are needed for the Salary table.

|  |  |
| --- | --- |
| Table: Author | |
| Attributes | **Size - Bytes** |
| PPSN | 9 |
| fName | 30 |
| lName | 30 |
| street | 70 |
| city | 20 |
| country | 50 |
| postcode | 8 |
| email | 30 |
| tel | 12 |
| biography | 3 |
| paymentId | 4 |
| Total: | **226** |

The publishing house works with fifteen authors.

15 \* 226 = **3,390** bytes are needed for the Author table.

**Manuscript – 3 options:**

60 written manuscripts were delivered to the publisher, 20 of them are Non-Fiction and 40 are Fiction. Below are the calculations and total bytes for each of the options:

1. Three tables

|  |  |
| --- | --- |
| Table: Manuscript | |
| Attributes | **Size - Bytes** |
| manuscriptId | 5 |
| manuscriptText | 4 |
| dateArrived | 10 |
| decisionDate | 10 |
| pagesNo | 4 |
| notesAttached | 3 |
| PPSN | 9 |
| printHouseId | 2 |
| Total: | **47** |

47 \* 60 = 2820

**2,820** bytes are needed for the Manuscript table.

|  |  |
| --- | --- |
| Table: NonFiction | |
| Attributes | **Size - Bytes** |
| manuscriptId | 5 |
| discipline | 18 |
| Total: | **23** |

23 \* 20 = 460

**460** bytes are needed for the Non-Fiction table.

|  |  |
| --- | --- |
| Table: Fiction | |
| Attributes | **Size - Bytes** |
| manuscriptId | 5 |
| genre | 18 |
| Total: | **23** |

23 \* 40 = 9200

**460** bytes are needed for the Fiction table.

2820 + 460 + 460 = 4200

**4,200** bytes are needed for the three-table option.

1. Two tables

|  |  |
| --- | --- |
| Table: NonFiction | |
| Attributes | **Size - Bytes** |
| manuscriptId | 5 |
| manuscriptText | 4 |
| dateArrived | 10 |
| decisionDate | 10 |
| pagesNo | 4 |
| notesAttached | 3 |
| discipline | 18 |
| PPSN | 9 |
| printHouseId | 2 |
| Total: | **65** |

65 \* 20 = 1300

**1,300** bytes are needed for the Non-Fiction table.

|  |  |
| --- | --- |
| Table: Fiction | |
| Attributes | **Size - Bytes** |
| manuscriptId | 5 |
| manuscriptText | 4 |
| dateArrived | 10 |
| decisionDate | 10 |
| pagesNo | 4 |
| notesAttached | 3 |
| genre | 18 |
| PPSN | 9 |
| printHouseId | 2 |
| Total: | **65** |

65 \* 40 = 2600

**2,600** bytes are needed for the Fiction table.

1300 + 2600 = 3900

**3,900** bytes are needed for the two-table option.

1. One table

|  |  |
| --- | --- |
| Table: Manuscript | |
| Attributes | **Size - Bytes** |
| manuscriptId | 5 |
| manuscriptText | 4 |
| dateArrived | 10 |
| decisionDate | 10 |
| pagesNo | 4 |
| notesAttached | 3 |
| discipline | 18 |
| genre | 18 |
| PPSN | 9 |
| printHouseId | 2 |
| Total: | **83** |

83 \* 60 = 4980

**4,980** bytes are needed for the one table option.

|  |  |
| --- | --- |
| Table: Book | |
| Attributes | **Size - Bytes** |
| ISBN | 13 |
| author | 62 |
| title | 100 |
| publisher | 5 |
| price | 6 |
| publicationDate | 10 |
| pagesNo | 4 |
| bookDescription | 3 |
| printHouseId | 2 |
| Total: | **205** |

Fifty-eight manuscripts have been put into print by the publishing house and as many titles have been printed.

58 \* 205 = **11,890** bytes are needed for the Book table.

|  |  |
| --- | --- |
| Table: BookSale | |
| Attributes | **Size - Bytes** |
| saleDate | 10 |
| dueDate | 10 |
| quantity | 4 |
| saleValue | 8 |
| termsCond | 8 |
| ISBN | 13 |
| bookshopId | 3 |
| Total: | **56** |

*The following changes were made to the above table:*

* *renamed the table from 'Book' to 'BookSale'*
* *two PKs were added (ISBN and bookshopId)*

The publishing house has sold fifty-eight titles to bookshops.

58 \* 56 = **3,248** bytes are needed for the Book Sale table.

For the purposes of this project, it has been assumed that the Manuscript table will be presented as an option with two sub-tables. The storage requirements for this project are shown below:

4,832 + 12,120 + 2,880 + 3,390 + 4,200 + 11,890 + 3,248 = **42,560** bytes are needed to provide the data storage requirements for this project.

# Table Design

## PrintingHouse

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| FIELD | TYPE | SIZE | NULL/not NULL | DEFAULT | Constraints | INDEX | Description |
| printHouseId | int | 2 | not NULL |  | Unique | PK | This is a unique printing house number starting with 10. |
| printHouseName | Varchar | 60 | not NULL |  |  |  | Name of the printing house |
| street | Varchar | 70 | not NULL |  |  |  | Street name of the printing house |
| city | Varchar | 20 | not NULL |  |  |  | The city where the printing plant is. |
| country | Varchar | 50 | not NULL |  |  |  | The country where the printing house is located. |
| postcode | Varchar | 8 | not NULL |  |  |  | Postal code of printing house. |
| email | Varchar | 30 | not NULL |  |  |  | Email of printing house. |
| tel | Varchar | 12 | NULL |  |  |  | Phone number of the printing house. |
| contactPerson | Varchar | 50 | NULL |  |  |  | Contact person at the printing house. |

## Bookshop

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| FIELD | TYPE | SIZE | NULL/not NULL | DEFAULT | Constraints | INDEX | Description |
| bookshopId | int | 3 | not NULL |  | Unique | PK | This is a unique bookshop number starting with 100. |
| bookshopName | Varchar | 60 | not NULL |  |  |  | Name of the bookshop. |
| street | Varchar | 70 | not NULL |  |  |  | Street name of the bookshop. |
| city | Varchar | 20 | not NULL |  |  |  | The city where the bookshop is. |
| country | Varchar | 50 | not NULL |  |  |  | The country where the bookshop is located. |
| postcode | Varchar | 8 | not NULL |  |  |  | Postal code of bookshop. |
| email | Varchar | 30 | not NULL |  |  |  | Email of bookshop. |
| tel | Varchar | 12 | NULL |  |  |  | Phone number of the bookshop. |
| contactPerson | Varchar | 50 | NULL |  |  |  | Contact person at the bookshop. |

## Salary

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| FIELD | TYPE | SIZE | NULL/not NULL | DEFAULT | Constraints | INDEX | Description |
| paymentId | Varchar | 4 | not NULL |  | Unique | PK | The first letter is a capital P, and the next three characters are a number starting with 001 which gives a unique ID. |
| payment | decimal | 12 | not NULL | 0 |  |  | The value of the payment transferred to the author of the manuscript. |
| paymentDate | Varchar | 10 | not NULL |  |  |  | Date of payment made to the author. |
| advancePayment | decimal | 12 | NULL | 0 |  |  | The value of the pre-payment transferred to the author of the manuscript. |
| advancePayDate | Varchar | 10 | NULL |  |  |  | Date of pre-payment made to the author. |

## Author

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| FIELD | TYPE | SIZE | NULL/not NULL | DEFAULT | Constraints | INDEX | Description |
| PPSN | Varchar | 9 | not NULL |  | Unique | PK | Unique PPSN number of the author. |
| fName | Varchar | 30 | not NULL |  |  |  | Author's first name. |
| lName | Varchar | 30 | not NULL |  |  |  | Author's surname. |
| street | Varchar | 70 | not NULL |  |  |  | The name of the street where the author lives |
| city | Varchar | 20 | not NULL |  |  |  | The city where the author lives |
| country | Varchar | 50 | not NULL |  |  |  | The country where the author lives |
| postcode | Varchar | 8 | not NULL |  |  |  | Author's postal code. |
| email | Varchar | 30 | not NULL |  |  |  | Author's email. |
| tel | Varchar | 12 | NULL |  |  |  | Author's phone number. |
| biography | Varchar | 3 | NULL |  | Only Yes or No |  | Has the author's biography been delivered to the publisher. ‘Yes’ or ‘No’? |
| paymentId | Varchar | 4 | not NULL |  | Unique | FK | Unique paymentId . It references to the table: Salary. |

## Manuscript

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| FIELD | TYPE | SIZE | NULL/  not NULL | DEFAULT | Constraints | INDEX | Description |
| manuscriptId | Varchar | 5 | not NULL |  | Unique | PK | The first letter is a capital M, and the next four characters are a number starting with 0001 which gives a unique ID and is generated automaticly. |
| manuscriptText | Varchar | 4 | not NULL | Part |  |  | This attribute tells you whether the submitted manuscript is complete or just a fragment. Full or part? Only complete manuscripts can be sent for printing. |
| dateArrived | Varchar | 10 | not NULL |  |  |  | Date of delivery of the Manuscript to the publisher. |
| decisionDate | Varchar | 10 | NULL |  |  |  | The date of the decision to print the book based on the manuscript. |
| pagesNo | int | 4 | NULL |  |  |  | Page number in the manuscript. |
| notesAttached | Varchar | 3 | Null | No | Only Yes or No |  | Have additional information and notes from the author been provided with the manuscript? ‘Yes’ or ‘No’? |
| PPSN | Varchar | 9 | not NULL |  | Unique | FK | Unique PPSN number of the author. It references to the table: Author |
| printHouseId | int | 2 | not NULL |  | Unique | FK | Unique printHouseId. It references to the table: PrintingHouse. |

## NonFiction

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| FIELD | TYPE | SIZE | NULL/not NULL | DEFAULT | Constraints | INDEX | Description |
| manuscriptId | Varchar | 5 | not NULL |  | Unique | PK | Unique manuscriptId . It references to the table: Manuscript |
| discipline | Varchar | 18 | NULL |  |  |  | The discipline to which the content of the book/textbook relates. |

## Fiction

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| FIELD | TYPE | SIZE | NULL/not NULL | DEFAULT | Constraints | INDEX | Description |
| manuscriptId | Varchar | 5 | not NULL |  | Unique | PK | Unique manuscriptId . It references to the table: Manuscript |
| genre | Varchar | 18 | NULL |  |  |  | The literary genre in which the book was written. |

## Book

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| FIELD | TYPE | SIZE | NULL/not NULL | DEFAULT | Constraints | INDEX | Description |
| ISBN | Varchar | 13 | not NULL |  | Unique | PK | A unique ISBN number assigned to books - an international standard. |
| author | Varchar | 62 | not NULL |  |  |  | Name of the author of the book. |
| title | Varchar | 100 | not NULL |  |  |  | The title under which the book was published. |
| publisher | Varchar | 5 | not NULL |  | Only Big A |  | Publisher's name: 'Big A'. |
| price | dec | 6 | NULL | 0 |  |  | Wholesale price to bookstores. |
| publicationDate | Varchar | 10 | not NULL |  |  |  | Publication date of the book. |
| pagesNo | int | 4 | not NULL |  |  |  | The number of pages in the book. |
| bookDescription | Varchar | 3 | not NULL |  | Only Yes or No |  | Is the book description on the cover of the book? ‘Yes’ or ‘No’? |
| printHouseId | int | 2 | not NULL |  | Unique | FK | A unique printHouseId. It references to the table: PrintingHouse. |

## BookSale

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| FIELD | TYPE | SIZE | NULL/not NULL | DEFAULT | Constraints | INDEX | Description |
| saleDate | Varchar | 10 | not NULL |  |  |  | The date of a single book sale to a bookstore. |
| dueDate | Varchar | 10 | not NULL |  |  |  | Date of required payment for sold books to the bookstore. |
| quantity | int | 4 | not NULL | 0 |  |  | The number of books delivered to the Bookshop in one order. |
| saleValue | dec | 8 | not NULL |  |  |  | The value of book sales resulting from the above transaction. |
| termsCond | Varchar | 8 | not NULL |  | Only Standard or Tailored |  | Type of sales terms agreed upon when negotiating the sale. ‘Standard’ or ‘Tailored’. |
| ISBN | Varchar | 13 | not NULL |  | Unique | PK, FK | A unique ISBN number It references to the table: Book. |
| bookshopId | int | 3 | not NULL |  | Unique | PK, FK | A unique printHouseId. It references to the table: Bookshop. |

*The following changes were made to the above table:*

* *renamed the table from 'Book' to 'BookSale'*
* *two PKs were added (ISBN and bookshopId*

# Security & View

There are several categories of users who need different levels of access to the database:

1. Trainees:
   * Access: Read-only access to the Author and Manuscript tables.
   * Reason: Trainees will only write reports for the writers' cooperation department, and they do not require updating privileges.
2. Office staff:
   * Access: Update access to the Printing House and Bookshop tables.
   * Reason: Office staff will supervise the status of book demand and the ordering process in the publishing house, and they need to update the tables accordingly.
3. Manager:
   * Access: Full access to all tables in the database.
   * Reason: The manager will be required to update all tables in the database.
4. Administrator:
   * Access: Full access to all tables in the database.
   * Reason: As the administrator, he needs full access to manage and maintain the database.

To implement these access rights, it is necessary to create different user accounts with the corresponding permissions assigned to them.

The view chosen is primarily designed for trainees and aims to combine the Author and Manuscript tables to provide a comprehensive view of authors and their manuscripts. This view can be selected because it simplifies the querying and retrieval of information about authors and their manuscripts, rather than combining these tables multiple times.

CREATE VIEW author\_manuscript\_view AS

SELECT a.PPSN, a.fName, a.lName, m.manuscriptId, m.manuscriptText, m.dateArrived, m.decisionDate, m.pagesNo, m.notesAttached

FROM Author JOIN Manuscript m ON a.PPSN = m.PPSN;

This query joins the Author and Manuscript tables on the PPSN column and selects specific columns from both tables to create a view with the combined data.

# Queries

1. List of authors living in Ireland
2. Search for all manuscripts cheaper than €20,000
3. Book title and price under the heading book
4. List of books when the full title is not known.
5. Book sales where a single purchase exceeded 200 books.
6. Summary of sales transacted in 2022 with the number and value of sold books.
7. Merging data from 3 tables: 'book', 'book sale' and 'bookshop'.
8. Average number of books sold.
9. Preparation of a summary of payments orders for banks in each city.
10. Book sales over the last year decreased, together with the name of the author, and the name of the printing house.

# Conclusion

The database created for this project provides much of the useful data needed for the publishing house. The Conceptual Data Model perfectly illustrates the relationships in the flow of information within the company. It allows it to be well managed and, with the help of the Logical Data Model, to know its logical dependencies and relationships, taking into consideration the various options for managing tables (Manuscript). Next, it is shown how to calculate the disk space requirements in the Storage Representation. In the Table Design, the data structure and constraints are presented, which is essential for the implementation of the whole project. In the Security & Views paragraph, the aspects of security and views created specifically for use by individual employees with the appropriate permissions are presented. Finally, the section on Querying presents a selection of queries that can provide useful data for publishing. The project provides a good tool to use in the early stages of the publishing business.

After I did the project in the third semester, I was surprised that we would continue it in the fourth term. However, during the process it became clear that it could be approached from a different perspective, as a result of which my logic model changed significantly. Besides, completely new topics like storage representation, table design, view and protection elements were introduced. The project became more comprehensive and more interesting. If the changes were made systematically, it did not cause difficulties, especially since we could always have the support of the course leader. There was enough time for everything. A little beyond the project itself, I want to mention that MCQs without negative marks are a fair and readable form of evaluating students' knowledge.