

2023 Spring CS504 Project

1. Introduction

Welcome to the final project for *CS504 Principles of Data Management and Mining*. This project aims to test your understanding of the concepts covered throughout the course, such as database design, implementation, and querying. You will be required to develop a functional database system for a given scenario, ensuring that the data is efficiently stored and easily accessible.

2. Project Scenario

2.1. Overview

The objective of this project is to design and implement a database management system for a public library. The system will help the library staff manage their resources, including books, magazines, digital media, and other materials. Additionally, it will provide efficient access to library member's information and facilitate borrowing and tracking of library materials. The library management system should ensure data integrity and minimize redundancy for both library staff and members.

2.2. Features

- **Materials Management:** The system should store and maintain information about all library materials, such as *books*, *magazines*, *e-books*, and *audiobooks*, including their *titles*, *authors*, *publication dates*, and *genres*.
- **Membership Management:** The system should store and manage information about library members, including their *names*, *contact information*, *membership numbers*, and *borrowing history*.
- **Borrowing:** The system should facilitate the borrowing process, allowing members to check out items, and providing library staff with the necessary information to manage the circulation of library materials. Once a material is checked out, a librarian should record its borrow date, anticipated due date. And once the material is returned, its return date should be updated.
- **Reporting and Analytics:** The system should generate reports on library usage, popular materials, and other relevant statistics, enabling the library staff to make data-driven decisions about resource acquisition and management.

2.3. Entities and Relationships

1. Material

Represents individual items available in the library, such as books, magazines, e-books, and audiobooks.

Attributes:

- Material_ID: A unique identifier for each material.
- Title: The title of the material.
- Publication Date: The date of publication of the material.
- Catalog_ID: A reference to the catalog entry for the material.
- Genre_ID: A reference to the genre of the material.

2. Catalog

Represents a record of library materials with information on their availability and location.

Attributes:

- Catalog_ID: A unique identifier for each catalog entry.
- Name: The name of the catalog.
- Location: The location of the material within the library.

3. Genre

Represents the various genres or categories of library materials.

Attributes:

- Genre_ID: A unique identifier for each genre.
- Name: The name of the genre.
- Description: The brief introduction of the genre.

4. Borrow

Represents the borrowing activity of library materials by members.

Attributes:

- Borrow_ID: A unique identifier for each borrowing transaction.
- Material_ID: A reference to the borrowed material.
- Member_ID: A reference to the member who borrowed the material.
- Staff_ID: A reference to the staff who processed the transaction.
- Borrow Date: The date the material was borrowed.
- Due Date: The date the material is due.
- Return Date: The date the material is returned.

5. Author

Represents authors who have created library materials.

Attributes:

- Author_ID: A unique identifier for each author.
- Name: The name of the author.
- Birth Date: The birth date of the author.

- Nationality: The nationality of the author.
6. Authorship
- Represents the relationship between authors and the materials they have created.
- Attributes:**
- Authorship_ID: A unique identifier for each authorship record.
 - Author_ID: A reference to the author.
 - Material_ID: A reference to the material authored.
7. Member
- Represents library members who can borrow and reserve materials.
- Attributes:**
- Member_ID: A unique identifier for each member.
 - Name: The name of the member.
 - Contact_Info: Email address (or phone number) of the member.
 - Join_Date: The date the member joined the library.
8. Staff
- Represents library staff who manage library resources and assist members.
- Attributes:**
- Staff_ID: A unique identifier for each staff member.
 - Name: The name of the staff member.
 - Contact_Info: Email address (or phone number) of the member.
 - Job_Title: The job title of the staff member (e.g., librarian, assistant librarian).
 - Hire_Date: The date the staff member was hired by the library.

3. Requirements

3.1. Database Design

1. Define the scope of the project and identify the entities and their relationships.
2. Create an Entity-Relationship (ER) diagram to represent the database schema.
3. (Optional) Normalize the database schema to ensure data integrity and minimize redundancy.

3.2. Database Implementation

1. Choose an appropriate Database Management System (DBMS) for the project. We recommend using Postgre.
2. Implement the database schema using SQL or the chosen DBMS's data definition language (DDL).
3. Populate the database with sample data.

3.3. Querying and Manipulation

Example queries are included in the Tests section below.

1. Develop a set of SQL queries or stored procedures to perform common tasks, such as searching, updating, inserting, and deleting records.
2. Demonstrate the use of advanced querying techniques, such as joins, aggregation, and subqueries.

3.4. Documentation and Presentation

1. Prepare a project report that covers the design and implementation
2. Include the ER diagram, SQL code (DDL and DML), and sample output for each query.

4. Tests and Evaluation

Your final project will be evaluated based on the following criteria:

1. (30pts) Database Design: Correctness and completeness of the ER diagram and constraints.
2. (10pts) Implementation: Functionality of the implemented SQL to define the database schema and insert sample data.
3. (40pts) Querying and Manipulation: Correctness of SQL queries or updates. Reasonable system designs.
4. (20pts) Documentation: Quality of the project report.

4.1. Queries/Updates

1. Which materials are currently available in the library?
2. Which materials are currently overdue?
Suppose today is 04/01/2023, and show the borrow date and due date of each material
3. What are the top 10 most borrowed materials in the library?
Show the title of each material and order them based on their available counts
4. How many books has the author Lucas Piki written?
5. How many books were written by two or more authors?
6. What are the most popular genres in the library?
7. How many materials have been borrowed from 09/2020-10/2020?
8. How do you update the "Harry Potter and the Philosopher's Stone" when it is returned on 04/01/2023?
9. How do you delete the member Emily Miller and all her related records from the database?
10. How do you add the following material to the database?

Title: New book

Date: 2020-08-01

Catalog: E-Books

Genre: Mystery & Thriller
Author: Lucas Pipi

4.2. Design

You are required to explain how to extend the existing database system to incorporate the following features. While you can provide SQL statements to illustrate your concepts, you don't need to execute the statements.

1. Alert staff about overdue materials on a daily-basis?
2. Automatically deactivate the membership based on the member's overdue occurrence (\geq three times). And reactivate the membership once the member pays the overdue fee.

5. Submission Guidelines

1. **Honor Code:** You are to work on this project alone and not in groups. All projects will be checked for plagiarism against each other and against the Internet using software. Suspicious projects will be sent to the honor committee for further review.
2. Submit your project report in **PDF (or docx) format**, including the description, ER diagram, constraints, answers to Section 4, including sample output.
3. Submit a compressed file in **tar (or zip) format**, containing the report, the project code, sample data, instructions to load/execute the code, and any additional resources used.
4. **Postgre** compatible SQL statements are preferred.

Please feel free to reach out to the instructor or TA if you have any questions or need clarification on any aspect of the project. Good luck!

6. Data

Material:

Material_ID	Title	Publication_Date	Catalog_ID	Genre_ID
1	The Catcher in the Rye	1951-07-16	1	1
2	To Kill a Mockingbird	1960-07-11	2	1
3	The Da Vinci Code	2003-04-01	3	2
4	The Hobbit	1937-09-21	4	3
5	The Shining	1977-01-28	5	4
6	Pride and Prejudice	1813-01-28	1	1
7	The Great Gatsby	1925-04-10	2	1
8	Moby Dick	1851-10-18	3	1
9	Crime and Punishment	1866-01-01	4	1
10	The Hitchhiker's Guide to the Galaxy	1979-10-12	5	3
11	1984	1949-06-08	1	5
12	Animal Farm	1945-08-17	2	5

13	The Haunting of Hill House	1959-10-17	3	4
14	Brave New World	1932-08-01	4	5
15	The Chronicles of Narnia: The Lion, the Witch and the Wardrobe	1950-10-16	5	3
16	The Adventures of Huckleberry Finn	1884-12-10	6	1
17	The Catch-22	1961-10-11	7	1
18	The Picture of Dorian Gray	1890-07-01	8	1
19	The Call of Cthulhu	1928-02-01	9	4
20	Harry Potter and the Philosopher's Stone	1997-06-26	10	3
21	Frankenstein	1818-01-01	6	4
22	A Tale of Two Cities	1859-04-30	7	1
23	The Iliad	1750-01-01	8	6
24	The Odyssey	1725-01-01	9	6
25	The Brothers Karamazov	1880-01-01	10	1
26	The Divine Comedy	1320-01-01	6	6
27	The Grapes of Wrath	1939-04-14	7	1
28	The Old Man and the Sea	1952-09-01	8	1
29	The Count of Monte Cristo	1844-01-01	9	1
30	A Midsummer Night's Dream	1596-01-01	10	7
31	The Tricky Book	1888-01-01	10	7

Catalog:

Catalog_ID	Name	Location
1	Books	A1.1
2	Magazines	B2.1
3	E-Books	C3.1
4	Audiobooks	D4.1
5	Journals	E5.1
6	Newspaper	F6.1
7	Maps	G7.1
8	Novels	H8.1
9	Sheet Music	I9.1
10	Educational	J10.1

Genre:

Genre_ID	Name	Description
1	General Fiction	Literary works with a focus on character and plot development, exploring various themes and human experiences.
2	Mystery & Thriller	Suspenseful stories centered around crime, investigation, or espionage with an emphasis on tension and excitement.

- 3 | Science Fiction & Fantasy | Imaginative works that explore alternate realities, futuristic concepts, and magical or supernatural elements.
- 4 | Horror & Suspense | Stories designed to evoke fear, unease, or dread, often featuring supernatural or psychological elements.
- 5 | Dystopian & Apocalyptic | Depictions of societies in decline or collapse, often exploring themes of political and social oppression or environmental disaster.
- 6 | Classics | Enduring works of literature that have stood the test of time, often featuring rich language and complex themes.
- 7 | Historical Fiction | Fictional stories set in the past, often based on real historical events or figures, and exploring the customs and experiences of that time.
- 8 | Epic Poetry & Mythology | Ancient or traditional stories and poems, often featuring heroes, gods, and mythical creatures, and exploring cultural values and beliefs.

Borrow:

Borrow_ID | Material_ID | Member_ID | Staff_ID | Borrow_Date | Due_Date | Return_Date

-----	-----	-----	-----	-----	-----	-----
1	1	1	1	2018-09-12	2018-10-03	2018-09-30
2	2	2	1	2018-10-15	2018-11-05	2018-10-29
3	3	3	1	2018-12-20	2019-01-10	2019-01-08
4	4	4	1	2019-03-11	2019-04-01	2019-03-27
5	5	5	1	2019-04-20	2019-05-11	2019-05-05
6	6	6	1	2019-07-05	2019-07-26	2019-07-21
7	7	7	1	2019-09-10	2019-10-01	2019-09-25
8	8	8	1	2019-11-08	2019-11-29	2019-11-20
9	9	9	1	2020-01-15	2020-02-05	2020-02-03
10	10	10	1	2020-03-12	2020-04-02	2020-03-28
11	1	11	2	2020-05-14	2020-06-04	2020-05-28
12	2	12	2	2020-07-21	2020-08-11	2020-08-02
13	3	13	2	2020-09-25	2020-10-16	2020-10-15
14	4	1	2	2020-11-08	2020-11-29	2020-11-24
15	5	2	2	2021-01-03	2021-01-24	2021-01-19
16	6	3	2	2021-02-18	2021-03-11	2021-03-12
17	17	4	2	2021-04-27	2021-05-18	2021-05-20
18	18	5	2	2021-06-13	2021-07-04	2021-06-28
19	19	6	2	2021-08-15	2021-09-05	2021-09-03
20	20	7	2	2021-10-21	2021-11-11	2021-11-05
21	21	1	3	2021-11-29	2021-12-20	
22	22	2	3	2022-01-10	2022-01-31	2022-01-25
23	23	3	3	2022-02-07	2022-02-28	2022-02-23
24	24	4	3	2022-03-11	2022-04-01	2022-03-28
25	25	5	3	2022-04-28	2022-05-19	2022-05-18
26	26	6	3	2022-06-22	2022-07-13	2022-07-08
27	27	7	3	2022-08-04	2022-08-25	2022-08-23
28	28	8	3	2022-09-13	2022-10-04	2022-09-28

29	29	9	3	2022-10-16	2022-11-06	2022-11-05
30	30	8	3	2022-11-21	2022-12-12	2022-12-05
31	1	9	4	2022-12-28	2023-01-18	
32	2	1	4	2023-01-23	2023-02-13	
33	3	10	4	2023-02-02	2023-02-23	2023-02-17
34	4	11	4	2023-03-01	2023-03-22	
35	5	12	4	2023-03-10	2023-03-31	
36	6	13	4	2023-03-15	2023-04-05	
37	7	17	4	2023-03-25	2023-04-15	
38	8	8	4	2023-03-30	2023-04-20	
39	9	9	4	2023-03-26	2023-04-16	
40	10	20	4	2023-03-28	2023-04-18	

Author:

Author_ID	Name	Nationality
----- ----- ----- -----		
1	Jane Austen	1775-12-16 British
2	Ernest Hemingway	1899-07-21 American
3	George Orwell	1903-06-25 British
4	Scott Fitzgerald	1896-09-24 American
5	J.K. Rowling	1965-07-31 British
6	Mark Twain	1835-11-30 American
7	Leo Tolstoy	1828-09-09 Russian
8	Virginia Woolf	1882-01-25 British
9	Gabriel Márquez	1927-03-06 Colombian
10	Charles Dickens	1812-02-07 British
11	Harper Lee	1926-04-28 American
12	Oscar Wilde	1854-10-16 Irish
13	William Shakespeare	1564-04-26 British
14	Franz Kafka	1883-07-03 Czech
15	James Joyce	1882-02-02 Irish
16	J.R.R. Tolkien	1892-01-03 British
17	Emily Brontë	1818-07-30 British
18	Toni Morrison	1931-02-18 American
19	Fyodor Dostoevsky	1821-11-11 Russian
20	Lucas Piki	1847-10-16 British

Authorship:

Authorship_ID	Author_ID	Material_ID
----- ----- -----		
1	1	1
2	2	2

3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	1	21
22	2	22
22	3	22
23	3	23
24	4	24
25	5	25
26	6	26
27	7	27
28	8	28
29	19	28
30	9	29
31	10	30
32	8	30
33	2	29

Member:

Member_ID	Name	Contact_Info	Join_Date
-----	-----	-----	-----
1	Alice Johnson	alice.johnson@email.com	2018-01-10
2	Bob Smith	bob.smith@email.com	2018-03-15
3	Carol Brown	carol.brown@email.com	2018-06-20
4	David Williams	david.williams@email.com	2018-09-18
5	Emily Miller	emily.miller@email.com	2019-02-12
6	Frank Davis	frank.davis@email.com	2019-05-25
7	Grace Wilson	grace.wilson@email.com	2019-08-15

8	Harry Garcia		harry.garcia@email.com		2019-11-27
9	Isla Thomas		isla.thomas@email.com		2020-03-04
10	Jack Martinez		jack.martinez@email.com		2020-07-01
11	Kate Anderson		kate.anderson@email.com		2020-09-30
12	Luke Jackson		luke.jackson@email.com		2021-01-18
13	Mia White		mia.white@email.com		2021-04-27
14	Noah Harris		noah.harris@email.com		2021-07-13
15	Olivia Clark		olivia.clark@email.com		2021-10-05
16	Peter Lewis		peter.lewis@email.com		2021-12-01
17	Quinn Hall		quinn.hall@email.com		2022-02-28
18	Rachel Young		rachel.young@email.com		2022-06-17
19	Sam Walker		sam.walker@email.com		2022-09-25
20	Tiffany Allen		tiffany.allen@email.com		2022-12-10

Staff:

Staff_ID	Name	Contact_Info	Job_Title	Hire_Date
-----	-----	-----	-----	-----
1	Amy Green	amy.green@email.com	Librarian	2017-06-01
2	Brian Taylor	brian.taylor@email.com	Library Assistant	2018-11-15
3	Christine King	chris.king@email.com	Library Assistant	2019-05-20
4	Daniel Wright	dan.wright@email.com	Library Technician	2020-02-01