

Databases

INTRODUCTION TO SQL



Izzy Weber

Curriculum Manager, DataCamp

Course goals

1. Understand databases and their structure → Chapter 1
2. Extract information from databases using SQL → Chapter 2

Introducing databases

patrons

| card_num | name | member_year | total_fine |
|----------|--------|-------------|------------|
| 54378 | Izzy | 2012 | 9.86 |
| 94722 | Maham | 2020 | 0 |
| 45783 | Jasmin | 2022 | 2.05 |
| 90123 | James | 1989 | 0 |

books

| id | title | author | genre | pub_year |
|-----|---------------------------|----------------|-------------|----------|
| 638 | Being Mortal | Atul Gawande | Non-Fiction | 2015 |
| 912 | Educated | Tara Westover | Non-Fiction | 2018 |
| 322 | Night | Elie Wiesel | Non-Fiction | 1956 |
| 156 | Where the Wild Things Are | Maurice Sendak | Childrens | 1963 |

checkouts

| id | start_date | due_date | card_num | book_id |
|-----|------------|------------|----------|---------|
| 567 | 2022-05-13 | 2022-05-27 | 54378 | 638 |
| 568 | 2022-06-10 | 2022-06-24 | 54378 | 322 |
| 569 | 2022-06-27 | 2022-07-11 | 45783 | 156 |
| 570 | 2022-08-14 | 2022-08-28 | 90123 | 912 |

Introducing databases

patrons

| card_num | name | member_year | total_fine |
|----------|--------|-------------|------------|
| 54378 | Izzy | 2012 | 9.86 |
| 94722 | Maham | 2020 | 0 |
| 45783 | Jasmin | 2022 | 2.05 |
| 90123 | James | 1989 | 0 |

Relational databases

- Define relationships between tables of data inside the database

patrons

| card_num | name | member_year | total_fine |
|----------|--------|-------------|------------|
| 54378 | Izzy | 2012 | 9.86 |
| 94722 | Maham | 2020 | 0 |
| 45783 | Jasmin | 2022 | 2.05 |
| 90123 | James | 1989 | 0 |

books

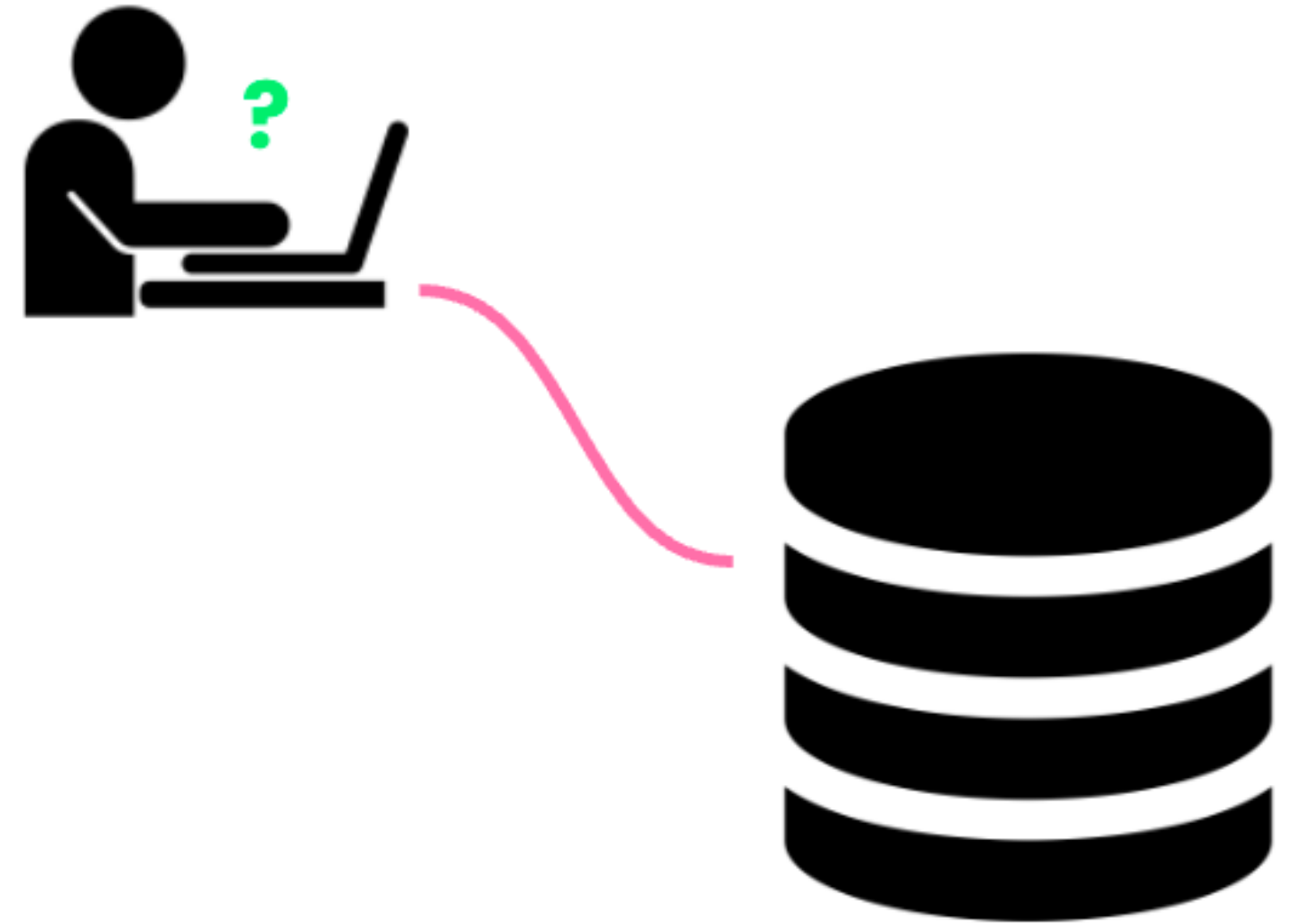
| id | title | author | genre | pub_year |
|-----|---------------------------|----------------|-------------|----------|
| 638 | Being Mortal | Atul Gawande | Non-Fiction | 2015 |
| 912 | Educated | Tara Westover | Non-Fiction | 2018 |
| 322 | Night | Elie Wiesel | Non-Fiction | 1956 |
| 156 | Where the Wild Things Are | Maurice Sendak | Childrens | 1963 |

checkouts

| id | start_date | due_date | card_num | book_id |
|-----|------------|------------|----------|---------|
| 567 | 2022-05-13 | 2022-05-27 | 54378 | 638 |
| 568 | 2022-06-10 | 2022-06-24 | 54378 | 322 |
| 569 | 2022-06-27 | 2022-07-11 | 45783 | 156 |
| 570 | 2022-08-14 | 2022-08-28 | 90123 | 912 |

Database advantages

- More storage than spreadsheet applications
- Storage is more secure



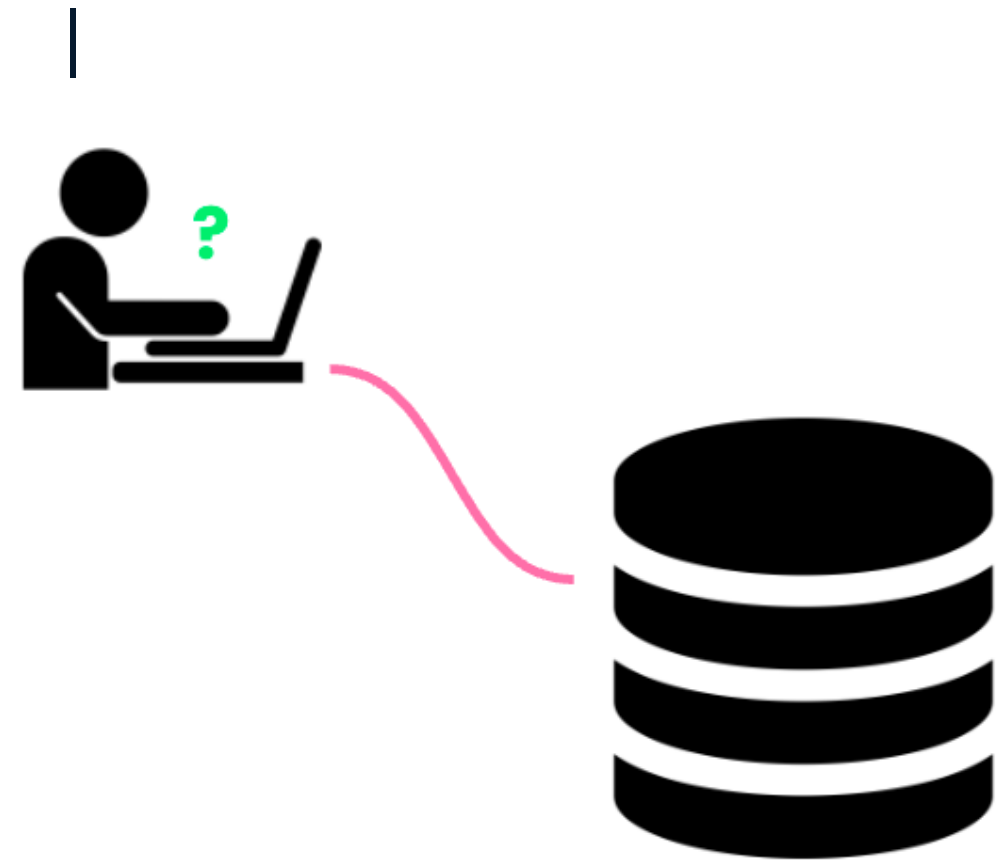
Database advantages



SQL

- Short for Structured Query Language
- The most widely used programming language for databases

```
SELECT *  
FROM patrons  
LIMIT 30
```



Let's practice!

INTRODUCTION TO SQL

Tables

INTRODUCTION TO SQL



Izzy Weber

Curriculum Manager, DataCamp

A seat at the table

- Table rows and columns are referred to as *records* and *fields*
- Fields are set at database creation; there is no limit to the number of records

patrons

| card_num | name | member_year | total_fine |
|----------|--------|-------------|------------|
| 54378 | Izzy | 2012 | 9.86 |
| 94722 | Maham | 2020 | 0 |
| 45783 | Jasmin | 2022 | 2.05 |
| 90123 | James | 1989 | 0 |

books

| id | title | author | genre | pub_year |
|-----|---------------------------|----------------|-------------|----------|
| 638 | Being Mortal | Atul Gawande | Non-Fiction | 2015 |
| 912 | Educated | Tara Westover | Non-Fiction | 2018 |
| 322 | Night | Elie Wiesel | Non-Fiction | 1956 |
| 156 | Where the Wild Things Are | Maurice Sendak | Childrens | 1963 |

checkouts

| id | start_date | due_date | card_num | book_id |
|-----|------------|------------|----------|---------|
| 567 | 2022-05-13 | 2022-05-27 | 54378 | 638 |
| 568 | 2022-06-10 | 2022-06-24 | 54378 | 322 |
| 569 | 2022-06-27 | 2022-07-11 | 45783 | 156 |
| 570 | 2022-08-14 | 2022-08-28 | 90123 | 912 |

Good table manners

Table names should...

- be lowercase
- have no spaces—use underscores instead
- refer to a collective group or be plural



patrons

| card_num | name | member_year | total_fine |
|----------|--------|-------------|------------|
| 54378 | Izzy | 2012 | 9.86 |
| 94722 | Maham | 2020 | 0 |
| 45783 | Jasmin | 2022 | 2.05 |
| 90123 | James | 1989 | 0 |

Laying the table: records

A record is a row that holds data on an individual observation

patrons

| card_num | name | member_year | total_fine |
|----------|--------|-------------|------------|
| 54378 | Izzy | 2012 | 9.86 |
| 94722 | Maham | 2020 | 0 |
| 45783 | Jasmin | 2022 | 2.05 |
| 90123 | James | 1989 | 0 |

a record

Laying the table: fields

A field is a column that holds one piece of information about all records

a field

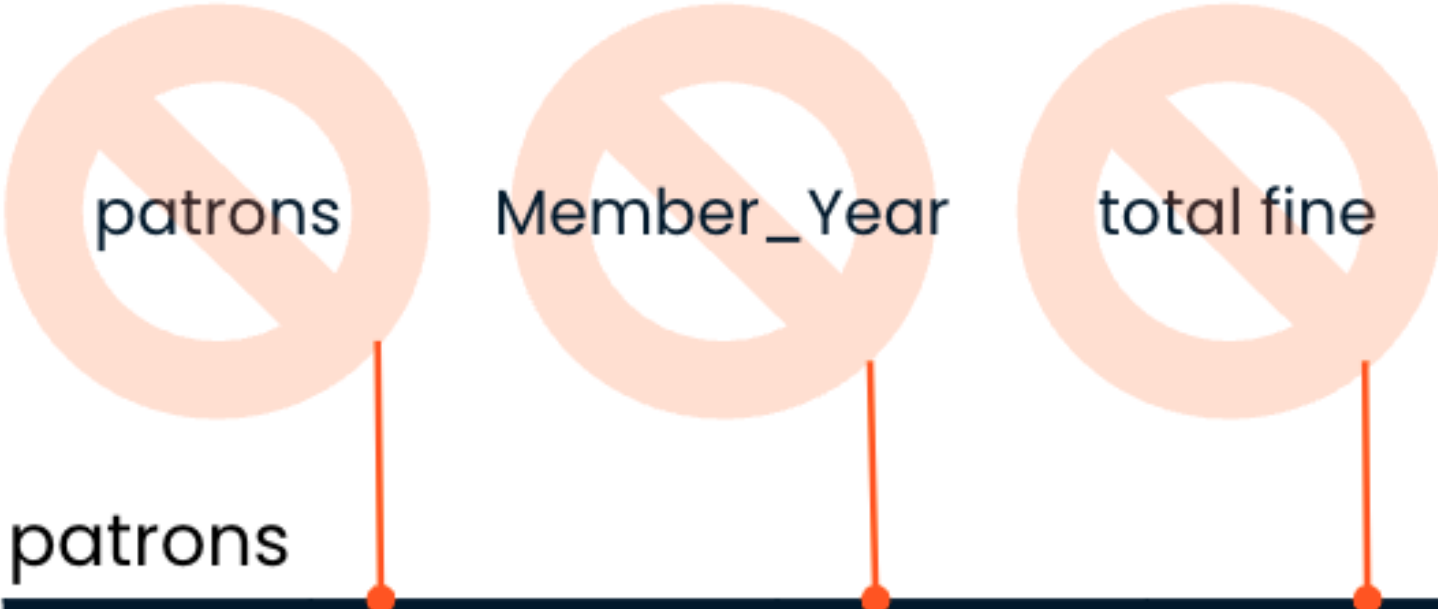
patrons

| card_num | name | member_year | total_fine |
|----------|--------|-------------|------------|
| 54378 | Izzy | 2012 | 9.86 |
| 94722 | Maham | 2020 | 0 |
| 45783 | Jasmin | 2022 | 2.05 |
| 90123 | James | 1989 | 0 |

More table manners

Field names should...

- be lowercase
- have no spaces
- be singular
- be different from other field names
- be different from the table name



The diagram shows three examples of bad field names, each inside a light orange circle with a diagonal slash (prohibition sign). The names are 'patrons', 'Member_Year', and 'total fine'. Red lines connect these names to the corresponding columns in the table below.

| patrons | | | |
|----------|--------|-------------|------------|
| card_num | name | member_year | total_fine |
| 54378 | Izzy | 2012 | 9.86 |
| 94722 | Maham | 2020 | 0 |
| 45783 | Jasmin | 2022 | 2.05 |
| 90123 | James | 1989 | 0 |

Assigned seats

- *Unique identifiers* are used to identify records in a table
- They are unique and often numbers

unique identifier

patrons

| card_num | name | member_year | total_fine |
|----------|--------|-------------|------------|
| 54378 | Izzy | 2012 | 9.86 |
| 94722 | Maham | 2020 | 0 |
| 45783 | Jasmin | 2022 | 2.05 |
| 90123 | James | 1989 | 0 |

The more the merrier

patrons

| card_num | name | member_year | total_fine |
|----------|--------|-------------|------------|
| 54378 | Izzy | 2012 | 9.86 |
| 94722 | Maham | 2020 | 0 |
| 45783 | Jasmin | 2022 | 2.05 |
| 90123 | James | 1989 | 0 |

checkouts

| id | start_date | due_date | card_num | book_id |
|-----|------------|------------|----------|---------|
| 567 | 2022-05-13 | 2022-05-27 | 54378 | 638 |
| 568 | 2022-06-10 | 2022-06-24 | 54378 | 322 |
| 569 | 2022-06-27 | 2022-07-11 | 45783 | 156 |
| 570 | 2022-08-14 | 2022-08-28 | 90123 | 912 |

patron_checkouts

| card_num | name | member_year | total_fine | checkout_id | start_date | due_date | book_id |
|----------|--------|-------------|------------|-------------|------------|------------|---------|
| 54378 | Izzy | 2012 | 9.86 | 567 | 2022-05-13 | 2022-05-27 | 638 |
| 54378 | Izzy | 2012 | 9.86 | 568 | 2022-06-10 | 2022-06-24 | 322 |
| 94722 | Maham | 2020 | 0 | | | | |
| 45783 | Jasmin | 2022 | 2.05 | 2022-06-27 | 2022-07-11 | 45783 | 156 |
| 90123 | James | 1989 | 0 | 570 | 2022-08-14 | 2022-08-28 | 912 |

Let's practice!

INTRODUCTION TO SQL

Data

INTRODUCTION TO SQL



Izzy Weber

Curriculum Manager, DataCamp

SQL data types

all one data type

all one data type

all one data type

all one data type

patrons

| card_num | name | member_year | total_fine |
|----------|--------|-------------|------------|
| 54378 | Izzy | 2012 | 9.86 |
| 94722 | Maham | 2020 | 0 |
| 45783 | Jasmin | 2022 | 2.05 |
| 90123 | James | 1989 | 0 |

- Different types of data are stored differently and take up different space
- Some operations only apply to certain data types

Strings

a string field

patrons

| card_num | name | member_year | total_fine |
|----------|--------|-------------|------------|
| 54378 | Izzy | 2012 | 9.86 |
| 94722 | Maham | 2020 | 0 |
| 45783 | Jasmin | 2022 | 2.05 |
| 90123 | James | 1989 | 0 |

- A string is a sequence of characters such as letters or punctuation
- `VARCHAR` is a flexible and popular string data type in SQL

Integers

an integer field

patrons

| card_num | name | member_year | total_fine |
|----------|--------|-------------|------------|
| 54378 | Izzy | 2012 | 9.86 |
| 94722 | Maham | 2020 | 0 |
| 45783 | Jasmin | 2022 | 2.05 |
| 90123 | James | 1989 | 0 |

- Integers store whole numbers
- `INT` is a flexible and popular integer data type in SQL

Floats

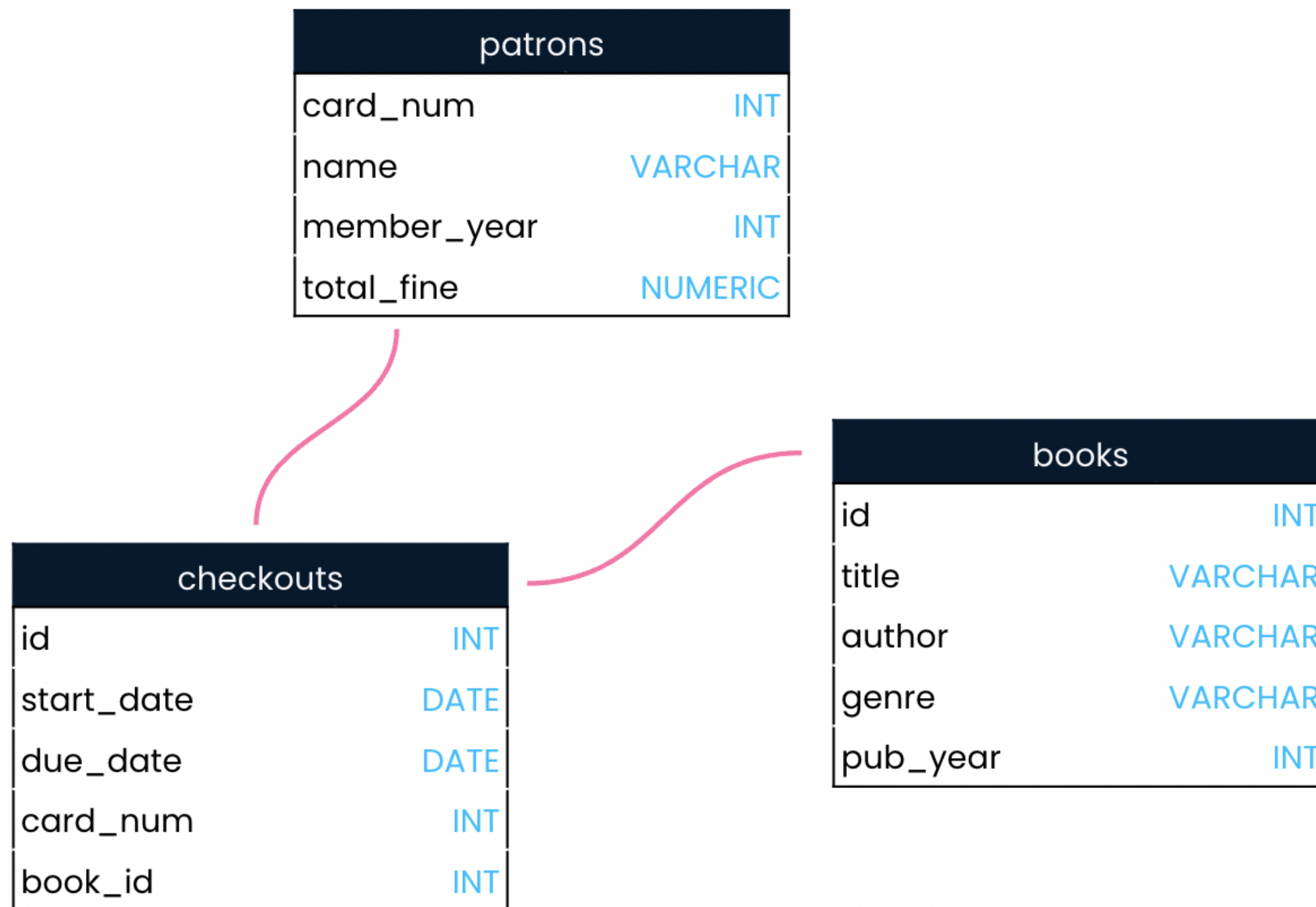
a float field

patrons

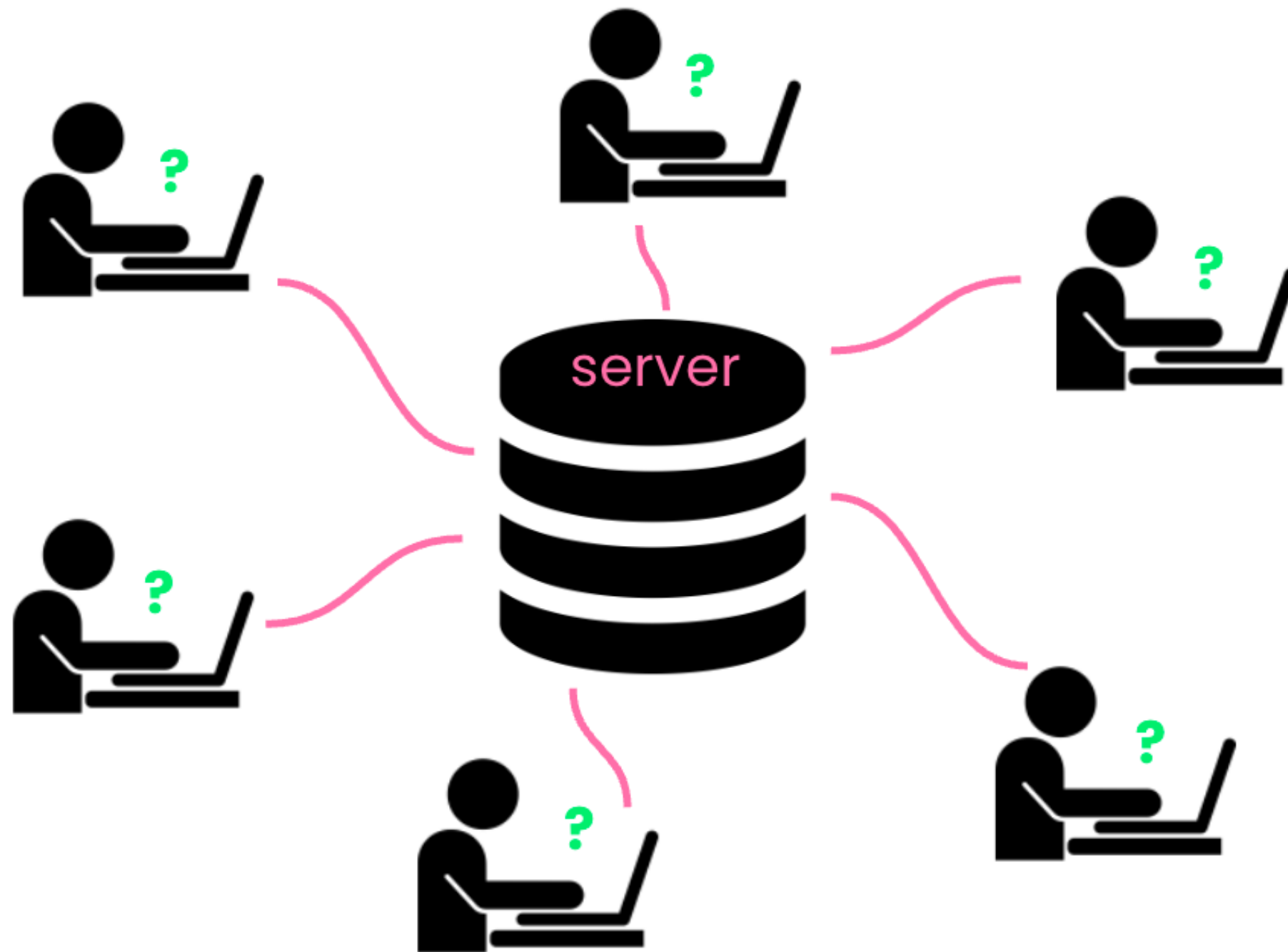
| card_num | name | member_year | total_fine |
|----------|--------|-------------|------------|
| 54378 | Izzy | 2012 | 9.86 |
| 94722 | Maham | 2020 | 0 |
| 45783 | Jasmin | 2022 | 2.05 |
| 90123 | James | 1989 | 0 |

- Floats store numbers that include a fractional part
- `NUMERIC` is a flexible and popular float data type in SQL

Schemas



Database storage



Let's practice!

INTRODUCTION TO SQL