Objectives

After completing this lab, you will be able to:

- Minimize data redundancy and inconsistency in a database by using normalization.
- Use keys to uniquely identify a record in a table, establish a relationship between tables, and identify the relation between them.
- Maintain data integrity in a relational data model using constraints.

Datasette is an open source multi-tool for exploring and publishing data.
BookShop.sql

```
DROP TABLE IF EXISTS BookShop;
DROP TABLE IF EXISTS BookShop_AuthorDetails;

-- Create the table

CREATE TABLE BookShop (
        BOOK_ID VARCHAR(4) NOT NULL,
        TITLE VARCHAR(100) NOT NULL,
        AUTHOR_NAME VARCHAR(30) NOT NULL,
        AUTHOR_BIO VARCHAR(250),
        AUTHOR_ID INTEGER NOT NULL,
        PUBLICATION_DATE DATE NOT NULL,
        PRICE_USD DECIMAL(6,2) CHECK(Price_USD>0) NOT NULL
        );
-- Insert sample data into the table

INSERT INTO BookShop VALUES
('B101', 'Introduction to Algorithms', 'Thomas H. Cormen', 'Thomas H. Cormen is the co-author of Introduction to Algorithms, along with Charles Leiserson, Ron Rivest, and Cliff Stein. He is a Full Professor of computer science at Dartmouth College and currently Chair of the Dartmouth College Writing Program.', 123 , '2001-09-01', 125),
('B201', 'Structure and Interpretation of Computer Programs', 'Harold Abelson', 'Harold Abelson, Ph.D., is Class of 1922 Professor of Computer Science and Engineering in the Department of Electrical Engineering and Computer Science at MIT and a fellow of the IEEE.', 456, '1996-07-25', 65.5),
('B301', 'Deep Learning', 'Ian Goodfellow', 'Ian J. Goodfellow is a researcher working in machine learning, currently employed at Apple Inc. as its director of machine learning in the Special Projects Group. He was previously employed as a research scientist at Google Brain.', 369, '2016-11-01', 82.7),
('B401', 'Algorithms Unlocked', 'Thomas H. Cormen', 'Thomas H. Cormen is the co-author of Introduction to Algorithms, along with Charles Leiserson, Ron Rivest, and Cliff Stein. He is a Full Professor of computer science at Dartmouth College and currently Chair of the Dartmouth College Writing Program.', 123, '2013-05-15', 36.5),
('B501', 'Machine Learning: A Probabilistic Perspective', 'Kevin P. Murphy', '', 157, '2012-08-24', 46);
```

-- Retrieve 1 records from the table
SELECT * FROM BookShop LIMIT 1

BOOK_ID	TITLE	AUTHOR_NAME	AUTHOR_BIO	AUTHOR_ID	PUBLICATION_DATE	PRICE_USD
B101	Introduction to Algorithms	Thomas H. Cormen	Thomas H. Cormen is the co-author of Introduction to Algorithms, along with Charles Leiserson, Ron Rivest, and Cliff Stein. He is a Full Professor of computer science at Dartmouth College and currently Chair of the Dartmouth College Writing Program.	123	2001-09-01	125

Task A: First normal form (1NF)

2 question for checking 1NF

1)Does the above table have unique rows? yes

2)Does each cell of the above table have single/atomic value? Yes

Then it is 1NF

Task B: Second normal form (2NF)

a relation is in second normal form if it is already in 1NF and does not contain any partial dependencies.

To transform 1 NF to 2 NF
the author information such as AUTHOR_ID, AUTHOR_NAME and AUTHOR_BIO out of the BookShop table into another table.THUS , reducing redundancy and increasing consistency of data. Thus 2NF is ensured. Author with multiple books will be problem then create new table

```
CREATE TABLE BookShop_AuthorDetails
(AUTHOR_ID INTEGER NOT NULL,AUTHOR_NAME VARCHAR(30) NOT NULL,
 AUTHOR_BIO VARCHAR(250),PRIMARY KEY (AUTHOR_ID)) ;

-----Insert the records of Bookshop to this table.

insert into BookShop_AuthorDetails select DISTINCT AUTHOR_ID, AUTHOR_NAME,
AUTHOR_BIO FROM BookShop;

select * from BookShop_AuthorDetails;
```

Task A: Primary Key

--Drop the table.

```
DROP TABLE IF EXISTS BookShop;
```

-----Recreate it with Primary Key -------

```
CREATE TABLE BookShop (
 BOOK_ID VARCHAR(4) NOT NULL,
 TITLE VARCHAR(100) NOT NULL,
 AUTHOR_NAME VARCHAR(30) NOT NULL,
 AUTHOR_BIO VARCHAR(250),
 AUTHOR_ID INTEGER NOT NULL,
 PUBLICATION_DATE DATE NOT NULL,
 PRICE_USD DECIMAL(6,2) CHECK(Price_USD>0) NOT NULL,PRIMARY KEY (BOOK_ID));

INSERT INTO BookShop VALUES
('B101', 'Introduction to Algorithms', 'Thomas H. Cormen', 'Thomas H. Cormen is the co-author of Introduction to Algorithms, along with Charles Leiserson, Ron Rivest, and Cliff Stein. He is a Full Professor of computer science at Dartmouth College and currently Chair of the Dartmouth College Writing Program.', 123 , '2001-09-01', 125),
('B201', 'Structure and Interpretation of Computer Programs', 'Harold Abelson', 'Harold Abelson, Ph.D., is Class of 1922 Professor of Computer Science and Engineering in the Department of Electrical Engineering and Computer Science at MIT and a fellow of the IEEE.', 456, '1996-07-25', 65.5),
('B301', 'Deep Learning', 'Ian Goodfellow', 'Ian J. Goodfellow is a researcher working in machine learning, currently employed at Apple Inc. as its director of machine learning in the Special Projects Group. He was previously employed as a research scientist at Google Brain.', 369, '2016-11-01', 82.7),
('B401', 'Algorithms Unlocked', 'Thomas H. Cormen', 'Thomas H. Cormen is the co-author of Introduction to Algorithms, along with Charles Leiserson, Ron Rivest, and Cliff Stein. He is a Full Professor of computer science at Dartmouth College and currently Chair of the Dartmouth College Writing Program.', 123, '2013-05-15', 36.5),
('B501', 'Machine Learning: A Probabilistic Perspective', 'Kevin P. Murphy', '', 157, '2012-08-24', 46);
```

-----Create another table BookShop_AuthorDetails with author id as the primary key

```
CREATE TABLE BookShop_AuthorDetails(AUTHOR_ID INTEGER NOT NULL,AUTHOR_NAME VARCHAR(30) NOT NULL,AUTHOR_BIO VARCHAR(250),PRIMARY KEY (AUTHOR_ID)) ;
```

-----Insert the records of Bookshop to this table.

```
insert into BookShop_AuthorDetails select DISTINCT AUTHOR_ID, AUTHOR_NAME, AUTHOR_BIO FROM BookShop;
```

select * from BookShop_AuthorDetails LIMIT 1

AUTHOR_ID	AUTHOR_NAME	AUTHOR_BIO
123	Thomas H. Cormen	Thomas H. Cormen is the co-author of Introduction to Algorithms, along with Charles Leiserson, Ron Rivest, and Cliff Stein. He is a Full Professor of computer science at Dartmouth College and currently Chair of the Dartmouth College Writing Program.

INSERT INTO BookShop VALUES

('B101', 'Introduction to Algorithms', 'Thomas H. Cormen', 'Thomas H. Cormen is the co-author of Introduction to Algorithms, along with Charles Leiserson, Ron Rivest, and Cliff Stein. He is a Full Professor of computer science at Dartmouth College and currently Chair of the Dartmouth College Writing Program.', 123 , '2001-09-01', 125)

a foreign key is a column that establishes a relationship between two tables. It acts as a cross-reference between two tables because it points to the primary key of another table. A table can have multiple foreign keys referencing primary keys of other tables.

Rules for defining a foreign key:

- **A foreign key in the referencing table must match the structure and data type of the existing primary key in the referenced table.**
- **A foreign key can only have values present in the referenced primary key**
- **Foreign keys do not need to be unique. Most often they are not.**
- **Foreign keys can be null.**

```
DROP TABLE IF EXISTS BookShop;

CREATE TABLE BookShop (
 BOOK_ID VARCHAR(4) NOT NULL,
 TITLE VARCHAR(100) NOT NULL,
 AUTHOR_NAME VARCHAR(30) NOT NULL,
 AUTHOR_BIO VARCHAR(250),
 AUTHOR_ID INTEGER NOT NULL,
 PUBLICATION_DATE DATE NOT NULL,
 PRICE_USD DECIMAL(6,2) CHECK(Price_USD>0) NOT NULL,PRIMARY KEY (BOOK_ID),
FOREIGN KEY (AUTHOR_ID)
REFERENCES BookShop_AuthorDetails(AUTHOR_ID)
ON UPDATE NO ACTION
ON DELETE NO ACTION);
```

Exercise 3: Constraints

1. **Entity Integrity Constraint**: Entity integrity ensures that no duplicate records exist within a table and that the column identifing each record within the table is not a duplicate and not null. The existence of a primary key in both the BookShop and BookShop_AuthorDetails tables satisfies this integrity constraint because a primary key mandates NOT NULL constraint as well as ensuring that every row in the table has a value that uniquely denotes the row.
2. **Referential Integrity Constraint**: Referential integrity ensures the existence of a referenced value if a value of one column of a table references a value of another column. The existence of the foreign Key (AUTHOR_ID) in the BookShop table satisfies this integrity constraint because a cross-reference relationship between the BookShop and BookShop_AuthorDetails tables exists. As a result of this relationship, each book in the BookShop table is linked to the relevant row in the BookShop_AuthorDetails table through the AUTHOR_ID columns.
3. **Domain Integrity Constraint**: Domain integrity ensures clarity of column purpose and consistency of valid values. The BookShop table adheres to this constraint through the specification of data types, length, date format, check constraints, and null constraints in its CREATE statement. This comprehensive approach guarantees that the values in each column are not only valid but also conform to the specified domain constraints.