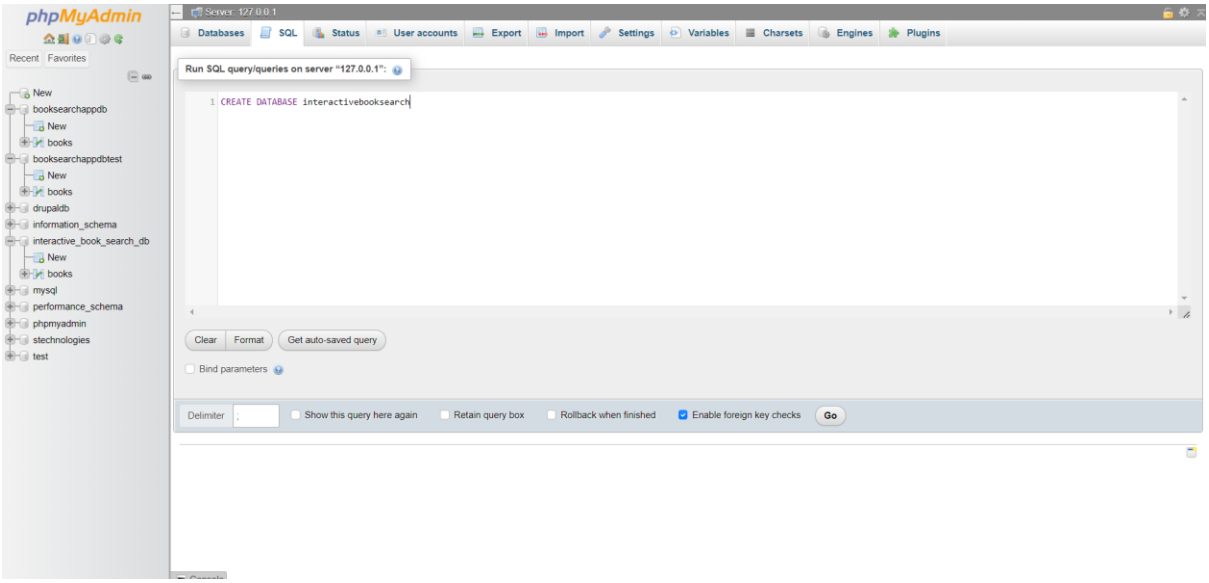
Part-1 (Database Setup)

1. Created the database using MySQL on my local database server (phpMyAdmin).

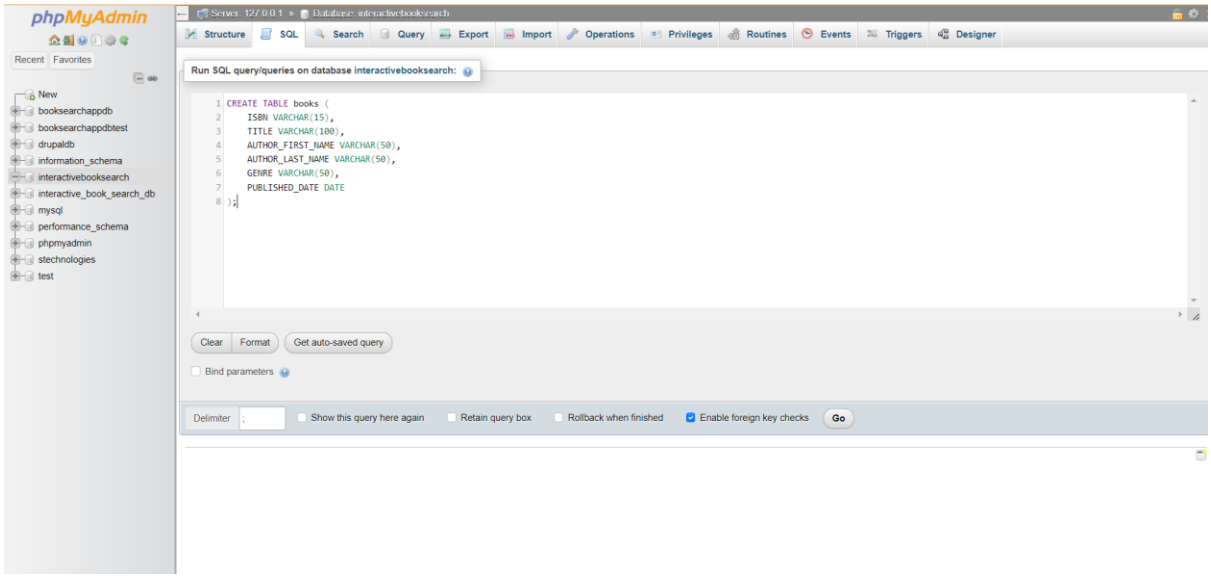
   Query used to create the database:

   “CREATE DATABASE interactivebooksearch”

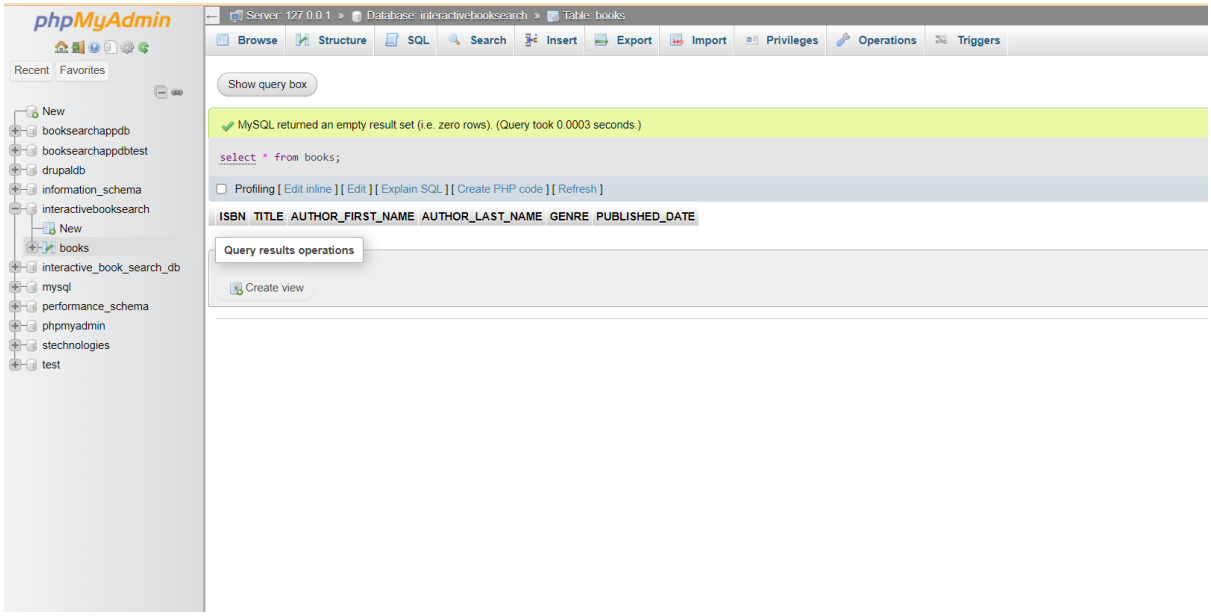


2. Created a table “books” in the same database using the below schema.

```
CREATE TABLE books (
    ISBN VARCHAR(15),
    TITLE VARCHAR(100),
    AUTHOR_FIRST_NAME VARCHAR(50),
    AUTHOR_LAST_NAME VARCHAR(50),
    GENRE VARCHAR(50),
    PUBLISHED_DATE DATE
);
```



3. After creating the table “books”, I ran the query “select * from books” and currently this table has no data.



Part-2 (Data Injection and SQL Queries)

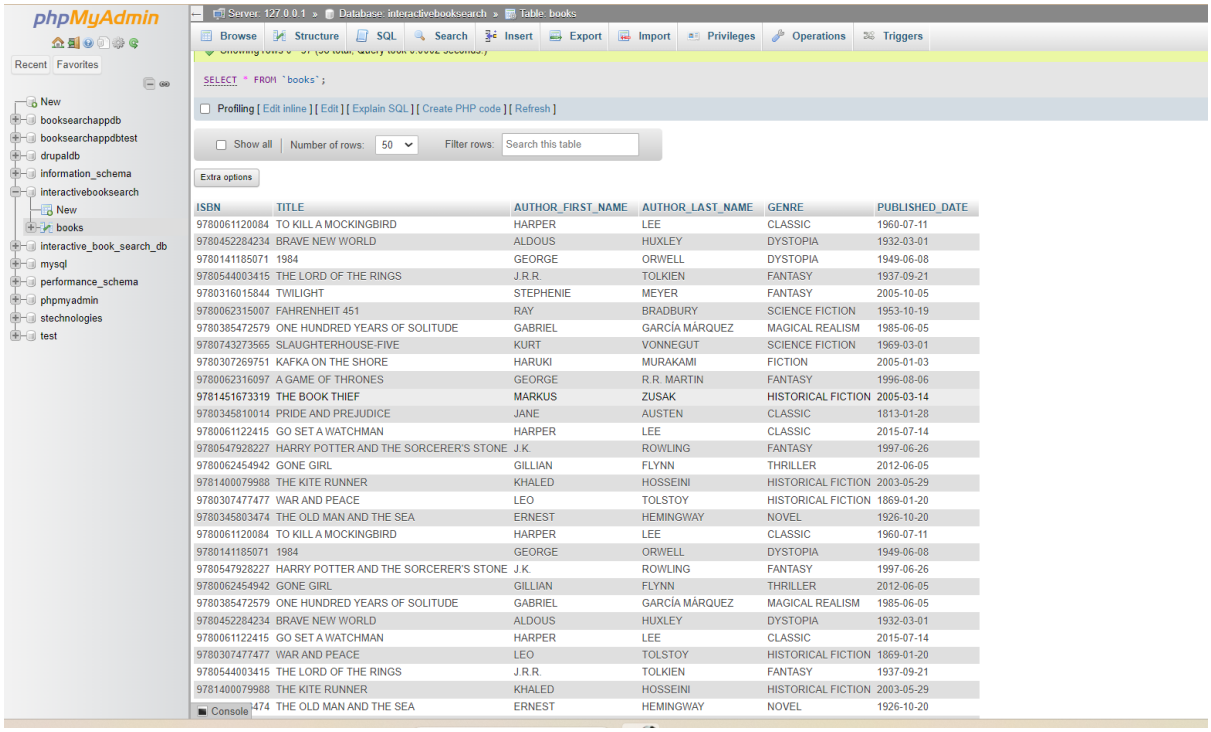
1. Wrote the below python script to import the data from the csv file to the “books” table on the local server. The name of my local script is “script_new.py” and I installed python to run this script.

Note: If you want to run the below script in python you have to change the csv_file_path “H:\\books.csv” according to your system. Also update “user” and “password” according to your localhost.

```
import pandas as pd
import mysql.connector

# MySQL database connection settings
db_config = {
    "host": "localhost",
    "user": "root",
    "password": "",
    "database": "interactivebooksearch"
}

# CSV file path
csv_file_path = "H:\\books.csv"

# Read CSV data using pandas
csv_data = pd.read_csv(csv_file_path)

# Establish MySQL database connection
connection = mysql.connector.connect(**db_config)
cursor = connection.cursor()

# Iterate through the CSV data and insert into the 'books' table
for index, row in csv_data.iterrows():
    query = "INSERT INTO books (ISBN, TITLE, AUTHOR_FIRST_NAME, AUTHOR_LAST_NAME, GENRE,
PUBLISHED_DATE) VALUES (%s, %s, %s, %s, %s, %s)"
    values = (row['ISBN'], row['TITLE'], row['AUTHOR FIRST NAME'], row['AUTHOR LAST NAME'],
row['GENRE'], row['PUBLISHED DATE'])

    cursor.execute(query, values)
    connection.commit()

# Close the database connection
cursor.close()
connection.close()

print("CSV data imported into 'books' table successfully!")
```

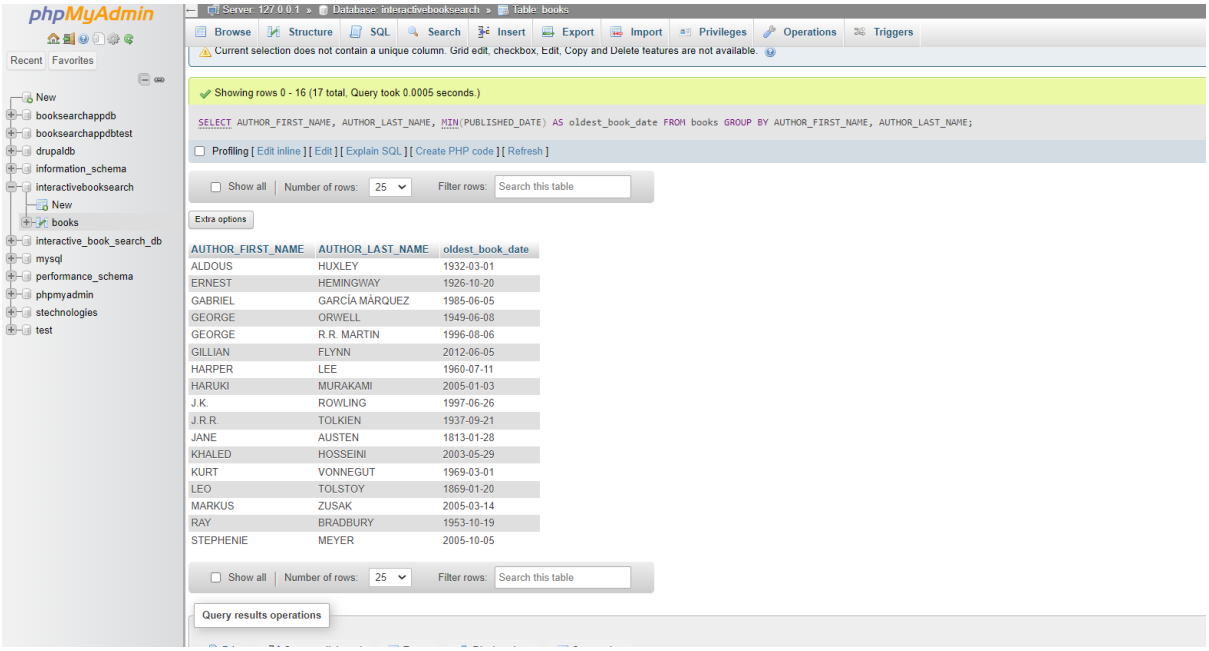
```
C:\Windows\System32\cmd.e

  File "H:\pythonscripttoinsertdata.py", line 5, in <module>
    db_connection = mysql.connector.connect(
                    ^^^^^^^^^^^^^^^^^^^^^^^^
  File "C:\Users\tajin\AppData\Local\Packages\PythonSoftwareFoundation.Python.3.11_qbz5n2kfra8p0\LocalCache\local-packag
es\Python311\site-packages\mysql\connector\__init__.py", line 179, in connect
    return MySQLConnection(*args, **kwargs)
           ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
  File "C:\Users\tajin\AppData\Local\Packages\PythonSoftwareFoundation.Python.3.11_qbz5n2kfra8p0\LocalCache\local-packag
es\Python311\site-packages\mysql\connector\connection.py", line 95, in __init__
    self.connect(**kwargs)
  File "C:\Users\tajin\AppData\Local\Packages\PythonSoftwareFoundation.Python.3.11_qbz5n2kfra8p0\LocalCache\local-packag
es\Python311\site-packages\mysql\connector\abstracts.py", line 713, in connect
    self.config(**kwargs)
  File "C:\Users\tajin\AppData\Local\Packages\PythonSoftwareFoundation.Python.3.11_qbz5n2kfra8p0\LocalCache\local-packag
es\Python311\site-packages\mysql\connector\abstracts.py", line 303, in config
    raise AttributeError("Unsupported argument '{0}'".format(key))
AttributeError: Unsupported argument 'ssl_mode'

H:\>mysqlimport --local --fields-terminated-by=',' --lines-terminated-by='\n' --ignore-lines=1 -u 6gw6r5chmdil69ouir5n -
p booksearchappdb H:\\books.csv
'mysqlimport' is not recognized as an internal or external command,
operable program or batch file.

H:\>python script_new.py
CSV data imported into 'books' table successfully!

H:\>python script_new.py
CSV data imported into 'books' table successfully!

H:\>
```

2. After running the python script data from the local file is now being inserted into the “books” table. See the output below.



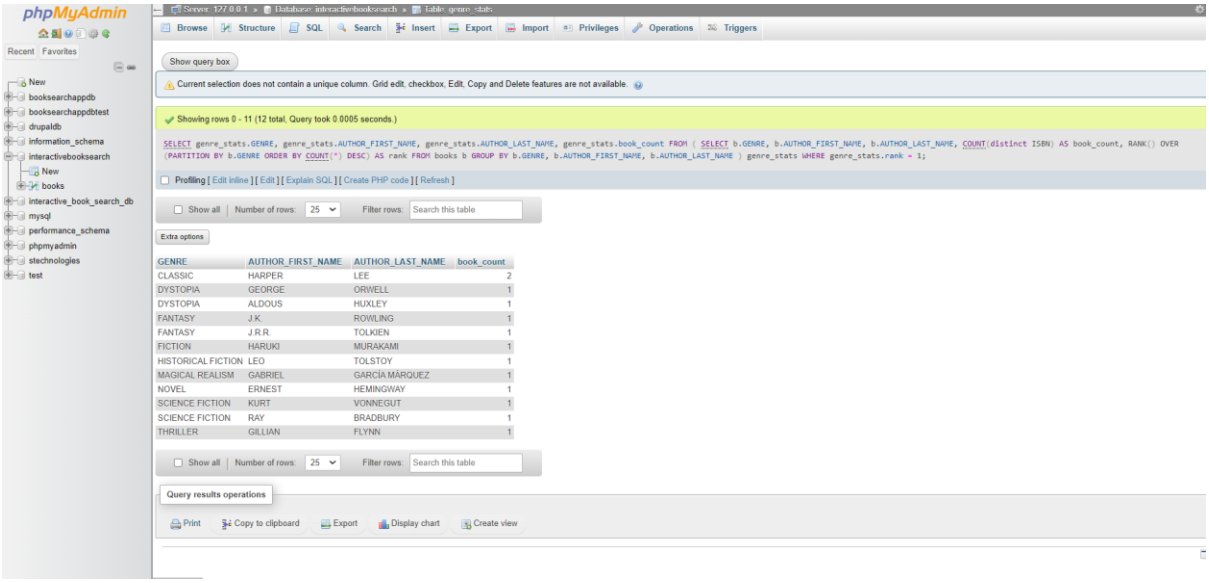
3. Query 1 (Find the oldest book for each author) – see output in the screenshot below.

```
SELECT AUTHOR_FIRST_NAME, AUTHOR_LAST_NAME, MIN(PUBLISHED_DATE) AS oldest_book_date
FROM books GROUP BY AUTHOR_FIRST_NAME, AUTHOR_LAST_NAME;
```



4. Query 2 (Find the authors who have written the most books in each genre.) – see output in the screenshot below.

```
SELECT genre_stats.GENRE, genre_stats.AUTHOR_FIRST_NAME,
genre_stats.AUTHOR_LAST_NAME, genre_stats.book_count
FROM (
    SELECT b.GENRE, b.AUTHOR_FIRST_NAME, b.AUTHOR_LAST_NAME, COUNT(distinct ISBN)
AS book_count,
        RANK() OVER (PARTITION BY b.GENRE ORDER BY COUNT(*) DESC) AS rank
    FROM books b
    GROUP BY b.GENRE, b.AUTHOR_FIRST_NAME, b.AUTHOR_LAST_NAME
) genre_stats WHERE genre_stats.rank = 1;
```



5. Query 3 (Calculate the total number of books published each year and display the results by year) – see output in the screenshot below.

   SELECT YEAR(PUBLISHED_DATE) AS publication_year, COUNT(distinct ISBN) AS book_count
   FROM books
   GROUP BY YEAR(PUBLISHED_DATE)
   ORDER BY publication_year;