

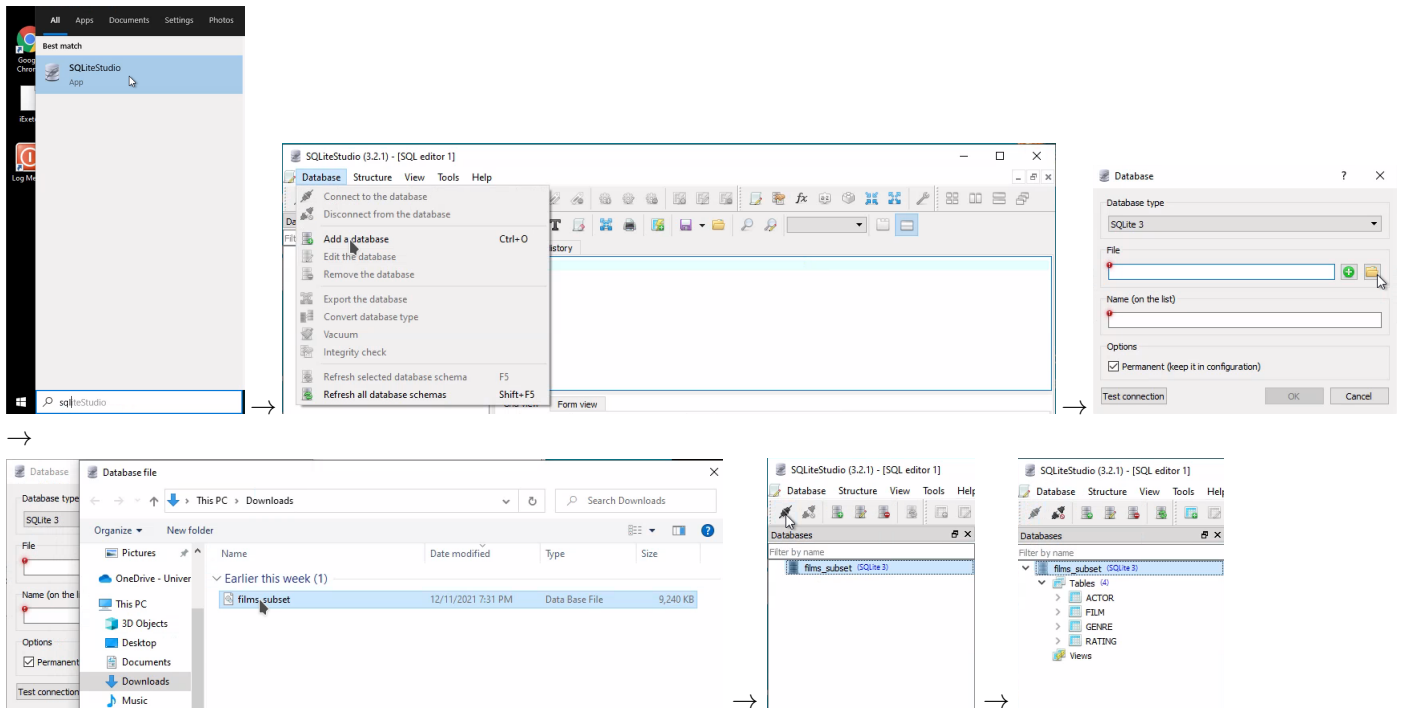
Database Technologies for Business Analytics

BEM2040

Practice – Week 2

The following instructions can be followed by using university computers, a **virtual desktop** or your personal computer, if the software has been installed.

1. Download the file **Week2.zip**. We will be using:
 - films_subset.db
 - practice_w2_start.ipynb
2. Decompress the files in your downloads folder.
3. Start SQLiteStudio and add the database:



4. On the left-hand side (highlighted in blue in the picture below) you will see the tables (groups of data) that exist in the database: *ACTOR*, *FILM*, *GENRE* and *RATING*.

Each table has several columns, as follows:

ACTOR:

actor_id,
actor_name,
actor_year_born,
actor_year_dead

FILM:

film_id,
film_title,
film_year_start,
film_year_end,
film_major_genre

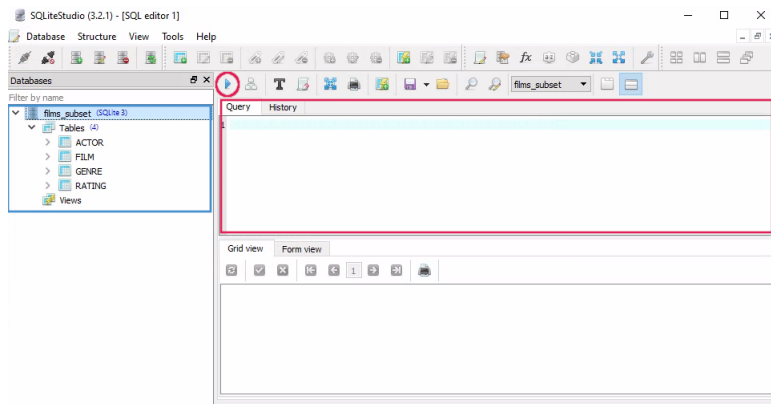
GENRE:

genre_id,
genre_name

RATING:

rating_film_id,
rating_average,
rating_num_votes

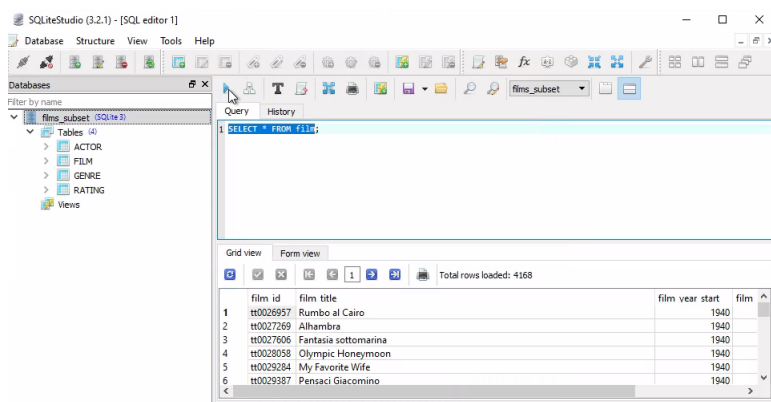
We will be using queries to operate on some of them.



5. The following query helps us to retrieve all data in the table called *film*:

```
select * from film;
```

We input it in the *query* tab (highlighted in red in the previous picture). After executing the query with the play button (also highlighted in red), we see the result in the section right below the query.



6. Let's now obtain films for a specific year (1940). The instruction is:

```
select * from film where film_year_start=1940;
```

7. Now, let's look at table *actor*. The instruction

```
select * from actor;
```

gives an empty output, as there are no rows.

8. Let's add some rows:

```
insert into actor(actor_id, actor_name, actor_year_born, actor_year_dead)
values ("a1", "Timothée Chalamet", 1994, null);

insert into actor(actor_id,actor_name,actor_year_born,actor_year_dead)
values ("a2", "Rebecca Ferguson", 1983, null);
```

9. Check the table *actor* after the inserts, using once more the instruction:

```
select * from actor;
```

10. We can change data in one row, or several. For now, let's change the year Timothée Chalamet was born from 1994 to 1995, based on the id for that row:

```
update actor set actor_year_born = 1995 where actor_id = "a1";
```

11. And again let's check the content:

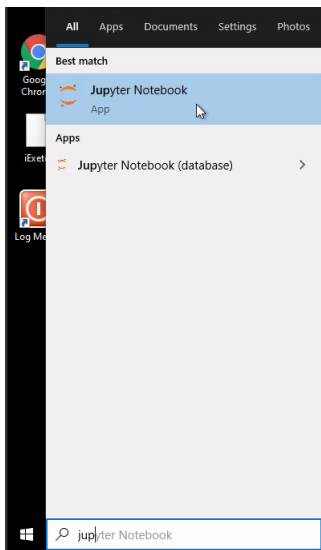
```
select * from actor;
```

The year in which actor Timothée Chalamet was born is now updated.

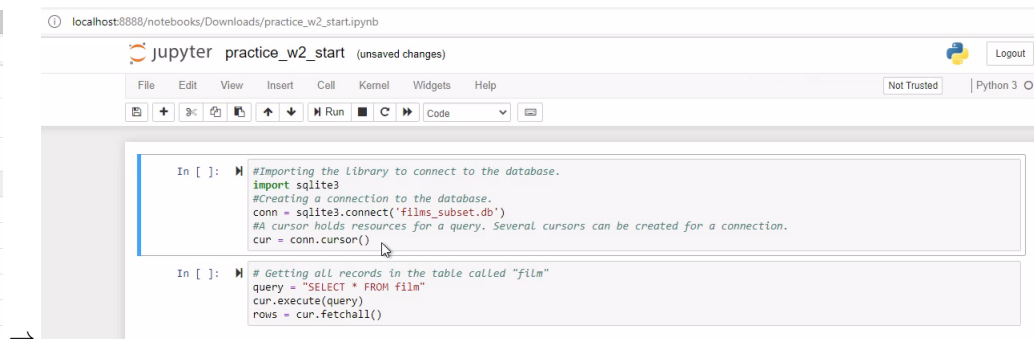
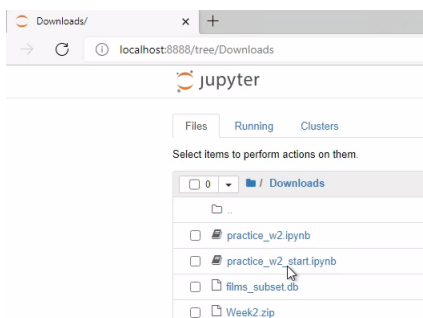
12. Rows in a table can be counted in the following way:

```
select count(*) from actor;
```

13. Start Jupyter (it will open in a web browser) and browse to the download folder.



14. Open the Jupyter notebook (practice_w2_start.ipynb).



15. Add or modify instructions to the notebook, according to guidance in the practice. Following is the code we are going to use.

```
#Importing the library to connect to the database.
import sqlite3
#Creating a connection to the database.
conn = sqlite3.connect('films_subset.db')
#A cursor holds resources for a query.
#Several cursors can be created for a connection.
cur = conn.cursor()

# Getting all records in the table called "film"
query = "SELECT * FROM film"
cur.execute(query)
rows = cur.fetchall()

for row in rows:
    print(row)

# Getting the films from 1940
query = "SELECT * FROM film where film_year_start=1940"
cur.execute(query)
rows = cur.fetchall()

for row in rows:
    print(row)

#Getting all the records in the table called "actor"
query = "SELECT * FROM actor"
cur.execute(query)
rows = cur.fetchall()

for row in rows:
    print(row)

#Adding some records to the table actor
query = "insert into actor( " + \
        "actor_id," + \
        "actor_name," + \
        "actor_year_born," + \
        "actor_year_dead)" + \
        "values (" + \
        "?," + \
        "?," + \
        "?," + \
        "?)"

cur.execute(query, ("a1", "Timothée Chalamet", 1994, None))

cur.execute(query, ("a2", "Rebecca Ferguson", 1983, None))

conn.commit()

#Listing all records after having added two.
query = "SELECT * FROM actor"
cur.execute(query)
rows = cur.fetchall()
for row in rows:
    print(row)
```

```
# Updating the birth year of actor with code "a1", that is, Timothée Chalamet
query = "update actor set actor_year_born = ? where actor_id = ?"
cur.execute(query, (1995, "a1"))
conn.commit()

#Listing actors after the update.
query = "SELECT * FROM actor"
cur.execute(query)
rows = cur.fetchall()
for row in rows:
    print(row)

#Counting the number of records in table actor.
query = "SELECT count(*) FROM actor"
cur.execute(query)
rows = cur.fetchall()
print(rows)
print(rows[0][0])

#Closing the connection
conn.close()
```
