Database Technologies for Business Analytics
BEM2040
Practice – Week 3

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 Week3.zip. We will be using:
   - films_subset.db
   - practice_w3_start.ipynb.

   The complete file with all the instructions (practice_w3.ipynb) can be used to practice at home.

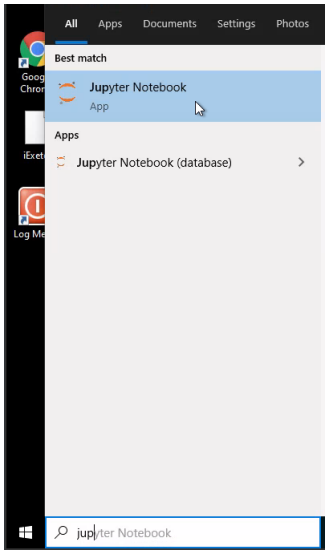
2. Decompress the files in your downloads folder.

   The database will contain several tables, as follows:

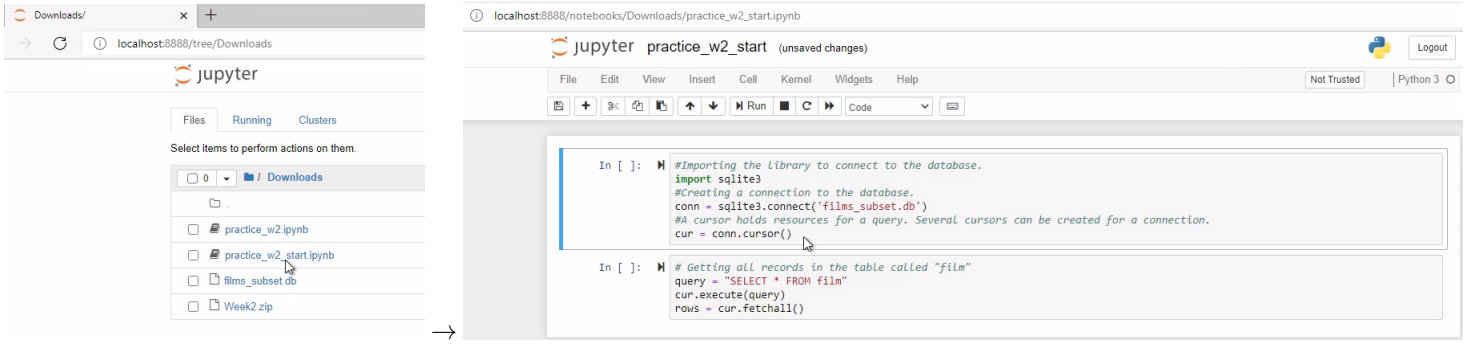
```
ACTOR:
 actor_id,
 actor_name,
 actor_year_born,
 actor_year_dead

FILM:
 film_id,
 film_title,
 film_year_start,
 film_year_end,
 genre_id

GENRE:
 genre_id,
 genre_name

RATING:
 rating_film_id,
 rating_average,
 rating_num_votes
```

   Python will be used to obtain data from some of those tables and to enter new data.

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



4. Open the Jupyter notebook (practice_w3_start.ipynb). To practice at home you can also use the complete notebook (practice_w3.ipynb)



5. 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()
```
