



Welcome! In order to provide you with the best experience we ask that you complete this tutorial before attending the workshop.

Questions: In case you are having any difficulties in following the tutorial please ask directly to our guys at [NDS \(https://www.facebook.com/NDS-Núcleo-de-Data-Science-FCT-UNL-769904580035480/\)](https://www.facebook.com/NDS-Núcleo-de-Data-Science-FCT-UNL-769904580035480/) through the messsenger chat.

Tutorial

Time to interact with a real database! For this course we will connect to an SQLite database, which is just a file! You can read more about that [here \(https://www.sqlite.org/onefile.html\)](https://www.sqlite.org/onefile.html). For this, we'll need an SQL client, which is a graphical interface to interact with the database. So let's install a SQL client and get started!

Here we're going to use the [DBeaver SQL client \(https://dbeaver.io/\)](https://dbeaver.io/) since it can work with different DBMSs and can be installed in Linux, Mac and Windows machines.

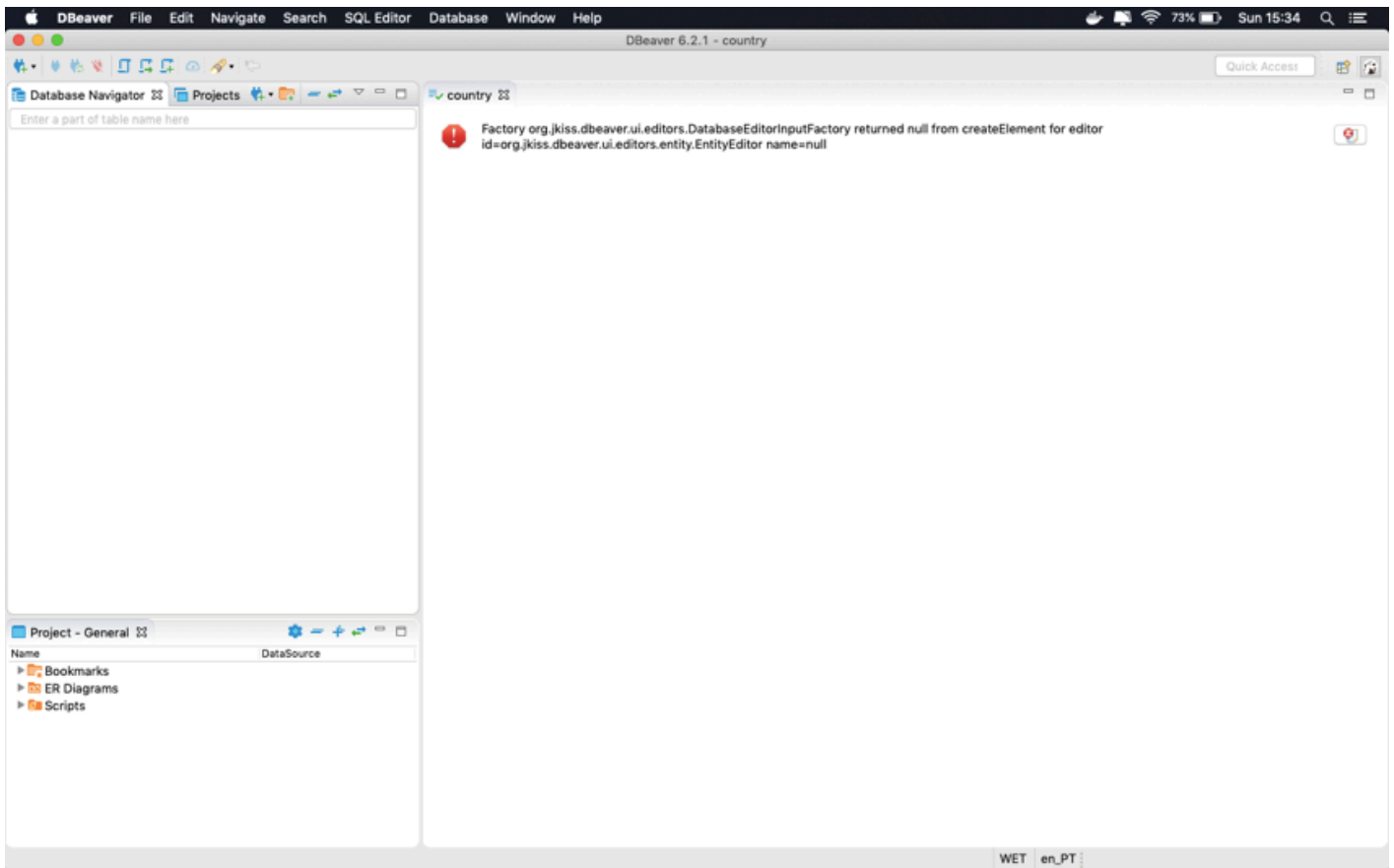
Step 1: Download and Install DBeaver client

[Download Link for DBeaver \(https://dbeaver.io/download/\)](https://dbeaver.io/download/)

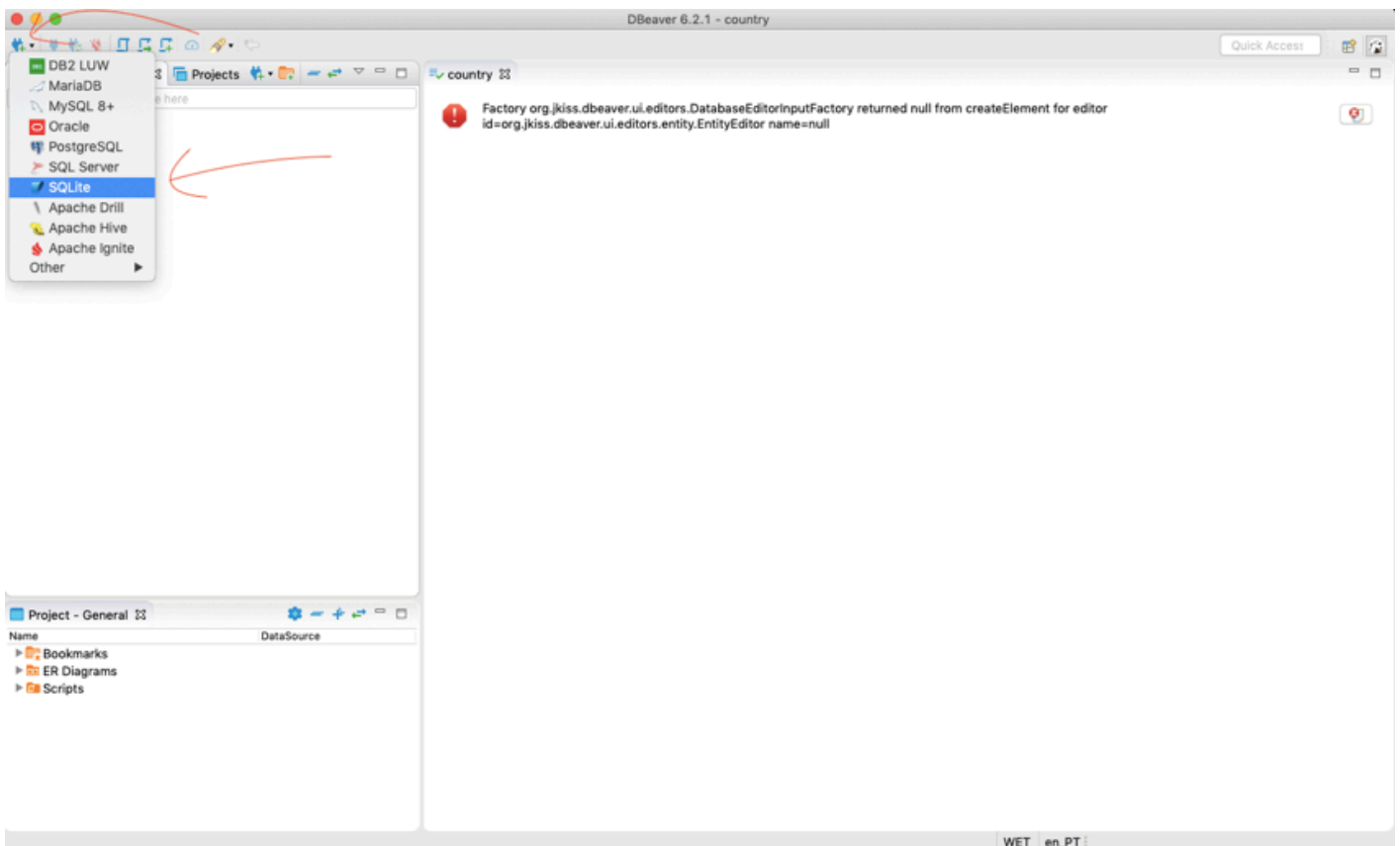
Step 2: Connect to local database

In case you did not notice, we actually have a database in this repository! If you look up in the root of the project for the folder "data" there's a file there named "WorkshopDB.db". Let's connect to it!

If you open your application DBeaver, this should be similar to the main screen you are seeing on your computer:



Now we want to create a new (SQLite) connection to our database!



This window should have popped up:

Connect to database

Generic JDBC Connection Settings

SQLite connection settings

General Driver properties SSH Proxy

JDBC URL:

Path:

Advanced settings:

Driver name:

< Back Next > Cancel Test Connection ... Finish

DBeaver is asking you to provide the full path on your computer of the database file. You should now press "browse" and find the the_movies.db file which is inside the workshop folder under "data": (you can see where mine is stored below, but your path is probably slightly different)

Connect to database

Generic JDBC Connection Settings

SQLite connection settings

General Driver properties SSH Proxy

JDBC URL: jdbc:sqlite:/Users/simaonovais/Documents/GitHub/NIEGI_NDS/data/the_movies.db

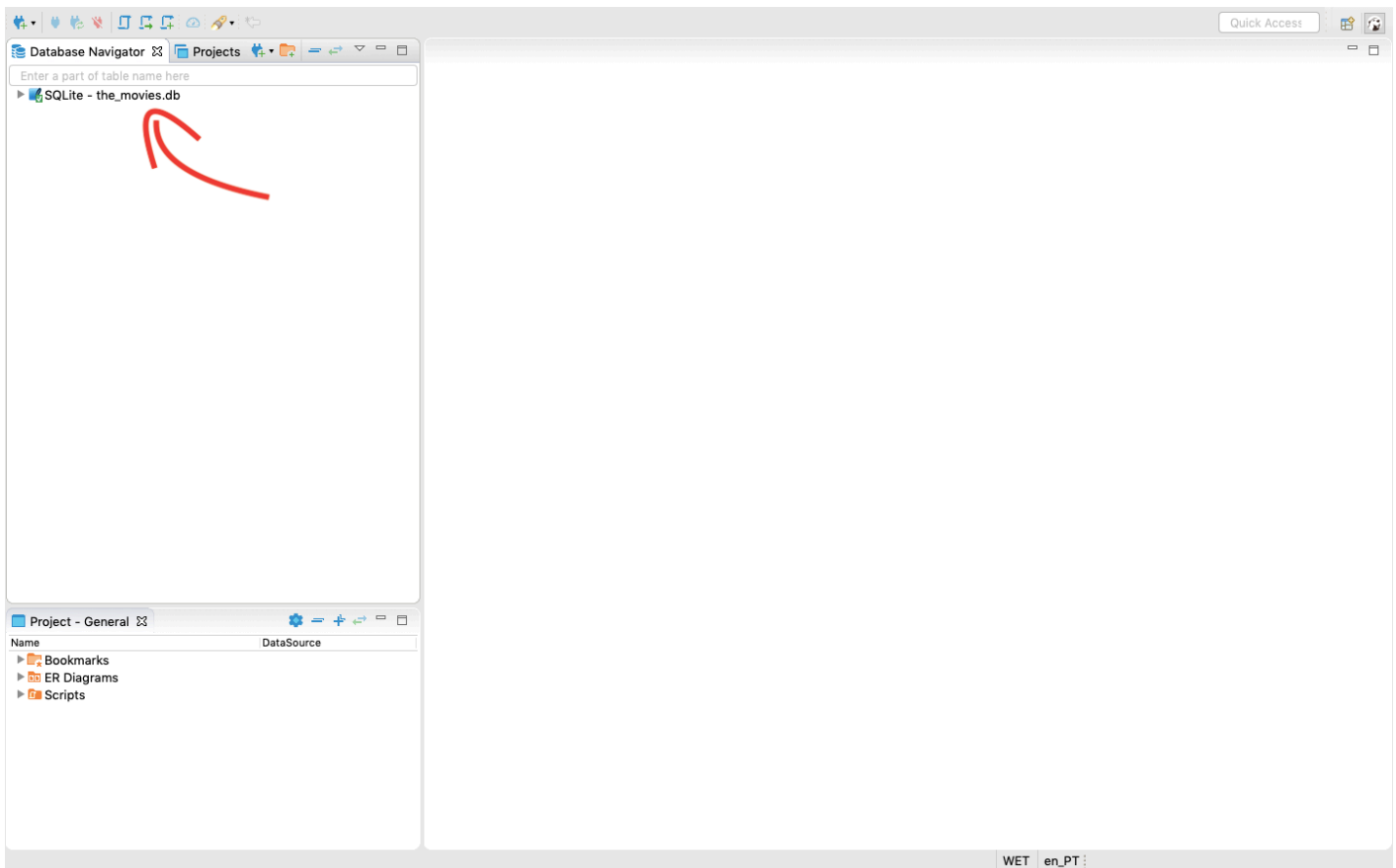
Path: /Users/simaonovais/Documents/GitHub/NIEGI_NDS/data/the_movies.db Browse Create

Advanced settings: Connection details (name, type, ...)

Driver name: SQLite Edit Driver Settings

< Back Next > Cancel Test Connection ... Finish

You can now press "Finish" and this should be what you get:



Step 3: Investigate the DB

You can now start looking at what's inside our database! We will let you explore this by yourself but here's something that you can do:

The screenshot shows the DB Browser for SQLite interface. On the left, the 'Database Navigator' pane shows the 'SQLite - the_movies.db' database with a tree view of tables. The 'movie' table is selected. The main window displays the 'Data' tab for the 'movie' table, showing a grid of 34 rows. The columns are: id, imdb_id, original_title, release_date, and budget. The status bar at the bottom indicates '200 row(s) fetched - 10ms (+17ms)'.

	id	imdb_id	original_title	release_date	budget
1	1	tt0114709	Toy Story	1995-10-30	30,000,000
2	2	tt0113497	Jumanji	1995-12-15	65,000,000
3	3	tt0113228	Grumpier Old Men	1995-12-22	0
4	4	tt0114885	Waiting to Exhale	1995-12-22	16,000,000
5	5	tt0113041	Father of the Bride Part II	1995-02-10	0
6	6	tt0113277	Heat	1995-12-15	60,000,000
7	7	tt0114319	Sabrina	1995-12-15	58,000,000
8	8	tt0112302	Tom and Huck	1995-12-22	0
9	9	tt0114576	Sudden Death	1995-12-22	35,000,000
10	10	tt0113189	GoldenEye	1995-11-16	58,000,000
11	11	tt0112346	The American President	1995-11-17	62,000,000
12	12	tt0112896	Dracula: Dead and Loving It	1995-12-22	0
13	13	tt0112453	Balto	1995-12-22	0
14	14	tt0113987	Nixon	1995-12-22	44,000,000
15	15	tt0112760	Cutthroat Island	1995-12-22	98,000,000
16	16	tt0112641	Casino	1995-11-22	52,000,000
17	17	tt0114388	Sense and Sensibility	1995-12-13	16,500,000
18	18	tt0113101	Four Rooms	1995-12-09	4,000,000
19	19	tt0112281	Ace Ventura: When Nature Calls	1995-11-10	30,000,000
20	20	tt0113845	Money Train	1995-11-21	60,000,000
21	21	tt0113161	Get Shorty	1995-10-20	30,250,000
22	22	tt0112722	Copycat	1995-10-27	0
23	23	tt0112401	Assassins	1995-10-06	50,000,000
24	24	tt0114168	Powder	1995-10-27	0
25	25	tt0113627	Leaving Las Vegas	1995-10-27	3,600,000
26	26	tt0114057	Othello	1995-12-15	0
27	27	tt0114011	Now and Then	1995-10-20	12,000,000
28	28	tt0114117	Persuasion	1995-09-27	0
29	29	tt0112682	La Cité des Enfants Perdus	1995-05-16	18,000,000
30	30	tt0115012	摇啊摇，摇到外婆桥	1995-04-30	0
31	31	tt0112792	Dangerous Minds	1995-08-11	0
32	32	tt0114746	Twelve Monkeys	1995-12-29	29,500,000
33	33	tt0114952	Guillaumet, les ailes du courage	1996-09-18	0
34	34	tt0112431	Babe	1995-07-18	30,000,000

Awesome! So you can actually see what types of columns each table has and even look directly at the data!

Step 4: Test a query

Select "SQL Editor" in the tool-bar and then click on "New SQL editor".

The screenshot shows the DBeaver 6.2.1 interface. The SQL Editor is active, and a dropdown menu is open, showing options like 'New SQL Editor', 'Toggle results panel', 'Maximize results panel', 'Switch active panel', 'Set active connection', 'Select active schema', 'Set connection from navigator', and 'Auto-sync connection with navigator'. The background displays a table of movie data with columns: id, imdb_id, original_title, release_date, budget, and value.

id	imdb_id	original_title	release_date	budget	value
7	tt0114319	Toy Story	1995-10-30	30,000,000	1
8	tt0112302	Jumanji	1995-12-15	65,000,000	
9	tt0114576	Grumpier Old Men	1995-12-22	0	
10	tt0113189	Waiting to Exhale	1995-12-22	16,000,000	
11	tt0112346	Father of the Bride Part II	1995-02-10	0	
12	tt0112896	Heat	1995-12-15	60,000,000	
13	tt0112453	Sabrina	1995-12-15	58,000,000	
14	tt0113987	Tom and Huck	1995-12-22	0	
15	tt0112760	Sudden Death	1995-12-22	35,000,000	
16	tt0112641	GoldenEye	1995-11-16	58,000,000	
17	tt0114388	The American President	1995-11-17	62,000,000	
18	tt0113101	Dracula: Dead and Loving It	1995-12-22	0	
19	tt0112281	Balto	1995-12-22	0	
20	tt0113845	Nixon	1995-12-22	44,000,000	
21	tt0113161	Cutthroat Island	1995-12-22	98,000,000	
22	tt0112722	Casino	1995-11-22	52,000,000	
23	tt0112401	Sense and Sensibility	1995-12-13	16,500,000	
24	tt0114168	Four Rooms	1995-12-09	4,000,000	
25	tt0113627	Ace Ventura: When Nature Calls	1995-11-10	30,000,000	
26	tt0114057	Money Train	1995-11-21	60,000,000	
27	tt0114011	Get Shorty	1995-10-20	30,250,000	
28	tt0114117	Copycat	1995-10-27	0	
29	tt0112682	Assassins	1995-10-06	50,000,000	
30	tt0115012	Powder	1995-10-27	0	
31	tt0112792	Leaving Las Vegas	1995-10-27	3,600,000	
32	tt0114746	Othello	1995-12-15	0	
33	tt0114952	Now and Then	1995-10-20	12,000,000	
34	tt0112431	Persuasion	1995-09-27	0	
		La Cité des Enfants Perdus	1995-05-16	18,000,000	
		搖啊搖，搖到外婆橋	1995-04-30	0	
		Dangerous Minds	1995-08-11	0	
		Twelve Monkeys	1995-12-29	29,500,000	
		Guillaumet, les ailes du courage	1996-09-18	0	
		Babe	1995-07-18	30,000,000	

Type "SELECT * FROM movie where original_title like 'Toy Story';" and check that you get the results (press CTRL+Enter or click the orange arrow to execute SQL statements)

The screenshot shows the DBeaver 6.2.1 interface with the SQL Editor. The query entered is: `select * from movie where original_title like "Toy Story"`. The results panel shows one row of data for the movie 'Toy Story'.

id	imdb_id	original_title	release_date	budget	value
1	tt0114709	Toy Story	1995-10-30	30,000,000	1

Congratulations you have completed the DBeaver tutorial! You may now have your well deserved rest and patiently wait until Wednesday the 16th. See you there!

