

### 3.3: SQL for Data Analysts

Step 1) I have used **SELECT \*** **FROM** category

The screenshot shows the pgAdmin 4 web interface. On the left, the 'Servers' tree is expanded to show the 'public' schema, with 'Tables (15)' selected. The 'Query Editor' tab is active, displaying the SQL query: `1 SELECT *` and `2 FROM category`. Below the query editor, the 'Data Output' tab shows the results of the query in a table format. The table has three columns: 'category\_id' (integer), 'name' (character varying (25)), and 'last\_update' (timestamp without time zone). The results list 16 categories, each with a unique ID and a timestamp of 2006-02-15 09:46:27.

category_id	name	last_update
1	Action	2006-02-15 09:46:27
2	Animation	2006-02-15 09:46:27
3	Children	2006-02-15 09:46:27
4	Classics	2006-02-15 09:46:27
5	Comedy	2006-02-15 09:46:27
6	Documentary	2006-02-15 09:46:27
7	Drama	2006-02-15 09:46:27
8	Family	2006-02-15 09:46:27
9	Foreign	2006-02-15 09:46:27
10	Games	2006-02-15 09:46:27
11	Horror	2006-02-15 09:46:27
12	Music	2006-02-15 09:46:27
13	New	2006-02-15 09:46:27
14	Sci-Fi	2006-02-15 09:46:27
15	Sports	2006-02-15 09:46:27
16	Travel	2006-02-15 09:46:27

**Step 2)** You're ready to add some new genres! Write an INSERT statement to add the following genres to the category table: Thriller, Crime, Mystery, Romance, and War:

- Copy-paste your INSERT commands into your answers document.

```
INSERT INTO category (name)
VALUES
('Thriller'),
('Crime'),
('Mystery'),
('Romance'),
('War');
```

- The CREATE statement below shows the constraints on the category table. Write a short paragraph explaining the various constraints that have been applied to the columns. What do these constraints do exactly? Why are they important?

```
CREATE TABLE category
(
  category_id integer NOT NULL DEFAULT nextval('category_category_id_seq'::regclass),
  name text COLLATE pg_catalog."default" NOT NULL,
  last_update timestamp with time zone NOT NULL DEFAULT now(),
  CONSTRAINT category_pkey PRIMARY KEY (category_id)
);
```

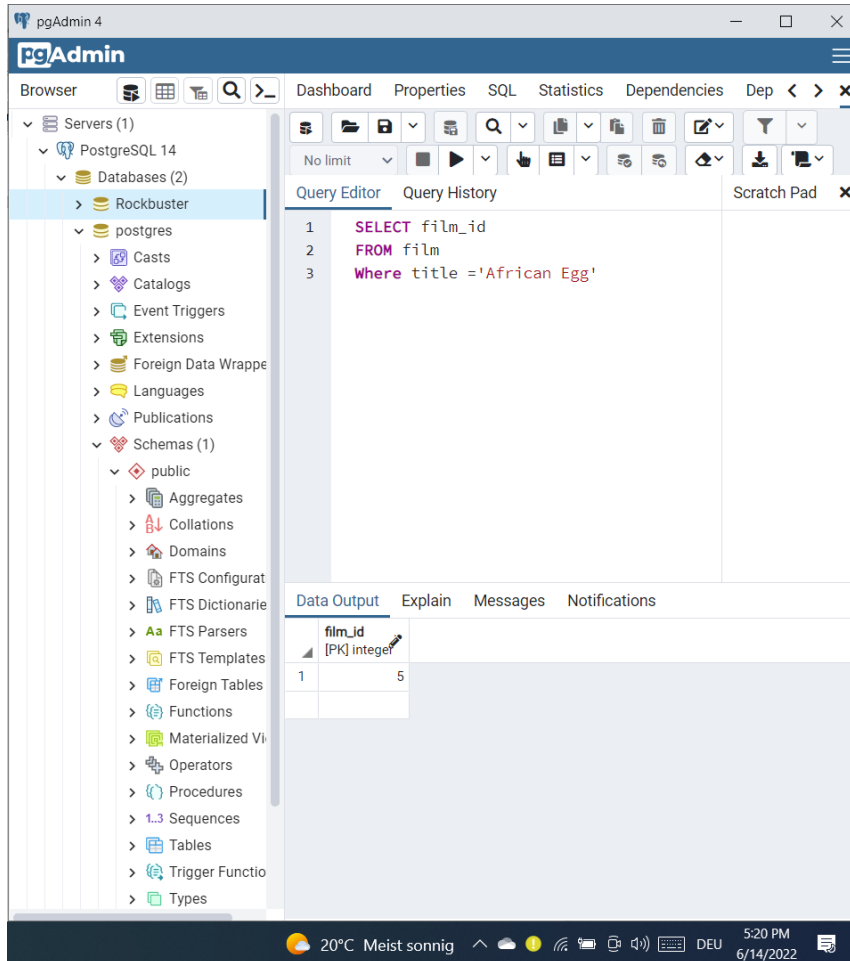
**NOT NULL Constraint** - This ensures that a column can't have any empty or missing values. Use NOT NULL if your table contains columns that should never be empty. When this constraint is applied correctly, an error message will appear if you try to insert empty values.

**PRIMARY KEY Constraint** - the primary key gives each record in a table a unique ID. The primary key column can't contain any null or duplicate values.

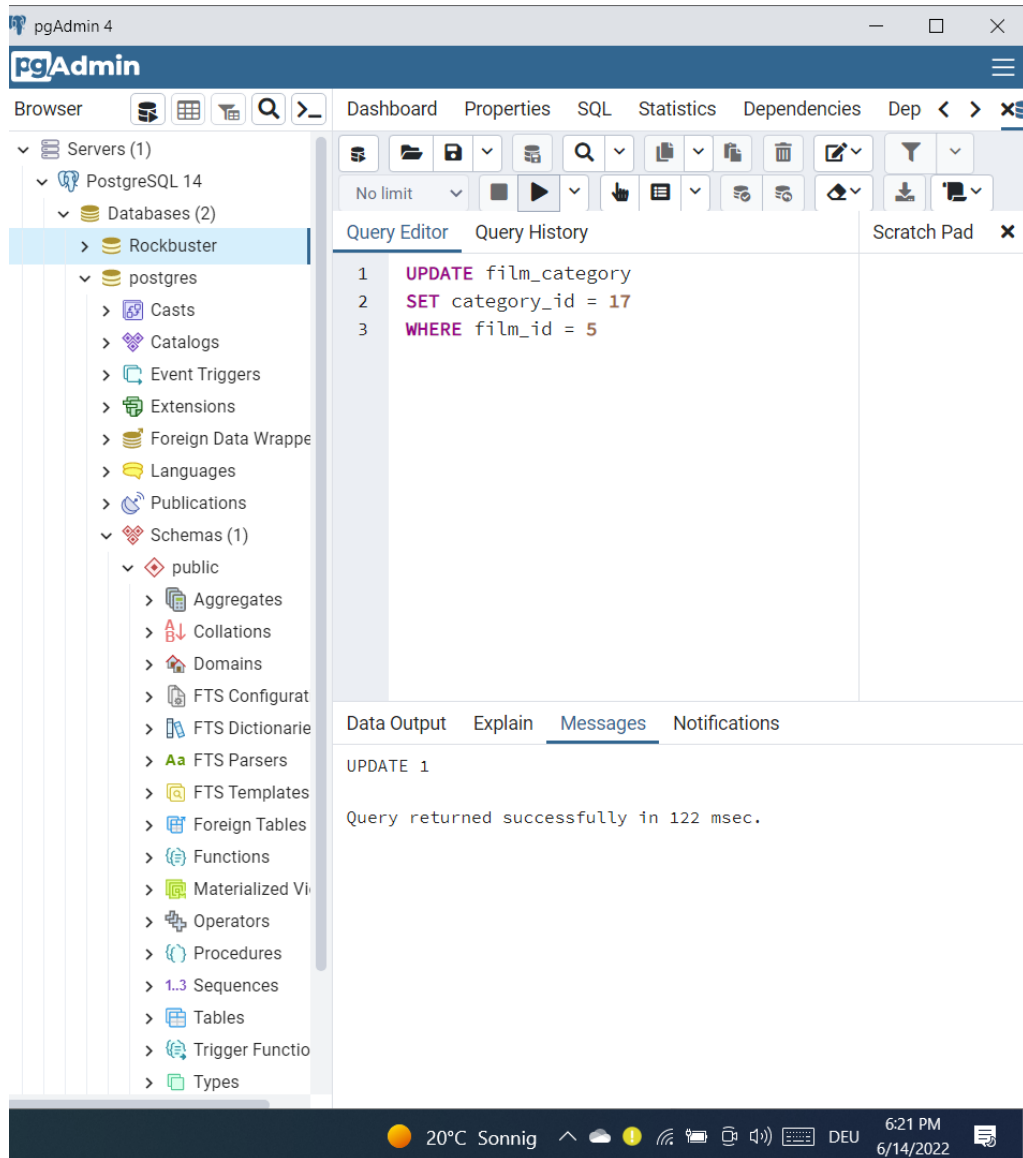
### Step 3)

The genre for the movie African Egg needs to be updated to thriller. Work through the steps below to make this change:

- Write the SELECT statement to find the film\_id for the movie African Egg.

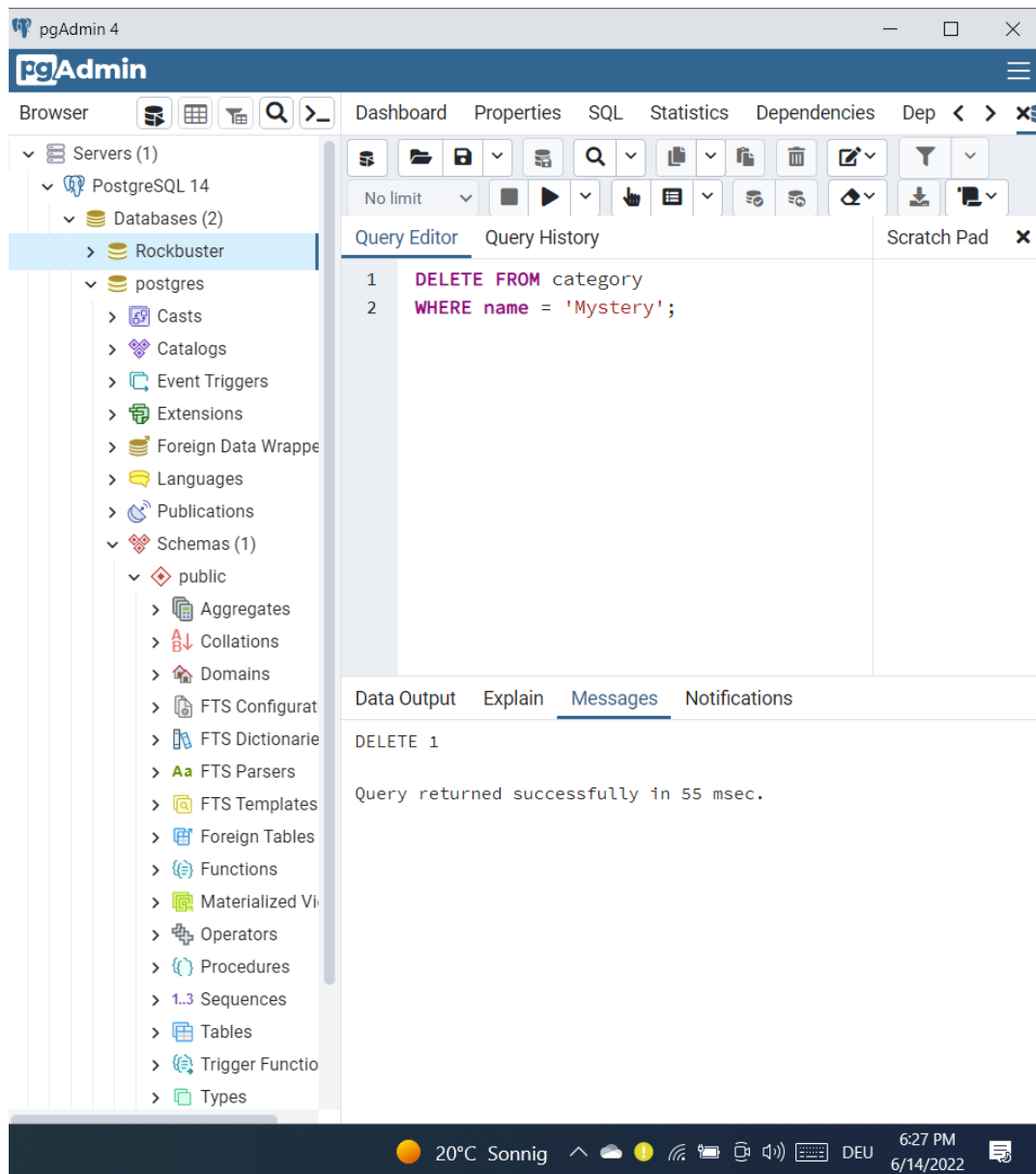


- Once you have the film\_ID and category\_ID, write an UPDATE command to change the category in the film\_category table (not the category table). Copy-paste this command into your answers document.



#### Step 4)

Since there aren't many movies in the mystery category, you and your manager decide to remove it from the category table. Write a DELETE command to do so and copy-paste it into your answers document.



## Step 5)

Based on what you've learned so far, think about what it would be like to complete steps 1 to 4 with Excel instead of SQL. Are there any pros and cons to using SQL? Write a paragraph explaining your answer.

I think that both programs have its own advantages. Excel would be more appropriate with working with smaller Databases, when SQL in other hands is more appropriate for larger Databases.