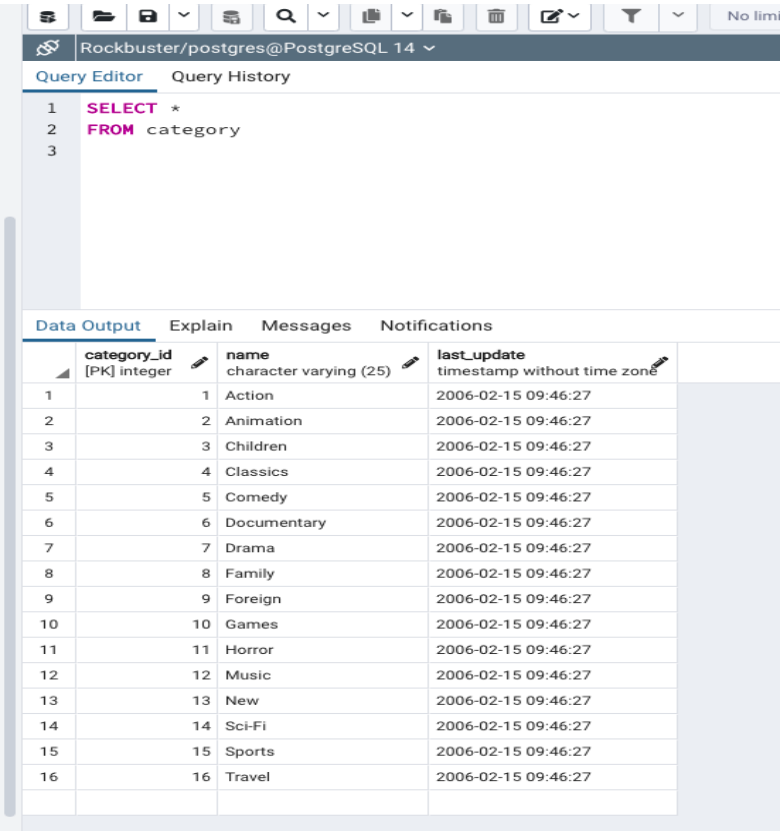
3.3 SQL For Data Analysts

Step 1

- Write a SELECT command to find out what film genres exist in the category table.
- Copy-paste the output into your answers document or write the answers out—it's up to you. Make sure to include the category ID for each genre.



	category_id [PK] integer	name character varying (25)	last_update timestamp without time zone
1	1	Action	2006-02-15 09:46:27
2	2	Animation	2006-02-15 09:46:27
3	3	Children	2006-02-15 09:46:27
4	4	Classics	2006-02-15 09:46:27
5	5	Comedy	2006-02-15 09:46:27
6	6	Documentary	2006-02-15 09:46:27
7	7	Drama	2006-02-15 09:46:27
8	8	Family	2006-02-15 09:46:27
9	9	Foreign	2006-02-15 09:46:27
10	10	Games	2006-02-15 09:46:27
11	11	Horror	2006-02-15 09:46:27
12	12	Music	2006-02-15 09:46:27
13	13	New	2006-02-15 09:46:27
14	14	Sci-Fi	2006-02-15 09:46:27
15	15	Sports	2006-02-15 09:46:27
16	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.

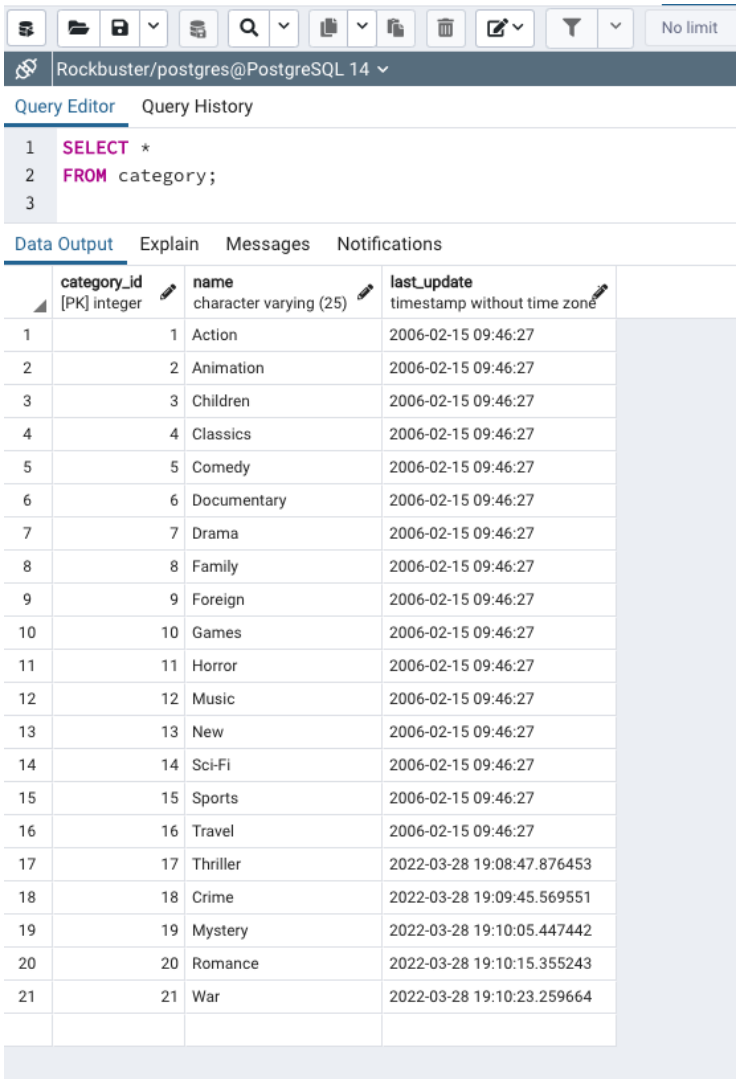
1. INSERT INTO category(name)
   VALUES('Thriller');

2. INSERT INTO category(name)
   VALUES('Crime');

3. INSERT INTO category(name)
   VALUES('Mystery');

4. INSERT INTO category(name)
   VALUES('Romance');

5. INSERT INTO category(name)
   VALUES('War');



	category_id [PK] integer	name character varying (25)	last_update timestamp without time zone
1	1	Action	2006-02-15 09:46:27
2	2	Animation	2006-02-15 09:46:27
3	3	Children	2006-02-15 09:46:27
4	4	Classics	2006-02-15 09:46:27
5	5	Comedy	2006-02-15 09:46:27
6	6	Documentary	2006-02-15 09:46:27
7	7	Drama	2006-02-15 09:46:27
8	8	Family	2006-02-15 09:46:27
9	9	Foreign	2006-02-15 09:46:27
10	10	Games	2006-02-15 09:46:27
11	11	Horror	2006-02-15 09:46:27
12	12	Music	2006-02-15 09:46:27
13	13	New	2006-02-15 09:46:27
14	14	Sci-Fi	2006-02-15 09:46:27
15	15	Sports	2006-02-15 09:46:27
16	16	Travel	2006-02-15 09:46:27
17	17	Thriller	2022-03-28 19:08:47.876453
18	18	Crime	2022-03-28 19:09:45.569551
19	19	Mystery	2022-03-28 19:10:05.447442
20	20	Romance	2022-03-28 19:10:15.355243
21	21	War	2022-03-28 19:10:23.259664

- 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)
);
```

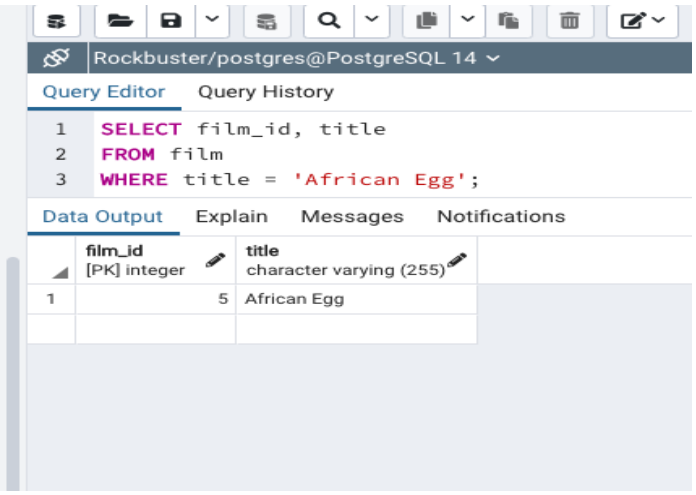
The NOT NULL constraint ensures that the column cannot have empty value (missing or blank). This has been assigned to the three columns in the table. Essentially, user must enter a value for the column else an error would be returned.

The PRIMARY KEY constraint assigned to the category_id column ensures that the value in this column must be unique for each row, can neither contain null value nor be modified if it is linked to another table.

Step 3

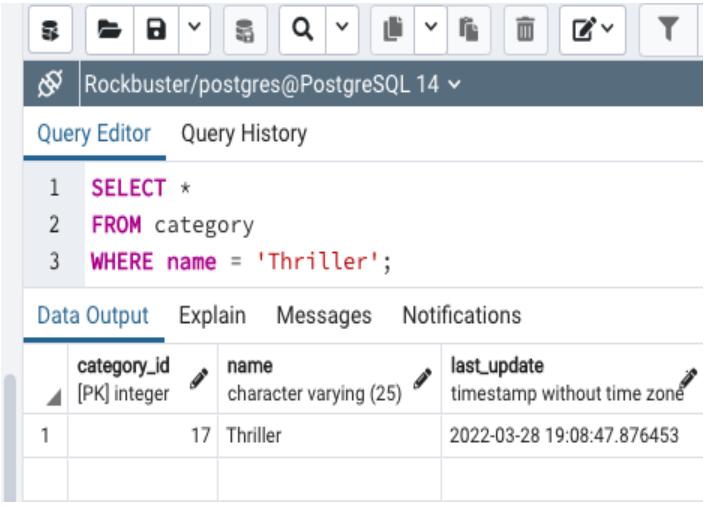
The genre for the movie *African Egg* needs to be updated to thriller. Work through the steps below to make this change:

- Write the SELECT statement to find the film_id for the movie *African Egg*.

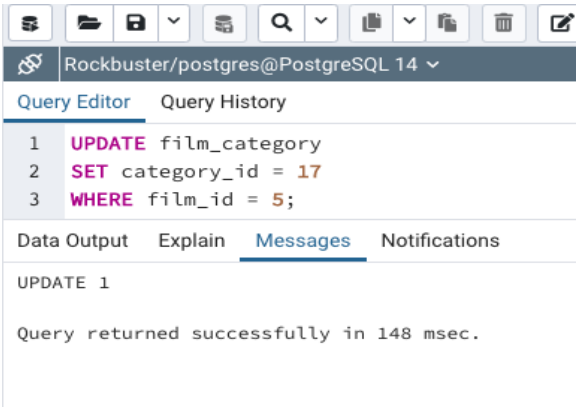


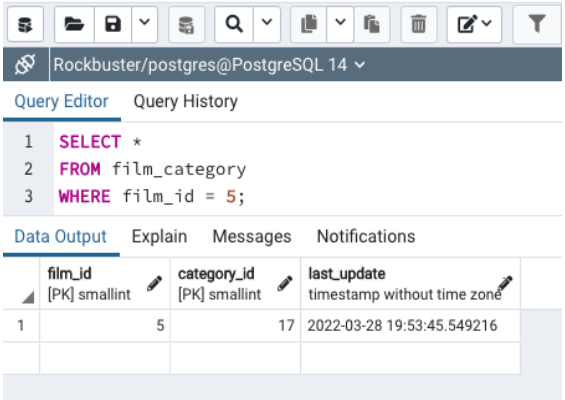
- Once you have the film_ID and category_ID, write an UPDATE command to change the category in the film_category table (not the category table). Copy-paste this command into your answers document.

  The film_id for African Egg is 5. However, we must know the category_id of the genre "Thriller" from the category table (because the film_category column data type in the film_category table is SMALLINT).



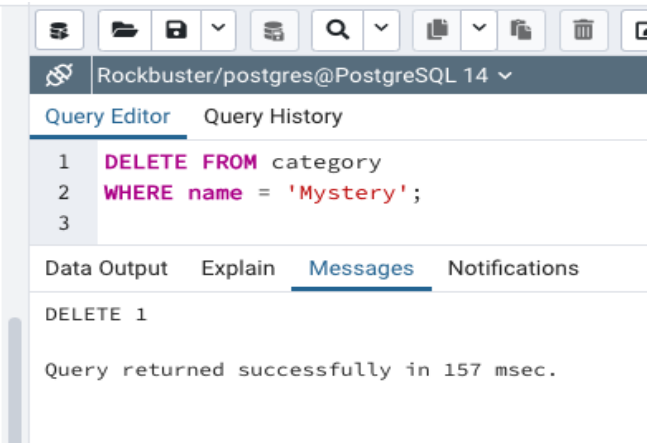
Finally, we can update the film_category using the category_id and film_id.

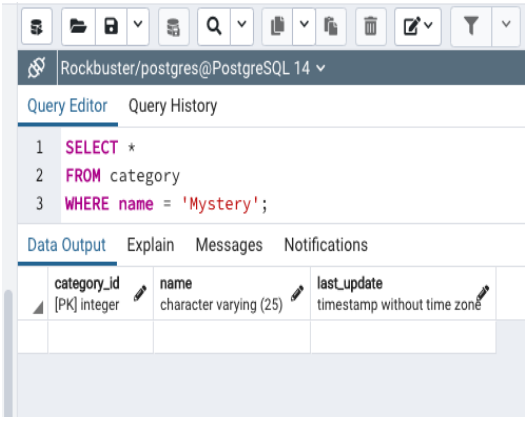




Step 4

- 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.

  SQL makes it easy to work with different table whereas it would require clicking many tabs on the Excel worksheet. Also, manipulating of data is a lot easier with SQL as compared with excel. However, excel is better for quick data visualization and working with small data size.

Bonus

- The SQL query below contains some typos. See if you can fix it based on what you've learned so far about SQL and data types; then try running it in pgAdmin 4. If the query works, copy it into your Answers 3.3 document.

If you get this you're a SQL champ!

Rockbuster/postgres@PostgreSQL 14

Query Editor | Query History

```sql
CREATE TABLE employees_3
(
employee_id VARCHAR(30) NOT NULL,
name VARCHAR(50),
contact_number VARCHAR(30) ,
designation_id INTEGER,
last_update TIMESTAMP NOT NULL DEFAULT now(),
CONSTRAINT employee_pkey PRIMARY KEY (employee_id)
)
```

Data Output | Explain | Messages | Notifications

```
CREATE TABLE

Query returned successfully in 93 msec.
```

Rockbuster/postgres@PostgreSQL 14

Query Editor | Query History

```sql
SELECT *
FROM employees_3;
```

Data Output | Explain | Messages | Notifications

employee_id	name	contact_number	designation_id	last_update
[PK] character varying (30)	character varying (50)	character varying (30)	integer	timestamp without time zone