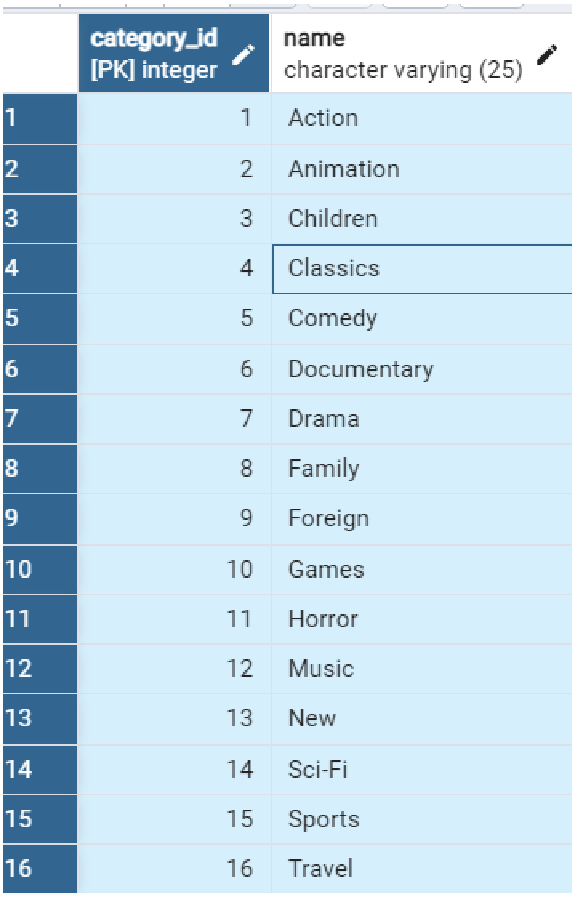
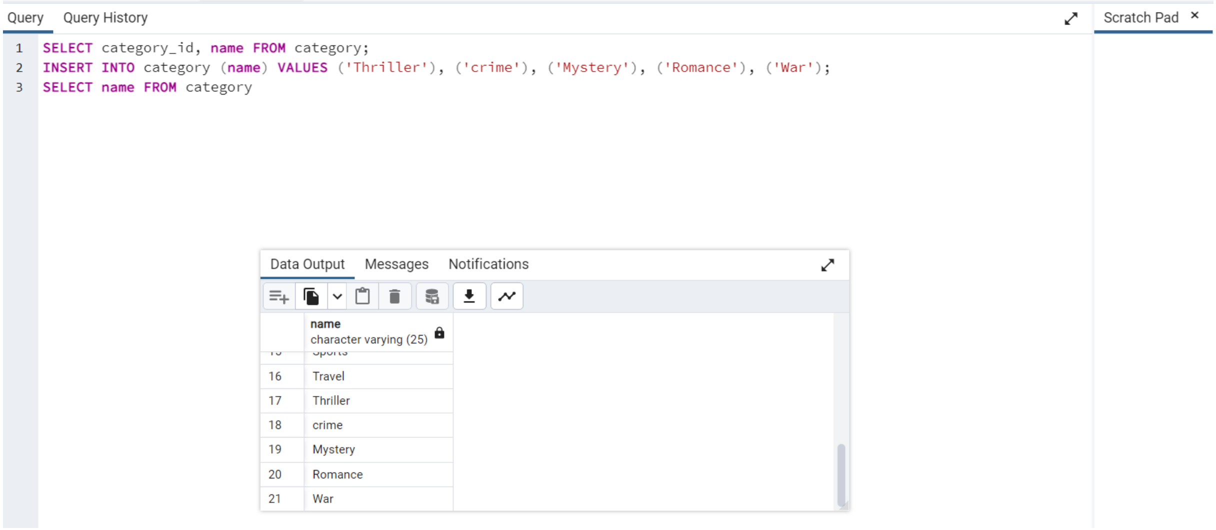
3.3: SQL for Data Analysis

1. Write a SELECT command to find out what film genres exist in the category table.



	category_id [PK] integer	name character varying (25)
1	1	Action
2	2	Animation
3	3	Children
4	4	Classics
5	5	Comedy
6	6	Documentary
7	7	Drama
8	8	Family
9	9	Foreign
10	10	Games
11	11	Horror
12	12	Music
13	13	New
14	14	Sci-Fi
15	15	Sports
16	16	Travel

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



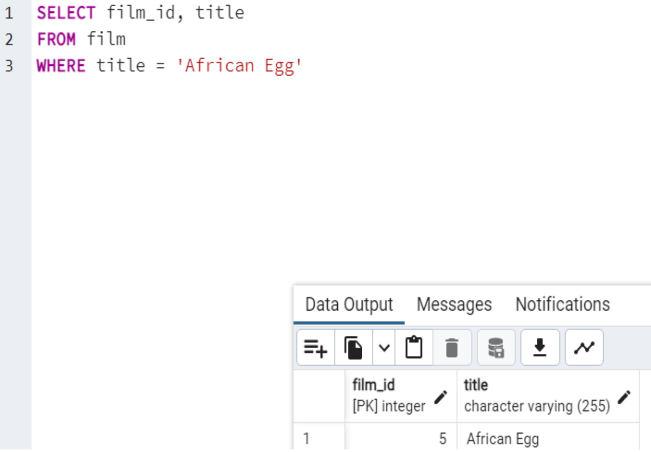
```
SELECT category_id, name FROM category;
INSERT INTO category (name) VALUES ('Thriller'), ('crime'), ('Mystery'), ('Romance'), ('War');
SELECT name FROM category
```

	name character varying (25)
15	Sports
16	Travel
17	Thriller
18	crime
19	Mystery
20	Romance
21	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?

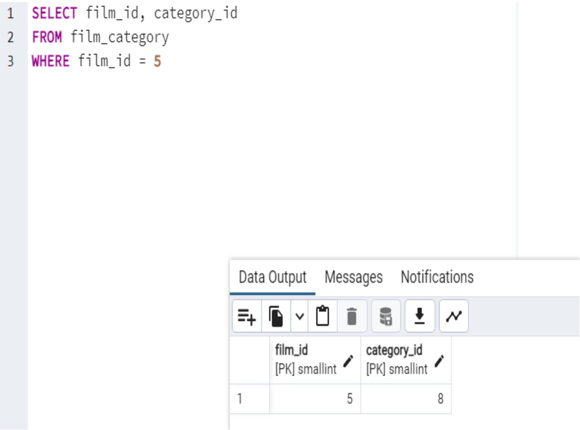
- NOT NULL – the constraints makes sure that there are no missing values in every column
- PRIMARY KEY – Category_id is the primary key column. The primary key column can`t contain any null or duplicate values.

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



```sql
SELECT film_id, title
FROM film
WHERE title = 'African Egg'
```

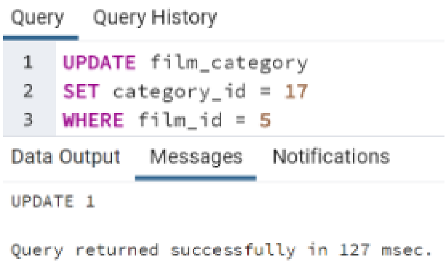
	film_id [PK] integer	title character varying (255)
1	5	African Egg



```sql
SELECT film_id, category_id
FROM film_category
WHERE film_id = 5
```

	film_id [PK] smallint	category_id [PK] smallint
1	5	8

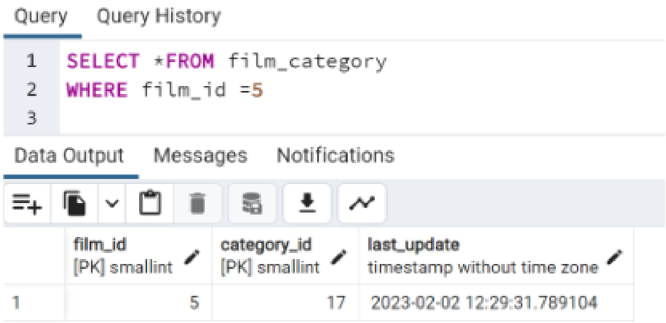
Once you have the film_ID and category_ID, write an UPDATE command to change the category in the film_category table (not the category table). Copy-paste this command into your answers document.



```sql
UPDATE film_category
SET category_id = 17
WHERE film_id = 5
```

```
UPDATE 1

Query returned successfully in 127 msec.
```

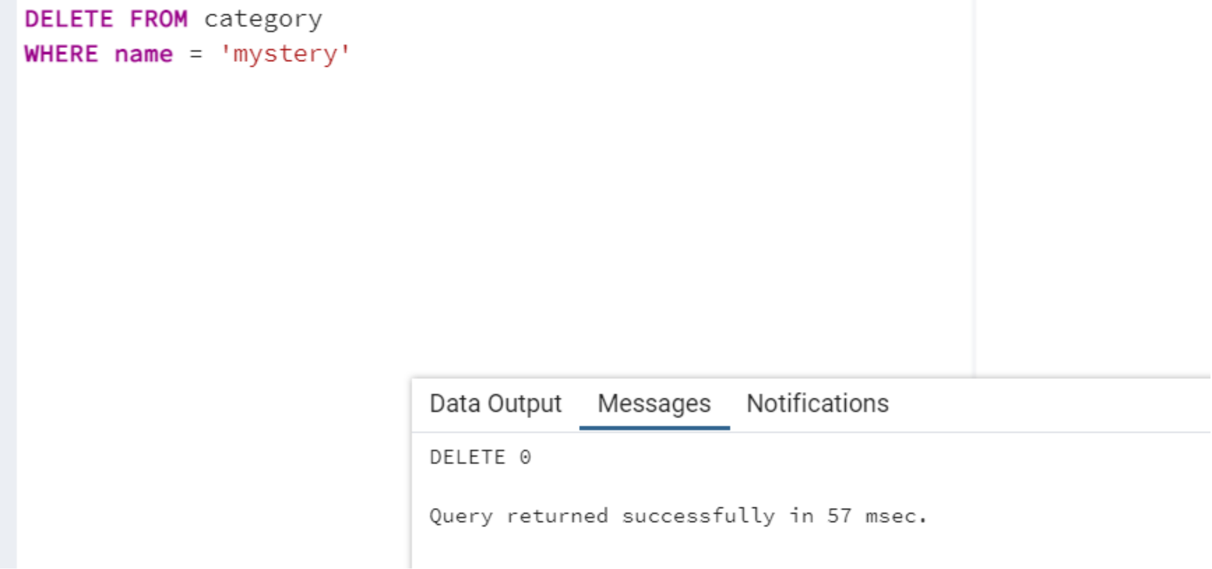


```sql
SELECT *FROM film_category
WHERE film_id =5

```

	film_id [PK] smallint	category_id [PK] smallint	last_update timestamp without time zone
1	5	17	2023-02-02 12:29:31.789104

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.



```
DELETE FROM category
WHERE name = 'mystery'
```

```
DELETE 0

Query returned successfully in 57 msec.
```

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.

Excel provides a easier insight in the dataset and is more intuitive.. However, SQL is easier to manipulate and to change, edit, and delete data. It is faster because it requires only a few commands, while in Excel you need to physically identify the values to delete them.

SQL is more beneficial for datasets with more rows where more scrolling would be required in Excel.