3.3- SQL for Data Analyst

Step 1

Your first task is to find out what film genres already exist in the category table:

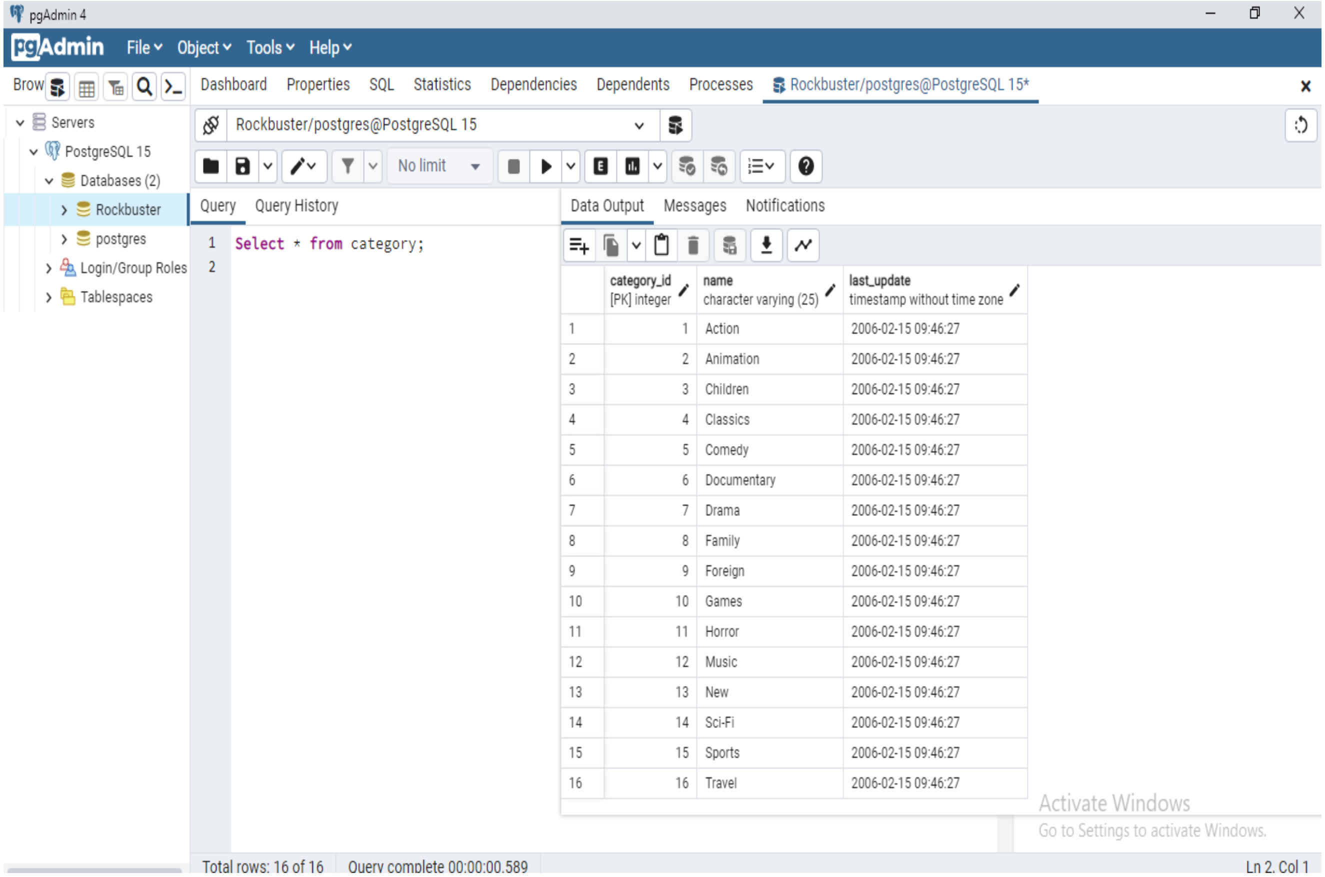
A. Open pgAdmin 4, click the Rockbuster database, and open the Query Tool.

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

C. Copy-paste the output into your answers document or write the answers out—it's up toyou. Make sure to include the category ID for each genre.

QUERY:

```
SELECT * FROM Category;
```



```
Select * from category;
```

	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

Total rows: 16 of 16 Query complete 00:00:00.589 Ln 2, Col 1

Step 2

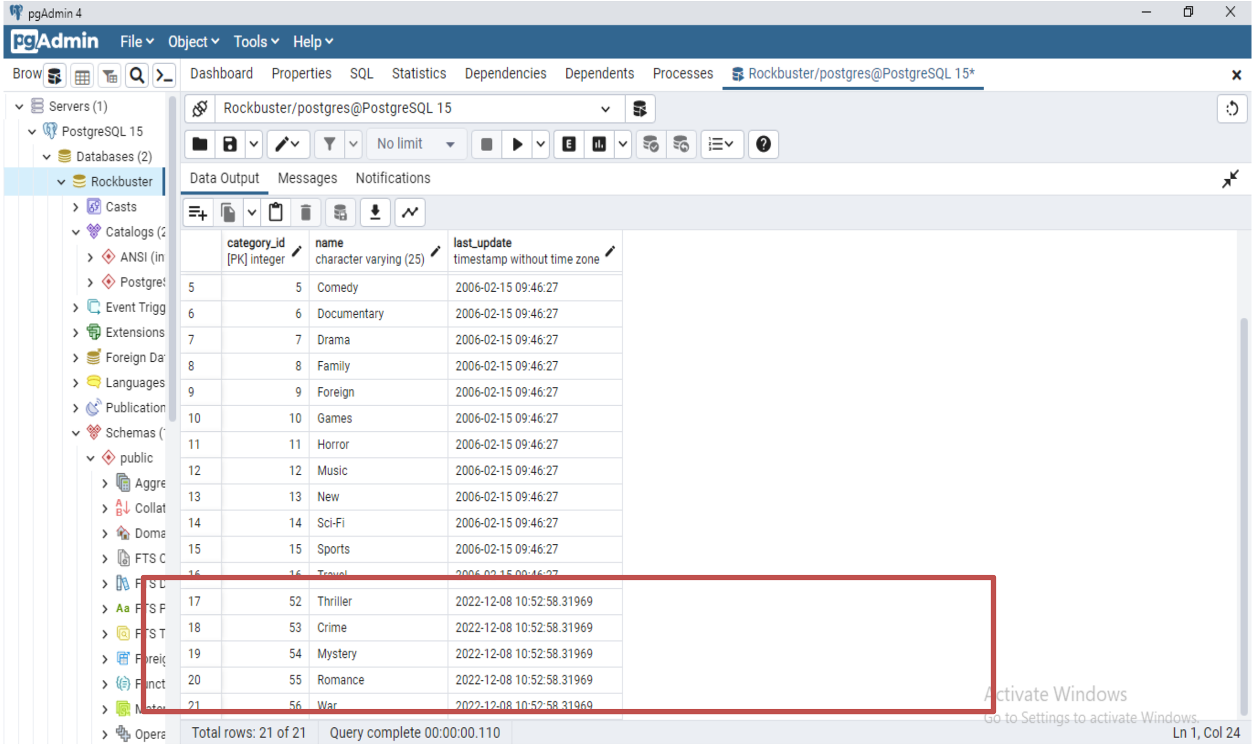
You’re ready to add some new genres! Write an INSERT statement to add the following genres to the category table: Thriller, Crime, Mystery, Romance, and War:

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

QUERY:

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

Screenshot



	category_id [PK] integer	name character varying (25)	last_update timestamp without time zone
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	52	Thriller	2022-12-08 10:52:58.31969
18	53	Crime	2022-12-08 10:52:58.31969
19	54	Mystery	2022-12-08 10:52:58.31969
20	55	Romance	2022-12-08 10:52:58.31969
21	56	War	2022-12-08 10:52:58.31969

Total rows: 21 of 21 Query complete 00:00:00.110

B.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)
);
```

CONSTRAINTS:

1. NOT NULL DEFAULT Constraint: This helps to ensure that a column cannot have any empty or missing. Here we used as:
   - Category_id- The data type should be an Integer and it cannot be null.
   - Name- The data shoul be in text and I cant null
   - Last_update- the data type should be the Timestamp with Time Zone and cannot be Null.
2. PRIMARY KEY: it to give each record in a table Unique id and Not null.

These Constraint help us to verify whether new Value meet the Conditions or not. It also make less mistakes in Data sets and automatically check data quality and let us to know when we make a mistakes.

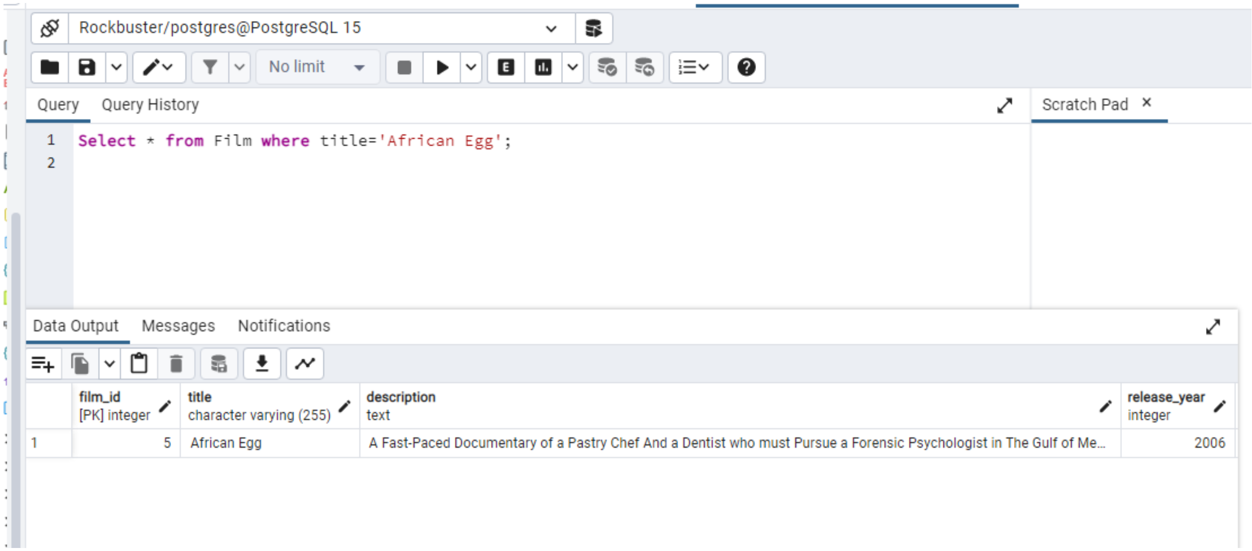
Step 3

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

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

QUERY:

```
SELECT * FROM Film WHERE title= ‘ African egg’;
```



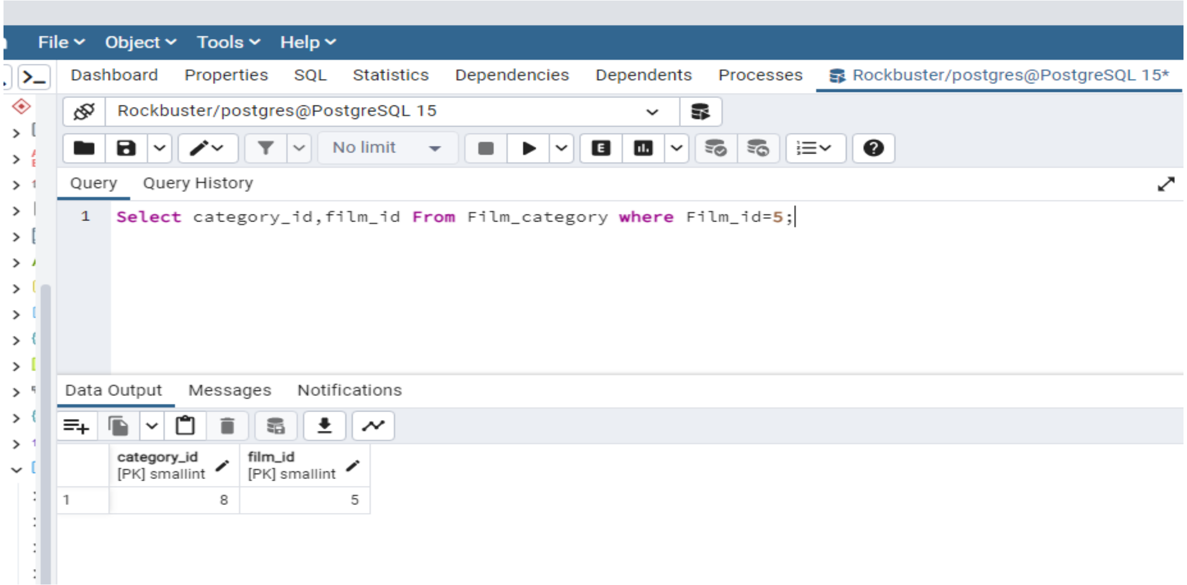
```
Select * from Film where title='African Egg';
```

	film_id [PK] integer	title character varying (255)	description text	release_year integer
1	5	African Egg	A Fast-Paced Documentary of a Pastry Chef And a Dentist who must Pursue a Forensic Psychologist in The Gulf of Me...	2006

B. 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)

QUERY:

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



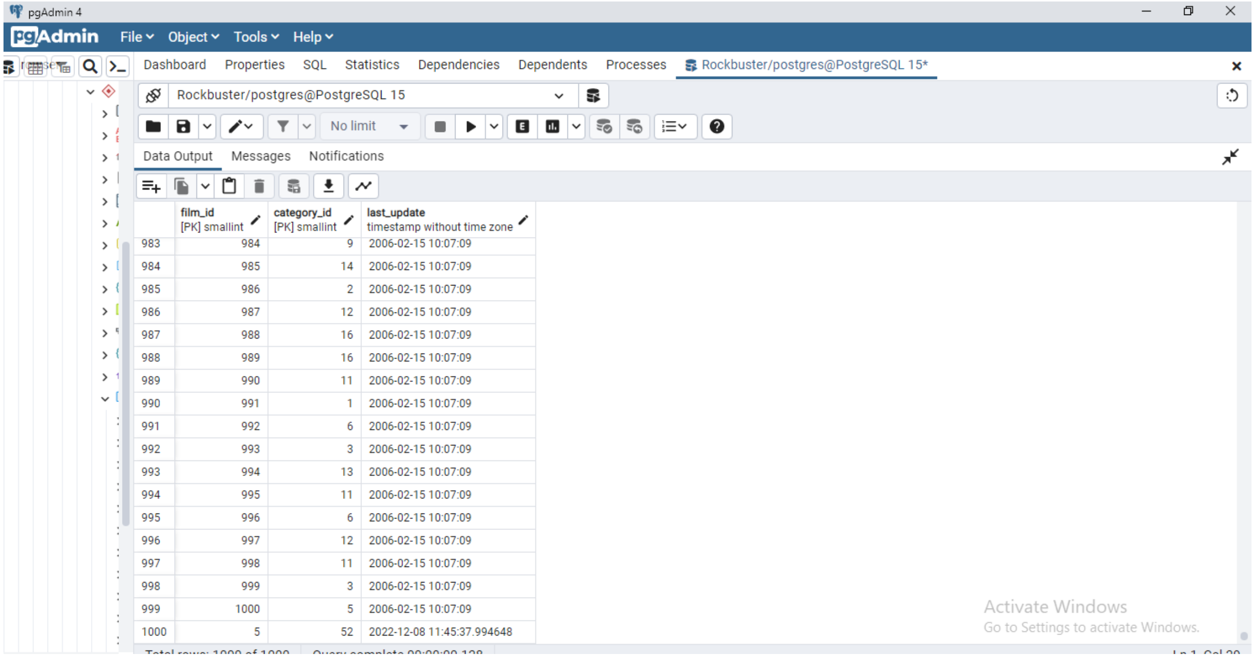
```sql
Select category_id,film_id From Film_category where Film_id=5;
```

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

Part 3B

QUERY

UPDATE film_category SET category_id=52 where film_id = 5;



	film_id [PK] smallint	category_id [PK] smallint	last_update timestamp without time zone
983	984	9	2006-02-15 10:07:09
984	985	14	2006-02-15 10:07:09
985	986	2	2006-02-15 10:07:09
986	987	12	2006-02-15 10:07:09
987	988	16	2006-02-15 10:07:09
988	989	16	2006-02-15 10:07:09
989	990	11	2006-02-15 10:07:09
990	991	1	2006-02-15 10:07:09
991	992	6	2006-02-15 10:07:09
992	993	3	2006-02-15 10:07:09
993	994	13	2006-02-15 10:07:09
994	995	11	2006-02-15 10:07:09
995	996	6	2006-02-15 10:07:09
996	997	12	2006-02-15 10:07:09
997	998	11	2006-02-15 10:07:09
998	999	3	2006-02-15 10:07:09
999	1000	5	2006-02-15 10:07:09
1000	5	52	2022-12-08 11:45:37.994648

Total rows: 1000 of 1000 Query complete 00:00:00.128 Ln 1, Col 20

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.

QUERY:

```
DELETE FROM category WHERE name=’Mystery’ ;
```

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?

Pros using SQL:

1. Easy to use with simple language Commands than Formulas learning in Excel.
2. Easy to maintain Large Data in Sql than Excel
3. SQl is faster than Excel.
4. Showing Your Mistakes with information about which likely the mistake is or where it is.

Cons using SQl

1. I think a person need to be perfect in sql and must have great basic knowledge otherwise Data gone, it is difficult to retrieve.