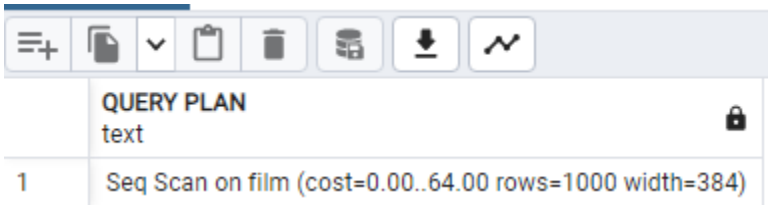
Task 3.4 – Database Querying in SQL

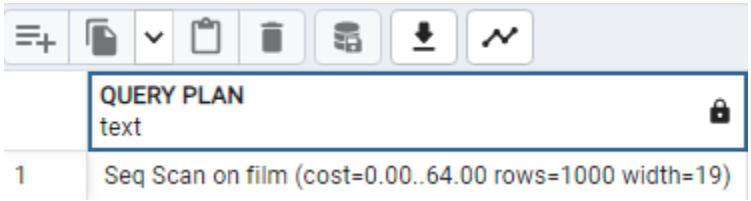
1. **SELECT film_id, title**
   FROM film

Original query



	QUERY PLAN text
1	Seq Scan on film (cost=0.00..64.00 rows=1000 width=384)

Revised query



	QUERY PLAN text
1	Seq Scan on film (cost=0.00..64.00 rows=1000 width=19)

Both queries "cost" the same, where returning all the rows is 64. The only difference I found was the time it took to query both. The revised query was faster because I am only querying two columns as opposed to all the columns.

2. **SELECT title, release_year, rental_rate**
   FROM film
   ORDER BY title, release_year DESC, rental_rate DESC

Data output | Messages | Notifications

	title character varying (255)	release_year integer	rental_rate numeric (4,2)
1	Academy Dinosaur	2006	0.99
2	Ace Goldfinger	2006	4.99
3	Adaptation Holes	2006	2.99
4	Affair Prejudice	2006	2.99
5	African Egg	2006	2.99
6	Agent Truman	2006	2.99
7	Airplane Sierra	2006	4.99
8	Airport Pollock	2006	4.99
9	Alabama Devil	2006	2.99
10	Aladdin Calendar	2006	4.99

3. **SELECT rating, AVG(rental_rate)**
   FROM film
   GROUP BY rating

Data output | Messages | Notifications

	rating mpaa_rating	avg numeric
1	R	2.9387179487179487
2	NC-17	2.970952380952381
3	G	2.888876404494382
4	PG	3.0518556701030928
5	PG-13	3.034843049327354

SELECT rating, MIN(rental_duration), MAX(rental_duration)
FROM film
GROUP BY rating

Data output | Messages | Notifications

	rating mpaa_rating	min smallint	max smallint
1	R	3	7
2	NC-17	3	7
3	G	3	7
4	PG	3	7
5	PG-13	3	7

4. Data engineers are the ones who handle the ETL (Extract, Transform, and Load) procedure. Data is collected from multiple data sources, and then converted into another format. Last, the transformed data is inserted or loaded into the new database.

   Analyzing the data before it's been loaded is like analyzing raw data. It may not make sense by itself, or it may not have the correct formatting, which can lead to confusion.

Ordering the Data: Link for title, release year, and rental rate

Grouping Data: Link for AVG rental rate

Grouping Data: Link for MIN/MAX rental duration