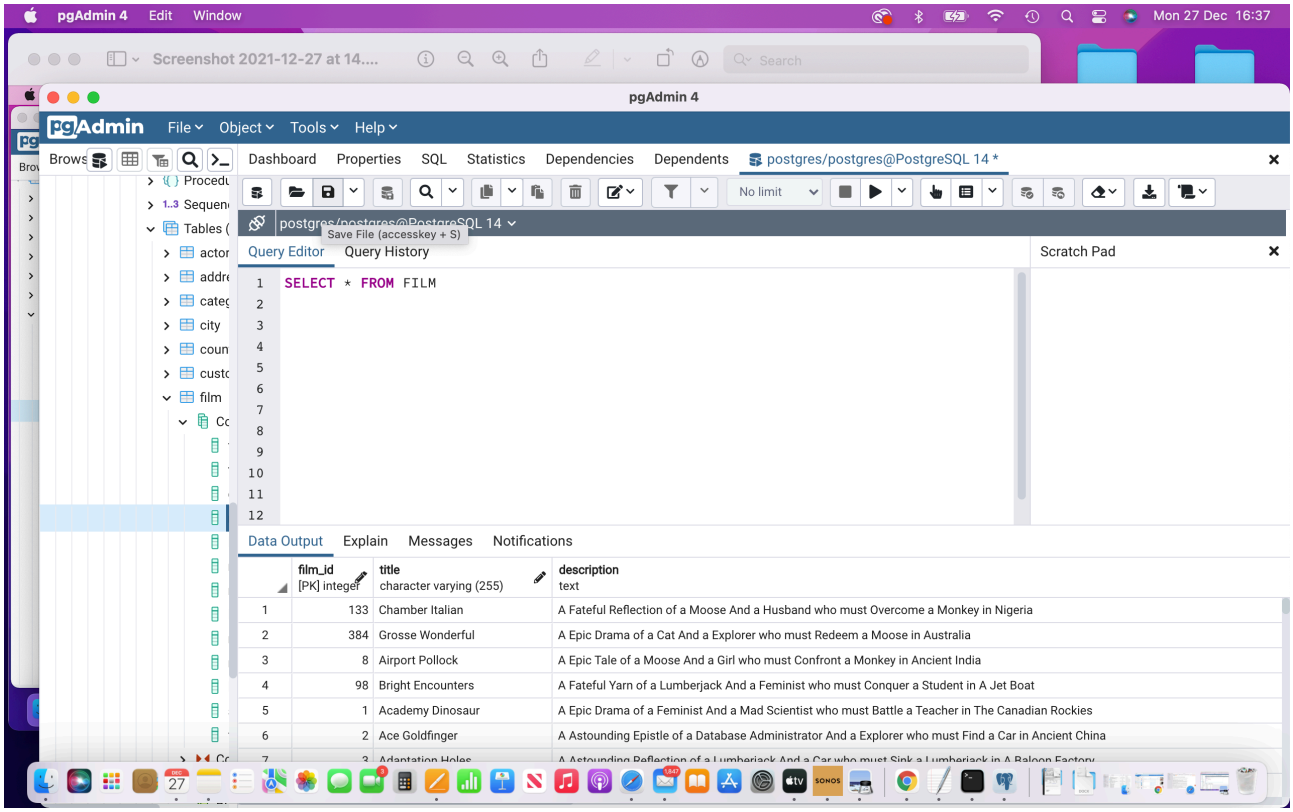
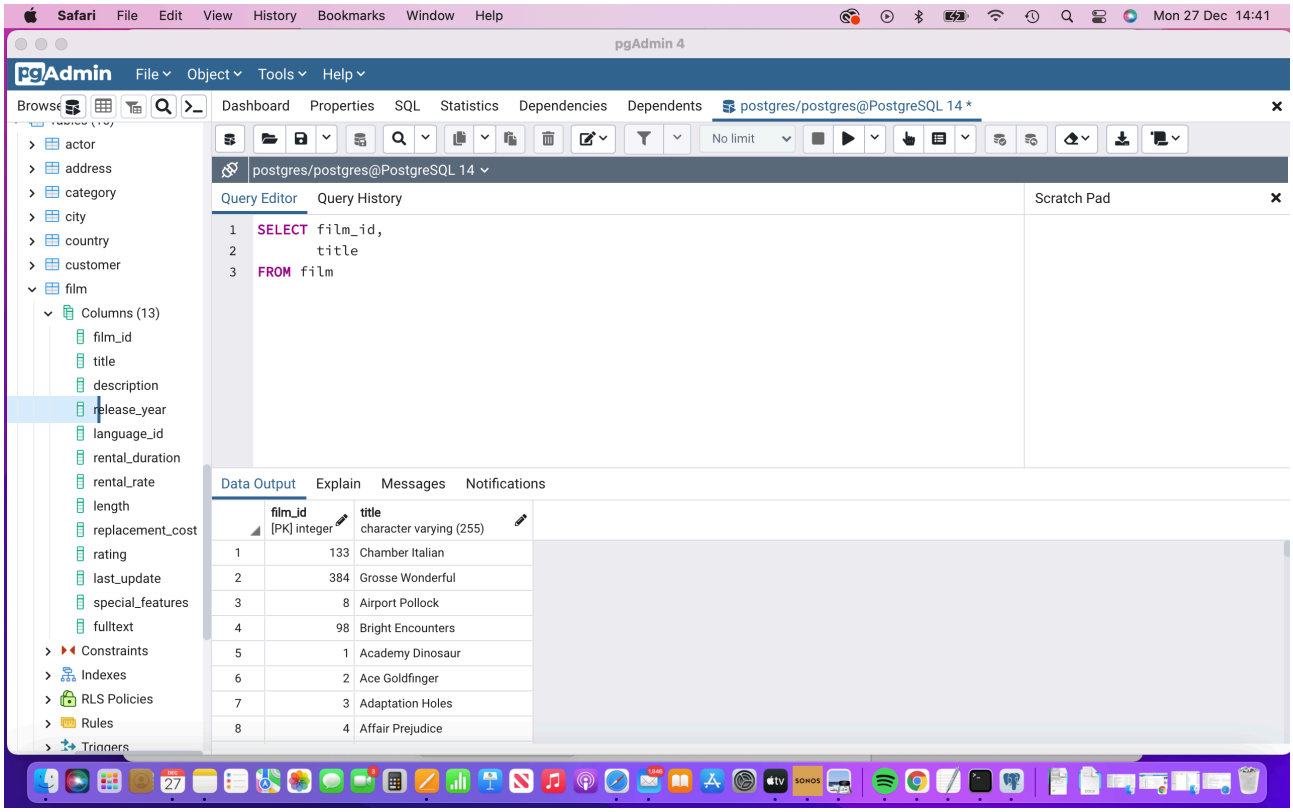
EXERCISE 3.4

Q1a.



Q1b.

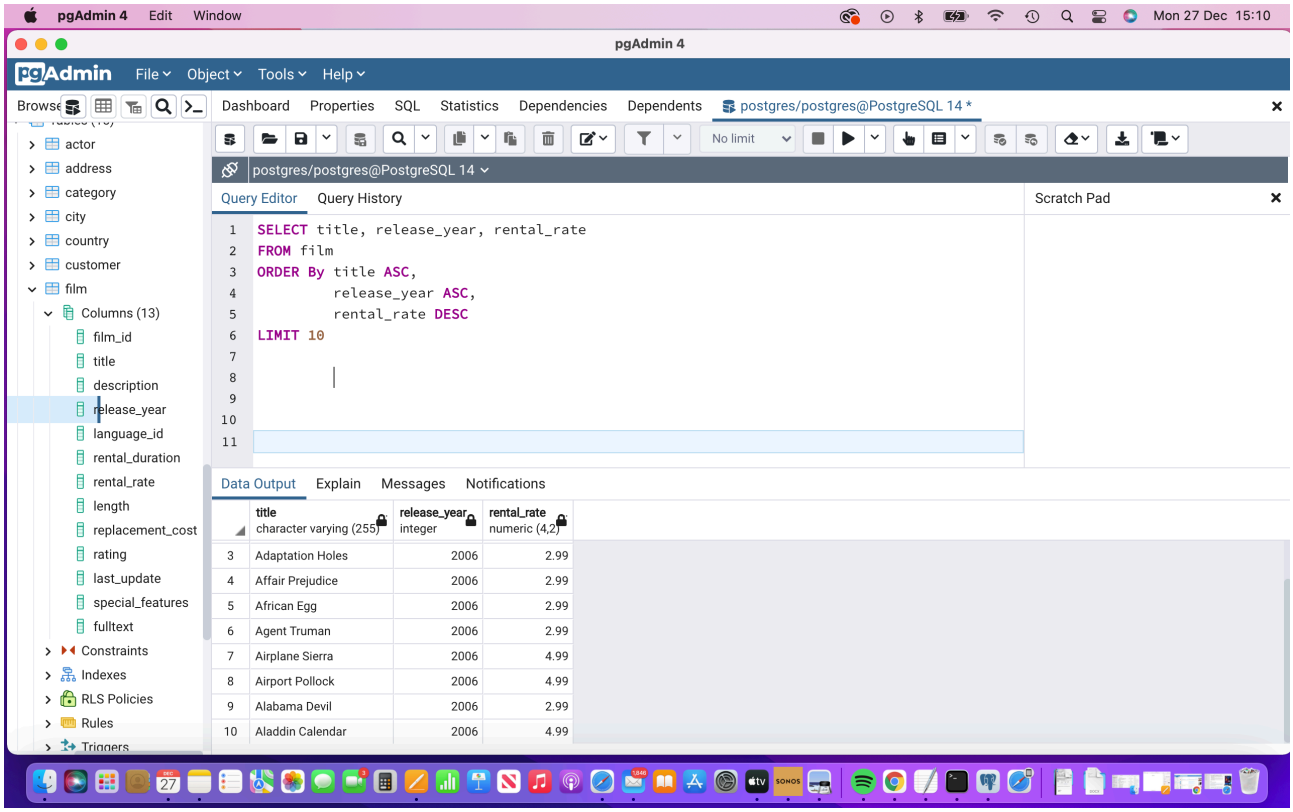


Q1 c.

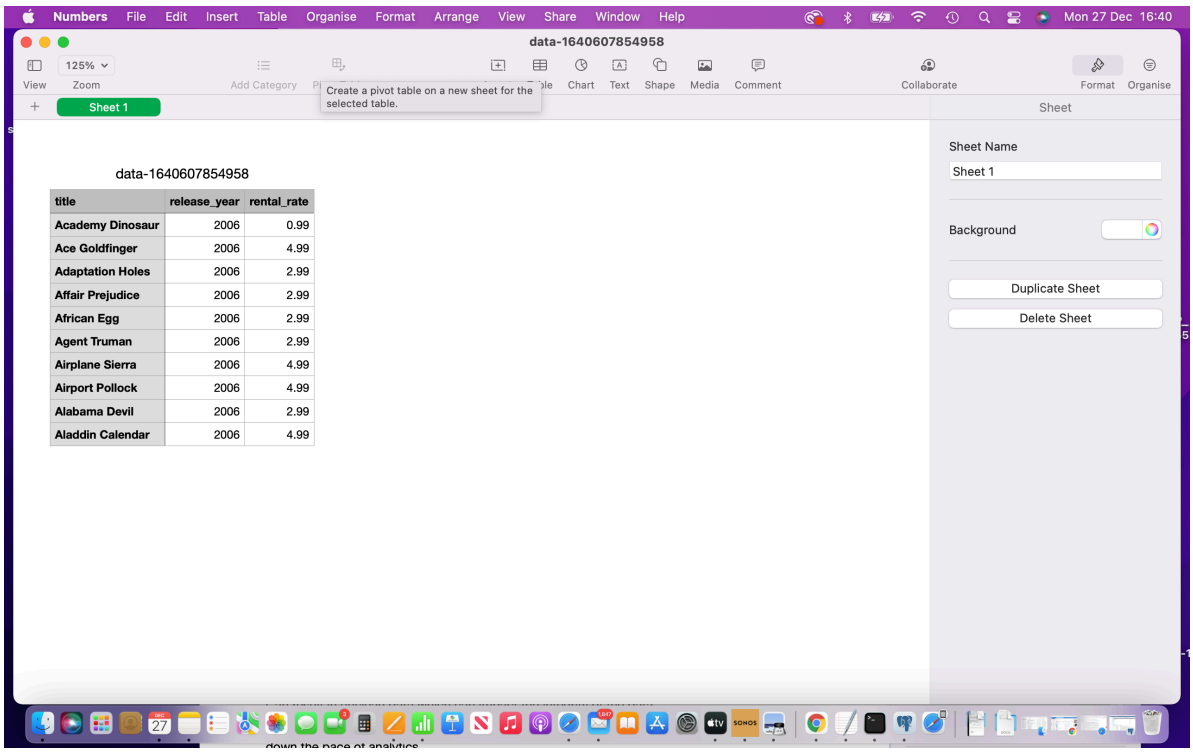
The Select * Command pulls all the data in the Film table, some of which is not needed or required.

Focusing on specific columns in my query, allowed me to be more efficient, and is much cleaner and only pulls the required information to meet needs and requirements.

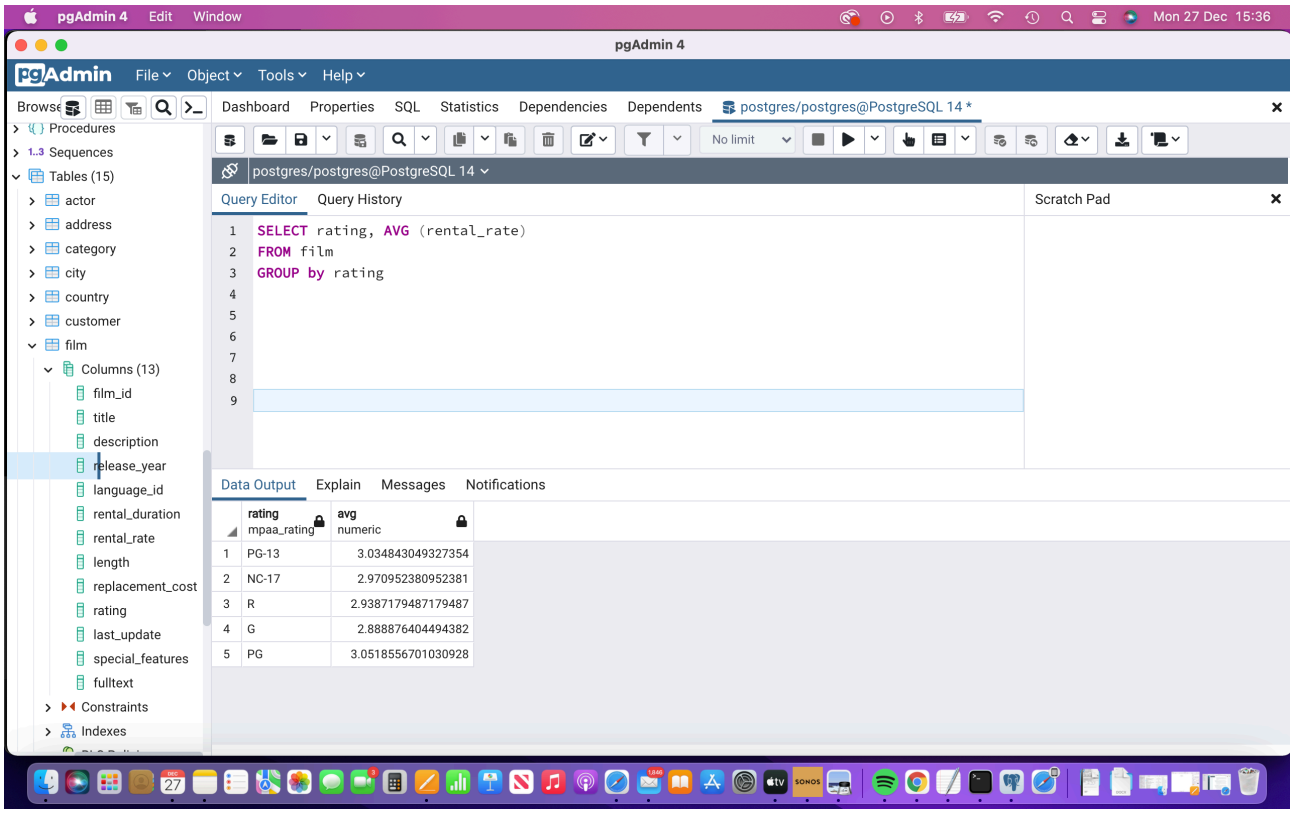
Q2a



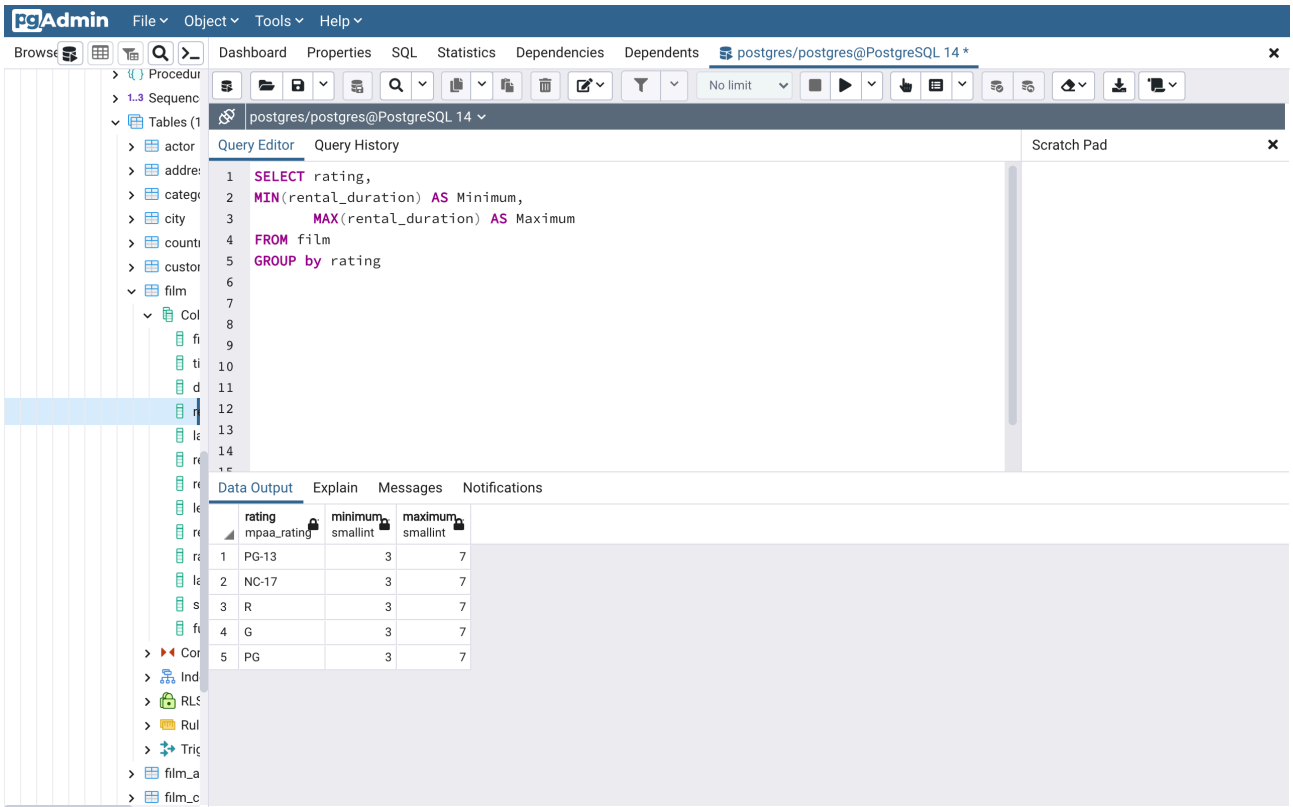
Q2b.



Q3.


pgAdmin 4
Edit
Window
Mon 27 Dec 15:36
pgAdmin 4
File
Object
Tools
Help
Dashboard
Properties
SQL
Statistics
Dependencies
Dependents
postgres/postgres@PostgreSQL 14 *
No limit
Procedures
Sequences
Tables (15)
actor
address
category
city
country
customer
film
Columns (13)
film_id
title
description
release_year
language_id
rental_duration
rental_rate
length
replacement_cost
rating
last_update
special_features
fulltext
Constraints
Indexes
Query Editor
Query History
Scratch Pad
1 SELECT rating, AVG (rental_rate)
2 FROM film
3 GROUP by rating
Data Output
Explain
Messages
Notifications
rating mpaa_rating
avg numeric
1 PG-13 3.034843049327354
2 NC-17 2.970952380952381
3 R 2.9387179487179487
4 G 2.888876404494382
5 PG 3.0518556701030928

Q3b.


pgAdmin
File
Object
Tools
Help
Dashboard
Properties
SQL
Statistics
Dependencies
Dependents
postgres/postgres@PostgreSQL 14 *
No limit
Query Editor
Query History
Scratch Pad
1 SELECT rating,
2 MIN(rental_duration) AS Minimum,
3 MAX(rental_duration) AS Maximum
4 FROM film
5 GROUP by rating
Data Output
Explain
Messages
Notifications
rating mpaa_rating
minimum smallint
maximum smallint
1 PG-13 3 7
2 NC-17 3 7
3 R 3 7
4 G 3 7
5 PG 3 7

Q4. Procedure for migrating data includes Extracting, Transformation, and Loading, which is commonly known as ETL.

Data engineers are usually responsible for ETL, however Data Analysts do play an involved role with ETL particularly in Startups.

Q4b. Can result in analysing unclean data which can impact the integrity of the data.

Data that hasn't gone through the ETL process impacts efficiency, increases time, cost and slows down the pace of analytics.