3.4

1. Refining your Query -

Query Editor Query History

```
1  SELECT film_id, title FROM film
```

Data Output   Explain   Messages   Notifications

	film_id [PK] integer	title character varying (255)
1	133	Chamber Italian
2	384	Grosse Wonderful
3	8	Airport Pollock
4	98	Bright Encounters
5	1	Academy Dinosaur
6	2	Ace Goldfinger
7	3	Adaptation Holes
8	4	Affair Prejudice
9	5	African Egg
10	6	Agent Truman

Query Editor   Query History

```
1  Explain
2  SELECT film_id, title FROM film
```

Data Output   Explain   Messages   Notifications

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

Query Editor   Query History

```
1  Explain
2  SELECT *
3  FROM film;
```

Data Output   Explain   Messages   Notifications

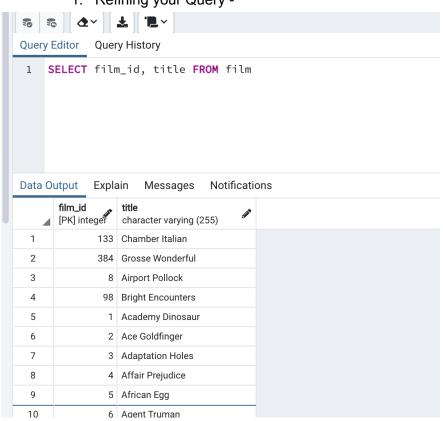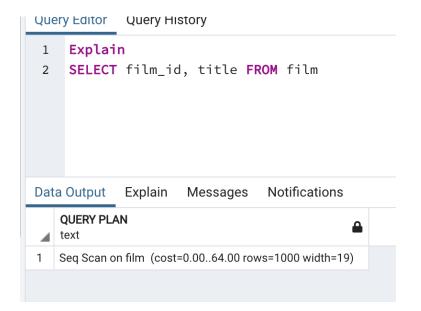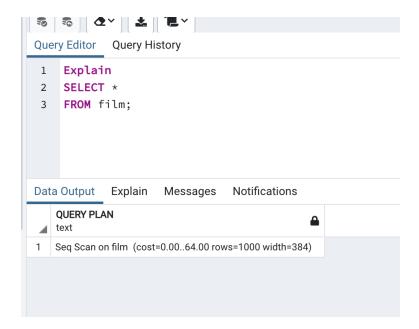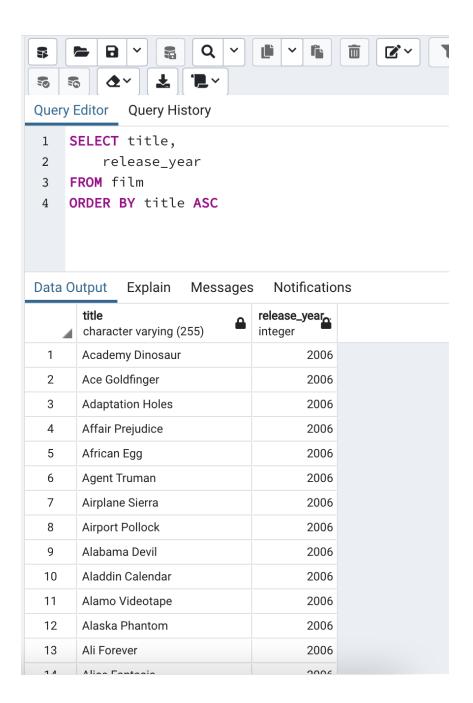| QUERY PLAN text	🔒	
1	Seq Scan on film (cost=0.00..64.00 rows=1000 width=384)	

It says they are both zero for the cost (or time) it took. But to optimize the time it takes can be adjusted by writing different SQL queries. You can do this by pinpointing the data you want with certain commands like SELECT.
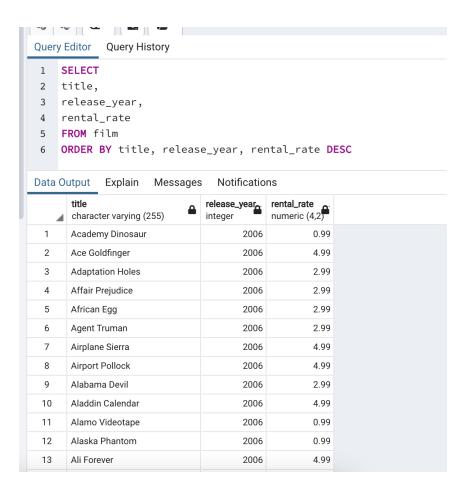
    2. Ordering your data:

a.

```
1   SELECT *
2   FROM film
3   ORDER BY title ASC
```

CREAT

(

categ

name

last_u

CONS

Data Output   Explain   Messages   Notifications

	film_id ✎ [PK] integer	title ✎ character varying (255)	description text
1	1	Academy Dinosaur	A Epic Drama of a Feminist And a Mad Scientist who must
2	2	Ace Goldfinger	A Astounding Epistle of a Database Administrator And a Ex
3	3	Adaptation Holes	A Astounding Reflection of a Lumberjack And a Car who mu
4	4	Affair Prejudice	A Fanciful Documentary of a Frisbee And a Lumberjack who
5	5	African Egg	A Fast-Paced Documentary of a Pastry Chef And a Dentist
6	6	Agent Truman	A Intrepid Panorama of a Robot And a Boy who must Escap
7	7	Airplane Sierra	A Touching Saga of a Hunter And a Butler who must Discov
8	8	Airport Pollock	A Epic Tale of a Moose And a Girl who must Confront a Mor
9	9	Alabama Devil	A Thoughtful Panorama of a Database Administrator And a
10	10	Aladdin Calendar	A Action-Packed Tale of a Man And a Lumberjack who mus
11	11	Alamo Videotape	A Boring Epistle of a Butler And a Cat who must Fight a Pas
12	12	Alaska Phantom	A Fanciful Saga of a Hunter And a Pastry Chef who must Va

b.

C.

```sql
1  SELECT
2  title,
3  release_year,
4  rental_rate
5  FROM film
6  ORDER BY title, release_year, rental_rate DESC
```

Data Output   Explain   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
11	Alamo Videotape	2006	0.99
12	Alaska Phantom	2006	0.99
13	Ali Forever	2006	4.99

Film_Collection.csv

3. Grouping the data

A.

```
1   SELECT
2   AVG (rental_rate) AS "Average Rental Rate",
3   rating AS "Rating"
4   FROM film
5   Group BY rating
6   Order BY rating;
```

Data Output   Explain   Messages   Notifications

	Average Rental Rate 🔒 numeric	Rating 🔒 mpaa_rating
1	2.888876404494382	G
2	3.0518556701030928	PG
3	3.034843049327354	PG-13
4	2.9387179487179487	R
5	2.970952380952381	NC-17

AVG_rental_rating.csv

B.

Query Editor   Query History

```
1   SELECT
2   MIN (rental_duration) AS "Minimum Rental Duration",
3   MAX (rental_duration) AS "Maximum Rental Duration",
4   rating AS "Rating"
5   FROM film
6   Group BY rating
7   Order BY rating;
```

Data Output   Explain   Messages   Notifications

	Minimum Rental Duration 🔒 smallint	Maximum Rental Duration 🔒 smallint	Rating 🔒 mpaa_rating
1	3	7	G
2	3	7	PG
3	3	7	PG-13
4	3	7	R
5	3	7	NC-17

Min_Max_duration_ofrentals.csv

4. Database migration:

Can you outline the procedure for migrating the data and who will be responsible for it?

The procedures are extract, transform and load. The software engineers are responsible for this task. You extract the data from multiple sources, then transform the data into a specific format and then load the data into the new database.

What problems do you foresee if you start analyzing the data before it's been loaded into the data warehouse?

If you analyze data that isn't in the correct format, it will be very difficult to analyze. The coding won't work and things like primary keys won't be recognized.