1)

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

Data Output   Messages   Explain   Notifications

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

```
1   EXPLAIN
2   SELECT
3   film_id,
4   title
5   FROM film;
```

Data Output   Messages   Explain   Notifications

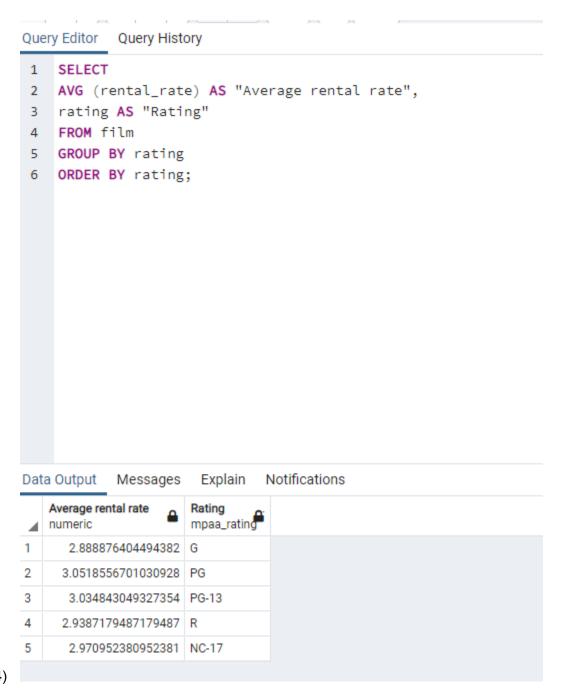| QUERY PLAN 🔒	
text	
1	Seq Scan on film (cost=0.00..64.00 rows=1000 width=19)

Both queries cost 0 and have a score of 64. If this was a larger database you would see more benefits of a refined query but overall the more specific query you can run the better.

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

Data Output   Messages   Explain   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
14	Alice Fantasia	2006	0.99

2)

3) https://docs.google.com/spreadsheets/d/1Zyz1NMTKhAKHwjjxLZP3HZOKVu2xu1AgEtw2rgHY3uI/edit#gid=1450602450

Query Editor   Query History

```
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   Messages   Explain   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

4) https://docs.google.com/spreadsheets/d/1qJ2X4NWq4QiK4uN_dSpDoc3iuEHyeLI0iw3EQ922KtI/edit#gid=1901467665

There is a process for migrating data and its called ETL. ETL means to extract, transform and load. Data engineers would source the data, then transform it into a compatible format and do any aggregations necessary and then load it into a database.

The problems I foresee if you start analyzing the data before it's loaded into a data warehouse is that it has not been properly transformed or cleaned so you are at higher risk of working with incorrect data. There is also a very high chance that you'd also be working with raw data which means it is dirty so there will be a huge margin of error with data results.