# 1a. query command SELECT film id, title FROM film;

## 1b Old query



Difference: The cost from both queries is the same 0.00-64.00. However, the actual query runtime differs that the new query takes faster than the original one.

2a. query Command
SELECT title, release\_year, rental\_rate
FROM film
ORDER BY title, release\_year, rental\_rate DESC;

=+		4	<u> </u>		
	title character varying (255)	â	release, integer	year 🔓	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
5	Agent Truman			2006	2.99
7	Airplane Sierra			2006	4.99
3	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
15	Alien Center			2006	2.99
16	Alley Evolution			2006	2.99
17	Alone Trip			2006	0.99
18	Alter Victory			2006	0.99
19	Amadeus Holy			2006	0.99
20	Amelie Hellfighters			2006	4.99
21	American Circus			2006	4.99
22	Amistad Midsummer			2006	2.99
23	Anaconda Confessions			2006	0.99
24	Analyze Hoosiers			2006	2.99
25	Angels Life			2006	2.99
26	Annie Identity			2006	0.99
27	Anonymous Human			2006	0.99
28	Anthem Luke			2006	4.99
29	Antitrust Tomatoes			2006	2.99
30	Anything Savannah			2006	2.99

### 3a. query command

SELECT rating, AVG(rental rate) FROM film GROUP BY rating;

	rating mpaa_rating	avg numeric
1	PG	3.051855670
2	NC-17	2.970952380
3	R	2.938717948
4	G	2.888876404
5	PG-13	3.034843049

3b. maximum rental durations for each rating category

Query command:

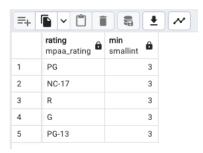
SELECT rating, MAX(rental duration) FROM film GROUP BY rating;



Minimum rental durations for each rating category

Query command:

SELECT rating, MIN(rental\_duration) FROM film GROUP BY rating;



4a. The first step is extract which involves collecting the data from user behavior data sources; then converted extracted data into another format that we are using; last inserting/loading the transformed data into new database. Usually, Data Engineer is one who are responsible for this.

4b. The data probably will not be in the same format or linked to current data in the data warehouse. It would make the process very manual and time consuming.

#### **BOUNES**

#### Query Query History 1 SELECT rating, MIN (replacement\_cost), MAX(replacement\_cost) 2 FROM film GROUP BY rating 3 ORDER BY CASE WHEN rating='G' then 1 4 WHEN rating='PG' then 2 5 WHEN rating = 'PG-13' then 3 6 WHEN rating = 'R' then 4 7 ELSE 5 8 END Data output Messages Notifications rating rating mpaa\_rating a min numeric a max numeric 9.99 2 PG 9.99 29.99 3 PG-13 9.99 29.99 4 R 9.99 29.99 NC-17 9.99 29.99