

Answers Ex 3.4

Database querying in SQL

- 1: While the **SELECT** film_id, title query is specific to our needs, it does take more time. (See images below.)

Query	Query History	Data Output
1	EXPLAIN SELECT * FROM film	QUERY text 1 Seq Sc
Total rows: 1 of 1		Query complete 00:00:00.049

Query	Query History	Data Output	Messages
1	EXPLAIN SELECT film_id, title FROM film	QUERY PLAN text 1 Seq Scan on film	
Total rows: 1 of 1		Query complete 00:00:00.062	

The only way I can think of optimising is to create a table with only these two values in it.

- 2: See below

	title	release_year	rental_rate
	character varying	integer	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

Total rows: 1000 of 1000 Query complete 00:00:00.092

- 3: Grouping Data

Query	Query History	Data Output	Messages	Notifications	Scratch Pad	×
1	SELECT rating, AVG (rental_rate),					
2	MIN (rental_duration),					
3	MAX (rental_duration)					
4	FROM film					
5	GROUP BY rating					
		rating mpaa_rating	avg numeric	min smallint	max smallint	
1		PG	3.0518556701030928	3	7	
2		R	2.9387179487179487	3	7	
3		NC-17	2.9709523809523810	3	7	
4		PG-13	3.0348430493273543	3	7	
5		G	2.8888764044943820	3	7	

- 4: Database Migration

We would need to follow ETL protocols. **Extract** from the Android app, **Transform** the data into the proper format for the database, **Load** the data into the new/correct database.

The Android team or Data Engineer for Android should be in charge of the task. It would also be good to have the Android DE and the Rockbuster DE sit down and agree on the standards/guidelines for the data going forward to not have any confusion etc. Always have a plan!

Analysing the data before migration can lead to bad processes. One team may prioritise certain data over another, which may not be accurate for other teams involved. Naming protocols etc should all be agreed upon and set down in a 'Style Guide' for everyone to follow. Otherwise you could have 'mpaa-ratings', 'film_ratings' and 'age_ratings' all showing the same information. Bloat is to be avoided at all costs.