

Databases & SQL for Analysts

Task 3.5: Filtering Data

Ngawang Dhundup

LINK NAME:

Directions:

1. Write some SQL queries to return a lists of films that meet the following conditions. (Hint: These queries are not building upon each other, they are separate! Hence, when running the query for 1b. you don't need to take into account 1a, for example.) Your results tables should include the columns "film_ID," "title," and "description".
 - Film title contains the word Uptown in any position
- `SELECT film_id, title, description`
 - `FROM film`
 - `WHERE title LIKE '%Uptown%';`
 - Film length is more than 120 minutes and rental rate is more than 2.99
 - `SELECT film_id, title, description, length AS film_length, rental_rate`
 - `FROM film`
 - `WHERE length > '120' AND`
 - `rental_rate > '2.99';`
 - Rental duration is between 3 and 7 days (where 3 and 7 aren't inclusive)
 - `SELECT film_id, title, description, rental_duration`
 - `FROM film`
 - `WHERE rental_duration BETWEEN 4 AND 6`
 - `ORDER BY rental_duration;`
 - Film replacement cost is less than 14.99
 - `SELECT film_id, title, description, replacement_cost`
 - `FROM film`
 - `WHERE replacement_cost < 14.99`
 - `ORDER BY replacement_cost;`
 - Film rating is either PG or G
 - `SELECT film_id, title, description, rating AS film_rating`
 - `FROM film`
 - `WHERE rating = 'PG' OR rating = 'G'`
 - `ORDER BY rating;`

2. Download your SQL queries outputs as CSV files using the pgadmin inbuilt functionality. Merge them into one Excel file (.xlsx) and create a separate sheet for each query (label them 1a, 1b, 1c, etc.). You'll use this file for all further questions in this Task too.
3. The query you wrote in step 1e returned a list of movies that meet certain criteria (film rating is either PG or G). The inventory team has asked for the following information about this list.
 - Count of the movies
 - SELECT rating AS film_rating,
 - COUNT(title)
 - FROM film
 - GROUP BY rating
 - HAVING rating IN ('PG', 'G')
 - Average rental rate
 - SELECT rating AS film_rating,
 - COUNT(title),
 - AVG(rental_rate)
 - FROM film
 - GROUP BY rating
 - HAVING rating IN ('PG', 'G')
 - Maximum rental duration and minimum rental duration
 - SELECT rating AS film_rating,
 - COUNT(title),
 - AVG(rental_rate),
 - MAX(rental_duration),
 - MIN(rental_duration)
 - FROM film
 - GROUP BY rating
 - HAVING rating IN ('PG', 'G');
4. To make the output easier for your coworkers to understand, give your aggregate columns the following aliases: “count of movies,” “average movie rental rate,” “maximum rental duration”, and “minimum rental duration”. Run the query and transfer the result into your Excel file on a new sheet as well as the code you used to get there.

Done

5. The customer team would like to see the fields you calculated in step 3 grouped by rating. The totals in your results table should look the same as in step 4, but broken down by the rating column. Copy-paste your query and its output in your answers on a new sheet.

Done

6. Save all of your answers in the Excel file you created in step 2 and upload it here for your tutor to review.