Homework 3

Due: 09/19/2024 by 11:59 PM

For questions 1-4, please use ***dvdrental*** database. For questions 5,6, & 7 use ***northwind*** database. Write the queries down and answer the questions. The submission should contain both the query as well as the answer if required. ***Please do not include screenshots.***

NOTE: This assignment is about SELECT statements. You should not use JOINs in this assignment.

# Dvdrental Database

1. The management is running a promotion to reward the top 5 customers with coupons. What are the “customer\_id” of the top 5 customers by total spend in the ***“payment”*** table?

*SELECT*

*customer\_id,*

*SUM(amount) AS total\_spend FROM payment*

*GROUP BY customer\_id ORDER BY SUM(amount) DESC LIMIT 5;*

The top 5 customers by total spend are **148, 526, 178, 137 and 144**

*Explanation: I used the SUM(amount) to calculate the total spend for each customer and grouped the results by customer\_id. The query is ordered in descending order of total spend to retrieve the top 5 customers.*

1. Write an SQL query to determine the maximum payment for each customer from the ***“payment”*** table. The “customer\_id” should be between 100 and 119. Return “customer\_id” and maximum amount.

*SELECT*

*customer\_id,*

*MAX(amount) AS highest\_amount FROM payment*

*WHERE customer\_id BETWEEN 100 AND 119 GROUP BY customer\_id;*

Explanation: The query uses MAX(amount) to find the highest payment for customers whose IDs are between 100 and 119. The GROUP BY clause is used to group results by customer\_id

1. Write a query to get the average replacement cost for each film rating from the ***“film”*** table, considering only films with a “rental\_rate” greater than or equal to $4.99. Display the results with the highest average replacement cost at the top.

*SELECT*

*rating AS film\_rating,*

*AVG(replacement\_cost) AS average\_replacement\_cost FROM film*

*WHERE rental\_rate >= 4.99 GROUP BY rating*

*ORDER BY average\_replacement\_cost DESC;*

*Explanation: I filtered films with a rental rate greater than or equal to $4.99 and then calculated the average replacement cost (AVG(replacement\_cost)) for each film rating. The results are ordered by the average replacement cost in descending order.*

1. Write a query to determine the maximum payment for 5 customers with customer\_ids (314, 12, 123, 234, 456) from the ***“payment”*** table. Return “customer\_id” and the maximum amount.

*SELECT*

*customer\_id,*

*MAX(amount) AS highest\_payment FROM payment*

*WHERE customer\_id IN (314, 12, 123, 234, 456)*

*GROUP BY customer\_id;*

*Explanation: The IN clause is used to filter for specific customer IDs. The MAX(amount) is used to get the highest payment for each of these customers.*

# Northwind Database

1. Write a query to find the distinct city names from the “ship\_city” column in the ***“orders”*** table, along with the number of orders placed for each city, and return the top three cities with the highest number of orders.

*SELECT*

*DISTINCT ship\_city AS shipped\_to\_cities, COUNT(\*) AS no\_of\_orders*

*FROM orders GROUP BY ship\_city*

*ORDER BY no\_of\_orders DESC LIMIT 3;*

1. *Explanation: I selected distinct ship\_city values and counted the number of orders for each city. The results are ordered by the count of orders in descending order, showing the top 3 cities.*The ***“orders”*** table has a column called *“ship\_via”*, which stores company IDs (encoded as numerical digits). Write a query that uses the *“ship\_via”* column to find which shipping company has the greatest number of orders. Then, manually look up the corresponding company name from the ***“shippers”*** table. Report the number of orders and company name.

*Hint:* First, find the *“ship\_via”* value with the most orders, and then refer to the

***“shippers”*** table to match the *“ship\_via”* ID with its company name.

Example Answer:

…SQL Query….

The company with the most orders is **A** with order number **B**.

*SELECT*

*COUNT(\*) AS no\_of\_orders, ship\_via AS shipping\_company FROM orders*

*GROUP BY ship\_via LIMIT 1;*

The company with the most orders is **Speedy Express** with number of orders **24**

*Explanation: I used the COUNT(\*) function to determine the number of orders for each shipping company (ship\_via). The query retrieves the shipping company with the most orders by grouping by ship\_via. The company name can be cross-referenced manually from the shippers table using the ship\_via value.*

1. List the categories (category\_id) that have more than 3 products in the “***products”*** table. Only include products with a unit price between $10 and $30. Then, refer to the **“categories”** table to match the category\_id with category name and report the names of these categories.

*SELECT*

*category\_id,*

*COUNT(\*) AS number\_of\_products FROM products*

*WHERE unit\_price BETWEEN 10 AND 30 GROUP BY category\_id*

*HAVING COUNT(\*) > 3;*

The categories that have more than 3 products in each are **Seafood, Beverages, Condiments and Confections**

*Explanation: The query retrieves categories that have more than 3 products with unit prices between $10 and $30. The HAVING clause ensures only categories with more than 3 products are included.*