

Ex 3.7

Prepared by: Nikita Kasrekar

1. Write a query to find the top 10 countries for Rockbuster in terms of customer numbers.

(Tip: you'll have to use GROUP BY and ORDER BY, both of which follow the join.)

The screenshot shows a pgAdmin interface with a query editor and a data output window. The query editor contains the following SQL code:

```
SELECT country,
       COUNT(customer_id) AS count_of_customer
  FROM customer AS customer
  INNER JOIN address AS add ON customer.address_id = add.address_id
  INNER JOIN city AS cit ON add.city_id = cit.city_id
  INNER JOIN country AS coun ON cit.country_id = coun.country_id
 GROUP BY country
 ORDER BY COUNT (customer_id) DESC
 LIMIT 10;
```

The data output window displays the results of the query:

country	count_of_customer
India	60
China	53
United States	36
Japan	31
Mexico	30
Brazil	28
Russian Federation	28
Philippines	20
Turkey	15
Indonesia	14

Showing rows: 1 to 11

To identify the top 10 countries by customer numbers, I followed the relational path between tables: customer → address → city → country. I used INNER JOINs to connect these tables because customer locations are stored indirectly through address and city. After linking the correct tables, I grouped the results by country to count how many customers belong to each country. Finally, I ordered the results from highest to lowest and limited the output to the top 10. This approach ensures that I accurately capture customer distribution across countries while keeping the query efficient and easy to interpret.

2. Next, write a query to identify the top 10 cities that fall within the top 10 countries you identified in step 1. (Hint: the top 10 cities can be in any of the countries identified—you don't need to create a separate list for each country.)

The screenshot shows the pgAdmin 4 interface with a query editor and a results grid.

Query Editor:

```
SELECT city, country,
       COUNT(cust.customer_id) AS count_of_cust
  FROM customer AS cust
 INNER JOIN address AS add ON cust.address_id = add.address_id
 INNER JOIN city AS cit ON add.city_id = cit.city_id
 INNER JOIN country AS coun ON cit.country_id = coun.country_id
 WHERE country IN ('India', 'China', 'United States', 'Japan', 'Mexico', 'Brazil', 'Russia Federation', 'Philippines', 'Turkey')
 GROUP BY city, country
 ORDER BY count(customer_id) DESC
 LIMIT 10;
```

Data Output Grid:

city	country	count_of_cust
America	United States	2
Almaty	Kazakhstan	1
Buenos Aires	Argentina	1
Brasilia	Brazil	1
Delhi (India)	India	1
Guangzhou	China	1
Helsinki	Finland	1
Ho Chi Minh City	Vietnam	1
London	United Kingdom	1
Mexico City	Mexico	1
Sao Paulo	Brazil	1

Total rows: 10. Query complete 00:00:00.486.

I already knew the top 10 countries from the previous query, so I used them directly in the WHERE clause to filter the cities. I then joined the customer, address, city, and country tables to connect each customer to their city and country. After filtering to the top countries, I grouped the results by city and country and counted the number of customers in each city. Finally, I ordered the results in descending order to find the top 10 cities overall. This approach is straightforward and avoids the need for a separate subquery.

- 3. Now write a query to find the top 5 customers from the top 10 cities who've paid the highest total amounts to Rockbuster. The customer team would like to reward them for their loyalty!**

The screenshot shows a PostgreSQL client interface with a query editor and a data output table.

```

SELECT cust.customer_id, cust.first_name,
       cust.last_name, cou.country, cit.city,
       SUM(pay.amount) AS total_payment
  FROM customer AS cust
 INNER JOIN payment AS pay ON cust.customer_id = pay.customer_id
 INNER JOIN address AS add ON cust.address_id = add.address_id
 INNER JOIN city AS cit ON add.city_id = cit.city_id
 INNER JOIN country AS cou ON cit.country_id = cou.country_id
 WHERE cit.city IN ('Aurora', 'Atlixco', 'Celaya', 'Adant', 'Guadalajara', 'Korashik', 'Pingxiang', 'Siyuan', 'Celaya')
 GROUP BY cust.customer_id, cust.first_name, cust.last_name, cou.country, cit.city
 ORDER BY total_payment DESC
 LIMIT 5;
  
```

Data Output:

customer_id	first_name	last_name	country	city	total_payment
1	Bru	Perry	Mexico	Ahuaz	128.79
2	Ella	Gabriel	Turkey	Sivas	108.75
3	Georgi	Stanislav	Mexico	Celaya	102.76
4	Clinton	Hughes	United States	Arizona	98.76
5	Adam	Gooch	India	Adant	92.80

To find the top 5 customers with the highest total payments in the top 10 cities, I first used the list of top cities identified in the previous query. I joined the customer, payment, address, city, and country tables so that each customer could be linked to their city and country. Then I filtered the results to include only customers living in the top 10 cities and top 10 countries. I grouped the data by customer and summed up their total payments. Finally, I sorted the totals in descending order and selected the top 5 customers. This approach ensures that only the most valuable customers within the most important cities are included.