

MPivec_Exercise 3.7: Joining Tables

View Database Tables: `SELECT * FROM pg_catalog.pg_tables`

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

Query

Query History

```
1 SELECT D.country,  
2     COUNT(customer_id)  
3 FROM customer A  
4 INNER JOIN address B ON A.address_id = B.address_id  
5 INNER JOIN city C ON B.city_id = C.city_id  
6 INNER JOIN country D ON C.country_ID = D.country_id  
7 GROUP BY country  
8 ORDER BY COUNT(customer_id) DESC  
9 LIMIT 10
```

Data output

Messages

Notifications

	country character varying (50)	count bigint
1	India	60
2	China	53
3	United States	36
4	Japan	31
5	Mexico	30
6	Brazil	28
7	Russian Federation	28
8	Philippines	20
9	Turkey	15
10	Indonesia	14

- First, I reviewed the ERD created in exercise 3.2 to see which tables I needed to combine and through which keys. I saw that the connection path would require combining a few columns of data from multiple tables (4), so I now know this will require an INNER JOIN. After combining the tables, I used GROUP BY country because we are looking at which countries hold the most customers. Then, to show an output in a specified order (top 10 countries), I used the ORDER BY COUNT of customer_id in DESCending order to show the highest customer count at the top of the list. Lastly, adding LIMIT 10 gave us an output of only the 10 highest countries' customer counts.

○

2. Write a query to find the top 10 cities within the top 10 countries identified in step 1.

Query Query History

```

1 SELECT C.city,
2     D.country,
3     COUNT(customer_id) AS customer_count
4 FROM customer A
5 INNER JOIN address B ON A.address_id = B.address_id
6 INNER JOIN city C ON B.city_id = C.city_id
7 INNER JOIN country D ON C.country_id = D.country_id
8 WHERE D.country IN ('India','China','United States','Japan','Mexico','Brazil','Russia')
9 GROUP BY C.city, D.country
10 ORDER BY COUNT(customer_id) DESC
11 LIMIT 10

```

Data output Messages Notifications

	city character varying (50)	country character varying (50)	customer_count bigint
1	Aurora	United States	2
2	Atlixco	Mexico	1
3	Xintai	China	1
4	Adoni	India	1
5	Dhule (Dhulia)	India	1
6	Kurashiki	Japan	1
7	Pingxiang	China	1
8	Sivas	Turkey	1
9	Celaya	Mexico	1
10	So Leopoldo	Brazil	1

-
- From step 1, I understood that this query would use an INNER JOIN of the same tables, so it doesn't need to change much. I chose to select the city, country (for clarity), and customer count in the output. I used the WHERE clause to include the top 10 countries from step 1. Then grouped by city and country, in descending order of customer count. Lastly limited to the top 10.

- Write a query to find the top 5 customers in the top 10 cities who have paid the highest total amounts to Rockbuster. The customer team would like to reward them for their loyalty!

Query Query History

```

1 SELECT A.customer_id,
2     A.first_name,
3     A.last_name,
4     C.city,
5     D.country,
6     SUM(E.amount) AS total_paid
7 FROM customer A
8 INNER JOIN address B ON A.address_id = B.address_id
9 INNER JOIN city C ON B.city_id = C.city_id
10 INNER JOIN country D ON C.country_id = D.country_id
11 INNER JOIN payment E ON A.customer_id = E.customer_id
12 WHERE C.city IN ('Aurora','Atlixco','Xintai','Adoni','Dhule (Dhulia)','Kurashiki','Pingxiang','Sivas','Celaya','So
13 GROUP BY A.customer_id, A.first_name, A.last_name, C.city, D.country
14 ORDER BY total_paid DESC
15 LIMIT 5

```

Data output Messages Notifications

	customer_id integer	first_name character varying (45)	last_name character varying (45)	city character varying (50)	country character varying (50)	total_paid numeric
1	84	Sara	Perry	Atlixco	Mexico	128.70
2	518	Gabriel	Harder	Sivas	Turkey	108.75
3	587	Sergio	Stanfield	Celaya	Mexico	102.76
4	537	Clinton	Buford	Aurora	United States	98.76
5	367	Adam	Gooch	Adoni	India	97.80

○

