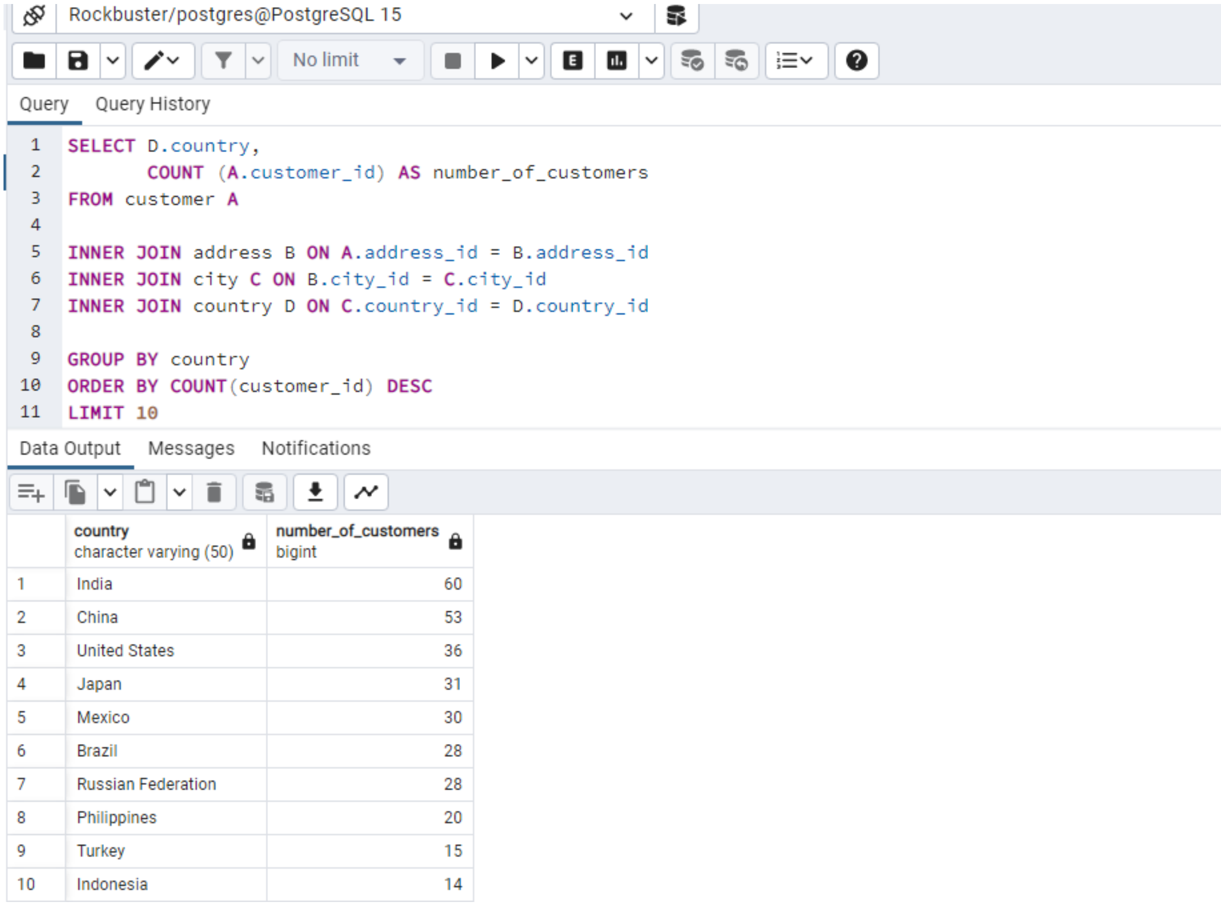
Joining Tables of Data - SQL Analysts

Ideal Kuci

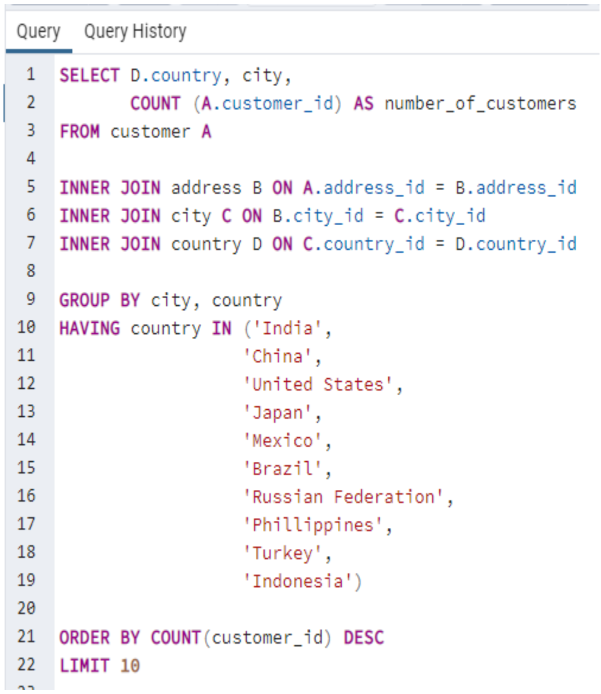
1. Top 10 countries for Rockbuster in terms of customer numbers:



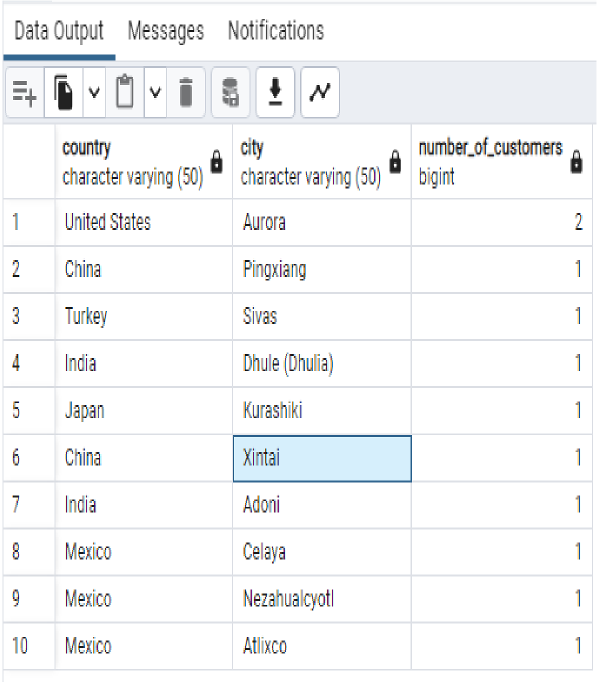
```sql
SELECT D.country,
       COUNT (A.customer_id) AS number_of_customers
FROM customer A

INNER JOIN address B ON A.address_id = B.address_id
INNER JOIN city C ON B.city_id = C.city_id
INNER JOIN country D ON C.country_id = D.country_id

GROUP BY country
ORDER BY COUNT(customer_id) DESC
LIMIT 10
```

	country character varying (50)	number_of_customers 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

2. Top 10 cities that fall within the top 10 countries I identified in step 1



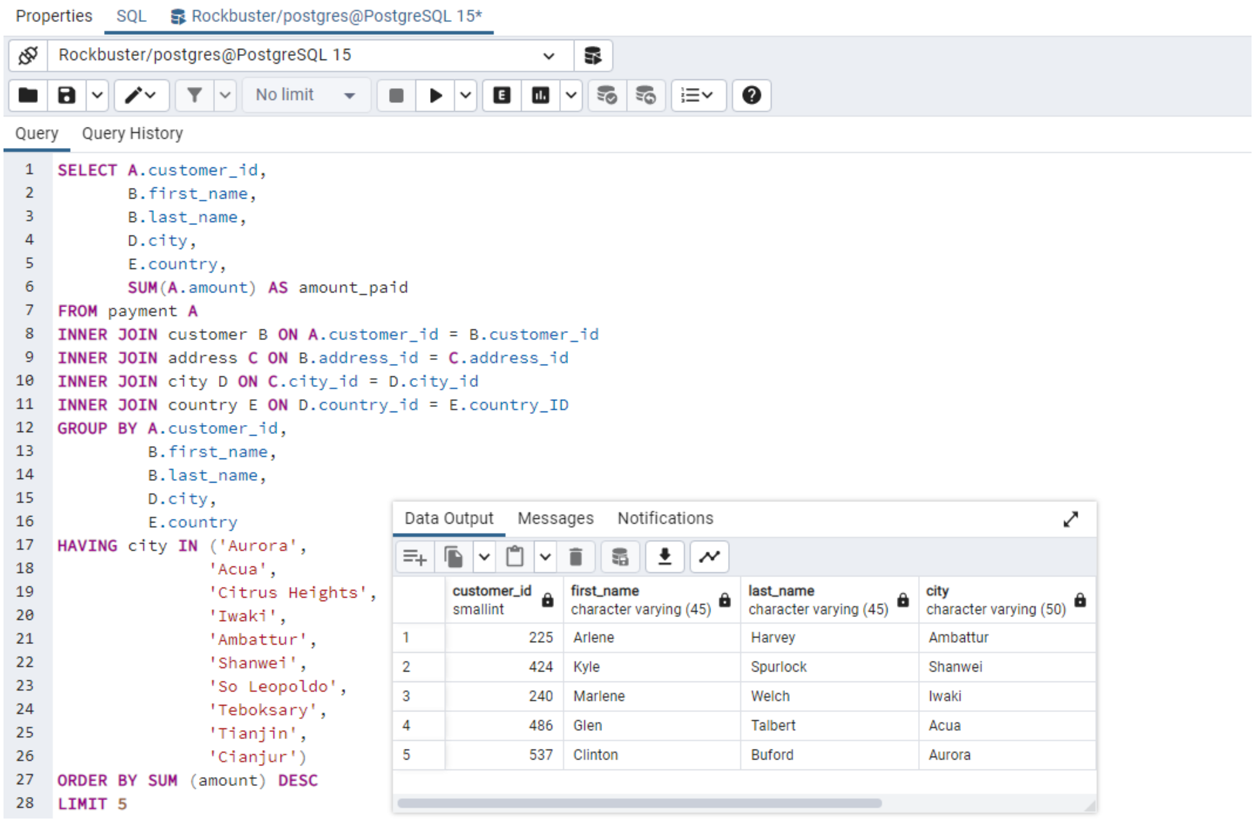
```sql
SELECT D.country, city,
       COUNT (A.customer_id) AS number_of_customers
FROM customer A

INNER JOIN address B ON A.address_id = B.address_id
INNER JOIN city C ON B.city_id = C.city_id
INNER JOIN country D ON C.country_id = D.country_id

GROUP BY city, country
HAVING country IN ('India',
                  'China',
                  'United States',
                  'Japan',
                  'Mexico',
                  'Brazil',
                  'Russian Federation',
                  'Phillippines',
                  'Turkey',
                  'Indonesia')

ORDER BY COUNT(customer_id) DESC
LIMIT 10
```



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

The step I took here is the same as I did in the previous one, only here I did use “HAVING” instead of “WHERE” because “WHERE” can’t be used to filter aggregate columns (after GROUP BY city, country).

3. Top 5 customers from the top 10 cities who've paid the highest total amounts to Rockbuster.



```sql
SELECT A.customer_id,
       B.first_name,
       B.last_name,
       D.city,
       E.country,
       SUM(A.amount) AS amount_paid
FROM payment A
INNER JOIN customer B ON A.customer_id = B.customer_id
INNER JOIN address C ON B.address_id = C.address_id
INNER JOIN city D ON C.city_id = D.city_id
INNER JOIN country E ON D.country_id = E.country_ID
GROUP BY A.customer_id,
         B.first_name,
         B.last_name,
         D.city,
         E.country
HAVING city IN ('Aurora',
               'Acua',
               'Citrus Heights',
               'Iwaki',
               'Ambattur',
               'Shanwei',
               'So Leopoldo',
               'Teboksary',
               'Tianjin',
               'Cianjur')
ORDER BY SUM (amount) DESC
LIMIT 5
```

	customer_id smallint	first_name character varying (45)	last_name character varying (45)	city character varying (50)
1	225	Arlene	Harvey	Ambattur
2	424	Kyle	Spurlock	Shanwei
3	240	Marlene	Welch	Iwaki
4	486	Glen	Talbert	Acua
5	537	Clinton	Buford	Aurora

Here I used a similar query to the one above, except here I added an extra table and after the "GROUP BY" in the "HAVING" part I used cities instead of countries - I used the results from the previous query. I used SUM to accumulate the amount paid.