

### 3.7: JOINING TABLES OF DATA

#### STEP 1

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

- We were looking for specific data from both tables (customer, country), therefore it was appropriate to use INNER JOIN.

From looking at ERD we knew that access to 4 tables was needed (customer, address, city, country). The tables were connected by common keys that can be seen in the query.

The results were grouped by *country* column and ordered in descending order.  
A LIMIT 10 returns only top 10 countries.

```
158 SELECT D.country,
159        COUNT(A.customer_id) AS number_of_customers
160 FROM customer A
161 INNER JOIN address B ON A.address_id = B.address_id
162 INNER JOIN city C ON B.city_id = C.city_id
163 INNER JOIN country D ON C.country_id = D.country_id
164 GROUP BY country
165 ORDER BY number_of_customers DESC
166 LIMIT 10
167
```

	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

Total rows: 10 of 10    Query complete 00:00:00.109

#### STEP 2

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

- I used the query from step 1 and added city as a column I am interested in SELECT.
- To specify which cities in which country I am looking for I added WHERE country IN ('India', 'China', 'United States', 'Japan', 'Mexico', 'Brazil', 'Russian Federation').

```
158 SELECT C.city,
159        D.country,
160        COUNT(A.customer_id) AS number_of_customers
161 FROM customer A
162 INNER JOIN address B ON A.address_id = B.address_id
163 INNER JOIN city C ON B.city_id = C.city_id
164 INNER JOIN country D ON C.country_id = D.country_id
165 WHERE country IN ('India', 'China', 'United States', 'Japan', 'Mexico', 'Brazil', 'Russian Federation')
166 GROUP BY city, country
167 ORDER BY number_of_customers DESC
168 LIMIT 10
169
```

	city character varying (50)	country character varying (50)	number_of_customers bigint
1	Aurora	United States	2
2	Acua	Mexico	1
3	Citrus Heights	United States	1
4	Iwaki	Japan	1
5	Ambattur	India	1
6	Shanwei	China	1
7	So Leopoldo	Brazil	1
8	Teboksary	Russian Federation	1
9	Tianjin	China	1

Total rows: 10 of 10    Query complete 00:00:02.896

### STEP 3

*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!*

- Because we are also looking for Total Amount paid by customers, a payment table needed to be added as well. The lettering of the tables has been shifted:

*A payment*

*B customer*

*C address*

*D city*

*E country*

- After JOIN we added WHERE city IN to specify which cities, we want to pull the customer data from.
- This time we are only searching for TOP 5 customers, therefore LIMIT 5.

The screenshot shows a SQL query editor with a query window and a results window. The query is as follows:

```
175 SELECT SUM(A.amount) AS total_amount_paid,
176         B.customer_id,
177         B.first_name,
178         B.last_name,
179         D.city,
180         E.country
181 FROM payment A
182 INNER JOIN customer B ON A.customer_id = B.customer_id
183 INNER JOIN address C ON B.address_id = C.address_id
184 INNER JOIN city D ON C.city_id = D.city_id
185 INNER JOIN country E ON D.country_id = E.country_id
186 WHERE city IN ('Aurora', 'Acua', 'Citrus Heights', 'Iwaki', 'Ambattur', 'Shanwei',
187               'So Leopoldo', 'Teboksary', 'Tianjin', 'Cianjur')
188 GROUP BY B.customer_id,
189          first_name,
190          last_name,
191          city,
192          country
193 ORDER BY total_amount_paid DESC
194 LIMIT 5
```

The results window shows the following data:

	total_amount_paid numeric	customer_id integer	first_name character varying (45)	last_name character varying (45)	city character varying (50)	country character varying (50)
1	111.76	225	Arlene	Harvey	Ambattur	India
2	109.71	424	Kyle	Spurlock	Shanwei	China