Katie Goyal
3.7 Joining Tables of Data

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

Query Query History

```
1  SELECT C.customer_id, A.address_id, CI.city_id, CO.* FROM rental R
2  JOIN customer C ON C.customer_id = R.customer_id
3  JOIN address A ON A.address_id = C.address_id
4  JOIN city CI ON CI.city_id = A.city_id
5  JOIN country CO ON CO.country_id = CI.country_id;
6  SELECT CO.country, COUNT(DISTINCT C.customer_id) AS number_of_customers, COUNT(R.customer_id) AS rental_frequency
7  FROM rental R
8  JOIN customer C ON C.customer_id = R.customer_id
9  JOIN store S ON C.store_id = S.store_id
10 JOIN address A ON A.address_id = C.address_id
11 JOIN city CI ON CI.city_id = A.city_id
12 JOIN country CO ON CO.country_id = CI.country_id
13 GROUP BY CO.country
14 ORDER BY number_of_customers DESC
15 LIMIT 10;
16
```

Data Output Messages Notifications

	country character varying (50)	number_of_customers bigint	rental_frequency bigint
1	India	60	1572
2	China	53	1426
3	United States	36	968
4	Japan	31	825
5	Mexico	30	796
6	Brazil	28	748
7	Russian Federation	28	713
8	Philippines	20	568
9	Turkey	15	388
10	Indonesia	14	367

- In this query, I created connections from individual rentals to their broader geographical origins, detailing each customer's location from their address up to their country to provide a full picture for analysis or customer insights. Then I highlighted the top countries where our rental service is most popular, by counting unique customers and total rentals, which informs us about our market size and demand. By focusing on the top ten countries, we distill our data into a clear picture of our most active markets, guiding potential business strategies and decisions.

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.)

Query | Query History

```
WITH TopCountries AS (
    SELECT CO.country_id
    FROM rental R
    JOIN customer C ON C.customer_id = R.customer_id
    JOIN store S ON C.store_id = S.store_id
    JOIN address A ON A.address_id = C.address_id
    JOIN city CI ON CI.city_id = A.city_id
    JOIN country CO ON CO.country_id = CI.country_id
    GROUP BY CO.country_id
    ORDER BY COUNT(DISTINCT C.customer_id) DESC
    LIMIT 10
)
SELECT CI.city, COUNT(DISTINCT C.customer_id) AS number_of_customers
FROM rental R
JOIN customer C ON C.customer_id = R.customer_id
JOIN address A ON A.address_id = C.address_id
JOIN city CI ON CI.city_id = A.city_id
JOIN TopCountries TC ON TC.country_id = CI.country_id
GROUP BY CI.city
ORDER BY number_of_customers DESC
LIMIT 10;
```

Data Output | Messages | Notifications

	city character varying (50)	number_of_customers bigint
1	Aurora	2
2	Adoni	1
3	Allende	1
4	Akishima	1
5	Akron	1
6	Allappuzha (Alleppey)	1
7	Ahmadnagar	1
8	Alvorada	1
9	Ambattur	1
10	Adana	1

- To identify the top cities within the previously determined top countries, I utilized a Common Table Expression (CTE) to first isolate the countries of interest. This approach simplifies the queries. By establishing the top countries in the CTE, it efficiently filters and focuses on the main query on cities within these countries. The main query focuses on the largest populations within these countries that are renting the most videos. This is how we came up with this list.

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!**

Query | Query History

```
    JOIN address A ON A.address_id = C.address_id
    JOIN city CI ON CI.city_id = A.city_id
    JOIN country CO ON CO.country_id = CI.country_id
    GROUP BY CI.city_id
    ORDER BY COUNT(DISTINCT C.customer_id) DESC
    LIMIT 10
),
TopCustomers AS (
    SELECT P.customer_id, SUM(P.amount) AS total_amount
    FROM payment P
    JOIN customer C ON C.customer_id = P.customer_id
    JOIN address A ON A.address_id = C.address_id
    JOIN TopCities TC ON TC.city_id = A.city_id
    GROUP BY P.customer_id
    ORDER BY SUM(P.amount) DESC
    LIMIT 5
)
SELECT C.customer_id, C.first_name, C.last_name, TC.total_amount
FROM TopCustomers TC
JOIN customer C ON C.customer_id = TC.customer_id
ORDER BY TC.total_amount DESC;
```

Data Output | Messages | Notifications

	customer_id [PK] integer	first_name character varying (45)	last_name character varying (45)	total_amount numeric
1	210	Ella	Oliver	134.70
2	512	Cecil	Vines	115.74
3	574	Julian	Vest	108.73
4	486	Glen	Talbert	100.77
5	297	Sherri	Rhodes	99.74

- The resulting list shows the top 5 customers from the top 10 cities who have contributed the highest total revenue.