Answers 3.7

1) Top 10 Countries by Customer Count

SQL Query:

```
SELECT
   co.country,
   COUNT(DISTINCT cu.customer_id) AS customer_count
FROM customer cu
JOIN address a   ON cu.address_id = a.address_id
JOIN city ci    ON a.city_id = ci.city_id
JOIN country co  ON ci.country_id = co.country_id
GROUP BY co.country
ORDER BY customer_count DESC
LIMIT 10;
```

Output:

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

Explanation (thought process):

I started from the customer table because the question is about the number of customers. To connect each customer to a country, I joined customer → address (address_id), address → city (city_id), and city → country (country_id). Then I grouped the results by country and

counted distinct customer_id to avoid double-counting. Finally, I ordered the counts from highest to lowest and limited the results to the top 10.

2) Top 10 Cities within the Top 10 Countries

SQL Query:

```
WITH top_countries AS (
  SELECT
    co.country_id,
    co.country,
    COUNT(DISTINCT cu.customer_id) AS customer_count
  FROM customer cu
  JOIN address a   ON cu.address_id = a.address_id
  JOIN city ci     ON a.city_id = ci.city_id
  JOIN country co  ON ci.country_id = co.country_id
  GROUP BY co.country_id, co.country
  ORDER BY customer_count DESC
  LIMIT 10
)
SELECT
  ci.city,
  tc.country,
  COUNT(DISTINCT cu.customer_id) AS customer_count
FROM customer cu
JOIN address a   ON cu.address_id = a.address_id
JOIN city ci     ON a.city_id = ci.city_id
JOIN top_countries tc ON ci.country_id = tc.country_id
GROUP BY ci.city, tc.country
ORDER BY customer_count DESC, ci.city
LIMIT 10;
```

Output:

city	country	customer_count
Aurora	United States	2
Probolinggo	Indonesia	1
Poos de Caldas	Brazil	1
Pontianak	Indonesia	1
Po	Brazil	1
Pjatigorsk	Russian Federation	1

Pingxiang	China	1
Peoria	United States	1
Pemalang	Indonesia	1
Patiala	India	1

Explanation (thought process):

I reused the same join path (customer → address → city → country) and first created a CTE (top_countries) to capture only the top 10 countries from step 1. Then I joined city to that CTE, grouped by city and country, counted distinct customers, and ordered the results to return the top 10 cities overall (across those countries).

3) Top 5 Customers (from Top 10 Cities) by Total Amount Paid

SQL Query:

```
WITH top_countries AS (
  SELECT
    co.country_id,
    COUNT(DISTINCT cu.customer_id) AS customer_count
  FROM customer cu
  JOIN address a   ON cu.address_id = a.address_id
  JOIN city ci     ON a.city_id = ci.city_id
  JOIN country co  ON ci.country_id = co.country_id
  GROUP BY co.country_id
  ORDER BY customer_count DESC
  LIMIT 10
),
top_cities AS (
  SELECT
    ci.city_id,
    ci.city,
    co.country,
    COUNT(DISTINCT cu.customer_id) AS customer_count
  FROM customer cu
  JOIN address a   ON cu.address_id = a.address_id
  JOIN city ci     ON a.city_id = ci.city_id
  JOIN country co  ON ci.country_id = co.country_id
  JOIN top_countries tc ON co.country_id = tc.country_id
  GROUP BY ci.city_id, ci.city, co.country
  ORDER BY customer_count DESC, ci.city
  LIMIT 10
```

```
)
SELECT
    cu.customer_id,
    cu.first_name,
    cu.last_name,
    co.country,
    ci.city,
    SUM(p.amount) AS total_amount_paid
FROM payment p
JOIN customer cu ON p.customer_id = cu.customer_id
JOIN address a   ON cu.address_id = a.address_id
JOIN city ci     ON a.city_id = ci.city_id
JOIN country co  ON ci.country_id = co.country_id
JOIN top_cities tc ON ci.city_id = tc.city_id
GROUP BY cu.customer_id, cu.first_name, cu.last_name, co.country, ci.city
ORDER BY total_amount_paid DESC
LIMIT 5;
```

Output:

customer_id	first_name	last_name	country	city	total_amount_paid
372	Steve	Mackenzie	Indonesia	Probolinggo	152.68
506	Leslie	Seward	Indonesia	Pontianak	123.72
487	Hector	Poindexter	Russian Federation	Pjatigorsk	119.74
254	Maxine	Silva	Brazil	Poos de Caldas	107.70
116	Victoria	Gibson	Indonesia	Pemalang	106.74

Notes:

This query first restricts the geography using CTEs for the top 10 countries and top 10 cities. Then it joins payment to customer and sums payment amounts per customer. The final output is ordered by total_amount_paid to identify the top 5 highest-paying customers from those top cities.