

Ex. 3.8 Performing Subqueries

Step 1: Find the average amount paid by the top 5 customers.

9	SELECT ROUND(AVG(total_amount_paid),2) AS Average
10	FROM
11	(SELECT
12	A.customer_id,
13	B.first_name,
14	B.last_name,
15	D.city,
16	E.country,
17	SUM(amount) AS Total_amount_paid FROM payment A
18	INNER JOIN customer B ON A.customer_id = B.customer_id
19	INNER JOIN address C ON B.address_id = C.address_id
20	INNER JOIN city D ON C.city_id = D.city_id
21	INNER JOIN country E ON D.country_id = E.country_id
22	WHERE country IN ('India','China','United States','Japan','Mexico','Brazil',
23	'Russian Federation','Philippines','Turkey','Indonesia')
24	AND city IN ('Aurora','Atlixco','Xintai','Adoni','Dhule (Dhulia)',
25	'Kurashiki','Pingxiang','Sivas','Celaya','So Leopoldo')
26	GROUP BY A.customer_id,B.first_name,B.last_name,D.city,E.country
27	ORDER BY Total_amount_paid DESC LIMIT 5) AS Average

Data Output	Explain	Messages	Notifications
average numeric			
1	107.35		

```
SELECT ROUND(AVG(total_amount_paid),2) AS Average
FROM
(SELECT
A.customer_id,
B.first_name,
B.last_name,
D.city,
E.country,
SUM(amount) AS Total_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
WHERE country IN ('India','China','United States','Japan','Mexico','Brazil',
'Russian Federation','Philippines','Turkey','Indonesia')
AND city IN ('Aurora','Atlixco','Xintai','Adoni','Dhule (Dhulia)',
'Kurashiki','Pingxiang','Sivas','Celaya','So Leopoldo')
GROUP BY A.customer_id,B.first_name,B.last_name,D.city,E.country
ORDER BY Total_amount_paid DESC LIMIT 5) AS Average
```

Step 2: Find out how many of the top 5 customers are based within each country.

```

1  SELECT
2  D.country,
3  COUNT(DISTINCT A.customer_id) AS all_customer_count,
4  COUNT(DISTINCT top_5_customers) AS top_customer_count
5  FROM
6  customer A
7  INNER JOIN address B ON A.address_id = B.address_id
8  INNER JOIN city C ON B.city_id = C.city_id
9  INNER JOIN country D ON C.country_id = D.country_id
10 LEFT JOIN
11 (SELECT
12  A.customer_id,
13  B.first_name,
14  B.last_name,
15  D.city,
16  E.country,
17  SUM(amount) AS Total_amount_paid
18  FROM payment A
19  INNER JOIN customer B ON A.customer_id = B.customer_id
20  INNER JOIN address C ON B.address_id = C.address_id
21  INNER JOIN city D ON C.city_id = D.city_id
22  INNER JOIN country E ON D.country_id = E.country_id
23  WHERE country IN ('India','China','United States','Japan','Mexico','Brazil',
24  'Russian Federation','Philippines','Turkey','Indonesia')
25  AND city IN ('Aurora','Atlixco','Xintai','Adoni','Dhule (Dhulia)',
26  'Kurashiki','Pingxiang','Sivas','Celaya','So Leopoldo'))
27  GROUP BY A.customer_id,B.first_name,B.last_name,D.city,E.country
28  ORDER BY Total_amount_paid DESC LIMIT 5) AS top_5_customers
29  ON D.country = top_5_customers.country
30  GROUP BY D.country
31  ORDER BY all_customer_count DESC
32
33
34
35

```

Data Output Explain Messages Notifications

	country character varying (50)	all_customer_count bigint	top_customer_count bigint	
1	India	60	1	
2	China	53	0	
3	United States	36	1	
4	Japan	31	0	
5	Mexico	30	2	
6	Brazil	28	0	
7	Russian Federation	28	0	
8	Philippines	20	0	
9	Turkey	15	1	

```

SELECT
D.country,
COUNT(DISTINCT A.customer_id) AS all_customer_count,
COUNT(DISTINCT top_5_customers) AS top_customer_count
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
LEFT JOIN
(SELECT
A.customer_id,
B.first_name,
B.last_name,
D.city,
E.country,
SUM(amount) AS Total_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
WHERE country IN ('India','China','United States','Japan','Mexico','Brazil',
'Russian Federation','Philippines','Turkey','Indonesia')
AND city IN ('Aurora','Atlixco','Xintai','Adoni','Dhule (Dhulia)',
'Kurashiki','Pingxiang','Sivas','Celaya','So Leopoldo')
GROUP BY A.customer_id,B.first_name,B.last_name,D.city,E.country
ORDER BY Total_amount_paid DESC LIMIT 5) AS top_5_customers
ON D.country = top_5_customers.country
GROUP BY D.country
ORDER BY all_customer_count DESC

```

Step 3:

Do you think steps 1 and 2 could be done without using subqueries?

- Yes, they could be run as two separate stand-alone queries the results of which could be added in excel. This would however be much less efficient.

When do you think subqueries are useful?

- To add a new column to a main query in order to compare different data points e.g., an amount to a total average. This would be done by adding a subquery to the 'Select' clause.
- To filter the results of a main query based on the results of a subquery. This is useful when subquery results will constantly be updated / change e.g., top 5 records. This would be done by adding a subquery to the 'Where' clause.
- To analyse the results of a complex query by adding a subquery to the 'From' clause.