

Task 3.9

Step 1:

WITH top_5_paid_cte (customer_id, first_name, last_name, country, city, total_amount_paid)
AS

```
(SELECT A.customer_id, B.first_name, B.last_name, E.country, D.city,  
       SUM (A.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 D.city IN ('Aurora','Tokat', 'Tarsus', 'Atlixco', 'Emeishan', 'Pontianak', 'Shimoga',  
'Aparecida de Goiania', 'Zalantun', 'Taguig')
```

```
GROUP BY A.customer_id, B.first_name, B.last_name, E.country, D.city, A.amount
```

```
ORDER BY total_amount_paid DESC
```







```
LIMIT 5)
```

```
SELECT customer_id, first_name, last_name, country, city, AVG(total_amount_paid) AS  
average_amount_paid
```

```
FROM top_5_paid_cte
```

```
GROUP BY customer_id, first_name, last_name, country, city
```

```
ORDER BY average_amount_paid DESC
```

Data Output		Explain	Messages	Notifications		
	 customer_id smallint	 first_name character varying (45)	 last_name character varying (45)	 country character varying (50)	 city character varying (50)	 average_amount_paid numeric
1	72	Theresa	Watson	Philippines	Taguig	44.9100000000000000
2	389	Alan	Kahn	China	Emeishan	39.9200000000000000
3	537	Clinton	Buford	United States	Aurora	39.9200000000000000
4	566	Casey	Mena	Turkey	Tokat	39.9200000000000000
5	93	Phyllis	Foster	China	Zalantun	34.9300000000000000

WITH top_countries_cte AS

```
(SELECT A.customer_id, B.first_name, B.last_name, E.country_ID, D.city,  
      SUM (A.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 D.city IN ('Aurora','Tokat', 'Tarsus', 'Atlixco', 'Emeishan', 'Pontianak', 'Shimoga',  
'Aparecida de Goiania', 'Zalantun', 'Taguig')
```

```
GROUP BY A.customer_id, B.first_name, B.last_name, E.country_ID, D.city, A.amount
```

```
ORDER BY total_amount_paid DESC
```

```
LIMIT 5)
```

```
SELECT country, COUNT (DISTINCT B.customer_id) AS all_customer_count, COUNT (DISTINCT  
F.customer_id) AS average_top_5_customers
```

```
FROM top_countries_cte F
```

```
INNER JOIN country E ON F.country_ID = E.country_id
```

```
INNER JOIN city D ON E.country_id = D.country_id
```

```
INNER JOIN address C ON D.city_id = C.city_id
```

```
INNER JOIN customer B ON C.address_id = B.address_id
```

```
GROUP BY country
```

```
ORDER BY average_top_5_customers DESC
```

	country character varying (50)	all_customer_count bigint	average_top_5_customers bigint
1	China	53	1
2	Mexico	30	1
3	Philippines	20	1
4	Turkey	15	1
5	United States	36	1

Essentially, for both CTEs I copied the inner query from 3.8. On the first query, I just had to add the average operation at the end of the code whether on the second query, I had to add an inner join that would pull the original data to the inner query.

Step 2:

I believe the CTEs perform better as they are faster to write, and faster to run.

Subqueries

```
"(cost=122.95..122.97 rows=5 width=38)" @ 45msec
```

```
"(cost=907.56..907.57 rows=5 width=25)" @ 44msec
```

CTEs

```
"Sort (cost=71.00..71.01 rows=5 width=65)" @44msec
```

```
"Sort (cost=99.78..99.85 rows=28 width=25)" @ 43msec
```

As we can see, CTEs are not only faster but more cost efficient than subqueries. It is all the more of a reason to do CTE's for its simplicity and its efficiency.

Step 3:

It was actually much easier to write a CTE than learning how to pencil down a subquery for the first time. The first challenge, however, precented in understanding that I needed a right join at the end of my CTE for the 2nd query. Although I had the CTE ready and all the numbers were running, they gave me the same answer for top 5 customers and all customer count. I had to realize that I needed to make another join to pull the original data in comparison to using the CTE on the table.