

## SQL Challenge 6

Marketing Analysis



<https://steeldata.org.uk/sql6.html>

# Introduction

You are a Marketing Analyst

The 'Sustainable Clothing Co.' have been running several marketing campaigns and have asked you to provide your insight into whether they have been successful or not. Analyse the following data and answer the questions to form your answer.

## Problem Statement

Key datasets for this case study:

- Sustainable\_clothing
- Transactions
- Marketing\_campaigns

## Tables Used

**sustainable\_clothing**

Product ID	Product Name	Category	Size	Price
1	Organic Cotton T-Shirt	Tops	S	\$29.99
2	Recycled Denim Jeans	Bottoms	M	\$79.99
3	Hemp Crop Top	Tops	L	\$24.99
4	Bamboo Lounge Pants	Bottoms	XS	\$49.99
5	Eco-Friendly Hoodie	Outerwear	XL	\$59.99
6	Linen Button-Down Shirt	Tops	M	\$39.99
7	Organic Cotton Dress	Dresses	S	\$69.99
8	Sustainable Swim Shorts	Swimwear	L	\$34.99
9	Recycled Polyester Jacket	Outerwear	XL	\$89.99
10	Bamboo Yoga Leggings	Activewear	XS	\$54.99
11	Hemp Overalls	Bottoms	M	\$74.99
12	Organic Cotton Sweater	Tops	L	\$49.99
13	Cork Sandals	Footwear	S	\$39.99
14	Recycled Nylon Backpack	Accessories	One Size	\$59.99
15	Organic Cotton Skirt	Bottoms	XS	\$34.99
16	Hemp Baseball Cap	Accessories	One Size	\$24.99
17	Upcycled Denim Jacket	Outerwear	M	\$79.99
18	Linen Jumpsuit	Dresses	L	\$69.99
19	Organic Cotton Socks	Accessories	M	\$9.99
20	Bamboo Bathrobe	Loungewear	XL	\$69.99

**transactions (first 10 shown)**

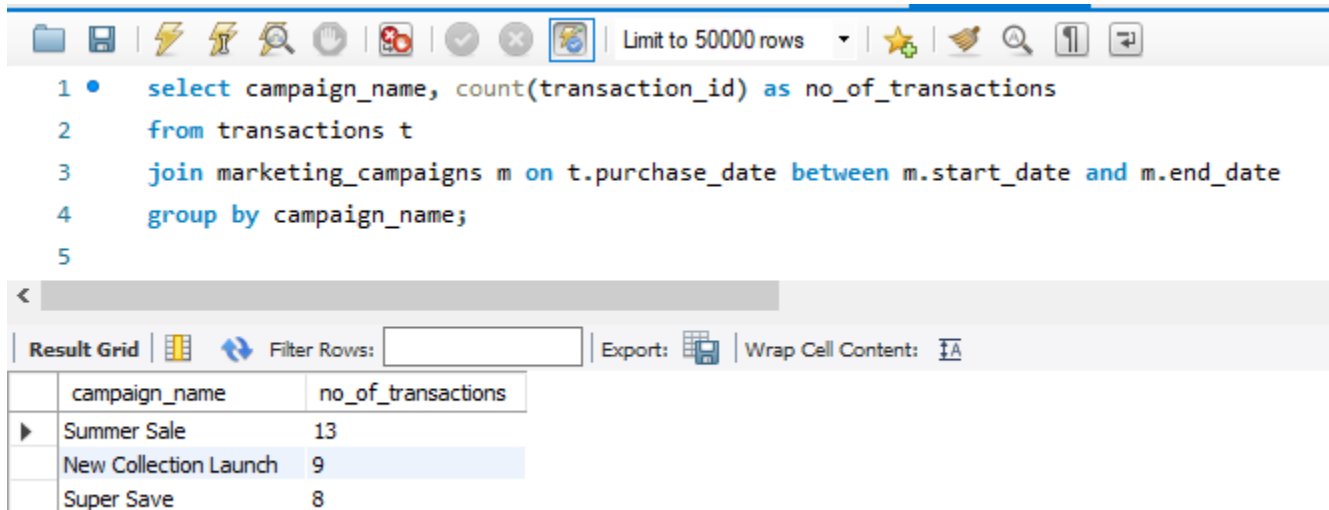
transaction_id	product_id	quantity	purchase_date
1	2	2	2023-06-02
1	14	1	2023-06-02
2	5	2	2023-06-05
3	2	1	2023-06-07
4	19	2	2023-06-10
5	2	1	2023-06-13
5	16	1	2023-06-13
6	10	2	2023-06-15
7	2	1	2023-06-18
8	4	1	2023-06-22
9	18	2	2023-06-26
10	2	1	2023-06-30
10	13	1	2023-06-30

**marketing\_campaigns**

campaign_id	campaign_name	product_id	start_date	end_date
1	Summer Sale	2	2023-06-01	2023-06-30
2	New Collection Launch	10	2023-07-15	2023-08-15
3	Super Save	7	2023-08-20	2023-09-15

## Case Study Questions

1. How many transactions were completed during each marketing campaign?



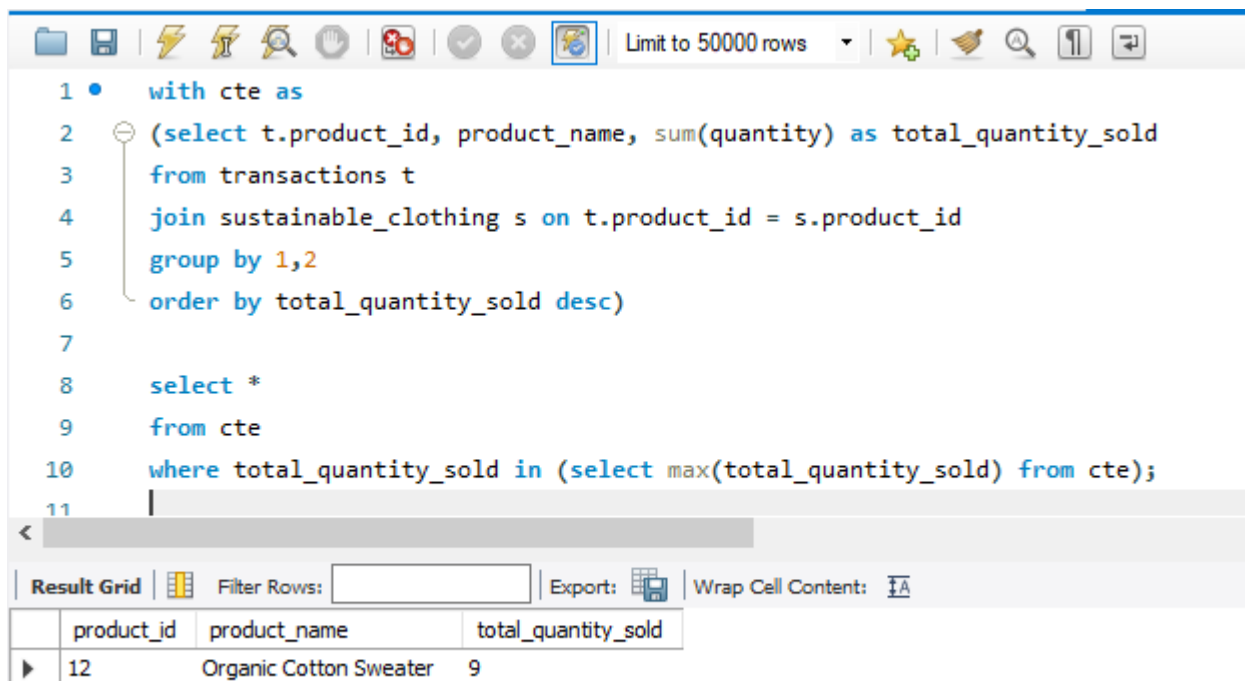
The screenshot shows a SQL query editor with a toolbar at the top. The query is as follows:

```
1 • select campaign_name, count(transaction_id) as no_of_transactions
2   from transactions t
3  join marketing_campaigns m on t.purchase_date between m.start_date and m.end_date
4  group by campaign_name;
5
```

Below the query editor is a "Result Grid" section. It includes a "Filter Rows" input field, an "Export" button, and a "Wrap Cell Content" checkbox. The result grid displays the following data:

campaign_name	no_of_transactions
Summer Sale	13
New Collection Launch	9
Super Save	8

2. Which product had the highest sales quantity?



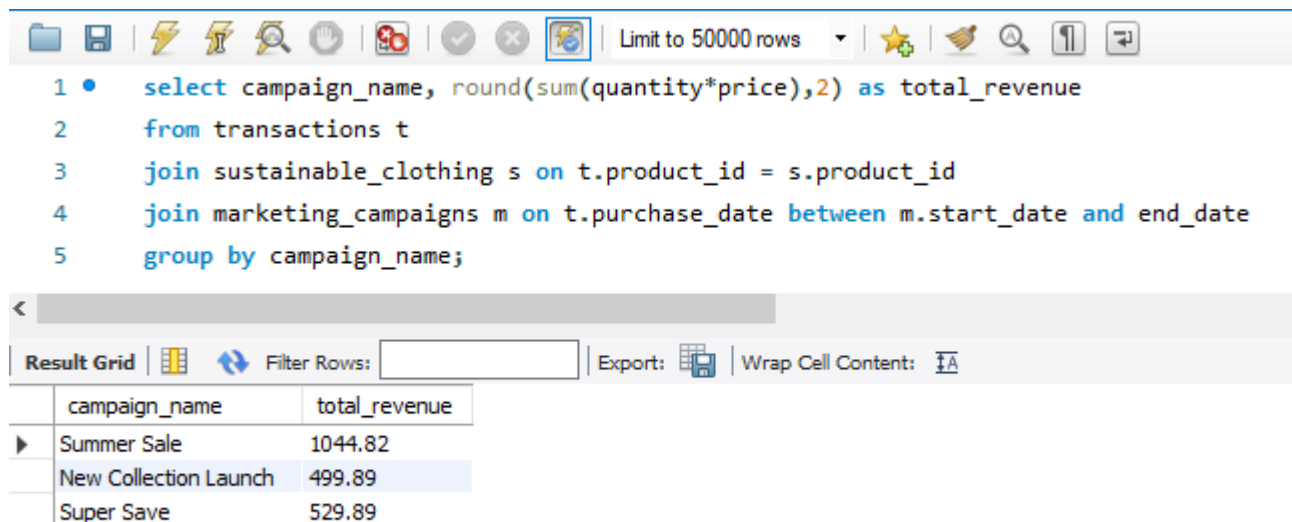
The screenshot shows a SQL query editor with a toolbar at the top. The query is as follows:

```
1 • with cte as
2   (select t.product_id, product_name, sum(quantity) as total_quantity_sold
3    from transactions t
4   join sustainable_clothing s on t.product_id = s.product_id
5   group by 1,2
6   order by total_quantity_sold desc)
7
8   select *
9   from cte
10  where total_quantity_sold in (select max(total_quantity_sold) from cte);
11
```

Below the query editor is a "Result Grid" section. It includes a "Filter Rows" input field, an "Export" button, and a "Wrap Cell Content" checkbox. The result grid displays the following data:

product_id	product_name	total_quantity_sold
12	Organic Cotton Sweater	9

3. What is the total revenue generated from each marketing campaign?

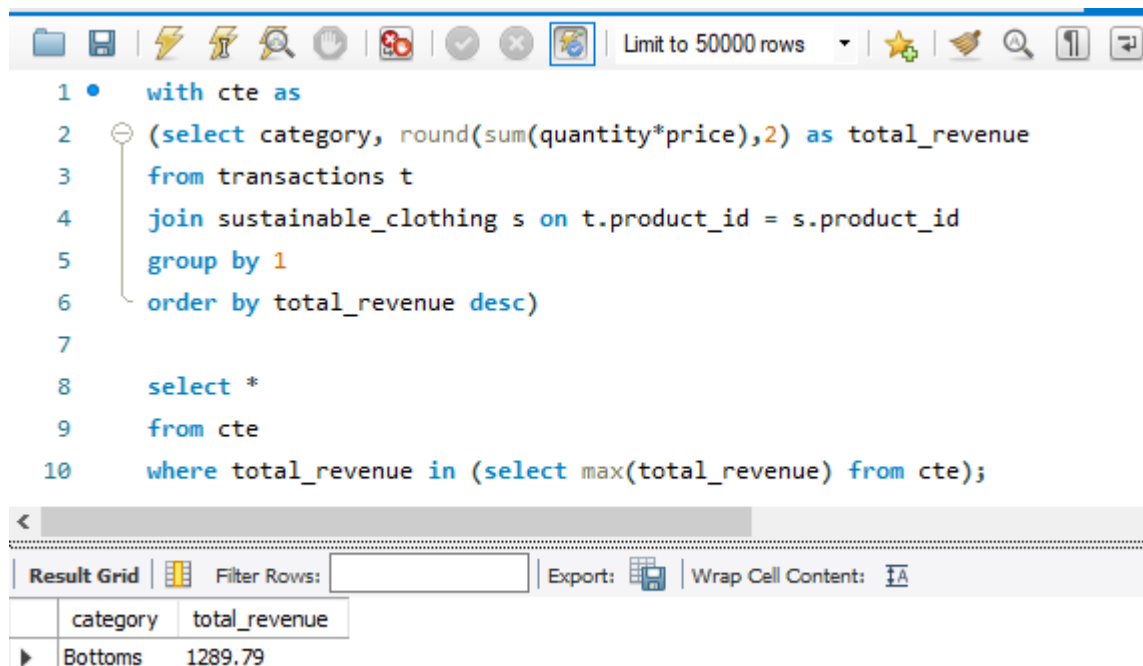


```
1 • select campaign_name, round(sum(quantity*price),2) as total_revenue
2   from transactions t
3  join sustainable_clothing s on t.product_id = s.product_id
4  join marketing_campaigns m on t.purchase_date between m.start_date and end_date
5  group by campaign_name;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: |

campaign_name	total_revenue
Summer Sale	1044.82
New Collection Launch	499.89
Super Save	529.89

4. What is the top-selling product category based on the total revenue generated?













```
1 • with cte as
2   (select category, round(sum(quantity*price),2) as total_revenue
3    from transactions t
4   join sustainable_clothing s on t.product_id = s.product_id
5   group by 1
6   order by total_revenue desc)
7
8   select *
9   from cte
10  where total_revenue in (select max(total_revenue) from cte);
```




Result Grid | Filter Rows: | Export: | Wrap Cell Content: |

category	total_revenue
Bottoms	1289.79

5. Which products had a higher quantity sold compared to the average quantity sold?






Limit to 50000 rows



```
1 • select t.product_id, product_name, quantity
2   from transactions t
3   join sustainable_clothing s on t.product_id = s.product_id
4   where quantity > (select avg(quantity) from transactions);
5
```

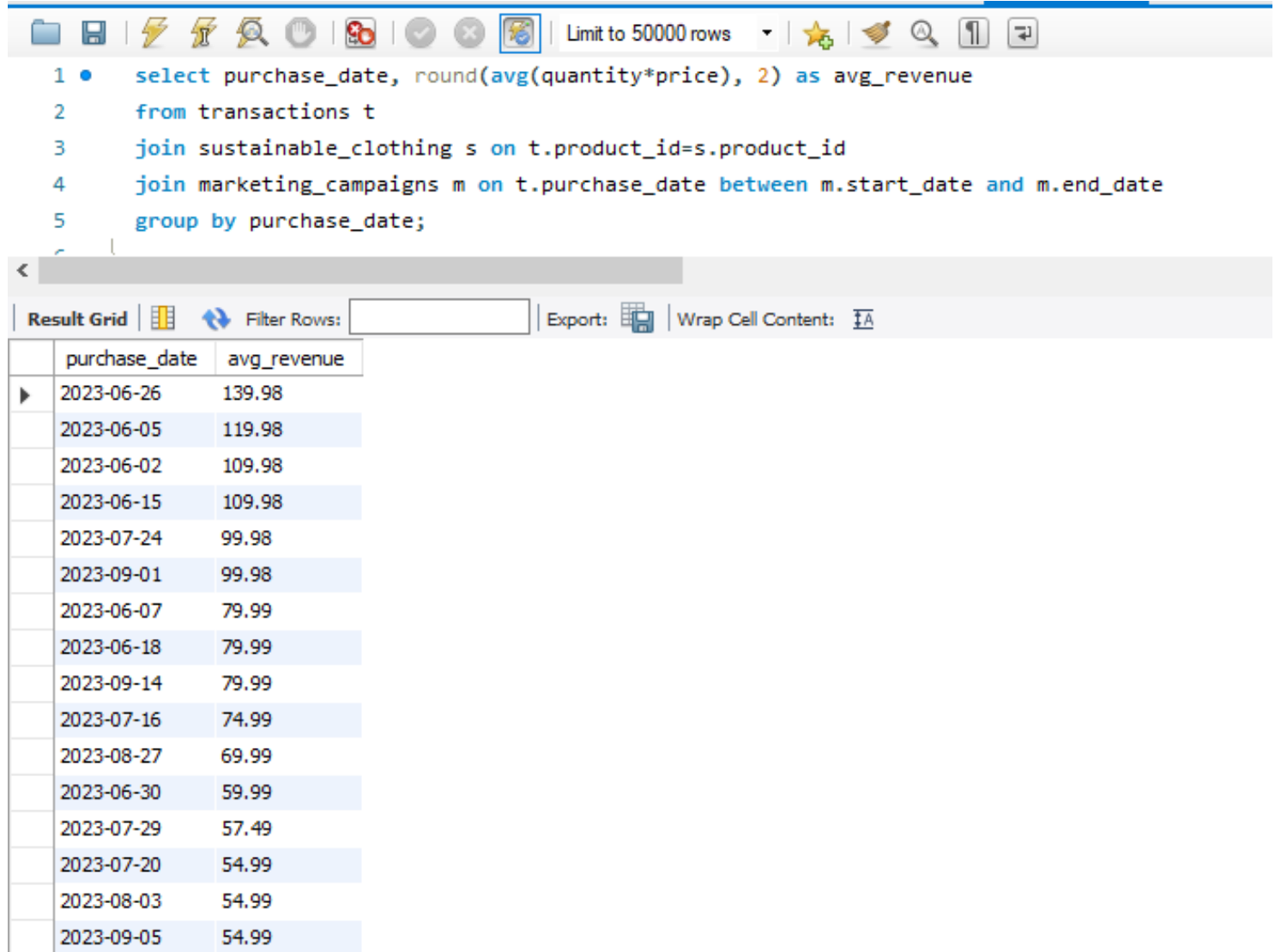
Result Grid

 Filter Rows:

Export:  Wrap Cell Content: 

	product_id	product_name	quantity
▶	2	Recycled Denim Jeans	2
	5	Eco-Friendly Hoodie	2
	19	Organic Cotton Socks	2
	10	Bamboo Yoga Leggings	2
	18	Linen Jumpsuit	2
	6	Linen Button-Down Shirt	2
	9	Recycled Polyester Jacket	2
	12	Organic Cotton Sweater	2
	19	Organic Cotton Socks	2
	16	Hemp Baseball Cap	2
	12	Organic Cotton Sweater	2
	15	Organic Cotton Skirt	2
	11	Hemp Overalls	2
	5	Eco-Friendly Hoodie	2
	12	Organic Cotton Sweater	2
	18	Linen Jumpsuit	2
	12	Organic Cotton Sweater	2
	4	Bamboo Lounge Pants	2
	8	Sustainable Swim Shorts	2

6. What is the average revenue generated per day during the marketing campaigns?



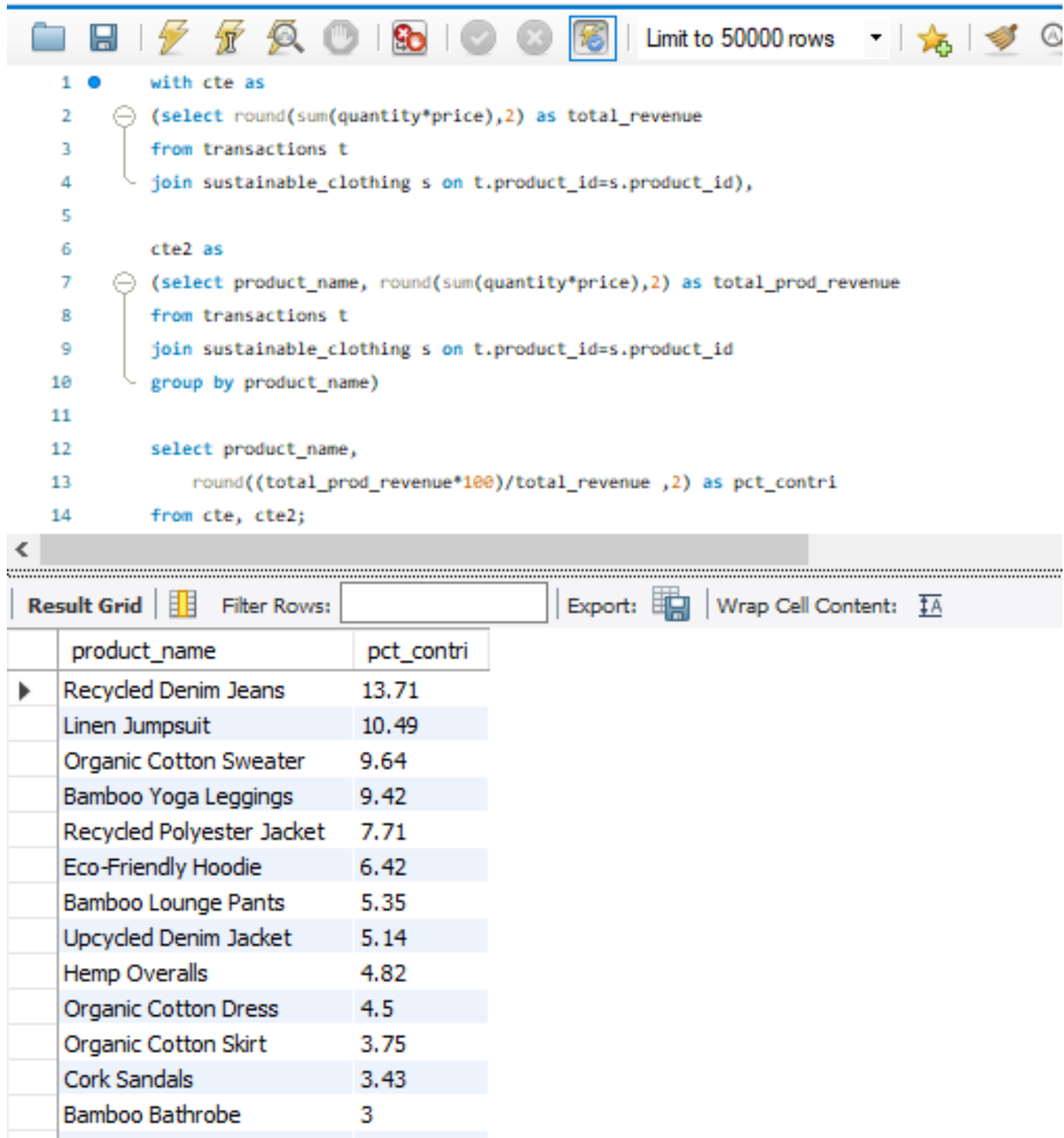
The screenshot shows a SQL query editor with a toolbar at the top. The query is as follows:

```
1 • select purchase_date, round(avg(quantity*price), 2) as avg_revenue
2   from transactions t
3   join sustainable_clothing s on t.product_id=s.product_id
4   join marketing_campaigns m on t.purchase_date between m.start_date and m.end_date
5   group by purchase_date;
```

Below the query editor is the "Result Grid" section, which includes a "Filter Rows:" input field, an "Export:" button, and a "Wrap Cell Content:" checkbox. The results are displayed in a table with two columns: "purchase\_date" and "avg\_revenue".

purchase_date	avg_revenue
2023-06-26	139.98
2023-06-05	119.98
2023-06-02	109.98
2023-06-15	109.98
2023-07-24	99.98
2023-09-01	99.98
2023-06-07	79.99
2023-06-18	79.99
2023-09-14	79.99
2023-07-16	74.99
2023-08-27	69.99
2023-06-30	59.99
2023-07-29	57.49
2023-07-20	54.99
2023-08-03	54.99
2023-09-05	54.99

7. What is the percentage contribution of each product to the total revenue?



The screenshot shows a SQL query editor with a toolbar at the top. The query is as follows:

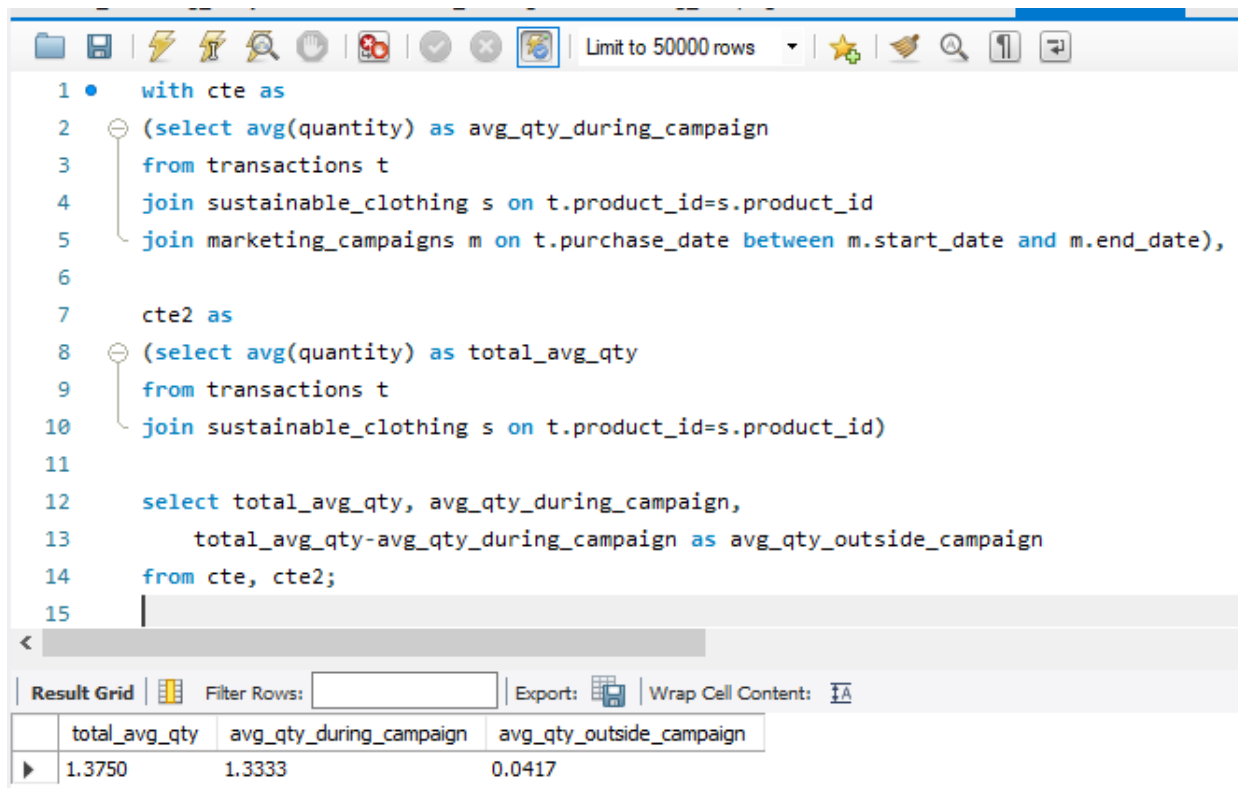
```
1 with cte as
2   (select round(sum(quantity*price),2) as total_revenue
3    from transactions t
4    join sustainable_clothing s on t.product_id=s.product_id),
5
6   cte2 as
7   (select product_name, round(sum(quantity*price),2) as total_prod_revenue
8    from transactions t
9    join sustainable_clothing s on t.product_id=s.product_id
10   group by product_name)
11
12   select product_name,
13          round((total_prod_revenue*100)/total_revenue ,2) as pct_contri
14   from cte, cte2;
```

Below the query editor is the "Result Grid" tab, which displays the following data:

	product_name	pct_contri
▶	Recycled Denim Jeans	13.71
	Linen Jumpsuit	10.49
	Organic Cotton Sweater	9.64
	Bamboo Yoga Leggings	9.42
	Recycled Polyester Jacket	7.71
	Eco-Friendly Hoodie	6.42
	Bamboo Lounge Pants	5.35
	Upcycled Denim Jacket	5.14
	Hemp Overalls	4.82
	Organic Cotton Dress	4.5
	Organic Cotton Skirt	3.75
	Cork Sandals	3.43
	Bamboo Bathrobe	3



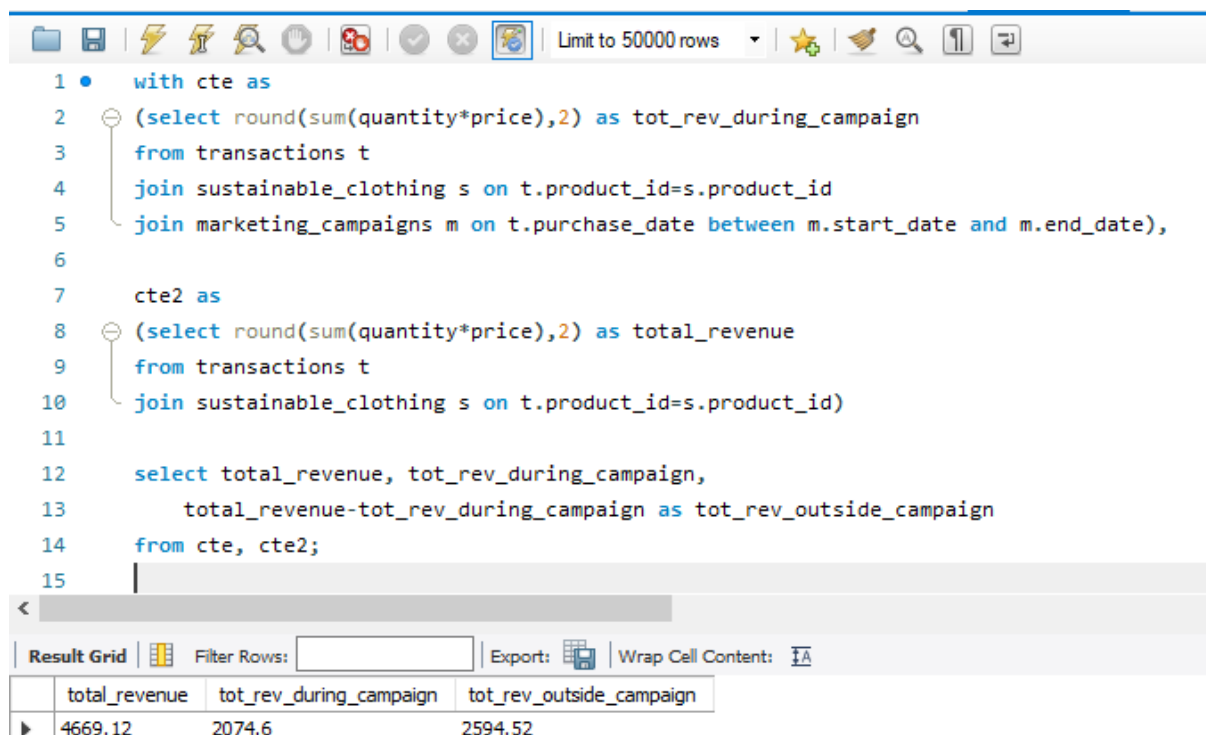
8. Compare the average quantity sold during marketing campaigns to outside the marketing campaigns



```
1 • with cte as
2   (select avg(quantity) as avg_qty_during_campaign
3     from transactions t
4     join sustainable_clothing s on t.product_id=s.product_id
5     join marketing_campaigns m on t.purchase_date between m.start_date and m.end_date),
6
7   cte2 as
8   (select avg(quantity) as total_avg_qty
9     from transactions t
10    join sustainable_clothing s on t.product_id=s.product_id)
11
12   select total_avg_qty, avg_qty_during_campaign,
13          total_avg_qty-avg_qty_during_campaign as avg_qty_outside_campaign
14   from cte, cte2;
```

total_avg_qty	avg_qty_during_campaign	avg_qty_outside_campaign
1.3750	1.3333	0.0417

9. Compare the revenue generated by products inside the marketing campaigns to outside the campaigns

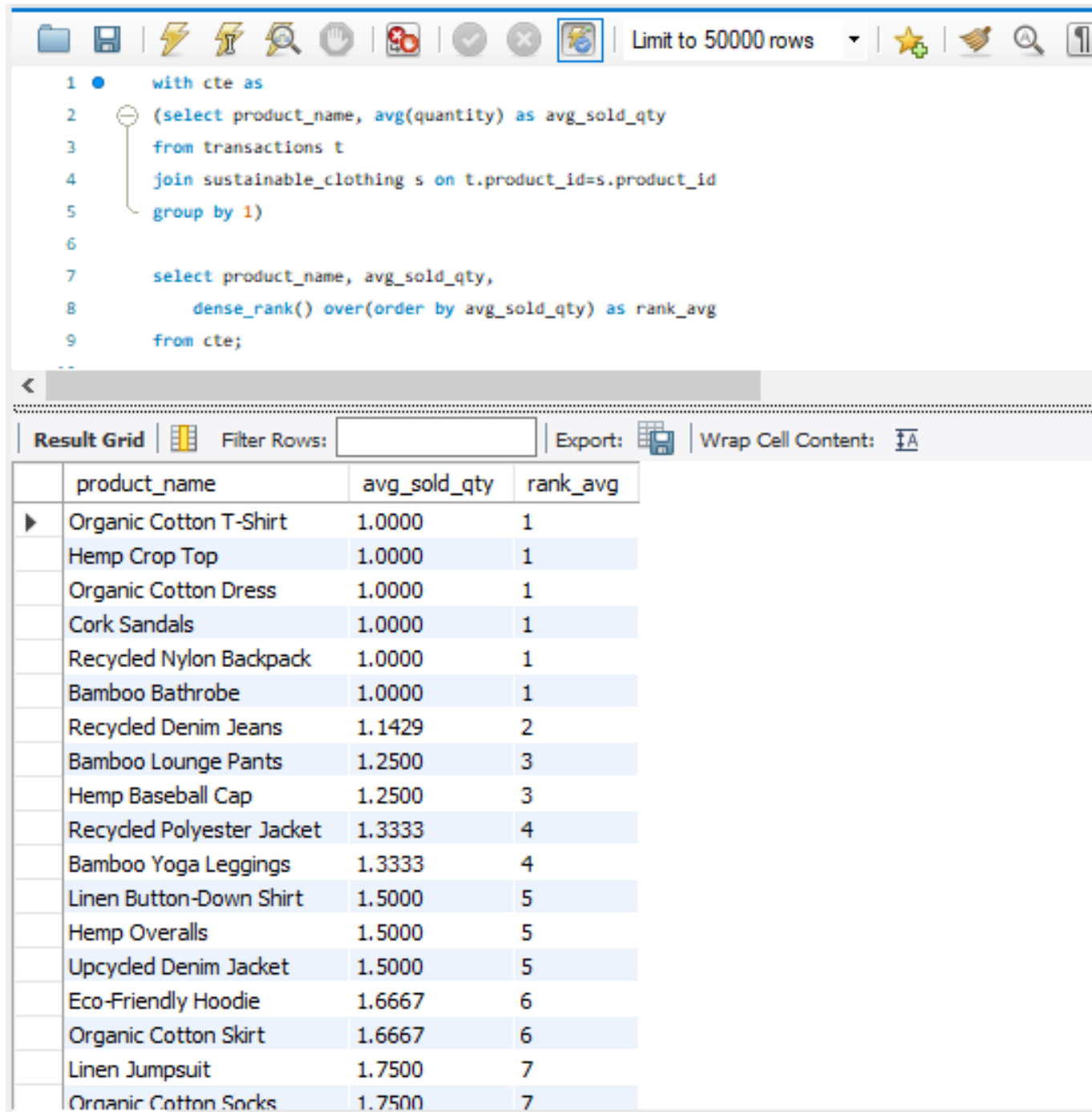


```
1 • with cte as
2   (select round(sum(quantity*price),2) as tot_rev_during_campaign
3     from transactions t
4     join sustainable_clothing s on t.product_id=s.product_id
5     join marketing_campaigns m on t.purchase_date between m.start_date and m.end_date),
6
7   cte2 as
8   (select round(sum(quantity*price),2) as total_revenue
9     from transactions t
10    join sustainable_clothing s on t.product_id=s.product_id)
11
12   select total_revenue, tot_rev_during_campaign,
13          total_revenue-tot_rev_during_campaign as tot_rev_outside_campaign
14   from cte, cte2;
```

total_revenue	tot_rev_during_campaign	tot_rev_outside_campaign
4669.12	2074.6	2594.52



## 10. Rank the products by their average daily quantity sold



The screenshot shows a SQL IDE interface. At the top, there is a toolbar with various icons and a dropdown menu set to "Limit to 50000 rows". Below the toolbar, a SQL query is written in a text editor. The query uses a Common Table Expression (CTE) to calculate the average quantity sold for each product and then ranks them. The results are displayed in a table below the query editor.

```
1 with cte as
2   (select product_name, avg(quantity) as avg_sold_qty
3    from transactions t
4    join sustainable_clothing s on t.product_id=s.product_id
5    group by 1)
6
7   select product_name, avg_sold_qty,
8          dense_rank() over(order by avg_sold_qty) as rank_avg
9   from cte;
```

Below the query editor, there is a "Result Grid" tab. The table below shows the results of the query, with columns for product\_name, avg\_sold\_qty, and rank\_avg. The products are ranked by their average quantity sold, with Organic Cotton T-Shirt, Hemp Crop Top, Organic Cotton Dress, Cork Sandals, Recycled Nylon Backpack, and Bamboo Bathrobe all having an average quantity sold of 1.0000 and a rank of 1. The products with an average quantity sold of 1.1429 and a rank of 2 are Recycled Denim Jeans. The products with an average quantity sold of 1.2500 and a rank of 3 are Bamboo Lounge Pants and Hemp Baseball Cap. The products with an average quantity sold of 1.3333 and a rank of 4 are Recycled Polyester Jacket and Bamboo Yoga Leggings. The products with an average quantity sold of 1.5000 and a rank of 5 are Linen Button-Down Shirt, Hemp Overalls, and Upcycled Denim Jacket. The products with an average quantity sold of 1.6667 and a rank of 6 are Eco-Friendly Hoodie and Organic Cotton Skirt. The products with an average quantity sold of 1.7500 and a rank of 7 are Linen Jumpsuit and Organic Cotton Socks.

product_name	avg_sold_qty	rank_avg
Organic Cotton T-Shirt	1.0000	1
Hemp Crop Top	1.0000	1
Organic Cotton Dress	1.0000	1
Cork Sandals	1.0000	1
Recycled Nylon Backpack	1.0000	1
Bamboo Bathrobe	1.0000	1
Recycled Denim Jeans	1.1429	2
Bamboo Lounge Pants	1.2500	3
Hemp Baseball Cap	1.2500	3
Recycled Polyester Jacket	1.3333	4
Bamboo Yoga Leggings	1.3333	4
Linen Button-Down Shirt	1.5000	5
Hemp Overalls	1.5000	5
Upcycled Denim Jacket	1.5000	5
Eco-Friendly Hoodie	1.6667	6
Organic Cotton Skirt	1.6667	6
Linen Jumpsuit	1.7500	7
Organic Cotton Socks	1.7500	7

# Insights

The following topics are completely covered in this case study:

- Joins in SQL
- Where clause
- Aggregate functions
- Group by clause
- Order by clause
- Limit in SQL
- Window Functions
- CTEs

The following insights can be gathered for this case study:

- The Summer Sale has the maximum number of transactions with the total revenue of 1044, i.e., it has the maximum success rate as compared to the other campaigns.
- Organic cotton sweater is the Hot Selling Product, while Recycled Denim jeans has the maximum percentage of share in the total revenue i.e., 13%.
- The maximum average revenue that has been generated in a day of campaign is 139.9
- The average number of quantities sold is more during campaigns but the total revenue generated is more outside the campaign as compared to the total revenue during campaign.