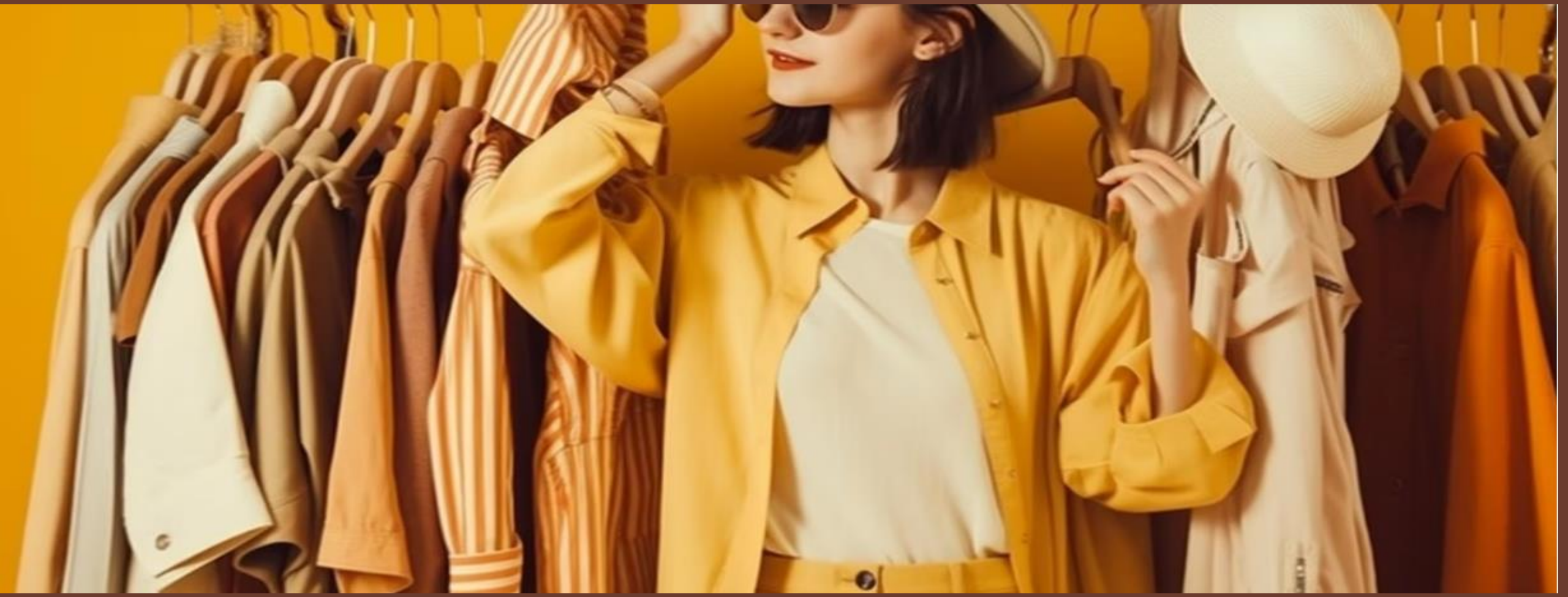
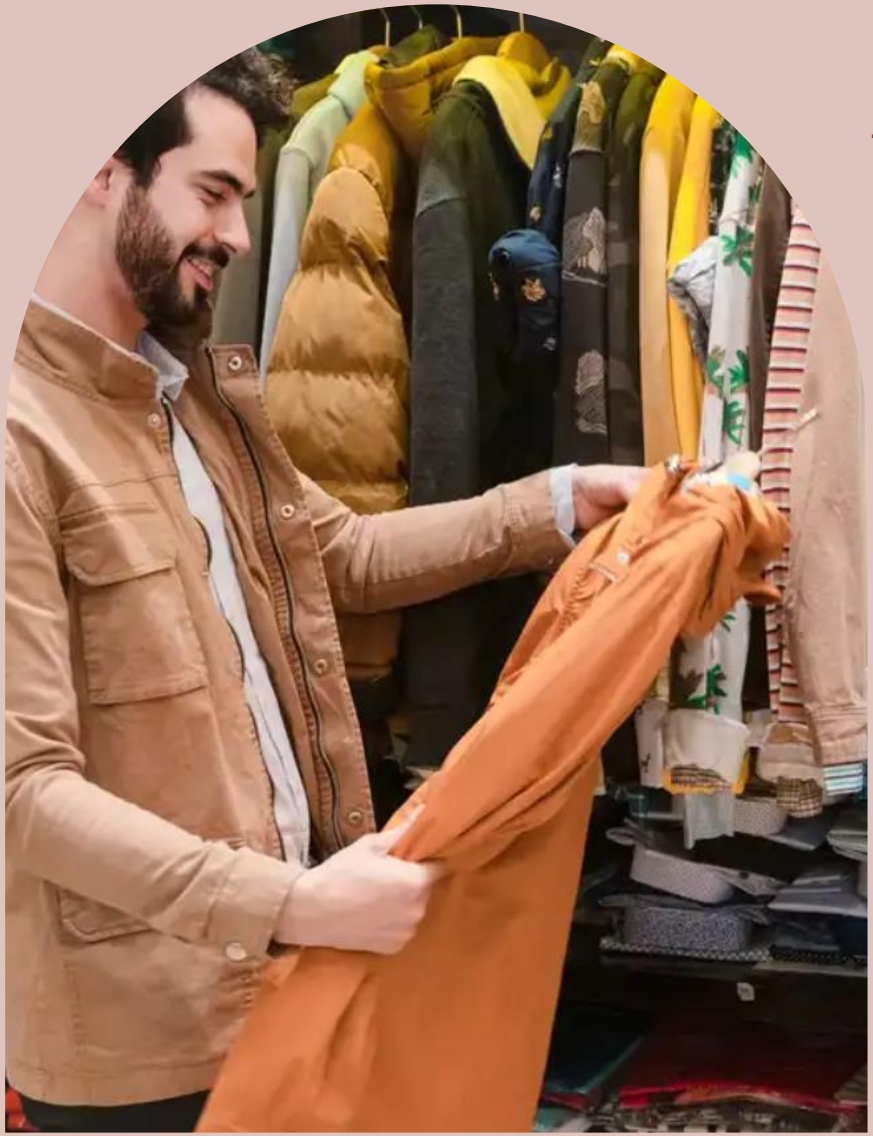
BALANCED TREE

CLOTHING TREE



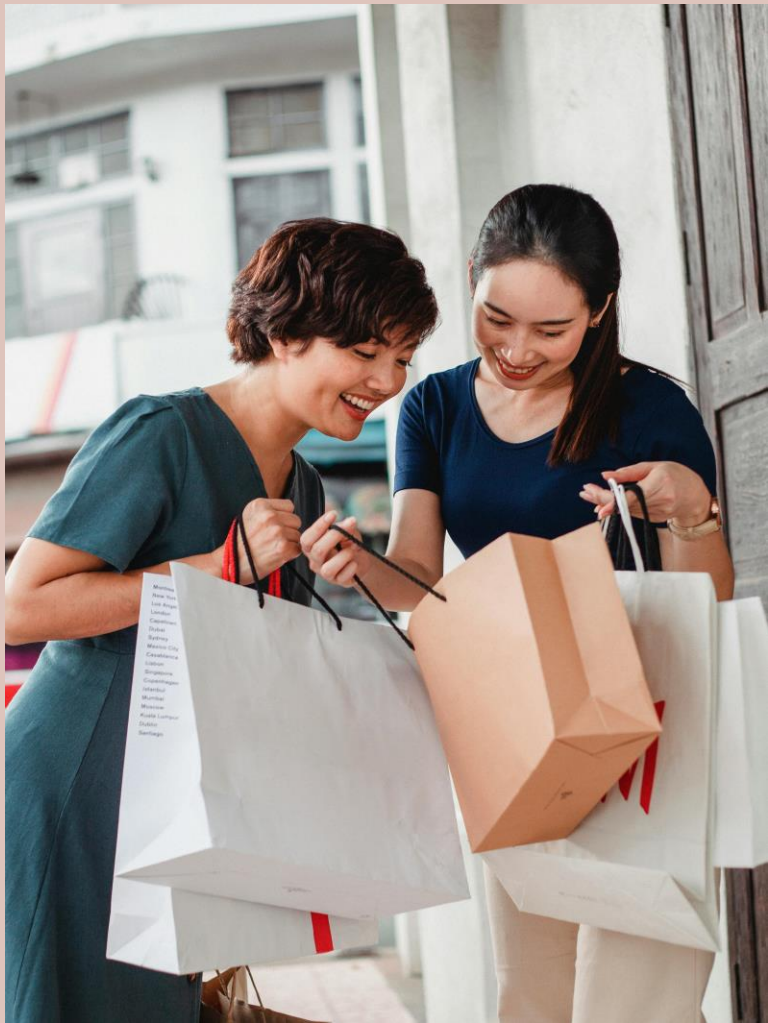
-SRABANA BAIDYA

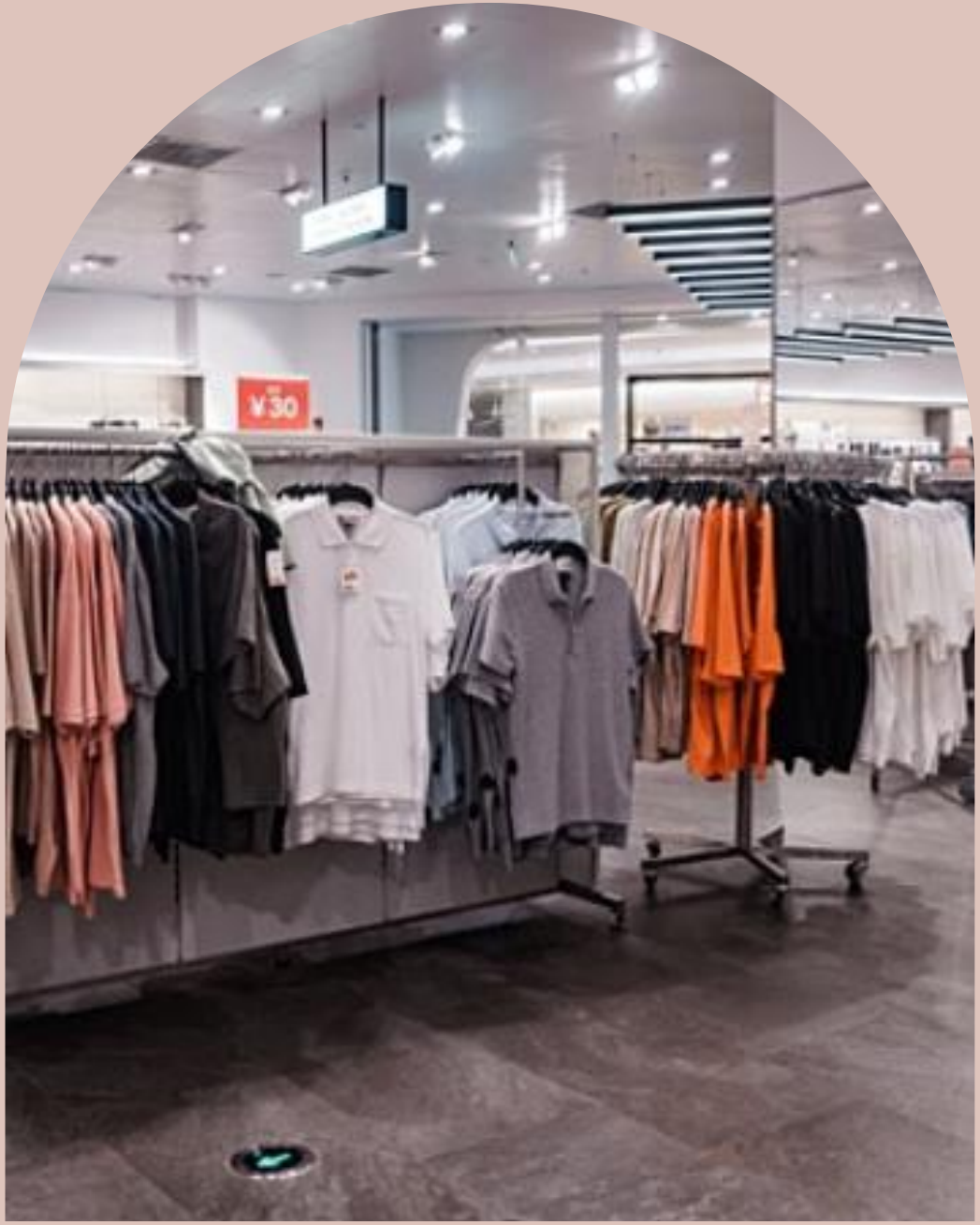


ABOUT COMPANY

Balanced Tree Clothing Company prides themselves on providing an optimised range of clothing and lifestyle wear for the modern adventurer!

Danny, the CEO of this trendy fashion company has asked you to assist the team's merchandising teams analyse their sales performance and generate a basic financial report to share with the wider business.





DATASET DETAILS

For this case study there is a total of 4 datasets for this case study - however you will only need to utilised 2 main tables to solve all of the regular questions, and the additional 2 tables are used only for the bonus challenge question!

PRODUCT DETAILS

balanced_tree.product_details includes all information about the entire range that Balanced Clothing sells in their store.

product_id	price	product_name	category_id	segment_id	style_id	category_name	segment_name	style_name
c4a632	13	Navy Oversized Jeans - Womens	1	3	7	Womens	Jeans	Navy Oversized
e83aa3	32	Black Straight Jeans - Womens	1	3	8	Womens	Jeans	Black Straight
e31d39	10	Cream Relaxed Jeans - Womens	1	3	9	Womens	Jeans	Cream Relaxed
d5e9a6	23	Khaki Suit Jacket - Womens	1	4	10	Womens	Jacket	Khaki Suit
72f5d4	19	Indigo Rain Jacket - Womens	1	4	11	Womens	Jacket	Indigo Rain
9ec847	54	Grey Fashion Jacket - Womens	1	4	12	Womens	Jacket	Grey Fashion
5d267b	40	White Tee Shirt - Mens	2	5	13	Mens	Shirt	White Tee
c8d436	10	Teal Button Up Shirt - Mens	2	5	14	Mens	Shirt	Teal Button Up
2a2353	57	Blue Polo Shirt - Mens	2	5	15	Mens	Shirt	Blue Polo
f084eb	36	Navy Solid Socks - Mens	2	6	16	Mens	Socks	Navy Solid
b9a74d	17	White Striped Socks - Mens	2	6	17	Mens	Socks	White Striped
2feb6b	29	Pink Fluro Polkadot Socks - Mens	2	6	18	Mens	Socks	Pink Fluro Polk...

PRODUCT SALES

balanced_tree.sales contains product level information for all the transactions made for Balanced Tree including quantity, price, percentage discount, member status, a transaction ID and also the transaction timestamp.

prod_id	qty	price	discount	member_m	txn_id	start_txn_time
c4a632	4	13	17	t	54f307	2021-02-13 01:59:43
5d267b	4	40	17	t	54f307	2021-02-13 01:59:43
b9a74d	4	17	17	t	54f307	2021-02-13 01:59:43
2feb6b	2	29	17	t	54f307	2021-02-13 01:59:43
c4a632	5	13	21	t	26cc98	2021-01-19 01:39:00
e31d39	2	10	21	t	26cc98	2021-01-19 01:39:00
72f5d4	3	19	21	t	26cc98	2021-01-19 01:39:00
2a2353	3	57	21	t	26cc98	2021-01-19 01:39:00
f084eb	3	36	21	t	26cc98	2021-01-19 01:39:00
c4a632	1	13	21	f	ef648d	2021-01-27 02:18:17
e83aa3	5	32	21	f	ef648d	2021-01-27 02:18:17
d5e9a6	1	23	21	f	ef648d	2021-01-27 02:18:17
72f5d4	1	19	21	f	ef648d	2021-01-27 02:18:17
5d267b	3	40	21	f	ef648d	2021-01-27 02:18:17
f084eb	4	36	21	f	ef648d	2021-01-27 02:18:17
b9a74d	4	17	21	f	ef648d	2021-01-27 02:18:17
c4a632	2	13	23	t	fba96f	2021-03-03 00:32:56

PRODUCT HIERARCHY AND PRODUCT PRICE

Thes tables are used only for the bonus question where we will use them to recreate the balanced_tree.product_details table.

id	parent_id	level_text	level_name
1	NULL	Womens	Category
2	NULL	Mens	Category
3	1	Jeans	Segment
4	1	Jacket	Segment
5	2	Shirt	Segment
6	2	Socks	Segment
7	3	Navy Oversized	Style
8	3	Black Straight	Style
9	3	Cream Relaxed	Style
10	4	Khaki Suit	Style
11	4	Indigo Rain	Style
12	4	Grey Fashion	Style
13	5	White Tee	Style
14	5	Teal Button Up	Style
15	5	Blue Polo	Style
16	6	Navy Solid	Style
17	6	White Striped	Style

id	product_id	price
7	c4a632	13
8	e83aa3	32
9	e31d39	10
10	d5e9a6	23
11	72f5d4	19
12	9ec847	54
13	5d267b	40
14	c8d436	10
15	2a2353	57
16	f084eb	36
17	b9a74d	17
18	2feb6b	29

HIGH LEVEL SALES ANALYSIS

- What was the total quantity sold for all products ?

```sql
SELECT SUM(qty) AS total_quantity
FROM balanced_tree.sales;
```

OUTPUT

total_quantity
45216

- What is the total generated revenue for all products before discounts ?

```sql
SELECT SUM(qty * price) AS total_revenue
FROM balanced_tree.sales;
```

OUTPUT

	total_revenue
▶	1289453

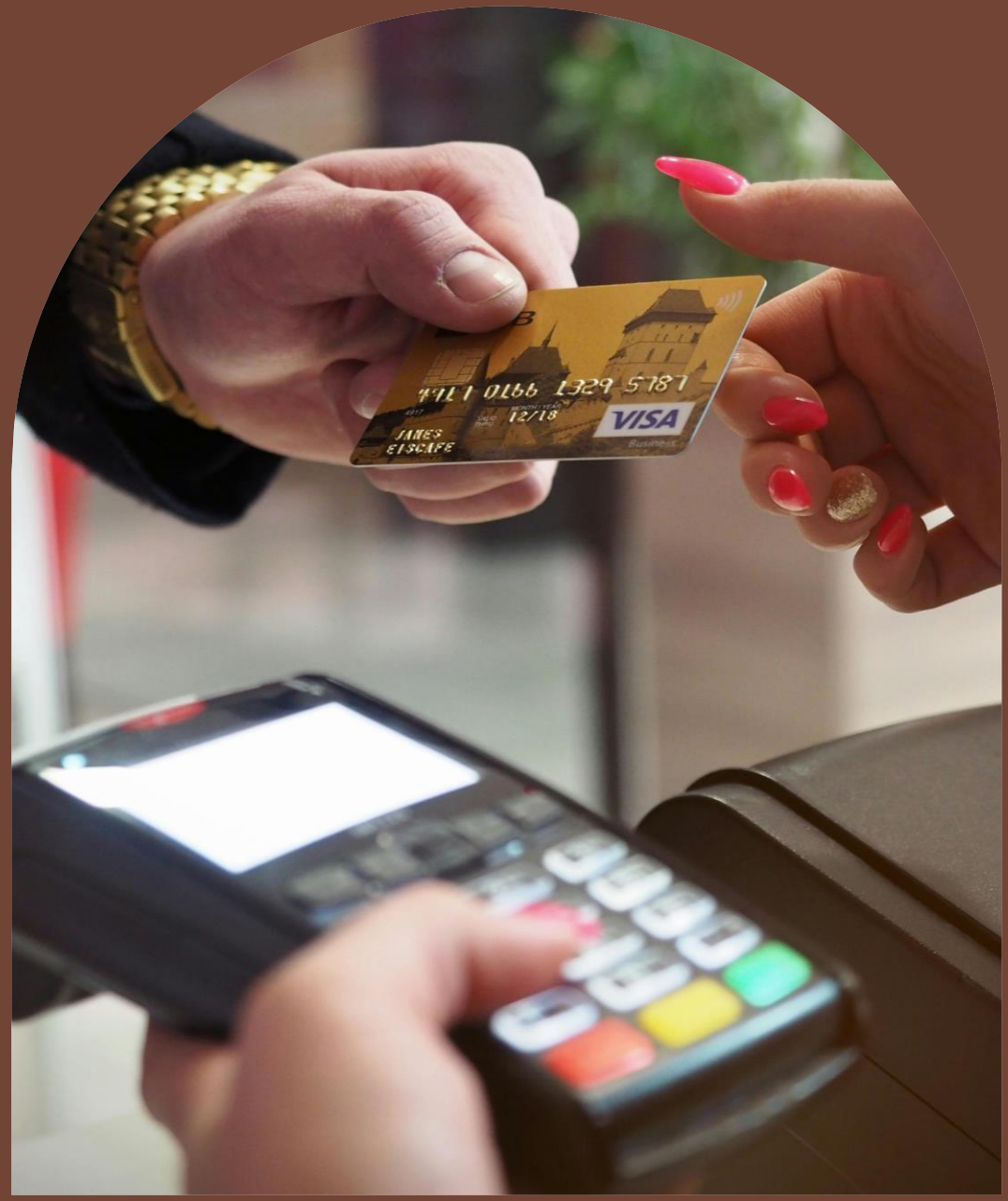
- What is the total discount amount for all products?

```sql
SELECT ROUND(SUM(qty * price * discount / 100), 2) AS total_discount
FROM balanced_tree.sales;
```

OUTPUT

total_discount
156229.14

TRANSACTION ANALYSIS



❑ How many unique transactions were there?

```sql
SELECT COUNT(DISTINCT txn_id) AS unique_transaction
FROM balanced_tree.sales;
```

OUTPUT

	unique_transaction
▶	2500

❑ What is the average unique products purchased in each transaction?

```sql
WITH cte AS (
    SELECT txn_id, COUNT(DISTINCT prod_id) AS unique_products
    FROM balanced_tree.sales
    GROUP BY txn_id
)
SELECT ROUND(AVG(unique_products), 1) AS avg_unique_products
FROM cte;
```

OUTPUT

	avg_unique_products
▶	6.0

❑ What are the 25th, 50th and 75th percentile values for the revenue per transaction ?

```sql
WITH cte_1 AS (
        SELECT txn_id, qty , price,
    ROW_NUMBER()OVER(PARTITION BY txn_id ORDER BY price) AS row_num,
    count(*)OVER(PARTITION BY txn_id) AS product_count
 FROM sales ) ,
cte_2 AS (
        SELECT * ,
        floor(product_count *0.25) AS 25th_percentile_index,
        floor(product_count *0.5)  AS 50th_percentile_index,
        floor(product_count *0.275) AS 75th_percentile_index
      FROM cte_1 ),
cte_3 AS (
         SELECT txn_id , (qty*price) AS 25th_percentile_revenue
         FROM cte_2
         WHERE row_num = 25th_percentile_index ),
cte_4 AS (
     SELECT txn_id , (qty*price) AS 50th_percentile_revenue
         FROM cte_2
     WHERE row_num = 50th_percentile_index ),
cte_5 AS  (
         SELECT txn_id , (qty*price) AS 75th_percentile_revenue
         FROM cte_2
     WHERE row_num = 75th_percentile_index )

SELECT c3.txn_id , 25th_percentile_revenue , 50th_percentile_revenue , 75th_percentile_revenue
FROM cte_3 AS c3
INNER JOIN cte_4 AS c4 USING(txn_id)
INNER JOIN cte_5 As c5 USING(txn_id) ;
```

OUTPUT

txn_id	25th_percentile_revenue	50th_percentile_revenue	75th_percentile_revenue
000027	30	34	30
000106	10	95	10
000dd8	10	26	10
003920	40	69	40
003c6d	26	57	26
003ea6	30	39	30
0053d3	40	68	40
00a68b	20	38	20
00c8dc	20	39	20
00d139	10	34	10
00ebd5	50	115	50

❑ What is the average discount value per transaction?

```sql
WITH cte AS (
    SELECT txn_id, ROUND(SUM(qty * price * discount / 100), 2) AS discount_value
    FROM balanced_tree.sales
    GROUP BY txn_id
)
SELECT ROUND(AVG(discount_value), 2) AS avg_discount_value
FROM cte;
```

OUTPUT

avg_discount_value
62.5

❑ What is the percentage split of all transactions for members vs non-members?

```sql
WITH SalesSummary AS (
    SELECT
        member_m,
        COUNT(*) AS member_count
    FROM
        sales
    GROUP BY
        member_m
)

SELECT
    IF(member_m = "t", "Members", "Non-members") AS type_of_members,
    ROUND(member_count * 100 / (SELECT COUNT(*) FROM sales), 1) AS prcnt
FROM
    SalesSummary;
```

OUTPUT

	type_of_members	prcnt
▶	Members	60.0
	Non-members	40.0

- What is the average revenue for member transactions and non-member transactions?

```sql
WITH cte_member_revenue AS (
    SELECT
        member_m,
        txn_id,
        SUM(price * qty) AS revenue
    FROM balanced_tree.sales
    GROUP BY
        member_m,
        txn_id
)
SELECT
    member_m,
    ROUND(AVG(revenue), 2) AS avg_revenue
FROM cte_member_revenue
GROUP BY member_m;
```

OUTPUT

	member_m	avg_revenue
▶	t	516.27
	f	515.04

PRODUCT ANALYSIS

❑ What are the top 3 products by total revenue before discount?

```sql
SELECT s.prod_id, d.product_name, SUM(s.qty * s.price)
 AS total_revenue
FROM balanced_tree.sales AS s
JOIN balanced_tree.product_details
AS d ON d.product_id = s.prod_id
GROUP BY s.prod_id, d.product_name
ORDER BY total_revenue DESC
LIMIT 3;
```

OUTPUT

prod_id	product_name	total_revenue
2a2353	Blue Polo Shirt - Mens	217683
9ec847	Grey Fashion Jacket - Womens	209304
5d267b	White Tee Shirt - Mens	152000

❑ What is the total quantity, revenue and discount for each segment?

```sql
WITH qty_details AS (
    SELECT segment_name, SUM(s.qty) AS total_qty
    FROM balanced_tree.sales AS s
    JOIN balanced_tree.product_details p ON p.product_id = s.prod_id
    GROUP BY segment_name
),
revenue_details AS (
    SELECT segment_name, SUM(s.qty * s.price) AS total_revenue
    FROM balanced_tree.sales AS s
    JOIN balanced_tree.product_details p ON p.product_id = s.prod_id
    GROUP BY segment_name
),
discount_details AS (
    SELECT segment_name,
     ROUND(SUM(s.qty * s.price * s.discount / 100), 2)
    AS total_discount
    FROM balanced_tree.sales AS s
    JOIN balanced_tree.product_details p ON p.product_id = s.prod_id
    GROUP BY segment_name
)
SELECT q.segment_name, q.total_qty, r.total_revenue, d.total_discount
FROM qty_details AS q
JOIN revenue_details AS r USING (segment_name)
JOIN discount_details AS d USING (segment_name);
```

OUTPUT

segment_name	total_qty	total_revenue	total_discount
Jeans	11349	208350	25343.97
Shirt	11265	406143	49594.27
Socks	11217	307977	37013.44
Jacket	11385	366983	44277.46

❑ What is the top selling product for each segment?

```sql
WITH cte1 AS (
    SELECT p.product_name, p.segment_name, SUM(s.qty) AS sold_quantity
    FROM balanced_tree.product_details AS p
    JOIN balanced_tree.sales AS s ON s.prod_id = p.product_id
    GROUP BY p.product_name, p.segment_name
),
cte2 AS (
    SELECT *,
           DENSE_RANK() OVER(PARTITION BY segment_name
            ORDER BY sold_quantity DESC) AS rnk
    FROM cte1
)
SELECT segment_name, product_name, sold_quantity
FROM cte2
WHERE rnk = 1;
```

OUTPUT

segment_name	product_name	sold_quantity
Jacket	Grey Fashion Jacket - Womens	3876
Jeans	Navy Oversized Jeans - Womens	3856
Shirt	Blue Polo Shirt - Mens	3819
Socks	Navy Solid Socks - Mens	3792

❑ What is the total quantity, revenue and discount for each category?

```sql
SELECT
    details.category_id,
    details.category_name,
    SUM(sales.qty) AS total_quantity,
    SUM(sales.qty * sales.price) AS total_revenue,
    ROUND(SUM(sales.qty * sales.price * sales.discount) / 100, 2)
     AS total_discount
FROM balanced_tree.sales AS sales
INNER JOIN balanced_tree.product_details AS details
    ON sales.prod_id = details.product_id
GROUP BY details.category_id, details.category_name
ORDER BY total_revenue DESC;
```

OUTPUT

category_id	category_name	total_qty	total_revenue	total_discount
2	Mens	22482	714120	86607.71
1	Womens	22734	575333	69621.43

❑ What is the top selling product for each category?

```sql
WITH cte1 AS (
    SELECT p.product_name, p.category_name, SUM(s.qty)
    AS sold_quantity
    FROM balanced_tree.product_details AS p
    JOIN balanced_tree.sales AS s ON s.prod_id = p.product_id
    GROUP BY p.product_name, p.category_name
),
cte2 AS (
    SELECT *,
           DENSE_RANK() OVER(PARTITION BY category_name
            ORDER BY sold_quantity DESC) AS rnk
    FROM cte1
)
SELECT category_name, product_name, sold_quantity
FROM cte2
WHERE rnk = 1;
```

OUTPUT

	category_name	product_name	sold_quantity
▶	Mens	Blue Polo Shirt - Mens	3819
	Womens	Grey Fashion Jacket - Womens	3876

❑ What is the percentage split of revenue by product for each segment?

```sql
WITH cte1 AS (
    SELECT segment_name, product_name, SUM(s.qty * s.price)
     AS total_revenue
    FROM balanced_tree.sales AS s
  JOIN balanced_tree.product_details AS p ON p.product_id = s.prod_id
    GROUP BY segment_name, product_name
    ORDER BY segment_name
),
cte2 AS (
    SELECT *,
           SUM(total_revenue) OVER(PARTITION BY segment_name)
           AS segment_rev
    FROM cte1
)
SELECT segment_name, product_name,
 ROUND(total_revenue * 100 / segment_rev, 1) AS prcnt
FROM cte2;
```

OUTPUT

segment_name	product_name	prcnt
Jacket	Grey Fashion Jacket - Womens	57.0
Jacket	Indigo Rain Jacket - Womens	19.5
Jacket	Khaki Suit Jacket - Womens	23.5
Jeans	Black Straight Jeans - Womens	58.1
Jeans	Cream Relaxed Jeans - Womens	17.8
Jeans	Navy Oversized Jeans - Womens	24.1
Shirt	Blue Polo Shirt - Mens	53.6
Shirt	Teal Button Up Shirt - Mens	9.0
Shirt	White Tee Shirt - Mens	37.4
Socks	Navy Solid Socks - Mens	44.3
Socks	Pink Fluro Polkadot Socks - Mens	35.5
Socks	White Striped Socks - Mens	20.2

❑ What is the percentage split of revenue by segment for each category?

```sql
WITH cte1 AS (
    SELECT category_name, SUM(s.qty * s.price) AS total_revenue
    FROM balanced_tree.sales AS s
JOIN balanced_tree.product_details AS p ON p.product_id = s.prod_id
    GROUP BY category_name
),
cte2 AS (
    SELECT SUM(total_revenue) AS total
    FROM cte1
)
SELECT category_name,
 ROUND(total_revenue * 100 / (SELECT total FROM cte2), 1) AS prcnt
FROM cte1;
```

OUTPUT

	category_name	segment_name	prcnt
▶	Mens	Shirt	56.9
	Mens	Socks	43.1
	Womens	Jacket	63.8
	Womens	Jeans	36.2

❑ What is the percentage split of total revenue by category?

```sql
WITH cte1 AS (
    SELECT category_name, SUM(s.qty * s.price) AS total_revenue
    FROM balanced_tree.sales AS s
    JOIN balanced_tree.product_details AS p
    ON p.product_id = s.prod_id
    GROUP BY category_name
),
cte2 AS (
    SELECT SUM(total_revenue) AS total
    FROM cte1
)
SELECT category_name, ROUND(total_revenue * 100 /
(SELECT total FROM cte2), 1) AS prcnt
FROM cte1;
```

OUTPUT

	category_name	prcnt
▶	Womens	44.6
	Mens	55.4

❑ What is the total transaction “penetration” for each product?

```sql
WITH cte AS (
    SELECT COUNT(DISTINCT txn_id) AS total_txn
    FROM balanced_tree.sales
)
SELECT p.product_name,
       ROUND(COUNT(DISTINCT s.txn_id) * 100 /
        (SELECT total_txn FROM cte), 1) AS penetration
FROM balanced_tree.sales AS s
RIGHT JOIN balanced_tree.product_details AS p
 ON p.product_id = s.prod_id
WHERE s.qty ⩾ 1
GROUP BY p.product_name;
```

OUTPUT

product_name	penetration
Navy Oversized Jeans - Womens	51.0
White Tee Shirt - Mens	50.7
White Striped Socks - Mens	49.7
Pink Fluro Polkadot Socks - Mens	50.3
Cream Relaxed Jeans - Womens	49.7
Indigo Rain Jacket - Womens	50.0
Blue Polo Shirt - Mens	50.7
Navy Solid Socks - Mens	51.2
Black Straight Jeans - Womens	49.8
Khaki Suit Jacket - Womens	49.9
Grey Fashion Jacket - Womens	51.0
Teal Button Up Shirt - Mens	49.7

What is the most common combination of at least 1 quantity of any 3 products in a 1 single transaction?

```sql
WITH cte AS (
    SELECT s.prod_id, p.product_name, s.qty, s.price, s.discount,
     s.member, txn_id, s.start_txn_time
    FROM balanced_tree.sales AS s
    INNER JOIN balanced_tree.product_details p ON
     p.product_id = s.prod_id
)
SELECT c1.product_name AS first_product, c2.product_name
AS second_product,
       c3.product_name AS third_product, COUNT(*)
        AS combination_count
FROM cte AS c1
INNER JOIN cte AS c2 ON c2.txn_id =
c1.txn_id AND c1.prod_id < c2.prod_id
INNER JOIN cte AS c3 ON c3.txn_id =
 c1.txn_id AND c2.prod_id < c3.prod_id
GROUP BY first_product, second_product, third_product
ORDER BY combination_count DESC
LIMIT 1;
```

OUTPUT

first_product	second_product	third_product	combination_count
White Tee Shirt - Mens	Grey Fashion Jacket - Womens	Teal Button Up Shirt - Mens	352

BONUS CHALLENGE

- Use a single SQL query to transform the product_hierarchy and product_prices datasets to the product_details table.

```sql
WITH t AS (
    SELECT h.id AS style_id, h.level_text AS style_name, t1.id
    AS segment_id,t1.level_text AS segment_name,
  t1.parent_id AS category_id, t2.level_text AS category_name
    FROM balanced_tree.product_hierarchy h
    LEFT JOIN balanced_tree.product_hierarchy t1
     ON h.parent_id = t1.id
    LEFT JOIN balanced_tree.product_hierarchy t2
     ON t1.parent_id = t2.id
    WHERE h.id BETWEEN 7 AND 18
)
SELECT p.product_id,
       p.price,
    CONCAT(t.style_name, ' ', t.segment_name, ' - ', t.category_name)
        AS product_name, t.category_id, t.segment_id,t.style_id,
 t.category_name,t.segment_name,t.style_name
FROM balanced_tree.product_prices p
LEFT JOIN t ON p.id = t.style_id;
```

OUTPUT

product_id	price	product_name	category_id	segment_id	style_id	category_name	segment_name	style_name
c4a632	13	Navy Oversized Jeans - Womens	1	3	7	Womens	Jeans	Navy Oversized
e83aa3	32	Black Straight Jeans - Womens	1	3	8	Womens	Jeans	Black Straight
e31d39	10	Cream Relaxed Jeans - Womens	1	3	9	Womens	Jeans	Cream Relaxed
d5e9a6	23	Khaki Suit Jacket - Womens	1	4	10	Womens	Jacket	Khaki Suit
72f5d4	19	Indigo Rain Jacket - Womens	1	4	11	Womens	Jacket	Indigo Rain
9ec847	54	Grey Fashion Jacket - Womens	1	4	12	Womens	Jacket	Grey Fashion
5d267b	40	White Tee Shirt - Mens	2	5	13	Mens	Shirt	White Tee
c8d436	10	Teal Button Up Shirt - Mens	2	5	14	Mens	Shirt	Teal Button Up
2a2353	57	Blue Polo Shirt - Mens	2	5	15	Mens	Shirt	Blue Polo
f084eb	36	Navy Solid Socks - Mens	2	6	16	Mens	Socks	Navy Solid
b9a74d	17	White Striped Socks - Mens	2	6	17	Mens	Socks	White Striped
2feb6b	29	Pink Fluro Polkadot Socks - Mens	2	6	18	Mens	Socks	Pink Fluro Polk...

THANK YOU

THANK YOU

THANK YOU

THANK YOU