High Level Sales Analysis

1. *What was the total quantity sold for all products?*

```
Select sum(qty) as total_quantity_sold_products
From sales;
```

total_quantity_sold_products
45216

2. *What is the total generated revenue for all products before discounts?*

```
Select sum(qty *price) as total_revenue
From sales;
```

total_revenue
1289453

3. *What was the total discount amount for all products?*

```
Select round(sum(qty * price * discount/100),2) as total_discount
From sales;
```

total_discount
156229.14

Transaction Analysis

4. *How many unique transactions were there?*

```
Select count(distinct(txn_id)) as number_of_transactions
From sales;
```

number_of_transactions
2500

5. *What is the average unique products purchased in each transaction?*

```
Select round(avg(unique_products),0) as average_unique_products
From (Select txn_id, count(distinct(prod_id)) as unique_products
From sales
Group by txn_id)temp;
```

average_unique_products
6

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

I'm doing this exercise using MySQL Workbench, so I can't use 'PERCENTILE_CONT' function. What I managed to come up with is:

```
Create View Revenue_transaction As
SELECT distinct(txn_id), SUM(qty*price) AS revenue
FROM sales
GROUP BY txn_id;
```

txn_id	revenue
54f307	338
26cc98	421
ef648d	547
fba96f	440
4e9268	306

```
Create View percentile_rank As
SELECT txn_id, revenue, ROUND(PERCENT_RANK() OVER (ORDER BY txn_id),2) as percentile_rank
FROM Revenue_transaction
Order by percentile_rank DESC;
```

txn_id	revenue	percentile_rank
fe7039	626	1
feaf08	627	1
fee9ce	631	1
fefff0	242	1
ff0bc8	451	1

```
Select *
From percentile_rank
Where percentile_rank>0.75
Group by txn_id;
```

txn_id	revenue	percentile_rank
fe7039	626	1
feaf08	627	1
fee9ce	631	1
fefff0	242	1
ff0bc8	451	1

613 row(s) returned

7. *What is the average discount value per transaction?*

```
Select round(avg(total_discount),2) as average_discount_per_transaction
From
(Select txn_id, sum(qty*price*discount /100) as total_discount
From sales
Group by txn_id) temp;
```

average_discount_per_transaction
62.49

8. *What is the percentage split of all transactions for members vs non-members?*

```
Select round(100*count(distinct case
when members = 1 Then txn_id Else 0 End)/count(distinct(txn_id)),2) as members_percentage,
round(100*count(distinct case when members= 0 Then txn_id Else 0 End)/count(distinct(txn_id)),2) as non_members_percentage
From sales;
```

members_percentage	non_members_percentage
60.24	39.84

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

```
Create View members_revenu As
Select members, txn_id, sum(qty*price) as revenue
From sales
Group by members, txn_id;

Select members, round(avg(revenue),2) as average_revenue
From members_revenu
Group by members;
```

members	average_revenue
1	516.27
0	515.04

Product Analysis

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

```
Select s. prod_id, d.product_name, sum(s.qty*s.price) as revenue
From sales as s
Join product_details as d On s.prod_id = d.product_id
Group by prod_id, d.product_name
Order by revenue DESC Limit 3;
```

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

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

```
Select d.segment_name, sum(s.qty) as total_quantity, sum(s.qty*s.price) as revenue,
round(sum((s.qty*s.discount*s.price)/100),2) as total_discount
 From sales as s
 Join product_details as d On s.prod_id = d.product_id
 Group by d.segment_name
 Order by d.segment_name;
```

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

12. What is the top selling product for each segment?

```
Select d.segment_name, d.product_name, sum(s.qty) as total_selling,
RANK() OVER (PARTITION BY d.segment_name ORDER BY SUM(s.QTY) DESC) AS rank_selling
From sales as s
Join product_details as d On s.prod_id = d.product_id
Group by d.segment_name, d.product_name
Order by rank_selling;
```

segment_name	product_name	total_selling	rank_selling
Jacket	Grey Fashion Jacket - Womens	3876	1
Jeans	Navy Oversized Jeans - Womens	3856	1
Shirt	Blue Polo Shirt - Mens	3819	1
Socks	Navy Solid Socks - Mens	3792	1
Jacket	Indigo Rain Jacket - Womens	3757	2

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

```
Select d.category_name, sum(s.qty) as total_quantity, sum(s.qty*s.price) as revenue,
round(sum((s.qty*s.discount*s.price)/100),2) as total_discount
 From sales as s
 Join product_details as d On s.prod_id = d.product_id
 Group by d.category_name
 Order by d.category_name;
```

category_name	total_quantity	revenue	total_discount
Mens	22482	714120	86607.71
Womens	22734	575333	69621.43

14. What is the top selling product for each category?

```
Select d.category_name, d.product_name, sum(s.qty) as total_selling,
RANK() OVER (PARTITION BY d.category_name ORDER BY SUM(s.QTY) DESC) AS rank_selling
From sales as s
Join product_details as d On s.prod_id = d.product_id
Group by d.category_name, d.product_name
Order by rank_selling;
```

category_name	product_name	total_selling	rank_selling
Mens	Blue Polo Shirt - Mens	3819	1
Womens	Grey Fashion Jacket - Womens	3876	1
Mens	White Tee Shirt - Mens	3800	2
Womens	Navy Oversized Jeans - Womens	3856	2

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

```
Create View revenue_segment As
Select d.segment_name, d.product_name, sum(s.qty*s.price) as revenue
From sales as s
Join product_details as d On s.prod_id = d.product_id
Group by d.segment_name, d.product_name
Order by revenue DESC;

 Create view total_revenue_segment As
 Select d.segment_name, sum(s.qty*s.price) as total_revenue
 From sales as s
 Join product_details as d On s.prod_id = d.product_id
 Group by d.segment_name
 Order by total_revenue DESC;

Select s.segment_name, s.product_name, round(100*s.revenue/t.total_revenue,2) as percentage_split_of_revenue
From revenue_segment as s
Join total_revenue_segment as t On s.segment_name=t.segment_name
Order by percentage_split_of_revenue DESC;
```

segment_name	product_name	percentage_split_of_revenue
Jeans	Black Straight Jeans - Womens	58.15
Jacket	Grey Fashion Jacket - Womens	57.03
Shirt	Blue Polo Shirt - Mens	53.60
Socks	Navy Solid Socks - Mens	44.33
Shirt	White Tee Shirt - Mens	37.43
Socks	Pink Fluro Polkadot Socks - Mens	35.50
Jeans	Navy Oversized Jeans - Womens	24.06
Jacket	Khaki Suit Jacket - Womens	23.51
Socks	White Striped Socks - Mens	20.18
Jacket	Indigo Rain Jacket - Womens	19.45
Jeans	Cream Relaxed Jeans - Womens	17.79
Shirt	Teal Button Up Shirt - Mens	8.98

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

```
Create View revenue_category As
Select d.category_name, d.segment_name, sum(s.qty*s.price) as revenue
From sales as s
Join product_details as d On s.prod_id = d.product_id
Group by d.category_name, d.segment_name
Order by revenue DESC;

Create view total_revenue_category As
Select d.category_name, sum(s.qty*s.price) as total_revenue
From sales as s
Join product_details as d On s.prod_id = d.product_id
Group by d.category_name
Order by total_revenue DESC;
```

```
Select c.category_name, c.segment_name, round(100*c.revenue/t.total_revenue,2) as percentage_split_of_revenue
From revenue_category as c
Join total_revenue_category as t On c.category_name=t.category_name
Order by percentage_split_of_revenue DESC;
```

category_name	segment_name	percentage_split_of_revenue
Womens	Jacket	63.79
Mens	Shirt	56.87
Mens	Socks	43.13
Womens	Jeans	36.21

17. What is the percentage split of total revenue by category?

```
Select category_name, round(100*total_revenue/(Select sum(total_revenue)
From total_revenue_category),2) as percentage_split_of_total_revenue
From total_revenue_category;
```

category_name	percentage_split_of_total_revenue
Mens	55.38
Womens	44.62

18. *What is the total transaction "penetration" for each product? (hint: penetration = number of transactions where at least 1 quantity of a product was purchased divided by total number of transactions)*

```
Select d.product_name, count(s.txn_id) as total_transaction, round(100 *count(s.txn_id)/
(Select count(distinct(s.txn_id)) From Sales as s),2) as penetration
From sales as s
Join product_details as d On s.prod_id = d.product_id
Group by d.product_name;
```

product_name	total_transaction	penetration
Blue Polo Shirt - Mens	1268	50.72
Navy Solid Socks - Mens	1281	51.24
Black Straight Jeans - Womens	1246	49.84
Khaki Suit Jacket - Womens	1247	49.88
Grey Fashion Jacket - Womens	1275	51.00
Teal Button Up Shirt - Mens	1242	49.68

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

First, we need a table with transaction and product columns.

```
Create View transaction_and_product As
Select s.txn_id, d.product_name
From sales as s
Join product_details as d On s.prod_id = d.product_id;
```

txn_id	product_name
54f307	Navy Oversized Jeans - Womens
54f307	White Tee Shirt - Mens
54f307	White Striped Socks - Mens
54f307	Pink Fluro Polkadot Socks - Mens
26cc98	Navy Oversized Jeans - Womens
26cc98	Cream Relaxed Jeans - Womens
26cc98	Indigo Rain Jacket - Womens

```
Select p1.product_name as product_1, p2.product_name as product_2, p3.product_name as product_3, count(*) as transactions
From transaction_and_product as p1
Join transaction_and_product as p2 On p1.txn_id = p2.txn_id and p1.product_name < p2.product_name
Join transaction_and_product as p3 On p1.txn_id = p3.txn_id
and p1.product_name < p3.product_name and p2.product_name < p3.product_name
Where p1.product_name is not null and p2.product_name is not null and p3.product_name is not null
Group by p1.product_name, p2.product_name, p3.product_name
Order by transactions DESC Limit 1;
```

product_1	product_2	product_3	transactions
Grey Fashion Jacket - Womens	Teal Button Up Shirt - Mens	White Tee Shirt - Mens	352