```sql
USE[SQL_Retail data analysis]

-- Q1: What is the total no. of rows in each of the 3 tables in the database?


SELECT
(SELECT COUNT(transaction_id)  FROM Transactions )as transaction_rows_count,
(SELECT COUNT(prod_cat_code)  FROM prod_cat_info)as prod_cat_rows_count,
(SELECT COUNT(customer_id)  FROM customer_table)as cust_details_rows_count;

-- Q2 What is the total number of transactions that have a return

SELECT COUNT(Qty) as Total_Return_transactions from Transactions where Qty<0;


-- Q3 Pls convert the date variables into valid date formats before proceeding
  ahead.

SELECT convert(date,tran_date,110) as date_of_txn from Transactions ;

--OR

SELECT CAST(tran_date as date) as date_of_txn from Transactions;

/*Q4 What is the time range of the transaction data available for analysis?
Show the output in number of days,months, years simultaneously in diff columns.
*/

select
DATEDIFF(year,min(tran_date),max(tran_date)) as timerange_years,
DATEDIFF(month,min(tran_date),max(tran_date)) as timerange_months,
DATEDIFF(day,min(tran_date),max(tran_date)) as timerange_days
from Transactions;

--Q5 Which product category does the sub-category "DIY" belong to?

select prod_cat from prod_cat_info
where prod_subcat='DIY';


--DATA ANALYSIS

--Q1 Which channel is most frequently used for transactions?

select top 1
count(store_type) as channel_txn_freq, Store_type
from Transactions
group by Store_type
```

```
order by channel_txn_freq desc;

--Q2 Count of Male and Female customers in data base?

select
count(Gender) as Count_Gender,Gender
from Customer_table
where Gender is not null
group by Gender;

--Q3 From which city do we have the maximum number of customers and how many?

select top 1
count(customer_Id) as count,city_code
from Customer_table
group by city_code
order by count desc;

--Q4 How many sub-categories are there under book category?

select
prod_cat='books', count(prod_subcat) as no_of_subcat
from prod_cat_info
where prod_cat='Books';


--Q5 What is the Maximum Quantity of products ever ordered?

select
max(Qty) as Max_Qty_Ever_Ordered
from Transactions;


--Q6 What is the net total revenue generated in categories  Electronics and Books?

SELECT
ROUND(SUM(total_amt),2) AS NET_TOTAL_REVENUE
FROM
Transactions AS T JOIN prod_cat_info AS PCI
ON T.prod_cat_code=PCI.prod_cat_code AND T.prod_subcat_code=PCI.prod_sub_cat_code
WHERE
prod_cat in( 'Electronics','Books');


-- Q7 How many customers have > 10 Txn with us, excluding returns?

SELECT COUNT (*) AS Count_Customers_more_than_10txn FROM
(select
distinct cust_id
```

```
from Transactions
WHERE Qty>0
group by cust_id
having count(transaction_id)>10
) as C


-- Q8 What is the combined revenue earned from the "Electronics" & "Clothing"
  category, from "Flagship stores"?

SELECT round(sum(total_amt),2) as Combined_Revenue
from
Transactions as T join prod_cat_info as P
on T.prod_cat_code=P.prod_cat_code AND T.prod_subcat_code=P.prod_sub_cat_code
where
(P.prod_cat='Electronics' and T.Store_type='Flagship store') OR
  (P.prod_cat='Clothing' and T.Store_type='Flagship store')


--Q9 What is the total revenue generated from 'Male' customers in 'Electronics'
  Category?
--   Output should display total revenue by prod sub-cat.

SELECT
SUM(total_amt) as Total_Revenue,
prod_subcat
from
Transactions as T join prod_cat_info as P
ON T.prod_cat_code=p.prod_cat_code and  T.prod_subcat_code= P.prod_sub_cat_code
JOIN
Customer_table AS C
ON T.cust_id=C.customer_Id
WHERE
Gender='M' AND prod_cat='Electronics'
GROUP BY prod_subcat


--Q10 What is percentage of Sales and Returns by prod-sub-cat ?
--    Display only Top 5 sub-cat in terms of sales

---- for Sales ----

select prod_subcat_code, concat(percent_sales,'%') as Percent_of_Sales
from
(
SELECT TOP 5
prod_subcat_code, round((sum(total_amt)/(SELECT SUM(TOTAL_AMT) FROM Transactions
  WHERE total_amt>0))*100,2) as Percent_Sales
from
```

```
Transactions
where total_amt>0
group by prod_subcat_code
ORDER BY Percent_Sales desc
) as pt

------ For Returns ----------
select prod_subcat_code, concat(Percent_Returns,'%') as Percent_of_Returns
from
(
SELECT TOP 5
prod_subcat_code, round((sum(total_amt)/(SELECT SUM(TOTAL_AMT) FROM Transactions
  WHERE total_amt<0))*100,2) as Percent_Returns
from
Transactions
where total_amt<0
group by prod_subcat_code
ORDER BY Percent_Returns desc
) as pr


----Q11. For all customers aged btw 25 to 35 years, find what is the net total
  revenue generated
-------  by these customers in last 30 days of transactions from max transaction
  date avaiable in data.


select sum(total_amt)
from
(
select *,
datediff(year,DOB,getdate()) AS Age,
datediff(day,tran_date,(select max(tran_date) from Transactions )) as Txn_days
from
Transactions as T JOIN Customer_table as CT
ON T.cust_id=CT.customer_Id
) as New
where
Txn_days <= 30 and
Age between 25 and 35


---Q12. Which product category has seen the max value of returns in the last 3
  months of transactions.

SELECT top 1
prod_cat, sum(total_amt) as returns_value
from
Transactions as T join prod_cat_info as P
```

```
on T.prod_cat_code=P.prod_cat_code AND T.prod_subcat_code=P.prod_sub_cat_code
where total_amt<0 and datediff(month,tran_date,(select max(tran_date)from
  Transactions))<3
group by prod_cat
order by returns_value


----Q13. Which store type sell the maximum products; by value of sales amt and by
  Qty sold?

select top 1 store_type, sum(total_amt) as sales_value,count(Qty) as qty_sold
from Transactions
where total_amt>0 and Qty>0
group by Store_type
order by sales_value desc, qty_sold desc


----Q14. What are the categories for which average revenue is above the overall
  average.


select prod_cat,round(avg(total_amt),2)as average
from
Transactions as T join prod_cat_info as P
on T.prod_cat_code=P.prod_cat_code AND T.prod_subcat_code=P.prod_sub_cat_code
group by prod_cat
having avg(total_amt)>(select avg(total_amt) from Transactions)


----Q15. Find the average and total revenue by each subcategory for the categories
  which are
-------- among top 5 categories in terms of Qty sold.


CREATE VIEW Topfive_cat as(
select prod_cat
from
(
select top 5 prod_cat,count(Qty) as Qty_Sold
from
Transactions as T join prod_cat_info as P
on T.prod_cat_code=P.prod_cat_code AND T.prod_subcat_code=P.prod_sub_cat_code
group by prod_cat
order by Qty_Sold desc
) as PP)


Select prod_cat,prod_subcat,avg(total_amt)as avg_revenue, sum(total_amt) as
  total_revenue
```

```
from
Transactions as T join prod_cat_info as P
on T.prod_cat_code=P.prod_cat_code AND T.prod_subcat_code=P.prod_sub_cat_code
where prod_cat in (select * from Topfive_cat)
group by prod_cat,prod_subcat
```