```sql
--DATA PREPARATION AND UNDERSTANDING

--1.what is the total number of rows in each of the 3 tables in the database?
select 'customer' as header, COUNT(*) as count_rows_ from [dbo].[Customer1]
union
select 'table2', COUNT(*) from [dbo].[prod_cat_info1]
union
select 'table3', COUNT(*) from [dbo].[Transactions1]
```

OUTPUT:

```
header   count_rows_
-------- -----------
customer 11294
table2   23
table3   23053
```

```sql
--2.what is the total number of transactions that have a return?
select count(case when [total_amt] < 0 then 1 else NULL end) as cnt from  [dbo].[Transactions1]
```

OUTPUT:

```
cnt
-----------
2177
```

```sql
--3. As you would have noticed, the dates provided across the datasets are not in a correct format.
as first steps, pls convert the data variables into valid date formats befor proceeding ahead
alter table customer1
alter column DOB datetime


--4. what is the time range of the transaction data available for analysis? show the output in
number of days, months and years simultaneously in different columns.
select DATEDIFF(day, (select MIN([tran_date]) from Transactions1), (select MAX([tran_date]) from
Transactions1)) as Date_Diff,
DATEDIFF(month, (select MIN([tran_date]) from Transactions1), (select MAX([tran_date]) from
Transactions1)) as Month_Diff,
DATEDIFF(year, (select MIN([tran_date]) from Transactions1), (select MAX([tran_date]) from
Transactions1)) as Year_Diff
```

OUTPUT:

```
Date_Diff   Month_Diff  Year_Diff
----------- ----------- -----------
1130        37          3
```

```sql
--5.which product category does the sub category "DIY" belong to?
select [prod_cat] from prod_cat_info1 where prod_subcat = 'DIY'
```

OUTPUT:

```
prod_cat
--------------------------------------------------
Books
```

```sql
--DATA ANALYSIS

--1. which channel is most frequently used for transactions?
select Store_type, Count([transaction_id]) as Store_cnt from [dbo].[Transactions1]
group by Store_type having Count([transaction_id])= ( select max(store) from (select Store_type,
Count([transaction_id]) as store from [dbo].[Transactions1]
group by Store_type) y)
```

OUTPUT:

Store_type	Store_cnt
e-Shop	9311

```
--2.what is the count of male and female customers and how many?
select [Gender], COUNT([Gender]) from [dbo].[Customer1] group by [Gender] having Gender IN ('M','F')
```

OUTPUT:

Gender	
F	5506
M	5784

```
--3.from which city do we have the maximum number of customers and how many?
select [city_code], COUNT( [customer_Id]) as cust_cnt from [dbo].[Customer1]
group by [city_code] having COUNT( [customer_Id]) = (select max(cust_cnt1) from( select [city_code],
COUNT( [customer_Id]) as cust_cnt1 from [dbo].[Customer1]
group by [city_code]) y)
```

OUTPUT:

city_code	cust_cnt
3	1190

```
--4.how many sub categories are there under the book category?
select [prod_cat], Count([prod_subcat]) as prod_cnt from [dbo].[prod_cat_info1] group by [prod_cat]
having prod_cat = 'Books'
```

OUTPUT:

prod_cat	prod_cnt
Books	6

```
--5. what is the maximum quantity of products ever ordered?
select COUNT( [Qty]) as qty from [dbo].[prod_cat_info1] a, [dbo].[Transactions1] b where
a.prod_cat_code = b.prod_cat_code
group by a.prod_cat_code having COUNT([Qty]) = (select max(qty1) from (select COUNT( [Qty]) as qty1
from [dbo].[prod_cat_info1] a, [dbo].[Transactions1] b where a.prod_cat_code = b.prod_cat_code
group by a.prod_cat_code)y)
```

OUTPUT:

qty
36414

```
--6. what is the net total revenue generated in categories Electronics and books?
select a.[prod_cat], SUM([total_amt]) as revenue from [dbo].[prod_cat_info1] a,
[dbo].[Transactions1] b where a.prod_cat_code = b.prod_cat_code
group by a.[prod_cat] having a.[prod_cat] in ('Books', 'Electronics')
```

OUTPUT:

prod_cat	revenue
Books	76936164.22966
Electronics	53612318.2006073

```sql
--7. how many customers have > 10 transactions with us, excluding returns?

select count(t.[cust_id]) from
(select [cust_id] from [dbo].[Transactions1]
where [Qty]>0
group by [cust_id]
having count([transaction_id])>10) t
```

OUTPUT:

```
-----------
6
```

```sql
--8. what is the combined revenue earned from the "Electronics" and "clothing" category, from
"flagship stores"?

select sum([total_amt]) totalrevenue
from [dbo].[prod_cat_info1] inner join [dbo].[Transactions1]
on [dbo].[prod_cat_info1] .[prod_cat_code] =  [dbo].[Transactions1] .[prod_cat_code]
where [prod_cat] in ('Electronics' , 'Clothing') and [Store_type] = 'Flagship store'
```

OUTPUT:

```
totalrevenue
----------------------
14658949.8962402
```

```sql
--9. what is the total revenue generated "male" customers in "electronics" category? output should
display total revenue by prod sub-cat.

select c.prod_subcat, sum( b.total_amt)as total_revenue from [dbo].[Customer1] a,
[dbo].[Transactions1] b, [dbo].[prod_cat_info1] c
where a.customer_Id = b.cust_id and b.prod_cat_code = c.prod_cat_code and a.Gender = 'M' and
c.prod_cat = 'Electronics'  group by c.prod_subcat
```

OUTPUT:

```
prod_subcat                                        total_revenue
-------------------------------------------------- ----------------------
Audio and video                                    11406218.8554993
Cameras                                            11406218.8554993
Computers                                          11406218.8554993
Mobiles                                            11406218.8554993
Personal Appliances                                11406218.8554993

(5 rows affected)
```

```sql
--10. what is the percentage of sales and returns by product sub category; display only top 5 sub
categories in terms of sales?

select top 5 [prod_subcat],
sum(case when [total_amt] > 0 then [total_amt] end)/(select SUM([total_amt]) from
[dbo].[Transactions1] where [total_amt] > 0)*100 [% of sales],
sum(case when [total_amt] < 0 then [total_amt] end)/(select SUM([total_amt]) from
[dbo].[Transactions1] where [total_amt] < 0)*100 [% of return]
from [dbo].[Transactions1] t1 inner join [dbo].[prod_cat_info1] t2 on t1.[prod_cat_code] =
t2.[prod_cat_code]
             and t1.[prod_subcat_code] = t2.[prod_sub_cat_code]
group by [prod_subcat]
order by 2 desc
```

OUTPUT:

prod_subcat	% of sales	% of return
Women	12.8917878874334	14.5195957404741
Mens	12.6820517494295	12.4872929575256
Kids	8.8270984876016	9.37352567368825
Mobiles	4.60692261673398	4.42527565476655
Fiction	4.5780039267682	4.43727733898276

```
--11. for all custemors ages btw 25 to 35 years find what is the net total revenue generated bye
these customers in last 30 days of trans for max transaction data vailable in the data?

select t2.[cust_id] , DATEDIFF(yy, t1.[DOB], t2.[tran_date]) "age",
sum([total_amt]) as total_revenue
from [dbo].[Transactions1] t2 inner join [dbo].[Customer1] t1
on t1.[customer_Id] = t2.[cust_id]
where DATEDIFF(yy, t1.[DOB], t2.[tran_date]) between 25 and 35
group by t2.[tran_date], [cust_id], [DOB]
having DATEDIFF(dd, t2.[tran_date], (select max(t2.[tran_date]) from  [dbo].[Transactions1] t2))<=30
order by age



--12.

select [prod_cat] from
(select top 1 t1.[prod_cat], sum(t2.[total_amt]) as total_ret
       from [dbo].[prod_cat_info1] t1 , [dbo].[Transactions1] t2
              where t1.[prod_cat_code] = t2.[prod_cat_code]
              and t1.[prod_sub_cat_code] = t2.[prod_subcat_code]
              and t2.tran_date > DATEADD(MONTH, -3, (select Max(t3.tran_date) from
[dbo].[Transactions1] t3))
              and t2.total_amt<0
                     group by t1.[prod_cat]
                           order by sum(t2.total_amt) ) y


OUTPUT:

prod_cat
--------------------------------------------------
Books



--13.
select Store_type from (select t2.[Store_type], sum(t2.[total_amt]) [ttl_sales], sum(t2.[Qty])
[ttl_qty]
from [dbo].[prod_cat_info1] t1 , [dbo].[Transactions1] t2
where t1.[prod_cat_code] = t2.[prod_cat_code]
and t1.[prod_sub_cat_code] = t2.[prod_subcat_code]
and t2.[total_amt] > 0 and t2.[Qty] > 0
           group by t2.[Store_type]) y , (select max(x.ttl_amt) TAmt, max(x.ttl_qty) TQty
                           from (select t2.[Store_type], sum(t2.[total_amt]) ttl_amt,
                                                  sum(t2.[Qty])[ttl_qty]
                           from [dbo].[prod_cat_info1] t1 , [dbo].[Transactions1] t2
                           where t1.[prod_cat_code] = t2.[prod_cat_code]
                           and t1.[prod_sub_cat_code] = t2.[prod_subcat_code]
                           and t2.[total_amt] > 0 and t2.[Qty] > 0
                           group by t2.[Store_type]) x)z
where y.ttl_sales =  z.TAmt and y.ttl_qty = z.TQty
```

OUTPUT:

```
Store_type
--------------------------------------------------
e-Shop
```

```sql
--14.

select [prod_cat] from
(select t1.[prod_cat], avg(t2.[total_amt]) AVG_Rev from [dbo].[prod_cat_info1] t1 ,
[dbo].[Transactions1] t2
where t1.[prod_cat_code] = t2.[prod_cat_code] and t1.[prod_sub_cat_code] = t2.[prod_subcat_code] and
t2.[total_amt]>0
group by t1.[prod_cat] having avg(t2.[total_amt]) > (select avg([total_amt]) from
[dbo].[Transactions1] where [total_amt] > 0)) x
```

OUTPUT:

```
prod_cat
--------------------------------------------------
Bags
Books
Clothing
Electronics
```

```sql
--15.
select t1.[prod_cat], t1.[prod_subcat],
avg(t2.[total_amt]) [avg_revenue], sum(t2.[total_amt]) [ttl_revenue]
from  [dbo].[Transactions1] t2 inner join [dbo].[prod_cat_info1] t1
on t2.[prod_cat_code] = t1.[prod_cat_code] and t2.[prod_subcat_code] = t1.[prod_sub_cat_code]
where t2.[prod_cat_code] in (select top 5 t3.[prod_cat_code] from [dbo].[Transactions1] t3
group by t3.[prod_cat_code]
order by sum(t3.[Qty]) desc)
group by t1.[prod_cat], t1.[prod_subcat]
```

OUTPUT:

```
prod_cat                                           prod_subcat
avg_revenue            ttl_revenue
-------------------------------------------------- --------------------------------------------------
- ---------------------- ----------------------
Books                                             Academic
2125.48521033586       2055344.19839478
Electronics                                       Audio and video
2247.96000075941       2140057.92072296
Home and kitchen                                  Bath
2059.84961563215       2107226.15679169
Electronics                                       Cameras
2165.87853154991       2133390.35357666
Books                                             Children
2136.66750498601       2211450.86766052
Books                                             Comics
2037.68001891616       2100848.09950256
Electronics                                       Computers
2181.74983576892       2090116.34266663
Books                                             DIY
2108.37240019703       2085180.30379486
Books                                             Fiction
2140.22078679963       2232250.28063202
Home and kitchen                                  Furnishing
2084.00695634859       2098595.00504303
```

```
Clothing                                          Kids
2136.59914373483       2110959.95401001
Footwear                                          Kids
2125.99262194449       2145126.55554199
Home and kitchen                                  Kitchen
2008.95819427284       2083289.64746094
Clothing                                          Mens  (……)
2128.26314327709       2058030.45954895
(21 rows affected)
```