--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

--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

--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

--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

-----------------------------------------------------------------------------------------------------------------------------------------------------------------

--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

--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

--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

--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)

--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

--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

--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)