**Analytical case study**

**Q1**

**Query (1)**

select country,count(stockcode),rank()over(order by count(stockcode)desc )from online\_retail

group by 1

**Story :**

we find united kingdom have The largest number of goods of various kinds which led to an increase in the number of customers coming here.

**---------------------------------------------------------------------------------------------------**

**Query (2)**

select distinct description, country,sum(Quantity\*unitprice)

over(partition by description,country rows between unbounded preceding and

unbounded following ) as sales

from online\_retail

order by 3 desc;

**story:**

during analysis , we find the product that achieve high sales rate , we find united kingdom

occupies the first position for products may be in united kingdom There are many branches / stores.

**Query (3)**

select \*,rank()over(order by sales desc)

from(select distinct country,sum(quantity\*unitprice)over(partition by country

rows between unbounded preceding and unbounded following) as sales

from online\_retail

order by 2 desc

)as base

**Story:**

during analysis this data , I find the most sales in total happens in

united kingdom , and the sales in this country is very high than other countries.

**----------------------------------------------------------------------------------------------------**

**Query (4)**

select country ,extract ( month from cast(invoicedate as date))as month,max(quantity\*unitprice)

over(partition by cast(invoicedate as date),country) as sales

from online\_retail

where country='United Kingdom'

order by 3 desc

**story :**

relation between months and sales .. we find the sales is very high ,in December comparing with the other months ,It may be according to the type of product that the customer buys, it may be related to the seasons of the year.

**Query (5)**

select \*

from (select distinct description, unitprice, count(\*)over (partition by description)

from online\_retail

where country='United Kingdom'

order by 3 desc)as base

**Story:**

we find unit price of the most sale product changes many times..

may be this discount happen so this attractive.. we can prepare more offers

which lead to improve the business.

**Query (6)**

with invoice\_cust\_sales as

(

select distinct customerid,invoiceno

from online\_retail

where country='United Kingdom'

order by customerid)

select \* ,percent\_rank()over(order by no\_of\_item)\*100 as percentrank

from(

select distinct customerid,count(\*)over(partition by customerid)as no\_of\_item

from invoice\_cust\_sales)as base

**Story:**

we find top customers , who makes a lot of invoices , so it improve business and profit

can suggest make a lot of offers for this kind of customer.

**Q2**

segment each customer based on the below groups Champions - Loyal Customers - Potential Loyalists – Recent Customers – Promising - Customers Needing Attention - At Risk - Cant Lose Them – Hibernating – Lost.

with RFM as (

select customerid,

(SELECT MAX(TO\_DATE(invoicedate,'DD Mm YYYY') ) FROM online\_retail)-

MAX(TO\_DATE(invoicedate,'DD Mm YYYY') )as recency,

count(\*)as frequency,

sum(quantity\*unitprice)as monetary

from online\_retail

group by customerid order by customerid

),calc\_rfm as (

select cal\_rfm.\*,ntile(5) over(order by recency desc)as R\_score,

ntile(5) over(order by frequency )as F\_score,

ntile(5) over(order by monetary )as M\_score

from RFM cal\_rfm)

select r.\*,

(case when R\_score=5 and (F\_score+M\_score)/2=5 then 'Champion'

when R\_score=5 and (F\_score+M\_score)/2=4 then 'Champion'

when R\_score=4 and (F\_score+M\_score)/2=5 then 'Champion'

when R\_score=5 and (F\_score+M\_score)/2=2 then 'Potential Loyalists'

when R\_score=4 and (F\_score+M\_score)/2=2 then 'Potential Loyalists'

when R\_score=3 and (F\_score+M\_score)/2=3 then 'Potential Loyalists'

when R\_score=4 and (F\_score+M\_score)/2=3 then 'Potential Loyalists'

when R\_score=5 and (F\_score+M\_score)/2=3 then 'Loyal Customers'

when R\_score=4 and (F\_score+M\_score)/2=4 then 'Loyal Customers'

when R\_score=3 and (F\_score+M\_score)/2=5 then 'Loyal Customers'

when R\_score=3 and (F\_score+M\_score)/2=4 then 'Loyal Customers'

when R\_score=5 and (F\_score+M\_score)/2=1 then 'Recent Customers'

when R\_score=4 and (F\_score+M\_score)/2=1 then 'Promising'

when R\_score=3 and (F\_score+M\_score)/2=1 then 'Promising'

when R\_score=3 and (F\_score+M\_score)/2=2 then 'Customers Needing Attention'

when R\_score=2 and (F\_score+M\_score)/2=3 then 'Customers Needing Attention'

when R\_score=2 and (F\_score+M\_score)/2=2 then 'Customers Needing Attention'

when R\_score=2 and (F\_score+M\_score)/2=5 then 'At Risk'

when R\_score=2 and (F\_score+M\_score)/2=4 then 'At Risk'

when R\_score=1 and (F\_score+M\_score)/2=3 then 'At Risk'

when R\_score=1 and (F\_score+M\_score)/2=5 then 'Cant Lose Them '

when R\_score=1 and (F\_score+M\_score)/2=4 then 'Cant Lose Them '

when R\_score=1 and (F\_score+M\_score)/2=2 then 'Hibernating '

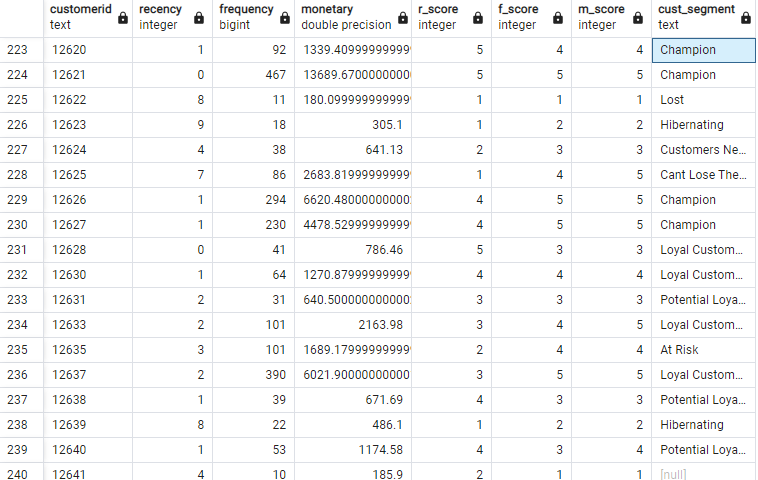
when R\_score=1 and (F\_score+M\_score)/2=1 then 'Lost'

end )as cust\_segment

from calc\_rfm r

order by customerid

**output**

****

**🡪 in last column in last row , the customer segment is null**

**Because the range in table not contains value of 2,1 for r\_score and avg f&m score.**