--frequency

select [order\_status],

Count( [order\_status] ) AS Frequency

  FROM [dbo].[F\_ORDER]

  GROUP BY [order\_status]

  ORDER BY COUNT([order\_status] ) DESC

--AVG and Total REvenue--

select avg([total\_ex\_tax]- [shipping\_cost\_ex\_tax]) as avgrevenue,  sum([total\_ex\_tax]- [shipping\_cost\_ex\_tax]) as totalrevenue, [month], [year], [month\_number]

from [dbo].[D\_DATE] d, [dbo].[F\_ORDER] f

where f.[order\_date\_key]=d.[date\_key]

group by [month], [year], [month\_number]

order by [year], [month\_number]

-- cast to convert decimal to numeric

select cast(avg([total\_ex\_tax]- [shipping\_cost\_ex\_tax])as numeric) as avgrevenue,  cast(sum([total\_ex\_tax]- [shipping\_cost\_ex\_tax]) as numeric) as totalrevenue, [month], [year], [month\_number]

from [dbo].[D\_DATE] d, [dbo].[F\_ORDER] f

where f.[order\_date\_key]=d.[date\_key]

group by [month], [year], [month\_number]

order by [year], [month\_number]

--Drop Time from Date--

select convert (date, [the\_date], 111)

from [dbo].[D\_DATE]

---http://www.w3schools.com/sql/func\_convert.asp---

--Customers with more than one order--

select [customer\_key], count(distinct [order\_ID]) as number\_of\_orders

from [dbo].[F\_ORDER] f, [dbo].[D\_DATE] d

where f.[order\_date\_key]=d.[date\_key]

and [year]=2015

and [order\_status]='Shipped'

group by customer\_key

having count (distinct  [order\_ID])>1

---avg number of orders of all the repeat customers--

select avg(a.number\_of\_orders)

from (select [customer\_key], count(distinct [order\_ID]) as number\_of\_orders

from [dbo].[F\_ORDER] f, [dbo].[D\_DATE] d

where f.[order\_date\_key]=d.[date\_key]

and [year]=2015

and [order\_status]='Shipped'

group by customer\_key

having count (distinct  [order\_ID])>1) a

--create table

use kaggle;

go

create table trans\_kag(

id bigint,

chain bigint,

dept bigint,

category bigint,

company bigint,

brand int,

date1 date,

productsize float,

productmeasure varchar(2),

purchasequantity bigint,

purchaseamount float

)

go

-calculations

select sum([total\_ex\_tax]-[shipping\_cost\_ex\_tax]) as totalrevenue , [month], [year], [month\_number]

from [dbo].[D\_DATE] d, [dbo].[F\_ORDER] f

where f.[order\_date\_key]=d.[date\_key]

--and [store\_credit\_amount] < 0

group by [month], [year], [month\_number]

order by [year], [month\_number]

--rounding

round and convert to big int

cast(ROUND([purchaseamount],0) as bigint ) as [purchaseamount]

--rounding

sql round

SELECT ROUND([purchaseamount],0)

from [dbo].[kaggle80]

--date format

select

 right('0000' + cast(datepart(year,[date1]) as varchar(4)), 4)

+ right('00' + cast(datepart(month, [date1]) as varchar(2)), 2)

from  [dbo].[kaggle80]

--create view

create view [kaggle80] as

SELECT TOP 900000 [id]

,[chain]

,[dept]

,[category]

,[company]

,[brand]

--,convert (date, [date1], 111) as yyyymm

,[date1]

,year([date1]) as year\_

,month([date1]) as month\_

--,LEFT(CONVERT(varchar, GetDate(),112),6) as date2

,cast(ROUND([productsize],0) as bigint ) as [productsize]

,[productsize] as [productsize\_original]

,[productmeasure]

,[purchasequantity]

,cast(ROUND([purchaseamount],0) as bigint ) as [purchaseamount]

, right('0000' + cast(datepart(year,[date1]) as varchar(4)), 4) + right('00' + cast(datepart(month, [date1]) as varchar(2)), 2) as [yyyymm]

FROM [kaggle].[dbo].[transactions]

drop view [kaggle80]

select \*

From [kaggle80]

select sum([purchasequantity]\*[purchaseamount]) as totalrevenue, sum([purchasequantity]) as totalquantity, avg([purchaseamount]\*[purchasequantity]) as averagepurchase, [date1]

from [dbo].[kaggle80]

group by [date1]

order by [date1]

---year format

--Select Query for MS Azure

SELECT Format([year],'d4')+'/'+Format([month\_number], 'd2') as yyyymm

      ,Sum([total]) as Total

      ,Sum([subtotal]) as SubTotal

      ,Sum([subtotal\_ex\_tax])

      ,Sum([subtotal\_inc\_tax])

      ,Sum([subtotal\_tax])

      ,Sum([total\_ex\_tax])

      ,Sum([total\_inc\_tax])

      ,Sum([total\_tax])

      ,Sum([total\_shipping])

      ,Sum([shipping\_cost\_ex\_tax])

      ,Sum([shipping\_cost\_inc\_tax])

      ,Sum([shipping\_cost\_tax])

      ,Sum([total\_discount])

  FROM [dbo].[F\_ORDER] as FO, [dbo].[D\_DATE] as DD

  Where DD.[date\_key]=FO.[order\_date\_key]

  Group By Format([year],'d4')+'/'+Format([month\_number], 'd2')

  Order By Format([year],'d4')+'/'+Format([month\_number], 'd2')

--decimal to numeric

select cast(avg([total\_ex\_tax]- [shipping\_cost\_ex\_tax])as numeric) as avgrevenue,  cast(sum([total\_ex\_tax]- [shipping\_cost\_ex\_tax]) as numeric) as totalrevenue, [month], [year], [month\_number]

--into TempTABLE

from [dbo].[D\_DATE] d, [dbo].[F\_ORDER] f

where f.[order\_date\_key]=d.[date\_key]

group by [month], [year], [month\_number]

order by [year], [month\_number]

--obs 10

Select \*

From hivesample table limit 10;

--row count

Select count(\*) from weblogs;