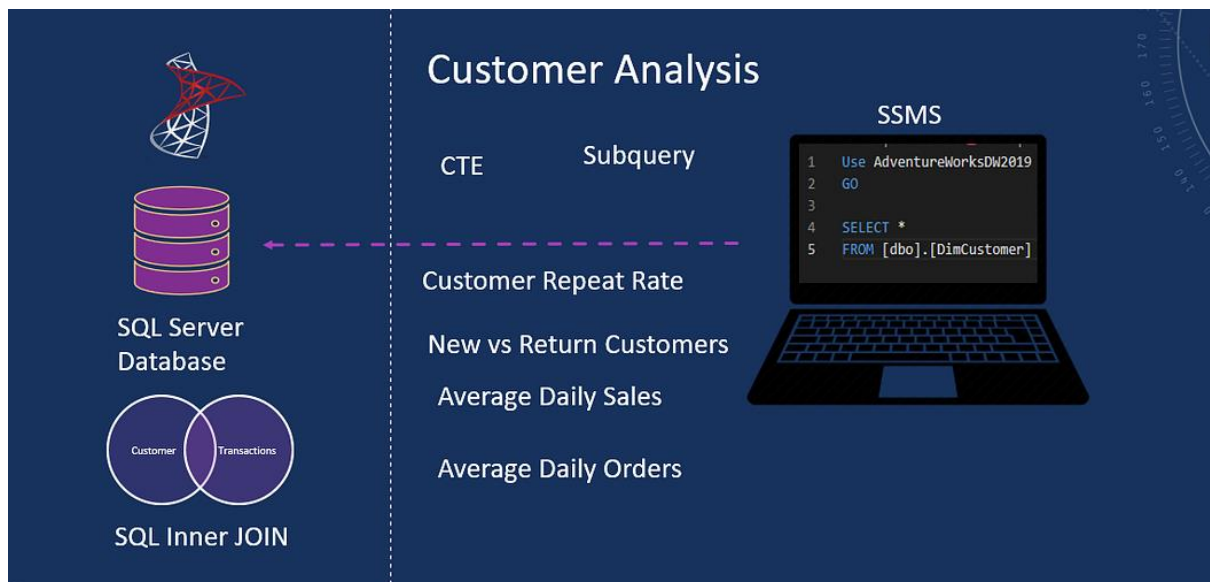


SQL Series: Customer and Sales Analysis using SQL



SQL Customer Analysis

Customer and customer behavior analysis is critical for business development. You should be familiar with your customer's demographic and their buying patterns. This can help you grow your business and keep in demand items in stock. Attracting new customers is only part of the story. The key to success is the Customer retention. This is a key metric and it plays big part in a business success. Therefore, it is essentials to have customer retention metrics handy. We will calculate the customer retention metrics along with sales using SQL.

It is easy to address a problem if you know it exists. If you have customer retention metrics clearly defined then it can help you identify if your business is struggling to retain existing customers. It is helpful to have a flag that identifies new vs repeat customer on the fly. Using this flag we can calculate the customer retention rate.

We will focus on the following key business aspects.

- Find out how many customers have single vs repeat purchases. Define new vs returning customers flag?
- What is our user repeat rate?
- On average, how many orders do we receive per day?
- What's our average daily sales?

New vs Returning Customer

Customer retention is crucial to any business looking to understand its customers and it can help explain the sales figures or the impact of a marketing campaign. The first step to calculating this metric is to have a flag that identifies the new and the returning customer. Or customers with single and repeat purchases. We can calculate this column using SQL. We create a grouping of the customers based on the customer's transactions.

```
with customers as (  
select  
CustomerKey,  
ROW_NUMBER() Over (partition by CustomerKey Order by CustomerKey) as rown_number  
FROM[dbo].[FactInternetSales] )
```

Based on this we can easily categorize new and the returning customers. We can filter on the “row_number” column and when it is greater than one then we will label this returning customer. These are the customers who have made more than one purchase. Similarly, we can set the filter to “row_number <= 1” and we will assume these are the new customers with a single purchase.

```
repeat_buyers as (  
select distinct  
CustomerKey  
FROM customers  
where row_number > 1), single_purchase as (  
select distinct
```

```
CustomerKey
FROM customers
where row_number <= 1)
```

Now we have the base to create a flag that identifies new and the repeat customers. We join the above CTE's on the customer key. "Single_purchase" CTE contains all customers while "repeat_buyers" only includes customers with multiple orders. We check if the *CustomerKey* from the "repeat_buyers" is not missing then we set the "repeat_new" column to repeat otherwise, we set this to new. This gives us the new vs returning customer flag. Also, we calculate the number of customers with the count function. This displays the number of new and returning customers.

```
select
case
  when r.CustomerKey IS NOT NULL
  then 'repeat'
  else 'new'
end as repeat_new,
count(*) as number_of_customer
FROM single_purchase f
left join repeat_buyers r on f.CustomerKey = r.CustomerKey
group by
case
  when r.CustomerKey IS NOT NULL
  then 'repeat'
  else 'new'
end
```

To calculate the customer's retention rate. We divide the new and repeat customer count with the total customers. We are calculating the total customers with the help of subquery from the initial "customers" CTE. This gives us the customer retention rate. In this case, the retention rate is 86.52% which is great. It is a good indicator that business is able to retain the lion share of the customers and it is a recipe for success.

```
FORMAT(cast(count(*) as decimal(18,2))/cast((select count(distinct customerkey) from
customers) as decimal(18,2)), 'P') as repeat rate
```

The next item on the agenda is to check how many orders we receive on a daily basis. And based on these orders what's the daily average sales. We can obtain these two metrics from the fact table. We define a CTE to give us a unique order count and total sales.

```
with ave_sales_per_day as (  
select OrderDate,  
count(distinct SalesOrderNumber) as order_count,  
sum(SalesAmount) as Sales  
FROM [dbo].[FactInternetSales]  
group by OrderDate)
```

In the select following the CTE we apply the average function on the sales and order count columns. This gives us the average daily sales and average orders per day.

```
select  
avg(Sales) as avg_sales_day,  
avg(order_count) avg_order_count  
FROM ave_sales_per_day
```

That's it folks. We have successfully created a flag to distinguish between new and repeat customers. Which in turn helped us calculate the customer retention rate. We have calculated the daily averages to check how the business is performing on a daily basis. This is all for now.

Conclusion

- We have identified and created key business metrics to help identify how the business is performing.
- We have successfully created a new and repeat customers flag.
- We calculated the customer retention rate.
- We calculated the daily averages to see how business is performing on a daily basis.