GROUP BY & HAVING

It's important to know when to use GROUP BY and HAVING by what your intended output will result. **You use HAVING to filter the GROUP BY. You can use a WHERE clause (filter) on columns before the GROUP BY to serve as a range.**

Problem: Take for instance a simple query such as this and you want to sum up by CustomerID the 11000's, 12000's, and 13000's (**douplicate CustomerID's**) and so on:

```
USE AdventureWorks2014
GO

SELECT CustomerID, TaxAmt AS CustomerTax
FROM Sales.SalesOrderHeader
ORDER BY CustomerID
```

```
CustomerID    CustomerTax
-----------------------
11000  271.9992
11000  187.3576
11000  200.5624
11001  269.9992
11001  193.5944
11001  47.1168
11002  271.9992
11002  183.5992
11002  193.5248
11003  271.9992
11003  185.5168
11003  193.6272
11004  271.9992
11004  190.1568
11004  193.5248
11005  269.9992
11005  188.9816
11005  190.7256
11006  271.9992
11006  186.7976
11006  190.7256
11007  271.9992
11007  191.356
11007  193.5248
11008  269.9992
11008  184.9808
11008  193.5248
11009  269.9992
11009  183.7824
11009  193.5248
--Continues…
```

```
SELECT CustomerID, SUM(TaxAmt) AS CustomerTax
FROM Sales.SalesOrderHeader
GROUP BY CustomerID
ORDER BY CustomerID
```

Note: When we use GROUP BY we used the SUM function and merged our duplicates into one row (In this case, CustomerID). It's important to note that, you must include all columns in the GROUP BY clause before the aggregated summed column.

```
CustomerID    CustomerTax
-----------------------
11000  659.9192
11001  510.7104
11002  649.1232
11003  651.1432
11004  655.6808
11005  649.7064
11006  649.5224
11007  656.88
11008  648.5048
11009  647.3064
--Continues to 11999

12000  654.64
12001  649.0408
12002  651.6264
12003  660.52
12004  649.1232
12005  558.7848
12006  11.5792
12007  561.9024
12008  6.42
12009  655.64
--Continues to 12999

13000  543.956
13001  543.9543
13002  23.5952
13003  18.42
13004  15.2216
13005  30.9568
13006  540.9719
13007  453.6992
13008  10.0408
13009  13.4936
--Continues on...
```

We can use GROUP BY to Sum up the rows in groups as shown and filter using HAVING:

```
SELECT CustomerID, SUM(TaxAmt) AS CustomerTax
FROM Sales.SalesOrderHeader
--You could put a WHERE clause here if needed.
GROUP BY CustomerID
HAVING SUM(TaxAmt) > 2000
ORDER BY CustomerID
```

Result:

```
CustomerID CustomerTax
-----------------------
29484       12725.2865
29485       11040.8575
29486       49933.5423
29487       7347.8108
29488       21817.206
29489       36492.5716
29491       8383.5414
29492       13633.7763
29496       9533.6265
29497       67870.8634
--Continues on...
```

Here, we are using GROUP BY, the WHERE Clause and HAVING keywords. Keep in mind that our WHERE clause is above the GROUP BY keyword and that the HAVING filter is only used with GROUP BY.

```
SELECT CustomerID, SUM(TaxAmt) AS CustomerTax
FROM Sales.SalesOrderHeader
WHERE TaxAmt > 3000 AND TaxAmt < 4000 --Filtered before grouping.
GROUP BY CustomerID
HAVING SUM(TaxAmt) > 21000 --Range of Summed TaxAmt.
ORDER BY CustomerID
```

Result:

```
CustomerID CustomerTax
-----------------------
29734      24253.3731
29834      23758.9157
29935      21316.6836
29950      21084.0621
30065      21968.3595
```