## Thomson Deals Business Intelligence Percentage and Ranking Calculation

*By Chris Zhao*

In mathematics, percentage calculation is dependent not only on its case scenario, but also on its perspective. A perspective represents the angle or the view you interpret the percentage.

Let us use a simple example to illustrate the points. Suppose we want to calculate gender ratio among working professionals. Specifically we want to calculate female percentage of working professionals. Assume the dimensions (parameters) that we are looking at include States / Cities, Industries, Job functions, and Employers.

Let us say, you are reviewing the percentage data, and you pull States into your rows and Industries into your columns. You locate the cell at the intersection with Row = ‘NY’ and Column = ‘Financial’, and you see the value at the cell is 35%.

There will be at least three different perspectives to interpret this percentage:

* the ratio of female working professionals in NY state and in Financial industry across all of the states and all of the industries;
* the ratio of female working professionals in NY state and in Financial industry across all of the industries in NY state;
* the ratio of female working professionals in NY state and in Financial industry across all of the states in Financial industry

Generally speaking, these three percentages should have different values, because they are logically different, even if sometimes they may coincide with each other. There wouldn’t be a single formula that meets all the three different logics. If one formula is defined based on one perspective, and your expectation falls into a different perspective, you will see that the percentage doesn’t look right.

To make the matter more complicated, if you filter by states and industries, and only choose two states – say NY and NJ - for rows, and two industries – say Financial and Health - for columns, then you will have another three perspectives of interpretation in addition to the three perspectives described above:

* the ratio of female working professionals in NY state and in Financial industry across the two states and the two industries selected;
* the ratio of female working professionals in NY state and in Financial industry across the two industries in NY state;
* the ratio of female working professionals in NY state and in Financial industry across the two states in Financial industry

Currently, the existing set of SOW and Market Share in Production follows the original definition of perspective. Their names are listed below.

* [Share Of Wallet (Must be used with Dimension Bank - Bank Name)]
* [Share Of Wallet EURO (Must be used with Dimension Bank - Bank Name)]
* [Share Of Wallet YEN (Must be used with Dimension Bank - Bank Name)]
* [Share Of Wallet STG (Must be used with Dimension Bank - Bank Name)]
* [Market Share (Must be used with Dimension Bank - Bank Name)]
* [Market Share EURO (Must be used with Dimension Bank - Bank Name)]
* [Market Share YEN (Must be used with Dimension Bank - Bank Name)]
* [Market Share STG (Must be used with Dimension Bank - Bank Name)]

The definition of these SOW% and Market Share follow the same logic, except for the difference in the currency and difference between Bank Fees and Deal Amount. SOW% is defined using Bank Fees, whereas Market Share is defined using Deal Amount.

**Definition of the old SOW%**

Let us look at the precise logic (perspective) of the original definition, using

[Share Of Wallet (Must be used with Dimension Bank - Bank Name)]

as an example. Let’s simply call it SOW%. This is defined as

([Measures].[Fees in USD]) / ([Banks].[Bank Name].[All], [Measures].[Fees in USD])

Namely, the ratio of

*Bank fees (in USD) earned by bank(s) implied by* ***the context***

*divided by*

*Bank fees (in USD) earned by all the banks within* ***the context***

Pay special attention to the phrase “the context”. For example, for each cell in a Pivot Table, **the context is determined by the filters selected in the filter area, the row and the column the cell is located**. The context always resolves to a particular bank, a set of banks, or all the banks. If it resolves to a particular bank, then SOW% is equal to

*Bank fees (in USD) earned by the particular bank implied by* ***the context***

*divided by*

*Bank fees (in USD) earned by all the banks within* ***the context***

If it resolves to a set of banks, then SOW% is equal to

*Bank fees (in USD) earned by the set of banks implied by* ***the context***

*divided by*

*Bank fees (in USD) earned by all the banks within* ***the context***

If it resolves to all the banks, then SOW% is equal to

*Bank fees (in USD) earned by all the banks within* ***the context***

*divided by*

*Bank fees (in USD) earned by all the banks within* ***the context***

Obviously this is 100%. If your context references no bank, for example, your filters, rows and columns do not reference Bank dimension at all, then the context resolves to all the banks. In this case, your SOW% would be 100%.

**Definition of Rank Order**

Now let us look at the definition of Rank Order. There are four rank orders defined.

* [Rank Order (Fees in USD)]
* [Rank Order (Fees in Euro)]
* [Rank Order (Fees in STG)]
* [Rank Order (Fees in Yen)]

They all follow the same logic, and their only differences lie in the currency. So we only need to talk about the first one. It is defined as

Rank([Banks].[Bank Name].CurrentMember, [Banks].[Bank Name].[Name].members, ([Measures].[Fees in USD]))

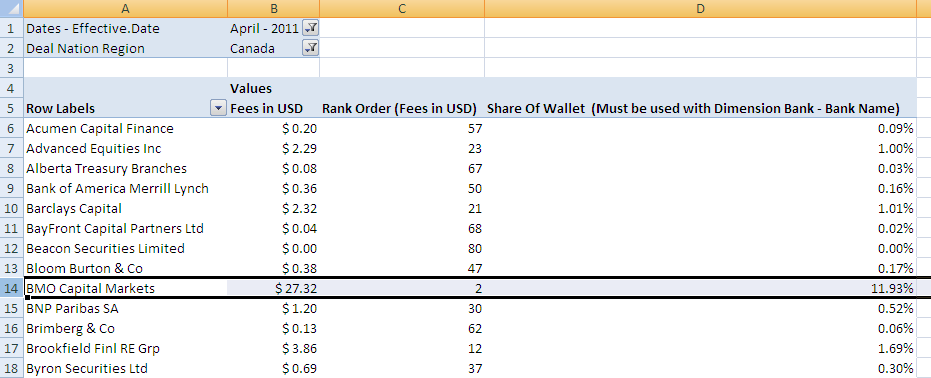
Let us try to understand this. Because of the reference of [Banks].[Bank Name].CurrentMember, your context must resolve to a particular bank. In this formula, the particular bank implied by your context is ranked among all the banks according to the bank fees earned by each bank - the measure [Measures].[Fees in USD].

If your context references more than one particular banks or references no bank at all, for example, in the pivot table filter area, you select two banks, or your filters, rows and columns do not reference Bank dimension at all, then you will not be able to use this rank order in your context.

**Interpretation of SOW% and Rank Order**

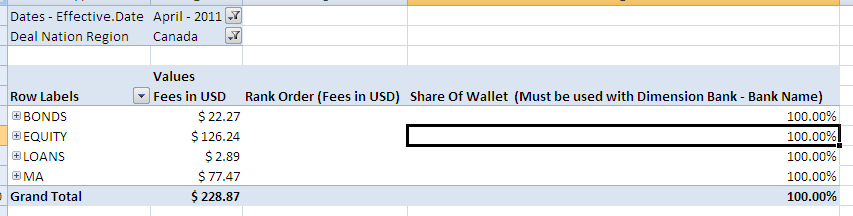
Let us use some example to illustrate how to understand and interpret the SOW% and Rank Order.

Let us say, you select the Bank dimension and the bank name hierarchy as the rows, and filter by Deal Nation Region and Effective Date, and you are looking at Fees in USD, Rank Order and SOW%. See the screenshot below.



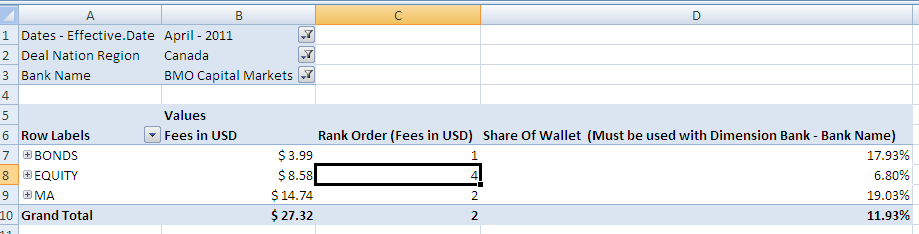
This would be a **perfect case** to use Rank Order and SOW%, because each cell within the context resolves to a specific bank. (Because of various cases which don’t fall into this perfect case, the need for a generic SOW% arises. See below.)

Now you replace Bank Name with Product Group in the row, and let us see what we get:



You get meaningless SOW%, and you don’t get Rank Order. That’s because each cell in the context does not resolve to a specific bank. In fact, it resolves to all the banks. That’s why you see 100% as the SOW%.

Now let us put an extra filter in the filter area, and filter by a particular bank – say BMO Capital Markets, you then get this screen:

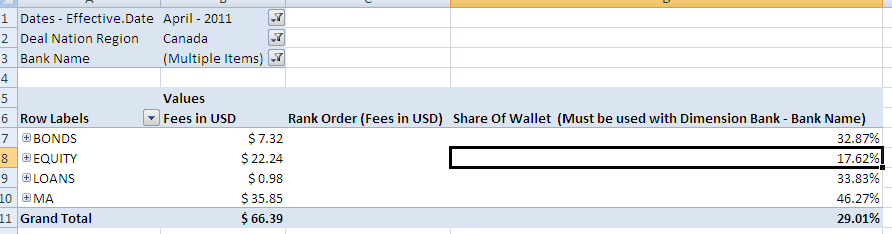


If you say: “Wait a minute. EQUITY has $8.58 million, more than double of BONDS, how come it is ranked 4 and has lower SOW% than BONDS? This doesn’t look right.”

Then you have a wrong perspective and interpreted the Rank Order and SOW% wrong.

The rank 4 means that BMO Capital Markets ranked number 4 among all the banks in the EQUITY deal category within the context (i.e. in April 2011 in Canada), and all the fees earned by BMO Capital Markets in the EQUITY deal category is 6.80% of all the fees earned by all the banks in the EQUITY deal category within the context (i.e. in April 2011 in Canada).

Now let us change the Bank Name filter to include two banks: BMO Capital Markets and RBC Capital Markets. You then get this:



Now you don’t get any rank order, because the rank order applies only to a particular bank. But you still get SOW%, although you may feel that these percentages do not make sense. “The Bank Fees in the Grand Total line is bigger. How could its SOW% be less than some SOW% at the row level?”

Read on, you will see that these percentages do make sense.

For example, 17.62% is the percentage of all the fees earned by BMO Capital Markets and RBC Capital Markets combined in the EQUITY deal category relative to all the fees earned by all the banks in the EQUITY deal category within the given context (i.e. in April 2011 in Canada). The reason why the SOW% at the Grand Total line is even less than the SOW% on some rows, is because 29.01% is the percentage of all the fees earned by BMO Capital Markets and RBC Capital Markets combined in all the deal categories relative to all the fees earned by all the banks in all the deal categories within the given context (i.e. in April 2011 in Canada).

Would the female percentage among working professionals across all industries be greater than the female percentage among working professionals in a specific industry? Not necessarily. Think about nursing industry.

That explains why SOW% at the Grand Total line could be less than some SOW% at individual rows.

**The Need for generic SOW%**

I have illustrated various case scenarios through simple examples above. On these various case scenarios, the SOW% would fall off what you expected, and then you could not understand these percentages.

Therefore a new need arises. That is, a generic SOW% which can be applied to whatever rows and filters selected by user.

Since the user can select any dimension hierarchies or attributes for the filters, rows, and columns. We need to create dynamic calculation that responds to the context selected by user.

To do so, we first create a dynamic named set, which defines the rowset. A rowset is a set of tuples that are existing at rows within the constraint of filters selected in the filter area and at row level. Then we aggregate the bank fees across this dynamic named set within the context defined by user.

We are just trying to generalize or stretch our percentage calculation as far as possible to cover as many case scenarios as possible. And we call it a generic percentage. But any percentage calculation implements one perspective only. If you use a perspective to interpret a percentage that is defined based on a perspective different from yours, you’ll think the percentage is off.

Currently there are eight types of generic percentage defined, and they are:

* [Share Of Wallet USD]
* [Share Of Wallet EURO]
* [Share Of Wallet YEN]
* [Share Of Wallet STG]
* [Market Share USD]
* [Market Share EURO]
* [Market Share YEN]
* [Market Share STG]

These have been pushed to the main cube on QA box and 37 box.

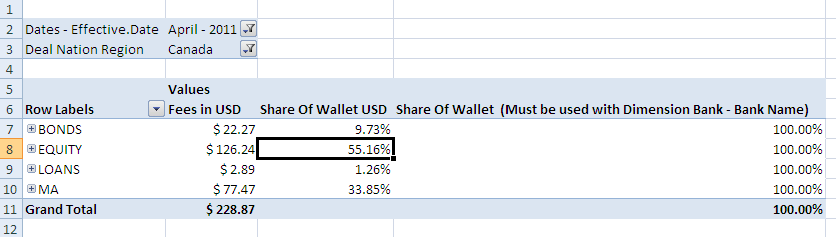
**Interpretation of generic SOW% & Comparison between the generic SOW% and the old SOW%**

Let us use the same examples illustrated above to compare the old and the generic SOW%.

Again, let us say, you select the Bank dimension and the bank name hierarchy as the rows, and filter by Deal Nation Region and Effective Date, and you are looking at Fees in USD, the generic SOW% and the old SOW%. See the screenshot below.

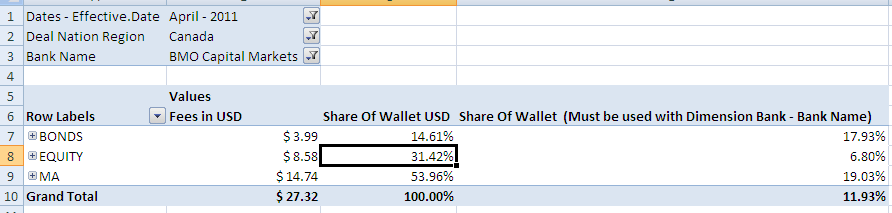


You will see that, in this perfect case, the generic and the old SOW% are identical. Now you replace Bank Name with Product Group in the row, and let us see what we get:



The old SOW% does not tell you anything. However, the generic SOW% does tell you something meaningful. For example, 55.16% is the percentage of all the fees earned by all the banks in the EQUITY deal category relative to all the fees earned by all the banks in all the deal categories within the given context (i.e. in April 2011 in Canada).

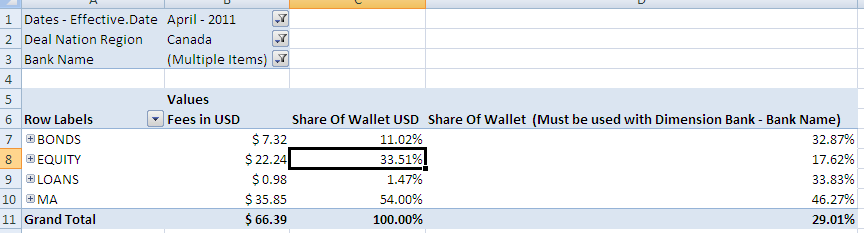
Now let us put an extra filter in the filter area, and filter by a particular bank – say BMO Capital Markets, you then get this screen:



You will have a hard time to understand and to compare the two percentages, since they are quite different.

Let us look at the row of EQUITY, and try to interpret the percentages. 31.42% is the percentage of all the fees earned by BMO Capital Markets in the EQUITY deal category relative to all the fees earned by BMO Capital Markets in all the deal categories within the given context, whereas 6.80% is the percentage of all the fees earned by BMO Capital Markets in the EQUITY deal category relative to all the fees earned by all the banks in the EQUITY deal category within the given context (i.e. in April 2011 in Canada).

Now let us change the Bank Name filter to include two banks: BMO Capital Markets and RBC Capital Markets. You then get this:



Again, let us look at the row of EQUITY, and try to interpret the percentages. 33.51% is the percentage of all the fees earned by BMO Capital Markets and RBC Capital Markets combined in the EQUITY deal category relative to all the fees earned by BMO Capital Markets and RBC Capital Markets combined in all the deal categories within the given context, whereas 17.62% is the percentage of all the fees earned by BMO Capital Markets and RBC Capital Markets combined in the EQUITY deal category relative to all the fees earned by all the banks in the EQUITY deal category within the given context (i.e. in April 2011 in Canada).

So now you see, **understanding a percentage always depends the perspective that you are using to interpret the percentage**. The definition of any percentage can only follow one perspective. It’s important to understand the perspective a percentage is defined on before you interpret it.