Crosstab

Alberto Ghedin
2011/09/19 14:20

Table of Contents

Crosstab

This functionality is available since SpagoBI 2.7.0.
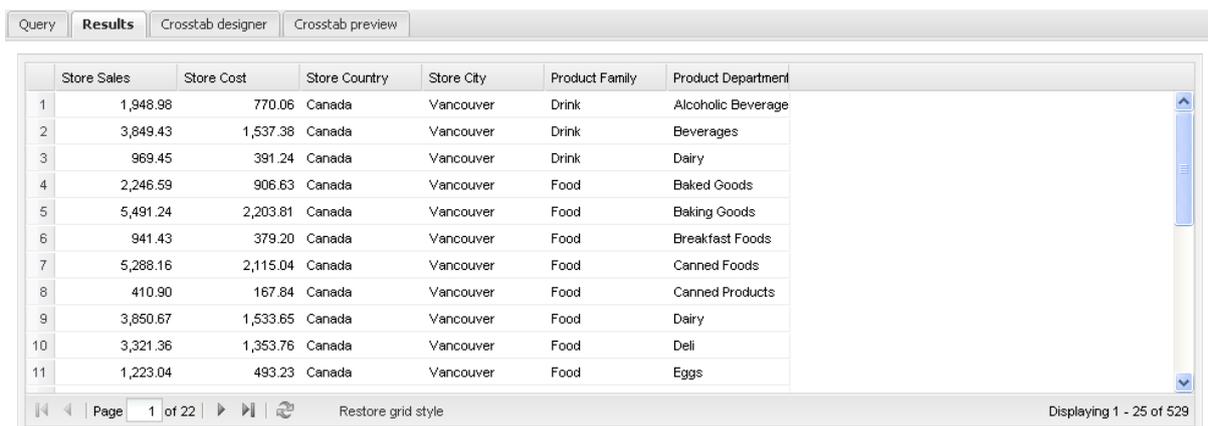
1 Use guide

Suppose you build a query that returns the sum of Store sales and the sum of Store costs grouped by store (country and city) and by product (family and department) like in figure 1:
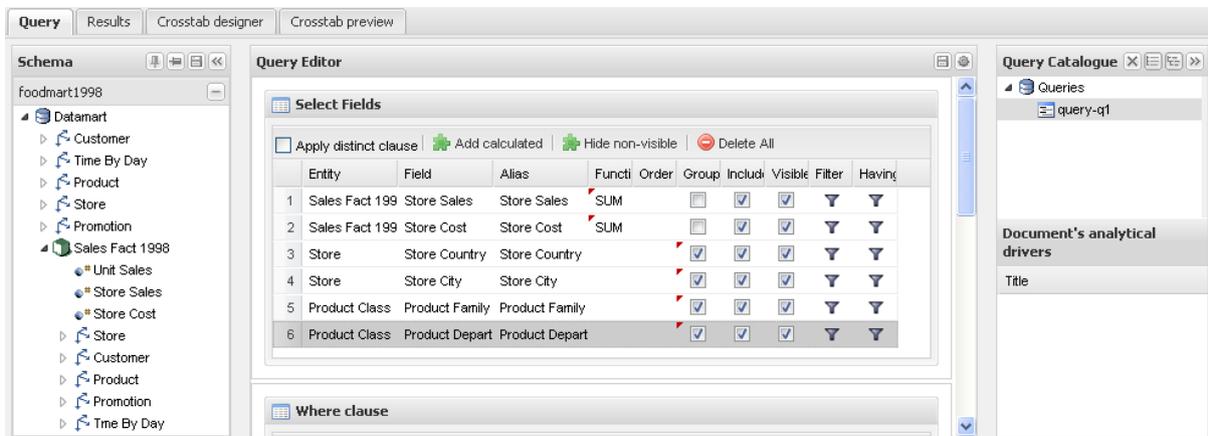
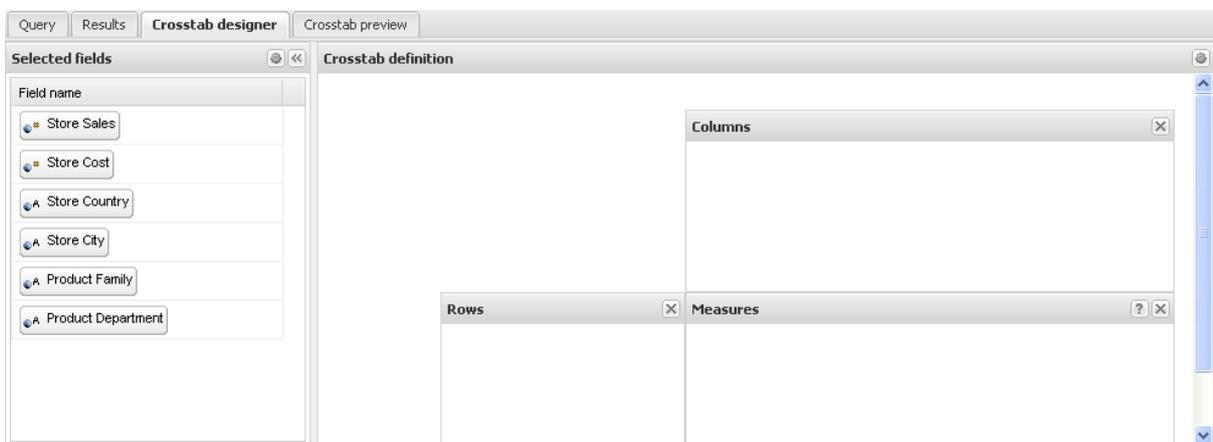… then you execute it ...

Now we want to define a crosstab displaying Store sales and Store costs, putting store's country and city into the columns and product's family and department into the rows. Click on the third Qbe tab, 'Crosstab designer':
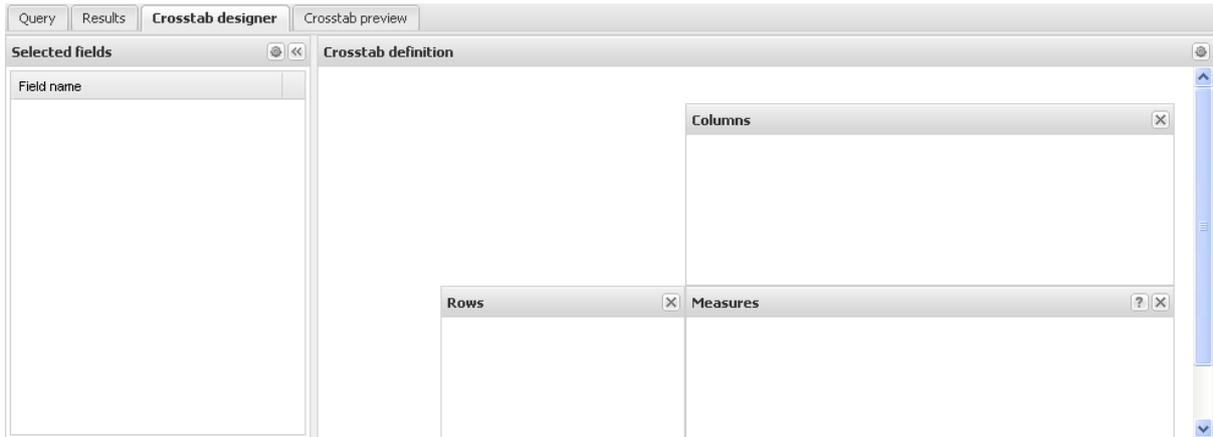
then retrieve the query fields clicking on the gear button on the "Selected fields" panel:

Then, drag and drop the measures and the grouping properties from the "Selected fields" panel into the crosstab schema on the right:
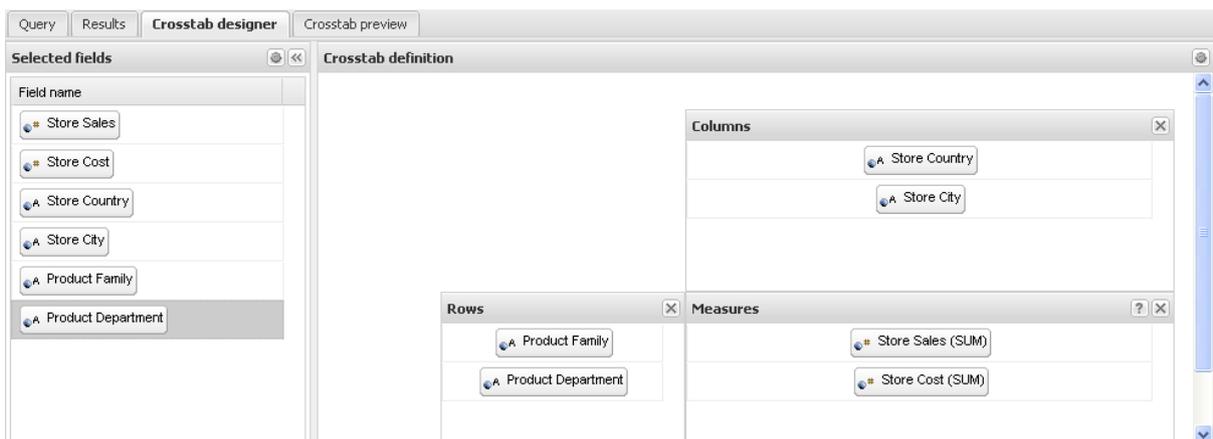
Note that the crosstab is produced by a 2-steps process:

1. base query (the one defined on figure 1) execution against the datamart;
2. crosstab query execution against the results of the base query.

Using the crosstab designer you are defining the crosstab query (second step), that can return measures with a higher granularity (than the base query); besides crosstab query can use different aggregation functions: as an example, if the base query is using COUNT aggregator, may be you want to use the SUM aggregator in the crosstab query. By default, the engine will consider,

for the crosstab query, the same aggregation function on measures as the ones used in the base query, but you can change the aggregation function double-clicking on the measure in the crosstab schema:
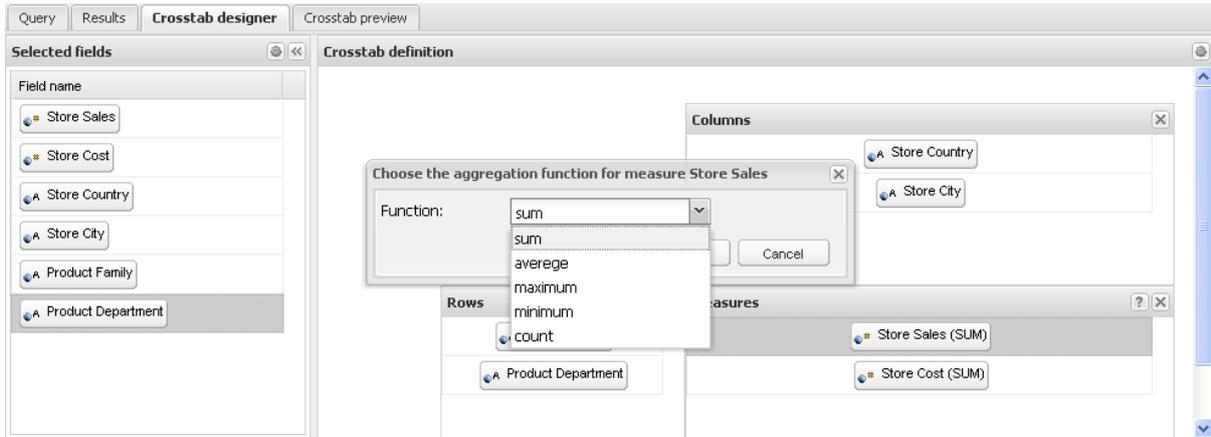
Clicking on the "?" button, you can also set some other properties of the crosstab: display measures on rows or columns, display totals or sub-totals on rows/columns:
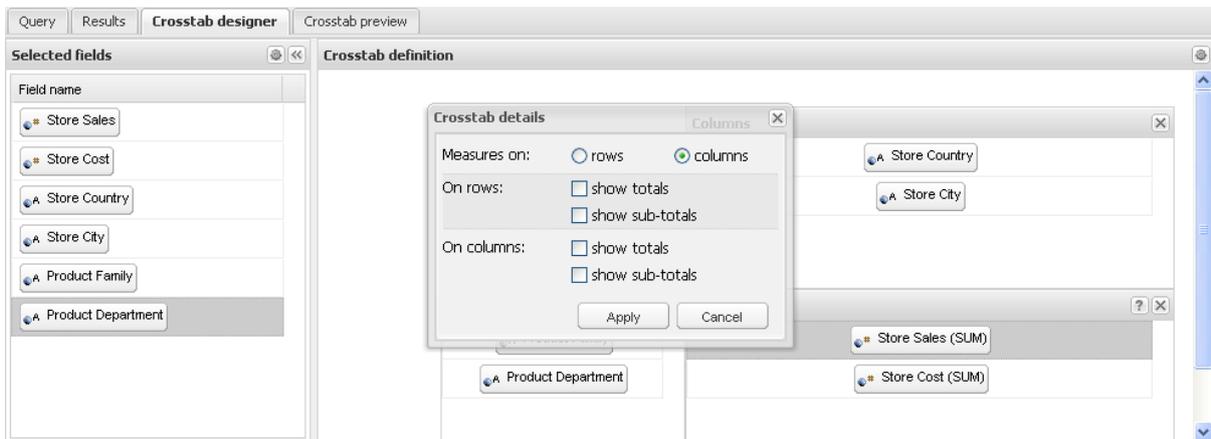
Then you can click on the gear button on the top-right corner and the crosstab will be loaded and displayed:

Then you can export the crosstab result into a XLS file, by opening export menu (on document execution toolbar), "Crosstab" submenu.

Example of crosstab displaying totals:

	Drink		Food		Non-Consumable		Total	
	Unit Sales	Store Sales	Unit Sales	Store Sales	Unit Sales	Store Sales	Unit Sales	Store Sales
Canada	4,354.00	8,669.84	32,691.00	70,028.98	9,112.00	19,346.64	46,157.00	98,045.46
Mexico	18,483.00	36,890.31	146,642.00	311,041.43	38,789.00	82,361.85	203,914.00	430,293.59
USA	24,117.00	48,182.01	186,590.00	397,065.39	49,209.00	105,561.02	259,916.00	550,808.42
Total	46,954.00	93,742.16	365,923.00	778,135.80	97,110.00	207,269.51		

Example of crosstab displaying sub-totals:

		Drink	Food	Non-Consumable	Total
		Unit Sales	Unit Sales	Unit Sales	Unit Sales
Canada	Vancouver	3,398.00	25,975.00	7,270.00	36,643.00
	Victoria	956.00	6,716.00	1,842.00	9,514.00
	Total	4,354.00	32,691.00	9,112.00	46,157.00
Mexico	Acapulco	2,299.00	16,428.00	4,499.00	23,226.00
	Camacho	2,088.00	17,205.00	4,459.00	23,752.00
	Guadalajara	197.00	1,514.00	413.00	2,124.00
	Hidalgo	4,236.00	34,499.00	9,015.00	47,750.00
	Merida	3,377.00	26,568.00	7,198.00	37,143.00
	Mexico City	933.00	7,880.00	2,067.00	10,880.00
	Orizaba	2,313.00	17,812.00	4,571.00	24,696.00
	San Andres	3,040.00	24,736.00	6,567.00	34,343.00
	Total	18,483.00	146,642.00	38,789.00	203,914.00
	Bellingham	174.00	1,478.00	332.00	1,984.00
	Beverly Hills	2,180.00	16,360.00	4,167.00	22,707.00
	Bremerton	2,227.00	17,259.00	4,583.00	24,069.00
	Los Angeles	2,302.00	17,223.00	4,536.00	24,061.00

On crosstab preview panel you can modify the crosstab itself:

1. you can change the position on the headers (both column headers and row headers) if the headers belong to the same level;
2. you can adjust the heigth of the column headers and the width of the row headers (just move the mouse into the header's border; if you are using IE, you have to move the mouse on the top right corner of the column header);
3. click on the header and a context menu will appear with the following options:

1. remove the calculated field (if it is a calculated field, see below)
2. modify the calculated field (if it is a calculated field, see below)
3. add a calculated field (see below)
4. hide the header: it will hide the selected header and its descendants
5. hide the header of that type: it will hide the selected header and all the other headers with the same name at the same level (and their descendants)
6. show hidden headers (a submenu will display all hidden headers)
7. show/hide measures

2 Calculated fields

If you select the "Add calculated field" option from the header's context menu, a wizard will appear: you have to fill the name and the expression of the calculated field. Type the calculated-field's name by hand; click on the headers (belonging to the same level of the header clicked when you opened the context menu) in order to define the operands (note that the final expression should be a javascript expression, therefore you can use operators such as "+", "*", "/", "-" and brackets).

The following figure shows 3 different calculated fields:

1. Profit, that is the difference between Store sales and Store costs: it was defined just once, and the engine automatically calculates Profit wherever Store sales and Store costs appear at the second level in the columns' hierarchy (since I opened the context menu on that level);
2. Drink & Food, that is the sum of measures for Food and Drink;
3. USA plus Canada, that is the sum of measures for USA and Canada, on rows.

Suppose that a calculated-field is the sum between X and Y and that they are high level members: the engine will merge all X and Y descendants and calculate the field for them: see example in the figure below:

| Query | Results | Crosstab designer | **Crosstab preview** |

	USA				Canada		USA plus Canada			
	Deluxe Supermarket	Gourmet Supermarket	Mid-Size Grocery	Small Grocery	Deluxe Supermarket	Mid-Size Grocery	Deluxe Supermarket	Gourmet Supermarket	Mid-Size Grocery	Small Grocery
	Unit Sales	Unit Sales	Unit Sales	Unit Sales	Unit Sales	Unit Sales	Unit Sales	Unit Sales	Unit Sales	Unit Sales
Drink	6,389.00	2,180.00	862.00	572.00	3,398.00	956.00	9,787.00	NaN	1,818.00	NaN
Food	51,176.00	16,360.00	7,048.00	4,614.00	25,975.00	6,716.00	77,151.00	NaN	13,764.00	NaN
Non-Consumable	13,225.00	4,167.00	1,800.00	1,069.00	7,270.00	1,842.00	20,495.00	NaN	3,642.00	NaN

3 Technical notes

3.1 Max cells' number

The crosstab has a maximum dimension: it can be configure into SpagoBIQbeEngine/WEB-INF/conf/qbe.xml, inside QBE-CROSSTAB-CELLS-LIMIT tag, the default value is 2000. If the crosstab cells' number exceeds that limit, a warning will be produced and the crosstab will be truncated to that limit.

```
<QBE-CROSSTAB-CELLS-LIMIT value=

"2000" isBlocking=
"false"/>
```

3.2 Temporary table

When executing the crosstab, the Qbe engine will execute the base query (the query defined in figure 1) and will save the results into a table; then the crosstab data are produced quering this intermediate table.
We call the intermediate table as "temporary table", even if it is an actual table created into you data warehouse. This mechanism improves performance during the crosstab query definition process, since, if the base query does not change, it is not re-executed; obviously, if the base query is changed, the temporary table is dropped and re-created. The temporary table's name is composed by:

1. a prefix;
2. the user-id;
3. a suffix.

SpagoBIQbeEngine will create a temporary table per user. The prefix and the suffix can be defined in SpagoBIQbeEngine/WEB-INF/conf/qbe.xml, inside QBE_TEMPORARY_TABLE tag, default values are "TMPSBIQBE_" and "" respectively.

```
<QBE_TEMPORARY_TABLE enabled=

"true" prefix=
"TMPSBIQBE_" suffix=
"" />
```

The user-id is cleaned from accents and non-letter characters. Of course SpagoBIQbeEngine has to be provided with a database connection user with table creation right; if it is not the case, you can disable the temporary table strategy by setting the "enabled" property to false, and this strategy will be substituted by another strategy that uses the SQL inline view construct.